

Chapter 3

Methodology

Since this study focuses on the identification of main sources of PM from a multidisciplinary point of view, the methodology used includes meteorological tools (e.g. back-trajectories, synoptic charts or model simulations), satellite observations, interpretation of time series of PM levels, chemical characterisation of PM₁₀, PM_{2.5} and grain size segregated fractions of PM and the monitoring of the grain size resolved PM levels in rural, urban and industrial environments. Thus, from the methodological point of view, this study is innovative given the number of tools simultaneously treated. As far as we know, this is the first study that simultaneously includes the above cited tools.

3.1 Interpretation of PM₁₀ and TSP time series

Time series of PM₁₀, TSP, NO_x, CO, SO₂ and O₃ levels recorded from 1996 to 2000 at nineteen stations forming part of several air quality monitoring networks (AQMN) were studied to assess the influence of meteorology on PM₁₀ and TSP levels in Eastern Spain. The aim of this assessment was to determine the role of the local and regional sources as well as the role of long range transport in PM levels. These air quality networks belong to the Autonomous Governments of Catalonia and Valencia, and to ENDESA (Empresa Nacional de Electricidad S.A.). The temporal variations of PM concentrations and the origin of the suspended particles were studied using the following tools:

1. Isentropic back-trajectories calculated with the HY-SPLIT model (Draxler, 1994), using meteorological data (190x190 kmxkm, 14 levels from the surface to 20hPa pressure level, and 6 hours, for horizontal, vertical and time resolution of the input data, respectively) supplied by the National Oceanic and Atmospheric Administration (NOAA). Five to ten days back-trajectories were calculated at 750, 1500 and 2500 m.a.s.l.

2. Synoptic meteorological charts for Europe provided by the U.K. Meteorological Office and the German Meteorological Service (DWD).

3. Maps of dust load and meteorological fields to identify the transport patterns of African dust to Eastern Spain. These maps of dust load were determined by simulations in the SKIRON forecast system (Kallos et al., 1997; Nickovic et al., 2001) performed during a stay at the University of Athens (Greece) from October to December 1999.

4. Maps of absorbing aerosol index to detect the occurrence of African dust outbreaks. These maps were supplied by the National Aero-Space Agency (NASA), and determined by means of TOMS-spectrometer satellite observations (Herman et al., 1997).

5. Correlation of PM₁₀ and TSP with primary (CO, SO₂ and NO_x) and secondary (O₃) pollutant gases, and analysis of the local meteorological variables measured simultaneously with the particulate and gaseous pollutants at the monitoring stations.

3.1.1 Measurements of PM₁₀ and TSP and gaseous pollutants in the air quality networks

At the AQMN, TSP measurements were performed by means of beta attenuation monitors (FAG and Dasiby at most stations forming part of the AQMN of Catalonia and Valencia, respectively), whereas PM₁₀ levels were monitored by means of beta (FAG and Dasiby at most stations belonging to the AQMN of Catalonia and Valencia, respectively), TEOM (Rupprecht and Patashnick at a station forming part of the AQMN of ENDESA) and a laser spectrometer (GRIMM at a station belonging to the AQMN of Valencia). Levels of gaseous pollutants were monitored by means of standard methods: absorption of IR radiation for CO, UV fluorescence for SO₂, chemiluminescence for NO_x and absorption of UV radiation for O₃.

Although the aim of the current study is to characterise PM behaviour and origin in Eastern Spain, data from other regions of the Iberian Peninsula (e.g. Andalusia, Castilla La Mancha or La Rioja, in Southern, Central and Northern Spain, respectively) were also used to support the interpretations.

It should be pointed out that the use of relatively "old" databases presents some limitations. Although the data were obtained without demonstrating the equivalence between the instrumentation used and the current standardised PM₁₀ and TSP measurement procedures, the quality of the data would be suitable for this study. Recent studies have demonstrated that automatic PM₁₀ instruments (TEOM and beta) underestimate PM₁₀ levels (down to 35%) when compared with the EN 12341 reference method (Allen et al., 1997; APEG, 1999; European Commission, 2001). This underestimation is due to the volatilisation of semi-volatile species owing to the heating of the sampled air to remove particle-bound water. This underestimation may also affect TSP (but to a lesser degree than PM₁₀) and PM_{2.5} (Ayers et al., 1999) measurements. Given that the demonstration of equivalence and the subsequent correction of the real time data were not mandatory until July 2001, the data presented in this study could include a slight underestimation regarding PM₁₀ and TSP measures obtained according to the requirements of the 1999/30/EC Directive. In this study, PM₁₀ sampling with the EN 12341 reference method was performed for one year (June 1999 - June 2000), and the potential underestimation was quantified.

