FORECASTING THE EFFECTS OF SNOW DROUGHT ON STREAMFLOW VOLUMES IN THE WESTERN U.S.

Joshua T. Sturtevant[1] and Adrian A. Harpold[1]

ABSTRACT

Mountain snowpack supplies critical water resources to natural ecosystems and downstream populations, particularly in semi-arid regions such as the Western U.S. Extremely low snowpack, or snow drought, can arise from inter-annual climate variability but is worsened by long-term declines in snowpack. Snow drought has negative implications for water availability as earlier, slower, and smaller snowmelt fundamentally changes runoff patterns including runoff efficiency and groundwater recharge rates. We explored the implications of snow drought on streamflow in the Western U.S. through the investigation of historical (1980-2014) observational data from the U.S. Geological Survey and Natural Resource Conservation Service for 59 watersheds. Our primary focus was evaluating the accuracy of April-July streamflow volume forecasts using the statistical method of Principal Component Regression (PCR). Early results indicate that forecast errors are highest during snow drought years, but that these errors vary across basins and are affected by precipitation received after the forecast issue date. We also show that forecast errors have a strong positive correlation with the inter-annual variability of runoff efficiency. Our objective is to refine statistically-based forecasting methods by introducing non-linearities into the forecast equations or sub-setting the data into dry and wet years in order to improve streamflow volume forecasts during snow drought years. (KEYWORDS: snow drought, streamflow, water supply forecasting, climate change, NRCS.)

INTRODUCTION

Water supply forecasting provides valuable information about anticipated spring runoff to a variety of end users including farmers, ranchers, dam operators, and municipalities. Historically, water supply forecasts (WSFs) in the Western U.S. have relied on manual snow course measurements supplemented by automated measurements from Snow Telemetry (SNOTEL) stations. Though there are several Federal agencies that provide seasonal WSFs, the U.S. Department of Agriculture's Natural Resource Conservation Service (NRCS) is the longest standing such forecaster. The NRCS employs a tried and true statistical regression technique that uses monthly snow water equivalent (SWE), accumulated precipitation, and antecedent streamflow to forecast April-July streamflow volumes for nearly 900 forecast points across the Western U.S. The NRCS method differs from other agencies who often rely on a newer, more physically-based method called Ensemble Streamflow Prediction forecasting. While a perennially valuable decision-making tool for water users, the NRCS WSFs are arguably most valuable during times of extreme drought and low streamflow. Significant reductions in mountain snowpack (Mote et al., 2018), increased winter rainfall (Knowles et al., 2006), and decreases in streamflow linked to decreasing snowfall (Barnett et al., 2005; Berghuijs et al., 2014) highlight how climate change is driving smaller snowpacks, or snow droughts. Differentiated from meteorological drought, snow droughts can both be driven by a lack of precipitation (dry snow drought) or from abnormally warm temperatures leading to increased winter rainfall and early melt (warm snow drought) (Harpold et al., 2017a). In the snowmelt dominated water supply infrastructure of the West, there is considerable incentive to understand how snow droughts affect streamflow generation processes. Yet, the systematic study of these same physical mechanisms on WSF errors has been limited in scope and geographic extent (Mantua et al., 2008; Harpold et al., 2017b; Lehner et al., 2017), with no current research explicitly considering the role of snow drought on WSFs. Our approach represents a novel attempt to use a large-scale, empirical method to highlight patterns in WSF error, identify its driving factors, and consider possible WSF improvement techniques.

METHODS

Fifty-nine watersheds of interest were selected from the Catchment Attributes and Meteorology for Large-Sample Studies (CAMELS) dataset (Addor et al., 2017; Newman et al., 2015). CAMELS, a research grade hydrologic dataset for minimally impacted watersheds in the contiguous U.S., provides valuable basin-scale information including streamflow data, daily forcing data, and catchment attributes. Site selection criteria were

Paper presented Western Snow Conference 2019
[1]Joshua Sturtevant, University of Nevada, Reno, Reno, NV, jsturtevant@nevada.unr.edu
[1]Adrian Harpold, University of Nevada, Reno, Reno, NV, aharpold@cabnr.unr.edu

threefold: 1) the site was in the CAMELS database, 2) the site had at least one SNOTEL station, and 3) the site was an active NRCS WSF point. These criteria provided a subset of geographically and hydroclimatically representative watersheds for the Western U.S., spanning the states of Washington, Oregon, Idaho, California, Nevada, Montana, Wyoming, Utah, and New Mexico. Arizona and Colorado did not have any sites that met the criteria.

