IAHS AISH ICSW STAHY ICWRS

# HW06 workshop
# Expert judgement versus statistical goodness-of-fit for hydrological model evaluation

# Introduction


**Convenors**

| | | |
|---|---|---|
| **Charles Perrin** | Cemagref | **France** |
| **Mark Thyer** | The University of Adelaide Australia | **Australia** |
| **Salvatore Grimaldi** | Università della Tuscia | **Italy** |
| **Hoshin Gupta** | The University of Arizona | **USA** |
| **Jean-Emmanuel Paturel** | IRD | **France** |


*On the original idea of Vazken Andréassian*


XXV IUGG General Assembly,
3-4 July 2011, Melbourne

# Context

- **Assessment of model outputs quality useful at various steps of the modelling process:**
  - **calibration**
  - **validation**
  - **a posteriori evaluation**

- **Depends on:**
  - **modelling objectives (simulation, forecasting, design, etc.)**
  - **experience of user, prior knowledge or availability of soft data**
  - **testing procedure**
  - **data at hand**
  - **criteria used (graphical or numerical) and how they are interpreted**

➢ **Some degree of subjectivity involved**

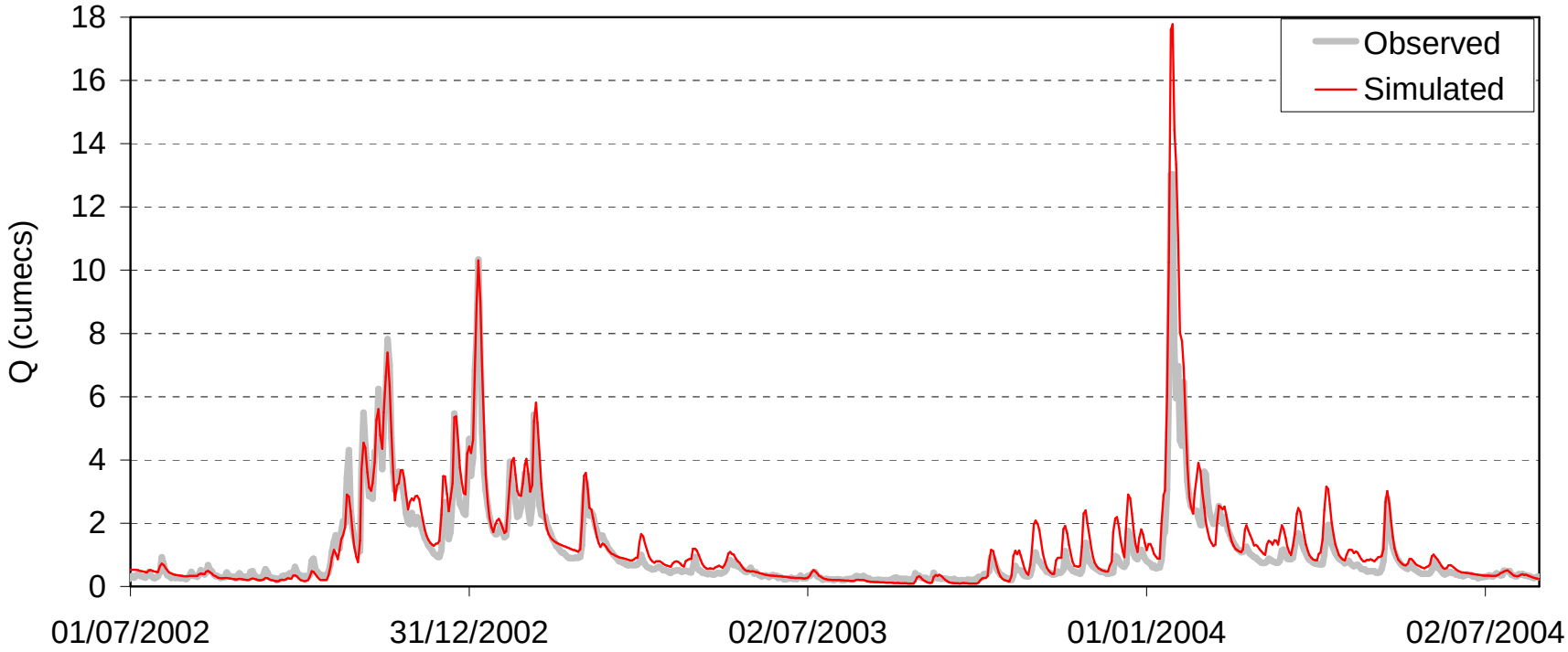
➢ **No consensus on how to evaluate models**

