A METHODOLOGY FOR THE CREATION OF METEOROLOGICAL DATASETS FOR LOCAL AIR QUALITY MODELLING AT AIRPORTS

Nicolas Duchene[1], James Smith[1] and Ian Fuller[2]

[1]ENVISA, Paris, France
[2]EUROCONTROL Experimental Centre (EEC), Brétigny-sur-Orge, France

Abstract: In order to properly estimate local air pollution concentrations at airports, several different dispersion models are routinely applied using a variety of different modeling approaches (Gaussian, Lagrangian or Eulerian). Common to all dispersion models is the requirement for accurate meteorological parameters. The paper outlines the benefits and risks of three separate approaches to obtain meteorological input for atmospheric dispersion models.

The preferred approach is based on directly-measured observations and the primary source of readily available observed data at airports is METAR. A typical METAR report contains observations of temperature, dew point, wind, precipitation, cloud cover, cloud heights, visibility, and barometric pressure. However, most dispersion models require information on atmospheric stability. Although stability is not directly reported in METAR data, a widely-available algorithm allows for the estimation of atmospheric stability class using measured values of wind speed and the observed cloud cover.

The next preferred option should be used when METAR or other observed data are not readily available or more sophisticated 3D gridded meteorological fields are required by the specific dispersion model. This second approach uses meso-scale numerical weather prediction (NWP) models. These models can produce high quality 'best guess' meteorological fields on a wide variety of time and distance scales. NWP models, however, require high-level meteorological expertise in order to run and are computationally intense. This may make the NWP approach impractical for use in routine applications or for large-scale studies which involve many different airports.

Finally, this paper outlines a third approach to obtaining meteorological data. This approach uses long-term, globally-archived, gridded meteorological analysis fields, such as REANALYSIS data, which are readily available and cover long-term time scales. Although less accurate than METAR and NWP models, this approach may be of benefit to those users who require 'good guess' meteorological data for air pollution studies in those cases where direct observations, such as METAR, are not available and NWP modelling is not a viable solution.

Key words: *METAR, REANALYSIS, MM5, WRF, numerical weather prediction, local air quality, dispersion, stability, data completion.*

1. INTRODUCTION

In order to properly estimate local air quality (LAQ) concentrations at airports, several different types of dispersion models are routinely applied using a variety of different modelling approaches (whether Gaussian, Lagrangian or Eulerian). Common to all dispersion models, is the requirement of accurate meteorological parameters. Critical parameters for local air quality studies are wind speed, wind direction and stability. Atmospheric stability is a measure of the amount of turbulence in the ambient atmosphere. This has a major effect on the dispersion of air pollution plumes because turbulence increases the mixing of unpolluted air into the plume and thereby acts to reduce the concentration of pollutants in the plume. In general, the more 'unstable' the atmosphere, the more turbulent it is and therefore the greater the amount of air pollution dispersion. It is therefore important to categorize the amount of atmospheric turbulence present at any given time.

Although meteorological data for LAQ studies can be provided by national weather services, such as the British Met Office or Meteo France, and also by specialized meteorological companies, the purpose of this paper is to describe freely available meteorological data sources and to outline a procedure to choose between them. The intended audience would be those LAQ modellers who may wish to use freely available sources for budgetary reasons or as part of an initial scoping study. This paper will outline the benefits and risks of three separate meteorological data sources that can be used in order to obtain meteorological input for LAQ atmospheric dispersion models: observed METAR data, numerical MM5/WRF forecast data, and archived Reanalysis data.

2. METEROLOGICAL SOURCES

METAR Reports

First of all, the preferred source of meteorological data for dispersion models is directly-measured observations. The primary source of readily and freely available observed data at airports is METAR[i]. METAR is a routine meteorological report documenting the observed weather conditions at an airport at the time of the report in an international standard format. International METAR weather reports are usually issued once per hour, but when there are significant changes in weather, the METAR reports are issued more frequently. In addition, many of the airports in Europe issue METARs every 30 minutes during the airport opening hours.

METAR reports are predominantly used by pilots in fulfilment of a pre-flight weather briefing but they are also used by meteorologists, who use aggregated METAR information to help forecast the weather. METAR reports usually come from airports (civilian and military) and are generally produced by automated sites whereby weather information is recorded by digital sensors and encoded via software. These automated observations are usually reviewed by certified weather observers or forecasters prior to being transmitted. In some rare cases weather observations may be taken by trained observers or forecasters who manually observe and encode their observations prior to their being transmitted.

