atmosphere

MDPI

Supplementary Materials

CCAM Model Performance Evaluation

Validation of the 2008 CCAM model is used to assess its ability to predict the meteorological conditions which drive the chemical transportation and transformation of pollutants in the NSW GMR. The key meteorological parameters used for the 2008 CCAM validation against a selection of NSW OEH and Bureau of Meteorology (BoM) monitoring stations (Table S1) are temperature and winds.

Table S1. OEH and BoM station locations.

Name	Site	Latitude	Longitude
Bargo	NSW OEH	–34.307	150.580
Bringelly	NSW OEH	–33.919	150.760
Chullora	NSW OEH	–33.894	151.050
Prospect	NSW OEH	–33.795	150.910
Randwick	NSW OEH	–33.933	151.240
Richmond	NSW OEH	–33.618	150.750
Wollongong	NSW OEH	–34.419	150.890
Newcastle	NSW OEH	–32.910	151.758
Badgerys Creek	BoM	–33.897	150.728
Bankstown Airport	BoM	–33.918	150.986
Camden Airport	BoM	–34.039	150.689
Richmond RAAF	BoM	–33.600	150.776
Sydney Airport	BoM	–33.947	151.173
Williamtown RAAF	BoM	–32.793	151.836

This is considered an operational evaluation where model estimates are compared to observations and deviations are quantified through statistical tests. The model performance for the entire period can be assessed with statistical metrics (see Table S2). For the CCAM validation modelling period the mean bias (MB—positive or negative deviation from the mean), mean (gross) error (MGE—overall deviation from mean) and index of agreement (IOA) are presented for each parameter at each station. Benchmarks can provide an acceptable range of values to measure model performance against. Due to the extent of the uncertainties in the modelling it is not a case of pass-fail and any knowledge of the biases or shortcomings provide the user with a measure of the range of uncertainty in the data. Benchmarks instead are a helpful tool to understand how good or poor modelling results are, relative to a range of other model applications [1]. The most commonly referenced meteorological benchmarks in the literature were established by [2]. The modelling conducted by [2] was over the eastern and mid-west of the US where the terrain is considered flat and "simple". For more complex terrain the benchmarks provided by [3] and [4] may be more appropriate. CCAM performance statistics for temperature and wind speed predictions in 2008 are shown in Table S3, where the results meet benchmark simple (complex) are highlighted with green (yellow).

Taylor diagrams allow a visual comparison of the performance of different stations, experiments or variables. The Taylor diagram presents the correlation coefficient (R—linear relationship) and centred RMSE (CRMSE—overall accuracy) as metrics of similarity and the standard deviation (σ_M and σ_O—spread from the mean) representing amplitude of the variation of model results versus observations on a single diagram. The Taylor diagrams in Figure S1 show all the stations for temperature, wind speed, U-component and V-component across seasons. The optimal performance is when the model results are closest to the observed (purple dot on the x-axis). Between the seasons for all variables there is only a little spread in the performance statistics, indicating that the model performance is consistent throughout the year.

Table S2. Summary of statistical metrics used in comparison.

Name	Equation	Perfect Agreement
Mean Bias (MB)	$MB = \frac{1}{n}\sum_{i=1}^{n}(M_i - O_i)$	0
Mean Gross Error (MGE)	$MGE = \frac{1}{n}\sum_{i=1}^{n}\lvert M_i - O_i\rvert$	0
Pearson Correlation Coefficient (R)	$R = \frac{\sum_{i=1}^{n}((O_i - \bar{O})(M_i - \bar{M}))}{\sqrt{\sum_{i=1}^{n}(O_i - \bar{O})^2 \sum_{i=1}^{n}(M_i - \bar{M})^2}} = \frac{\sum_{i=1}^{n}((O_i - \bar{O})(M_i - \bar{M}))}{\sigma_O \sigma_M}$	1
Index of Agreement (IOA)	$IOA = 1 - \frac{\sum_{i=1}^{n}(M_i - O_i)^2}{\sum_{i=1}^{n}(\lvert M_i - \bar{O}\rvert + \lvert O_i - \bar{O}\rvert)^2}$	1
Centred RMSE (CRMSE)	$CRMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}((M_i - \bar{M}) - (O_i - \bar{O}))^2}$	0

The notations represent the number of data point pairs is denoted by N, M, and O reference the model and observation parameters or concentrations, respectively and the index i represents the period that the evaluation is covering.

