BASIC ENVIRONMENTAL MECHANISMS

Affecting Cultural Heritage

Understanding deterioration mechanisms for conservation purposes

edited by
Dario Camuffo, Vasco Fassina, John Havermans

NARDINI EDITORE

cost

COST Action D 42
Chemical Interactions between Cultural Artefacts and Indoor Environment (EnviArt)

kermesquaderni

COST
European Cooperation
in Science and Technology

COST Action D 42
Chemical Interactions between Cultural Artefacts
and Indoor Environment (EnviArt)

BASIC ENVIRONMENTAL MECHANISMS

Affecting Cultural Heritage

Understanding deterioration mechanisms for conservation purposes

edited by Dario Camuffo, Vasco Fassina, John Havermans

NARDINI EDITORE

kermesquaderni

COST Office, 2010

Neither the COST Office nor any person acting on its behalf is responsible for the use which might be made of the information contained in the present publication. The COST Office is not responsible for the external web sites referred to in the present publication.

COST Office copyright: *No permission to reproduce or utilize the contents of this book by any means is necessary, other than in the case of images, diagrams or other material from other copyright holders.*

Authors and Printer copyright: All rights reserved; no part of this publication may be translated, reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the written permission of the publisher. Authors hold the copyright of their intellectual property and are authorized to make other use of their contributions

This book should be cited as:

D. Camuffo, V. Fassina, J. Havermans (Editors) Basic environmental mechanisms affecting cultural heritage. Understanding deterioration mechanisms for conservation purposes. COST Action D 42: CHEMICAL INTERACTIONS BETWEEN CULTURAL ARTEFACTS AND INDOOR ENVIRONMENT (ENVIART).

This publication is supported by COST

Contact
COST Office
Erwan Arzel
149 Avenue Louise
1050 Brussels
Belgium

Front page: Formella in gilded bronze (1403-1415) with some corrosion, by Lorenzo Ghiberti, Northern door of the Baptistery, Florence. By courtesy of the Opificio Pietre Dure, Florence.

ISSN 2036-1122 ISBN 978-88-404-4334-8

Impaginazione e redazione: Massimo Rubino
con la collaborazione di Elena Nazzari

© 2010 per l'edizione:
Nardini Editore, Firenze
www.nardinieditore.it
info@nardinieditore.it

Stampato a Firenze nel 2010
presso Nuova Grafica Fiorentina

Contents

Foreword

Dario Camuffo, Vasco Fassina, John Havermans

For ethical reasons, the conservation of cultural heritage is a duty for all nations. Slowly, decision makers are beginning to understand that caring about cultural heritage and especially about museum, library and archival collections is a valuable long-term investment for their economy and in the interests of their citizens. The accessibility of movable heritage depends not only on its direct conservation but also on preventive conservation because the quality of the indoor environment is crucial for the preservation of a collection. Sensitive materials, displayed in an aggressive environment may suffer from chemical attack of pollutants, leading to irreversible damage after only a few weeks of inappropriate exposure.

The interpretation of results on the impact of pollutants on the degradation of artefacts (in combination with other environmental parameters, such as humidity and temperature) and consequently, any appropriate measure to prevent damage, requires close collaboration between multidisciplinary key players: chemists concerned with environmental effects and material degradation, physicists concerned with microclimate and physical deterioration mechanisms, conservators, conservation scientists, art historians, curators, environmental engineers, show case manufacturers, and even politicians and decision makers concerned with international standards.

Within the EU Research Initiatives (from PF2, 1986 till PF7, 2010) over 100 projects have been dedicated to cultural heritage. Among these projects, more than 20 years of European Research Initiatives are being carried out in the field of Cultural Heritage. The main goal is to reinforce the scientific basis for the establishment of measures and methodologies for the protection and rehabilitation of European Cultural Heritage.

But how can we reach the stakeholders? One of the methods is networking which takes place not only in the EU Framework programmes but also in COST activities.

What is COST?

The *European Cooperation in Science and Technology* (COST, www.cost.eu) is one of the longest-running projects supporting co-operations among scientists and researchers across Europe. The *European Science Foundation* (ESF, www.esf.org/) is the legal entity that provides and manages the scientific, administrative and technical secretariat for COST. Within COST the *Cultural Heritage Interest Group* (HIG) is a part of the *European Cultural Heritage Network* (ECHN, www.echn.net/echn/) and is an umbrella covering COST Actions 625, A27, C17, C20, G1, G7, G8, IE0601 and D42.

Within these COST Actions, not only topics are discussed on researching the materials of our cultural heritage in art, archaeology, the built environment and conservation, but also on how take the correct precautions to have our heritage accessible for the future.

For example, COST Action G1 *'Application of Ion Beam Analysis to Art or Archaeological Objects'* was launched in 1995 and is the first COST Action specifically devoted to cultural heritage research. This Action ran for 5 years aiming for example to promote exchanges between the various laboratories or scientists involved in this activity throughout Europe.

COST Action G7 was dedicated to Artwork Conservation by Laser and has been set up to address challenges in three main areas: Laser systems for investigation and diagnosis; Laser systems for real-time monitoring of environmental pollution and Laser Systems for cleaning applications.

The main objective of COST Action G8 was to achieve a better preservation and conservation of our cultural heritage by increasing the knowledge in museum objects through non-destructive analysis and testing.

This book is an outcome of cooperation in COST Action D42.

What is COST D42 – ENVIART?

On June 28, 2006, COST Action D42 *'Chemical Interactions between Cultural Artefacts and Indoor Environment'* (EnviArt, www.echn.net/enviart/) was born. The aim of COST D42 is to *explore chemical interactions between cultural artefacts and typical indoor environmental conditions* through field studies and laboratory experiments and to transfer the results into preventive conservation practice. COST Action D42 established the links between the old and new European Research Initiatives and broaden it with new sections and co-operation initiatives. 26 countries are involved, covering not only Europe but also abroad as the US.

Within this Action three working groups (WG) were established, each of them devoted to specific activities, and subdivided into Task Groups (TG) as follows.

WG1 *"Preservation"* is composed of TG1 *'Degradation and Stabilization'* and TG2 *'Prevention of Artefacts'*.

WG2 *"Analysis"* is composed of TG1 *'Material Composition and Deterioration of the Objects'* constituting the tangible cultural heritage and TG2 *'Indoor Air Quality and Environment'*.

WG3 *"Guidelines"* is building bridges between developed techniques and the application. It is composed of TG1 *'Methods'* and TG2 *'Storage and Health'*.

All WGs have a common focus: fundamental research and education in order to safeguard our cultural heritage and prevent it from deterioration by environmental factors.

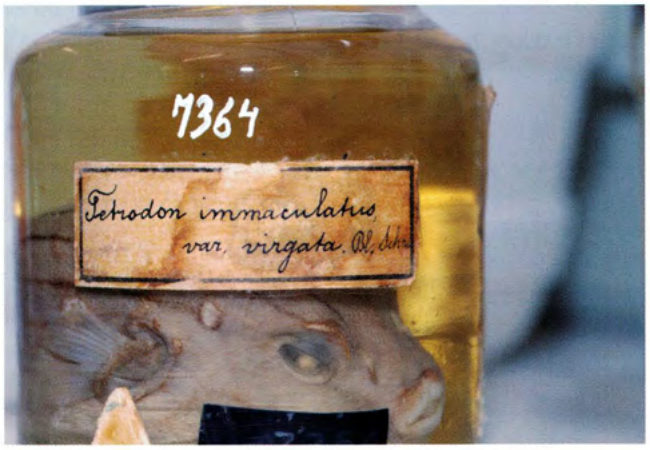

Fig. 1 - How complex can environmental deterioration be? The micro environment inside the vessel keeps the artefact preserved while the macro environment deteriorates the information on the object (photo by C. van den Berg, Naturalis).

What is CEN TC 346 and the ongoing standardization on cultural property?

In 2001 the Italian Standardisation Body *"Ente Italiano di Unificazione"* (UNI, www.uni.com/it/), presented a request to the *European Committee for Standardization* (CEN, also *Comité Européen de Normalisation*, www.cen.eu/cen/NTS/Pages/default.aspx) to create a new Technical Committee (TC), i.e. TC346, dealing with the conservation of cultural property. The scope of CEN/TC 346 concerns the *standardisation in the field of definitions and terminology, methods of testing and analysis, the characterisation of materials and deterioration processes of movable and immovable heritage, and the products and technologies used for their conservation, restoration, repair and maintenance.*

CEN TC 346 started his activity in June 2004 and since the beginning of 2005 five Working Groups (WG) were operative, as follows.

WG1 *"General guidelines and terminology"* is aimed to draft guidelines and standards on conservation planning, including conservation monitoring, terminology dealing with movable and immovable components, deterioration processes and its graphic and symbolic documentation.

WG2 *"Materials constituting cultural property"* is aimed to define tests and methodologies for the analysis and characterization of materials and the evaluation of the state of conservation.

WG3 *"Evaluation of methods and products for conservation"* is aimed to draft documents and criteria to select methods and/or products and to select operative/working conditions in relation to the conservation/restoration, repair, maintenance and preventive conservation. Finally, to draft documents on the evaluation of operative methodologies.

WG4 *"Environment"* is aimed to draft guidelines and standards concerning the assessment, measurement and control of environmental factors, including indoor climate, air quality, HVAC, lighting, exhibition and storage conditions, environmental risk and any other real or potential interaction between the environment and cultural property.

WG5 *"Transportation and packing"* is aimed to draft guidelines and standards on packing and transportation of cultural property.

Advantages from standardization will derive from:

i) improving diagnostic tools, reducing costs, with a subsequent better management of funding and benefits to conservation and related business activity;

ii) suggesting diagnostic studies, sharing results and avoiding to repeat expensive research when not necessary, promoting conservation;

iii) helping to develop and improve products, materials, equipment and technologies to be specifically used for conservation;

iv) increasing the durability of conservative interventions, reducing costs by better focusing, planning and optimizing the subsequent actions.

Users of the guidelines and standards developed by CEN/TC 346 include: international bodies concerned with cultural heritage conservation (e.g. UNESCO, ICCROM, ICOMOS, IIC, ICOM); national governmental and non-governmental bodies with the same aim (e.g. Ministries of Culture and Education, Governmental Agencies, Heritage preservation bodies); ecclesiastical bodies; restoration/ conservation schools; public and private analytical laboratories; professionals in the field of conservation, restoration, exhibition, transportation and packaging; distributors and manufacturers of materials and tools used in restoration; HVAC planners and installers, personnel in charge of museums, galleries, libraries and archives.

The CEN TC 346 activity in the first five years of its life can be summarized as follows.

Already published documents: (1) EN 15801-Determination of water absorption by capillarity (WG3); (2) EN 15802-Measurement of static contact angle (WG3); (3) EN 15803-Determination of water vapour permeability (δ) (WG3).

Documents expected in 2010: (1) prEN 15886-Colour measurement of surfaces (WG3); (2) pr EN 15898 - Main general terms and definitions concerning conservation of cultural property (WG1); (3) prEN 15757-Specifications for temperature and relative humidity to limit climate-induces damage in organic hygroscopic materials (WG4); (4) prEN 15758-Procedures and instruments for measuring temperatures of the air and the surfaces of object (WG4); (5) prEN 15759-Specification and control of indoor environment - Heating of places for worship-Part 1 (WG4).

Documents expected in 2011: (1) pr EN 16096-Condition survey of immovable heritage (WG1); (2) pr EN 16095-Condition report of movable heritage (WG1); (3) prEN 15999-Guidelines for management of environmental conditions - Recommendations for showcases used for exhibition and preservation of cultural heritage (WG4); (4) prEN 15946-Packing methods (WG5).

Liasons established with International Institutions

ICOM CC - *International Council of Museums*. Represents 22.000 members worldwide (liason established in 2007).

IIC - *International Institute for Conservation of Historic and Artistic Works*, representative of 2.300 individuals in 65 countries (liason established in 2008).

ECCO - *European Confederation of Conservator/Restorers Organization*, representative of 5000 members in 16 European countries (liason established in 2008).

UIA - *International Union of Architects*, representative of 1.300.000 architects worldwide (liason established in 2009).

Why this book?

Conservation is a relatively novel science based on several disciplines having strong synergy between them. Preventive conservation requires a thorough knowledge of every deterioration mechanism, a careful evaluation of all potential risks and, finally, a conscious intervention that considers pros and cons of any strategy. Material science, physics, chemistry and

biology are strictly related to each other, and microclimate plays a fundamental role in governing chemical and other deterioration mechanisms. Pollutant deposition is governed by microclimate. Again, pollutants become more and more aggressive with increasing temperature (i.e. Arrhenius equation) and humidity levels. Lowering temperature, humidity or gaseous pollutant concentration are different strategies to obtain the same effect, e.g. to reduce corrosion. In the presence of already deposited particulate matter, only humidity and temperature govern corrosion. Colour fading, equilibrium moisture content, shrinkage and swelling of hydrophilic materials are mainly governed by relative humidity. Briefly: conservation science is based on a holistic approach including all disciplines.

Many years ago, conservators were obliged to use tools and instruments devised for other purposes, e.g. industry, agriculture, health, meteorology. In the last decades, an effort has been made to develop new sensors to measure the environmental factors responsible for deterioration in order to take observations without any risk of damaging objects, and possibly, without any contact with them. More recently, specific sensors have been invented to directly measure the damage. The advantage was to get a preventive alert and to gather data to calculate damage functions able to establish the individual or synergistic role of any environmental factor.

As always, research and professional application proceed at different speeds. Many universities have created courses to teach conservation science. Tradition and innovation are often at different levels and these differences generate some confusion. Despite this, we had to proceed with research (especially supported by the EU) and networking in order to discuss, disseminate ideas and share consensus (thanks to COST). The next step is standardization, in order to share knowledge and experience and understand each other by using same methodologies (CEN TC346). Standards go straight to the objective and provide strict terms of reference. However, the long path and the cultural background behind them remain almost unknown to most people.

The above reasons suggested the idea of writing this book: "*Basic environmental mechanisms affecting cultural heritage - Understanding deterioration mechanisms for conservation purposes*" in order to reach a wider audience and make it aware of the above cultural background.

This book shows that the environment should not be regarded as the sum of structural, chemical and physical factors surrounding an object and affecting it (e.g. envelope, walls, light, air, pollutants, humidity and temperature), but that environment and objects constitute a global, interactive system, including feedbacks and synergisms between any object and its surroundings and no object can be percieved separately from its surroundings and the conditions in which it is living.

The first part of this book is dedicated to review the fundamental concepts of molecular physics and chemistry, with particular emphasis on how energy is partitioned and exchanged at molecular and macroscopic level. Microclimate is also extensively discussed in relation to chemical and physical deterioration mechanisms. In any conservative intervention, the first care and attention should be paid to the environment, considered the primary cause of any deterioration mechanisms, and the second to the object, which is experiencing the synergistic effect of microclimate (e.g. temperature, humidity, air motions, electromagnetic radiation) and atmospheric pollution or other factors. Any restoration aimed to mask effects without eliminating their cause (e.g. adverse microclimate and/or poor air quality) is not only useless, but also dangerous.

The second part of this book is dedicated to both indoor and outdoor airborne pollutants, either gaseous or particulate matter, and how to measure their concentration in air with the help of active or passive samplers. It goes insight to chemical degradation processes and ends with examples concerning some specific materials, their degradation with VOC emission, soiling and the perception of damage.

Atmospheric and absorbed water, environmental monitoring and conservation strategies are essential features of this book and are helpfully illustrated with some case studies which always constitute a useful bridge between theory and practice.

This book is both basic and advanced. The aim is to provide cutting edge information about science, technology and ongoing standardisation applied to the environment, materials and the conservation of cultural property. It is a useful tool to increase understanding of deterioration mechanisms and to suggest the most appropriate strategies to avoid or reduce them.

This book is primarily intended for undergraduate and graduate students, young scientists, architects, conservators and whoever is culturally or professionally concerned with preventive conservation or restoration. Finally, it is aimed at those who are specialists in a specific field but are willing to broaden their vision to a wider multidisciplinary approach to the conservation of our tangible cultural heritage.

The Editors: Dario Camuffo, Vasco Fassina, John Havermans

Chapter 1
The role of temperature and moisture

Dario Camuffo

Part 1. Temperature

1. Exchanges of Heat in a Room

Heat is the thermal energy transfer from a body, or a system, to another. Heat passes from warm bodies, i.e. which have a higher thermal level, to cold bodies, i.e. with a lower thermal level, until an equilibrium is reached. Once heat enters a body, it controls the internal energy, e.g. the potential and/or kinetic energy of the molecules, with changes of phase or increases in temperature. In a gas or a liquid, the temperature represents the average kinetic energy of molecular motions, in a crystal, the vibration of atoms around their position in the crystal lattice.

Heat is exchanged via three mechanisms: (i) conduction, i.e. the transfer of thermal energy from a body at a higher temperature to another body at a lower temperature to equalize temperature differences; (ii) convection, i.e. heat entering a fluid, e.g. air, generates a three-dimensional fluid motion (via expansion and buoyancy force), while it also causes heat and mass to be transported by this bulk motion; (iii) radiation, i.e. the thermal electromagnetic radiation spontaneously emitted from the surface of bodies at any temperature. At room temperature all bodies emit radiation in the infrared (*IR*) band of the spectrum. Any temperature unbalance implies heat transfer and diffusive or convective air motions. This may be influential to some deterioration mechanisms, and to soiling due to the increased deposition rate of airborne particulate matter.

It is clear that all bodies, and the air, interact with each other in order to reach equilibrium. However, it is extremely difficult for a stable equilibrium to be reached, except for a few specific cases like burial rooms kept for centuries below a thick layer of insulating soil. The fact is that many sources and sinks of heat naturally exist, fed by seasonal and diurnal cycles, e.g. air leakages, heat losses through the building envelope, the presence of people, the operation of lighting or HVAC systems, or other disturbing factors. Any body responds in a different way to the microclimate forcing, for its thermal inertia or other different physical properties that affect in a different way any specific exchange of heat. In practice, the general case is composed of bodies and air continually interacting with each other in view of reaching a final equilibrium that is incessantly updated and changed.

Any room that contains a collection, the indoor air, and any object of this collection, will take an active part in pursuing their equilibrium with ceaseless heat and moisture exchanges. Our task is to be aware of what is happening and to understand, control or mitigate the mechanisms that in the long run may be dangerous to the preservation of tangible cultural heritage.

2. The two definitions of Temperature and their consequences

What is temperature? If we want to measure it, what should we know? Everybody knows that temperature is a physical variable, relevant in conservation, and that it is measurable with a thermometer. Thermometers show thermal levels. However, not everybody has a clear idea of the physical meaning of this variable and of the risks to collect thermometer readings with no physical value, which may lead to inappropriate interpretations or decisions. We can define temperature in two ways, an empirical and a theoretical one, as follows.

The **empirical definition** is very simple: *"temperature is the physical variable we can measure with a thermometer"*. This definition is similar to the definition of time, i.e. "the physical variable we can measure with a clock". Any physical variable should be objectively measured with instrumental readings. Any other definition of time looses any physical meaning but concerns metaphysics, philosophy, psychology, or other disciplines. The measurable time is therefore the time elapsed between two clock readings, and the future time will be similarly established with other clock readings. In the same way a thermometer can solve our problem about temperature, and our variable is known after a simple instrumental reading indicating a thermal level. However, a problem is that a thermometer indicates only the temperature of its bulb, or of any other electronic sensor, and we more or less arbitrarily assume that it is at the same thermal level as the air, or the target surface of the object we want to measure.

We can take advantage of this empirical definition only when we have an appropriate physical idea of the theoretical definition, i.e. what we should measure, the perturbation factors we should avoid, how to proceed and shield the thermometer, what is the mechanism we are investigating, where to place the sensors to perform a correct measurement. To this aim, we must be sure that the instrument is calibrated and appropriately used, and we should avoid that the measurement itself will affect the temperature we wish to investigate.

The **theoretical definition** given by Thermodynamics is based on the energy of molecular motions, mainly vibrations of atoms in solids, and molecular speed in liquids and gases as we already said. The kinetic theory of gases establishes that *"the temperature (T) is proportional to the average Kinetic Energy (E_k) of the molecules"*.

More exactly, $E_k = 1/2\ kT$ per degree of freedom of any molecule. The degrees of freedom are 3 for a mono-atomic molecule, 5 for a diatomic molecule, 7 for a three-atom molecule, etc. The proportionality factor k is the Boltzmann con-

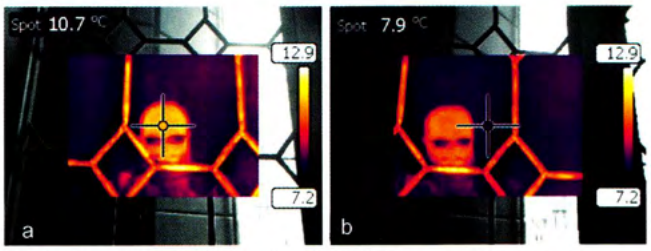

Fig. 1 - Glass reflects part of the IR radiation in the spectral band used by radiometers or thermocameras. In this example, in which thermal images are placed in the centre of visual images, we see the reflected IR image of who is taking the infrared picture of a stained glass window. The spot measurement with the pointer located in the virtual image of the face (left side picture) gives 10.7°C, outside it gives 7.9°C (right side).

stant and the temperature T is expressed in Kelvin (K). For any kind of molecule we have

$$E_{kin} = \tfrac{3}{2} kT$$

if we make reference to the mass centre and the three-dimensional translational motion (i.e. excluding rotation).

If we remember that the physical definition of Kinetic Energy is

$$E_{kin} = \tfrac{1}{2} m v^2$$

where m is the mass of a molecule and v its speed, we recognize that at the same temperature T the average speed is higher for lighter molecules and slower for heavier ones. More precisely, if we consider two molecular species, label 1 and 2, we have: $m_1 v_1^2 = m_2 v_2^2$, hence the speed of the molecule label 1 is

$$v_1 = \sqrt{\frac{m_2}{m_1}}\, v_2$$

i.e. proportional to the speed of the molecule label 2, but the proportionality coefficient is the square root of the inverse ratio of their masses. The mass of H_2O, N_2 and O_2, are $m_{H2O} = 18$, $m_{N2} = 28$ and $m_{O2} = 32$ respectively, and the H_2O and N_2 molecules are 1.33 and 1.07 times faster than O_2, respectively. This explains why the diffusion coefficient of water vapour into Oxygen, i.e. $D_{H2O\text{-}O2} = 0.420$ cm^{2}s^{-1} at T=20°C is larger than that of Nitrogen, i.e. $D_{N2\text{-}O2} = 0.219$ cm^{2}s^{-1}, under the same conditions. As a consequence, molecules of water vapour penetrate and diffuse inside porous materials at a faster rate than dry air.

The practical consequence of the above thermodynamic definition of T in terms of E_{kin} is that the air or the surface temperature is proportional to the average kinetic energy of the molecules and this excludes any radiant contribution from third bodies, which should be considered as an additional, external, perturbing energy. However, it is impossible to work in the absence of *IR* radiation and theoretically the measurement should be performed in a room where all surfaces have the same temperature as the target surface, i.e. the *IR* impinging on the target surface and reflected by it should be the same as the *IR* emitted by it. If the power of the reflected *IR* is higher, the target surface appears warmer, if it is lower, colder. Sometimes it is possible to find environments where all surfaces have the same temperature; more often it is not so. This means that the observation should be made by carefully shielding the sensor from either direct or reflected radiation, including ultraviolet (*UV*), visible or infrared (*IR*) part of the spectrum, coming from the sky or other bodies.

The above thermodynamic definition excludes any contribution from radiant energy from other sources, considered noise. However, for the same physical mechanism, i.e. the thermal emission of bodies, we can focus on the *IR* emitted by our target surface and deduce its temperature from accurate *IR* measurements. It is well known that bodies emit radiant energy whose power P is proportional to the fourth power of their surface temperature, i.e. the Stefan-Boltzmann law

$$P = \varepsilon \sigma T^4$$

whence

$$T = \sqrt[4]{\frac{P}{\varepsilon \sigma}}$$

where ε is the emissivity of the surface, ranging between 0 for a perfect mirror and 1 for a fully emitting-absorbing surface, and σ is the Stefan constant. Polished metals have a very low emissivity, i.e. $\varepsilon<0.1$. On the other hand, water, graphite, asphalt, terracotta, brick, plaster, concrete, wood, paper, marble, limestone, granite, sandstone, gypsum and stucco have a high emissivity, i.e. $\varepsilon>0.9$. At ambient temperature all bodies emit in the *IR* part of the spectrum, and we could know their thermal level by measuring the *IR* emitted power with a radiometer, if we are able to exclude or correct for the *IR* arriving from the exterior or from other bodies. It is clear that we cannot measure the temperature of a mirroring surface simply because we do not see it, but only the thermal images of bodies reflected by it. This methodology can be used only for bodies having low reflectivity and high emissivity; fortunately this applies to most materials used for cultural property. The correction can be simply made by adjusting the appropriate value of ε in the instrument, if we know the emissivity of our target surface, or by totally shielding our target surface from any external radiation, as we will see in the chapter concerning how to measure temperature.

Although the measuring principle is the same, we have two classes of instruments. The first is the radiometer that evaluates the average *IR* power emitted from a spot or a wide area, but is unable to give any thermal image or to specify where the target area is colder or warmer. The second is the infrared camera which reproduces in false colours the thermal image viewed with its electronic eye sensitive to *IR*, typically in the 7-14 μm band interval. A problem is that the emissivity is dependent on the *IR* spectral wavelength. It may happen that a material has high bulk emissivity, e.g. glass whose ε = 0.91-0.94 making an average over all emitted wavelengths, but has low emissivity and high reflectivity or transmissivity in correspondence to the specific spectral band used by the measuring instrument (Fig. 1).

The *IR* based methodology cannot be applied to surfaces having low emissivity and/or high reflectivity or transmissivity in the measuring spectral band, e.g. polished metals and glass, and measuring the radiant temperature of stained glass in historic buildings is a problem. However, we can reach our goal with a trick. If we stick a small piece of self-adhesive paper on the reflecting, or the transparent target surface, the thermal capacity of the piece of paper is negligible and the paper practically assumes the temperature of our target surface. We know the emissivity of paper and we can assume that the *IR* emitted from our paper sample is representative of the temperature of the reflecting surface onto which the paper is sticking.

The same trick can be used if we wish to know the air temper-

ature, although air is almost fully transparent to *IR*. A sheet of paper soon reaches thermal equilibrium with the air and the *IR* emitted by the paper is representative of air temperature in the case conductive air motions and *IR* from heaters or other extraneous heat sources can be excluded. An example is shown in Fig.2, where the vertical profile of the air temperature in a church has been evidenced by a strip of paper raised from the floor up to the ceiling. All the main surfaces in this church have similar emissivity values, i.e. plaster on walls and ceiling: $\varepsilon = 0.91$; soot on surfaces: $\varepsilon = 0.91$-0.94; wooden pews, doors, presbytery behind the altar, choir loft etc.: $\varepsilon = 0.90$; stucco decorations: $\varepsilon = 0.93$; floor made of dolomite limestone: $\varepsilon = 0.96$ and paper: $\varepsilon = 0.93$. Setting the instrument to the paper emissivity, we obtain in the same picture the air temperature and in good approximation the surface temperature of most objects inside the church. It is very important to measure the air and surface temperatures with the same instrument because only in this case are we sure that all differences between a thermal level and another are due to real temperature differences. If we use two or more distinct instruments or sensors, we fall within the unavoidable cross-comparison uncertainties due to differences in calibration and/or response.

Strips of materials having ε close to 1 and low thermal capacity, i.e. black-body strips, are very useful to obtain an indirect measurement of thermal profiles in the air, representing the equilibrium derived from air temperature, convective motions and *IR* radiation. They have been used in the EU Friendly-Heating project (Camuffo et al., 2007, 2010a) and have been proposed in the CEN standardisation (CEN prEN15758, 2010).

Once we have in mind a clear idea of when, where, and how to perform a temperature measurement, and we have an instrument appropriate to do it, then we can use our thermometer according to the theoretical definition and all related thermodynamic consequences. Only at this point will our thermometer reading represent our physical variable. Lacking this, our readings will be a mere collection of numbers deprived of any usefulness or, worse, they may be input for inappropriate decisions, e.g. concerning the room temperature control.

3. Why measure Temperature for Cultural Heritage Preservation

Temperature is a key variable in conservation, because it exerts a direct or indirect role in generating internal tensions, in driving interactions between environment and objects or in determining the rate of some deterioration mechanisms. The surface temperature is generally different from body to body and, on the same body, from point to point. Absolute values, space gradients, time rates, cycles (either seasonal or daily) and fluctuations are all relevant factors.

Materials change physical characteristics as they absorb or release heat. The energy that has been absorbed with heat increases the vibrations of molecules that will require a wider space around them. The cumulative effect will be a three-dimensional expansion of objects, or a contraction if heat will be released. In the event internal stress exceeds the material resistance, it may generate yield and failure.

When the temperature rises above some critical levels, some materials melt, or lose their strength, as it happens for some types of glue used to restore objects. Other materials become rigid and brittle if the temperature falls below the glass transition temperature. Temperature may affect some minerals and masonry crystallisation, especially in the case of daily cycles around the critical level for change of state. The problem is similar to the freezing-thawing cycles when temperature swings around 0°C.

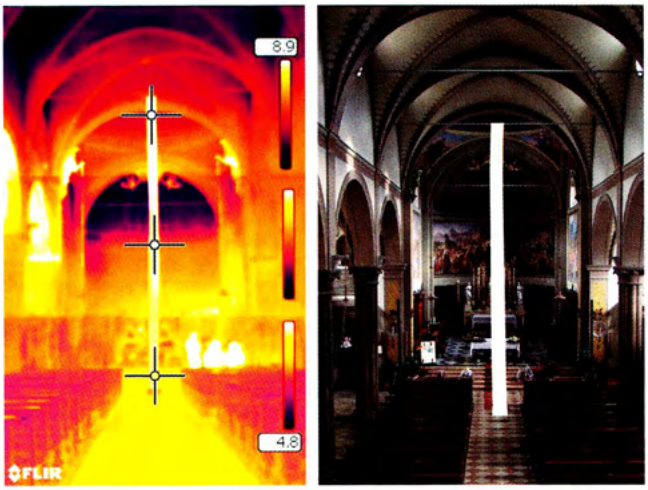

Fig. 2 - Temperature distribution inside a church taken with a thermocamera. The same view is reported in both IR and visible light. The vertical profile of air temperature is made visible with a strip of paper raised from the floor to the ceiling and evidenced with three cross pointers. Reading on the strip the spot temperature at several heights from the floor it is possible to obtain a vertical temperature profile and at the same time to know the temperature of all objects in the church. The IR image is composed of three superimposed pictures, all having the same range, i.e. 4.8° to 8.9°C. In the church, a warm-air heating system operates, with emission grilles on the left nave. This justifies the high temperature of the left arches, and of the two wooden angels supporting lamps from the ceiling, on both sides of the altar. Wood is an insulating material, and the surface temperature of the two angels rises faster than the plaster in the vaults. Three standing persons are visible on the right.

Another key role of temperature emerges from the Arrhenius equation which describes the effect of temperature on the reaction rate originated from the variability of the so-called "equilibrium constant" $k = k_1/k_2$, where k_1 and k_2 are the rate "constants" for the forward and reverse reactions respectively. We underline that k is improperly named "constant" because it is temperature dependent, although it is repetitive, i.e. it returns to the same value when the temperature returns to the same level, and it is constant when the temperature remains unchanged. The Arrhenius equation is

$$k = A \exp(-E_a/\Re T)$$

where the term A is the pre-exponential factor which derives from integration. This is related to the frequency of molecular collisions and the entropy of the system. E_a is the activation energy, i.e. the energy barrier that the reactants should pass for a reaction to occur. $\Re$ is the constant of gases. In practice, the Arrhenius equation establishes that the reaction rates have an exponential increase with increasing temperature. Any chemical deterioration mechanism, e.g. chemical transformations, metal oxidation, corrosion, glass leaching, hydrolysis and degradation of cellulose (paper, textiles), is accelerated at higher temperatures.

Making faster chemical kinetics implies that metabolism and in particular microbial growth and infestation too are accelerated. The temperature level influences the activity of fungi and insects responsible for the biodeterioration of organic materials. However, the microbiological growth soon reaches a physiological threshold above which a too warm environment has an adverse effect and begins to counteract any growth.

Another key role of the air temperature is the (partial) control of the Relative Humidity (*RH*). Actually, *RH* is controlled by two factors, i.e.: air temperature and amount of moisture dispersed in the air. In the event no moisture is added to, or subtracted from the atmosphere, the only governing factor is temperature. A rise in temperature causes lowering in *RH*, which results in drying

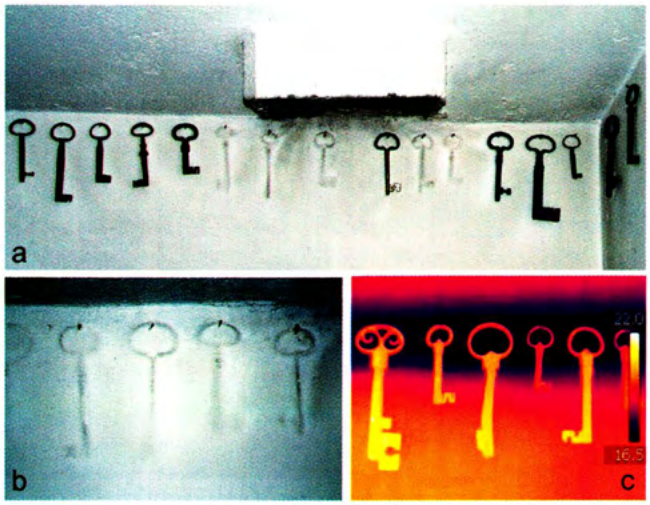

Fig. 3 - A number of keys are hung to this wall. The air and wall temperature are generally different and responsible for endless air motions. Also the keys have a different temperature, intermediate between air and wall. These differences are responsible for aerodynamic particle deposition with wall soiling reproducing the shape of the key. In addition, when the keys are warmer than the wall, another deposition mechanism, i.e. thermophoresis, increases soiling. The figures report: (a) the series of keys with the soiled image of removed ones; (b) removed keys show the detail of selective soiling behind them; (c) a thermocamera image showing the difference in temperature between wall and keys in winter, the room being heated. The horizontal dark colder area is the edge between wall and ceiling, hardly reached (and warmed) by indoor air movements.

the moisture absorbing materials like wood, paper or leather. Such drying may cause shrinkage and embrittlement. On the contrary, a drop in temperature forces *RH* to rise, and hydrophilic materials undergo moisture sorption and swelling. For some materials thermal expansions are a critical factor, e.g. the stone in the leaning Tower of Pisa. For the above-mentioned absorbing materials the direct effect, i.e. the thermal expansion or contraction, is negligible compared to the indirect effect, i.e. shrinkage or swelling due to the change induced on *RH* that in turn alters the equilibrium moisture content (*EMC*). Some feedbacks are negative, but not all. Sometimes a competition between positive and adverse effects is generated, as in the case of metal corrosion: a rise in temperature should increase the chemical reaction rate but, on the other hand, the lowering in *RH* reduces the layers of water covering the surface and the corrosion rate. When the radiation from sun, lamps or radiant heaters reaches objects, the consequent temperature rise causes a local drop in *RH* even when the *RH* of the surrounding air remains constant. Whatever is the air temperature and *RH*, the water vapour may condense on cold surfaces if their temperature drops below the dew point.

Also space gradients and uneven distributions of temperature in a room are relevant. Colder floor and warmer ceiling form internal thermal layering and still air. Gravitational settling of large particles is favoured. The opposite situation, i.e. warmer floor and colder ceiling will feed ceaseless convective motions with accelerated aerodynamic deposition rate of airborne particles on the ceiling and walls. When walls are colder, or warmer than the ambient air, the air in contact with them changes density and buoyancy, forming a layer of uprising air, or noisy downdraughts, both having the effect of increasing the aerodynamic deposition of airborne particles on walls, accelerating soiling. Walls, and especially external walls, are in general warmer or colder than indoor air, depending on the season and the time of day. In practice, any room may have thermal unbalances that act as a powerful motor to create internal convective motions and particle deposition. Heaters or lamps may increase very much such mechanism. Internal edges, corners or cavities cannot be reached by the ceaseless airflow and have a different temperature, e.g. colder if air is warmer. This leads to marked temperature gradients with a number of local consequences on the microscale, e.g. anomalous particle deposition rate, formation of *RH* gradients in the air and *EMC* gradients on the masonry, sometimes constituting a localised favourable habitat for microorganisms.

An example is shown by a number of keys hanging on a wall (Fig.3). The air and wall temperature are generally different. In winter, the wall is colder and the indoor air warmer; vice-versa in summertime. In the figure, the edge between the wall and ceiling, hardly reached by the convective movements of the warm indoor air, remains colder (it is evidenced in the picture by the blue horizontal belt) and, therefore, with a higher *RH* level. A colder and moister microarea may favour the development of microbiological life and mould infestation. The same happens in all less ventilated edges between walls, or between walls and floor, and in corners. Ventilation does not kill microorganisms, but favours a better heat distribution avoiding the formation of bio-favourable habitats. The air-wall temperature difference generates down- or up-draughts along the wall, and airborne particles arriving in proximity of the keys undergo trajectory deviations for local micro-turbulence. When the inertia forces of particles exceed the viscous drag of the air, an aerodynamic particle deposition is created, with wall soiling reproducing the shape of the key. In addition, in winter the temperature of the air and the keys is some degrees higher compared with that of the wall. When a surface is colder than the air, another adverse deposition mechanism is generated, i.e. thermophoresis. A temperature gradient is established in the air close to the wall. Suspended particles receive from each side continual impacts of air molecules bouncing on them. The most energetic hits are from air molecules impinging from the warmer side, with the result that suspended particles are displaced towards the colder side, and in their motion they will eventually reach the cold wall surface. Warmer air also means faster molecular motions, more energetic impacts on suspended aerosols and enhanced Brownian motions. The efficiency of this third type of deposition is greater on fine particles and increases with temperature. More details on particle deposition mechanisms, soiling damage and perception can be found elsewhere (Camuffo, 1998; Brimblecombe, 2010).

In the field of environmental diagnostics or for conservation purposes, measurements of air and object temperature are often needed. Regular or spot monitoring are necessary to verify whether the preservation is made under favourable, or risky conditions. Surface temperature observed under stationary, or dynamic conditions, are a useful tool to individuate structural parts affected by heat loss (Fig.4) or air leakage (Fig.5), or even suspected water percolation (Fig.6), or capillary rise; also to find cracks or internal detachment of plasters or frescoes, or to investigate other structural discontinuities (Fig.7) or to recognize parts made of materials which have a different thermal response (Fig.8). It should be remembered, however, that while temperature measurements constitute a useful investigation tool, they are not per se sufficient to draw any conclusion or to take any decision without the support of other specific investigations.

4. Some facts that we should keep in mind when we perform temperature measurements

When we perform temperature measurements, we should keep in mind a number of theoretical and practical problems that may affect our results. Not always is a solution possible, but hav-

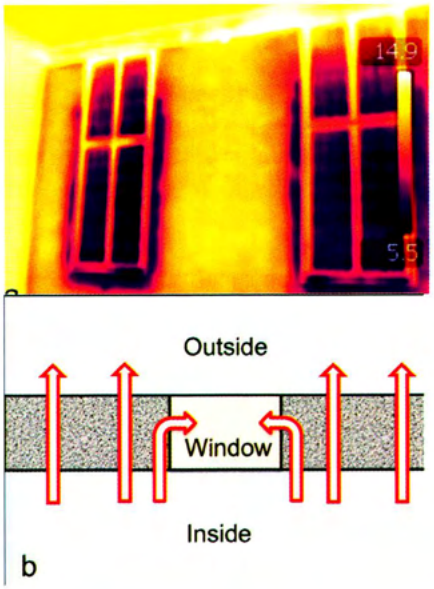

Fig. 4 - (a) Around the windows, a colder blue area is visible. In a heated room the heat flow across the exterior wall is inversely proportional to the wall thickness. In proximity of the window openings, the heat finds a shorter oblique path to go through the window posts. The local heat flow becomes higher and the internal surface, affected by higher dissipation, colder. (b) Scheme of how the heat (red arrows) flows across the cross section of a thick wall (grey) and finds a shorter path across the window posts. Pale brown indicates the window opening.

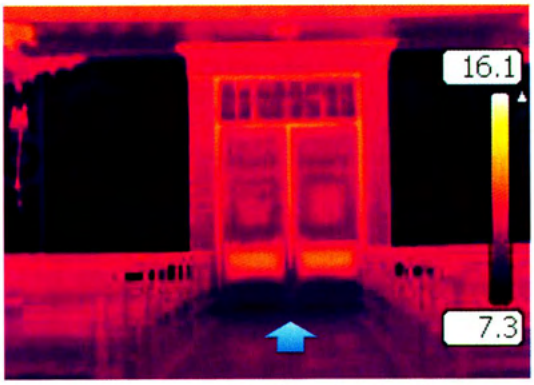

Fig. 5 - Air leakage through the fissure between the entrance door of a church and the floor. In winter, cold air blows in through the lowest openings, which appear colder, and goes out through the upper ones, which appear warmer. The cold air leakage is the blue area on the floor in front of the door, evidenced with an arrow.

Fig. 6 - Water percolation. The portion of wall below the window has been moistened by rainwater percolated inside. Moisture has increased wall conductivity and lowered the thermal level of the indoor moist surface. The different conductivity between damp mortar and bricks makes the internal structure of the wall visible.

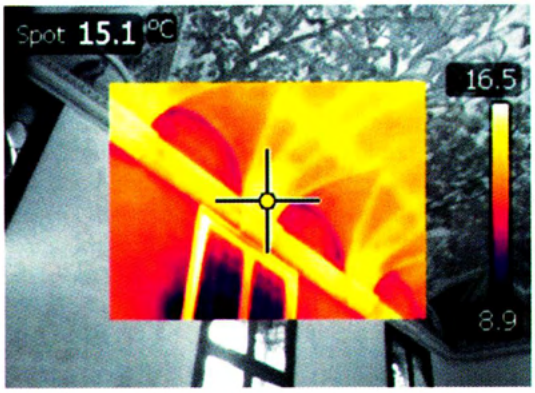

Fig. 7 - Structural discontinuities made visible in a thermal image. Yellow indicates the warmest surfaces, e.g. the wooden frame of the window, the cornice in stucco just below the vaulted ceiling, the wooden frames supporting the vaults. Orange is also warm, but less than yellow. Blue indicates the coldest surfaces, i.e. the window panes and the wall adjacent to the window opening. Violet is an intermediate cold level, e.g. between orange and blue, or typical of the areas hardly reached by warm air motions, e.g. the borders of the lunettes, the edge between the cornice and the wall.

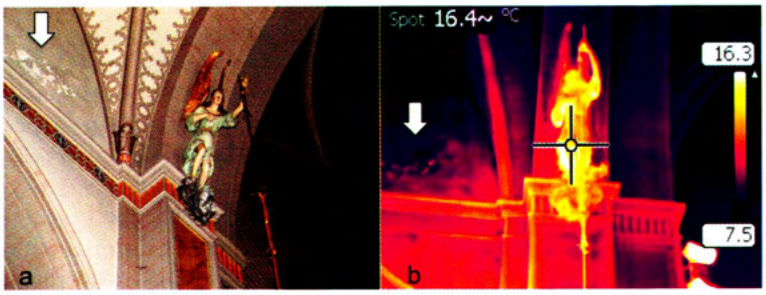

Fig. 8 - Visible and IR pictures of an angel supporting a lamp on the top of the transect pillar. In this church, when the warm-air heating is on, a thermal layering is formed and the wooden angel overheats much more (i.e. some 10°C) than the thick wall structure. The next material to heat is the metal chain of the lamp supported by the angel, then the horizontal stucco decoration along the wall. The deterioration of the plaster (arrow) in the left-side lunette is also visible.

ing a clear idea of them helps to reach the most accurate results.

❑ *A thermometer measures the temperature of its bulb or sensor*, not that of the air or a surface. It is only *assumed* they are in equilibrium, and hence equal.

❑ The heat capacity of the sensor is much greater than the specific heat of the air. In micro indoor environments, any *measurement perturbs the original temperature of the air. The same can be said if we use a contact sensor to measure the surface temperature of an object that has a small heat diffusivity or capacity.* For instance, if we put a sensor into contact with a sheet of paper, or a parchment, or a painting on canvas, we will find that in the contact point both the sensor and the object surface will reach a common thermal level that is intermediate between the temperatures the sensor and the object had before. Contact sensors cannot be used in this specific case, especially if they have a large heat capacity compared with the target area. This drawback is similar to the *Uncertainty Principle* in Quantum Mechanics where any measurement perturbs the original state of the system.

❑ *We should measure the average kinetic energy of the molecules of the air, or a surface, without the perturbation of IR from other bodies.* Although it is impossible to remove all the *IR, the sensor should be always shielded with a reflecting (metal) screen.*

❑ Thermometer calibration is performed under controlled conditions, with sensors immersed in liquid baths where the *IR* radiation is fully absorbed. As opposed, when we perform a

real measurement, the sensor is always hit by some radiation, and may absorb or reflect it. *Thermometers perfectly calibrated in the laboratory, but having different IR reflectivity, have different outputs*, i.e. different temperatures.

- In the general case in which the air is characterized by ceaseless motions and turbulence, the temperature fluctuates continually. Any sensor has a characteristic time of response that introduces a time lag and a smoothing to the output. Hence, *thermometers which are perfectly calibrated in the laboratory, but have different characteristic time, give different outputs*, i.e. different temperatures.
- If we wish to know *the temperature of a surface hit by solar radiation*, or which is dewing, evaporating, or having other exchanges with the air, the contact sensor will shadow or cover the surface on the sampling point and generate bias. Contact sensors cannot be used in this specific case.
- In general, the air and the surface temperature are different and affected by gradients. *The choice of representative sampling points is a crucial problem.*
- In principle, *artworks should be preserved from any contact with measuring instruments. Non-contact monitoring should be preferred.* Two types of non-contact monitoring exist, both based on *IR* monitoring, i.e. Remote Sensing in which the *IR* sensor is far away from the target, and the Quasi-Contact Sensing, in which the sensor is close to the target, but not in contact with it. Non-contact sensing is possible only for good *IR* emitters, e.g. non-metallic surfaces, or heavily oxidised metals. In the case of surfaces with low emissivity, e.g. polished metals, any methodology based on *IR* sensing is impossible, and contact sensors should necessarily be used.

PART 2. MOISTURE

1. What do we know about water?

A problem in conservation science is that many people continue to use some basic ideas derived from a simplified presentation typical at the high school or even at the university level. However, Nature is very complex and some popular models are unable to explain what is happening in the real world and inappropriate ideas might lead to misleading thoughts and solutions. This becomes clear when advancing in professionalism. The aim of this chapter is to review our basic concepts in line with the problems we will meet in conservation science. Many readers will be tempted to skip this elementary, but fundamental chapter. My suggestion is: please have a look through the following basic sentences. Each of them may be either true or false. You will find the answer at the end of this section.

1. The water vapour can be approximately considered as a (perfect) gas, except when it changes of phase.

2. All the water molecules have the same physical properties and exert the same electromagnetic forces.

3. In the vapour phase, all the water molecules behave in the same way.

4. In the liquid phase, all the water molecules behave in the same way.

5. The corrosion depends on the number of water molecules impinging, per unit time and unit surface, on a metal surface.

6. At high relative humidity levels, air becomes saturated of moisture and cannot support the addition of new molecules of water vapour.

7. In wet conditions objects are covered with water and pores are filled; in dry conditions (i.e. low relative humidity) they are not.

No one of the above sentences is true. In the case not all your answers are "false", it might be useful to dedicate some time in reading the chapter and understand why we have provided a different answer.

2. The key moisture variables

Before to start our discussions about moisture and preservation needs, we should shortly present and define the key moisture variables. We will use in such definitions the symbols officially used by the World Meteorological Organisation (WMO, 1966). However, in order to avoid that the reader should remember too many symbols, we will use throughout the text the acronym of each variable, that will be established at the end of any definition. Formulae for calculations are reported in the Appendix at the end of this Chapter. Readers interested to know more about variables, physical meaning, how they are mathematically derived, related graphs etc. can find more in other papers (Camuffo, 1998; 2004).

The Mixing Ratio of moist air (i.e. dry air and water vapour) (WMO symbol ***r***) is the (dimensionless) ratio of the mass of water vapour m_v to the mass of dry air m_a

$$r = \frac{m_v}{m_a}$$

r represents the ponderal mixture of these two gaseous substances, i.e. the proportion in which water molecules are mixed with dry air. As the above ratio is very small, *r* is conventionally multiplied by 10^3 and is expressed as g kg^{-1}. In the following we will refer to this variable with the acronym *MR*.

This variable is independent from temperature, barometric pressure and volume. Therefore, *MR* is a conservative variable for physical processes unless some moisture is subtracted to the air (condensation on a surface) or added to it (e.g., evaporation from a surface, people breathing) or the air is mixed with other, external air. This property is very useful in environmental diagnostics for the following applications.

The *MR* in a room may vary with air exchange, or when some moisture is released by visitors, or damp walls, or is released or absorbed by hydrophilic materials. An indoor *MR* increase, not justified by external air leakage, may be an index of evaporation (Fig.9).

MR is a tracer to see the path and the diffusion of external air (Fig. 10). To obtain maps with the *MR* distribution in a room, it is necessary to calculate this variable from direct observations of air temperature and relative humidity, or wet-bulb temperature or dew point (see later).

The gradient of *MR* in proximity of a surface is useful to detect exchanges of moisture between the ambient air and the surface. When the surface is condensing, the air near to it is locally deprived of vapour, i.e. *MR* shows a minimum. When the surface is evaporating, the air near to it is locally enriched with vapour, i.e. *MR* shows a maximum (Fig.11). This fact provides a useful tool to perform diagnostic analyses in historical buildings, aimed to detect areas suffering for moisture exchanges. The methodology is easy: it only requires to compare observations taken close to the surface and far from it and see whether the surface is locally increasing or decreasing the moisture content. In still air,

when condensation (or evaporation) occurs on a wall, the excess (or the deficiency) moisture concentration in front of the surface decreases exponentially with the distance and 63% of the departure from the undisturbed level is attenuated at 9 mm distance from the wall (Camuffo and Giorio, 2003). This distance, in which a variable is reduced by 1/e (where e is the Neper number, i.e. e=2.7182818...) is called damping depth. In still air the damping depth is determined by the diffusivity of moisture. In the case of air movements or turbulence the damping depth is reduced.

The Specific Humidity of moist air (WMO symbol ***q***) is the (dimensionless) ratio of the mass of water vapour m_v to the mass of moist air m_a+m_v, and this ratio represents the ponderal dilution of the vapour in the atmosphere, i.e.

$$q = \frac{m_v}{m_a + m_v}$$

As the mass of moisture is small compared to the mass of air, the value of this variable is very close to the mixing ratio.

As the above ratio is very small, q is conventionally multiplied by 10^3 and is expressed as g kg^{-1}. This variable is independent from temperature, barometric pressure and volume. In the following we will refer to this variable with the acronym *SH*.

The Absolute Humidity (WMO symbol $\boldsymbol{d_v}$) is the density of the water vapour, i.e. the mass of vapour m_v contained in the unit volume V of moist air. It is also called *mass concentration* or *moisture content*. Its value is close to r.

$$d_v = \frac{m_v}{V}$$

It is used to evaluate the total mass of water vapour contained within a volume (e.g., room, showcase) and potentially available after condensation or after the permanence of people. A person releases about 50 g of water vapour per hour. It is easy to obtain the increase in moisture in a room by multiplying the unit release of moisture by the number of persons and their time of permanence and divide by the room volume.

As this ratio is very small, it is conventionally multiplied by 10^3 and is expressed as g m^{-3}. This variable is dependent on temperature, barometric pressure and volume. In the following we will refer to this variable with the acronym *AH*.

The Relative Humidity (WMO symbol $\boldsymbol{U_w}$) (popularly termed '*humidity*') is the (non-dimensional) ratio between the actual partial pressure of the vapour e_v and its saturation vapour pressure $é_w$:

$$U_w = \frac{e_v}{é_w}$$

Below 0°C, the saturation point is conventionally computed with reference to supercooled water (subscript: *w*). If the temperature is far below 0°C, reference is made to ice (subscript: *i*).

This is also defined as the ratio x_v/x_{vw} of the actual vapour mole fraction x_v to the largest possible vapour mole fraction x_{vw} (called "saturation fraction") which the air would have with respect to water at the same pressure p and temperature t. *Relative Humidity represents the actual degree of saturation of the water vapour*. Please note: saturation of the *vapour*, not of the *air*. The definition "ratio of the vapour mole fraction x_v to the vapour mole fraction x_{vw} which the air would have if it were saturated with respect to water at the same pressure p and temperature t" is popular but formally incorrect because the air is never saturated by mixture with another gas, as two gases can be mixed in whatever proportion and the partial pressures remain independent of each other (Dalton law). Relative humidity can be equally defined for water vapour alone, even in the absence of air, from the ratio of the actual to the saturation vapour pressure as previously discussed. The presence or absence of air is irrelevant.

In 1947, the International Meteorological Organisation (IMO), later World Meteorological Organisation (WMO), defined the Relative Humidity in terms of Mass of water (List, 1966) for some theoretical reasons discussed elsewhere

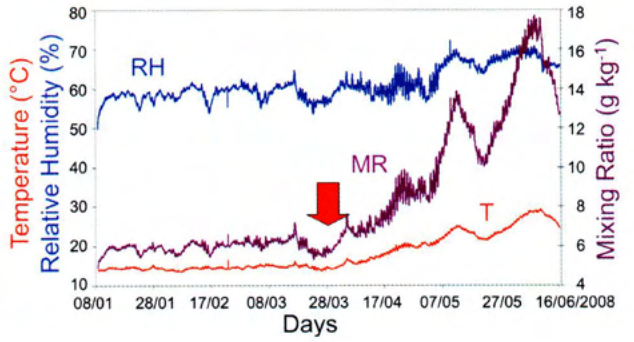

Fig.9 - Air Temperature (T), Relative Humidity (RH) and Mixing Ratio (MR) in a room. After a rainfall (arrow), the exterior wall dampens with some inside evaporation. This is shown by an increase in MR, i.e. the moisture content in air.

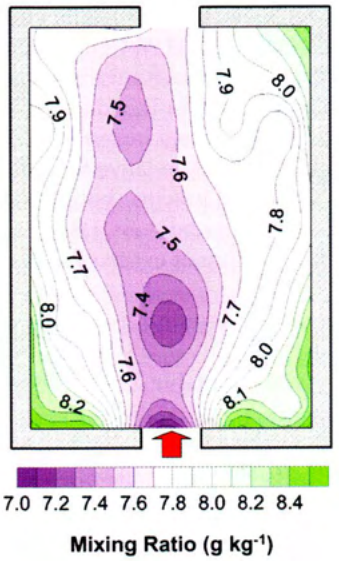

Fig.10 - Map of humidity Mixing Ratio (g kg-1) pointing out the penetration and diffusion of external air entering through the door at the bottom (red arrow).

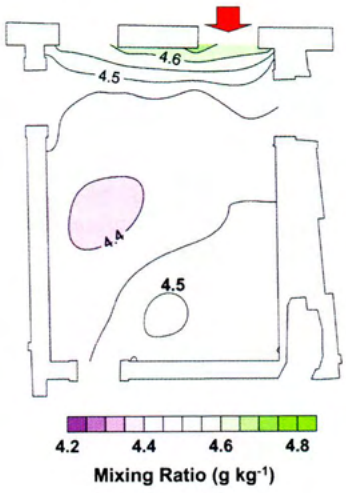

Fig.11 - Map of humidity Mixing Ratio (g kg-1) pointing out the evaporation from the wall on the top (negative gradient near the wall, red arrow). (After Camuffo et al., 2010b)

(Camuffo, 1998, 2004) and because this definition is useful for a number of formulae.

$$(U_w)_{1947} = \frac{m_v}{m_{vw}} = \frac{\frac{m_v}{m_a}}{\frac{m_{vw}}{m_a}} = \frac{r}{r_w} = \frac{\frac{m_v}{V}}{\frac{m_{vw}}{V}} = \frac{d_v}{d_{vw}} \approx \frac{\frac{m_v}{m_a + m_v}}{\frac{m_{vw}}{m_a + m_{vs}}} = \frac{q}{q_w}$$

According the 1947 definition, the *Relative Humidity is the ratio between the mass of vapour m_v, or the mixing ratio, or the absolute humidity etc. actually present in whatever volume of atmosphere, to the largest amount possible* (i.e. under saturation, subscript *w*) *of the same variable at the same temperature*. In 1966, however, IMO returned to the original definition in terms of ratio of partial pressures or, which is the same, in terms of mole fractions. At very high temperature and low *RH*, the *RH* levels according the 1947 definition have a minor departure from those obtained from the 1966 definition in terms of vapour partial pressure. This difference vanishes at saturation and low temperatures but increases with increasing temperature and decreasing relative humidity; however it is scarcely relevant in the practice because it reaches 10% under extreme hot and dry conditions, i.e. 50°C and 20% relative humidity. At 20°C it cannot exceed 2% (for details see Camuffo, 2004).

