Bias Adjustments Based on Climate and Temporal Resolution for a Precipitation Intensity Factor Used in Soil Erosion Modeling

Andrew Fullhart[1], Mark Nearing[1], Mark Weltz[2]

[1]USDA-ARS Southwest Watershed Research Center, [2]USDA-ARS Great Basin Rangelands Research; Corresponding Author Email: Andrew.Fullhart@usda.gov

Overview

BACKGROUND: Intensity is critical in modeling. Erosion rate may quadruple when rainfall intensity is doubled (Meyer, 1981). Furthermore, peak runoff is approximately proportional to intensity when the entirety of a watershed is contributing to flow (Nash, Halliwell and Cox, 2002).

PROBLEM: Calculating rainfall intensity during a storm event requires <u>sub-hourly</u> precipitation data, but widely available datasets are <u>hourly and daily</u>.

METHOD: Create predictive statistical models for intensity using regression techniques that take lower resolution data including hourly and daily data.

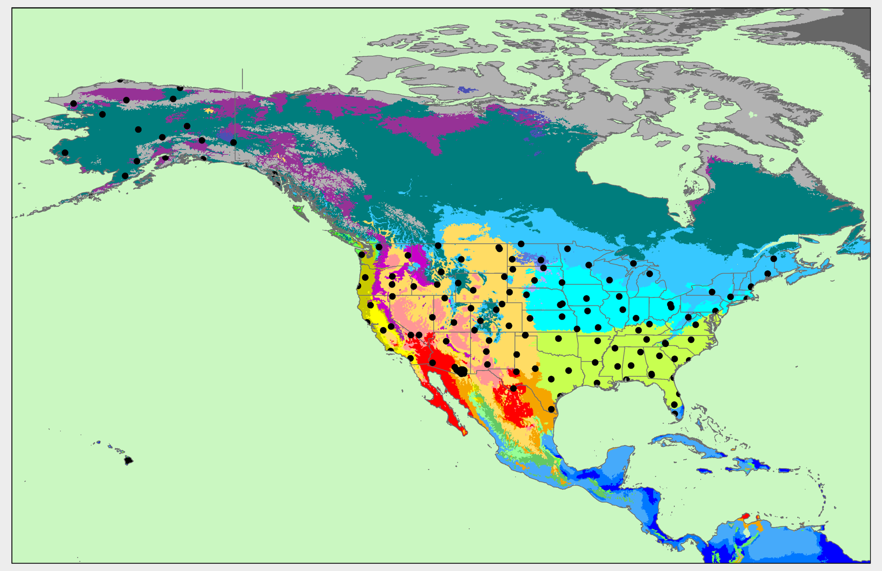
RESULT: Intensity can be estimated using information from lower resolution datasets and geographical location including climate type.

Precipitation and Climate Data

The precipitation data used in this study comes from NOAA's Climate Reference Network (CRN).

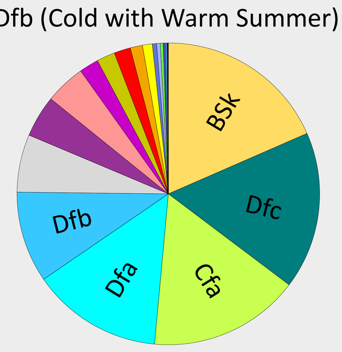
- 140 stations distributed across the United States.
- 5-min resolution.
- 11 year record durations on average.

Köppen-Geiger climate classification is used to group stations and as a categorical variable for regressions in which multiple predictor variables are considered (Beck et al, 2018). Köppen-Geiger climate classification often aligns with differences in weather patterns and vegetation biomes.



