UNIVERSITÀ DEGLI STUDI DI PADOVA

Facoltà di Ingegneria
Dipartimento di Tecnica e Gestione dei sistemi industriali

Tesi di Laurea Magistrale

CORSO DI LAUREA IN INGEGNERIA GESTIONALE

QUESTIONNAIRE-BASED STATISTICAL SURVEYS: METHODOLOGICAL ASPECTS AND SOME CASE STUDIES

Relatore: Ch.mo Dott. Livio Corain

Correlatore: Ch.mo Prof. Luigi Salmaso

Laureando: Walter Palazzin

ANNO ACCADEMICO 2011-2012

TABLE OF CONTENTS

SUMMARY

Via an interdisciplinary approach, this study aimed to collect data on *qualitative* features of rather diverse case studies, and subsequently to perform analyses with *quantitative* statistical methods. A sound research should, first and foremost, provide us with deeper knowledge on the "universe" under scrutiny, and consequently change our way of viewing it. This change brings about new objectives and new researches, ultimately setting off a virtuous cycle that expands our competence.

The *longitudinal* approach clearly emerges at the end of the previous paragraph: this is the second leitmotiv in this work. Most of us accept today that longitudinal information is necessary, and it seems especially the case for causal studies on individual behavior. This acceptance rests on the understanding that longitudinal studies can show the nature of growth, trace patterns of change, and possibly give a true picture of cause and effect over time.

The topics discussed are:

1. QUESTIONNAIRE METHODOLOGY: Questionnaires are the tool commonly used in three different cases; each questionnaire is specifically tailored for the various circumstances through the use of methods and techniques.
2. CARRYING CAPACITY: a whole chapter is devoted to a field of research that tries to answer a simple question: how much use can ultimately be accommodated in national parks and related areas? Environmental and social aspects are brought up to discuss the issue.
3. TWO CASE STUDIES are presented after the discussion on Carrying Capacity because they concern environmental/naturalistic experiences: a survey on the Dolomites district of Sesto, and a study on didactics in several ski schools in Alto Adige.
4. A STUDY ON MALODOUR: after a brief literature review on malodour, a case study is presented where a questionnaire helps understand the impact of bad smells from different sources on the population of two towns in the province of Padua. Along with the panel of "sniffers" (people trained to detect odours), a special device is used, called "electronic nose".

INTRODUCTION

Questionnaires are used every day for different purposes and in a wide range of applications, spanning from clinical research to customer satisfaction surveys, and beyond. They embody one means by which planning agencies may "hear" the constructive views of the public. One should bear in mind that a questionnaire is not just a list of questions, but a scientific instrument for measurement and for correlation of particular kinds of data; therefore it has to be specially designed. Issues involved with questionnaire design include: what are the primary goals of collecting data from individuals? Who are the target respondents? What should you consider a representative sample? Which method should be used to reach the respondents? Also you should develop the question wording carefully, which provides a typical example of pros and cons presented to the questionnaire designer when you limit your choice to open-ended or closed-ended questions:

Open-ended questions	**Closed-ended questions**
Elicit 'rich' qualitative data	Elicit quantitative data
Encourage thought and freedom of expression	Provide an easy way of indicating an answer, without need for articulation
May discourage response from less literate respondents	Should be easy for all literacy levels to respond to
Take longer to answer and may put some people off, to the extent of preventing completion	Are easy to answer and may improve your response rate

Generally, statistical methods are applied once the data has been collected, in order to extract information which more often than not surprises researchers themselves, and may spawn further studies.

The cases we examined, involving human behaviors and opinions, seemed particularly suitable to be analyzed with a longitudinal approach. Most generally, longitudinal studies collect data about the same subjects relating to multiple time points. Subjects may be individual people or other entities, e.g. organizations such as firms. Longitudinal research provides an understanding of social change, of the trajectories of individual life histories and of the dynamic processes that underlie social and economic

life, not possible from research based on cross-sectional data. We can identify a range of issues and types of research where longitudinal approaches seem especially appropriate, in particular they are essential when phenomena of interest directly concerned with individuals change over time.

After an extensive chapter on questionnaire methodology, we moved on to discuss an emerging environmental issue in the last decades, Carrying Capacity. A whole, extensive chapter is dedicated to this topic, which has numerous facets, and helps introduce two studies related to naturalistic/environmental issues.
Finally, a statistical sensory and device-based survey on odour perceptions is presented in the final chapter, to evaluate the impact of odorous sources located in the territory of two towns, Este and Ospedaletto Euganeo. Using a panel of "sniffers", that is a group of people trained to detect odours, and a sensory device called "electronic nose", data on olfactory perceptions is collected for a whole year, with several objectives in mind. The main ones are: measure the temporal evolution of the phenomenon, depending on atmospheric and climatic changes; provide a map of the perceptions, taking into account the area involved and the seasonal period; "quantify" the subjective and objective impact of the perceived problem, differentiating the annoyance, particularly by manifestation area. All this is carried out with statistical analyses, trying to compare measurements from the panel of judges and the device.

CHAPTER 1. Questionnaire Design and Techniques

Finding out what people think and know, how they live and behave, is interesting and, sometimes, crucial to take decisions in many situations. Even if it is often possible to directly observe behaviors and individual characteristics – the house or car they possess, how they spend their spare time, etc. – the best way to collect such information is almost always by asking questions and recording the answers. Therefore many researches aim at identifying, listing or explaining individuals' characteristics using questionnaires or interviews. They are also used as tools in evaluating personal characteristics, such as personality traits and attitudes (typically in job interviews).
This chapter intends to examine problems and issues relevant in those contexts where, more or less directly, questions are asked and answers are recorded. It is paramount to develop methodological abilities and critical skills, not only to evaluate and interpret results of someone else's research, but mainly to learn what to do and not to do, in order to collect quality data, and then extract valuable results.

1.1 Survey General Approach

The discussion on questionnaire techniques concerns many disciplinary, theoretical and methodological aspects, which constitute very broad research fields. Here we will focus on topics relevant in some types of research, while barely touching specific elements (e.g. measuring attitudes, research methods in social sciences, theories and techniques behind the implementation of a test).

1.1.1 Types of Research

Researches based on interviews or questionnaires can be characterized according to different aspects, some of which are pointed out:

- the *purposes* of a research, , which range from simply describing a phenomenon in a group of individuals, to identifying and explaining the reasons underneath;
- the type of *population* examined and the *sampling* method used to select a sample in that population (this phase is crucial to make sure that the individuals selected for the sample are representative of a certain population);
- the *type of objects*, phenomena or variables examined (opinions, behaviors, etc.), and the way they are examined;

- (social, cultural, and geographical) *degree of generalization* over time and space of the phenomena examined, that is how much these are due to causes invariant over time, or in relation to the culture where the individuals live, or in relation to their personal story and characteristics;
- The *degree of standardization* of the whole process of data collection. This aspect concerns not only the questions asked and their answers, but the whole research process (presenting the questionnaire to subjects, encoding methods, and analyzing data).

1.1.2 Questions, Answers and their Objects

A research is carried out to examine the properties of a specific matter, event or phenomenon, an objective which we pursue by asking questions. The answers given by the individuals contacted provide the data which will be analyzed to describe the phenomena and possibly draw conclusions.

To analyze the answers from a quantitative point of view (as well as qualitative), verbal data is to be converted into numeric data, which takes place in a phase called *data encoding*. The value assigned to an answer can reflect its value (for example, with age or weight), or arbitrarily identify a specific option (for example, "0" if you do not watch TV, "1" if you watch it occasionally, "2" if you watch it every day).

Every question has an "object", that is it concerns a certain topic or event that the researcher is interested in. It is possible to describe a systematic classification of the principal objects, that is which are the categories of subjects or topics involved.

In social sciences, objects investigated with questions are often categorized as *objective* or *subjective*. Typical examples of objective data or events are the weight of a person, his age, race, nationality, income, the number of times the individual reads newspapers in a year. Subjective data are his opinions, his emotions, his intentions (where shall I go on holiday?), etc.

Some kinds of objects investigated with questions are:

- Attitudes, values, inclinations, preferences: how favorable (or unfavorable) an individual feels towards an object, such as a political party, immigrants, his job, health issues, etc. ;
- Beliefs, opinions, perceptions, expectations, prejudices: what the individual believes to be true or false, right or wrong, probable or improbable, including the credibility of a politician, the reliability of an appliance, etc. ;

- Behavioral intentions: how the individual thinks he will behave towards a certain object, such as voting for a party, choosing a faculty, employing new workers, dealing with a foreign neighbor;
- Emotions, sensations, moods: for example, how often the individual feels depressed, happy, elated, stressed; if he feels lonely, satisfied with his job, in good/bad terms with his colleagues;
- Information regarding non personal facts or events (also called *knowledge questions*): who is the president of a certain country, how many foreign unemployed people live in Italy, at what age (on average) people get married, what responsibilities a certain position entails, etc.;
- Actions carried out: for example, food and beverages a person had yesterday, the frequency of sexual intercourses, the books read in a year, the way new personnel is selected in a company, reactions to stressful situations;
- Personal, socio-demographic characteristics: age, sex, income, schooling, religion, nationality, etc.

A question can have a varying degree of intrusiveness: *high* (when it concerns topics which are considered personal or highly confidential, such as sexual behavior, the party voted for in the latest elections, the reasons for a very negative mood); *medium* (questions are considered mildly intrusive, not very aggressive – for example inquiring about the health status of an individual without serious illnesses); *low* (when topics would be present in everyday conversations, such as the weather during holidays, the last movie watched, who won the football championship). A question can also have a temporal dimension: *past* (regarding a behavior/attitude/value the individual had weeks, months or years ago); *present* (what the individual does or believes nowadays); *future or hypothetical* (what the individual might do or think in a month or more, but also how he would act in a hypothetical situation).

1.1.3 Variables Examined in a Research

In scientific terms, *variables* indicate the phenomena assessed in a research. Variables can relate to different characters: psychological (such as the emotional well-being of a person), social (his network of relationships), economic (such as his income), political (such as the party he votes for), etc.

A variable clearly describes a phenomenon which can vary over time, and we are usually interested in measuring its variations in a population. Such variations are classified operatively in categories, which can be verbal labels, numbers or values. For these categories to be informative, they must be *exhaustive* and *mutually exclusive*:

the set of categories must include all the relevant variations, and each answer should fit precisely in only one option. Identifying the categories is a conceptual process which highly depends on the definition of the variable that the researcher holds from a theoretical standpoint. A researcher will provide the categories which are most useful for the objectives of his research, and will define operative criteria to include objects in a certain category (for example, a minimum and maximum values should be specified to identify the category Low Income).

Variables can be *continuous*, that is they can take any value in a numeric scale (e.g., the weight or height of a person), or *discrete*, when they vary discontinuously and can only take certain values (e.g. sexual gender, color of the eyes, the model of car owned, etc.).

Variables can be measured at different levels, that is using measurement scales with a different degree of precision. The type of measurement (that is, the set of rules we use to assign numeric values to certain characteristics) are important because they influence the operations we can carry out on the values. Measurement scales normally used in statistical analysis are:

- *Nominal scales*: they are also called categories, and are the least precise and they show variations of a variable, but do not measure specific attributes. The categories do not have a rank or order, and values are purely arbitrary (for example, we can assign 1 to Male and 2 to Female, but the opposite would be equally acceptable). The values assigned to each category does not indicate a measure of quality, it simply identifies that option. The distance or difference between two categories holds no meaning. When data are expressed on a nominal scale, we can only measure the frequency of a certain category (for example, considering the customers of a shopping mall over a week, 63% were female, 37% male).
- *Ordinal scales*: in this case, the variations of a phenomenon can be ranked along a certain dimension. The values allow to classify (that is, to determine a hierarchical order for) the objects, depending on the ordinal position. For example, a medical condition (satisfactory, poor, serious), the ability of an individual, the comfort of a car, etc. With these scales, we can identify which object ranks first, second, etc. but we cannot measure the degree of an attribute in each position, and the distance between one position and the next. Basically, other than rough order, no precise measurement is possible.

- *Interval scales*: variables can be ordered along a continuum of values by using equal intervals. Typical examples are the temperature scale in degrees, or the height in centimeters. This type of scales provides more information because differences between numerical values are meaningful, allowing to compare more accurately. The 0 value is arbitrary, purely conventional. Interval scales are effective not only to rank two objects, but also to measure how much one object is superior to another for a certain quality. It should be noted that in social sciences, these are the most precise scales, yet the assumption that equal differences in values indicate equal variations in degree does not always hold true (for example, for a student it is clearly better to get 8 as a mark rather than 4, but one cannot claim that a student receiving 8 is twice as prepared as one receiving 4, or that the difference in knowledge between these students is the same as two students getting 6 and 10).
- *Ratio scales*: this is similar to a nominal scale, but the difference between any two values accurately measures their distance, and the value 0 normally signifies the absence of a certain quality. Examples of ratio or proportional scales are: the frequency of a certain event occurring, the income in Euros, the age in years, the talent of a swimmer measured in seconds necessary to complete a race.

To avoid dealing with too many values, ratio and interval scales often use classes of numbers, that is intervals that hold particular interest for the researcher. For example, we are not interested in the exact age of an individual, but rather the range he belongs to: 1 = between 13 and 17; 2 = between 18 and 25; 3 = between 26 and 35, etc.

A variable can be conceptualized and measured at different levels, depending on the objectives of a research. For example, when asking an individual if he takes medications, we could simply accept nominal answers like Yes and No, or we could provide ordered alternatives, defining an ordinal or ratio scale (0 = never, 1 = once every 6 months, 2 = once a month, 3 = once a week, etc.)

A research usually examines *relationships* between variables: a *dependent* variable is that whose variations are affected by different values of another variable, called *independent*. Dependent variables indicate those aspects which we want to observe, describe, explain or ultimately control. Typical independent variables are sex, age or income of an individual, while opinions, beliefs and behaviors are assumed to be dependent. Questionnaires and interviews include questions on both variables.

Almost any variable can be dependent or independent, depending on the objectives and the hypotheses of the research. For example, the marital status of a person is a classical socio-demographic variable normally assumed to be independent. Researchers would often be interested in different opinions or behaviors of individuals, depending on their marital status. Yet, during their lives individuals tend to change their status, going from unmarried to married, divorced, etc. This natural variation of status may be exploited to examine this variable as dependent: for example, the likelihood of being married or unmarried could be calculated based on information such as age, or income, or craving for children.

An independent variable is defined as experimental when they can be manipulated or controlled by the researcher to monitor what effects different levels of this variable have on a dependent variable. For example, conducting a research on the quality of data collected in a poll, the level of interviewers' training is often an independent experimental variable, while the quality of the answers represents the dependent variable, measured for example in terms of accuracy (high, medium, low). The idea is to verify the hypothesis that a thorough training leads to more accurate answers.

The objectives of a research are normally to expand our knowledge on a certain field: the frequency of a behavior in a certain population, the distribution of a characteristic across income classes, etc. On a more complex level, a research may aim at examining the relationship between two or more variables in terms of correlation, which measures the degree of simultaneous presence of two variables. If A is always present every time B appears, we speak of positive correlation and use the value 1; if A is missing every time B is present, we speak of negative correlation and use the value -1; we use the value 0 when two variables are not correlated; a partial correlation between variables is expressed with a value between 0 and 1. Examples of variables whose correlation has been proved in many empirical researches, are mood and weather, or income and level of education. It should be noted that, despite noticing meaningful correlations between variables, this level of investigation cannot help draw conclusions on the causes of the relationship and its direction.

On a more complex level of investigation, researchers try to determine a cause-effect relationship between variables. Such a relationship presumes that the cause exists before the effect, and more often than not a linear model is adopted. Unfortunately, the case can often be more complex, with many variables yielding the effect, or with a non-linear model: for example, a certain cause A (age), may produce a behavior B only for people between 20 and 30, but not older.

1.1.4 Population, Sampling, Data Quality and Validity

A relatively small sample, as long as it is adequately defined, allows to derive valid inferences on the entire population, or universe. This explains the popularity of exit polls, since they produce accurate predictions on electoral results when based on representative samples. The target population can be all the citizens of a nation, the households, the employees in the banking sector, car-drivers, newly married couples, etc. but also not people, such as discotheques, elementary schools, companies with specific characters, etc. Inferences based on a representative sample can only be generalized to the population that the sample was extracted from. An adequate sample is a random or probabilistic sample, which represents a miniaturized model of the target population. The size of the sample, relative to the size of the population, is not important since a thousand individuals can be an adequate sample for many millions or just a few millions. It is generally accepted that the sample is probabilistic (that is all members of the population have the same chance to be included in the sample) and representative (that is the distribution of certain characters such as age or income, reflects the distribution in the entire population).

Once the target population has been clearly defined, with unequivocal criteria to include or exclude an element, probabilistic samples should then be extracted. The most general and known method to extract a probabilistic sample is by *simple random sampling*: from the list of all members of a population, x elements are extracted randomly. Variations on this approach may take into account information on the population to make the sample more representative: for example, when studying psychologists, it might occur that 60% are male, and 40% are female, therefore it would seem reasonable to reproduce these percentages in the sample. *Multistage sampling* is used when the population can be naturally divided into subgroups (geographical areas, level of income, schooling, etc.) and the sample is formed by extracting elements from each group bearing in mind proportions.

Ideally, a *sampling frame* (that is a complete list of the members of the population) would be available, but that is rarely the case. For example, phone directories are often used but this can be distorting the data in many ways: for example some areas might have fewer users, in percentage, than others, because of economic or geographical reasons. Therefore, to make sure the sampling method is valid, information on the sampling frame should be collected.

There are sampling strategies which are not probabilistic, for example when the subjects contacted come from specific groups such as the interviewer's acquaintances,

the customers of a supermarket, through the ads on a noticeboard. There is specific literature on these forms non-probabilistic sampling (e.g., quota sampling, snowballing, convenience sampling, etc.) which are used in particular circumstances, especially when selecting a truly probabilistic sample would be too expensive in terms of resources needed. Data collected this way cannot be generalized since we have no guarantee that the interviewees are representative of the population, still sometimes this is the only option, and also some first-hand information can be extracted to give a rough idea of the characteristics of the population, useful for further examination. More on non-probabilistic samples will be discussed later on in this paragraph.

Another important aspect to determine the quality of collected data is the *response rate*, that is the proportion of subjects, amongst those initially selected to form the sample, who actually provided the information. It is nearly impossible to contact all the members of the sample, and to obtain their cooperation. The total non-response rate is often labeled *missing subjects* or *no answer*. The problem is, as many researches have shown, that missing answers are not distributed randomly, thus creating important distortions in the data (for example, non-response rate is often higher in towns than in villages). For surveys based on samples, the response rate that is unanimously considered as necessary is equal to or above 75%. Besides, even with a high response rate, sometimes some questions (maybe the critical ones) are left unanswered, which may prevent from a valid statistical analysis.

Many strategies can be adopted to increase the chances of an acceptable response rate: sampling more individuals than necessary (to make up for those who won't reply), using incentives (typically, monetary rewards), changing the lay-out of the questionnaire (which may involve complex psychological reasoning), training interviewers, using different methods to collect data (for example, postal forms usually have the lowest response rate).

Several research projects conducted by private companies use convenience samples, that is samples whose members are easy to contact, or expected to be cooperative, possibly because this way not much time will be wasted, or money spent, to reach the size of the sample. Oftentimes the reason is simply laziness or ignorance, rather than limited resources. Clearly a convenience sample examines a specific group (for example, uni students, or customers of a supermarket, or hospitalized patients), which cannot be representative due to their condition of "captivity".

When a research analyzes traits which we may expect to belong homogenously to any human being, we have no reason to fear that results may be invalidated by the

peculiarities of a certain sample. When instead we can hypothesize that the phenomena under scrutiny might have relationships with social, cultural or demographic variables (age, sex, level of education, income, etc.), the possibility of generalizing these results becomes crucial. Basically, a sound method of sampling is a prerequisite to be able to infer any conclusion on the whole population. Optimal design and implementation still are not enough to guarantee the validity of a research, since many factors can reduce the quality of data, introducing distortions. Although any distortion can ultimately be ascribed to the strategies used by the researcher to design and implement the project, here we provide the main causes:

- Low quality of the tool used to measure data (for example, inadequate questions, incomplete checklists, poor presentation of the project).
- The survey takes place in the wrong context (for example, lack of sufficient privacy, noisy environment, etc.).
- Various psychological processes may interfere with the validity or truthfulness of the answers (for example, the subject wishes to give a socially acceptable image of himself, or it may simply be difficult to recall certain information);
- Lack of expertise or professionalism by the interviewer, which may have many drawbacks: subjects are less cooperative, questions are rephrased poorly changing their meaning, ambiguous answers are interpreted instead of inquiring further, etc. An interviewer doing a poor job increases the amount of missing data, where "missing" here means not accurate, badly recorded, and consequently unusable.
- A high total non-response rate of the subjects sampled, which may be due to lack of cooperation, or simply wrong information to contact them (phone/address/email/etc. are not correct). A low response rate usually implies that the sample is no longer representative.
- The partial non-response rate (that is, the number of missing or unusable responses for specific questions) can cause distortion. It is normally denoted as missing answers, and may differ a lot from question to question (usually higher for intrusive ones).
- Poor encoding and analysis of data.

The validity of data is a broad concept with several facets. First of all, statistical validity: the idea is to correctly identify a relationship between the variables observed, more specifically which factors are the source of variability of a certain phenomenon. The validity of any conclusion regarding the presence of a statistically meaningful

relation between variables, can be compromised by two types of error: concluding that there is a relationship when in fact there is not, or concluding there is no relationship between variables when in fact they influence each other. Statistical validity might not be guaranteed because of many factors: the general assumptions to use a certain statistical test are violated, measurements are not reliable, the sample is too small, presence of variables which have not been correctly identified, etc.

Statistical validity is closely connected to internal validity, which has the following requirements: a) variables under examination have been properly defined; b) no confounding variables are present (those which may be the source of confusion/misunderstanding, or wrong relations); c) relationships between variables must be properly defined, in particular the direction of any causal relationship hypothesized. Internal validity is guaranteed when it is possible to prove that the observed relationship is causal, that is when variations of the independent variables are the actual cause of variations in the dependent variable.

External validity expresses the quality of those results which can be generalized to different (temporal, cultural, etc.) contexts. External validity has three elements: population validity (results can be generalized to the whole population, therefore the sample is presumably representative); temporal validity (how stable the results are over time, and not due to seasonal effects); ecological validity (results are applicable to real life situations, and are not "artificial", that is they are not the consequence of the particular context where the research has taken place).

Finally, *conceptual validity*, which refers to psychological and behavioral variables, usually rather complex, which cannot be measured directly but only inferred from behaviors observed. This validity also refers to the process of theorization or conceptualization, that is how well a variable represents reality, and is a valid indicator of a situation. Because of its subtle psychological and theoretical implications, this validity is beyond the scope of this paper, and is only briefly mentioned.

1.2 Main Components of a Research

This paragraph will examine the various "ingredients" of a research, trying to define and analyze them in an unambiguous way, although there is always a certain degree of arbitrariness.

1.2.1 The Actors of a Research

As mentioned before, a research aims at collecting information from subjects and it serves someone's objectives, like a researcher, a public body, or a company. We could use the term "actors" to define the individuals involved:

- *Clients, researchers and "performers" of the research*. It may seem unnecessary to distinguish between the *client*, who decides to carry out a certain survey, and the *researcher*, who translates these intentions into a plan by designing each phase and monitoring its evolution, and the person who actually puts it all into action, such as an interviewer. Sometimes these roles are performed by the same person, for example when economic resources are low, but that does not imply lower quality results. On other occasions, specialized agencies are used for each activity. Yet again, an actor may be absent, for example when postal forms are utilized. Still, it is important to distinguish the client (and his objectives) from the research designer, who implements it: you may have excellent objectives with very poor implementations, and vice versa. Another important aspect is the identities of client and researcher, because normally the latter's can be revealed since it is perceived as neutral, whilst the client's identity may influence the answers, therefore is often kept secret.
- *The interviewer*. This actor normally has a subordinate role to the person who plans the research, yet he actually implements it, and in some cases, like a job interview, his role overlaps with that of the researcher. Interviewers can also be vital in the initial phases of design to evaluate the clarity and structure of the questions. Also, the interviewer's training and professionalism is paramount to guarantee the quality of the data being collected.
- *The interviewees or subjects*. The main actor is obviously the person being interviewed, because he is to provide the information we are interested in. Sometimes he acts as a *proxy respondent*, that is he gives information on others, like their spouse, children or colleagues. When planning the research, you should bear in mind the social, demographic and psychological characteristics of the interviewees, in order to obtain as much cooperation as possible.

1.2.2 Overall Design of a Research

With research design we mean all the aspects regarding the planning and implementation of a research, including the phase of data collection, analysis and final presentation. We could present the linear model often used to describe surveys:

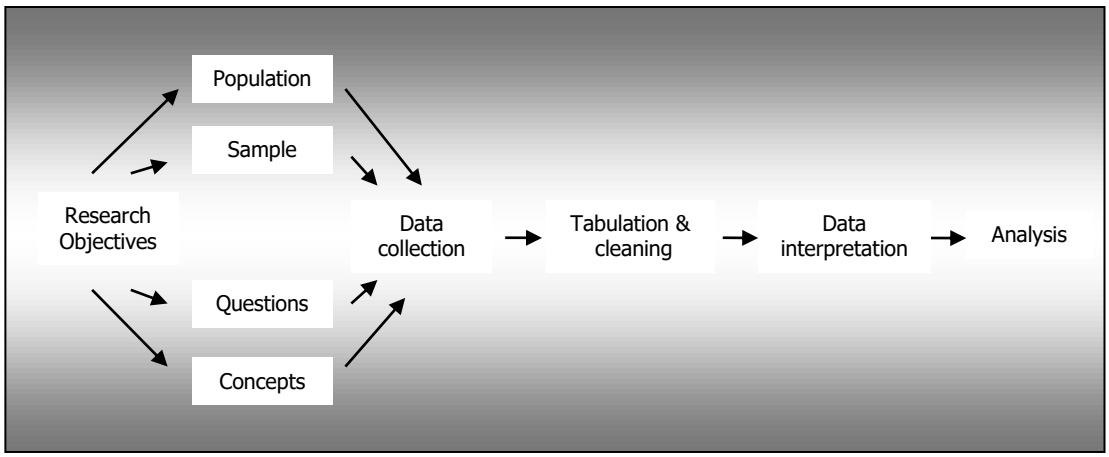


Figure 1. Sequence of the phases when designing a research.

According to the model, specifying the objectives is the first and foremost phase; this phase is connected to a second phase where the target population is defined, a sample is selected, and questions are formulated, depending on the concepts and objects which are to be examined. Data collection can be implemented in many ways, which will be later presented. The final three phases include tabulation, with encoding and cleaning of data, interpretation and (statistical) analysis, often followed by a final report or publication. The whole process, although present as linear here, should be recursive: decisions in each phase may influence all the other phases because they are logically connected, so many authors use the *total design method* where all elements are interconnected.

During the initial phases, a designer should decide whether to collect fresh data by interviewing new people, or to re-analyze data previously collected and available in databases, archives, etc. This second, indirect way may be very effective, but is obviously not always applicable, for example when current information is required.

1.2.3 Resources and Constraints

Resources and constraints of a research can affect one or more aspects during the planning of a research. Theoretically, we can classify resources and constraints into strictly financial ones, and non-financial ones, which we may define technical.

a) *Financial resources and constraints*. Financial resources affect decisions in every phase of a research, including for example the preparation of materials for tests or questionnaires, and the final analysis and presentation. Clearly they affect the way data is collected (face-to-face interviews are more expensive than postal forms to be filled in, professional interviewers cost more than student-interviewers for a university project).
b) *Temporal constraints*. There is usually a deadline to complete the research, otherwise data would be less relevant or even useless, or interviewees might not be available any longer. These constraints can be particularly severe when choosing the most appropriate way of collecting data.
c) *Constraints regarding the sample or the subjects examined*. Constraints here depend on the characteristics of the individuals being asked questions: how large is the sample? How dispersed geographically? How easy to contact? Some groups of people have special requirements, which limit the way information can be extracted (kids who cannot read would need visual tools; students can only be contacted at school and under their teachers' approval and supervision, etc.). Occasionally, it would be necessary to use proxy respondents or key informants, that is people well informed who provide information on others.
d) *Constraints relative to questions and categories in the answers*. Questions are the tool normally used to serve the research purposes. Occasionally, literature provides many such tools, like sets of questions, or nominal scales, that proved to be effective. If an adequate tool cannot be found in other researches, building a questionnaire can be costly and time-consuming. Several decisions must be taken when preparing this tool, as numerous options are available (how many questions, which order, professionalism of interviewers, etc.)
e) *Constraints and resources regarding the analysis of data*. The type and complexity of the analysis must serve the purposes of a research, and affects previous phases (for example, a large sample and open questions can only be used if enough people may be used to extract and encode all the information). Too often inexperienced researchers collect loads of data they do not have the time or skill to examine, or realize only too late that the questions posed have the wrong format for the analysis planned.
f) *Constraints regarding main research designs - measures repeated over time, and unique measures*. An important factor when planning a research is to decide whether only one measurement is needed (*one-shot design*), or more

measurements of the same subjects would be required over time (*longitudinal design*). A longitudinal design (or *panel*) is normally used: (1) to monitor the evolution of some variables in order to find meaningful trends, or (2) to record the values of some independent variables over a period of time, to see if they allow predicting the values of a supposedly dependent variable. Longitudinal designs have specific constraints: for example, we should be able to reasonably assume that the sample will be available and willing to collaborate more times. The time interval between measurements should also be carefully planned, depending on the objects investigated, and the purposes pursued (for example, when evaluating the impact of TV adverts, it would be advisable to measure spectators' attitudes once before the advertising campaign, and at least twice afterwards).

1.2.4 The Context

Any research is carried out at a certain moment in time and within a precise cultural context – variables which are often overlooked in research manuals. Depending on the culture, certain topics can or cannot be addressed, the wording would be more or less appropriate, and results can be generalized. If you bear in mind this element, all phases of the design will be affected. Another element of context would be the location where the exchange of information takes place (in a company, at the interviewee's home, inside a shopping mall, etc.), which can affect the willingness of the individual to cooperate, especially on confidential information. A final contextual element is the way the research is presented: is information given to clarify its purposes and confidentiality? Are the questions really clear? What are the look and voice tone of the interviewer, his general demeanor?

1.2.5 Measurements

By measurements we mean the tools used to collect data, but the term also refers to the data itself. So typical measurements would be the sets of questions and answers used in questionnaires or interviews, but this kind of measurements can be used with other ways of collecting information, like direct observation or the analysis of archives.

Questions do not always come in the form of an interrogative sentence, such as "How old are you?" Another form would be a statement, like "The price of fuel is too high", where the subject should agree or disagree.

Occasionally, you may have pictures or drawings, to present objects or situations. Also, questions may be more or less direct, that is they may reveal explicitly to the subject what object or dimension is being examined.

Questions can have a closed-ended format (where there are only two possible answers like Yes/No, True/False, Agree/Disagree), or open-ended questions (where the subject can articulate whole sentences in response). Another common format is multiple-choice questions where the subject is given a list of possible answers to choose from.
Several elements ought to be considered when building a list of questions: number and variety of topics (multipurpose questionnaires refer to different topics, while a focused questionnaire analyzes different aspects of the same topic); the area of investigation (social, political, commercial, etc.); the number of questions (the overall total, as well as the subtotal for each topic); the objects of the questions (attitudes, behaviors, etc.); the format of questions (closed-ended or open-ended, but also multiple-choice with a selection of alternatives); the degree of standardization (with reference to the wording of questions, the alternatives presented, the interviewer's behavior, etc.)
Measurements obviously include the answers provided, and they are the data from which to extract the information we are looking for. The expression *quality of data* is a phrase which refers to a rather complex idea: answers can be better or worse depending on their being complete, truthful, unambiguous, generalizable to other members of the population, etc.

1.2.6 Methods of collecting data

Information can be obtained from the interviewees in different ways, but essentially there are three options:

1. A questionnaire filled in by the user, where a typical application would be a mail survey or form; a more recent variation of this is a survey conducted on the web, where the subject replies using keyboard/mouse/touchscreen/etc.
2. The phone interview, which has the advantage of being rather economical but often scores low on response rate, and cannot use visual aids;
3. The face-to-face interview, clearly the most expensive option, but the psychological ability of the interviewer can play a critical role for the quality of the answers

Along with these "sheer" cases, mixed formats are possible: a questionnaire is presented to a group of subjects by an interviewer who provides explanations, the purposes of the research and other generic information, and is available during the whole process in case subjects need further elucidations.
Interviews can be more or less standardized and structured, that it to say how much the following items have been determined in advanced: 1. topics that will be examined; 2. the order in which questions are posed; 3. the wording used for each

question. A *structured* interview takes place when the interviewer has no freedom to change topics or the order of questions (for example, when a journalist interviews a politician or scientist, digressions can happen, and only a loose topic guide may be followed). When the formulation of each question is exactly the same, the interview is also standardized, and subjects are exposed to the same stimulus.

In order to collect data useful for generalizations and predictions, structured and standardized interviews are needed, and these interviews are fully comparable to a questionnaire, the only difference is that the interaction is oral, not written.

Another possibility is the semi-structured interview, where there is a topic list to be followed but the order and the wording of questions can change depending on the answers. This is typically used in face-to-face interviews, and can be particularly useful in the preliminary stages of a research when designers have only a superficial knowledge on the topics. Unfortunately, this type of interview largely depends on the ability of the interviewer, and is only rarely used by researchers who need generalizable data.

It is also possible to combine the three main ways of collecting data, using a mixed format that is suitable for specific situations. For example, in a face-to-face interview you could pose most questions orally, but for some delicate issues or where privacy is crucial, a form is filled in by the subject on his own; also, a mail survey may be supported with phone interviews to motivate the subjects to answer, provide explanations, or just make sure the form has been received. Finally, there exist some particular techniques to collect data, which go beyond the scope of this paper (examples are the "thinking aloud" procedure, focus groups, or keeping a diary).

