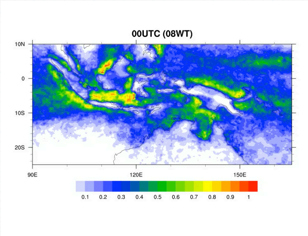


Evaluation of clouds in large-scale models

What? Why? How?

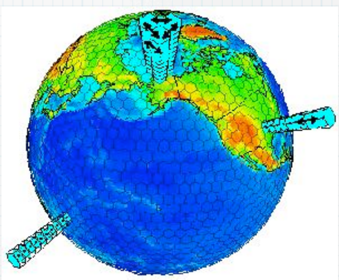
Christian Jakob & Jean-Louis Dufresne

The purpose of the three lectures

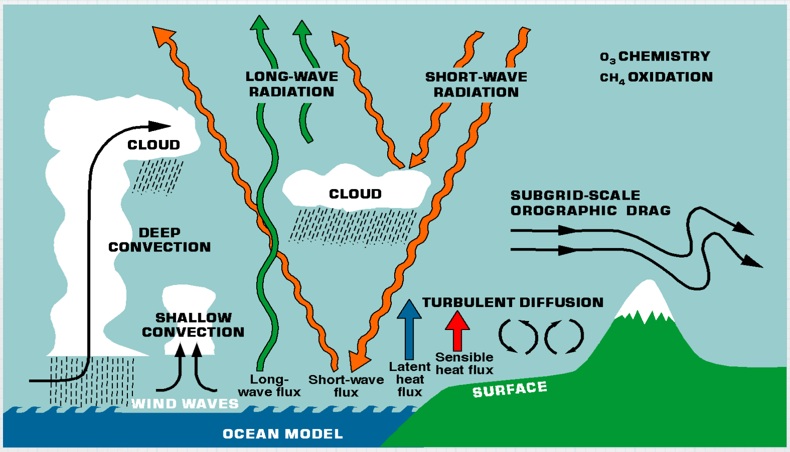
* **Motivate** model evaluation and diagnosis.
* Provide insight into **key approaches and techniques** of large-scale model evaluation with a focus on **concepts rather then details**.
* Provide **some results on the current state-of-the-art** in modelling cloud-related variables in climate models.

Motivation

* **We have built a model by**
  * choosing the equations
  * choosing coordinate systems
  * choosing numerical methods
  * choosing physical parametrizations
* Here we **limit ourselves to global models in which clouds and precipitation are parametrized.**



Parametrizations



What is model evaluation?

* Models are used in many different ways: Predictions at many time scales, research tools, ...
* In the broadest sense, model evaluation is about measuring how fit for purpose a particular model is.

What is fit for purpose?





Questions for model evaluation

* Does the model meet the **needs of the application** and to what extent, e.g., prediction?
* What are the **limits of its applicability**, e.g, ENSO?
* What are the **key errors influencing the performance** in the model application?
* What are the **causes of these errors**?

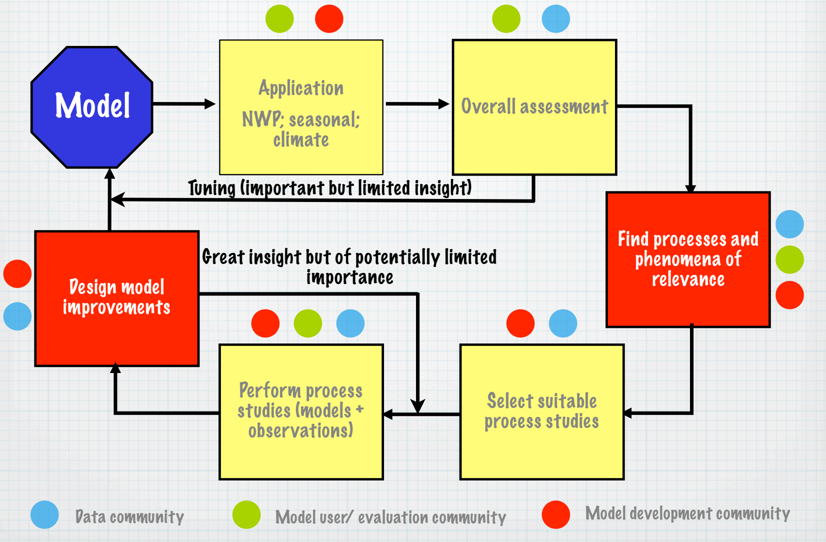
What is the truth?

