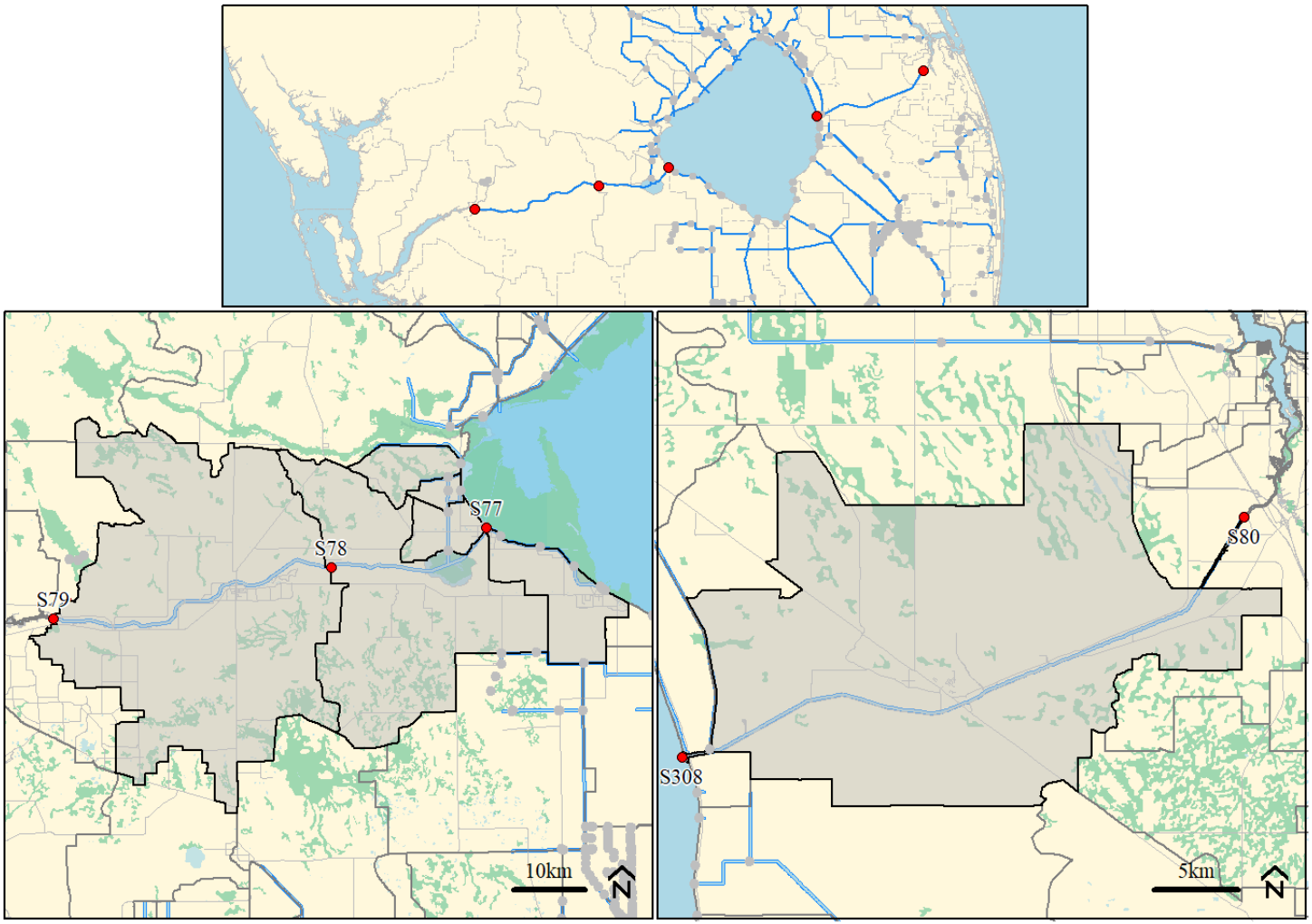
Lake Okeechobee System Operating Manual (WQ Subteam)