3.1.2 Study of the African dust transport patterns by simulations in the SKIRON forecast system

Simulations using the SKIRON (Kallos et al., 1997; Nickovic et al., 2001) forecast system were performed to assess the occurrence of African dust outbreaks. Furthermore, synoptic charts were also used to support the back-trajectory and SKIRON analysis, and to identify the synoptic pressure systems giving rise to the transport pattern of mineral dust from North Africa.

The SKIRON forecasting system is a version of the ETA/NCEP weather forecasting model developed at the University of Athens (Kallos et al., 1997, Nickovic et al., 1988a). The 'core' of the system is the ETA/NCEP model initially developed at the university of Belgrade and the Yugoslav Hydrometeorological Service and further advanced at NOAA/NCEP.

The system is fully portable and contains a series of modules for pre- and post-processing of data. A major feature of the SKIRON system is the modules describing the dust cycles in the atmosphere: uptake-transport-deposition (Nickovic and Dobricic, 1996; Nickovic et al., 1998b).

The infrastructure for the SKIRON simulations (model with the dust cycle module, access to meteorological data, computers, etc..) was provided by the "Atmospheric Modelling and Weather Forecasting Group" at the University of Athens. The simulations were performed under the supervision of Professor George Kallos and Dr Olga Kakaliagou during a stay at the University of Athens from October to December 1999.

Simulations with the SKIRON system were carried out for selected high PM₁₀ and TSP events recorded at rural stations and previously identified as potential African dust intrusions by back-trajectory and the synoptic charts analysis. Output files were taken every 6 hours for winds at standard pressure levels, dust load (the vertically integrated dust concentration in g/m²) and precipitation data. Studied episodes showing good agreement between the transport patterns obtained from SKIRON simulations and from direct satellite observations of African intrusions over the Mediterranean are reported by Nickovic et al. (1997, 1998b) and Özsoy et al. (2001).

For the SKIRON simulations, the ECMWF analysis data fields (0.5°x0.5° latitude x longitude) at standard pressure levels were used for initiation and for nudging the boundary domain every 6 hours. Moreover, climatological sea-surface temperature fields (1°x1° SST), topography data base (30"x30"), soil textural class data set (1°x1°) and gridded vegetation data (30"x30") were used. The modelling domain covered an area ranging from the Western Atlantic to the Central Mediterranean, and from Central Europe to Southern Algeria. The defined mesh had a grid increment of approximately 24 km horizontal (0.25°x0.25°) and 32 vertical levels up to about 16 km. The simulations had a duration of at least 5 days.

Data were treated as follow: *Step-1*, High PM₁₀ and TSP events simultaneously recorded at distant rural stations were selected (pre-selection). *Step-2*, Air back-trajectories were computed for these events and episodes showing a North African long range transport pattern were finally selected for subsequent analysis. *Step-3*, the impact of African dust on the PM₁₀ and TSP levels at urban and industrial stations for the selected events was studied to

assess the relative importance of the African dust input versus the anthropogenic emissions. *Step-4*, SKIRON simulations were run for most of the selected events to study the meteorological patterns giving rise to the African dust outbreaks and to correlate the areas with high PM₁₀ and TSP levels with those affected by the North African dust.

3.1.3 Assessment of the influence of the meteorology on PM levels

Initially, the daily (24-hours) mean PM concentrations time series from urban, industrial and rural sites were compared between January 1996 and December 2000. This enabled us to identify the periods with simultaneous high or low PM concentrations at urban/industrial and rural sites, and those with high urban/industrial and low rural PM concentrations. In the latter case, the high PM concentrations measured at urban/industrial sites were attributed to local pollutant sources. However, simultaneous high PM concentrations at urban/industrial and rural sites may be associated with regional episodes (transport from the urban/industrial to the rural sites) or with "external" contributions by long range transport of particulate matter from North Africa (natural mineral dust) or Europe (particulate pollutants). The identification of the "regional" and "external" sources (North Africa or Europe) was performed by means of meteorological analysis based on the study of: a) synoptic charts, b) meteorological variables measured at the air quality monitoring stations, c) air Hysplit back-trajectories, d) TOMS maps of absorbing aerosol index determined by satellite observations, e) dust load maps determined by simulations in SKIRON forecast system and f) correlation of PM with gaseous pollutants (NO_x, SO₂, O₃ and CO). Although, these methods are useful tools for the interpretation of the PM source origin, they suffer from some limitations. Thus, TOMS measurements do not detect African dust events if the dust transport occurs at altitudes < 1 km (Herman et al., 1997). In these cases, the meteorological analysis and SKIRON maps of dust load are complementary tools for detecting these events. The Hysplit back-trajectories cannot represent the actual air mass movements activated by local and regional processes given the limitations in the resolution of the input data. Bearing in mind the results from earlier studies on summer meteorology in this region (Millán et al., 1997, 2000; Salvador et al. 1999; Soriano et al., 2000), the analysis of synoptic charts and the local meteorological variables recorded at the air quality monitoring stations are complementary tools for detecting these regional events.

