NAME III Document
MD13/1(version 2)
08/04/05

Validation of NAME III (version 2.0) against the Kincaid data set

Andrew R Jones, Helen N Webster, Matthew C Hort and David J Thomson

1 The Kincaid experiment

The Kincaid field experiment was a very comprehensive experimental campaign carried out during 1980 and 1981 (see [1] for an overview of the Kincaid campaign). The experiment has been used to validate a number of dispersion models (such as HPDM [5], ADMS [2], AERMOD [2] and NAME [6]). The Kincaid power plant is located in Illinois, USA (39.59°N, 89.49°W) in an area of flat farmland with some lakes. The terrain has an approximate roughness length of 0.1 m and is at an elevation of approximately 180 m above mean sea level. During the campaign, a buoyant plume containing the tracer SF_6 was released from a power station stack of height 187 m and diameter 9 m. Arc-wise maxima of the hourly-averaged ground level concentrations were recorded along measurement arcs at distances ranging from 500 m to 50 km from the stack. A range of model validation tools are available for use with the Kincaid data set as part of the Model Validation Kit [4].

2 Met data for the validation study

During the Kincaid experiment, hourly met data was measured by a 100 m meteorological tower at the power plant site, augmented by estimates of boundary-layer depth derived from radiosonde ascents. All these basic meteorological data are included in the files presented as part of the Model Validation Kit [4]. The input met data used for our NAME III validation study consists of:

- wind speed and direction at 10 m and at 100 m[1],
- ambient air temperature at 10 m,
- cloud cover, and
- boundary-layer depth (optional use).

Note that observations of wind speed and direction at 30 m and 50 m are also available in the Model Validation Kit, but are not used here. Met data is available for each hour of the actual experiments, but is also supplied for hours prior to the start of

[1]The main validation exercise presented in Section 4.2 uses 100 m winds (following the approach adopted in previous NAME II and NAME III validation work); but we also repeat the runs using wind information at 10 m (emulating the approach taken in ADMS validation studies).

each trial. This extended data set provides modellers with an opportunity to use the history of the meteorology in determining boundary-layer parameters using their own met preprocessors. We have made use of this full data set here in modelling of the boundary-layer depth. A small number of the meteorological observations are missing from the data set (six hours of data are affected during the experiment itself, with four of these hours occurring consecutively on 09/05/80). We have estimated missing values by interpolating (in time) from the available data. Our use of interpolated data in the small number of missing cases would be unlikely to have a substantial effect on the validation results (except possibly for the case of 09/05/80 where the absence of observations is more severe) although, of course, the actual impact in these cases is impossible to assess here.

Some of the met variables are presented as hourly-averaged quantities referenced to the end of each hour (e.g. hourly-averaged data for hour 16 would be an average over the time period from 15:00 to 16:00). All times are quoted in Central Standard Time (CST) equivalent to GMT-6. As noted in our previous Kincaid validation reports, it is by no means obvious from the validation kit documentation as to which meteorological values are hourly averages and which are instantaneous spot values. Hence we have used 'hour' (meaning 'end of hour') in the meteorological input files to indicate the time for which the data is valid but acknowledge that, for time-averaged parameters, half an hour earlier could be more representative. However, provided that the meteorology is not changing too rapidly, the impact of any differences here should be small. It would appear reasonable to assume that the wind speed and direction data are time averages, whereas air temperature, cloud cover information and boundary-layer depths are probably spot values.

The boundary-layer depth is an observed value which was determined manually by interpretation of radiosonde data from on-site soundings. Previous NAME validation studies have used this observed value of boundary-layer depth, but the single-site met preprocessor in NAME III now gives us the option of enabling the model to determine its own value for the boundary-layer depth from other meteorological input. The primary validation results presented in this report (see Section 4.2) are based on boundary-layer depth as calculated by the met preprocessor (in other words, the observed values given in the Kincaid met files are *not* used in this instance). However, some parallel validation runs have been performed using the observed values for comparison, and these results are presented in Section 4.3. Other meteorological parameters such as the Monin-Obukhov length scale are derived using the single-site met preprocessor in NAME III. Again this is in contrast with our previous validation activities which used modelled values of the Monin-Obukhov length provided in the Kincaid data files. (However the Kincaid documentation warns that these suggested values of the derived parameters should be used with caution because of slight modifications subsequently made to the boundary-layer formulae in [3] which were originally used to derive these quantities. It instead recommends that modellers use their own preprocessing methods, but this was not possible with the NAME II model.)

3 The NAME III model set up

The validation study presented in this report assesses version 2.0 of the NAME III model (frozen on 07/10/2004). An example of one of the input files for the validation exercise is included in Appendix A. The example input file provides full information on the configuration of NAME III used in the validation study, but a further discussion on certain aspects of the model set-up may be useful here.

As in the previous NAME III validation study (for version 1.3), we model plume dispersion using puffs (in preference to particles), as the puff scheme is intended to be used for short-range applications. The two main benefits of using puffs rather than particles at short ranges are better computational speed (because we are following a relatively small number of puffs in comparison with the large number of particles that would be needed) and reduced levels of statistical noise. A further benefit of using puffs is that it enables us to calculate the concentration at a specific point in space (whereas some form of spatial averaging over a sample volume is necessary when particles are used). Thus we are able to calculate a ground-level concentration field that can be compared more precisely against the ground-level concentration measurements provided in the Kincaid data. Note that validation of NAME II considered modelled concentrations averaged over the layer from 0 m to 40 m above ground level, and the previous NAME III validation exercise attempted to emulate the original approach by calculating concentration at a height of 20 m. However for an elevated stack release, any differences between the concentrations at ground level and at an elevation of 20 m are likely to be small in most situations (especially when hourly-averaged concentrations are considered).

Ground-level concentrations are calculated on a horizontal (latitude-longitude) grid of resolution $\sim$ 0.5 km centred on the stack location and extending 51.0 km in each of the N-S and E-W directions. The output grid was defined here to agree with the grids which were used in previous validation work (although we have calculated a more precise grid definition that will, in reality, give very slightly different grid points). It is worth noting that the validation exercise actually requires concentrations along arcs at various downwind distances from the stack. Currently these are derived by a post-processing of the lat-long concentration fields but it is recognised that direct output on a polar coordinate system centred on the stack would enable this extra processing step to be avoided. NAME III can define horizontal grids based on polar coordinate systems but the full functionality to enable their use for output has not yet been developed. For future consideration, it would be worth developing NAME III to realise this potential for polar-coordinates output.

The tracer measurements in the Kincaid data set are hourly-average concentrations observed over the preceeding hour (that is, the validity time refers to the end of the hour). Thus we wish to calculate an hourly-average concentration field for each hour. At the time of the previous NAME III validation study, the model was unable to produce time-averaged concentration predictions and so instantaneous concentrations at the half hour were adopted as a proxy for the hourly averages. Incidentally, the earlier validation of NAME II was able to consider the correct hourly-averaged quantities. The NAME III model now has the ability to calculate time-averaged concentration fields. For comparison with the field measurements (and following the approach used in NAME II), we calculate an hourly-average concentration for each hour by averaging the instantaneous concentrations at four individual times within that hour. Specifically, the time-averaged concentration is taken to be the average of the values at T-00:45, T-00:30, T-00:15 and T-00:00.

The plume emitted from the Kincaid stack was buoyant; this requires the modelling of plume rise to be included in the NAME III runs. The plume rise calculations are based on information about the emission temperature and emission velocity from the stack over each hour (included as part of the Kincaid data set). The experiment involved the release of the passive tracer SF_6 into the plume. We model this in our NAME III runs as an inert tracer (that is, there are no deposition or decay processes acting on the plume). The effects of turbulence and meander are both modelled. The random-walk model for the plume dispersion includes velocity memory for all travel times. The inhomogeneous turbulence scheme is used within the boundary layer with turbulence profiles (velocity variance and Lagrangian timescales) being a function of height. The present validation exercise uses a puff interval (i.e. the time between successive puff releases) of 1 minute, in contrast with the longer puff interval of 5 minutes considered in the previous work on NAME III (this earlier work also experimented with puff intervals extending from 2 minutes up to an hour). The shorter puff interval is believed to give better overall results (less patchy plumes) and does not appear to be significantly more expensive in terms of run times. A synchronisation interval of 5 minutes is applied to each model run.

One of the 'global parameters' was modified in the NAME III source code to perform the validation study: `MaxPuffs` was reset to 750,000. This change was applied to ensure that the model had enough memory available to store the large number of puffs required in some of the simulations (such as in those cases where the boundary layer collapses at sunset and there is a rapid increase in the number of puffs needed to adequately represent the plume). However, the increased value of `MaxPuffs` remains insufficient to adequately cope with the modelling requirements of all cases, and some of the model runs eventually failed due to the demand for too many puffs. These failures occurred only when NAME III determined its own boundary-layer depths, and, although we could have re-run the simulations with a larger value of `MaxPuffs`, we have simply fixed up any missing concentrations resulting from the failed runs with the corresponding predictions of the runs using observed boundary-layer depths.

In the main validation exercise presented in Section 4.2, the NAME III runs failed during the final hour on 11/07/80 and during the penultimate hour on 13/07/80, with all other model runs successfully completed.[2] The parallel validation runs considered in Section 4.3.3 use an increased A1 parameter (A1 = 200 rather than the default value A1 = 50) and failed on four individual cases (01/05/80, 11/07/80, 13/07/80 and 25/07/80). The increased frequency of failures is not surprising here, since the increased value of the A1 parameter would be expected to lead to an approximate four-fold increase in the number of computational puffs that are used to resolve the plume in each run (and thus we would expect the puff number to reach its limiting value sooner than in the default case). Any future validation work of NAME III should consider adopting a larger value of `MaxPuffs` (we suggest that two million puffs might be suitable) with

[2]These failed runs result in one missing value in our set of NAME III predictions of arc-wise maximum concentration for quality 3 data, and six missing values for quality 2 or 3 data.

the aim of avoiding a repeat of the modelling difficulties encountered in the present validation study.

