TAPM V4. Part 2: Summary of Some Verification Studies.

Peter Hurley, Mary Edwards and Ashok Luhar
CSIRO Marine and Atmospheric Research Paper No. 26
October 2008

Abstract

Air pollution predictions for environmental impact assessments usually use Gaussian plume/puff models driven by observationally-based meteorological inputs. An alternative approach is to use prognostic meteorological and air pollution models, which have many advantages over the Gaussian approach and have now become a viable tool for performing year-long simulations. Continuing rapid increases in computing power have brought this approach well within the reach of a desktop PC. This paper provides some verification studies of The Air Pollution Model (TAPM) for two US tracer experiments (Kincaid and Indianapolis) used internationally for model inter-comparison studies, for several annual US dispersion datasets (Bowline, Lovett and Westvaco), for annual meteorology and/or dispersion in various regions throughout Australia (Anglesea, Kwinana, Kalgoorlie, Melbourne, Sydney, Brisbane and Perth).

The meteorological results show that TAPM performs well in a variety of regions (e.g., coastal, inland and generally complex terrain for sub-tropical to mid-latitude conditions). The pollution results show that TAPM performs well for a range of important phenomena (e.g. nocturnal inversion break-up fumigation; stable, neutral, convective and building wake dispersion; shoreline fumigation; and general dispersion in complex rural and urban conditions). In particular, TAPM performs very well for the prediction of extreme pollution statistics, important for environmental impact assessments, for both non-reactive (tracer) and reactive (nitrogen dioxide, ozone and particulate) pollutants for a variety of sources (e.g. industrial stacks and/or general surface or urban emissions).

2

1 Introduction

Air pollution models that can be used to predict pollution concentrations for periods of up to a year are generally semi-empirical/analytic approaches based on Gaussian plumes or puffs. These models typically use either a simple surface-based meteorological file or a diagnostic wind field model based on available observations. The Air Pollution Model (TAPM – see Hurley et al., 2005a, for details on TAPM V2 and Hurley et al., 2005b, for details on TAPM V3) is different to these approaches in that it solves the fundamental fluid dynamics and scalar transport equations to predict meteorology and pollutant concentration for a range of pollutants important for air pollution applications. It consists of coupled prognostic meteorological and air pollution concentration components, eliminating the need to have site-specific meteorological observations. Instead, the model predicts the flows important to local-scale air pollution, such as sea breezes and terrain-induced flows, against a background of larger-scale meteorology provided by synoptic analyses.

This paper follows on from the technical description of the model (TAPM V4) in part 1 of this series (Hurley, 2008) and updates and summarises a number of model verification studies presented previously in the literature using earlier versions of the model. In Section 2 TAPM is verified for two US tracer datasets (Kincaid and Indianapolis) used for international model inter-comparison. Sections 3, 4 and 5 show verification studies for annual meteorology and point source dispersion for the US studies at Bowline (building wakes), Lovett and Westvaco (complex terrain). Sections 6 and 7 present model verification for the Anglesea (Victoria) and Kwinana (Western Australia) regions for annual datasets of sulfur dioxide dispersion. Section 8 examines the performance of the model for annual upper-level meteorology, by comparing wind predictions to SODAR data for Kalgoorlie in Western Australia. Section 9 looks at annual meteorology and photochemically reactive dispersion for urban sources in Melbourne (Victoria). Section 10 examines model performance for annual meteorology in urban regions of Australia (Perth, Brisbane, Sydney and Melbourne). All of these verification datasets have well known emission characteristics and good quality monitoring data with well quantified uncertainties, allowing a meaningful evaluation of model performance to be made.

2   International model inter-comparison datasets

In 1991, The Joint Research Centre of the European Commission launched an initiative for increased cooperation and standardisation of atmospheric dispersion models for regulatory purposes. As part of the initiative, a series of conferences on "Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes" was organised to promote the use of new-generation models within atmospheric dispersion modelling, and in general improve "modelling culture" (for more details see Olesen, 1995). A set of four short-range field data sets, called the Model Validation Kit, was prepared to facilitate a standard and uniform comparison of model results. The two most comprehensive of the four field datasets on point-source plume dispersion are the 1980–81 Kincaid (rural) dataset, and the 1985 Indianapolis (urban) dataset, both taken in the US and widely used for model evaluation purposes. These two datasets were taken in relatively simple orography (i.e. flat terrain, no coastal influences), where simple plume models should be expected to perform well.

In this Section, the performance of TAPM V4 is evaluated for the Kincaid and Indianapolis datasets. In order to be consistent with other dispersion models used for these intercomparison studies, here TAPM was run with wind data assimilation – for more detail, see Luhar and Hurley (2003).

Since the above two data sets precede the global synoptic meteorological data supplied with TAPM, which are given from 1997, we used the National Center for Environmental Prediction (NCEP)/National Center for Atmospheric Research (NCAR) reanalysis data (Kalnay et al., 1996) on horizontal wind components, temperature and moisture, to obtain the required synoptic fields for the model. These data have a horizontal resolution of 2.5° and a temporal resolution of 6 h, while the vertical levels are in a pressure coordinate system with the lowest five levels being 1000, 925, 850, 700 and 600 hPa.

2.1   Kincaid

EPRI's (Electric Power Research Institute) Kincaid field study was conducted in 1980 and 1981 (Bowne et al., 1983). It involved sulfur hexafluoride (SF_6) tracer releases from the 187 m stack (with diameter 9 m) at the Kincaid power plant in Illinois, USA. The power plant is surrounded by relatively flat farmland with some lakes (roughness length of about 0.1 m).

Most meteorological measurements were taken from 10-m and 100-m towers located on a central site about 650 m east of the power plant together with solar and terrestrial radiation measurement gear. The wind observations were taken at the 10, 30, 50 and 100 m AGL levels, while the temperature measurements were taken at the 10, 50 and 100 m levels. There are also data from a National Weather Service station 30.6 km northwest of the source, and profiles from routine radiosonde releases 120 km north of the source.

Hourly-averaged concentrations of SF_6 due to buoyant power station plumes were observed by ground-level monitors on a maximum of 12 arcs at distances 0.5, 1, 2, 3, 5, 7, 10, 15, 20, 30, 40 and 50 km from the stack. There are three measurement periods: 20 April−9 May 1980, 10−25 July 1980, and 16 May−1 June 1981. Figure 2.1 shows the locations of the stack, meteorological sites and the tracer monitors on 22 May 1981. A total of 171 hours of tracer data are available in the Model Validation Kit, representing mostly daytime convective cases. Arc-wise maxima were calculated from the crosswind concentration variation, and a quality indicator was assigned to each value. It is recommended that only the data with quality indicator 2 (maxima identified) and 3 (maxima well defined) be used for model comparison. Out of a total of 1284 arc-hours of data, 585 are quality 2 and 3, and 338 are quality 3.
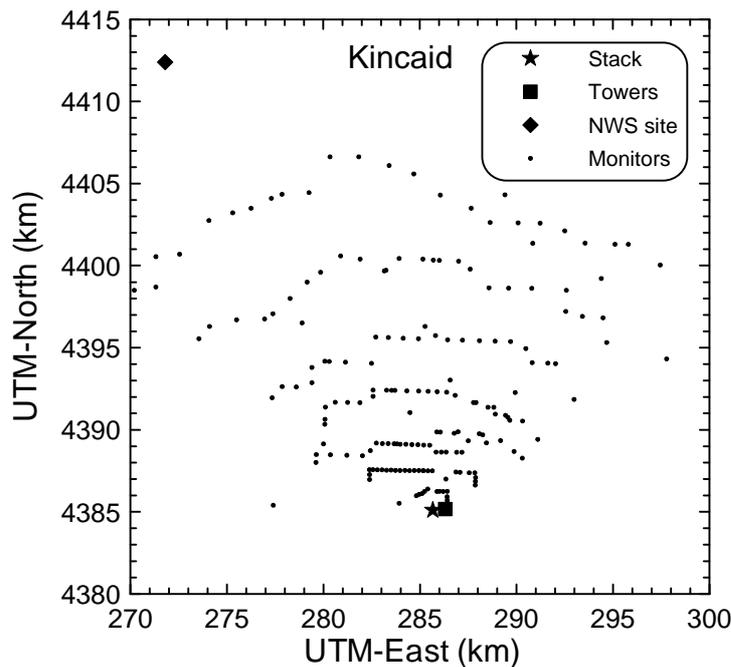
Figure 2.1. Concentration monitors (22 May 1981), and meteorological sites for the Kincaid study.

TAPM was run for the above three data periods separately with an extra spin-up day at the start in each run. Three nested domains of 31×31 horizontal grid points at 16-km, 4-km, and 1-km spacing for the meteorology, and 61×61 horizontal grid points at 8-km, 2-km, and 0.5-km spacing for the pollution, both centred on the stack coordinates, were used. There were 25 vertical levels: 10, 25, 50, 100, 150, 200, 250, 300, 400, 500, 600, 750, 1000, 1250, 1500, 1750, 2000, 2500, 3000, 3500, 4000, 5000, 6000, 7000 and 8000 m. The model was run in Lagrangian mode to capture the near-source dispersion more accurately. Given that no advice based on local measurements was given about surface moisture availability, we used the model default deep soil moisture content of 0.15 kg kg^{-1}. The wind speeds and directions observed at the tower site at four levels, namely 10, 30, 50 and 100 m AGL, were assimilated in model calculations. The hourly average pollution predictions on the 0.5 km spaced grid were processed to obtain ground-level concentration maxima at the 0.5, 1, 2, 3, 5, 7, 10 and 15 km arcs while those on the 2-km spaced pollution grid were processed to obtain the maxima at the 20, 30, 40 and 50 km arcs.

Table 2.1 lists model performance statistics, including RHC_R and MAX_R that indicate performance for the extreme end of the concentration distribution. The results show that TAPM performs well for both the mean and for the extreme performance statistics for the Kincaid dataset.

Figure 2.2 shows TAPM Quantile-Quantile (Q-Q) plots of the (sorted) predicted versus the (sorted) observed concentrations for both Quality 2&3 and Quality 3 data. These plots again illustrate that TAPM performs well for extreme concentrations, and does slightly better for the higher quality data, except for the maximum concentration.

Table 2.1. Model performance statistics for Kincaid (a) for Quality 2&3 data and (b) for Quality 3 data.

(a) Q2&3	MEAN	STD	NMSE	COR	IOA	RHC$_R$	MAX$_R$
OBSERVED	41.0	39.3	0.00	1.00	1.00	1.00	1.00
TAPM	53.2	48.8	1.40	0.27	0.50	1.17	1.08

(b) Q3	MEAN	STD	NMSE	COR	IOA	RHC$_R$	MAX$_R$
OBSERVED	54.3	40.3	0.00	1.00	1.00	1.00	1.00
TAPM	56.6	44.7	0.86	0.27	0.53	0.91	0.77

Key: MEAN = Arithmetic Mean (ng m^{-3} (g s^{-1})$^{-1}$), STD = Standard Deviation (ng m^{-3} (g s^{-1})$^{-1}$), NMSE = normalised mean square error, COR = correlation coefficient, IOA = Index of Agreement; RHC$_R$ = Ratio (Predicted/Observed) of Robust Highest Concentration (RHC); MAX$_R$ = Ratio (Predicted/Observed) of Maximum Concentration.
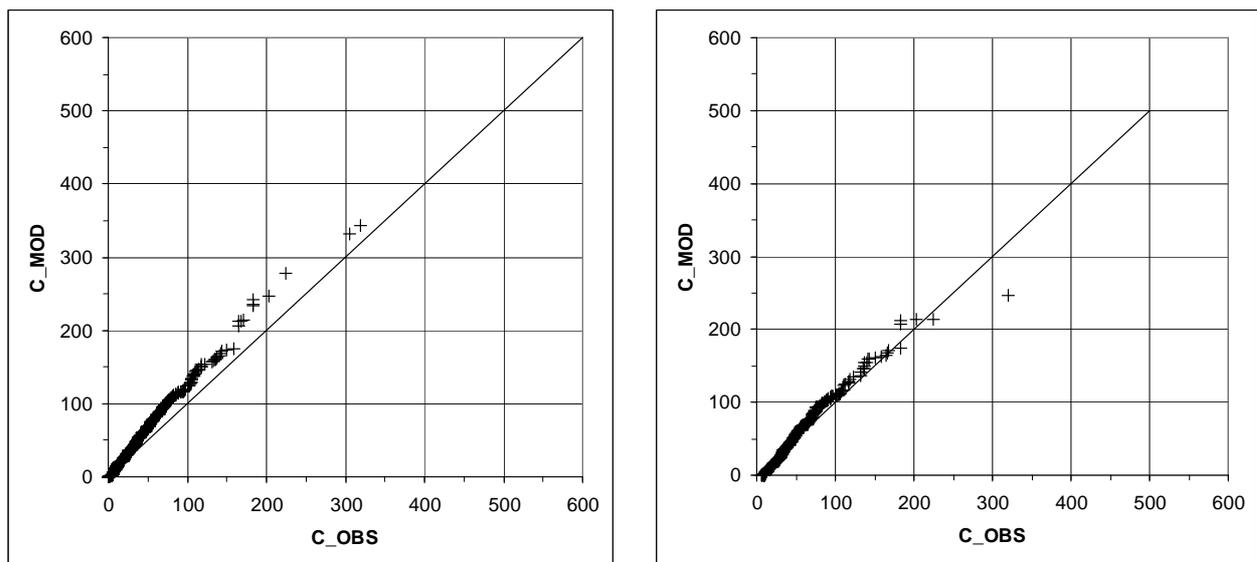

Figure 2.2. Kincaid Quantile-Quantile plots of scaled concentration for Quality 2&3 (left) and Quality 3 (right) data.

