BAYESIAN THEORY of ENSEMBLE FORECASTING

Roman Krzysztofowicz

University of Virginia

USA

Presented at the

HEPEX Meeting on

Uncertainty Post-Processing for Streamflow Forecasts

Delft, The Netherlands

23 – 25 June 2008

Acknowledgments:

National Science Foundation, Grant No. ATM–0641572

Henry D. Herr, Coire J. Maranzano

BAYESIAN THEORY

(Krzysztofowicz, 1999)

- Derived from : Total Probability Law
  Bayes Theorem

- Applicable to any deterministic hydrologic model
- Fuses optimally information (=> informativeness)
- Quantifies total uncertainty
- Outputs probabilistic forecast (=> calibration)

- Basis for: system decomposition
  mathematical models
  computation methods

1. Bayesian Forecasting System (BFS)

   Analytic – Numerical
   For headwater basin

2. Ensemble BFS (EBFS)

   Ensemble / Monte-Carlo simulation
   For any basin

DECOMPOSITION OF UNCERTAINTY

OPERATIONAL UNCERTAINTY (exterior to forecasting theory)

- Missing data
- Processing errors
- .
- .
- .

✓ INPUT UNCERTAINTY (dominant uncertainty)

- Future precipitation

✓ HYDROLOGIC UNCERTAINTY (all other uncertainties)

- Models
- Parameters
- Initial conditions
- Deterministic inputs
- .
- .
- .

TOTAL: INPUT and HYDROLOGIC

BAYESIAN THEORY

(Krzysztofowicz, 1999)

Decomposition Theorem

$$\left[\begin{array}{c}\text{Total}\\ \text{Uncertainty}\end{array}\right] = \int \left[\begin{array}{c}\text{Hydrologic}\\ \text{Uncertainty}\end{array}\right] \cdot \left[\begin{array}{c}\text{Input}\\ \text{Uncertainty}\end{array}\right]$$

↑ Ensemble RSF — ↑ [Bayesian HUP] — ↑ Ensemble QPF

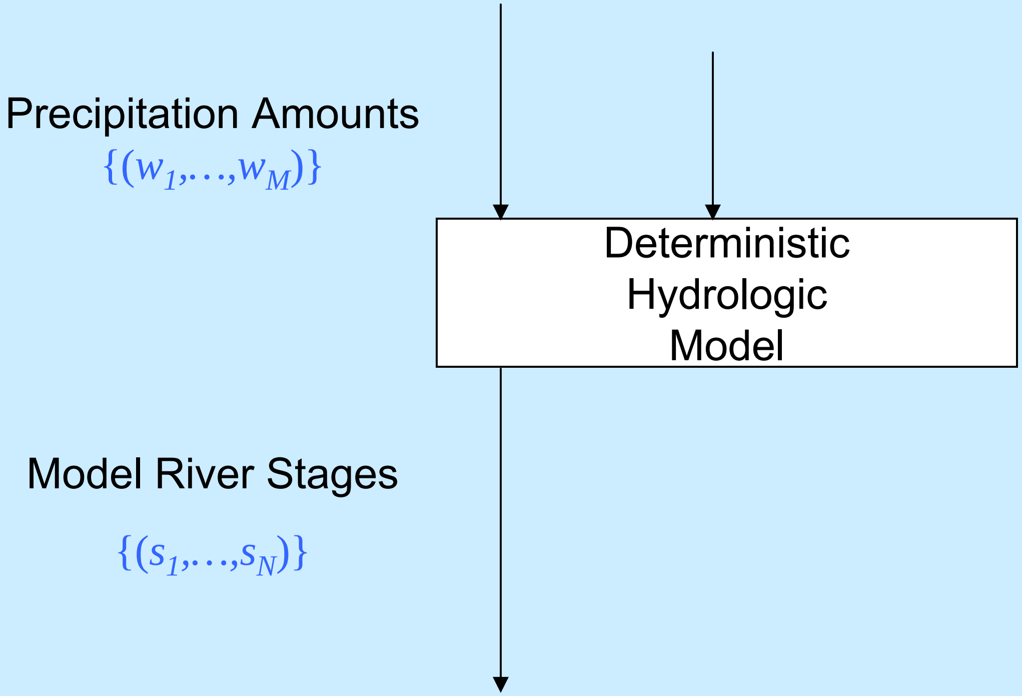
Calibration Theorem

In a *Bayesian Forecasting System* (BFS)
with any deterministic hydrologic model,
if the input distribution is *well calibrated*,
then the output distribution is *well calibrated*.

Inferred Requirements on Input

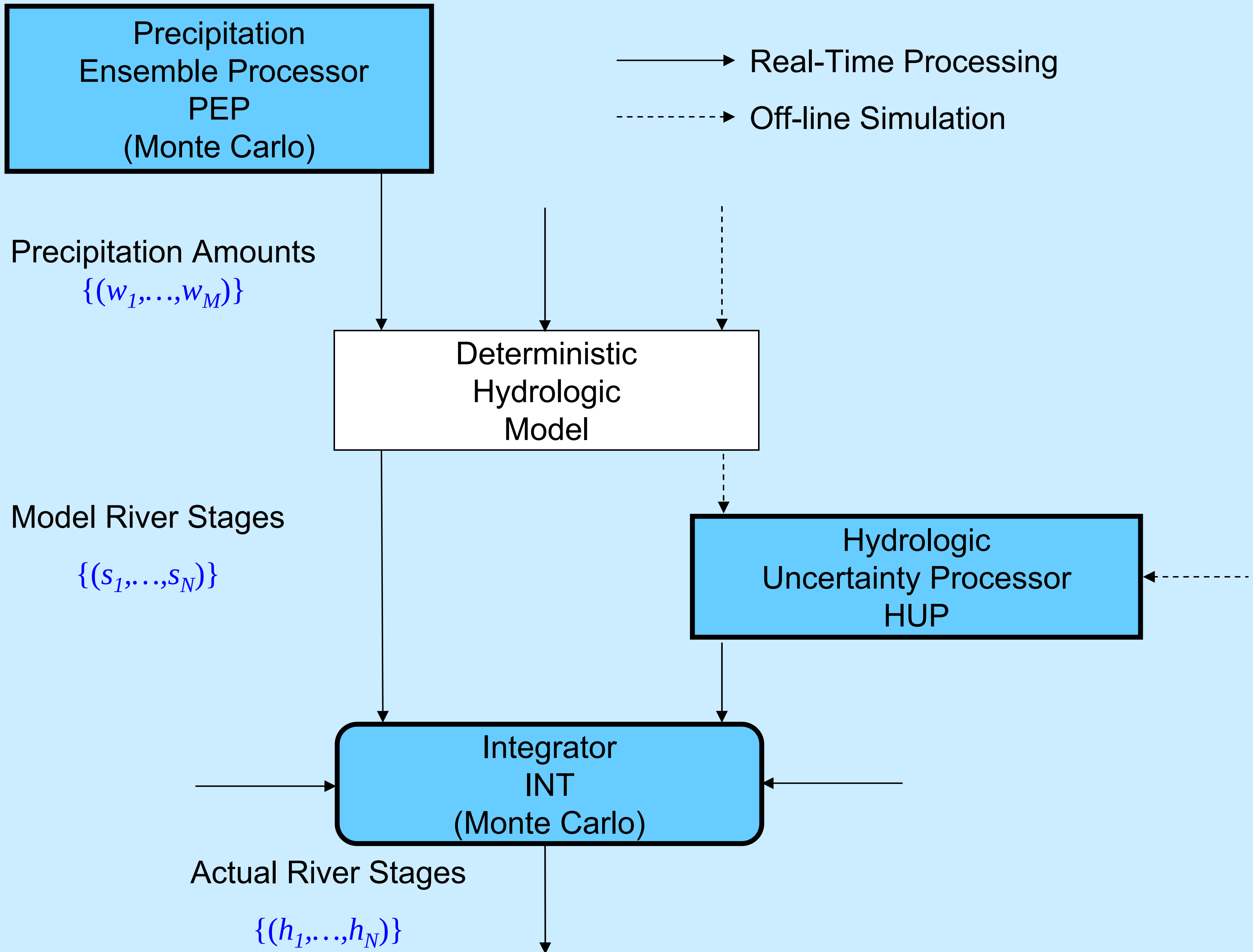
1. Ensemble QPF must be *well calibrated*.
2. Ensemble size must be *large enough*.

