UNIVERSITY OF ROME "TOR VERGATA"
DEPARTMENT OF ELECTRONIC ENGINEERING

DOCTORATE IN

SENSORIAL AND LEARNING SYSTEMS ENGINEERING

TITLE OF THE THESIS

SAMPLING AND SENSING STRATEGIES FOR NOVEL APPLICATIONS WITH AN ARTIFICIAL OLFACTORY SYSTEM

2003-2004

Candidate

Giorgio Pennazza

Supervisors

Prof. Arnaldo D'Amico

Prof. Corrado Di Natale

UNIVERSITÀ DEGLI STUDI DI ROMA "TOR VERGATA"
DIPARTIMENTO DI INGEGNERIA ELETTRONICA

DOTTORATO IN

INGEGNERIA DEI SISTEMI SENSORIALI
E DI APPRENDIMENTO

TITOLO DELLA TESI

STRATEGIE DI SAMPLING E SENSING PER NUOVE APPLICAZIONI DI UN SISTEMA OLFATTIVO ARTIFICIALE

2003-2004

Dottorando

Giorgio Pennazza

Relatori

Prof. Arnaldo D'Amico

Prof. Corrado Di Natale

'La pazienza è ciò che nell'uomo
più assomiglia al procedimento
che la natura usa nelle sue creazioni.'

Honoré De Balzac

'Patience, in a man, is
the most similar thing to the processes
followed by nature in its creation'

A Nonna Emma

Acknowledgments.

First of all, I would like to thank to my supervisors Prof. Arnaldo D'Amico and Prof. Corrado Di Natale for this opportunity. Their excellent expertise has given me the answers to learn more, their passion for science has provoked the questions to try a step over.

I would say thank to Prof. Riccardo Marino, Doctorate Coordinator, for his kindly support during this period.

"Travelling with great expectations is better than arriving"
Robert Louis Stevenson

This quote is a good expression to tell about these three years. A good travel asks for good guidance and good fellows, and I have many of them to whom I have to show my gratitude.

I would like to thank again to Prof. Corrado Di Natale, for all his precious teachings, for the example of tireless work, for he trusted in me, for the sharing of his knowledge, for his friendship. Very often we only need someone telling us we can do it, I found that!

I would like to thank again to Prof. Arnaldo D'Amico, because his guidance was the basis of which avail to face these three years, his suggestions were the shortest way to benefit from hard work.

I would like to thank Prof. Roberto Paolesse, for his precious support and his prompt helpfulness in chemical subjects..

I would say thank to Dr. Eugenio Martinelli, a friend, a precious colleague, a mine of ideas, an excellent 'fellow traveller'.

I would like to express my gratitude to Dr. Antonella Macagnano, a friend on whom I can rely. Thanks for her suggestions in designing complex experimental sets, thanks for her support.

I would like to thank to Prof. Claudio Roscioni, for his perseverance and for his precious expertise.

I would like to express my gratitude to Emiliano Zampetti, because his ability in producing electronics is faster than his thinking of doing it. Thanks again for his friendship. I would express

my gratitude to Simone Pantalei, because his ability in producing software and 3D-models was fundamental. Thanks again for his friendship.

I would like to thank to Marco Santonico, a mine of solutions for application designing. His friendship was a decisive support.

I would say thank to Marco Valdoni, because he is a good philosopher, a desirable quality in a friend.

I would say thank to Nina, a smart fellow, a mine of cheerfulness.

Thanks to Alexandro, Fabio, Emiliano, Andrea, Emanuela, Christian, Alessia, Mara, Massimiliano. Thanks to all my further colleagues at University, which have not been mentioned namely, for their diverse contributions on the way to accomplish this thesis.

I would like to say thanks to Michele Zarlenga, Giovanni Petrocco and Luigi Maita, for their helpfulness, their expertise and their patience.

This morning I woke up to write these acknowledgements and to take my thesis to the bookbinder's. I felt something that can be synthesized in the following quotation:

'I awoke this morning with devout thanksgiving for my friends, the old and the new' R.W.Emerson

I can't evade this thanksgiving, so let me spend another page.

Thank you, mum, dad, Simona and nonna Emma for you trusted in me. Thanks a lot for our serenity.

Thank you Paolo, for your presence and your encouragements!

Thank you Francesco, Loredana and Ludovica, for your support.

Thank you Riccardo, for your friendship.

Thank you Aurora, Emanuela, Francesca, Michela, Enrico, Luca, Marco, Paolo, my gratitude for you is greater than a simple list of name.

Too many people? Probably I'm lucky, so thanks to God for His Grace!

Preface

Sometimes we have problems which need answers; so the answer is the solution. This is a good way to solve problems. Other times, we have problems which need questions; so the question is the solution. This is a good way to research.

Everyday, everyone of us uses the five senses, very often without consciousness at all of using them, but simply exploiting the results. The results consist in an interpretation of the real world around us.

Olfaction is probably one of the senses humans have lost potentiality to exploit during evolution, but odour is a very important source of data, because of the power of synthesis of a lot of interactions and parameters we are not able to collect, read and elaborate at the same time.

An existing technology, the Electronic Nose of the Sensors and Microsystems Group of the University of Rome 'Tor Vergata', is the objective of this Thesis. The aim is to ask fundamental questions about it, and by studying the possible solutions, reach the whole system improvement.

Index

CHAPTER 4

Lung cancer diagnosis

CHAPTER 5

Environmental Monitoring

CHAPTER 6

Static sampling

Appendix

Multivariate Data Analysis

Motivations

The great potentialities of artificial olfactory systems can be exploited in many fields. This multidisciplinary range of applications asks for designing of dedicated sampling systems depending on the different scope.

Although the working principle and the sensitive material are the basis on which these devices are developed, the sampling, within the measure chain, is a fundamental step to optimize the whole system performances.

According to these considerations the best way for designing "ad hoc" experimental set is to specialize a specific sampling procedure for each application.

In this thesis three different studies are considered: medical and environmental applications, and odour sampling in static conditions.

OLFACTION AND ELECTRONIC NOSE

'Però gli bastarono quei pochi secondi di scuro assoluto e di silenzio per fargli percepire un odore non usuale che, ne era certo, aveva sentito un'altra volta[…]. Siccome gli veniva naturale, sin da quando era nicareddru, di dare un colore ad ogni odore che lo colpiva, si disse che questo era di colore verde scuro.'

Da 'Il cane di terracotta' di Andrea Camilleri

'But in a black darkness he took few seconds of absolute silence to percieve an unusual odour that, he was sure, he had already smelt. […] Considering that, since he was a child, he used to associate a colour to each particular odour which fscineted him, he told to himself the colour of this odour was dark green.'

From 'Il cane di terracotta' by Andrea Camilleri

1

1.1 Olfaction and the other senses

All living beings have the ability to interact with the environment by means of suitable interfaces commonly called senses. Traditionally human senses are five: sight, hearing, touch, smell and taste.

Senses are part of a complex mechanism, which involves different cognitive processes, giving rise to perceptions, namely the conscience of the external world.

The task of the perceptual system is to gather information of the environment and use this information, usually aiming at optimizing some vital functions.

The crucial points of human sensing are: the nature of the signal, the content of information, the understanding of processes elaborating registered information.

The senses operating processes can be divided in four steps: object sensing and filtering, signal transduction, perception, object definition and classification by mean of language and concepts.

In the last two steps, perception and definition-classification, the source of information which is in the object of senses operation discloses its informative content. In the first, perception, the signal is received and recognized as the command for other signals, something like a sort of instinctive reaction. This is common to animals and human beings. In the second case, the definition-classification level, the sensed environment is interpreted, classified, recorded, expressed and compared among humans.

Complexity of translation from perception to interpretation is due to differences in filtering processes, in definition of shared classifications schemes and lack of objectivity in case of particular detected quantities. Considering the nature of the sensed quantities, senses can be divided into two groups: those detecting physical quantities and those detecting chemical quantities. Light, sound and pressure are physical quantities that can be defined in simple terms and referenced against standards [1]. On the other hand, chemical interfaces even if described in literature, present some aspects of physiological working principle that are still unclear [2]. The functional organization of the olfactory system is

similar to other sensory systems but, in this case, the sensory input is provided by molecules.

As far as the nature of the input is concerned, these two groups bring to to a a first important distinction. As a consequence of the different signal source, the decisive point is the relationship between signal and information: in the case of physical senses this relationship is almost linear, but this is not the case with the chemical ones.

Actually, the perception of a physical or chemical quantity by the five senses, can be considered a linear operation, if the relationship between the perceived object and an existent reference is linear.

The functional relationship between the structure of odorant molecules and their olfactory impact is poorly understood. This is mainly due to the high degree of signal processing that takes place in the olfactory system, so that compounds of similar chemical structure can give completely different olfactory responses [3].

The following example is useful to understand: the description of a landscape to a person who has never seen it; he knows the meaning of words like tree, lake, hill and so on, and it is able to imagine a picture not so different from the real place. On the other hand, the description of odour and taste to a person who doesn't know it; he knows the meaning of words like tomato, bacon, and so on, but it is very difficult to imagine the odour and taste produced by their interaction, because the recipe is a list of possible odour but not the odour of the final food.

There is an additional constraint, that not all individuals respond to olfactory signals in the same manner. These differences can be genetics as well historical due to our adaptive response to certain olfactory events, and different living creatures experience completely different odours [1].

At this point we have to remark an important difference in the use of chemical information between living organisms, and about real utility of this kind of information; indeed, while the information from the physical senses can be adequately elaborated, verbally expressed, memorized and communicated, chemical information, coming from nose and tongue, are surrounded by

vagueness and this is the poor description and memorization capacity in reporting olfactory and tasting experience. Chemical information is of primary importance for most animals; for many of them, indeed, chemistry is the only way to communicate with nature, on the contrary, for human beings, evolution has enhanced the physical interfaces, living little detail of the chemical ones, if we exclude unconscious acquisition and side behaviours [4].

For this reason, in many practical applications, where a quantitative and objective recording of odors is required, this task cannot be realized satisfactorily by any individual natural nose. An olfactory system really capable to make a synthesis of 'odor knowledge' (data recording, fusion and treatment) and giving objective, meaningful, appropriate information for a qualitative or quantitative evaluation is decisive in many applications (see table 1.1.1).

Field	Application	Information	Employ	Existing technology requirements
Food	Fruits quality and ripeness, beverage, products certification	qualitative and quantitative	On and off line	Non destructive analysis, objective aroma evaluation, consumer satisfaction and expectation studies
Medicine	Illness diagnosis, pharmaceutical product evaluation	qualitative and quantitative	On and off line	Prevention, non invasive analysis
Environmental monitoring	Water, air and land pollution monitoring	qualitative and quantitative	On and off line	Malodor identification, odor source discrimination
Process monitoring	Quality control in industrial production	qualitative and quantitative	On line	Non destructive analysis, overall quality control
Automotive	Hazardous compounds monitoring, passenger safety	qualitative and quantitative	On and off line	Odor monitoring
Explosive detection	Security	qualitative and quantitative	On line	Sensitivity enhancement, avoid human life risk

Table 1.1.1: field of application of an artificial olfactory system

To summarize the different needs of all presented fields of application, the real goal is to reach the ability to produce and manage a chemical portrait of odours of specific gas volumes and of chemical ensembles in liquid volumes, as

useful knowledge vehicles to be used for the construction of paradigms oriented to the chemical information storage [5], to the general purpose control and adaptive control optimization, to data comparison purposes and to remote odour sensing and actuation.

A panel test has a real perception of odor but lack of objectivity; at the same time technology trying to mimick olfaction very often spends its efforts towards an objective and precise list of compounds loosing the reality of odor perception. In the middle, probably, we can found a solution. Until now there are two main directions to approach olfaction mimicking: GC-MS and chemical sensors (so-called electronic noses).

Especially in food industries, the characterization of odors is performed by a panel of well-trained persons. Drawbacks of this approach concern for example the high costs, the risk of subjective responses from the involved 'human sensors', the large statistical errors because of generally small numbers of really comparable human sniff tests [1].

The other approaches concern the use of individual chemical or biochemical sensors and the use of instruments from analytical chemistry (usually expensive and complex). Drawbacks of both approaches are based on the fact, that the experimental results are not directly linked with an odor characterization. For a food aroma more than 1500 different individual GC-MS peaks can be recorded; the problem is that they cannot be correlated unequivocally with the odor space of human odor perception. There are several reasons for this: the number of peaks is large and contains fractions of molecules which are not separated, mass to charge ratio is not directly related to odors or chemical structures, and so on.

Chemical sensor arrays, dubbed as electronic noses are probably an answer to these problems. In analogy with natural olfaction an array of non-specific, cross-sensitive sensors is utilized together with a suitable data processing method.

An important effort to spend for the future of this technology is the real comprehension of the olfaction mechanism and the choice of the right working principles to match to mimick these processes. In few words we need to understand better sensing, transducing and signal processing in natural olfaction to rebuilt an artificial chain obtaining the same results.

1.2 Natural olfaction, an overview

Trying to mimick nature is a difficult challenge with an ambitious goal; the real understanding of the natural processes is the difficult but most obvious way to engage to realize this task.

In the case of olfaction, it is very important to define the real meaning of words like odors and odorant, to study the complex biological processes of olfactive signal perception and transmission, and trying to individuate the elaboration activities taking place in the brain for odor recognition and memorization. To avoid a long dissertation not useful for our task, we have focused on the following crucial points.

Odors and odorants

Odor sensations result from the interaction of odorants with specialized neurons called receptors arranged the olfactory epithelium at the top of the nasal cavity (fig. 1.2.1) [6]; odorants are volatile, hydrophobic compounds whose molecular weights may reaches 300 daltons [7].

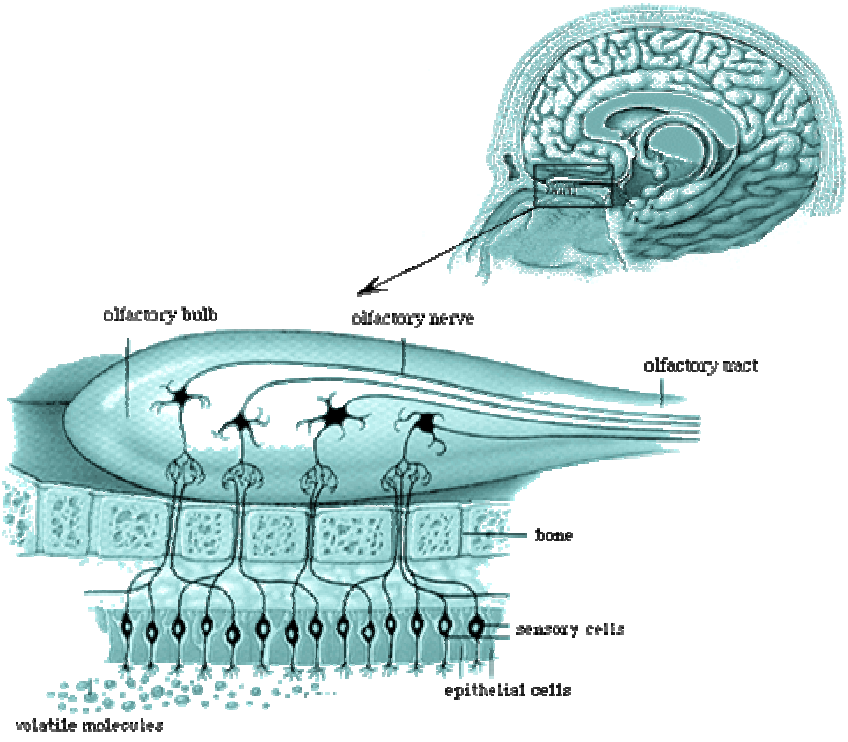


Fig. 1.2.1: olfactory epithelium in the top of the nasal cavity

Signal and sensitivity

Interaction of odorants with olfactory receptors produces signals that are transmitted through the olfactory bulb to the brain [6].

Humans have several hundred distinct genes that encode olfactory receptor proteins and rodents have upwards of 500 to 1000 separate genes, that is, as much as 1% of the genome [12,13].

The sense of smell is a remarkably sensitive system able to respond to very low concentrations of chemicals. It is estimated that only 2% of the volatile compounds available in a single sniff will reach the olfactory receptors, but this is sufficient to perceive an odor [8]. The sensitivity of the smell is given by the human detection threshold. Many compounds can be detected by the largest part of individuals at concentrations in lower than parts-per-billion (ppb, one molecul over 10^9) and for some compounds in the parts-per-trillion (ppt, one molecule over 10^{12}) range.

Mixture and odor real meaning

Most real odor sensations are produced by mixtures of hundreds of odorants rather than by a single compound. Humans have limited capacity to identify single odorants in mixtures with three to four components being maximum [9]. *(power of synthesis, lack of objectivity and need to use a parallel system to know the recipe)*

Descriptors

From a verbose point of view odors are characterized either by general descriptors, or by their source. The number of distinct odors is enormous, and a skilled perfume chemist can identify from 8000 to 10000 different substances on the basis of their odor [10,11].

Odor classification: two main methods.

1) Adjective Descriptors

2) Chemical Properties

1) Classification systems based on adjective descriptors have been historically used to organize the thousands of different odors in a limited number of groups. Modern odor classification methods are based on an extensive array of adjective descriptors selected for their relevance for specific applications [6]. Descriptive classification methods can be general (e.g. for the broad range of odors encountered in everyday life) or specific (e.g. relevant particular applications in the food or fragrance industry).

Classification schemes for odor quality are beset, however, by a variety of limitations: label such as pleasant, delightful, disgusting, and revolting are common associations with odors, and this subjective evaluations can influence the choice of descriptors of odor quality; emotional responses; individual differences in the actual perception of odor based on genetic differences [14]; differences in the use of odor descriptors even among trained panelists; the vocabulary of most languages lacks words that describe the full range of odor sensations.

For this reason, measures of similarity rather than adjective descriptors have been used to quantify odor quality by arranging odor sensations in multidimensional spaces [6].

2) Although much progress has been made in our knowledge of olfactory physiology and biochemistry, the fundamental relationship between odor quality and molecular properties is still poorly understood. Even slight alterations in the chemical structure of an odorants can induce profound changes in odor quality. Current structure-activity models in olfaction are, for the most part, simplye collections of disparate facts with no unifying theme.

There are several reasons for the lack of progress in classifying odors on the basis of chemical properties: first, it is not yet possible to model odorant-receptor interactions because the three-dimensional protein structures of the receptor sites are not known. Second, there are vast numbers of agonist types (thousands of odorant structures) as well as thousands of different odor sensations. Third, identical molecules may activate different receptor types depending on the orientation of the molecule at the receptor. A fourth problem is that there are not

standard methods for quantifying odor quality for use in structure-activity studies.

Brief history of structure-activity studies of olfaction

In spite of these limitations, a variety of attempts have been made to relate odors to physicochemical parameters; among them it is worth to mention the following.

Amoore proposed that the shape and size of molecule are the physicochemical parameters that determine odor quality [15].

Wright indicated that it is inappropriate to represent complex 3D molecular shape by an index consisting of a single number because many different 3D profiles could give the same molecular shape index [16]. Wright suggested that the mechanism for stimulation of olfactory receptors is low-energy molecular vibrations, and that molecules with similar vibrational frequencies patterns should have similar odor quality.

Dravnieks and Laffort suggested that four factors related to intermolecular interaction forces (an apolar factor, a proton receptor factor, an electron factor and a proton donor factor) could predict both quantitative and qualitative odor discrimination in human beings [17].

Molecular parameters and odor threshold

In addition to odor quality, attempts have been made to determine the relationship between odor threshold and molecular parameters. Considered variables include among the others molecular weight, cross-sectional area, adsorption constants at oil-water interface, hydrophobicity, molar volume, satured vapour pressure, polarizability, hydrogen bonding, air/water partition coefficients.

Although it is possible to develop techniques that weight a series of parameters to predict odor quality, this is of little practical use in understanding

the physiological basis of odor. A more complete understanding of structure-activity relationships in olfaction will occur when the molecular structure of the odorant receptor is brought into the model along with the structure of the odorant.

Physiology and Anatomy of Olfaction

***General Structure*:**

The complex sequence of the biological mechanisms and operations taking place in the olfaction can be reported in a flow chart where odorant molecules are input and odors are the output.

This scheme 'reduces' the physiology of natural olfaction to a series of steps of a chain. For each step, it is possible to point out intermediate inputs and outputs, in order to provide a paradigmatic sequence of mechanical or electrical operations. Looking at the scheme the synergy of many different scientific disciplines appears, this suggests the interdisciplinary character of an artificial design of the olfactory system. Moreover it is easier to illustrate the different efforts spent in various research fields to improve the different parts of this chain, because the real state of the art of the electronic nose, is the state of the art of chemistry, physics, electronic, data analysis, applied to the problem solving of the the difficult matter of nature mimicking. Figure 1.2.2 shows the flow chart of natural olfactory system. In the next section the analogous scheme for an artificial olfactory system will be discussed.

The first step of the flow chart is **sampling**. To understand the importance of odor capture and delivery we have to recall again the difference between physical and chemical senses. The functional organization of the olfactory system is similar to other senses. In this case, the input is provided by molecules. This is important in the translation of this system by natural to artificial, and it is the great part of the work presented in this thesis.

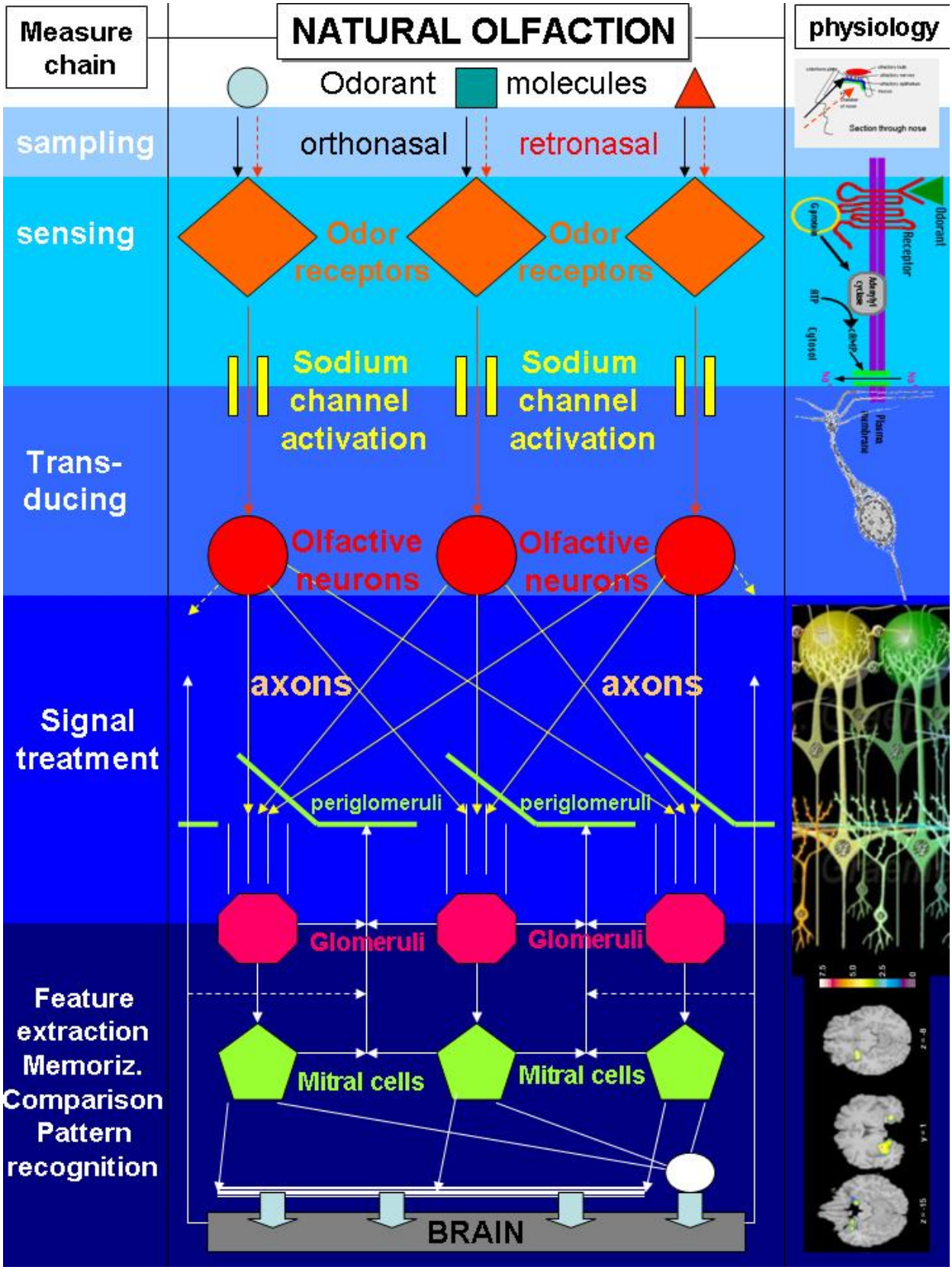


Fig. 1.2.2: 'Flow chart' of the natural olfaction system

In the human nose, odorant molecules do not have immediate access to olfactory receptors, and have to traverse airways covered with the extensive sorptive surface of the respiratory epithelium to reach the olfactory tissue (Hornung & Mozell, 1981).

Olfactory sensation occurs when flow of odorous air passes through the olfactory airway in the upper posterior portion of the nasal cavity [18].

There are two way odorants can reach the nasal cavity: via orthonasal transport through the nostrils (e.g. sniffing), or via retronasal transport from the oral cavity (e.g. chewing food).

Several variables can influence the convective-diffusion transport of odorants from ambient air to the mucus lining of the nasal cavity, such as anatomy of the nose, velocity field in the airways, diffusivity of odorants in air and mucus, solubility of odorants in mucus, and the thickness of the mucus layer [18].

Finally, the odorous molecules reach the roof of each nostril, where there is a region called nasal mucosa, and where **sensing** takes place.

The nasal mucosa consists in the sensory epithelium , covered by mucus. The area of this olfactory region is 5 cm^2 in humans and 25 cm^2 in cats.

The epithelium contains the odorant receptors cells which are the 'bridges' between the sensing and the signal transduction and elaboration part of olfaction. Actually, these are bipolar neurones which possess cilia projected into the mucus to interact with the inspired air, and axons projected to the olfactory bulb to deliver the signal to elaborate.

The receptor cells in humans are about 10 millions, they possess a terminal enlargement (a 'knob') that projects above the epithetial surface, from which extends about 8-20 olfactory cilias that contains the olfactory receptors. In addition some proteins in the olfactory mucus bind odorants and they have been termed the odorant binding proteins. The proteins on the ciliary membranes of olfactory neurons, are the receptors in charge to recognize odorants. Odorants dissolve in the aqueous/lipid environment of the mucus and then bind to odorant binding proteins (OBP).

(OBP) facilitate the transfer of lipophilic ligands (odorants) across the mucus layer to the receptors, and also increase the concentration of the odorants in the layer, with respect to air.

The large family of odorant receptors may be encoded by as many as 1000 different genes. This is a huge amount and accounts for about 2% of the human genome. In humans, however, most are inactive pseudogenes and only around 350 code for functional receptors. Receptor proteins are members of a well known receptor family called 7-transmembrane domain G-protein coupled receptors.

The binding of an odorant molecule gives rise to a chain of biochemical events that results in a stimuli reaching the brain. The main steps of this chain are the following.

After the binding of the odorant is activated the G Protein is coupled to the receptor on its cytoplasmatic side. This releases adenylyl cyclase, enzyme which is embedded in the plasma membrane of the cilia. Adenylyn cyclase catalyzes the conversion of ATP to the 'second messenger' cyclic AMP (cAMP) in the cytosol; cAmp opens up ligand-gated sodium channels for the facilitated diffusion of Na+ into the cell; the influx of Na+ reduces the potential across the plasma membrane. If the depolarization reaches threshold, it generates an action potential. This action potential is conduced back along the olfactory nerve to the brain.

This action started from axon, which begins to transmit the odor received. The perception depends on the simultaneous, cooperative activity of millions of neurons spread throughout expanses of the cortex. Such global activity can be understood only if one adopts a macroscopic view alongside the microscopic one. There is an analogy to this approach in music. To appreciate the beauty in a choral piece, it is not enough to listen to the individual singers sequentially. One must hear the performers together, as they modulate their voices and timing in response to one another [19].

Since the molecules enter the epithelium decoding processes starts. The number of activated receptors indicates the intensity of the stimulus, and their location in the nose conveys the nature of the scent. So, we can say that scent is expressed by a spatial pattern of receptor activity, which is transmitted to the bulb.

The synthesis of the message is contained in the input pattern which is analyzed by the bulb. This message is then transmitted to the olfactory cortex. From there, new signals are sent to many parts of the brain, in particular in an area called the entorhinal cortex, where the signals are combined with those from other sensory systems. The result is a meaning-laden perception, that is unique to each individual.

Such scheme, simple in general, leaves a number of unanswered questions. For istance, how does the the brain distinguish one scent from all others that accompany it? How does the brain achieve what is called generalization-over-equivalent receptors? The real point is that every neuron in the bulb participates in generating each olfactory perception. In other words, the salient information about the stimulus is carried in some distinctive pattern of bulbwide activity, not in a small subset of feature-detecting neurones excited by a particular odor.

Axons of the bipolar olfactory sensory neurons fasciculate together and start the 'path of information' through the olfactory connections. The main nets of this connections consist of mitral cells, glomeruli, periglomerular cells, granule cells.

Mitral cells are the principal neurons in the olfactory bulb; their number in the bulb of an adult human is about 50000. They receive inputs from the olfactory receptors by mean of the glomerulus, connected to the cell by primary apical dendrites.

Glomeruli are roughly spherical bundles of dendritic processes, some 25 mitral cells may send their primary dendrites to a single glomerulus. From the other side the information continue its way to the brain, with other axons starting from the mitral cells and merging together to form the lateral olfactory tract. They possess colaterals, involved in negative feedback and positive feed-forward. Periglomerular cells are involved in lateral inhibition at the level of the glomeruli.

Granule cells are inhibitory interneurons. They receive both contra and ipsi lateral input.

Because individual olfactory sensory neurons can respond to multiple odorants, it follows that the pattern across multiple glomeruli provides the basis for discrimination of olfactory quality. The distinct spatial patterns of glomerular

activation by specific odorants can be visualized using optical imaging techniques. Application of positron emission tomography (PET) and magnetic resonance imaging (MRI) give the possibility to have a deeper insight in brain involvement in olfactory perception [20]. For example it is demonstrated that sniffing and smelling engage separate subsystems in the human olfactory cortex; both passive perception or familiar odorants, and active judgement of odor familiarity seem to involve semantic circuits located outside the olfactory core region. Anyway, when interpreting olfactory imaging data, it is important to consider the ability of odorants to elicit immediate associations and judgements of stimulus characteristics. This condition makes it difficult to disentangle the effects of the chemical stimulus from the effects of immediate associations with this stimulus.

1.3 Electronic Nose, state of the art

In this section a review of Gas Sensors Array will be presented. According to the International Union of Pure and Applied Chemistry (IUPAC) nomenclature, '*a chemical sensor* is a device that transform chemical information, ranging from the concentration of a specific sample component to total composition analysis, into an analytically useful signal'.

Can be a Gas sensor array named electronic nose? Of course they are gas sensors, so their goal is gas presence identification and quantification, but is this odor perception? Odor perception is an action whose object is odor. But behind this action there are series of different operations and in the meaning of the word 'odor' there are many chemical concepts. So this considerations show odor perception to be a complex sensorial process.

Sensorial processes are naturally organized in a series of steps consisting in complex path which leads the interaction with the external world to become a sensation in the core of human brain. This transduction is called perception.

In the previous section we illustrated the fundamental steps to unroll this chain. Now we can exploit this sort of flow chart to create a measure chain. Artificial olfaction consists in the translation of each of these single steps in a chemical, electronics, and mathematical processes in order to measure instead of perceive.

So we have to define the conditions to identify a gas sensor array as an artificial olfactory system, and the only way is the study of all the steps of the measure chain.

The following 'rule' is a good approach to discuss the similarity between each artificial function compared with the corresponding in the natural olfaction: a GAS SENSOR ARRAY is something like an artificial olfactory system if it is a non selective and enough sensitive SENSOR ARRAY of large dimension capable to respond to mixture of GAS. In few words, to quote the definition by Gardner and Bartlett, electronic nose is "an instrument, which comprises an array of electronic chemical sensors with partial specificity and an appropriate pattern-recognition system, capable of recognising simple or complex odours" [21]. Discussing this

definition Mielle (et. Al.) observed that such an analytical instrument is obviously electronic, but not nose', and actually the biological sense of smell is still far superior over today's artificial odour recognition systems.

This consideration is important to understand the need to reduce a complex natural process to a flow chart, because the real goal is the 'informative content' of the final result: odor, a synthesis of the analyte properties. For this reason, although the sensing material is less sensitive than natural odor receptors, the main goal is to tailor its selectivity and to investigate the potentiality of it sensitivity. Although transducing mechanism is less efficient than sodium channel activation and axons transmission, the main objective consists in a deeper understanding of its principle, to better control the process.

There are three decisive steps designing artificial olfactory systems: the sensing material, which must 'catch' the odorant molecules, the transducer, which transform the chemical information in a physical one, and the basic device, which translate this physical quantity variation in an electronic signal.

The chemical interactive material is the open door between the environment and the basic device.

The basic device is the 'stair' from the physical quantity 'floor' to the electrical signal transmission one. The electrical signal is the starting point of the mesurand properties extraction and elaboration.

The figure 1.3.1 illustrates a flow chart for an artificial olfaction system design.

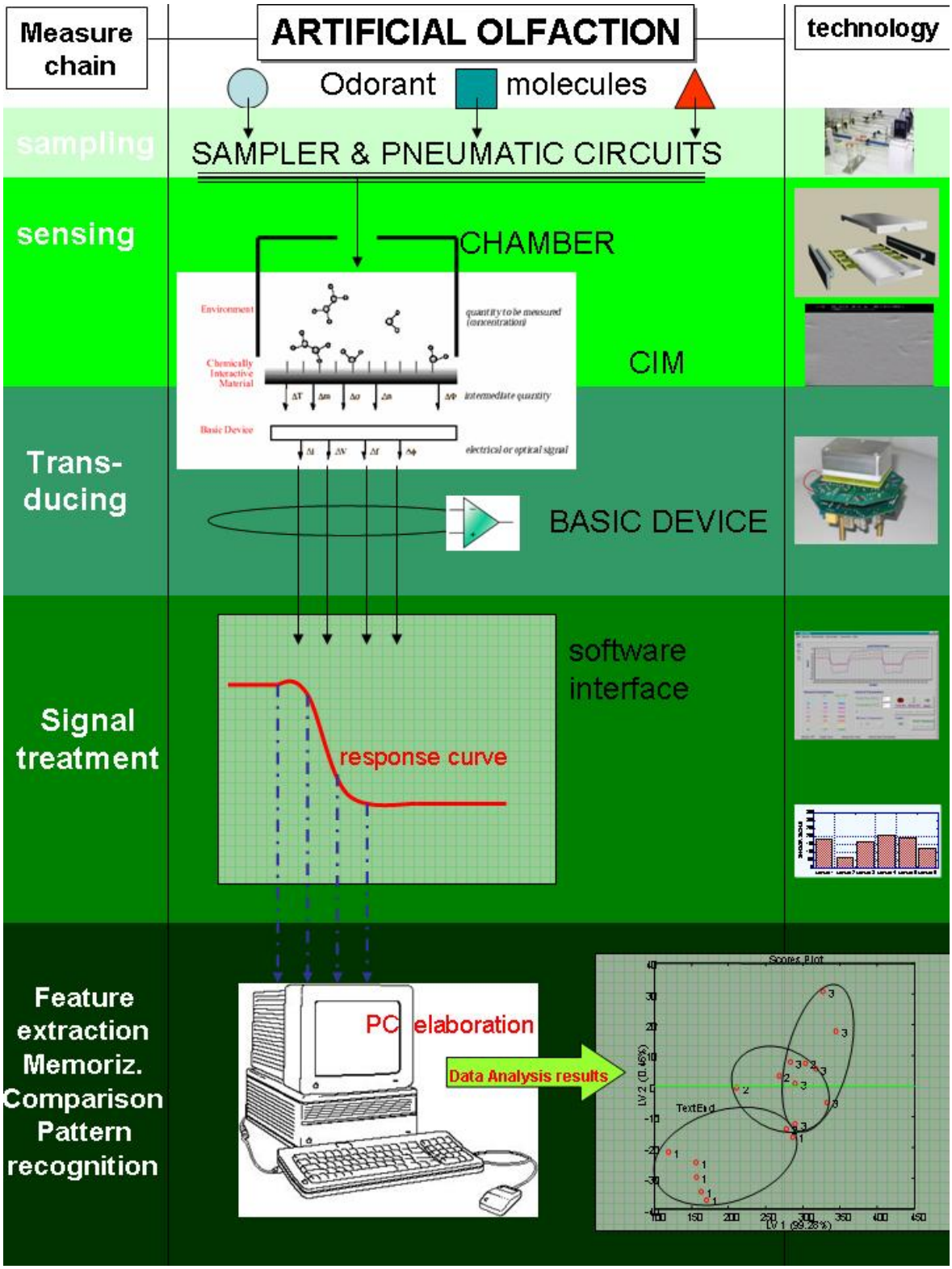


Fig. 1.3.1 : 'flow chart' of the artificial olfaction system

Transducers and basic devices

The transducer part of the sensor translate the information to a signal that is suitable for further processing. There is a wide range of transducers available; the following list is based on the main physical quantity considered:

- ✓ Change in Work Function: Chemfet, Light Addressable Potentiometric S. [25,26,27,28]
- ✓ Change in Conductivity: Chemoresistance [29-36,53-54]
- ✓ Change in Ionic Current: Amperometric Gas Sensor [37-41]
- ✓ Change in Dielectric Constant: Chemocapacitor [42]
- ✓ Change in Temperature: Thermopile, Pellistor Catalytic Sensor [43,44]
- ✓ Change in optical spectrum: Colorimeter and spectrophotometer [45]
- ✓ Change in Fluorescence: Optical Fiber [46]
- ✓ Change in Refraction index: Optical Fiber, Surface Plasmon Resonance [47-51]
- ✓ Change in Mass: Cantilevers, Surface Acoustic Wave, Quartz Crystal Microbalance. [2,4,5,24,52]

The list above, provides a good review of chemical gas sensors, nontheless two important additional features makes them sensors for electronic noses: a particular choice of the chemical interactive material (CIM) to use as coating of these devices to obtain the ability to be sensitive to chemical compounds, and the array strategy with different CIM, to have patterns of different responses that represents something similar to odours.

The concept of using sensor array for chemical analysis offers advantages over individual sensors concerning sensitivity to a wider range of analytes, improved selectivity and the capability for recognition of both single and complex analytes. According to this philosophy the odor could be described not by a sum of its individual components but by some abstract representation, a sort of *chemical image,* a virtual fingerprint.

An alternative way to define the electronic nose, is to consider the subject of its investigation: complex samples characterized by the simultaneous presence of many compounds.

The "quality" of these samples are sometimes not related to few specific compounds but rather to the global composition.

A strategy to measure these samples is to use instruments sensitive to the globality of the compounds.

An electronic nose is an array of individual sensors different one each other but globally selective according to the principle that each sensor senses more compounds and each compound is sensed by more sensors.

Chemical Interactive Material

The chemical recognition translates information from the chemical domain, into a physical output signal with a defined sensitivity. The main purpose of the recognition system is to provide the sensor with a unique profile of selectivity for the analyte to be measured [22].

Considering the different interactions taking place in the binding processes between odorant molecules an CIMs, we can observe two main types of solid gas interaction:

- Physisorption
- Chemisorption

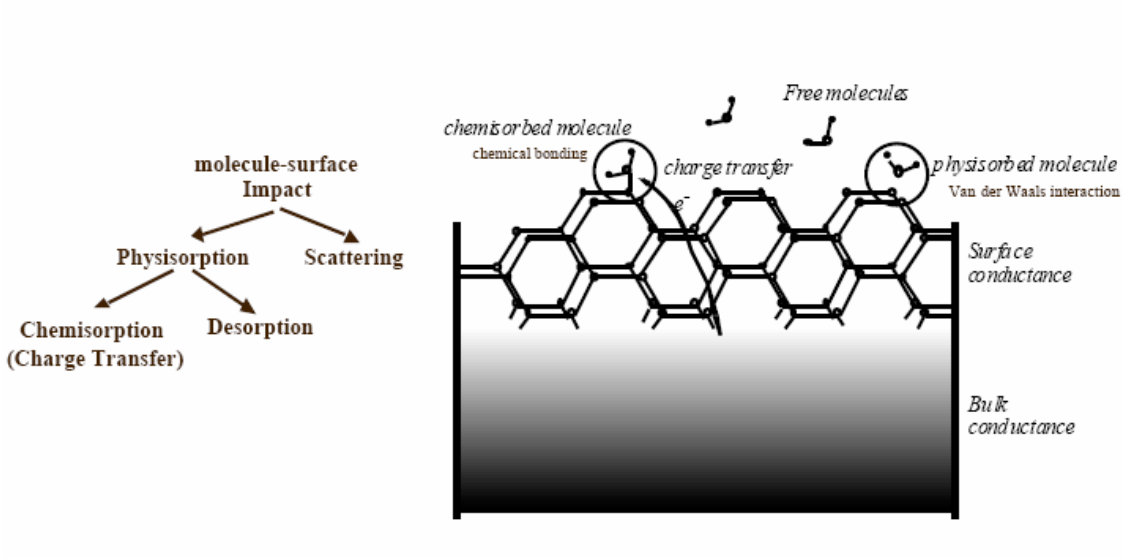


Fig. 1.3.2: Schematic representation of sorption phenomena.

The contribution of each of the above mentioned mechanisms depends on the nature of receptors centers, thickness of the sensitive layer and film structure.

The forces involved in Chemisorption are the valence forces involved in the formation of chemical compounds. The forces involved in Physisorption are intermolecular forces, with not significant change in the electronic orbital patterns of the species involved.

Between Chemisorption and physisorption a not absolutely definite distinction can be made because of the relevance of intermediate cases; for istance adsorption involving strong hydrogen bonds or weak charge transfer.

There are a number of factors which determine the strength and specificity of intramolecular interaction, such as molecular size and symmetry; polarizabilty of the molecular skeleton; substituents type, number and polarity; degree of freedom of individual molecular components; hydrophilic-hydrophobic properties of interacting components.

Usually, the higher is the interaction specificity, the stronger is the interaction. But for sensor applications reversibility of the sorption process is a must. In these processes the adsorbed molecule is held on the solid surface by the forces of predominantly electrostatic nature. These are Van Der Waals, Hydrogen, donor-acceptor interactions.

The Van Der Waals forces are a special kind of weak non-valence interactions. The dispersion force is the most general example of these type of interactions. The main features of these interactions is that they are not oriented somewhere.

The polar molecules demonstrate interaction not only through the dispersion forces but also by means of dipole and induction forces of electrostatic origin.

Hydrogen bonding, another type of intermolecular interaction characteristic of the reversible adsorption processes, presents an energy much more over Van Der Waals one. It determine a stronger interaction of higher specificity.

A weak donor-acceptor bonding also can be a reason for physical adsorption.

An understanding of the sorption and selectivity is important to design such sensors, this choice, anyway, depending not only on the CIM but also on the target analyte. Tailoring the physical and chemical properties of coating materials used in the sensors can control the selectivity and sensitivity of each sensor [23].

A possible list of these materials, sometimes strictly connected with particular transducers, are reported above:

- ✓ organic sensitive materials:
  - Conductive and non conductive polymers;
  - Molecular Films;
- ✓ Inorganic sensitive materials:
  - Catalytic Metals;
  - Metal Oxide;
  - Solid State Ionic Conductor;

A Brief summary

To conclude this overview with a clear schematic representation, is necessary to build a table to merge all chemical, physical and electronic features in a flow chart, to classify the existing enose technologies.

The sensing material is the feature on the basis of which we can organize a sort of classification. This is because of the importance of the choice of the CIM in the definition of the transduction mechanism and the basic device. To make an example the 'chemical language' by means of which our sensor interact with the environment is decided on the basis of the chosen CIM, as a consequence we can choice the correct 'dictionary': the transducer.