2.2   Indianapolis

A full description of EPRI's Indianapolis field study, conducted during 16 September to 12 October, 1985, is given in TRC (1986). It involved SF$_6$ tracer releases from the 83.8-m stack (with diameter 4.72 m) at the Perry K power plant on the southwest edge of Indianapolis, Indiana, USA. The stack is located in a typical industrial/commercial/urban complex with many buildings within one or two kilometres (roughness length of about 1 m), and relatively flat local terrain.

Meteorological observations were taken from a 94 m height at the top of a bank building in the middle of the urban area, from two 10-m towers in suburban and rural areas, and an 11-m

tower at an urban location. In addition, vertical meteorological profiles were also taken. Hourly-averaged concentrations were observed on a network of up to 160 ground-level monitors on 12 arcs at distances 0.25, 0.5, 0.75, 1.0, 1.5, 2, 3, 4, 6, 8, 10 and 12 km from the stack. To sample the plume, the network of monitors was moved so that it was downwind of the source. Data were taken in 8 or 9 hour test blocks with 19 such blocks altogether. Figure 2.3 shows the locations of the stack, meteorological sites and the tracer monitors corresponding to the test block 9. A total of 170 hours of tracer data is available, representing all stability classes and most wind speed ranges. Arc-wise maxima were calculated from the crosswind concentration variation, and a quality indicator was assigned to each value. It is recommended that only the data with quality indicator 2 (maxima identified) and 3 (maxima well defined) be used for model comparison. Out of a total of 1511 arc-hours of data, 1216 are quality 2 and 3, and 469 are quality 3.

TAPM was run for the period 15 September to 12 October, 1985, with four nested domains of 30×30 horizontal grid points at 30-km, 10-km, 3-km and 1-km spacing for the meteorology, and 101×101 horizontal grid points at 7.5-km, 2.5-km, 0.75-km and 0.25-km spacing for the pollution, both centred on the stack coordinates. The 25 vertical levels were the same as in the Kincaid case. The model was run in Lagrangian mode to capture the near-source dispersion more accurately. For Indianapolis data, a value of 0.5 is recommended for the moisture availability factor, which is defined as the ratio of the surface latent heat flux to the total surface heat flux. To match this value, we used a deep soil moisture content of $0.3 \, \text{kg} \, \text{kg}^{-1}$. The wind speeds and directions observed at the Urban tower (10 m AGL) and the Bank building (94 m) were assimilated into model calculations. The hourly average pollution predictions on the 0.25-km spaced grid were processed to obtain ground-level concentration maxima at the 12 arcs.
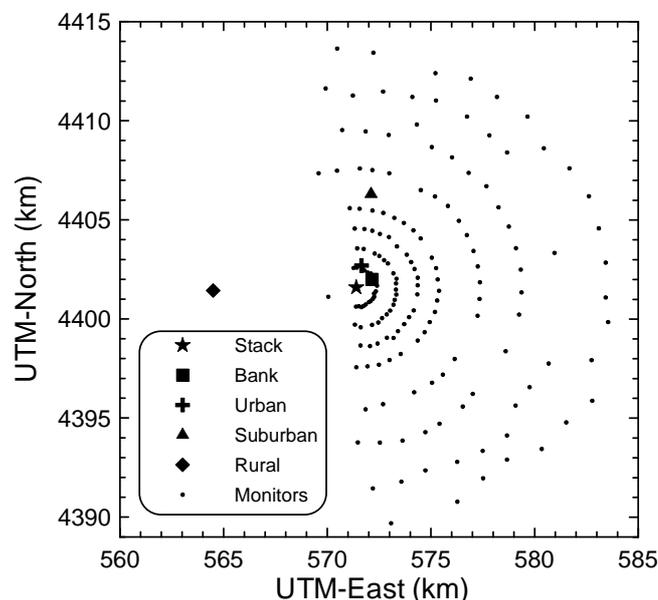

Figure 2.3. Tracer monitors (Test 9), and meteorological sites for the Indianapolis study.

Table 2.2 lists model performance statistics for TAPM. The results show that for the Indianapolis dataset, TAPM performs well for mean and extreme performance statistics, with little bias shown by TAPM for extreme statistics.

Table 2.2. Model performance statistics for Indianapolis (a) for Quality 2&3 data and (b) for Quality 3 data.

(a) Q2&3	MEAN	STD	NMSE	COR	IOA	RHC$_R$	MAX$_R$
Observed	258	222	0.00	1.00	1.00	1.00	1.00
TAPM	207	256	1.16	0.49	0.70	0.91	0.97

(b) Q3	MEAN	STD	NMSE	COR	IOA	RHC$_R$	MAX$_R$
Observed	352	221	0.00	1.00	1.00	1.00	1.00
TAPM	324	286	0.62	0.49	0.69	0.93	1.05

Key: MEAN = Arithmetic Mean (ng m^{-3} (g s^{-1})$^{-1}$), STD = Standard Deviation (ng m^{-3} (g s^{-1})$^{-1}$), NMSE = normalised mean square error, COR = correlation coefficient, IOA = Index of Agreement; RHC$_R$ = Ratio (Predicted/Observed) of Robust Highest Concentration (RHC); MAX$_R$ = Ratio (Predicted/Observed) of Maximum Concentration.

Figure 2.4 shows TAPM Quantile-Quantile (Q-Q) plots of the (sorted) predicted versus the (sorted) observed concentrations for both Quality 2&3 and Quality 3 data. These plots again illustrate that TAPM shows good performance for extreme concentrations, although there is some underprediction of the frequency of low concentrations that occur mostly at night.
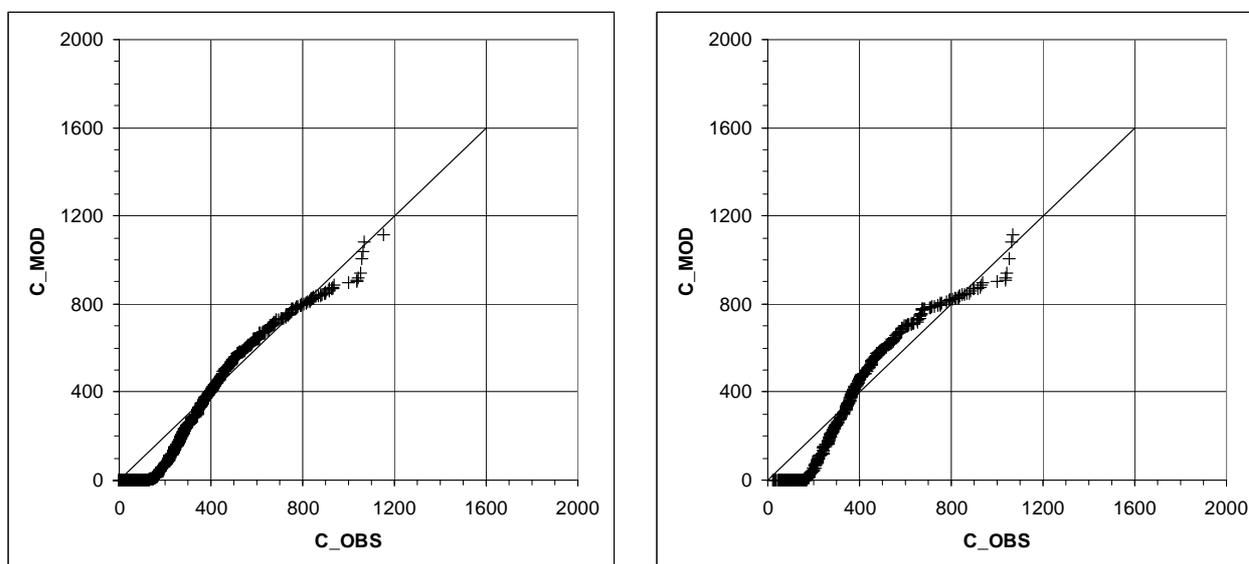

Figure 2.4. Indianapolis Quantile-Quantile plots of scaled concentration for Quality 2&3 (left) and Quality 3 (right) data.

3 Bowline

The Bowline 1981 annual dataset consists of SO_2 measurements at four monitoring sites situated to the south of a power station in a river valley in New York (U.S.) (Schulman and Hanna, 1986). The plant consists of two 87 m high stacks, separated by a distance of 90 m, and situated about 27 m to the east of the main building, which is about 66 m high, 140 m wide and 76 m deep. Peak concentrations due to building downwash effects occurred mainly at the two monitoring sites situated to the southeast of the plant (see the caption in Figure 3.1 for site location details). Source characteristics, emission data, building information and monitoring data were obtained from the Bowline dataset on the AERMOD evaluation web site.

TAPM was configured for the Bowline dataset using nested grids of $25 \times 25 \times 25$ points down to a resolution of 1000 m for meteorology, and using nested grids of $21 \times 21 \times 21$ points down to a resolution of 100 m for pollution. Default model settings and databases, with the exception of the NCEP synoptic reanalyses data, were used. The Lagrangian particle module was used to represent near-source dispersion. TAPM was run without local meteorological observations (i.e. meteorological data assimilation was not used). Building dimensions and characteristics for five building blocks were used by TAPM.

Wind and temperature data from the 100 m level of a nearby (on-site) tower are compared to TAPM predicted meteorology in Table 3.1, with statistical results (see the Appendix) showing little bias, low RMSE, and high IOA for winds (averaged over wind speed and the horizontal wind components) and temperature of 0.87 and 0.98 respectively.

Table 3.1. Statistics for TAPM simulation of Bowline 1981 tower data at 100 m for wind speed at 100 m above the ground (WS); the west-east-component of the wind (U); the south-north-component of the wind (V); and temperature (T).

