Colorado Ground Water Analysis

Projecting water levels based on RCP8.5 scenarios

Garret Flowers, Shivam Pandey

December 2023

Abstract

In this study we investigate historical precipitation, temperature, and groundwater levels in the front range region of Colorado and how those variables may be impacted by climate change. We apply a machine learning approach using convolution neural networks to the three main aquifer monitoring stations servicing the Denver, Boulder, Pueblo, and Colorado Springs areas to assess groundwater level trends under RCP4.5 and RCP8.5 scenarios. The RCP8.5 scenario for Muddy Creek indicates a drastic drop in groundwater levels in 2067, but otherwise a relatively stable mean throughout the 21st century. The Pueblo site shows a neutral trend. Overall, this suggests that a worst-case climate change scenario may not impact Colorado's front range as dramatically as other studied regions[2,9]

Contents

1 Introduction

As the planet warms, the likelihood of climate related disasters increases [1]. Greenhouse gas emissions from human activities accelerates sea level rise, intensifies weather patterns, impacts water and food security, decreases healthfulness of the population, and continues to drive economic loss [2]. While the outlook appears bleak, any incremental offset to greenhouse gases will drastically decrease the long-run economic impact, therefore placing importance on the use of prediction to analyze and calculate these risks [3].

The responsibility of macro-level climate research is largely driven by the World Climate Research Programme (WCRP) and its partnerships with governmental organizations across the globe. Within the WCRP is the Working Group on Coupled Modelling (WGCM), which is responsible for the development and review of the Coupled Model Intercomparison Project (CMIP) and its resulting climate ensemble data. While CMIP data encompasses a global view of climate, there are different regional applications that can be explored. One example of climate modeling research using regional CMIP data can be seen in a study examining climate change in Germany. The authors focused on groundwater and examined its relationship between precipitation and temperature, then modeled its potential future state using representative concentration pathway (RCP) scenarios [2,8]. These RCP scenarios are published by the Intergovernmental Panel on Climate Change (IPCC) and represent various climate change scenarios given certain atmospheric densities of greenhouse gases into 2100 [8]. A similar study was done which utilized precipitation, temperature, and aquifer measurements in Iran along with CMIP models under an RCP8.5 scenario to measure climate change impacts to groundwater levels [9].

This framework of existing research fosters the opportunity to continue exploring new regions. With three aquifer measurement sites and thorough climatic data in the Colorado front range, we will assess groundwater levels, precipitation, and temperature under RCP4.5 and RCP8.5 scenarios, thus driving more discussion and context around localized climate change impacts.

2 Methodology

2.1 Data

The four pieces of information needed to complete the study are: Groundwater levels, Temperature, Precipitation, and Representative Concentration Pathways (RCP) scenarios. We will then discuss feature engineering to prepare the data for modeling, and the usage of RCP for projections.

2.1.1 Groundwater

The United States Geological Survey (USGS) hosts an abundant amount of water data with a good historical record. While examining the state of Colorado, we found thousands of sites documenting various water readings including river levels, aquifers, and wells. While the availability of water data is good, there were only three official USGS sites that had data from two aquifers and one drought well in the front range, which were most relevant to our study as visible in Figure1.

Two of the three sites contain data back to 2016 and one site contains data from 2017. The measurement for the three sites is represented in feet from ground to water surface as in Figure 2. Due to the incompleteness of the Kit Carson data, we ended up dropping it from our analysis. Figures 3 and 4 depict a closer inspection of groundwater levels for our selected sites.

2.1.2 Temperature Precipitation

The next two pieces needed are temperature and precipitation which can serve as proxies for evapotranspiration and water recharge, respectively [2]. Conveniently, the same three stations offering groundwater readings also have daily temperature and precipitation readings. By using the station number associated with each station, we can pull via the USGS.

