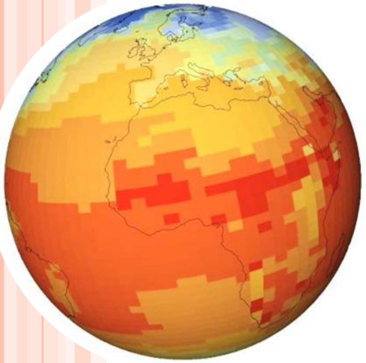


# ISIMIP2a Global Water Sector

Simon Gosling (sector coordinator), Rutger Dankers (sector coordinator), jacob Schewe, Jamal Zaherpour, Nick Mount, Sebastian Ostberg, Naota Hanasaki, Catherine Morfopoulos, Tian Zhou, Yoshimitsu Masaki, Stephanie Eisner, Hannes Müller Schmied, Tobias Stacke, Yoshihide Wada, Yusuke Satoh, Yadu Pokhrel, Qiuhong Tang, Guoyong Leng, Rene Orth Lamprini Papadimitriou, Ted Veldkamp, Jinfeng Chang and Philippe Ciais

# Outline

- Summary of modelling progress for ISI-MIP2.1A
- Evaluation of global water models
- Weighting of models within an ensemble
- Dual-scale ensembles of climate impacts from 1, 2 and 3$^{o}$C global warming

# Summary of progress for ISIMIP2.1A

# 12 Models Participating

| Model | Institution |
|---|---|
| CLM | Pacific Northwest National Laboratory (USA) |
| DBH | IGSNRR, Chinese Academy of Sciences (China) |
| H08 | Institute for Environmental Studies (Japan) |
| JULES | University of Exeter, CEH and Met Office (UK) |
| LPJmL | PIK (Germany) |
| MATSIRO | IIS, University of Tokyo (Japan) |
| MPI-HM | Max-Planck-Institute for Meteorology (Germany) |
| PCR-GLOBWB | University Utrecht (Netherlands) |
| VIC | University of Washington (lead institution) with Norwegian Water Resources and Energy Directorate (assisting) |
| WaterGAP2 | Universität Kassel and Universität Frankfurt (Germany) |
| SWBM | CUNY (USA) |
| ORCHIDEE | Laboratoire d'Océanographie et du Climat, IPSL (France) |

ISIMIP
Inter-Sectoral Impact Model Intercomparison Project

# The Scenarios

- All simulations are for present-day under **three main scenarios**:

| | nosoc | pressoc | varsoc |
|---|---|---|---|
| **Scenario:** | **Naturalized runs** (no human impact):<br>• Climate varies<br>• No population<br>• No GDP<br>• No irrigation & dams<br>• Time varying land-cover | **Human impact runs**:<br>• Climate varies.<br>• **Constant year 2000** values for:<br>o Population<br>o GDP<br>o Land-cover<br>o Irrigation & dams | **Human impact runs:**<br>• Climate varies<br>• **Time-varying** values for:<br>o Population<br>o GDP<br>o Land-cover<br>o Irrigation & dams |
| **Related research questions:** | **Compared with varsoc to demonstrate human impacts on water resources.** | Used to quantify **adaptation pressure** under current socio-economic conditions. | **Used for model evaluation.** |

**What value can multi-model ensembles provide?**

- Also running with different climate input datasets to explore sensitivity to inputs:
  - PGMFD v.2 (Princeton)
  - GSWP3
  - WATCH (WFD)
  - WATCH+WFDEI.GPCC

# Progress

| | nosoc | | | | pressoc | | | | varsoc | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | gswp3 | princeton | watch | wfdei | gswp3 | princeton | watch | wfdei | gswp3 | princeton | watch | wfdei |
| H08 | | | | | | | | | | | | |
| LPJmL | | | | | | | | | | | | |
| PCR-GLOBWB | | | | | | | | | | | | |
| WaterGAP2 | | | | | | | | | | | | |
| MPI-HM | | | | | | | | | | | | |
| MATSIRO | | | | | | | | | | | | |
| DBH | | | | | | | | | | | | |
| JULES-TUC | | | | | | | | | | | | |
| JULES-UoE | | | | | | | | | | | | |
| VIC | | | | | | | | | | | | |
| CLM | | | | | | | | | | | | |
| ORCHIDEE | | | | | | | | | | | | |
| SWBM | | | | | | | | | | | | |

# PAPERS ADDRESSING THE RESEARCH QUESTIONS

## Model evaluation

- Xingcai Liu et al. (Chinese Academy of Sciences, China) "***Validation of multimodel simulations of river flow in China***"
  - Using observed records outside of the GRDC database.

- Zaherpour, Gosling, Mount et al. (University of Nottingham, UK) "***Multi-model evaluation of river runoff simulations from global water models***"
  - The first consistent evaluation of several global water models.

# PAPERS ADDRESSING THE RESEARCH QUESTIONS

**The effect of water management on global water resources**

- Xingcai Liu et al. (Chinese Academy of Sciences, China) "***How does human impact influence the uncertainty in multi-model simulations of river flows***?"
  - An evaluation of uncertainties introduced by parametrisations of human impact processes (dams, irrigation etc.).

- Ted Veldkamp et al. (Vrije Universiteit, Netherlands) "***Human impacts on water scarcity in the 20th century: a multi-model and multi-forcing analysis***"
  - How humans have compounded water scarcity threats.

- Yoshimitsu Masaki et al. (National Institute for Environmental Studies, Japan) "***Intercomparison of regulated river discharge among multiple hydrological models under multiple forcings — (II) Multiple models***".
  - The first multi-model inter-comparison study on river flow regulated by dams.

# PAPERS ADDRESSING THE RESEARCH QUESTIONS

**Enhanced value from multi-model analysis**

- Zaherpour, Gosling, Mount et al. (University of Nottingham, UK) "***Weighting global-scale climate change impact models using machine learning algorithms***"
  - An ensemble summary metric that performs better than the ensemble mean over the historical period.

- Fred Hattermann et al. (PIK, Germany) "***Cross-scale inter-comparison of climate change impacts simulated by regional and global hydrological models in eleven large river basins***"
  - Comparative evaluation of model performance of global- and catchment-scale hydrological models and projections under climate change.

- Simon Gosling et al. (University of Nottingham, UK) "***A comparison of changes in river runoff from multiple global and catchment-scale hydrological models under global warming scenarios of 1°C, 2°C and 3°C***"
  - Comparative impacts assessment of global warming effects on runoff across global- and catchment-scale hydrological models.

# Evaluation of global water models

# Approach

- 5 (soon to be 6) global water models evaluated over 40 catchments.
  - All catchments are >100,000 km$^2$ and have >25 years of observed GRDC data.
  - Located in different hydroregions.

