Ensemble Data Assimilation for Watershed Water Quality Forecasting

Sunghee Kim[1],Hamideh Riazi[1], D.-J. Seo[1], Changmin Shin[2] ,Kyunghyun Kim[2]

[1]Dept. Of Civil Eng., The Univ. of Texas at Arlington, Arlington, TX, USA
[2]Water Quality Control Center, National Institute of Environmental Research, Incheon, Korea

In this presentation

- Water quality forecasting
  - Needs, challenges for DA
- Maximum likelihood ensemble filter for Hydrologic Simulation Program – Fortran (MLEF-HSPF)
- Hindcasting experiment
- Operational implementation
- Conclusions and research questions

Why watershed water quality forecasting?

Example: Algal control in Youngsan River (1)

- Background
  - Severe algal blooms in September 2011
  - Water temperature > 30 ℃
  - Chl-a conc. > 123.6 ug/L

- **Water temperature and Chl-a data at Seongcheon weir in Sept. 2011**

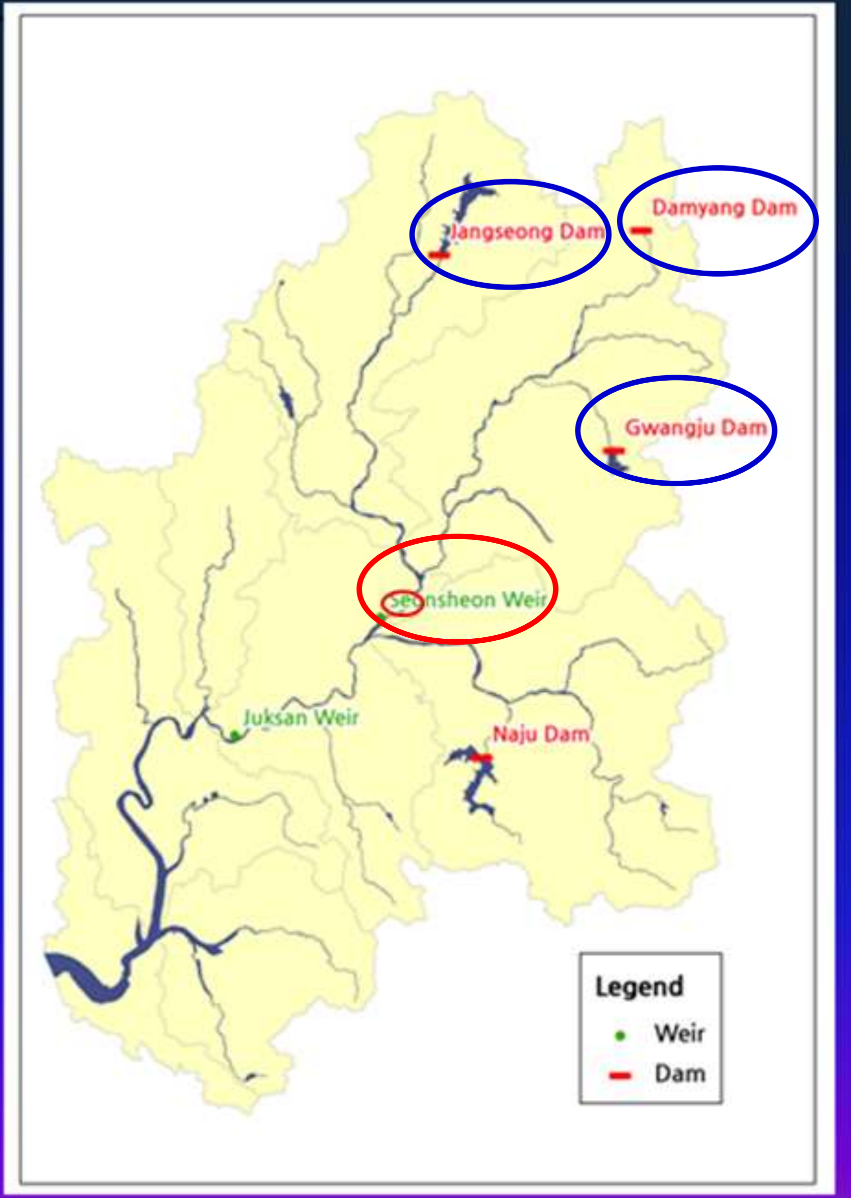
Location		08-30	09-01	09-02	09-03	09-04	09-05	09-06	09-07
500m upstream of Seongcheon weir	Water Temp.	-	30.4	28.9	30.8	30.3	29.0	28.4	27.9
	Chl-a	-	22.7	54.5	44.9	64.9	122.8	123.6	-
500m downstream of Seongcheon weir	Water Temp.	34.1	30.0	28.7	30.8	34.1	29.4	27.8	27.2
	Chl-a	124.9	32.8	35.8	67.0	58.5	183.7	102.4	-

Example: Algal control in Youngsan River (3)