This variable is dependent on temperature, barometric pressure and volume.

In the following we will refer to this variable, expressed in percent (%), with acronym *RH*.

Wet-Bulb Temperature: the temperature of evaporation

The wet-bulb temperature (WMO symbol $\boldsymbol{T_w}$) is the temperature an air parcel would have if cooled adiabatically to saturation at constant pressure by evaporation of water into it, all latent heat being supplied by the parcel. This is the temperature directly measured by the wet bulb of a psychrometer.

From the thermodynamic point of view, T_w is the temperature that an air parcel would have when some liquid water is supplied gradually, in very small quantities and at the same temperature as the environmental air, and then this water is evaporated into the air adiabatically (i.e. the latent heat being supplied by the air) at constant pressure, until the saturation is reached. The saturation is reached for the combined action of two factors due to the evaporation: the increase in *MR* and the drop in *T*. As the WMO symbol is easy to remind, we will continue to use it in the following.

Dew Point: the temperature of condensation

The Dew Point Temperature (WMO symbol $\boldsymbol{T_d}$), commonly termed Dew Point, is the temperature to which a sample of moist air should be cooled at constant pressure and constant water vapour content in order for saturation to occur. The Dew Point can also be defined as the temperature at which the actual pressure of the vapour contained in an air sample equals the saturation pressure, under constant pressure and mixing ratio. Although it is popularly called dew point of the '*air*', it is a property of the *vapour* that might be extended to the air sample.

The Dew Point is the critical temperature around which the transitions between the liquid and the vapour phases change direction. If liquid water and vapour are present in the same room, and both have the same temperature *T*, the water molecules will tend to pass from the liquid to the vapour state (evaporation) if $T > T_d$, and in the opposite direction (condensation) if $T < T_d$. At $T = T_d$, the equilibrium between the two phases is reached. The air temperature is always $T > T_d$. As opposed, the temperature of liquids or solids may be higher, equal, or lower.

Evaporation and condensation should be considered from the statistical point of view, i.e. they represent the net flux resulting from a balance in which some fast molecules may escape from the liquid water and other may arrive from the gas phase. In the previous section we have seen that this flux can be considered as the result of the gradient in vapour pressure established across the liquid-air interface.

Condensation and evaporation happen at two distinct temperatures. By measuring the surface temperature of a wet wall we can know whether it is condensing because it is cold, or is evaporating after capillary suction. T_d depends only on the moisture content in air expressed in terms of Mixing Ratio of humid air. Should it be expressed in terms of Absolute Humidity, the air temperature should be included as well. Always the air Temperature exceeds the Wet-Bulb Temperature T_w, which in turn always exceeds the Dew Point, all of these variables having the same value when saturation is reached. In formula:

$$T_d \leq T_w \leq T$$

This means that the temperature of an evaporating surface cannot reach the temperature of condensation. As the WMO symbol is easy to remind, we will continue to use it in the following.

3. The concept of Dose, the Reciprocity Principle and their application to water vapour

Many physical or chemical systems, or biological processes, are governed by agents, either physical quantities or chemical substances, which produce certain effects. The *cumulative effect* is the result obtained after the agent has been repeatedly applied to the system. The cumulative effect may be, or may be not, equal to the sum of all the individual effects after any single application. The total quantity of the physical or chemical agent that has been acting on, or assumed by, a body is defined "*total dose*" and the partial portion supplied at any time is a *fraction of this dose*. Considering a selected body, the final effect produced by an agent may be the dependent on, or independent from how it has been distributed over the time, i.e. concentrated in a short period or diluted over a longer one. Radiations, gases, medicines, drugs, are examples of substances that furnish doses to some bodies, with various effects that depend upon the type of dose, body and, in some cases, how the dose is supplied over the time. For instance, a bottle having 1 litre capacity is filled when we have poured 1000 g of water, independently from the pouring number, and 1000 g are the total dose. The amount of marble that in a moist environment will react with SO_2 forming gypsum is dependent on the dose of SO_2. If the amount of SO_2 is doubled, also the gypsum is doubled. As opposed, in medicine some drugs can be fractioned in some ways up to a certain limit; changing the time distribution or the total dose the effect may be different.

In the case the cumulative effect equals the sum of the single contributions, the *Reciprocity Principle* applies. The Reciprocity Principle establishes that the same total dose, and the same final effect, can be reached in different ways, e.g. subdividing the individual concentration of the substance as a function of how many times we want to supply it. The only relevant factor is the product: intensity by time, or quantity by

number of times, disregarding how, or for many times, or how long, the total dose is partitioned or supplied. For instance, we can consider a fraction of dose (or a reduced concentration) multiplied by the number of times the fractioned dose is supplied; the same for a reduced concentration (or a reduced intensity) multiplied by the time of exposure. Concerning light or radioactivity: the same (total) dose can be reached doubling the time of exposure but with radiation at halved intensity; concerning drugs, the same (total) dose can be reached supplying per three times the same substance but with concentration reduced to one third.

Now let us consider hydrolysis, corrosion, oxidation, or whatever other deterioration mechanism related to dampness. It would seem logical to expect that the impact of water vapour on objects is *proportional to the concentration of water vapour dispersed in the air* or, better, *proportional to the dose of water supplied to the object*, i.e. to the number of water molecules that have so far impinged on the object surface. It would seem logical to expect that *the probability that a water vapour molecule is absorbed by a hydrophilic material* (e.g. wood, paper), or reacts with a gaseous pollutant (e.g. to form corrosion) *is proportional to the density of population of water molecules*, i.e. to their concentration in air.

However, this is not true. We know that the uptake of vapour from air that causes wood swelling, or the moisture content in hydrophilic materials, or the oxidation of metals and other chemical reactions are only related to *RH*, not to absolute or specific humidity. The same holds for corrosion or mould infestation. In winter the moisture content in air is much less than in summer and on the ground of the number of impacts we should expect that any oxidation or corrosion should occur only in summer, but this is not true. Again, keeping unchanged the temperature, according the Reciprocity Principle, a surface exposed in air for one month at *RH* = 90% experiences the same number of impacts with water molecules as it were exposed for three months at *RH* = 30% or nine at *RH* = 10%. We can easily argue that after one month exposure at *RH* = 90% our surface will be likely affected by biological infestation, or oxidation, or corrosion, or other problems depending on the material nature, but this will not happen after nine months at *RH* = 10%, i.e. with the same dose. Again, in the vapour phase water occupies a volume some 10^3 times larger than in the liquid phase. According to the Reciprocity Principle, to keep an object in contact with water vapour for one millennium would be equivalent to keep it immersed in liquid water for one year. This is evidently not true.

The main conclusions of this section are:

- If the interactions between water vapour and materials are governed by *RH*, then *the number of the water molecules impinging on a surface, the moisture content in air and the exposure time are insufficient to describe the mechanism*.
- The effect of *water vapour is independent from the dose and the Reciprocity Principle cannot be applied* to water vapour;
- If in some occasions (e.g. at high *RH*) few molecules of water may cause relevant effects, and in other occasions (e.g. at low *RH*) a large number of them are unable to reach the same results, *not all the molecules of water vapour have the same behaviour, but under certain circumstances they change character*.

The above conclusions are hard to accept if we keep in mind the classic model of perfect gases where all molecules are identical between them and always have the same behaviour. However, the experience certainly shows that the water molecules sometimes are very active and have strong interactions with materials, i.e. when temperature is low and *RH* is high; sometimes they are (almost) totally inert and simply bounce on the object surface as they were a dry gas, and this typically happens when temperature is high and *RH* is low. This fact is even more surprising because the Arrhenius equation suggests that we should expect higher reactivity at higher temperatures. Something should change character to the water molecule in the vapour phase. What? Why? When? We will see it later, but before we should introduce some other concepts useful to our aims.

4. PVT diagram showing the gas and the liquid phases and the phase transitions

The well-known phase diagram showing the simultaneous presence of the liquid and the gaseous phases in a thermodynamic representation of Pressure (*P*) versus Volume (*V*) and a number of isotherms at various temperature levels (*T*), is represented in Fig.12. The thin lines are isotherms. The thick lines separate the areas with different physical state: Liquid, Liquid + Vapour, Vapour and Gas. The vapour and the liquid can be simultaneously present in the dotted bell ABC, and the point B, on the top, represents the critical point and the isotherm per B it is the critical temperature T_c. The gas is in the region above the critical temperature and the vapour is below it, to the right of the saturation line BC.

In a first approximation, the law of perfect gases can be applied to the water vapour when the water molecules are far from each other (to the extreme right side of the diagram), or when the temperature is high, approaching T_c. The perfect gas approximation cannot be used near saturation (i.e. near the line BC). The saturation depends on moisture concentration (you can reach BC going from right to left) and temperature (going down in the diagram). The moisture content represents the density of the population constituted by the molecules of water vapour mixed to air and, definitely, the minimum distance to which the molecules can approach each other at a selected temperature.

In Fig.12, some blue lines are indicated which correspond to selected *RH* levels, i.e. 100, 90, 80,...%. In this diagram, we see that *RH* is determined by two physical factors: the temperature (alias kinetic energy, alias molecular speed) and intermolecular distance. This holds at any given pressure, e.g. the sea level barometric pressure. A constant pressure corresponds to a horizontal line in the graph. In practice we see that in the molecular water vapour population, *RH* is determined by two independent factors: temperature (or speed) and intermolecular distance. These two factors have similar effects: temperature (or speed) to reduce the probability that a molecule is captured, the distance to reduce intermolecular forces, with the same effect. *This means that, under the same physical conditions (i.e. the same RH and force interactions), speedy (i.e. warm) molecules can approach each other at a shorter distance than slow (i.e. cold) molecules can do.*

For each temperature we will find a critical concentration above which saturation will start to remove water molecules from the gaseous state and keep their number not exceeding the critical level. Saturation can be considered either in terms of pressure or vapour density, also called saturation absolute humidity (Fig.13). The same concept can be expressed making reference to the amount of air mixed with the water vapour, and in this case it will be expressed in terms of saturation mixing ratio or other similar variables. Per each temperature (*T*),

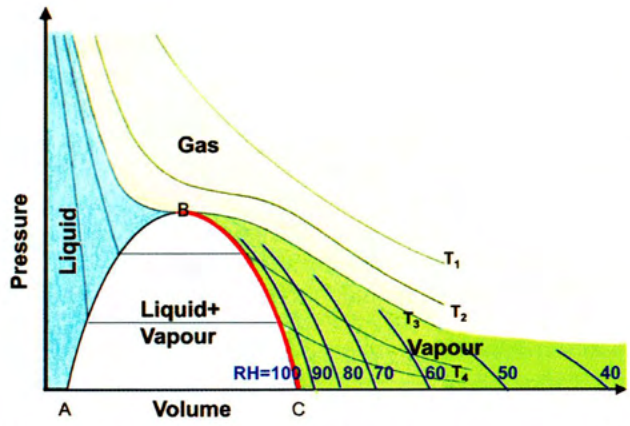

Fig.12 - PVT Phase diagram with Pressure versus Volume, and isotherms (Tn, thin lines), showing the gas, vapour and liquid phases. The red line indicates the saturation border line, when vapour starts condensation, i.e. RH=100%. Blue lines indicate other RH levels, decreasing as the volume and the intermolecular distances increase.

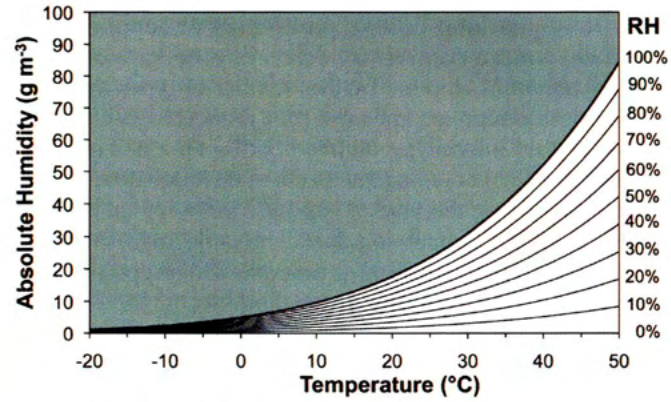

Fig.13 - Maximum concentration of water molecules, i.e. Absolute Humidity, at saturation (thick line) and at selected RH levels, i.e. RH= 10, 20, 30…%. The grey area is not allowed for a stable equilibrium and may host the simultaneous presence of the liquid and the gaseous phases in transition between each other.

water molecules cannot exceed a certain critical concentration in a given volume; this means that a minimum allowable distance between the molecules is established as a function of *T*. Should this intermolecular distance be reduced, i.e. super-saturation, some molecules will pass to the liquid state to be removed from the atmosphere and re-establish the sustainable maximum density.

Physically, we should remind that fast molecules correspond to high temperature, i.e. high energy, and slow to low temperature, i.e. low energy. Therefore, at any vapour concentration a minimum allowable temperature is found, below which saturation will start. This critical temperature is named "Dew Point" (*DP*). High vapour density (or pressure, etc.) is compatible with fast (and warm) molecules having high dissipative potential. Every time two molecules will approach to each other, the higher the temperature, the higher the molecular energy and, secondarily, the shorter the interaction time. Briefly, an energetic particle needs strong forces to be entrapped; as a consequence, high kinetic energy reduces the probability that a molecule is entrapped in a minimum of potential of attractive forces and, ultimately, the passage to the liquid phase. Vice-versa, once the molecules have a certain average temperature, the maximum sustainable molecular density is consequently established.

The vapour saturation and the transition from the gaseous to the liquid phase are mechanisms related to the nature of the vapour, and it is totally irrelevant if the vapour is mixed with dry air. Air is never saturated for being mixed with another gaseous substance. The same properties can be verified with water vapour mixed with whatever unreactive gaseous species, or even in the total absence of air. In the absence of air we can equally define *AH, RH,* T_d; however, we cannot use *MR* and *SH* simply because they explicitly consider the proportion of mixture between water and air, the latter being missing.

It is clear that in the gaseous phase the vapour can exist provided that its concentration and its temperature are compatible between them, otherwise the excess moisture passes to the liquid phase. Similarly, the liquid phase remains in equilibrium with the vapour under the condition of saturation, otherwise it evaporates until this condition is established, or the population in the liquid phase is extinguished. This means that the transfer of molecules from the liquid to the vapour phase, or vice-versa, is determined by a physical mechanism.

When a molecule of water vapour impacts on the surface of liquid water, it remains entrapped in the liquid phase, releasing the latent energy corresponding to a change of potential. However, apart some casual impacts, a statistically relevant vapour-liquid transition (i.e. condensation) occurs when the vapour pressure exceeds saturation or, which is the same, the temperature drops below the dew point.

The first idea is that, in the presence of both the liquid and the gas phases, a continuous flow of vapour molecules is transferred from the air to the liquid (condensation) and another goes in the opposite direction (evaporation). In practice, we can recognise only the net flux, i.e. the difference between the two. Below saturation, the evaporation is dominant and above saturation, the condensation occurs. The process continues as far as the equilibrium is reached, i.e. when the net flux vanishes, just at saturation. This is well known, while less the reason why, and a possible explanation is here reported.

5. The liquid state, the Maxwell distribution and the vapour inside and outside liquids

The physical states are determined by a balance between the Potential Energy generated by the attractive forces that tend to form an ordered structure, and the internal Kinetic Energy due to the thermal molecular motions that tend to destroy any order and create freedom from chaos. The solid state is characterized by a strong potential and a small kinetic energy; the gaseous state by weak potential and large kinetic energy. The liquid state is characterised by molecules with potential and kinetic energy having the same order of magnitude. Any passage of state requires some latent heat to adjust the critical equilibrium typical of each state. In the liquid state, the water molecules have an intermediate character between the freedom that characterizes the gas phase and the rigidity of the crystal lattice of ice. Molecules are still bound to one another, but the binding energy is less than in the solid state. The concept of a physical structure, when applied to the liquid state, should be regarded in a special way. For the ceaseless molecular motions, the long-range and the long-term molecular order are lost. Any order should be concentrated in time and space, e.g. temporary aggregations of some molecules that will be soon destroyed and reformed in other ways.

At subatomic, atomic and molecular scale the situation is very complex and Quantum Mechanics has abandoned the deterministic approach in order to get a better mathematical description of the wave-particle duality of energy and matter at

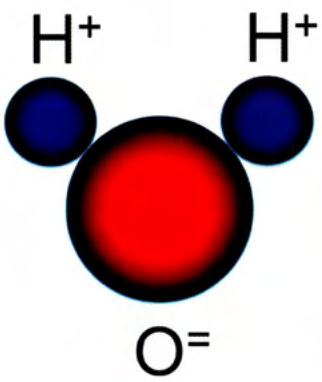

Fig. 14 - The water molecule (H_2O) is an electric dipole. The two Hydrogen atoms (blue) have positive charge, the Oxygen (red) negative. The molecule is asymmetrically charged and behaves as an electric dipole.

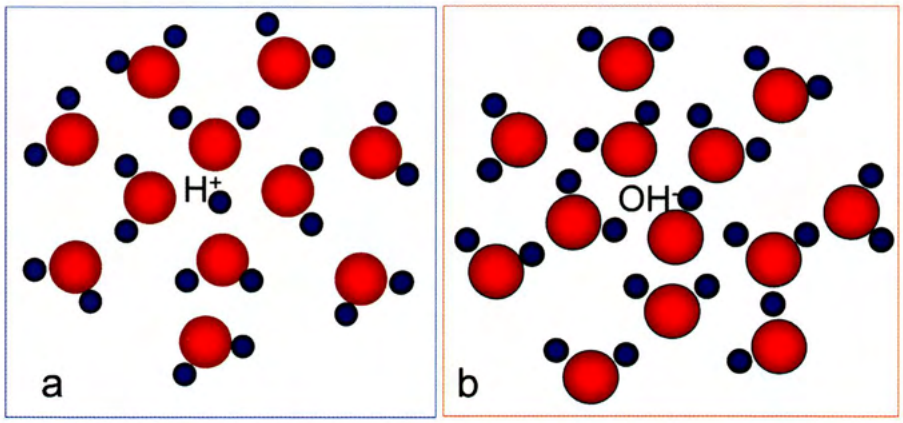

Fig. 15 - Radicals present in water may act as aggregation nuclei. (a) Aggregation with radial symmetry around a H^+ radical, i.e. a single positive charge in the centre. Most water molecules will be oriented with the Oxygen facing the centre. (b) Aggregation with axial symmetry around an OH^- radical, i.e. a negatively charged dipole in the centre. Most water molecules will follow the orientation of the radical in the centre.

the small scale. The physical state of particles is expressed as the sum of two operators, i.e. two mathematical tools, one corresponding to Potential Energy and the other to Kinetic Energy. The Potential Energy exists when there is a force that tends to pull a particle back towards some lower energy positions. Potential Energy represents energy stored within a system and constitutes a physical quantity associated to each point of the space-time continuum, called "*Field*". The mathematical formulation is based on waves describing the density of probability of finding particles in some positions, with energy governed by some quantized levels and the uncertainty principle. Quantum Mechanics provides a unified description with better results than classical physics can do, at least when the number of involved particles is small. When the system is composed of too many particles, the mathematical approach becomes too much complex and hard to solve. Quantum Mechanics applied to theoretical chemistry, i.e. Quantum Chemistry, is aimed to describe the complex interactions of atoms and simple molecules on a larger, but still limited scale. When Quantum Mechanics, or Quantum Chemistry, is able to reach a goal, the result is better than that obtained with classical physics, but the description remains abstract, based on probability, with some obscure or non-intuitive items. For this reason, we will take benefit of some results of Quantum Chemistry, but we will try to present them in the more understandable terms of cause-effect relationships governed by physical laws, typical of classical physics.

Quantum Chemistry has suggested a cluster model for liquid water (Nemethy and Scheraga, 1962) that has been further developed in the course of years. According to this model, liquid water is composed of organised *clusters*, i.e. molecular aggregates composed of a variable number of molecules, also called "*coherent water*", and individual isolated molecules called "*incoherent water*". The adjectives "coherent" or "incoherent" mean that there is, or there is not, any connection, common behaviour or coordination among molecules. Coherent water is composed of a number of molecules temporarily bound among them by intermolecular forces; as opposed, incoherent water is composed of totally independent, lone molecules.

In the liquid phase, water molecules aggregate to each other forming large or small clusters with an internal rather ordered structure that resists for a very short time, e.g. until they impact on another cluster or a freely moving water molecule. It is known that a water molecule is an electric dipole (Fig.14) and we will show an example based on Hydrogen bonds. This will lead to an unstable, short-lived situation. Clusters will preferably form around charged radicals naturally present in water, e.g. H^+ and OH^- that will act as aggregation nuclei. For instance, a H^+ radical (Fig.15a) is the simplest case, based on a radial symmetry, originated by a single positive charge. The Hydrogen ion will attract other water molecules from the Oxygen side, so that the Hydrogen atoms will be external to the first aggregation layer. These external Hydrogen atoms will in turn attract other water molecules, always from the Oxygen side, repeating the same electrical structure until a number of "errors" will make unstable the whole aggregate. The same will happen around the OH^- radical, but with axial symmetry (Fig.15b). The situation is more complex because the OH^- radical is a negatively charged dipole. On one side of the OH^- radical, the Oxygen will attract a Hydrogen atom of some water molecule passing in proximity of it; on the opposite side of the radical, the Hydrogen will attract the Oxygen atom of some other water molecule. This will determine an unstable situation that makes harder the build-up of external layers. Whatever their origin, clusters can freely move within the liquid where they are included, and for this reason the liquid has not a specific shape, but it adapts to the container.

The liquid state exists only in the presence of an external pressure, either generated by an atmosphere of dry gases, or its own (saturation) pressure or a mechanical one. This pressure is necessary to keep together the whole molecular population, as an additional help to the attractive forces to keep together the mass of liquid water counteracting part of the disaggregation activity operated by internal motions. If the external pressure is too small, the liquid boils to pass to the gaseous phase. The phase transition does not only occur at the interface, but also within the liquid, where bubbles of vapour are formed.

In either the vapour or the liquid phase, the water molecules have a ceaseless motion, the average value being determined by the ambient temperature. However, almost all of the molecules individually depart from the average. The Maxwell distribution of molecules by velocities (Maxwell, 1868) holds for both the gaseous and the liquid state, we will tray to apply the above thermodynamic model to liquid water too looking for possible analogies to Quantum Chemistry. The thermodynamic approach will make easier the comprehension of the problem.

The Maxwell law of distribution of molecular speeds

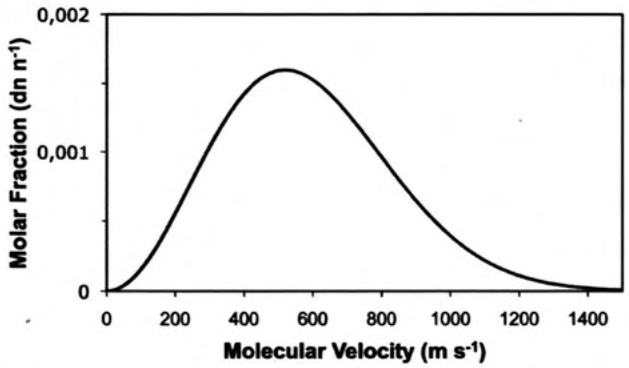

Fig.16 - The Maxwell bell-shaped distribution of molecules by velocities calculated for water vapour at ambient temperature (T=20°C).

(Fig.16) establishes that molecules have diverse individual velocities from very slow to very fast, and explains how the speed is distributed. The result is a bell shaped distribution with the mode corresponding to a kinetic energy in equilibrium with the ambient temperature. In fact, temperature is proportional to the kinetic energy, and to the square of velocities. Under this point of view, a "*fast*" molecule is "*hot*" and a "*slow*" one is "*cold*". If we remember the formula relating the temperature to kinetic energy E_{kin} used in the first definition of temperature, and that the kinetic energy is defined as

$$E_{kin} = \tfrac{1}{2}\, m\, v^2$$

we see that the Maxwell distribution can be represented in terms either of speed or temperature.

This Chapter will discuss the most general aspects of the problem for liquids and later for gases, making reference either to speed or temperature depending on which of them is more convenient for an easier explanation. Liquid and gases have the same distribution but with small differences, e.g. the energy corresponding to the latent heat of vaporization. More details about the Maxwell distribution, the molecular impact frequency on surfaces, the energy transfer and related consequences can be found elsewhere (Camuffo and Pagan, 2006).

In the case of a liquid, differences in velocities constitute a problem for keeping a long-term binding. Clusters can only be established between molecules that move together at the same velocity or within a narrow range of velocities (the most probable velocities are around the mode) and for the limited period in which the internal departures do not accumulate exceeding differences. Looking at the Maxwell bell-shaped distribution of molecules by velocities the population forming the liquid can be ideally divided into three parts as follows.

The first part is composed of some *slow* (i.e. *cold*) molecules near the origin of the distribution. Slow molecules, either singularly or clustered, provide a small contribution because their energy is small. Their role will be greater in forming freezing nuclei at the passage from the liquid to the solid state.

The second part is the central part of the bell, with a huge number of molecules having a *velocity close to the mode* (i.e. *temperate*), which constitute most of the liquid. Many molecules have a speed close to the peak value and can therefore aggregate, at least temporarily, continually forming and renewing *clusters*, i.e. *coherent water*. The probability of forming clusters, their number and size depends on the relative abundance of molecules having the same velocity. The most frequent and massive clusters fall near the mode. When the number of molecules having the same velocity decreases, clusters are progressively reduced in number and size.

The third part is the final tail of the bell, composed of the *fastest* (i.e. *hot*) molecules. In the Maxwell distribution, the number of the most energetic molecules decreases with increasing individual speeds. The minority of fast molecules cannot remain bound to the majority of molecules because of the differences in velocity. Very energetic molecules are necessarily *isolated* and *move freely* within the liquid. This swarm of *vapour-like molecules* can be compared with the lone molecules called "*incoherent water*" in Quantum Chemistry. Within the liquid, these molecules do not longer preserve their identity. When they collide with clusters the violent collision destroys clusters and they loose their elevated speed, and may be included in newly forming clusters. When they collide with other lone water molecules they exchange their momentum, energy and role. Although fast molecules are individually short-lived, their total number remains constant as long as the temperature of the liquid remains the same. Their number increases when liquid temperature rises, and vice-versa. Their continual collisions destroy any form of internal organisation within the liquid. When the liquid temperature drops their number exponentially decreases and at 0°C, when their number is insufficient to avoid any strong organisation within the liquid, water passes from the liquid to the solid phase. The swarm of freely moving molecules can be perceived as a *'vapour' within the liquid*, where it exerts some pressure. At any temperature, the population density of the *'internal vapour'* can be calculated by considering that *at equilibrium the pressure that the internal vapour exerts inside the liquid equals the saturation pressure outside it*. At ebullition, it equals the external atmospheric pressure.

By dividing the number of the vapour-like molecules within the liquid (highly dependent on temperature) by the volume occupied by the liquid (poorly dependent on T), we recognise that the density of the vapour-like molecules is determined by the average liquid temperature. We can also easily extend this concept to the pressure exerted, internally to the liquid, by this gas-like population: things behave as some vapour molecules were dissolved within the liquid, like the ordinary water vapour dissolved into the atmosphere.

The gas-like swarm of fast molecules freely moving within the liquid exerts an own pressure within the volume occupied by the liquid and the fastest molecules have the highest probability of escaping from the liquid. Should this occur, they require some energy to cross the potential at the interface (i.e. latent heat for evaporation). Also the molecules of vapour dispersed in the air, outside the liquid, exert a partial pressure. Of course, in the bulk liquid, the internal pressure of this gas-like population always tends to reach equilibrium with the external vapour, and the driving force is the gradient in pressure across the liquid-gas interface. Should the internal vapour-like pressure be dominant, a flow of vapour molecules would be transferred from the liquid to the atmosphere, and vice-versa. The net flow is governed by the partial pressure gradient across the interface. At saturation the net flow stops. As a consequence, also the pressure gradient will vanish, and the two pressures will be exactly the same. This means that the partial pressure exerted by the vapour-like molecules within the liquid equals the saturation pressure $é_w$ in the air. Also, the saturation pressure $é_w$ is the maximum reachable pressure at that temperature, both within the air and within the liquid.

Once we know that the pressure exerted within the liquid by the swarm of vapour-like molecules is the same as the vapour

saturation pressure $é_w$ measured in the air, we can interpret in a different way the definition of relative humidity, i.e.

$$RH = 100\frac{e}{é_w}$$

i.e. *RH is the ratio between the actual partial pressure of the vapour in the air and in the water.*

In this simple model, the exchange of molecules between air and water is regulated at the interface by the difference in the internal (in the liquid) and the external (in the air) partial pressures although we should remember that, when molecules cross the interface, they release/gain the latent heat for the phase transition. When the internal vapour pressure dominates, the net flux is outwards (evaporation); vice-versa, the excess of external vapour pressure determines condensation. At any instant and at every equilibrium temperature, the external vapour pressure has its own actual value *e*, and the internal pressure is $é_w$ (saturation). We can express the same concept not in terms of pressure difference but of pressure ratio, i.e. *RH*. From this point of view, *RH* is ultimately related to *the intensity and direction of the vapour pressure gradient at the liquid-air interface, i.e. the tendency to evaporate or to condense.*

In the air, the temperature at which saturation tension is reached is T_d. In equilibrium, to every liquid temperature, which equals the air temperature, a specific value of the saturation pressure is associated. Therefore, it makes no difference whether the problem is handled in terms of pressure or temperature.

In conclusion, liquid water too contains a swarm of fast, freely moving molecules that behave like a vapour within the liquid and reach equilibrium with the external vapour. Exchanges between 'internal' and 'external' vapour, i.e. the net flow of mass, are determined by the gradient in partial pressure at the liquid-air interface. *RH represents the ratio between the actual value of the 'external' vapour pressure e to the 'internal' one* $é_w$ (which coincides with the external at saturation), *and gives an idea of the net flow of molecules that may pass from the liquid to the gaseous state* (*RH*<100%), *or vice-versa* (*RH*>100%).

As the saturation pressure is regulated by temperature, and the saturation pressure is reached at the Dew Point, the same concept can be expressed in terms of vapour temperature. We can suppose that the vapour-like molecules within the bulk water are at saturation and that, whatever is the evaporation rate on the free water surface, their loss is always compensated because the water constitutes a huge reservoir to supply such kind of molecules, so that they always remain at saturation. The same can be said in the case of condensation: a number corresponding to the excess of molecules entering the liquid passes to the liquid phase, leaving unchanged the partial pressure of the vapour-like molecules within the liquid. Therefore, the vapour-like molecules within the liquid are at saturation at the actual temperature *T* of the liquid; on the other hand, the partial pressure of the vapour molecules in the air reaches the saturation level only when the air temperature reaches the Dew Point (T_d). Should it be above, the vapour would be over-saturated; below, it would be under-saturated. Therefore, the transition between the two phases is regulated by the Dew Point Spread (*DPS*), i.e. the temperature difference $T - T_d$.

When $DPS > 0$ (i.e. $T > T_d$; e.g., under-saturation in the gaseous phase, or liquid with temperature above the level needed for equilibrium) the transition is from the liquid to the vapour phase, i.e. *evaporation*. The liquid requires the latent heat to evaporate molecules, which will pass to a higher energetic level. The air is enriched with new molecules of water vapour. In the case of an open environment where the vapour molecules are continually dispersed in an infinite space, the evaporation continues indefinitely, or as far as the water reservoir is finished. In the case of a closed space, the evaporation increases the density of the H_2O molecules in the gaseous phase until the saturation equilibrium is reached.

When $DPS = 0$ the equilibrium is reached and the net flux across the interface vanishes.

When $DPS < 0$ (e.g., super-saturation in the gaseous phase, or liquid with temperature below the equilibrium level) the transition is from the gaseous to the liquid phase, i.e. *condensation,* and this releases energy, passing to a lower energetic level.

Somebody might be surprised in finding suggested that vapour-like molecules are dispersed within the liquid. However, this hypothesis is supported by further evidence and some phenomena can hardly be explained without it.

For instance, ebullition occurs when the partial pressure of the water vapour equals the atmospheric pressure. If one supposes that the bulk liquid is composed of molecules in the liquid phase only, and that the vapour is only external to the liquid, in the air, the ebullition can hardly be explained. Why the liquid inside a pot should change behaviour looking at what happens outside the pot? What is the mechanism that relates the violent transformation of phase within the liquid with the external pressure of the water vapour? Why it should be identical to the air pressure? The explanation is simple when one hypothesises the presence of vapour-like molecules inside the liquid, which are continually related to the external (partial) pressure of the vapour. At the boiling point, the (saturation) partial pressure of the vapour in the water equals the pressure of the external air. The bulk liquid is not anymore compressed and is controlled from the exterior by an overwhelming atmospheric pressure. The internal pressure of the vapour is the same as outside the liquid and the whole of the liquid tends to be transformed into vapour and be transferred to the air. This violent and chaotic transformation from the liquid to the gaseous phase within the bulk liquid constitutes the ebullition.

Again: why the ebullition starts from the internal surface of the pot and begins with a characteristic crackling sound? The Kelvin theory (Thomson, 1870) for the formation of droplets, bubbles or micropore filling establishes a relationship between the vapour partial pressure in equilibrium with a curved meniscus of water and the radius of curvature. A plane meniscus is in equilibrium at the ordinary saturation pressure. For a positively curved meniscus, i.e. a droplet, the equilibrium pressure is above the ordinary saturation level, i.e. super-saturation. On the other hand, for a negatively curved meniscus, e.g., the meniscus of water into a micropore, or the meniscus of a bubble into water, the equilibrium pressure is below the ordinary saturation level, i.e. under-saturation.

When water approaches the boiling temperature, a number of microbubbles will form, which are initially flat and adherent to the pot surface. This because the radius of curvature of the flat meniscus is extremely large, nearly infinite, in equilibrium with the saturation pressure. After a short time, new molecules of vapour will feed the microbubbles which will assume a hemispherical shape with a small, negative, radius of curvature. This situation is unstable because a negatively curved meniscus is in equilibrium with a pressure below saturation.

The change in the meniscus curvature will displace the vapour from a condition of equilibrium to that of excess. This vapour will return dispersed within the liquid phase, the microbubbles will collapse and the water will violently invade the space left empty. Every time a microbubble implodes, the water crashes against the pot surface and emits a burst sound and the result is the typical crackling.

The microbubble can grow only in the case that some air originally dissolved into the liquid water enters the bubble at the early stage, forming a stable configuration. This is because the inward pressure of the liquid water, which tends to implode the bubble, is counteracted by the outward pressure of the compressed air pocket. This bubble can grow, being fed by new vapour molecules and without the possibility of bursting inwards.

Similarly, in the case some external particles are introduced in water at the boiling point, e.g., NaCl crystals, the size of the crystals is large compared with the size of the meniscus of the microbubbles in which the effect of the curvature of the meniscus, i.e. the Kelvin effect (Thomson, 1870), is dominant. The crystals will be nuclei for the formation of bubbles, and a violent ebullition will immediately start.

The key conclusions of this section that will be useful in the following are:

- *The liquid phase is a dynamic state composed of temporary aggregates and loose molecules.*
- A swarm of *vapour-like molecules is present within the liquid phase.*
- *The net transfer of molecules between the two phases, i.e. condensation and evaporation, is governed by the gradient in vapour pressure at the gas-liquid interface.* The same mechanism can be expressed in terms of temperature, below the dew point (condensation) or above it (evaporation);
- The saturation pressure within the liquid is governed by the same equation as in the atmosphere, and is a function of temperature only. At the equilibrium, *the same saturation pressure is found both in the gaseous phase external to the liquid, and in the vapour-like molecules inside of the liquid.*

6. The water molecule and its changeable forces

The key point is to understand that the water molecule may change character because it exerts a variety of forces responsible for a number of consequences. When the water molecule is in the gaseous, liquid or solid phase, it changes intermolecular distances and type of forces for the limitations imposed to its moving mode. It is not easy to describe this complex matter and especially to be understandable to people without an adequate background in Physics. To be clear, we will abandon Quantum Mechanics and we will perform crude approximations. We apologize for that. However, we will do our best in using the most elementary concepts of the classical electrodynamics everybody has studied at the high school. The aim of this paper is to help people interested to conservation to understand how moisture acts in deterioration mechanisms.

The water molecule is composed of two Hydrogen atoms and one of Oxygen, i.e. H_2O. The electrons of the two Hydrogen atoms will fill the external Oxygen shell. Each Hydrogen atom will be positively charged (H^+) and the Oxygen will assume two negative charges ($O^=$). This simple model represents the water molecule as an electric dipole that exerts electrostatic and electromagnetic forces that are stronger the shorter the distance.

We start considering a molecule in a fixed position. From the electrostatic point of view, the Coulomb law regulates the electric forces F between two electric charges, i.e.

$$F = \frac{1}{4\pi\varepsilon_o}\frac{q_1 q_2}{r^2}$$

where ε_o is the dielectric constant, q_1 and q_2 the electric charges interacting between them, and r their distance. At long distances the molecule is perceived as neutral (equilibrium between positive and negative charges), but at short distances the position of charges is dominant and this determines how water molecules can interact between them, or with other materials. This is a weakly attractive (or repulsive) electric force that decreases with an inverse proportion of the square of the distance.

From the chemical point of view, the attractive forces can be described in terms of *Hydrogen bond*. The Hydrogen bond is a special type of dipole-dipole bond that exists between an electronegative atom and a Hydrogen atom bound to another electronegative atom. Two molecules of water can form a Hydrogen bond between them. When a number of H_2O molecules are present, e.g. liquid water or vapour near saturation, more bonds are possible because the Oxygen of one water molecule has two lone pairs of electrons, each of which can form a Hydrogen bond with another Hydrogen atom of another water molecule. Hydrogen bonds can repeat from one molecule to another, and form some coordination between water molecules.

When a water molecule adheres to a surface, it exerts in addition the Van der Waals force, derived by the fact that the Hydrogen atoms are deprived of their electrons, but not permanently. In the molecular orbit the electric charge sometimes returns around the Hydrogen causing a continual fluctuation of the electric field that is responsible for this short-distance force.

When the water molecule is in the gaseous phase, free to move, we should pass from the electrostatic to the electrodynamic model. An electric charge q^+ in rectilinear motion with speed v is equivalent to an electric current i on a wire and, if negative, i.e. q^- to $-i$ (Lorentz law). It exerts an electromagnetic force perpendicular to the movement of the charge, which is equivalent to a current on a rectilinear wire, which is inversely proportional to the distance (Biot and Savart Law). In this specific case, the electromagnetic forces generated by the rectilinear motion of the two positive H^+ are counteracted by the opposite field generated by the $O^=$ which has double charge (Fig.17). No net electromagnetic force results from the rectilinear motion of this molecule.

However, the kinetic energy is equally partitioned between all degrees of freedom, and rotation has the same energy as translation. The mass of the Hydrogen atom is 1, of the Oxygen 16 and the H_2O molecule 18. The centre of mass of the H_2O molecule is very close to the centre of mass of the Oxygen atom. As a consequence, when the H_2O molecule rotates, the Oxygen does not describe any orbit, or a very small one, around the centre of mass of the H_2O molecule. As opposed, the two Hydrogen atoms will have orbits (only a finite number is possible in Quantum Mechanics), depending on the H_2O rotation axis, and their actual position in the rotating molecule. About the position, if we assume as zenith the axis of rotation, we range from the equatorial position, i.e. both Hydrogen atoms rotate on the same horizontal plane describing the largest orbit (Fig.18a), to the polar position, in which one Hydrogen atom is exactly on the axis of rotation, i.e. no orbit at all, and the other Hydrogen atom is describing a small orbit (Fig.18b) The strongest magnetic field is created by the equa-

torial rotation and the weakest by the polar one. The magnetic field generated by each of the two Hydrogen atoms on their orbits around the rotating water molecule is equivalent to a magnetic lamina, with Southern and Northern faces determined by the clockwise or counter-clockwise rotation. The intensity of the field depends on the speed of each atom on its orbit and the orbit diameter as well.

Any H_2O molecule in the gaseous phase has its own magnetic field, larger or smaller, positive or negative, depending on how it rotates, i.e. rotation axis and speed. The result is that any H_2O molecule exerts attractive or repulsive electromagnetic forces with other H_2O molecules and materials as well, but the intermolecular forces depend on the matching of the motion (especially rotation) modes. This may induce a stronger or weaker coordination between the molecules of the water vapour. When intermolecular forcers are weaker, also the coordination is weaker, and the water vapour may behave in a way similar to the well-known approximation for perfect gases. This situation is typically found when the water molecules are very distant from each other, or when some other factor will work against the coordination forces, as we will see later. Ambient temperature, frequency of molecular collisions, thermal disturbance, intermolecular forces and the average electromagnetic coordination are all related to each other as well as to different *RH* levels. *RH* is a physical variable reflecting how much the potential energy of intermolecular interaction governs the molecular population as a whole, including weak clustering and mutual orientation of molecules. When two H_2O molecules are close to each other, the interaction deflects their path. The balance between the inertia force of flying molecules and the weaker or stronger intermolecular forces determine the *trajectory deflections*. When the inter-molecular forces become stronger, the water vapour changes character and becomes an aggressive swarm of coordinated molecules that may exert many adverse deterioration mechanisms.

In conclusion, we have seen that:

- *Water molecules change type of forces when they are free to fly and rotate in the air or have a limited motion because they are adsorbed on a material surface.*
- In the vapour phase the molecular rotation makes impossible Hydrogen bonds.
- In the vapour phase, the forces exerted by water molecules depend on the rotation mode; they may be stronger or weaker, with a number of important consequences.

7. The gaseous state, the Maxwell distribution and the three classes of vapour

The water vapour molecules in the Maxwell distribution can be subdivided into three classes that reflect the three phases of water, i.e.: frozen, liquid, gaseous (Fig.19). Each class is based on the consideration of what would happen if a molecule impinges on a surface at the same thermal level, as follows.

The first class, the 'Frozen Type' (FT), covers the interval 0-228K where water exists only in the ice phase. The upper limit (228K) is the lowest temperature to which the properties of liquid water can be extrapolated (Pruppacher e Klett, 2000), at the limit of the supercooled water in cloud droplets (Mason, 1971) and in porous materials (Fagerlund, 1973).

FT is composed of molecules with very low kinetic energy and may potentially pass to the ice phase if they hit any surface at the same temperature. These molecules will hardly interact between each other and with materials (Warnatz et al., 2001) for four reasons.

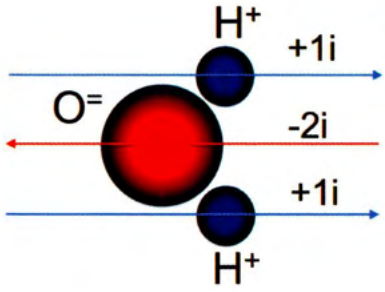

Fig. 17 - Equivalent electric currents generated by a water molecule (H_2O) in a rectilinear motion, in this drawing from left to right. The two Hydrogen atoms (blue), each with a positive charge, are equivalent to two wires each with current +i; the Oxygen atom (red) with double negative charge to -2i. The induced electromagnetic forces compensate for each other and the net effect is null.

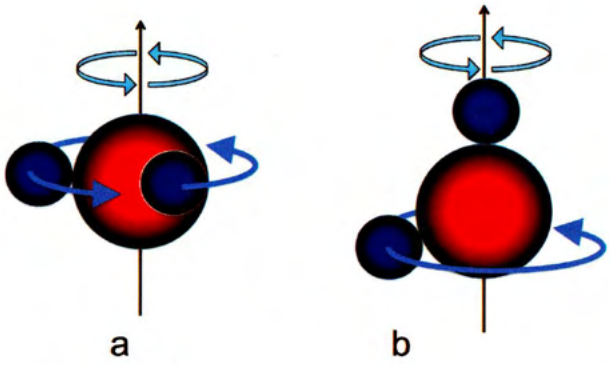

Fig. 18 (a) - Rotation of the water molecule (H_2O) in the Equatorial Mode, i.e. with the two Hydrogen atoms (blue) describing the largest orbits around the Oxygen (red).
(b) - Rotation in the Polar Mode, in which a Hydrogen atom is on the rotation axis (no orbit), and the other Hydrogen describes an orbit around the Oxygen. Orbiting charged particles are equivalent to magnets. Rotations evidenced with arrows.

i) Having a low kinetic energy, the same holds for the molecular rotation and the related field of electromagnetic forces generated around them. Having a weak electromagnetic field, they will hardly interact between each other and with materials.

ii) The rate of chemical reactions *k* given by the Arrhenius equation depends to a large degree on temperature. Common reactions at ordinary temperatures are highly improbable in the FT zone, because in many cases the molecular energy falls below the activation threshold, i.e. the energy barrier which has to be overcome during a reaction.

iii) The adsorption rate k_{ads} also decreases with decreasing temperature:

$$k_{ads} = S\sqrt{\frac{\Re T}{2\pi M}}$$

where S is a sticking coefficient, $\Re$ the gas constant and M the mass of the sticking particle.

iv) Molecular viscous forces tend to slow down velocities in the laminar sub-layer around bodies (Kundu, 1990). Vapour molecules with very low kinetic energy are unlikely to cross this sub-layer and to reach the body surface.

FT covers 49% of the whole molecular population at 20°C but is composed of hardly reactive molecules, which pass substantially unobserved. Of course, the actual number of molecules per cubic metre is obtained by multiplying their normalised number by the absolute humidity at saturation, and then by *RH*.

In comparison with the two following classes, which are characterised by higher temperatures, the molecules in FT have

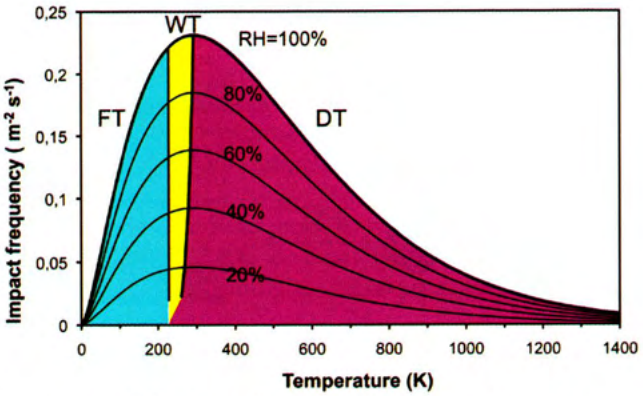

Fig. 19 - Maxwell distribution of molecular hits. Every second any surface experiences a number of impacts with water vapour molecules. The number of hits per unit time and surface has been calculated with the Maxwell distribution of gases by thermal energy for every molecule of the whole water vapour population. According to this distribution, hitting molecules can be divided into three classes: Frozen (FT, cyan), Wet (WT, yellow) and Dry (DT, magenta) Type. The number of impacts and the individual effects change with the class type. The top thick line is for RH=100%, lower thin lines for RH = 80, 60, 40, 20%, respectively. This example is calculated for water vapour at 20°C room temperature.

the lowest kinetic energy and the lowest interaction capability. Their low energy falls below the activation threshold for many physical, chemical and biological processes.

The second Class, the 'Wet Type' (WT) ($228K<T<T_d$), lies in the central part of the distribution and is the least populated class, i.e. generally less than 12%, with the number of molecules strongly dependent on *RH*. On the other hand, the density of population is very high. The upper limit of this class is constituted by the Dew Point T_d which depends on the amount of vapour mixed to air.

Due to the disturbance generated by thermal motions, molecules in WT have an electromagnetic orientation capability that is smaller than in class FT, but is still consistent. Class WT is composed of potentially liquid, unstable molecules. When they become into contact with a surface at the same temperature, they change phase. For this reason WT is the most relevant class for interactions with materials.

A variation of T_d implies a similar variation in the number of metastable molecules. The Maxwell distribution shows that the central part, i.e. the most densely populated part, including molecules with root mean square values of T in equilibrium with the ambient, plays a fundamental role in determining the character of the whole ensemble. T_d ranges just within this critical region, so that even any small displacement of T_d has a relevant impact in changing the nature of the interactions (i.e. wet type or dry type) with materials. Consequently, *RH* is related to a rivalling balance between the liquid-like and the gas-like fraction of molecules impinging on materials.

Two mechanisms may induce a reduction of the WT molecular population at the passage from saturation to lower *RH* levels. On one hand, moving from up to down, the concentration decreases proportionally to the lowering of *RH*. In fact, at ambient temperature and pressure constant, the proportionality is good, e.g. at T =293K it is better than 2% in the interval $20\leq RH \leq 100\%$ (Camuffo, 2004). On the other hand, moving from right to the left, it depends on the actual value of T_d. In the span $10\leq RH\leq 100\%$ the molecules in class WT lie between 6.8-11.8% of the whole population.

The third Class, the 'Dry Type' (DT) ($T > T_d$), is densely populated. In the humidity range $10\leq RH \leq 100\%$, class DT covers 39.2-44.2% of the whole population, the lower limit is T_d. Electromagnetic forces increase with the molecular speed v, but the interaction time is inversely proportional to v and the kinetic energy increases with the square of v. These fast molecules have the typical behaviour of the gaseous phase, with neutral equilibrium, the disturbances by violent thermal motions being the dominating factor. As a result, DT molecules behave like a gas, bouncing on surfaces with a very low probability of reacting with materials.

More details about these three classes, molecular impinging frequency and energy transfer can be found elsewhere (Camuffo and Pagan, 2006).

In conclusion, from the Maxwell distribution, the key emerging factors are:

- The *properties of the physical states are determined by the balance between kinetic and potential molecular energy*.
- *Not all of the molecules within the water vapour population behave in the same manner.*
- The "coldest" part of the molecular population has an irrelevant role, the "warmest" part behaves like a gas and is also irrelevant; *only a small part of molecules in the mid of the distribution is in metastable equilibrium, potentially available to pass from the gaseous to the liquid phase and react with materials*. Only these molecules are responsible for the various deterioration mechanisms.

8. Relative Humidity and intermolecular distances

In the case of real gases some weak intermolecular forces are exerted at short distances and a potential is defined to describe this situation. Two molecules cannot approach each other at a distance smaller than an effective value ID_o (positive potential for repulsion). At intermediate distances molecules weakly attract each other (negative potential). Only when molecules are far away from each other (i.e. rarefied gas) or the kinetic energies are very large (i.e. hot gases) intermolecular forces vanish (null potential) and the perfect gas approximation becomes possible. In the case of water vapour, intermolecular forces increase in strength with *RH* level, favouring intermolecular interactions between flying molecules. Even some clusters (dimers, trimers and higher-order polymers) may form, though in small concentrations only (Zeidler, 1983; Pruppacher and Klett, 2000).

Intermolecular forces are strongest when the vapour approaches saturation and transition to the liquid phase. Cooling the vapour, i.e. slowing down the molecular motions, the instability of the vapour system is augmented. In the liquid phase, intermolecular forces strengthen to hold together molecules, although with a certain mobility. In addition, kinetic energies are comparable with potential energies. In the solid phase, intermolecular forces are much stronger, potentials become dominant and motions are forbidden except for vibrations of particles around their equilibrium position in the crystalline lattice (Atkins, 1990).

In water vapour, a balance is established between kinetic and potential energy. On one extreme, far from saturation, kinetic energy dominates, molecules are rarefied and/or with a high thermal level, intermolecular forces are very weak and the vapour behaves like a gas. In the opposite case, the potential establishes mutual interactions that weakly involve the molecular population as a whole and may form some internal clustering. Intermolecular forces slightly strengthen with the degree of saturation and have an inverse dependence on the thermal level. When molecules approach each other a weak correlation between them is established.

The potential energy of inter-molecular interaction is a com-

plicated function of the angle between molecules; the probability of a particular orientation is given by the Boltzmann distribution, related to the temperature. At higher temperatures, greater thermal motions disturb or overcome the mutual orientating effects of dipoles. At lower temperatures and shorter inter-molecular distances (*ID*), interactions are stronger. Near saturation, when *ID* are shortest and/or kinetic energies lowest, intermolecular forces reach their maximum, the dipole-dipole and magnet-magnet correlation is highest, the molecular population as a whole is characterized by a certain internal organisation and some molecules may undergo a phase transition. This explains why high *RH* levels favour chemical reactions.

For every temperature *T*, the minimum *ID*, i.e. ID_w is established from the smallest volume available per each molecule V_{unit}, i.e. the total volume V_t divided by the highest number of molecules. V_{unit} is calculated from saturation absolute humidity $AH_w(T) = m_{vw}/V_t$, which represents the highest density the vapour can reach at that temperature. It is possible to pass from mass density to population density by multiplying $AH_w(T)$ by N_A/M_{mol}, where $N_A = 6.023\ 10^{23}$ molecule mol^{-1} is the Avogadro number and M_{mol}=18.01 g mol^{-1} the molar mass. ID_w is obtained as the cubic root of $V_{unit}(T)$, i.e.

$$ID_w(T) = \sqrt[3]{\frac{M_{mol}}{AH_w(T)\,N_A}}$$

ID_w has an inverse dependence on temperature. ID_w has been calculated with *AH* at saturation, i.e. AH_w when *RH*=100%, but it can be obtained in correspondence of the various *RH* levels if we multiply AH_w in the above formula by the actual fraction represented by the *RH*, i.e.:

$$ID(T,RH) = \sqrt[3]{\frac{M_{mol}}{AH_w(T)\,RH\,N_A}}$$

The minimum allowable intermolecular distance for water vapour is shown for various *RH* levels (Fig.20). If we remember that H_2O has 3.1 Å molecular diameter (1 Å = 10^{-8} cm), the ID ranges from 60 molecular diameters at *T* = 273K, i.e. 0°C, to 23 diameters at *T*=323K i.e. 50°C. We can also write *ID* in the simplified form (Camuffo and Pagan, 2006), i.e.:

$$ID(T,RH) = \frac{A(T)}{\sqrt[3]{RH}}$$

where *A(T)* is a temperature-related coefficient:

$$A(T) = -0.0015T^3 + 0.2473T^2 - 18.968T + 845.18$$

and in the interval 273-323K *A(T)* ranges from *A*(273K) = 845.44 to *A*(323K) = 328.36. Writing the above equation in terms of *RH*, i.e.

$$RH = \frac{A^3(T)}{ID^3(T,RH)}$$

we recognize that *RH is a variable that has an inverse dependence on the third power of intermolecular distances and reflects, for each temperature, to which extent the vapour is approaching the critical intermolecular distance for phase transition*. This justifies the relationship between *RH* and the reactivity with chemical substances in the gaseous phase.

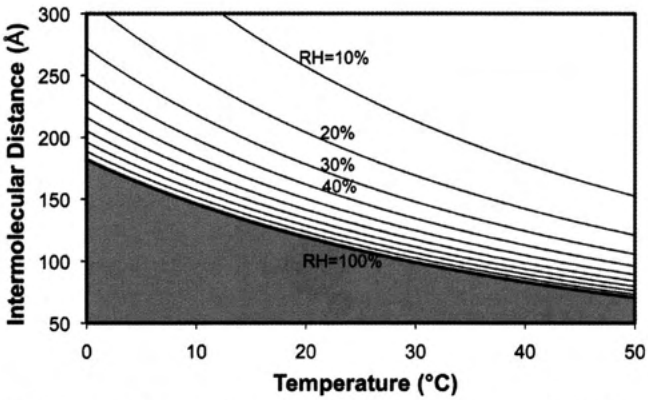

Fig. 20 - Intermolecular distance of water molecules at saturation (thick line) and at selected RH levels, i.e. RH= 10, 20, 30...%. The grey area is not allowed for a stable equilibrium and may host the liquid and the gaseous phases in transition between each other. 1 Å = 10^{-8} cm.

In conclusion:

- *The closer the H_2O molecules, the stronger the attraction forces between each other.*
- *The H_2O molecules cannot approach to each other below a given threshold*, which is determined by the balance between attractive forces (i.e. the potential) that tend to establish some order, and kinetic energy (i.e. the temperature) that tends to generate chaos and destabilise any order.
- The shortest intermolecular distance has been calculated for any temperature, as well as the average distance that water molecules will have at the various *RH* levels. *RH is the physical variable that represents, for each temperature, how much the average intermolecular distance of water vapour is approaching the threshold for changing phase.*

9. The formation of water monolayers on the object surface

A simple electrostatic model has been useful to describe how a water molecule is attracted by the surface of materials composed of real or virtual dipoles, and how it forms layers of adsorbed water on such surfaces. Now we will show a few examples of how monolayers of solid water build-up on a hydrophilic (i.e. wettable) surface.

