Hindcasting Water Levels for EDEN Gaging Stations, 2000-2006

Paul Conrads
USGS South Carolina Water Science Center
Ed Roehl

Advanced Data Mining

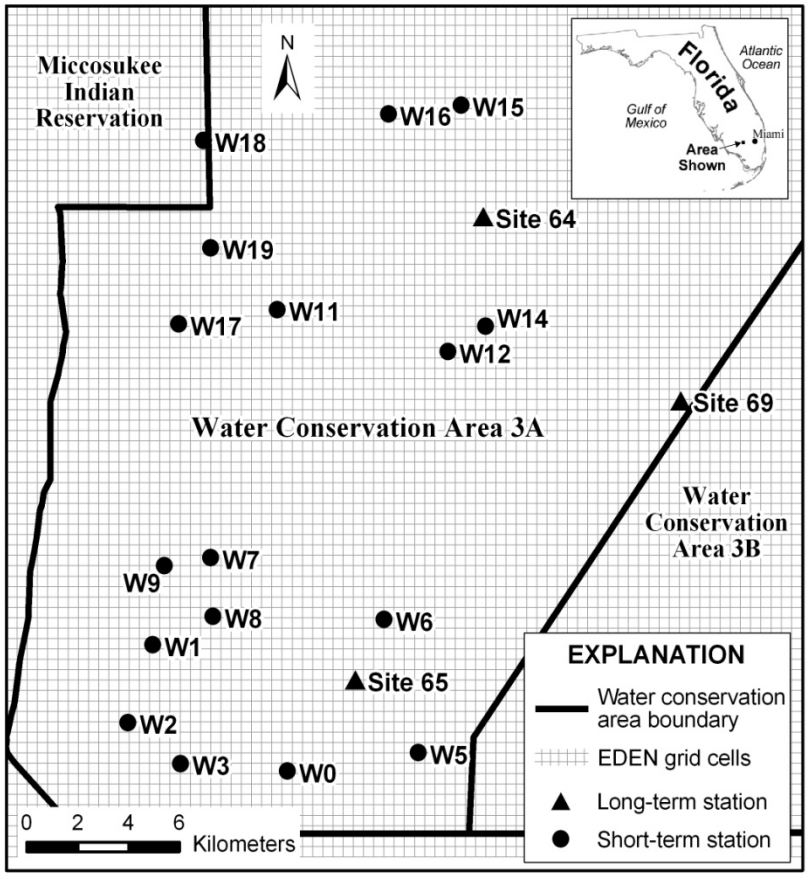
GEER 2008
Naples, FL
July 30, 2008

What's on tap?

- Methods development in WCA3a
- Application to EDEN
- Quiet desperation
- The fix
- Results and discussion

Problem : How to Estimate Water Depths at Ungaged Sites

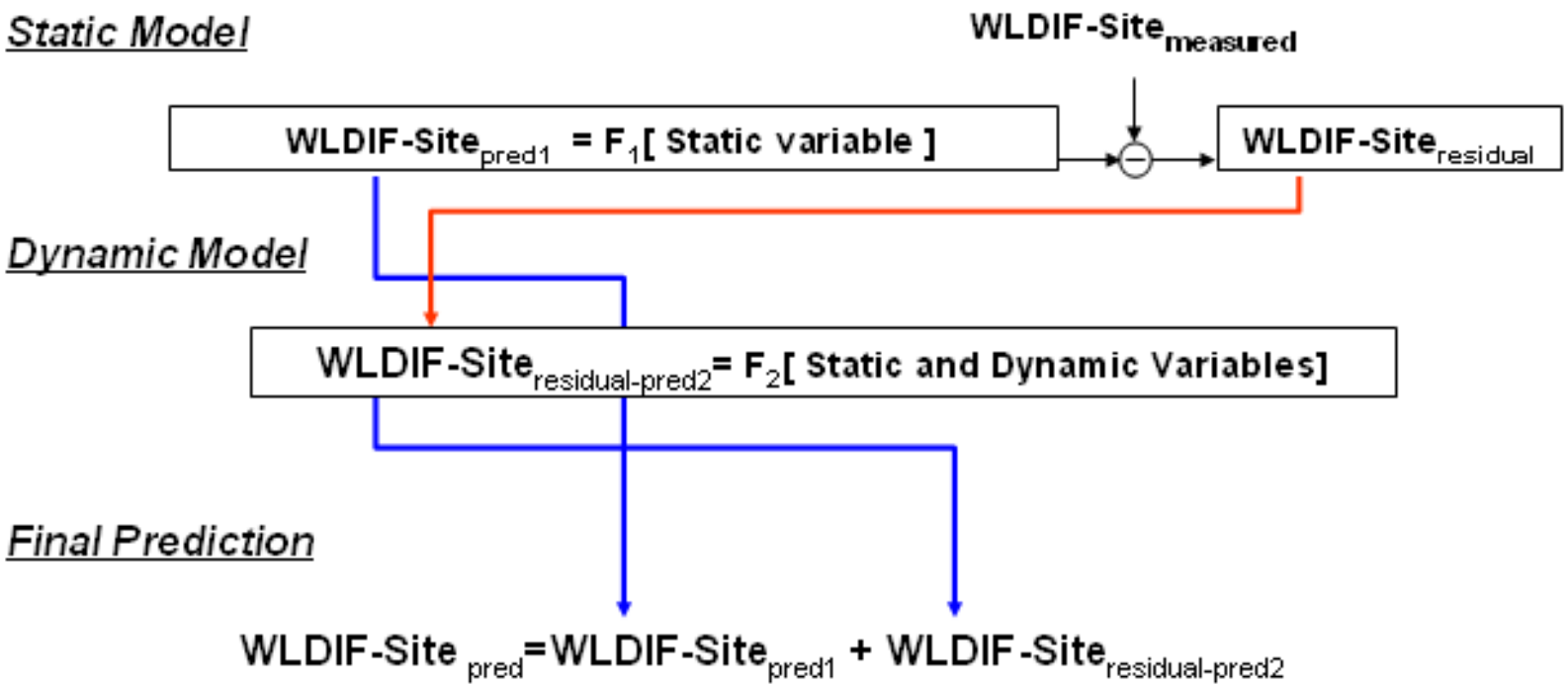
- Dataset – WCA 3a
  - Water-level from 3 sites
  - Water-depth data from 17 sites
  - EDEN grid and vegetation attributes
  - % prairie
  - % sawgrass
  - % slough
  - % upland
  - UTM North
  - UTM South



