

Phase 7 WSM Development – Water Quality Calibration (Step 1)

Modeling Workgroup Quarterly Meeting – July 2024

Gopal Bhatt¹, Isabella Bertani², Lewis Linker³, Robert Burgholzer⁴

¹ Penn State, ² UMCES, ³ US EPA, ⁴ VA DEQ – Chesapeake Bay Program Office

Presentation Outline

Phase 7 Dynamic Watershed Model (DWM)

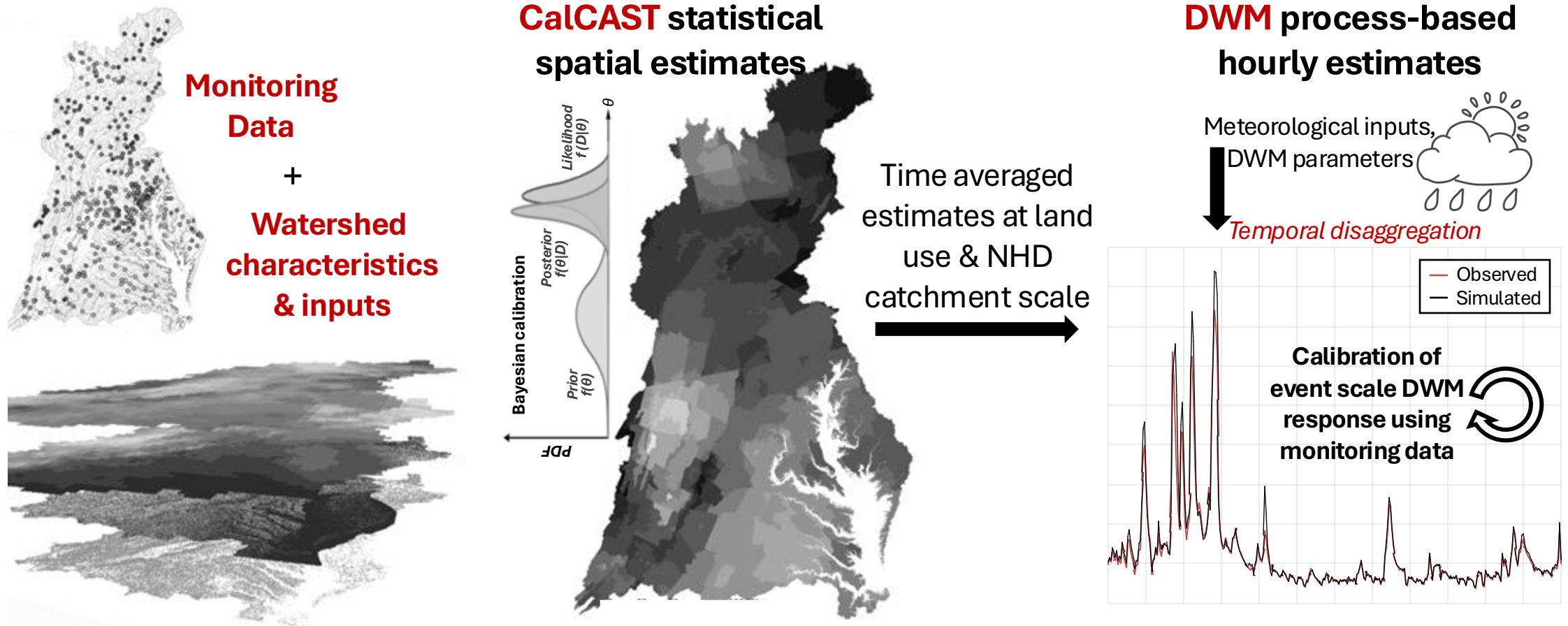
1. Dynamic Watershed Model Overview
2. Review of prior model development progress
3. Water quality calibration
 - Step 1: mechanics of water quality calibration
 - Step 2: revisions to the calibration methods
 - Small stream Q-C relationship
4. Summary and next steps

Purpose

NHD Scale Dynamic Watershed Model (DWM)

- Inputs for the estuarine models (MBM/MTMs)
- Watershed model calibration and scenario applications
- Support research and collaboration activities

Framework: Statistical Model (CalCAST) → Dynamic Watershed Model (DWM)



- Data-driven CalCAST informs DWM parameters and responses.
- NHD-scale DWM prototype is now using CalCAST *average annual* (a) total flow, (b) stormflow, (c) sediment erosion and delivery factors, and (d) total nitrogen and total phosphorus loads and delivery factors.

Dynamic Watershed Model (DWM) Development

- Year 2022: NHD-scale model structure and prototypes for hydrology, sediment, and nutrients.
 - Operational prototypes with reasonable runtime and on the graph paper model results.

Development Milestones (2022)

100K NHD	NHD-scale model structure; Hydrology prototype; Expanded simulation period 1985 to 2020; ^{[1][2]}
HYDROLOGY	Hydrology calibration (CalCAST→DWM) method updates; initial testing of numerical simplifications for flow routing; ^[3]
SEDIMENT	Sediment model; Hydrology model calibration updates with respect to stormflow; ^[4]
NUTRIENTS	Nutrient (Nitrogen and Phosphorus) model; Updated sediment model; ^[5]

[1] <https://d18ev1ok5eia.cloudfront.net/chesapeakebay/documents/Progress-in-Phase-7-WSM-Development-GopalBhatt-Penn-State-10.4.2022.pdf>
[2] <https://d18ev1ok5eia.cloudfront.net/chesapeakebay/documents/Progress-in-Phase-7-WSM-Development-GopalBhatt-Penn-State-10.4.2022.pdf>
[3] <https://d18ev1ok5eia.cloudfront.net/chesapeakebay/documents/Progress-in-Phase-7-WSM-Development-GopalBhatt-Penn-State-10.4.2022.pdf>
[4] <https://d18ev1ok5eia.cloudfront.net/chesapeakebay/documents/Progress-in-Phase-7-WSM-Development-GopalBhatt-Penn-State-10.4.2022.pdf>
[5] <https://d18ev1ok5eia.cloudfront.net/chesapeakebay/documents/Progress-in-Phase-7-WSM-Development-GopalBhatt-Penn-State-10.4.2022.pdf>

- Year 2023: Incremental refinements of model prototypes in terms of model segmentation, CalCAST→DWM linkage, and simulation of the small streams.

... incremental improvements during 2023

Segmentation	We performed re-segmentation and tested the revised model. <ul style="list-style-type: none">▪ tidal percent attribute was updated using new shoreline layer▪ all databases (river mainstem, topology, etc.) were updated▪ segmentations in the tidal watershed were improved
Model	Improvements on overcoming the 'aggregation effect' in the simulation of river mainstems (CalCAST → DWM) <ul style="list-style-type: none">➢ Non-iterative hydraulic routing for small 100K NHDplus streams➢ WQ routing (TN) for small 100K NHDplus streams
Simulation	Testing on Amazon AWS and MS Azure cloud HPC environments with various node type and size configurations

[1] <https://d18ev1ok5eia.cloudfront.net/chesapeakebay/documents/20230401-BHATT-Phase-7-WSM-Development-Dynamic-Model-Development-2023Q1.pdf>
[2] <https://d18ev1ok5eia.cloudfront.net/chesapeakebay/documents/Progress-in-Phase-7-WSM-Development-GopalBhatt-Penn-State-6.20.2023.pdf>
[3] <https://d18ev1ok5eia.cloudfront.net/chesapeakebay/documents/Progress-in-Phase-7-WSM-Development-GopalBhatt-Penn-State-10.17.2023.pdf>
[4] <https://d18ev1ok5eia.cloudfront.net/chesapeakebay/documents/20230109-BHATT-Phase-7-WSM-Development-Dynamic-Model-Development-2023Q4.pdf>

- Year 2024 Q1: small stream routing based on β parameters

[1] <https://d18ev1ok5eia.cloudfront.net/chesapeakebay/documents/Progress-in-Phase-7-WSM-Development-GopalBhatt-Penn-State-CBPO-4.2.2024.pdf>

Water Quality Calibration

Step 1: mechanics of water quality calibration

- We implemented a draft of the calibration framework through development of scripts according to P7 modules and data flow.
- We tested the calibration framework by running it for two model setups:
 1. Output = Input x stream transport factor
[Output = Input x STF]
 2. Small stream routing for flow and water quality
[Beta Parameters]

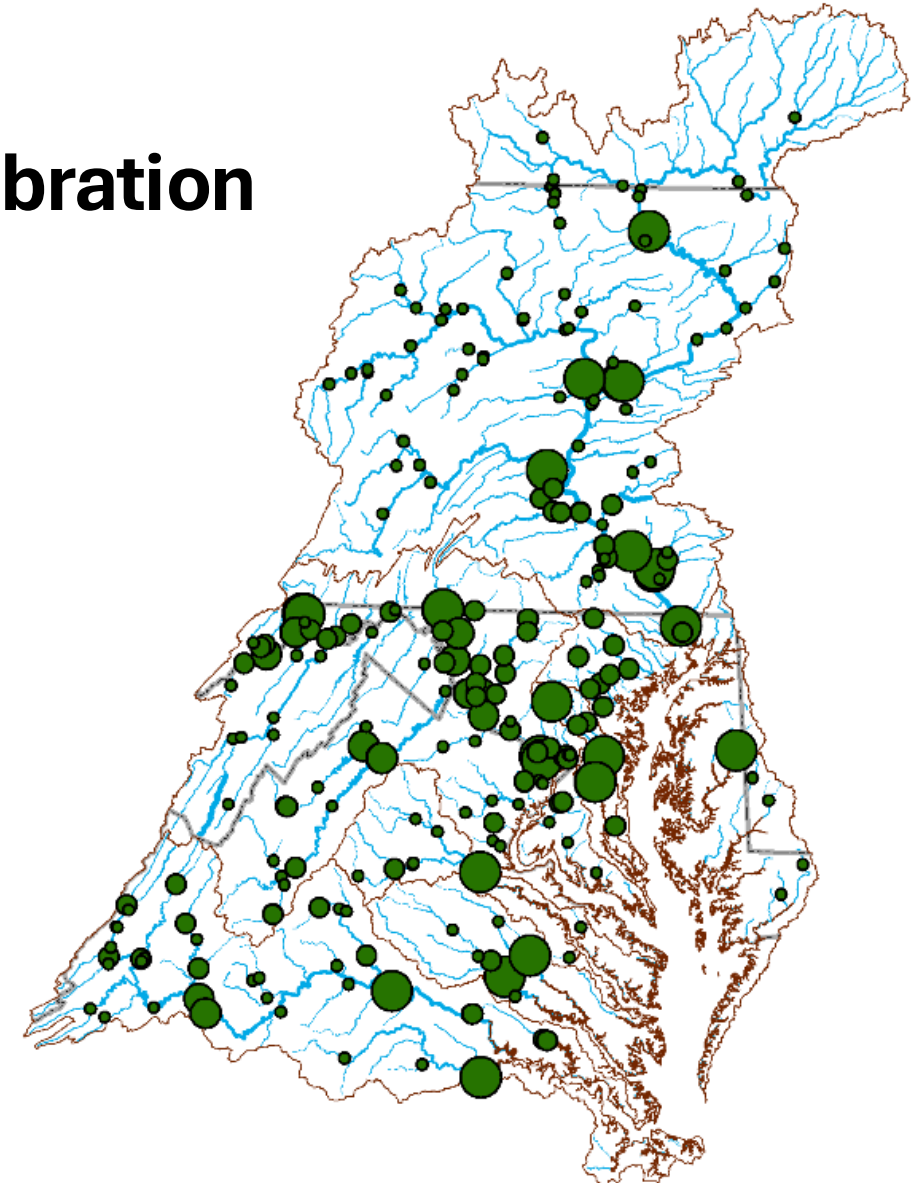
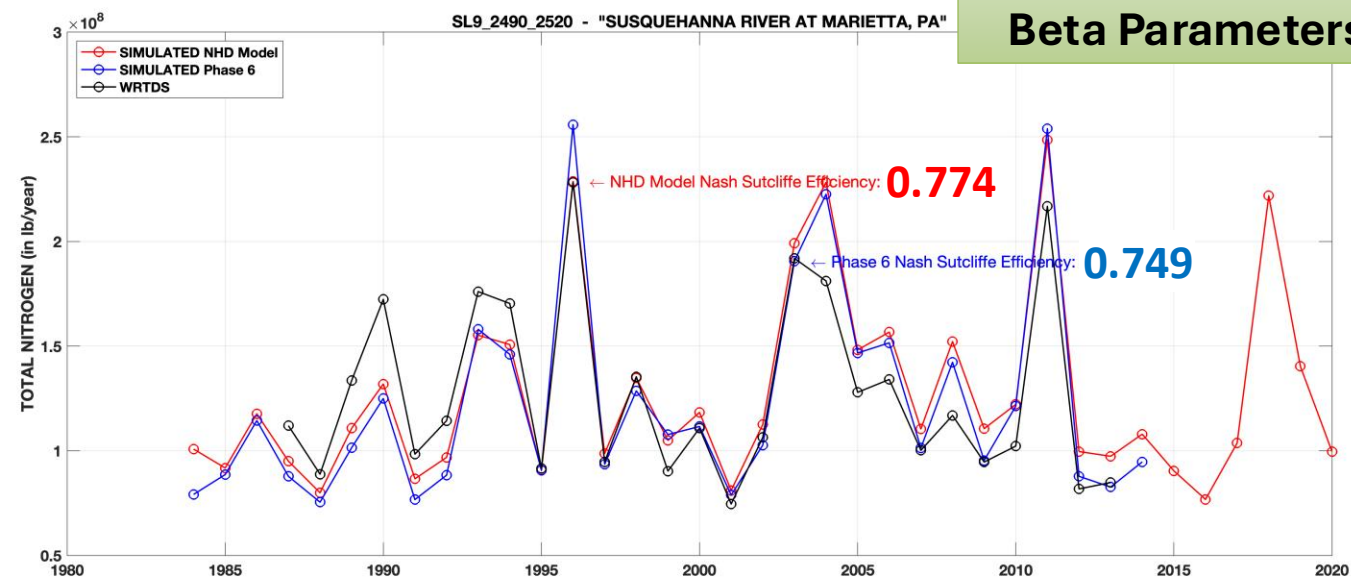
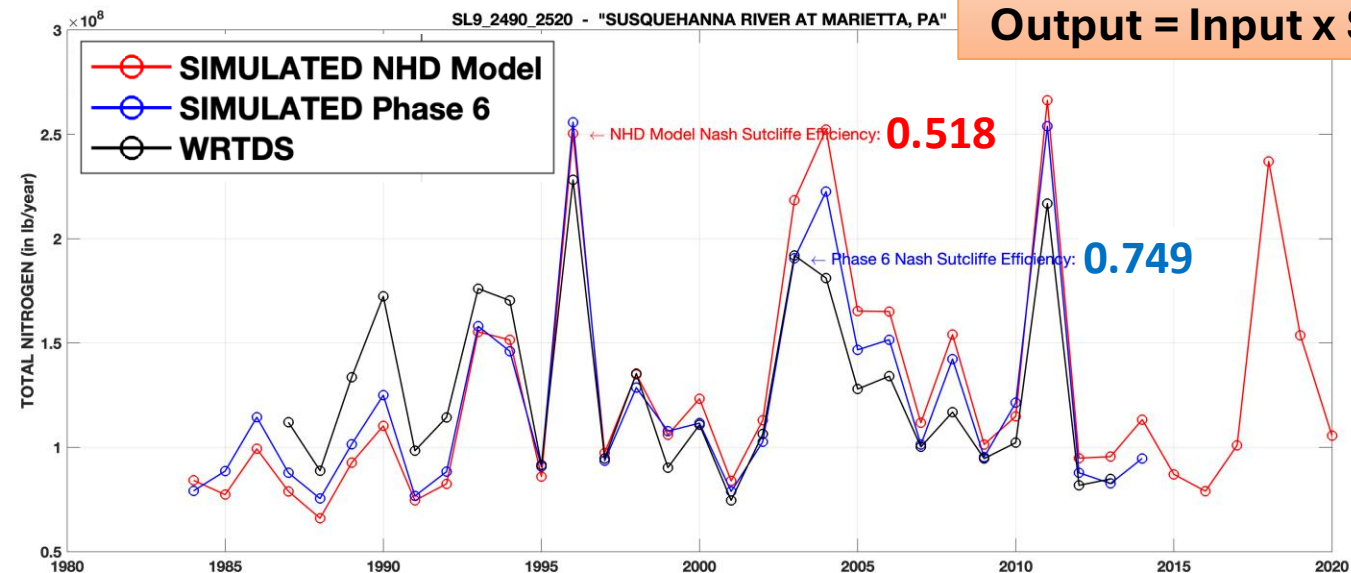
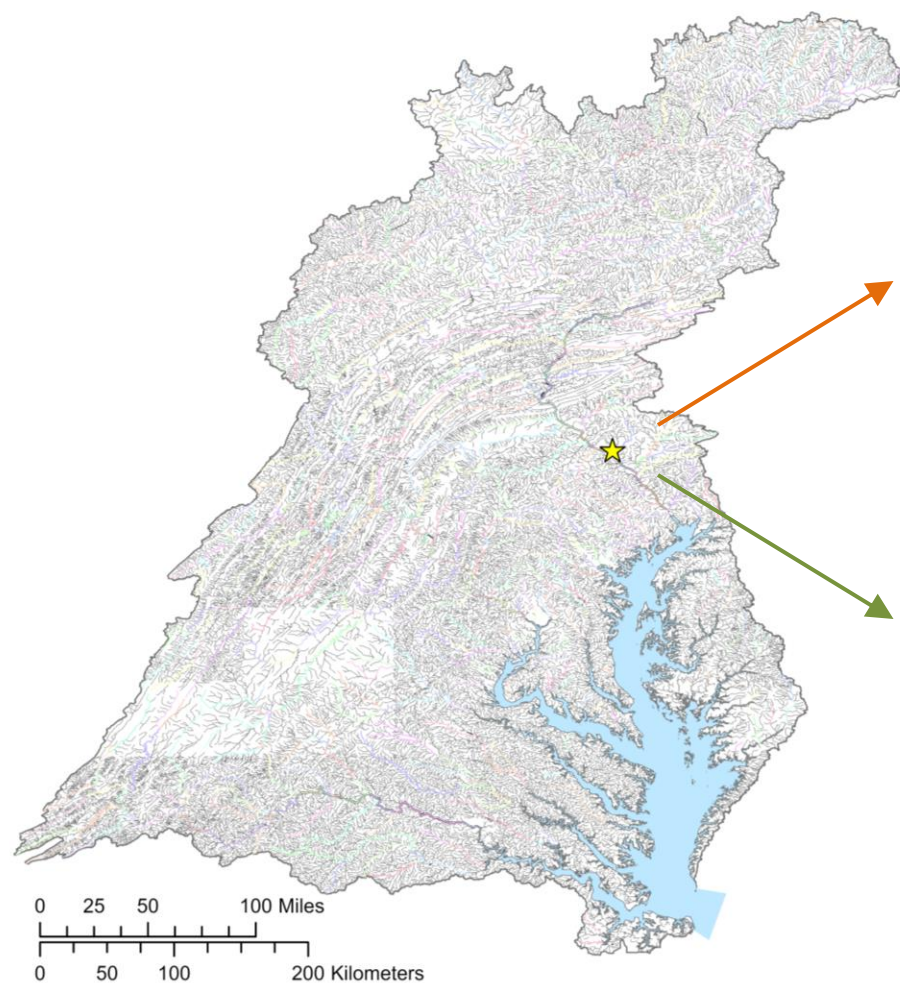


