An assessment of NRCS seasonal streamflow forecast performance in Idaho

Christina Andry

A report prepared in partial fulfillment of
the requirements for the degree of

Master of Science
Earth and Space Sciences: Applied Geosciences

University of Washington

March 2018

Project mentor:
Danny Tappa, USDA NRCS Snow Survey

Reading committee:
Juliet Crider
Nicoleta Cristea

MESSAGe Technical Report Number: 062

Executive Summary

This report summarizes recent NRCS seasonal streamflow forecast performance in Idaho. Considering only April 1st forecasts for the April-July period, I first attempted to quantify forecast performance from a users' perspective, using three simple statistical parameters: (1) the standard error derived from each forecast equation, (2) the percent error of the median forecast in relation to the observed flow, and (3) a count of how frequently the observed streamflow volume fell outside the forecasted range of values. Regional differences in these forecast error metrics were apparent, with points in northern basins showing lower standard errors and percent errors but higher out-of-range counts, and points in southern basins showing higher standard errors and percent errors and generally lower out-of-range counts. The high out-of-range counts across the state in certain years can be attributed to extreme spring weather events that could not be predicted using the current statistical forecast techniques. Future weather remains the largest source of uncertainty in seasonal streamflow forecasts.

I then used the Nash-Sutcliffe skill score (NS), which accounts for the annual variability of observed streamflow and allows for direct comparison of forecast skill between basins with highly-variable streamflow and basins with more annual consistency in streamflow volumes. Regional differences in forecast error were filtered out using the NS skill score, and points with lower scores were identified, possibly revealing more systematic errors in the forecasts. I then attempted to isolate influences on forecast quality by comparing basin- or region-scale metrics to forecast performance. Of the factors considered in this analysis, climatic region and the occurrence of extreme spring weather events remain the strongest influences on streamflow forecast quality and performance.

Table of Contents

Introduction

In 2015, water supply shortages across Idaho resulted in a $1.3 billion decline in agriculture cash receipts from the previous year (Idaho State Department of Agriculture, 2016). The Idaho Water Resources Board (IWRB) attributed some of these losses, mostly in the Snake River Plain region, to water management inefficiencies that may have been avoided with improved streamflow forecasts and snowpack data. IWRB identified a data deficit at mid-elevations, where snowmelt occurred much earlier than at higher elevations. Most data sites are situated at higher elevations. The early spring peak flows in the Upper Snake River Basin came largely as a surprise to downstream water managers, who were unable to plan for the unexpected and rapid increase in reservoir inflows. The IWRB called for an investigation of these mid-elevation data gaps in the Natural Resources Conservation Service (NRCS) SNOTEL network and intends to fund the installation and maintenance of several new sites to improve streamflow forecasts. The NRCS is currently performing a data network analysis to determine whether the addition of automated snow monitoring sites will improve forecasts. A component of this investigation involves identifying basins in the state in which seasonal streamflow forecast performance is poor or inconsistent.

With over 3.3 million acres of irrigated land and one of the largest agricultural economies in the western United States, Idaho depends on water supply forecasts to make informed water management decisions, especially in water sparse regions (Idaho State Department of Agriculture, 2016). The arid Snake River Plain is home to most of the state's agricultural land and relies on runoff from the surrounding mountainous regions for water supplies. Most municipal water supplies are also fed by streamflow originating in the mountains. Snowmelt-derived runoff accounts for an estimated 50% of total annual runoff in the western U.S., and up to 70% of total runoff in high-elevation mountainous regions (Li et al. 2017). Consistent, real-time information related to snowpack and mountain weather conditions is critical for quantifying future streamflows and for subsequent water management decision-making.

In this report, I perform a cursory assessment of streamflow forecasts in Idaho to identify basins or regions where forecast performance is poor. Though I do not determine if or where forecasts could be improved with the addition of new data sites, the results of this report will guide the broader IWRB-funded data network analysis project mentioned above. After calculating a series of error metrics that offer several perspectives of forecast performance, I begin by addressing the question: do basins with more SNOTEL sites correspond to better forecasts? I then attempt to isolate physiographic and climatic influences on forecast quality by comparing basin- or region-scale metrics to forecast performance. Because this is the first statewide review of streamflow forecasts, this study also provides NRCS Idaho staff with a better understanding of recent forecast performance and initiates efforts to better communicate the limitations of streamflow forecasts to end users.

Background

The NRCS Snow Survey data network

The USDA Natural Resources Conservation Service (NRCS) Snow Survey operates a network of high-elevation snow monitoring stations throughout the mountainous western United States for the purpose of water supply forecasting. Over 800 of these data collection sites are automated SNOTEL (snow telemetry) stations that transmit hourly measurements of snow depth, snow water equivalent, precipitation, temperature, and other meteorological conditions to a publicly available web database. An additional 1400 sites serve as locations for monthly manual snowpack measurements ("snow courses"), many of which were established in the early 20th century prior to the transition to the automated SNOTEL network in the 1970s. Primary users of the Snow Survey data products include reservoir operators, water managers, irrigators, scientific researchers, and recreationists.