Several modifications to the puff scheme have occurred since the previous validation exercise using version 1.3, with the intention of resolving the 'blobby plume' issue highlighted in that earlier validation report. For instance, the puff output calculation has been revised to spread concentration along the distance traversed by the puff in a time step, with a new puff parameter, A7, being introduced to control whether to use this 'error function formula' or a Gaussian approximation of this value in the output calculation (based on the distance moved by the puff over the time step). Two other parameters of the puff scheme have also been changed, with A1 being reduced from 200.0 to 50.0 (effectively reducing the number of puffs created via the puff splitting process), and A5 being reduced from 4.0 to 3.0 (so that less of the Gaussian tail of each puff is included in the concentration calculation). We also identified two bugs in the puff model which have subsequently been corrected. These modifications to the puff scheme have largely resolved the 'blobby plume' problem evident in the NAME III runs performed as part of the previous Kincaid validation exercise, although it is worth mentioning that the plume is still slightly patchy in a few of the new cases (once again, these occur in stable conditions and often when the plume travels within the free troposphere before mixing down to ground level). Some examples are given in Appendix B.

4 Validation results

We present here the results of our comparison of arc-wise maximum concentrations predicted by NAME III against the observed values that are provided in the Kincaid data set. Section 4.1 describes the procedures used to analyse the data, and then the results are discussed in Sections 4.2 and 4.3.

4.1 Overview of analysis procedure

Before we analyse the results, the set of arc-wise maximum concentrations (observed and predicted) are filtered to remove those arcs where the modelled plume has not had sufficient time to reach the detectors. The explanation for this step lies in a strange discrepancy of the data provided in the Kincaid files. On each day, the tracer release commenced some time before the start of the experiment (presumably to allow the plume to reach the various measurement arcs) and, although meteorological information is available for these pre-experiment hours, the source information (release rate and other stack emission properties) is not provided. Thus we have to start the model run at the beginning of the first measurement hour, and then allow time for transport of the tracer to the detector arcs. In general, we should therefore expect the NAME III predictions of hourly-average concentration to be initially lower than the observations (at least, in those cases where the plume has not reached the monitoring arc by the first contribution time of the hourly average (15 minutes here)). We resolve this issue by applying a filter to remove any detector arcs which are likely to be affected. The filtering used here is identical to that applied in previous NAME validation studies and, in fact, is implicitly performed during the construction of the input files for the analysis packages.

Two model analysis packages, BOOTW and RESIDUW, and the plotting package SIGPLOT are distributed as part of the Model Validation Kit. The BOOTW package assesses the performance of the model in a statistical manner. Performance measures calculated by the BOOTW program are the mean ($\overline{c}$), standard deviation (σ_c), bias, normalised mean square error (NMSE), correlation (r), fractional bias (FB), fractional standard deviation (FS) and proportion of values within a factor of two (FA2) of the observed concentrations. These performance measures are calculated using the following formulae:

$$
\begin{aligned}
\text{BIAS} &= \overline{c_o} - \overline{c_p}, \\
\text{NMSE} &= \frac{\overline{(c_o - c_p)^2}}{\overline{c_o}\,\overline{c_p}}, \\
r &= \frac{\overline{(c_o - \overline{c_o})\,(c_p - \overline{c_p})}}{\sigma_{c_o}\,\sigma_{c_p}}, \\
\text{FB} &= \frac{\overline{c_o} - \overline{c_p}}{0.5\,(\overline{c_o} + \overline{c_p})}, \\
\text{FS} &= \frac{\sigma_{c_o} - \sigma_{c_p}}{0.5\,(\sigma_{c_o} + \sigma_{c_p})}, \\
\text{FA2} &= \text{fraction of data for which } 0.5 \le \frac{c_p}{c_o} \le 2,
\end{aligned}
$$

where c_o is the observed arc-wise maximum concentration and c_p is the predicted arc-wise maximum concentration (and similarly for the standard deviations). The overbar denotes the average over all hours and all distances (i.e. downwind arcs). Note that a negative bias denotes over-prediction and a positive bias denotes under-prediction. The BOOTW program uses the blocked bootstrap resampling method to obtain confidence intervals on the model statistics.

The ratio of (arc-wise maximum) predicted concentrations to observed concentrations, `C MOD/C OBS`, known as a residual, should not demonstrate any dependence on variables such as downwind distance, stability and wind speed in a good model. The RESIDUW package analyses the residuals to give a scientific evaluation of the model.

The plotting package SIGPLOT displays various measures of the model performance as scatter plots, quantile-quantile plots and box plots (although we have used our own PV-WAVE plotting routines to generate scatter plots and quantile-quantile plots in a more flexible format for this report). A scatter diagram of observed versus predicted concentrations together with quantile-quantile plots and other statistical measures is essential information in model assessment. A quantile-quantile plot shows the same information as a scatter diagram but with model predictions and observed concentrations ordered in magnitude. That is, the largest observed concentration is plotted against the largest model prediction regardless of whether they occurred at the same time, same distance from source, etc. Quantile-quantile plots show whether a model has a general tendency to under-predict or to over-predict. A collection of box plots display the behaviour of the residuals with respect to various parameters. An example is shown in Figure 3 below. Here box plots have been produced with the data grouped according to a number of physical and meteorological variables, namely:

- distance from source,
- friction velocity u_*,
- stability (as characterised by z_i/L, where z_i is the boundary layer depth and L is the modelled value of the Monin-Obukhov length),
- boundary layer depth z_i,
- time of day, and
- quality index of the observational data (as defined in [4]).

The cumulative distribution function of the residuals (`C MOD/C OBS`) within each group is expressed here as the 5th, 25th, 50th, 75th and 95th percentiles, and these five significant points are plotted in a box format. For a good model, these residual boxes should be small and should not deviate too much from unity (the solid central horizontal lines in these box plots). The dashed horizontal lines define the boundaries of the region within which modelled values lie within a factor of two of observations. A lower threshold of 0.01 and an upper threshold of 100.0 have been imposed on the residuals `C MOD/C OBS`. A filter to reduce noise can be implemented when both observed and predicted values fall below a certain limit. In our case, no filter was used (although it may be worth reviewing this decision in the future, for instance, by discarding any concentrations below the response threshold of the instrumentation).

4.2 Discussion of the main validation exercise

The Kincaid data does not fully constrain the configuration of the NAME III model in this validation study (for example, wind information is available at a number of measurement heights in the data set). Instead, there is an element of flexibility in how we choose to set up the model runs to use the data that is provided. Our primary validation exercise uses the wind speed and direction at 100 m, air temperature at 10 m and cloud cover information. It does not use the observed boundary-layer depth provided in the Kincaid data files, but uses the met preprocessor in NAME III to determine its own boundary-layer depth. The results of the main validation study using this preferred model set up are now discussed in this section.

Table 1 shows the statistics generated for observational data of quality 3 (as defined in [4]). Results from the earlier validation of NAME II are also included for comparison.

	$\overline{c}$ μg m^{-3}	σ_c μg m^{-3}	bias μg m^{-3}	NMSE	r	FB	FS	FA2
Observations	0.69	0.51	–	–	–	–	–	–
NAME III	0.74	0.52	-0.05	0.56	0.473	-0.072	-0.022	0.758
NAME II	0.58	0.58	0.11	1.07	0.306	0.180	-0.113	0.667

Table 1: Model performance statistics for quality 3 data

The NAME III predictions compare satisfactorily with observations and are a significant improvement on NAME II. Furthermore, the performance of NAME III is comparable to other leading short-range dispersion models[3] (as discussed by, for example, [2] and [5]). The results show a small over-prediction of the mean concentration with NAME III (in contrast to the under-prediction experienced with NAME II). Spread in the predicted concentrations of NAME III is in good agreement with the observed spread. The normalised mean square error has been substantially reduced in the new model. Similarly, both correlation and the proportion of modelled concentrations within a factor of two of observed values have been improved in NAME III. The improvements seen in all of these statistical measures are partly a consequence of the new model formulation (i.e. using a puff scheme, etc.) but also partly due to inclusion of the meteorological pre-processor (which allows NAME III to determine its own values of boundary layer depth), see Section 4.3 for further details.

Figure 1 shows the corresponding scatter plot for data of quality 3. Observations are plotted against model predictions and the one-to-one line denoting perfect agreement between modelled and observed concentrations is included here for reference. A quantile-quantile plot for the same data is given in Figure 2.

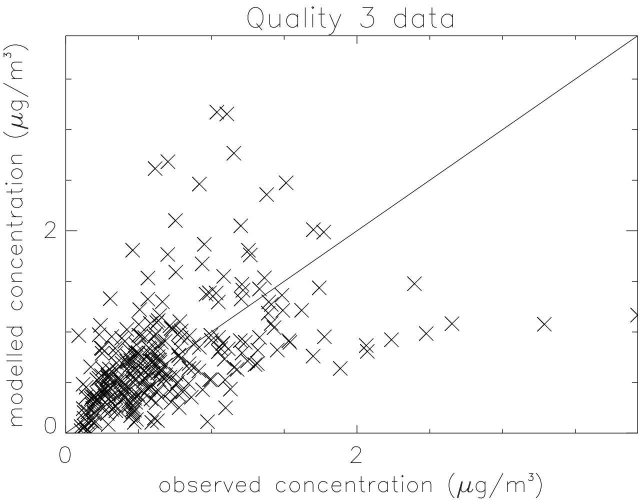

Figure 1: Scatter plot for data of quality 3

The behaviour of the ratio of modelled concentrations to observed concentrations for observational data of quality 3 is displayed in Figure 3. The residual `C MOD/C OBS` decreases with increasing downwind distance so there is a tendency to over-predict near source and under-predict further downwind. In cases with shallow boundary layers, concentrations are under-predicted. This is probably due in some part to instances when the lower part of the puff extends into the boundary layer but is not mixed down

[3]A direct comparison of NAME III with these other models is not possible in the current report because of differences in the statistical analysis procedures (our analysis of the NAME III predictions does not consider the source-strength scaling of concentrations that is used elsewhere), although we believe that such scaling does not strongly influence the validation statistics.

