

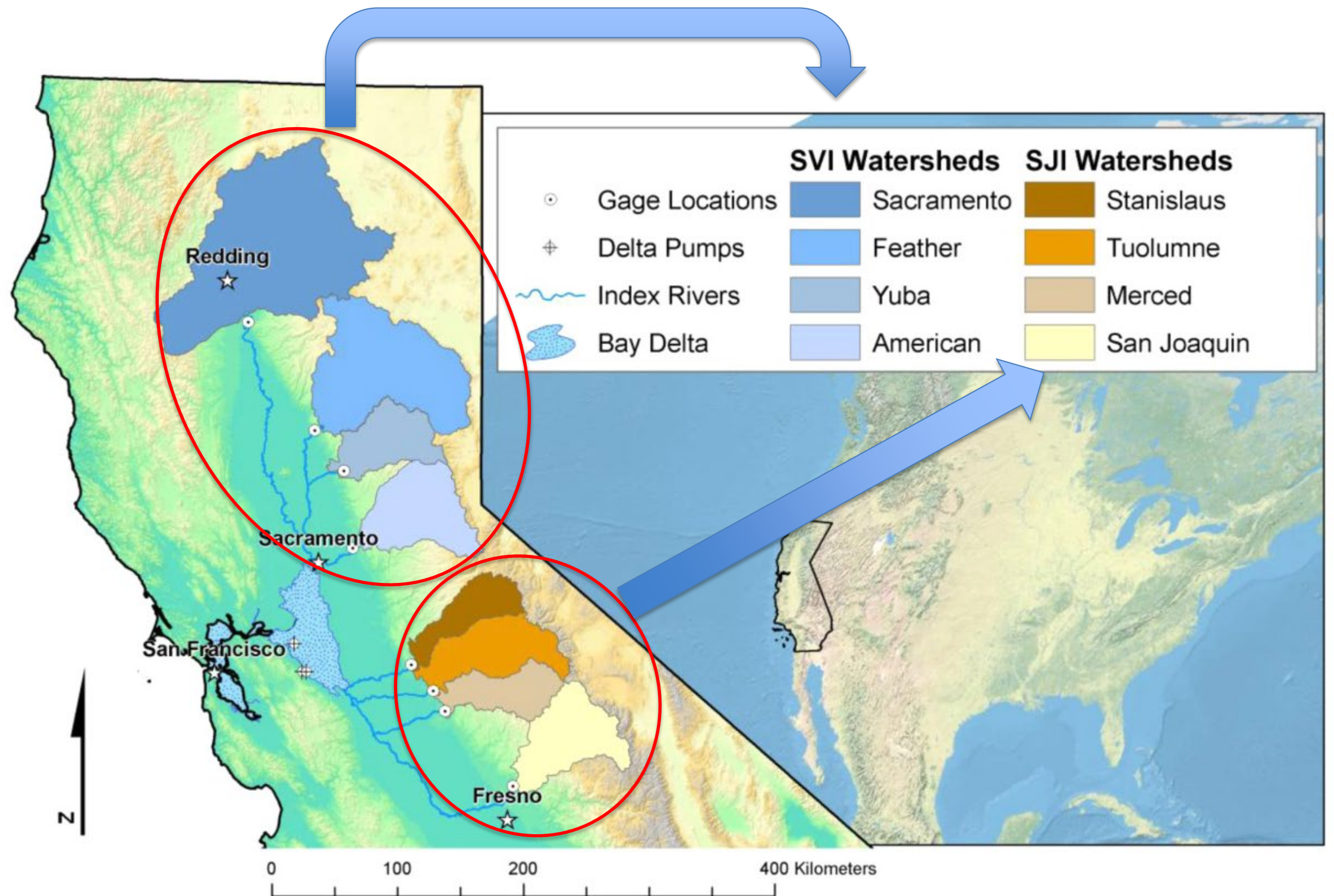
Applying AI/Machine Learning Algorithms for Forecasting California Water Year Types

Tariq Kadir Ph.D., P.E.
Civil Engineer (retired)

CWEMF 2026 Annual Meeting
Session: Machine Learning Applications in
Water Resources (Part 1)
February 22, 2026

Topics

1. Background
2. Prior work by presenter
3. Bulletin 120 Dec 1st Forecasts (WY2020-WY2026)
4. Applying ML/DL timeseries models
5. Investigating emerging AI timeseries forecasting models/frameworks
6. Final thoughts



Sacramento River above Bend Bridge (SBB)
 Feather River at Oroville (FTO)
 American River at Nimbus Dam (AMF)
 Yuba River near Smartsville (YRS)
 Stanislaus River at Goodwin (SNS)
 Tuolumne River at La Grange Dam (TLG)
 Merced River near Merced Falls (MRC)
 San Joaquin River Below Faint Dam (SBF)

Calculating Unimpaired Outflows (Runoff) *Key California Central Valley Upper Watersheds*

Key Indicator for California Hydrology is Water Year Type Index: *Wet, Above Normal, Below Normal, Dry, Critical*

SACRAMENTO VALLEY WATER YEAR TYPE INDEX 40-30-30 (SVI)

Index = 0.4 * Current Apr-Jul Runoff (1)
+ 0.3 * Current Oct-Mar Runoff (1)
+ 0.3 * Previous Year's Index (2)

Notes:

(1) Runoff is the sum of unimpaired flow in million acre-feet at:

- Sacramento River above Bend Bridge
- Feather River at Oroville (aka inflow to Lake Oroville)
- Yuba River near Smartville
- American River below Folsom Lake

(2) Maximum 10.0 for previous year index term

Year Type Classification: Index based on flow in million acre-feet:

Wet	Equal to or greater than 9.2
Above Normal	Greater than 7.8, and less than 9.2
Below Normal	Greater than 6.5, and equal to or less than 7.8
Dry	Greater than 5.4, and equal to or less than 6.5
Critical	Equal to or less than 5.4

This index, originally specified in the 1995 SWRCB Water Quality Control Plan, is used to determine the Sacramento Valley water year type as implemented in SWRCB D-1641.

<i>Sacramento Valley Index</i>
W >= 9.2
7.8 < AN < 9.2
6.5 < BN <= 7.8
5.4 < D <= 6.5
C <= 5.4

SAN JOAQUIN VALLEY WATER YEAR TYPE INDEX 60-20-20 (SJI)

Water Year Index based on flow in million acre feet

Index = 0.6 * Current Apr-Jul Runoff (1)
+ 0.2 * Current Oct-Mar Runoff (1)
+ 0.2 * Previous Year's Index (2)

Notes:

(1) Runoff is the sum of unimpaired flow in million acre-feet at:

- Stanislaus River below Goodwin Reservoir (aka inflow to New Melones Res.)
- Tuolumne River below La Grange (aka inflow to New Don Pedro Reservoir)
- Merced River below Merced Falls (aka inflow to Lake McClure)
- San Joaquin River inflow to Millerton Lake

(2) Maximum 4.5 for previous year index term

Year Type Classification: Index based on flow in million acre-feet:

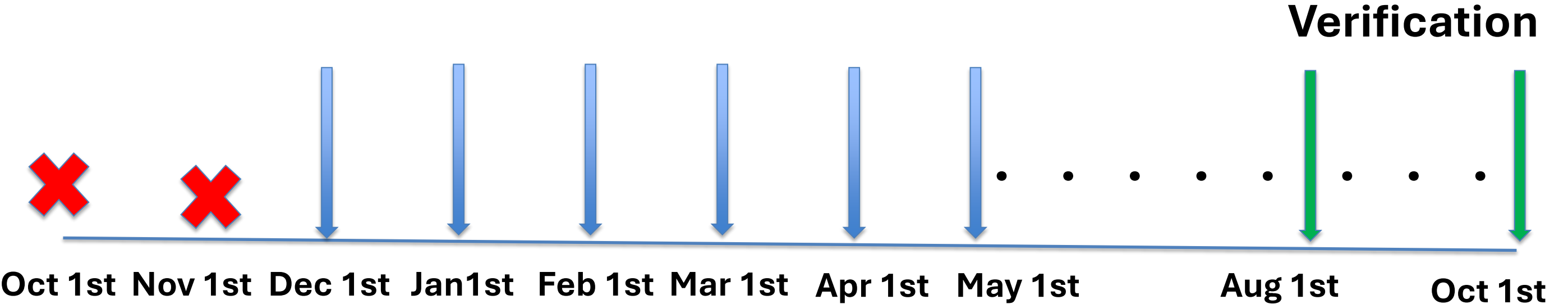
Wet	Equal to or greater than 3.8
Above Normal	Greater than 3.1, and less than 3.8
Below Normal	Greater than 2.5, and equal to or less than 3.1
Dry	Greater than 2.1, and equal to or less than 2.5
Critical	Equal to or less than 2.1

This index, originally specified in the 1995 SWRCB Water Quality Control Plan, is used to determine the San Joaquin Valley water year type as implemented in SWRCB D-1641.

<i>San Joaquin Valley Index</i>
W >= 3.8
3.1 < AN < 3.8
2.5 < BN <= 3.1
2.1 < D <= 2.5
C <= 2.1

All values in millions of acre-feet (MAF)

DWR's Bulleting 120 Water Supply Index Forecast Schedule



Potential Applications:

1. Improving Bulletin 120 forecasts.
2. Real time reservoir operations.
3. Planning (projects, water districts, SGAs, farmers).
4. Modeling (e.g., in reservoir operations, or forecasting next year crop types)

Prior Work by TK (circa 2020) for Water Year Type Using Deep Learning ANNs Using Last 120 months for Forecasting next 12 Months of Streamflows

Older research by presenter

Sacramento Valley Index

WY	Reported	Predicted (3 years back data)	Predicted (10 years back data)
1976	C	AN	
1977	C	C	
1978	AN	D	
1979	BN	BN	
1980	AN	AN	
1981	D	BN	
2004	BN	BN	
2005	AN	D	
2006	W	AN	
2007	D	AN	
2008	C	AN	
2009	D	D	
2010	BN	D	BN
2011	W	W	AN
2012	BN	AN	BN
2013	D	AN	D
2014	C	BN	C
2015	C	C	C
2016	BN	C	BN
2017	W	AN	W
2018	BN	AN	BN
2019	W	W	W
2020	D	D	D

% correct 36.4 90.9

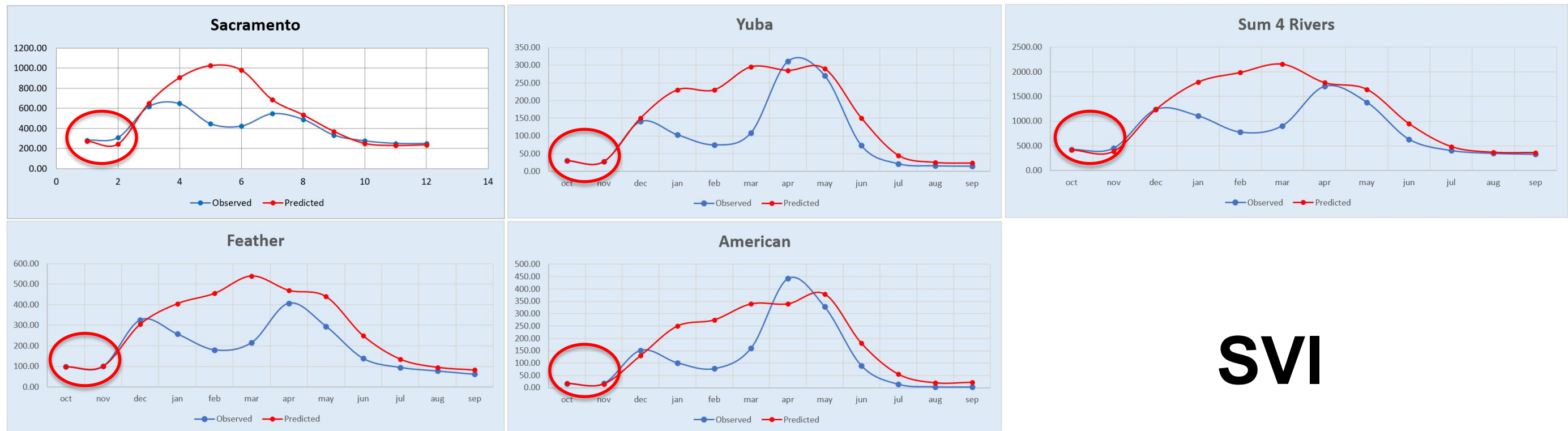
San Joaquin Valley Index

WY	Reported	Predicted (3 years back data)	Predicted (10 years back data)
1976	C	BN	
1977	C	C	
1978	W	D	
1979	AN	AN	
1980	W	W	
1981	D	AN	
2004	D	BN	
2005	W	BN	
2006	W	W	
2007	C	AN	
2008	C	AN	
2009	BN	BN	
2010	AN	BN	AN
2011	W	W	W
2012	D	AN	D
2013	C	BN	C
2014	C	D	C
2015	C	C	C
2016	D	C	D
2017	W	D	W
2018	BN	BN	BN
2019	W	W	W
2020	D	BN	D

% correct 45.5 100.0

Bulleting 120 using December 1st Forecasts

WY2020 (based on Dec 1st 2019 forecast)



SVI

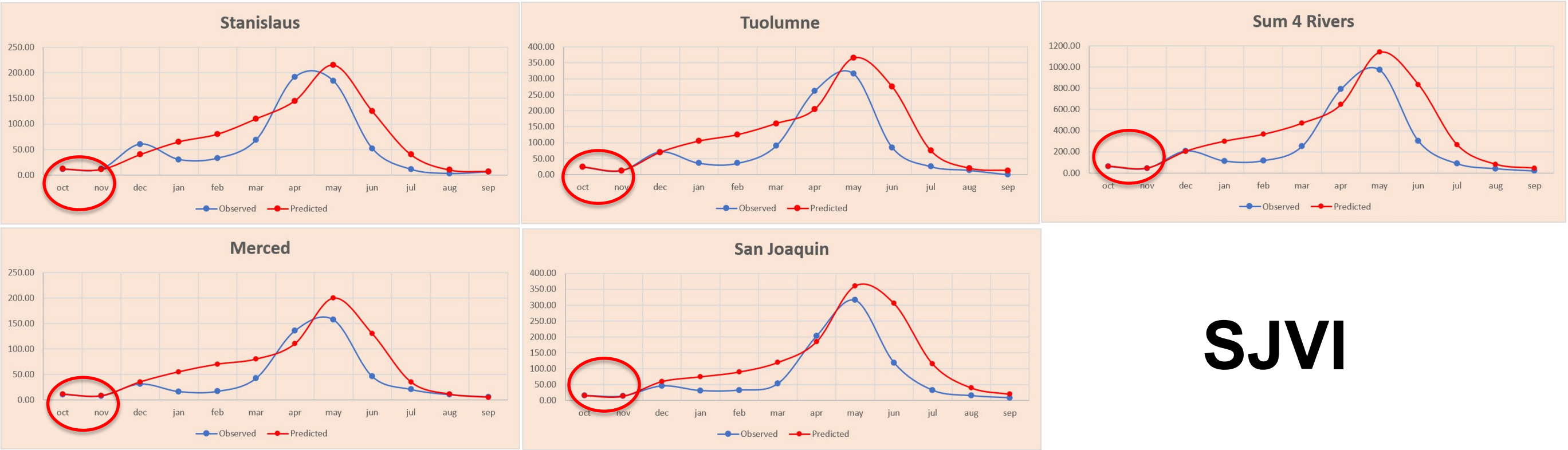
mon	Observed	Predicted
Reported PrevYrIndex	10.34	10.34
Calculated Index	6.12	7.34
Calculated WY Class	D	BN

Year Type Classification: Index based on flow in million acre-feet:

Wet	Equal to or greater than 9.2
Above Normal	Greater than 7.8, and less than 9.2
Below Normal	Greater than 6.5, and equal to or less than 7.8
Dry	Greater than 5.4, and equal to or less than 6.5
Critical	Equal to or less than 5.4

Monthly flows in TAF
Index in MAF

WY2020 (based on Dec 1st 2019 forecast)



SJVI

mon	Observed	Predicted
Reported PrevYrIndex	4.94	
Calculated Index	2.35	2.92
Calculated WY Class	D	BN

Year Type Classification: Index based on flow in million acre-feet:

Wet	Equal to or greater than 3.8
Above Normal	Greater than 3.1, and less than 3.8
Below Normal	Greater than 2.5, and equal to or less than 3.1
Dry	Greater than 2.1, and equal to or less than 2.5
Critical	Equal to or less than 2.1

Monthly flows in TAF
Index in MAF

Summary of Bulletin 120 Dec 1st Forecasts of Water Year Type

SVI		
WY	Actual	Forecasted
2020	<i>D</i>	<i>BN</i>
2021	<i>C</i>	<i>D</i>
2022	<i>C</i>	<i>D</i>
2023	<i>W</i>	<i>D</i>
2024	<i>AN</i>	<i>BN</i>
2025	<i>AN</i>	<i>AN</i>
2026	?	AN

SJVI		
WY	Actual	Forecasted
2020	<i>D</i>	<i>BN</i>
2021	<i>C</i>	<i>D</i>
2021	<i>C</i>	<i>BN</i>
2023	<i>W</i>	<i>BN</i>
2024	<i>AN</i>	<i>BN</i>
2025	<i>AN</i>	<i>BN</i>
2026	?	BN

Forecasting Streamflows (and WY Types) for WY2020 through WY2025(6)

Model Development Guidelines Followed for Forecasting:

1. Forecast is on October 1st of a water year.
2. Use only past streamflow data to forecast next 12 months of streamflows for a stream.
3. Keep model structure the same for all years.
4. Sacramento Valley Index: forecast each of 4 streams - Sacramento (SBB), Feather (FTO), American (AMF), Yuba (YRS)
5. San Joaquin Valley Index: forecast each of 4 streams – Stanislaus (SNS), Tuolumne (TLG), Merced (MRC), San Joaquin below Friant (SBF)
6. Use SWRCB formulae to compute Water Year Index, and then classify Water Year Type
7. Compare results to Bulletin 120 December 1st forecasts.