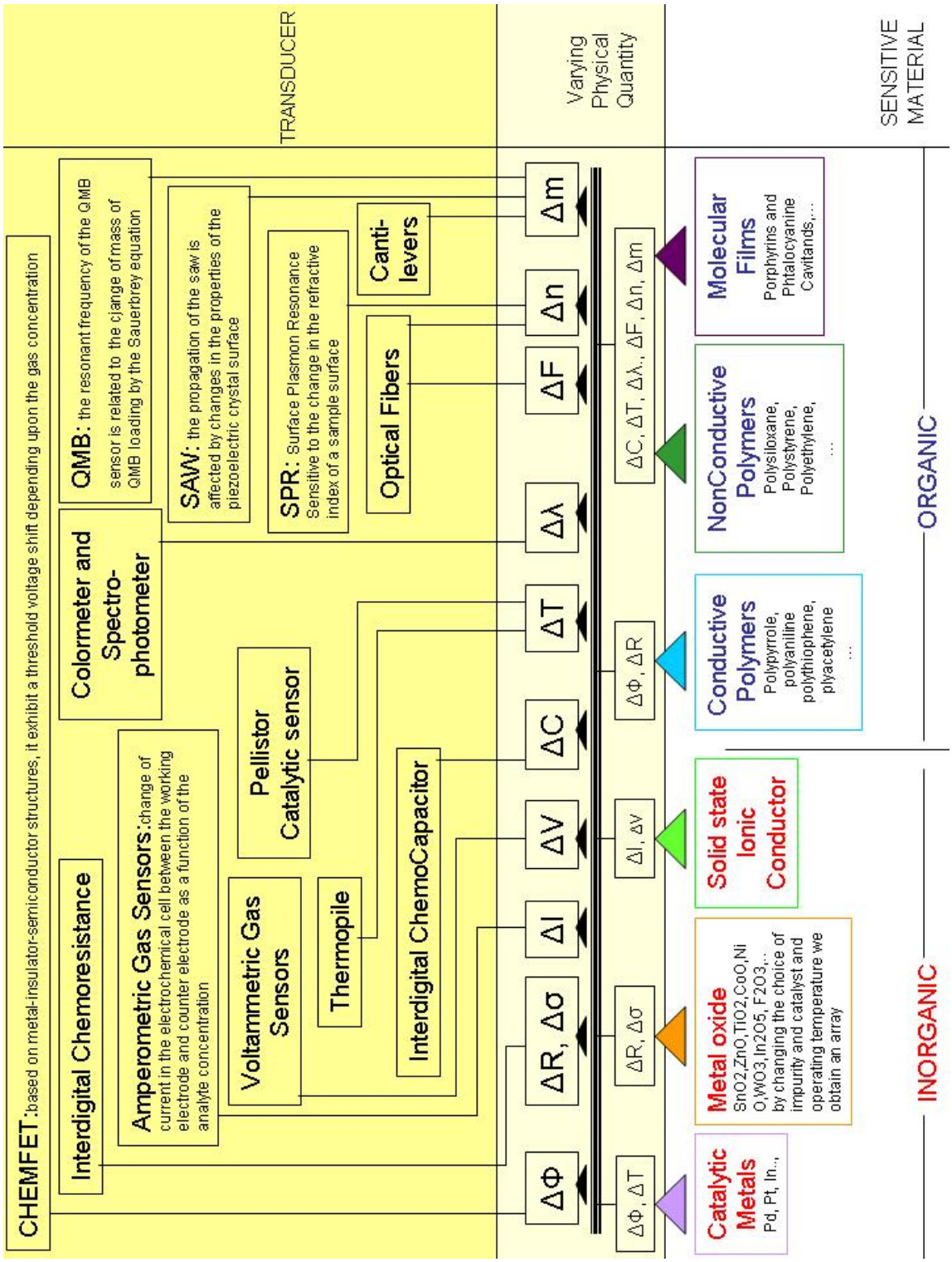


Fig. 1.3.3 **: Enose existing technology scheme**

Figure 1.3.3, provides an attempt to classify the current e-nose technology.

1.4 Tor Vergata Electronic Nose

The first prototype of the electronic nose of the University of Rome Tor Vergata was built in 1994. The name of this device, Libra Nose, says something about its working principle: 'it weighs the odor'. It means that the mass is the physical quantity varying during the measurement. Looking at the scheme in table1.3.1, there are a number of different transducers for variation of mass; QMBs are the transducers used in the Libra Nose.

Among the mass sensitive transducers the quartz microbalances are the most simple. These sensors consist in a thin slab of crystalline quartz, cut along a certain symmetrical axis in order to obtain a material able to sustain bulk electroacoustical oscillation at certain frequencies . The most commonly utilized is the so-called AT cut from which oscillating frequencies from 5 to 30 MHz are obtained [24].

As for any mechanical oscillator, the resonant frequency of the quartz is inversely proportional to the mass graviting on its surfaces. This characteristic, well known since the beginning of the 1960s, can turn a quartz into a chemical sensor when a chemically interactive material, able to capture molecules from the environment, coats the quartz surfaces. Such sensors are often called Quartz Microbalances (QMB) .

The frequency-mass relationship is described, at a first approximation, by a relation known as Sauerbrey Law:

$$\Delta f_0 = -\cdot \frac{2 f_0^2}{A\sqrt{\mu_q \rho_q}} \Delta m_s \qquad (1.3.1)$$

where A is the coated area, ρ_q the quartz density, μ_q is the shear stiffness, f_0 is the fundamental frequency.

MetalloPorphyrins are the chemical interactive material used in the Libra Nose as coating of the quartzes.

Porphyrins are necessary for life: their functions as complexing ligands or redox catalysts are essential for all organisms.

A large number of modified tetrapyrroles is present in nature or has been synthesised in laboratories. The properties of these macrocycles can be tuned by simple modifications on the basic molecular framework and these skeletal variations are used by nature to optimise the biological activities of different tetrapyrroles.

From the point of view of their use for sensor applications, their main feature is the dependence of the sensing properties (in terms of selectivity and sensitivity) on the nature of both the central metal and peripheral substituents of the macrocycles. With small variations in the synthetic process it is possible to get sensors with different behaviour.

In Libra Nose different metal complexes of tetraphenylporphyrin (TPP) have been used, moreover molecule have also been functionalised with four alkylic chains ($O(CH_2)_3CH_3$) in order to ensure the necessary porosity of the solid state film for an optimal analyte diffusion through the absorbing film.

All together these characteristics increase the versatility of these molecules and different transducers have been proposed for porphyrin-based chemical sensors, all showing outstanding properties of these materials in terms of stability, chemical sensitivity and reproducibility.

Along a period of ten years the Libra Nose has shown to be an instrument of practical use in numerous fields of application, such as medicine, food industry, environmental monitoring (see respectively chapter 3 and 4 in this book). Each of these different experimentations has asked for ad hoc modification in the sampling procedures, in the chemistry of the sensing material and of course in the sensors themselves, so that the Libra Nose has gone through a series of different improvements resulting in four following versions of the instrument. The complete evolution is illustrated in the figure 1.4.1.

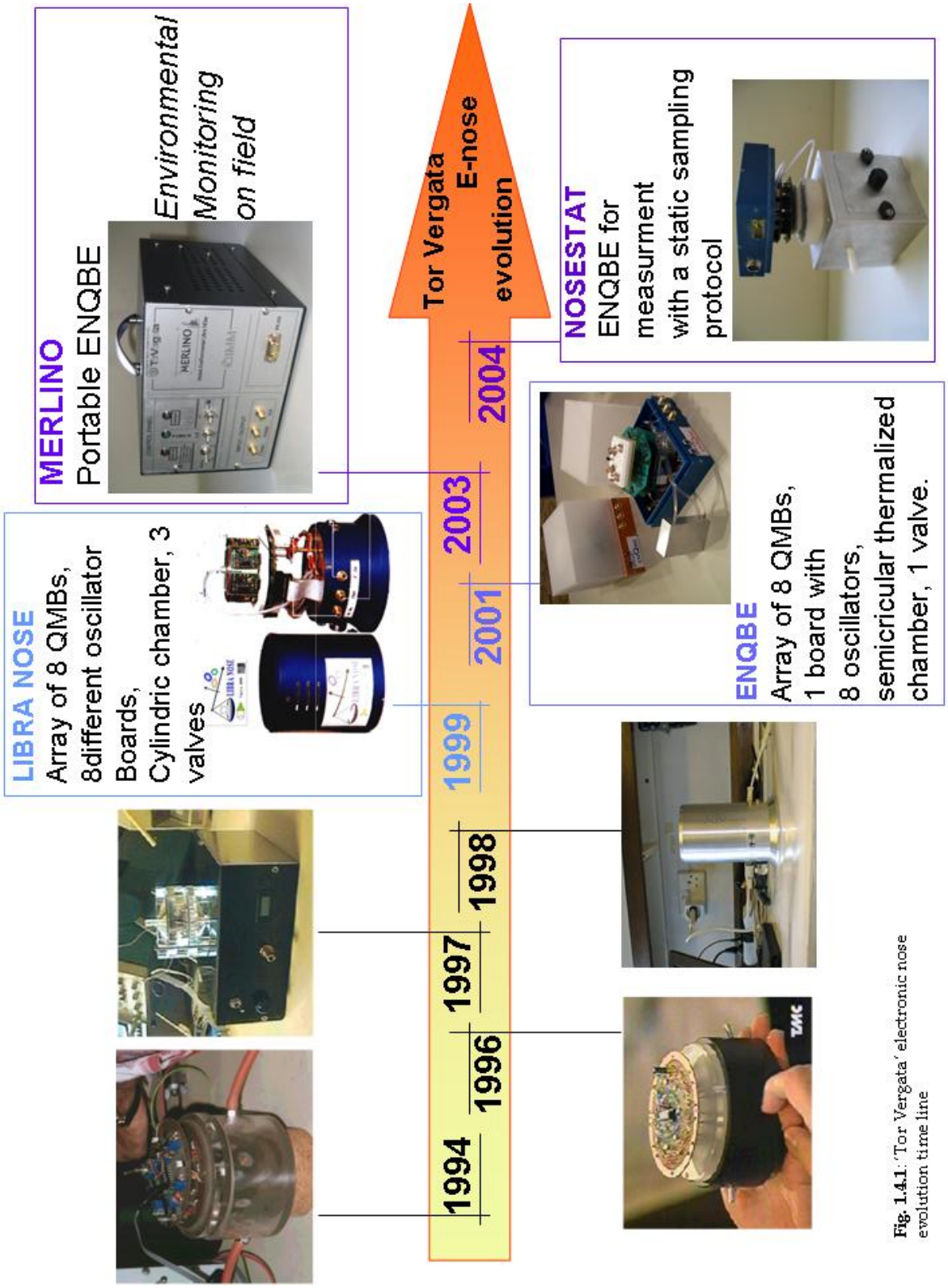


Fig. 1.4.1: 'Tor Vergata' electronic nose evolution time line

1.5 References

[1] W. Gopel, 'Chemical Imaging: I. Concepts and visions for electronic and bioelectronic noses', *Sensors and Actuators B* 52 (1998) 125-142.

[2] C. Di Natale, A. Macagnano, R. Raolesse, E. Tarizzo, A. D'Amico, et al., 'A comparison between an electronic nose and human olfaction in a selected case study', *Proceedings of Transducers '97*.

[3] W. Gopel, Ch. Ziegler, H. Breer, D. Schild, R. Apfelbach, J. Joerges, R. malaka, 'Bioelectronic noses: a status report, Part I', *Biosensors and Bioelectronics* 43 (1998) 479-493

[4] A. D'Amico, C. Di Natale, A. Macagnano, F. Davide, A. Mantini, E. Tarizzo, R. Paolesse, T. Boschi, ' Technologies and Tools for mimicking olfaction: status of the Rome "Tor Vergata" electronic nose', *Biosensors and Bioelectronics* 13 (1998) 711-721.

[5] A. D'Amico, C. Di Natale, R. Paolesse, ' Portraits of gasses and liquids by arrays of nonspecific chemical sensors: trends and perspectives', *Sensors and Actuators B* 68 (2000) 324-330.

[6] T.C. Pearce, S. S. Schiffman, H. T. Nagle, J. W. gardner, '*Handbook of machine Olfaction*' , Wiley-VCH (2003).

[7] G. Ohloff, 'Chemistry of Odor Stimuli', *Experientia* 1986, 42, 271-279.

[8] H. DeVries, M. Stuiver, 'The absolute sensitivity of the human sense of smell', Sensory Communication, John Wiley and Sons, New York, 1961, Chapter 9, pp 159-167.

[9] A. Jinks, D. G. Laing, 'The Analysis of odour mixtures: evidence for a configurational process', *Physiol. Behav*. 2001, 72, 51-63

[10] R. Axel, 'The molecular logic of smell', *Sci. Am.* 1995, 273, 154-159.

[11] L. Buck, R. Axel, 'A novel multigene family may encode odorant receptors: a molecular basis for odor recognition', *Cell*, 1991, 65, 175-187.

[12] P. Mombaerts, 'Molecular biology of odorant receptors in vertrebates', *Annu. Rev. Neurosci*., 1999, 22, 487-509.

[13] P. Mombaerts, ' Seven-transmembrane proteins as odorant and chemosensory receptors', *Science*, 1999, 286, 707-711.

[14] D. Whissell-Buechy, J. E. Amoore, 'Odour blindness to musk: simple recessive inheritance', *Nature* 1973, 242, 271-273.

[15] E. J. Amoore, 'The stereochemical specificities of human olfactory receptors', Perfum. Essent., *Oil Rec*. 1952, 43, 321-323.

[16] R. H. Wright, 'Stereochemical and vibrational theories of odour', *Nature*, 1972, 239, 226.

[17] A. Dravnieks, P. Laffort, 'Physico-chemical basis of quantitative and qualitative odor discrimination in humans', *Olfaction and Taste* vol.4, 1972, pp. 142-148.

[18] K. Keyhani, P. W. Scherer, M. M. Mozell, 'A Numerical Model of Nasal Odorant Transport for the Analysis of Human Olfaction', *J.theor. Biol*. (1997) 186, 279-301.

[19] W. J. Freeman, 'The physiology of Perception', *Scientific American* (1991) vol.264, (2) pp. 78-85.

[20] I. Savic, 'Imaging of brain avtivation by odorants in humans', *Current Opinion in Neurobiology* 2002, 12: 455-461.

[21] J.W. Gardner , P. N Bartlett. 'A Brief History of Electronic Noses'. *Sensors and Actuators B*., 18, 211–220 (1993)

[22] B. A. Snopok, I. V. Kruglenko, ' Multisensor system for chemical analysis: state of the art in electronic nose technology and new trends in machine olfaction', *Thin Solid Films* 418 (2002) 21-41.

[23] B. Adhikari, S. Majumdar, 'Polymers in sensor applications', *Prog. Polym. Sci*. 29 (2004) 699-766.

[24] C. Di Natale, R. Paolesse, A. Macagnano, A. Mantini, P. Mari, A. D'Amico, 'Qualitative structure-sensitivity relationship in porphyrins based QMB chemical sensors', *Sensors and Actuators B* 68 (2000) 319-323.

[25] I. Lundstrom, S. Shivaraman, C. Svensson, L. Lundkuist, *Appl. Phys. Lett*., 1975, 26, 55-57.

[26] I. Lundstrom, E. Hedborg, A. Spetz, H. Sundgren, F. Winquist, Sensors and Sensory Systems for an E-nose, Eds. J. W. Gardner and P. N. Bartlett, *NATOASI Series* (Kluwer, Dordrecht), 1992, 212, 303-319.

[27] S. L. Rose-Pehrsson, J. W. Grate, D. S. Ballantine, P. C. Jurs, *Anal. Chem.* , 1988, 60, 2801-2811.

[28] Y. Morita, K. Nakamura, C. Kim, 'Langmuir analysis on hydrogen gas response of palladium-gate FET', *Digest of technical papers, Transducers '95 / Eurosensors IX*, 1995, Stockholm, vol.I, pp 734-737.

[29] P. N. Bartlett, S. K. Ling-Chung, *Sensors and Actuators*, 1989, 20, 287-292.

[30] J. W. Gardner, P. N. Bartlett, *Nanotechnology*, 1991, 2, 19-33.

[31] J. R. Stetter, S. Zaromb, M. W. Findlay Jr., *Sensors and Actuators A*, 1984, 6, 269-288.

[32] H. Nanto, Y. Yokoi, T. Mukai, J. Fijioka, E. Kusano, A. Kinbara, Y. Douguchi, *Mater. Sci. Eng*., 2000, 12, 43-48.

[33] S. Kurosawa, N. Kamo, D. Matsui, Y. Kobatake, *Anal. Chem*., 1990, 62, 353.

[34] U. Lampe, J. Gerblinger, H. Meixner, 'Comparison of transient response of exhaust-gas sensors based on thin film of selected metal oxides', *Sensors and Actuators B*, 7 (1992) 787-791.

[35] N. Taguchi, 'Gas detecting element and methods of making it', *US Patent*, No. 3644795 (1972).

[36] U. Kirner, K. D. Schierbaum, W. Göpel, B. Leibold, N. Nicoloso, W. Weppner, D. Fischer, W. F. Chu, 'Low and high temperature TiO2oxygen sensors', *Sensors and Actuators B* (1990) 103-107.

[37] (a) J. R. Stetter, S. Zaromb, M. W. Findlay, U.S. PAtent 5055266, 1991. (b) J. R. Stetter, S. Zaromb, W. R. Penrose, *U.S. Patent* 4670405, 1987. (c) J. R. Stetter, Chemical Sensor Arrays: Practical Insights and Examples, in Sensors and Sensory Systems for an e-nose, Eds. J. Gardner and P. N. Bartlett, (Kluwer Academic Publishers). 1992, 273-301.

[38] J. R. Stetter, J. *Colloid Int. Sci*., 1978, 65 (3) , 432-443.

[39] F. P. Winquist, P. Wide, I. Lundstrom, *Anal. Chim. Acta*, 1997, 357, 21-31.

[40] S. C. Chang, J. R. Stetter, C. S. Cha, *Talanta*, 1993, 40 (4), 461-467.

[41] Z. Cao, W. J. Buttner, J. R. Stetter, *Electroanalysis*, 1992, 4, 253-266.

[42] A. Koll, S. Kawahito, F. Mayer, C. Hagleitner, D. Scheiwiller, O. Brand, H. Baltes, *Proc. SPIE*, 1998, 3328, 223-232.

[43] J. N. Zemel, *Sensors and Actuators A*, 1996, 56 (1-2) 57-62.

[44] W. J. Buttner, J. R. Stetter, G. J. Maclay, *Sens. Mater*., 1990, 2, 99-106.

[45] D. S. Ballantine Jr., D. Callahan, G. J. Maclay, J. R. Stetter, *Talanta*, 1992, 39(12), 1657-1667.

[46] J. White, J. S. Kauer, T. A. Dikkinson, D. R. Walt, *Anal. Chem*., 1996, 2191-2202.

[47] B. Liedberg, C. Nylander, I. Lundstrom, *Sensors and Actuators*, 1983, 4, 299-302.

[48] E. Kretschmann, *Z. Phys*., 1971, 241, 313.

[49] K. S. Johnston, S. R. Karlson, C. Jung, S.S. Yee, *Mater Chem. Phys*., 1995, 42, 242.

[50] B. Chadwick, M. Gal, *Jpn. J. Appl. Phys*., 1993, 32, 2716.

[51] S. Nelson, K. S. Johnston, S. S. Yee, *Sensors and Actuators B*, 1996, 35/36, 187.

[52] T. Hamacher, J. Niess, P. S. Lammers, B. Diekmann, P. Boeker, 'Online measurement of odorous gases close to the odour threshold with a QMB sensor system with an integrated preconcentration unit', *Sensors and Actuators B* 95 (2003) 39–45.

[53] P. Ivanov, E. Llobet, X. Vilanova , J. Brezmes , J. Hubalek , X. Correig, 'Development of high sensitivity ethanol gas sensors based on Pt-doped SnO2 surfaces' *Sensors and Actuators B* 99 (2004) 201-206.

[54] S. Capone, P. Siciliano, F. Quaranta, R. Rella, M. Epifani, L. Vasanelli,' Analysis of vapours and foods by means of an electronic nose based on a sol–gel metal oxide sensors array', *Sensors and Actuators B* 69 2000 230–235.

SAMPLING

'In the middle of every difficulty lies opportunity'
Albert Einstein

'Nel mezzo di ogni difficoltà riposa un opportunità'

2

2.1 Sampling

In the first chapter an overview on natural and artificial olfaction has been presented. About the state of the art of artificial olfaction a survey on existing technologies has also been illustrated. In spite of the great variety of working principles and sensitive materials, an important matter is common to each of these devices when practical applications are designed: the sampling problems.

Performing odor measurements with an electronic nose, means to face a lot of problems strictly connected to the nature of the mesurand (odor) which is very difficult to handle and deliver.

Actually, a complete definition of any system mimicking natural olfaction, beside sensor array, data acquisition system and pattern recognition algorithm, has to contain the sampling system.

The role of the sampling system is to collect and convey the volatile sample to the sensors, then to restore previous conditions by means of a cleaning procedure. Considering that the interaction between sensors and odours is the first fundamental step of the data acquisition process, since its execution influences all successive steps, the way to put the sensor in contact with the analyte is decisive [2].

In order to measure with a good stability and repeatability as well as high amplitude signals and fast sensor responses, the sampling device has to be designed in such a way that all factors capable of influencing sensor responses are optimised and kept under control, so that the only variable left is the odour composition [8].

To analyze any substance with an e-nose, the sample has to be brought into the sensor chamber, and it could seem very simple, but for simple operations such as taking up of the sample, conditioning it, and transferring it in the measure chamber, a number of different methods exists.

The development of sampling systems and of sample pre-treatment are fundamental ways to increase the utility and performance of instrumental analysis. While sampling for classic analytical instruments evolved to a variety of sophisticated methods to gain optimal instrument performance, the potential

sampling for e-nose has not yet thoroughly been investigated. By simply taking over established sampling techniques for classic analytical instruments, the special requirements and potentialities of e-nose have often been neglected [5].

To design an optimized sampling protocol it is necessary to evaluate the possible factors influencing the reproducibility of the experiment. The lack of reproducibility is, indeed, a recurrent problem for aroma quality control. In a great extent experimental condition fluctuations leads to different sensor responses. Among the influent parameters the most common are temperature, hygrometry and the headspace sampling technique itself.

Anyway the problem of sampling odor is so articulated that it is not possible to be exhaustively treated with a general overview, this is because it is strictly connected to the mesurand rather than to the sensor itself.

For this reason in this chapter a general review of sampling techniques theory and problems will be presented, while designing and testing of particular sampling protocols will be treated in the following three chapters, each of them being referred to each different experiment, concerning objectives very different.

2.2 Theoterical aspects of sampling

There are two aspects of a sampling protocol that can affect the efficiency of an experimental set. Considering that a sampling technique is a strategy to put in contact a sample with a sensing material, the correct way to face the problem is to approach it from two different points of view: the sample, and the sensor. It is then necessary to illustrate the principles of the physics of evaporation, looking from the point of view of the sample, and then the principles of the sorption theory, looking at sampling problems from the point of view of the sensor.

Principles of the physics of evaporation

The sample tested by e-nose is a gas mixture. These gases very often consists of odorants evaporated from liquid or solid. It is therefore important to know the physiscochemical behaviour of evaporation to relate the properties of gases with those of solid and liquid samples.

A key point strictly connected with the nature, composition and concentration of the sample under measure, is the fact that saturated vapour pressure is dependent on temperature.

Although thermodynamics calculations related to phases equilibrium ruled by the Clapeyron law are available, several of the assumptions made in the derivation of the equation fail at high pressure and near the critical point, and under those conditions the Clausius-Clapeyron equation gives inaccurate results.

To counteract the possible deviation from theory, the *Antoine equation* is used. It consists in an empirical 3-parameter fit to experimental vapor pressures measured over a restricted temperature range:

$$\ln(P) = A - \frac{B}{C+T} \qquad (2.2.1)$$

where A, B and C are experimentally measured constants, different for each compound.

The vapour concentration should be kept below the maximum corresponding to the saturated vapour pressure, otherwise the excess of the

vapour pressure above the saturated point leads to its condensation into liquid drops.

It is important to remark that the saturation pressure of a compound with high odour intensity is tipically small whereas highly volatile compounds have high saturated vapour pressures.

Antoine law applies for pure compounds a situation that is not found in practical applications where samples are composed by mixtures.

In a mixture of volatile compounds, evaporating from a complex matrix, the relative concentration of each compound in the vapour above the sample at equilibrium is determined by both the partial vapour pressure and the amount of compound present in the matrix. According to Raoult's law:

$$P_A = N_A P_A^0 \qquad (2.2.2)$$

Where P_A is partial pressure of compound A,
N_A the molar ratio of that compound in the solution,
P^0_A the vapour pressure of the pure compound

a small amount of compound with a high vapour pressure (a highly volatile compound) will contribute more to the headspace vapour than a similar amount of compound with a lower vapour pressure. This means that the relative proportion of the compounds in the vapour or 'air space' or 'head space' surrounding the sample may not be the same as the proportions of the same compounds within the solid or liquid matrix.

Anyway this equation is valid in the case of ideal solutions. Most compounds, however, are non-ideal solutions. In non-ideal solutions, the equation above is replaced by:

$$P_A = \gamma_A N_A P_A^0 \qquad (2.2.3)$$

Where γ_A is an activity coefficient dependent upon N_A. Since for interaction between the components occurring in the non-ideal solution, the superposition theorem for the compound mixture is not valid.

Elements about the theory of sorption

When a solid, such as the chemical interactive material of a sensor, is exposed to a gas at some definite pressure, it begins to adsorb gas molecules and

(if the solid is the coating for example of a QCM) it can be seen that the process is accompanied by an increase in the weight of the solid and a decrease in the pressure of the gas. After some time the pressure becomes constant at the value *p*, say, and correspondingly the weight graviting on sensor ceases to increase any further [9].

The phenomenon described above is termed *Adsorption*. Adsorption is the selective accumulation of a chemical at the interface between two phases. Vice versa *Desorption* is the reverse of adsorption. The substance that adsorbs is called the *adsorbate* and, if it binds at a solid/liquid interface, the solid is called *adsorbent*; *adsorptive* is the general term for the material in the gas phase which is capable of being adsorbed.

The adsorption is brought about by the forces acting between the solid and the molecules of the gas. These forces are of two main kinds – physical and chemical – and they give rise to physical (or Van der Waals) adsorption, and chemisorption respectively. To be more precise using the different terms, *adsorption*, a surface or interfacial process, can be different from *absorption*, in which a substance is transferred from one phase into the interior of another. However, it is often almost impossible to distinguish between adsorption and absorption, so the two processes are sometimes treated jointly. In such cases, the terms defined above are written as *sorption, sorbate, sorbent*.

The quantity of gas taken up by a sample of solid is proportional to the mass of the sample, and it depends also on the temperature *T*, the pressure *p* of the vapour, and the nature of both the solid and the gas.

If n is the quantity of gas adsorbed expressed in moles per gram of solid,

$$n = f(p, T, gas, solid) \tag{2.2.4}$$

For a given gas adsorbed on a particular solid maintained at a fixed temperature, the eaquation (2.2.4) simplies to

$$n = f(p)_{T,gas,solid} \tag{2.2.5}$$

If the temperature is below the critical temperature of the gas, the alternative form

$$n = f\left(\frac{p}{p^0}\right)_{T,gas,solid} \tag{2.2.6}$$

is more useful, p^0 being the saturation vapour pressure of the sorptive.

Equations (2.2.5) and (2.2.6) are expressions of the adsorption isotherm, i.e. the relationship, at constant temperature, between the amount of gas adsorbed and the pressure, or relative pressure, respectively.

In the literature of subject there are recorded tens of thousands of adsorption isotherms, measured on a wide variety of solids. Nevertheless, the majority of those isotherms which result from physical adsorption may conveniently be grouped into five classes, referred to a classification proposed by Brunauer, Emmett and Teller (BET).

2.3 Sampling techniques: an overview

Looking at the wide list of sampling techniques of classic analytical chemistry, it must be taken into account, again, the fact that odour is the sampled mesurand. A sample of such a headspace might be expected to approximate the composition of the vapour in field conditions, in the real case in which an odour can be perceived. The suite of compounds detected may reflect the odour experienced by any other organisms in the environment, but with an objective evaluation.

So the overview of different strategies is a list of protocols similar to the 'classic one' (GC analysis) [4], but sometimes modified and adapted to the particular target of 'sniffing' in spite of analyzing.

The different techniques for odour handling and delivery can be divided in two main categories: without and with preconcentration. The second strategy, with preconcentration, is more used in case of GC systems, while the first is common in 'real odour measurements', so with electronic noses.

All the techniques performed without preconcentration can be divided again in other two classes referred to two different sampling systems: the sample flow system and the static system. In the first case, of sample flow system, the measurement is performed placing the sensor in the vapour flow; this technique presents the advantage of rapid exchange of vapour, and of the fact that many samples can be measured within a short time. In the case of the static system the sensors are exposed to vapour at a constant concentration, without any vapour

flow. The scheme shown in figure 2.3.1 is useful to illustrate the 'tree' of sampling possibilities, it also needs in introducing the rest of this thesis, because each of the following chapters is about an experimentation exploiting and adapting one of these techniques to the three real case study.

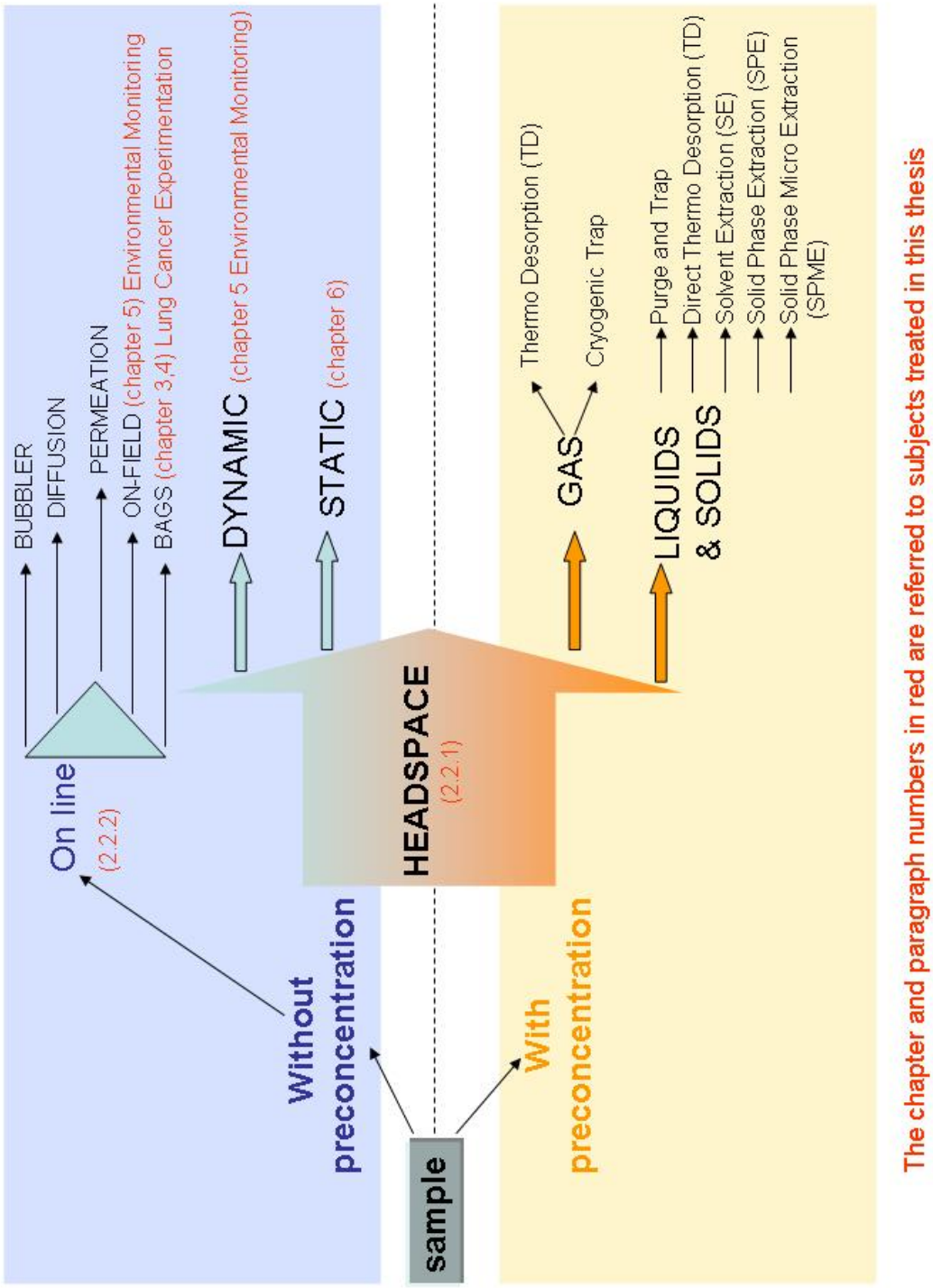


Fig. 2.3.1 : sample techniques overview

Looking at the scheme of fig. 2.3.1 the first distinction can be done between sampling systems with or without preconcentration.

Although the discrimination capabilities of an odor-sensing system using a sensor array are determined by the combination of the sensors, it is possible to achieve further improvement if different information can be combined with the sensor array. The main way for this target is the use of a preconcentrator usually used to compensate sensor sensitivity since the vapor desorption temperature varies according to the vapour kind [11]. It was also reported that chromatographic behavior of the preconcentrator also helped the sample discrimination when multi-way pattern recognition technique was applied [7].

Although the increase in sensitivity and in data acquisition, all the methods without preconcentration are the most commonly used for sampling in electronic nose applications.

Anyway, there are two main methods to treat the sample object of the measurement, and the differences in these two strategies, are mainly due to a different point of view; in the case of headspace sampling the 'sampling action' is to delivery the sample to the e-nose, in the case of on-line measurement the 'sampling action' is to delivery the e-nose in the sample environment.

2.3.1 Headspace sampling

To study the headspace sampling is necessary to start from a description of the 'sampling action': the sample headspace (HS), i.e. the gas volume above the substance to be analysed, or a representative part of it has to be brought into the sensor chamber [1].

The measure of the headspace using chemical sensors reveals information about the nature of the sample. This sampling process consists of taking up the sample, conditioning it, and transferring it to the analytical equipment.

There are several possible methods for performing it, but the need of operating with maximum possible efficiency without altering the composition of the headspace, is obviously a must for the objectivity of the measurement.

Headspace analysis is an extraction technique for semi volatile and volatile compounds. When the system is in equilibrium, the composition of the vapour phase is in quantity and quality representative of the composition of the original sample and, recalling the considerations about the Raolut's Law, it can be understood why the word 'representative' is used in spite of 'the same'. An aliquot is taken from the gas phase and transferred to analysis.

For routine analysis with electronic noses headspace analysis is the most common sampling method. A correct and effective application of this method is obtained with a careful monitoring of a number of parameters defined below.

With the concentrations of the analyte i in the gas phase $C_{i(g)}$ and liquid/solid sample phase $C_{i(s)}$, the partition coefficient K is given by

$$K = \frac{C_{i(s)}}{C_{i(g)}} \qquad (2.3.1)$$

The area of the signal peak A_i for component i is proportional to the concentration in the gas phase $C_{i(g)}$ and proportional to the original concentration in the sample.

To determine the concentration of an analyte in the headspace, there is the need to calculate the partition coefficient. Samples must be prepared to maximize the concentration of the volatile components in the headspace, and minimize unwanted contamination from other compounds in the sample matrix.

Compounds that have low K values will tend to partition more readily into the gas phase, and have relatively high responses and low limits of detection.

An example of this would be hexane in water: at 40°C, hexane has a K value of 0.14 in an air-water system.

Compounds that have high K values will tend to partition less readily into the gas phase and have relatively low response and high limits of detection.

An example of this would be ethanol in water: at 40°C, ethanol has a K value of 1355 in an air-water system. Partition coefficient values for other common compounds are shown in table 2.3.1.

Solvent	K Value
cyclohexane	0.077
n-hexane	0.14
tetrachloroethylene	1.48
toluene	2.82
benzene	2.90
dichloromethane	5.65
n-butyl acetate	31.4
n-butanol	647
isopropanol	825
ethanol	1355

Table 2.3.1 : Partition Coefficient values for some common compounds

Another important parameter to match with the partition coefficient, is the phase ratio (ß), which is defined as the relative volume of the headspace compared to volume of the sample in the sample vial.

$$\beta = \frac{V_g}{V_s} \qquad (2.3.2)$$

Lower values for ß (i.e., larger sample size) will yield higher responses for volatile compounds.

Partition coefficients and phase ratios work together to determine the final concentration of volatile compounds in the headspace of sample vials. The concentration of volatile compounds in the gas phase can be expressed as

Cg = Co / (K+ß)

(where Cg is the concentration of volatile analytes in the gas phase and Co is the original concentration of volatile analytes in the sample). Striving for the lowest values for both K and ß will result in higher concentrations of volatile analytes in the gas phase and, therefore, better sensitivity.

The sample uptake requires the control of different parameters such as temperature, pressure, volume and headspace formation time being the most critical ones. The performance of the headspace Analysis can be taken under

control, by the monitoring of some parameters: the partition coefficient, the vial volume determining the maximum sample amount, and the matrix.

In a closed or static system the gas or vapour phase will be in equilibrium with the condensed phase. Analytes are distributed between the condensed phase matrix and the vapour phase. In some cases conditions can be adjusted so that the analyte distribution favours the vapour phase (to increase the concentration of the headspace in volatile compounds).

Headspace Analysis consists of two steps. In the first step, the sample is placed in a vial. The closed vial is thermostated and, if necessary, shaken for a defined time until equilibrium between phases (solid/gas or liquid/gas) is reached. In the second step, the vial is pressurized and subsequently vented so that an aliquot of the vial's headspace is introduced into the carrier gas stream and transferred to analysis (to the sensor chamber). This is performed either by a balanced-pressure system or over a pressure/loop system. In the balanced-pressure system the sample headspace is injected over a given time by the carrier gas using the pressure of the analytical instrument inlet. In the pressure/loop system then is opened toward a sample loop and equilibrated against ambient pressure. Then, the loop is flushed by carrier gas transferring the sample to the e-nose [4].

The carrier gas such as dry air is supplied at the inlet of the bottle where the headspace has been generated and the vapour evaporated at the liquid surface carried by the carrier gas is supplied to the sensors. The simplicity of this protocol represent id major advantage. Nevertheless there are many attentions to take into account and some problematic objectives.

The distance between the liquid surface and the tips of the tool used to suck the headspace should be kept constant since the vapour in the headspace is often unsaturated and its concentration varies according to the level.

Although many samples can be measured within a short time, the supplied vapour concentration is not known and varies during the vapour supply. The vapour concentration at the outlet of the bottle gradually changes until it reaches the liquid vapour equilibrium.

The vapour concentration profile sometimes influences the waveform of the sensor response, which is a convolution of this profile and the sensor impulse response. When a sufficiently narrow vapour pulse is supplied to the sensor, the sensor response is not influenced by the concentration profile. The pulse vapour supply method can therefore be used to ignore the influence of the concentration variation during vapour supply [12]. This method of the pulse measurement is, really, a trade-off between a static and a dynamic measure of the headspace. Actually, the delivering of the sample is driven by a carrier gas flow, but the measure is performed in an almost static condition.

The sensor response is sometimes influenced by its position inside the cell, especially when dense vapour with a high boiling point is supplied. The cell should therefore have an internal volume as small as possible to minimize any effect due to the sensor location. The sensor response however, is sometimes slow enough to ignore that effect. Anyway, this last consideration, put in evidence the importance of a chamber geometry which allows a sensor response independent of location [12].

The last problem, not the least, is temperature; if the sensor responses are likely to be influenced by temperature, the sensor cell can be kept at constant temperature by use of a peltier device or immersion in a thermo bath.

2.3.2 On-line sampling

The need of on-line sampling can be due to different reasons. Sometimes it is an answer to one of the problems reported before, sometimes it is due to the peculiarity of certain samples, which need on-line measurement within their proper environment, in order to continue to be representative of their classes. This is the case of breath sample, collected and measured in a bag. Moreover there is the environmental monitoring application, which sometimes asks for direct sampling in the object location. These two cases are treated in the chapters three and four respectively.

In the following, some of these on line techniques are reported.

In the diffusion method, diffusion of vapour from a tube of accurately known dimensions is measured. Low concentration are usually measured [13].

The liquid in the reservoir is allowed to evaporate and the vapour slowly diffuses from a reservoir through the diffusion tube into a flowing gas stream at a constant rate. The resultant mixture concentration is determined by the ratio of the diffusion rate to that of the flowing gas stream. The actual concentration sometimes deviates from the formula now described, when the vapour above the liquid is not saturated. An alternative method is to precisely measure the mass change of the liquid reservoir during the constant period using a balance. The reduction in the amount of liquid over a certain time indicates the diffusion rate.

The permeation method is similar to the diffusion method, except that a permeation tube is used. Liquefied gas, when enclosed in an inert plastic tube, may escape by dissolving in and permeating through the walls of the tube.

In the bubbler method vapour is generated by bubbling in a bottle containing the sample to measure. A carrier gas such as dry air is passed through the liquid in the bottle, and takes away the generated vapour. Although it is easy to obtain the vapour by this method, several point should be taken into account. The headspace over the liquid sample sometimes does not saturate [14,15].

Another useful method makes use of sampling bags. The use of this technique depends on the kind of sample. In case of gas samples, such as breath or environmental air, either the patient or a pumping system fills the bag before to connect it directly to the e-nose. In case of liquid sample, vapour is generated after the sample is injected to the large air volume sampling bag by syringe, and then evaporated. In the two case the vapour in the bag is then sucked out using a pump, and introduced into a sensor cell. The material of the sampling bag should be carefully selected to avoid water and other molecules permeating through. The generation of vapours can also occur in certain type of plastic bag. Adsorption of the sampled vapour inside the bag cannot be ignored in case of low concentration. There are many materials considered as inert which present the quoted characteristic of non interacting with the sample and of non adsorbing or desorbing vapour, such as telfon, tedlar, fluorine-containing resin bag [16,17].

The flow-type system, with headspace formation or on-line, mentioned above are closed system, with the exception of environmental monitoring application. This particular case of on-line measurement, described in chapter five, is used for odour mapping of critical areas such as an industrial district or a zone surrounding a wastewater treatment plant.

Moreover, direct exposure of the sensor to the gas, is often performed when the rapid concentration change in an open system should be captured. However, the sensor response does not correspond to instantaneous gas concentration due to its response delay, even if it is open to the ambient atmosphere. The sensor dynamics can be analyzed when both sensor response and gas concentration change are simultaneously obtained.

2.3.3 Static sampling system

The fundamental static system measures the steady-state response of a sensor to a vapour at constant concentration and at a constant temperature. This will be the object of chapter six, in which designing, realization and testing of an electronic nose for static sampling is described. Anyway, in the following of this paragraph, a very short synthesis of this method is illustrated.

The sample could be a liquid, a gas or solid. In each of these cases the sample is inserted in a tight closed chamber and the measure is performed connecting the sensor chamber with the sample chamber at the moment the equilibrium is reached. In case of liquid, gas or headspace of solids, there is also the possibility to insert a sample at fixed concentration by the mean of a syringe.

Anyway, the manual method of syringe injection is a basic method, but it is possible to automate such a sampling method using a mass flow controller.

The chamber is typically made of inert material, to avoid vapour adsorption onto the internal wall.

2.3.4 Adopted sampling methods

In the adopted sampling techniques scheme of figure 2.3.1, the strategies experimented in this thesis are underlined. In figure 2.3.2 the four experimental sets are illustrated.