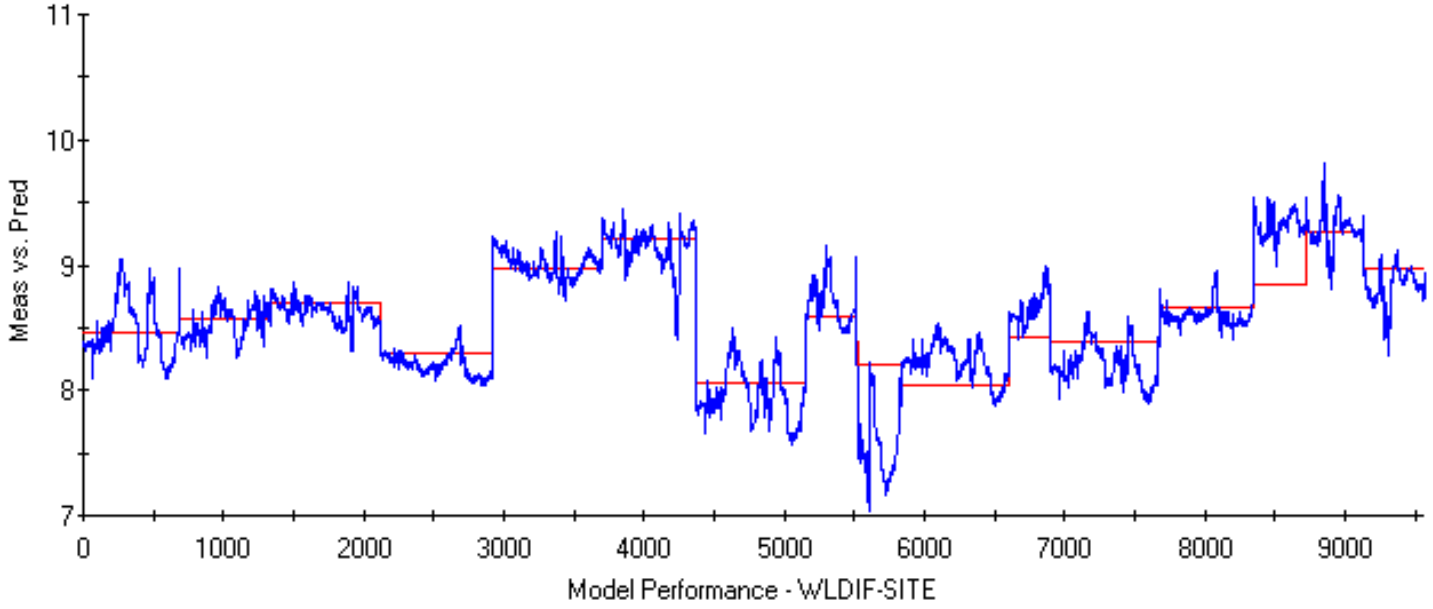
Approach

- Two step ANN model
  - First step: estimate mean water-depths using static model – “spatially interpolating” ANN scheme
  - Second step: estimate water-depths variability using dynamic variables

