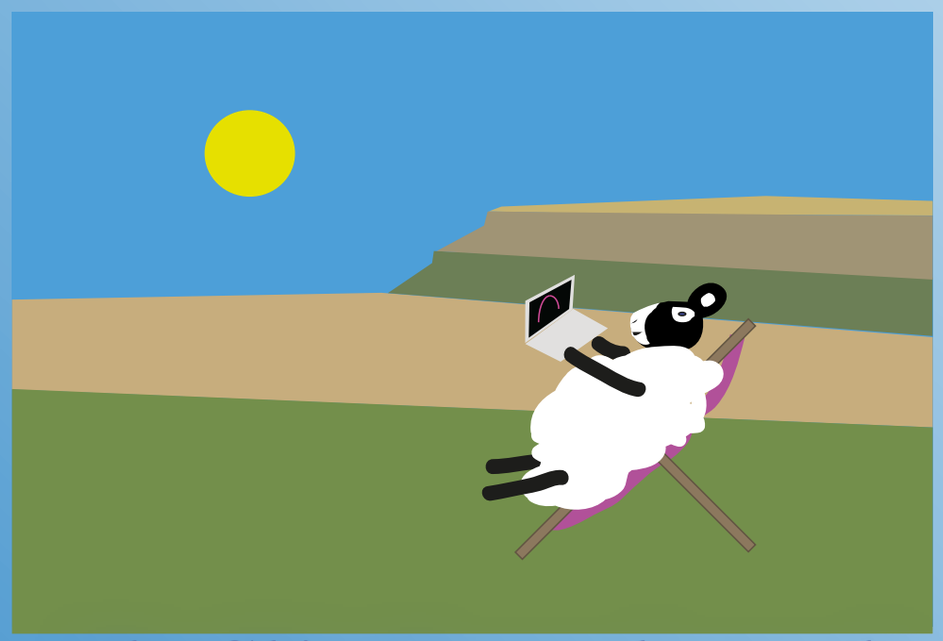
Model, models, models…

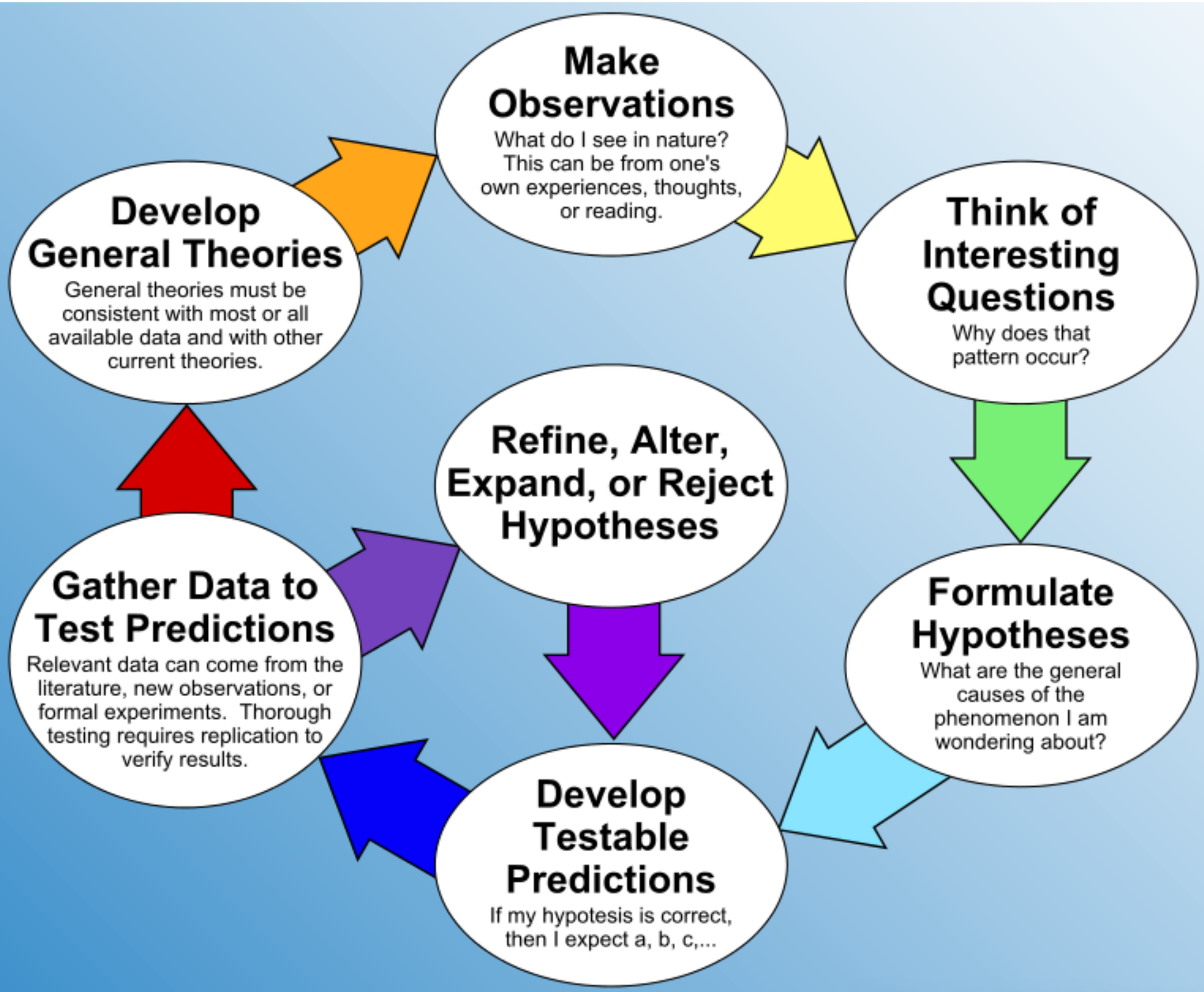
Julia Hargreaves, Chief Executive Editor of
Geoscientific Model Development and second best
Scientist at BlueSkiesResearch.org.uk



Outline
(the product of a fevered mind)

- What are models, anyway?
- Perceived problems with models
- Models through the GMD filter
- Conclusions

A model of the scientific method (from Wikipedia)



schematic by Archon Magnus

A model of the scientific method (from Wikipedia)

Alternatively: "Science is the process of gathering, comparing, and evaluating proposed models against observables." (also from Wikipedia)

schematic by Archon Magnus

Geoscientific Model Development

Geoscientific Model Development
An interactive open-access journal of the European Geosciences Union

| EGU.eu | EGU Journals | Contact | Imprint |

Documenting model development

BlueSkiesResearch.org.uk

Geoscientific Model Development

Geoscientific Model Development
An interactive open-access journal of the European Geosciences Union

| EGU.eu | EGU Journals | Contact | Imprint |

Documenting model development

- In GMD a model is the computer code that is run with certain boundary conditions to make the prediction - same meaning as assumed in this session(?)

Where do models fit into the Geosciences?

Notice something

Think of Interesting Questions

Formulate Hypotheses

Develop Testable Predictions

Gather Data to Test Predictions

Refine, Alter, Expand, or Reject Hypotheses

Develop General Theories

Where do models fit into the Geosciences?

Notice something

Think of Interesting Questions

Formulate Hypotheses

Develop Testable Predictions

Numerical Models

Gather Data to Test Predictions

Refine, Alter, Expand, or Reject Hypotheses

Develop General Theories

Geoscientific Model Development

Geoscientific Model Development
An interactive open-access journal of the European Geosciences Union

| EGU.eu | EGU Journals | Contact | Imprint |

Documenting model development
- since mid-2015 code must be made available

BlueSkiesResearch.org.uk

Geoscientific Model Development

Geoscientific Model Development
An interactive open-access journal of the European Geosciences Union

| EGU.eu | EGU Journals | Contact | Imprint |

Documenting model development
- since mid-2015 code must be made available

- In GMD a model is the computer code that is run with certain boundary conditions to make the prediction - same meaning as assumed in this session(?)