Models: Statistical, Machine Learning, and Deep Learning

(WY2020-2026)

Statistical Forecasting Models

Autoregressive Integrated Moving Average **ARIMA** (1970's)

Seasonal Autoregressive Integrated Moving Average **SARIMA** (1970's)

Holt-Winters' Exponential Smoothing (circa 1957-1960)

Theta Method (2000)

Machine Learning Models

Light Gradient Boosting Machine **LightGBM** (2017)

Deep Learning Models

Neural Hierarchical Interpolation for Time Series Forecasting **N-HiTS** (2022)

Neural Basis Expansion Analysis for Interpretable Time Series forecasting **N-BEATS** (2019)

Time-series Dense Encoder **TiDE** (2023)

Normalized Linear **N-Linear** (2022)

Decomposition Linear **D-Linear** (2022)

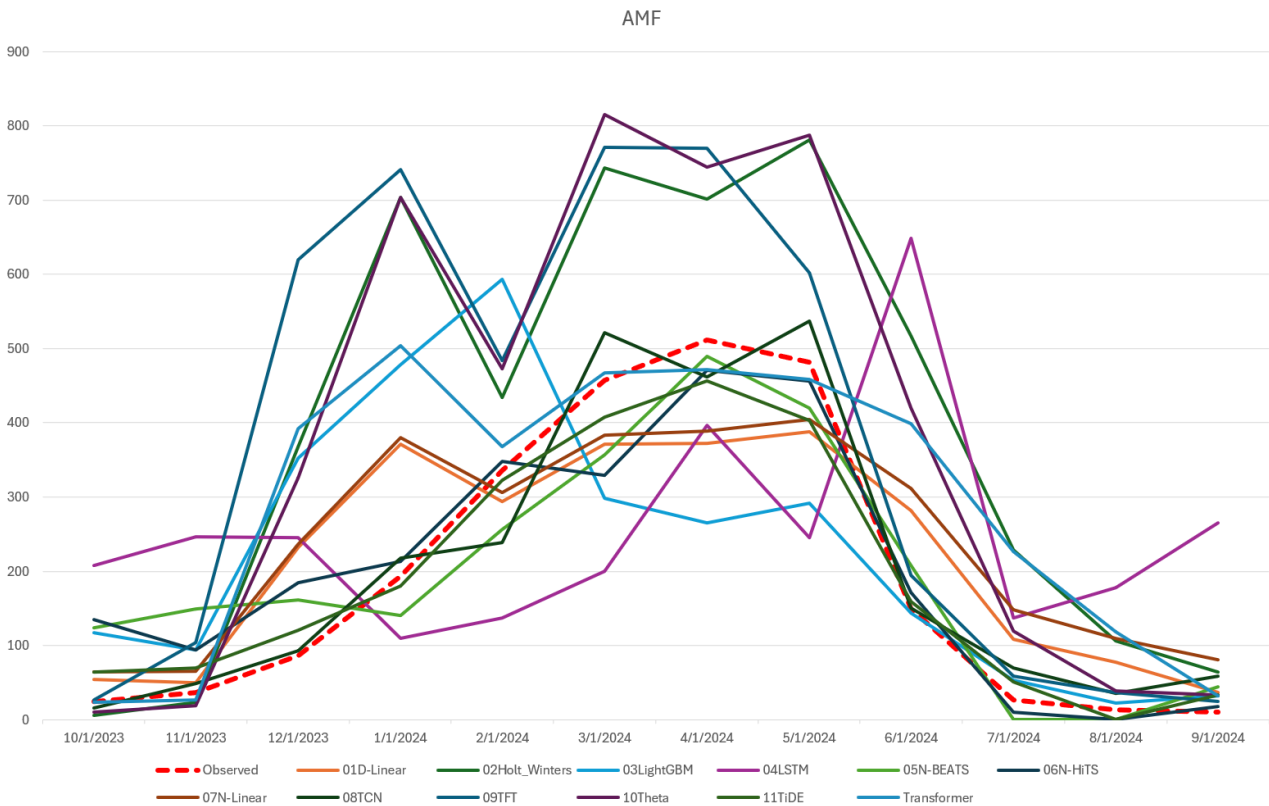
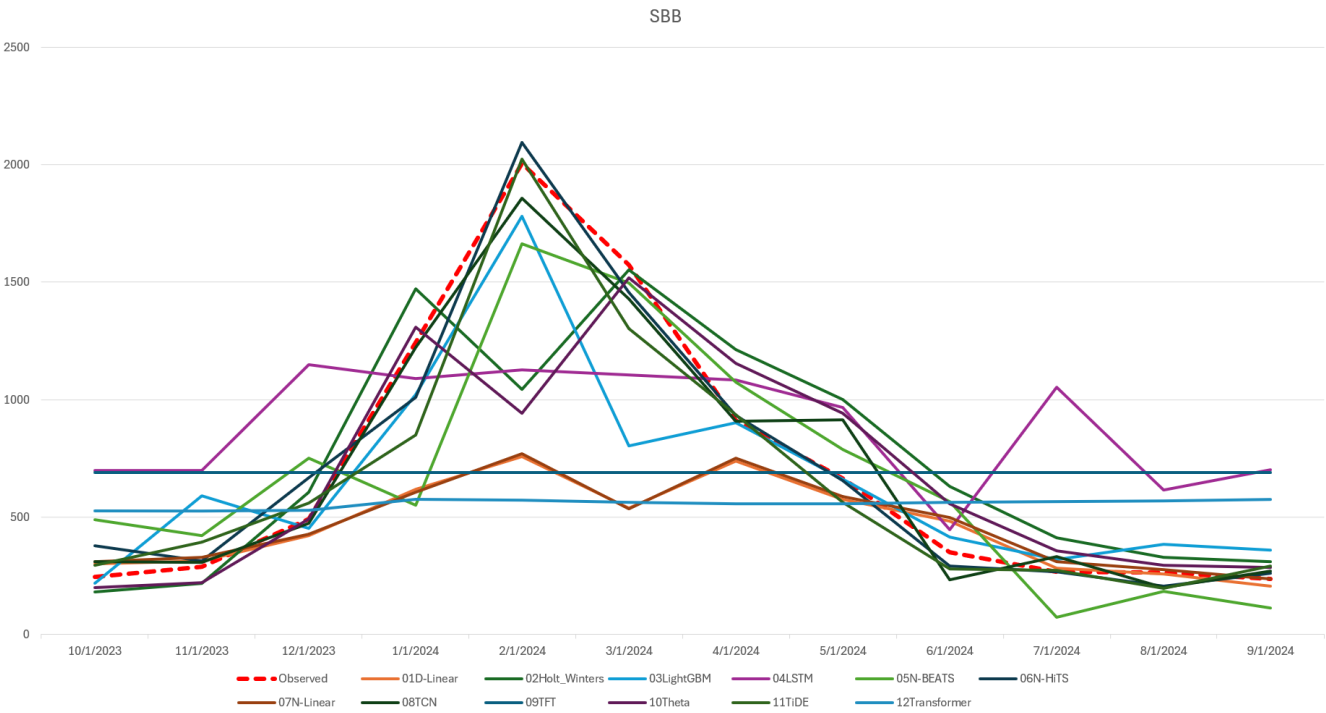
Long Short-term Memory **LSTM** (1997)

Temporal Convolutional Network **TCN** (2018)

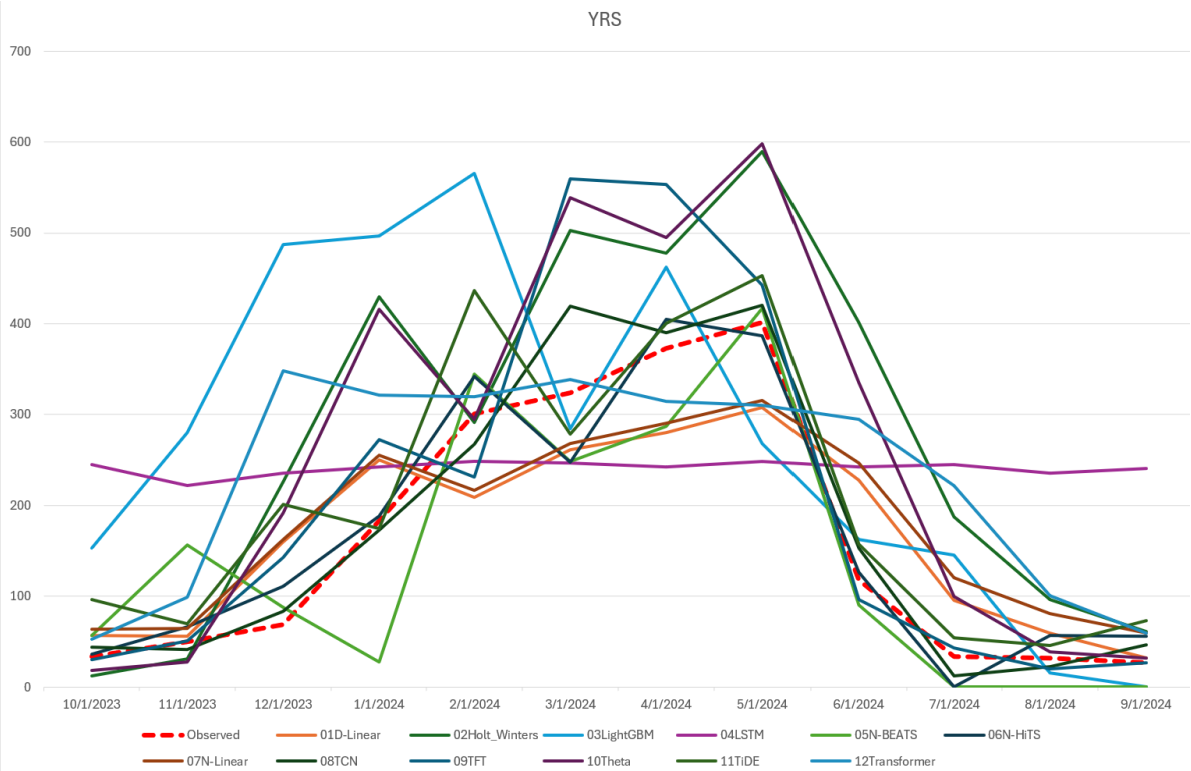
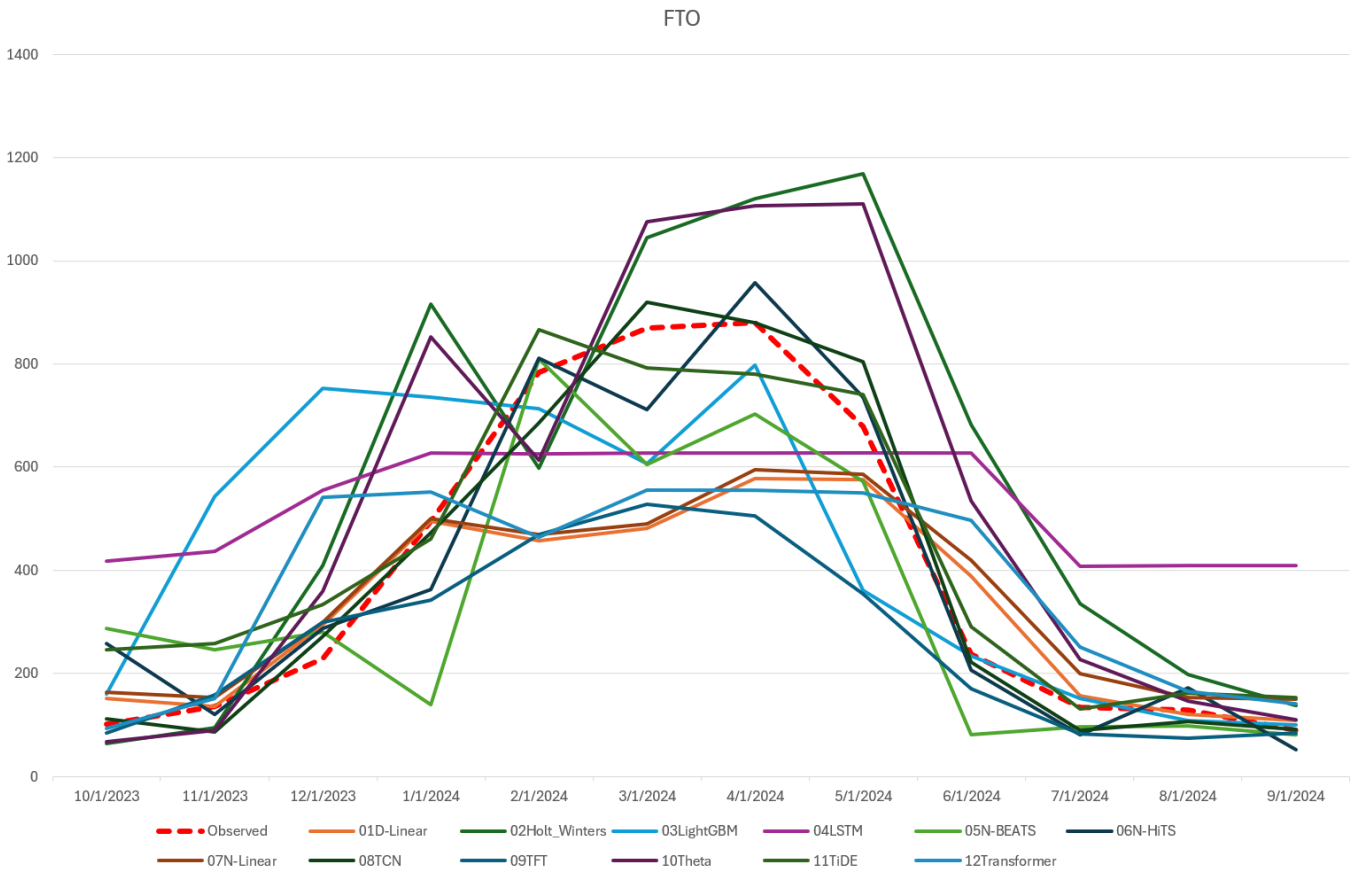
Transformer Model **Transformer** (2017)

Temporal Fusion Transformer **TFT** (2021)

WY2024

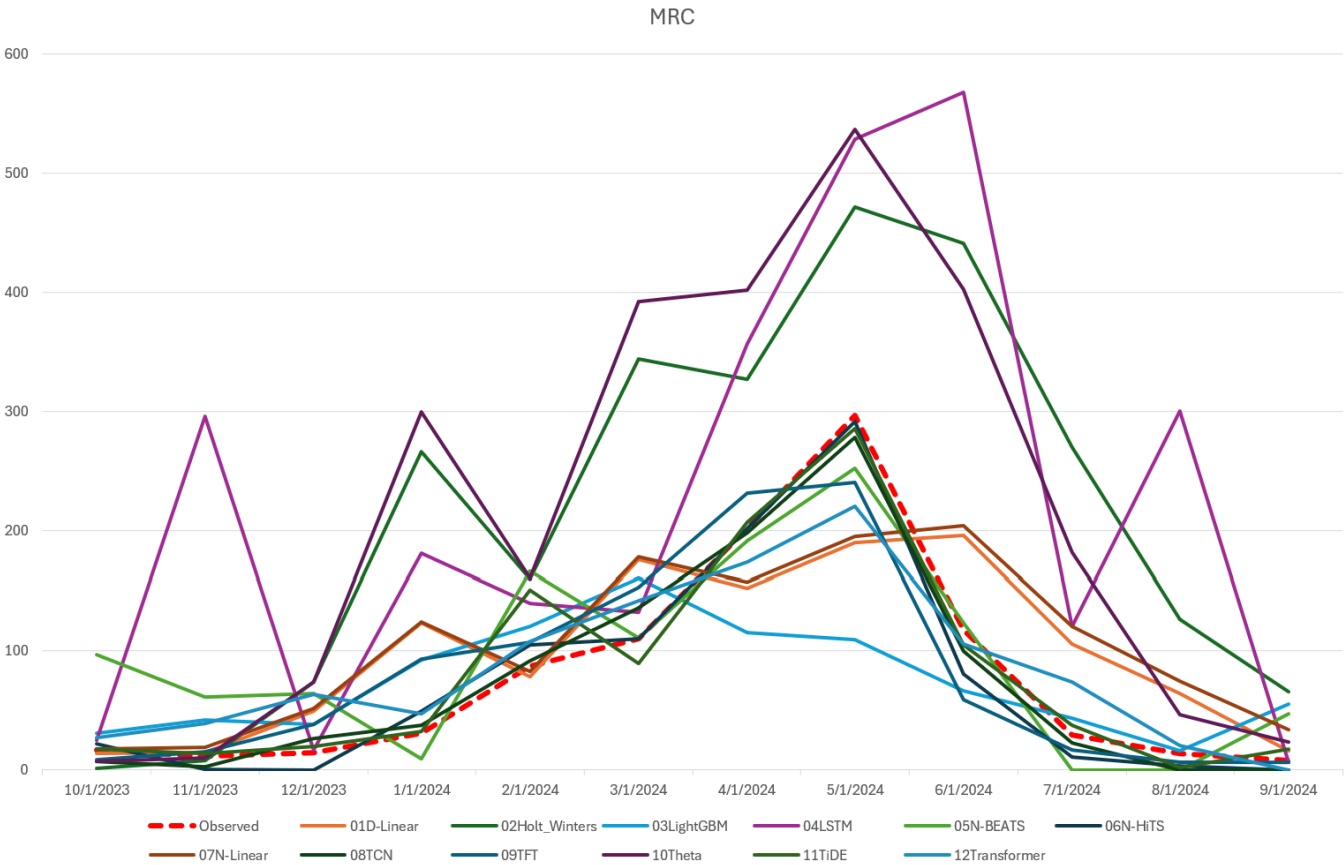
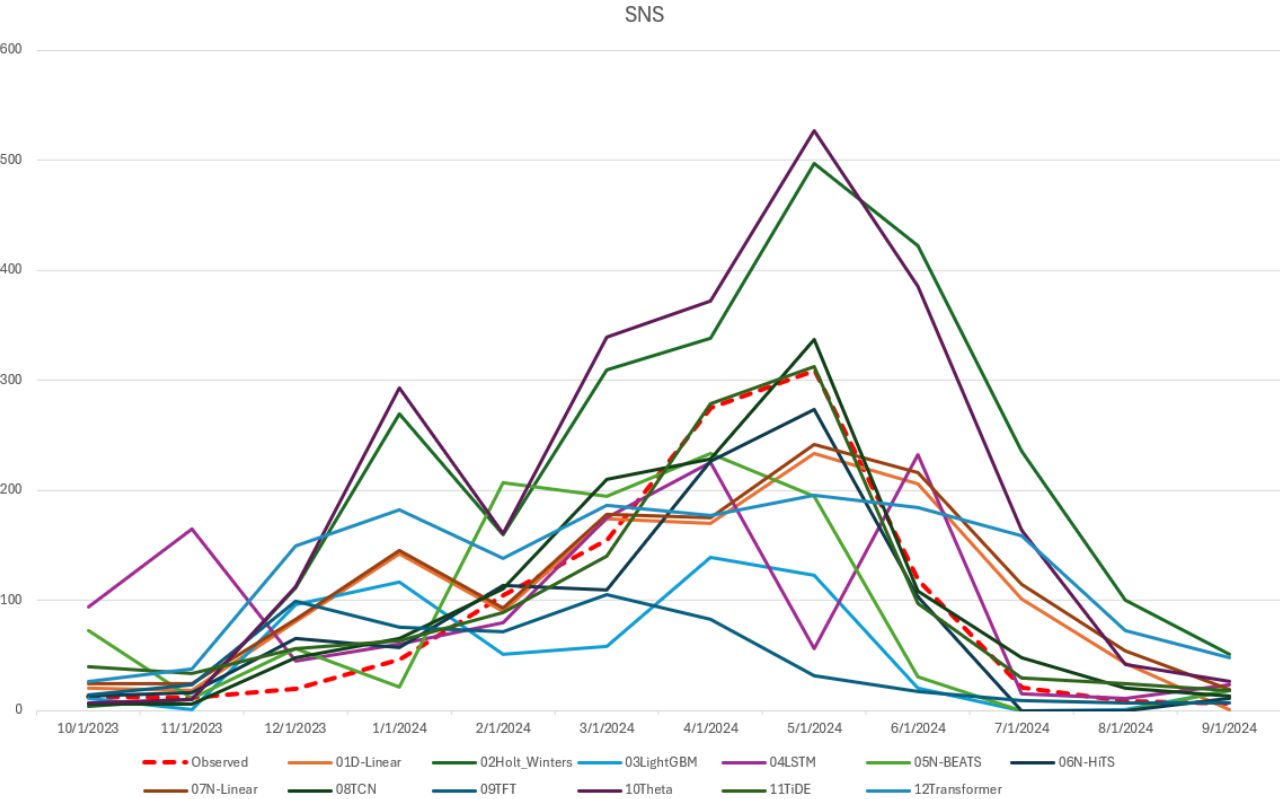


