Met Office

Upper-Air Physics parallel session

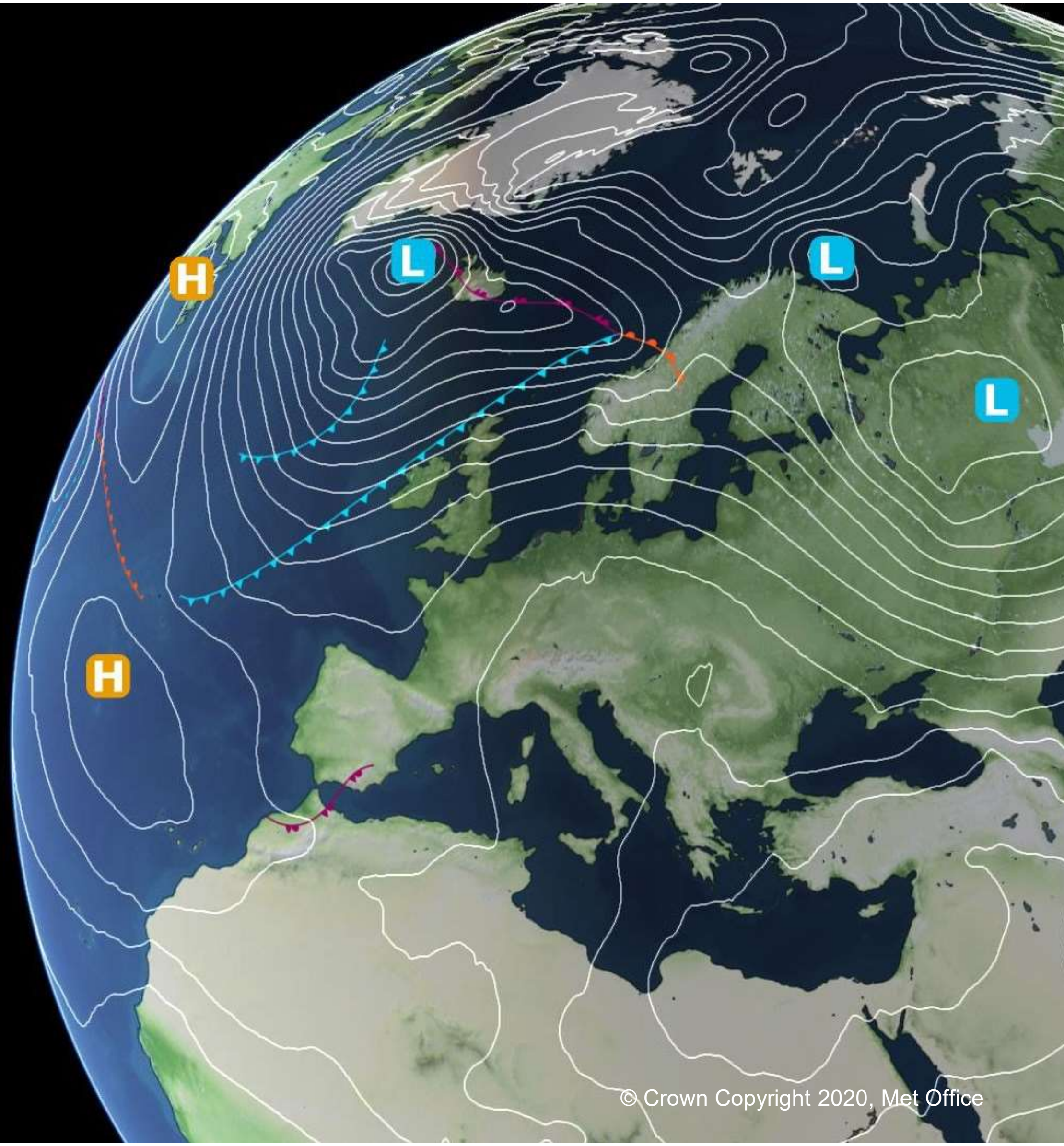
45th EWGLAM and 30th SRNWP meeting,
Reykjavík, Iceland

27th September 2023

Chaired by Mike Bush

With thanks to many colleagues including:

Lewis Blunn…



Proposed discussion topics

- **"Uncertainty of modelling components and their impact“ – how does this relate to model physics?**
- Urban scale modelling
- Scale-aware parametrizations
- AI/ML approaches
- Unified physics
- Given finite resources, how do we prioritise these activities?