* Model evaluation usually **needs an estimate of the "truth"...**

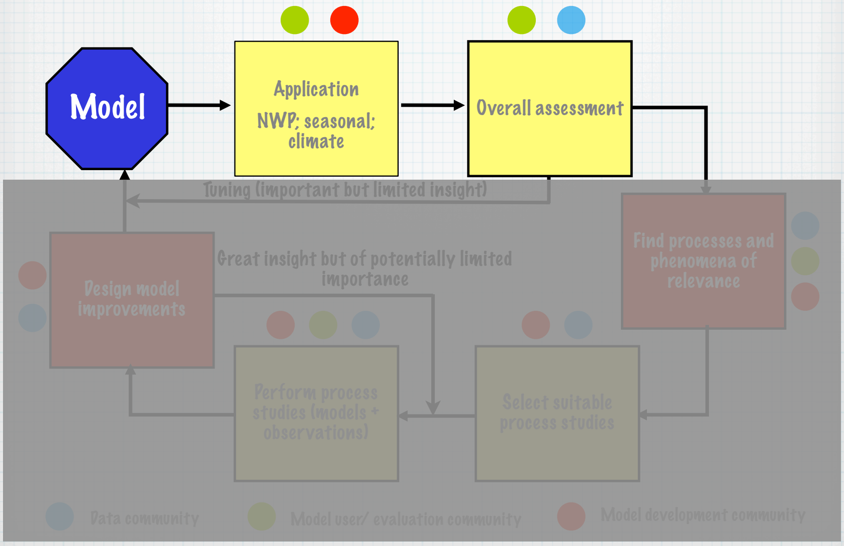
Limits of model evaluation

* Model evaluation usually **needs an estimate of the truth** and is hence limited by observations and knowledge.
* There is **no one-fits-all method** of model evaluation. Many **different approaches need to be combined** depending on purpose.
* **Connecting model error to model formulation** is one of the most **difficult** but often most desirable tasks of evaluation.

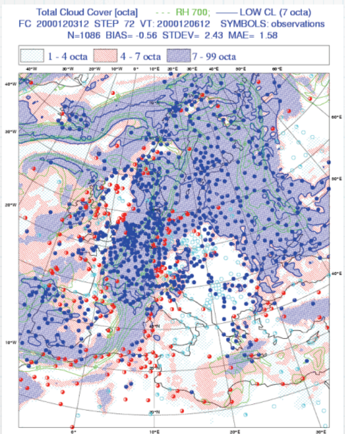
Approaches to model evaluation



Approaches to model evaluation



3-day cloud forecasts



ECMWF 72-h cloud forecast for 3 Dec 2000

Mean errors

For example:
f_n - Forecast at point n
o_n - Observation/Analysis at point n

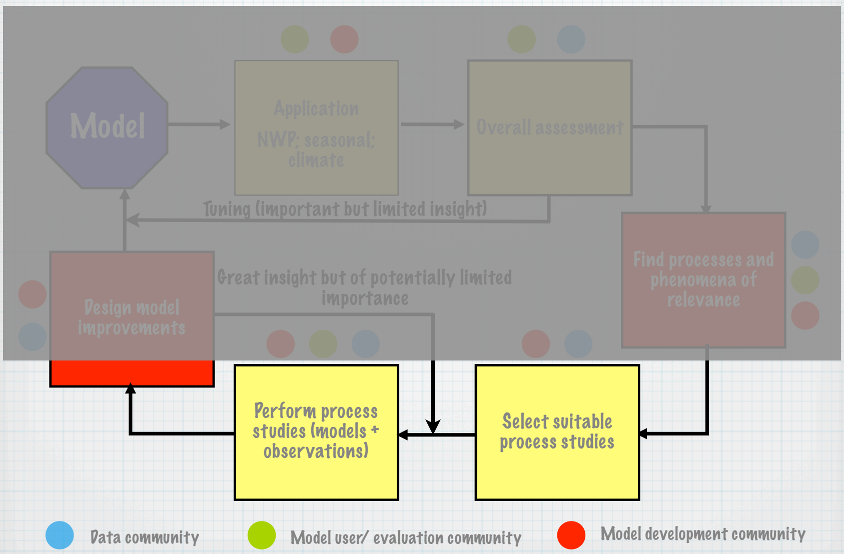
Mean error (Bias) $$ME = \frac{1}{N}\sum_{n=1}^{N}(f_n - o_n)$$