SVI

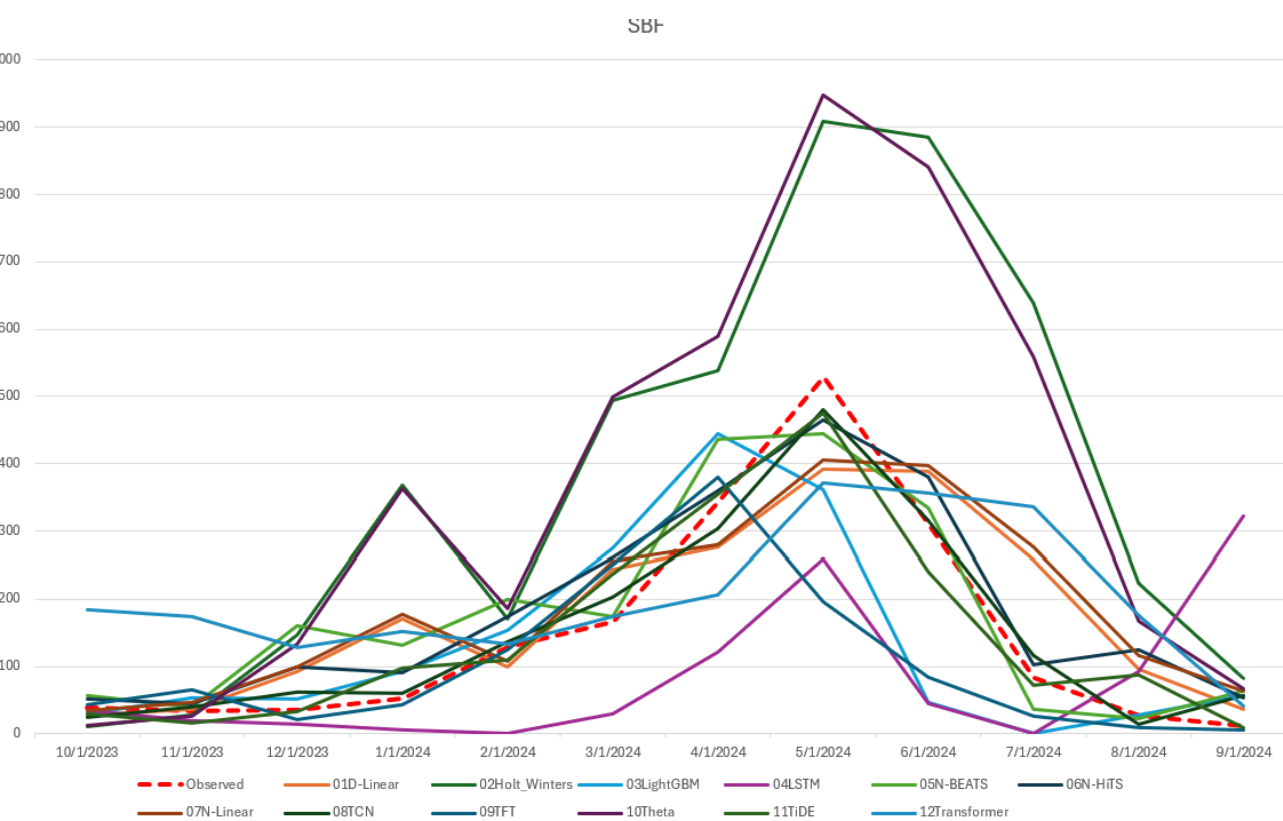
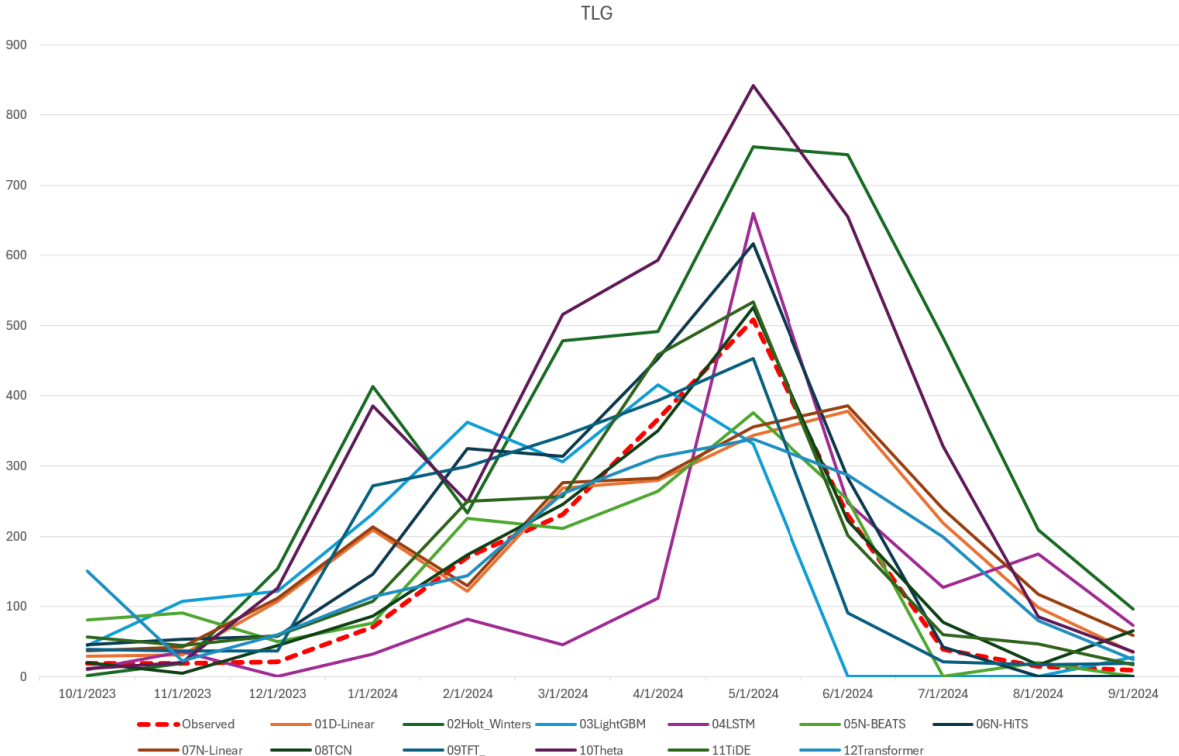


All Monthly flows in TAF

WY2024 (cont.)



SJVI



All Monthly flows in TAF

What Model(s) to Choose?

QUANTIFICATION OF ACCURACY IN WATERSHED SIMULATIONS

D. N. Moriasi, J. G. Arnold, M. W. Van Liew, R. L. Bingner, R. D. Harmel, T. L. Veith

ABSTRACT. Watershed models are powerful tools for simulating the effect of watershed processes and management on soil and water resources. However, no comprehensive guidance is available to facilitate model evaluation in terms of the accuracy of simulated data compared to measured flow and constituent values. Thus, the objectives of this research were to: (1) determine recommended model evaluation techniques (statistical and graphical), (2) review reported ranges of values and corresponding performance ratings for the recommended statistics, and (3) establish guidelines for model evaluation based on the review results and project-specific considerations; all of these objectives focus on simulation of streamflow and transport of sediment and nutrients. These objectives were achieved with a thorough review of relevant literature on model application and recommended model evaluation methods. Based on this analysis, we recommend that three quantitative statistics, Nash-Sutcliffe efficiency (NSE), percent bias (PBIAS), and ratio of the root mean square error to the standard deviation of measured data (RSR), in addition to the graphical techniques, be used in model evaluation. The following model evaluation performance ratings were established for each recommended statistic. In general, model simulation can be judged as satisfactory if $NSE > 0.50$ and $RSR \leq 0.70$, and if $PBIAS \pm 25\%$ for streamflow, $PBIAS \pm 55\%$ for sediment, and $PBIAS \pm 70\%$ for N and P. For PBIAS, constituent-specific performance ratings were determined based on uncertainty of measured data. Additional considerations related to model evaluation guidelines are also discussed. These considerations include: single-event simulation, quality and quantity of measured data, model calibration procedure, evaluation time step, and project scope and magnitude. A case study illustrating the application of the model evaluation guidelines is also provided.

Keywords. Accuracy, Model calibration and validation, Simulation, Watershed model.

Computer-based watershed models can save time and money because of their ability to perform long-term simulation of the effects of watershed processes and management activities on water quality, water quantity, and soil quality. These models also facilitate the simulation of various conservation program effects and aid policy design to mitigate water and soil quality degradation by determining suitable conservation programs for particular watersheds and agronomic settings. In order to use model outputs for tasks ranging from regulation to research, models should be scientifically sound, robust, and defensible (U.S. EPA, 2002).

Sensitivity analysis is the process of determining the rate of change in model output with respect to changes in model inputs (parameters). It is a necessary process to identify key parameters and parameter precision required for calibration (Ma et al., 2000). Model calibration is the process of estimating model parameters by comparing model predictions (output) for a given set of assumed conditions with observed data for the same conditions. Model validation involves running a model using input parameters measured or determined during the calibration process. According to Refsgaard (1997), model validation is the process of demonstrating that a given site-specific model is capable of making “sufficiently accurate” simulations, although “sufficiently accurate” can vary based on project goals. According to the U.S. EPA (2002), the process used to accept, reject, or qualify model results should be established and documented before beginning model evaluation. Although ASCE (1993) emphasized the need to clearly define model evaluation criteria, no commonly accepted guidance has been established, but specific statistics and performance ratings for their use have been developed and used for model evaluation (Donigan et al., 1983; Ramanarayanan et al., 1997; Gupta et al., 1999; Motovilov et al., 1999; Saleh et al., 2000; Santhi et al., 2001; Singh et al., 2004; Bracmort et al., 2006; Van Liew et al., 2007). However, these performance ratings are model and project specific. Standardized guidelines are needed to establish a common system for judging model performance and comparing various models (ASCE, 1993). Once established, these guidelines will assist modelers in preparing and reviewing quality assurance project plans for modeling (U.S. EPA, 2002) and will increase accountability and public acceptance of models to support

Submitted for review in May 2006 as manuscript number SW 6494; approved for publication by the Soil & Water Division of ASABE in March 2007.

The authors are **Daniel N. Moriasi, ASABE Member Engineer**, Hydrologist, USDA-ARS Grazinglands Research Laboratory, El Reno, Oklahoma; **Jeffrey G. Arnold**, Supervisory Agricultural Engineer, USDA-ARS Grassland Soil and Water Research Laboratory, Temple, Texas; **Michael W. Van Liew, ASABE Member Engineer**, Environmental Science Specialist, Water Quality Planning Bureau, Montana Department of Environmental Quality, Helena, Montana; **Ronald L. Bingner, ASABE Member Engineer**, Agricultural Engineer, USDA-ARS Watershed Physical Processes Research Unit, Oxford, Mississippi; **R. Daren Harmel, ASABE Member Engineer**, Agricultural Engineer, USDA-ARS Grassland Soil and Water Research Laboratory, Temple, Texas; and **Tamie L. Veith, ASABE Member Engineer**, Agricultural Engineer, USDA-ARS Pasture Systems and Watershed Management Research Unit, University Park, Pennsylvania. **Corresponding author:** Daniel N. Moriasi, USDA-ARS Grazinglands Research Laboratory, 7207 W. Cheyenne St., El Reno, OK 73036-0000; phone: 405-262-5291, ext. 263; fax: 405-262-0133; e-mail: dmoriasi@spa.ars.usda.gov.

Transactions of the ASABE

Vol. 50(3): 885–900

2007 American Society of Agricultural and Biological Engineers ISSN 0001-2351

885

RSR: RMSE-observations Standard Deviation Ratio

$$RSR = \frac{RMSE}{STDEV_{obs}} = \frac{\sqrt{\sum_{i=1}^n (Y_i^{obs} - Y_i^{sim})^2}}{\sqrt{\sum_{i=1}^n (Y_i^{obs} - Y^{mean})^2}}$$

Percent Bias (PBIAS)

$$PBIAS = \frac{\sum_{i=1}^n (Y_i^{obs} - Y_i^{sim}) * (100)}{\sum_{i=1}^n (Y_i^{obs})}$$

Nash_Sutcliffe Efficiency (NSE)

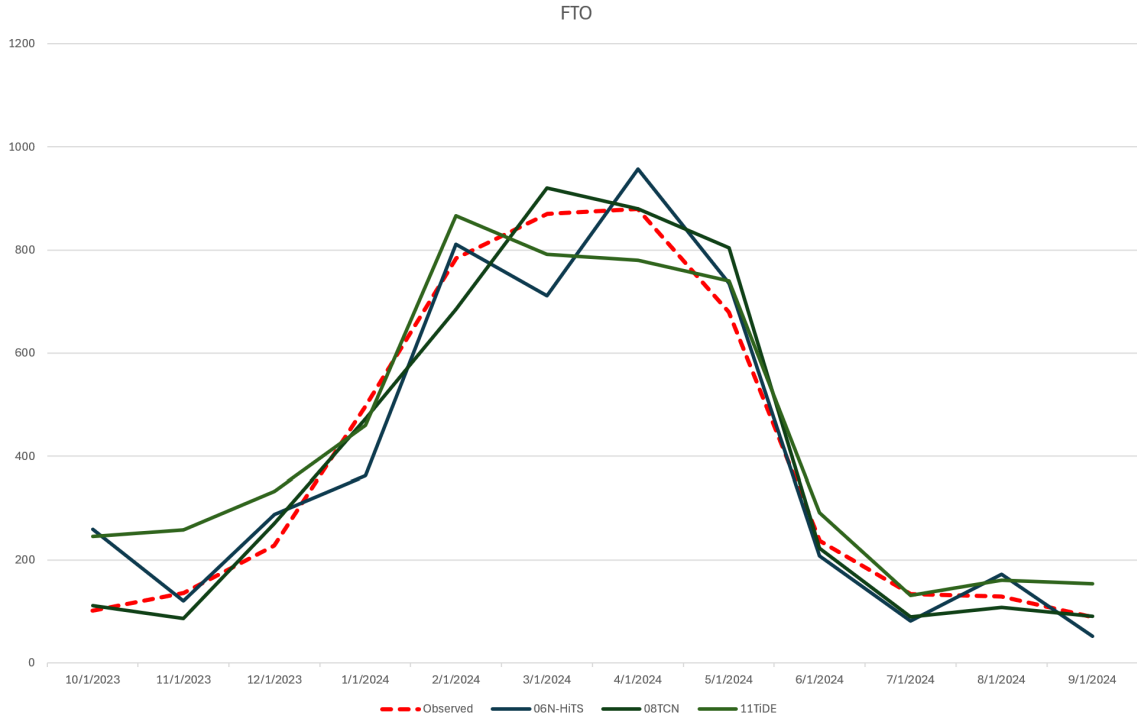
$$NSE = 1 - \frac{\sum_{i=1}^n (Y_i^{obs} - Y_i^{sim})^2}{\sum_{i=1}^n (Y_i^{obs} - Y^{mean})^2}$$

Table 4. General performance ratings for recommended statistics for a monthly time step.

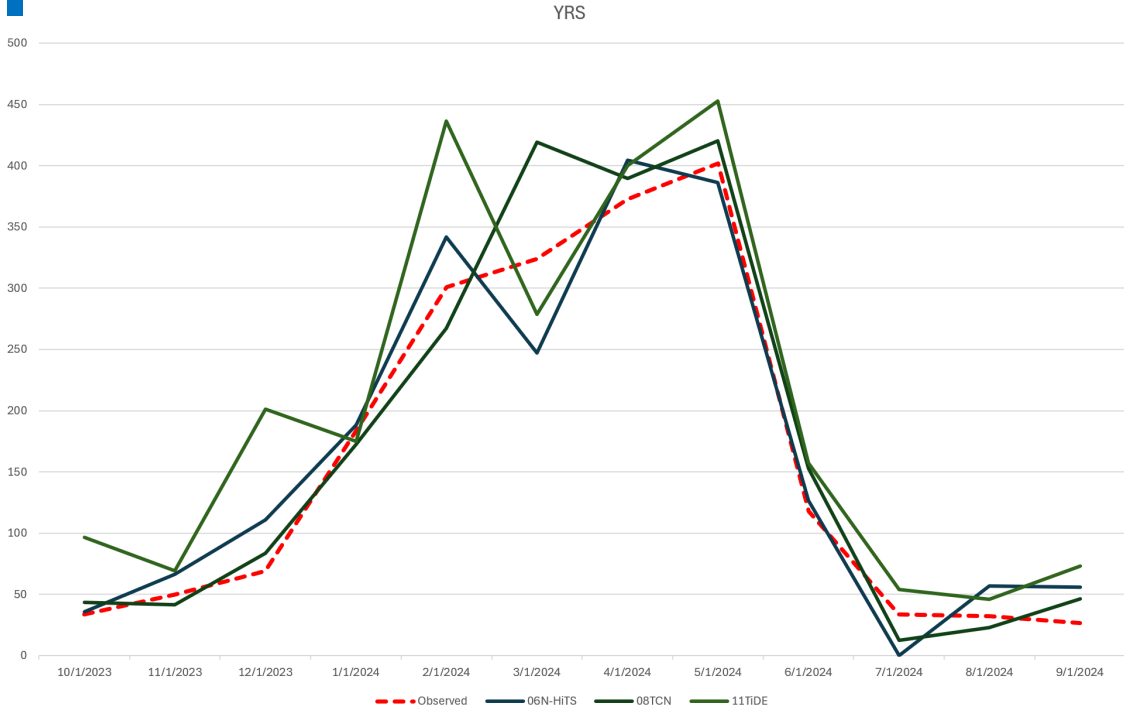
Performance Rating	RSR	NSE	PBIAS (%)		
			Streamflow	Sediment	N, P
Very good	$0.00 \leq RSR \leq 0.50$	$0.75 < NSE \leq 1.00$	$PBIAS < \pm 10$	$PBIAS < \pm 15$	$PBIAS < \pm 25$
Good	$0.50 < RSR \leq 0.60$	$0.65 < NSE \leq 0.75$	$\pm 10 \leq PBIAS < \pm 15$	$\pm 15 \leq PBIAS < \pm 30$	$\pm 25 \leq PBIAS < \pm 40$
Satisfactory	$0.60 < RSR \leq 0.70$	$0.50 < NSE \leq 0.65$	$\pm 15 \leq PBIAS < \pm 25$	$\pm 30 \leq PBIAS < \pm 55$	$\pm 40 \leq PBIAS < \pm 70$
Unsatisfactory	$RSR > 0.70$	$NSE \leq 0.50$	$PBIAS \geq \pm 25$	$PBIAS \geq \pm 55$	$PBIAS \geq \pm 70$