DRAFT - Estuary Nutrient Loading Model

FDEP - Office of Water Policy and Ecosystem Restoration

September 18, 2020

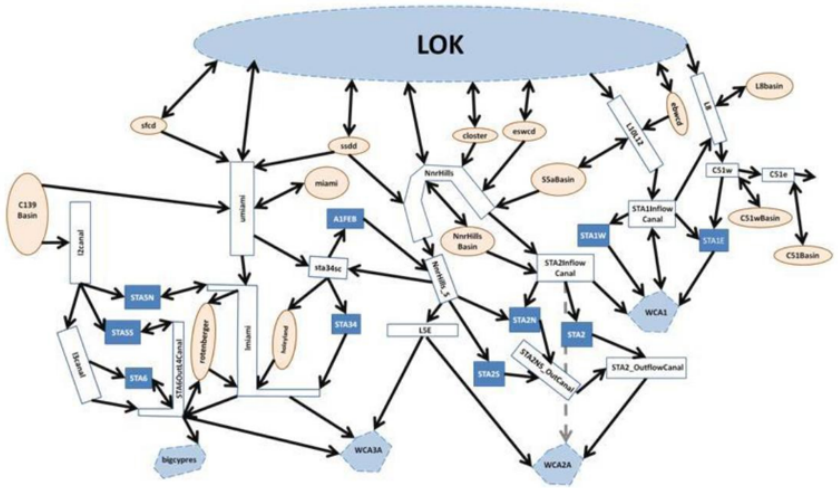

S77
S78
S79
S80
S308
10km
5km
N
N

S-79 Water Quality Model

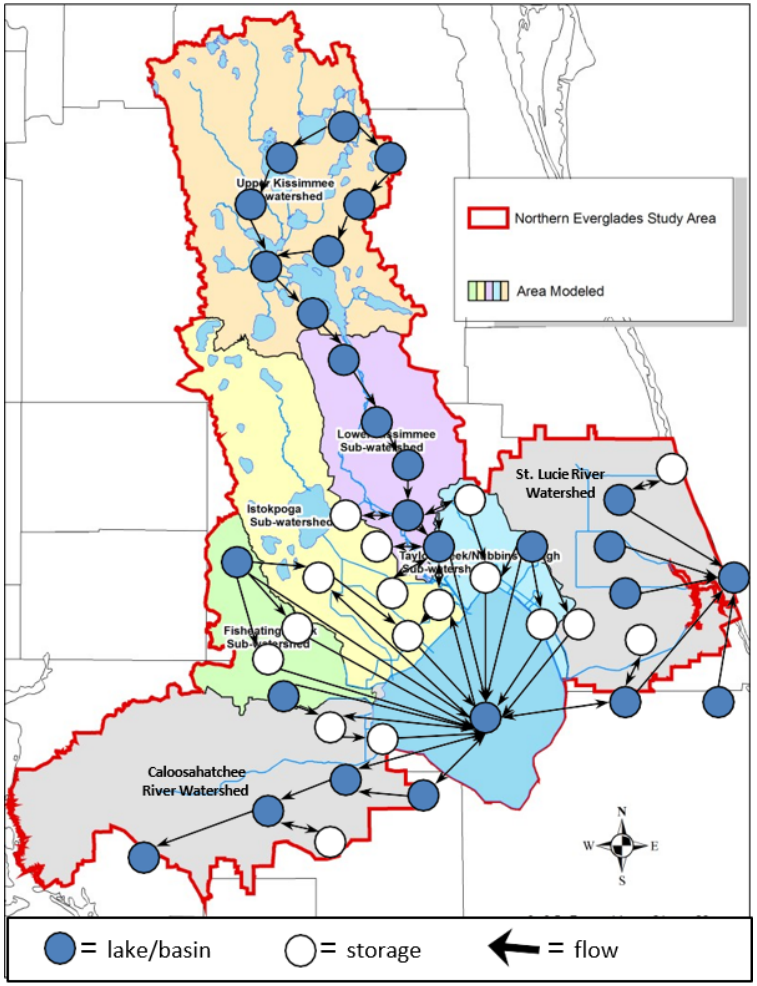
- **Goal:** Develop a series of water quality models based on hydrodynamic indicators to be used in planning model scenario evaluation for LOSOM using RSMBN.
- **Period of Record:** May 1980 – April 2019 (WY1981 – 2019)
    - Based on available data.
    - Years with major hurricanes were excluded.
- **Parameters of Interest:** Total Phosphorus and Total Nitrogen.
- **Predictor Variables:** Discharge (S77, S78, S79, C43 Basin) converted from $ft^3\ s^{-1}$ to Acre-Ft d^{-1} and Lake Okeechobee stage elevation were considered.
- **Statistical Modeling:**
    - Multiple regression models using training and testing datasets (70:30).
        - Training dataset: randomly sampled 70% of monthly data
        - Testing dataset: remaining 30% was used for model testing
    - Verified with k-fold cross-validation linear modeling.

Regional Simulation Model - Basins (RSMBN)

- A link-node application of the Regional Simulation Model (RSM) specific to Lake Okeechobee and Basins.
  - >100 basins/lakes/canals represented
  - >150 connections represented
- Used in prior planning efforts.

