Met O (PMSR) Turbulence and Diffusion Note No. 260

Investigating the Use of Ensemble Forecasts in Atmospheric Dispersion Modelling

by

V. Lucas

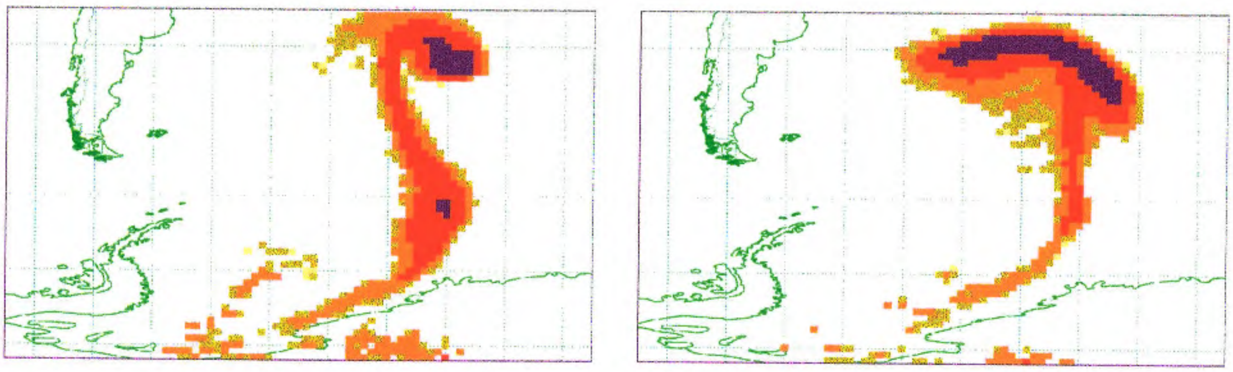
Met O (PMSR)
(Public Met Service Research)
Meteorological Office
London Road
Bracknell
Berks, RG12 2SZ

Report on work supported by the Met Office Core Research Programme

Note

This paper forms part of the requirements for the Master of Science Degree at Surrey University. Permission to quote from it should be obtained from the Branch Head, Public Met Service Research Division, Met O (PMSR), Meteorological Office, London Road, Bracknell, Berkshire, RG12 2SZ.

Investigating the use of Ensemble Forecasts in Atmospheric Dispersion Modelling



Air concentration of an airborne pollutant, developed by two members of an ensemble forecast. The simulated release began at the Falklands and is shown four days later.

Vicky Lucas

Meteorological Office TDN 260
MSc Dissertation Project
Radiation and Environmental Protection
Surrey University
September 1999

1 Abstract

Ensemble NWP forecasts have been applied to atmospheric dispersion modelling. A global ensemble forecast of 27 members was used as the meteorological input for NAME, the atmospheric dispersion model of the UK Met Office. Releases of inert tracer were simulated by NAME and the resulting distribution of tracer concentrations were followed over 132 hours. Techniques have been developed to investigate the ensemble dispersion data, including validation against the equivalent NAME solutions that utilised the operational forecast and the analysed meteorological fields. It has been demonstrated that ensemble techniques can produce probability dispersion forecasts.

This is a brief preliminary study and whilst ensemble information has been shown to add useful probabilistic information, more experimentation and objective evaluation is required.

2 Acknowledgements

Roy Maryon and Derrick Ryall of the Atmospheric Dispersion group, UKMO, provided valued guidance during the study. Ken Mylne of the Predictability and Ensemble Forecasting group, UKMO, proposed the original idea for this project and has continued to be a great source of encouragement and knowledge of ensemble prediction. Richard Barnes of ECMWF was instrumental in exporting the UKMO data from ECMWF and Alison Malcolm of the UKMO showed great patience in helping to convert the data to a format compatible for use in NAME. Thanks are also due to Ed Morton of Surrey University for his assistance in producing this work.

Contents

Investigating the use of Ensemble Forecasts in Atmospheric Dispersion Modelling

3 Introduction & Background

The development of ensemble forecasting is a new area of research at the UK Meteorological Office; in parallel to the introduction and development of ensemble weather forecasting products it has been proposed that atmospheric dispersion modelling may also benefit from the ensemble technique. This report is a preliminary study designed to establish some tools for the analysis of ensemble atmospheric dispersion and to assess the effectiveness and value of ensemble methods. The scope of the study is vast as it combines two entirely independent areas of meteorology. Initial investigations into a novel discipline in atmospheric dispersion are presented in this study.

3.1 Ensemble Forecasting

An ensemble forecast is a collection of two or more forecasts that are valid for the same time frame. These forecasts start from different initial conditions and/or are based on different forecasting procedures. The various ensemble forecasts all represent possible physical solutions, given the uncertainties associated with formulating the forecast.

Most weather forecasts are derived from a single, deterministic NWP solution. A deterministic solution is one in which there is only one outcome (Lorenz 1996, 7), so deterministic weather forecasts can only suggest one sequence of events. Ensemble forecasts use large numbers of these deterministic forecasts in order to produce a range of solutions, which is a probability distribution of outcomes (Molteni 1996).

The purpose of ensemble forecasting is to recognise the inherent uncertainty of weather forecasting and to indicate which meteorological scenarios may be possible. Its goals are to increase average forecast accuracy, to estimate the likelihood of various events and to estimate the decay of forecast skill with increasing lead time (Sivillo et al. 1997).

The Development of Ensemble Forecasting

During the 1960s Edward Lorenz investigated fundamental aspects of atmospheric predictability. He found that errors develop nonlinearly in numerical weather predictions and he went on to suggest that the atmosphere might have a finite prediction time, of the order of 10 days (Sivillo et al. 1997). Lorenz proposed that errors of different spatial scales grow at different rates, with on average the fastest error growth occurring at small scales (Lorenz 1993, 182-183). This means that whilst mesoscale detail (for example a thunderstorm) may be lost within hours of the start of a forecast, synoptic and larger scale features (such as positions of cyclones and anticyclones) are predictable for several days. This slower growth of large features makes weather forecasting possible (Lorenz 1993, 106).

In 1964 Jule Charney presented a paper on the possibility that the atmosphere behaves chaotically and he persuaded modellers to perform numerical experiments in which pairs of forecasts originating from slightly different conditions would be examined for sensitive dependence (Lorenz 1993, 103-104). Then, in 1974 Leith demonstrated that an ensemble of roughly ten forecasts seemed to be large enough to make real improvements in six to ten day forecasts, a method he labelled 'Monte Carlo Forecasting' (Sivillo et al. 1997). Today, with increased computing power, ensemble forecasts of many members can be run operationally out to 10 days.

Ensemble Forecasting Philosophy

Sources of error that limit NWP performance are (Molteni et al. 1996):

1. Incomplete observations of initial conditions leading to errors in the specification of the initial model state. According to Lorenz's research, these errors amplify, no matter how small, and so impose a limit on how far into the future a skilful weather forecast is possible.
2. Imperfections in the NWP model formulation, including incomplete mathematical representation of the dynamical and physical equations that govern the atmosphere and further approximations associated with converting differential equations into forms that can be solved by the computer within the available time.

These errors mean that perfect solutions are not possible. In ensemble forecasting, a finite number of different initial conditions are used and a nonlinear model integration is then carried out from each of these states. The properties of the spread of solutions to the forecast are assumed to be described by the statistics computed from the ensemble (Molteni et al. 1996).

Current operational ensemble forecasting at both NCEP (National Center for Environmental Prediction, USA) and ECMWF (the European Centre for Medium-Range Weather Forecasts) is focused on the consequences of initial value errors. However, some work is being undertaken to form multi-model analyses in order to overcome the differences in different NWP model calculations and parameterisations (Evans et al. 1998).

The initial perturbations used in creating the array of different ensemble members will always be small compared with the infinite number of possible atmospheric states, so the initial perturbations are chosen as those that are likely to cause the largest forecast errors. ECMWF computes those errors that, if the dynamics were linear, would grow most rapidly during the first 48 hours of the forecast (Sivillo et al. 1997). It can be shown that for the first day or two forecast errors at synoptic and larger scales are often governed by linear processes, the size of the forecast error is directly proportional to the size of the initial error. The ECMWF method also assumes that all ensemble solutions are equally likely (Sivillo et al. 1997).

The ensemble of 50 forecasts is comprised of 25 pairs of initial conditions which are created by adding and subtracting these perturbations from the global analysis. These changes imposed on the original state are called 'singular vectors' or 'optimal perturbations' (Ehrendorfer & Tribbia 1997). Once the 50 members with different conditions have been made, fully nonlinear NWP models are used to compute the forecasts.

Perturbations

The behaviour of the NWP model atmosphere over a very long time period reveals a certain restricted set of states which are allowed and those which are not possible. The allowed states are known as the model attractors or model climate (Anderson 1996, Lorenz 1993 39-55). Ideally, the model climate will exactly match the real global climate, but in reality this condition is not met. If the NWP model was perfect, any forecast produced by the model would lie on the model attractor dictated by the model climate and the initial conditions.

Ensembles use the same model climate for each forecast but the model atmosphere is able to evolve into different states due to differences imposed on the initial conditions. Using skilful perturbations is crucial to the success and usefulness of an ensemble forecast (Ehrendorfer & Tribbia 1996). The ensemble solutions should suggest the range of possible states that the atmosphere might attain, but these solutions should not be so widely spread so that they lie in any part of the model climate. Therefore the skill of choosing initial perturbations is to faithfully reflect the variability of the atmosphere and so cover the full range of possible solutions, whilst restricting the perturbations and not making them so large that arbitrary solutions to the model atmosphere are created.

The Development of Operational Ensemble Forecasting

In December 1992 NCEP began computing 10-day ensemble forecasts on an operational basis (Sivillo et al. 1997). Later that month ECMWF began ensemble forecasts, which in turn became operational in May 1994 with 32 ensemble members (Molteni et al. 1996). Advances in computer power allowed the ECMWF system to expand to 50 members in December 1996. Ensemble forecasting has now begun in Japan and South Africa and an ensemble prediction forecast evaluation using different global NWP models is employed operationally in Australia (Leslie et al. 1994).

Ensemble Forecasts of the UKMO Unified Model at ECMWF

The operational NWP model of the UK Meteorological Office is implemented as a single, deterministic solution, producing one 144 hour forecast every 12 hours. However, research is being conducted into running an ensemble forecast of the UKMO Unified Model (UM) (Evans et al. 1998). Due to the operational commitments at the Met Office and the expense of computer time, the ensembles are run at ECMWF.

To create an ensemble the UKMO Unified Model (UM) version 4.4 is run on the Fujitsu at ECMWF, with a resolution of 1.25^{o} longitude and 0.83^{o} latitude and 30 vertical levels. Each of the singular vectors is calculated using the ECMWF scheme (described in Molteni et al. 1996) and each is added and subtracted to the UKMO operational analysis field in order to form the initial conditions for the UM ensemble runs (Evans et al. 1998). 27 members of the UKMO model are run, comprising of the control plus the first 13 pairs of perturbations. The horizontal resolution is approximately half that used by the UKMO for the operational forecasts, which are run on the UKMO's Cray T3E.

The Application of Ensembles to Probability Forecasting

Since the atmosphere is a chaotic system deterministic prediction is not always possible and it is desirable to seek probabilistic solutions. Ensemble predictions provide a practical approach to creating probability estimates which still recognise the inherent nonlinearlity of the atmosphere and are not simply a probabilistic interpretation of a deterministic solution. The variation between forecasts in an ensemble suggest a range of possibilities whose probabilities can be estimated (Anderson 1996).

3.2 NAME – The Atmospheric Dispersion Model of the Met Office

NAME was developed by the UK Meteorological Office (UKMO) in order to model the long-range transport of pollutants. The desire for such a capability came in response to the Chernobyl incident of 1986. The original purpose of NAME (Nuclear Accident Model) was to provide guidance to government and other agencies in the event of a nuclear accident either originating from or affecting the UK. NAME is capable of providing this guidance by simulating the transport of pollutants in the atmosphere and their deposition to the ground.

Current applications of NAME still include the emergency response capability; NAME fulfils the UK Met Office's international obligations for providing model forecasts in the event of major atmospheric releases, including radionuclide emergencies and volcanic ash events. In addition to the operational requirements of NAME, it is also the subject of continual research and development. NAME is used as a research tool for investigating long-range transport of aerosols and gases, and it is also being applied to air quality modelling.

NAME is used to simulate the transport of airborne pollutants at distances from the source of the order of tens of kilometres to thousands of kilometres. The model provides estimates of air concentrations and dosages, and of deposition of pollutants to the ground by both wet and dry deposition processes. The meteorological data on which NAME is based are provided by Version 4.4 of the global UKMO Numerical Weather Prediction Model, known as the Unified Model (UM). The UM calculates meteorological data on a grid covering the globe, which can be read by NAME.

Modelling Dispersion

The NAME model is of a Lagrangian, Monte Carlo type in which emissions are modelled by releasing large numbers of particles into the model atmosphere. The particles are carried along passively by the ambient three-dimensional wind flow, with turbulent dispersion simulated by random walk techniques. Wind direction is the single most important predictor of the spread of airborne material. In NAME the particles are advected each model time step (e.g. 15 minutes) using the mean wind, plus a turbulent velocity component. The mean wind is taken from the UM.

Each particle represents a mass or an activity of a pollutant, and the number of particles in each model grid box is used to calculate the concentration of the pollutant at that point. The mass of a particle is reduced over time by wet and dry deposition processes and radioactive decay. The model also has facilities for modifying the plume spread and estimating source strengths from observational data.

The Boundary Layer and Mechanisms of Dispersion

The boundary layer is the lowest layer of the atmosphere, in which the flow properties are affected by the underlying surface. The surface causes frictional effects and exchanges of momentum, as well as exchanges of heat and moisture with the air above. The dispersion of aerosols and gases released into the atmosphere near the ground depends on effective mixing in the boundary layer (Pasquill & Smith 1983). Turbulent and convective motions produce mixing in the boundary layer; turbulent mixing is created by mechanical overturning of the air as it passes over surfaces and convective mixing is created by heat fluxes produced by warm surfaces. The height of the boundary layer evolves continuously in response to spatial and temporal changes in surface conditions, varying from tens of metres at night up to 2000 metres or more during summer afternoons. Above the boundary layer the airflow is relatively smooth.

The correct determination of the boundary layer depth is crucial for modelling the dispersion of airborne pollutants and model results are very sensitive to it (Maryon & Best 1992). Turbulence is usually greater in the boundary layer than in the free troposphere, resulting in more rapid horizontal and vertical diffusion, but usually slower transports. In NAME, boundary layer depths are directly calculated from Unified Model wind and temperature profiles. Material is extracted from the atmosphere by wet and dry deposition and material can be entrained into the free atmosphere above either by variations in the boundary layer height or by an entrainment parameterisation.

For releases which are tracked over hundreds of kilometres and on time scales of more than a day, the direction and dispersion of the plume is dominated by the changing large-scale wind patterns.

Defining Releases

Each particle can represent one or more species and the species that can be modelled include CFCs, sulphur dioxide, radioactive gases and aerosols. The release can be defined as a function of height and time, which allows for continuous releases or variable rates of release.

Meteorological Data

NAME utilises meteorological data from the UM, the operational numerical weather prediction (NWP) system of the UKMO. The meteorological data that are required as input to NAME are:

- wind in three dimensions,
- cloud and precipitation,
- temperature,
- heat and momentum fluxes.