Model Assessment Based on Moriasi Table (WY2024)

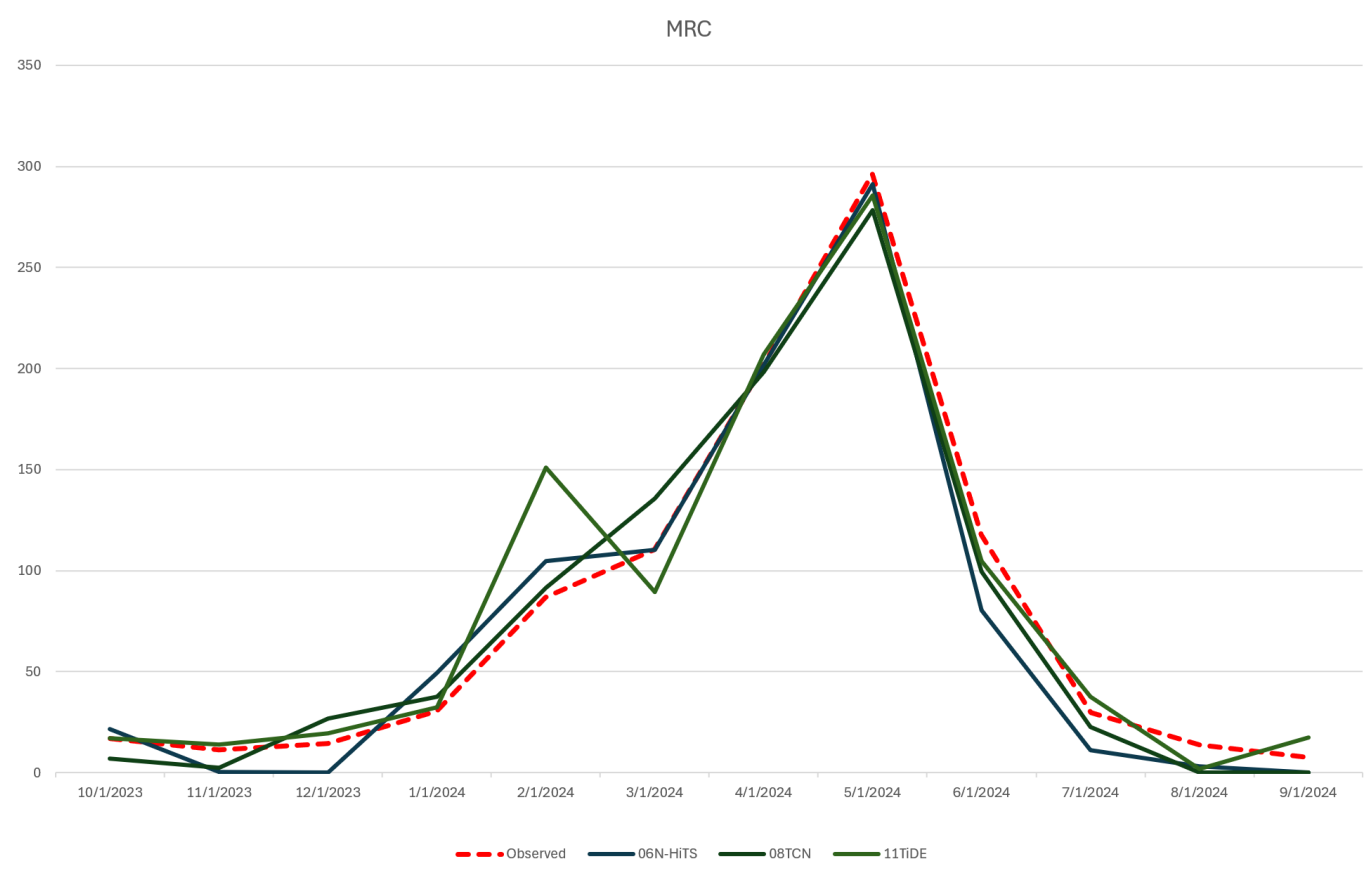
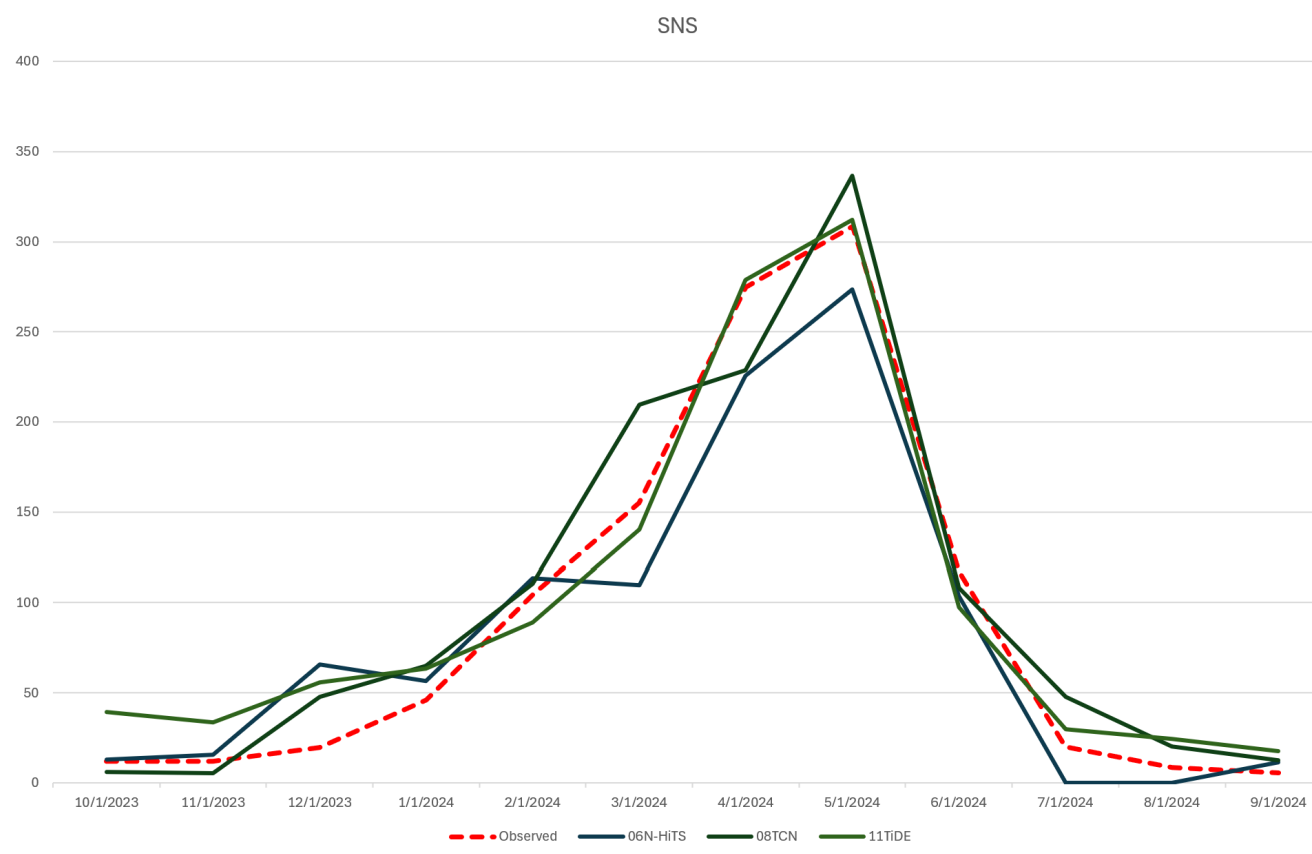
	01D-Linear	02Holt_Winters	03LightGBM	04LSTM	05N-BEATS	06N-HITS	07N-Linear	08TCN	09TFT	10Theta	11TIDE	12Transformer
SBB	Unsatisfactory		Very Good	Unsatisfactory	Very Good	Very Good	Unsatisfactory	Very Good			Very Good	
FTO		Unsatisfactory		Unsatisfactory		Very Good	Good	Very Good			Very Good	Satisfactory
AMF	Good	Unsatisfactory		Unsatisfactory	Very Good	Very Good		Very Good	Unsatisfactory	Unsatisfactory	Very Good	Unsatisfactory
YRS	Very Good	Unsatisfactory	Unsatisfactory	Unsatisfactory		Very Good	Good	Very Good		Unsatisfactory		Unsatisfactory
SNS	Satisfactory	Unsatisfactory	Unsatisfactory			Very Good			Unsatisfactory	Unsatisfactory	Very Good	Unsatisfactory
TLG	Satisfactory	Unsatisfactory			Very Good			Very Good		Unsatisfactory		
MRC		Unsatisfactory		Unsatisfactory		Very Good	Unsatisfactory	Very Good	Very Good	Unsatisfactory	Very Good	Very Good
SBF		Unsatisfactory	Satisfactory	Unsatisfactory				Very Good	Unsatisfactory	Unsatisfactory	Very Good	Unsatisfactory



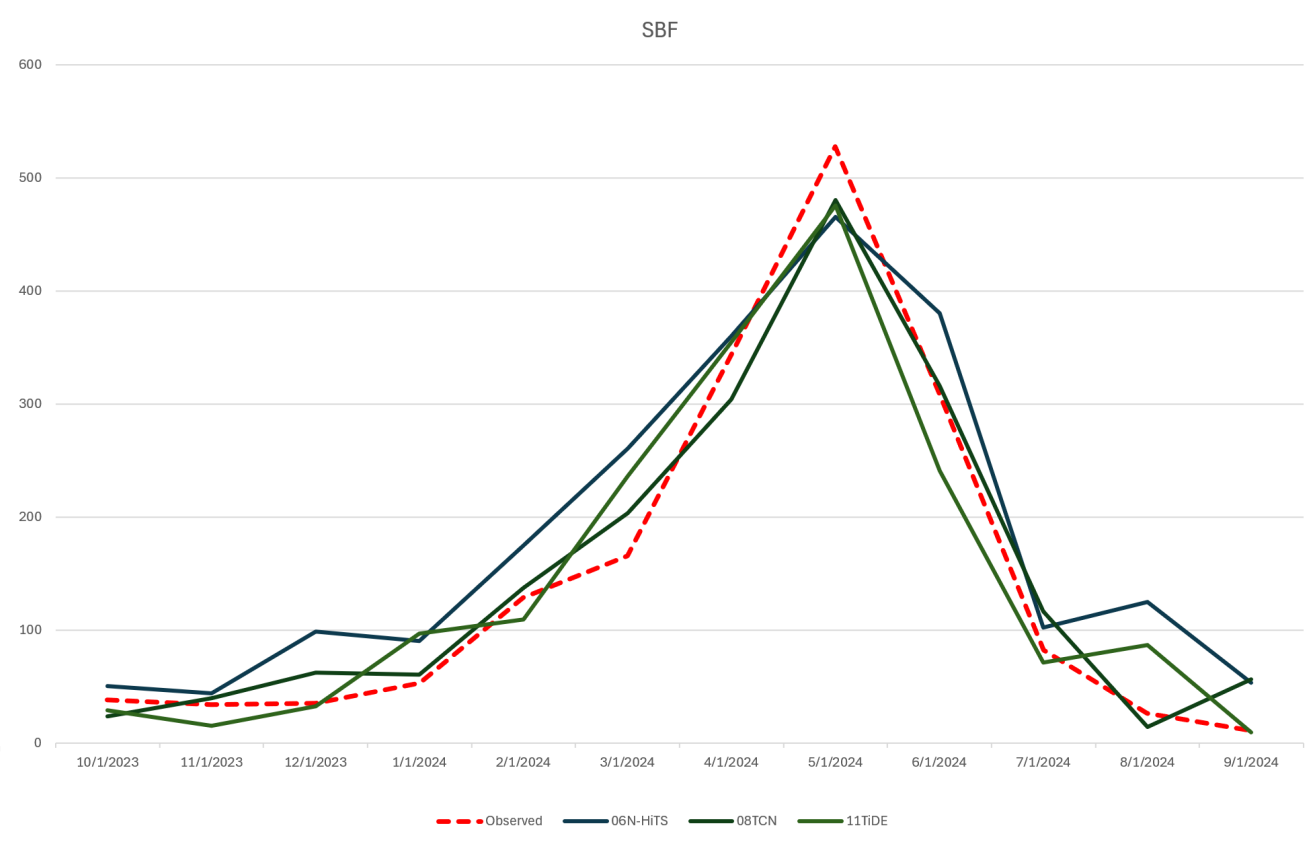
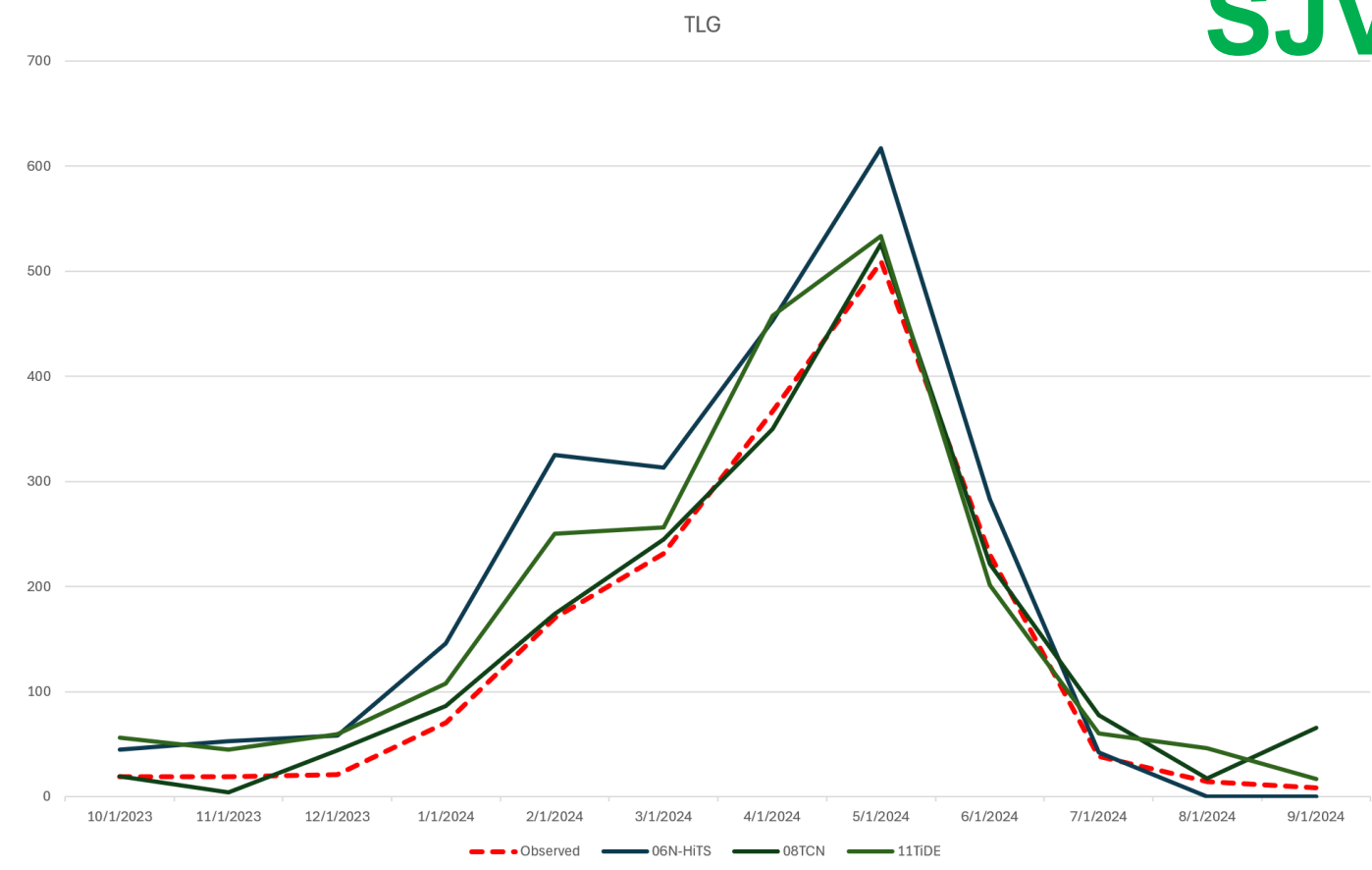
SVI



All Monthly flows in TAF



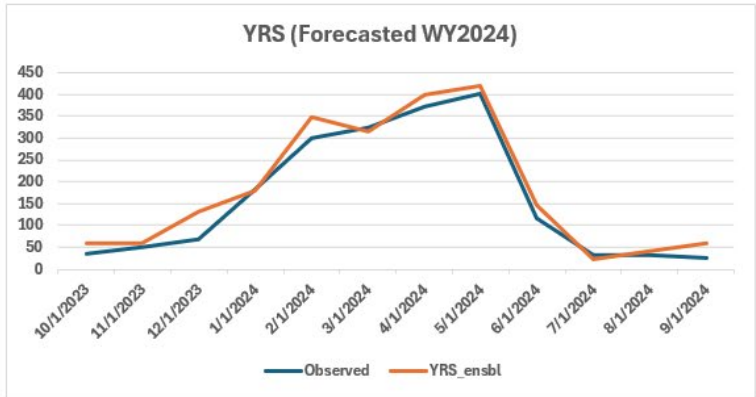
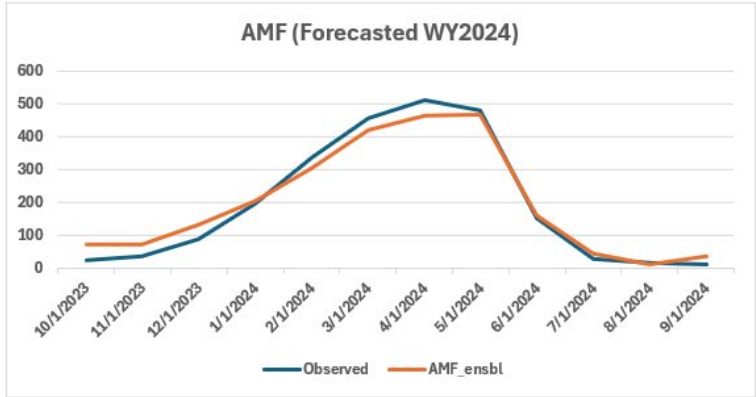
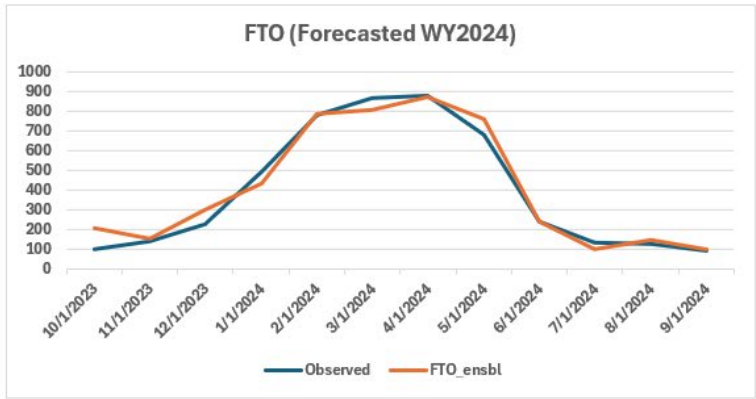
SJVI



