Forecasting Rio Grande Streamflow

Annie Miller, Hannah Abraham, James Wang, Jasmine Zhang

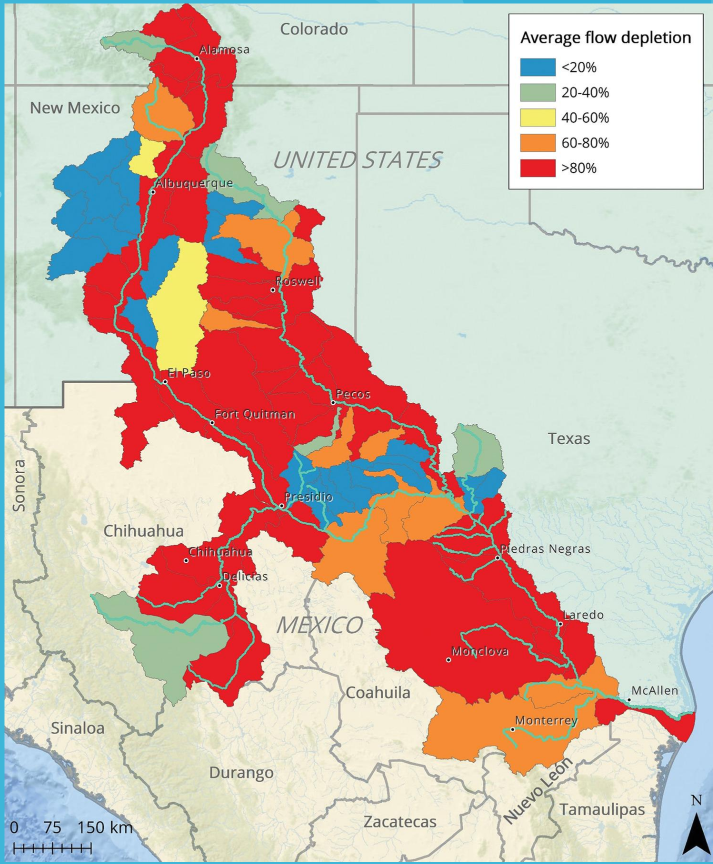
The Rio Grande's Critical Water Challenge

Key Drivers of Water Crisis:

- Increasing Water Consumption
- Existing Water Rights Exceed Sustainable Supplies
- Climate Change Impacts Reducing Streamflow
- Declining River Water Volumes

What is Streamflow?

- The volume of water moving through a river channel at a given time
- Measured in cubic feet per second (CFS)
- Influenced by:
  - Precipitation
  - Snowmelt
  - Groundwater
  - Human water usage



(Rio Grande River Flow Depletion during April-September 2000-2015 *USGS,NOAA*)

Water Allocation Challenges

- **What is Water Allocation?**
  - Distributing Limited Water Resources Among Competing Needs
  - Balancing Urban, Agricultural, and Environmental Requirements

- **Cascading Challenges:**
  - Difficulty Predicting Streamflow
  - Risks to Rio Grande Compact Deliveries
  - Increased Management Uncertainty
  - Higher Risk of Water Misallocation

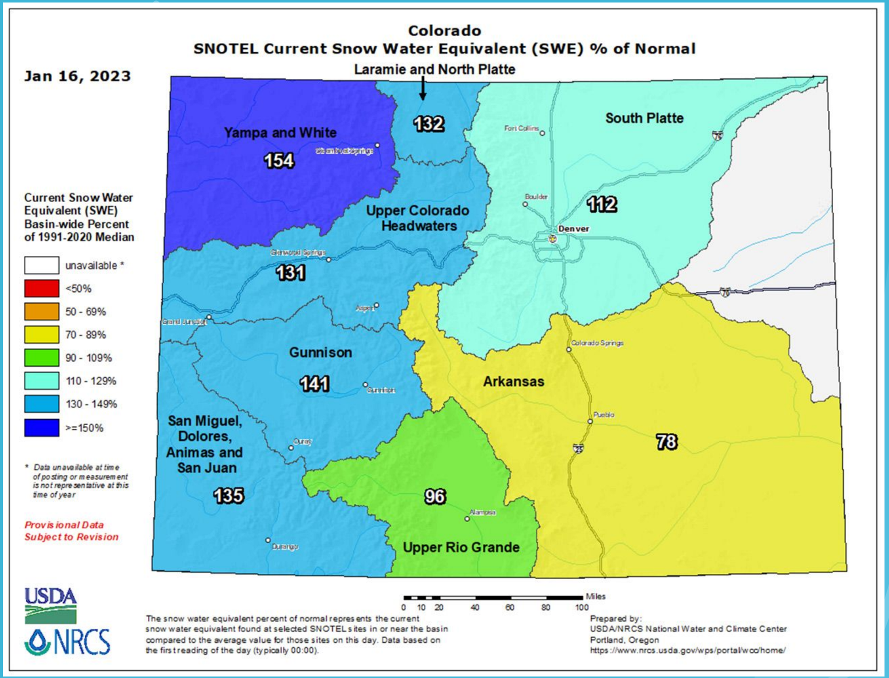
Limitations in Current Water Forecasting

Data Challenges:

- Incomplete Information on:
  - Snow Levels
  - Runoff Patterns
  - Vegetation Impact

Predictive Constraints:

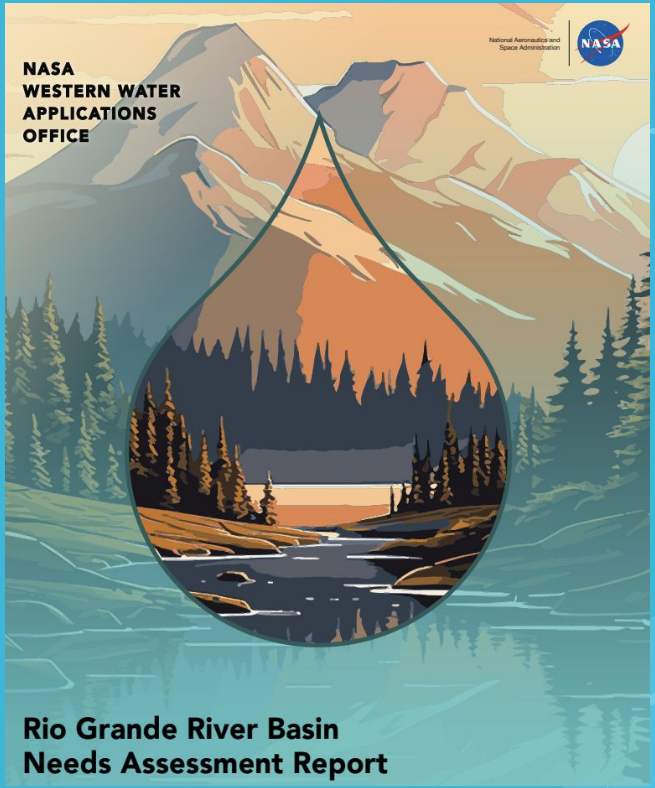
- Water Managers Struggle to Accurately Predict Streamflow
- Limited Understanding of Complex Water Systems
- Inability to Anticipate Extreme Water Events