Two-step Model



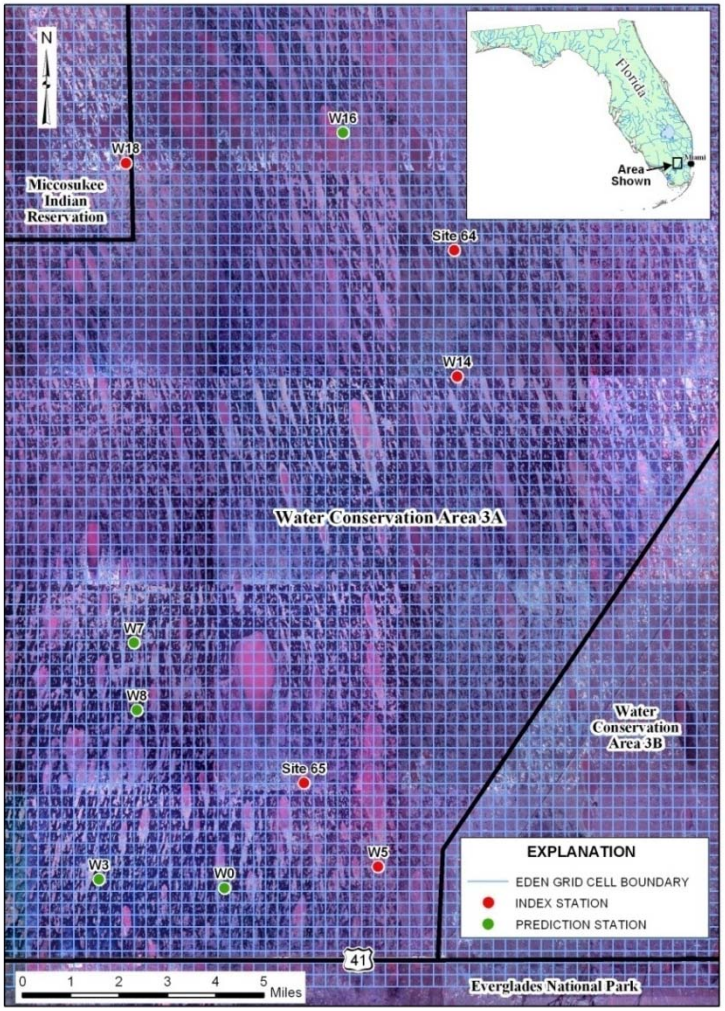
Static Model Results



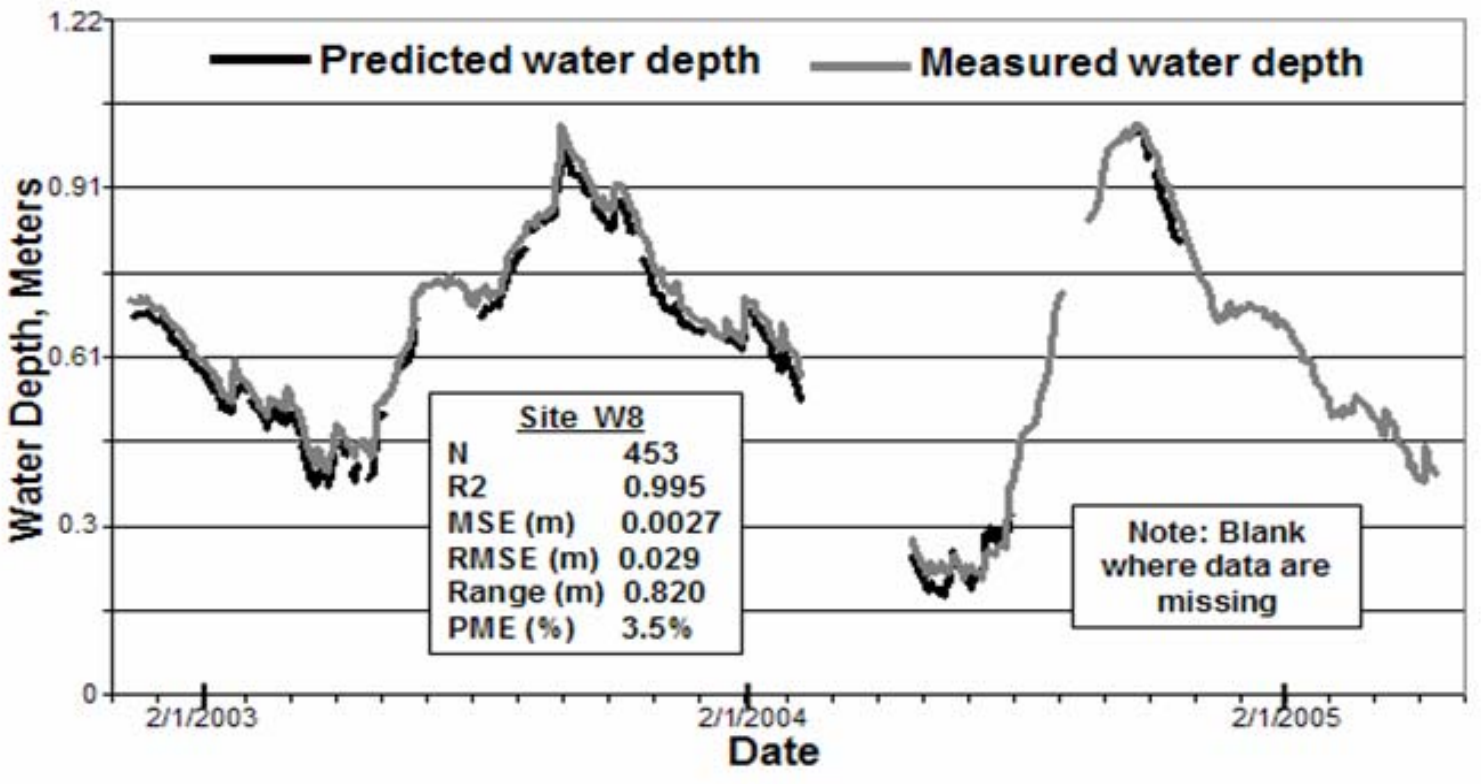
- Each “step” represents a different site
- Model able to generalize water level difference but not the variability

Dynamic Model

- 5 “index” stations (red dots)
- Combination of static and dynamic data
- 5 validation stations (green dots)



Final Model Results



Hindcasting 25 EDEN Sites

Existing EDEN sites : 7 years (61,400 data points)

New EDEN sites : 4 -12 months of data (925- 8,760 data points)

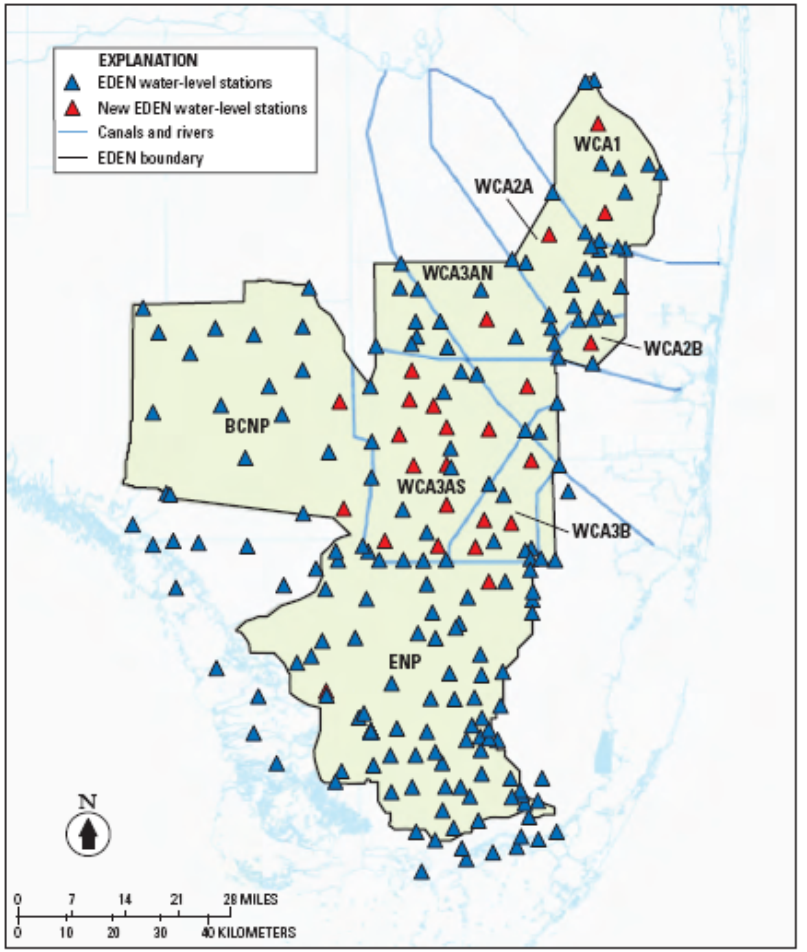
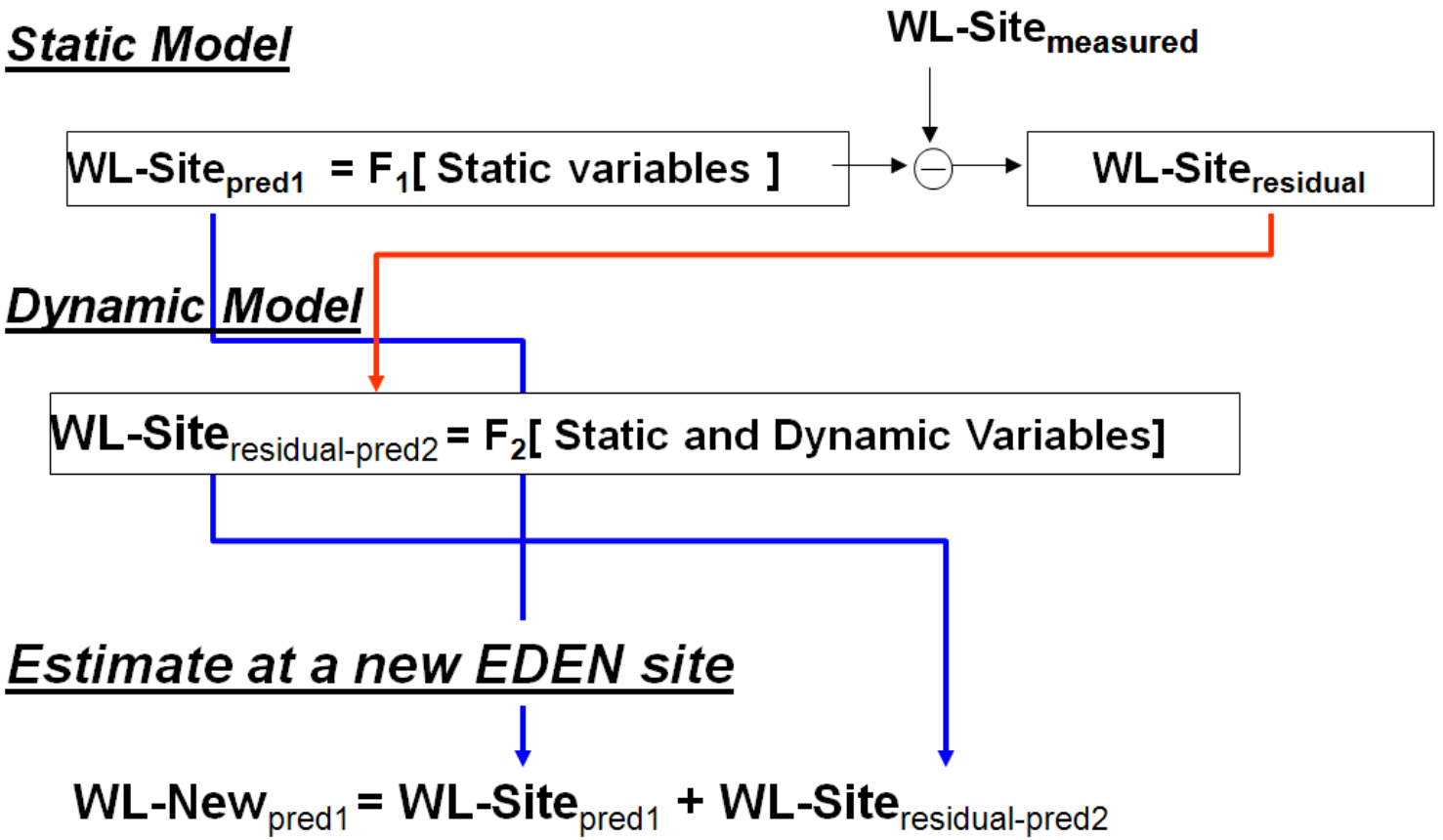