3.2 Chemical characterisation of PM

PM sampling for chemical characterisation was performed by means of a set of PM₁₀ and PM_{2.5} high vol. samples and cascade impactors. The sampling period was not the same at all the monitoring sites. At some sites, systematic sampling of PM₁₀, PM_{2.5} and cascade impactor was performed for one year, whereas at other sites this was limited to intensive campaign sampling.

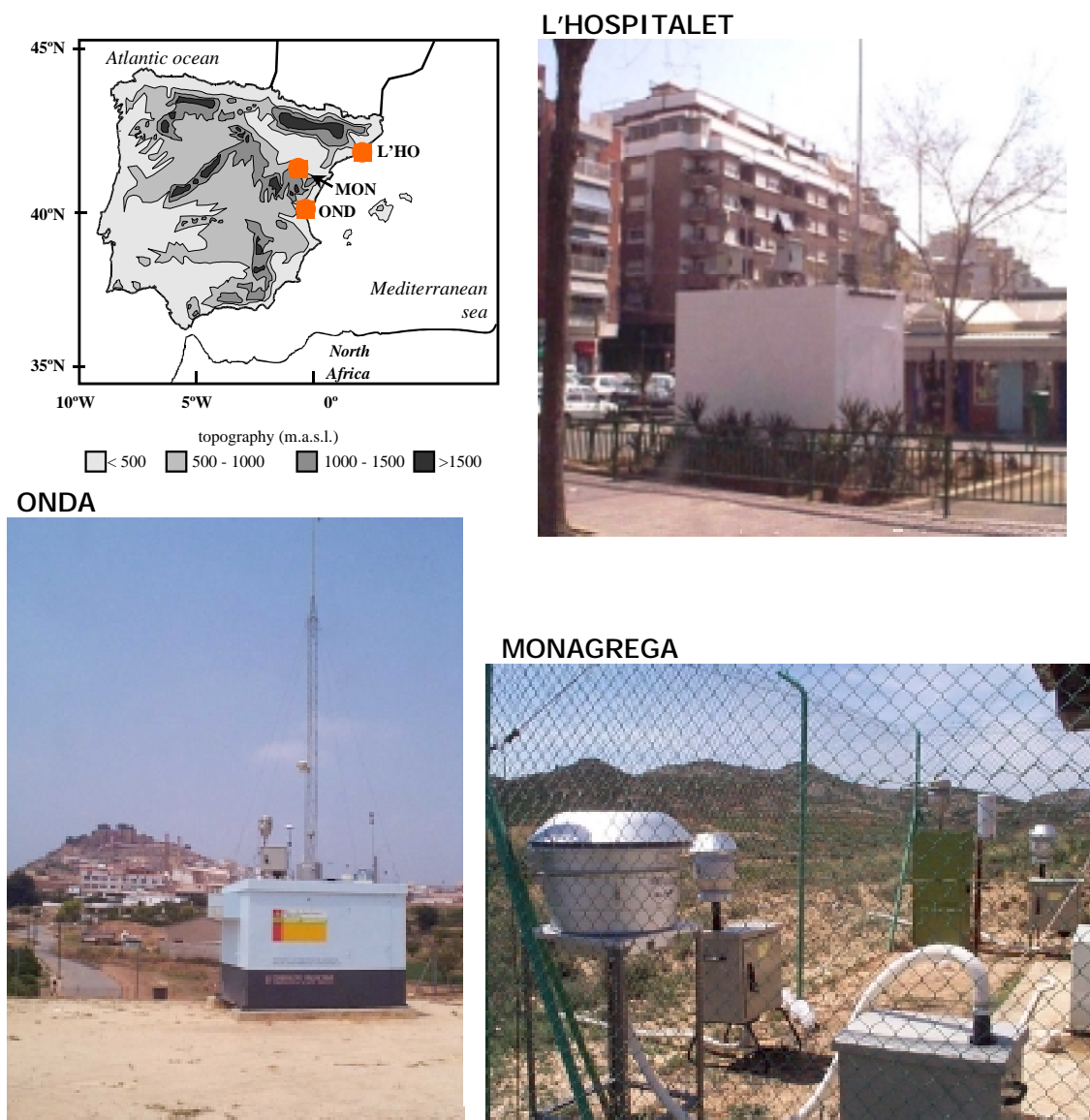


Figure 3-1. L'HOSPITALET (L'HO), MONAGREGA (MON) and ONDA (OND) air quality monitoring stations in Eastern Spain.

3.2.1 PM₁₀ and PM_{2.5} sampling

PM sampling was carried out at three sites (Figure 3-1) that were representative of different types of environments: an urban site in the Metropolitan Area of Barcelona (L'HOSPITALET station), a rural site in the Southern part of the Ebro basin (MONAGREGA station) and an industrial site (ONDA station) affected by primary anthropogenic emissions.