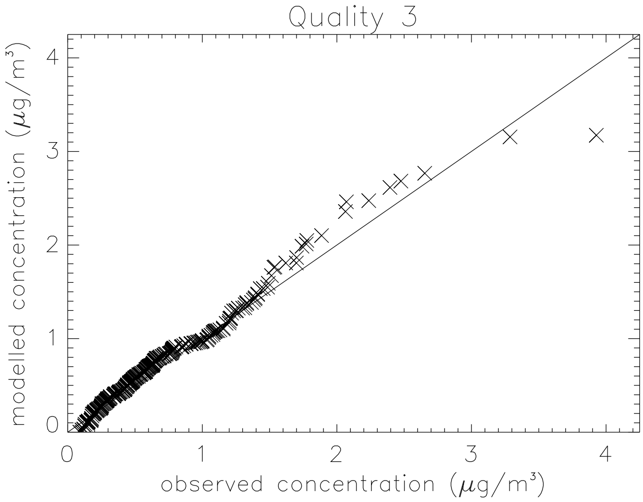

Figure 2: Quantile-quantile plot for data of quality 3

to the ground since turbulence levels are calculated at the puff mean height within the free troposphere.

Table 2 presents the statistics generated by the BOOTW program for observational data of either quality 2 or 3. NAME III over-predicts concentrations (more so than NAME II, in fact) although this is to be expected to some extent because, for the observations of quality 2, the true maximum values are believed to be greater than those observed. As with the quality 3 results, the statistics for NAME III again demonstrate a significant improvement over those for NAME II.

	$\overline{c}$ μg m^{-3}	σ_c μg m^{-3}	bias μg m^{-3}	NMSE	r	FB	FS	FA2
Observations	0.53	0.50	–	–	–	–	–	–
NAME III	0.65	0.52	-0.12	0.86	0.455	-0.197	-0.047	0.649
NAME II	0.61	0.67	-0.07	1.82	0.160	-0.127	-0.284	0.579

Table 2: Model performance statistics for quality 2 and 3 data

Figure 4 shows the corresponding scatter plot of observations versus model predictions. A quantile-quantile plot for data of quality 2 and 3 is given in Figure 5.

Box plots displaying the behaviour of the residuals, `C MOD/C OBS`, for data of quality 2 and 3 are shown in Figure 6. Comparing Figures 3 and 6, we see that the residual boxes are much larger for data of quality 2 and 3, and extend further outside of the factor of two range. At small distances near to the source, NAME III over-predicts against observations of quality 2 and 3. At low wind speeds ($u_* < 0.2$ ms^{-1}), there is some evidence of under-prediction of concentrations. However, since there are only five observations in this category, all of which are of quality 2, it is debatable whether this is a reliable signal. There is under-prediction of concentrations in stable conditions and,

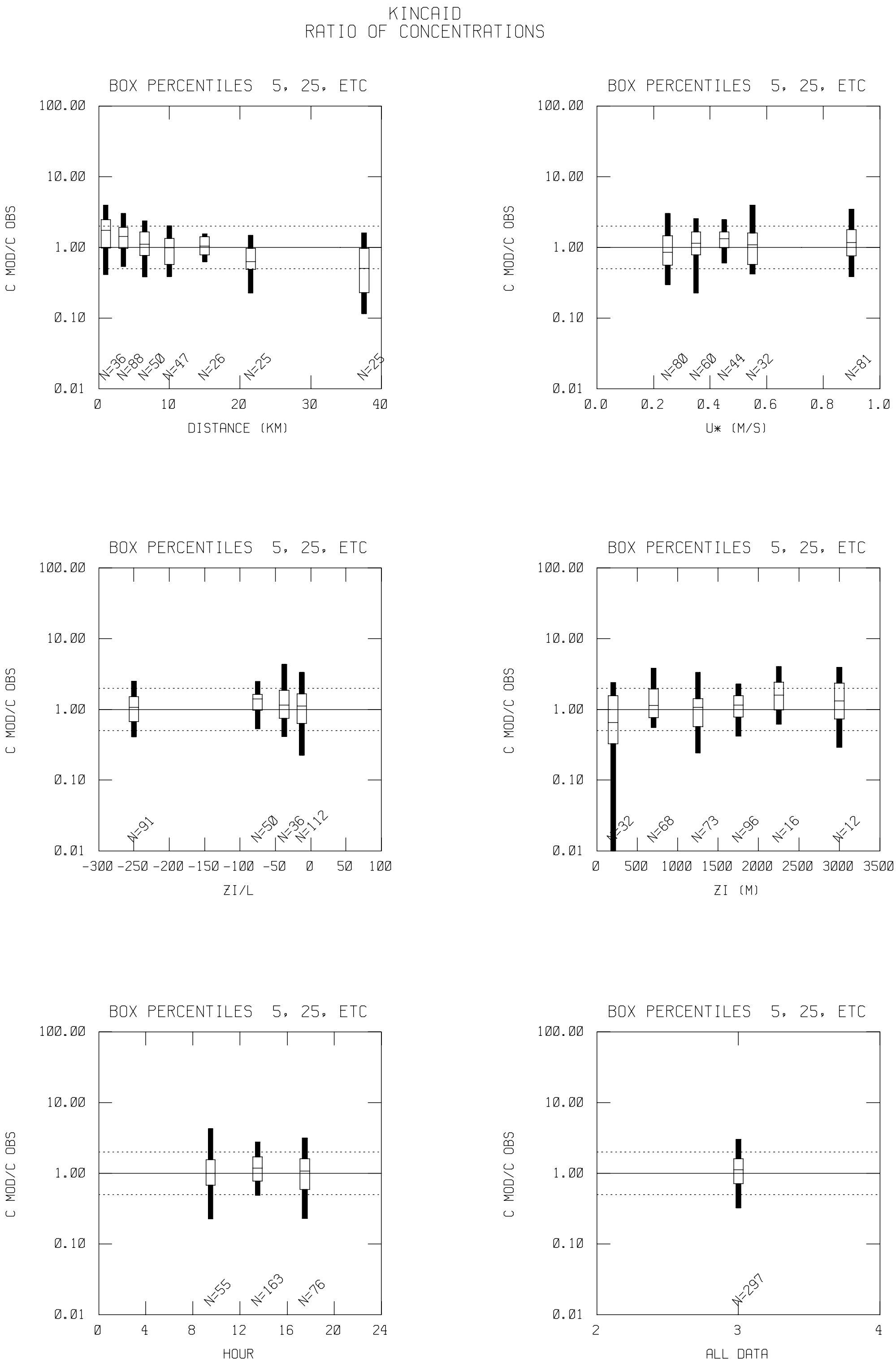

Figure 3: Residual box plots using quality 3 data

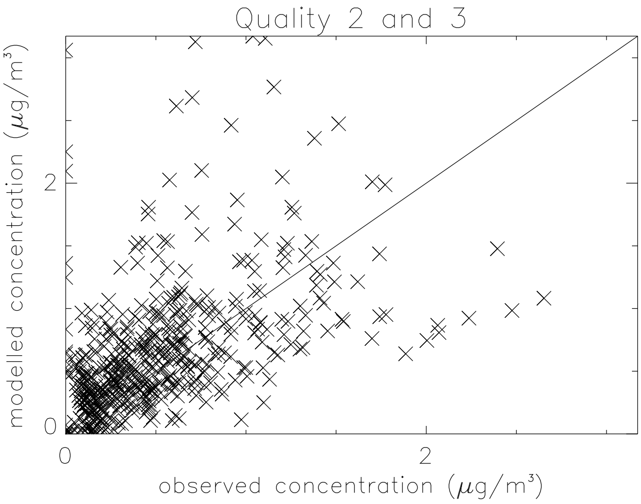

Figure 4: Scatter plot for data of quality 2 and 3

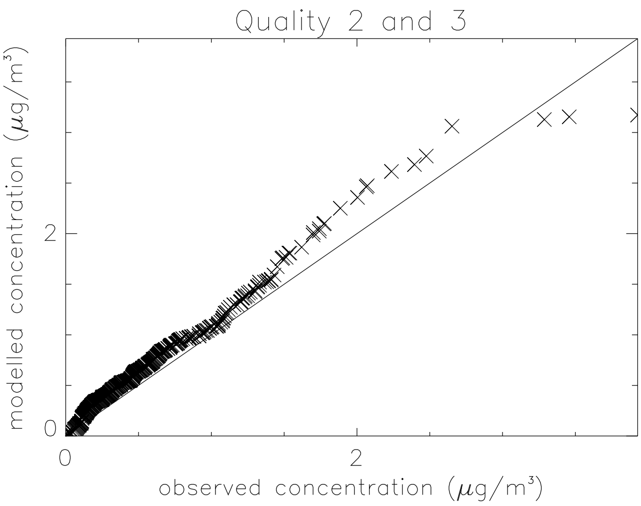

Figure 5: Quantile-quantile plot for data of quality 2 and 3

related to this, in the early morning and evening. These cases are probably influenced by shallow boundary layers where the lower part of the puff is within the boundary layer but is not mixed down to the ground since turbulence levels for the entire puff are those determined at its mean height.