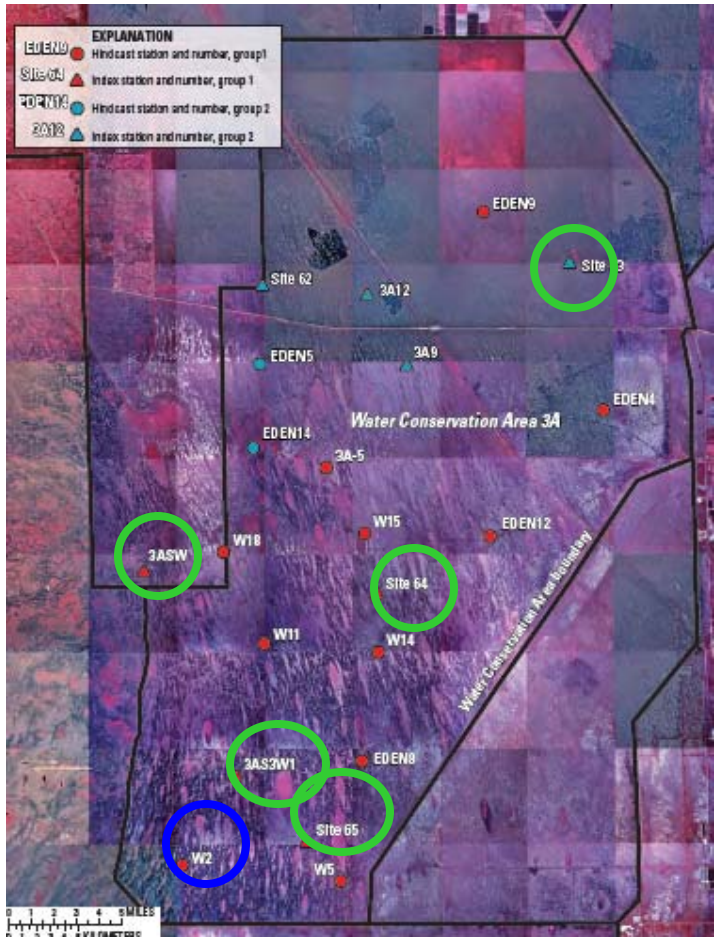
Figure 1. Locations of Everglades Depth Estimation Network (EDEN) gaging stations in southern Florida (modified from Pearlstine and others, 2007).