The L'HOSPITALET traffic/urban station, located in the Barcelona Metropolitan Area (Gornal avenue, 41.2°N, 2.0°E, 70m.a.s.l.), was selected for PM₁₀ and PM_{2.5} sampling. The monitoring station belongs to the Department of the Environment of the Autonomous Government of Catalonia. This site was selected because it is representative of stations of the AQMN of Eastern Spain, which are located at kerbsides. The L'HOSPITALET station is located <1km westward of the Mediterranean shoreline. The site is representative of a typical urban

environment of the coastal area, highly influenced by traffic emissions. 24-hour PM sampling was performed using MCV high volume captors (30 m³/h) equipped with DIGITEL PM10 and PM2.5 inlets. The sampling was carried out at a rate of two and one daily sample per week for PM10 and PM2.5, respectively, during the study period (June 9 1999 to June 29 2000). A total of 115 PM10 and 63 PM2.5 daily samples were collected using the above procedure.

The MONAGREGA rural station was selected for PM10 sampling from March 24 1999 to July 9 2000. From June 1999 to June 2000 sampling was carried out simultaneously with that at the L'HOSPITALET urban station. MONAGREGA is located in a semi-arid area (Calanda desert) in the Southwestern Ebro basin (40.5°N, 0.2°W, 600m.a.s.l.), 85km westward of the Mediterranean coast, and close to the Iberian range. The MONAGREGA station is located in the vicinity of the Teruel power plant (7.6Km in a SE direction), which is a large source of SO₂ emissions. Relatively low primary PM emissions are expected from this power plant due to the use of high efficiency electrostatic precipitators. The MONAGREGA monitoring station was selected because it is an adequate site for comparing the PM10 load of secondary pollutants (e.g. sulphate) versus natural crustal components. 24-hour PM10 sampling was performed using DIGITEL high volume captors (30 m³/h) equipped with a DIGITEL PM10 inlet. A total of 132 PM10 samples were collected in the study period. Additionally, PM10 and PM2.5 samples were collected during an intensive campaign in summer 2000.

The ONDA industrial station was selected for PM10 sampling from June 15 to November 2 1999. This sampling was performed simultaneously with that at the rural and urban stations. ONDA (39.9°N, 0.2°W, 167m.a.s.l.) is located in the Millars valley and is an industrial background station belonging to the Department of Environment of the Autonomous Government of Valencia. This area is affected by primary anthropogenic emissions from the ceramic industry of Castelló. The site is located 20 Km westward of the Mediterranean shoreline. In addition to the ceramic industry, the Millars valley is an important area of orange production. Activities derived from the management of orange plantations such as the burning of orange tree wastes may also have a significant impact on the airborne particulate composition in the valley. 24-hour PM10 sampling was performed using MCV high volume captors (30 m³/h) equipped with DIGITEL PM10 inlet, and 48 samples were obtained. In addition, a set of simultaneously collected PM10 and PM2.5 samples were obtained during an intensive sampling campaign in summer 1999.

3.2.2 Cascade impactor sampling

TSP size segregated samples were collected at the different sampling sites by means of a Retsch PI-1 cascade impactor with 7 stages from 0.3 to >20 µm and without sampling substrate. The sampling period ranged from 4 to 10 days.

At the L'HOSPITALET urban site, a set of 7 cascade impactor samples homogeneously distributed from June 1999 to June 2000 were collected.

At the MONAGREGA rural site, cascade impactor sampling was performed during the intensive campaign in summer 2000.

At the ONDA industrial site, a cascade impactor sample was collected during the intensive campaign in summer 1999.

3.2.3 Chemical analyses

The 24-hour sampling of PM₁₀ and PM_{2.5} was performed using 15cm quartz fibre filters (QF20 Schleicher and Schuell). After obtaining the PM₁₀ and PM_{2.5} concentrations by gravimetric standard methods, one half of each filter plus the insoluble fractions of the different stages of the cascade impactor (previously filtered with MQ cellulose membrane filters) and blank filters were acid digested. The digestion method is a modified version of a method devised by Querol et al. (1996), which is summarised as follow:

- a) a fraction of the filter is introduced into a 60ml PFA digestion bomb, and 2.5 ml HNO₃ + 4ml HF are added.
- b) The closed bomb is heated overnight at 90°C.
- c) After cooling the bomb is opened and 2.5 ml of HClO₄ are added.
- d) The resulting acidic digestion is driven to dryness on a hot plate at 250°C.
- e) The residue is dissolved with 2.5 ml HNO₃ and is transferred to a graduated flask, after which milliQ grade water is added to attain a volume of 50ml.

