

Data assimilation in meteorological pre-processors: Effects on atmospheric dispersion simulations

[E. Davakis^a](#), [S. Andronopoulos^b](#), [I. Kovalets^c](#), [N. Gounaris^b](#), [J.G. Bartzis^d](#), [S.G. Nychas^a](#)

^a Department of Chemical Engineering, Aristotle University of Thessaloniki, University Box 453, 54124, Thessaloniki, Greece

^b Environmental Research Laboratory, Institute of Nuclear Technology and Radiation Protection, NCSR ‘Demokritos’, 15310 Aghia Paraskevi, Greece

^c Department of Environmental Modelling, Institute of Mathematical Machine & System Problems, National Academy of Sciences of Ukraine, pr. Glushkova-42, Kiev 03187, Ukraine

^d Department of Engineering and Management of Energy Resources, University of Western Macedonia, Bakola & Sialvera str., 50100, Kozani, Greece

* Corresponding author: Tel. +30 210 650 3426, Fax. +30 210 652 5004, e-mail:

sandron@ipta.demokritos.gr

E-mail addresses: stratos@ipta.demokritos.gr (E. Davakis), ik@env.com.ua (I. Kovalets),

gounaris@ipta.demokritos.gr (N. Gounaris), Bartzis@uowm.gr (J.G. Bartzis),

nychas@vergina.eng.auth.gr (S.G. Nychas)

Abstract

In previous work (Kovalets et al., 2004) the authors have developed data assimilation (DA) procedures and implemented them in the frames of a diagnostic meteorological pre-processor (MPP) to enable simultaneous use of meteorological measurements with Numerical Weather Prediction (NWP) data. The DA techniques were directly validated showing a clear improvement of the MPP output quality in comparison with meteorological measurement data. In the current paper it is demonstrated that the application of DA procedures in the MPP, to combine meteorological measurements with NWP data, has a noticeable positive effect on the performance of an atmospheric dispersion model (ADM) driven by the MPP output. This result is particularly important for emergency response systems used for accidental releases of pollutants, because it provides the possibility to combine meteorological measurements with NWP data in order to achieve more reliable dispersion predictions. This is also an indirect way to validate the DA procedures applied in the MPP. The above goal is achieved by applying the Lagrangian ADM DIPCOT driven by meteorological data calculated by the MPP code both with and without the use of DA procedures to simulate the first European Tracer Experiment (ETEX I). The performance of the ADM in each case was evaluated by comparing the predicted and the experimental concentrations with the use of statistical indices and concentration plots. The comparison of resulting concentrations using the different sets of meteorological data showed that the activation of DA in the MPP code clearly improves the performance of dispersion calculations in terms of plume shape and dimensions, location of maximum concentrations, statistical indices and time variation of concentration at the detectors locations.

Keywords: Data Assimilation, Atmospheric Dispersion Models, Meteorological Pre-processors, Model Evaluation, ETEX

1. INTRODUCTION

Atmospheric Dispersion Models (ADMs) are widely used for the assessment of air quality or for the prediction of pollutants dispersion following accidental releases. It is well known that the performance of these models depends crucially on the meteorological driving data. The latter are obtained from prognostic meteorological models and/or from meteorological stations, after they have been processed by Meteorological Pre-Processors (MPPs) or diagnostic wind flow models (e.g., Seaman, 2000). This processing is necessary to calculate eventually missing variables and to adapt the original data to the computational grid of the ADM (e.g., to assure mass consistency of the wind field on the ADM grid).

The Numerical Weather Prediction (NWP) models require significant computational resources (time and space) and it is not practical to run them on request each time a dispersion calculation needs to be performed. This is especially true for emergency response systems (ERSs) like RODOS (the Real-time On-line DecisiOn Support system for nuclear emergency management in Europe; (Raskob and Ehrhardt, 1999)) that require very fast predictions (less than ten minutes) of pollutant dispersion. Therefore ERSs rely on NWP model output produced regularly on specific time intervals (e.g., new prognostic data become available every six hours) and possibly on meteorological measurements made in the area of interest during the time interval in question.

The issue of using simultaneously both types of data (i.e., from NWP model and from measurement stations) in the same MPP has been addressed by the authors (Kovalets et al.,

2004). This was done to enable exploiting all the meteorological information available at a certain time, and to avoid the differences between ADMs results produced with the use of only NWP models forecasts or locally measured meteorological data. The problem has been tackled by developing and introducing 3-Dimensional Data Assimilation (3DDA) procedures in the MPP. For the assimilation of scalar variables (surface temperature, cloud cover, net radiation, precipitation) the objective analysis procedure known as “iterations to optimal solution” (IOS - Daley, 1991) has been adopted, while for the wind velocity, either the IOS or a multivariate optimal interpolation was used (Kovalets et al., 2004).

4-Dimensional Data Assimilation (4DDA) is already used in the NWP codes that provide the original data for the ADMs. However, as it was discussed in Kovalets et al. (2005), if there are additional measured meteorological data then it is necessary to use DA procedures inside the MPPs too, otherwise the above information will be lost. This is particularly important in the MPPs that are incorporated into emergency response systems which require the most updated information. Moreover, the scales of the atmospheric movements resolved by the NWP model and the MPP are significantly different because the grid scale of the MPP is usually finer than that of the NWP model. Hence, the small-scale movements present in the measurements that are treated as “noise” when used in the NWP model are to be resolved by the MPP. Thus, even if some particular meteorological observations were already used for the calculation of the NWP forecast, they should be used again if these forecasts are being pre-processed on the finer grid of the MPP.

The authors validated the developed methods by comparing the MPP output against the meteorological measurements performed during the European Tracer Experiments (Kovalets et al., 2004). The comparisons showed that the DA procedures improved the finally calculated

wind velocity field in terms of better agreement with the experimental data. However, it is important — especially in the frames of emergency response systems — to examine if the adoption of DA procedures in the MPPs improves also the atmospheric dispersion model calculations driven by the MPPs output. Other authors (e.g. Seaman, 2000, Hurley et al., 2003, Luhar and Hurley, 2004, Deng et al., 2004) have presented similar studies, where data assimilation was introduced in prognostic meteorological codes, which directly provided data to the ADM.

Therefore the aim of this work is to investigate whether the performance of an ADM is improved when it is driven by the output of an MPP that applies meteorological DA to merge NWP data with meteorological measurements in comparison to the case when the MPP does not use DA. This is achieved by running an ADM code for a real case driven by data produced by an MPP with and without DA procedures. In this context, a modified version of the Lagrangian particle dispersion model DIPCOT (Davakis et al. 2000, 2003 2005) was applied to simulate the ETEX-I (Grazianni et al., 1998). The ETEX-I has been selected because of the availability of forecast data from the NWP model of the European Medium-Range Weather Forecast Centre (ECMWF), meteorological measurements and of concentration measurements. Three applications of the DIPCOT model were performed, using different meteorological data sets. In the first application we used the MPP output obtained by interpolating only the prognostic meteorological fields to the ADM grid. In the second application, the MPP output was obtained interpolating both NWP and measurement data on the ADM grid without applying DA, while in the third application we used the MPP output obtained by applying DA techniques to combine the meteorological measurements and the prognostic fields from the ECMWF. In all three applications a divergence minimising procedure was applied to the wind velocity output of the MPP to assure mass conservation.

The ADM results for the abovementioned cases were statistically and qualitatively compared with the observations and between themselves. The analysis of the results showed that the performance of DIPCOT was acceptable in all cases; however, the activation of DA to merge NWP data with meteorological measurements clearly improved the results of dispersion. This result validates also indirectly the use of DA procedures in the MPP code. It is also very important regarding emergency response systems used for accidental releases of pollutants, which require updated input information and should provide reliable dispersion predictions.

2. MODEL DESCRIPTION

2.1 The MPP code

The MPP used in this work is a diagnostic meteorological model (FILMAKER - Andronopoulos et al., 2005) that produces gridded data sets of variables such as wind velocity, temperature, mixing layer height, atmospheric stability, etc., based on meteorological prognostic data and on measurements. The output of the MPP is used to drive atmospheric dispersion calculations. The horizontal computational grid of the MPP is Cartesian while the vertical is terrain-following, both non-equidistant. The meteorological variables for which prognostic data or observations exist are calculated on the computational grid by spatial $1/r^2$ interpolation in the horizontal direction from the NWP model grid or from the observation points. For the variables without available observations (sensible heat flux and other), semi-empirical relations are used (Hanna and Chang, 1993, Zannetti, 1990, IAEA, 1980, Seibert et al., 1997, Desiato and Palmieri, 1988). In the vertical direction logarithmic, power-law, linear or exponential functions are used for interpolation, depending on the variable (wind velocity, temperature, pressure, humidity).

The DA procedures that were developed and incorporated in the above MPP have been based on techniques developed earlier for the needs of weather prediction models. These DA methods have been modified taking into account that the grid of the MPP is usually finer than the NWP models, and it should resolve the micro to meso-scale features of the wind flow. After having calculated the “first guess” meteorological variables fields by $1/r^2$ interpolation on the MPP grid from the NWP grid, assimilation of the scalar variables measurements is performed, followed by assimilation of the wind velocity measurements. For the scalar variables (surface temperature, cloud cover, net radiation, precipitation) the 3DDA procedures implemented in the MPP were based on the objective analysis procedure known as “iterations to optimal solution” (IOS - Daley, 1991). Assimilation of the wind velocity observations can be performed by two alternative methods: the IOS or a multivariate optimal interpolation (OI) algorithm. The latter has been used in this study. Finally, the rest of the meteorological variables (for which routine measurements are not usually available, such as mixing layer height, sensible heat flux, stability category, friction velocity, Monin-Obukhov length) are affected indirectly by the DA procedures, through their dependence on the wind velocity and net radiation. A detailed description of assimilation procedures used in the MPP code is given by Kovalets et al. (2004).

2.2 The ADM code

The ADM used in this study is a modified version of the Lagrangian particle dispersion model DIPCOT (DisPersion Over Complex Terrain); DIPCOT is a 3D air pollution model, which simulates atmospheric dispersion estimating particle trajectories based on Langevin equation.

The particles are advected at each time step using

$$x_i^{n+1} = x_i^n + (\bar{u}_i + u_i' + u_{ti}')\Delta t \quad (1)$$

where x_i is the position of the particle in the i^{th} direction, n is the time step, $\overline{u_i}$ is the mean wind component in the i^{th} direction at the n^{th} step, provided by an MPP code, u'_i represents the turbulent velocity fluctuations and u'_{ii} the velocity the low frequency horizontal meandering and Δt the time step. Assuming i) that the velocity and the position of a particle evolve as a Markov process and ii) mutual independence of the three velocity components, u'_i is estimated using the one-dimensional Langevin equation.

$$du'_i(t) = \alpha_i dt + \sqrt{C_o \varepsilon(z)} dW_i \quad (2)$$

where dt is the time increment, α_i is the acceleration – drift term in the three directions ($i = x, y, z$), $\varepsilon(z)$ is the ensemble-average rate of dissipation of turbulent kinetic energy and dW_i are one-dimensional increments of a Wiener process – a Gaussian random forcing with zero mean and variance dt . The dissipation rate ε is computed based on the assumption that it is a height-dependent function (Rodean, 1994). The constant C_o is a universal constant, associated with the Lagrangian structure function (Rodean, 1991, Shuming, 1997). There is a considerable uncertainty about its value. Different investigators indicated different values ranging between 1 and 10. In this work we set the value of C_o equal to 3. This value was suggested as a mean by Shuming (1997) and it gave best fit between several values that were examined using our laboratory data (Davakis et al., 2005). The deterministic acceleration term α_i is derived using the “well mixed condition”, proposed by Thomson (1987).