1.3 Questions and Alternative Answers

As mentioned briefly before, questions can be classified according to different parameters: the object being investigated (attitudes, general knowledge, etc.); the medium used to retrieve the information (for example, a questionnaire on paper or an interviewer); psychological dimensions (such as intrusiveness and importance for the subject); the format of a question, which we will articulate further.

1.3.1 Open-ended or Closed-ended Questions

There is a wide variety of questions, but a fundamental classification divides them into two formats: open-ended and closed-ended questions. This greatly influences the kind of answers provided, in terms of ambiguity and precision, but also in terms of work

needed to encode the data. The decision regarding which format to use depends on many factors, including financial resources, quality of data, and topic.

The open-ended format gives the interviewee more freedom to articulate, but requires more analysis afterwards. The closed-ended format is adequate to measure the relative importance and frequency of a series of issues. It can be used when a) a scale can be built to clearly identify the degree of a certain characteristic; b) a list of categories can be defined, where each category represents a relevant variation for the variable observed.

A mixed format is not unusual, to meet specific needs or circumstances. It is noteworthy that, in closed-ended questions, researchers often add "Other" as a final category: this is common when the listed alternatives do not cover all the options, because this is simply impossible due to too many options, or the topic is still being explored. With such a category, the question almost becomes an open-ended one, since more options could be specified, and later analyzed.

1.3.2 The Objective of a Question

Another criterion to classify questions refers to the objective, or cognitive function, performed. Four classes have been identified:

- *Substantial questions*. These are the questions that can actually provide valuable data, for which the research was designed (examples are: how many books did you read the last year? Are you happy in your marriage? Etc.)
- *Interactive or introductory questions*. These are normally present at the beginning, to introduce a new topic from a general point of view, or simply to create a feeling of ease and trust (e.g.: Are you generally satisfied with TV programmes? Do you support the government policy on immigration?)
- *Filter or branching questions*. They serve the purpose of discriminating between subgroups, or selecting through a salient character. The typical format is dichotomous (Do you have any children? Do you drive a car?), but sometimes more than two options are available. Depending on the answer, only the logically connected questions are asked, while other questions or entire sections are skipped.
- *Buffer or filler questions*. The object or topic of these questions is not relevant for the research; the buffer question serves only the psychological purpose of distracting the interviewee's attention, making him forget the previous questions, and preparing him with the right attitude. Sometimes, instead of buffer questions, other actions may be asked, like reading a text or counting

backwards. On rare occasions, answers to these questions could help other researches, but their psychological effect should be carefully assessed.

1.3.3 The Number of Choices when Answering

Another criterion is to classify questions on the number of items available when answering. That is possible with open-ended questions as well, when you clearly mention how many elements are required (e.g., Which are the *three* most important qualities you appreciate in a fellow worker?). Regarding closed-ended questions, this criterion counts the alternatives available, and two options are normally used: *forced-response* questions, where only one item can be selected, and *multiple response* questions, where a checklist is presented and more choices can be checked by the subject.

With checklists, a (minimum) number of items is required to be selected, and the category "Other" or "Don't Know" are present. The items do not exclude each other, but should be exhaustive for the variable examined. Checklists can vary considerably in their structure: statements, lists of adjectives or nouns, ranges of values, etc. Frequently they require the candidate to rank the answers in order, that is to somehow specify a classification from most to least relevant. The number of items has psychological importance, since a list can be too long and be perceived as complex or boring. The order in which items are listed has relevance since, when "mark all that apply" is used, very often only the first choices are selected.

The forced-response format is very frequent, and in many cases is dichotomous (e.g., Do you drive a car? Are you married? Do you support government policy on housing?). Occasionally it has more than two items, but they must be mutually exclusive (What is your marital status? Unmarried/Married/Divorced/Widowed). In sociological and psychological researches, nominal categories are normally used, like the example in the previous sentence. Answers can also belong to a judgement scale of another kind (ordinal, interval or ratio). The individual chooses his answers among *n* items, ranked in terms of preference, importance, size, etc. Research shows that 6/7 items are to be preferred in written questionnaires, while 3/5 options are more suitable for phone interviews where the respondent's memory gets involved. Graphical versions of the scale are also possible. Each position in the scale is identified by a verbal label: for example, when evaluating the frequency of an event, we could use categories such as Always, Often, Sometimes, Seldom and Never. These categories do not univocally quantify or measure the frequency, except for Never which clearly stands for 0. The quantity associated with a category is not a trivial or simple decision.

A final format is the *attitude scale*, used in social researches: a group of forced-response questions, relative to a complex topic, where items are typically rated with an interval scale (e.g., from 1 = totally agree to 7 = totally disagree, or defining the importance a topic has). This scale aims at measuring the attitude of an individual towards a certain object, where an attitude is usually assumed to have three components: behavioral, cognitive and emotional. The individual's disposition can be examined in all components, therefore the attitude scale is normally used to take multiple assessments on a topic, with several questions posed.

1.3.4 Response Categories with Multiple Choice Questions

The alternatives presented to an individual influence his understanding of the question, and ultimately the quality of data. Linguistically, the answers fall into two categories: statements or complete sentences regarding the object (e.g., "There are too many foreigners living in my country"); single words, labels or expressions, with a clear meaning. The subtypes of this final category are detailed hereunder:

1. *Single words*. Items denote events, values, people, emotions, etc. They can also denote frequency, agreement, importance of/with a certain object (e.g. Excellent/Good/Sufficient/Bad/Terrible).
2. *Numeric labels*. They can denote single values, or intervals, and they can measure the frequency of an event, as well as its magnitude (age, income, etc.).
3. *Expressions which indicate an assessment*. To clearly denote a status or event, an emotion or behavior, it is sometimes necessary to use expressions which imply an assessment or evaluation: "a happy marriage", "a gratifying job", "a political party with a clear economic plan". Notice that the object is not presented in a neutral way, but with a qualifying term, therefore the interviewee judges a specific status or quality.
4. *Words or phrases linked with numeric labels*. This format is very commonly used, and associates a numeric value to each answer, with phrases to better define the meaning of an item. Typically, you can choose between 7 options, numbered from 1 to 7, and a phrase to better define the meaning may be on all items, only on the two extremes, or on the extremes and the middle.
5. *Visual elements, with or without labels*. The scale to represent different degrees of intensity may be visual, to measure size, probability, agreement with a statement, etc. Normally only the extremes are labeled, and the subject expresses, with a position along the scale, his inclination.

When using graphical scales, a criterion must be chosen to assign a value. If, for example, a straight line is used, with one end representing 0 and the other 10, the distance in centimeters from the origin seems a reasonable pick. Often instructions should be provided for the subject to avoid ambiguous answers.

1.3.5 Intrusiveness, Salience and Temporal Dimension

Other criteria to classify questions concern elements such as: intrusiveness (how embarrassing or personal the question might be), salience (how important and relevant the topic is perceived), and temporal dimension (whether the question relates to present aspects or something in the past or even in the future). Research has shown that, to motivate accurate and truthful questions, it is preferable to pose salient, unintrusive questions regarding the present, the recent past or the immediate future.

Questions regarding the future – just like those which refer to a hypothetical situation – are called *hypothetical* and are often related to a behavioral intention. They are subject to contextual circumstances, thus the individual cannot give a really informed or truthful answer: intuitively what the person answers *now* might not match what he will actually do, as his mood, emotions, level of knowledge on the topic might have changed substantially.

The degrees of intrusiveness and salience of a question are tough to measure in general, since they depend on the individual, the historical and cultural context, etc. In fact, a question may be very relevant for a person and uninteresting for another, embarrassing for someone and perfectly acceptable for someone else. Money, sex and health are usually salient but delicate matters, and many topics may be added to the list. Along with the topic, other factors can influence the interviewee's perception:

- The wording of a question: compare "Do you get drunk?" with "Do you sometimes happen to drink more than you intended to?"
- The format of a question: a closed-ended question may for example appear less intrusive if the alternative answers, including a certain behavior, help the subject feel "normal" or less embarrassed.
- The way questions are posed: normally, a written questionnaire and phone interviews reduce the degree of perceived intrusiveness, compared to a face-to-face interview, since the subject feels the answers are more confidential.
- The way a question – or the whole research – is presented: for example, thanks to an adequate introduction, any perplexity or reticence, that would usually be present otherwise, may vanish. An introduction can also make interviewees more aware of the importance of their answers.

Intrusiveness and salience are therefore crucial aspects to bear in mind when designing questions. These issues can only be partially treated considering: (a) results of previous researches on the same topic being investigated; (b) what is suggested by common sense, that is which norms apply for the socio-cultural context in examination; (c) what results emerged, using different questions, during pretests or trial phases, utilizing tools such as focus groups.

1.4 Building and Assessing the Tool Used to Collect and Analyze the Data

A research is carried through for one or more objectives, typically to find relationships between variables, which relate to characteristics of the interviewees. To achieve its goals, a research should be carefully designed and implemented in a series of steps or phases: some refer to the design and construction of the measurement tool, other to collect and analyze the data. These phases are interrelated, sometimes representing recursive cycles, therefore an "overall design approach" is advisable, where specific roles and interconnections are borne in mind.

1.4.1 Different Phases of a Research Project

Here we will examine in more detail the phases constituting a complete research:

a) Examining the topic from a conceptual and theoretical point of view: researches examine specific topics or issues, often investigated in their different facets. It is paramount to clearly focus on *what* should be studied, and *why*. It is often indispensable to get as much information on the issue as possible, to avoid redundancy. This entails examining any literature on the topic, and possibly meeting with people considered to be experts. This preliminary phase allows knowing what has already been discovered, and which tools are available.

b) Desirable qualities and technical characteristics of the tool, the sampling phase and data collection: Once the relevant variables have been identified, hypotheses and variables should be examined more analytically. So, considering any financial constraints, the research should produce a list of variables (dependent and independent), the alleged relationships, and the most adequate methods and levels of measurement. Literature can provide a lot of support for decisions in this phase, for example which scales proved to be more accurate in a certain context. Theoretical and methodological decisions taken in this phase depend on and affect the following phases, particularly the next one.

c) Choosing the sample and the tool to collect data: existing financial and temporal constraints, and available resources affect decisions in this phase, and such decisions can be extremely restrictive on subsequent phases. Here the researcher decides which is the most appropriate way to collect data (whether he needs an interviewer, or questionnaires should be used, etc.), which is the target population, the sampling method and its size. The researcher should bear in mind all information available on the topic and the population, and given all the constraints, find the best compromise.

d) Building the measurement tool: Once the sample has been identified, along with the method to "extract" data from the sample, a draft version is created, in three stages: writing down the questions, specifying all prominent characteristics of the questionnaire and the way it is posed, anticipating how data will be analyzed. Often questions are separated in modules, each containing *n* questions on a specific variable or aspect. To build every question adequately, some factors should be clearly defined: its *content* (which object is investigated, which categories are possibly listed, the scale used to measure, whether visual aids are utilized, etc.); its *objectives* (in relation to the general hypotheses and the relevant variables of the research, how does this question contribute? Also, what is its specific role: substantial, interactive, filter or buffer question?); its *format* (open-ended or closed; coherent with other questions in the same module or not); the *level of measurement* (in case of closed-ended questions, which scale is adopted: nominal, ordinal, etc.); *the number of possible responses* (this only applies to forced-response and multiple response questions); *the type of categories* (verbal or numeric labels? Should all intervals be labeled or only the extremes? Is the option "Don't Know" present?); *the level of specificity* (a generic introductory question, often followed by a more specific one); the order of items (how categories, or simple statements, are listed); *the length of a question* (research shows that brevity and concision are effective: the question is understood and not skipped by the subject; longer sentences can be used to give more time to think, and can be interpreted as a sign of salience/importance); *the wording* (this is the final touch and has psychological and informative effects on the subject). Sometimes focus groups or panel interviews with experts take place in this phase.

e) Verifying the adequacy of the tool and building the definitive version: once the tool has been designed, a pretest can be done on a small group, similar to the

final sample. This can provide plenty of information on corrections and adjustments to the tool: ambiguous or wordy questions may be identified, checklists may be too long or incomplete, cooperation may be much lower than expected due to wording or intrusiveness, instructions provided for some questions may be misunderstood, etc. The results of a pretest are normally utilized to build the definitive version, but occasionally modifications are so significant that more pretests are carried out.

f) Using the tool to collect data: in this phase the final version of the questionnaire is used, and all the procedures planned to guarantee a high quality of data are put in place. Any mistakes in the design of the research can only be rarely corrected since they would be too costly, in terms of time and money. Typical problems would be: a very low rate of response on certain questions, an obsolete sampling list (e.g., an old phone directory), insufficient training of interviewers. Making up for these problems requires data to be analyzed as it is collected, from the moment it starts to its completion, in order to set off corrective measures.

g) Data encoding, tabulation and cleaning: answers must be converted in a format that allows analysis in order to draw conclusions. This implies that even verbal answers are to be transformed into *n* given categories, each identified by a numerical value. Often tables of data are created in this phase, where each row identifies a subject of case, and each column is assigned values of one of the variables being examined. The categories of socio-demographic variables typically identify subgroups within the sample (males versus females, people under of above 25, etc.). Values attributed to different answers represent the measures for which the research was set up, and allow quantitative analysis, in order to find relationships between variables. This analysis normally uses computer-aid statistical packages, which allow for most analyses (frequency, mean, variance, correlation, factorial, etc.). Encoding the data must follow clear and explicit rules, which should be specified in writing when the process is not standard and somewhat complex (e.g., How often do you read newspapers? Every day =3, Twice a week = 2, Once a fortnight = 1, Never = 0). Sometimes, rules should contemplate a special answer, like "Don't Know" or "Other". Rules should take into account the analyses that will be done, and must indicate how to deal with missing answers (notice that missing answers can be due to different causes, for example when the subject doesn't know or doesn't want to

give a response, so designers should decide whether to encode this differently or not).

h) Data analysis, interpretation and final report: first, verification should be done to make sure the encoding was accurate (for example, verifying that only acceptable values are present, or that the sum of frequencies for each value equals to total of answers). A "cleaning" could also take place, depending on preliminary results: for example, some categories may be joined because they occur too rarely, some subjects may be taken out because they answered casually, some variables may be removed because they are not very informative or suitable for analysis, etc. The analysis follows, and it is a crucial phase for many reasons. It is not uncommon to see weaknesses emerge only at this stage, which may lead to collect more data, or refine the analysis with different objectives and variables. Finally, a report is written, which may become an article in a scientific magazine, or be used confidentially by the client who has requested the research.

1.4.2 Analyzing the Quality of Data

As mentioned before, during all phases of a research the quality of data should be guaranteed, that is to say making sure measurements are valid and truthful, avoiding any distortion as much as possible. This is achieved with several practical techniques which vary from project to project, but there are some general parameters of quality that emerge in literature, which will only be briefly discussed here, without going into technical details.

- Total response rate: this indicator typically measures the ratio between those who decided to collaborate and give answers, and the total of subjects contacted. The rate depends on many factors, and a well-designed research takes all possible measures to achieve a high level of cooperation. For example, face-to-face interviews have the highest rate, mail questionnaire the lowest; clear questions, or even explanations on the objectives and importance of a research, can increase the rate. A low rate generally indicates a problem (too many questions, intrusive wording, unclear instructions, etc.), and may lead to a distortion of the characteristics of the sample. This indicator is one of the main parameters used to evaluate the ability of an interviewer, as it is assumed that a professional and well-trained interviewer will obtain a lot of cooperation.
- Response rate for specific items: the general assumption here is that a well-designed tool guarantees to obtain an answer for every substantial question,

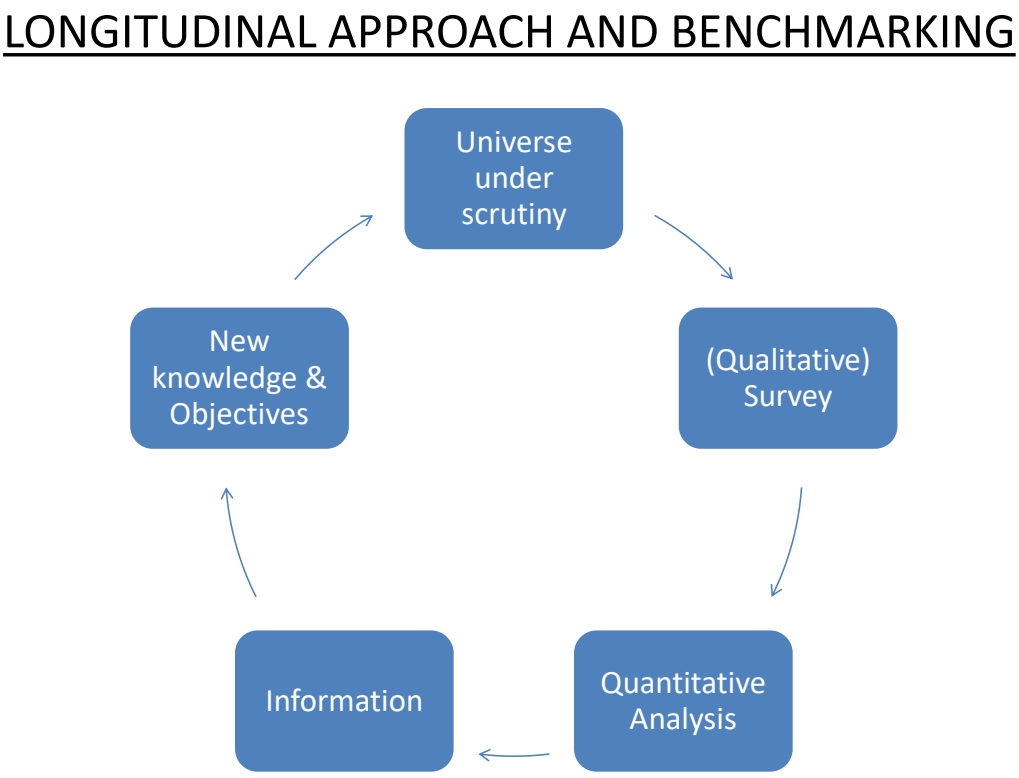
with very few irrelevant or missing data. If the rate of missing answers is high (it is difficult to provide a general percentage here), the reason should be identified and taken care of (ambiguous wording, delicate matter, etc.). Intuitively a high rate of missing answers drastically reduces the quality of the information that can be extracted from a specific answer.

- Profuseness of an answer: this parameter refers to multiple response questions, and counts the average number of statements selected, or the average number of concepts specified in an open-ended answer. The typical assumption is that a higher average indicates more motivation and willingness to cooperate, thus providing complete and accurate answers. This assumption is to be verified case by case, and the minimum number of selected items which shows cooperation depends on several factors (culture, personal involvement/interests, lifestyle, etc.).
- Relevance of an answer: this is normally applied to open-ended questions, where the content of the answer is examined, too see how pertinent it is. A high number of irrelevant answers usually indicates a badly-designed tool (unclear questions, poorly trained interviewers, etc.). Typical occurrences in this context are digressions, or any answer which gives personal details instead of information related with the topic. Research shows that the role of the interviewer, when present, is paramount in these situations: he can clearly identify irrelevant answers, and take corrective measures to obtain valuable data (by helping interpret the answers, by rephrasing the question, by offering explanations or emphatic comments, etc.).
- Distribution of answers in closed-ended questions: when a research design contemplates two or more experimental conditions, where aspects of the measurement vary (interview versus questionnaire; order of questions; wording; etc.), possible distortions can be identified. For example, similar distributions of various categories of an answer in two samples who were given different formulations of a question, allow to conclude that wording does not influence the answer. Analyzing the distribution of the answers in different conditions allows to check for their statistical ”normality”, and the effect of elements which were not originally borne in mind (such as privacy giving answers or presence of distractions, like when being interviewed in a quiet room versus a public place, etc.)

- Follow-up interviews or questionnaires: quality of data can be assessed indirectly by asking how the tool is perceived by the subjects (were questions clear? Did you get bored? Etc.) These particular questions are normally present in a separate questionnaire which the interviewee can fill in confidentially. Naturally, follow-up tools should also be carefully designed and analyzed, with procedures similar to those used with the main questionnaire, in order to draw conclusions and bring any necessary adjustments.

1.5 A New Vision

From the linear approach, outlined in the first chapter, both examples in literature and the experience drawn from the case studies presented have driven us to a different scheme: cyclic and longitudinal.



CHAPTER 2. Carrying Capacity

By the early 1990s, visitors to the U.S. national park system had topped 250 million a year, continuing its historic upward trend. The interest of so many people was something to celebrate, but also presented serious challenges: national parks should be protected from visitors who, in such large numbers, threaten their integrity by trampling vegetation, polluting water and air, disturbing wildlife, disrupting soundscapes, etc. Besides, the quality of the visitor experience was being diminished through crowding and congestion, and the aesthetic consequences of possible resource degradation. The issue of how much use should be accommodated in protected areas is conventionally called *carrying capacity*, and around 20 years ago a group of planners in the U.S. resolved to address this issue. Scientists, university professors and government members devised a framework called *Visitors Experience and Resource Protection* (VERP) to identify and protect what is valuable about parks. VERP defines indicators for park resources and the quality of the visitor experience, provides procedures for monitoring conditions, and requires management actions to ensure standards are maintained. The framework was progressively applied to all units of the national park system, and applications are supported by continuous research. The diversity of the park system in the U.S. allowed adopting, adapting and applying an array of theory and methods from a host of academic disciplines, such as sociology, ecology, statistics, landscape architecture and computer science.

National parks are a manifestation of the challenges associated with managing common property resources, causing a concern which derives from the most fundamental question of all in conservation: how much can we use the environment without spoiling it? In contemporary terminology, carrying capacity is now morphing into sustainability and is expanding into many sectors of environmental management.

2.1 The tragedy of the Commons

The historical lineage of environmental issues can be traced back through centuries, but in contemporary literature they probably best emerge in Garrett Hardin's paper, *The Tragedy of the Commons*, published in 1968. He asserted that without deliberate management action, human use of common property resources would exceed carrying capacity and lead to tragic consequences. Hardin began his paper with an illustration using perhaps the oldest and simplest example of an environmental commons, a

shared pasture: each herdsman will try to keep as many cattle as possible on this commons, since he receives all the proceeds from the sale of any additional animal, with a positive utility of nearly +1… Instead the effects of overgrazing are shared by all the herdsmen, therefore the negative impact for any of them is only a fraction of -1. The rational herdsman concludes that the only sensible course for him is to keep adding animals, without limit, in a limited world… therein is the tragedy.

Hardin identified other examples of environmental commons, in particular national parks and protected areas: at his times, they were open to all without limit, and the values that visitors seek in the parks are steadily eroded, so management should soon cease to treat parks as commons.

This original paper has been republished in over one hundred policy-related anthologies, stimulating research and writing. The work has been applied to a growing list of commons-related resources, such as wildlife and fisheries, surface and ground water, range lands, forests, climate, biodiversity and population. Its conceptual foundation has even been extended to a growing array of public resources that are not necessarily environmentally related, such as education and medicine.

The issue of managing common property has a long history, even Aristotle noted that what is common to the greatest number gets the least amount of care as men pay most attention to what is their own. The first modern expression of the commons issue is credited to Lloyd (1833) who suggested the environmental degradation caused by unfettered population growth and the inability of the Earth to support very large numbers of humans. More contemporary and scientific explications of the commons were offered in the 1950s in the context of ocean fisheries.

Common property resources can be defined as having several characteristics: first, ownership is held in common, often by a large number of owners with independent rights to use the resource; second, control of access is problematic for different reasons (large area, its pervasive character, its migratory nature,…); third, the level of exploitation by one user adversely affects the ability of other users to exploit the resource; fourth, in conventional common property resources users can extract tangible (e.g. forage, fish) and intangible (e.g. enjoyment) benefits. Harding also noted the existence of "reverse" commons, in which pollution is deposited into a resource that is owned in common, such as the oceans and the atmosphere.

How can this tragedy be averted? Note that there are no technological solutions: increased efficiency of resource use might only postpone the need to address this issue, but some limitations will eventually be required. Hardin believed that only two

forms of managements could address the tragedy of the commons: private or government ownership. Private ownership internalizes benefits and costs of exploitation (benefits and costs are both borne by the owner). Government ownership allows for a long-term management perspective, focused on the welfare of society as a whole, thus offering protection for resources important to society.

Western countries, and the U.S. in particular, rely on private ownership to guide production and consumption of goods and resources. This approach in inherent in the capitalist system, and is supported by the concept of "the invisible hand" proposed by Adam Smith: decisions of individuals in a free market economy lead to outcomes that benefit society at large. Yet there are notable exceptions where government action is required. For example, the full costs of pollution are sometimes not paid by producers, who may overproduce, resulting in pollution levels harmful to society. In other cases, such as national parks, private entities cannot capture the full benefits of producing such goods and services, leading to undersupply. These are examples of market failures, where social action is required to regulate. (e.g. laws against pollution, and a national park system).

These types of social actions are manifestations of "mutual coercion, mutually agreed upon": they are limitations on resource use that apply to all potential users, in order to protect the greater welfare of society.

There is a fundamental assumption underlying the tragedy of the commons: increasing exploitation of resources will lead to unacceptable environmental degradation and undermine the ability of the natural environment to support some minimum quality of life. Most discussions of carrying capacity date its modern emergence to an essay published by Malthus in 1798, where he hypothesized that human population tends to grow in an exponential fashion, while the food production is limited to arithmetic growth, as illustrated in Figure 2.

The supply of food presents an ultimate limit to population growth, and if these limits are not respected, the results will be substantial human misery. Malthus's ideas about limits to population and economic growth have become fundamental concepts of the contemporary environmental movement. Popular books were written on these ideas and, based on this lineage, contemporary environmentalists are sometimes referred to as "neo-Malthusians".

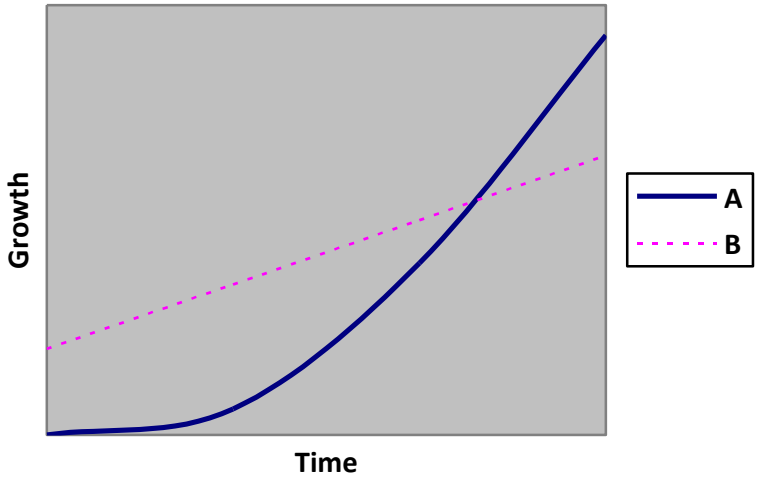


Figure 2. Malthus's model of exponential population growth (A) versus arithmetic growth of food resources (B)

Considering other quantitative treatments of carrying capacity, an early important paper theorized that population growth can be characterized by a sigmoid curve defined by the following equation:

$$\frac{dN}{dt} = rN\frac{(K-N)}{K}$$

where N = population size, t = time, r = rate of population growth, K = an asymptote (a tangent to a curve).

A curve can be derived from this equation, as illustrated in Figure 3.

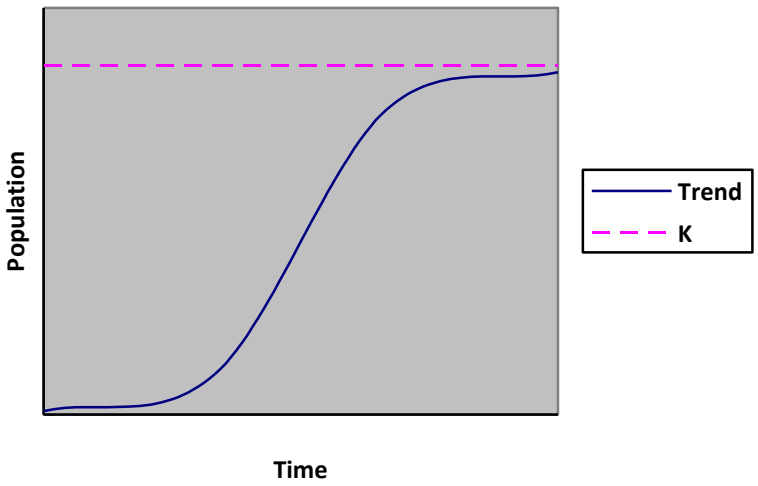


Figure 3. Logistic population growth curve

This formulation specifies that population grows slowly at first, then faster and faster until it reaches an inflection point associated with approaching environmental limits. Afterwards, population grows more and more slowly as it approximates a horizontal asymptote. This asymptote, often denoted as K, represents carrying capacity and is based on some ultimately limiting factor in the environment (e.g. food, space).

Publication of this paper by Pearl and Reed in 1920, sparked immediate interest in carrying capacity and its formulation in several contexts, such as range and wildlife management, but also ecology. It is noteworthy to mention that in laboratory experiments with simple life forms, population growth tends to follow the dictated of the logistic model as expansion is limited by factors such as food and space; however with higher life forms, findings are more variable, in that population growth tends to be mediated by a number of factors, including interspecies competition.

The complexity of carrying capacity increases even further as it addresses issues of human population. It is now widely recognized that carrying capacity, in the context of humans, is mediated by social and institutional issues. For example, one would have to consider questions such as: what level of material well-being should be maintained? How should this material well-being be distributed? What level of environmental protection should be achieved? What social and political institutions should be applied?

Human carrying capacity is not devoid of natural constraints, but human values and related choices must be considered as well. Thus, carrying capacity applied to humans is less mechanistic and deterministic than models such as the logistic growth curve. Recent treatments suggest that one should consider the impacts (*I*) on the environment that human population and related economic growth have, and these impacts are what ultimately dictate acceptable growth. These maximum acceptable levels are largely a function of human values, so carrying capacity analysis and management is evolving from its traditional emphasis on defining maximum population size (*K*) to defining conditions under which this population chooses to live (*I*). A way of describing this in thinking is by redefining K. In its original context, K represents the environmental limits of population, at some subsistence-related level of existence. However, humans might choose to live at higher levels of material and environmental well-being, and such value-based choices might be symbolized by variations of K. This notion is illustrated in Figure 4, which uses the symbols of Kb and Ks to represent biophysical and social carrying capacity, respectively. Ks represents a conscious choice to stabilize or manage population and related economic at a level that is lower than that at the margins of ecological limits. Such a choice is presumably related to a desire

for some minimum quality of life. A similar notion has been found in anthropological studies where human populations have been found to stabilize at a point that is below what is ecologically possible.

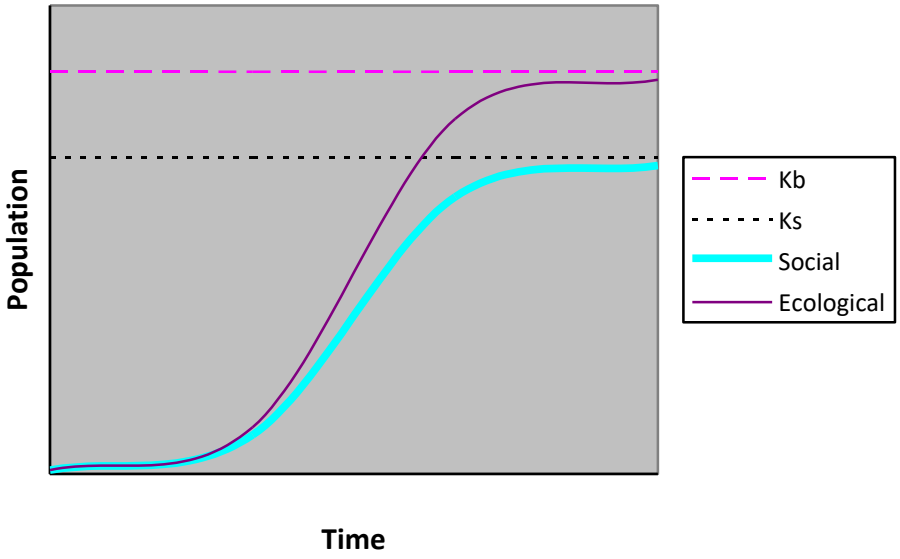


Figure 4. Biophysical (Kb) and social (Ks) carrying capacity

Carrying capacity has been subject to considerable investigation, both theoretical and empirical, and many authors conclude that carrying capacity is vague and controversial, especially when applied to human population, to the extent that it does not offer any empirical guidance. However, recent conceptual models address factors that affect human impacts on the environment, as well as judgments about the acceptability of those impacts and social conditions, and the level of human population that might best be maintained given desired environmental living conditions. Societal norms and values provide a theoretical and empirical foundation for defining the environmental and related social conditions upon which carrying capacity must be determined and management actions needed to avert the tragedy of the commons. This contemporary approach to carrying capacity and the commons is being applied in a number of professional fields, included management of parks and protected areas, and this work is described in the following paragraph.

2.2 Carrying capacity of parks and protected areas

Expanding use of national parks and growing popularity of outdoor recreation have created concern about appropriate use levels of parks, forests, lakes, etc. Most parks and related areas have been established for public use and appreciation, however they must be protected. National parks, as mandated by law at the beginning of the 20th

century, are subject to a two-fold mission, which manifests an inherent tension: parks should be managed to conserve scenery, natural treasures and wildlife, but also provide for the enjoyment of the same in such manner as will leave them unimpaired for the enjoyment of future generations.

Since the 1930s, the number of visits to U.S. national parks has shown a clear upward trend. The question is: how much use can ultimately be accommodated in national parks and related areas? What is the *carrying capacity* of these resources?

As suggested previously, the issues underlying the concept of carrying capacity have a long history in human affairs. The term has received wide use in wildlife and range management where it can be generally defined as the number of animals of any one species that can be maintained in a given habitat. Perhaps the first suggestion for applying the concept of carrying capacity to parks and related areas was recorded in the mid-1930s: a report on policy recommendations for parks in California posed the question "How large a crowd can be turned loose in a wilderness without destroying its essential qualities?" Later in the report, the actual phrase of carrying capacity is used. Only in the 1960s a commission incorporated the concept more formally in the management of parks and the outdoor recreation field. In this decade we also witnessed the first rigorous scientific application of carrying capacity to parks: a monograph by Wagar expanded the dominant emphasis on environmental concerns to a dual focus including social and experiential considerations. He argued that as more people visit a park or similar outdoor recreation area, not only are the environmental resources of the area affected, but also the quality of the recreation experience. Wagar illustrated the effects of increasing use on recreation quality by means of relationships between use level and visitor satisfaction.