Depletion Processes

The removal of a pollutant from the atmosphere to the ground is a function of both the physical and chemical properties of the pollutant and the meteorological conditions. Wet deposition is removal by precipitation, either by washout, when the pollutant is scavenged out of the air by passing raindrops, or by the more efficient process of rainout which occurs as the pollutant is incorporated into rain within clouds. Wet deposition dominates over dry deposition and the effect of localised rainfall is to create hot spots of pollutant on the ground. Such wet deposition caused the Caesium-137 contamination in North Wales during the Chernobyl accident. Dry deposition is a result of the pollutant coming into contact with surfaces and adhering to them by impaction and inception. The dry deposition rate is proportional to the air concentration in the boundary layer.

Radioactive half lives can be imposed on appropriate species as a form of depletion and sedimentation can be applied to massive substances.

Further Information on NAME

Only a brief summary of NAME is presented here and further information on the structure, capabilities and parameterisations within NAME can be found in (Ryall & Maryon 1996).

4 Method

In order to run the NAME dispersion model it requires either

1. meteorological forecast data provided by a NWP model *or*
2. analysed meteorological data recorded on a NWP grid.

Forecast Data

For this investigation an ensemble NWP forecast was required, which was obtained by using a single ensemble forecast of the UM run at ECMWF. A set of ensemble forecasts out to T+132 hours was acquired, which provided 27 possible evolutions of the global atmosphere (subject to the limitation of the inherent model climatology). Each of the 26 ensemble members, plus the 27th unperturbed control member, were used as separate data inputs for NAME. The forecast began at 1200UTC on 1st March 1999.

The UKMO operational forecast taken from the full resolution UM was also obtained for a starting time (T+0) of 1200UTC on 1st March 1999 and run out to T+144. This provided the forecast information available at the time, and would have been the data that NAME would have been run on operationally if an emergency release had occurred on the afternoon or evening of 1st March 1999.

Analysed Meteorological Data

An analysed field of meteorological data is obtained by collecting measurements of wind velocity, temperature and humidity at different sites and blending them into a model field. These real data are assimilated into a model field taken from the most recent forecast, which is valid for the same time at which the observed information was recorded. The assimilation process is complex, as the dynamical balance has to be preserved, so the model is carefully 'nudged' in the direction of the measured data. In this way the regular grid of meteorological data is constructed as the initial frame (T+0) for a NWP forecast. The analysed field does not consist entirely of real data, but contains a background field of existing model information.

The archived analysed fields were used to run NAME for the same period as the forecasts, in order to produce the 'best' solution. This 'best' solution would differ from reality in two ways, from errors introduced by inaccuracies in the analysed field and errors in the dispersion as modelled by NAME.

Simulating a Release

Once the ensemble forecast, operational forecast and archived meteorological fields had been obtained, NAME was used to simulate a six hour release of pollutant. The release started at 1800UTC on 1st March 1999 in each of the respective model atmospheres and was followed for 138 hours. As each ensemble and the operational model atmosphere evolved into different states over the 138 hours, 28 different plumes were created from the same defined release. Identical releases were simulated within the analyses meteorological fields in order to produce a 'real' solution. The ensemble plumes were compared to the operational plume and the analysis plume in order to establish if the ensemble could add useful information compared to the operational solution alone.

A table summarising the release criteria and the NAME set up is shown in the Appendix, page 45.

Choosing Release Sites

A number of different release sites were used in order to make maximum use of the single global ensemble. The global ensemble made it possible to investigate different types of weather and also to try and find areas of high and low predictability, areas which should have been highlighted by the diversity of the ensemble solutions. The sites of release that have been used as case studies in this report are Stockholm, the Falklands and Sydney.

The perturbations which created the ensemble members were calculated by the ECMWF singular vector scheme. These singular vectors are maximised for the northern hemisphere extratropics, the perturbations chosen are those which create differences of maximum amplitude north of 30°N after 48 hours (Molteni et al. 1996). No perturbations are added to the tropics as the method for producing singular vectors uses a 'dry' model which therefore can not represent the latent heat fluxes which dominate events in the tropics.

5 Data Analysis

An unavoidable feature of ensemble forecasts is the vast amount of data that are produced and it is important to find techniques to consolidate and present this information so that it can be easily understood, without losing useful detail. It is also desirable to avoid simply reverting to deterministic interpretations of the ensemble distribution and therefore neglecting the probability information that can be gained from it. The reduction of 27 ensemble plumes into one plot has the benefit of reducing the amount of data to take in, but at the same time this one solution should still express the variability between ensemble members as it is their inherent advantage.

The data analysis methods used here rely on elementary statistics and the use of graphical output. The data from NAME is usually output graphically, as concentration contours superimposed on a map. The importance of the graphics in emergency situations is to unambiguously highlight the area(s) most at risk from the pollutant and this aim was also followed in the ensemble analysis.

The data analysis has been divided into two distinct areas due to the different uses of forecasts at each stage and the availability of information:

1. operational analysis,
2. retrospective assessment.

The operational analysis assesses the data only with regard to information that would be available during a real time run of NAME, namely, the ensemble members and the operational forecast. The retrospective assessment has the 'definitive' information provided by the archived meteorological data. This report addresses both types of information, but the emphasis is towards the hindsight assessment, since the purpose of this report is to analyse the performance and extra information offered by the ensemble towards the 'best' NAME solution.

In the operational analysis section the emphasis has been to address the valuable information that a forecaster or emergency planner might ideally gain from the ensemble, as opposed to the deterministic NAME data alone:

1. Which area(s) is (are) most at risk from high concentrations and do the ensemble indicate different or similar areas as being at risk?
2. The confidence level or probability that may be attached to any particular location suffering from the risk, with this confidence level also indicating the range of concentrations that might be expected at a site.

Deterministic operational forecasts can address the former requirement of NAME and the latter can be objectively assessed via the ensemble approach.

Basic Analysis Tools

The NAME output calculates concentration values for each cell of a regular grid which covers the globe. Cells that have a concentration greater that zero create a 'plume', which is the area of the material as distributed according to the meteorology fields and NAME dispersion parameterisations.

Due to the large number of ensemble members it is time consuming and potentially confusing to examine each plume individually. Instead, the following schemes have been developed to highlight objective information on the area, concentration and position of the ensemble plumes. The basic analysis methods that were employed are described here. The first four methods can be used purely operationally or in retrospective assessment of ensemble performance. The other methods labelled as 'retrospective' identifies those methods that perform a validation against the analysis plume.

i) Ensemble Mean

$$\overline{\mathrm{E}} = \frac{1}{n}\sum_{0}^{n} e$$

n = total number of ensemble members,
e = the value of concentration for the same cell in each ensemble member

The mean ensemble concentration was found for each cell and then plotted as a plume. The mean was then compared to the corresponding concentrations as determined by the operational forecast and also to the analysis plume. The ensemble mean included the unperturbed member.

ii) Intermittency

$$I = \frac{i}{n} x100\%$$

i = number of members present in the same grid cell for each ensemble member

The intermittency analysis is one of the simplest of the data analysis methods to interpret as it relies only on a hit or miss criteria and it disregards plume concentrations. If the plume has a concentration value greater than zero in a grid square, that grid square is scored as a hit. The number of the members present is then converted into a percentage for that square. The purpose of the intermittency is to indicate a confidence level as to whether the plume is likely to affect any particular grid square, but without giving an indication of how high the air concentration of the pollutant might be.

Intermittency can also be used in hindsight, to assess how well the ensemble covered the spread of the analysis plume. In particular it can be adapted to indicate where the entire ensemble has missed the analysis plume.

iii) Intermittency with an Imposed Threshold

The intermittency analysis can be turned into a probability of exceeding a certain value by applying a threshold to the ensemble concentration data. The threshold can be chosen in order to highlight the areas most at risk from the pollutant. The intermittency tool can then be applied in order to find the percentage of the ensemble that predicts concentrations above a certain value, which corresponds to a probability.

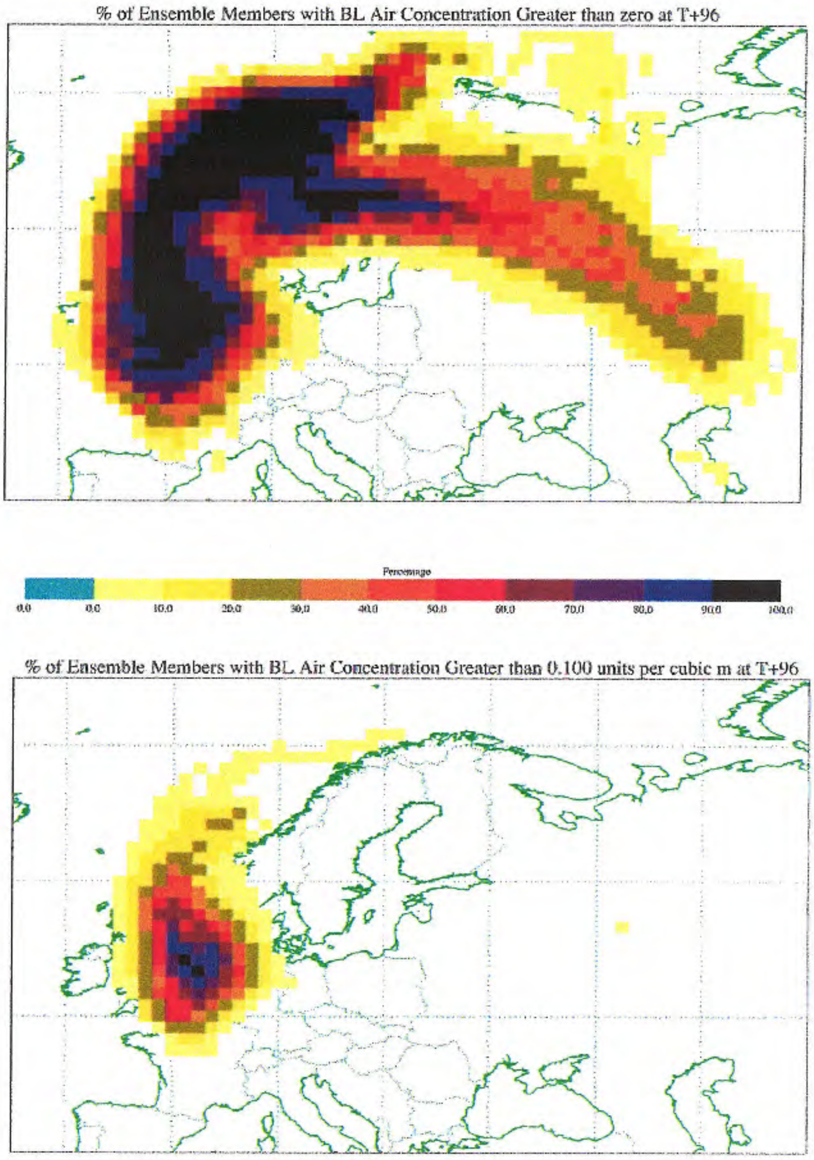


Figure 1: Intermittency, the percentage of members present. Top, with no threshold and below with a threshold of 0.1 units. Stockholm release at T+96.

iv) Expected Concentration in a Cell

The intermittency analysis can be considered to be a probability forecast, but only in terms of hit or miss. In order to apply probabilities or confidence levels to concentrations within the plume, but still produce understandable results, some information must be cut out. The method employed here considers the predicted concentration in each grid box. The expected concentrations for a certain cell is determined by each ensemble member and put into a histogram, and this can give real time estimates of the range of concentrations that might reach any particular location. For verification purposes, the concentrations can be compared to that for the analysis plume, noting that on most occasions the analysis should fall within the ensemble values if the ensemble has successfully captured the variability of the atmospheric state.

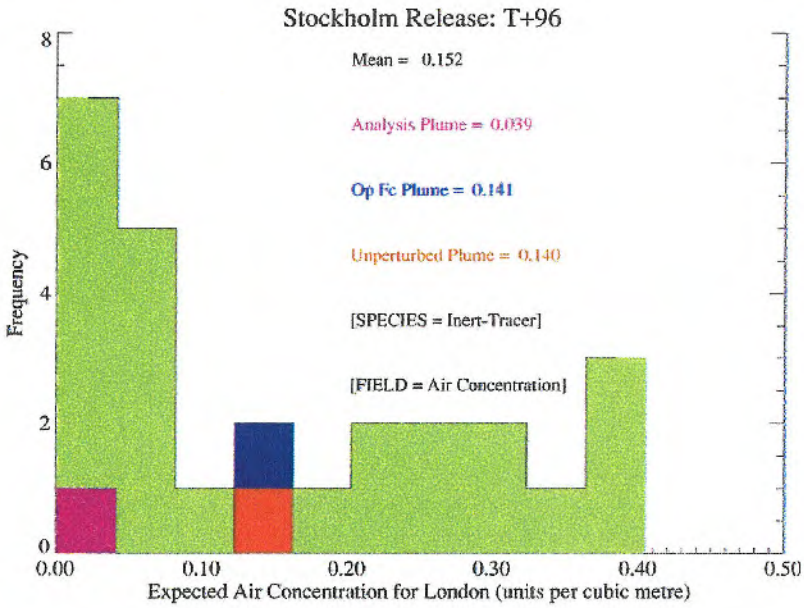


Figure 2: Example of Expected Concentration in a Cell

v) Percentage Overlap

The overlap of each ensemble plume compared with the analysis plume was also used as a data analysis method, simply using the area (number of cells) that the plume covered, not considering the value of the concentration. The cells overlapped by both plumes, the intersection, was divided by the entire number of cells covered by both plumes, the union. This quantity is also known as the figure of merit in space, *FMS* (Ryall & Maryon 1998) and has been used in previous atmospheric dispersion verification schemes.

If A = total number of cells covered by the analysis plume

M = total number of cells covered by each ensemble member

$$FMS = \frac{A \cap M}{A \cup M}$$

In the scheme presented here the *FMS* was converted to a percentage.

Comparing each ensemble plume to the analysis plume allowed histograms of the percentage overlap to be constructed. The operational forecast plume was also compared to the analysis plume.

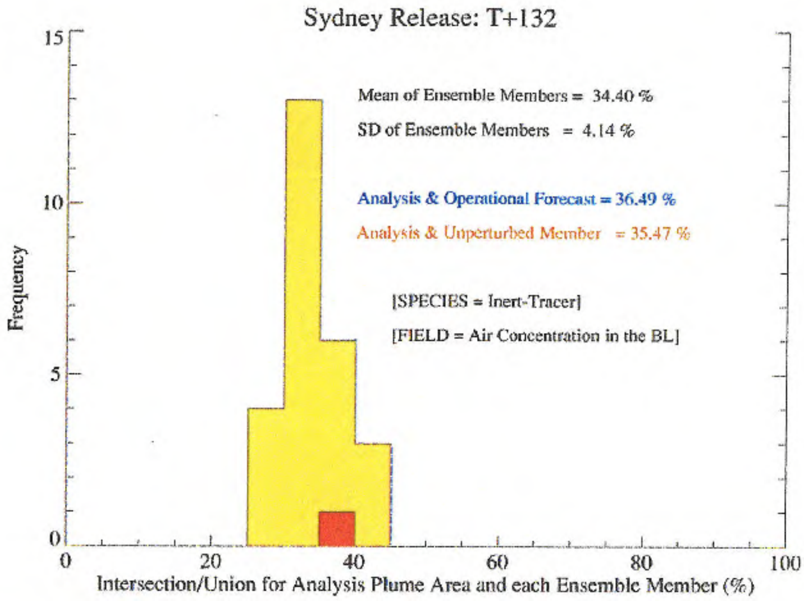


Figure 3: A histogram of the percentage overlap for Sydney at T+132.

vi) Best Fit Member

The percentage overlap method described in section 5 was also used to find the 'best fit' member, defined here as the ensemble member that produced a plume with the greatest *FMS*. This method demonstrates that suitably perturbed members, even at half horizontal resolution, can be more skilful than the operational forecast, although this skill test is of a very limited nature as it disregards concentration values.

vii) False Alarm Concentration

$$F = \overline{E} - a$$

a = the air concentration in the boundary layer of the analysis in each cell

The false alarm concentration, F, in each cell was found by subtracting the analysis concentration from the ensemble mean. For simplicity, any negative values where the analysis concentration was greater than the ensemble concentrations were ignored.

viii) Added Value

This is possibly the most complicated of all the analysis methods, but in graphical format it qualitatively emphasises the areas where ensemble members captured the analysis plume more accurately than the operational forecast. For each grid box it shows the percentage of ensemble members that had a greater concentration than the operational forecast, whilst not exceeding the concentration suggested by the analysis plume.