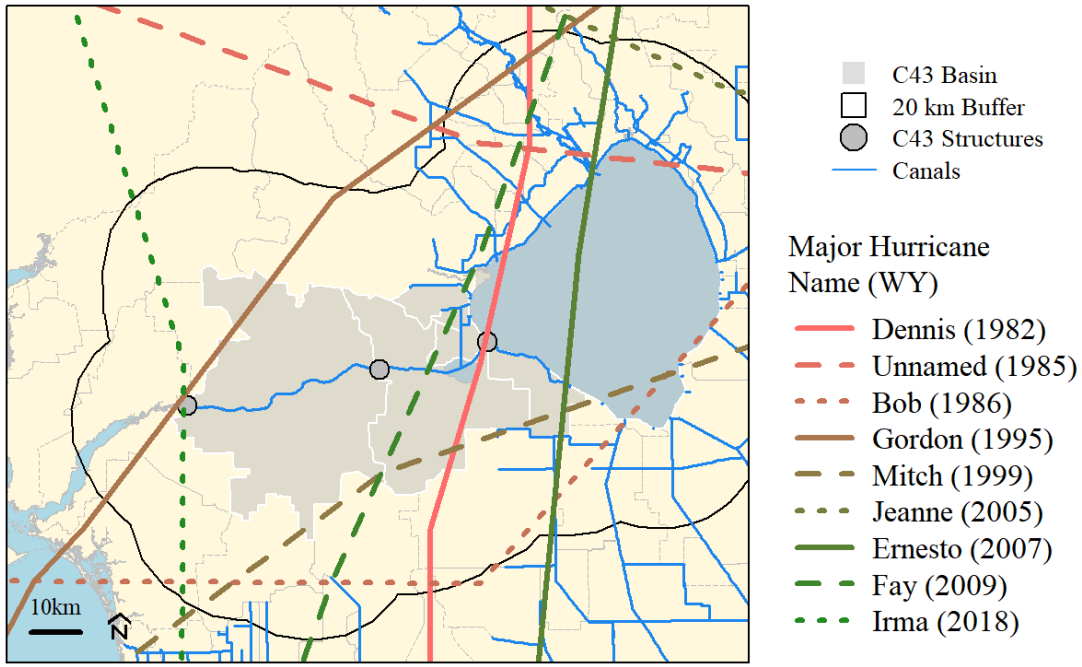


RSMBSN Link-Node Routing Diagram: Initial Operating Regime Baseline Simulation

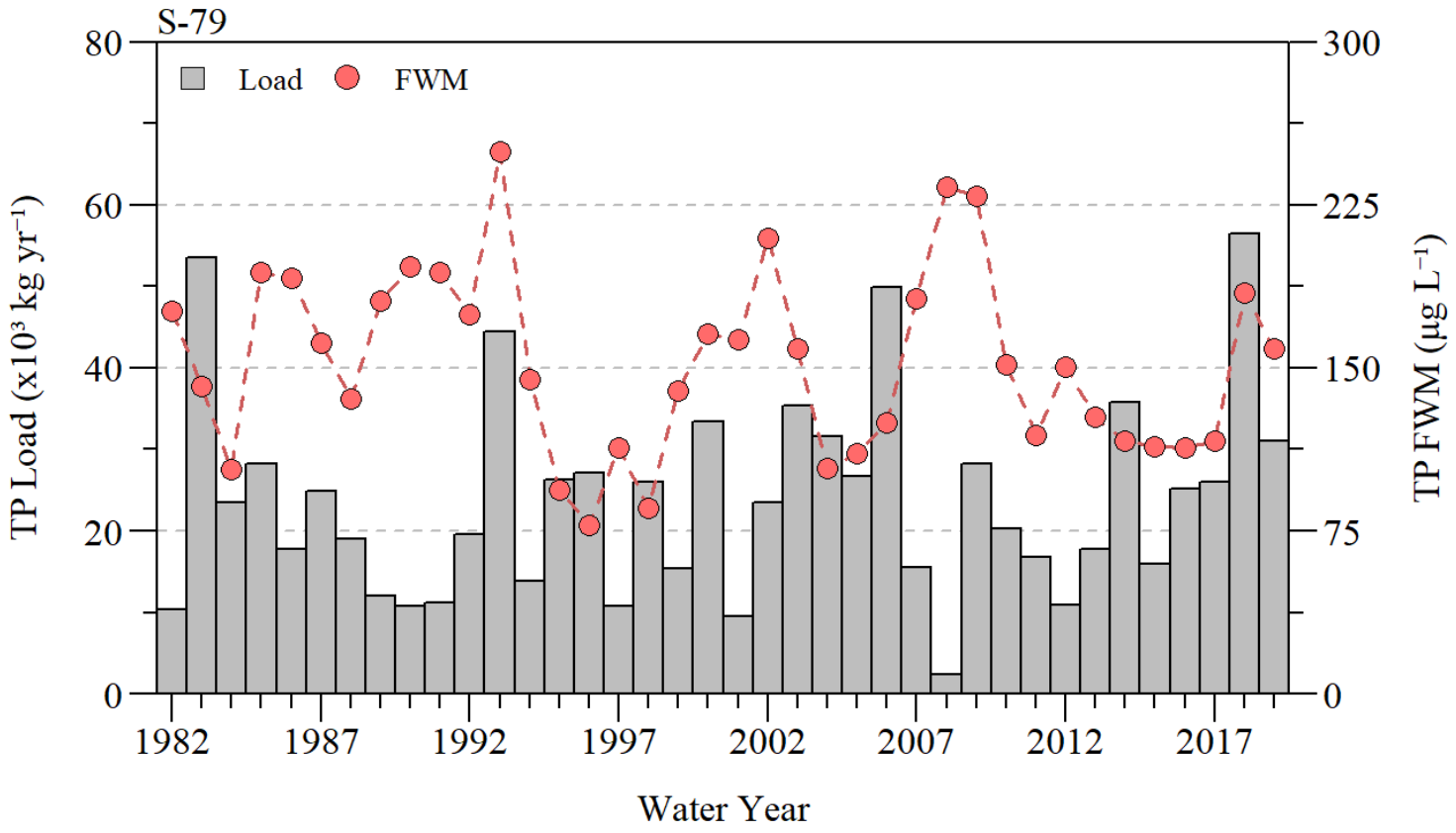


Geographical representation of RSM-BN

- Annual TP and TN loads were estimated by interpolating water quality concentration daily from grab samples collected at each respective structure during days of observed flow (consistent with SFWMD Nutrient Load Program).
    - Period of Record was restricted to WY 1982 – 2019.
- Major hurricane years were excluded from the dataset.
    - Major hurricane years were identified by major storms centerline that passed within 20 km of the C43 Basins and Lake Okeechobee.