❖ Actions taken base on Scenario #3

1. Water was released from agricultural dams

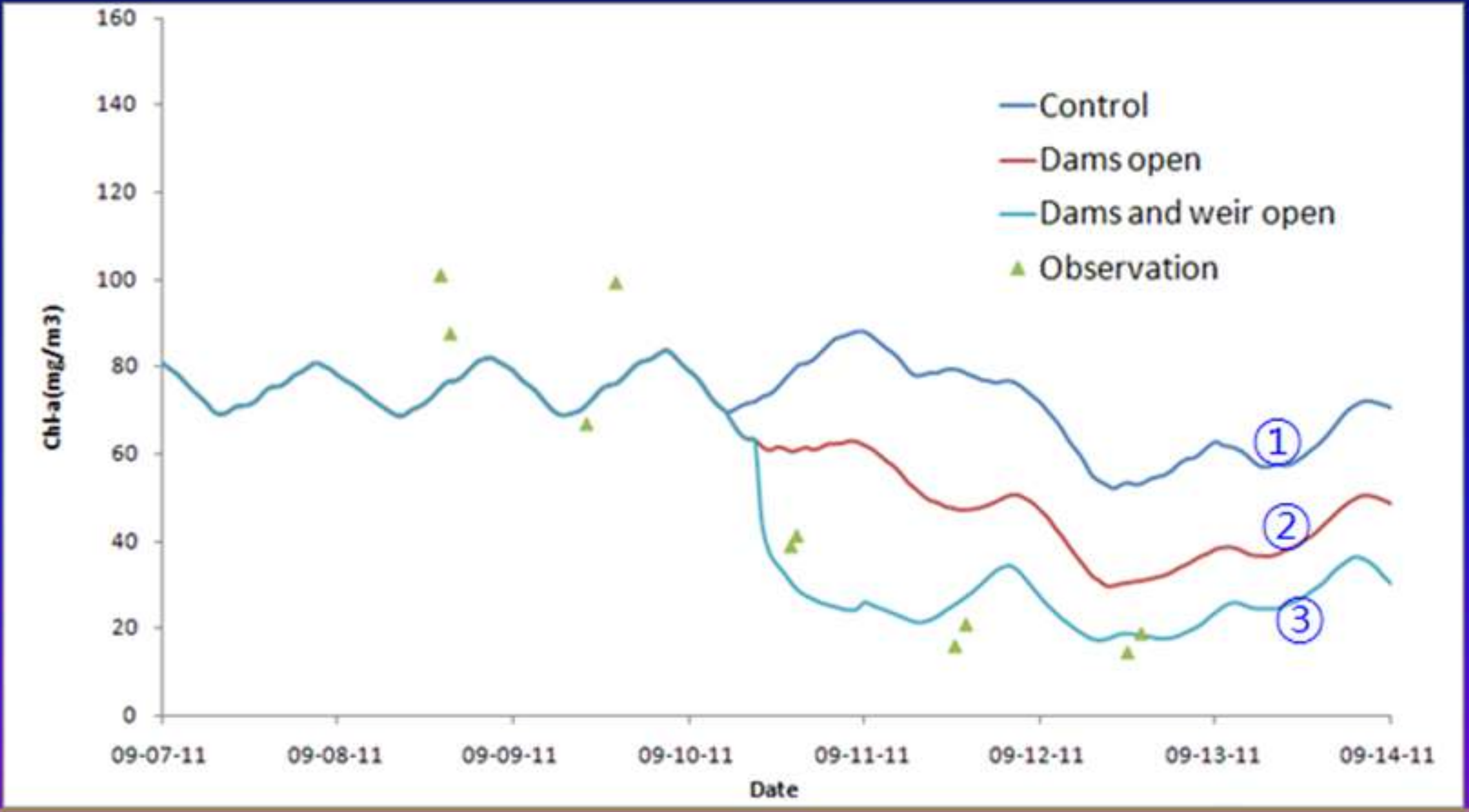
2. The water level was reduced in Seongcheon weir



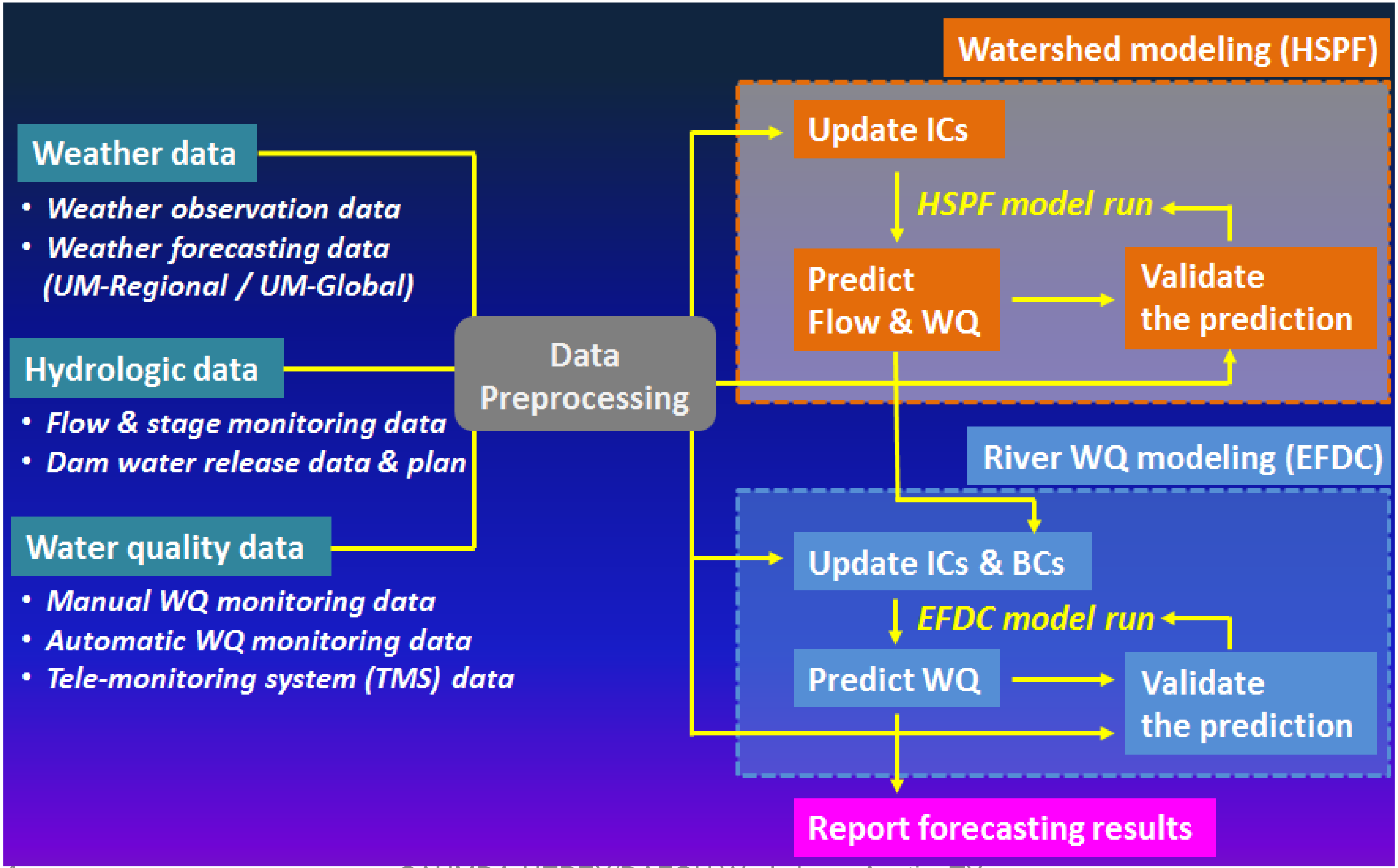
Example: algal control in Youngsan River (4)