75% of U.S. Surface Area is Comprised of 5 Climate Types:

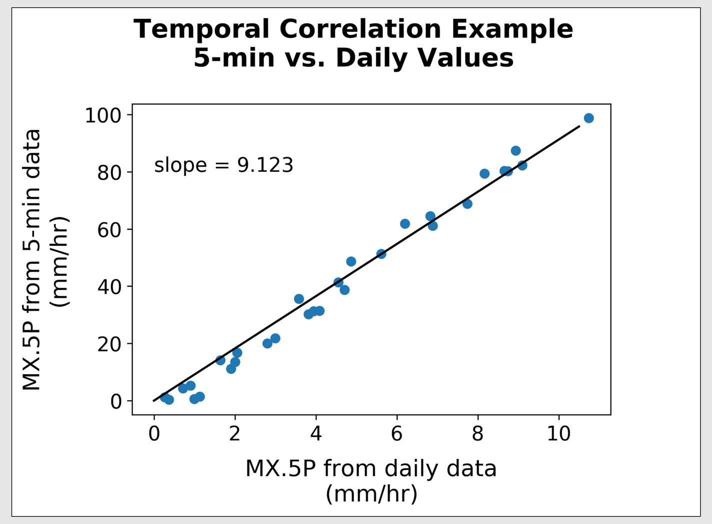
BSk (Semi-Arid Cold)
Dfc (Cold with Cold Summer)
Cfa (Temperate, Wet, Hot Summer)
Dfa (Cold with Hot Summer)
Dfb (Cold with Warm Summer)



Methodology

- An Intensity factor that considers maximum 30-min precipitation (I_{30}) is analyzed. The factor is known as **MX.5P**, which is the mean monthly maximum I_{30}. Each station has 12 MX.5P values.
- MX.5P is biased by temporal resolution when I_{30} values are determined by time-averaging intensity given the accumulation of a measurement interval. This necessitates temporal correlations.
- Multiple regressions and machine learning regressions use the additional predictor variables below:

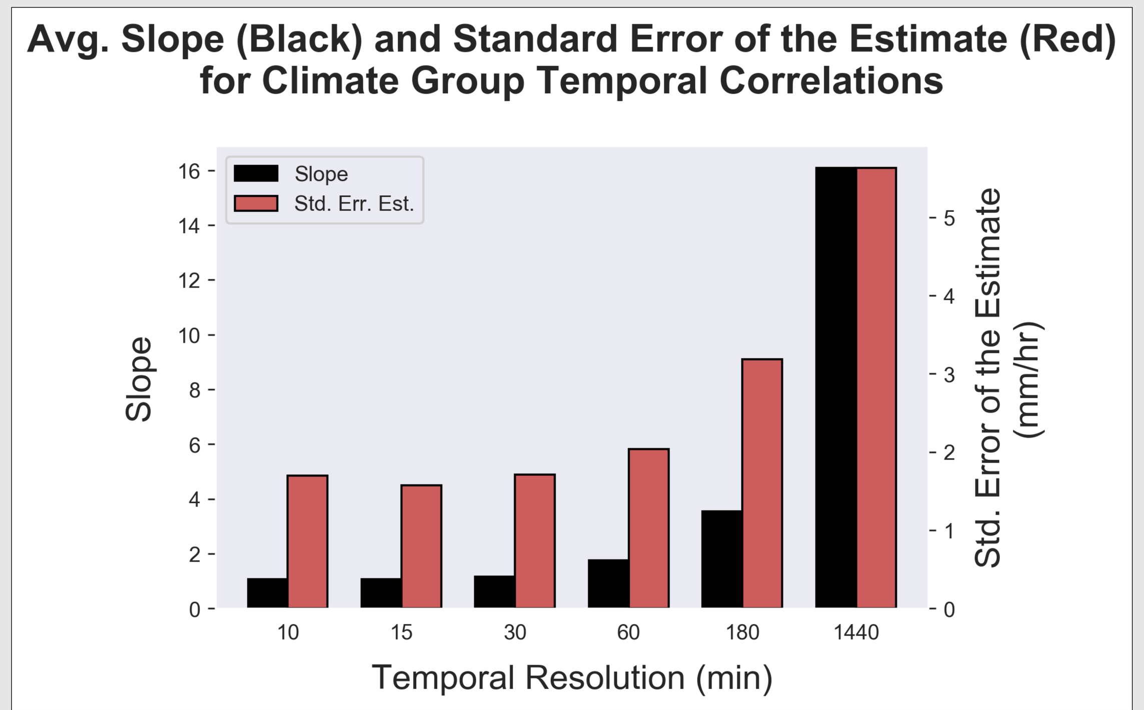
Variable	Label	Unit	Values per station
Monthly mean maximum 30-min intensity	MX.5P	mm/hr	12
Modified Fournier index	Fournier Coeff	mm	1
Average daily rainfall for wet days in the month	MEAN P	mm	12
Standard deviation for daily rainfall for wet days in the month	S DEV P	mm	12
Skewness for daily rainfall for wet days in the month	SKEW P	-	12
Monthly transition probability of a wet day given a wet day	P(W/W)	-	12
Monthly transition probability of a wet day given a dry day	P(W/D)	-	12
Station elevation	Elev	m	1
Station latitude	Lat	deg.	1
Station coastal proximity	Coastal Prox	km	1
Calendar month (categorical variable)	Month	-	12
Köppen climate (categorical variable)	Climate	-	1