Mean absolute error $$MAE = \frac{1}{N}\sum_{n=1}^{N}|f_n - o_n|$$

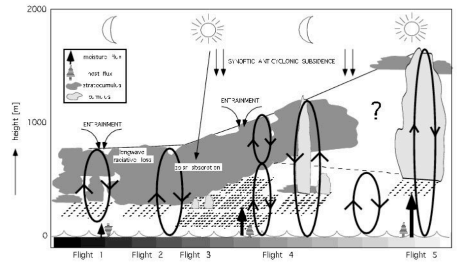
Model assessment in its application

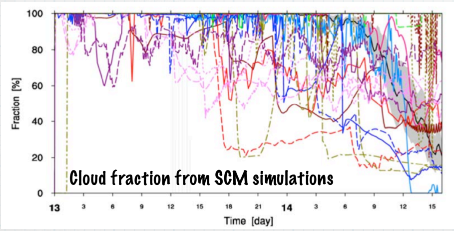
* **Pros:**
  * assesses what we care about
  * provides insight into limits of applicability of the model
* **Cons:**
  * provides very little insight into causes of model error

Approaches to model evaluation



"Process" studies



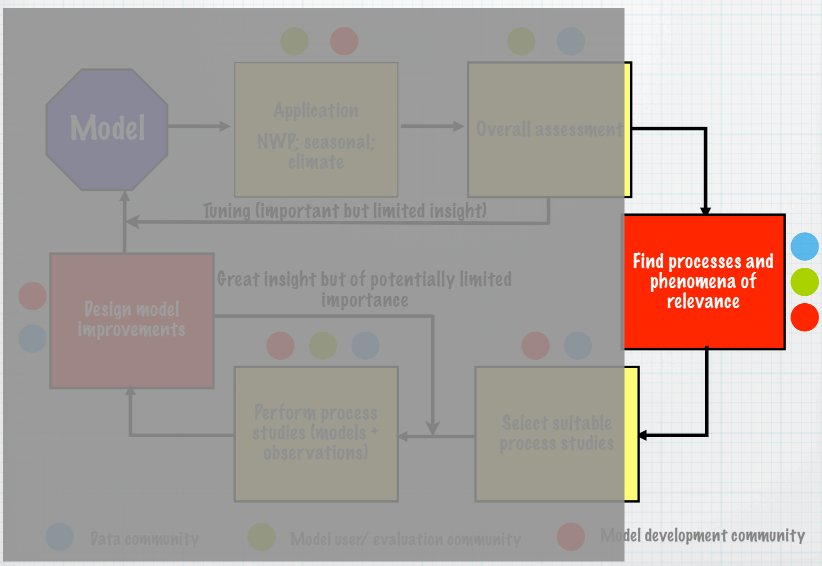


* Process studies aim to **isolate parts of the system** to study it in detail.
* "Truth" usually comes from **field experiments** in which many detailed observations are available.
* **Complex cloud models** as well as so-called **single column models** are often applied to simulate parts of the observed period
* Detailed comparisons at the "process level" can then be performed.

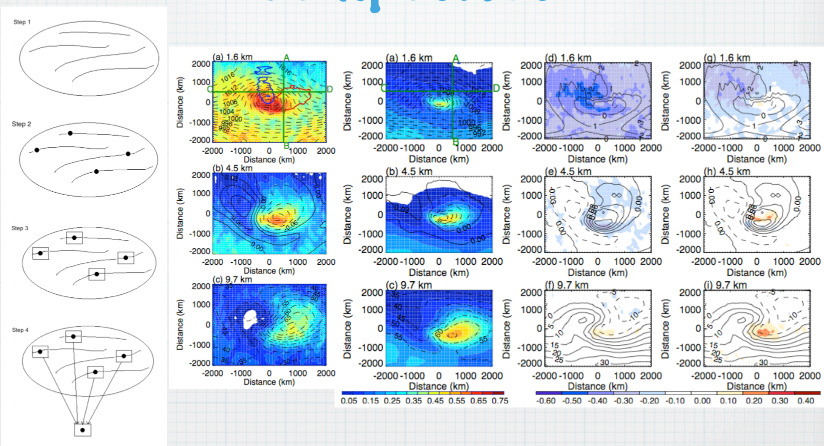
Process studies

* **Pros:**
  * allow detailed assessment of individual model processes
  * might provide great insight as to why things are happening
* **Cons:**
  * link to overall model errors is not always clear
  * better process representation does not guarantee better model results