4.3 Additional validation results investigating sensitivities of the performance statistics to the run configuration

This section contains a brief discussion of some parallel validation runs which were performed using different model configurations and choices for the input meteorological data. These parallel runs were carried out in an effort to explore sensitivity of the validation statistics to choices in the set up of the validation exercise. Here we consider three aspects of the sensitivity: (a) differences emerging from the use of observed boundary-layer depths in preference to using boundary-layer depths calculated within NAME III by the meteorological pre-processor; (b) differences resulting from the use of wind speed and direction information at different heights; and (c) identificaton of any manifest dependency of the statistical results on the level of statistical noise (as controlled by the value of the A1 parameter in the NAME III puff scheme). These issues are by no means intended to be an exhaustive list, although it would seem reasonable to suggest that they represent some of the main elements of the uncertainty in the set up of these runs, and that good agreement in the performance measures between these cases would increase our confidence in the results. Other aspects of the modelling set up which might be worth investigating in the future are the use of alternative met variables (such as heat flux instead of cloud cover) or changes to the model configuration (such as modifying other 'puff parameters' or 'dispersion options').

	z_{wind} m	$\overline{c}$ μg m^{-3}	σ_c μg m^{-3}	bias μg m^{-3}	NMSE	r	FB	FS	FA2
Observations		0.69	0.51	–	–	–	–	–	–
Mod-BL	100	0.74	0.52	-0.05	0.56	0.473	-0.072	-0.022	0.758
	10	0.73	0.50	-0.03	0.55	0.466	-0.049	0.017	0.724
Obs-BL	100	0.65	0.53	0.04	0.83	0.315	0.059	-0.035	0.694
	10	0.64	0.50	0.05	0.82	0.307	0.079	0.020	0.697

Table 3: Comparison of performance statistics using alternative run configurations (for quality 3 data)

4.3.1 Boundary layer depth (observed versus modelled values)

The depth of the boundary layer is often one of the most critical factors controlling dispersion of the plume, and needs to be given careful attention. The main validation runs used modelled values of the boundary-layer depth calculated by the NAME III met preprocessor, but it is now valuable to repeat the validation exercise using the observed values provided as part of the Kincaid met files. It should be noted that, although we talk here of an 'observed' boundary-layer depth and the Kincaid data set provides

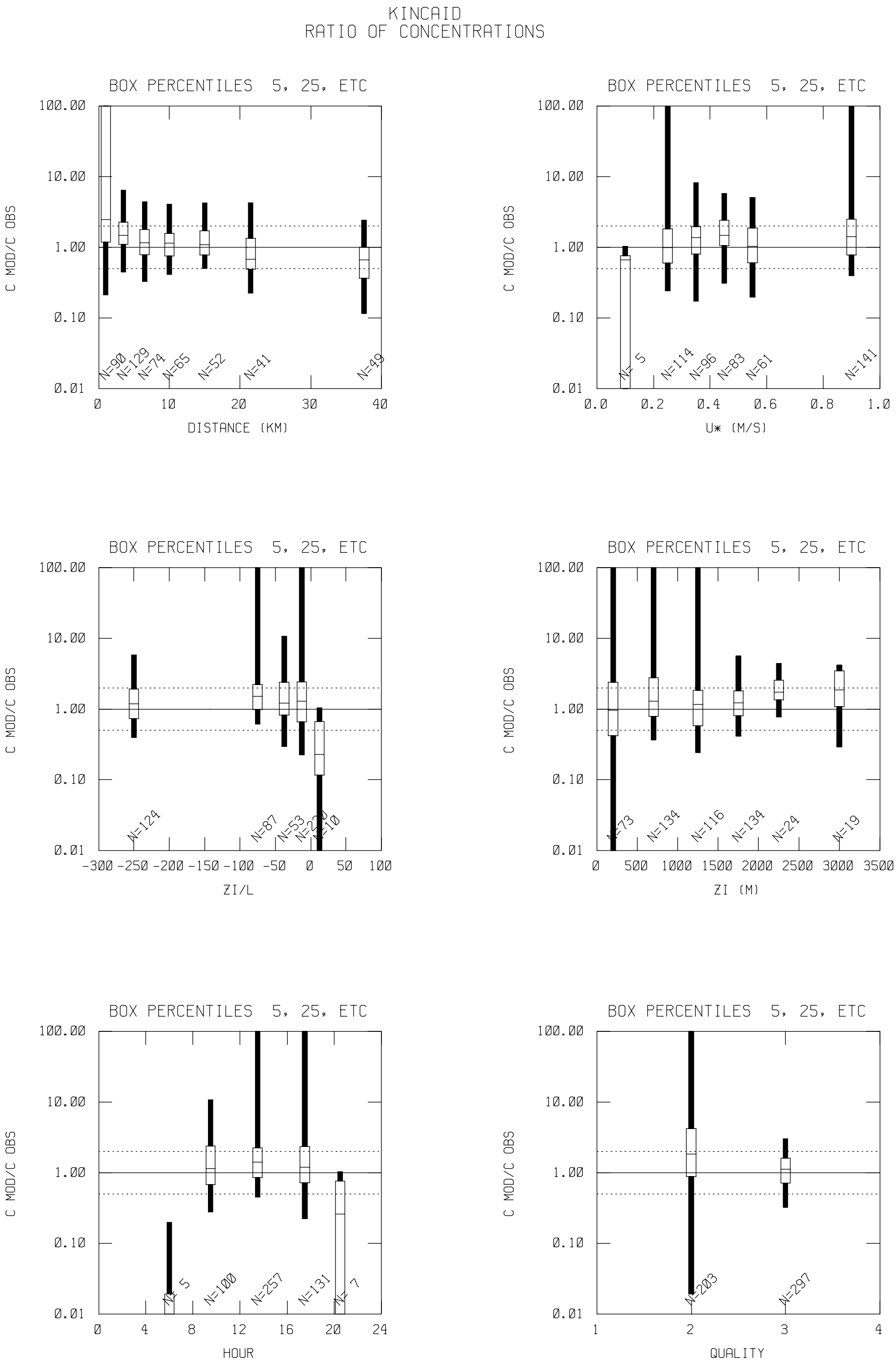

Figure 6: Residual box plots using data of quality 2 and 3

‘observations’ of this depth, in reality boundary-layer depth is not a directly observable quantity but needs to be inferred in some way from the measurements. Various definitions can be adopted and these use different techniques to estimate the depth of a boundary layer. The Kincaid data is based on a manual interpretation of radiosonde profiles. It is generally recognised that different techniques will provide different estimates, and the following analysis should be placed in the context of this uncertainty.

It is reported in the Kincaid documentation that observations of boundary-layer depth should generally be regarded as an upper estimate of the mixing depth. Such an interpretation would suggest that boundary-layer concentrations would tend to be underestimated by models using these observations, except for shallow boundary layers where an overestimate of z_i could result in the plume remaining in the boundary layer and leading to an overestimate of concentrations. However the relationship between the observations of boundary-layer depth and the modelled values determined by the preprocessor is not quite as simple as that suggested here. Figure 7 is a scatter plot comparing the model estimate of boundary-layer depth with the observations. Generally, it appears that the modelled values have much less spread than the observations (with the preprocessor frequently returning values in the range 500–1500 m, whereas observations cover a wider range from almost zero up to over 2000 m). This results in a tendency for NAME III predictions of boundary-layer depth to be significantly larger than observed values for shallow boundary layers (esp. stable cases) but also to be smaller than observations for deep convective boundary layers. As a consequence, concentrations obtained using observed z_i will tend to be less than those obtained using the NAME III predictions of z_i, and the results tend to support this conclusion.

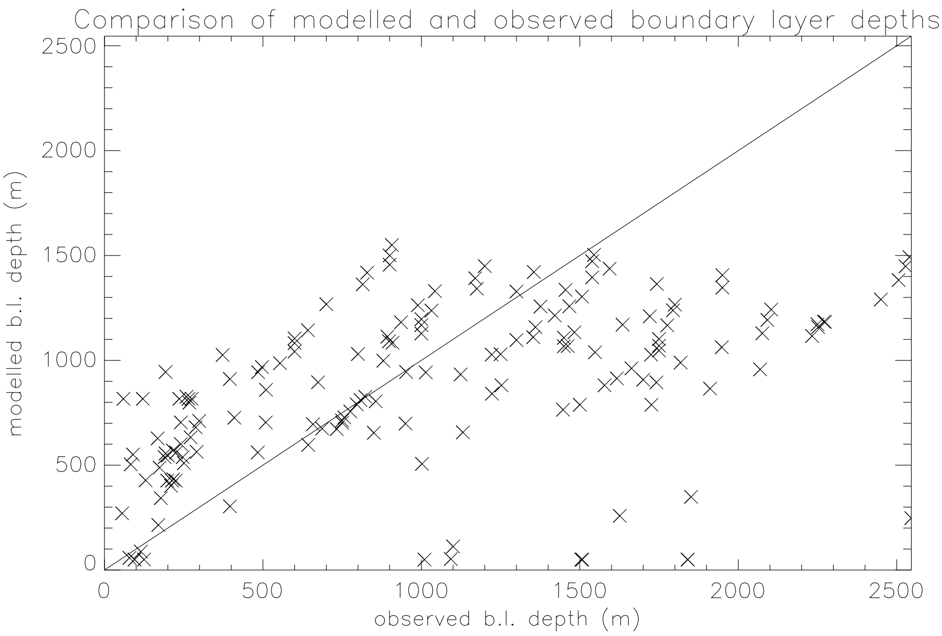

Figure 7: Scatter plot of observed boundary-layer depth against the modelled value provided by the NAME III meteorological preprocessor

A comparison of the validation statistics produced by model runs based on observed boundary-layer depth with those statistics produced by runs using model predictions of boundary-layer depth is given in Table 3. The under-estimate of the mean concentration when observations of boundary-layer depth are used (hinted at in the above discussion) is evident here, compared with an over-prediction of similar magnitude obtained with modelled values. The spread (as represented by σ_c) is similar in the two cases. However, some of the other measures show a significant deterioration when observed values of the boundary-layer depth are used, with the normalised mean square error, correlation and proportion of values within a factor of two all showing poorer performance. Similar relative results are obtained when wind speed and direction at 10 m are used in place of the 100 m wind data (see next section).

Figure 8 contains box-plots of the residuals (for data quality 3) obtained from the model runs based on observations of boundary-layer depth and 100 m wind data (c.f. Figure 3). At first glance, the box-plots confirm the message provided by the poorer performance statistics given above. The box widths are generally much greater and the under-estimate of concentrations in some situations is clearly evident. However, closer inspection of the fourth plot (based on boundary-layer depth) reveals an interesting insight into the results. It appears that the deterioration in the performance is particularly associated with cases of shallow boundary layers when the model severely under-estimates ground level concentrations (presumably because the plume is not mixed down to the ground in these cases with very shallow boundary layers). In fact, it could be argued that using observed mixing heights actually yield better results in other situations. This observation might be worth further investigation in the future.