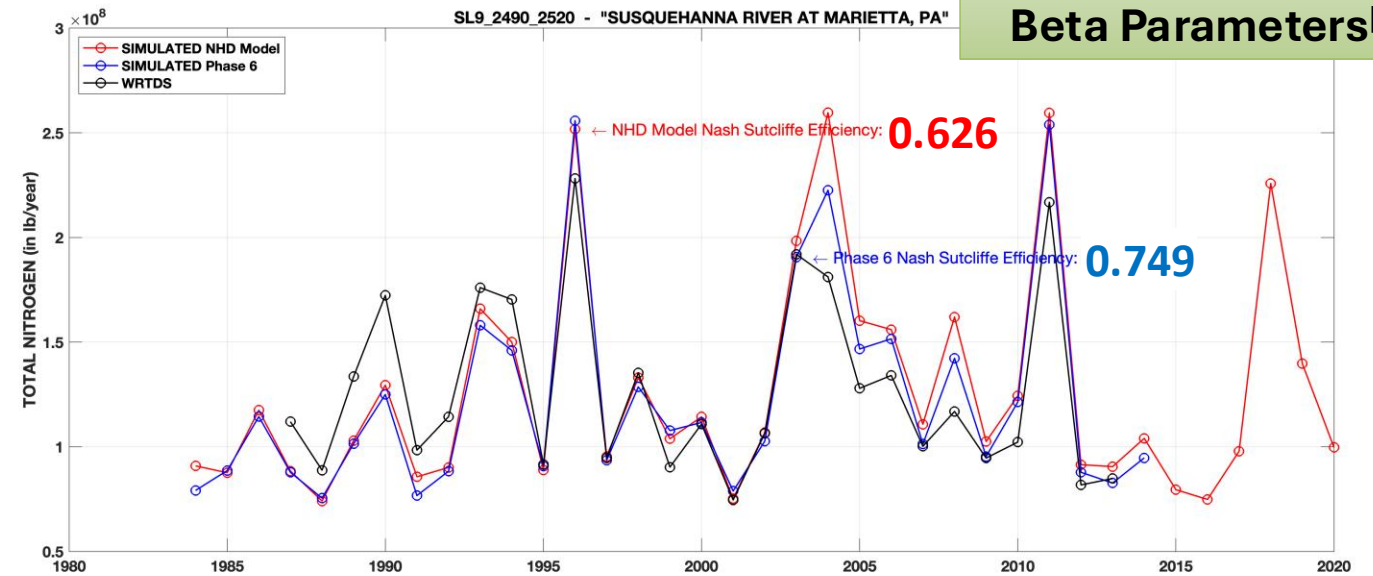
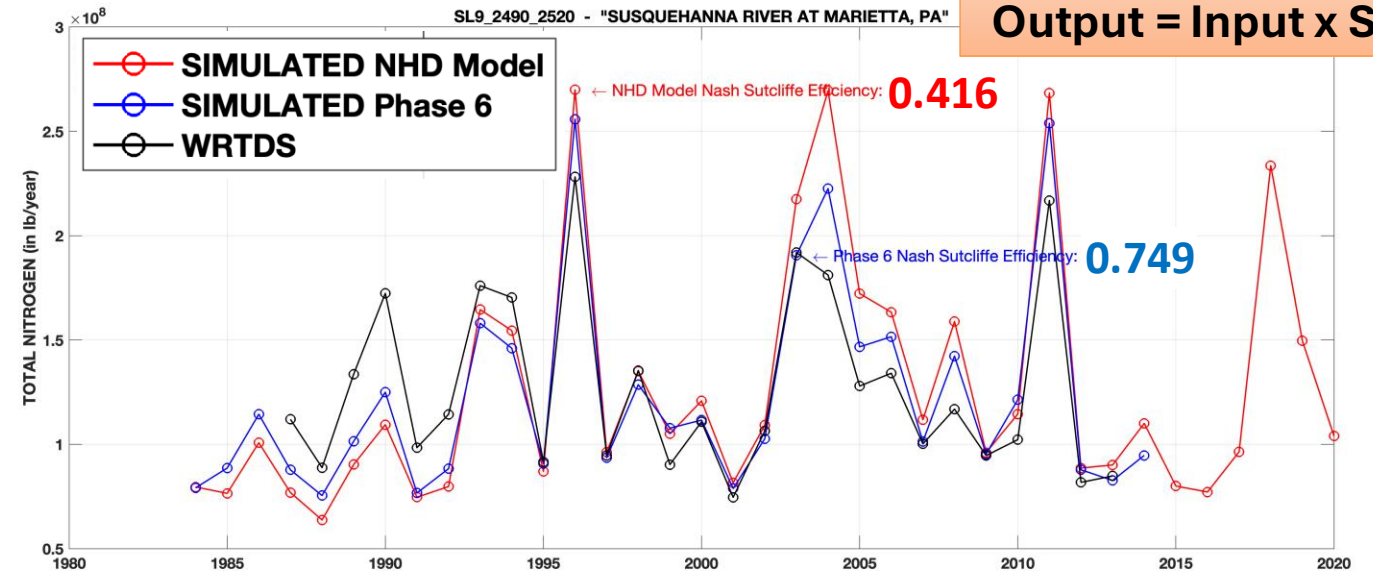
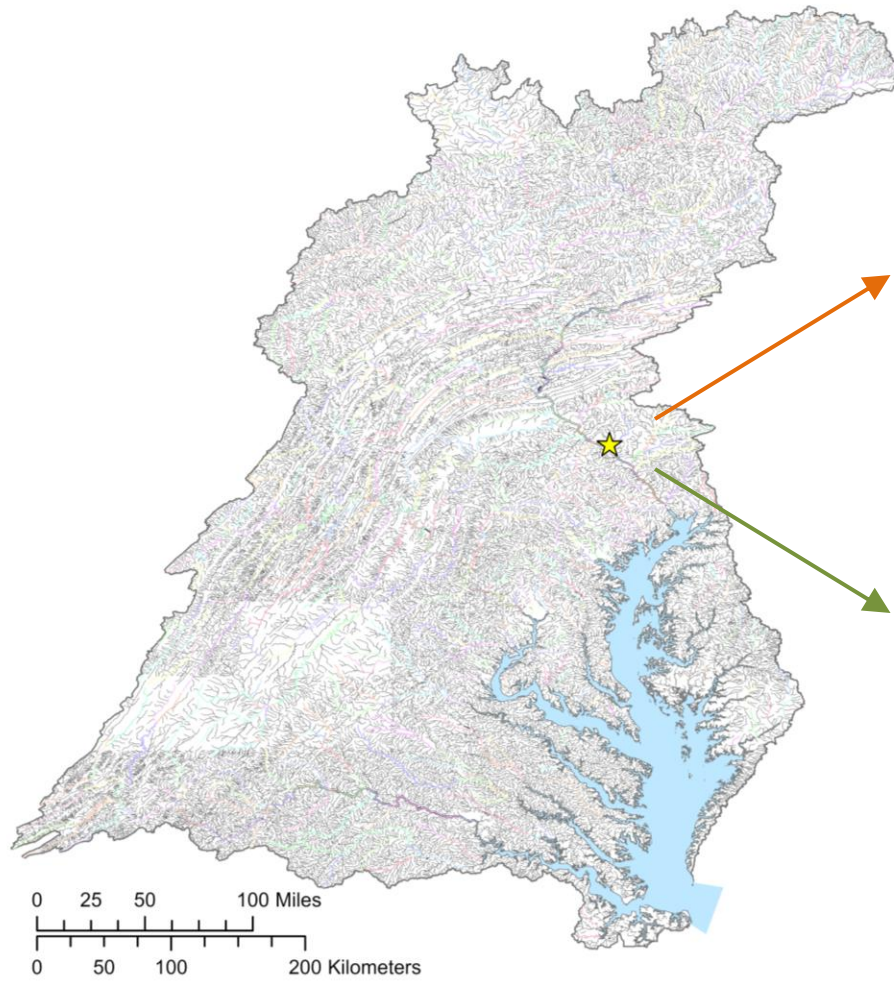
Figure: Monitoring stations used in the model calibration. Number of samples of total nitrogen are shown.

Nitrogen – P6 Parameters



[1] Watershed average beta parameters for all streams

Nitrogen – P7 Calibration^[2]

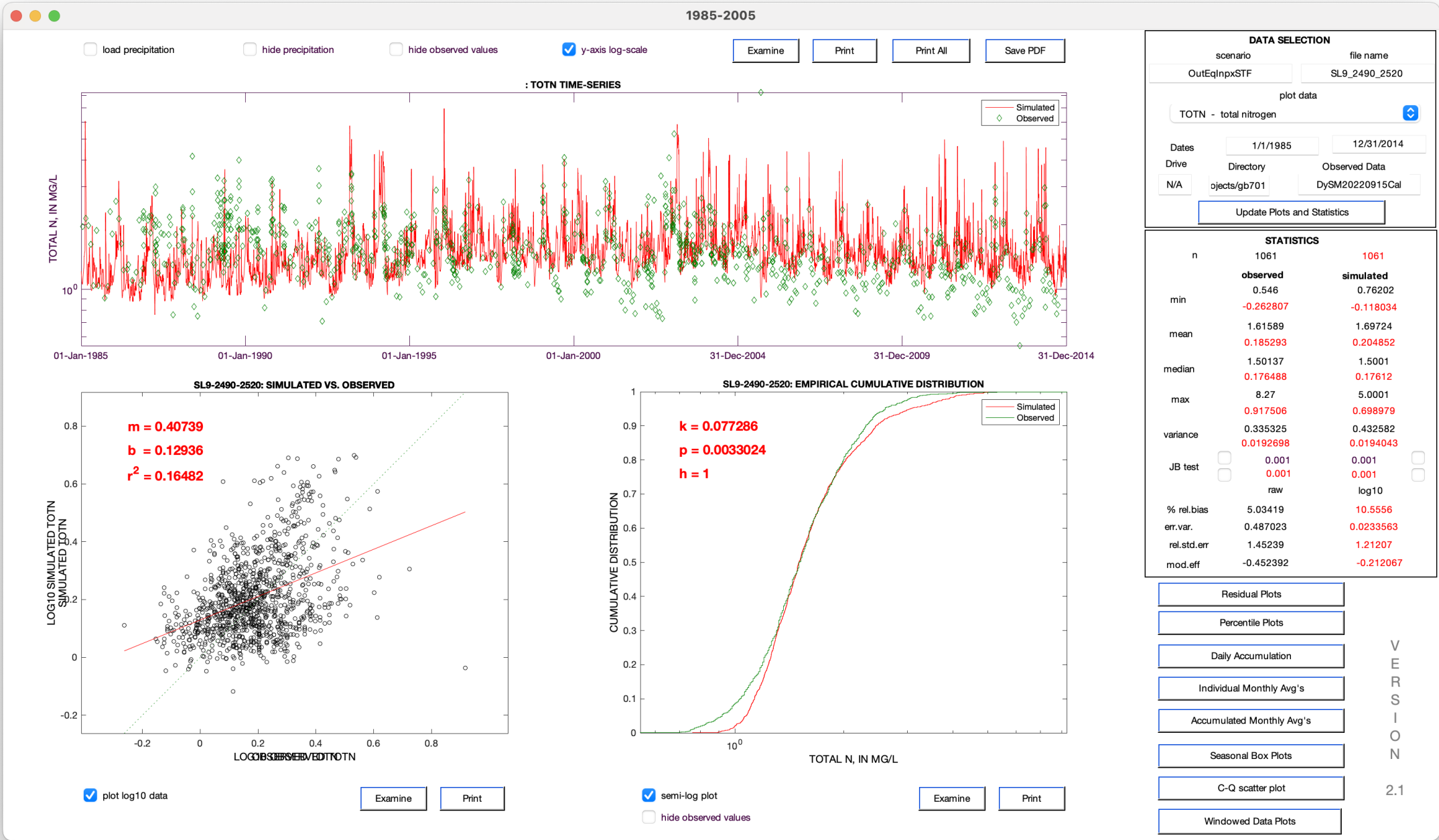


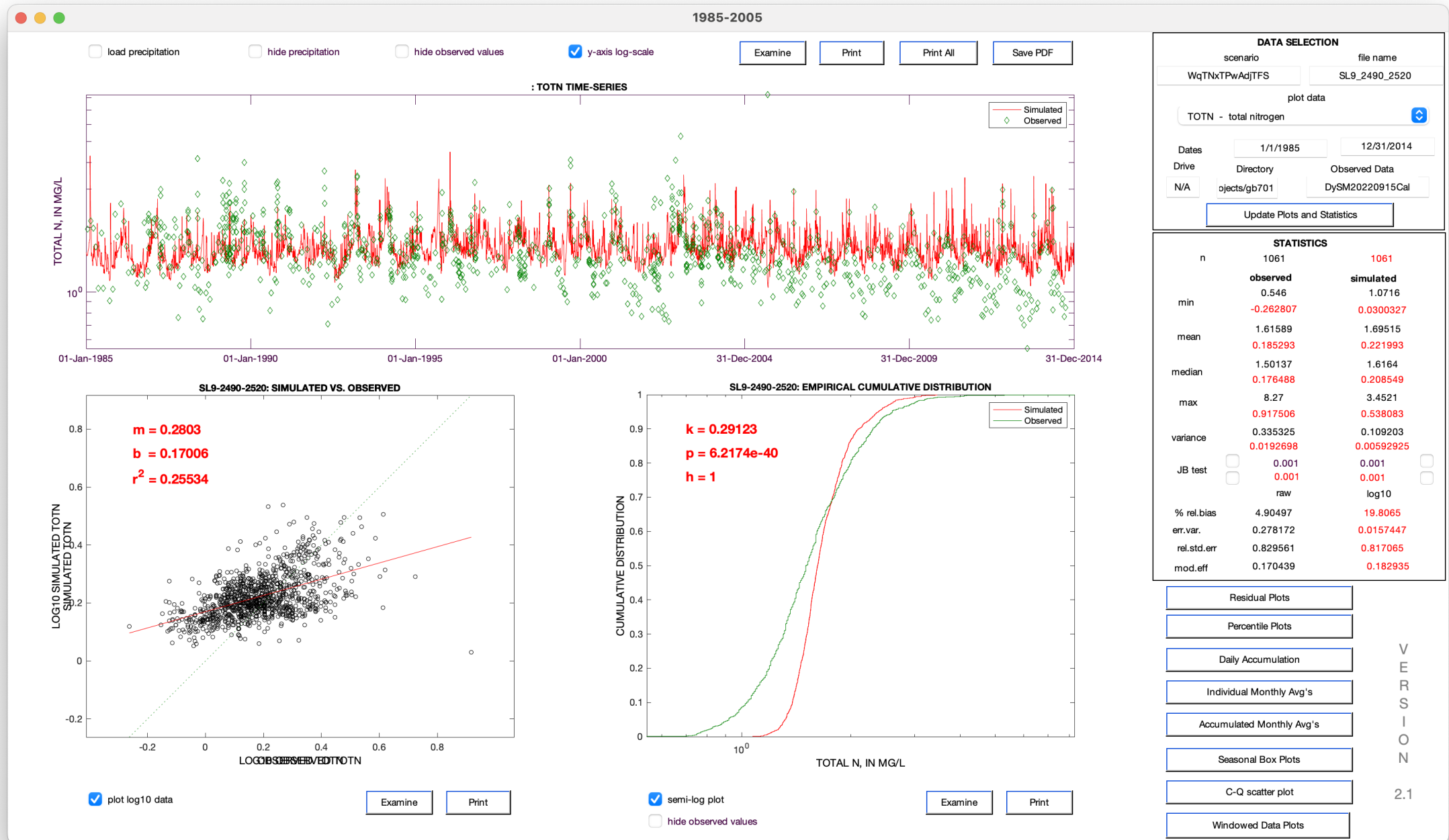
- [1] Watershed average beta parameters for all streams
[2] initial drafting and testing of calibration mechanics only