Geoscientific Model Development

Geoscientific Model Development
An interactive open-access journal of the European Geosciences Union

| EGU.eu | EGU Journals | Contact | Imprint |

Documenting model development
- since mid-2015 code must be made available

- In GMD a model is the computer code that is run with certain boundary conditions to make the prediction - same meaning as assumed in this session(?)
- As well as model descriptions, GMD also publishes papers about practically any sort of code related to numerical models, some of which could also be thought of as models:
  - data analysis - transforming measurements into understood variables - for model input or comparison with output
  - evaluation methods
  - visualisation methods
  - data assimilation methods

Geoscientific Model Development

Geoscientific Model Development
An interactive open-access journal of the European Geosciences Union

| EGU.eu | EGU Journals | Contact | Imprint |

Documenting model development
- since mid-2015 code must be made available

- In GMD a model is the computer code that is run with certain boundary conditions to make the prediction - same meaning as assumed in this session(?)
- As well as model descriptions, GMD also publishes papers about practically any sort of code related to numerical models, some of which could also be thought of as models:
  - data analysis - transforming measurements into understood variables - for model input or comparison with output
  - evaluation methods
  - visualisation methods
  - data assimilation methods
- Models that GMD does not (usually) publish
  - Models that are trivial to code and that can be fully described in the same paper as the results
  - Physical, Chemical or Biological models - like wave tanks, lab coccoliths or lab rats
  - Mathematical models - models expressed in maths (there are some of these in GMD!)
  - Conceptual models - models expressed in words and pictures

Geoscience according to GMD!

Develop General Theories

Make Observations

Think of Interesting Questions

Formulate Hypotheses

Refine, Alter, Expand, or Reject Hypotheses

Develop Testable Predictions

Gather Data to Test Predictions

Geoscience according to GMD!

Notice something

Think of Interesting Questions

Formulate Hypotheses

Refine, Alter, Expand, or Reject Hypotheses

Develop General Theories

Develop Testable Predictions

Numerical Models

Gather Data to Test Predictions

Geoscience according to GMD!

Notice something

Think of Interesting Questions

Formulate Hypotheses

Develop Testable Predictions

- Model Development
- Numerical Models

Design and run Experiment

- Model boundary conditions

Gather Data to Test Predictions

- Data Analysis Methods
- Data Assimilation Methods

Refine, Alter, Expand, or Reject Hypotheses

- Visualisation Methods
- Evaluation Methods

Develop General Theories

Geoscience according to GMD!

Notice something

Think of Interesting Questions

Formulate Hypotheses

Develop Testable Predictions

Design and run Experiment

Model boundary conditions

Gather Data to Test Predictions

Visualisation Methods

Evaluation Methods

Data Analysis Methods

Data Assimilation Methods

Develop General Theories

Model Development

Mathematical Models

Numerical Models

Geoscience according to GMD!

Notice something

Think of Interesting Questions

Formulate Hypotheses

Develop Testable Predictions

Design and run Experiment

Model boundary conditions

Gather Data to Test Predictions

Visualisation Methods

Evaluation Methods

Data Analysis Methods

Data Assimilation Methods

Refine, Alter, Expand, or Reject Hypotheses

Develop General Theories

Conceptual Models

Mathematical Models

Numerical Models

Model Development

Physical, Chemical and Biological Models

Problems of large models

We get stuck in one or two pentagons, and don't have time for the other stuff that makes it science.

> "The fundamental goals of Science are to formulate generalizations based on observations and testable hypotheses (induction), and to use these generalizations to make statements about particular cases (inference), in support of decision-making."

Many models are so complex that no one person can understand them.

More problems?