ENSEMBLE (NAIVE) SYSTEM

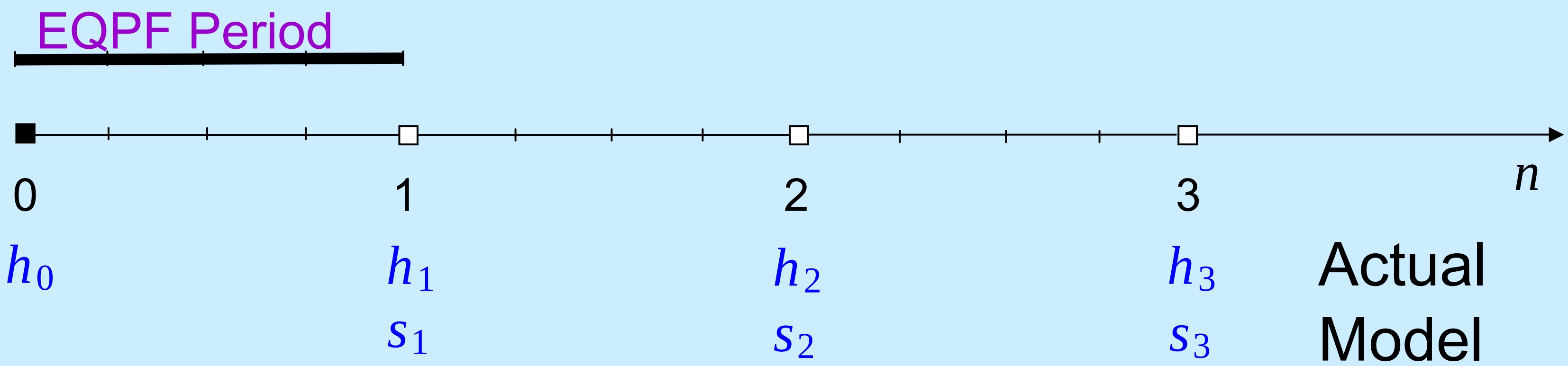


How to implement Bayesian Theory?

ENSEMBLE BAYESIAN SYSTEM



NOTATION



v – indicator of precip. occurrence during EQPF period:

$v = 1$ yes, $v = 0$ no

n – index of times

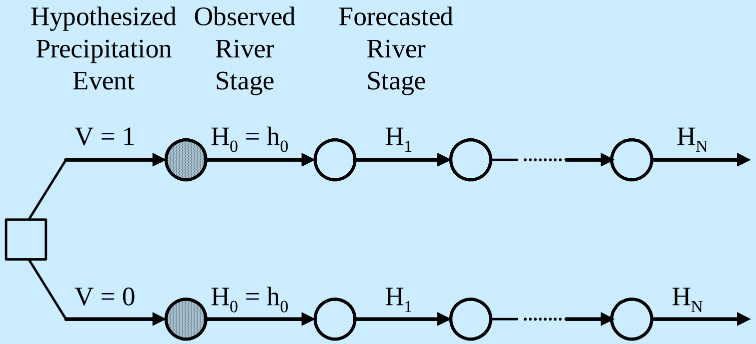
h_n – actual river stage

s_n – model river stage

induced by actual precipitation during EQPF period

→ no precip. uncertainty

HYDROLOGIC UNCERTAINTY PROCESSOR



Process: Two-branch, Non-stationary, Markov

Two families of joint conditional densities:

$$\phi_v(\mathbf{h}_N | \mathbf{s}_N, h_0) \qquad v = 0, 1$$

Bayesian Formulation

Meta-Gaussian Model:

- Marginal distributions: Any form
- Dependence structure: Non-linear
  Heteroscedastic

HYDROLOGIC UNCERTAINTY PROCESSOR

Prior 1-step Transition Density

$$r_{nv}(h_n|h_{n-1}) \qquad n = 1,\ldots,N \qquad v = 0,1$$

- **stochastic model** of river stage process

Likelihood Function

$$f_{nv}(s_n|h_n, h_{n-1}, h_0) \qquad n = 1,\ldots,N \qquad v = 0,1$$

- stochastic representation of **deterministic model**

→ Posterior 1-step Transition Density

$$\phi_{nv}(h_n|s_n, h_{n-1}, h_0) \propto f_{nv} \cdot r_{nv} \qquad n = 1,\ldots,N \qquad v = 0,1$$

- Joint Conditional Density

$$\phi_v(\mathbf{h}_N|\mathbf{s}_N, h_0) = \prod_{n=1}^{N} \phi_{nv}(h_n|s_n, h_{n-1}, h_0) \qquad v = 0,1$$

HUP: BAYESIAN REVISION

Conditional Expected Density

$$\kappa_{nv}(s_n|h_{n-1},h_0) = \int_{-\infty}^{\infty} f_{nv}(s_n|h_n,h_{n-1},h_0)\, r_{nv}(h_n|h_{n-1})\, dh_n$$

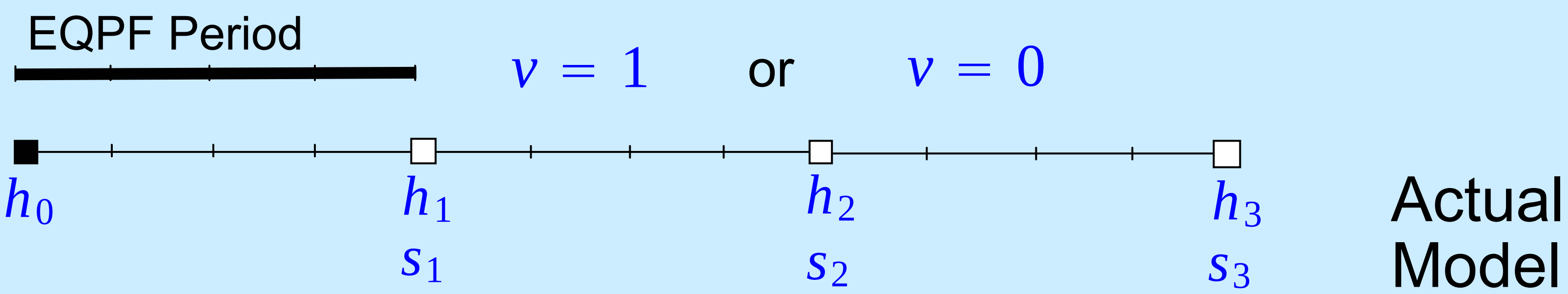
$$n = 1,\ldots,N \qquad v = 0,1$$

Posterior 1–step Transition Density

$$\phi_{nv}(h_n|s_n,h_{n-1},h_0) = \frac{f_{nv}(s_n|h_n,h_{n-1},h_0)\, r_{nv}(h_n|h_{n-1})}{\kappa_{nv}(s_n|h_{n-1},h_0)}$$

$$n = 1,\ldots,N \qquad v = 0,1$$

SAMPLES



Prior Distribution ← historical record (long)

$\{(v; h_0, h_1, h_2, h_3)\}$

Likelihood Function ← simulation experiment (short)

$\{(v; s_1, s_2, s_3, h_0, h_1, h_2, h_3)\}$

Hydrologic model re-initialized (as in real-time)