**Should we evaluate evaluation criteria?**



Cemagref
Sciences, eaux & territoires

# Vizualizing hydrographs



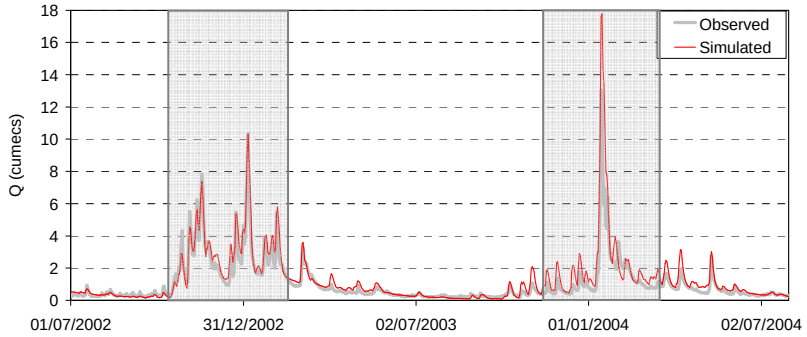

**Is my model good/average/poor given my simulation objectives?**

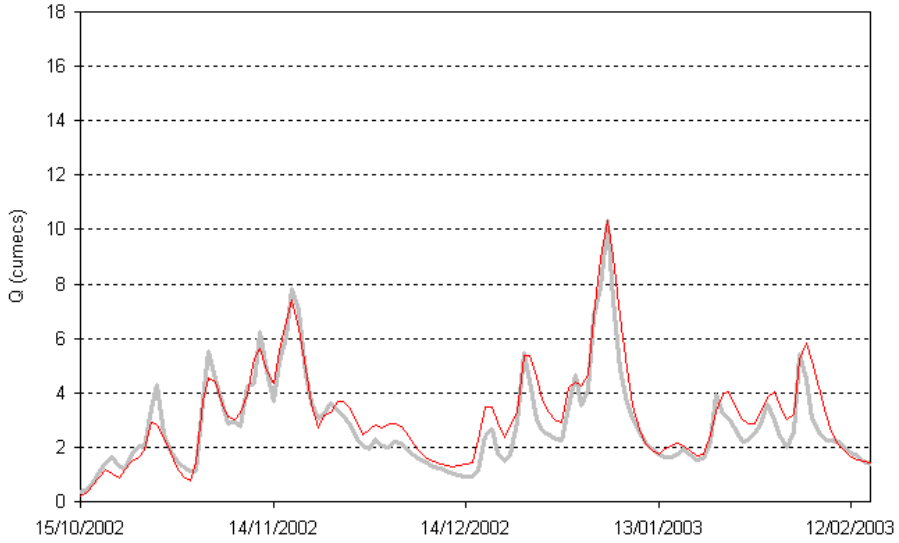


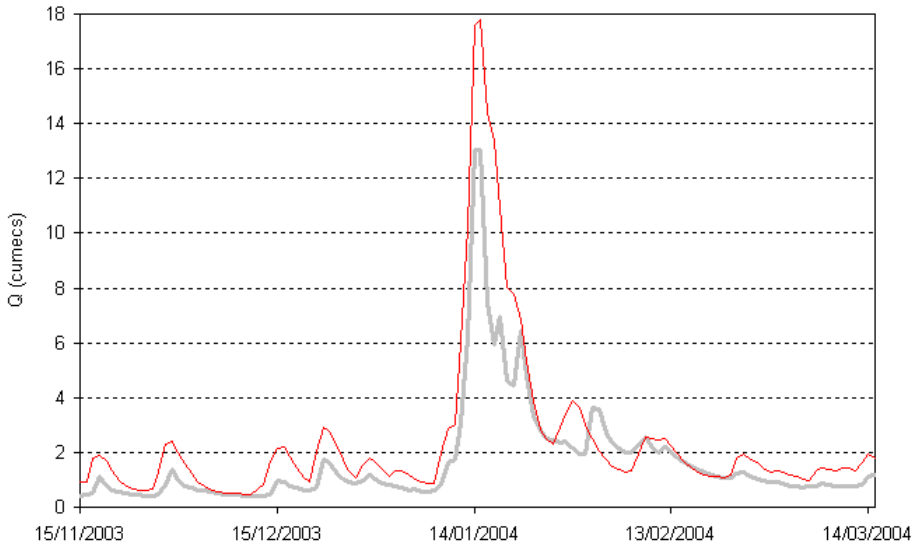
Cemagref