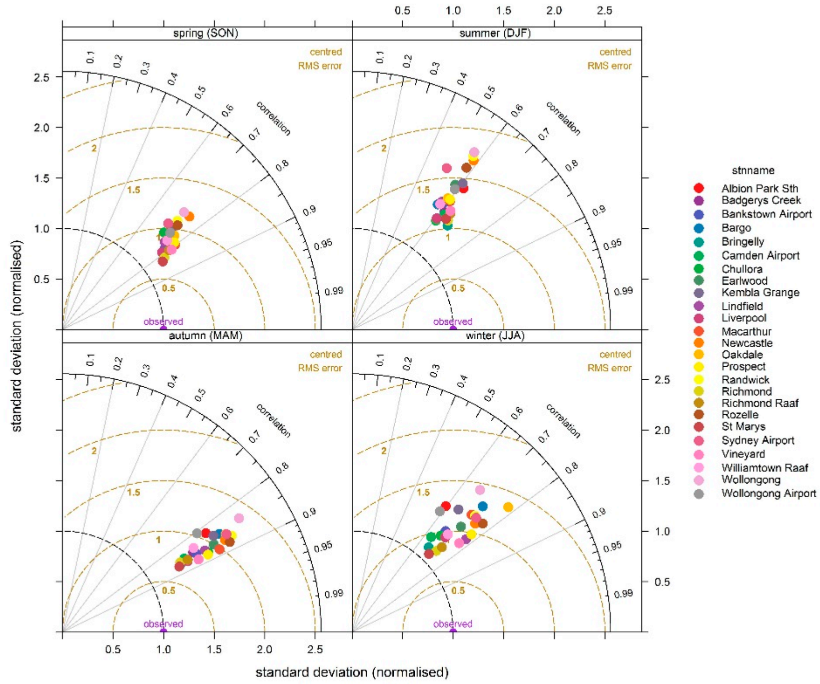
Table S3. Performance statistics for CCAM temperature and wind speed.

Parameter	Temperature			Wind Speed		
Stations	**MB (°C)**	**MGE (°C)**	**IOA**	**MB(m/s)**	**MGE(m/s)**	**IOA**
Ideal value	0	0	1	0	0	1
Benchmark (simple)	≤±0.5	≤±2	≥0.8	≤±0.5	≤±2	≥0.6
Benchmark (complex)	≤±1	≤±3	-	≤±1.5	≤±2.5	-
Bargo	0.94	1.93	0.95	1.71	1.81	0.63
Prospect	1.13	1.91	0.96	1.86	2.03	0.59
Newcastle	1.76	2.22	0.91	3.50	3.62	0.42
Wollongong	1.12	1.73	0.94	2.20	2.34	0.53
Badgerys Creek	1.63	2.43	0.94	1.34	1.89	0.70
Bankstown Airport	1.82	2.41	0.93	0.72	1.69	0.74
Camden Airport	1.66	2.44	0.94	1.61	2.05	0.72
Richmond RAAF	1.84	2.58	0.94	1.40	1.97	0.73
Sydney Airport	1.12	1.77	0.95	-0.45	1.56	0.83