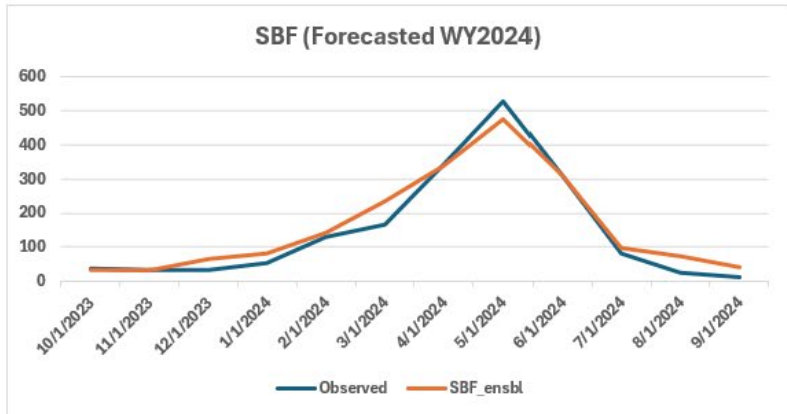
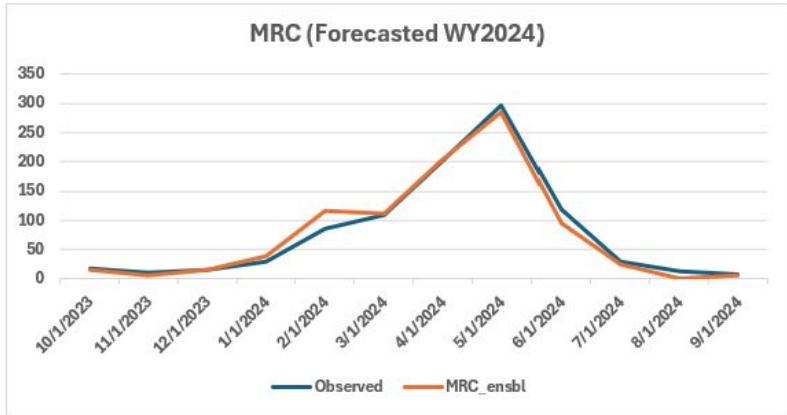
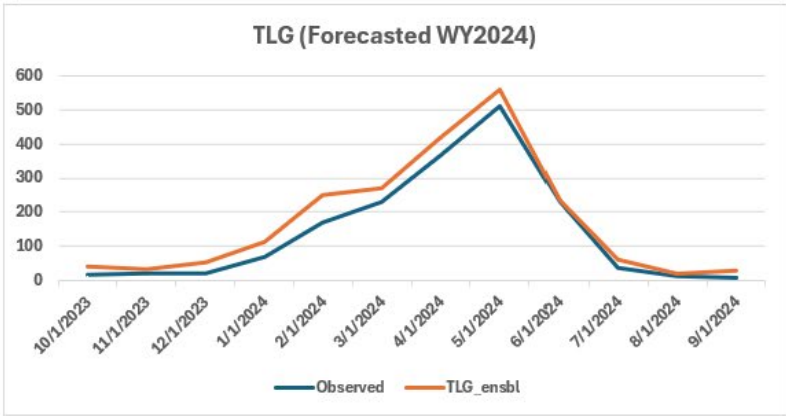
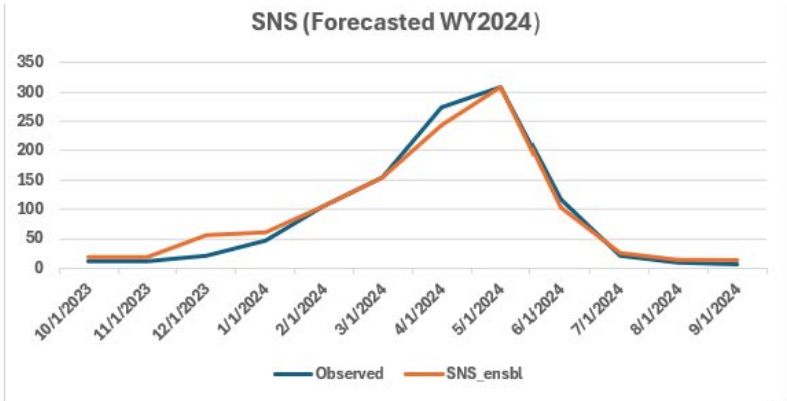
All Monthly flows in TAF

Choose Ensemble : Average of N-HiTS, TCN & TiDE

SVI

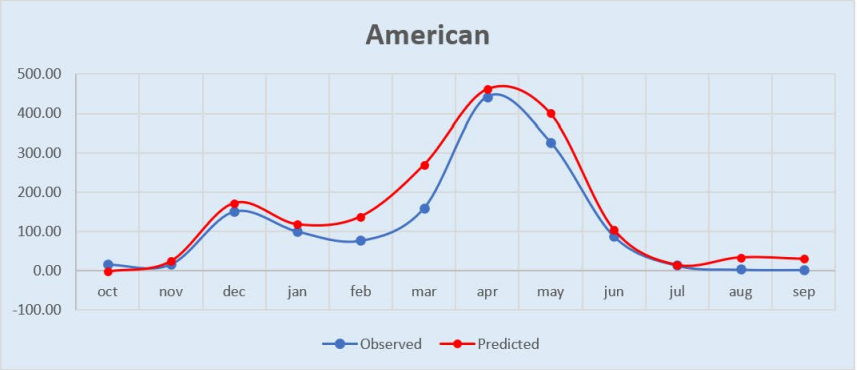
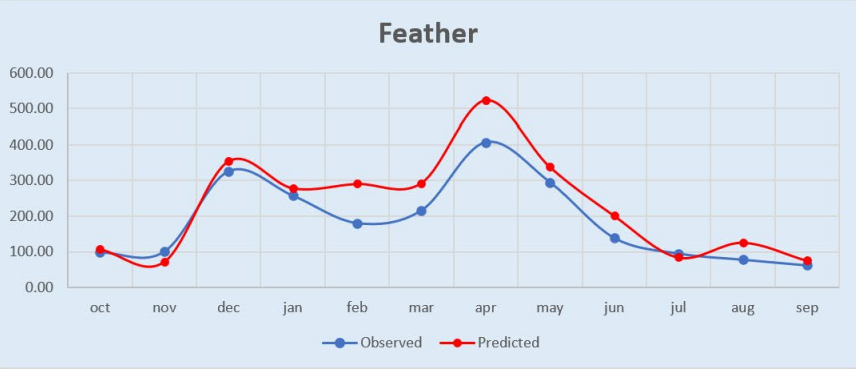
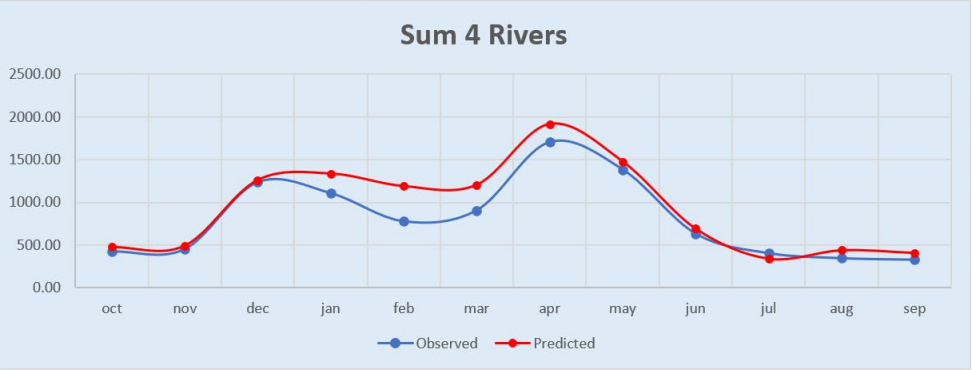
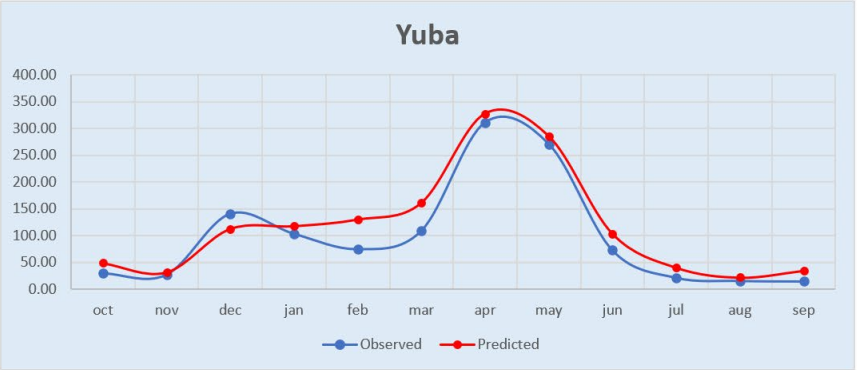
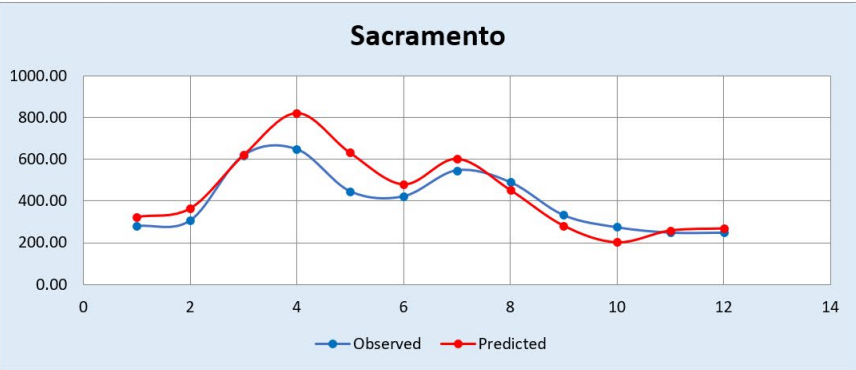


SJVI



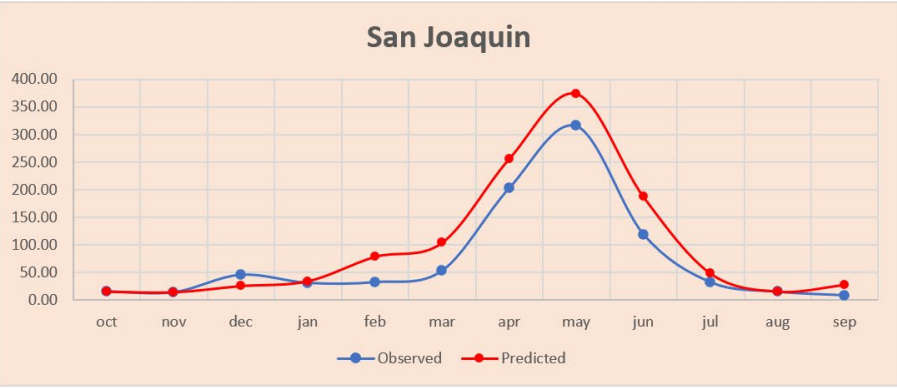
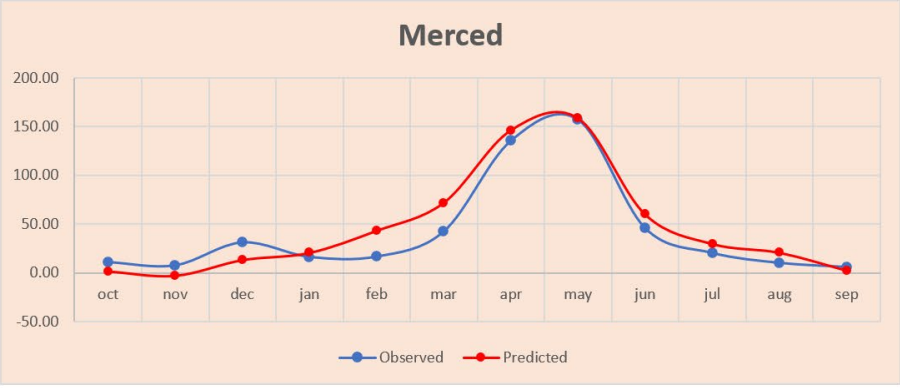
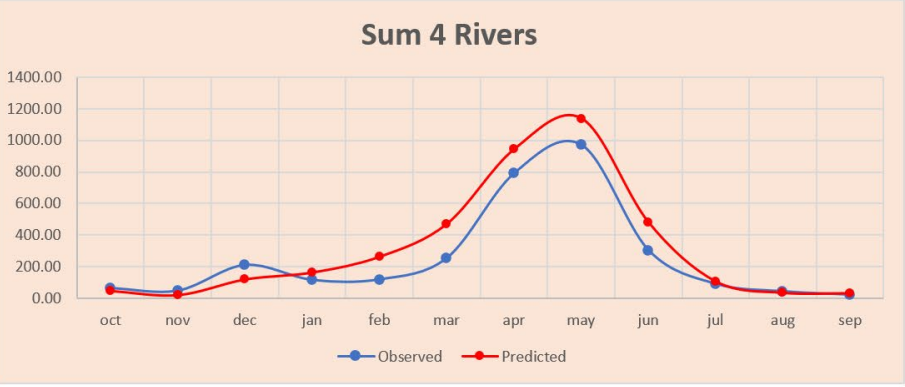
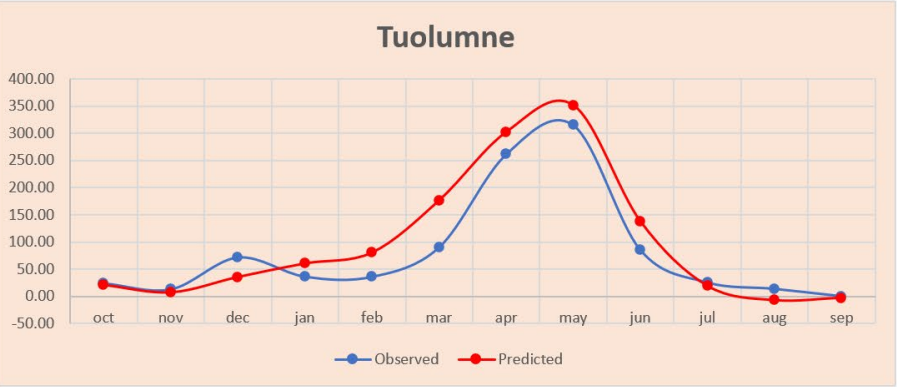
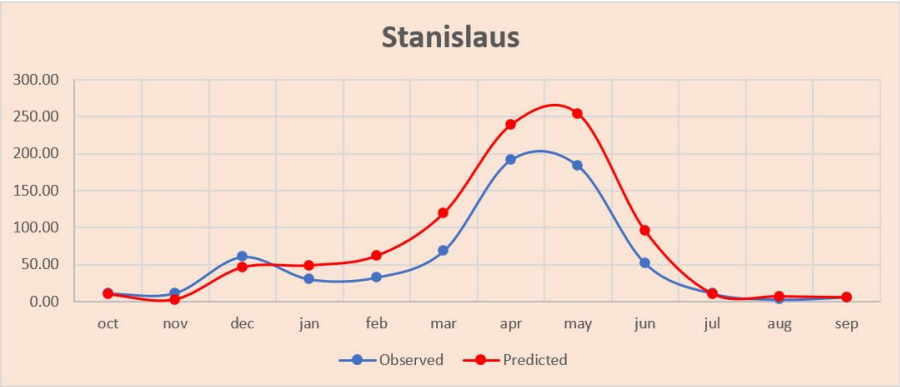
All Monthly flows in TAF

WY2020 using Ensemble



SVI

mon	Observed	Predicted
Calculated Index	6.12	6.56
Calculated WY Class	D	BN



SJVI

mon	Observed	Predicted
Calculated Index	2.35	2.72
Calculated WY Class	D	BN

**Monthly flows in TAF
Index in MAF**

Summary of "Ensemble" October 1st Forecasts of Water Year Type

SVI		
WY	Actual	Forecasted
2020	<i>D</i>	<i>BN</i>
2021	<i>C</i>	<i>C</i>
2022	<i>C</i>	<i>C</i>
2023	<i>W</i>	<i>AN</i>
2024	<i>AN</i>	<i>AN</i>
2025	<i>AN</i>	<i>AN</i>

SJVI		
WY	Actual	Forecasted
2020	<i>D</i>	<i>BN</i>
2021	<i>C</i>	<i>D</i>
2021	<i>C</i>	<i>D</i>
2023	<i>W</i>	<i>W</i>
2024	<i>AN</i>	<i>AN</i>
2025	<i>AN</i>	<i>AN</i>

More Recent AI Time Series Models

(TK work in progress....)