As an example, a preliminary attempt to estimate the recreation carrying capacity of the Boundary Waters Canoe Area, Minnesota, followed shortly, and researchers found that perceptions of crowding varied by different user groups. Paddling canoeists were found to be more sensitive to crowding than any other visitors on boats. A range of carrying capacities was estimated depending upon these different relationships.

Wagar's original conceptual analysis hinted at a third element of carrying capacity, noting how carrying capacity might vary according to the amount and type of management. For example, the durability of park resources might be increased through practices such as fertilizing and irrigating vegetation and periodic rest and rotation of impact sites. Similarly, the quality of the recreation experience might be enhanced in the face of increasing use by means of more even distribution of visitors,

as well as appropriate rules and regulations, additional visitor facilities, and educational programs designed to encourage desirable user behavior. Thus carrying capacity, as applied to parks and related areas, has been expanded to a three-dimensional concept by the addition of management consideration, as illustrated in Figure 5.

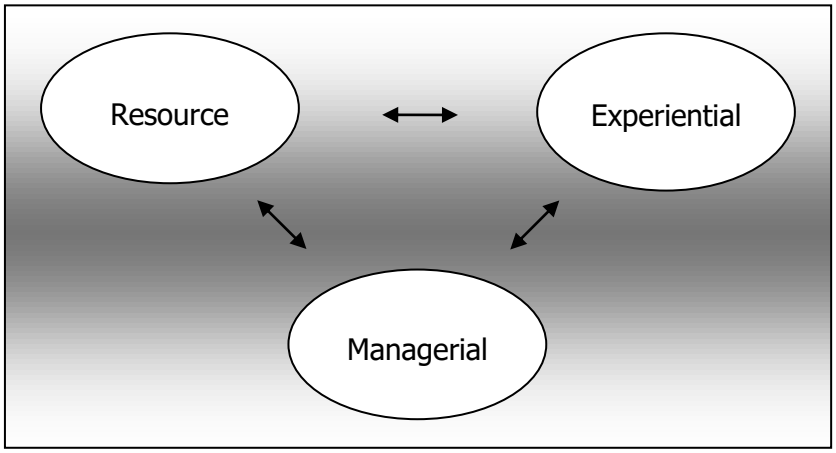


Figure 5. Three dimensions of carrying capacity

Carrying capacity has attracted much focus as a research and management concept in parks and outdoor recreation. Several books and papers have been published on the issue but despite the impressive literature base, efforts to apply carrying capacity in the field have had little success. The main difficulty lies in determining how much impact should be allowed within the three components that make up the carrying capacity concept: environmental resources, quality of recreation experience, and extent and type of management actions.

The growing research base in parks and outdoor recreation indicates that increasing recreation use often causes impact or change, especially with regard to park resources. The ecological impacts of outdoor recreation can be extensive and wide ranging, including soil erosion, trampling of vegetation, water pollution, soundscapes disruption and disturbance of wildlife, as reported in many studies. Similarly, social science research has documented impacts of increasing visitor use on the quality of the recreation experience through crowding, conflict, and the aesthetic implications of resource degradation. Finally, research suggests that increasing recreation use can change the management environment through development and implementation of more intensive management practices. Despite increasing knowledge about park use and resulting impacts, the critical question remains: how much change should be allowed?

This issue is often referred to as the "limits of acceptable change": with increasing use of parks and related areas, some change in the recreation environment – park resources, the visitor experience, the management context – is inevitable. Yet sooner or later the amount of change may become unacceptable. What determines the limits of acceptable change?

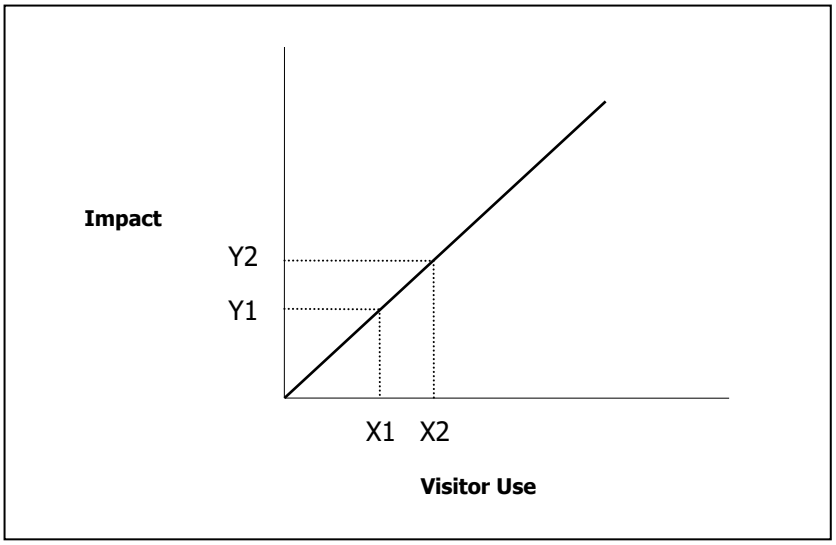


Figure 6. Hypothetical relationship between visitor use and impact to parks and related areas

Figure 6 graphically shows a possible relationship between visitor use and impacts to the resource, experience, and management components of parks. This relationship suggests that increasing recreation use can cause increasing impacts in the form of damage to fragile soils and vegetation, crowding and conflicting uses, and more intensive recreation management actions. However it is not clear when carrying capacity has been reached. Clearly, X1 and X2 are two alternative levels of visitor use which result in corresponding levels Y1 and Y2 of impact, but which of these two points (or maybe another along the vertical axis) represents the maximum amount of acceptable impact?

To further clarify this issue, some studies have suggested distinguishing between descriptive and evaluative (or prescriptive) components of carrying capacity. The descriptive component focuses on factual, objective data, such as the relationships shown in the graph above. The evaluative component concerns the more subjective issue of how much impact or change is acceptable.

2.2.1 Management Objectives, Indicators and Standards

Recent experience with carrying capacity suggests that answers to evaluative/prescriptive questions can be found through formulation of management objectives (or desired conditions) and associated indicators and standards. This

approach focuses on defining the level of resource protection to be maintained and the type of visitor experience to be provided. Management objectives are broad, narrative statements defining the type and quality of park conditions. Indicators are more specific, measurable variables reflecting the essence of management objectives, and are quantifiable measures. Indicators may include elements of the resource, experiential and management environments that are important in determining park conditions. Standards define the minimum acceptable condition of indicator variables.

An example may help illuminate these ideas. Review of the Wilderness Act of 1964 suggests that areas to be preserved, designated by the Congress, are to be managed to provide opportunities for visitor "solitude". So providing opportunities for solitude is an appropriate desired condition for most wilderness areas, but solitude is a somewhat abstract concept that is difficult to measure. Research on wilderness use suggests that the number of other visitors encountered is important in defining solitude. Thus, trail and camp encounters are potentially good indicators because they are measurable and serve as a proxy for the objective of wilderness solitude. Research also shows that visitors may have standards about how many trail and camp encounters can be experienced before opportunities for solitude decline to an unacceptable degree (for example, some studies suggest that no more than five groups per day encountered along trails are acceptable). Therefore, a maximum number of encounters per day may be good standards for managing the carrying capacity of wilderness areas.

Management objectives, and associated indicators and standards, should be formulated bearing in mind some considerations, which can be organized into three broad categories:

- Resource: the ecological characteristics of the natural resource help determine the degree of change in the environment that results from recreation use. Resource characteristics should be studied and may become important guides in formulating management objectives
- Experiential: needs and desires of society are important in determining appropriate park and outdoor recreation opportunities. Studies of visitors to outdoor recreation areas may suggest appropriate types and levels of recreation use and associated impacts.
- Managerial: legal directives, mission statements and other policy-related guidelines may suggest management objectives and related indicators. Financial, personnel and other management resources may suggest the types and levels of park and recreation use that are feasible.

The information above is important to formulate objectives, but there is also a value-based element of park and recreation carrying capacity that must be addressed. Research can illuminate the relationships between increasing use levels and change in the recreation environment (the descriptive component of carrying capacity) as illustrated in Figure 6. Moreover, research on the standards of park visitors and other stakeholders can help inform the prescriptive component of carrying capacity. Some element of management judgment will be needed to integrate resource, experiential and managerial components of carrying capacity into informed management objectives; therefore several frameworks have been developed to help guide this process.

2.2.2 Carrying Capacity Frameworks

An operational definition of carrying capacity, along with several frameworks for analyzing and applying it to parks, has been defined over the years. Examples of frameworks include Limits of Acceptable Change (LAC), Visitor Impact Management (VIM), Visitor Experience and Resource Protection (VERP). All of these frameworks provide a rational, structured process for conducting carrying capacity analysis and management. While terminology and sequencing may vary, these frameworks share a common underlying logic. Core elements include:

1. Definition of park conditions to be maintained. These conditions should be defined in terms of management objectives and associated indicators, and should address the resource, experiential and managerial components of parks.
2. Monitoring of indicator variables to determine if existing park conditions meet the specified standards.
3. Application of management practices to ensure standards are maintained.

Recreation-related carrying capacity includes resource, experiential and managerial considerations, descriptive and evaluative components, desired conditions and associated indicators and standards. Therefore there is no one carrying capacity for a park. Rather, it depends upon how the components of the concept are fashioned together. This complexity has caused some disillusionment, with characterizations such as "slippery", "elusive" and "illusive" in literature. According to surveys amongst park and wilderness managers, even though they suspect that recreational use of their areas has exceeded carrying capacity, managers have not yet established proper policies.

Some authors, noting weaknesses and shortcomings of carrying capacity, point out that the term may imply a single "magic number" for each recreation area, and that this is misleading and obscures the role of management judgment.

Others have argued that the very term *carrying capacity* seems to imply an undue emphasis on use limitations: management practices might be used to meet objectives aside from use limitations. While management objectives for some areas may well set relatively low carrying capacities and thus ultimately require use limits, other areas will properly have relatively high carrying capacities and may not require use limits.

Finally, even the author of the original conceptual analysis of recreation carrying capacity has suggested that borrowing the term from range and wildlife management may not have been a wise choice. The close association between carrying capacity and resource and ecological considerations tends to divert attention from the equally important experiential and managerial concerns which must be a part of carrying capacity as applied to parks and outdoor recreation.

All of these points are valid criticisms. However the term carrying capacity is now deeply entrenched in the field of parks and outdoor recreation (and in environmental management more broadly), and recent legislation and institutional directives have made it a formal part of park management. More important, carrying capacity represents a vital issue of growing urgency, and a specific manifestation of the tragedy of the commons, therefore requiring informed and explicit management action.

Carrying capacity can be useful as an outdoor recreation concept when viewed in proper perspective – as an organizing framework for analyzing, defining and managing appropriate park conditions. The carrying capacity frameworks developed in the literature and their successful application in the field prove that it is a powerful concept for managing parks and related areas.

2.3 Indicators and Standards

Indicators are measurable, manageable variables that help define the quality of parks and outdoor recreation areas. Standards define the minimum acceptable condition of indicator variables. Carrying capacity can be managed by monitoring indicator variables and implementing management actions to ensure standards are maintained. This chapter discusses indicators and standards in more detail, describing desirable characteristics and providing examples for both.

2.3.1 Characteristics of Good Indicators

Several studies explored what defines a good indicator, which helps further understand the role of indicators and standards, and assist in evaluation and selection of potential indicator variables. Characteristics of good indicators include the following:

- Specific: indicators should define specific rather than general conditions. For example, "solitude" is too general, while "the number of other groups encountered per day along trails" would be a better indicator variable.
- Objective: indicators should be objective rather than subjective. That is, indicator variables should be measured in absolute, unequivocal terms. Variables that are subjective, expressed in relative terms, or subject to interpretation make poor indicators. "The number of people at one time at Delicate Arch" is a good choice because it refers to an absolute number that can be readily counted. However "the percentage of visitors who feel crowded at Delicate Arch" is a subjective indicator as it is subject to interpretation by visitors (e.g. it depends on the types of visitors making the judgment).
- Reliable and repeatable: the indicator measurement must yield similar results under similar conditions. This criterion is particularly important when monitoring of indicator variables is conducted by more than one person. Monitoring should take place at regular intervals over a long period of time.
- Related to visitor use: indicators should be related to at least one of the following attributes of visitor use: level of use, type of use, location of use, or behavior of use. A major role of indicators is to help determine when management action is needed to control the impacts of visitor use.
- Sensitive: indicators should be sensitive to visitor use over a relatively short period of time. If an indicator changes only after impacts are substantial, it will not serve as an early warning mechanism, allowing managers to react in a timely manner.
- Manageable: indicators should be responsive to, and help determine the effectiveness of, management actions. Indicators should be maintained within prescribed standards, so they must be manageable.
- Efficient and effective to measure: indicators should be easy and cost-effective to measure. Indicators must be monitored regularly, therefore the more expertise, time, equipment and staff needed to take such measurements, the less desirable a potential indicator may be.

- Integrative or synthetic: potentially there are many desired conditions to be achieved in parks. As noted in the previous chapter, these might apply to park resources, quality of visitor experience, and type and level of management. Yet it is impractical to monitor large numbers of indicator variables. Therefore synthetic variables (those which are proxies for more than one component of protected areas) are especially useful. For example an indicator of the level of visitor use may be useful as a measure of crowding and associated resource and social impacts.
- Significant: perhaps the most important characteristic of indicators is that they help define the quality of park resources and visitor experience. It does little good to monitor the condition of a variable irrelevant for such issues.

It may be useful to incorporate these characteristics within a matrix, as shown below, to evaluate potential indicators. Potential indicators can then be rated as to how well they meet those characteristics: indicators that receive the highest aggregate ratings may have the greatest value in measuring and managing carrying capacity.

Potential Indicators	Criteria for good indicators							
	Specific	Objective	Reliable and repeatable	Related to visitor use	Sensitive	Manageable	Efficient and effective to measure	Significant
Indicator 1								
Indicator 2								
Indicator 3								
Indicator …								

2.3.2 Potential Indicator Variables

A body of research has focused on identifying potential indicators for a variety of recreation areas and activities. The aim is to determine variables important to visitors the quality of park resources and the recreation experience. There exist a plethora of examples of notable indicators, depending on the characteristics of a resource.

These studies have addressed a variety of recreation areas and activities and utilized several study methods open- and closed-ended questions and surveys of visitors, interest groups, and scientists. Several general conclusions might be derived.

First, potential indicators appear to be wide ranging. It may be useful to employ the three-fold framework of carrying capacity (resource, experiential and managerial) described in the previous chapter when thinking of potential indicators.

Second, most studies have found some indicator variables to be more important than others. For example, litter and other signs of use impacts appear to be universally important. Level of visitor use appears important too, but how it is manifested may be even more significant. For example, the type of visitors encountered (e.g., hikers encountering bikers, floaters encountering motor boaters) may be just as relevant as the number of encounters. In other park contexts, the impacts of level of use may be manifested, for example, in terms of waiting times or completion for access.

Third, visitors to wilderness may be generally more sensitive to a variety of potential indicators than visitors to more highly used areas, but research may have not yet identified which indicators possess this characteristic.

2.3.3 Characteristics of Good Standards

Several studies have explored what defines good standards, and these are the characteristics that should be incorporated:

- Quantitative: since indicators are specific and measurable variables, standards should be expressed in an unequivocal, quantitative manner. For example, if an indicator is "the number of encounters with other groups per day along the river", then the standard might be "an average of no more than three encounters". In contrast, "low numbers of encounters with other groups per day along the river" would be a poor standard because it does not specify the minimum acceptable condition.
- Time- and space-bounded: a time- and space-bounded element in the standard expresses how much of an impact is acceptable and how often or where such impacts can occur. It is often desirable for standards to have a time period associated with them, which is particularly relevant for crowding-related issues.

For examples, in the above example, the standard for encounters with other groups was expressed in terms of "per day". Other qualifiers might be "per night", "per trip", "per hour", "at one time", depending on circumstances.

- Expressed as a probability: often it is advantageous to include a tolerance for some percentage of the time that the desired condition may not be met. For example, a standard may specify that hikers will have "no more than three encounters with other groups per day along trails for 90% of the days in the summer season". The 90% probability of conditions meeting or exceeding the standard allows for 10% of the time that unusual events might prevent management from maintaining these conditions. This allows for the complexity and randomness inherent in park-use patterns. In the example of encounters along a trail, on peak-use days several hiking parties might depart from a trailhead at closely spaced intervals, therefore increasing the likelihood to encounter each other several times during the day. So it might be wise to incorporate a tolerance in standards for holiday weekends or other days of exceptionally high visitation. The amount of tolerance needed depends on the unpredictability of each individual situation and the degree to which management can consistently control conditions.
- Impact-oriented: standards should focus on the impacts that affect the quality of park resources and the visitor experience, not the management action employed to keep impacts from violating standards. For example, an appropriate standard might be "no more than ten encounters with other groups along the river per day". This standard focuses direct on the impact that affects the quality of visitor experience. Alternatively, "a maximum of twenty groups per day floating the river" would not be as good a standard because it does not focus as directly on the impact of concern; visitors experience encounters with other groups more directly than they experience total use levels! Basin standards on management actions rather than on impacts can also limit consideration of the potential range of useful management practices. For example, limiting the number of boats to twenty per day might ensure fewer encounters per day, but other actions could also ensure an acceptable encounter rate and could be less restrictive on the level of visitation.
- Realistic: standards should generally reflect conditions that are realistically attainable. Standards that limit impacts to extremely low levels may set up

unrealistic expectations in the minds of visitors, may be infeasible to maintain, and may unfairly restrict visitor use to very low levels.

2.3.4 Potential Standards

An increasing number of studies have been conducted to help define standards for parks and related areas. These studies have addressed a variety of park and outdoor recreation areas and potential indicators. They have also used alternative question formats and wording, different response scales, and other methodological variations. Numerous potential standards are listed in literature, depending on the type of resource being monitored. Several general conclusions might be derived from this growing body of literature.

First, standards can be measured for a variety of potential indicators. While many studies have addressed encounter and crowding-related indicator variables, other studies have measured standards for widely ranging variables which are capable of representing all three of the components of carrying capacity (resource, experiential and managerial).

Second, visitors tend to report standards more often in wilderness situations than in more developed areas. Also, there tends to be more agreement about wilderness-related standards (this issue is often called *crystallization*). For example, standard deviations of encounter standards for floaters on three rivers were found to increase as the recreation opportunity described moved from "wilderness" to "semi-wilderness" to "undeveloped recreation".

Third, standards tend to be less tolerant in wilderness than in more developed areas. For example, visitors to wilderness areas tend to want to camp out of sight and sound of other groups while visitors to developed campgrounds can tolerate relatively large numbers of other groups.

Fourth, there may be some consistency in standards within similar types of outdoor recreation areas. For instance, a study of visitor standards for a variety of potential indicators found broad agreement across the four geographically diverse wilderness areas included in the study. Moreover, some studies suggest that standards for encountering other groups along trails during a wilderness experience are quite low and that many wilderness visitors prefer to camp out of sight and sound of others.

Fifth, standards of visitors can vary from those of managers. For example, a study of standards for wilderness campsite impacts found that visitors reported more restrictive standards regarding the presence of fire rings than did managers.

2.4 Alternative Management Practices

The literature on management of parks and outdoor recreation has identified a range of management practices that might be applied to issues such as crowding, conflict or environmental degradation. It is useful to classify these practices, in order to illustrate the broad spectrum of alternatives available.

2.4.1 Management Strategies

One classification system defines alternatives on the basis of management strategies. Management strategies are basic conceptual approaches to management, related to achievement of desirable objectives. Four main strategies can be identified for managing outdoor recreation: increase supply, reduce impact of use, limit use, increase durability of resource. Two strategies deal with supply and demand: the supply of recreation opportunities may be increased to accommodate more use, or the demand for recreation may be limited through restrictions. The other two basic strategies treat supply and demand as fixed and focus on modifying either the character of recreational use to reduce its adverse impacts or the resource base to increase its durability.

Within each basic management strategy, there are a number of sub-strategies. For example, the supply of protected areas can be increased in terms of both space (e.g., adding new areas) and time (e.g., shifting timetables). Within the strategy of limiting demand, restrictions might be placed on the total number of visitors allowed or on their length of stay.

The third strategy suggests reducing the social or environmental impacts of existing use. This might be accomplished by modifying the type or character of use, or by dispersing or concentrating use.

A final management strategy involves increasing the durability of park resources. This might be accomplished by hardening the resource itself through intensive maintenance, or development of facilities to accommodate use more directly.

2.4.2 Management Tactics

A second system of classifying management alternatives focuses on on-the-ground management practices, which are direct actions to accomplish management strategies. Restrictions on length of stay, differential fees, and use permits, for example, are management practices designed to accomplish the strategy of limiting recreation demand. These actions can be classified as direct management practices (when they act directly on visitor behavior, leaving little or no freedom of choice) and indirect management practices (when they attempt to influence the decision factors upon

which visitors base their behavior). A conceptual diagram illustrating direct and indirect park and recreation management practices is shown in Figure 7

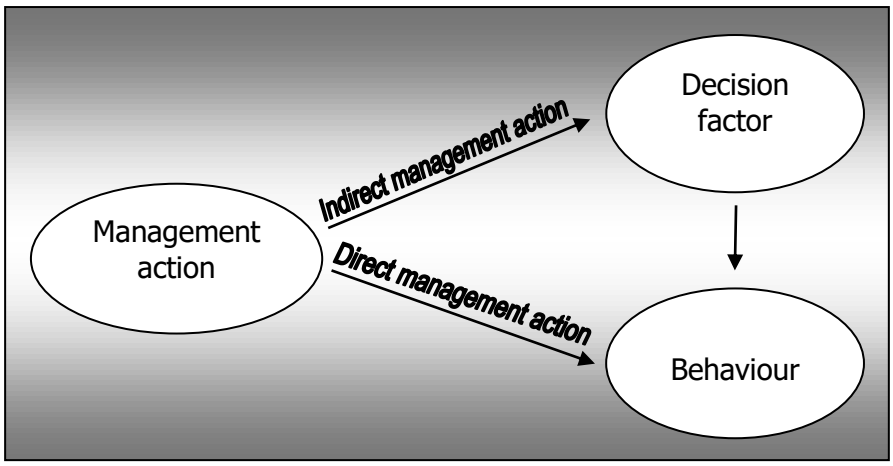


Figure 7. Diagram of direct vs indirect management tactics

As an example, a direct management practice aimed at reducing campfires in a wilderness environment would be a regulation barring campfires. An indirect management practice would be an education program designed to inform visitors of the undesirable ecological and aesthetic impacts of campfires and to encourage them to carry and use portable stoves instead.

Relative advantages and disadvantages of direct and indirect management practices have been evaluated in the literature. Indirect practices are generally favoured when they are believed to be effective, but also for other reasons. First, legislation and management agency policies applied to wilderness often emphasize provision of "unconfined" visitor opportunities. Thus, direct regulation of visitor behavior may be inconsistent with such objectives. Second, recreation is a leisure activity with freedom of choice in thought and actions, therefore regulations designed to control visitor behavior can be seen as antithetical to the very nature of recreation. Third, many studies reveal that, given the choice, visitors prefer indirect management practices. Finally, indirect practices may be more efficient because they do not entail the costs associated with enforcement of rules and regulations.

Emphasis on indirect management practices has not been uniformly endorsed, since some authors believe they may be ineffective. Some visitors, for example, will ignore management efforts to influence their behavior, and the action of a few may hamper the attainment of management objectives. Some have argued that a direct regulatory approach to management can ultimately lead to more freedom rather than less: when

all visitors are required to conform to mutually agreed-upon behavior, management objectives are more likely to be attained and a diversity of park and recreation opportunities preserved. Empirical evidence suggests that, under certain circumstances, direct management practices can enhance the quality of the visitor experience. Sometimes visitors are surprisingly supportive of direct management practices when they are needed to control the impacts of recreation use.

An analysis of management problems caused by visitors suggests that both direct and indirect management practices can be applicable. There are several reasons why visitors may not conform to standards of behavior, reasons ranging from lack of knowledge about appropriate behavior to willful rule violations. Indirect measures, such as education programs, seem most appropriate in the case of the former, while direct practices are usually needed in the case of the latter.

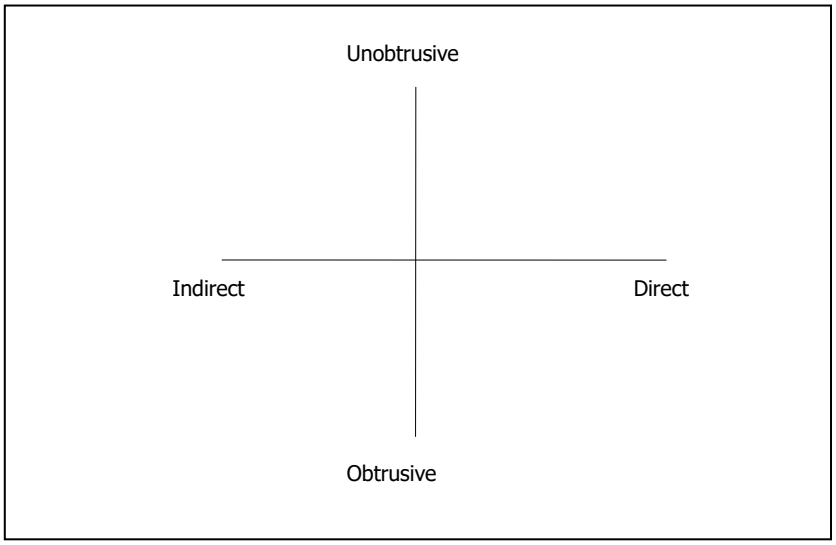


Figure 8. Two dimensions of park and recreation management practices

Some suggested there is a continuum of management practices that range from indirect to direct. For example, an educational program on the ecological and aesthetic impacts of campfires would be found toward the indirect end of a continuum. On the other hand, aggressive enforcement of a regulation with uniformed rangers would clearly be a very direct management practice. Not only can management practices be direct or indirect, they can also be implemented in an obtrusive or unobtrusive manner. It has also been suggested that direct and indirect measures are not mutually exclusive but can, in fact, often complement each other.

Some authors have classified the main recreation management practices, outlining basic strategies such as: reduce use of an area, modify the timing of use, modify type

of use and visitor behavior, or increase the resistance of the resource. Several tactics are available for each strategy.

2.5 Evaluating the Effectiveness of Management Practices

Given the vital role of management in contemporary carrying capacity frameworks, it is important to test the potential effectiveness of alternative management practices. A growing body of literature has focused on evaluation of selected management practices, including visitor information, education programs, use rationing and allocation, and other widely used measures.

2.5.1 Information and Education

Information/education is seen as an indirect, light-handed management tool, designed to persuade visitors to adopt behaviors compatible with management objectives without regulating visitors directly. Research shows that it is effective, and a set of principles for its application is emerging.

Problem behaviors of park visitors can be classified into 5 basic types, and the effectiveness of information/education on each varies considerably. At the two ends of the spectrum, problem behaviors can be seen as either deliberately illegal or unavoidable, and in these instances information/education has limited effectiveness. The other three types of problem behavior (careless actions, unskilled actions and uninformed actions) may be a lot more amenable to information/education programs.

Another approach refers to the "mindfulness" or "mindlessness" of visitors. A mindful visitor processes new information, and consciously thinks about appropriate ways to behave. Therefore strategies to enhance mindfulness can facilitate learning and better decision making.

Another conceptual approach to the application of information/education suggests that people progress through stages of moral development, ranging from being very self-centered to highly altruistic. Since a park visitor may be at any of the stages of moral development, information/education should be designed to reach any such stage. For example, to reach visitors who tend to be selfish, managers might emphasize rewards or punishments for selected types of behavior. However, communicating with visitors at higher levels of moral development might be more effective by appealing to a sense of altruism, justice and fairness.

Also, communication theory suggests that the potential effectiveness of information/education depends upon variables associated with the content and delivery

of messages to visitors. For example, visitor behavior is at least partially driven by attitudes and believes, so information/education aimed at connecting with or modifying relevant attitudes, beliefs and norms may be successful in guiding or changing visitor behavior. Moreover, the media by which messages are delivered may also influence the effectiveness of information/education programs.

Finally, from a theoretical standpoint, information/education can be seen to operate through three basic models. *Applied behavior analysis* focuses directly on visitor behavior. For example, visitors can be informed of rewards or punishments administered, dependent upon their behavior. This model does not address underlying cognitive or behavioral variables, so its effectiveness may be short term. In the *central route to persuasion*, relevant beliefs of visitors are modified through substantive messages. Modified beliefs then lead to desired changes in behavior. The *peripheral route to persuasion* model emphasizes nonsubstantive elements of information/education messages, such as message source and medium. For example, messages from sources considered to be authoritative by visitors may be more influential. This model appears to be especially useful where it is difficult to maintain visitors' attention, such as at visitor centers, entrance stations, and bulletin boards. However, this model may not influence antecedent conditions of behavior, and therefore may not have lasting effects.

Empirical studies have examined the effectiveness of a variety of park and recreation-related information/education programs. These can be described as (1) studies to influence visitor-use patterns; (2) studies focused on enhancing visitor knowledge; (3) studies aimed at influencing visitor attitudes toward management policies; and (4) studies that address depreciative behavior, such as littering.

Visitor-use patterns in parks and related areas often have uneven spatial and temporal distribution. Visitor-caused impacts may be reduced if use patterns could be changed: early studies explored the effectiveness of providing visitors with information on current use patterns as a way to alter future use patterns. For example, visitors with permits for heavily used entry points were mailed an information packet, noting in particular heavily used areas and times. A survey of a sample of this group who again visited the area found that most respondents found the information useful, and about one-third were influenced in their choices during subsequent visits. In another study, a brochure explained resource impacts associated with concentrated camping and showed the location of other nearby camping areas: a group which was given the brochure, dispersed their camping activity.

Prior to visiting, a group of hikers in Yellowstone National Park was given a guidebook that described the attributes of lesser-used trails: over one-third of the group selected one of the lesser-used trails. Results also show that the earlier the information is received, the more influence it had on behavior. Moreover, employing computer-based approaches (e.g. touch-screen programs) have been found to be effective.

Wilderness rangers are also used as a source of information/education, but a study showed that only 20% of visitors reported that the information they received from rangers influenced their destination. However, less experienced visitors admitted they were more likely to be influenced when returning to the study area.

Another study in a wilderness area in Montana illustrated potential problems in using information/education to influence visitor use: brochures describing current recreation-use patterns were distributed. There was little effect on subsequent use patterns, possibly due to three limitations on this program effectiveness: many visitors did not receive the brochure; most of those who did receive it, got it too late to affect their choice; some visitors doubted the accuracy of the information provided.

A second category of studies has focused primarily on enhancing visitor knowledge to reduce ecological and social impacts. For example in Rocky Mountain National Park, information was provided on low-impact camping practices through a series of media (slide/sound exhibits, brochures and trailhead signs); exposure to a trailhead sign and brochure was not found to be very effective.

Another study on day hikers to subalpine meadows near Washington showed the effectiveness of providing a short, personal program on reasons for complying with guidelines for off-trail hiking. Most of those visitors who received the interpretative program complied with the guidelines illustrated.

A study of day hikers to a national park in Arizona, found that an aggressive campaign featuring the message "heat kills, hike smart" presented in the park newspaper and on trailhead posters, influenced safety-related hiking practices (carrying sufficient water, starting hikes early in the day, etc.). Bulletin boards have been found to be effective in enhancing visitor knowledge, however increasing the number of messages posted beyond two had little or no effect.

Workshops and special programs delivered to organizations also can be effective in influencing intentions to follow low-impact practices. For example, Leave No Trace (LNT) is a public/private national educational initiative that integrates outdoor recreation research into park and outdoor recreation education. LNT establishes a collaborative framework connecting managers and researchers and providing visitors

with current minimum-impact skills and information. The effectiveness of these programs has been demonstrated in several studies, as well as training commercial guides and outfitters to deliver education programs to clients, or using trail guide booklets.

Not all research has found information/education programs to be as effective as indicated in the previous studies. There was no difference in knowledge about general backcountry policies between visitors exposed to the park's interpretative services and those who were not, although the former group scored higher on knowledge of park-related hazards. Visitors requesting information on wilderness permits for an area in Minnesota, were mailed specific brochures, but only a very small minority reported altering their actual or intended behavior.

A third category of studies has examined visitor attitudes towards a variety of recreation management policies, and has found that information/education can be effective in modifying their behavior. For example, visitors to Yellowstone National Park were exposed to interpretative messages about fire ecology and the effects of controlled-burn policies: these messages influenced both beliefs about these issues, and attitudes based on those beliefs.

A fourth category of studies has focused on depreciative behavior, especially littering: information/education has been found effective in reducing littering and even cleaning up littered areas. For example, samples of visitors to a developed campground were given three different treatments: a brochure describing the costs and impacts of littering, the brochure plus personal contact with a park ranger, and these two treatments along with a request for assistance in reporting depreciative behavior to park rangers. The brochure plus the personal contact was the most effective treatment, with a significant reduction in the number of groups littering or damaging natural resources. Types of messages and related purposes found to be effective in several studies include incentives to visitors to assist with cleanup efforts, and the use of rangers as role models for cleaning up litter.

Several other types of studies, while not evaluating the effectiveness of information/education, also suggest its potential for park and recreation management. First, studies of visitor knowledge through questionnaires indicate that marked improvements are possible, occasionally leading to improved visitor behavior. It should be noted that, on a few occasions, there were significant differences among types of respondents, types of knowledge, and the accuracy of various sources of information,

providing indications of where and how information/education should be channeled most effectively.

Second, several studies indicate that information/education programs could be substantially improved. Evaluation of literature mailed in response to visitor requests has identified several areas of needed improvements, including more timely response, more focus on management issues, greater personalization, more visual appeal, and reduction of superfluous materials.