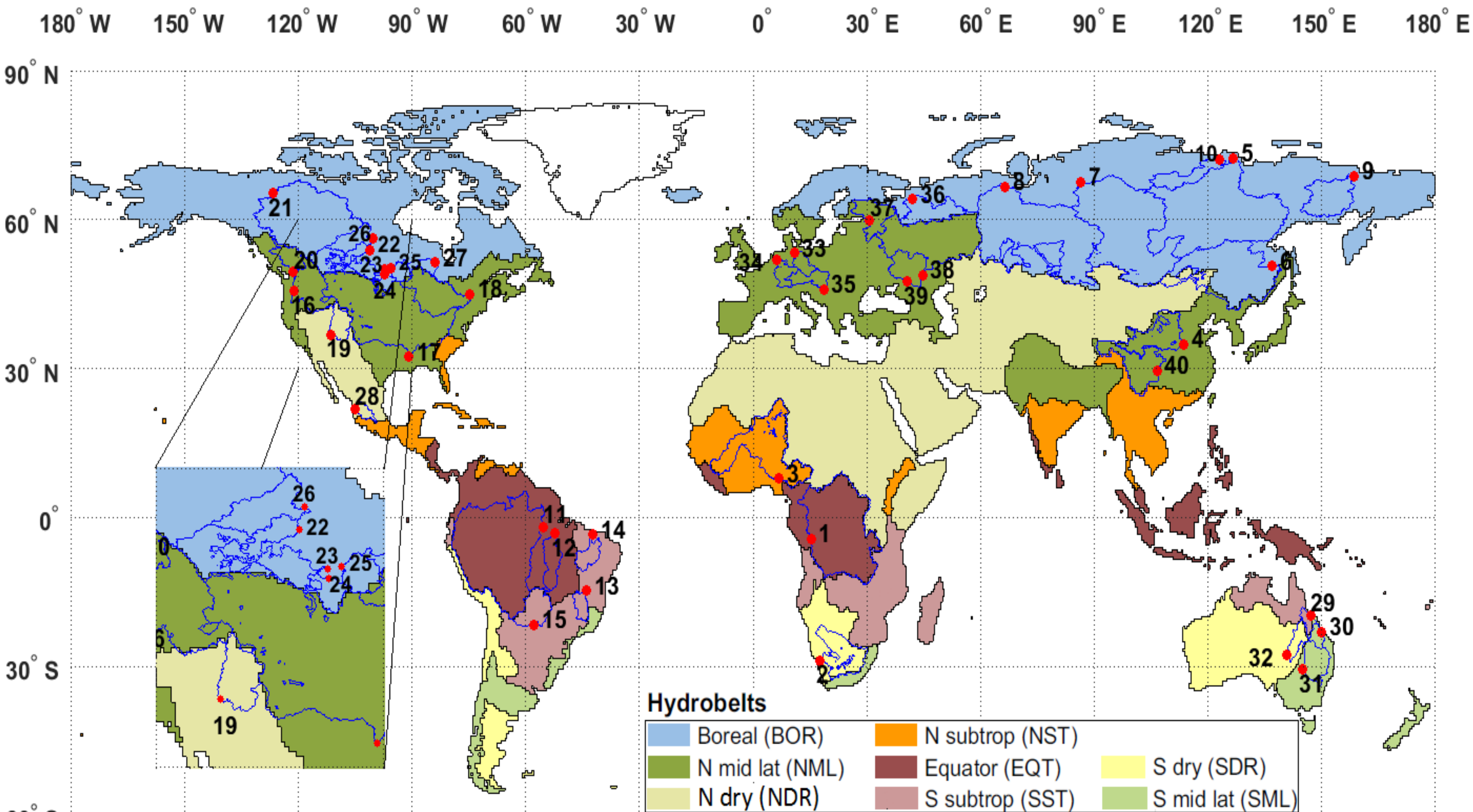
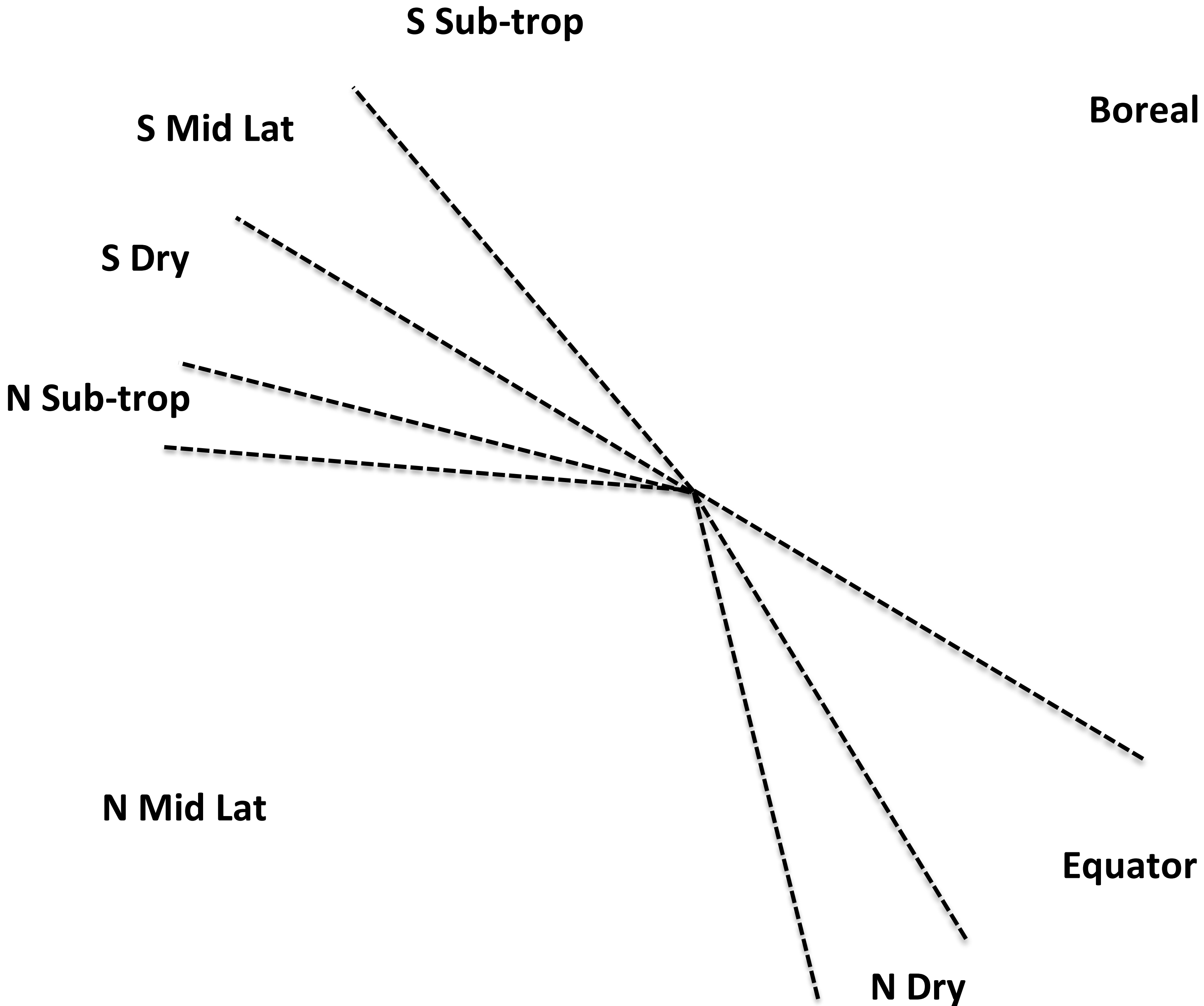



Figure source: Jamal Zaherpour, University of Nottingham

# Results


S Sub-trop
Boreal
S Mid Lat
S Dry
N Sub-trop
N Mid Lat
Equator
N Dry

# Results

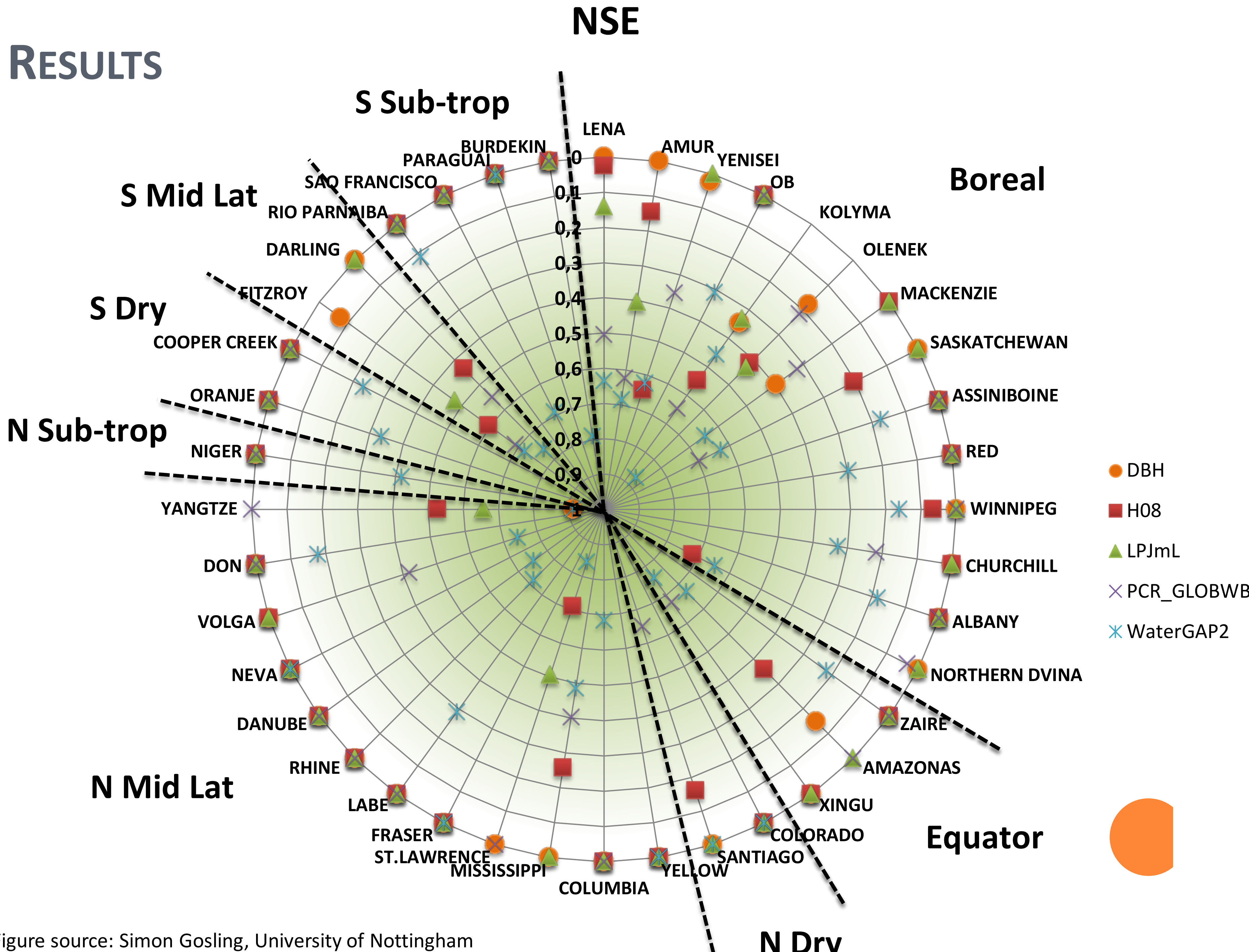



Figure source: Simon Gosling, University of Nottingham

# Results

|  | DBH | H08 | LPJmL | PCR_GLOBWB | WaterGAP2 | EM |
|---|---|---|---|---|---|---|
|  | % | % | % | % | % | % |
| NSE >= 0.90 | 3 | 0 | 0 | 0 | 3 | 0 |
| NSE >= 0.80 | 3 | 0 | 0 | 0 | 8 | 3 |
| NSE >= 0.70 | 3 | 5 | 0 | 0 | 25 | 10 |
| NSE >= 0.60 | 3 | 8 | 3 | 15 | 48 | 23 |
| NSE >= 0.50 | 3 | 15 | 5 | 20 | 48 | 28 |