The DIPCOT model has been modified for this application in order to account for the low frequency horizontal meandering. The perturbation u'_{ii} in equation (1) is taken from a formula exactly analogous to Langevin equation, but substituting the meander statistics. Maryon (1997) has introduced this methodology, which was applied by Ryall and Maryon (1998) for the ETEX data set. The velocity variances that this parameterisation generates are similar in

magnitude to the turbulent velocity variances. The range of horizontal motions lying between the scales resolved by the numerical model (say $2T_f$, where T_f is the interval between the wind fields) and the scales parameterised through the turbulent diffusion constitute the meander range. Meander variances and timescales were estimated from monthly spectra selected from the Meteorological Research Unit at Cardington, England (Maryon, 1997).

The trajectories of the released particles are computed using the above-mentioned equations. The pollutant concentrations are calculated by considering the number and the masses of particles present at a certain time in predefined grid cells that cover the area of interest (Davakis et al., 2003).

3. THE EXPERIMENTAL DATA SET

The European Tracer Experiment (ETEX - Grazianni et al., 1998) was performed in October 1994 and involved the release of a passive, non-reactive and non-depositing gas (perfluoro-methyl-cyclo-hexane - PMCH) from a source located in western France and the subsequent dispersion over North Europe. The gas release started at 16:00 UTC on 23 October 1995 approximately 35Km west of Rennes, at Monterfil, in Brittany, France, 90m above sea level at a constant rate of 7.95 g s^{-1} . The release duration was 12 hours. For a time period of 72 hours after the release a ground net of 168 sites in 17 European countries east of the release site performed concentration measurements and produced 3hours-averages. Besides the tracer concentrations measurements, the experimental database contains ground and upper air meteorological observations collected from the start time of the tracer release and three days ahead, as well as prognostic meteorological fields from the ECMWF (Gryning et al., 1998,

Straume and Nodop, 1997). Therefore it is a suitable case for the purposes of the current study.

The measurements performed during the experiment covered a domain of almost 1500 x 2000 Km². This is much more extended than the computational domain of the MPP used in this work, which is designed for calculations on the local to meso-scales. So the calculations performed in this work covered a domain of dimensions approximately 1000 x 700 Km² around the point of tracer release in the ETEX (Monterfil, see Fig. 1). The computational domain was discretised by 5 x 5Km² horizontal grid and the models were applied for 33 hours after the release start. From the 168 concentration receptors 66 fall into the selected domain following the suggestions of Dubois et al., 2005 (see Figure 2).

From the ETEX-I meteorological database (Straume and Nodop, 1997) 120 observations stations were used. Specifically data from 47 ground-based synoptic weather stations located near the source (over France), along with surface observations from other 66 similar stations over Europe, presented in Figure 1 were used. The last ones contributed in better estimates of net radiation and stability category, because these stations contained, in addition to measurements of surface wind velocity, temperature and pressure, measurements of cloud cover, not available at the other ground stations. Vertical wind and temperature profiles were also included through the use of seven sodar stations. The prognostic, non-analysed, ECMWF meteorological data have been processed in this work by the MPP as the meteorological model forecast (NWP data). These data are available at the surface and at four upper levels for the day of the tracer release plus the next three days at the grid points shown in Figure 1. At the surface, pressure, total cloud cover, 10-metre wind (u, v) and 2-metres temperature are

given. At the upper levels (1000, 850, 700, 500 hPa), geopotential, temperature, wind speed (u, v, w) and humidity are given.

4. MODEL APPLICATION

Three applications of the MPP code were performed based on different input data sets and computational methods. In the first case only the NWP data were used (from now on referred to as “NWP case”). In the second case the NWP data and the measured variables from the meteorological stations were used simultaneously in an equal way, without using DA (from now on called “AM case”). The third application, referred to as “DA case”, was based again on both data sets but this time data assimilation procedures were switched on. In the latter case, the observed meteorological data from all the measurement stations was used to improve the first-guess fields through the data assimilation procedure.

The DIPCOT model performed dispersion calculations for the three meteorological data sets produced by the MPP in the above cases. Regarding the model particles release rate, a sensitivity study using 300 and 4000 particles per hour gave almost identical results. In the rest of the paper the presented results were obtained with the former release rate. The ADM was applied for 33 hours after the release start. For the evaluation of the model performance the model predictions were compared with the observations from 66 receptors that fall in the selected domain of the ETEX-I. For this purpose, the calculated gridded concentrations were linearly interpolated to the locations of the concentration detectors. The predicted concentrations were compared with concentrations from the so-called “Global analysis” data set (Mosca et al., 1998). This means that the experimental data set consists of all the zero values six hours before the cloud arrival and six hours after the cloud departure at any location as well as all the zeros between non zero values. The ETEX experimental data set

contains a significant amount of zero measured values (34%), which influence the statistics. For these reasons, in all the indices and graphs involving the logarithm of the predicted and observed concentration values, the zeros have been set equal to 0.01 ngr/m^3 , which is the sensitivity limit of the receptors (Mosca et al., 1998).

5. EVALUATION PROCEDURE

In this section the methods used for the evaluation of the ADM performance under the three input data sets are listed.

The ADM predictions are compared with the experimental values using well-known statistical indices (e.g. Hanna, 1993, Mosca et al., 1998) as: the Fractional Bias (*FB*), the Geometric Mean bias (*MG*), the Normalized Mean Square Error (*NMSE*) and the Geometric mean Variance (*VG*). The *FB* and *MG* are measures of the deviations between model predictions and observations. Since the *FB* is calculated from differences while the *MG* from ratios between C_p and C_o (C_o and C_p being the observed and predicted concentration values respectively), *FB* gives more weight to large values of concentrations, while *MG* gives equal weight to all pairs. A perfect model would have *FB* equal to zero and *MG* equal to one. Values of *FB* less than zero and *MG* less than unit mean that the model over-predicts the observations. The statistical measures of *NMSE* and *VG* give information on the spread of the deviations values (not on the over- or under-prediction) and thus they are always positive. The closer the *NMSE* value to zero is the better is the model performance. For *VG*, a perfect model would have a value of unit. Differences on peak values between the predicted and the observed concentrations have a higher influence on the value of *NMSE*, while *VG* assigns the same weight to all pairs of values with the same ratio. All the above indices were estimated by the MEteorological and DIspersion STATistics code (Deligiannis et. al., 1997). The 95%

confidence interval limits for FB and MG were also estimated, based on the Bootstrap re-sampling method (Hanna, 1989). It consists in re-sampling for a number of times (1000 in this case) the set of pairs, with possible repetitions of pairs, and each time re-estimate the statistical indices. From the distribution of these values, the values corresponding to 2.5% and 97.5% cumulative probability are taken as the limits of the confidence interval. In a “statistically correct” model the 95% confidence limits for FB and VG should include zero and unit respectively. The values of FB with its 95% confidence limits are plotted against the related $NMSE$ values and the MG values against the VG .

Another quantitative index of ADM performance with the three meteorological data sets is the “factor-of- x ” (FACT x), which is the percentage of (C_p, C_o) pairs for which $1/x \leq C_p/C_o \leq x$. Obviously the larger this factor is, the better is the agreement between model predictions and measurements. Usual values of x in model evaluation studies are 2, 5 and 10.

In order to examine any model bias over the concentration distribution the Quantile–Quantile (Q-Q) plot has been drawn. All predicted concentrations were grouped and then sorted according to magnitude, and so were also the observed concentrations. The sorted predicted concentrations from the three meteorological data sets (NWP, AM and DA) have been plotted against the sorted observed values in logarithmic scale. There is no time or position correspondence between the concentrations of the two-sorted groups, and the ranked pairs of concentrations that are plotted do not necessarily correspond to the same events.

As it was mention before, the Global Analysis set of the ETEX contains a great number of zero concentrations. In particular almost 37% of the experimental concentrations used in this study were zero. Thus, for a model that would not estimate any concentrations an evaluation

procedure based only on the above-mentioned statistical indices would produce an incorrect view of its performance. For this reason the percentage of the non-zero observed concentrations (*%nz*) that the ADM predicts in each case is estimated (there are 150 non-zero measured concentrations in all the receptors for the selected computational domain and time).

The model performance is also evaluated qualitatively using plots where contours of the experimental ground level concentrations are overlaid on the predicted ones at specific times (12, 24 and 33 hours after the release start). These plots show the time and space evolution of the predicted plume in relation to the spread of the real plume giving a qualitative view of the effect that the different meteorological data set have on dispersion calculations. Finally time series of the predicted against the measured concentrations are plotted at several stations placed along the real plume centreline. This shows how well the ADM predicts in each case the plume arrival and departure times, as well as the peak concentrations.

6. RESULTS AND DISCUSSION

The concentration contour plots (Figures 3 to 5) give a view of the overall effects that the different meteorological data sets used in this study had on the ADM calculations. During the first 12 hours after the release start there are some differences in the plume centreline direction, downwind extent and maximum concentration values. When meteorological observations are taken into account in addition to the NWP data, the plume centreline is slightly shifted towards the north and also the plume east front is extended towards east-northeast. This behaviour is further enhanced when the DA procedure is activated (see Figure 3c in comparison to 3b and 3a). At the same time under the influence of DA the north-west front of the plume is shifted further to the west. Thus the wind field stretches the cloud in horizontal direction in the case of DA. That stretching is better revealed on the Figures 4, 5 in

the cases of both: DA and measurements. By looking at the contour values at the Fig. 3 it can be noticed that the plume calculated using the NWP data only underestimates the maximum concentration observed experimentally, while the plume predicted in the AM and DA cases overpredicts the observed maximum values. However the above differences are rather small.

As the plume departs from the source, the effects of including the meteorological measurements in the calculations become more obvious. At 24 hours after the release start (Figures 4a-c) the south-west- part of the plume in the DA case is closer to the observed one than the other two, concerning both the location of the main concentration maximum over the north of France and the direction of the plume centreline first to the east and then to north-east. In contrast, the NWP case predicted plume has a direction north to south, with a concentration maximum over the Netherlands. From the contour values it is concluded that the calculated plumes underestimate the maximum observed concentrations in all cases, however this underestimation is smaller in the DA case.

As it is seen from the Fig. 4 the most essential influence of the DA algorithms is in the west-south part of the cloud, which is closer to the source. This is possibly due to the higher density of the meteorological measurements in that part of the domain. The quality of the wind field closer to the source can be especially important for the quality of the subsequent dispersion calculations because initially the concentration gradients in the cloud are large and small variations in wind velocity close to the source can lead to large concentration deviations even far from the source.

At 33 hours after the release start (Figures 5a-c) the effects of the meteorological DA procedures are also noticeable, resulting in a calculated plume for the DA case that has better

overlap with the observed one than the other two, especially in the location of the area of maximum concentration, that was advected from the west-south corner of the domain (Fig. 4). The values of the contour lines indicate that in all cases the calculated plumes underestimate the maximum observed concentration values. The underestimation is smallest when only the NWP data are used. Nevertheless the differences in maximum concentrations values between the calculated plumes are very small.