Monthly streamflow forecasts

In addition to the collection and quality control of data, the NRCS generates monthly Water Supply Outlook Reports (WSORs), which summarize current snowpack conditions, current reservoir levels, and forecasts of future streamflow volumes. As the snow accumulation season progresses through the winter and into early spring, the monthly streamflow forecasts are updated to reflect current conditions. Forecasts are presented as cumulative seasonal volumes – typically April through July or April through September – and as a distribution of five values based on probability of occurrence or exceedance (Table 1). For example, the minimum streamflow volume presented in a forecast corresponds to a 90% chance of occurrence or exceedance and the maximum volume corresponds to a 10% chance. With an understanding of the uncertainty inherent in streamflow forecasting, end users can then assess the range of provided values and make decisions according to their individual needs.

NRCS relies on statistical forecasting techniques (as opposed to conceptual or physically-based models) and presently employs a principal components regression method, developed by David Garen, the Development Hydrologist at the NRCS National Water and Climate Center (NWCC) (Garen, 1992). An overview of the NWCC streamflow forecasting procedure can be found in Appendix A. In snowmelt-dominant watersheds, this statistical approach has effectively produced reliable forecasts for most years. However, in mixed rain-and-snow basins and in years of above or below average precipitation or temperature, the relationship between snowpack and seasonal streamflow volumes is much more difficult to predict using statistical models.

Forecast points correspond to gaged points on major rivers or tributaries and are chosen based on the needs of end users. Information from SNOTEL sites or snow courses in or near the contributing drainage area to the forecast point are used as the input data for the forecast model. Current water year precipitation and snow water equivalent are the most commonly used predictor variables. Additional predictor variables that are less commonly used include antecedent streamflow or climate teleconnection (e.g. ENSO). Thirty or more years of continuous precipitation, snowpack, and streamflow data are typically required to calibrate a forecast model. For cases in which one or more years of data are missing, an alternate statistical method, a Z-score regression, is applied (see Appendix A for details). Models are recalibrated every few years, or as frequently as the forecaster deems necessary.

Once the input data (predictor variables) are run through the final equation, the model returns a streamflow volume that serves as the median forecast value (50% chance of exceedance). The standard error of the regression equation is then used to calculate the minimum (90% exceedance), lower (70% exceedance), upper (30% exceedance), and maximum (10% exceedance) forecast volumes, which are typically normally distributed. The greater the variability in the relationship between snowpack and streamflow in past years, the larger the standard error of the forecast equation, and the greater the spread between exceedance forecast values.

Final forecasts are published in the state's monthly WSOR and distributed to the public (https://www.wcc.nrcs.usda.gov/state_outlook_reports.htm). In each WSOR, an introductory statement outlines the forecasting method and explains how users should interpret the range of values presented in each forecast. It also emphasizes that there are several sources of uncertainty: the equations themselves, the data used to create the equations and generate forecasts, and, most importantly, the unpredictable future weather conditions.

Forecast basins of Idaho

This analysis is focused on forecast points included in the Idaho WSORs. Figure 1 displays these 69 forecast points, their associated contributing drainage areas, and the nine greater forecast areas. These include the Panhandle region, the Clearwater Basin, the Salmon River Basin, the West Central Basins (which includes the Weiser, Payette, and Boise River Basins), the Lost and Wood River Basins, the Upper Snake region, the Bear River Basin, the Southside Snake region, and the Owyhee River Basin. A complete list of forecast points, their associated forecast region, and basin metrics can be found in Appendix B.

Each forecast area is associated with unique physiographic and climatic characteristics. Figure 2 reveals the varying topography and complex annual precipitation patterns across the state. Together, the topography and climate dictate land use, resource availability, and economic activities of each region. In general, southern Idaho is more arid than northern Idaho, but is home to most of the state's agricultural lands. Elevations in northern Idaho are lower on average and annual precipitation totals are higher. Water shortages in Idaho are rarely an issue in normal years, but southern basin water supplies are more sensitive to a below-average snowpack.

The average characteristics of annual snowpack vary regionally. Most of Idaho is classified as an intermountain snowpack regime (Trujillo and Molotch, 2014). An intermountain snowpack exhibits characteristics that fall between those of the higher density, higher volume maritime snowpack of the Cascades and Sierra Nevada and the lower density, lower volume continental snowpack of the Rockies. The snowpack in the headwaters of the Snake River is classified as continental (Trujillo and Molotch, 2014).

Contracted GIS basin analysis (BAGIS) project

A forecast performance review was intended to supplement a broader project assessing data gaps in the Idaho SNOTEL network and identifying new snowpack monitoring sites. This year-long project is funded by IWRB and is led by Kara Ferguson, an independent contractor to NRCS and graduate student at Boise State University. The procedure for identifying data gaps and new data sites was standardized in the Basin Analysis GIS (BAGIS) system, developed by researchers at Portland State University in conjunction with the NRCS-NWCC (Duh, 2011). The BAGIS procedure performs a series of spatial computations in GIS and generates maps and reports detailing the results. The analysis optimizes locations for new monitoring sites based on PRISM data, elevation, slope, aspect, accessibility, and proximity to existing data sites. The goal of the BAGIS project is to produce information on basins or regions where recent forecast performance is poor or inconsistent, to help guide final decisions regarding establishment of new SNOTEL sites to improve water supply forecasts.