The soluble fraction of ¼ of each filter and the different grain-size stages (rings) obtained with the cascade impactor were extracted with distilled water at 80°C for the analysis of soluble ions. The last untreated ¼ of each filter was employed to determine the total C content using an elemental C analyser (LECO methodology based on the thermal oxidation of a carbonaceous aerosols and the measure of CO₂ by IR). The solutions obtained from the bulk digestion and leaching of the different sections of each filter were analysed by: a) Inductively Coupled Plasma Atomic Emission Spectrometry (ICP-AES) for major elements (Ca, Al, Fe, Mg, Na, K, Ti, Sr, Mn), b) Inductively Coupled Plasma Mass Spectrometry (ICP-MS) for trace elements (Pb, V, Ni, Cr, Cu, Cd, As), c) Capillary Electrophoresis (CEF) for anions (SO₄²⁻, NO₃⁻ and Cl⁻), and d) Colorimetry-Flow Injection Analysis (FIA) for NH₄⁺. The digestion of blank filters with 2 to 3 mg of NBS-1633a reference material was used to ensure analysis quality for the same levels of the sample digestion concentrations. Relative analytical errors were between 3 and 10 % for the elements studied. The ICP-AES and ICP-MS analyses were performed at the Institute of Earth Science "Jaume Almera" (CSIC) in Barcelona, the CEF analyses were carried out at the Institute of Coal-Chemistry in Zaragoza (CSIC) and the LECO and FIA analyses at the Technical and Scientific Services of the University of Barcelona. Moreover, fifteen rural and 15 urban PM₁₀ samples homogeneously distributed throughout the study period were selected for the analysis of organic and elemental carbon concentrations by thermal-optical techniques (Pio

et al., 1994). These analyses were carried out during a stay at the University of Aveiro (Portugal) from November to December 2000. The facilities for these analyses were provided by the "Departamento de Ambiente e Ordenamento" of the University of Aveiro, and the analyses were performed under the supervision of Professor Casimiro Pio and Mr. Abel Carvalho. Indirect determinations from analytical data were obtained for: a) levels of non-sea-salt sulphate (nss-SO₄⁼) were obtained by subtracting the sea salt-SO₄⁼ (indirectly determined by stoichiometry from the soluble Na levels: SO₄⁼/Na⁺=0.25 (in mass) in sea water; Warneck, 1988) from total SO₄⁼; b) The contents of SiO₂ and CO₃⁼ were determined from the concentrations of Al, Ca and Mg on the basis of prior experimental equations (3*Al₂O₃ = SiO₂ and 1.5*Ca + 2.5*Mg = CO₃⁼, mass ratios). The Si and Al equations were obtained in earlier studies (e.g. Bergametti et al., 1989; Dulac et al., 1992; Molinaroli et al., 1993, Chiapello et al., 1997). The Ca, Mg and CO₃⁼ equation is based on stoichiometric relations, and has also been corroborated by previous studies in Barcelona. The fraction of non mineral carbon, or organic + elemental carbon (OC+EC), was determined by subtracting the load of mineral carbon (deduced by stoichiometry from CO₃⁼ levels) from total carbon.

Levels of carbon are partitioned between the organic carbon + elemental carbon (OC+EC) and mineral carbon (MC) fractions by assuming that all Ca molecules occur as calcium-carbonate. However, a fraction of nitrate may also occur as calcium-nitrate and its formation induces the volatilisation of a fraction of MC: CaCO₃ + 2HNO₃ → Ca(NO₃)₂ + CO₂(g) + H₂O. The reaction of other acidic species, such as sulphuric acid, with CaCO₃ also leads to the volatilisation of a fraction of MC. This overestimation in the MC levels would induce an underestimation in the OC+EC levels. However, this is relatively low given that levels of OC+EC recorded in this study are much higher than those of MC (discussed in chapter 5). If 50% of Ca occurs as Ca(NO₃)₂, the underestimation in the OC+EC levels at L'HOSPITALET is estimated at 0.4µg/m³, which is only 3.7% of the previously determined OC+EC levels (on the assumption that 100% of Ca is present as CaCO₃). In the rural environment where the MONAGREGA station is located, a minor fraction of Ca is expected to occur as Ca(NO₃)₂ because of the minor NO_x, acid nitric and nitrate pollution. If 30% and 50% of Ca occur as Ca(NO₃)₂ in PM₁₀ at MONAGREGA, the underestimation in the OC+EC levels is estimated at 0.15µg/m³ and 0.09µg/m³, which is only 6.5% and 1.2% of the previously determined OC+EC levels (on the assumption that 100% of Ca is present as CaCO₃), respectively. The underestimation is much lower in PM_{2.5} than in PM₁₀ because Ca mainly occurs in the fraction >2.5µm (discussed in chapters 5 and 6).

This PM₁₀ and PM_{2.5} sampling procedure meets the EU Directive 1999/30/EC requirements given that samplers equivalent to the EN12341 reference equipment were used. However, the limitations should be considered. The filters (quartz fibre) employed do not produce significant positive artefacts, but semi-volatile species such as ammonium-nitrate and

organic species may be evaporated from the filter, resulting in negative artefacts (Schaap et al., 2001).