Met Office

Uncertainty of modelling components and their impact (thanks to Anne Mccabe)

- **1) Which aspects of our modelling systems are known to be uncertain?**

- Mechanisms for identifying model uncertainty:
  - Experience of model developers
  - Feedback from forecasters and end users
  - Forensic evaluation of our model (processed based evaluation, statistical evaluation of analysis increments and model tendencies)
- How do these uncertainties vary between models?
- Are (for example) the uncertainties predominantly in particular schemes for all models, or do different models have significantly different uncertainty characteristics?

Met Office

Uncertainty of modelling components and their impact (thanks to Anne Mccabe)

2) How much does the uncertainty impact the predictability of the system?

• Are the uncertainties systematic (mean biases) or random?

• Which aspects of the forecast are we never getting right? Which aspects do we sometimes get right and sometimes not?

• How does error growth associated with the uncertainty grow on different time and spatial scales?

• Do the known uncertainties in a parametrization have the expected impact on the predictability of the system? E.g. is the model as sensitive to the known uncertainties as we would expect?

Met Office

Uncertainty of modelling components and their impact (thanks to Anne Mccabe)

3) Which processes do we know to be inherently chaotic and need to be represented as such?

•Where should we focus our attention when developing stochastic physics schemes?

Met Office

Urban Workshop with Special Focus on Anthropogenic Heat Emissions (thanks to Lewis Blunn)

- There was an urban (hectometric scale) workshop on $11^{th}/12^{th}$ September 2023 held at the Bureau of Meteorology, Melbourne
- It was suggested that urban-scale (25-300 m) models are currently not appropriate for modelling screen-level meteorological variables for the purpose of Urban Climate Services, since the model parametrisations break down at sub-km scales.
- The input building information is often still not available at urban-scale resolution. The blending height is at the bottom model level (i.e., ~2 m) which is too low because it means e.g., the 1.5 m air temperature is approximately the same over each land cover tile.
- Also, to properly represent the screen-level air temperature a vertically distributed urban canopy scheme needs to be developed.

Urban Workshop with Special Focus on Anthropogenic Heat Emissions (thanks to Lewis Blunn)

- Proposal for a new AHE (anthropogenic heat emission) model for the UM/JULES.
- The new AHE model should be globally applicable, flexible so that input data can be updated, and "dynamic" in that it responds to and feeds back on the outdoor environment.
- If all the desired data were globally available, the model should split AHEs into six components related to human metabolism, vehicles, industry, non-temperature related domestic energy consumption, heating energy consumption, and cooling energy consumption.
- In the first stage of the AHE model development the six components will be combined.

Met Office

Urban Workshop with Special Focus on Anthropogenic Heat Emissions (thanks to Lewis Blunn)

- There was discussion on whether machine learning (ML) might be a better AHE modelling approach, due to AHE models being largely data driven rather than physics driven.
- However, there was push back since global datasets of the required data to train ML models are not available. It was argued that it is better to make simple well-understood assumptions to achieve a global AHEs model.
- It was suggested that a ML approach might be worth exploring for limited regions (e.g., for Australia) where high-quality data is available.

Met Office

Scale-aware parametrizations (thanks to Humphrey Lean)

- In the UK we are just starting a 4 year joint Met Office/academic programme (funded by NERC) called "Turbulent Processes"
- This will aim to answer questions (i.e. development of grey zone parameterisations and including their effect on predictability).
- The Met Office side of this is led by Humphrey Lean and Alison Stirling.

Met Office

"Fusing simulations and data science"

Worth clarifying what we each mean. I propose defining:

Cyril Morcrette

Type 1:

Simulations appended by data science.