A typical METAR report contains data for wind speed and direction, temperature, dew point, cloud cover, cloud height, visibility, and barometric pressure. A METAR report may also contain information on precipitation amounts, lightning, and other information that would be of interest to pilots or meteorologists. In addition, a short period forecast called a 'Trend' may be added at the end of the METAR covering likely changes in weather conditions in the two hours following the observation.

A European METAR signal may comprise up to 10 components - preceded by a 3 component identifier.

Identifier Components:
1. Report Type - METAR
2. Location Indicator - ICAO four letter aerodrome code.
3. Time - The time of the observation in UTC in hours and minutes followed by the letter Z (the universal indication that UTC time is being reported).

METAR Weather Components:
1. Wind speed and direction
2. Visibility
3. RVR (Runway visual range)
4. Weather
5. Cloud
6. Temperature/Dew Point
7. QNH (barometric pressure extrapolated to sea level)
8. Recent Weather
9. Wind Shear
10. Trend

Unlike wind speed or temperature, which is directly measured and reported in METAR, stability class is not routinely measured at observing stations, and so is not part of the reported data. Therefore in order to use METAR data for LAQ studies, stability parameters must be estimated using algorithmic procedures based on the available METAR variables. An example of such an algorithmic procedure is the standard EPA 'Turner' method for determining the 'Pasquill-Gifford' stability class[ii]. This method estimates the effects of wind speed and net radiation on stability from data which is routinely available in METAR; cloud cover, wind speed, and solar altitude (a function of time of day, time of year, and lat/lon location). If required by the specific LAQ model used, 'Pasquill-Gifford' stability categories can then be converted into other stability class categories, such as the 'Klug-Manier' stability class (used by such LAQ models as ALAQS and LASPORT), by a simple conversion.

While stability class parameters can be estimated using METAR data, other stability parameters, required by certain LAQ models cannot be so easily estimated using METAR data; parameters such as Monin-Obukhov length, boundary layer height, and roughness length (used by such LAQ models as AERMOD and ADMS). Given these stability estimation considerations, the user will need to decide whether METAR will be sufficient for their use, depending on the estimation scheme adopted, the LAQ model used and the allowed error toleration. If not then one of the two remaining meteorological sources described in this paper, MM5/WRF or REANALYSIS, can be tried.

A further caution about using METAR needs to be addressed. Although being observed data, which usually means there will invariably be missing data, METAR data has very few missing data gaps as METAR is used directly by the aviation community on a continuous basis. As a rule of thumb, the larger the airport, the less likely that there will be missing data. The user is cautioned, however, that even METAR data must be checked for gaps before use. Further, METAR data is reported only during opening hours. This is not a problem for airports such Heathrow (EGLL) which is open 24 hours each day of year. Other airports, such as London City (EGLC), are open only for part of each day. These time gaps need to be filled either by substituting data from another airport which is open and close by or using data from either MM5/WRF or REANALYSIS.

MM5/WRF Numerical weather prediction model

Secondly, in cases where the use of METAR (or other observed data) is not readily applicable, either because these data are not available, or more sophisticated 3D gridded meteorological fields are required by the specific dispersion model (models such as CAMx or CMAQ), the use of meso-scale numerical weather prediction (NWP) models is the next preferred option. These NWP models can produce high quality 'best guess' meteorological fields on a wide variety of time and distance scales. These models are typically the weather forecast products available from national weather services. One such product, freely available and widely used for data collection, is MM5[iii] which, with its successor WRF[iv], was developed by the National Centre for Atmospheric Research (NCAR).

MM5/WRF has been used for a broad spectrum of theoretical and real-time studies, including applications of both predictive simulation and four-dimensional data assimilation and can provide all the surface and upper air meteorological data needed for LAQ studies including stability and boundary layer parameters. MM5/WRF produces very accurate data sets on very fine spatial and temporal grids. However, MM5/WRF needs specialized training in order to run and requires intensive CPU time to run which limits the simulation period. As a rule of thumb, approximately one hour of real CPU time is required for each day of model simulation. This means that long-term local air quality studies (of the scale of one year, for instance) may be prohibitively expensive in terms of effort expended. Another consideration if to be taken into account is the fact that a separate MM5/WRF study would need to be conducted for each airport of interest so that multi-airport studies would usually be prohibitively time consuming to conduct.