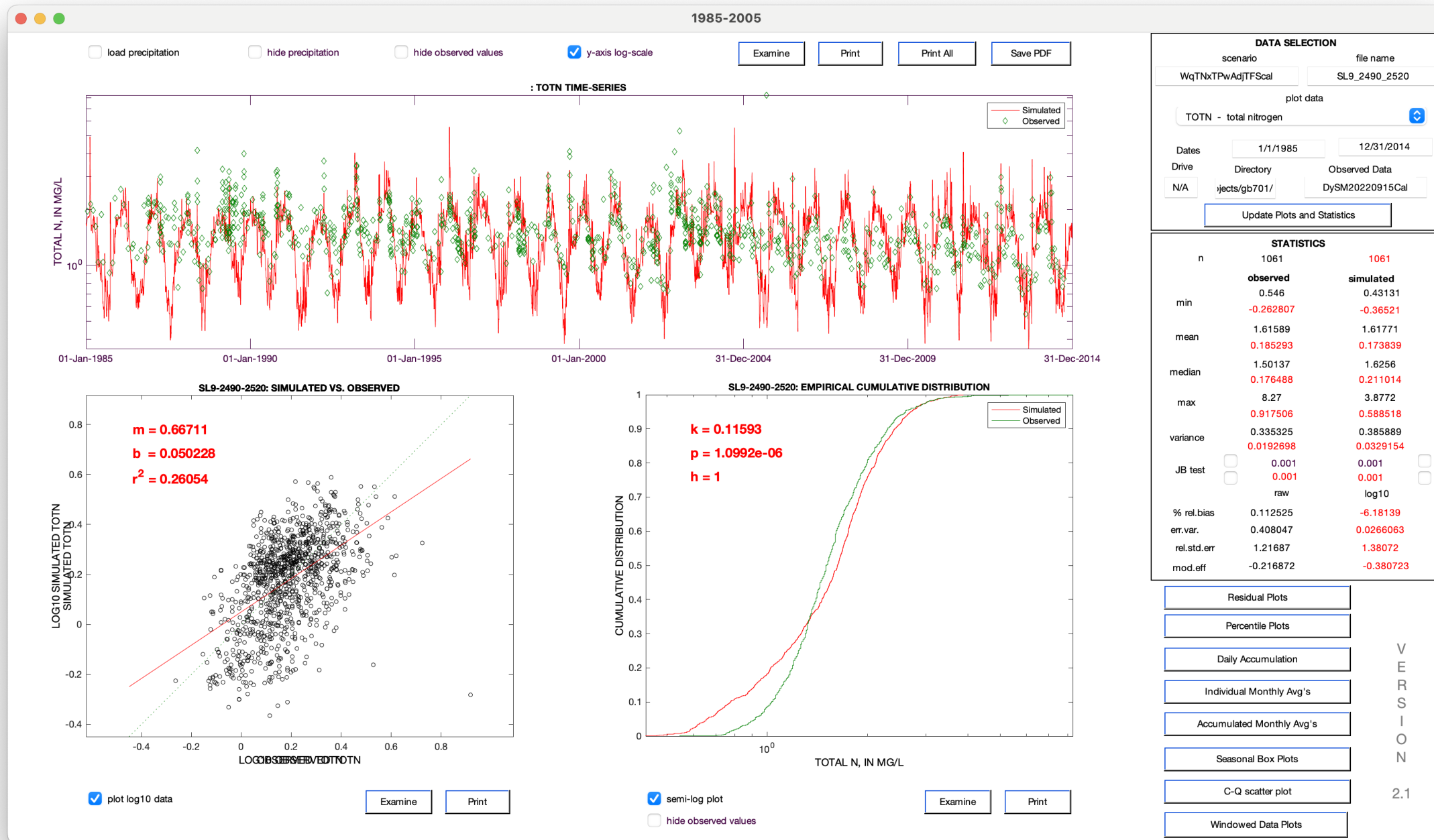
Output = Input x STF

Susquehanna at Marietta

Nitrogen







Step 2: revisions to the calibration methods (2024 Q3)

Information available for the calibration dynamic model parameters:

- Observations of daily concentrations for nutrient species at monitoring stations
- Transport factors for TN, TP, and SS for each river mainstems (CalCAST)
- QC relationships (generalized β parameters for river mainstems)

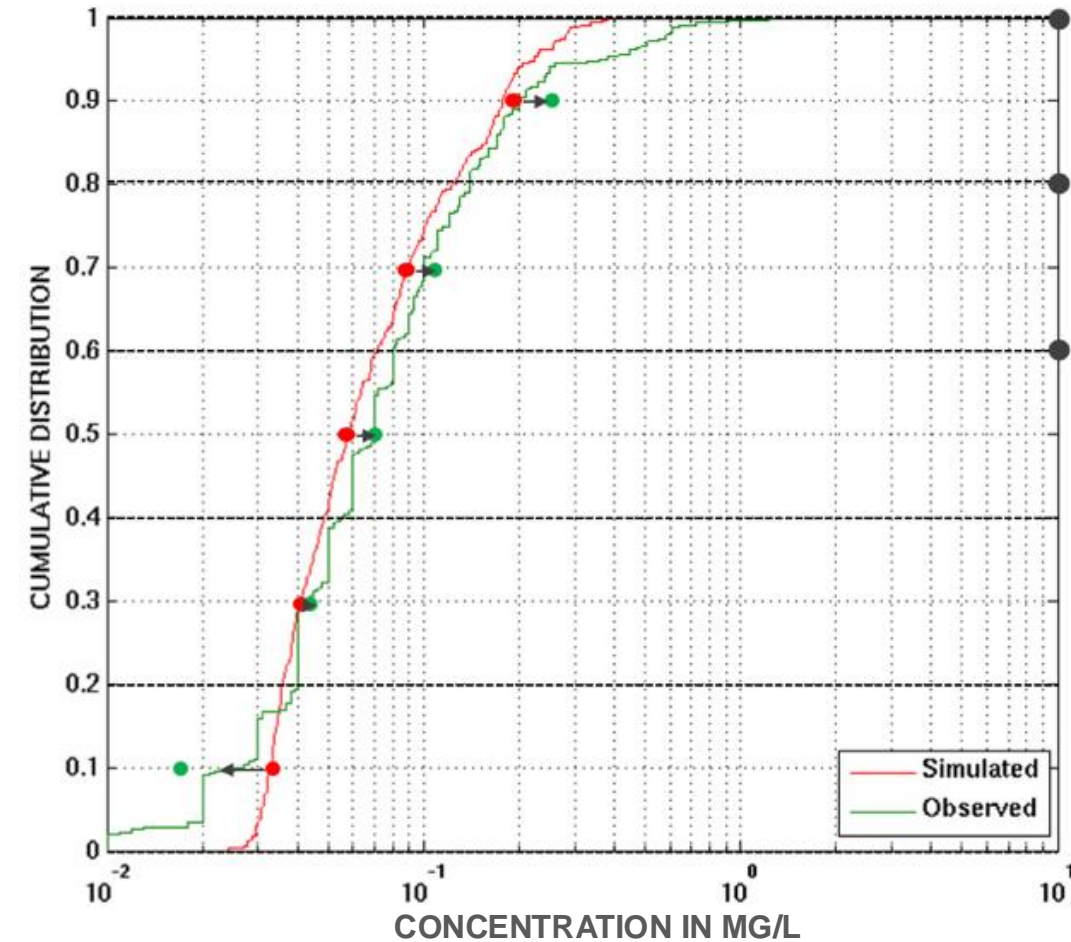


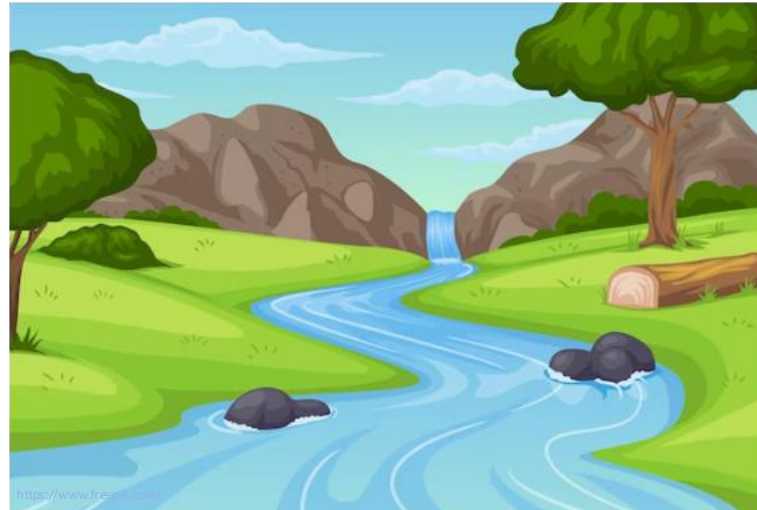
Figure: Cumulative distribution of concentrations at a monitoring station.

Small-stream flow and concentration (Q-C) relationship



HSPF: hourly surface and groundwater hydrology of land uses

UNEC: annual surface and groundwater concentrations as a function of input history and estimates of lag-times

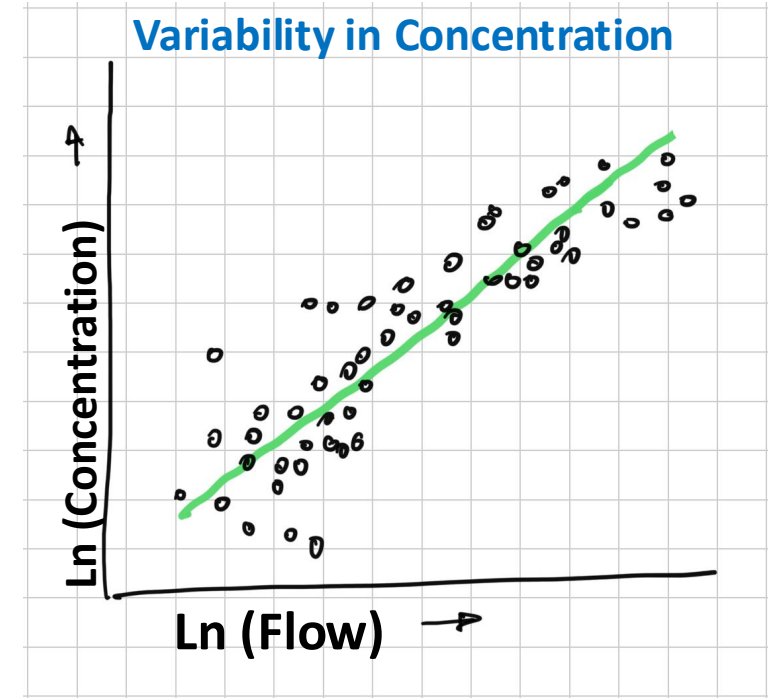


Biogeochemical processing,
Storage/deposition, Scour
→ Fate and Transport

Stream Transport Factor (STF)



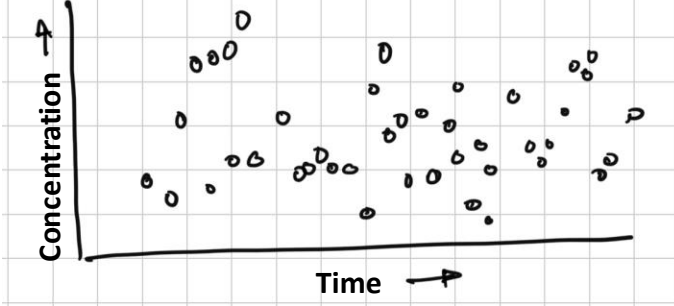
Variability in Concentration



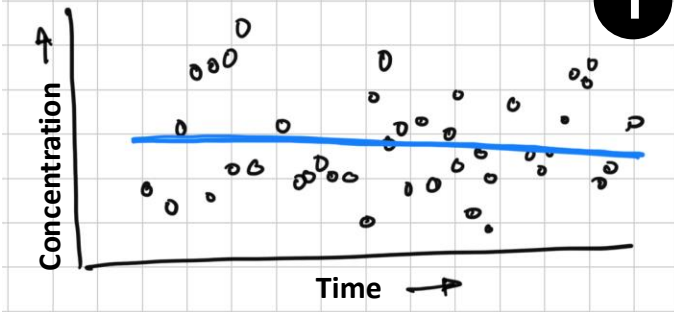
? ... emergent behavior

Small-stream flow and concentration (Q-C) relationship

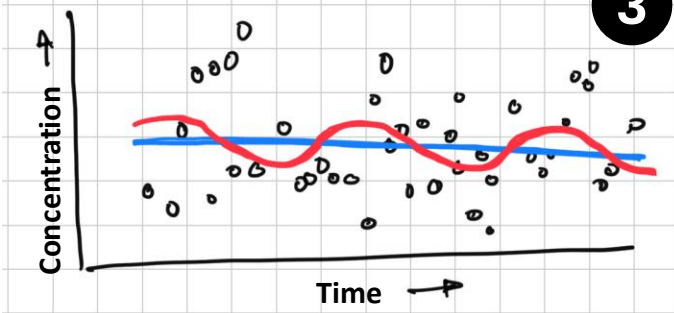
> data at a monitoring station



$$\ln(c_{trend}) = \beta_x + \beta_1 t$$



$$\ln(c_{season}) = \beta_3 \sin(2\pi t) + \beta_4 \cos(2\pi t)$$



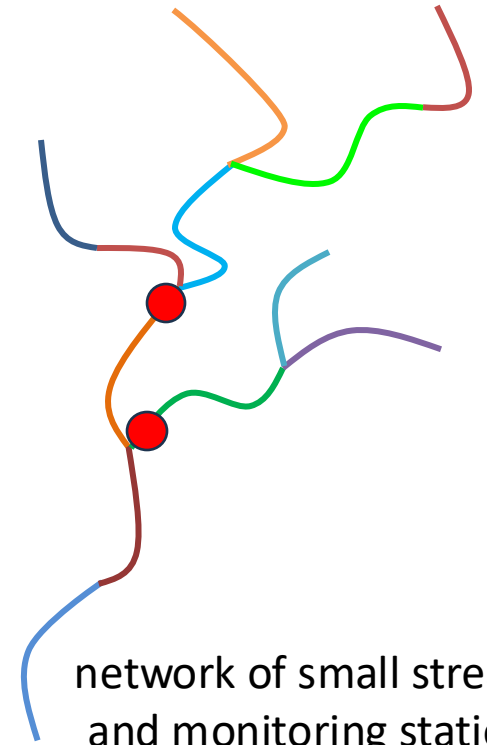
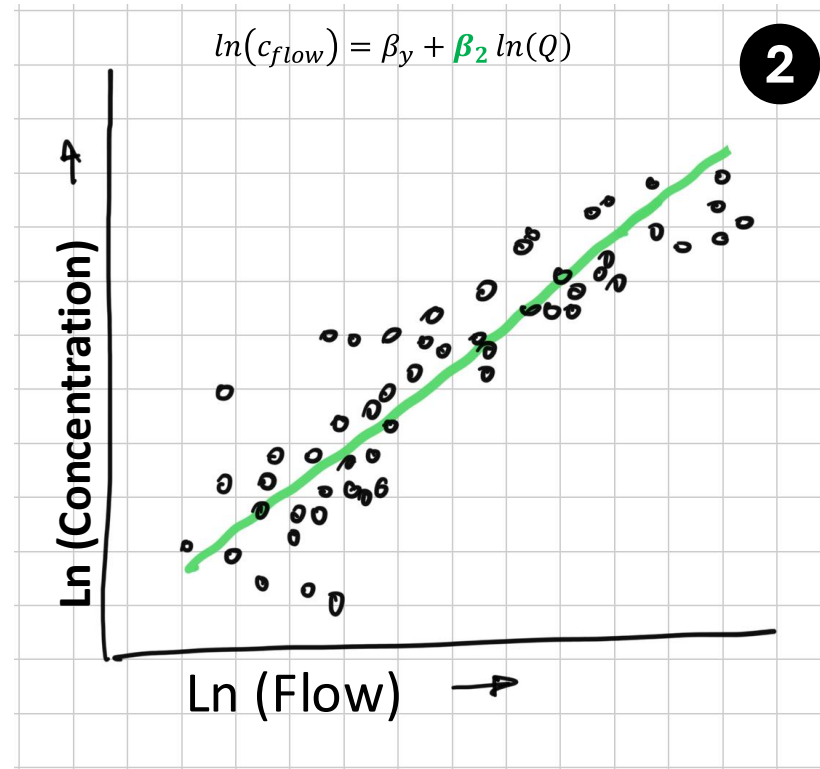
$$\ln(c) = \underbrace{\beta_o}_{1} + \underbrace{\beta_1 t}_{2} + \underbrace{\beta_2 \ln(Q) + \beta_3 \sin(2\pi t) + \beta_4 \cos(2\pi t)}_{3} + \varepsilon$$

