Why asking this question?

- We need models for climate change studies
- We use them for testing scientific hypotheses
- We use them as decision support tools
What is a ‘good’ hydrological model?
What is a ‘good’ hydrological model?

Other hydrological models...

Static models

\[ Q = aP + b \]

\[ Q = aP^b \]

\[ Q = aP^b E_0^c \]

\[ Q = aP + bE_0 + c \]

\[ Q = \frac{1 - \frac{1}{P}}{1 + \left( \frac{P}{E_0} \right)^n} \]

eetc…

Dynamic models

Elementary building blocks for dynamic rainfall-runoff models
What is a ‘good’ hydrological model?

Why asking this question?

- No hydrological model has achieved global supremacy over the others
- The number of models produced by hydrologists keeps rising

\[ one \ \text{hydrologist} \approx one \ \text{model} \]

- Producing one’s own model is certainly considered as a way to demonstrate scientific status and independence...
Why asking this question?

- We have to get used to live in a world of large model diversity…
Why asking this question?

- We have to get used to live in a world of large model diversity…

- … but need to be able to recognize their essential properties, qualities and drawbacks, the basic principles on which they are built

- Which hypothesis can we test with which model?

- Which decision can we base on which model?
What did famous scientists say on this topic?

(a personal selection)
A few citations…

« Make your model as simple as possible, but no simpler »
Albert Einstein (1879-1955)

- The level of complexity of a model should be kept as low as possible
- At equal performance, prefer the simplest model
- There is however a limit below which a model will not be useful any more
« A model can be falsified [refuted] but never validated »

Karl Popper (1902-1994)

- A model cannot be proven true: at most we can say that it corroborates observed facts
- A model can and should be challenged by decisive experiments
- If is good practice to keep the model as testable as possible
« All models are wrong, some are useful »

Georges Box (1919-2013)

A model is not meant to be true and represent perfectly the reality, but to serve a given purpose.

A model is neither good nor bad:
- a model is just adapted (or not) to the use we have in mind
- efficient (or not) in simulating the real system we want to reproduce
What famous hydrologists have written

- **Klemeš (1986):** no validation, eventually an ‘operational adequation’
  

- **Konikow & Bredehoeft (1992):** models cannot be validated
  

- **de Marsily et al. (1992):** validation is not the point: hydrologists do not aim for certainty and perfection... they just do the best they can
  

- **Oreskes et al. (1994):** models can only be evaluated in relative terms
  

- **A discussion of these papers is available in:**
  

What Keith Beven told me when I asked him

- a good model is one that is fit for purpose

- It is partly because we have many different purposes in hydrology that we have ended up with so many models –
  - and partly because since every PhD graduate can see that it surely be possible to do better, that we have had so many theses proposing new model constructs
A good hydrological model shall:

1. Be based on a sound scientific understanding of the hydrological system.
2. Perform well on a period outside the one used for calibration
3. Be transferable to different physiographic and climatological settings
4. Have a complexity that can be justified by available data and other information about the catchment
5. Have a complexity that is justified by model performance
6. Be understandable by non-specialists in hydrology
What is a ‘good’ hydrological model?

- The perfect model does not exist
- It is more appropriate to speak of models that are ‘better able to describe reality, compared to others’, than to speak of good model.
Let us now summarize…

*(my personal summary in seven propositions)*
Proposition 1

A ‘good’ model is one that gets a good (numerical) grade
A ‘good’ model is one that gets a good (numerical) grade

Many numerical criteria are available…

\[ \alpha = \frac{\sigma_{obs}}{\sigma_{sim}} \quad \beta = \frac{\mu_{obs}}{\mu_{sim}} \quad \rho \text{ – correlation coefficient} \]

\[ KGE = 1 - \sqrt{(1 - \alpha)^2 + (1 - \beta)^2 + (1 - \rho)^2} \]

\[ NSE = 1 - \frac{\sum_{n}(Q_{obs}^i - Q_{cal}^i)^2}{\sum_{n}(Q_{obs}^i - \overline{Q}_{obs})^2} \]

etc…
A grade brings a relative information (an element of comparison)

How to combine different grades which do not necessarily converge?
Choix multicritère

- Multicriteria choices
- Pareto fronts

Taylor diagrams

What is a ‘good’ hydrological model?
Proposition 2

A ‘good’ model is one that is deemed good by experts

What is a ‘good’ hydrological model?