Approach: Similar 2-Step Model

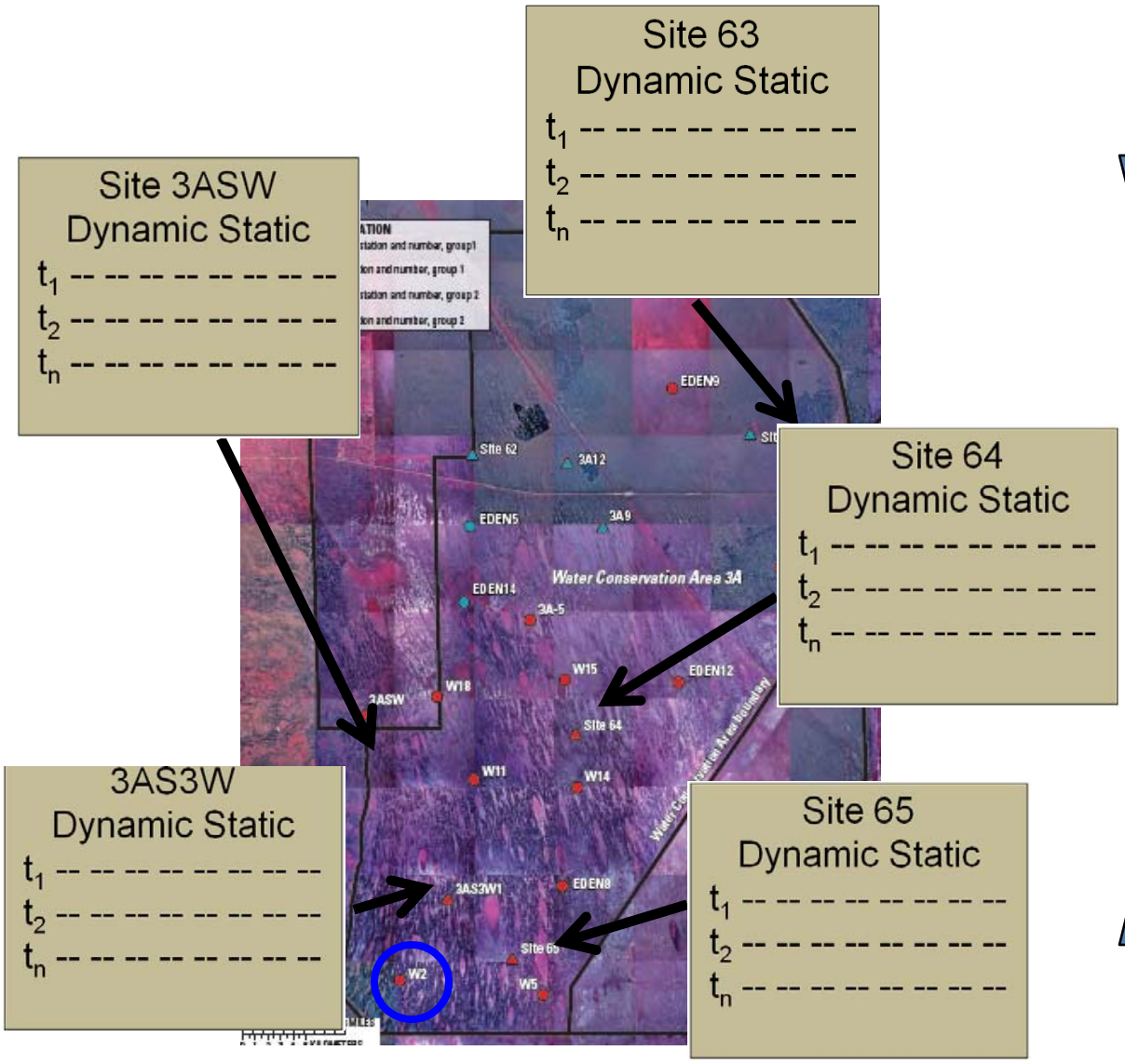


Approach: Similar 2-Step Model

- Separate models for each area
- Number of potential inputs sites reduced using dynamic clustering
- Hindcast example – W2
- Input Sites
  - Sites 63, 64, 65, 3ASW3, and 3AWS



Spatially Interpolating ANN Model



Stacked dataset

Site 63
Dynamic Static
t_1 -- -- -- -- -- -- -- --
t_2 -- -- -- -- -- -- -- --
t_n -- -- -- -- -- -- -- --

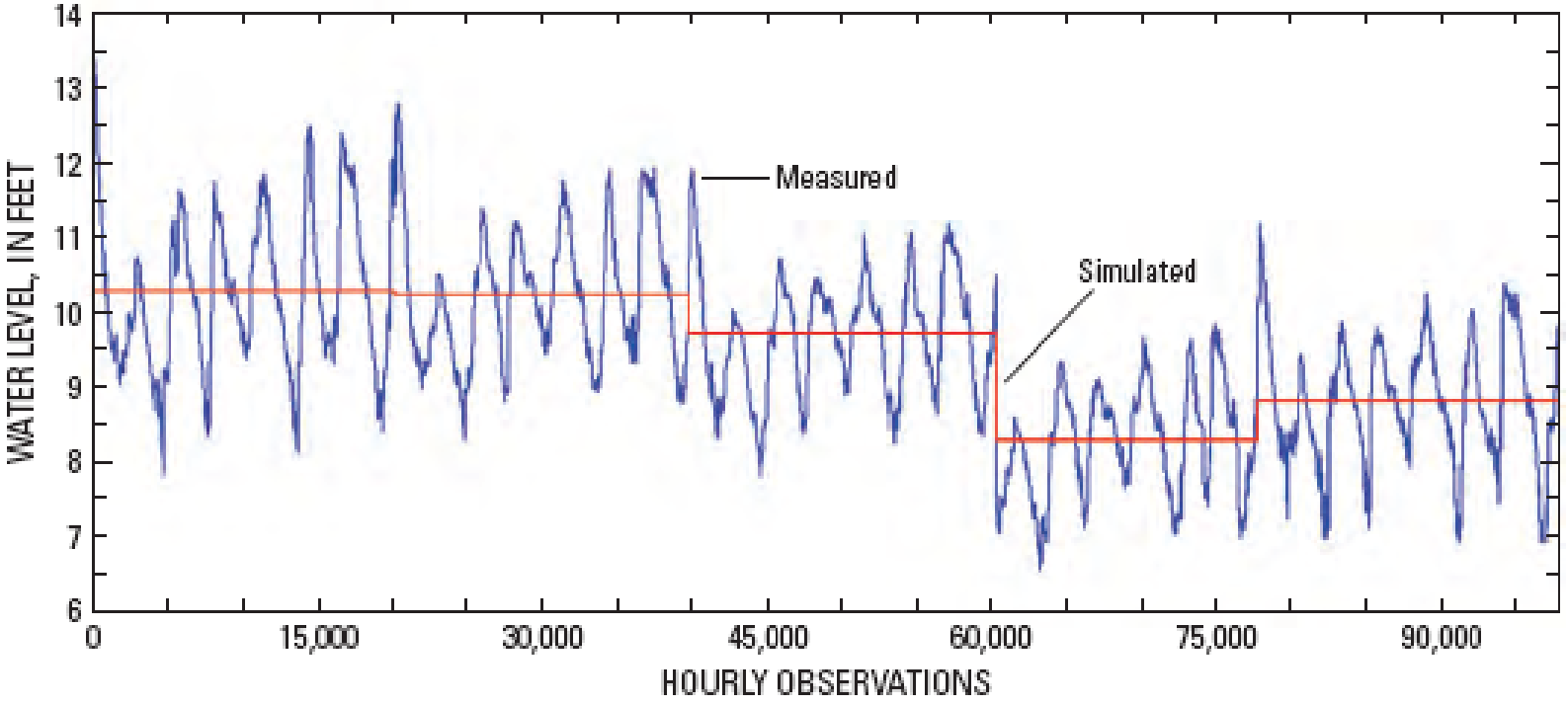
Site 64
Dynamic Static
t_1 -- -- -- -- -- -- -- --
t_2 -- -- -- -- -- -- -- --
t_n -- -- -- -- -- -- -- --

Site 65
Dynamic Static
t_1 -- -- -- -- -- -- -- --
t_2 -- -- -- -- -- -- -- --
t_n -- -- -- -- -- -- -- --

Site 3ASW
Dynamic Static
t_1 -- -- -- -- -- -- -- --
t_2 -- -- -- -- -- -- -- --
t_n -- -- -- -- -- -- -- --