NASA's Western Water Applications Office

- **Key Recognition:**
  - Identified Rio Grande Streamflow Prediction as **"Most Important Priority"**

- **Significance:**
  - Validates Critical Nature of Research
  - Highlights National Importance
  - Demonstrates Alignment with Scientific Priorities



Streamflow Forecasting Impact

Resource Allocation

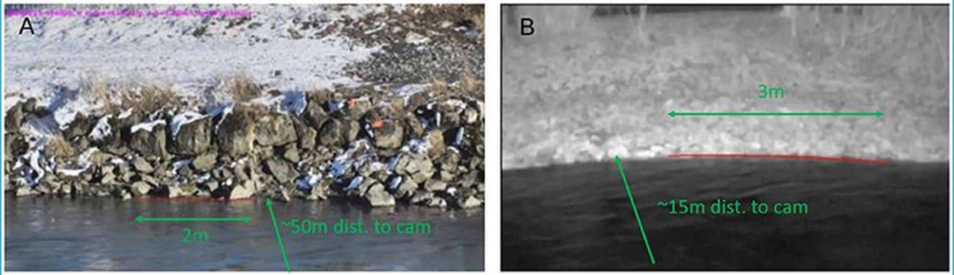
- Capture critical flow volumes
- Mitigate drought risks

Predictive Insights

- 30-day forecasts guide decisions
- Support community planning

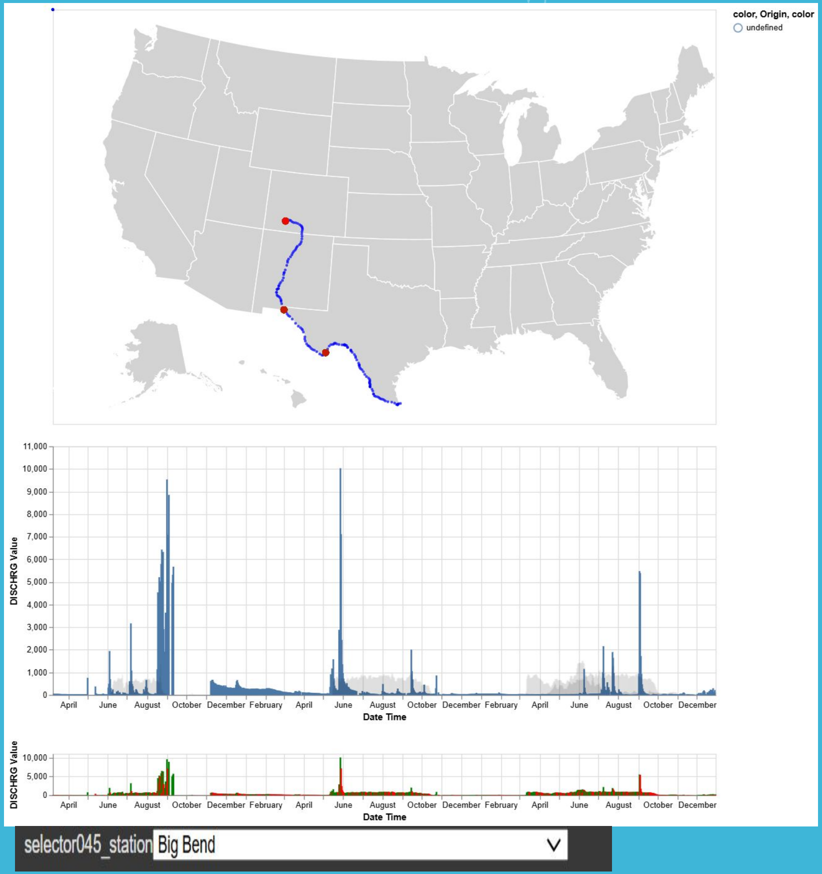
Ecosystem Protection

- Anticipate climate change impacts
- Understand environmental shifts



MVP

Minimum Viable Product



Dashboard that predicts river conditions at specific gauges along the Rio Grande.
It will provide varying forecasts (1-day, 4-day, 7-day, 30-day) of streamflow.

Live Demo!

User Questions/MVP Value

Q: What are current and future Rio Grande river conditions?

A: Users can see any immediate or future short-term risks posed by critical flow volume, drought, or flooding.

Q: How do we Identify important streamflow trends over time?

A: Community planners can guide water decisions based on previous and forecasted streamflow trends.

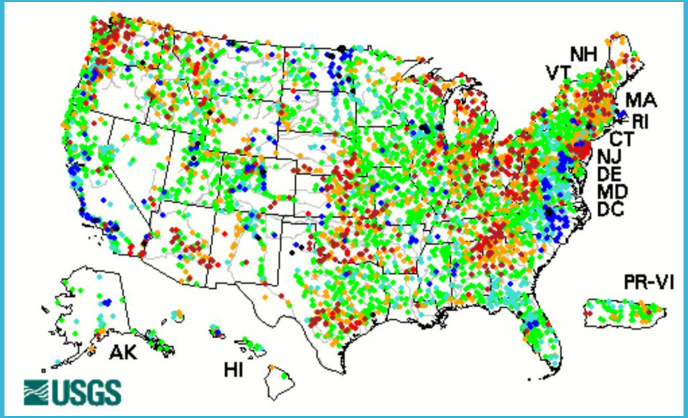
Q: Can we get accurate streamflow predictions?

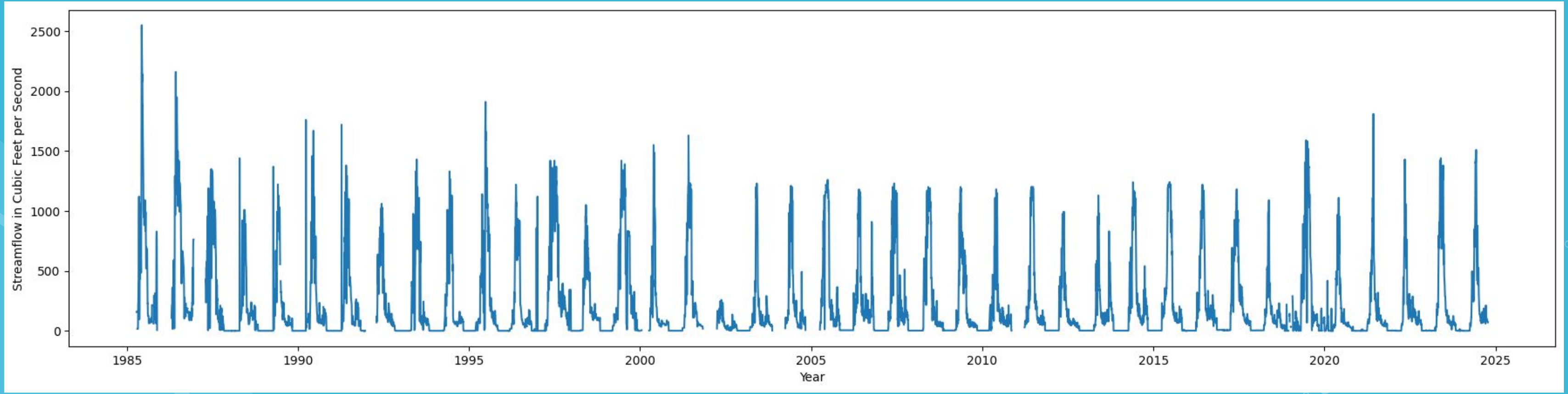
A: With our modeling approach there is potential to be more accurate than NOAA and USGS (long-term goal).

