Evaluation of regional air quality models over Sydney, Australia: surface ozone and PM2.5

## Abstract

## Introduction

Air quality models are valuable tools to investigate the complex and dynamic interactions between meteorology and chemistry leading to poor air quality episodes

This inter-comparison exercise was designed to promote policy-relevant research on regional air quality modelling in Australia, specifically over the greater Sydney area (similar to AQMEII but our models are not operational – except OEH and we are modelling a much smaller region)

Models were run for three periods for which detailed characterisation of atmospheric composition is available with the aims to:

* Determine ability of models to reproduce observed surface concentrations of criteria pollutants, especially ozone and PM2.5, on an hourly basis (and on a rolling 4-hour average basis for ozone and on a daily basis for PM2.5)
* Investigate causes of discrepancies between models and observations and in between models (e.g. biogenic emissions/mixing ratios, chemistry?)
* Discuss implications for modelling framework to test policy scenarios

## Methods

The modelling has been conducted for the same geographical domains, grid resolution and measurement campaign periods

**2.1 Description of models**

For AQ evaluation: 3 modelling groups, three models, 4 configurations (UoM: CMAQ, WRF-Chem, CSIRO, OEH)

Domain/grid

Schemes

Emissions – not covered by Khalia

\*\*Need a table – sum over the domain for CO, NOx, etc + range of total emissions over the domain for isoprene (as in Im et al 2015)

Also, map of NOx emissions overlaid with stations that measure O3 and map of PM2.5 emissions overlaid with stations that measure PM2.5 (as per Im et al 2015)

Boundary conditions

Table 1: Overview of the configuration of the air quality models (the meteorological models were described in Monk et al)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model specifications | Model | WRF-Chem | WRF-CMAQ | CCAM-CTM (CSIRO) | CCAM -CTM (OEH) |
| Model Version |  |  |  |  |
| Domain | Number of nests |  |  |  |  |
| Horizontal resolution (for each nest) |  |  |  |  |
| Number of x grid points (per nest) |  |  |  |  |
| Number of y grid points (per nest) |  |  |  |  |
| Number of vertical layers |  |  |  |  |
| Height of first layer |  |  |  |  |
| Initial & Boundary conditions | Chemical BCs |  |  |  |  |
| Emissions | Anthropogenic |  |  |  |  |
| Biogenic |  |  |  |  |
| Sea salt |  |  |  |  |
| Dust |  |  |  |  |
| Chemical parameterisations | Gas-phase mechanism |  |  |  |  |
| Aerosol modules |  |  |  |  |
| Phytolysis schemes |  |  |  |  |

**2.2 Description of observations**

OEH network – ozone, (NOx), PM2.5

PM2.5 speciation?

Campaigns (SPS1 and 2, MUMBA) – detailed characterisation of surface atmospheric composition, including VOCs

* + - SPS1 and 2
      * When and why was this campaign held – Table 2 for dates
      * References: SPS report
      * Data sources:
        + <http://doi.org/10.4225/08/57903B83D6A5D>
        + <http://doi.org/10.4225/08/5791B5528BD63>
    - MUMBA
      * When and why was this campaign held
        + Table 2
      * References: MUMBA overview paper – ESSD – in progress
      * Data sources: PANGAEA – in progress

we have two locations with detailed measurements within our domain - this is an advantage (despite the fact that we have two summer campaigns – less relevant for met, but might be useful for AQ parameters)

Table 2: Overview of measurement campaigns

|  |  |  |  |
| --- | --- | --- | --- |
| Campaign | Period | Data source | Publication |
| SPS1 | 02/2011 - 03/2011 |  |  |
| SPS2 | 04/2012 - 05/2011 |  |  |
| MUMBA | 21/12/2012 – 15/02/2013 |  |  |

**2.3 Statistical analyses**

(include equations)

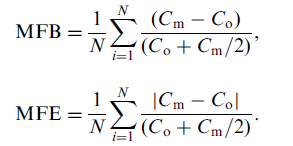
2.3.1 Ozone

r, RMSE, NMSE and NMB as per Im et al 2015 Part 1: Ozone

2.3.2 PM2.5

r, RMSE, MFE and MFB as per Boylan and Russell 2006

“The mean fractional bias (MFB) and mean fractional error (MFE) normalize the bias and error for each model-observed pair by the average of the model and observation before taking the average (does not assume the observations are the absolute truth)



...the mean fractional error and bias are the least biased and most robust of the various performance metrics“

But also test NMSE and NMB as per Im et al 2015 Part 2: PM

C:\Users\eag873\Notebooks\Notes\Work\pasted_image.png

PCC (r) is a measure of associativity and allows gauging whether trends are captured, and it is not sensitive to bias; RMSE is a measure of accuracy and, because it is squared, is sensitive to large departures

Table 3: Benchmarks for model performance for air quality parameter – THIS TABLE NEEDS UPDATING

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Species | Metric | Criteria | Goal | Source |
| PM2.5 | MFE | 75% | 50% | Boylan and Russell (2006), US-EPA (2007) |
| MFB | ±60% | ±30% |
| O3 | MFE | 35% | |
| MRB | 15% | |
| NOx, CO, PM10 | FAC2 | Half points within -0.3<FB<0.3 | | Chang and Hanna (2004) |
| FB |
| NMSE | <4 | |

## Model evaluation results

**3.1 Ozone**

3.1.1 Region/domain-wide analysis

Table with r, RMSE, NMSE and NMB for each model for ozone (hourly and rolling 4-hour average, also max daily ozone?)

