Global Flood Partnership Conference 2019
11-13 June 2019, Guangzhou, China

河海大學
HOHAI UNIVERSITY

Hydrological Ensemble Prediction – A New Paradigm in Hydrological Forecasting

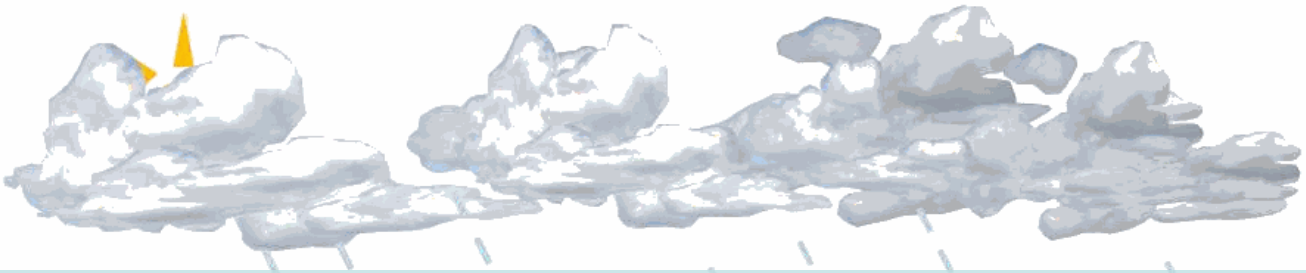
Qingyun Duan

College of Hydrology and Water Resources
Hohai University

June 11, 2019

What is Hydrological Forecasting?

河海大學
HOHAI UNIVERSITY

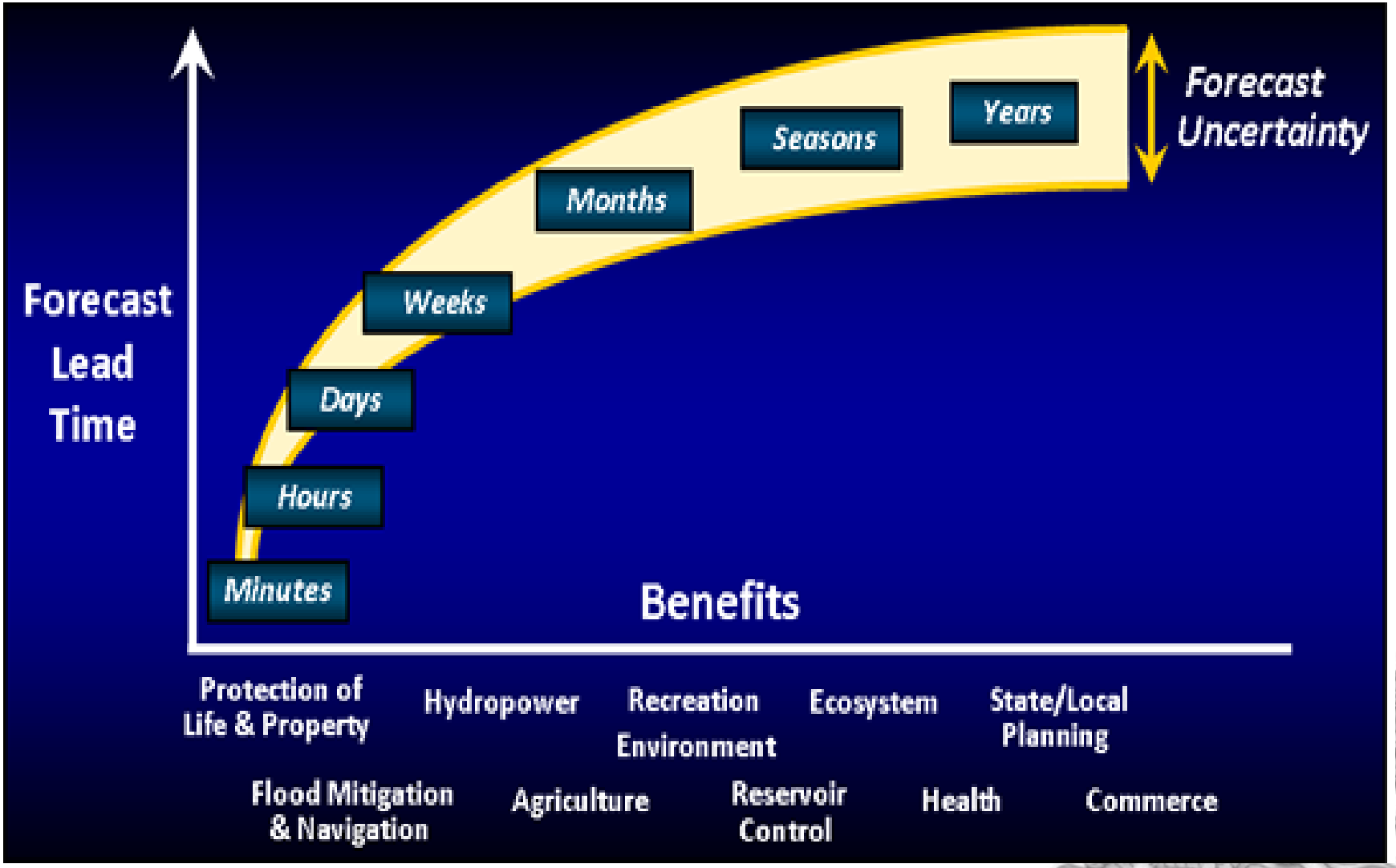


Hydrological forecasting addresses those questions:

- Where does water flow?
- How much water is there?
- What is the chance that my house would be flooded?

Hydrological Forecasts and Societal Benefits

河海大学
HOHAI UNIVERSITY



From NOAA Website

Where Do Uncertainties Come From?

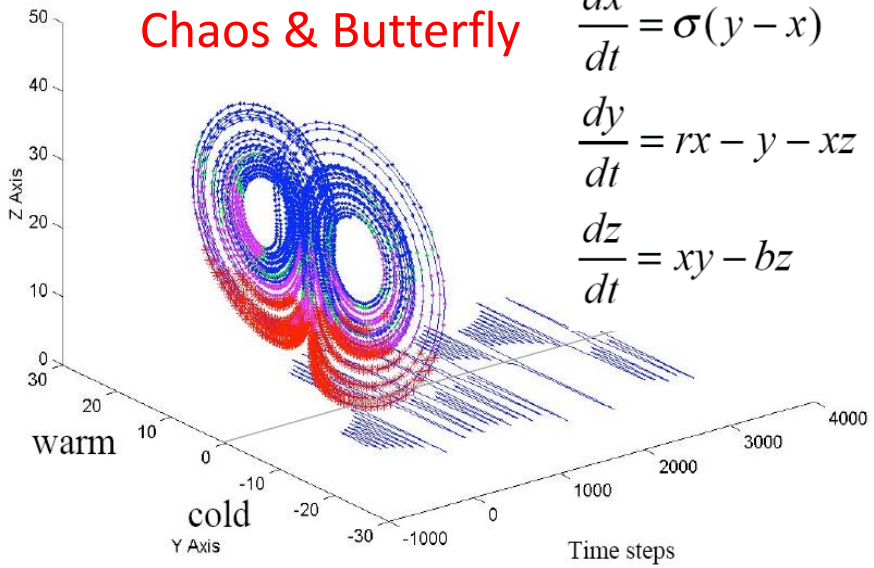
河海大学 HOHAI UNIVERSITY

Chaos & Butterfly

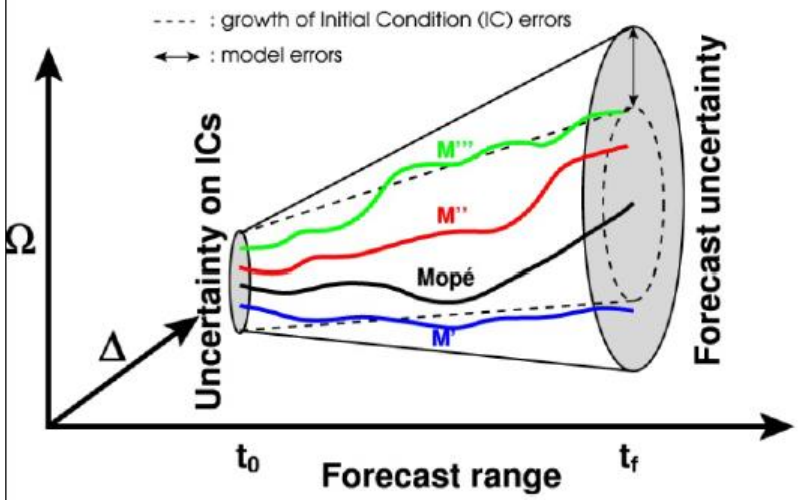
$$\frac{dx}{dt} = \sigma(y - x)$$

$$\frac{dy}{dt} = rx - y - xz$$

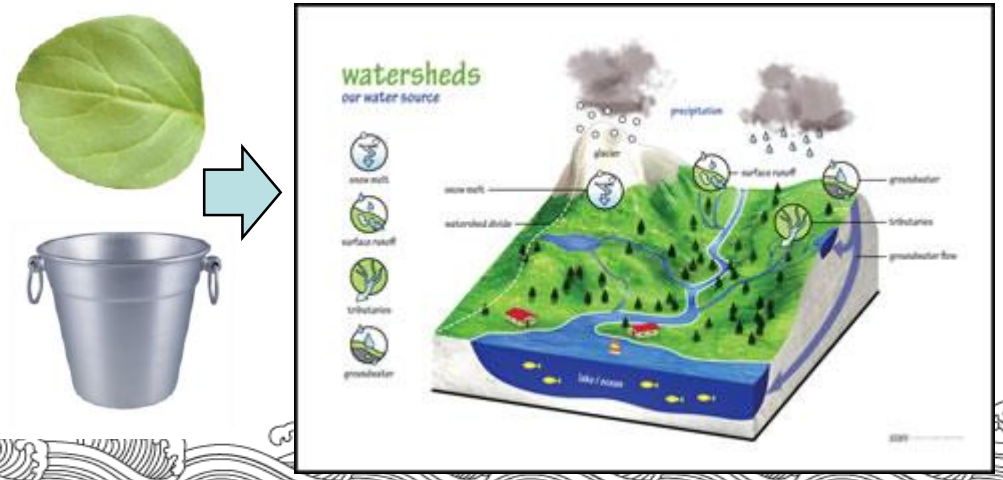
$$\frac{dz}{dt} = xy - bz$$



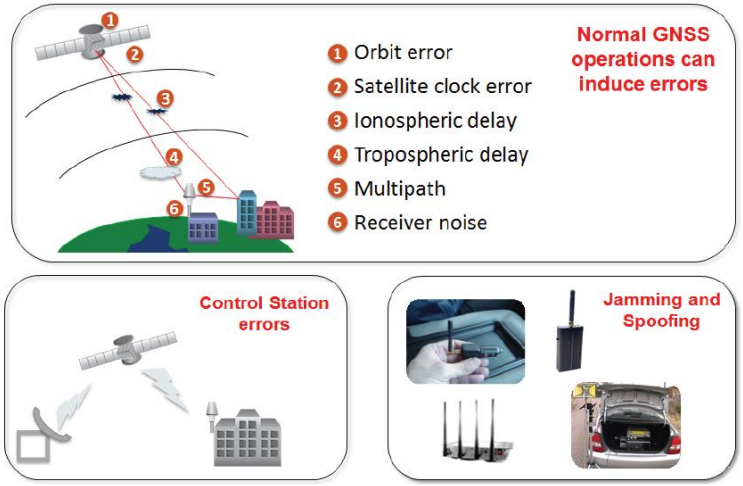
Initial Condition Uncertainty



Model Uncertainty



Observation Uncertainty

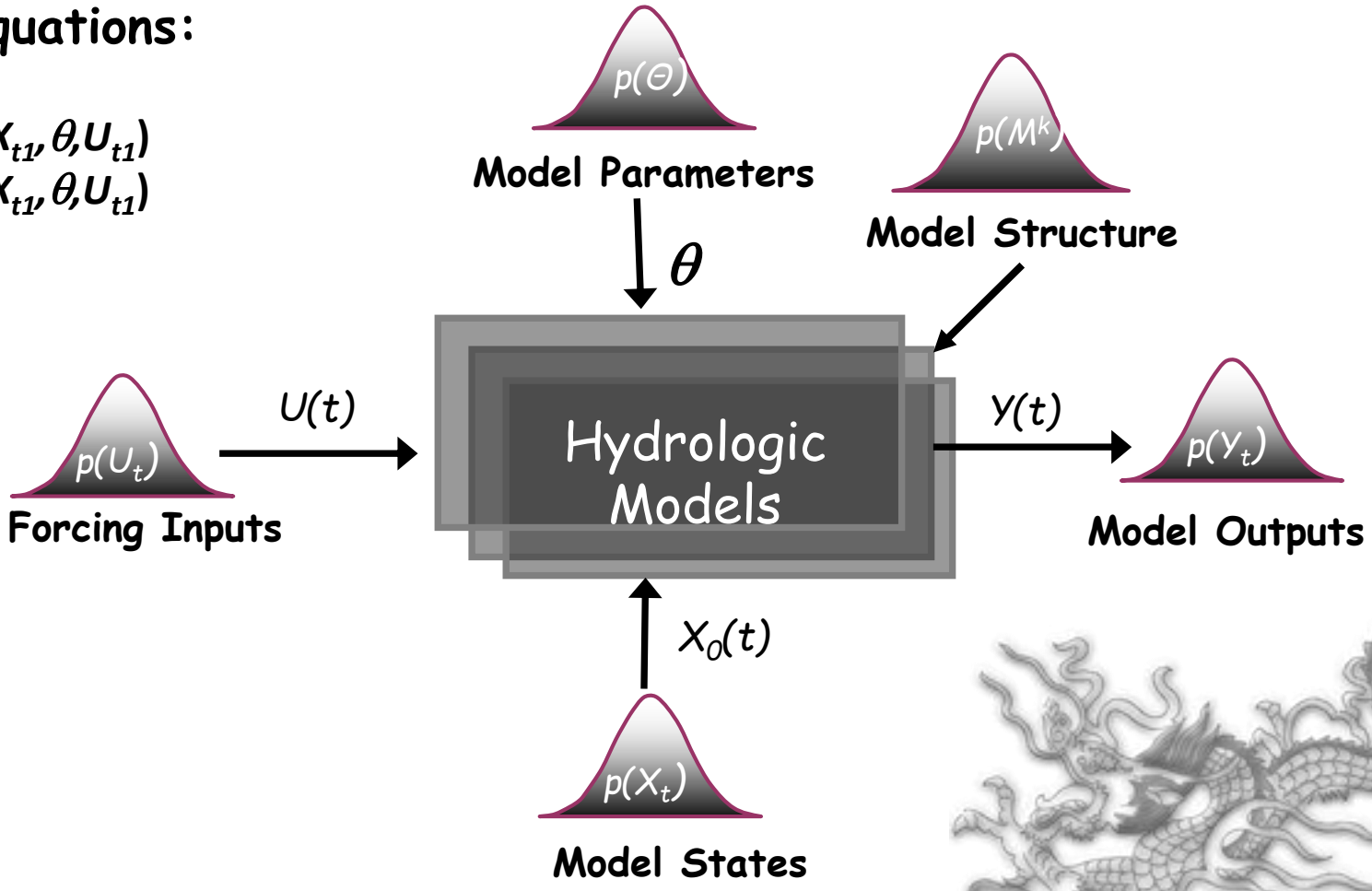


Uncertainties Are Prevalent in Hydrologic Forecasting

河海大學 HOHAI UNIVERSITY

Model Equations:

$$X_{t2} = F(X_{t1}, \theta, U_{t1})$$
$$Y_{t2} = G(X_{t1}, \theta, U_{t1})$$



How to Handle Uncertainties in Hydrologic Forecasts

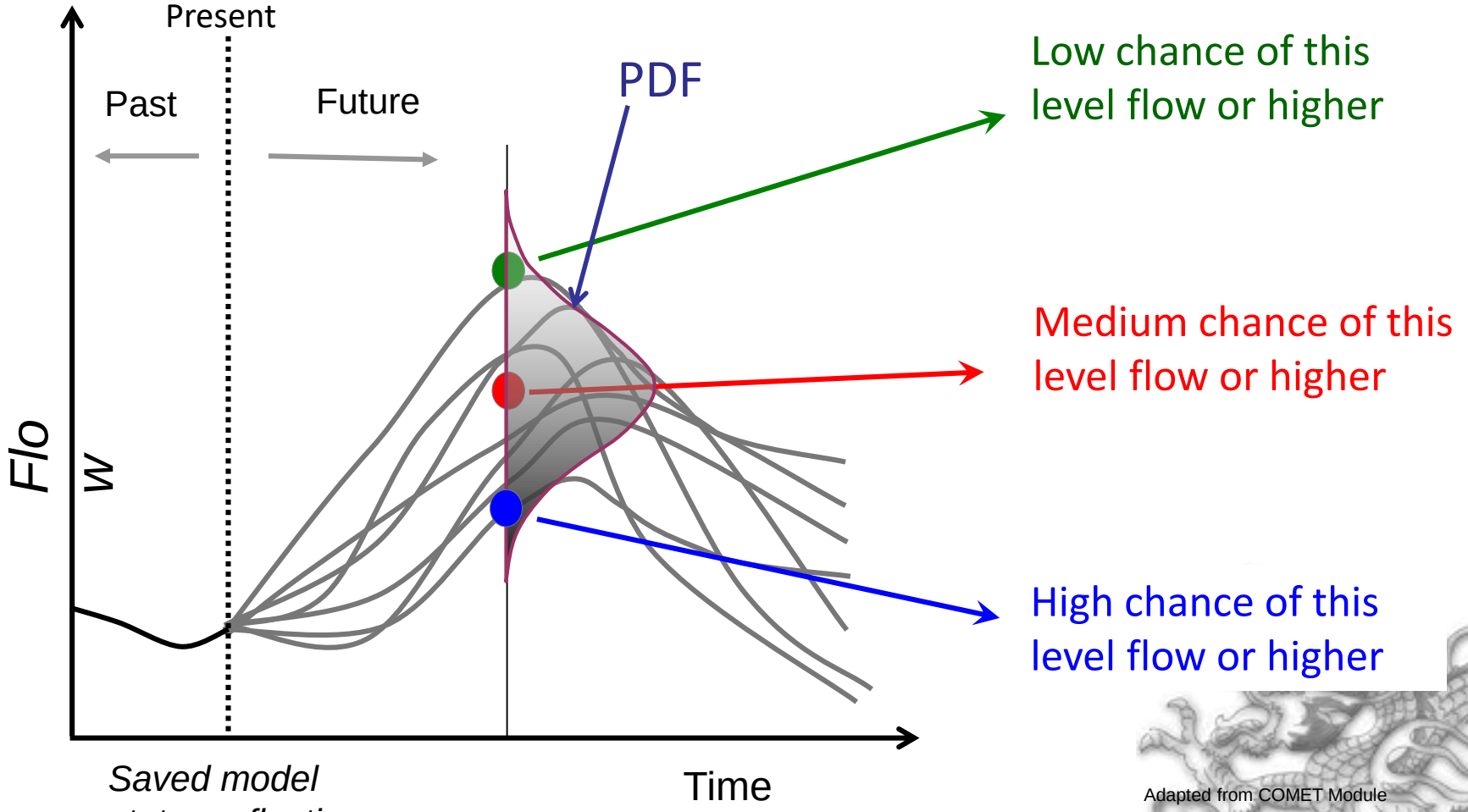
河海大學 HOHAI UNIVERSITY

- Theoretically the most direct way to handle the uncertainties is to account for them using **Stochastic dynamical equations and solve them analytically or numerically**
  - However, it is not **practical** !!!
- The only practical way to quantify the uncertainties today is to employ **Ensemble Forecasting methods**

What Is An Ensemble Forecast?