NRCS forecasts were recreated for 59 watersheds during 34 water years (WY 1981 to 2014) at four different lead times (Jan. 1st, Feb. 1st, Mar. 1st, and Apr. 1st) for a 50% exceedance April-July total streamflow volume, largely mimicking the NRCS WSF publishing format. WSFs were developed using the principal component regression (PCR) method as outlined by Garen (1992). The predictor variables in the principal component analysis (PCA) included the SNOTEL-measured SWE and accumulated precipitation, both for the day prior to the forecast date. In all cases, only the first principal component (PC1) was retained for the PCA. The PC1 was then linearly regressed against the U.S. Geological Survey (USGS) measured April-July streamflow volumes, i.e. hindcasting. Equation variables for each forecast were obtained from the actual NRCS forecasts, available on the National Water and Climate Center Air and Water Database. Data gaps for the predictor variables and/or the predictand led to the recreation of n=1,703 April 1 forecasts out of a possible n=2,006. Antecedent streamflow and snow course measurements were not included as predictor variables since not all WSF equations used these variables.

Predicted streamflow volumes from the PCR forecasts were compared to actual April-July streamflow volumes as measured by USGS stream gauges and compiled by the CAMELS dataset (Newman et al., 2015). Forecast error metrics were calculated for all forecasts at all lead times. Standard error metrics calculated included relative root mean squared error (relative RMSE, %), mean absolute error (MAE), mean absolute percent error (MAPE, %), and relative bias (%); only relative RMSE and relative bias are presented here. Normalized error metrics were used to standardize for mean streamflow volume. Error metrics were calculated for above and below average snow years, defined by the long-term mean peak SWE. This assured that sample size remained consistent between the two populations. The below-average snow years classification served as a generalization for snow drought; in practice, snow drought is often defined as <75% mean peak SWE (Harpold et al., 2017a).

RESULTS AND DISCUSSION

In considering the effects of snow drought on WSF skill, our results demonstrate that below average, or "dry", snow years (red line, Figure 1a) have forecast errors that are roughly one-third higher than during above average, or "wet", snow years (blue line, Figure 1a). Specifically, we found that during dry years, the January 1st (April 1st) relative RMSE was 35.5% (28.9%) higher than errors during wet years. Errors also decrease throughout the forecast season as more of the water year unfolds and a greater percentage of the annual precipitation accumulates. The coefficient of variation of the forecast error was 30.5% higher during dry years compared to wet years, highlighting the site-specific sensitivity of forecast error to snow drought.

Higher forecast errors during dry years are likely indicative of how low snowpack accumulation and early ablation fundamentally change the mechanisms of streamflow generation and how WSFs statistically internalize these changes. However, in characterizing the streamflow generation process, it is important to recognize that there is also a data limitation signal to disentangle. A study of WSFs for Lake Powell from 1947 to 1984 found that on January 1, nearly 80% of forecast error was due to unknown precipitation after the forecast issue date, only decreasing to about 50% by April 1 (Schaake and Peck, 1985). For this study the change in forecast error from early to late season (Figure 1a) remains fairly consistent from wet to dry years, suggesting that there is some baseline level of error derived from these data limitations. More work is needed, however, to fully quantify how unknown precipitation derived error propagates forward during wet versus dry years.

Another limitation of statistically-based WSFs is the behavior of forecasts during significant accumulation anomalies. Trained on a dataset of long-term historical conditions, WSFs are generally very robust when predicting normal to near-normal conditions. However, significant accumulation anomalies, particularly early in the forecasting season, produce larger forecast biases (Figure 1b), where dry years systematically over predict streamflow and wet years systematically under predict streamflow. This suggests that large accumulation anomalies are inherently challenging to predict, particularly if they exist outside of the historical extremes. Together, these results illustrate what should be a growing concern for water managers: not only are streamflow volumes lower following snow droughts, but there are higher forecast errors, greater forecast error variability, and a systematic overprediction of streamflow, all driving an increase in WSF uncertainty.