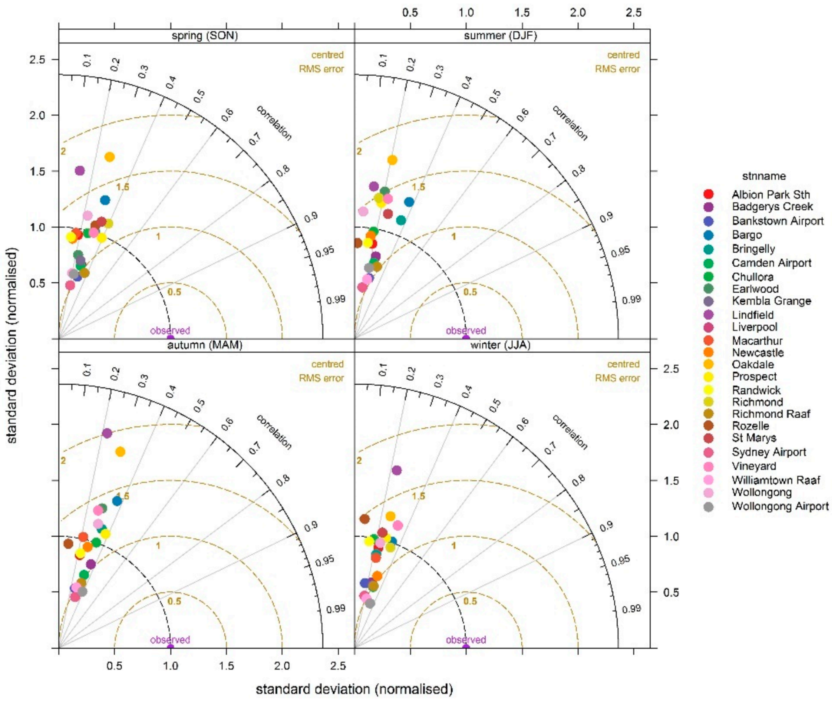
The performance of CCAM for temperature is good with correlations around 0.9 and CRMSE well below 1. The standard deviations of the model are also close to the centre line indicating a similar amplitude of variation between CCAM and observations. The model performance for each station is clustered close together which indicates there are not any spatial biases in CCAM ability to simulate temperature across the Sydney basin.

The performance of the wind speed is relatively similar between stations, particularly during spring and winter. Lindfield is a clear outlier with consistently lower correlations, higher CRMSE and standard deviations. For most stations the correlations are between 0.4–0.6 and the CRMSE are predominantly under 2.0, except for Lindfield. The performance is worst at Lindfield and Oakdale, while it is best at Sydney and Bankstown Airports.

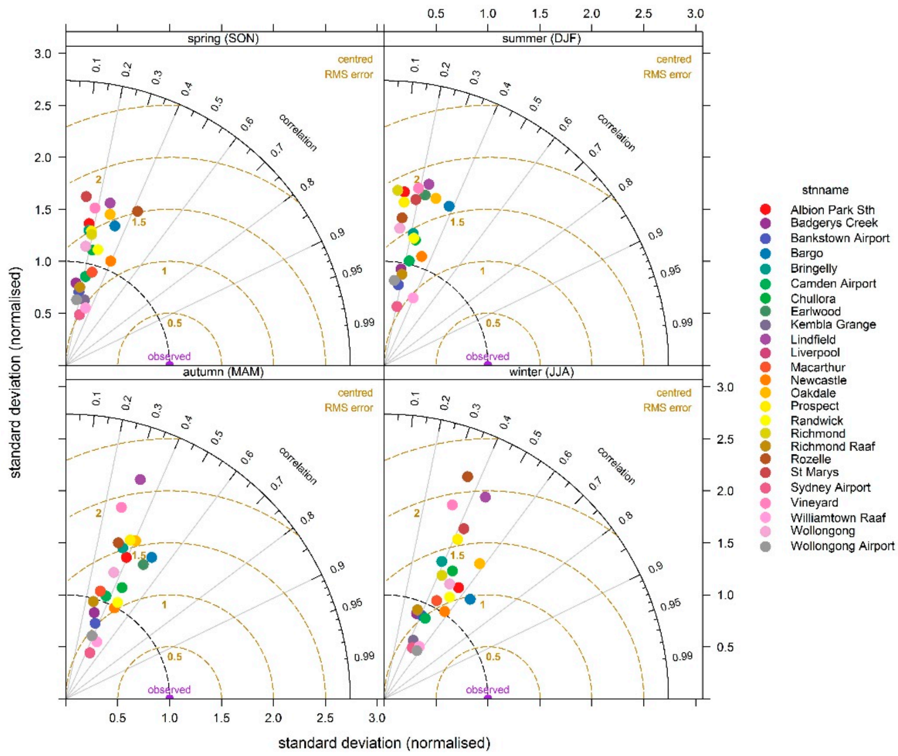
The zonal and meridional winds have a greater spread across the Taylor Diagrams and it appears that the best performance is seen during spring and summer and worst during autumn and winter. The correlations are between 0.6–0.8 and the CRMSE are less than 2, except for Lindfield as seen in the wind speed plots. As with wind speed the best representation of winds is at Sydney and Bankstown Airports. This would largely be influenced by the higher wind speeds recorded at these stations as CCAM has much strong wind speeds, as seen in the previous analysis.