This graph is introduced here in order to show a first attempt at an overall assessment of the skill of the ensemble above the operational forecast; an assessment which is more sophisticated than the basic percentage overlap method. The 'added value' graph itself may not be useful, as it is a largely qualitative demonstration. However, the idea behind 'added value', which is to produce an objective assessment of the ensemble which includes both concentration and the position of that concentration is the point that is being illustrated.

6 Results

Ideally, the errors within NAME should be consistent between forecasts, in reality they are not as NAME is likely to produce different biases in different meteorological conditions. For simplicity in the validation of results, the errors in NAME have been ignored; the differences between plumes have been entirely attributed to the differences in the input meteorological data.

The results have largely been presented between T+72 to T+132 hours after the start of release. In weather forecasting this period is considered 'medium range' and is the timescale at which ensemble prediction systems are generally used. Due to these relatively lengthy time scales it may be misleading to look at specific details and values of results. Instead, where possible, order of magnitude comparisons have been made, along with an evaluation of the general pattern of results produced by the ensemble.

Case Studies

Several releases were investigated, but for brevity, only three case studies are presented here. These releases have been chosen to explain different techniques of presenting the ensemble information and to give a preliminary assessment of the probabilistic information offered by the ensemble. All of the data analysis methods have been applied to the Stockholm release at T+96 in order to give a comprehensive example. The Falklands and Sydney releases have only a selection of analysis methods presented in order to highlight points of interest.

The graphical information relating to the examples is contained in the Appendix and mean sea level pressure charts are included for reference.

Release Site	**Latitude**	**Longitude**
Falklands *(UK, South Atlantic)*	51.70S	58.00W
Stockholm *(Sweden)*	59.30N	18.00E
Sydney *(Australia)*	33.90S	151.50E

Figure 4: Table of the positions of the case study releases.

6.1 Stockholm Release

All graphical information for the Stockholm release can be found in the Appendix page 46 to 53.

The Meteorological Situation

The analysed pressure field for 1800UTC 01/03/99 is shown on page 46. In the analysed pressure field the simulated release was in the vicinity of a low pressure system. After 36 hours the low had tracked east into Russia, whilst a further area of low pressure moved onto the UK from the eastern Atlantic. A west to southwest airflow covered mainland Europe. The analysed pressure field 96 hours after the start of release, indicates that the low over the UK weakened as it moved onto Denmark and northern Germany and in its wake a north to northwest wind developed over the UK, France and Spain.

Description of the Analysis Plume

Basic plume data valid at T+96 hours after the start of the release are presented on page 47, comprising the plumes created by the analysed meteorological fields, the operational forecast, the unperturbed ensemble member and the best fit member of the ensemble. All individual plumes are plotted as the air concentration, in the boundary layer, of the inert tracer in arbitrary units. The concentrations are averaged over the six hours preceding the frame time.

The analysis plume, page 47(a) shows that some of the pollutant was captured by the low over Scandinavia and moved east over Russia. The area of higher concentration was taken up in the cyclonic flow that developed over the UK and by T+96 this maximum concentration was being drawn across the UK and into France in the northwesterly winds.

Description of the Operational Analysis Techniques

The operational forecast plume is shown on page 47(b) and the operational analysis techniques are presented on page 48. The operational forecast has similarities to the analysis plume in its general shape, including the position of the maximum concentrations of the tracer over the UK and with a tendril of lower concentrations stretching east from Scandinavia. Differences between the operational plume and the analysis include; the failure of the operational forecast to take the tracer far enough south over France at T+96, the plume does not cover southern Norway and it has not transported the pollutant far enough east, as the analysis plume indicates a low concentration of material as far east as the Caspian Sea at T+96.

i) Ensemble Mean

The mean ensemble air concentration, shown on page 48(a), demonstrates that the ensemble broadly captures the same plume shape as the operational forecast, page 47(b). The subsequent plot of mean ensemble air concentration minus the operational forecast is designed to highlight the areas where higher concentrations of the pollutant might be expected. The mean ensemble plume extends further east into Russia than the analysis plume and there is an indication of higher concentrations up the northern coast of Norway in the ensemble.

ii) Intermittency

The intermittency, or percentage of ensemble members present, confirms that all plumes are present in the higher concentration areas such as over the UK and the northern North Sea. The plume that extends into Russia has between 10 and 50% of the ensemble members in support of that scenario. The plot of the intermittency outside the operational forecast plume is shown on page 48(d). This plot removes the operational forecast area (taken at any concentration), from the intermittency plot in order to highlight information being indicated by the ensemble which at the same time is not supported by the operational forecast plume at all. In this case this clearly demonstrates the 10 to 50% probability of material being taken into Russia and a bias to keep the plume further south over Norway and Sweden.

iii) Intermittency with an Imposed Threshold

A time series of intermittency with an imposed threshold is presented on page 49. This chart highlights the percentage of ensemble members present that exceed the threshold, which corresponds to a percentage probability of exceeding the threshold if all ensemble member solutions are considered equally likely. The threshold has been set at 0.01 units per m^3 air concentration of the inert tracer in the boundary layer. The T+96 frame can be compared with the same plot with zero threshold on page 47(c).

iv) Expected Concentration in a Cell

Expected concentration in a cell is shown on page 50. Four time frames have been presented for the predicted air concentration in the boundary layer at Oslo, which is situated 400km west of Stockholm. The histogram can be converted from a frequency distribution to a probability distribution by assuming that the 27 ensemble members cover all possibilities and that each member is equally likely. Any number of probability forecasts can then be made using the ensemble information; for example, at Oslo at T+36 there is 15% probability that the air concentration of the inert tracer would be above 0.4 units per m^3, and by T+72 this has fallen to 4%. Furthermore, in both the T+72 frame and the T+96 frame the air concentration suggested by the analysis is higher than the operational plume prediction, but the concentration is still captured within the ensemble envelope.

The highest concentration value at T+36, 72 and 96 are attributed to the same ensemble member, but at T+132 the out-lying member is different. It should also be noted at T+132 that the highest concentration is more than twenty times higher than the next highest prediction. The ability to capture this unusual development is one of the key skills of an ensemble forecast. The analysis concentration was more than two orders of magnitude less than this out-lying member, but the highest probability would have placed the analysis plume at the lower concentration.

The mean and standard deviation is noted on each graph. The fact that the histograms do not appear to approximate to Gaussian distributions is the main reason why the standard deviation of plume concentrations has not been pursued as an analysis tool.

Description of the Retrospective Assessment of the Ensemble

The validation techniques are shown on page 51. Charts (a) and (b) assess the ensemble members against the analysis, and (c) and (d) asses the additional correct information that the ensemble signalled in comparison with the operational forecast.

i) Intermittency within the Analysis Plume Area

The percentage of ensemble members within the analysis area takes the usual intermittency plot and restricts it to within the coverage of the analysis. If the analysis plume was present in any concentration in a grid box then the percentage of ensemble members present in that grid box is presented. A comparison with the analysis plume, page 47(a), confirms that the intermittency has been plotted within the outline of the analysis.

The plot on page 51(a) shows that much of the plume was covered by between 70 and 100% of members. The failure of the ensemble, however, is indicated over France, Switzerland and northern Italy. The light blue colour in the plot indicates that no ensemble members were present in around thirty grid boxes, or 100,000 square kilometres. The intermittency accounts for any level of concentration, although in this area as described by the analysis plume, concentrations not captured by the ensemble reach values of up to 0.1 unit per cubic metre. The extension of the plume into Russia is captured by the ensemble, although its southerly extent across Norway and Sweden is only captured by between 10 and 60% of the ensemble.

The use of thresholds in retrospective analysis is presented on page 52. The analysis plume has been plotted for concentrations greater than 0.01 units per m^3 and the corresponding intermittency within that analysis area has also been plotted, using ensemble member concentrations of greater than 0.01 units per m^3. The charts emphasise that the ensemble plumes fail to capture the southerly extent of the analsis area and all ensemble members (shown in black) only cover a small portion of the total area.

ii) False Alarm of the Mean Ensemble Air Concentration

The mean false alarm of the ensemble air concentration field, page 51(b), is designed to emphasise areas of over prediction. For the Stockholm release at T+96 it indicates that the area with the greatest over prediction of concentration was in eastern England, the North Sea and Belgium. This false alarm example demonstrates that the maximum concentration of the ensemble was too broad and extended too far east in comparison with the analysis plume maxima.

iii) Analysis Plume Not Captured by the Operational Forecast Plume

The air concentration of the analysis plume that was not captured by the operational forecast plume is shown on page 51(d). The operational forecast had high enough (or too high) concentrations over the North Sea, but as in the ensemble forecast, failed to take the plume far enough south over France and Italy.

iv) Added Value

The added value chart, page 51(c), is a qualitative assessment of how the ensemble plumes improved on the information offered by the operational forecast plume alone. Added value draws attention to areas where the ensemble members gave a correct signal of greater concentration, to within an order of magnitued. The greatest improvement offered by the ensemble was over the Norwegian Sea where between 50 and 100% of ensemble members correctly exceeded the operational forecast concentration.

v) Percentage Overlap

Histograms of the intersection divided by the union for the area of each forecast plume compared to the analysis plume is presented on page 53. The histograms show that in this example the distributions tended to be centred around the value achieved by the operational forecast and the unperturbed member. It can also be seen that the spread resembles a normal distribution and the standard deviations do not vary a great deal during the forecast period. A comparison of the mean values reveals that the area of overlap of the plumes was lower at T+36 than at the other time frames shown.

6.2 Falklands Release

The Meteorological Situation

The analysed pressure field for 1800UTC 01/03/99 is shown on page 54. In the analysed pressure field the simulated release was in northerly winds produced by an area of low pressure to the west and a ridge of high pressure to the east. The low moved southeast to reach the Antarctic Peninsula by

T+36. In the following days the ridge of high pressure weakened in the mid South Atlantic and a strong westerly flow developed by T+96.

Summary of the Falklands Results

The analysis and operational plumes are shown on page 55(a,b) and two intermittency plots are also presented (c,d) in order to highlight the ensemble distribution. A comparison of the analysis plume and operational plume reveals that the position of highest concentrations within the analysis remained to the north of the Falklands, whilst the operational forecast suggested that concentrations of up to 0.1 units per m^3 of the inert tracer would be taken south to Antarctica. The intermittency plot, which displays the percentage of all ensemble members present in each grid box, indicates that whilst all ensemble members are present in the northern part of the plume, only 40-80% of members are present in the section of the plume that extends to the Antarctic. The plot of the intermittency within the analysis area emphasises that the ensemble coverage is high in the northeastern section, but fewer ensemble members cover the south and west. The plot indicates that the ensemble is, in general, too far east, suggesting that the westerly winds were represented as stronger in the ensemble than in the analysed fields.

6.3 Sydney Release

The Meteorological Situation

The analysed pressure field for 1800UTC 01/03/99 is shown on page 56. According to the analysed pressure field, the simulated release was in northeasterly winds, a consequence of an anticyclone situated in the Tasman Sea. Between T+24 and T+36 the high moved onto New Zealand and a low developed off the east coast of Australia. The low pressure deepened as it crossed the Tasman and cleared southern New Zealand shortly after T+96. Winds remained light over central Australia as ridge of high pressure over western Australia strengthened in the Australian Bight. By T+132 an anticyclone was situated to the west of Tasmania with a northeasterly flow over southern and eastern Australia.

A comparison with the pressure charts from the unperturbed member of the ensemble forecast indicated the synoptic developments were broadly similar, although the low in the Tasman Sea became deeper and slower moving than in the analysed fields and stronger northeasterly winds developed over Australia after T+96.

Summary of Sydney Results

The analysis charts for the Sydney release at T+132 hours are shown on page 57. It can be seen from the analysis plume (a) that most of the inert tracer was kept over the Tasman, with only low concentrations penetrating into Queensland. The operational forecast (b) predicts a narrower plume with air concentrations of an order of magnitude higher over Queensland and material extending further into central Australia. The mean false alarm concentration is a plot of the ensemble mean minus the analysis concentration. This chart reveals that mean concentrations were too high by around 0.1 units per m^3 around Brisbane and extended too far east by approximately 1800km. The intermittency chart (d) also highlights that all ensemble members were responsible for taking material into the centre of Australia.

A time series of the expected concentration at Brisbane is shown on page 58. The histograms demonstrate that the prediction at T+72 has a wide range of possible concentrations. The analysis plume concentration, supported by 50% of the ensemble, is for a concentration which is three orders of magnitude smaller than the highest of the ensemble predictions. It is interesting to note that at both T+72 and T+96 the unperturbed member has a notably different concentration to the majority of the ensemble, which is an indication that the perturbations are producing a variable range of solutions. At T+132 it can be seen that there is a spread of possible concentrations and the analysis plume produced a concentration represented by 30% of the ensemble.

6.4 Summary of All Results

The Stockholm release at T+96 demonstrated that the ensemble was able to capture relatively low concentrations of the pollutant which were taken around three thousand kilometres in the opposite direction to the main area of higher concentrations of the inert tracer. Not all ensemble members took the pollutant east in this manner, suggesting that its occurrence was with a probability of between 10 and 50%. The use of the histograms of expected concentrations in a cell suggested that the air concentrations of the tracer demonstrated by the analysed fields were often captured within the ensemble concentration distribution. The expected concentration histograms can also be converted into probability distributions by assuming that all the ensemble outcomes are equally likely. Finally, the Stockholm release showed that the ensemble members failed to capture part of the analysed plume, indicating that the individual ensemble members were not sufficiently different from one another to mimic the analysed fields; the analysis plume was outside the envelope suggested by the ensemble.

The Falklands release at T+96 highlighted that the ensemble scheme could improve on the guidance of the operational forecast alone. A comparison of the analysis plume and that produced by the operational forecast revealed that the operational forecast over predicted concentrations of the inert tracer in one area. The addition of ensemble information in the form of an intermittency analysis indicated that not all ensemble members were present and implied that any level of concentration being moved there was with a probability of between 40 and 80%.

The Sydney release at T+132 indicated that the ensemble could give misleading information. The intermittency analysis demonstrated that all ensemble members took some of the inert tracer into central Australia at this time, whilst the analysis plume remained mainly over the Tasman Sea with only low concentration affecting Australian land. The addition of the ensemble mean false alarm information showed that average concentrations taken into central Australia were around two orders of magnitude too high in places.