In the case of ionic compounds, e.g. marble, limestone, carbonatic sandstone and mortar the material structure is characterized by the dipole Ca^{2+} $CO_3^{=}$. The water molecule is attracted from the Oxygen side (i.e. $O^{=}$) by Ca^{2+}, and from the two Hydrogen side (i.e. H^{+}) by $CO_3^{=}$, the latter being more probable. The first monolayer is likely based on the attraction between H^{+} and $CO_3^{=}$ (Fig.21) or, alternatively, between Ca^{2+} and $O^{=}$. Similarly, granite, silica sandstone, basalt, glass have a structure based on the dipole Si^{2+} $O_2^{=}$, where we can repeat the above mechanism.

Metals are atomic compounds without pairs of charges, but the water molecule may generate a mirror image of it for electromagnetic induction. The mirror image has inverted polarity, and this generates an attraction between the water molecule and its mirror image. Once the first monolayer has been formed, its surface is electrically charged and will build-up other layers.

As opposed, a hydrophilic surface can be transformed into a hydrophobic (i.e. non-wettable) one with the help of substances composed of molecules what are polar from one side and not from the other one, e.g. siloxane $Si_2O(CH_3)_4$ (Fig.22). In such example, the siloxane molecule is attracted from the O^{-} side towards any material with a polar surface and remains

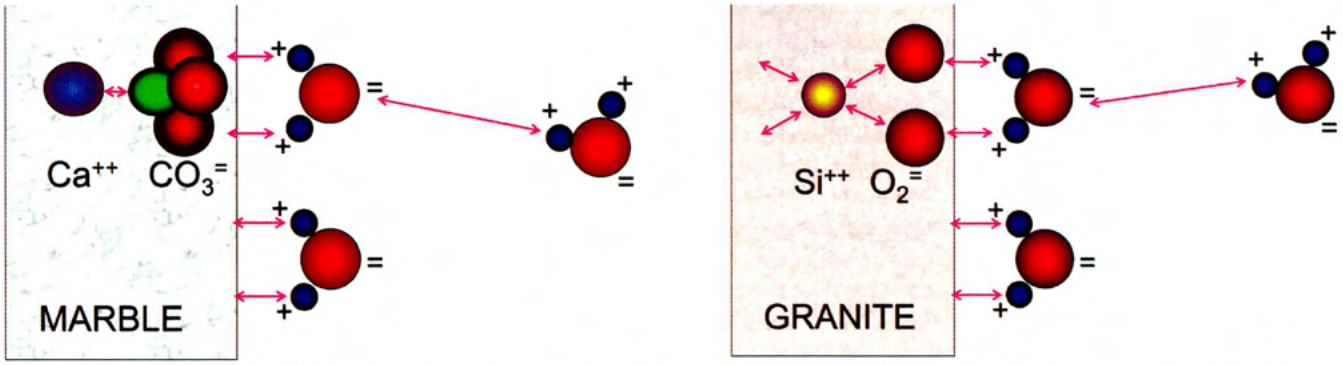

Fig. 21 - (a) Formation of the first H_2O monolayer on a marble, limestone or plaster sample. This example is based on the attraction between H^+ and CO_3=. Once the first layer is formed, other H_2O molecules will be attracted by the back charges, and so on until equilibrium is reached between adsorbed water and ambient RH. (b) As in (a), but for a silica (SiO_2) based stone (e.g. granite, silica sandstone, basalt) and glass.

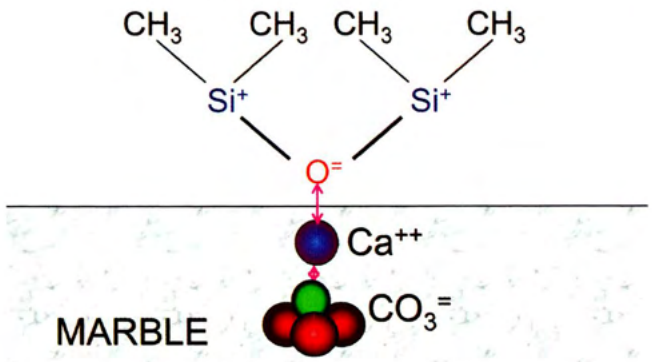

Fig. 22 - How siloxane $Si_2O(CH_3)_4$ transforms a hydrophilic surface (e.g. marble) into a hydrophobic one. The polar marble surface attracts the siloxane from the O= side. The Oxygen atom remains bound to Calcium included in the stone lattice; the methane radical side (CH_3), not electrically charged, remains external with no attractive forces. The surface was with polar charges, after treatment it is not.

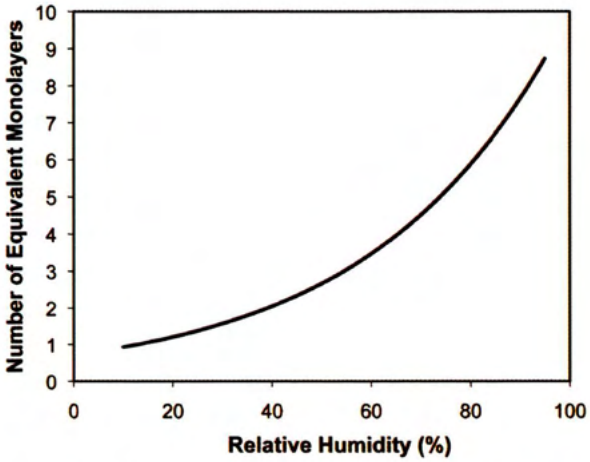

Fig. 23 - Number of Equivalent Monolayers of water forming a film on Silver at 25°C (drafted after experimental data by Leygraf and Graedel, 2000).

bound to it. The non-polar organic side composed of organic radicals, e.g. methane (CH_3), remains external, oriented towards the air and unable to attract water molecules, being not electrically charged. The surface of the treated material is covered with a non-polar barrier, has no more tendency to form Hydrogen bonds with water molecules and cannot be wetted.

When the first monolayer of solid water is formed on a hydrophilic surface, e.g. with dipole based structure, the water molecule can repeat the surface charge of the material and its electric attraction, although with weaker intensity on each additional molecular layer of bound water that will be formed. The forces between the bound water in the first layers and the material surface are so strong, and the H_2O motions so limited, that we should consider the adsorbed water in the *solid phase*, at least when we deal with the monolayers closest to the surface. The very first monolayers are not reactive and behave as a varnish coating on the surface. When a number of molecular layers have formed on a surface, the attractive forces between such layers and the surface become weaker and weaker and *the water gradually passes from the solid to the liquid state*.

We can easily imagine that the first monolayer of solid water forms on a polar surface (i.e. a surface with electric charges) because the water molecule is an electric dipole that is attracted and then bound by another dipole having opposite charge, either real or induced. The first monolayer repeats the same charges on the external layer, so that the mechanism might indefinitely continue and at first sight it is not clear why the number of layers is limited and related to the ambient *RH*. However, the solution is fund if we remember that the *RH* level involves two physical properties of the water molecule population, i.e. average molecular speed and distance, which in other terms determine the capture efficiency of flying molecules. High *RH* means slow molecules or high molecular concentration, or a combination of the two and, consequently, high probability of capturing molecules from the gaseous phase and transforming them in an ordered layer on the surface. Low *RH* means fast molecules or a low concentration, or a combination of the two and, consequently, low probability of capture. If we add the loss of surface charge for a growing disorder as a consequence of the building up of external monolayers, we easily recognize that the building up process is destined to be limited to a number of water layers. Their number depends on the actual *RH* level, and an endless deposition of water is possible only on surfaces at temperature below the dew point, i.e. at super-saturation conditions.

The experimental studies conduced by Leygraf and Graedel (2000) on metals show that the number of equivalent monolayers (*NEM*) of water grows exponentially with the ambient *RH*. For instance, the equation we have calculated for *NEM* on silver at 25°C (Fig.23) is

$$NEM = 0.7153\exp(0.0263RH)$$

If we express *RH* in terms of *MR* and temperature, the above equation can be written in explicit form of the temperature dependence (symbol *t*, expressed in °C), i.e.:

$$NEM = \frac{0.7153\exp(0.0263\ MR\ 10^{-at\,(h+t)})}{0.03795}$$

where the exponential coefficients are $a = 7.5$ and $b = 237.3$°C, i.e. the Magnus coefficients (Camuffo, 1998, 2004).

We see that silver is covered from 1 to 8.7 monomolecular layers of water in the 10 to 90% *RH* range, making easy tarnishing at high humidity. Under such conditions, the most external "liquid" layers are able to capture any aerosol impacting on them, furnish water to form a solution and start corrosion. Similarly, atmospheric Oxygen may dissolve on them and start oxidation.

A similar behaviour has been observed on glass. When we supply light to an optic fibreglass, the light beam is transported with multiple reflexions inside the fibre and leaves it from the opposite side, i.e. the distal end where the fibre is truncated. However, a small portion of light is reflected back from the distal end surface. When some water monolayers form on this surface, the backlight intensity decreases (Fig.24). The backlight attenuation starts to become visible at *RH* >50% and is more and more relevant in the 75% to 99% *RH* interval. Condensation causes a stronger backlight attenuation, which drops to about 50% of the initial intensity level it had in dry conditions (Bergsten et al, 2010).

Again, water-absorbing materials with internal moisture content in equilibrium with ambient *RH*, at high *RH* levels will have enough free water adsorbed on the surface, or filling pores, to allow microbiological life and be covered with moulds.

In conclusion, the key to interpret many deterioration mechanisms is to remind that *hydrophilic materials are always covered with a number of monolayers of water whose number and reactivity is related to RH*. In addition, *porous or water-adsorbing materials have a further capacity of accumulating water inside their microcavities*.

At low RH adsorbed water is a small quantity but strongly bound at the solid state, hence not reactive. As far as ambient *RH* increases, the amount of adsorbed and free water grows. The excess water approaches the character of the liquid phase and is available for many physical, chemical and biological deterioration mechanisms.

When the material and the air have the same temperature, the moisture content of the material is governed by the equilibrium with the *RH* in air. However, when their temperature is different, the process is governed by the *RH* resulting from the temperature of the object, not of the air. If we want to reduce the water amount and the *RH* affecting an object, *we should increase the temperature of our object*; an increase of air temperature not followed by a similar increase in the object temperature is a misleading remedy. Condensation on cold surfaces is an extreme example of the same problem. *The very problem is not the moisture and the RH in the air, but the equilibrium moisture content affecting our material, and the physical state of the water on the material, both determined by the temperature of this material.*

The state of "equilibrium", i.e. air and all objects having the same temperature, is rarely met and often misleading because the general case is characterized by a "*dynamic equilibrium*" in which air and objects are continually exchanging heat because they are at different thermal levels and tend to establish a thermal homogeneity, almost never reached.

When we perform environmental monitoring, we should remember that *our task is to preserve objects, not the "air"*. The first aim is to assess: how much the temperature of individual objects departs from the air temperature; how homogeneous the indoor temperature distribution is; how much indoor temperature changes over time causing unbalances in the thermal inertia of objects and structures or in the internal air circulation. At this point, once the moisture content in the air is known, it is possible to consider which *RH* level will be in equilibrium with any object. The knowledge of such a level and its variability is useful to assess if objects are preserved in risky or safe conditions.

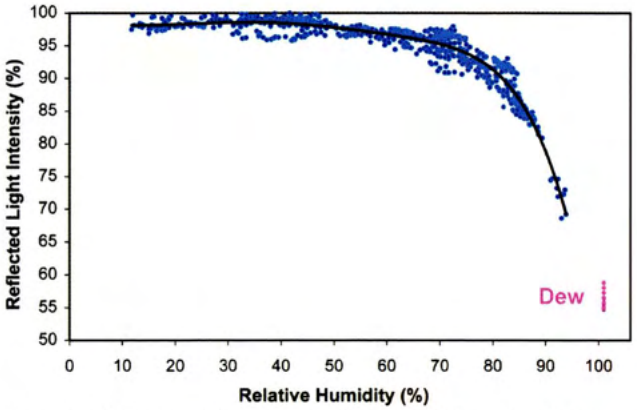

Fig. 24 - Backlight from an illuminated fibreglass. Backlight readings before saturation: blue dots; backlight after condensation (dew): pink dots. Black continuous line: 6th order polynomial interpolation of blue dot readings. The decrease in the backlight is due to the formation and growth of monolayers of water, starting at RH >50% and dropping at RH>75%. Condensation causes a deeper drop reaching about 50% of the dry level.

10. Condensation of liquid water on the object surface

Condensation of liquid water occurs when the surface temperature (T_s) drops below the dew point. The period of time in which T_s-T_d <0 is called Time-of-Wetness (*ToW*). On non-absorbent materials condensed water is visible in form of drops and dewing. Porous materials absorb condensed water and suck it inside the capillary fringe and condensation is not visible.

The amount of water (W_c) that condenses on a cold surface with $T_s < T_d$ is

$$W_c = A(U)\, MR \times (T_d - T_s) \times ToW$$

where $A(U)$ is a proportionality coefficient that varies with the ventilation speed U. In still air $A(U)$ was determined to be $A(0) = 0.194$; in stirred air, $A(U) = 0.756$. Used units as follows. W_c: mg cm^{-2}, *MR*: g kg^{-1}, T_d: °C, T_s: °C, *ToW*: hr. This means that the water condensed on a surface is proportional to the concentration of moisture in the air, the drop of temperature below the dew point and the duration of the time-of-wetness. Air motions favour exchanges and increase the amount of condensed water (Camuffo and Giorio, 2003).

Indoor condensation typically occurs on cold surfaces, e.g. windowpanes or north facing walls, especially in crowded rooms as a consequence of the moisture released by people. Outdoors, the most typical case is the dewing for nocturnal cooling in the presence of clear sky. Any surface loses and receives heat via *IR* radiation. However, the thermal balance is not the same for all surfaces. We will distinguish the opposite cases of horizontal and vertical surfaces.

Horizontal surfaces emit *IR* radiation on the vertical direction. Most of this radiation crosses the atmosphere and is dispersed outside it, in the interstellar space. They emit much more *IR* than they receive back and cool more rapidly than surrounding air. The balance is a net loss. Their temperature may drops some 4° or 5°C below the air temperature and the dew point, if the air temperature reaches T_d.

As opposed, vertical surfaces emit *IR* on the horizontal direction. The emitted *IR* is soon or later absorbed by aerosols or the landscape and is almost balanced by other *IR* travelling in the opposite direction, and originated from aerosols or other bodies at the same or similar temperature. The balance is a

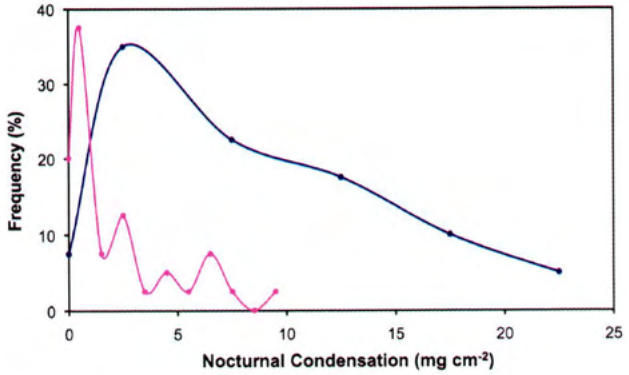

Fig. 25 - Water condensing in clear nights on a horizontal (blue) or a vertical (pink) surface. Observations are reported in terms of the frequency in which selected condensation intervals have been found on specimens of Carrara Marble.

weak loss, for the colder soil influence. Vertical surfaces remain in a closer thermal equilibrium with the surrounding air, e.g. some 2°C below air temperature.

For this reason nocturnal cooling and dewing are heavier on horizontal surfaces, almost negligible on vertical ones. With a clear sky, nocturnal condensation observed in Italy on specimens of various types of marble, limestone and bricks was generally in the range of 10 to 20 mg cm^{-2} on horizontal specimens and from 0 to 10 mg cm^{-2} on vertical ones (Fig.25) (Camuffo and Giorio, 2003).

The most intense dewing in Italy was observed in the cold season when the warm and humid Sirocco wind, blowing from Northern Africa, meets cold monuments. The same situation, but with smaller temperature spread, is typical in spring when monuments are still cold for their long thermal inertia, and the seasonal winds are reach in moisture.

11. Temperature and Relative Humidity to preserve collections

A number of papers, handbooks and standards have recently been written (e.g. Michalski, 2000; 2009; Tétreault, 2003; ASHRAE, 2007, CEN prEN15757, 2010) or are under development concerning the microclimate recommended for preservation of cultural heritage. An evolution has been noted in the last years and also some convergence, although some conflicting ideas still survive.

The key problem is to determine the most appropriate temperature and *RH* levels, and their allowed variability, to preserve collections. The answer is not easy, especially because collections may include a number of materials having different needs. In such a case we should find a compromise, or have a choice to protect the most vulnerable, or the most valuable item. In practice, the correct question is not *"what is the best?"* but *"what solution minimizes damage?"*. This necessarily requires a holistic analysis of the problem, and a subjective final decision.

However, some objective answers can be found on the ground of the above findings.

From the Arrhenius equation we know that low temperatures reduce chemical deterioration, e.g. hydrolysis (e.g. cellulose acetate or nitrate), oxidation, corrosion (e.g. metals) and any other chemical reaction. In addition, *RH* should be kept at low levels, because in dry conditions hydrolysis and any other chemical process proceed at an extremely slow rate. The same can be said for any microbiological deterioration. In conclusion, a cold and dry environment is the best situation against both chemical and microbiological decay, and such microclimate conditions should be preferred when these mechanisms are the key problem.

A number of soluble salts become deliquescent at some relative humidity levels, e.g. NaCl at *RH*=70%, and humidity cycles around the deliquescence threshold may cause dangerous dissolution – crystallisation cycles. This problem is particularly relevant for masonry when heating-cooling cycles force crystallisation – deliquescence cycles in the mortar. When this is the problem, a dry environment is the best. However, if the "normal" situation is humidity above the threshold, the problem becomes how to heat the building, because any heating operation lowers the relative humidity and may induce a dangerous cycle.

A number of minerals should be preserved under some specific conditions or intervals, to avoid critical temperature and relative humidity thresholds for transformation from one crystalline species to another, maybe with a different number of crystallisation water, e.g. gypsum ($CaSO_4 \cdot 2H_2O$) and anhydrite ($CaSO_4$).

The most difficult situation is found with a number of hygroscopic materials , e.g. wood, for which the risk of physical damage, e.g. swelling-shrinkage, is more relevant that any other form of chemical or microbiological deterioration. Adsorption isotherms, strain-stress diagrams, acoustic emission and other laboratory test can recognize the well-being, the alert and the risky areas in temperature and relative humidity diagrams in order to individuate recommended average levels and variability range. However, the problem can easily be solved for the preservation of the *material samples* used for the laboratory tests. The problem regarding the preservation of *real objects* is different. For instance, a painting on panel kept for centuries in the same building has adapted to that particular microclimate with some internal deformations. Some cracks may also have formed to respond to the specific environmental variability as expansion joints. Any change in average levels, or in the variability range or frequency, may have a severe impact. We should also be reminded that the damage does not necessarily appear visible to the naked eye, but starts in the material's interior with microcracks that eventually grow, propagate and finally appear on the surface after a number of cycles or after some time. For this reason, for such hygroscopic materials, it is recommended to keep the so-called *"historic climate"* unchanged, i.e. the natural microclimate conditions in which the object has been preserved for a long time and has adapted to.

The most critical situation is found in heated buildings, especially in the case of churches. In cold regions continuous heating lowers too much relative humidity, and air humidification is not convenient because added moisture condenses on cold walls generating moulds and soluble salt mobilisation. Also occasional heating is dangerous for the enhanced temperature and humidity cycles. It has been demonstrated that thermal comfort and cultural heritage preservation are conflicting needs. For this reason a compromise between the two needs is necessary, that means a limitation in heating (Camuffo et al., 2007; 2010a; CEN prEN15759, 2010). On the other hand, cultural heritage preservation and energy saving are strictly related to each other because when a moderate heat is supplied, the natural microclimate will suffer for a limited perturbation. The problem can be mitigated with local heating, i.e. by leaving the building structure unaffected and keeping heat concentrated in the area where people stay.

Acknowledgements

This paper was made possible thanks to the European Science Foundation funding and the scientific activity performed within the COST action D42: *Enviart*. The paper reports results achieved during EU funded projects, namely AER (ENV4-CT95-0088), FRIENDLY-HEATING (EVK4-2001-00007), SENSORGAN (EVK4-2001-00007) and *Climate for Culture* (GA 226973). Special thanks are due to all my coworkers and, in particular, for the most recent period: Dr Emanuela Pagan for Maxell calculations, Dr. Roberta Giorio for tests on condensation, Dr Antonio della Valle and Dr Chiara Bertolin for field surveys. Special thanks to Prof. Arnaldo Longhetto, University of Turin, for useful discussions and suggestions.

APPENDIX

Equations to compute relative humidity and related variables

The following equations enable the calculation of relative humidity and other related variables from three basic parameters, including atmospheric pressure p. In the case of approximate calculations and in the absence of a barometric reading, p can be substituted with 1013 hPa. Details on the derivation, explanation and accuracy of the formulae and the transmission of errors can be found elsewhere (Camuffo, 1998, 2004).

1. Instruments: Psychrometer, barometer

Input readings: air temperature t (°C), wet bulb air temperature t_w (°C), atmospheric pressure p (hPa).

- Water vapour pressure e (hPa)

$$e = 6.112 \times \left(10^{\frac{7.65 t_w}{243.12 + t_w}} - 1.068 \times 10^{-4} p (t - t_w) \right)$$

- Mixing ratio (g/kg)

$$MR = 3801.5 \times \frac{10^{\frac{7.65 t_w}{243.12 + t_w}} - 1.068 \times 10^{-4} p (t - t_w)}{p - 6.112 \times \left(10^{\frac{7.65 t_w}{243.12 + t_w}} - 1.068 \times 10^{-4} p (t - t_w) \right)}$$

- Absolute humidity (g/m^3)

$$AH = 1344.6 \times \frac{10^{\frac{7.65 t_w}{243.12 + t_w}} - 1.068 \times 10^{-4} p (t - t_w)}{273.15 + t}$$

- Relative humidity (%)

$$RH = 100 \times \frac{10^{\frac{7.65 t_w}{243.12 + t_w}} - 1.068 \times 10^{-4} p (t - t_w)}{10^{\frac{7.65 t}{243.12 + t}}}$$

- Dew-point temperature (°C)

$$t_d = \frac{243.12 \times \ln \left(10^{\frac{7.65 t_w}{243.12 + t_w}} - 1.068 \times 10^{-4} p (t - t_w) \right)}{17.62 - \ln \left(10^{\frac{7.65 t_w}{243.12 + t_w}} - 1.068 \times 10^{-4} p (t - t_w) \right)}$$

2. Instruments: RH hygrometer, thermometer, baromete

Input readings: air temperature t (°C), relative humidity RH (%), atmospheric pressure p (hPa).

- Water vapour pressure (hPa).

$$e = 0.06112 \times 10^{\frac{7.65 t}{243.12 + t}} \times RH \quad (hPa)$$

- Mixing ratio (g/kg)

$$MR = 38.015 \times \frac{10^{\frac{7.65 t}{243.12 + t}} \times RH}{p - \left(0.06112 \times 10^{\frac{7.65 t}{243.12 + t}} \times RH \right)}$$

- Absolute humidity (g/m^3)

$$AH = 13.44 \times \frac{10^{\frac{7.65 t}{243.12 + t}}}{273.15 + t} \times RH$$

- Dew-point temperature (DP) (°C)

$$t_d = \frac{243.12 \times \ln \left(10^{\frac{7.65}{243.12 + t}} \times \frac{RH}{100} \right)}{17.62 - \ln \left(10^{\frac{7.65}{243.12 + t}} \times \frac{RH}{100} \right)}$$

3. Instruments: Dew-point hygrometer, thermometer, barometer

Input readings: air temperature t (°C), dew point temperature t_d (°C), atmospheric pressure p (hPa).

- *Water vapour pressure (hPa)*

$$e = 6.112 \times 10^{\frac{7.65 t_d}{243.12 + t_d}}$$

- Mixing ratio (g/kg)

$$MR = 3801.5 \times \frac{10^{\frac{7.65 t_d}{243.12 + t_d}}}{p - 6.112 \times 10^{\frac{7.65 t_d}{243.12 + t_d}}}$$

- Absolute humidity (g/m^3)

$$AH = 1344.6 \times \frac{10^{\frac{7.65 t_d}{243.12 + t_d}}}{273.15 + t}$$

- Relative humidity (%)

$$RH = 100 \times 10^{\left(\frac{7.65 t_d}{243.12 + t_d} - \frac{7.65 t}{243.12 + t} \right)}$$

References

Arrhenius, S., 1889: On the Reaction Velocity of the Inversion of Cane Sugar by Acids, Zeitschrift für Physikalische Chemie 4, 226-232.

ASHRAE, 2007: Museums, Galleries, Archives and Libraries, Chapter 21, in: ASHRAE Handbook – HVAC applications, ASHRAE - American Society of Heating, Refrigerating, and Air-Conditioning Engineers Inc.

Atkins, P.W., 1990: Physical Chemistry. Oxford University Press, Oxford, 995 pp.

Bergsten, C.J.,Odlyha, M., Jakiela, S., Slater, J., Cavicchioli, A., de Faria, D.L.A., Niklasson, A., Svensson, J-E., Bratasz, L., Camuffo, D., della Valle, A., Baldini, F., Falciai, R., Mencaglia, A., Senesi, F., 2010: Sensor system for detection of harmful environments for pipe organs (SENSORGAN). E-Preservation Science (in print)

Brimblecombe, P., 2010: Soiling damage and perception. This volume.

Camuffo, D., 1998: Microclimate for Cultural Heritage. Developments in Atmospheric Science 23, Elsevier, Amsterdam, 415 pp.

Camuffo, D., 2004: Thermodynamics for Cultural Heritage, pp. 37-98 in M. Martini, M. Milazzo and M. Piacentini (eds.): Physics Methods in Archaeometry. International School of Physics 'Enrico Fermi', Varenna. IOS Press, Amsterdam.

Camuffo, D. and Giorio, R., 2003: Quantitative Evaluation of Water Deposited by Dew on Monuments. Boundary Layer Meteorology, 107, 665-672.

Camuffo, D. & Pagan E., 2006: What is behind Relative Humidity? Why it is so a relevant variable in the conservation of Cultural Heritage? Pp. 21-38 in R.A. Lefèvre (ed.) The materials of cultural heritage in their environment. European University Centre for Cultural Heritage, Ravello, Edipuglia, Bari.

Camuffo, D., Pagan, E. Schellen, H., Limpens-Neilen, D., Kozlowski, R., Bratasz, L., Rissanen, S., Van Grieken, R., Spolnik, Z., Bencs, L., Zajaczkowska-Kloda, J., Kloda, P., Kozarzewski, M., Mons. Santi,G., Chmielewski, K., Jütte, T., Haugen, A., Olstad, T., Mohanu, D., Skingley, B., Sáiz-Jiménez, C., Bergsten, C.J., Don Russo, S., Bon Valsassina, C., Accardo, G., Cacace, C., Giani, E., Giovagnoli, A., Nugari, M.P, Pandolci, A.M., Rinaldi, R., Acidini, C., Danti, C., Aldrovandi, A., Boddi, R., Fassina, V., Dal Prà, L., Raffaelli, F., Bertoncello, R., Romagnoni, P., Camuffo, M. & Troi, A., 2007: Church heating and preservation of the cultural heritage: a practical guide to the pros and cons of various heating systems. Electa Mondadori, Milano, 240 pp.

Camuffo, D., Pagan, E., Rissanen, S., Bratasz, Ł., Kozłowski, R., Camuffo M., della Valle, A., 2010a: An advanced church heating system favourable to artworks: a contribution to European standardisation Journal of Cultural Heritage 11, 205-219. Doi: 10.1016/j.culturher.2009.02.008.

Camuffo, D., della Valle, A., Bertolin, C., Leorato, C., Bristot, A., 2010b: Umidità e Diagnostica Ambientale in Palazzo Grimani, Venezia. Proceedings of the Symposium: "Edifici storici e destinazione museale: conservazione degli edifici e delle opere d'arte progetti per il restauro e integrazione di impianti esistenti". Milan, 1 – 2 April 2010.

CEN, 2010a: Draft European Standard CEN/TC346 prEN 15758, Conservation of Cultural Property - Indoor Climate - Procedures and instruments for measuring temperatures of the air and of the surfaces of objects. European Committee for Standardisation, Brussels.

CEN, 2010b: Draft European Standard CEN/TC346 prEN15757, Conservation of Cultural Property - Specifications for temperature and relative humidity to limit climate-induced damage in organic hygroscopic materials. European Committee for Standardisation, Brussels.

CEN, 2010c: Draft European Standard CEN/TC346 prEN15759, Conservation of Cultural Property - Indoor Climate - Heating Places of Worship. European Committee for Standardisation, Brussels.

Fagerlund, G., 1973: Determinations of Pore-Size Distribution from Freezing Point Depression, Materiaux et Constructions, 6, 215-225.

Leygraf, C. & Graedel, T.E., 2000: Atmospheric Corrosion, Whiley-Interscience, New York.

List, R.J., 1966: Smithsonian Meteorological Tables. 6 Edition, Vol 114. Smithsonian Institution Press, Washington D.C.

Mason, B.J., 1971: The Physics of Clouds, Clarendon Press, Oxford, 671 pp.

Maxwell, J.C., 1868: On the Dynamical Theory of Gases, Phil. Mag. vol.xxxv, p.129 , Feb.1868 and p.185, Mar.1868.

Michalski, S., 2000: Guidelines for humidity and temperature in Canadian Archives. Canadian Conservation Institute, Ottawa.

Michalski, S., 2009: The ideal climate, risk management, the ASHRAE chapter, proofed fluctuations, and towards a full risk analysis model. 19 pp in: Contribution to the Experts' Roundtable on Sustainable Climate Management Strategies, held in April 2007, in Tenerife, Spain. The Getty Conservation Institute, www.getty.edu/conservation/science/climate/climate_experts_roundtable.html

Nemethy, A. & Scheraga, H. A., 1962: Structure of water and hydrophobic bonding in proteins. A model for the thermodynamic proprieties of the liquid water. J. Chem. Phys., 36, 3382-3400.

Pruppacher, H.R. & Klett, J.D., 2000: Microphysics of Clouds and Precipitation. Kluwer, Dordrecht, 954 pp.

Thomson, W. (Lord Kelvin), 1870: On the Equilibrium of Vapour at a Curved Surface of Liquid. Proc. Roy. Soc. Edinburgh, 7, 63-69.

Tétreault, J., 2003: Airborne Pollutants in Museums, Galleries and Archives: Risk Assessment, Control Strategies and Preservation Management. Canadian Conservation Institute, Ottawa.

Warnatz, J., Maas, U. & Dibble R.W. (eds), 2001: Combustion. Springer, Berlin, 299 pp.

WMO, 1966: International Meteorological Vocabulary. World Meteorological Organisation No 182 TP 91, Geneva, 276 pp.

Zeidler, M.D., 1983: NMR Spectroscopic Studies, pp. 363-413 in F. Franks (ed.): Water, a Comprehensive Treatise. Vol.1. The Physics and Physical Chemistry of Water. Plenum, New York.

Chapter 2

How to measure temperature and relative humidity Instruments and instrumental problems

Dario Camuffo, Vito Fernicola

1. Introduction

The European Standardisation Committee (CEN), Brussels, set up a technical committee to deal with cultural heritage (i.e. CEN/TC346) and the related preservation issues. A specific working group, i.e. WG4, was also set up on environmental measurement and control in museums and exhibition sites. WG4 has in charge the development of new draft standards for indoor and outdoor climate and air quality monitoring. These standards are becoming useful tools for conservators and technical personnel of the wide cultural heritage community sharing knowledge, expertise and providing several suggestions on instrumentation type and use. The input to these standards comes from many discussions carried out among researchers, technical personnel, end users and manufactures about monitoring applications and related problems.

Standards are necessarily concise and strictly limited within the declared scopes. This chapter is intended to provide wider information and illustration of the topics included into standards under discussion at CEN/TC346 WG4 concerning the temperature and humidity measurements, i.e.:

Conservation of Cultural Property - Indoor Climate - Procedures and instruments for measuring temperatures of the air and of the surfaces of objects (CEN 2010a, prEN 15758)

Conservation of cultural property – Procedures and instruments for measuring humidity in the air and moisture exchanges between air and cultural property (CEN 2010b, Adoption n.185).

Ideally, this paper would make friendly the interpretation of the above documents and the spirit behind them.

Any measurement is the synergistic effect of two separate contributions, i.e. the operational modality of the observer, and the instrument used. The former includes the approach chosen to solve the problem, how, when and where to observe. The latter refers to the used instrument that should be adequately chosen and calibrated. Both factors are extremely relevant, and the former can be even more critical than the latter because it includes the general strategy and the individual choices to met all the specific goals of the study.

The aim of this chapter is limited to the instruments and their performances, except a few introductory words on measurements concerning tangible cultural heritage. It is clear that objects belonging to the tangible cultural heritage need special care, and that observation modalities and instruments popular in other fields cannot be applied for the potential risk the contact may produce. For this reason, observational methodologies, instruments and sensors should be appropriate to this specific aim, and guarantee the highest intrinsic safety to any object. Under this point of view, any direct contact between sensor and object surface should be preferably avoided and remote or non-contact observations should be preferred whenever possible.

Temperature and humidity are responsible, in a direct or indirect way, of a number of deterioration mechanisms, as documented in some chapters of this book. In conserving tangible cultural heritage, the most outstanding effect is the visual appearance of damage or other deterioration products, that may be more or less easily restored, but the key problem is to recognize the causes that have generated the damage and undertake mitigation actions to stop or reduce the deterioration rate. Any object interacts with the surrounding environment, e.g. exchanging heat, moisture, gaseous species and particulate where the micro-environmental conditions can be favourable to the acceleration of chemical reactions or microbiological activity. In any case, microclimate plays a fundamental role and the key physical variables should be measured to know their relevance in the ongoing decay and to control them at levels more convenient for preventive conservation. In other terms, we should have a clear idea of what may be happening and investigate environmental forcing and object response to verify if they fill within the range of sustainability and acceptable preservation.

2. The most common errors in environmental monitoring

Air temperature and humidity measurements should be part of an overall plan of environmental diagnostics. Air temperature and humidity measurements should be carried out by specialized people capable of investigating the actual environmental conditions in order to avoid the perturbation or the bias caused by operational errors or an inappropriate choice of sampling conditions, an inappropriate choice of sensors or the use of instruments not having the characteristics required by the problem.

It is evident that any measurement should represent the actual quantity we wish to monitor, removing all the disturbing factors that may affect the readings. The most common and often almost unavoidable disturbance is infrared radiation (*IR*) which supplies external heat to the sensor, with the result that the affected temperature sensors show a higher value and the calculated relative humidity (*RH*) show a lower value. Any control of the room temperature or humidity level based on inputs affected by errors may lead to unpredictable, risky conditions.

Particular care should be taken to shield the air temperature and relative humidity sensors from direct radiation sources, e.g.

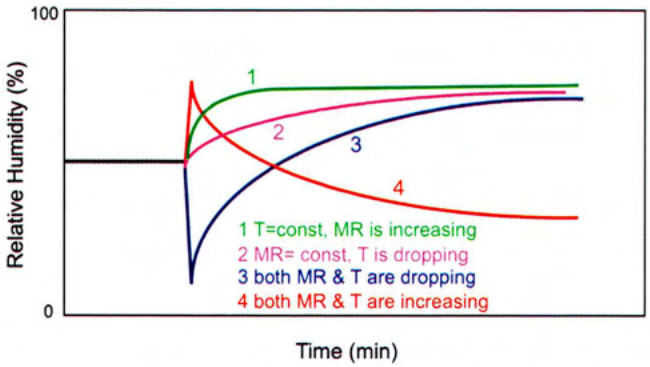

Fig.1 - Output of a capacitive or a resistive hygrometer for a sudden change in RH for the different time response of the humidity sensor and the case/probe. Case (1, green): T is constant, MR is increasing; (2, violet) MR is constant, T is dropping; (3, blue) both MR & T are dropping; (4, red) both MR & T are increasing.

direct solar radiation, incandescent lamps or radiant heaters. Screens should be made of reflecting materials and should have adequate natural or forced ventilation as described e.g. in EN ISO 7726, Annex A.

Another common source of measurement error is the time needed by instruments to reach the equilibrium. A thermometer, or a hygrometer, placed in a given environment does not immediately indicate the air temperature, or humidity, but needs some time to reach equilibrium. The "*response time*" is defined as the time interval between the instant when the measure is subjected to a specified abrupt change and the instant when the response of the measuring instrument reaches and remains within specific limits around its final steady value. The response time is typically expressed as the time needed to reach $1/e = 63.2$ % of the final values and, in this case, it is called "*time constant*", where $e = 2.7182818...$ is the Neper number. Other frequently used response time parameters correspond to 90 % or 95 % of the change and they are 2.3 and 3 times longer than the time constant, respectively. For commercial reasons instrument makers often do not declare the characteristics of their instruments in terms of 90 % or 95 % response time, however who performs the observations should keep in mind these longer time intervals necessary for a reasonably accurate sampling. The response time indicates how fast the instrument is responding. It is a specific characteristic of any instrument and it is independent from the relative amplitude of the change. On the other hand, the absolute error for a too short sensor exposure, i.e. for the instrument having not yet reached the equilibrium, is proportional to the change. For instance, if after an abrupt temperature or humidity change the instrument is read after an elapsed time interval equal to one time constant, the output gives only 63.2 % of the change and the actual value is underestimated by about 37 %. When the time interval corresponding either to 90 % or 95 % is used, the underestimation error is limited either to 10 % or 5 %, respectively.

The microclimate is subject to a continuous variability induced by external factors with sudden short peaks or drops generated by windows or door openings, temporary use of HVAC systems or other disturbances. Such peaks or drops may be very risky for preservation, but a slow response instrument may cut them off, providing a smooth diagram leaving totally obscure the risk. Therefore, if we need accurate measurements, we should choose instruments having a 95 % response time equal to or lower than one third of the shortest expected peak or drop to be monitored. This means that the time constant of the instrument should around be one tenth of the peak/drop duration. In the case of routine measurements, the minimum requirement for an acceptable reading is that a measurement should not be made before an interval of at least 2 times the time constant of the probe.

Before buying or using an instrument, we should carefully read the data sheet with the instrument characteristics and check whether the response time is referred to the sensor alone or to the probe assembled in a protecting case. If this information is unknown, measurement errors are possible. Metal cases or probes are massive and need time to reach equilibrium; those made of plastics have difficulty in exchanging heat with the air and in reaching equilibrium, so that they may even need a longer time.

In the case of temperature sensors with a slow response, a change from a lower to a higher temperature causes the output to gradually rise from the initial to the final level. In the case of a *RH* sensor, things are different. *RH* is governed by both temperature (*T*) and moisture content in the air, which can be appropriately expressed in terms of its mixing ratio (*MR*) (see Chapter 1). The two measured quantities may show different inertia, with the possibility of positive or negative synergism. The thermal inertia of the case/probe is generally higher than that of the *RH* sensor, e.g. 20 min and 5 min, respectively. We will consider some different cases, as follows (Fig. 1).

a. T remains unchanged, but *MR* changes, e.g. from a lower to a higher level. The *RH* output will gradually and rapidly rise from the initial lower level to the final higher one (case 1).

b. *MR* remains unchanged, but *T* changes, e.g. from a higher to a lower level. The *RH* output will gradually and slowly rise from the initial higher level to the final lower one (case 2).

c. Both *MR* and *T* change, the former from a lower to a higher level, and the latter from a higher to a lower level or vice-versa for both of them. The two partial inertias go in the same direction and the output has a gradual transition. The response is similar to case (1) and is characterised by a slow *RH* increase in the first case, or a slow decrease in the opposite case.

d. Both *MR* and *T* will change, both on the same direction, from a lower to a higher level, or vice-versa. This is the most complex case, for opposite synergistic effects developed at different response time of the humidity and temperature sensors. The initial *RH* level may be either higher or lower than the final one, but the output will not gradually pass from the initial to the final level. We will first suppose that both *MR* and *T* are dropping with opposite contributions to *RH*. The *RH* output will first respond to the change in *MR*, i.e., immediately dropping with an undershoot, but later it will be governed by the change in *T*, gradually rising to the final level (case 3). Now we will suppose that both *MR* and *T* are rising. The *RH* output will first respond to the change in *MR*, i.e., immediately peaking with an overshoot, but later it will be governed by the change in *T*, gradually descending to the final level (case 4). In both cases all the output readings made before the final equilibrium is reached are affected by a large bias, and in particular they can show an initial change in the opposite direction compared with the final one.

In practice, the microclimate is often affected by large variability and we cannot exclude that an instrument, although of good quality and carefully calibrated in laboratory under stationary conditions, once it is operating in the field under dynamic environmental conditions may be adversely affected for the above reasons with large output errors.

3 Measuring the effective temperature including radiant contribution

A thermometer measures the equilibrium temperature that it reaches with its surrounding environment. The equilibrium is reached via conductive, convective and radiative heat exchanges. At ambient temperature, radiative exchanges are mostly due to the visible and infrared (*IR*) part of the electromagnetic wave spectrum. Such exchanges are mostly conditioned by the physical properties of the surface, compared with the wavelength (λ) of the incoming radiation.

Incoming radiation is partly reflected, absorbed or transmitted (Fig. 2). The sum of the reflected, absorbed and transmitted fractions equals the incoming flow. A mirror is a typical example of a reflecting surface in the visible range, and a polished metal in the *IR* range; asphalt or graphite are good absorbers in both visible and *IR* ranges; glass is transparent in the visible, but only partially in the *IR* range; a thin layer of water is transparent in the visible, but opaque in the *IR* range. A good *IR* absorber is also a good emitter at each wavelength λ the ratio between absorbing $\alpha(\lambda)$ and emissive power $\varepsilon(\lambda)$ is established to be equal to 1 (Kirchhoff law).

The emissive power varies in the range $0 \leq \varepsilon(\lambda) \leq 1$. The lower limit, i.e. 0, is for a perfect mirroring surface, the upper limit, i.e. 1, for a total absorber, i.e. a black-body cavity. Most of the materials used to build tangible cultural heritage are opaque, and in such a case the incoming radiation is either absorbed or reflected. When we measure the *IR* power emitted by a surface, we should consider the bulk emissivity (ε) to which the instrument is sensitive, generally in the 7-14 μm *IR* spectral window. Stone, brick, plaster, wood, parchment, paper, textiles, oxidised metals, deposited smoke and soot are characterized by $\varepsilon > 0.9$; on the other hand most polished metals have $\varepsilon < 0.1$. If an opaque surface has emissivity $\varepsilon = 0.9$, this means that an *IR* detector aimed at the surface will receive approximately 90% of the radiation from the target surface and another 10% of the radiation reflected from surrounding bodies. If surrounding bodies have a higher temperature, the reflected fraction will apparently raise the thermal level of our target and the temperature will be overestimated; the opposite for surrounding bodies having a lower temperature.

The intensity (I) of the *IR* radiation emitted from a surface is highest (I_o) in the direction normal to the surface and vanishes in the plane tangent to it, following the Lambert's cosine law for any angle θ from the normal (Fig. 3), i.e.

$$I(\theta) = I_o \cos\theta$$

However, when we observe the surface with an increasing angle θ from the normal, the decrease of $I(\theta)$ is exactly compensated by an equal increase of the emitting surface S that falls within our optical angle, exactly what happens with the shadow when the Sun is lowering towards the horizon. For the same reason the Moon or the Sun have the appearance of flat discs and not of spheres. Any surface for which such compensation happens is called a Lambertian surface, and in such a case there is no problem for measuring the power emitted at a tilted angle. This fact helps avoiding the error in measuring the *IR* reflected by the observer or other hot bodies located in proximity (Fig. 4). However, sometimes we need to know the effective temperature reached by our target surface when an intense light beam (e.g. from the sun or a spot light) or an *IR* radiation (e.g. from a heater or another hot surface) hit it.

The effective temperature is measured with the so-called black-globe thermometer which absorbs both visible and *IR*

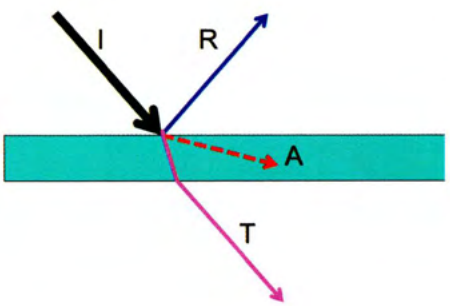

Fig.2 - Incoming (I), reflected (R), absorbed (A) and transmitted (T) radiation impinging on a partially transparent slab.

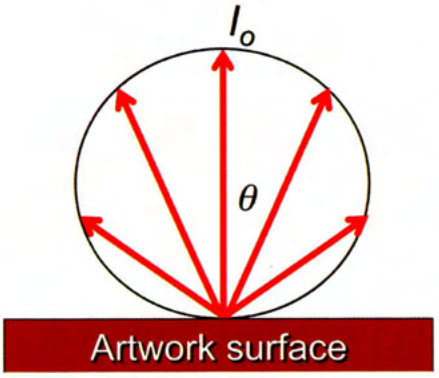

Fig.3 - Lambert's cosine law for the emission of IR radiation from a body surface.

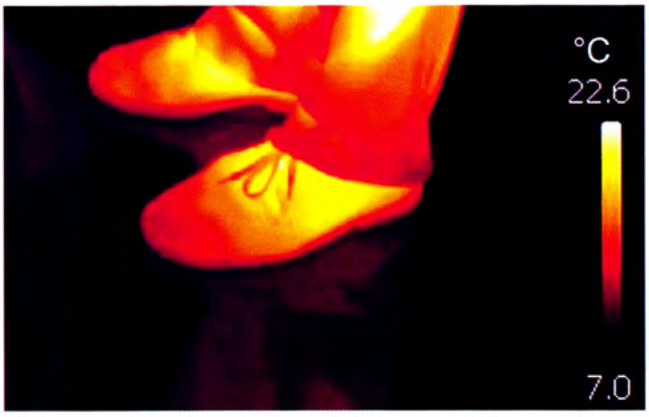

Fig.4 - Reflected thermal image of shoes and legs on a cold limestone floor (e = 0.90) in an unheated church. The reflected image is a weak shadow (i.e. 10% intensity compared with the real image) that becomes visible when the floor is much colder compared with the body on it.

parts of the spectrum. This temperature results from a thermal balance among the air temperature, the radiation effect from the target object and other surrounding bodies, and the convection effect. The black-globe thermometer has a spherical shape, because it was originally introduced for measuring the amount of radiant heat received by a human body from different directions. Its shape is not appropriate for measurements of objects with largely different geometries, for example flat paintings on canvas, wooden panels or tapestry. Its response time is generally long, more than 15 min, depending on the globe size and the ambient conditions. The black-globe-thermometer and the measurement of the effective temperature are discussed in details in the EN ISO 7726 standard.

Lighting a painting on canvas, or a wooden panel, may cause a sharp rise in temperature. It may be useful to know how much a spot lamp, or a heater may raise the surface temperature of an object, or the thermal feeling a person may have when enter a historical building heated by *IR* radiators (Fig. 5). This measurement can be done with a black strip target located in the position planned for the object, or placed close to the object but without contacting it. The black strip target will simulate what happens to

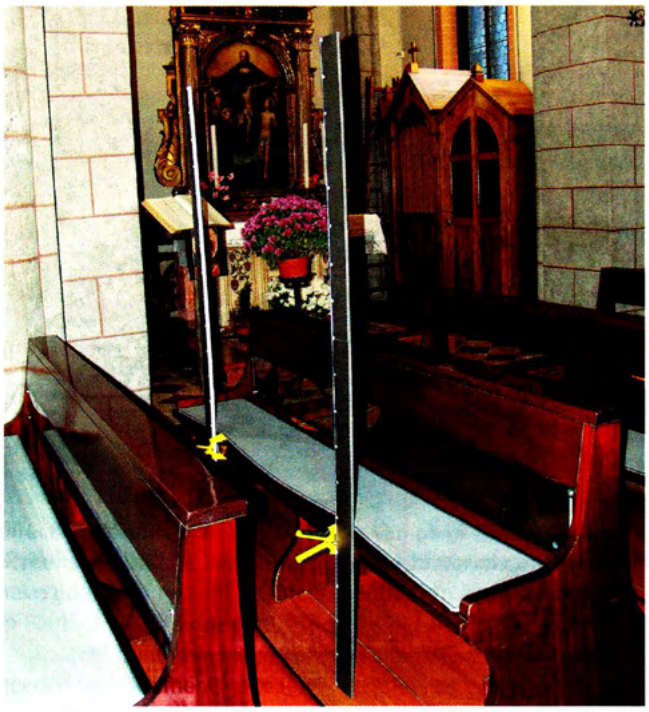

Fig.5 - Black vertical strip utilzed as a target to monitor the combined effect of air temperature, incoming radiation and air motions.

the object without making experiments with the object itself. This simulation is representative of most of the materials used for cultural property, except for polished metals that have a very low emissivity and reflect most of the incoming radiation. The method is also useful to assess the level of the thermal comfort for visitors to museums, galleries, churches, historical buildings etc., where heating should be reduced to a minimum acceptable level and/or to restricted areas.

The effective temperature resulting from a thermal balance between air temperature and radiation is determined with a fast responding method, if a low thermal inertia, blackbody target, like a strip of black textile is used. The equilibrium temperature of the blackbody strip target should be accurately measured using the procedures for the measurement of surface temperature described below. A black strip target is useful to measure gradients from vertical or horizontal profiles, with a high spatial resolution.

4. Measuring the surface temperature

The surface temperature can be measured by contact thermometers or *IR* radiation thermometers either remote or "quasi-contact". A radiation thermometer can be monochromatic, i.e. it detects the radiation in a very small portion of the spectrum, or it can be a total radiation thermometer where it detects a very wide portion of the incoming radiation spectrum.

4.1 Thermometers with contact sensors

Contact measurements should be made when necessary and with caution because they are generally obtained by exerting a pressure against the object surface or using an adhesive or a conductive paste to improve the heat exchange between the surface and the sensor in order to reach the thermal equilibrium. Such a measurement is potentially dangerous to the object. Contact measurements should be undertaken only after consultation with an appropriate professional, e.g in the case of building surfaces with a qualified conservation architect and in the case of other heritage items with a qualified, experienced, conservator.

Contact temperature sensors are usually small in size and enclosed in small probes. Such probes have a flat surface on one side to favour a good thermal contact with the object surface, and are insulated or shielded on the other side, so that the probe is poorly affected by air temperature, infrared or visible radiation. Three types of sensors are generally used for contact sensor: platinum resistor, thermistor and thermocouple. The elemental sensors performance and specifications for such sensors are discussed in published standards, such as the EN IEC 60751, EN IEC 60584 and ASTM E879-01.

As the probe covers the object surface in the measurement area, it may shield the surface from exchanging heat between the object and the environment. The problem becomes relevant when the object surface is exposed to intense radiation because the surface temperature is measured over an area shadowed by the probe and the temperature is thus underestimated. This is especially a problem on materials with relatively low thermal conductivity. Heat capacity of the contact thermometers may also alter the surface temperature of objects with low thermal mass like paper sheets, photographs and paintings on canvas.

Contact measurements are technically possible only on smooth, flat surfaces which provide a good thermal contact with the probe. In the case the surface is rough, jelly-like substances are added between the target surface and the probe to improve thermal contact. This methodology involves the risk of staining the surface and can be applied only in some particular cases. In order to measure the surface temperature of polished metals, a contact sensor approach is the only one that may guarantee reliable results. The emissivity of polished metals is very low and this excludes the possibility of using the next two methodologies. Again great care must be taken not to mar the surface at end.

4.2 Infrared thermometers

An *IR* thermometer is an energy detector that enables non-contact, remote, measurement of surface temperature. It has a special advantage in case of surfaces difficult to reach or mobile objects. The *IR* radiation emitted from the target surface enters the *IR* thermometer either directly or through a filter, and is focused onto a detector that produces an output signal proportional to the incident radiation. The emission temperature of the target surface is calculated from the power of the measured *IR* flux that is proportional to the surface emissivity and to the fourth power of the surface absolute temperature. The surface emissivity should be known or estimated in advance. Once the instrument has reached thermal equilibrium with the environment, measurements are very fast.

Depending on the type of detector used, *IR* thermometers can be divided into three categories:

- Radiation pyrometers, i.e. devices that measure the power density of the *IR* emitted by the target surface over a given range of wavelengths. They are subdivided into total radiation pyrometers, i.e. operating on a very wide range of wavelengths, narrow band pyrometers, i.e. sensitive to a narrow spectral band, or ratio pyrometers, i.e. determining the temperature after the ratio of the *IR* power emitted by two wavelength bands. Particularly relevant are optical pyrometers in the 7-14 μm band. Sensors include photo detectors, e.g. photo-conductive or photo-voltaic, that are wavelength sensitive. The output is not directly caused by heat, but by the electric charges released at the detector and collected by an electrode.
- Thermal detectors which generate an output that is related to the thermal *IR* energy they absorb. The sensor can be

made of thermopiles, thermocouples, metal- or semiconductor-type bolometers or pyroelectric detectors. The output is directly caused by heat.

We can also distinguish *IR* thermometers in terms of the output produced, i.e. imaging and non-imaging.

An imaging *IR* thermometer, also known as thermal camera, produces a picture that is a thermal picture of what is in front of the camera, reproduced in false colours that correspond to the thermal level of the imaged surface. The false colours are coded according to a temperature scale. The detection spectral range is in the band from 7 μm to 14 μm. False colour images are particularly useful in environmental diagnostics to give a general view of the temperature distribution, to recognize heat losses, thermal bridges, overheated objects, space gradients, structural discontinuities, water percolation etc. They can be used either to monitor the actual surface temperature of a target in a given instant or to detect temperature discontinuities under dynamic regimes, i.e. during warming or cooling cycles.

A non-imaging *IR* thermometer delivers an output that is proportional to the average temperature of a defined area on the target surface. The optical angle of view of the detector defines the size of the target area.

Remote sensing is generally characterized by a lower accuracy than contact temperature sensing because the detected radiation may also include contributions due to the reflected radiation from the surface of the object or emitted radiation by other nearby objects. The lack of knowledge of the surface emissivity is also another cause of error. As a general rule, the measurements uncertainty increases when the surface emissivity decreases; metals have very low emissivity and may even reflect the apparent thermal image of the operator causing large measurement errors. This procedure is not appropriate for polished metals, but it can be used for heavily oxidized or painted metals. Some materials, like glass or plastic, are transparent at the visible wavelengths but may be opaque at the *IR* radiation. A key issue is that for many materials the transmittance and the absorption coefficients change with the wavelength and it is thus important to know how the spectral emissivity of the target material is matched with the spectral band of the *IR* thermometer.

4.3 Quasi-contact thermometers

Quasi-contact thermometers generally comprise a concave mirror which focuses the *IR* radiation emitted by the target surface onto a detector, either a pyroelectric, thermal, or photo detector (Fig. 6). In operation, the device is placed close to, but not in contact with the surface of the object. The external surface of the mirror shields the detector from the surrounding incoming radiation. The internal surface of the mirror focuses the direct and diffused *IR* radiation emitted by the object onto the sensor. Under these conditions the sensor receives almost the whole amount of radiation emitted. The surface emissivity can thus be approximated to one whatever is the actual emissivity of the material in the absence of the concave mirror, ensuring higher accuracy than other radiation measurements.

Generally quasi-contact thermometers are well suited for measurements of the surface temperature of cultural properties because there is no physical contact with the surface, which is particularly advantageous in ensuring safety to delicate objects. Thanks to its small size, it can be used to measure the temperature of curved and rough surfaces or materials composed by loose fibres. This is particularly advantageous when measuring the temperature of soft materials (e.g. textiles) or of objects with a very low thermal capacity (e.g. a sheet of paper) without the risk of introducing perturbing errors due to the heat flux between the sensor and the object.