TSMixer (Google Research 2023)

- High performance, lightweight, and efficient multivariate forecasting model
- Uses MLP architecture that alternates between mixing time and feature dimensions to capture both temporal patterns and cross-variate dependencies

AutoGluon (Amazon Research 2019-2020)

- Open –source AutoML library
- Supports both “global” deep learning models (e.g., Temporal Fusion Transformer, DeepAR) and “local” statistical models (e.g., ARIMA, Error-Trend-Seasonality (ETS))
- Probabilistic forecasting models
- Multiple related time series (results from 10 included in this presentations)

Prophet (Meta “Facebook” 2017)

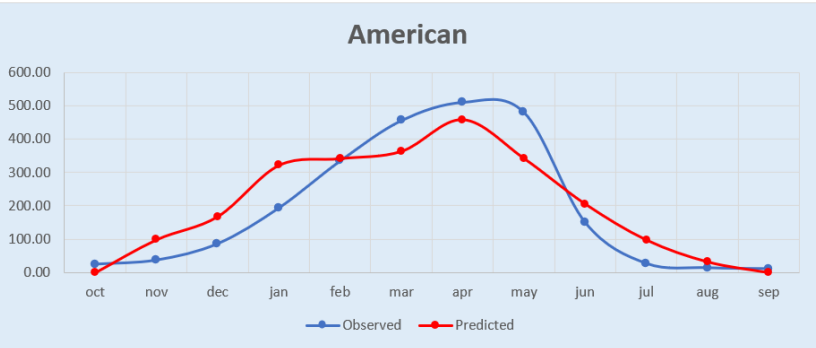
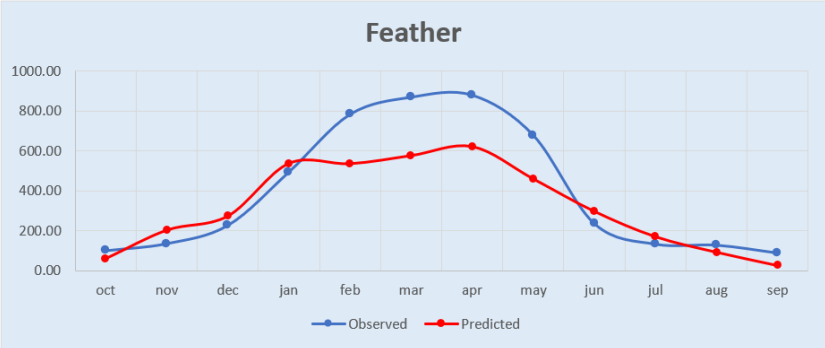
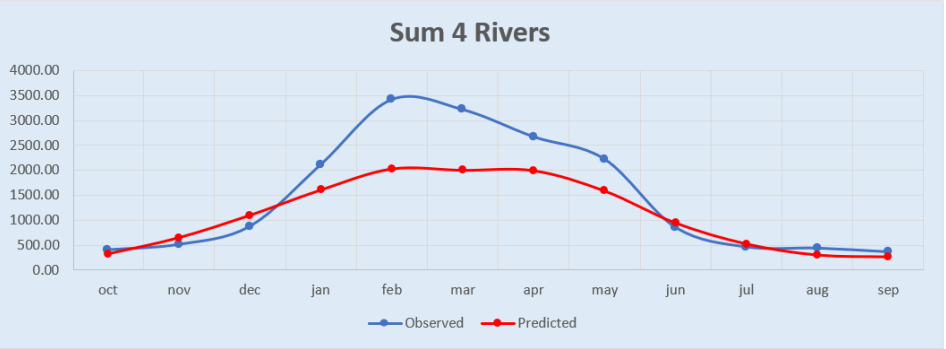
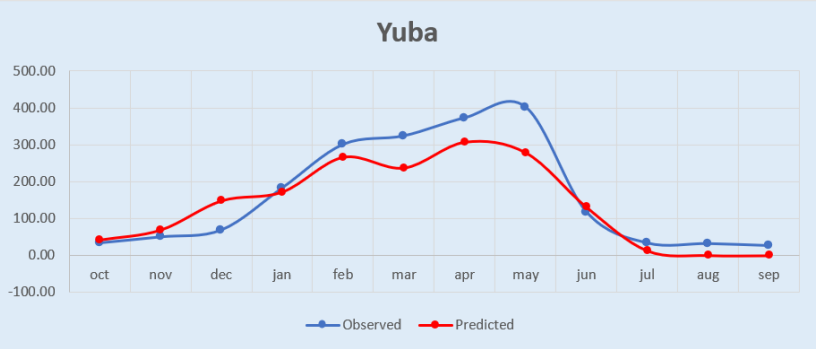
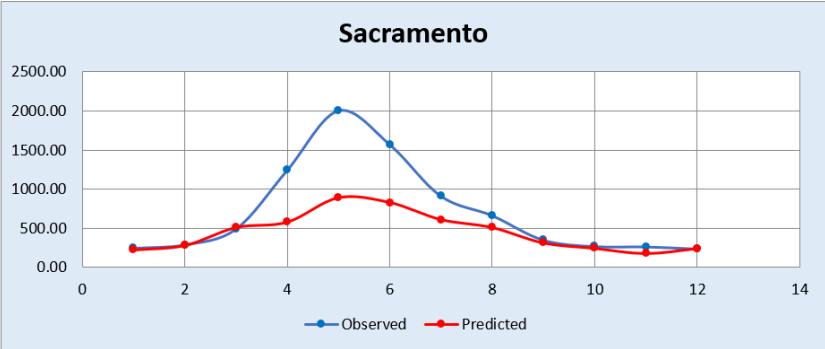
- Additive time-series forecasting model
- Ease of use, speed, and robustness to missing data, outliers, and abrupt trend changes
- Decomposes time series into trend, seasonality (yearly, weekly, daily), and holidays

Silverkite (Linkedin 2021)

- Flexible, fast, open source
- Designed to handle complex, real-world business data with various trends, multiple types of seasonality, holidays, and external regressors
- Greykite Python library

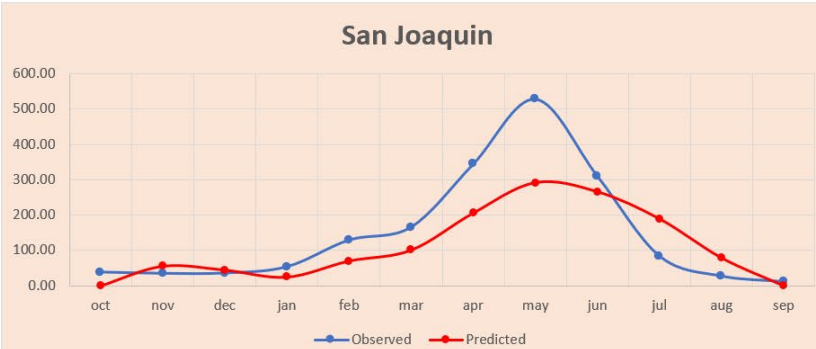
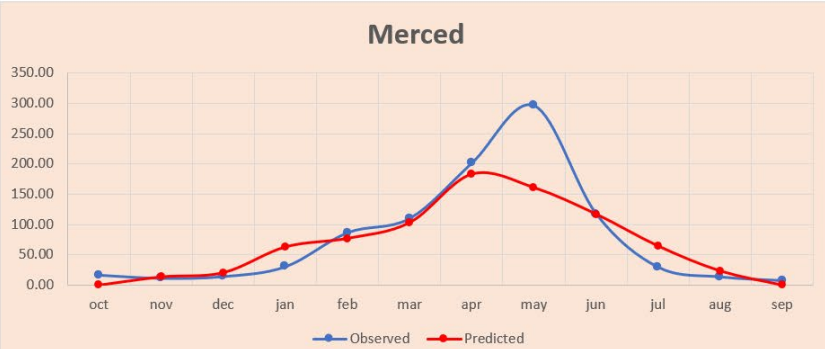
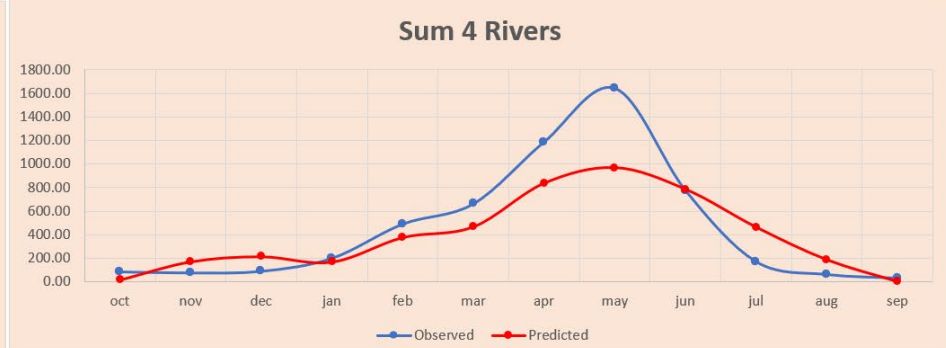
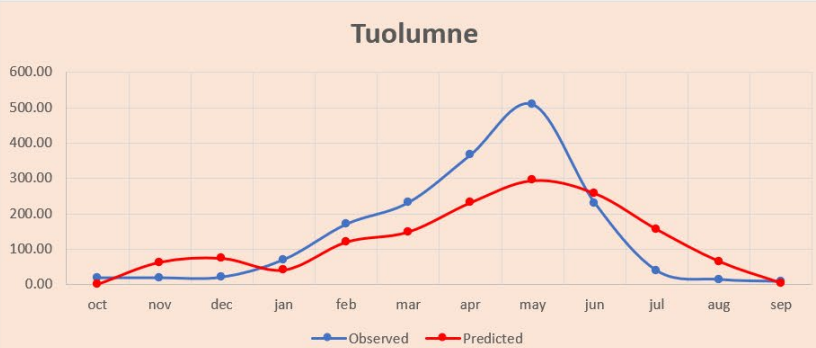
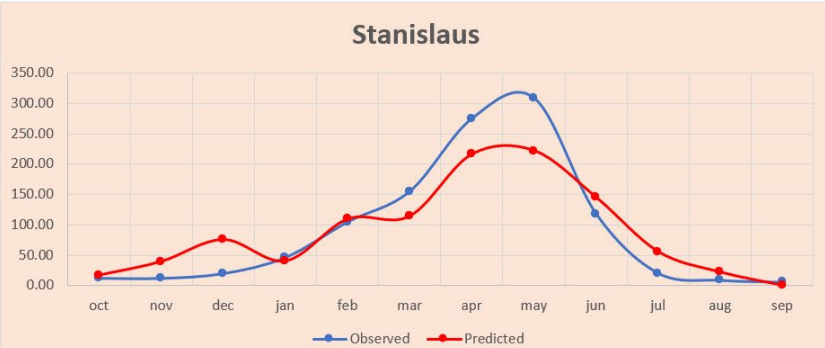
Results for TSMixer

WY2024



SVI

mon	Observed	Predicted
Calculated Index	8.45	7.15
Calculated WY Class	AN	BN

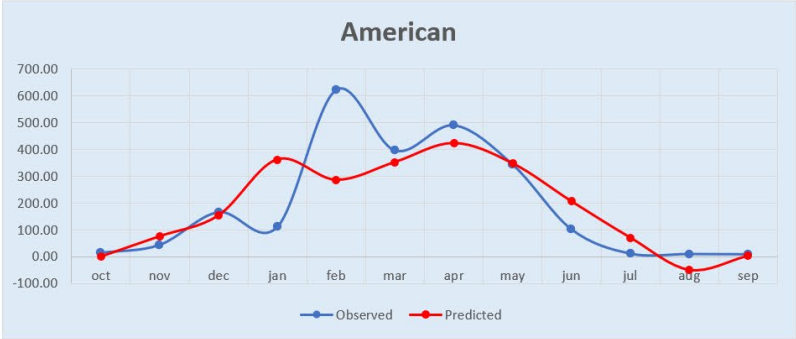
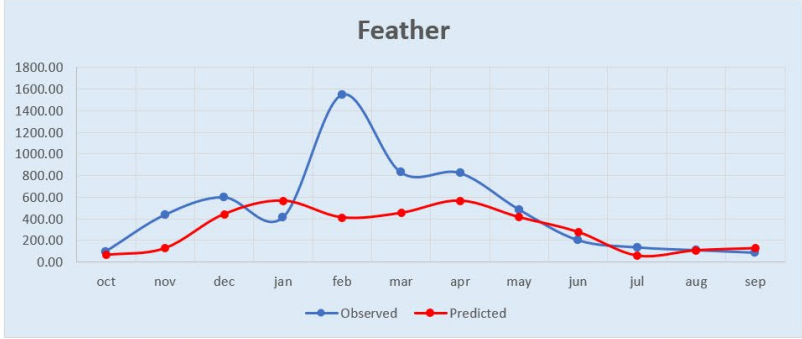
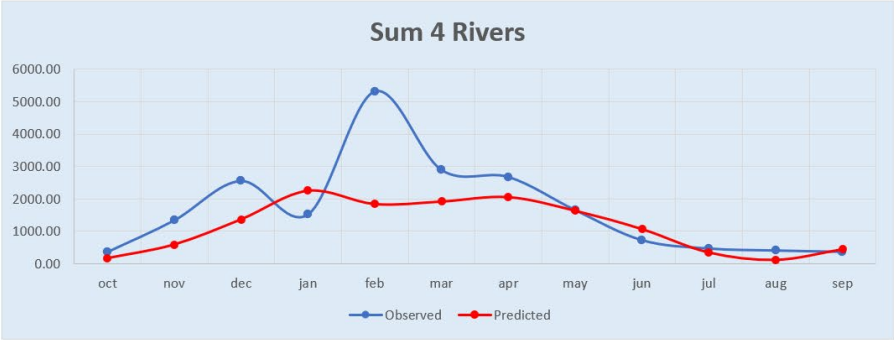
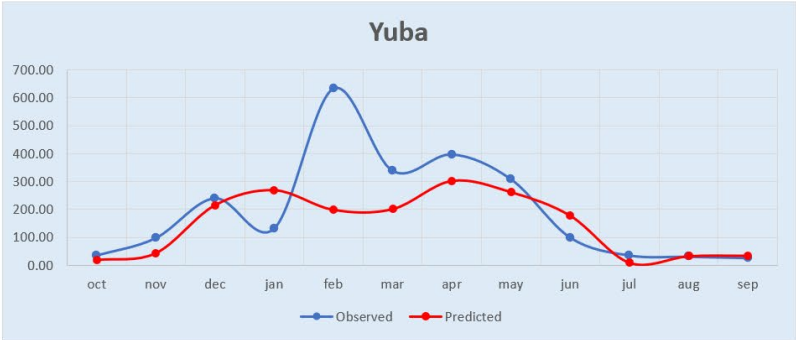
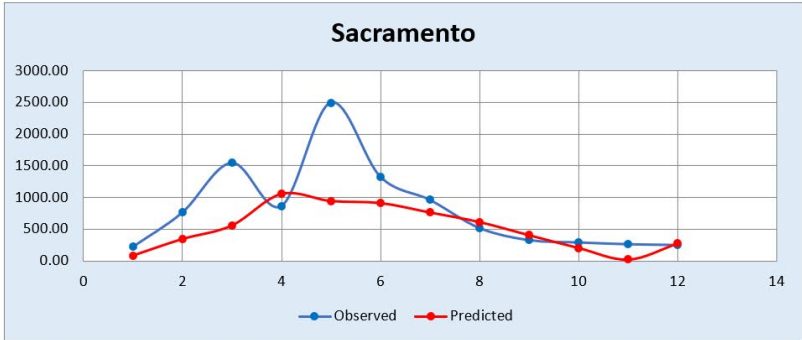


SJVI

mon	Observed	Predicted
Calculated Index	3.49	3.02
Calculated WY Class	AN	BN

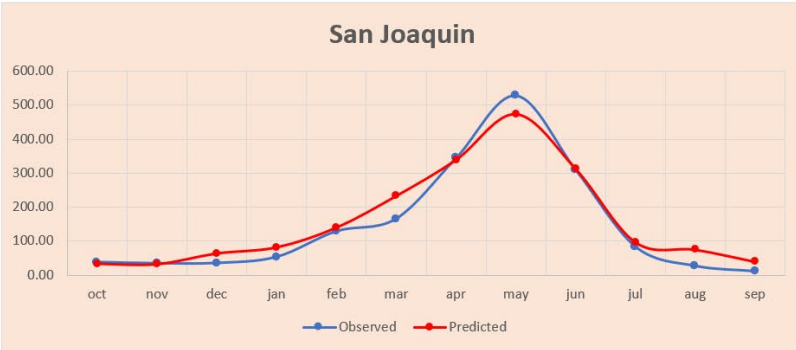
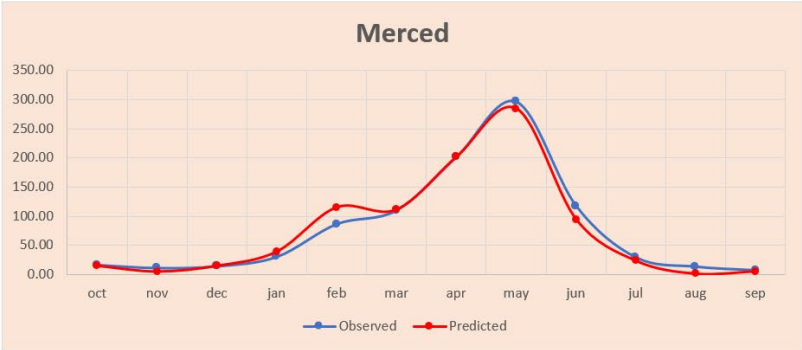
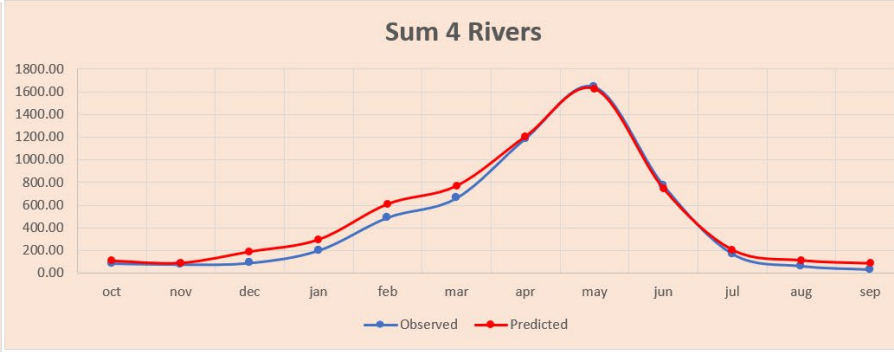
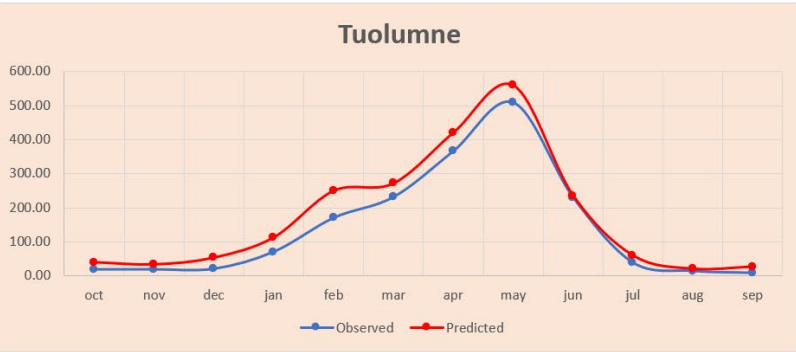
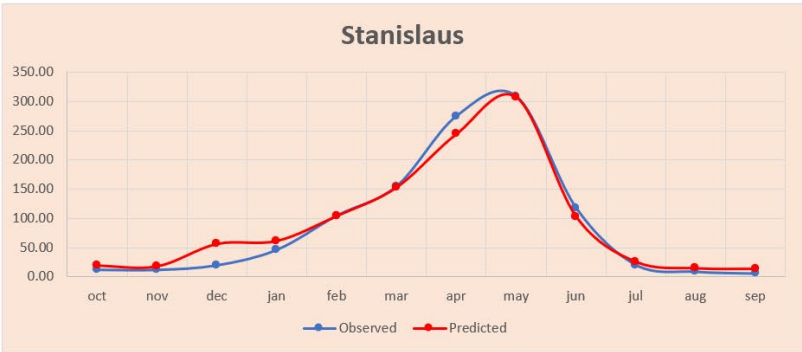
Monthly flows in TAF
Index in MAF