Data & Modeling Pipeline

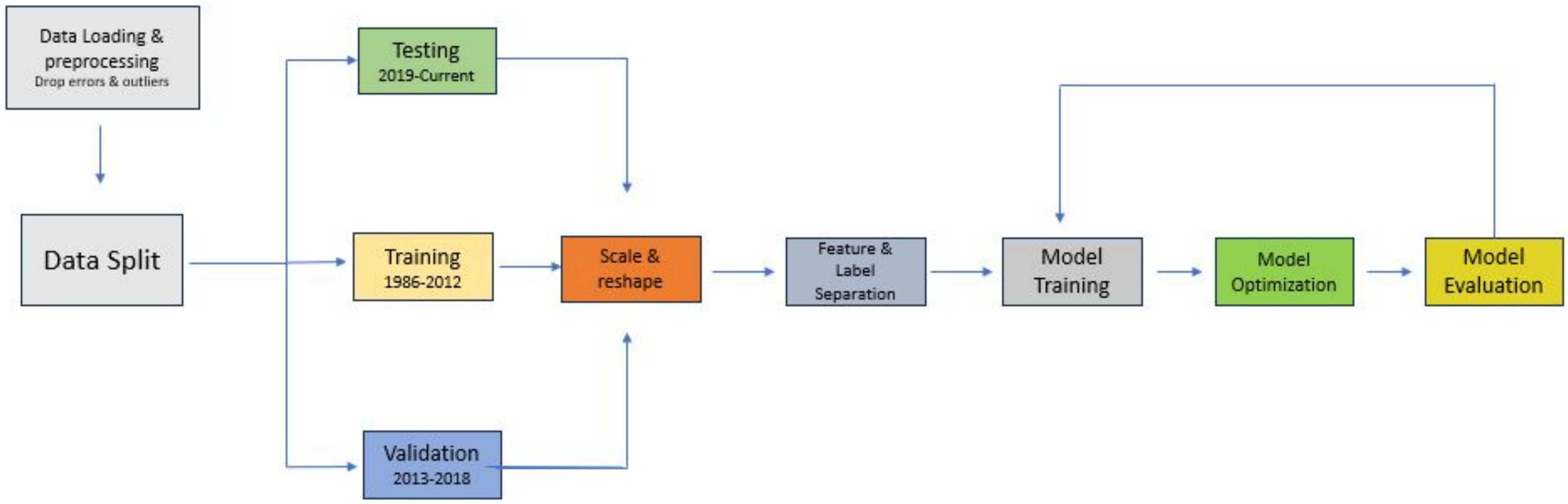
Data Sets

- Historical streamflow data
- USGS, Colorado DWR, USIBWC
- Seasonality patterns





Model Architecture



Modeling

Modeling Overview

- Used RIOMILCO Station Data– Rio Grande at Thirty Mile Bridge Near Creede for all evaluation purposes
- 23 unique models were tested
- Model types experimented with were:
    - Univariate/Multivariate LSTM
    - Multilayer Perceptron
    - Prophet
    - Transformers
    - N-Beats

Model Type Evaluation

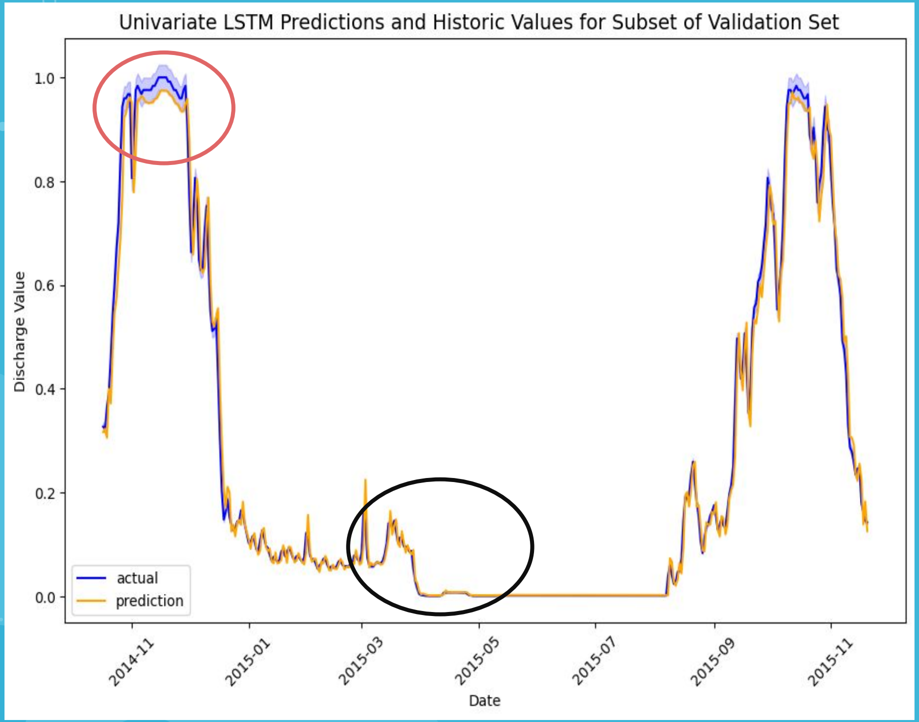
Model Type	Pros	Cons
LSTM (Long Short-Term Memory)	- Captures long-term dependencies in sequential data - Proven for hydrological applications	- Requires substantial computational resources - Longer training time
Transformer	- Effective for long sequences due to self-attention mechanism - Captures complex relationships	- Needs a large dataset for effective training - Complex architecture
Multilayer Perceptron (MLP)	- Works well with smaller datasets. - Less computational cost	- Poor at handling temporal dependencies - Less effective for sequential data
Prophet	- Easy to use, interpretable results - Handles missing data and seasonality well	- Limited for highly non-linear data and complex relationships - Struggles with long-term dependencies
N-BEATS	- Strong performance for univariate time series - Interpretable outputs	- Limited for multivariate data - Can be computationally expensive for long-term forecasts