2.1.3 CORDEX RCP Scenarios

To validate our results, we pull downscaled CMIP data regional to North America (NA-CORDEX) under the RCP4.5 and RCP8.5 scenarios [6]. RCP4.5 is generally seen as a baseline scenario, which requires that carbon dioxide emissions begin to decline around 2045 and that methane emissions stop increasing by 2050

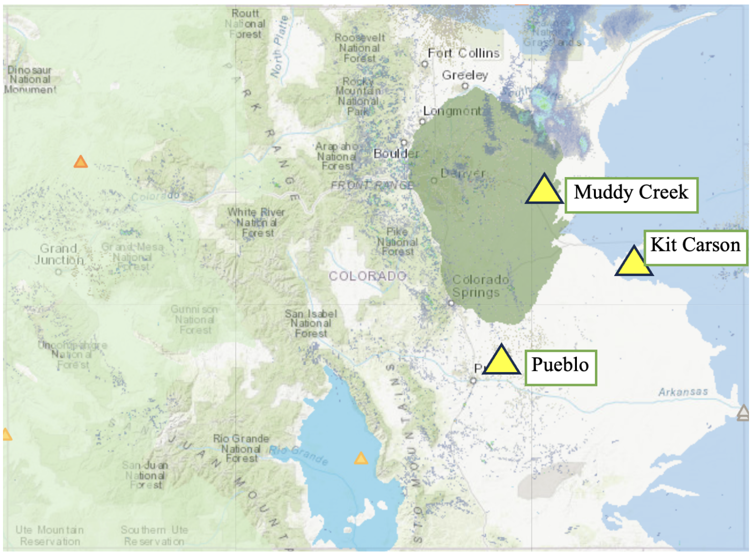

Figure 1: Selected sites with Aquifer Overlay in Colorado

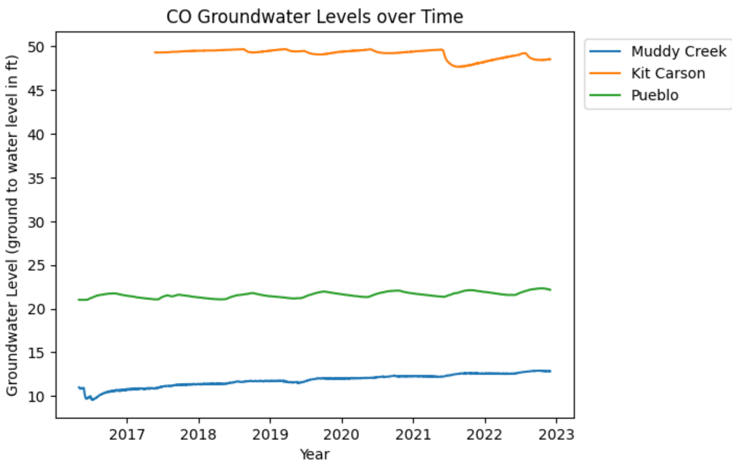

Figure 2: Ground water level data for selected 3 sites in Colorado

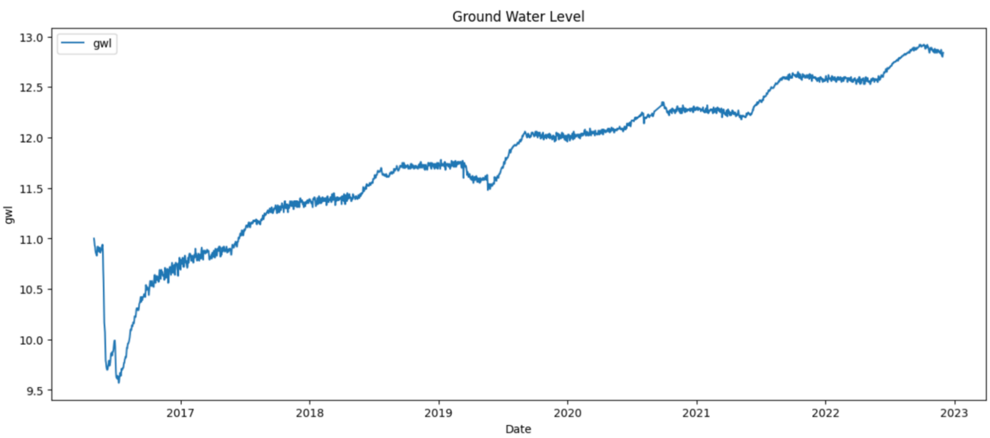

Figure 3: Ground water level data for Pueblo

[8]. This would likely result in a global mean temperature increase between 2C to 3°C by 2100 [8]. The RCP8.5 scenario is seen as a worst-case scenario, where greenhouse gas emissions continue to rise into 2100, likely resulting in a global mean temperature increase of 5°C.