Third, a survey of wilderness managers identified the extent to which different visitor information/education practices were used: only six of them (brochures, personnel at agency offices, maps, signs, personnel in the backcountry, and displays at trailheads) were used in a majority of wilderness areas. Managers were also asked to rate the perceived effectiveness of such practices, noting that personnel-based practices are generally considered to be more effective than media-based ones.

Finally, several studies have examined the sources of information/education used by park and recreation visitors for trip planning: many respondents report using sources not directly produced by management agencies (such as outdoor clubs, professional outfitters, outdoor stores, guidebooks, travel agents). This suggests that linkages with selected private and commercial organizations may be an effective approach.

Despite the fact that studies are diverse in terms of geographic areas, methods and issues addressed, a number of principles for using information/education are emerging from literature:

- Information/education programs may be most effective when applied to problem behaviors characterized by careless, unskilled or uninformed actions.
- Information/education programs should be designed to reach visitors at multiple stages of moral development.
- Information/education programs designed to "connect" with or modify visitor attitudes and beliefs, are likely to be most effective in the long term.
- Using multiple media to deliver messages can be more effective than using a single medium.
- Brochures, personal messages and audiovisual programs may be more effective than signs.
- Messages may be more effective when delivered early in the visitor experience, such as during trip planning.
- Messages from sources judged highly credible may be especially effective.

- Strongly worded messages and aggressive delivery can be an effective way of enhancing "mindfulness" of visitors and may be justified when applied to issues such as visitor safety and protection of critical or sensitive resources.
- Computer-based information systems can be an effective means.
- Training of volunteers, outfitters and commercial guides can be an effective and efficient way of communicating information/education.
- Non-agency media (such as magazines and guidebooks) can be an effective and efficient way of communicating information/education.
- Information on impacts, costs and consequences can be an effective information/education strategy.
- Role modeling by rangers and volunteers can be an effective strategy.
- Personal contact with visitors by rangers and other employees can be effective in communicating information/education.
- Messages should be targeted to specific audiences: receptive target audiences include those who request information in advance, and those who are least knowledgeable.
- Messages should be targeted at issues that are least known or well understood by visitors.

Studies on information/education suggest that it can be an effective and desirable tool, and should employ a variety of message types and media, addressing a variety of management issues and target audiences.

2.5.2 Use Rationing and Allocation

Substantial attention has been focused on the strategy of limiting the use that parks and protected areas receive. Use rationing is often considered a management approach of "last resort", because it runs counter to the basic objective of granting public access. However, limits on use are surely needed at some times, to protect the integrity of critical resources and to maintain the quality of the recreation experience.

Five basic management practices have been identified to ration recreation use:

1. reservation systems, where visitors are to reserve a space or permit in advance of their visit;
2. lotteries, which require visitors to request a permit in advance but allocate permits on a purely random basis;

3. first-come, first-served, or queuing, which requires potential visitors to "wait in line" for available permits;
4. pricing, whereby visitors are to pay a fee for a permit, which may "filter out" those unwilling or unable to pay;
5. merit, a system which requires potential visitors to "earn" the right to a permit by virtue of demonstrated knowledge or skill.

Each of these management practices has potential advantages and disadvantages. For example, reservation systems may favor visitors who are willing to plan ahead, but these systems may be difficult to administer. Lotteries are often viewed as "fair", but hey can be cumbersome to administer. First-come, first-served systems are relatively easy to administer, but may favor those who have more leisure time. Pricing is a commonly used practice to allocate scarce resources, but may discriminate against visitors with low incomes. Merit systems are rarely used but may lessen environmental and social impacts of use.

Several guidelines have been suggested for considering these practices. First, emphasis should be placed on the environmental and social impacts of recreation use, since some types of recreation use may cause more impacts than others. To the extent that such impacts can be reduced, rationing might be avoided, or postponed. Second, as noted, rationing use should be considered a practice of last resort: less "heavy-handed" management practices would seem more desirable when they are proven effective. Third, good information is paramount to implement these practices: managers must be certain that environmental problems dictate use rationing, and that visitors are understood well enough to predict the effects of alternative allocation systems. Fourth, combinations of use-rationing systems should be considered. Given pros and cons of each practice, hybrid systems may have special application. For example, half of the permits might be allocated with a reservation system, and half on a first-come, first-served basis. This would serve the needs of those who can plan vacations in advance, as well as those whose lifestyles do not allow for this. Fifth, use rationing should establish a linkage between the probability of obtaining a permit and the value of the recreation opportunity to potential visitors. In other words, visitors who value the opportunity highly should have a chance to "earn" a permit through pricing, advance planning, waiting time or merit. Finally, use-rationing practices should be monitored and evaluated to assess effectiveness and fairness: they are relatively new for parks and related areas, so they are likely to be controversial.

A critical element of use-rationing and allocation practices is fairness. Recreation areas administered by federal and state agencies are public resources, therefore allocation practices are expected to be both efficient and equitable. But how are concepts like equity and fairness defined? Several studies have developed important insights into this issue, outlining different dimensions.

One study identified four dimensions of an overall theory of *distributive justice*, defined as an ideal whereby individuals obtain what they "ought" to have. A first dimension is *equality*, which suggests that all individuals have an equal right to a benefit. A second dimension is *equity*, which suggests that benefits be distributed to those who "earn" them through some investment of time, money or effort. A third dimension is *need*, which suggests that benefits be distributed on the basis of unmet needs. A final dimension is *efficiency*, which suggests that benefits be distributed to those who place the highest value upon them.

Insights into these dimensions of distributive justice were developed in a survey of river runners in Idaho: they were asked to rate the five use-allocation practices on the basis of four criteria: perceived chance of obtaining a permit, perceived fairness of the practice, acceptability of the practice, and willingness to try the practice. Results suggest that visitors evaluate use-rationing practices using fairness and pragmatism (that is, the perceived ability on the part of the respondent to obtain a permit), the latter having the strongest effect on willingness to try a practice. These findings suggest that managers have to convince potential visitors that proposed use-allocation practices are not only fair but they will provide them with a reasonable chance to obtain access.

A second series of studies has examined an extended taxonomy of equity dimensions, applicable to a broad spectrum of park and recreation opportunities. Eight potential dimensions of equity are identified. A first dimension is compensatory and allocates benefits on the basis of economic disadvantage. The second two dimensions are variations of equality (equal benefits; equal impacts) and allocate benefits to all individuals equally. The fourth and fifth dimensions are based on demand, and allocate benefits to those who make greatest use of them (demonstrated use) or those who advocate most effectively for them (advocacy). The final three dimensions are market driven and distribute benefits based on amount of taxes paid, the price charged for services, or the least-cost alternative for providing recreation services. A sample of California residents were asked to indicate the extent to which they agreed or disagreed with each dimension of equity as a principle for allocating public park

services. A majority of the sample agreed with only three of the dimensions, which were (in decreasing order) demonstrated use, price paid and equal benefits.

Despite the complex and controversial nature of use rationing, there appears to be considerable support for many management practices among outdoor recreation visitors. Even most individuals who have been unsuccessful at obtaining a permit continue to support the need for use rationing. A study of visitors to wilderness areas in Oregon found that support for use restrictions was based on concerns for protecting both resource quality and the quality of visitor experience (in particular, there showed concern with crowding and environmental impacts).

Preferences among alternative use-rationing practices have been found to be highly variable, based on both location and type of user. Support for a particular practice appears to be related primarily to which practices respondents are familiar with, and the extent to which they believe they can obtain access.

In keeping with the generally favorable attitude toward use limitation, most studies have found visitor compliance rates for mandatory permits to be high, with most areas in the 90% range. Also, permits that have incorporated trailhead quotas have been found to be effective in redistributing use both spatially and temporally.

A practice that has received special attention in literature is pricing, which is the primary means of allocating scarce resources in a free market economy. Economic theory suggests that higher prices will result in less consumption of a given good or service, so pricing may be an effective approach to limiting use of outdoor areas. However, recreation services in the public sector traditionally have been priced at a nominal level or even provided free of charge. The basic philosophy underlying this policy is that access to parks is important to all people, and no one should be "priced out".

Studies on pricing have tended to focus on several issues related to its potential as a management practice. First, to what extent does pricing influence use of parks and related areas? Several studies have found an inverse relationship between price and use: some studies have shown little or no effect, whilst in a research 40% of respondents reported they would no longer use these areas if a fee was instituted. The literature suggests that the influence of fees on park and recreation use is dependent upon several factors:

- The *elasticity* of demand, which refers to the slope of the demand curve defining the relationship between price and quantity consumed (or visitation). If the demand for some recreation area is relatively elastic, it means that a

change in price has a comparatively large effect on visitation, while an inelastic demand means that a change in price has small effect on visitation.

- The significance of the park or recreation area. Parks of national significance are likely to have a relatively inelastic demand, suggesting that pricing is not likely to be effective in limiting use unless price increases are dramatic.
- The percentage of total cost represented by the fee. Pricing is likely to be a more effective use-limiting approach only in cases where the fee charged represents a relatively high percentage of the total cost of visiting a park.
- The type of fee instituted. Pricing structure can be a potentially important element in determining the effectiveness of fees (for example, a daily-use fee might be more effective in limiting total use than an annual pass that allows unlimited use opportunities for a flat fee).

A second issue addressed in literature is the acceptability of fees to potential visitors: study findings often suggest that there is a substantial willingness to pay for recreation services, but the acceptability of fees is partially dependent on several factors.

- Fees are judged to be more acceptable to park visitors when revenues derived from fee programs are retained by the collecting agency and reinvested in recreation facilities and services.
- Public acceptance of new fees where none were charged before tends to be relatively low compared to increases in existing fees.
- Local visitors tend to be more resistant to new fees than nonlocal visitors, probably because fees represent a larger percentage of the total cost of visiting a recreation area for those living locally. Moreover, local residents are more likely to visit a given recreation area more than once.
- Visitor acceptance of a fee is likely to be greater when information is provided on the costs of substitute recreation opportunities, and when visitors are made aware of the costs of providing recreation opportunities.

A third issue concerns the potential for pricing to discriminate against certain groups in society, particularly those with low incomes and minority ethnic groups. Again, research is mixed, showing in some cases no discriminatory effect when an entrance fee was initiated, while in many studies lower-income visitors appeared to have a more elastic demand curve than did high-income users.

A final issue concerns the use of differential pricing to influence park and recreation use patterns. Differential pricing consists of charging higher or lower fees at selected times and locations. Research demonstrates that outdoor recreation tends to be characterized by relatively extreme "peaking", that is certain areas or times are used very heavily, while other times or areas are lightly used. Can pricing be used to even out such recreation use patterns? Research is suggestive of this potential use of pricing.

2.5.3 Other Park and Recreation Management Practices

A number of other practices are available, most of which tend to be direct management practices. Beyond information/education and limiting use, four broad categories of management practices are addressed in the literature:

1) *Rules and Regulations*: they are a commonly used practice, albeit controversial at times. Common applications in parks and outdoor recreation include group-size limitations, assigned campsites and/or travel itineraries, and length of stay limitations. The importance of encouraging visitors to comply with rules and regulations is emphasized in studies that found that visitors not complying caused extensive damage. As noted before, visitors are often unaware of rules, therefore managers should effectively communicate rules and regulations using the principles and guidelines described in the section on information and education. In particular, visitors should be informed of the reasons why rules and regulations are necessary, sanctions associated with failure to comply, and alternative activities and behaviors that can be substituted for those not allowed. The literature suggests most visitors support limitations on group size but group types should also be considered, and group-size limits should not be set so low that they affect primary groups of visitors (e.g. families or close friends). Regulations requiring the use of assigned campsites (or fixed itineraries) in wilderness or backcountry are generally not supported by visitors. Studies on the effectiveness have found that visitor compliance rates are relatively low, where noncompliance primarily meant using campsites other than those specified or staying more or fewer nights than originally agreed. Research on regulations closing selected areas to public use suggest they are supported by visitors if the underlying reason is clear and justified (e.g., ecological reasons).
2) *Law Enforcement*: Little research has been conducted on law enforcement in parks. Most of the literature discusses the controversial nature of law enforcement in this context, but one study focused on the use of uniformed rangers to deter off-trail hiking, which was found to be significantly effective. Moreover, visitors tended to

react positively to this management practice when they understood that the presence of a ranger was needed for information dissemination, visitor safety and resource protection.

3) *Zoning*: in its most generic sense, zoning simply means assigning certain recreation activities to selected areas (or restricting activities *from* areas). Zoning can be applied in a temporal dimension as well as in a spatial sense. Zoning is also widely applied as a way of creating different types of outdoor recreation opportunities: for example, "rescue" and "no-rescue" zones have been proposed for wilderness areas. Finally, zoning is used in outdoor recreation to restrict selected recreation activities from environmentally sensitive areas and to separate conflicting recreation uses.
4) *Site Design and Management*: recreation areas can be designed and manipulated to "harden" them against impacts. For example, boardwalks can be built to concentrate use in developed areas, and facilities can be constructed along trails to channel use. Campsites can also be designed to minimize social and ecological impacts. However, most of these management practices involve resource management and activities that may not be appropriate in some protected areas.

2.5.4 Status and Trends in Park and Recreation Management

Which park and recreation management practices are used most often? How effective do managers think these practices are? What are the trends in park and recreation management? Several studies offer insights into these questions, studies focusing on wilderness areas and involving periodic surveys of protected-area managers. A study in the early 90's explored recreation management practices in the national park system, asking managers to indicate which of more than one hundred recreation management practices were currently used, and which were judged most effective.

Comparisons across the studies can provide some insights into trends in park and recreation management problems and practices. Although the areas, management agencies and methods varied among these studies, their primary objectives were similar: to assess recreation management problems and/or practices in parks and recreation areas. These studies provide benchmarks at 5 different moments over a period of 25 years, from 1979 to 2004, and suggest several basic trends.

First, environmental impacts, primarily on trails and campsites, are the dominant problems perceived by managers. In all studies, managers tended to report site deterioration as the most frequently occurring recreation management problem.

Second, social problems of crowding and conflicting uses appear to have increased over time. The initial study in 1979 revealed no crowding problems, and user conflict

was cited as a problem by a minority of managers and was associated with nonconforming uses of wilderness, such as grazing by domestic livestock. More recent studies report substantial and increasing levels of crowding and conflict among recreation users.

Third, carrying capacity has become a pervasive but largely unresolved issue. The initial study in 1979 did not report carrying capacity as a significant issue. In subsequent studies, recreation use was judged to exceed carrying capacity "sometimes" or "usually", by an increasing percentage of managers. Also, despite the apparent seriousness of the carrying capacity issue, most managers have not yet addressed it adequately: for example a large portion of them admitted being unable to estimate carrying capacity for any portions of their areas, or did not base their estimates on scientific studies.

Fourth, implementation of both direct and indirect recreation management practices have tended to increase over time. For example, overnight permits for backcountry camping increased considerably over 10 years, as well as minimum-impact education programs, suggesting a shift in management practices.

Fifth, day use is an emerging issue that warrants more management attention, although in National Park Service areas the percentage of day users has remained relatively constant. Two factors exacerbate this issue: first, many management problems are attributed by managers to day users, who are held more responsible than overnight visitors; second, day users often are not targeted for management actions.

Finally, management of parks and outdoor recreation is becoming more complex and more sophisticated. This trend is reflected in the nature of the five studies examined. The original study in 1979 was primarily an exploratory study, asking managers to describe their important problems. The second study focused primarily on recreation management practices across several land-management agencies. The third study adopted several objectives, such as recreation-use patterns and recreation-related problems. The fourth study incorporated the preceding objectives and added others, like investigating the perceived causes of management problems or the effectiveness of practices. The fifth and most recent study, in 2004, monitored trends in backcountry/wilderness management and expanded this study to focus more directly on the emerging issue of day use.

Studies on alternative recreation management practices are beginning to be marshaled into handbooks and guidelines that can be used by park managers. In addition to suggesting which recreation management practices might be applied to a series of

recreation-related problems, handbooks typically offer basic information on understanding and applying each practice identified. Prototypes of computer-based "expert-systems" also are being developed to provide park and recreation managers with guidance based on the scientific and professional literature.

2.6 Indicators and Standards of Sustainability

The conceptual frameworks presented in this chapter offer approaches to analyzing and managing the carrying capacity of parks and protected areas, but they might be applied more broadly in at least two ways. First, these approaches are fundamentally management-by-objectives systems: specific management objectives are formulated, in measurable terms; relevant conditions in parks are monitored; management practices are applied to help ensure that the objectives are achieved and maintained. This offers a thoughtful approach to managing many environmental and social issues in parks and related areas.

Second, this approach can be applied to broad-ranging environmental issues and areas. Carrying capacity addresses the fundamental tension between use of these areas and protection of important resources. This is a specific manifestation of the broader tension between (1) the degree to which we can use the environment for a host of purposes, and (2) protecting what we find valuable about it.

The framework outlined previously can guide environmental management in an array of manifestations and contexts. In fact, it is now being integrated into many applications: indicators (and to a lesser degree, standards) now guide planning in many environmental fields, and evolving contemporary management concepts – such as ecosystem management, adaptive management, and sustainability – are highly compatible with these conceptual and research approaches. While definitions and operational procedures for these concepts are still being defined, several principles can be isolated that might be broadly applicable to environmental management. First, the integration of ecology and society must be addressed: the integrity of important ecological processes must be protected, but natural and environmental resources must be managed for the benefits of society. Thus, ecosystem management has been defined as "regulating… ecosystem structure and function… to achieve socially desirable conditions, integrating… ecological relationships within a complex sociopolitical and values framework".

Second, managing the environment for the benefits of the present generation should not preclude the ability of future generations to attain similar benefits. This principle is at the heart of the emerging concept of sustainability.

Third, environmental management should be conducted within a framework that identifies goals and objectives and uses a program of monitoring: this principle is fundamental to the evolving concept of adaptive management, which emphasized the role of ongoing monitoring and evaluation. The carrying capacity-related frameworks outlined in this chapter - Visitor Experience and Resource Protection (VERP) and Limits of Acceptable Change (LAC) – offer a procedural approach that emphasizes (1) the development of management objectives (often called desired conditions in the context of parks and protected areas) and their expression in the form of quantitative indicators and standards: (2) a long-term commitment to monitoring indicator variables; (3) a program of management that responds to monitoring data, to ensure that standards for indicator variables are maintained. This procedural framework addresses both resource (or ecological) and experiential (or social) aspects of environmental management. Finally, this framework was developed explicitly to address the inherent tension between resource use and protection, or the underlying issue of sustainability, so it may offer a useful approach to guiding broader environmental management.

Research designed to help formulate indicators and standards can be applied in an array of environmental contexts. Society has a potentially important role to play in identifying indicators of environmental quality. Research on inherent tradeoffs between resource use and protection are nearly universal in their application to issues of sustainability. Innovative approaches to monitoring (such as simulation modeling) are needed to guide management actions and to assess their effectiveness.

2.6.1 Environmental Indicators and Standards

Application of the emerging concepts and principles of environmental management is clearly manifested in the use of environmental and social indicators and, to a lesser degree, standards. The contemporary scientific and professional literature contains thousands of references to the expanding use of indicators. There is evidence that early humans relied on environmental indicators such as migratory animal movements for information about changing natural conditions. However, modern scientific use of environmental indicators can be traced to the work of Clements (1920), who laid the foundation for the use of plants as indicators of ecological conditions and processes. Environmental indicators have expanded to include a host of measures other than

observation of plant and animal species, and they sometimes use indexes comprising multiple variables.

Social indicators also have a relatively long history of use. An early example is the work of H. Odum (1936), who developed a large suite of indicators of socioeconomic conditions in the southern United States for purposes of regional planning. Economic indicators such as unemployment rate, interest rate, and gross national product (GNP), along with social indicators such as crime rate, literacy, and life expectancy have been central to economic and social planning in the US for many years. Emergence of the concept of ecosystem management has emphasized the connections between the environment and society, and this has suggested that environmental management should include indicators of both ecological and associated social conditions.

Contemporary emphasis on the use of indicators is tied to the concept of sustainability, a direct outgrowth of the United Nations Conference on Environment and Development (popularly known as the Earth Summit) held in Rio de Janeiro in 1992. This conference prepared a plan of action to achieve sustainability on a global basis and called for identification of "indicators of sustainable development". The Commission on Sustainable Development was established to help ensure effective follow-up. To monitor the implementation of the plan, the commission established over one hundred broad-ranging indicators, including three types of variables: environmental (e.g. ambient concentration of air pollutants in urban areas), social (e.g. population with access to safe drinking water) and institutional (e.g. implementation of national sustainable development strategy).

The work of the Commission on Sustainable Development has been extended to many areas of environmental management by a host of organizations. For example, one of the more highly developed applications of indicator-based approaches to environmental management is the current program of sustainable forestry. Since the early 1990's, several international seminars and workgroups developed criteria and specific indicators to guide sustainable forestry at the country or national level. The criteria are analogous to management objectives or desired conditions as conceived in contemporary carrying capacity frameworks (e.g. the first criterion is *conservation of biological diversity*). The indicators are measurable, manageable variables that can be used as proxies for these criteria or objectives (e.g. an indicator of the first criterion is the number of forest-dependent species). The criteria and indicators are intended to provide a commonly agreed-upon understanding of what is meant by sustainable forest management and to be a mechanism for evaluating a country's success at achieving

sustainability at the national level. Given substantial differences among nations regarding basic forest-related conditions, standards for indicator variables are left to the discretion of countries that choose to endorse the program. These countries are expected to monitor indicators on a regular basis, with resulting data suggesting the degree to which sustainability in forest management is being achieved.

2.7 Conclusion

How much can we use the environment without spoiling it? This is the most fundamental question in all of environmental management. We have wrestled with this question across the span of human history, and it is now firmly on our environmental agenda in the form of common property resources, carrying capacity, and the emerging notion of sustainability. Manifestations and applications of this question continue to grow in number, scale and urgency: this chapter has addressed how this question applies to parks and protected areas, with special attention to the U.S. national park system.

Parks and protected areas are also examples of common property resources and subject to the tragedy of the commons: they are especially vulnerable to overexploitation because the environmental (and related social) degradation caused by their use is not borne fully by individual users (but is borne by society at large). There is a built-in incentive for "rational" individuals to overexploit common property resources. In the case of parks and protected areas, people will continue to visit these areas because – at the level of the individual – the benefits they receive outweigh the costs they must pay. However, at the greater societal level, the parks are nonetheless degraded and their value is ultimately diminished.

Concern over the tragedy of the commons is driven by an assumption that there are limits to our use of the environment. This issue is often considered within the rubric of carrying capacity: in its broadest manifestation, it is applied to the population of humans that can ultimately be accommodated in a given area or even on the planet as a whole.

More recent thinking suggests that the tragedy of the commons is less deterministic than originally conceived: for example, the rigid notion of rationality often assumed might be tempered by some degree of altruism or enlightened self-interested to protect what is important to society as a whole. This interpretation suggests that there are (or at least can be) social values and related norms that guide environmental management, including our use of common property resources and carrying capacity.

Such norms can be the basis of "mutual coercion, mutually agreed upon", the social action that Hardin suggests is required to resolve these issues.

Another shift in thinking suggests that population growth will cease before environmental constraints are reached (that is, no availability of natural resources), because other non-material conditions might begin to matter (e.g., economic well-being, level of environmental quality). There is intuitive and growing scientific understanding that increasing population and associated economic growth can lead to an array of environmental impacts and related social costs. The operative questions associated with carrying capacity, managing common property resources, and environmental management more broadly then become (1) what levels of environmental impacts are acceptable? and (2) what type of environmental and related social conditions do we want to maintain?

Contemporary application of carrying capacity to parks has followed this line of thinking, and the concept of "limits of acceptable change" was introduced into the outdoor recreation literature decades ago. Also, carrying capacity can be determined only as it relates to environmental and associated social objectives. For example, what type of visitor experience should be provided? These management objectives are sometimes called *desired conditions*. Moreover, they should be expressed in quantitative terms – generally called indicators and standards – so that conditions can be measured empirically. Indicators are measurable, manageable variables that help define the quality of parks. Standards define the minimum acceptable condition of indicator variables. With this conceptual framework, carrying capacity can be defined in an operational way as the level and type of visitor use that can be accommodated without violating standards for relevant indicator variables.

This approach to carrying capacity has been designed into several management frameworks, such as the Visitor Experience and Resource Protection (VERP) developed by the US National Park Service. However, all of the contemporary carrying capacity frameworks are built upon the conceptual foundation described earlier and function through a similar core sequence of steps:

1. Establish management objectives / desired conditions and associated indicators and standards.
2. Monitor indicator variables.
3. Apply management practices to ensure that standards are maintained.

As applied to parks and related areas, carrying capacity has resource, experiential and managerial components. Desired conditions and associated indicators and standards

should be considered for each of these components. The descriptive component addresses relationships between levels and types of visitor use, and resulting impacts to park resources, experiences and management. The prescriptive component addresses the seemingly more subjective issue of how much impact should be allowed. This component is often the most challenging. Finally, carrying capacity determination will always require some element of management judgment: if alternative carrying capacities are possible, than some judgment will have to be rendered as to which is the most appropriate.

Carrying capacity is an inherently interdisciplinary concept, affected by both natural and social science research. A program of natural science-based research on the ecological impacts of outdoor recreation – often called *recreation ecology* – has generated an increasing body of knowledge to help formulate indicators and associated standards. A program of social-science research has also evolved that includes theoretical and methodological approaches and a body of knowledge about experiential impacts that can also help formulate indicators and standards. Taken together, this research can address all components of carrying capacity, and, where needed, can be integrated to address the inevitable nexus between these components. For example, at what point do resource-related impacts of recreation degrade the quality of the visitor experience, or how can visitors be encouraged to mitigate their environmental and experiential impacts?

There is a growing notion that carrying capacity may be largely a social issue driven by the needs and wants of society, therefore its social aspect was emphasized in this chapter. It would be foolish to deny that there are environmental constraints, however they might often be wide ranging. In such cases, society will play a vital role in determining carrying capacity, and social science (integrated with natural science where appropriate) will facilitate this process.

A range of social science research methods have been adapted and applied to carrying capacity analysis. Perhaps the most important is normative theory and methods. If carrying capacity is a normative rather than deterministic concept as is suggested by recent thinking, then social norms are at the heart of measuring and managing carrying capacity. Research suggests that visitors have normative standards about appropriate environmental and experiential conditions in parks. These norms can often be measured and help inform the development of desired conditions and associated indicators and standards. Several forms of tradeoff analysis can be used to help ensure that normative questions and the answers they elicit are as informed as possible about

potentially competing and even conflicting objectives. Qualitative and quantitative survey research can help identify salient societal indicators of resource and experiential conditions in parks. Computer simulation modeling of visitor use can help develop important baseline data on visitor-use levels and patterns, can help monitor variables that are difficult to observe, and can help predict the potential effectiveness of alternative management practices.

Many case studies outlined in research are examples of the ways in which these methods are being applied and carrying capacity is being measured, analyzed and managed in a variety of contexts. A range of indicators of all three components of carrying capacity are being identified, associated standards are being formulated, and indicators are being monitored so that management actions can be taken to ensure that standards are maintained.

Efforts to address carrying capacity would ring hollow without feasible and effective management practices. Fortunately, there is a range of possibilities for parks and protected areas. When the number and/or size of protected areas cannot be increased to accommodate more visitors, we can limit demand through restrictions on the amount of use, or educate visitors in ways that will limit their environmental and social impacts. Sometimes, we can even harden resources to impacts with specific practices. It is advisable to consider the full range of management alternatives, and research in parks has only begun to assess the potential effectiveness of alternative practices.

This chapter addresses carrying capacity in the context of parks and protected areas but the concepts, principles and approaches might be equally applicable to the broader field of environmental management. Both carrying capacity and the newer concept of ecosystem management stress the relationships between the environment and society that must be addressed. Also there are obvious parallels between carrying capacity and the emerging concept of sustainability as both address the inherent tension between use of the environment and protection of its integrity. Indicators of environmental and related social conditions are a cornerstone of contemporary carrying capacity frameworks, and for similar reasons indicators are becoming cornerstone of environmental management in many of its applied fields.

Armed with a conceptual foundation and related set of terminology, an associated planning/management framework, a growing set of supporting research approaches, an array of management alternatives, and a number of hopeful case studies, it is viable to engage the carrying capacity of parks and protected areas more deliberately. Applying these tools will be challenging and sometimes even contentious. But failure to

do so will be even more painful in the long run. Management of parks and protected areas – often the crown jewels of a nation's natural and cultural heritage – should be conducted by design, not by default. By choosing not to manage parks and protected areas, we are implicitly deciding that their current conditions are acceptable, and that trends in use and related impacts are not worrisome. Management of parks – and of broader environmental issues – should be based on societal values and related norms, not on privilege bestowed by power or even scientific knowledge. Engaging the public in decisions about managing parks builds trust, ownership, and the "social capital" that engenders public enthusiasm and support. Parks and protected areas often have national and even international significance. Thus environmental management should increasingly be conducted at a global level. Some research methods, especially those that employ representative sampling approaches, can be useful at these higher scales. Management decisions that inherently limit personal freedoms are likely to be contentious, but they are more apt to endure if they are built upon the values and norms of those they most directly affect.

Despite advances in theory and related empirical methods, some measure of management judgment will remain inescapable. However, when this judgment is rendered in the context of a rational, transparent, conceptual and planning framework, and when it is supported by informed research and related public engagement, it will lead to a program of management that protects both the environment and the public good. If freedom is truly the recognition of necessity, then it is time to move ahead in the management of parks and protected areas and the broader field of environmental management.

CHAPTER 3. Questionnaire-based surveys: two case studies

A research group, who includes members from the Universities of Padua, Bologna, Naples and Florence, has been working at a national level using statistical methods to evaluate products and services. This is a new trend, since more and more clients ask for services to be evaluated quantitatively. Besides, there are not many studies involving advanced methods to analyze natural reserves, so the tourist association of Sesto and the Park managers were contacted to carry out a study using an articulate questionnaire.

It is paramount, when you provide a service, to monitor quality. Monitoring quality means to clearly understand how it is perceived by users, which is the main goal of this study. Only by understanding their perception, can we improve the performance. Nowadays it is vital to collect data on such performance, in a statistically correct way, and take action based on the analysis of this data, not simply on tourists' opinions. Several aspects of quality can be studied, but customer satisfaction seemed the most important element when offering a touristic service.

A natural reserve, or protected area, being a system that delivers goods and services, is required to monitor its own quality, and set goals for improvement. Therefore it shall implement monitoring tools, and for example *satisfaction surveys* can be such a tool, with reference to the quality *perceived* by users. Of course a park, a naturalistic area, like the Three Peaks Park, has specificities that must be taken into account, when you set up a project to monitor quality.

Inside a naturalistic, protected area, a natural tension is generated between the use of resources on one hand, and the need to protect them on the other. The question is how much we can exploit the environment without reducing its ecological and social value unacceptably. Of course the ever growing flow of visitors to different natural reserves, and alpine resorts (therefore also the increasing demand for excursions), determines a relevant tension, since this growing influx can bring modifications to the territory and the environment. This makes it necessary to monitor the touristic development, meaning that it is always necessary to have information on the tourist's experience, and how he relates to the environment. So managing a natural park has recently become ever so complex because, beyond specifically naturalistic aspects, you should bear in mind social aspects, with reference to local communities and the

experience of the visitors. Over the last years, conceptual/managerial/managing orientations have emerged (people speak of *sustainable tourism*), referring not only to the optimal use of natural resources, but also to respect the sociocultural identities of local communities. We talk about the idea of Carrying Capacity, that is the need to measure the capability of an area to provide recreational opportunities and deliver services, which are directed to everyone, taking into account that you cannot reduce the naturalistic and sociological values of the system.

Management becomes more and more complex, and local bodies, who are in charge of natural regions so rich and therefore attractive for tourists, are required to integrate specifically naturalistic aspects, with social aspects, meaning that it is necessary to protect ecological/naturalistic resources and processes, but these resources and processes must be directed to the advantage of society. The necessity arises for local bodies, to have information not only to implement policies to protect/support the environment, but also policies to promote tourism, and the cultural integrity of local populations. In this more and more complex setting, the role of monitoring statistical projects/surveys, both of environmental and social processes present in touristic activities, becomes crucial, in order to have some information to support the decisions of administrative bodies who are in charge of the territory. These surveys allow defining a series of *indicators*, that is objective statistical measures; they are not subjective opinions but have an important degree of objectivity, and provide information for various aspects of the management of environmental heritage. For example, examining the experience undergone by visitors, an indicator could be the percentage of tourists who consider the signposting of mountain paths more than satisfactory. So these surveys can give information on several specific aspects, and provide values for such indicators. The goals, as well as actions to improve quality, can entail reaching specific target values, so called *targets*. For example, a local body could decide, as a goal, to reach the value of 80%, within 2 years, for the indicator mentioned above. Naturally, surveys which monitor quality may evaluate different aspects in the management of the district, and in the relationship between the mountainous environment and touristic activity, so surveys may also concern the evaluation of how services are organized, in terms of what is being offered, or may involve the operators who produce goods and services supporting tourism, and may evaluate the experience and opinion of tourists.

3.1 Sesto Nature Survey

Statistical surveys such as the Sesto Nature Survey can provide invaluable information to support projects that aim to monitor and improve quality, and also to support policies directed at evaluating naturalistic and social aspects, involving modern concepts such as Sustainable Tourism and Carrying Capacity. The survey took place in July-August-September of 2010. Information was collected from a total of 262 respondents. During those 3 months sampling weeks were identified, and a network of interviewers was organized, and positioned in strategic spots along paths of the district. A questionnaire was prepared which consisted of 5 sections:

1. General Information
2. Information on the daily trip inside the district
3. Walks or excursions along the paths of the district
4. Iron ways or rock climbs in the district
5. Improvements of services and protection of the district

Notice the final section, which aimed to collect information on suggestions directly from the tourists, regardless of their characteristics.