1

2

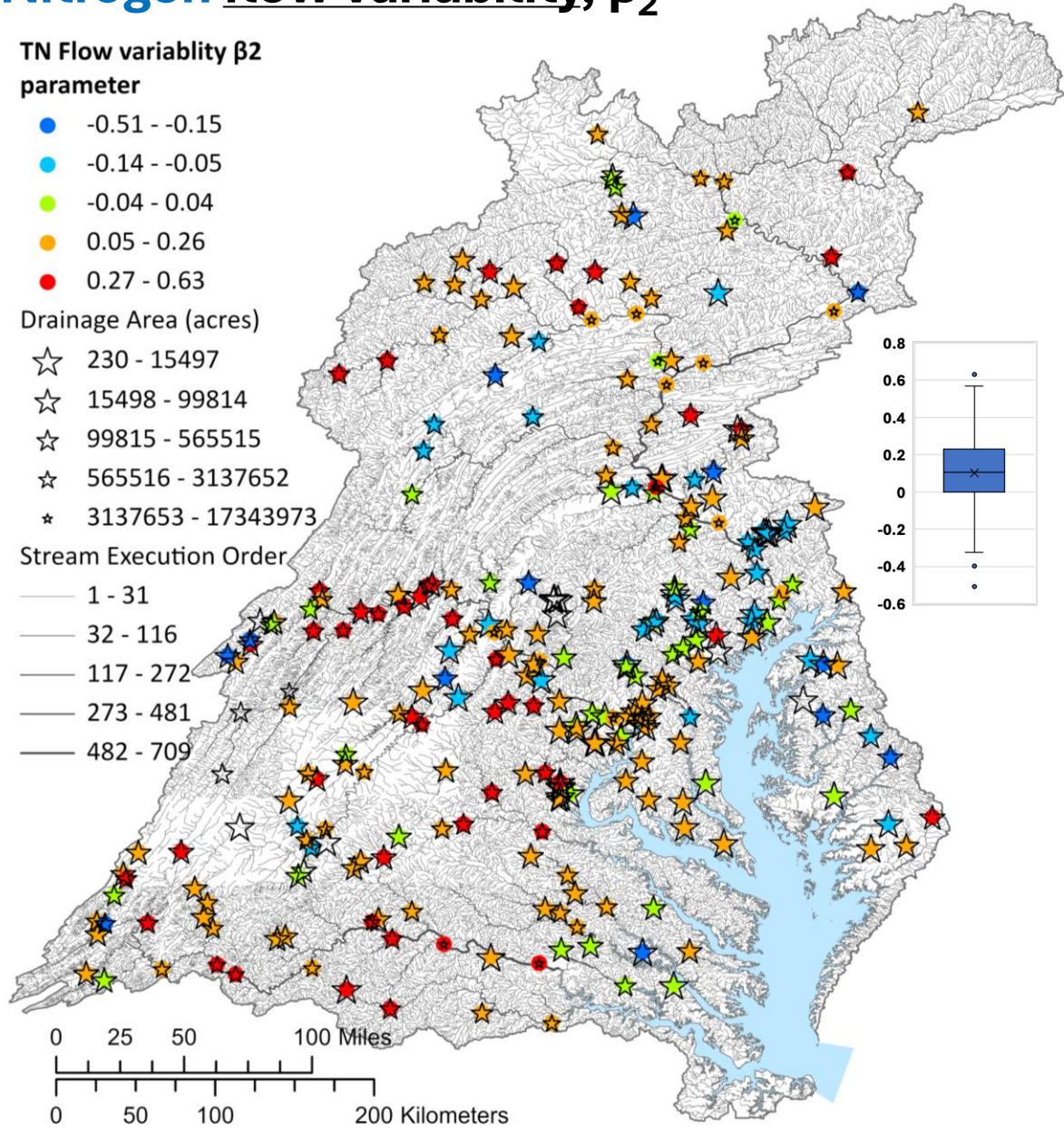
3

$$\ln(c_{flow}) = \beta_y + \beta_2 \ln(Q)$$

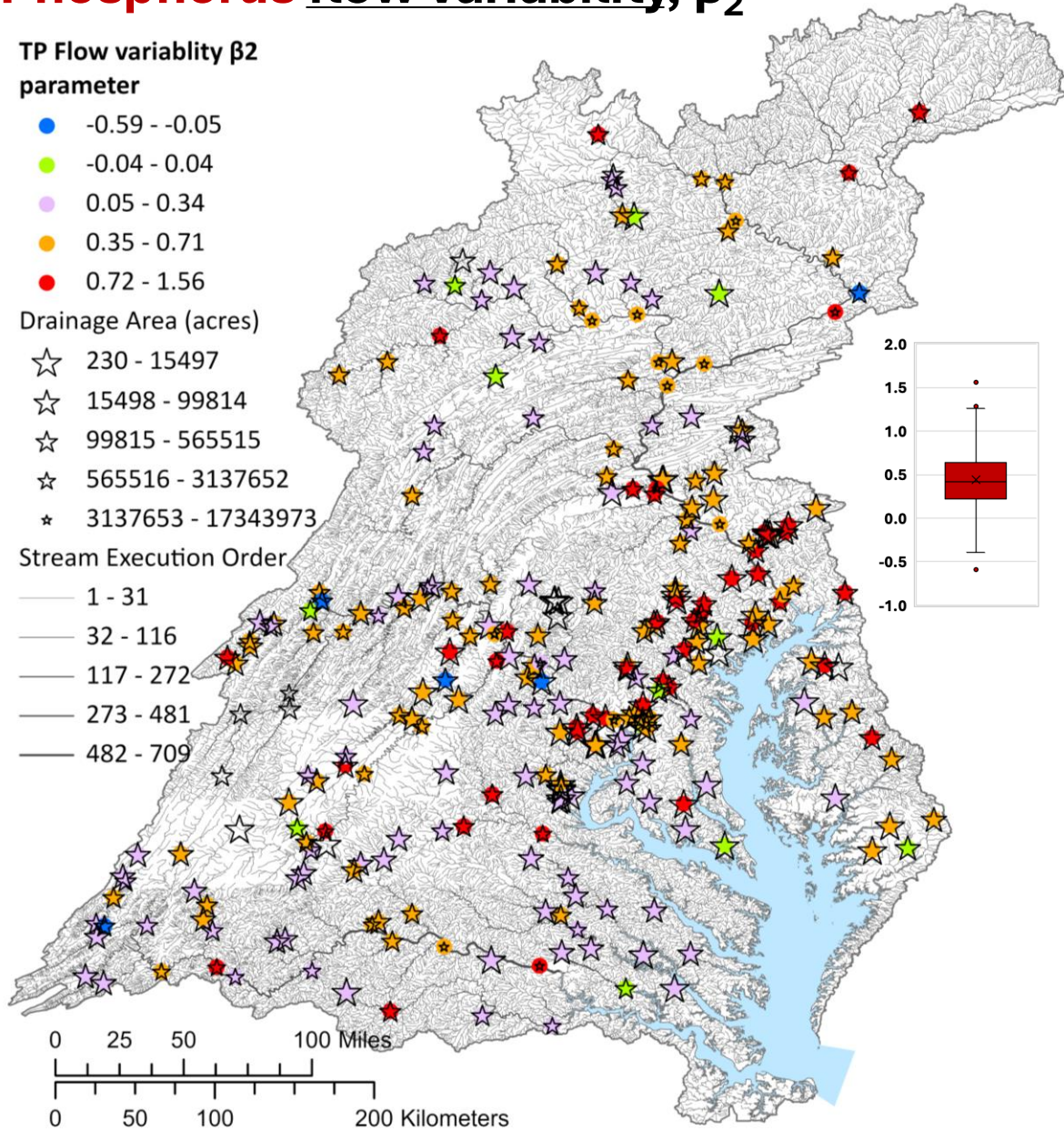


network of small streams
and monitoring stations
with water quality and
streamflow data

Nitrogen flow variability, β_2



Phosphorus flow variability, β_2

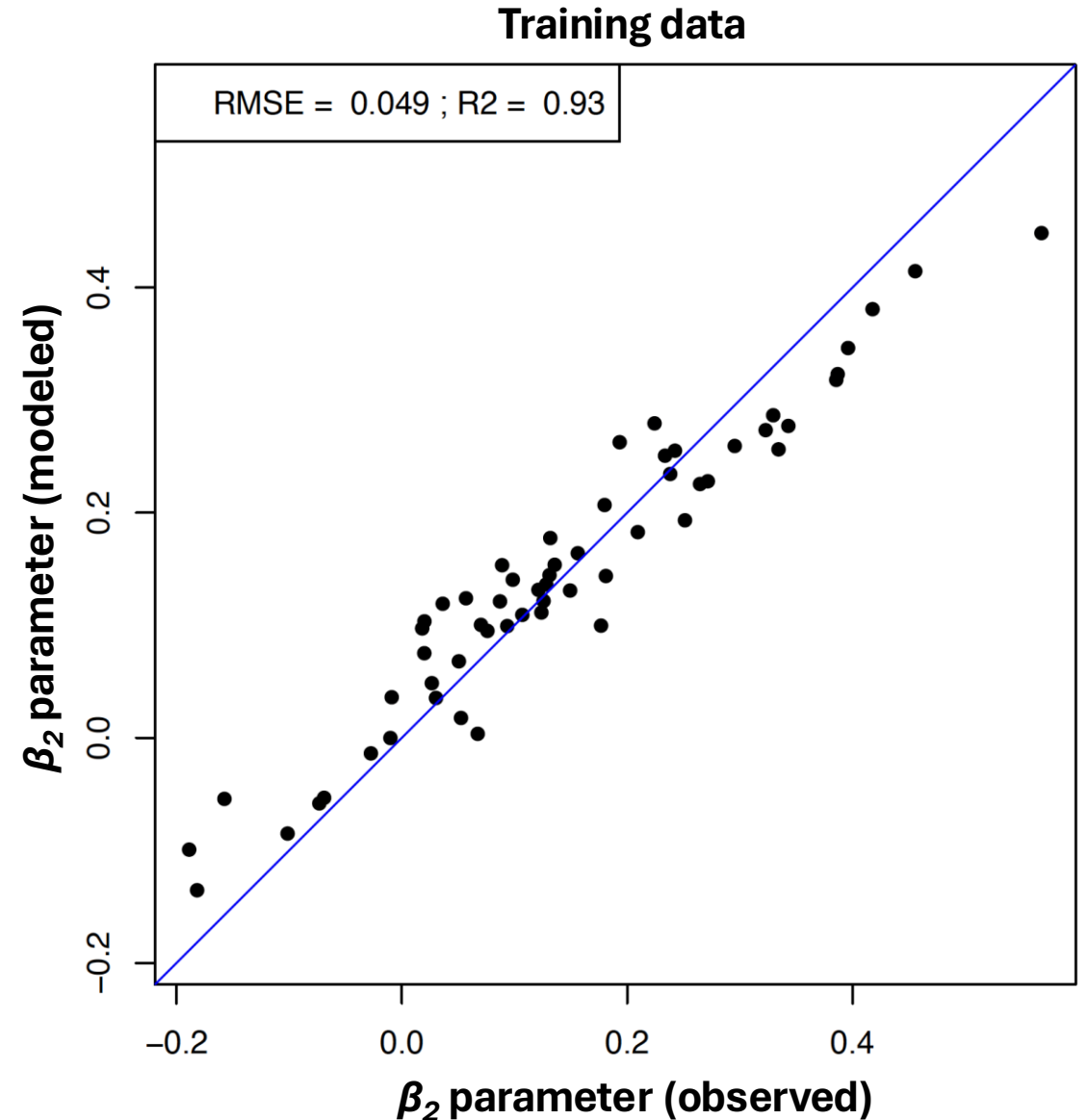


$$\ln(c) = \beta_o + \beta_1 t + \beta_2 \ln(Q) + \beta_3 \sin(2\pi t) + \beta_4 \cos(2\pi t) + \varepsilon$$

Generalization of β parameters – Qian Zhang

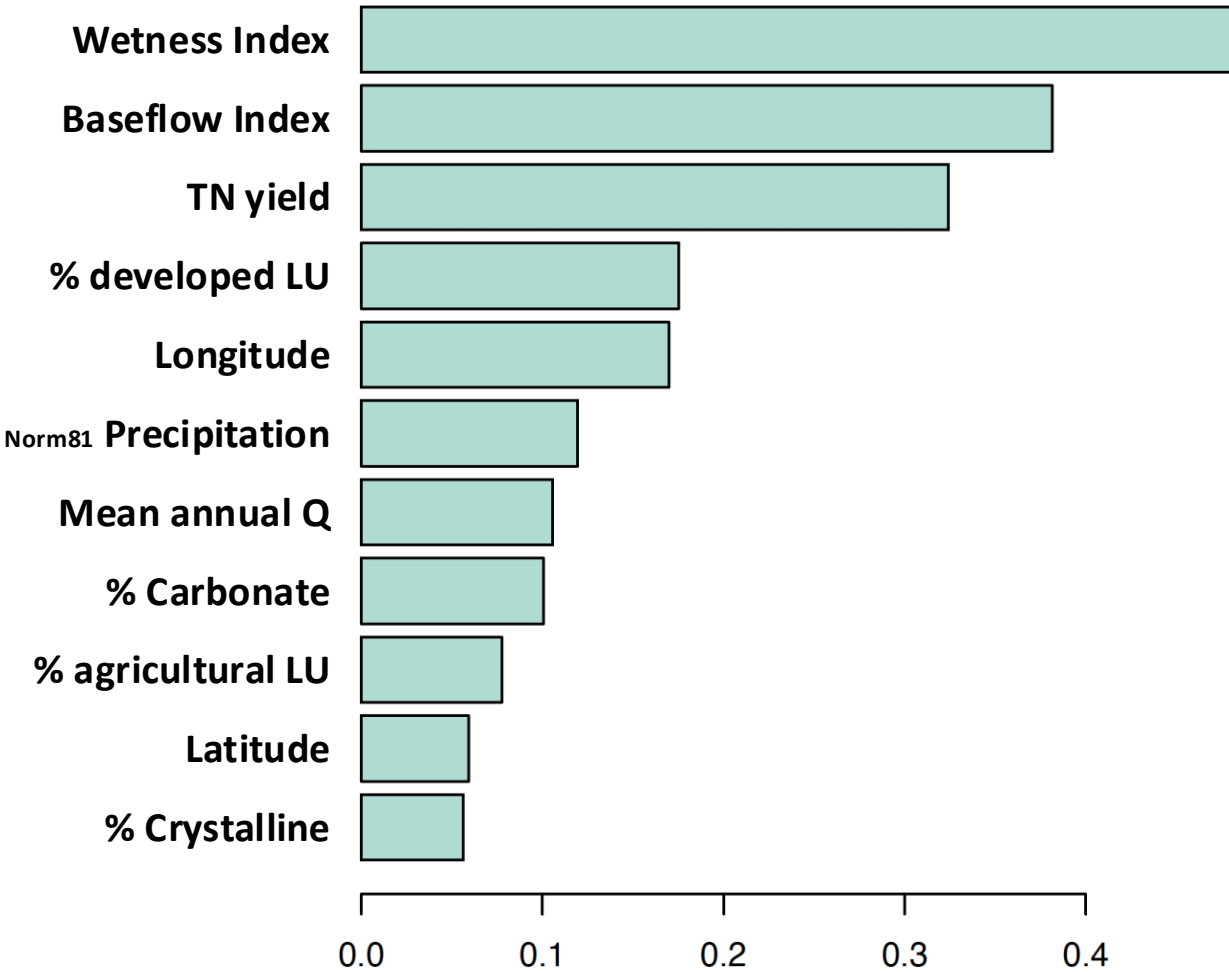
An initial *Random Forest* model was developed:

- The model links β_2 parameter with watershed attributes
- Performance of the model for training data is shown in the figure
- Currently using data for NTN stations; an expansion would require watershed attributes for additional stations

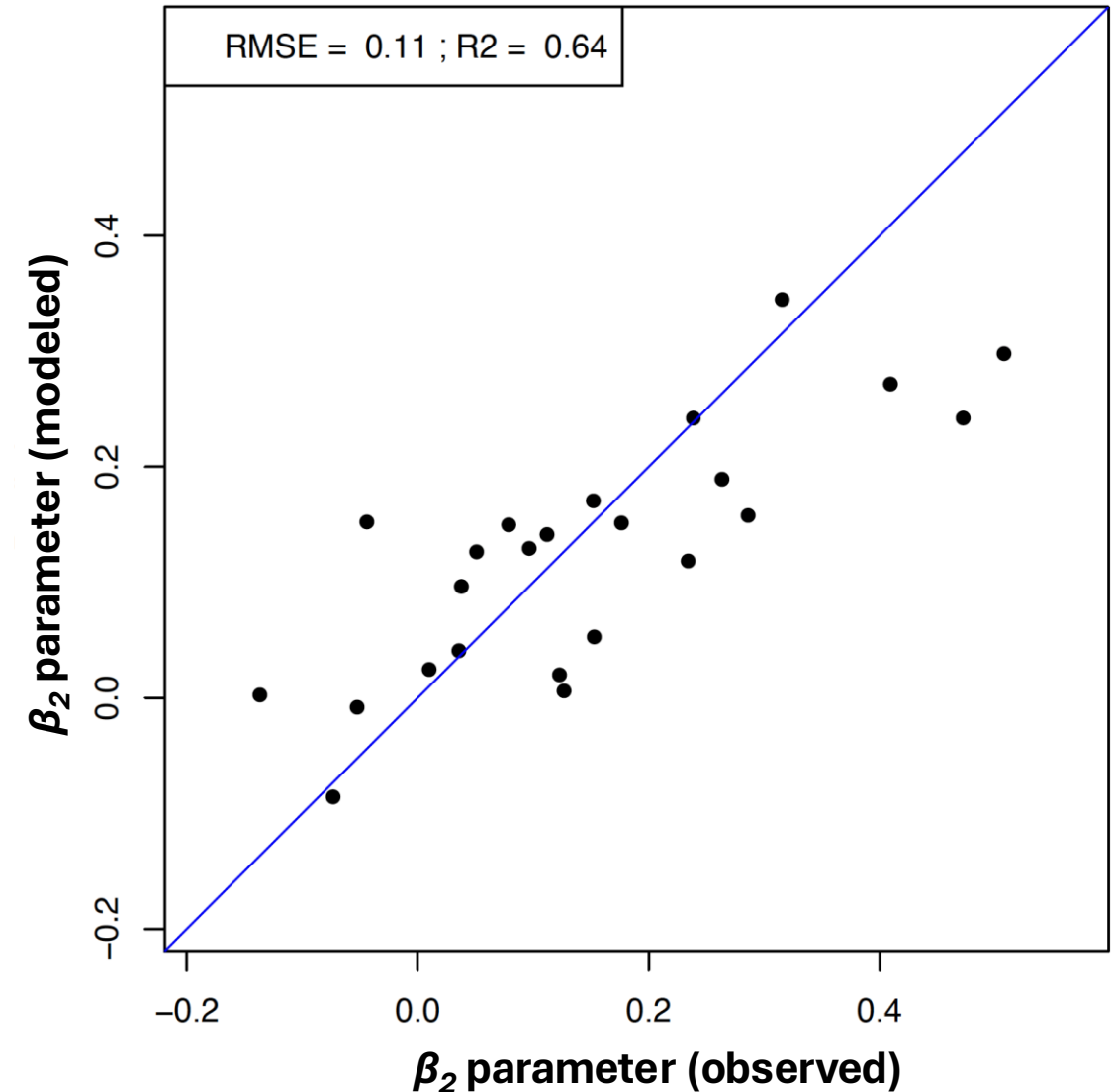


Generalization of β parameters – Qian Zhang

Importance of watershed attributes



Validation data



Summary

1. We are focusing our efforts on water quality calibration for linking watershed model flows and loads with the estuarine model