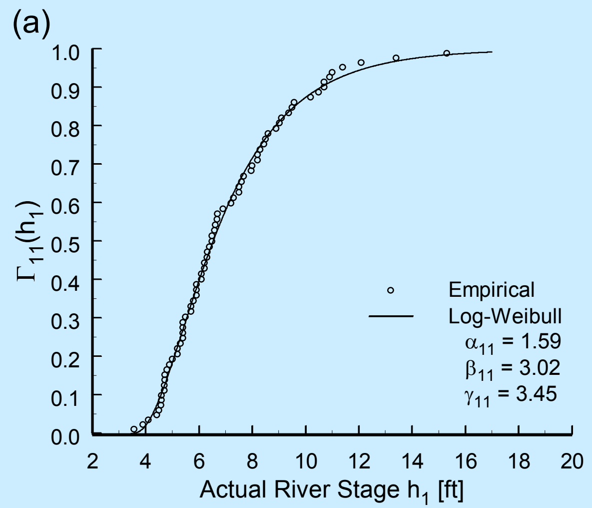
Real-time inputs, except precipitation

Actual precip. during EQPF period

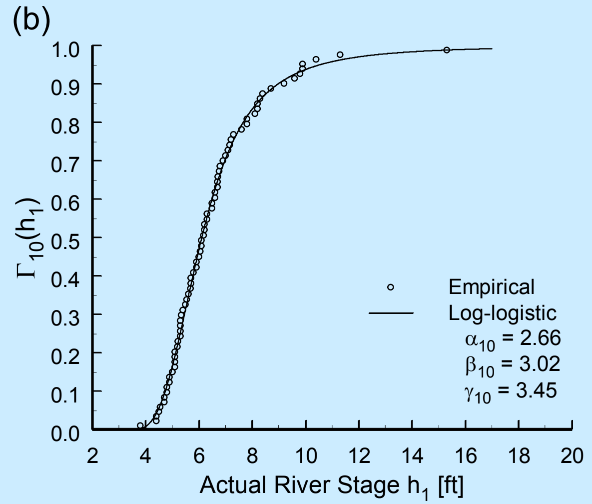
→ No precip. uncertainty

Marginal Distribution Functions of H_1

$V = 1$

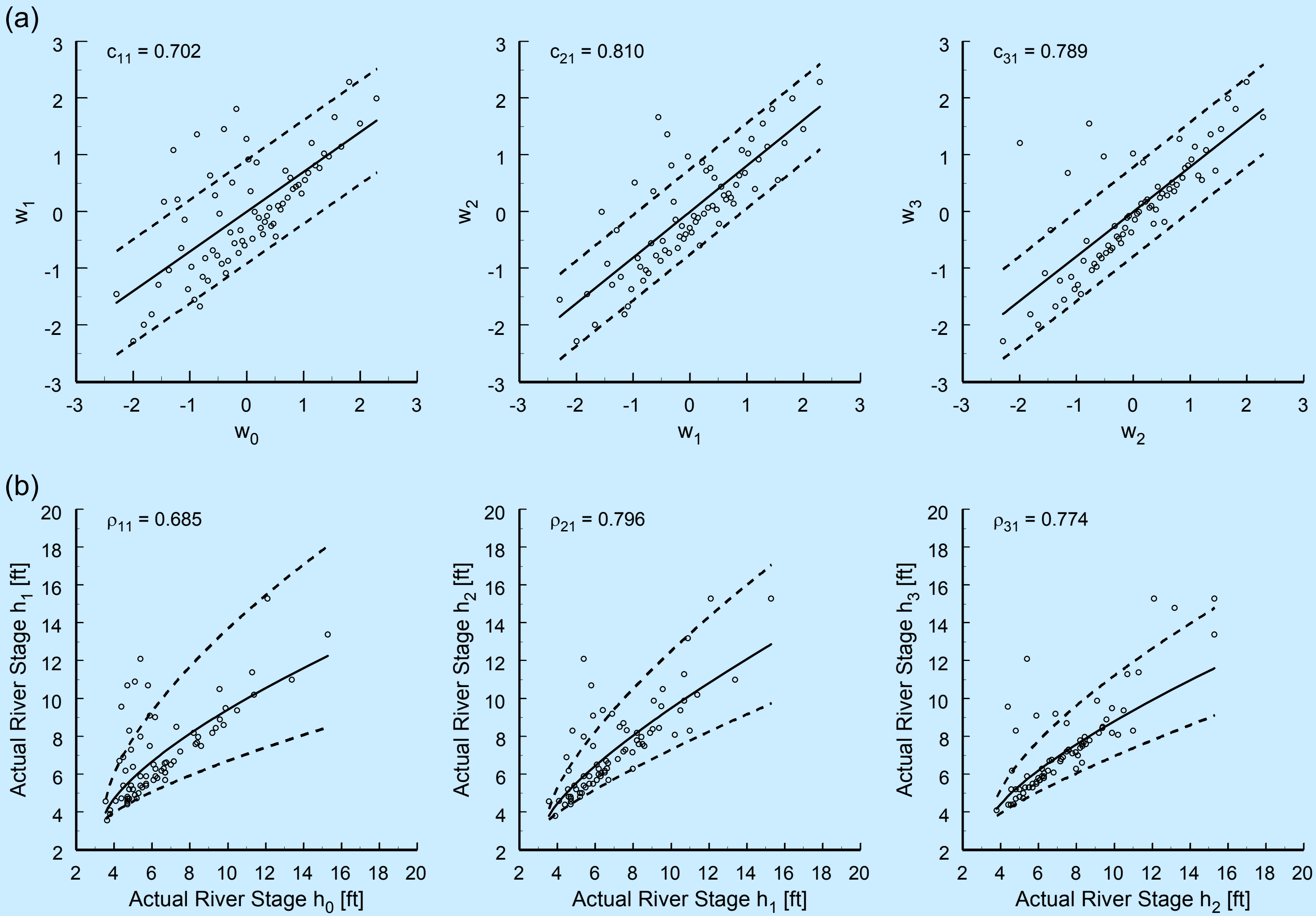


$V = 0$



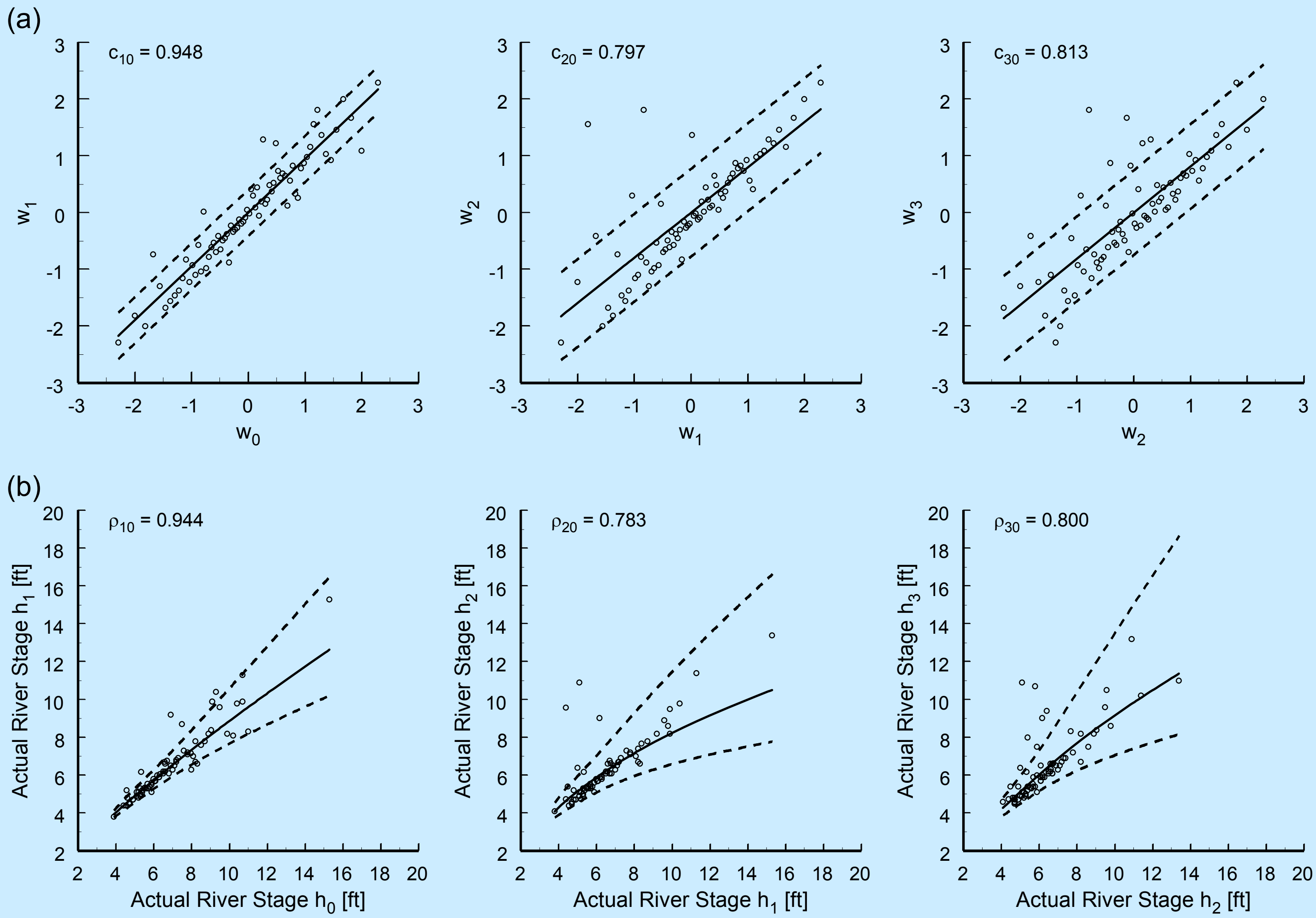
Prior Dependence Structure: $H_n | H_{n-1}, \quad V = 1$