VARIABLE	NUMBER	MEAN_OBS	MEAN_MOD	STD_OBS	STD_MOD	CORR	RMSE	RMSE_S	RMSE_U	IOA	SKILL_E	SKILL_V	SKILL_R
WS	8294	4.7	4.6	2.8	2.2	0.70	2.04	1.29	1.58	0.82	0.56	0.78	0.72
U	8294	1.7	2.0	3.6	2.9	0.81	2.11	1.25	1.70	0.89	0.48	0.81	0.59
V	8294	-1.0	-0.7	3.7	3.7	0.83	2.16	0.73	2.03	0.91	0.55	0.99	0.58
T	8086	11.4	11.0	9.9	10.0	0.97	2.51	0.45	2.47	0.98	0.25	1.01	0.25

KEY: OBS = Observations, MOD = Model Predictions, NUMBER = Number of hourly-averaged values used for the statistics, MEAN = Arithmetic mean, STD = Standard Deviation, CORR = Pearson Correlation Coefficient (0 = no correlation, 1 = exact correlation), RMSE = Root Mean Square Error, RMSE_S = Systematic Root Mean Square Error, RMSE_U = Unsystematic Root Mean Square Error, IOA = Index of Agreement (0 = no agreement, 1 = perfect agreement), SKILL_E = (RMSE_U)/(STD_OBS) (<1 shows skill), SKILL_V = (STD_MOD)/(STD_OBS) (near to 1 shows skill), SKILL_R = (RMSE)/(STD_OBS) (<1 shows skill).

Figure 3.1 summarises the extreme concentrations (RHC and MAX) for each of the monitoring sites, and for the average (AVG) and RMSE over all sites, for observations versus TAPM. The results show that TAPM performs acceptably, with good prediction of AVG, and moderately low RMSE, although the model shows a slight under-prediction (~20%), mainly as a result of some under-prediction at site 1 (the site furthest from the facility).
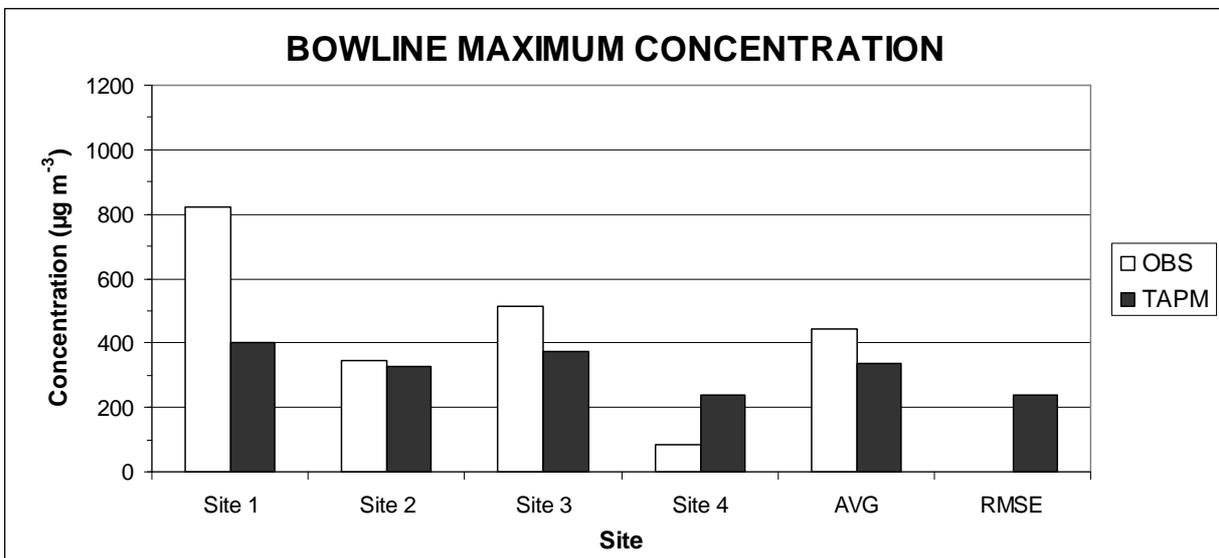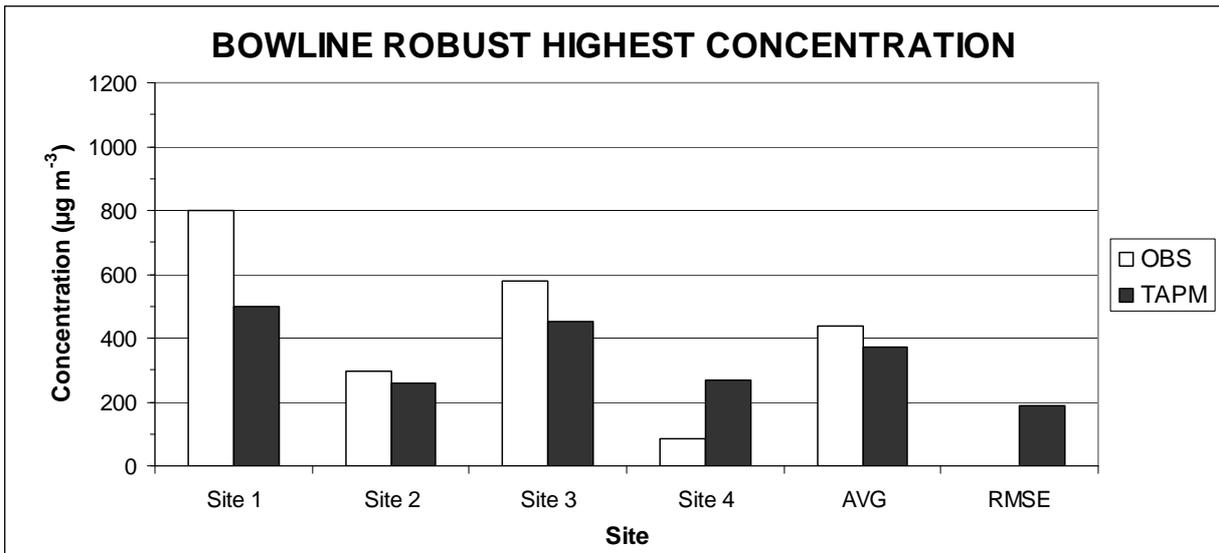
BOWLINE ROBUST HIGHEST CONCENTRATION

BOWLINE MAXIMUM CONCENTRATION

Figure 3.1. Bowline annual extreme concentration (μg m^{-3}) at each of the monitoring sites for the observed (OBS – white bars) concentrations and those predicted by TAPM for RHC (top) and MAX (bottom). Note that AVG is an average over all sites and RMSE is the root mean square error over all sites. The monitoring sites are numbered from 1 to 4, with approximate locations relative to the mid-point of the two stacks: Site 1 850 m to the southeast; Site 2 250 m to the south; Site 3 400 m to the southeast; and Site 4 850 m to the west-southwest.

4   Lovett

The Lovett 1988 annual dataset consists of SO_2 measurements at a number of monitoring sites situated mainly on Dunderberg Mountain to the north of the Lovett power station in southeastern New York state (U.S.) (Paumier et al., 1992). The difference in terrain height between the base of the Lovett stack and the highest sites on the mountain was approximately 300 m, and the monitoring sites ranged in distance from 2–4 km from the stack. Peak concentrations at the monitoring sites occurred under southerly winds and mainly during stable conditions at night, although some daytime events were also observed. Source characteristics, emission data and monitoring data were obtained from the Lovett dataset on the AERMOD evaluation web site.

TAPM was run for the Lovett dataset using nested grids of $25 \times 25 \times 25$ points down to a resolution of 300 m for both meteorology and pollution, using default model settings and databases, with the exception of 3-second resolution USGS terrain data and NCEP synoptic reanalyses data. The Lagrangian particle module was used to represent near-source dispersion. Note that the simulation of these datasets was performed without using local meteorological observations (i.e. meteorological data assimilation was not used).

Wind and temperature data from the 100 m level of a nearby (on-site) tower are statistically compared to TAPM predicted meteorology in Table 4.1, with acceptable results considering the complexity of the local terrain – the index of agreement for winds (averaged over wind speed and the horizontal wind components) and temperature was 0.77 and 0.98 respectively. Interestingly, the mean model wind speed was over-predicted by 1.3 m s^{-1}, but this hasn't translated to poor performance for extreme concentrations (see below).

Table 4.1. Statistics for TAPM simulation of Lovett 1988 tower data at 100 m for wind speed at 100 m above the ground (WS); the west-east-component of the wind (U); the south-north-component of the wind (V); and temperature (T).

VARIABLE	NUMBER	MEAN_OBS	MEAN_MOD	STD_OBS	STD_MOD	CORR	RMSE	RMSE_S	RMSE_U	IOA	SKILL_E	SKILL_V	SKILL_R
WS	8665	3.5	4.8	2.2	2.3	0.50	2.61	1.67	2.00	0.67	0.91	1.05	1.19
U	8665	1.3	2.5	2.7	2.9	0.66	2.61	1.49	2.15	0.78	0.79	1.05	0.96
V	8665	-0.4	-0.6	2.8	3.7	0.76	2.42	0.17	2.42	0.85	0.86	1.31	0.86
T	8694	10.4	10.8	10.2	10.7	0.96	2.87	0.37	2.84	0.98	0.28	1.05	0.28

KEY: OBS = Observations, MOD = Model Predictions, NUMBER = Number of hourly-averaged values used for the statistics, MEAN = Arithmetic mean, STD = Standard Deviation, CORR = Pearson Correlation Coefficient (0 = no correlation, 1 = exact correlation), RMSE = Root Mean Square Error, RMSE_S = Systematic Root Mean Square Error, RMSE_U = Unsystematic Root Mean Square Error, IOA = Index of Agreement (0 = no agreement, 1 = perfect agreement), SKILL_E = (RMSE_U)/(STD_OBS) (<1 shows skill), SKILL_V = (STD_MOD)/(STD_OBS) (near to 1 shows skill), SKILL_R = (RMSE)/(STD_OBS) (<1 shows skill).

Figure 4.1 summarises the extreme concentrations (RHC and MAX) for each of the monitoring sites, and for the average (AVG) and RMSE over all sites, for observations versus TAPM. It is clear from these plots that TAPM performs well with good average (AVG) and small RMSE for each of the statistics.
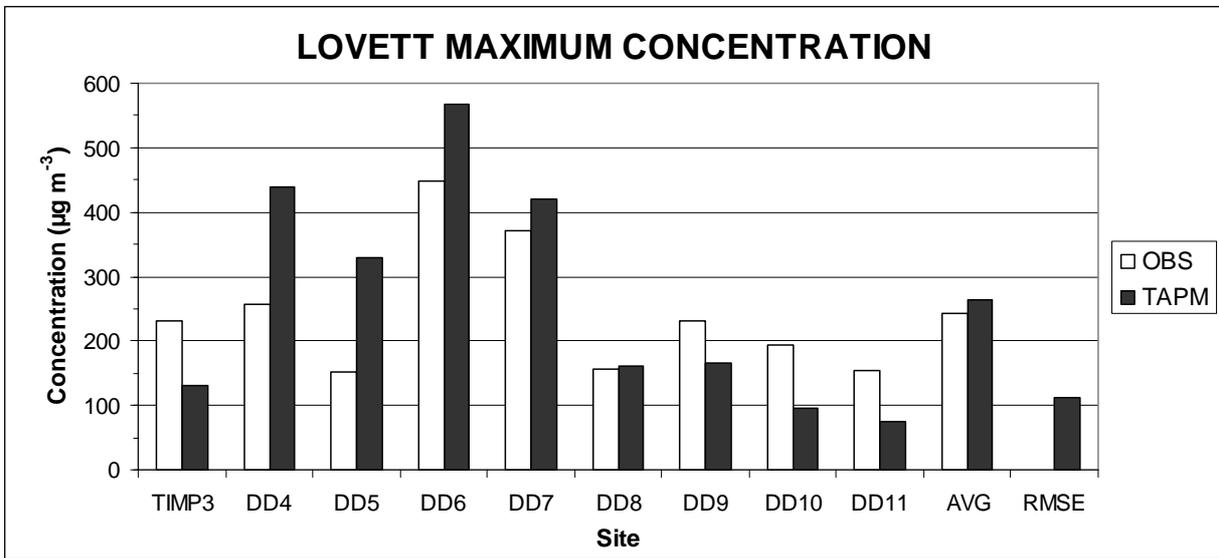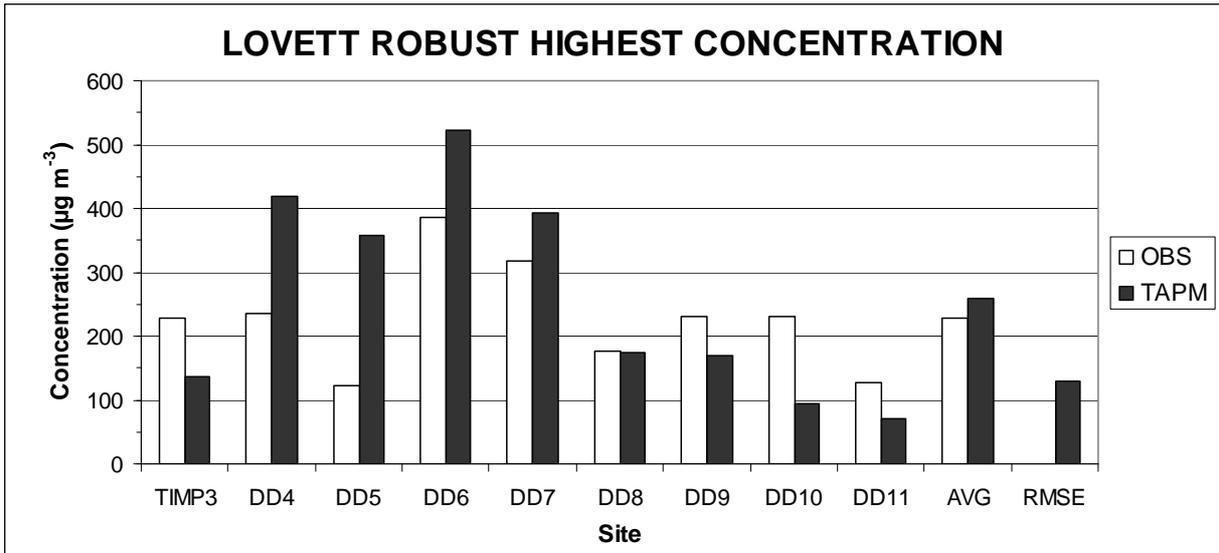
Figure 4.1. Lovett annual extreme concentration ($\mu g\,m^{-3}$) of sulfur dioxide at each of the monitoring sites for the observed (OBS – white bars) concentrations and those predicted by TAPM (black bars) for RHC (top) and MAX (bottom). Note that AVG is an average over all sites and RMSE is the root mean square error over all sites. The monitoring sites are numbered from 3 to 11, with TIMP3 located on The Timp and DD4-DD11 located on Dunderberg Mountain.