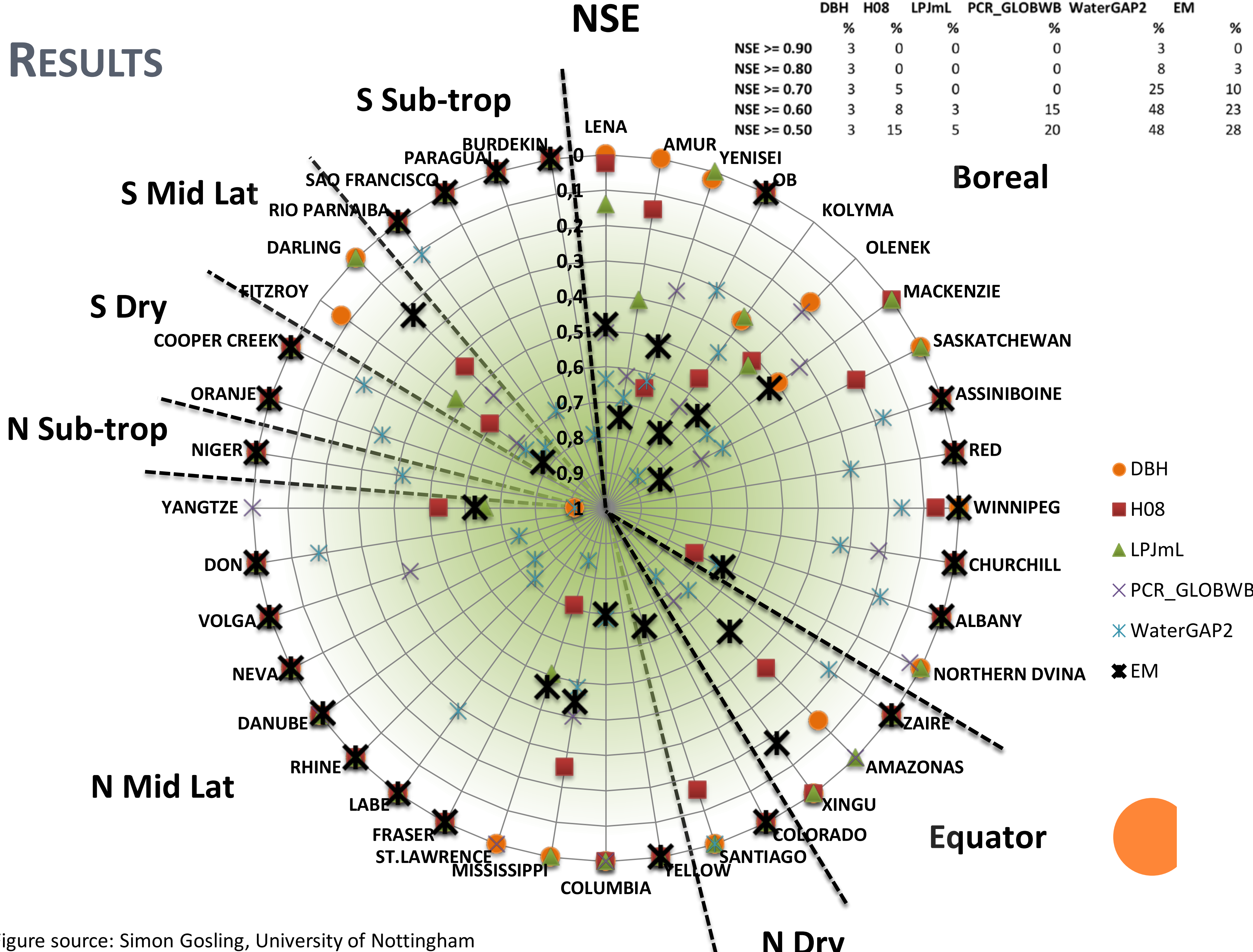
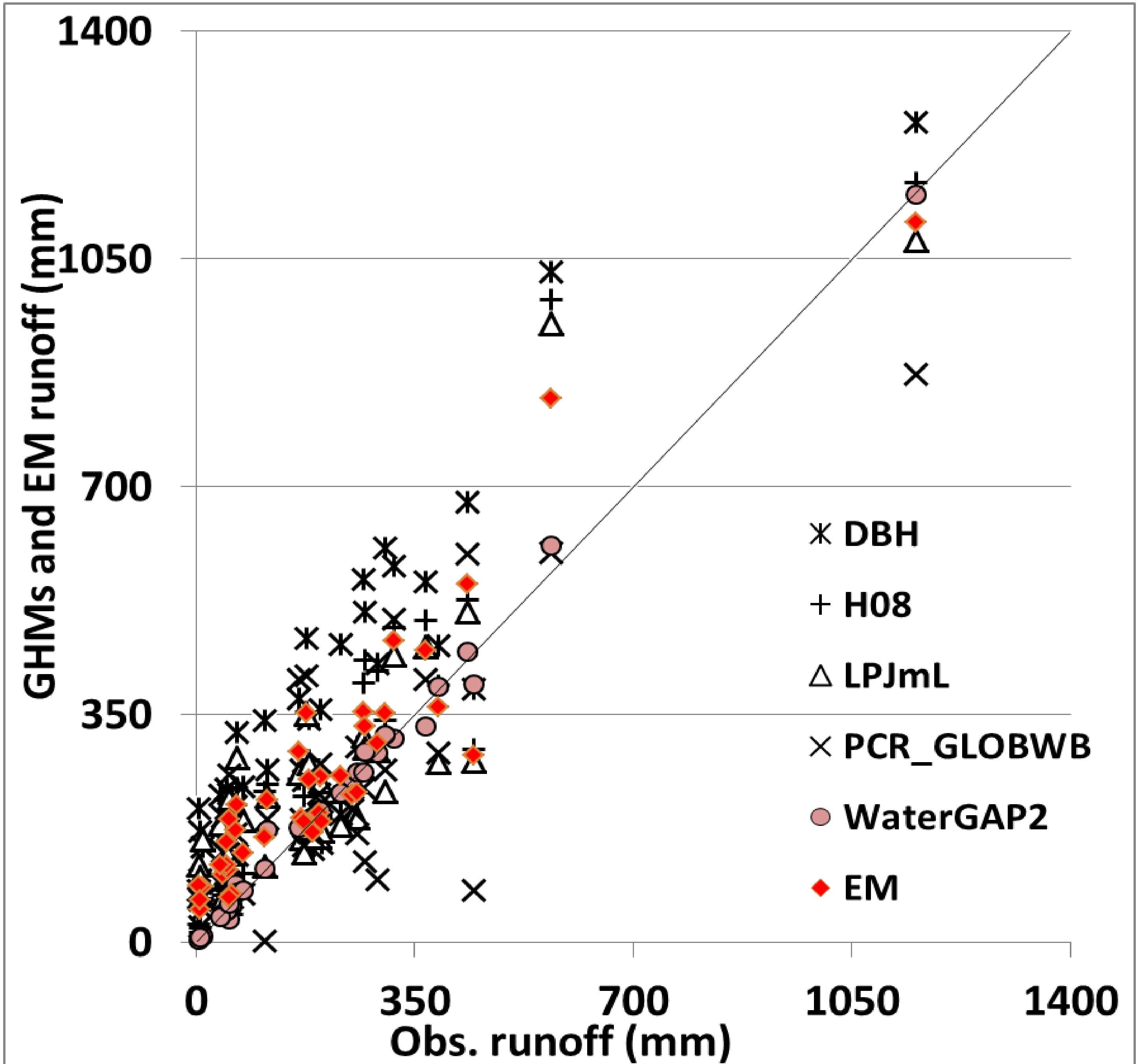



Figure source: Simon Gosling, University of Nottingham

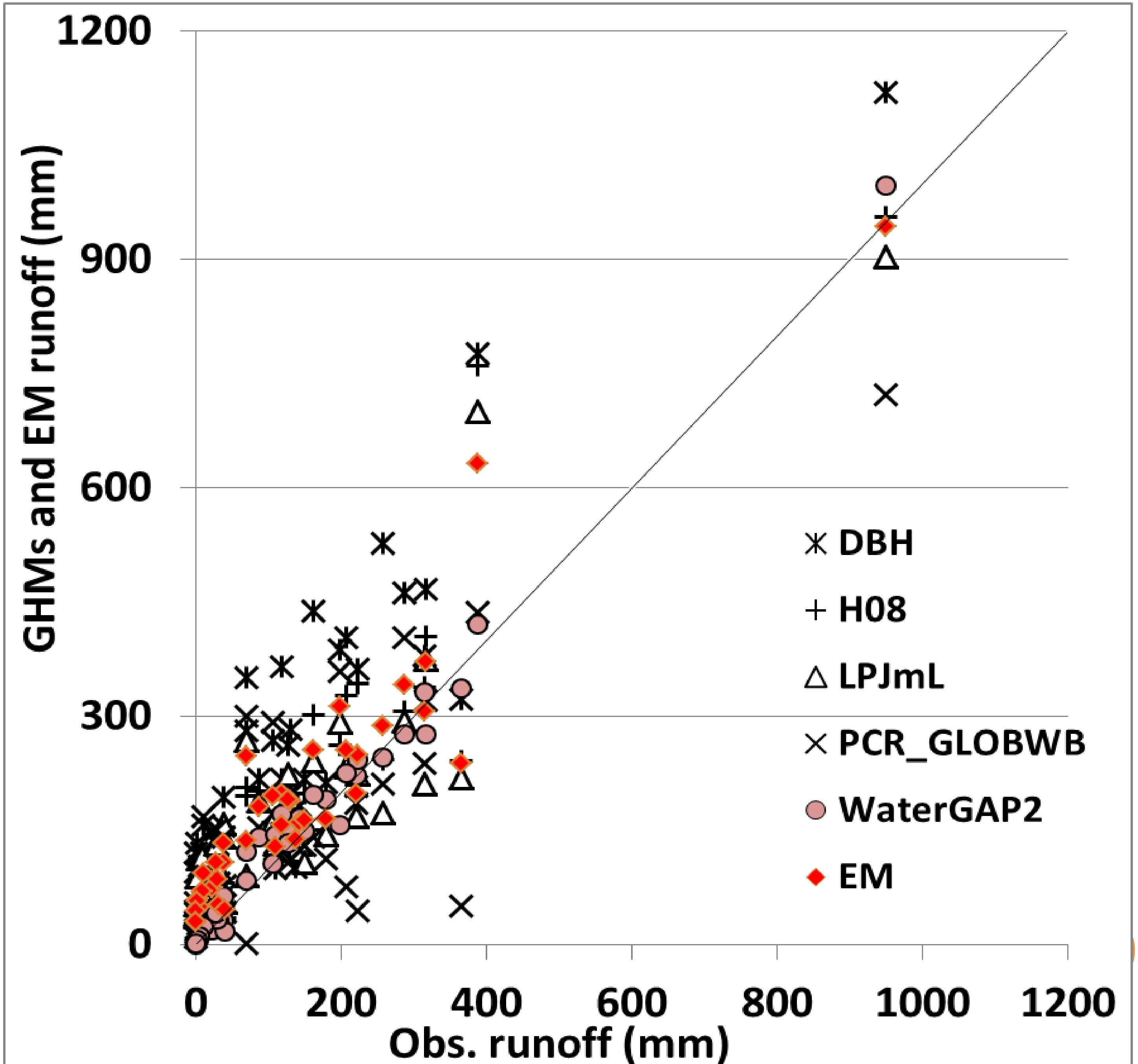
# Results

- Mean annual runoff:



# Results

- Q95:



1200
900
600
300
0
GHMs and EM runoff (mm)
0
200
400
600
800
1000
1200
Obs. runoff (mm)
DBH
H08
LPJmL
PCR_GLOBWB
WaterGAP2
EM

# Weighting models within an ensemble

# THE CHALLENGE

**Global Hydrological Model (GHM) simulations for Yangtze catchment at Cuntan (sample from present-day)**

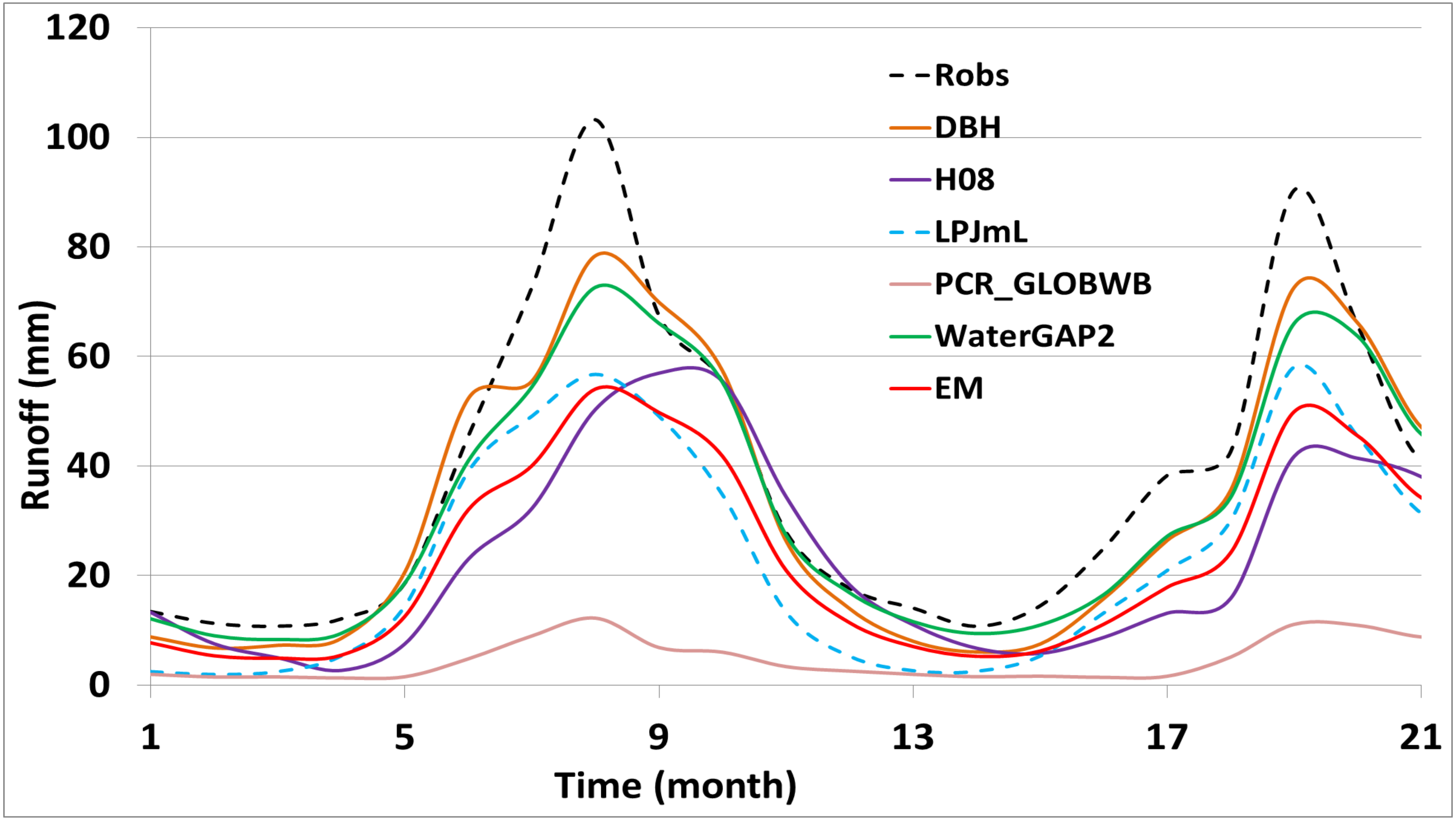



Figure source: Jamal Zaherpour, University of Nottingham

# The Research Question

Can we develop an ensemble summary indicator that is **informed by model performance and model weighting** that is a more **robust** indicator than treating all models equally (e.g. ensemble mean)?

# Our Approach

- Conducted a **proof of concept** study to demonstrate potential for **use of the method by other sectors**.
- We computed **multi-model combination (MMC)** estimates of runoff for the historical period using varsoc simulations from 5 Global Hydrological Models (GHMs) run in the ISI-MIP2A project.
- These MMCs are **an alternative to the ensemble mean**.
- We used **machine learning algorithms** (MLAs) to create the MMCs.
- MLAs enable the development of statistical data-driven models by iteratively defining programs/functions that represent the input data.
  - E.g. using symbolic regression:

**MMC = cos(M1) + 3*M2 + 0.5**

Dependent variable (e.g. runoff) — Independent variables (e.g. impact models)

- **The independent variables (impact models) are given <u>different weights</u> according to their performance.**

# Our Approach

- We evaluated individual impact model fitness, for 40 catchments individually, using the Ideal Point Error (IPE) statistic:

$$IPE = \left\{0.333 * \left[\left(\frac{\mathrm{RMSE_i}}{\max(\mathrm{RMSE})}\right)^2 + \left(\frac{\mathrm{MARE_i}}{\max(\mathrm{MARE})}\right)^2 + \left(\frac{\mathrm{CE_i} - 1}{\min(\mathrm{CE}) - 1}\right)^2\right]\right\}^{\frac{1}{2}}$$

**RMSE**: Root mean squared error

**MARE**: Mean absolute relative error

**CE**: Coefficient of efficiency (NSE)

All metrics are weighted equally

***i***: *ith* participating model

- IPE = 0 if model fit to observations is perfect.

# Results

- The MLAs result in a set of **candidate functions/programs** of different sizes (complexity) and fit.
- They are developed from calibration data and tested on separate data.
- Individual impact models are **weighted according to their performance**.
- Typical results, **for one catchment** look like:

GHMs

| Size | Solution |
|---|---|
| 3 | $Robs = pcrglobwb - 0.38$ |
| 5 | $Robs = pcrglobwb + 0.13\,h_{08}$ |
| 7 | $Robs = 0.61\,pcrglobwb + 0.38\,vic$ |
| 9 | $Robs = vic + 0.32\,pcrglobwb - 0.47\,watergap$ |
| 11 | $Robs = 0.96\,vic + 0.36\,pcrglobwb - 0.45\,watergap$ |
| 13 | $Robs = 1.8 + vic + 0.013\,pcrglobwb^2 - 0.5\,watergap$ |
| 15 | $Robs = 3.2 + 2.6\,wbm - 1\,watergap - 0.016\,h_{08}\,wbm$ |
| 19 | $Robs = 1.8 + 2.4\,wbm + 0.19\,pcrglobwb - 0.87\,watergap$ |