河海大學 HOHAI UNIVERSITY

Definition: A set of forecasts of hydrologic events for pre-specified lead times, generated by perturbing different uncertain factors



Adapted from COMET Module

Illustration of Probabilistic Ensemble Forecast Products

河海大学 HOHAI UNIVERSITY

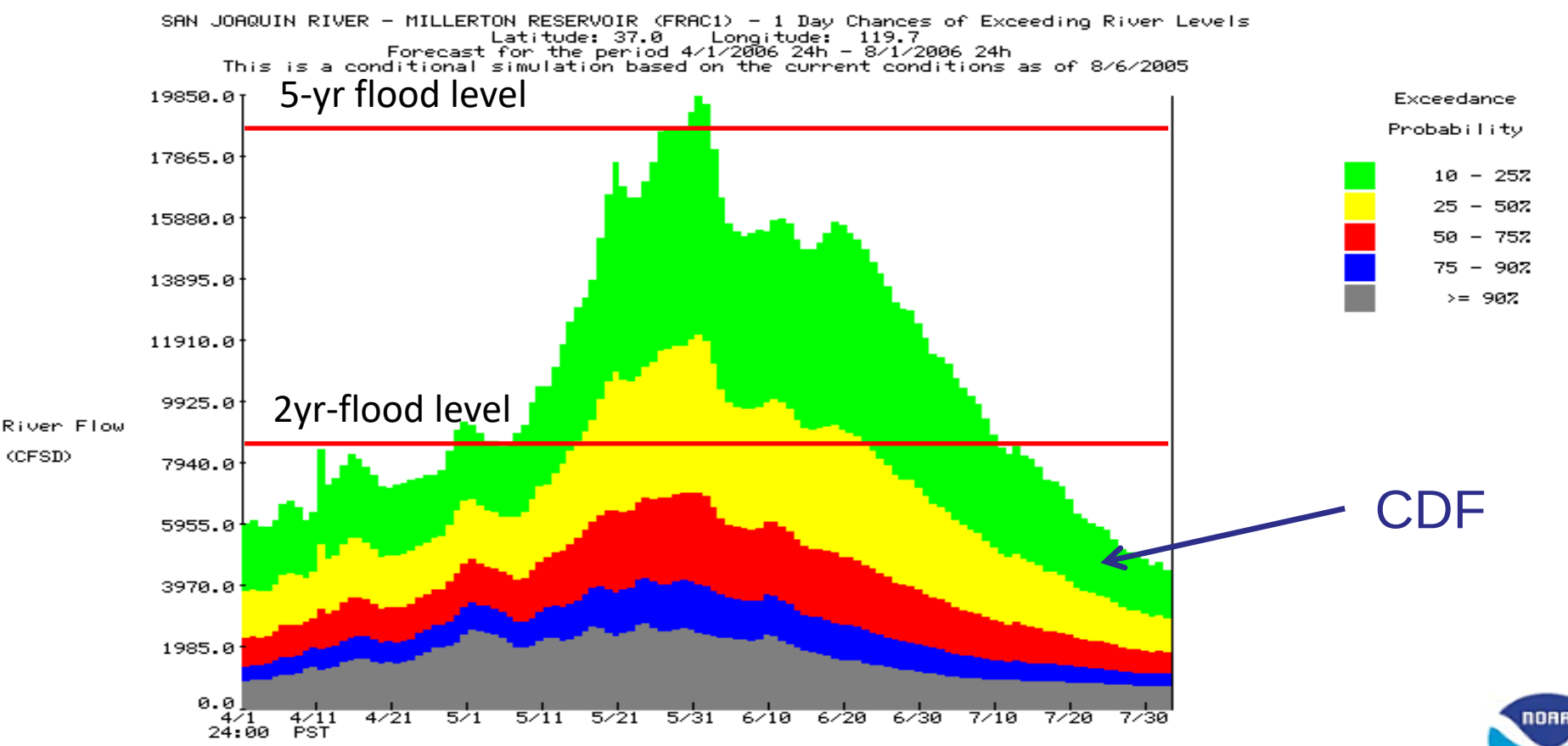
SAN JOAQUIN RIVER - MILLERTON RESERVOIR (FRAC1)

Latitude: 37.00° N Longitude: 119.71° W Elevation: 581 Feet

Location: Fresno County in California

Monitor Stage: N/A *Flood Stage: 578.0 Feet*

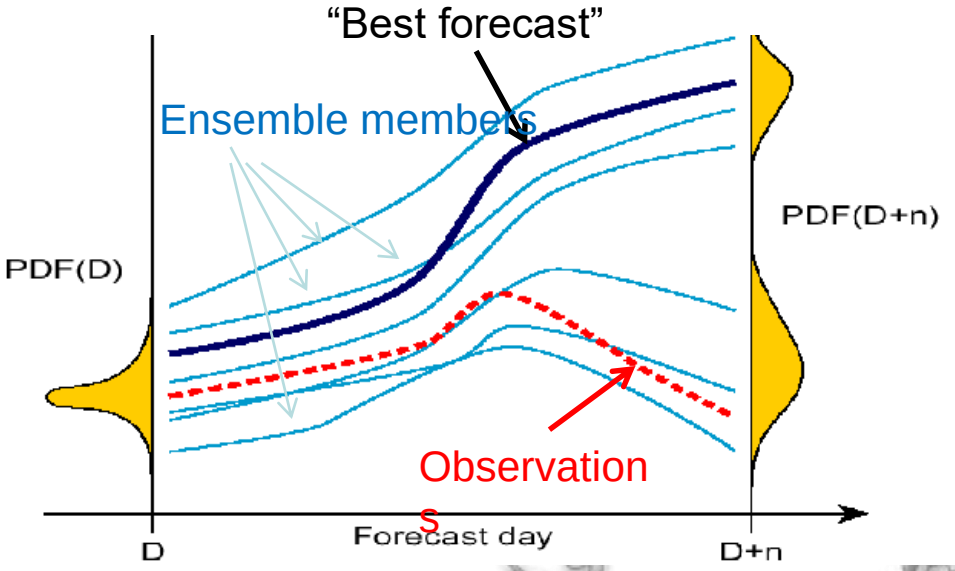
1 Day Chances of Exceeding River Levels

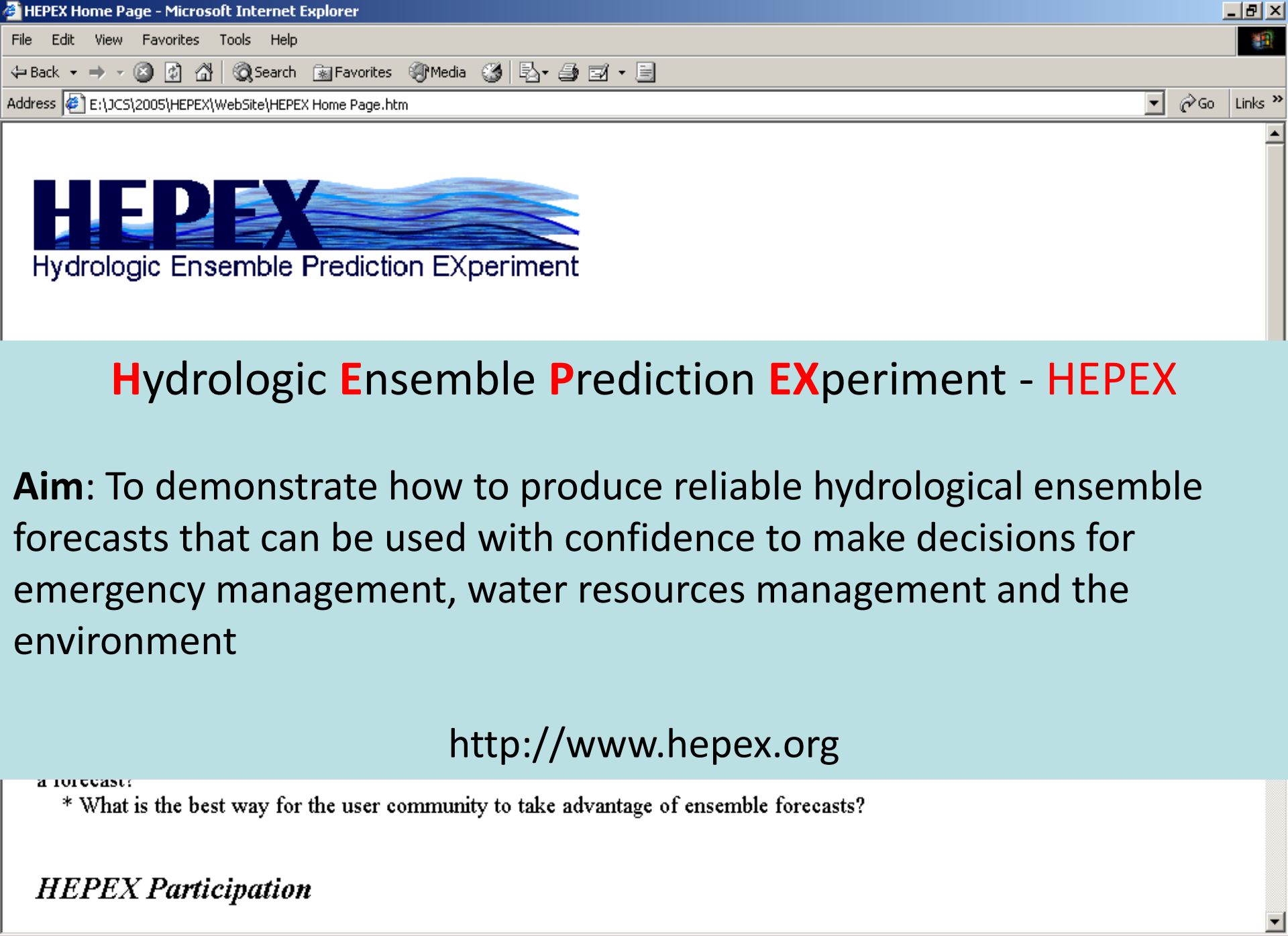


Advantages of Ensemble Forecasts

河海大學 HOHAI UNIVERSITY

- To provide quantitative uncertainty information：
  - Confidence information (for forecaster)
  - User-specified risk information (for user)
- To improve forecast accuracy
  - The average performance of ensemble predictions is better than any single prediction
- To extend forecast lead times
  - Meteorological predictions contain large uncertainties. Single valued predictions cannot express the uncertainty information. Therefore, they have shorter lead times

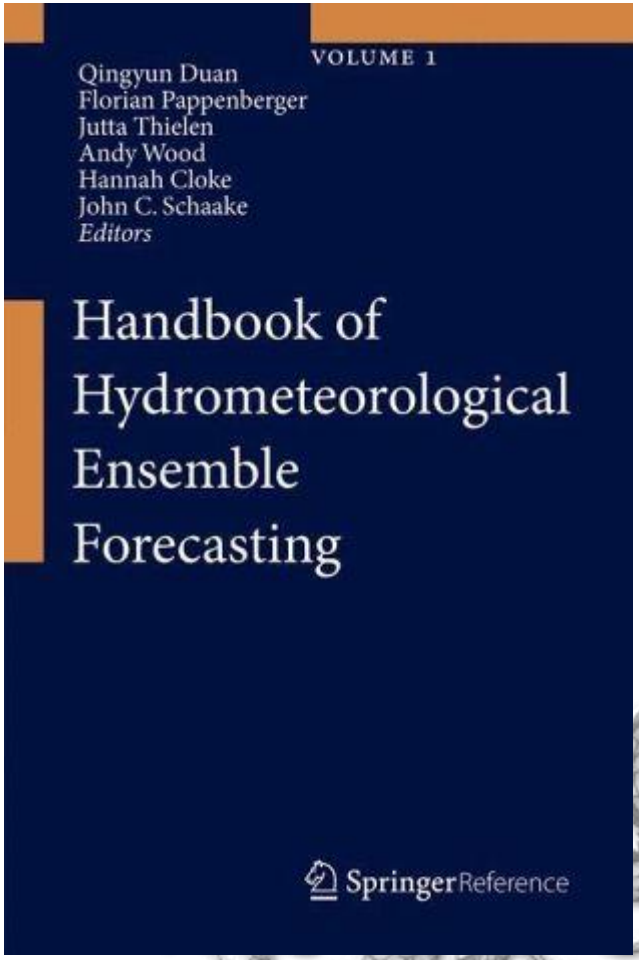



HEPEX Home Page - Microsoft Internet Explorer
File Edit View Favorites Tools Help
Back Search Favorites Media
Address E:\JCS\2005\HEPEX\WebSite\HEPEX Home Page.htm
Go Links
HEPEX
Hydrologic Ensemble Prediction EXperiment
Hydrologic Ensemble Prediction EXperiment - HEPEX
Aim: To demonstrate how to produce reliable hydrological ensemble forecasts that can be used with confidence to make decisions for emergency management, water resources management and the environment
http://www.hepex.org
a forecast?
* What is the best way for the user community to take advantage of ensemble forecasts?
HEPEX Participation
Done
2019/8/7
Internet

Handbook of Hydrometeorological Ensemble Forecasting

河海大學
HOHAI UNIVERSITY

- Editor-in-Chief： Qingyun Duan et al.
- Publisher： Springer-Nature
- Publication series： Major Reference Books
- Publication date： Jan. 9, 2019