Results:

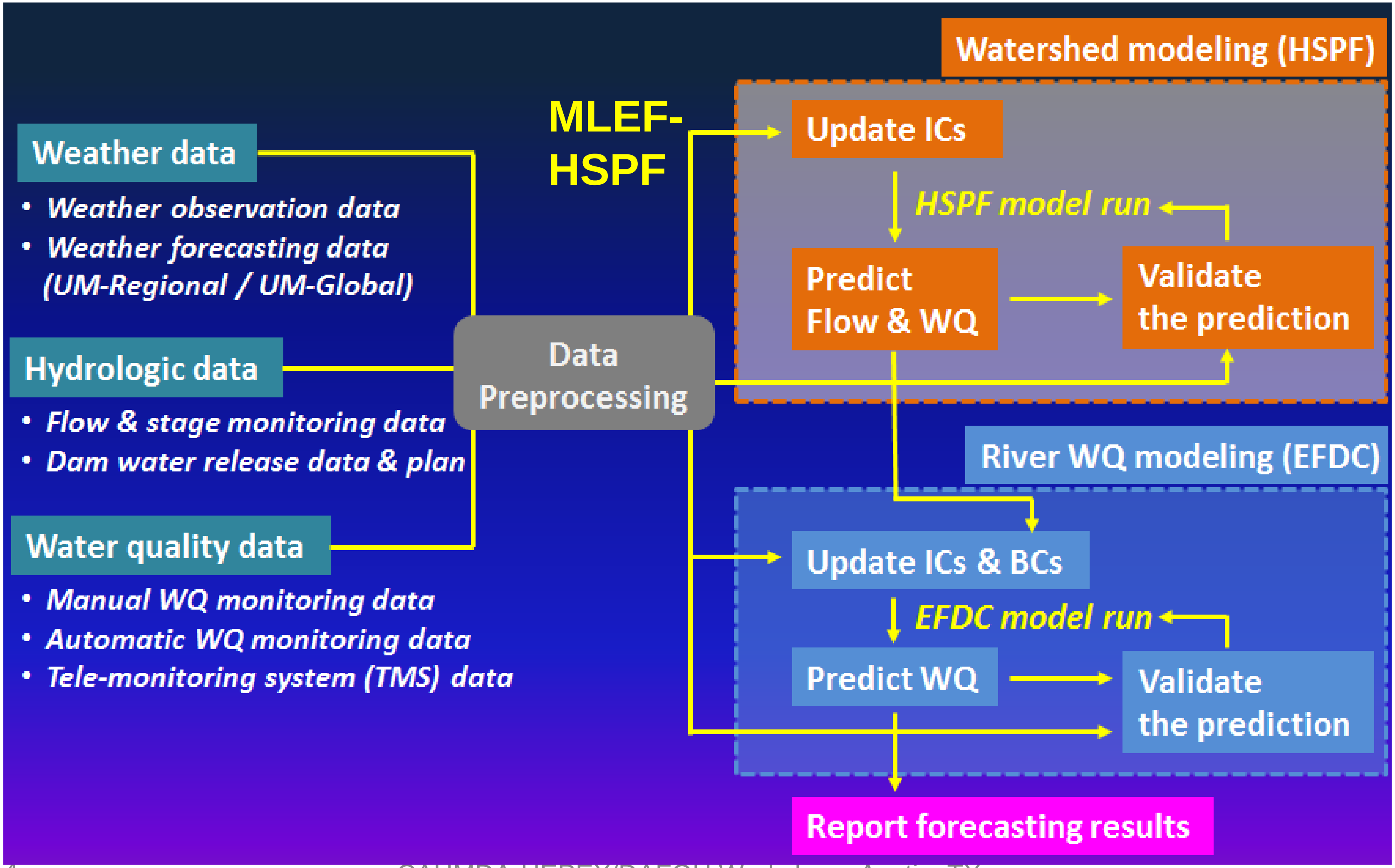
Observations showed decreased concentration up to 19 ug/L on Sept. 12 (prediction 20 ug/L)



Operational WQ forecasting in NIER, Korea



Operational WQ forecasting in NIER, Korea



DA for HSPF – Challenges

- High dimensionality
  - A large number of control variables (28 state variables, 31 model segments → 331 control variables)
- Sparse observations
  - In-river variables only
  - Weekly only
- Nonlinear bio-physiochemical processes
- Nonlinear observations

