

University of Texas Rio Grande Valley

ScholarWorks @ UTRGV

---

Civil Engineering Faculty Publications and  
Presentations

College of Engineering and Computer Science

---

6-11-2019

## Incorporating Antecedent Soil Moisture into Streamflow Forecasting

Abdoul Oubeidillah

*The University of Texas Rio Grande Valley*

Glenn Tootle

Thomas Piechota

Follow this and additional works at: [https://scholarworks.utrgv.edu/ce\\_fac](https://scholarworks.utrgv.edu/ce_fac)



Part of the [Civil Engineering Commons](#), and the [Earth Sciences Commons](#)

---

### Recommended Citation

Oubeidillah, A., Tootle, G., & Piechota, T. (2019). Incorporating Antecedent Soil Moisture into Streamflow Forecasting. *Hydrology*, 6(2), 50. <https://doi.org/10.3390/hydrology6020050>

This Article is brought to you for free and open access by the College of Engineering and Computer Science at ScholarWorks @ UTRGV. It has been accepted for inclusion in Civil Engineering Faculty Publications and Presentations by an authorized administrator of ScholarWorks @ UTRGV. For more information, please contact [justin.white@utrgv.edu](mailto:justin.white@utrgv.edu), [william.flores01@utrgv.edu](mailto:william.flores01@utrgv.edu).

Article

# Incorporating Antecedent Soil Moisture into Streamflow Forecasting

Abdoul Oubeidillah <sup>1,\*</sup>, Glenn Tootle <sup>2</sup> and Thomas Piechota <sup>3</sup>

<sup>1</sup> Department of Civil Engineering, University of Texas-Rio Grande Valley, Edinburg, TX 78539, USA

<sup>2</sup> Department of Civil, Construction and Environmental Engineering, University of Alabama, Tuscaloosa, AL 35487, USA; gatootle@eng.ua.edu

<sup>3</sup> Office of Research, Chapman University, Orange, CA 92866, USA; piechota@chapman.edu

\* Correspondence: abdoul.oubeidillah@utrgv.edu; Tel.: +1-865-243-0342

Received: 22 February 2019; Accepted: 4 June 2019; Published: 11 June 2019



**Abstract:** This study incorporates antecedent (preceding) soil moisture into forecasting streamflow volumes within the North Platte River Basin, Colorado/Wyoming (USA). The incorporation of antecedent soil moisture accounts for infiltration and can improve streamflow predictions. Current Natural Resource Conservation Service (NRCS) forecasting methods are replicated, and a comparison is drawn between current NRCS forecasts and proposed forecasting methods using antecedent soil moisture. Current predictors used by the NRCS in regression-based streamflow forecasting include precipitation, streamflow persistence (previous season streamflow volume) and snow water equivalent (SWE) from SNOTEL (snow telemetry) sites. Proposed methods utilize antecedent soil moisture as a predictor variable in addition to the predictors noted above. A decision system was used to segregate data based on antecedent soil moisture conditions (e.g., dry, wet or normal). Principal Components Analysis and Stepwise Linear Regression were applied to generate streamflow forecasts, and numerous statistics were determined to measure forecast skill. The results show that when incorporating antecedent soil moisture, the “poor” forecasts (i.e., years in which the NRCS forecast differed greatly from the observed value) were improved, while the overall forecast skill remains unchanged. The research presented shows the need to increase the monitoring and collection of soil moisture data in mountainous western U.S. watersheds, as this parameter results in improved forecast skill.

**Keywords:** streamflow; antecedent soil moisture; snowpack

## 1. Introduction

Streamflow forecasting is the estimation of seasonal volumes of water at a specific site (gauge) at a specific time. In mountainous regions of the western U.S., the season of interest is the spring-summer season when natural supply levels decrease and demand increases due to seasonal influences. The Natural Resource Conservation Service (NRCS), in cooperation with the National Weather Service (NWS), issue water supply (or streamflow) forecasts for over 750 points in the western U.S., near the first of the month between January and June each year. These forecasts assist water managers/users for future planning according to the forecasted amount of water available. While these forecasts are produced monthly, this study focuses on forecasting the cumulative April-May-June-July streamflow volume.

Water managers operate with a smaller margin of error, increasingly complex and competing demands, and try to retain flexibility to adapt to hydro-climatic conditions [1]. The primary objective of these forecasts is to minimize any risk and uncertainty for water managers and to create a more efficient use of a scarce resource.

Thorough understanding of forecast performance helps decision makers determine when and how much to rely upon forecasts, as well as how to respond to expected climatic anomalies. Over-allocated supplies and increasing demands require efficient management of water resources. While the NRCS has been forecasting water supplies for close to 80 years, it is evident that the physical and demographic landscapes of the western U.S. are changing. Reference [2] reveals how streamflow forecasts can be more effectively used if scientists look at it from the user's perspective.

While the NWS uses a comprehensive set of models and hydrologic techniques, NRCS forecasts are produced using statistical approaches such as multiple linear regression models. These regression-based forecasts rely on measurements of current snowpacks, antecedent streamflow and autumn precipitation that simplify complicated hydrologic processes [3]. The regression models suggest a relationship between predictor variables (precipitation, snow water equivalent, antecedent streamflow, etc.) and the predicted (streamflow volume of interest). Several techniques were developed by Garen [4] to significantly improve forecast accuracy when using regression models. These techniques include: (1) Basing the regression model only on data known at the forecast time (no future data); (2) principal components regression; (3) cross validation; and (4) systematic searching for optimal or near-optimal combinations of variables [4]. Historical practice in forecasting often included variables in regression equations that described future precipitation amounts. Previous research [4–6] proved that use of future variables (variables that describe future snow accumulation or precipitation) and the substitution of averages reduced forecasting accuracy, especially early in the forecasting season. The research presented here does not use future variables, but only variables known at the time of the forecasting process.