Reanalysis data

Finally, this paper outlines a third approach to obtaining meteorological data when direct observations, such as METAR, are not available and when NWP modelling is not a viable solution. This approach is the use of long-term, globally-archived, gridded meteorological fields which are readily available and cover long-term time scales. One such product is the National Center for Environmental Protection (NCEP) and NCAR Reanalysis-2 project[v]. This product uses a state-of-the-art analysis/forecast system to perform data assimilation using past global meteorological data from 1948 to the present. There are over 80 different variables available in the Reanalysis data set, (including geopotential height, temperature, relative humidity, U and V wind components) on 17 pressure levels (heights) on 2.5x2.5 degree grids, four times daily. Reanalysis-2 was created to improve upon the earlier NCEP/NCAR Reanalysis by fixing the errors and by updating the parameterizations of the physical processes. NCAR's Reanalysis data is usually stored in two formats NetCDF and GRIB both of which are commonly encountered meteorological formats and several specialized freely available products are available to 'de-code' theses formats[vi].

As in the case of METAR, stability class is not stored in the Reanalysis-2 data. This parameter needs to be estimated using an algorithm like the Turner method mentioned above. Another common stability parameter, the Monin-Obukhov length, can be calculated using Reanalysis data from the momentum flux, from which the friction velocity is determined, the sensible heat flux and the temperature. These parameters are all stored in the Reanalysis-2 data. Reanalysis-2 has the advantage that long term data sets can be extracted quickly and data can be extracted for many different sites (globally) at the same time. Although less accurate than METAR or NWP models this approach may be of benefit to those users who require 'good guess' meteorological data for air pollution studies but are unable to obtain observed data where observations and cannot use NWP.

The user should be warned, however, that for the purpose of LAQ modelling Reanalysis data has the disadvantage that the data is stored on a very course spatial and temporal grid. This means that the data usually cannot be used directly but must be spatially and temporally interpolated. While this would not be so severe a restriction in areas such large, homogenous grasslands, the problem is especially disconcerting when performing LAQ studies in areas where the meteorological and stability conditions can change rapidly over distance and time (such as in mountainous terrain). In such high change areas, Reanalysis data should be used with caution. Spatial interpolation schemes to be tried in these areas could include a Cressman-style or Multiquadric analysis. Temporal schemes could include a bi-cubic spline interpolation applied to the data. It is beyond the scope of the paper to go into the details of these data interpolation methods.

3. METHODOLOGY

Table 1 summarises the suitability (benefits and problems) of each of the three meteorological data sources described above for the application of LAQ modelling.

Table 1. Comparison of the suitability of using the three proposed meteorological sources for local air quality modelling.

Suitability of Met source for LAQ studies	**METAR**	**MM5/WRF**	**Reanalysis**
Directly observed data?	**Yes**	**No** (see 1. below)	**No** (see 1. below)
Gaps in data?	**Yes** (see 3. below)	**No**	**No**
Freely available?	**Yes**	**Yes**	**Yes**
Requires specialized expertise to use?	**No**	**Yes**	**No**
Suitable for single airport study?	**Yes**	**Yes**	**Yes** (see 2. below)
Suitable for multiple airport study?	**Yes**	**No** (see 4. below)	**Yes** (see 2. below)
Suitable for models requiring only single location for met source (e.g. as required by ALAQS, ADMS or AERMOD, etc.)?	**Yes**	**Yes**	**Yes**
Suitable for models requiring gridded met input (e.g. as required by CAMx or CMAQ, etc.)?	**No**	**Yes**	**No**
Suitable for short time steps (e.g. hourly or half-hourly)?	**Yes**	**Yes**	**Yes** (see 3. below)
Suitable for long-term study (e.g. yearly or seasonal)?	**Yes**	**No** (see 5. below)	**Yes**
Stability parameters explicitly provided (e.g. P-G stability class or Monin-Obukhov length, etc.)?	**No** (see 6. below)	**Yes**	**No** (see 6. below)

1. Data is forecast data based upon observed data
2. May require spatial interpolation of data
3. May require temporal interpolation of data
4. Requires multiple runs (one for each airport)
5. Requires significant time and CPU resources
6. However, routine estimation methods can be used

A methodology which can be used to identify which of the three data sources described above would best be used for a specific study is the following.