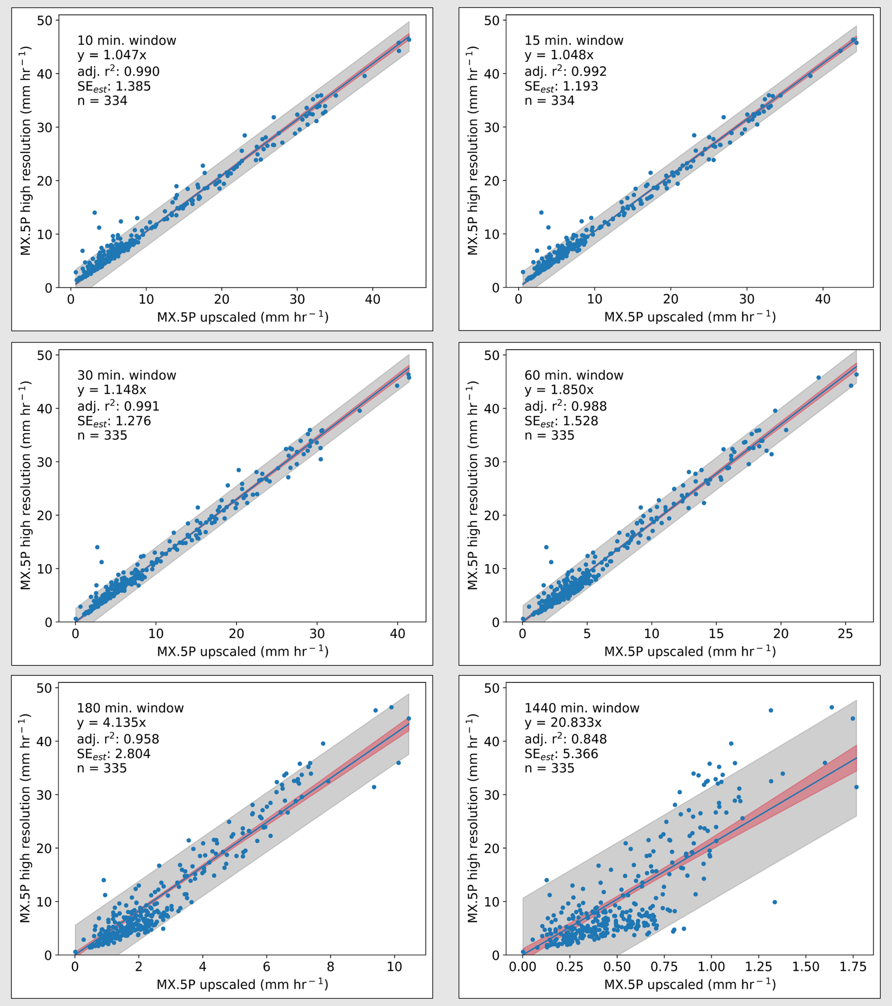
Hypothetical Temporal Correlation Example: The MX.5P values derived from daily data are underestimated by a factor of 9.123. The slope can be used as an adjustment factor for daily-derived values. In the real example below, variance is higher.