Prior to developing a forecast model, it is vital to analyze predictors. This includes creating and maintaining a high quality historical dataset, subjected to rigorous screening and data quality testing [7]. As stated by Garen [4], "A more robust, accurate and consistent forecasting equation can be obtained by having several sites for the same data type and time in the equation." Currently, the predictors obtained for NRCS streamflow forecast models are gathered from remote sensing data sources. Due to the relative newness of these remote sensing sites, the period of record used by the NRCS to develop a streamflow forecast is relatively short.

The motivation of this research evolved through discussions with NRCS regarding the forecasting of streamflow in the Upper North Platte River Basin. First, the NRCS stated that "The Upper North Platte River Basin was one of the more challenging regions to forecast in the western United States." An additional question posed was "Is there a way to increase forecast skill for years in which current NRCS forecasts result in a 'poor' forecast (i.e., NRCS forecast is much different than actual) while maintaining overall model skill?" The challenge posed to researchers is to achieve these two objectives (maintain overall model skill and improve "poor" year forecasts) while, still being constrained to using current NRCS forecasting methods (principal components stepwise linear regression).

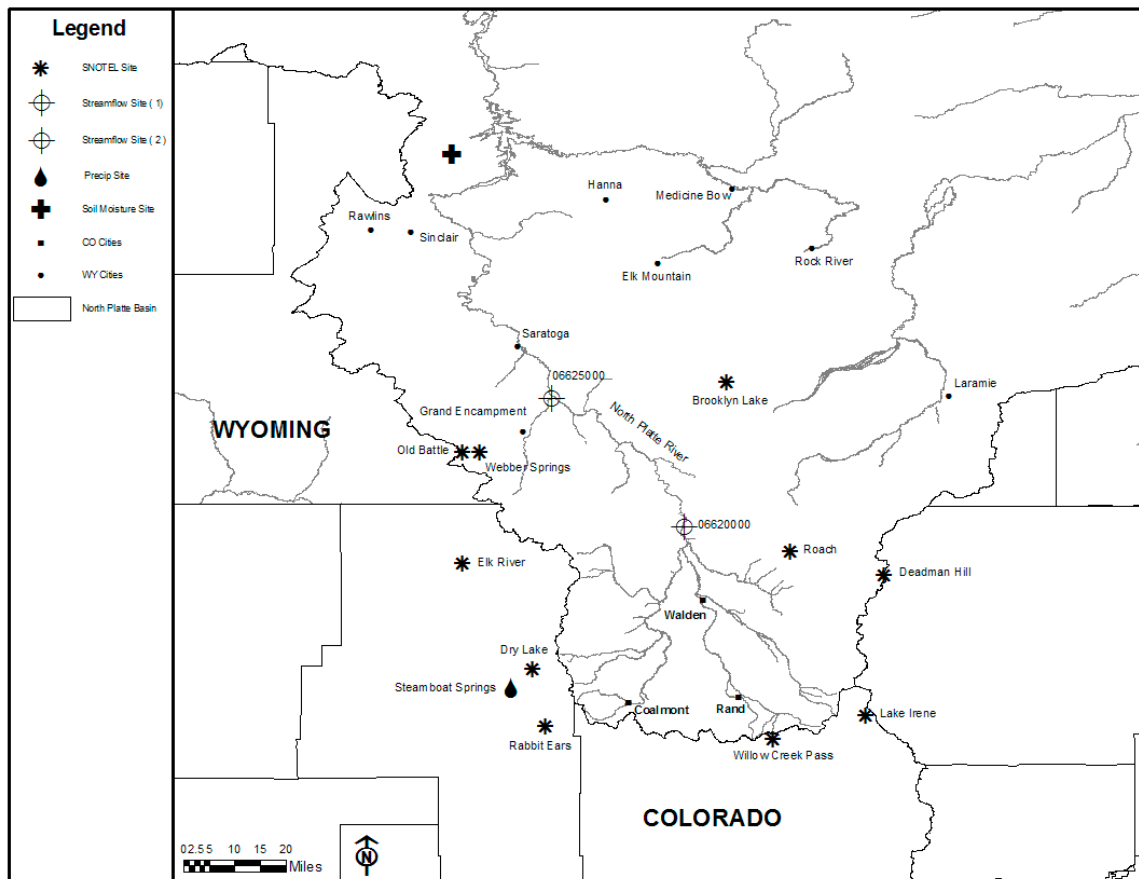
The purpose of this study is to propose an improved streamflow forecast model using antecedent soil moisture. This is done through the use of traditional predictors (snow water equivalent, precipitation, and antecedent streamflow) currently used in streamflow forecast models. In addition, through the incorporation of the National Oceanic and Atmospheric Administration (NOAA) climate division soil moisture data. Past research has looked at the importance of soil moisture in streamflow forecasting [8–14]. In addition to incorporating (adding) antecedent soil moisture (ASM) as a predictor, this research proposes a novel approach in the development of a decision system based upon ASM. This decision system is based on segregating ASM data into three specific categories: Wet, normal, and dry. Each category has its own regression equation, utilizing current NRCS methods (principal components stepwise linear regression).

The contribution of this research is the identification of a valuable predictor (ASM) and a new framework (decision system based on ASM) for improving poor NRCS streamflow forecasts while maintaining overall model skill in the North Platte River Basin. The results support the need to increase monitoring and the collection of soil moisture data, especially in mountainous western U.S. watersheds, in which the snowpack is the primary driver of streamflow runoff.

The collection of soil moisture data will ultimately provide a useful database to improve streamflow forecasts in these regions.

## 2. Watershed Description

The North Platte River is a tributary of the Platte River, which is approximately 1094 km in length. The North Platte River originates in Colorado where it flows north into Wyoming, and then flows east to Nebraska (Figure 1). The three major reservoirs in Wyoming along the North Platte River are Seminoe, Pathfinder, and Glendo. Present use and future development of water resources in the North Platte River Basin are controlled by the 1945 Supreme Court Decree for the North Platte River.