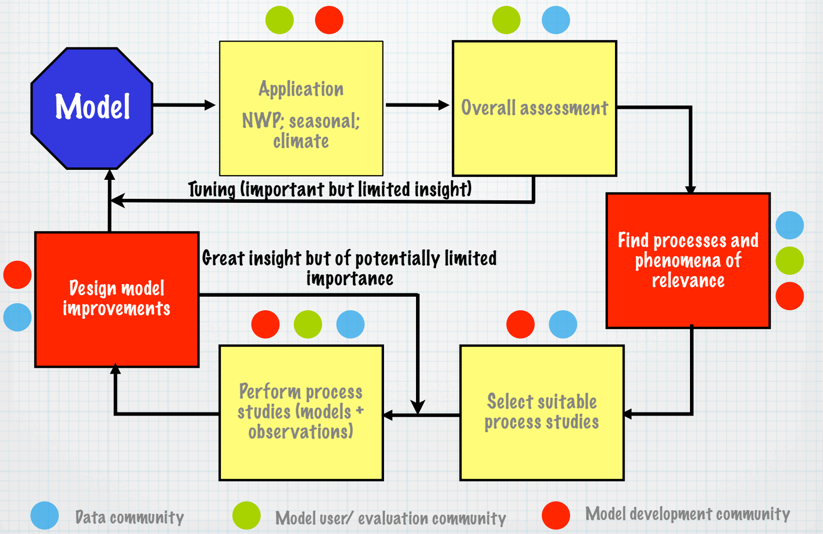
Approaches to model evaluation



Example: Cyclone Composites



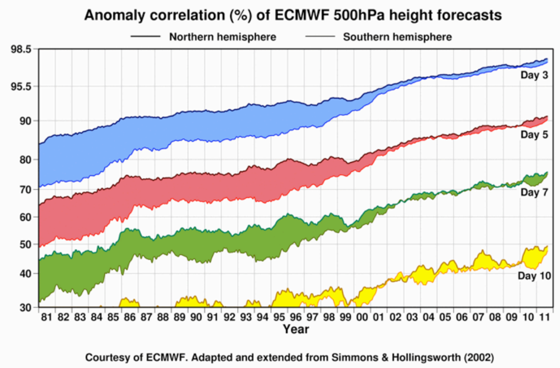
Model evaluation is a community effort



Qualitative vs quantitative evaluation

* Model evaluation can be **"qualitative"** or **"quantitative"**. Both are useful.
* **"Qualitative": Compare plots and make subjective assertions** on the quality of the model. Useful when a priori knowledge can be included into the assessment. Subjective part may introduce extra problems (e.g., rose-tinted glasses.)
* **"Quantitative": Calculate objective skill measures.** Useful when comparing models or model versions to each other and when monitoring the evolution of a model over long periods of time. Can mislead if used unqualified.

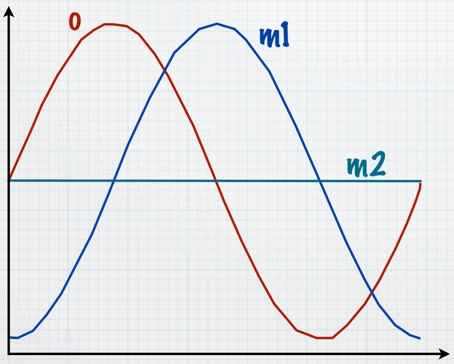
"Quantitative" Evaluation



500 hPa geopotential height anomaly correlation score for the ECMWF model
updated from Simmons and Hollingsworth (2002)

Quantitative evaluation

... is not always as simple as it seems!



$$o = sin(x) \text{ for } 0 < x < 2\pi$$

$$m1 = sin(x - \frac{\pi}{2}) \text{ for } 0 < x < 2\pi$$

$$m2 = 0$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(m_n - o_n)^2}$$

$RMSE_1 = 1$
$RMSE_2 = 0.7$

So which is the better model?