Sciences, eaux & territoires

# Vizualizing hydrographs

- **Focus on high flow periods**

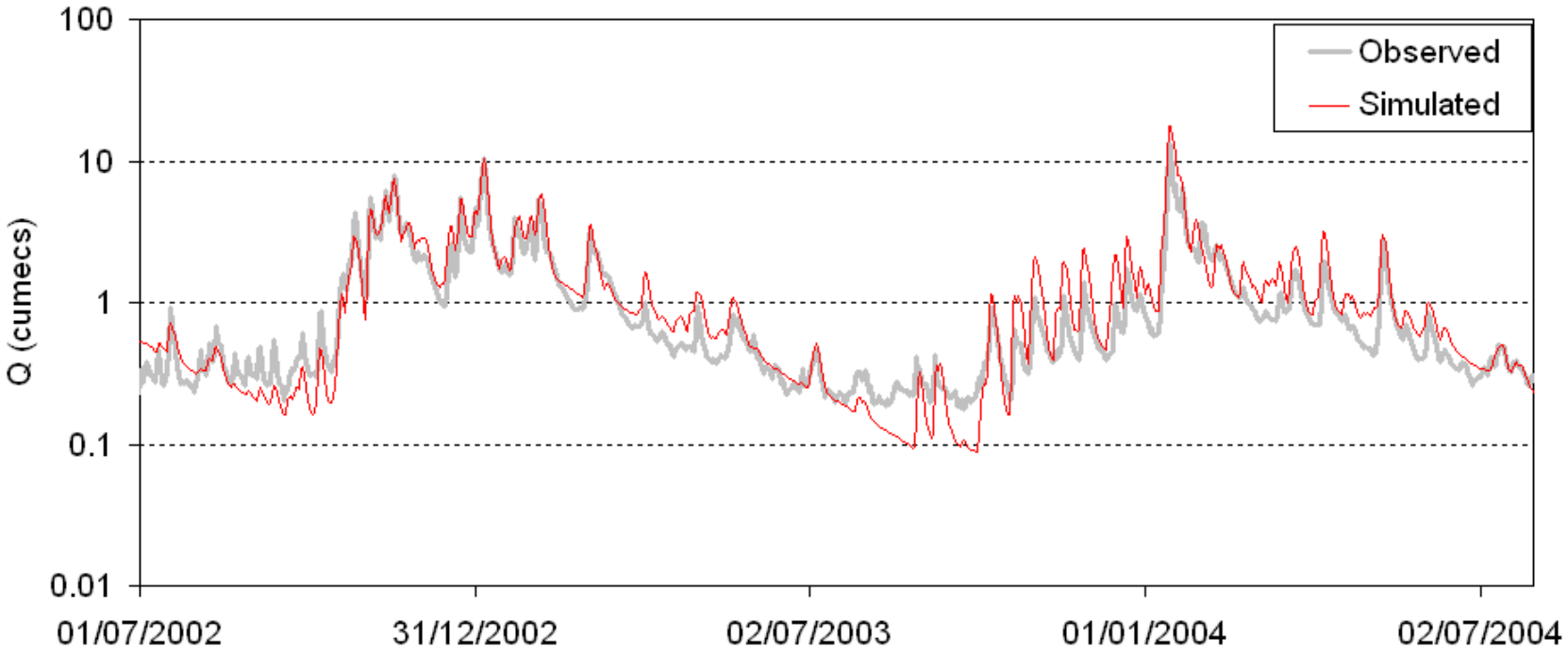


# Vizualizing hydrographs

- **Focus on low flow periods**



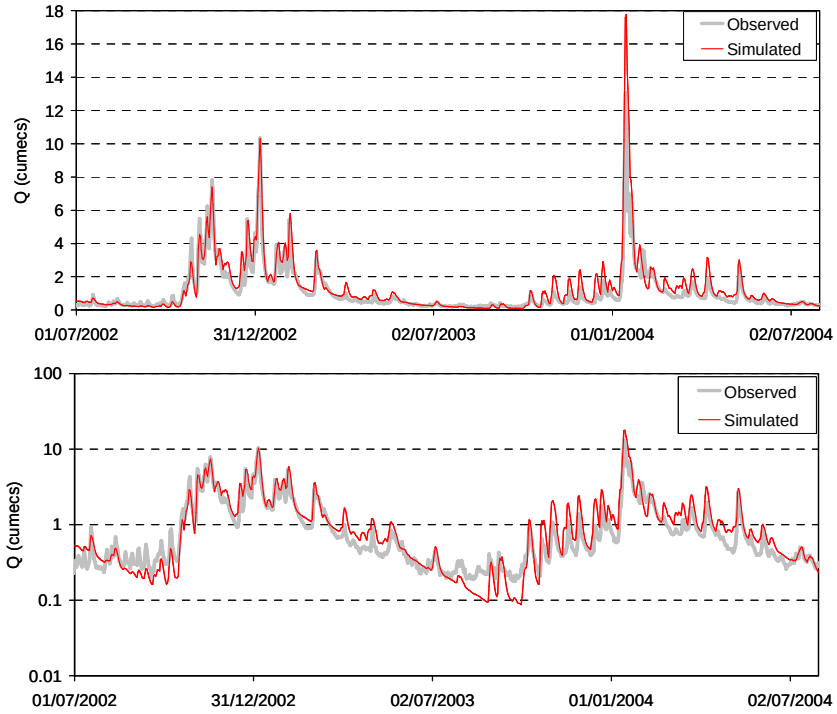



Cemagref
Sciences, eaux & territoires

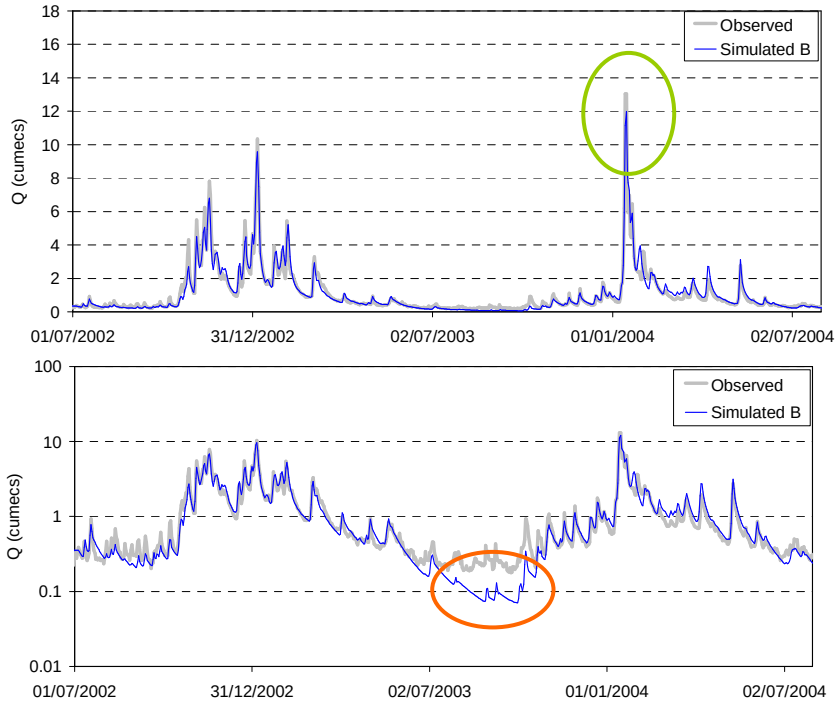
# Vizualizing hydrographs

- Is my model good, average or bad?
- Beyond the difficulty to answer this question:
  - Can my judgement be considered reliable?
  - Would my judgement be different if another model was also tested?

**Model A**




**Model B**




  - Would someone else provide a different answer?