From the above it is apparent that when meteorological observations in addition to the NWP data are taken into account by the MPP the resulting data set improves the predictions of the ADM regarding the plume direction and spread. The effect of DA becomes more pronounced at greater distances from the tracer release location. Furthermore, the activation of the DA procedures in the MPP has an additional beneficial effect on the performance of the ADM, resulting in a plume that overlaps better with the observed one and also agrees more in the location and magnitude of maximum concentration. Even though the meteorological stations are more dense close to the gas source (see Figure 1) and the corrections in the meteorological fields produced through the DA of the meteorological measurements are relatively small – as can be concluded by the small differences between the predicted plumes at early stages (Figure 3a-c) – the effects of the above corrections on the ADM predictions are significant at later stages of the dispersion (Figure 4 and 5).

The first quantified measure that is examined for the overlap between predicted and observed plumes is the percentage of the non-zero observed values (*%nz*) that the ADM predicts. This index takes the values of 66.67%, 62 % and 59.33 % respectively for DA, AM and NWP cases (that in absolute numbers stands for 100, 93 and 89 predicted non-zero concentrations

over a total of 150 values). This demonstrates that, when meteorological DA is used, the predictions of both the plume direction and the plume spread are improved.

The $FACT_x$ ($FACT_2$, $FACT_5$ and $FACT_{10}$) indices are presented for the DIPCOT predictions in Figure 6, using the 3 meteorological data sets that were produced by the MPP. It can be seen that when the meteorological observations are taken into account (AM, DA cases) in addition to the NWP data (NWP case), the indices increase, meaning a better agreement of the ADM predictions with the observations. The maximum values for all factors are achieved when the DA procedures are activated in the MPP (DA case).

In figure 8, the $NMSE$ of the ADM predictions for the three cases is plotted against the FB . It is noted that the performance of the ADM is best when it is driven by the meteorological data set produced with the DA procedures activated in the MPP. The DA case presents an FB value close to zero (which is included in its 95% confidence interval) with the smallest $NMSE$ value. The FB values of the ADM predictions in the other two cases are larger and the zero value is not included in their 95% confidence intervals, meaning that they are significantly different than zero at the 95% confidence level. It is reminded that the FB by definition gives more weight to the large concentration values since it is constituted by the differences between predicted and observed values.

On the other hand the Geometric Mean Bias (MG) is constituted by the ratios between predicted and observed concentration values and it therefore attributes equal weight to all data pairs. The MG of the ADM predictions is plotted against the corresponding Geometric Mean Variance (VG) values in Figure 8 for the three cases of meteorological data sets produced by the MPP. The MG of the ADM predictions is closest to unit for the case of the meteorological

data calculated with activated the DA procedures. However the *MGs* of all cases include the value of unit in their 95% confidence intervals, meaning that they do not differ significantly than unit at the 95% confidence level. Nevertheless the *MG* 95% confidence interval for the DA case is the smallest of the three cases. The smallest value of *VG* resulted in the NWP case (meaning the smallest spread of values), however the differences between the *VG* values are rather small between the three cases. Therefore and as long as the *MG* values are concerned there is no clear distinction of the ADM performance when using the three different meteorological data sets.

The conclusion from the analysis of the statistical indices is that when the meteorological observations are used together with the NWP inside the MPP (case AM), the resulting data set improves the performance of the ADM in comparison to the case when only the NWP data are used (case NWP). The *FACT2*, *FACT5* and *FACT10* are increased and the *MG* value is closer to unit. The *FB* is not affected and is different than zero at the 95% confidence level, although the *NMSE* is reduced. When in addition the DA procedures are activated inside the MPP (case DA), the ADM performance is further improved as indicated by the increase of *FACTx*'s. The *FB* value goes very close to zero and the *NMSE* is further reduced. The *MG* is not drastically affected but still is even closer to unit and its 95% confidence interval is reduced. Therefore the DA procedures in the MPP have a more pronounced effect on the larger concentration values predicted by the ADM. This is in agreement with the conclusions drawn from the concentration contour plots, where it was observed that when the ADM was driven by the DA data set it gave a better prediction of the plume maximum concentration areas.

The Q-Q plot presented in Figure 9 indicates that there is a bias in the DIPCOT results to under-predict the experimental concentration distribution for both NWP and AM cases. This tendency is greater in the NWP case. This is in agreement with Figure 7, where the *FB* values for the cases NWP and AM are greater than zero and also different than zero at the 95% confidence level. When the ADM uses the DA data set then its predictions are improved in comparison to the AM case at the higher concentrations, verifying also the conclusions of the statistical analysis.

The effects from using the different meteorological data sets in the ADM calculations are also shown in Figure 10, where the calculated and experimental concentrations are plotted against time from the release start at the locations of six sensors along the ETEX plume path. The locations of the sensors can be seen in Figure 2. The sensors in Figure 10 are listed with their distance from the release location, with sensor 40 being the closest to the source and sensor 15 the farthest. It can be seen that the activation of DA procedures in the MPP decreases the discrepancies between the ADM predictions and the experimental data at stations 40, 46, 5 and 15 (Figure 9 a-d) which are located near the predicted plume centreline. For the sake of completeness, in Figure 9 two other stations (12 and 39) are included, which lie at the edge of the calculated plume, where the ADM predictions are better when the DA procedures are not activated in the MPP.

The above analysis demonstrates the degree to which the predictions of an ADM are improved – in terms of better agreement with observations – when 3DDA procedures are used in the meteorological pre-processor that provides the driving meteorological information. This is also an indirect confirmation of the more accurate calculation of the meteorological fields calculated by the MPP when DA is used to combine measurements with NWP data.

7. SUMMARY AND CONCLUSIONS

In previous work (Kovalets et al., 2004) the authors have developed and incorporated in a meteorological pre-processor (MPP) 3-dimensional Data Assimilation (DA) procedures in order to exploit simultaneously and in an optimal way both the Numerical Weather Prediction (NWP) data and the meteorological measurements. In the above-mentioned study the DA procedures have been directly evaluated, showing a clear improvement of the MPP output quality in terms of better agreement with the meteorological measurements. The present paper investigates whether and to what extent the results of an atmospheric dispersion model (ADM) are improved when the latter is driven by the output of the MPP produced with the DA techniques activated. In this respect, the MPP FILMAKER was used, along with the Lagrangian particle dispersion model DIPCOT to simulate the ETEX-1 for which an extended database exists, including NWP data, meteorological measurements and concentration measurements.

The MPP code was run using prognostic data from the NWP model of the ECMWF and meteorological measurements from stations located in the computational domain. Three cases of MPP calculations have been performed: i) using only NWP data, ii) using both NWP data and meteorological measurements in an identical way, i.e., without activating DA procedures (i.e., in that case all available meteorological data are treated equally as measurements, as described in Andronopoulos, et.al., 1998), and iii) using both NWP data and meteorological measurements with DA procedures activated. The three output data sets were provided to the ADM to perform dispersion simulations, and the ADM results have been inter-compared and compared with the concentration measurements to evaluate the effect of the DA procedures. The following means have been used for the evaluation: overlaid contour plots of near-ground

calculated and measured concentrations, statistical indices (fractional bias, geometric mean bias, normalised mean square error, geometric mean variance, factors of 2, 5 and 10, and percentage of non-zero concentrations predicted), quantile - quantile (Q-Q) concentration plot and finally plots of concentration time histories at stations locations.

All the above comparisons have demonstrated that the combination of meteorological measurements with the NWP data inside the MPP results in data sets that improve the performance of the ADM. In particular the results of the ADM show the best agreement with the experimental data when the DA procedures are activated in the MPP. This ADM performance improvement is due to the improvement of the quality of the driving meteorological fields produced by the MPP with the DA procedures activated. When the dispersion calculations are driven by meteorological data produced by the MPP with DA activated, the values of the factors of 2, 5 and 10 for concentrations increase, the value of the fractional bias approaches zero as does also the normalised mean square error. The improvement is also apparent in the better overlap between the calculated and experimental concentration contour plots, especially in the areas of maximum concentrations. Moreover the plume spread and the calculated plume motion appears closer to the observed one. The Q-Q plot reveals that the DA in the MPP affects more the high concentration values calculated by the ADM. This is also confirmed by the geometric mean bias that is not affected as much as the fractional bias by the DA, although it approaches the value of unit, including it at the 95% confidence level. Therefore it appears that the effect of the DA procedures is more pronounced for the higher concentrations, which is of greater interest in dispersion calculations for emergency response systems. However, there is also improvement in the estimation of the smaller concentrations. Finally, the majority of concentration time series showed also that when DA procedures were employed in the MPP code the performance of

the ADM model was improved. The above constitute also an indirect evaluation of the DA methods themselves revealing that, when they are applied in the MPP to combine NWP data with existing meteorological measurements, the quality of the calculated meteorological fields is improved.

The obtained conclusions are consistent with the results of (Jackson, et.al., 2006), which showed, that “hybrid meteorological fields, developed by merging the results of objective analysis and prognostic models, can improve episodic air quality model performance over using prognostic meteorological fields alone.”

The above work is a first evaluation step that has demonstrated that the 3DDA procedures incorporated in MPPs have a noticeable positive effect on subsequent dispersion simulation calculations. Further tests must be carried out using other real case data, to investigate remaining issues such as the effects of the spatial scale of the problem, the effects of the monitoring network density, and the influence of the terrain complexity. Also further tests must be carried out using data sets with more upper air meteorological measurements. Nevertheless the above work is a very positive sign towards the operational application of DA procedures in MPPs used in the frames of emergency response systems.

Acknowledgements

Partial funding by the European Commission FP6 (Contract No. FI6R-CT-2004-508843 “EURANOS” and the EURATOM grant RODOS/METADM, contract № 516492 (FI6R)) is gratefully acknowledged by the authors.

References

Andronopoulos S., Davakis E., Goumaris N., Bartzis J.G., Nychas S.G., 2005. Dispersion modelling of radioactive pollutants: Application of the Demokritos Transport Code System

for Complex Terrain (DETRACT) to the Hanford Purex Scenario. *International Journal of Environment and Pollution*, 25, 1/2/3/4, 33–47.

Andronopoulos S. et al., 1998, FILMAKER: a computer model for pre-processing of the weather data. RODOS report WG2_TN98_08, <http://www.rodos.fzk.de>

Daley R., 1991. *Atmospheric Data Analysis*. Cambridge University Press, Cambridge.

Davakis E., Bartzis J.G., Nychas S.G., 2000. Validation of DIPCOT-II model based on Kincaid experiment. *International Journal of Environment and Pollution*, 14 (Nos 1-6), 131-142.

Davakis E., Andronopoulos S., Bartzis J.G., Nychas S.G., 2003. Validation study of the Lagrangian dispersion particle model DIPCOT over complex topographies using different concentration calculation methods. *International Journal of Environment and Pollution*, 20, no. 1-6, 33-46.

Davakis E., Andronopoulos S., Sideridis G. A., Kastrinakis E.G., Bartzis J.G., Nychas S.G., 2005. Evaluation of the Lagrangian particle dispersion model DIPCOT against data from wind tunnel simulations of quasi two-dimensional turbulent flow. *International Journal of Environment and Pollution*, 24, 1/2/3/4, 114–126.

Deligiannis P., Bartzis J.G., Davakis E., 1997. Complex terrain modelling exercise. *International Journal of Environment and Pollution*, 8, 367-377.

Desiato F., Palmieri S., 1988. Review on some methods for the evaluation of meteorological input parameters for dispersion models. Report DISP/ARA/MUT(1988)02, Comitato Nazionale per la Ricerca e per lo Sviluppo dell' Energia Nucleare e delle Energie Alternative.