3.3 PM₁₀ and PM_{2.5} source apportionment

Receptor modelling techniques were used for the source apportionment of PM₁₀ and PM_{2.5}. These techniques were applied to PM₁₀ and PM_{2.5} data collected at L'HOSPITALET, MONAGREGA and ONDA.

The source contribution to ambient PM levels have frequently been based on dispersion models in which emission inventories for various sources are used as input data to predict ambient PM concentrations. These dispersion models do not always reproduce experimental data. Alternative procedures based on Lagrangian trajectory models have been developed to simulate the main processes involved in aerosol emission, formation, transport and deposition (Eldering and Cass, 1996; Kleeman and Cass, 1998). Although these procedures are consistent with the experimental data, they require a detailed emission inventory that is not always available.

The receptor modelling techniques are based on the evaluation of data acquired at receptor sites, and most of them (those known as multivariable receptor models) do not require prior identification of emission sources (Henry et al., 1984). These models have played a key role in the evaluation of PM sources with respect to national air quality standards in certain countries. In the United States, the Chemical Mass Balance model (e.g. Gertler et al., 1995; Chow et al., 1996) has been widely used, whereas in Europe receptor modelling techniques have mainly been based on methodologies that do not require chemical profiles of previously identified source emissions (e.g. Harrison et al., 1997a, 1997b; Pio et al., 1996, 1998; Marcazzan et al., 2001).

Receptor models are based on the hypothesis that the bulk mass of PM can be considered to be the sum of independent contributing components (Hopke, 1985, 1991):

$$PM = \sum_{i=1}^n PM_i \quad (PM_i = \text{mass concentration from the } i^{\text{th}} \text{ source; } n = \text{number of sources}).$$

Two main stages are distinguished in the multivariable receptor models: 1) source identification, and 2) estimation of source contribution. In the source identification process, a number of groups of chemical species of PM are differentiated on the basis of the degree of association or correlation between the analysed PM components. As regards the chemical species, each group is associated with a PM source. The source identification is usually based on the application of statistical tools, such as Principal Component Analysis (e.g. Sexton et al., 1985; Thurson and Spengler, 1985; Pio et al., 1996; Prati et al., 2000), Factor Analysis (Alpert and Hopke, 1980; Henry et al., 1984; Gordon, 1988) or Cluster Analysis (Sánchez and Ramos, 1987; Saucy et al., 1985). The quantitative determinations of the source contributions are usually based on multilinear regression analysis, in which the bulk PM concentration is used as a dependent variable, and a tracer of each source is employed as an independent variable. Chemical species

of known origin or the absolute score factors obtained from the Principal Component Analysis are frequently used as source tracers (e.g. Thurson and Spengler, 1985; Hopke, 1991; Castro, 1997).

3.3.1 Theoretical foundations of the receptor modelling

A summary of the main foundations of the multivariable receptor model used in this study is presented. For details on the mathematical developments, readers are addressed to the excellent reviews of receptor models presented by Henry et al. (1984), Thurson and Spengler (1985), Gordon (1988), Hopke (1985, 1991) or Castro (1997).

Receptor models are based on the equation of mass conservation:

$$C_{ij} = \sum_{k=1}^m a_{ik} \cdot S_{kj} \quad (\text{Eq. I})$$

($i= 1.....n$; $j=1.....N$) n = number of chemical species constituting PM; N = number of samples (equivalent to days for 24-hours samples); m = total number of PM sources; C_{ij} = concentration of the i^{th} component in the sample j^{th} ; S_{kj} = mass of particles associated with the k^{th} source during the sampling of the j^{th} sample; a_{ik} = mass fraction of the i^{th} component in the k^{th} source.

The equation is expressed in terms of matrix:

$$C = A \cdot S \quad (\text{Eq. II})$$

where C , A and S are matrixes of dimension $n \times N$, $n \times m$ and $m \times N$, respectively. The aim is to determine the matrixes A and S . The multivariable receptor models assume that the variations in the PM component concentrations (C_i) are caused by variations in the contributions of each source. The model defines the most probable combination of sources which accounts for the variability in PM composition at the receptor sites. Thus, these models offer the advantage that they do not require prior knowledge of the number of PM sources or its chemical composition. However, limitations of these receptor models should be taken into account. Sources are identified on the basis of the correlation between the variations in the PM components levels. Thus, if any PM source presents a constant contribution during the study period, it is not detected. However, this is not likely if a relatively long sampling period is considered. Another limitation to be taken into account is the potential influence of factors which may induce correlated variations between PM components which do not result from the same source. In this case, the model would group these PM components in the same source, e.g. PM emissions from two sources located in the same direction with respect to the sampling site will be simultaneously transported to the receptor-sampling site under the same wind direction. Another drawback is that a relatively high number of PM samples are necessary, at least a minimum 30 samples in order to have statistical significance.

The Principal Component Analysis (PCA) is based on the variability of the PM component concentrations at the receptor site. This technique seeks to determine the matrixes A and S (Eq.