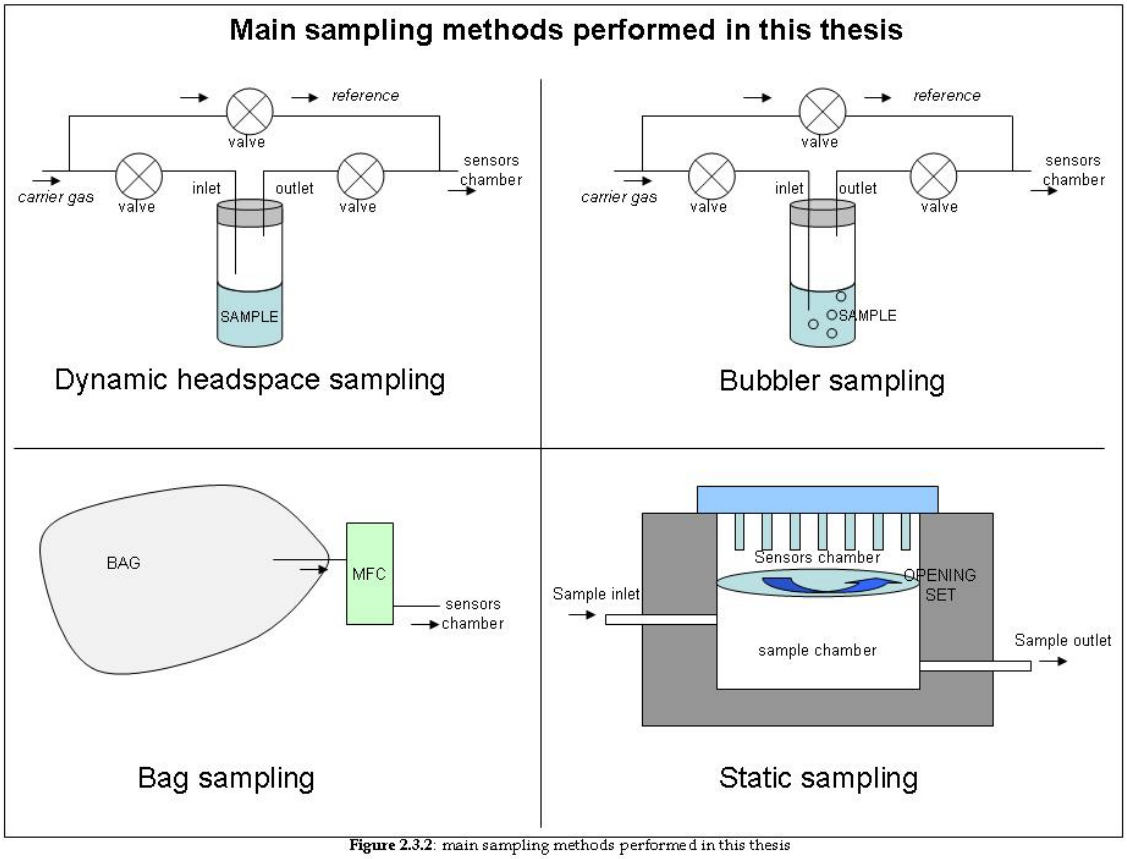


Figure 2.3.2: main sampling methods performed in this thesis

2.4 Sampling protocol chain

Before starting to treat each sampling protocol applied to the three particular situations, and after an overview of normally used techniques, there is the need to discuss the sampling chain. Actually, to better understand how the methods listed above can be specialized to different experimentations, it is necessary to underline all the critical aspects that ask a particular sampling strategy to be modified or adapted to carry out the application requirements.

There are different steps in a sampling protocol, starting from the sample source, passing through the sample itself and its preservation and enclosing, to finish with its delivering in the chamber and the effective measurements. All this steps are underlined in figure 2.4.1, which puts in evidence all the critical aspects. Their consideration makes possible to elaborate a method to refine a sampling protocol for each experimental necessity.

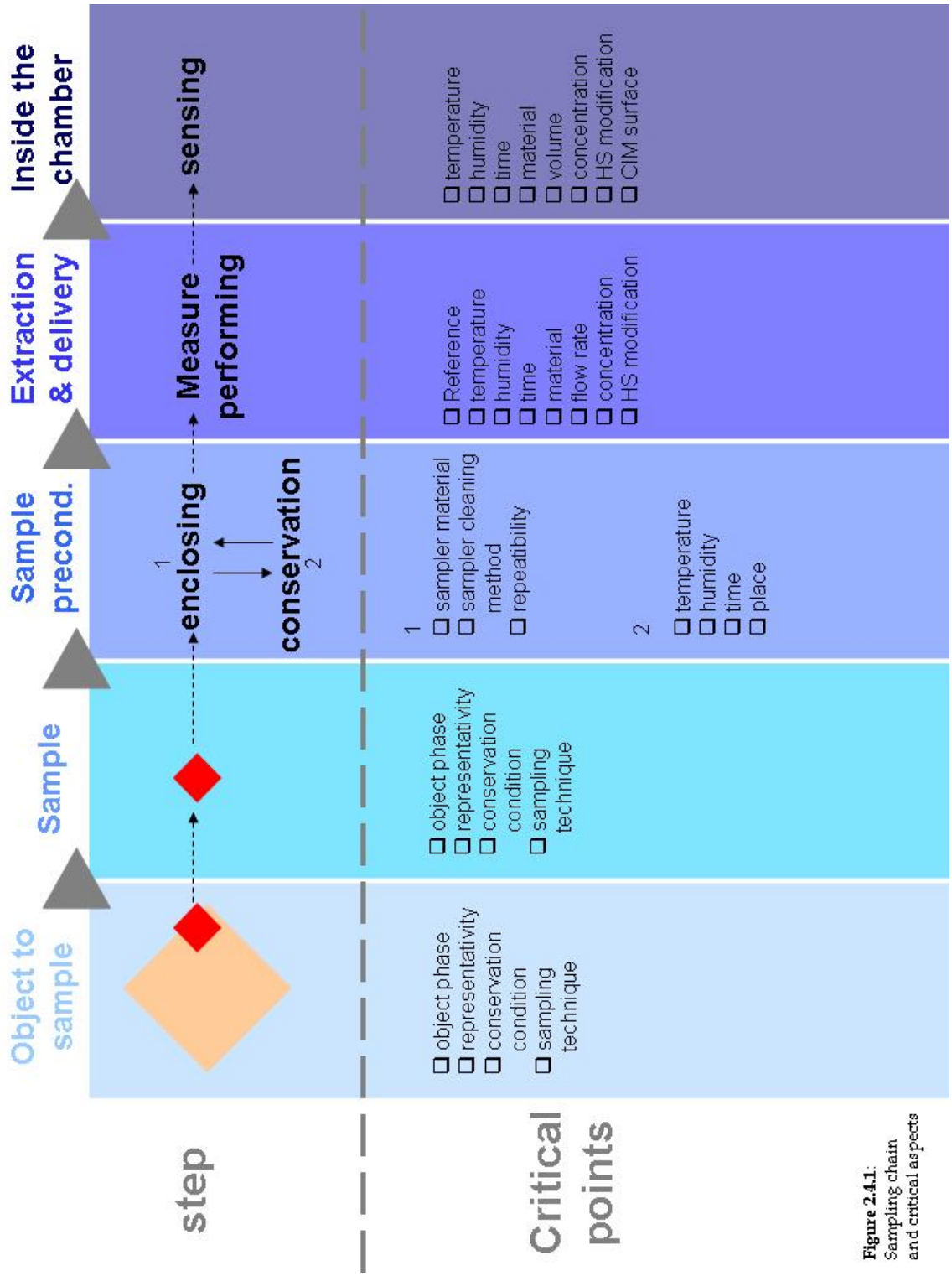


Figure 2.4.1: Sampling chain and critical aspects

The first two steps are fundamental for the object of the measurements. It is possible to control all the physical and chemical parameters during the measure, it is possible to choice the correct materials and to perform a measure in optimal conditions of time and flow, but if the sample is not representative of its class the instrument is not measuring something useful for the experimentation target.

So it is important to choice a 'source', from which the sample is extracted, which presents all the characteristics of the class at which the experiment is interested in, and without peculiarity that could hide the elements necessary for a correct identification inside the target of the experiment. This could seem trivial, but it is obtained only after a synergy between the sensorist performing the measurements and the samples supplier, very often experts in various fields very far from the sensorist normal experience.

To satisfy these needs, it is necessary to take under control environmental parameters like temperature and relative humidity, to ensure the repeatability of the correct procedure for the sample extraction, and the goodness of sample preservation conditions.

In these first steps two crucial parameters like temperature and relative humidity has been quoted. They are so important to be critical in all the enrolled steps of the sampling chain, their changing being fundamental in modifying peculiarity of the sample, delivering and measuring conditions.

The phase of the sample is another basic characteristic conditioning all the steps. Gas, liquid or solid sample ask for specialized extraction methods, particular sampler, dedicated control and delivery systems, precise controlled conditions and sampling strategies for the final measurement.

Looking at the phase of the sample it is important to remember that the instrument object of this thesis is a gas sensor array, so, obviously the sample flowing or diffusing to the sensors chamber is always a gas sample. This characteristic makes the phase of the real object of the measurement more and more critical. Actually, depending on the nature of the measurand, the choice of the tool to enclose of the analyte, means the choice for its sampler, and the selected strategy for the measurement (static, dynamic, on-line), are important to ensure again that the gas measured is really representative of the original sample.

If the objective of the experimentation is to characterize and classify the odour of solid or liquid samples, the relative gas sample is expected to mirror the concentration and composition of these sample; moreover, considering that this gas sample is expected to represent the 'odour' of the sample, the best choice for

sampling and delivery systems is to put the measurand in the better condition to release its natural aroma.

A last important thing to be considered is the reference gas. Almost all sensors used indeed performs a differential measurements; it means that the response due to a certain sample is given as the difference between the sensor response in presence of the sample and the steady state reached before the starting of the measure when the sensor is exposed to a reference gas. This steady state is a reference chosen at the beginning of the experimental session and it must remain the same for all the measures that has to be compared.

There are three main choices that can be done for the reference: the response to a certain gas chosen as reference gas (often an inert gas), the response to environmental air, the response to another sample respect to which the difference has to be evaluated. In the first two cases the response to the reference is called cleaning phase.

The choice of an inert gas as reference, is also due to the need of desorption to recover sensors initial unloaded state. The possible reference gases are Nitrogen, environmental air filtered either by silica gel or CaCl, and, in particular cases, environmental air added with a fixed content of relative humidity. This last case is useful when the relative humidity contribute in the measure of the sample has to be completely eliminated, for example in case of environmental air quality monitoring. The principle is based on the fact that if dry air is used, its content of humidity is very low, so in the following differential measurement there is a large contribute (not interesting) due to the sample humidity and a small contribute (interesting) due to the odour of the sample. This could be not so critical in case of sample in which RH is under control, but it becomes a problem in case of environmental measurement on-field, where humidity is connected to the metereological conditions, so out of control. Figure 2.4.3 illustrates the reference methods just discussed.

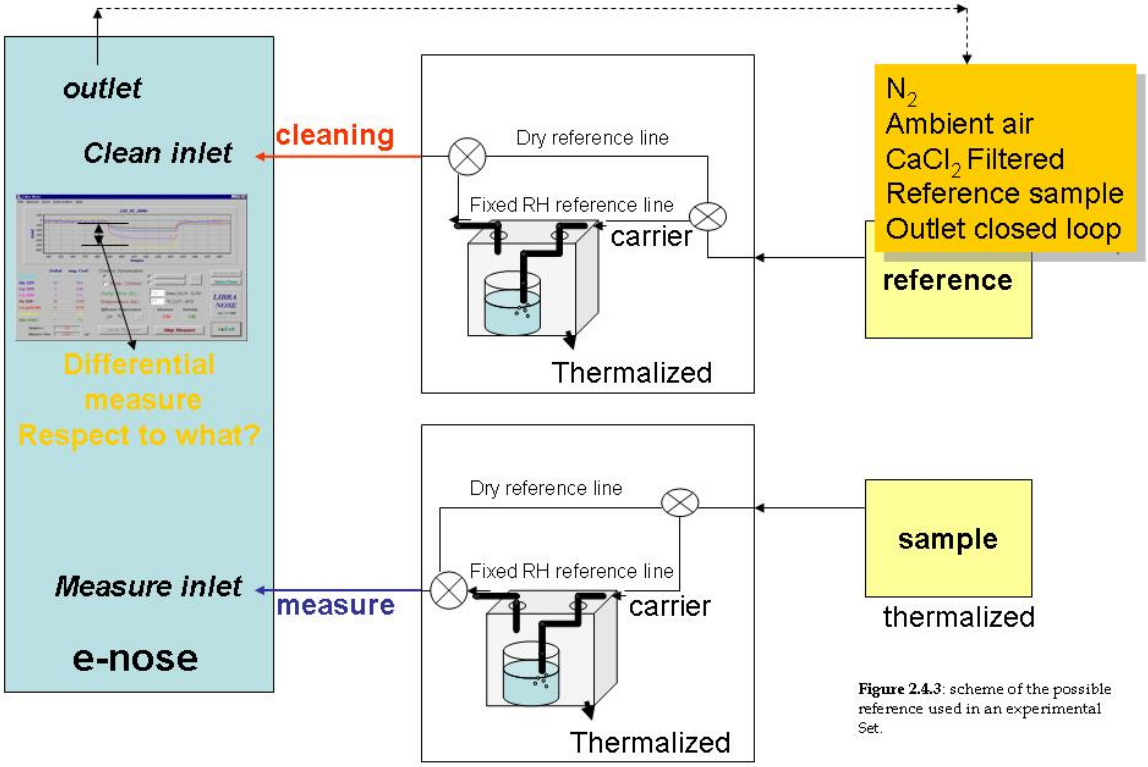


Figure 2.4.3: scheme of the possible reference used in an experimental Set.

All these considerations cannot lead to a general lists of solutions, because of their dependence on a series of characteristics strictly connected to particular situations and experimentation. Anyway, the evaluation of these problems, lead to the drafting of a 'chart of rules' to take in account for the elaboration of the optimal sampling protocol. The actualization of these rules to the running experiment transform this list in a sort of a measure identity card (MIC: Measured Identity Card).

This is a list of all the critical points emerged in the discussion of the sampling chain. So it contains all the problems to face in the different steps of the experiment, and it will be used in the following chapters to synthesized the elaborated solutions. A MIC is a good means for the starting and for the ending phase of an experimental session. At the beginning, when it is only a list of problems, it ensures that any possible critical point is taken into account. A the end of the experiment, when it consists in a list of applied solutions, it ensures the possibility to repeat the same experiment to test its reproducibility.

In figure 2.4.4 the MIC structure at the starting phase is reported.

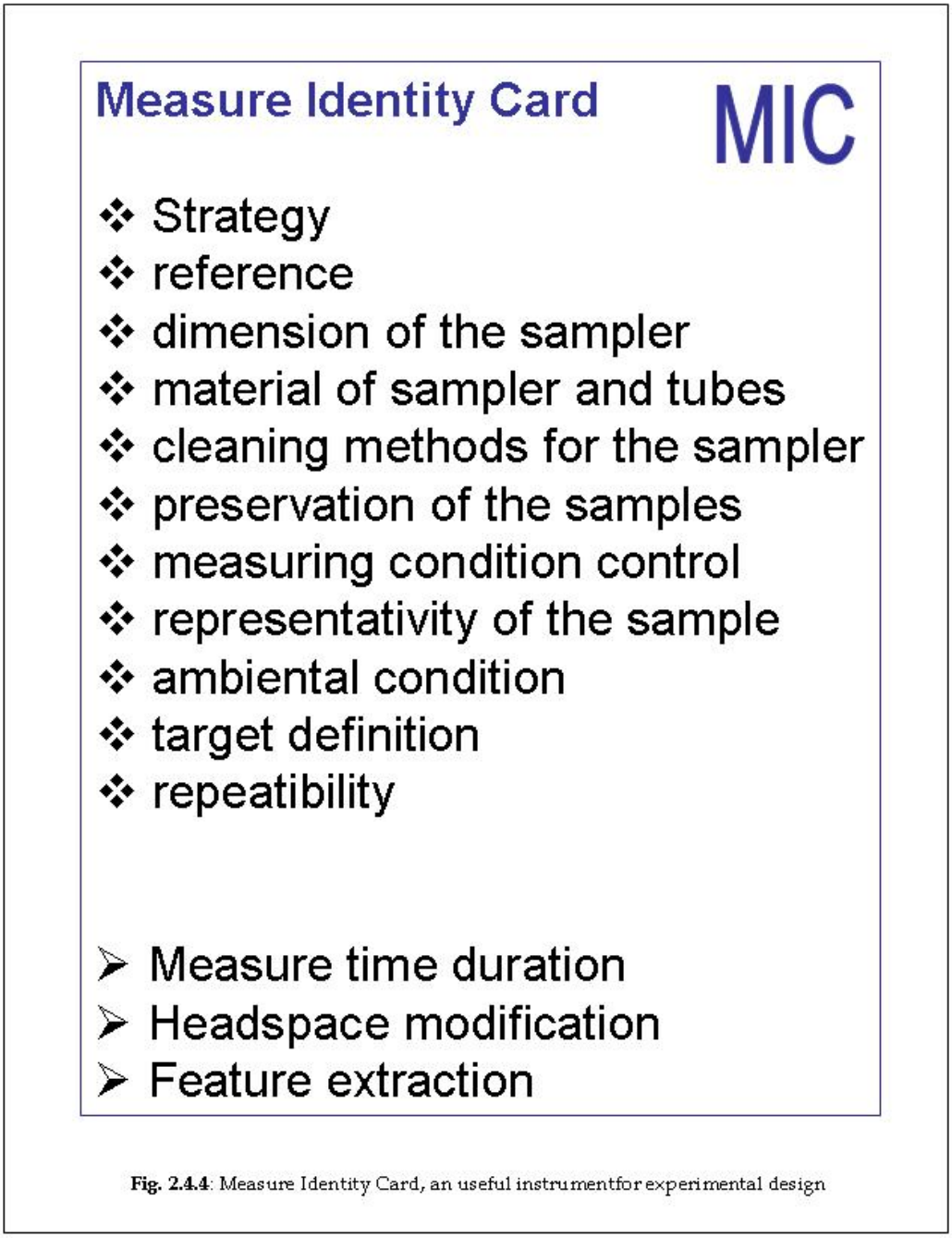


Fig. 2.4.4: Measure Identity Card, an useful instrumentfor experimental design

2.5 References

[1] T.C. Pearce, S. S. Schiffman, H. T. Nagle, J. W. Gardner, '*Handbook of machine Olfaction*' , Wiley-VCH (2003).

[2] T. Nakamoto, A. Iguchi, T. Moriizumi, 'Vapor supply method in odor sensing system and analysis of transient sensor responses' , *Sensors and Actuators B* 71 (2000) 155-160.

[3] M. Markelov, O. A. Bershevits, 'Methodology of quantitative headspace analysis using vapour phase sweeping', *Analytica Chimica Acta* 432 (2001) 213-227.

[4] N. H. Snow, G. C. Slack, 'Head-space analysis in modern gas cromatography', *Trends in analytical chemistry*, vol.21, 2002.

[5] P. Mielle, F. Marquis, 'An alternative way to improve the sensitivity of electronic olfactometers', *Sensors and Actuators B* 58 (1999) 526-535.

[6] P. Mielle, F. Marquis, C. Latrasse, 'Electronic noses: specify or disappear', *Sensors and Actuators B* 69 (2000) 287-294.

[7] T. Hamacher, J. Niess, P. S. Lammers, B. Diekmann, P. Boeker, 'Online measurement of odorous gases close to the odour threshold with a QMB sensor system with an integrated preconcentration unit', *Sensors and Actuators B* 95 (2003) 39–45.

[8] S. Roussel, G. Forsberg, P. Grenier, V. Bellon-Maurel, 'Optimisation of electronic nose measurements. Part II: Influence of experimental parameters' , *Journal of Food Engineering* 39 (1999) 9-15

[9] S. J. Gregg, K. S. Sing, 'Adsorption, Surface Area and Porosity' , Academic Press, 1967.

[10] S. Roussel, G. Forsberg, P. Grenier, V. Bellon-Maurel, 'Optimisation of electronic nose measurements. Part I: Methodology of output feature selection', *Journal of Food Engineering 37,* 207-222.

[11] Nakamoto I., Isaka Y., Ishige T., Moriizumi T., 'Odor-sensing system using preconcentrator with variable temperature', *Sensors and Actuators B*, 69 (2000) 58-62.

[12] A. Iguchi, T. Nakamoto, T. Moriizumi, *Sensors and Actuators B* (2000) 155.

[13] T. Nakamoto, T. Fukuda, T. Moriizumi, *Sensors and Actuators B*, 3 (1991) 1.

[14] M. Ohnishi, T. Ishibashi, K. Kijima, C. Ishimoto, J. Seto, Sensors and Materials, 1 (1992) 53.

[15] C. Di Natale, A. Macagnano, E. Martinelli, R. Paolesse, G. D'Arcangelo, C. Roscioni, A. Finazzi Agrò, A. D'Amico 'Lung cancer identification by analysisof breath by means of an array of non-selective gas sensors'. *Biosensors and Bioelectronics* 18, 1209–1218 (2003).

[16] Phillips M. 'Method for he collection and assay of volatile organic compounds in breath'. Anal Biochem 1997;247: 272–8.

ODOUR AND MEDICAL DIAGNOSIS

Some men see things as they are and say, "Why?" I dream of things that never were and say, "Why not?'

Gorge Bernard Shaw

"Molti vedono le cose così come sono e si domandano il perché.
Io sogno di cose che non sono mai state e mi domando "Perché no?"

3

3.1 Medical diagnosis and odour

Nowadays prevention in medicine has gained a strategic importance. Patients are demanding rapid and early qualitative diagnosis of diseases. As an example early discrimination between different infections is important to facilitate rapid treatment as part of a preventative health strategy [1].

To this regard it is well known that microbial species produce a range of volatile compounds. These are the components of an odour that could be perceived and recognized as characteristic of particular disease and used for monitoring the health of individuals.

The importance of odours in humans health was well understood by ancient medicine, such as Chinese medicine, that used to diagnose diseases analyzing the body odor. Conventional western medicine, on its own, recognized that some pathologies produced unpleasant characteristic odors.

Many studies, based on analytical tools such as gas chromatography (GC) or GC linked with mass spectrometry (GC–MS), have shown that microorganisms produce many volatile organic compounds, including alcohols, aliphatic acids and terpenes, some of which have characteristic odours.

Other work has shown that the type of culture media and the age of the culture, as well as the microbial species, all influence the amounts and patterns of the volatile compounds that are produced.

As these patterns are characteristic for certain types of infectious microorganisms, they can potentially be used as markers of disease, becoming an alternative diagnostic method.

In the past years, GC and GC–MS techniques have already been used to monitor the production of volatile compounds as an aid to the clinical diagnosis of aerobic and anaerobic bacterial infections and cardiopulmonary disease, and have been used to analyse several substrates, including human pus, urine, blood plasma and alveolar air.

About human skin, it is known that its smells unquestionably influence human behaviour as well. Human odor results by the combined action of both

skin gland and sectating organic compounds and whose regulation is subject to human hormonal control, and bacterial populations localized at skin surfaces, which live by metabolizing and transforming organic compounds that they are able to absorb from their external environment [2].

The potentiality of GC/GC–MS in medical diagnosishas been thwarted by a number of limitations: such as the cost of analytical equipments, the expertise required to operate such instruments and the length of time required to obtain results. Early in this research area the question arose as to whether chemical reactions between volatile markers and various sensors could be amplified and be sensitive enough to enable qualitative differences to be measured between the markers at relatively low concentrations. It is a crucial point already treated in the previous chapters: odour could be a meaningful synthesis of the properties of a disease revealed by a particular interaction or 'cooperation' of some components sometimes present in a too low concentration.

For prevention and early treatment strategies to be implemented, it is essential to obtain relevant rapid results in a useable format. Moreover, in many medical applications, a non-invasive investigation method is desirable. Can volatile fingerprinting and electronic-nose systems detect and discriminate between pathogens? In the two following tables there are listed the basis on which giving a positive answer to this question. In table 3.1.1 a list of key volatiles associated with different disease types according to different GC-MS analysis, illustrates the potentiality of such a diagnostic means, and table 3.2.2, summarizing the *in vitro* and *in situ* results obtained using this technology, confirms the capability of the methods.

In vitro studies have shown the possibility to discriminate between different aerobic bacteria, such as *Helicobacter pylori*, *Escherichia coli* and *Enterococcus* species that are present in samples, both alone and as a mixture of the three species, on the basis of differences in the amounts of terpenes, trimethylamine and ketones produced [18].

Mycobacterium tuberculosis, the causative agent of tuberculosis, has been detected in cultured sputum samples either directly or following treatment with enzymes to enhance *M. tuberculosis* growth and volatile production [19].

SAMPLE	DISORDER/INFECTION	VOLATILE COMPOUNS
Microorganism-associated disorders		
Urine	Urinary tract infection	Isovaleric acid, alkanes [4]
Intraperitoneal fluid	Aerobic Gram-negative bacteria	Terpenes, ketones [5]
Intraperitoneal fluid	Anaerobic bacterial infections	Acetic, butyric acids [6]
Human pus	-	Isobutyric, isovaleric, isocaproic acids [7]
Other disorders		
Human breath	Breast cancer	Alkanes, monomethylated alkanes [8]
Human breath	Lung cancer	Alkanes, monomethylated alkanes [9]
Human breath	Acute asthma	Pentane [10]
Urine	Metabolic disorders	Isovaleric acids [11]
Alveolar air	Hepatic coma	Methyl-mercaptan [12]
Alveolar air	Rheumatoid arthritis	Pentane [13]
Alveolar air	Schizophrenia	Pentane, carbon disulphide [14]
Alveolar air	Ketosis	Acetone [15]
Alveolar air	Cardiopulmonary disease	Acetone, ethanol [16]
Blood plasma, cerebrospinal fluid	Hepatic encephalopathy	3-methylbutanol [17]

Table 3.1.1: Summary of key volatiles associated with different diseasetypes analysis by GC-MS (table extracted from the article [1])

Anaerobic bacteria such as *Clostridium* species and *Bacteroides fragilis* have been successfully differentiated in culture on the basis of the discrimination of volatiles, such as isobutylamine, acetic acid and butyric acid.

In other work, samples from patients with urinary tract infections (UTIs) and tuberculosis were identified correctly in 90–99% of cases compared with traditional culture techniques.

Additionally, six different bacterial species responsible for eye infections – *E. coli*, *Haemophilus influenzae*, *Moraxella catarrhalis*, *Pseudomonas aeruginosa*, *Staphylococcus aureus* and *Streptococcus pneumoniae* – were successfully discriminated into six different classes (98% success) using data that were obtained with a handheld portable electronic nose [20].

A case study on bacterial vaginosis in the United Kingdom has also shown that a conducting polymer- based sensor array successfully diagnosed 89% of test subjects as being positive or negative for both bacterial and yeast infections [21].

Sample type	**references**
In vitro samples	
Bacterial classes successfully discriminated by groth phase and type	[22,23]
Helicobacter pylori and gastroeosophageal isolates cross-validation successfull	[18]
Anaerobic bacteria differentiated on agar media using PCA	[24]
Bacteria, yeasts and filamentous fungi differentiatedc using PCA and cluster analysis	[25,26]
Bacteria detected in blood cultures	[27]
Fungal spoilage detected more rapidly than traditional techniques	[28]
Toxigenic strains of moulds successfully separated on media and on bakery prodsucts	[29,30]
Different agaricus species discriminated successfully	[31]
In situ clinical samples	
Contact dressing from leg ulcers: bacterial infection detected	[32]
Vaginal swabs: bacteria detected	[33]
Urine: urinary tract infection detected	[19]
Vaginal swabs: bacterial vaginosis diagnosed	[21]
Sputum: tuberculosis diagnosed	[34]
Urine containing blood: haematuria detected	[35]
Breath samples: respiratory infection detected	[36]
Breath samples: lung cancer detected	[37]
Breath samples: diabetes detected	[38]

The Electronic Nose developed at the University of Rome 'Tor Vergata') has been used in four different medical applications.

One of these studies has shown that patients with kidney disorders produce characteristic volatile compounds, which can be a useful tool in the diagnosis and control of renal dialysis [35].

Additionally studies, by the means of the same device, have shown that lung cancer can be detected by breath analysis using non-selective gas sensors; this last experimentation is the central object of this chapter [37].

Another medical application using the Libra Nose has reached good results in the study of Schizophrenia.

The case study here reported is about human skin odor. Human skin is known to be colonized by a huge number of bacteria that live as commensals on the surface and within the follicles [2]. It is possible to describe the basic pattern of colonization of a healthy human skin. Variations of this pattern may be observed: dry skin supports a low level of colonization, while moist areas provided with sebaceous and apocrine glands are the most heavily populated.

The microbial flora usually localized on the skin appear to have several functions, the most important of which is probably the defense against pathogenic bacterial and micotic infections.

In an adult subject, skin living microorganisms may be mainly observed, at different concentrations, in the following sites: the nasal vestibule, the external auditory meatus, the axilla, the perineum and the groin, the scalp, the face and the limbs. The number of microorganisms changes with age, sex and race. Sweat is sterile and mostly odourless by its own when secreted. The main role of the bacteria in the odor formation appears to be the breakage of the precursor–odorant complex and the cleavage of the covalent bonds holding the acid molecules to the precursors. As an example, axillary odor is a distinctive malodorous scent of adults generated when Gram-positive microorganisms interact with the apocrine sweat. Recently numerous studies demonstrated that the characteristic human axillary rated aliphatic acids, alcohol, carbonyls and some steroids.

To conclude this survey on olfaction diagnostic potentiality, it is necessary to underline the importance of such a technology, because of its peculiarity being less invasive, often cheaper and probably more indicated for early diagnosis, than the classical diagnostic tools.

These considerations lead to the challenge to exploit this techniques in that medical fields in which an early, cheaper and non invasive diagnosis is desirable, such as the lung cancer, objective of this chapter. For this reason, in the following chapter, the overview on 'smell diagnosis' will focus on breath analysis, the obvious means to study for lung cancer detection.

3.2 Breath analysis, an overview

Within the possible biological means on which performing an olfactive diagnostic test, breath is one of the less invasive to sample, and it is a rich information source about the most complex and at the same time harmful diseases such as lung cancer. Indeed, analysis of exhaled air enables the observation of many biochemical processes in the body [39].

The ancient Greek medicine practitioners already knew that some diseases could be diagnosed from the characteristic odour of patients' breath. Modern breath analysis started in the 1970s when Pauling et al. determined more than 200 components in human breath using gas chromatography [40].

This list of compounds opened a window on a rich source of clinical information. Nonetheless, until now, breath analysis has faced, along more than thirty years, problems such as substance separation and identification, correlation of the identified compounds with certain diseases and the elaboration of an effective sampling method.

In the first years after Pauling experiments, the main problems consisted of substance separation and identification. The technical progress of analytical methods achieved in the 1980s and 1990s partially solved the problem; and issues concerning the physiological meaning of the volatile substances and correlations of breath markers with patients' clinical conditions became more and more clear [41]. Anyway, the simple finding of a possible new diagnostic tool such as breath analysis, is far from its introduction in the clinical practice, and this is yet a challenge of today.

The fundamental components of breath consist in a mixture of nitrogen, oxygen, carbon dioxide, water, and inert gases; In the long list of components occurring in concentrations in the nmol/l–pmol/l (ppbv–pptv) range, has to be searched the individual healthstate information, that results well hidden in the bulk matrix of breath . Far more than 500 of these compounds have been described until now.

These volatile substances may be generated in the body or may be absorbed as contaminants from the environment. Both these two groups of substances, exogenous and endogenous ones, has an informative content useful to be investigated for different scopes. Endogenous molecules, especially halogenated organic compounds, may be analyzed for environmental or expositional issues to assess compound specific uptake into the body and elimination from the body. In order to monitor metabolic or any pathologic processes in the body, endogenous substances have to be determined.

The following of the paragraph will focus on endogenous compounds and in particular on alveolar breath. Anyway, before penetrating the investigation of alveolar VOCs, a survey on endogenous compounds will be presented. These endogenous compounds include inorganic gases, such as NO, CO, volatile organic compounds (VOCs) such as ethane, pentane, acetone, isoprene, and other normally nonvolatile substances such as isoprostanes, peroxynitrite or cytokines that can be determined in breath condensate.

During the last two decades, NO has been recognized as a mediator of numerous physiological processes and as a marker of airway inflammation. Exhaled NO can now be determined by commercially available devices and has therefore been investigated in a large number of studies. Exhalation kinetics and relationships with airway diseases have been described in detail. A number of normally nonvolatile substances, such as isoprostanes, cytokines, leukotrienes, or hydrogen peroxide can be found in breath condensate. Quantitative analysis of breath condensate is hampered by a number of serious problems. There is no clear relationship between assumed alveolar or airway concentrations and substance concentrations in the condensate. Furthermore, some of these compounds only

have limited stability. Volatile organic substances such as ethane, pentane, isoprene, or acetone can provide insights into different biochemical processes in the healthy and the diseased human body. In contrast to the non-volatile substances in breath condensate exhalation, kinetics of volatile organic substances can be approximated according to substance solubilities. In addition, for most of the exhaled organic compounds, there is no problem of stability.

Alveolar breath VOCs origin and pathway

Moreover, speaking of breath VOCs, a remark has to be specified about breath source and its potential meanings. Actually, at a certain point we referred to breath with the specification of 'alveolar'. This is a fundamental distinction from the point of view of healthstate monitoring of an individual.

Alveolar breath is a distinctive gas whose chemical composition differs markedly from inspired air. Volatile organic compounds are either subtracted from inspired air (by degradation and/or excretion in the body) or added to alveolar breath as products of metabolism. Some features of this transformation has been well understood form many years [42].

Endogenous markers, which are commonly used for diagnostic purposes, are hydrocarbons like ethane, pentane and isoprene; oxygen-containing compounds like acetone, acetaldehyde, methanol, ethanol, and 2-propanol; sulphur-containing compounds like dimethylsulfide, methyl, and ethyl mercaptanes; and carbon disulfide and nitrogen containing substances like ammonia and dimethyl/ trimethylamine. In order to assess the physiological meaning and the diagnostic potential of these substances, the biochemical pathways of generation have to be known. For this reason, the following figure 3.2.1 represents this pathways: gaseous and capillary VOCs equilibrate rapidly in the pulmonary alveoli, and the dominant process varies with the phase of respiration. During the inspiratory phase, room air VOCs equilibrate with pulmonary venous blood, while during the expiratory phase, pulmonary arterial blood equilibrates with VOCs in alveolar breath. Extra-pulmonary input of VOCs

is primarily from endogenous synthesis, and extra-pulmonary output of VOCs is predominantly by metabolism in the liver and excretion in the kidneys.

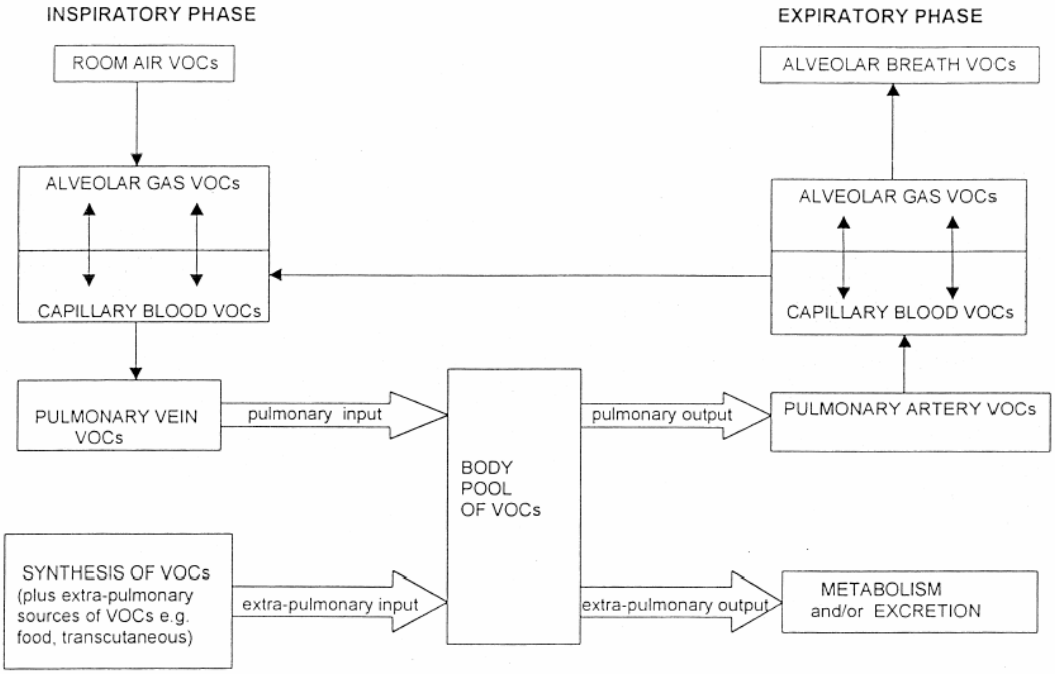


Fig. 3.2.1: Pathways of VOCs through body compartments [M. Phillips et al., Journal of Chromatography, 729 (1999), figure 3, page 84]

To complete this illustration, the concept of the 'body pool of VOCs' illustrated in figure 3.2.2 is introduced. The VOC kinetic is described with a water flow analogy: a VOC enters the body pool either from the inspired air or from synthesis in the body (ignoring minor inputs such as VOCs in foodstuffs). The VOC leaves the body pool either by clearance (metabolism and/or excretion) or else in the breath. If the VOC is neither synthesized nor cleared from the body, then the amount leaving in the breath must equal the amount entering from inspired air, and the alveolar gradient (amount in breath minus amount in air) will be zero. If the VOC is synthesized in the body but not cleared, more leaves in the breath than is inspired from the air, and the alveolar gradient becomes positive. Conversely, if the VOC is cleared from the body but not synthesized, less leaves in the breath than is inspired from the air, and the alveolar gradient becomes negative. Hence, if a VOC is both synthesized and cleared in the body, the alveolar

gradient will vary with their combined effect: positive if synthesis is greater than clearance, and negative if clearance is greater than synthesis.

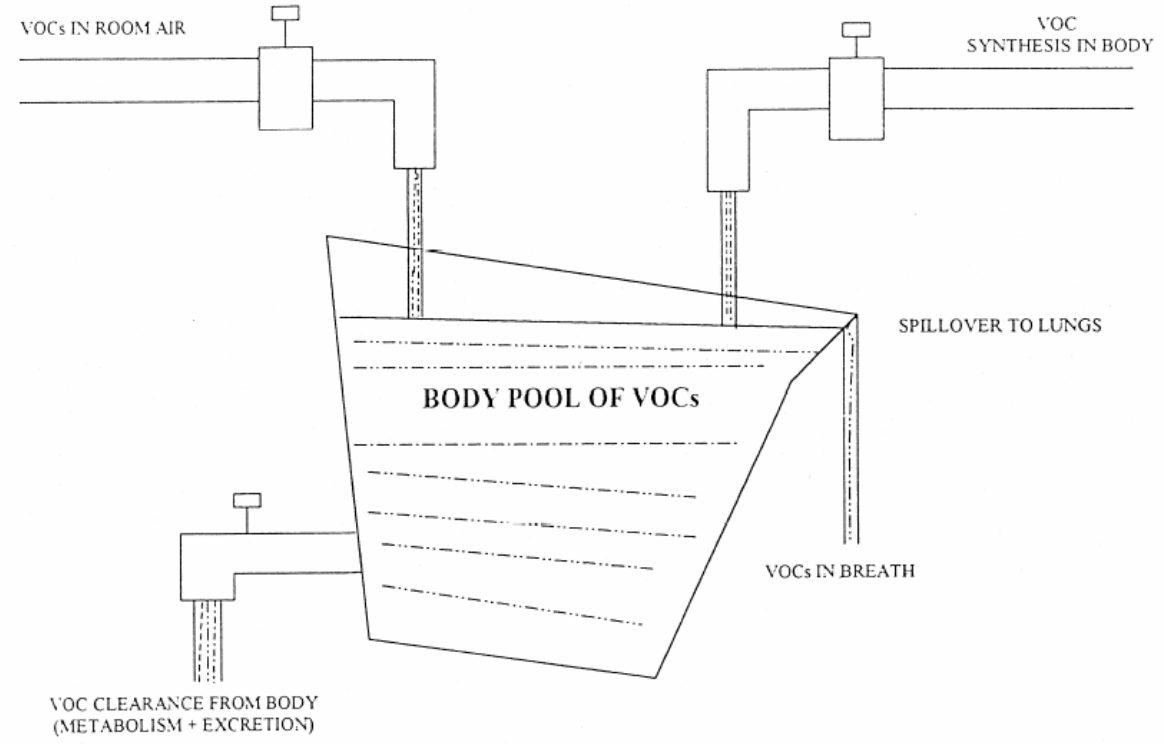


Fig. 3.2.2: Body pool of VOCs [M. Phillips et al., Journal of Chromatography, 729 (1999), figure 4, page 85]

Studying the particular pathway and production mechanisms of the different VOCs, it is possible to connect their presence or their concentration to a particular disease. In the following some examples will be presented. Anyway it is useful to remember that in the case study illustrated below the experiment expectation is the confirmation of some theories elaborated on the disease and the VOCs related with its mechanisms. Often the scope of new technique consists in the finding of something of unusual, as an abnormal concentration or the presence of an unusual compound. This is the case of lung cancer, the main matter of this chapter, and this is the reason of a separated discussion of lung cancer case in the next paragraphs.

First of all a general treatment of origin and properties of endogenous volatile organic biomarkers will be studied, in order to trace a connection between these properties and certain disease. This is a study from the point of view of the VOCs. As consequence, an overview from the point of view of the disease will be

presented, with different investigations from cross-sectional studies and from longitudinal studies.

Animal and clinical studies demonstrated a close correlation between clinical conditions with high peroxidative activity and the exhalation of ethane and pentane [43].

Breath markers often proved to be more sensitive than the serum markers. Although there are other potential sources of hydrocarbons in the body, such as protein oxidation and colonic bacterial metabolism, these are apparently of limited importance and do not interfere with the interpretation of the hydrocarbon breath test for ethane and pentane [44].

Exhaled pentane concentrations should be interpreted cautiously whenever liver function varies between patients or during the sampling period.

Hydrocarbons as stable end products of lipid peroxidation show only low solubility in blood and are therefore excreted into breath within minutes of their formation in tissues. Hence, exhaled concentrations of ethane and n-pentane can be used to monitor the degree of oxidative damage in the body [45].

The parallel decrease in isoprene secretion suggests that breath isoprene is derived from the cholesterol synthesis pathway in humans in vivo [46]. A small fraction of exhaled isoprene may be of bacterial origin [47].

There is experimental evidence that isoprene exhalation may be related to oxidative damage to the fluid lining of the lung [48] and the body [49].

Concentrations of breath isoprene seem also to be age dependent, indeed they are significantly lower in children.

Acetone is one of the most abundant compounds in human breath. Its concentrations are increased in patients with (uncontrolled) diabetes mellitus [50].

Ethanol concentrations in breath of human subjects are normally very much lower than the levels found in human breath after alcohol ingestion. The potential source of endogenous ethanol is the intestinal bacterial flora [51].

Under normal conditions, concentrations of sulphur-containing compounds in human blood and breath are very low. Impairment of liver function increases the level of sulphur-containing compounds.

The origin of breath methanol may also be the intestinal bacterial flora. Compounds like ethyl mercaptane, dimethylsulfide, or dimethyldisulfide are responsible for the characteristic odour in the breath of cirrhotic patients.

The characteristic odour of uremic breath due to elevated levels of dimethylamine and trimethylamine has been known for a long time. In 1977, amines were conclusively identified and quantified in human breath [52].

Ammonia could also be identified in the breath of normal and in higher concentrations in uremic patients [53].

The validity of the correlation between the great number of new marker substances found in exhaled breath and patients' clinical conditions is checked by the explanation of the pathways of generation of these VOCs. In the cases mentioneed above the origin of compounds is known and the measure is a confirmation; sometimes the process is the inverse, and this is often the case of cross-sectional and longitudinal studies, of which a brief survey is reported below.

A great number of clinical studies have been undertaken to put in evidence correlations between chemical composition of exhaled air and clinical status. Main targets of these investigations were different lung diseases, inflammatory and malignant processes in the body. As mentioned before two possible experimental typologies exist in clinical studies: longitudinal and cross-sectional studies. In cross-sectional studies exhaled biomarkers are studied as a function of disease. A control group is compared to a patient or diseased group. Breath markers are analyzed in order to identify qualitative or quantitative differences between the two groups. Differences established in this way should be large enough to enable a clinically relevant predictive use of breath markers. In longitudinal studies breath markers are observed during the course of a disease or an intervention within one patient group. Variation of results is a lesser problem than in cross-sectional investigations, because each subject can act as his/her own control.