Forecast performance analysis

My role in the BAGIS project is to identify basins in Idaho in which streamflow forecast performance is poor or inconsistent. A statewide forecast performance analysis has not previously been conducted for Idaho. In this analysis, I consider only April 1st forecasts for the April-July period. I first attempted to quantify forecast performance from a users' perspective, using three statistical parameters: (1) the standard error derived from each forecast equation, (2) the percent error of the median forecast in relation to the observed flow, and (3) a count of how frequently the observed streamflow volume fell outside the forecasted range of values (either below the minimum forecast or above the maximum forecast). Each of these three measures of error is illustrated in Figure 3. All three parameters were calculated for the most recent 10-year period (2008-2017). Though ten years is an insufficient duration to represent long-term trends in forecasts and observations, it still captures annual variability in snowpack conditions and allows for identification of recent patterns and inconsistencies in forecast performance.

To assess the predictive power or the "skill" of forecasts, I used the Nash-Sutcliffe coefficient of efficiency (Nash and Sutcliffe, 1970) for NRCS streamflow forecasts over the last twenty years. The Nash-Sutcliffe (NS) skill score accounts for annual variability of streamflow observations and allows for comparison between forecasts in basins with high local variability and those with more consistent seasonal streamflow volumes. I then compared the standard error and the NS skill score for each forecast point to a series of forecast basin metrics to identify possible influences on the quality of forecasts.

(1) Normalized standard error

The first "user perspective" error metric I assess was the standard error of each April 1 forecast model equation. Though the standard error does not indicate how a forecast compares to actual observed flows, it represents the historic annual variability of snowpack-streamflow relationships in a basin and expresses the level of uncertainty in the forecast. I calculated the standard error of each forecast for each year by taking the difference between the maximum forecast volume and the median forecast volume (Figure 3a). To allow for comparison between large rivers and smaller streams, I normalized the standard errors from each forecast to the 30-year average streamflow for the same April-July period. I then took the ten-year mean of the normalized standard errors for each forecast basin and used this value to identify basins and regions with the highest uncertainty in their forecasts.

The 10-year mean normalized standard error for each forecast basin is displayed on a map in Figure 4 and listed in a table in Appendix C. Basins with the highest mean standard errors are located in the southern portion of the state. Forecast regions with the highest standard errors include the Owyhee, Bear, and Southside Snake River Basins, all of which line Idaho's southern border. Forecast regions with the lowest standard errors include the Clearwater Basin, Panhandle Region, and the Upper Snake River Basin. Forecast regions in the central portion of the state make up the middle range of standard errors.

(2) Mean percent error

To determine how forecasts relate to observations from a user's perspective, I calculated the percent error of the median April 1 forecast in relation to observed flows for each year. The percent error is equal to the difference between the observed volume and the median forecast, divided by the observed volume, and multiplied by 100. Again, I took the ten-year (2008-2017) mean of the calculated percent errors for each forecast basin. The percent error indicates how close the observed flow is to the median forecasted volume (Figure 3b), but does not indicate whether the observed flows fall within the forecasted range of volumes.

A map of mean percent error by basin reveals a similar pattern to the map of standard errors: basins with highest percent errors of the median forecast are in the southern portion of the state and line the Snake River Plain (Figure 5). Basins with the highest percent errors fall within the Owyhee River Basin, Lost and Wood River Basins, and lower portions of the Upper Snake River Basin. Appendix D lists the complete range of mean percent errors by basin.

(3) Count of observed flows outside forecast range

To determine how observed flows compare to the entire forecasted range of volumes (not just the median volume), I performed a simple count of how many times in the ten-year period that observed flows fell outside this range (Figure 3c). Statistically, observed flows should fall below the minimum forecast value or above the maximum forecast value 20% of the time, or 2 out of the 10 previous years. If observed flows fall outside this range more frequently, the standard error used to calculate the maximum and minimum values is likely too small and should be adjusted to reflect greater uncertainty in the forecast.

The spatial distribution of this out-of-range count is nearly opposite of that for standard error and percent error, with highest counts in the northern portion of the state (Figure 6). The Clearwater Basin and Panhandle Region show the highest counts of observed flows outside of the forecasted range, at 4-5 years out of 10 in the Clearwater and 4-6 years out of 10 in the Panhandle. All other forecast regions in the state contain at least one basin with 3 or more years of out-of-range observed flows, despite wider ranges and larger formal uncertainties of forecasts. Table 2 lists the forecasted flows, observed flows, and bin classification for the basin with the highest out-of-range counts and Figure 7 illustrates the forecasts and observations in a boxplot.

In three out of the ten years, more than 30 basins out of 69 across the state saw observed flows that were out of the forecast range (Figure 7). In 2010 (40 basins) and 2011 (49 basins), all flows were greater than the associated maximum forecasted volume. In 2016, all 30 basins with out-of-range observed flows were below the minimum forecasted volume. In every other year in the 2008-2017 period, at least 9 and as many as 19 basins saw flows outside of the forecast range.