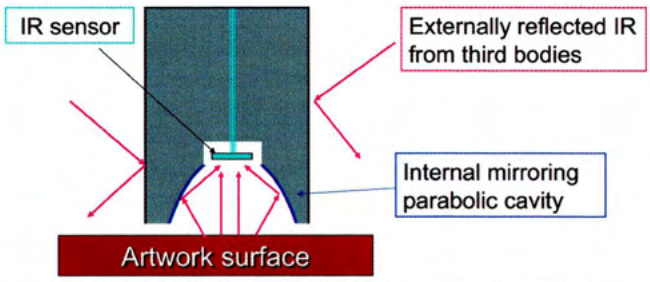

Fig.6 - Operating principle of a Quasi-Contact Thermometer. The sensor is hit by the total flux, i.e. direct and diffuse IR radiation emitted from the target surface. The diffuse radiation hits the sensor after having been reflected by the internal mirroring parabolic cavity. The external surface of the cylinder is also mirroring and reflects away the IR coming from third bodies.

The surface temperature of polished metals, because of their extremely low emissivity, cannot be measured with this method and should necessarily be measured with contact sensors.

5. Space and time variability

5.1 Aspects related to temperature variations in space

Outdoor and indoor environments, historic buildings or simple objects show generally temperature distributions due to a complex energy balance, including the presence of heat sources like sunlight, heaters, spotlights; heat sinks like walls or ceilings; air leakage due to opened doors or windows. For these reasons, it may be necessary to monitor the temperature or the temperature gradients of the environment surrounding the object, the surface of the building or the object itself and maybe inside cavities or spaces internal to the building or the object.

When the environment is thermally heterogeneous, temperature should be measured at one or several locations according to a precise measurement design which takes into account locations of objects exposed to a thermally induced risks as well as local thermal and air flow pattern. Specific problems to be studied should be defined. Complex heterogeneous spaces can be divided into several limited zones for which different measurement plans are defined. Temperature monitoring often requires an array of sensors. To insure a prompt comparison of the measurement performance and easy interchangeably of the sensors, it is suggested to use sensors of the same type.

5.2 Aspects related to temperature variations in time

Changes in temperature may be characterised as cyclic or irregularly variable, on the long, medium and short term. Cycling may be seasonal or diurnal or a combination of the two. The variability may be due to weather and may be complicated by patterns of use, such as the entry of visitors and the running of lighting on and off.

When the environment cannot be considered stationary, temperature variability should be measured as a function of time according to a precise plan, which would take into account the specific problems and quantities to be highlighted. Measurements over a full year are most satisfactory but more limited seasonal periods can be considered depending on the problem to be studied. Sampling frequency should be adapted to the time scale, the dynamics and the fluctuations of the phenomena under investigation; so that the shortest variation of interest is well documented.

6. Performance of instruments to measure temperature

The measuring ranges, uncertainty, resolution, response time and stability of the commercially available sensors for air and surface temperature are summarized in Table 1 and correspond to the quality of the average thermometer used for indoor measurements, e.g. museums, galleries, libraries, or inside showcases. Outdoor measurements are often more problematic and are carried out with larger uncertainties.

Table 1. *Typical thermometer performance.*

Instruments	Measuring range	Uncertainty	Resolution	Typical response time	Stability
Thermometer for air temperature	Outdoors -40 °C to 60 °C Indoors -20 °C to 60 °C	0.5 °C	0.1 °C	The shortest possible; not longer than 1 min	0.2 °C/year
Black globe thermometer	Outdoors -40 °C to 100 °C Indoors -20 °C to 100 °C	0.5 °C	0.1 °C	The shortest possible; not longer than 20 min	0.2 °C/year
Black strip thermometer	Outdoors -40 °C to 100 °C Indoors -20 °C to 100 °C	1.0 °C	0.1 °C	The shortest possible; not longer than 3 min	0.2 °C/year
Surface temperature (contact or proximity sensor)	Outdoors -40 °C to 100 °C Indoors -20 °C to 80 °C	1 °C + 0.01 T	0.1 °C	The shortest possible; not longer than 3 min	0.2 °C/year
Surface temperature (remote sensor)	Outdoors -40 °C to 100 °C Indoors -20 °C to 80 °C	1 °C + 0.01 T	0.1 °C	The shortest possible; not longer than 1 min	0.2 °C/year

It should be noted, however, that a number of specific applications may require better performances than the averages listed in Table 1 that represent the minimum acceptable level.

Instruments should have readings expressed in degrees Celsius (°C).

7. Historical touch: popular instruments used to measure humidity in the past

Historically, relative humidity (*RH*) of the air was mostly measured by two instruments: the hair hygrometer and the psychrometer.

The hair hygrometer was invented in 1783 by Horace Bénédict de Saussure. The hair hygrometer was the best hygrometer developed in the 18th century, and was able to provide reasonably correct, direct reading on an analogue scale. This made possible routine monitoring with continuous records on a strip chart by means of an ink pen fixed to its mobile arm connected to the hair whose elongation was proportional to the equilibrium *RH*. In weather stations, but also in other application fields the thermo-hygrograph, i.e. a combination of a bimetal thermometer and a hair hygrometer became very popular. Air temperature and *RH* were thus plotted with two ink pens on weekly or monthly charts fixed onto a rotating drum. The most famous producers in the 19th and early 20th century were Richard Frères in France.

In 1888-1892, Adolf Richard Assmann developed an instrument for accurate measurement of atmospheric humidity and temperature, based on two thermometers shielded from solar radiation, one of them having a wet bulb. The hygrometric quantities were obtained from the two readings with the help of a diagram or a tables. The instrument, called the psychrometer, was soon produced by the factory of Rudolf Fuess in Germany.

The most accurate measurements were carried out with the Assmann-type psychrometer which was made of two identical mercury-in-glass thermometers, ventilated by a mechanical fan. The observation required manual operations and readings were possible only when the wet bulb temperature was above 0 °C, which corresponds to an air temperature above 8 °C in a dry environment (e.g. *RH* =10 %). Continuous *RH* monitoring was not possible and any reading of the dry and wet bulb temperatures required the help of tables or diagrams to obtain *RH*. In practice, such mechanical psychrometer was rarely used in the field, confined to laboratory use for calibration of lower accuracy instruments, such as the hair hygrometer.

Although hair hygrometers performed less accurate measurements compared with the Assmann psychrometer, they were largely used until the 1970s. The thermo-hygrograph became very popular in museums, galleries, archives and libraries. At a first sight, the routine use of such instrument is very simple, as it is sufficient to replace the strip chart at fixed intervals and the ink pens when necessary. However, this instrument requires periodic care and maintenance because the hair suffers the contamination from dust and from the air pollution which could cause severe instrument drift. To overcome such limitations, it needs periodic cleaning and suitable moistening; after such maintenance operations it should be calibrated again.

Hair hygrometers are also affected by a large hysteresis as they reacts to humidity changes in a different way when they are moist or dry, the less accurate readings being in moist conditions. For the above reasons, the use of the hair hygrometer in museums and other historical buildings, where important collections are preserved is highly inappropriate. A comparison of the instrument output against a calibrated reference standard may show, under some circumstances, departures of 20 % to 30 % from the initial calibration.

With the advances in electronics and material science, in the late 1970s new electrical sensors and instruments were developed and made available on the market, with higher performances and reliability, at a lower costs. Polymers-based capacitive or resistive *RH* sensors were soon adopted by national weather services, in the industry, in the agriculture and in other field applications. Some museums, galleries etc. updated their instrumentation; others did not. Today the situation is rather inhomogeneous, with a large amount of environmental data originated from instruments with different levels of accuracy.

The aim of this section is to assist the personnel in charge of conservation to carry out humidity measurements at an appropriate level of accuracy based on adequate knowledge of methodology and instrumentation.

8. Electronic psychrometer

The psychrometer consists of two thermometers and a fan to ensure their ventilation at a sufficient air speed. One of the thermometers is used to measure the air temperature. This is referred to as 'dry-bulb' temperature as opposed to 'wet-bulb' temperature indicated by another thermometer. The latter has a bulb wrapped by a wet wicking. Evaporation from the wet bulb lowers its temperature, i.e. the 'wet-bulb' thermometer becomes colder than the dry-bulb thermometer where the temperature difference decreases with the degree of saturation of the air moisture.

Modern electronic psychrometers consist of two platinum resistance thermometers read by digital electronics and data processing which enable real-time calculation of *RH* and other moist air parameters. When thermometers are regularly calibrated and the instrument is appropriately used, accurate humidity measurements are possible. The electronic psychrometer is more reliable than the mechanical type and can easily be used for on site check of other filed hygrometers. Nevertheless its use requires trained personnel.

The main limitation of this instrument is that it cannot be used at wet-bulb temperature below 0 °C because the wet wicking freezes. In addition, when the ambient temperature drops below 0 °C, also the water in the reservoir freezes. The water reservoir capacity imposes a limit to the duration of unattended monitoring. The air flow generated by the fan and evaporation from the wet wicking may slightly perturb the surrounding environment. For that reason this instrument is not suitable for measurement in small volumes, e.g. showcases.

The instrument operates in a wide measurement range from about 15 % to 100 % *RH*. The relative humidity is calculated with formulae which include the atmospheric pressure. The latter should be independently measured when it departs too much from the standard sea level pressure, e.g. in mountain sites (Fig. 7). However, at moderate altitudes, under usual barometric pressure variability, in the lack of direct pressure measurements, it is possible to assume a pressure of 1000 hPa as an acceptable approximation; the resulting measurement error would be around 1 % *RH*.

For accurate readings, care should be paid to the following aspects:

- the air flow speed should be kept constant, in the range from 3 m/s to 5 m/s. The fan should be operated well in advance so that both sensors and the wet muslin may reach a thermal equilibrium. This operation can be adequately provided by an electronic psychrometer with an electrically-driven fan. The old-type mechanical psychrometer with spring-operated fan had the ventilation speed variable with the spring winding and may provide incorrect readings. The ageing of the fan should be periodically checked.
- the temperature sensors should be shielded against radiation with a suitable radiation shield;
- the wet wicking should be clean; this is particularly relevant in coastal (sea spray) or polluted areas.
- the water reservoir should be clean and filled with pure, distilled or deionised water.

For accurate measurements with a psychrometer, a key precaution is to ensure that both thermometers give exactly the same readings when both bulbs are dry. The *RH* measurement error due to a difference in the readings between the two thermometers in dry conditions is as much relevant as the uncertainty in the temperature measurement.

An example of the errors done when a difference of 0.1 °C is found between the two thermometers, or when a reading is under- or over-estimated by 0.1°C in the case of ambient temperature between 0 °C and 50°C and *RH* = 50 % is shown in Fig. 8.

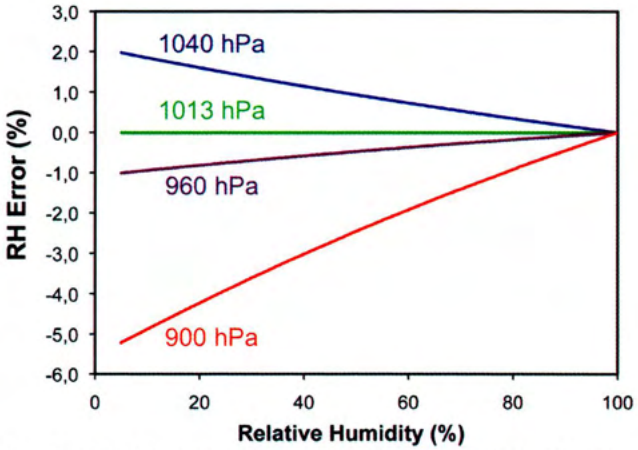

Fig.7 - Error made in computing RH from psychrometric readings neglecting the atmospheric pressure variability, i.e. setting the atmospheric pressure p=1000 hPa in the formula for some selected cases, i.e.: 1040 hPa, top high pressure level at mean sea level (blue line); 1013 hPa average pressure (green line); 960 hPa, lowest pressure level at mean sea level; 900 hPa, average pressure at 1000 m above mean sea level for a standard atmosphere (red line). The error in the usual range of weather variability at sea level is in the band between the blue and the violet lines.

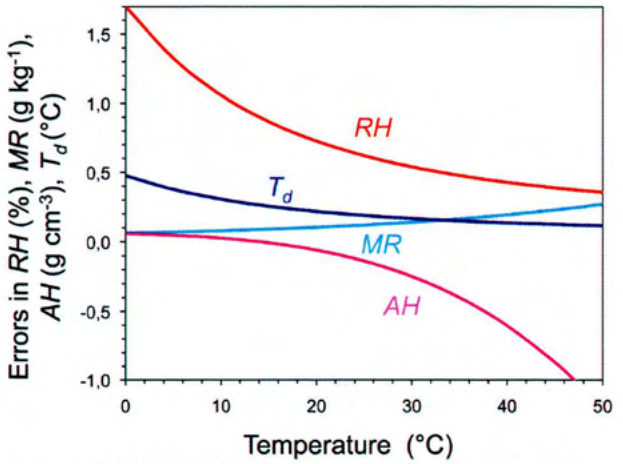

Fig.8 - Example of errors generated in computing Relative Humidity (RH), Mixing Ratio (MR), Absolute Humidity (AH) and Dew Point (T_d) when the dry or the wet bulb temperature T_w is measured with an error of -0.1°C, at ambient RH = 50%. The error is reported in the ordinates per each variable, and refers to the corresponding curve (after Camuffo, 1998).

Platinum resistance temperature sensors offer the best performances for such an instrument. When the sensors comply with the EN IEC 60751 industrial standard, accuracy and interchangeability is also assured.

9. Electronic hygrometer with a capacitance-based sensor

The sensor is made of a hydrophilic material that reaches equilibrium with the ambient relative humidity; the equilibrium moisture content affects the capacitance dielectric constant as a consequence of the absorbed water vapour. The most popular kinds of sensors are based on metal oxides or polymers.

Metal oxide capacitive sensors are based on a porous aluminium oxide deposited on a conductive aluminium substrate with a very thin film of gold that will act as electrode. Water vapour molecules are absorbed by the porous aluminium oxide, changing the capacitance of the system. The main advantages are a wide range and small size. The accuracy how-

ever is not that high and the sensor needs periodical recalibration to compensate for contamination and ageing.

Polymer capacitive sensors are based on a polymer dielectric. The polymer sensor is largely immune to condensed water, is less affected by pollutant contamination, provides better linearity, and is more stable than a metal oxide sensor. It is often used in conjunction with a temperature sensor, e.g. a platinum resistance detector or a thermistor, to obtain a compact air temperature and *RH* sensor for environmental monitoring. The resulting sensor probe can be small and can easily achieve the accuracy needed in the field for long-term monitoring.

Key factors and main limitations of capacitance-based sensors are:

- typical measurement range is from about 10 % to 95 % *RH*;
- a protection is needed against mechanical shocks with a shield, often of plastic material. This induces a thermal inertia that might lead to errors in case of fast temperature or humidity changes;
- in the case of outdoor measurements, the humidity sensor should be shielded against radiation with a metallic shield;
- a long-term exposure in polluted or coastal environments may poison the sensor.

10. Electronic hygrometer with a resistance-based sensor

The sensor is made of a hydrophilic material that quickly reaches equilibrium with the ambient relative humidity; the equilibrium moisture content determines its electrical resistance. The most popular sensors are based on solution resistance cells or on polystyrene film cells.

The solution resistance cell, also called the Dunmore Cell, consists of a wire grid on an insulating surface that is coated with a solution, e.g. lithium chloride. The lithium chloride solution is hydrophilic and absorbs water vapour molecules from moist air until the solution reaches equilibrium with the ambient *RH*. The main limitation is that its measurement range is quite narrow. As any electrolyte-based material, it may be damaged after a long dampness period.

The polystyrene film cell, also known as the Pope Cell, is based on a polystyrene film used as a substrate for the conductive wire grid. The polystyrene is treated with sulphuric acid to increase the affinity with water. When the ambient *RH* varies, also the moisture content in the polystyrene layer changes, changing the impedance of the system. The equilibrium is reached in a short time making fast sampling possible; however when such sensor is encased in a probe, the thermal time constant of the probe itself can limit the response time. The polystyrene sensor is quite stable and may provide long-term monitoring on site. The small sensor size allows its use in small sampling volumes. It shows the same limitations and key factors of the above capacitive-based sensors.

11. Dew-point hygrometer

The dew point hygrometer is based on the detection of the temperature at the point at which water condensation forms on a suitable cooled surface. The temperature at the condensing surface is read by an embedded resistance thermometer and controlled by an electronic feedback circuit to maintain a dynamic equilibrium between evaporation and condensation, thus closely following the dew-point temperature of the air. The early formation of water droplets or a thin film of water on the chilled surface can be detected either optically or electrically.

The optical detection is accomplished by detecting the reflected and/or the scattered light signal of visible or *IR* radiation from a metal mirror. The reflected light intensity decreases when water condensation occurs on the chilled mirror surface. The embedded thermometer reads the temperature at which condensation occurs.

The surface conductivity dew-point hygrometer is similar, but water condensation is detected by electrical means. The chilled surface is made of an insulating material between two gold-coated electrodes; the electric current between the electrodes increases with the thickness of the water layer on the surface.

Both operating principles are reliable for dew-point measurement with some advantage in compactness for the electrical detection but with a better accuracy for the optical detection. In fact, the chilled-mirror hygrometer provides the most accurate humidity measurements and for its performance is widely used in test and calibration laboratories.

The dew point hygrometer is a stable and reliable instrument and covers a very wide humidity range. For this reason it can be used as a reference standard for the calibration of other (secondary) hygrometers. It can be also used to detect the frost point. Its use requires well-trained personnel.

Key factors and main limitations are:

- the dew-point hygrometer operates in the full range, from about 0 % up to nearly 100 % *RH*;
- it is rarely used for on-site monitoring because of its size, even if newly development instruments are more compact and lightweight enabling selected field applications;
- its high accuracy requires periodic control of all parts, including cleaning of the metal mirror and the optical components, or the electrically sensitive surface.

12. Hair hygrometer and cellulose hygrometer

The hair hygrometer operation was preliminary presented in the historical touch (Section 7). The sensor operating principle is based on the elongation of hair or synthetic fibres. A mechanical linkage is used to amplify the hair elongation for the readout device that can be a mechanical leverage with an ink pen plotting a line on a strip chart recorder (mechanical type) or an electrical detector like a potentiometer (electro-mechanical type).

The hair sensor suffers from a large hysteresis (Fig. 9) and its output can depart significantly from linearity. The readout can be based on a linear or a non linear scale on the strip chart (Fig.10a,b). The instrument needs continuous maintenance and frequent calibration from trained personnel; without this care the hair hygrometer is prone to large measurement errors. Therefore, it is not recommended for the environmental diagnostics in the cultural heritage field. Nowadays, after the introduction of modern electronic sensors, that are more reliable

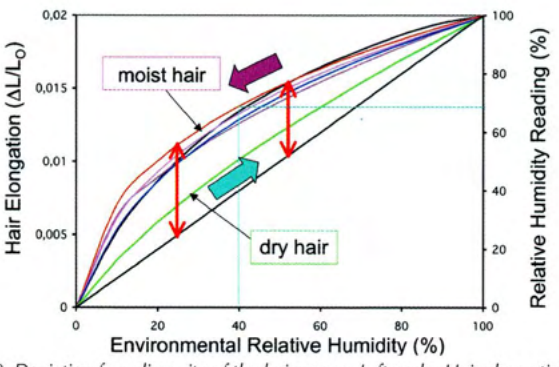

Fig.9 Deviation from linearity of the hair sensor. Left scale: Hair elongation (DL) as a fraction of the initial length Lo. Right scale: actual RH reading (output). Abscissa: reference RH in the calibration chamber.

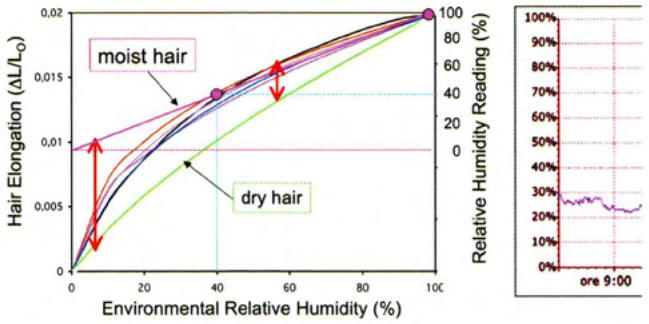

Fig. 10a - Linear output solution: reduced scale, mid-range best fit.

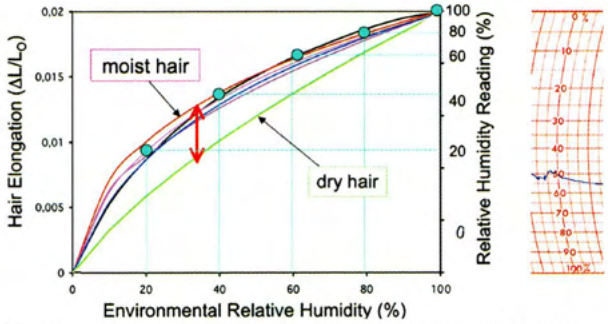

Fig. 10b - Non-linear output solution: the scale is non-linear and follows the curvature with variable resolution.

and less expensive, its use is not longer justified.

A cellulose strips hygrometer changes dimensions with its moisture content, which in turn is in equilibrium with the ambient *RH*. Cellulose sensors are commonly used to build dial for heating, ventilating and air conditioning (HVAC) indicators, transmitters or hygrostats. Measurement characteristics, problems and use are similar to the hair hygrometer.

13. General considerations concerning humidity measurements

An accurate *RH* determination requires particular care because measurements are affected by the equilibrium between the temperature of the air and the actual instrument temperature and by the content of water vapour in the air.

The measurement of air humidity must be conceived in the frame of a specific monitoring campaign that considers not only its absolute value but also the thermal field in the surrounding environment as well as close to the artefact.

The locations of the measuring points must be selected in such a way that they are representatives of the environment under investigation. In general, each room shows variations of temperature and *RH* from point to point, therefore temperature and *RH* of the air that interact with the object should be measured at a close distance to the target surface. *RH* should be measured also in the free air, i.e. in a location not affected by the target (preferably, at a distance of one metre or in the middle of the room). From these two measurements it is possible to establish whether the surface is exchanging moisture with the air or not.

In case the target surface temperature is different from the air temperature, the air layer in contact with the surface reaches a different *RH*, which is difficult to measure. The actual *RH* at the interface between the air and the surface should be calculated from the actual surface temperature and the humidity mixing ratio of the air in the proximity, the latter to be derived from the values of air temperature and wet-bulb or dew-point temperature or, equivalently, from air temperature and *RH*.

Measurements in locations affected by disturbing factors like heaters, ventilation grids, windows or doors, surfaces having a different temperature should be avoided. The measuring instruments should be placed at the level of the object if air temperature stratification is suspected.

The effect of temperature on relative humidity is particularly important and must be taken into account in order to avoid measurement errors equal to or larger than the estimated measurement uncertainty. A common source of error in relative humidity measurements is the temperature difference between the air and the instrument.

The response time of an *RH* hygrometer must be considered. Meaningful results can be obtained after the sensor attains the equilibrium at a given temperature and relative humidity, which requires a time interval longer than twice its response time. If the probe is likely to be exposed to solar radiation, intense artificial light sources or infrared radiation from heaters, it should be carefully shielded.

Relative humidity can be measured both directly, by means of sensors whose output is proportional to it, and indirectly through the measurement of the air temperature and the dew-point or wet-bulb temperatures. In the latter case, one have to take into account the atmospheric pressure, in particular for measurements performed at high altitudes in mountain sites.

14. Performances of instruments to measure humidity

The typical measuring ranges, uncertainty, resolution and response time of commercially available instruments for measuring the air humidity are summarized in Table 2.

Table 2. *Typical hygrometer performances.*

	Dew-point hygrometer	**Electronic psychrometer**	**Capacitive and resistive electronic hygrometer**
Measuring range	-20 ° to 50 °C	5 % to 95 % *RH* (0 °C ≤ *t* ≤ 50 °C)	5 % to 95 % *RH* (-10 °C ≤ *t* ≤ 50 °C)
Typical measurement uncertainty	0.5 °C	3 % *RH*	3 % *RH*
Resolution	0.1 °C	0.1 % *RH*	0.1 % *RH*
Typical response time	-	≤ 2 min	≤ 5 min
Stability	≤ 0.2 °C/year	≤ 2 %/year	≤ 2 %/year

It should be noted, however, that a number of specific applications may require better performances than those listed in Table 2.

Inaccurate measurements performed either by unqualified personnel, unaware of adequate measuring procedures, or done with uncalibrated, low quality, instruments should be rejected because they may suggest inappropriate remedial actions that might have a negative impact on the preservation.

Acknowledgements

This work was made possible thanks to the European Science Foundation funding and the scientific activity performed within the action COST D42: *EnviArt*. The paper reports results achieved during previous EU funded projects, namely AER, FRIENDLY-HEATING, SENSORGAN and *Climate for Culture* (GA 226973). Many useful discussions have been made with all the WG4 colleagues when preparing drafts for standards at the European Committee for Standardisation (CEN TC346 WG4) as well as UNI, the Italian Standardisation Body.

References

Actis A. and Fernicola V., 1999: A Reference Instrument for Accurate Measurements of Air Temperature, pp. 191-196 in Proc. XXV World Congress of the International Measurement Confederation (IMEKO XV), Vol. VII, Osaka, Japan.

Arpino F., Fernicola V., Frattolillo A., Rosso L., 2009: A CFD Study on a Calibration System for Contact Temperature Probes, Int. J. Thermophysics, Vol. 30. pp. 306-315.

Ballard, L. F., 1973: Instrumentation for Measurement of Moisture: Literature Review and Recommended Research. Highway Research Board, National Research Council.

Bentley, R.E., 1998: Handbook of Temperature Measurement: Temperature and humidity measurement. Springer Verlag, Berlin.

Camuffo, D., 1998: Microclimate for Cultural Heritage. Developments in Atmospheric Science 23, Elsevier, Amsterdam.

Camuffo, D., 2004: Thermodynamics for Cultural Heritage, pp. 37-98 in M. Martini, M. Milazzo and M. Piacentini (eds.): Physics Methods in Archaeometry. International School of Physics 'Enrico Fermi', Varenna. IOS Press, Amsterdam.

Camuffo, D., Pagan, E., Rissanen, S., Bratasz, Ł., Kozłowski, R., Camuffo, M., della Valle, A., 2009: An advanced church heating system favourable to artworks: a contribution to European standardisation Journal of Cultural Heritage 11, 205-219. Doi: 10.1016/j.culturher.2009.02.008.

Carr-Brion, K., 1986: Moisture Sensors in Process Control. Elsevier Applied Science, Amsterdam.

Childs, P. R. N., 2001: Practical Temperature Measurement. Butterworth-Heinemann, London.

Davey, F.K., 1965: Hair Humidity Elements, pp. 571-573 in A. Wexler (ed.): Humidity and Moisture Vol. 1: Principles and Methods of Measuring Humidity in Gases. Rehinold, New York.

Fernicola, V., 2005: Measuring Air Humidity, pp. 97-112 in: The Chalcolithic Mummy, Vol. 3, Chap. 6, Collana del Museo Archeologico dell'Alto Adige, Folio Verlag, Vienna and Bolzano.

Fernicola, V., Banfo, M., Rosso, L., Smorgon, D., 2008: Investigations on Capacitive-based Relative Humidity Sensors and Their Stability at High Temperature, Int. J. Thermophysics, Vol. 29, pp. 1668-1677.

Liptak, B.G., 2003: Instrument Engineers' Handbook. Process Measurement and Analysis, Vol.1. CRC Press, Boca Raton, London, New Your, Washington D.C.

Macleod, K.J., 1983: Relative Humidity: Its Importance, Measurement and Control in Museums. Canadian Conservation Institute, Ottawa.

Michalski, L., Eckersdorf, K. and McGee, J., 1991: Temperature Measurement. Wiley, New York.

Michalski, S., 2000: Guidelines for humidity and temperature in Canadian archives. Canadian Conservation Institute, Ottawa.

Nicholas, J.V. and White, D.R., 1994: Traceable Temperatures -An Introduction to Temperature Measurement and Calibration. Wiley, New York.

Pande, A., 1975: Handbook of Moisture Determination and Control, Dekker.

Rosenhow, W.M., Hartnett, J.P., and Ganic', E.N., 1985: Handbook of Heat Transfer Applications, Mc Graw-Hill, New York.

Saucier, W.J., 1989: Principles of Meteorological Analysis, Dover, New York.

Thomson, G., 1986: The Museum Environment. Butterworths.

UK Meteorological Office, 1981: Handbook of Meteorological Instruments - Vol.2 Measurement of Temperature. Her Majesty's Stationery Office, London.

UK Meteorological Office, 1981: Handbook of Meteorological Instruments - Vol.3 Measurement of Humidity. Her Majesty's Stationery Office, London.

Wexler A. and Ruskin, R.E., 1965: Humidity and Moisture, Measurement and Control in Science and Industry, Vol 3: Fundamentals and Standards. Rehinold, New York.

Wolfe, W.L. and Zissis, G.J., 1989: The Infrared Handbook. Environmental Research Institute of Michigan.

World Meteorological Organisation, 1983: The Guide to Instrument and Methods of Observation, WMO, Technical Publication No 8, Geneva.

World Meteorological Organisation, 1986: Compendium of Lecture Notes on Meteorological Instruments for Training Class III and Class IV Meteorological Personnel, WMO Technical Publication No 622, Geneva.

World Meteorological Organization, 1987: International Cloud Atlas, Vol.II. WMO, Geneva.

CEN standards under discussion related to temperature and humidity measurements

CEN, 2010a: Draft European Standard CEN/TC346 prEN 15758, Conservation of Cultural Property - Indoor Climate - Procedures and instruments for measuring temperatures of the air and of the surfaces of objects. European Committee for Standardisation, Brussels. At present at the final formal vote.

CEN, 2010b: Draft European Standard CEN/TC346, Conservation of cultural property – Procedures and instruments for measuring humidity in the air and moisture exchanges between air and cultural property. European Committee for Standardisation, Brussels. Adoption n.185 – Munich 30 Oct 2009.

Published standards related to temperature and humidity measurements

EN ISO 7726: 2001, Ergonomics of the thermal environment -Instruments for measuring physical quantities

EN ISO/IEC 17025: 2005, General requirements for the competence of testing and calibration laboratories

EN IEC 60751: 1995 + EN IEC 60751/A2: 1995, Industrial Platinum Resistance Thermometer Sensors.

EN IEC 60584-1: 1995, Thermocouples – Part 1: Reference Tables.

EN IEC 60584-2: 1995, Thermocouples -Part 2: Tolerances.

ASHRAE, 2003: 2003, HVAC APPLICATIONS Handbook (SI), Chapter.21: Museum, libraries and archives., Atlanta: American Society of Heating, Refrigeration and Air-Conditioning Engineers. pp. 21-1, 21-16.

ASTM: 1999, Annual Book of ASTM Standards Section 14: General Methods and Instrumentation:

Temperature Measurement (Annual Book of A S T M Standards Volume 14.03), ASTM Intl, Vol. 14.

ASTM E879-01, Standard Specification for Thermistor Sensors for Clinical Laboratory Temperature Measurements. Benedict, R., 1984: Fundamentals of Temperature, Pressure, and Flow Measurements. Wiley, New York.

ASTM D4230-02: 2007, Standard Test Method of Measuring Humidity with Cooled-Surface Condensation (Dew-Point) Hygrometer.

ASTM E337-02: 2007, Standard Test Method for Measuring Humidity with a Psychrometer (the Measurement of Wet- and Dry-Bulb Temperatures).

CEN prEN 15757: 2008, Conservation of cultural heritage - Specifications for temperature and relative humidity to limit climate-induced mechanical damage.

EN ISO/IEC 17025: 2005, General requirements for the competence of testing and calibration laboratories.

ISO 4677/1: 1985, Atmospheres for conditioning and testing; Determination of relative humidity; Part 1: Aspirated psychrometer method.

ISO 7726, Ergonomics of the thermal environment - Instruments for measuring physical quantities.

ISO/IEC Guide 98, Guide to the expression of uncertainty in measurement (GUM).

ISO/IEC Guide 99, International vocabulary of metrology - Basic and general concepts and associated terms (VIM).

PD ISO/TR 18931: 2001, Imaging materials. Recommendations for humidity measurement and control.

UNI 11131: 2005, Cultural heritage - Field measurement of the air humidity.

DIN 50012-1: 1986, Climates and their technical application; methods of measuring humidity; general.

DIN 50012-2: 1986, Climates and their technical application; methods of measuring humidity; psychrometers.

NF X15-117: 1999, Measurement of air moisture. Mechanical hygrometers.

NF X15-119: 1999, Measurement of air moisture. Salt solution humid air generators for the calibration of hygrometers.

NF X20-521: 1981, Gas analysis. Determination of the water dew point of natural gas. Cooled surface condensation hygrometers.

Chapter 3
Microclimate monitoring in a Church

Dario Camuffo, Chiara Bertolin, Vasco Fassina

Introduction

The physical variables and the instruments needed to monitor air and surface temperature and relative humidity have been presented in the previous chapters. It may be useful to see an example of the use of some instruments in a case study, i.e. controlling the performances of a warm-air heating system in the church of Santa Maria Nascente in Agordo, Italian Alps (Fig.1). The aim of this chapter is to comment some basic data collection and analysis over a short winter monitoring, from December to February. The impact of heating on historical buildings and places of worship has already been discussed in other papers (Camuffo, 1998; Camuffo et al., 2007, 2009). For this reason we will not draw any conclusion about the HVAC performances and related damages to objects or structures but we will restrict ourselves to illustrating a sampling methodology.

The church monitoring was divided into two parts. Permanent monitoring was made with two automatic data loggers that sampled air and surface temperature and relative humidity (RH) at different levels, from the floor to the ceiling, in two positions. Sampling was programmed every ten minutes to get reliable information during the services, when the church is heated and crowded. Spot surveys were made for an extra acquisition of data taken with manual instruments. Such kind of monitoring, which required the action of two observers, was conduced on business days, without celebrations, but with the heating system on for the same time and under the same conditions as on Sunday, in order to get data without disturbing celebrations.

Instruments and methodology

Routine air and surface wall temperature monitoring was made with platinum resistance sensors (PT100, accuracy ±0.1°C) and RH was measured with capacitive polymer sensors (accuracy ±2%). Psychrometric measurements were excluded because of the cold season and the height above mean sea level that would have caused the wet bulb and the water reservoir to freeze.

Vertical temperature and RH profiles were monitored near the choir loft over the whole period. Sensors were hung to a rope to form a chain with sampling at selected heights. All sensors were removed from cases or probes, to avoid any thermal influence from the supporting structure. They were left unshielded in the free air to allow better ventilation in the absence of shortwave radiation emitters, i.e. no need for shielding.

Manual observations as follows. Air temperature and RH distribution in a horizontal cross section of the church, 1.5 m above the floor was directly measured with a Pt 100 sensor (accuracy ±0.1°C) and capacitive polymer sensors (accuracy ±2%). It was important to use the same fast-response instrument displacing it from a sampling location to another, because in this way the thermal map is reconstructed in a differential way, considering only differences from one sampling point to another. In the case of simultaneous readings with different instruments, the problem is to be certain that all instruments respond exactly in the same way, following the same calibration curve. More about problems and solutions concerning this methodology can be found elsewhere (Camuffo, 1998):

Fig.1 - The church of Santa Maria Nascente in Agordo, Italian Alps.

Vertical temperature profiles were also measured with the blackbody strip methodology, i.e. a strip made of a good absorber material with low thermal hystheresis in order to respond quickly and reach thermal equilibrium with ambient air and IR radiation.

To assess churchgoers' thermal comfort a number of blackbody strips was located in selected positions between the rows of pews in different parts of the church. The height of the strips was 180 cm, like a standing man. The blackbody strip temperature was monitored with a precision radiometer every 10 cm height in order to obtain a high-resolution thermal profile.

The precision radiometer is a modified Raytek Raynger MX4+ with a chip included to compensate for the error generated by the temperature difference between target and sensor. After this compensation the instrument provides accuracy better than ±0.1°C as verified with a calibration made in the range from -20° to +300°C. The instrument works in the spectral bandwidth 8 to 14 μm and the sensor is a thermopile. This radiometer was also used to monitor the wall and other surface temperature in selected points. Pointing is laser assisted.

Thermal images of the church interior, and of some selected surfaces were taken with a FLIR B400 infrared camera.

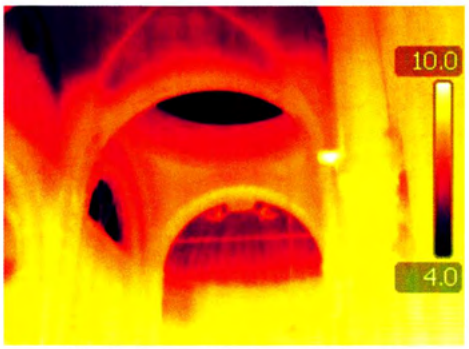

Fig.2 - Infrared picture of the nave and the vaults. Temperature scale on the right side. Coldest parts are blue and warmest yellow.

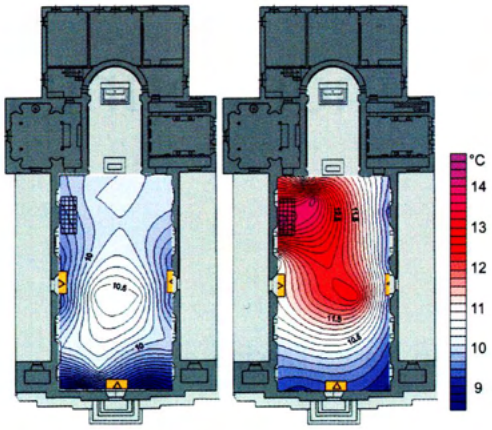

Fig.3 - Air temperature in a church cross section 1.5 m above the floor. On the left side the unperturbed situation before heating; on the right the temperature distribution after one hour heating operation. Doors indicated with orange rectangles and emission grille with a grid on the left side corner. Measurements 11-12-2009, 16:30 and 17:30 respectively.

Fig.4 - (a) The main door is provided with two series of slits to allow the exhaust of internal air. Red arrows indicate internal air going out and blue ones external air leakage. (b) Infrared image of the door and the cold wall of the counterfacade. (c) Detail of the door slits and in/out air leakages.

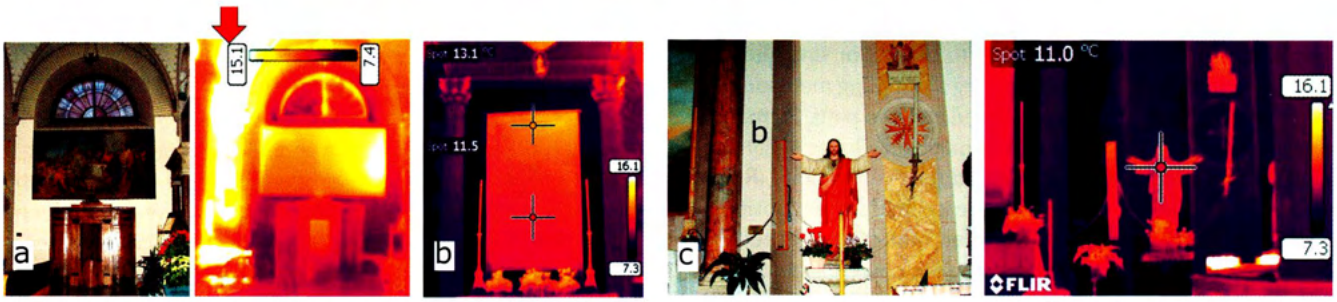

Fig.5 - (a) Infrared picture of a painting on canvas over a wood confessional. The warm air blown from the floor grille overheats the painting. The red arrow shows the wall overheated by the warm air jet. (b) A painting on canvas above a side altar between two chandeliers and two columns. The canvas is affected by a temperature gradient, i.e. 1.5°C/m. The wall behind it remains cold. (c) Visible and IR pictures of a wooden statue of Jesus Christ.

This camera works in the range 7.5 to 13 μm, 0.05°C at +30°C thermal sensitivity, ±2°C or ±2% reading accuracy, 1.3 megapixel visual images and a number of interchangeable optical lenses. This thermocamera was useful to see the impact of warm air on the surface temperature of objects made of diverse materials, e.g. masonry, wood, paintings on canvas, tapestry. In addition, having a very narrow angle for spot measurements, i.e. 1.36 mrad spatial resolution, it was possible to monitor the temperature of a strip of paper raised from the floor to the ceiling in order to obtain a vertical profile of the air temperature. Pointing is visible in the camera and laser assisted.

The quasi-contact temperature measurements have been performed with a total radiation sensor produced by FINGER, ±0.1°C accuracy. This instrument is useful because it shields the target surface from external radiation and converges on a photosensor all the direct and diffuse *IR* radiation emitted by the surface, as it had emissivity equal to 1. In addition, no physical contact is necessary with the target surface.

All measurements were made following the recommendations of the Standard prEN15758 (CEN, 2010a).

Data discussion and analysis

In winter, when the heating is off, the natural situation inside the church shows warmer floors and colder ceilings. The unperturbed temperature difference between floor and ceiling is however modest, e.g. 1° to 2°C in the masonry structure, and much more (5° to 6°C) in correspondence of the heat dissipated through the roof, and particularly through the vaults and the lunettes for the very poor insulation they have (Fig.2). A warm floor and a cold ceiling are the motor for a ceaseless, slow convective mixing of the indoor air. After heating is operated, warm air is blown in from a grille in the floor, reaches the upper part of the nave and spreads in the church. At churchgoer level, it is possible to monitor the warm air spread, and the expected thermal comfort, by displacing a fast temperature sensor in a series of different positions in a grid representative of a horizontal cross-section of the church at 1.5 m above the floor. The initial situation and the temperature distribution after 60 min of this operation are reported in Fig.3. It is evident that people located near the grille and in the centre of the nave benefit from the heating; people near the opposite wall or in the rear part of the church suffer from the cold structures and the air

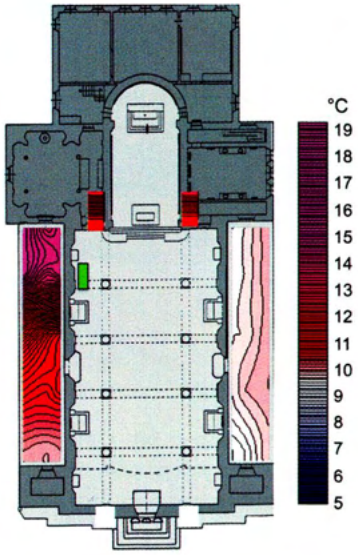

Fig.6 - Temperature distribution on the internal side of the church walls from1.5 to 3 m above floor level. Emission grille indicated with green rectangle. Measurements taken 11-12-2009, after one hour heating operation.

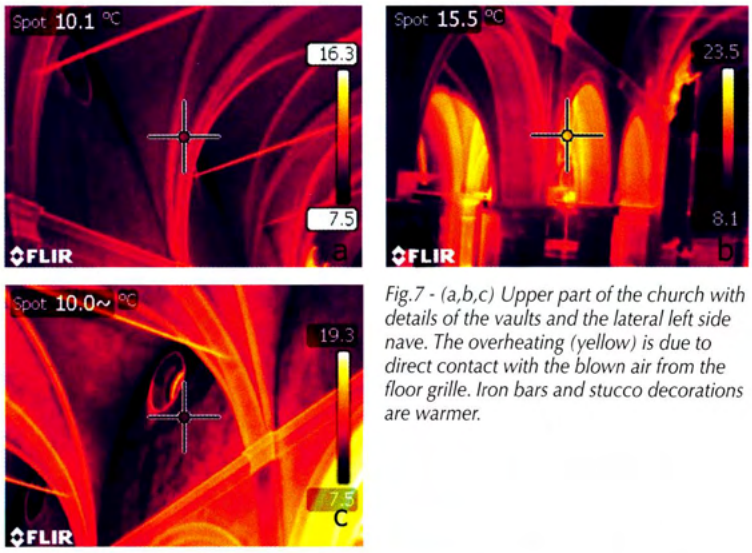

Fig.7 - (a,b,c) Upper part of the church with details of the vaults and the lateral left side nave. The overheating (yellow) is due to direct contact with the blown air from the floor grille. Iron bars and stucco decorations are warmer.

leakage through the main door. The main door (Fig.4a) is provided with two series of slits to allow the exhaust of internal air that is continually renewed with warm air blown in by the heater fan. The thermal images of the door (Fig.4b, c) show that the two series of slits and all the leakages of warm indoor air through the door are overheated; on the other hand, cold external air, having a higher density, enters through the fissure between door and floor, lowering the temperature in the rear part of the church.

The warm air dispersed in the lower part of the church for the churchgoers' comfort also affects paintings on canvas, wooden structures and wall temperature. Paintings on canvas show the quickest warming, followed by wood and then by masonry. An example of how such surface temperatures are increased is reported in Fig.5a,b,c. The wall temperature distribution (Fig.6) repeats in some way the air temperature distribution in Fig.3, being highest near the emission grille and lowest in the counterfacade. In the upper part of the church, the vault above the emission grille is the warmest part (Fig.7a). The other vaults remain colder, although ribvaults, metal chains and stucco decorations are quickly heated (Fig.7b,c).

Before the heating system is operated, the organ (Fig.8a,b) is in equilibrium, at a thermal level midway between the warmer floor and the colder ceiling. When heating is operating, the contact with warm air immediately raises the temperature of wooden parts and the organ's metal pipes. The organ becomes the warmest object in the upper part of the church. The temperature increase causes a drop in RH and the delicate organ needs to be regularly checked and tuned.

The vertical temperature profile has been measured with the thermocamera readings in the spot measurement mode, sampling on a vertical blackbody strip (Fig.9a,b,c). The blackbody strip was obtained by unrolling a 27-cm-wide roll of paper and raising the paper strip from the floor to the ceiling with a rope. It is assumed that the blackbody strip, made of a paper, is in thermal equilibrium with the air. Once the paper strip reached the required vertical position, it was possible to sample the paper temperature at a number of levels in order to obtain the required profile. A further advantage of this methodology is that any picture shows, in addition to the air temperature, the surface temperature of ceiling, walls and whatever is included in the optical view angle. This is possible because most of the surfaces have emissivity values close to each other, i.e. paper: 0.92-0.93; plaster: 0.91-0.92; soot: 0.91-0.94; wooden pews and furniture: 0.90-0.93; oil paintings on canvas: 0.94. By setting the instrument emissivity at 0.92, all the surfaces are investigated with the appropriate emissivity to a close approximation (Fig.10a). From the temperature and *RH* measurements taken at 1.5 m above the floor, it is possible to calculate the mixing ratio and calculate how the thermal layering is influential on the *RH* distribution at different heights. This is shown in Fig.10b and we see that the upper part is affected by 10°C temperature increase and 23% RH drop; the lower part by 6°C and 15% RH.

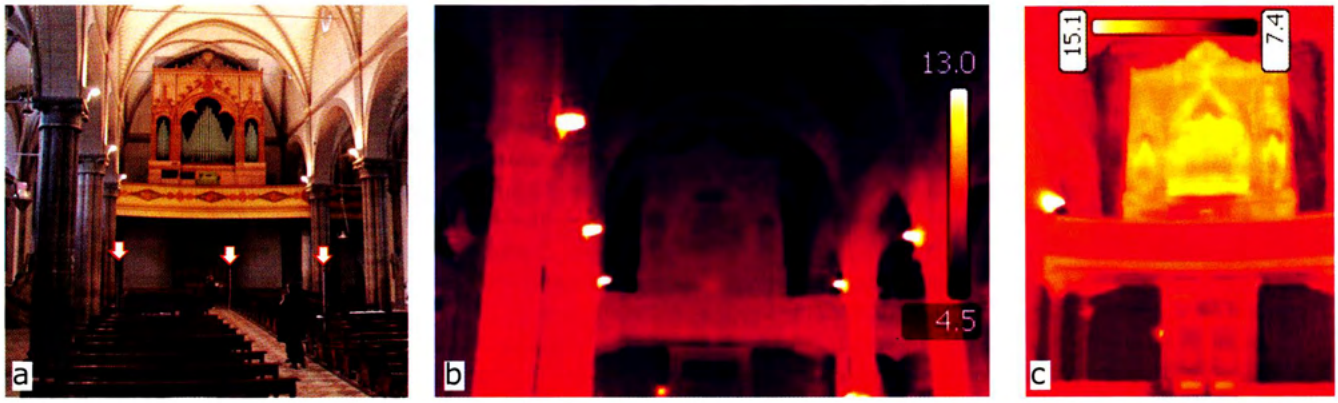

Fig.8 - Visible (a) and infrared (b,c) pictures of the rear part of the church with the organ. The organ is initially (b) in thermal equilibrium with the church at around 6°C; after one hour operation (c) it becomes the warmest part and approaches 15°C.

Fig.9 - Visible (a) and infrared (b,c) pictures of the rear part of the church with the organ and the blackbody strip from the floor to the ceiling made with white paper roll. Figures b and c at mid heating and after one hour operation, respectively. In (c) the paper strip shows a clear thermal layering. The organ temperature follows closely the air temperature.

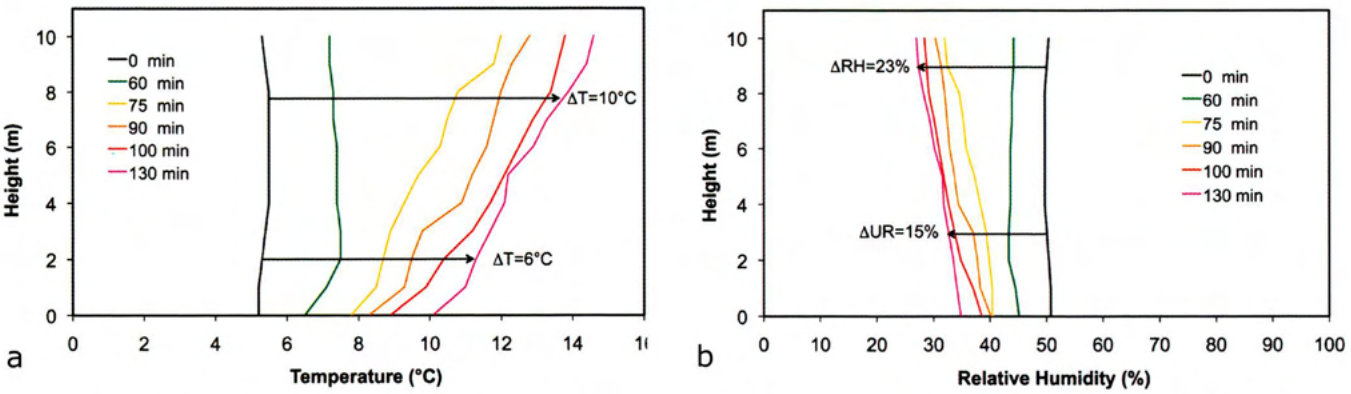

Fig.10 - (a) Vertical temperature profile in the church before operating (0 min, blue) and at different time intervals up to 130 min (violet) operation. (b) The same but for relative humidity.

The vertical profiles of air temperature for human comfort were also measured with the blackbody strip method (Fig.11a,b) in different parts of the pew area in church, i.e. near the main entrance, in the middle on the right and left sides and in the front pews, not far from the emission grille. The result (Fig.12) is that the front position benefits from 4°C at face level and 1.5°C at foot level; the two central positions from about 1.5°C at face and 0°C at foot level; the rear position suffered from some door opening and the leakage counteracted the face heating and cooled the foot level by 1.5°C. The balance is that 1.5°C average face heating and 0°C foot heating implies 10°C warming and 23% RH dropping in the upper part, that is not beneficial to the organ.

Conclusions

This case study has shown that the same variable, e.g. air or surface temperature, can be measured with different methodologies and instruments depending on the aims, materials and local constraints. The objective is to investigate some specific mechanisms that may cause potential damage to cultural heritage and to be prepared to monitor the key features useful to assess the environmental risk on objects. This implies using instruments in the most convenient way, or finding solutions able to respond to some specific needs. Any measurement should be planned only after we clearly have in mind what we want to know, how to reach this goal, with what kind of instruments, when and where to monitor.

In this specific case, the goal was to know the performances of the warm-air heating system and the potential impact on the organ and other works of art. After this monitoring, the conservator's knowledge of the situation is improved, and he knows better what to do, and what not to do, to provide a better preservation. For instance, when warm air is blown in to heat the church, paintings on canvas have the same temperature as air, wood a bit less, the masonry is heated much less, except in the area directly hit by the warm airflow.

Adding moisture to the blown air would balance the *RH* level in the air and it would be good for canvas; however, added water would risk condensing on walls, stucco decorations and frescoes with a much lower temperature level, creating dampness and moulds. We are not supposed to preserve the air, but works of art, and we should know how they react when we use or enjoy them. For this reason, measurements limited to air temperature and relative humidity alone are of little help, and a holistic view of the whole problem is necessary.

Departing too much from the historical climate is dangerous and is not recommended by the Standard prEN15757 (CEN, 2010b). In reality, it is hard, or even impossible, to

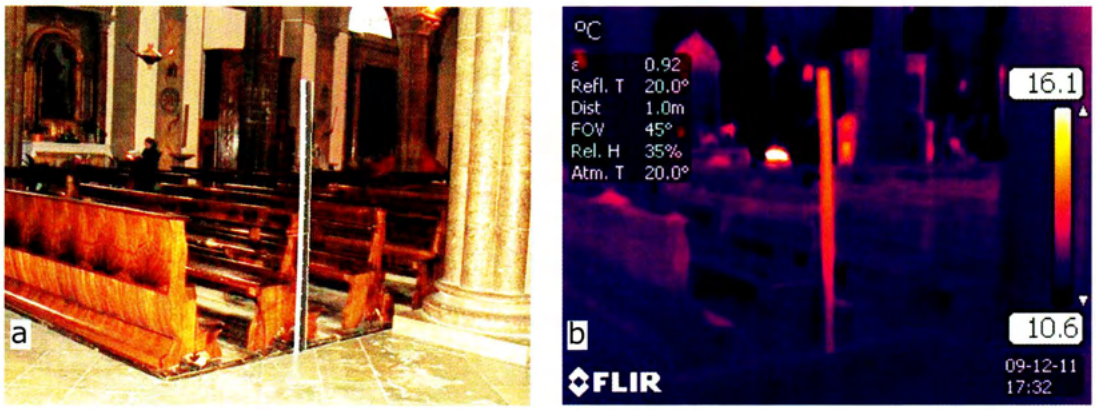

Fig.11 - Visible (a) and infrared (b) pictures of the blackbody strip used to assess churchgoer thermal comfort. The IR image shows thermal layering.

counteract the negative effects of any form of central heating, and also to stop warm air reaching artworks disseminated in various parts of the church. The best strategy is determined by two factors, i.e.: to use localised heating, i.e. keeping heat concentrated in the churchgoer area, and to limit heat emission, i.e. accepting a compromise between the churchgoers' thermal comfort and the cultural heritage conservation needs (Camuffo et al., 2007, 2009). This is recommended by the Standard prEN15759 (CEN, 2010c).

Acknowledgements

This paper was made possible thanks to the European Science Foundation funding and the scientific activity performed within the COST action D42: *Enviart*. The paper unfolds results achieved during EU funded projects, namely AER (ENV4-CT95-0088), FRIENDLY-HEATING (EVK4-2001-00007), SENSORGAN (EVK4-2001-00007) and *Climate for Culture* (GA 226973). Thanks are due to the Paris Church for having facilitated our work.

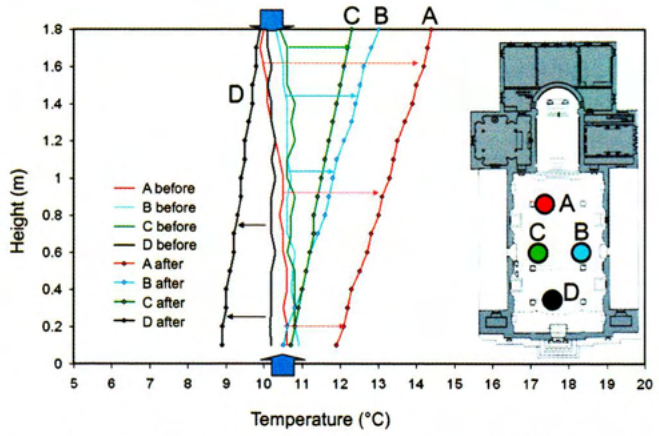

Fig.12 - Vertical profiles of the air temperature in various parts of the church, position and colour codes on the right hand church map. Blue arrows indicate the temperature before heating. Measurements 11-12-2009, 16:30 and 17:30.

References

Camuffo, D., 1998: Microclimate for Cultural Heritage. Developments in Atmospheric Science 23, Elsevier, Amsterdam, 415 pp.