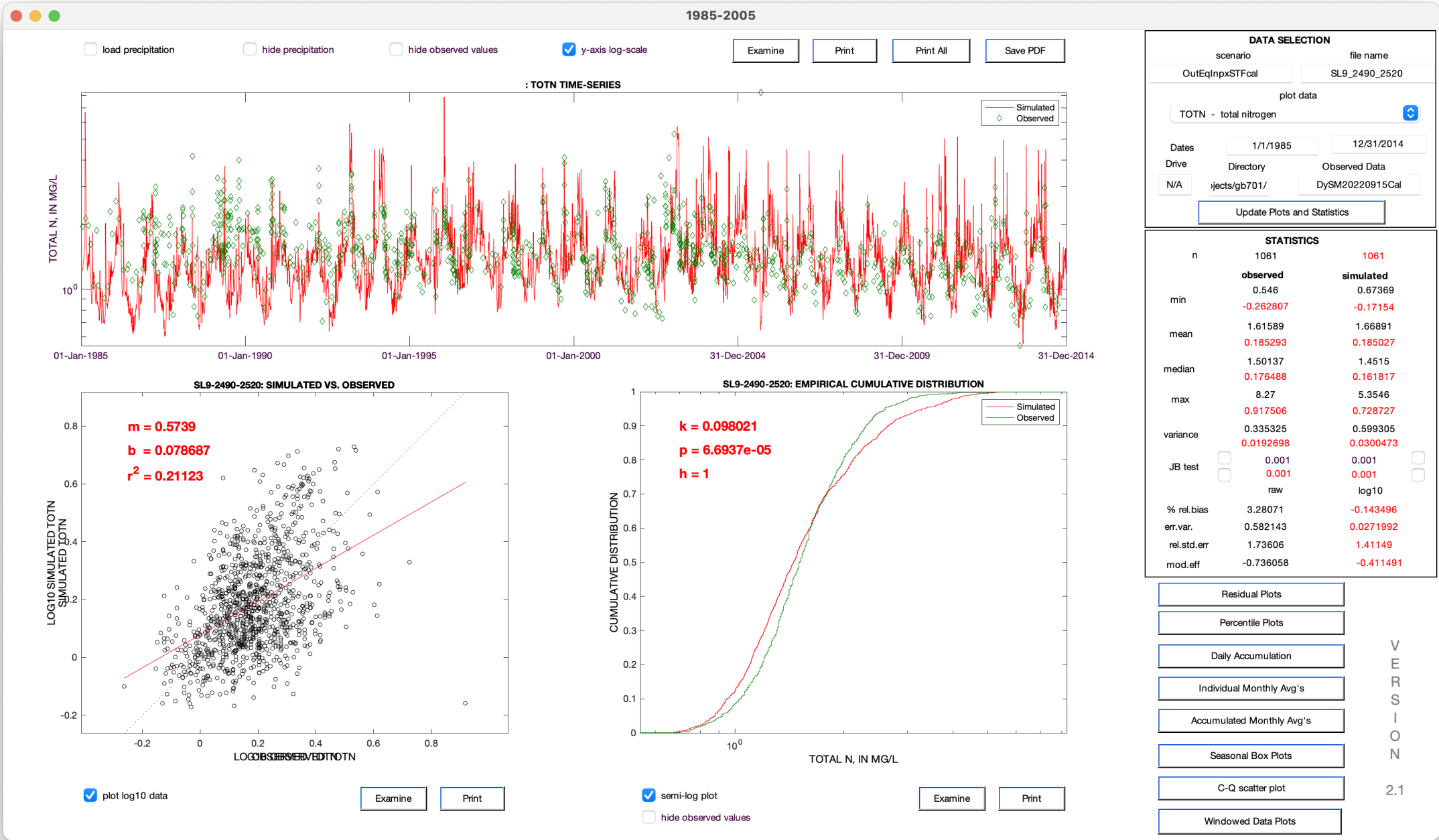
- we successfully implemented and tested Step 1 covering mechanics of water quality calibrations
- over the next quarter we will work on revisions to the calibration methods (Step 2)

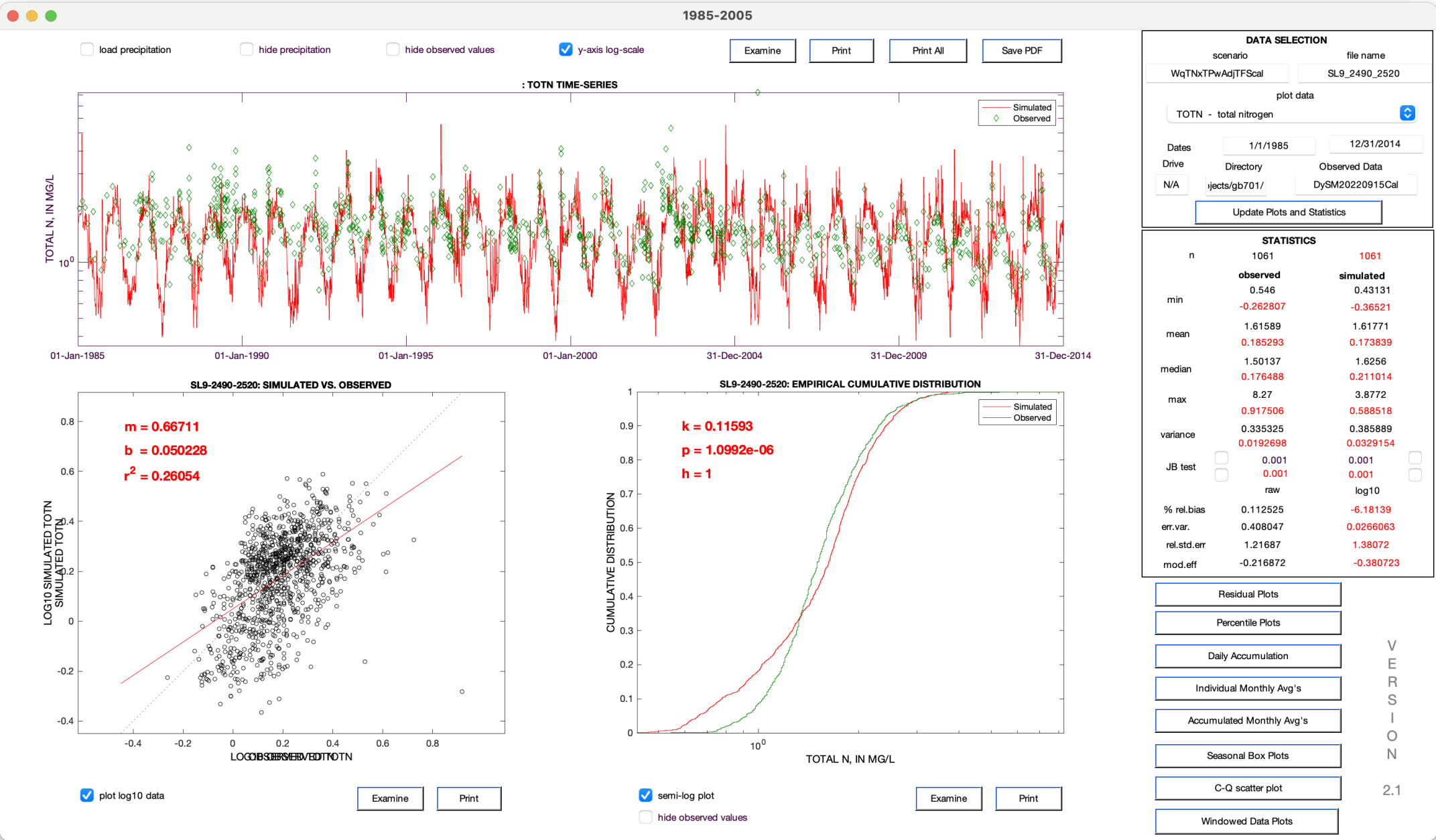
2. Initial model for the generalization of β_2 parameter for TN performed quite well. But it needs additional work for improving the model, including use of relevant explanatory variables.

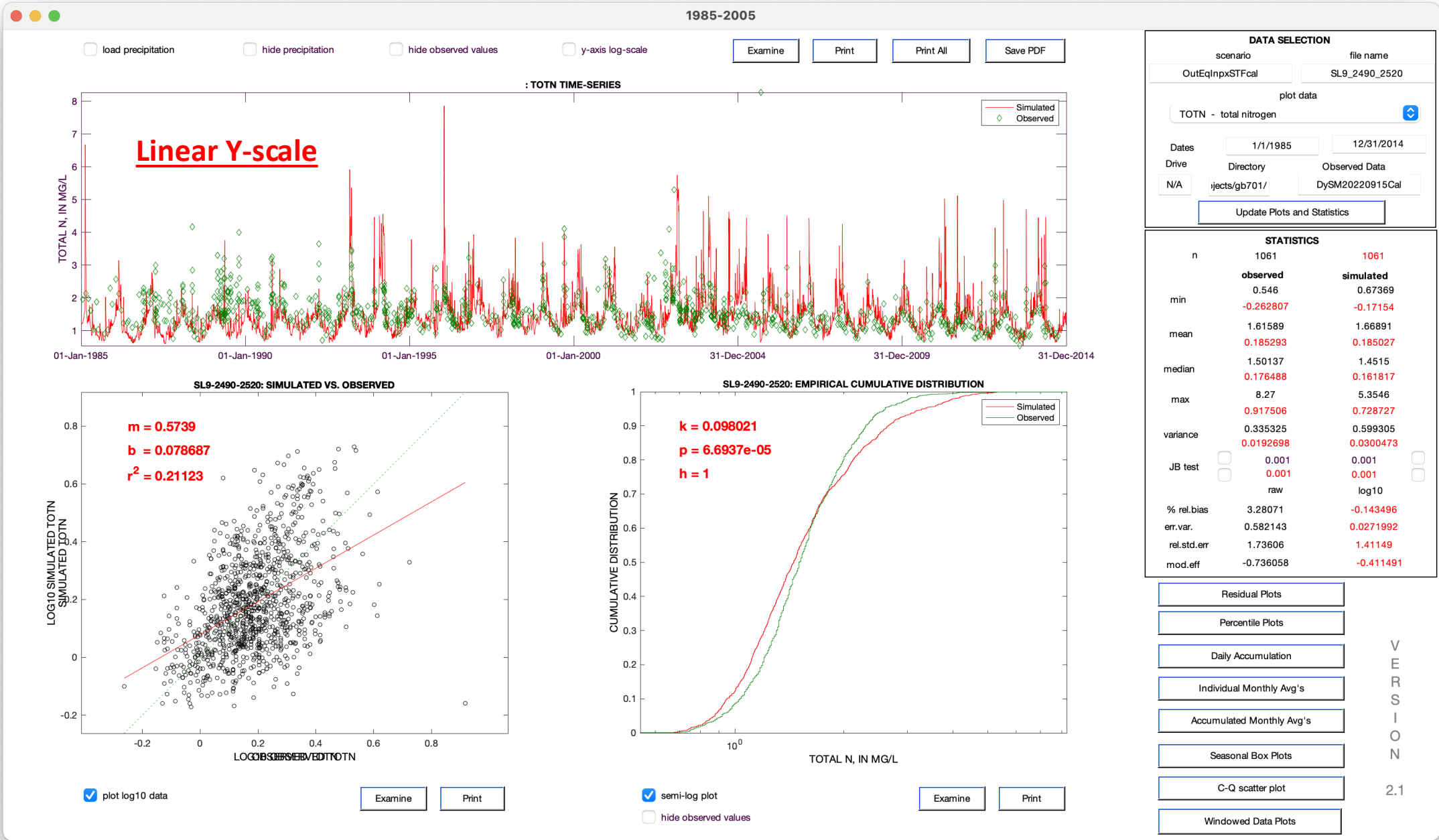
>> Next Steps for the Phase 7 Dynamic Watershed Model (DWM)

3. (a) Testing and revisions to calibration methods, and (b) incorporation of new data where appropriate.

Plots comparing the calibration of both (a) $\text{Output} = \text{Input} \times \text{stream transport factor (STF)}$ vs. (b) Small-stream routing for flow and water quality







DATA SELECTION

scenario

file name

OutEqInpxSTFcal

SL9_2490_2520

plot data

TOTN - total nitrogen

Dates

1/1/1985

12/31/2014

Drive

Directory

Observed Data

N/A

jects/gb701/

DySM20220915Cal

Update Plots and Statistics

STATISTICS

n

1061

1061

observed

simulated

min

0.546

0.67369

max

-0.262807

-0.17154

mean

1.61589

1.66891

median

0.185293

0.185027

variance

1.50137

1.4515

JB test

0.176488

0.161817

raw

log10

% rel.bias

8.27

5.3546

err.var.

0.917506

0.728727

rel.std.err

0.335325

0.599305

mod.eff

0.0192698

0.0300473

err.var.

0.001

0.001

rel.bias

0.001

0.001

err.var.

3.28071

-0.143496

rel.std.err

0.582143

0.0271992

mod.eff

1.73606

1.41149

mod.eff

-0.736058

-0.411491

Residual Plots

Percentile Plots

Daily Accumulation

Individual Monthly Avg's

Accumulated Monthly Avg's

Seasonal Box Plots

C-Q scatter plot

Windowed Data Plots

VERSION

2.1

☒ plot log10 data

Examine

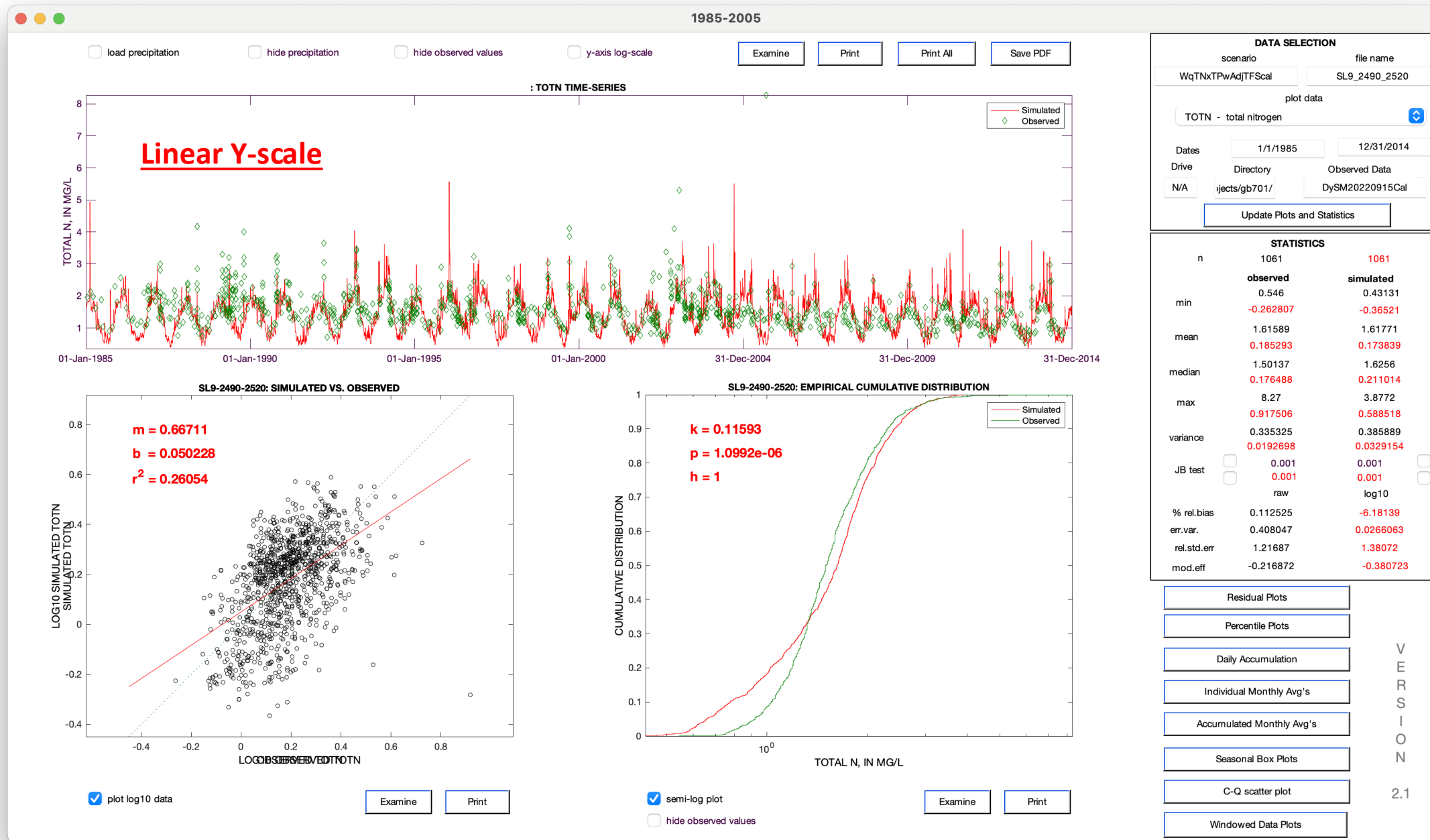
Print

☒ semi-log plot

☐ hide observed values

Examine

Print

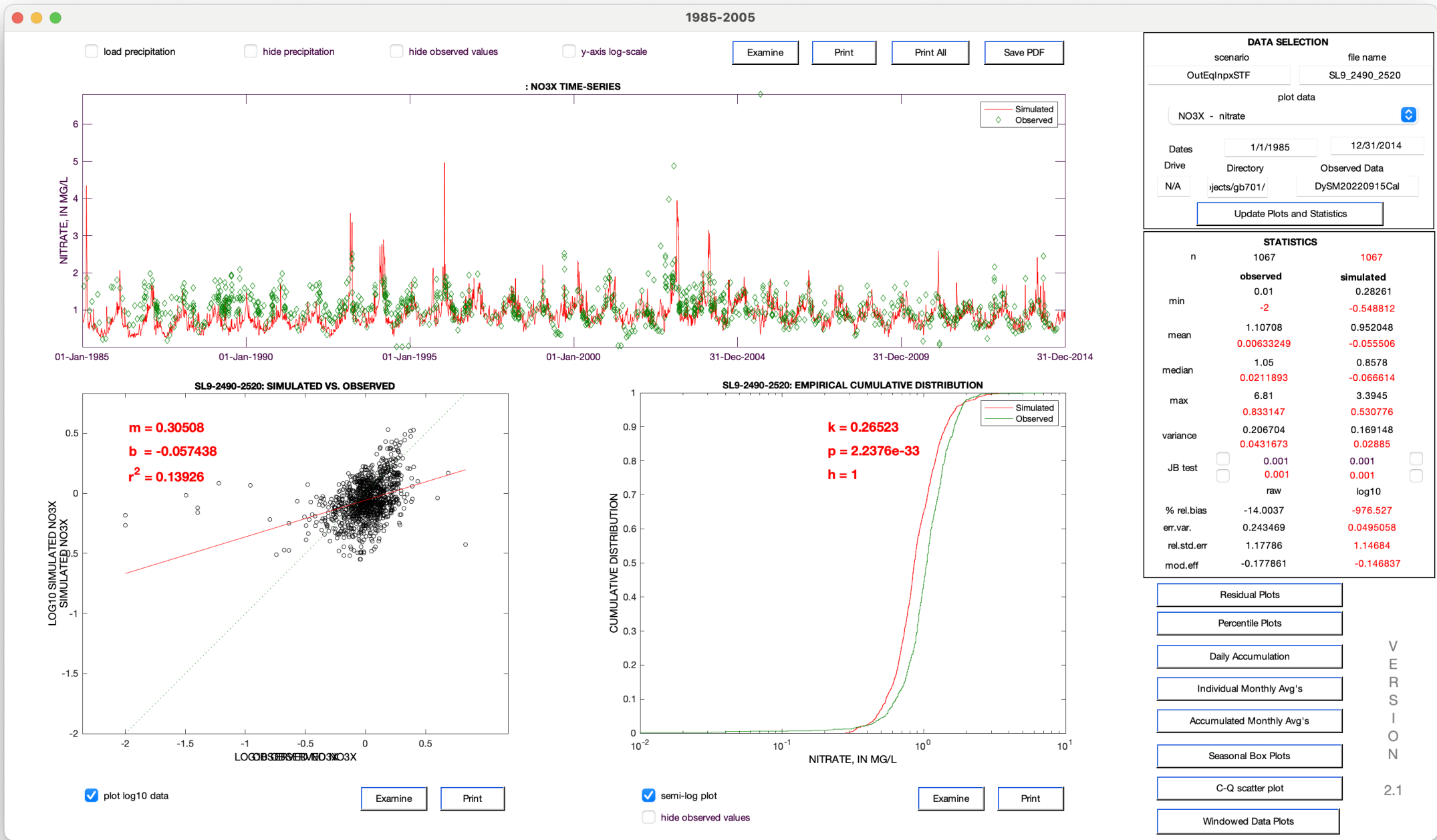


Calibration plots for the nitrate (before and after calibration)