S-79 Water Quality Model (Total Phosphorus)



S-79 Water Quality Model (Total Phosphorus)

- TP load was square-root transformed to fit the assumptions of linear modeling
- Model assumptions and verified
  - GVLMA (Global Stats = 4.37, ρ=0.36)
- Variance inflation factors (VIF) evaluated for model
- Residuals check for autocorrelation (Breusch-Godfrey test)
  - Breusch-Godfrey (LM test = 0.27, df = 1, ρ=0.60)
- Final Model:

$$\sqrt{TPLoad_{S79}} = Q_{C43Basin} + Q_{S77} + MeanLakeStage$$

S-79 Water Quality Model (Total Phosphorus)

$$\sqrt{TPLoad_{S79}} = Q_{C43Basin} + Q_{S77} + MeanLakeStage$$

- Model Output:

Parameter	Estimate	Standard Error	t-value	ρ-value	VIF
(Intercept)	30.68	181.80	0.17	0.87	---
Q_{C43}	2.17×10^{-4}	7.33×10^{-5}	2.95	<0.01	2.96
Q_{S77}	4.16×10^{-5}	3.07×10^{-5}	1.36	0.19	1.59
Mean Lake Stage	17.07	15.70	1.09	2.89	2.89

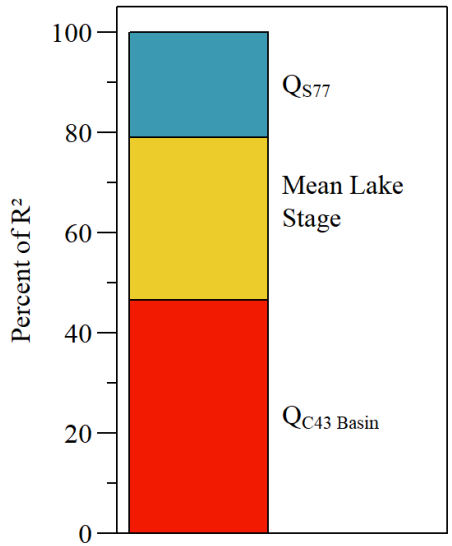
Residual standard error (σ): 65.71 on 16 degrees of freedom

Multiple R^2: 0.79; Adj R^2: 0.75

F-statistic: 20.33 on 3 and 16 DF; ρ-value<0.01

S-79 Water Quality Model (Total Phosphorus)

$$\sqrt{TPLoad_{S79}} = Q_{C43Basin} + Q_{S77} + MeanLakeStage$$



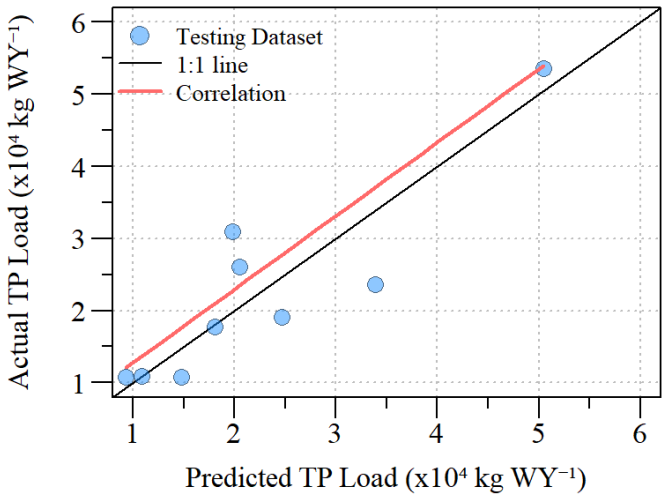
Relative importance of each predictor calculated by partitioning R^2 by averaging sequential sums of squares over all orders of regressors (Lindeman et al 1979). All metrics are normalized to a sum of 100%.

Relative Importance Metrics for the S79 TP Load annual model.

Predictor	Percent of R^2
Q_{C43}	46.5
Q_{S77}	21.0
Mean Lake Stage	32.3

S-79 Water Quality Model (Total Phosphorus)

$$\sqrt{TPLoad_{S79}} = Q_{C43Basin} + Q_{S77} + MeanLakeStage$$



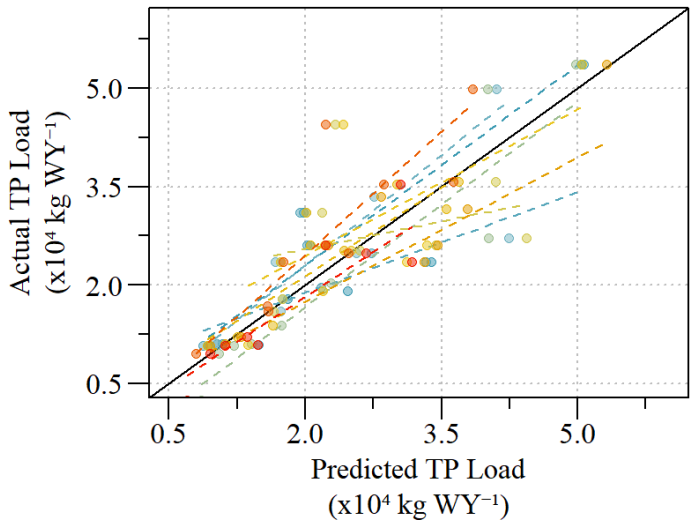
Actual versus predicted TP loads at S-79 based on predictive model. Actual and predicted concentration were highly correlated (Spearman's correlation: r=0.83, ρ<0.01).