Camuffo, D., Pagan, E., Rissanen, S., Bratasz, Ł., Kozłowski, R., Camuffo, M., della Valle, A., 2010: An advanced church heating system favourable to artworks: a contribution to European standardisation Journal of Cultural Heritage 11, 205-219 Doi: 10.1016/j.culturher.2009.02.008 (in press).

Camuffo, D., Pagan, E., Schellen, H., Limpens-Neilen, D., Kozlowski, R., Bratasz, L., Rissanen, S., Van Grieken, R., Spolnik, Z., Bencs, L., Zajaczkowska-Kloda, J., Kloda, P., Kozarzewski, M., Mons. Santi, G., Chmielewski, K., Jütte, T., Haugen, A., Olstad, T., Mohanu, D., Skingley, B., Sáiz-Jiménez, C., Bergsten, C.J., Don Russo, S., Bon Valsassina, C., Accardo, G., Cacace, C., Giani, E., Giovagnoli, A., Nugari, M.P., Pandolci, A.M., Rinaldi, R., Acidini, C., Danti, C., Aldrovandi, A., Boddi, R., Fassina, V., Dal Prà, L., Raffaelli, F., Bertoncello, R., Romagnoni, P., Camuffo, M. & Troi, A., 2007: Church heating and preservation of the cultural heritage: a practical guide to the pros and cons of various heating systems. Electa Mondadori, Milan, 240 pp.

CEN, 2010a: Draft European Standard CEN/TC346 prEN 15758, Conservation of Cultural Property - Indoor Climate - Procedures and instruments for measuring temperatures of the air and of the surfaces of objects. European Committee for Standardisation, Brussels.

CEN, 2010b: Draft European Standard CEN/TC346 prEN15757, Conservation of Cultural Property - Specifications for temperature and relative humidity to limit climate-induced damage in organic hygroscopic materials. European Committee for Standardisation, Brussels.

CEN, 2010c: Draft European Standard CEN/TC346 prEN15759, Conservation of Cultural Property - Indoor Climate - Heating Places of Worship. European Committee for Standardisation, Brussels.

Chapter 4

Acceptable and non-acceptable microclimate variability: the case of wood

Łukasz Bratasz

Hierarchical structure of wood

Cellulose is one of the major constituents of the wood cell wall and confers rigidity to the plant. Cellulose molecules are chains of sugar units, the basic unit being formed by two sugar rings. The native form of cellulose in wood cell walls are thin microfibrils which contain nanocrystals of the cellulose. Typically, the hexagonal axis of the crystal structure follows the fibril axis.

A crystal of cellulose present in the wood cell wall of spruce (Picea abies) is shown in Figure 1c. The size of the cellulose nanocrystals is species-dependent and is 2.5 nm in spruce wood. Cellulose fibrils are wound helically in the dominant cell-wall layer S2 (Fig. 1b) and a microfibril angle μ is an important parameter responsible for the mechanical property of the wood cell and its shrinkage on drying. Two fracture surfaces from wood cells with different microfibril angle μ are visible in Figure 1b. Wood is formed by an assembly of cells as shown in Figure 1a, where the cell-wall thickness is larger in latewood (LW) than in earlywood (EW). The detailed structure of the wood cell wall has been determined with the use of various sophisticated techniques of structural analysis such as synchrotron scanning X-ray microdiffraction or high resolution microscopies (Fratzl 2003).

The concentric arrangement of annual growth rings as well as the longitudinal orientation of the wood cells provide basis for designating three anatomical directions within the wood: longitudinal, or parallel to the grain, radial and tangential. The three fundamental planes of wood are the transverse (cross sectional), radial and tangential planes (Fig. 2). Many wood parameters are dramatically different in the three anatomical directions.

Water vapour sorption

Cellulose is a hygroscopic material gaining moisture (adsorption) when the relative humidity (RH) is high, or losing moisture (desorption) when the surrounding air is dry. The moisture content in wood exposed to a given temperature and RH eventually attains a constant level termed the equilibrium moisture content (EMC). The relationships between the RH and EMC at constant temperatures are termed water adsorption or desorption isotherms.

There has been general agreement that several mechanisms of wood – water interactions prevail at different levels of RH. The most comprehensive picture of these interactions is obtained when the desorption is carried out from the fully saturated state in which all cell cavities are filled with water and cell walls are fully swollen with water molecules adsorbed within the fibrillar structure. Such state is characteristic of the "green" wood of living trees but it can be produced also in laboratory by saturating a wood specimen with water for example by use of the pressure membrane procedure (Almeida & Hernandez 2007). Such desorption curve, shown schematically in Figure 4 for lime wood (*Tilia sp.*), is termed the boundary des-

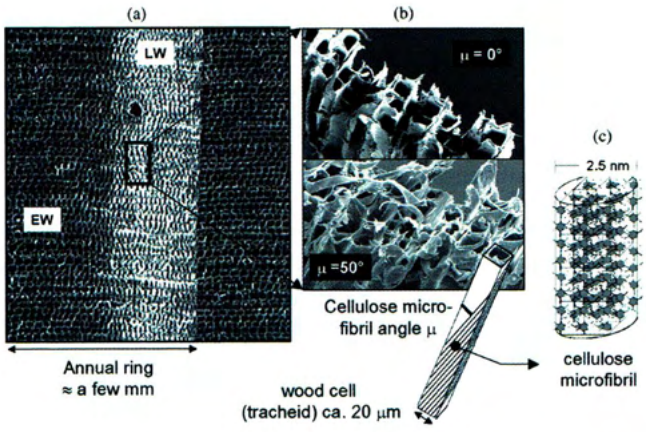

Fig. 1 - Hierarchical structure of spruce wood. (a) Cross-section through the stem showing earlywood (EW) and latewood (LW) within an annual ring. (b) Scanning electron microscopic pictures of fracture surfaces with two different microfibril angles. One of the wood cells (tracheids) is drawn schematically showing the definition of the microfibril angle between the spiralling cellulose fibrils and the tracheid axis. (c) Sketch of the crystalline part of a cellulose microfibril (from Fratzl 2003, courtesy of Elsevier).

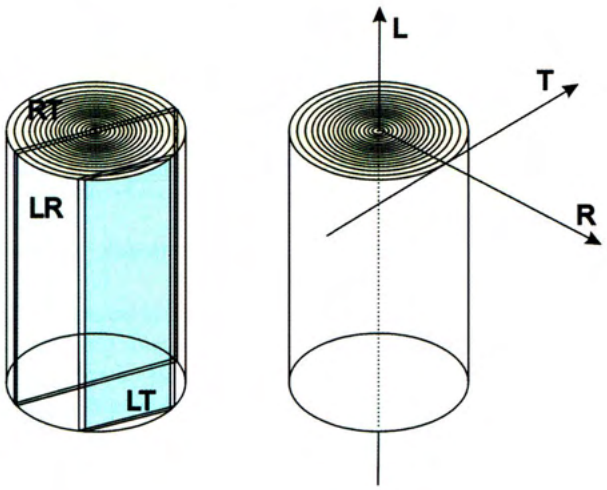

Fig. 2 - The fundamental anatomical directions (L-longitudinal, T-tangential, R-radial) and three structural planes within the wood (RT-cross sectional, LR-radial and LT-tangential).

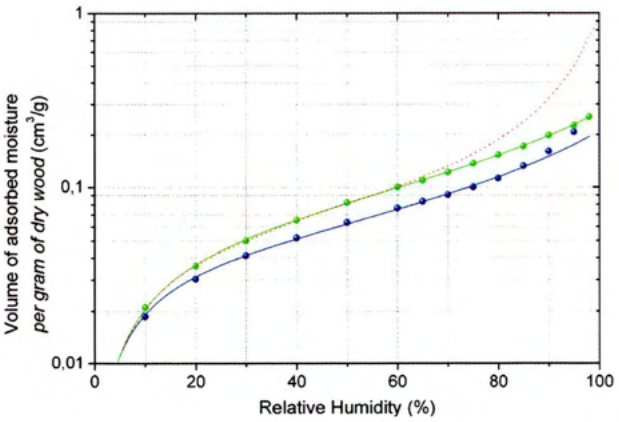

Fig. 3 - Sorption isotherms of lime wood at 24°C; blue and green lines show prediction of the GAB equation for adsorption and desorption branches.

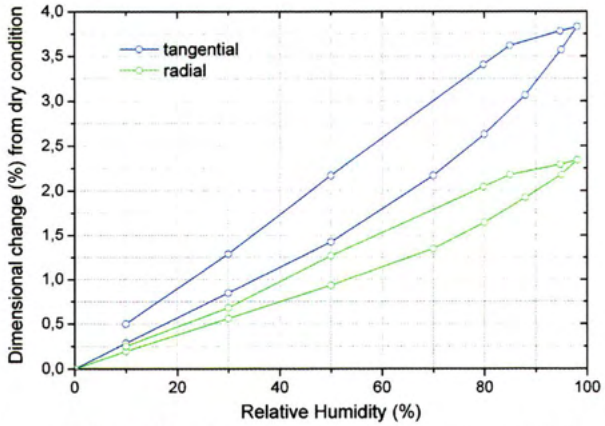

Fig. 4 - Swelling and shrinkage isotherms for walnut wood in the radial and tangential directions.

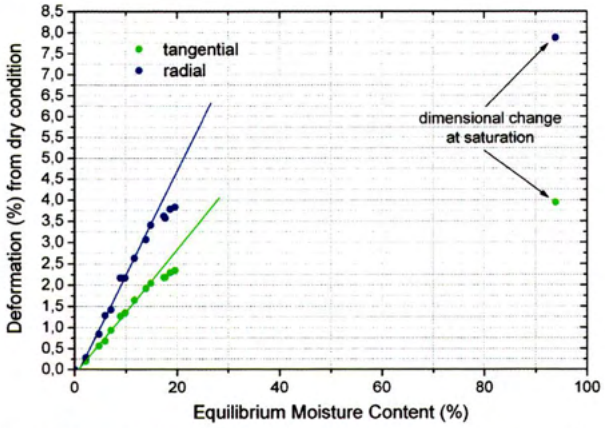

Fig. 5 - Dimensional change of walnut wood in the radial and tangential directions plotted as a function of the equilibrium moisture content.

orption curve and is plotting the maximum EMC expected for each RH condition on desorption.

Two types of water are present in wood at saturation: 'bound water' adsorbed on the in the cell walls and 'liquid or lumen water' present in the capillaries of the wood microstructure. On decreasing RH, the capillaries are gradually emptied until Fiber Saturation Point (FSP) is reached which is defined as EMC at which the cell walls are saturated with bound water, with no liquid water in the lumens. However, it was demonstrated that a loss of bound water can occur at EMC well above the FSP with liquid water entrapped in the smallest capillaries and channels in wood. The loss of liquid water is usually complete at RH levels below 60%.

After a complete drying of wood, an increase in RH leads first to a monolayer and then to multilayer physical sorption of water molecules. An adsorbed water layer progressively thickens as the vapour pressure is increased up to the saturation pressure. The desorption curve is located below the boundary desorption curve at high RH values as generally no liquid water is introduced by the adsorption process. The two curves join together at about 60% RH. A distinct hysteresis effect is observed i.e. higher moisture content during desorption when compared to that during adsorption at any given RH value. This phenomenon is associated with the swelling of a non-rigid wood structure in the course of adsorption so that the effect is in fact the elastic hysteresis of wood.

ùThe relationships between the amount of the bound water in the wall cells, and RH is described by a sigmoid shape of the type II isotherm in the IUPAC 1985 classification (Sing et al. 1985). In recent years, Guggenheim-Andersen-de Boer (GAB) sorption equation has been used to interpret the physisorption of water on wood as it is capable of describing the shape of the type II isotherm in a RH broad range of between 5 and 80-90% (Fig. 3) and yields meaningful physical parameters as monolayer capacity and energy constant related to the net molar heat of adsorption (Timmermann 2003). Furthermore, it has been shown that the GAB equation is equivalent to the physically inaccurate Hailwood and Horrobin equation (H-H) widely used to describe the moisture sorption by wood (Hailwood &Horrobin 1946).

Dimensional change

Dimensional change is perhaps the most important consequence of moisture sorption by wood. Wood shrinks as it loses moisture and swells when it gains moisture. Wood is anisotropic and its moisture-related dimensional changes vary in three principal anatomical axes; wood is dimensionally stable in the longitudinal direction and least so in the tangential direction.

The relationships between the RH and the relative dimensional change at constant temperatures are termed swelling or shrinkage isotherms. Such isotherms for the walnut wood (*Juglans regia*) are shown in Figure 4.

A convenient way of expressing swelling and shrinkage behaviour of individual wood species is to plot dimensional change as a function of EMC. Such relationships are shown in Figure 5 for the tangential and radial directions in the walnut wood. All the experimental points for the EMC values between 0 and 20% were obtained by adsorption and desorption of water vapour at varying RH whereas the last point corresponds to a wood specimen fully saturated under distilled water. Generally, the measurements did not show any differences between the dimensional change values at the same EMC but obtained after adsorption or desorption.

The analyzed relationship is described by a sigmoid function. The function shows an early exponential growth at low EMC, then slows to a linear growth for the broad medium part between approximately 2-15%, and then approaches a constant value as the capillary system of wood is saturated with water. The dimensional change coefficients are determined by the linear fit across the linear range of the plot and for the investigated wood are 0.15 and 0.25 for radial and tangential directions, respectively.

Measurements of the dynamic dimensional response of authentic historic wooden objects have been made possible through the use of a range of sensors or optical methods. Three case studies are described to illustrate the issues.

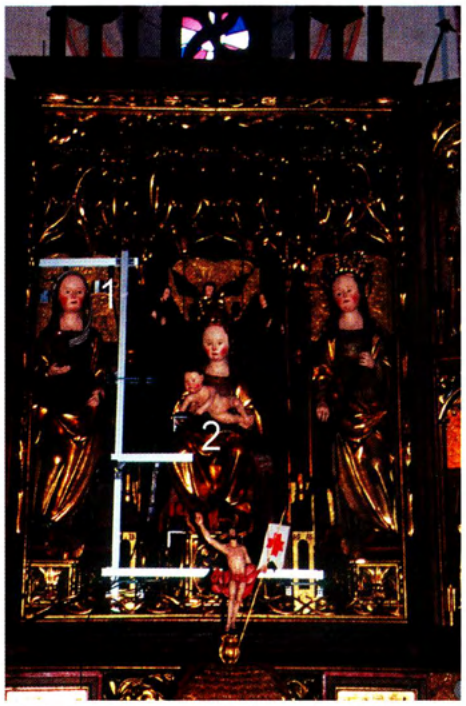

Fig. 6 - The main altar of the church of Santa Maria Maddalena in Rocca Pietore, Italy. The numbers indicate the position of wooden elements monitored for dimensional change: (1) the saint's head, (2) the Child's finger.

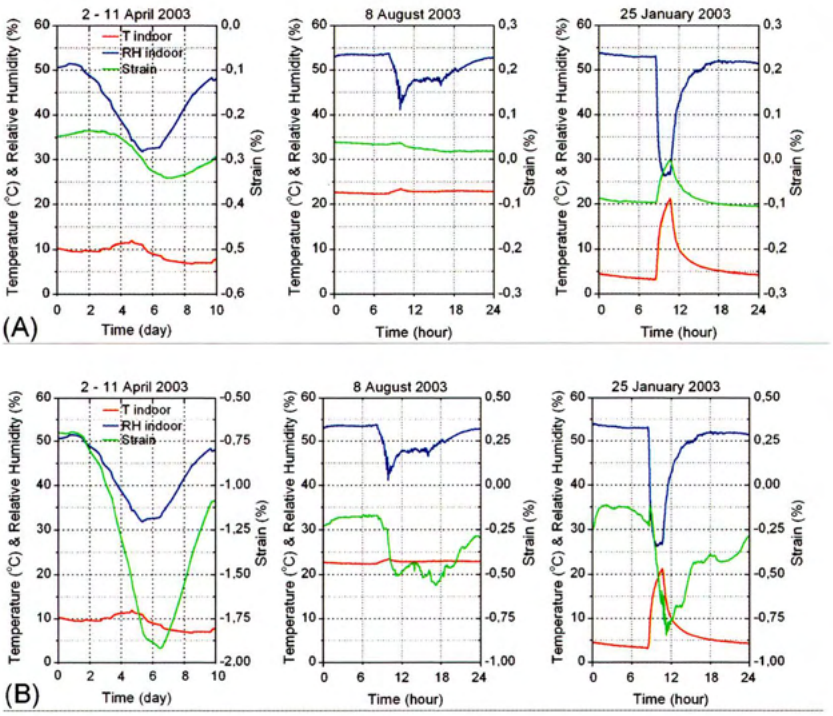

Fig. 7 - Relative dimensional changes (l/lo) of the saint's head (A) and the Child's finger (B) in response to three short variations of climate: weekly variation due to a change of weather outside, indoor diurnal cycle in summer and fluctuation caused by a heating episode in winter involving an inflow of warm air. Records of indoor T and RH are shown for comparison.

Case study 1. The altarpiece in the church of Santa Maria Maddalena in Rocca Pietore, Italy

Triangulation laser displacement sensors (Fig. 6) were used to measure on site the dimensional response of the altarpiece, a medieval polychrome and gilded wooden triptych, to variations of RH and temperature (Bratasz & Kozłowski 2005). The altar was carved in lime wood. The monitoring was part of a broader European research project Friendly Heating (2002-2005). The project developed a novel heating system consisting of low temperature radiant sources located in pews which provided direct confined heat just to people sitting and leave undisturbed the church as a whole (Camuffo et al. 2009). Previously, the church had an intermittent heating system based on a forced inflow of warm air which had negative impact on the church fabric and its contents as it generated short-term temperature peaks accompanied by drops in RH.The elements of the altar responded to the variability of indoor T and RH (Bratasz et al. 2007). The moisture content in the wood varied according to whether RH increased or decreased in the proximity of the altar. The variation in the moisture content caused dimensional changes in the wood; the response, however, was characterised by ranges and rates which varied considerably with the thickness of the wooden elements. A fine wooden element 0.5 cm thick, the Child's finger, expanded and contracted quickly and completely even during short-term RH changes (Fig. 7). Therefore, the external layer of all wooden elements of the altar, at least several millimetres thick, will be strongly affected by any, even rapid, change in RH.

On the other hand, the overall dimensional reaction of the saint's head – a massive wooden element 15 cm thick - was much slower as the uptake or release of the water vapour was too slow to produce uniform moisture content through wood. For short-term cycles, natural or due to heating episodes, practically no change in the overall dimension of this element was observed in association with changes in RH. Only thermal expansion due to an increase in temperature was observed. The slow overall moisture movement out of or into the wood, combined with an immediate response of the external layer, led to strong radial gradients in moisture content.

Case study 2. Differential dimensional response across a specimen, simulating a panel painting

There are two sources of restraint, and resulting stress, in the panel paintings subjected to variations of RH in their environment. As usually, an internal restraint is experienced by a panel as the moisture diffusion is not instantaneous and the outer part of wood dries more quickly than the interior on the reduction of RH. However, the response is more complex, when compared to a massive wooden cylinder considered above, due to a non-symmetric gradient across the thickness of the panel. Generally, only rear, uncoated face of the panel is able to exchange moisture with the environment, the other face coated with gesso, paint and varnish being largely impermeable to water vapour diffusion. Figure 8 illustrates the response of such specimen imitating a panel painting.

In the experiment, one face and the edges of a panel 15 cm

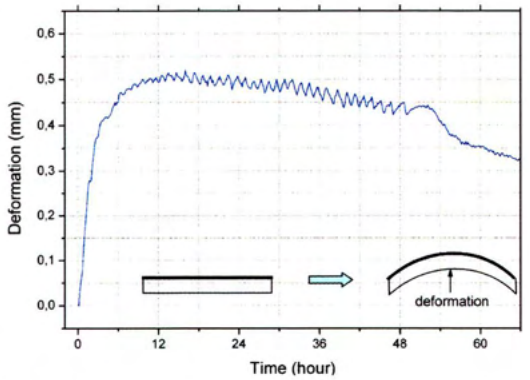

Fig. 8 - Cupping deformation of a specimen imitating a panel painting in response to a step RH variation from 70 to 35%.

Fig. 9 - The Mazarin Chest (412-1882). Courtesy of V&A Images.

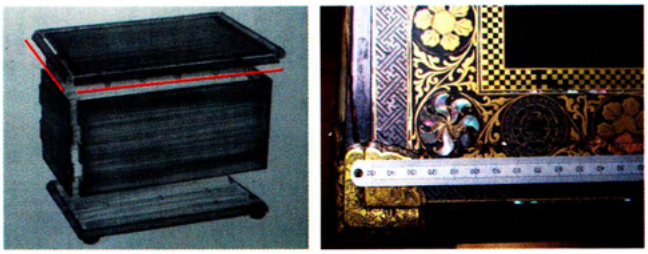

Fig. 10 - Construction of the Mazarin Chest with lines of restraint indicated (left). Damage areas in the lid (right). Courtesy of V&A Images.

(longitudinal to the grains) x 20 cm (perpendicular to the grains) x 1 cm (thickness) were coated by a silicone sealant to create the worst case condition, a completely asymmetrical diffusion of the water vapour through one face of the panel. The panel was stabilized at 70%rh inside a climatic chamber. Then it was removed from the climatic chamber, placed under 35%rh and the development of its deformation (cupping) was measured every 5 minutes using a laser scanner. The non-symmetrical evolution of the moisture gradient profile induced a differential dimensional response between the two faces resulting in a growing deformation; at the maximum of this unbalance the largest deformation was recorded and then faded out. If deformation is prevented, for example by cradling, stress is produced in the panel.

Stress development

A noteworthy effect of the dimensional change of wood may be the resulting high stress within the material, which can cause significant damage. If a relatively dry wood is restrained from free swelling in high RH regions, it can experience plastic deformation in compression with resulting crushing of the internal structure and buckling. In turn, if moist wood is restrained from free shrinkage on return to low RH regions, it will experience tension leading to irreversible stretching and eventual cracking.

The described mechanism would not be pertinent if the wood were unrestrained. But in real-world conditions the restraint is universally present. Restraint may result from the rigid construction restricting movement, or by joining wood elements with different mutual orientation of their fibre direction. Wood can also experience internal restraint as the moisture diffusion is not instantaneous and the outer part of wood will dry more quickly than the interior on the reduction of RH. The dry outer part will be restrained from the shrinkage by the still wet core beneath, which will result in mechanical stress: the outer shell will go into tension and the core into compression. Two case studies are described to illustrate damage caused by the external and internal restraint of the dimensional response in wooden objects, respectively.

Case study 3. The Mazarin Chest, Victoria & Albert Museum in London

The Mazarin Chest, renowned as one of the finest pieces of Japanese export lacquer to have survived from around 1640, is a star item in the V&A's collection of Japanese art (Fig. 9). The areas of restraint in the Mazarin Chest correspond to assemblies of cross grained wooden elements. A particular adverse condition is created by the cleated (hashibami) construction of the lid (Fig. 10). The cleats have acted as a restraint to the shrinkage of planks, which is the underlying cause of cracks in the wood and lacquer layers. Detailed analysis of the problem has indicated that the object must have been subjected to large, uncontrolled fall in ambient RH in the past; the RH variations in the current display environment do not endanger the chest (Bratasz et al. 2008).

Case study 4. Stress field resulting from the restrained differential dimensional response across a wooden cylinder, simulating sculptures

A case of a massive wooden cylinder, simulating a wooden sculpture, was considered, as it constitutes one of the worst cases in terms of risk of the climate-induced damage to wooden objects (Jakieła et al. 2008a). As mentioned above, the stress development in such an object is the result of slow moisture diffusion in

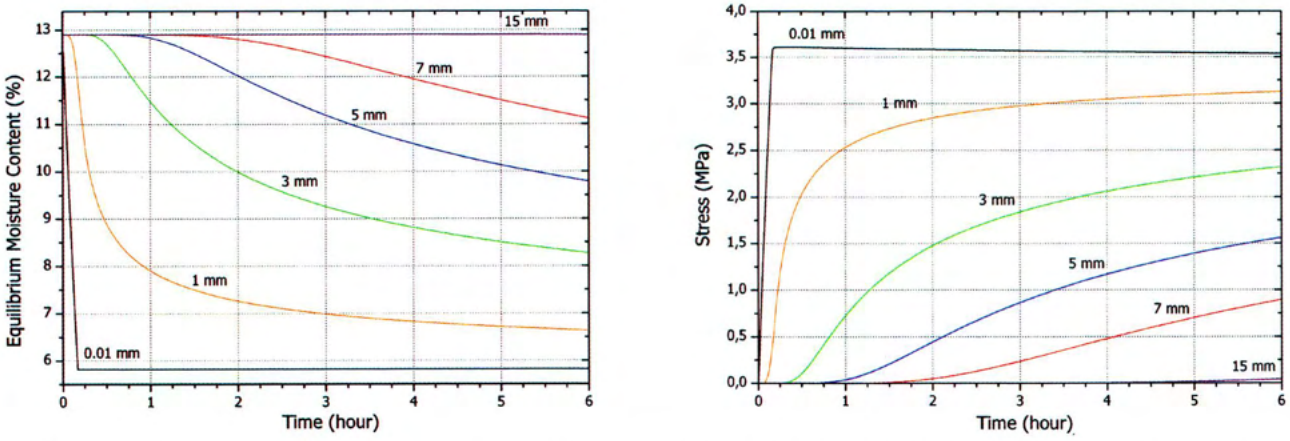

Fig. 11 - Change in the distribution of moisture content (left) and tangential stress developing in wood at selected distances from the external surface of a wooden cylinder for a step RH variation from 70 to 30%.

response to the change in ambient RH. Water vapour movement into or out of wood in response to variations of T and RH in its environment, and the evolution of the moisture content gradient across wood and the stress field thus created were analyzed by numerical modelling. The results for a step RH variation, in which RH was lowered instantaneously from 70 to 30%, are shown below for lime wood in Figure 11.

As anticipated, early in drying the change of moisture content is concentrated in the outer part of the cross-section. The surface shell, the first ring of the width of 10 μm, attains the new EMC almost immediately. However, the core of the cylinder, lying deeper than 1 cm into the wood, has not experienced any change in the moisture content before 3 h. After 24 h, the change in the EMC at this distance from the surface is only 30% of the change at the surface. The presented strong gradient in the moisture content gives rise to a considerable drying stress, due to the differential shrinkage restrained. The plots mirror the changes in the moisture contents at different strata of wood. The maximum tension of 5.5 MPa is attained very quickly at the surface layer. The stress decreases very slowly as moisture gradient gradually vanishes on a progressive drying of the interior layers in wood.

Critical levels of strain at which physical damage appears - allowable variations of RH

Relationships between stress and deformation (strain) must be determined to evaluate the critical levels of strain at which the materials begin to deform plastically or fail mechanically. Figure 12 shows such relationship for lime wood.

The relationship is linear at low strain values for which the material deforms elastically and will return to its original shape when the applied stress is removed. The slope corresponds to the elasticity modulus of the wood. It is noteworthy that wood is typically 50 times more rigid in the longitudinal direction than in the tangential one, and 25 times than in the radial one. The yield point corresponds to strain necessary to produce permanent (plastic deformation) and to the deflection at the stress-strain curve from a straight line. Though testing should determine the yield point for any specific wood samples of interest, Mecklenburg, Erhardt & umosa (1998) assumed that the yield point of wood is generally around 0.4% and postulated that the allowable RH variations should not cause strains exceeding this critical value. Therefore, the allowable RH variations can be determined at a first approximation by using the dimensional change coefficients reported in literature and simply calculating the change in RH that will cause a strain of 0.4%. As the dimensional change coefficients depend strongly on RH, the allowable RH variations have to be presented as functions of starting RH levels.

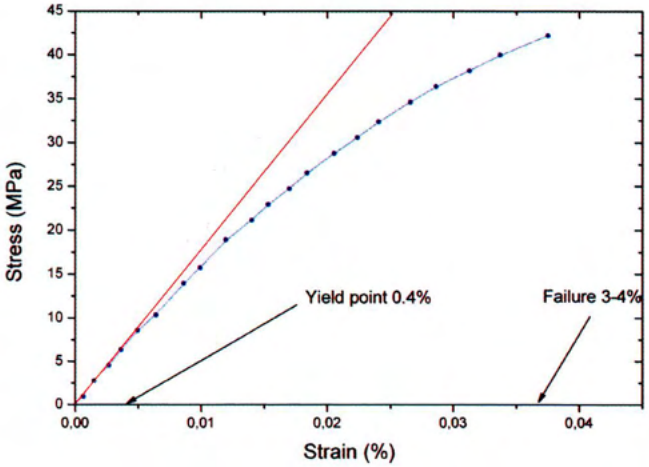

Fig. 12 - The stress-strain curve for lime wood loaded in radial direction.

An example of such map, calculated for a wooden cylinder of a diameter of 13 cm for two different time scales of the RH variations - instantaneous and diurnal - are shown in Figure 13 (Jakieła et al. 2008a).

The map of stress levels induced by the instantaneous RH variations shows that any variation exceeding 10-15% produces stresses exceeding the elastic limit, in which case wood can undergo damaging plastic deformation. The risk of failure appears already for the RH variation of 40% if the change occurs from a high initial RH level of above 95%. The variations centred on 50% RH can be considered the safest – a finding that confirms the established observations of the conservation practice. The reason for that finding is that the least slope of the dependence between RH and moisture content in wood, and the largest strength of wood, is in the mid-region of the water vapour adsorption isotherm.

The map of stress levels induced by slower, diurnal variations indicates an increase of the domain of RH variations that produce stress levels within the elastic limit safe for the material. RH variations potentially leading to failure have been practically eliminated. The calculations illustrate the old wisdom of preventive conservation that even a significant RH change can be safe if the adaptation time for a wooden object is long enough.

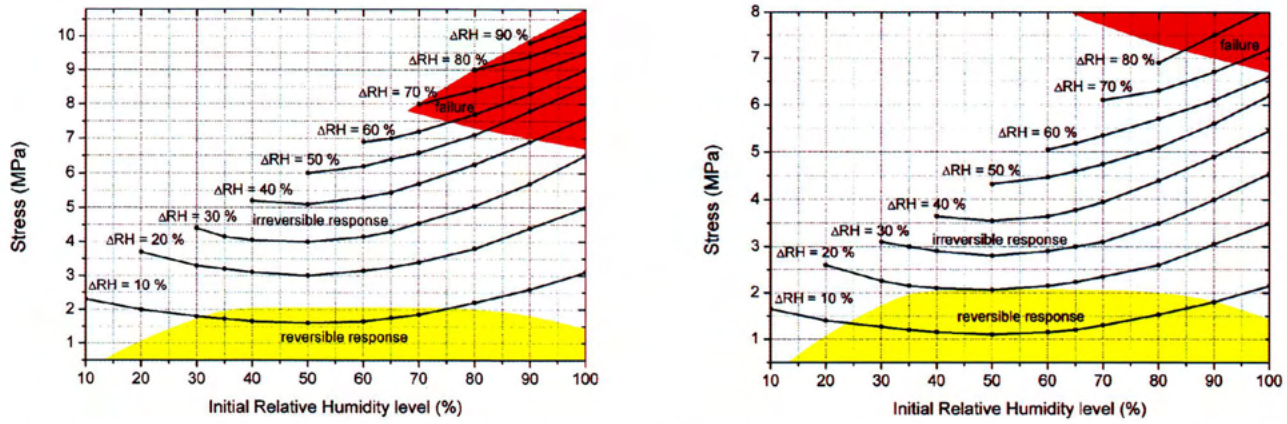

Fig. 13 - Stress induced by RH variations between 10% and 90%, instantaneous (left) and lasting 24 h (right), plotted as a function of the initial RH level from which the variation starts. Domains of RH variations endangering the wood by irreversible response (deformation) or complete failure are marked, as is the domain of tolerable variations that produce safe, reversible response of the wood.

Damage of decorative layers on wood and allowable variations in RH

Painted wooden objects are complex multi-layer structures composed of humidity-sensitive materials – wood, animal glue, gesso and paints. Moreover, each material responds differently to the loss and gain of moisture, which induces high stress in the different layers of painting leading to damage if the strength of a given material is exceeded. Substantial investigations have been undertaken to quantify mechanical properties and swelling response of materials that constitute painted wooden objects (Mecklenburg, Erhardt & Tumosa 1998; Richard, Mecklenburg & Tumosa 1998). The yield points have been determined as 0.4% for paints and glues, and 0.25% for brittle gesso found in historic panel paintings. Much higher strains – 0.9% or greater – were found necessary to cause failure. Swelling responses of the panel materials were measured for a full range of RH and expressed as rates of dimensional response, that is, variations in strain per unit change in RH. While these rates are low in the range of 40–60% RH, the wood and glue show dramatic increases outside the range when compared to far less responsive gesso and paint. The mismatch in the response of gesso and wood, especially in the most responsive tangential direction of the wood, has been identified as the worst case condition: upon desiccation, the shrinkage of wood overrides that of the gesso which experiences compression, whereas upon wood swelling the gesso layer experiences tension. If the uncontrolled changes in the dimensional change go beyond a critical level, the gesso can crack or delaminate. The forenamed authors have assumed that the strains induced by the mismatch in the dimensional response should not exceed a yield point for gesso either in tension or compression. Therefore, allowable changes in RH were calculated as those inducing in the gesso layer a strain level of 0.25%. As the rates of dimensional changes depend strongly on RH, the allowable RH variations were presented as functions of starting RH levels; for example, the allowable increase of RH is 12% for a painting on a substrate of cottonwood (Populus sp.) equilibrated at 50% RH before tensile yielding in gesso occurs, while the allowable decrease in RH is 17% before the compression yielding.

The structural analysis of painted wood has allowed proposing environmental specifications for collections of historic objects based on a continuous control of RH to as constant level as possible. The optimum RH set-point is usually specified within the 50-60% range and the recommended amplitude of RH variations should be less than 15%. The National Trust Specifications for Conservation Climate Control (Staniforth et al. 1994) are a well known example:

Control to constant RH

set-point 58% (may need to be stepped down from a higher level when bringing the building under control)

Alarms

First level: RH below 50% or above 65%

Second level: RH below 40% or above 75%

The RH control was to be achieved by adjusting heat input or by dehumidification. RH was the priority control variable unless a temperature limit was reached: a lower temperature limit was specified of 5 °C to prevent water freezing, and upper temperature limits of 18 °C in winter and 22 °C in summer were specified to limit the energy consumption and visitor discomfort.

Other influential climate specifications for museums and archives are published by the American Society of Heating, Refrigerating and Air-Conditioning Engineers - ASHRAE (Museum 2007). Until recently, they recommended 21°C and 50%rh, with minimal fluctuations. Since 1999, a separate 'Museums Libraries and Archives' chapter has appeared which included a large table called 'Temperature and relative humidity specifications for museum, gallery, library, and archival collections'. The table recommended more relaxed set points: 50%rh (or historic annual average for permanent collection) and temperature set point between 15 - 25 °C. A wide range of permissible RH fluctuations was considered and, for each option, the table provided collection risks/benefits: the AA class of control, corresponding to the tight ideal specifications allowed for mere ±5%rh with no seasonal change but guaranteed no risk of mechanical damage to most artefacts and paintings. RH fluctuations of ±5% with the allowed seasonal change of ±10%rh or RH fluctuations of ±10% with no seasonal change (both class A) should produce small risk of mechanical damage to high-vulnerability artefacts, whereas ±10% with the allowed seasonal change of ±10%rh (class B) should produce moderate risk to high-vulnerability artefacts but tiny risk to most paintings.

The described specifications point clearly to an inherent difficulty in establishing universal specifications for vulnerable objects, like painted wood, based on the understanding of their mechanical behaviour. The reason has been awareness of the uncertainty of the properties of the individual material components, accumulated defects and the physical interaction between the layers which cannot be fully recognised for individual works of art. The investigations of the material properties are carried out usually on samples of new materials which not necessarily reflect material properties of the historic ones, moreover generally adapted over many decades to a particular indoor environment within which they have been preserved. Such adaptation might have involved an unknown level of permanent change, like ageing, deformation or fracturing, making historic materials different in terms of their vulnerability to damage processes.

The acclimatisation concept

The ASHRAE specification described above recognized that collections can acclimatize to an average annual RH other than the central value of 50%. Michalski (2009) coined the term of a 'proofed fluctuation' defined as the pattern of largest RH or temperature fluctuations to which the object has been exposed in the past. It was assumed that the risk of further mechanical damage (beyond that already accumulated in the past) from fluctuations which do not go beyond the proofed pattern is extremely low. The proofed fluctuation concept eliminates any need for elaborate mechanical response calculations and offers a risk assessment based just on past climate records. Traditionally the 'acclimatisation' concept was basis for the recommendations that the past climate conditions should be retained when vulnerable objects are moved from their usual location for restoration or exhibition. With the growing use of electronic monitoring systems, long-term surveys to understand RH and temperature levels and fluctuations have become easier and can be undertaken on a wider scale. The accumulated data can be mathematically processed to establish more quantitative target microclimates suitable for the preservation of vulnerable objects, like painted wood, by specifying average levels and variability of temperature and RH. These are statistically represented in terms of: average level over a selected period, e.g. one year, a seasonal cycle and short-term fluctuations. The procedure is illustrated by a case study of a historic church (Bratasz et al 2007).

Case study 5. Wooden church of Saint Michael Archangel, Debno, Poland. This building is unheated – a relatively open structure with an indoor climate strongly governed by the outdoor weather due to a high rate of air exchange between inside and outside

Figure 14 shows plots of indoor RH for the Debno church. The RH data (the jagged black line) were sampled every fifteen minutes for one year, beginning in June and ending in May the subsequent year. The yearly average, calculated as the arithmetic mean of the RH readings, is 71%. The seasonal cycle is obtained by calculating, for each reading, the central moving average (MA) which is the arithmetic mean of all the RH readings taken in a 30 day period composed of 15 days before and 15 after the time at which the average is computed. MA smoothes out the short–term fluctuations and highlights longer-term trends or cycles. For the case study investigated, an increase in winter up to 80% and a decrease in warm period down to 55% are observed.

The seasonal variability is quite considerable when it is compared with the museum standards for indoor climate stability, exemplified in this paper by the National Trust (NT) specifications. The upward deviation observed exceeds the allowable increase by 7%, specified as the first alarm level by NT. Moreover, in absolute terms, the maximum seasonal deviation of RH in the church studied attains the upper limit of 80% above which attention should be paid to the risk of mould growth. The downward deviations during the spells of dry weather in spring or summer come close to the decrease by 18%, specified as the second alarm level by NT.

The short-term RH fluctuations, superimposed on the seasonal variations, are shown in Figure 15. They were extracted from the raw RH data by subtracting MA from the instantaneous RH. A fluctuation is calculated relatively to MA, i.e. the seasonal cycle rather than the yearly average level, to take into account both the natural seasonal variability and the stress relaxation time constants of the materials and, thus, to reflect better the risk of their physical damage. In this way, the fluctuations calculated with reference to the seasonal cycle are correctly smaller, and in consequence less risky, than those calculated relatively to the yearly average. On the other hand, fluctuations that are counter to the seasonal cycle are correctly calculated as larger and more risky than relative to the yearly average.

The lower and upper limits of the tolerable band, marked as horizontal lines, correspond to the 7th and the 93d percentiles of the fluctuations recorded in the monitoring period, respectively. In the case the fluctuations follow a Gaussian distribution, these limits correspond to -1.5 and +1.5 Standard Deviation, respectively. In this way 14% of the largest, most risky fluctuations are excluded, the cuts being equally applied to peaks and drops in RH, yielding excessively moist or dry environments.

The 'acclimatisation' concept was also explicitly expressed in the standards on choice and control of the indoor environmental conditions favouring conservation of sensitive historic materials (Italian 2002, Draft 2008). The two approaches can be also combined so that maintaining the past microclimate in terms of levels, seasonal cycles and fluctuations of temperature and RH is recommended on one hand, but the 'absolute' allowable variations based on the mechanical behaviour of paintings are defined on the other – as a result very stable past microclimates do not lead to unnecessarily strict future targets for climate control.

It should be stressed at this point that the harmlessness of the pre-existing climatic conditions has been a key assumption in the approach. The assumption has to be carefully checked in each case as physical damage can be cumulative rather than catastrophic, therefore fluctuations, even if not exceeding the historic levels, can involve risk of damage. Furthermore, conservation treatments can erase safety margins of objects achieved by their acclimatisation to the historic conditions.

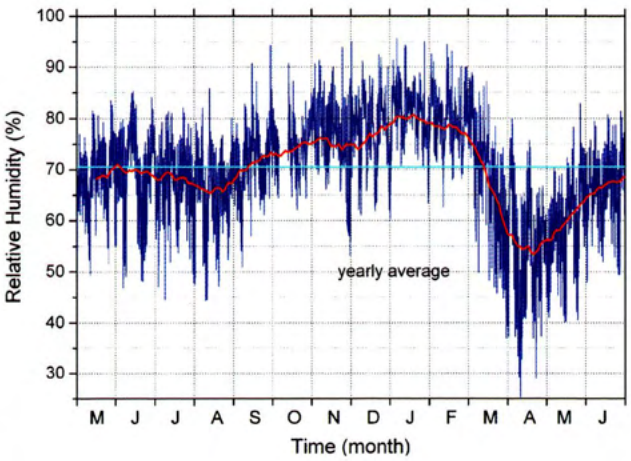

Fig. 14 - RH readings measured during one year (the jagged blue line) and a seasonal RH cycle (smooth red line) obtained by calculating the 30-day central moving average of the readings. The yearly RH average is marked by a horizontal line.

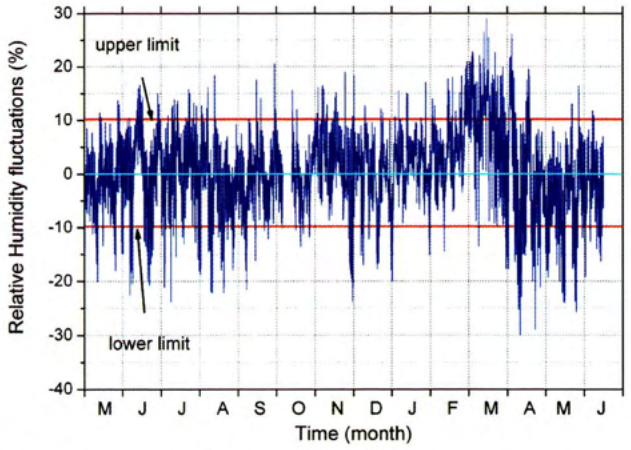

Fig. 15 - Target range of RH fluctuations; the lower and upper limits of the range are calculated as the 7th and the 93d percentiles of the fluctuation magnitudes, respectively.

Direct tracing of climate-induced damage

The limitations of the described two approaches to predicting risk of physical damage of a concrete historic object in its specific environment have led to search for scientific methods of direct tracing climate-induced damage: non-invasive, simple, economic and capable of operating in the real-world conditions in museums or historic buildings. The idea is to record an observable related to damage (damage indicator) at an object in a continuous way or at a specified time interval rather than to monitor the environment which affects the object.

Acoustic emission (AE) for tracing fracturing in wood

AE monitoring has been particularly successful in direct tracing of the fracturing intensity in wooden cultural objects exposed to variations in temperature and RH. Acoustic emission is defined as energy released due to micro-displacements in a structure undergoing deformation. The energy passes through the material as ultrasound and sound waves, and is typically detected at the surface using a piezoelectric transducer which converts the surface vibration to an electrical signal. As mechanical failure of many

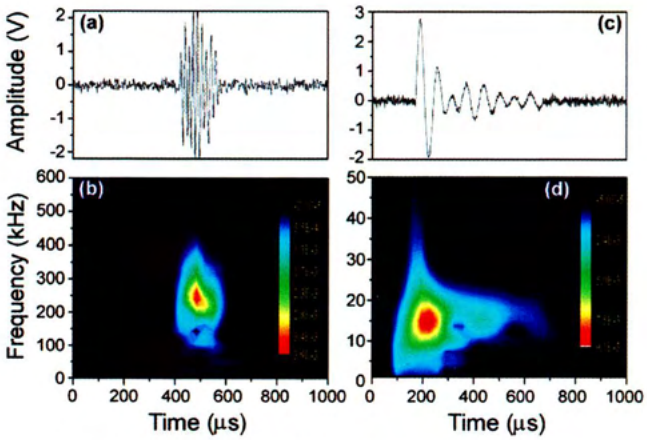

Fig. 16 - Raw acoustic emission signals acquired in a massive wooden cylinder during the large variation of: (a) relative humidity; (c) temperature. The predominant characteristic frequencies in the two signals are shown as red regions in (b) and (d) respectively.

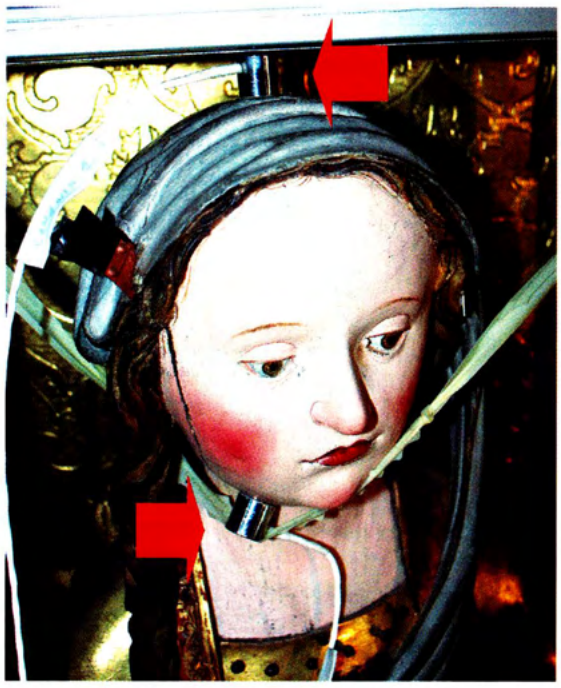

Fig. 17 - Details of two AE sensors coupled to a statue's head in the church of Santa Maria Maddalena in Rocca Pietore, Italy

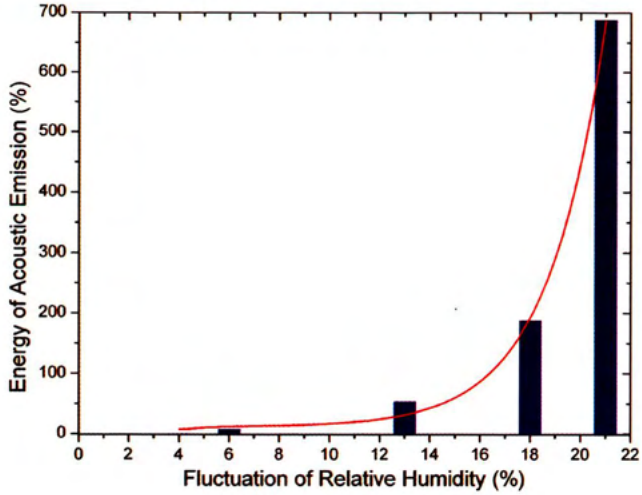

Fig. 18 - Total energy of AE events recorded on the wooden altarpiece as a function of the amplitude of the fluctuation in relative humidity.

materials is preceded by a discernible level of AE activity, AE monitoring has become an important non-destructive tool in material science and engineering, capable of predicting macro-damage and tracing crack propagation accurately in space and time due to digital capture and real-time processing of individual AE events.

Despite its unique characteristics and popularity, only recently was the AE monitoring applied to tracing directly climate-induced stress in wooden historic artefacts (Jakieła 2007b, 2008). An important feature making the AE monitoring a very promising non-destructive tool for indicating the risk of physical damage to wooden objects is a frequency signature in the AE signals related to damage. The AE events recorded in massive wooden cylinders, imitating sculptures, subjected to damaging climate-induced stress could be divided into two time-frequency bands: one of low frequency between 5-30 kHz and long duration between 500-2000 s, and the other of high frequency between 60-300 kHz and short duration between 20-450 s. Raw signals of such events together with their wavelet transforms into the time-frequency domain are illustrated in Figure 16.

Only low-frequency events were observed when specimens were subjected to non-damaging change in temperature alone or when two pieces of wood were rubbed one against another. In contrast, high-frequency events were related to damage of wooden samples when subjected to critical stress. The association of the short, high-frequency events with fracturing of the wood structure made possible their extraction from raw AE signals and filtering off all other effects, including the ambient noise.

The frequency pattern of AE signals for the new and ancient wood proved similar, confirming the applicability of the AE technique for tracing the fracturing intensity in wooden cultural objects exposed to variations in temperature and relative humidity (RH). The on-site monitoring of the AE activity of the wooden mediaeval altarpiece in the church of Santa Maria Maddalena in Rocca Pietore, Italy (Fig. 17), complemented the extensive programme of monitoring the climatic parameters and the dimensional response of the wooden sculptures (sease study 1 above).

The results indicated that the acoustic signals from routine church activities, such as the organ music or the sounds of bells, can be filtered out. The accumulated energy of high-frequency AE components related to fracturing of wood depended on the magnitude and rate of RH variations. The AE activity became negligible below the allowable magnitude for rapid RH variation established by numerical modelling, or when the time interval allowed for the RH variation was long enough. This is illustrated in Figure 18, which shows the dependence of the total energy of the AE signals on the amplitude of the variation in relative humidity (RH) produced during heating episodes in the church. It is evident that the threshold in the magnitude of the RH variations above which AE activity appears is not a discrete value. The increasing energy of events is recorded at increased loading as fracturing occurs in locally weakened areas in wood. A visible, macroscopic failure of wood is therefore preceded by the progressive evolution of damage at the micro-level which can be traced accurately by the AE analysis.

Portable digital speckle pattern interferometry (DSPI) for surveying surfaces of painted wood

DSPI makes use of speckles – a granular pattern of light and dark – produced whenever a rough surface is illuminated by a laser light. Speckles are a result of interference between the illuminating and the reflected light. Like a fingerprint, the speckle image is inherent to the investigated surface. Superimposing a

reference light, which is split out of the same laser source, on the light reflected from the surface results in the creation of an interferogram. When the object's surface is deformed - mechanically or thermally - the speckle interferogram also changes. Subtraction of interferograms recorded for the same surface before and after loading results in a fringe pattern, which reveals the displacement of the surface during loading as contour lines of the deformation. The technique is non-contact and gives full-field information on almost any surface and any material. When a high resolution CCD camera is used for recording the speckle pattern, mapping surface displacements to a fraction of a micrometer is possible.

The potential of DSPI technique to record profiles of polychrome wood, especially to trace non-recoverable alterations or damages to the painted surface like discontinuities, hair-line fractures, deformed or delaminated areas has been recognised early. The state-of-the-art and the evolution of DSPI, various features of the technique as well as examples of the applications in diagnosing panel paintings are thoroughly described in a review paper by Ambrosini & Paoletti (2004) whereas Bernicola et al. (2009) report on recent use of the technique in monitoring dimensional changes in wooden panel paintings due to climate-induced alterations.

A case study is described to illustrate use of DSPI for the condition surveys of painted wood on site in a historic church (Lasyk et al. in press).

Case study 6. The altarpiece in the church in Hedalen, Norway

A left bottom panel of a right wing of the altarpiece, marked in Figure 19, was selected for the analysis. The selected part of the panel was divided into 44 fields measuring 25 x 30 mm which were individually analysed by DSPI. Figure 20 shows the size of the field illuminated by a laser beam during the measurement. The relatively small size of the fields analysed, increased image quality and produced fringe patterns of good resolution in the range of millimetres. The images optimally revealing various features of the painted surface were combined digitally, so that they could be compared with photographs of the panel.

DSPI has revealed irregularities in the painted surfaces as lost paint, discontinuities in the surface between retouching and the original paint, cuppings, or irregularities caused by the painting technique. In this way, the technique has helped a painting conservator to identify the flaws that should be treated or monitored. A particularly helpful feature of the technique has been revealing areas of paint damage at an incipient stage which were so minor that the paintings conservator hardly would have seen them without the guidance of DSPI as illustrated in Figure 21.

In the most refined version of DSPI, sound is used to induce vibration of the surface. The approach is particularly suitable for identifying delaminated areas of paint or gesso. It provides precise information on the characteristic vibration frequency of the delaminated area, its size and the spatial distribution of the vibration amplitude as illustrated in Figure 22.

Acknowledgements

The research on the use of portable digital speckle pattern interferometry for surveying surfaces of painted wood was carried out within the project supported by Financial Mechanism of the European Economic Area, grant PL0086 'Establishing standards for allowable microclimatic variations for polychrome wood'.

Many research results contained in this chapter were presented and discussed during the meetings of COST Action D42. The author thanks the Action for its generous support.

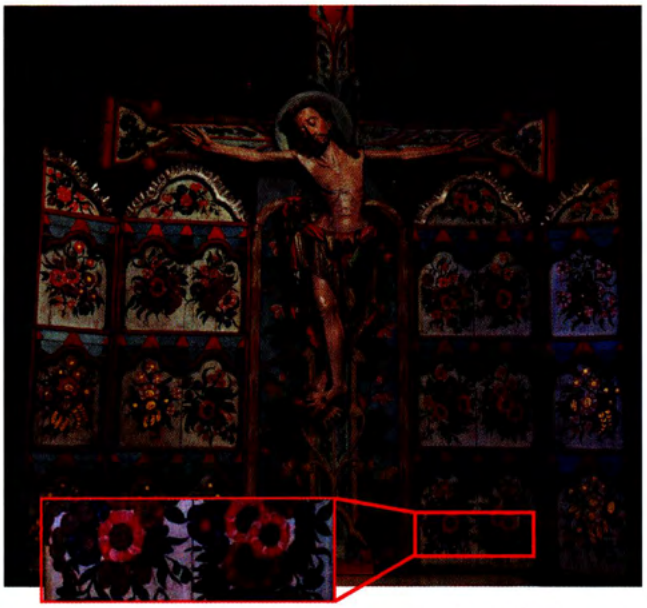

Fig. 19 - Main altar of the church in Hedalen, Norway. The conditions surveys were performed for the painted wooden panel marked.

Fig. 20 - DSPI system on the altar during the measurements - the field illuminated by a laser beam is visible.

Fig. 21 - Irregular or problematic areas on the painting as detected by a traditional conservation survey and by DSPI, marked by rectangles and ovals, respectively.

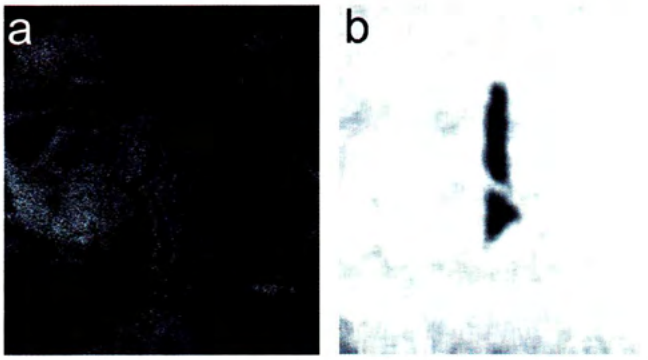

Fig. 22 - A detachment of the paint layer at an incipient stage: (a) a contour of the detachment revealed by a cluster of distorted fringes; (b) the analysis of sound-induced vibration of the surface.

References

Almeida, G., Gagné, S. &, Hernández, R.E., 2007. A NMR study of water distribution in hardwoods at several equilibrium moisture contents. Wood Science and Technology, 41: 293-307.

Ambrosini, D. & Paoletti, D., 2004. Holographic and speckle methods for the analysis of panel paintings. Developments since the early 1970s. Reviews in Conservation 5: 38-48.

Bernikola, E., Nevin, A. & Tornari, V., 2009. Rapid initial dimensional changes in wooden panel paintings due to simulated climate-induced alterations monitored by digital coherent out-of-plane interferometry. Applied Physics A 95: 387-399.

Bratasz, Ł. & Kozłowski, R., 2005. Laser sensors for continuous in-situ monitoring of the dimensional response of wooden objects. Studies in Conservation, 50: 307-315.

Bratasz, Ł., Kozłowski, R. & Camuffo, D., 2007. Target microclimate for preservation derived from past indoor conditions, In T. Padfield & K. Borchersen (eds) Contributions to the Museum Microclimates Conference, The National Museum of Denmark, Copenhagen, 129-134.

Bratasz, Ł., Kozłowski, R., Camuffo, D. & Pagan, E., 2007. Impact of indoor heating on painted wood. Monitoring the altarpiece in the church of Santa Maria Maddalena in Rocca Pietore, Italy. Studies in Conservation, 52: 199-210.

Bratasz, Ł., Kozłowski, R., Kozłowska, A. & Rivers, S., 2008. Conservation of the Mazarin Chest: structural response of Japanese lacquer to variations in relative humidity, In J. Bridgland (ed.), ICOM Committee for Conservation, 15th Triennial Conference, New Delhi: Preprints, 933-940.