MLEF-HSPF

- Based on MLEF (Zupanski, 2005)
  - Model error added via state augmentation

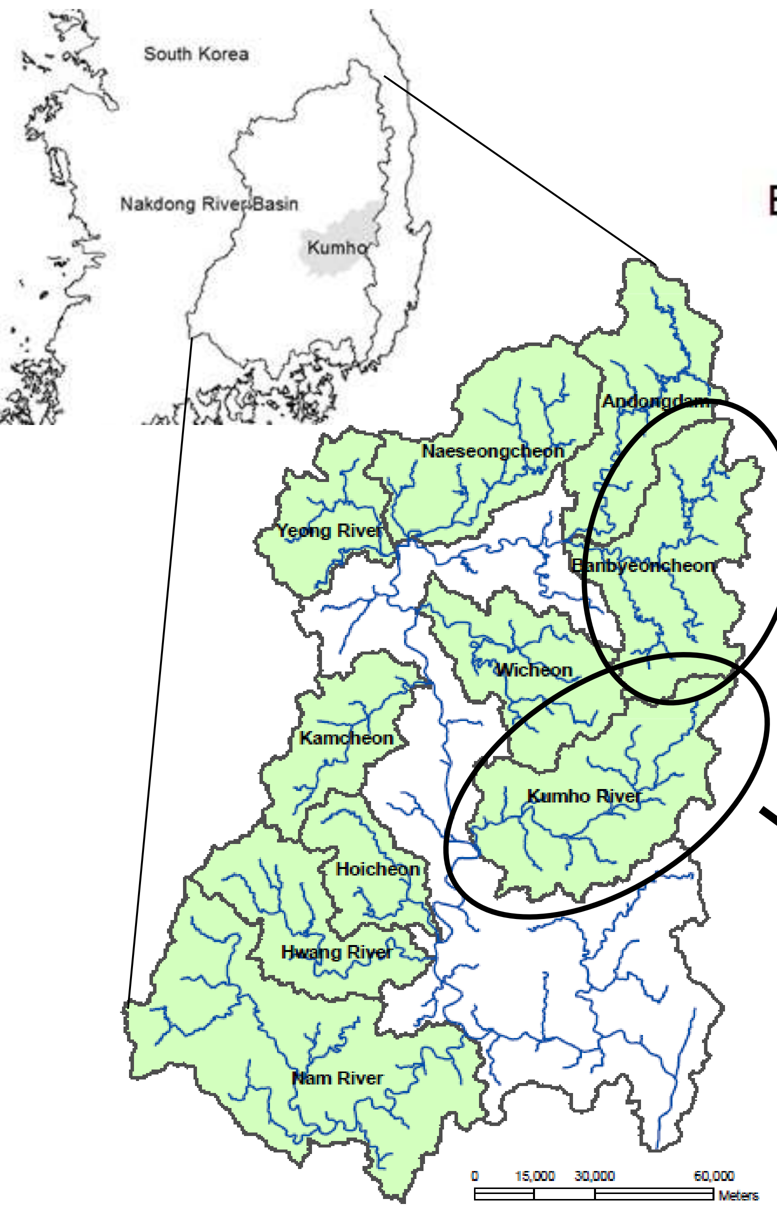
$$P_f(k) = M_{k-1,k} P_a(k-1) M_{k-1,k}^T + Q(k-1)$$
$$= \{M_{k-1,k} P_a(k-1)\}^{1/2} \{M_{k-1,k} P_a(k-1)\}^{T/2} + Q^{1/2}(k-1) Q^{T/2}(k-1)$$
$$P_f^{1/2}(k) = [\{M_{k-1,k} P_a(k-1)\}^{1/2} \quad Q^{1/2}(k-1)]^T$$

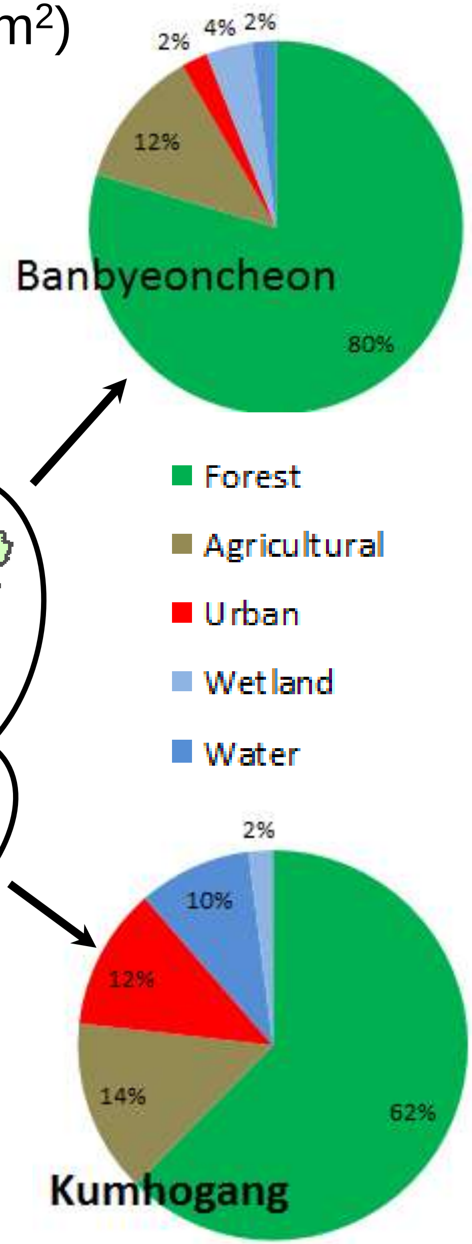
- Formulated as a fixed-lag smoother
- Bias correction added in the observation equations to remove/reduce systematic biases
  - Conditional bias-penalized linear regression (Seo 2012)

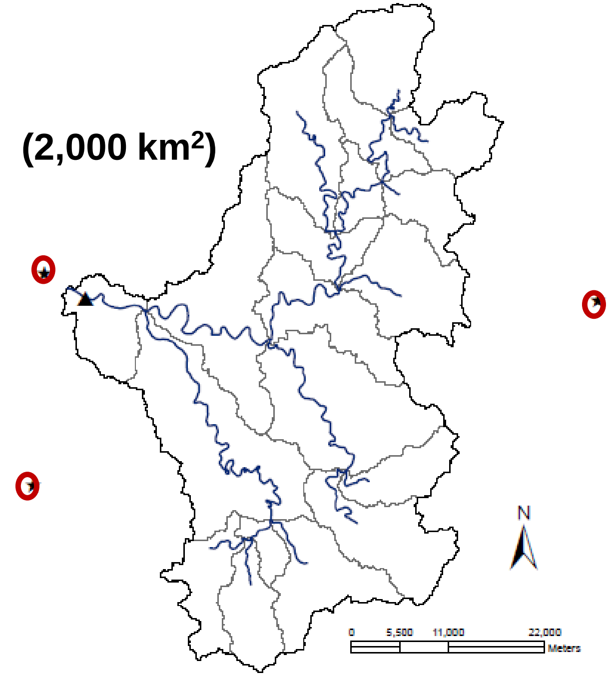
Hindcasting experiment

- 2 catchments in the Nakdong River Basin, Korea, for 2008-2009
  - Kumho, Banbyeon
- Assimilate weekly observations at the outlet and up to 4 interior monitoring stations
  - Flow, TW, NH_4, NO_3, PO_4, CHL-a, TN, TP, TOC, BOD, DO
- Ensemble size of 9 based on sensitivity and eigenvalue spectrum analysis (see poster by Riazi et al.)
- No. of control variables: 333 for Kumho, 316 for Banbyeon