7 Discussion

7.1 Assessment of the Data Analysis Techniques

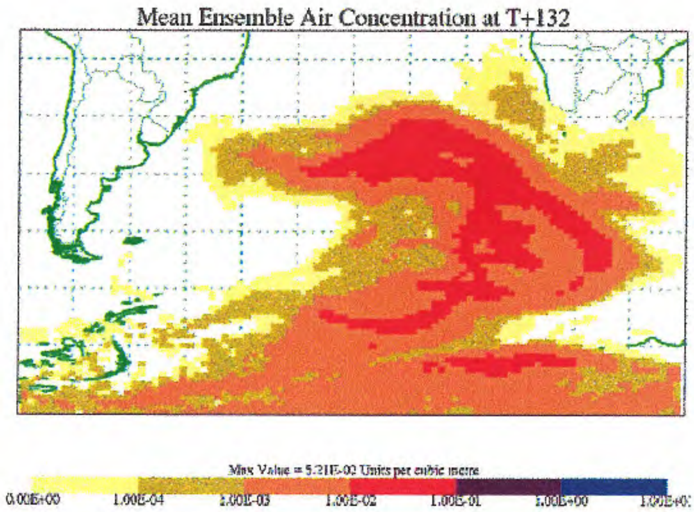
The emergency response requirement of the NAME dispersion model is to highlight areas at risk following accidental releases of airborne pollutants. The data analysis techniques presented in this study focus on this operational aspect of atmospheric dispersion modelling. An ensemble forecast has been used in conjunction with NAME to predict a range of concentrations that might be expected from a specified release and to also give an estimate of the confidence in that prediction. By comparison with NAME run on archived meteorological data, the data analysis is also designed to assess the effectiveness the ensemble plumes in indicating the position, coverage and concentration of a 'real' plume. The data analysis should also be easy to interpret and to address this requirement the techniques have graphical outputs.

The current purpose of ensemble forecasting is to add predictive skill to weather forecasts in the medium range, which is 3 to 10 days after the start of the forecast (Molteni et al 1996). This additional skill is achieved by using ensemble predictions as a practical means of estimating the range of solutions possible in the period where error evolution has become nonlinear. The emphasis on the medium range advantage of ensembles has been considered in the assessment of the data analysis techniques and the period from T+72 to T+132 has been studied more closely than earlier parts of the forecast.

i) Ensemble Mean

In ensemble weather forecasting the ensemble mean is a chart which is an average of all the ensemble fields and it is a technique that is regularly used; it can be shown that in general the ensemble mean has more predictive skill than any of its individual members, partly because the average filters out extreme elements (Mittelstadt 1995).

The mean ensemble concentration uses all the individual weather forecasts to produce different plumes which are then averaged. The mean ensemble concentration is attractive as a simple scheme for interpreting the whole ensemble, it does result in a solution which cannot convey the variation of the ensemble or provide any probability estimates.



The mean ensemble air concentration for the Falklands release at T+132 hours shows one of the difficulties of interpreting the ensemble mean after the plume has been allowed to develop for a few days. (Some of the spread can be attributed to the projection of the most southerly latitudes onto a regular grid, distorting the lower part of the plume by apparently spreading it over a wider area.) The averaging of the plume results in a loss of detail and disguises where high concentrations from some members are cancelled by low concentrations from others. A semilog plot, producing more contour intervals, does not give any appreciably greater detail than the log concentration plot shown.

The mean ensemble concentration is useful to examine as it indicates the entire spread of the plumes. When the mean is used in conjunction with the intermittency plot it then has the potential for estimating areas most at risk from high concentrations. Areas of high average concentration can be cross-referenced to check what percentage of the ensemble produce that high concentration and so indicate a first estimate of the areas most at risk, namely, high concentration with high probability. It should also be noted that the mean spreads the plume and increases the false alarm area.

A final comment is that the ensemble mean is not a real solution of the model climate. If a perfect NWP model is assumed, the ensemble mean is not a state that the atmosphere will evolve to, although in early parts of the forecast some close approximations are possible.

ii) Intermittency

Intermittency, like the mean, is useful for establishing the entire envelope that the ensemble plumes cover. The intermittency plot can also indicate the most likely position of the plume, but without the benefit of concentration values. It can provide estimations of the confidence of where the plume will cover and it becomes more powerful when used in conjunction with the ensemble mean or an imposed threshold.

The charts of mean concentrations and intermittency appear broadly similar, especially up to T+96 (for an example see page 48). This similarity occurs partly because the overall shape of both plots is identical as both techniques cover the entire area of all the plumes. It can also be inferred that the ensemble plumes have similarly shaped concentration distributions as it suggests that the high concentrations occur in the same place for many members and that the lowest concentrations are more widely dispersed by low numbers of ensemble members. Confusion between average concentration and percentage of the ensemble present can therefore occur, and analysis techniques should seek to prevent this.

iii) Intermittency with an Imposed Threshold

The addition of thresholds to the intermittency method makes it a flexible and useful technique for determining the probability of any area exceeding a certain concentration level. The introduction of a threshold draws attention to areas potentially most at risk from the higher concentrations in the plume. The thresholds presented in this study are arbitrary, but it would be possible to set the threshold at pre-determined levels depending on the hazard that a pollutant presents.

iv) Expected Concentration in a Cell

Despite cutting out all the information regarding the coverage of the plumes, the histograms of expected concentration give a succinct and easily interpreted presentation of the data at a chosen location. The concentration distribution described by the histogram indicates the level of concentration that may be experienced at any individual location and the probability of exceeding any threshold can also be estimated. Access to the entire range of concentration that the ensemble predicts is useful for discovering extreme predictions such as the T+132 frame, page A7, where the maximum concentration is twenty times higher than indicated by the next highest.

The inability to compare large areas in one diagram is a notable limitation and producing several charts for different locations is possible, but may result in confusion. The appeal of this technique is that it faithfully retains all of the ensemble information at each location as contributed by each member.

v) Percentage Overlap

The percentage overlap technique is an estimate of the skill that each forecast had with respect to the analysis. Skill in this context is a measure of how well the position and area of the forecast plumes compare with the analysis, also known as figure of merit in space (*FMS*).

Despite the limitation of not including any concentration values in this assessment of the ensemble, it is a technique which is simple and easily visualised, and it could be the basis of an objective assessment of the ensemble performance. The ability to be able to convert the overlap to a percentage is also attractive and easy to convey as a test of the ensemble. The overlap technique could be used in conjunction with a threshold in order to remove very low concentration areas, which in the current method carry as much weight as high concentration areas. From the examples that have been studied it is unusual to find an overlap score of 70% or higher, although no great importance is attached to this number.

In addition to the values of percentage overlap of the ensemble, the relative skill of the operational forecast and unperturbed member were also determined. The unperturbed member typically showed an overlap percentage which was in the centre of the ensemble values. Given that the error growth in atmospheric prediction is linear for the first 48 hours (Sivillo et al. 1997) this result is to be expected, but at later times as errors become nonlinear this should not necessarily be the case. The Stockholm example, page A10, also indicated that the overlap distribution remained a largely similar shape up to T+132, but again, this need not necessarily be the case when error growth becomes nonlinear. A final observation of the percentage overlap technique demonstrated that the operational forecast member often achieved a greater coverage of the analysis plume, especially early in the forecast. An example is shown in figure 6. The greater skill of the operational forecast is explained by its higher resolution.

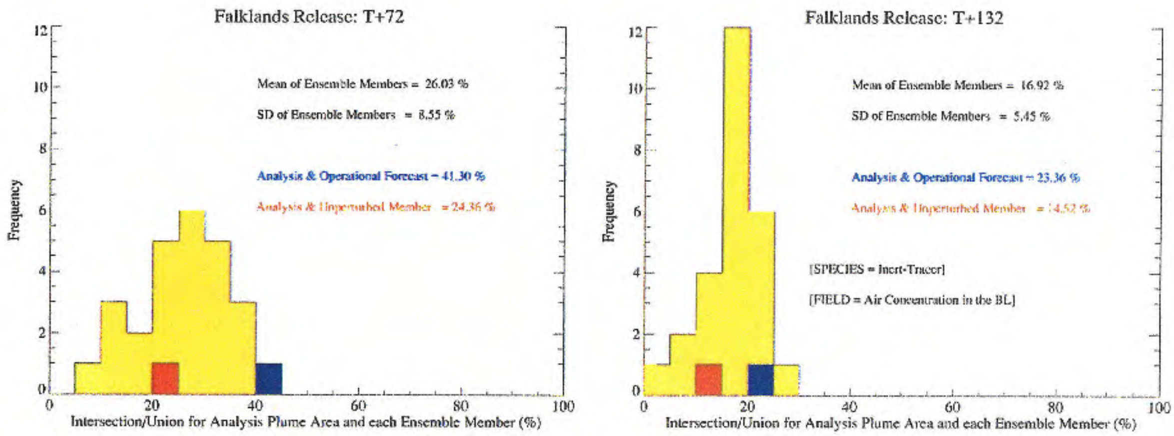


Figure 6: histograms of percentage overlap, showing the greater skill of the operational forecast plume (shown in blue) compared to most ensemble members.

vi) Best Fit Member

A technique often used in ensemble forecasting is to find the member that retained most skill during the forecast. The method used for determining the best fit member, in this instance, was the overlap method discussed above. The shortcomings of the overlap method are therefore inherent in choosing the best fit member.

The best fit member demonstrates that, at medium range periods, a perturbed member of a lower resolution ensemble can retain more skill than the operational forecast, illustrated in the T+132 frame of figure 6. Verifications of the ECMWF ensemble system have shown that for individual ensemble members in the medium range, about one third of perturbed forecasts were closer to the verifying analysis than the higher resolution unperturbed forecasts (Molteni et al. 1996). Whilst the concept of best fit members is interesting to observe, its benefit to ensemble atmospheric dispersion modelling at this time is limited.

vii) False Alarm Concentration

The modification of the mean ensemble concentration by subtracting the analysis concentration field gives a first approximation to the areas that would have suffered most over-prediction of concentration following a release. In practical situations false alarm is an important consideration where mobilisation of emergency services is necessary. The disadvantage of losing the details of the individual ensemble member contributions is also a limitation of this analysis technique, as discussed previously in the ensemble mean section.

viii) Added Value

Added value identifies, for each grid box, the percentage of ensemble members that had a greater concentration than the operational forecast, whilst not exceeding the concentration of the analysis plume by more than an order of magnitude. The inclusion of this analysis technique is to demonstrate a system that could be adapted to an impartial skill score to compare different releases. It should, however, be acknowledged that this added value system has many faults, including the bias of only judging the concentrations greater than the operational forecast and not accounting for the concentrations correctly captured by the ensemble that are less than the operational forecast. It is likely that any skill scoring system is likely to be highly complex and would require a great deal of development and assessment in order to establish its objectivity.

7.2 Comments on the Ensemble Prediction System

Ensemble Spread

In ensemble prediction, ensemble spread refers to the amount that the ensemble member solutions deviate from the unperturbed member. It is often taken as the 75^{th} percentile of a certain quantity, such as the root mean square 500hPa height difference between the perturbed ensemble members and the unperturbed. The underlying importance of the ensemble spread is that it should be sufficient to include the actual development of the atmosphere whilst, not being so great that it produces arbitrary forecasts that are random solutions to the model climate.

The ensemble spread is a qualitative measure of the influence of the meteorological input data on the plume dispersion (Straume 1998). Arguably, it can also be assumed that plume dispersion is also a measure of ensemble spread; if the ensemble plume does not entirely cover the area that the real plume creates then the actual meteorological situation has not been captured within the variations of the ensemble forecast and this indicates that the ensemble spread is not sufficient. It should be noted at this point that an actual plume description could only be obtained by experimental data recorded after a real release. The analysis plume used for verification in this study is only an approximation to reality, depending on the mixture of observed meteorological data assimilated into a model field analysis.

Insufficient ensemble spread is known as an occasional shortcoming of ensemble prediction, the spread sometimes does not provide solutions close to the atmospheric state. Reasons for this weakness are attributed to the sometimes adverse contribution of model errors and underestimation of certain flow types in the model climate (Molteni et al. 1996).

Ensemble spread is a function of the selection of the ensemble members, so ensemble spread is a result of the initial perturbations imposed on the analysis field. The perturbations used in the UKMO system are those which are calculated by the ECMWF scheme and since the study presented here was undertaken, the initial perturbations imposed for the UKMO ensemble have been increased by 20%.

Discussion of Examples Relating to Ensemble Spread

The Stockholm case study at T+96, page 51, reveals an instance where the analysis plume is not captured within the envelope of the ensemble area. This implies that either:

- the ensemble spread was insufficient and so failed to disperse the pollutant enough, or
- the analysis fields produced errors which were unrealistic and therefore were not captured in the ensemble, or
- errors in NAME were not consistent between the ensemble members and the analysed meteorological fields.

It should be noted that the Stockholm example was the most marked case of insufficient coverage of the analysis plume of all the cases that were studied, as it persisted for around 36 hours from T+72 until T+108. Lack of coverage by the entire ensemble was also occasionally found to occur up to T+48 in the plume development due to minor differences in the direction of the analysis plume compared to the ensemble member trajectories which was not offset by diffusion.

The Sydney release demonstrated that it is possible for all the ensemble members to take their plumes to an area not covered by the analysis, giving a strong signal which was a false alarm. In this example it was apparent that the ensemble members seemed to be very similar to the general signal given by the operational forecast plume, which took the plume in a similarly erroneous direction. The Sydney release is an example, also found in other examples, that on occasions the ensemble members appear to be more like each other than they were like the analysis member. A reason for this shortcoming of the ensemble in Australia could be attributed to the fact that ensemble perturbations are maximised for the northern hemisphere extratropics, so sufficient variability may not be included in parts of the southern hemisphere.

Despite the occasional inadequacies of ensemble spread, the ensemble scheme has still demonstrated that it can provide information additional to that of the operational forecast. The Falklands release showed that the ensemble could indicate that the operational forecast had arrived at a solution supported by between 40 and 70% of the ensemble. The ensemble added useful information by showing some contradiction to the operational forecast plume. This is an important result to note as it demonstrates the strength of the ensemble prediction system.

To be in a position to comment with full authority on the effectiveness of ensembles it would require the objective assessment of a number of different ensemble forecasts. A large-scale assessment would in turn require the development of a scoring system so that the performance of different ensembles could be compared. The limited examples explored in this study have shown that: retrospective assessment demonstrates that useful information can be added to areas not already covered by the operational forecast (or unperturbed member) and that areas covered by the operational forecast can be of a lower probability within the ensemble.

Data Handling and Storage

The single most important impediment to using ensembles routinely is the computing resources that would be required, firstly to run the ensemble forecast at the Met Office and secondly to implement NAME in ensemble mode. Data handling was a significant part of this study as the single ensemble forecast amounted to 12GB and each set of NAME generated plumes for each release site took 3GB. The ability to run an ensemble and then manipulate these data in real time would be a significant undertaking. Prospects for increased computing power over the coming years offers hope for a remedy to the problem, and if this can be overcome then the opportunity for developing a real time ensemble atmospheric dispersion model increases dramatically.

Differences in Model Resolution

The operational forecast was approximately twice the resolution of the ensemble forecast and the analysis field was also diagnosed at this higher resolution. The consequence of the resolution differences is that the analysis field was run with more information than the ensemble members and so could be a source of differences between the ensemble plumes and analysis plume.

7.3 Suggestions for Further Work

Comparison with Real Data

Since this study only examined one ensemble forecast, in order to obtain a more objective view of the ensemble performance it would be desirable to perform more ensemble NAME simulations and compare them to analysis plumes.

In order to overcome the errors of the analysis meteorological field, it would be beneficial to carry out a comparison with observed data from a real release. A suitable dataset exists in the form of the European Tracer Experiment (ETEX). The experiment entailed the release of an inert atmospheric tracer from northern France, which was then followed over the next 72 hours by detecting and recording its air concentration at a network of 170 stations across Europe.

NAME has already been verified against these data (Ryall & Maryon 1998) and was found to perform well, successfully predicting the overall spread and timing of the plume across Europe. However, in common with most other models, NAME tended to over-predict the observed concentrations. It would be interesting to use the Met Office analysis data for the start of the release in order to generate a UKMO ensemble forecast for that period (which was prior to the introduction of ensemble forecasting).