Single Predictor Results

- Grouping by climate reduced the standard error of the estimate in half for daily-derived MX.5P values.
- Still, error is too large for daily-derived values.
- When MX.5P datapoints are grouped by climate, error for hourly-derived MX.5P values was 2.04 mm/hr on average but 5.63 mm/hr for daily-derived values.

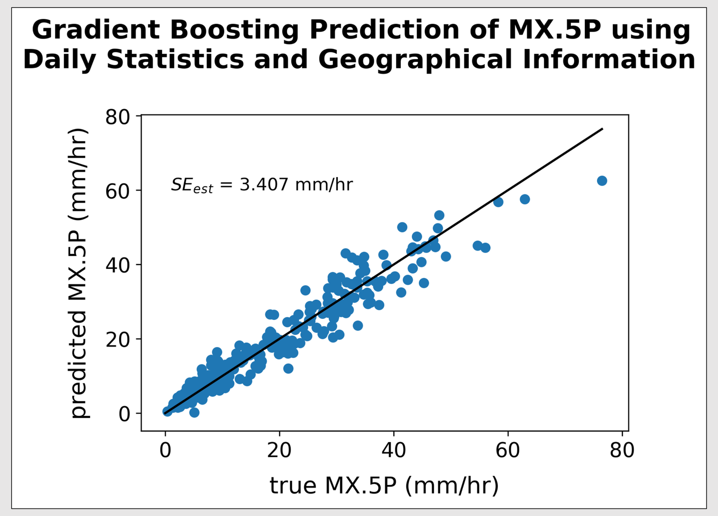


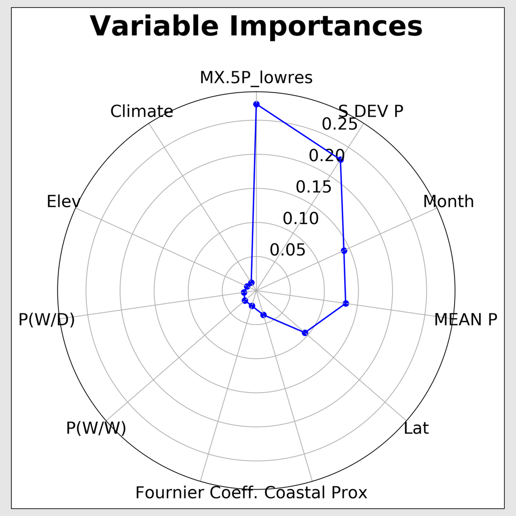
BSk climate correlations for 10, 15, 30, 60, 180 and 1440-min relative to 5-min values.



Multiple Predictor Results

- The additional predictor variables may be determined from daily data and geographical information, allowing them to be used to predict MX.5P values at any of the resolutions.
- The data is pooled to ensure a large sample size, and climate is used as a categorical variable.
- Random Forest and Gradient Boosting regressions gave comparable results and reduced error by half relative to Multiple Regression.
- For the daily correlation, error is reduced to 3.41 mm/hr and variance is considerably more homogeneous.





Conclusions

A stochastic weather generator, CLIGEN, is capable of using MX.5P and information from daily data to produce tp-ip data (time-to-peak, peak intensity) needed for modeling. The ability to estimate MX.5P from daily data means that daily data is the only requirement for production of a tp-ip dataset for modeling.

The MX.5P factor is harder to estimate from lower resolution precipitation data.

Use of additional predictor variables and machine learning regressions improves the daily temporal correlations.

This project is made possible by a USDA-ARS Headquarters Grant.

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future Köppen-Geiger climate classification maps at 1-km resolution. Scientific data, 5, 180214.

Meyer, L. D. (1981). How rain intensity affects interrill erosion. *Transactions of the ASAE*, *24*(6), 1472-1475.

Nash, D., Halliwell, D., & Cox, J. (2002). Chapter 10: Hydrological Mobilization of Pollutants at the Field/Slope Scale. *Agriculture, hydrology and water quality*, 225.