Nash-Sutcliffe efficiency skill score

The Nash-Sutcliffe efficiency skill score (NS) is a metric commonly used to assess the predictive power of hydrologic models (Nash and Sutcliffe, 1970). It essentially takes the squared errors of the median forecasts for each year and normalizes them to the variance of observations for the period of analysis:

$$NS = 1 - \frac{\sum_{i=1}^{N}(f_i - o_i)^2}{\sum_{i=1}^{N}(\bar{o} - o_i)^2}$$

where f_i and o_i are the forecast and observation for the ith year in a total of N years, and $\bar{o}$ is the mean of the observations for N years (Pagano et al, 2004). A NS score of 1 indicates perfect skill, 0 indicates no skill, and a score less than 0 indicates negative skill. When a forecast has negative skill, the observation for that year is closer to the mean of observations for the period record than it is to the forecast. This skill score has previously been used for NRCS streamflow forecasts, and Figure 8 displays a map produced by Pagano (2005) for NRCS forecast points across the western U.S., calculated for the period of record through 2005.

To update the Pagano NS skill score map, I calculated a new NS skill score for forecast points in Idaho with at least 20 years of forecast-observation pairs (1998-2017). 46 of the 69 forecast points had complete data for this period. Figure 9 displays the updated NS skill score ranges for each forecast point and reveals that skill scores in Idaho for this period range from 0.42 to 0.86. Appendix E includes a table of all forecast points and their associated NS skill scores.

Identifying relationships between error parameters and basin characteristics

I compared both the normalized standard error and the NS skill score for each forecast point to three basin characteristic metrics: contributing drainage area, mean basin elevation, and percent forest cover. All metrics were obtained from either the USGS Streamstats database (U.S. Geological Survey, 2012) or from the NRCS BAGIS (Basin Analysis GIS) program (Duh, 2011). I chose to use both the NS skill score and the standard error because the two error parameters are not strongly correlated, as displayed in Figure 10, and may reveal different relationships to the chosen basin metrics. For example, if a basin metric shows significant correlation to the standard error, the metric may explain the historic variability in the snowpack-streamflow relationship and the resulting forecast uncertainty. On the other hand, if a basin metric shows significant correlation to the NS score, it may reveal that the metric has notable influence on forecast performance.

I first plotted each error parameter against contributing drainage area. As shown in Figure 11, there is no clear relationship between contributing drainage area and either of the error parameters.

Figure 12 displays each error metric against mean basin elevation. Figure 12a reveals that in 6 of the 9 greater forecast areas, mean basin elevation seemed to correlate with the normalized standard error. I fit a regression line to each of these six forecast area groupings and for five of the areas (Clearwater, Panhandle, West Central, Lost/Wood, and Upper Snake), observed a clear decrease in standard error with an increase in mean basin elevation. In the Salmon River Basin, standard error increases with an increase in mean basin elevation. In each of the remaining three forecast areas (Southside Snake, Bear, and Owyhee), there is either an insufficient number of forecast points to detect a trend or no clear correlation between standard error and mean basin elevation.

A relationship between mean basin elevation and NS skill score is less apparent (Figure 12b). In the West Central area, forecast skill seems to increase with an increase in mean basin elevation, but in each of the other forecast areas, there is either an insufficient number of points to detect a trend or no clear correlation between standard error and mean basin elevation.

Figure 13 displays each error metric against percent forest cover. There is overall a clear relationship between percent forest cover and standard error (Figure 13a). Standard error seems to decrease with an increase in forest cover. Within most composite forecast areas, the same relationship seems to apply: individual forecast basins with higher percent forest cover seem to have lower standard errors than nearby basins with lower forest cover. The only forecast area that does not show this correlation is the Southside Snake, which has some of the lowest forest cover and some of the highest standard errors. There does not seem to be a relationship between percent forest cover and NS skill score.

Identifying relationships between error parameters and data availability

To determine if more data sites in a basin inherently yields better forecasts, I compared each error parameter to three measures of data availability: (1) the number of data sites (SNOTEL or snow course) used as predictor variables in the forecast equation, (2) the number of data sites (SNOTEL only) per unit area available in each forecast basin, and (3) the area of the elevation band represented by SNOTEL sites in the basin. I used both metrics (1) and (2) because in many cases, not all data sites in a basin are used in its forecast and not all data used in a forecast are from sites within the forecast basin boundaries (an example is shown in Figure 14). Appendix A includes more information on how data sites are selected for use in a forecast model. If more input data (more predictor variables) result in forecasts with lower errors, or if more data sites per unit area are associated with lower forecast errors, adding data sites to a basin may improve its forecast performance. If larger elevational distributions of SNOTEL sites in a basin are associated with lower forecast errors, adding data sites across more representative elevations may improve forecast performance.

The number of data sites used in each forecast equation were obtained from the forecast model summaries (Figure 15) and normalized to the corresponding basin area to obtain site per unit area. The number of sites per unit area was plotted against each error parameter. Figure 16 reveals no apparent relationship between number of sites used in a forecast per unit area and forecast performance or quality.

The number of data sites available in each basin were obtained using overlapping shapefiles of SNOTEL site locations and delineated forecast basins in ArcMap. Only SNOTEL sites were counted because snow courses are more rarely used as predictor variables in a forecast and because any site additions to the network would likely be automated. The count of data sites in each basin were then normalized to the basin's area to obtain site per unit area. The count of available sites per unit area was plotted against each error parameter. Figure 17 reveals no apparent relationship between the number of available SNOTEL stations per unit area and forecast performance or quality.

I obtained the percent area represented by the elevations of available SNOTEL sites using an area-elevation curve for each basin overlain with points representing the elevation of each in-basin data site (Figure 18). These plots are typical products of the BAGIS program and were generated by Kara Ferguson. I took the difference between the percent area below the highest elevation SNOTEL site and the percent area below the lowest elevation SNOTEL site and called this value the percent area represented by available sites. I plotted the percent area against each of the error parameters. Area-elevation curves were not available for every basin, so some basins were excluded from this portion of the analysis.