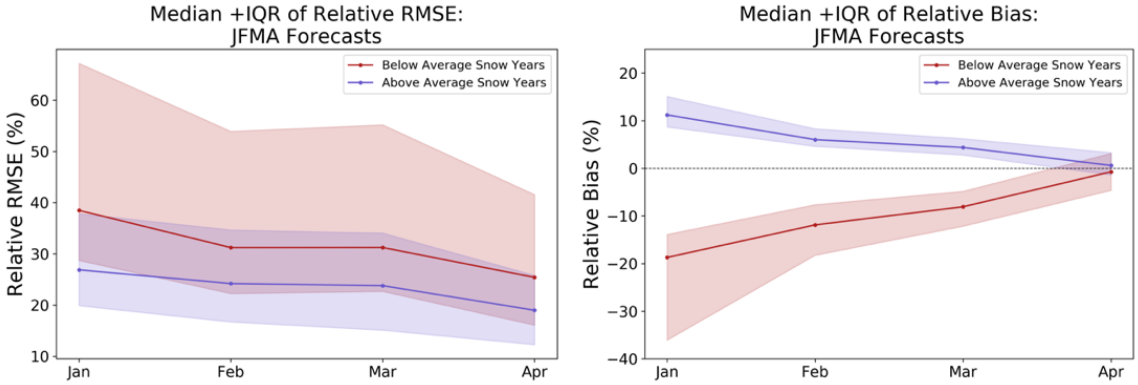

Figure 1. a) Relative RMSE (%) for WSFs during below (red, top) and above average (blue, bottom) snow years at all lead times for water years 1981 to 2014; solid lines are the median values bound by the interquartile range (25th to 75th percentiles) in the shaded region, b) Mean relative bias (%); dashed black line equals zero, or no forecast bias (above, blue, top, and below, red, bottom).

Snow drought may challenge the methodology of PCR forecasting, but what physical mechanisms drive forecast error hydrologically? By considering late-season forecasts only, we can minimize (though not eliminate) the exposure of our analyses to the issue of unknown precipitation. In most of the Western U.S., peak SWE has occurred by April 1 and drier weather begins to dominate, with the notable exception of the Southwest's spring monsoon. Analyzed spatially, a strong regional pattern in mean April 1 WSF error emerges: higher errors exist in the Southwest and Sierra Nevada/Great Basin, contrasted by lower errors in the Northwest extending into the Northern Rockies (Figure 2). Preliminary results suggest that the sites most proximal to the monsoon are not subject to any additional uncertainty or error from spring water inputs as compared to other sites, though more research is needed.

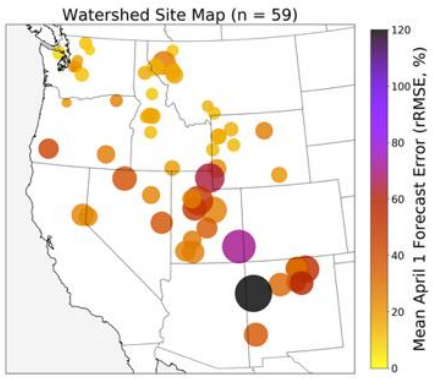

Figure 2. Watershed site map (n = 59) colored by mean April 1 forecast error (rRMSE, %) where the size of the markers is proportional to the coefficient of variation of the runoff efficiency for each watershed.

The assessment of WSF error is, in effect, an evaluation of how well a linear model represents the streamflow response to a unit precipitation on a watershed scale. Runoff efficiency (RE), the ratio of streamflow to precipitation, is a metric conceptually embedded within the WSF assessment question: WSFs rely on a RE to measure precipitation and predict streamflow. Furthermore, by using a linear regression model, standard statistically-based forecasts assume a single RE. Through assessing WSFs, we test this assumption and ask: do dry and wet years yield streamflow equally efficiently? Our results highlight the dependence of statistically-based WSFs on the single RE: we found that the inter-annual variability (coefficient of variation) of RE is strongly correlated with the mean April 1 forecast error (r^2=0.82), where a site with a more variable RE has a larger mean forecast error.

Inter-annual variability of RE encapsulates two problems for WSFs. First, a variable RE reflects variable streamflow (and variable precipitation, though to a lesser extent). Intuitively, a flashy stream, driven by basin attributes, land cover, precipitation phase, and the rate of snowmelt, is simply a more challenging condition to forecast. Second - and more problematic in the face of climate change - is that a more variable RE can also signify a non-linear streamflow response to water inputs. For example, under such a regime runoff is generated less

efficiently during dry years (low RE) and more efficiently during wet years (high RE). Such complexity is oversimplified by the linear regression within most statistically-based WSF. Our findings suggest snow drought drives more variable RE through both of these mechanisms, and hints towards challenges ahead for the assumptions of stationarity and linearity. Despite these challenges, statistically-based WSFs will continue to be an invaluable tool for water managers, but will require creative solutions to address large-scale, non-stationary RE.