WY2025



SVI

mon	Observed	Predicted
Calculated Index	8.94	7.02
Calculated WY Class	AN	BN



SJVI

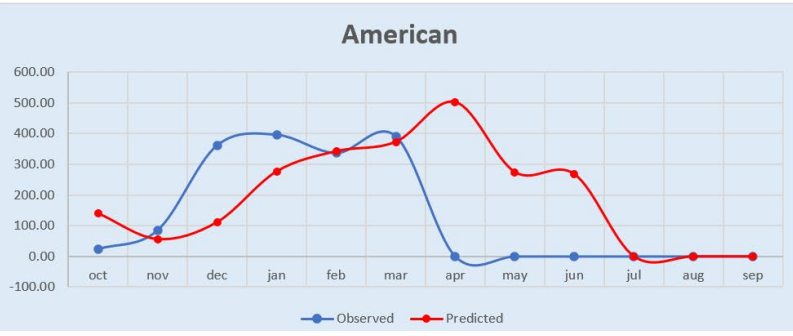
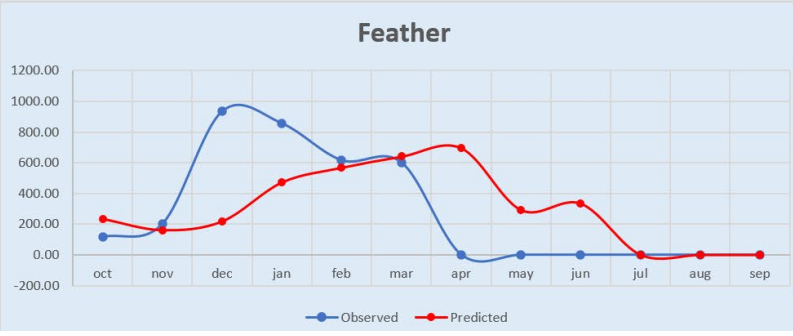
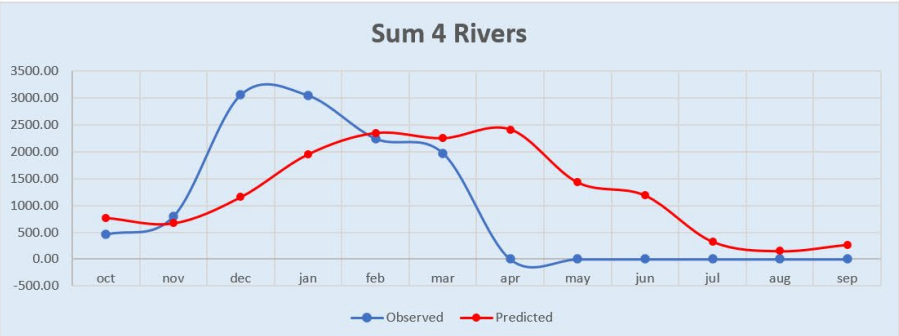
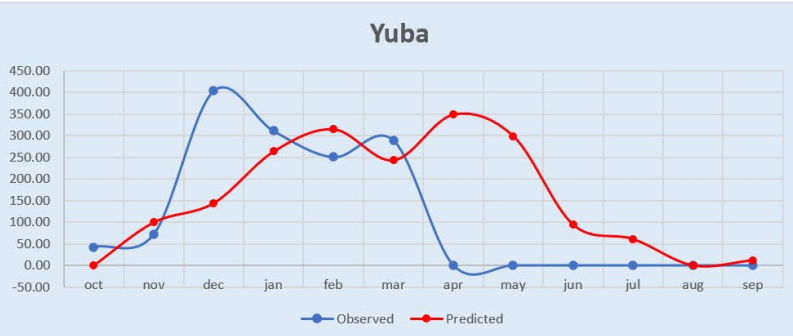
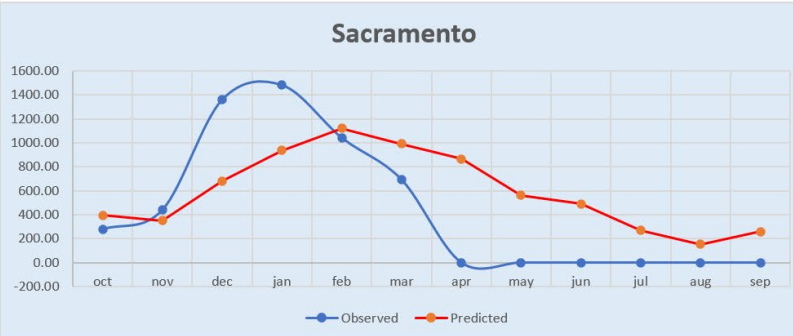
mon	Observed	Predicted
Calculated Index	3.28	3.38
Calculated WY Class	AN	AN

Monthly flows in TAF
Index in MAF

**Note: Observed data for April through Sep, 2026
not available yet!**

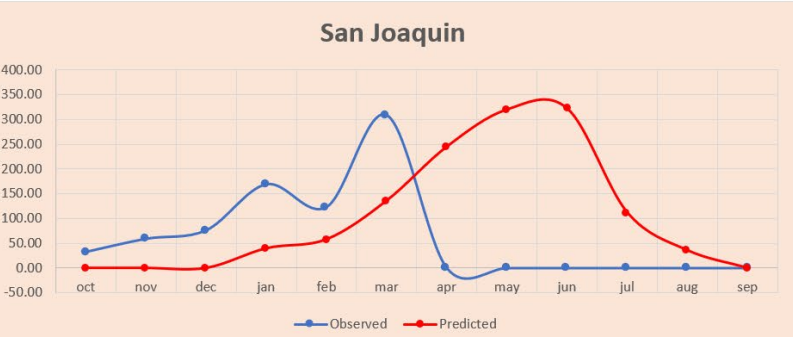
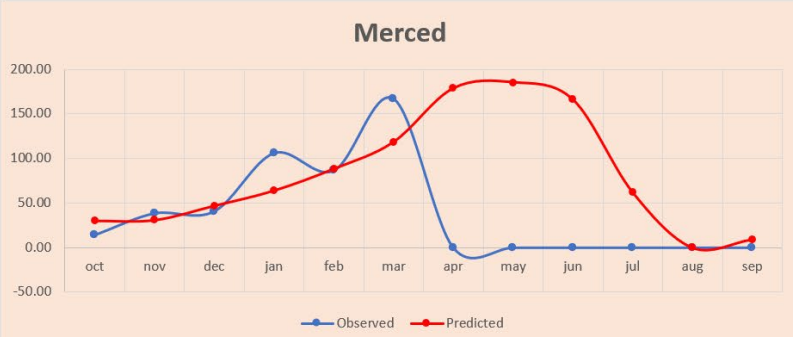
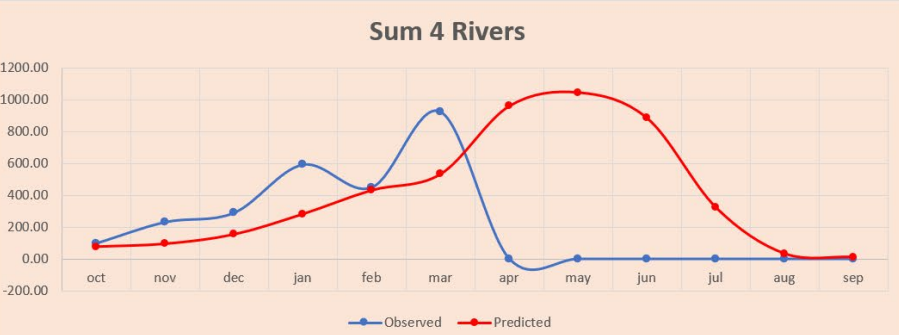
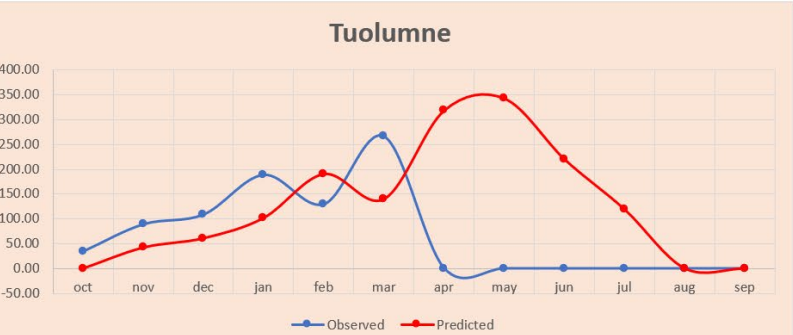
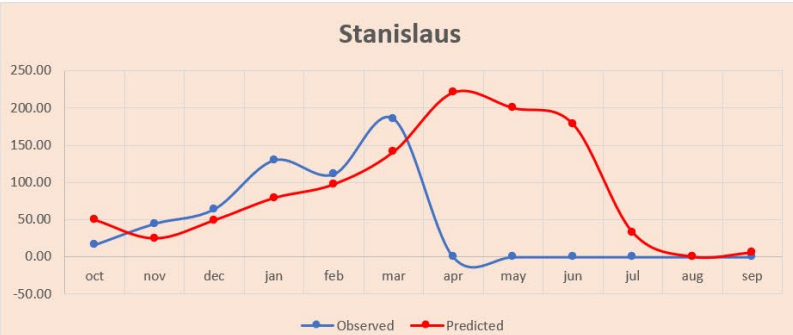
WY2026

Monthly flows in TAF



SVI

**Calculated Predicted Index = 7.57 MAF (BN)
Bul120 April 1st forecast = 7.44 MAF (BN)**



SJVI

**Calculated Predicted Index = 2.76 MAF (BN)
Bul120 April 1st forecast = 2.05 MAF (C)**

Summary of TSMixer October 1st Forecasts of Water Year Type

SVI		
WY	Actual	Forecasted
2020	<i>D</i>	<i>AN</i>
2021	<i>C</i>	<i>BN</i>
2022	<i>C</i>	<i>D</i>
2023	<i>W</i>	<i>D</i>
2024	<i>AN</i>	<i>BN</i>
2025	<i>AN</i>	<i>BN</i>
2026	?	<i>BN</i>

SJVI		
WY	Actual	Forecasted
2020	<i>D</i>	<i>AN</i>
2021	<i>C</i>	<i>AN</i>
2021	<i>C</i>	<i>AN</i>
2023	<i>W</i>	<i>AN</i>
2024	<i>AN</i>	<i>BN</i>
2025	<i>AN</i>	<i>AN</i>
2026	?	<i>BN</i>

Bul120 April 1st 2026 forecast = 7.44 MAF (BN)

Bul120 April 1st 2026 forecast = 2.05 MAF (C)

Results for AutoGluon

WY2020

#	model	score_val	pred_time_val	fit_time_marginal	fit_order
0	WeightedEnsemble	-258.8	0.86974	10.0	11
1	ChronosWithRegressor[bolt_small]	-270.6	0.03270	7.7	8
2	TemporalFusionTransformer	-274.6	0.02100	297.3	9
3	Chronos2	-279.4	0.04432	18.4	5
4	DirectTabular	-291.8	0.05860	5.0	3
5	Chronos2SmallFineTuned	-305.2	0.04228	290.5	6
6	DynamicOptimizedTheta	-306.5	0.06903	3.7	4
7	DeepAR	-311.4	0.06981	152.1	10
8	AutoETS	-314.0	0.68513	13.3	7
9	SeasonalNaive	-348.5	0.02752	3.3	1
10	RecursiveTabular	-359.3	0.07175	7.7	2

WY2021

#	model	score_val	pred_time_val	fit_time_marginal	fit_order
0	WeightedEnsemble	-248.2	0.77867	10.6	11
1	ChronosWithRegressor[bolt_small]	-255.7	0.03194	8.2	8
2	Chronos2	-265.1	0.04307	9.0	5
3	TemporalFusionTransformer	-265.5	0.02132	300.0	9
4	DirectTabular	-275.9	0.06362	7.1	3
5	Chronos2SmallFineTuned	-289.5	0.04353	292.8	6
6	DeepAR	-298.9	0.07253	154.0	10
7	DynamicOptimizedTheta	-299.3	0.07021	3.9	4
8	AutoETS	-303.9	0.61458	16.0	7
9	SeasonalNaive	-338.8	0.02997	3.3	1
10	RecursiveTabular	-347.3	0.06743	8.3	2

WY2022

#	model	score_val	pred_time_val	fit_time_marginal	fit_order
0	WeightedEnsemble	-249.9	0.76521	10.8	11
1	ChronosWithRegressor[bolt_small]	-257.2	0.03253	10.5	8
2	Chronos2	-263.1	0.04245	29.3	5
3	TemporalFusionTransformer	-270.0	0.02111	296.1	9
4	DirectTabular	-279.5	0.06325	5.4	3
5	Chronos2SmallFineTuned	-285.5	0.04653	287.8	6
6	DeepAR	-293.9	0.07076	151.9	10
7	DynamicOptimizedTheta	-295.0	0.07888	5.6	4
8	AutoETS	-299.6	0.57343	14.5	7
9	SeasonalNaive	-330.0	0.02774	3.8	1
10	RecursiveTabular	-353.8	0.06568	6.2	2

WY2023

#	model	score_val	pred_time_val	fit_time_marginal	fit_order
0	WeightedEnsemble	-244.5	0.79978	10.6	11
1	ChronosWithRegressor[bolt_small]	-251.6	0.03039	11.0	8
2	Chronos2	-257.5	0.04202	28.7	5
3	TemporalFusionTransformer	-264.8	0.01822	266.1	9
4	DirectTabular	-273.4	0.06021	5.3	3
5	Chronos2SmallFineTuned	-278.8	0.04128	287.9	6
6	DeepAR	-288.5	0.07090	131.6	10
7	DynamicOptimizedTheta	-289.6	0.09920	5.5	4
8	AutoETS	-294.2	0.61680	14.9	7
9	SeasonalNaive	-322.5	0.02834	3.2	1
10	RecursiveTabular	-349.2	0.06413	6.2	2