**Figure 1.** Location map of the Wyoming North Platte River Basin.

The North Platte River watershed is predominately located in mountainous regions of Colorado and Wyoming. Most of the annual streamflow is attributed to melting snowpack that has accumulated during the winter and early spring months in mountainous headwater regions. Reference [3] suggests that snowmelt provides approximately 80 percent of the streamflow in the western United States. The delay between the snow accumulation and snowmelt creates the opportunity to generate an estimate of the actual amount of runoff.

## 3. Data

Available datasets used to forecast streamflow include antecedent streamflow (streamflow persistence), snow water equivalent, precipitation and ASM.

### 3.1. Streamflow Data

The data used in this study comes from two streamflow stations (USGS 06620000 and USGS 06625000), which are located the Upper North Platte River Basin. The data was obtained from the United States Geological Survey (USGS) NWIS website [15]. Each of these stations is unimpaired [16], and a current forecast is provided by the NRCS. USGS streamflow station 06620000 is the most upstream (southern) station. The station's elevation is 2380 m above sea level, and has a drainage area of 3706 square kilometers. USGS station 06625000 is located on a downstream tributary (Encampment River) and is 2124 m above sea level with a drainage area of 686 square kilometers. See Figure 1 for a detailed location map covering the region of study. The USGS provides daily, monthly, and annual mean streamflow in cubic feet per second (cfs). Total monthly streamflow in cubic meters for April-May-June-July (AMJJ) was calculated using appropriate conversions. Antecedent (January-February) streamflow volume, a commonly used predictor in NRCS forecasts, was also utilized.

### 3.2. Snow Water Equivalent (SWE) Data

The NRCS National Water and Climate Center provides snow water equivalent data (in inches) for the western United States [17]. Snow water equivalent data was distinguished into 2 groups—snow course & SNOTEL. Early SWE data (snow course) was recorded manually, and SNOTEL data is published in real-time through the use of remote sensing stations. Snow course data in the western U.S. dates as far back as 1906, and SNOTEL data in the North Platte River Basin dates back to the early 1970s, depending upon when the digital sensors in the station were installed. Within and adjacent to the North Platte River Basin are a total of ten SWE stations (Figure 1). These stations provide accumulated precipitation, snow depth, snow water equivalent, and temperature data. April 1st SWE (converted from inches to centimeters) was used as a predictor in the current research.

### 3.3. Precipitation Data

Precipitation data was available from the Western Regional Climate Center (WRCC) website [18] which publishes monthly precipitation totals. One precipitation station has data dating back to the year of interest (1940) for the proposed analysis. This station is located in Steamboat Springs, Colorado and has monthly data dating back to 1908. Average precipitation data was obtained for the Steamboat Springs, Colorado station (converted from inches to centimeters) for the period of October through December of the previous year and January through March of the forecasted year.

### 3.4. Antecedent Soil Moisture Data

Data for soil moisture was obtained from the National Oceanic and Atmospheric Administration (NOAA) website [19]. NOAA soil moisture data is estimated by a one-layer hydrological model [20,21]. The model takes observed precipitation and temperature and calculates soil moisture, evaporation and runoff. The soil moisture is modeled data, so there is one value for the entire climate division. Of the 344 climate divisions in the U.S., one soil moisture dataset is within the North Platte River Basin. This station is located in climate division 10 within the state of Wyoming, and the data covers an area of 61.6 square kilometers (Figure 1). Accessible monthly soil moisture data is available from 1932 to 2005 (74 years). Average ASM for this station in mm (for the period of October through December of the previous year and January through March of the forecasted year) was obtained.

The authors acknowledge that NOAA climate division soil moisture contains many uncertainties; however, the primary hypothesis of this research is that ASM is a useful predictor for streamflow forecasting.

## 4. Forecast Methodology

### 4.1. Current NRCS Methods

The NRCS has developed a Visual Interactive Prediction and Estimation Routine (VIPER) to forecast streamflow. This forecast application gathers all data on a monthly to seasonal time scale in real-time directly from the source. Linked with both historical and real-time data, the hydrologist specifies a list of predictor sites for a specified streamflow gage, the type of analysis desired (principal components stepwise linear regression is the most common NRCS method to forecast streamflow in the North Platte River Basin), and equations are automatically developed and the forecast produced in real-time. The methodology used for these forecasts and the current research replicates these methods. This real-time approach is very efficient, but has the disadvantage of using only data that has been recorded by digital sensors (i.e., limited record). Data of this type varies in relation to when the SNOTEL site was installed. The first of these SNOTEL sites in the North Platte River Basin was established in the early 1970s, which limits the digital data that can be used in producing forecasts. For varying streamflow stations, this period of record varies depending upon the available SNOTEL data. Within the VIPER interface, various types of streamflow transformations can be applied to improve forecast skill. Current NRCS forecasts do not incorporate antecedent soil moisture as a predictor.

### 4.2. Identifying Predictors

Initially, predictors (e.g., SWE stations) were identified by correlating the yearly April 1 SWE values with yearly seasonal (AMJJ) streamflow. Next, moving time (10, 20, 25, and 30 years) window correlations between streamflow and SWE were performed [22]. This ensures that reliable and consistent SWE data sets are used (i.e., stability throughout the record), given the uncertainties (e.g., prolonged equipment malfunction, equipment calibration, human error) in the collection of SWE data for various periods of record. Finally, correlation values between snow course/streamflow and SNOTEL/streamflow are analyzed. This will (or will not) confirm that the relationship between snow course/streamflow is similar to the relationship between SNOTEL/streamflow. A minimum difference between snow course/streamflow and SNOTEL/streamflow correlation values is essential because stability throughout the period of record is needed to utilize data in the early 1940s.