5 Westvaco

The Westvaco December 1980 – November 1981 annual dataset consists of SO_2 measurements at a number of monitoring sites situated mainly on a mountain range to the southeast of the Westvaco power station in Maryland state (U.S.) (Strimaitis et al., 1987). The difference in terrain height between the base of the Westvaco stack and the highest sites on the mountain was approximately 350 m, and these monitoring sites ranged in distance from 0.8–1.6 km from the stack. Peak concentrations at the monitoring sites occurred under northwesterly winds and stable conditions at night. There were three other monitoring sites (S02, S10 and S11 – see the key in Figure 5.1 for site names) that were at varying distances to the north, northeast and northwest of the power station respectively. Source characteristics, emission data and monitoring data were obtained from the Westvaco dataset on the AERMOD evaluation web site.

TAPM was run for the Westvaco dataset using nested grids of $25 \times 25 \times 25$ points down to a resolution of 300 m for both meteorology and pollution, using default model settings and databases, with the exception of 3-second resolution USGS terrain data and NCEP synoptic reanalyses data. The Lagrangian particle module was used to represent near-source dispersion. Note that the simulation of these datasets was performed without using local meteorological observations (i.e. meteorological data assimilation was not used).

Wind and temperature data from the 100 m level of a nearby (on-site) tower are statistically compared to TAPM predicted meteorology in Table 5.1, with good results, especially considering the complexity of the local terrain – the index of agreement for winds (averaged over wind speed and the horizontal wind components) and temperature was 0.87 and 0.98 respectively.

Table 5.1. Statistics for TAPM simulation of Westvaco tower data at 100 m for wind speed at 100 m above the ground (WS); the west-east-component of the wind (U); the south-north-component of the wind (V); and temperature (T).

VARIABLE	NUMBER	MEAN_OBS	MEAN_MOD	STD_OBS	STD_MOD	CORR	RMSE	RMSE_S	RMSE_U	IOA	SKILL_E	SKILL_V	SKILL_R
WS	8568	6.3	6.8	3.7	3.4	0.75	2.56	1.23	2.25	0.86	0.61	0.92	0.69
U	8568	3.9	5.5	5.7	4.5	0.87	3.23	2.35	2.22	0.90	0.39	0.79	0.57
V	8568	-0.1	0.0	2.4	2.8	0.74	1.93	0.32	1.91	0.85	0.79	1.19	0.80
T	7787	9.2	10.6	10.4	10.8	0.98	2.65	1.36	2.27	0.98	0.22	1.03	0.25

KEY: OBS = Observations, MOD = Model Predictions, NUMBER = Number of hourly-averaged values used for the statistics, MEAN = Arithmetic mean, STD = Standard Deviation, CORR = Pearson Correlation Coefficient (0 = no correlation, 1 = exact correlation), RMSE = Root Mean Square Error, RMSE_S = Systematic Root Mean Square Error, RMSE_U = Unsystematic Root Mean Square Error, IOA = Index of Agreement (0 = no agreement, 1 = perfect agreement), SKILL_E = (RMSE_U)/(STD_OBS) (<1 shows skill), SKILL_V = (STD_MOD)/(STD_OBS) (near to 1 shows skill), SKILL_R = (RMSE)/(STD_OBS) (<1 shows skill).

Figure 5.1 summarises the extreme concentrations (RHC and MAX) for each of the monitoring sites, and for the average (AVG) and RMSE over all sites, for observations versus TAPM. It is clear from these plots that TAPM under predicts the AVG by about 45% for the RHC, while the model performs better for the MAX with good average (AVG) and small

RMSE for each of the statistics, although there is still an overall under-prediction by about 25%. The level of model underprediction shown by TAPM is typical of other dispersion models for the Westvaco dataset when site-by-site statistics (as opposed to site-independent statistics that are commonly used) are considered.
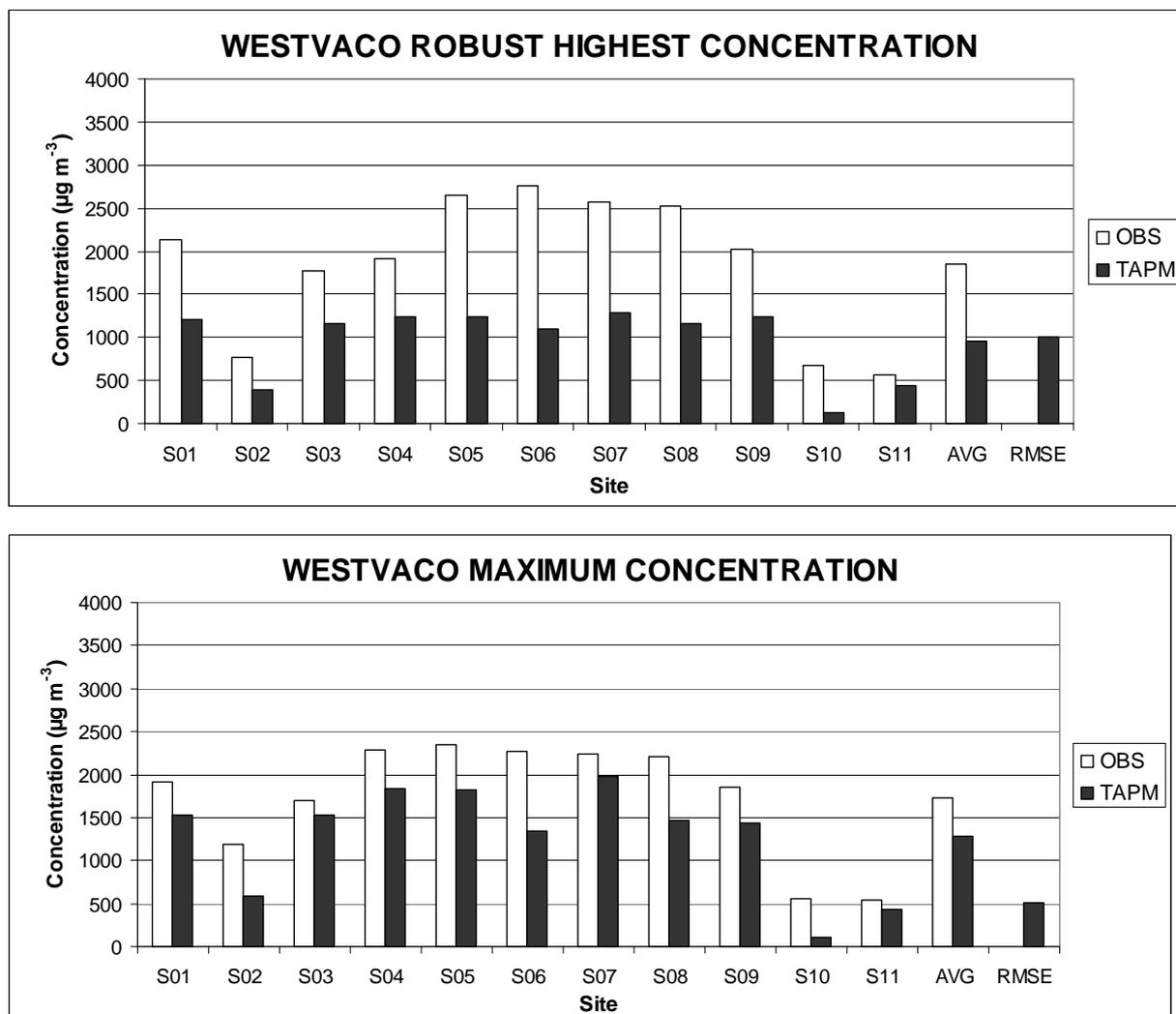

Figure 5.1. Westvaco annual extreme concentration (μg m^{-3}) of sulfur dioxide at each of the monitoring sites for the observed (OBS – white bars) concentrations and those predicted by TAPM (black bars) for RHC (top) and MAX (bottom). Note that AVG is an average over all sites and RMSE is the root mean square error over all sites. The monitoring sites are: MET TOWER (S01), LUKE HILL (S02), FOLLY (S03), TREEHOUSE (S04), SLIPPERY HILL (S05), RIDGE END (S06), OVERLOOK (S07), SOUTH SIDE (S08), COW PASTURE (S09), STONY RUN (S10) and BLOOMINGTON (S11).

6 Anglesea

Alcoa of Australia operates a 160 MW power station within the Anglesea Heath in Victoria. This area comprises the catchment of two creeks and forms a large basin mostly surrounded by higher ground. It is well used for recreation and has high conservation value due to outstanding botanical diversity. The station has been a significant point source of sulfur emissions and a potential source of acidic deposits to the surrounding ecosystems since 1968. The power station is located at an elevation of 10 m within the naturally vegetated Anglesea River catchment. It sits, in effect, near a break in the side of a sloping 8,500 ha basin mostly surrounded by higher ground. Stack height is 110 m, which puts the point of emission roughly on a level with the edge of the basin, 3 km to the north and south. The sea is 3 km to the east. This Section summarises the performance of TAPM from the latter study, compared to monitoring data at two sites: School – located to the south of the power station; and Mt. Ingoldsby – located south-west of the power station.

TAPM was run for the Anglesea dataset using nested grids of $25 \times 25 \times 25$ points down to a resolution of 300 m for both meteorology and pollution, using default model settings and databases, with the exception of local high resolution databases of terrain and land use. The Lagrangian particle module was used to represent near-source dispersion. Note that the simulation of these datasets was performed without using local meteorological observations (i.e. meteorological data assimilation was not used). The power station stack emission characteristics input into the model included time-varying exit velocity, exit temperature and sulfur dioxide emission rate (calculated from coal use, moisture, ash and sulphur content and hourly load). The load-based emission technique used for this study has been verified to be accurate to within approximately 10% compared to in-stack emission measurements.