Once the questionnaire collection was completed, the data was summarized and analyzed along different dimensions. Eight parts can be identified: (1) General characteristics of respondents' sample, (2) Characteristics of the sojourn in the scenic Dolomites district of Sesto, (3) Characteristics of the sojourn in the scenic Dolomites district of Sesto for holidays lasting a few days, (4) Images evoked by the scenic Dolomites district of Sesto and reasons of interest for the visit, (5) Routes/paths completed during the visit to the scenic Dolomites district of Sesto, (6) Walks or excursions on foot or by mountain-bike along paths of the scenic Dolomites district of Sesto, (7) Iron ways and rock climbs in the scenic Dolomites district of Sesto and (8) Improvement on services and protection of the scenic Dolomites district of Sesto. In each section, respondents have been grouped, depending on the aspect that was being studied (e.g. nationality, presence/absence of children, reasons why they visited the area, etc.).

All tables for this survey are presented in Appendix A.

It would be advisable to carry out surveys like this periodically: they provide valuable information at first, to support decisions regarding environmental and social aspects. Surely, the longitudinal approach (that is, repeating the survey at different times) has more than one advantage: it keeps the situation under control by taking several pictures of its status; it provides more robust information from a statistical standpoint

(each sample is always subject to fluctuations), it can show the effectiveness of the actions taken to improve the naturalistic experience, all this by using a tailored system of quality indicators. If this survey were to be repeated in the future, there would be room for several improvements: for example, it could be extended to include other aspects, such as the opinion of the workforce employed in the area.

3.2 Survey on Alto Adige Ski Schools

A research group, involving a few Italian universities, has been working on the evaluation of didactics, in a classical sense (mainly schools and colleges). Prof. Salmaso coordinated this survey on the satisfaction of students attending a ski school in Alto Adige. More and attention has been focused on sport activities, because it seems that monitoring and improving quality are becoming crucial.

Collaboration started with several schools in Alto Adige, but other organizations were involved, and the outlook is promising. This study is innovative at a national level: it is the first systematic study conducted in different schools, with qualitative evaluation, using a questionnaire scientifically designed to measure satisfaction and quality *perceived* by the users. A separate publication on this project will be issued, and the survey will be presented at an international conference - held by the prestigious American Statistical Association - where most innovative researches will be discussed, on new study areas such as didactics in skiing.

This study was conducted in selected weeks of January, February and March 2011, contacting 38 schools and handing out questionnaires both for kids and adults. We identified 3 phases of the service, each with specific quality dimensions:

1. *Booking service*, with quality measured by: adequate opening times; clarity & completeness of informative brochures; staff (clarity & completeness of information provided, courtesy and helpfulness);
2. *Course organization,* with the following quality dimensions: ways to organize courses; skill homogeneity of groups after selection; events planned with courses (torchlit descents, competitions, …); slope enrichment (inflatables, snow sculptures, …);
3. *Carrying out classes*, with quality measures based on: effective teaching (clarity of notions, courtesy and helpfulness of teachers); safety (adequate slopes and lifts, subjective perception of safety); users' general satisfaction (enjoyment & fun, increased passion for skiing, kids' comfort, …).

Each dimension was investigated with specific questions and adequate scales.

It should be noted that performance of ski schools appears to be similar during peak periods and low season, which is an indication of quality since the level perceived does not seem to be affected by seasons but remains pretty high.

A note of excellence emerges when observing the courtesy and helpfulness of teachers, so professionalism is indisputable. As a consequence, parents would recommend the school to others, which determines a very positive word of mouth, even internationally.

On a lower note, course selection and clarity of brochures and web-site are aspects that do not score high in terms of perceived quality. For adults only, the perceived improvement in their skiing abilities seems a lever to bear in mind in the future.

To fully exploit its potential, a survey like this should be repeated in different years: only by comparing indicators in different moments in time, can we understand trends of improvement/constant level/decline, thus monitoring and improving quality continuously. Therefore, we evaluated only one school in two following years. This is particularly important to assess quantitatively whether actions taken by the school managers have been appreciated by the clients.

A coherent system of indicators should be developed to measure the progress of these schools. Still, another possible development would be to define indicators for the whole touristic package offered to skiers, not just to evaluate ski schools, however important. Ultimately, one could monitor the quality perceived by those who do not ski, and yet visit these mountain resorts. This perspective suggests a much broader vision where, along with the activities of a ski school, other services are monitored: facilities like hotels and accommodation, entertainment for non-skiers, tours. Finally, quality could be analyzed in terms of perception by tourists, but also listening to the opinion of those who provide the service: teachers on courses, technical staff on slopes and ski facilities, etc. This leads to the fundamental concept of total quality, once again evaluated in terms of progress over time.

All charts for this survey are presented in Appendix B.

CHAPTER 4. A literature review on odour emissions

Here we present relevant extracts from papers on odour emissions, examining in particular issues regarding their measurement.

4.1 Estimation of odor emission rate from landfill areas using the sniffing team method (Nicolas/Craffe/Romain)

Unpleasant smells can cause serious nuisance in the vicinity of sanitary landfills. Odors of different kinds are released by the fresh deposits of municipal solid waste, by the landfill gas (LFG), by the leachate treatment plants, by flares and by some waste treatment works, like composting facilities. Concerning solely the waste odor, there is a wide variety of emission sources, conveniently separated into the specific activities that liberate odorous compounds such as the active tipping of waste itself, but also the waste transportation by disposal trucks, the intermediate storage or the handling process after the garbage deposit.

Consequently, controlling odors from landfill sites has become an important regulatory issue, requiring accurate and reproducible sampling and measurement. But the monitoring of the odor annoyance generated by a landfill area is difficult. Problems appear already at the sampling level.

The most important sources at landfill sites are indeed passive area sources that are remarkably large. Very often, it is not possible to sample more than 1% of the total area, so one must assume that the distribution of the specific emission rate is homogeneous, which is not realistic.

Many authors mention also that the main odor problem of a landfill is caused by the handling of the fresh waste. As this is an intermittent activity, the sampling of the gas that is emitted at the landfill working face is particularly problematic. Some additional problems arise at the analysis level.

Recently, some attempts were made to use the electronic nose for the field monitoring of the landfill odor. Such technique leads to very promising results at the research stage, while its routine use to monitor on-site odor remains challenging. It notably entails the improvement of the quality of the used sensors: limit of detection, drift, influence of water content, etc.

For the above mentioned reasons, one of the most representative and the most frequently used way to assess the overall odor level still remains the sensory measurement using a panel of judges. Usually, the measurement goal is the determination of the mean odor emission rate from the whole landfill area, expressed in odor unit per second (ouE/s: the "E" stands for "European", as defined by the European standard EN13725 (2003), later on, this subscript will be used only if that European standard method is applied). Such outcome can be used for further evaluation of odor concentration percentiles prevailing for typical climatic conditions. That long term exposure is quantified in terms of a frequency of occurrence of hourly averaged concentrations above a certain limit odor concentration.

For example, the 98-percentile for a given odor concentration, e.g. 5 ouE/m^3 (odor unit per cubic metre), represents the contour line delimiting the zone at the ground level where that concentration is exceeded more than 2% in the year. In short notation: C98, 1 h = 5 ouE/m^3. Here, "1 h" means that the concentrations are hourly averaged. This measure of exposure is calculated from the estimated or measured odor emission rate from the source, using an atmospheric dispersion model. Knowing that 1 ouE/m^3 corresponds to the odor detection threshold, in particular, the C98, 1 h = 1 ouE/m^3 percentile shows the limit of the area beyond which the odor is perceived less frequently than 2% of the time.

A first way to estimate the overall odor emission rate from a diffuse source, like the fresh deposits of municipal waste, is to use an isolation flux chamber or a portable wind tunnel placed on the landfill surface to collect gases, which are then transferred to a Tedlar bag for subsequent testing by olfactometry. Dynamic olfactometry (e.g. European standard EN13725) is the method by which different dilutions of the gas sample are dynamically presented to trained odor assessors to determine the odor concentration of the original sample (in ouE/m^3). The combination of surface sample collection and olfactometry provide both the concentration of odor (in ouE/m^3) and the volume air flow (in m^3/s), the product of which is the specific odor emission rate (in ouE/s). However, as already mentioned, such point samplings over the large and heterogeneous area of the landfill site pose the problem of the representativeness of the resulting emission rate. Moreover, it is impossible to estimate the flux of the odor emitted when handling the solid waste, or the one generated by the waste truck traffic by this method. Hence, the result provided by the method is only a part of the total odor emission rate.

Alternatively, the determination of the odor emission rate can be based on global field measurement, taking account of the real perception of the odor in the environment in the surroundings of the source. A possible approach is the method of the sniffing team observations, which utilises experienced people to evaluate the maximum distance from the source at which the odor is perceived. "Experienced people" means operators with reliable olfactory performance who always apply the same sniffing procedure. The results of a dozen of such measurements allow calculation of the typical odor emission rate with a dispersion model. Sniffing team methods have some advantages over instrumental and olfactometric measurements. The main advantage is that they involve field measurements, by which the global impact of the source is evaluated, allowing consideration of diffuse, surface and less clear sources, such as waste handling or transportation. Furthermore, these methods reflect the actual perceptibility of the odor in the environment.

However, the sniffing team observation method also presents many limitations. That is chiefly the aim of this paper to describe them in detail. Firstly, it makes fundamental assumptions: it is valid only if both the meteorological situation and the odor emission do not vary too much during the measurement period. The waste odor around a landfill site is actually emitted as discontinuous puffs, depending on the activities on the landfill tipping face. Moreover, often accessibility problems in the surroundings of the site do not allow very quick observations. If the measurement takes about 1 h, both the meteorological conditions and the emission rate can vary significantly.

The present paper discusses the applicability of the sniffing team observation method to estimate the annoyance zone around landfill areas. It is based on 52 measurements made on five different municipal solid waste landfill sites in Wallonia, in the South of Belgium. The causes of the estimation bias are identified and the relative errors are estimated by a sensitivity analysis. The main topic of the paper is the discussion of the applicability of the methodology to diffuse and discontinuous odor sources. The results are only supplied for illustration purposes.

4.1.1 Methods and Operating Conditions

The sniffing method, as applied by the Department of Organic Chemistry at the University of Gent, is described in detail in Van Langenhove and Van Broeck (2001). One or two observers are firstly familiarised with the odor emitted by the source. If necessary, the olfaction performance of new and unexperienced observers may be checked against n-butanol, considered as a standard reference odorant, like for dynamic olfactometry. Then, they detect the odor at different points, by a zig–zag

movement around the axis of the plume. The transitional stages from no odor perception to odor perception are recorded on a detailed map, so that the odor area can be plotted and the maximum odor perception distance can be determined. By definition, the odor concentration at this maximum is 1 ou/m^3. Strictly, this definition is only valid when the rigorous procedure of olfactometry, according to EN13725 standard, is respected. The detection threshold may be different in the environment than in laboratory conditions. However, we will consider that 1 ou/m^3 corresponds also to the perception threshold for the present method. To avoid any confusion, we will not use the "E" subscript of ouE/m^3, valid only for the European standard method.

As the size of the odor perception area also depends on the meteorological situation at the time of the measurement, the wind direction, the wind speed and the solar radiation (or cloudiness) are simultaneously recorded. The two last parameters allow determination of the atmospheric stability using the Pasquill stability class system (Pasquill, 1974). Then, a bi-Gaussian model, adapted to simulate the odor perception, is used with the average values of these meteorological data. The emission rate entered into the model is adjusted until the simulated average isopleth for 1 ou/m^3 at about 2 m height (the height of the human nose) fits the measured maximum perception distance.

The sniffing team method was applied to five landfill areas in Wallonia (South of Belgium), which is a region characterised by quite homogeneous climatic conditions, with prevailing wind directions NE and SW. Landfill sites (Mont-Saint-Guibert, Hallembaye, Champ-de-Beaumont, Cour-au-Bois, and Froidchapelle) are different in size (capacity from 0.8 to 5.3 million m^3), in topography (from almost flat environment to slight hills) and in neighbourhood (always in rural areas, but from almost none to about 500 dwellings in a circular zone of 1 km radius around the active tipping area). Typically 100,000 m^3 of waste are deposited per year on the landfill areas. All of the landfill sites predominantly receive municipal solid waste, which is immediately spread and compacted with suitable engines. There are no other odor sources in the immediate surroundings of the studied sites, except in the case of Hallembaye where the odor emissions of a hen house cannot, however, be confused with those of the fresh waste from the landfill.

Observers were trained by the same person prior to measurement campaigns, in order to be sure that all of them use the same procedure: detecting the same odor quality, considering the same minimum puff duration before acknowledging an odor point, staying about the same duration at each location, etc.

Although the landfill gas (LFG) odor was sometimes perceived on some sites, the sniffing concerned exclusively the fresh garbage odor, which was, by far, the strongest odor during activity periods and which generally corresponded to the complaints in the surrounding area. LFG emissions are actually due to imperfectly airtight extraction wells. For all of the investigated sites, LFG collection networks are very efficient and LFG odor was exceptional and only locally detected. However, that shows the importance of the familiarisation of the observers with the typical smell of the source. In our case, many different odor characters could locally be perceived on the same site: the sour smell of the fresh waste, or the sickly sweet smell of the LFG, or the "rotten-egg" odor of leachates, or in some cases, the odor of the compost used as capping material or the one of sewage sludge temporarily stored on the site. After a short training, the observer can easily distinguish all of the odor qualities.

The research group ENVOC, from Gent University in Belgium experienced the sniffing team observation method in a very great number of different cases and they conclude that one single observer is sufficient since the difference in observed maximum perception distance for different observers is only 10–15%. Our research confirms such values: we investigated a similar method with student teams in different cases, comparable to the one of landfill site, and the range of the estimated values of the maximum perception distance among all the observers was always within 10%.

Extensive analysis has led us to identify various errors and their effect on the adjusted results, for the 52 studied cases on landfill areas. Briefly, we noticed two types of errors: estimation errors (that is: field sensitive perception, height of measurement of wind speed and direction, estimation of stability class, release height and plume rise, size and shape of the diffuse emission) and methodological errors (choice of a bi-Gaussian model, choice of an algorithm to simulate the odor, frequency of meteorological observations entered into the model, choice of the isopleth identified to the perception limit, reflection on mixing layer).

4.1.2 Conclusion

Among the above listed errors, those which are due to methodological options are by far the most important ones. It is clear that such a method, involving notably the use of a model, leads only to approximate measurement results since it is based on a crude representation of the reality. Such methodological errors must be pointed out to show the limits of the approach and to relativize the importance of the field observations. Nevertheless, they should not be considered in the calculation of a confidence interval around the estimated percentile. If all the methodological options are well argued, they

define the exact frame of the used method and one must only care about the biases induced by deviations with respect to this reference method.

Those considerations do not prevent the method to be improved when it is possible. More particularly, a significant finding is that the Pasquill stability classes scheme does not offer a sufficiently fine resolution to be used in the Gaussian-type models (or a least in most computer codes) when a single hourly averaged meteorological observation is used to adjust an odor emission rate.

The validation of such a method is not easy: the emission is discontinuous and diffuse. The validation for all of the weather situations and various emission types should request substantial amount of work and money.

The results of other similar studies could be used for an attempt of validation of our own method, but few scientific papers or technical reports provide all the data required to apply the model. In the appendices of a study concerning livestock odors carried out by the University of Gent (De Bruyn et al., 2001), a table gives the maximum distance of odor perception, the average meteorological conditions prevailing during the measurement period and the odor emission rate as adjusted by the model.

To sum up, we may conclude that the proposed method is proved reliable for the determination of percentiles of the odor perception threshold exceeding for typical climatic conditions. All errors induced on the final result, especially the methodological biases, are very reduced if the intermediate result, i.e. the odor emission rate, is not exploited as output variable and if the same model, with the same hypothesis, are used both to adjust the emission rate and to calculate the percentiles. Although the described method is particularly well adapted for perturbed climatic conditions and fluctuating odor emissions, its reliability is maximum for rather steady situation and when the detection of limit points is carried out as fast as possible.

4.2 Community modelling: a tool for correlating estimates of exposure with perception of odour from municipal solid waste (MSW) landfills (Sarkar/Longhurst/Hobbs)

Assessing odour emissions from large landfill sites is a significant problem. The odour comes from operations and processes on site exposing mixtures of volatile organic compounds present in the landfill gas, leachate, and treatment systems, as well as waste odours from sludges and solids. Symptoms of these problems are normally associated with reports of annoyance from neighboring premises and increasing complaint rates from the community.

To meet the requirements of the local planning authority and comply with the EU Directive (85/337) on environmental impact assessment (EIA), odour was assessed as part of an application for permission to extend an existing landfill site, located to the south west of Bedfordshire in the United Kingdom. The County Council, as the Planning Authority responsible for determining the application, recognised concerns about the risk of odour problems arising from the proposed extension and requested a review of the potential for odour annoyance to assess the extension of the site and its operation. This work formed part of an on-going research study within the School of Industrial and Manufacturing Science (SIMS) at Cranfield University on the assessment of odour impacts on the communities surrounding solid waste disposal sites. Where complaints had been received of malodour from the site, researchers in SIMS were provided with these records. Alongside this data, an ongoing survey by the University of landfill odours detected within the last five years allowed an analysis of the potential for complaints under differing operational and meteorological conditions.

A quantitative model was developed at the College of Aeronautics, Cranfield University to assess the perception of odorous emissions from the landfill site on the surrounding community. The objective was to test the major components of the model namely, assessment of odorous emissions, dispersion and reception by the surrounding community around the landfill site. Community modelling was used as a tool for linking the last two components of the model—dispersion and perception of odour from the landfill site. This was also used as a validation step for the results of predictive dispersion modeling.

The standard method to assess odour impacts from a site includes an assessment of emissions and the use of dispersion estimates as a predictive model of community exposure to determine the dose-effect relationship indicating annoyance. This later phase of relating the analysis of exposure to the predicted perception of dispersed odours when received by residents in the surrounding community is key in interpreting dispersion model results. Perception has previously been analysed with four well-known psychophysical models. This paper describes the use of community modelling to link the calculated exposure, from dispersion analysis, with the perception reported by the community surrounding the site.

The test site, located in Bedfordshire, normally accepts between 2 and 300 vehicle loads of waste/day though it is licensed for more than twice this number. The site receives a wide range of controlled wastes of domestic, commercial and industrial origin. Major sources of emission were identified as: the transfer and filling edge; gas

extraction and pre-treatment process prior to combustion; extraction and re-circulation of leachate; the residual fugitive emissions from the covered landfill surface.

4.2.1 Community Modelling

This study was designed to identify the extent and intensity of odours attributable to emissions from the site. Therefore, where records were available, odours detected from the site were correlated with dispersion calculations using estimates of site emissions. Monitors were recruited by the University to report on a daily basis whether odours were detected from the landfill site. The location of monitors was specified from the history of complaints made to the County Council and an initial screening for the likelihood of maximum impacts was carried out using weather data from a nearby meteorological station.

Monitors were selected to regularly report on odours based on guidelines, previously recorded by Hitchin (1998). Each monitor was tested using jar dilution tests to ensure a minimum level of sensitivity—individual threshold, a discrimination test between five different odours, and a category scaling to data. Lists of the location for monitors within the neighboring villages are presented. Each monitor was requested to report any incidence of odour detected within the day. Reports were normally from the same location, preferably from one particular area of their residence. Each monitor reported the scale of odour, probable source and a certainty level for the origin of the source.

It is to be noted that all odour measurements on site were carried out as per the then draft European standard of odour (EN13725), and therefore traceable to the reference value of 123 mg m^{-3} n-butanol (40 ppb) panel threshold.

All positive records, reported by the community monitors, were analysed to judge the consistency of each monitor in terms of intensity scaling. Records reporting odours were then compared with the results from the dispersion model predictions. Here, separate intensity–concentration plots have been fitted for each of the monitors whose reports were found to be logically consistent with regard to the intensity scaling.

Community panels were used as these provide a more reliable tool than the complaint histories, which are prone to fluctuation, to identify long-term trends in exposure to odours arising from the site. Odour records were collected from 1994 and positive records for the year 1997 are listed for 10 of the 42 monitors.

4.2.2 Conclusion

Community modelling is useful in analysing the correlation between exposure predictions from dispersion modeling with the analysis of perception of odour from specified sites. Community modelling quantitatively integrates two components of a

model for the analysis of odour, namely the exposure to odour from a dispersed source and the perceived intensity. An ongoing limitation of this approach is likely to be a lack of sufficient data, particularly where reported odours are also used to gain information on how to reduce site emissions. Successful use of this method requires measurements that account for the variety of operations on a site and an understanding of any changes in the intensity of emissions that may result.

Where additional information can be gained about the change in intensity of an odour with concentration, this in turn will enhance knowledge of site emissions and complaint reports. This may be gained from increasing the number of dilution levels in olfactometric experiments where intensity–concentration measurements are used. Two separate experiments may be required for an equivalent sample where an olfactometer is restricted on its dilution range.

4.3 Odour from municipal solid waste (MSW) landfills: A study on the analysis of perception (Sarkar/Hobbs)

Odours from landfill wastes comprise complex mixtures of a large number of volatile compounds. Odour concentration is a measure of the detectability of the odour as assessed by a panel of people. Odour intensity is defined as the perceived magnitude of a stimulus. Odour intensity and offensiveness are subjective measures of the strength and unpleasantness of an odour as assessed by a panel of people. Odours of equal concentration will not necessarily be of equal perceived intensity or offensiveness. Although the intensity can be perceived directly without any knowledge of the odour concentration, it is necessary when used in conjunction with dispersion modelling, in terms of comparing the resultant odour concentrations at the receptors (locations of potential complaints), as obtained from the dispersion analysis, with those obtained by reducing the intensity scales of the odour complaints to odour concentration levels. The idea could also be utilised by legislators to establish minimum separation distances between the landfill site and zones of potential complaints based on objective criteria.

In this paper, the main focus will be given to the selection of various psychophysical models and estimation of their parameters with suitable techniques. Afterwards, the models will be evaluated with statistical analysis.

The results will be discussed afterwards, with nine samples taken from various locations within a municipal solid waste (MSW) landfill site. There will be an attempt to rank the models according to their performance and one or two model(s) will be selected as the basis for community nuisance analysis. One of these psychophysical

models, already discriminated on the basis of its performance, will be used to convert the intensity scales reported by the community sniffers to odour concentration (ou/m3), which may be used to validate the results from dispersion analysis.

4.3.1 Methodology

The development of methods will include:

- Measurement of odour concentration and intensity,
- Selection of the psychophysical models and estimation of the respective parameters,
- Evaluation of the models with statistical analysis.

Odour intensity and threshold odour concentration were measured simultaneously by using a dynamic dilution forced-choice olfactometer. Several experiments were carried out and data of odour intensity and threshold odour concentration were obtained.

Samples of odour emissions were collected from the knock-out-pots (KOPs) and from the waste surface. The KOPs and gas well heads had conveniently fixed gas sampling ports, a sampling tube was connected to this and the gas sample sucked from the port into a Nalophan odour bag, contained in a barrel, using the lung principle. Duplicate samples were taken from each well or KOP. Samples from waste surfaces, freshly tipped and those 1-day-old, were collected using a Lindvall hood. With this equipment, a controlled flow of air is passed over the surface. The flexible air inlet hose is positioned at least 10 m upwind of the sampling area, the air passes through the fan and then through an activated charcoal filter to eliminate the odour of the inlet air. The hood covers 1.5 m^2 and the air velocity is of the order of 0.1 m/s. At each sampling position, duplicate samples of inlet and outlet air were taken over a period of about 10 min immediately after the fan was started.

Olfactometry is an objective method of expressing the strength, concentration or intensity, etc., of an odour. The method used determines how many times a sample must be diluted with odour-free air to be at the threshold of detection by 50% of the panel. The number of required dilutions defines the odour concentration in odour units per cubic meter (ou/m3). These tests are carried out inside an odour-free, clean laboratory with trained and selected panelists.

In this study, odour concentration was measured using an "Olfactomat" dynamic dilution olfactometer (Project Research, Amsterdam). A sample was presented to an odour panel using the forced-choice method. Six dilutions of each sample, differing from each other by a factor of two, were presented to the panelists three times.

Dilutions were made using odour-free air supplied by a compressor fitted with carbon filters and an air dryer. The olfactometer has two sniffing ports, one containing the diluted sample air and the other odour-free air. For each presentation, panelists indicated via a keyboard which port delivered the odorous air. In order to put greater confidence on the panelists' responses, they were also asked to indicate whether their choice was a "guess" (as it would have to be if the odour presented was below their personal threshold level), whether they had an "inkling" that their choice was correct (when the odour was close to the threshold level) or whether they were "certain" that their choice was correct. The mean threshold value for each sample was calculated using Dravniek's method (Cheremisinoff and Young, 1975).

The assessment of odour intensity indicates the effect of differing odour dilutions on the likely smell sensation for an individual. Measurements of intensity are determined by the "sniffing" panel using a subjective scale (usually 0–6) from no odour to extremely strong. Depending upon odour type and selection of the panel, high confidence levels can be achieved from these qualitative judgements. Odour intensity was measured using a category estimation technique. Following the determination of odour concentration, ranges of suprathreshold dilutions were presented in random order. The panelists were required to indicate their perception of intensity at each dilution. Mean intensity scores were obtained at each dilution presented to the panel. The concentration of the odour at each dilution was calculated as the sample concentration divided by the dilution factor.

Various psychophysical functions, based on different empirical laws, were chosen to demonstrate the relationship between perceived intensity and odour concentration for the samples drawn from the landfill site.

4.3.2 Conclusion

The analysis of perception of odour samples from a MSW landfill site was done using various well-known psychophysical models and respective parameters for each of the models were estimated and the overall performance of the model was tested against sets of data from the olfactometry analysis.

It could be concluded that for odour samples from various KOPs and areas of freshly tipped wastes of the landfill site, Model 1 (based on the Weber–Fechner law), could demonstrate the intensity–concentration relationship best. In the above analysis, Model 1 (based on Weber–Fechner law) was ranked 1 in case of five out of nine samples and it has been found more representative of the less intense odour samples.

The Weber–Fechner law performed better than Power Law since the scaling technique used was category estimation and not magnitude estimation.

Model 4 (based on Laffort's equation) could correlate the intensity with odour concentration very well for samples from the horizontal wells. Laffort's equation has specifically represented the intensity–concentration relationship better for comparatively more intense odour samples.

In case of the particular samples analysed, it has been found that frequency of intensity scales reported have been mostly in the lower range. Hence, the performance of Model 1 could be tested with much more data in comparison to Model 4.

Depending on the nature of the odour sample and its range of intensity levels, each of Model 1 or 4 could be selected to find out the concentration of odour at a particular receptor location and the dispersion modelling results could be validated.

4.4 Appropriateness of selecting different averaging times for modelling chronic and acute exposure to environmental odours (Drew/Smith/Gerard et al.)

The emission of odour from landfill sites and industrial processes is a recurrent problem for operators and regulators, who have to deal with complaints from the public. Population growth and housing needs have resulted in increasing numbers living within close proximity to these odour sources. In the UK, 80% of the population live within 2 km of either a closed or active landfill site and therefore, the potential for exposure to odours is high. Odour at landfill sites is primarily caused by the anaerobic decomposition of biodegradable waste. The exact nature of odour emissions is therefore dependent on waste characteristics, such as composition and age.

The sequence of events leading to odour annoyance has been described as: formation of the odour at source; emission from source; transport to receptor; and perception by receptor, who then makes a judgement as to whether the odour causes an annoyance or not. Transport of the odour is affected by factors such as the season, time of day and the atmospheric conditions influencing dispersion of the odour (e.g. turbulence, wind speed and wind direction). Detectability and annoyance potential will influence the response by receptors. Perception of odour may therefore be affected by the combination of odorous compounds released during formation, as well as the characteristics of the odour itself, such as duration and frequency of emission.

Odour intensity and hedonic (the pleasantness or unpleasantness of the odour) properties experienced by the population may be interpreted as strong or offensive, respectively, in place of faint or not-unpleasant. Research shows that the hedonic tone

or pleasantness of the odour has an effect on the annoyance people feel. Hedonic tone has a clear impact on the annoyance felt by receptors at low concentrations, with pleasant odours having significantly lower annoyance potential than neutral or unpleasant odours. The hedonic tone of unpleasant odours at higher concentrations does not affect the annoyance potential of these odours, and odour frequency is sufficient to predict odour annoyance from unpleasant odours.

Factors including personal health, social status and previous exposure to odours may all influence how a person perceives an odour. A person's response to odour can further be influenced by the context of exposure, such as the presence of other odours and the reactions of people around them. Some members of the population are more predisposed to complain, while others may adopt alternative coping strategies. If the odour is perceived to be associated with a potential health risk, the probability of concern and increased annoyance is higher.

The primary concern during monitoring and measuring odours is determining the threshold at which an odour becomes a nuisance. Two terms used to define the response of the public to odour emissions are annoyance and nuisance. Annoyance is defined by Lindvall and Radford (1973) as the negative response associated with exposure to an agent or event that is believed to cause harm to the individual, and thus requires a coping strategy. A nuisance is commonly defined in law as the threshold at which a population experiences annoyance, from repeated incidents of exposure. These may be translated into law as a statutory limit.

The difficulty in predicting perception and response to odour at different concentrations is problematic for the definition of emission limits with which to regulate industries causing odour. Two metrics are commonly used to define annoyance: the sensory metric of odour concentration or 'dose' to which a receptor is predicted to be exposed, and the time or duration of exposure.

Authors reviewed standards for various USA state authorities, as well as European and Pacific Rim countries. Their review shows wide variations as to what is considered acceptable across these authorities. The pattern that emerges from studying odour regulations across the world is that less densely populated countries, such as Australia and the USA, have more stringent regulations than more densely population countries. The logic behind such stringent regulations is that if there is no odour, there will be no complaints and therefore no problem. However, stringent limits such as these can result in high remedial costs to the process operators. Most European countries seek to

regulate less stringently by providing quantitative limits aimed at reducing annoyance to an acceptable level at an acceptable cost.

No regulations are imposed by the European Union with respect to odours, except for a standard for the measurement of odours, developed by the European standardization committee (CEN, 1995, 2003). Individual countries have national regulations.

Odour emissions are episodic, characterized by periods of high emission rates and interspersed with periods of low emissions. The human olfactory sense responds within seconds to a stimulus. Odours therefore create a response in the receptor quicker than most other atmospheric pollutants (Irish Environmental Protection Agency, 2001). Greater annoyance is caused by more short periods of odour than by longer lasting odour emissions, as the olfactory sense is able to adapt to persistent odours, thereby reducing annoyance. However, the short term, high peak concentrations may still be detected and considered an annoyance. In other words, it is frequently the fluctuations from the mean concentration, and not the actual mean itself, that determine how the odour is perceived. However, odour regulations are currently expressed as hourly average concentrations.

Dispersion modelling has frequently been used to assess the potential dispersion of odour from industrial sources. Two approaches to modelling odour nuisance for regulation can be adopted. The first option aims to model the "real life" situation and is an attempt to model and understand the odour concentrations that may cause annoyance, or in other words, the concentration average over a certain time period, usually 1 h. This is the approach often used by regulators and is acceptable as long as exposure is not underestimated and a "tolerable level" is defined.

The use of concentrations averaged over such periods effectively filters out peak and short term fluctuations, resulting in conservative results with respect to maximum concentration levels. While a single peak may not result in annoyance, repeated high peaks at times of high exposure could be missed by using averages. Authors considered it unlikely that an odour will be a nuisance until it is detectable for longer periods of time, typically longer than 3 min.

The second modelling approach involves the use of short averaging times. In this way, it is possible to capture concentration peaks, and thereby obtain a more accurate prediction of odour dispersion. New generation air dispersion models can be run at averaging times of less than 1 h, although they are typically not used for short interval averaging times by regulators. Furthermore, the most frequently available atmospheric input data for these dispersion models are hourly averaged variables.

Any model will require simplifying assumptions to be made and will have built-in uncertainties, uncertainty being a measure of the reliability that can be associated with the results of a model. In particular, uncertainties associated with source term measurements, for example, instrument failure or incomplete data recording, will be carried over into modeling studies. If the magnitudes of measured results are considered as a Gaussian distribution, the "tails" of the distribution, representing relatively low sample numbers, are associated with a higher margin of statistical error. Furthermore, odours are commonly the results of a release of several odorous compounds, but they are generally modelled as a single indicator compound, usually with a low odour threshold and a high emission rate. Taken with the regulatory approach of modelling the hourly average concentration, this can mean that total odour concentration peaks could be seriously underestimated, resulting in annoyance and complaints.

Odour concentration measurements within a laboratory alone, using olfactometry, or instrumental analysis, fail to capture the properties of the odour as perceived by a community as it does not capture the other characteristics of the odour such as hedonic tone, which influence the way the odour is perceived by the public. Hedonic tone assessments can also be carried out in the laboratory. Authors analysed the link between odour dispersion and the perception of odour from a landfill site, using data from a monitoring programme within a community. The response of the community was found to vary greatly.

Odour emissions are episodic, and it is the infrequent, high concentration peaks that cause annoyance. Dispersion modelling is accepted as a useful tool for odour impact assessment and guidance exists for odour dispersion modeling. However, little attention has been paid to the appropriate definition of averaging time when attempting to understand off-site amenity impacts.

This study attempts to assess the appropriateness of using different averaging times to model the dispersion of odour from a landfill site. These results will be compared with a community monitoring programme database. We aim to examine the perception of the odour in the community in conjunction with the modelled odour dispersal.

4.4.1 Material and Method

The landfill site studied, located in Bedfordshire, is licensed to receive up to 600 waste vehicles a day, although it usually accepts about half that number. These vehicles contain commercial, household and industrial waste. This site has been studied for approximately 10 years by researchers at Cranfield University.

In order to determine if annoyance was being caused by the landfill site, two indications of odour annoyance were used:

1. Complaints to the operators from all members of the community.
2. Daily monitoring records made by selected members of the local community.

Site inspections carried out by the Community Liaison Officer from the company are used to understand the causes of incidents and assess the control the operator has over these incidents.