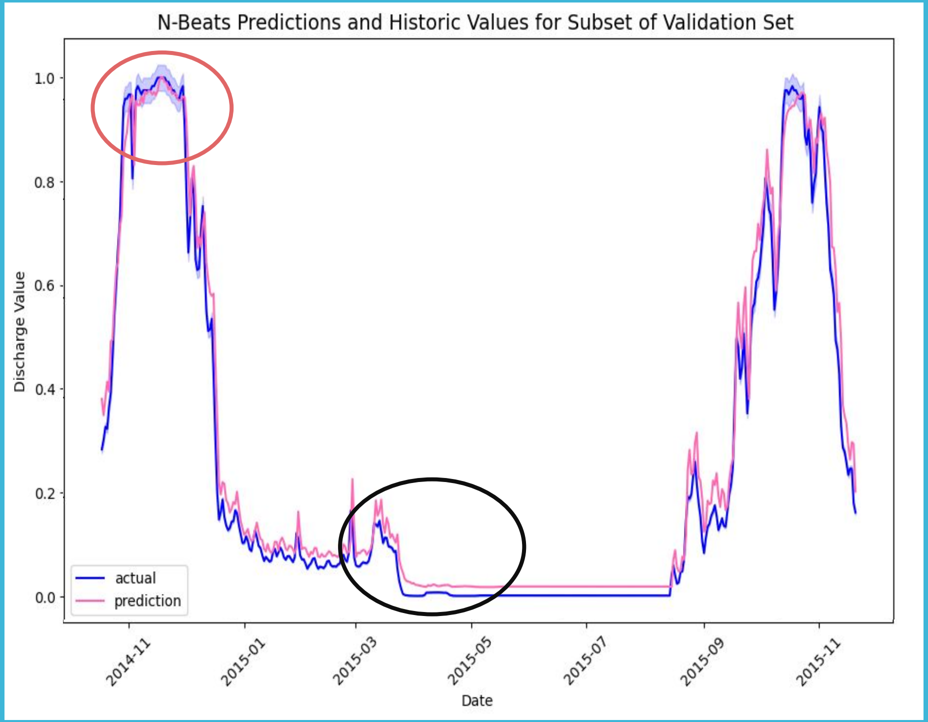
Model Evaluation Overview

Num	Model Type	Hyperparameters	Highest Performing Model's RMSE
1 - 10	LSTM 1-10	1 day forecast, 1-10 day lag	0.030505 - 7 Day lag
11-13	N-beats 1-3	1 day forecast, [15 day backcast, 30 day backcast, 35 day backcast]	0.01017988 - 30 day backcast
14	N-beats 4	4 day forecast, 30 day backcast	0.02774635
15	Multi-Layer Perceptron 1	4 day forecast	4.35
16	Transformer	4 day forecast, 100 day lag	0.02959
17	Multi-Layer Perceptron 2	7 day forecast	13.19
18	N-beats 5	7 day forecast, 30 day backcast	0.018327575
19-22	N-beats 6-9	30 day forecast, [30 day backcast, 60 day backcast, 70 day backcast, 90 day backcast]	0.04392118 - 60 day backcast

- N-Beats models outperformed all other models for 1, 4, 7, and 30 day prediction time periods

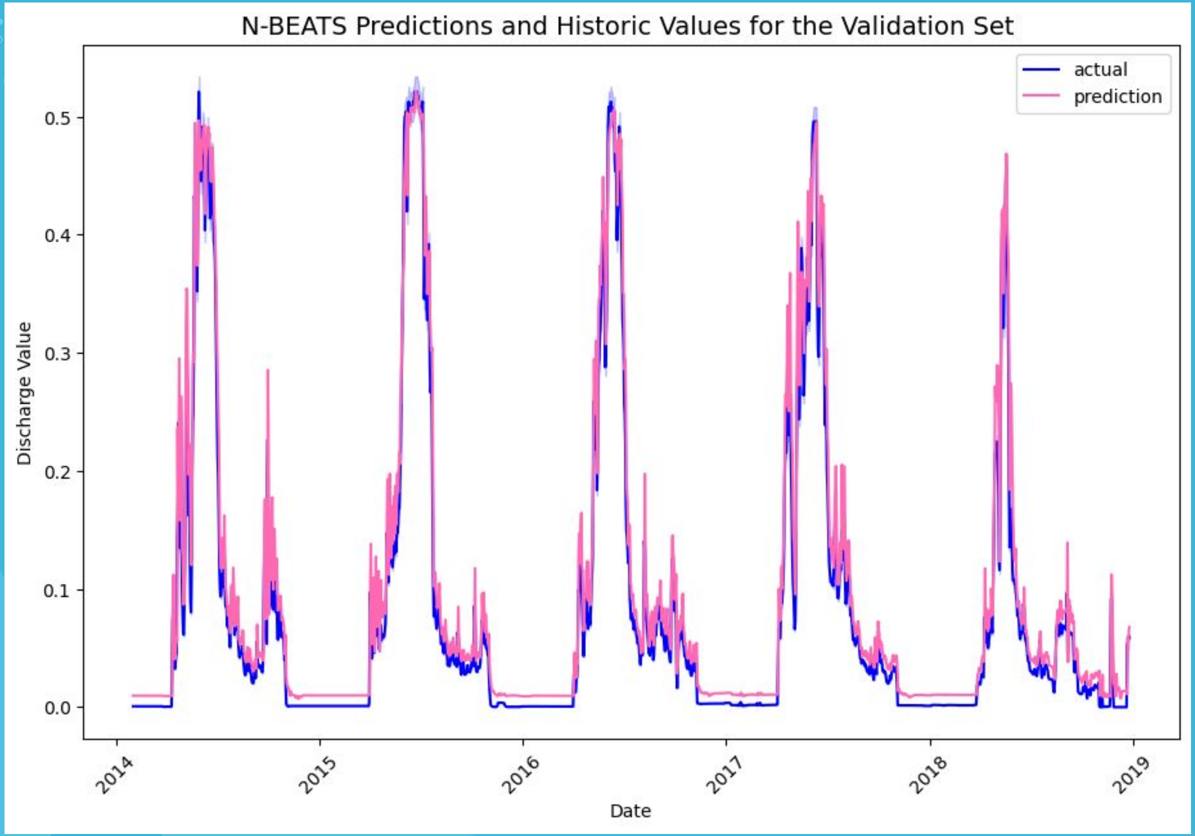
Comparison of Univariate LSTM & N-BEATS





- LSTM better follows known data and lower streamflow values
- N-BEATS better captures highest values of streamflow

N-Beats Performance



- Accurately predicts highest streamflow values
- Better predicts extreme values at the beginning of the training set
- Backcast blocks help to capture changes in seasonality over time

Technical Takeaways

1. Ability to predict univariate models within federally accepted thresholds

2. N-BEATS can accurately capture seasonal variation and changes in seasonality

3. Strong performance on 30 day prediction

Modeling & User Takeaways

1. N-BEATS models accurately capture seasonality and highest streamflow rates

2. Only slight trade-offs in accuracy for 30 day predictions

3. Robust to changes in patterns of streamflow over time due to external factors

Ethics and Privacy

- No major concerns from ethics and privacy audit
- Potential concerns for the future:
    - Biases of using historical data to predict the future
    - Placement of monitoring stations

Top Roadmap Items

- Expand forecasts to include more stations along the Rio Grande
- Generalize modeling to include other endangered rivers across the U.S.
- Aid water allocation decisions

MISSION

- Improve water resource management by providing precise, real-time river forecasts
- Empower decision-makers to protect water resources, support communities, and preserve endangered river ecosystems

Thank You!