1) First determine whether the dispersion model requires gridded meteorological input or not (for instance, ALAQS, LASPORT, ADMS or AERMOD do not require gridded input, whereas CAMx and CMAQ do require it). Of the three sources described in this paper only MM5/WRF can readily produce gridded input and so if gridded data is required then an NWP model such as MM5/WRF would be the best (and only) choice for the study.
2) If single position (that is, non-gridded) meteorological data is required then the user needs to determine whether METAR data is the best source to use. First the question to ask is whether METAR is available for the required time period (checking for missing data and airport opening hours). The next issue the user needs to determine is whether the stability parameter required by the LAQ model can be estimated using METAR. This usually reduces to the question of whether stability 'class' is sufficient for the input stability requirements (either directly, or by a simple algorithmic procedure) since stability class can be estimated from METAR data using the above mentioned Turner method. An affirmative answer to both these questions suggests that METAR is the best source for the study.
3) Next, in a similar way as the previous step, the user needs to determine is whether the stability parameter required by the LAQ model can be estimated using Reanalysis data. Stability class, Monin-Obukhov length, and boundary layer height can all be estimated using Reanalysis data. If other stability parameters (or if the estimation procedure has been determined to be too 'inaccurate') then the best option would be MM5/WRF
4) Otherwise if Reanalysis stability data is sufficient and the study to be performed is a long-term (e.g. a full year or season) study, then REANALYSIS would probably be the best option available. It must be kept in mind, however, that Reanalysis data usually needs to be spatially and temporally interpolated to be useful.

The following diagram outlines this decision procedure:

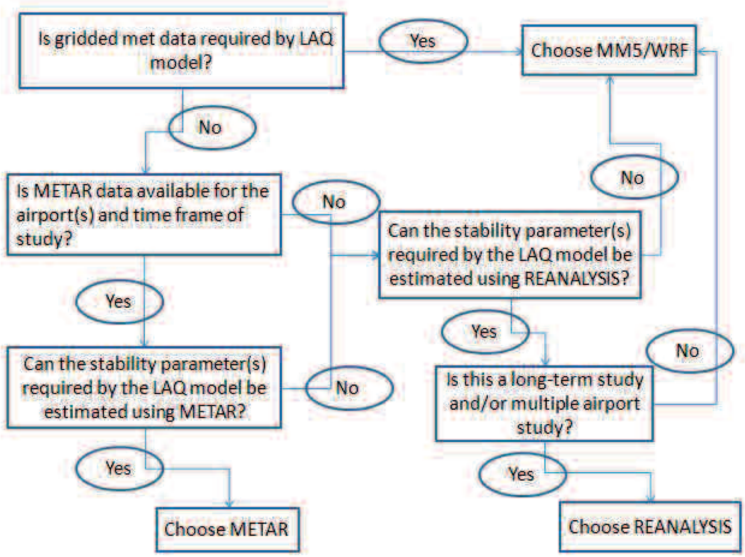

Figure 1. Flow chart of decision procedure for LAQ meteorological data source.

4. SUMMARY AND CONCLUSIONS

Three freely available meteorological data sources have been identified along with a discussion the pros and cons of each data source. A methodology has been presented to help the LAQ modeller, who may wish to use these sources, to choose which source would be the best choice for a specific LAQ study. METAR, an observed data source, should preferencially be used except in cases where the LAQ model requires more sophisticated requirements such as gridded input data. In cases of gridded data requirements, the choice of meteorological data source would be to use a numerical weather prediction model such as MM5/WRF. However, these models require significant investment in time and effort. A third option, archived Reanalysis data, can be used as a last resort but this option requires spatial and temporal interpolation in order to be useful.

NOTES

---

[i] One of the sites where global METAR data can be found is http://weather.noaa.gov/weather/metar.shtml
[ii] A clear description of the turner method of estimating Pasquill-Gifford stability class can be found at http://www.webmet.com/met_monitoring/641.html
[iii] The MM5 homepage is http://www.mmm.ucar.edu/mm5/
[iv] The WRF homepage is http://www.wrf-model.org/index.php
[v] Reanalysis-2 data can be found in NetCDF format at http://www.cdc.noaa.gov/cdc/data.ncep.reanalysis2.html and in GRIB format at http://nomads.ncdc.noaa.gov/data.php
[vi] For instance see the Integrated Data Viewer (IDV) at http://www.unidata.ucar.edu/software/idv/