Camuffo, D., Pagan, E., Rissanen, S., Bratasz, L., Kozlowski, R., Camuffo, M. & della Valle, A., 2009. An advanced church heating system favourable to artworks: A contribution to European standardisation. Journal of Cultural Heritage, doi:10.1016/j.culher.2009.02.008.

Draft European Standard CEN/TC346 prEN 15757, 2008. Conservation of Cultural Property - Specifications for temperature and relative humidity to limit climate-induced damage in organic hygroscopic materials, European Committee for Standardisation, Brussels.

Fratzl, P., 2003. Cellulose and collagen: from fibre to tissues. Current Opinion in Colloid and Interface Science, 8: 32–39.

Hailwood, A.J., Horrobin, S., 1946. Absorption of water by polymers. Analysis in terms of a simple model. Transactions of the Faraday Society, 42B: 84-102.

Italian Standard UNI 10969, 2002. Cultural Heritage - Environmental conditions for conservation. General principles for the choice and the control of the indoor environmental parameters. Part 1 Microclimate, Ente Nazionale Italiano di Unificazione.

Jakieła, S., Bratasz, Ł. & Kozłowski, R., 2007. Acoustic emission for tracing the evolution of damage in wooden objects. Studies in Conservation 52: 101-109.

Jakieła, S., Bratasz, Ł. & Kozłowski, R., 2008a. Numerical modelling of moisture movement and related stress field in lime wood subjected to changing climate conditions. Wood Science and Technology, 42: 21-37.

Jakieła, S., Bratasz, Ł. & Kozłowski, R., 2008b. Acoustic emission for tracing fracture intensity in lime wood due to climatic variations. Wood Science and Technology 42: 269-279.

Lasyk, Ł., Łukomski, M., Olstad, T.M. & Haugen, A., 2010. Electronic speckle pattern interferometry (ESPI) for the condition surveys of painted wood: monitoring the altarpiece in the church in Hedalen, Norway. Studies in Conservation submitted for publication.

Mecklenburg, M.F., Tumosa, C.S. & Erhardt, D., 1998. Structural response of painted wood surfaces to changes in ambient relative humidity. In V. Dorge & F.C. Howlett (eds.) Painted Wood: History and Conservation, The Getty Conservation Institute, Los Angeles, 464-483.

Michalski, S., 2009. The ideal climate, risk management, the ASHRAE chapter, proofed fluctuations, and towards a full risk analysis model, In F. Boersma (ed.) Proceedings of Experts' Roundtable on Sustainable Climate Management Strategies, Tenerife, 2007, Getty Conservation Institute, Los Angeles www.getty.edu/conservation/science/climate/climate_expertsroundtable.html (accessed 15 December 2009)

Museums, Galleries, Archives and Libraries, Chapter 21, 2007. In ASHRAE Handbook – HVAC applications, ASHRAE - American Society of Heating, Refrigerating, and Air-Conditioning Engineers Inc.

Richard, M., Mecklenburg, M. F. & Tumosa, C.S., 1998. Technical considerations for the transport of panel paintings. In K. Dardes & Andrea Rothe (eds) The structural conservation of panel paintings: proceedings of a symposium at the J. Paul Getty Museum, April 1995, The Getty Conservation Institute, Los Angeles, 525-556.

Sing, K.S.W., Everett, D.H., Haul, R.A.W., Moscou, L., Pierotti, R.A., Rouquerol, J. & Siemieniewska, T., 1985. Reporting physisorption data for gas solid systems with special reference to the determination of surface area and porosity. Pure and Applied Chemistry, 57: 603-619.

Staniforth, S., Hayes, B. & Bullock, L., 1994. Appropriate technologies for relative humidity control for museum collections housed in historic buildings. In A. Roy & P. Smith (eds) Preventive Conservation Practice, Theory and Research, International Institute of Conservation, London, 123-128.

Timmermann, E.O., 2003. Multilayer sorption parameters: BET or GAB values? Colloids and Surfaces A: Engineering Aspects 220: 235-260.

Chapter 5
The role of light

Mauro Bacci, Costanza Cucci

Light and light sources

Light is a fundamental ingredient for enabling vision in human beings and is thus a key element of museum environments and expositive spaces. Nowadays, lighting is indeed considered to be a fundamental issue in designing sites intended for hosting exhibitions of artworks. It is common knowledge, in fact, that only proper illumination can guarantee the total visual satisfaction needed in order to fully appreciate and comprehend artworks on display (Shaw 2002, Wilson Kesner 2008). On the other hand, the terms *'light'* and *'lighting'* in the conservation field are often linked more frequently to topics such as deterioration, damage and environmental risk factors, and are therefore treated as matters of preventive conservation and risk assessment (Thomson 1994, Cuttle 1996, Cuttle 2007).

In fact, this twofold connotation of the concept of light cannot actually be simplified, since both the aspects mentioned are closely correlated and are often in contrast to each other (Fig. 1). However, some basics of the interaction between light and matter could be better elucidated by a brief review, without any claim to being exhaustive. We would do best to start from the definition of the word 'light', which in itself is not totally unambiguous.

In general, the term *light* is used to indicate the electromagnetic radiation within the so-called *visible* region, that is, the narrow spectral interval spanning from 380nm to 780nm. This, in fact, is the range of wavelengths that can be perceived by the photoreceptors of the human retina, which has a variable degree of sensitivity to the different colours of the spectrum, with a maximum response at 555nm in daylight conditions. The average response of the human eye to light is described by tabulated functions that report luminous eye efficiency versus the wavelength. These functions have been extrapolated on the basis of statistical studies aimed at reconstructing the visual perception of a 'Standard Observer', and are adopted by the CIE (Commission Internationale de l'Éclairage) as standard references for describing all phenomena related to human vision. Since eye sensitivity also depends on the average lighting conditions, different sensitivity curves are available: namely, the *photopic curve*, which is suitable for describing human vision under well-lit conditions, and the *scotopic* curve, which refers to (human) eye response under low illumination levels (below luminance values of about 10^{-2} cd/m^2) These few basics can help to understand how greatly environmental lighting affects the visual perception of objects on display. Consequently, by changing the illumination levels as well as by varying the spectral content of the overall lighting, different results can be achieved in terms of visual comfort and colour rendering (Judd &Wyszecki 1975, Wade & Swanston 1991).

Back to the definition of the term 'light', another point of view, that is not anthropocentric but focuses instead on the characteristics of the light source, is also often considered. In many contexts, the term *light* is actually used in a broader sense to indicate all of the radiation emitted by an optical source. It thus ranges from the Ultraviolet (UV) to the Infrared (IR) part of the electromagnetic spectrum, and includes the wavelengths from about 100nm to about 1 mm.

No matter what the extension of the spectral interval assumed is, the fundamental concept is that light is electromagnetic radiation, and can therefore be represented in terms of propagating waves, which are *univocally identified* by their wavelength (λ), or, alternatively, by their frequency (ν). The product $\lambda \cdot \nu$ provides the speed of the wave in the medium passed through, and if the wave travels in the vacuum, it corresponds to the universal constant $\lambda_0 \cdot \nu_0 = c = 2.9979 \cdot 10^8$ m·s^{-1}. As is well known, propagating waves transport energy, which can be transferred or exchanged with the matter on which the light

Fig. 1 - Lighting in museums: example of opposite needs, protection of the work of art and visitor visual satisfaction, in the Uffizi Gallery, Pollaiolo's room, Florence.

impinges. The latter is a fundamental aspect in explaining the role of light as a deterioration factor.

This concept can be better emphasized by considering the *corpuscular model* for describing light and the phenomena related to its interaction with materials. Within this scheme, the electromagnetic radiations are represented in terms of *quanta* (discrete packets) of energy called *photons*. Each photon behaves like a particle with energy $E=h\nu = hc/\lambda$, (h = 6.626×10^{-34} J s is the Plank constant), and ν is the frequency of the corresponding electromagnetic wave. According to this picture, a beam of light is thus represented as a flux of photons that can interact with other particles, for example with atoms or molecules constituting the materials, by means of different scattering processes. These interactions occur with energy exchanges that can cause structural alterations in the molecules involved, thus inducing alterations in the materials, as we shall elucidate in greater detail here as follows.

The twofold (i.e. wave-like and corpuscular) behaviour of light is known as *wave-corpuscle dualism*, and was proven by several fundamental experiments conducted in the early 20th century. These remarkable studies are considered to be milestones in the history of physics and to have laid the foundations of modern quantum theory. It is, of course, beyond the scope of the present discussion to illustrate the numerous implications of the dual nature of light. However, this concept has been mentioned here in order to emphasise the fact that, no matter what scheme is adopted, *light is a form of energy*. When a light beam strikes a given material, a certain amount of energy is transferred to it and several different phenomena may occur, depending on both the energy of the light beam and on the kind of molecules involved. This point is crucial in elucidating why light is often considered to be a dangerous deteriorative agent for artefacts and is, therefore, regarded as a factor to be minimised and controlled, rather than exploited, in order to improve the public's enjoyment of the object.

As can be easily concluded by considering the Plank relation $E= h\nu$ the higher the frequency (or, analogously, the shorter the wavelength), the higher the energy of the photon (and hence its capability to induce chemical reactions) is. This fact roughly explains why UV radiation is generally more effective in causing deterioration in materials, and why it is considered to be very hazardous for artefacts.

All the above considerations highlight the importance of a careful choice of appropriate light sources and illumination levels in museum environment, where lighting design should be result in the best compromise between the safety of objects on display and the visual comfort of the observer.

Generally speaking, a light source is the physical emitter of the electromagnetic radiation needed to initiate visual perception. It is univocally characterised by the Spectral Power Distribution (SPD) curve, which reports the relative radiant power versus the wavelength. The shape of SPD contains information on the overall appearance of the light emitted, and hence on the colour-rendering ability of the source, as well as on the content of damaging radiation (e.g. UV) for the materials exposed to it. A variety of light sources are available to meet the different expositive requirements, and the spectral characteristics of emitted radiation should be always taken into account in lighting design (Schaeffer 2001).

Another way of classifying light sources is based on their *colour temperature*, which descends from the blackbody concept. A blackbody is an ideal object that absorbs all electromagnetic radiation impinging on it, without any transmission or reflection. A blackbody thus appears to be black when it is perfectly cold (0 K), but it emits radiation (called thermal radiation) when heated to a certain temperature. It can be observed from daily experience that any heated material emits radiation, and that, if the temperature can be sufficiently raised without destroying the material, it will glow-spanning from red to white as it become hotter. As far as the ideal blackbody, as described by the Plank theory, is concerned, it can be demonstrated that the distribution of thermal radiation emitted, known as the blackbody spectrum, is a bell-shaped continuous curve, and depends only on the temperature reached. The latter is the colour temperature: its value is expressed in Kelvin and it always corresponds to a given spectral distribution of the thermal radiation. The higher the temperature, the shorter the wavelength of the peak and the higher the maximum energy emitted. Hence, with an increase in the temperature, the light shifts from reddish towards bluish colours. It should be noted that this can be somewhat confusing with respect to ordinary terminology, which usually indicates red-yellow colours as being "warm" and those tending towards blue tones as being "cold".

For real light sources, colour temperature is defined as the temperature value of a black-body radiator at which the colours of the light source and the black body are identical. It must be stressed here that, whereas incandescent lamps emit a continuous spectrum of wavelengths and behave similarly to the ideal black-body, other sources (e.g. fluorescent and discharge lamps) have more articulated spectral distributions, where lines and peaks are present whereas some spectral bands are omitted. In these cases, recourse to the blackbody model would be not completely correct and the *correlated colour temperature (CCT)* is thus introduced for characterising lamps with a discontinuous spectrum.

As discussed above, light sources can have different physical features resulting in very different lighting effects. In order to provide *standardised models* of the different types of light sources, the CIE has established various kinds of *Illuminants* by providing tabulated values of their SPD curves. Illuminants are, therefore, numerical data sets that represent ideal, but not necessarily existing, light sources. Illuminants are used, for example, in colorimetric calculations for envisaging the colour appearance of an object if it is illuminated with a given light source. For example, the CIE illuminants of series D represent natural illumination, with D65 being the most commonly-used source for simulating daylight to a mid-day sun (colour temperature: 6500 K). Instead, illuminant A represents the typical behaviour of a tungsten light source, whereas illuminants of the F series are used to simulate fluorescent lights (Fairchild, 2005).

The question of colour rendering is obviously fundamental when dealing with lighting design in museums (Thomson 1961, Scuello et al. 2004, Pinto et al. 2008). Therefore, related concepts deserve a certain attention in the present discussion.

The Colour Rendering Index (CRI) is a specifically defined parameter for quantifying the capability of real light sources to faithfully reproduce the same colour effects of an ideal source that is represented by a standard illuminant (e.g. D65 simulating daylight). The CRI value, which is indicated with the symbol R_a, varies from 0 (very poor rendering) to 100 (high-performance rendering). For lamps used in a museum context, R_a should normally range from 80 to 100.

The most common light sources for the museum environment will be reviewed, by way of a conclusion to this section.

The first case to be considered is natural *daylight*. In principle, it is the most comfortable for vision and, thanks to the uni-

form energy distribution over the whole visible range, offers the best conditions for chromatic perception. However, it is well-known that daylight contains a considerable portion of UV as well as of IR radiation, both of which are highly damaging to artefacts on display. Direct sunlight illumination is strongly discouraged, therefore, by all conservation guidelines. Moreover, daylight varies with weather conditions, hour of the day, and season of the year, and its colour temperature can accordingly cover a very wide range: from 2500 K (the sun at sunset) up to 20,000K (clear blue sky). This strong variability makes it difficult to control the overall lighting conditions and to fulfil recommendations for illumination of highly sensitive artefacts (see below). For these reasons, there are two different approaches regarding the use of daylight in museums: the first one requires the use of an artificial lighting system that will ensure stable and constant illumination conditions; the second asserts that natural lighting can be profitably exploited in museum environments by means of adequate architectural and design solutions, possibly as a complement to artificial illumination, without risks for the artefacts and with a great improvement in public comfort and spatial quality. In particular, the use of neutral density filters on the glass windows, as well as resorting to appropriate arrangements of the exhibits by exploiting the structural and architectural features of the room or building, are common strategies adopted for the use of natural lighting (Shaw 1995, Oliveira & Correia Guedes 2006, Oliveira & Steemers 2008, Ezrati 2008).

As far as artificial lighting is concerned, the source typologies most commonly used in museums are tungsten-halogen bulbs and fluorescent bulbs. In addition, in recent years the use of white LED has also become relatively frequent in exhibition spaces, especially in conjunction with optical-fibre-based illumination systems.

Tungsten bulbs belong to the class of incandescent sources and are characterised by a continuous spectrum of radiation, which is emitted by the heated tungsten filament within the glass bulb. Typically these lamps have a colour temperature in the 2600-3000K range. Most of the radiation emitted by tungsten lamps is in the near infrared region, and only a small fraction of radiation falls within the ultraviolet region (about 2% of the visible radiation). In tungsten-halogen lamps, the light bulb is filled with a halogen gas in order to increase the life time of the lamp. In these cases the UV portion of emitted radiation may be greater than in models without the halogen gas; however, the availability of UV-stop bulbs manufactured with treated glass so as to filter-out the UV radiation, makes possible the widespread use of this type of lamp in the conservation field. The presence of halogen gas also affects the colour temperature, which is about 3000-3400K. The CRI is excellent (around 100).

Fluorescent lamps are based on the discharge of electricity between a pair of electrodes in a tube filled with low-pressure vapour or gas (typically Hg). The tube is coated with phosphors, which re-emit radiation in the visible region when they are impinged by the UV radiation generated by the electrical discharge. The spectral output is a continuous spectrum over which the typical emission lines of the discharge gas (e.g. mercury) are superimposed. Models designed for low – UV emission exist also for these types of lamps. The colour temperature and the colour rendering index of fluorescent lamps depend on the blend of phosphors. As a rule, the colour temperature ranges from 3000 K up to about 6500 K and more (7500 K, 8800 K and 15000 K), while CRI is in the 80 - 98 range.

Light-emitting diodes (LEDs) consist of semiconductors that emit light when electron-hole recombination occurs after the application of a suitable voltage. LEDs usually emit quasi-monochromatic coloured light, the colour of which can be altered by changing the composition of the semiconductor, and, hence, its band-gap. However, it is possible to obtain white light if the surface of a UV or blue LED is coated with fluorescent materials that emit in the green or red region so as to compensate the original emitted colour and to produce a clean white light. The colour temperature of LEDs can vary between 4000 K and 12000 K, with the colour temperature of common white LEDs being about 6500 K. Cool white LEDs have a CRI of around 70, while warm white LEDs have a CRI that is above 85. At present, LEDs are more and more widely used for lighting in museums and galleries, for energy saving reasons and thanks to the ease in coupling them to optical fibres. In fact, the use of optical fibres has many advantages: a) they have little aesthetic impact on the objects on display; b) their flexibility makes it possible to illuminate objects of limited accessibility (in showcases, for instance); c) optical fibres can lessen UV and IR radiation and reduce heating, and the light source, whichever it may be, can be placed far from the objects; d) external maintenance of the light source adds security to the system, because frequent opening of the display cases is thus avoided.

2. How to measure radiation

There are two possible approaches for introducing the topic of light measurements that overlap only partially, because they have been developed within different contexts and for different purposes (McCluney & McCluney, 1994). The first is the *radiometric approach*, which assumes as its starting point the fact that light is a form of radiant energy, and can therefore be measured in terms of radiating power. In the radiometric approach, the fundamental physical quantity is the *radiant flux*, which is the radiant energy transported per time unit by the electromagnetic wave into a region of space. The radiant flux is measured in Watt, W. Another quantity of interest for our purposes is *irradiance*: it corresponds to the power per unit area, is measured in W/m^2, and is used to measure the rate at which light energy passes through a unit surface. Irradiance is, therefore, a flux density ($d\Phi/dS$). The radiant flux divided by the solid angle is measured in W/steradian, and is referred to as *intensity*.

The second approach is the *photometric approach*, which instead is focussed on the perceptive point of view and is thus involved in describing only the phenomena related to effects of the *visible portion (380-780 nm)* of the electromagnetic spectrum. Since it has been specifically developed for measure of light as perceived by the human eye, the photometric scheme is the one usually adopted in conservation and in applications to the visual arts. All the physical entities within the photometric framework take into account the sensitivity of the human eye, and therefore are all weighted by the photopic curve that describes the spectral response of the eye. Based on this, in the photometric scheme the fundamental quantity is the *luminous flux*, which is the radiant flux of visible radiation weighted by the human sensitivity curve. Luminous flux is measured in *lumen* (lm). A derived quantity is *illuminance*, which is the luminous flux through a surface that is measured in *lux (=lumen/m^2)*. A very important quantity for applications within the museum context is the *total luminous exposure*, which results from the product of the illuminance multiplied by the exposure time, and represents the cumulative light-dose received by an item. It is measured in lux.hours. In fact, the

photometric system is quite complex, and includes many other quantities that can be related to their corresponding entities in the radiometric system. The above definitions should be sufficient for a basic understanding of the instrumentation used for light monitoring in museums, galleries and archives, whereas the reader is referred to the numerous texts on the topic for a more in-depth review of photometric and radiometric schemes.

According to accepted guidelines and recommendations (reported in detail in Section 5) for museum lighting, the parameters to be monitored are illuminance levels (sometimes referred as light intensity), total luminous exposure, (that is the total light-dose), and the presence of UV radiation.

The instruments for measuring illuminance are luxmeters (or light-meters), which are sensitive to the electromagnetic radiation in the 380 - 780 nm range and provide a value, expressed in lux, of the intensity of visible light impinging on the sensor. Several types of devices for measuring indoor light levels are available. For all practical purposes, the measuring range should be from 0.1 to 10^6 lux. Many models of these commercially-available electronic devices are suitable for museum applications, since they are compact, aesthetically non invasive and are easy to use. The most recent models are data-loggers, which are compact devices capable of simultaneously measuring various environmental parameters, such as visible light, UV portion, temperature, and relative humidity. Data-loggers can be set to acquire data at the desired rate over prolonged periods of time (months or years), which makes it possible to fully characterize the average environmental conditions. As far as visible light measurement is concerned, in addition to the values of instantaneous illuminance, many devices can supply the total luminous exposure by integrating the data over time. Alternatively, the latter can be easily calculated once the sampling measurement rate is known.

Correct light measurement should take into account the fact that lighting conditions can be extremely variable in time and that they also depend on the particular position in a room. Thus, a mean illuminance value must be calculated, and different measurements in strategic positions in the room need to be performed. Hence, a light monitoring programme should include not only an assessment of the average daily illumination, but also a control of the peak illuminance values, since these may be much higher than the recommended levels for museum environments. When the lighting conditions of a specific object need to be monitored, the luxmeter should be placed as close as possible to the object, and the sensor surface must be oriented in the same direction as the object itself. Should the object not be flat, the sensor must be parallel to the principal plane of the object or to the more exposed surface. Measurements should be carried out on different points in the area of interest and then be averaged. Ultraviolet (UV) radiation is more energetic than visible light and, as already explained, can be very effective in initiating deterioration in materials. Therefore, the UV component must be eliminated, or at least be reduced to the lowest possible level. UV cannot be perceived by the human eye and its presence has to be tested by means of specific devices. Two ways exist in which to measure, or more precisely, to prove its presence. The first method involves the use of a UV-meter, which is a device that gives quantitative values in microwatts per square centimetre ($\mu W/cm^2$). Different manufacturers produce their own UV sensors. The result is a multitude of UV sensors, all showing different sensitivities curves. For practical uses in the conservation field, the correct choice is 365 nm at the maximum of the sensitivity curve. The second method is based on the use of a "UV monitor", which has been specially developed for museum use and measures the proportion of UV radiation present in the light. The measurements are expressed in microwatts per lumen ($\mu W/lm$), which is not included in the international system of measurements. As far as practical applications are concerned, to verify the presence of UV radiation using an UV-meter, two subsequent measurements are necessary. The first one is accomplished with a UV filter in front of the sensor, while the second is achieved without this. If the difference is significant (i.e. not less than 50%), the light source emits a significant portion of UV which and needs to be eliminated. Instead, with the special "UV monitor" for museum use, elimination of the UV source is necessary if the value exceeds $75\mu W/lm$ (value measured on incandescent light source). The elimination of UV for vulnerable objects and collections will have an effect only if the visible light is controlled at the same time.

Although a number of models of data-loggers and devices specifically designed for indoor light monitoring are commercially available at present, these tools require a certain care in handling data, and the extraction of effective information for practical purposes may thus not be straightforward. For these reasons, other strategies for implementing large-scale controls and reducing light damage risks are currently being encouraged in the conservation community. These alternative methodologies are based on the use of calibrated indicators that indicate cumulative light damage by means of a colorimetric variation. The Blue-Wool standard (BWS) card is a well-established light indicator that works on this principle. BWS includes a set of eight samples of blue-dyed wool that have different light-fastness levels that fade progressively – at different rates - under exposure to light. As a rule, each BWS card has a colour-fastness that is twice that of the previous one, since BWS n.1 is the most sensitive one and BWS n.8 is the most stable one. In practice, a BWS card set including all eight levels should be placed in the location to be controlled. One half of the BWS card should be covered and protected from light, so that it can be used as a reference. The card must be checked periodically, by registering the date of the check and comparing the exposed part of the card with the protected half. Since every BWS level corresponds to different light-sensitivity categories of materials (ISO classification; see also Table 2), the BWS samples that have undergone appreciable colour change will indicate that materials of the same category, exposed in the same place and for the same time, are subject to light damage of the same entity (Feller, R. L. & Johnston-Feller, R. M. 1979a, b, Bullock & Saunders 1999).

Another tool, which works on a principle similar to the BWS, has recently been developed within the framework of the EC funded Project 'Lido', and has subsequently been marketed as LightCheck® (Bacci, 2005, Dupont et al. 2008). This system includes two light indicators, LightCheck® Ultra (LCU) and LightCheck® Sensitive (LCS), which are designed for responding to very low and moderate light doses, respectively. LCU covers the [0-100 kluxh] range, whereas LCS covers the [100-400kluxh] range. These light-dose intervals cannot be monitored using BWS, because not even BWS n. 1 is sensitive enough to react appreciably. LightCheck® reacts to light by spanning different colour hues, and so a simple visual check makes possible a semi-quantitative evaluation of luminous exposure. Indeed, the dosimeter is provided with a reference colour scale in which the light doses corresponding to each

colour are reported (Fig.2). If used appropriately (i.e. no direct exposure to sunlight or to excessively high illuminances), LightCheck® provides a reliable semi-quantitative indication of the luminous exposure received after a given period. The correct use of the dosimeter is described in the data-sheet included in the kit.

3. Basics on photosensitive surfaces and deterioration mechanisms

The background needed in order to understand the causes of light-induced damage observed on art objects has its bases in *photochemistry*, the science that studies the chemical reactions induced in materials by light (Brill, 1980, Cuttle 2007).

When a photon of frequency ν impinges on a molecular system M, different phenomena can occur, depending on both the energy of the incident radiation and the molecular structure of the system invested. When the energy of the photon $E_p = h\nu = h\lambda/c$ is equal to the energy difference between two energy levels of the molecule, that is, if $h\nu = |E_{m1} - E_{m0}|$ the photon is absorbed by the molecule, the radiation energy is transferred to the molecular system, and a re-arrangement of the electronic structure takes place. Hence, in following the light absorption, the molecule changes its energy state and makes a transition to an excited level M^*, which is characterized by higher energy. The subsequent decay from an excited state to ground level can occur in several ways. The molecule can return to a lower energy level with a conversion to heat of the excess energy (*internal conversion*), or alternatively it can lose the excess energy by means of a radiative process, i.e. with the re-emission of light (fluorescence and phosphorescence). Another possibility is that it can transfer its energy to an adjacent molecule. In some cases, the excess energy may break bonds within the molecule. Regardless of which decay mechanism is favoured, it can be stated that when molecules are in an excited state (M^*), they are often more reactive than they are in their ground state. In fact, their electron distribution and polarizability are modified, and geometrical changes are also possible. Therefore, after the absorption of light, a certain number of molecules in the material become available to interact chemically with other species in the surroundings and to undergo a change in their chemical structure. Therefore, the probability that a photochemical reaction may occur depends not only on the chemical properties of the molecular system, but also on the surrounding chemical context. The presence of water, free oxygen, or other chemical species in the surrounding environment can facilitate the occurrence of reactions in the material when it is exposed to light. This fact points out the reason why preventing, or at least slowing down, photo-induced deterioration in art objects is recommended both to ensure adequate lighting conditions, and also to maintain all the environmental parameters (relative humidity, temperature, possible pollutants) within the advised levels (Feller, 1974).

It should be emphasised that photochemical changes in materials are only a possible consequence of light absorption, and that their occurrence is regulated by a certain probability, the entity of which depends on: a) the chemical properties of the material, b) the features (namely, intensity and spectral distribution) of the light interacting with the material; and c) the surrounding environmental conditions.

From a theoretical point of view, the probability of photochemical reactions occurring is closely related to the concept of *quantum yield*. This is the probability that a certain decay process, among the possible ones, will take place after a photon has been absorbed by the molecular system. To better comprehend the concept of quantum yield it has to be taken into account that when a light beam impinges on a given material a huge number of absorption processes occur, since every photon of the beam interacts with a single molecule. As stated above, after the light absorption the molecule can return to the ground state by following different paths (fluorescence, internal conversion, etc), and these processes are competitive with each other. Each process has, in fact, a certain probability of occurring, and this probability can be measured in terms of quantum yield. The latter can be calculated as the ratio between the number of events and the number of photons absorbed.

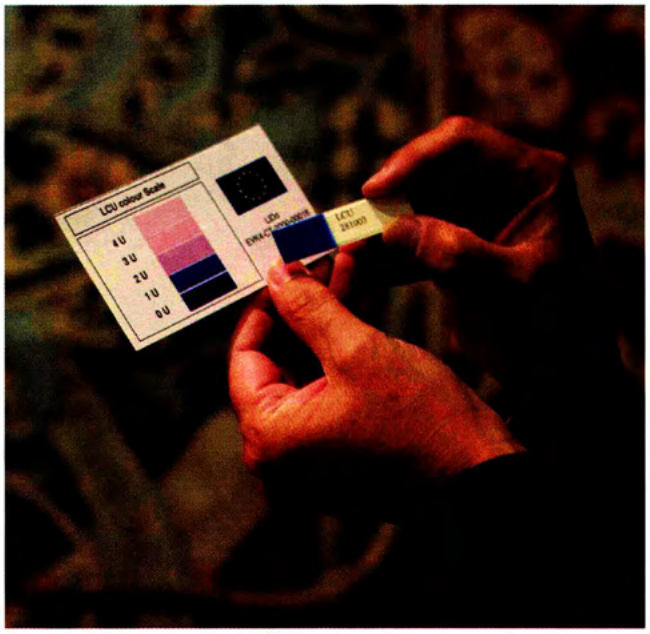

Fig. 2 - Example of use of LightCheck® Ultra: after exposure, the dosimeter has to be compared with the colour chart for assessing the light dose received (Image published in the paper: A.-L. Dupont, C. Cucci, C. Loisel, M. Bacci, B. Lavédrine, "Development of LightCheck® Ultra, a novel dosimeter for monitoring lighting conditions of highly photosensitive artefacts in museums" - Studies in Conservation - Vol. 53, n. 2, (2008) pp.49-72. Photograph by Maja Kardum, reproduced with permission of the Trustees of the V&A Museum, London.

The quantum yield is a number between 0 and 1, and the sum of all quantum yields for a particular absorption event should be 1. Although the quantum yield for pure substances can, in principle, be determined by means of accurate measurements, in practical applications this parameter is rarely available, especially for composite materials like those of interest in the conservation field. Thus, in most cases, empirical methods are adopted instead, in order to assess the risk of chemical changes. For example, artificial ageing tests under controlled lighting conditions are usually performed on samples made of materials of interest, so as to estimate their rate of fading or chemical change.

The tendency of a given material to undergo chemical changes upon being exposed to light greatly depends on the nature of its constitutive chemical species and on the reactivity of its excited states. At the macroscopic level, this tendency to undergo photochemical reactions is expressed in terms of *lightfastness* or, alternatively, the photosensitivity level of the material. Since different materials have differing stability to light, the measurements for a preventive conservation of cultural items have to be differentiated on the basis of the average photosensitivity of the object in question. This aspect will be detailed further in Section 5.

Fig. 3 - Raphael: 'The Miraculous Draught of Fishes', 1515-16. Body colour on paper mounted on canvas (tapestry cartoon), 302 x 309cm, on loan from H.M. The Queen. (Image courtesy of The Royal Collection (©) 2010 Her Majesty Queen Elizabeth II).

A practical criterion for establishing whether a material is susceptible to undergoing deterioration when exposed to a given light source is that of considering its *absorption spectrum*. As stated above, the absorption of radiation by a molecule occurs only when the radiation energy (hc/λ) matches the difference between two energy levels of the molecular (or atomic) system. Every material has a characteristic absorption spectrum that corresponds to the relative intensity of incident radiation absorbed by the said material at each wavelength (or frequency). Acknowledge of the typical absorption spectrum of a given material can help to establish the most suitable light sources to be used for illuminating it. Practically, the absorption spectrum of the material must be compared with the spectral distribution of the light used to illuminate it. For example, it is well known that organic materials have wide absorption bands in the UV spectral region, and that they are indeed highly prone to photo-deterioration induced by exposure to sources having a significant content of short wavelengths (200-400nm). Another aspect to be taken into consideration in establishing the risk of photo-deterioration is the role of the intensity of light. In fact, since the number of photoreactions occurring is proportional to the number of impinging photons, the stronger the light intensity emitted for every wavelength, the higher the risk of photo-deterioration. For these reasons, guidelines for the prevention of light-induced deterioration in objects on display in museums set limits on the average intensity of visible light admitted for every light-fastness category of the materials (see Section 5).

As for possible photochemical reactions, many different kinds of reactions can occur. Those of greatest interest as far as artistic materials are concerned are:

a) *photo-reduction*, which occurs for example when an excited molecule of dye reacts with molecules of the substrate (e.g. fibres or paper), acquiring from it electrons or hydrogen atoms that alter the molecular structure of dye;

b) Photo-oxidation, which occurs in many organic materials, such as polymers and cellulose compounds, as well as in dyes. This phenomenon, which is induced by light, results from the interaction of excited molecules of the materials with free oxygen or ozone present in the air, which give rise to chain reactions (often enhanced by the presence of water in the air) that cause the loss of original molecular species, the breaking of the polymeric chains, and cross-linking reactions;

c) *photo-fragmentation,* which results in rearrangement or fragmentation after an initial bond cleavage.

All these processes give rise to different, clearly visible and not reversible, alteration phenomena, such as fading (that is, loss of colour), chromatic variation (i.e. changes in the original colour), mechanical deteriorations (i.e. brittleness, fibre destruction, etc.), and the yellowing of colourless materials (Feller, 1964, Michalski 1987, Saunders, D. & Kirby J. 1994a,b, CIE 1990, Schaeffer 2001).

A meaningful and spectacular example of colour-balance alteration due to light action is provided in Figure 3, where a chromatic incongruence can be observed between Christ's mantle, which now looks white, and its red reflection on the water. The dye of the robe, probably of natural origin and different from that of its reflected image, has, therefore, completely faded, as also testified by the copy of the same scene realized on a Flemish tapestry, dated as being realized a few years later (1517-1519 ca.) and conserved in the "Pinacoteca" of the Vatican City Museum.

Another light effect, which is very often encountered, is the yellowing of the varnish layer that again modifies the overall colour balance, as shown in the example is reported in Figure 4. In such case the removal of the old varnish and a new varnishing can lead to a good recovery of the appearance of the painted layer.

4. Doses and effects

The above discussion has pointed out that all objects on display can be liable to a progressive deterioration caused by their interaction with light. Photochemical damage is cumulative and irreversible, and, therefore, the light-ageing of objects can be minimized but not completely eliminated. This fact can be summarized by a fundamental law, known as *reciprocity principle*, which states that *"the net photochemical effect is the result of the total light exposure which an object receives, which is the product of the light-intensity (I) multiplied by the exposure time (t)"*. According to this principle, an object would undergo the same damage if it were exposed either for 10 hours to a light intensity of 500 lux or for 100 hours at 50 lux. The physical concept behind the reciprocity principle is that changes induced in materials by radiation are determined by the total amount of energy absorbed, and do not depend on the rate at which the photons are absorbed. Different studies reported in the literature discuss the applicability of the reciprocity principle to conservations problems (Thomson 1994, Saunders, D., 1995, Saunders D. & Kirby, J. 1996, Chin et al. 2005). All these studies show that the reciprocity law holds for most cases of interest to museum applications, although under certain circumstances it might fail (for example, for certain materials, or at very high values of illuminance). It therefore makes sense to refer to the reciprocity principle in order to establish the main criteria for a lighting

Fig. 4 - Predella by Luca Signorelli, oil painted panel, early years of XVI century, Uffizi Gallery (Florence). Effect on the appearance of the yellowed varnish layer: The panel before (top image) the removal of the yellowed varnish layer, and after the restoration.

policy in museums and expositive spaces. Indeed, the guidelines for preventing risks of light damage to collections recommend that, among the other factors (see section 5), the total light-dose D=I·t received by an object is a fundamental parameter to be minimized. Based on this, preventive strategies often aim not only at controlling the intensity of illumination, but also at reducing the overall exposure time of artworks on display. For example, many important museums and institutions adopt the criterion of rotation for displaying the various items of a given collection.

Another aspect that should be pointed out regards photo-induced alterations and their dependence on the dose of light received. Indeed, as stated in the previous section, several photochemical reactions are caused by light, but are enhanced or accelerated by other factors, such as gaseous pollutants or chemical species present in the surrounding environment, or by particular thermo-hygrometric conditions. Consequently, the photochemical damage observed for an object cannot be attributed only to the exposure to light, but should be considered to be the result of a synergistic action between light and other environmental factors (e.g. temperature, relative humidity, pollutants). In the past decade, several scientific projects aimed at investigating a new approach to environmental monitoring in museums have been carried out, based on the use of passive sensors (Bacci et al. 2000, Odlyha et al. 2007). The latter are sensors, or indicators, made of sacrificial materials that mimic – in some way or other – the behaviour of real works of art and are able to record the overall environmental effects with a change of some physical property (e.g. colour). Within this framework the concept of *Equivalent Light Dose* (ELD) has also been introduced, with the aim of providing a semi-quantitative indication of the amount of the cumulative environmental damage (Bacci et al. 2003, Mignani et al 2003). By means of experiments both *in situ* and in the laboratory, it has indeed been shown that the colour change obtained by exposing samples of photosensitive materials (for example, the BWS) to a certain light dose in a controlled environment could be also achieved by exposing the same material to a smaller light dose but under highly variable environmental conditions (for example, with high fluctuations in relative humidity) (Bacci et al. 2004). The ELD can be defined as the dose of light capable of producing in a given material – exposed in an uncontrolled environment – the same spectral variation as that measured in the same material exposed in well-defined and controlled environmental conditions. The *ELD* is thus used to take into account the uncontrollable contributions of environmental factors that accelerate the light-induced reactions.

5. Lighting photosensitive cultural items

The most important museums and galleries usually adopt their own lighting policies, and these are aimed at synthesizing different aspects such as the typology of the collections to be exhibited, the architectonic features of the expositive spaces (e.g. in ancient or modern buildings), the comfort of the public, the economic resources available, and so forth. All these variables can have different types of significance, and may result in a variety of diverse practical solutions for environmental lighting in the said museums. Nevertheless, a few well-established criteria are extensively accepted, and have been adopted as a basis in implementing lighting strategies for all collections of cultural items (Thomson 1994, Michalski, S. 1990, Derbyshire & Ashley-Smith 1999, Ashley-Smith et. al. 2002, Tétreault 2003). These basic guidelines are listed here as follows:

a) Illumination levels should be as low as possible, consistent with the needs and visual comfort of the public.

b) Every material has a different response to light. Therefore, a categorisation, based on their sensitivity to light, of materials occurring in collections of artefacts is needed.

c) As a general rule, the total light dose received by objects should be minimised. This can be done both by lowering the intensity of the light and also by reducing the light-exposure time to the necessary minimum.

d) The UV component of light should be always avoided

To implement the above recommendations in a practical way, materials of interests in museums collections have been grouped into four different light-sensitivity classes (see Table 1). Correspondingly, limits on the illumination levels and on the total annual light exposure have been set for each class. The limits recommended by CIE are reported in Table 2 (CIE 2004). It should be pointed out that the most recent indications on the lighting of cultural items recommend the complete elimination of non visible radiation (that is, UV and IR), taking into account the fact that nowadays many technological solutions (filters, specific lighting components, optical fibres) are available for realizing these conditions. However, many other sources in the literature provide less strict indications, ones that permit the presence of minimal levels of UV. The maximum acceptable limit for UV is commonly set at 75 μW/lm, except for items belonging to the high-sensitivity class.

Table 1. *Four category classifications of materials according to sensitivity to visible light (from CIE 157:2004).*

Category of light sensitivity	Description
1. irresponsive	The object is entirely composed of materials that are **insensitive to light**. Examples; most metals, stone, most glass, genuine ceramic, enamel, most minerals
2. Low responsivity	The object includes durable materials that are **slightly responsive to light**. Examples; oil and tempera painting, fresco, undyed leather and wood, horn, bone, ivory, lacquer, some plastics
3. Medium responsivity	The object includes fugitive materials that are **moderately responsive to light**. Examples; most textiles, watercolours, pastels, prints and drawings, manuscripts, miniatures, paintings in distemper media, wallpaper, photographic documents, most natural history objects, including botanical specimens, fur and feathers
4. High responsivity	The object includes **highly light-responsive** materials. Examples; silk, colorants known to be highly fugitive, graphic art and photographic documents

Table 2. *Recommended limits for illumination and annual light exposure for the different categories of materials (from CIE 157:2004).*

Category of light sensitivity Corresponding BWS Category (ISO classification)	Illuminance limit (Lux)	Total annual light exposure (lux.hours per year)
1.irresponsive BWS cat. : none	No limit	No limit
2.Low responsivity BWS cat.: 7, 8	200	600000
3.Medium responsivity BWS cat.: 4, 5, 6	50	150000
4. High responsivity BWS cat. : 1,2,3	50	15000

6. Comments to CEN

From everything that has been said above, it is evident that the problem of lighting in places where art objects are exhibited requires particular care and skill on the part of those who are in charge of the general arrangement of the rooms dedicated to the exhibit. In fact, the challenge is to find an appropriate compromise between the needs of visitors and the appropriate conditions for safe conservation. Accordingly, many countries have produced their own recommendations for correct lighting inside museums. In 2004, the CIE drew up a Technical Report that was aimed at controlling damage to museum objects by means of optical radiation (CIE 2004). To date, this is the most complete document on the matter. However, European countries are still awaiting general guidelines and standards that would provide a unification of the procedures and management for the correct conservation of works of art. Indeed, within the European Community the European Committee for Standards (CEN) is recognized as the sole body authorized to develop and set technical standards (Havermans & Adriaens 2008; Fassina 2008). In 2004, the creation of a new European Technical Committee 346 for Cultural Heritage (CEN/TC 346) was launched, and was assigned the task of proposing standards for the best conservation practices regarding movable and immovable cultural artefacts. The structure of CEN/TC 346 consists of 5 Working Groups (WGs), which correspond to the various main areas for which technical development work has to be done. The fourth WG (WG4) *"Environment"* is responsible for the drafting of guidelines for the control of environmental variables and of standards for the measurement of indoor and outdoor conditions, as well as for the control of exhibition and storage premises, and for monitoring the interaction of artefacts with the environment. It is precisely within this context that WG4 is preparing a document concerning the exhibition-related lighting of cultural properties, which will be submitted to public enquiry in the next few months. The aim of this document is to define both the procedures and the means for implementing good lighting with regard to conservation policies, as well as to conditions of visibility and exhibition design. It intends to provide a tool for setting up a common European policy and a helpful guide for curators, conservators and project managers so as to provide architects and designers with a correct lighting program that has a European point of reference. Lastly, it should be noted that the document in preparation deals with lighting only for permanent and temporary exhibitions in museums and galleries, and does not consider lighting within other cultural heritage contexts. For instance, it does not deal with hypogeal sites, for which - to the best of our knowledge - few data are available in the literature, and much more research will have to be done in order to protect them from irreversible damage mainly due to biodeterioration.

Acknowledgements

The authors would like to thank the Uffizi Gallery in Florence for the collaboration and the kind hospitality offered in several research programmes devoted to the study of the museum environment.

The COST Action D42 is also gratefully acknowledged for supporting the present work in a fruitful context of research networking.

References

Ashley-Smith, J., Derbyshire, A.; Pretzel, B., 2002, The continuing development of a practical lighting policy for works of art on paper and other object types at the Victoria and Albert Museum In: Preprints of ICOM Committee for Conservation 13th Triennial Meeting, Rio de Janeiro, 22-27 September 2002, James & James Publisher, 1, 3-8.

Bacci, M., Cucci, C., Mencaglia, A. A., Porcinai, S., Mignani, A.G., 2003, Optical fibers for safer exhibition conditions in museums: the measurements of equivalent light dose. In: Renzo Salimbeni (Ed.) Optical Metrology for Arts and Multimedia. Proceedings of SPIE 2003, 5146, 170-174.

Bacci, M., Picollo, M., Porcinai, S. & Radicati, B., 2000, Evaluation of the museum environmental risk by means of tempera-painted dosimeters Thermochimica Acta ,65(1-2); 25-34.

Bacci, M., Cucci, C., Mencaglia, A. A., Mignani A. G. & Porcinai, S., 2004, Calibration and use of photosensitive materials for light monitoring in museum: the Blue Wool Standard 1 as a case study, Studies in Conservation, 49(2), 85-98.

Bacci, M., Cucci, C., Dupont, A.-L., Lavédrine, B., Loisel, C., Gerlach, S., Roemich, H., & Martin, G., 2005, "LightCheck®, new disposable indicators for monitoring lighting conditions in museums" In: Proceedings ICOM Comittee for Conservation 14th Triennal Meeting The Hague Preprints (2005), Vol. II, 569-573.

Brill, T. B., 1908, Decomposition of Organic materials by light In: Light Its interaction with art and antiquities, Plenum Press, New York, 173-194.

Bullock, L.& Saunders, D., 1999, Measurement of cumulative exposure using Blue Wool standards, In: ICOM Committee for Conservation Preprints triennial meeting (12th), Lyon, 29 August-3 September 1999: James & James, London, UK, Vol. 1, 21-26.

Cassar, M., 1995, The monitoring Programme in Environmental Management Guidelines for museums and galleries, Ed. Routledge, 62-74

CIE, 1990, On the deterioration of exhibited museum objects by optical radiation, CIE Technical Collection, CIE No. 89/3.

CIE, 2004, Control of damage to museum objects by optical radiation, CIE No 157: 2004.

Cuttle, C., 1996, Damage to museum objects due to light exposure Lighting Research and Technology, 28 (1), 1-9.

Cuttle, C., 2007, Light for art's sake: lighting for artworks and museum displays, Ed. Butterworth-Heinemann, Oxford UK.

Derbyshire, A. & Ashley-Smith, J., 1999, A proposed practical lighting policy for works of art on paper at the V&A In: Preprints of ICOM Committee for conservation, 12th triennial meeting, Lyon, 29 August-3 September 1999, James & James Publisher, 1, 38-41.

Dupont A.-L., Cucci, C., Loisel, C., Bacci M. & Lavédrine, B., 2008, Development of LightCheck® Ultra A novel dosimeter for monitoring lighting conditions of highly photosensitive artefacts in museums, Studies in Conservation, 53, 49 –72.

Ezrati, J., 2008, Daylighting for Museum, A Good Choice? In: Proceedings of the 9th Int. Conference on Non Destructive Testing (NDT) of Art 2008 Jerusalem, Israel, 25-30 May 2008.

Fassina, V., 2008, European Technical Committee 346 – Conservation of Cultural Property – Updating of the Activity After a Three Year Period In: Proceedings of the 9th Int. Conference on Non Destructive Testing (NDT) of Art 2008 Jerusalem, Israel, 25-30 May 2008.

Feller, R. L., 1964, Control of deteriorating effects of light upon museum objects. In: Paul Whitmore (ed.) Contributions to Conservation Science: A Collection of Robert Feller's Published Studies on Artists' Paints, Paper and Varnishes, Carnegie Mellon University Press, Pittsburg 2002, 53-87.

Feller, R.L., 1974, Speeding up photochemical deterioration. In: Contributions to Conservation Science: A Collection of Robert Feller's Published Studies on Artists' Paints, Paper and Varnishes, ed. P.M. Whitmore, Carnegie Mellon University Press, Pittsburgh 2002, 107–119.

Feller, R. L. & Johnston-Feller, R. M., 1979a, Use of the International Standards Organization's blue-wool standards for exposure to light. I. Use as an integrating light-monitor for illumination under museum conditions In: Paul Whitmore (ed.) Contributions to Conservation Science, Contributions to Conservation Science: A Collection of Robert Feller's Published Studies on Artists' Paints, Paper and Varnishes, Carnegie Mellons University Press, Pittsburgh 2002, 121-127.

Feller, R. L. & Johnston-Feller, R. M., 1979b, Use of the International Standards Organization's blue-wool standards for exposure to light. II. Instrumental measurement of fading, In: Paul Whitmore (ed.) Contributions to Conservation Science, Contributions to Conservation Science: A Collection of Robert Feller's Published Studies on Artists' Paints, Paper and Varnishes, Carnegie Mellons University Press, Pittsburgh 2002, 135-141.

Fairchild M.D., 2005, Colorimetry in Colour appearance models Second Edition John Wyley and Sons, West Sussex, England, 53-82.

Havermans, J. & Adriaens A., 2008, COST D42 - ENVIART: A European Network In Conservation Research In: Proceedings of the 9th Int. Conference on Non Destructive Testing (NDT) of Art 2008 Jerusalem, Israel, 25-30 May 2008.

Judd, D.B. & Wyszecki, 1975, Basic Facts in Color in Business, Science and Industry, Third Edition, John Wiley and Sons, New York USA, 5-91.

McCluney, WR. & McCluney R., 1994, Chapter 1 and Chapter 2 in:Introduction to Radiometry and Photometry, Ed. Artech House Publishers Norwood,MA, 1-59.

Michalski, S., 1987, Damage to museum objects by visible radiation (light) and ultraviolet radiation (uv) In: Lighting in museums, galleries and historic houses. Papers of the conference. Bristol, 9-10th April 1987, The Museums Association Publisher, London, 3-16.

Michalski, S., 1990, Towards specific lighting guidelines.In: Preprints of the 9th Triennial Meeting of the International Council of Museums Committee for Conservation. Dresden.583-588.

Mignani, A.G., Bacci, M., Mencaglia, A.A. & Senesi, F., 2003, Equivalent light dosimetry in museums with blue wool standards and optical fibers, IEEE Sensors Journal 3(1), 108 -114.

Odlyha, M., Theodorakopoulos, C., Thicket, D., Ryhl-Svendsen, M., Slater J.M. & Campana, R., 2007, Dosimeters for indoor microclimate monitoring for cultural heritage, in: Museum Microclimates In: Proceedings of the Conference Copenhagen 19-23 Nov. 2007, Padfield, T and Borchersen, K. (eds.) National Museum of Denmark Pulisher, 73-79.

Oliveira, F. & Correia Guedes M., 2006, Daylighting museums – a case study In Lisbon Proc. of PLEA2006 The 23rd Conference on Passive and Low Energy Architecture, Geneva, Switzerland, 6-8 September 2006, I 263-269 available on-line at:

http://www.unige.ch/cuepe/html/plea2006/Vol1/PLEA2006_PAPER130.pdf.

Oliveira, F. & Steemers, 2008, K. Daylighting Museums – a survey on the behaviour and satisfaction of visitors. In Paul Kenny, Vivienne Brophy & J. Owen Lewis Ed. Proceedings of PLEA 2008 – 25th Conference on Passive and Low Energy Architecture, Dublin, 22- 24 October 2008 Published by University College Dublin.– available on-line at: http://www.plea2008.org/.

Pinto, P. D., Linhares, J. M. M. & Cardoso Nascimento, S. M., 2008, Correlated color temperature preferred by observers for illumination of artistic paintings, Journal of the Optical Society of America A, 25(3), 623-630.

Saunders, D., 1995, Photographic Flash: Threat or Nuisance?, The National Gallery Technical Bulletin, Publisher: National Gallery, 16(1) 66-72.

Saunders, D. & Kirby J., 1994a, Wavelength-dependent fading of artists' pigments In: A. Roy, and P. Smith (eds.) Preventive Conservation: Practice, Theory and Research, London, International Institute for Conservation, 190–194.

Saunders, D. & Kirby, J., 1994b, Light-induced colour changes in red and yellow lake pigments, National Gallery Technical Bulletin, 15, 79–97.

Saunders, D. & Kirby, J., 1996, Light-induced damage: investigating the reciprocity principle In Proceedings ICOM committee for conservation, 11th triennial meeting in Edinburgh, Scotland, 1-6 September 1996, James & James Publisher, London, 87-90.

Chin, J., Nguyen, T., Byrd E. & Martin J., 2005, Validation of the reciprocity law for coating photodegradation Journal Journal of Coatings Technology and Research, 2(7), 499-508.

Schaeffer, T.T., 2001, Effects of Light on Materials in Collections, Getty Publications, Los Angeles.

Scuello, M., Abramov, I., Gordon, J. & Weintraub S., 2004, Museum lighting: Optimizing the illuminant, Color Research & Application 29 (2), 121 – 127.

Shaw, K., 1995, Museum lighting, a Three-Part Lecture on Aspects of Museum Lighting originally given to members of Forum for Exhibitors, Norrkoping, Sweden - available on-line at http://www.kevan-shaw.com/pdf/museums_art.pdf.

Shaw, K., 2002, Lighting. In Barry Lord y Gail Dexter Lord (eds.). The Manual of Museum Exhibitions, AltaMira Press, 207-214 .

Tétreault, J., 2003, Airborne Pollutants in Museums, Galleries and Archives: Risk Assessment, Control Strategies and Preservation Management". Canadian Conservation Institute: Ottawa; 168.

Thomson, G., 1961, A New Look at Colour Rendering, Level of Illumination, and Protection from Ultraviolet Radiation in Museum Lighting, Studies in Conservation, 6(2/3), 49-70.

Thomson, G., 1994, Light Part I and Part II In: The museum environment, Second Edition, Butterworth-Heinemann - Ed. Elsevier, Oxford.

Wilson Kesner, C., 2008, Analysis of the Museum Lighting Environment, Journal of Interior Design, 23 (2), 28-41.

Wade, N.J. & Swanston, M., 1991, Visual perception: an introduction; Second Edition Routledge, London.

Chapter 6

Basic chemical mechanisms indoors

David Thickett

There are major differences between indoor and outdoor environments and therefore the dominant deterioration mechanisms. The main difference is availability of water through precipitation or condensation. Leaching and corrosion reactions are generally much slower. Corrosion outdoors is dominated by the time of wetness and indoors this is often close to zero. The lack of precipitation also means that deterioration products cannot be washed away and accumulate on surfaces, which has consequences. The other major difference is that certain pollutant gases, such as sulfur dioxide and ozone, do not, on the whole propagate far into buildings. They are often present at much lower concentrations in many buildings than outside of them. The nitrogen oxides and reduced sulfur gases do generally ingress into buildings. As the RHs are generally lower indoors and the deposition of these gases is a strong function of RH, reactions tend to be much slower indoors.

Depending on the material and reaction, the fundamental parameters indoors are relative humidity, light doses (especially ultraviolet radiation), temperature (and near infra-red radiation), pollutant gases and particles. Light damage will be considered in the chapter Light and Lighting and is not discussed in any detail here. Oxygen is required for many deterioration reactions, but in most display situations the oxygen concentration is not controlled, as this is generally expensive. Iconic artifacts such as: the US Declaration of Independence; the Japanese National Museum's collection of swords, some of the British Natural History Museum's collections have been stored or displayed in low oxygen environments.

Oxidation

Most organic materials undergo oxidation reactions, through a free radical mechanism (Fig. 1). The elasticity of natural rubber originates from double bonds in its structure. Oxidation removes these double bonds eventually generating a rigid, brittle material

Initiation	R-R	→	2R*
	ROOH	→	RO* + HO*
	2ROOH	→	RO* + RO2* + H2O
Propagation	RO2* + RH	→	ROOH + R*
	R* + O2	→	RO2*
Termination	2R*	→	R-R
	R* + RO2*	→	ROOR
	2RO2*	→	Nonradical products + O_2

Fig. 1 - Free radical mechanism elucidated by Natural Rubber Producers Association.

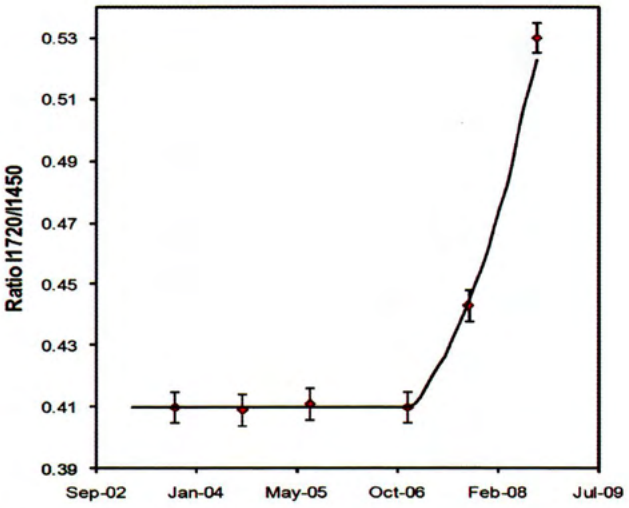

Fig. 2 - Progress of oxidation of natural rubber measured with FTIR at Dover Secret Wartime Tunnels (squares) and general curve (full line).

(Fig. 2). The free radical oxidation mechanism, below, was first elucidated for rubber.

High relative humidities, ultra violet radiation and the presence of ozone all increase oxidation rates of rubber. There is a strong synergistic effect between RH and ozone, with ozone reaction rates increasing dramatically above 50% RH (Ryan 1992). Natural rubbers contain a range of natural antioxidant species and the rubber oxidizes very slowly until these are depleted, at which point the oxidation rate increases dramatically. All materials deteriorate, to some extent, over time and the conservation profession is beginning to develop concepts of acceptable change. In many outdoor situations acceptable change may be defined by material failure with structural, performance or safety implications. Such ideas are much less developed for indoor cultural heritage. For rubber initial work based on suitability for display in historic interiors has developed a preliminary scale for rubber deterioration (Fig. 3), which is predominantly oxidation (Thickett 2008, Rhyl Svensen 2008).