# Results

Mean annual runoff (MAR) performance over 40 catchments

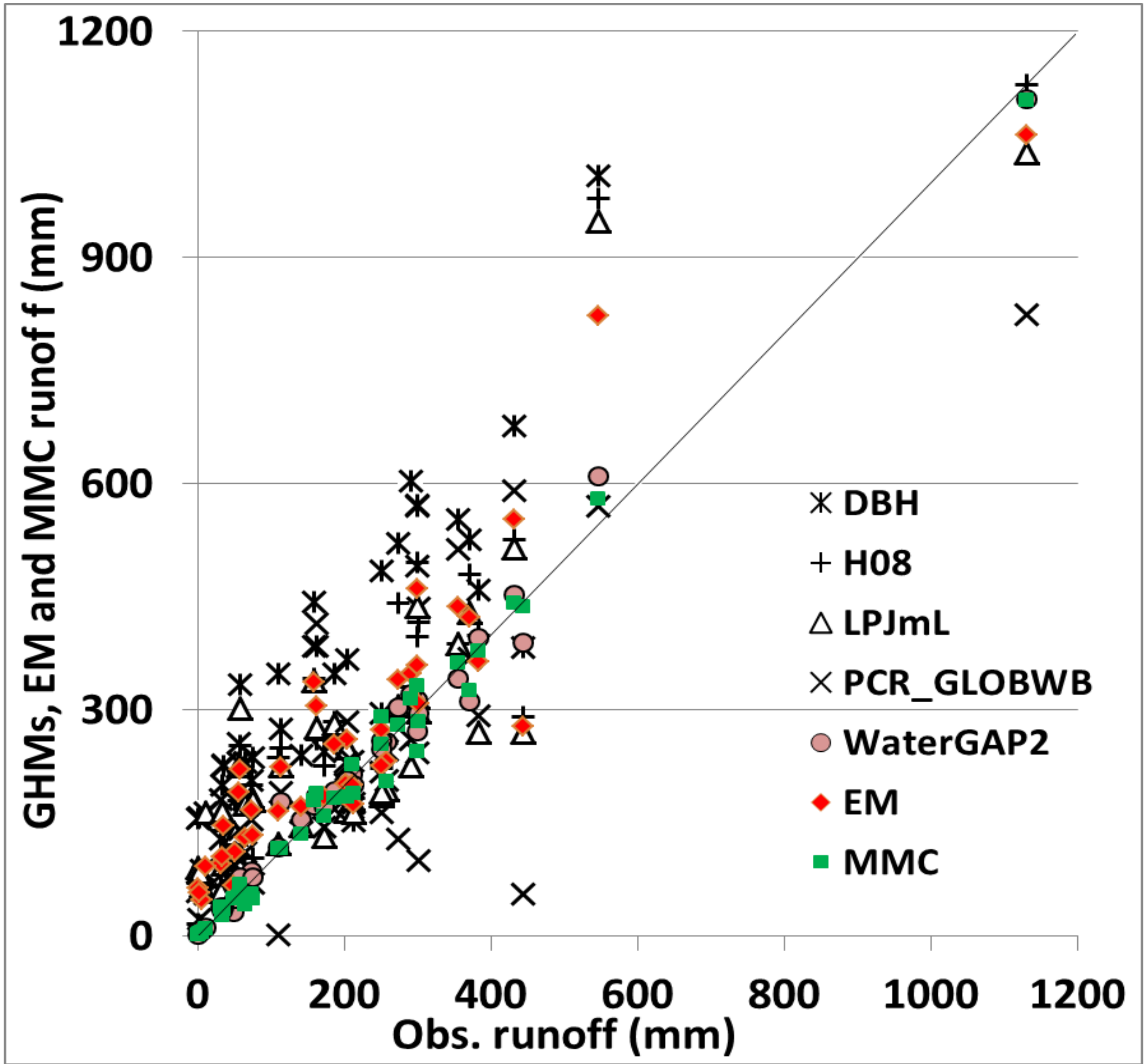



Figure source: Jamal Zaherpour, University of Nottingham

# Results

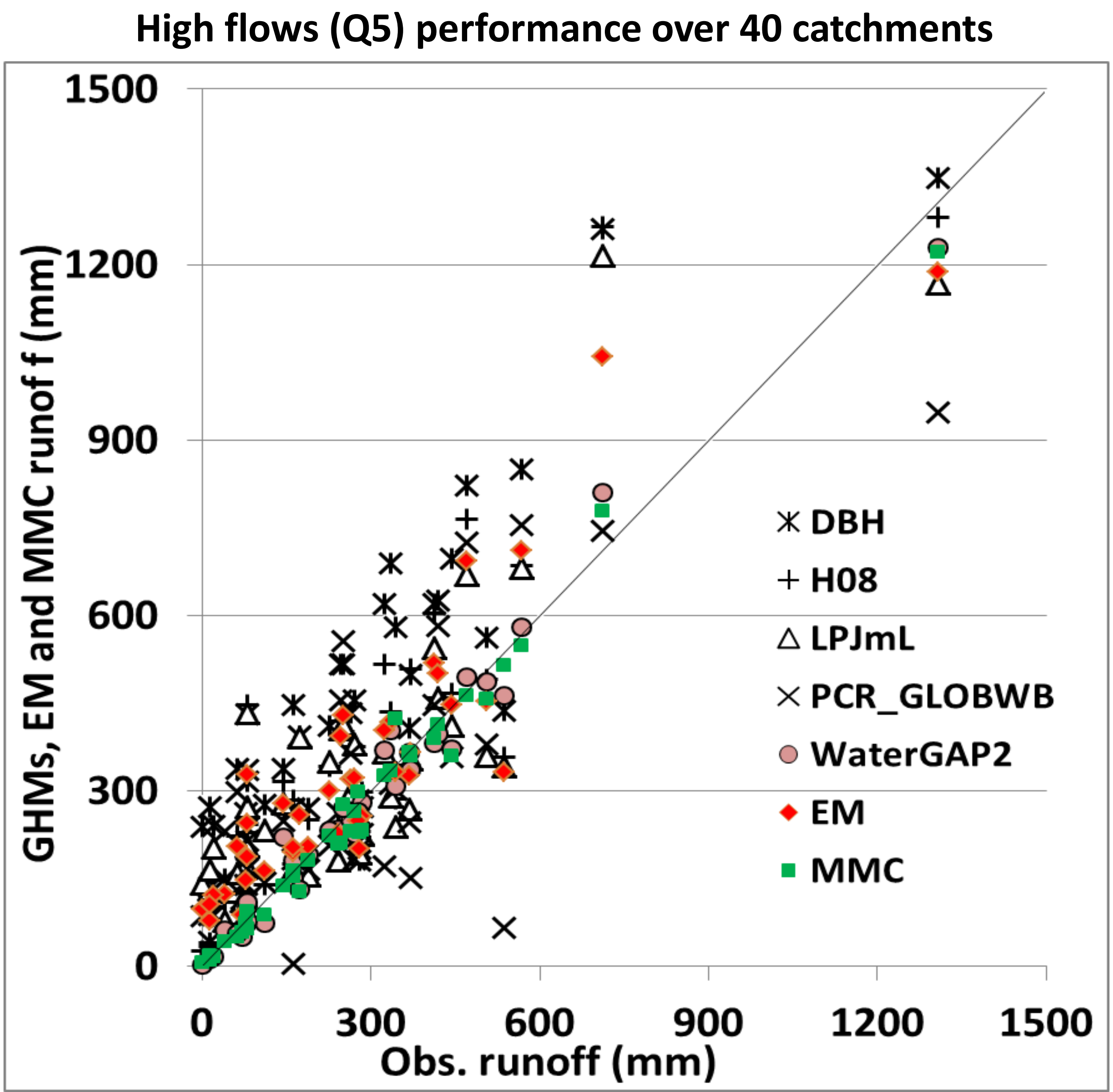



Figure source: Jamal Zaherpour, University of Nottingham

# RESULTS

- MMC performs better than EM and individual models in most catchments:

**"Best performers" of mean annual runoff for each catchment (GHMs, EM and MMC)**

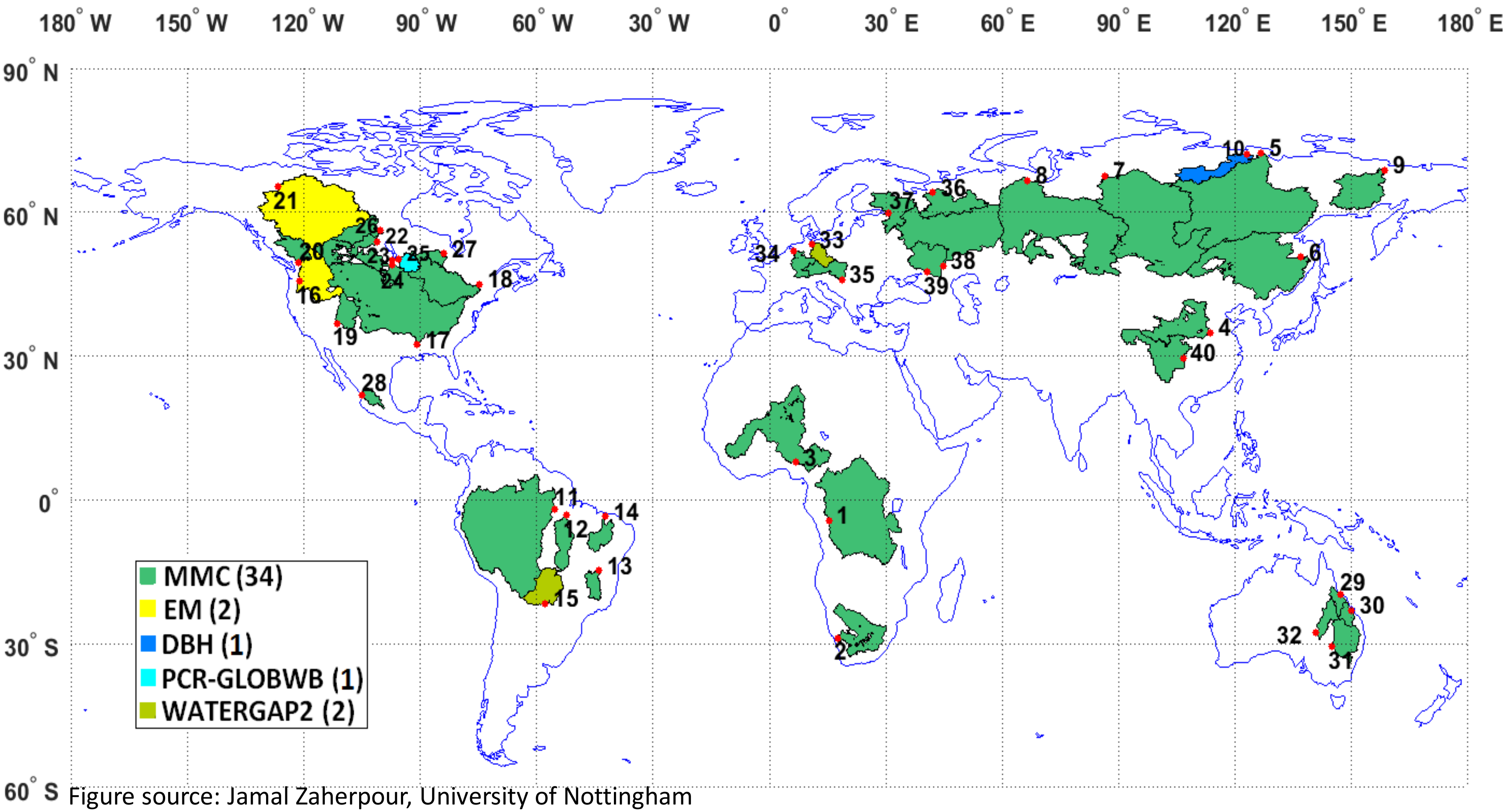



Figure source: Jamal Zaherpour, University of Nottingham

# Results

- Quantifying simulation improvement by MMC over the best GHM and EM:

| Hydrobelt | IPE improvement vs. best GHM (%) | IPE improvement vs. EM (%) |
|---|---|---|
| BOR | -37.3 | -47.1 |
| NML | -19.0 | -63.3 |
| NDR | -12.3 | -71.3 |
| NST | -23.2 | -88.4 |
| EQT | -30.8 | -50.7 |
| SST | -27.4 | -87.4 |
| SDR | -25.1 | -94.6 |
| SML | -9.6 | -96.6 |
| **Mean** | **-29.3 *** | **-58.1 *** |

