Using WRF downscaling and self-organising maps to investigate particulate pollution in the Sydney region.

Supplementary Information

JAGODA CRAWFORD[1*], ALAN GRIFFITHS[1], DAVID D. COHEN[1], NINGBO JIANG[2], EDUARD STELCER[1]

[1]*Australian Nuclear Science and Technology Organisation, Locked Bag 2001 Kirrawee DC NSW 2232, Australia*

[2]*Office of Environment and Heritage, NSW Department of Premier and Cabinet, Sydney, Australia*

1 Corresponding author. Tel: +61 2 9717 3885, Fax: +61 2 9717 9260; e-mail Jagoda.Crawford@ansto.gov.au

1 The Weather Research and Forecasting model (WRF)

The Advanced Research (AR) WRF (WRF hereafter) model version 3.5.1 (Skamarock et al., 2004; Skamarock and Klemp, 2008) was used to carry out meteorological dynamical downscaling. The main advantages of using dynamical downscaling is that some processes which are on subgrid scales in the global model will be simulated explicitly by the regional model and more frequent output intervals can be specified. If the regional model simulates sea breezes and the influence of topography more realistically, trajectories computed from its output will also be more accurate.

A vast number of WRF configuration options are possible. The current configuration is similar to that used by Angevine et al. (2013). The model was configured with 12 km horizontal resolution, 50 vertical levels, with greatly enhanced resolution near the surface (half of the levels below 1 km and a 20 m thick near-surface level). Subgrid vertical turbulence was modelled with the Mellor-Yamada-Janjic (MYJ) boundary layer and surface layer schemes (Janjić, 1994), longwave and shortwave radiation was modelled with RRTMG (Iacono et al., 2008), cloud microphysics was modelled with the Lin (Purdue) scheme (Lin et al., 1983) and subgrid moist cumulus was modelled using the Grell-Freitas scheme (Grell and Freitas, 2014).

The model was run for one month at a time with five days of spin-up before the start of each month with boundary-conditions from the half-degree Climate Forecast System (CFS) and CFSv2 (Saha et al., 2010, 2013). The change to CFSv2 occurs during 2011.

Validation of the WRF wind fields (to observations) was limited to ensuring that the regional model provides improvements in the root mean squared error (rmse) and bias when compared with the CFSR or the ERA-Interim (Dee et al., 2011) reanalyses. The improvement in error scores, based on observations within the model domain during the year 2013 is shown in Table 1. For the 15 Australian Bureau of Meteorology (BOM; Table 2) meteorological observation sites, the wind speeds from the WRF simulations had the smallest rmse, although at night a positive bias was seen. The BOM stations were chosen to be well within the WRF domain and close to the coast.

Table 1: Root mean squared error (rmse) and bias in wind speed for 15 BOM stations.

Case	Time of day	rmse	bias
cfsr	**day**	3.35	-1.84
gdas1	**day**	2.77	-1.02
WRF	**day**	2.22	0.17
cfsr	**night**	2.79	0.32
gdas1	**night**	2.94	0.88
WRF	**night**	2.29	1.16
cfsr	**combined**	3.00	-0.57
gdas1	**combined**	2.86	0.10
WRF	**combined**	2.25	0.72

Table 2: The selected BOM sites.

WMO ID	Lat	Lon	Height (m)	Site name
95753	-33.6004	150.7761	19	Richmond RAAF
95940	-35.0936	150.8049	85	Jervis Bay (Point Perpendicular AWS)
95748	-34.5638	150.7900	8	Albion Park (Wollongong Airport)
95770	-33.2814	151.5766	18.8	Norah Head AWS
94765	-33.9181	150.9864	6.5	Bankstown Airport AWS
94776	-32.7932	151.8359	9	Williamtown RAAF
95779	-32.7103	152.1612	25	Nelson Bay (Nelson Head)
94767	-33.9465	151.1731	6	Sydney Airport
94937	-35.9093	150.1532	17	Moruya Heads Pilot Station
94755	-34.0391	150.6890	73.9	Camden Airport AWS
94938	-35.3635	150.4827	35.7	Ulladulla AWS
94752	-33.8969	150.7281	81.2	Badgerys Creek AWS
95771	-32.7886	151.3377	61	Cessnock Airport AWS
95937	-35.9004	150.1437	4	Moruya Airport AWS
94768	-33.8607	151.2050	39	Sydney (Observatory Hill)

