Evaluation of Earth System Models*

Dominick Spracklen
University of Leeds

Please ask questions as we go along.....there is no such thing as a “stupid question”!

*apologies for heavy bias to the work of my group
The need for Earth System Models to understand atmospheric composition

Biogeochemical cycles

Climate

Human Health

Global premature deaths from selected environmental risks: Baseline, 2010 to 2050

Note: * Child mortality only.
Source: OECD Environmental Outlook Baseline; output from IMAGE.
Some general pointers on model evaluation

*Evaluation* rather than *Validation*

**Avoid** using models does a “good job”, “model compares well”, “model in reasonable agreement” .......

Where possible make comparisons *quantitative*

Go beyond model bench marking

Comparison with observations should help improve our understanding of *processes* in the atmosphere
How do we quantify model-observation agreement? : Metrics

Test of Metrics (Continued) : 11 models for nss-SO$_4^{2-}$

<table>
<thead>
<tr>
<th>Models</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
<th>K</th>
<th>L</th>
<th>M</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{O}$</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
<td></td>
</tr>
<tr>
<td>$\bar{M}$</td>
<td>0.35</td>
<td>1.37</td>
<td>1.19</td>
<td>1.34</td>
<td>1.22</td>
<td>1.16</td>
<td>1.19</td>
<td>1.02</td>
<td>0.79</td>
<td>1.23</td>
<td>0.67</td>
<td>0.00</td>
<td>1.96</td>
<td>+∞</td>
</tr>
<tr>
<td>$N$</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>$r$</td>
<td>0.295</td>
<td>0.240</td>
<td>0.737</td>
<td>0.277</td>
<td>0.239</td>
<td>0.769</td>
<td>0.953</td>
<td>0.977</td>
<td>0.609</td>
<td>0.692</td>
<td>0.767</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

**Difference**

- $B_{MB}$: -0.63, 0.40, 0.21, 0.37, 0.24, 0.18, 0.21, 0.05, -0.19, 0.25, -0.31, -0.98, +0.98, +∞
- $E_{RMSSE}$: 0.63, 0.46, 0.42, 0.52, 0.34, 0.42, 0.24, 0.14, 0.42, 0.52, 0.41, 0.98, +0.98, +∞
- $E_{NMBE}$: 0.79, 0.55, 0.52, 0.70, 0.49, 0.48, 0.37, 0.16, 0.58, 0.63, 0.55, 0.98, +0.98, +∞

**Relative Difference**

- $B_{RMB}$: -0.65, 1.23, 0.91, 0.38, 0.38, 0.70, 1.40, 0.34, 0.33, 0.19, 0.75, -0.06, -1.00, +1.00, +∞
- $E_{RMSSE}$: 0.65, 1.26, 1.01, 0.60, 0.80, 1.58, 0.39, 0.39, 0.59, 0.94, 0.52, 1.00, +1.00, +∞
- $B_{NMB}$: -0.64, 0.41, 0.22, 0.38, 0.25, 0.18, 0.21, 0.05, -0.20, 0.26, -0.32, -1.00, +1.00, +∞
- $E_{NMBE}$: 0.64, 0.47, 0.43, 0.53, 0.34, 0.43, 0.25, 0.15, 0.44, 0.53, 0.42, 1.00, +1.00, +∞
- $B_{FMB}$: -1.00, 0.53, 0.37, 0.16, 0.30, 0.35, 0.22, 0.16, -0.04, 0.30, -0.24, -2.00, +0.67, +∞
- $E_{FMBE}$: 1.00, 0.56, 0.48, 0.45, 0.43, 0.56, 0.27, 0.24, 0.47, 0.53, 0.53, 2.00, +0.67, +∞
- $B_{NMBF}$: -1.80, 0.41, 0.22, 0.38, 0.25, 0.18, 0.21, 0.06, -0.24, 0.26, -0.46, -∞, +1.00, +∞
- $E_{NMBF}$: 1.80, 0.47, 0.43, 0.53, 0.34, 0.43, 0.25, 0.15, 0.54, 0.53, 0.61, +∞, +1.00, +∞