A visual inspection of the streamflow and SWE correlations resulted in the following “rules” for the inclusion of the SWE station as a predictor. First, the overall correlation value between SWE and streamflow must exceed 99% significance to be included as a predictor. Second, if any of the moving time window correlations resulted in a negative value, the SWE station was not included. Finally, when comparing snow course/streamflow and SNOTEL/streamflow, the correlation values must exceed 95% significance in each comparison (while maintaining an overall 99% significance per above).

Precipitation records dating back to the period of interest were limited in the North Platte River Basin. One location (Steamboat Springs, Colorado) provided precipitation data dating back to the early 1940s. Finally, streamflow persistence (JF streamflow) is also correlated with AMJJ streamflow.

After determining the most appropriate predictor variables, the same methodology used by NRCS (principal components stepwise linear regression) was performed to account for potential cross correlation. For this initial analysis, ASM was not included because this evaluation was designed to replicate current NRCS forecast methodology.

### 4.3. Applying Current NRCS Methods Incorporating ASM

The next analysis adds ASM as a predictor into the principle components stepwise linear regression analysis. The forecast timeline is kept consistent and previous predictor variables identified are not changed, and the same forecast methodology was used. This process determines if incorporation of ASM as a predictor results in improved streamflow forecast skill.

#### 4.4. Decision System Incorporating ASM

The final analysis in this research segregates predictor variables (i.e., decision system) based solely on ASM (e.g., wet, dry, normal). This requires performing a simple statistical analysis to determine the average and standard deviation of ASM data for the season and record of interest. Wet years are defined as those whose soil moisture is 1.25 standard deviations ( $\sigma$ ) above average, and similarly, dry years are defined as those whose soil moisture is 1.25 $\sigma$  below the overall average. The 1.25 $\sigma$  (above and below) the overall average corresponds approximately to the 90th and 10th quartiles. Remaining years are considered normal years. After grouping the data into the appropriate categories using ASM conditions (wet, normal, dry), principle components stepwise linear regression was performed individually on the three sets of data (wet, normal, dry). This analysis results in three separate regression equations that are used to forecast streamflow based on ASM conditions.

#### 4.5. "Poor" NRCS Forecasts

A "poor" streamflow forecast was one that predicts a streamflow volume that was much different than the actual (observed) volume. For this research, a "poor" forecast was determined by ranking (worst to best) each year and selecting the upper quartile (25%) of worst forecasts and defining them as "poor" forecasts. There are a variety of reasons that lead to a "poor" streamflow forecast for a particular year. They include: Data unavailability, unexpected precipitation, unforeseen drought conditions and climate change. Any one of these or a combination of them were possible reasons a "poor" forecast ended up being produced. Implications of "poor" streamflow forecasting include: Inefficient management/allocation of water, water managers having little confidence in forecasts, and reduced credibility of the forecaster. The hypothesis of this research was that the incorporation of ASM into a decision system framework will result in improved streamflow forecasts for "poor" years while maintaining overall model skill. The physical basis of this hypothesis was that ground surface conditions (wet, dry, normal) will influence the amount of runoff. The same snowpack and precipitation for wet ground surface conditions will produce more runoff than dry ground surface conditions.

#### 4.6. Statistical Analysis

In forecasting, problems with intercorrelation arise when predictor variables are highly correlated with other predictor variables. For example, antecedent streamflow correlates highly with precipitation and snow water equivalent. The most statistically rigorous way to deal with intercorrelation, and the method applied in this study, is to use principle components regression [4]. Principle components regression is a useful technique for addressing multicollinearity problems, and can yield better predictors [23]. An important property of the principle components was that they were uncorrelated [24]. Thus, there were no problems with multicollinearity. The number of components retained in the equation depends upon how many of the components have statistically significant regression coefficients. It was also necessary to determine which principal components to use in the regression equation [4]. Reference [4] used a standard t-test to determine the significance of the regression coefficient for the component. A similar method, presented in this study, was the use of forward stepwise linear regression to determine the number of principle components to include in the regression model. Forward stepwise regression determines which of the predictors explain a significant amount of the variance, starting with the predictor that explains the most variance, while adding/removing any predictors that do/do not significantly improve the fit.

For this study, a stepwise linear regression F-value of 4 is used. For a two-sided test with  $\alpha = 0.05$  (95% significance) and sample sizes of 20 or more, the critical value for the standard t-test is close to 2. Squaring this t-value produces a critical partial F-value near 4.

Numerous predictive statistics can be calculated to determine the skill of a principle component regression model. These include the standard error of the regression,  $R^2$ , adjusted  $R^2$ , the PRESS statistic, and the predicted  $R^2$ . The standard error of the regression (S) was used to describe model fit,

and was equivalent to the square root of the mean squared error.  $S$  represents the cumulative distance between the data and the fitted regression line. Thus, a lower value of  $S$  indicates better prediction of the response from the fitted regression equation.  $R^2$  was a function of  $S$  that was scaled to be between 0 and 1. Thus,  $R^2$  measures the proportion of variation in the response that is accounted for by the predictor variables. A higher  $R^2$  indicates a better fit of the model to the data.

Adjusted  $R^2$  also describes the variation of the response variable due to the relationship between the response variable and one or more predictor variables. The relationship was adjusted based upon the number of predictors in the model.  $R^2$  values will always increase when a new predictor is added to the model. However, using adjusted  $R^2$  prevents the model from appearing better due to adding marginally important predictor terms.

It is well known that the prediction ability of the model as measured by the previous criteria can provide an overly optimistic measure of the true forecasting performance [4]. In order to achieve a closer representation of forecasting ability, cross validation procedures were recommended. Cross validation creates a validation series by dropping observations corresponding to the years, creating a regression equation for the remaining observations, and then predicting values for those years that were dropped. The PRESS (prediction sum of squares) statistic was such a measure of the predictive ability of the model. PRESS was based upon a leave-one-out cross-validation in which a single year or observation was removed when fitting the model. As a result, the prediction errors are independent of the predicted value at the removed observation [4]. For selecting a model when the primary interest was in prediction (forecasting), the model with the smaller PRESS was preferable [25]. The PRESS value was also used to calculate the predicted  $R^2$  statistic, which was an “ $R^2$ -like” statistic that reflects the prediction capability of the model [26]. Thus, predicted  $R^2$  ranges from 0 to 1.0. PRESS was on the same scale as the residual sum of squares (squared units).