2 Self Organising Map (SOM) classification

The SOM (Kohonen 2001) is an unsupervised method for data classification and visualisation. An important property of the SOM is related to the topologically ordered display of input data in the output space: similar data items (e.g., weather maps) from the input space are projected onto the nearby SOM nodes (representative vectors) and dissimilar data items onto the SOM nodes further apart (Jiang et al., 2013a). In this study, a synoptic classification was derived through a two-phase batch SOM classification procedure (CP2) detailed in Jiang et al. (2012) and Jiang et al. (2014). The data used for the classification consisted of twice daily (0000 and 1200 UTC, i.e., 10:00 and 22:00 AEST) National Centre for Environmental Prediction (NCEP)/National Centre for Atmospheric Research (NCAR) 1000hPa geopotential height reanalysis for 1958-2013. As in Jiang et al. (2012), the study domain covered latitudes 15°S-50°S and longitudes 130°E-170°E at a 2.5°×2.5° resolution, with NSW (area of interest) located near the centre. The first phase was to capture a rough estimation of the global patterns in the data, while the second phase was to fine tune the mapping to achieve local optimal and thus obtain the final groupings. A weather map (geopotential height field) is assigned to an SOM node from which it has the smallest squared Euclidean distance. Relatively large variability in geopotential height occurs at high latitudes. Hence, CP2 was implemented on the standardised time series so as to suppress the effects of spatial heterogeneity (in variability) on the classification results (Jiang et al., 2014). The final synoptic patterns were obtained by de-standardising the SOM nodes after the implementation of CP2. Several SOM map sizes (i.e., total number nodes/types on the SOM grid) were considered for the geopotential height dataset. It was found that the 5x4 SOM mapping (i.e., 20 types) was sufficient to reproduce the typical synoptic types identified in the literature for east Australia (Jiang et al., 2012). A similar classification was used in Jiang et al. (2013a), but focussing on visualisation of climate and air quality (ozone) data for the study region.

3 Locations of Power stations

Table 3: Centres and extent of rectangles representing each of the eight power stations. Mount Piper and Wallerawang are of particular interest in this study as they are located on the western side of the Great Dividing Range.

Number	Power Station	Lat (°S)	Long (°E)
1	Bayswater	32.3953±0.1	150.9491±0.1
2	Redbank	32.5784±0.1	151.0345±0.1
3	Liddell	32.3719±0.1	150.9783±0.1
4	Mount Piper	33.3934±0.1	149.9705±0.1
5	Wallerawang	33.4040±0.1	150.0845±0.1
6	Eraring	33.0623±0.1	151.5214±0.1
7	Vales Point B	33.1596±0.1	151.5428±0.1
8	Munmorah	33.2077±0.1	151.5393±0.1

4 Back Trajectory Coverage

To illustrate that more low altitude GDAS back trajectories travelled further west than WRF back trajectories, back trajectory density maps were generated for back trajectories starting at 300 m above ground level. A back trajectory density map was generated, one for each the GDAS and WRF back trajectories, by subdividing the region into 0.5 ° by 0.5 ° grid cells, and then counting the number of times a trajectory end point (i.e. the location of the trajectory at 30 min intervals) landed in each grid cell. Then for each grid cell, the count under WRF was calculated as a percentage of the count under GDAS (presented in Figure 1). Back trajectories corresponding to the high sulfur days were used, and only those days were chosen when air masses had originated east of Richmond. Further the grid cell count was incremented only if the back trajectory's altitude was below 1000 m above ground level, as we were only interested in low altitude back trajectories and the aim was to illustrate that more low altitude GDAS back trajectories travelled further west than WRF back trajectories. As can be seen from Figure 1, generally to the west of the Great Dividing Range the grid cell percentages are below 50%, with a few exceptions.