$n = 1$, 24 h $\qquad$ $n = 2$, 48 h $\qquad$ $n = 3$, 72 h

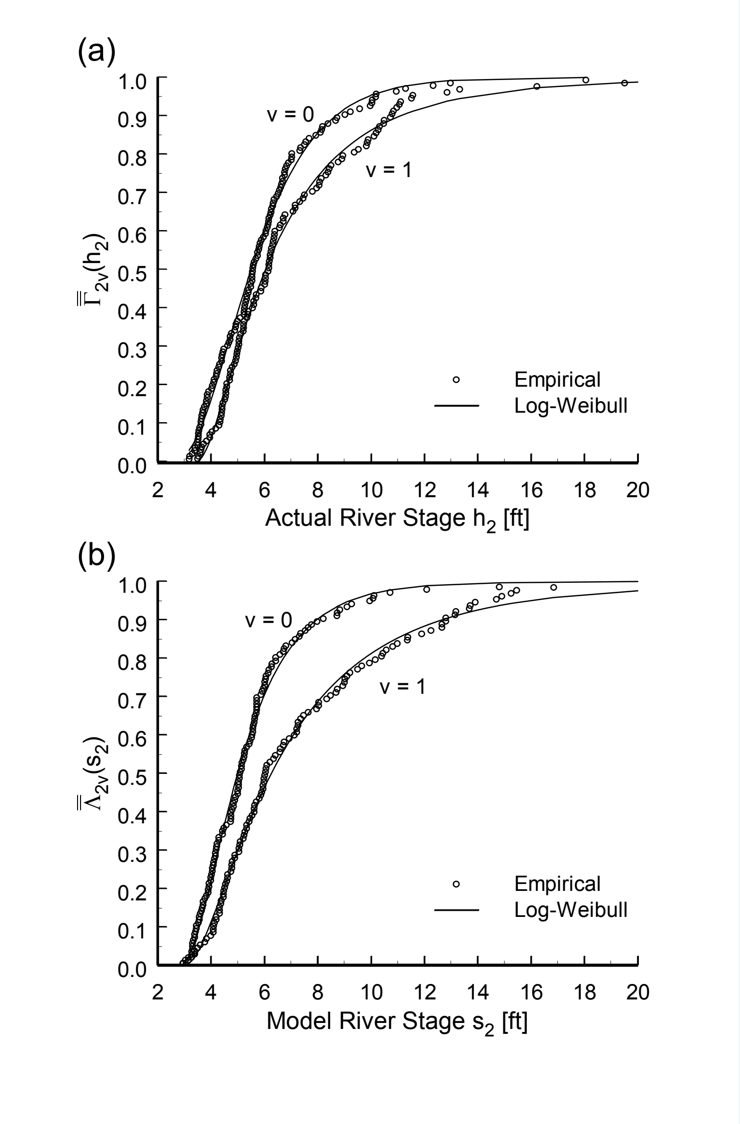


Prior Dependence Structure: $H_n | H_{n-1}, \quad V = 0$

$n = 1$, 24 h $\qquad$ $n = 2$, 48 h $\qquad$ $n = 3$, 72 h

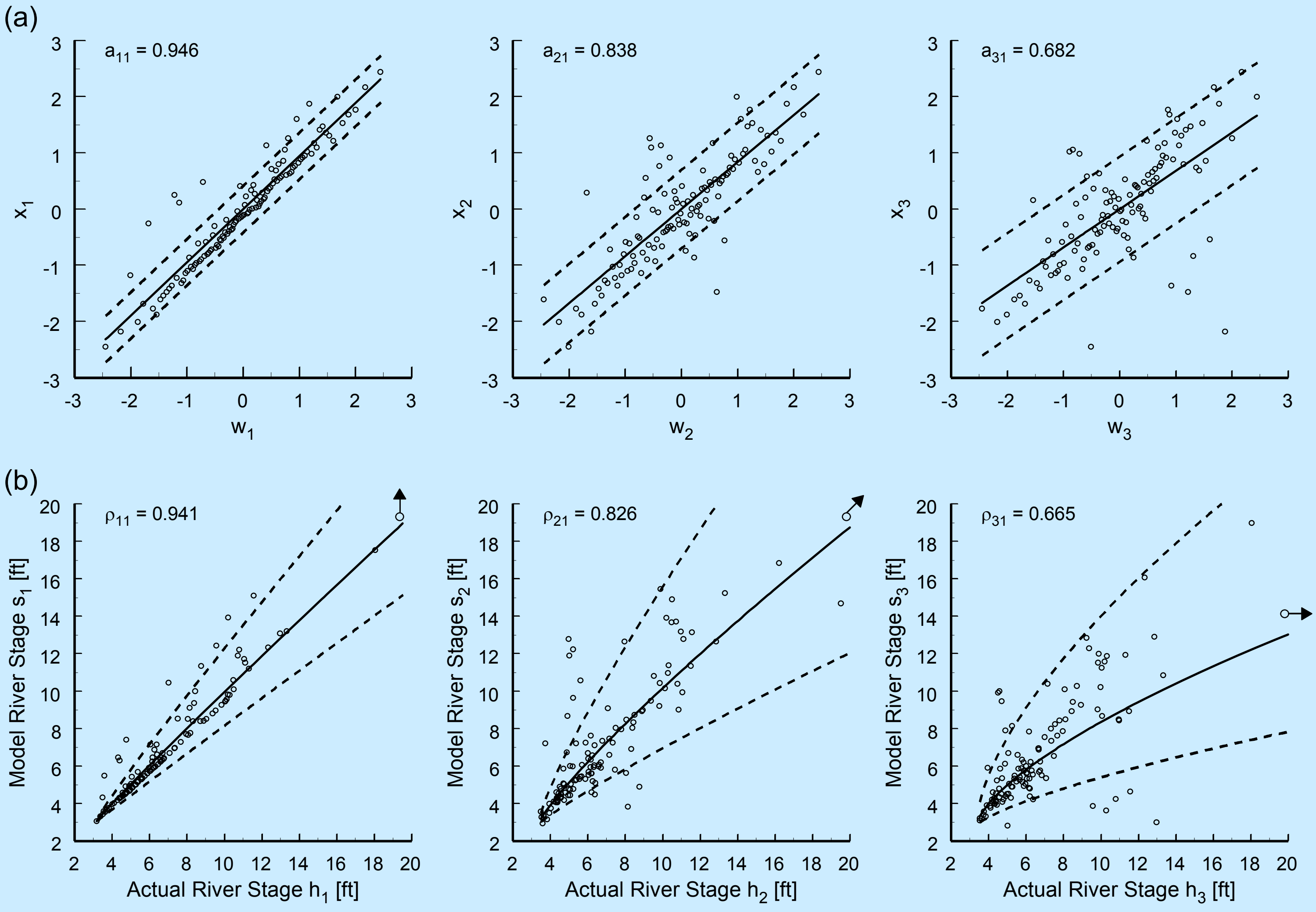


Marginal Distribution Functions of H_2 and S_2



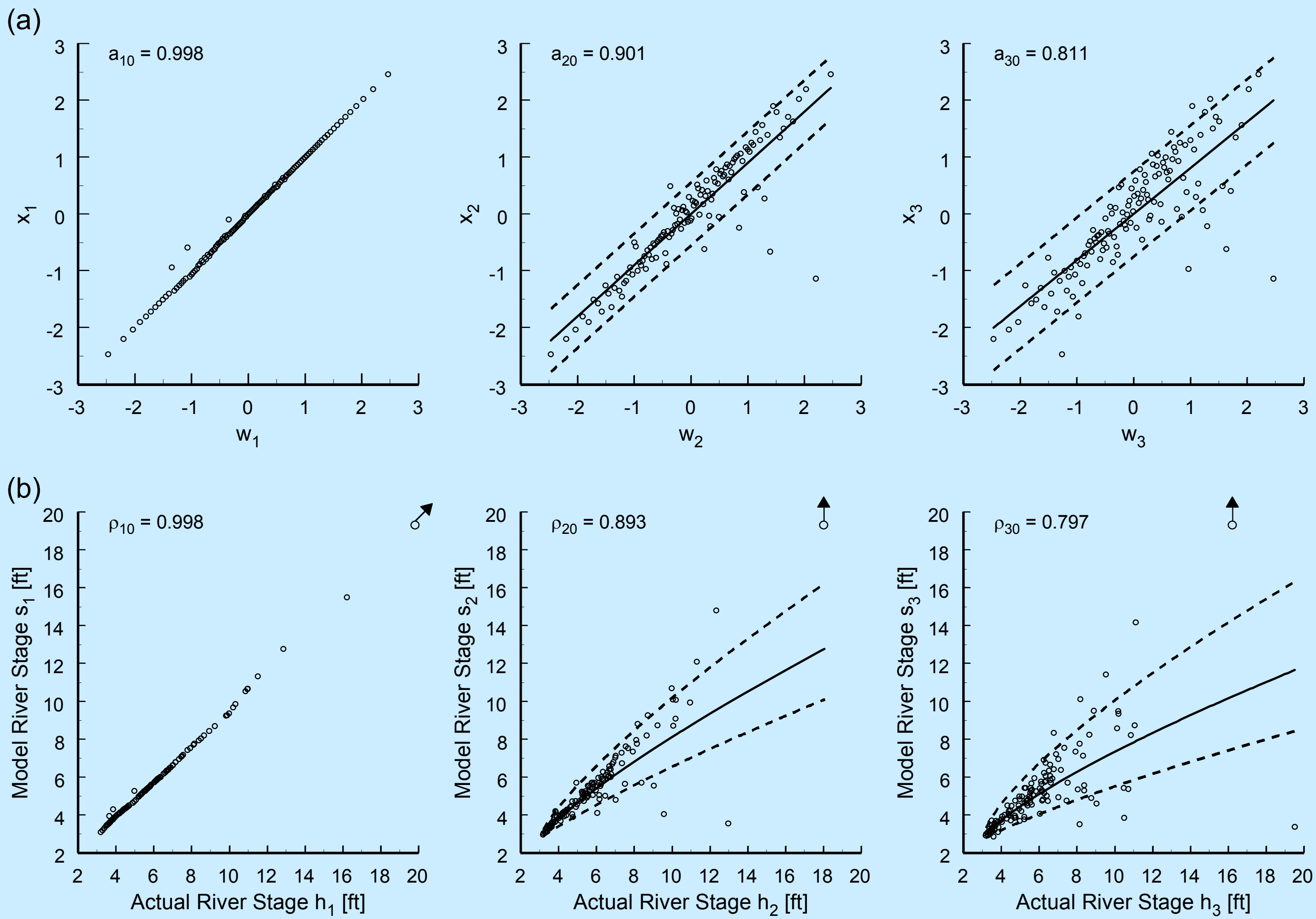