Train:Test

- R^2: 0.79
- RMSE: 65.71
- Mean Absolute Present Error: 21 %
- Min-Max Accuracy: 82 %

S-79 Water Quality Model (Total Phosphorus)

$$\sqrt{TPLoad_{S79}} = Q_{C43Basin} + Q_{S77} + MeanLakeStage$$



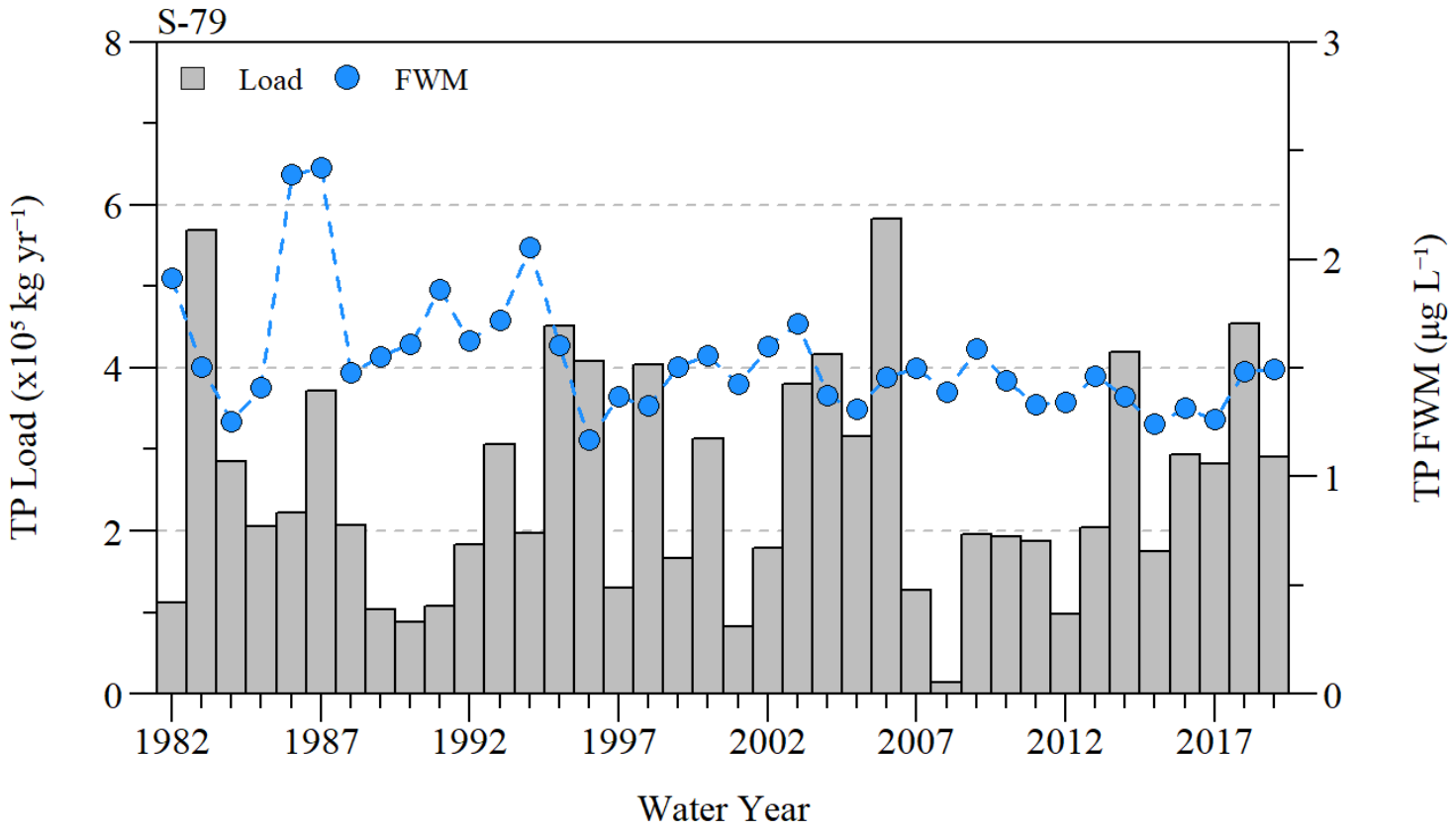
Actual versus predicted TP loads at S-79 with each k-model presented.

k-fold (k=10)

Cross-validation error (average k errors)

- R^2: 0.79
- RMSE: 61.0
- Mean Absolute Present Error: 24 %
- Min-Max Accuracy: 82 %

S-79 Water Quality Model (Total Nitrogen)



S-79 Water Quality Model (Total Nitrogen)

- No response variable transformation performed
- Model assumptions and verified
  - GVLMA (Global Stats = 4.77, ρ=0.31)
- Variance inflation factors (VIF) evaluated for model
- Residuals check for autocorrelation (Breusch-Godfrey test)
  - Breusch-Godfrey (LM test = 0.73, df = 1, ρ=0.39)
- Final Model:

$$TNLoad_{S79} = Q_{C43Basin} + Q_{S77} + MeanLakeStage$$

S-79 Water Quality Model (Total Nitrogen)

$$TNLoad_{S79} = Q_{C43Basin} + Q_{S77} + MeanLakeStage$$

- Model Output:

Parameter	Estimate	Standard Error	t-value	ρ-value	VIF
(Intercept)	5.20×10^4	1.14×10^6	-0.05	0.96	---
Q_{C43}	2.42	0.46	5.27	<0.01	2.96
Q_{S77}	1.16	0.19	6.02	<0.01	1.59
Mean Lake Stage	2.35×10^3	9.83×10^4	0.02	0.98	2.89