Acknowledgements

Thank you to Dr. Erin Urquhart and NASA's western water office for their guidance on our project.

Thank you to Joyce Shen, Morgan G. Ames, our classmates, and the 210 teaching team for their continuous feedback and support throughout the semester.

Resources

1. https://waterdata.usgs.gov/nwis/rt - USGS Current Water Data for the Nation
2. https://www.weather.gov/wgrfc/ - NOAA river forecasts
3. https://waterdata.ibwc.gov/AQWebportal/Data/Location/Summary/Location/08377200/Interval/Latest - Rio Grande Streamflow Data
4. https://wwao.jpl.nasa.gov/water-portfolio/ - NASA Western Water Applications Office Water Portfolio Page
5. Muhamad Nur Adli Zakaria, Ali Najah Ahmed, Marlinda Abdul Malek, Ahmed H. Birima, Md Munir Hayet Khan, Mohsen Sherif, Ahmed Elshafie, Exploring machine learning algorithms for accurate water level forecasting in Muda river, Malaysia, Heliyon, Volume 9, Issue 7, 2023, e17689, ISSN 2405-8440, https://doi.org/10.1016/j.heliyon.2023.e17689.
6. Zalenski, G., W. F. Krajewski, F. Quintero, P. Restrepo, and S. Buan, 2017: Analysis of National Weather Service Stage Forecast Errors. Wea. Forecasting, 32, 1441–1465, https://doi.org/10.1175/WAF-D-16-0219.1.
7. Naresh Kedam, Deepak Kumar Tiwari, Vijendra Kumar, Khaled Mohamed Khedher, Mohamed Abdelaziz Salem, River stream flow prediction through advanced machine learning models for enhanced accuracy, Results in Engineering, Volume 22, 2024, 102215, ISSN 2590-1230, https://doi.org/10.1016/j.rineng.2024.102215.
8. Rahimzad, Maryam & Moghaddam Nia, Alireza & Zolfonoon, Hosam & Soltani, Jaber & DANANDEH MEHR, Ali & Kwon, Hyun-Han. (2021). Performance Comparison of an LSTM-based Deep Learning Model versus Conventional Machine Learning Algorithms for Streamflow Forecasting. Water Resources Management. 35. 1-21. 10.1007/s11269-021-02937-w.
9. Francis Yongwa Dtissibe, Ado Adamou Abba Ari, Hamadjam Abboubakar, Arouna Ndam Njoya, Alidou Mohamadou, Ousmane Thiare, A comparative study of Machine Learning and Deep Learning methods for flood forecasting in the Far-North region, Cameroon, Scientific African, Volume 23, 2024, e02053, ISSN 2468-2276, https://doi.org/10.1016/j.sciaf.2023.e02053.

Model Results Table

Num	Model Type	Hyperparameters	RMSE
1	LSTM 1	1 day forecast, 1 day lag	0.07466
2	LSTM 2	1 day forecast, 2 day lag	0.06070
3	LSTM 3	1 day forecast, 3 day lag	0.046717
4	LSTM 4	1 day forecast, 4 day lag	0.0459
5	LSTM 5	1 day forecast, 5 day lag	0.0414
6	LSTM 6	1 day forecast, 6 day lag	0.0398
7	LSTM 7	1 day forecast, 7 day lag	**0.030505**
8	LSTM 8	1 day forecast, 8 day lag	0.052649
9	LSTM 9	1 day forecast, 9 day lag	0.041738
10	LSTM 10	1 day forecast, 10 day lag	0.10599
11	N-beats 1	1 day forecast, 15 day backcast	0.0149606
12	**N-beats 2**	**1 day forecast, 30 day backcast**	**0.01017988**
13	N-beats 3	1 day forecast, 35 day backcast	0.0161285
14	**N-beats 4**	**4 day forecast, 30 day backcast**	**0.02774635**
15	Multi-Layer Perceptron 1	4 day forecast	4.35
16	Transformer	4 day forecast, 100 day lag	0.02959
17	Multi-Layer Perceptron 2	7 day forecast	13.19
18	**N-beats 5**	**7 day forecast, 30 day backcast**	**0.018327575**
19	N-beats 6	30 day forecast, 30 day backcast	0.06750578
20	**N-beats 7**	**30 day Forecast, 60 day backcast**	**0.04392118**
21	N-beats 8	30 day Forecast, 70 day backcast	0.04525362
22	N-beats 9	30 day forecast, 90 day backcast	0.0452625

technical discussion

. For example: summarize data pipeline (30 sec), overall architecture and ML pipeline (1 min), discuss model(s) and technique(s) in compelling details so that the audience (class) understand what you are building “under the hood” and why. Overall, show and discuss the data science techniques used, how exactly you are applying the techniques in your project, challenges, trade-offs, current results. Put your CTO, data architect, data engineer, and data scientist hats on.

Presentation Rubric

Requirement	Done?
• Data Analysis Data collection and identification: Data resources are clearly identified and are adequate for addressing the problem space and proposed solution (or have been identified as inadequate, and a plan to pivot has been introduced).	
• Data exploration: Early data has been explored and insights from EDA have been identified that confirm approach or inform a pivot strategy.	
• EDA is appropriate, complete, and informs decisions and plans for methods/models. *(Note that some teams may be taking on a project space that requires more or less EDA, setting up data collection apparatus, etc.) Evaluation of this criterion are dependent on scope and details that vary project to project.*	
• Clear identification of the problem, impact, target user segment, and explanation of how MVP will serve the target users and use case	
• Cloud-based infrastructure for data science pipeline and model options are discussed. Reasons for pursuing specific architecture and models and methods are articulated. Methods and potential models that are being / to be constructed are explicitly tied to data collection, analysis, and to contextual details of problem space and use case.	
• At least one model is being built and evaluated. The team has at least “line of sight” to the evaluation methodology	
• Options and decisions are articulated.	

Presentation 2 guidance

Learning Objective 1

Teams will be assessed on their **decision making**: what options they explored, if their decisions are driven by data and evidence, and how well they can articulate the alignment of decisions and data, analysis, and contextual details of their problem space.

Learning Objective 2

Teams will be assessed on the **Technical Work** that is in progress: each team should have a more defined idea of what the MVP looks would entail and at least one ML/DL/Generative model completed with preliminary evaluation methodology formed.

Learning Objective 3

Teams will be assessed on their ability to **articulate their plan for the work that remains** in their project: What the final project will look like, how feasible it will be to implement their plan, what obstacles they may face, and what steps still need to be taken in order to achieve their MVP for the end of term.

Capstone Presentation 2 Guidance

2 minutes -- overview of the customer / user problem you are solving and why it is an important problem to solve -- thematic problem, problem you are specifically solving, target user, market opportunity, impact. (you should have nailed this part). Anyone and everyone should immediately understand why you are building what you are building. Strong and clear motivation to the problem.

3 minutes -- describe the MVP and key features that directly relate back to the problem for the target user. Anyone and everyone should understand what you intend to deliver and why it is the right and valuable deliverable for your target user.