河海大学
HOHAI UNIVERSITY

Hydrologic Ensemble Forecast System

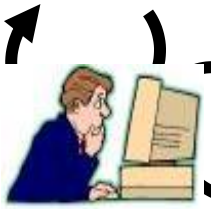
Atmospheric Ensemble Pre-Processor

Ensemble Data Assimilator

Hydrology and Water Resources Models

Parametric Ensemble Processor

Hydrologic Ensemble Post-Processor

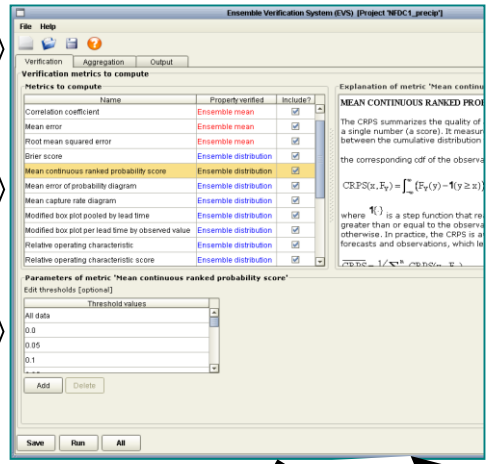


Hydrology and Water Resources Ensemble Product Generator

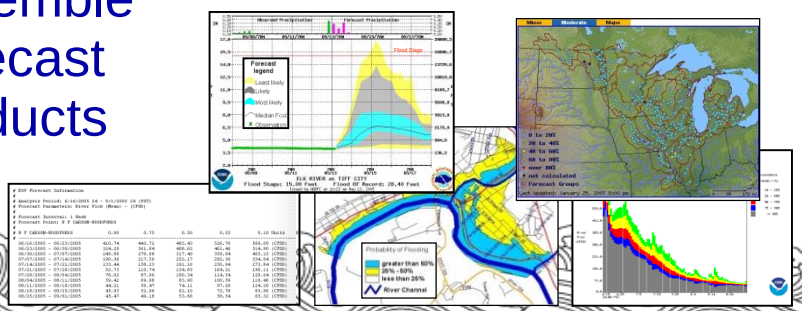
Forecasters

Users

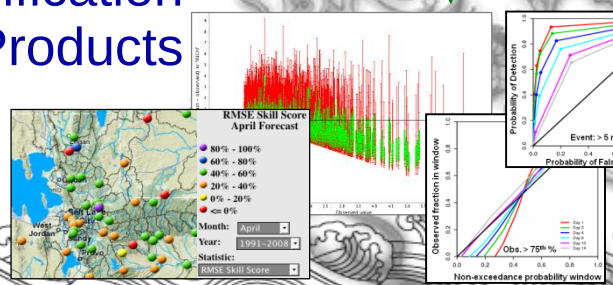
Ensemble Verification System



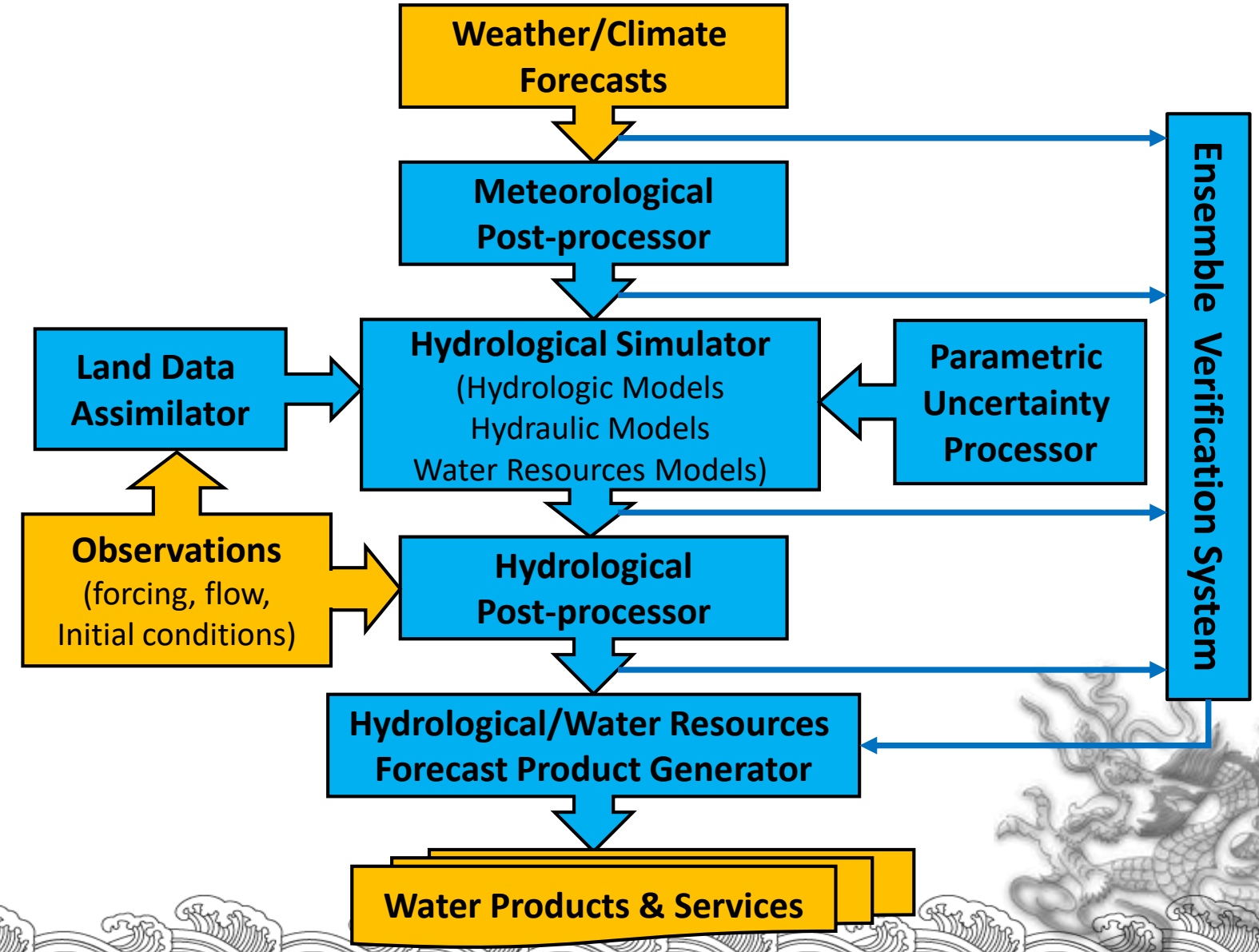
Ensemble Forecast Products



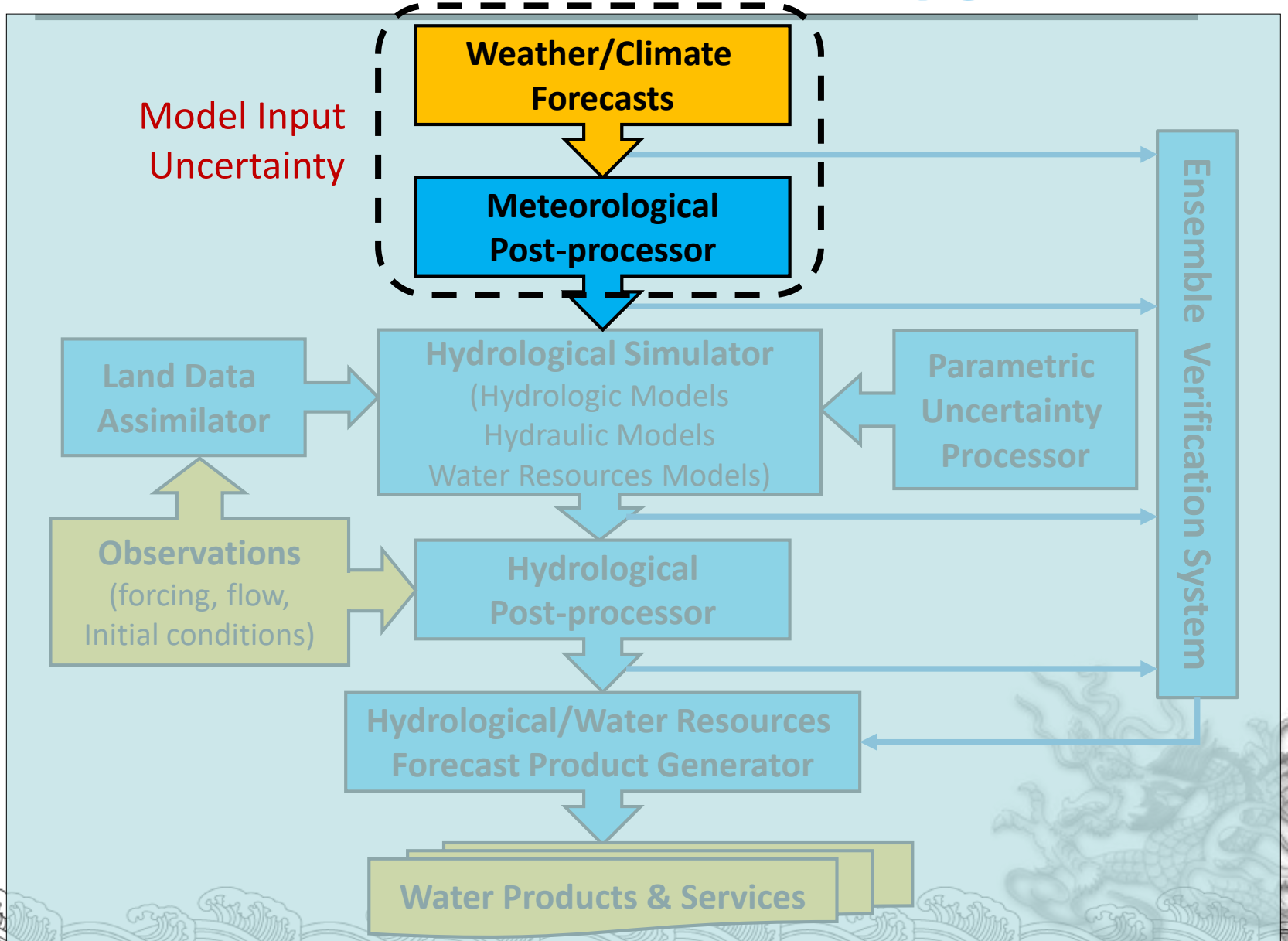
Verification Products



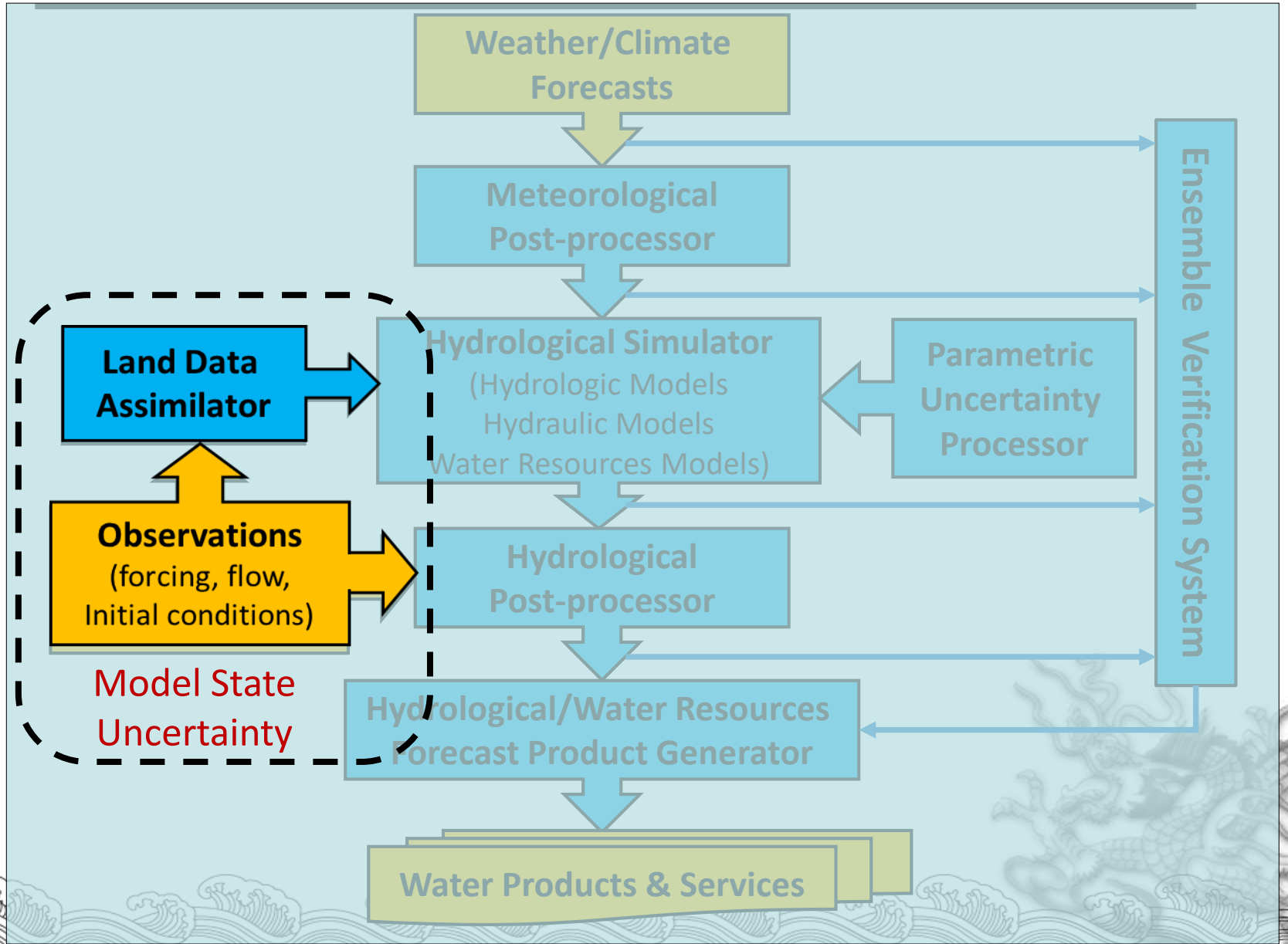
The Hydrologic Ensemble Prediction Experiment (HEPEX) Framework



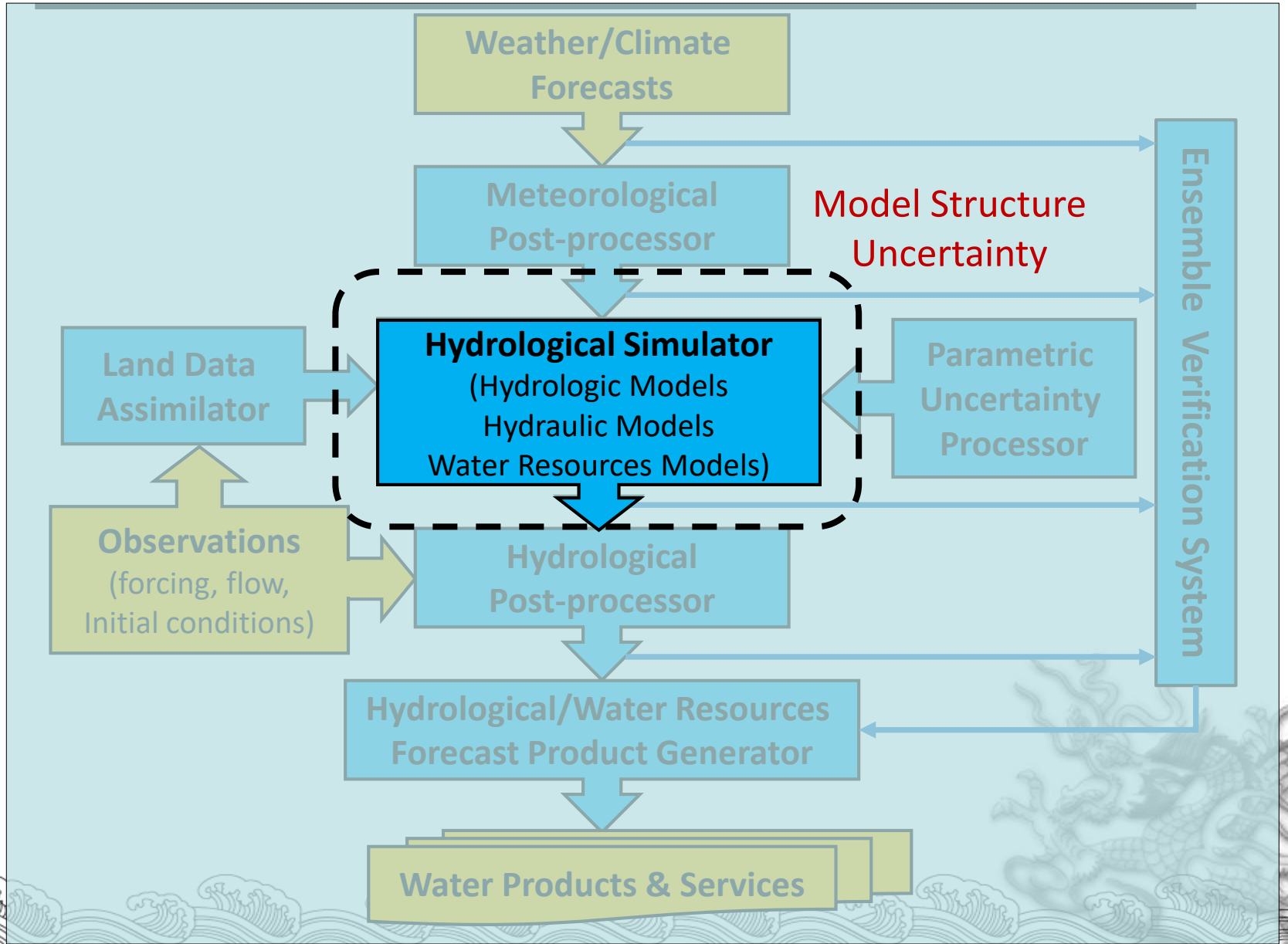
Confronting Uncertainties at Their Sources



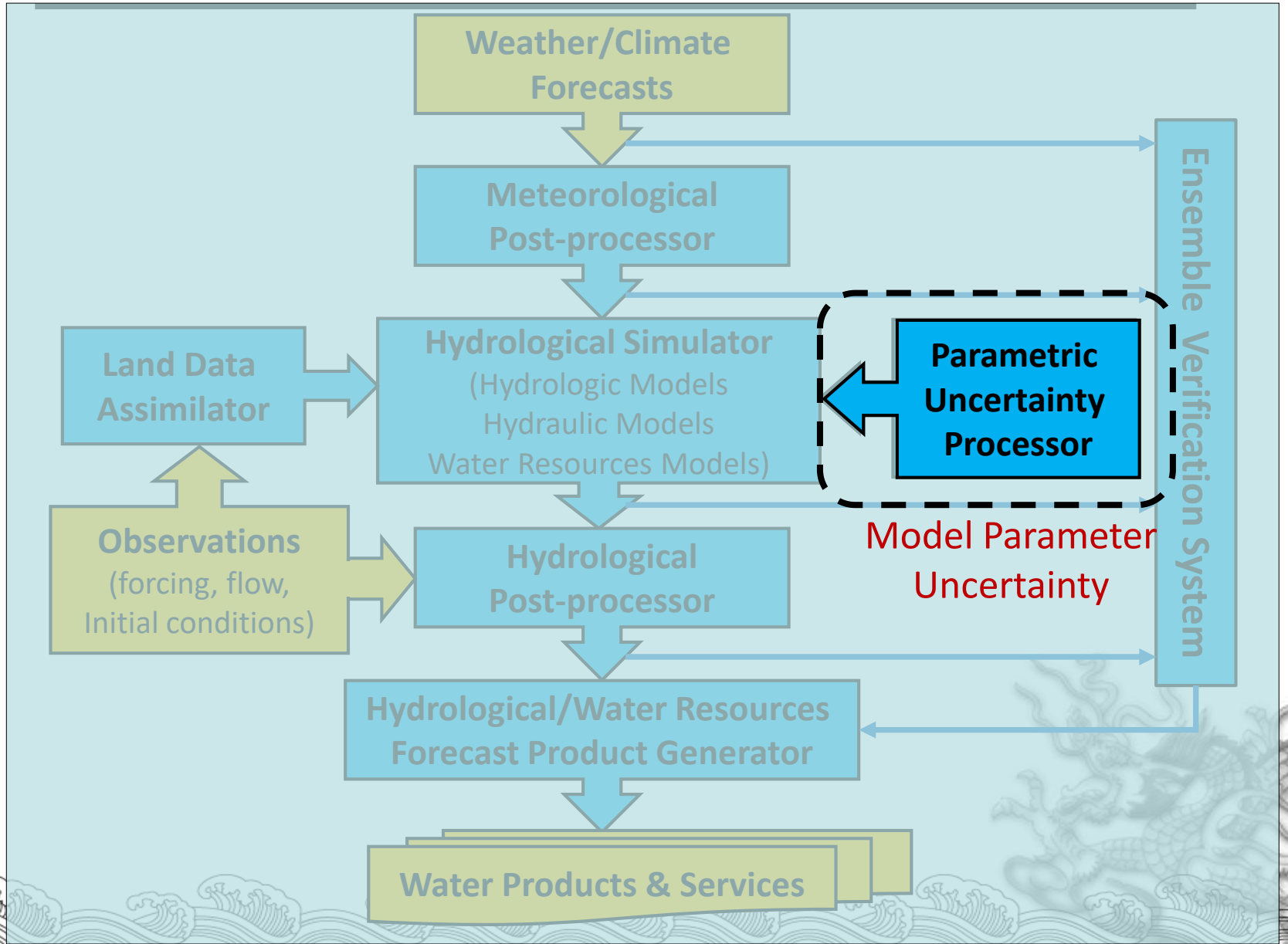
Confronting Uncertainties at Their Sources