Output = Input x STF

Susquehanna at Marietta

Nitrate



DATA SELECTION

scenario

file name

OutEqInpxSTF

SL9_2490_2520

plot data

NO3X - nitrate

Dates

1/1/1985

12/31/2014

Drive

Directory

Observed Data

N/A

jects/gb701/

DySM20220915Cal

Update Plots and Statistics

STATISTICS

n	1067	1067
	observed	simulated
min	0.01	0.28261
	-2	-0.548812
mean	1.10708	0.952048
	0.00633249	-0.055506
median	1.05	0.8578
	0.0211893	-0.066614
max	6.81	3.3945
	0.833147	0.530776
variance	0.206704	0.169148
	0.0431673	0.02885
JB test	<input type="checkbox"/> 0.001	<input type="checkbox"/> 0.001
	<input type="checkbox"/> 0.001	<input type="checkbox"/> 0.001
	raw	log10
% rel.bias	-14.0037	-976.527
err.var.	0.243469	0.0495058
rel.std.err	1.17786	1.14684
mod.eff	-0.177861	-0.146837

Residual Plots

Percentile Plots

Daily Accumulation

Individual Monthly Avg's

Accumulated Monthly Avg's

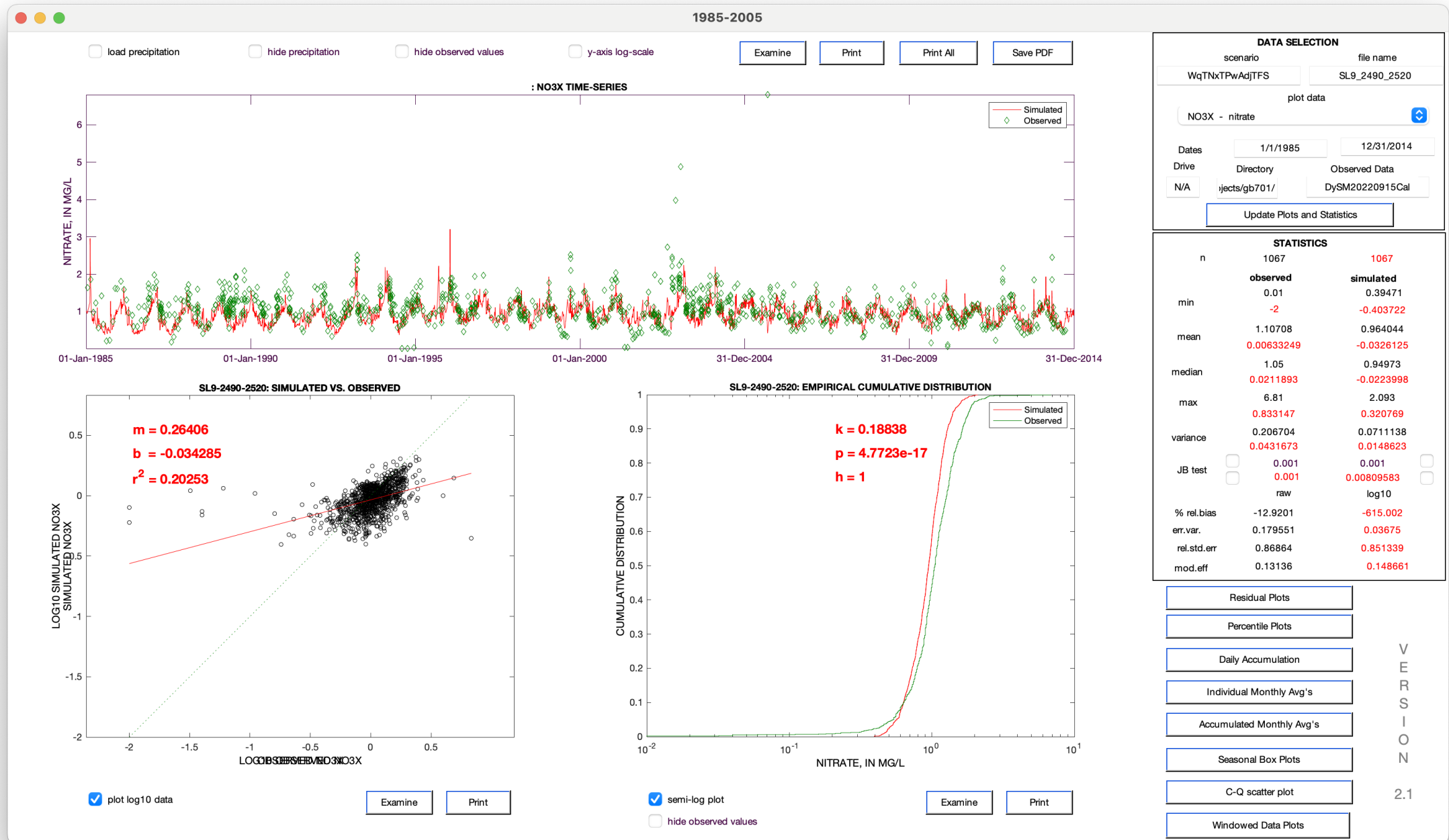
Seasonal Box Plots

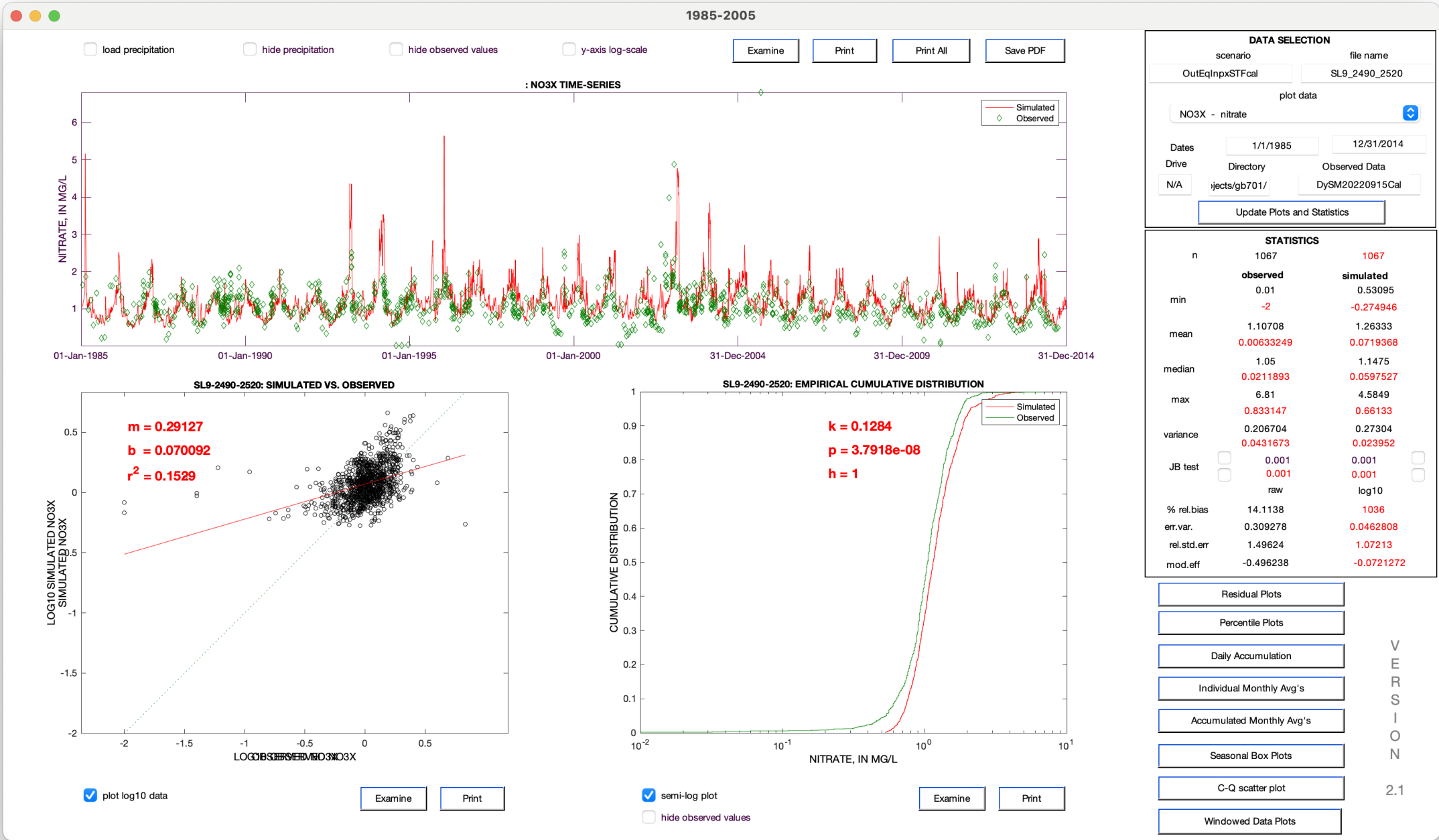
C-Q scatter plot

Windowed Data Plots

VERSION

2.1





DATA SELECTION

scenario

file name

OutEqInpxSTFcal

SL9_2490_2520

plot data

NO3X - nitrate

Dates

1/1/1985

12/31/2014

Drive

Directory

Observed Data

N/A

jects/gb701/

DySM20220915Cal

Update Plots and Statistics

STATISTICS

n	1067	1067
	observed	simulated
min	0.01	0.53095
	-2	-0.274946
mean	1.10708	1.26333
	0.00633249	0.0719368
median	1.05	1.1475
	0.0211893	0.0597527
max	6.81	4.5849
	0.833147	0.66133
variance	0.206704	0.27304
	0.0431673	0.023952
JB test	<input type="checkbox"/> 0.001	<input type="checkbox"/> 0.001
	<input type="checkbox"/> 0.001	<input type="checkbox"/> 0.001
	raw	log10
% rel.bias	14.1138	1036
err.var.	0.309278	0.0462808
rel.std.err	1.49624	1.07213
mod.eff	-0.496238	-0.0721272