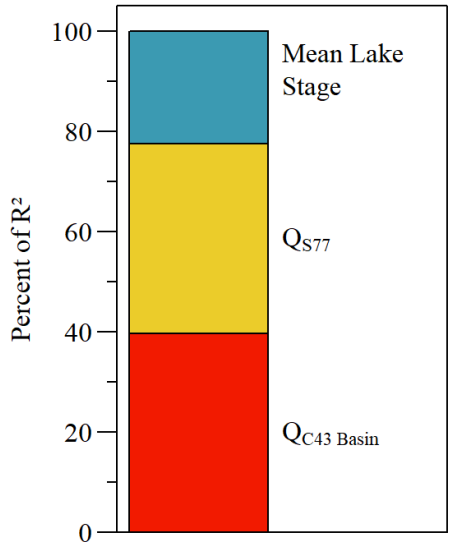
Residual standard error (σ): 4.11×10^5 on 16 degrees of freedom

Multiple R^2: 0.93; Adj R^2: 0.92

F-statistic: 73.87 on 3 and 16 DF; ρ-value<0.01

S-79 Water Quality Model (Total Nitrogen)

$$TNLoad_{S79} = Q_{C43Basin} + Q_{S77} + MeanLakeStage$$



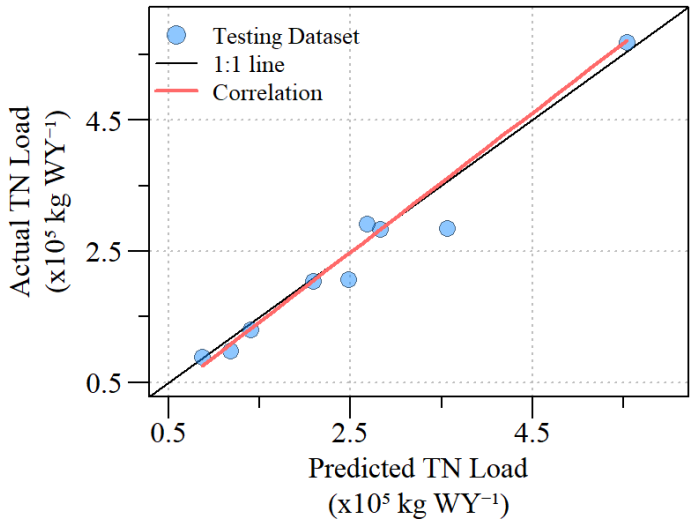
Relative importance of each predictor calculated by partitioning R^2 by averaging sequential sums of squares over all orders of regressors (Lindeman et al 1979). All metrics are normalized to a sum of 100%.

Relative Importance Metrics for the S79 TN Load annual model.

Predictor	Percent of R^2
Q_{C43}	39.6
Q_{S77}	37.9
Mean Lake Stage	22.5

S-79 Water Quality Model (Total Nitrogen)

$$TNLoad_{S79} = Q_{C43Basin} + Q_{S77} + MeanLakeStage$$



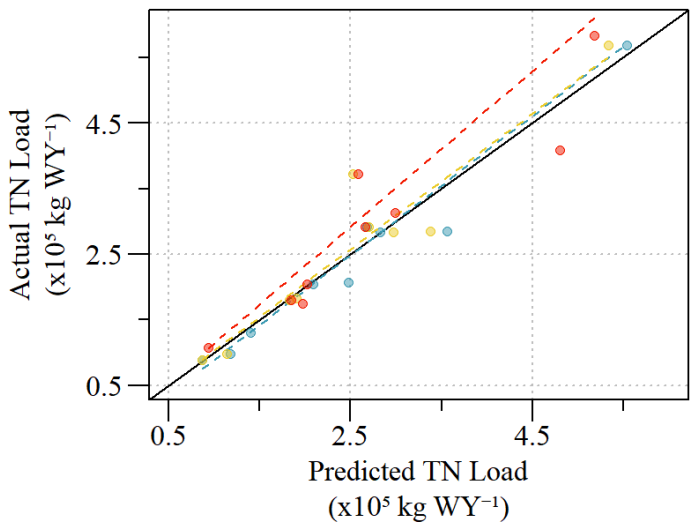
Actual versus predicted TP loads at S-79 based on predictive model. Actual and predicted concentration were highly correlated (Spearman's correlation: r=0.95, ρ<0.01).

Train:Test

- R^2: 0.93
- RMSE: 4.11×10^5
- Mean Absolute Present Error: 10 %
- Min-Max Accuracy: 92 %