Deng A., Seaman N.L., Hunter G.K., Stauffer D.R., 2004. Evaluation of Interregional Transport Using the MM5-SCIPUFF System. *Journal of Applied Meteorology*, 43, 1864-1886.

Dubois G., Galmarini S., Saisana M., 2005. Geostatistical investigation of ETEX-1: Structural analysis. *Atmospheric Environment*, 39, 1683-1693.

Grazianni G., Klug W., Mosca S., 1998. Real-time, long-range dispersion model evaluation of the ETEX first experiment. Report EUR 177754 EN, EC.

Gryning S.E., Batchvarova E., Schneiter D., Bessemoulin P., Berger H., 1998. Meteorological conditions at the release site during the two tracer experiments. *Atmospheric Environment*, 32, 4123–4137.

Hanna S.R., 1989. Confidence limits for air quality model evaluations, as estimated by bootstrap and jackknife resampling methods. *Atmospheric Environment*, 23, 1385-1398.

Hanna S.R., 1993. Uncertainties in air quality model predictions. *Boundary Layer Meteorology*, 62, 3-20.

Hanna S.R., J.C. Chang, 1993. Hybrid plume dispersion model (HPDM) improvements and testing at three field sites. *Atmospheric Environment*, 27A, 1591-1508.

Hurley P., Manins P., Lee S., Boyle R., Ng L.Y., Dewundege P., 2003. Year-long, high-resolution, urban airshed modelling: verification of TAPM predictions of smog and particles in Melbourne, Australia. *Atmospheric Environment*, 37, 1899-1910.

IAEA, 1980: Atmospheric Dispersion in Nuclear Power Plant Siting. International Atomic Energy Agency, Safety Series, 50-SG-S3.

Jackson, B., Chau D., Gurer K., Kaduwela A., 2006. Comparison of ozone simulations using MM5 and CALMET/MM5 hybrid meteorological fields for the July/August 2000 CCOS episode. *Atmospheric Environment*, 40, 2812-2822.

Kovalets I., Andronopoulos S., Bartzis J.G., Gounaris N., Kushchan A., 2004. Introduction of data assimilation procedures in the meteorological pre-processor of atmospheric dispersion models used in emergency response systems. *Atmospheric Environment*, 38, 457-467.

Kovalets I., Andronopoulos S., Bartzis J.G., 2005. About the need to develop the data assimilation procedures for the meteorological pre-processors of the emergency response systems. Short communication. *Atmospheric Environment*, 39, 3369-3372.

Maryon R.H., 1997: Determining cross-wind variance for low-frequency wind meander. *Atmospheric Environment*, 32, 115-121.

Mosca S., Graziani G., Klug W., Belasio R., Bianconi R., 1998. A statistical methodology for the evaluation of long-range dispersion models: An application to the ETEX exercise. *Atmospheric Environment*, 32, 4307-4324.

Luhar A.K., Hurley P., 2004. Application of a prognostic model TAPM to sea-breeze flows, surface concentrations, and fumigating plumes. *Environmental Modelling & Software*, 19, 591-601.

Raskob W., Ehrhardt J. (Eds.), 1999. The RODOS System: Decision Support for Nuclear Off-Site Emergency Management in Europe. RODOS report GEN TN99 02, <http://www.rodos.fzk.de>.

Rodean C.H., 1991. The universal constant for the Lagrangian structure function, *Physics of Fluids*, A 3 (6), 1479-1480.

Rodean C.H., 1994. Notes of the Langevin model for turbulent diffusion of "Marked" Particles. Technical Report UCRL-ID-115869, Lawrence Livermore National Laboratory.

Ryall D.B., Maryon R.H., 1998. Validation of the UK Met. Office's NAME model against the ETEX dataset. *Atmospheric Environment*, 32, 4265-4276.

Seaman N., 2000. Meteorological modelling for air-quality assessments. *Atmospheric Environment*, 34, 2231-2259.

Seibert P., Beyrich F., Gryning S.-E., Joffre S., Rasmussen A., Tercier Ph., 1997. Mixing Height Determination for Dispersion Modelling. Report of Working Group 2, in COST Action 710 – Final Report, EUR 18195 EN, European Commission.

Shuming D., 1997. Universality of the Lagrangian velocity structure function constant (C_0) across different kinds of turbulence. *Boundary Layer Meteorology*, 83, 207-219.

Straume A. G., Nodop K., 1997. Meteorological observations, collected during the ETEX. Environment Institute, JRC, Italy, <http://rem.jrc.cec.eu.int/etex/>

Thomson D.J., 1987. Criteria for the selection of stochastic models of particle trajectories in turbulent flows. *Journal of Fluid Mechanics*, 180, 529-556.

Zannetti P., 1990. Air Pollution Modeling. Theories, Computational Methods and Available Software. Computational Mechanics Publications, pp. 444.

Figures Captions

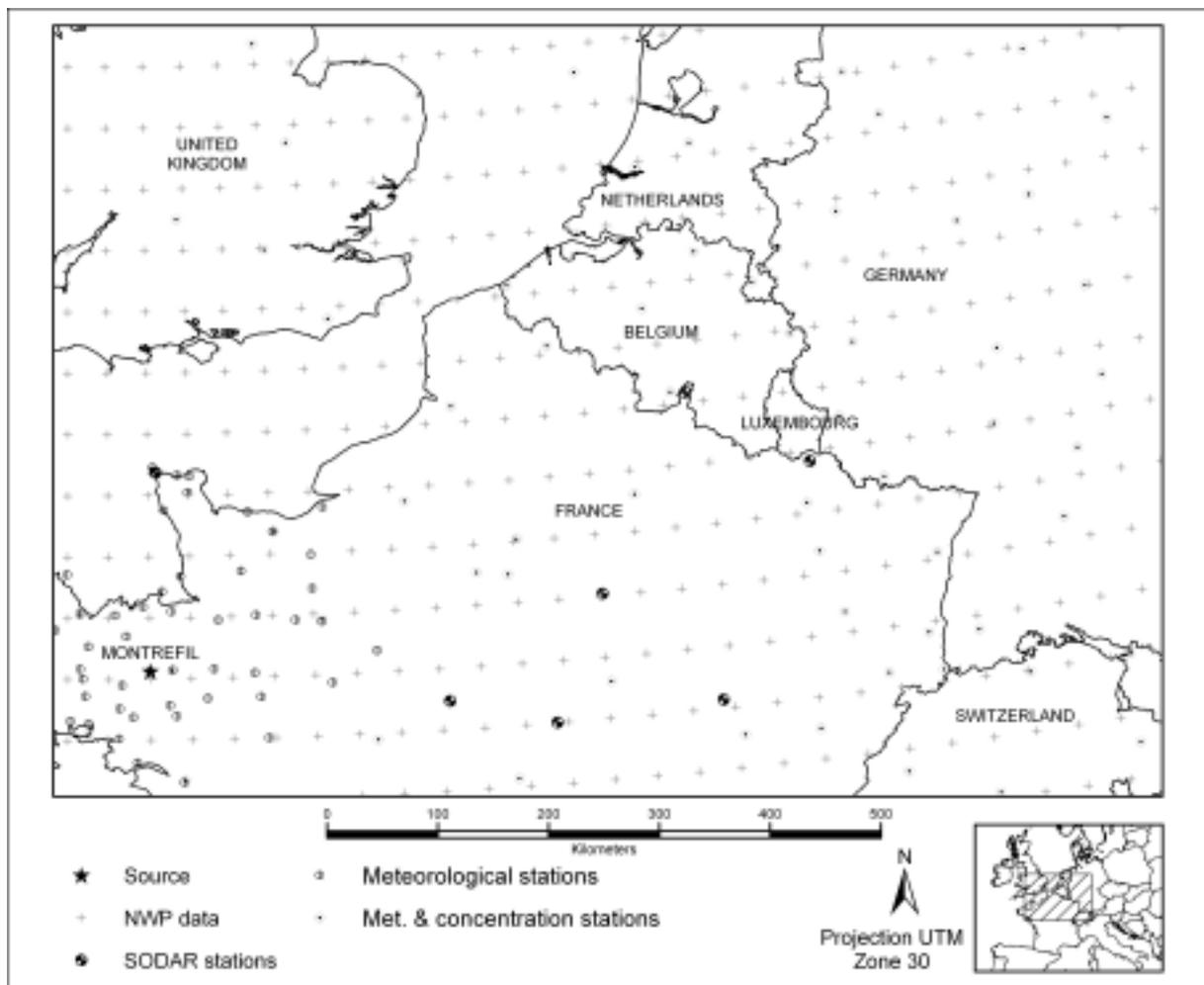


Figure 1: Computational domain with the tracer release location, the observation points' locations from the ETEX database and the grid of the Numerical Weather Prediction model