Likelihood Dependence Structure: $S_n | H_n, \quad V = 1$

$n = 1$, 24 h $\quad$ $n = 2$, 48 h $\quad$ $n = 3$, 72 h



Likelihood Dependence Structure: $S_n | H_n, \quad V = 0$

$n = 1$, 24 h $\qquad$ $n = 2$, 48 h $\qquad$ $n = 3$, 72 h



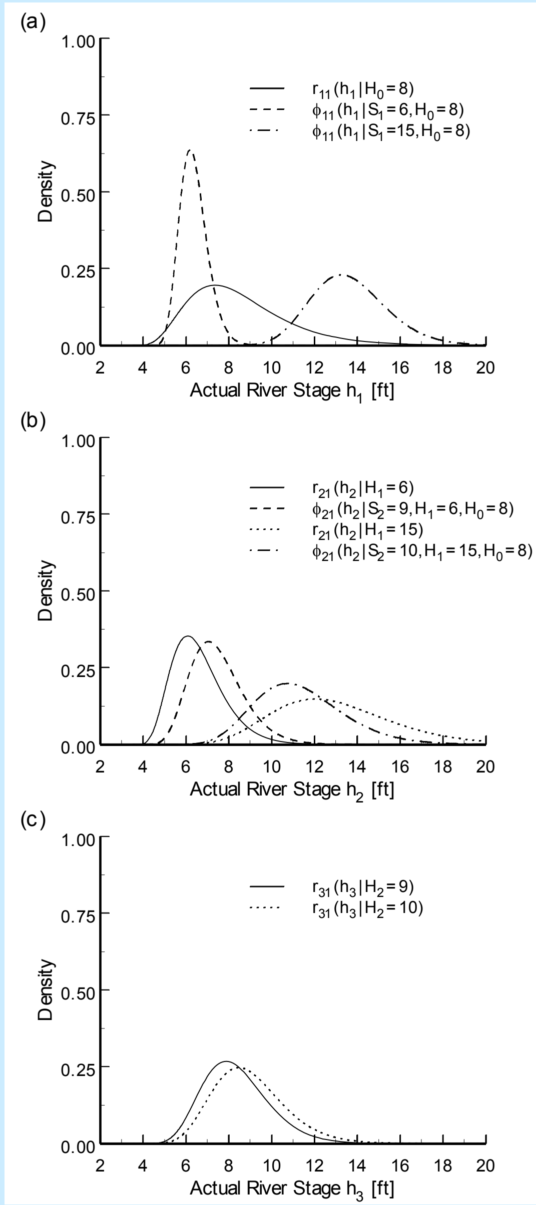
Prior and Posterior 1-Step Transition Densities