Nakdong River Basin (23,817 km^2)

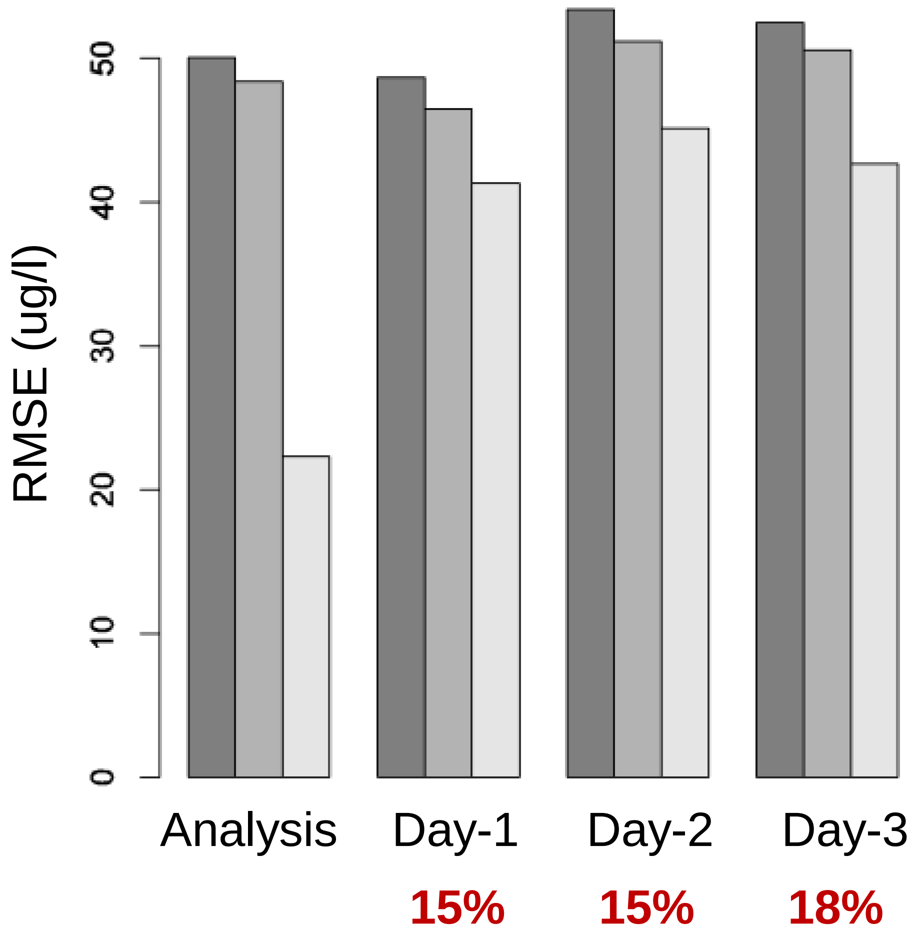






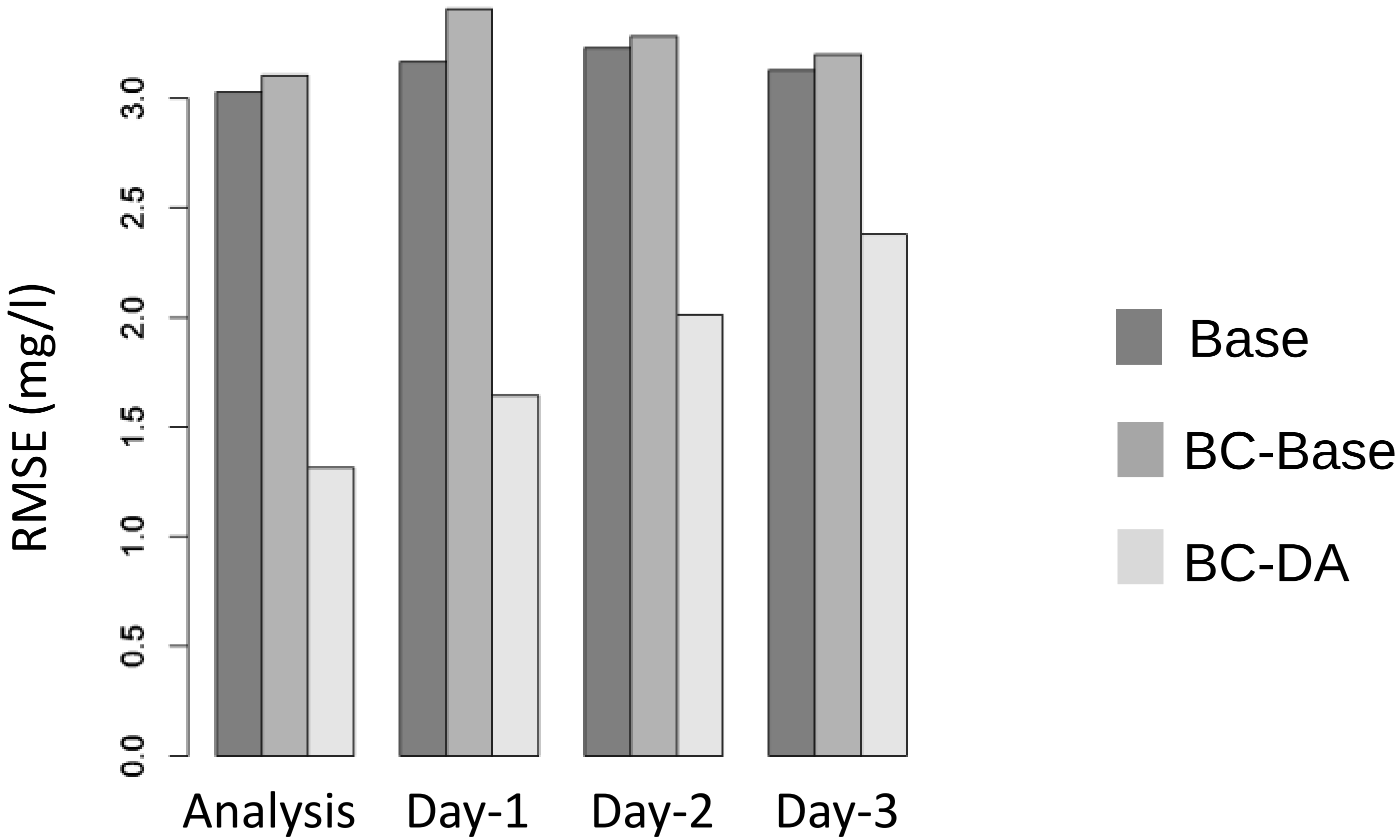
Meteorological stations

Kumho, CHL-a, outlet only