* Calculated over individual catchments

# Results

- MMC also yields improvements in timeseries of simulations:

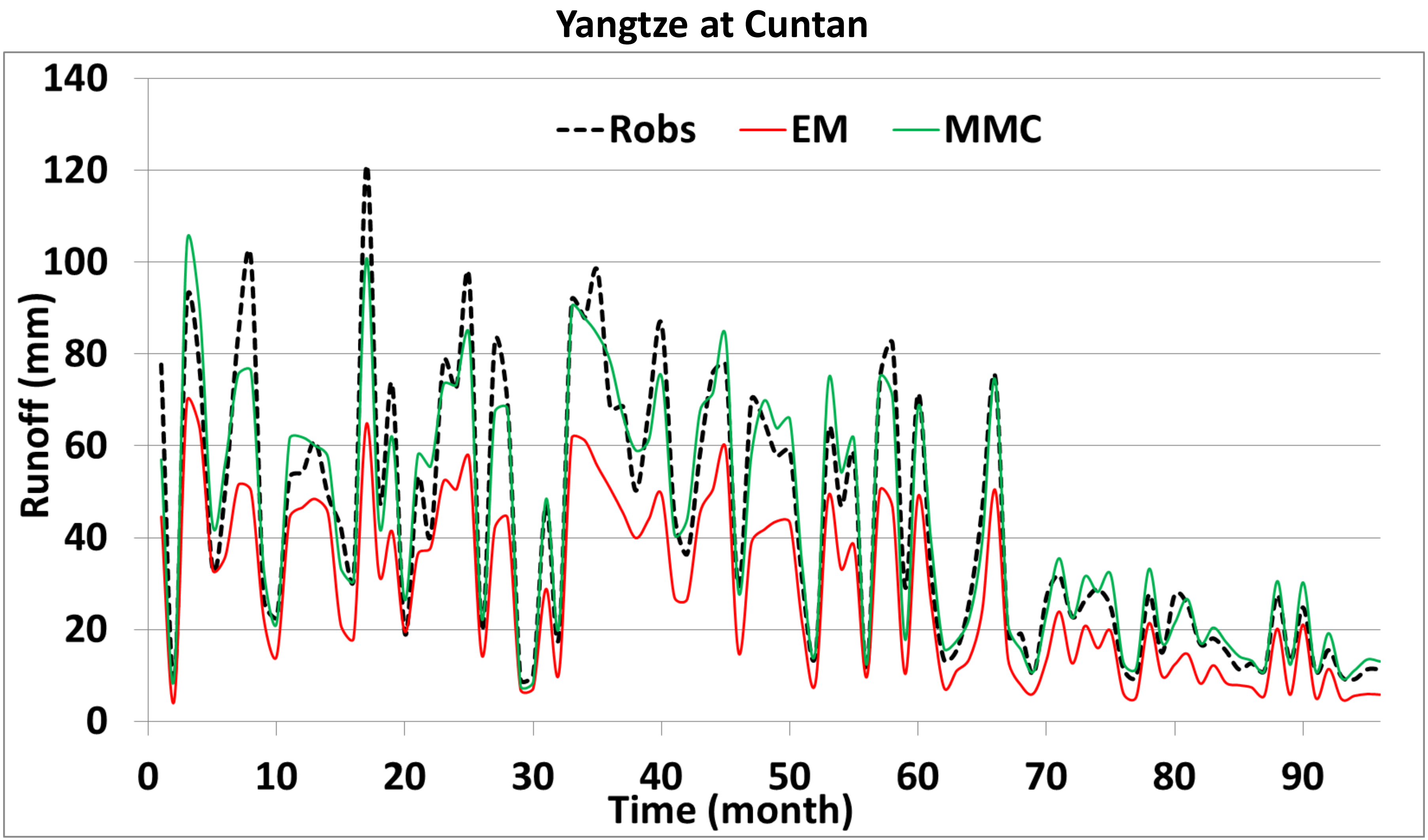



Figure source: Jamal Zaherpour, University of Nottingham

# Conclusions and future work

- We see significant improvements over the ensemble mean in simulations with MLAs.

- Ensemble spreads should still be presented alongside the MMC – we have not reduced the uncertainty!
  - But we have provided a better informed summary of an ensemble than the ensemble mean.

- Our proof of concept is being submitted to the ISI-MIP2 *ERL* Special Issue.

- Ongoing work is using the developed MMCs with climate change projections.

# Dual-scale ensembles of climate impacts from 1, 2 and 3°C global warming

# THE ENSEMBLES

- What do we mean by "dual-scale"?
- Two broad types of hydrological model are:

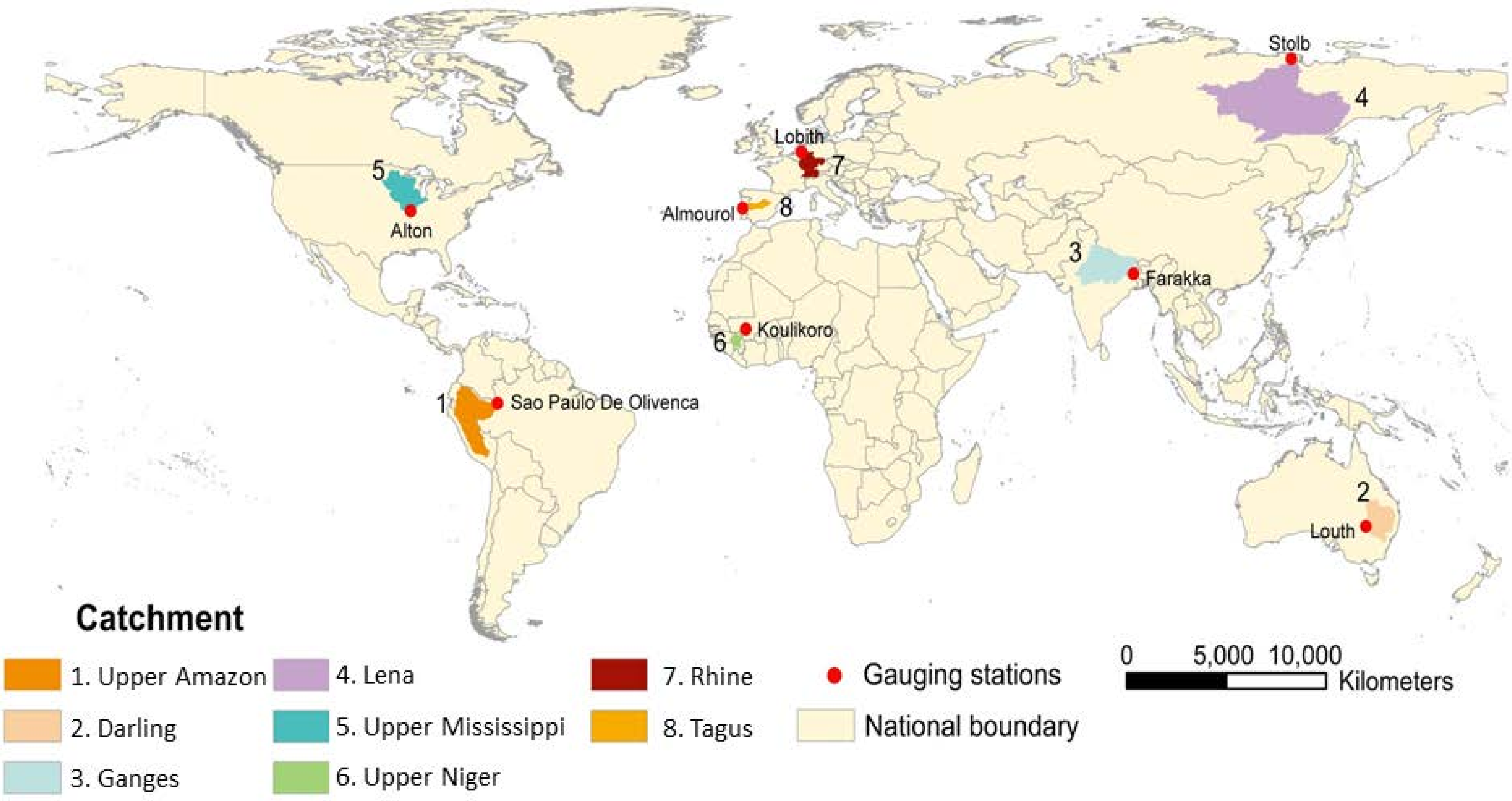
| | Global-scale | Catchment-scale |
|---|---|---|
| Abbreviation: | Glob-HM | Cat-HM |
| Spatial domain: | Global | Catchment |
| Approach to calibration: | Spatially generalised parameters across the globe (can change spatially in some models) | Model parameters tuned so that simulated flow matches observed flow |

# RESEARCH QUESTION

Are there systematic differences between projections of runoff change from two ensembles that are comprised of different types of hydrological model?