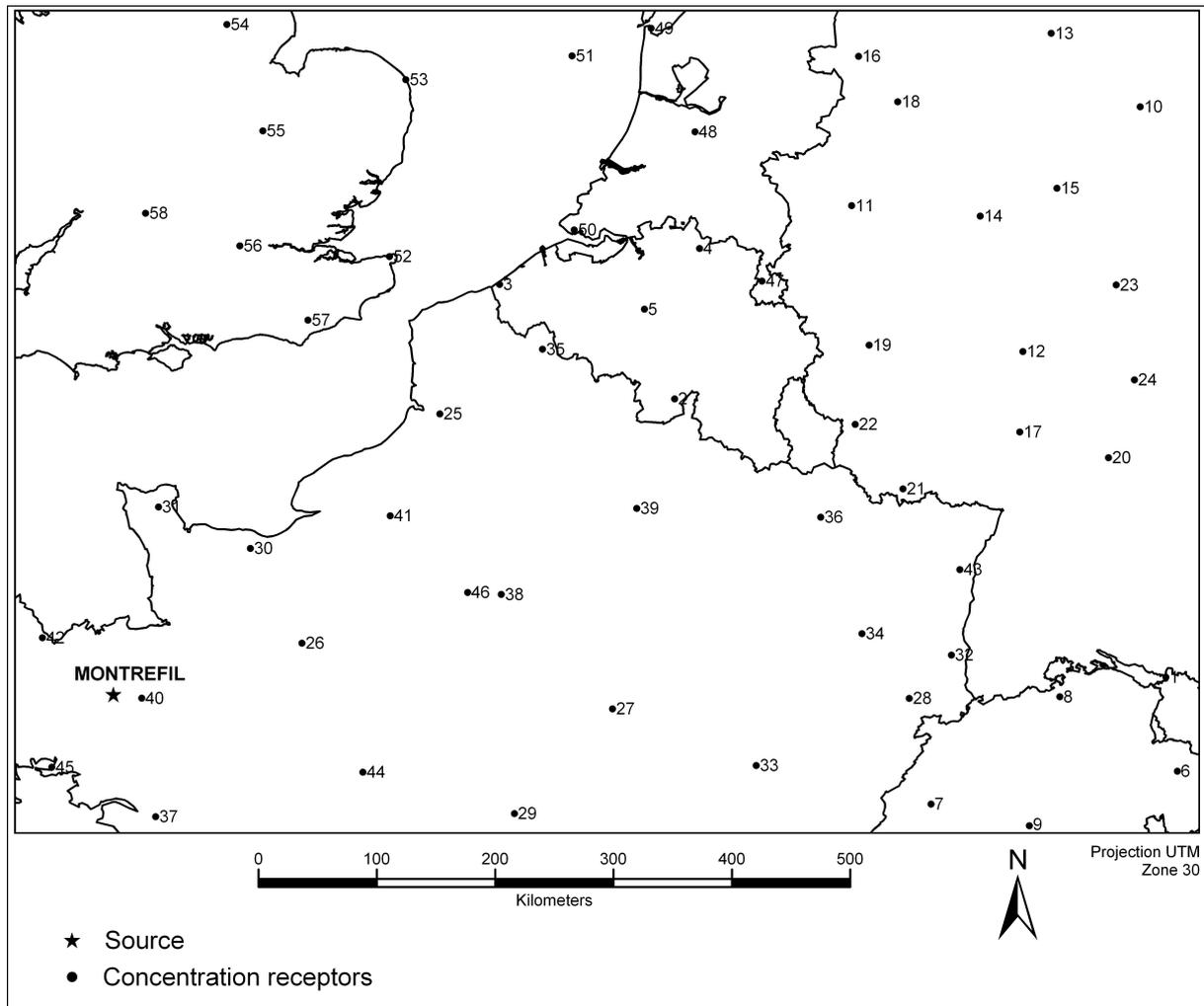


Figure 2: Concentration receptors in the computational domain

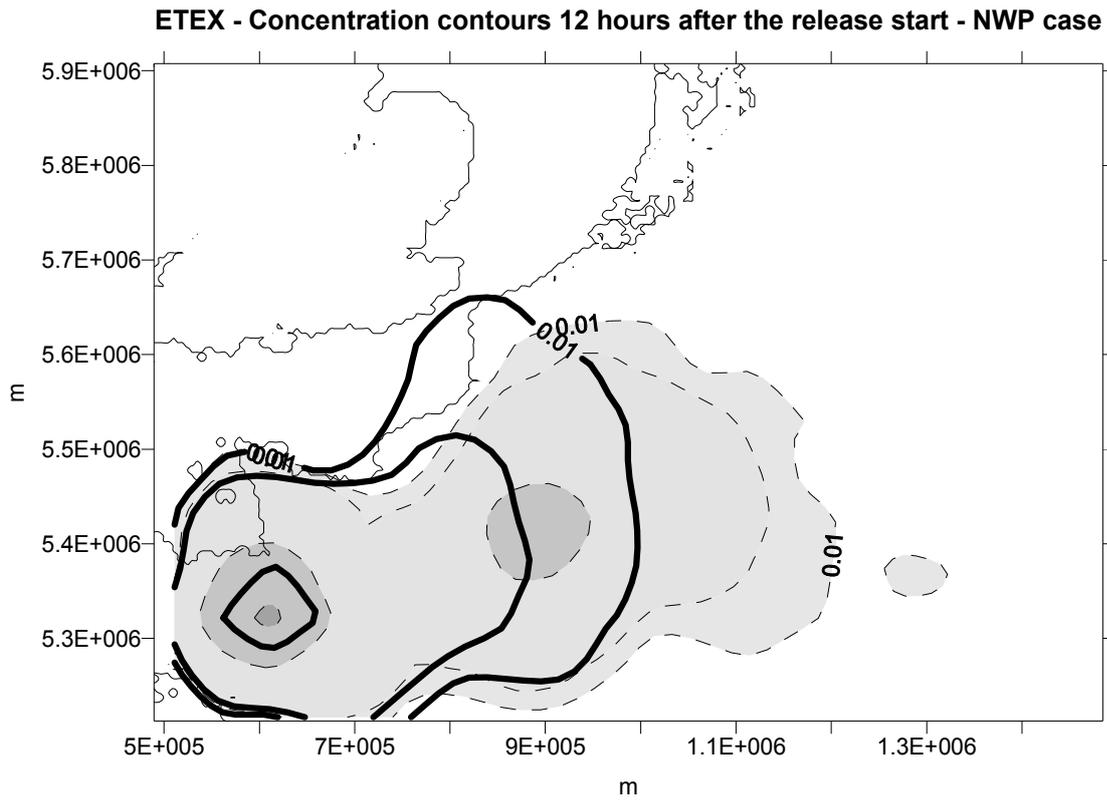


Figure 3a: Contour plots of experimental (dashed line) and predicted (solid line) concentrations 12 hours after the release start for NWP meteorological data set; contour values: 0.01, 0.1, 1, 2, 3 ngr/m³; computational domain as in Figures 1-2; axes coordinates: UTM Zone 30

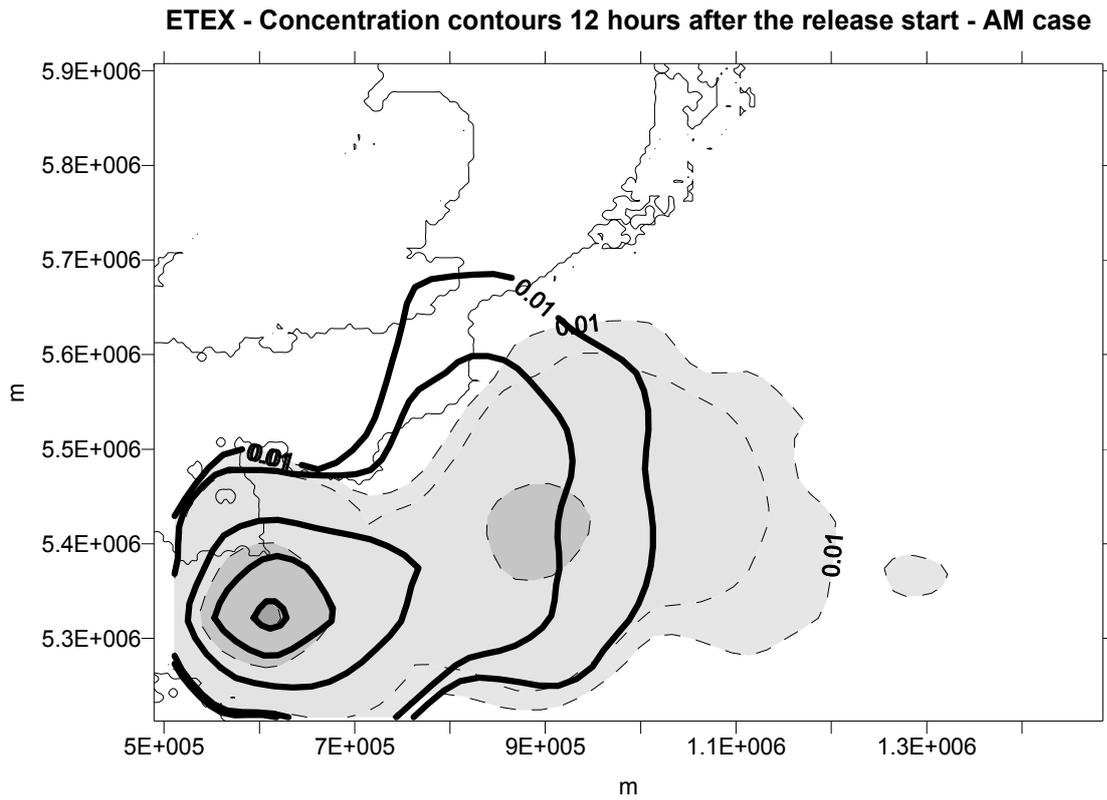


Figure 3b: Contour plots of experimental (dashed line) and predicted (solid line) concentrations 12 hours after the release start for AM meteorological data set; contour values: 0.01, 0.1, 1, 2, 3 ngr/m³; computational domain as in Figures 1-2; axes coordinates: UTM Zone 30

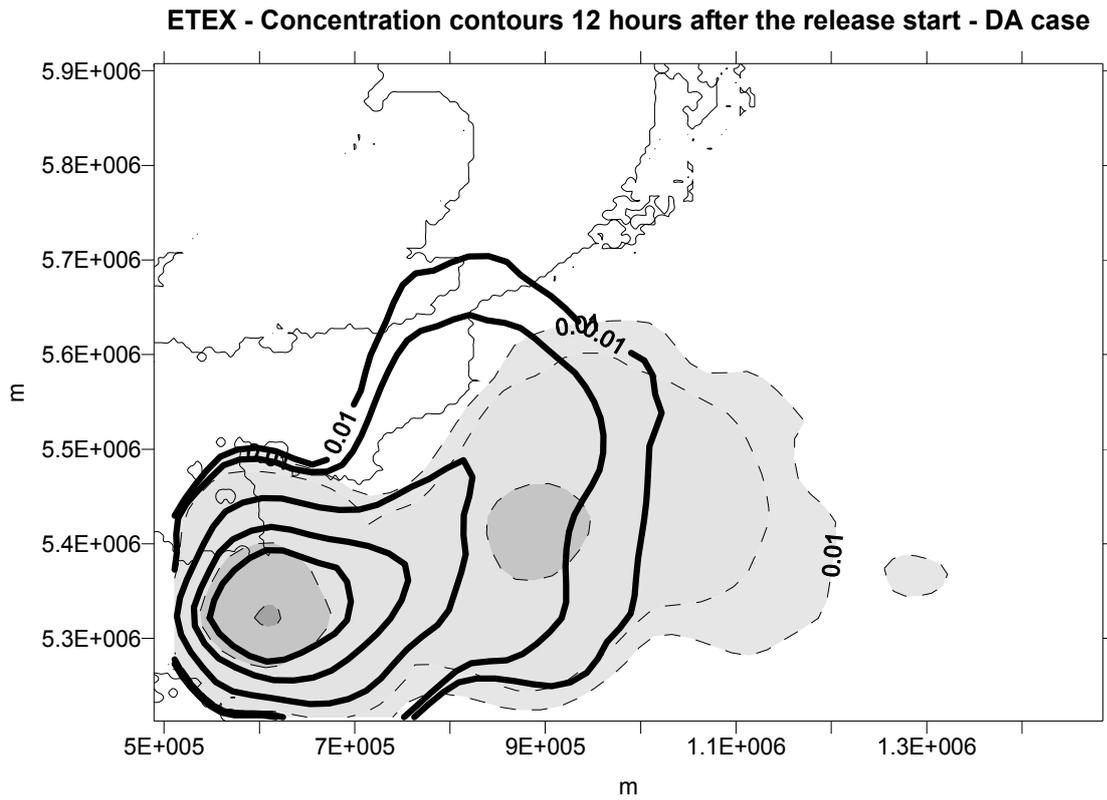


Figure 3c: Contour plots of experimental (dashed line) and predicted (solid line) concentrations 12 hours after the release start for DA meteorological data set; contour values: 0.01, 0.1, 1, 2, 3 ngr/m³; computational domain as in Figures 1-2; axes coordinates: UTM Zone 30