WY2024

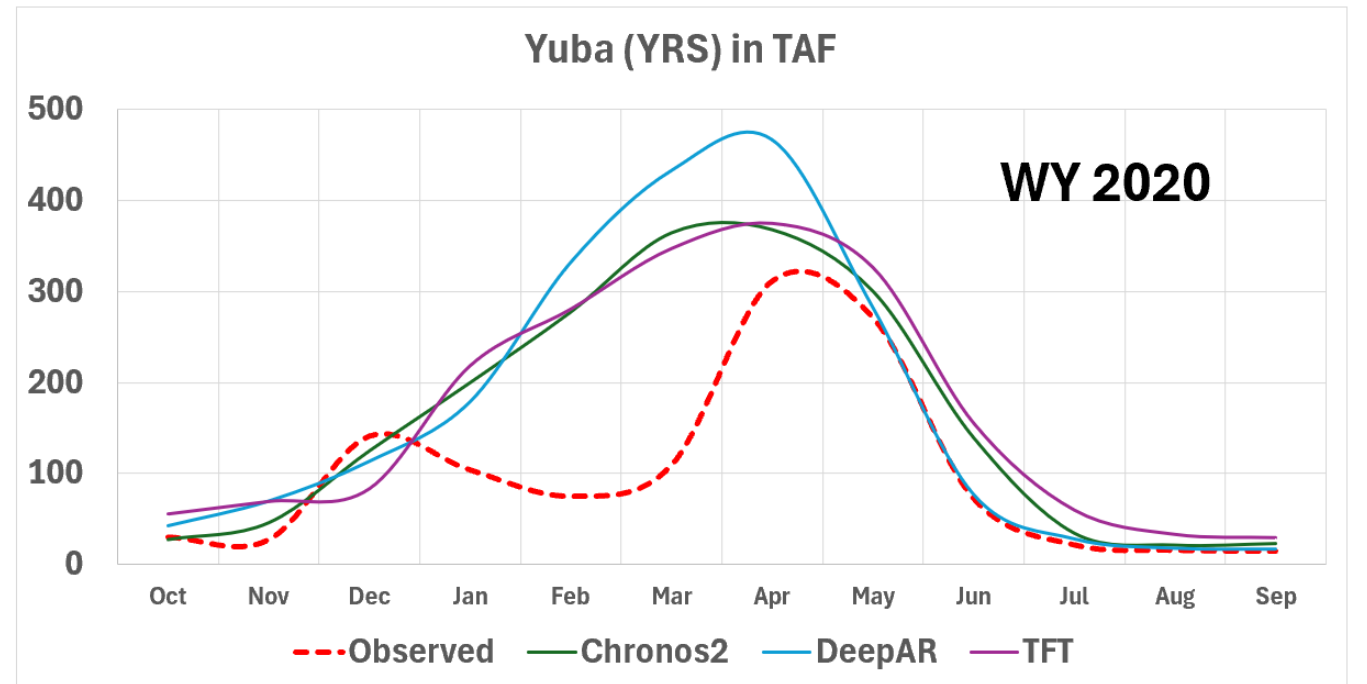
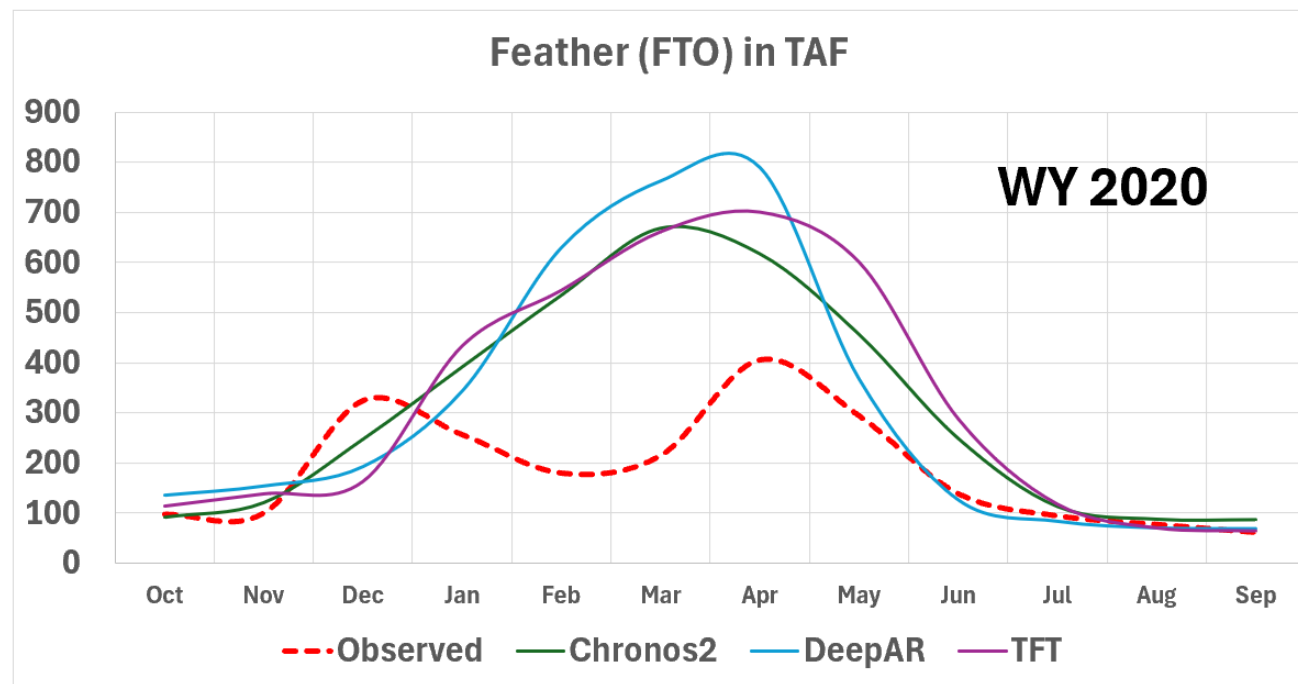
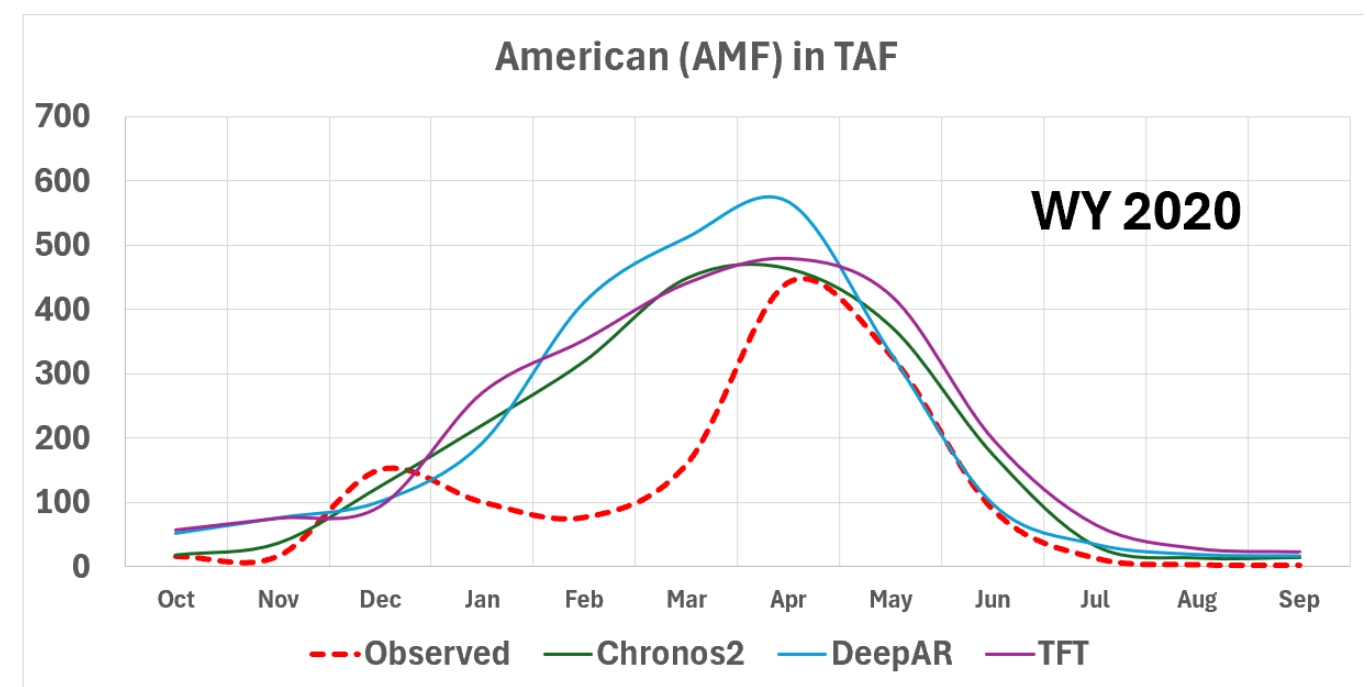
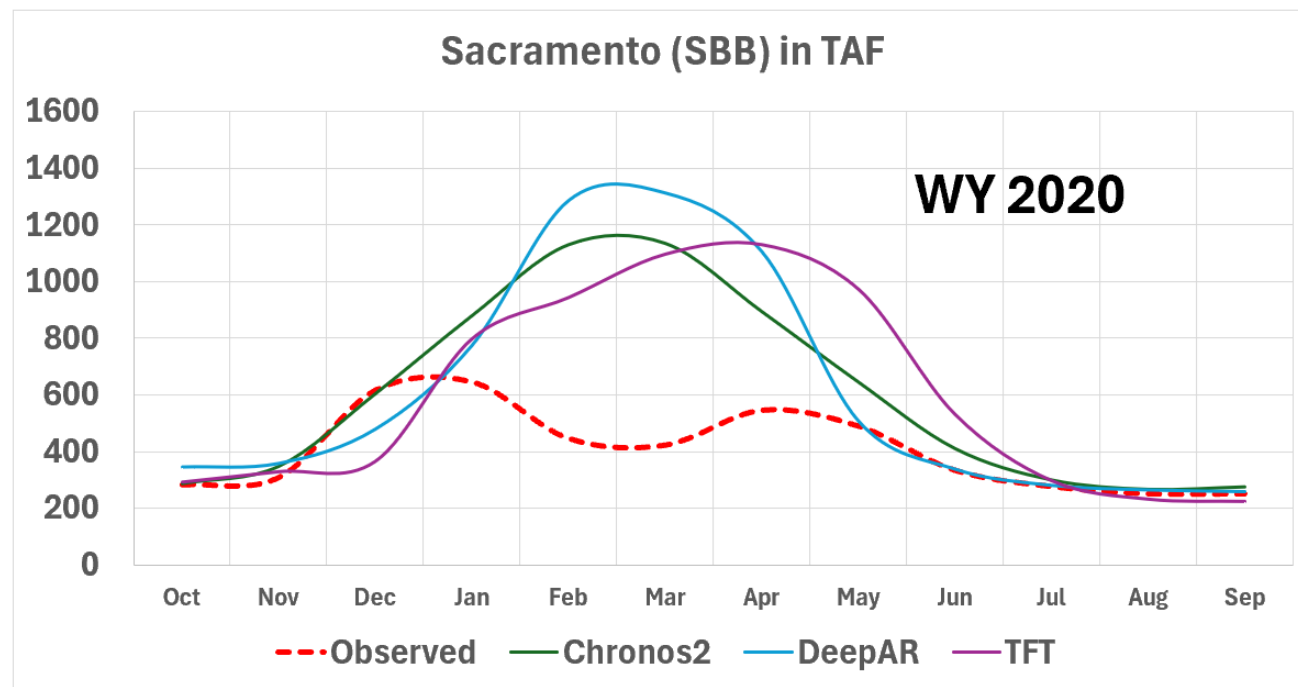
#	model	score_val	pred_time_val	fit_time_marginal	fit_order
0	WeightedEnsemble	-249.7	1.86797	11.0	11
1	ChronosWithRegressor[bolt_small]	-254.6	1.13048	6.7	8
2	Chronos2	-261.9	0.04169	18.0	5
3	Chronos2SmallFineTuned	-265.4	0.04370	290.3	6
4	TemporalFusionTransformer	-280.4	0.02164	293.6	9
5	DeepAR	-283.5	0.06991	118.4	10
6	DynamicOptimizedTheta	-291.5	0.08092	4.1	4
7	AutoETS	-295.2	0.57311	15.5	7
8	DirectTabular	-309.6	0.05908	7.0	3
9	SeasonalNaive	-325.7	0.02687	3.3	1
10	RecursiveTabular	-377.3	0.06548	7.2	2

WY2025

#	model	score_val	pred_time_val	fit_time_marginal	fit_order
0	WeightedEnsemble	-245.4	0.57651	11.6	11
1	ChronosWithRegressor[bolt_small]	-251.5	0.03084	10.5	8
2	Chronos2	-257.9	0.04574	27.8	5
3	Chronos2SmallFineTuned	-262.0	0.04159	287.4	6
4	TemporalFusionTransformer	-276.4	0.02003	293.4	9
5	DeepAR	-280.7	0.06923	118.5	10
6	DynamicOptimizedTheta	-290.7	0.09816	5.8	4
7	AutoETS	-295.5	0.37857	16.2	7
8	DirectTabular	-306.6	0.06072	7.3	3
9	SeasonalNaive	-325.3	0.02826	3.4	1
10	RecursiveTabular	-371.9	0.06346	7.0	2

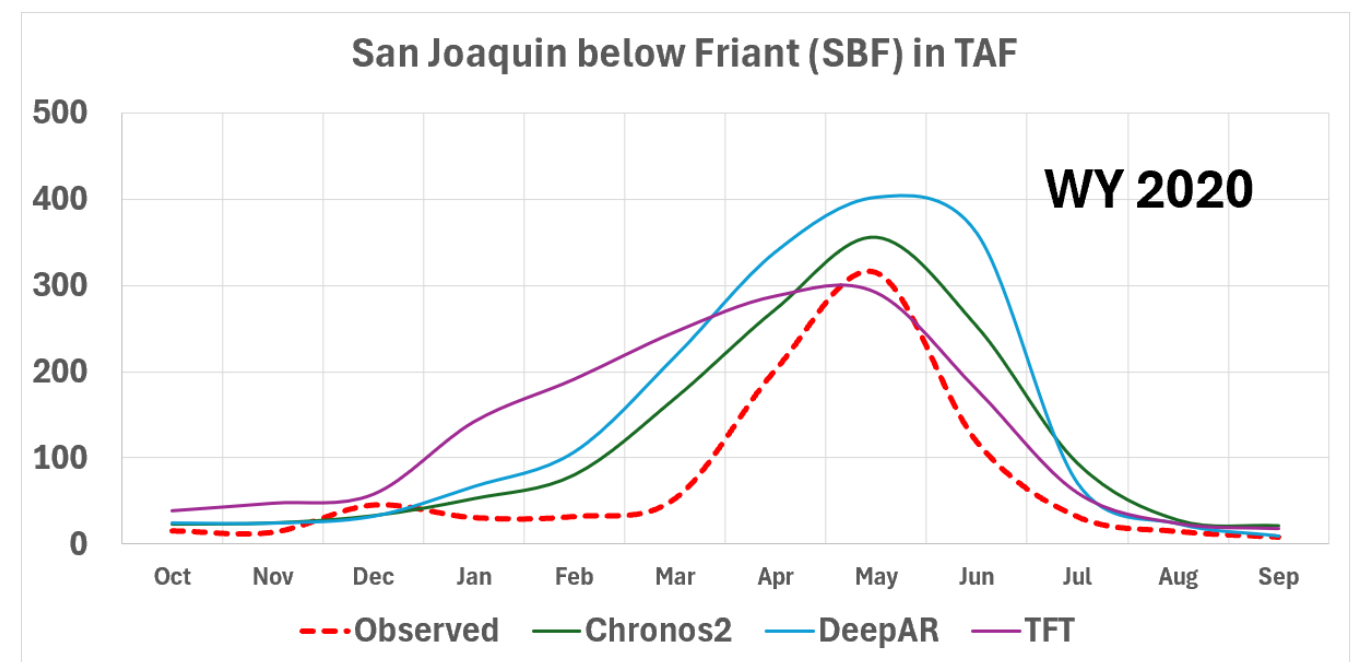
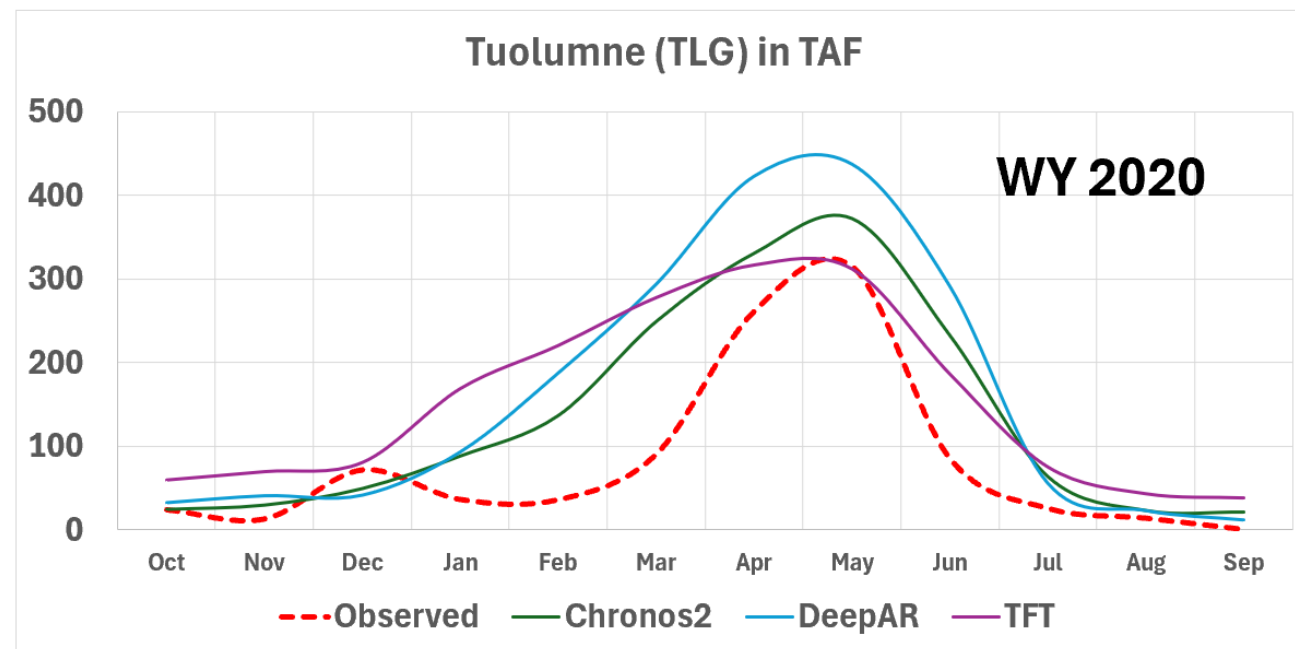
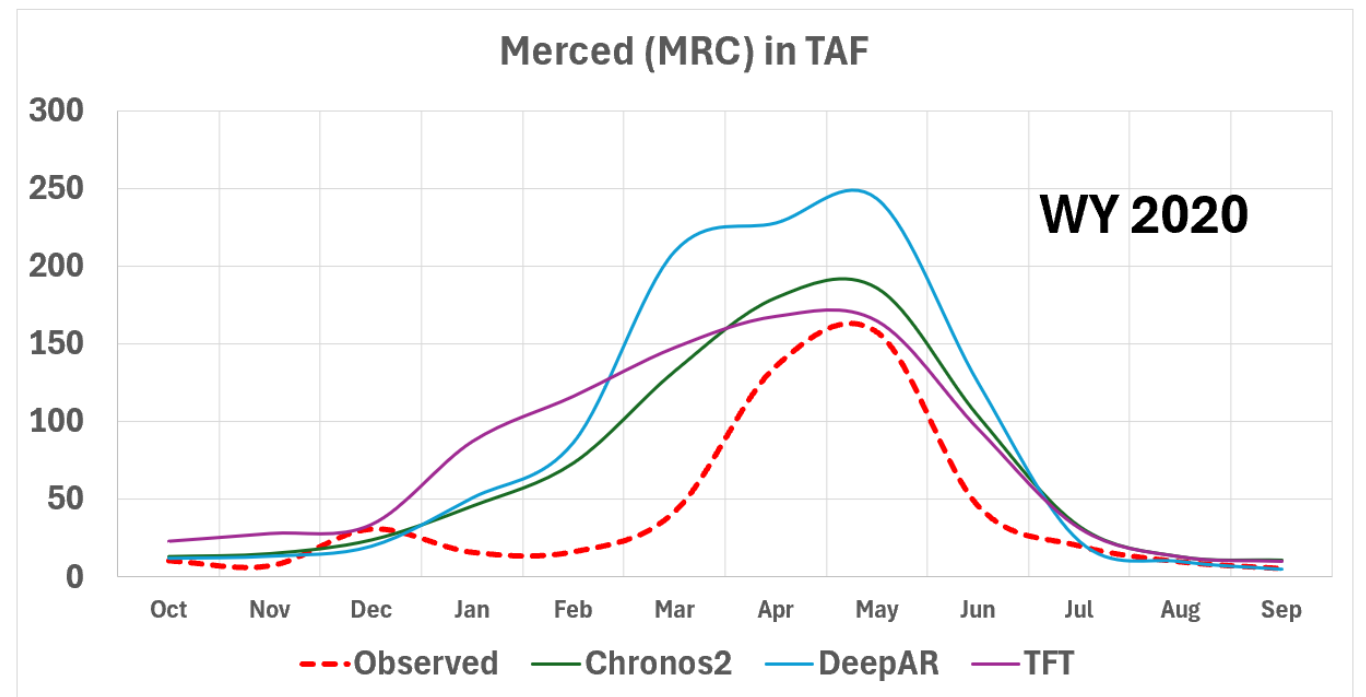
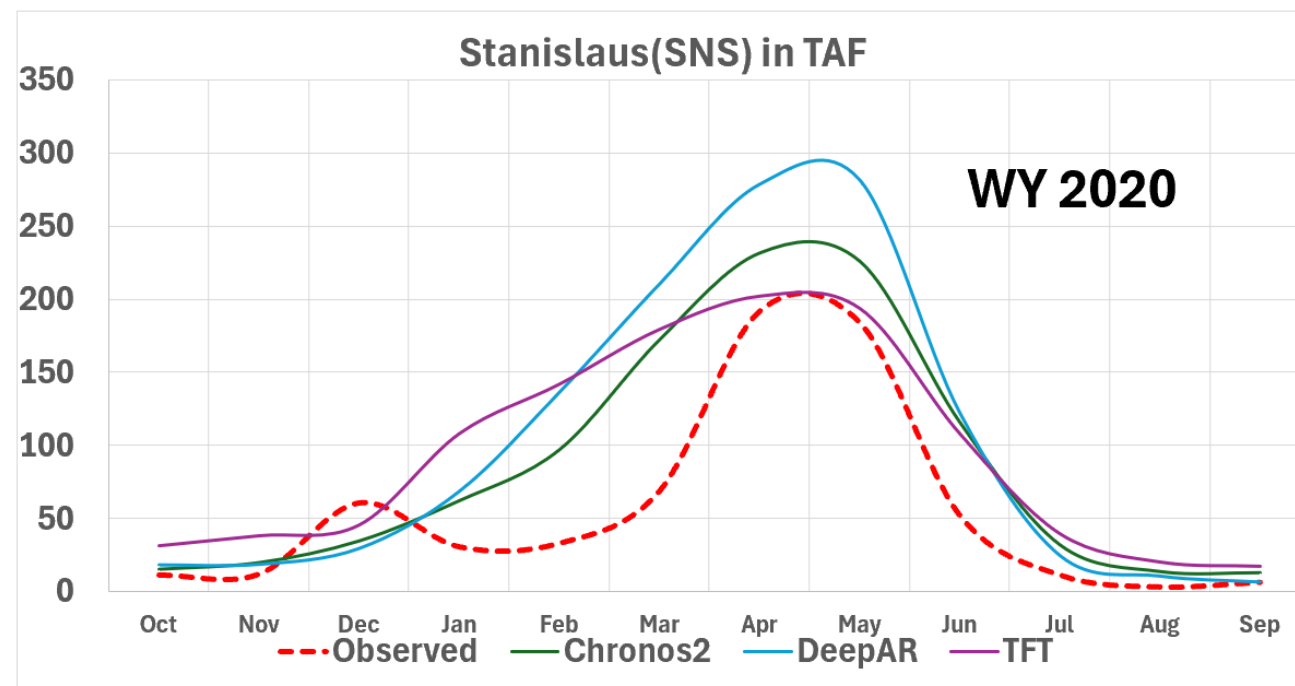
AutoGluon (top 3)

SVI



AutoGluon (top 3)

SJVI



Final Thoughts

- Application of ML and emerging AI techniques promising
- Many, many timeseries forecasting models
- Optimizing (tuning) hyperparameters important
- No one model fits all
- Classification directly (W, AB, BN, D, C) vs. numerical (via streamflows)
- Include covariates (precipitation, temperature, etc))
- Multi-year forecasting
- Online accelerators (GPUs) may be necessary (e.g., Google A100)!

Thank you!

(tkcalwater@gmail.com)