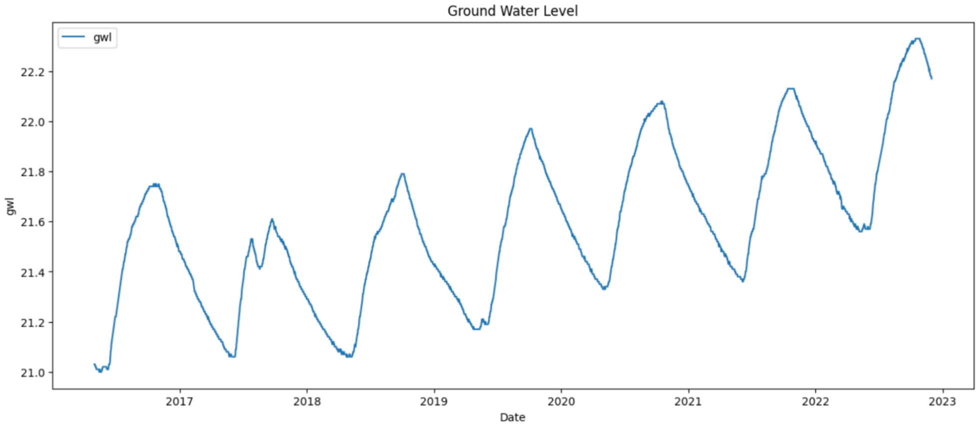

Figure 4: Ground water level data for Muddy Creek

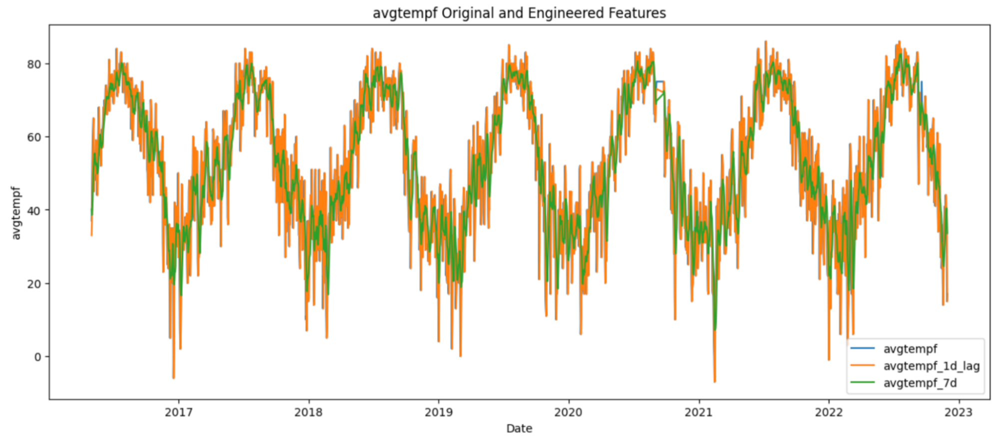

Figure 5: Temperature data for Muddy Creek

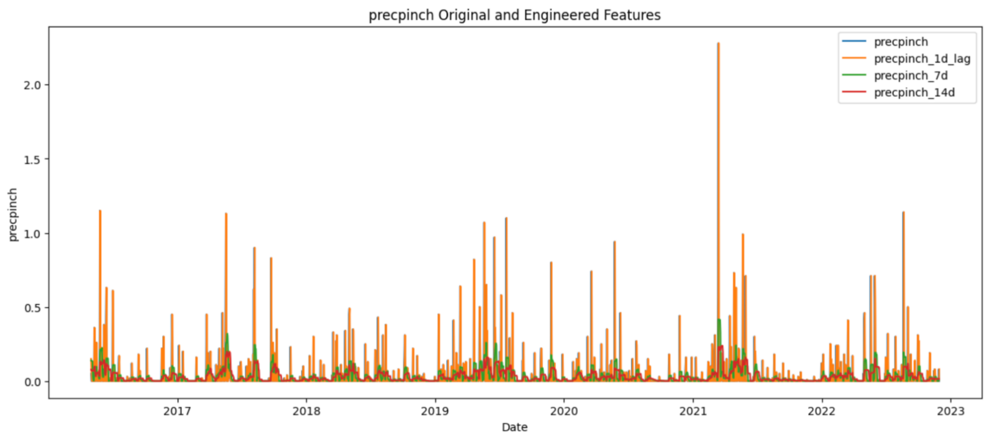

Figure 6: Precipitation data for Muddy Creek

2.2 Modeling

2.2.1 Data Collection

Climate data was collected from the NOAA website and the ground water level via the USGS with the help of APIs. The data is then cleaned and transformed to have useful features; datetime, average temperature, average precipitation and average ground water level at daily level. The data availability for both the sites is shown in Table 1 and 2.

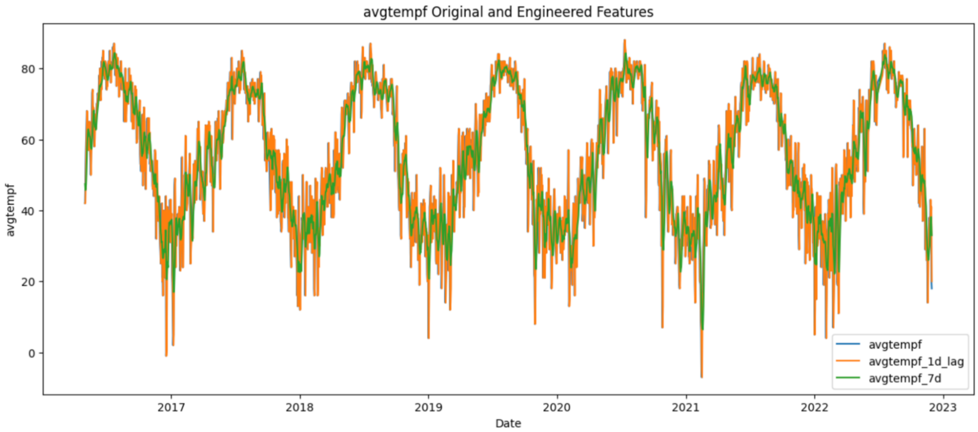

Figure 7: Temperature data for Pueblo

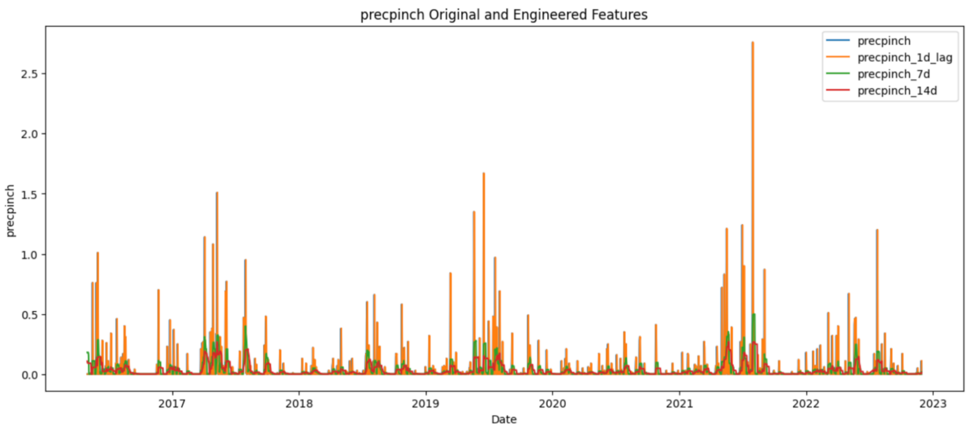

Figure 8: Precipitation data for Pueblo

Table 1: Data availability overview: Muddy Creek and Pueblo respectively

Feature	Min.Date	Max. Date	Feature	Min.Date	Max. Date
Temperature	2015-01-01	2022-12-31	Temperature	2015-01-01	2022-12-31
Precipitation	2015-01-01	2022-12-31	Precipitation	2015-01-01	2022-12-31
Ground water level	2016-05-01	2022-11-30	Ground water level	2016-05-01	2022-11-30

2.2.2 Feature Engineering

Data being time-series is passed through a feature engineering pipeline to generate additional engineered features suitable for analysis and modeling phase. The additional engineered features are; past 7 days rolling mean for temperature and precipitation, past 14 days rolling mean for temperature and precipitation and 1 day lagged data for temperature and precipitation. These features were then scaled using MinMax scaler to ensure that the features are on some unit scale. These new generated engineered features were then passed on to the modeling pipeline.