After model run, output is read by external algorithm and post-processed for end-user.

This can be done from multiple output times.

But the ML does not feed-back on the physical simulation.

Type 2:

A two-way coupling process,

with millions of model columns, distributed across many HPC processors,

calling an ML algorithm to predict something,

passing the information back into the weather/climate model,

updating the model state,

and that process repeating over thousands of model timesteps.

Type 3:

Full end-to-end emulation.

Of "just" the atmospheric model, or of the data assimilation process as well.

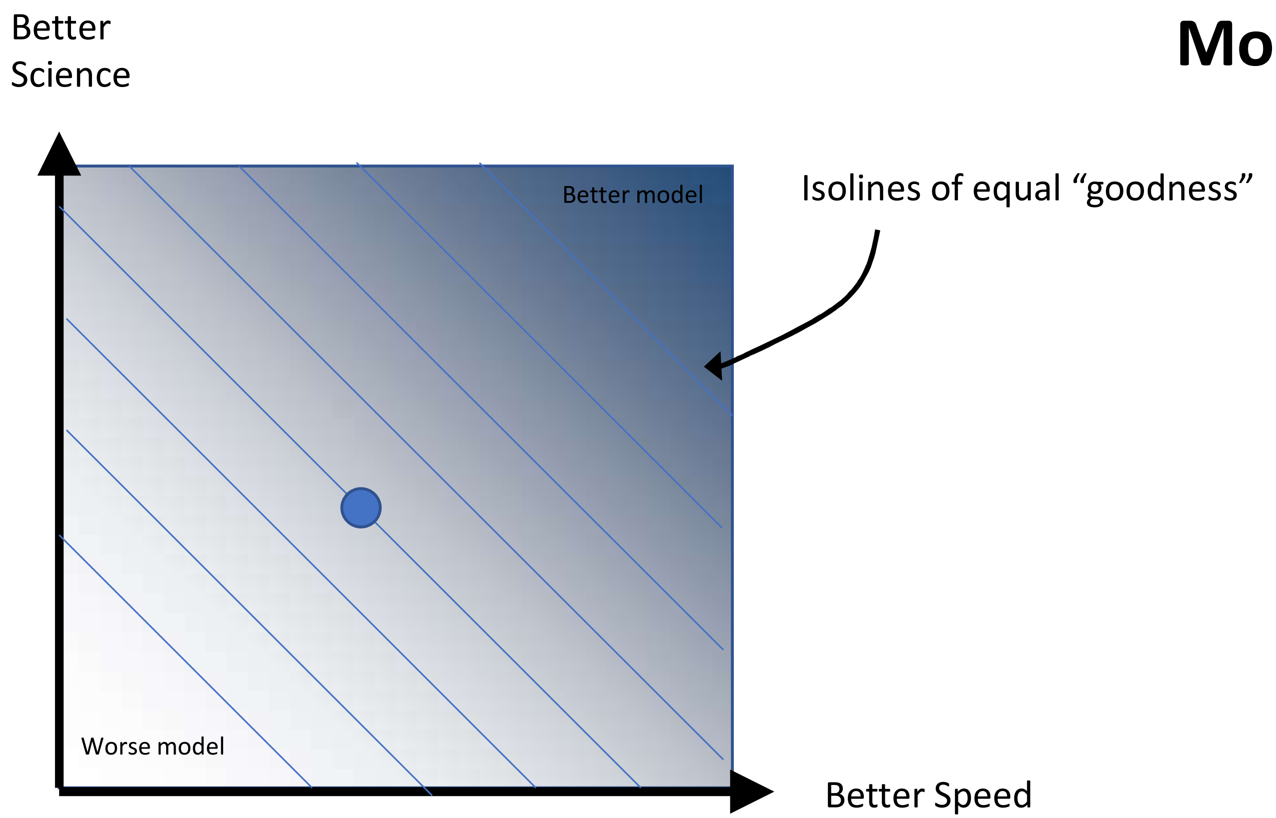
Cyril
Morcrette

Type 2:

A two-way coupling process,
with millions of model columns, distributed across many HPC processors,
calling an ML algorithm to predict something,
passing the information back into the weather/climate model,
updating the model state,
and that process repeating over thousands of model timesteps.