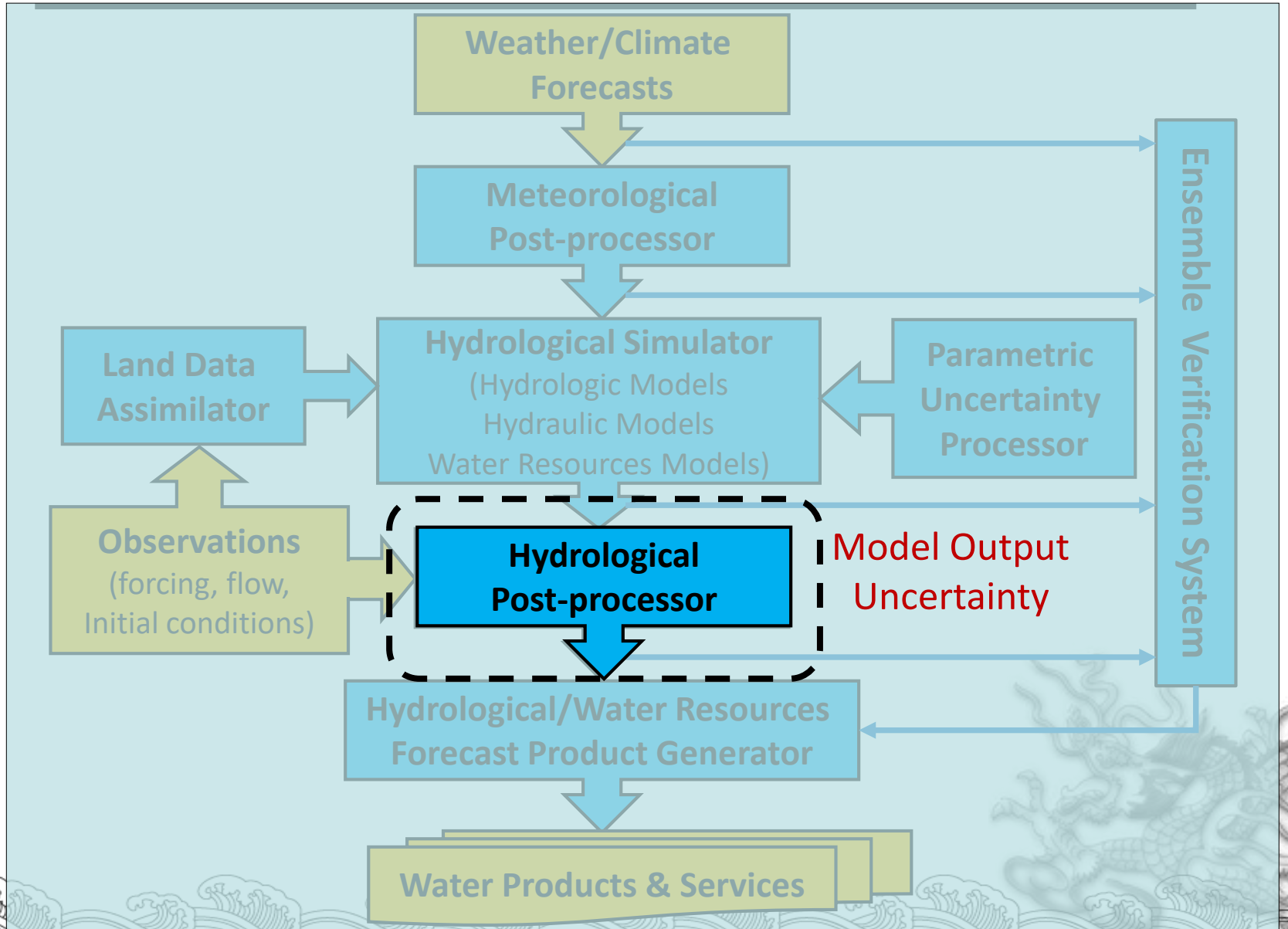
Confronting Uncertainties at Their Sources



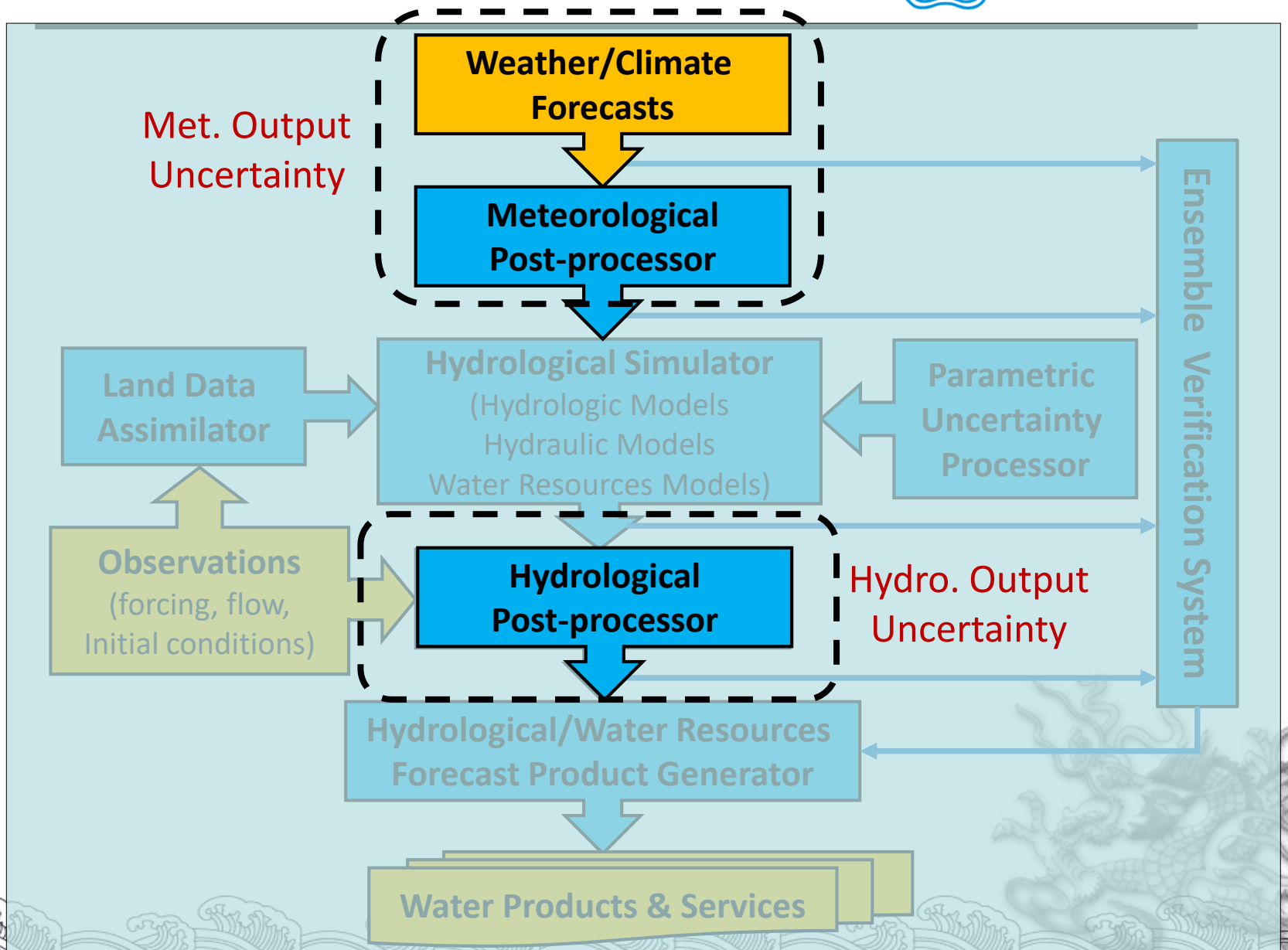
Confronting Uncertainties at Their Sources



Confronting Uncertainties at Their Sources



Confronting Model Output Uncertainties

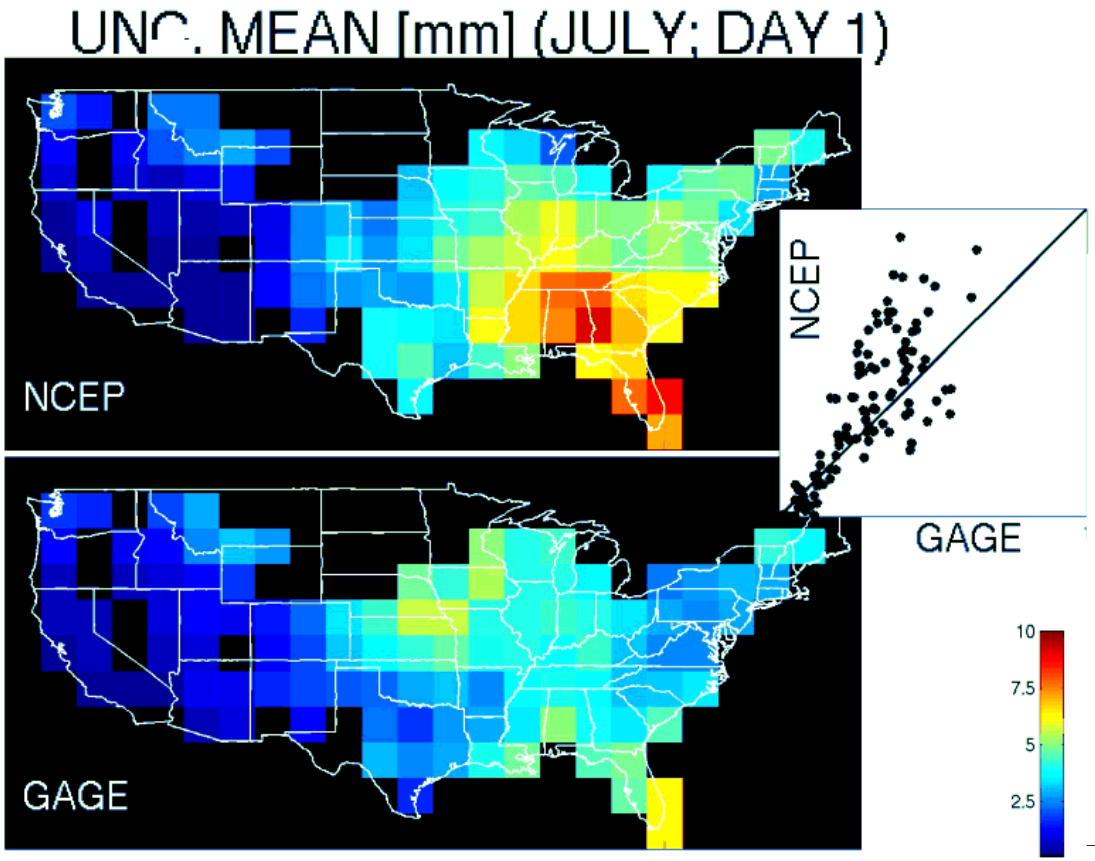


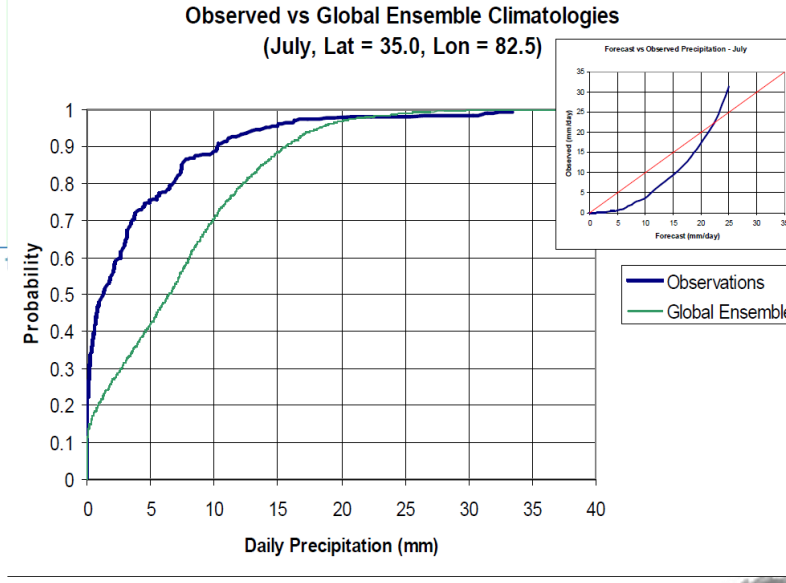
Confronting Model Output (Forecast) Uncertainty
Statistical Post-Processors

- Statistical post-processors are **statistical models** based on past samples of forecast-observation relationships to produce bias corrected, downscaled space-time series of hydrometeorological variables.

- The means include all kinds of statistical methods including big data, machine learning, deep learning, AI, etc.

Why Post-Processing?





Schaake, 2004

Problems: Skill varies with lead times;

Small events overestimated while large events underestimated

Heteroscedasticity: variances change with magnitude

Non-Gaussian distribution

Post-processing Methods for Meteorological Forecasts

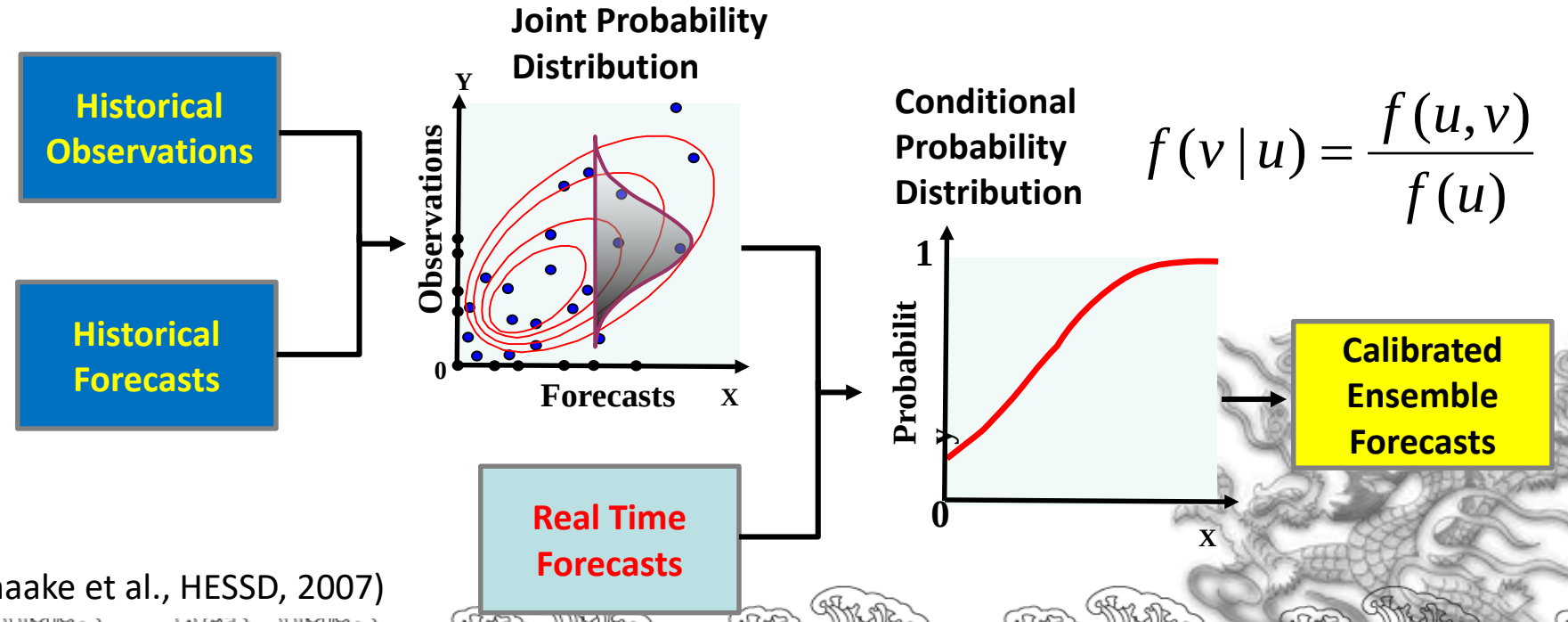
Types:

- Simple, unconditional methods: quantile mapping…
- Non-parametric methods:
  - Analog method
  - Kernel density methods (Ensemble dressing)…
- Parametric methods:
  - Condition distribution-based: BPO, EPP…
  - Regression-based methods: EMOS, logistic regression, quantile regression…

Ensemble Pre-Processor (EPP)

河海大学 HOHAI UNIVERSITY

- Ensemble Pre-Processor: assume the joint distribution of transformed observations and forecasts follow a bivariate Normal distribution, and obtain the conditional distribution given a certain forecast.
- Generate ensemble members from the conditional distribution and apply Schaake shuffle to preserve space-time dependency structure



(Schaake et al., HESSD, 2007)

Post-processing Methods for Hydrological Forecasts

河海大學 HOHAI UNIVERSITY

- "Post-Processor": Statistical models based on past samples of hydrologic forecast-observation relationships to produce bias corrected, space-time series of hydrologic variables of interest. It has the following functions:
  - **Correct spread** problems in hydrologic ensembles
  - **Remove systematic and random bias** in hydrologic forecasts
  - **Preserve space-time variability** and uncertainty structure

- As strong temporal autocorrelation exists in hydrological quantities, past recent observations or forecasts should also be included in statistical post-processing models

Regression-based Methods: General Linear Model Post-Processor

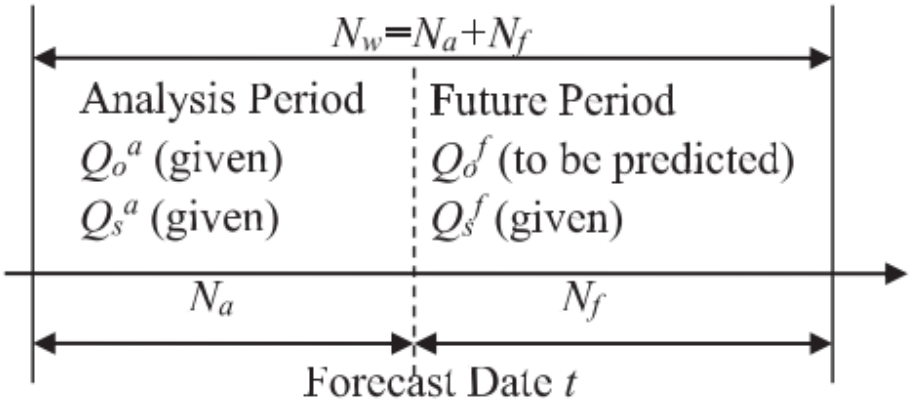
河海大学 HOHAI UNIVERSITY

- GLMPP: a linear regression model
- Advantages: include multiple recent past observations conveniently

$$\mathbf{Y} = \mathbf{A} \cdot \mathbf{X} + \mathbf{B} \cdot \mathbf{E}$$

$$\mathbf{Y} = \tilde{\mathbf{Q}}^{f}_{obs} \qquad \mathbf{X} = [\tilde{\mathbf{Q}}^{f}_{sim}, \tilde{\mathbf{Q}}^{a}_{sim}, \tilde{\mathbf{Q}}^{a}_{obs}]^{T}$$

observations *simulations* *past recent observations*



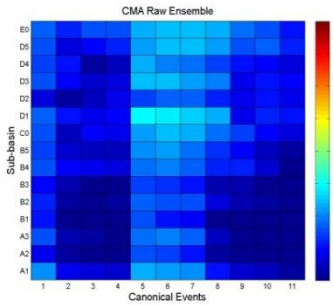
(a) A schematic of the forecast date, the analysis period and forecast period