- Identify key takeaways/quotes from talking to target users and/or domain experts and connect these feedback.
- Identify 1-2 key questions that your users have, and how your deliverable will help answer the specific questions.
- If it is not an MVP but a paper, describe how your paper will add value to your end user's knowledge base and work.

7-8 minutes -- technical discussion. For example: summarize data pipeline (30 sec), overall architecture and ML pipeline (1 min), discuss model(s) and technique(s) in compelling details so that the audience (class) understand what you are building "under the hood" and why. Overall, show and discuss the data science techniques used, how exactly you are applying the techniques in your project, challenges, trade-offs, current results. Put your CTO, data architect, data engineer, and data scientist hats on.

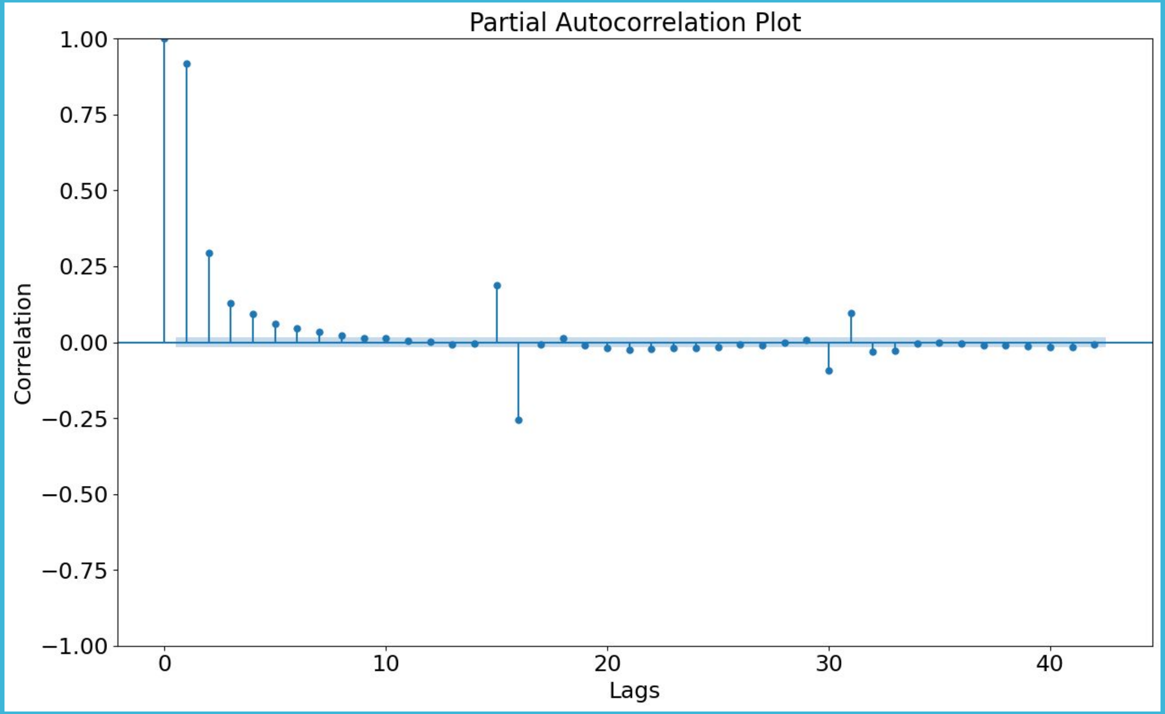
30 seconds -- Highlight the remaining key areas to tackle and project management for the remainder of the semester

Presentation outline

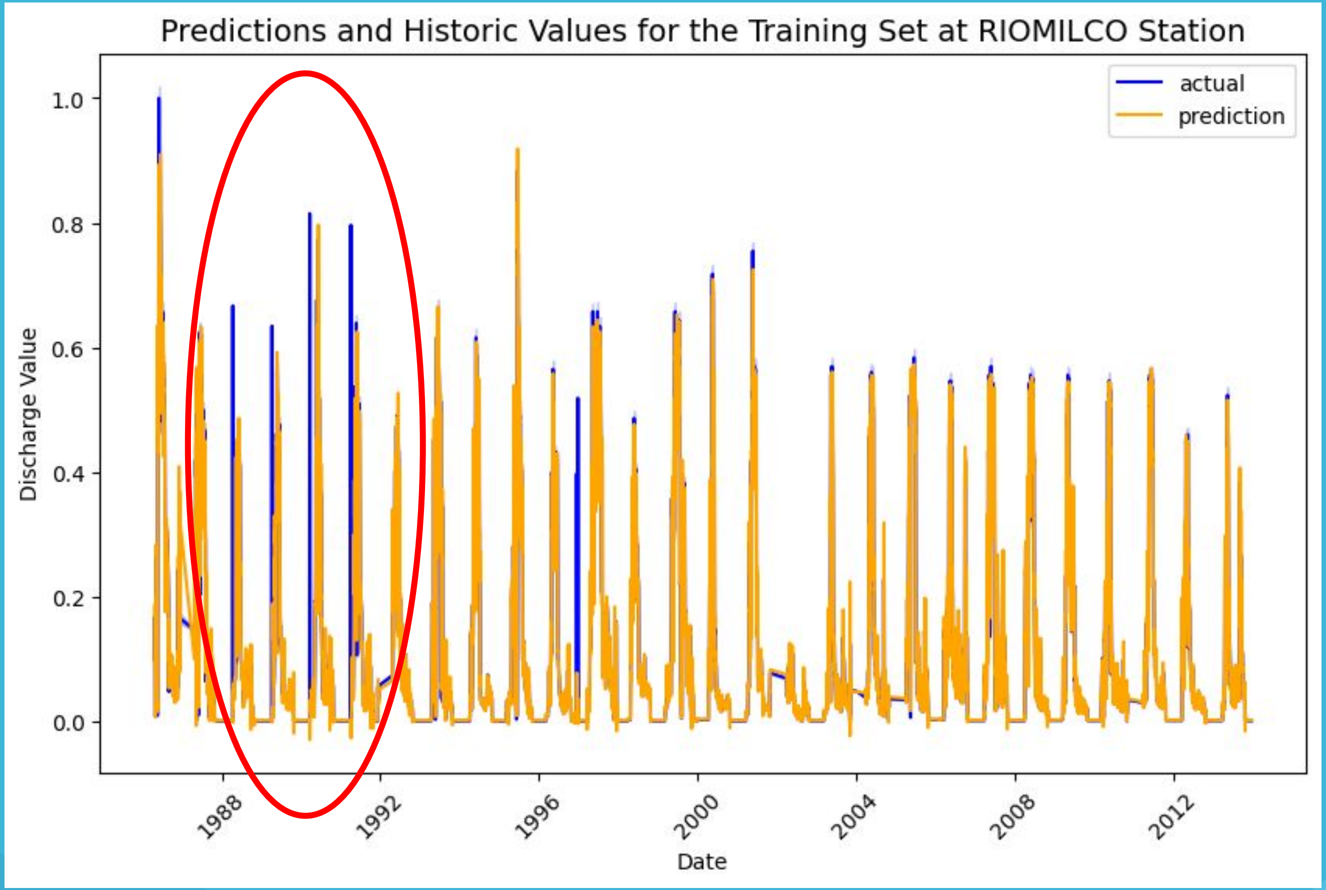
- Overview of the problem (James) 2 minutes
  - What is the problem
  - Why is it important
  - Target user
- MVP (Hannah) 3 Minutes
  - Describe MVP
  - Key Features
    - Why do these features matter to our customer
  - Identify 1-2 key questions that our users may have and how our deliverable will answer these questions
- Technical Discussion (Jasmine, Hannah, Annie, James) 8 Minutes
  - Data Summarization & collection – Jasmine (30 second)
  - Early EDA summary – Jasmine (30 seconds)
  - Summarize Data pipeline – James (30 seconds)
  - Overall architecture and ML Pipeline – James (30 seconds)
  - Discuss models and techniques – At least one model built and evaluated – Hannh & Annie (3 minutes each)
    - Very detailed discussion of technical specifics
  - Technical Challenges, Trade offs and current results
    - Make sure to show a baseline