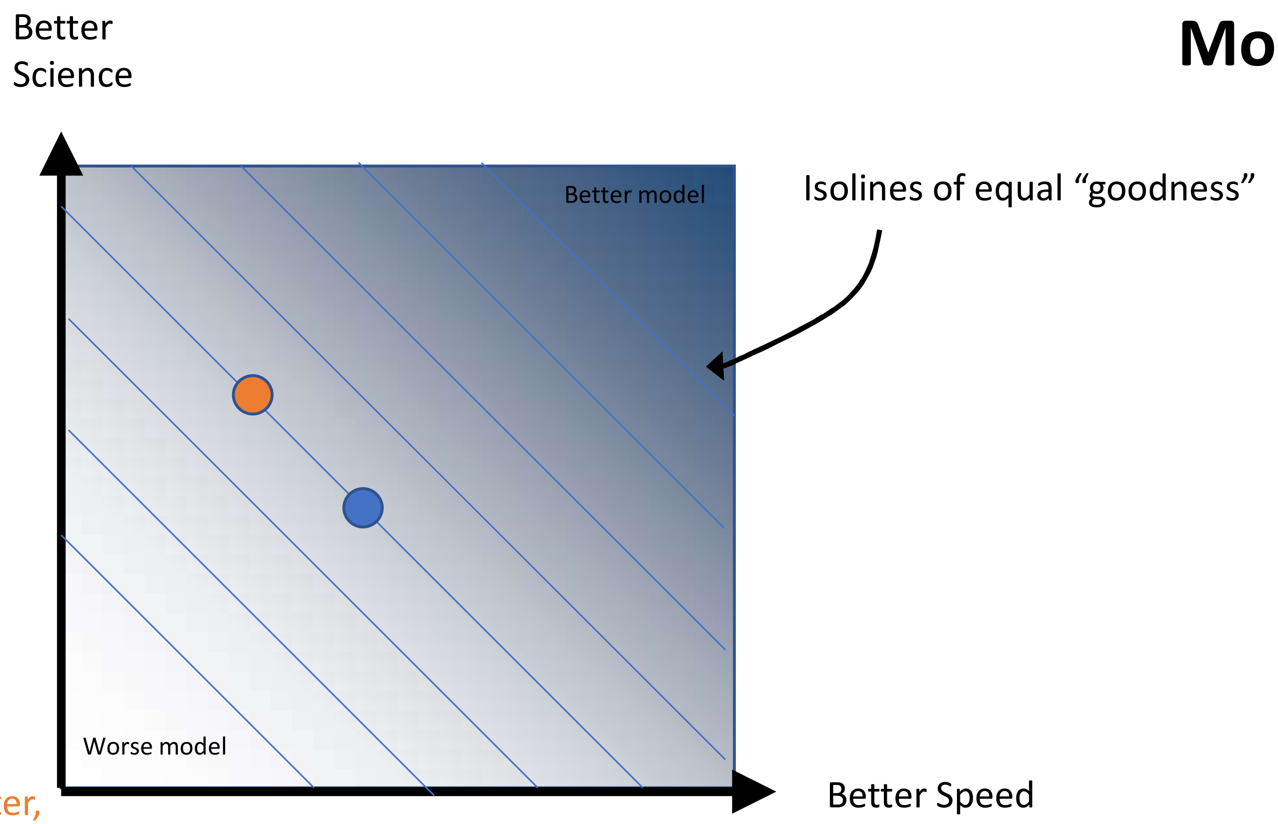
How can Type 2 fusion make the model "better" ?