Figure 19 displays plots of each error parameter against the percent area represented by the elevations of available data sites in a basin. Figure 19a shows that there is not a strong correlation between standard error and percent area represented. There may be a loose correlation between NS skill score and percent area represented, but the number of points available in this plot (n = 29) make it difficult to definitively make this claim (Figure 19b).

Discussion

Trends in user-perspective forecast error

The clustering of basins with highest forecast uncertainties and percent errors in the southern portion of the state was anticipated because of the region's higher variability in April-July streamflow volumes (Figure 20). The Snake River Plain and southern Idaho are the most arid regions in the state and nearly all of the annual precipitation falls at high elevations in the mountains (Figure 2). Relative to northern basins, lower total snowfall accumulation makes the snowpack of southern Idaho more sensitive to spring rains or warm spring temperatures. Any significant rain-on-snow event can melt out a large portion of the snowpack and result in peak flows at times that vary from year to year, depending on the earliest occurrence of such a storm.

The lowest forecast uncertainties and percent errors in the northern portion of the state and in the uppermost headwaters of the Snake River reflect the higher total annual precipitation (Figure 2) and snowpack distributions of those regions and the more consistent streamflows (Figure 20). These are the wettest basins in the state, where the relatively thick snowpack is not as susceptible to melt out by a single spring rain-on-snow event. David Garen (developer of the NRCS forecast technique) also reported that data sites in the northern basins show a relatively high degree of collinearity. This reflects a more spatially uniform relationship between point snowpack accumulation at any given data site and streamflow.

The third user error metric – the count of years that the observed flows fall outside the forecasted range – demonstrates a different dimension of forecast performance. Despite their lower forecast uncertainties and lower percent errors of the median forecasts, the northern Idaho basins had the highest out-of-range counts in the state. While the observed flow volumes were overall closer to the median forecast (lower percent errors), they more frequently fell below the minimum or above the maximum forecasts. Nearly half of Idaho's forecast basins saw three or more years in the last decade with observed flows falling outside of the forecasted range of volumes, despite the statistical expectation that this should only occur 20% of the time, or twice in the last decade. The high out-of-range counts in 2010, 2011, and 2016 (Figure 21) can be attributed to extreme weather events in the spring of each year. In 2010, a large snowstorm in April recharged the snowpack, left peak SWE values much higher than the April 1 measurement used in forecasts, and resulted in higher than anticipated April-July streamflow volumes. In 2011, the snow accumulation period continued past April 1 and through the end of spring, also yielding higher than expected streamflows. In 2016, a warmer than average April melted out snow much earlier than expected, resulting in lower than forecasted streamflows. While statewide spring-season weather events may explain poor forecast performance for these three years, it is unclear why many northern basins are out-of-range during other years.

So which error metric indicates that a forecast is "good" or "better" than the others? Is it a lower percent error of the median forecast? When the observed flows fall more closely to this middle-of-the-range volume? Or is it a less frequent occurrence of observed flows falling outside the forecasted range of volumes? The answer depends on how an individual is utilizing the forecast. If the user is making decisions based on the median volume (50% exceedance), then forecasts in the Panhandle, Clearwater, and Salmon River basins might be considered the "best." If the user is making decision based on either the minimum (90% exceedance) or the maximum (10% exceedance) forecasted volume, then the "best" forecasts are certainly not the same ones as considered by the median user; but are some of the ones considered the "worst" by the median user where volumes fall within the forecasted range but are relatively further from the median. Ideally, observed flows will fall close to the median volume *and* within the forecasted range, as intended by the forecaster.

Predictive power of forecasts

The Nash-Sutcliffe coefficient of efficiency provides a better, or more fair assessment of forecast quality than the above user-perspective error metrics. When the local variability of observations is accounted for, it allows for direct comparison of forecast skill between basins with highly-variable streamflow and basins with more annual consistency in streamflow volumes.

The updated NS skill score map reveals significant improvement in skill for many points across the state from Pagano's (2005) previous analysis (Figure 8 and Figure 9). Forecast points with skill scores above 0.8 are mostly located in the West Central basins (Figure 9). The forecast points with the lowest NS skill scores for the 1998-2017 period were scattered across several forecast areas. The absence of spatial patterns in NS skill score confirms the usefulness of the metric by ensuring that regional climatic differences are accounted for in the skill score calculation.

The Deadwood Reservoir Inflow forecast point in the West Central basins had the highest NS score at 0.86. The Priest R nr Priest forecast point in the Panhandle had the lowest NS score at 0.42. Deadwood and Priest had nearly identical normalized standard errors (0.211 and 0.212, respectively), but show significantly different predictive capabilities for the recent twenty-year period. Figure 22 shows the 2008-2017 forecast-observation pairs for Deadwood and Priest and provides a visual comparison of forecast performance. In five out of ten years in the Priest River basin, observed flows fell outside the forecasted range and overall, the differences between the observations and the median forecasts were greater than for Deadwood. In only one out of ten years, observed flows at the Deadwood Reservoir Inflow point fell outside the forecasted range.