One of the main problems of cross-sectional studies is the large variation of results. Due to interindividual differences, concentrations of some volatile markers vary considerably even in the nondiseased state [42]. Nevertheless, some correlations between patients' clinical status and breath marker concentrations

have been described. If one looks upon more than one substance, specific sets of volatile markers can be defined which may have predictive value even in complex diseases or disease states, such as oxidative stress [9], lung cancer [54], cystic fibrosis [55], liver disease [56], or coronary heart disease [57]. As far as the clinicalrelevance of such marker sets is concerned, statistical and analytical problems have to be meticulously analyzed.

Sampling problems

One of the main problems of breath sampling is reliability . Indeed, large variation of concentrations between different studies, probably due to dilution and contamination of samples with dead space gas and loss of analytes during the sampling procedure, suggests that the collection of gaseous samples represents a key-point.

In principle, there are two basic approaches to breath collection:

1 Mixed expiratory sampling (*total breath including dead space air*)
2 Alveolar sampling (*pure alveolar gas*)

Each of these two methods present problematic and positive aspects:

Method 1)

Advantages: easy to perform in spontaneously breathing subjects and requires non additional equipments.

Disadvantages: effect of diluition and contamination by dead space.

Method 2)

Advantages: concentrations of endogenous volatile substances in alveolar air are two three times higher than those found in mixed expiratory samples. alveolar breath samples have the lowest concentration of contaminants.

Disadvantages: Difficulties in the individuation of the correct volume.

It is evident that the main problem in breath sampling is to distinguish between endogenous substances and exogenous contaminants. There are three approaches normally used to overcome this problem:

1) Breath concentrations correction:

Substance concentrations in ambient or inspiratory air are measured, and exhaled concentrations are corrected by subtracting inspiratory from expiratory concentrations or by calculating 'alveolar gradients'[58],($AUC_{VOC\text{in breath}}/AUC_{\text{internal standard}} - AUCVOC_{\text{in air}}/AUC_{\text{internal standard}}$, where AUC is Area Under Curve, referred to GC analysis) by the means of which an endogenous VOC has a positive alveolar gradient, while an exogenous one has a negative alveolar gradient.

But these subtraction methods, although easy to perform, do not take into account the complexity of pulmonary adsorption and exhalation of volatile substances. Indeed, expired samples may be diluted or contaminated by inspiratory and/or dead space gas depending on the ratio of alveolar and dead space ventilation, which itself depends on the breathing pattern [59]. In addition, excretion and intake of volatile substances depend on the ventilation/perfusion ratio in the lung and on the alveolar concentration gradients of the substances [60]. This is enough to understand the simple subtraction of concentrations is not a complete effective method to solve such a complex problem, at least not investigating about identification and separation of very low concentrations of particular VOCs.

2) Pure air breathing:

This methods consists in the attempt to eliminate ambient concentrations by having patients or volunteers to breathe pure air for a certain time before measurements [45].

But this approach is cumbersome and time consuming, and will not be applicable for clinical routine purposes.

3) Dead space elimination:

This is of course the most simple method in principle, but it presents two problematic aspects referred to the experiment target and to the dead space definition. It consists in the elaboration of a particular protocol allowing the collecting of the first part of the breath in a bag apart, continuing to collect the second (alveolar breath) in a bag for the measurement, and to complete these two operation inside the same breath collecting action.

But the first problem is related to the goal of the investigation. Indeed, if the experiment aims to separate and identify the single different VOCs in the alveolar breath, the problem is quite the same of the classical dilemma about opportunity to sample the alveolar phase of the expiratory action. Anyway, the central objective of this chapter is the use of an electronic nose for lung cancer detection, so, the aim is to measure not the single compounds, but in the whole combination of these compounds, that could probably represent the 'global marker' of a lung cancer disease. From this point of view the alveolar breath collected with this method is a sort of headspace of the blood-lung exchange taking place at the alveolar level, the preferable sample for e-enose measurements. The whole technique will be described better in the following of the chapter, when the evolution of the sampling protocol will be illustrated during the six years of experimentation.

The other problem related to this method is the determination of the dead space volume to eliminate. There are many different theories about dead space volume. Some fixing a particular number included in a range not over 0.5 liters, other choices, more reasonable, indicating a this volume as the result of a calculus proportional to the lung capacity of the individual. This second method seems to be the most precise, but it presents problems of feasibility very similar to the second techniques reported above. Again, the application of this experience is postponed to the following of the chapter.

The three sampling techniques now described, can be synthesized in figure 3.2.3.

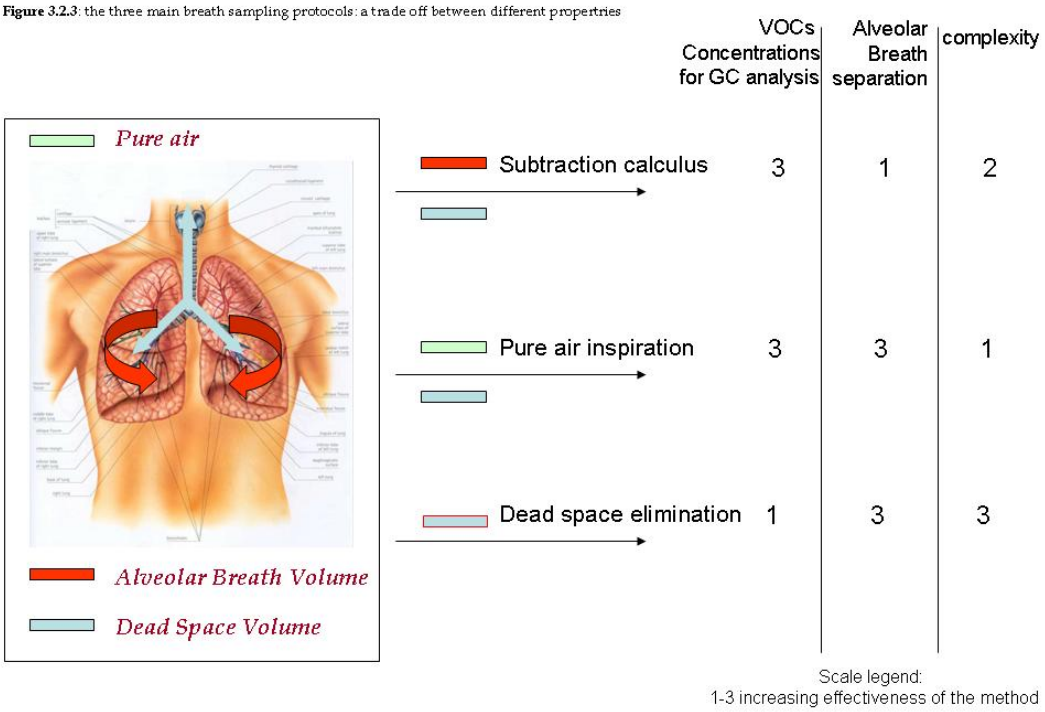


Figure 3.2.3: the three main breath sampling protocols: a trade off between different propertries

Anyway, in most of the classic techniques, preconcentration is a must, most substance concentrations in exhaled air falling in the nmol/l – pmol/l (ppbv-pptv) range. Three different methods have been applied until now:

1) Sorbent traps:

Due to the different boiling points of the VOCs, adsorbents in sorbent traps have to be selected carefully to avoid breakthrough as well as memory effects. Organic polymers (e.g., Tenax TA) [61], activated charcoal [62], different types of graphitized carbon (e.g., Carbopack X), and carbon molecular sieves (e.g., Carboxen 1021) have been used.

2) coated fibers (solid phase microextraction SPME):

Breath volatiles can also be preconcentrated by means of SPME [63]. Due to the physical properties of the available fibres, the number of substances that can be adsorbed is limited. Desorption of volatile compounds from these coated fibres is done by direct heating in the gas chromatography (GC) inlay. This adsorption technique is not affected by high water contents of the samples, and SPME may be automated for adsorption as well as for desorption.

3) Direct cryofocussation

Volatile substances can be preconcentrated directly into the GC inlay. This has been used as an independent sampling method as well as an enhancement of thermodesorption.

As anticipated in the lines above, the two analytical techniques used for breath analysis are: Gaschromatograph and the so called electronic nose. To conclude this chapter, it is necessary to have a very short look to the technological state of the art in this field.

The complexity of the sample, breath, still requires sophisticated equipments and excellent skill for sampling, preconcentration and analysis. Anyway there is a trend, well supported by technological progress, that foreshadows a simplification analytical techniques which could be employed for routine use in the near future:

Sampling can be standardized by automatic devices being able to separate alveolar air from dead space and inspiratory gas [64]. Preconcentration of some volatile substances has already been simplified considerably by the introduction of solid phase microextraction (SPME) [63] facilitating automatic preconcentration as well as automatic desorption.

A very promising new technique is the membrane extraction with sorbent interface (MESI), recently developed by Pawliszyn et al. [64]. This method integrates sampling and preconcentration in one step. It is based on a selective membrane acting as the interface between the respiratory circuit and the analytical system (GC).

Promising developments of real time determination of volatile substances are going on. Ethane can already be detected in the pmol/l (pptv) range via laser spectroscopy [65], other compounds were detected and identified without relevant delay by means of selected ion flow tube mass spectrometry (SIFT) [66].

Nonspecific sensor arrays technology (electronic noses) [37] may gain importance for diagnostic purposes when specific patterns of disease markers have been defined and disturbing factors can be excluded. This is the case of the lung cancer detection experience illustrated since paragraph 3.4.

3.3 Lung cancer diagnosis

3.3.1How is lung cancer diagnosed

When lung cancer develops, it tends to spread from the original cancer site itself to the lymph glands, and then, either at the same time or sequentially, to

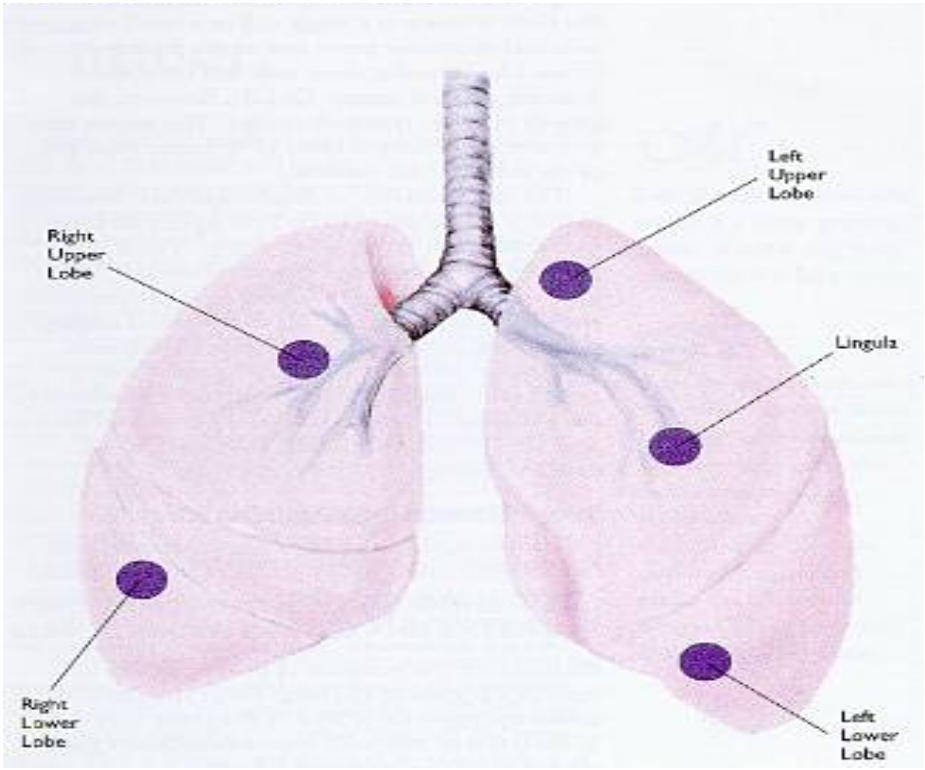


Figure 3.2.1: common sites within the lung whetre canvcer occurs

other areas of the body. Figure 3.3.1 illustrates the common sites within the lung where cancer occurs. The most common sites for lung cancer spread (metastasis) are the brain, bones, liver, adrenal glands (where *adrenalin* is produced), and any other organ with a high rate of blood flow. It is this process of metastasis that leads to fatality in most patients. Therefore, when we think about lung cancer, we need to think about the symptoms caused by the primary cancer, the symptoms caused by regional metastasis within the chest, and the symptoms that may be caused by the cancer's spread to distant areas of the body. When a physician first

discovers a cancer on physical examination or on any of our diagnostic tests, it is at least 1 cm (1/2 inch) in size. Below this size, we are not quite sure, either on physical exam or on an x-ray, whether it is an abnormal finding or not. The physical examination and diagnostic tests that follow a suspected lung cancer are all directed toward ascertaining whether the cancer is still localized, or whether it has become regional or distant. The physical examination and diagnostic tests that follow a suspected lung cancer are all directed toward ascertaining whether the cancer is still localized, or whether it has become regional or distant.Using a stethoscope, the physician will listen to the chest for any evidence of pneumonia, or collapse of a portion of the lung. He can also detect the presence of fluid in the chest cavity (pleural fluid), and, with experience, detect fluid around the heart (pericardial fluid).The physician will also look for signs of increased pressure on the veins in the chest, which will present as readily apparent distended neck veins. He will also search for spread to the lymph glands in the area just above the clavicle (collarbone). These glands are known as the supraclavicular lymph nodes. A presence of cancer in this area represents a very extensive regional spread of cancer.

The initial history, physical examination, and chest x-ray all determine the types of diagnostic tests that will then be performed. If the physician feels that the cancer is still localized or regionalized to the chest, diagnostic tests will then be directed toward proving the patient's disease can be safely removed and cured by surgery. On the other hand, if the initial examination suggests widespread disease, diagnostic tests are then performed to rapidly verify the condition. In this case, nonsurgical therapies would be considered.

Tests for Presumed Local or Regional Disease

If the patient's history, **physical examination**, and **chest x-ray** suggest local or regional disease, a series of tests will be initiated to verify that the disease is surgically resectable and curable. The first of these tests is called a chest ***CAT scan***, also known as a *CT scan* or *computed tomography*. A CT scan allows the radiologist to look inside the chest and determine whether the primary cancer is pressing

against any other vital structure(s). It also identifies whether the lymph glands in the middle of the chest have been invaded. Additionally, because most *CT scans* include the upper abdomen, the images also suggest whether or not the cancer has affected the liver and adrenal glands. The next most commonly performed procedure is a ***bronchoscopy***. Under local anesthesia, a flexible tube is placed down a patient's airway to search for evidence of cancer. Oftentimes, a very fine needle will be introduced into the chest to take a ***biopsy*** (*fine-needle aspiration*).

Tests for Presumed Distant (Widespread) Disease

If the initial studies reveal that the patient may have widespread disease, it makes no sense to perform complex tests to determine whether a patient can be safely resected. Instead, the most rapid and definitive procedure should be performed to confirm distant disease. As part of the initial work-up, a set of blood tests will be performed. One of them, the *alkaline phosphatase*, is often a harbinger of spread to the liver or bone. The liver is usually well visualized on the CT scan of the chest. If the liver is normal on CT scan, but the alkaline phosphatase level is elevated, a bone scan should be obtained, even if there are no symptoms of spread to the bone. Bone pain, too, should be investigated with a *bone scan*. A bone scan involves the injection of a mildly radioactive phosphorus into newly forming bone. Wherever a cancer destroys a bone, the body tries to repair the bone directly next to it and it is this repair process that shows up on the bone scan. A symptom of back pain could also indicate the presence of a condition known as bony metastasis. Here, the possible compression of the spinal cord by a tumor is not well diagnosed by the bone scan or routine x-rays. An MRI of the spine is the definitive way to be sure that *bony metastasis* is not present. In the presence of headaches, double vision, and/or confusion, the preferred test is a *magnetic resonance imaging* (MRI) of the brain. Often, a CT scan of the head with contrast is substituted; but it is not quite as effective.

To conclude it is useful to present the general classification accepted nowadays, which follows the indications of the OMS:

Two main different kinds:

- Small-Cell Lung Cancer (15%)
- Non-Small-Cell Lung Cancer (of which:)
    - Epidermoid Carcinoma or Squamous (40%)
    - Adenocarcinoma (30%)
    - Anaplastic Carcinoma (15%)

This dissertation is enough to understand the reason of an investigation by the means of breath odour to diagnose lung cancer. The reason consists in a list of very interesting advantages: non-invasive exams, possibility of early diagnosis, cheaper medical tests.

3.3.2 A survey on reported experiments

Two main studies about lung cancer diagnosis by the means of breath analysis are reported in literature: the first, in chronological order, is the experiment by O'Neill et al. [67], the second is the cross-sectional study by Phillips et al. [9,54]. Both these two investigations aimed to find a set of VOCs present in unusual concentrations in the alveolar breath of patients affected by lung cancer, and to identify, in this set, a list of possible markers of the disease.

In the first study O'Neill and his co-workers individuated a list of 28 VOCs, which seem to have an anomalous concentrations related with the presence of lung cancer; the main of this group are reported in table 3.3.1.

ALKANES	Hexane Methylpentane
BENZENE derivates	O-toluidine Aniline

Table 3.3.1: main lung cancer markers individuated by O'Neill.

In the successive study Phillips collected breath samples from 108 patients with an abnormal chest radiograph who were scheduled for bronchoscopy. Lung cancer was confirmed histologically in 60 of these patients. A list of 67 possible VOCS was identified, in this ghroup a list of 22 was chosen for the discriminant analysis of data. The general list of compounds identified as lung cancer possible markers is not so different from the compounds found in the breath of a healthy person; the difference consists in the anomalous values of their concentration.

The 22 VOCs individuated by Phillips are quite identical to the list given by O'Neill; in particular in both the two lists there are Isoprene, benzene and four derivates of Benzene. The origin of the presence of benzene in human breath is unknown.

Unfortunately it does not seem to emerge any relation between the anomalous values of concentration of these VOCs and the stage of advance of the tumor. In table 3.3.2 the list of these 22 VOCs is reported.

Styrene (ethenylbenzene)
Heptane,2,2,4,6,6-pentamethyl
Heptane,2-methyl
Decane
Benzene,propyl
Undecane
Cyclopropane,1-methyl-2-pentyl
Methane,trichlorofluoro-
Benzene
Benzene,1,2,4-trimethyl-
1,3-butadiene, 2-methyl-(isoprene)
Octane, 3-methyl
1-hexene
Nonane,3-methyl-
1-heptene
Benzene,1,4-dimthyl
Heptane,2,4-dimethyl
Hexanal
Cyclohexane
Benzene,1-methylethenyl
Heptanal

Table 3.3.2: list of the 22 VOCs identified as possible lung cancer markers by Phillips et al.

The sampling procedure was performed by the use of an ad hoc Breath Collecting Apparatus (BCA), which allowed to collect 10 litres of breath during a 5 minutes continuous breathing of the patient. The method used for the identification of the endogenous VOCs was the calculus of the alveolar gradient; by this calculus a mean number of 204 VOCs were traced in the sampled population.

Cross-validation criterion gave a percentage of correct classification of 71.1% for the hill patient and of 66.7% for the healthy individuals.

The great potentiality of these studies is in the finding of a peculiarity of the breath as consequence of the tumor presence, the problem is the correct identification of this peculiarity. Indeed, the identification of these possible markers is a sort of 'clue' that breath can tell something about the health status of the lung, but it is clear that all these compounds are not markers by themselves. This consideration of course suggest two important concepts: the list is probably not complete because of the difficult in the quantification of very low concentrations; moreover the fact that the contemporary presence of these compounds represents an indication of tumor cells, suggest to look for a synthetic source of information capable to reveal the global action of all these compounds as symptom of disease. This synthetic information could be represented by breath odour, and the e-nose is the device used to trace this odour. These instruments are basically arrays on non-selective sensors. In fact, the sensors response is not univocally correlated with the concentration of a single compound (as in the classical analytical chemistry) but rather it is a sort of combination of all the chemical information contained in each sample. This is obtained through a broad selectivity, like a sort of chemical filtration.

3.4 References

[1] A. P. F. Turner, N. Magan, 'Electronic noses and disease diagnostic', *Nature review*, Microbiology, vol.2, Feb. 2004, pp. 161-166.

[2] C. Di Natale, A. Macagnano, R. Paolesse, E. Tarizzo, A. Mantini, A. D'Amico, 'Human skin odour Analysis by means of an electronic nose', *Sensors and Actuators B* 65 (2000) 216-219.

[3] Persaud, K. & Dodd, G., 'Analysis of discrimination mechanisms in the mammalian olfactory system using a model nose'. *Nature* 299, 352–355 (1982).

[4] Aathithan, S., Plant, J. C., Chaudry, A. N. & French, G. L.,'Diagnosis of bacteriuria by detection of volatile organic compounds in urine using an automated headspace analyser with multiple conducting polymer sensors'. *J. Clin. Microbiol*. 39, 2590–2593 (2001).

[5] Grametbauer, P., Kartusek, S. & Hausuer, O.' Diagnosis of aerobic Gram negative bacteria by the detection of volatile metabolites using gas chromatography'. *Cesk Epidemiology Mikrobiology Immunology* 37, 216–223 (1988).

[6] Vitenberg, A. C., Stolbova, A. V., Loffe, B. V., Kocherovets, V. I. & Tsibul'skaia, I. A. 'Headspace gas chromatography analysis in the rapid diagnosis of anaerobic infections'. *Zh Mikribiology Epidemiology Immunobiology* **1**, 20–24 (1986).

[7] Socolowsky, S, Hohne, C. & Sandow, D. 'The direct detection of volatile fatty acids by gas chromatography in microbiological diagnosis'. *Zeitschrift Med. Lab. Diagn*. 31, 445–452 (1990).

[8] Phillips, M. *et al.'* Volatile markers of breast cancer in the breath'. *Breast J.* 9, 184–191 (2003).

[9] Phillips, M. *et al.* 'Detection of lung cancer with volatile markers in the breath'. *Chest* 123, 1788–1792 (2003).

[10] Olopade, C. O., Zakkar, M, Swedler, W. I. & Rubinstein, I.' Exhaled pentane levels in acute asthma'. *Chest* 111, 862–865 (1997).

[11] Guernion, N., Ratcliffe, N. M., Spencer-Phillips, P. T. & Howe, R. A. 'Identifying bacteria in human urine: current practice and the potential for rapid, near-patient diagnosis by sensing volatile organic compounds'. *Clin.Chem. Lab. Med.* **39**, 893–906 (2001).

[12] Kaji, H., Hisamura, M., Saito, N. & Murao. M. 'Gas chromatographic determination of volatile sulphur compounds in expired alveolar air in hepatopathic patients'. *J. Chromatogr.* 145, 464–468 (1978).

[13] Humad, S., Zarling E., Clapper, M. & Skosey, J. L. 'Breath pentane excretion as a marker of disease activity in rheumatoid arthritis'. *Free Radicle Res***. 5**, 101–106 (1988).

[14] Phillips, M., Sabas, M. & Greenberg, J. 'Increased pentane and carbon disulphide in the breath of patients with schizophrenia'. *J. Clin. Pathol*. 46, 861–864 (1993).

[15] Dobbelaar, P. *et al.* 'Detection of ketosis in dairy cows by analysis of exhaled breath'. *Veterinary Quality* 18,151–152 (1996).

[16] Skrupskii, V. A. 'Gas chromatographic analysis of ethanol and acetone in the air exhaled by patients'. *Clin. Lab.Diagn***. 4**, 35–38 (1995).

[17] Goldberg, E. M., Blendis, L. M. & Sandler, S. 'A gas chromatographic–mass spectrometric study of profiles of volatile metabolites in hepatic encephalopathy'. *J. Chromatogr*. **226**, 291–299 (1981).

[18] Pavlou, A. *et al.* 'An *in vitro* rapid odour detection and recognition model in discrimination of *H. pylori* and other gastroeosophageal pathogens'. *Biosensors and Bioelectronics* 15, 333–342 (2000).

[19] Pavlou, A. *et al.* 'Use of an electronic nose system for diagnoses of urinary tract infections *in vivo*' . *Biosensors and Bioelectronics* 17, 893–899 (2002).

[20] Dutta, R., Hines, E. L., Gardner, J. W. & Boilot, P. 'Bacteria classification using Cyranose 320 electronic nose'. *BioMedical Engineering Online* 1, 1–7 (2002)

[21] Persaud, K. C., Pisanelli, A. M. & Evans, P. in *Handbook of Machine Olfaction: Electronic Nose Technology* (eds Pearce, T. C., Schiffman, S. S., Nagle, H. T. & Gardner, J. W.) 445–460 (Wiley, 2002).

[22] Gibson, T. D., Prosser, O., Hulbert, J., Marshall, R. W. & Li, Z.'Detection and simultaneous identification of microorganisms from headspace samples using and electronic nose'. *Sensors and Actuators B* 44, 413–422 (1997).

[23] Gardner, J. W., Craven, M., Dow, C. & Hines, E. L. 'The prediction of bacteria type and culture growth phase by an electronic nose with a multi-layer perceptron network'. *Measurement Sci. Technol.* 9, 120–127 (1998).

[24] Pavlou, A., Turner, A. P. F. & Magan, N. 'Recognition of anaerobic bacterial isolates *in vitro* using electronic nose technology'. *Lett. Appl. Microbiol.* 35, 366–369 (2002).

[25] Magan, N., Pavlou, A. & Chrysanthakis, I. 'Milke sense: a volatile sensory system for detection of microbial spoilage by bacteria and yeasts in milk'. *Sensors and Actuators B*, 72, 28–34 (2001).

[26] Needham, R. & Magan, N. 'Detection and differentiation of toxigenic and non-toxigenic *Penicillium verrucosum* strains on bakery products using an electronic nose'. *Aspects Appl. Biol.* 68, 217–222 (2003)

[27] Lykos, P., Patel, P. H., Morong, C. & Joseph. A. 'Rapid detection of bacteria from blood culture by an electronic nose'. *J. Microbiol*. 39, 213–218 (2001).

[28] Keshri, G., Vosey, P. & Magan, N. 'Early detection of spoilage moulds in bread using volatile production patterns and quantitative enzyme assays'. *J. Appl. Microbiol.* 92, 165–172 (2002).

[29] Keshri, G. & Magan, N. 'Detection and differentiation between mycotoxigenic and non-mycotoxigenic strains of *Fusarium* spp. using volatile production profiles and hydrolytic enzymes'. *J. Appl. Microbiol.* 89, 825–833 (2000).

[30] Needham, R. & Magan, N. 'Detection and differentiation of microbial spoilage organisms of bakery products *in vitro* and *in situ'*. *Proceedings of the Ninth International Symposium on Olfaction and Electronic Nose* (eds D'Amico, A. & Di Natale, C.) 385–388 (Rome, Italy, 2003).

[31] Keshri, G., Challen, M. P., Elliot, T. J. & Magan, N. 'Differentiation of *Agaricus* species and other homodasidiomycetes based on volatile

production patterns using an electronic nose system'. *Mycol. Res.* 107, 609–613 (2003).

[32] Parry, A. D. & Oppenhaim, B. 'Leg ulcer odour detection identifies β-haemolytic streptococcal infection'. *J. Wound Care* 4, 404–406 (1995).

[33] Chandiok, S. *et al.* 'Screening for bacterial vaginosis: a novel application of artificial nose technology'. *J. Clin.Pathol*. 50, 790–795 (1997).

[34] Pavlou, A. 'Novel intelligent gas-sensing in diagnosis of infectious diseases'. PhD Thesis, Cranfield Univ. (2003).

[35] Di Natale, C. *et al.* 'Electronic nose analysis of urine samples containing blood'. *Physiol. Measurement* 20, 377–384 (1999).

[36] Hanson, C. W. & Steinberger, H. A. 'The use of a novel 'electronic nose' to diagnose the presence of intrapulmonary infection'. *Anesthesiology* 87, A269 (1997).

[37] C. Di Natale, A. Macagnano, E. Martinelli, R. Paolesse, G. D'Arcangelo, C. Roscioni, A. Finazzi Agrò, A. D'Amico 'Lung cancer identification by analysisof breath by means of an array of non-selective gas sensors'. *Biosensors and Bioelectronics* 18, 1209–1218 (2003).

[38] Ping, W., Yi, T., Haibao, X. & Farong, S. 'A novel method for diabetes diagnosis based on electronic nose'. *Biosensors and Bioelectronics* 12, 1031–1036 (1997).

[39] W. Miekisch, J. K. Schubert, G. F. E. Noeldge-Schomburg, 'Diagnostic potential of breath analysis-focus on volatile organic compounds', *Clinica Chimica Acta* 347 (2004) 25-39.

[40] Pauling L, Robinson AB, Teranishi R, Cary P. 'Quantitative analysis of urine vapour and breath by gas – liquid partition chromatography'. Proc Natl Acad Sci U S A 1971;68:2374– 6.

[41] Manolis A. 'The diagnostic potential of breath analysis'. *Clin Chem* 1983;29:5–15.

[42] M. Phillips et al., 'Variation in volatile organic compounds in the breath of normal humans', *Journal of Chromatography B* 729 (1999) 75-88.

[43] Aghdassi E, Allard JP. 'Breath alkanes as a marker of oxidative stress in different clinical conditions'. *Free Radic Biol Med* 2000;28:880– 6.

[44] Kneepkens CM, Lepage G, Roy CC. 'The potential of the hydrocarbon breath test as a measure of lipid peroxidation'. *Free Radic Biol Med* 1994;17:127– 60.

[45] Risby TH, Sehnert SS. 'Clinical application of breath biomarkers of oxidative stress status'. *Free Radic Biol Med* 1999;27:1182–92.

[46] Stone BG, Besse TJ, Duane WC, Evans CD, DeMaster EG. 'Effect of regulating cholesterol biosynthesis on breath isoprene excretion in men'. *Lipids* 1993;28:705–8.

[47] Kuzma J, Nemecek-Marshal M, Pollock WH, Fall R. 'Bacteria produce the volatile hydrocarbon isoprene'. *Curr Microbiol* 1995;30:97– 103.

[48] Foster MW, Jiang L, Stetkiewicz PT, Risby TH. 'Breath isoprene: temporal changes in respiratory output after exposure to ozone'. *J Appl Physiol* 1996;80:706– 10.

[49] Mendis S, Sobotka PA, Euler DE. 'Expired hydrocarbons in patients with acute myocardial infarction'. *Free Radic Res* 1995;23:117– 22.

[50] Lebovitz HE. 'Diabetic ketoacidosis'. *Lancet* 1995;345: 767– 72.

[51] Cope K, Risby T, Diehl AM. 'Increased gastrointestinal ethanol production in obese mice: implications for fatty liver disease pathogenesis'. *Gastroenterology* 2000;119:1340– 7.

[52] Simenhoff ML, Burke JF, Saukkonen JJ, Ordinario AT, Doty R. 'Biochemical profile or uremic breath'. *N Engl J Med* 1977;297:132–5.

[53] Davies S, Spanel P, Smith D. 'Quantitative analysis of ammonia on the breath of patients in end-stage renal failure'. *Kidney Int* 1997;52:223–8.

[54] Phillips M, Gleeson K, Hughes JMB, Greenberg J, Cataneo RN, Baker L, et al. 'Volatile organic compounds in breath as markers of lung cancer: a cross-sectional study'. *Lancet* 1999;353:1930– 3.

[55] McGrath LT, Patrick R, Mallon P, Dowey L, Silke B, Norwood W, et al. 'Breath isoprene during acute respiratory exacerbation in cystic fibrosis'. *Eur Respir J* 2000;16:1065– 9.

[56] Sehnert SS, Jiang L, Burdick JF, Risby TH. 'Breath biomarkers for detection of human liver diseases: preliminary study'. *Biomarkers* 2002;7:174 – 87.

[57] Phillips M, Cataneo RN, Greenberg J, Grodman R, Salazar M. 'Breath markers of oxidative stress in patients with unstable angina'. Heart Dis 2003;5:95–9.

[58] Phillips M. 'Method for he collection and assay of volatile organic compounds in breath'. *Anal Biochem* 1997;247: 272–8.

[59] Cope KA, Watson MT, Foster M, Sehnert SS, Risby TH. 'Effects of ventilation on the collection of exhaled breath in Humans'. *J Appl Physiol* 2004;96:1371–9 [Electronic publication 2003 Dec 12].

[60] Wagner PD, Saltzman HA, West JB. 'Measurement of continuous distributions of ventilation– perfusion ratios: theory'. *J Appl Physiol* 1974;36:588– 99.

[61] Tangerman A, Meuwese-Arends MT, van Tongeren JH. 'New methods for the release of volatile sulfur compounds from human serum: its determination by Tenax trapping and gas chromatography and its application in liver diseases'. *Lab Clin Med* 1985;106(2):175–82.

[62] Mueller W, Schubert J, Benzing A, Geiger K. 'Method for analysis of exhaled air by microwave energy desorption coupled with gas chromatography– flame ionization detection–mass spectrometry'. *J Chromatogr B Biomed Sci Appl* 1998;716:27– 38.

[63] Grote C, Pawliszyn J. 'Solid-phase microextraction for the analysis of human breath'. *Anal Chem* 1997;69:587– 96.

[64] Schubert JK, Spittler K-H, Braun G, Geiger K, Guttmann J. 'CO2-controlled sampling of alveolar gas in mechanically ventilated patients'. *J Appl Physiol* 2001;90:486– 92.

[64] Lord H, Yu Y, Segal A, Pawliszyn J. 'Breath analysis and monitoring by membrane extraction with sorbent interface'. *Anal Chem* 2002;74:5650–7.

[65] Dahnke H, Kleine D, Hering P, Mu¨rtz M. 'Real-time monitoring of ethane in human breath using mid-infrared cavity leak-out spectroscopy'. *Appl Phys B* 2001;72:971– 5

[66] Spanel P, Smith SD. 'Selected ion flow tube: a technique for quantitative trace gas analysis of air and breath'. *Med Biol Eng Comput* 1996;34:409– 19.

[67] O'Neil, H.J., Gordon, S.M., O'Neil, M.H., Gibbons, R.D., Szidon, J.P., 1988. 'A computerized classification technique for screening for the presence of breath biomarkers in lung cancer'. *Clinical Chemistry* 34, 1613_ 1618.

LUNG CANCER DIAGNOSIS

'Se vuoi sapere quanto buio hai intorno, devi aguzzare lo sguardo sulle fioche luci lontane'.
Da 'Le città invisibili' di Italo Calvino

'When you want to know how deep is the darkness surrounding you, you must sharpen your look at the faint far lights'

4

4.1 Detection of Lung cancer by the means of electronic nose.

On the a basis of the literature reported in the previous paragraph, a first set of experiments with the electronic nose of 'Tor Vergata' on lung cancer detection started in 1998, at the Forlanini Hospital in Rome. Since that a total number of four experiments have been performed. The objective of this paragraph is to give a global overview of the results obtained along these years, especially pointing out the problems emerged in each experience and the consequent improvements adopted in the successive. Particular attention will be devoted to the last experiences (2004), as result of the whole knowledge acquired along the six years and as a suggestion for the next step.

In figure 4.1.1 a general overview of these whole experiment, divided in four fundamental steps, is illustrated. On the basis of this scheme, in the next pages, all these experiments will be discussed in its fundamental parts: objective, protocol, results and partial conclusions.

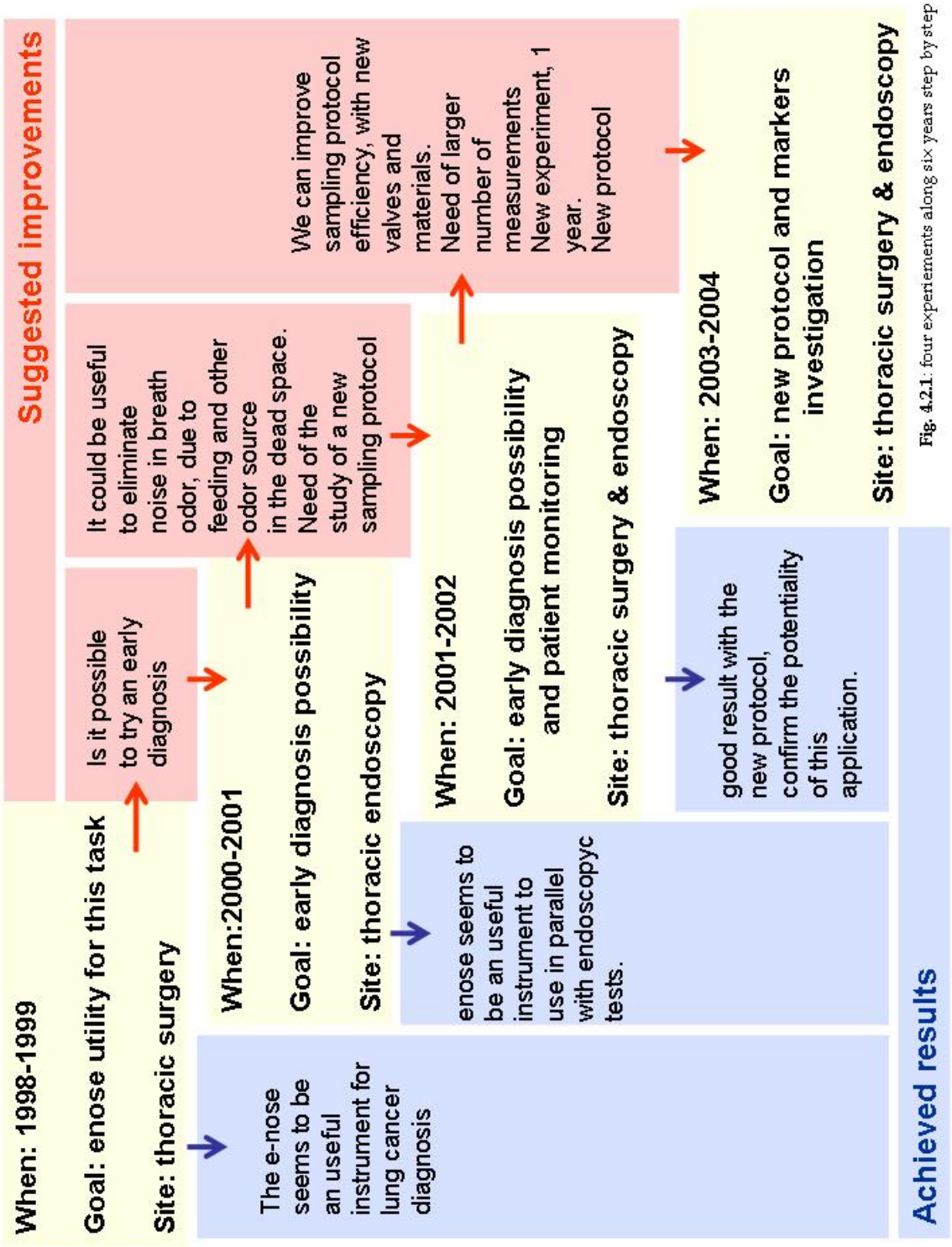


Fig. 4.2.1. four experiements along six years step by step

4.2 Previous experiments

First experiment (1998-1999)

This experiment involved 42 volunteers, all affected by various forms of lung cancer, have been recruited at the C. Forlanini Hospital in Rome [1]. Thirty-five of them were hospitalised waiting for a surgical treatment. Nine patients have been checked after a surgical removal of the tumour mass from the lung. Two patients were measured before and after the surgical operation. Eighteen volunteers have been recruited among the medical and nurse staff of the hospital, as reference. These controls were not affected by any apparent disease, and were not taking any drug. The cancer patients did not show of any different pathology.

Each subject was required to follow the same diet and the same procedure for mouth hygiene. Measurements have been performed in the morning before any food intake. Individuals were required to deeply breathe in a sterile and disposable bag (volume of about 4 l). The bag was endowed with tight valves in order to prevent any diffusion of external air inside the bag. The sampled breaths were then analysed with the electronic nose with a delay time of maximum 5 min from the sampling time.

All measurements were performed on-site. Each subject was measured twice, and the average measurement was used in the data analysis. In order to minimise the influence of possible instrumental drift, the measurement sequence was randomised. The whole experiment lasted 5 weeks. Post-surgical patients were checked about 1 month after the operation, in the occasion of a periodical control at the hospital site.

The content of the bag was flown at the constant speed of 0.4 ml/s into the sensor chamber. Fig. 1 shows examples of the temporal behaviour of the sensor signals A stabilisation of the sensor response, corresponding to the achievement of a thermodynamic equilibrium between absorbed and desorbed molecules, is reached in about 80 s. In order to be sure that in any measurement the steady-state is reached, the duration of each single measurement has been fixed at 200 s, with a consumption of 1.3 l of sampled breath.

The measurements have been performed as a difference between the sample and synthetic air. The difference between the frequency variation measured in synthetic air flux and in sample flux has been considered as the sensors response and it has been utilised for successive evaluations.

The data have been linearly normalised in order to remove as much as possible any concentration effects in the sample.

In this study, we were concerned with testing the capability of the electronic nose to correctly classify the groups of subjects. For this reason a supervised technique has been chosen. The simplest technique that can be used is the discriminant analysis, where a linear model between sensor data and classes is built. The use of PLS helps avoiding the drawbacks due to sensors correlation. Furthermore, PLS provides a decomposition of the sensors data in latent variables that can be plotted to provide a visual representation of the classification properties (Massart et al., 1988).

As in any supervised classification techniques, the classes have to be chosen a-priori. The natural choice for the samples in our experiment was to choose three classes including the patients with lung cancer, the healthy subjects, and the patients after the surgery. With this classification scheme a PLS-DA model has been built.

As it can be seen (fig. 4.2.2), a clear separation between the data related to patients with lung cancer and the other samples is observed. On the otherhand, the samples related to post-surgery and healthy reference show some overlap A numerical evaluation of the classification properties can be obtained by considering the cross-validation of the PLS-DA method according to a 'leave-one-out' technique. The results are shown in the table inserted in fig. 3.4.2 in form of a confusion matrix. 100% of the lung cancer patients have been correctly classified. 94% of the controls has been correctly classified and 6% of them has been classified as belonging to the post-surgery group. Concerning the samples of the post-surgery group 44% have been classified as belonging to an autonomous class, while 56% has been classified as healthy controls. It is worth mentioning the data related to the two patients measured twice, before and after surgery. In fig. 4.2.2 it

can be seen the migration of the data points from the class of lung cancer diseasesfrom the healthy controls.

Confusion matrix of the PLS-DA methods, leave-one-out validation method has been utilised

True/estimated	Cancer diseased	Post-surgery patients	Control sample
Cancer diseased	35	0	0
Post-surgery patients	0	4	1
Control sample	0	5	17

True classes are read along the columns and estimated classes along the rows. A total accuracy of 90.3% is achieved, but it is worth to consider that 100% of patients with lung cancer are correctly identified, and no misclassification between this group and the others occurs.

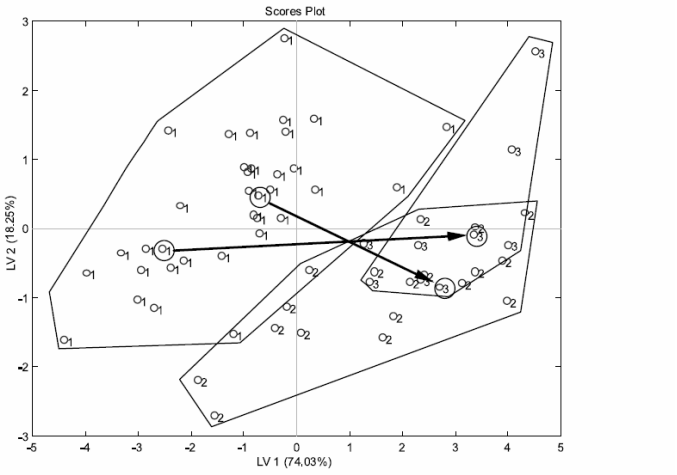


Fig. 4.2.2: Scoreplot of the first two latent variables of the PLS-DA Model built on te data collected; the confusion matrix is also reported.Labels: 1: Cancer disease, 2: Post-surgery patients, 3: Control sample

In this study all patients were roughly at the same stage of development of cancer. In perspective, the relation between cancer stadiation and sensor signal shall be studied in order to investigate the resolution of the method. In particular a possible application of electronic nose for early diagnosis of cancer will be much-needed goal. This has been the objective of the second experiment.