Another method to measure forecast skill was the linear error in probability (LEPS) score [27,28]. The LEPS score was originally developed to assess the position of the forecast and the position of the observed values in the cumulative probability distribution. Reference [28] describes the advantages of the LEPS score over traditional skill measurements such as root-mean-square error. The LEPS score ( $S''$ ) and the average skill (SK) are defined in reference [29]. A LEPS SK score of greater than +10% is generally considered “good skill”. The LEPS SK score has been previously utilized as a measure of skill in streamflow forecast models [30–32].

## 5. Results

### 5.1. Identifying Predictors

Of the nine SWE stations within the North Platte River Basin, four were selected based on the “rules” established in the Methods section. The correlation coefficient between precipitation (October to March) at Steamboat Springs and streamflow (AMJJ volume) for stations 06620000 and 06625000 exceeded 99% significance, respectively, and both displayed stability throughout the extended period of record. This shows that the precipitation records from Steamboat Springs were adequate, and will be used in the study. Based on correlation values and current NRCS methodologies, streamflow persistence was used as a predictor for USGS 06620000, but not as a predictor for USGS 06625000. As mentioned earlier, streamflow data was first recorded in this region in the early 1940s. For USGS streamflow station 06620000, the forecasted period was 1940–2005 (66 years inclusive) while 1941–2005 (65 years inclusive) was the forecasted period used for USGS station 06625000.

### 5.2. Comparison of Current NRCS Methods with and without ASM

A comparison was made between current NRCS methods and current NRCS methods incorporating ASM as a predictor. Table 1a,b shows the numerical measures of skill. For both streamflow stations, there was a slight increase in overall forecast skill when incorporating soil moisture. The LEPS SK scores for USGS 06620000 were 63.5 without soil moisture, and 64.6 incorporating ASM.



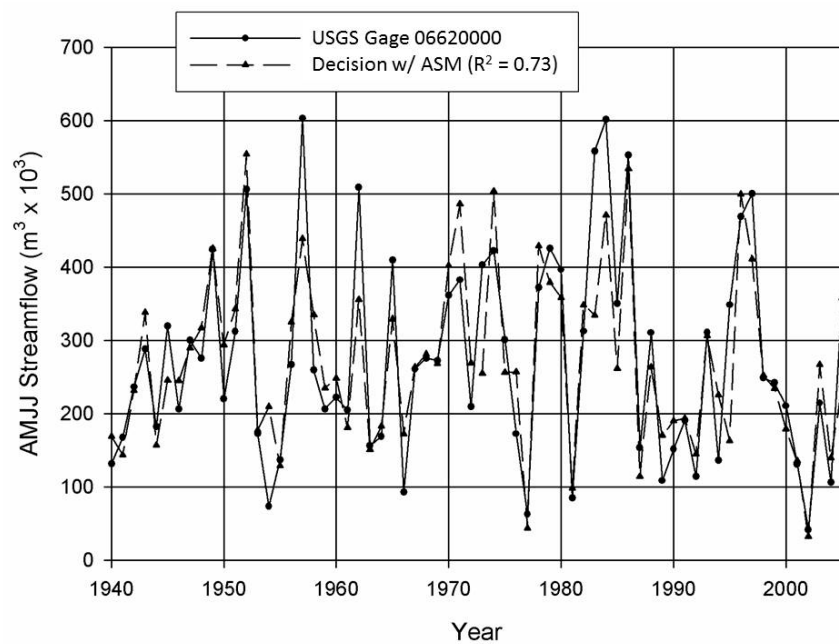
LEPS SK scores were 66.0 (without ASM) and 66.9 (with ASM) for USGS 06625000. While the authors acknowledge the increases in skill were small, all forecasts produce a higher  $R^2$ , predicted  $R^2$ , adjusted  $R^2$ , and LEPS SK values when incorporating ASM. More importantly, the decrease in PRESS values show that the models incorporating ASM were preferable. Finally, it is important to note that the ASM data used in this research is modeled data that represents an entire climate division (i.e., a large spatial area). This ASM data was most likely a poor reflection of upper watershed soil moisture. The collection of soil moisture data, from improved land-based equipment or satellites, spatially upstream from the streamflow station(s) will most likely result in even greater improvement in overall model skill.