CONCLUSIONS

Snow drought presents significant challenges to the assumptions of linearity and stationarity on which statistically-based WSFs are predicated. Across the 1980 to 2014 period of record, we demonstrated that below average snow years have forecast errors that are nearly one-third (29.2%) higher than their counterpart. Furthermore, we were able to show that long lead time forecasts are challenged by accumulation anomalies, in particular, negative accumulation anomalies (below average precipitation) which drives WSF over prediction. Strong regional patterns of forecast error exist, and our results demonstrate that mean April 1 forecast errors are well-correlated (r^2=0.82) with the coefficient of variation of runoff efficiency (RE). These findings point towards challenges ahead for statistically-based WSFs and likely for physically-based forecasts as well. However, improving such WSFs is a daunting task that requires navigating significant institutional inertia for plausibly marginal gains (Harpold et al., 2017b; Lehner et al., 2017). Non-linear transformations of predictor variables, running a two-step PCR, or sub-setting the historic training data to create dry, normal, and wet year forecasts are some considerations that could address some of the shortcomings of the PCR-based WSF models in forecasting streamflow during low snow years.

REFERENCES

Addor, N., Newman, A. J., Mizukami, N., & Clark, M. P. 2017. The CAMELS data set: catchment attributes and meteorology for large-sample studies. *Hydrol. Earth Syst. Sci*, *21*, 5293–5313.

Barnett, T. P., Adam, J. C., & Lettenmaier, D. P. 2005. Potential impacts of a warming climate on water availability in snow-dominated regions. *Nature*, *438*(7066), 303–309.

Berghuijs, W. R., Woods, R. A., & Hrachowitz, M. 2014. A precipitation shift from snow towards rain leads to a decrease in streamflow. *Nature Climate Change*, *4*(7), 583–586.

Garen, D. C. 1992. Improved Techniques in Regression-Based Streamflow Volume Forecasting. *Journal of Water Resource Planning Management*, *118*(6), 654–670.

Harpold, A., Dettinger, M., & Rajagopal, S. 2017. Defining Snow Drought and Why It Matters. *Eos*, *98*.

Harpold, A. A., Sutcliffe, K., Clayton, J., Goodbody, A., & Vazquez, S. 2017. Does Including Soil Moisture Observations Improve Operational Streamflow Forecasts in Snow-Dominated Watersheds? *Journal of the American Water Resources Association*, *53*(1), 179-196.

Knowles, N., Dettinger, M. D., & Cayan, D. R. 2006. Trends in Snowfall versus Rainfall in the Western United States. *Journal of Climate*, *19*, 4545–4559.

Lehner, F., Wood, A. W., Llewellyn, D., Blatchford, D. B., Goodbody, A. G., & Pappenberger, F. 2017. Mitigating the Impacts of Climate Nonstationarity on Seasonal Streamflow Predictability in the U.S. Southwest. *Geophysical Research Letters*, *44*(24), 12, 208-12,217.

Mantua, N., Dettinger, M.D., Pagano, T.C., Wood, A.W., Redmond, K., & Restrepo, P. 2008. A Description and Evaluation of Hydrologic and Climate Forecast and Data Products that Support Decision-Making for Water Resource Managers. In: Decision-Support Experiments and Evaluations using Seasonal-to-Interannual Forecasts and Observational Data: A Focus on Water Resources. *U.S. Climate Change Science Program and the Subcommittee on Global Change Research.*, 29-64.

Mote, P. W., Li, S., Lettenmaier, D. P., Xiao, M., & Engel, R. 2018. Dramatic declines in snowpack in the western US. *Npj Climate and Atmospheric Science*, *1*(1), 2.

Newman, A. J., Clark, M. P., Sampson, K., Wood, A., Hay, L. E., Bock, A., … Duan, Q. 2015. Development of a large-sample watershed-scale hydrometeorological data set for the contiguous USA: data set characteristics and assessment of regional variability in hydrologic model performance. *Hydrol. Earth Syst. Sci*, *19*, 209–223.

Schaake, J.C. & Peck, E.L. 1985. Analysis of water supply forecast accuracy. *Proceedings of the 53rd Annual Western Snow Conference,* 44-53.