Diurnal cycles

* is timing of max ozone right?
  + Is amplitude OK? It tends to be larger in models that simulate a more stable and shallow nocturnal boundary layer (Im et al 2015 Part 1)

Box plots (overall and by campaign, or by campaign only? 4 vs. 3 plots – no big deal)

Taylor diagrams (one for each campaign -or all campaigns on one – if legible):

* + - How spread are the results?
    - Are there any outliers?
    - How do they vary between campaigns?
    - Are there any obvious seasonal differences?

(Note: Taylor plot only show random errors, not systematic bias – Boylan and Russell 2006)

check how the model biases vary with ozone levels – bin ozone (quartiles?) and check NMB for each bin– do this for both time bases (Im et al 2015 Part 1) -all campaigns combined?

Investigate role of boundary conditions for ozone – does the global model overpredicts/underpredicts over the greater Sydney region?

(I will need these boundary conditions from the modellers)

3.1.2 Spatial analysis

Bubble plots of statistical parameters: r, RMSE, NMSE, NMB for each model for each campaign

(hourly only? Or rolling 4-hour and max also? – plot them all, maybe don’t include them all)

|  |  |  |
| --- | --- | --- |
| CMAQ |  |  |
| WRF-Chem |  |  |
| OEH |  |  |
| CSIRO |  |  |
| SPS1 | SPS2 | MUMBA |

Figure 1: Spatial evaluation of model performance – bubble plot – correlation coefficient

|  |  |  |
| --- | --- | --- |
| CMAQ |  |  |
| WRF-Chem |  |  |
| OEH |  |  |
| CSIRO |  |  |
| SPS1 | SPS2 | MUMBA |

Figure 2: Spatial evaluation of model performance – bubble plot – RMSE

|  |  |  |
| --- | --- | --- |
| CMAQ |  |  |
| WRF-Chem |  |  |
| OEH |  |  |
| CSIRO |  |  |
| SPS1 | SPS2 | MUMBA |

Figure 3: Spatial evaluation of model performance – bubble plot – NMSE

|  |  |  |
| --- | --- | --- |
| CMAQ |  |  |
| WRF-Chem |  |  |
| OEH |  |  |
| CSIRO |  |  |
| SPS1 | SPS2 | MUMBA |

Figure 4: Spatial evaluation of model performance – bubble plot – NMB

**3.2 PM2.5**

3.2.1 Region/domain-wide analysis

Table with r, RMSE, NMSE and NMB (or MFE and MFB) for each model for PM2.5 (hourly and calendar day average)

FOR PM, daily means a calendar day average - so it makes sense to look at the data on that basis - there is no hourly standard, and no clear diurnal cycle in most cases - but you can check - obs vs. models might let you know whether the timing of emissions is off/weird/wrong

Box plots (overall and by campaign, or by campaign only?)

Taylor diagrams:

* + - How spread are the results?
    - Are there any outliers?
    - How do they vary between campaigns?
    - Are there any obvious seasonal differences?

3.2.2 Spatial analysis – less relevant, only 5-6 stations

Daily averages only

|  |  |  |
| --- | --- | --- |
| CMAQ |  |  |
| WRF-Chem |  |  |
| OEH |  |  |
| CSIRO |  |  |
| SPS1 | SPS2 | MUMBA |

Figure 5…: Spatial evaluation of model performance – bubble plots

**3.3 PM2.5 speciation?**

Same stats as for PM2.5

Species and time-basis will depend on observations:

* we have OC and EC at MUMBA, OC and EC, SO42-, NO3-and NH4+ for SPS1 and SPS2 (filters: 5- or 8-hour resolution)

Plot timeseries

Table with stats

## Discussion

* + Do models meet benchmarks for performance? Overall? Site-by site?
  + Do models capture exceedances, if any?
  + Are there City vs rural? / Inland vs coastal? biases
  + What drives the differences between models / with the observations? (look at other parameters such as NOx, isoprene, etc.)
  + dry deposition rate (or totals) for each model for ozone? Chemistry schemes?
  + NO2 levels (to interpret ozone)
  + Refer back to meteorological analysis – wind speed, mixing height
  + Can model reproduce intensive campaign results (anthropogenic VOCs, biogenic tracer, PM2.5 composition)?
  + What does this tell us about the models/inputs/etc?
  + What is the influence of natural sources vs. anthropogenic sources?
  + Also, how do our models fare compared to AQMEII, etc

## Summary and conclusions

* Does one model outperform the others?
* Main reason(s) for biases
* Going forward – modelling air quality in the Greater Sydney area – implications for policy/modelling framework