2.2.3 Model Building

The engineered data is split into 2 parts; training data and validation data with stratified splitting considering 80:20 as the ratio. The validation data is used to ensure the model is not overfitting along with finding out the optimal sets of hyperparameters.The modeling approach was built on the existing findings and the model structure in the Germany based similar analysis[2]. Convolution Neural Net (CNN)[10] are usually used for image identification but some studies show that they perform really good for time-series analysis also[11]. Two CNNs were created separately for Muddy Creek and Pueblo due to the difference in the range of ground water level in both wells along with the nature of both wells. CNN model used in this analysis is a 1D convolutional neural network with fixed kernel size of 2, with 32 filter size for both models (Muddy Creek and Pueblo). This was followed by a Max Pooling 1D layer and a fixed 20% dropout layer to prevent overfitting. This is followed by a dense layer with 16 and 32 neurons respectively for both models with leaky

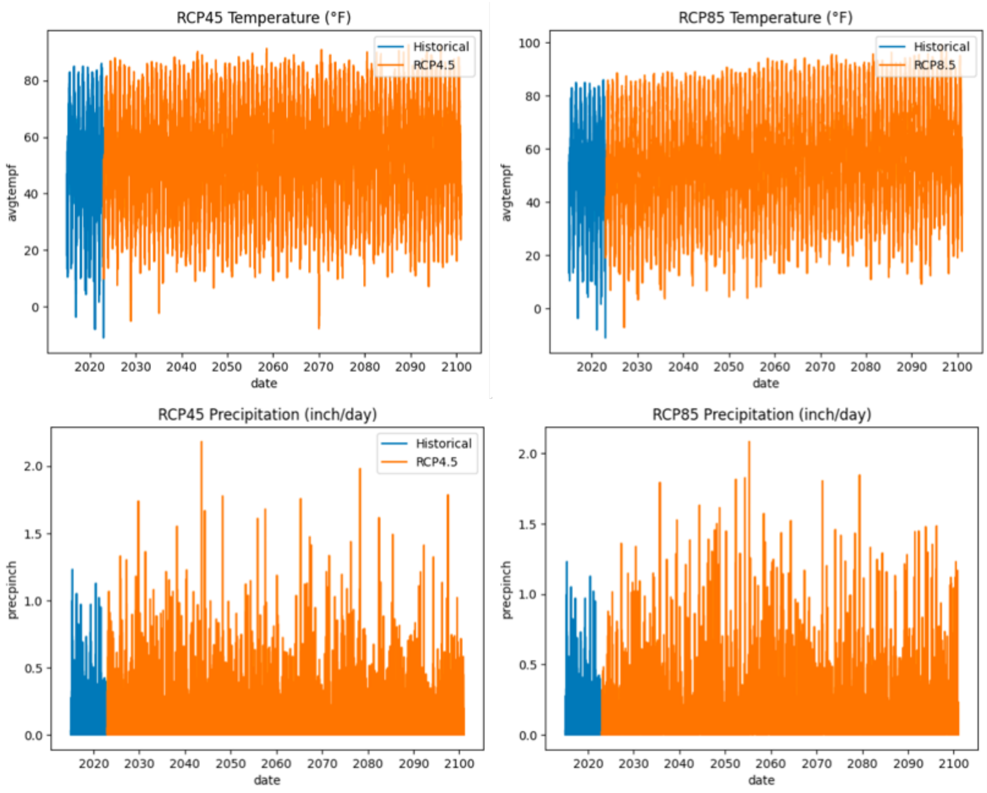

Figure 9: Combined snapshot of both historical & modeled CMIP data

rectified unit layer as the activation function, followed by a single output neuron layer. To overcome the exploding gradient issue, Adam optimiser is used along with a learning rate of 0.001. Early stopping is also applied as a regularization to overcome overfitting. All above are the final values of the hyperparameters after multiple iterations and analyzing the training vs validation data loss curve generated by considering mean squared error (MSE) as loss function.