4.3.2 Wind measurement height

The main validation exercise in Section 4.2 used wind data measured at 100 m. Wind information is also available at a height of 10 m in the Kincaid met files. Each measurement height has its own advantages and disadvantages; for instance, the 100 m wind is likely to be more representative of the meteorology at the stack height, but is also occasionally above the boundary layer (and the NAME III single-site met module then has difficulty in interpreting the information). The 100 m wind was chosen for the main exercise to coincide with the approach taken in previous NAME validation work, but we now repeat the modelling exercise using wind speed and direction recorded at 10 m.

Table 3 presents a comparison of the validation statistics produced by model runs based on wind measurements at 100 m with those based on wind measurements at 10 m. We focus attention on those runs where the model determines its own boundary-layer depth using the met preprocessor (i.e. the Mod-BL cases). The differences in the statistics are small, with neither the 100 m nor the 10 m runs providing a systematic improvement in all of the performance measures. For example, the 10 m case demonstrates a marginal improvement in terms of the bias and normalised mean square error but with a slight worsening of the correlation value and the proportion of data within a factor of two. Box plots analysing the residuals (for data of quality 3) are presented in Figure 9. A comparison of these box plots with those in Figure 3 shows only nominal differences, although it could be argued that the 10 m wind provides a better performance in cases with shallow boundary layers (this would support the comment made about the 100 m

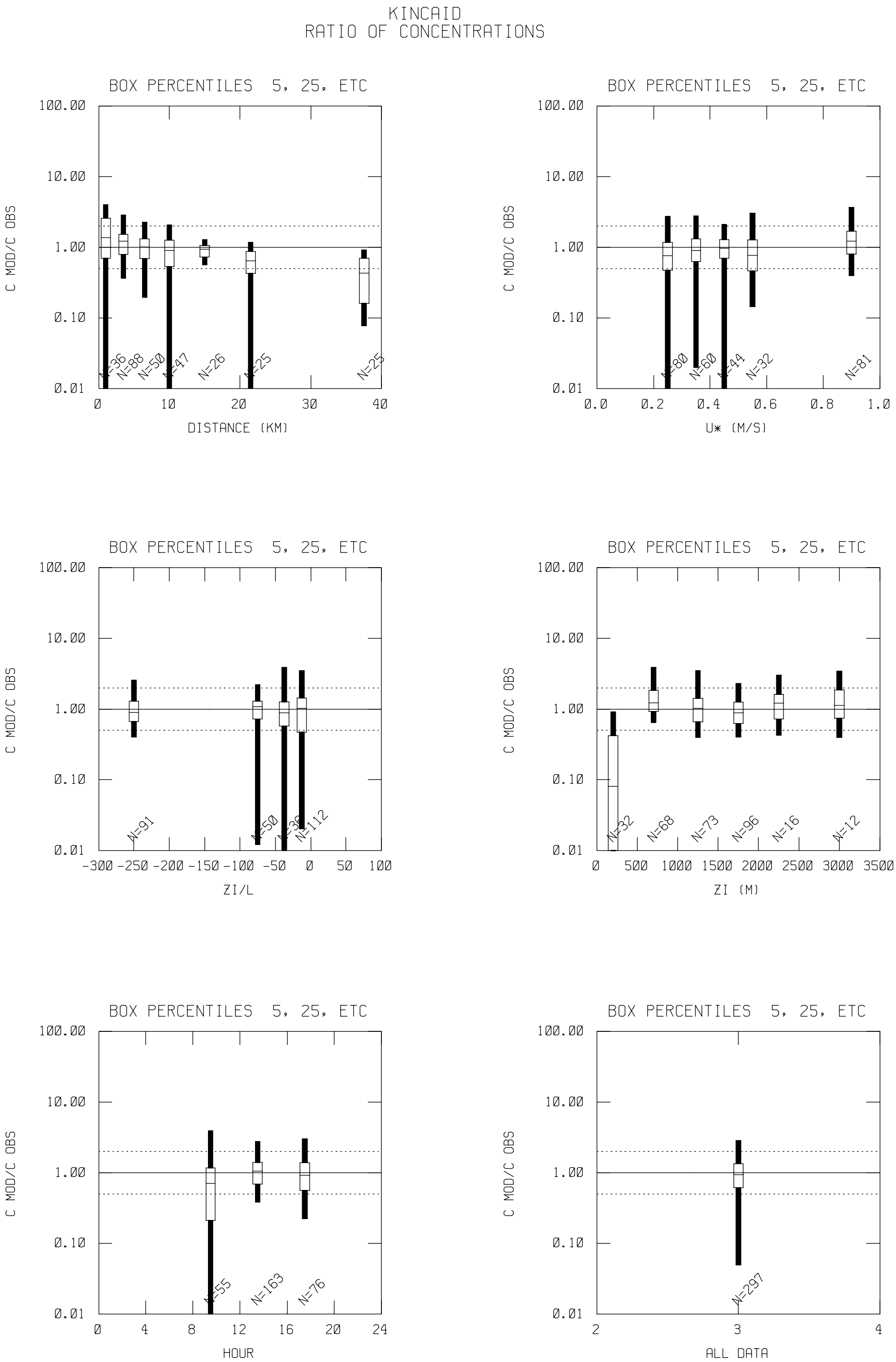

Figure 8: Residual box plots based on runs using observed values of boundary layer depth and 100 m wind information (for quality 3 data)

measurement height sometimes being above the boundary layer). Similar conclusions can be drawn for those runs where the boundary-layer depth has been determined using the Kincaid observations (i.e. the Obs-BL cases). Here a small reduction in mean concentration when using 10 m wind data leads to a slight worsening in the model bias.

4.3.3 Puff numbers (statistical noise)

The final sensitivity test aims to establish that the level of statistical noise in these validation runs is acceptable (i.e. does not have a major influence on the statistical results). Here the 'level of statistical noise' is related to the number of individual computational puffs created in the puff-splitting step and how well this multitude of puffs represents the overall structure of the plume. Inspection of the graphical output (mean concentration fields) would suggest that statistical noise is not an issue (except perhaps in a few cases with shallow boundary layers where the plume is slightly noisy; see Appendix 4.3.3). However it is prudent to confirm this suggestion by studying the statistical results over all cases. The model runs discussed in the previous section (that is, using model values of boundary-layer depth with 10 m winds) were repeated with the puff parameter A1 increased from 50 (the default value) to 200. We would expect this change to produce an approximate four-fold increase in the number of puffs.

The statistics are presented in Table 4, with only nominal differences evident between the two cases (supporting the view that the default value of A1 = 50 is sufficient here). Similarly, the box plots of residuals in Figure 10 compare well with the corresponding plots in Figure 9 (indeed, it is difficult to resolve any differences at all).

	$\overline{c}$ μg m^{-3}	σ_c μg m^{-3}	bias μg m^{-3}	NMSE	r	FB	FS	FA2
Observations	0.69	0.51	–	–	–	–	–	–
A1 = 50	0.73	0.50	-0.03	0.55	0.466	-0.049	0.017	0.724
A1 = 200	0.73	0.51	-0.04	0.55	0.465	-0.053	0.010	0.727

Table 4: Comparison of performance statistics for different levels of statistical robustness (for quality 3 data)

References

[1] Bowne, N.E. and Londergan, R.J. (1983). Overview, results and conclusions for the EPRI plume model validation and development project: plains site, *EPRI report EA-3074*.

[2] CERC (1999). ADMS validation summary, Cambridge Environmental Research Consultants Ltd.

[3] Hanna, S.R. and Paine, R.J. (1989). Hybrid Plume Dispersion Model (HPDM) development and evaluation, *J. Appl. Meteorol.*, **28**, pp. 206-224.