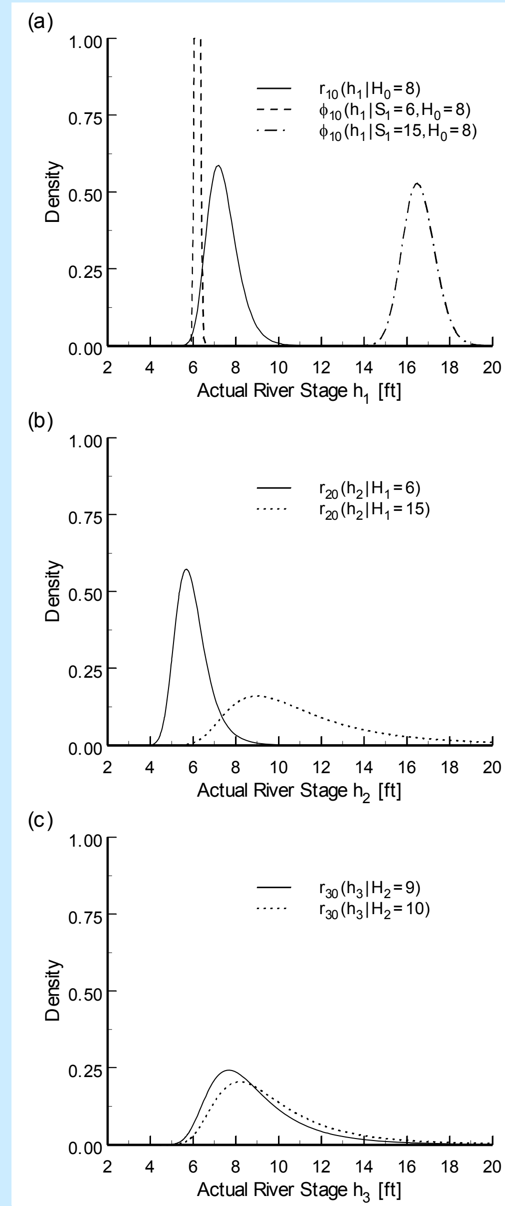
$V = 1$ | $V = 0$

$n = 1$ — 24 h

$n = 2$ — 48 h

$n = 3$ — 72 h





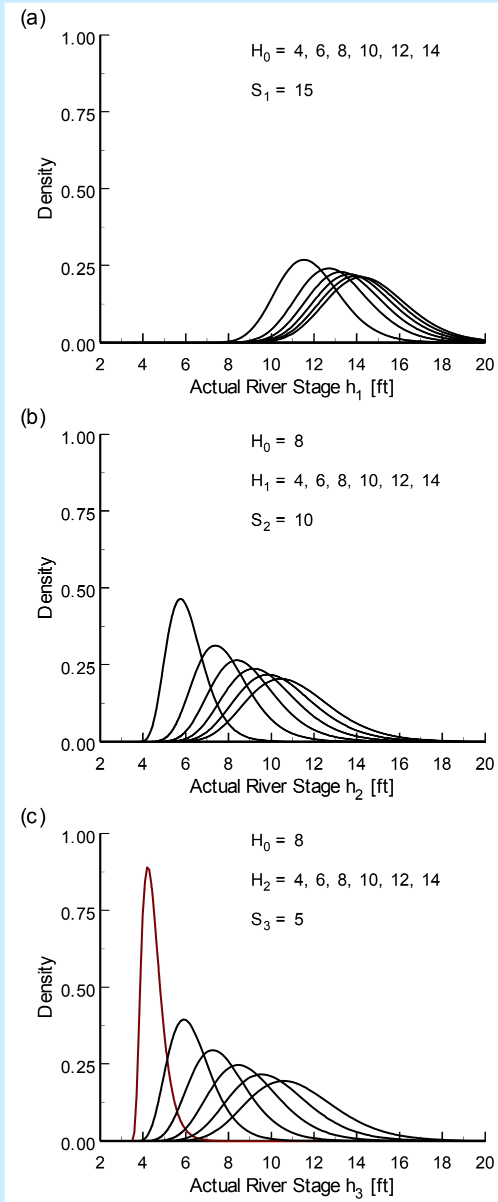
Posterior 1-Step Transition Densities

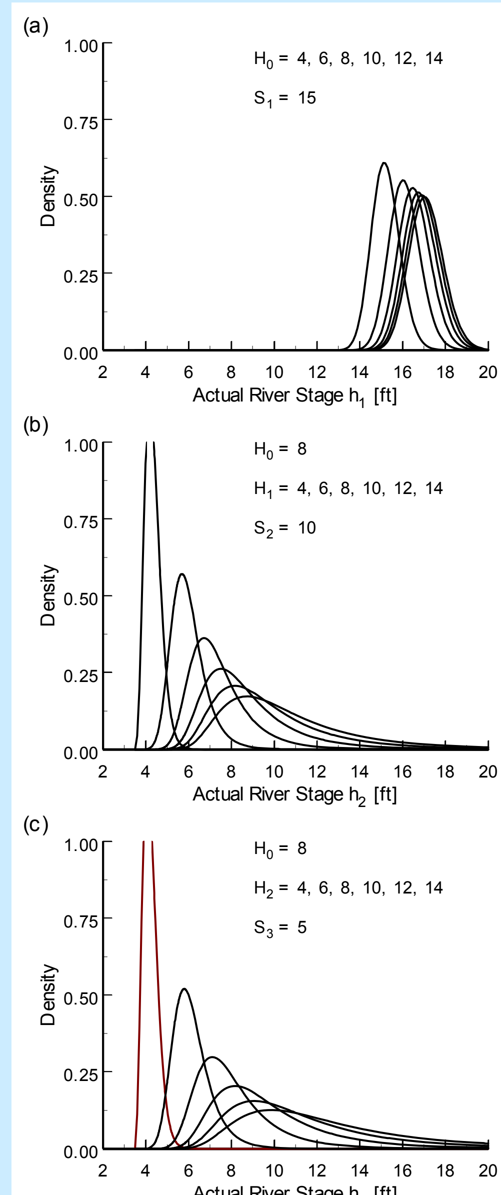
$V = 1$ | $V = 0$

$n = 1$
24 h

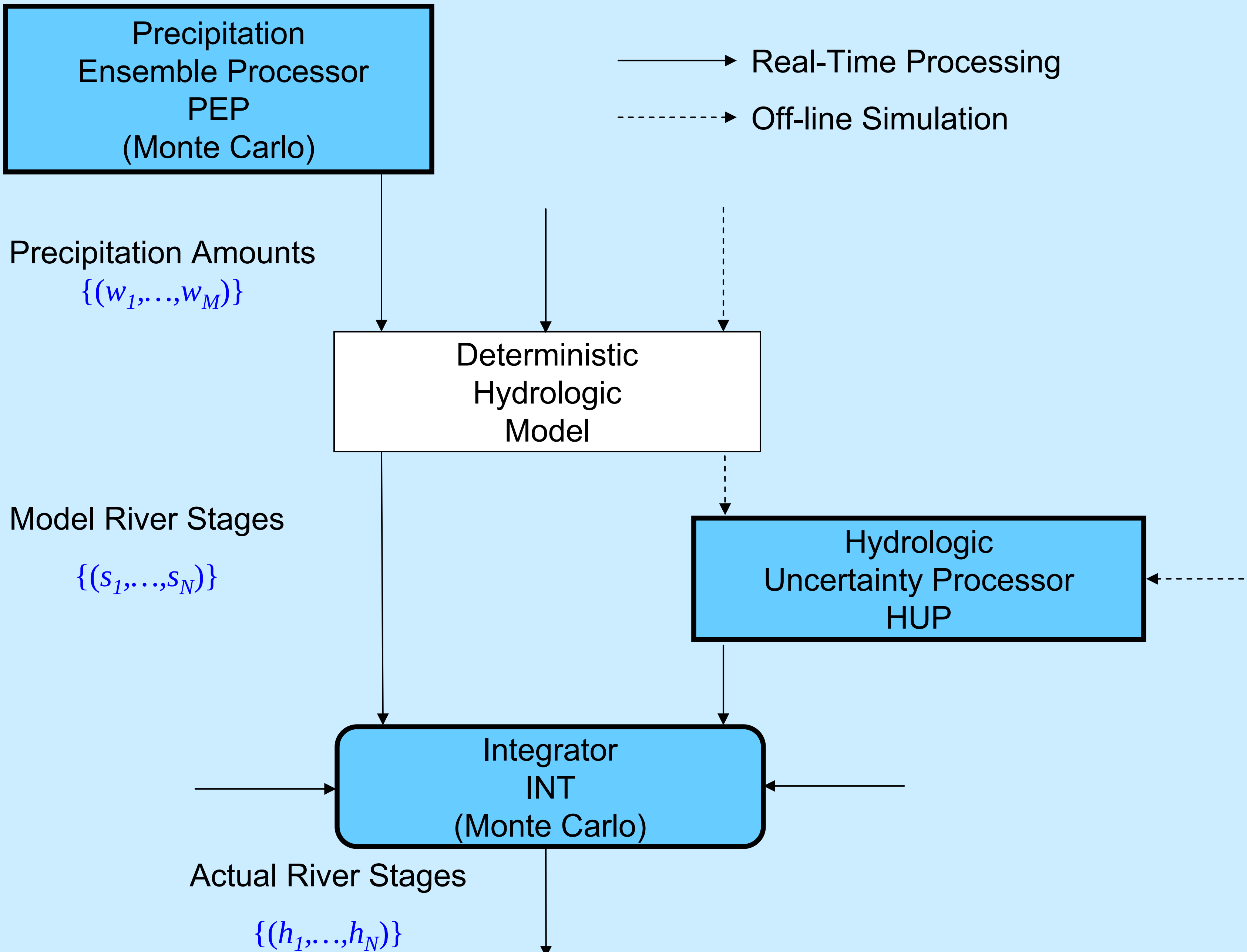
$n = 2$
48 h

$n = 3$
72 h





ENSEMBLE BAYESIAN SYSTEM



INTEGRATOR (Monte Carlo)