2.2.4 Model Testing

Final models were tested by fresh 2023 data. The new data is passed through the same set of feature engineering and scaling pipeline used in training. Interestingly, both model perform better on H2 2023 data than compared to H1 with mean absolute percentage error being 2.8% and 2.87% respectively for Muddy Creek and Pueblo based models.

2.2.5 Muddy Creek based model

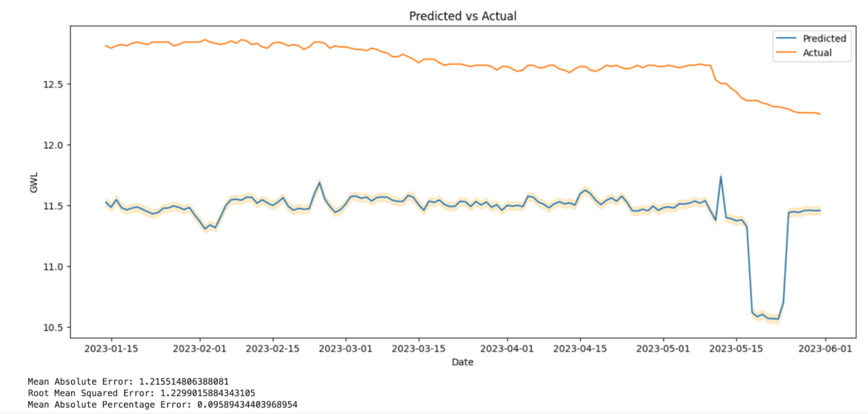

Figure 10: H1 2023 test: Muddy Creek

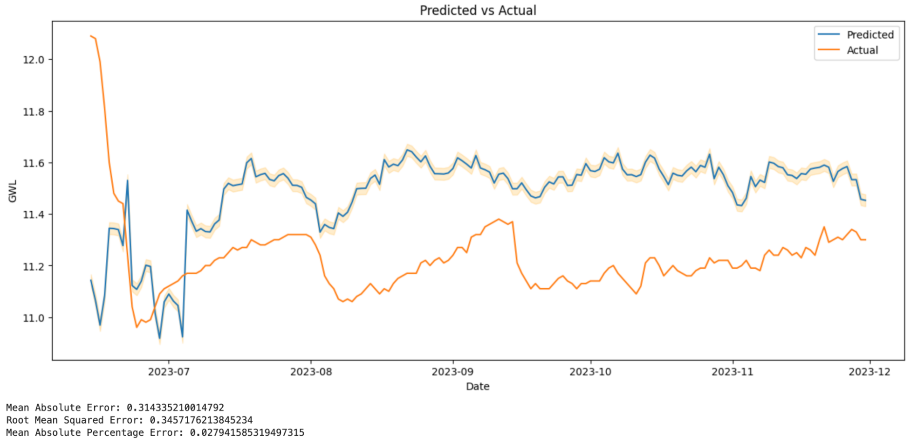

Figure 11: H2 2023 test: Muddy Creek

2.2.6 Pueblo based model

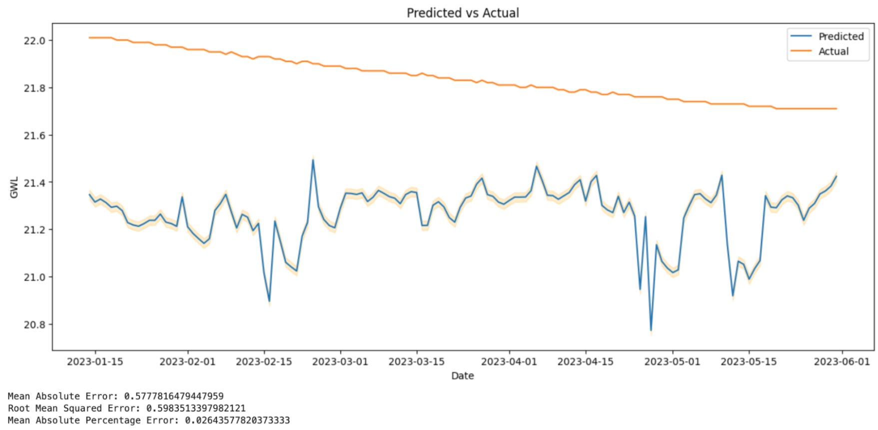

Figure 12: H1 2023 test: Pueblo

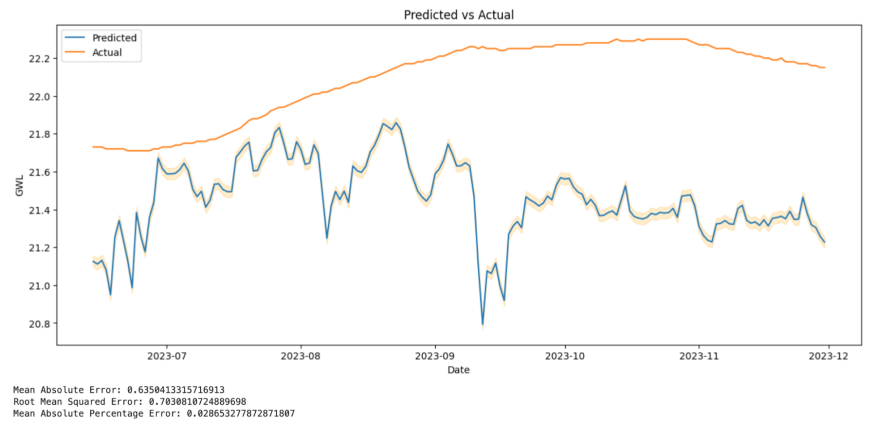

Figure 13: H2 2023 test: Pueblo

3 Results

Colorado specific data is filtered from the CORDEX climate projection data for 2024 and beyond, which is used as input to the models (Muddy Creek and Pueblo) for projecting the ground water levels under RCP8.5 scenarios as in Figure x and y respectively (red dotted line shows the average ground water level over the entire projection).

Ground water level in Pueblo as in Figure 3 has not dropped below 21 historically, but the projections show drop up to 20.25 in upcoming years and the water level is projected to be decreased below the current levels under the RCP8.5 (which is the worst case scenario). While for Muddy Creek, the projections shows decrease in average daily ground water level under RCP8.5 scenarios with a major decrease in somewhere between 2065-70