Trends in basin characteristics

As a whole, forecast quality or performance did not seem to correlate with basin size. Mean basin elevation did seem to correlate with standard error, but only within each forecast region. If the plot in Figure 12 did not separate points by forecast region, there would be no clear relationship between elevation and forecast uncertainty. For example, there is a clear decrease in the standard error with an increase in elevation in the Clearwater Basin. But mean basin elevations in the Clearwater are some of the lowest in the state so the relationship between elevation and forecast uncertainty are relative to each forecast region (or climatic region). On the other hand, the Lost and Wood River basins are home to the highest peaks in Idaho and have some of the highest mean basin elevations in the state. However, forecasts in these basins also have overall higher uncertainties. Even the highest elevation basin in the Lost/Wood region has a higher standard error than the lowest elevation basin in the Clearwater. As stated before, this is a reflection of the variability in snowpack conditions. Thus, a generalization that basins with lower (or higher) mean elevations are connected to forecasts with higher (or lower) uncertainties cannot be made.

The Salmon River Basin is the only forecast area that exhibits an increase in standard error with an increase in mean basin elevation. This can likely be explained by the topographic and climatic variations within the Salmon River Basin forecast region (Figure 2). The highest peaks (just northwest of the Lost and Wood River valleys) are the driest areas in the forecast region, while the lower-elevation subbasins receive more annual precipitation.

As a result, snowpack and streamflows are more variable in the higher-elevation subbasins of the greater Salmon River Basin than in the lower-elevation subbasins.

Mean basin elevation did not seem to correlate with the NS skill score. The correlation of mean basin elevation to standard error and not NS skill score suggests that mean basin elevation controls the variability of snowpack and streamflow, but is not a significant factor in the forecast performance as it relates to observed flows.

Trends in data availability

Incorporating more data sites (or predictor variables) into a forecast does not necessarily improve the forecast performance. The principal components regression method used for these forecasts minimizes the number of predictor variables used in a forecast model by identifying highly intercorrelated variables and summarizing them in fewer uncorrelated "components" (Garen, 2011). In basins like the Panhandle and Clearwater, where many of the sites exhibit this collinearity, the resulting forecast models may incorporate very few predictor variables and still exhibit lower standard errors and uncertainties. In basins like the Owyhee or Bear, where the snowpack exhibits higher spatial variability, the resulting forecast models may incorporate many predictor variables and still exhibit higher standard errors and uncertainties.

There does not seem to be a correlation between standard error and any of the three data availability metrics. NS skill score does not show considerable correlation to either the number of data sites used in a forecast or the number of data sites available in a basin. There may be a correlation between NS skill score and the percent area represented by available site elevations, but the sample size for this comparison is limited to 29 basins.

The question of whether more snowpack data inherently yields better forecasts cannot be answered by this analysis alone. The BAGIS project will include a more complete assessment of the spatial distribution of data sites in relation to basin-scale precipitation patterns. Figure 18 is an example product of the BAGIS program that aids in determining whether current data sites are representative of precipitation distribution patterns within the basin. Using the results of the BAGIS project and the results presented in this study, NRCS and IDWR can better understand influences on forecast quality and determine if or where the addition of data sites might improve forecasts.

Limitations of the forecast analysis

The simple statistical parameters chosen for this study offer a very limited perspective of the quality of the forecast models. As expressed by Pagano et al. (2004), "the robustness and scientific rigor of forecast evaluation techniques are inversely proportional to their accessibility and understandability by the lay forecast user." A comprehensive verification of probabilistic forecasts requires more statistically complex methods such as the use of the Brier score (Brier, 1950), the ranked probability score (Epstein, 1969), or a distributions-oriented approach (Wilks, 1995). Though these methods are beyond the scope of this limited analysis, they should be considered in future NRCS forecast verification studies.

The basin metrics I considered in this analysis do not provide a comprehensive review of potential influences on forecast quality and performance. Some additional metrics that might be considered in future studies include: average aspect of slopes; total basin relief; percent forest cover; percentage of annual precipitation falling in months before April 1; percentage of annual precipitation fall in the April-July period; distribution of data sites in relation to the snowline; mean, minimum, and maximum winter temperatures; mean, minimum, and maximum temperatures during the forecast period,

I chose to use all forecast basins included in the Idaho WSORs and did not consider degree of human regulation within a basin, which may affect both the initial forecast uncertainty and the final adjusted observed streamflow volumes. Observed flows are adjusted to represent "natural flows," or what the flows would be without upstream interference. Diversions and reservoir storage and release volumes are used to calculate these adjusted flows in an internal NRCS database. Errors in reporting the degree of human interferences to make these calculations can result in inaccurate adjustments of observed flows.

I also did not assess the frequency of model recalibration or the differences in forecasters across the state and through time. Models are recalibrated as needed, typically every few years. The 10-year mean errors were also used to reflect possible changes to the forecast model equation. For Idaho, there is currently one main forecaster for most basins and two additional forecasters that cover fewer than five of the basins that are included in the state's WSOR. No information on the history of which forecasters were assigned to which basins was considered for this analysis. All forecasters use the same modeling technique and the same forecast production interface, but are ultimately making the decision of which model output to use. Because forecast production depends somewhat on the judgment of the forecaster, a regional bias may be inherent in forecasts and in the results of this analysis.