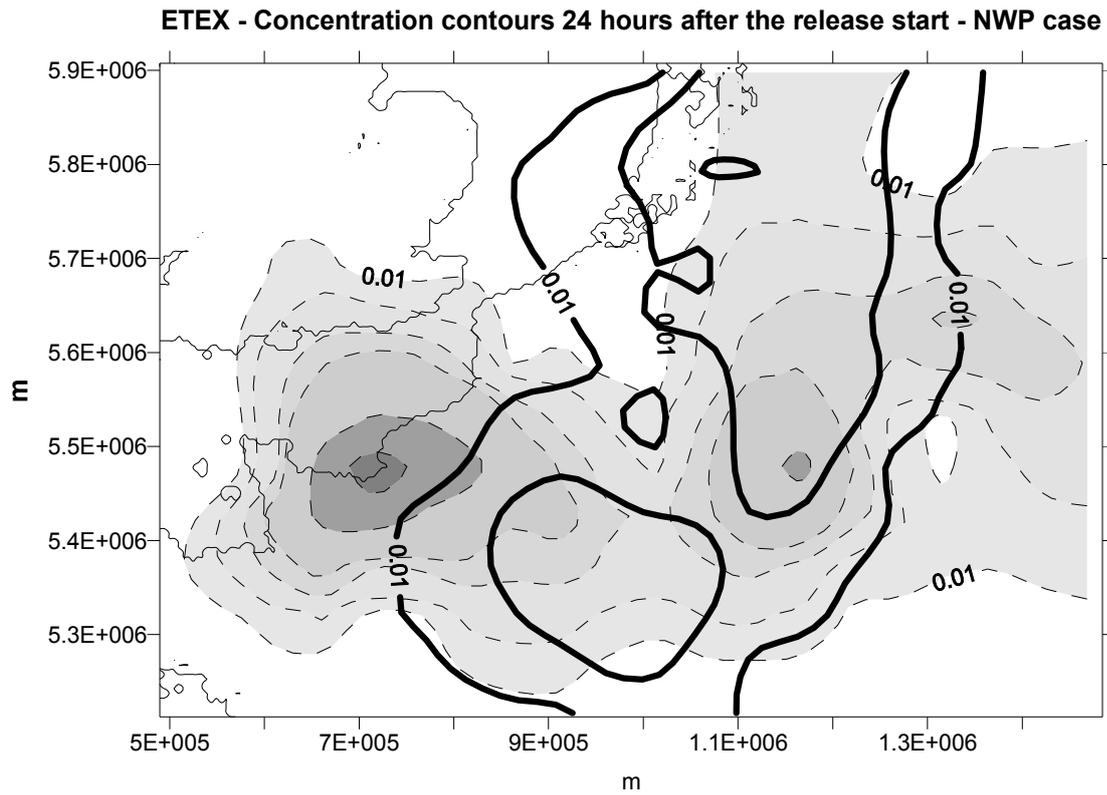


Figure 4a: Contour plots of experimental (dashed line) and predicted (solid line) concentrations 24 hours after the release start for NWP meteorological data set; contour values: 0.01, 0.1, 0.3, 0.5, 1.5, 2 ngr/m³; computational domain as in Figures 1-2; axes coordinates: UTM Zone 30

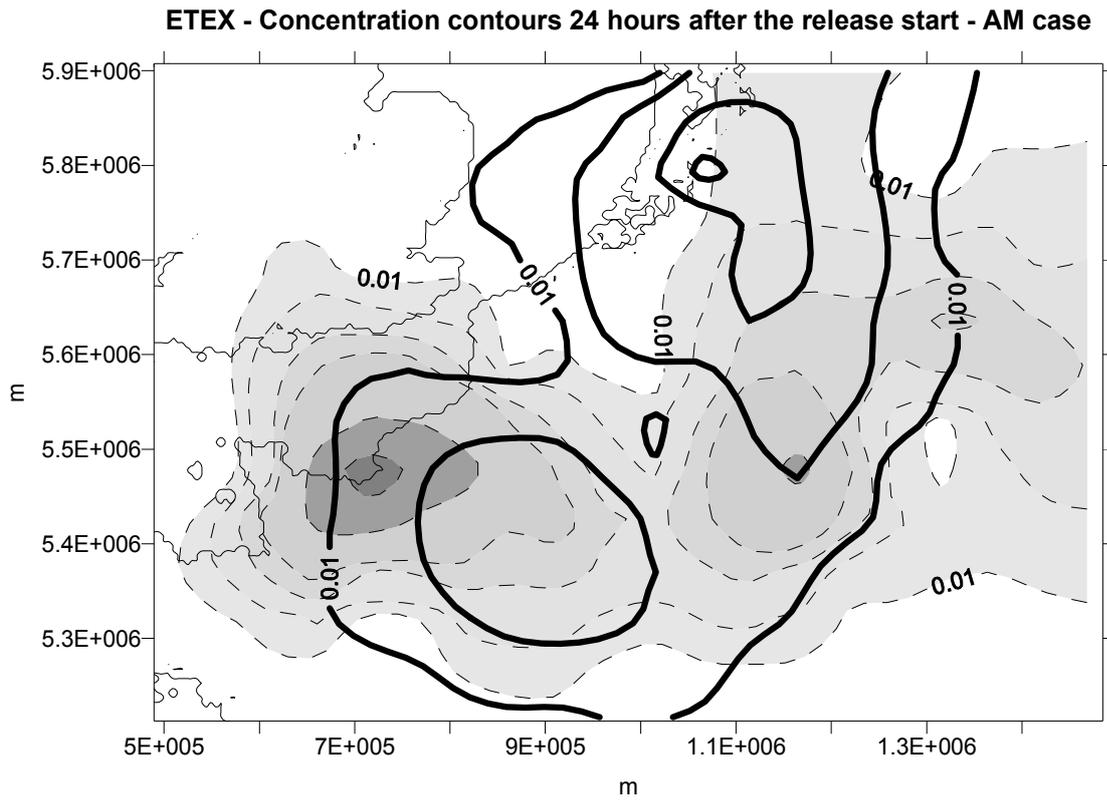


Figure 4b: Contour plots of experimental (dashed line) and predicted (solid line) concentrations 24 hours after the release start for AM meteorological data set; contour values: 0.01, 0.1, 0.3, 0.5, 1.5, 2 ngr/m³; computational domain as in Figures 1-2; axes coordinates: UTM Zone 30

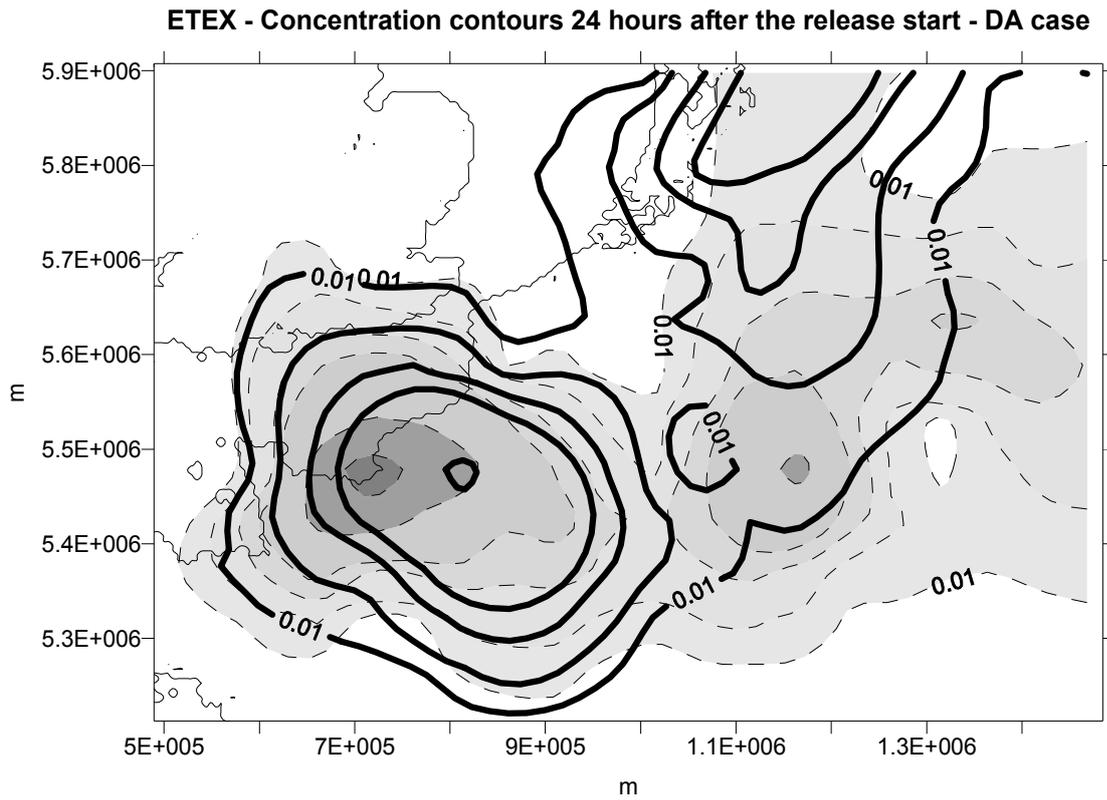


Figure 4c: Contour plots of experimental (dashed line) and predicted (solid line) concentrations 24 hours after the release start for DA meteorological data set; contour values: 0.01, 0.1, 0.3, 0.5, 1.5, 2 ngr/m³; computational domain as in Figures 1-2; axes coordinates: UTM Zone 30

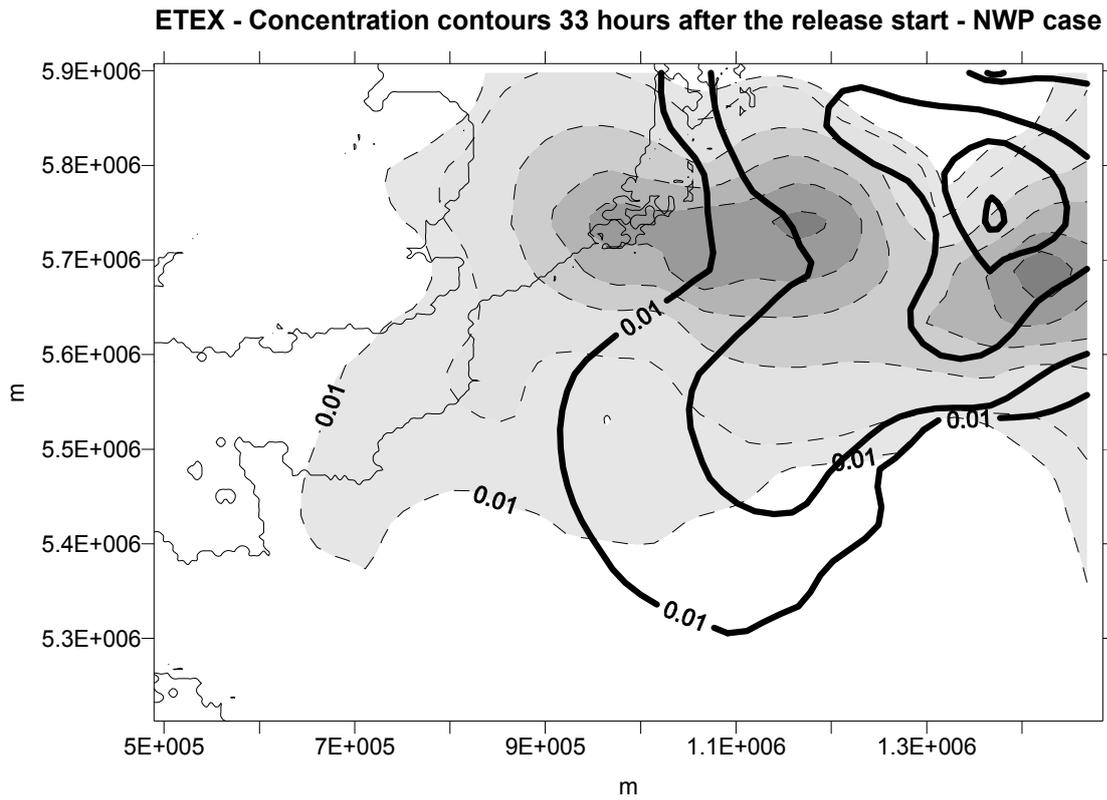


Figure 5a: Contour plots of experimental (dashed line) and predicted (solid line) concentrations 33 hours after the release start for NWP meteorological data set; contour values: 0.01, 0.1, 0.5, 1, 1.5, 2 ngr/m³; computational domain as in Figures 1-2; axes coordinates: UTM Zone 30