S-79 Water Quality Model (Total Nitrogen)

$$TNLoad_{S79} = Q_{C43Basin} + Q_{S77} + MeanLakeStage$$



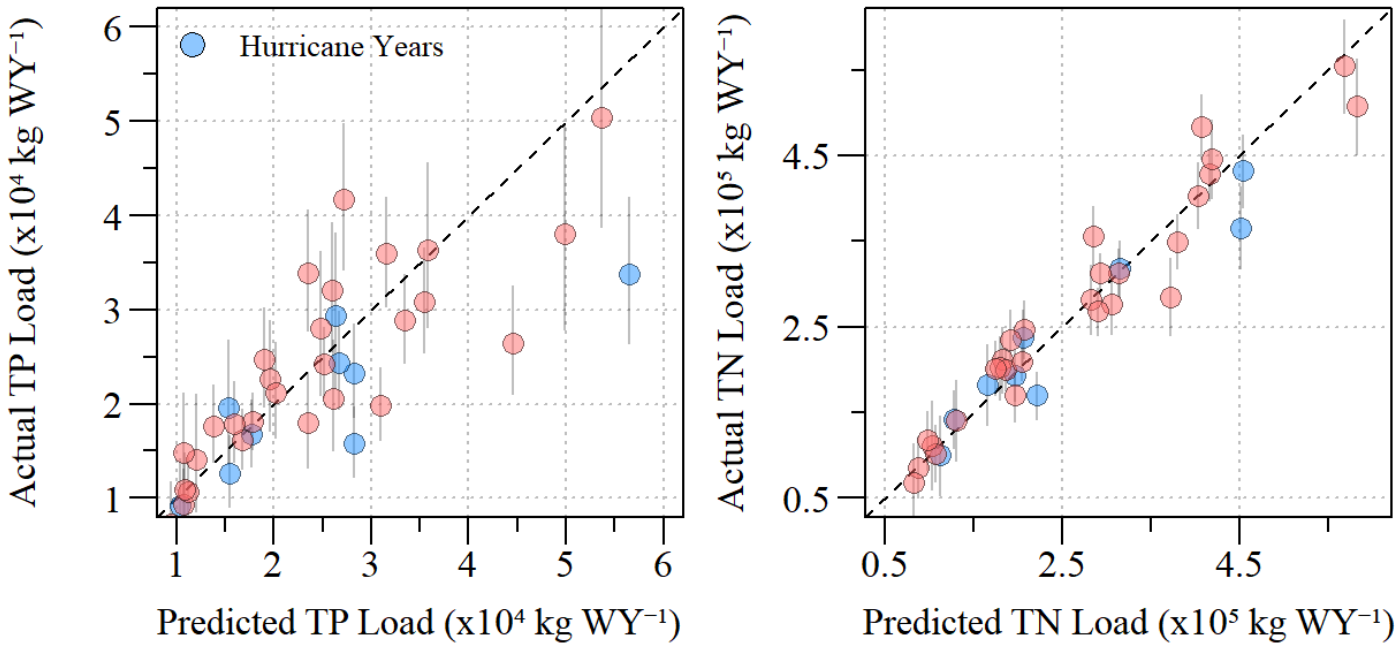
Actual versus predicted TP loads at S-79 with each k-model presented.

k-fold (k=10)

Cross-validation error (average k errors)

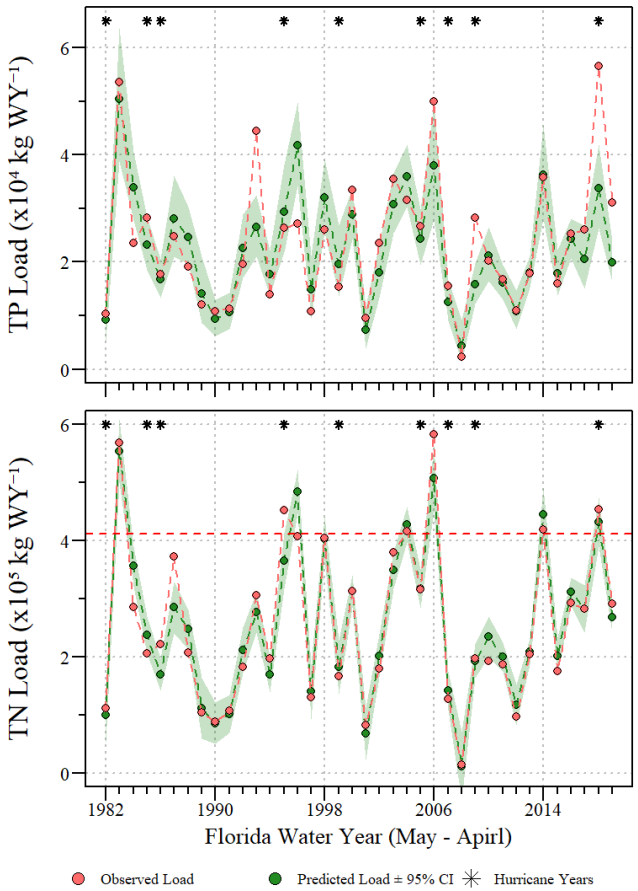
- R^2: 0.94
- RMSE: 3.57×10^5
- Mean Absolute Present Error: 11 %
- Min-Max Accuracy: 90 %