Another atmospheric dispersion model, the Severe Nuclear Accident Program (SNAP) model has been used with the corresponding ECMWF ensemble forecast and the ETEX observations for this time (Straume et al. 1998).

Improvement of Data Analysis Techniques

The list of improvements to the analysis techniques, both operational and retrospective, is numerous and only some of the more important areas to address are included here.

The application of thresholds to the plume concentration data was shown to be a useful tool in establishing the probability of exceeding a specified concentration. The extended use of thresholds would be desirable in the future, especially in operational analysis of data, in order to simplify the presentation of ensemble information.

The analysis of ensemble and operational forecast performance up to T+72 has indicated that the operational forecast is often more skilful at this time, as illustrated by figure 6. The greater skill of the operational forecast is derived from its increased resolution over the ensemble members. A suggestion for an operational analysis technique is to use the operational forecast plume as an ensemble member, but giving the operational forecast extra weight in the early period of the forecast. The bias to the operational forecast could then be removed at medium range. This technique is currently applied in the NCEP (USA) ensemble weather forecasting scheme (Mittelstadt 1995). The added importance of the operational forecast in the early part of the forecast may help to add skill to the ensemble performance at this time and would reduce the amount of data to present.

For the operational emergency response using ensemble atmospheric dispersion techniques, a key factor to address is the amount of data generated. To avoid confusion and ambiguities, any future techniques developed should aim to consolidate and present the information without losing the probability estimates provided by the ensemble. It is acknowledged in ensemble prediction in general that there is much work to be done in the development of user-orientated products (Molteni et al. 1996).

Clustering is a technique currently used in ensemble forecasting to reduce the amount of forecast data to evaluate. In weather forecasting, clustering is applied in order to group ensemble members into a small number of similar solutions and the number of ensemble members in any cluster suggests the probability of that scenario occurring. By trying to approximate real solutions, clustering is a more advanced form of the simple ensemble mean approach. The cluster forecast fields could be used to run NAME to produce just a few solutions. This would help to reduce the large amounts of data associated with ensembles by degrading meteorological data into smooth, but still useful fields. Clustering in this way would restrict any probabilities to those of the original clustering.

Clustering, however, could be introduced as a post-processing tool to NAME (as all of the other data analysis techniques presented here). The ensemble of plumes produced by NAME could be clustered into a few different solutions by grouping like plumes as determined, say, by comparing the areas of the maximum plume concentration. The application of a clustering algorithm to the NAME output (rather than NAME input) would be preferable, as it would leave the reduction of information to the latest stage. However, clustering is a deterministic approach to ensemble information and so it is only desirable either in conjunction with other truly probabilistic techniques or if data interpretation is too much of a burden without the reduction of information that clustering can offer.

The analysis techniques described in this report could be applied to different fields and species available within NAME, such as wet deposition and dry deposition to the ground. For simplicity, this study used an inert tracer, which does not undergo any deposition to ground and is only removed from the boundary layer by entrainment into the free atmosphere above. The deposition fields, especially wet deposition, would entail a more complex analysis. Wet deposition would rely on both the windfields investigated in this study, but also different rainfall patterns between ensemble members.

The final recommendation for future work is the creation of an objective verification system for the comparison of different ensemble forecasts. The task of producing a measure of skill would be complex, as it would have to ensure that the plume position, area of coverage and concentration distribution were all accounted for in the plume assessment. It would also be important to normalise the score so that it was unaffected by the increase in plume area with time, so that plumes of different ages could be compared.

Improvement of Ensemble Spread

Since obtaining the ensemble data used here the perturbations imposed on the analysis field have been increased by 20%; a greater ensemble spread would now be likely. It would be interesting to investigate the greater spread on the dispersion of the inert tracer.

8 Conclusions

The ability to produce probabilistic dispersion forecasts by using an ensemble NWP forecast and an atmospheric dispersion model has been demonstrated. The ensemble prediction system of the UKMO was applied to NAME to produce simulations of atmospheric transport. The same configuration of NAME was used with the higher resolution operational forecast and the analysed meteorological fields in order to perform some validation of the ensemble method. These comparisons have shown that:

1. Some ensemble members correctly indicated the transport of the pollutant to areas that were not covered by the operational forecast but that were covered by the analysis plume.
2. An area falsely predicted by the operational forecast to be affected by the pollutant, was not affected by all ensemble member plumes (less than 100% probability).
3. On occasions the ensemble can provide misleading information, by indicating that all ensembles were present in an area not affected by the analysis plume.
4. The analysis plume was not always entirely covered by the ensemble envelope and on these occasions it appeared that all the ensemble members resembled the spatial distribution of the operational forecast.

It is desirable to find practical techniques to add reliability to forecasts or extend the time period over which they remain skilful; ensemble prediction coupled with atmospheric dispersion modelling shows promise by offering a feasible way of producing objective probability estimates. The use of probabilities is beneficial in the medium range and ensemble dispersion techniques could be a valuable planning tool for emergency response.

Ensemble prediction is data intensive in both input and output. The quantity of data that can be produced must be handled carefully and forecast products must be designed with the end user in mind. The introduction of a comprehensive validation scheme would be useful for the intercomparison of future ensemble dispersion trials.

9 References

1. Anderson, L.J., 1996 'Selection of Initial Conditions for Ensemble Forecasts in a Simple Perfect Model Framework'. Journal of the Atmospheric Sciences, vol 53 no 1, 22-36.
2. Ehrendorfer, E. and Tribbia, J.J., 1997 'Optimal Prediction of Forecast Error Covariances through Singular Vectors'. Journal of the Atmospheric Sciences, vol 54 no 2, 286-313.
3. Evans, R.E., Mylne, K.R. and Harrison, M.S.J, 1998 'Preliminary Results from Quasi-Operational Multi-Model Multi-Analysis Ensembles on Medium-Range Timescales'. UKMO Forecasting Research Technical Report Number 258.
4. Leslie, L.M. et al., 1994, 'The Performance of Four Global Models over the Australian Region'. Weather and Forecasting, vol 9 number 2, pg 229-240.
5. Lorenz, E.N., 1993, 'The Essence of Chaos', University of Washington Press.
6. Maryon, R.H. and Best, M.J., 1992, 'NAME, ATMES and the Boundary Layer Problem'. UK Met Office Turbulence and Diffusion Note 204.
7. Maryon, R.H., 1994, 'Modelling the long range transport of radionuclides following a nuclear accident'. Nuclear Energy, vol 33 no 2, 119-128
8. Mittelstadt, J., 1995, 'Introduction to Ensemble Forecasting'. Western Region Technical Attachment no 95-29. http://nimbo.wrh.noaa.gov/wrhq/95TAs/TA9529/ta95-29.html
9. Pasquill, F. and Smith, F.B., 1983, 'Atmospheric Diffusion'. Ellis Horwood Limited, Chichester.
10. Ryall, D.B. and Maryon, R.H., 1996, 'The NAME 2 dispersion model: a scientific overview'. UK Met Office Turbulence and Diffusion Note 217b.
11. Ryall, D.B. and Maryon, R.H., 1998, 'Validation of the UK Met. Office's NAME model against the ETEX dataset'. Atmospheric Environment, vol 32 no 24, 4265-4276.
12. Sivillo, K., Ahlquist, J.E. and Toth, Z., 1997 'An Ensemble Forecasting Primer'. Weather and Forecasting, vol 12 no 4, 809-818.
13. Straume, A.G., Koffi, E.N. and Nodop, K., 1998, 'Dispersion Modeling Using Ensemble Forecasts Compared to ETEX Measurements', Journal of Applied Meteorology, vol 37 no 11, 1444-1456.

10 Appendix

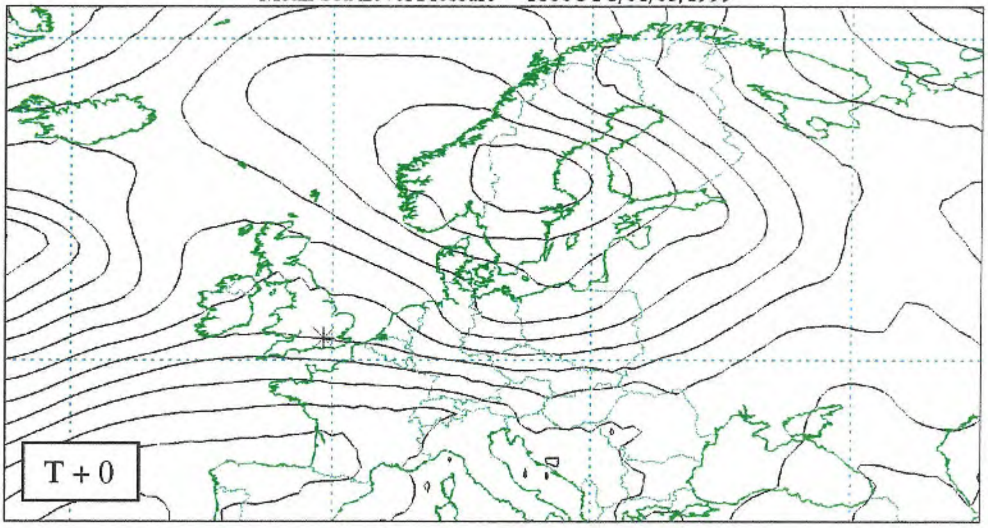
Details of the simulated releases.

The operational forecast and analysis member had increased input resolutions, but the NAME set up, release criteria and output resolution were all identical to that of the ensemble members.

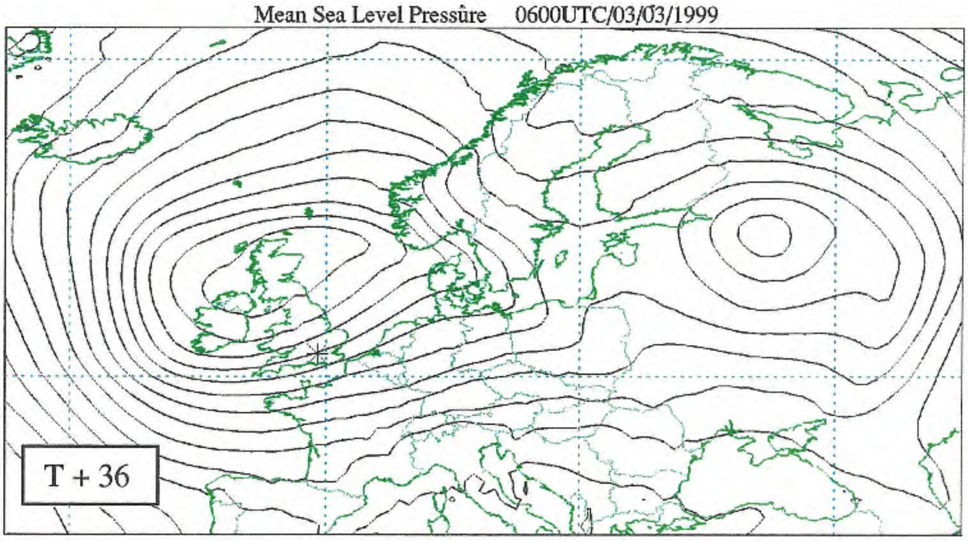
Release Details	**Ensemble Members**	**Operational Forecast**	**Analysis Member**
Resolution of Meteorological Data Input to NAME (Global)	$1.25^0 \times 0.83^0$	$0.83^0 \times 0.55^0$	$0.83^0 \times 0.55^0$
Resolution of NAME Output	$1.25^0 \times 0.83^0$		
NAME Version	4.3		
Start of Forecast	1200UTC 01/03/1999		
Start of Release	1800UTC 01/03/1999		
Duration of Release	6 hours		
Length of Plume Forecast	132 hours		
Release Height	0 – 50m		
Species Released	Inert Tracer		
Release Rate	10^{14} units h^{-1}		
Number of particles emitted to model the release	120,000		
Advection time step	900 seconds		
Output Concentrations Averaging Period	6 hours		
Output field	Mean Boundary Layer Air Concentration		

Analysed mean sea level pressure fields for T+0, T+36 and T+96 hours for the Stockholm release

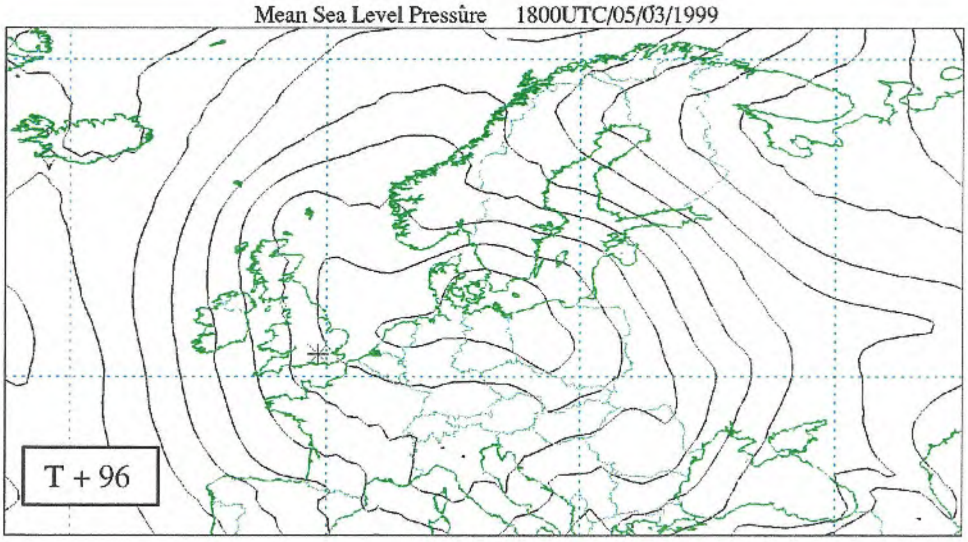
Mean Sea Level Pressure 1800UTC/01/03/1999