Complaints to the landfill operators can be made through a number of routes, either directly to the site, to the operator's Community Liaison Officer, to the local authority (County Council), to Environmental Health Officers (EHOs), or to the Environment Agency, which incorporates the previous Waste Regulation Authority. Each of these parties ensures that the operator and local authority are informed of the complaint. The details of the complaints include the location, name (where given) of complainant, number of people complaining, the nature of the problem, the time the odour occurred, the time of reporting, result of the investigation as to the cause, and the weather conditions recorded from the automatic weather station on-site. The most common recorded complaints are from the following sources:

- Landfill gas emissions caused by methanogenesis taking place within the landfill cell.
- The construction of liquid waste disposal trenches, which expose existing waste deposits prior to appropriate covering.
- Refuse recently delivered to the landfill site and not yet placed within the cell structure.
- Odours arising from certain types of waste (e.g. chemical treatment waste or malodorous waste) as it is delivered.

In addition to the complaints data, a system of odour monitoring by selected members of the surrounding community has been established since 1994. These daily reports record all odour types, coded into four categories: local odours (e.g. bonfires), landfill odours, odour from a neighbouring brick-works and agricultural odours. Community monitors have all volunteered to take part in the study and are anonymous to the landfill operators. In addition to their monitoring role are encouraged to report complaints to the operators as any other member of the community would normally do.

The number of monitors has varied throughout the period of the study, from 13 to 25, with 43 individuals contributing since 1994. An average of 17 monitors have recorded

odours within the area during the 10 years of the study. Each person monitoring is given guidance in the procedures and tested for specific anosmia (lack of sensitivity) to the mercaptan family of odours. This data is of help in interpreting the significance of the complaint data. In addition to describing the odour, monitors are asked to quantify the scale and offensiveness of the odour, the time of monitoring, and general weather conditions. The monitors are trained to assess the likely cause of the odour and record the certainty of the assessments as part of the process. By comparing the timing and location of complaints with the recorded incidents, an indication as to the extent to which complaints reflect the recorded experience of landfill odours can be found. In addition, the monitoring and complaints data provide real life evidence of where odour occurs and where annoyance results.

4.4.2 Conclusions

This study has examined the influence of different averaging times on modelled odour dispersion from a landfill site. These modelled results were compared with a community monitoring database that reports incidents of odour detected in the areas surrounding the landfill site. We have shown that the current regulatory method of dispersion modelling, using hourly averaging times, is less successful at capturing peak concentrations, and does not capture the pattern of odour emission as indicated by the community monitoring database. The use of short averaging times produces a modelled pattern of dispersal that more closely matches the observed database. This approach is therefore of greater value in predicting the likely nuisance impact of an odour source and in framing appropriate regulatory controls.

CHAPTER 5. Statistical Survey on malodour in the area of Este (Padua)

5.1 Foreword

The University of Padua, in collaboration with eAmbiente Ltd, has carried out a statistical sensory and device-based survey on odour perceptions, to evaluate the impact on local population of various odorous sources located in the territory of two towns, Este and Ospedaletto Euganeo.

Specifically the objective of this survey is to monitor, from a statistical viewpoint, odour perceptions of local citizens in order to:

- Determine the impact of different odour categories on citizens' perception;
- Measure the temporal evolution, also depending on atmospheric and climatic changes;
- Provide a map of this phenomenon, taking into account the area involved and the seasonal period;
- "Quantify" the subjective and objective impact (or "weight") of the perceived problem, differentiating the annoyance, particularly by manifestation area.

The survey lasted 12 months, starting on February 15 2010 and ending on February 6 2011.

5.2 Territorial Boundary

As mentioned above, this study takes place in the territories of two towns, Este and Ospedaletto Euganeo. In Figure 5.1 the actual boundaries of the area under study are shown with an orthophoto.

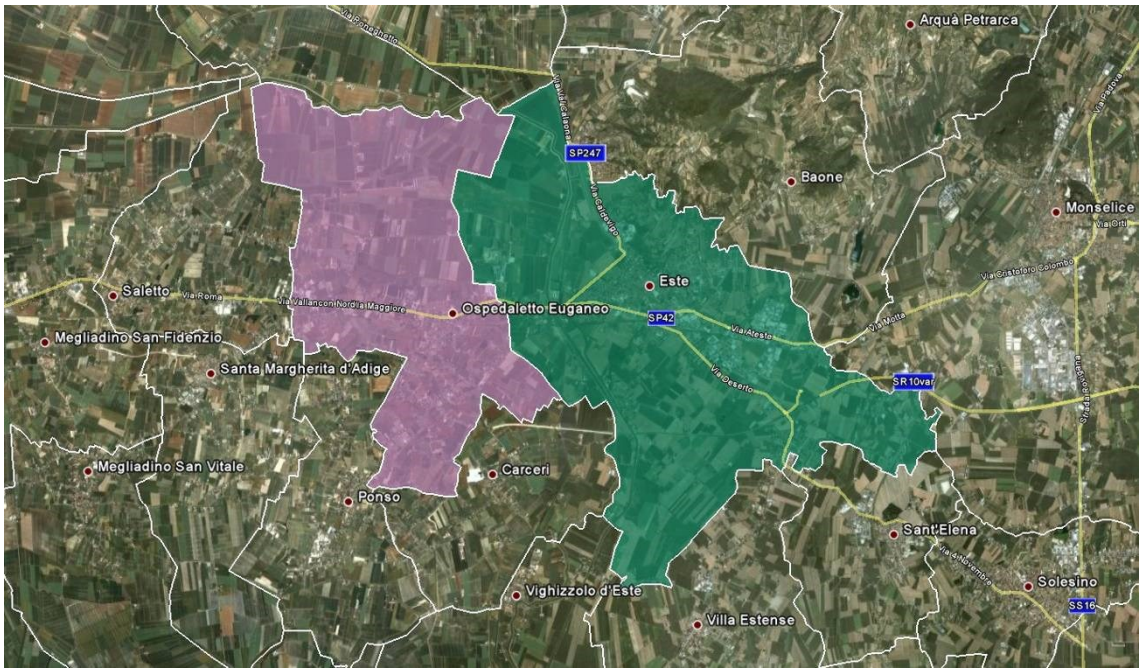


Fig. 5.1 Orthophotographic boundaries of Este and Ospedaletto Euganeo (source Google Earth)

5.3 Odour sources in the area investigated

The odour sources which have an impact on the territory being studied, come from livestock farms and from plants operating locally.

Regarding the first sources, many stock raising farms were identified (broilers, ducks, cattle, rabbits, pheasants, guinea fowls, sheep, turkeys, geese and swine), and a couple of farms for weaning. Overall there are 49 livestock farms, where 28 in the municipal area of Este, and 21 in the municipal area of Ospedaletto Euganeo.

Notice that the census of the farms mentioned above was obtained from the Veterinary Office of the local National Health Service department.

Regarding the second odour sources, the main sources identified are a feed mill, a cement factory and a solid waste treatment landfill.

5.4 Description of Activities

The current survey consisted of two main activities:

1. Statistical sensory analysis based on the reporting of odour perceptions by the citizens of the town of Este who took part in the survey (these people are called *sniffers* from now on). The reporting happens by filling in a questionnaire specifically prepared by the University of Padua and available online at a website (alternatively, a printed document is available).
2. Olfactometric campaigns carried out in the municipal territories of Este and Ospedaletto Euganeo, in the way explained later. This activity was done by

eAmbiente Ltd, who also elaborated meteorological data for the climatic characteristics of the area examined.

Table 5.1. Division of municipal provinces of Este and Ospedaletto Euganeo into survey areas

Town	Survey area N°	Description
Este	1	Este – City Centre
	2A	Meggiaro Basso - Este nuova
	2B	Meggiaro Alto
	3	Pilastro - Salute - Torre
	4A	Motta – Industrial and manufacturing area
	4B	Schiavonia
	5	Deserto
	6	Prà
Ospedaletto Euganeo	7	Ospedaletto – Centre
	8	Palugana - Tresto – Peagnola
	9A	Vallancon
	9B	Santa Croce – Dossi

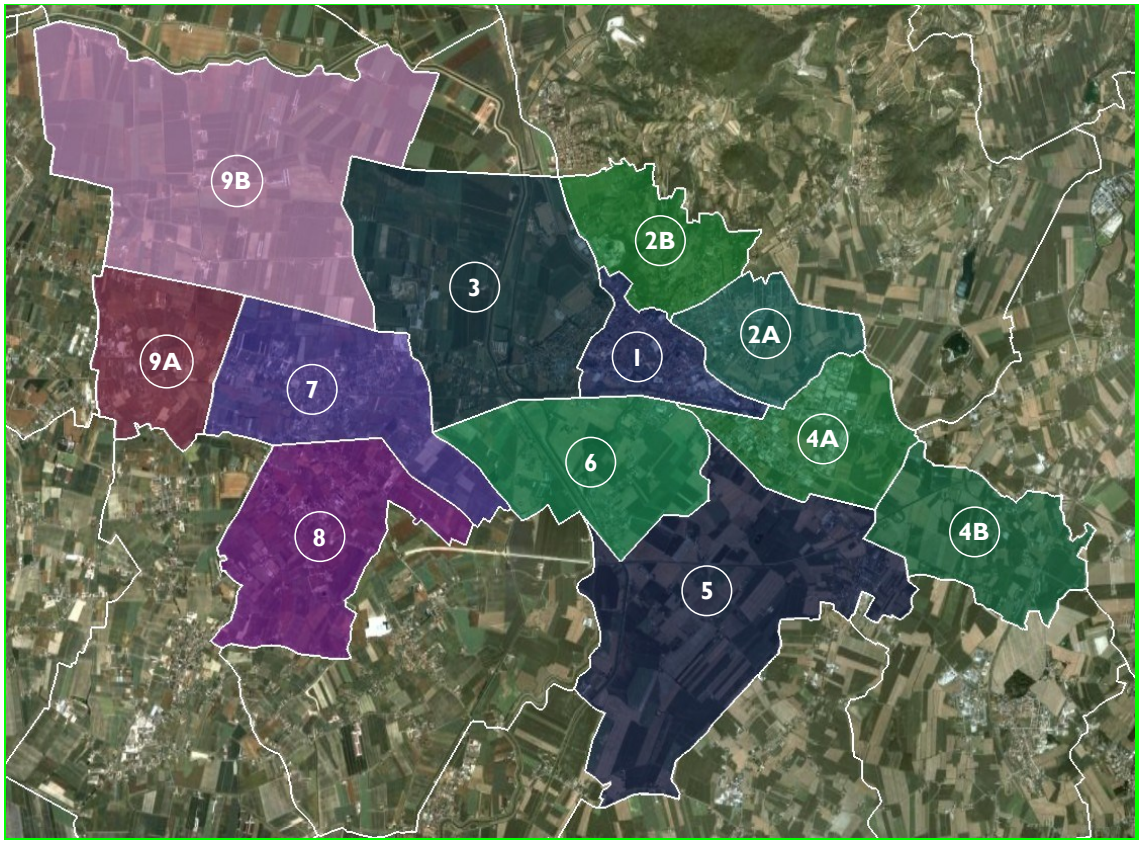


Fig. 5.2 Division of municipal provinces of Este and Ospedaletto Euganeo into survey areas

To realize the olfactometric campaigns, the territories of the two towns examined have been divided into 12 survey areas, listed and described in table 5.1, while in Figure 5.1 their exact locations are shown in an orthophoto. In every area we identified the appropriate location for a sensory analyzer ("electronic nose"), in order to continuously analyze, over the period of a week, the atmospheric air, and identify its characteristics in terms of odorous substances. The location is made available by a private citizen, or by the City Council. The planned duration for this study is one year, beginning in mid-February 2010, and divided into 4 trimesters. Every trimester is then subdivided into survey weeks. During the whole duration of the study, sniffers must continuously and periodically signal the presence of odours, and report a series of information on odorous perceptions, by filling in a questionnaire. The device, instead, will be initially installed in the chosen location of area n.1, where it will stay for a week. Then, the analyzer will be moved to the location in area n.2, where it will stay another week, and so on until all area in Table 5.1 are completed.

Partly beforehand, and partly during the sensorial investigation, the electronic nose has been "trained", that is the main sources of odour in the area (stock farms, feed mills, the waste landfill and the cement factory) have been characterized and memorized in the device itself. This training is necessary to be able to compare air samples analyzed in the different locations where the sensor is positioned, with typical odour sources, and determine possible matches. In the analysis and evaluation of results of the olfactometric study, we have also taken into account the local meteo-climatic conditions which characterize the area (wind profile).

5.4.1 Survey Scheduling

In table 5.2 the official survey schedule is presented, divided by areas. The areas belonging to the town of Ospedaletto Euganeo have been highlighted in azure.

Table 5.2. Survey schedule (continues onto next page)

FIRST TRIMESTER				SECOND TRIMESTER			
1	15-21 Feb	Area 1	Este – City Centre	1	17-23 May	Area 1	Este – City Centre
2	22-28 Feb	Area 2A	Meggiaro Basso–Este nuova	2	24-30 May	Area 2A	Meggiaro Basso–Este nuova
3	01-07 Mar	Area 3	Pilastro–Salute–Torre	3	31-06 Jun	Area 3	Pilastro–Salute–Torre
4	08-14 Mar	Area 2B	Meggiaro Alto – Hilly Area	4	07-13 Jun	Area 2B	Meggiaro Alto – Hilly Area
5	15-21 Mar	Area 4A	Motta– Industrial Area.	5	14-20 Jun	Area 4A	Motta– Industrial Area

6	22-28 Mar	Area 4B	Schiavonia	6	21-27 Jun	Area 4B	Schiavonia
7	29-04 Apr	Area 5	Deserto	7	28-04 Jul	Area 5	Deserto
8	05-11 Apr	Area 6	Prà	8	05-11 Jul	Area 6	Prà
9	12-18 Apr	Area 7	Ospedaletto - Centre	9	12-18 Jul	Area 7	Ospedaletto - Centre
10	19-25 Apr	Area 8	Palugana–Tresto–Peagnol	10	19-25 Jul	Area 8	Palugana–Tresto–Peagnol
11	26-02 May	Area 9A	Vallancon	11	26-01 Aug	Area 9A	Vallancon
12	03-09 May	Area 9B	Santa Croce–Dossi	12	02-08 Aug	Area 9B	Santa Croce–Dossi
THIRD TRIMESTER				**FOURTH TRIMESTER**			
1	16-22 Aug	Area 1	Este – City Centre	1	15-21 Nov	Area 1	Este - City Centre
2	23-29 Aug	Area 2A	Meggiaro Basso–Este nuova	2	22-28 Nov	Area 2A	Meggiaro Basso–Este nuova
3	30-05 Sept	Area 3	Pilastro–Salute–Torre	3	29-05 Dec	Area 3	Pilastro–Salute–Torre
4	06-12 Sept	Area 2B	Meggiaro Alto– Hilly Area	4	06-12 Dec	Area 2B	Meggiaro Alto– Hilly Area
5	13-19 Sept	Area 4A	Motta– Industrial Area	5	13-19 Dec	Area 4A	Motta – Industrial Area
6	20-26 Sept	Area 4B	Schiavonia	6	20-26 Dec	Area 4B	Schiavonia
7	27-03 Oct	Area 5	Deserto	7	27-02 Jan	Area 5	Deserto
8	04-10 Oct	Area 6	Prà	8	03-09 Jan	Area 6	Prà
9	11-17 Oct	Area 7	Ospedaletto Centro	9	10-16 Jan	Area 7	Ospedaletto – Centre
10	18-24 Oct	Area 8	Palugana–Tresto–Peagnol	10	17-23 Jan	Area 8	Palugana–Tresto–Peagnol
11	25-31 Oct	Area 9A	Vallancon	11	24-30 Jan	Area 9A	Vallancon
12	01-07 Nov	Area 9B	Santa Croce–Dossi	12	31-06 Feb	Area 9B	Santa Croce–Dossi

5.5 Survey Methodology

5.5.1 Sensory Statistical Survey on Odour Perceptions

To better present and summarize the survey results, some areas have been joined, so the Este province has been divided into 9 areas, where 6 belong to the town of Este and 3 to the town of Ospedaletto Euganeo (see Table 5.3 and Figure 5.3).

This unification was suggested by the division of the province of Este in districts and villages, bearing in mind:

1. homogeneity with regard to the issue of odours (similar exposure to odorous sources);
2. morphology of the territory;

3. density of local population.

Table 5.3. Division of the provinces of Este and Ospedaletto Euganeo into 9 areas

Town	Area N°	Description
Este	1	Este – City Centre
	2	Meggiaro - Este Nuova
	3	Pilastro - Salute - Torre
	4	Motta - Zona industriale e artigianale - Schiavonia
	5	Deserto
	6	Prà
Ospedaletto Euganeo	7	Ospedaletto – Centre
	8	Palugana - Tresto – Peagnola
	9	Vallancon - Santa Croce - Dossi

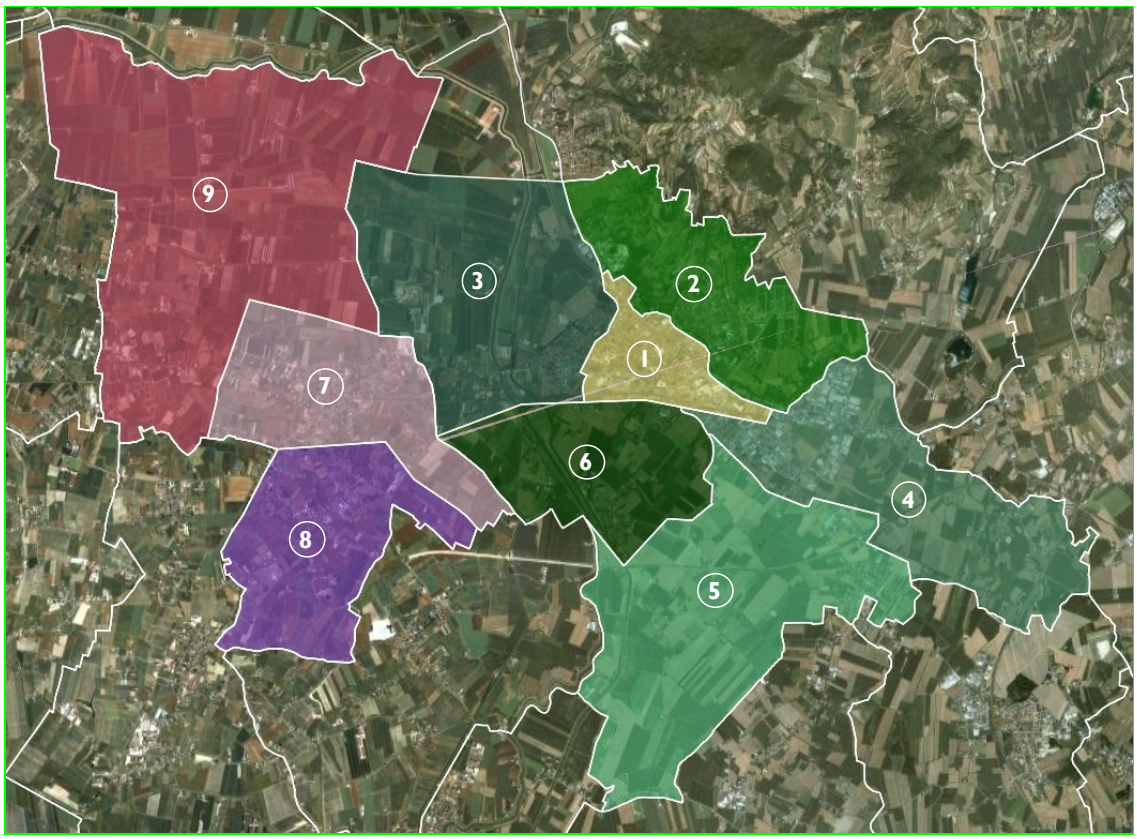


Fig. 5.3 Division of the provinces of Este and Ospedaletto Euganeo into 9 areas

In Table 5.4 you can see the detailed list of statistical units who took part in the survey (sniffers), for each of the 9 areas under examination. Notice that each sniffer represents a household, therefore it consists of at least a person, but normally there would be more than one housemate.

Table 5.4. Number of sniffers in each of the 9 areas examined

Area N°	Area	Count of statistical units (sniffers)
1	Este – City Centre	78
2	Meggiaro–Este nuova	45
3	Pilastro–Salute–Torre	48
4	Motta– Industrail Area – Schiavonia	16
5	Deserto	22
6	Prà	12
	TOTAL FOR ESTE	**221**
7	Ospedaletto – Centre	43
8	Palugana–Tresto–Peagnola	6
9	Vallancon–Santa Croce–Dossi	15
	TOTAL FOR OSPEDALETTO EUGANEO	**64**
	OVERALL TOTAL	**285**

ODOUR PERCEPTION INDEX (IPO)

In order to better study the phenomenon of odour perceptions, to evaluate their temporal trend and to compare reports done at different times in different areas, we have defined a synthetic indicator and called it Odour Perception Index (IPO from the Italian acronym). The formula is:

$$\text{IPO (week x, area y)} = \frac{\sum(\textit{duration of perceptions in week x and area y})}{\textit{N of active sniffers in week x and area y}}$$

where ∑ indicates sum.

By "number of active sniffers" we mean the number of sniffers who actually file a report (even when no odour is detected), taking into account that the number of active sniffers is affected by the presence of the electronic nose in the area. In fact we noticed that, when all other factors are unchanged, during the week when the device is present in an certain area, the relevant sniffers tend to file a greater number of reports, compared with the weeks when the nose is not positioned in the area.

To better understand the meaning of this indicator, we present some possible values for IPO, and their explanation:

- **IPO=0** (*minimum value*): there was no report during the whole week (the duration sum is 0); therefore, as soon as there is at least one report, IPO has a value greater than zero.

- **IPO=0.5**: the sum of the duration of all odour reports during a week is equivalent to half a day (12 hours).
- **IPO=1.0**: the sum of the duration of all odour reports during a week is equivalent to one day (24 hours).
- **IPO=7** (*maximum value*): all sniffers reported the presence of malodour for an overall duration of 7 days out of 7 (uninterruptedly, for every hour of every day of the week).

Briefly, the greater IPO is, the longer the odour perception lasted. Notice that IPO takes into account the number of reports, their durations and the number of sniffers present and really active in the area under scrutiny.

5.5.2 Instrumental Sensory Survey

SENSORY MEASUREMENTS

Instrumental sensory analysis enables to evaluate odour objectively via an electronic and mathematical simulation that reproduces the human process of olfactory appraisal. The use of this technique allows detecting and classifying odours caused by volatile chemical substances coming from various sources.

Instrumental sensory measures were taken with Airsense analyzer, model PEN3 (see Figure 5.4), able to determine odorous prints of samples in aerial, liquid or solid state.

Fig. 5.4 Pictures of the sensory analyzer ("electronic nose") used in the olfactometric survey

The device simulates the mental process of memorization and identification of the human olfactory system, via 10 sensors coated with a metallic film, thermostatically treated at different specific temperatures, sensitive to different classes of compounds.
More specifically, the instrument uses MOS-type sensors (Metal Oxide Sensors), properly treated (Sn, Pd or Ir metal layer, or a blend of these metals) in order to be sensitive to different classes of compounds.
A very important and peculiar characteristic of the measurement system used in PEN Airsense, is the programming of each sensor at a different and specific temperature (ranging from 150 to 500° C); this allows broadening the range of perception of volatile substances having an odorous effect.
The use of this measurement technique aims to classify odours (memorizing the digital print of the sample), which are the result of a mixture of innumerable compounds, often present in very low concentration but with a high olfactory threshold.
In other words, the array of sensors, being sensitive to different substances, can "read" the sensory impact of the sample, showing the result of contemporary measures of ten sensors.
The sampling is carried out by drawing the air to be analyzed with a pump, into a measure cell. The measurement cycle has an overall duration of 500 sec, and consists of two phases:

- The **washing** phase: 400 sec (the air pumped is purified with an activated carbon filter);
- The **reading** phase: 100 sec.

Before starting any measurement, the operator is supposed to select and analyze odorless air, which represents the "blank" state.
Thanks to the simultaneous measures of all sensors, sent as a digital signal to the software for processing, the sensory print of the odour is calculated, and compared for matching with the air samples of a database previously populated. This way it is possible to correlate the odour measured by the device with the relative source.
The instrument is equipped with an anemometer to measure wind speed and direction (specifically, the anemometer takes a measurement every 30 seconds).

DEFINING THE PATTERN

The results on all samples measured with the sensory analyzer, are then compared and correlated with the "patterns", that is the odour prints previously measured in the vicinity of examined sources and recorded onto the device.

This comparison enables us to match or exclude the origin of odours from a source being examined.

TRAINING THE SENSORY ANALYZER

In order to "train" the sensory analyzer ("electronic nose"), so as to able to identify odour prints from the main sources in the municipal territories of Este and Ospedaletto Euganeo, eAmbiente Ltd carried out specific survey campaigns, with the presence of a veterinary from the local NHS department.

These campaigns consisted in taking environmental air samples near the sources identified (farms/plants). Within the 24 hours following each collection, samples were subjected to olfactometric analysis in a lab, using the sensory analyzer.

Three campaigns were completed, and more specifically executed in May 2010, August 2010 and November 2010; as a consequence, a total of 25 odour sources were fully characterized.

Table 5.5 shows the list of sources sampled during the survey campaigns to train the sensory analyzer. Notice that plants/farms located in the province of Ospedaletto Euganeo are highlighted in blue. The remaining sources are located in the province of Este.

Figures 5.5 and 5.6 show the locations of such sources in orthographic photos.

Regarding the Solid Waste Treatment plant (source n.6), we investigated the sites considered more meaningful, that is the composting plant (6A), the biofilter used with the maturing compost (6B), the biofilter used in the oxidization of compost (6C) and the area utilized for maturing green produce (6D).

With reference to source n.9, samples were taken both in the area where poultry manure had just been spread (9A), and in the area with raising animals (9B).

Finally, even for the feed mill two odour sources were selected and characterized, one where the smell of oil was prevalent (10A), and another where the smell of fat was prevalent (10B).

Table 5.6 shows main characteristics of raised animals which have been sampled: number of units present in the farm, number of units in the area surveyed and duration of stay for the animals.

Table 5.5. Description of sampled odour sources

Town	Source N°	Source Type
Este	1	Cattle (calves)
	2	Poultry livestock (turkeys)
	3	Poultry livestock (turkeys)
	4	Poultry livestock (turkeys)
	5	Broilers (poults)
	6A	Composting plant
	6B	Biofilter with maturing compost
	6C	Biofilter with oxidizing compost
	6D	Area for maturing green produce
	7	Cement factory
Ospedaletto Euganeo	8	Swine breeding
	9A	Poultry manure spreading
	9B	Poultry livestock (turkeys)
	10A	Feed mill (oil)
	10B	Feed mill (fat)
Este	11	Cattle (baby-beef)
	12	Poultry livestock (turkeys)
	13	Poultry livestock (turkeys)
	14	Poultry livestock (turkeys)
Ospedaletto Euganeo	15	Poultry livestock (turkeys)
	16A	Poultry stock (guinea fowls)
	16B	Poultry stock (ducks)
	17	Poultry livestock (turkeys)
	18	Poultry livestock (broilers)
	19	Swine breeding

Table 5.6. Main characteristics of sampled animals

Source N°	Source Type	Total Units	Units present in the sampled area	Duration of animals' stay
1	Cattle (calves)	not avail.	120 (calves)	90 days
2	Poultry stock (turkeys)	8.000	4.000	125 days
3	Poultry stock (turkeys)	26.900	4.400	127 days
4	Poultry stock (turkeys)	19.000	4.500	130 days
5	Poultry Broilers (poults)	40.300	40.300	Male: 40 days Female: 60 days
8	Swine breeding	400 sows + 1.000 piglets	40 sows	3 years
9A	Poultry manure just swept	-	-	After 5 months of stay
9B	Poultry stock (turkeys)	not avail.	not avail.	150 days
11	Cattle (baby-beef)	300	150	1 year
12	Poultry stock (turkeys)	15.000	7.500	15 days
13	Poultry stock (turkeys)	16.000	4.500	70 days
14	Poultry stock (turkeys)	11.000	5.500	90 days
15	Poultry stock (turkeys)	13.000	6.500	75 days
16A	Poultry stock (guinea fowls)	80.000	20.000	8 days
16B	Poultry stock (ducks)	-	-	Animals being moved
17	Poultry stock (turkeys)	15.000	7.500	92 days
18	Poultry stock (broilers)	26.000	14.000	36 days
19	Swine breeding	2.000	300	4-5 months

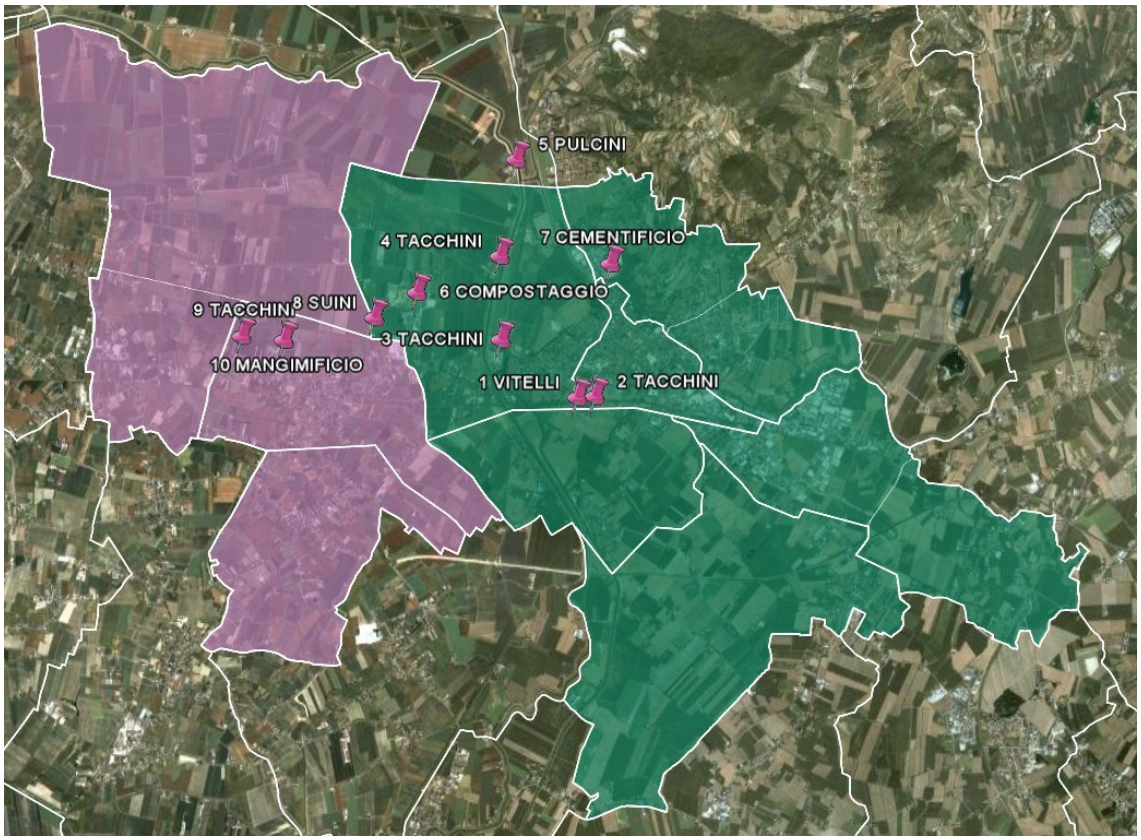


Fig. 5.5 Location of sampled odour sources with orthophoto (sources 1-10)

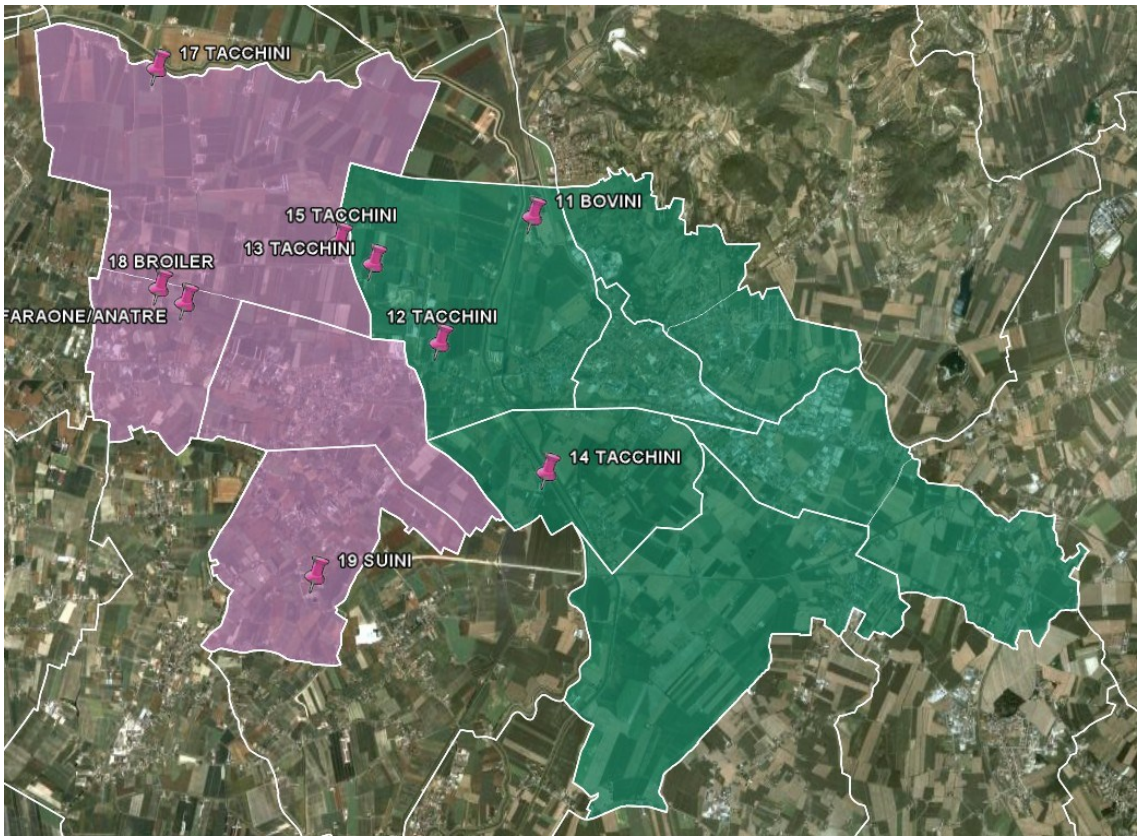


Fig. 5.6 Location of sampled odour sources with orthophoto (sources 11-19)

MEASUREMENT POSITIONS

Table 5.6 shows the list, for each surveyed area, of the measurement positions where the sensory analyzer was located. The table also reports the exact address and period of sampling. Positions located in the province of Ospedaletto Euganeo are highlighted in blue.