Cyril Morcrette



Met Office

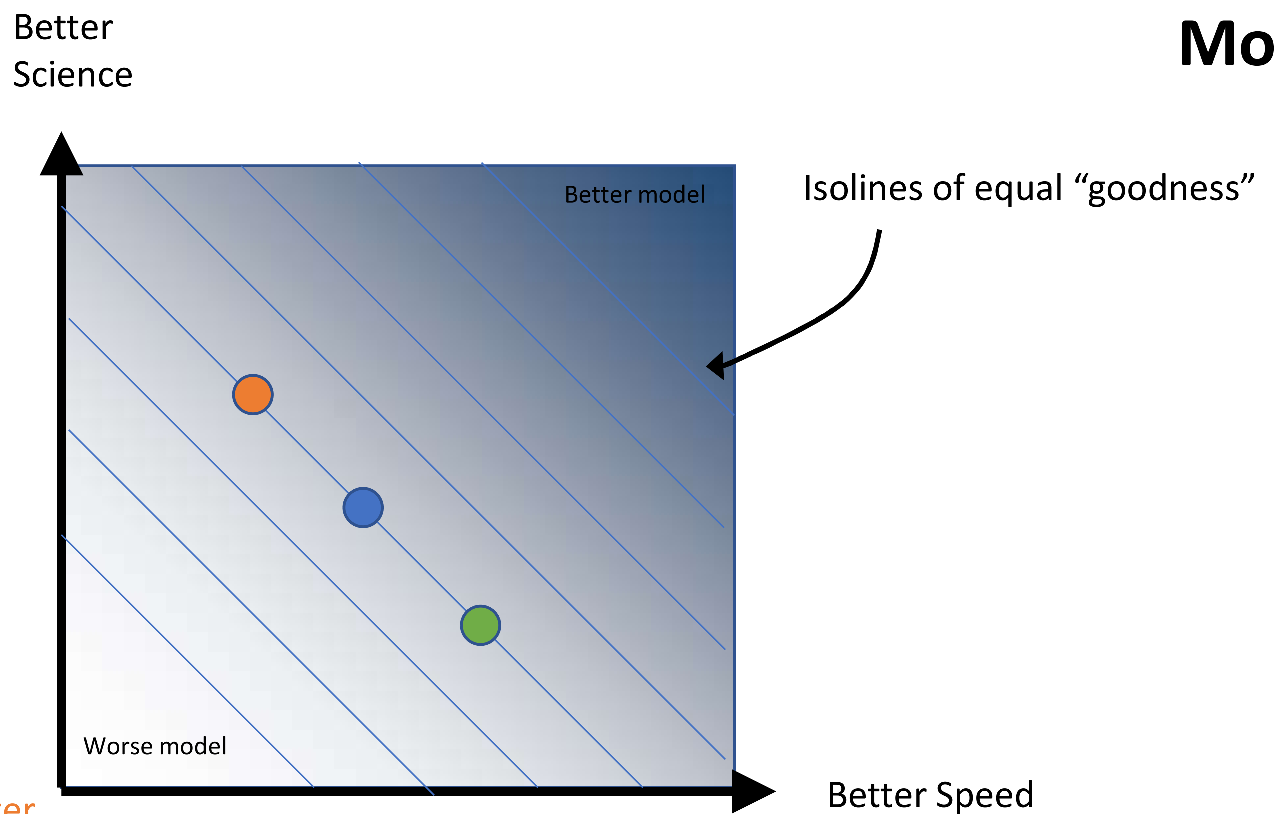
Cyril Morcrette



You can make the science better, but the model is more expensive and so overall, it is still just as "good".