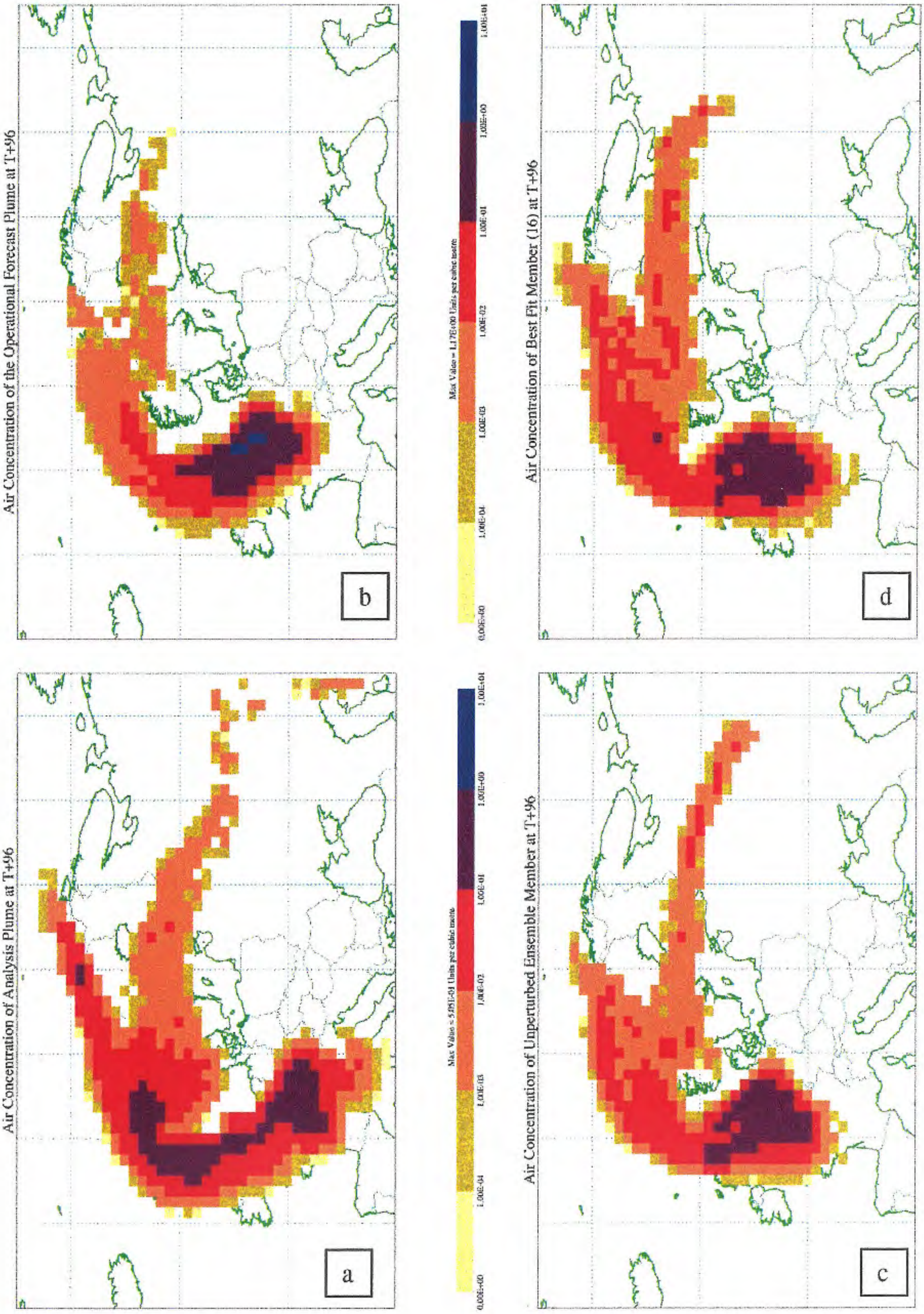
Mean Sea Level Pressure 0600UTC/03/03/1999



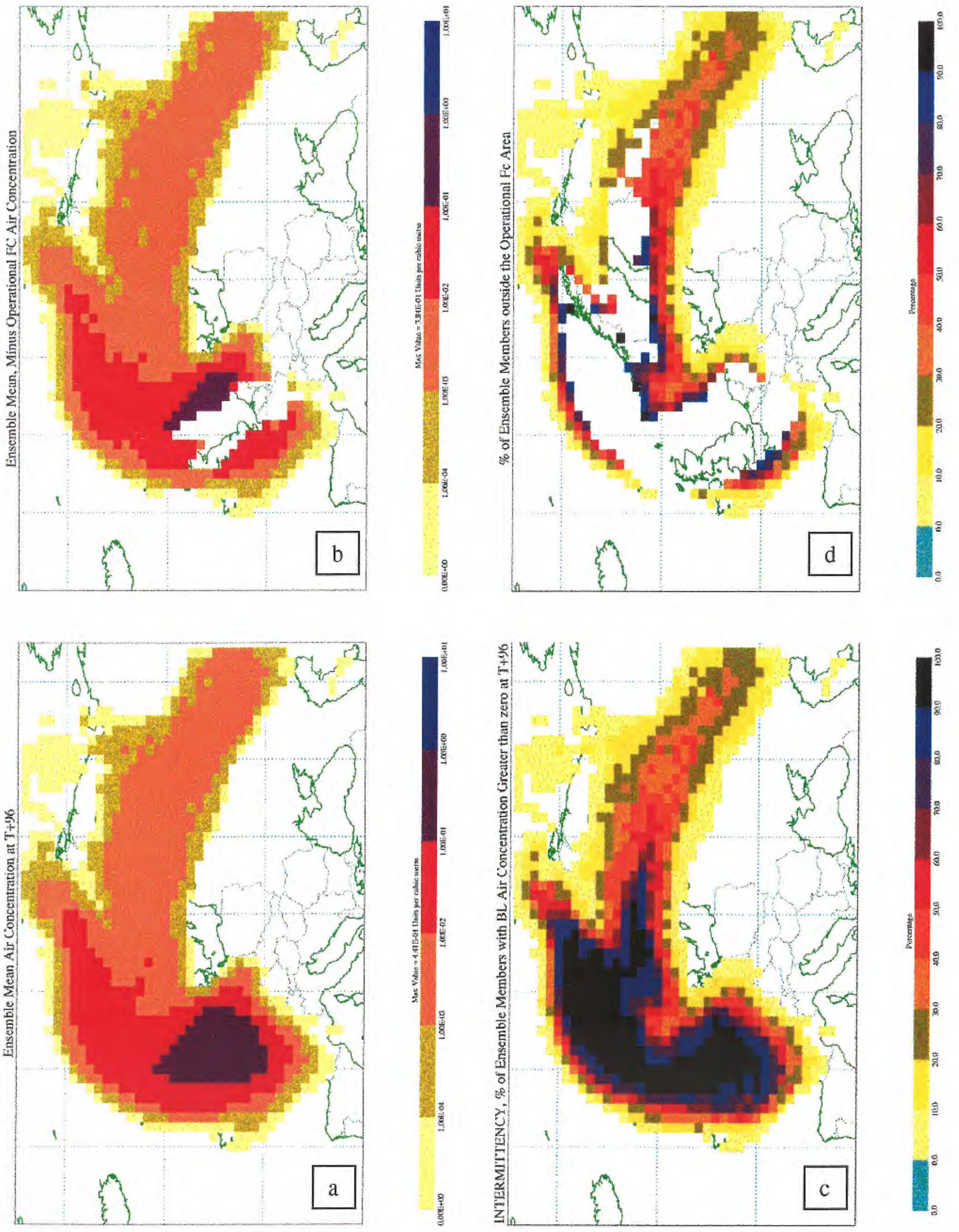
Mean Sea Level Pressure 1800UTC/05/03/1999



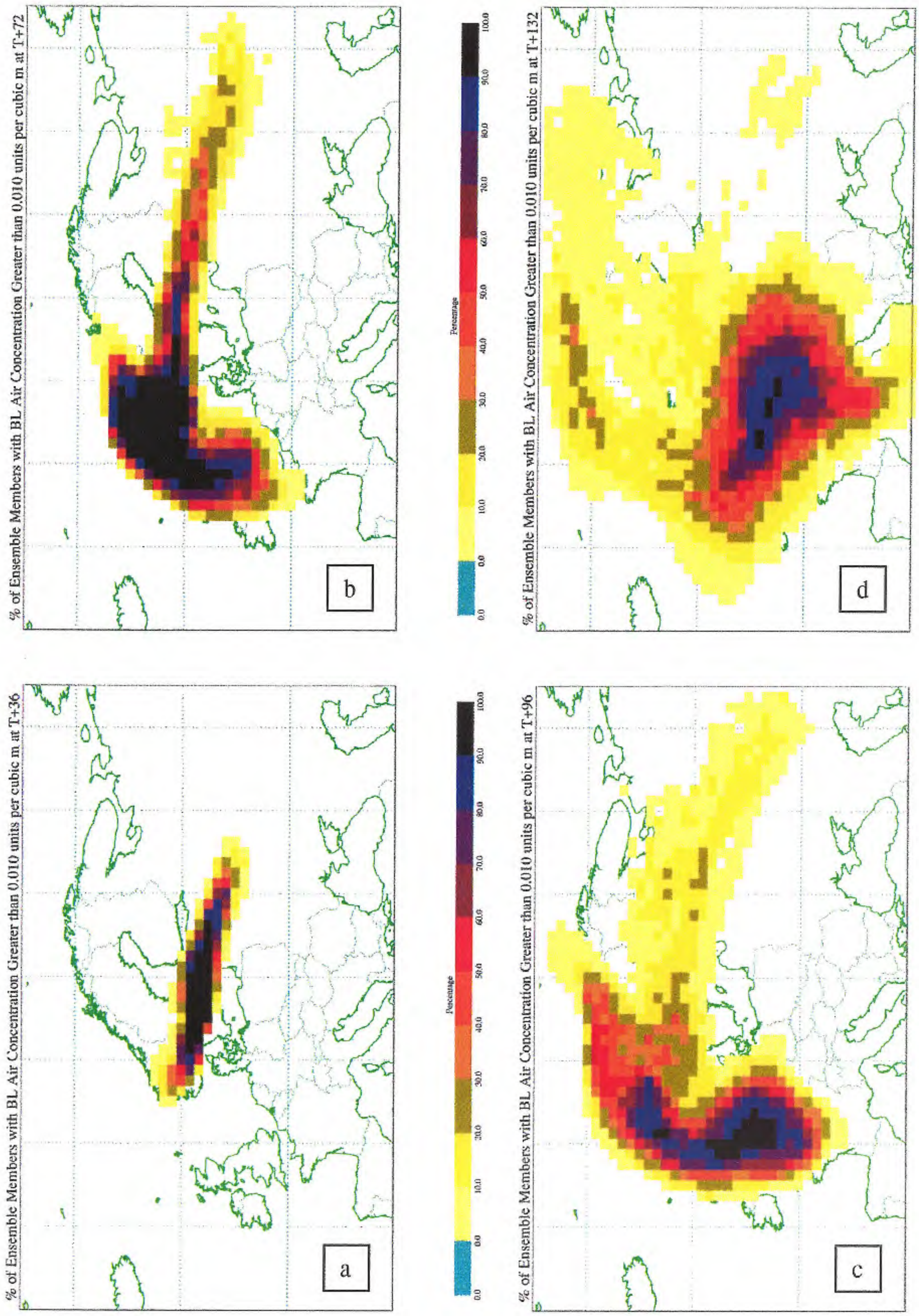
Air concentration of the inert tracer in the boundary layer at T+96 hours after the start of release from Stockholm. The analysis, operational forecast, unperturbed ensemble member and the 'best fit' member of the ensemble are shown.



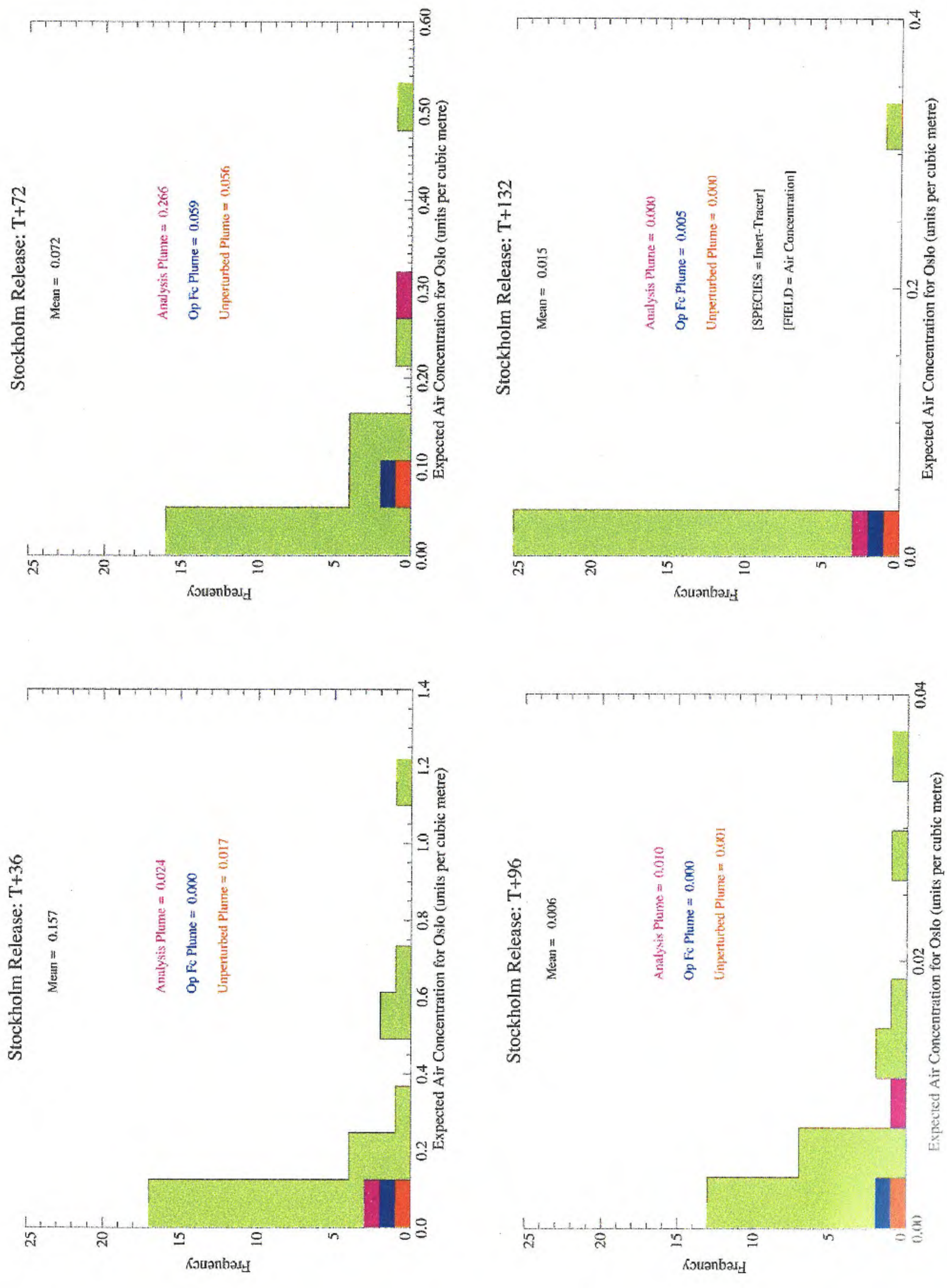
Operational analysis techniques for the Stockholm release after T+96 hours.