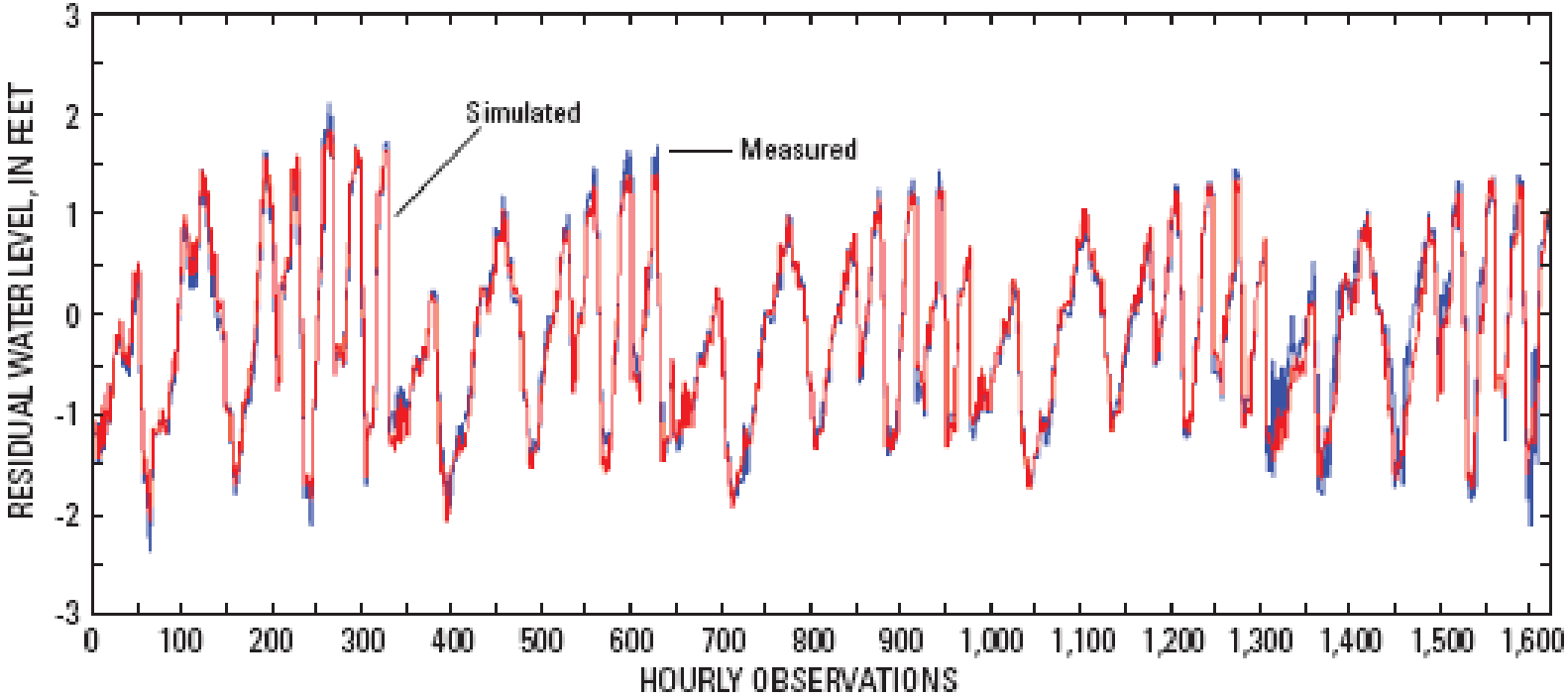
3AS3W
Dynamic Static
t_1 -- -- -- -- -- -- -- --
t_2 -- -- -- -- -- -- -- --
t_n -- -- -- -- -- -- -- --

Static Model Results



Site 63 Site 64 Site 65 3AS3W 3ASW

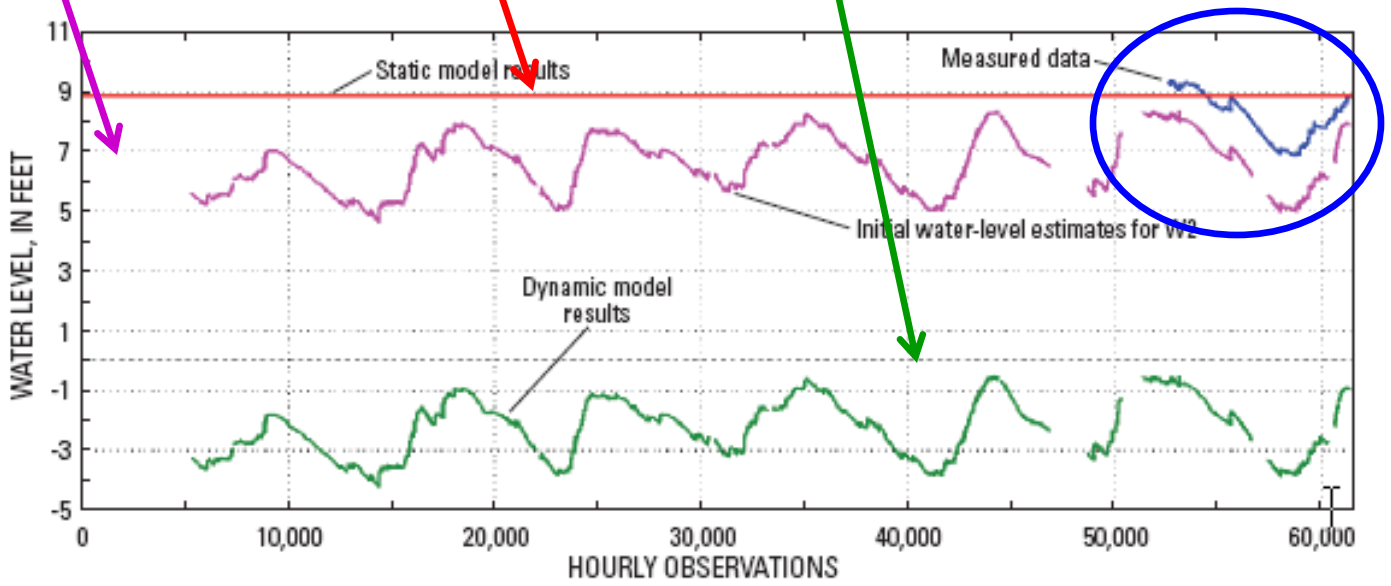
Dynamic Model Results



Site 63 Site 64 Site 65 3AS3W 3ASW

Initial Water Level Estimate – W2
"Quiet Desperation"

$$\text{WL-New}_{\text{pred1}} = \text{WL-Site}_{\text{pred1}} + \text{WL-Site}_{\text{residual-pred2}}$$