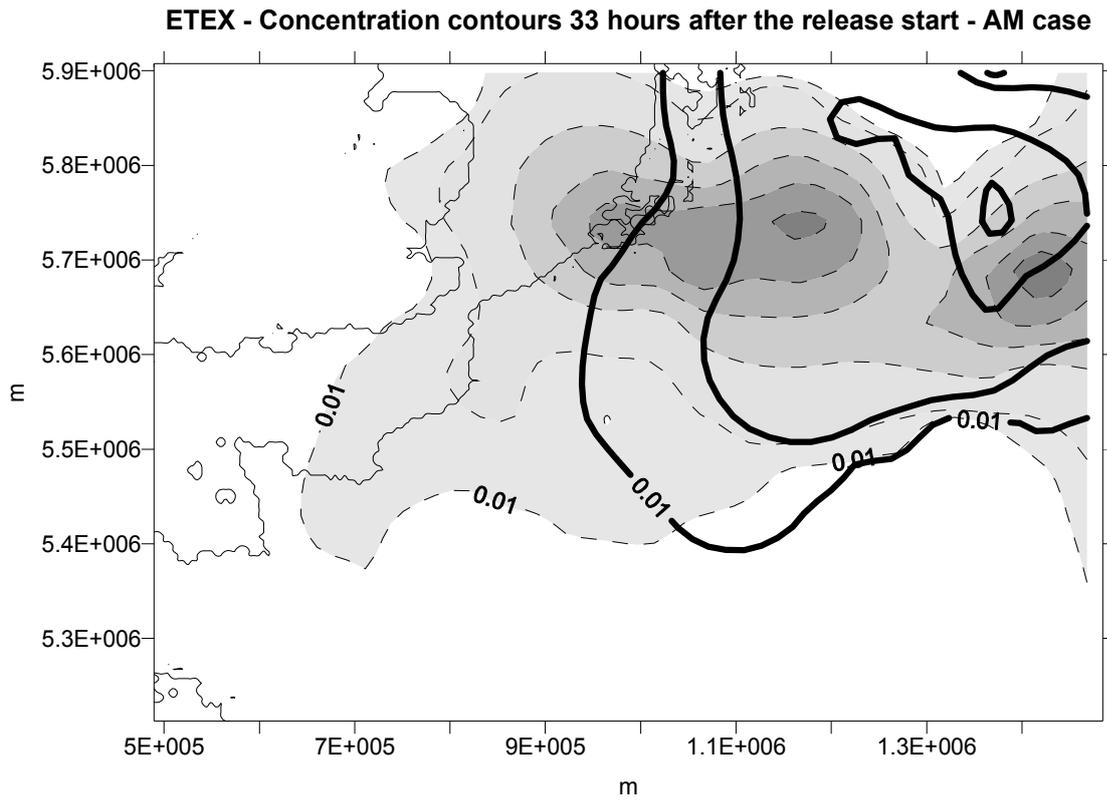


Figure 5b: Contour plots of experimental (dashed line) and predicted (solid line) concentrations 33 hours after the release start for AM meteorological data set; contour values: 0.01, 0.1, 0.5, 1, 1.5, 2 ngr/m³; computational domain as in Figures 1-2; axes coordinates: UTM Zone 30

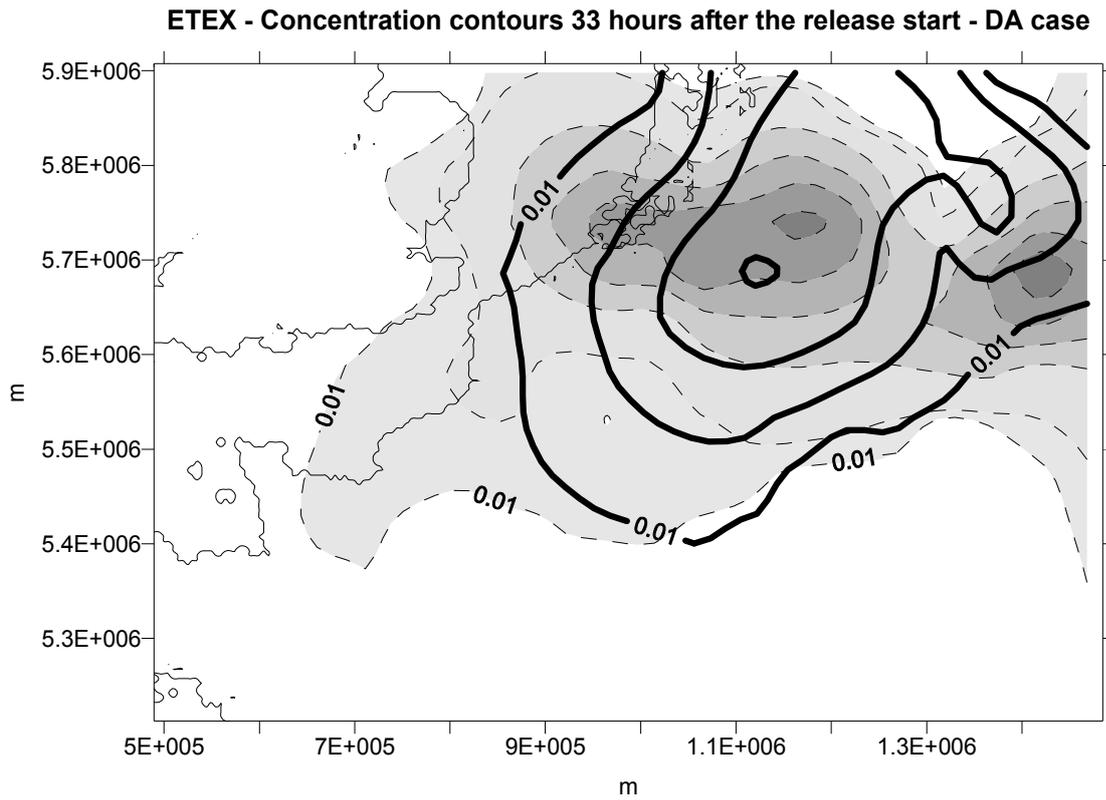


Figure 5c: Contour plots of experimental (dashed line) and predicted (solid line) concentrations 33 hours after the release start for DA meteorological data set; contour values: 0.01, 0.1, 0.5, 1, 1.5, 2 ngr/m³; computational domain as in Figures 1-2; axes coordinates: UTM Zone 30

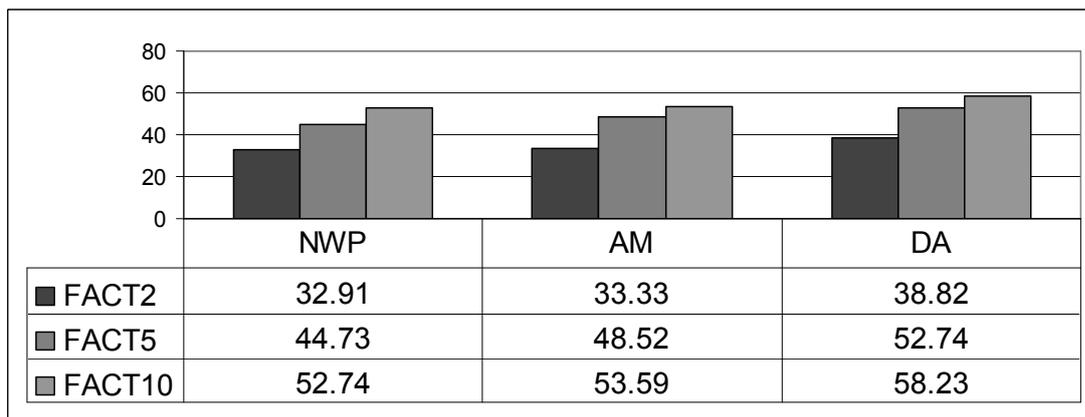


Figure 6: FACT2, FACT5 and FACT10 histogram for ADM predictions calculated with NWP, AM, and DA meteorological data sets

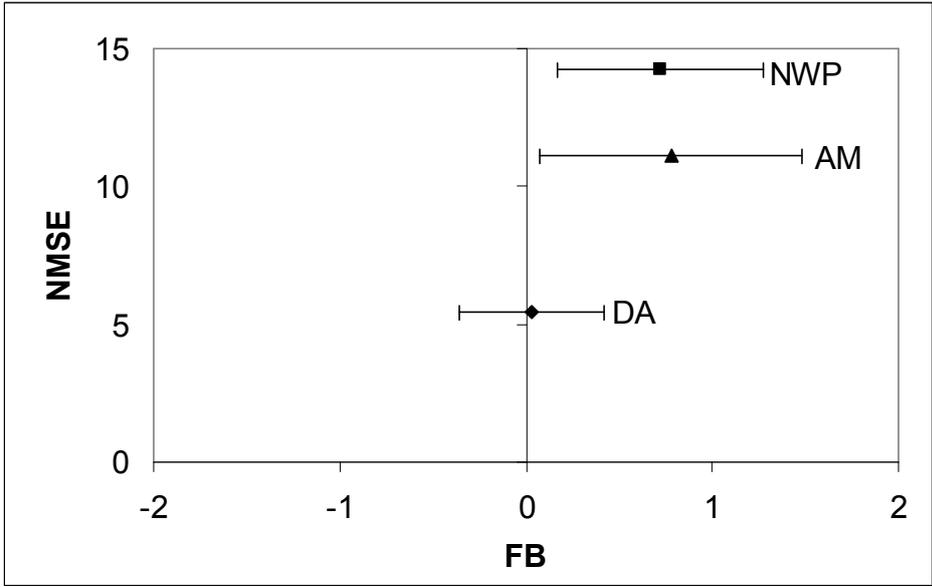


Figure 7: Normalised Mean Square Error plotted versus Fractional Bias for ADM predictions calculated with NWP, AM, and DA meteorological data sets

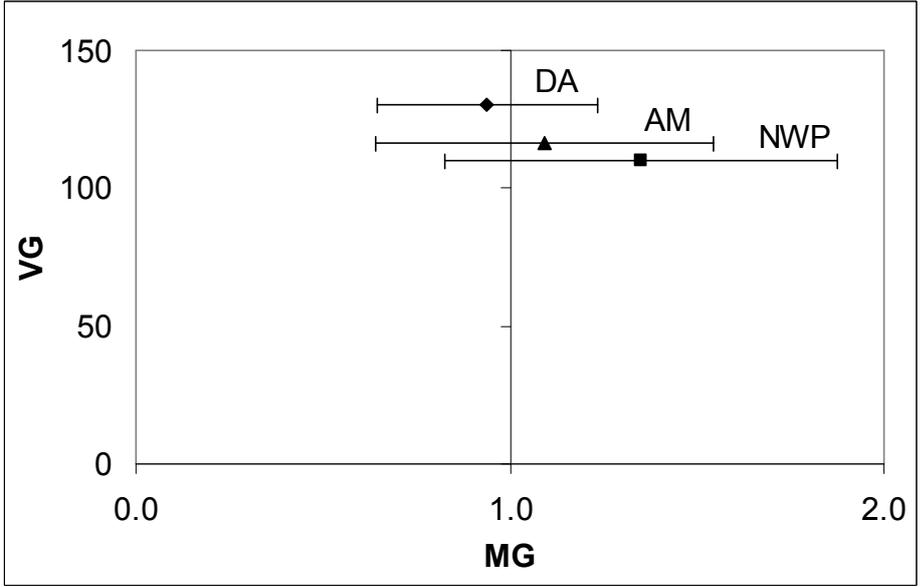


Figure 8: Geometric mean Variance plotted versus Geometric Mean bias for ADM predictions calculated with NWP, AM, and DA meteorological data sets