Met Office

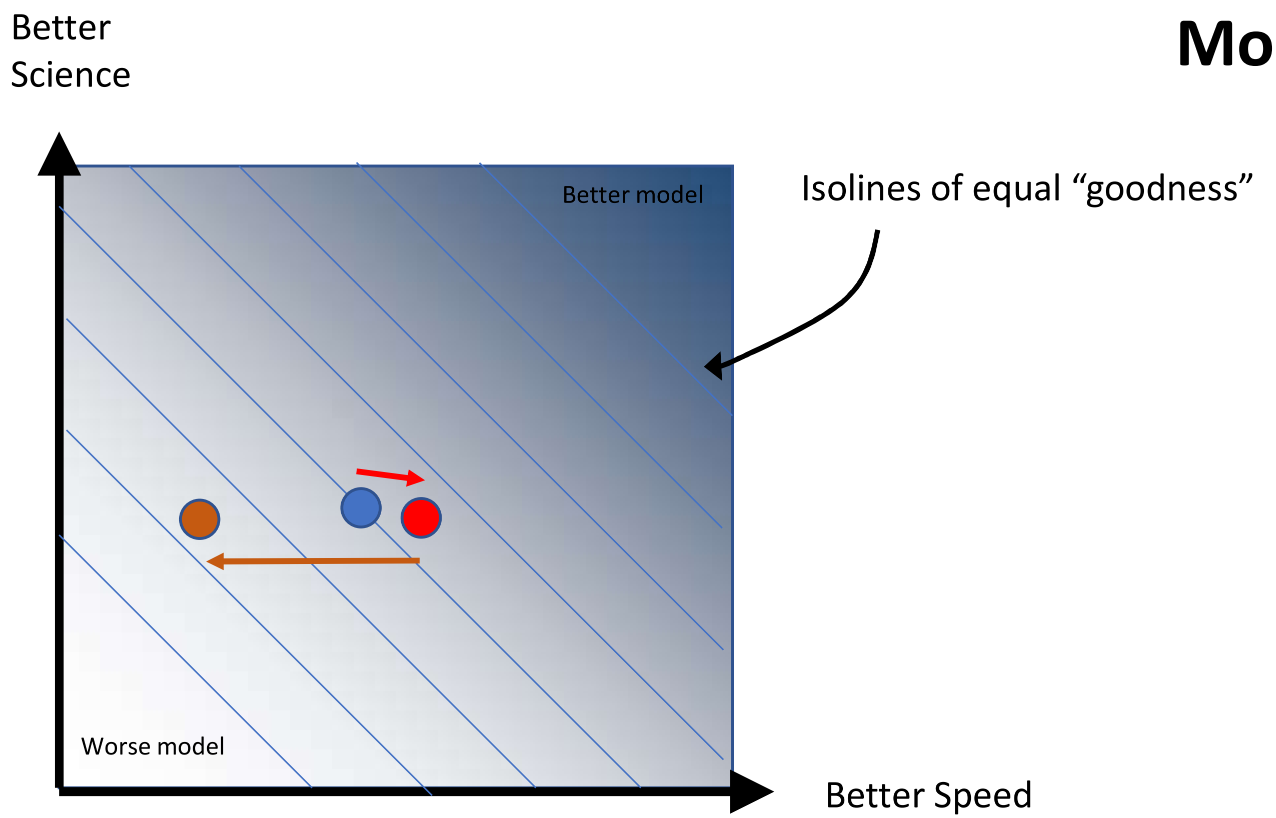
Cyril Morcrette



You can make the science better, but the model is more expensive and so overall, it is still just as “good”.

You can make the model “faster” by simplifying the science and by some metric it is still as “good”.