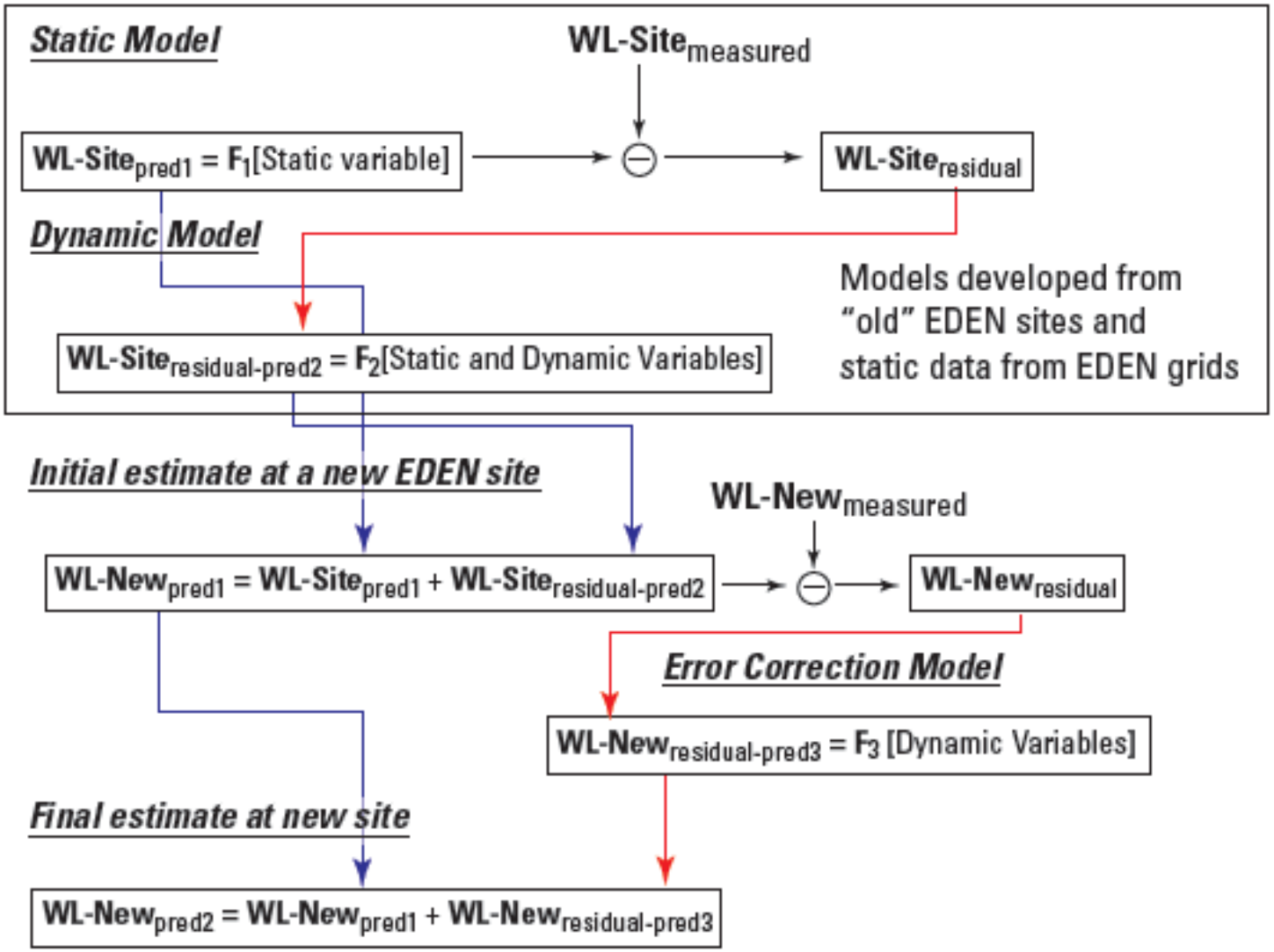


Data for W2 not used to train models

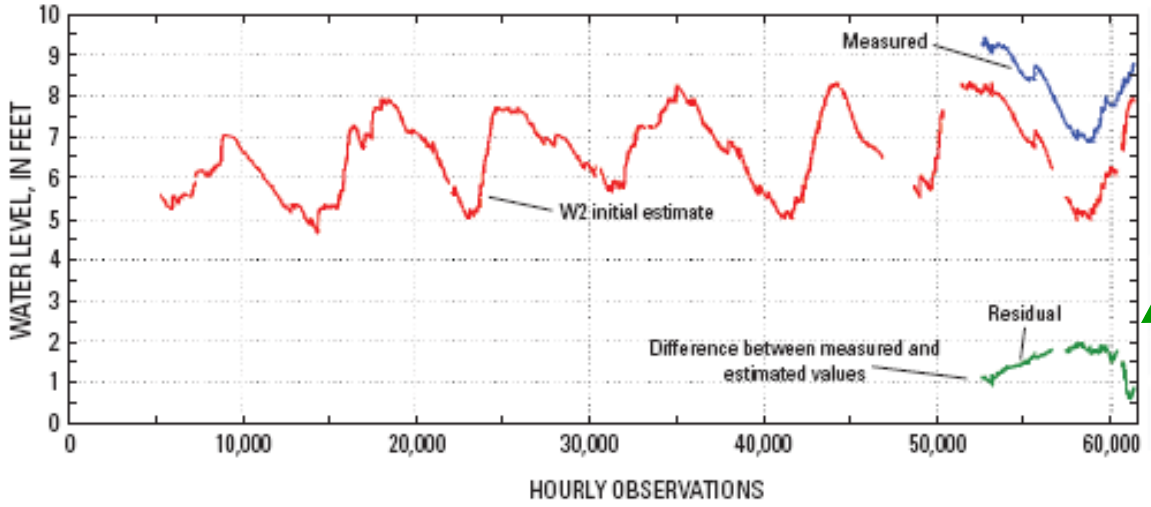
Possible Causes

- Datum confusion
- Gage offset
- W2 location with respect to input sites
- Limited information content of static variables
- Did not use difference from a standard signal