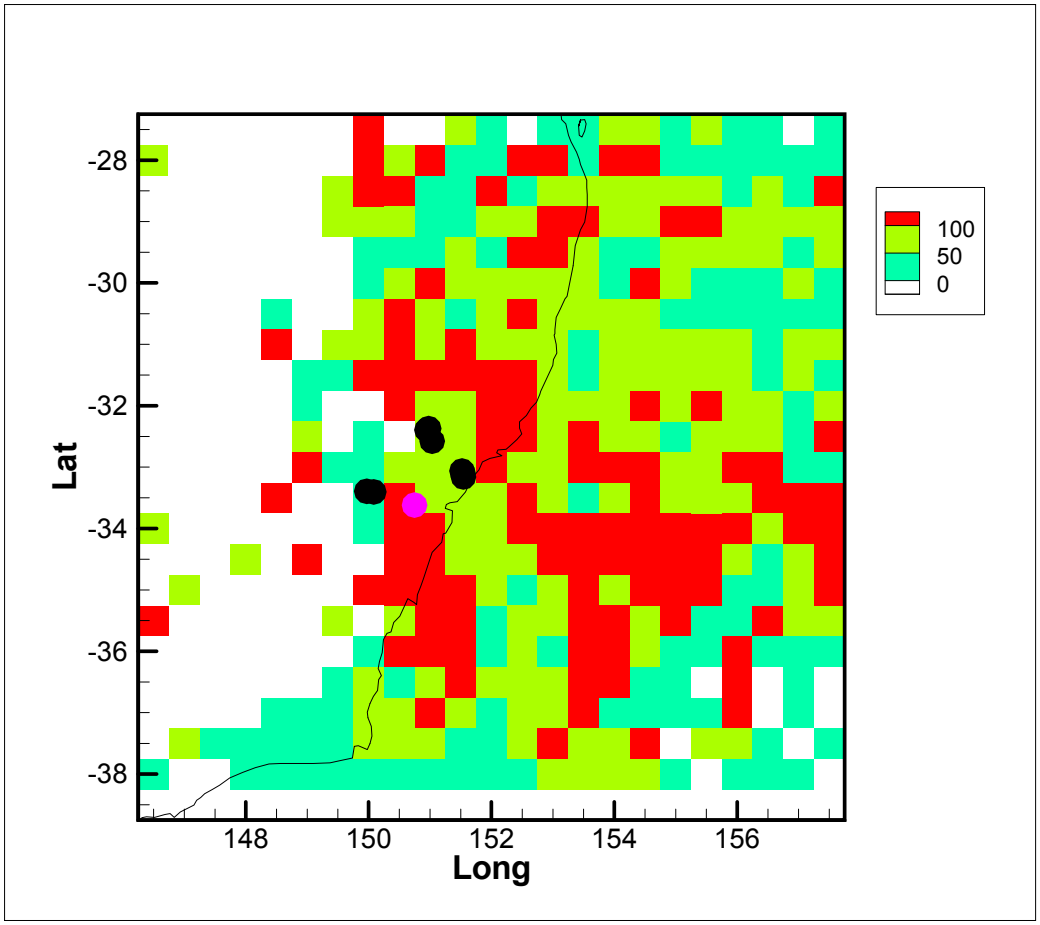

Figure 1: WRF back trajectory density map represented as a percentage of the GDAS back trajectory density map. The Richmond site is represented by the purple circle and the power station are represented by black circles.

5 Sea Breeze Example

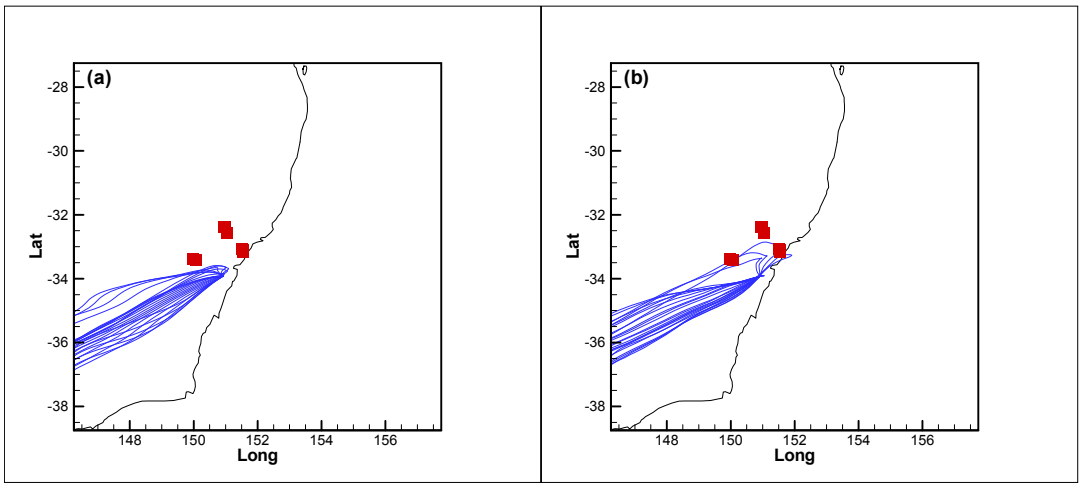

Figure 2: Difference between GDAS (left) and WRF (right) back trajectories for 20/1/2010, arriving at 300m agl, at Liverpool, illustrating effects of sea breezes under the WRF data. The red squares are the power stations.