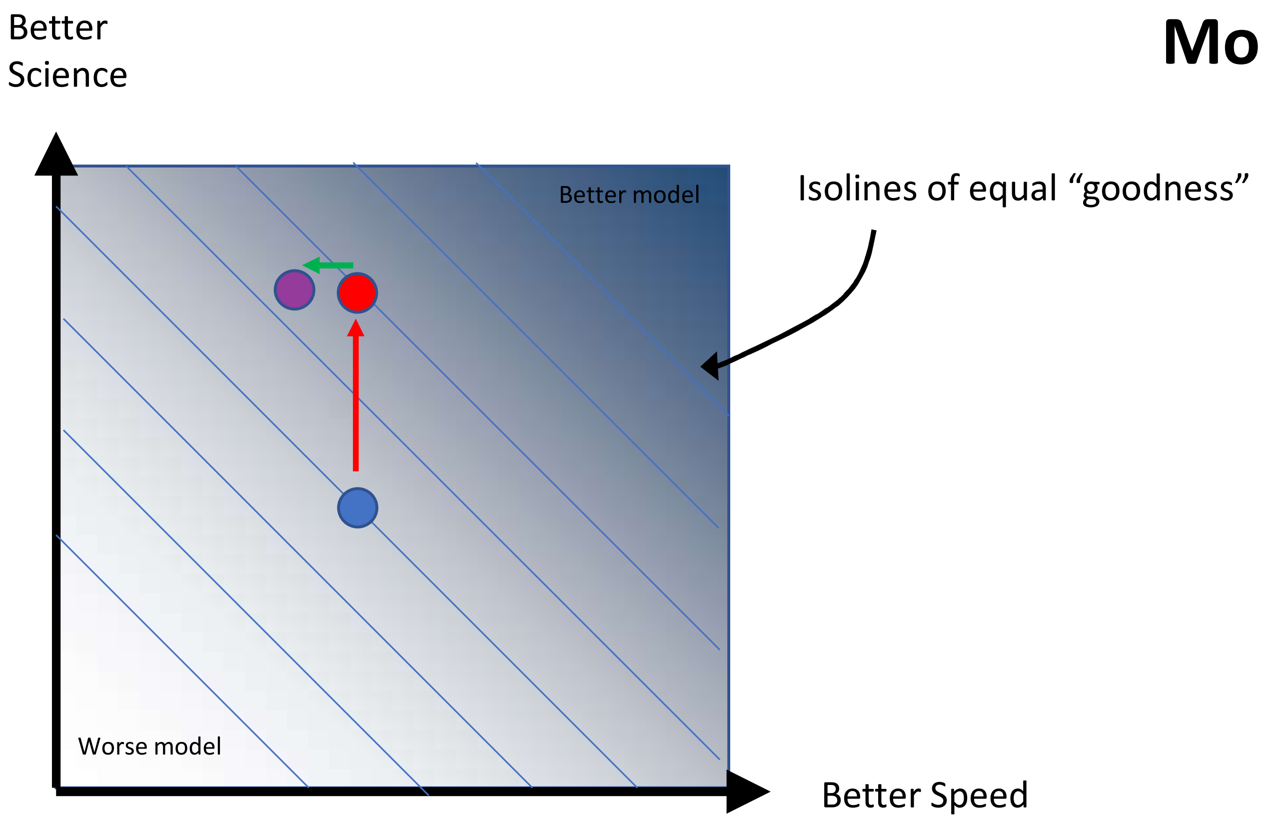
Cyril
Morcrette



Attempt to make the model better by coupling to an imperfect ML emulator of an existing scheme.
But the coupling is very inefficient and takes longer than the time gained by emulation.
Model is "worse" overall.

Met Office

Cyril Morcrette



Attempt to make the model better by using ML to represent a new process within the Fortran code.

Code is slightly more expensive, but no coupling overhead.

Model is “better” overall.

Met Office

Cyril Morcrette

Type 1:

Simulations appended by data science.

After model run, output is read by external algorithm and post-processed for end-user.

This can be done from multiple output times.

But the ML does not feed-back on the physical simulation.