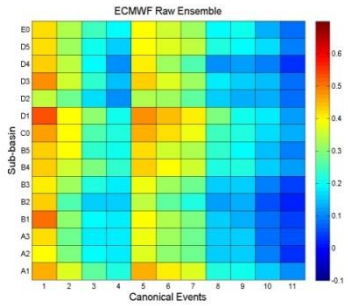
Ref. Zhao, et al. 2009; Ye et al., 2015

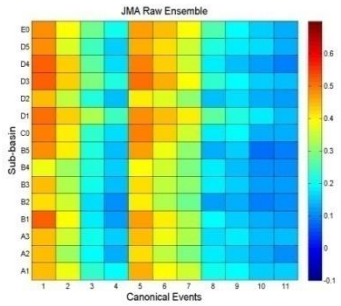
A Comparison of CRPSS Scores of the Raw Forecasts and Post-processed Forecasts

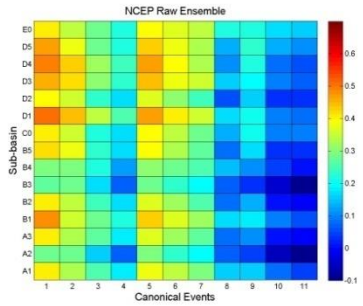
CRPSS of raw forecasts

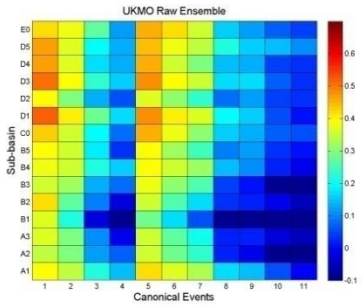
Tao, et al., J. Hydrol. 2014



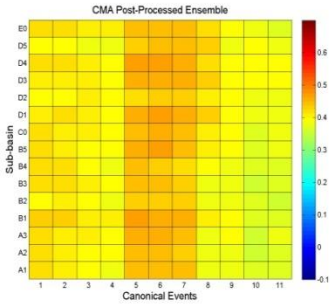


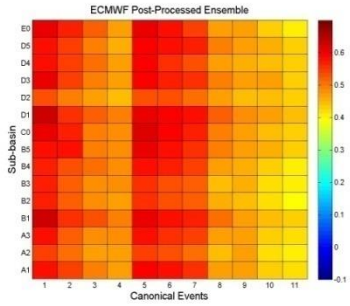


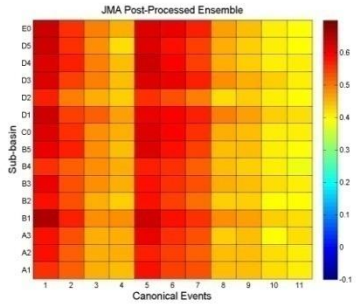


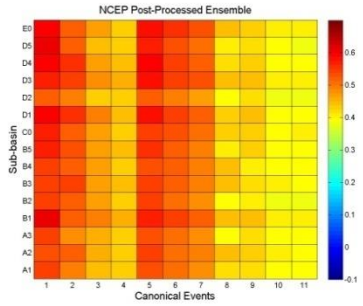


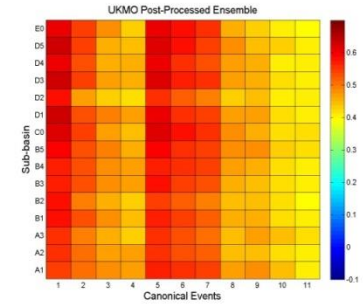
CRPSS of post-processed forecasts









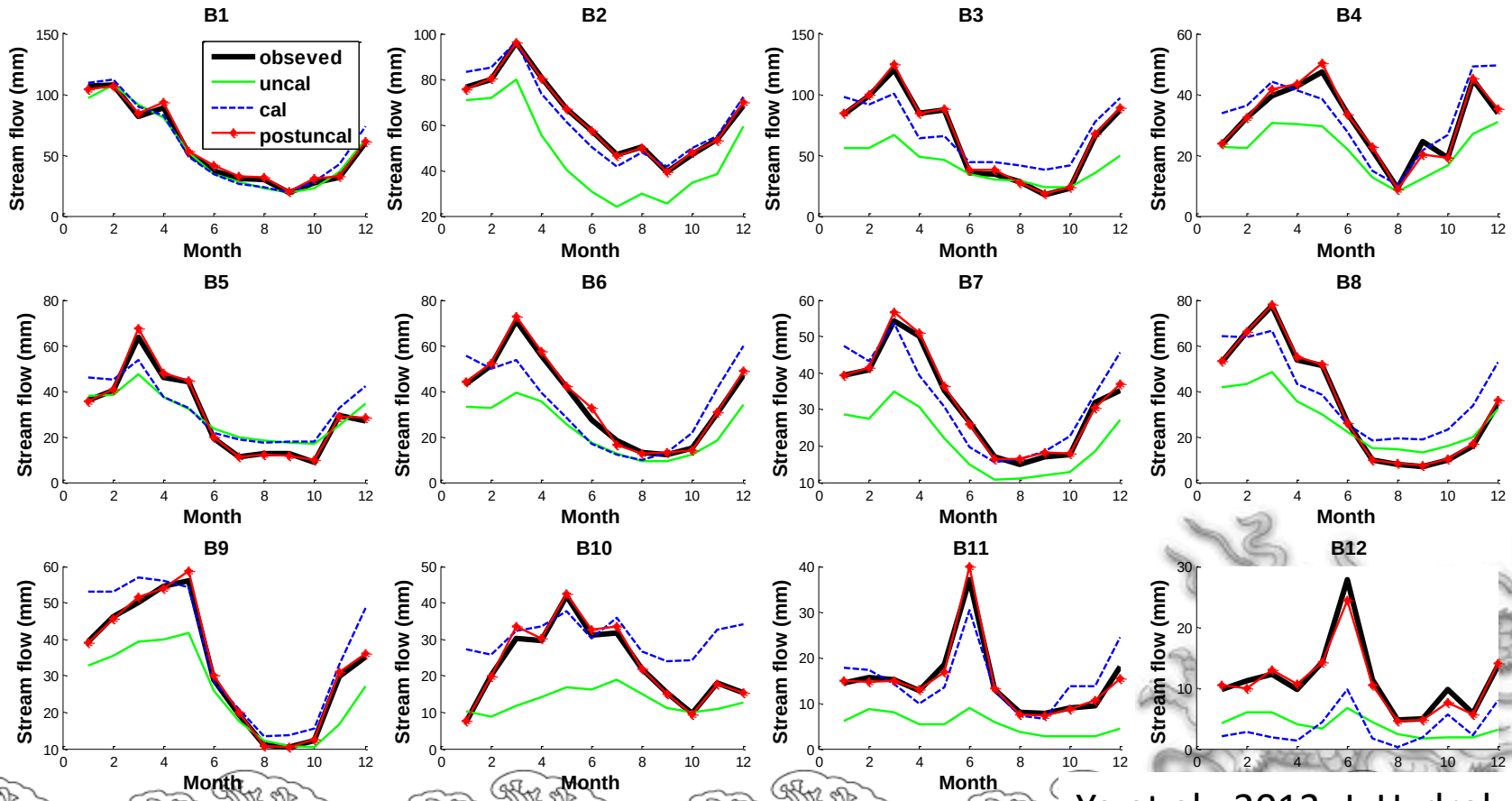


CMA ECMWF JMA NCEP UKMO

Period	1	2	3	4	5	6	7	8	9	10	11
Forecast days	Day 1	Day 2	Day 3	Day 4	1 – 2 days	1 – 3 days	1 – 4 days	5 – 6 days	7 – 9 days	5 – 9 days	1 – 9 days

A Comparison of Streamflow Forecasts Before / After Post-processing

Uniqueness of Hydrological Post-processing: Because of strong temporal autocorrelation in hydrological quantities, past recent observations or forecasts must be included in any statistical post-processing model for hydrological quantities



Ye et al., 2013, J. Hydrol

河海大学
HOHAI UNIVERSITY

Advanced Review

A review on statistical postprocessing methods for hydrometeorological ensemble forecasting

Wentao Li,[1,2] Qingyun Duan,[1,2]* Chiyuan Miao,[1,2] Aizhong Ye,[1,2] Wei Gong[1,2] and Zhenhua Di[1,2]

Computer simulation models have been widely used to generate hydrometeorological forecasts. As the raw forecasts contain uncertainties arising from various sources, including model inputs and outputs, model initial and boundary conditions, model structure, and model parameters, it is necessary to apply statistical postprocessing methods to quantify and reduce those uncertainties. Different postprocessing methods have been developed for meteorological forecasts (e.g., precipitation) and for hydrological forecasts (e.g., streamflow) due to their different statistical properties. In this paper, we conduct a comprehensive review of the commonly used statistical postprocessing methods for both meteorological and hydrological forecasts. Moreover, methods to generate ensemble members that maintain the observed spatiotemporal and intervariable dependency are reviewed. Finally, some perspectives on the further development of statistical postprocessing methods for hydrometeorological ensemble forecasting are provided. © 2017 Wiley Periodicals, Inc.

How to cite this article:
WIREs Water 2017, e1246. doi: 10.1002/wat2.1246

Confronting Model State Uncertainty

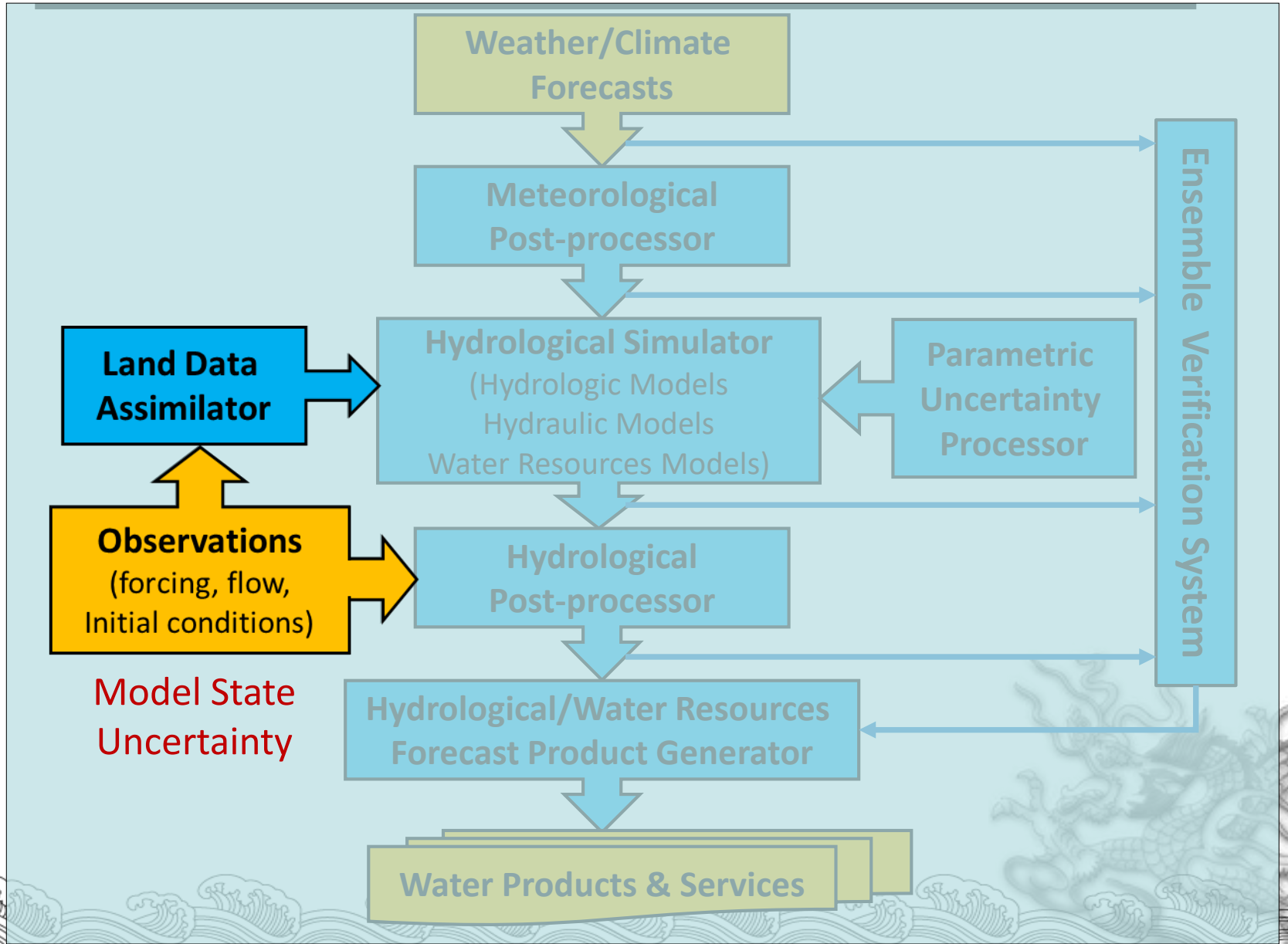
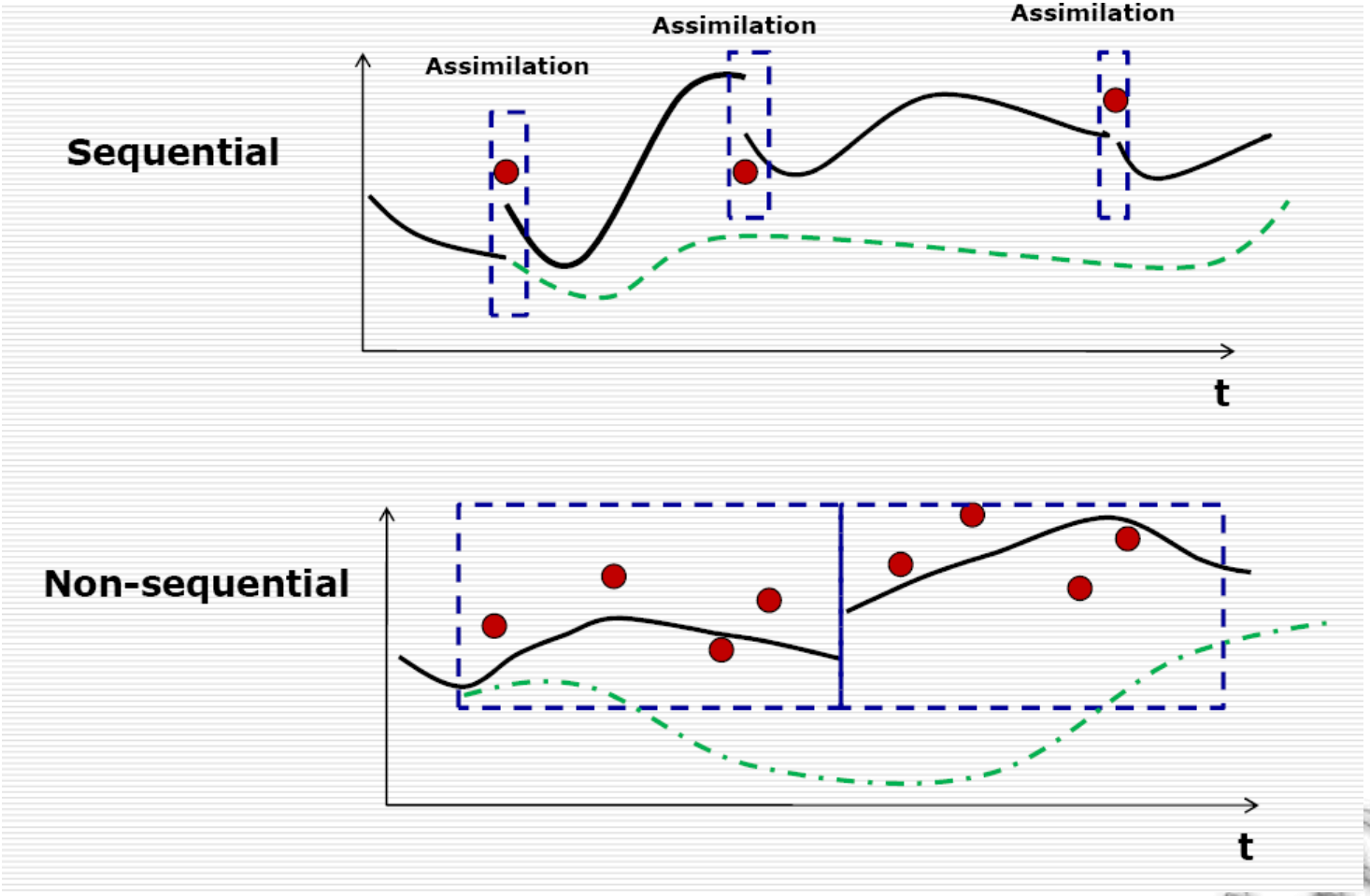


Illustration of Data Assimilation



Filter

Smoothers

Data assimilation aims to improve model simulation by merging model state variables with corresponding observations