- Multitudes of: data, approaches for building and testing of models and ways to estimate and handle uncertainty.
  - Obstructs communication across disciplines and with it the building of interdisciplinary models to address interdisciplinary questions.
- Progress towards a commonly applicable framework for model building and application along the following questions:
  - • How to evaluate the appropriateness (generality, parsimony) of models in a generalized way?
  - • How to evaluate the interplay of data-, model structure- and predictive uncertainty, i.e. the flow of information from data through models to decision-makers?
  - • How to learn from the encounter of models and data; i.e. how to detect, diagnose and correct model structural errors?
  - • How to build complex interdisciplinary models that remain falsifiable?
  - • How to assess limits of predictability of models and the underlying theoretical concepts?

How to build complex interdisciplinary models that remain falsifiable?

How to build complex interdisciplinary models that remain falsifiable?

- If you are predicting future climate, even if your model is super-simple, this is a real problem!

How to build complex interdisciplinary models that remain falsifiable?

- If you are predicting future climate, even if your model is super-simple, this is a real problem!
  - Models are trivially falsified in that they are all wrong - just look at the last century simulations.

How to build complex interdisciplinary models that remain falsifiable?

- If you are predicting future climate, even if your model is super-simple, this is a real problem!
  - Models are trivially falsified in that they are all wrong - just look at the last century simulations.
  - The future is always out-of-sample, so you will never know if your model is good enough.

How to build complex interdisciplinary models that remain falsifiable?

- If you are predicting future climate, even if your model is super-simple, this is a real problem!
  - Models are trivially falsified in that they are all wrong - just look at the last century simulations.
  - The future is always out-of-sample, so you will never know if your model is good enough.
- But you can check that your model behaves as expected (In GMD, this is called “evaluation”).

How to build complex interdisciplinary models that remain falsifiable?

- If you are predicting future climate, even if your model is super-simple, this is a real problem!
  - Models are trivially falsified in that they are all wrong - just look at the last century simulations.
  - The future is always out-of-sample, so you will never know if your model is good enough.
- But you can check that your model behaves as expected (In GMD, this is called “evaluation”).
- And you can compare models to observations of past and present and to each other to adjust your confidence in your model’s ability (these papers mostly in GRL, JClim, ClimDyn, CP etc).

How to build complex interdisciplinary models that remain falsifiable?

How to build complex interdisciplinary models that remain falsifiable?

- For other fields this can be less of a problem, as we believe that the thing we are predicting is quite like the thing we use to evaluate the model (eg. yesterday's weather forecast, last year's oil spill, last decade's tsunami, El Nino, or volcanic eruption).

How to build complex interdisciplinary models that remain falsifiable?

- For other fields this can be less of a problem, as we believe that the thing we are predicting is quite like the thing we use to evaluate the model (eg. yesterday's weather forecast, last year's oil spill, last decade's tsunami, El Nino, or volcanic eruption).
- -> diversity of approaches depending on the problem.

Geoscientific Modelling through the filter of GMD

Geoscientific Modelling through the filter of GMD

- Looking through papers submitted in 2015 and already published, and focussing on questions raised
  - Approaches to modelling, Evaluation of models and Interdisciplinary collaboration

Geoscientific Modelling through the filter of GMD

- Looking through papers submitted in 2015 and already published, and focussing on questions raised
  - Approaches to modelling, Evaluation of models and Interdisciplinary collaboration
- At GMD, code must be made available

Geoscientific Modelling through the filter of GMD

- Looking through papers submitted in 2015 and already published, and focussing on questions raised
  - Approaches to modelling, Evaluation of models and Interdisciplinary collaboration
- At GMD, code must be made available
- Tone of almost all GMD papers is encouraging others to collaborate and to use the code that is presented. Many suggest their code is useful for varied applications.