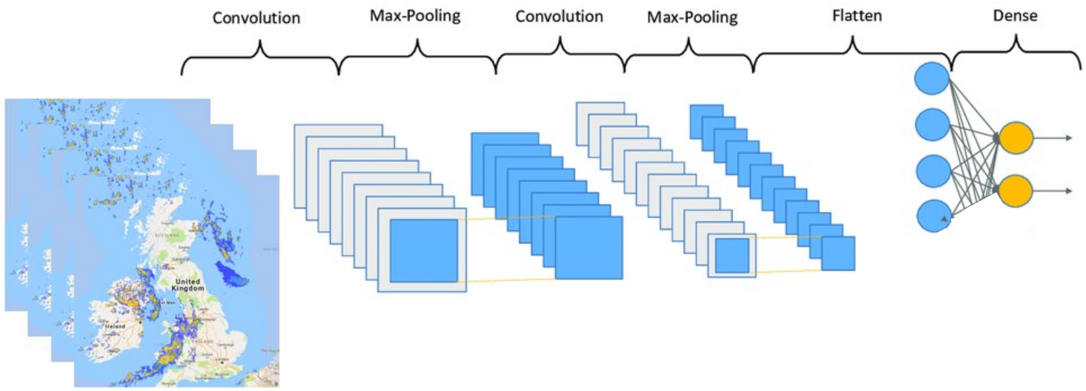
Example of Type 1 task

Met Office

Correcting forecast precipitation fields

Cyril Morcrette

Inputs:
Labelled imaged of either:
a) radar rainfall or
b) model rainfall.



1. Train a convolutional neural network (CNN) to distinguish a real radar map from a model rainfall field.

2. Then keep the neural network weights fixed.

3. Take a RAL3 model rainfall map, work back and forth through the network modifying the input model rainfall map (not the NN weights) until the network classifies it as real radar data.

4. This is a form of bias correction, but unlike quantile-quantile regression, it could e.g. reduce heavy rain-rate only on cold side of front (rather than everywhere), it could fill in gaps in squalls and it could create low rain-rate out of nothing (for scattered showers behind a front).

5. Obviously does not "fix the physics" and this post-processing won't feedback and affect model evolution, and could lead to physical inconsistencies, but it could be a way of getting packages of changes in. Allowing lots of new functionality and science to go in while being less hindered by compensating errors and interactions in new schemes.

Cyril Morcrette

		>10 years ago	Last ~10years	Present
		HG2	**GA3-8**	**GA9**
Global	Climate	Smith+CCA	PC2	Bi-modal initiation?
	NWP			

Kilometre-scale	UKV → RAM	Smith+EACF+ACF	Smith+EACF+ACF	Bi-modal?
	Singapore → RAT		PC2	Bi-modal initiation?

What we 'care' about	
Radiative balance, cloud-rad effects, …	
Large-scale flow, increasingly surface weather.	

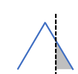
Surface weather (**cld**,**base**,T2M,ppn,wind,vis)	Smith and now bi-modal good for **cld** & **base**
Mainly **ppn**	PC2 good for **ppn**

Short-term	Medium Term
Keep things unified within GA	Then aim to unify GA and RA
Try to use hybrid scheme to unify… (i.e. bi-modal for liquid cloud, PC2 for ice, **hence get good cld, base AND ppn**).	

CCA=convective cloud amount
EACF= empirically adjusted cloud fraction
ACF=area cloud fraction

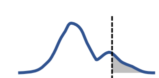
Diagnostic schemes: use **<u>current</u>** T,q,p to split qT into qv, qcl, qcf and find C.
Prognostic scheme (e.g. PC2): looks at **<u>all</u>** processes modifying T,q,p and calculates **<u>changes to</u>** qv, qcl, qcf, C.

Smith scheme
(Smith, 1990)



T,q variability represented using a single symmetric triangular function.

Bi-modal cloud scheme
(Van Weverberg et al. 2021)



T,q variability allowed to be skewed and with 2 peaks.

What is the “best” car?

This is the best car for reliability.

This is the best car for speed.

This is the best car for boot space.

… but I only want to run and maintain 1 car.

Cyril Morcrette

- If we try to unify and have a SINGLE car that is inevitably going to mean a car that is not quite a fast, not quite as good boot space and not quite a fuel efficient.
- If our metrics are the same as they were before, then the unification effort just seems like a compromise and a step backwards. By all measures the model is worse.
- It is only if we carefully, quantify the BENEFIT of unification (in terms of reduced staff cost from maintaining and testing multiple models) that we can see that unification is a benefit.
- One of the problems though is that the units are different. How do you compare RMSE of precip getting worse to FTE of staff cost of maintaining multiple models to bias in temperature?
- How do we have a combined metric that quantifies ALL aspects of what we care about, given maintenance cost etc is not something we measure or really consider when accessing whether to accept a new model configuration?

Met Office

Let the discussion begin!

www.metoffice.gov.uk