**Table 1.** Prediction and validation statistics for (a) USGS 06620000 (b) USGS 06625000.

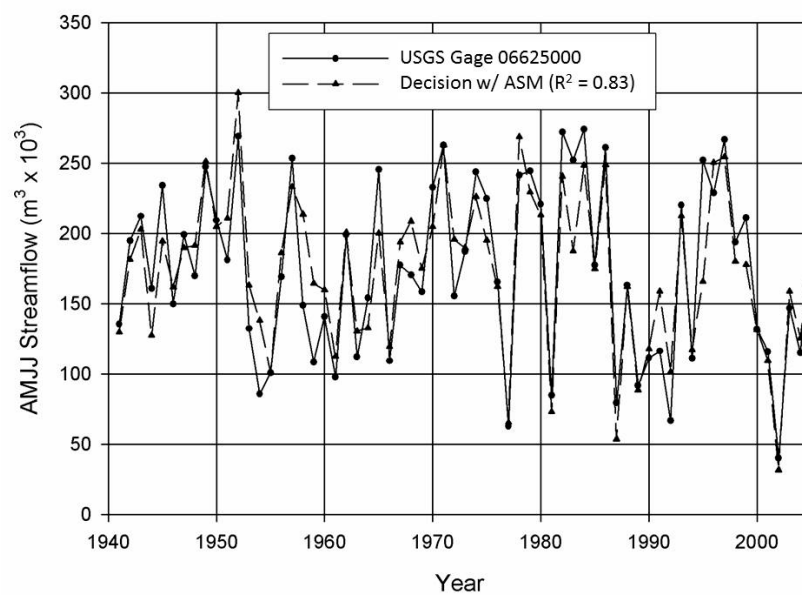
(a)			
	USGS 06620000 w/o ASM	USGS 06620000 w/ASM	USGS 06620000 Decision System
Period of Record	1940–2005 (66)	1940–2005 (66)	1940–2005 (66)
$R^2$	0.67	0.69	0.73
$R^2(\text{adj})$	0.67	0.69	0.72
$R^2(\text{pred})$	0.65	0.67	0.70
PRESS	286,191	269,145	248,396
S	65.0	63.5	59.4
$S'''$	41.9	42.6	43.3
SK	63.5	64.6	65.5
(b)			
	USGS 06625000 w/o ASM	USGS 06625000 w/ASM	USGS 06625000 Decision System
Period of Record	1941–2005 (65)	1941–2005 (65)	1941–2005 (65)
$R^2$	0.73	0.77	0.81
$R^2(\text{adj})$	0.72	0.76	0.80
$R^2(\text{pred})$	0.70	0.74	0.78
PRESS	48,715	42,312	36,253
S	26.5	24.6	22.4
$S'''$	42.8	43.5	45.7
SK	65.9	66.9	70.2

### 5.3. Decision System Incorporating Antecedent Soil Moisture

ASM was used in a decision system in which three individual forecasts are developed for wet, dry and normal conditions (Figure 2a,b). The number of years for each category are as follows: USGS streamflow station 06620000 (7 dry years, 50 normal years, 9 wet years); USGS streamflow station 06625000 (7 dry years, 49 normal years, 9 wet years). The decision system maintains overall forecast skill when compared to applying NRCS methods (with and without incorporating ASM) for both USGS streamflow stations. The decision system LEPS SK scores are 65.5 and 70.2 for USGS 06620000 and USGS 06625000, respectively. The decision system model was the preferred one based on the PRESS statistic. Additional statistics are displayed in Table 1a,b.



(a)



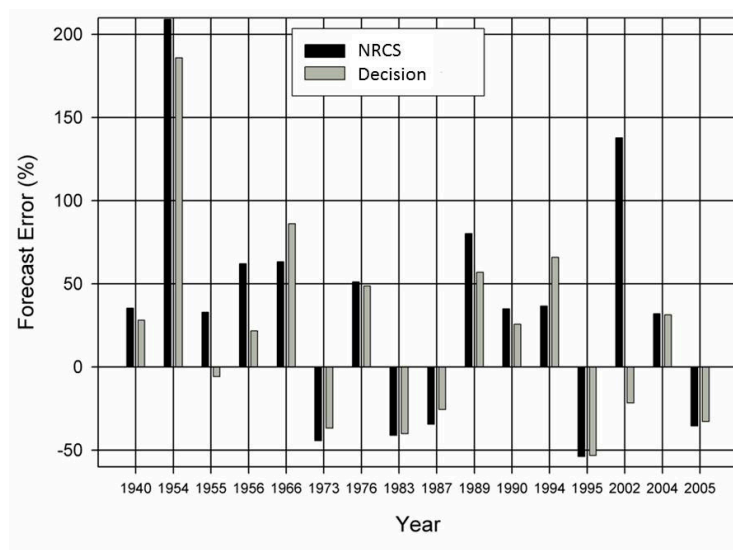
(b)

**Figure 2.** Plot of Gage Record and Forecast (Decision System with ASM) for (a) USGS 06620000 (b) USGS 06625000.

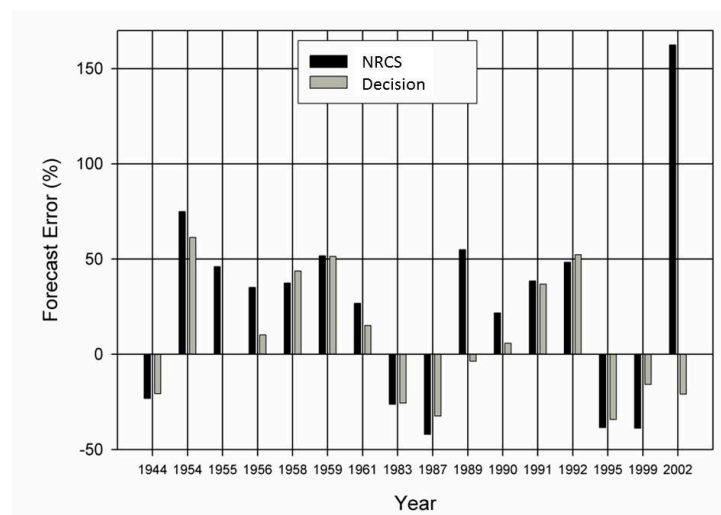
#### 5.4. Improving Poor Forecasts

As previously mentioned, the motivation of this research was to improve “poor” forecasts. To accomplish this, NRCS has sought improvement in the modeling strategy that will maintain the overall skill of the model while specifically providing a better prediction of poor forecasts, or those years in which the model provided an extremely poor prediction, as previously defined. Slight to moderate increase in measures of overall skill have been previously demonstrated (Section 5.3). A “poor” forecast was determined by ranking (worst to best) each year and selecting the upper quartile (25%) of worst forecasts and defining them as “poor” forecasts.