GMD research areas

GMD research areas

Atmospheric Sciences
Climate and Earth System Modeling
Biogeosciences
Numerical Methods
Cryosphere
Hydrology
Oceanography
Solid Earth
Earth and Space Science Informatics
Solar-terrestrial Science
Total

GMD research areas

	number of papers published
Atmospheric Sciences	277
Climate and Earth System Modeling	245
Biogeosciences	88
Numerical Methods	48
Cryosphere	38
Hydrology	30
Oceanography	28
Solid Earth	14
Earth and Space Science Informatics	12
Solar-terrestrial Science	4
Total	784

GMD research areas

	number of papers published	percentage rejected/ withdrawn
Atmospheric Sciences	277	18%
Climate and Earth System Modeling	245	16%
Biogeosciences	88	17%
Numerical Methods	48	27%
Cryosphere	38	10%
Hydrology	30	30%
Oceanography	28	24%
Solid Earth	14	52%
Earth and Space Science Informatics	12	56%
Solar-terrestrial Science	4	0%
Total	784	20%

A multitude of modelling approaches

A multitude of modelling approaches

- Are necessary: we **need** different sorts of models for different purposes.

A multitude of modelling approaches

- Are necessary: we **need** different sorts of models for different purposes.
- **CranSLIK v2.0: improving the stochastic prediction of oil spill transport and fate using approximation methods (Rutherford et al 2015)**
  - Oceanography: **Stochastic model** for oil spills. Other models are fluid dynamics based. This one is fast and makes big ensembles, for **fast response** when oil spills really happen!

A multitude of modelling approaches

- Are necessary: we **need** different sorts of models for different purposes.
- **CranSLIK v2.0: improving the stochastic prediction of oil spill transport and fate using approximation methods (Rutherford et al 2015)**
  - Oceanography: **Stochastic model** for oil spills. Other models are fluid dynamics based. This one is fast and makes big ensembles, for **fast response** when oil spills really happen!
- **Modelling the dispersion of particle numbers in five European cities (Kukkonen et al 2016)**
  - Atmosphere: For practical reasons **different local systems** for each city, from meteorological models through to data-based.

A multitude of modelling approaches

- Are necessary: we **need** different sorts of models for different purposes.
- **CranSLIK v2.0: improving the stochastic prediction of oil spill transport and fate using approximation methods (Rutherford et al 2015)**
  - Oceanography: **Stochastic model** for oil spills. Other models are fluid dynamics based. This one is fast and makes big ensembles, for **fast response** when oil spills really happen!
- **Modelling the dispersion of particle numbers in five European cities (Kukkonen et al 2016)**
  - Atmosphere: For practical reasons **different local systems** for each city, from meteorological models through to data-based.
- **r.randomwalk v1, a multi-functional conceptual tool for mass movement routing (Mergill et al 2015)**
  - Solid Earth: **Monte Carlo approach** for, “a flexible open-source tool for backward and forward analyses of mass movement propagation.” Landslides etc.

A multitude of modelling approaches

- Are necessary: we **need** different sorts of models for different purposes.
- **CranSLIK v2.0: improving the stochastic prediction of oil spill transport and fate using approximation methods (Rutherford et al 2015)**
  - Oceanography: **Stochastic model** for oil spills. Other models are fluid dynamics based. This one is fast and makes big ensembles, for **fast response** when oil spills really happen!
- **Modelling the dispersion of particle numbers in five European cities (Kukkonen et al 2016)**
  - Atmosphere: For practical reasons **different local systems** for each city, from meteorological models through to data-based.
- **r.randomwalk v1, a multi-functional conceptual tool for mass movement routing (Mergill et al 2015)**
  - Solid Earth: **Monte Carlo approach** for, "a flexible open-source tool for backward and forward analyses of mass movement propagation." Landslides etc.
- **The Met Office Global Coupled model 2.0 (GC2) configuration, (Williams et al 2015)**
  - Climate: **General Circulation Model;** Seasonal Prediction to Climate Prediciton

Translating between different models and data

Translating between different models and data

- **GO2OGS 1.0: a versatile workflow to integrate complex geological information with fault data into numerical simulation models (Fischer et al 2015)**
  - "**Conversion workflows** are needed as a way of communication between the diverse tools of the various disciplines. Our approach offers an **open-source**, platform-independent, robust, and comprehensible method that is potentially useful for a **multitude of environmental studies**."