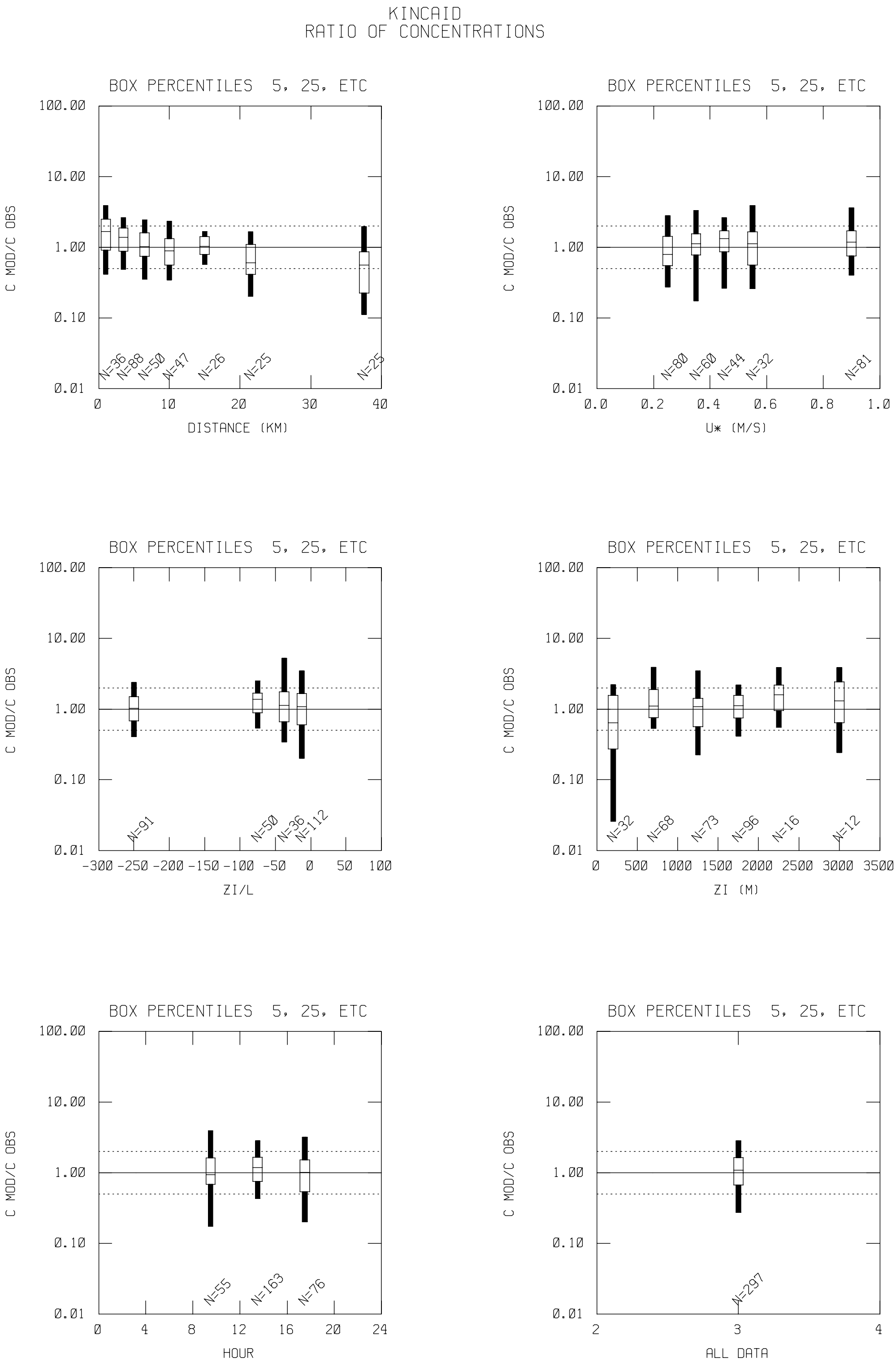

Figure 9: Residual box plots based on runs using model values of boundary layer depth and 10 m wind information (for quality 3 data)

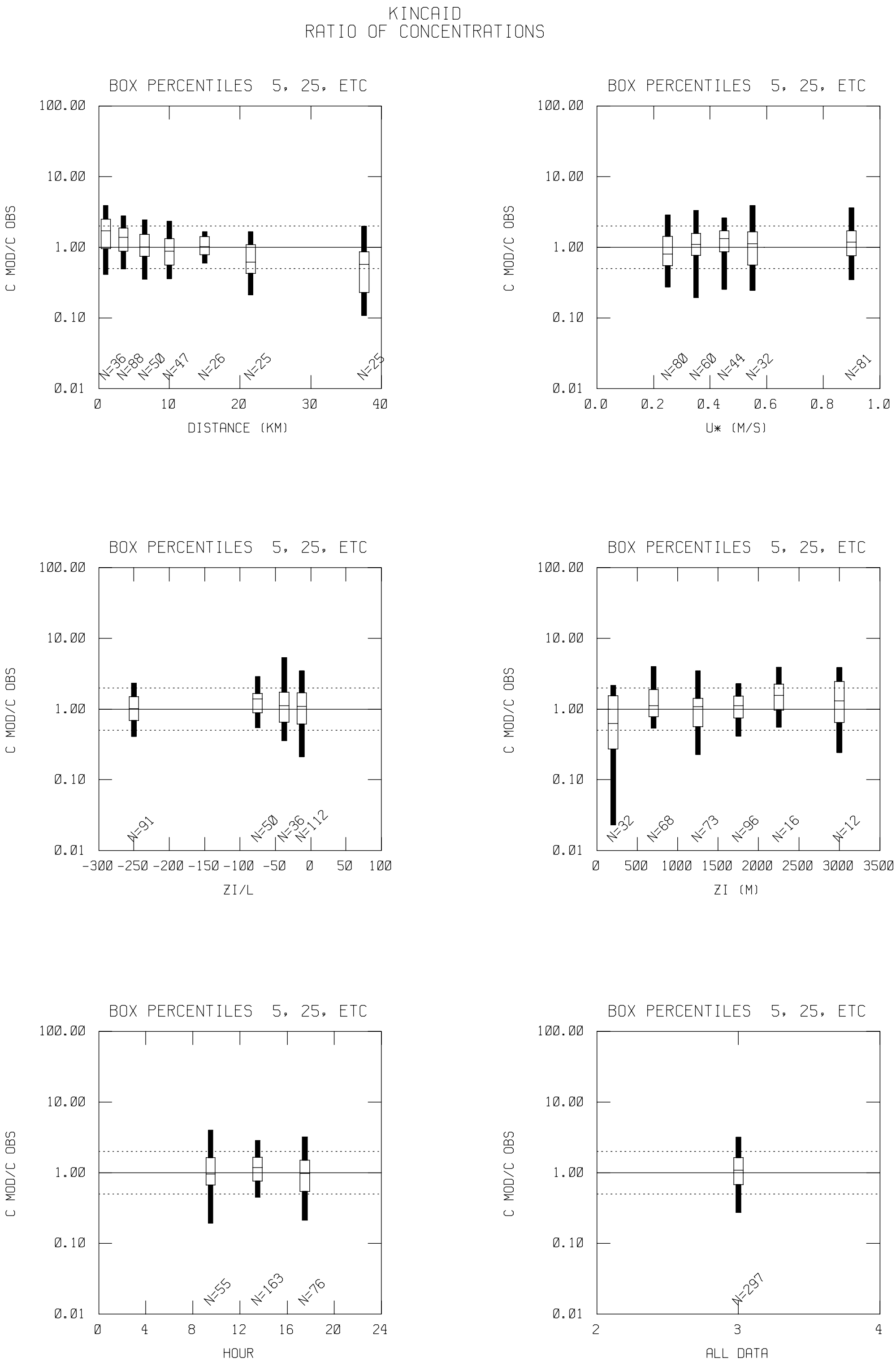

Figure 10: Residual box plots based on runs with the puff parameter A1 set to 200 (using model values of boundary layer depth and 10 m wind information, and showing quality 3 data, c.f. Figure 9)

[4] Olesen, H.R. (1994). Model Validation Kit for the workshop on 'Operational short-range atmospheric dispersion models for environmental impact assessments in Europe', Nov 1994, National Environmental Research Institute, Denmark.

[5] Olesen, H.R. (1995). The model validation exercise at Mol: overview of results, workshop on 'Operational short-range atmospheric dispersion models for environmental impact assessment in Europe', Nov 1994, published in *Int. J. Environment and Pollution*, **5**, pp. 761-784.

[6] Webster, H.N. (2001). Validation of NAME and its plume rise scheme against the Kincaid data set, *Met O (GMR) Turbulence and Diffusion Note No. 279.*