These poor forecasts are examined for the period of record at both streamflow stations. Incorporating ASM with the decision system approach resulted in considerable improvement in forecasting “poor” years. An improved forecast was achieved for 14 out of the 16 “poor” years for both USGS streamflow stations (Figure 3a,b) when compared to current NRCS methods. For the “poor” forecast years, the overall average error for USGS 06620000 was reduced from 35.4% (when using existing NRCS methods) to 20.9% (when using the ASM based decision system). For USGS 06625000, the overall average error for the “poor” forecast years was reduced from 26.8% to 7.7%. Additionally, when applying correlation techniques and the student’s t-test, significance levels increased from 95% for the NRCS forecast to 99% for the ASM decision system forecast for both streamflow stations. The current research resulted in a significant increase in forecast skill for the “poor” years. When comparing the NRCS forecast to the ASM decision system forecast, the LEPS score for USGS 06620000 was almost unchanged (41.2% to 39.0%). However, the LEPS score for USGS 0662500 increased from 34.6% to 56.6%.



(a)



(b)

**Figure 3.** Plot comparing forecasts for the worst quartile of “poor” forecasts using current forecasting methods (NRCS) with proposed methods (decision system) for (a) USGS 06620000 (b) USGS 06625000.

Interestingly, six out of the 16 “poor” forecast years were considered extreme (wet or dry) for USGS 06620000, and five of 16 were extreme based on ASM conditions for USGS 06625000. Therefore, since the majority of NRCS poor forecast years were considered to be normal based on the decision system approach, the process of removing extreme years (both dry and wet) results in improving “normal year” forecast of “poor” years.

## 6. Conclusions & Future Work

The incorporation of ASM into NRCS streamflow forecasting models in the North Platte River Basin achieved both goals set forth by the NRCS. First, the overall model forecast skill was maintained. The model that incorporates ASM was shown to be a preferred model with a slight improvement in skill. Second, a notable improvement was observed when incorporating the new approach using ASM for “poor” forecasts. The extended period of record resulted in 28 out of 32 “poor” forecasts being improved. The development of a decision system, based on ASM, is a novel approach that reveals the importance of ASM in streamflow forecasting. While this research provides a basis to consider ASM in streamflow forecasting, there is a considerable void in soil moisture data availability, both in the length of record and the accuracy/precision of the data. Enhanced SNOTEL stations in the North Platte River Basin, that measure soil moisture digitally and in real-time, may result in further increased forecast skill when incorporated as predictors. Inexpensive instrumentation, such as conductivity devices and tensiometers can also provide soil moisture data. For both calibrated conductivity-based devices and well-maintained tensiometers, the user can expect measurement accuracies of up to 90 to 95 percent [33]. Further research may also incorporate NASA MODIS snow cover data in addition to SNOTEL data. MODIS data will reflect a spatial coverage of snowpack in the basin versus current (SNOTEL) point data. However, one limitation of using MODIS technology includes the limited availability of data and digital images. With the increased importance of producing accurate streamflow forecasts, more soil moisture models (and arguably more accurate) are being created. Incorporating this soil moisture data from various available models and instrumentation may prove to be an important predictor in future streamflow forecasts.

**Author Contributions:** A.O. led research efforts in data acquisition, analysis, and development of the manuscript, G.T. assisted in research efforts, providing mentoring and serving as Lead PI on both research awards (see Funding). T.P. provided mentoring and project Quality Control.

**Funding:** This research was funded by the University of Wyoming Water Research Program, the United States Geological Survey (USGS), the Wyoming Water Development Commission, and the University of Wyoming. Additional funding was provided by the National Science Foundation Paleo Perspectives for Climate Change program award AGS-1003393.

**Acknowledgments:** The authors wish to thank Greg Kerr, Larry Pochop of the University of Wyoming; Cody Moser of the University of Wyoming and of the University of Tennessee; and Barry Lawrence of the Wyoming Water Development Commission for their assistance and support of this research.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

1. Pagano, T.C.; Pasteris, P.; Dettinger, M.; Cayan, D.; Redmond, K.T. Water year 2004: Western water managers feel the heat. *EOS Trans. Am. Geophys. Union* **2004**, *85*, 385–400. [[CrossRef](#)]
2. Hartmann, H.C.; Pagano, T.C.; Sorooshian, S.; Bales, R. Confidence builders: Evaluating seasonal climate forecasts from user perspectives. *Bull. Am. Meteorol. Soc.* **2002**, *83*, 683–698. [[CrossRef](#)]
3. Pagano, T.C.; Garen, D.C. Climate variations, climate change, and water resources engineering. In *Integration of Climate Information and Forecasts into Western US Water Supply Forecasts*; ASCE Publications: Reston, VA, USA, 2006; pp. 86–102.
4. Garen, D.C. Improved techniques in regression-based streamflow volume forecasting. *J. Water Resour. Plan. Manag.* **1992**, *118*, 654–670. [[CrossRef](#)]