Figure 6.1 summarises the extreme concentrations for each of the monitoring sites and years, and for the average (ALL) and RMSE over all sites, for observations versus TAPM. It is clear from these plots that TAPM performs well with good average (ALL) and small RMSE for RHC, although the model over-predicts the RHC and especially the MAX for the Mt. Ingoldsby site in 2002 (ING02).
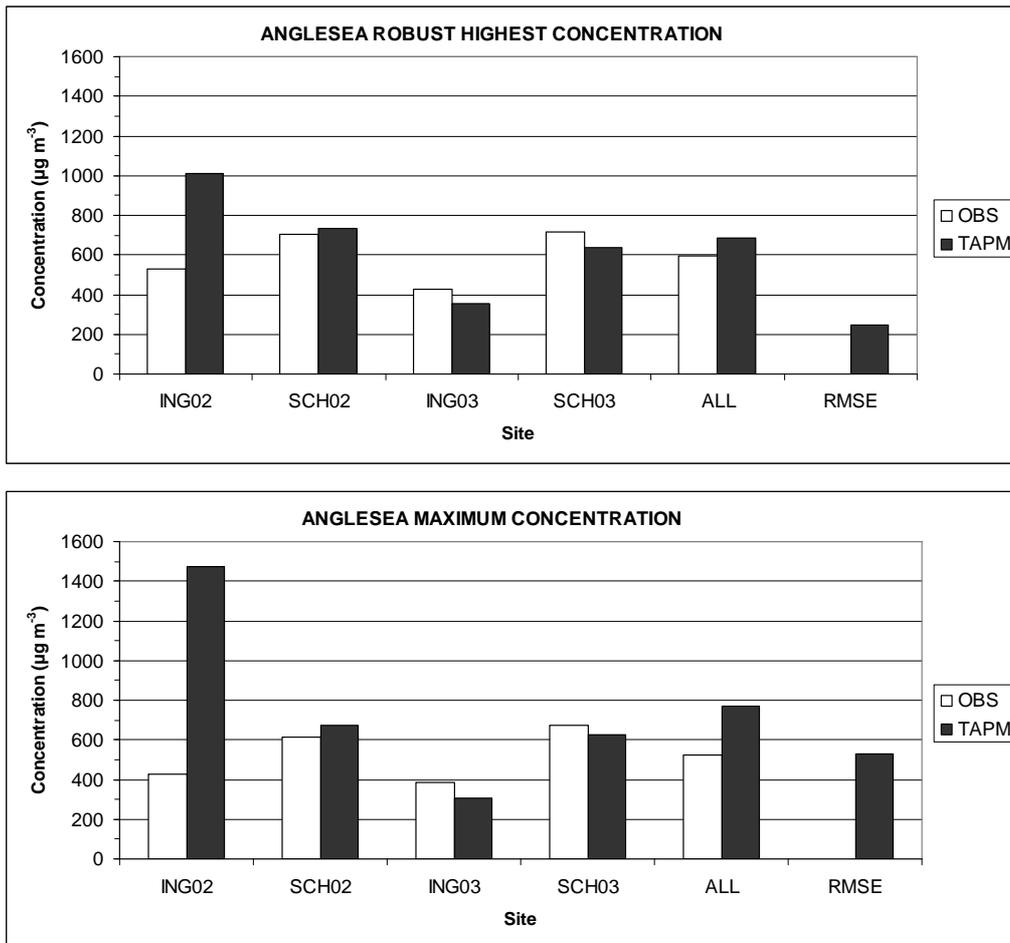
Figure 6.1: Anglesea annual extreme concentration ($\mu g\, m^{-3}$) at the School (SCH) and Mt. Ingoldsby (ING) monitoring sites for 2002 and 2003 for the observed (OBS – white bars) concentrations and those predicted by TAPM (black bars) for RHC (top), and MAX (bottom). Note that ALL is an average over all sites and RMSE is the root mean square error over all sites.

7 Kwinana

Kwinana is a major heavy industrial area 30 km south of Perth, Western Australia. It is a coastal area with sea to the west (Indian Ocean) and land to the east, an approximately north-south coastline, and relatively flat local terrain. The Kwinana region includes industries such as power generation, refineries (oil, alumina and nickel), iron smelting, cement works, and titanium dioxide and fertilizer plants. Most of the twenty point sources are on the coast, and they have plume heights that vary from tens of metres up to a few hundred metres. The plumes generally fumigate to ground in the strong south-westerly sea-breeze flow, resulting in relatively high concentrations at distances of up to several kilometres from the coast.

The Kwinana Industries Council (KIC) was established in the course of Policy development to provide a forum for negotiations between industries and to form a single body to represent industry's viewpoint. The strong regulatory framework, and government and industry cooperation, have produced a high quality hour-by-hour emissions inventory for the industrial sources in the region. The Kwinana air monitoring network, measuring both near-surface meteorology and air pollution, was designed to capture the maximum ground level concentrations outside of the industrial zone boundaries, and near local towns, under sea-breeze fumigation conditions. The detailed emission inventory, coupled with the extensive monitoring network data, provide an excellent framework for model development and/or verification.

Table 7.1. Statistics for TAPM simulation of 1997 in Kwinana (Hope Valley) for 10 m wind speed (WS), the west-east-component of the wind (U), the south-north-component of the wind (V) and temperature (T).

VARIABLE	NUMBER	MEAN_OBS	MEAN_MOD	STD_OBS	STD_MOD	CORR	RMSE	RMSE_S	RMSE_U	IOA	SKILL_E	SKILL_V	SKILL_R
WS	8482	4.0	3.7	1.7	1.7	0.72	1.29	0.55	1.17	0.84	0.69	1.00	0.76
U	8482	-0.2	-0.5	3.2	3.1	0.87	1.64	0.54	1.55	0.93	0.49	0.99	0.52
V	8482	1.1	0.5	2.8	2.5	0.83	1.63	0.88	1.37	0.90	0.50	0.90	0.59
T	8728	18.2	18.2	5.2	5.0	0.94	1.79	0.53	1.71	0.97	0.33	0.96	0.34

KEY: OBS = Observations, MOD = Model Predictions, NUMBER = Number of hourly-averaged values used for the statistics, MEAN = Arithmetic mean, STD = Standard Deviation, CORR = Pearson Correlation Coefficient (0 = no correlation, 1 = exact correlation), RMSE = Root Mean Square Error, RMSE_S = Systematic Root Mean Square Error, RMSE_U = Unsystematic Root Mean Square Error, IOA = Index of Agreement (0 = no agreement, 1 = perfect agreement), SKILL_E = (RMSE_U)/(STD_OBS) (<1 shows skill), SKILL_V = (STD_MOD)/(STD_OBS) (near to 1 shows skill), SKILL_R = (RMSE)/(STD_OBS) (<1 shows skill).

TAPM was used to model 1997 meteorology in Kwinana using nested grids of $25 \times 25 \times 25$ points at 30-km, 10-km, 3-km and 1-km spacing for meteorology, and $41 \times 41 \times 25$ points at 15-km, 5-km, 1.5-km and 500-m spacing for pollution. Default model options were used. Tracer mode was used for sulfur dioxide, with the four dominant point sources (of the twenty point sources in the region) run in Lagrangian (LPM) mode. More detail on the Kwinana region can be found in Hurley et al. (2001).

Model predictions were extracted at the nearest grid point to the Hope Valley monitoring site on the inner grid (1-km grid spacing) at the lowest model level (10 m above the ground) for winds and temperature. Performance statistics are given in Table 7.1. The statistics used were

based on the recommendations of Willmott (1981), as described in the Appendix. The results suggest that both winds and temperature are predicted well.

Figure 7.1 summarises the extreme concentrations (RHC and MAX) for each of the monitoring sites and years, and for the average (ALL) and RMSE over all sites, for observations versus TAPM. It is clear from these plots that TAPM performs well with very good average (ALL) and small RMSE for each of the statistics.
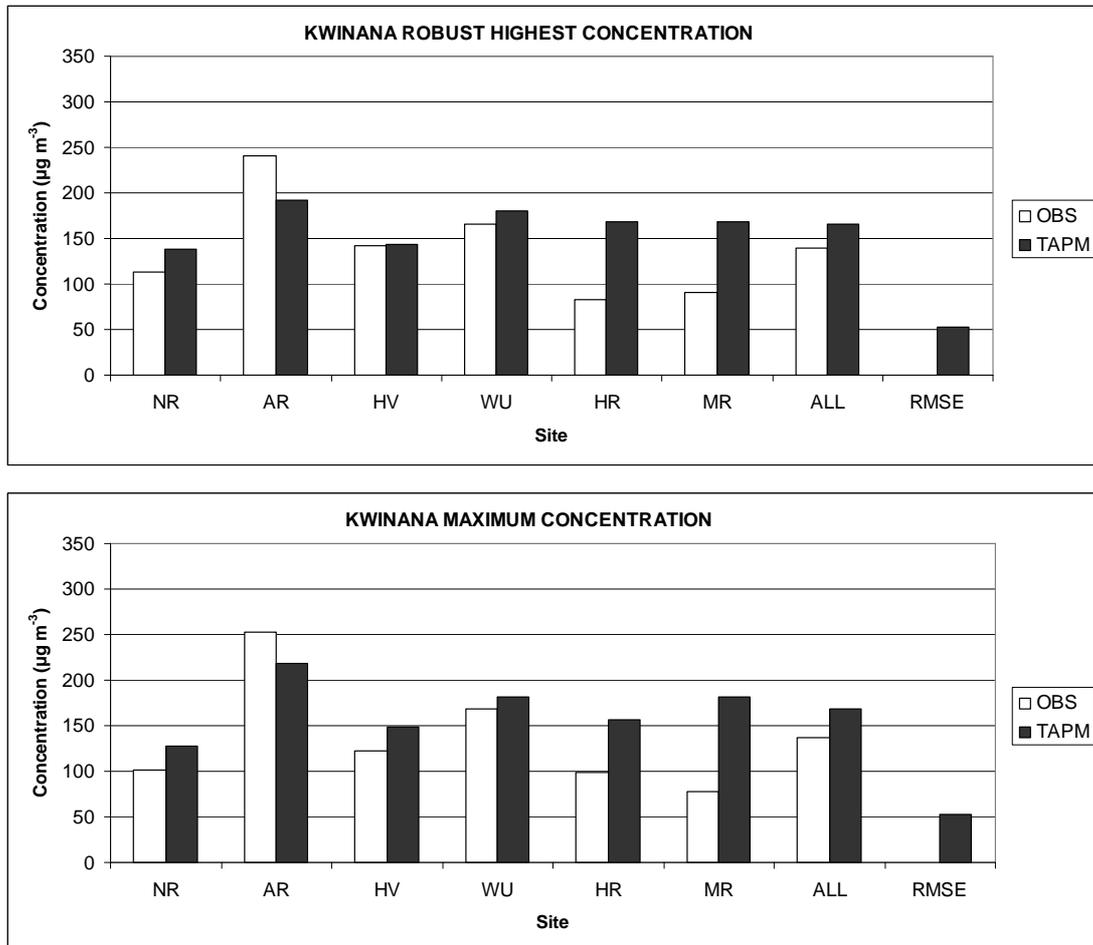

Figure 7.1: Kwinana annual extreme concentration ($\mu g\ m^{-3}$) at all monitoring sites for 1997 for the observed (OBS) concentrations and those predicted by TAPM for RHC (top), and MAX (bottom). Note that ALL is an average over all sites and RMSE is the root mean square error over all sites.

8   Upper-Air Meteorology in Kalgoorlie

The Kalgoorlie region of southern Western Australia is an inland industrial region with sparsely vegetated, relatively flat terrain. Western Mining Corporation (WMC) has a SODAR system in the region, which is used as part of a reactive pollution control strategy to minimise the impact of industrial emissions on the local township. Edwards et al. (2004) modelled year-long meteorology in this region for 2000, and compared results with upper-level winds from a tower and a SODAR. Results were presented for annual, seasonal and diurnal statistics, as well as for a case study when a cold front passed through the region. A summary of the updated annual results for Kalgoorlie using the current version of TAPM are presented here.