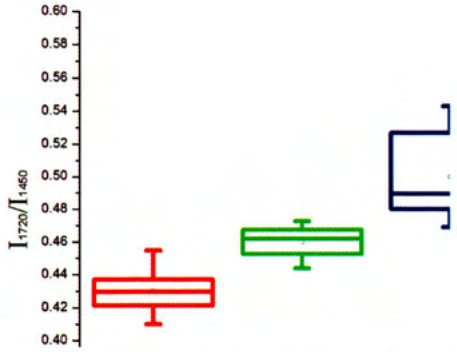

Fig. 3 - Correlation of FTIR analysis for rubber with curatorial perception of suitability for display.

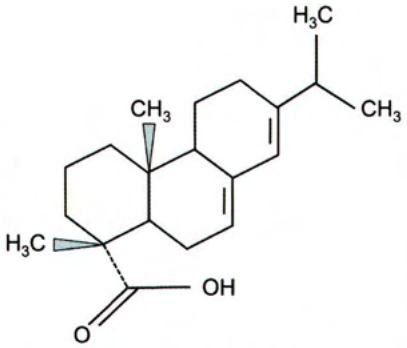

Fig. 4 - Structure of Baltic amber with vulnerable methylene groups in bold.

Fig. 5 - Damage to amber surface after approximately one hundred years around 75% RH.

Amber is often highly polished and oxidation of the surface occurs slowly. Initially the polish is lost, followed by surface cracking and, in some instances powdering (Shashoua 2002). The deteriorated surface often changes colour and refractive index, initially clouding, and then turning the amber opaque. The oxidation proceeds, initially, with loss of methylene groups, which can be readily followed by infra-red spectroscopy (Beck 1992).

In the nineteenth century amber with inclusions (especially insects) was popular and the pieces were highly polished to provide a good view of the inclusions. With time oxidation obscures the inclusions and fundamentally alters the aesthetic of the object. When humidities are around 75% this process is complete within a hundred years for at least some ambers.

When silver tarnishes the first stage is the formation of silver oxide 'islands' on the silver surface (Fig. 6), often under the influence of the oxidizing pollutants nitrogen dioxide and ozone. On continued exposure the islands coalesce to form a more or less continuous film and then react with sulfur gases (not sulfur dioxide) and chloride to form silver sulfide or chloride or mixtures depending on the atmosphere (Costa 2001, Ankersmit 2004). For silver sulfide films the initial reaction rate is relatively rapid and then quickly slows to give characteristic parabolic corrosion kinetics (Fig. 7).

This means the tarnish film is protective, it retards the passage of ions or current to slow the corrosion process. Since silver tarnish films are protective and rarely very thick, cleaning needs to be carefully considered. In many instances little metal is corroded and the effect can be considered essentially aesthetic and monitoring with colorimetry provides a convenient method to follow the tarnish. However in some instances pitting has been reported under thick corrosion layers, which is much more damaging.

Whilst the effects of RH on silver tarnish are well studied at high hydrogen sulfide and carbonyl sulfide concentrations, the difficulty of producing atmospheres near ambient concentration has limited studies to these artificially polluted atmospheres (Graedel 1992). A significant change in kinetics has been reported moving from such highly polluted atmospheres towards ambient levels likely to be encountered in cultural heritage (Pope 1968). Only one paper has been published to date at pollution levels comparable to museum atmospheres (Kim & Payer 1999) and this would be a relatively high level.

Corrosion

In almost all circumstances metals corrode electrolytically, with the corroding metal as anode, another reaction, often oxygen reduction at the cathode and an electrolyte. The electrolyte is generally a water solution of corrosion products, salts or pollutant gases. High concentrations of ionic species increase the electrolyte conductivity and the corrosion rate. As the major source of water indoors is vapour, metal corrosion is a very strong function of RH. All surfaces have water molecules on them at any RH. At low RHs there is less than three complete layers of water on the surface and the binding with the surface changes the nature of the water so it cannot act as a solvent. Above three layers, the water molecules act as free water and corrosion increases dramatically. For iron this occurs at approximately 70% and this value is known as the critical RH (Phipps 1979). The presence of salt ions, dust particles and corrosion products can lower this critical RH. Many salts can also act to some extent as electrolytes when they are solids but have a large amount

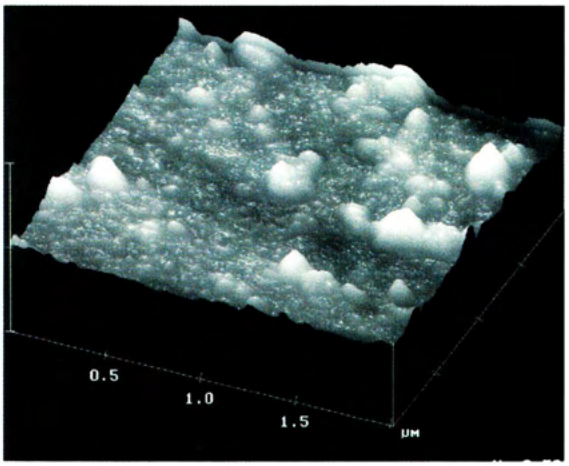

Fig. 6 - AFM image of oxide islands on silver surface (courtesy of Bart Ankersmit, ICN).

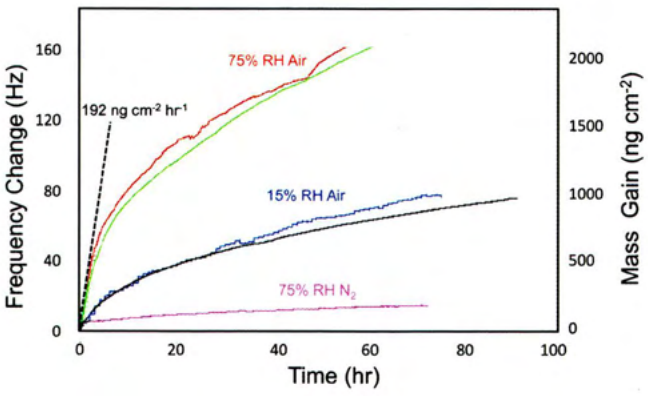

Fig. 7 - Parabolic corrosion of silver at 100 ppb of hydrogen sulfide, bottom line in nitrogen, middle line air at 15%, top line air at 75%. Δf is proportional to mass gain (data from Kim & Payer).

of absorbed water, which lowers the critical RH below the deliquescent RH of the salt (Yang 2001). The rate of corrosion is often limited by the movement of electrons or ions or gases in the electrolyte. Pollutant gases and salt particles dissolve in the electrolyte and increase its conductivity, increasing the corrosion rate. Whilst thirty years ago sulfur dioxide concentrations dominated the corrosion processes, environmental policy has now reduced concentrations in Western Europe to such an extent that nitrogen oxides and ozone now often have an effect of the same order of magnitude, but much lower than previously. If wet and dry cycles do occur indoors then the corrosion rate is often very rapid at the end of the electrolyte drying as the layer is very thin and oxygen can diffuse quickly through it, removing the limit on the cathodic reaction. The reduction of lepidocrite has been proposed as an alternative cathodic reaction, to explain the faster than expected atmospheric corrosion of some ferrous alloys (Stratmann1990). Recent mapping and identification of the species present in the atmospheric corrosion of steels has elucidated this important reaction (Antonya 2005).

There are a large number of ferrous alloys available, they are grouped into wrought iron, cast irons and steels. The composition and metallurgy has a very strong effect on the corrosion rate and mechanisms. The graphite flakes in grey cast irons and cementite in steels act as cathodic corrosion sites and can dramatically increase corrosion rates.

Archaeological iron generally contains chlorides from burial which favour the formation of the corrosion product akaganeite (Fig. 8). On exposure to the ambient atmosphere, after excavation, ferrous chloride is found. This dramatically lowers the RHs above which deterioration occurs (Fig. 9) (Selwyn 1999).

Lead and its alloys are very strongly affected by ethanoic acid, the mechanisms for these reactions are dealt with in the chapter dedicated to VOCs released by wood and metal corrosion.

High concentrations of ethanoic acid can react with calcareous stones, shells and salt laden ceramics. This problem has only received limited study and many aspects of the mechanism require further elucidation, particularly the risk at different realistic ethanoic acid concentrations and RHs for materials other than shell (Brockerhof 1996, Gibson 2005). These reactions can produce a variety of calcium ethanoates and calcium ethanoate chloride nitrate has been widely reported.

Calcium ethanoate chloride nitrate requires chloride and nitrate from soluble salt in the stone, shell or ceramics pores (Halsberghe 2005). This explains the differential reactivity in some collections. Only a small number of over 2500 limestone sculpture, stored in glass fronted oak boxes, expressed the calcium ethanoate chloride nitrate salt, whilst the reaction rates in other collections are reported to be much higher. High RHs appear to be related to at least some reported instances. In a collection of approximately 10,000 cuneiform tablets, very low instances were reported until a heating failure generated RHs above 80%, after which over 20% of the collection had the salt efflorescence. This may also be related to higher RHs generating greater ethanoic emission rates from storage materials such as wood products and cardboard.

Leaching

Liquid water should not be encountered indoors in a well maintained museum, however the presence of salts on surfaces makes some materials susceptible to leaching at ambient RHs. Glasses with insufficient stabilizers are susceptible to leaching reactions. Water vapour reacts to form hydroxonium ions that undergo ion exchange with sodium or potassium in the glass. This results in a surface layer rich in silica and depleted of these elements. Sodium or potassium salts often form on the glass surface,

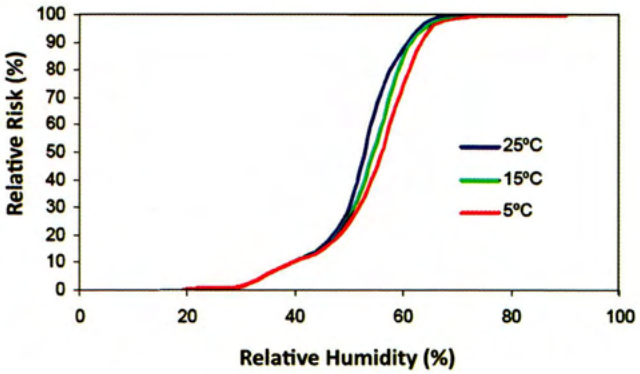

Fig. 8 - Relative risk to archaeological iron at different relative humidities and temperatures.

Fig. 9 - SEM micrograph of akaganeite crystals growing from iron surface, courtesy of Lyndsey Selwyn CCI, Bessie Argyropoulou, TEIATH, © All rights reserved. Reproduced with the permission of the Canadian Conservation Institute of the Department of Canadian Heritage, 2010.

most frequently sodium methanoate. This salt deliquesces below 50% RH and droplets of salt solution can be observed on some surfaces, so called 'weeping' glass. The silica gel layer formed is prone to cracking (Fig. 10), exposing fresh glass to attack and encouraging condensation in the narrow cracks formed. If these modified glasses, with thick gel layers, are exposed to low RHs they can catastrophically crizzle (Fig. 11).

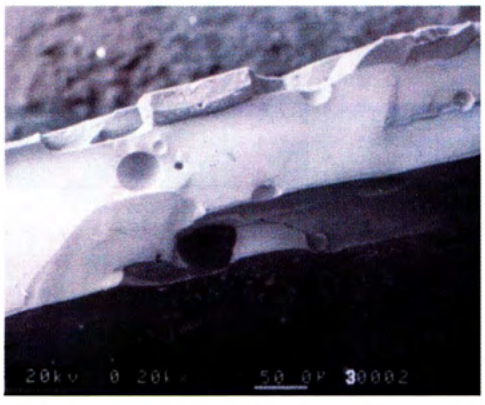

Fig. 10 - SEM Micrograph of gel layers on enamel.

$$Si\text{-}O\text{-}X + H2O = Si\text{-}O\text{-}H + XOH$$

If pH >9 then

$$Si\text{-}O\text{-}Si + XOH = Si\text{-}O\text{-}H + Si\text{-}O\text{-}X$$
$$Si\text{-}O\text{-}X + H2O = Si\text{-}OH + XOH$$

can occur

Fig. 11 - Ion exchange mechanism for gel Limoge layer formation on glass.

Fig. 12 - Damaged cellulose acetate plotting sheet.

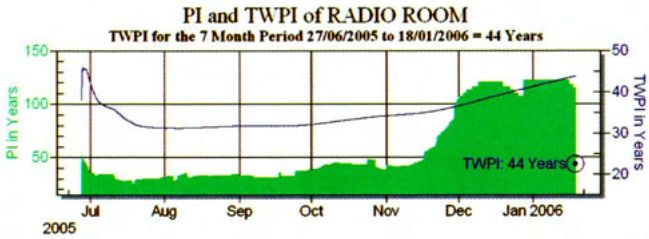

Fig. 13 - Time Weighted preservation index for York Royal Observer Corps Bunker.

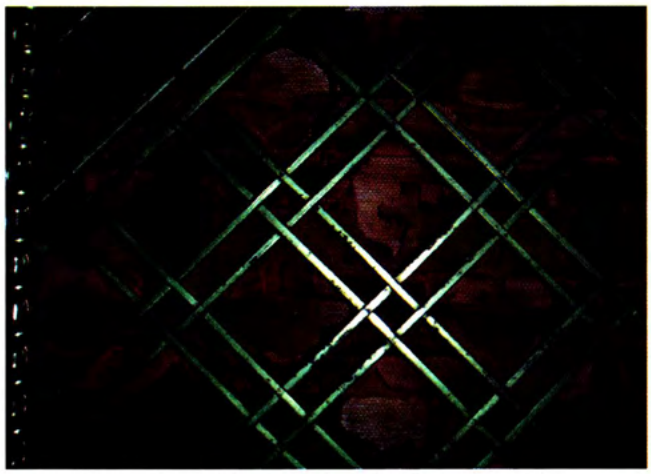

Fig. 14 - Deteriorated silk fabric at Brodsworth Hall, near Doncaster.

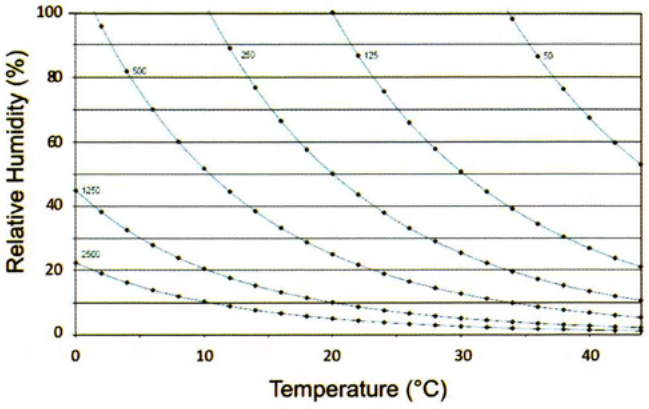

Fig. 15 - Isoperms for silk deterioration.

Glasses with too small a concentration of stabilizers (calcium oxide, lead oxide, barium oxide) are prone to such reactions (Newton 1989). Some enamel formulations, especially purples and blues, are also affected. Increasing RH, increases the rate of ion exchange and gel layer formation (Ryan 1995). Many susceptible pieces are kept in controlled RH conditions, with the RH as low as possible to retard this reaction. The drying of the gel limits the lower RH and often an RH range around 35-40% is used. The difficulty in identifying glasses likely to undergo these reactions limits this approach. Whilst x-ray fluorescence spectroscopy can readily quantify the stabilizers and potassium, air path instruments struggle to accurately quantify silica, and sodium cannot be detected by many instruments. This means a large uncertainty in the overall chemistry of the glass. Glasses often deteriorate to a great extent before the phenomena are visible to the naked eye and the weeping will disappear when the RHs drops. Some glass formulations are susceptible to accelerated attack by ethanoic and methanoic acid and methanal. Copper containing glasses are particularly susceptible to methanal with a conservation to methanoic acid via the Canazzaro reaction. Sodium methanoates are frequently identified on the surface of deteriorating glasses.

Hydrolysis

Hydrolysis is often the major mechanism for paper deterioration aswell as number of other polymers. It can occur under both acid and alkaline conditions, although acid hydrolysis is more commonly encountered. Acid hydrolysis is autocatalytic with the reaction accelerating rapidly when a certain pH is reached (3.7 for cellulose acetate), giving a similar reaction progress graph to figure 2. Paper is dealt with in a separate chapter and is not considered here. The deterioration of cellulose acetate has been extensively studied due to its importance in film media. Careful and extensive aging studies have quantified the relationship between temperature, RH and deterioration.

$$\text{Preservation Index (years)} = \exp(E_{act}^{-134.9[RH(\%)]/(8.134[T(°C)])})$$

Where preservation index is the time until the autocatalytic point is reached (years); Eact is the activation energy of the reaction; T is temperature (°C) and RH is relative humidity (%).

The equation is an empirical fit of the deterioration to the data and the end point lifetime is defined as when the autocatalytic point is reached and deterioration accelerates rapidly. Acid hydrolysis can be very heavily dependant on composition and for paper, one author has written that degradation is 90% due to manufacturing and 10% due to the environment. The formulation of polymers has a dramatic effect, as does the presence of plasticizers and stabilizers. Pollutant gases can have a strong effect, especially the acid pollutants such as sulfur dioxide and nitric acid. As cellulose acetate deteriorates (Fig. 12) it liberates ethanoic (acetic) acid. If this is enclosed, the acid accelerates the subsequent deterioration (Fig. 13).

The deterioration rate can be slowed by controlling the temperature or relative humidity or by removing the ethanoic acid from the system, although cellulose acetate has a strong affinity to ethanoic acid and some zeolite absorbers cannot compete with the cellulose acetate.

Silk also undergoes acid hydrolysis which is dramatically accelerated at high RHs. The composition and processing of the silk have significant effects on its deterioration. Degumming stabilizes the silk, whilst dying, mordanting and weighting can all deteriorate the silk during that process (Fig. 14), and also make the silk more susceptible to light or RH damage subsequently (Fig. 15) (Luxford 2010, Quye 2009).

References

Ankersmit, H. A., Doménech, C.A. & Tennent, N. H., 2004. Tarnishing of silver: evaluation by colour measurements, in Metal 2001: proceedings of the international conference on metals conservation, Santiago, Chile, 2001: 157-166.

Antonya, H., Legranda, L., Maréchalb, L., Perrinb, S., Dillmannc Ph. & ChausséaDillman, A., 2005. Study of lepidocrocite - FeOOH electrochemical reduction in neutral and slightly alkaline solutions at 25 °C, ElectroChimica Act: 51, 745-753.

Beck, C.W., 1986. Spectroscopic investigations of amber, Applied Spectroscopy Reviews, 22: 57-110.

Brokerhof, A.W. &, Van Bommel, M., 1996. Deterioration of calcareous materials by acetic acid vapour: a model study, in Preprints of ICOM committee for conservation, 11th triennial meeting in Edinburgh, Scotland, James & James (Science Publishers) Ltd, London: 769-775.

Costa, V., 2001. The deterioration of silver alloys and some aspects of their conservation, Reviews in Conservation, 2: 18-37.

Evans, U.R., 1981. An Introduction to Metallic Corrosion, Arnold, London.

Gibson, L., 2005. The mode of formation of thecotrichite, a widespread calcium acetate chloride nitrate efflorescence, Studies in conservation, 50: 284-294.

Graedel, T.E., 1992. Corrosion mechanisms for silver exposed to the atmosphere, Journal of the Electrochemical Society, 139 (7): 1963-1970.

Halsberge, L., Gibson, L., & Erhardt, D., 2005. A collection of ceramics damaged by acetate salts: conservation and investigation into causes, in Preprints of Fourteenth Triennial Meeting of ICOM-CC, The Hague: 131-138.

Kim, H. & Payer, J.H., 1999. Tarnish process of silver in 100ppb H_2S containing environments, Journal of Corrosion Science and Engineering, 1: 1-18.

Luxford, N., 2010. Reducing the risk of open display: optimising the preventive conservation of historic silks, unpublished doctoral thesis, University of Southampton.

Newton, R. & Davison, S., 1989. Conservation of Glass, Butterworths, London.

Phipps, P.B.P and Rice, D.W., 1979. The role of water in atmospheric corrosion, in Corrosion Chemistry, ed. G.R Brubaker and P.B.P Phipps, ACS, Washington DC: 235-261.

Pope, D., Gibbens, H.R. & Moss, R.L., 1968. The tarnishing of Ag at naturally occurring H_2S and SO_2 levels, Corrosion Science, 8: 883-887.

Quye, A., Hallett, K. & Carettaro, C.H., 2009. 'Wroughte in gold and silk': preserving the art of historic tapestries, National Museums of Scotland, Edinburgh

Rhyl Svendsen, M., 2006. 'An elastomer dosimeter for monitoring ozone exposures in museum storage rooms'. http://iaq.dk/iap/iaq2006/2006_contents.htm [accessed 2 June 2007].

Ryan, J.L., 1995. The Atmospheric Deterioration of Glass. Unpublished PhD dissertation, Imperial College of Science, Technology and Medicine, University of London.

Ryan, P.B, Koutrakis, P., Bamford, S. et al., 1992. 'Ozone reactive chemistry in indoor microenvironments', in Tropospheric Ozone and the Environment. Pittsburgh, PA: Air and Waste Management Association.

Scott, D. and Eggert, G., 2009. Iron and Steel in Art, Archetype Publications Ltd, London.

Selwyn, L.S. and Sirois, P.J., 1999. The corrosion of excavated archaeological iron with details on weeping and akaganeite. Studies in Conservation 44: 217-232.

Shashoua, Y., 2002. Degradation and inhibitive conservation of Baltic amber in museum collections, unpublished report, National Museum of Denmark.

Stratmann, M., 1990. Ber. Bunsenges,. Physical Chemistry. 94: 626-639.

Thickett, D. & Richardson, E., 2008. Preventive conservation research for plastics on open display, Postprints of Plastics: Looking at the Future & Learning from the Past, London: 89-98.

Yang, L., Pabalan, R.T. & Browning, L., 2001. Experimental determination of the deliquescence relative humidity and conductivity of multicomponent salt mixtures, in Scientific Basis for Nuclear Waste Management XX, MRS Proceedings, volume 713: JJ1.4.

Chapter 7
Basic chemical mechanisms outdoors

Vasco Fassina

1. Introduction

Stone, in its broadest sense, includes all those inorganic non-metallic, hard solids that are, or have been, employed by humans for recording, embellishing, advertising or preserving the facts of their lives and their spiritual and artistic visions, as well as to satisfy their practical needs for shelter, protection, and secure storage (Lewin, 1983).

Stone decay is not a new phenomenon. It starts as soon as built heritage has been completed, and continues progressively for as long as the object is in contact with any kind of environment according to the Second Law of Thermodynamics.

According to Lewin (1983) the deterioration of stone of outdoor sculptures and buildings has been evident since ancient times, at least over the past 3000 years, and Greek and Roman writers reported in their manuscripts of continuous efforts to intervene in the decay processes. The oldest reference to the chemical deterioration of rocks is probably the report of Herodotus of Halicarnassos whose source was a book written by a certain Hekataios 2500 years ago (Krumbein, 1988). According to him the deterioration of Cheops pyramids in Egypt was caused by the corrosive attack of salts and salt solutions climbing upwards into the rocks of the pyramids. In addition he also mentioned a natural content of soluble salt within the sandstones and limestones the pyramids were built from. He derived this natural salt content from the fact that the rocks were a product of the sea, namely sedimentary rocks deposited in highly saline waters. Vitruvius gave us not only a precise notion about physical and chemical attack on building stones but also about remedial techniques.

In England, in the middle of nineteenth century, some important buildings such as the Houses of Parliament showed some dramatic manifestations of stone decay to enlist the best scientists of that time to study the causes of deterioration. Since that time increased the consciousness to make any effort to stop the deterioration processes by using the new products synthesized by the arising modern chemistry. According to Lewin (1983) since the second and third decades of the twentieth century we can speak of a science of stone conservation as beginning to take form.

According to Schaffer (1932) no stone resists to the action of atmospheric agencies indefinitely. In nature, the surface of the earth suffers a continuous cycle of changes: the weathering of the igneous rocks is followed by the deposition as sediments of the products of their disintegration: these sediments become hardened and consolidated to form the sedimentary rocks, and the sedimentary rocks in their turn are again subjected to the manifold action of atmospheric agents. Stone in buildings is subject to the action of those same agencies. The influence of any one atmospheric agent varies with the locality; thus climatic and atmospheric conditions have an important effect in controlling the rate of weathering of any particular material. The study of the problem is rendered more complex owing to the difficulty of separating the effects of the various agents of decay. No weathering agent acts alone; the relative importance of each is influenced by the concurrent effect of other agents, or exposure to the action of one may render the material more susceptible to the subsequent action of another. Some agents are of predominating importance in relation to the decay of certain materials, but have no significant action on others. In some cases it is clear that more than one factor is necessary to bring about an observed effect. But the study of weathering in buildings is not limited to the consideration of the action of natural atmospheric agencies. Under modern conditions, pollution of the atmosphere is an important factor in causing decay (Fig. 1).

Fig. 1 - On the façade of the Santa Maria dei Miracoli church a strong decay due to atmospheric pollution was visible.

2. The weathering environment

Weathering and *decay* are frequently used to describe similar phenomena. *Weathering* is used to describe the action of the weather and should include all those changes desirable and undesirable, which are suffered by materials exposure either indoors or outdoors.

Decay on the other hand, usually refers to undesirable changes which are associated with dangerous or ugly forms of weathering. Although this distinction has certain conveniences, it is preferable that the term weathering should retain its wider significance, while the term decay can still be used specifically to describe ugly or dangerous forms of weathering.

The evolution and the transformation of rocks is mainly the result of physical-chemical disparity between the initial formative environment of the rock, and the new environment to which it becomes exposed at the surface of the lithosphere. The whole sequence of reactions of the lithosphere[1] with the atmosphere[2], the hydrosphere[3], and the biosphere[4], constitutes the process of weathering in the widest sense of the word.

The alteration of rocks in the lithosphere is produced by the various continental agents (temperature, water, erosion, earth movements, biological agents etc.) as well as by their intrinsic properties, i.e. mineralogy, texture and structure. The change produced by weathering in fresh rocks are governed by thermodynamic laws, and can be ascribed to partial or complete decomposition of some minerals and the partial or complete migration of both major and minor chemical elements (Amoroso & Fassina, 1983).

The phenomena of monument decay and stone deterioration embrace those factors which operate to alter the appearance, strength coherence, dimensions, or chemical behavior of the material, either as individual elements, or as parts of structures (Lewin, 1983).

Among mechanisms generally recognized as responsible for the deterioration of exposed stone are:

-Chemical attack, i.e. etching, erosion, and dissolution, of alkaline stones by acidic substances, both natural (atmospheric CO_2, volcanic gases, rainwater) and man-made (combustion products, industrial emissions), or leaching and dissolution of one or more minerals constituting the stone by water or acidic solutions.

-Physical and mechanical stresses caused by expansive and disruptive forces generated in pores, channels, and cracks by the formation and growth of crystals or by freezing of imbibed water.

-Disfigurement, due to migration into stone of colored matter from adjacent materials (e.g. rust and copper staining), or alteration of the original colour or texture by selective leaching from the stone of one of its components, or by the etching and roughening of the polished surface.

-Abrasion attrition and stress-cracking due to wind-driven particulates, seismic shocks, vibrations induced by vehicular traffic, accidents, human contacts, etc.

-Disfigurement as well as chemical and mechanical disruption resulting from the biological activity of microorganisms, fungi, algae, mosses, and higher organisms including, inter alia pigeons (Lewin, 1983).

Weathering is an exceedingly complex concept, divided traditionally into physical, chemical and biological weathering. It is often hard to tell which type predominates in stone materials. Frequently what we see is the result of a combination or succession of these main types (Lofvendhal, 1995).

Mechanical or physical weathering involves the breakdown of rocks and soils through direct contact with atmospheric conditions, such as heat, water, ice and pressure.

Chemical weathering involves the direct effect of atmospheric chemicals or biologically produced chemicals (also known as *biological weathering*) in the breakdown of rocks, soils and minerals.

Biological weathering refers to attack/degradation through physical chemical and chemical influence of vegetation and bacteria.

Physical weathering is the only process that causes the disintegration of rocks without chemically changing it. The primary process in physical weathering is abrasion (the process by which clasts and other particles are reduced in size). However, chemical and physical weathering often go hand in hand. For example, cracks exploited by physical weathering will increase the surface area exposed to chemical action. Furthermore, the chemical action at minerals in cracks can aid the disintegration process.

Physical weathering is a very wide-ranging concept and includes water uptake with attendant swelling through ion exchange reactions and/or uptake of crystal water crystallization/phase transitions and many other different processes resulting in exfoliation, salt weathering, cracking as a result of varying thermal expansion, detachment due to crystallization pressure and so on (Lofvendahl, 1992). These may consist of clay minerals, calcite or other easily weathered material, after which the stone disintegrates. Frost weathering is a common and troublesome problem at northern latitudes, with the temperature fluctuating around the freezing point of water for a large part of the year. At freezing point, the conversion of water into ice increases its volume by about 10%. When this happens, rocks which are more or less completely filled with moisture are liable to break because the gluing matrix will disintegrate. Temperature fluctuations and variations in relative humidity frequently give rise to phase transitions between salts with different quantities of crystal water (and thus occupying different volume). These transitions cause variations of crystal pressure which can burst the rock. Physical weathering also includes mechanical wear from wind-borne particles and other types of mechanical, often anthropogenic, damage (Lofvendahl, 1996). The physical

[1] The lithosphere includes the outer part of the solid Earth. In general, the term refers to minerals encountered in the Earth's crust and to the complex and variable mixture of minerals encountered in the Earth's crust and to the complex and variable mixture of minerals, organic matter, water, and air making up soil. Insofar as environmental chemistry is concerned, the soil is the most significant part of the lithosphere.

[2] The atmosphere is the envelope of gases surrounding the Earth. The atmosphere is subdivided into different regions depending on altitude. Atmospheric chemistry varies a great deal with altitude, exposure to solar radiation, pollutant load, and other factors.

[3] The hydrosphere refers to water in its many forms: It includes the oceans, lakes, streams, reservoirs, snowpack, glaciers, the polar ice caps and water under the ground (groundwater). For the study of environmental chemistry, however, liquid water and the reactions of the chemical species in it are of predominant importance.

[4] The biosphere refers to life. It includes living organisms and their immediate surroundings. The biosphere is influenced tremendously by the chemistry of the environment and, in turn, exerts a powerful influence upon the chemistry of most environments, particularly the lithosphere and hydrosphere. Moreover, biological activity is responsible for the present composition of the atmosphere (specifically high oxygen level and low carbon dioxide level).

properties of natural rocks are important and control their resistance to degradation outdoors. Porosity and permeability are probably the most important physical factors of all (Winkler, 1973). The porosity of building stone materials varies widely, from less than 1% in the case of granites and other magmatic and metamorphic rocks up to 30 % and more for sandstones and limestones. The degradation rates is affected, not only by porosity but also by the pore size distribution of the material. Rocks with very fine, coherent pores have a very high capillary water absorption capacity. In damp conditions, rocks of this kind are almost completely saturated with water. A stone with a porosity of a few percent or more, and liable to become completely saturated with water is seriously in danger of congelifraction (disintegration by freezing) in a temperate climate (Hudec, 1989). Permeability increases in line with pore size, so long as the pores are open. Thermal properties such as the coefficient of thermal expansion of minerals are another important factor of damage. Electrical and magnetic properties, on the other hand, play a negligible part in weathering processes.

Physical processes cannot be simulated and quantified in the same way as chemical weathering (the solubility of mineral in water). Many of the physical processes are entirely random and in some cases discrete. Some processes, however, can be simulated in the laboratory, e.g. freeze/thaw cycles, load, mechanical strength or swelling through water absorption in clays.

Some physical weathering processes can be mentioned: thermal expansion, frost disintegration, salt-crystal growth.

Thermal expansion, also known as exfoliation, insolation weathering or thermal shock, often occurs in areas, like deserts, where there is a large diurnal temperature range. The temperatures are high in the day, while dipping greatly at night. As the rock heats up and expands by day, and cools and contracts by night, stress is often exerted on the outer layers. The stress causes the peeling off of the outer layers of rocks in thin sheets. Though this is caused mainly by temperature changes, thermal expansion is enhanced by the presence of moisture.

Frost disintegration can also be called frost shattering, frost-wedging or freeze-thaw weathering. This type of weathering is common in mountain areas where the temperature is around the freezing point of water. Moist soils expand or frost heave upon freezing as a result of water migrating along from unfrozen areas via thin films to collect at growing ice lenses. This same phenomena occurs within pore spaces of rocks. The ice accumulations grow larger as they attract liquid water from the surrounding pores. The ice crystal growth weakens the rocks which, in time, break up. It is caused by the expansion of ice when water freezes, so putting considerable stress on the walls of containment. Freeze induced weathering action occurs mainly in environments where there is a lot of moisture, and temperatures frequently fluctuate above and below freezing point–that is, mainly alpine areas. Rock susceptible to frost action has generally many pore spaces for the growth of ice crystals. When water that has entered the joints freezes, the ice formed strains the walls of the joints and causes the joints to deepen and widen. When the ice thaws, water can flow further into the rock. Repeated freeze-thaw cycles weaken the rocks which, over time, break up along the joints into angular pieces. The splitting of rocks along the joints into blocks is called block disintegration. The blocks of rocks that are detached are of various shapes depending on rock structure.

Fig. 2 - Salt weathering of building stone on the island of Gozo, Malta.

Fig. 3 - Salt weathering of sandstone near Qobustan, Azerbaijan.

Salt-crystal growth. Salt crystallization, otherwise known as haloclasty, causes disintegration of rocks when saline solutions seep into cracks and joints in the rocks and evaporate, leaving salt crystals behind. These salt crystals expand as they are heated up, exerting pressure on the confining rock. Salt crystallization may also take place when solutions decompose rocks (for example limestone) to form salt solutions of sodium sulfate or sodium carbonate, of which the moisture evaporates to form their respective salt crystals. The salts which have proved most effective in disintegrating rocks are sodium sulfate, magnesium sulfate, and calcium chloride. Some of these salts can expand up to three times or even more. It is normally associated with arid climates where strong heating causes strong evaporation and therefore salt crystallization. Honeycomb is a type of tafoni, a class of cavernous rock weathering structures, which likely develop in large part by chemical and physical salt weathering processes (Figs. 2,3).

Biological weathering. Living organisms may contribute to mechanical weathering (as well as chemical weathering). Lichens and mosses grow on essentially bare rock surfaces and create a more humid chemical microenvironment. The attachment of these organisms to the rock surface enhances physical as well as chemical breakdown of the surface microlayer of the rock. On a larger scale seedlings sprouting in a crevice and plant roots exert physical pressure as well as providing a pathway for water and chemical infiltration. Burrowing animals and insects disturb the soil layer adjacent to the bedrock surface thus further increasing water and acid infiltration and exposure to oxidation processes.

Plants generate a variety of chemical substances. Lichens,

for example, secret lichen acids, often more or less endospecific, which attack the stone and extracts nutrients necessary for lichen growth. The body of crustose lichen, moreover, consists to a great extent of oxalates, and these in an acid environment correspond to oxalic acid, which attacks the minerals. Endolithic lichens (i.e. those growing inside the stone) create cavities beneath the surface of the stone, thereby facilitating subsequent physical attacks. Other processes such as swelling/shrivelling in connection with the absorption/desorption of water and root bursting can also be important. Biological colonisation of stone surface frequently conforms to a certain succession, beginning with bacteria (which are always present in stone material), continuing with algae and lichens and concluding with moss and plants, which retain degraded matter and trap flying particles. This eventual results in the formation of a continuous layer of soil. Biological colonisation is heavily dependent on moisture and occurs mainly on the north side of buildings or in the shade of trees. Algae and lichens retain moisture causing longer wetness periods and, consequently, mineral dissolution. Damp-dry cycles cause the organism to expand and contract, producing mechanical wear and stresses in the surface on which they are growing.

The importance of this type of weathering compared with physical and chemical weathering remains fairly unclear. Biologists and microbiologists consider this very important, whereas as chemists and physicists often neglect it. The effect of lichen on mineral surfaces are complex and very often destructive (Wilson, 1995).

Chemical weathering

As a chemical phenomenon, weathering can be viewed as the result of the tendency of the rock/water/mineral system to reach equilibrium. Chemical weathering involves the change in the composition of rocks, often leading to a break down in its form. This is done through a combination of water and various chemicals to create an acid which directly breaks down the material. Chemical weathering is a gradual and ongoing process as the mineralogy of the rock adjusts to the near surface environment. New or secondary minerals develop from the original minerals of the rock.

The process involved in chemical weathering may be divided into the following major categories: dissolution-precipitation, acid-base reactions, hydration-dehydration, hydrolysis, oxidation-reduction and complexation.

Weathering is strongly influenced by water and its rate is increasing by several orders of magnitude in the presence of moisture. Water holds weathering agents in solution such that they are transported to chemically active sites on rock minerals and contact the mineral surfaces at the molecular and ionic level.

Prominent among such weathering agents are CO_2, O_2, sulphur acids (SO_2 (aq), H_2SO_4), and nitrogen acids (HNO_3, HNO_2). Water provide the source of H^+ ion needed for acid-forming gases to act as acids as shown by the following:

$$CO_2 + H_2O \rightleftharpoons H^+ + HCO_3^- \quad (1)$$
$$SO_2 + H_2O \rightleftharpoons H^+ + HSO_3^- \quad (2)$$

Dissolution

Rainfall is acidic because atmospheric carbon dioxide dissolves in the rainwater producing weak carbonic acid. In unpolluted environments, the rainfall pH is around 5.6. Acid rain occurs when gases such as sulphur dioxide and nitrogen oxides are present in the atmosphere. These oxides react in the rain water to produce stronger acids and can lower the pH to 4.5 or even 3.0. Sulfur dioxide, SO_2, comes from volcanic eruptions or from fossil fuels, can become sulfuric acid within rainwater, which can cause solution weathering to the rocks on which it falls. Therefore, rainwater is chemically aggressive towards some kinds of mineral matter, which it breaks down by a process called chemical weathering. One of the most well-known solution weathering processes is carbonation, the process in which atmospheric carbon dioxide leads to solution weathering. Carbonation occurs on rocks which contain calcium carbonate, such as limestone. This takes place when rain combines with carbon dioxide or an organic acid to form a weak carbonic acid which reacts with calcium carbonate (the limestone) and forms calcium bicarbonate. This process speeds up with a decrease in temperature, not because low temperatures generally drive reactions faster, but because colder water holds more dissolved carbon dioxide gas.

The reactions as follows:

$$CO_2 + H_2O \rightleftharpoons H_2CO_3 \quad (3)$$
$$H_2CO_3 + CaCO_3 \rightarrow Ca(HCO_3)_2 \quad (4)$$

Carbonation on the surface of well-jointed limestone produces a dissected limestone pavement which is most effective along the joints, widening and deepening them.

Dissolution with hydrolysis as occurs with the hydrolysis of carbonate ion when mineral carbonates dissolve:

$$CaCO_3\,(s) + H_2O \rightarrow Ca^{2+}\,(aq) + HCO_3^-\,(aq) + OH^-\,(aq) \quad (5)$$

Hydration

Mineral hydration is a form of chemical weathering that involves the rigid attachment of H^+ and OH^- ions to the atoms and molecules of a mineral. When rock minerals take up water, the increased volume creates physical stresses within the rock. For example iron oxides are converted to iron hydroxides and the hydration of anhydrite forms gypsum. A freshly broken rock shows differential chemical weathering (probably mostly oxidation) progressing inward.

$$CaSO_4\,(s) + 2H_2O \rightarrow CaSO_4 \bullet 2H_2O\,(s) \quad (6)$$
$$Fe_2O_3\,(s) + (3+2x)\,H_2O \rightarrow 2Fe(OH)_3 \bullet xH_2O\,(s) \quad (7)$$

Dehydration

$$2Fe(OH)_3 \bullet xH_2O\,(s) \rightarrow Fe_2O_3\,(s) + (3+2x)\,H_2O \quad (8)$$

Dissolution

$$CaSO_4 \bullet 2H_2O\,(s)\,(water) \rightarrow Ca^{2+}\,(aq) + SO_4^{2-}\,(aq) + 2H_2O \quad (9)$$

Oxidation

Within the weathering environment chemical oxidation of a variety of metals occurs. The most commonly observed is the oxidation of Fe^{2+} (iron) and combination with oxygen and water to form Fe^{3+} hydroxides and oxides such as goethite, limonite, and hematite. This gives the affected rocks a reddish-brown coloration on the surface which crumbles easily and weakens the rock. This process is better known as rusting.

Oxidation, such as occurs in the dissolution of pyrite:

$$4FeS_2\,(s) + 15\,O_2(g) + (8+2x)\,H_2O \rightarrow 2Fe_2O_3 \bullet xH_2O + 8\,SO_4^{2-}\,(aq) + 16\,H^+\,(aq) \quad (10)$$

or in the following example in which dissolution of an iron (II) mineral is followed by oxidation of iron (II) to iron (III):

$$Fe_2SiO_4(s) + 4\,CO_2\,(aq) + 4H_2O \rightarrow 2Fe^{2+} + 4HCO_3^- + H_4SiO_4 \quad (11)$$

$$4Fe^{2+} + 8HCO_3^- + O_2\,(s) \rightarrow 2Fe_2O_3\,(s) + 8\,CO_2 + 4H_2O \quad (12)$$

The second of these two reactions may occur at a site some distance from the first, resulting in a net transport of iron from its original location. Iron, manganese, and sulphur are the major elements that undergo oxidation as part of the weathering process.

Hydrolysis

Hydrolysis is a chemical weathering process affecting silicate and carbonate minerals. In such reactions, pure water ionizes slightly and reacts with silicate minerals. The major means by which silicates undergo weathering as shown by the following reaction of olivine (forsterite).

$$Mg_2SiO_4 + 4H^+ + 4OH^- \rightleftharpoons 2Mg^{2+} + 4OH^- + H_4SiO_4 \quad (13)$$

This reaction results in complete dissolution of the original mineral, assuming enough water is available to drive the reaction. However, the above reaction is to a degree deceptive because pure water rarely acts as a H^+ donor. Carbon dioxide, though, dissolves readily in water forming a weak acid and H^+ donor.

$$Mg_2SiO_4 + 4CO_2 + 4H_2O \rightleftharpoons 2Mg^{2+} + 4HCO_3^- + H_4SiO_4 \quad (14)$$

This hydrolysis reaction is much more common. Carbonic acid is consumed by silicate weathering, resulting in more alkaline solutions because of the bicarbonate. This is an important reaction in controlling the amount of CO_2 in the atmosphere and can affect climate. The weathering of silicates yields soluble silicon as species such as H_4SiO_4 and residual silicon-containing mineral (clay minerals).

Aluminosilicates (e.g. orthoclase, aluminosilicate feldspar) when subjected to the hydrolysis reaction produce a secondary mineral (kaolinite, a clay mineral) rather than simply releasing cations.

$$2KAlSi_3O_8 + 2H_2CO_3 + 9H_2O \rightleftharpoons Al_2Si_2O_5(OH)_4 + 4H_4SiO_4 + 2K^+ + 2HCO_3^- \quad (15)$$

Acid hydrolysis which accounts for the dissolution of significant amounts of $CaCO_3$ and $CaCO_3MgCO_3$ in the presence CO_2-rich water:

$$CaCO_3\,(s) + H_2O + CO_2\,(aq) \rightarrow Ca^{2+}\,(aq) + 2HCO_3^-\,(aq) \quad (16)$$

Complexation as exemplified by the reaction of oxalate ion, $C_2O_4^{2-}$ on muscovite $K_2(Si_6Al_2)Al_4O_{20}\,(OH)_4$

$$K_2(Si_6Al_2)Al_4O_{20}\,(OH)_4\,(s) + 6C_2O_4^{2-}\,(aq) + 2OH^- \rightarrow 6AlC_2O_4^+\,(aq) + 6Si(OH)_4 + 2K^+ \quad (17)$$

Acid hydrolysis, especially, is the predominant process that release elements such as Na^+, K^+, and Ca^{2+} from silicate minerals.

All rocks disintegrate under natural atmospheric conditions. Weathering speeds vary considerably, from extremely resistant quartzites (made up of the mineral quartz) to very soluble salt deposits. All minerals are, at least to some extent, soluble in water, from the un-reactive mineral quartz to highly soluble minerals like gypsum or halite salt (Table 1).

Table 1. *Solubility of minerals in pure water, 25 °C (from Lovendhal, 1992).*

Mineral	Formula	Solubility (g/l)	Source
Potassium feldspar	$KAlSi_3O_8$	0.8×10^{-4}	Berner 1978
Albite (Na-feldspar)	$NaAlSi_3O_8$	1.6×10^{-4}	Berner 1978
Anortite	$CaAl_2Si_2O_8$	5.0×10^{-4}	Estimated value
Calcite	$CaCO_3$	1.4×10^{-2}	Berner 1978
Gypsum	$CaSOx2H_2O$	2.4	Berner 1978
Halite	NaCl	357	Berner 1978
Potassium nitrate	KNO_3	2810	Weast 1976

Water in its natural state contains a large number of dissolved components and particles as well. Therefore, in nature minerals are not only dissolved, but also react to form new mineral phases. Thus feldspars are transformed to various clay minerals, such as kaolinite.

Important factors influencing the solubility of minerals in natural water include carbon dioxide and oxygen. The latter is a powerful oxidant which, in the presence of water, converts dead organic material to carbon dioxide, among other things, and metallic sulphides to sulphate. High concentrations of carbon dioxide/increase the solubility of most rock-forming minerals, that of calcite being particularly important. Under natural conditions at the earth's surface, minerals follow a weathering scheme (Stumm and Morgan, 1981) which shows the order in which they are dissolved and transformed (Table 2).

Table 2. *Mineral weathering sequence (with weathering increasing from above to below the table), (from Lovendhal, 1992).*

Mineral	Formula
Hematite	Fe_2O_3
Kaolinite (clay mineral)	$Al_2Si_2O_5(OH)_4$
Muscovite	$KAl_3Si_3O_{10}(OH)_2$
Quartz	SiO_2
Albite (Na-feldspar)	$NaAlSi_3O_8$
Potassium feldspar	$KAlSi_3O_8$
Anortite (Ca-feldspar)	$CaAl_2Si_2O_8$
Calcite	$CaCO_3$

The main reason for mineral weathering is that conditions at the earth's surface differ from conditions where they were formed; they are thermodynamically unstable. Many minerals were formed at high pressure and temperature. When, a result of geological transfer, rocks enter new environments, they disintegrate or are converted to other minerals phases, frequently clay minerals. This process is perfectly natural and has been going on for as long as we have had a oxygenated atmosphere with free water on Earth. Most rocks contain several different

minerals with different resistance to weathering. As a result, rock surfaces undergo selective dissolution/weathering, the less resistant minerals disappearing fastest, only leaving a small pit behind, while resistant ones remain. The stone materials worked and used are widely diverse. We distinguish between igneous rocks (formed from molten fluid, magma) sedimentary ones (which have hardened from loose sediments) and metamorphic ones (converted from the first two as a result of increased pressure and temperature). Rock metamorphosis includes, for example:

granite→gneissic, granite→gneiss
sandstone→quartzite
marlstone→limestone→marble

Sedimentary rocks such as limestones or calcareous sandstones are a good deal more prone to dissolution/weathering than a pure silicate rocks like granites or pure sandstones. A granite, however, consists of several different minerals, mainly quartz, potassium feldspar, plagioclase and micas. Because these weather at different rates (table 2), an initially smooth (e.g. glacially worked) surface will in time become rougher; its specific area increases. The result is a continuously accelerating weathering and degradation.

For highly soluble minerals, rain exposure is crucially important. Gypsum crusts frequently form on highly calcareous rocks. These become permanent in rain shadow while surfaces exposed to rain are washed clean, gypsum being highly soluble (table 1).

3. Environmental pollution in relation to stone decay

During the XIX century the expansion of urban conglomerations and the concentration of adjoining industries have resulted in a formidable increase in the local aggressive airborne impurities. The combined action of natural weathering agents and atmospheric pollutants has resulted in a worrying disfiguration and decay of historical buildings and monuments.

In recent decades the acceleration of stone deterioration in urban environments has been attributed to the sharp increase of air pollutants.

To illustrate the extent and significance of air pollution a comprehensive review of atmospheric pollutants that directly or indirectly affect the stone will be given.

Pollutants include almost any natural or artificial matter capable of being airborne (Amoroso & Fassina, 1983). Air pollution is composed of particles, aerosols and gases. Larger particles settle near the source but can be transported by winds over some distance. They are mainly of local importance. Aerosols are liquid or solid particles of small sizes that remain suspended in the air and can be transported over long distances. Gases of course mix thoroughly with air (Rosvall, 1988).

Most of the components usually called pollutants are already present in clean air, their concentration, however, is increased significantly by human activity. Human activity influences the quality of the atmosphere particularly in dense urban areas and near large emissions sources. Most gases, however, are mainly products of industrial activities, including heating and traffic. A major part of the pollution comes from burning processes: domestic heating, industrial heating, power stations and combustion engines. Most fuels, except natural gas, contain sulphur, so sulphur dioxide is one of the main gaseous by-products of combustion. At high temperatures nitrogen and oxygen combine to nitrogen oxides which are therefore also normal by-products. Particles often contain soot from incomplete combustion, tarry material, trace of metals like iron, manganese and vanadium as well as adsorbed SO_2 and water. Particles can also arise from asphalt roads, vehicles tyres, and brakes (Rosvall, 1988). Gaseous components of air pollution are oxide of sulphur and nitrogen, ozone, hydrocarbons and, to a lesser extent, hydrogen sulphide and other inorganic and organic substances. Of primary interest when considering building stones are some acid gaseous pollutants of anthropogenic origin, i.e. sulphur oxides, nitrogen oxides and carbon dioxide which in contact with atmospheric water form acids: some of these are strong acids (e.g. sulphuric and nitric acid) while others are weak (e.g. carbonic acid). It should also be kept in mind that many processes can be speeded up by catalytic action by, for example, metallic air pollutants and that synergistic effects may thus occur. Another important catalytic action in oxidation processes is developed by carbonaceous particles which are generally deposited on sheltered areas and are also responsible for the blackening of stone surfaces.

The effects of acidic pollutants on calcareous stones depend very much on the immediate environment of the stone. If the stone is in an exposed position where it is regularly washed by rain, the reaction products are washed away and the surface of the stone gradually recedes. If, however, the stone is in a relatively sheltered position, the reaction products accumulate and may form a dense black crust on the surface (Price, 1996).

Taken on a world-basis the total mass of pollutants emitted by nature exceeds those emitted by man. According to estimation of Cullis and Hirschler (1979) pollution sources for sulphur dioxide in the Northern Hemisphere surpassed (60% of the atmospheric sulphur was coming from anthropogenic sources) in 70'ies of XX century the contribution ascribed to natural sources. In urban areas the situation is more serious than in rural regions, because sulphur emissions from anthropogenic sources are several orders of magnitude higher than natural sources, and therefore the deterioration of building stones are rapidly increased.

For deterioration and conservation problems it is simplest to regard any constituent of the atmosphere present in trace quantities, whether of human origin or not, which might harm common materials as a pollutant.

First of all we will consider the sources, abundance and chemical transformations of pollutants within the atmosphere including acid rain and successively we will focus our attention on the stone-atmosphere interface.

4. Sources, abundance and chemical transformations of sulphur compounds

The principal sulphur compounds in the atmosphere are hydrogen sulphide, sulphur dioxide and sulphate aerosols and mists. Sulphur compounds are continuously emitted in the atmosphere, transported between the different phases and removed from the atmosphere according to the cycle proposed by Cullis and Hirshler (1979) (Fig. 4).

The main natural sources of atmospheric sulphur are geothermal activity, sea-spray and biogenic decay (over oceans, land and in coastal areas). Particularly in urban and industrial areas, sulphur is also released as a result of human activities,

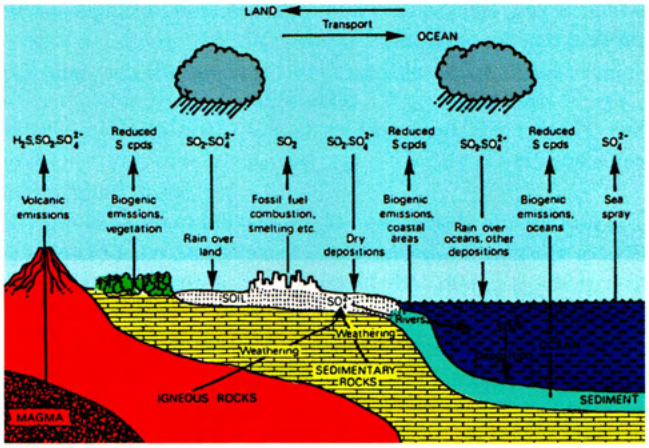

Fig. 4 - The atmospheric sulphur cycle (courtesy of Cullis and Hirschler, 1979).

especially the combustion of fossil fuels and the smelting of sulphur-containing ores. Sulphur is removed from the atmosphere by precipitation processes (involving mainly sulphate) and by dry deposition (principally sulphur dioxide).

Anthropogenic sources of atmospheric sulphur are mainly the combustion of coal and petroleum. Sulphur is present in coal and petroleum as impurity so that during combustion processes it is emitted in the atmosphere, mainly as sulphur dioxide. The sulphur content in coal is generally between 1 and 3%, while in petroleum it is typically between 0.75 and 2.5%. A minor anthropogenic source of sulphur is the smelting of non-ferrous ores.

Throughout the past 100 years sulphur dioxide emitted by human activities has increased markedly (Table 3). In the last decades this increase was even more marked. The contribution of coal has decreased from over 97% in 1880 to under 60% in 1976 (Table 3). In contrast, the contribution of petroleum over the same period has increased from less than 1% to 28%.

Table 3. *Increase in sulphur dioxide emissions over the past years (Tg SO_2 per year).*

	Coal	Petroleum	Non ferrous ores	Others	Total
1880	12.2 (97.4%)	0.07 (0.6%)	0.26		12.5
1910	42.1 (90.5%)	0.7 (1.5%)	3.7		46.5
1930	51.3 (85.5%)	3.1 (5.2%)	4.9	0.66	60.0
1950	66.0 (79.4%)	8.3 (10%)	7.4	1.4	83.1
1960	95.7 (73.6%)	19.9 (15.3%)	12.3	2.2	130.1
1965	102.0 (68.5%)	28.5 (19.1%)	15.7	2.7	148.9
1970	110.4 (64.1%)	39.6 (23%)	18.8	3.2	172.2
1974	115.2 (61.5%)	47.2 (25.2%)	21.1	3.7	187.3
1975	120.8 (64.2%)	43.8 (23.3%)	20.0	3.6	188.2
1976	123.8 (59.7%)	58.3 (28.1%)	21.4	3.7	207.2

Table 4 shows that the Northern Hemisphere generates 94% of the sulphur emissions of anthropogenic origin. This reflects the fact that almost all industrialized societies are located on this half of the globe, which also contains a larger proportion of land area.

Table 4. *Estimates of emissions of atmospheric sulphur (Tg S per year) in the '70 ies.*

Sources	1970 Northern & Southern Hemispheres	1976 Northern & Southern Hemisphere	1976 Northern & Southern	1976 Northern Hemisphere
Natural				
Volcanoes	5	5	3 (60%)	2 (40%)
Sea-spray	44	44	19 (43%)	25 (57%)
Sub total non biogenic	**49**	**49**	**22 (45%)**	**27(55%)**
Biogenic (land)	48	48	32 (67%)	16 (33%)
Biogenic (ocean)	50	50	22 (44%)	28 (56%)
Subtotal biogenic	**98**	**98**	**54 (55%)**	**44 (45%)**
Total natural	**147**	**147**	**76 (52%)**	**71 (48%)**
Anthropogenic	**86**	**104**	**98 (94%)**	**6 (6%)**
Total	**233**	**251**	**17 (69%)**	**77 (31%)**
Anthropogenic (% of natural)	**59**	**71**	**129**	**9**
Anthropogenic (% of total)	**37**	**41**	**56**	**8**

4.1. Sulphur dioxide reactions in the atmosphere: oxidation of sulphur dioxide

Many factors, including temperature, humidity, light intensity, atmospheric transport, and surface characteristics of particulate matter, may influence the atmospheric reactions of sulphur dioxide. Like many other gaseous pollutants, sulphur dioxide undergoes chemical reactions resulting in the formation of particulate matter, which can settles or is scavenged from the atmosphere by rainfall or other processes.