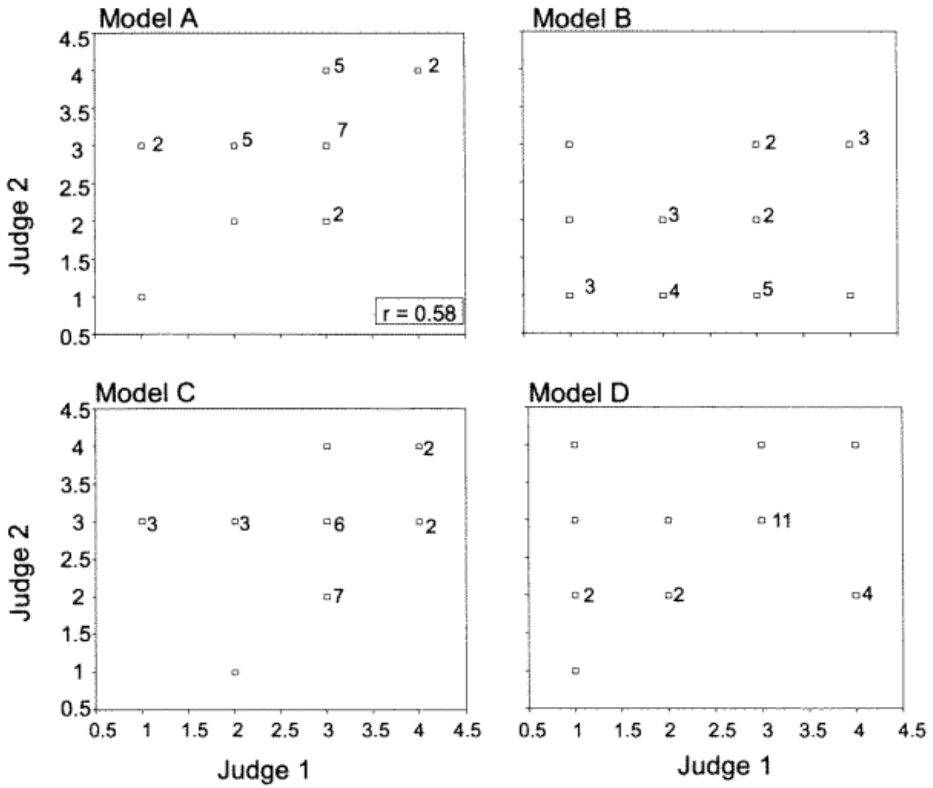
# Case study

## Comparison of expert judgement by Houghton-Carr (HSJ, 1999)

- Two experts asked to evaluate daily hydrographs simulated by four models (A to D)
- Marks from from 1 to 4
- No prior guidance
- Set of 25 catchments

➢ Considerable spread in model evaluation

➢ Judgement differed significanctly if FDC was shown instead of daily hydrographs




(Fig. 6 in Houghton-Carr, 1999)

# Quantitative criteria as an alternative to qualitative criteria?

- **Seen as more objective tools than graphical criteria**
- **But…**
  - **a large number of criteria exist (continuous, event-based, etc.)**
  - **quantitative criteria can also disagree**

| | | | |
|---|---|---|---|
| Criterion 1 | Model A | > | Model B |
| Criterion 2 | Model A | < | Model B |

  - **behaviour of criteria not always fully understood (still ongoing discussions on the 40-year old Nash-Sutcliffe efficiency index)**



Cemagref
Sciences, eaux & territoires

# Qualitative and quantitative criteria: an easy or difficult wedding?

- Largely used as complementary tools for the evaluation of model outputs
- Are they actually complementary?
- To which extent do they match?



Cemagref
Sciences, eaux & territoires

# How can model evaluation be made more systematic?

- Are there too many criteria?
- Can we define a set of reference criteria?
- Should we prefer absolute or relative evaluation?
- Do we need reference models (“benchmarks”)?
- Should hydrologists be better trained for evaluating the quality of models’ outputs?



Cemagref
Sciences, eaux & territoires

# Objectives of the workshop

- **To better understand the links between expert judgement and numerical criteria, based on a survey and an inventory of numerical criteria**

- **To better understand the role of criteria in**
  - **model calibration**
  - **model evaluation procedures for various objectives (simulation, forecasting)**
  - **model comparisons**
  - **diagnosis of model failures**
  - **estimation of uncertainties**

- **To give time to exchanges and share experience on the problems associated with the evaluation of hydrological models**