# STUDY APPROACH

- Investigated the research question for eight study catchments:




Source: Gosling et al. (under review, *Climatic Change*)

# STUDY APPROACH

- The two ensembles:

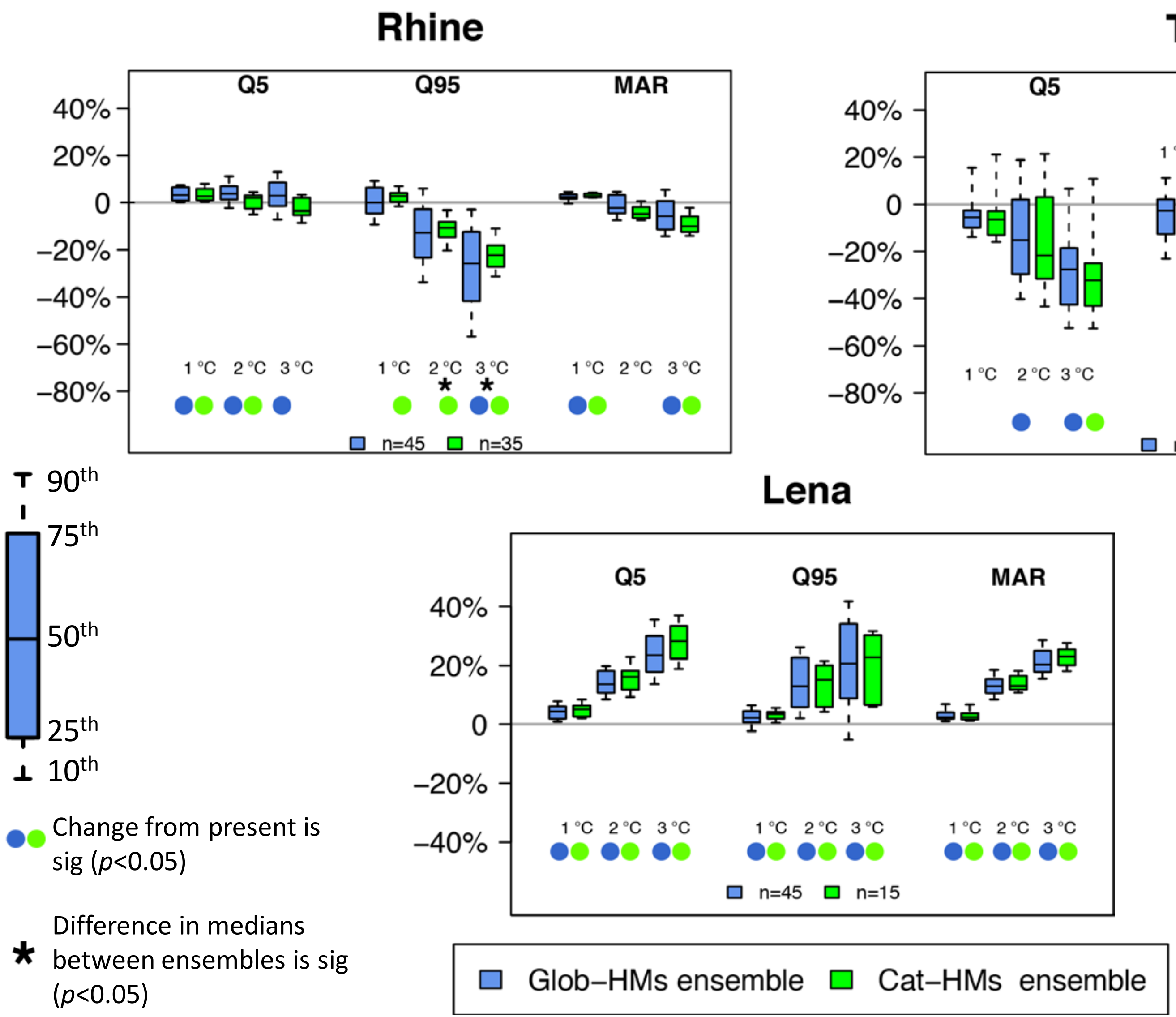
| | Catchment | Upper Amazon | Darling | Ganges | Lena | Upper Mississippi | Upper Niger | Rhine | Tagus |
|---|---|---|---|---|---|---|---|---|---|
| | Gauge | Sao Paulo De Olivenca | Louth | Farakka | Stolb | Alton | Koulikoro | Lobith | Almourol |
| | **Upstream drainage area (km$^2$)** | 990,781 | 489,300 | 835,000 | 2,460,000 | 444,185 | 120,000 | 160,800 | 67,490 |
| | | | | | | | | | |
| **Cat-HM ensemble** | **HBV** (Bergstrom and Forsman 1973) | | | | | | | | |
| | **HYMOD** (Boyle 2001) | | | | | | | | |
| | **mHM** (Kumar et al. 2013) | | | | | | | | |
| | **SWIM** (Krysanova et al. 1998) | | | | | | | | |
| | **VIC-Cat-HM** * (Liang et al. 1994) | | | | | | | | |
| | **WaterGAP3** + (Verzano 2009) | | | | | | | | |
| | **HYPE** (Lindstrom et al. 2010) | | | | | | | | |
| | **SWAT** (Arnold et al. 1993) | | | | | | | | |
| | Number of Cat-HM simulations | 7 | 3 | 6 | 3 | 7 | 6 | 7 | 3 |
| | | | | | | | | | |
| **Glob-HM ensemble** | **LPJmL** (Bondeau et al. 2007) | | | | | | | | |
| | **Mac-PDM.09** (Gosling and Arnell 2011) | | | | | | | | |
| | **MPI-HM** (Hagemann and Dümenil 1997) | | | | | | | | |
| | **PCR-GLOBWB** (Wada et al. 2014) | | | | | | | | |
| | **WBMplus** (Wisser et al. 2010) | | | | | | | | |
| | **H08** (Hanasaki et al. 2008) | | | | | | | | |
| | **MATSIRO** (Pokhrel et al. 2011) | | | | | | | | |
| | **VIC-Glob-HM** * (Liang et al. 1994) | | | | | | | | |
| | **DBH** (Tang et al. 2007) | | | | | | | | |
| | Number of Glob-HM simulations | 9 | 9 | 9 | 9 | 9 | 9 | 9 | 9 |
| | | | | | | | | | |
| | **Total number of simulations** | **16** | **12** | **15** | **12** | **16** | **15** | **16** | **12** |

Source: Gosling et al. (under review, *Climatic Change*)

- Used the 31-year periods corresponding to 1, 2, and 3° C global-mean warming relative to pre-industrial from 5 GCMs.

# Results

- **Finding 1**: both ensembles generally show significant changes with global warming relative to present-day for some catchments:

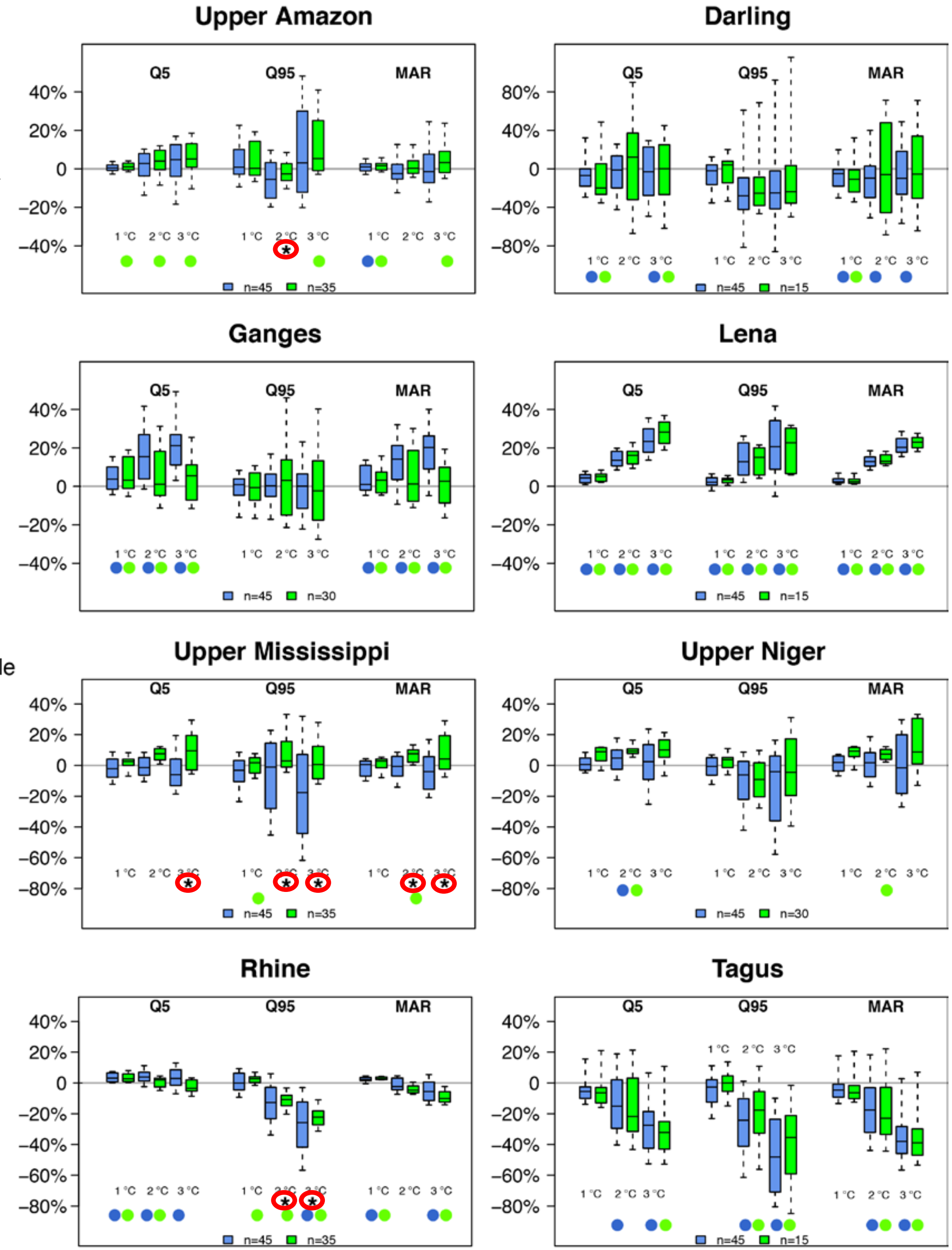



Source: Gosling et al. (under review, *Climatic Change*)