5. Stedinger, J.R.; Grygier, J.; Yin, H. Seasonal streamflow forecasts based upon regression. In *Computerized Decision Support Systems for Water Managers, Proceedings of the 3rd Water Resources Operations and Management Workshop, Fort Collins, CO, USA, 27–30 June 1988*; ASCE: New York, NY, USA, 1988; pp. 266–279.
6. Koch, R.W. *Influences of Climate Variability on Streamflow Variability: Implications in Streamflow Prediction and Forecasting*; Final Report for Grant Award 14-08-0001-G1316; U.S. Geological Survey: Washington, DC, USA, 1990.
7. Pagano, T.C.; Erxleben, J.; Perkins, T. Operational simulation modeling at the NRCS National Water and Climate Center. In *Proceedings of the Western Snow Conference, Great Falls, MT, USA, 11–14 April 2005*; pp. 87–100.
8. Day, N.G. Extended streamflow forecasting using NWSRFS. *J. Water Resour. Plan. Manag.* **1985**, *111*, 157–170. [[CrossRef](#)]
9. Aubert, D.; Loumagne, C.; Oudin, L. Sequential assimilation of soil moisture and streamflow data in a conceptual rainfall-runoff model. *J. Hydrol.* **2003**, *280*, 145–161. [[CrossRef](#)]
10. Berghuijs, W.R.; Woods, R.A.; Hutton, C.J.; Sivapalan, M. Dominant flood generating mechanisms across the United States. *Geophys. Res. Lett.* **2016**, *43*, 4382–4390. [[CrossRef](#)]
11. Chen, X.; Kumar, M.; McGlynn, B.L. Variations in streamflow response to large hurricane-season storms in a Southeastern US watershed. *J. Hydrometeorol.* **2015**, *16*, 55–69. [[CrossRef](#)]
12. Silvestro, F.; Rebora, N. Impact of precipitation forecast uncertainties and initial soil moisture conditions on a probabilistic flood forecasting chain. *J. Hydrol.* **2014**, *519A*, 1052–1067. [[CrossRef](#)]
13. Pathiraja, S.; Westra, S.; Sharma, A. Why continuous simulation? The role of antecedent moisture in design flood estimation. *Water Resour. Res.* **2012**, *48*, W06534. [[CrossRef](#)]
14. Berthet, L.; Andréassian, V.; Perrin, C.; Javelle, P. How crucial is it to account for the antecedent moisture conditions in flood forecasting? Comparison of event-based and continuous approaches on 178 catchments. *Hydrol. Earth Syst. Sci.* **2009**, *13*, 819–831. [[CrossRef](#)]
15. U.S Geological Survey. National Water Information System Data. 2019. Available online: <http://waterdata.usgs.gov/nwis/rt> (accessed on 10 January 2018).
16. Wallis, J.R.; Lettenmaier, D.P.; Wood, E.F. A daily hydroclimatical data set for the continental United States. *Water Resour. Res.* **1991**, *27*, 1657–1663. [[CrossRef](#)]
17. USDA Natural Resources Conservation Service. National Water and Climate Center Data. 2019. Available online: <http://www.wcc.nrcs.usda.gov/snow/> (accessed on 10 February 2018).
18. Western Regional Climate Center. Recent Climate in the West. Temperature and Precipitation Data. 2019. Available online: <http://www.wrcc.dri.edu> (accessed on 21 January 2018).
19. NOAA National Weather Service Climate Prediction Center. Soil Moisture Monitoring Data. 2019. Available online: [http://www.cpc.ncep.noaa.gov/products/Soilmst\\_Monitoring](http://www.cpc.ncep.noaa.gov/products/Soilmst_Monitoring) (accessed on 19 February 2018).
20. Huang, J.; van den Dool, H.; Georgakakos, K.P. Analysis of model-calculated soil moisture over the United States (1931–93) and application to long-range temperature forecasts. *J. Clim.* **1996**, *9*, 1350–1362. [[CrossRef](#)]
21. Van den Dool, H.; Huang, J.; Fan, Y. Performance and analysis of the constructed analogue method applied to U.S. soil moisture over 1981–2001. *J. Geophys. Res.* **2003**, *108*, 8617. [[CrossRef](#)]
22. Biondi, F.; Waikul, K. DENDROCLIM2002: A C++ program for statistical calibration of climate signals in tree-ring chronologies. *Comput. Geosci.* **2004**, *30*, 303–311. [[CrossRef](#)]
23. Khattree, R.; Naik, S. *Multivariate Data Reduction and Discrimination with SAS Software*; SAS Institute Inc.: Cary, NC, USA, 2000; pp. 60–61.
24. Anderson, T.W. *An Introduction to Multivariate Statistical Analysis*, 3rd ed.; John Wiley & Sons: Hoboken, NJ, USA, 2003; pp. 461–462.
25. Montgomery, D.C.; Peck, E.A.; Vining, G.G.; Myers, R.H. *Introduction to Linear Regression Analysis*, 4th ed.; John Wiley and Sons: Hoboken, NJ, USA, 2006; pp. 141–142.
26. Myers, R.H. *Classical and Modern Regression with Applications*, 2nd ed.; Duxbury Press: Belmont, CA, USA, 1990; p. 171.
27. Ward, N.M.; Folland, K.K. Prediction of seasonal rainfall in the North Nordeste of Brazil using eigenvectors of sea-surface temperature. *Int. J. Climatol.* **1991**, *11*, 711–743. [[CrossRef](#)]
28. Potts, J.M.; Folland, C.K.; Jolliffe, I.T.; Sexton, D. Revised ‘LEPS’ scores for assessing climate model simulations and long-range forecasts. *J. Clim.* **1996**, *9*, 34–53. [[CrossRef](#)]

29. Tootle, G.A.; Singh, A.K.; Piechota, T.C.; Farnham, I. Long lead-time forecasting of U.S. streamflow using partial least squares regression. *Am. Soc. Civ. Eng. J. Hydrol. Eng.* **2007**, *12*, 442–451. [[CrossRef](#)]
30. Piechota, T.C.; Chiew, F.H.S.; Dracup, J.A.; McMahon, T.A. Seasonal streamflow forecasting in eastern Australia and the El Niño—Southern Oscillation. *Water Resour. Res.* **1998**, *34*, 3035–3044. [[CrossRef](#)]
31. Piechota, T.C.; Dracup, J.A. Long range streamflow forecasting using El Niño–Southern Oscillation indicators. *J. Hydrol. Eng.* **1999**, *4*, 144–151. [[CrossRef](#)]
32. Tootle, G.A.; Piechota, T.C. Suwannee River long range streamflow forecasts based on seasonal climate predictors. *J. Am. Water Resour. Assoc.* **2004**, *40*, 523–532. [[CrossRef](#)]
33. Murphy, T.A. Sensing the elements: Moisture sensors and weather stations. *Irrig. J.* **1996**. Available online: <http://www.sowacs.com/sensors/sensingaug.html> (accessed on 3 March 2018).



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).