Second experiment (2000-2001)

Bronchoscopy ambulatory is the most obvious place where to perform a second measurements campaign aiming to investigate about a possible early diagnosis of lung cancer.

The same sampling protocol of the first experiment was used.

To test the capability of the instrument to give an early diagnosis, the need was to study the sensitivity to the different stadiations of the disease in order to evaluate the resolution. For this reason volunteers was recruited at their first visit, also planning he possibility to follow the patients during the disease evolution.

Anyway, this choice led to measure many spurious cases, namely other pulmonary and thoracic pathologies, present in a to little number to have a statistical relevance.

The total number of measures performed were 39. By a PLS-DA data analysis considering three classes (tumor, endoscopy negative reports, healthy) a percentage of correct classification of 81% has been calculated.

In figure 4.2.3 the scoreplot of the first two Principal Components of the PCA model elaborated is reported. The data set considered in this model does not include the measures relative to the other pulmonary pathologies.

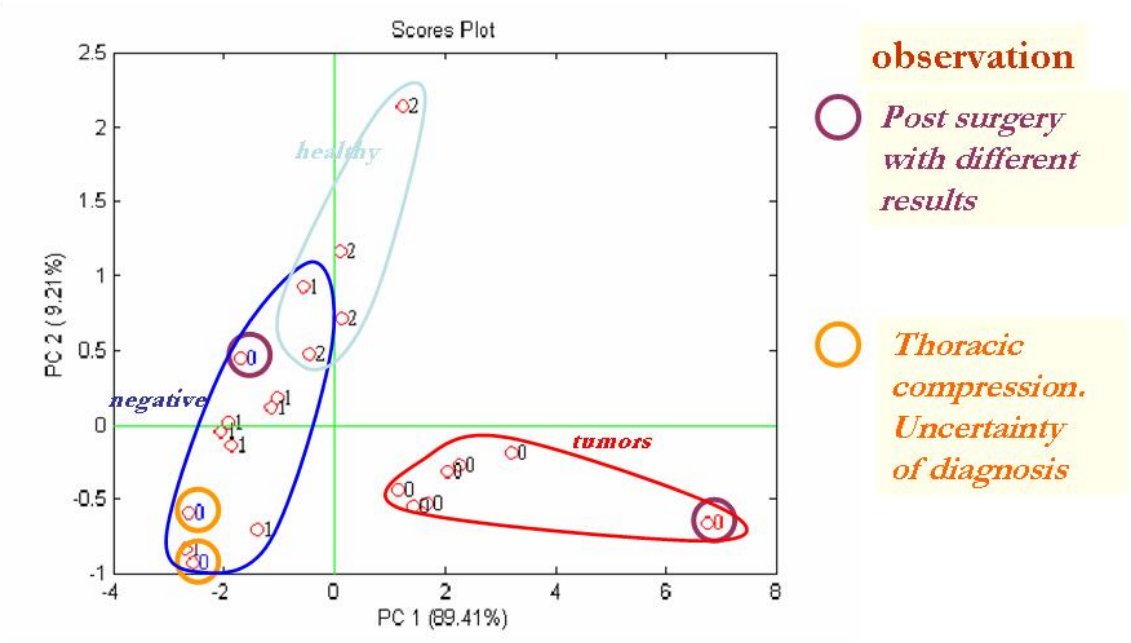


Figure 4.2.3: scoreplot of the first two PCs of the PCA model elaborated on tumors, negatives and healthy data set

Comparing e-nose data with the reports of thoracic endoscopy, an interesting observation has been pointed out from the different classification sets: two patients classified by the e-nose as negative, presented an uncertain diagnosis after the endoscopic exam. These two cases are pointed out in figure 4.2.3 by the means of an orange circle. This is why thoracic endoscopy is not only an invasive medical test, but its potentialities are also limited by some physical problems; indeed the technicians performing this test use to say that the endoscopy is able to

see everything in the branches of the bronchial tree, but not what is the 'leaves'. Another remark emerging from the scoreplot reported in figure 4.2.3 is the case of a different classification for two post-surgical samples. The e-nose has shown the capability to discriminate between the different results obtained in the two operations; either the individual yet affected by tumor and the patient in which the tumoral mass had been completely removed were correctly classified (the two samples are pointed out in figure 4.2.3 by the means of violet circles).

It emerged from this experiment that the e-nose performance are sufficient to use the instrument in parallel with the endoscopyc test. As a consequence a new problem was pointed out. Indeed, while in the first experience patients have been measured during their period in hospital, in this case, very often, patients were still leaving at home; this means lack of control on their habits and feeding, two important noise influencing the e-nose measurements effectiveness.

To eliminate these disturb sources, a new sampling protocol was elaborated in the third experimental phase.

Third experiment (2001-2002)

In this measurements campaign protocol, breath sampled volume, patient treatment and data analysis techniques was still the same used in the first and second experience, and well described in the paragraph referred to the first one. The improvement in the sampling apparatus is a new 3-way valve, able to separate the inlet air into two different variable volumes. This system, opportunely inserted at the inlet of the sampling apparatus, has allowed the separation between the headspace and the alveolar breath [2], and the successive elimination of the first one from the measured air sample. In fig. 4.2.4 is illustrated the simple working principle of this valve and the two different ways along where the two samples of breath have been conveyed.

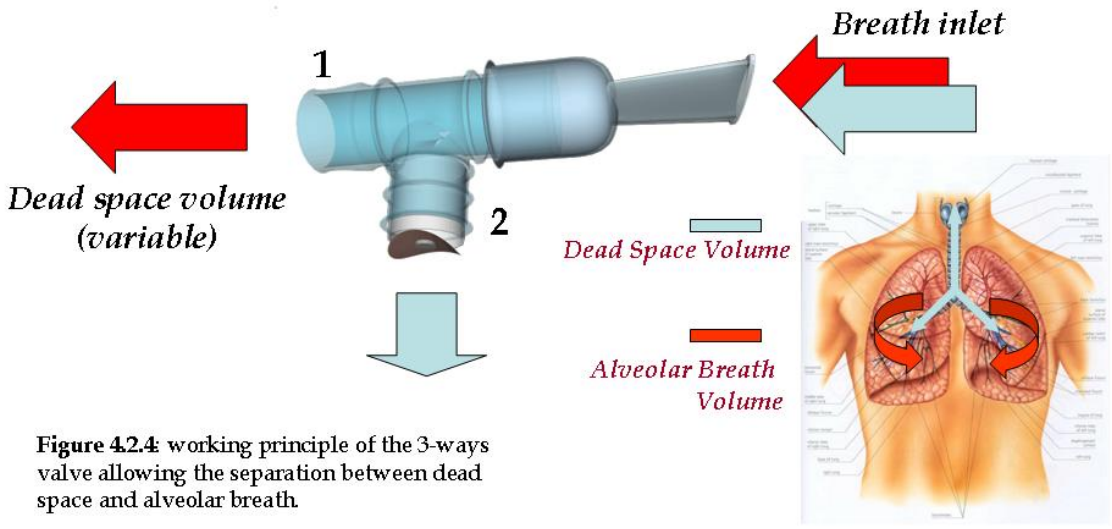


Figure 4.2.4: working principle of the 3-ways valve allowing the separation between dead space and alveolar breath.

Looking at the figure it is easy to understand the mechanism: the 'total breath' expired by the individual, entering the valve, finds only one way out (1). By this way the first part of the breath, the dead space, goes to fill a first bag of fixed volume. This bags completely filled results in a closed set, so the second part of the breath, the alveolar one, goes to open the second way filling another bag, then used for the measurement.

At the beginning of the experiment a variable volume has been chosen for the dead space beg. According to the theory of a dead space proportional to the individual lung capacity, a spirometric test was performed as first, before the breath sampling procedure for the e-nose. This technique, anyway, being very difficult to practice with patients affected by lung disease, decreased the importance of one of the main principle of this experimentation: non-inasivity. The natural consequence of this consideration has been the choice for a fixed volume of 0.5 litres for the dead space bag, of course a quantity superior to a normal dead space.

A second variation respect to the previous experiment, has consisted in the extension of the investigation to both the groups of the individuals, the

hospitalised patients of thoracic surgery and the patients addressed to thoracic endoscopy. The measurements campaign involved 63 volunteers, 15 from the ambulatory of bronchoscopy and 48 from thoracic surgery.

The two sets of measurements have been analyzed both separately and together with PLS-DA method.

The highest percentage of correct classification, 93%, has been obtained elaborating the endocopy data treated apart, the three classes were healthy, tumor diseases and negative. This good result is obviously due to a tested good ability of the e-nose in this matter, anyway the low number of measurements has contributed to the high value of correct prediction. The scoreplot of the first two Latent Variables of the PLS model elaborated is reported in figure 4.2.5.

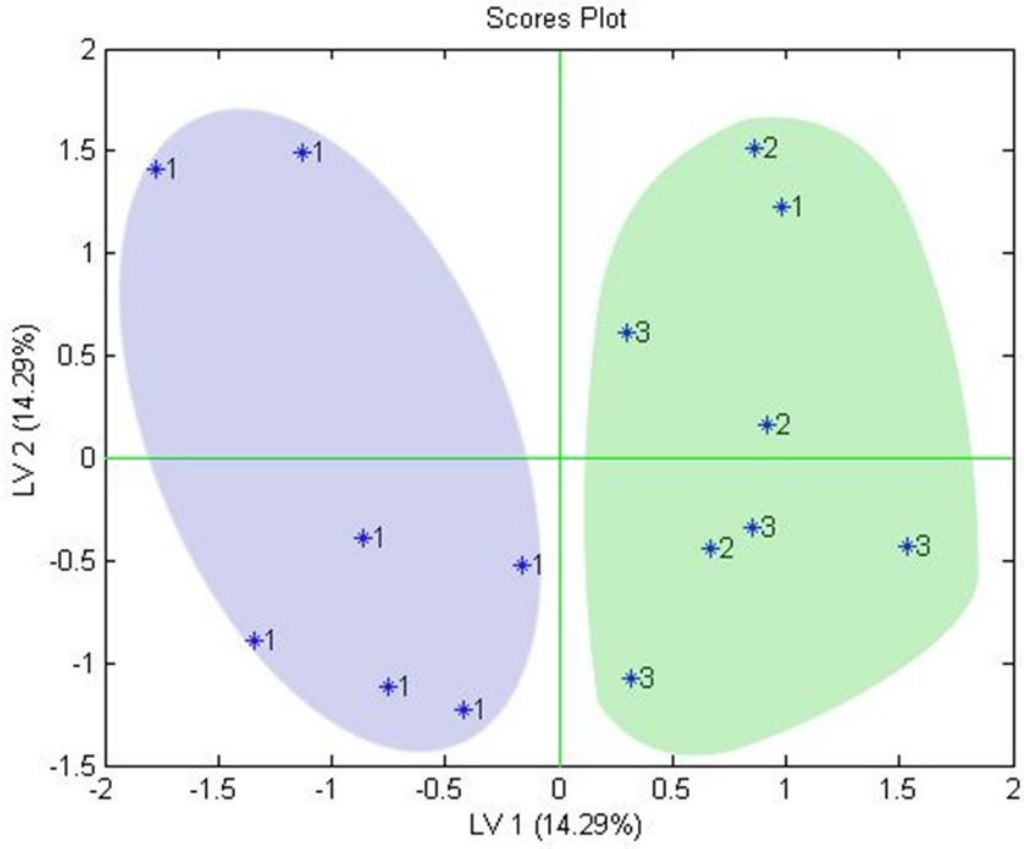


Fig. 4.2.5: scoreplot of the first two LVs of the PLS-DA model elaborated on the broncoscopy data set. (labels: 1.tumors, 2. healthy, 3. negatives)

The treatment of Thoracic surgery data has given a percentage of correct classification of 83% in the discrimination between the two classes of tumor diseases and healthy persons. The absence of negative individual is due to the fact

that persons hospitalised in a surgery department are waiting for the surgical operation, so the presence of the tumoral mass in these cases is already assessed.

The scoreplot of the first two Latent Variables of the PLS model elaborated is reported in figure 4.2.6.

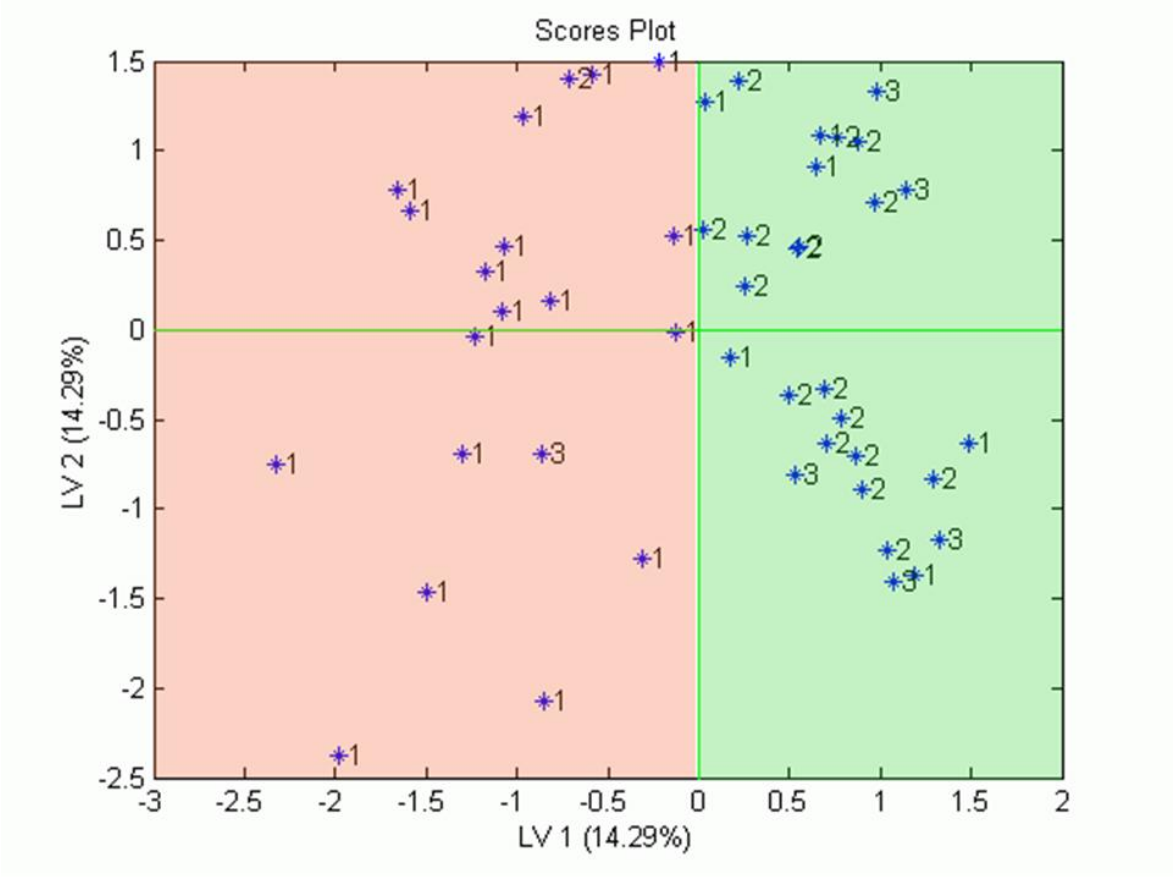


Fig. 4.2.6: scoreplot of the first two LVs of the PLS-DA model elaborated on the thoracic surgery data set. (labels: 1.tumors, 2. healthy, 3. negatives)

To conclude the whole set of 63 measures has been analyzed, using three classes: tumor diseases, negative reports, healthy persons. The correct classification percentage obtained, 75%, is a 'clue' that the potentiality of the instrument are yet good, even if hardly tested by a so confused situation. Actually, it has to be considered the presence of three main problematic new aspects. The two experimental data, mixed together, give a more general overview on lung cancer stadiation. The tested volunteers, in the case of endoscopy were not so controlled in their habit as the hospitalised patients. The individuals classified as negative reports by the endoscopyc exam did not really belong to a same class,

considering that this negativity very often meant another pathology regarding lung and bronchial tubes. Anyway, the number of measurements performed in the ambulatory of bronchoscopy, did not allowed a more detailed analysis of the different pathologies.

These considerations, as usual, becomes suggestions for the next, and last until now, experiment.

A more sophisticated sampling protocol is a must to avoid an excessive influence of the not well monitored habits of the patients, so it is important to design again the apparatus for the breath collecting, following the results emerged in the previous experiments, and trying to increase the resolution of the instruments by the elimination of certain noises added by some non-optimized elements in the protocol.

The use of the tested 3-ways valve for dead space elimination, is confirmed to reduce the effects of feeding in the breath analysis. Moreover, the use of simple connection non opposing resistance to breath flow, and the need to build every part of the sampling apparatus with inert materials, are expected to be considered as the only way to optimize a well trained but not perfect collecting apparatus.

Moreover It is necessary to perform a high number of measurements, in order to increase the statistic value of the experimentation and to allow a discrimination of the different pathologies inside the usual three classes.

4.3 Fourth experiment (2003-2004)

This last experimentation has taken place in the period from December 2003 to July 2004, in two departments of C. Forlanini Hospital in Rome: C.U.B.E. (Centro Universitario Broncopatie Emergenti), and in the ambulatory of Bronchoscopy.

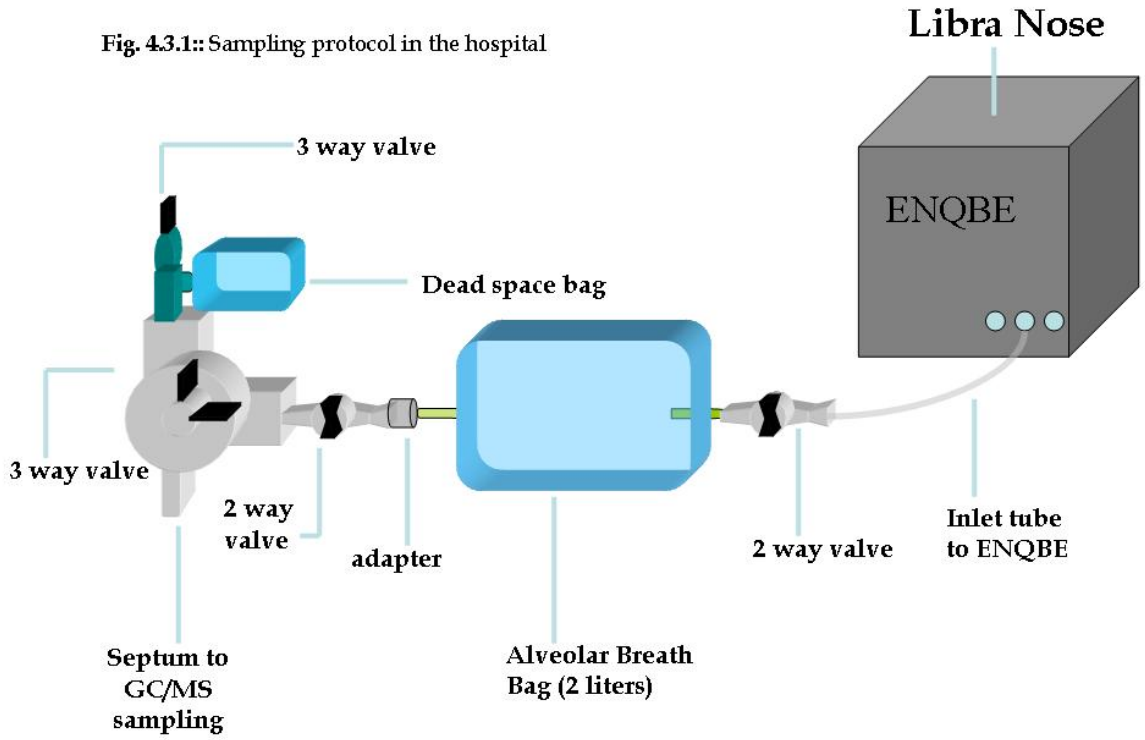
In parallel with the measurements in the hospital, another measurement campaign has been performed in the laboratories of CNR-IMM ('Tor Vergata' Rome), to study the sampling protocol and to investigate about the influence of some of the possible lung cancer markers reported in literature.

The structure of this complex experience is reported in this paragraph by a subdivision into three different parts, as follows: the elaboration of the new sampling protocol, the measurements in the hospital, the measurements in the laboratory.

4.3.1 Elaboration of the sampling protocol

The ultimate sampling protocol used in this experimentation is illustrated in figure 4.3.1. On the basis of this figure it will be possible to present all the components, the materials and the potentiality of this apparatus.

Fig. 4.3.1:: Sampling protocol in the hospital



There are some elements to point out in this figure, in order to better explain the novelty of this method.

The 3-ways valve is the same already used and validate in the previous experience. The dead space bag is connected to this valve without a supplementary valve, because there is no need to collect and preserve it.

The alveolar breath, instead, is conveyed in another bigger valve, made by teflon, an inert material, with two way out. This valve has two tasks: the first is to allow to stop the filling of the bag in case of problems due to the difficulties of a lung diseased individual; the second is the possibility to take a small volume of the collected alveolar breath, by the means of a needle connected to the septum of the valve, for extra analysis to perform in parallel with the e-nose, such as a GC-MS measure.

The bag for the collecting of the alveolar breath has two way; the reason of this shrewdness is due to different needs: to use two different way for the inlet and for the output of the breath sample and the possibility to disconnect the bag from the whole system and preserve it waiting for the measurement.

4.3.2 Measurements in the hospital

The two departments chosen in the 'C. Forlanini' hospital are strategical for this experimentation. Actually, besides the ambulatory of bronchoscopy, crucial for an early diagnosis, the other department was the C.U.B.E. (Centro Universitario Broncopatie Emergenti, *University Centre for Emerging pathologies of the Bronchial Tubes*) where all the patients affected by lung disease at different stadiations (at their first date, after a surgical operation, during a medical therapy) are addressed to the correct place in the hospital.

The total number of measurements performed in the hospital is 96. Of this 96 volunteers 36 are healthy individuals, 60 are ill patients. This 60 ill patients can be divided into two big groups: 31 patients affected by lung cancer and 29 persons affected by other pathologies. It can be observed that the three main classes have quite the same numerical weight in the total population.

Each of the last two groups of tumors and other pathologies are divided again in different pathologies and different tumoral formations. Anyway this list of sets and subsets is very articulated and it could be quite useless to write it here, so the situation is well described in figure 4.3.2, where a 'three' of the

classifications is reported, with some anticipation about the prediction potentiality of the e-nose in the discrimination between all these classes.

Looking at this scheme the obtained results can be presented as a path of data analysis elaboration at the different levels of the classification 'tree'.

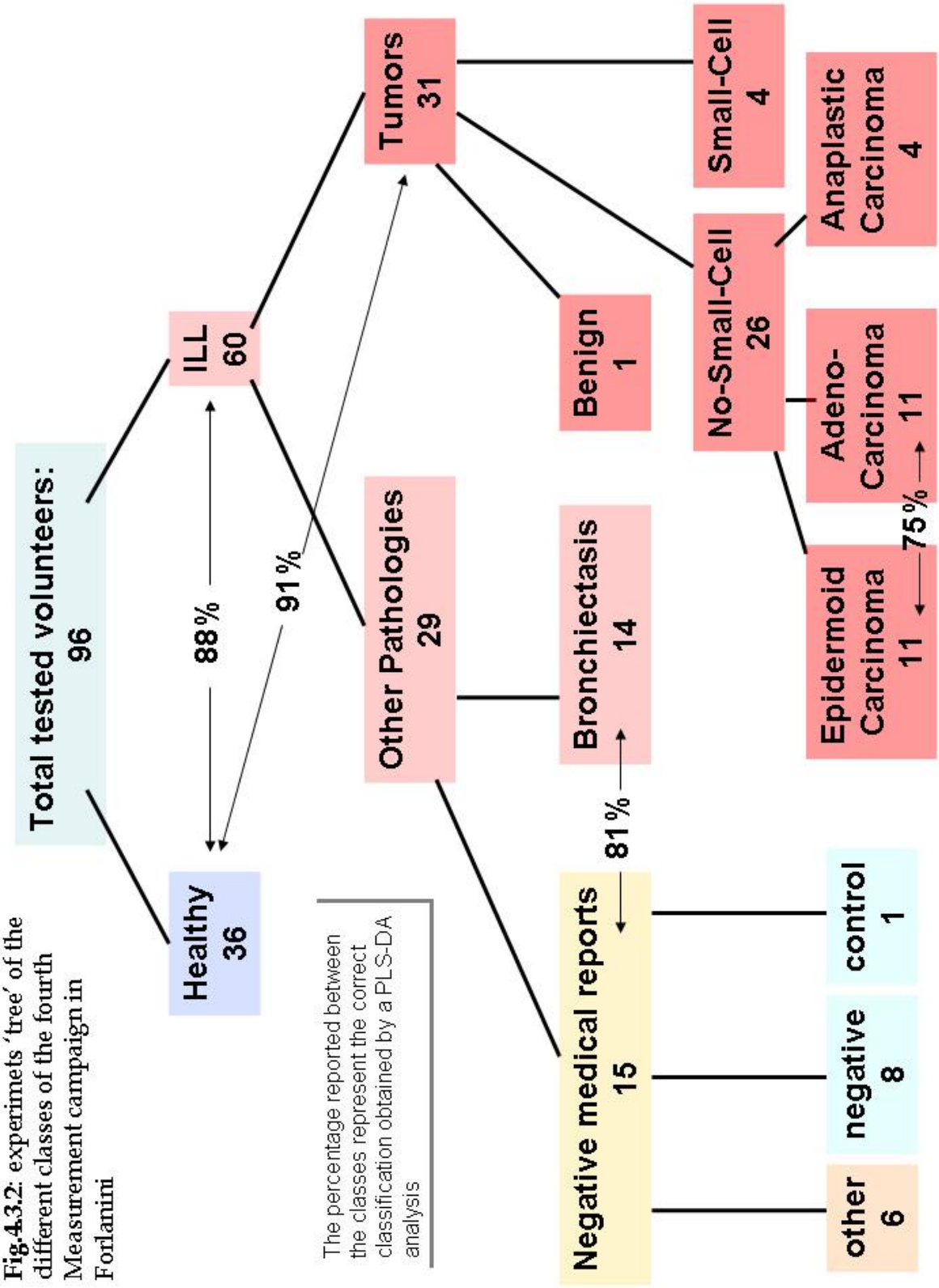


Fig.4.3.2: experimets 'tree' of the different classes of the fourth Measurement campaign in Forlanini

The data analysis here presented has been performed by partial least square-discriminant analysis (PLS-DA). All Data Analysis have been performed in MATLAB, PLS-DA has been calculated using the MATLAB PLS TOOLBOX 2.0 (Eigenvector Research Inc.).

4.4 Measurement in the hospital: results

The scheme in figure 4.3.2 and the classification studies reported below have been obtained after a parallel comparison between the medical report given by the hospital staff and a form filled by the volunteers before the measurement.

Healthy-ill Discrimination:

The first level in this analysis, is the discrimination between healthy persons (reference control chosen in the hospital and in the university staffs), and ill patients (a big class including individual affected by different lung pathologies and patients whose medical report is negative, at least about lung diseases).

The scoreplot of the first two latent variables of the PLS-DA model elaborated is reported in figure 4.4.1. The label 0 is for healthy, 1 for ill.

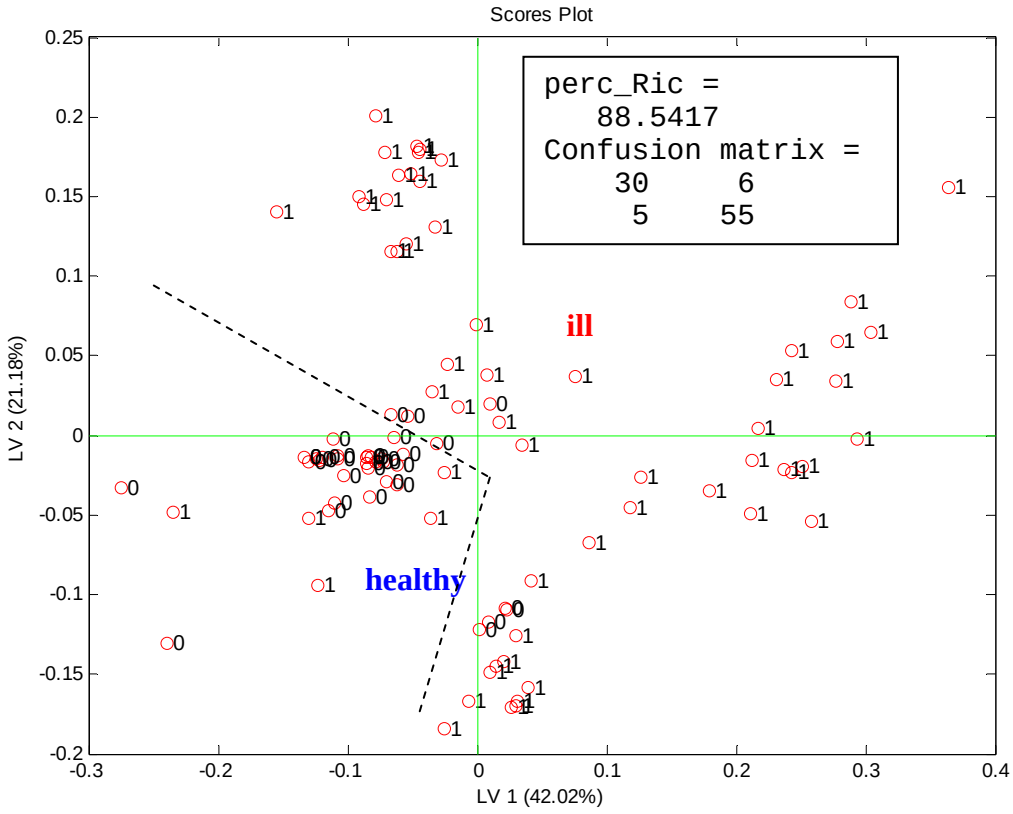


Figure 4.4.1: scoreplot of the first two latent variables of the PLS-DA model; two classes: 0.healthy, 1:ill.; confusion matrix; percentage of correct classification.

In the scoreplot a good discrimination between the two classes can be observed. The percentage of correct classification is high but not so near to 100%. Moreover, in the two clusters it is evident a subdivision in other different subsets. This last observation is useful to understand the 11 misclassified cases. Actually, it must be considered that inside the ill clusters many different classes are grouped, as can be seen in figure 3.4.5. Besides the presence of different stadiations and different typologies of the tumor cells, there are many other pathologies and a little group of negative medical reports which it is not so difficult to imagine very similar to a healthy sample.

This result suggests to study the discrimination of the different subsets individuated to better understand the abilities of the e-nose, by deeper analyzing the levels of the cited classification 'tree'. So the next step consists in the elaboration of another PLS model on two different sets of data both including the healthy reference group: the first comparing the tumors with the other pathologies, the second including the tumors and the negative medical reports.

Healthy-tumor-negative-other pathologies Discrimination:

The results obtained in these two groups are not so good in term of prediction. There are two reasons: the low number of measurements in the negative and in the other pathologies datasets respect to the healthy and the tumor ones (lower than half), and the peculiarity, expecially of the negative cases, to be misclassified as healthy. To put this assertion to the test, the negative measures has been given as healthy in the elaboration of the PLS-DA model, with the result of an increase of the percentage of correct classification from 69% to 75%.

The successive question is about the capability of discrimination of the e-nose in between healthy and lung cancer, means reducing the ill patients dataset to the measurements relative to the tumor cases.

Healthy-tumor Discrimination

The whole set analyzed includes 67 measurements, of which 36 are healthy individuals and the others are 31 patients affected by lung cancer. In figure 4.4.2 is reported the scoreplot of the first two latent variables of the PLS-DA model elaborated.

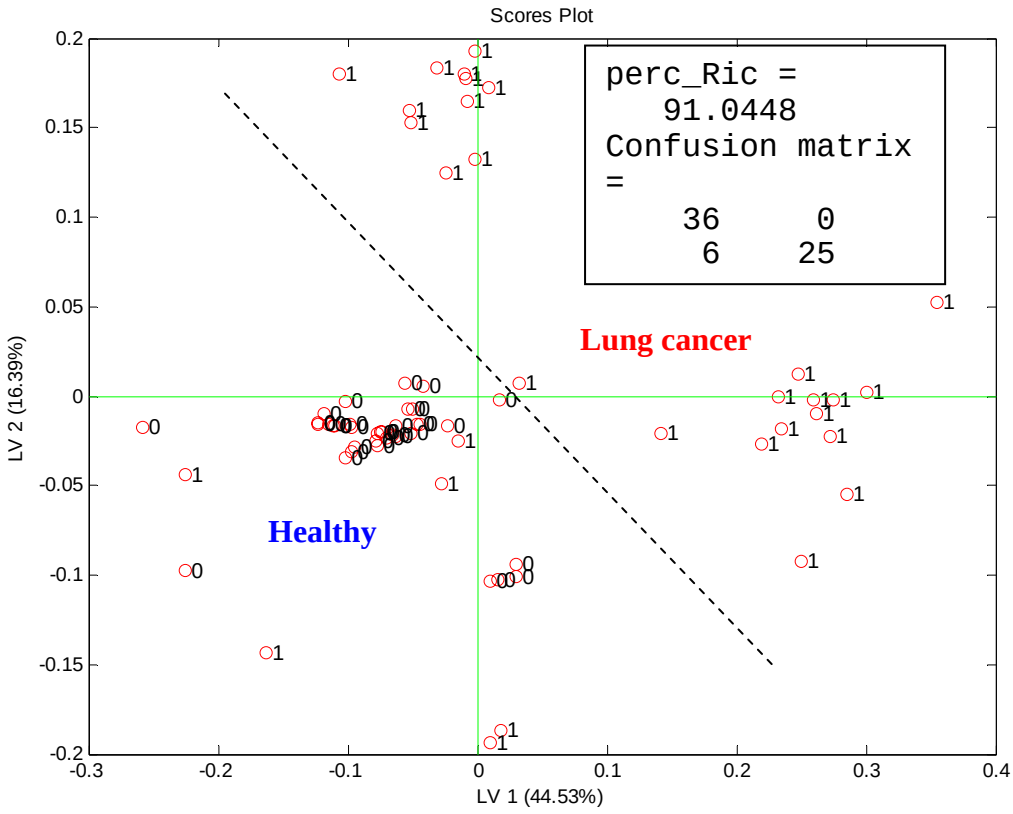


Figure 4.4.2: scoreplot of the first two latent variables of the PLS-DA model; two classes: 0.healthy, 1:lung cancer.; confusion matrix; percentage of correct classification.

Many interesting things can be observed in this scoreplot. The discrimination between the two set is clearer than in the previous representation and this is due to the reduction in the generalization of the ill patients group. In particular, inside the ill set, all the measurements referred to ambiguous negative medical reports has been eliminated, these cases representing a forced grouping in a too general class of ill patients. The obvious result is the 100% correct prediction of the healthy individuals. The correct classification of the healthy is also evident in the figure, in which the cluster of the 0 labels looks very concentrated, instead of

the spread distribution of the lung cancer samples in the different subsets already noted. The suggestion of these considerations for a successive analysis is the need to proceed further into two different studies: one relative to the other pathologies, the other relative to the different Lung cancer forms.

Discrimination between different tumoral forms

It could be very interesting to test the ability of the e-nose in discriminating between the two main typologies of lung cancer: small-cell and no-small-cell; the problem is the fact that the number of the small-cell tuomrs is very low respect to the no-small-cell.

So, the next step is to test the possibility to discriminate the different classes inside the no-small-cell group. On a data set of 22 measurments (11 classified as epidermoid carcionoma and 10 as adenocarcinoma) a PLS-DA model has been elaborated. In figure 4.4.3 the fisrt two latent variables of this model are reported.

In the figure a good discrimination between the two typology can be observed, with a percentage of correct classification of 76%. The third typology, anaplastic carcinoma, has not be inserted in the dataset because the reported case is only one

Moreover it is possible to observe particular situations for some of the misclassified samples. Indeed, very often, the presence of a tumor promotes the raising up of many thoracic complications which can influence the measure.

Even if the presence of this further pathologies is not present in a so considerable number to have a statistical relevance, it could be a reason to explain some misclassifications.

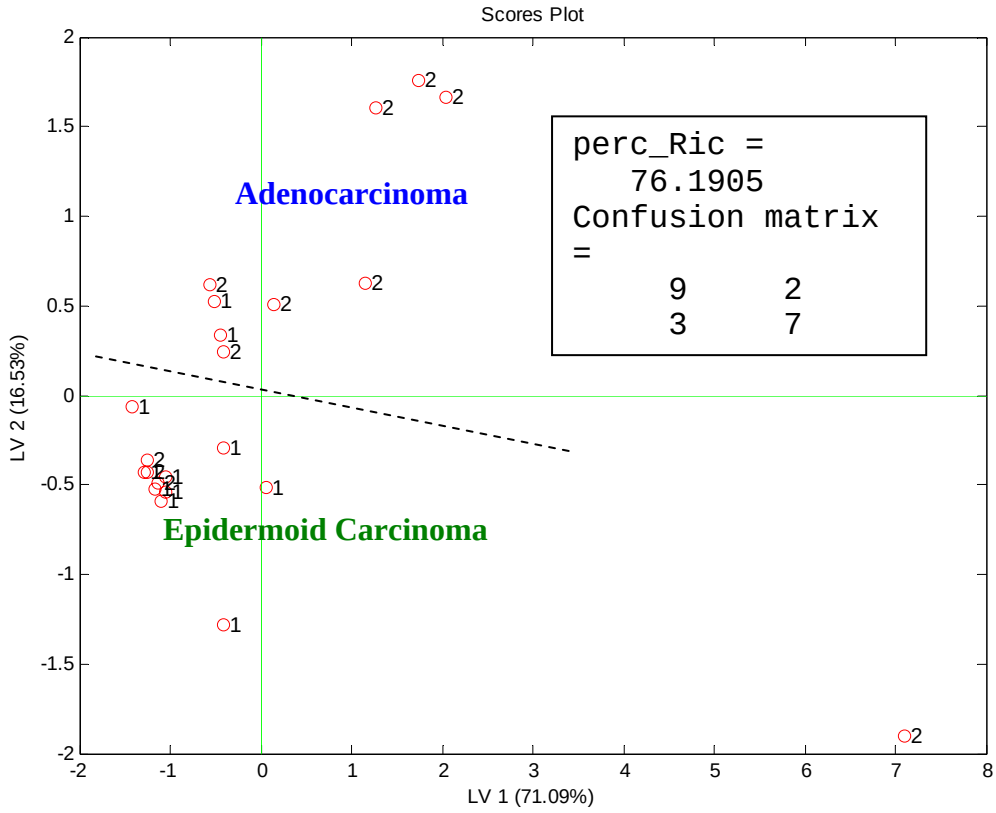


Figure 4.4.3: scoreplot of the first two latent variables of the PLS-DA model; two classes: 1.Epidermoid carcinoma, 2: Adenocarcinoma.; confusion matrix; percentage of correct classification.

Healthy - Negative medical report – bronchiectasis Discrimination

The last study performed is about the left part of the classification 'tree'. In the elaboration of this PLS-DA model the same problem pointed out by the study with healthy, negative and tumor emerges. Actually the percentage of correct classification is very low, about 67%. But the misclassified samples are the negative ones, identified as healthy. The test performed modifying this classification from negative to healthy, not so far from reality, gives a percentage of correct classification of 81%.

4.5 Measurements of possible Lung Cancer Markers

As mentioned before in the introductive sections of this chapter, in many cross sectional studies about lung cancer identification by the means of breath analysis, a list of possible markers has been idenytified [3,4]. In this list five markers have been selected for this experiment: anlinine, o-toluidine, benzene, cyclopentane, hexane.

The experimental set was not so different from the protocol used in the hospital. Indeed the individual had to fill the bag with the same collecting apparatus. After the filling of the bag, a volume at a fixed concentration of the marker was added to the bag from the second way positioned at the opposite side respect to the inlet of the breath.

The sampling protocol used is illustrated in figure 4.5.1

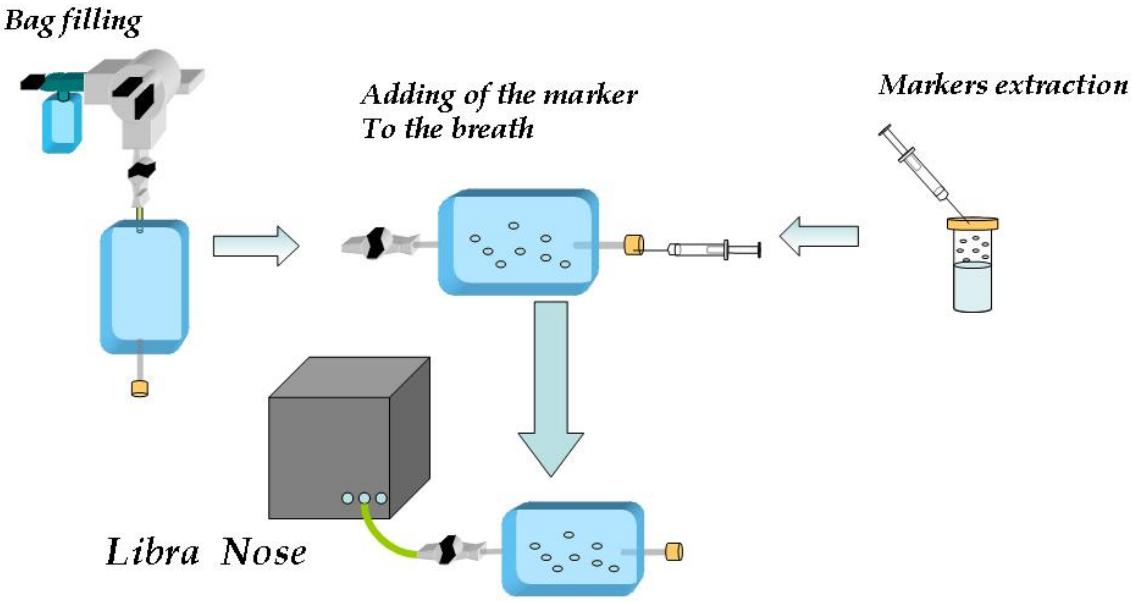


Fig. 4.5.1: experimental set for the measurements performed in the laboratory with the possible tumor markers.

The strategy aimed to study the classification of a healthy breath added with the possible tumor markers. The whole data set of 96 was the good statistical basis on which testing this situation.

The range of concentrations of the added volumes were 200-400 ppb, as reported in literature.

The markers has been added alone and together. In figure 4.5.2 is reported the scoreplot of the first two latent variables of the PLS-DA model elaborated on the whole data set of measurements in the hospital plus the measurements performed in the laboratory with the markers.