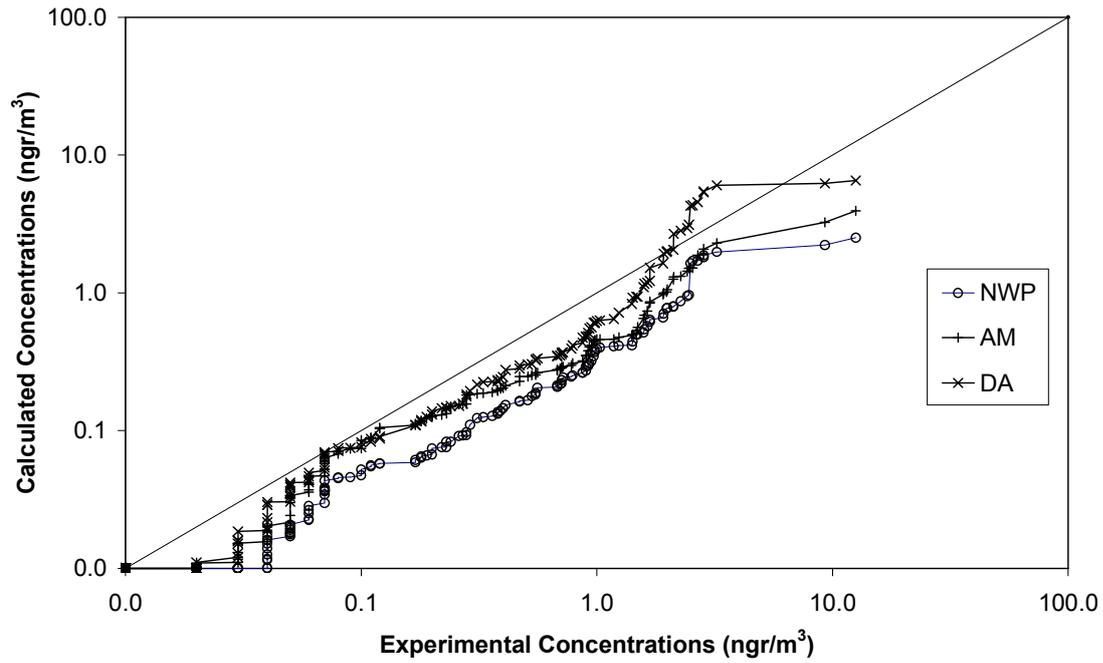
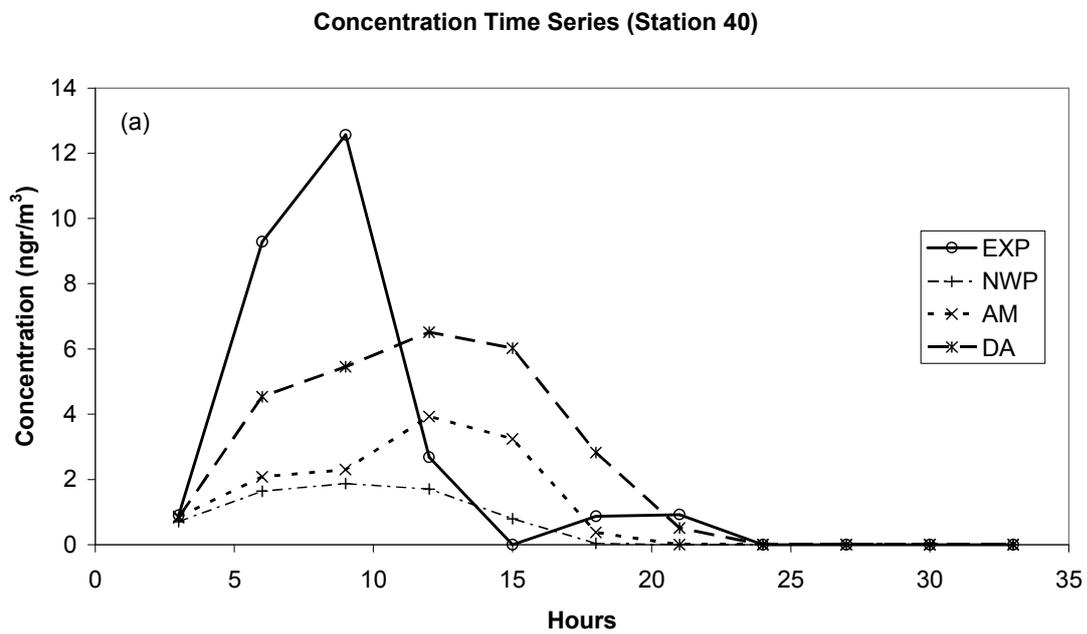
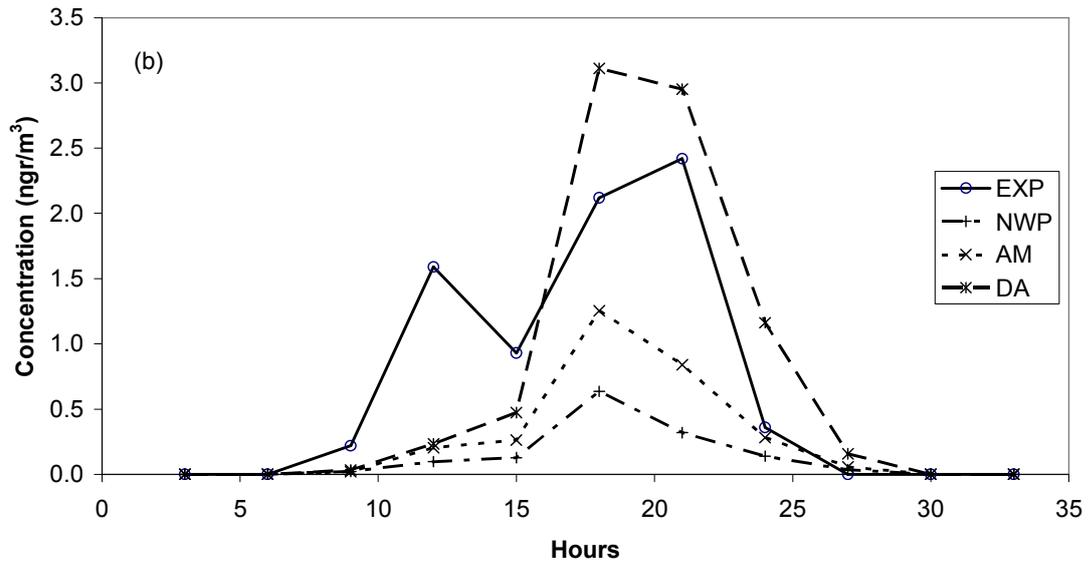


Figure 9: Quantile–Quantile plot of the predicted vs. observed concentrations for NWP, AM, and DA meteorological data sets.



Concentration Time Series (Station 46)



Concentration Time Series (Station 5)

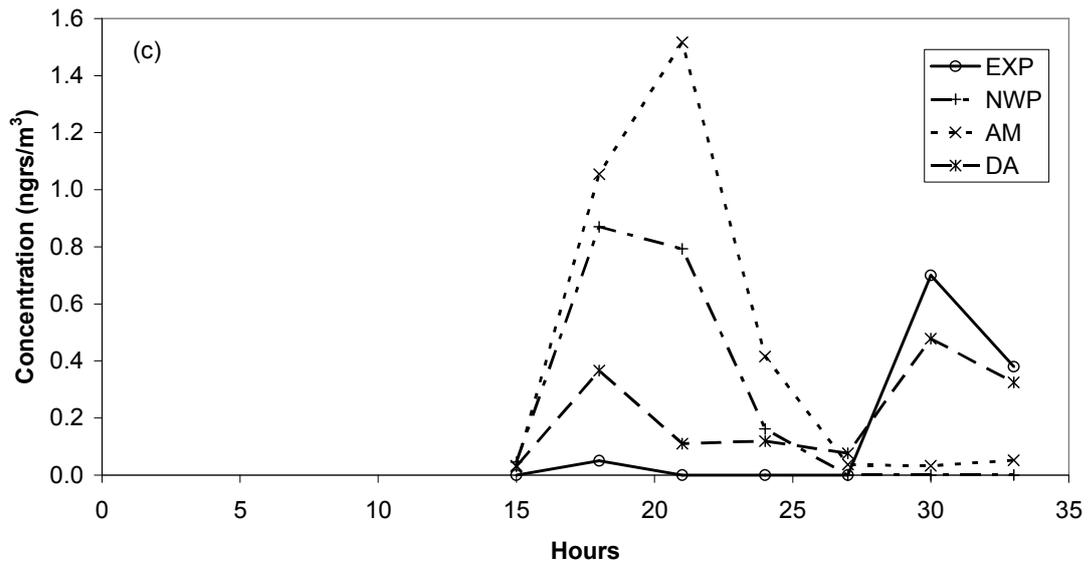


Figure 10: Concentration time histories at 6 observation points (beginning)

Concentration Time Series (Station 15)

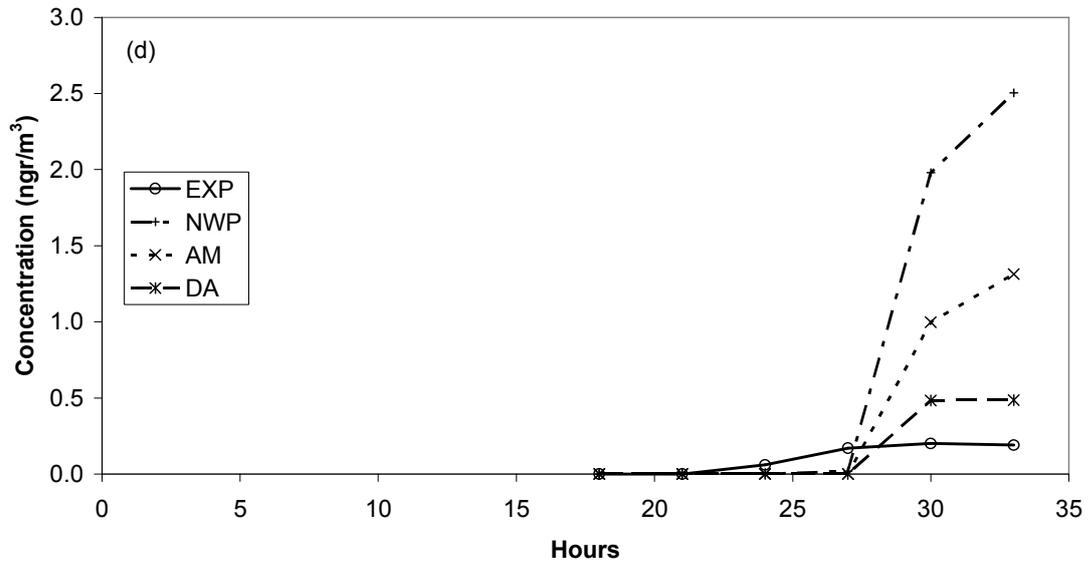


Figure 10: Concentration time histories at 6 observation points