- A: Expert A: mine
- B: Expert B: mine
- C: Expert C: mine
- D: Expert D: mine
What is a ‘good’ hydrological model?

A study made in 2011

Which model provides the better simulation?
- A
- B
- Equivalent

How would you rate the best simulation?
- Very Good
- Good
- Slightly Good
- Average
- Slightly Poor
- Poor
- Very Poor
An experiment where 150 hydrologists were asked to compare
- In relative terms streamflow simulations
- In absolute terms the quality of models

There is no universal conclusion
- Among hydrologists
- Between numerical criteria
- Between numerical criteria and expert judgement
Proposition 3
A ‘good’ model is one that can be used in extrapolation

(robustness issue)
Extrapolation

What is a ‘good’ hydrological model?
What is a ‘good’ hydrological model?

Generalized differential split-sample test to assess the extrapolation capacity of models

« generalized split-sample test » (GSST)

Calibration on sub-periods

<table>
<thead>
<tr>
<th>SP01</th>
<th>SP02</th>
<th>SP03</th>
<th>SP04</th>
<th>SP05</th>
<th>SP06</th>
<th>SP07</th>
<th>SP08</th>
<th>SP09</th>
<th>SP10</th>
<th>SP11</th>
<th>SP12</th>
<th>SP13</th>
<th>SP14</th>
</tr>
</thead>
<tbody>
<tr>
<td>θ₁</td>
<td>θ₂</td>
<td>θ₃</td>
<td>θ₄</td>
<td>θ₅</td>
<td>θ₆</td>
<td>θ₇</td>
<td>θ₈</td>
<td>θ₉</td>
<td>θ₁₀</td>
<td>θ₁₁</td>
<td>θ₁₂</td>
<td>θ₁₃</td>
<td>θ₁₄</td>
</tr>
</tbody>
</table>

‘control’ of the model: transfer of parameters on independent sub-periods

In this example: 18 years of observations, 5-year sub-periods => 90 split-sample tests
Generalized differential split-sample test to assess the extrapolation capacity of models

What we would like to see:
no correlation between model error in validation and temperature anomaly

What is a ‘good’ hydrological model?
Generalized differential split-sample test to assess the extrapolation capacity of models

The kind of result which shows a lack of robustness: anomalies in simulated discharge are correlated to the anomalies in temperature
What is a ‘good’ hydrological model?

This concept should be extended to ‘changing’ catchments.

Proposition 4
A ‘good’ model is one that can work in degraded mode

Robustness again
A ‘good’ model is one that can work in degraded mode
A ‘good’ model is one that can work in degraded mode.

Proposition 5

A ‘good’ model is one whose limits are known
A ‘good’ model is one whose limits are known

- Finding the limits requires testing the model beyond them

- Knowing the limits requires testing the model on a large data set
After all, what makes modern cars so safe?

‘Crash testing’ models is a necessity and the publication of these crash tests should be encouraged.

Proposition 6

A ‘good’ model is one which is numerically sound
A ‘good’ model is one which is numerically sound

Most of the hydrological models require calibration (at least partially): sound numerical behavior is a prerequisite for smooth multidimensional response surface and efficient calibration

- Modellers should avoid thresholds
- But integrated schemes are not as bad as publicized in the literature
Proposition 7
A ‘good’ model is one which contains the right equations
A ‘good’ model is one which contains the right equations

- This proposition is debated among hydrologists (I do not subscribe to it unconditionally)

  - « right » equations, « physically-based » equations

  - The « physical purity » of a model is an utopia: we do not know the ‘true’ catchment-scale equations, and upscaling physical properties from the lab-scale to the catchment (or computation unit) scale is a matter of faith

- For equivalent efficiency, one can (should?) favour formulations which have appropriate physical justifications
Conclusion
In lieu of conclusion
Alternative models for the Ju-52

Which model is the most faithful?

Which one flies?

Adapted from Sten Bergström, SMHI