Translating between different models and data

- **GO2OGS 1.0: a versatile workflow to integrate complex geological information with fault data into numerical simulation models (Fischer et al 2015)**
  - "**Conversion workflows** are needed as a way of communication between the diverse tools of the various disciplines. Our approach offers an **open-source**, platform-independent, robust, and comprehensible method that is potentially useful for a **multitude of environmental studies**."
- **An open and extensible framework for spatially explicit land use change modelling: the lulcc R package (Moulds et al 2015)**
  - Breaking barriers with **open-source code.** "It is envisaged that lulcc will enable future model development and comparison within an open environment.

Interactions between models - coupling

Interactions between models - coupling

- **EwE-F 1.0: An implementation of Ecopath with Ecosim in Fortran 95/2003 for coupling and integration with other models (Akoglu et al 2015)**
  - For coupling and integration with biogeochemical and hydrodynamic models, "a base for **complex interdisciplinary modelling integrations**."

Interactions between models - coupling

- **EwE-F 1.0: An implementation of Ecopath with Ecosim in Fortran 95/2003 for coupling and integration with other models (Akoglu et al 2015)**
  - For coupling and integration with biogeochemical and hydrodynamic models, "a base for **complex interdisciplinary modelling integrations**."
- **Implementation of the Community Earth System Model (CESM) version 1.2.1 as a new base model into version 2.50 of the MESSy framework (Baumgaertner et al 2016)**
  - Building a Super-Model! The Modular Earth Submodel System, now has two alternative GCMs - CESM and ECHAM5!

Interactions between disciplines 1

Interactions between disciplines 1

- **A fully coupled Atmosphere-Ocean Wave modeling system (WEW) for the Mediterranean Sea: interactions and sensitivity to the resolved scales and mechanisms (Katsafados et al 2016)**
  - **Atmosphere+Ocean:** Coupled atmosphere-ocean wave modelling

Interactions between disciplines 1

- **A fully coupled Atmosphere-Ocean Wave modeling system (WEW) for the Mediterranean Sea: interactions and sensitivity to the resolved scales and mechanisms (Katsafados et al 2016)**
  - **Atmosphere+Ocean:** Coupled atmosphere-ocean wave modelling
- **Prediction of cloud condensation nuclei activity for organic compounds using functional group contribution methods (Petters et al 2016)**
  - **Atmosphere+Biogeosciences:** organic compounds and clouds.

Interactions between disciplines 1

- **A fully coupled Atmosphere-Ocean Wave modeling system (WEW) for the Mediterranean Sea: interactions and sensitivity to the resolved scales and mechanisms (Katsafados et al 2016)**
  - **Atmosphere+Ocean:** Coupled atmosphere-ocean wave modelling
- **Prediction of cloud condensation nuclei activity for organic compounds using functional group contribution methods (Petters et al 2016)**
  - **Atmosphere+Biogeosciences:** organic compounds and clouds.
- **FORest Canopy Atmosphere Transfer (FORCAsT) 1.0: a 1-D model of biosphere-atmosphere chemical exchange (Ashworth et al 2015)**
  - **Atmosphere+Biogeosciences:** Forest organic compounds released into atmosphere

Interactions between disciplines 1

- **A fully coupled Atmosphere-Ocean Wave modeling system (WEW) for the Mediterranean Sea: interactions and sensitivity to the resolved scales and mechanisms (Katsafados et al 2016)**
  - **Atmosphere+Ocean:** Coupled atmosphere-ocean wave modelling
- **Prediction of cloud condensation nuclei activity for organic compounds using functional group contribution methods (Petters et al 2016)**
  - **Atmosphere+Biogeosciences:** organic compounds and clouds.
- **FORest Canopy Atmosphere Transfer (FORCAsT) 1.0: a 1-D model of biosphere-atmosphere chemical exchange (Ashworth et al 2015)**
  - **Atmosphere+Biogeosciences:** Forest organic compounds released into atmosphere
- Anthropogenic modelling: Agriculture in land surface models **(Wu et al 2016)** Wind farms in atmospheric models. **(Volker et al 2015)** Agricultural trees in ecosystem model **(Fader et al 2015)**