Residual Plots

Percentile Plots

Daily Accumulation

Individual Monthly Avg's

Accumulated Monthly Avg's

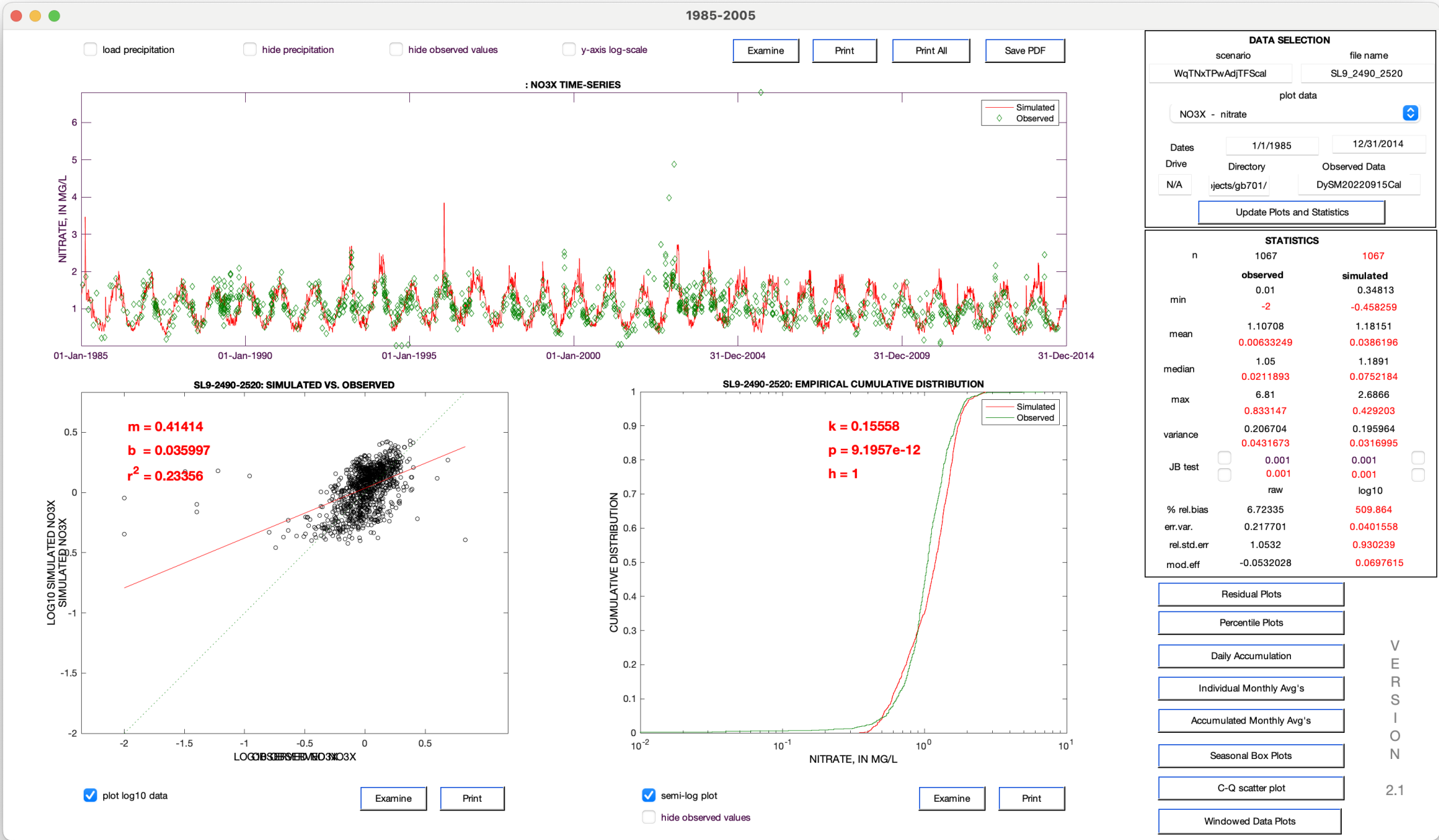
Seasonal Box Plots

C-Q scatter plot

Windowed Data Plots

VERSION

2.1

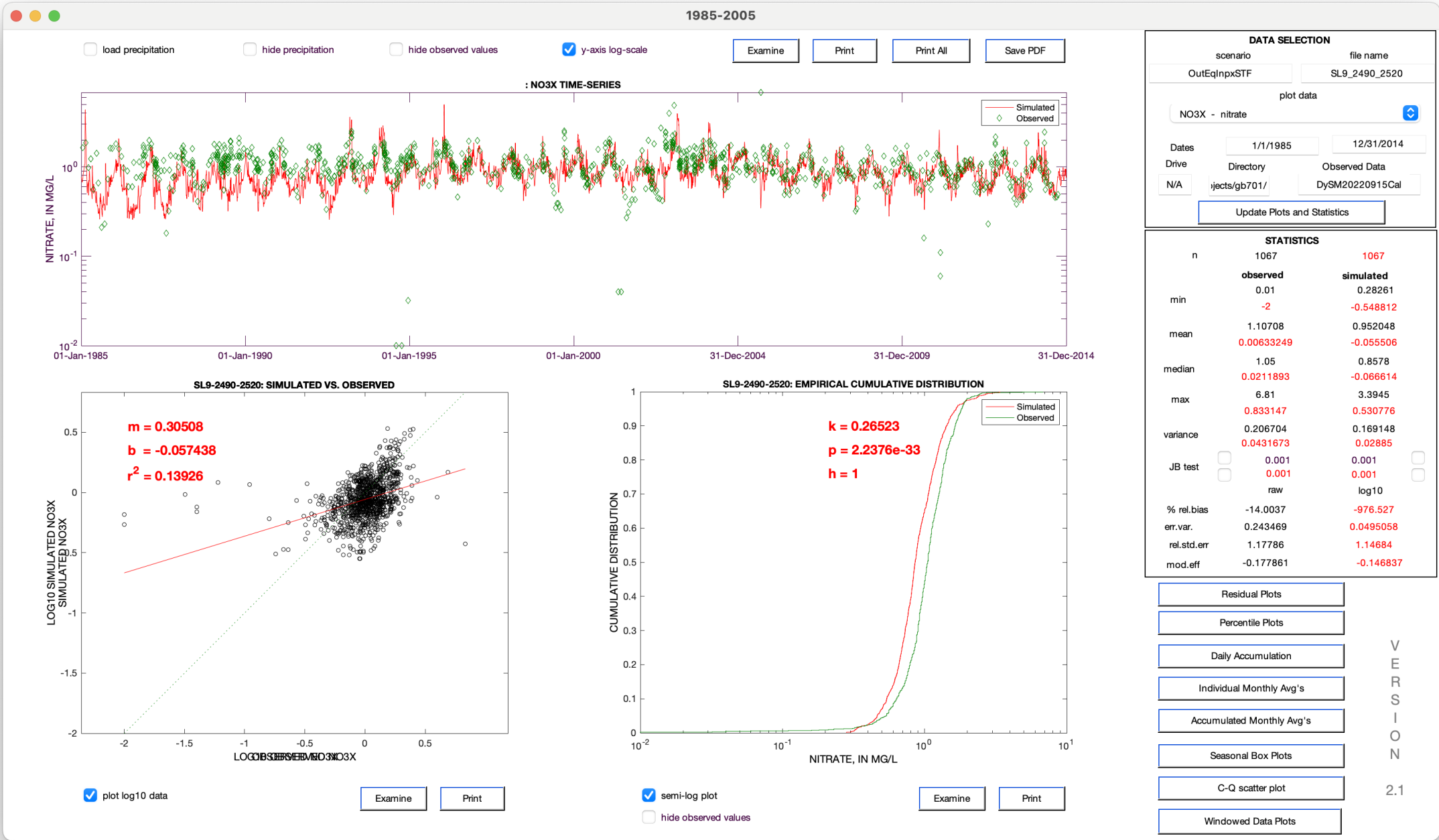


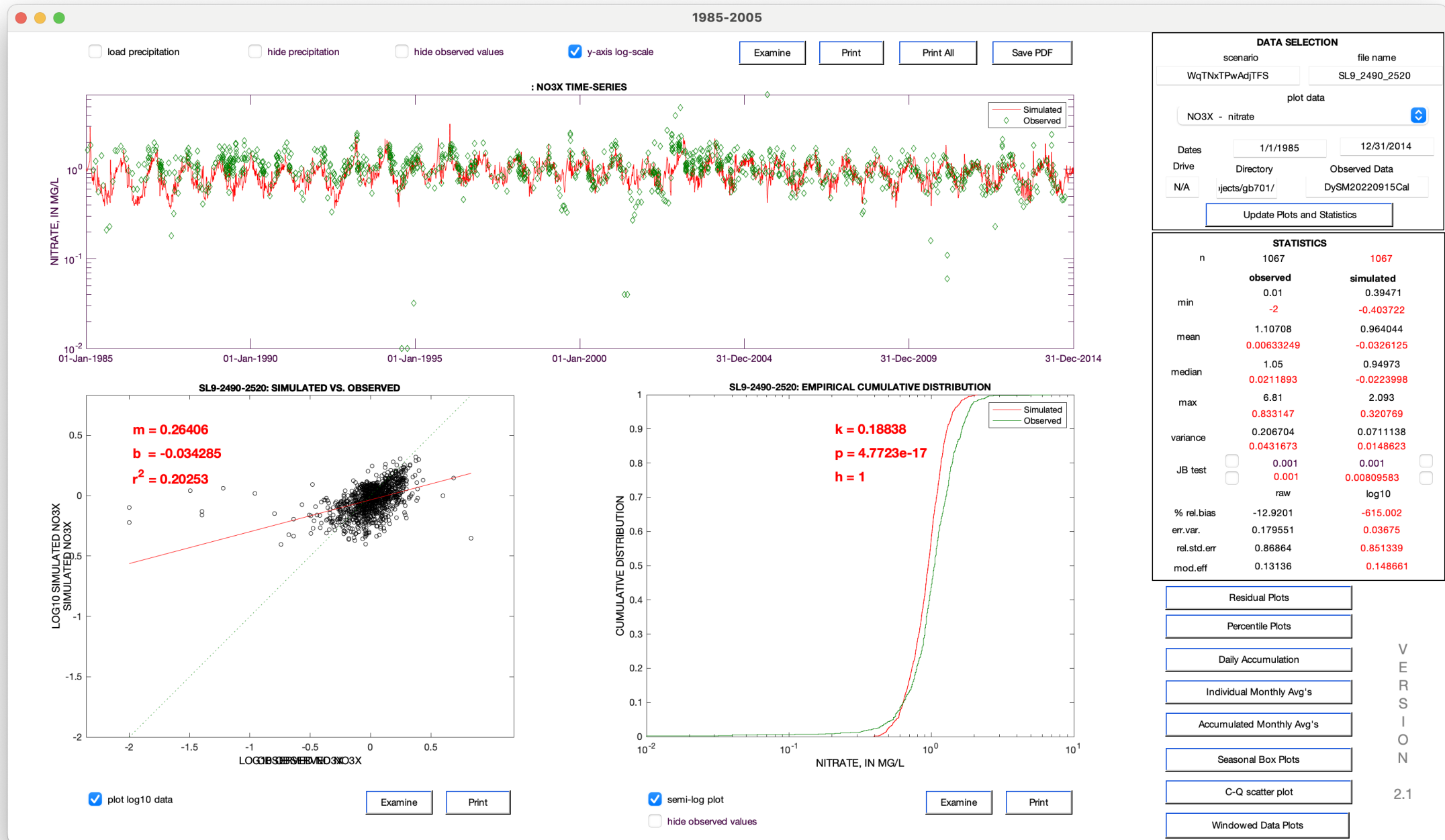
Calibration plots for the nitrate (before and after calibration)