S-79 Water Quality Models



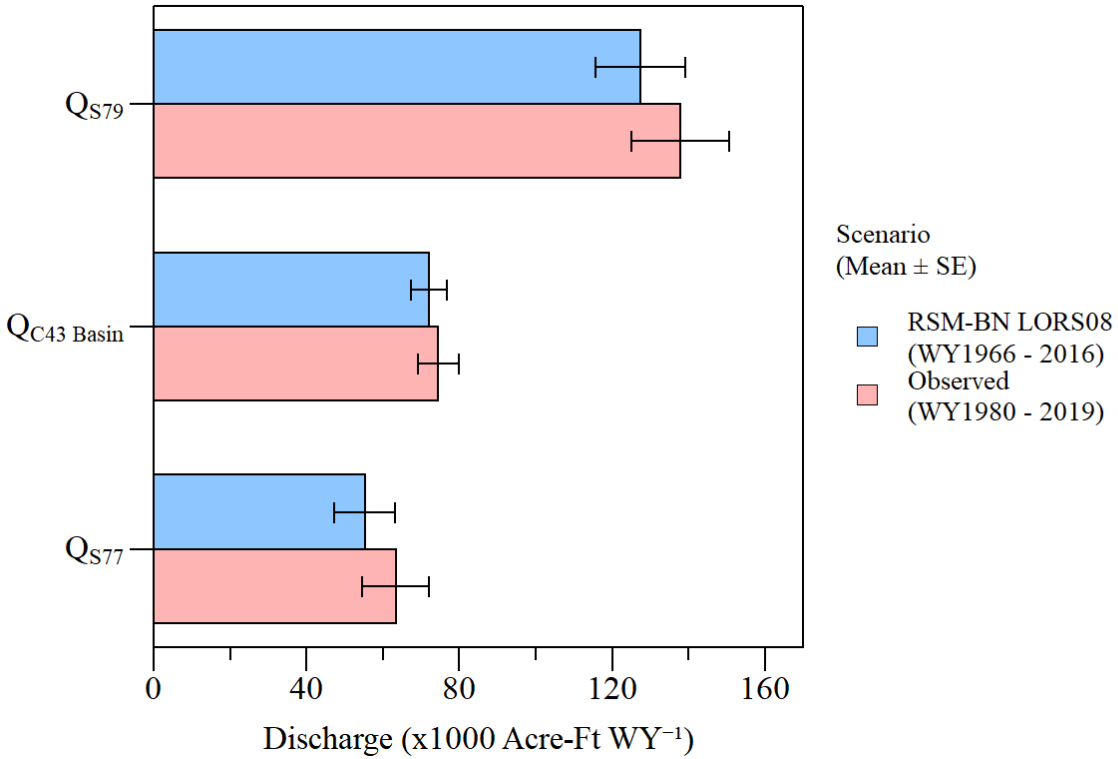
Annual observed versus predicted (± 95% CI) S-79 load during the period of record (WY1982 – WY 2019) with hurricane years identified.

S-79 Water Quality Models



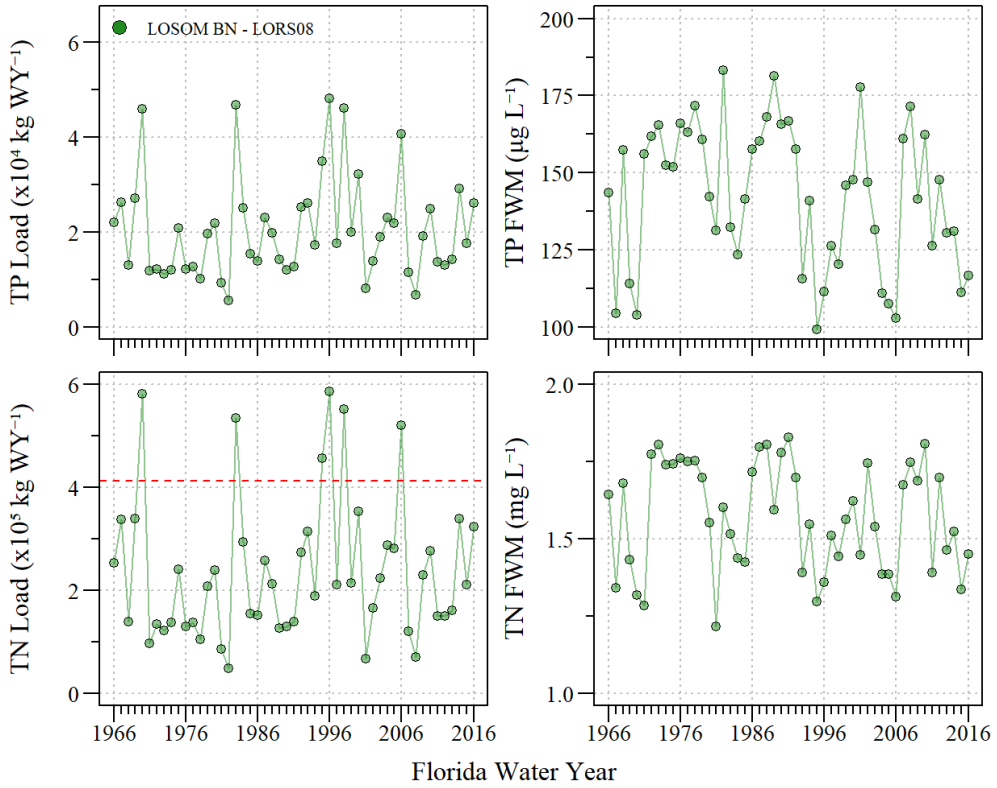
Annual observed versus predicted (± 95% CI) S-79 load during the period of record (WY1982 – WY 2019) with hurricane years identified.

Application of model with RSM-BN outputs[1]



Provisional RSM BN outputs with POR extension. For demonstration/testing purposes only.

Application of model with RSM-BN outputs[1]



Provisional RSM BN outputs with POR extension. For demonstration/testing purposes only.

Acknowledgements

Data

South Florida Water Management District (DBHYDRO)

Slides

- HTML Slide deck © Julian (2020) CC BY
- RMarkdown Source

Draft FDEP Work Product