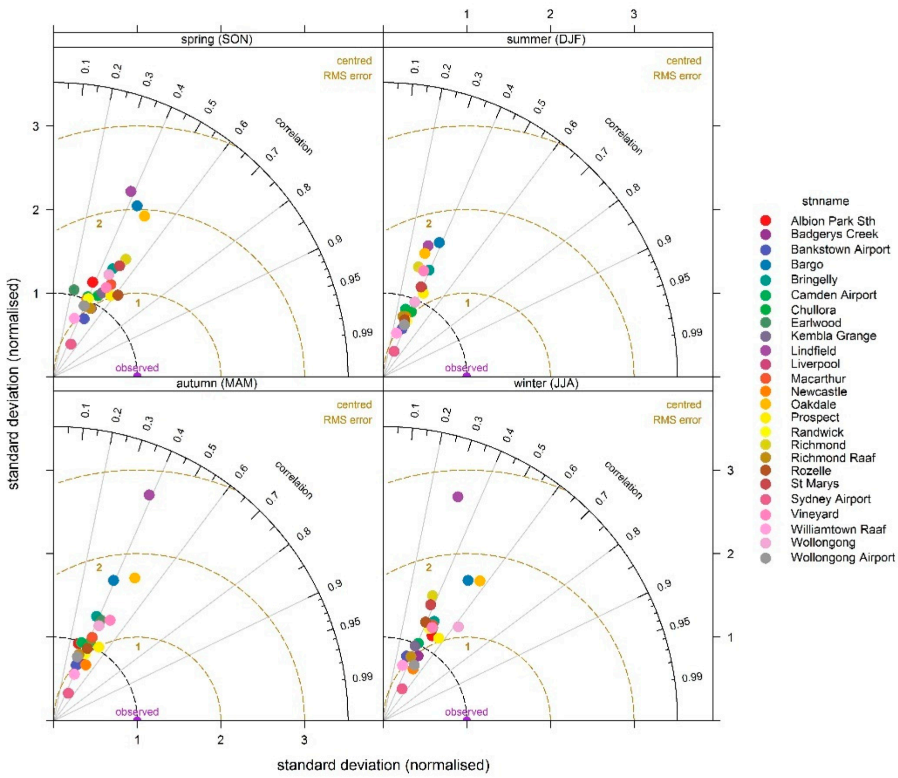
(a) Temperature



(b) Wind speed



(c) U wind component



(d) V wind component

Figure S1. Seasonal Taylor diagrams for CCAM (**a**) temperature, (**b**) wind speed, (**c**) U wind component and (**d**) V wind component.

References

1. Tesche, T.W.; McNally, D.E.; Tremback, D.C. Operational Evaluation of the Mm5 Meteorological Model over the Continental United States: Protocol for Annual and Episodic Evaluation. U. S. EPA: Washington, DC, USA, 2002.
2. Emery, C.; Tai, E.; Yarwood, G. Enhanced Meteorological Modeling and Performance Evaluation for Two Texas Ozone Episodes. ENVIRON International Corporation: Novato, CA, USA, 2001.
3. McNally, D.E. 12km MM5 Performance Goals. Alpine Geophysics, LLC: Arvada, CO, USA, 2009.
4. Kemball-Cook, S.; Jia, Y.; Emery, C.; Morris, R. Alaska MM5 Modeling for the 2002 Annual Period to Support Visibility Modeling. ENVIRON International Corporation: Novato, CA, USA, 2005.