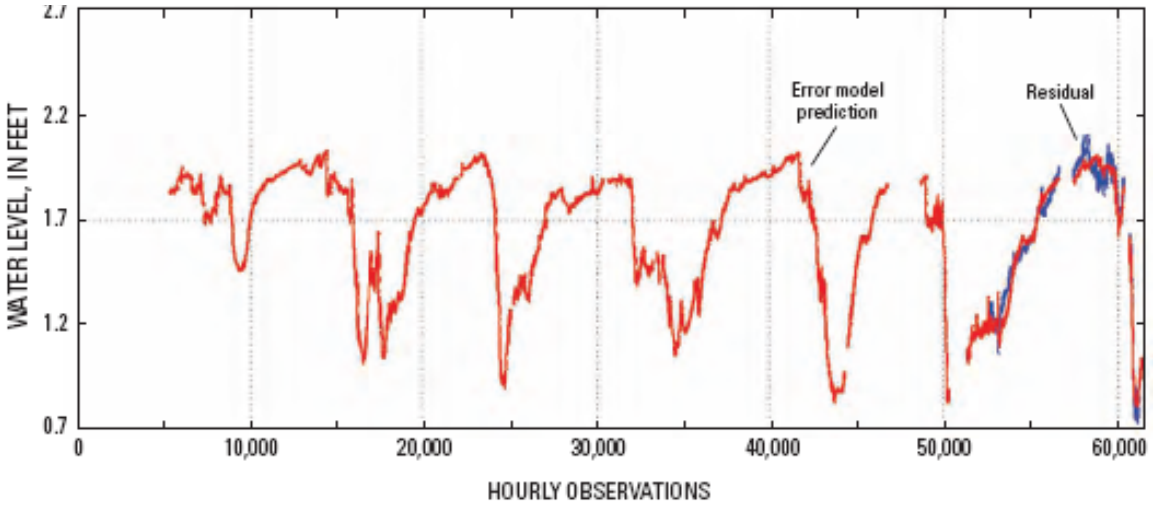
Third-step Model: Error correction



Error Correction Model

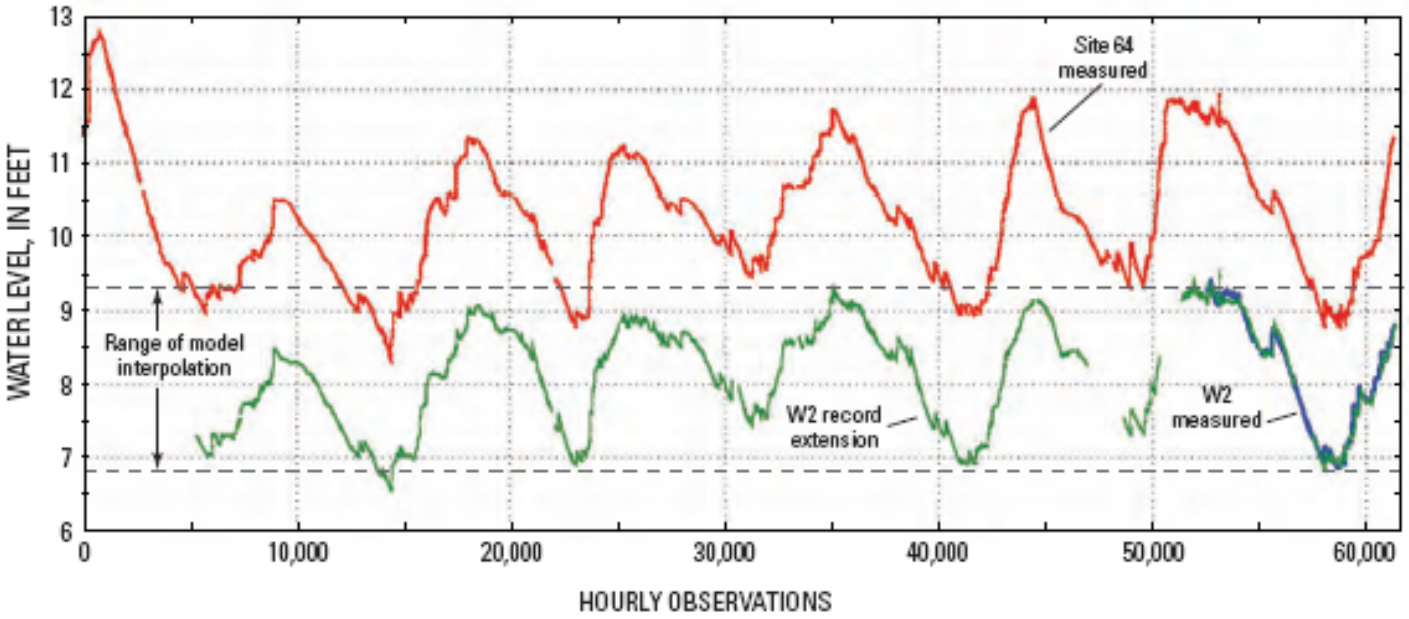


Model initial estimate error



Only model to use data from the site

Final Water-Level Prediction



Hindcast Results

Summary Statistics

- n: 925 – 43,409
- R^2: 0.821 – 0.999
- Mean error: -.06 to .03 ft
- RMSE: .001 - .134 ft
- Percent Model Error: 0.1 – 5.2%

Site	n	R	R^2	Data range Min, in ft	Data range Max, in ft	ME	RMSE	PME
Water Conservation Area 1 (fig. 15)								
North_CA1	41,721	0.943	0.889	13.82	16.38	-0.046	0.134	5.2
South_CA1	43,409	0.991	0.983	12.85	15.90	0.008	0.088	2.9
Water Conservation Area 2 (fig. 17)								
EDEN11	2,637	0.950	0.902	10.97	13.19	-0.054	0.044	2.0
EDEN13	1,968	0.955	0.912	7.07	7.69	-0.018	0.006	1.0
Water Conservation Area 3A (fig. 20)								
3A-5	5,684	0.998	0.995	8.15	10.10	0.002	0.001	0.1
EDEN4	2,399	0.999	0.998	6.96	10.30	0.001	0.002	0.1
EDEN5	1,653	0.999	0.999	8.01	10.20	0.004	0.001	0.1
EDEN8	2,442	0.999	0.999	6.79	9.19	0.001	0.001	0.0
EDEN9	925	0.999	0.998	7.69	10.55	-0.006	0.004	0.1
EDEN12	7,648	0.999	0.998	6.84	9.59	-0.002	0.001	0.0
EDEN14	969	0.906	0.821	8.66	9.59	0.011	0.010	1.1
W2	7,648	0.998	0.995	6.86	9.42	0.008	0.003	0.1
W5	7,648	0.999	0.998	6.84	9.59	-0.001	0.001	0.0
W11	7,628	0.999	0.998	7.08	10.08	-0.012	0.002	0.1
W14	[illegible]	0.998	0.997	7.00	10.00	-0.013	0.003	0.1
W15	[illegible]	0.999	0.998	7.47	9.67	-0.001	0.001	0.0
W18	7,628	0.998	0.996	7.92	10.21	-0.000	0.002	0.1
Water Conservation Area 3B (fig. 26)								
[illegible]	5,684	0.996	0.993	4.59	6.16	0.002	0.001	0.1
11-9	5,684	0.996	0.992	5.29	6.42	0.009	0.001	0.1
EDEN7	2,419	0.998	0.996	5.24	6.96	0.008	0.001	0.1
[illegible]	[illegible]	0.995	0.990	[illegible]	6.32	0.007	0.002	0.1
Big Cypress National Preserve (fig. 29)								
EDEN1	3,864	0.960	0.921	7.13	7.82	0.018	0.003	0.4
EDEN6	1,591	0.984	0.968	8.76	10.73	0.030	0.007	0.4
Everglades National Park (fig. 31)								
EDEN3	5,294	0.989	0.978	0.07	1.75	-0.042	0.006	0.4
Met1	1,238	0.994	0.989	5.23	5.76	0.000	0.000	0.0

Summary

- ANNs able to accurately predict water levels at the new EDEN sites
- Hindcasts accuracies could be improved with additional data
- Use of static and dynamic variable produce a multi-variate “kreiging” of water levels
- Additional “static” categorical variables may eliminate need for the error correction model

Questions

Paul Conrads

USGS-South Carolina Water Science Center

pconrads@usgs.gov

USGS
science for a changing world