II) accounting for the variability of PM component concentrations at the receptor site. To this end, the variables are normalised:

$$Z_{ij} = \frac{C_{ij} - \overline{C}_i}{\sigma_{C_i}} \quad (\text{Eq. III})$$

$$\text{being: } \overline{C}_i = \frac{1}{N} \sum_{j=1}^N C_{ij} \quad (\text{Eq. IV})$$

$$\sigma_{C_i} = (\overline{C_i^2} - \overline{C}_i^2)^{1/2} \quad (\text{Eq. V})$$

\overline{C}_i and σ_{C_i} average and standard deviations of the component C_i in the N samples.

Thus, equation II is expressed in terms of normalised variables:

$$Z = L \cdot F \quad (\text{Eq. VI})$$

where Z , L and F are matrixes of dimension $n \times N$, $n \times m$ and $m \times N$, respectively.

The normalised concentrations z_{ij} are adimensional variables, with an average value equal to zero and a standard deviation equal to one. Once normalised, the influence resulting from the different scales and value ranges of PM component concentrations is removed with the result that PM components at high concentrations do not displace the potential influence of the PM components at low concentrations. If this normalisation were not performed, the PCA would be based on the magnitude of the PM component concentrations, and not on the variability of the PM component concentrations. Thus, the PM components present at low and high concentrations exert the same influence on the PCA.

Two factors should be taken into account when normalisation is applied. First, trace species of PM at concentrations very close to the detection limit of the analytical instruments should not be introduced into the PCA. For these PM trace species, the magnitude of the variations of their concentrations in ambient air may be very close to the magnitude of the analytical errors. In this case, normalisation should amplify the analytical errors, interfering in the PM source identification. Second, normalisation hampers the quantitative determination of the PM source contribution. This is due to the fact that equation VI is resolved instead of II. A method for obtaining qualitative determination is discussed below.

Normalisation introduces new variables.

$$f_{pj} = \frac{S_{pj} - \overline{S}_p}{\sigma_{S_p}} \quad (\text{Eq. VII})$$

$$z_{ij} = \sum_{k=1}^m \alpha_{ik} \cdot f_{ki} \quad (\text{Eq. VIII})$$

($i = 1, \dots, n$; $j = 1, \dots, N$).

f_{pj} are known as score factors, and are related to the normalised contribution of the p^{th} source in the j^{th} source.

α_{ij} are known as factor loading, and represent the correlation between the normalised concentrations z_{ij} and the normalised factor f_p (principal components) associated with each PM source emission (it is mathematically demonstrable that the $r_{C_iS_p} = r_{z_i f_p} = \alpha_{ip}$).

The matrix L (Eq. VI) is mathematically related to the matrix R constituted by the correlation between the normalised concentrations of the PM components ($R=L \cdot L^t$). In the PCA, the $n \times m$ factors loading α_{ip} , which correlate the n variables z_i with the m factors f_p , define a m -dimensional space. Geometrically, these components f_p represent orthogonal axes in which the variables are plotted.

The PCA determines the m principal components which verify two conditions:

1) The factors should be orthogonal:

$$\sum_{j=1}^N f_{pj} \cdot f_{qj} = \delta_{pq} \quad (\delta \text{ Kronecker delta}) \quad (\text{Eq. IX})$$

2) Maximisation of the explained variance. Each principal component should account for the maximum possible total-variance.

The m principal components (or factors f_p) of the PCA define a m -dimensional space. Any linear combination of these m principal components is also a solution to the PCA. Thus, there are an infinite number of solutions arising from rotations (linear combinations) of the m principal components, but not all solutions have a physical meaning given that most of them present high factor loading in a number of PM components. It would be desirable to find a solution in which each principal component presents high factor loading only for a few PM components, which may be easily attributed to any PM source. This solution would be found after applying orthogonal rotations. The varimax rotation is a method used in most studies. In this method, rotation is performed to find a solution with a maximum simplicity, selecting a combination of vectors so that the variances of each variable are distributed in the minimum number of principal components. Thus, after applying the varimax rotation, each principal component (or factor) presents high factor loading for only a few PM components. After this treatment it is easy to associate each principal component with a PM source.

Currently, this complex matrix treatment may be performed by means of any of the statistical package available in software. In the present study the StatSoft® statistical package was used.

3.3.2 Source apportionment methods used in this study

In the current study, the source identification was performed by means of Principal Component Analysis (PCA) followed by varimax rotation. As described above, the PCA followed by varimax rotation is a technique which attempts to explain the statistical variance in a given data set in terms of a minimum number of significant components. The PCA was applied only to independent variables. Thus, indirect chemical determinations are not introduced in the PCA.

The source contribution was calculated by two methods based on Multi-linear Regression Analysis and Mass Balance Analysis, respectively.