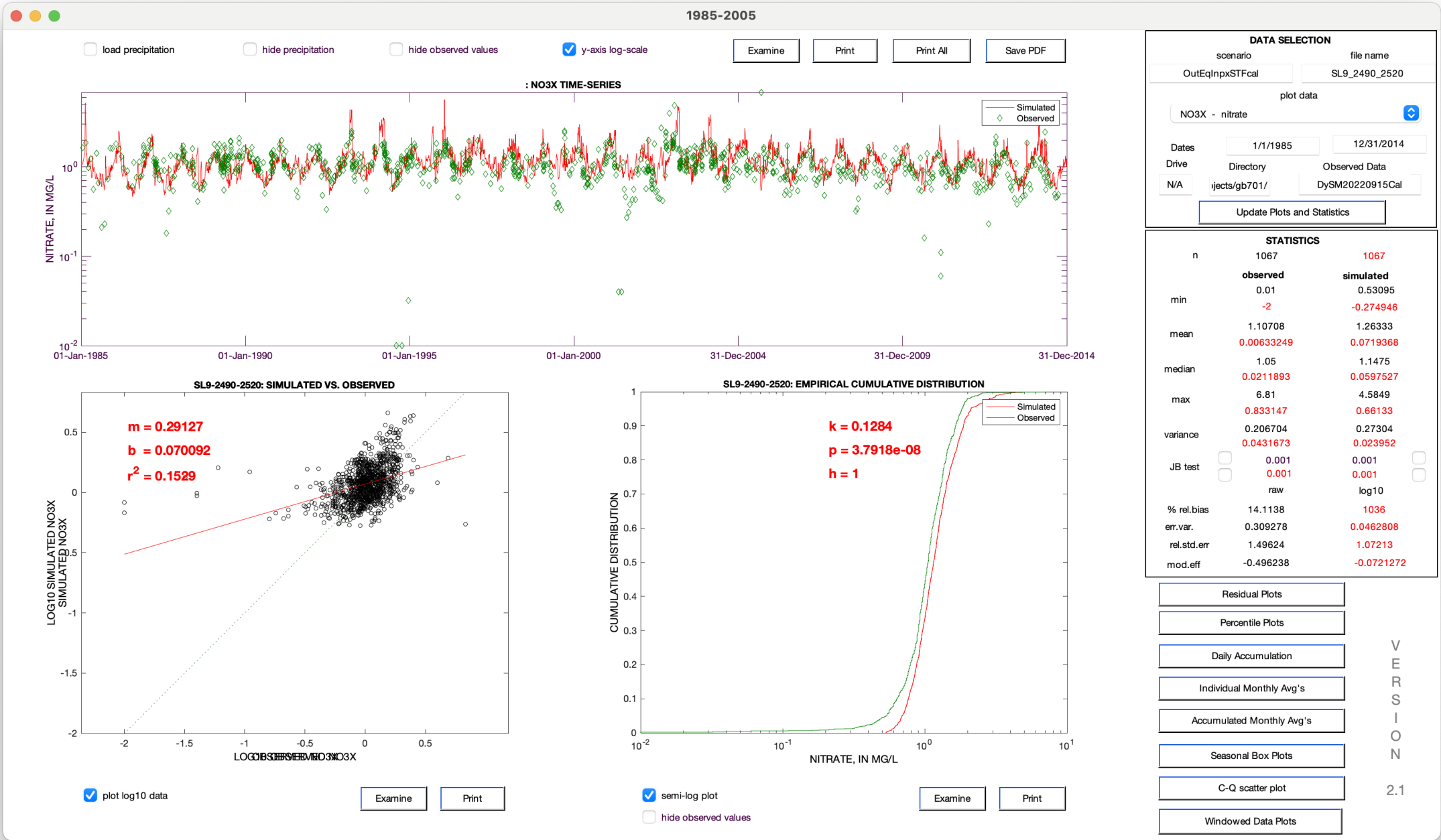
Output = Input x STF

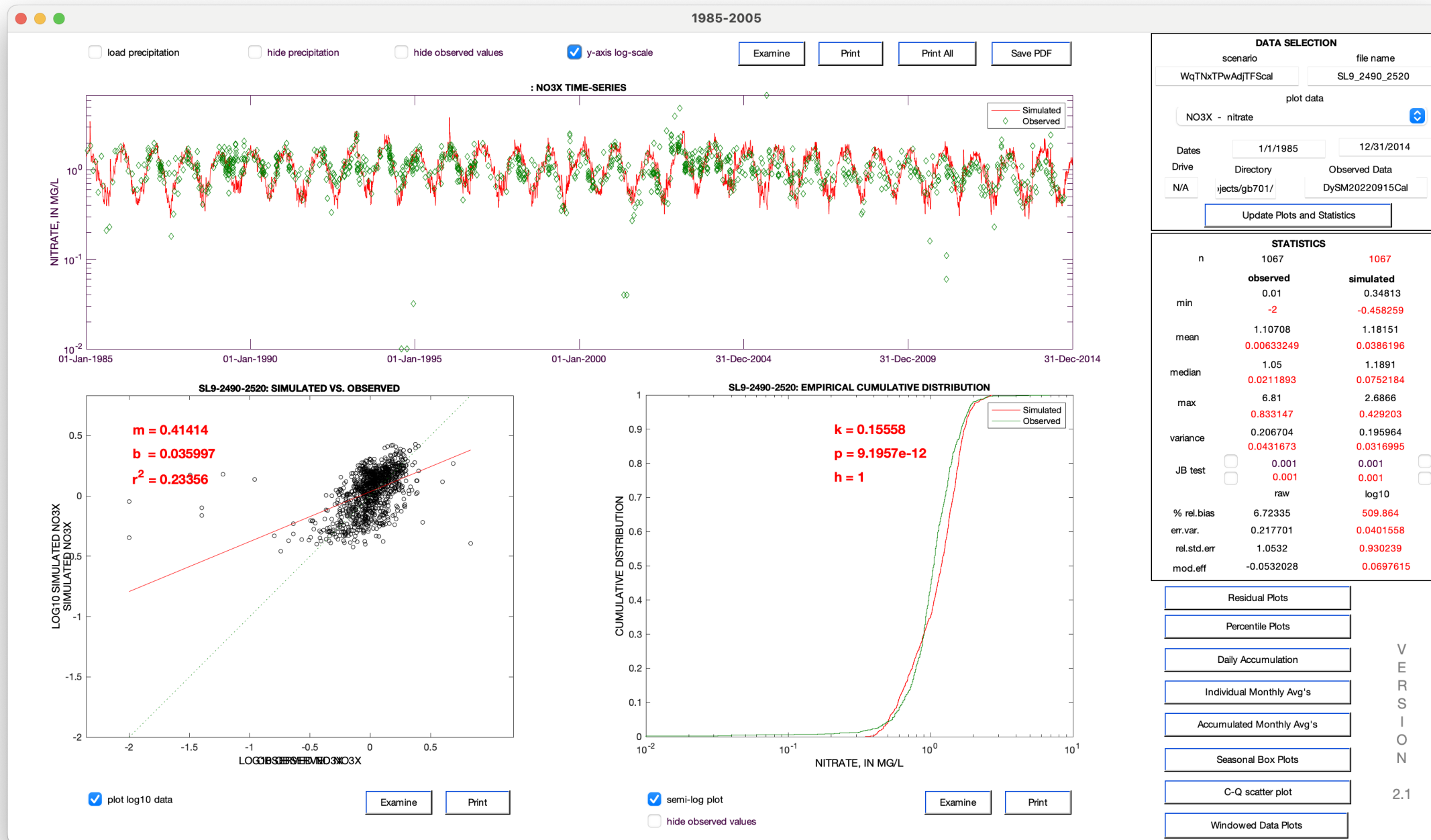
Susquehanna at Marietta

Nitrate









Scale and Simulation of Small Streams

**nested model
segmentation of streams
and river mainstems →**

100K NHD
STREAMS

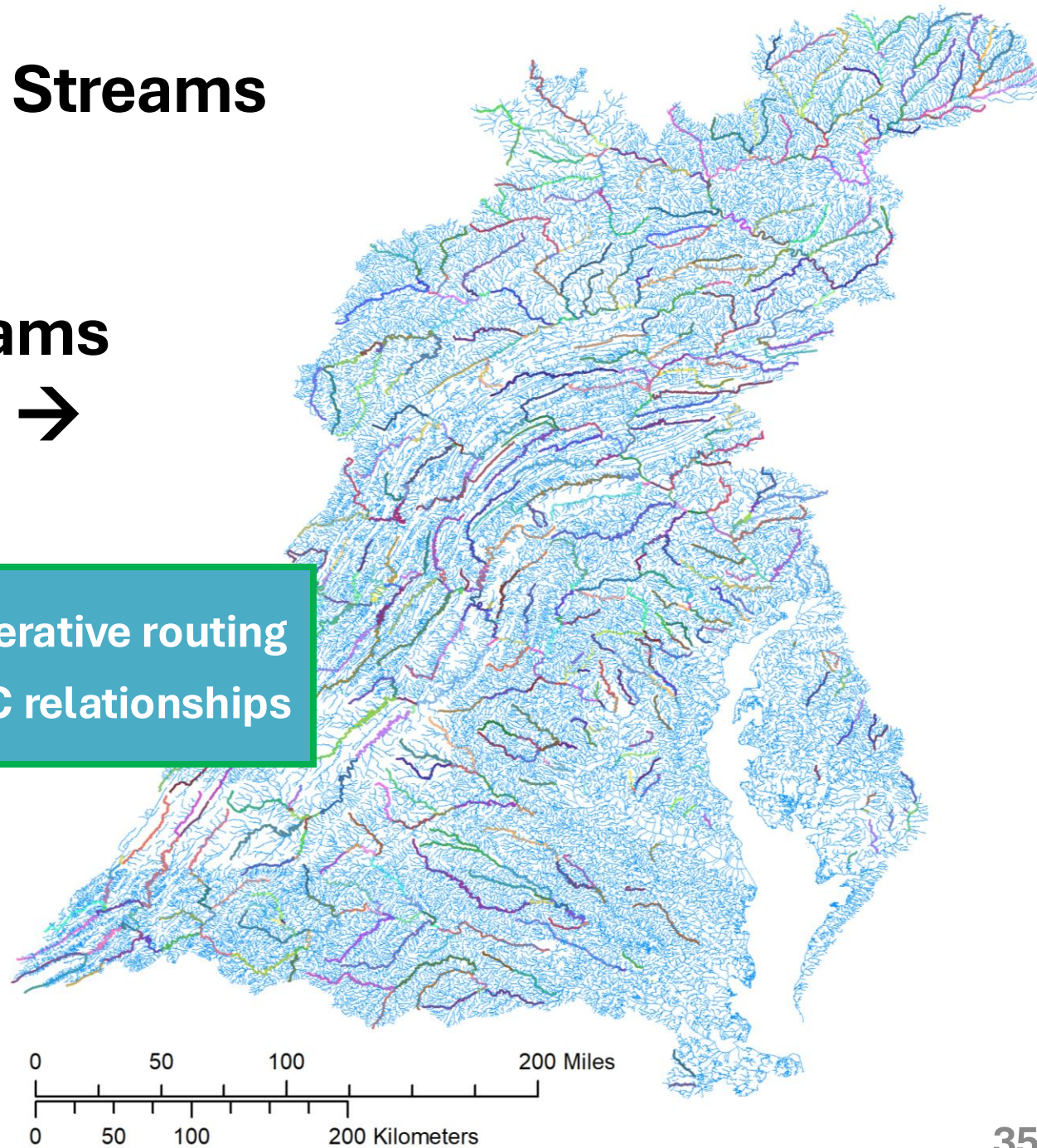


Hydrology: Non-iterative routing
Water Quality: Q-C relationships

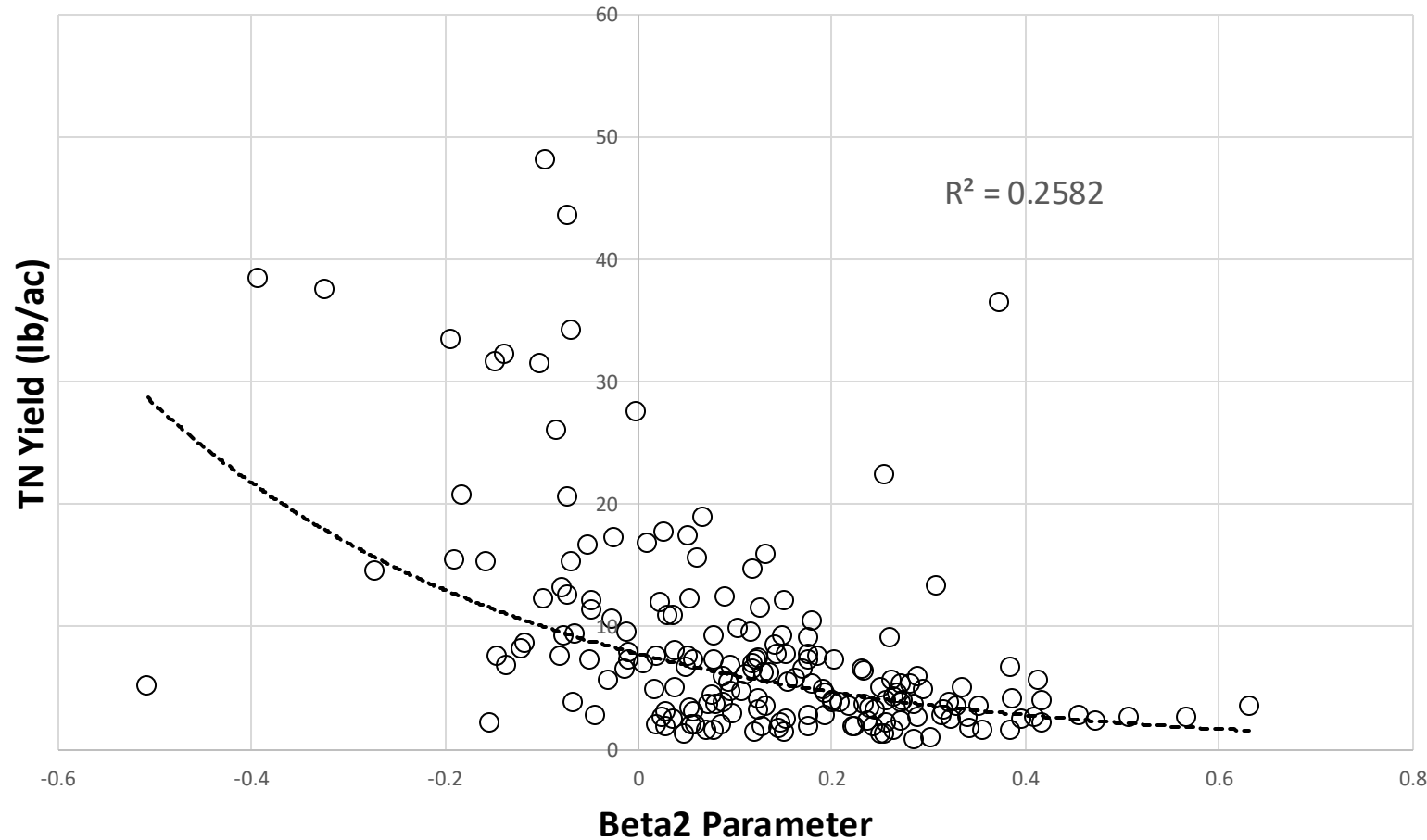
RIVER
MAINSTEM



HSPF



Explaining variability in β_2 parameter

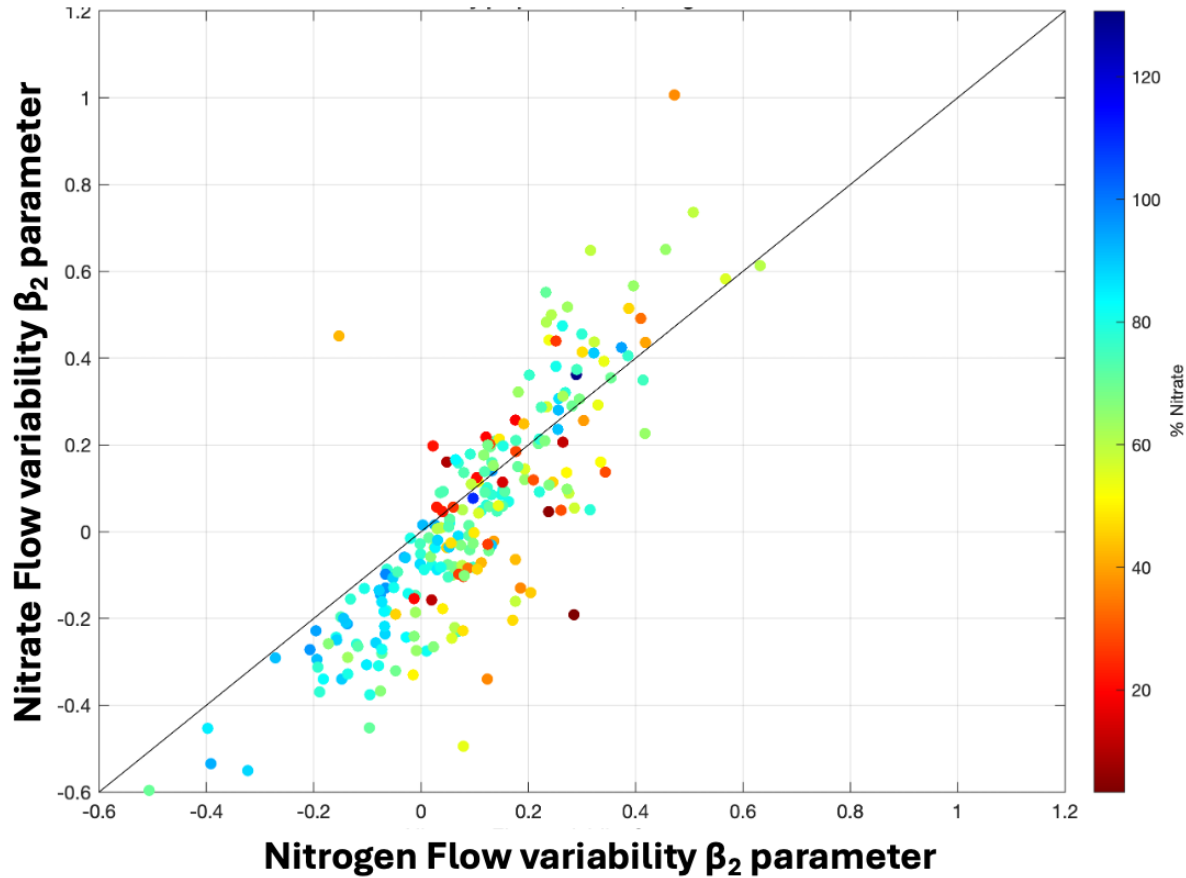


❖ We have not fully investigated explaining the variability in β_2 parameters.

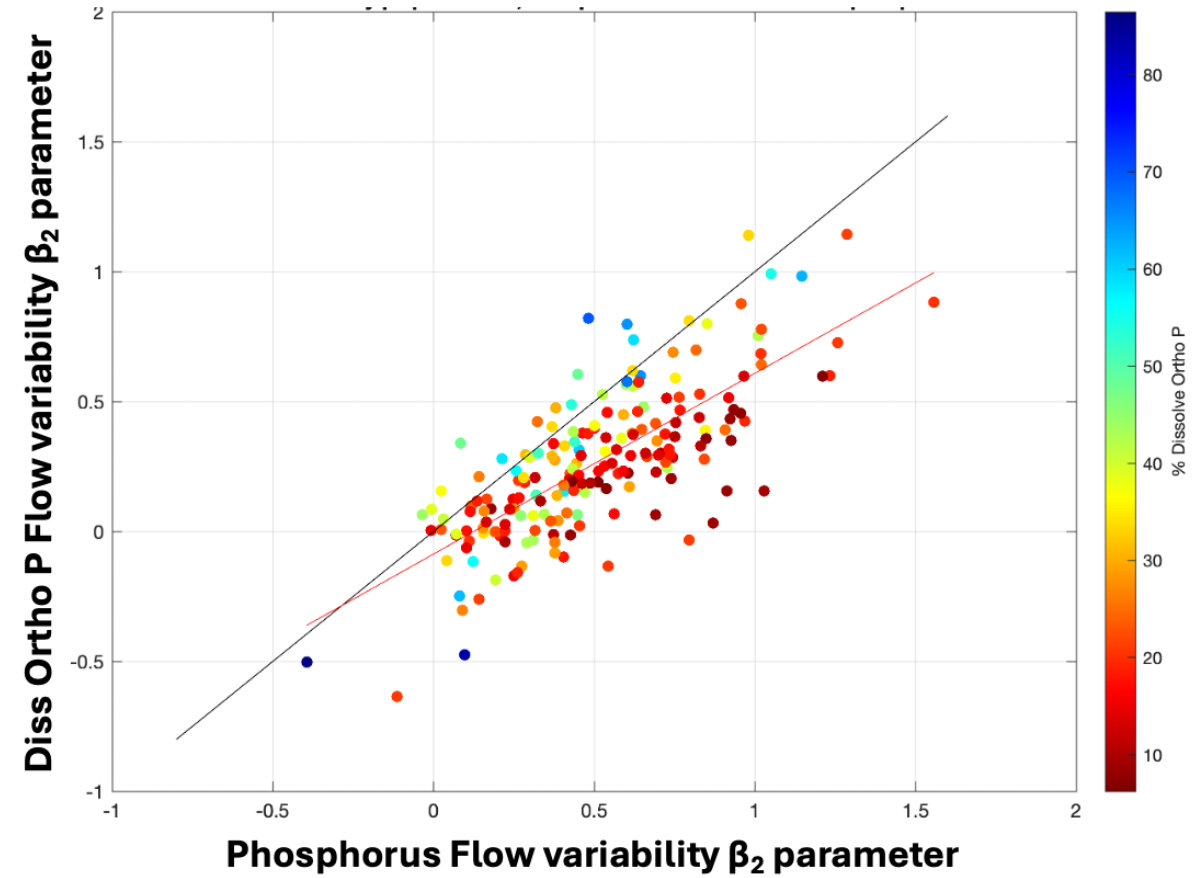
❖ But spatial patterns are intriguing and so we think use of watershed properties and better analysis tools should be successful.

Flow variability, β_2 parameter

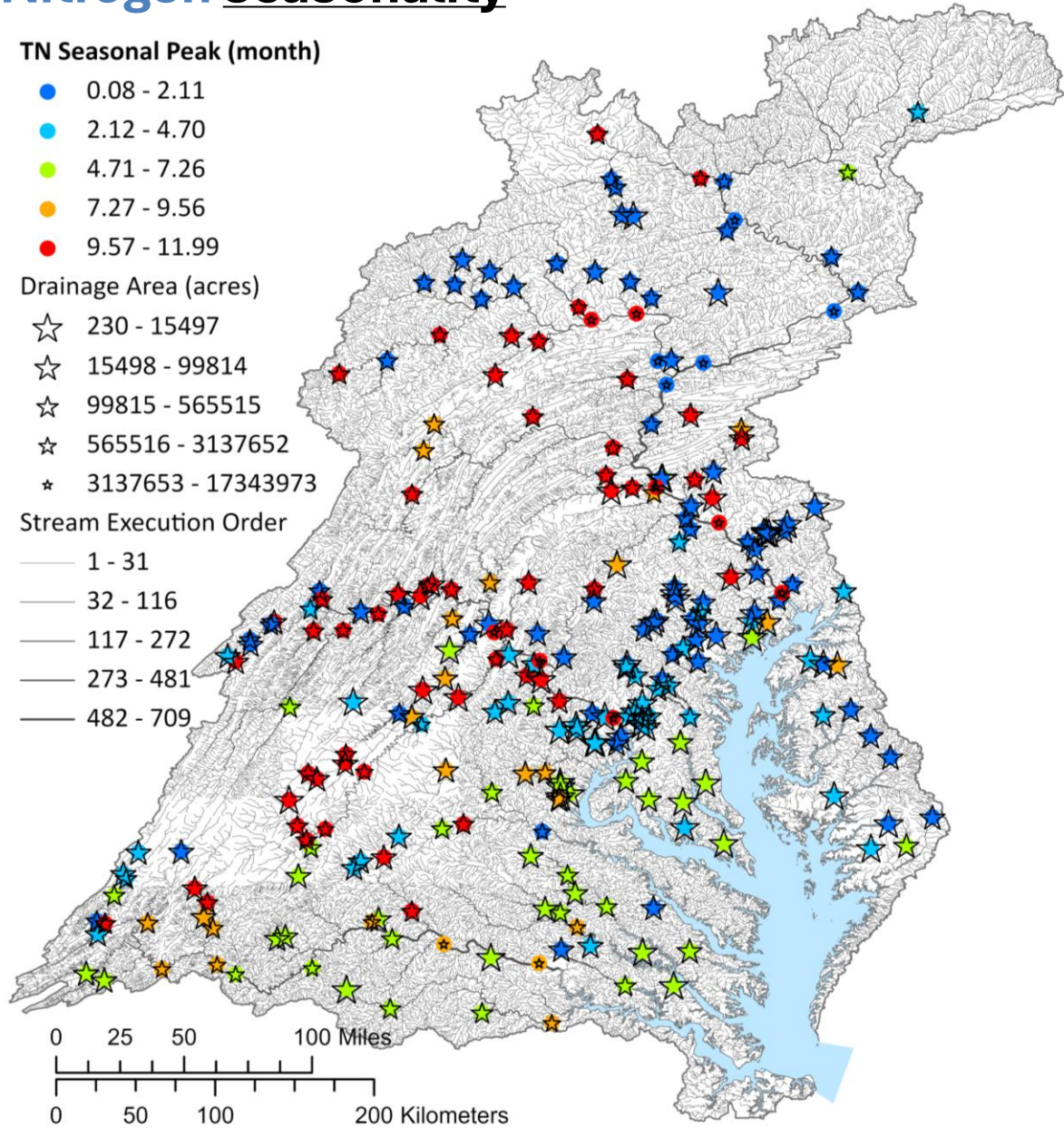
Nitrogen vs. Nitrate



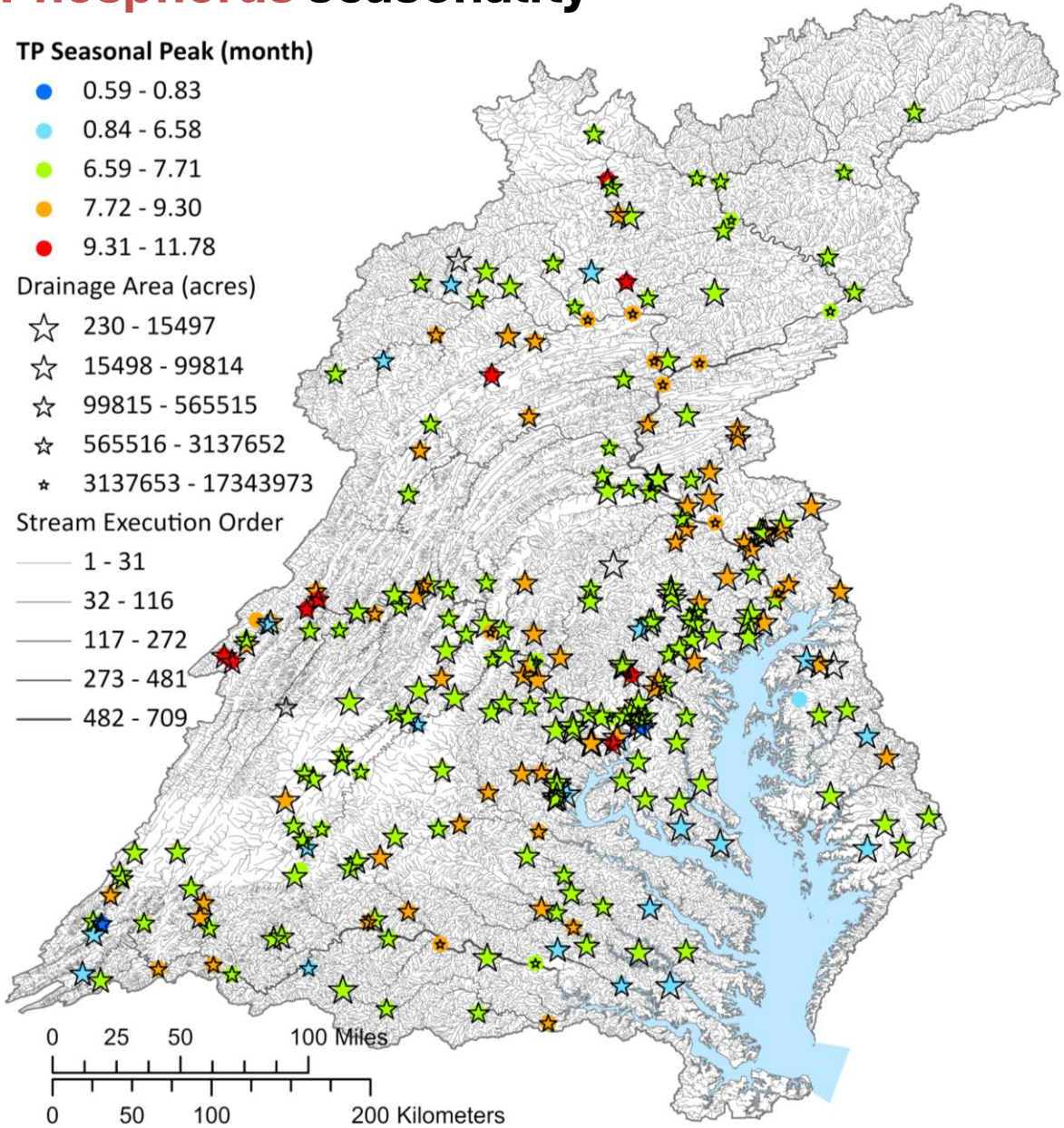
Phosphorus vs. Dissolved Orthophosphate



Nitrogen seasonality



Phosphorus seasonality



$$\ln(c) = \beta_o + \beta_1 t + \beta_2 \ln(Q) + \beta_3 \sin(2\pi t) + \beta_4 \cos(2\pi t) + \varepsilon$$