3.0.1 Muddy Creek based projections

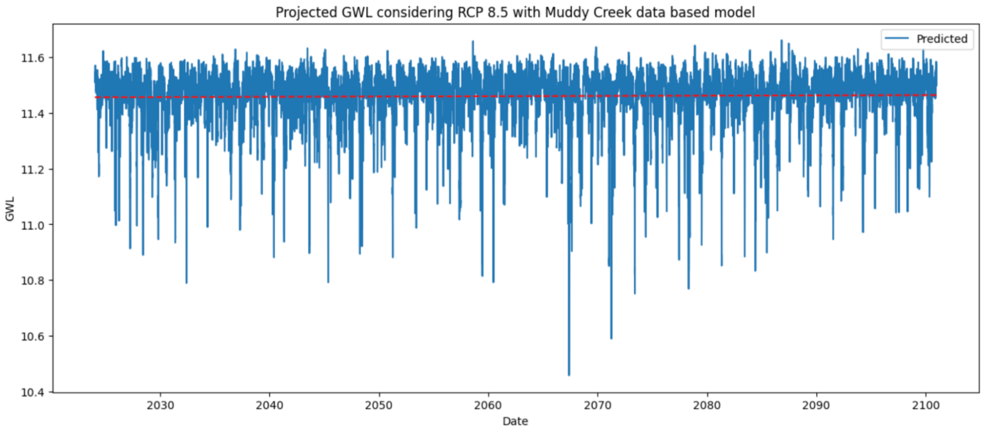

Figure 14: H2 2023 test: Pueblo

3.0.2 Pueblo based projections

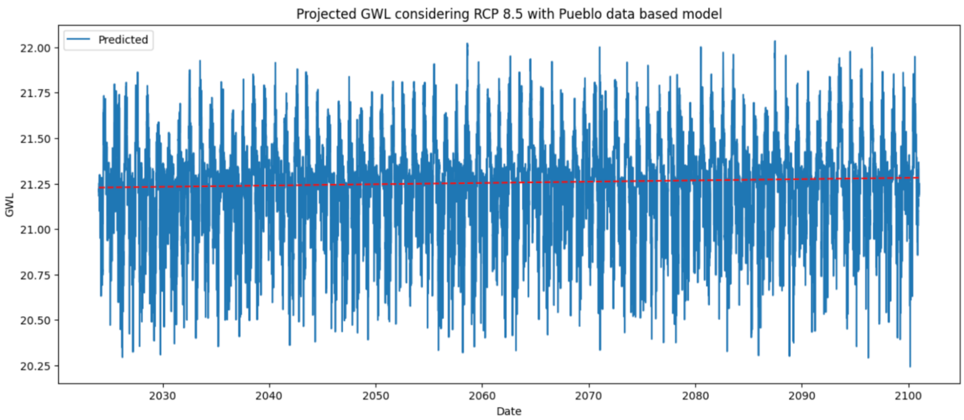

Figure 15: H2 2023 test: Pueblo

4 Limitations

The model is using 2015-22 data for training and validation and projection for close to 7.5 decades which might cause bias. The CORDEX data is adjusted for the biases, but still it doesn't capture real-time elements. Projections up to 2030 should be considered more accurate, and to make it further robust the model needs to be updated with new training data over years and then use it for projections. Another aspect that needs to be addressed is, consideration of various factors like the type of well, withdrawal rate from well, region where well lies, other climate data points like greenhouse gasses etc and their interconnections with the analysis.

5 Code Availability and Video Availability

All the relevant codes are available in this GitHub repository and here is the link for the video explaining the project

6 References

1. Jay, A.K., A.R. Crimmins, C.W. Avery, T.A. Dahl, R.S. Dodder, B.D. Hamlington, A. Lustig, K. Marvel, P.A. Méndez-Lazaro, M.S. Osler, A. Terando, E.S. Weeks, and A. Zycherman, 2023: Ch. 1. Overview: Understanding risks, impacts, and responses. In: Fifth National Climate Assessment. Crimmins, A.R., C.W. Avery, D.R. Easterling, K.E. Kunkel, B.C. Stewart, and T.K. Maycock, Eds. U.S. Global Change Research Program, Washington, DC, USA. https://doi.org/10.7930/NCA5.2023.CH1.