The first method used to calculate the source contribution is based on a Multi-linear Regions Analysis (MLRA) using the absolute score factors as source tracers (Thurston and Spengler, 1985):

$$PM_j = \sum_{i=1}^n (PM_i)_j = \sum_{i=1}^n (\alpha_i \cdot b_i)_j + b_{oj} = (\alpha_1 \cdot b_1)_j + (\alpha_2 \cdot b_2)_j + \dots + (\alpha_n \cdot b_n)_j + b_{oj}$$

(Eq. XI)

where:

i = source; j = sample (equivalent to day for 24-hours samples)

$(PM_i)_j$ = mass concentration of particles associated with the i^{th} source in the j^{th} sample

α_i = slope associated with the i^{th} source, determined in the in the MLRA

b_i = absolute score factor associated with the i^{th} source, determined in the PCA

b_{oj} = intersection constant in the MLRA

n = number of sources

As explained above, the results of the PCA provide the "normalised" score factors (f_{pj} in eq. VII). In order to determine the zero of the score factors, Thurston and Spengler (1985) developed a method in which a fictitious sample with PM component concentrations equal to zero is introduced into the PCA. The score factors associated with each principal component in this fictitious sample are considered to be the real zero. Thus, the absolute score factors are calculated by subtracting the zero from the score factor of each sample.

In the MLRA the interception constant represents the contribution from other "non identified sources". A negative interception constant has no physical meaning. In this case the multi-linear regression should be made to pass through the origin (Castro, 1997; Pio et al., 1998). A negative contribution from any identified PM source (i.e. negative slope in the multi-linear regression) has a physical meaning. This negative contribution is caused by an anti-correlation between the components from the PM source and the bulk PM concentrations. The anti-correlation may be caused by chemical reactions or by meteorological patterns (Castro, 1997). For example, the reaction of sulphuric acid with sodium chloride ($H_2SO_4 + NaCl \rightarrow Na_2SO_4 + HCl(g)$) induces the volatilisation of Cl. Thus, a negative contribution from the source of sulphuric acid to the Cl levels is expected. At an urban coastal site, strong winds from the sea will favour the dispersion and dilution of particulate pollutants but may increase the sea spray components. In this case, an anti-correlation between the bulk PM levels and sea spray components may be observed.

This MLRA technique was applied to this study given that it is the most widely used method for source apportionment studies in Europe (Harrison et al., 1996, 1997a, 1997b; Pio et al., 1991, 1996, 1998; Castro, 1997).

The second method employed for calculating source contribution is based on Mass Balance Analysis (MBA) of the components associated with each principal component. Thus, the contribution from each source is estimated as the sum of the PM components associated with each factor. For the PM components associated with several factors, the mean percentage of each PM component associated with each factor (or source) was determined by MLRA (using the absolute score factors as independent variables). In this method the chemical components indirectly determined were considered in the source contribution (SiO_2 and Al_2O_3 replace to Al; $\text{CO}_3^{=}$ associated with Ca; OC+EC replace to C; ss- $\text{SO}_4^{=}$ associated with Na; nss- $\text{SO}_4^{=}$ replace with $\text{SO}_4^{=}$). This MBA method was employed owing to the fact that the chemical determinations performed in this study accounted for almost all the PM10 and PM2.5 concentrations (75-85% of bulk PM levels at each sampling site, discussed in chapter 5). To our knowledge, most studies which have used the aforementioned source apportionment MLRA technique have reached lower chemical determinations of bulk PM levels, usually in the range 30-60% (Harrison et al., 1997a; Pio et al., 1991, 1996, 1998; Castro, 1997). Finally, a comparison between the source apportionment results obtained by the two methods was performed.

These receptor modelling techniques are very versatile tools. They can be applied to chemical PM data on a seasonal basis or during given periods of the year. They can also be applied to PM data taken at different sites on a regional scale during relatively short time (~days) periods (Hopke, 1991). Moreover, source apportionment can be applied to specific PM components (Castro, 1997), e.g. "industrial fraction of PM" or a specific PM chemical species (e.g. nitrate).

3.4 Grain size resolved measurements of PM levels

Automatic grain size resolved measurements of PM levels were performed with the laser spectrometer dust monitor GRIMM 1108 (Lambert Technik GmbH & Co. KG). This dust monitor allows the determination of particulate levels in 15 different grain-size channels from 0.3 to >20 μm . The period of measurements was different at each site:

At the L'HOSPITALET urban station, PM grain size distribution measurements were performed continuously from June 1999 to June 2000.

At the MONAGREGA rural station, and at other sites in the Ebro basin, PM grain size distribution measurements were carried out during a field campaign in summer 2000.

At the ONDA industrial station, and at other sites in the Millars valley, PM grain size distribution measurements were performed during a field campaign in summer 1999.

Inter-comparison between PM levels recorded with the GRIMM laser spectrometer and high vol. samplers was carried out to assess the accuracy of this new PM monitoring equipment.