Examples of DA on Hydrologic Simulations

河海大學 HOHAI UNIVERSITY

Sun et al., J Hydrol., 2016

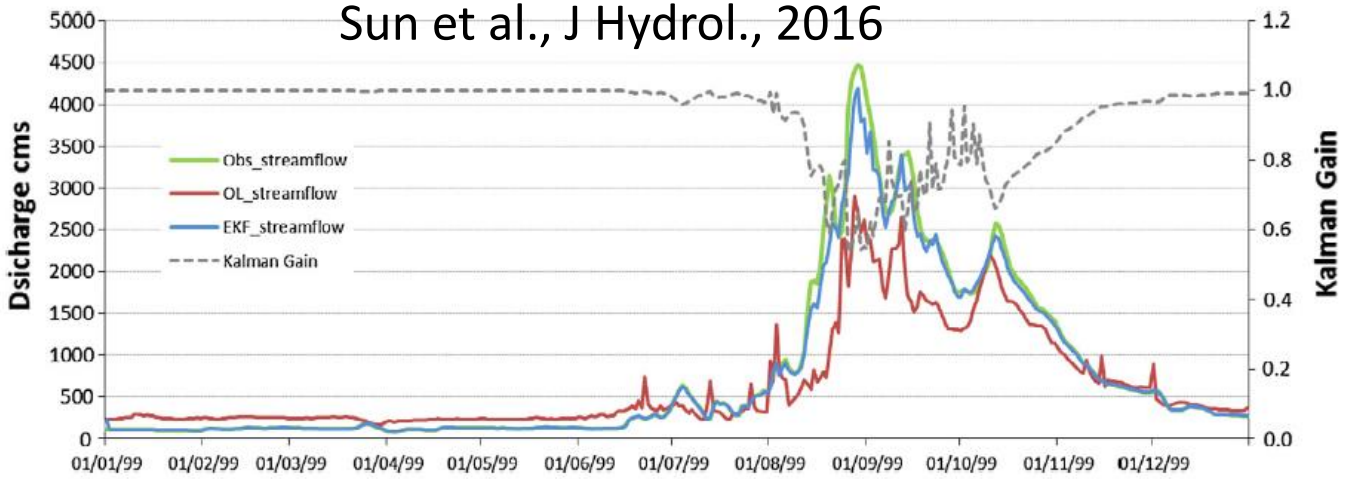
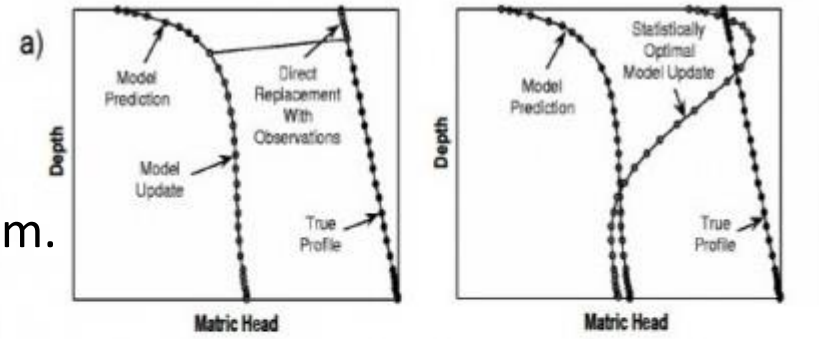
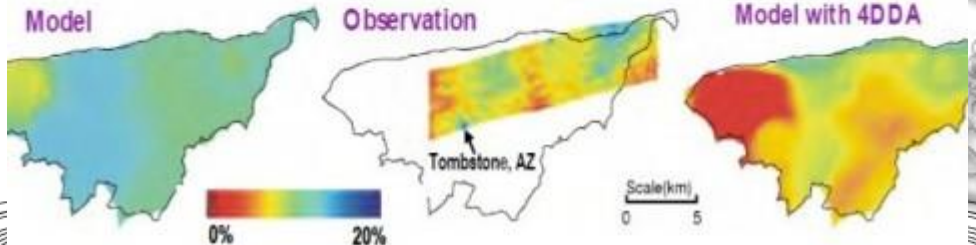


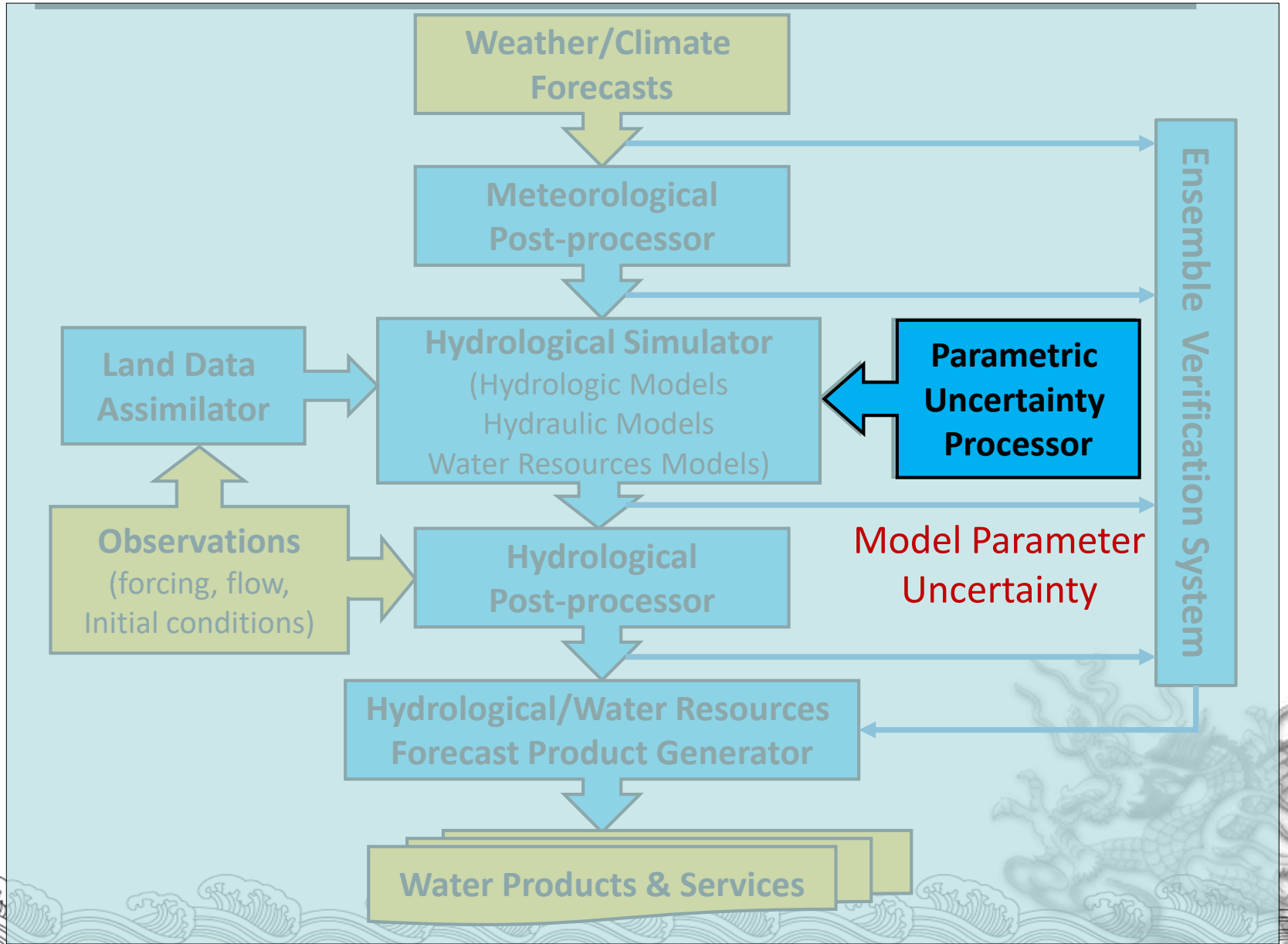
Fig. 4. Bakel station 1999 streamflow EKF assimilation results compared to observation (Obs) and Open loop (OL).

P.R. Houser, prhouser.com.



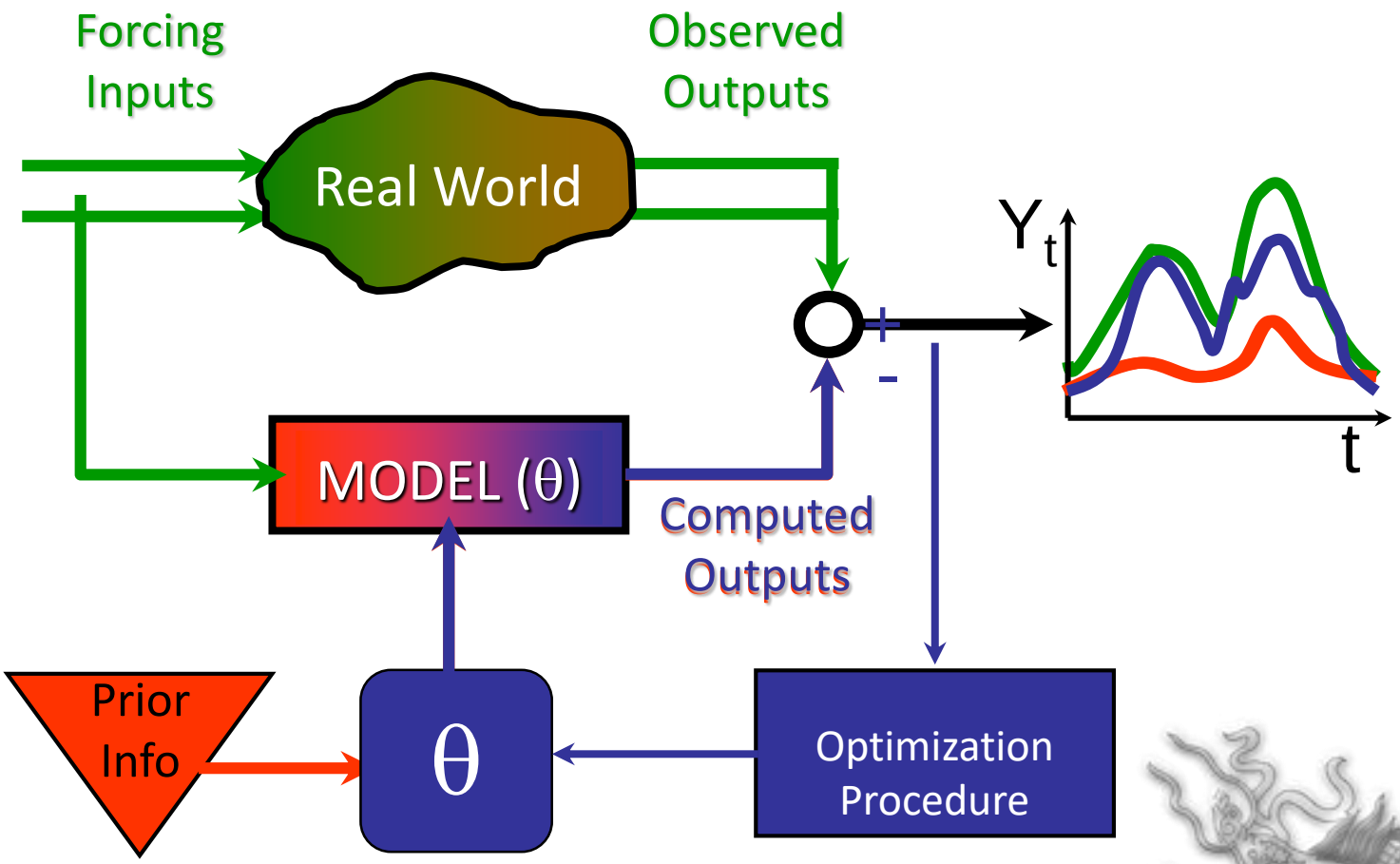


Confronting Parametric Uncertainty



Confronting Parametric Uncertainty - Model Calibration

河海大學 HOHAI UNIVERSITY



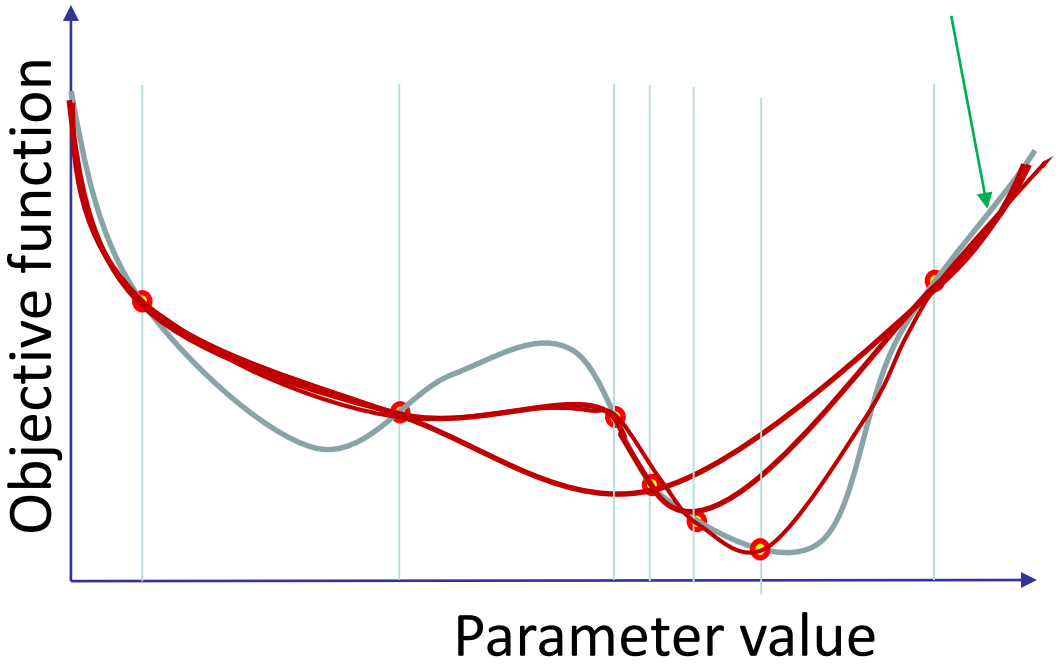
"**Calibration**: constraining the model simulations to be consistent with observations by tuning model parameters"

Global Search Algorithms

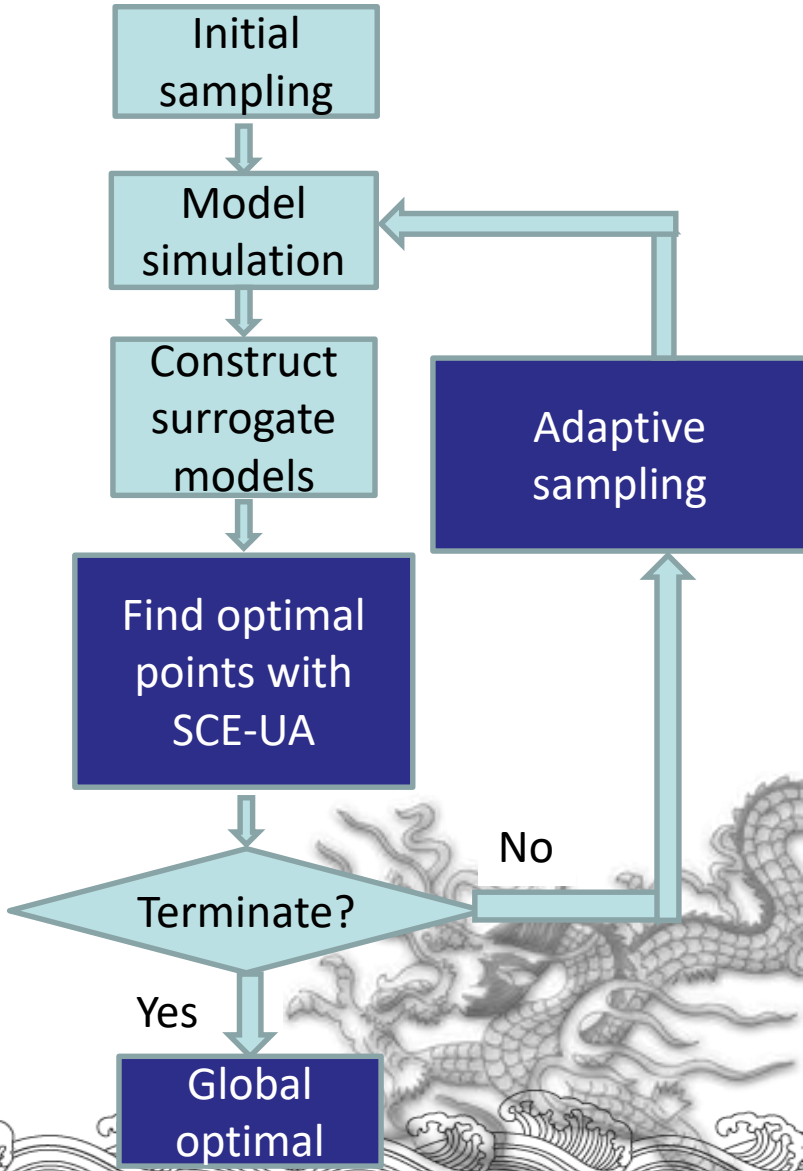
- **Evolutionary algorithms:**
  - Genetic algorithm (GA), Simulated annealing (SA), Particle swarming (PS), Frog-leaping (FL), …
- **Heuristic algorithms:**
  - Dynamically dimensioned search algorithm (DDS), Robust Gauss-Newton (RGN), …
- **Surrogate modeling based optimization methods:**
  - Optimization by radial basis function interpolation in trust-regions (ORBIT), Multiple surrogate efficient global optimization (MSEGO), Adaptive surrogate modeling-based optimization (ASMO), …