Statistical forecast models in an era of extremes

The last three water years (2015, 2016, and 2017) have set state snowpack records in many of Idaho's basins. In 2015, many SNOTEL stations and snow courses across the state reported record low April 1 measurements of SWE, and all basins were well below their median values for that date. After a season of near-normal snow accumulation, the spring of 2016 saw some of the earliest peak SWE and melt out dates in recorded history. 2017 was a record-high snowpack year, with some basins in the West Central mountains nearing 250% of normal SWE on April 1. Anomalous snowpack conditions that have not previously been observed in the data record make it difficult to use this record to predict future streamflows. Additionally, extreme spring weather events greatly affect forecast performance, as demonstrated in the analyses included in this study.

To address the challenges of using a statistical forecast model in years of unprecedented snowpack conditions, forecasters may need to include years such as 2015, 2016, and 2017 in their calibration datasets. Many of the forecast models reviewed in this study did not include these recent years for calibration and could be updated for future years.

Future weather remains the largest source of uncertainty in seasonal streamflow forecasts. Because accurate weather predictions cannot be made with a three-month lead time, incorporating weather into seasonal streamflow forecasts is not realistic. NRCS April-July streamflow forecasts are already updated on May 1 and June 1, but may require more frequent updates when large, impactful weather events (e.g. rain-on-snow, rapid or extreme warm-ups) occur during this period. Many forecast end users make important long-term decisions based on early-spring water supply forecasts, including crop selection by irrigators or loan allotment by agricultural bankers, so these additional mid-month forecasts would not entirely solve such problems of unexpected weather. Uncertainty is inherent in long-range forecasts and NRCS must continue to communicate this to users, while making realistic efforts to improve forecast models where possible.

Recommendations and future work

Regarding the IWRB-initiated network analysis project, the decision to install a new data site or automate an existing snow course comes with significant costs and a long-term commitment to data quality monitoring and site maintenance. While funding and support may be available for the early stages of site development, long-term financial and personnel requirements must be taken into consideration as well. Additionally, a site must collect a minimum of ten years of continuous data for it to be incorporated into a statistical forecast model (Garen, 2011).

If the Snow Survey program envisions a future that involves continued use of the current forecast model technique, adding SNOTEL sites in basins with significant data gaps may improve forecasts to some degree. However, the challenges discussed in the previous section should be considered, as they cannot be eliminated by the addition of data sites. Alternatively, if the program anticipates a change to the forecasting approach – for example, a switch to a conceptual or physically-based model – it may be more feasible to add sites according to the data requirements of these other techniques.

A full assessment of data gaps and the usefulness of site additions will occur upon completion of the greater basin analysis study by Kara Ferguson. These next steps will likely begin by late summer 2018. Combining the results of the BAGIS project and this forecast performance analysis, NRCS personnel and the IWRB can make more informed decisions regarding if and where new data sites will be installed. A discussion centered on the future role of the NRCS Snow Survey amidst technological and modeling advancements should also be considered in the process.

Conclusions

Of the factors considered in this analysis, climatic region and the occurrence of extreme spring weather events seemed to most impact streamflow forecast quality and performance. With the high variability observed in snowpack and weather conditions in the last decade, it is difficult for forecasters to predict how future events will affect streamflows. As long as NRCS uses a statistical approach to forecasting, these recent years should be included in forecast model calibrations to better prepare for possible extremes in the future.

It remains unclear whether an increase in data availability will improve statistical forecasts in a timeframe useful to end users. Regardless, NRCS forecasters and hydrologists have a responsibility to communicate the error in a streamflow forecast to the public, so that users can make informed and cautious decisions. NRCS also has a responsibility to communicate the limitations and inherent uncertainty in streamflow forecasts in order to manage the expectations of end users.

References

Brier, G. W., 1950, Verification of forecasts expressed in terms of probability, Monthly Weather Review, v 78, p 1–3.

Duh, J. D., 2011, Basin Analysis GIS - A Spatial Decision Support System for Selecting Snowpack Monitoring Sites, February 26, 2011, Oregon Academy of Science Conference, Portland, Oregon.

Epstein, E. S., 1969, A scoring system for probabilities of ranked categories. Journal of Applied Meteorology, v 8, p 985–987.

Garen, D. C., 1992, Improved techniques in regression-based streamflow volume forecasting. Journal of Water Resources Planning and Management, 118, p 654–670.

Garen, D., 2011, Water Supply Forecasting, Chapter 7 of Part 622 Snow Survey and Water Supply Forecasting National Engineering Handbook, 210-VI-NEH, Amend. 41, March 2011.

Homer, C.G., Dewitz, J.A., Yang, L., Jin, S., Danielson, P., Xian, G., Coulston, J., Herold, N.D., Wickham, J.D., and Megown, K., 2015, Completion of the 2011 National Land Cover Database for the conterminous United States-Representing a decade of land cover change information. Photogrammetric Engineering and Remote Sensing, v. 81, no. 5, p. 345-354.

Idaho State Department of Agriculture, 2016, Agriculture Facts 2016, agri.idaho.gov.

Li, D., Wrzesien, M.L., Durand, M., Adam, J., Lettenmaier, D.P., 2017, How much runoff originates as snow in the western United States, and how will that change in the future?, Geophysical Research Letters, v. 44, no. 12, p. 6163-6172.