- Model H: best; Model A: worst
- Models E, G, H: acceptable
- If criteria: ±25% ($B_{NMBF}$), 35% ($E_{NMBF}$)
Difficulties evaluating atmospheric composition in Earth System Models

- **Sparse observations** (mostly surface or column, some aircraft, satellite)
- **Errors in meteorology** (GCM: doesn’t use ‘real’ meteorology)
- Resolution
- **Sub-grid processes**
- **Compensating errors** (e.g., emissions vs removal)
- **Representativeness** of observations
  - “Special objectives” of “mission” observations introduces bias
- **Temporal processes** rarely observed (e.g., nucleation, growth, scavenging)
Uncertainty analysis to determine sensitivity of model diagnostics to different processes

CCN sensitivity

Spracklen et al., 2005

Remote

Polluted

Lee et al., 2011

Spracklen et al., 2005

Aqueous oxidation diameter

ACC_COEF

NUC_THRESH

NUCRT SIZE

SO2_PART

SCAV DIAM

SO2 emissions

SS EMS
Aerosol sources and processes: New particle formation

- **New particle formation**
- **Secondary particles**
- **Primary particles, soot, soil dust, sea salt**
- **SO$_2$, NO$_x$, NH$_3$, Organics**

**Processes**
- Condensation
- Coagulation
- Oxidation
- Wet deposition
- Dry deposition

**Cloud condensation nuceli**
- Cloud drops
Evaluating particle formation mechanisms in a global model

Observations

Binary Homogeneous Nucleation (BHN)

BHN + Primary particle

BHN + Primary particles + Boundary Layer Nucleation (BLN)

BLN Formation rate of 1 nm particles: \( J_1 = A [H_2SO_4] \)

Spracklen et al., ACP, 2006
Synthesis of observations of total particle number concentration

Locations of total particle number observations

Crucial importance of those making observations – involve early in analysis and acknowledge with co-authorship when appropriate

Spracklen et al., ACP, 2010
Model evaluation to elucidate importance of different processes for total particle number

Spracklen et al., ACP, 2010
Particle formation, particle growth and cloud condensation nuclei

Observations
- BHN+BLN (A=2e-8s⁻¹)
- BHN+BLN (A=2e-6s⁻¹)
- BHN+BLN (A=2e-6s⁻¹) ; SOAx5

BHN
- BHN+BLN (A=2e-7s⁻¹)

BLN
Formation rate of 1 nm particles:
\[ J_1 = A \left[ H_2SO_4 \right] \]

Spracklen et al., GRL, 2008
Use the evaluated model to estimate the importance of particle formation

Careful design of model experiments required

Combinations of model experiments were carried out to obtain the relative contribution from primary particles (PR), upper tropospheric nucleation (UTN), and boundary layer nucleation (BLN) to particle number concentrations:

1. PR: Runs with only primary particles from particulate emissions with and without ultrafine sea-salt. We can calculate the contribution of ultrafine sea-salt to CN and CCN from these runs.

2. PR+UTN: Runs with particulate emissions (without ultrafine sea-salt) and UTN represented with the BHN (Kulmala et al., 1998) nucleation parameterization.

3. PR+UTN+BLN: Runs with particulate emissions (without ultrafine sea-salt), UTN represented with BHN (Kulmala et al., 1998), and BLN represented by the activation nucleation parameterization (Kulmala et al., 2006).

These simulations were carried out by running the model over the full year 2000 with 3-month spin-ups.