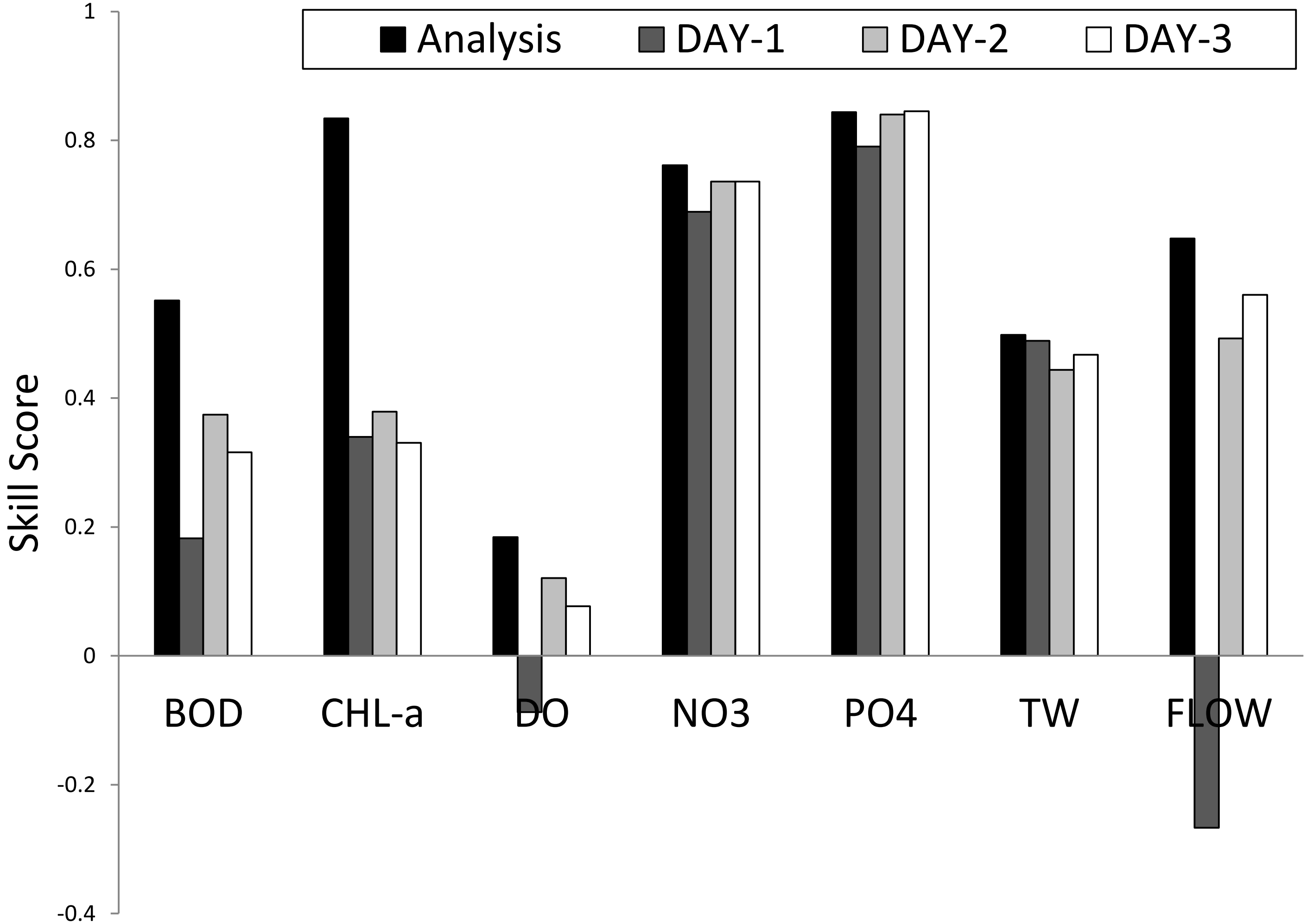


Base	HSPF simulation without Bias correction or DA
BC-Base	HSPF simulation with bias correction
BC-DA	HSPF simulation with bias correction and DA