Univariate LSTM

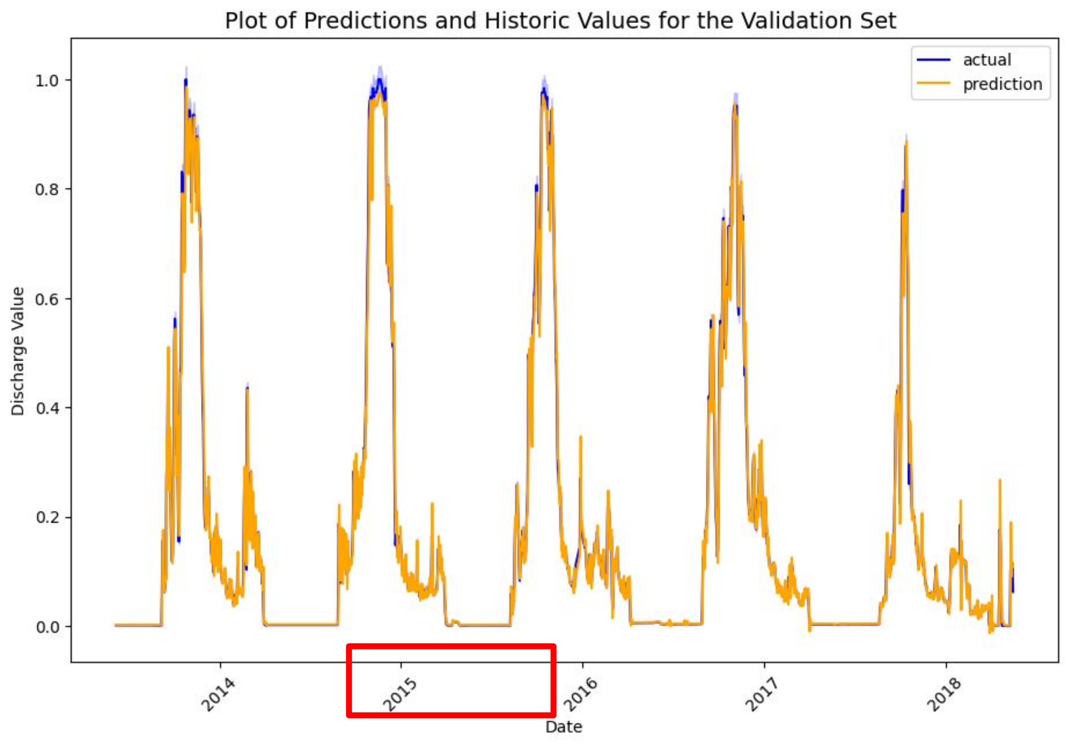
- Predicting Streamflow rates from past streamflow only