Interactions between disciplines 2

Interactions between disciplines 2

- **Open-source modular solutions for flexural isostasy: gFlex v1.0 (Wickert 2016)**
  - **Solid Earth + Cryosphere**: "gFlex may be used to better understand how ice sheets, glaciers, large lakes, sedimentary basins, volcanoes, and other surface loads interact with the solid Earth."

Interactions between disciplines 2

- **Open-source modular solutions for flexural isostasy: gFlex v1.0 (Wickert 2016)**
  - **Solid Earth + Cryosphere**: "gFlex may be used to better understand how ice sheets, glaciers, large lakes, sedimentary basins, volcanoes, and other surface loads interact with the solid Earth."
- **FPLUME-1.0: An integral volcanic plume model accounting for ash aggregation Folch et al 2016)**
  - **Solid Earth+Atmosphere:** Volcanic plumes - solid earth meets atmospheric science

Interactions between disciplines 2

- **Open-source modular solutions for flexural isostasy: gFlex v1.0 (Wickert 2016)**
  - **Solid Earth + Cryosphere**: "gFlex may be used to better understand how ice sheets, glaciers, large lakes, sedimentary basins, volcanoes, and other surface loads interact with the solid Earth."
- **FPLUME-1.0: An integral volcanic plume model accounting for ash aggregation Folch et al 2016)**
  - **Solid Earth+Atmosphere:** Volcanic plumes - solid earth meets atmospheric science
- **SHIMMER (1.0): A novel mathematical model for microbial and biogeochemical dynamics in glacier forefield ecosystems (Bradley et al 2015)**
  - **Biogeosciences+Cryosphere**: Ecosystem development in exposed glacier forefields

Open source tools

Open source tools

Open source tools

- **Distributed visualization of gridded geophysical data: The Carbon Data Explorer, version 0.2.3 (Endsley et al 2016)**
  - Manage, aggregate, visualize, and share gridded data sets. Web-based, “It is hoped that it will inspire the future development of similar tools.”

Open source tools

- **Distributed visualization of gridded geophysical data: The Carbon Data Explorer, version 0.2.3 (Endsley et al 2016)**
  - Manage, aggregate, visualize, and share gridded data sets. Web-based, “It is hoped that it will inspire the future development of similar tools.”
- **PRACTISE – Photo Rectification And ClassificaTIon SoftwarE (V.2.1) (Härer et al 2016)**
  - Finding snow in photos. “The presented method is not limited to the sensor system or the threshold used in here but offers **manifold application options for other scientific branches.**”

Open source tools

- **Distributed visualization of gridded geophysical data: The Carbon Data Explorer, version 0.2.3 (Endsley et al 2016)**
  - Manage, aggregate, visualize, and share gridded data sets. Web-based, "It is hoped that it will inspire the future development of similar tools."
- **PRACTISE – Photo Rectification And ClassificaTIon SoftwarE (V.2.1) (Härer et al 2016)**
  - Finding snow in photos. "The presented method is not limited to the sensor system or the threshold used in here but offers **manifold application options for other scientific branches.**"
- **NCAR_Topo (v1.0): NCAR global model topography generation software for unstructured grids (Lauritzen et al 2015)**
  - **Open-source** tool for generating topography

Comparing models to observations and to each other

Comparing models to observations and to each other

- Models published in GMD must be shown to work and many papers contain a lot of model evaluation.

Comparing models to observations and to each other

- Models published in GMD must be shown to work and many papers contain a lot of model evaluation.
- **The GRENE-TEA Model Intercomparison Project (GTMIP): Overview and experiment protocol for Stage 1 (Miyazaki et al 2015)**
  - A Japanese-funded project on the Arctic which aims to, "enhance communication and understanding between the modelling and field scientists".