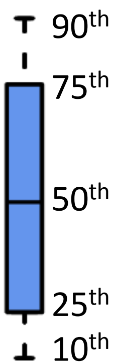
Source: Gosling et al. (under review, *Climatic Change*)

# Results

- **Finding 2**: projections are generally similar between ensembles.
- Only in 8 cases out of 72 is the difference between ensemble projections significant (*)




Glob–HMs ensemble   Cat–HMs ensemble



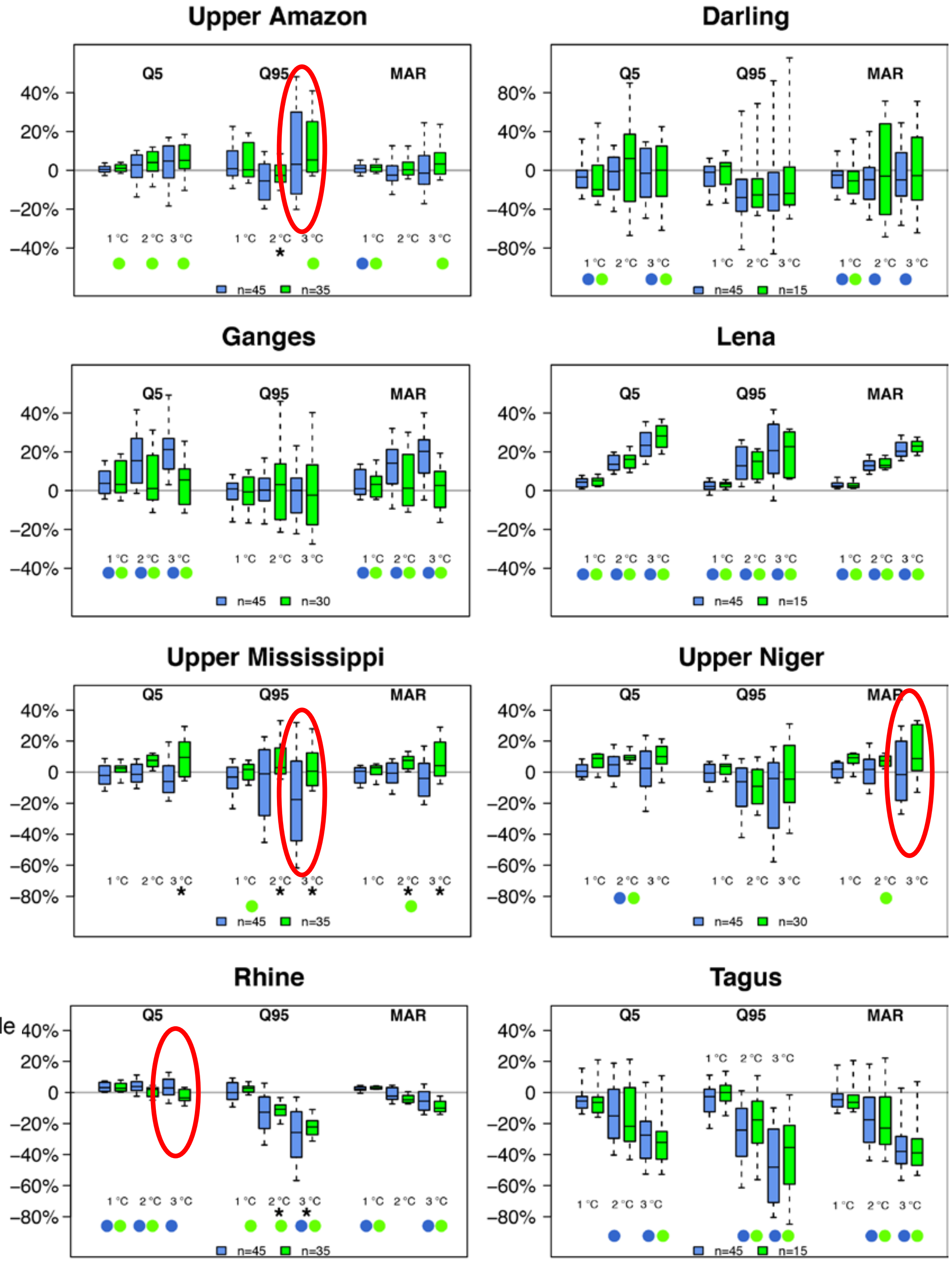

Change from present is sig ($p$<0.05)

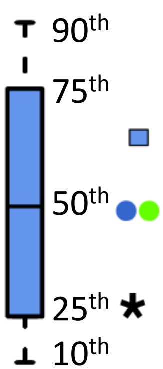
(*) Difference in medians between ensembles is sig ($p$<0.05)

Source: Gosling et al. (under review, *Climatic Change*)

# RESULTS

- **Finding 3**: the Glob-HM ensembles generally have a wider spread than the Cat-HM ensembles.
- Implies we could have more confidence in the Cat-HM ensemble.
- But this confidence could be misplaced because:
  - Strictly, spreads only show model agreement.
  - Models are being run outside of their calibration zone.
  - If all models miss out a key process the true outcome could lie outside the ensemble spread.



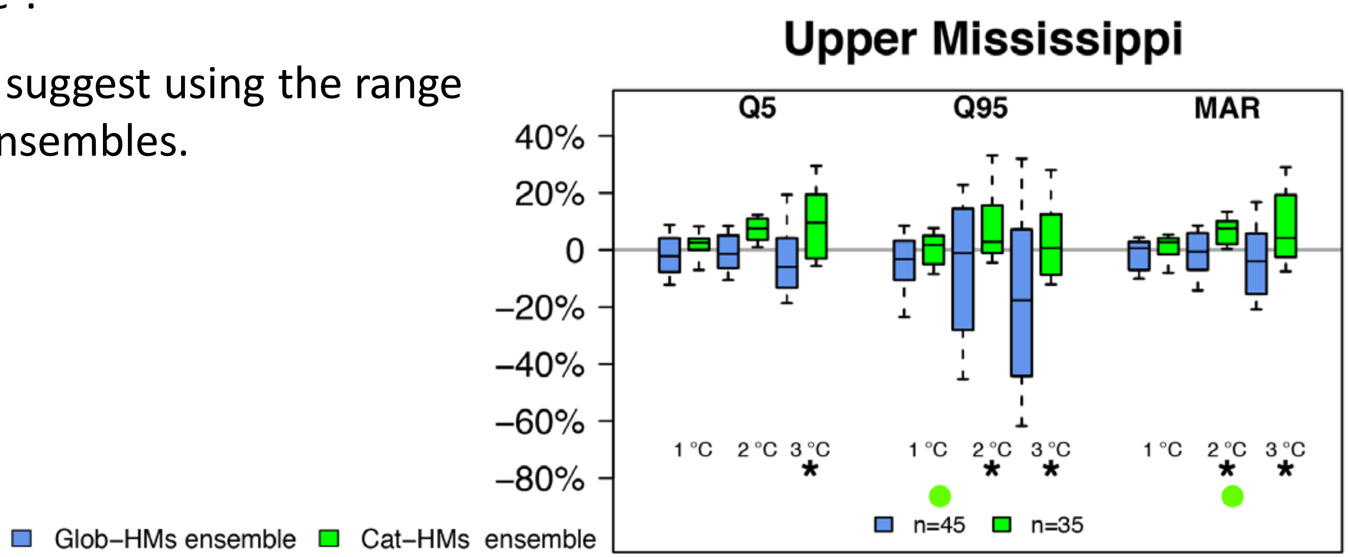

Glob–HMs ensemble Cat–HMs ensemble

Change from present is sig ($p$<0.05)

Difference in medians between ensembles is sig ($p$<0.05)

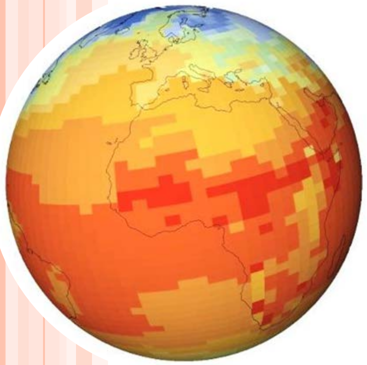
# CONCLUSIONS

- The two ensembles generally show similar median changes in response to climate change, with only a few distinct differences.
- Limiting global warming to 1.5° C in line with the Paris Agreement could avoid significant impacts in some catchments.
- A decision-maker might ask: if I am interested in the effect of climate change on the Mississippi, which model type should I use?
  - We advise against using one over another.
  - Although the Cat-HMs perform better in present-day, both ensembles are based upon models regularly used in the scientific community, run outside of their "comfort zone".
  - Therefore we suggest using the range across both ensembles.



# Future directions for the water sector and thanks

- All of the research referred to here would not have been possible without the many modellers who contributed global- and catchment-scale simulations to the ISI-MIP project.
- Thank you to my joint coordinator for the global water sector, Rutger Dankers (Met Office), who is stepping down from the role.
  - A new joint coordinator will be announced after the workshop.
- Future directions include:
  - Exploring synergies between the different spatial scales of modelling approaches – use this to understand model differences.
  - Working towards addressing the ISI-MIP2B protocol.



ISIMIP
Inter-Sectoral Impact Model Intercomparison Project

# Questions


**Simon Gosling**

Director of Research and Associate Professor in Climate Risk

School of Geography, University of Nottingham, UK

simon.gosling@nottingham.ac.uk

@simon_gosling