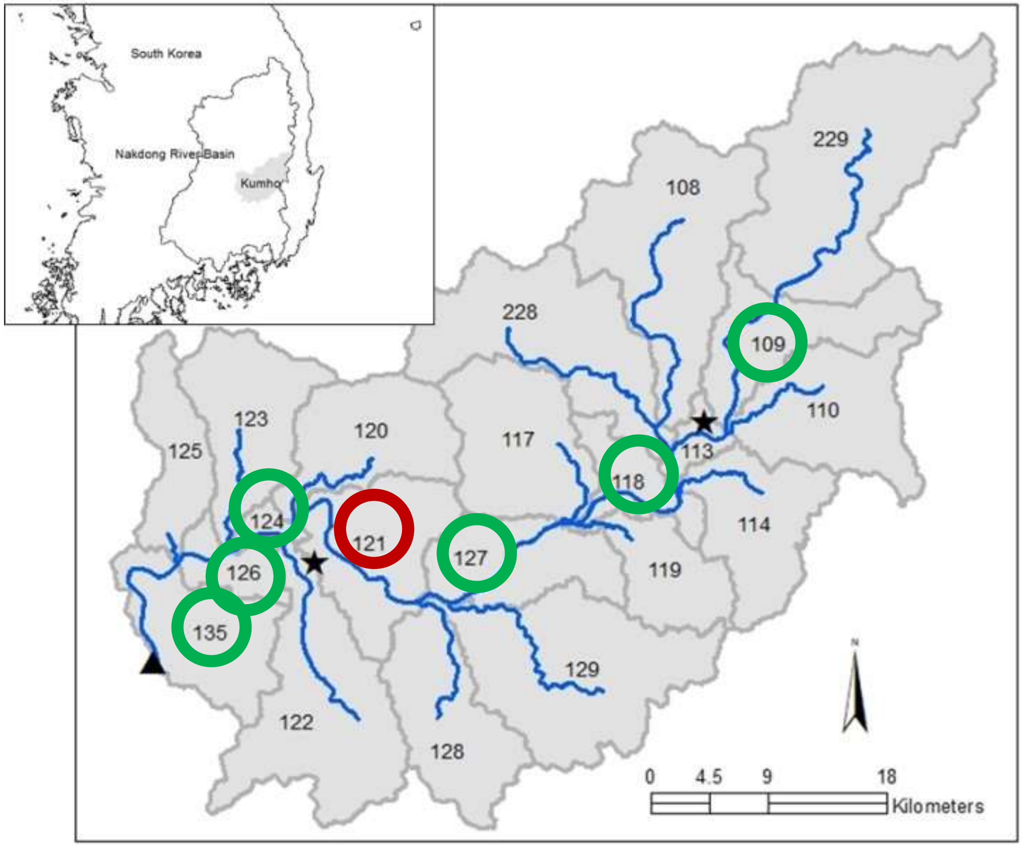
Banbyeon, DO, outlet only



MSE skill score - Kumho



Kumho



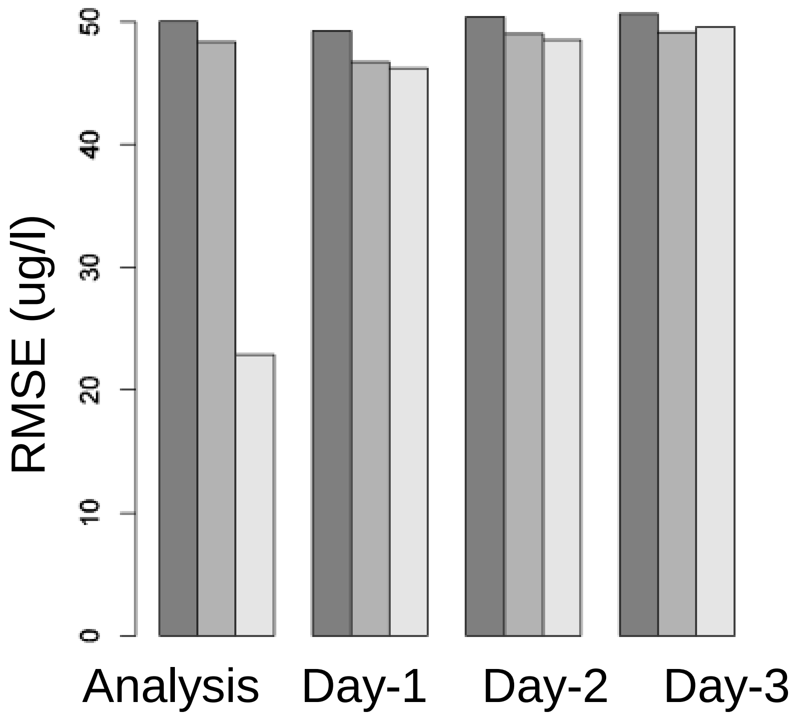
Reach No.	Sampling frequency (sample size)	Missing observations
109	Weekly (96)	Flow
118	Weekly (96)	NO_3, PO_4 CHL-a in 2008
121	Weekly (96)	NO_3, PO_4
124	Weekly (96)	Flow
126	Weekly (96)	Flow
127	Weekly (96)	NO_3, PO_4 CHL-a in 2008
135	Weekly (96)	------

Reach 135: outlet

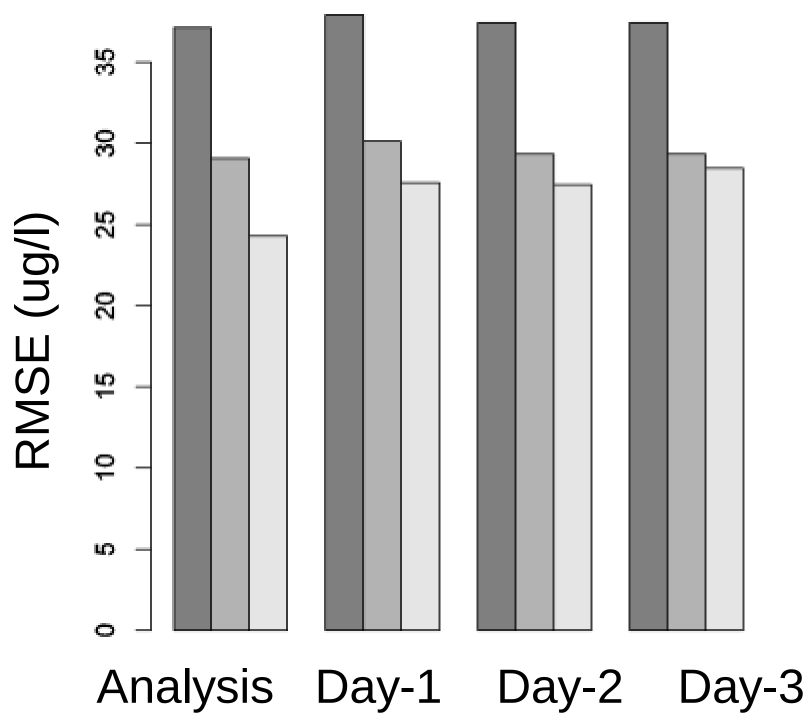
CHL-a at Reach 121

(NO_3, PO_4 observations missing)