Merikanto et al., ACP, 2009

Primary particles only

Primary particles + Binary homogeneous nucleation (BHN or UTN)

ALL = Primary particles + BHN + Boundary Layer Nucleation (BLN)
Particle formation accounts for \(~50\%\) of global CCN number

**Fraction of CCN from particle formation (gas-to-particle formation)**

Merikanto et al., ACP, 2009
Evidence from ambient observations for a role of organics in particle formation

A particle formation mechanism including organics results in better representation of vertical profiles of aerosol number

Binary Homogeneous nucleation

\[ J = A[H_2SO_4] \]

\[ J = k[H_2SO_4]^2 \]

\[ J = k[H_2SO_4][oxorg] \]
Aerosol sources and processes: Biomass burning

- **SO\(_2\), NO\(_x\), NH\(_3\), Organics**

- **Oxidation**

- **Coagulation**

- **Condensation**

- **Nucleation**

- **Secondary particles**

- **Cloud condensation nuclei**

- **Primary particles, soot, soil dust, sea salt**

- **Cloud drops**

- **Wet deposition**

- **Dry deposition**
Impacts of tropical biomass burning on surface air quality (PM2.5)

Aerosol optical depth (AOD) underestimated. Surface PM2.5 in Amazonia is well simulated

Normalised mean bias for PM2.5

Normalised mean bias for AOD at (440 nm)

Kaiser et al., 2012; Marlier et al., 2012; Tosca et al., 2013

PM2.5 observations c/o Paulo Artaxo

Different fire emissions
Aerosol profiles over the Amazon provide additional evidence of model performance

AMS observations during SAMBBA

Observations c/o Will Morgan

Photo c/o Jim McQuaid
Issues with emissions: biomass burning emissions from satellite

Total PM2.5 at Santarem NE Amazon

PM2.5 measurements c/o of P. Artaxo

Aerosol Optical Depth in SE Asia

Fire emissions from FINN give better model performance than GFED and GFAS particularly in regions with small fires.
Aerosol sources and processes: Secondary organic aerosol formation

- **Condensation**
  - Nucleation
  - Oxidation
  - Condensation

- **Coagulation**

- **Secondary particles**
  - Cloud condensation nuceli

- **Primary particles, soot, soil dust, sea salt**

- **Wet deposition**

- **Dry deposition**

- **SO₂, NOₓ, NH₃, Organics**
Organic aerosol is dominant in the atmosphere – but sources are very poorly understood

Zhang et al., GRL, 2007
Global budget of secondary organic aerosol (SOA) is particularly uncertain.

Kanakidou et al. (2005)

Goldstein & Galbally (2007)

Hallquist et al. (2009)

Tsigaridis et al., ACP, (2014)
Understanding sources and formation of SOA

Spracklen et al., ACP, 2011

GLOMAP model with ~100 Tg a⁻¹
SOA source matches surface AMS observations
Using aircraft observations – standard model underestimates organic aerosol

Aircraft observations

GEOS-CHEM Model
26 Tg SOA a⁻¹

Heald et al., 2011
Understanding sources and formation of SOA

Aircraft observations

GEOS-CHEM Model
Increased SOA source: 126 Tg SOA a⁻¹

Heald et al., 2011
Model-observation analysis provides improved estimates of SOA budget
Atmospheric processes: Oxidation

- Oxidation
- SO$_2$, NO$_x$, NH$_3$, Organics

**Secondary particles**
- Primary particles, soot, soil dust, sea salt
- Cloud condensation nuceli

**Coagulation**
- Condensation
- Nucleation

**Primary particles**
- Cloud drops
- Wet deposition
- Dry deposition
Going beyond evaluation of the model mean.....

Daytime change in ozone (09:00 - 17:00 UT) Cape Verde

Without halogen chemistry the model underestimates daytime destruction of ozone

Read et al., 2008
Role of halogens in ozone destruction

Read et al., 2008
Atmospheric processes: Deposition

- Atmospheric processes:
  - Deposition
  - Oxidation
  - Nucleation
  - Condensation
  - Coagulation

- Secondary particles:
  - Cloud condensation nuceli
  - Wet deposition
  - Dry deposition

- Primary particles, soot, soil dust, sea salt

- SO$_2$, NO$_x$, NH$_3$, Organics

- Cloud drops
But comparisons can also be interpreted as test of emissions rather than deposition.