2. Wunsch, Andreas, Tanja Liesch, and Stefan Broda. "Deep Learning Shows Declining Groundwater Levels in Germany until 2100 Due to Climate Change." Nature Communications 13, no. 1 (March 9, 2022): 1221. https://doi.org/10.1038/s41467-022-28770-2.

3. "The Economics of Climate Change — Swiss Re," April 22, 2021. https://www.swissre.com/institute/research/topics-and-risk-dialogues/climate-and-natural-catastrophe-risk/expertise-publication-economics-of-climate-change.html.

4. "CMIP." Accessed December 13, 2023. https://www.wcrp-climate.org/wgcm-cmip.

5. Karl E. Taylor, Martin Juckes, V. Balaji, Luca Cinquini, Sébastien Denvil, Paul J. Durack, Mark Elkington, Eric Guilyardi, Slava Kharin, Michael Lautenschlager, Bryan Lawrence, Denis Nadeau, and Martina Stockhause. Google Docs. "CMIP6_global_attributes_filenames_CVs." Accessed December 13, 2023. https://docs.google.com/document/d/1h0r8RZr_f3-8egBMMh7aqLwy3snpD6_MrDz1q8n5XUk/edit?usp=sharing&usp=embed_facebook.

6. Mearns, L.O., et al., 2017: The NA-CORDEX dataset, version 1.0. NCAR Climate Data Gateway, Boulder CO, accessed December 13, 2023. https://doi.org/10.5065/D6SJ1JCH

7. "EURO-CORDEX." Accessed December 13, 2023. https://www.euro-cordex.net/index.php.en.

8. IPCC, 2014: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 pp.

9. Dehghani, Reza, Hassan Torabi Poudeh, and Zohre Izadi. "The Effect of Climate Change on Groundwater Level and Its Prediction Using Modern Meta-Heuristic Model." Groundwater for Sustainable Development 16 (February 1, 2022): 100702. https://doi.org/10.1016/j.gsd.2021.100702.

10. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. Nature 521, 436–444 (2015)

11. Kløve, B. et al. Climate change impacts on groundwater and dependent ecosystems. J. Hydrol. 518, 250–266 (2014)