TAPM was used to model the meteorology of the year 2000 in the Kalgoorlie region using a nested grid of $25 \times 25 \times 25$ points at 30-km, 10-km, 3-km and 1-km horizontal grid spacing, and 25 vertical levels. Default model options were used. Model predictions of hourly-averaged winds at model levels closest to the SODAR levels were extracted at the nearest grid point to the site for the 1-km spaced inner grid. The corresponding SODAR levels were at 50, 110, 140, 200, 260, 290, 410, 500 and 590 m, and were all within 10 m of the nearest model level. Note that the number of valid data points was reduced by almost a factor of two by the 590-m height level, and by even more above this level where the data were thought to be outside the valid range of the instrument, and so were not used. Typically, this instrument is less reliable at heights above 400 m.

Tables 11.1a–c show statistics for observed (OBS) versus model predicted (MOD) winds compared to the SODAR data. The results show that the SODAR winds are modelled very well at all heights. The RMSE increases from 1.7 m s^{-1} for the 50-m level to 2.9 m s^{-1} at the 600-m level for wind speed (slightly higher for the wind components), and the average IOA are 0.88 for wind speed, 0.97 for the west-east (u) component and 0.94 for the south-north (v) component.

Table 11.1a. Wind speed (m s^{-1}) statistics for TAPM simulation of 2000 at the SODAR site.

HEIGHT	NUMBER	MEAN_OBS	MEAN_MOD	STD_OBS	STD_MOD	CORR	RMSE	RMSE_S	RMSE_U	IOA	SKILL_E	SKILL_V	SKILL_R
50 m	7846	5.6	4.8	2.1	1.9	0.73	1.69	1.06	1.32	0.82	0.63	0.92	0.80
100 m	7850	6.7	6.1	2.6	2.2	0.75	1.85	1.11	1.48	0.85	0.57	0.86	0.71
150 m	7713	7.1	7.1	2.9	2.7	0.77	1.88	0.79	1.70	0.88	0.59	0.94	0.65
200 m	7263	7.8	8.0	3.3	3.2	0.78	2.14	0.83	1.97	0.88	0.60	0.97	0.65
250 m	6762	8.3	8.6	3.7	3.6	0.81	2.25	0.76	2.12	0.90	0.58	0.99	0.61
300 m	6486	8.5	9.1	3.8	4.0	0.82	2.44	0.77	2.31	0.90	0.60	1.04	0.63
400 m	5514	9.1	9.6	4.2	4.5	0.82	2.62	0.68	2.53	0.90	0.60	1.06	0.62
500 m	4935	9.6	9.9	4.4	4.7	0.82	2.74	0.59	2.68	0.90	0.60	1.06	0.62
600 m	4483	10.2	10.0	4.6	4.9	0.81	2.93	0.66	2.85	0.90	0.62	1.06	0.64

KEY: OBS = Observations, MOD = Model Predictions, MEAN = Arithmetic mean, STD = Standard Deviation, CORR = Pearson Correlation Coefficient (0=no correlation,1=exact correlation), RMSE = Root Mean Square Error, RMSE_S = Systematic Root Mean Square Error, RMSE_U = Unsystematic Root Mean Square Error, IOA = Index of Agreement (0=no agreement, 1=perfect agreement), SKILL_E = (RMSE_U)/(STD_OBS) (<1 shows skill), SKILL_V = (STD_MOD)/(STD_OBS) (near to 1 shows skill), SKILL_R = (RMSE)/(STD_OBS) (<1 shows skill).

Table 11.1b. West-east (*u*) component of the wind (m s^{-1}) statistics for TAPM simulation of 2000 at the SODAR site.

	NUMBER	MEAN_OBS	MEAN_MOD	STD_OBS	STD_MOD	CORR	RMSE	RMSE_S	RMSE_U	IOA	SKILL_E	SKILL_V	SKILL_R
50 m	7846	-1.1	-0.8	4.6	4.1	0.93	1.77	0.89	1.53	0.96	0.33	0.88	0.39
100 m	7850	-1.3	-1.1	5.5	5.0	0.93	2.03	0.82	1.86	0.96	0.34	0.92	0.37
150 m	7713	-1.4	-1.4	5.8	5.8	0.93	2.11	0.36	2.08	0.97	0.36	1.00	0.36
200 m	7263	-1.1	-1.6	6.4	6.6	0.94	2.33	0.52	2.27	0.97	0.36	1.04	0.37
250 m	6762	-1.5	-1.8	6.9	7.2	0.94	2.40	0.31	2.38	0.97	0.35	1.05	0.35
300 m	6486	-1.5	-1.9	7.1	7.7	0.95	2.54	0.39	2.51	0.97	0.35	1.08	0.36
400 m	5514	-1.5	-1.8	7.8	8.3	0.95	2.72	0.33	2.70	0.97	0.35	1.07	0.35
500 m	4935	-1.4	-1.5	8.3	8.7	0.95	2.81	0.17	2.81	0.97	0.34	1.05	0.34
600 m	4483	-1.2	-1.2	8.8	9.0	0.94	2.97	0.33	2.96	0.97	0.34	1.02	0.34

KEY: Same as for Table 11.1a.

Table 11.1c. South-north (*v*) component of the wind (m s^{-1}) statistics for TAPM simulation of 2000 at the SODAR site.

HEIGHT	NUMBER	MEAN_OBS	MEAN_MOD	STD_OBS	STD_MOD	CORR	RMSE	RMSE_S	RMSE_U	IOA	SKILL_E	SKILL_V	SKILL_R
50 m	7846	0.6	0.6	3.6	3.1	0.87	1.81	0.97	1.53	0.92	0.42	0.84	0.50
100 m	7850	0.7	0.6	4.4	3.9	0.87	2.21	1.05	1.94	0.93	0.44	0.88	0.50
150 m	7713	0.7	0.5	4.8	4.6	0.87	2.41	0.83	2.26	0.93	0.47	0.96	0.50
200 m	7263	0.9	0.3	5.3	5.2	0.88	2.67	0.94	2.49	0.93	0.47	0.97	0.50
250 m	6762	0.5	0.2	5.8	5.6	0.89	2.67	0.79	2.55	0.94	0.44	0.98	0.46
300 m	6486	0.5	0.1	5.9	6.0	0.90	2.73	0.67	2.64	0.95	0.45	1.01	0.46
400 m	5514	0.6	0.1	6.2	6.2	0.90	2.84	0.77	2.73	0.95	0.44	1.01	0.46
500 m	4935	0.7	0.0	6.5	6.4	0.89	3.03	0.98	2.86	0.94	0.44	0.99	0.47
600 m	4483	0.7	0.0	6.7	6.5	0.88	3.31	1.23	3.07	0.93	0.46	0.97	0.49

KEY: Same as for Table 11.1a.

9 Annual urban meteorology and air pollution in Melbourne

Melbourne is a coastal city in the southern part of Victoria, Australia, with ocean to the south and mountains to the north. The EPA Victoria operates an air monitoring network covering the urban region of Melbourne, and measures both near-surface meteorology and air pollution. This Section summarises the results from TAPM modelling of year-long (July 1997 to June 1998) meteorology and photochemical air pollution in Melbourne (Hurley *et al.*, 2003a), which was an update/extension of the verification component of the work performed as part of the EPA Victoria Air Quality Improvement Plan.

TAPM was configured with three (nested) grids of $25 \times 25 \times 25$ points at 30-km, 10-km and 3-km spacing for meteorology, and $41 \times 41 \times 25$ points at 15-km, 5-km and 1.5-km spacing for pollution. Default model options were used. Land-use was dominated by urban characteristics for most of the nine Melbourne monitoring sites.

The emissions inventory covered Melbourne and Geelong, as well as major point sources in the Latrobe Valley about 100 km east of Melbourne. It represents emissions for all pollutants on approximately a 1-km spaced grid for vehicle, commercial and domestic emissions, as well as a point-source inventory, and a biogenic-emission inventory on a 3-km spaced grid (nitrogen oxides and hydrocarbons only). The emission inventory was developed for the Melbourne region by EPA Victoria for both general EPA use, and for use by the Australian Air Quality Forecasting System (AAQFS). Non-zero background concentrations were used for ozone (15 ppb), smog reactivity (Rsmog = 0.5 ppb) and particles (10 µg m^{-3} for PM$_{10}$ and 5 µg m^{-3} for PM$_{2.5}$, to account for dust and sea salt emissions not in the inventory), and the standard reactivity coefficient for Rsmog of 0.0067 was used for all VOC emissions.

Table 9.1. Statistics for TAPM simulation without meteorological data assimilation of July 1996 – June 1997 in Melbourne (averaged over eight sites) for wind speed at 10 m above the ground (WS); the west-east-component of the wind (U); the south-north-component of the wind (V); and screen-level temperature (T).

VARIABLE	NUMBER	MEAN_OBS	MEAN_MOD	STD_OBS	STD_MOD	CORR	RMSE	RMSE_S	RMSE_U	IOA	SKILL_E	SKILL_V	SKILL_R
WS	8295	3.3	2.7	1.9	1.4	0.75	1.37	1.00	0.93	0.81	0.49	0.74	0.74
U	8295	0.8	0.9	2.1	1.8	0.77	1.36	0.72	1.14	0.86	0.56	0.87	0.66
V	8295	0.3	0.0	3.0	2.3	0.87	1.55	1.06	1.12	0.91	0.37	0.76	0.52
T	8621	14.6	14.9	5.9	5.5	0.94	2.17	0.90	1.96	0.96	0.33	0.94	0.37

KEY: OBS = Observations, MOD = Model Predictions, NUMBER = Number of hourly-averaged values used for the statistics, MEAN = Arithmetic mean, STD = Standard Deviation, CORR = Pearson Correlation Coefficient (0 = no correlation, 1 = exact correlation), RMSE = Root Mean Square Error, RMSE_S = Systematic Root Mean Square Error, RMSE_U = Unsystematic Root Mean Square Error, IOA = Index of Agreement (0 = no agreement, 1 = perfect agreement), SKILL_E = (RMSE_U)/(STD_OBS) (<1 shows skill), SKILL_V = (STD_MOD)/(STD_OBS) (near to 1 shows skill), SKILL_R = (RMSE)/(STD_OBS) (<1 shows skill).

In order to examine the sensitivity of pollution results to predicted winds when simulating urban air pollution, the model was run twice in Hurley et al. (2003a). The first run was in normal mode without meteorological data assimilation, and the second run was in data assimilation mode using Melbourne and Geelong monitoring site 10-m level wind speed and

direction data. The second run used data assimilation site characteristics with a 20-km radius of influence, and assimilated the observed winds for the lowest two model levels (10 m and 25 m). The results showed that the inclusion of wind data assimilation for Melbourne did not improve annual statistics of pollution (e.g. mean, 99.9 percentile, robust highest concentration or maximum concentration). The results summarised below are for TAPM without wind data assimilation.