Comparing models to observations and to each other

- Models published in GMD must be shown to work and many papers contain a lot of model evaluation.
- **The GRENE-TEA Model Intercomparison Project (GTMIP): Overview and experiment protocol for Stage 1 (Miyazaki et al 2015)**
  - A Japanese-funded project on the Arctic which aims to, "enhance communication and understanding between the modelling and field scientists".
- Climate Model Inter-comparison Project: CMIP6 (special issue: **Coupled Model Intercomparison Project Phase 6 (CMIP6) Experimental Design and Organization**)
  - >20 MIPs plus data for forcing and comparison to models.

Data assimilation across Geoscience

Data assimilation across Geoscience

- Our best predictions combine models and data (includes some pre-2015 examples)

Data assimilation across Geoscience

- Our best predictions combine models and data (includes some pre-2015 examples)
- **Development and application of the WRFPLUS-Chem online chemistry adjoint and WRFDA-Chem assimilation system (Guerrette et al 2015)**
  - Atmosphere - chemistry
- **Adjoint Global sensitivity analysis, probabilistic calibration, and predictive assessment for the data assimilation linked ecosystem carbon model (Safta et al 2014)**
  - Biogeochemistry - ecosystem
- **Streamflow data assimilation for soil moisture analysis (Warrach-Sagi et al 2010)**
  - Hydrology - streamflow
- **Physically Based Data assimilation (Levy et al 2010)**
  - Cryosphere - sea-ice as a test case
- **ECCO version 4: an integrated framework for non-linear inverse modeling and global ocean state estimation (Forget et al 2015)**
  - Ocean - state estimation using an adjoint

Conclusions

Conclusions

- GMD publications show that scientists are already doing amazing things in the perceived problem areas of Geoscientfic modelling.

Conclusions

- GMD publications show that scientists are already doing amazing things in the perceived problem areas of Geoscientfic modelling.
- Do we need a generalised frameworks for model building and evaluation?

Conclusions

- GMD publications show that scientists are already doing amazing things in the perceived problem areas of Geoscientfic modelling.
- Do we need a generalised frameworks for model building and evaluation?
  - No - different problems need different approaches.

Conclusions

- GMD publications show that scientists are already doing amazing things in the perceived problem areas of Geoscientfic modelling.
- Do we need a generalised frameworks for model building and evaluation?
  - No - different problems need different approaches.
  - But within these approaches, discussion of frameworks is appropriate.

Conclusions

- GMD publications show that scientists are already doing amazing things in the perceived problem areas of Geoscientfic modelling.
- Do we need a generalised frameworks for model building and evaluation?
  - No - different problems need different approaches.
  - But within these approaches, discussion of frameworks is appropriate.
- Many of the problems may be reduced by making code open source, and describing models and other code in peer reviewed publications. Keen scientists will do the rest!

Conclusions

- GMD publications show that scientists are already doing amazing things in the perceived problem areas of Geoscientfic modelling.
- Do we need a generalised frameworks for model building and evaluation?
  - No - different problems need different approaches.
  - But within these approaches, discussion of frameworks is appropriate.
- Many of the problems may be reduced by making code open source, and describing models and other code in peer reviewed publications. Keen scientists will do the rest!
  - These days, developers of cool code tend to be very open to collaboration - they want their code to be used and developed.

Conclusions

- GMD publications show that scientists are already doing amazing things in the perceived problem areas of Geoscientfic modelling.
- Do we need a generalised frameworks for model building and evaluation?
  - No - different problems need different approaches.
  - But within these approaches, discussion of frameworks is appropriate.
- Many of the problems may be reduced by making code open source, and describing models and other code in peer reviewed publications. Keen scientists will do the rest!
  - These days, developers of cool code tend to be very open to collaboration - they want their code to be used and developed.
  - **Please join in: read the papers, and use the code!**