Whatever the processes involved, much of the sulphur dioxide in the atmosphere ultimately is oxidized to sulphuric acid and sulphate salts, particularly ammonium sulphate and ammonium hydrogen sulphate.

The oxidation of atmospheric sulphur dioxide to sulphate may occur either by *homogeneous gas-phase oxidation* or *by heterogeneous oxidation*.

The quantitative understanding of the reaction rates of homogeneous gas-phase oxidation of SO_2 is rather well developed (Cox and Penkett, 1972; Cox, 1974; Eggleton and Cox, 1978; Calvert et al., 1978).

By contrast, the quantification of the rates of the heterogeneous reaction remains rather primitive, despite many studies carried out by numerous scientists in the last decades. This is due in large part to the fact that the heterogeneous formation of sulphate may occur by a large number of separate and distinct paths.

In the aqueous phase (clouds or aerosols) the *heterogeneous oxidation of SO_2* may take place by (Beilke and Gravenhorst, 1978; Hegg and Hobbs, 1978):

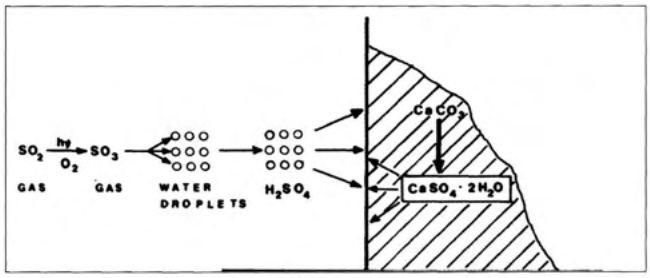

Fig. 5 - Homogeneous gas-phase oxidation of sulphur dioxide in the atmosphere and subsequent attack of H_2SO_4 on stone surface (from Amoroso & Fassina, 1983).

(i) transition-metal catalyzed reaction (Barrie and Georgii, 1976);
(ii) strong oxidizing agents such as O_3 or H_2O_2 (Penkett et al., 1979b; Larson et al., 1978);
(iii) oxygen in the absence of catalysts (Penkett et al, 1979a).

Heterogeneous oxidation of SO_2 adsorbed to solid particles (gas-liquid) or deliquesced salt particles (gas-saline solution) can take place (Chang et al., 1978; Liberti et al., 1978; Barbaray et al., 1977). This process involves the adsorption of SO_2 to solid particles, as a first step, followed by the formation of sulphurous acid in the presence of water, where partial neutralization to sulphites may occur, and subsequent oxidation to sulphuric acid and sulphates takes place.

In a real atmosphere the above oxidation mechanisms may proceed at the same time. The mechanism and the rate of oxidation are difficult to predict because of the complicated nature of these reactions and because of the great influence of many factors: such as humidity, sunlight, temperature and the presence of compounds acting as intermediates or catalysts. The prevalence of one of these mechanisms depends on the particular atmospheric conditions but very probably it is necessary to take into account a complex, synergistic, oxidation of SO_2, by O_2, O_3 and H_2O_2 in the presence of metal catalysts, basic substances, such as ammonia, and organic compounds (Amoroso & Fassina, 1983).

4.1.1. Homogeneous gas-phase oxidation of sulphur dioxide

Homogeneous gas-phase oxidation of atmospheric sulphur dioxide can take place by reactive intermediates, e.g. atoms, free radicals and excited molecular species which are generated photochemically. In polluted air the phenomenon known as photochemical smog involves the atmospheric oxidation of hydrocarbons and related substances and proceeds by a free-radical chain process.

The photochemically generated species which play an important role in the atmospheric oxidation of SO_2 are $\bullet OH$, $HO_2\bullet$ and $RO_2\bullet$ arranged in decreasing order of importance. In urban air, during summer months, the total rate of SO_2 oxidation by $\bullet OH$ and $RO_2\bullet$ species is calculated to be between 2.2 and 6.5% h^{-1}. In an unpolluted atmosphere the calculated SO_2 oxidation rate in winter lies between 0.2 and 1.1% h^{-1} and in summer between 1.4 and 2.2% h^{-1}.

In studies carried out in Western Europe over four years it was found that homogeneous oxidation of SO_2 by free radicals generated photochemically is significant for the conversion of SO_2 to SO_4 in the lower atmosphere.

Simulations of SO_2 oxidation in polluted urban atmospheres suggest that these reactions can lead to removal rates of as high as 5% h^{-1}(Sidebottom et al., 1972). Airborne sulphate thus formed can dissolve in atmospheric water droplets which are subsequently deposited on stone surfaces (Fig. 5). Part of the sulphate in the droplets can be present as sulphuric acid, which is very aggressive on limestone. The main stages in the process may be outlined as follows:

$$\underset{\text{oxidation}}{SO_2(\text{gas}) \rightarrow SO_3(\text{aerosol})} \underset{\text{dissolution}}{\rightarrow H_2SO_4(\text{droplets}) + \text{limestone}} \rightarrow CaSO_4\,2H_2O \quad (18)$$

The probable overestimation of the maximum removal rate for SO_2 of close to 5% h^{-1} indicates that the rate of photo-oxidation of SO_2 is significantly below this value and is probably less important under atmospheric conditions.

According to Beilke (1985) it is now accepted that the most effective homogeneous gas-phase reaction is the reaction of OH radicals with SO_2:

$$SO_2 + \bullet OH \xrightarrow{M} HOSO_2\bullet \xrightarrow{O_2} SO_3 + HO_2\bullet \quad (19)$$

There is substantial evidence that $HOSO_2\bullet$ and SO_3 formed by gas-phase SO_2 oxidation ultimately lead to the generation of sulfuric acid (Cox, 1974). Due to its very low equilibrium partial pressure H_2SO_4 cannot exist in measurable quantities as a trace gas under atmospheric conditions. Sulfuric acid gas may undergo nucleation processes in the course of which mixed aggregates of H_2SO_4 and H_2O molecules are formed. These aggregates may grow by coagulation processes into the accumulation mode. Another reaction pathway is a direct deposition of H_2SO_4 gas onto pre-existing aerosol particles which leads to an enhancement of sulfate content of those particles.

According to Cox and Penkett (1983), maximum SO_2-oxidation rates in central Europe are of order of 2% h^{-1} at full sunlight. In winter the corresponding rates are expected to be slower by a factor of 3 to 5. Daily average rates (averaged over 24 hours) during summer are of the order of 0.5% h^{-1}.

4.1.2. Heterogeneous oxidation of sulphur dioxide in the aqueous phase

In the atmosphere, the oxidation of sulphur dioxide in aqueous systems involves dissolution of SO_2 in water droplets which may be present in clouds or fog, followed by the formation of sulphite (SO_3^{2-}) which is oxidized by air or other oxidants and is partly neutralized by ammonia or metal oxides. Several laboratory investigations reproducing compounds commonly found in the atmosphere have been carried out recently but nearly all these researches differ greatly from the actual atmospheric conditions because of experimental difficulties.

In a real atmosphere a number of mechanisms must be considered simultaneously:

-liquid-phase oxidation of SO_2 by oxygen in the presence of catalysts

-liquid-phase oxidation of SO_2 by oxygen, ozone and hydrogen peroxide in the absence of catalysts.

The contribution of each mechanism to the overall process should be determined under the given atmospheric conditions.

Before describing the liquid-phase oxidation of SO_2 by different agents let us consider the first step of reaction; that is, the dissolution of SO_2 in water droplets.

First step-dissolution of gaseous SO2 in aqueous systems

In a heterogeneous system gaseous SO_2 in the air is in equilibrium with SO_2 dissolved in distilled water according to the following chemical equilibria (Barrie and Georgii, 1976):

Across the gas-liquid interface

$$SO_2\ (gas) + H_2O\ (liquid) \leftrightarrow SO_2 \bullet H_2O\ (liquid) \quad K_H = 1.24\ mol\ l^{-1}\ atm^{-1}\ at\ 25°C \quad (20)$$

physically dissolved SO_2

In the liquid phase

$$SO_2 \bullet H_2O \leftrightarrow H^+ + HSO_3^- \quad K_1 = 1.74\ 10^{-2}\ mol\ l^{-1}\ at\ 25°C \quad (21)$$

$$HSO_3^- \leftrightarrow H^+ + SO_3^{2-} \quad K_2 = 6.24\ 10^{-8}\ mol\ l^{-1}\ at\ 25°C \quad (22)$$

SO_2 exists in solution as physically dissolved SO_2 ($SO_2 \bullet H_2O$), bisulphate ion (HSO_3^-) and sulphite ion (SO_3^{2-}). The chemical equilibria indicate that the concentrations of the dissolved SO_2 species are strongly dependent on pH. The bisulphite ion concentration increases with increasing pH and becomes the most abundant species - over 90% of total dissolved SO_2 - in the pH range of precipitated atmospheric droplets (pH between 3 and 6) (Fig. 6).

Second step-oxidation of SO_2 in the droplet phase

As mentioned before, the liquid-phase oxidation of SO_2 can involve different oxidizing agents in the presence or absence of catalysts. Laboratory and theoretical studies for both the uncatalyzed and catalyzed reactions will be reviewed to evaluate the conditions under which they should be of importance in the atmosphere.

i) Liquid-phase oxidation of SO_2 by oxygen in the presence of catalysts

Traces of metal ions which are commonly found in the atmosphere catalyze the oxidation of SO_2 dissolved in atmospheric droplets. The catalysts are mainly transition metals of the 4th period, especially manganese and iron.

Many laboratory studies of the catalytic oxidation of SO_2 by oxygen in water carried out in the past are not applicable to the interpretation of the oxidation of SO_2 in the atmosphere because the concentrations used in the experiments were several orders of magnitude higher than those expected in atmospheric droplets. In fact, some authors used high concentrations of dissolved SO_2 or high catalyst concentrations, or most commonly both high SO_2 and catalyst concentrations (Hegg and Hobbs, 1978). In addition, most studies were carried out with sodium sulphite solutions, which have a pH range between 8 and 9, while the atmospheric pH range is between 3 and 6. It is more significant to examine studies carried out with concentrations of SO_2 and catalysts similar to those in a real atmosphere. For instance, in dilute solutions such as fog and cloud drops the catalyst concentrations are between 10^{-4} and 10^{-8} M and the SO_2 concentration is less than 1 ppm. Studies carried out using bulk solutions (homogeneous system) (Coughanour and Krause, 1965; Brimblecombe and Spedding, 1974; Beilke et al., 1975) or large droplets (Johnstone and Coughanour, 1958; Barrie and Georgii, 1976) yield valuable information about the reaction mechanism of catalytic oxidation, but, representing a one-phase system they do not adequately simulate the physical characteristics of the two-phase drop-air system. The reaction rate coefficients reported in these studies are not appropriate for modeling the oxidation of SO_2 in aerosols because the rate of reaction was limited by the extent of mixing and diffusion in the liquid or by inter-phase mass transfer. The liquid-phase oxidation of SO_2 in the presence of catalysts has been found to be independent of the concentration of oxygen (at least for pH below 8).

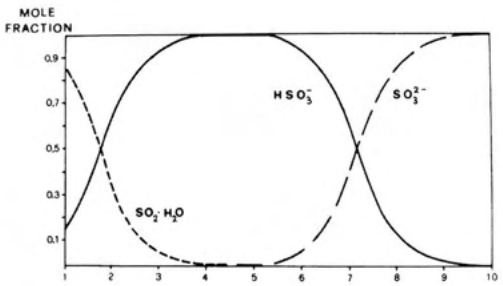

Fig. 6 - Distribution of dissolved SO_2 species versus pH (Courtey of Beilke & Gravenhorst, 1979).

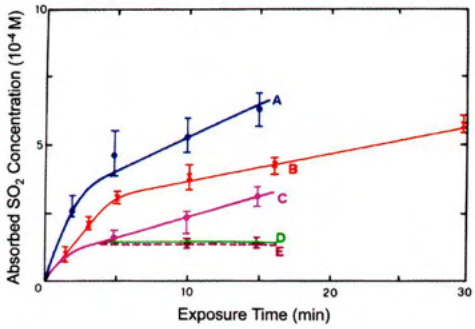

Fig. 7 - Effect of manganese ions on SO_2 absorption by distilled water droplet. Curve A: 10^{-4} M $MnCl_2$ solution and 900 ppb of SO_2; curve B: 10^{-4} M $MnCl_2$ solution and 450 ppb of SO_2; curve C: 10^{-5} M $MnCl_2$ solution and 900 ppb of SO_2; curve D: 10^{-6} M $MnCl_2$ solution and 900 ppb of SO_2; curve E: distilled water and 900 ppb of SO_2. (Courtesy of L.A. Barrie and H.W. Georgii, 1976).

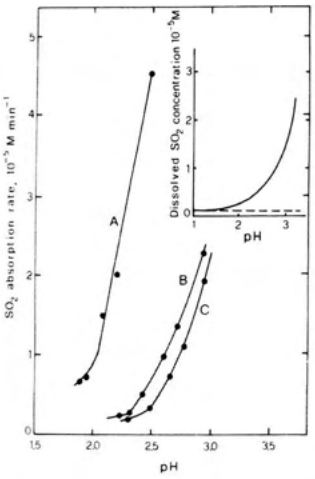

Fig. 8 - SO_2 absorption rate as a function of pH at 25°C and 700 ppb of SO_2. Curve A: 10^{-4} M $MnCl_2$+10^{-4} M $FeCl_3$ solution; curve B: 10^{-4} M $MnCl_2$ solution; curve C: 10^{-4} M $FeCl_3$ solution. Insert: the solubility of SO_2 in distilled water as a function of pH at 25°C and 700 ppb SO_2. (Courtesy of L.A. Barrie and H.W. Georgii, 1976).

Experiments of Barrie and Georgii (1976) carried out by absorbing SO_2 in a dilute $MnCl_2$ solution show that the absorption of SO_2 by distilled water droplets is strongly dependent on manganese chloride concentration (Fig. 7).

Curves D, C and A, measured at constant SO_2 concentration (900 ppb), show that the absorption of SO_2 by distilled water droplets increases with increasing $MnCl_2$ concentrations. During the experiment, the pH of the droplets gradually decreased as a result of oxidation, the sulphur content increased and the rate of SO_2 absorption by the drops decreased and stopped when a pH of about 2 was reached.

The rate of SO_2 absorption as a function of pH for three different solutions is shown in fig.8. In all cases the rate of SO_2 absorption decreased very rapidly until a pH of 2.2 for iron and manganese and a pH of 2 for the mixture. The strong decrease of SO_2 solubility as function of pH causes a reduction of the absorption rate, which is proportional to the amount of SO_2 reactant in the droplet. Below pH 2 there is so little SO_2 dissolved in the droplet that oxidation proceeds very slowly, as can be seen in the insert of fig. 8.

The mixture of $MnCl_2$ and $FeCl_2$ produces an oxidation rate about one order of magnitude greater than that for the same molar concentrations of $MnCl_2$ of $FeCl_2$ alone. This synergistic effect was also observed by other authors (Penkett et al., 1979a; Lenelle, 1975). Other experiments by Barrie and Georgii (1976) showed that the rate of SO_2 absorption by a solution of 10^{-5} M $MnCl_2$ was increased much more by adding a 10^{-5} M $FeCl_2$ solution than by increasing the concentration of $MnCl_2$ to 10^{-4} M. Very probably, iron and manganese interact to promote oxidation. They also found that there was no difference if ferric (Fe^{3+}) or ferrous (Fe^{2+}) ions were added to solutions of Mn^{2+}. It is evident that the catalytic effectiveness of Mn is enhanced by the addition of Fe^{2+} or Fe^{3+}. On decreasing the temperature of manganese solution from 25 °C to 8 °C, the SO_2 absorpion rate decreases five to ten times, while for the iron-manganese solution the decrease is very limited.

Comparison of the uncatalyzed oxidation rate constant obtained by Beilke et al. (1976) (ca. $10^{-3}s^{-1}$ in the pH range 5 to 6 at 25 °C) with the catalyzed oxidation rate constants of Table 5, shows that the presence of catalysts enhances the oxidation rate constants by six to seven orders of magnitude (Barrie and Georgii, 1976). In the real atmosphere catalyst concentration are lower by a factor of at least 10, making the difference much smaller.

Table 5. *Pseudo-first order oxidation rate constants for three solutions at 25 °C.*

Solution	K(s^{-1})
10^{-4} M $MnCl_2$+10^{-4} M $FeCl_2$	140 x 100
10^{-4} M $MnCl_2$	9.40 x100
10^{-4} M $FeCl_2$	5.80 x100

ii) Liquid-phase oxidation of SO_2 by ozone, hydrogen peroxide and oxygen in the absence of catalysts.

SO_2 oxidation by strongly oxidizing agents (H_2O_2, O_3) seems to be by far the most important process of sulfuric acid formation in droplets under atmospheric conditions (Penkett et al., 1979a; Penkett et al., 1979b). Two reactions are of importance:

$$S(IV) + O_3 \rightarrow S(VI) + O_2 \quad (23)$$
$$S(IV) + H_2O_2 \rightarrow S(VI) + H_2O \quad (24)$$

For H_2O_2, measured concentrations in the liquid phase were used (1 to 100 μm). In spite of large uncertainties in the concentrations of H_2O_2 in cloud water (Daum et al., 1983; Römer et al., 1983) this reaction seems to be by far the most important sulfuric acid formation process because the pH of fog and cloud droplets in central Europe is generally measured to be below 5.

Both H_2O_2 and O_3 can have their origin in homogeneous gas-phase reactions. In addition, H_2O_2 could be formed within the droplets by photolytically induced radical reactions. SO_2 oxidation by O_2 in the absence of catalysts contributes only to a negligible extent to sulfuric acid and sulfate production. Penkett et al. (1979b) calculated the rate of sulphate formation in cloud water from oxidation of SO_2 by ozone, hydrogen peroxide and oxygen (Fig. 9). For instance at pH 5 the rates of reaction of SO_2 with oxygen, ozone and hydrogen peroxide are $3.1x10^{-4}$, 0.24 and 2.2 μg/ml-min, respectively, and thus it would take approximately 5 days, 10 min and 2 min respectively, for cloud water to acquire the minimum amount of sulphate found at Harwell (2.2 μg/ml). From fig. 5 it is evident that the influence of oxygen on the SO_2 oxidation is negligible compared with ozone or hydrogen peroxide. In fact the rate of sulphate formation from SO_2 and ozone or hydrogen peroxide is 1000 times faster than that with oxygen.

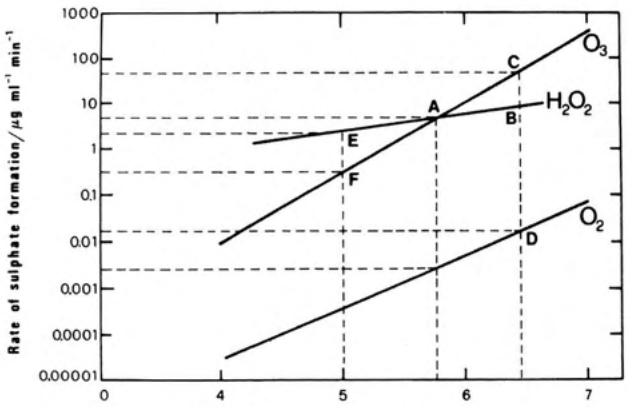

Fig. 9 - Rate of sulphate formation in cloud water from oxidation of SO_2 by oxygen, ozone and hydrogen peroxide (Courtesy of Penkett et al. 1979b).

Summarizing considerations on heterogeneous oxidation of SO_2 in droplet phase (rainwater, urban fogs , urban-polluted and rural-unpolluted areas).

In rural atmospheres uncatalyzed O_2 and O_3 oxidation and mixed-salt-catalyzed oxidation rates are all competitive for 3<pH<4.5 (Figs. 10 a,b). The mixed-salt reaction is probably not very important, since the metal concentrations is too low catalytic activity. Therefore the uncatalyzed oxidation of SO_2 by ozone may dominate over the O_2 reaction in clouds with a pH 5 at typical O_3 concentrations of 50 ppb. For example, according to Penkett et al (1979) the rate of reaction of SO_2 with ozone is 1000 times faster than that with oxygen and the time required is ten minutes for O_3 and 5 days for O_2.

For urban fog uncatalyzed O_2 and O_3 oxidation and iron-catalyzed oxidation are competitive for 3<pH<5 if the concentration of manganese in solution is negligible. The manganese-catalyzed oxidation proposed by Barrie and Georgii is 10 times faster and the mixed-salt-catalyzed SO_2 oxidation is 1000 times faster than the uncatalyzed oxidation or the iron-catalyzed oxidation.

As a consequence of these considerations the following generalizations can be made:

a) the catalytic effectiveness of a heavy metal of the 4th period depends on the pH. It is generally accepted that at low pH (below pH ca. 2.0). SO_2 is practically not oxidized despite the presence of metal catalysts; however atmospheric NH_3 keeps the pH of the water droplets sufficiently high during the formation of H_2SO_4 from SO_2, to allow SO_2 oxidation to continue (Blokker, 1978),

b) the catalytic effectiveness of a metal ion depends on the presence of other metals. A synergistic effect was observed by many authors (Barrie and Georgii, 1976; Penkett et al, 1979; Lenelle, 1975). Manganese catalyzes SO_2 oxidation in dilute aqueous solutions at a rate that depends on catalyst concentration, pH and temperature (Barrie and Georgii, 1976). The SO_2 oxidation rate obtained with a mixture of manganese and iron is higher (by one order of magnitude) than the sum of the rates with each ion present separately (Barrie and Georgii, 1976),

c) ammonia does not play a catalytic role in SO_2 oxidation but has the function of increasing the pH of the droplet, permitting further dissolution and oxidation of SO_2,

d) the rate-limiting steps in the overall oxidation in cloud and fog droplets are oxidation mechanisms, and not diffusion and transport processes in the gas and liquid phases.

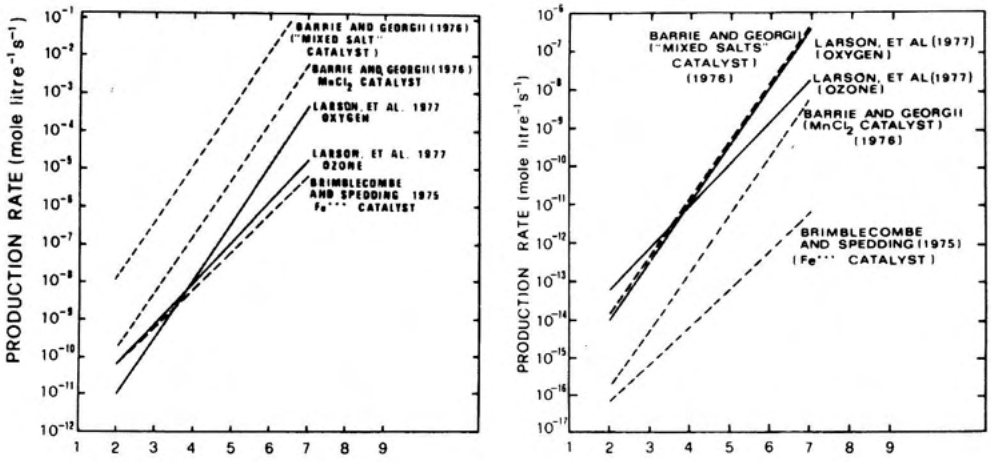

Fig. 10 - Rates of sulphate formation in the droplet phase as a function of the pH of the solution for uncatalyzed (solid lines) and catalyzed (dashed lines) laboratory measurements. (a) Polluted atmosphere, (b) unpolluted atmosphere (Courtesy of D.A. Hegg & P.V. Hobbs, 1978).

4.1.3. HETEROGENEOUS OXIDATION OF SO_2 ADSORBED TO SOLID PARTICLES

As previously said this process involves the adsorption of SO_2 to solid particles, as a first step, followed by the formation of sulphurous acid in the presence of water, where partial neutralization to sulphites may occur, and subsequent oxidation to sulphuric acid and sulphates takes place.

In the size range of 0.6 to 2 μm, sulphite is a significant fraction of the particulate sulphur fraction of aerosols. Measurements carried out in Pasadena, California, showed that the SO_4^{2-}/SO_3^{2-} ratio varied between 0.32 and 1,27 in the range 0.6-2 μm, and between 1.49 and 1.85 in the range 2-5 μm.

Laboratory experiments carried out on the adsorption and desorption of SO_2 on different types of dust showed that considerable adsorption takes place, but that conversion with urban atmospheric dusts to sulphate at room temperature did not occur to a significant extent.

5. Sources, abundance and chemical transformations of nitrogen oxides

Among the nitrogen oxides only nitrous oxide (N_2O), nitric oxide (NO) and nitrogen dioxide (NO_2) are found in minimal amounts in the atmosphere. Nitrous oxide (N_2O) is present in the natural environment in relatively large amounts, between 0.25 and 0.60 ppm with an average of 0.32 ppm (Beilke and Gravenhorst, 1979).

Nitrous oxide is produced in soils and waters by microbiological processes (Pratt et al., 1977). Other processes are agricultural activities (e.g. the use of nitrogen fertilizer) (Crutzen, 1979), combustion processes (Weiss and Craig, 1976; Pierotti and Rasmussen, 1976) and photochemical reactions in the stratosphere (the upper atmosphere). Nitric oxide (NO) is produced by combustion engines (Pratt et al., 1977), by lightning (Noxon, 1978; Chameides et al., 1977), and in soils (Galbally and Roy, 1978).

Under the high temperature in a combustion zone nitrogen reacts with oxygen in the air to form nitric oxide

$$N_2 + O_2 \leftrightarrow 2NO \qquad (25)$$

In exhaust gases some of the nitric oxide is oxidized to nitrogen dioxide according to the reaction

$$2\,NO + O_2 \rightarrow 2\,NO_2 \qquad (26)$$

Nitrogen dioxide (NO_2) varies from 0.5 to 10% of the NO present (EPA, 1971). In air, nitric oxide is oxidized more rapidly by atmospheric ozone (Cadle and Allen, 1970; Abeles et al., 1971) and photochemical processes than by oxygen to form nitrogen dioxide (NO):

$$NO + O_3 \rightarrow NO_2 + O_2 \qquad (27)$$

In polluted areas NO concentrations are typically 100 ppb, while in remote marine areas they are below 0.01 ppb. For NO_2 the range is between some 100 ppb in urban areas and less than 0.01 ppb in remote marine areas (Beilke and Gravenhorst, 1979). In the industrial mid-latitudes more than 70-80% of the total NOx ($NO+NO_2$) is emitted by anthropogenic sources.

5.1 HOMOGENEOUS GAS-PHASE OXIDATION OF NOX

According to Beilke (1985) in contrast to the generation of sulphuric acid, nitric acid can also be formed with an appreciable rate by homogeneous gas-phase oxidation at night. During daytime NO is oxidized in a reversible reaction by ozone (O_3), peroxyradicals ($RO_2\bullet$) and oxygen (O_2) to NO_2. For most conditions occurring in the lower atmosphere, oxidation of NO by O_3 is by far the most effective process for NO_2 generation:

$$NO + O_3 \rightarrow NO_2 + O_2 \qquad (28)$$

Nitrogen dioxide is further oxidized by •OH radicals to nitric acid gas (HNO_3 gas):

$$NO_2 + \bullet OH \xrightarrow{M} HNO_3 \qquad (29)$$

The rate-determining reaction (29) for nitric acid formation proceeds ca. ten times faster than the corresponding reaction of SO_2.

Maximum NO_2 oxidation rates in central Europe are of the order of 20% h^{-1} at full sunlight. Daily average rates in summer are of the order of 5% h^{-1}. In winter, due to the reduced photochemical activity, the oxidation rates will decrease by a factor of 3 to 5.

During nighttime, Beilke (1985) stated that the most effective gas-phase reaction is an irreversible oxidation of NO by O_3 to NO_2, followed by oxidation of NO_2 through reaction with O_3. The NO_3 radical formed by the NO_2-O_3 reaction is in equilibrium with N_2O_5 (Platt et al., 1981; Platt and Perner, 1985).

During daytime NO_3 concentrations are very low due to photolytical destruction and reaction with NO. According to Calvert and Stockwell (1983) NO_3 reacts with aldehydes

(formaldehyde, acetaldehyde) at an appreciable rate, leading to the formation of HNO_3 gas:

$$\bullet NO_3 + CH_2O \rightarrow HNO_3 + HC\bullet O \quad (30)$$
$$\bullet NO_3 + CH_3CHO \rightarrow HNO_3 + CH_3C\bullet O \quad (31)$$

Another possible reaction path for HNO_3 generation is a gas-phase reaction of N_2O_5 with water vapor (Calvert and Stockwell, 1983; Tuazon et al., 1983):

$$N_2O_5 + H_2O \rightarrow 2\ HNO_3 \quad (32)$$

According to these authors the estimated night-time HNO_3 formation rates via reaction (32) are rather high and should be included in long-range-transport computer models. It must be noted, however, that most chemists favor a heterogeneous reaction of N_2O_5 with atmospheric droplets or other wet surfaces rather than a homogeneous gas phase reaction. There is good evidence that the overall HNO_3 formation rate at night through a homogeneous gas-phase reaction proceeds as fast as the formation of this strong acid during daytime. In contrast to sulfuric acid, nitric acid can exist as a trace gas at measurable quantities due to its equilibrium vapor pressure.

5.2 Heterogeneous oxidation of NOx in aqueous phase

From a thermodynamic point of view, nitric acid concentrations in cloud water can be quite high at representative ambient gas concentrations of NOx. The generation of nitrous acid (HNO_2) and nitric acid (HNO_3) through absorption of NO and NO_2 by cloud droplets takes place according to the following reactions:

$$NO\ (g) \leftrightarrow NO\ (aq) \quad (33)$$
$$NO_2\ (g) \leftrightarrow NO_2\ (aq) \quad (34)$$
$$NO\ (aq) + NO_2\ (aq) + H_2O \leftrightarrow 2\ H^+ + 2NO_2^- \quad (36)$$
$$2\ NO_2\ (aq) + H_2O \leftrightarrow 2\ H^+ + NO_3^- + NO_2^- \quad (37)$$

For a cloud liquid water content of 1 g/m^3 and typical atmospheric mixing ratios of NOx, removal rates of this gas of the order of 10^{-4} to 10^{-5} % h^{-1} were calculated. These slow NOx removal rates are in contrast to the high nitrate content measured in cloud and rainwater. Therefore other mechanisms than the absorption and the chemical reaction of NOx can occur. There are strong indications that the rather high nitrate content of cloud and rainwater has its origin in an absorption of gaseous HNO_3, NO_3/N_2O_5 and/or in an incorporation of nitrate-containing aerosols. These processes are, however, difficult to quantify.

As far as NOx is concerned, a possible reaction of importance for acid generation is the conversion of NOx on aerosol surfaces (Kessler and Platt, 1984). Most likely this is a surface-catalyzed reaction.

The presence in the atmosphere of such a strong acid as nitric acid could be very harmful for building stones. Until a few years ago nitric acid had never been detected in the urban atmosphere. Adequate methods were developed only recently and nitric acid was found to be present at a concentration of 100 ppb. The role of nitric acid in the deterioration processes of stone is not very clear. In regions, where free nitric acid occurs its prolonged action will affect the building material by converting calcium carbonate of limestone into the more soluble calcium nitrate:

$$CaCO_3 + 2\ HNO_3 \rightarrow Ca(NO_3)_2 + H_2O + CO_2 \quad (38)$$

The low amounts of nitrates generally found in black crusts of monuments seem to indicate that nitric acid attack is not very important. It is difficult to establish whether nitric acid is attacking the limestone because the high solubility of calcium nitrate causes its removal from the surface and consequently this may be the reason for the low amounts found.

Another effect of nitrogen oxides is their catalytic role in the oxidation of sulphur dioxide. The oxidation of SO_2 on carbon particles in dry and humid air is enhanced by the presence of trace quantities of gaseous NO_2 (Cofer et al., 1980). Experiments carried out at 20 ppm SO_2 and 5 ppm NO_2 gave positive results. At low concentrations (0.1 ppm SO_2 and 0.02 ppm NO_2) the SO_2 oxidation in water droplets was also investigated. The chemistry of reactions between SO_2 and HNO_2 in acqueous solution is very complicated. The total stoichiometry of the reaction can be expressed as:

$$2\ HNO_2 + 2\ HSO_3^- \rightarrow N_2O + 2\ HSO_4^- + H_2O \quad (39)$$

It is difficult to estimate the rate of the nitrite–SO_2 reaction in the atmosphere, because little is known about the amount of nitrous acid in the troposphere, and nitrite levels in the aqueous phase of aerosols have not been reported (Martin and Damschen, 1981).

During a smog episode, acid concentration in the gas phase was estimated to be 2-4 ppb. This specie is present only at night. The rate of sulphate production due to nitrite-SO_2 reaction is small and is roughly comparable with the rate of sulphate production from the iron-catalyzed aqueous molecular oxygen reaction at pH 2 and 3 if the iron concentration is 10^{-8} mol/l. It is much slower than rates estimated when significant amounts of ozone or hydrogen peroxide are present.

In conclusion: in the absence of light, NOx is unlikely to be a fast oxidant for SO_2 in atmosphere aerosols; and the new source of N_2O is likely to be small in comparison with biological sources. However, this process can be very efficient close to sources and in heavily polluted urban areas where the concentrations of NO/NO_2 are high and the pH of aqueous droplet is low. Studies carried out by Johansson, Lindquist and Mangio (1988) showed that in humid atmospheres NO catalytically oxidizes SO_2, increasing the corrosion rate of calcareous stone. The net reaction is:

$$NO_2 + SO_2 + H_2O \rightarrow H_2SO_4 + NO \quad (40)$$

6. Sources, abundance and chemical transformations of carbon dioxide

Carbon dioxide is a normal constituent of the atmosphere and it is not generally considered an air pollutant. Carbon dioxide is by far the most abundant of all atmospheric gases generated by human activities. Although the total amount of gas emitted in this way represents only about 2% of that produced by Nature, it nevertheless interferes with the natural balance (Hirschler, 1981).

While natural emission and removal processes are more or less in equilibrium, man's contribution has been more than sufficient to explain the increases in atmospheric concentrations of carbon dioxide in recent years (Robinson and Robbins, 1972).

Carbon dioxide is one of the end products of the combustion of organic materials and is therefore produced from coal, petroleum, gas, the incineration of waste (urban and rural) and forest fires. The total amount of carbon dioxide emitted directly by man into the atmosphere has been increasing markedly ever since fossil fuels started to be employed as a source of energy. Another anthropogenic source of carbon dioxide is deforestation and the burning of woods, especially in tropical areas of South America and South Asia. Deforestation will indirectly induce the release into the atmosphere, as carbon dioxide, of a proportion of the carbon previously contained in the humus and protected by the trees from atmospheric erosion and now undergoing oxidation.

As far as the influence of carbon dioxide on the decay of limestone is concerned, we must take into account the dissolution of carbon dioxide in water and the dissociation equilibrium constant. Carbon dioxide in air dissolves in rain water, forming in the first instance $CO_2 \bullet H_2O$ (physically dissolved carbon dioxide) as a result of the reaction:

$$CO_2\ (gas) + H_2O\ (liquid) \leftrightarrow CO_2 \bullet H_2O\ (liquid) \quad K_H = 3.4\ 10^{-2}\ mol\ l^{-1}\ atm^{-1}\ at\ 25°C \quad physically\ dissolved\ CO_2 \qquad (41)$$

The equilibrium concentration of carbon dioxide in air is proportional only to undissociated carbonic acid $CO_2 \bullet H_2O$ (physically dissolved CO_2) and is described by Henry's law CO_2 H_2O in water dissociates into HCO_3^- and CO_3^{2-}through the following equations

$$CO_2 \bullet H_2O \leftrightarrow H^+ + HCO_3^- \quad K_1 = 4.45\ 10^{-6}\ mol\ l^{-1}\ at\ 25°C \qquad (42)$$

$$HCO_3^- \leftrightarrow H^+ + CO_3^{2-} \quad K_2 = 4.68\ 10^{-11}\ mol\ l^{-1}\ at\ 25°C \qquad (43)$$

The above system of equations requires that with an increase of $[CO_2]$ and correspondingly $[CO_2\ H_2O]$, $[H^+]$ must also increase. According to Junge (1963) the effect of this change in the pH value is that the increase of total carbon dioxide content in rainwater is only 1% for every 12.5% increase in $[CO_2]$ atmosphere.

The weakly acidic solution formed by dissolution of CO_2 in rainwater dissolves the calcium carbonate in limestone, marble, lime mortars and plasters because it forms the much more soluble calcium bicarbonate:

$$CaCO_3 + CO_2 \bullet H_2O \leftrightarrow Ca(HCO_3)_2 \qquad (44)$$

insoluble soluble

Calcium bicarbonate is about a hundred times more soluble (1.1 g/l) than calcium carbonate (1.4×10^{-2} g/l).

The amount of calcium bicarbonate dissolved by rainwater depends mainly on the water temperature and the partial pressure of CO_2 in the air, which in turn is related to physically dissolved CO_2 H_2O calculated as ppm CO_2 by Henry's law. An increase of water temperature causes a decrease in the amount of carbon dioxide dissolved. For instance, at temperatures near the freezing point, water dissolves nearly twice as much carbon dioxide as at 25 °C (Winkler, 1970) (Fig. 11).

A rise of the partial pressure of atmospheric CO_2 increases the amount of CO_2 dissolved. For instance, if the water temperature is 25 °C, in rural areas, where the average concentration is about 300 ppm, the CO_2 dissolved is about 0.25 ppm. In urban areas, when peak concentrations of 3000 ppm are reached, the CO_2 dissolved becomes ten times higher, 2,5 ppm. The increasing CO_2 concentration in the air forces the equilibria in eqns (41)-(44) towards the right-hand side, thus producing more calcium carbonates.

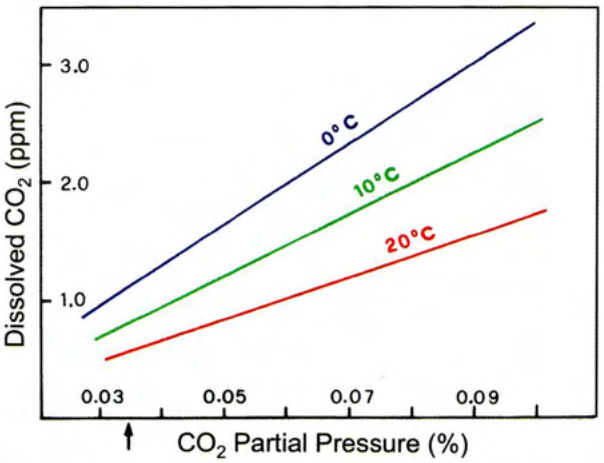

Fig. 11 - Equilibria of dissolved CO_2 ($CO_2 \bullet H_2O$) in rainwater at different atmospheric CO_2 partial pressures according to Henry's laws.

In rural areas the average concentrations of CO_2 vary by only about 10% from the normal 320 ppm, while in urban and industrial areas it may be two or three times as high as in rural regions. It is interesting to note that in urban areas peak concentrations of up to 3000 ppm can occur. This concentration, which is ten times as high as in rural areas, influences the pH of rainwater. According to Amoroso and Fassina (1983) in a polluted area with 3000 ppm of CO_2 the pH of rainwater shows a decrease of 0.5 compared to the pH values in a rural atmosphere. As a consequence, the attack by CO_2 proceeds more rapidly in urban areas than in rural ones, because both the CO_2 concentration may be increased and the pH decreased. The leaching action of slightly acidic rainwater on marble surfaces is frequently associated with the re-crystallization of water-soluble calcium bicarbonate. In fact, the carbon dioxide dissolved tends to evaporate when the solution temperature increases, and as a consequence the following reaction takes place:

$$Ca(HCO_3)_2 \rightarrow CO_2 + H_2O + CaCO_3\ \text{re-crystallization} \qquad (45)$$

The calcite formed is characterized by large crystals and a porous structure whereas the original calcite is microcrystalline and non-porous. The increase of porosity caused by the re-crystallization of water-soluble calcium bicarbonate is harmful to the stone (Fig. 12), since it allows acidic solutions containing sulphuric acid and soluble salts (e.g. sulphates and chlorides) to penetrate more deeply into the stone and to accelerate the decay.

7. Dry and wet surface deposition of air pollutants

In order to clarify the mechanism of stone decay and the formation of black crusts it is very important to focus our attention on the stone-atmosphere interface to estimate the deposition of sulphur dioxide and small (potentially acidic) particles, the condensation of water vapour at the surface and on already deposited particles, and the characteristics of pollutants delivered in precipitation.

The deposition of pollutants is only the final step in a chain of events of transportation and transformation between sources and receptor surface. Wet and dry deposition are important sinks for reactive atmospheric gases and aerosol particles.

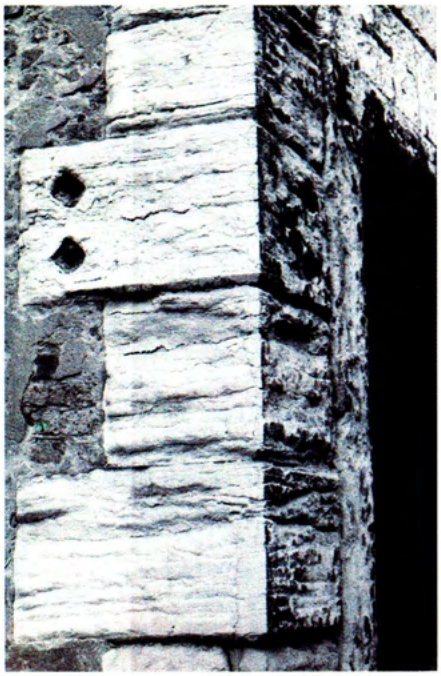

Fig. 12 - White deposit of calcite re-crystallization on Istrian stone according to reaction (44).

DRY DEPOSITION

The dry deposition phenomenon results from the transfer of airborne pollutants from the atmosphere to a surface in the absence of rain. Airborne pollutants are gases and/or particles, including aerosols. Dry deposition generally originates from nearby sources and is therefore called short-range deposition (Torraca, 1988). Gases are the most important contributors to dry deposition, and their arrival at the surface is governed by molecular diffusion or turbulence (Charola, 2002). Gases can react with both the surface of the stone and with aerosols particles. The mechanism of transport depends mainly on the particle size: gravitational setting increases with particle diameter, while Brownian diffusion decreases. For sub-micron particles, less than 0.1 μm, approaching the surface the most important motion is molecular diffusion. Smaller particles show a high deposition rate, since the surface act as a perfect sink, and once they enter in contact with the surface they are efficiently retained by Van der Waals forces (Hicks, 1981). The deposition velocity shows a minimum for particles between 0.1 μm and 1 μm where none of the transport processes are efficient. As the deposition velocity depends on the particle size, estimation of the rate of deposition of aerosols requires a knowledge of the size distribution. From several measurements of the size distribution of sulphate aerosols most of the sulphate mass falls between 0.1 and 1μm diameter, although a minor fraction is sometimes found between 1 and 20 μm (sea-salt sulphate or windblown soil sulphate). Submicron sulphate particles are mainly formed by oxidation of sulphur dioxide.

The entire process of particle deposition will be modified if particles carry an electrostatic charge. These processes are strongly affected by the presence of temperature gradients, leading to thermophoresis, or water vapour gradients, if condensation or evaporation processes are operative.

The actual deposition process of gases and larger particles occurs in the boundary layer at the surface and is influenced by the nature of the substrate, its conditions and its microenvironment. As is to be expected, the concentration of gases and particles affects their rate of deposition which in turn is enhanced by an increase in concentration of the pollutants, air turbulence, roughness and heterogeneity of the receiving surface, chemical affinity of the surface for the pollutants, and surface moisture. The concentration of pollutants found within a few centimetres above the receiving surface, together with microclimatic conditions, is most relevant to an analysis of dry deposition. These conditions may be fairly different from those in the nearby atmosphere, at a distance of some meters from the receiving surface, which in turn vary with daily and seasonal climatic changes (Charola, 2002).

Measurement of deposition rates of sulphur compounds for outdoor exposure with the aim to evaluate the reactivity of different stones were carried out by Furlan and Girardet from 1983 to 1996. In laboratory experiments, using an atmospheric simulation chamber, Ausset et al. (1996) and Girardet et al. (1996) exposed some trials of Jaumont limestone and Berne molasse 'naked' and covered with either fly-ash or soot particles. After one year of chamber exposure in real-life conditions (125 ppb SO_2, 50 ppb NO_2 at 79% RH and 13°C) the results obtained were comparable to those observed in site tests. The presence of fly-ash appeared to increase the SO_2 deposition velocity for the Jaumont stone after the first half-year, but not that for the Berne molasse. Soot particles appeared to protect the surface from sulphur deposition. When sulphur does deposit it penetrates into the stone to a depth of about 0.8 mm (Girardet & Furlan, 1996). Further laboratory studies with a simulation chamber confirmed that the reactivity of stones to SO_2 can increase significantly with increase relative humidity and to a smaller extent with increased temperature, but that these increases depend on the nature of the stone (Charola, 2002). In urban areas, calcareous sandstone trials, exposed in protected areas, show a double amount of sulphur deposited with respect to that found in rural areas (1.6 and 0.8 gS m^{-2} a^{-1} respectively). They also confirm that the flux of the *dry sulphur* was far more important than the flux of *humid sulphur*. Atmospheric aggressiveness of various cities in Europe and Washington, DC, based on the annual deposition of *dry sulphur* on a calcareous sandstone was evaluated on the basis of sulphur surface deposit that built up over the year (Fig. 13).

As outdoor atmosphere found in Milan revealed to be the most aggressive the researchers decided to use Milan as a pilot site for the exposure of ten different stones ranging from siliceous ones, like gneiss and sandstones, to calcareous sandstones, limestones and marble (Furlan & Girardet, 1992). The amount of *dry sulphur* deposited varied significantly among the stones and three categories of reactivity could be established. The most reactive stones proved to be the molasses, limestones and calcareous sandstone; Carrara

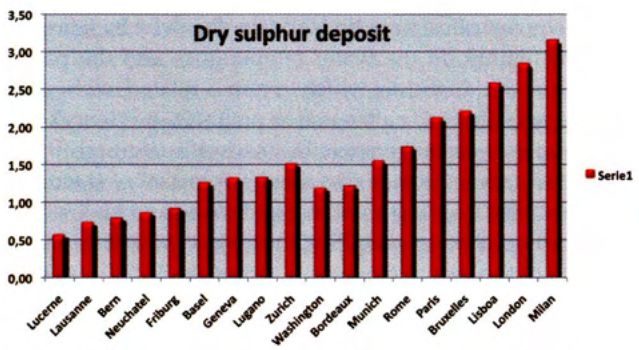

Fig. 13 - Atmospheric aggressiveness of various cities in Europe and Washington, DC, based on the annual deposition of dry sulphur on a calcareous sandstone, molasse de Villalord (modified from Furlan & Girardet, 1992)..

marble was found to be less reactive, while the siliceous sandstones and gneiss could be classified as non reactive (Charola, 2002).

Other studies such as National Materials Exposure Programme were based on the long-term field exposure studies of different materials at different sites, following the weight change procedure used for metal corrosion with the aim of obtaining damage functions. From these long-term exposure studies it was found that the developed damage functions did show some predictive capabilities (Butlin et al., 1992). Although the model over predicted stone loss for a very wet Scottish site (Webb et al., 1992). For Portland stone, which was used as reference material for this study, the natural solubility of the stone in water was the dominant term in the model.

WET DEPOSITION

Wet deposition comprises the incorporation of trace substances in cloud droplets (*rainout*) and removal by falling precipitation (*washout*). Wet deposition generally plays a lesser role than dry deposition in the acid deterioration of stone. However, in rural areas with low pollution wet deposition can be just as important as dry deposition (Furlan & Girardet 1983a). Incorporation of certain trace substances into cloud and precipitation elements leads to the formation of acid precipitation, a phenomenon observed for a long time in Northern Europe, United States and Canada. Commonly acid precipitation is called *acid rain*. The pollutants incorporated in wet deposition are usually produced by distant sources (Torraca, 1988). The efficiency of rainout and washout depend on the intensity of precipitation, the drop-size spectrum, the pH value and the temperature of the drops, as well as on the vertical distribution of trace substances in the atmosphere (Amoroso & Fassina, 1983). Chemical conversion processes within the atmosphere play an important role in precipitation chemistry and in wet deposition. If the acidity of precipitation is below the equilibria pH value with atmospheric CO_2 (a pH value of 5.6) the lowered pH is usually found to be associated with elevated sulphate and nitrate levels. In general, small rain drops <1mm diameter are saturated with SO_2, while larger droplets do not achieve saturation.

In remote areas, where gaseous emissions are at low concentrations, aerosol formation has its most direct influence on precipitation chemistry. According to Glover et al. (1980) the composition of precipitation would be expected to correlate relatively well with that of aerosols and the scavenging processes occur mainly by rainout, since washout of aerosols is comparatively inefficient.

In less remote areas, with higher levels of gaseous emissions, the composition of precipitation is controlled by a mixture of rainout and washout involving both local and distant sources, and the composition of precipitation does not necessarily correlate with local gaseous and aerosol compositions.

SCAVENGING OF SULPHUR DIOXIDE AND SULPHATE AEROSOL

Scavenging processes related to sulphur dioxide and sulphate aerosol should be considered to clarify their major importance in the wet removal of sulphur from the atmosphere. The general wet removal processes can be decomposed into several individual pathways. These include direct sulphur dioxide scavenging, direct sulphate scavenging, and combined scavenging and chemical reaction as described in Fig. 14.

Pathway 1-3: absorption of SO_2 (gas) in hydrometeors by precipitation.

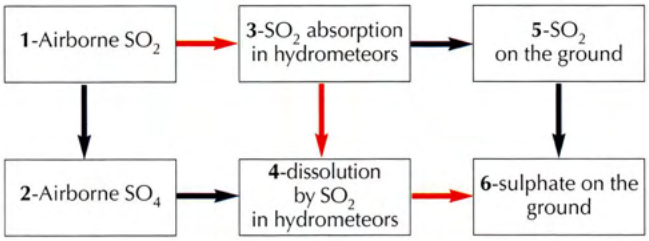

Fig. 14 - Schematic representation of atmospheric interactions between SO_2 gas, sulphate aerosol and precipitation.

At point 3 there are two possibilities:
Patway 3-5: deposition of SO_2 (aq) to the surface
Patway 3-4: oxidation of SO_2 (aq) to SO_4 and subsequent deposition to the surface
Pathway 1-2 oxidation of SO_2 (gas) to SO_4 (aerosol)
Pathway 2-4 dissolution of SO_4 in hydrometeors
Pathway 4-6 deposition of SO_4 (aq) onto the surface

Reverse processes are of course possible, and can contribute significantly to scavenging behavior under appropriate atmospheric conditions (Hales 1972). For example sulphate aerosol may be regenerated via droplet evaporation (step 4-2) and sulphate aerosol may be resuspended from the surface to the atmosphere (step 6-4). This is the case of sulphate input into the atmosphere by evaporation of sea-spray.

According to Hales pathways 1-2-4-6 and 1-3-4-6 constitute the most important wet removal routes for atmospheric sulphur, since the preponderance of anthropogenic sulphur emitted to the atmosphere is released in the form of sulphur dioxide, and since the majority of sulphur in precipitation occurs as sulphate.

Although the acidity of rain is often blamed for stone decay, particularly in calcareous stone, it has been found that acidity frequently decreases as rainfall continues. Wet deposition will affect only the exposed surfaces of a building, while dry deposition can affect all surfaces of a building.

Rain is more efficient in scavenging airborne pollutants than cloud droplets. Acidity in rainfall, as well as turbulent runoff, will attack the stone surface. But a strong rain will also wash away accumulated dry deposit from the stone surface (Charola, 2002).

The efficiency of acid rain depends on the droplet formation and the intensity of precipitation as well as on the vertical distribution of trace substances in the atmosphere. It would appear that the main source of sulphate in rainwater results from the oxidation of anthropogenically generated SO_2 via the entrapment of sulphate-containing particles and sulphuric acid-containing aerosols.

The main mechanisms that will define the final chemical composition of a droplet are:

a) *nucleation scavenging*, i.e. when suspended hygrophilic particles absorb water forming embryos that then grow into droplets;

b) *scavenging of gas and particles* during the evolution of the dispersed system, i.e. when droplets absorb gases and capture suspended particles;

c) *chemical reactions in the dispersed liquid phase*, i.e. reactions that occur within the droplet once pollutants have entered them; (see heterogeneous SO_2 oxidation). (see heterogeneous SO_2 oxidation);

d) *microphysical evolution of the system* as a function of atmospheric conditions .i.e. Stephan flow, thermophoresis, aerodynamic impaction, Brownian movements, coalescence and other such microphysical mechanisms that alter the chemical composition or the dimension of the droplets (Camuffo, 1990).

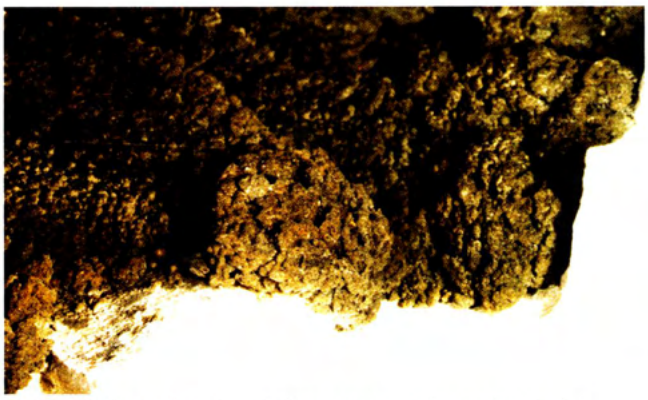

Fig. 15 - Black crust formation. Carbonaceous particles are deposited on gypsum crystals.

Fig. 16 - Running water on exposed areas slowly dissolves limestone.

Fig. 17 - Black crusts occurred on sheltered areas from direct rainfall and are mainly formed by gypsum.

Fig. 18 - Marked contrast between washed and unwashed areas is observed.

Condensation processes

In daytime, fixation of pollutants on stone will be greatest on the coolest parts of the exposed surface. The presence of temperature (thermophoresis) and humidity gradients (Stefan effect) near the surface can also promote or hinder the deposition of particles. When condensation is taking place at the stone surface, both particles and gas fluxes will be increased. The nearby air is cleaned as most pollutants pass into the liquid, including liquid and solid particles suspended in air (acid gases, aerosol dust and carbonaceous particles).

The presence of moisture at the stone surface is very dangerous, mainly for two reasons:

a) the pollutant concentration is very high since the amount of water involved is small;

b) these highly concentrated pollutant solutions and their reaction products can remain for a long time on the surface because they are not washed away.

Wetting phase. During the wetting phase of the surface, impinging particles will have a greater chance of adhering, and soluble trace gases will be more readily captured. In the wetting process the acid solutions penetrate into the stone up to a depth which depends on the porosity of the stone and on the amount of solution available per unit of surface.

The chemical nature of the surface is important; if reaction rates with deposited pollutants are rapid, then surfaces can act as nearly perfect sinks. When sulphur dioxide, sulphuric acid and airborne particulates enter into contact with a calcite surface, sulphation takes place and some of the solid is dissolved.

Drying phase. Wetting is followed by drying in the warm hours. Water evaporates and crystals of the soluble salts such as gypsum are deposited on the stone surface. Repeated wet and dry cycles cause water absorbed into the stone to open up channels in the crust as it penetrates down the calcite grains (Lewin & Rock, 1979). During the drying phase, dissolved matter is re-deposited both within the channels and on the surface. Gypsum crystals are repeatedly etched and eroded by water due to their partial high solubility. The gypsum crust incorporates the airborne soot, tar and fly-ash particles; it thickens but remains porous (Fig. 15). Consequently, the crust can thicken indefinitely, albeit at a decreasing rate.

Marking Patterns

The characteristic marking patterns frequently observed on building surfaces caused by local concentrations of runoff are mainly associated with the facade geometry. The surface geometry is related to the building's facade, such as: structural columns, spandrel beams, window heads and sills, cornices, canopies, external shading devices and drip devices (Robinson and Baker, 1975).

On limestone building surfaces rainwater acidified by carbon dioxide and sulphur dioxide from the air slowly dissolves the limestone, enabling dirt to be washed away (fig. 16). In places sheltered from the rain, dirt can accumulate as incoherent stratification or as an incoherent powder which sticks to the stone, or lastly as incrustations strongly bound to the surface with calcium carbonate and calcium sulphate (fig. 17). These different situations result in a marked contrast between washed and unwashed areas (fig. 18). Deposition or rain is usually several times greater at the top and sides of an exposed face of a building than over the remainder of the walls, due to deflection of air and rain.

Water deposited on building surfaces will tend to be absorbed at first by porous materials, but as soon as the rate