Model predictions of meteorology were extracted at the nearest grid point to each of the eight meteorological monitoring sites on the inner grid (3-km spacing) at 10 m above the ground for winds and at screen-level for temperature. Statistics of observations and model predictions are shown in Table 9.1, and are based on the recommendations of Willmott (1981) – see the Appendix for details. The results suggest that both winds and temperature are predicted well, with small biases, low RMSE and high IOA. The average RMSE and IOA for winds are 1.4 m s^{-1} and 0.86, and for temperature are 2.2 C and 0.96, which are very good results.
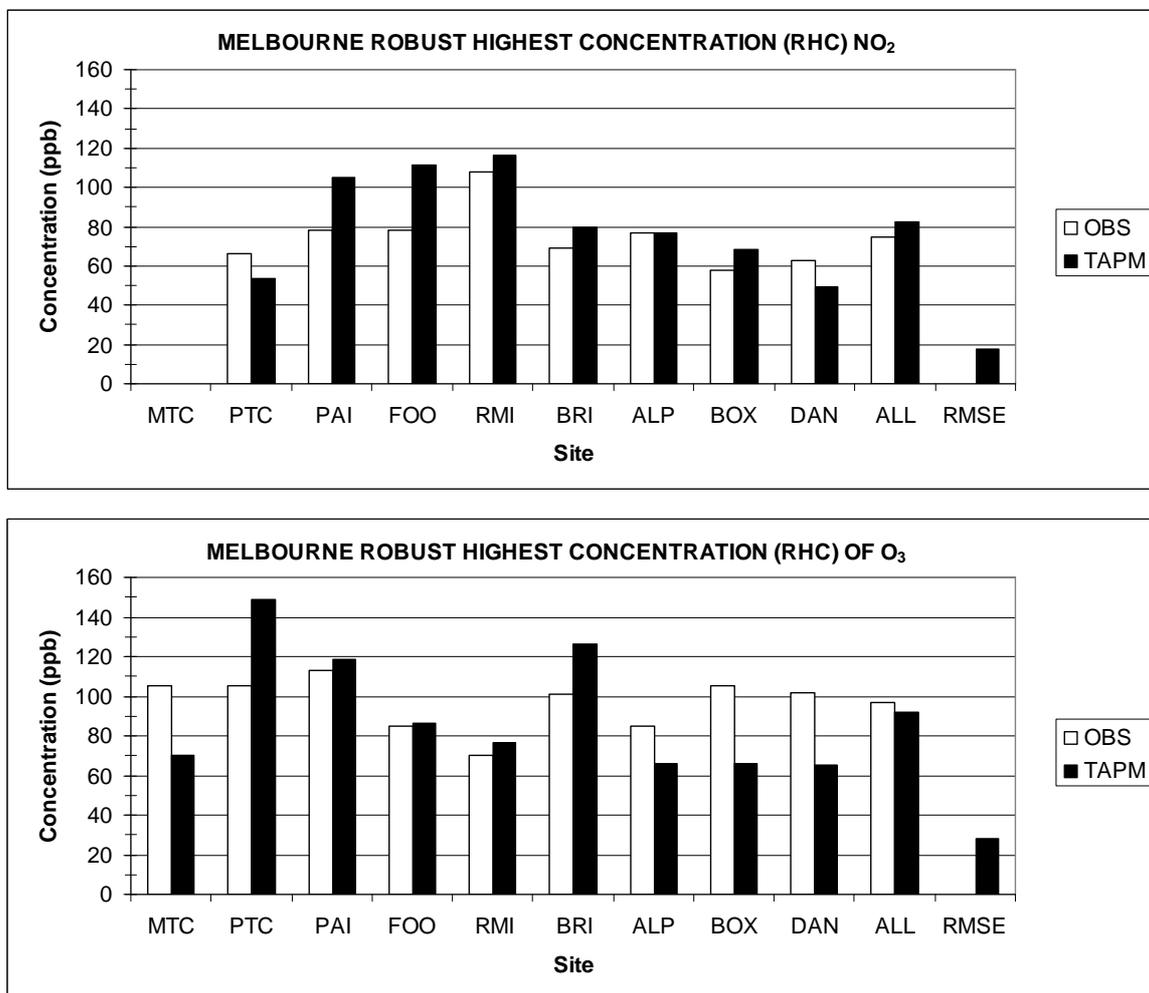
Figure 9.1. Nitrogen dioxide (top) and ozone (bottom) Robust Highest Concentration (RHC) values for each monitoring site and for the average over all sites, for TAPM. The EPAV monitoring sites correspond to: Alphington (ALP), Box Hill (BOX), Brighton (BRI), Dandenong (DAN), Footscray (FOO), Mt. Cottrell (MTC), Paisley (PAI), Pt. Cook (PTC) and RMIT (RMI).

Ground-level pollution results for hourly average NO_2 and O_3, and for daily average PM_{10} and $PM_{2.5}$, were extracted from the nearest grid point to each of the monitoring sites on the inner grid (1.5-km spacing), and are summarised for Robust Highest Concentration (RHC) in Figures 9.1 and 9.2 respectively. These results show that TAPM simulations of the extreme concentrations are very good for all species, although data at only three sites were available for $PM_{2.5}$.
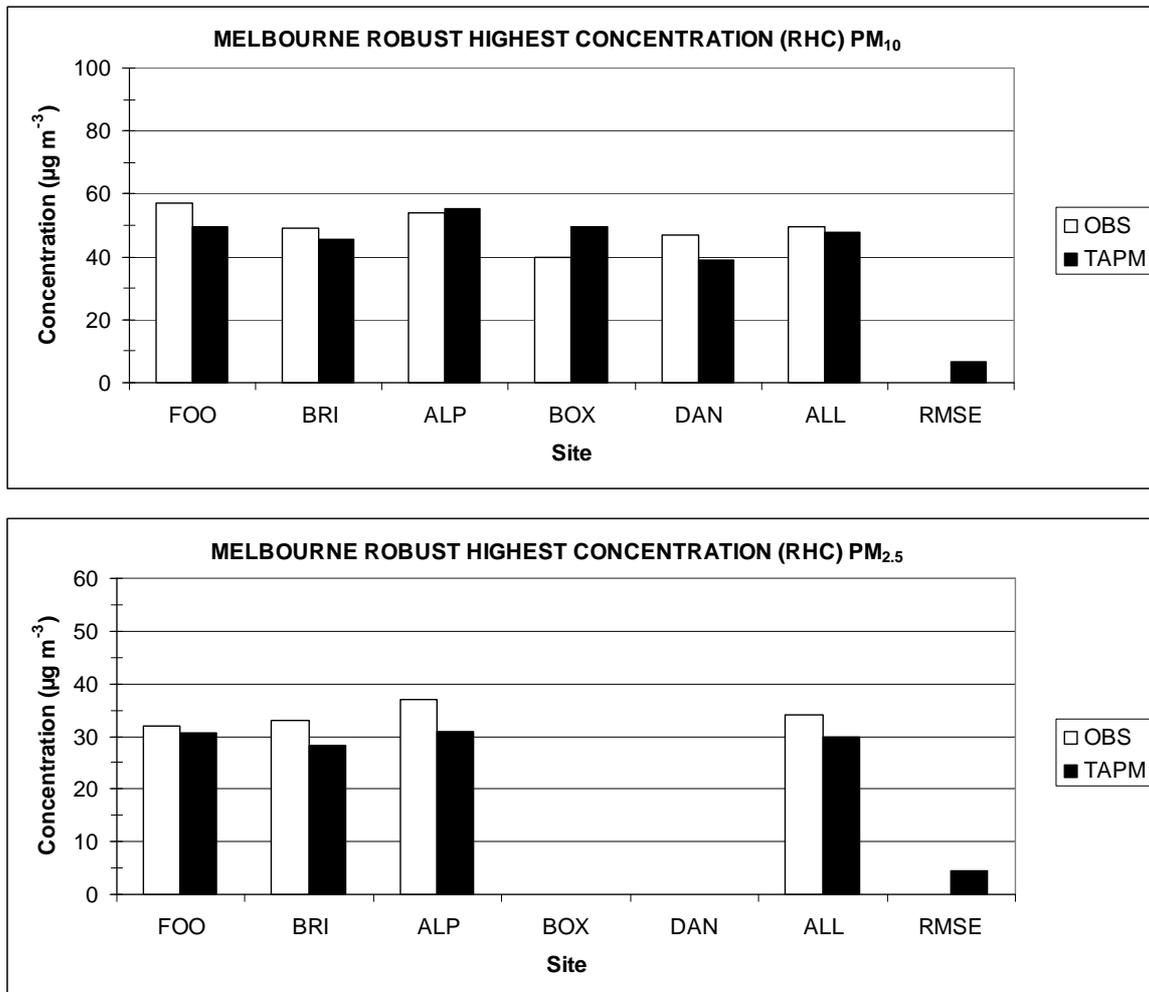

Figure 9.2. PM_{10} (top) and $PM_{2.5}$ (bottom) Robust Highest Concentration (RHC) values for each monitoring site and for the average over all sites, for TAPM. The EPAV monitoring sites correspond to: Alphington (ALP), Box Hill (BOX), Brighton (BRI), Dandenong (DAN), Footscray (FOO), Mt. Cottrell (MTC), Paisley (PAI), Pt. Cook (PTC) and RMIT (RMI).

23

10 Urban Meteorology

In order to explore the variation of TAPM performance in predicting urban meteorology, TAPM has been used to model the annual meteorology for urban sites is several Australian capital cities where data is available: Perth 1999, Brisbane 2001, Sydney 2003 and Melbourne 2003. For each region, the model was configured with three (nested) grids of $25 \times 25 \times 25$ points at 30-km, 10-km and 3-km spacing for meteorology. Default model options were used.

Model predictions of meteorology were extracted at the nearest grid point to each of the meteorological monitoring sites on the 3-km spaced grid at 10 m above the ground for winds and at screen-level for temperature. Statistics of observations and model predictions are shown in Table 10.1–10.4 for each city, and are based on the recommendations of Willmott (1981) – see the Appendix for details. The results suggest that the model predicts the winds and temperatures well for each of the cities, with overall results shown in Table 10.5 for the average over all cities indicating that both winds and temperature are predicted well, with no significant biases, low RMSE and high IOA. The average RMSE and IOA for winds are 1.4 m s^{-1} and 0.85, and for temperature are 2.2 C and 0.96, which are very good results.

The contribution of the urban scheme in TAPM to the results for Perth, Brisbane, Melbourne and Sydney, along with potential problems in site representation of monitors (especially for anemometers) and variability of urban characteristics not accounted for in the model resolution and input datasets in TAPM, contribute to uncertainty in wind predictions, and to a lesser extent for temperature predictions. In particular, AS2923 (1987) says that an anemometer (at 10 m) should be distant from any obstruction by at least 10 times the height of the obstruction, which can be difficult to achieve in urban areas. Given this complexity and uncertainty, we are encouraged by the results, especially given that the RMSE and IOA values for these urban regions are comparable to those obtained for non-urban regions for both winds and temperature.

Table 10.1. Statistics for TAPM simulation of 1999 in Perth (two sites) for wind speed at 10 m above the ground (WS); the west-east-component of the wind (U); the south-north-component of the wind (V); and screen-level temperature (T).

VARIABLE	NUMBER	MEAN_OBS	MEAN_MOD	STD_OBS	STD_MOD	CORR	RMSE	RMSE_S	RMSE_U	IOA	SKILL_E	SKILL_V	SKILL_R
WS	8301	3.9	3.5	1.9	1.6	0.72	1.44	0.89	1.13	0.83	0.59	0.85	0.75
U	8301	-0.5	-0.6	3.5	2.9	0.87	1.67	0.94	1.38	0.93	0.40	0.84	0.49
V	8301	0.8	0.4	2.5	2.4	0.85	1.39	0.60	1.25	0.91	0.50	0.95	0.56
T	8601	18.7	19.0	5.5	5.2	0.94	1.83	0.72	1.68	0.97	0.31	0.93	0.33

KEY: OBS = Observations, MOD = Model Predictions, NUMBER = Number of hourly-averaged values used for the statistics, MEAN = Arithmetic mean, STD = Standard Deviation, CORR = Pearson Correlation Coefficient (0 = no correlation, 1 = exact correlation), RMSE = Root Mean Square Error, RMSE_S = Systematic Root Mean Square Error, RMSE_U = Unsystematic Root Mean Square Error, IOA = Index of Agreement (0 = no agreement, 1 = perfect agreement), SKILL_E = (RMSE_U)/(STD_OBS) (<1 shows skill), SKILL_V = (STD_MOD)/(STD_OBS) (near to 1 shows skill), SKILL_R = (RMSE)/(STD_OBS) (<1 shows skill).

Table 10.2. Statistics for TAPM simulation of 2001 in Brisbane (three sites for wind and one site for temperature) for wind speed at 10 m above the ground (WS); the west-east-component of the wind (U); the south-north-component of the wind (V); and screen-level temperature (T).