Cemagref
Sciences, eaux & territoires

# Programme of the workshop

## This afternoon

| | | | | | |
|---|---|---|---|---|---|
| HW06S1 | **13:30-15:00** | **Session HW06S1<br>Comparing expert judgement and numerical criteria: experiment design<br>*Chairman : Charles Perrin*** | | | |
| | 13:30-13:45 | | Introduction: Expert judgement and criteria in hydrological modelling | | |
| | 13:45-14:00 | | Interactive survey on expert judgement | | |
| | 14:00-14:15 | | | | |
| | 14:15-14:30 | 2237 | Man vs. Machine: a Swedish Experiment on Hydrological Model Performance Assessment | Dr Jonas Olsson | SMHI |
| | 14:30-14:45 | | | | |
| | 14:45-15:00 | | **Discussion** | | |
| | 15:00-16:30 | *Coffee break and poster session* | | | |
| HW06S2 | **16:30-18:00** | **Session HW06S2<br>Do numerical criteria tell what we expect?<br>*Chairman: Salvatore Grimaldi*** | | | |
| | 16:30-16:45 | 1461 | Automated Detection and Categorization of Hydrograph Errors for Updating Operational Hydrologic Models | Mr Ashenafi Gragne | Norwegian University Of Science And Technology |
| | 16:45-17:00 | 4041 | Use of various types of evaluation criteria at different stages of model implementation | Dr Olga Semenova | State Hydrological Institute |
| | 17:00-17:15 | 4092 | On the significance of the performance criteria of event-based flood models | Mr Roger Moussa | Institut National De La Recherche Agronomique |
| | 17:15-17:30 | 3370 | Verification Methods for Probabilistic Streamflow Forecasts | Dr David Robertson | CSIRO |
| | 17:30-17:45 | 531 | On the Criteria of Model Performance Evaluation for Hydrological Forecasting Models | Prof Ke-Sheng Cheng | National Taiwan University |
| | 17:45-18:00 | | **Discussion** | | |

# Programme of the workshop

## Tomorrow morning

| | | | | | |
|---|---|---|---|---|---|
| HW06S3 | **08:30-10:00** | **Session HW06S3<br>Can we build better models? Criteria for calibrating and diagnosing models<br>*Chairman: Mark Thyer*** | | | |
| | 08:30-08:45 | 5901 | The effect of choice of calibration Criteria on Hydrological Model Performance | Pr Zoubeida Bargaoui | Laboratoire Modelisation en Hydraulique et Env., Tunis |
| | 08:45-09:00 | 5419 | Using entropy as a descriptor of flow duration for rainfall-runoff model calibration | Dr Ilias Pechlivanidis | Victoria University of Wellington |
| | 09:00-09:15 | 2883 | The value of considering uncertainty in design of model performance indicators | Dr Barry Croke | Australian National University |
| | 09:15-09:30 | 3357 | Use of Multiple Data Types to Evaluate and Diagnose Weaknesses in Hydrological Models | Dr Hilary Mcmillan | National Institute of Water and Atmospheric Research |
| | 09:30-09:45 | 3578 | Impact of modeller's attitude in catchment modelling: stage from priori prediction to calibration | Dr Hartmut Hollander | State Authority of Mining, Energy and Geology |
| | 09:45-10:00 | | **Discussion** | | |
| | 10:00-10:30 | *Coffee break* | | | |
| HW06S4 | 10:30-12:00 | **Session HW06S4<br>Reconciling expert judgement and numerical evaluation. Results and perspectives in hydrological modelling<br>*Chairman : Vazken Andréassian*** | | | |
| | 10:30-10:45 | 524 | Performance of a conceptual model in a complex river system and with scarce data: Case of Letaba River | Mr Zacharia Katambara | University of The Witwatersrand |
| | 10:45-11:00 | 3315 | The Effectiveness of Hydrological Models Sreamflow Prediction (...) | Mr Erwin Jeremiah | University Oo New South Wales |
| | 11:00-11:15 | 5550 | Expert judgement versus statistical goodness-of-fit for hydrological model evaluation: Results of experiment | Mr Charles Perrin | Cemagref |
| | 11:15-11:30 | | | | |
| | 11:30-11:45 | | **Discussion and perspectives** | | |
| | 11:45-12:00 | | | | |

Posters

# Important reminders

- **IAHS Banquet on Wednesday**
  **Tickets can be purchased at the IAHS counter on ground floor**

- **Red Books**
  **For those who ordered proceedings at registration, collect your issues at the IAHS counter**

- **IAHS Frisbee Contest**
  **Join one of the teams:**
  **“Deterministic modellers” vs. “Probabilistic modellers”**

# With wishes for a lively and fruitful workshop!



Cemagref
Sciences, eaux & territoires