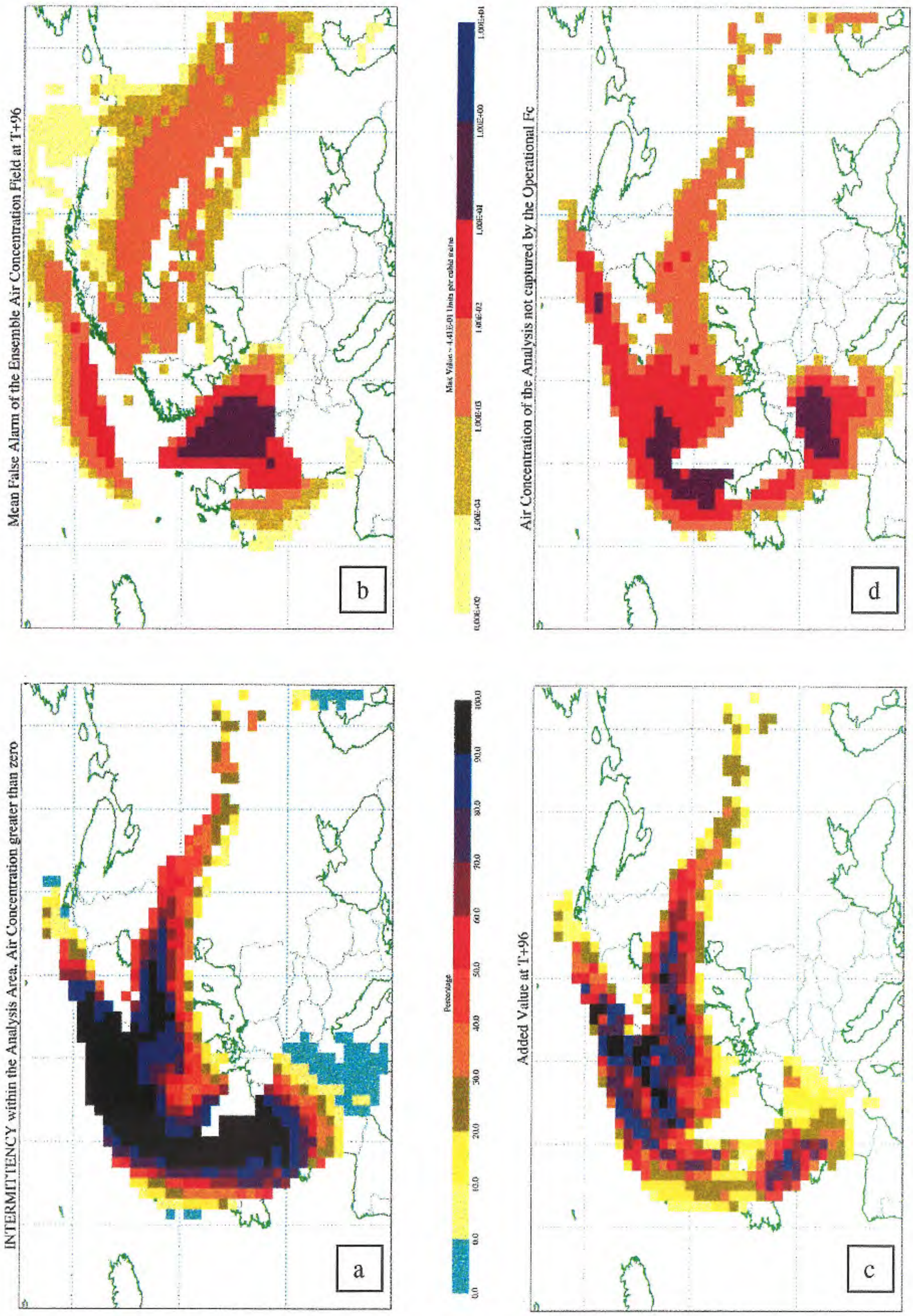
A times series showing the intermittency of ensemble members, with an imposed threshold of 0.01 units per cubic metre air concentration in the boundary layer. Stockholm release.



Four time frames of histograms for the expected air concentration in the boundary layer of the inert tracer at Oslo, following the Stockholm release.

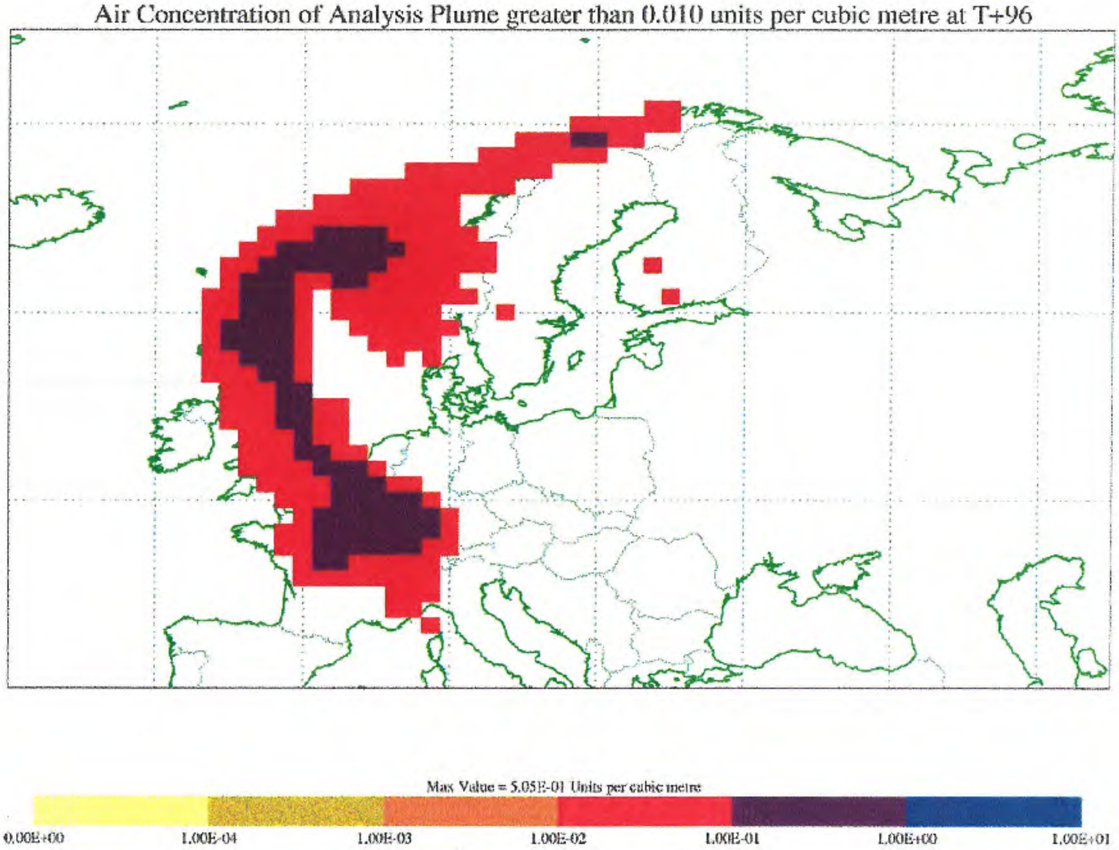


Retrospective analysis techniques for the Stockholm release at T+96 hours.

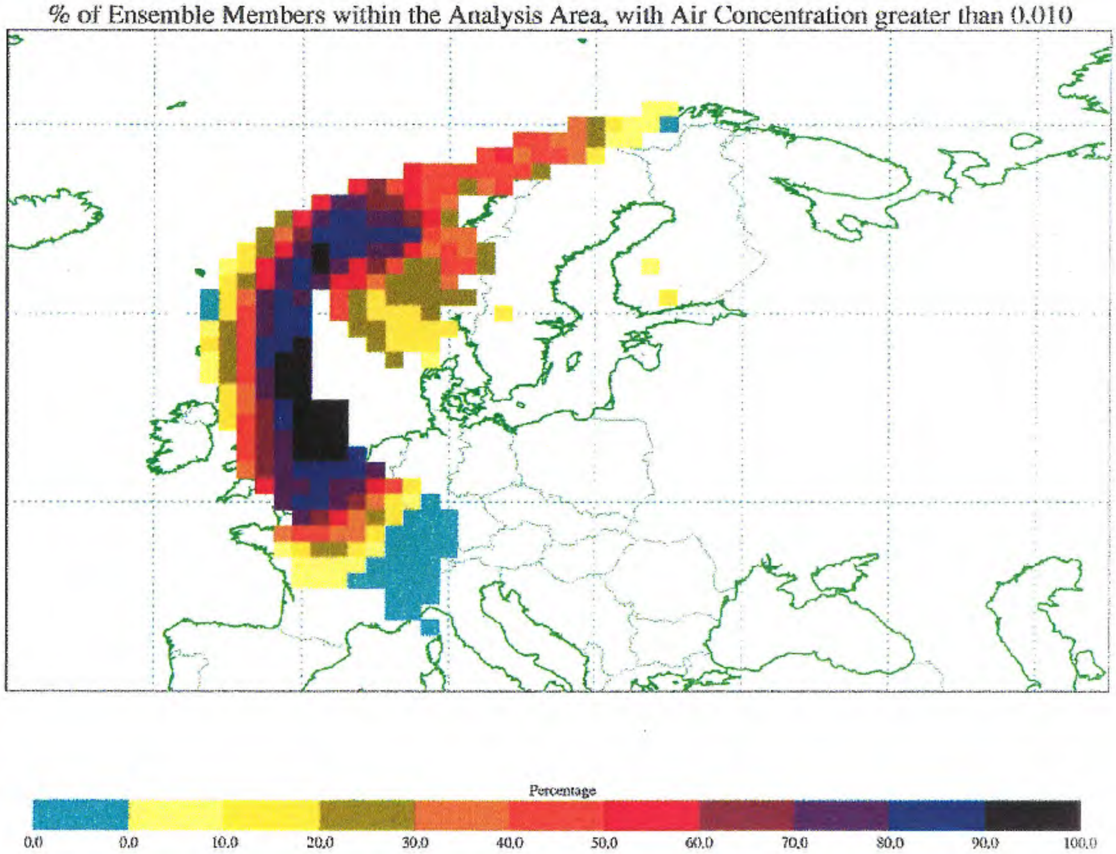


A comparison of the analysis plume at air concentrations in the boundary layer of greater than 0.01 units per m³ against the percentage of ensemble members present with concentrations greater than 0.01 units per cubic metre. Stockholm release at T+96 hours.

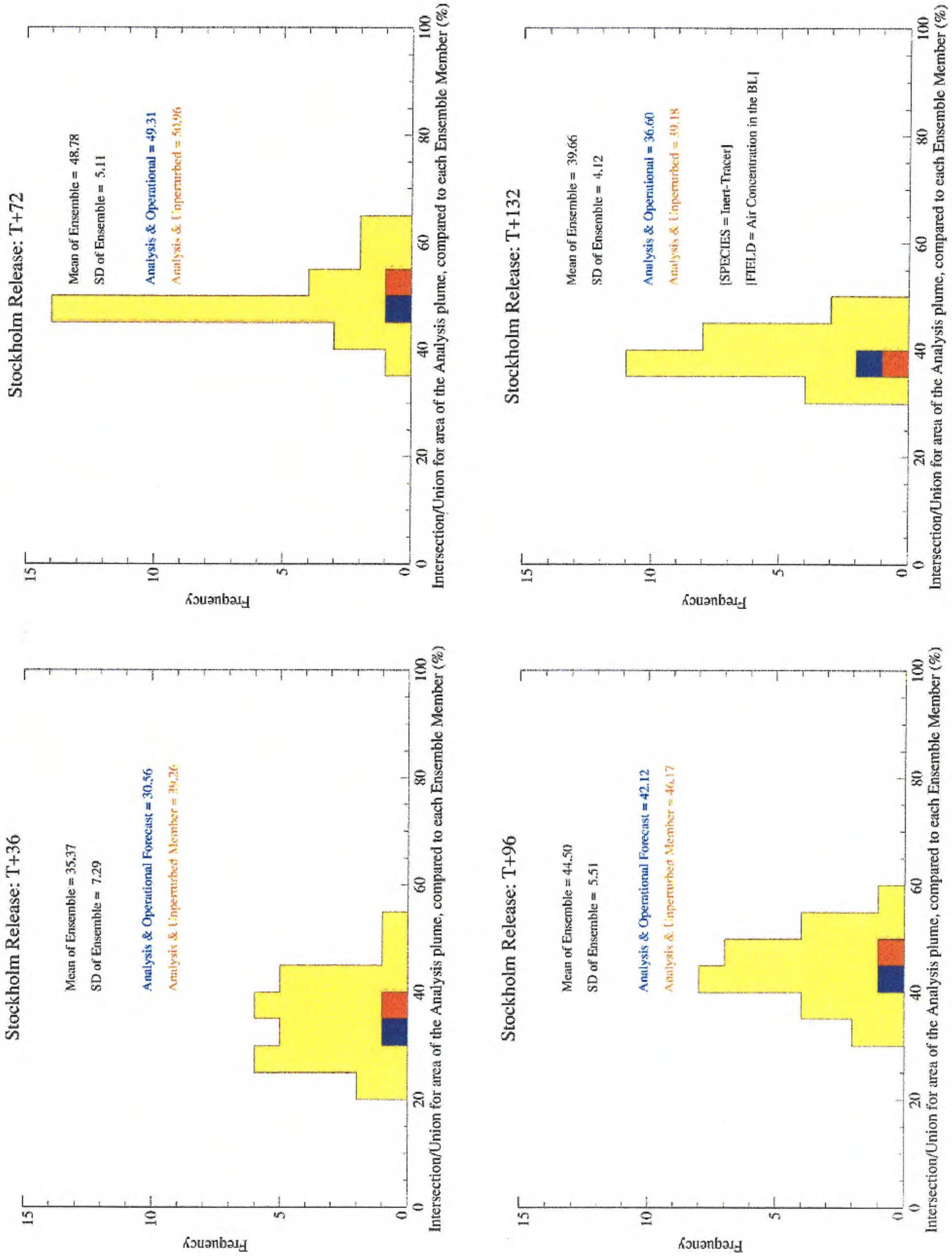
Air Concentration of Analysis Plume greater than 0.010 units per cubic metre at T+96



% of Ensemble Members within the Analysis Area, with Air Concentration greater than 0.010

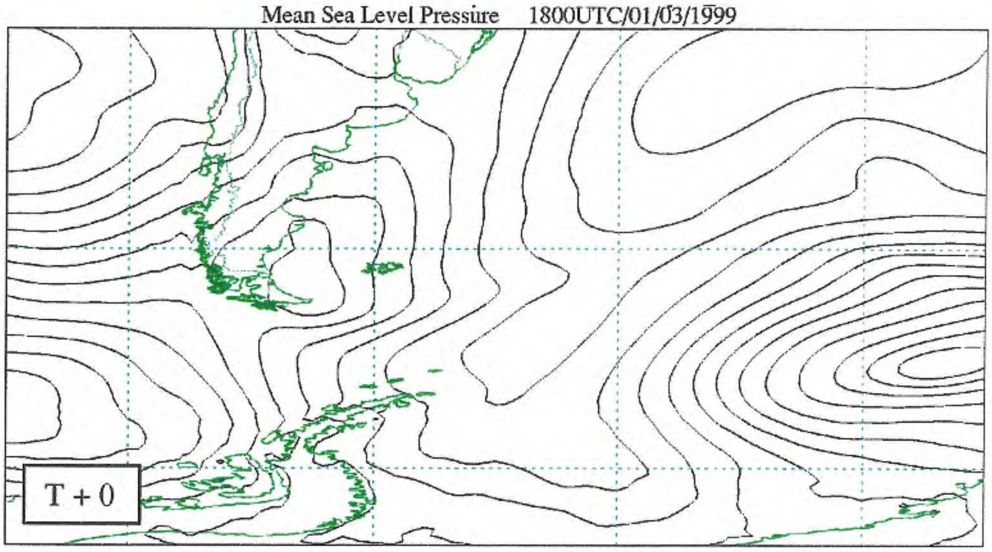


Four time frames of histograms showing the percentage overlap of all ensemble members and the operational forecast plume, in comparison with the analysis plume. Stockholm release.

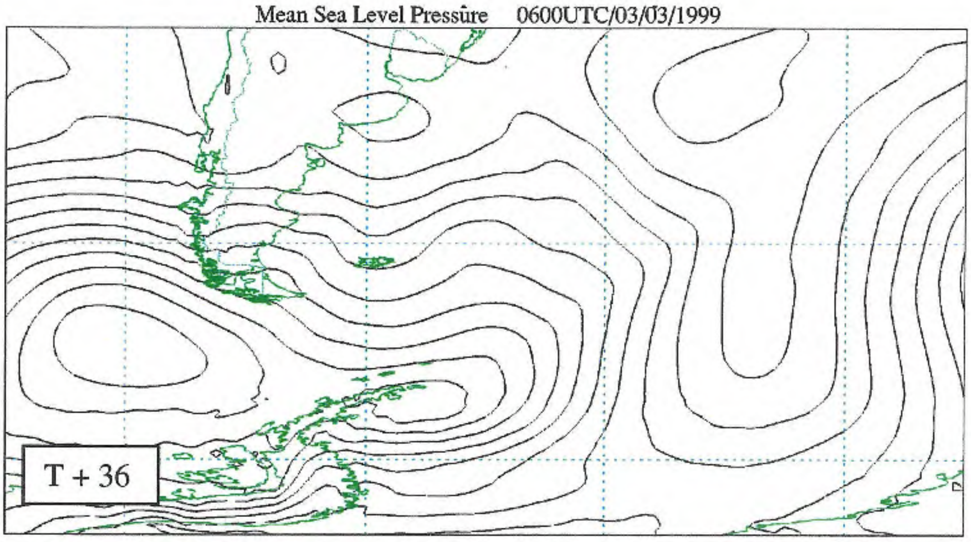


The analysed mean sea level pressure fields for T+0, T+36 and T+96 hours after the start of release from the Falklands.