Obs from 1station (121) assimilated

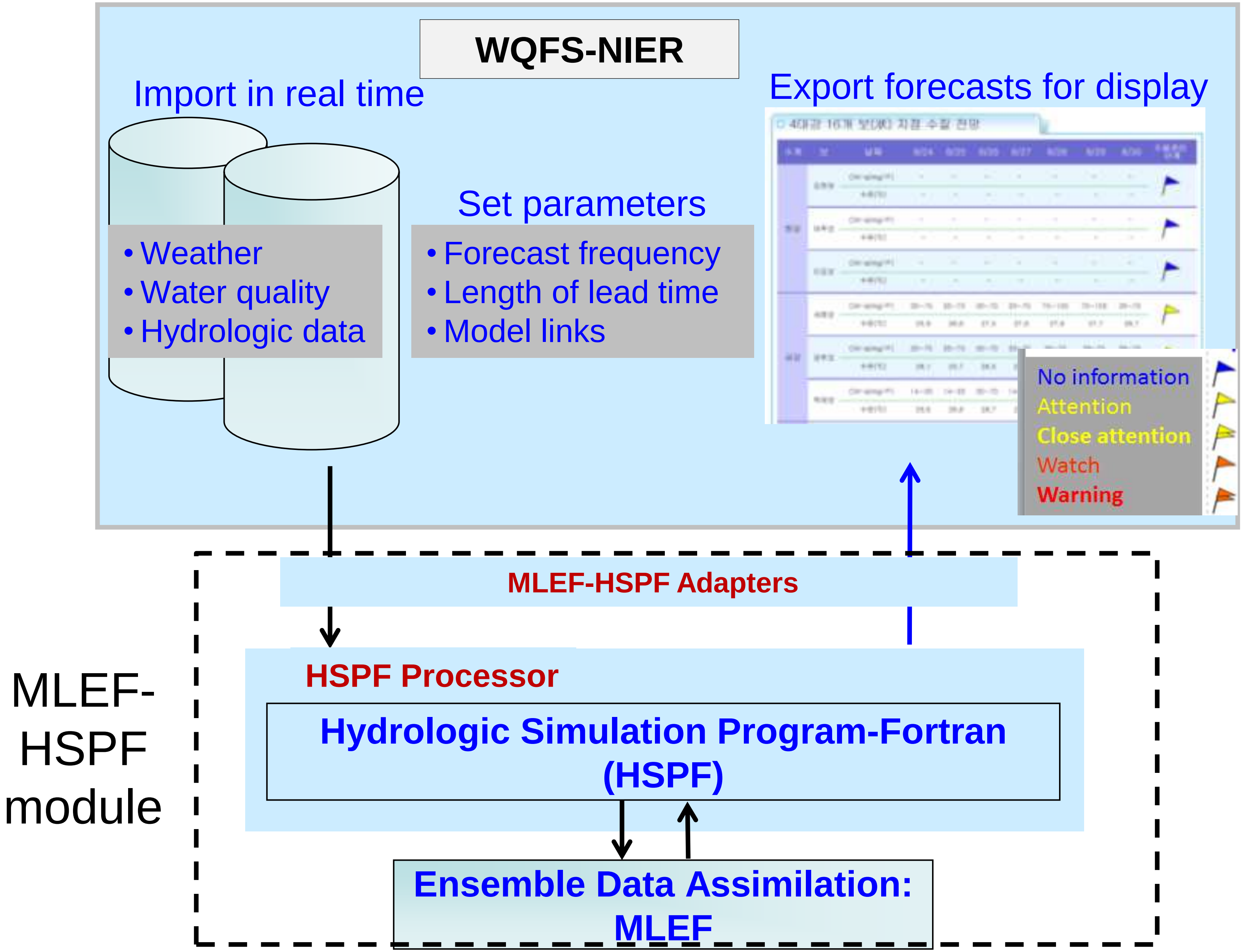


Obs from 4 (121, 109, 118, 127) stations assimilated



MLEF-HSPF

- Plugin module for the Water Quality Forecast System at the National Institute of Environmental Research (WQFS-NIER)
  - FEWS-based
- Module components
  - MLEF-HSPF program
  - HSPF processor
    - Interfaces MLEF-HSPF with HSPF
  - MLEF-HSPF adapter
    - Interfaces MLEF-HSPF with WQFS-NIER


WQFS-NIER
Import in real time
• Weather
• Water quality
• Hydrologic data
Set parameters
• Forecast frequency
• Length of lead time
• Model links
Export forecasts for display
No information
Attention
Close attention
Watch
Warning
MLEF-HSPF Adapters
HSPF Processor
Hydrologic Simulation Program-Fortran (HSPF)
Ensemble Data Assimilation: MLEF
MLEF-HSPF module

Ongoing work

- Multi-catchment, multi-basin evaluation
  - Data analysis
  - Hindcasting
- Operational implementation of MLEF-HSPF
  - Configuration for WQFS-NIER

River basin	Number of Catchment	Number of monitoring stations
Han	9	86
Keum	7	21
Yeongsan	6	24

Conclusions & research questions

- DA is generally effective in improving accuracy in water quality prediction
  - Correction of model biases as part of the observation equation is important
- MLEF handles nonlinear observation equations very well
- Improvement is larger for Banbyeon (natural) than Kumho (urbanized)
- How underdetermined is the inverse problem? How to reduce?
- Toward ensemble forecasting
  - Verify and improve the quality (in particular, reliability and spatio-temporal consistency) of updated ensemble ICs

Thank you

For more information:

sunghee@uta.edu