Appendix A: Example NAME III input file

[Note that this is an edited version of the actual input file. The '--->' notation denotes where long input blocks have been split across multiple lines as an aid to the reader.]

```
NAME III input file for Kincaid case: kin200480
Input file created on Mon Oct 11 08:17:05 BST 2004

! Kincaid input file - part 1.

Main Options:
Absolute or relative time?, Fixed met?, Time of fixed met, Flat Earth?,          Run Name,
                  absolute,         No,                  ,          No, Kincaid: kin200480,

 --->  Restart Case Interval, Restart Interval, Random Seed
 --->                      ,                 ,       Fixed

Multiple Case Options:
Multiple Cases?, Multiple Sets of Dispersion Options?
No,                                                  No

Output Options:
Folder
/data/local/aprj/Kincaid/kin200480

Horizontal Coordinate Systems:
Name,    Type, Pole Lat, Pole Long, Angle, x-Origin, y-Origin, x-unit, y-unit
LatLong,    1,     90.0,       0.0,   0.0,      0.0,      0.0,    1.0,    1.0

Vertical Coordinate Systems:
Name,  Type, Unit
m agl,    1,  1.0

Horizontal Grids:
Name,        H-Coord,  nx,  ny,          dx,          dy,         x0,      y0
HGrid500m,   LatLong, 205, 205,  0.00583481,  0.00449645, -0.0851506, 39.1314
HGrid50m,    LatLong, 205, 205, 0.000583481, 0.000449645,   0.450485, 39.5441
KincaidSite, LatLong,   1,   1,      0.0001,      0.0001,      00.51, 39.59

Vertical Grids:
Name,    Z-Coord,  nz,   dz,   z0
GLConc1,   m agl,   1, 40.0, 20.0
GLConc2,   m agl,   1,  1.0,  0.0

Temporal Grids:
Name,    nt,        dt,        t0
TGrid1,   3,  01:00:00,  20/04/1980 13:30:00
TGrid2,   3,  01:00:00,  20/04/1980 14:00:00

Sources:
Name,         Shape, H-Coord, Z-Coord,      X,     Y,     Z,   dX,  dY,  dZ, Angle, Metres?,
kincaid1, Ellipsoid, LatLong,   m agl,  00.51, 39.59, 187.0,  9.0, 9.0, 0.0,   0.0,     Yes,
kincaid2, Ellipsoid, LatLong,   m agl,  00.51, 39.59, 187.0,  9.0, 9.0, 0.0,   0.0,     Yes,
kincaid3, Ellipsoid, LatLong,   m agl,  00.51, 39.59, 187.0,  9.0, 9.0, 0.0,   0.0,     Yes,

 --->            Source Strength, Time Dependency, Plume Rise?, Temperature,  Flow Velocity,
 --->  INERT-TRACER  36720.0 g,                ,         yes,       416.0,           14.6,
 --->  INERT-TRACER  29520.0 g,                ,         yes,       416.0,           14.6,
 --->  INERT-TRACER  30240.0 g,                ,         yes,       416.0,           15.0,

 --->  # Particles,  Max Age, Top Hat,          Start Time,           Stop Time
 --->        20000, 12:00:00,     Yes, 20/04/1980 13:00:00, 20/04/1980 14:00:00
 --->        20000, 12:00:00,     Yes, 20/04/1980 14:00:00, 20/04/1980 15:00:00
 --->        20000, 12:00:00,     Yes, 20/04/1980 15:00:00, 20/04/1980 16:00:00
```

```
! Kincaid input file - part 2.

Species:
Name,         Half Life, Molecular Weight, Deposition Velocity, Wet Type, Category,  Material Unit
INERT-TRACER,    Stable,                1,                 0.0,        0,   TRACER,              g

Domains:
Name,             H-Coord,     X Min,    X Max,     Y Min,    Y Max, H Unbounded?, Z-Coord,
DispersionDomain, LatLong,       0.0,      0.0,       0.0,      0.0,          Yes,   m agl,
MetDomain,        LatLong,      -0.7,      1.7,      38.6,     40.6,           No,   m agl,

 --->    Z Max, Z Unbounded?, Start Time, End Time, Max Travel Time
 --->  10000.0,           No,  -infinity, infinity,        24:00:00
 --->  10000.0,           No,  -infinity, infinity,        24:00:00

Output Requirements - Fields:
Quantity,                   Species, Source,      H-Grid,  Z-Grid, T-Grid, BL Average, T Av Or Int,
Air Concentration,    INERT-TRACER,       ,   HGrid500m, GLConc2, TGrid2,         No,          Av,
Boundary Layer Depth,             ,       , KincaidSite, GLConc2, TGrid2,         No,          No,

 --->  Av Time, # Av Times, Sync?, Graph?, Screen?, Disk?, Stat?, Across, Separate File,
 --->    01:00,          4,    No,     No,      No,   Yes,    No,     TZ,             T,
 --->         ,           ,    No,     No,      No,   Yes,    No,       ,              ,

 --->  Output Format, Output Group
 --->             IA, Fields_grid2
 --->              A, MetData

Output Requirements - Fields:
Quantity,          Species, Source, T-Grid, H-Coord, Z-Coord, T Av Or Int, Sync?, Graph?, Screen?,
# Particles,              ,       , TGrid2,        ,        ,          No,   Yes,     No,      No,
# Particle Steps,         ,       , TGrid2,        ,        ,          No,   Yes,     No,      No,
# Puffs,                  ,       , TGrid2,        ,        ,          No,   Yes,     No,      No,
# Puff Steps,             ,       , TGrid2,        ,        ,          No,   Yes,     No,      No,

 --->  Disk?, Stat?, Output Format, Output Group
 --->    Yes,    No,             A, Numbers
 --->    Yes,    No,             A, Numbers
 --->    Yes,    No,             A, Numbers
 --->    Yes,    No,             A, Numbers

Sets of Dispersion Options:
Skew Time, Velocity Memory Time, Inhomogeneous Time, DeltaOpt, Puff Time, Sync Time,
    00:00,             infinity,           infinity,        1,  infinity,  00:05:00,

 --->  Computational Domain, Puff Interval, Deep Convection?, Radioactive Decay?, Agent Decay?,
 --->      DispersionDomain,      00:01:00,               No,                 No,           No,

 --->  Dry Deposition?, Wet Deposition?, Turbulence?, Meander?, Chemistry?
 --->               No,              No,         Yes,      Yes,         No

Single Site Met Module Instances:
Name,       H-Coord,   Long,   Lat, Height,  z0, z0d, Representative?,
Kincaidmet, LatLong,  00.51, 39.59,  100.0, 0.1, 0.1,             Yes,

 --->                        Met File, Ignore Fixed Met Time?
 --->  ./Met/kincaidmet_extended.met,                      No

Single Site Flow Module Instances:
Name,         Met Module,        Met, Domain
Kincaid, Single Site Met, Kincaidmet, MetDomain
```

```
Flow Order: Update
Flow Module,      Flow
Single Site Flow, Kincaid

Flow Order: Convert
Flow Module,      Flow
Single Site Flow, Kincaid

Flow Order: Flow
Flow Module,      Flow
Single Site Flow, Kincaid

Flow Attributes:
Name,    Flow Order
Update,  Update
Convert, Convert
Flow,    Flow
```

Appendix B: Revisiting the 'blobby plume' issue – some examples

The previous validation study (NAME III, version 1.3) identified a problem with patchy or 'blobby' plumes, see Figure 11 for an example from the morning of 05/05/80. This issue was especially evident in cases with shallow boundary layers where the plume travels some distance within the free troposphere before mixing down to ground level. Figure 12 is an updated version of this case, and illustrates that the problem is not as severe in version 2.0. The situation is further improved when the NAME III met preprocessor is used to predict the boundary-layer depth (see Figure 13) such that the problem is now largely resolved in this instance. The problem remains a minor issue in a few other cases (even when boundary-layer depth is determined within the model); for instance, Figure 14 shows a similar situation during the morning of 01/06/81.

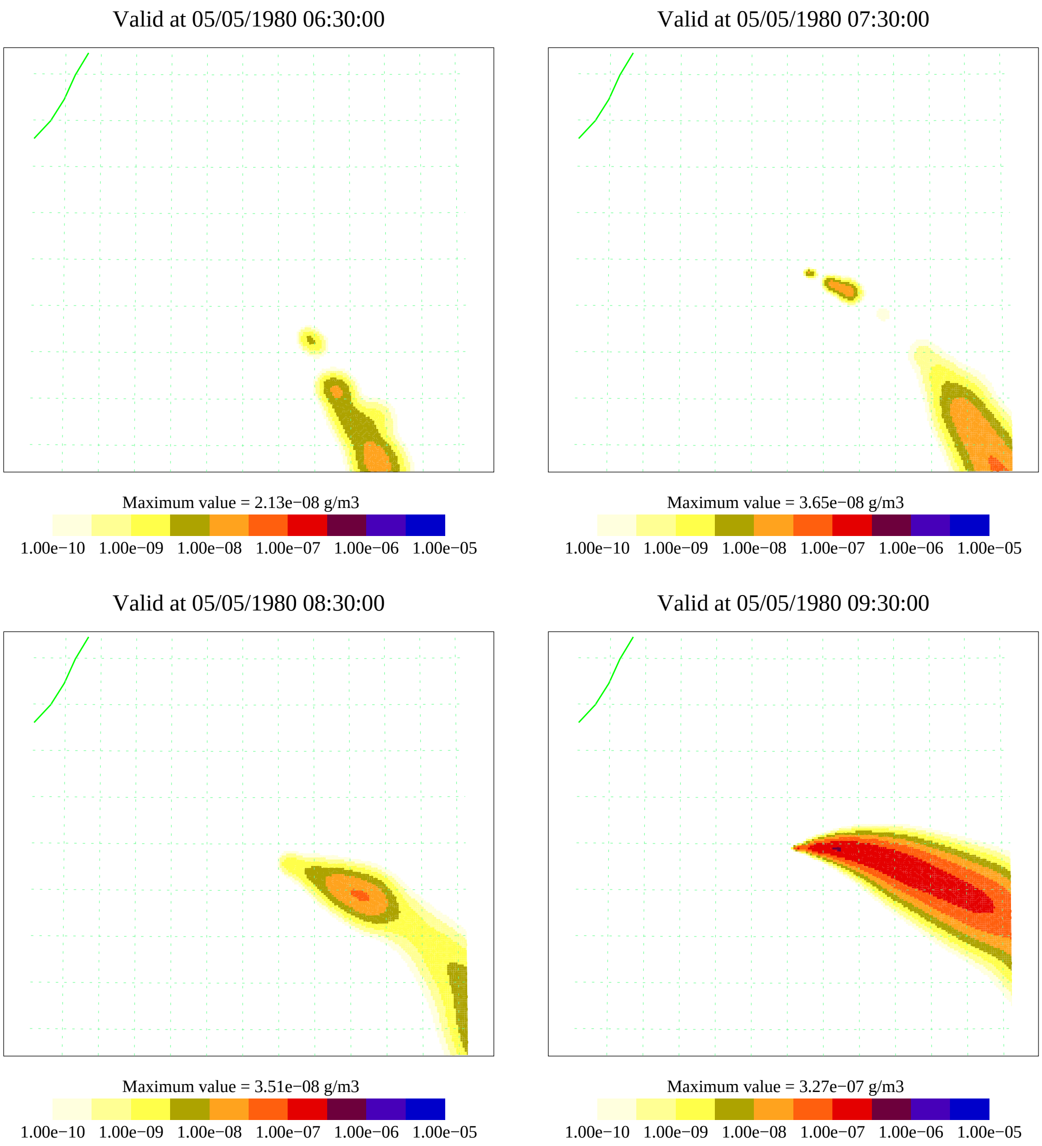

Figure 11: Example of a 'blobby' plume in NAME III (version 1.3)