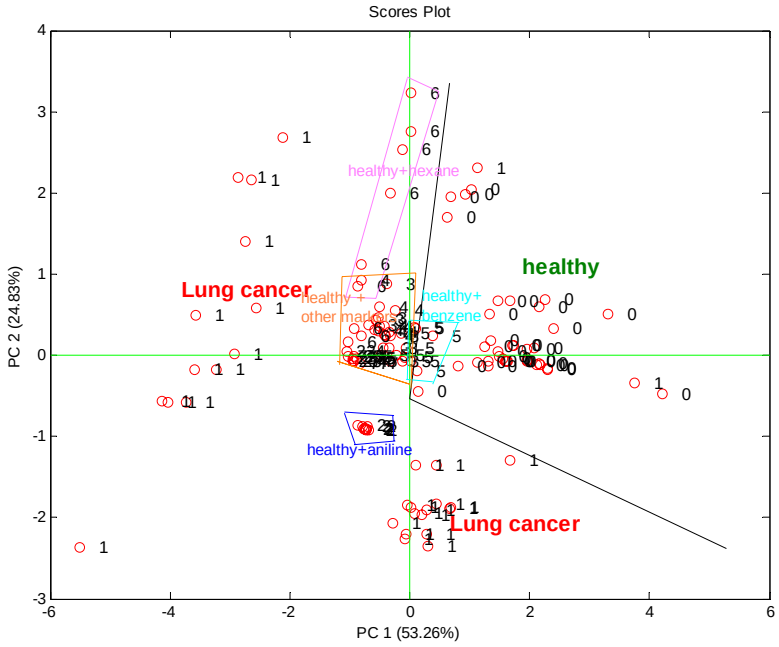


Fig. 4.5.2: scoreplot of the first two latent variables of the PLS-DA model elaborated on the two data set of the hospital and of the laboratory experiment about lung cancer markers

In figure a very important confirmation can be founded: the breath of a healthy individual, added with the lung cancer reported in literature, in the concentration range considered as abnormal, is classified by the e-nose as a tumor sample, moving from the original region of healthy individuals measures.An important implication of this result is the observation of the great utility of the e-nose used in parallel with a GC-MS apparatus. The contemporary presence of these compounds shows the GC-MS has identified a group of VOCs which probably cooperate together to reveal the presence of the tumor, but this cooperation appears only as clue in the anomalous concentrations of each single compound, while it is evident its effect when measured as odor, in its globality.

At the same time the GC-MS shows to be an useful instrument to give the reasons for the obtained odor classification.

A future perspective of this experiment is of course the use of GC-MS to measure all the samples tested with the e-nose, and to complete the set of measurements relative to whole list of possible markers identified (table 3.3.2).

4.6 References

[1] C. Di Natale, A. Macagnano, E. Martinelli, R. Paolesse, G. D'Arcangelo, C. Roscioni, A. Finazzi Agrò, A. D'Amico 'Lung cancer identification by analysisof breath by means of an array of non-selective gas sensors'. *Biosensors and Bioelectronics* 18, 1209–1218 (2003).

[2] Phillips M. 'Method for he collection and assay of volatile organic compounds in breath'. *Anal Biochem* 1997;247: 272–8.

[3] Phillips, M. *et al.* 'Detection of lung cancer with volatile markers in the breath'. *Chest* 123, 1788–1792 (2003).

[4] O'Neil, H.J., Gordon, S.M., O'Neil, M.H., Gibbons, R.D., Szidon, J.P., 1988. 'A computerized classification technique for screening for the presence of breath biomarkers in lung cancer'. *Clinical Chemistry* 34, 1613_ 1618.

ENVIRONMENTAL MONITORING

Il Sole, con tutti i pianeti che gli ruotano attorno e da esso dipendono, può ancora maturare un grappolo d'uva come se non avesse nient'altro da fare nell'Universo.
Galileo Galilei

The sun, with all those plants revolving around it and dependent upon it, can still ripen a bunch of grapes as if it had nothing else in the universe to do
Galileo Galilei

5

5.1 Environmental monitoring: existing technologies.

The problem of the environmental monitoring is articulated. The complexity is due to many reasons, but mainly to the different sources of pollution (industrial district, rubbish dump, wastewater treatment plant), and to the different targets of monitoring.

This chapter will present case studies relative to the whole set of situations concerning with the pollution sources above mentioned.

About the target of the monitoring, there are two main objectives in a environmental air quality monitoring of an area including one of the three cited polluting elements: to detect the presence of specific volatile compounds and their quantification (*Pollution Monitoring*), and to monitor the presence of olfactive nuisance due to plants or industries malfunctioning (*Olfactive Impact Monitoring*).

Pollution Monitoring

Although the numerous instrumental controls available to reduce the emission of polluting compounds such as CH_4, CO_2, H_2S, the problem of the pollution monitoring is of primary importance, because of the necessity of more precise systems able to keep the emission concentrations under the security thresholds [1].

At present there are many inadequacies associated with the most common methods used in environmental monitoring. Let us start our illustration from the wastewater treatment plants. These techniques are in particular devoted to the measurement of biodegradable organic matter by the 5-day biochemical oxygen demand test [2]. Normally the quality of treated wastewater is monitored by the measurement of global parameters such as Biological Oxygen Demand (BOD), Chemical Oxygen Demand (COD), Total Organic Carbon (TOC), and Total Suspended Solids. There is a list of problems related to these techniques: the time required to complete the test (5 days), the difficulty in achieving reproducible measurements, the lack of real time monitoring and control of sewage treatment works, the requirement of constant contact with the wastewater (resulting in instrument fouling, requiring frequent cleaning and re-calibration of the

monitoring system). These problems can be summarized as the necessity of a reproducible and non-invasive devices.

There is a great quantity of alternative techniques and technologies available for monitoring changes in organic load in wastewater. Biosensors, although very sensitive and selective, have short lifetimes, from a few days to a few months, which limits their application to continuous on-line monitoring. Optical sensors present the advantage of rapidity, versatility, low running-cost, absence of chemicals and limited or absence of sample handling, but also introduce problems like biofouling of the probe tips, calibration stability and selectivity [1].

The electronic nose, being in principle non-invasive and versatile, has a great potential for real-time and on-line monitoring of wastewater and air quality[2-7].

Olfactive Impact Monitoring

The increasing of complaints about odour emission in industrial districts meets the lack of acknowledged methods to solve olfactive nuisance quarrels; this situation stimulated interests in odor measurements to define objective methods to certify odor emissions.

Current methods uses to evaluate odor intensity exploiting the reactions of expert panellists. Sometimes H_2S concentration is used as a stimulant to determine odor strength. More widely used is the evaluation of threshold odor numbers (TON), defined as the number of dilutions at which 50% of panellists can detect no odor. Some experiences in odor environmental monitoring have shown there is not a clear relationship between TON and the concentration of H_2S. These differences could originate from the fact that different sewage receive strong industrial component in their flow and this may account for differences in H_2S emission rates, so H_2S may be an unsuitable surrogate for the determination of odor concentrations. Moreover, methods used for TON determination are time consuming, labour intensive, expensive and subject to large variation between panellists and laboratories.

E-nose in environmental monitoring: a short overview

In this paragraph an overview focusing on efforts to employ an electronic nose to monitor airborne volatile organic compounds that are released when waste products are dumped in water, land, or air is given.

Water monitoring:

Baby et al. [7] used the hybrid technology MOSES II e-nose to measure contaminating residues of insecticides and products from leather manufacture that are often offloaded in into streams and rivers.

Dewettinck et al [3] employed an e-nose consisting of 12 metal-oxide sensors to monitor volatile compounds in the effluents of a domestic wastewater treatment plant. In another study by the same group, Van Hege et al. [3] explored the application of evaporative technology as an alternative desalination technique for wastewater treatment plant effluents.

Bourgeois and Stuetz [8] reported the use of a 12 polypirrole conducting-polymer sensors array to analyze wastewater sampled sparged with N_2 gas in a temperature-controlled flow-cell. They also examined the use of a real-time sensors and array system for monitoring global organic parameters such as biochemical oxygen demand and total organic carbon [1,2].

Di Natale et al. [5] used a sensor array of ion-sensitive electrodes to analyze polluted water. This kind of devices that use sensor arrays to test liquid samples are usually called electronic tongues rather than electronic noses.

Gardner et al. [4] and Shin et al. [9] developed a system for detecting cyanobacteria in potable water.

Land monitoring:

There have been few research studies in this area. One example is the experiment by Biey and Verstraete [10] about the use of a 5-W UV lamp, generating ozone for seven hours per day, to reduce the odours produced by the decomposition of kitchen and vegetables waste. The efficency of this method has been tested, claiming good results, by an Alpha M.O.S. FOX 3000 e-nose.

Air monitoring:

Air quality has been the primary target of e-nose research projects in environmental monitoring [11,12]. Indeed, although in most cases annoying atmospheric emissions do not menace public health, they do greatly reduce the quality of life [13].

Many experimentations has been performed in this field, with an equally number of different tasks.

Odour abatement and control in municipal sewage treatment facilities: Gostelow et al. [14] reviewed various sensory, analytical and e-nose methods for monitoring sewage facility emissions. Stuetz et al. [2,15] employed a Neotronics NOSE to investigate the same matter considering also the effects of biofilters and the evaluation of Hydrogen sulphide concentrations.

Indoor air quality: Schreiber and Fitzner [16,17] studied the perception of the quality of indoor air by building inhabitants. Delpha et al. [18,19] investigated the use of an e-nose using metal-oxide TGS sensors for the detection of a leaking refrigerant gas (Forane R134a) in an air-conditioned atmosphere. Sarry and Lumbreras [20], investigated the detection of carbon dioxide, Forane R134a, or their mixtures by the means of a five tin-dioxide sensors. Ramalho [21] analyzed the characteristics of indoor paints and their effect on perceived indoor air quality.

Odourous emissions from animal production facilities: Hobbs et al. [22] correlated e-nose measurements of pig manure odours to those of a human panel. Four of the principle odourous compounds in pig manure were selected for the study.

Willers et al studied the efficiency of biofiltration of mice cages air with QMB based electronic nose [23].

All these studies obtained good results in terms of potentiality of discrimination and in the correlation with the classical methods results, panel test included.

In the next paragraph a short presentation of the mentioned methods of olfactometric tests by the means of panellists are reported, comparing this techniques with the e-nose in terms of capability, sampling problems and objectivity.

5.2 Olfactometry and e-nose

Olfactometry consists in the evaluation of the olfactive annoyance (in this application) of a mixture by a jury of selected persons (panellists). The olfatcometric analysis, on the contrary of the chemical one, does not provide any identification of substances, but the odor units of the gaseous mixture. This method quantifies the olfactive impact translating this concept from the subjective sensation into an objective number. As written in chapter 1, electronic nose performs a similar task with a difference between them about the objectivity of the assessors. Indeed, while chemical sensors directly produce an array of number, panellists give an answer that has to be translated in numbers. This translation is usually done considering the number of dilutions necessary to reach the limit of perception of the panellists, this number is the odour unit for the volume (OU/m^3).

The criterion reported in the normative law for the panellists recruitment is based on their sensitivity to normal-butanol. Considering that the mean concentration value of this substances for human is 40 ppb (123 $\mu g/m^3$), the sensitivity to the normal-butanol of a panellist should be in the range 20÷80 ppb (62÷246 $\mu g/m^3$). The used unit measures are reported below:

1 EROM (European Reference Odour Mass) corresponding to 123 $\mu g/m^3$ *of normal-butanol.*

1 OU_E *(European Odorimetric Unit) consisting in 1 EROM evaporated in 1* m^3 *of neutral gas in standard conditions.*

Before starting the studies about environmental monitoring the electronic nose has been tested as panellist. It has shown to be a good panellist, as can be seen in figure 5.2.1.

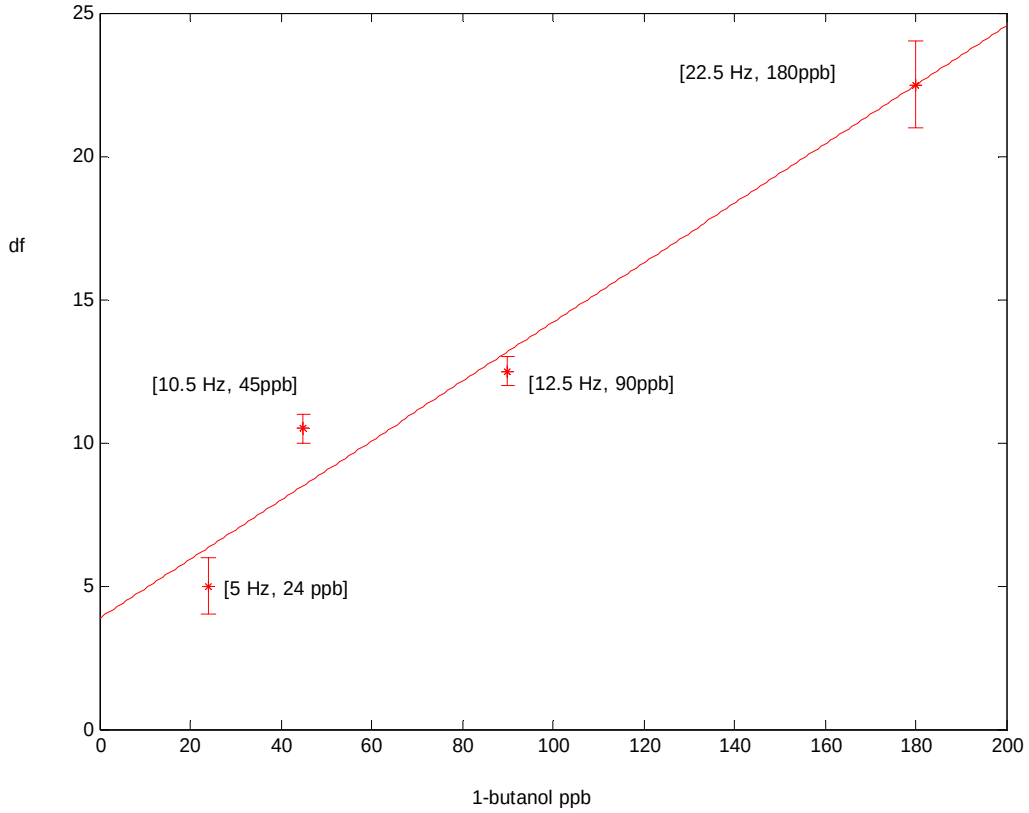


Figure 5.2.1: Registered frequency shift for one sensor in the array, calculated as the mean of the values obtained in three measurements
(the measurements were performed with a static sampling protocol (see chapter 6)).

A response of 6 Hertz by the electronic nose is low but anyway significant, and very close to the inferior limit of the established range for a panellist.

The response to the average concentration considered in the calculus of the OU_E is of 10 Hertz. It is worth to remark that for concentrations after the value of 24 ppb of 1-butanol no appreciable frequency shift has been registered.

The measurements were performed using N_2 as reference gas. This gas was used to clean the sensors and to fill the measure chamber. In the measure chamber the fixed concentration of 1-butanol was inserted by the mean of a GC-syringe. After a certain time interval the sensor chamber and the measure chamber were put in communication to perform the measurement.

The e-nose used for the 1-butanol calibration was the NoseStat, a version of the e-nose designed for measurements in static conditions. Designing, realization and measure protocol of the static measures are reported in details in chapter 6.

For environmental application Libra Nose has been evolved into MERLINO (MOBILE ENVIRONMENTAL LIBRA NOSE) (see Figure 5.2.2). A portable electronic nose optimized for on-field measurements.

The novelty of this version of the electronic nose of the University of 'Tor Vergata' consists in the fact that the enose core (sensors, electronic and pneumatic), the supply voltage and the battery pack are included in a box of the size of a shoes box.

Figure 5.2.2: MERLINO (Mobile EnviRonmental LIbra Nose), the mobile version of the electronic nose of the University of 'Tor Vergata'.

The weight of the whole system is about 5 Kg. The autonomy of the battery is about 24-30 hours. Power supply can be easily switched from battery to "net".

5.3 'MERLINO' application in environmental monitoring

'MERLINO' e-nose has been applied to monitor the odours of sludge, water and air in a wastewater treatment plant. These three case cover a vast overview of the pollution sources: water and sludge measurements for the monitoring of the processes of a wastewater treatment plant, air quality monitoring in an industrial district including a wastewater treatment plant, air quality monitoring in an area surrounding a rubbish dump.

5.3.1 Wastewater treatment plant: processes monitoring

These measurements were concerned with a large Wastewater treatment plant serving an industrial district of tanneries.

In figure 5.3.1 the map of the plant is illustrated, with some data on the size of the installation.

Fig. 5.3.1: map of the wastewater treatment plant; 150000 mq area, dayly 30000 mc of waste materials; 450 Km Waterworks net; 280 Km Sewers ; 8 Sludges dumps.

Samples have been collected in two of the three main lines of the treatment processes: industrial water and sludge line.

In figure 5.3.2 a general scheme of the three main treatment lines of the plant is reported.

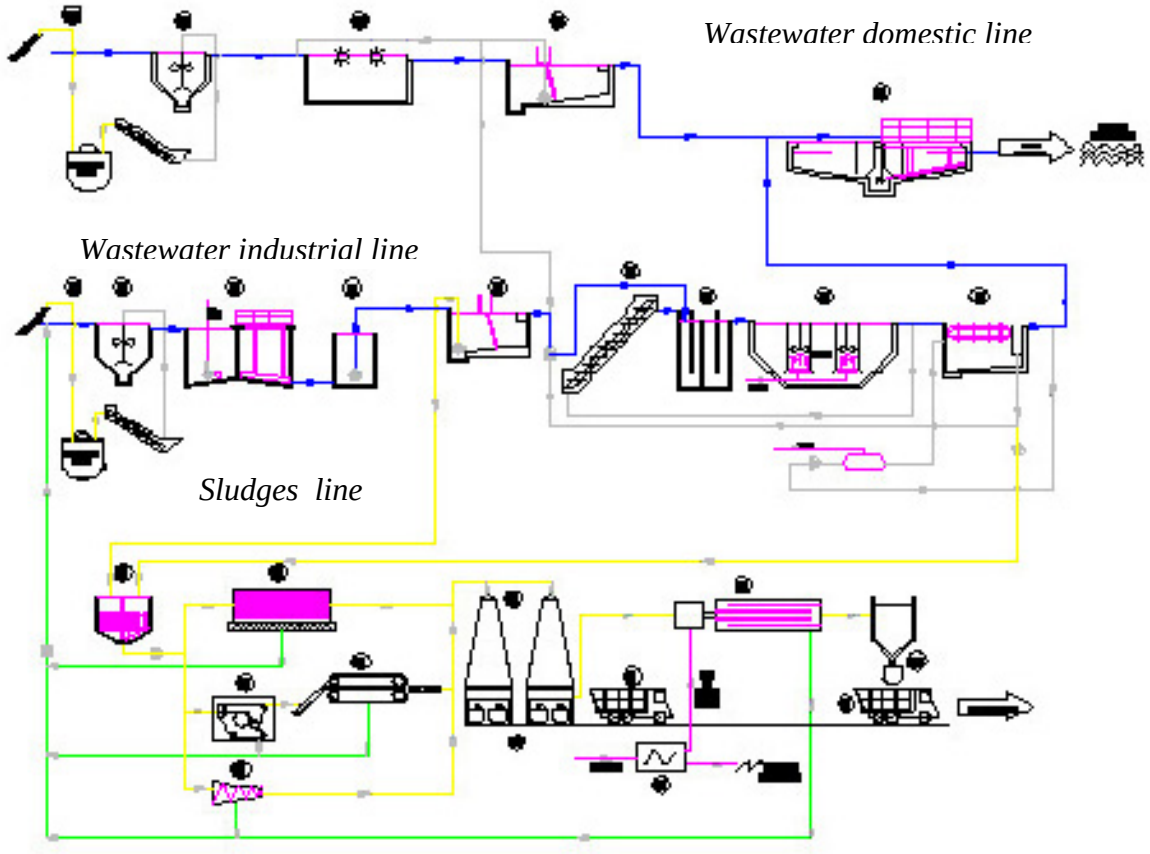


Fig. 5.3.2: scheme of the wastewater and sludge treatment lines in the plant.

Wastewater treatment processes monitoring

Four sample were collected at the wastewater treatment line: at the inlet of the plant, at the output of the grit removal, at the output of the homogenizer and of the nitrification processes. Samples were collected by the plant staff, directly on field during the normal functioning procedures of the treatment, using the tools for the routine analysis. The samples have been extracted from the volume collected, enclosed in glass vials and measured within a maximum time of 15 minutes.

All the measurements relative to the processes monitoring have taken place in the laboratories of the plant. The sensors chamber has been thermalized in order to perform all the measurements at the constant temperature of 30°C.

Frequency shift were calculated with respect to a reference of dry air used for the cleaning phase.

figure 5.3.3 illustrates the four sites, inside the plant, where the measured sample were taken

Fig. 5.3.3: the four sites of wastewater industrial line inside the plant, where the samples for the e-nose measurements have been collected.

In figure 5.3.4 the scheme of the sampling protocol is illustrated: the same source of dry air is connected both to the clean inlet of the e-nose and to the inlet of the sample bottle to use it as carrier gas.

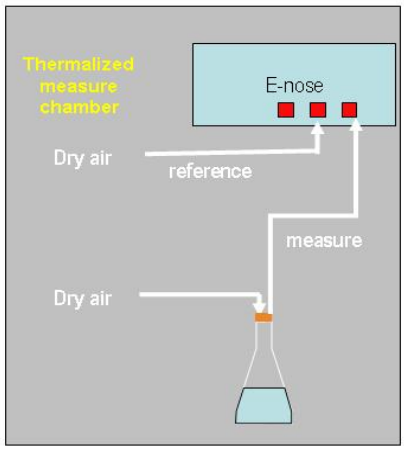


Figure 4.3.4: scheme of the sampling protocol for wastewater measurements

In figure 5.3.5 the MIC (Measure Identity Card) of this experiment is reported .The data analysis has been performed on the collected data, using the PLS technique.

The obtained results are well synthesized in the scoreplot of the first two Latent Variables of the model elaborated (fig. 5.3.6).

w

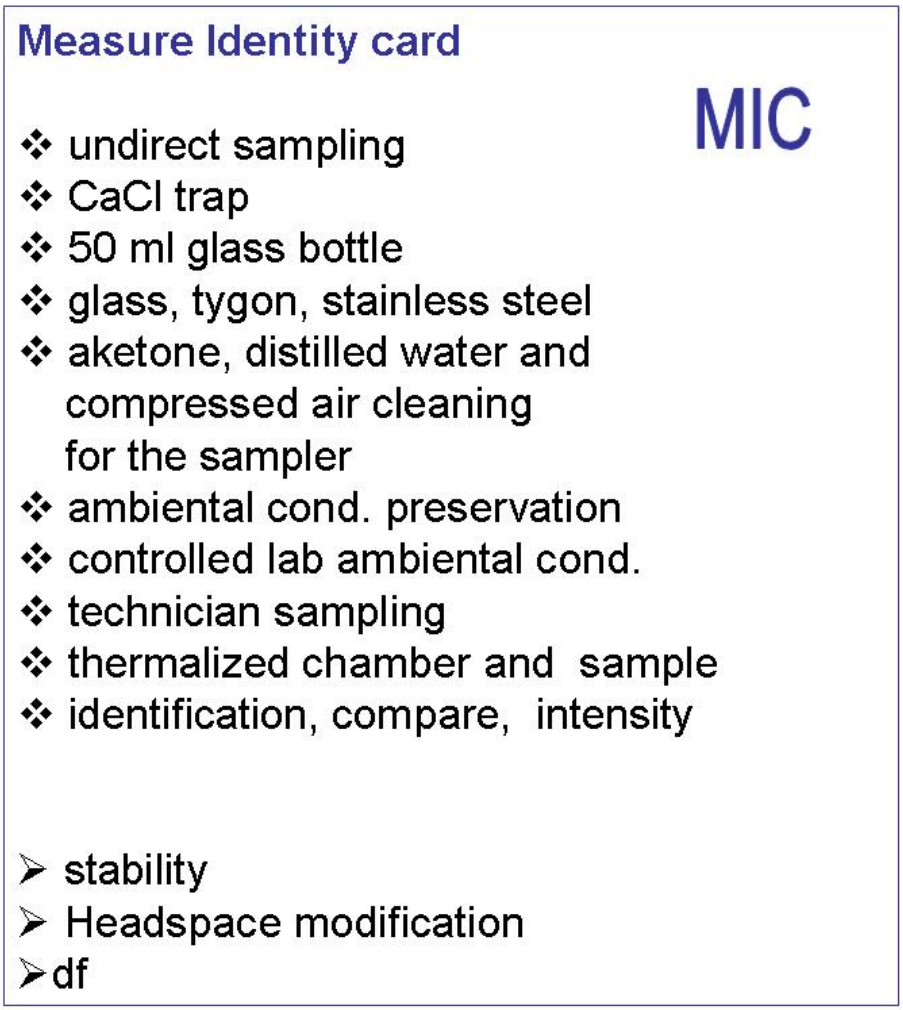


Fig. 5.3.5: measure Identity Card of wastewater measurements

The scoreplot shows that there is clear discrimination between the nitrification outlet samples and the other classes. Meanwhile a partial superpostition between the classes belonging to the large class of plant processes inlet and the homogenizer is observed. This is probably due to the absence of rain the month before the measurement campaign, reducing the most natural source of mixing product for the correct working of the homogenizer.

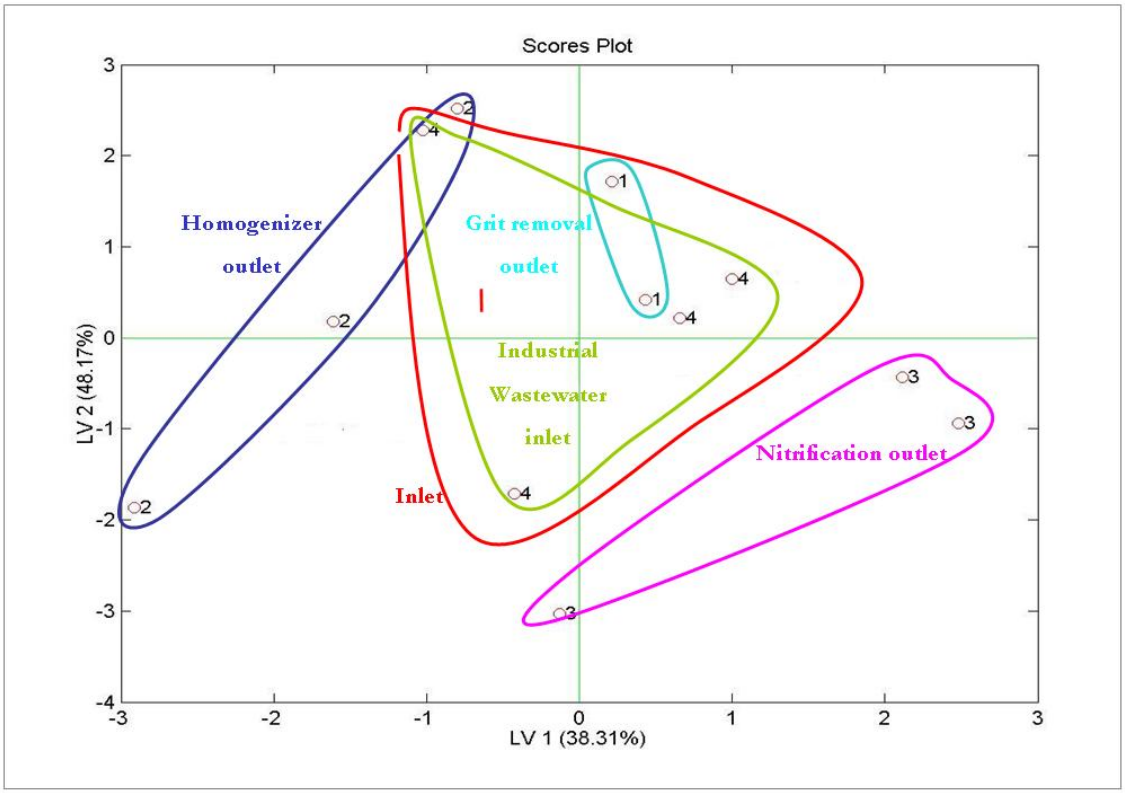


Fig. 5.3.6: scoreplot of the first two latent variable of the PLS model elaborated on the wastewater measurements.

Dewatering processes monitoring

The main treatment performed in the sludge line, is the dewatering process. This treatment takes place in three parallel lines, with three different methods considered as equivalent in terms of results.

The objectives of the e-nose investigations on sludge line, consisted in the evaluation of the efficency of the three methods, comparing the samples collected from the three lines with the samples collected from the dewatering processes common final outlet. Moreover, it has been interesting to compare the three processes, trying to evaluate their olfactive impact. In figure 5.3.7 a detailed view of the sludge line is reported

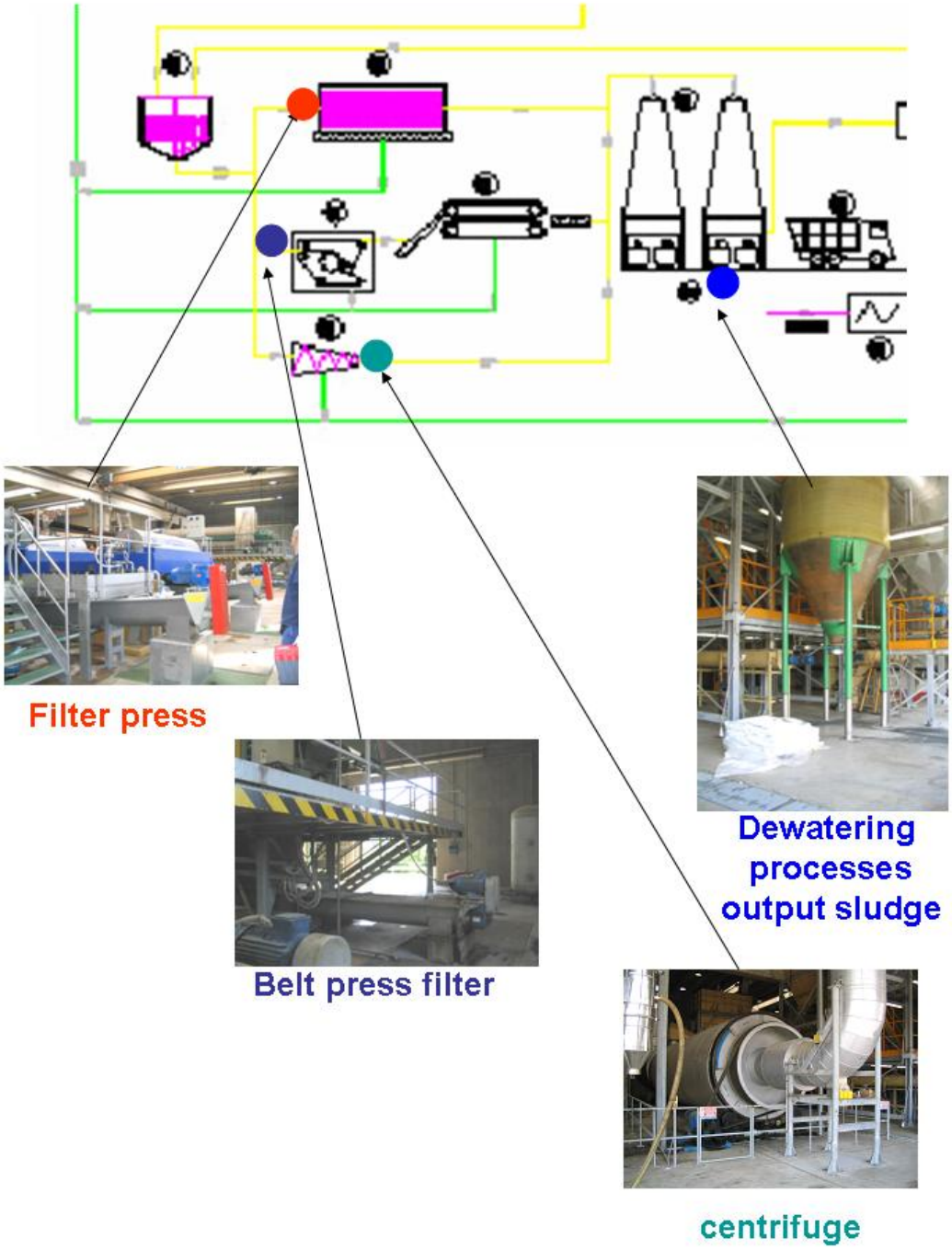


Fig. 5.3.7: Sludges line different dewatering processes.

The protocol used for this set of measurements is reported in the MIC of figure 5.3.8.

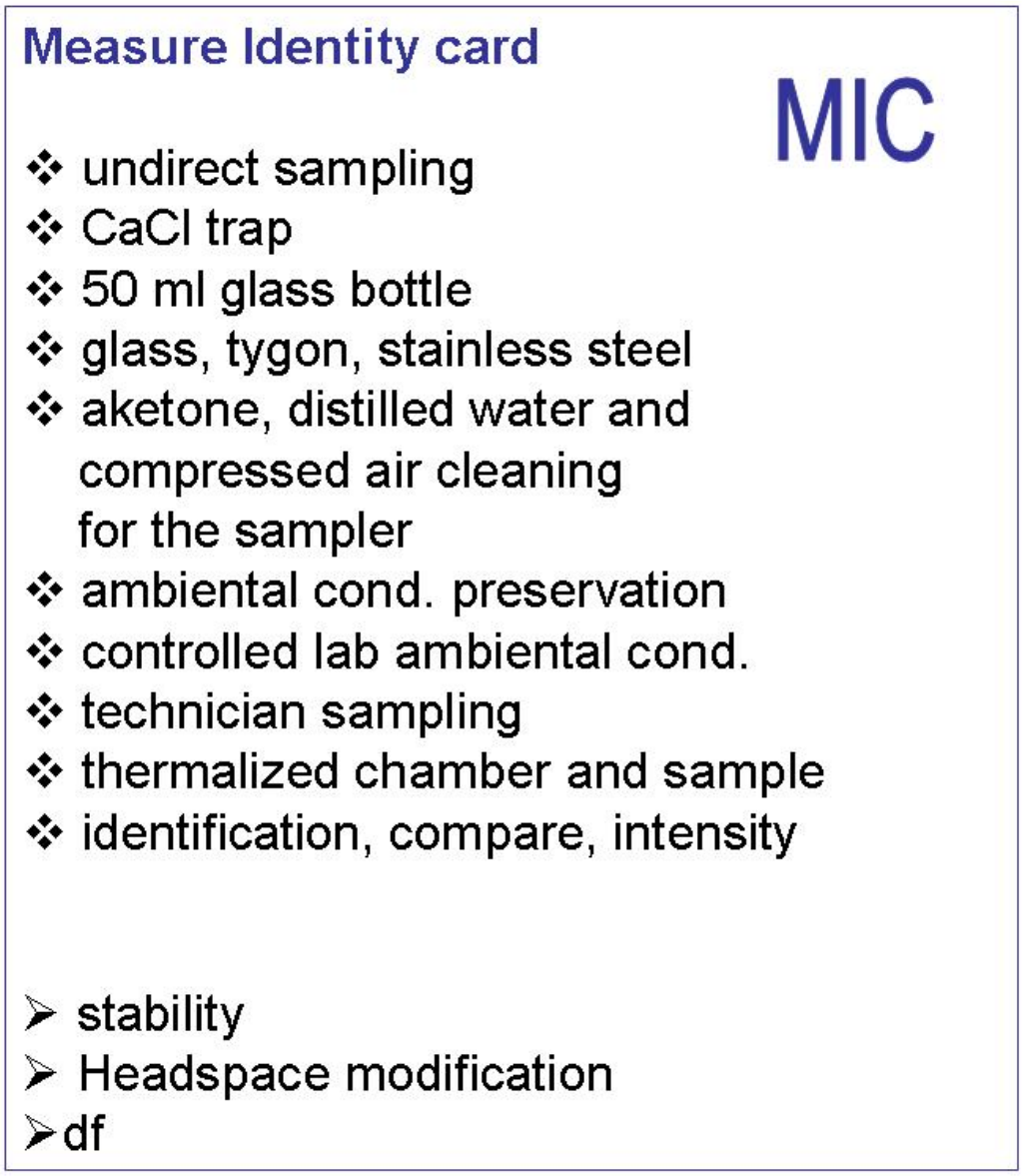


Fig. 5.3.8: MIC (Measure Identity Card) of the Sludge measurements set

E-nose data were analyzed for both quantitative and qualitative purposes.. The Qualitative analysis has been elaborated using the PLS-DA method, and the

obtained results are reported in the scoreplot of the first two latent variables of the model (fig. 5.3.9).

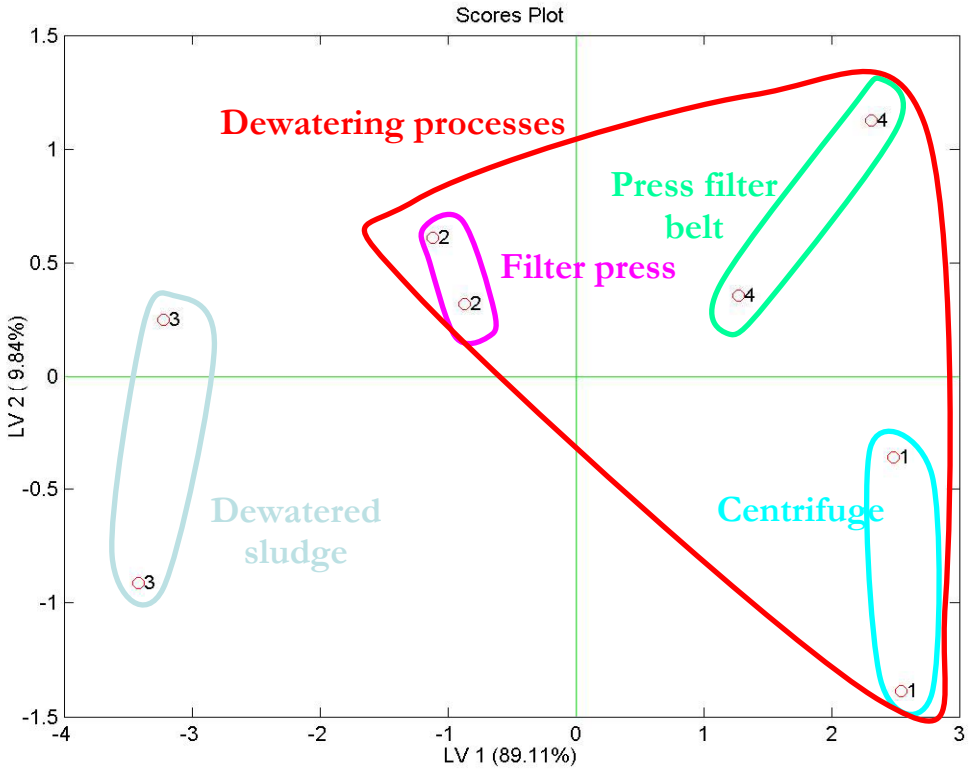


Fig. 5.3.9: scoreplot of the PLS-DA model built on sludge measurements data set.

The scoreplot provides a good discrimination between the final product of the three dewatering processes, and the three processes themselves can be observed. Moreover, inside the well defined clusters representing the three techniques, each of the different methods can be easily distinguished.

The quantitative analysis had the task to evaluate the odour intensities of the measured samples, in order to establish their olfactive impact.

The used calculus is a straightforward application of the Sauerbrey formula. Being the frequency shift of the qmb proportional to the mass graviting on the sensor surface a good estimate of odor intensity could be the sum of the eight mean df, calculated as mean values for each sensor between the measurements relative to the same sample.

According to such index, as expected the lowest olfactive intensity is obtained for the final products of the dewatering processes.

Comparing the estimated odour intensities of the three parallel processes the press filter odor impact is lower than the other. This is far from assessing a preference of this processes with respect to the others, but it provides an interesting suggestion for a deeper investigation.

5.3.2 Air quality monitoring of an industrial district

In this experiment e-nose was used to map odour in an industrial district populated with tanneries and including a big wastewater treatment plant.

As in the previous case both qualitative and a quantitative analysis have been performed. The main tasks of the study was twice: to characterize different odours inside the plant and in the different sites of the industrial district.

The map of the industrial district is reported in figure 5.3.10.

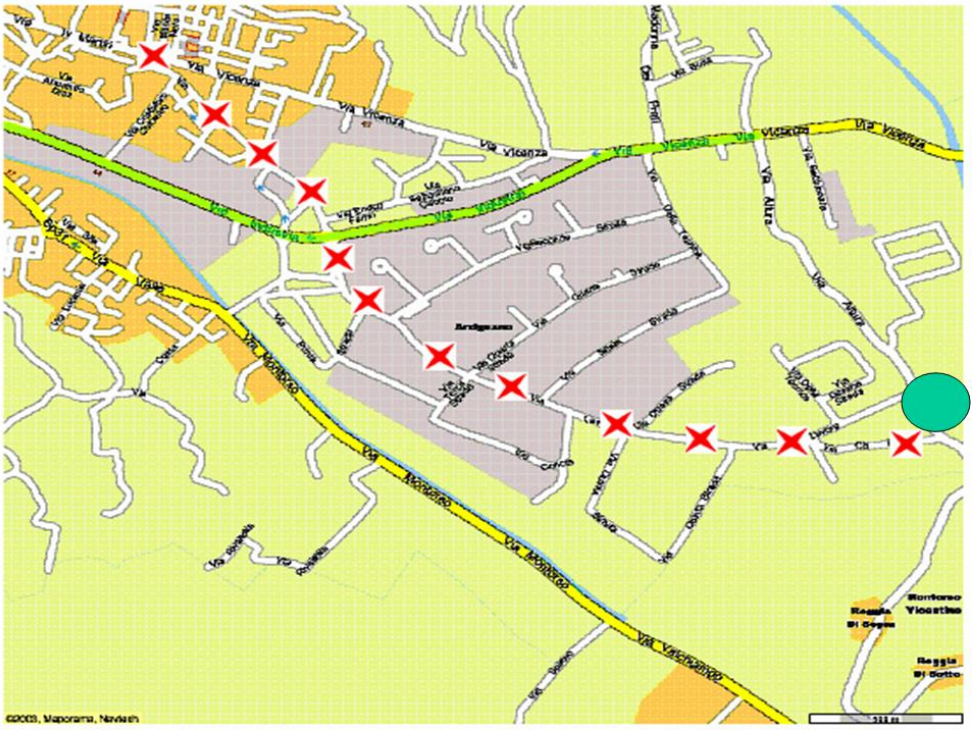


Fig. 5.3.10: map of the industrial district monitored by the means of the e-nose; in the map the measure points are indicated with a series of red crosses, the plant is indicated by a green point.

The MIC of this investigation is reported below in figure 5.3.11.

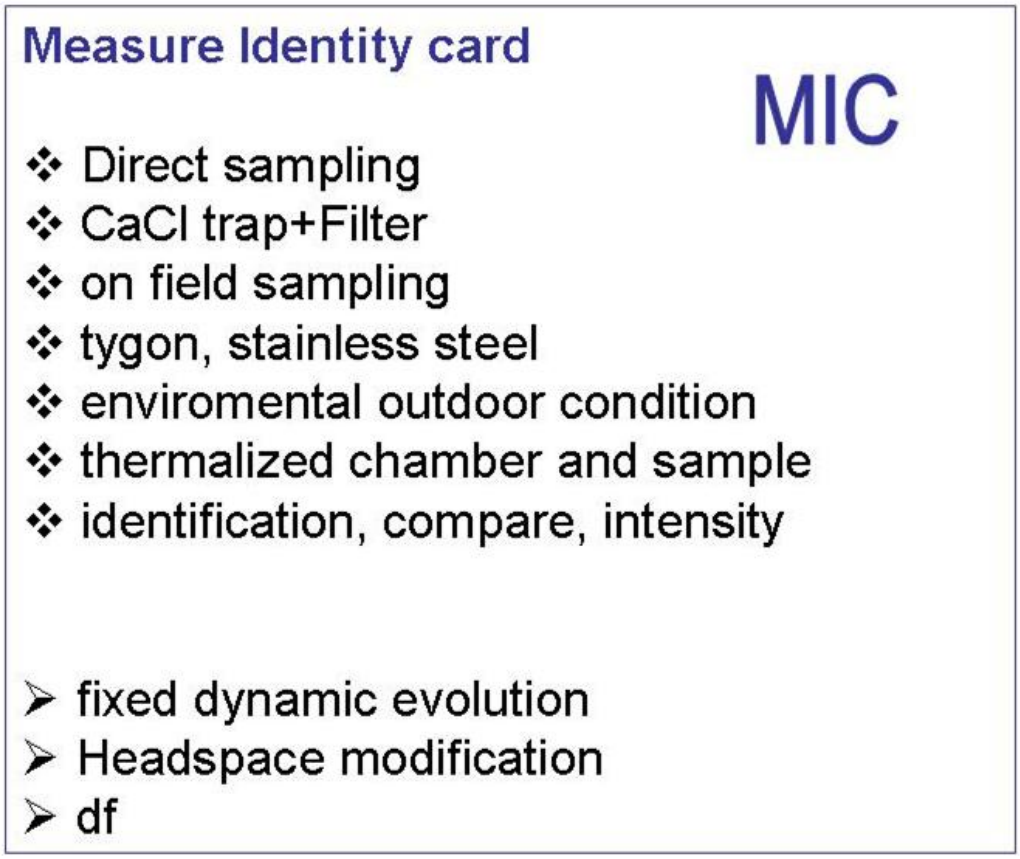


Fig. 5.3.11: MIC of the investigation aiming to monitoring an industrial district

The investigation presented a series of peculiar problems different from usual experimental set up used in the laboratory.