VARIABLE	NUMBER	MEAN_OBS	MEAN_MOD	STD_OBS	STD_MOD	CORR	RMSE	RMSE_S	RMSE_U	IOA	SKILL_E	SKILL_V	SKILL_R
WS	8497	2.2	2.4	1.6	1.2	0.67	1.26	0.86	0.91	0.78	0.57	0.77	0.79
U	8497	-0.3	-0.3	2.0	1.9	0.80	1.29	0.53	1.16	0.89	0.59	0.98	0.65
V	8497	0.4	0.5	1.8	1.8	0.72	1.38	0.55	1.25	0.84	0.70	1.01	0.77
T	8679	21.3	21.4	5.0	4.7	0.94	1.75	0.61	1.64	0.97	0.33	0.94	0.35

KEY: As for Table above.

Table 10.3. Statistics for TAPM simulation of 2003 in Sydney (ten sites) for wind speed at 10 m above the ground (WS); the west-east-component of the wind (U); the south-north-component of the wind (V); and screen-level temperature (T).

VARIABLE	NUMBER	MEAN_OBS	MEAN_MOD	STD_OBS	STD_MOD	CORR	RMSE	RMSE_S	RMSE_U	IOA	SKILL_E	SKILL_V	SKILL_R
WS	8063	2.0	2.2	1.5	1.3	0.65	1.27	0.82	0.94	0.77	0.64	0.86	0.86
U	8063	0.1	0.3	1.8	1.9	0.73	1.46	0.51	1.33	0.83	0.78	1.13	0.85
V	8063	0.2	0.1	1.8	1.6	0.74	1.26	0.64	1.05	0.84	0.62	0.92	0.73
T	8524	17.0	17.9	6.0	5.6	0.90	2.76	1.29	2.42	0.94	0.41	0.95	0.47

KEY: As for Table above.

Table 10.4. Statistics for TAPM simulation of 2003 in Melbourne (eight sites) for wind speed at 10 m above the ground (WS); the west-east-component of the wind (U); the south-north-component of the wind (V); and screen-level temperature (T).

VARIABLE	NUMBER	MEAN_OBS	MEAN_MOD	STD_OBS	STD_MOD	CORR	RMSE	RMSE_S	RMSE_U	IOA	SKILL_E	SKILL_V	SKILL_R
WS	8419	3.4	2.9	1.9	1.5	0.77	1.37	0.93	0.98	0.84	0.50	0.78	0.70
U	8419	0.6	0.7	2.1	2.0	0.78	1.37	0.60	1.22	0.88	0.59	0.94	0.66
V	8419	0.0	0.0	3.2	2.6	0.88	1.56	0.95	1.20	0.92	0.38	0.81	0.49
T	8509	14.9	15.2	6.0	5.8	0.93	2.28	0.71	2.13	0.96	0.36	0.97	0.38

KEY: As for Table above.

Table 10.5. Statistics for TAPM simulations of four Australian cities for wind speed at 10 m above the ground (WS); the west-east-component of the wind (U); the south-north-component of the wind (V); and screen-level temperature (T).

VARIABLE	NUMBER	MEAN_OBS	MEAN_MOD	STD_OBS	STD_MOD	CORR	RMSE	RMSE_S	RMSE_U	IOA	SKILL_E	SKILL_V	SKILL_R
WS	8320	2.9	2.8	1.7	1.4	0.70	1.33	0.87	0.99	0.80	0.58	0.81	0.77
U	8320	0.0	0.0	2.3	2.2	0.80	1.44	0.64	1.27	0.88	0.59	0.97	0.66
V	8320	0.3	0.3	2.3	2.1	0.80	1.40	0.69	1.19	0.88	0.55	0.92	0.64
T	8578	18.0	18.4	5.6	5.3	0.93	2.16	0.83	1.97	0.96	0.35	0.95	0.38

KEY: As for Table above.

Acknowledgments

References

AS2923 (1987). 'Ambient Air – Guide for the Measurement of Horizontal Wind for Air Quality Applications', Standards Australia Limited.

Bowne N. E., Londergan R. J., Murray D. R. and Borenstein H. S. (1983). 'Overview, results and conclusions for the EPRI plume model validation project: plain site', EPRI Report EA-3074, EPRI, Palo Alto, CA 94304.

Cox W. and Tikvart J. (1990). 'A statistical procedure for determining the best performing air quality simulation', *Atmos. Environ.*, **24A**, 2387-2395.

Edwards M. , Hurley P. and Physick W. (2004). 'Verification of TAPM meteorological predictions using SODAR data in the Kalgoorlie region', *Aust. Met. Mag*, **53**, 29-37.

Hanna S. (1989). 'Confidence limits for air quality model evaluations, as estimated by bootstrap and jackknife resampling methods', *Atmos. Environ.*, **23**, 1385-1398.

Hurley P. (2008). 'TAPM V4. Part 1: Technical Description', CSIRO Marine and Atmospheric Research Paper No. 25. 59 pp.

Hurley P., Blockley A., and Rayner K. (2001). 'Verification of a prognostic meteorological and air pollution model for year-long predictions in the Kwinana region of Western Australia', *Atmos. Environ.*, **35**, 1871-1880.

Hurley P., Manins P., Lee S., Boyle R., Ng Y. and Dewundege P. (2003). 'Year-long, high-resolution, urban airshed modelling: Verification of TAPM predictions of smog and particles in Melbourne, Australia', *Atmos. Environ.*, **37**, 1899-1910.

Hurley P., Physick W. and Luhar A. (2005a). 'TAPM – A practical approach to prognostic meteorological and air pollution modelling', *Environmental Modelling & Software*, **20**, 737-752.

Hurley P., Edwards M., Physick W. and Luhar A. (2005b). 'TAPM V3 – Model description and verification', *Clean Air* **39**, 32–36.

Kalnay E. et al. (1996). 'The NCEP/NCAR 40-Year Reanalysis Project', *Bull. Amer. Meteorol. Soc.,* **77**, 437–71.

Luhar A., and Hurley P. (2003). 'Evaluation of TAPM, a prognostic meteorological and air pollution model, using urban and rural point-source data', *Atmos. Environ.*, **37**, 2795-2810.

Olesen H. R. (1995). 'The model validation exercise at Mol: overview of results', *Int. J. Environ. Pollut.* **5**, 761–84.

Paumier J.O., Perry S.G. and Burns D.J. (1992). 'CTDMPLUS: A dispersion model for sources near complex topography. Part II: Performance characteristics', *J. Appl. Meteor.*, **31**, 646-660.

Pielke R.A. (2002). 'Mesoscale Meteorological Modelling', Academic Press, 676 pp.

Schulman L.L. and Hanna S.R. (1986). 'Evaluation of downwash modifications to the Industrial Source Complex model', *JAPCA*, **36**, 258-264.

Strimaitis D.G., Paine R.J., Egan B.A., and Yamartino R.J. (1987). 'EPA complex terrain model development: Final report', U.S. EPA Report No. EPA/600/3-88/006, 486 pp.

Thatcher M. (2008). 'Processing global land surface datasets for dynamical downscaling with CCAM and TAPM', CSIRO Marine and Atmospheric Research report.

TRC (1986). 'Urban Power Plant Plume Studies', EPRI Report EA-5468, EPRI, 3412 Hillview Ave, Palo Alto, CA 94304.

Willmott C.J. (1981). 'On the Validation of Models', *Phys. Geography*, **2**, 184-194.

Appendix

The statistics used to measure meteorological model performance in this paper are based on those used by Willmott (1981) and Pielke (1984), as described below.

Predicted Mean $P_{mean} = \sum_{i=1}^{N} P_i$, where P_i are the predictions.

Observed Mean $O_{mean} = \sum_{i=1}^{N} O_i$, where O_i are the observations.

Predicted Standard Deviation $P_{std} = \sqrt{\dfrac{1}{N-1}\sum_{i=1}^{N}(P_i - P_{mean})^2}$.

Observed Standard Deviation $O_{std} = \sqrt{\dfrac{1}{N-1}\sum_{i=1}^{N}(O_i - O_{mean})^2}$.

Pearson Correlation Coefficient $r = \dfrac{N\left(\sum_{i=1}^{N} O_i P_i\right) - \left(\sum_{i=1}^{N} O_i\right)\left(\sum_{i=1}^{N} P_i\right)}{\sqrt{\left[N\left(\sum_{i=1}^{N} O_i^2\right) - \left(\sum_{i=1}^{N} O_i\right)^2\right]\left[N\left(\sum_{i=1}^{N} P_i^2\right) - \left(\sum_{i=1}^{N} P_i\right)^2\right]}}$.

Root Mean Square Error $RMSE = \sqrt{\dfrac{1}{N}\sum_{i=1}^{N}(P_i - O_i)^2}$.

Systematic Root Mean Square Error $RMSE_S = \sqrt{\dfrac{1}{N}\sum_{i=1}^{N}(\hat{P}_i - O_i)^2}$.

Unsystematic Root Mean Square Error $RMSE_U = \sqrt{\dfrac{1}{N}\sum_{i=1}^{N}(\hat{P}_i - P_i)^2}$.

Index of Agreement $IOA = 1 - \dfrac{\sum_{i=1}^{N}(P_i - O_i)^2}{\sum_{i=1}^{N}(|P_i - O_{mean}| + |O_i - O_{mean}|)^2}$.

Measures of Skill $SKILL_E = \dfrac{RMSE_U}{O_{std}}$, $SKILL_V = \dfrac{P_{std}}{O_{std}}$ and $SKILL_R = \dfrac{RMSE}{O_{std}}$.

Note that N is the number of observations and $\hat{P}_i = a + bO_i$ is the linear regression fitted formula with intercept (a) and slope (b).

The Robust Highest Concentration $RHC = C(R) + (\overline{C} - C(R))\ln((3R-1)/2)$ pollution statistic is from Cox and Tikvart (1990), with $C(R)$ the Rth highest concentration and $\overline{C}$ the mean of the top R-1 concentrations. The value of $R = 11$ is used here so that $\overline{C}$ is the average of the top-ten concentrations, which is an accepted statistic for evaluation of model performance (Hanna, 1989). The RHC is preferred to the actual peak value because it mitigates the undesirable influence of unusual events, while still representing the magnitude of the maximum concentration (unlike percentiles or averages over the top-percentiles). The statistical performance measures used here for pollution with the Kincaid and Indianapolis datasets are the same as those used by Olesen (1995) in the Model Validation Kit.
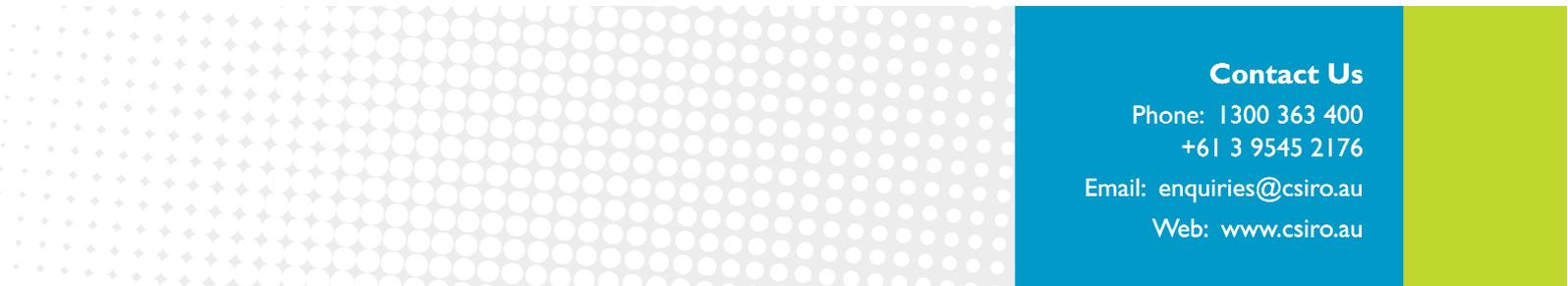
Your CSIRO

Australia is founding its future on science and innovation. Its national science agency, CSIRO, is a powerhouse of ideas, technologies and skills for building prosperity, growth, health and sustainability. It serves governments, industries, business and communities across the nation.