In areas 2B and 3 the sampling could not be carried out, since the instrument was being repaired. Therefore those areas will start to be sampled from the second trimester.

Figure 5.7 indicates the location of each measurement position in an orthophoto.

Table 5.6. Location of measurement positions

Town	Area N°	Area description	Address of measurement position	Period of sampling
Este	1	Este – City centre	piazza Maggiore, 6	18/2 - 22/2/2010
	2A	Meggiaro Basso - Este nuova	via G. Di Vittorio, 43	22/2 - 1/3/2010
	3	Pilastro - Salute - Torre	(*)	1/3 - 8/3/2010
	2B	Meggiaro Alto – Hilly area	(*)	8/3 - 15/3/2010
	4A	Motta – Industrial and manufacturing areas	via Rana Borgofuro, 6	15/3 - 22/3/2010
	4B	Schiavonia	via Bosco Crosara, 1	22/3 - 29/3/2010
	5	Deserto	via Adige, 11	29/3 - 6/4/2010
	6	Prà	via Guola Larga, 10	7/4 - 12/4/2010
Ospedaletto Euganeo	7	Ospedaletto - Centre	via IV Novembre, 4	13/4 - 19/4/2010
	8	Palugana - Tresto - Peagnola	strada Carceri, 10	19/4 - 26/4/2010
	9A	Vallancon	via Vallancon Nord, 63	26/4 - 3/5/2010
	9B	Santa Croce – Dossi	Via Boccadespin, 14	3/5 - 10/5/2010

(*) electronic nose undergoing maintenance

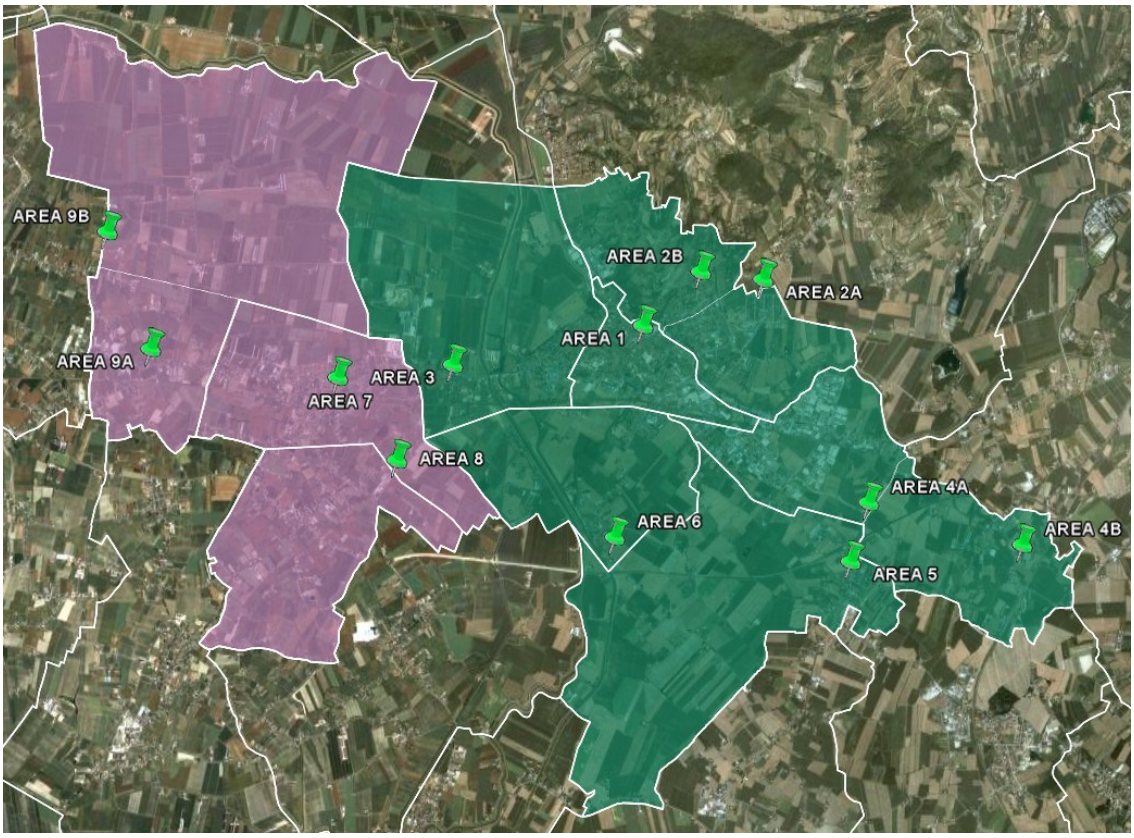


Fig. 5.7 Location of measurement positions on an orthophoto

INDEX OF AFFINITY (IA)

In order to show in a clear way the results obtained with the sensory analyzer, we calculated a specific daily mean index, designated **Index of Affinity** (IA). This parameter, calculated with statistical analysis, is an indicator of the discriminating capability of the instrument: it indicates how much the odour prints of air analyzed with the device in each period can be related to, and thus match the prints of the odour sources recorded. The Index of Affinity, calculated in relation to each odour source previously characterized, is a non-dimensional number ranging from 0 to 1.

5.6 SUMMARY RESULTS ON ODOUR REPORTS IN THE FIRST TRIMESTER

This section contains the main summary results on odour reports recorded in the first trimester of the survey, thus from February 15 2010 until May 9 2010. In particular, the number of reports in each town is reported, divided by odour taxonomy and potential source of the malodour. Then, in order to better study the phenomenon of odorous perceptions, to appreciate their temporal evolution and to be able to compare reports done at different times in different areas, we will show a synthetic indicator called Odour Perception Index (IPO). Finally, as an effective and brief tool to summarize odorous perceptions in the span of the whole trimester, we will show a series of bubble charts which enable a quick comparison among different areas.

5.6.1 Number of reports, Odour Type and Source

The first synthetic survey result consists in reporting the number of odour perceptions, divided by town and week of accounting (see Table 5.7). In order to better describe and analyze the temporal trend of the detections, Table 5.7 also indicates duration and average intensity of recorded reports.

The average duration was calculated by assigning a score to each answer concerning the duration of an odorous perception, when the odour is detected continuously. The scale is: 1="less than a minute", 2="a few minutes", 3="a few hours", 4="half a day", 5="the whole day", 6="the whole night". We calculated the arithmetic mean of the scores, and the average duration is always between 2 and 3, that is between a few minutes and a maximum of some hours.

Similarly, we created a scale for the intensity: 1="light", 2="moderate", 3="strong", 4="very strong". Then we calculated the arithmetic mean which always presents a value between 2 and 3, so from "moderate" to "strong", except for the second week where, in Ospedaletto Euganeo the average intensity was between "strong" and "very strong".

It is important to notice how stable the situation is: in fact, the phenomena observed fall into a precise interval, both in terms of duration and intensity.

On the contrary, a datum that has a lot of variability is the number of reports: to assess this measure, many factors should be considered, for all the number of sniffers (see Table 5.4). To take into account the diversity in the number of sniffers per area, we will use an indicator called Odour Perception Index (IPO).

Table 5.8 shows the number of odour reports, divided by type and possible source, and also separately for each town.

Table 5.7. Odour reports: number, average duration and intensity, week by week in each town

Town	Week	N° of reports	Avg duration	Avg intensity
Este	15/02/2010 - 21/02/2010	83	2,96	2,32
	22/02/2010 - 28/02/2010	54	2,56	2,24
	01/03/2010 - 07/03/2010	52	3,02	2,27
	08/03/2010 - 14/03/2010	44	3,02	2,34
	15/03/2010 - 21/03/2010	74	2,78	2,54
	22/03/2010 - 28/03/2010	47	2,57	2,45
	29/03/2010 - 04/04/2010	54	2,69	2,46
	05/04/2010 - 11/04/2010	63	2,62	2,43
	12/04/2010 - 18/04/2010	33	2,58	2,64
	19/04/2010 - 25/04/2010	68	2,55	2,59
	26/04/2010 - 02/05/2010	29	2,80	2,69
	03/05/2010 - 09/05/2010	15	2,92	2,93
Este Total		**616**	**2,75**	**2,45**
Ospedaletto Euganeo	15/02/2010 - 21/02/2010	4	2,00	2,50
	22/02/2010 - 28/02/2010	5	2,20	3,20
	01/03/2010 - 07/03/2010	5	2,20	2,20
	08/03/2010 - 14/03/2010	10	2,10	2,30
	15/03/2010 - 21/03/2010	28	2,21	2,82
	22/03/2010 - 28/03/2010	25	2,16	2,84
	29/03/2010 - 04/04/2010	22	2,09	2,82
	05/04/2010 - 11/04/2010	40	2,23	2,63
	12/04/2010 - 18/04/2010	86	2,38	2,67
	19/04/2010 - 25/04/2010	102	2,29	2,77
	26/04/2010 - 02/05/2010	76	2,12	2,70
	03/05/2010 - 09/05/2010	44	2,48	2,89
Ospedaletto Euganeo Total		**447**	**2,26**	**2,73**
Overall Total		**1.063**	**2,56**	**2,57**

Table 5.8. Number of odour reports, divided by type and potential source

Town	Potential Source	Odour Type								Total
		Animal related	harsh, stinging	ammo-niacal	Hay / forage	toasting	Animal manure	putrid, rotten	other smell	
Este	NOT identified	19	72	6	1	-	70	53	64	**285**
	Identified	28	107	3	1	6	43	70	73	**331**
	Poultry stock	11	6	-	-	-	11	9	11	**48**
	Cattle	3	2	-	-	-	4	1	2	**12**
	Swine breeding	1	2	-	-	-	2	1	1	**7**
	Poultry droppings spread on cropland	-	1	-	-	-	1	1	1	**4**
	Manure spread on cropland	1	3	-	-	-	1	3	1	**9**
	Municipal waste treatment plant	6	40	2	-	4	14	46	23	**135**
	Feed mill	1	-	-	-	-	3	2	2	**8**
	Cement factory	-	14	-	-	-	-	-	1	**15**
	Traffic	4	8	1	-	1	-	4	12	**30**
	Other source	1	31	-	1	1	7	3	19	**63**
Este Total		**47**	**179**	**9**	**2**	**6**	**113**	**123**	**137**	**616**
Ospedaletto Euganeo	NOT identified	14	32	1	-	5	29	41	23	**145**
	Identified	27	99	6	1	7	27	97	38	**302**
	Poultry stock	7	3	2	-	1	11	1	2	**27**
	Cattle	-	1	-	-	-	-	1	-	**2**
	Swine breeding	6	2	2	-	-	4	5	3	**22**
	Poultry droppings spread on cropland	-	2	-	-	-	3	-	2	**7**
	Manure spread on cropland	1	-	-	-	-	1	-	1	**3**
	Municipal waste treatment plant	10	86	1	1	1	7	83	18	**207**
	Feed mill	1	1	-	-	5	1	2	3	**13**
	Cement factory	1	2	1	-	-	-	2	1	**7**
	Traffic	-	-	-	-	-	-	1	1	**2**
	Other source	1	2	-	-	-	-	2	7	**12**
Ospedaletto Euganeo Total		**41**	**131**	**7**	**1**	**12**	**56**	**138**	**61**	**447**
Overall Total		**88**	**310**	**16**	**3**	**18**	**169**	**261**	**198**	**1063**

By looking at Table 5.8, we can draw two conclusions:

1. For a considerable number of reports, the source is not identified (46% in Este, 32% in Ospedaletto Euganeo); among those reports where the source is identified, the item most commonly indicated is "municipal waste treatment plant" (41% for Este, and even 69% for Ospedaletto Euganeo).
2. Regarding the type of odour, the most frequent categories are:
    i. "harsh, stinging" (29% both for Este and Ospedaletto Euganeo);
    ii. "putrid, rotten" (20% for Este, 31% for Ospedaletto Euganeo);
    iii. "animal manure" (18% for Este, 13% for Ospedaletto Euganeo).

It should also be noted that the category reported was "other smell" for 22% of reports in Este, and 14% in Ospedaletto Euganeo.

By intersecting the results on potential sources, with the types of smell, we can see that the municipal waste plant, reported as the main odour source, is believed to be the cause of a wide range of odour types, even very different ones. In other words, the municipal waste treatment plant is often indicated as the odour source, regardless of the type of smell perceived.

5.6.2 Odour Perception Index (IPO)

Figures 5.8 and 5.9 show the weekly values for IPO, separately for the different areas in the towns of Este and Ospedaletto Euganeo. At first, it appears that some areas almost constantly present the highest values of IPO, that is the city centre and Pilastro-Salute-Torre in Este, and the north area in Ospedaletto Euganeo.

Later, we will further investigate the index peak values in order to assess the intensity and consistency between bordering areas.

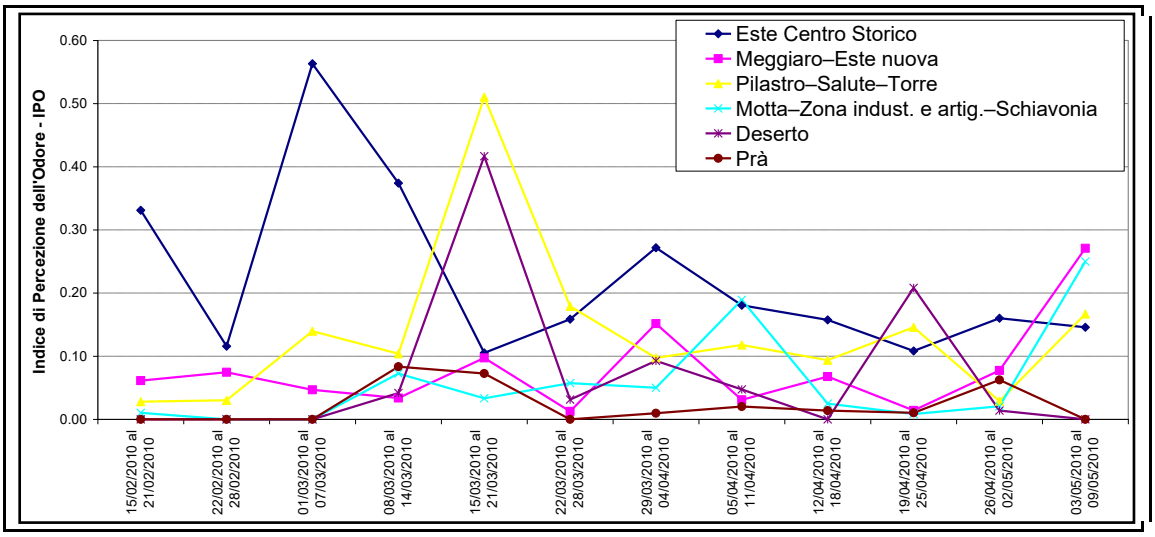


Fig. 5.8 IPO (Odour Perception Index) weekly values, for areas in the province of Este

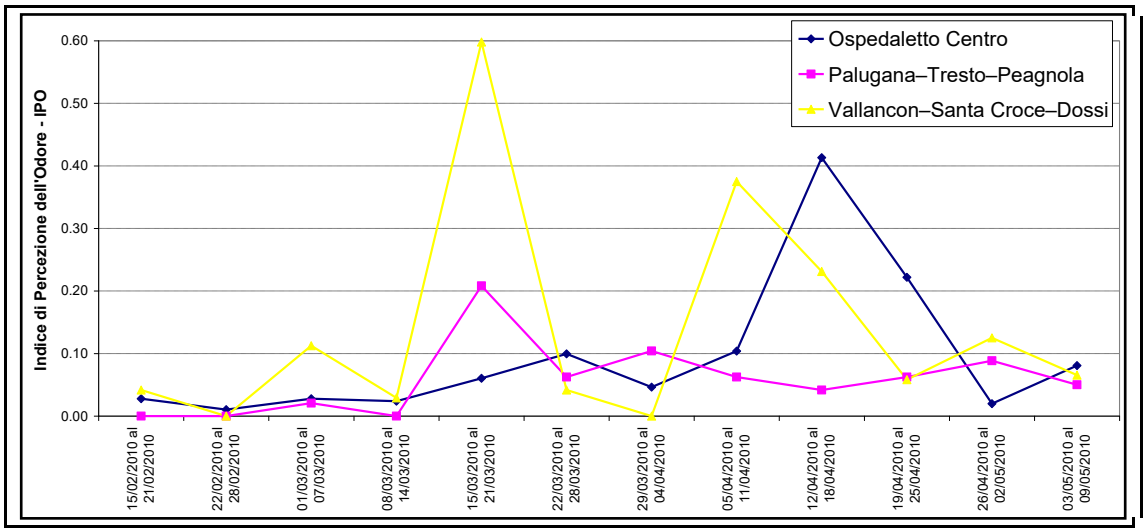


Fig. 5.9 IPO weekly values, for areas in the province of Ospedaletto Euganeo

It could be interesting to examine the IPO temporal evolution, separately for the three odour types which most frequently appear in the reports (see Table 5.8). So Figures from 5.10 to 5.13 plot the IPO weekly temporal series for the main odour types, separately for the 4 areas in the province of Este with the highest population density.

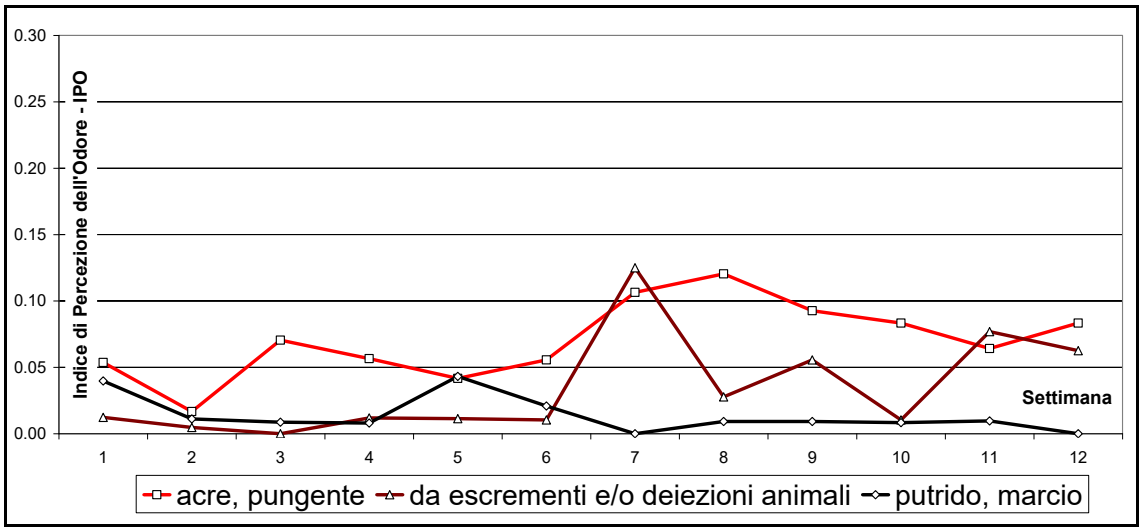


Fig. 5.10 IPO weekly values, for main odour types, in Este City Centre

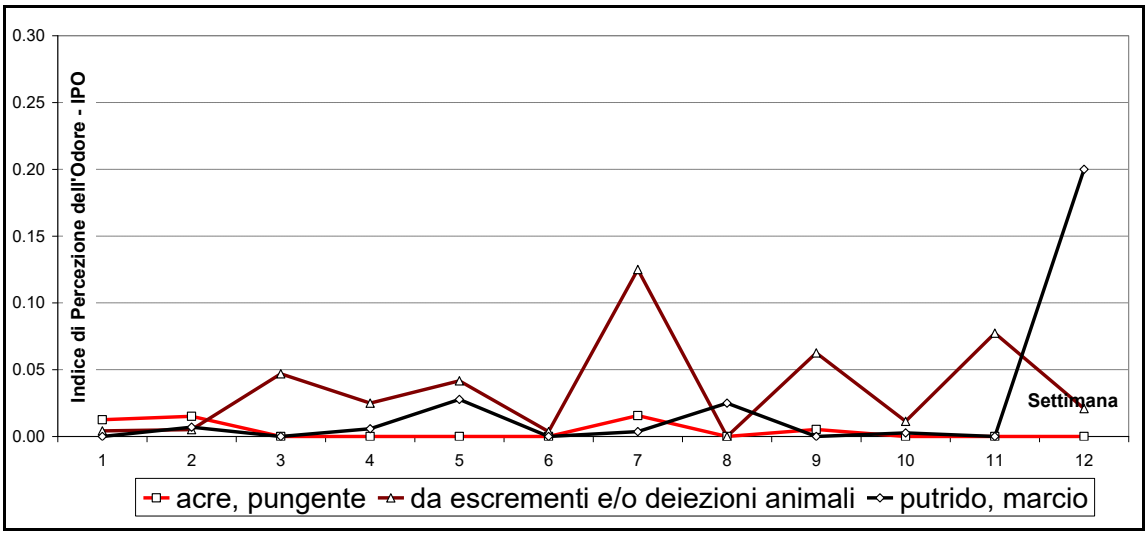


Fig. 5.11 IPO weekly values, for main odour types, in Meggiaro-Este nuova

For the area of Este – City Centre we observe that the type "harsh/stinging" (acre/pungente) has higher average values than any other category, with peaks for odours coming from animal dejections. Regarding the other areas of Este, reports seem to be oriented, in most cases, towards animal dejections instead.

The odour perceptions in the area of Ospedaletto – Centre mostly indicate both the categories mentioned above.

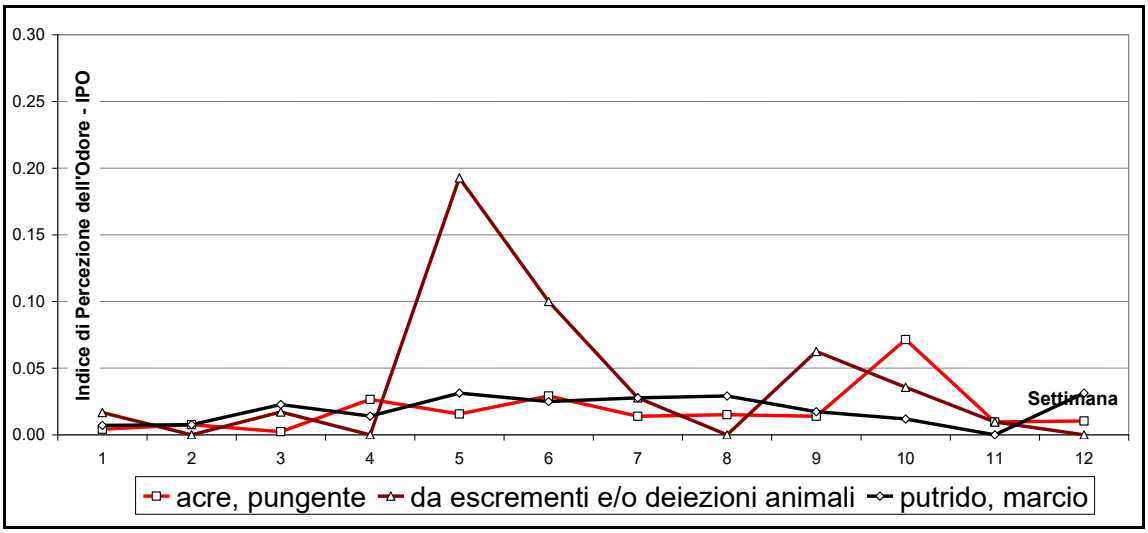


Fig. 5.12 IPO weekly values, for main odour types, in Pilastro-Salute-Torre

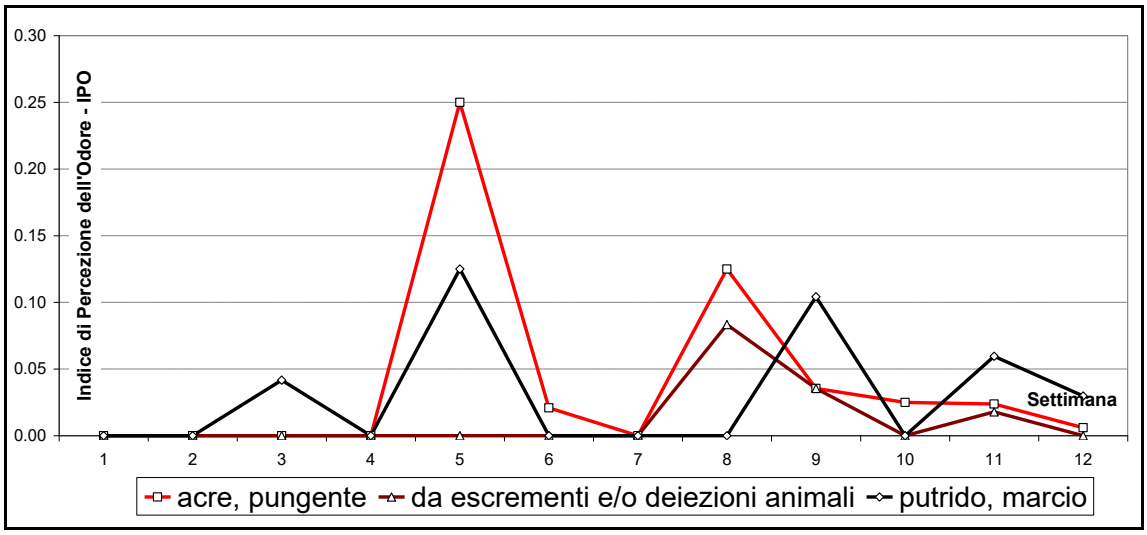


Fig. 5.13 IPO weekly values, for main odour types, in Ospedaletto Centre

5.7 Area Comparison and Cartographic Description for the First Trimester

To describe in an effective and synthetic manner the phenomenon of odour perceptions during the entire trimester, also providing the possibility of comparing areas, it is convenient to create a bubble chart, both with an overall view (Figures 5.14 and 5.15), and separately for each of the three main odour types (Figures 5.16-5.18). Such a chart should be interpreted with the following guidelines:

- along the X axis the average value for intensity is plotted, regarding all reports, or a specific odour type; therefore, the further to the right a bubble is, the more intense the odour perception;
- along the Y axis the average duration of the odour is plotted; therefore, the higher the bubble is, the longer the duration of a perception;
- the bubble diameter is proportional to the IPO index, so the larger a bubble is, the greater the impact/importance of the odour in the perception issue.

The chart is divided into 4 quadrants to easily read it. In fact:

- in the bottom left quadrant there are bubbles with low intensity and short duration; that is the "best" situation, when odours have a limited impact on perceptions;
- the bottom right quadrant contains bubbles with high intensity and short duration; this is an intermediate situation, odours are intense but short-lived;
- the top left quadrant presents bubbles with low intensity and long duration; this is another intermediate situation, with persistent but not intense odours;

- finally, the top right quadrant contains bubbles with high intensity and long duration; this is the "worst" situation, when smells have a high impact on perception, both in terms of intensity and duration.

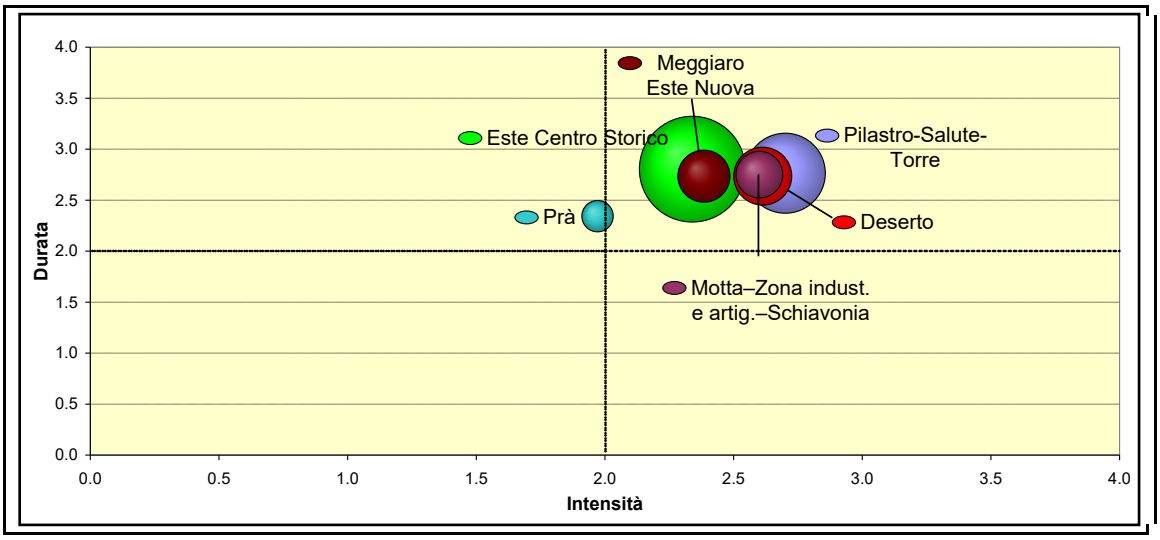


Fig. 5.14 Bubble chart for comparison in the trimester, between areas of Este

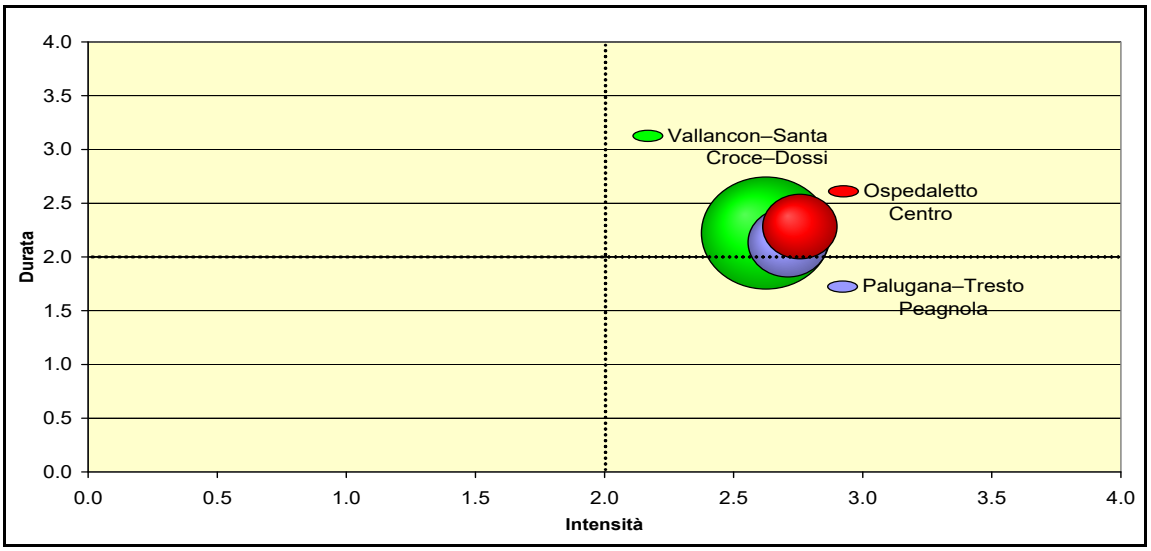


Fig. 5.15 Bubble chart for comparison in the trimester, between areas of Ospedaletto Euganeo

By analyzing Figures 5.16-5.18 we notice that different types happen to concentrate in specific areas: this allows associating an odour type with a source, as it is perceived by the population. The "harsh/stinging" type appears mostly in Este city centre; the predominance of the "putrid/rotten" category was perceived by the sniffers in the area of Meggiato-Este nuova. Thanks to the examinations that will follow, it will be interesting to note that the distribution of perceptions due to animal dejections is more homogenous in the whole territory.