ASMO: Adaptive Surrogate Modeling-based Optimization

河海大學 HOHAI UNIVERSITY

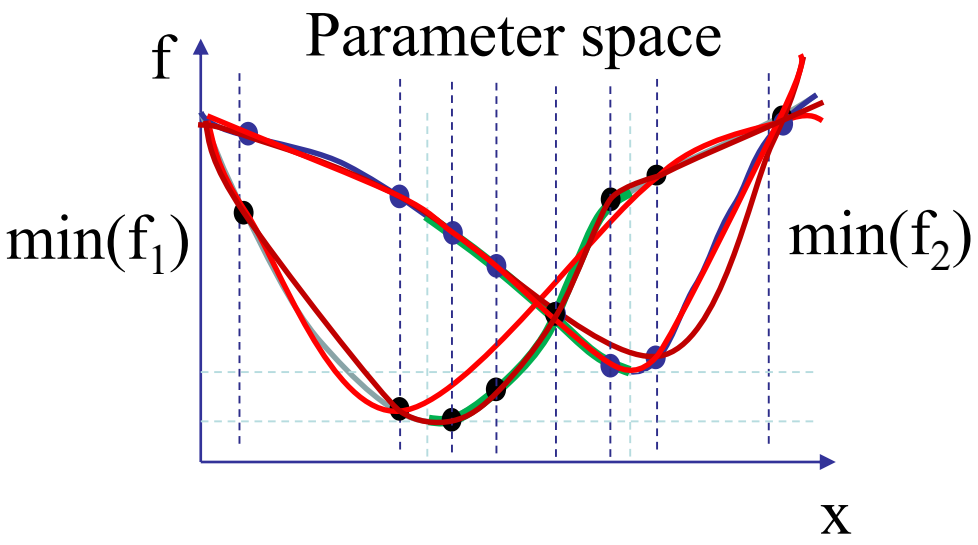


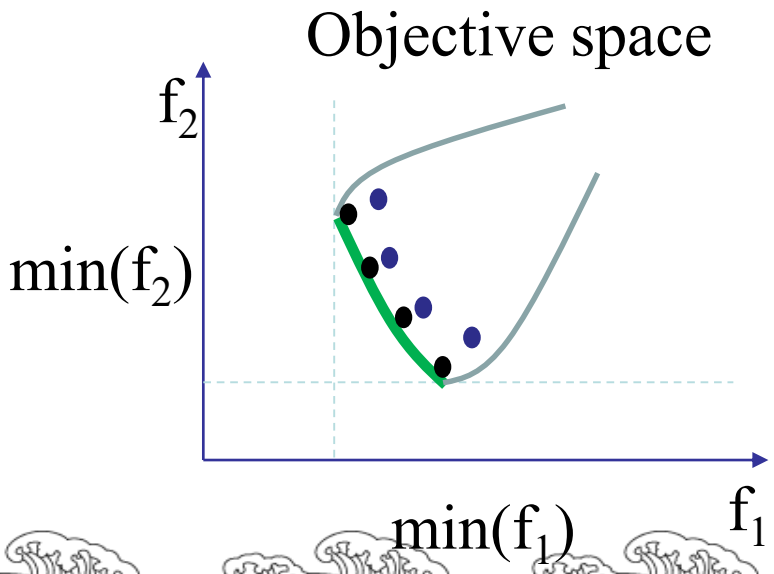
[Chen Wang et.al. 2013, EMS]

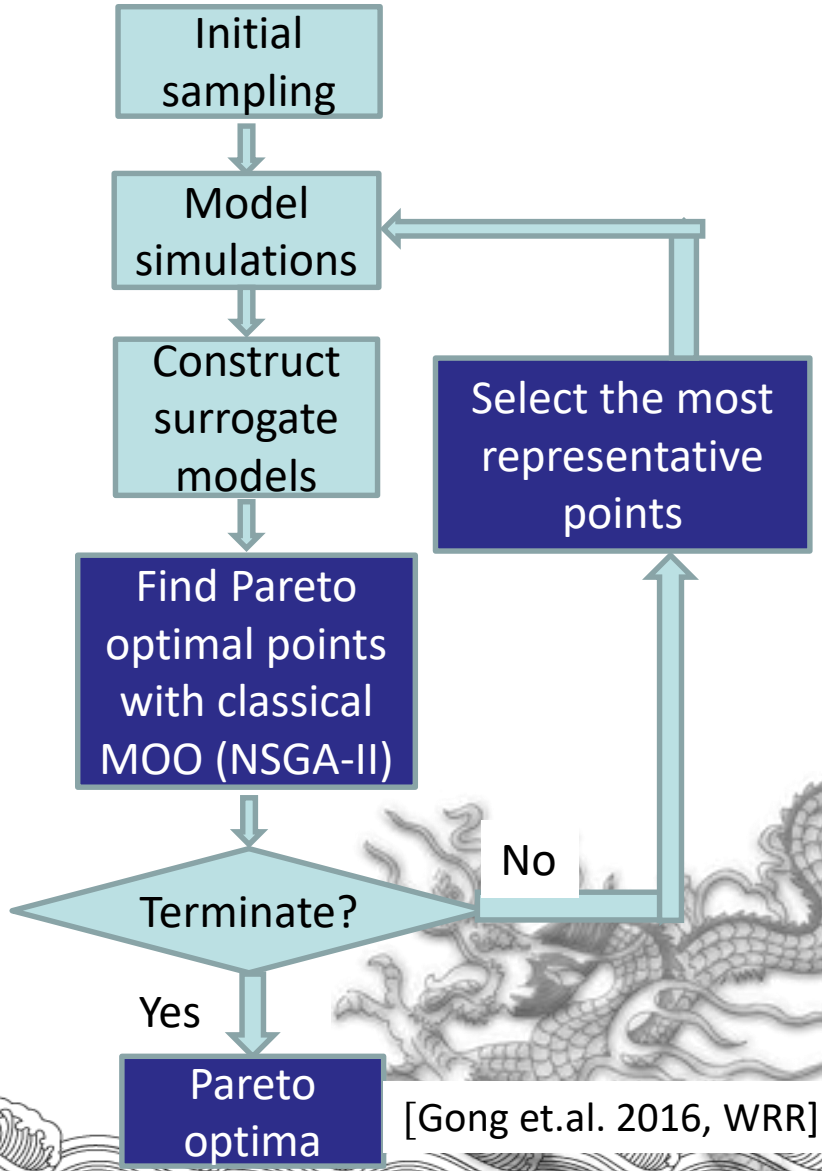


MO-ASMO: Multi-Objective ASMO

河海大學 HOHAI UNIVERSITY

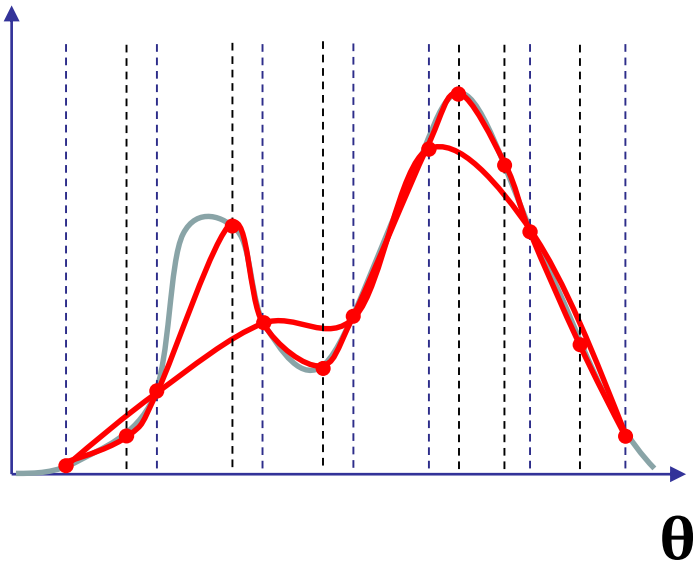


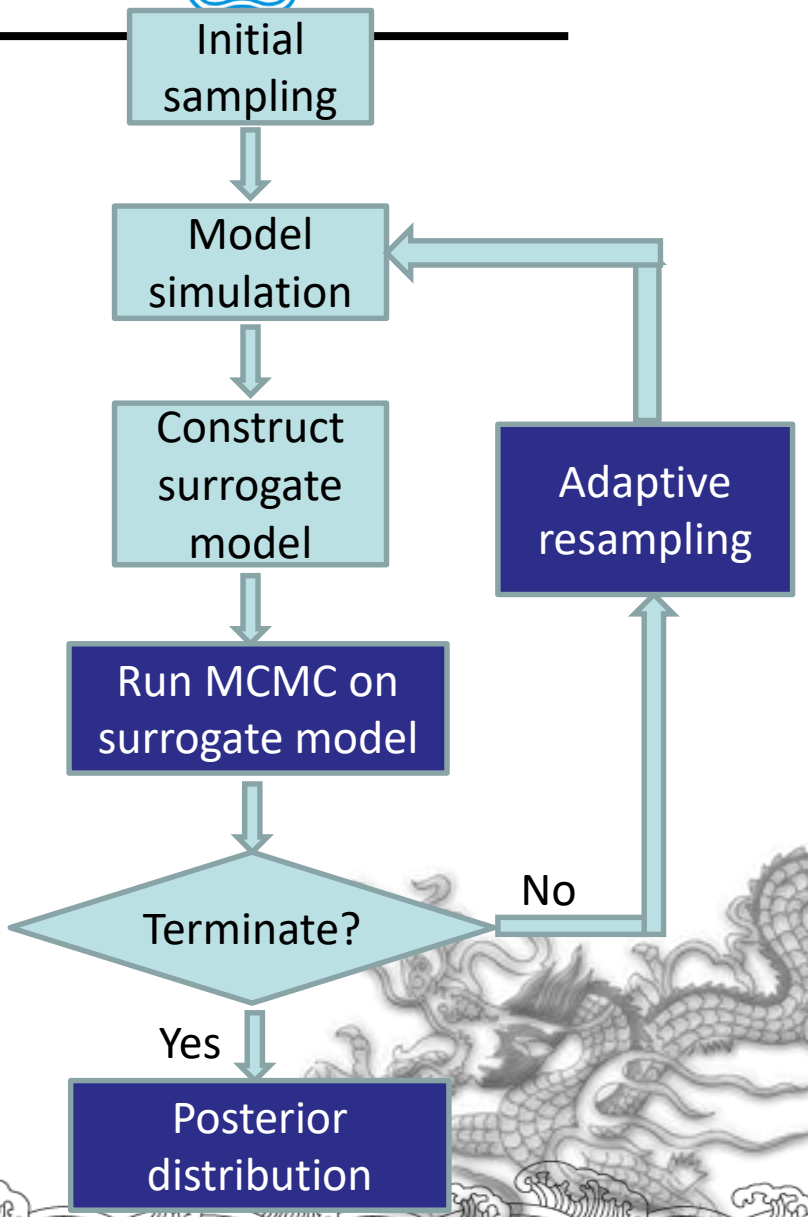




[Gong et.al. 2016, WRR]

ASMO-PODE: Parameter Optimization and Distribution Estimation





[Gong & Duan 2017, EMS]

Key Testing Results with ASMO, MO-ASMO and ASMO-PODE

河海大學
HOHAI UNIVERSITY

- ASMO is as effective as SCE, but more efficient:
  - **~200 vs ~1000**
- MO-ASMO is as effective as NSGA-II, but much more efficient:
  - **~800 vs ~25000**
- ASMO-PODE is as effective as MCMC Metropolis, but much more efficient:
  - **~2000 vs ~50000**

Uncertainty Quantification Python Laboratory (UQ-PyL)

河海大學 HOHAI UNIVERSITY

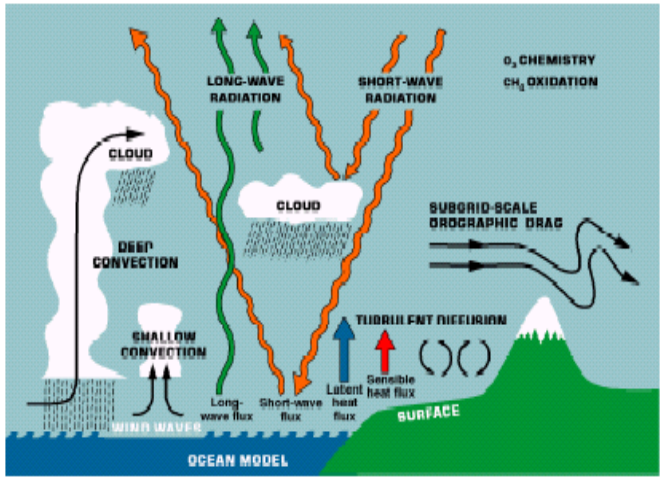
- A new, general-purpose, cross-platform UQ framework with a GUI
- Made of several components that perform various functions, including
  - *Design of Experiments*
  - *Statistical Analysis*
  - *Sensitivity Analysis*
  - *Surrogate Modeling*
  - *Parameter Optimization*
- Suitable for parametric uncertainty analysis of any computer simulation models

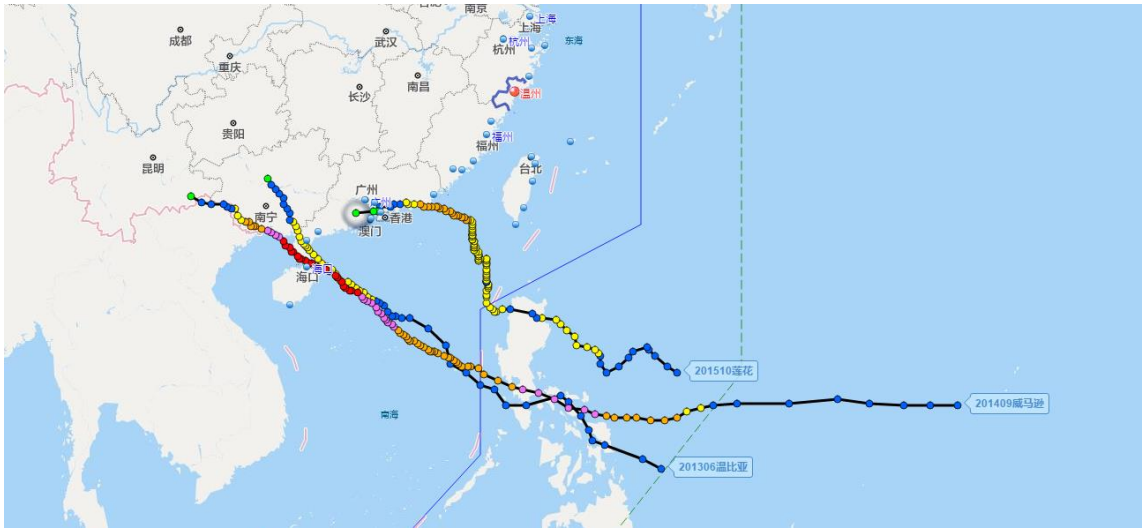
http://uq-pyl.com

(see Wang et al., EMS, 2016)

Optimization of the WRF Model Parameters

河海大学 HOHAI UNIVERSITY





Outer Grid: 18km： 211×178
Inner Grid: 6km: 178×190
Vertical Layers： 38
Model Version： WRFV3.6.1

3 Typhoon Cases：
#1306： 2013-06-30_18:00:00—2013-07-04_00:00:00
#1409： 2014-07-17_18:00:00—2014-07-21_00:00:00
#1510: 2015-07-05_18:00:00—2015-07-09_00:00:00
Forecast Lead Time: 78-hr，
First 6-hr for spinup， last 3 day used for analysis

Forcing Data： NCEP Reanalysis（1° x 1° ）
Calibration Data：
Precipitation: CMA CMORPH hourly（0.1° x 0.1° ）data
Wind speed: CMA Shanghai Typhoon Institute, Northwest Pacific typhoon dataset

WRF Model Parameters To Be Examined

河海大学 HOHAI UNIVERSITY

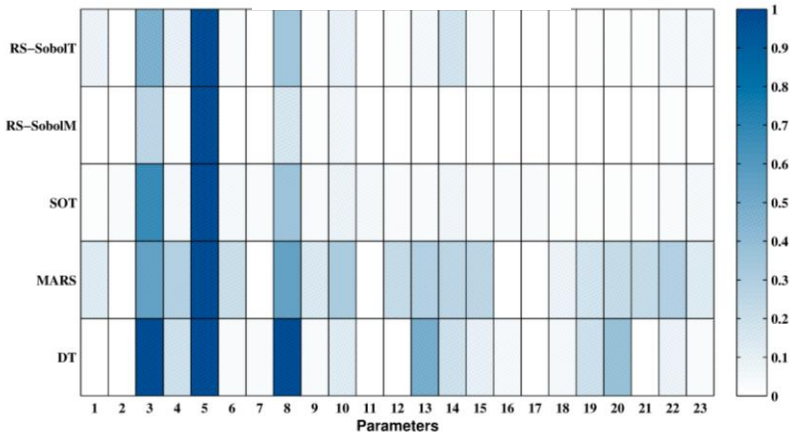
number	scheme	name	Default	range	description
1	Surface layer (module_sf_sfclay.F)	xka	0.000024	[0.000012 0.00005]	The parameter for heat/moisture exchange coefficient
2		CZO	0.0185	[0.01 0.037]	The coefficient for coverting wind speed to roughness length over water
3	Cumulus (module_cu_kfeta.F)	pd	0	[-1 1]	The coefficient related to downdraft mass flux rate
4		pe	0	[-1 1]	The coefficient related to entrainment mass flux rate
5		ph	150	[50 350]	Starting height of downdraft above USL
6		TIMEC	2700	[1800 3600]	Compute convective time scale for convection
7		TKEMAX	5	[3 12]	the maximum turbulent kinetic energy (TKE) value between the level of free convection (LFC)and lifting condensation level (LCL)
8	Microphysics (module_mp_wsm6.F)	ice_stokes_fac	14900	[8000 30000]	Scaling factor applied to ice fall velocity
9		n0r	8000000	[5000000 12000000]	Intercept parameter rain
10		dimax	0.0005	[0.0003 0.0008]	The limited maximum value for the cloud-ice diameter
11		peaut	0.55	[0.35 0.85]	Collection efficiency from cloud to rain auto conversion
12	short wave radiation (module_ra_sw.F)	cssca	0.00001	[0.000005 0.00002]	Scattering tuning parameter in clear sky
13		Beta_p	0.4	[0.2 0.8]	Aerosol scattering tuning parameter
14	Longwave (module_ra_rrtm.F)	Secang	1.66	[1.55 1.75]	Diffusivity angle
15	Land surface (module_sf_noahlsm.F)	hksati	0	[-1 1]	hydraulic conductivity at saturation
16		porsl	0	[-1 1]	fraction of soil that is voids
17		phi0	0	[-1 1]	minimum soil suction
18		bsw	0	[-1 1]	Clapp and hornbereger "b" parameter
19	Planetary Boundary Layer (module_bl_ysu.F)	Brcr_sbrob	0.3	[0.15 0.6]	Critical Richardson number for boundary layer of water
20		Brcr_sb	0.25	[0.125 0.5]	Critical Richardson number for boundary layer of land
21		pfac	2	[1 3]	Profile shape exponent for calculating the momentum diffusivity coefficient
22		bfac	6.8	[3.4 13.6]	Coefficient for prandtl number at the top of the surface laer
23		sm	15.9	[12 20]	Countergradient proportional coefficient of non-local flux of momentum moh 2002