6 References

Angevine W. M., Brioude J., McKee, S., Holloway J. S., Lerner B. M., Goldstein A. H., Guha A., Andrews A., Nowak J. B., Evan S., Fischer M. L., Gilman J. B. and Bon D., 2013. Pollutant transport among California regions, Journal of Geophysical Research, Atmospheres, 118(12), 6750–6763, doi:10.1002/jgrd.50490.

Dee D. P., Uppala S. M., Simmons A. J., Berrisford P., Poli P., Kobayashi S., Andrae U., Balmaseda M. A., Balsamo G., Bauer P., Bechtold P., Beljaars A. C. M., van de Berg L., Bidlo, J., Bormann N., Delsol C., Dragani R., Fuentes M., Geer A. J., Haimberger L., Healy S. B., Hersbach H., Hólm E. V., Isaksen L., Kållberg P., Köhler, M., Matricardi M., McNally A. P., Monge-Sanz B. M., Morcrette J.-J., Park B.-K., Peubey C., de Rosnay P., Tavolato C., Thépaut J.-N. and Vitart F., 2011. The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Q. J. R. Meteorol. Soc., 137(656), 553–597, doi:10.1002/qj.828.

Grell G. A. and Freitas S. R., 2014. A scale and aerosol aware stochastic convective parameterization for weather and air quality modeling, Atmos Chem Phys, 14(10), 5233–5250, doi:10.5194/acp-14-5233-2014.

Iacono M. J., Delamere J. S., Mlawer E. J., Shephard M. W., Clough S. A. and Collins W. D., 2008. Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models, J. Geophys. Res. Atmospheres, 113(D13), D13103, doi:10.1029/2008JD009944.

Janjić Z. I., 1994. The Step-Mountain Eta Coordinate Model: Further Developments of the Convection, Viscous Sublayer, and Turbulence Closure Schemes, Mon. Weather Rev., 122(5), 927–945, doi:10.1175/1520-0493(1994)122<0927:TSMECM>2.0.CO;2.

Jiang, N., Cheung, K., Luo, K., Beggs, P.J. and Zhou, W. (2012). On two different objective procedures for classifying synoptic weather types over east Australia. International *J. of Climatol* 32, 1475-1494. DOI: 10.1002/joc.2373

Lin, Y.-L., Farley, R. D. and Orville, H. D. (1983).Bulk Parameterization of the Snow Field in a Cloud Model*, J. Clim. Appl. Meteorol*., 22(6), 1065–1092, doi:10.1175/1520-0450(1983)022<1065:BPOTSF>2.0.CO;2.

Skamarock W. C., 2004. Evaluating mesoscale NWP models using kinetic energy spectra, Monthly Weather Review, 132(12), 3019–3032.

Skamarock W. C. and Klemp J. B., 2008. A time-split nonhydrostatic atmospheric model for weather research and forecasting applications, J. Comput. Phys., 227(7), 3465–3485, doi:10.1016/j.jcp.2007.01.037.

Skamarock W. C., Klemp J. B., Dudhia J., Gill D. O., Barker D. M., Wang W. and Powers J. G., 2014. A description of the advanced research WRF version 2, NCAR Technical Note.,

2005.National Centers for Environmental Prediction/National Weather Service/NOAA/U.S. Department of Commerce, 2008: NCEP ADP Global Upper Air and Surface Weather Observations (PREPBUFR format), May 1997 - Continuing. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory, Boulder, CO. [Available online at http://rda.ucar.edu/datasets/ds337.0/.] Accessed 7 Aug 2014.

Yerramilli, A., Dolda, V.B.R., Challa, V.S., Myles, L., Pendergrass, W.R., Vogel, C.A., et al. (2012). An integrated WRF/HYSPLIT modelling approach for the assessment of PM2.5 source regions over the Mississippi Gulf Coast region. *Air Quality Atmos Health*, 5, 401-412. Doi 10.1007/s11869-010-0132-1.