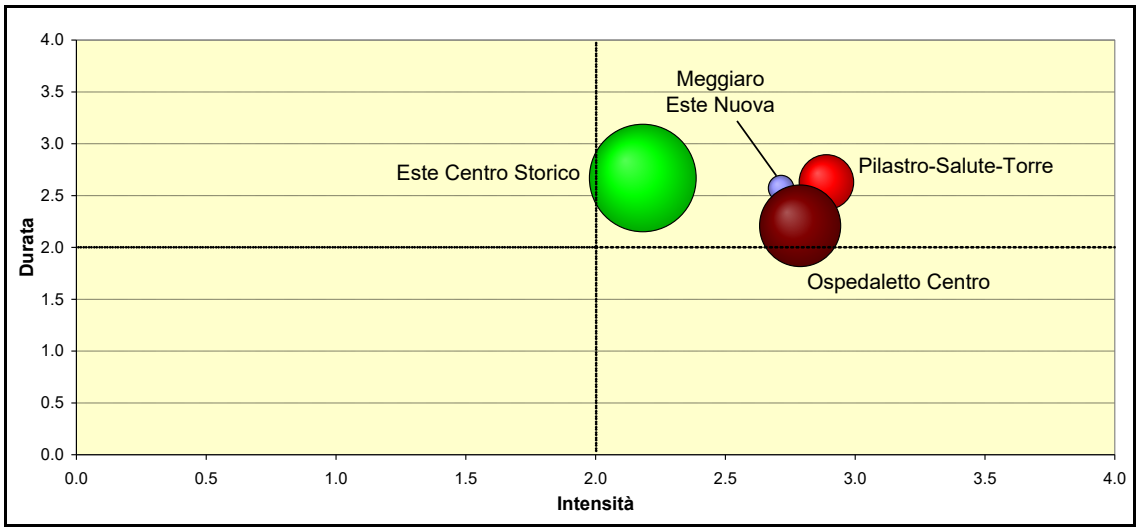


Fig. 5.16 Harsh smell: Bubble chart to compare areas in the trimester

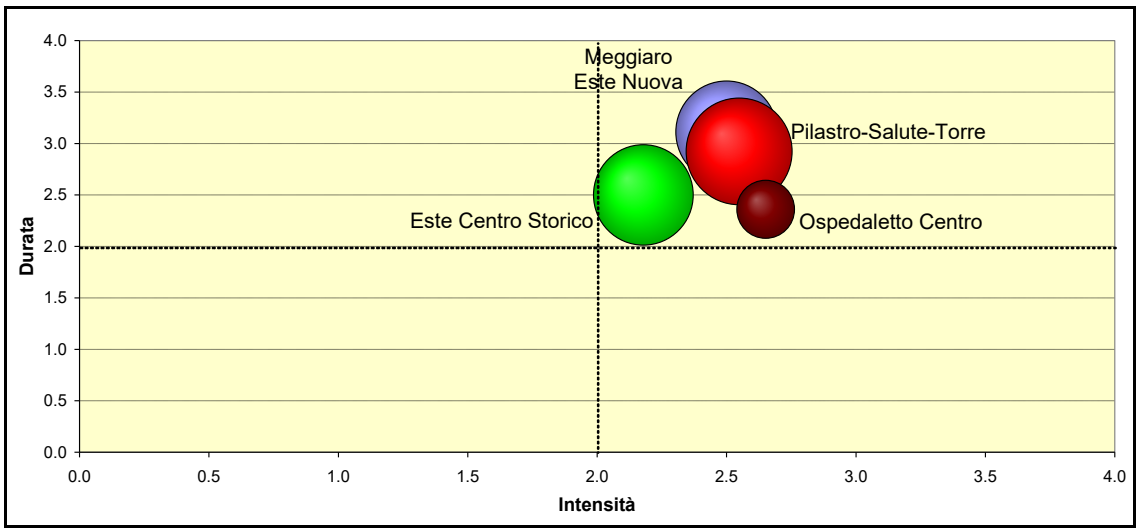


Fig. 5.17 Dejection smell: Bubble chart to compare areas in the trimester

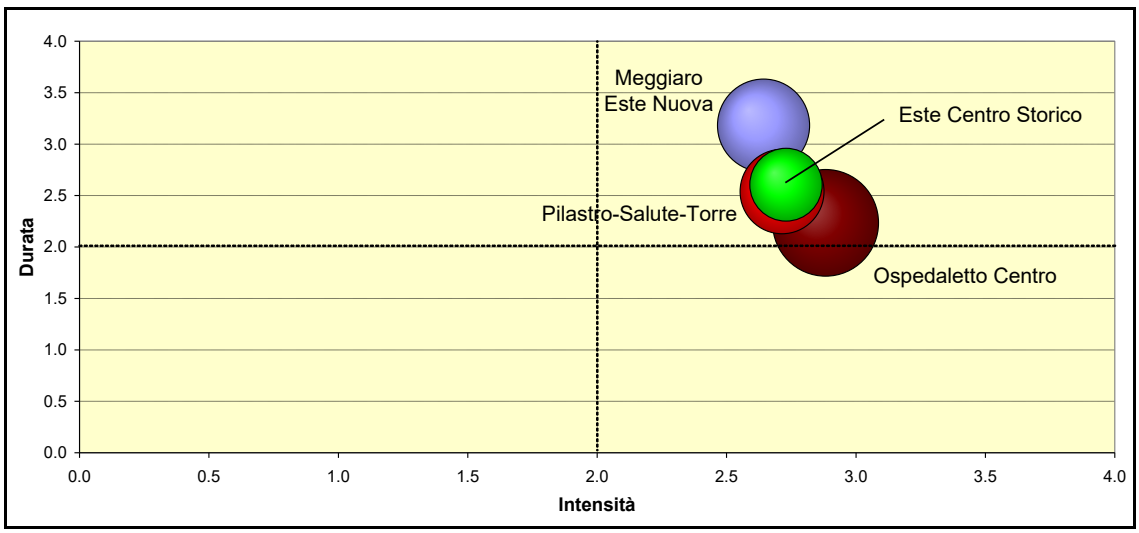


Fig. 5.18 Putrid smell: Bubble chart to compare areas in the trimester

The evolution of the IPO index for the trimester examined (weekly average values in the trimester) can be plotted with the chart in Figure 5.19.

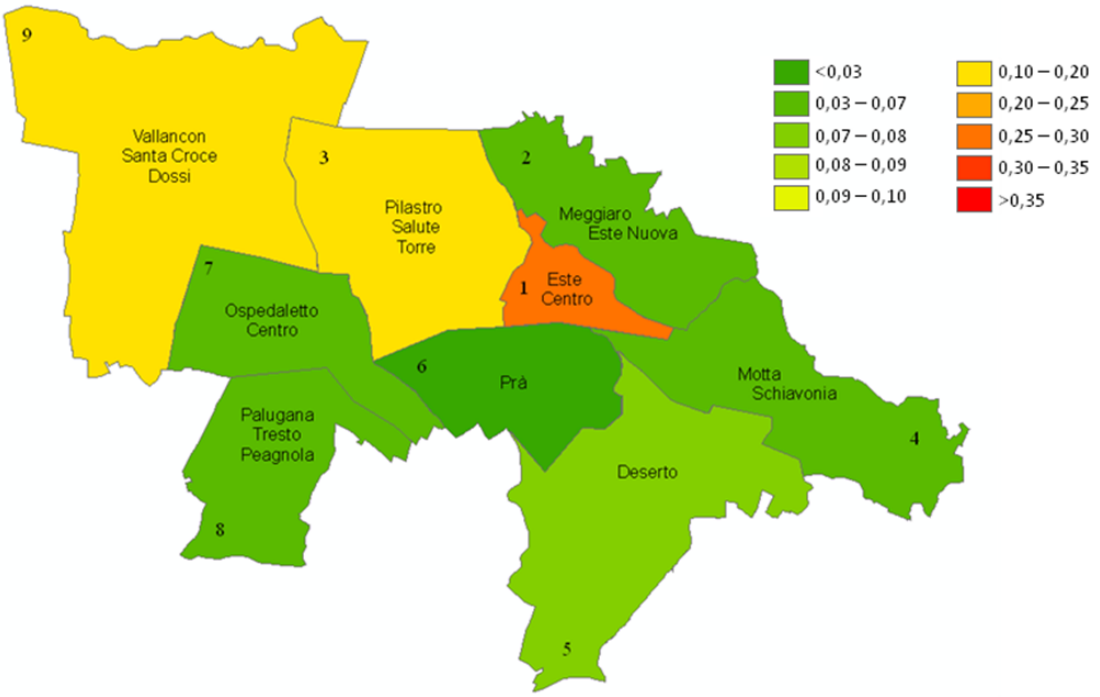


Fig. 5.19 Cartogram of IPO values in the trimester, for Este and Ospedaletto Euganeo

Studying the trimester report, it is clear that the areas where odour emissions were more intensely perceived are located in the north-east corner of the map examined, with the only exception of the centre of Este. To better understand the phenomenon, we carried out an analysis concerning type, intensity and source of perceptions.

In fact it is necessary to examine all the following three cartograms together to realize what the IPO index consists of. Analyzing a single cartogram would be misleading.

Figure 5.20 shows the most relevant odour types reported by the sniffers.

We noticed that the distribution of reports, grouped by odour type, is concentrated mainly on four items: harsh/stinging, putrid/rotten, from animal dejections and all the answers grouped under "other smell".

In particular, we noticed how the three critical areas, Vallancon-Santa Croce-Dossi, Pilastro-Salute-Torri and Este Centre, have a direct relation to the harsh/stinging category (see Fig. 5.20).

Via Figure 5.21, instead, it is possible to see how intense the perceived odours actually were. Notice the contrast between the areas of Este Centre and Ospedaletto Centre: apparently the latter presents more intense odour perceptions, but the low value of IPO underlines how short-lived they were, therefore perceived as less serious. In fact,

we would like to remind that, when calculating the IPO index, the duration of the perception has a relevant weight.

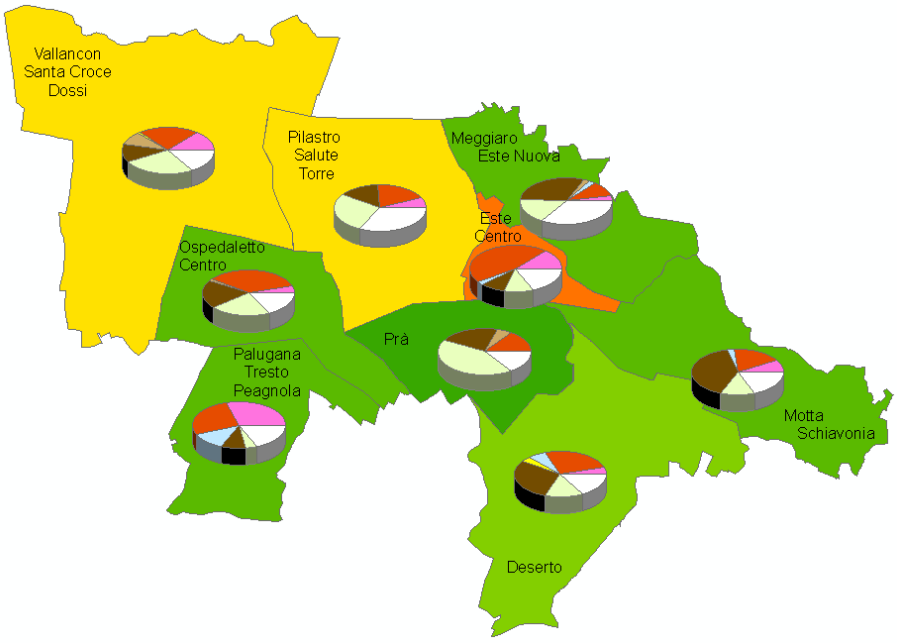


Fig. 5.20 Cartogram of IPO values in the trimester, for odour types, in the towns of Este and Ospedaletto Euganeo

Fig. 5.21 Cartogram of IPO values in the trimester and intensity, in the towns of Este and Ospedaletto Euganeo

Finally, in Figure 5.22, we show the charts plotting the frequencies of reports for each source. Ambiguity here is almost absent: the source identified by most sniffers as

potential cause of perceived smells it the municipal waste treatment plant. The data is rather homogenous in the whole territory. Two other key categories are poultry stocks and the residual item "other source". Specifically in the area of Este Centre, the latter plays a very important role with the IPO.

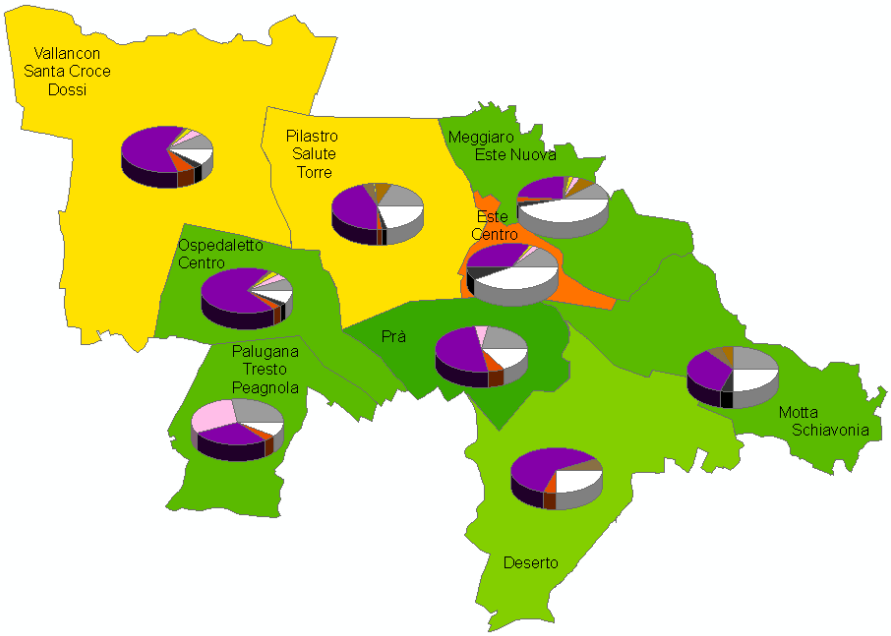


Fig. 5.22 Cartogram of IPO values in the trimester and sources, in the towns of Este and Ospedaletto Euganeo

5.8 Final Considerations for the First Trimester

The phenomenon of odorous perceptions in the towns of Este and Ospedaletto Euganeo, in the trimester studied (from February 15 2010 to May 9 2010), appeared to be relevant and widespread. Table 5.9 summarizes the intensity of the phenomenon in every single week and for each area in which the territory was divided.

The area most affected by the phenomenon is Este – City Centre, where the peak moments happened in the first, third and fourth weeks, but the situation never went under the minimum warning threshold, represented by the value of 0.10 for the Odour Perception Index (IPO), for the whole period. The phenomenon was also somewhat persistent, although less relevant and homogeneous, in the areas of Pilastro-Salute-Torri (province of Este) and Vallancon-Santa-Croce-Dossi (in the province of Ospedaletto Euganeo). Despite the pervasiveness of the phenomenon (only in the area of Prà were the values of IPO always under 0.10), it is clear that its intensity is rather varying on the territory. The odour phenomena are not perceived homogeneously in the different areas, both in terms of intensity and of time duration.

Focusing the analysis on the temporal dimension, we noticed that the peak moment in the whole period happened in the fifth week. Between March 15 and 21 six of the nine areas (four in the province of Este and two in Ospedaletto Euganeo) were affected with the phenomenon and three of them with high intensity (Pilastro-Salute-Torre, Deserto and Vallancon-Santa Croce-Dossi). Generally, odour presence was characterized by a lot of variability from week to week, possibly due to the different activities done in the stock farms and plants on the territory.

Odour types and potential sources identified by the sniffers were numerous, and again variable, depending on the area or the week examined, which suggests that possible corrective measures should not focus on a specific source but rather act in more than one direction.

Instrumental measurements supplied with the electronic nose presented some regularity, not always coherent with the evolution of odour perceptions. We believe that mainly depends on positioning the instrument in a new area every week, which prevented from measuring possible variations in odour perceptions in different areas for the same period.

More specifically, by a deeper analysis day by day, it seems that the electronic nose might show, in the area examined, a commixture of different smells originating from the different activities present on the territory. In other words, the instrument does not seem able to identify one odour type univocally, instead it identifies a mixture of smells from several sources. This observation partially agrees with what emerged from the sensory statistical survey on odour perceptions.

Still, these technical issues made it difficult to superimpose instrumental measurements with data provided by the sniffers.

Therefore we believe that electronic nose measures should be used as additional data, and not to draw comparisons with what the sniffers perceived, limitingly to the nose ability to make out only odour sources as perceived during training and specifically for the area being monitored, bearing in mind atmospheric conditions, especially with reference to wind.

In conclusion, in the first trimester of the survey we observed some intensity and diffusion of this phenomenon, but also a lot of heterogeneity both along the geographical and the temporal dimensions. The lack of homogeneity concerned not only its intensity, duration and number of reports, but also the types of smells perceived and the potential sources identified.

Table 5.9. Intensity of odour perceptions, calculated using IPO (Odour Perception Index), divided by area and week in the first trimester

Area	15-21 Feb	22-28 Feb	01-07 Mar	08-14 Mar	15-21 Mar	22-28 Feb	29-04 Apr	05-11 Apr	12-18 Apr	19-25 Apr	26-02 May	03-09 May
Este Centre												
Meggiaro-Este Nuova												
Pilastro-Salute-Torri												
Motta-Schiavonia												
Deserto												
Prà												
Ospedal. Centre												
Palugana-Tresto-Peagnola												
Vallancon-S.Croce-Dossi												

: index value between 0,10 & 0,20; : index value between 0,20 & 0,30; index value greater than 0,30.

5.9 Final Considerations for the Second Trimester

During the second trimester, spanning from May 17 to August 8 2010, reports completed by the sniffers (the citizens of Este and Ospedaletto Euganeo recruited to evaluate the impact of the phenomenon on locals, during the sensory statistical survey) showed qualitative and quantitative variations from the data collected in the first trimester.

Table 5.10 summarizes the evolution in the values of the Odour Perception Index (IPO), area by area and week by week. Odour reports, compared to the previous trimester, are fewer almost everywhere in the territory, with some exceptions and with

only one situation clearly in countertendency, Vallancon-Santa Croce-Dossi, in the province of Ospedaletto Euganeo. In this area, IPO was greater than 0.30 in three weeks: in the middle of June, and in the second half of July. In the first and third week of the trimester, that is from May 17 to May 23 and from May 31 to June 6, the entity of odour perceptions was intermediate, while in all other weeks IPO ranged between 0.10 and 0.20, apart from the period from June 28 to July 19 when perceptions were not acute.

In all other areas in the territory, IPO values were almost always less than the warning threshold of 0.10, but even when they surpassed that value, they were never greater than 0.20.

The area of Este - City Centre, which was the most critical in the previous period, witnessed a different evolution of the phenomenon which was, anyway, much less relevant in the province of Este than in the area of Ospedaletto Euganeo.

Analyzing the data on the temporal dimension, contrary to the previous trimester, we cannot identify a critical week, when odour perceptions might be more intense or geographically more widespread than other weeks. Yet there seems to be a greater number of reports in the first half of the trimester, from May 17 to June 27.

Even during this trimester types of smells and potential sources reported by the sniffers were numerous, and again variable, depending on the area or week being considered, which confirms that possible corrective measures should not focus on a specific source but rather act in different directions simultaneously.

Again, instrumental measurements with the electronic nose showed less variability, not completely consistent with the evolution of odour reports. As pointed out in the first trimester, this seems to depend mainly on positioning the device in a different area every week. Also on this occasion, deeper daily analyses show the perception of a "mixture" of various smells originating from different activities operating in the territory, which is partially in agreement with what emerged in the sensory statistical survey. These issues made it difficult to superimpose instrumental measures with the results from the sensory statistical survey.

Therefore, even for this trimester, we believe that measures recorded with the electronic nose should be considered as an additional element, not to draw a comparison with what the sniffers found, limitingly to the nose ability to make out only odour sources as perceived during training and specifically for the area being monitored, bearing in mind atmospheric conditions at the moment, especially with reference to wind.

In conclusion, the second trimester of the survey showed an evolution of the phenomenon different to the first trimester, especially in the second half. Only exception is the area Vallancon-Santa Croce-Dossi. Great heterogeneity of the phenomenon is confirmed, both geographically and temporally. The lack of homogeneity concerned not only its intensity, duration and number of reports, but also type of odours perceived and potential sources.

Table 5.10. Intensity of odour perceptions, calculated using IPO (Odour Perception Index), divided by area and week in the second trimester

Area	17-23 May	24-30 May	31-06 Jun	07-13 Jun	14-20 Jun	21-27 Jun	28-04 Jul	05-11 Jul	12-18 Jul	19-25 Jul	26-01 Aug	02-08 Aug
Este City Centre												
Meggiaro-Este Nuova												
Pilastro-Salute-Torri												
Motta-Schiavonia												
Deserto												
Prà												
Osped. Centre												
Palugana-Tresto-Peagnola												
Vallancon-S.Croce-Dossi												

: index value between 0,10 & 0,20; index value between 0,20 & 0,30; index value greater than 0,30.

5.10 Final Considerations for the Third Trimester

Table 5.11 summarizes the evolution in the values of IPO (Odour Perception Index), with reference to specific areas and weeks of the trimester. Odour perceptions are less frequent than the previous trimester, but there are rather important cases which

happened in some areas and in specific, short periods. The area of Este Centre is where IPO presented the largest values, greater than 0.30 in the last week, and settling between 0.20 and 0.30 in two other weeks of the trimester. Another area which recorded high values of the index was Pilastro-Salute-Toore where IPO surpassed the warning threshold of 0.10 in the first two weeks, and the critical value of 0.30 in the 34th week. Other notable situations (with the index below 0.30) were in Prà during week 32, and Vallancon-Santa Croce-Dossi in weeks 28 and 36. A special remark goes to Ospedaletto Centre, which records a value greater than 0.20 in one week only (the last one), but frequently (for six weeks out of twelve) the value surpasses the warning limit of 0.10. As a general trend, we can conclude that, compared to the previous period, the phenomenon has been reduced in the province of Ospedaletto Euganeo (particularly in the area of Ospedaletto Centre), while it has consolidated in the area of Este Centre.

If we analyze the phenomenon from a temporal perspective, surely the most critical weeks were the 29th (Sept 13 2010 – Sept 19 2010), the 32nd (Oct 4 2010 – Oct 10 2010), but most of all the 34th (Oct 18 2010 – Oct 24 2010) and the 36th (Nov 01 2010 – November 07 2010).

Again in this trimester types of smells and potential sources identified by the sniffers were numerous, and also variable, depending on the area and week examined, which confirms that any corrective measures should not focus on a specific source but act in more than one direction.

Again in this trimester instrumental measurements with the electronic nose were less variable, not completely coherent with the evolution of odour perceptions. We believe that, as pointed out in previous trimesters, this is due to positioning the instrument in different areas every week. On this occasion again, deeper daily analyses show the perception of a "mixture" of various smells originating from different activities operating in the territory, which is partially in agreement with what emerged in the sensory statistical survey. These issues made it difficult to superimpose instrumental measures with the results from the sensory statistical survey.

Therefore, even for this trimester, we believe that measures recorded with the electronic nose should be considered as an additional element, not to draw a comparison with what the sniffers found, limitingly to the nose ability to make out only odour sources as perceived during training and specifically for the area being monitored, bearing in mind atmospheric conditions at the moment, especially with reference to wind.

In conclusion, the third trimester of the survey showed a perception of the phenomenon different to the first two trimesters, and a strong heterogeneity has been confirmed, both geographically and temporally. The lack of homogeneity concerned not only its intensity, duration and number of reports, but also type of odours perceived and potential sources.

Table 5.11. Intensity of odour perceptions, calculated using IPO (Odour Perception Index), divided by area and week in the third trimester

Area	16-22 Aug	23-29 Aug	30-5 Sept	6-12 Sept	13-19 Sept	20-26 Sept	27-3 Oct	4-10 Oct	11-17 Oct	18-24 Oct	25-31 Oct	1-7 Nov
Este Centre												
Meggiaro -Este Nuova												
Pilastro-Salute-Torre												
Motta-Schiavonia												
Deserto												
Prà												
Ospedal. Centre												
Palugana-Tresto-Peagnola												
Vallancon -S.Croce -Dossi												

: index value between 0,10 & 0,20; index value between 0,20 & 0,30; ndex value greater than 0,30.

BIBLIOGRAPHY

Akaike H., 1973, "Information theory and an extension of the maximum likelihood principle", in *Second International Symposium on Information Theory*, pp. 267-281, Petrov B.N., Czaki F. Editors, Budapest, HUN: Akad, Kiadò.

Anderson M.J., ter Braak C.J.F., 2003, "Permutation tests for multi-factorial analysis of variance", *Journal of Statistical Computation and Simulation,* vol. 73, n. 2, pp. 85-113.

Burnham K.P., Anderson D.R., 2002, *Model selection and multimodel inference: a practical information-theoretic approach*, *2nd Edition*, New York, NY: Springer-Verlag.

Anzion C.J.M., Dragt A., Van Kuijk A.H.J., Post J.G., *Documentmeten en rekenengeur,* DenHaag, Ministerie VROM, 1994

Ballesta P.P., *The uncertainty of averaging a time series of measurements and its use in environmental legislation*, Atmospheric Environment 39, 2003–2009.

Best P.R., Lunney K.E., Killip C.A., *Statistical elements of predicting the impact of a variety of odour sources,* Water Science and Technology 44 (9), 157–164, 2001

Bertacchi M., Capuano F., Fornaciari S., Franzoni C., Poluzzi V., Renna E., Vvivi B., Bertolini E., Meglioli E., *Control of the air quality and research of smelling substances in landfills*, Sardinia 1997, Cagliari, Italy, 1997.

Bradley A.D., Cook D.J., Edwards J.S., Johnston A.G., Linforth R.S.T., Taylor A.J., *The control and measurement of landfill odours*, Sardinia 2001, Cagliari, Italy, 2001

Bowly S.W., *An assessment of current methods for quantify in landfill odours*, Sardinia 2003, Cagliari, Italy, 2003.

Both R., Sucker K., Winneke G., Koch E., *Odour intensity and hedonic tone—important parameters to describe odour annoyance to residents?* Water Science and Technology 50 (4), 83–92, 2004.

Carruthers D.J., Holroyd R.J., Hunt J.C.R., Weng W.S., Robins A.G., Apsley D.D., Thompson D.J., Smith F.B., *UK-ADMS: a new approach to modelling dispersion in the earth's atmospheric boundary layer,* Wind Engineering and Industrial Aerodynamics 52, 139–153, 1994.

Cavalini P.M., Koeter-Kemmerin, L.G., Pulles, M.P.J., *Coping with odour annoyance and odour concentrations: three field studies*, Journal of Environmental Psychology 11, 123–142, 1991.

CEN, *Odour concentration measurement by dynamic dilution olfactometry,* CEN/TC264wg2 Odours. CEN Kommission Reinhaltung der Luft, im VDI und DIN, Postfach 10 11 39, 40002. Dusseldorf, Germany, 1995

CEN, *Air quality—determination of odour concentration by dynamic olfactometry,* CEN EN 13725, 2003.

CERC, *ADMS 3 Technical Specification*, Cambridge Environmental Research Consultants, 2003.

Christensen T.H., Cossu R., Stegmann R., *Landfilling of Waste: Biogas*, Spon, London, 1996.

Clarkson C., *Smells and planning: the use of olfactometry in planning*, Institute of Waste Management, UK, 2000.

Dalton P., *Review: upper airway irritation, odor perception and health risk due to airborne chemicals*, Toxicology Letters 140–141, 239–248, 2003.

De Bruyn, G., Baron, M., Van Langenhove, H., Hendriks, J., Andries, A., Saevels, P., Leribaux, C., Vranken, E., Vinckier, C., Berckmans, D., *Development of a simple procedure for the determination of odor and of ammoniaemissions from farm buildingsaimingat the adaptation of environmental policy in Flanders. Part 2. Measurement procedure for ammonia and odoremissions from farm buildings*, Gent, KUL & RUG, 166, 2001.

Drew G.H., Smith R., Gerard V., Burge C., Lowe M., Kinnersley R., Sneath R., Longhurst P.J., *Appropriateness of selecting different averaging times for modeling chronic and acute exposure to environmental odours*, Atmospheric Environment 41, 2870–2879, 2007.

Elliott P., Morris, S., Briggs D., de Hoogh C., Hurt C., Kold Jensen T., Maitland I., Lewin A., Richardson S., Jon G.H., Wakefeld J., Jarup L., *Birth outcomes and selected cancers in populations living near landfill sites*, Report to the Department of Health, UK, 2001.

EN13725:2003, *Air Quality Determination of Odor Concentration by Dynamic Olfactometry*, CEN, Brussels, April 2003.

Environment Agency, *Integrated Pollution Prevention and Control (IPPC) Draft Horizontal Guidance for Odour Part 1—Regulation and Permitting*, Environment Agency, Bristol, 2002.

Frechen F.B., *A new model for estimation of odour emissions from landfill and composting facilities*, Sardinia 95, Cagliari, Italy, 1995.

Gendebien A., Pauwels M., Constant M., Ledrut-Damanet M.J., Nyns E.J., Willumsen H.C., Butson J., Fabry R., Ferrero G.L., *Landfill gas. From environment to energy*, Luxembourg, Commission of European Communities, 1992.

Gibson J.J., *The Senses Considered as Perceptual Systems*, Houghton Mifflin Co., Boston, 1966

Gifford F.A., *Statistical properties of a fluctuating plume dispersion model*, Adv.Geophys. 6, 117–138, 1959.

GOAA, *Guideline on odour in ambient air - Determination and Assessment of Odour in Ambient Air (1994)*, Länderausschuss für Immissionsschutz, LAI-Schriftenreihe No. 5, Berlin, 1999.

Gostelow P., Longhurst P.J., Parsons S.A., Stuetz R.M., Sampling for Measurement of Odours, IWA Publishing, London, UK, 2003

Gostelow P., Parsons S.A., Lovell M., *Integrated odour modelling for sewage treatment works*, Water Science and Technology 50 (4), 169-176, 2004.

Hamideh, S.A., A review of the literature regarding non-methane and volatile organic compounds in municipal solid waste landfill gas, MSW Manage. 12, 2002.

Hanna S.R., Egan B.A., Purdum J., Wagler J., *Evaluation of the ADMS, AERMOD and ISC3 dispersion models with the Optex, Duke Forest, Kincaid, Indianapolis, and Lovett field data sets*, International Journal of Environment and Pollution 16 (1–6), 301–314, 2000.

Irish Environmental Protection Agency, *Odour Impacts and Odour Emission Control Measures for Intensive Agriculture*, R&D Report Series No. 14.OdourNet UK Ltd, Bath, 2001.

Ittelson W.H., Environment Perception and Contemporary Perceptual Theory, *Percezione dell'ambiente e teoria contemporanea della percezione*, in S. Bagnara, Misiti R. (a cura di), Psicologia ambientale, Il Mulino, Bologna, 1978.

Jiang K., Kaye R., *Comparison study on portable wind tunnel system and isolation chamber for determination of VOCs from areal source*, Water Sci. Technol. 34, 583–589, 1996.

Karnik M., Parry C., *Landfill odour control. A practitioner's experience,* In: Christensen T.H., Cossu R., Stegmann R. (Eds.), Eighth International Waste Management and Landfill Symposiums Proceedings, vol. II, Leachate and Landfill Gas, Sardinia 2001, pp. 691–700, 2001.

Lindvall T., Radford, T.P., *Measurements of annoyance due to exposure to environmental factors*, Environmental Research 6, 1–36, 1973.

McGahan E., Nicholas P., Watts P., *Nuisance criteria for impact Assessment. Application of emission rate data in modelling and assessment*, Paper 5877/22, 13 November 2002.

Mahin T., Pope R., McGinley C., *When is a smell a nuisance? An overview of different approaches taken around the world in setting odor-control regulations*, Water Environment and Technology 12 (5), 49–53, 2002.

Manning Robert E., 2007, *Parks and Carrying Capacity*, Island Press.

McIntyre A., *Application of dispersion modelling to odour assessment: a practical tool or a complex trap?*, Water Science and Technology 41 (6), 81–88, 2000.

Miedema H.M.E., Walpot J.I., Vos H., Steunenberg C.F., *Exposure-annoyance relationships for odour from industrial sources*, Atmospheric Environment 34, 2927–2936, 2000.

Moortgat M., Schamp N., Van Langenhove H., *Assessment of odour nuisance problem in Flanders: a practical approach,* In: Dragt, A.J., van Ham, J. (Eds.), Biotechniques for Air PollutionAbatement and Odour Control Policies, Elsevier Science Publishers BV, 1992.

Mussio P., Gnyp A.W., Henshaw P.F., *A fluctuating plume dispersion model for the prediction of odour-impact frequencies from continuous stationary sources*, Atmospheric Environment 35, 2955–2962, 2001

Nicolas J., Craffe F., Romain A.C., *Estimation of odor emission rate from landfill areas using the sniffing team method*, Waste Management 26, 1259–1269, 2006.

Nicolas J., Romain A.-C., Monticelli D., Maternova J., Andre P., *Choice of a suitable E-nose output variable for the continuous monitoring of an odour in the environment,* ISOEN2000, Brighton, 2000.

Odotech, *Atmospheric emissions characterization and odor impact assessment of Argenteuil Deux Montagnes landfill area*, Montre´al, Canada, OdotechInc, 2001.

Pagé, T., Guy, C., *Odor dispersion modeling*, Air and Waste Management Association's 90th Annual Meeting and Exhibition, June 8–13, Toronto, Canada, 1997.

Pardo M., Sberveglieri G., Bergonzoni M., Ghizzoni F., Cortellini L., *The Pico-1 electronic nose for monitoring malodors in the environment*, Sardinia 2001, Cagliari, Italy, 2001.

Pasquill F., *Atmospheric Diffusion*, second ed. Chichester, England, 1974.

Reinhart D.R., Copper D.C., Walker B.L., *Flux chamber design and operation for the measurement of municipal solid waste landfill gas emissionrates*, J. Air Waste Manage. Association 42, 1067–1070, 1992.

Sarkar U., Hobbs S.E., Longhurst P., *Dispersion of odour: a case study with a municipal solid waste landfill site in North London*, United Kingdom. Journal of Environmental Management 68 (2), 153–160, 2003.

Sarkar U., Longhurst P., Hobbs S.E., *Community modelling: a tool for correlating estimates of exposure with perception of odour from municipal solid waste (MSW) landfills*, Journal of Environmental Management 68 (2), 133–140, 2003.

Sarkar U., Hobbs S.E., Longhurst P., *Odour from municipal solid waste (MSW): a study on the analysis of perception,* United Kingdom. Environment International 27, 656–662, 2002.

Schulz T.J., van Harreveld A.P., *International moves towards standardisation of odour measurement using olfactometry*, Water Science and Technology 34 (3–4), 541–547, 2006.

Sheridan B.A., Hayes E.T., Curran T.P., Dodd V.A., *A dispersion modelling approach to determining the odour impact of intensive pig production units in Ireland*, Bioresource Technology 91, 145–152, 2004.

Simms K.L., Wilkinson S., Bethan S., *Odour nuisance and dispersion modelling: An objective approach to a very subjective problem*, Water Science and Technology 41 (6), 89–96, 2000.

Sironi S., Capelli L., Centola, P., Del Rosso R., Il Grande M., *Odour emission factors for assessment and prediction of Italian MSW landfills odour impact*, Atmospheric Environment 39, 5387–5394, 2005.

Sironi S., Rossi A.N., Del Rosso R., Céntola P., Il Grande M., *Odour impact assessment using dispersion modeling: a case study of an operating landfill*, Sardinia 2003, Cagliari, Italy, 2003.

Stretch D., Laister G., Strachan L., Saner M., *Odour trails from landfill sites*, Sardinia 2001, Cagliari, Italy, 2001.

Termonia A., Termonia M., *Characterization and on-site monitoring of odorous organic compounds in the environment of a landfill site*, Int. J.AnalyticalChem. 73, 43–57, 1999.

Zammuner V.L., 1998, *Tecniche dell'intervista e del questionario*, Bologna: il Mulino.