Sensitivity Analysis Results

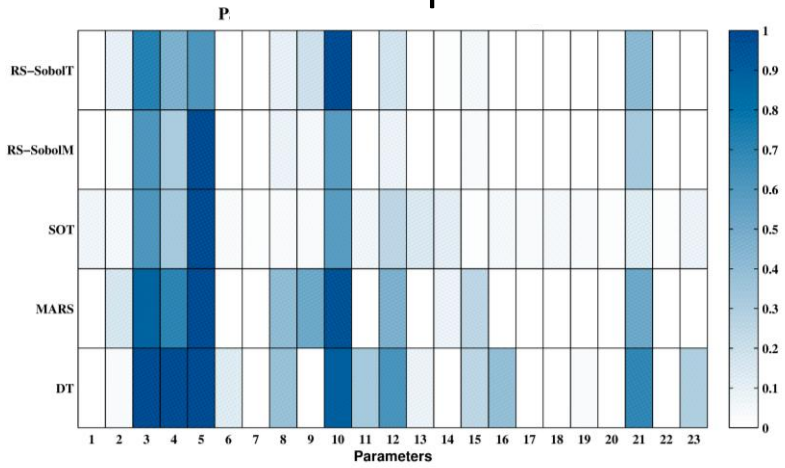
河海大學 HOHAI UNIVERSITY

Sensitivity Analysis Methods: DT, MARS, SOT, RSSOBOL(main and total effects)
Objective Functions: Threat Score (TS)， Root Mean Square Error (RMSE)

Precipitation



Wind Speed



Sensitive Parameters Identified: P3, P4, P5, P8, P10, P12, P21

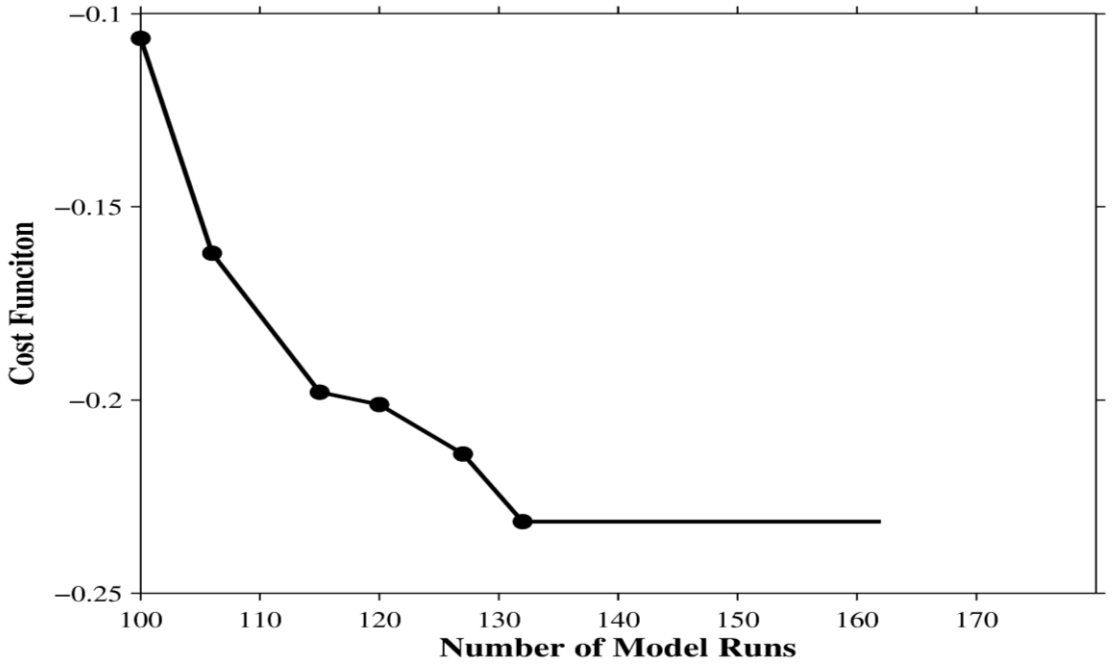
The Optimization Results

河海大學 HOHAI UNIVERSITY

Calibration Criterion:

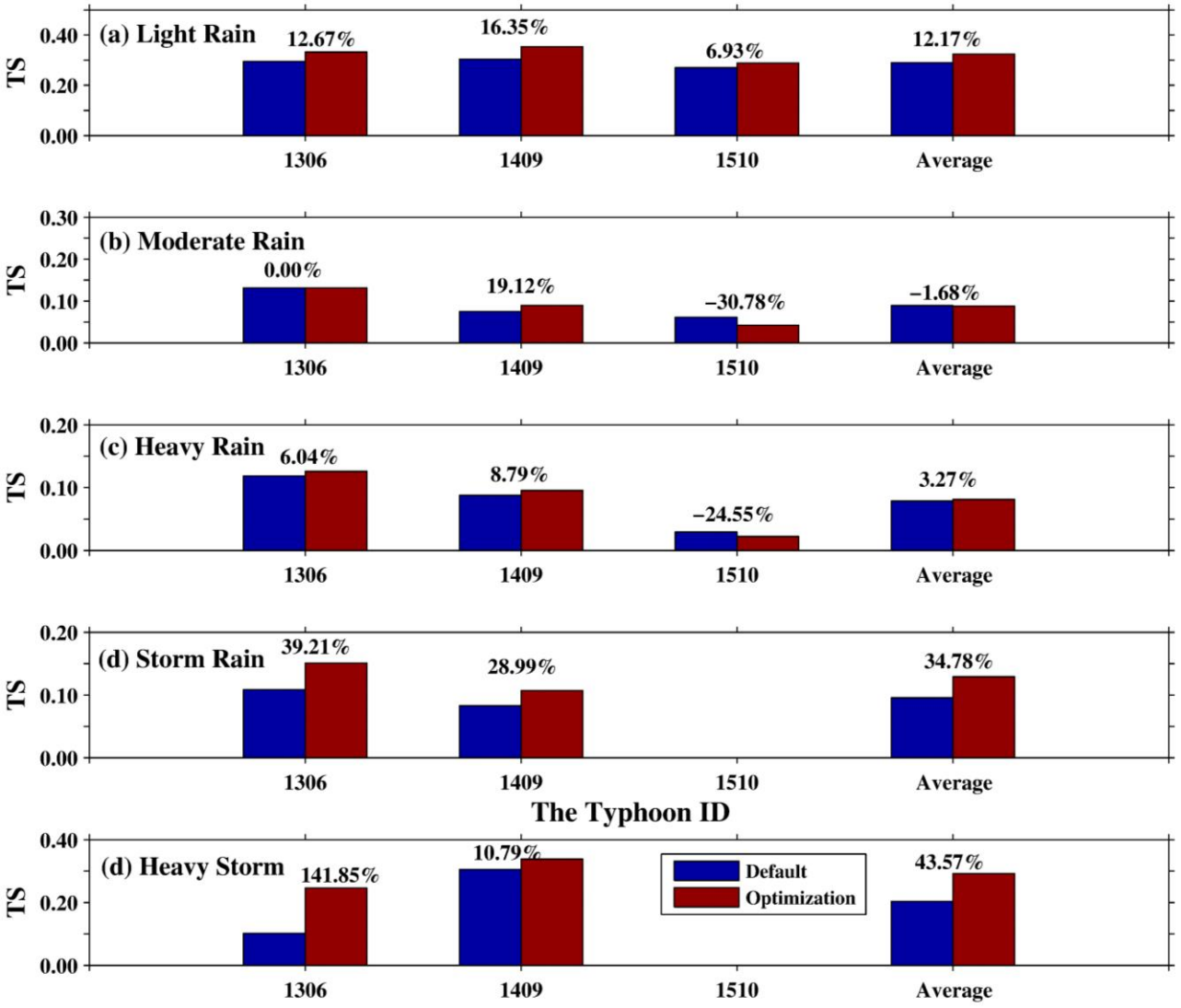
$$F(\theta) = \frac{1}{2} \bullet \left(\frac{RMSE(\theta)}{RMSE(\theta)_{def}} - \frac{1}{5} \sum_{j=1}^{5} \frac{TS_i(\theta)}{TS_{def}(\theta)} \right)$$

i=1: Light rain
i=2: Moderate rain
i=3: heavy rain
i=4: Storm rain
i=5: Heavy Storm



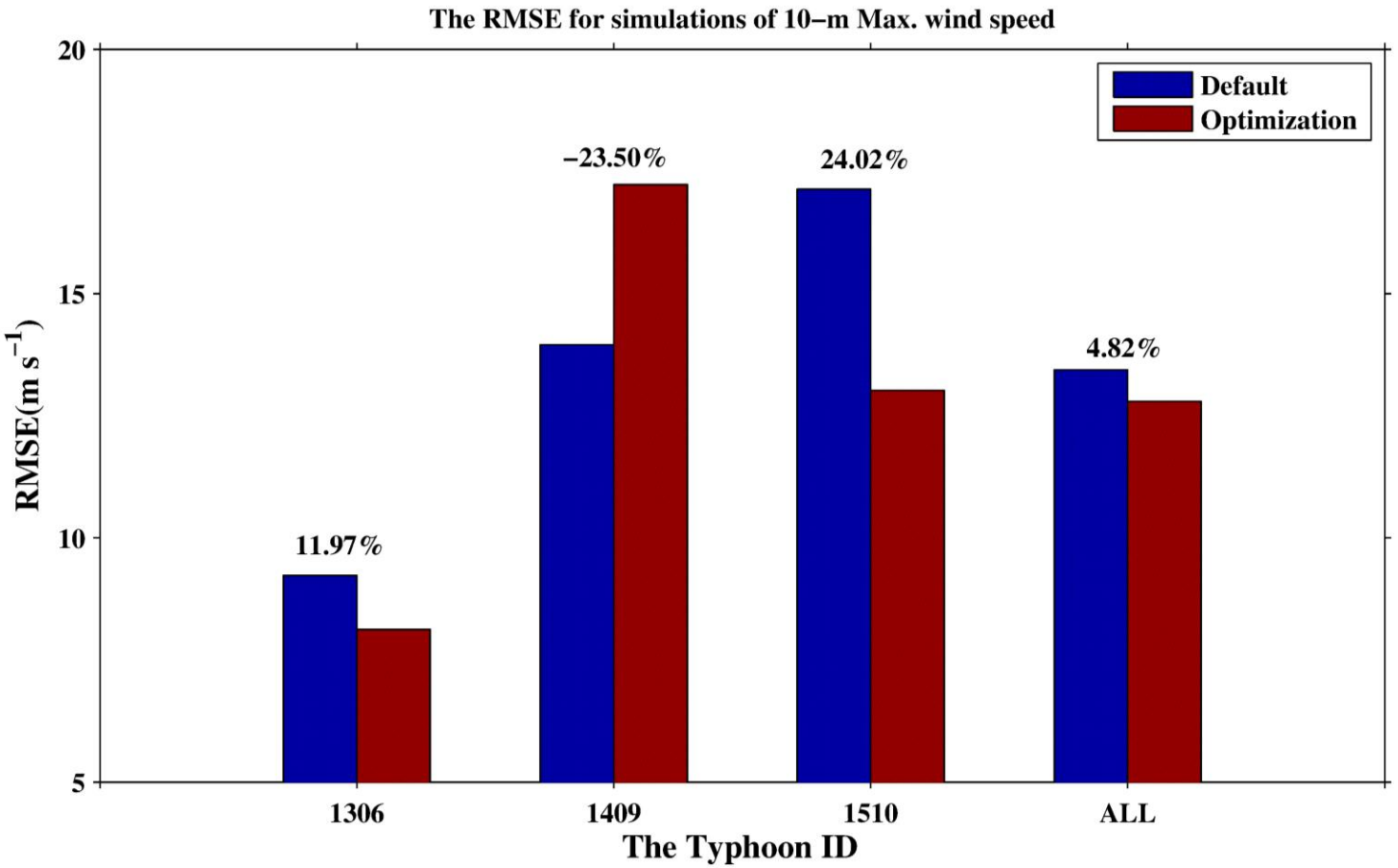
Comparison of Precipitation Forecast Skills of Optimized Forecasts and Default Forecasts

河海大學 HOHAI UNIVERSITY



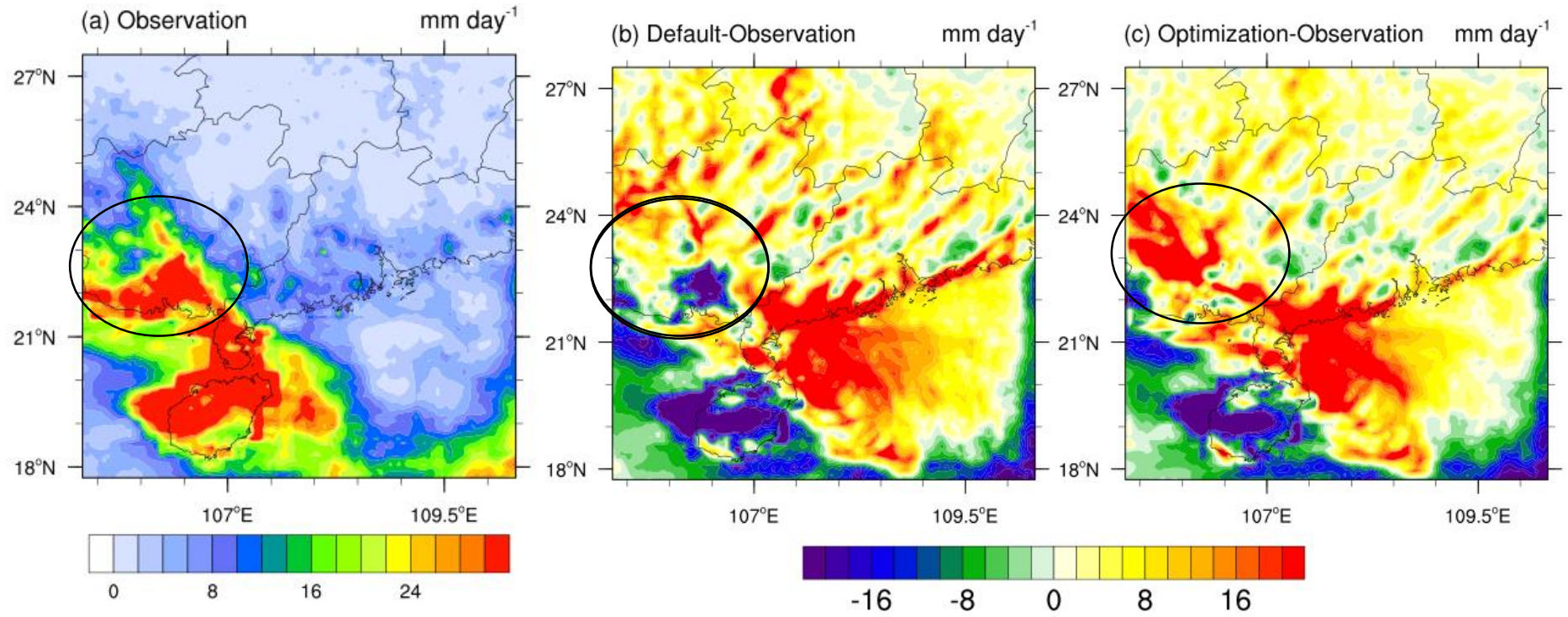
Comparison of Wind Forecast Skills of Optimized Forecasts and Default Forecasts

河海大學
HOHAI UNIVERSITY



Spatial Comparison of Cumulative Precipitation Forecasts Against Observations

河海大學
HOHAI UNIVERSITY



AUTOMATIC MODEL CALIBRATION

A New Way to Improve Numerical Weather Forecasting

Q. Duan, Z. Di, J. Quan, C. Wang, W. Gong, Y. Gan,
A. Ye, C. Miao, S. Miao, X. Liang, and S. Fan

Automatic model calibration improves numerical weather forecasting by tuning the numerical weather prediction model parameters to match model predictions with observations.

Weather forecasting skill has been improving steadily over the years. The improvement is due mostly to advances in the representation of physical processes by numerical weather prediction (NWP) models, observational systems, analyses techniques and forecasting methods (e.g., data assimilation, statistical postprocessing, and ensemble forecasting), new computational capability, and effective communications and training. There is an area that has received less attention so far but can bring significant improvement to weather forecasting—the calibration of NWP models. Model calibration refers to a process in which model parameters are tuned to match model predictions with corresponding observations. This process may be done using a manual, unsystematic "trial and error" approach, or using an optimization algorithm to tune model parameters automatically to minimize the difference between model predictions and observations (Duan et al. 2006). Automatic model calibration is a common practice in many fields including hydrology, biology, communications, and finance. It focuses on reducing errors resulting from the specification of model parameters. This is different from other popular approaches. For example, a common approach is to reduce model structural error by developing better physical parameterization schemes (Stensrud 2007). Multimodel ensemble approaches have been employed to account for model structural uncertainty (Krishnamurti et al. 1999). Data assimilation methods are commonly used to reduce errors in model initial conditions (Kalnay 2003).

There are several reasons that the automatic calibration of NWP models is not practiced as widely as in other fields. First, a typical NWP model has many

AFFILIATIONS: Duan, Di, Quan, Wang, Gong, Ye, and C. Miao—State Key Laboratory of Earth Surface Processes and Resource Ecology, Faculty of Geographical Science, Beijing Normal University, Beijing, China; Gan and Liang—State Key Laboratory of Severe Weather, Chinese Academy of Meteorological Sciences, Beijing, China; S. Miao, and Fan—Institute of Urban Meteorology, China Meteorological Administration, Beijing, China
CORRESPONDING AUTHOR: Qingyun Duan, qyduan@bnu.edu.cn

The abstract for this article can be found in this issue, following the table of contents.
DOI:10.1175/BAMS-D-15-00104.1

A supplement to this article is available online (10.1175/BAMS-D-15-00104.2)

In final form 24 August 2016
©2017 American Meteorological Society

河海大學
HOHAI UNIVERSITY

Summary and Discussion

河海大學
HOHAI UNIVERSITY

- Different ensemble forecasting methods reviewed:
  - Post-processing of model outputs
  - Land data assimilation
  - Parameter optimization

- Raw forecasts can be improved tremendously by using different ensemble forecasting methods

- Further challenges:
  - How do we consider all sources of uncertainties in an integrated manner?
    - How do we attribute uncertainties?
    - How different uncertainties interact?
  - How to demonstrate the usefulness of ensemble forecasting in water resources applications?

So far as the laws of mathematics refer to reality, they are not certain.

And so far as they are certain, they do not refer to reality.

Albert Einstein
Geometry & Experience

Questions ?