- Input: one realization

Model River Stages $(s_1, s_2, s_3, \ldots, s_N)$

Indicator of Precip. v

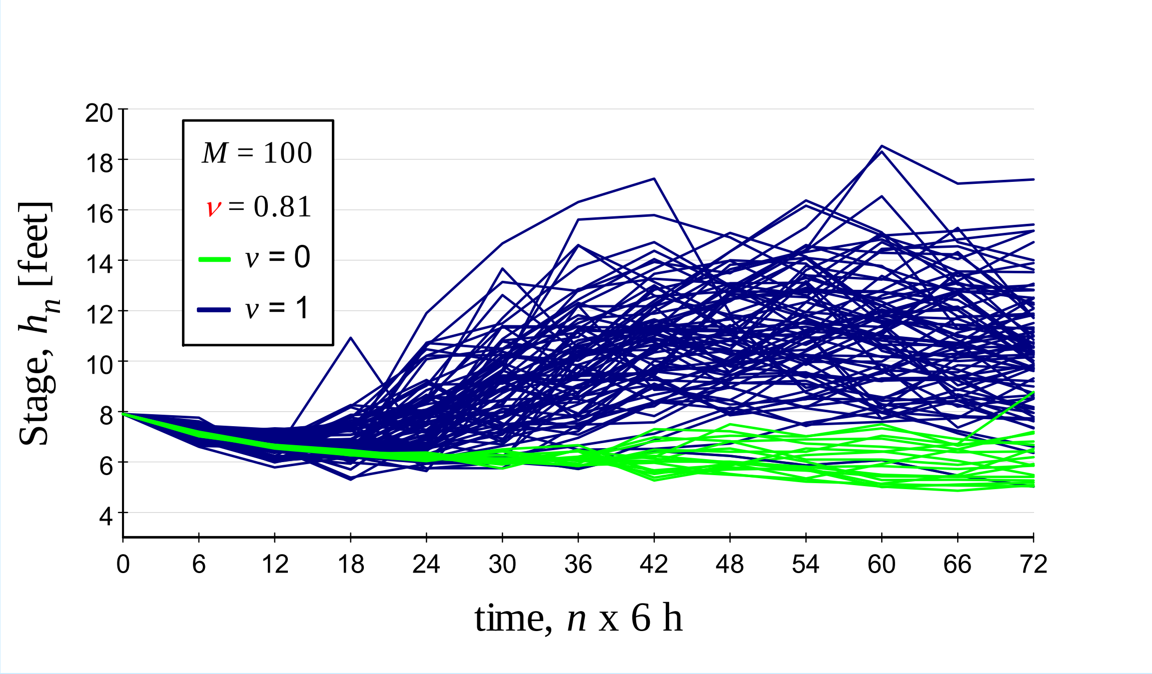
Initial River Stage h_0

Given	Generate	From Posterior Distribution
s_1	h_1	$\Phi_{1v}(\cdot \mid s_1, h_0)$
s_2, h_1	h_2	$\Phi_{2v}(\cdot \mid s_2, h_1, h_0)$
s_3, h_2	h_3	$\Phi_{3v}(\cdot \mid s_3, h_2, h_0)$
⋮	⋮	⋮

- Output: multiple realizations

Actual River Stages $\{(h_1, h_2, h_3, \ldots, h_N)\}$

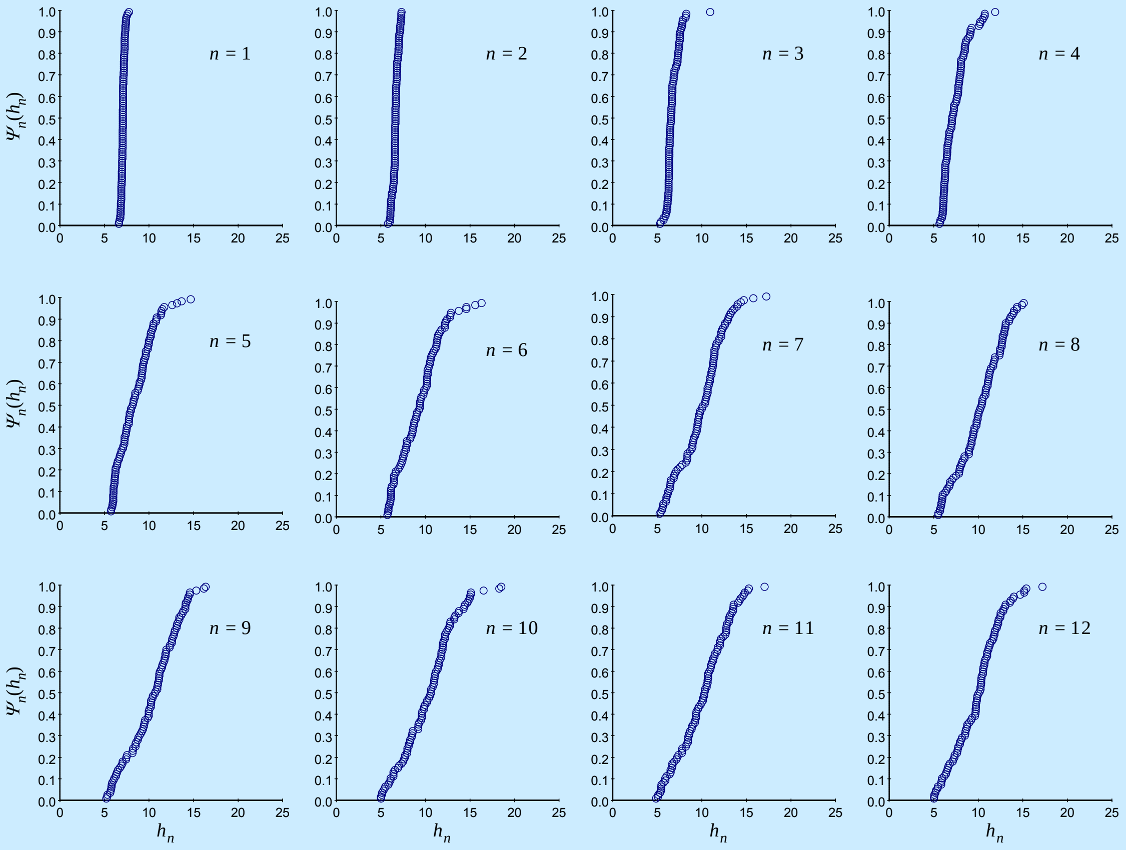
Bayesian Ensemble Forecast of River Stages



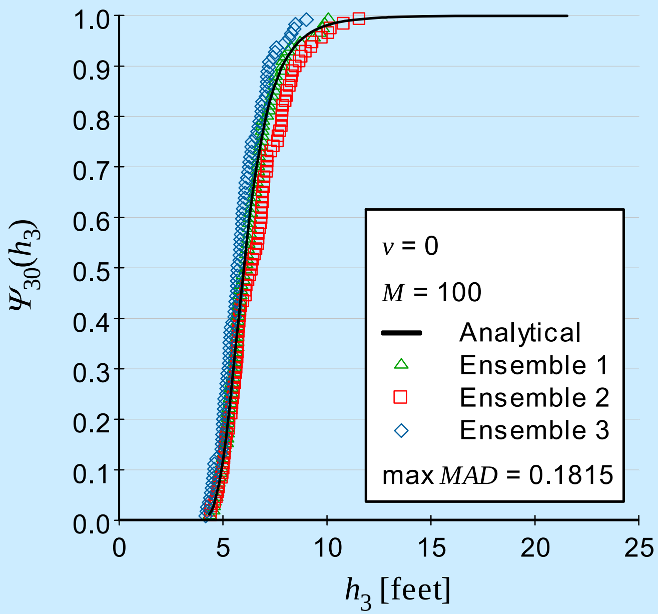
Predictive n-Step Transition Distributions