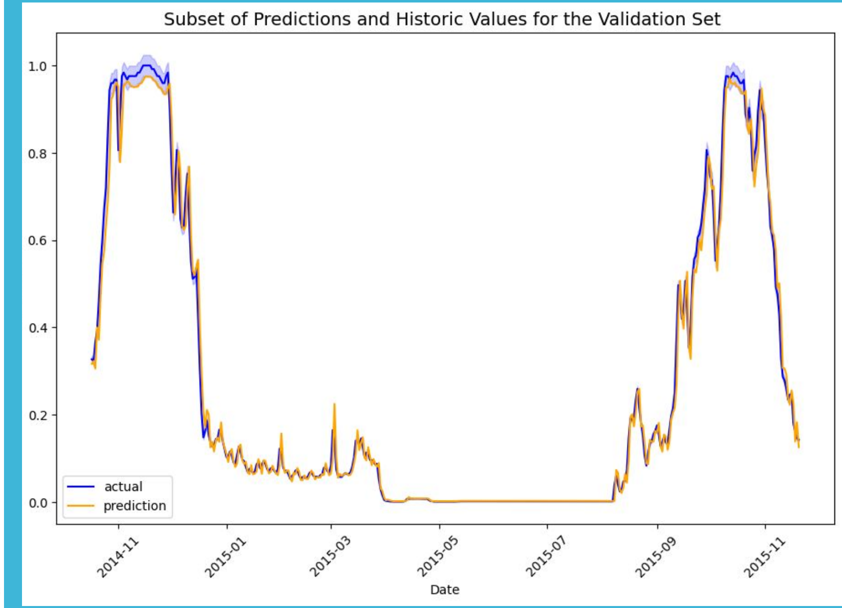


Analysis of Univariate LSTM

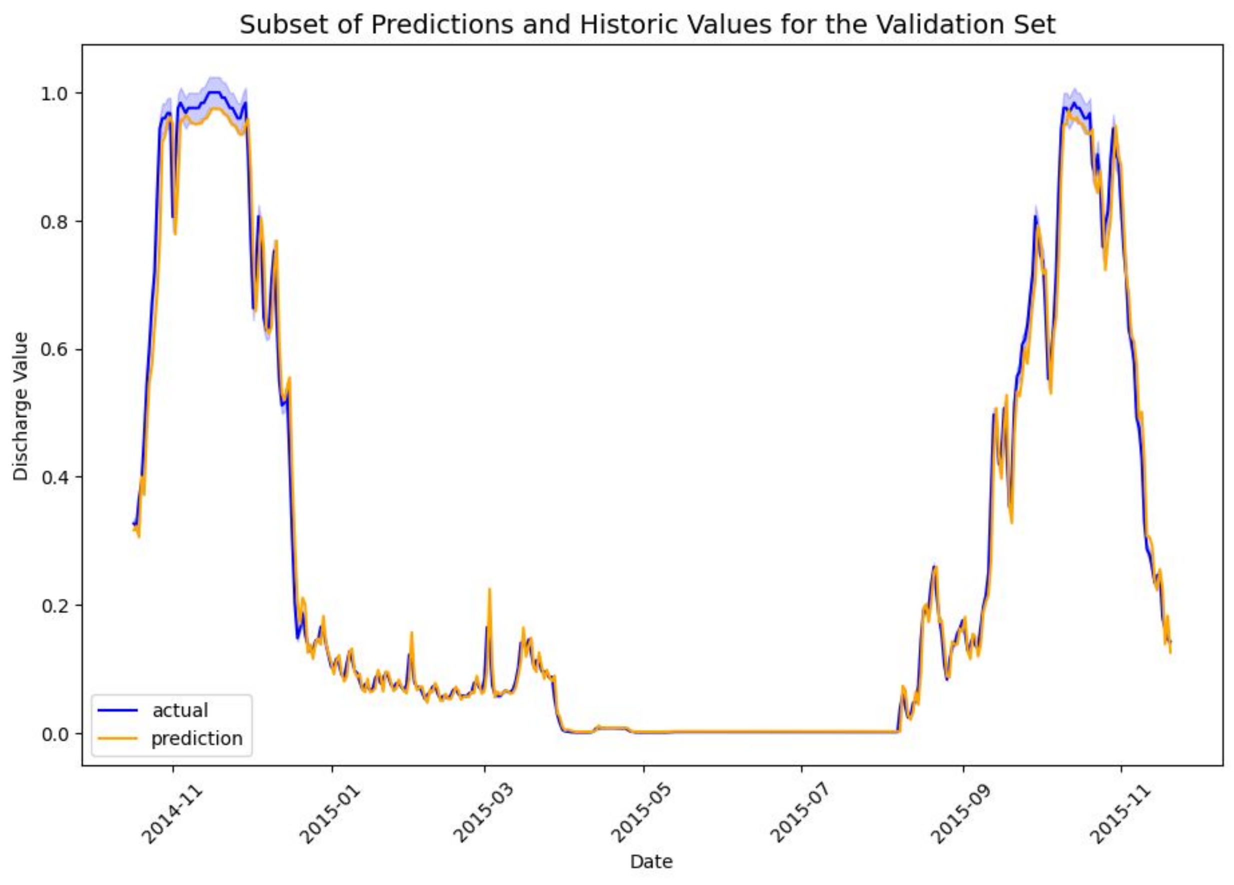


Analysis of Univariate LSTM Validation Data

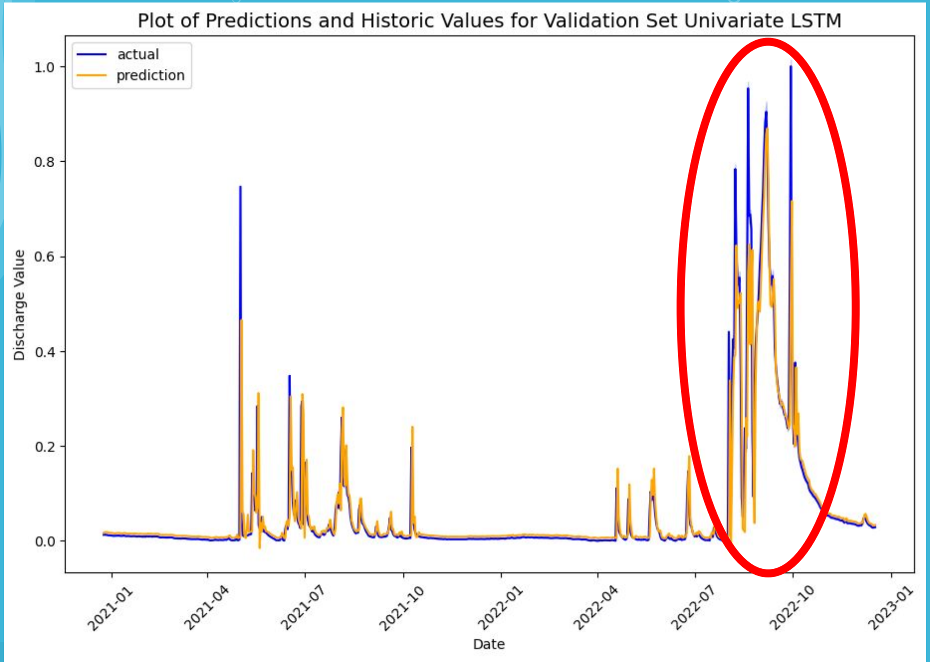


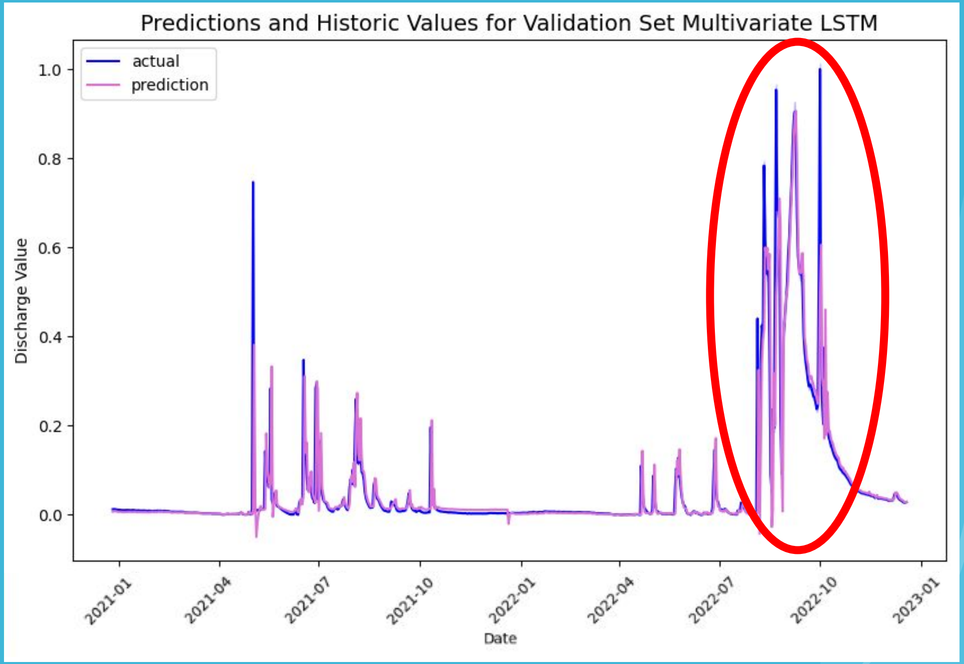


Analysis of Univariate LSTM



Multivariate LSTM Comparison





Multi-layer Perceptron

- Motivated to experiment with this model to capture non-linear trends in the time series streamflow data
- Experimenting with model inputs and outputs
  - 4-day and 7-day forecasts
  - How many lags to include in inputs
  - Differencing the data

Model	RMSE
Multi-Layer Perceptron 1 Hidden layer 7 day forecast	13.19
Multi-Layer Perceptron 1 Hidden layer 4 day forecast	4.35

N-Beats Performance

Model Name	Forecast	Backcast	RMSE
N–beats 1	1	15	0.0149606
N–beats 2	1	30	0.01017988
N–beats 3	1	35	0.0161285
N–beats 4	4	30	0.02774635
N–beats 5	7	30	0.018327575
N–beats 6	30	30	0.06750578
N–beats 7	30	60	0.04392118
N–beats 8	30	70	0.04525362
N–beats 9	30	90	0.0452625