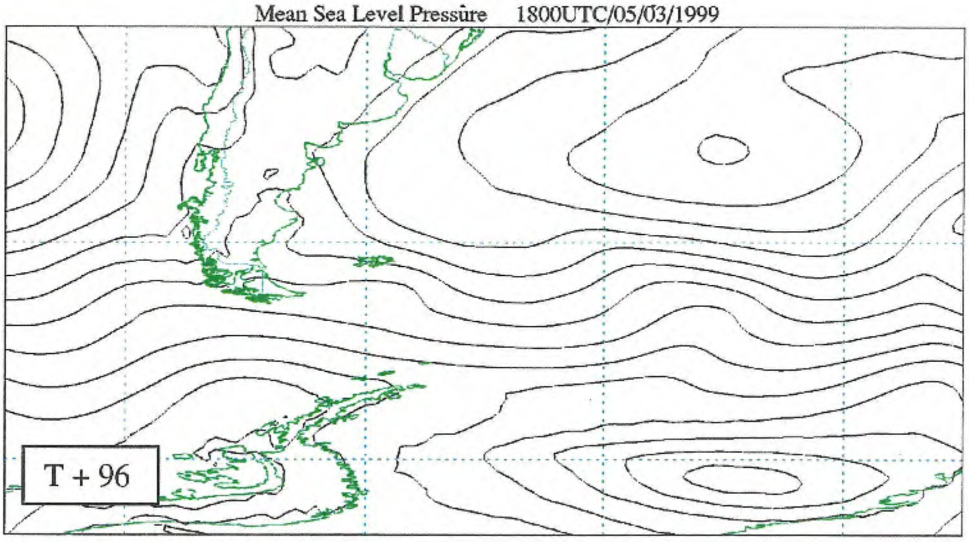
Mean Sea Level Pressure 1800UTC/01/03/1999



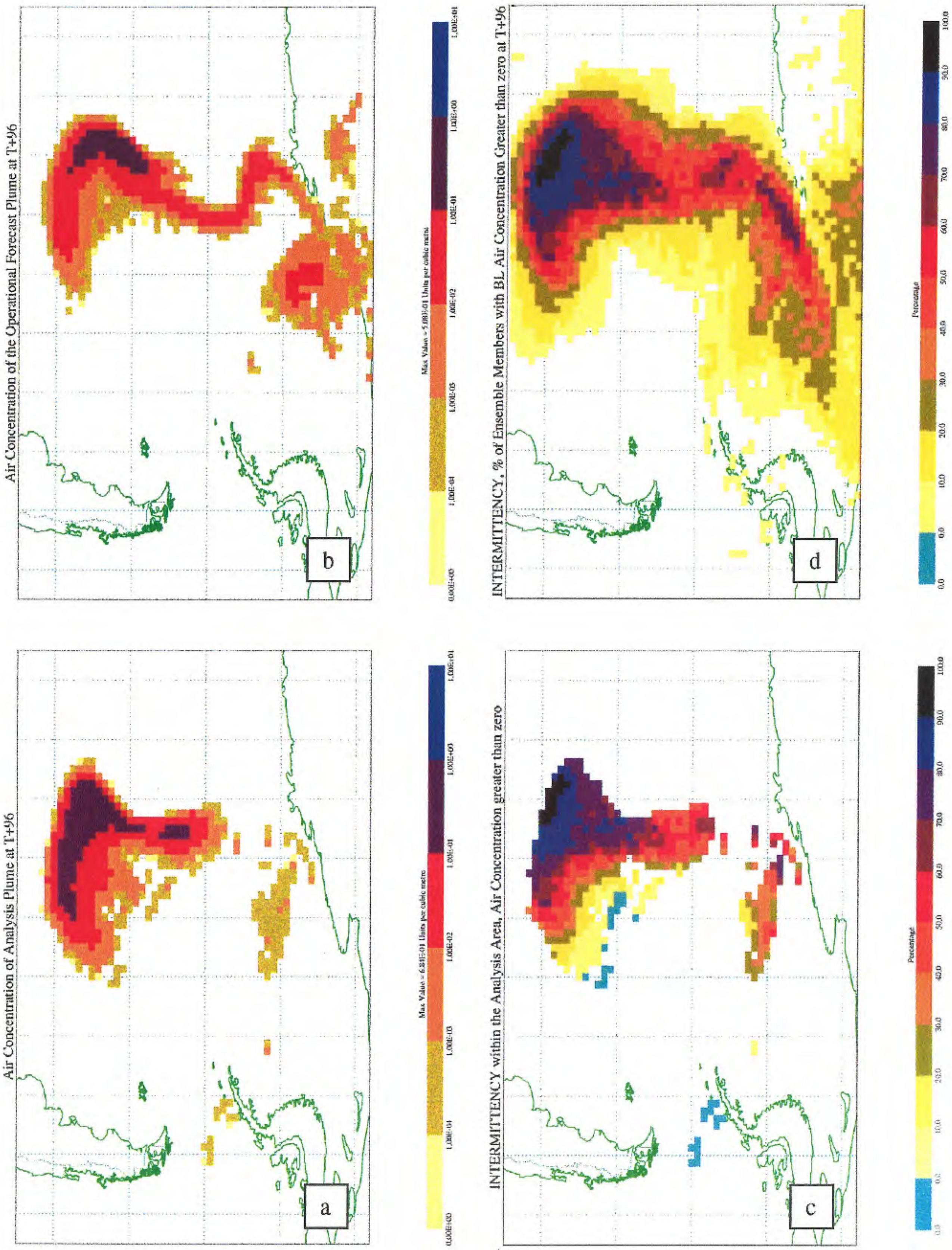
Mean Sea Level Pressure 0600UTC/03/03/1999



Mean Sea Level Pressure 1800UTC/05/03/1999

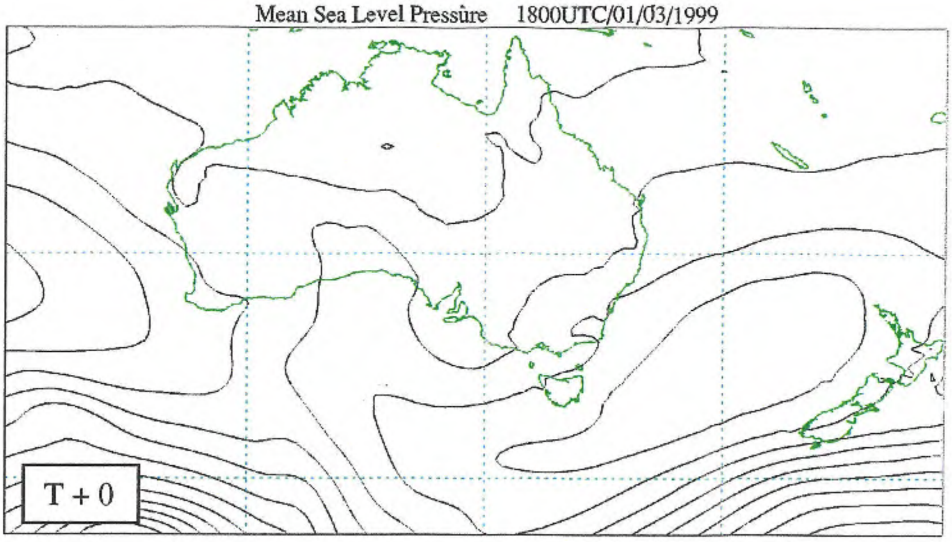


Analysis charts for the Falklands release at T+96 hours.

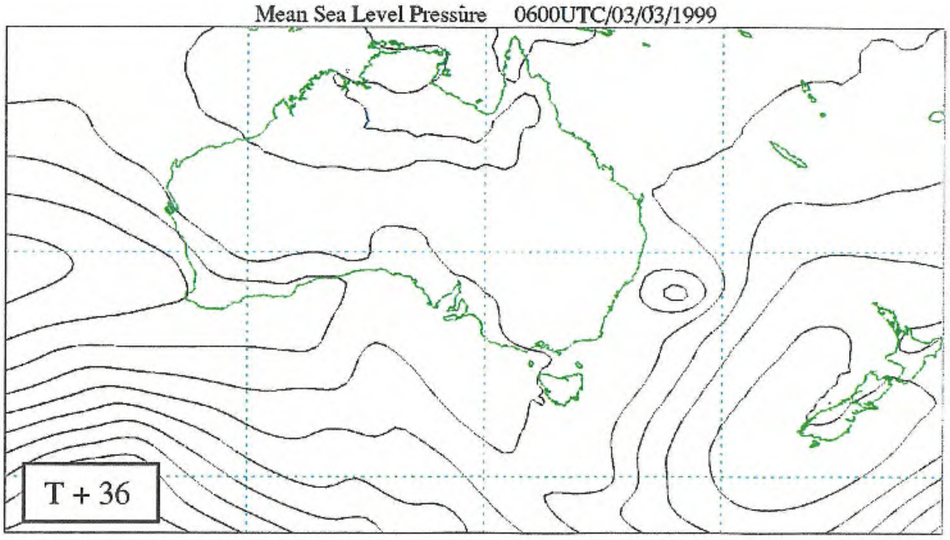


Analysed mean sea level pressure charts for T+0, T+36 and T+132 after the start of release for Sydney.

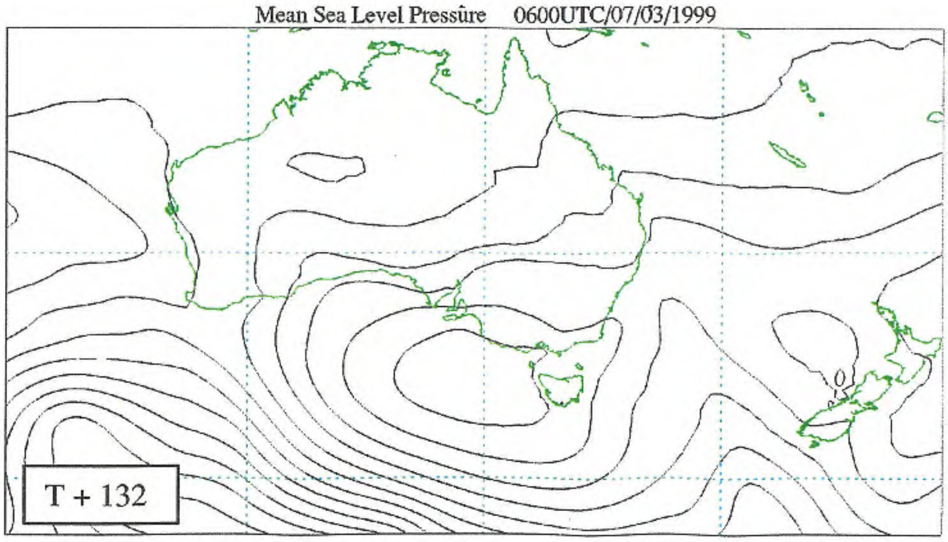
Mean Sea Level Pressûre 1800UTC/01/03/1999



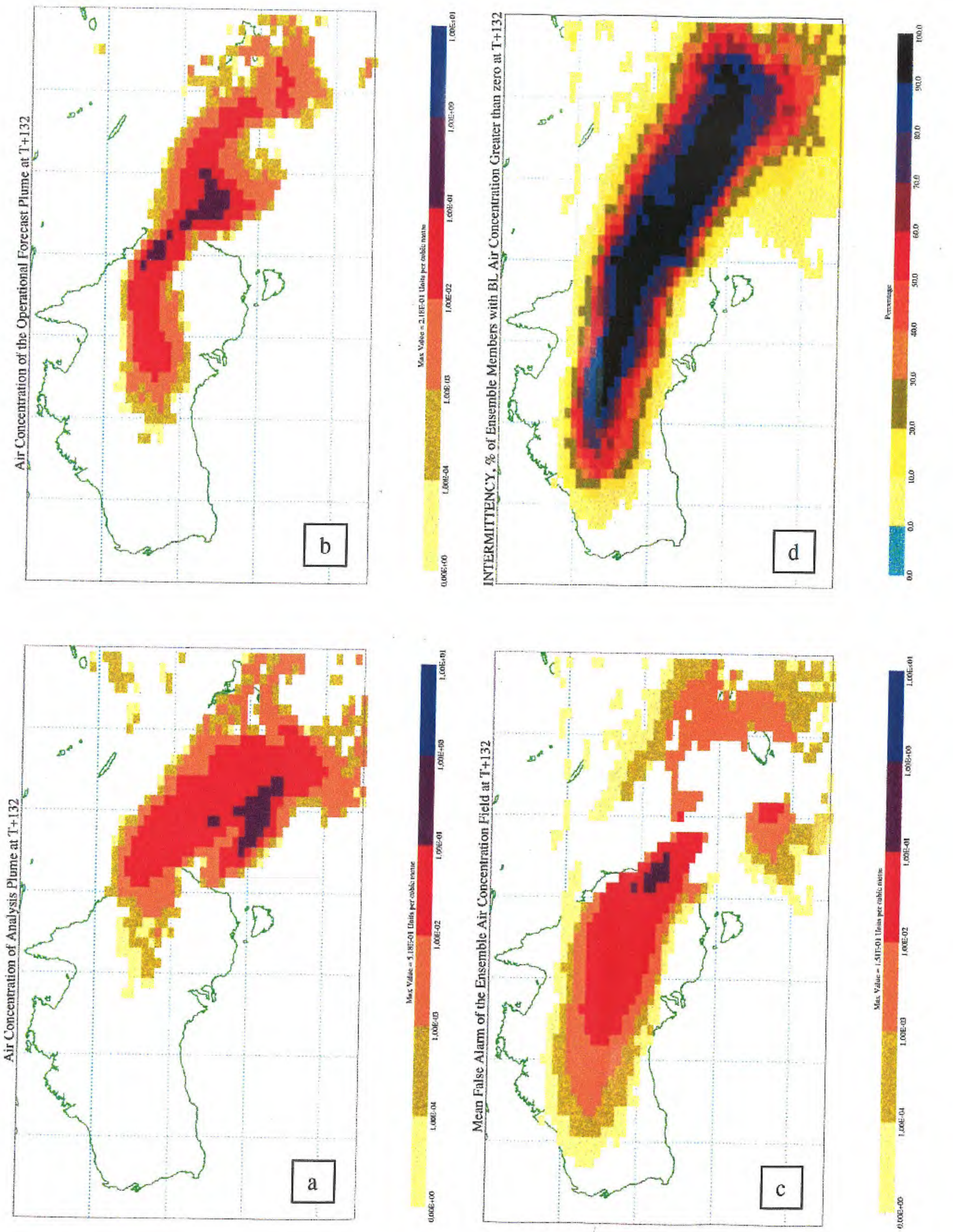
Mean Sea Level Pressûre 0600UTC/03/03/1999



Mean Sea Level Pressûre 0600UTC/07/03/1999



Analysis charts for the Sydney release at T+132 hours.



A time series for Brisbane of the expected air concentration in the boundary layer of the inert tracer, following the Sydney release. The first chart shown is T+48, when the tracer is first predicted to reach Brisbane.