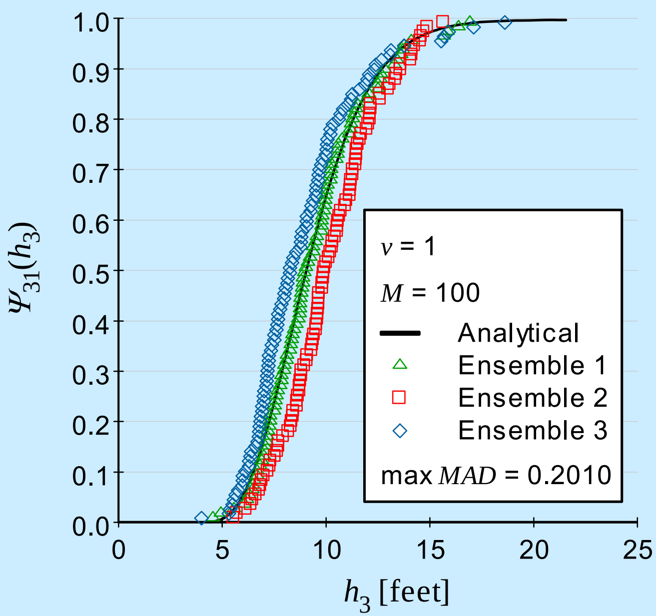
Empirical Distributions

Ensemble Size M = 100



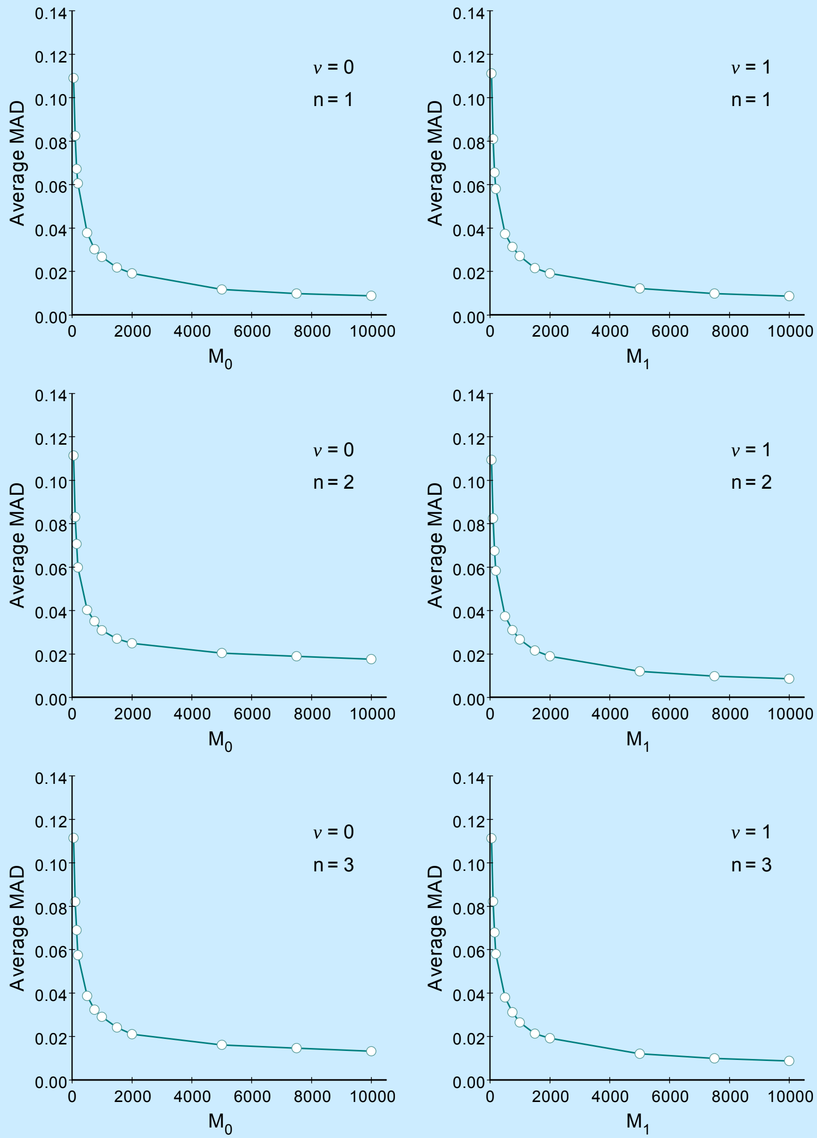
ERROR IN DISTRIBUTION Due to Ensemble Size





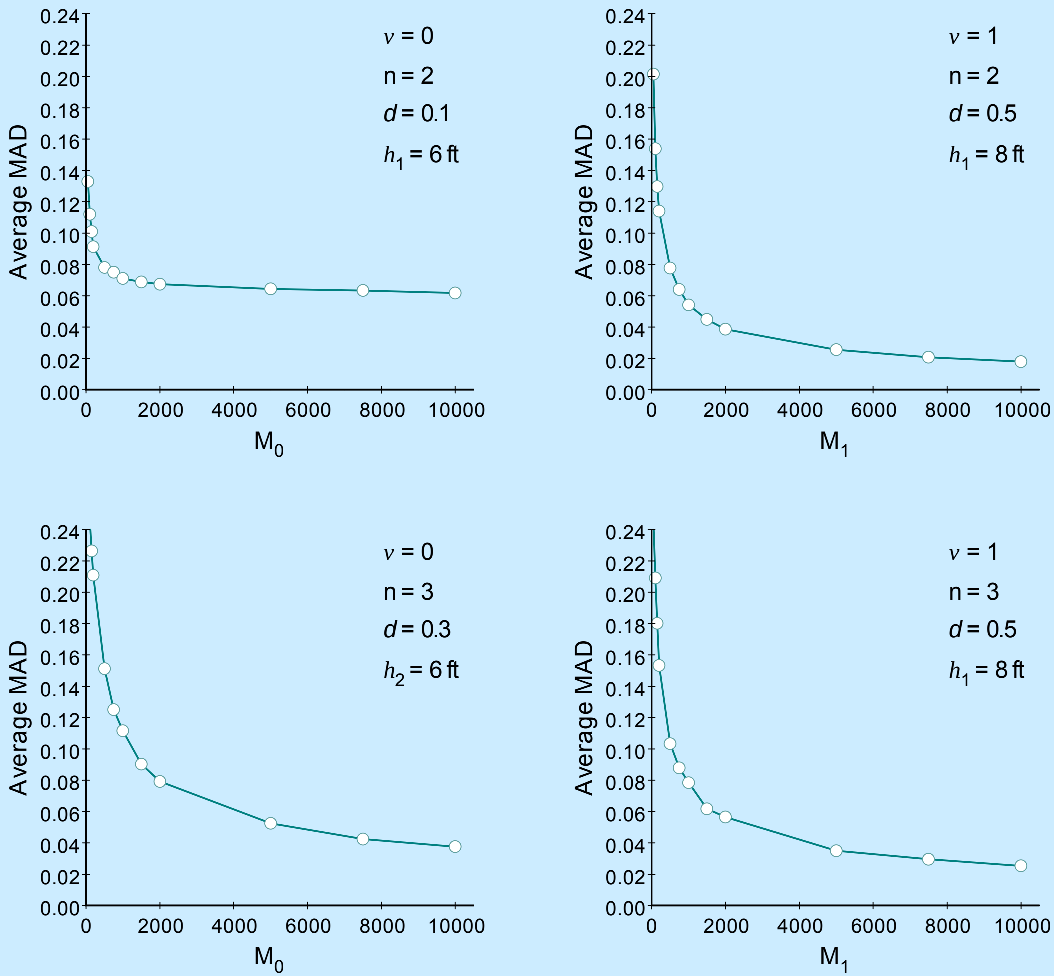
MAD Between Ensemble $\hat{\Psi}_{nv}(h_n)$ and Analytical $\Psi_{nv}(h_n)$

Average *MAD* across 500 ensembles



MAD Between Ensemble $\hat{\Theta}_{nv}(h_n|h_{n-1})$ and Analytical $\Theta_{nv}(h_n|h_{n-1})$

Average *MAD* across 500 ensembles



BAYESIAN THEORY: SUMMARY

Ensemble Bayesian Forecasting System (EBFS)

0. Deterministic Hydrologic Model

1. Precipitation Ensemble Processor
    - calibrated ensemble QPF

2. Hydrologic Uncertainty Processor
    - all hydrologic uncertainties
    - correct theoretic structure (Bayesian)
    - model fitting data (meta-Gaussian)

3. Integrator
    - Monte Carlo re-generation (ensemble size)

Forecast Products

- calibrated distributions

REFERENCES

Krzysztofowicz, R., Bayesian Theory of Probabilistic Forecasting via Deterministic Hydrologic Model, *Water Resources Research*, 35(9), 2739–2750, 1999.

Krzysztofowicz, R. and Herr, H.D., Hydrologic Uncertainty Processor for Probabilistic River Stage Forecasting: Precipitation-Dependent Model, *Journal of Hydrology*, 249(1–4), 46–68, 2001.

Maranzano, C.J. and Krzysztofowicz, R., Identification of Likelihood and Prior Dependence Structures for Hydrologic Uncertainty Processor, *Journal of Hydrology*, 290(1–2), 1–21, 2004.

Krzysztofowicz, R. and Maranzano, C.J., Hydrologic Uncertainty Processor for Probabilistic Stage Transition Forecasting, *Journal of Hydrology*, 293(1–4), 57–73, 2004.

Krzysztofowicz, R. and Maranzano, C.J., Bayesian System for Probabilistic Stage Transition Forecasting, *Journal of Hydrology*, 299(1–2), 15–44, 2004.

Herr, H.D. and Krzysztofowicz, R., Bayesian Ensemble Forecast of River Stages and Ensemble Size Requirements, *Research Paper* RK–0801, July 2008.

*   *   *