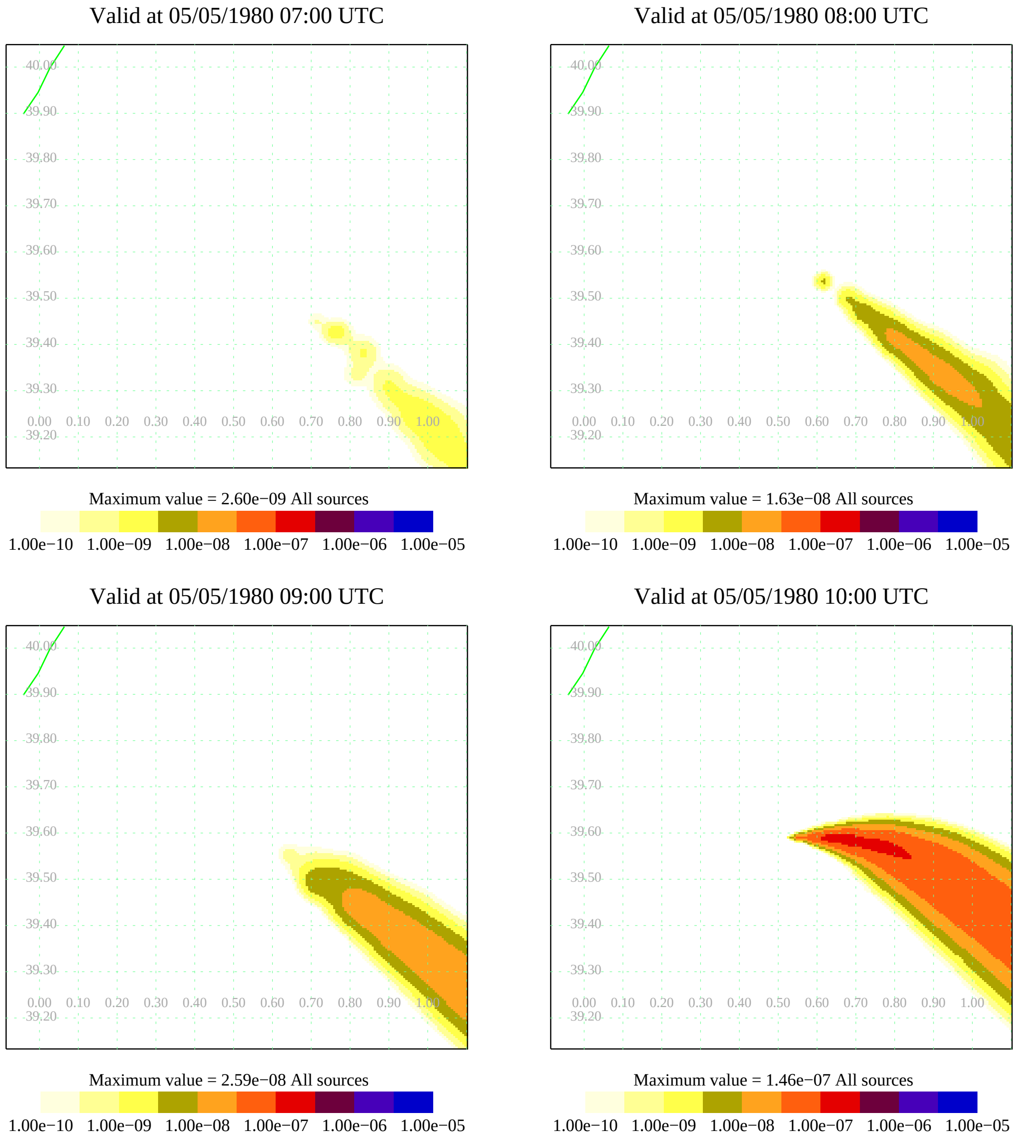

Figure 12: Revised output for NAME III (ver 2.0) using observed boundary-layer depth

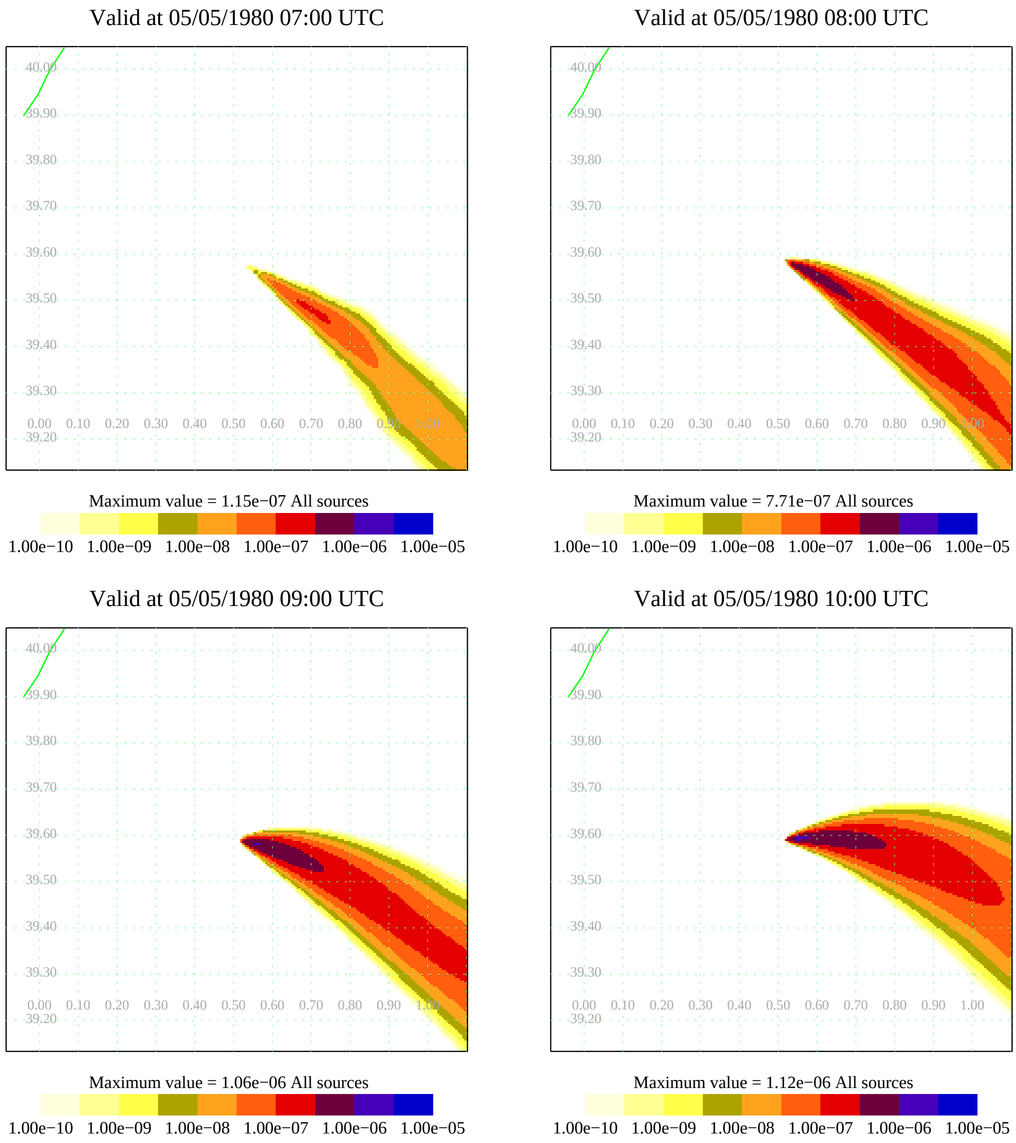

Figure 13: Revised output for NAME III (ver 2.0) using modelled boundary-layer depth

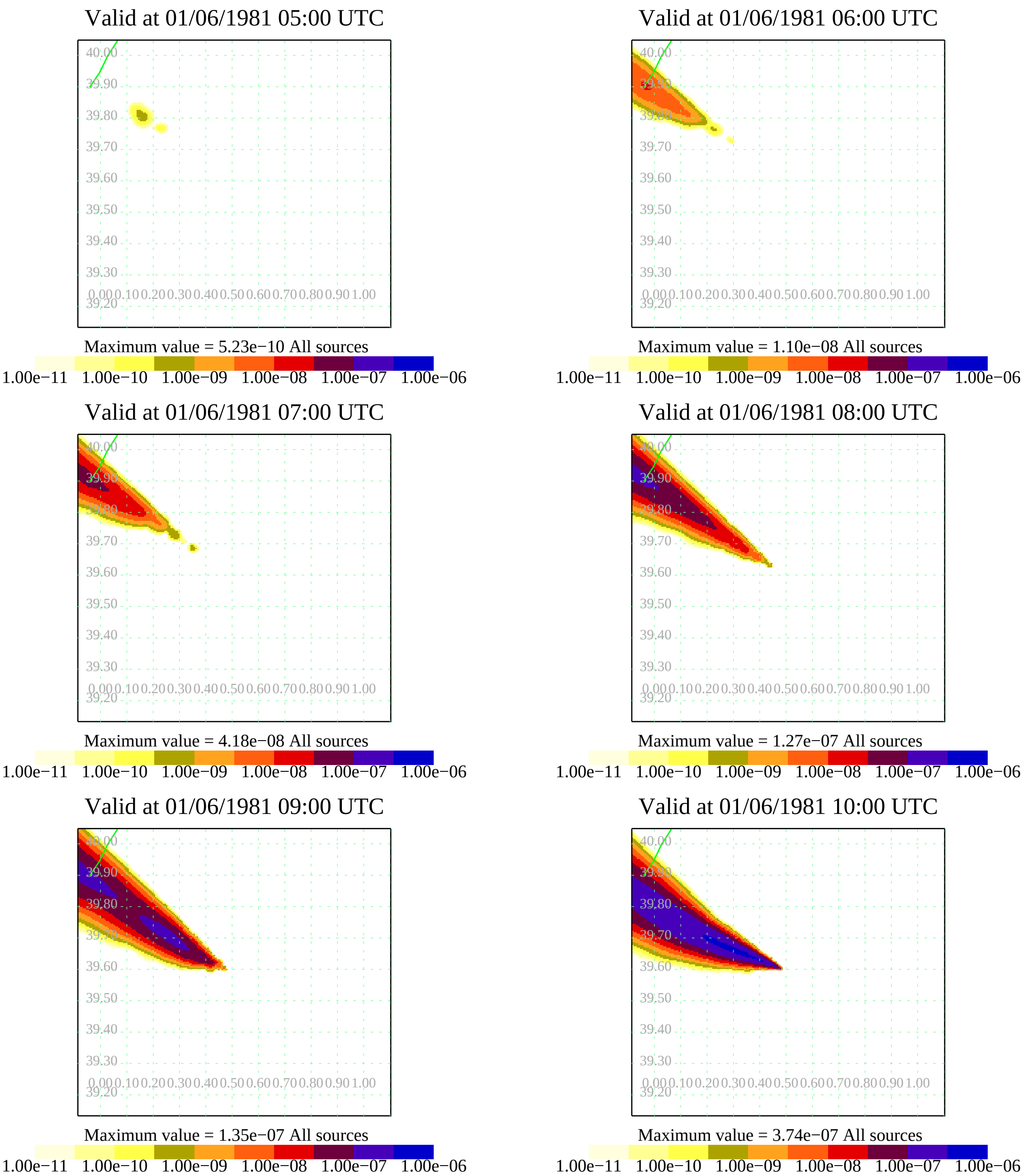

Figure 14: Example of a minor ‘blobby’ plume in NAME III (version 2.0)