Fisher et al., 2011
Multi-model evaluation creates additional complexity

Using observations of atmospheric composition in the Arctic to evaluate a range of chemistry-climate models

Monks et al., ACPD, 2014
Developing methods to synthesise evaluation of multiple models against the same data

Monks et al., ACPD, 2014
Satellite observations of CO and multi-model bias

Monks et al., ACPD, 2014
Evaluating “long-term” trends in atmospheric composition

Leibensperger et al., 2012

Observations

Model

Lamarque et al., 2010
Exploring processes that contribute to interannual variability of atmospheric composition

Western US regional mean Jun-Aug organic carbon (OC) aerosol

Model captures interannual variability driven by changing emissions

Observations (IMPROVE)

Global model (Observed fires)

Global model Climatological fires

Spracklen et al., GRL. 2007
Exploring the impacts of trends in anthropogenic sources of aerosol

**European mean anthropogenic emissions (MACCity)**

![Graph showing annual emissions of various pollutants from 1960 to 2010](image)

- Large changes in anthropogenic pollution sources over last 50 years.

**HadGEM3-UKCA simulated European mean AOD**

![Graph showing aerosol optical depth from 1960 to 2010](image)

Turnock et al., in prep
Sulfate aerosol trends over Europe – the importance of cloud pH

Model underestimates wintertime sulfate. Increasing cloud water pH increases importance of aqueous phase oxidation by ozone and improves model in winter

Turnock et al., in prep
Trends in surface shortwave radiation

Surface shortwave radiation across Europe

Model with aerosol radiative effects (ARE) captures observed brightening from ~1980 onwards

Re-analysis product impacts model performance before 1980

Turnock et al., in prep
Coupling to the Earth System: impacts of atmospheric composition on ecosystems

Long-term observations of forest biomass in the Amazon show a long-term increase in biomass.

The cause is not known: CO$_2$, T, aerosol?
Coupling and evaluating model components

Fire emissions GFEDv3

Deforestation emissions

Meteorology

Meteorological data ECMWF

JULES

Clouds, albedo, GHGs

ISCCP, ECMWF

GLOMAP
(Spracklen et al., 2005)
(Mann et al., 2010)

Aerosol model

Edwards-Slingo
(Edwards & Slingo, 1996)
(Rap et al., 2013)

Radiation model

EU WATCH

Meteorology

Land-surface model

OBSERVATIONS

JULES
(Mercado et al., 2007)
(Best et al., 2011)
(Clark et al., 2011)

Coupling and evaluating model components

OBSERVATIONS
Evaluating the impacts of aerosol on diffuse radiation and photosynthesis

2006-2007 GPP [μmol C m⁻² s⁻¹] at French Guyana (9am to 5pm only) - All switches

Gross primary productivity
(μmol C m⁻² s⁻¹)

Photosynthetically active radiation
(μmol m⁻² s⁻¹)

Observations
Model

Diffuse sunlight
Direct sunlight

(Kanniah et al., 2012)

Alex Rap
Diffuse radiation from deforestation fire increases Amazon basin carbon storage by 60 Tg C a\(^{-1}\) (100 kg C ha\(^{-1}\) a\(^{-1}\)), accounting for 15% of the observed carbon sink.

*Rap et al. submitted*
Summary

Model evaluation is a critical component of our work.

Model evaluation should be quantitative. Ideally it should inform and improve our understanding of atmospheric processes.

Careful experimental design and model-observation analysis is required.

As more processes are coupled into ESMs – evaluation will become more and more necessary (and difficult)!

We rely on observational data for our science. Involve observational scientists in evaluation from an early stage. Acknowledge the crucial role through co-authorship where appropriate.