Direct sampling scope is to record the real situation in the district. This method removes the sampling problems, but it raises the hitch of the influence of the environmental conditions. To counteract the temperature variations the sensors chamber has been kept at atmospheric temperature. The relative humidity elimination is a more complex matter; in an open environment, relative humidity disturbance can shadow the presence of other compounds. In a differential measurement where dry air is the reference gas, the response includes the atmospherical humidity content. The implemented solution consisted in the use of a particular filter, sensitive to some polluting compounds.

Moreover there was the necessity of a continuous sampling to measure the largest number of possible site inside the area and to have an amount of data

enough to monitor the change due to the atmospherical conditions. For this reason, the normal technique used for the measurements with the e-nose has been inverted: the e-nose measured continuously inside the industrial district from the plant to the town and vice versa, while a series of cleaning phases has been performed in the sites evidenced in the map (fig. 5.3.10).

A PCA model elaborated has given two important representations of the situation. In the first scoreplot, reported in figure 5.3.12 it is possible to trace the same path followed in the real trip around the industrial district.

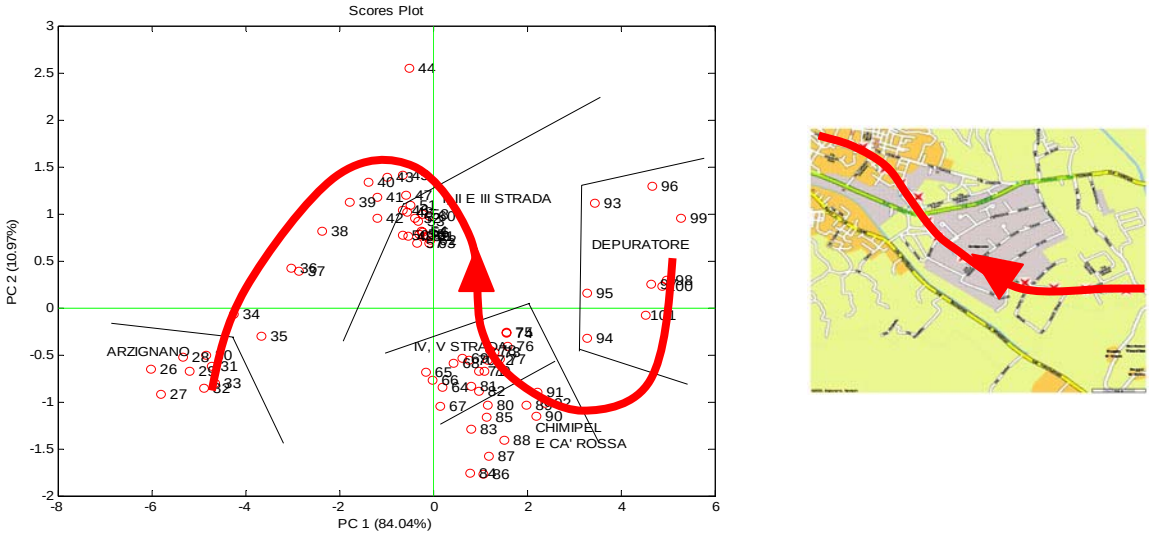


Fig. 5.3.12: Characterization of the different odour sources present in an industrial district, obtained tracking on the score plot of the PCA model built on the e-nose data, the same path followed during the measurements.

In Figure 5.3.13, it is possible to distinguish different clusters, grouping the different measurements on the basis of the nearest tannery. While the scoreplot above has been obtained analyzing the whole data set relative to the continuous sampling, the next scoreplot results from the measurements of the nine selected points. The scoreplot of the first two LV presented, in this case, is the result of a PLS-DA model. As already mentioned, it is evident a good separation between the different sampled site. Moreover, representing also the third LV (figure 5.3.14) a separation of two fundamental group is put in evidence along this variable: the cluster of the town respect to the group of clusters referred to the industrial district.

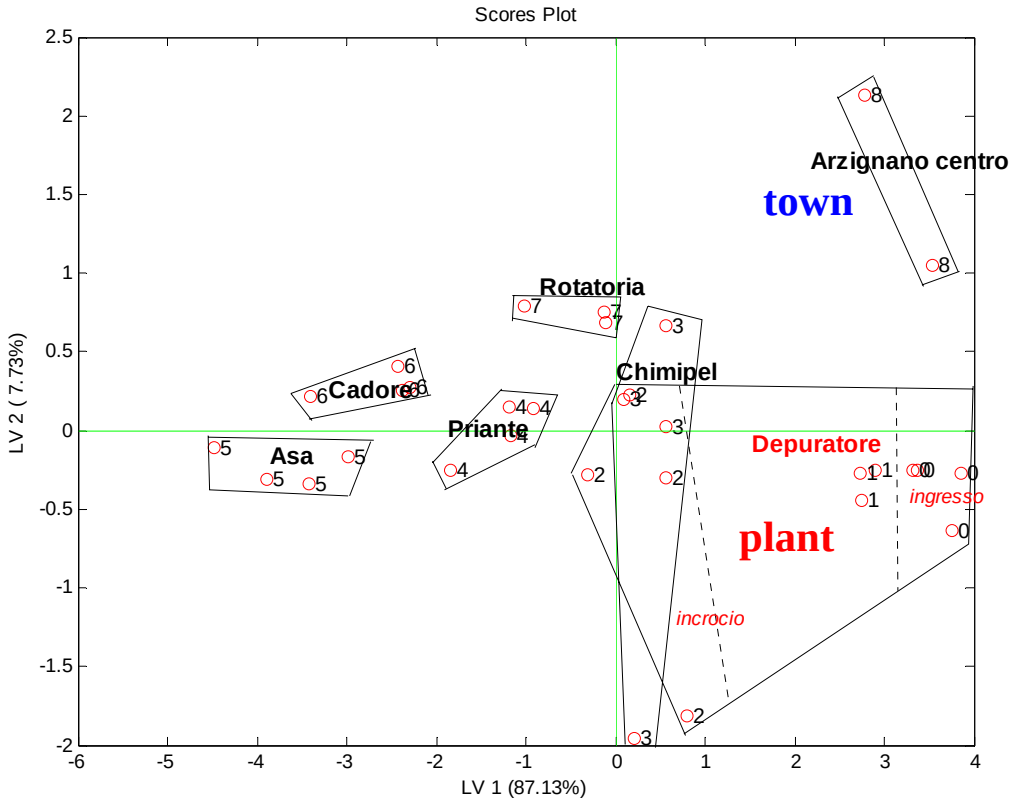


Fig. 5.3.13: Characterization of the different odour sources present in an industrial district; each point is labelled with the nearest industry to the sampling point.

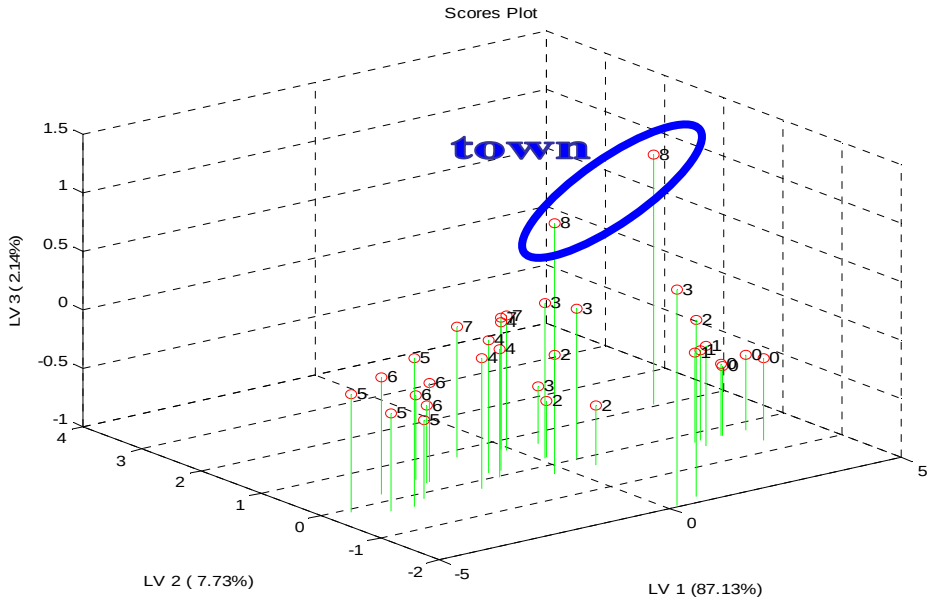


Fig. 4.3.14: scoreplot of the first three LV of the PLS-DA model elaborated aiming to distinguish the different odour sources inside the industrial district, the plant and the town.

In Figure 5.3.15 a color map of the olfactive intensities is represented. The map is calculated adding the mean Δf for the measurements collected for each of the nine selected sites. The results show the town and wastewater treatment plant are characterized by the lowest olfactive intensities. This result support a correct functioning of the plant. Nonetheless on the basis of collected data it is possible to confirm that odors from the plant were not present inside the town.

The different odours registered along the way inside the industrial district are probably due to different working phase taking place in the different tanneries positioned near the sampled site. Actually, each of these factories worked for the same final product performing different phases of the tanning process.

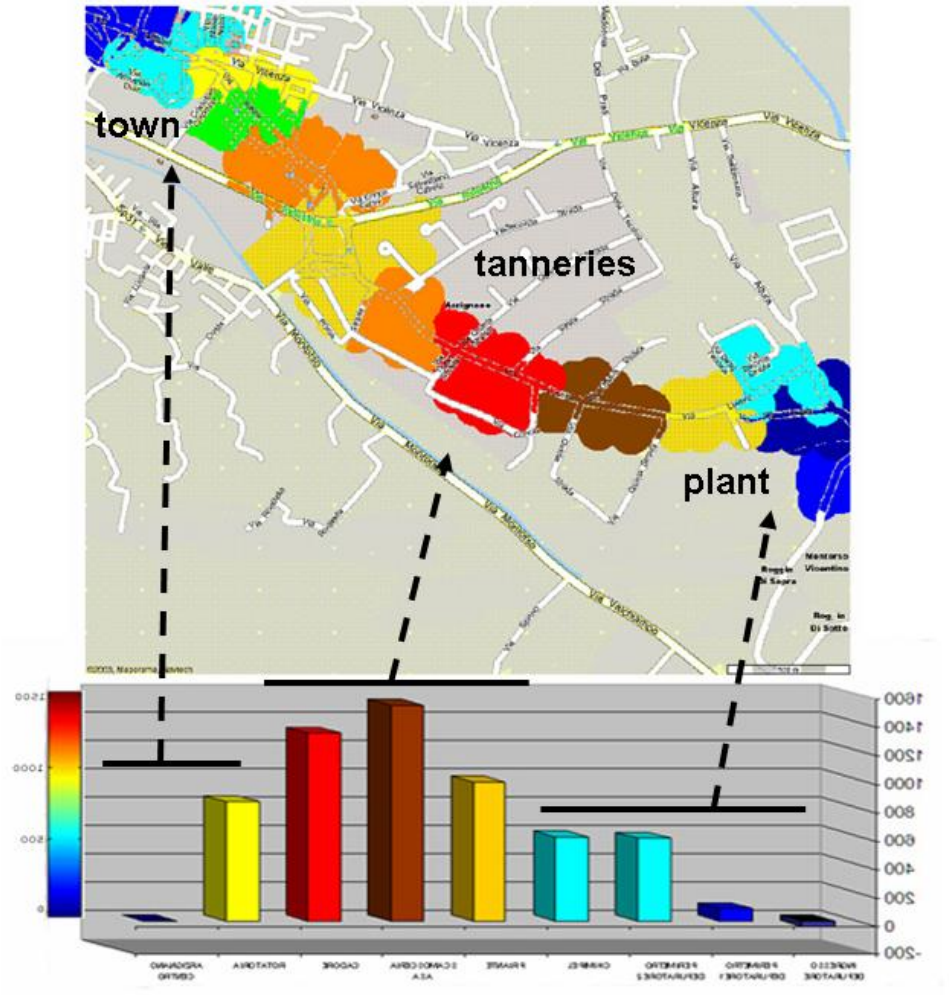


Fig. 4.3.15: Coloured map of the industrial district representing the olfactive impacts calculated for each measured point.

5.3.3 The rubbish dump case

The odour surrounding a large rubbish dump serving a big town was the last case investigated. The area under monitoring was very large because of the extension of the areas for tip collecting.

Moreover, the area where the rubbish dump is positioned is a zone with a high pollution risk, because of the contemporary presence of a refinery, a fuel deposit, an incinerator and a plant for dumping of hospital rubbish.

The e-nose was shown to be able to discriminate the characteristic odors of several different areas (indicated in figure 5.3.16): the collecting zone, the incinerator and the built-up area close to the rubbish dump (3km)).

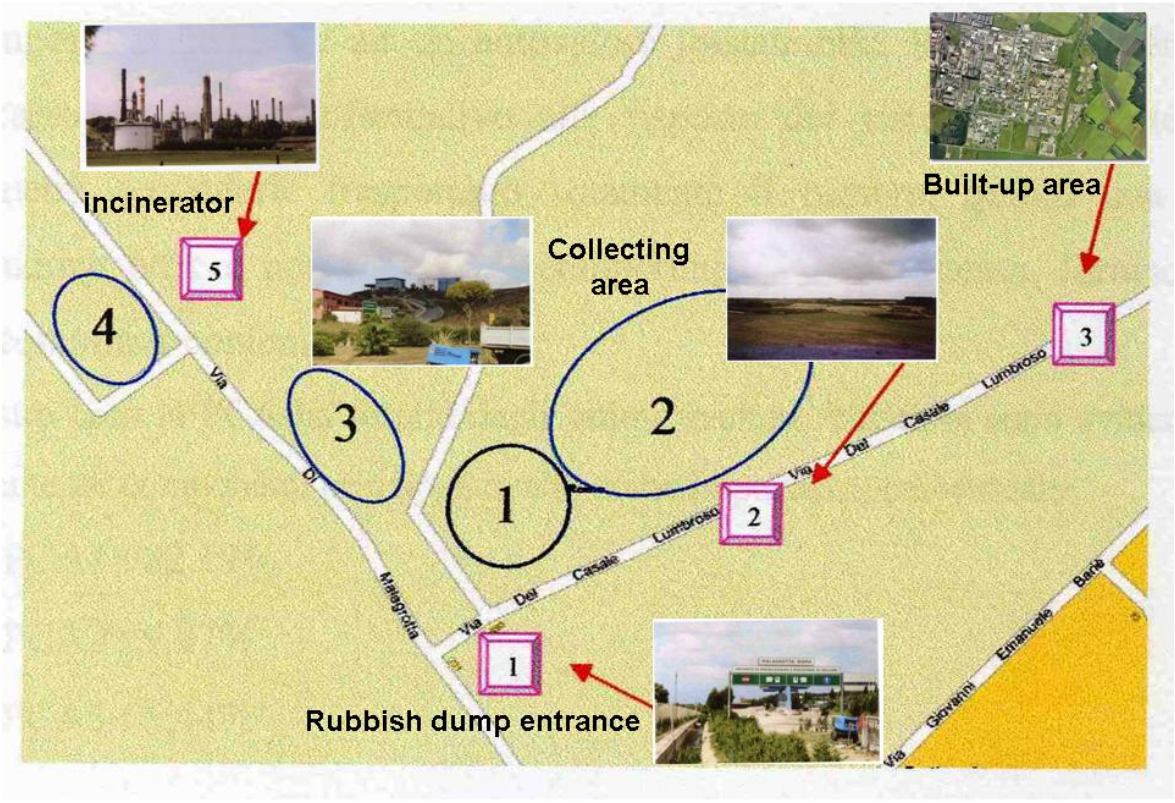


Fig.4.3.16: Map of the rubbish dump area. The sites pointed up are the places in which the odour has been sampled by means of the e-nose.

The ability of the e-nose in discriminating the different odour sources inside the area is shown in figure 5.3.17, where the scoreplot of the first two PCs of the

PCA model is plotted. In an in-set of the same figure a scoreplot including the third PC is also reported; the third dimension allows a better discrimination of the two sites of the built-up area and of the collecting zone.

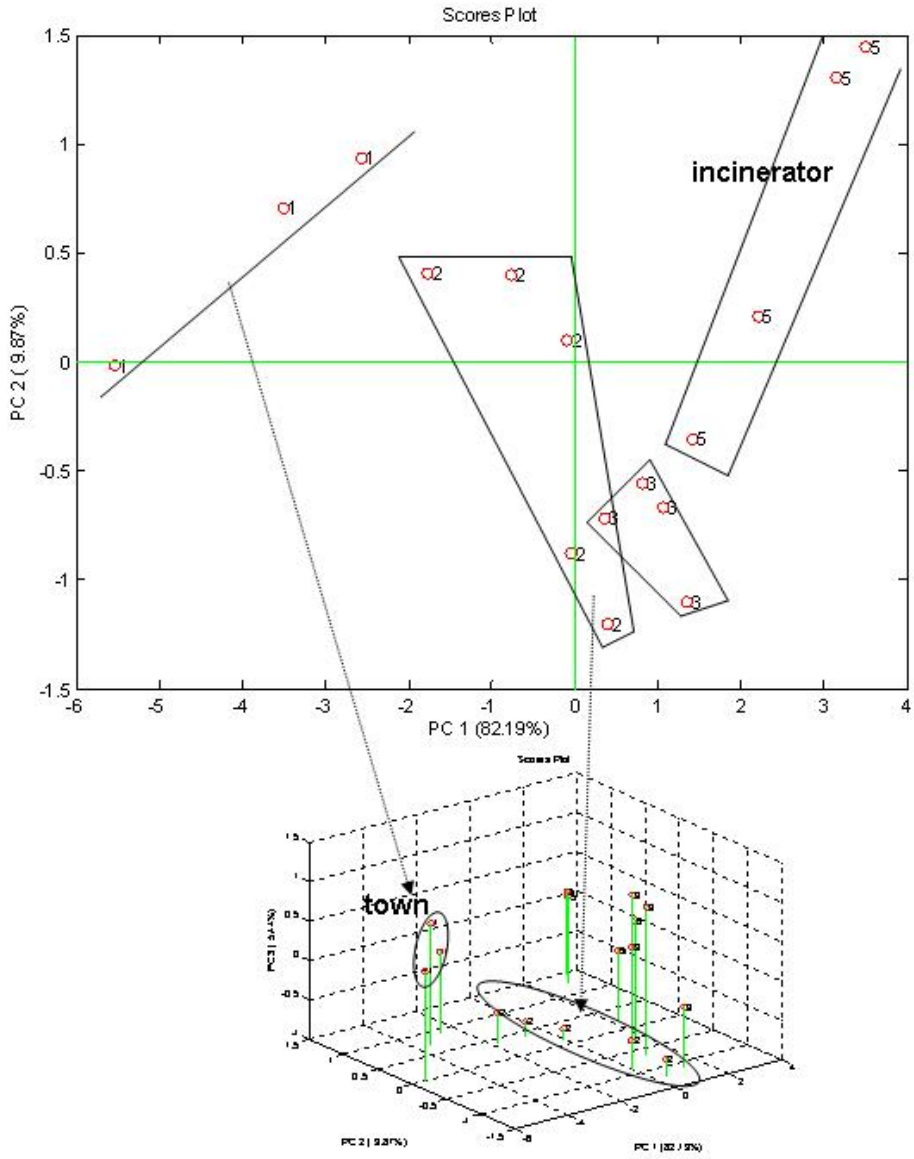


Fig. 5.3.17: scoreplots of the first two and of the first three Principal Components (PC) of the PCA model elaborated on the data collected inside the rubbish dump area.

Moreover, the measurments by GC-MS of air sampled in this area confirmed a common pattern of VOCs, with some particular concentration of some compounds characteristic for each of the three areas. From this experiment it is possible to conclude that e-nose, measuring a characteristic pattern of VOCs for

different points of sampling, is able to classify the samples as different, by detecting the differences in the concentration values of identical compounds.

The GC chromatograms are reported in figure 5.3.18.

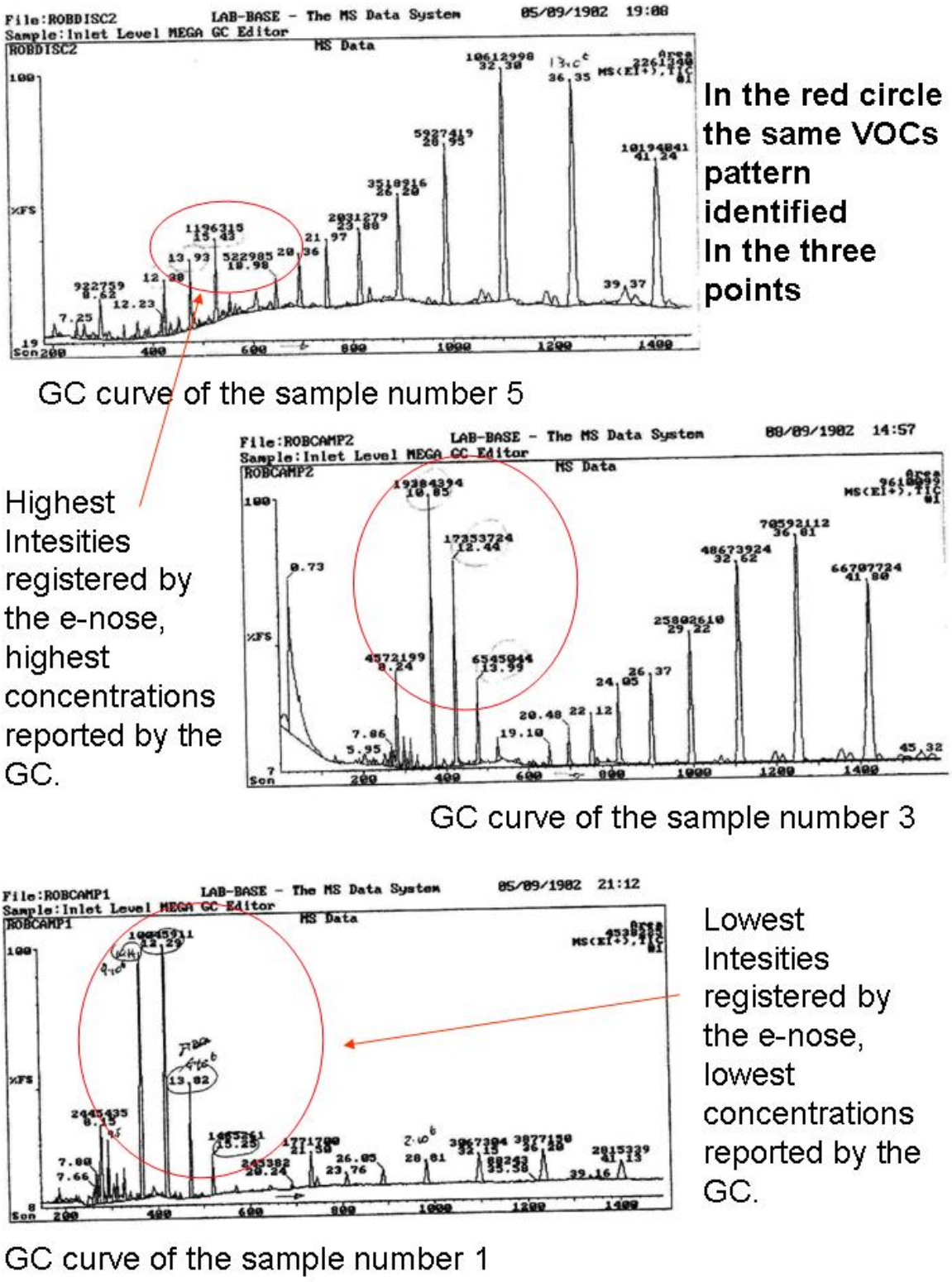


Fig. 5.3.18: GC curves relative to the air sampled in the different points in the rubbish dump area.

5.4 References

[1] W. Bourgeois, J. E. Burgess, R. M. Stuetz, 'On-line monitoring of wastewater quality: a review', *Journal of Chemical Technology an Biotechnology* 76: 337-348 (2001).

[2] R. M. Stuetz, R. A. Fenner, G. Engin, 'Characterisation of wastewater using an electronic nose', *Wat. Res.* Vol. 33, No. 2, pp. 442-452, 1999.

[3] T. Dewettinck, K. Van Hege, W. Verstraete, 'The electronic nose as a rapid sensor for volatile compounds in treated domestic wastewater', Wat. Res. Vol. 35, No. 10, pp. 2475–2483, 2001

[4] J. W. Gardner, Hyun Woo Shin, E. L. Hines, C. S. Dow, 'An electronic nose system for monitoring the quality of potable water', *Sensors and Actuators B* 69 (2000) 336–341.

[5] C. Di Natale, A. Macagnano, F. Davide, A. D'Amico, A. Legin, Y. Vlasov, A. Rudnitskaya, B. Selezenev, 'Multicomponents analysis on polluted waters by means of electronic tongue' *Sensors and Actuators B* 44 (1997) 423-428.

[6] J. Nicolas, A.C. Romain, V. Wiertz, J. Maternova, Ph. André, 'Using the classification model of an electronic nose to assign unknown malodours to environmental sources and to monitor them continuously' *Sensors and Actuators B* 69 (2000) 366–371.

[7] R. E. Baby, M. Cabeza, E. N. W. de Reca, ' Electronic nose: a useful tool for monitoring environmental contamination', *Sensors and Actuators B* 69 (3): 214-218 0ct 25 2000.

[8] W. Burgeois, R. M. Stuetz, 'Measuring wastewater quality using a sensor array: prospects for real time monitoring', *Water Sci. Technol*. 41 (12): 107-112 (2000).

[9] H. W. Shin, E. Llobet, J. W. Gardner, E. L. Hines, C. S. Dow, ' Classification of the strain and growth phase cyanobacteria in potable water using an electronic nose system', *IEE P-Sci. Meas Tech* 147 (4) : 158-164 (2000).

[10] E. M. Biey, W. Verstraete,' The use of a UV lamp for control of odour decomposition of kitchen and vegetables waste', *Environ Technol*. 20 (3): 331-335 (1999).

[11] P. E. Keller, R. T. Kouzes, L. J. Kangas, 'Three neural network based sensor systems for environmental monitoring', *IEEE Electro 94 Conference Proc.* Boston, MA, 377-382, (1994).

[12] C. Mouche, 'Electronic nose sniffs out, classifies contamination', *Pollut Eng* 31 (2): 31 (1999).

[13] S. S. Schiffman, E. A. Satterly-Miller, M. S. Suggs, B. G. Graham, 'The effect of environmental odors emanating from commercial swine operations on the mood of nearby residents', *Brain Res Bull* 37: 369-375 (1995).

[14] P. Gostelow, S. A. Parsons, R. M. Stuetz, 'Odour measurements for sewage treatment works', *Water Res*. 35 (3) (2001).

[15] R. M. Stuetz, G. Engin, R. A. Fenner, 'Sewage odour measurements using a sensory panel and an electronic nose', *Water Scs Technol* 38 (3): 331-335 (1998).

[16] F. W. Schreiber, K. Fitzner, 'Electronic nose investigation of the perceived air quality in indoor environments', *Indoor Air* 99, Vol. 2, 624-629, Edinburgh (1999).

[17] F. W. Schreiber, K. Fitzner, 'Investigation of the perceived air quality in an office building with an electronic nose', *Healthy Buildings* 2000, Helsinki (2000).

[18] C. Delpha, M. Siadat, M. Lumbreras, 'Discrimination of refrigerant gas in a humidity controlled atmosphere by using modelling parameters', *Sensors and Actuators B* 62 (3): 226-232 (2000).

[19] C. Delpha, M. Siadat, M. Lumbreras, 'An electronic nose for the identification of forane R134a in an air-conditioned atmosphere', *Sensors and Actuators B* 69 (3): 243-247 (2000).

[20] F. Sarry, M. Lumbreras, 'Gas discrimination in an air-conditioned system', *IEEE T Instrum MEAS* 49 (4): 809-812 (2000).

[21] O. Ramalho, 'Correspondences between olfactometry, analytical and electronic nose data for 10 indoor paints' *Analysis* 28 (3): 207-215 (2000).

[22] P. J. Hobbs, T. H. Misselbrook, M. S. Dhanoa, K. C. Persaud, 'Development of a relationship between olfactory response and major odorants from organic wastes', *J Sc Food Agr* 81 (2): 188-193 (2001).

[23] H. Willers, P. de Gijsel, N. Ogink, A. D'Amico, E. Martinelli, C. Di Natale, N. van Ras, J. Van der Waarde, 'Monitoring of biological odour filtration in closed environments with olfactometry and an electronic nose', *Water Science Technology* vol.50 No4 pp 93-100 (2004).

HEADSPACE STATIC SAMPLING

'Riesce in tutto chi si dà da fare mentre aspetta.'
T.A. Edison

'Who is able to work while waiting, can be able to do everything'

6

6.1 Static Headspace Sampling.

The term 'odour' indicates the biological, physical and psychological effects caused by the interaction between chemical stimulants – aromas and fragrances – and olfactive systems of living creatures [1]. In order to obtain the same effects with an artificial olfactory system it is necessary to preserve the natural conditions of odour delivery and measuring. Designing a suitable sampling procedure is mandatory to guarantee these conditions.

Procedures based on the manipulation of the headspace in contact with odorous materials are popular and very suitable for chemical analysis, especially for fragrances and aromas investigations. The different approaches normally used in GC-MS analysis or with chemical sensors such as electronic noses have been illustrated in the second chapter.

The simplest way to assess the chemical composition of an aroma is direct analysis (by GC or e-nose or other instruments) of a portion of the air in contact with the odour source, without any other sample-treatment-step.

The simplicity of the method, the opportunity to perform real on-line measurements, and the possibility to follow the dynamic of the sorption processes without any influences of a vapour sweeping procedure, make this system suitable for many different investigation objectives.

For example, in environmental monitoring, the scope is the investigation about the presence of some VOCs as contaminants in ground, water and soil. A static headspace method has been shown to be effective in analyzing compounds in contaminated water and soil samples [2]. The method is particularly useful in obtaining real-time data on-field using portable gas chromatographs.

About the quoted possibility to detail study different theoretical aspects of the chemistry of the mesurand or the interaction mechanisms between the headspace and the sensors performing the measure, numerous examples could be found in literature, exploiting a static headspace for determining equation constants or parameters, or to trace characteristic curves of the analyte under measurement.

A work presents a method for determining Henry's Law constants, applicable to the static headspace method, in which neither the exact concentration of the volatile compound, nor its matrix need to be known. Experimentally this method involves measuring by gas chromatography the equilibrium headspace peak areas of one or more compounds from aliquots of the same solution in three separate enclosed vials having different headspace-to-liquid volume ratios [3].

In other experiments this sampling method has shown to be useful, together with a full evaporation technique, to compensate for matrix effects in tracing small quantities of substances [4].

For these reasons an electronic nose based on static headspace sample has been developed: this argument revert the sampling bringing the sensors in contact with the sample. The main advantage of a static headspace sampling experimental set is the simplification of the whole apparatus, by the means of the elimination of the pneumatic system and of many of the most usual intermediate sampling tools.

This simplification results in an optimization of some problematic questions concerning the repeatability and the effectiveness of the measurements such as the influence of the sampling material and the geometry of the sensors chamber.

Moreover, in many applications, the static headspace sampling, guarantee a measure of the analyte in its proper natural environment, without conditioning or handling procedures.

The main opportunity of this method is the possibility to deeper study the binding interactions mechanisms taking place between the chemical interactive material and the mesurand, with the unique disadvantage of a longer measure time and a lowest intensity of the sensor response.

6.2 Libra Nose for static measurements: NoseStat

The design of a protocol for odour static sampling is a task which cannot be realized by the simple modification of tools and procedures for the collecting, the handling and the delivery of the sample. Actually, considering that the core of the study is to take the sensors in the sample headspace, the nose itself, and in particular the sensors chamber, became part of the protocol.

These indications suggest to design a dedicated e-nose for static measurements Satisfying the following requirements:

- ✓ Chamber has to act as measure chamber and sampler
- ✓ A mechanism to control the communication between the sample chamber and the sensors chamber is required
- ✓ The sampler chamber has to accomodate liquid, solid or gaseous samples.

The NoseStat (The static version of the Libra Nose) has been designed satisfying the objective above listed.

In order to reduce contamination of measurements it is necessary to use adequate materials repellent to odours. For these reasons Teflon, Derlin and Stainless steel have been used to construct NoseStat.

To satisfy the first two tasks, a sampling chamber of cylinder shape with a volume of 0.75 litres has been designed. It is the core of the whole device. The cylindric chamber is divided into two parts: the bottom, approximately 3/4 of the total volume, is the sampling chamber (made of stainless steel) while the upper part is the sensor chamber, and it is made of teflon. Figure 6.1.1 gives an overview of the whole instrument.

Mother board with the control and communication circuits, the oscillator board and the pneumatic system are the same of other e-nose versions.

Concerning the pneumatic system, there is no necessity of switching valve, because the flow driven by the pump is used only for the cleaning of the sensors chamber.

The two chambers are separated by a movable septum, controlled by the operator.

The measure chamber design allows for two possible sample insertion. The first is an airtight window, for the insertion of solid sample or of a vial containing a liquid. The presence of two tubes in two opposite sides of the measure chamber gives the opportunity to inlet controlled concentrations of VOCs, by the means of Mass Flow Controllers, or simply by means of syringe.

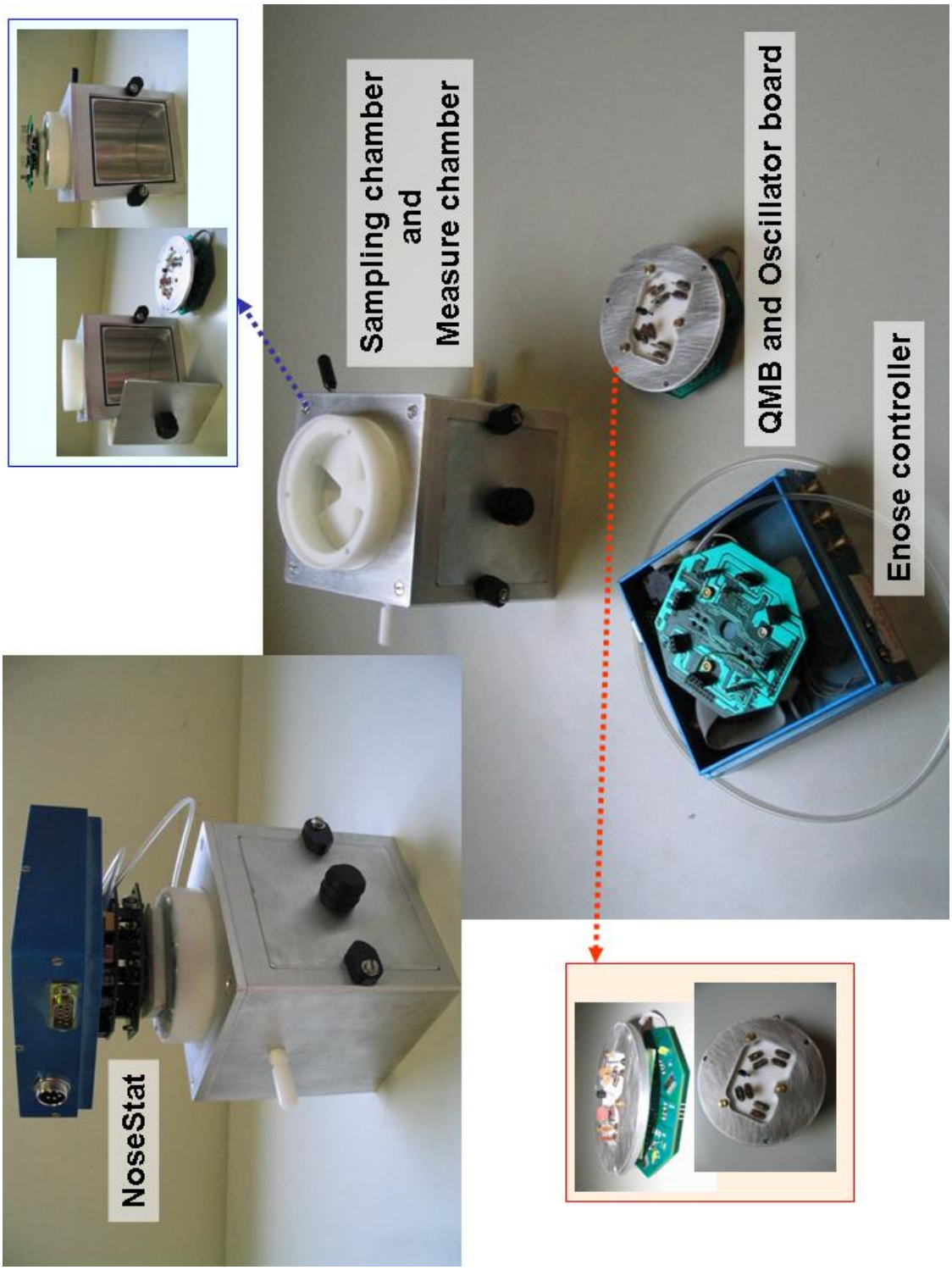


Figure 6.1.1: NoseStat dismounted in all its fundamental components.

6.3 Protocol definition

Protocol has been determined through a series of trials. Measurement sequence is outlined in figure 6.3.1.

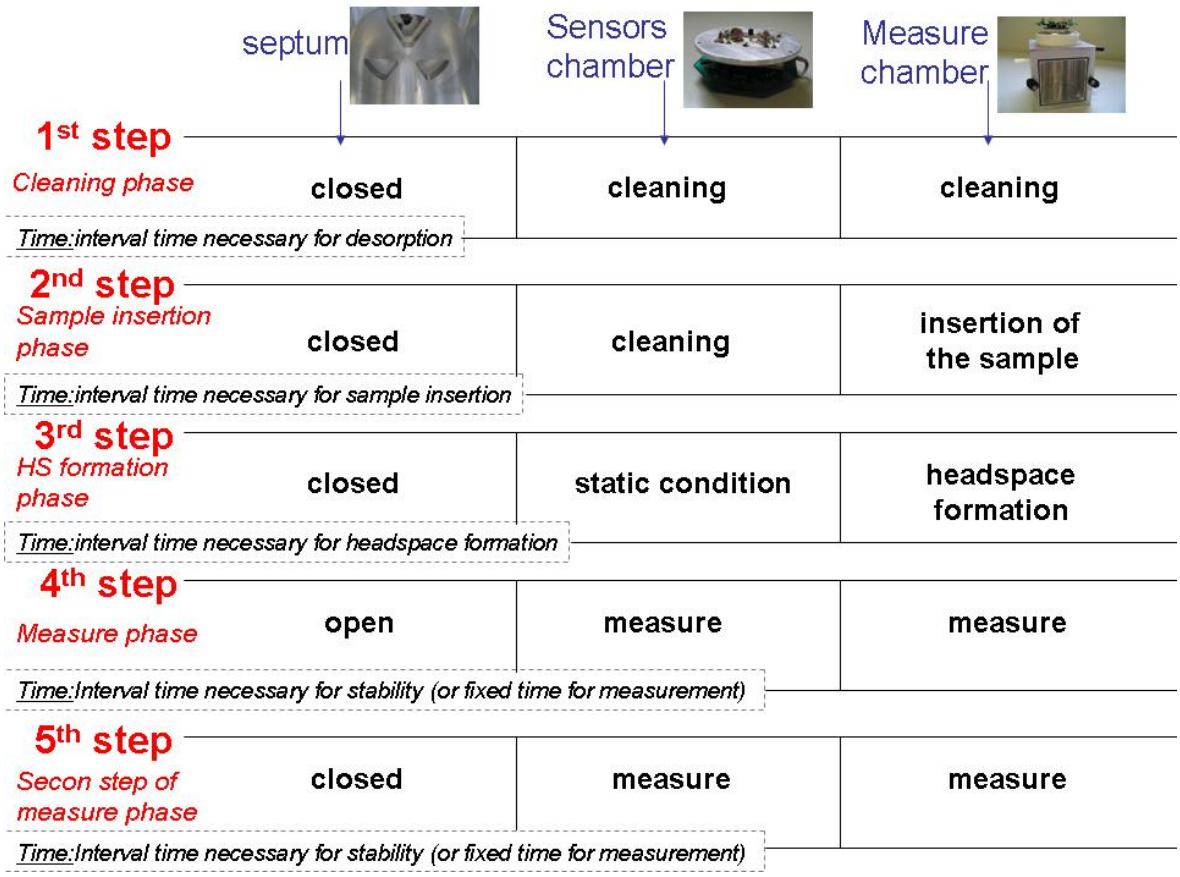


Fig. 6.3.1: Sequence of operations to perform a measurement with the NoseStat.

The two key points in the protocol are the time for the headspace formation, and the optional fifth step.

The amount of time left for the formation of the headspace is of fundamental importance. Usually a fixed time is predefined to obtain a good repeatability, in case of static condition it is necessary to consider the headspace situation before starting the measure.

Indeed, for a repeatable and optimal static measure, the measure chamber, used as sampler, has to be completely saturated. If saturation is not achieved time and sensitivity of the measurement result affected, as can be observed in figure 6.3.2.

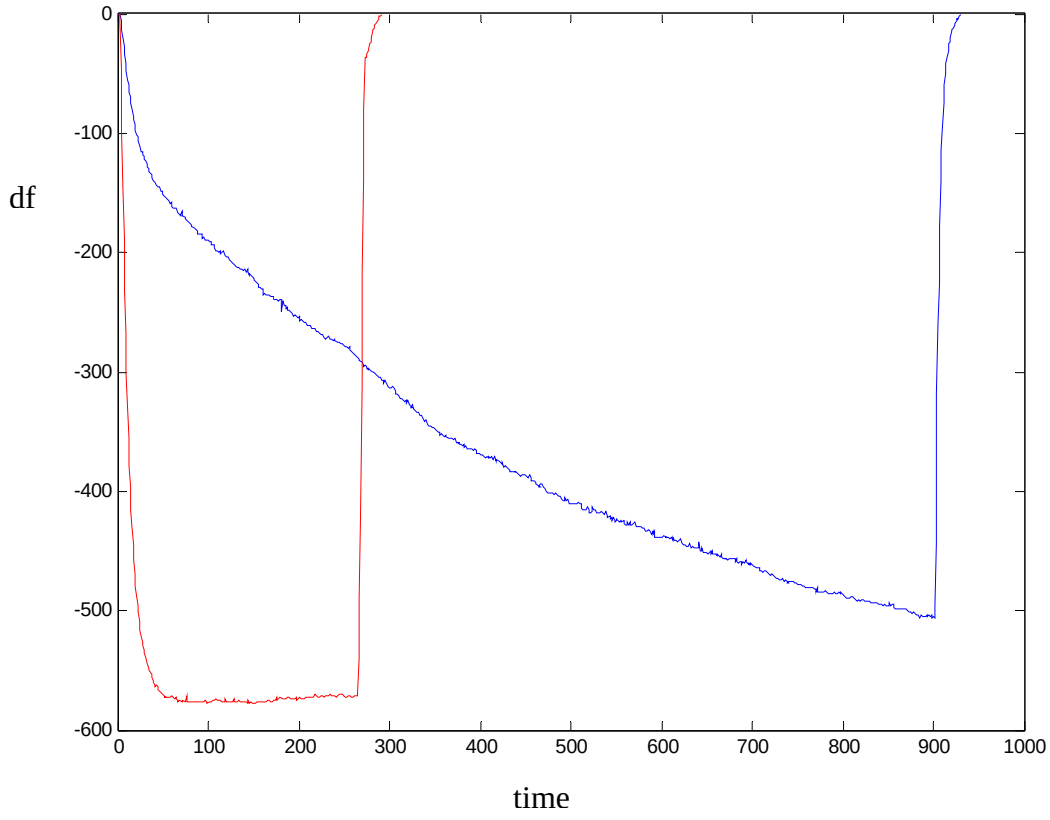


Fig. 6.3.2: sensor responses (one of the eight in the array) to ethanol; the blue curve is relative to the response to a headspace formed in 15 min., the red curve is relative to the response to a headspace formed in 30 min.

In this figure two responses of the same sensor to the same sample (ethanol) are reported, relative to two protocol differing in headspace formation time.

The protocol consist in a 15 minutes of HS formation time in case of the blue curve, and 30 minutes in case of the red. The differences in the measure time and in the response intensities are evident. The response to a 30 minutes headspace formation is more than four times faster than the response relative to a half time headspace formation (15 min.). Moreover, the frequency shift is of 100 hertz higher in the red curve (30 min. HS form.) than in the blue one (15 min. HS form.).

The second fundamental choice in the protocol definition is concerned with the opportunity of the fifth step. The reason of this additional step is the reduction of the volume of the measurement chamber and the elimination of the variability of the headspace during the measurements. It is obtained excluding the sample

from the sensor chamber immediately after the headspace expansion over the two chambers.

Many trials have been performed to test the validity of this step. This variable does not affect both response time and response magnitude, as shown in figure 6.3.3.

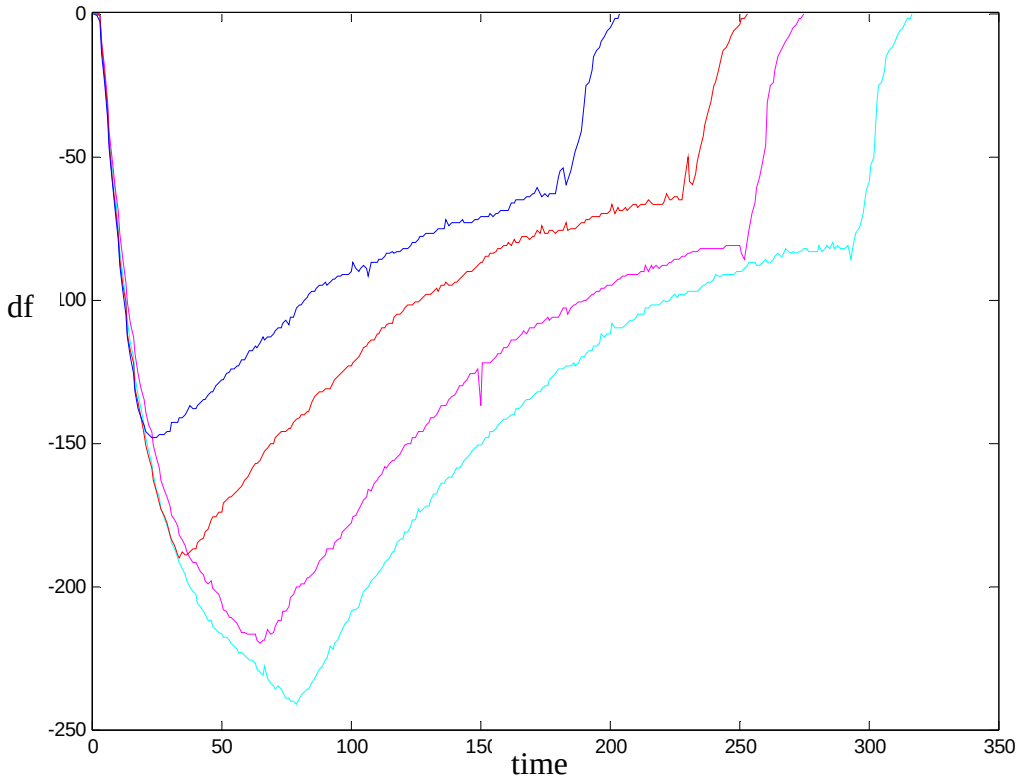


Fig. 6.3.3: the five response curves of one of the 8 sensors of the array to the same compound, ethanol, but with different closing time of the communication septum (from the highest (blue) to the lowest (cyan): 5 min., 7 min., 12 min., 15 min.).

In term of response intensity, of course, this method is less useful than the tested techniques in which the closure of the septum occurs at the stability. The reason is in the fact that the different curves in figure give the impression of a not completed dynamic response.

Although the intensity of the response is fundamental and strategical to satisfy discrimination tasks, the method suggested in the fifth optional step has shown an interesting potentiality. During the measure of different compounds, some differences in the response curves have emerged. These differences have been observed in the second part of the curve, when a new dynamic process of

equilibrium achieving takes place in the sensors chamber. Probably, the contemporary actions of adsorption and desorption processes gives a particular behaviour due to the interaction mechanism between the CIMs and the analytes, which is characteristic for each mesurand.

These considerations open the way to a set of features to be studied for the characterization of different substances by the means of the e-nose. Besides the usually evaluation of the frequency shift at the response steady state (which is a static feature) a series of dynamic features [5-9] can be investigated to discriminate different VOCs. The most evident of these has been the simple shape of the curves.

Ethanol and Isopropilic alcohol have been tested to support this conjecture.

To better represent the peculiarity of each of the two substances respect to the other, a PCA model has been elaborated on the basis of the whole set of point registered during the measurement. The result is a representation of the response curves of the two analytes in the scoreplot (fig. 6.3.4) of the first three principal components of the elaborated PCA model.

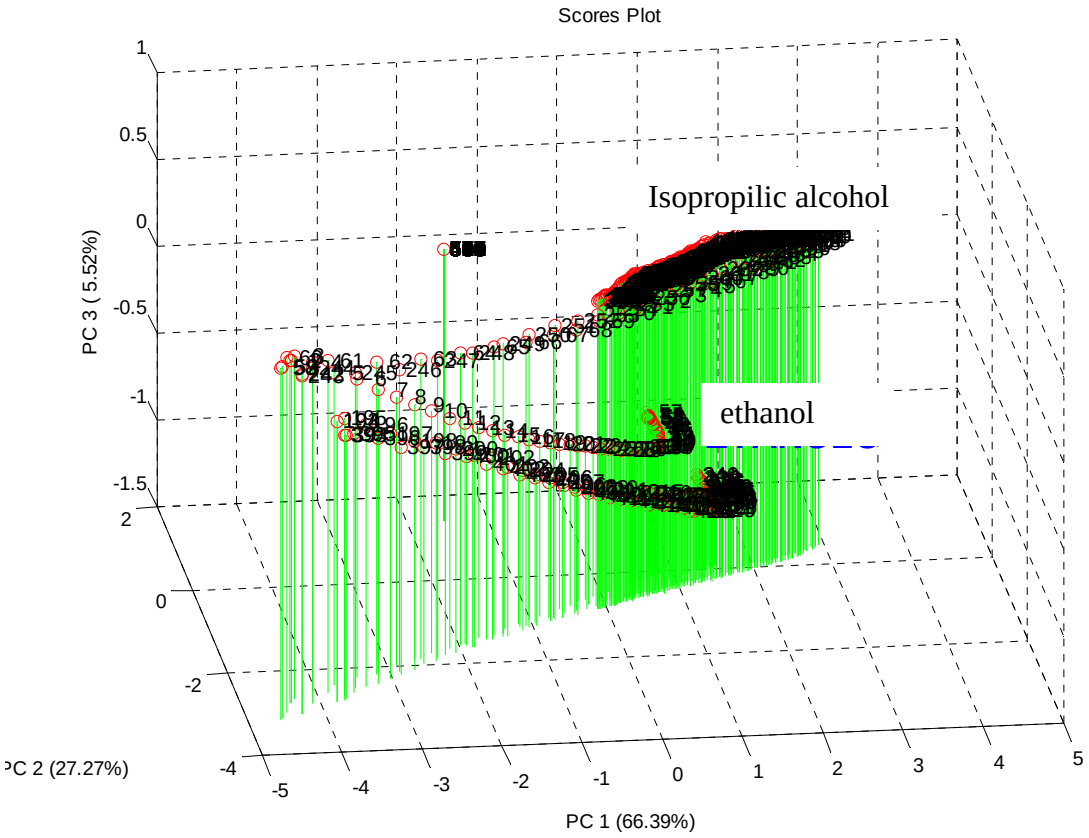


Fig. 6.3.4: scoreplot of the first three PC of the PCA model elaborated on the whole data set of the static measurements of ethanol and isopropilic alcohol.

Two important things emerge from this figure. First the ethanol achieves equilibrium in a shorter time. Moreover, shape, minimum value corresponding to the turning point, and the value corresponding to the final steady state contribute to the discrimination between the two compounds.

6.4 Results and conclusions

This experimentation can be divided into two steps: the first consisted in design and the realization of an apparatus to perform odour measurements in static conditions. The built instrument can be optimized by the automation of some of the operations taking place in the measure protocol. Anyway, it was effective for the preliminary trials performed until now.

The second step concerns the testing of a methodology, to evaluate potentialities and applications. This test phase requires for many measurements of many different compounds, so it is still a work in progress.

On the basis of the measurements performed, a series of conclusions can be drawn:

- The technique showed a good efficiency in the discrimination of different substances.
- Static sampling protocol allows to study the binding interaction mechanism and to characterize the different compounds under measurement.
- Static approach presents some advantage and some disadvantage if compared to dynamic method.

The ability of this method to discriminate different classes has been tested in a set of measurements of ethanol and isobutilic alcohol at different concentration (from 200 to 1500 ppm).

In figure 6.4.1 the scoreplot of the first two principal components of the PCA model calculated on the data set shows a good discrimination between the two groups.

The superposition of the two clusters for low concentrations is an indication about the resolution of this apparatus.

The position of two measured samples relative to mixtures of the two compounds, and reported in an intermediate region between the two classes, suggests the ability to identify the contemporary presence of different volatiles as a complex odor representing the global effect of their interaction.

In the scoreplot a direction can be identified along which the concentration increases.

The protocol used in this experiment is the one mentioned before, without the fifth step and with a time of 30 minutes for the headspace formation.

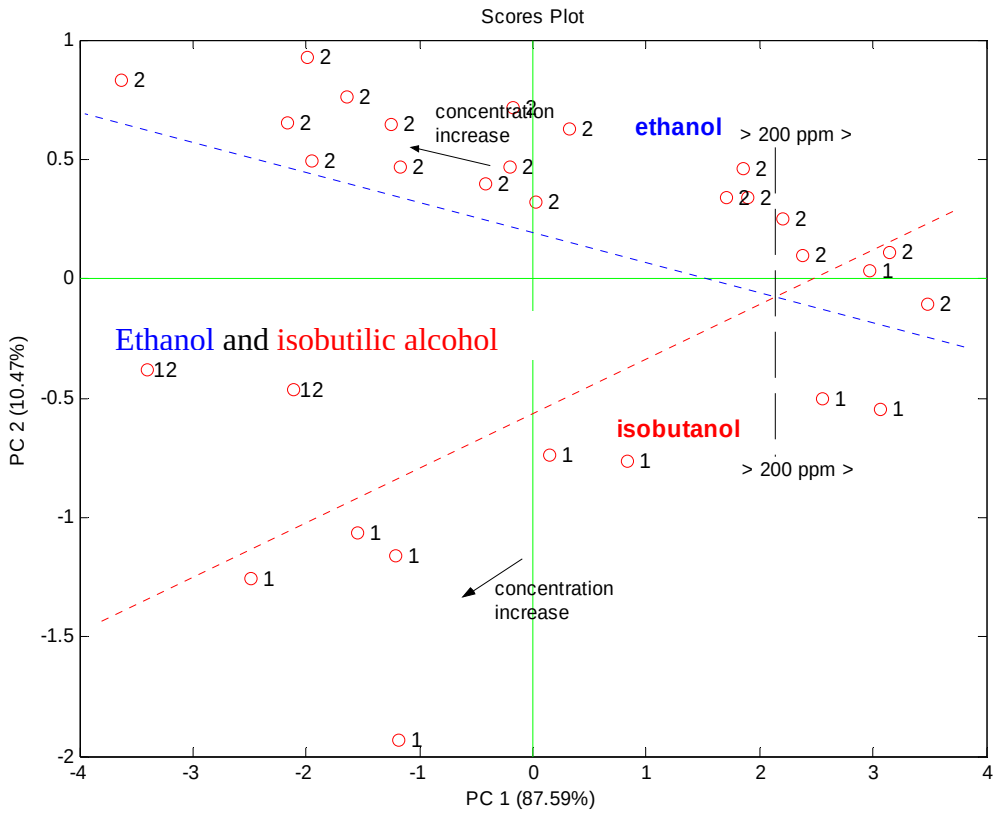


Fig. 6.4.1: scoreplot of the first two PCs of the PCA model elaborated on the data set relative to the measurements of different concentrations of ethanol and isobutilic alcohol.

Beside the mentioned opportunity to study the dynamic of the sensor response by the analysis of the curves during the fifth step of the protocol, there is also the possibility to characterize the measured compounds, because of the fixed concentrations measured.

An example is reported in figure 6.4.2, where the fitting calculated on the frequency shift results in an isotherm of the measured compound (in this case ethanol).

To conclude this dissertation the classic dynamic measurement technique and the static one has been compared. Main features are listed in table 6.4.1.

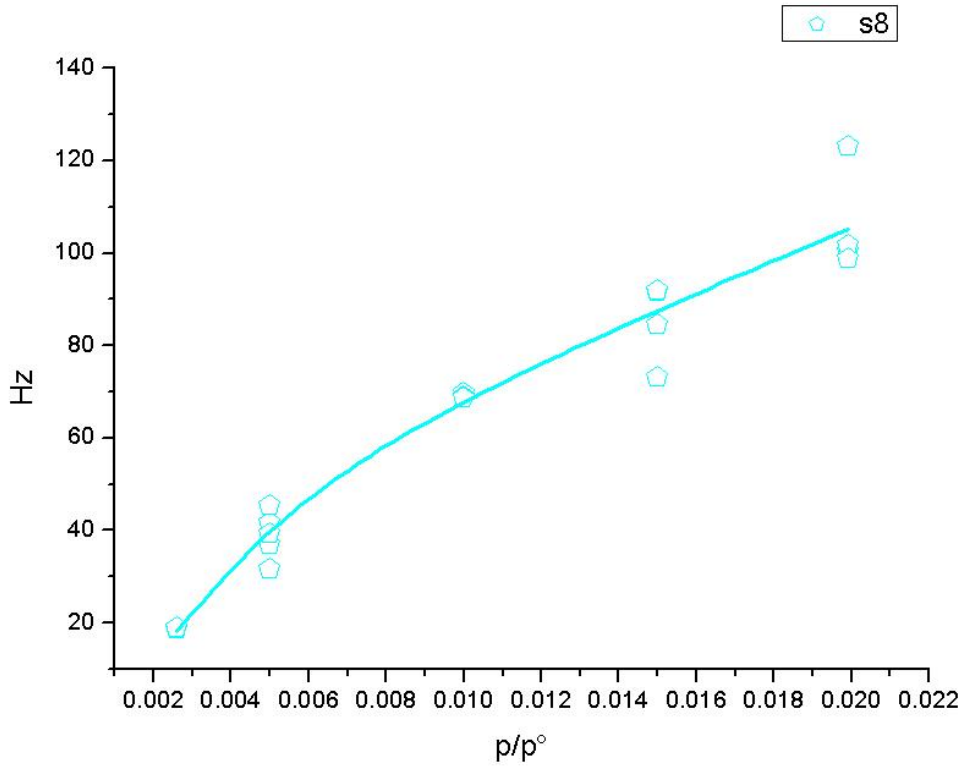


Fig. 5.4.2: ethanol isotherm elaborated on the data set obtained by the means of a static sampling protocol.

	STATIC	DYNAMIC
REPEATIBILITY	BEST	>>
FREQUENCY SHIFT	<<	BEST
MEASURE TIME	BEST	>>
CLEANING TIME	>>	BEST
DF FEATURE CLASSIFICATION	=	=
CURVE SHAPE FEATURE CLASSIFICATION	YES	NO

Table 6.4.1: Comparison between static and dynamic measure protocol. To perform this comparison the same measurements has been performed contemporary with the two different methods.

6.5 References

[1] F. Augusto, A. Leite e Lopes, C. Alcaraz Zini, 'Sampling and sample preparation for analysis of aromas and fragrances', *Trends in Analytical Chemistry*, Vol. 22, No.3, 2003.

[2] V. D. Roe, M. J. Lacy, J. D. Stuart, G. A. Robbins, Anal. Chem. 1989, 61, 2584-2585.

[3] G. A. Robbins, S. Wang, J. D. Stuart, 'Using the Static Headspace Method To Determine Henry's Law Constants', *Anal. Chem.* 1993, 3113-3118.

[4] S. Strassnig, E. P. Lankmayr, 'Elimination of matrix effects for static headspace analysis of ethanol', *Journal of Cromatography A*, 849 (1999) 629-636.

[5] T. Eklöv, P. Mårtensson, I. Lundström; *Analytica Chimica Acta* (1997) 353 pp. 291-300

[6] S. Roussel, G. Forsberg, P. Grenier, V. Bellon-Maurel, 'Optimisation of electronic nose measurements. Part I: Methodology of output feature selection', *Journal of Food Engineering 37,* 207-222.

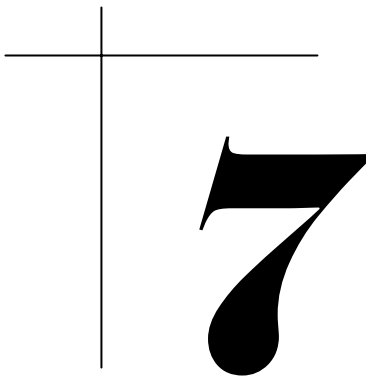
[7] E. Martinelli, G. Pennazza, C. Di Natale, A. D'Amico, 'Chemical sensors clustering with the dynamic moments approach', *Sensors and Actuators B* 101 (2004) 346-352.

[8] R. Guiterrez-Osuma, H.T. Nagle, S.S. Schiffmann, 'Transient response analysis of an electronic nose using multi-exponential models', *Sens. Actuators B* 61 (1999) 61 170–182.

[9] E. Llobet, J. Brezmes, X. Vilanova, J. E. Sueiras, X. Correig, ' Qualitative and quantitative analysis of volatile organic compounds using transient and steady-state responses of a thick film tin oxide gas sensor array', *Sensors and Actuators B* 41 (1997) 13-21.

Appendix

Multivariate Data Analysis



Appendix

7.1 Multivariate Data Analysis and electronic nose

Artificial olfaction systems are attempts to reproduce the functions of the natural olfaction sense. Currently, a sort of standard electronic nose model common to the available implementations of artificial olfaction may be identified. In this model three major sub-systems may be identified: sampling, sensor array, and data analysis. While sensor arrays are oriented to exploit the analogies with natural receptors (e.g. cross-selectivity and redundancy) the data analysis does not take advantage of natural data treatment paradigms. The principles of data processing of natural olfaction receptors signals are still rather unknown and the differences in terms of sensors signals make them hardly applicable to process sensors data, although the topic is evolving and first attempts to derive analysis methodologies from natural paradigms have already been proposed [1].

Electronic nose data analysis is rather based on techniques borrowed from other disciplines and chosen among that vast class of algorithms known as pattern recognition.

Pattern recognition in chemistry appeared at the end of the sixties. Among the first applications it has to be mentioned the use of the Learning Machine of behalf of the analytical chemists at the Washington State University. They classified molecules according to molecular structural categories used to form patterns [2].

In gas sensing area, multivariate methods of pattern recognition are commonly required when sensor arrays, composed of real, non-selective, cross-sensitive sensors, are utilised. Pattern recognition, exploiting the cross-correlation, extracts information contained in the sensor outputs ensemble.

Formally, pattern recognition may be defined as "the mapping of a pattern from a given pattern space into a class-membership function" [3].

A pattern can be defined as an ordered set of real numbers. In analytical chemistry, the response of a multi-channel instrument to one sample (e.g. gas chromatographs and spectrophotometers) forms a pattern. For a sensor array, the responses of the sensors to a sample takes the name of sensors pattern.

Class-membership space represents either quantitative or qualitative set of quantities. When quantification is considered the class-membership function turns out to be a vectorial function defined in a metric space. In the case of quality recognition, the class-membership is an ensemble of abstract sets at which each measured sample is assumed to belong. In the first case we prefer to talk of multicomponent analysis while the term pattern recognition is usually referred to the second case.

Mathematically, a pattern is a vector belonging to its own vector space. The vector space is the one where the sensors responses are represented in the sensor space. The simplest of these representations considers one scalar response for each sensor. This response is called sensor feature and feature extraction is the operation aimed at extracting from a sensor response a number of synthetic descriptors to be used to form patterns.

When a sensor (independently from its working principle) is exposed to a gas or a mixture of gases, it gives a response depending on the nature of the gas and its concentration. Almost the totality of the examples reported in literature considered, as the sensor response either the absolute or the relative change in sensor signal measured in two steady-state conditions: in absence and in presence of the gas and when all the transients are ended. This definition leaves not exploited the dynamics behaviour of the sensor signal. Several authors studied the optimisation of feature extraction in order to maximise the array performance, among them Eklöv et al [4] showed a method to optimise the information extraction from an array of MOSFET sensors. In their paper, the dynamical behaviour of sensor signals has been represented by seventeen different features, partially correlated one each other. The optimisation were carried out maximising the performances of an array of MOSFET sensors in the quantitative estimation of mixtures of hydrogen and ethanol. Although the evident restriction of the studied case this paper shows that quantification error is minimised when descriptors considering the dynamical behaviour are taken into account. On the other hand, it has to be mentioned that the dynamical behaviour is very sensitive to the fluctuations of sample delivery system.

Data Pre-processing: Scaling and Normalisation

An usual procedure in pattern recognition is the scaling of the data. Instead of using raw data, two main scaling procedures are widely used: zero-centred and autoscaling.

Zero-centred data means that each sensor is shifted across the zero value, so that the mean of the responses is zero. Zero-centred scaling may be important when the assumption of a known statistic distribution of the data is used. Supposing a normal distribution, zero-centred data are completely described only by the covariance matrix, this is the case of Principal Component Analysis.

Autoscaling means to scale each sensor to zero-mean and unitary-variance. This operation equalises the dynamic of the sensor responses avoiding that a sensor, with a wider dynamic, may obscure the contribution of other sensors dynamically limited. Further, autoscaling makes the sensor responses dimension-less, this feature is necessary when sensors whose signals are expressed in different units are joined. This is the case of hybrid arrays (different sensor technologies in the same array) and when electronic noses are fused with other instruments, e.g. the fusion of electronic noses and electronic tongues [5]

Normalisation

Horner and Hierold [6] have showed that the application of a simple normalisation of sensors data can greatly help in removing the quantitative information and putting in evidence the qualitative aspects of the data.

The cross-selectivity of the sensors make their responses ambiguous, so that different samples, due to a combination of qualitative and quantitative aspects may give rise to similar sensor response.

A simple way to disentangle the information can be applied if the relationship between sensor response and concentration of a given compound (in this example referred to an array of TSMR sensors) is linear, such as:

$$\Delta f_i = K_{ij} \cdot c_j \qquad (7.1.1)$$

where Δf_i is the response of the i-th TSMR, K_{ij} is the sensitivity of the i-th sensor towards the j-th compounds, and c_j is the concentration of the j-th compound.
The normalisation consists in dividing each sensor response by the sum of all the sensor responses to the same sample, so that the concentration information disappears.

$$\Delta f_i \Rightarrow \frac{K_{ij} \cdot c_j}{\sum_m K_{mj} \cdot c_j} = \frac{K_{ij}}{\sum_m K_{mj}} \qquad (7.1.2)$$

The application of linear normalisation to an array of linear sensors should produce, on the PCA score plot, one point for each compound, independent of its concentration, and achieving the highest possible recognition. Deviations from ideal behaviour are due to the presence of measurement errors, and more important, to the non-linear relationship between sensor response and concentration. The presence of specific interactions gives rise to the dispersion of the data related to the same compound. This dispersion could also be interpreted as a measure of the specificity of the interaction, but paying attention to the fact that the non-linearity only occurs in a range of concentrations behind which the relations become linear.

In the quoted paper, Horner and Hierold, treated also the common case of power law, valid for instance for metal-oxide semiconductor gas sensors. Eq.7.1.2 can be extended to sensors described by a power-law ($z = c^{\alpha}$) simply linearising, through the logarithm, the sensor response.

Normalisation is, in principle, useful to counteract possible fluctuations in the sample concentration. These fluctuations are due to temperature fluctuations of the sample, and to instabilities of the sampling systems that may bring to variations of the dilution factor of the sample with the carrier gas. Of course, normalisation is of limited help due to the fact that the previous assumptions hold for simple samples and they are faded to fail when mixtures of compounds are measured. Furthermore, fluctuations in temperature do not induce only a general concentration shift, but since compounds have different boiling temperatures,

each component of a mixture changes differently so that both quantitative (concentration shift) and qualitative (pattern distortion) variations occur.

7.2 Explorative and Regression methods.

Explorative methods allow to 'explore' the sensors responses set in a particular application, looking for relationship between these data set and possible existing structure or classification.

Regression methods, instead, are devoted to establish correlations between an input and an output data set.

An important distinction between these two techniques consists in the unsupervised character of the explorative methods. This means that explorative methods use only the sensors data set in the elaboration of the model; anyway information about the real environment in which measurements have been performed are necessary in a second step to interpret data structures emerged in the elaborated model.

In the regression techniques, instead, knowledge of the ambiental condition of the measurements are considered in the elaboration of the model.

7.2.1 Principal Component Analysis [7,8]

Principal Component Analysis (PCA) is a method for decomposition of multivariate data in uncorrelated components. In principal component analysis, the scope is to maximize the variance of a linear combination of the variables.

The main applications of PCA techniques are: to reduce the number of variables and to detect structure in the relationships between variables, that is to classify variables. Therefore, PCA is applied as a data reduction or structure detection method

In some applications, the principal components are an end in themselves and may be amenable to interpretation. More often they are obtained for use as input to another analysis.

To define the principal components it is necessary to consider the statistic of the observable space. Given a multivariate measure data set, consisting in the rows

of an array *X*, they can be represented as arrays in a space of diemsion equal to the number of columns of *X*.

PCA technique is based on the hypothesis that the multivariate data set has a Normal distribution, and the calculation method of the Principal components is based on this hypotesis.

To define a multivariate Probability Distribution Function (PDF) a generalization of the monodimensional statistical descriptors is necessary.

Two descriptors are used to define a Gaussian PDF: mean and variance. The multivariate mean is the mean array of the data set. In the case of variance the generalization is more complex. Indeed variance becomes the covariance matrix in the multivariate case. This matrix is Defined as the expected value of the quadratic form of x-m, where x is the data set and m is the mean value. This is, formally, a definition equivalent to the monodimensional variance, in which the fact that x is an array transform the square in a quadratic form.
The complete expression is:

$$cov(X) = \Sigma = E\left[(x-m)^T \cdot (x-m)\right] \qquad (7.2.1)$$

Covariance matrix is a square and simmetric matrix. This matrix defines the covariance degree of each single variable, so the elements of the matrix consist in a measure of how the variables 'vary together' (Co-variate).
The elements of the diagonal are the variances of the single variables, while the others are proportional to the correlation coefficient.

$$\Sigma_{ii} = \sigma_i^2 \quad ; \quad \Sigma_{ik} = \rho_{ik}\sigma_i\sigma_k \qquad (7.2.2)$$

The multivariate PDF expression is:

$$PDF(x) = \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{\Sigma}} exp\left[-\frac{1}{2}(x-m)^T \Sigma^{-1}(x-m)\right] \qquad (7.2.3)$$

Figure 7.2.1 represent the 2-dimensional case of multivariate Gaussian PDF

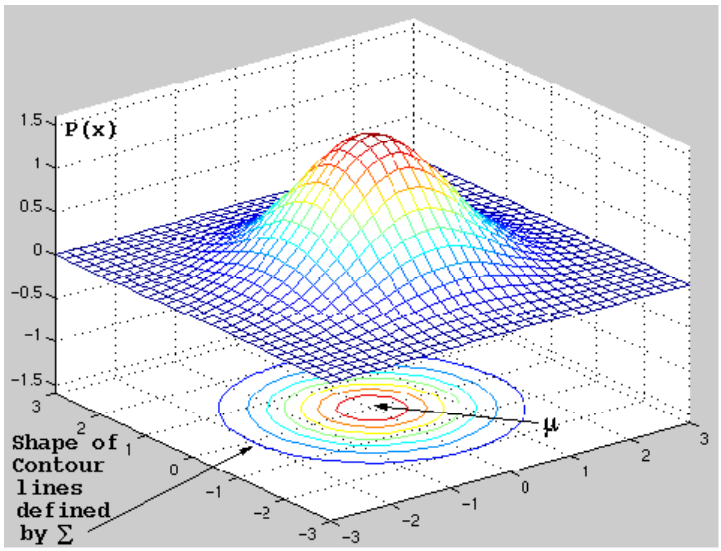


Fig. 7.2.1: 2-dimensional case of multivariate Gaussian PDF

Defining Y as the variables array and X as the observable array, it is possibile to estimate Y by the mean of the inversion (pseudoinversion) of the X array. It is possibile to perform this inversion if the array X has a maximum rank, it means that the number of linearly independent columns is the same of the number of columns of this array. This means that all the observables are lianearly indipendent.

If a partial dependance exists the numeric inversion could involve large calculation errors. This effect is named 'co-linearity', because X columns tend to co-variate. This colinearity is quantitative expressed by the covariance matrix. Indeed the non-diagonal terms of this matrix are different from zero in case of co-linearity.

To perform the inversion of the matrix it is necessary to remove the co-linearity, and this consists in transforming the covariance matrix in its diagonal form, by the mean of the introduction of new latent variables. This transformation is possibile by a suitable change of the reference system. This operation consists in the writing of the covariance matrix in its canonical form.

Before the calculation of these new latent variables it is useful to present some examples about covariance matrix. In order to graphically represent these matrixes a 2-dimension case is considered. Figure 7.2.2 shows the shapes of the isoprobability ellipses and the relative covariance matrixes.

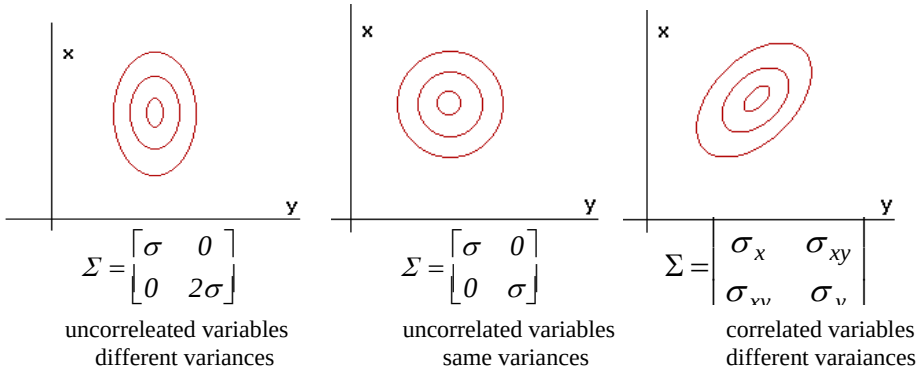


Figure 7.2.2: Shapes of the isoprobability ellipses and the relative covariance matrixes

The reference system which allows to put the ellipse in a canonical form consists of the eigenvectors of the covariance matrix, that means the principal axes of the ellipse represented as quadratic form of the covariance matrix itself. The new variables correspond to the first case reported in figure 7.2.2 and the PDF is obtained as the product of monovariate PDFs.

The variables of the new reference system are not physical observable, subject of the performed measuements, but they are a linear combination of the original data.

These new variables are called Principal Components, and the sequence of operations, calculus and interpretation is called Principal Component Analysis (PCA).

An example of the calculation for the coordinates in the 2-diemnsional case is reported below:

$$a \cdot x^2 + 2b \cdot xy + c \cdot y^2 = \begin{bmatrix} x & y \end{bmatrix} \cdot \begin{bmatrix} a & b \\ b & c \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix}$$
$$\Rightarrow \lambda_1 \cdot u^2 + \lambda_2 \cdot w^2 = \begin{bmatrix} u & w \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 & 0 \\ o & \lambda_2 \end{bmatrix} \cdot \begin{bmatrix} u \\ w \end{bmatrix} \qquad (7.2.4)$$

This is the mathematical context of the PCA; while its scope is the representation of a data set with a non-diagonal covariance matrix and of dimension N, in a new space with a reduced number of lower dimensions respect with N and a diagonal covariance matrix.

This diagonalization is performed by the means of the coordinates rotation in the basis of the eigenvectors (PCs). An eigenvalue is associated to each eigenvector. The eigenvalue describes the 'elongation' of the ellipse along the direction of the associated PC, and this elongation is relative to the variance of this PC.

It is worth to remark that if the original variables are partially correlated some of the eigenvalue have a neglectable value. In this case the eigenvectors assiociated to these eigenvalues can be neglected reducing the dimension of the final representation.Only in case of Normal PDF the PCs are indipendent, while in the opposite case the PCs are only uncorrelated.

Given a matrix X the PCs are the eigenvectors (**P**) of the covariance matrix which can be expressed as: X^TX. The coefficients of the eigenvectors in the observables original basis are named loadings.

The X matrix Data are transformed in the eigenvectors basis by the means of a projection in this new reference space. These coordinates are named scores and are obtained as: T=XP.

The X array results in a decomposition operated by the product of the **P** (loadings) and the coordinates of the original patterns in the new basis of the **T** (scores): $X=TP^T$.

7.2.2 Partial Least Square [7,8]

PCA is a solution to the Multiple Linear Regression problem. Indeed the calculation of the principal components eliminates the co-linearity of data allowing a correct calculation of the regression matrix.

Figure 7.2.3 shows the procedure of the regression method named PCR (Principal Component Regression).

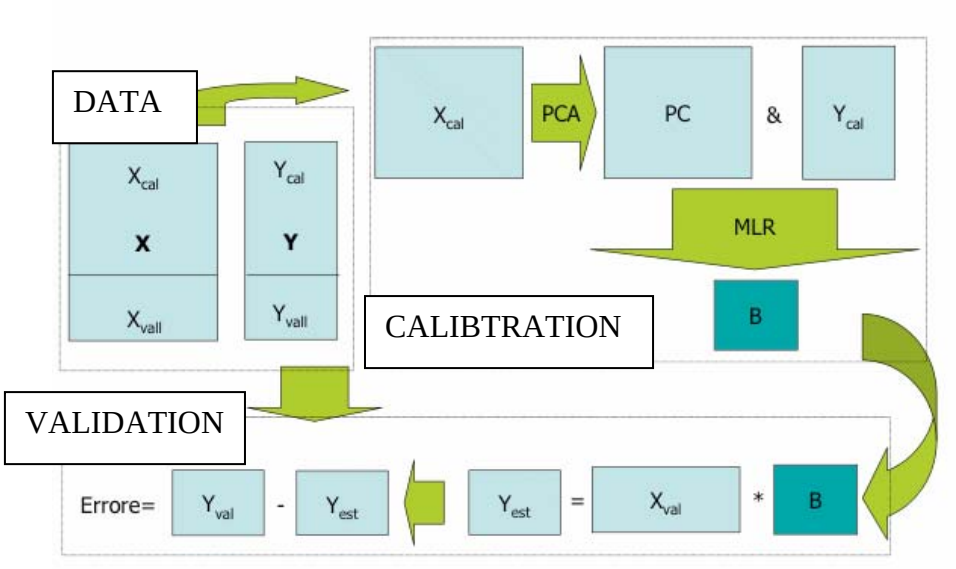


Fig. 7.2.3: PCR scheme

The PCR is not the optimal solution of the regression problem. Indeed in the calculation of the PCs the tipology and a possibile classification of the variables are not considered.

A step over consists in the application of the PCA calculation to the matrix X and to the matrix Y at the same time. In this way the PCs of Y are rotated to maximaze the correlation with the PCs of X and this method is named **Partial Least Squares** (PLS). These rotated components are different from the original PCs and are named Latent Variables.

The PCs are the direction of maximum variance of the data, but these direction are non-dependent on the scope (indicated in the matrix Y) cause they are univocally determined by the matrix X.

PLS algorithm is a data analysis technique which combines the PCA characteristics with the linear regression. This method is useful when there is the need to predict a set of dependent variables on the basis of a high number of indipendent variables.

PLS calculates the set of components of X, named latent variables, which represent a decomposition of the matrix, and at the same time maximaze their correlation with dependent variables of Y. The formula is reported below

$$X = T \cdot P^T + E \quad con \quad P \cdot P^T = I \qquad (7.2.5)$$

where I is the identity matrix , T is the scores matrix, P the Loadings matrix and E the residuals matrix.
In the same way Y can be represented as:

$$Y = T \cdot C^T + F \quad con \quad C \cdot C^T = I \qquad (7.2.6)$$

where C is the loadings matrix and F the residual matrix.
Using the 7.2.5 and the 7.2.6 it is possibile to write:

$$Y = X \cdot W \cdot C^T + F = X \cdot B + F \quad con \quad B = W \cdot C^T \qquad (7.2.7)$$

where W is the defined as:

$$T = X \cdot W \qquad (7.2.8)$$

PLS scope is to find the matrixes T and C which allows to maximaze the covariance between X and Y.

In PLS performing an overfitting problem can occur, so the number of latent variables to use must be optimized by the mean of a cross-validation technique.

Overfitting is originated by the fact that given the k latent variable the k+1 variable is obtained by the fitting of the subspace of dimension k+1; considering that the latent variables are not orthogonal, it exists the possibility to percfectly fit calibration data with a number of latent variables equal to N.

The goodness of the obtained model can be evaluated by the calculus of the Root Mean Square Error Calibration:RMSEC during the calibration phase, and by e the Root Mean Square Error Cross Validation :RMSECV during the validation step.

RMSEC decreases with the increasing of the number of LVs, while in the validation the phenomena is the same for the first LVs but it starts to increase after it has reached the minimum error.

The RMSEC monotone curve is due to the fact that with the increasing of the LVs the system begins to modellize also the noise included in the data set. Instead, the minimum of the RMSECV curves represents the trade-off between the ability of the model to generalize and its performances. So the number of the LVs to use in the model building is that correspondent to the RMSECV minimum.

7.3 References

[1] T. Pearce; *Biosystems*, 1997 41 pp. 43-67.

[2] B.R. Kowalski and S. Wold; in *Handbook of statistics Vol. 2*, P.R. Krishnaiah and L.N. Kanal eds., North Holland Publ. (Amsterdam, The Netherlands), (1982) pp. 673 – 697.

[3] C. Di Natale, F. Davide, A. D'Amico; *Sensors and Actuators B* (1995) 23 pp 111.

[4] T. Eklöv, P. Mårtensson, I. Lundström; *Analytica Chimica Acta* **1997** 353 pp. 291-300.

[5] C. Di Natale, R. Paolesse, A. Macagnano, A. Mantini, A. D'Amico, A. Legin, L. Lvova, A. Rudnitskaya, Y. Vlasov; *Sensors and Actuators B* (2000) 64 pp. 15-21.

[6] G. Horner, C. Hierold; *Sensors and Actuators B*, (1990) 2 p.p. 173-184.

[7] C. Di Natale, Lectures Notes of Sensors and Detectors at the University of Rome 'Tor Vergata'.

[8] A. C. Rencher, 'Methods of Multivariate Analysis', Wiley-Interscience (2002).

List of Publications

Papers related to the thesis subject

1) C. Di Natale, E. Martinelli, G. Pennazza, A. Macagnano, R. Paolesse, A. D'Amico, 'Artificial olfaction systems in medicine and telemedicine', IATICE 2002 Melbourne (Australia) Mar 2002.
2) Giorgio Pennazza, Antonella Macagnano, Eugenio Martinelli, Roberto Paolesse, Corrado Di Natale, Arnaldo D'Amico, 'Classification of complex mixtures with an electronic nose: the case of pharmaceutical products', AISEM 2003 Trento Feb 2003.
3) C. Di Natale, R. Paolesse, A. Macagnano, E. Martinelli, G. Pennazza, A. D'Amico, 'Optimization of molecular recognition in thickness shear mode resonator chemical sensors', Electrochemical Society Meeting Paris, France Apr, 2003.
4) G. Pennazza, E. Martinelli, R. Paolesse, C. Di Natale, A. D'Amico, E. Scaglia, B. Vialetto, J. Adzet, M. Jorba, N. Pritsos, 'Monitoring tannery processes using an array of porphyrins based QMB chemical sensors', 10th International Symposium Olfaction and Electronic Nose (ISOEN 03) Riga (Latvia) 25-27 Jun – 2003.
5) C. Di Natale, E. Martinelli, G. Pennazza, A. Macagnano, R. Paolesse, A. D'Amico, H.C. Willers, P. de Gijsel, N. Ogink, 'Air quality monitoring of closed environments using an electronic nose', 10th International Symposium Olfaction and Electronic Nose (ISOEN 03) Riga (Latvia) 25-27 Jun – 2003.
6) G. Pennazza, A. Valenti, C. Di Natale, R. Paolesse, A. Macagnano, E. Martinelli, A. D'Amico, 'Monitoring of Environmental Odours by an Electronic Nose: Waste Water Treatment Plant and Rubbish Dump, Three Case Study', AISEM 2004 Ferrara Feb 2004.
7) A. Macagnano, E. Sgreccia, C. Di Natale, R. Paolesse, S. Pantalei, E. Zampetti, G. Pennazza, A.D'Amico, 'Studying the binding processes

between porphyrins and organic vapours', IMCS 2004, Tsukuba (Japan) July 2004.

8) G. Pennazza, A. Valenti, C. Di Natale, R. Paolesse, A. Macagnano, E. Martinelli, A. D'Amico, 'Monitoring of Environmental Odours by an Electronic Nose: Waste Water Treatment Plant and Rubbish Dump, Three Case Study', IMCS 2004, Tsukuba (Japan) July 2004.
9) E. Martinelli, G. Pennazza, C. Di Natale, R. Paolesse, E. Milian, J. Albiol, F. Godia, A. D'Amico, 'Online monitoring of indoor air quality by an electronic nose', IEEE Sensors Conference 2004 Vienna (Austria) 24-27 Oct 2004.

Papers related to other scientific activities

Peer reviewed journals

1) C. Di Natale, G. Pennazza, A. Macagnano, E. Martinelli, R. Paolesse, A. D'Amico, 'Thickness shear mode resonator sensors for the detection of androstenone in pork fat', *Sensors and Actuators B* 91 (2003) 169-174 .
2) E. Martinelli, G. Pennazza, C. Di Natale, A. D'Amico, 'Chemical sensors clustering with the dynamic moments approach', *Sensors and Actuators B* 101 (2004) 346-352.
3) C. Di Natale, R. Paolesse, M. Bugio, E. Martinelli, G. Pennazza, A. D'Amico, 'Application of metalloporphyrins based gas and liquid sensor arrays to the analysis of red wine', Analitica Chimica Acta 513 (2004) 49-56.

International conferences

1) C. Di Natale, E. Martinelli, G. Pennazza, A. Macagnano, M. Noferino, G. Costa, A. D'Amico, 'Fusion of NIR spectroscopy and electronic nose data in apples and kiwi fruits quality estimation', AISEM 2002 Bologna Feb 2002.

2) C. Di Natale, E. Martinelli, G. Pennazza, A. D'Amico, M. Noferino, G. Costa, 'Stima della qualità di mele e kiwi tramite la data fusion di spettroscopia NIR e naso elettronico', Giornate Società Ortofrutticola Italiana Spoleto Apr, 2002.
3) C. Di Natale, A. Macagnano, G. Pennazza, E. Martinelli, R. Paolesse, A. D'Amico, F. Mencarelli, D. De Santis, 'Application of an electronic nose to the investigation of the effects of 1-MCP and propylene treatments of peaches', 9th International Symposium Olfaction and Electronic Nose (ISOEN 02) Rome 29 Sep - 2 Oct 2002.
4) E. Martinelli, G. Pennazza, C.Falconi, C.Di Natale, A.D'Amico, 'Comparison between two novel feature extracion methods for chemical sensors array', AISEM 2003 Trento Feb 2003.
5) E. Martinelli, G. Pennazza, C. Falconi, A. D'Amico, C. Di Natale, 'Sensor feature extraction by morphological descriptors of trajectories in phase space', Eurosensors XVI Guimaraes (Portugal) Sep 21-24, 2003.