Nash, J. E., and J. V. Sutcliffe, 1970: River flow forecasting through conceptual models. Part I: A discussion of principles. Journal of Hydrology, v 10, p 282–290.

Pagano, T., Garen, D., Sorooshian, S., 2004, Evaluation of Official Western U.S. Seasonal Water Supply Outlooks, 1922-2002, Journal of Hydrometeorology, v. 5, p. 896-909.

Pagano, T., 2005, ForecastSkillNSApr1, https://www.wcc.nrcs.usda.gov/ftpref/data/ water/ forecast/.

PRISM Climate Group, Oregon State University, http://prism.oregonstate.edu, created 4 Feb 2004.

Trujillo, E. and Molotch, N. P., 2014, Snowpack regimes of the Western United States, Water Resources Research, v. 50, no. 7, p. 5611-5623.

U.S. Geological Survey, 2012, The StreamStats program, online at http://streamstats.usgs.gov.

Wilks, D. S., 1995, Statistical Methods in the Atmospheric Sciences: An Introduction, Academic Press, 467 pp.

Figures

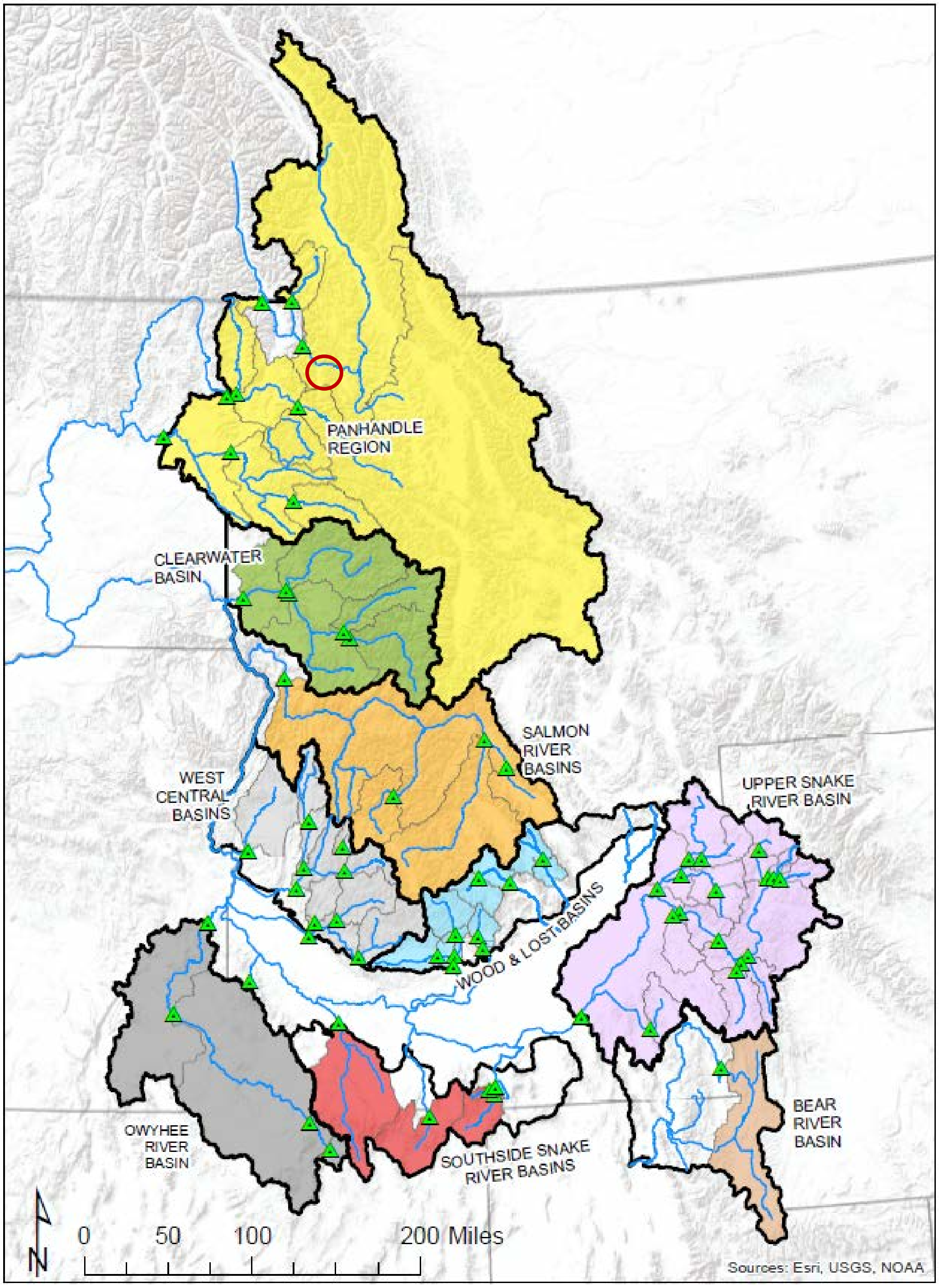


Figure 1: Map of Idaho forecast points (green triangles) and the contributing drainage areas to these points. Composite forecast areas are color coded, outlined in bold, and labeled. Major river networks are also shown as blue lines. Forecast data from the Kootenai at Leonia forecast point (circled in red) in the Panhandle region are presented in Table 2.

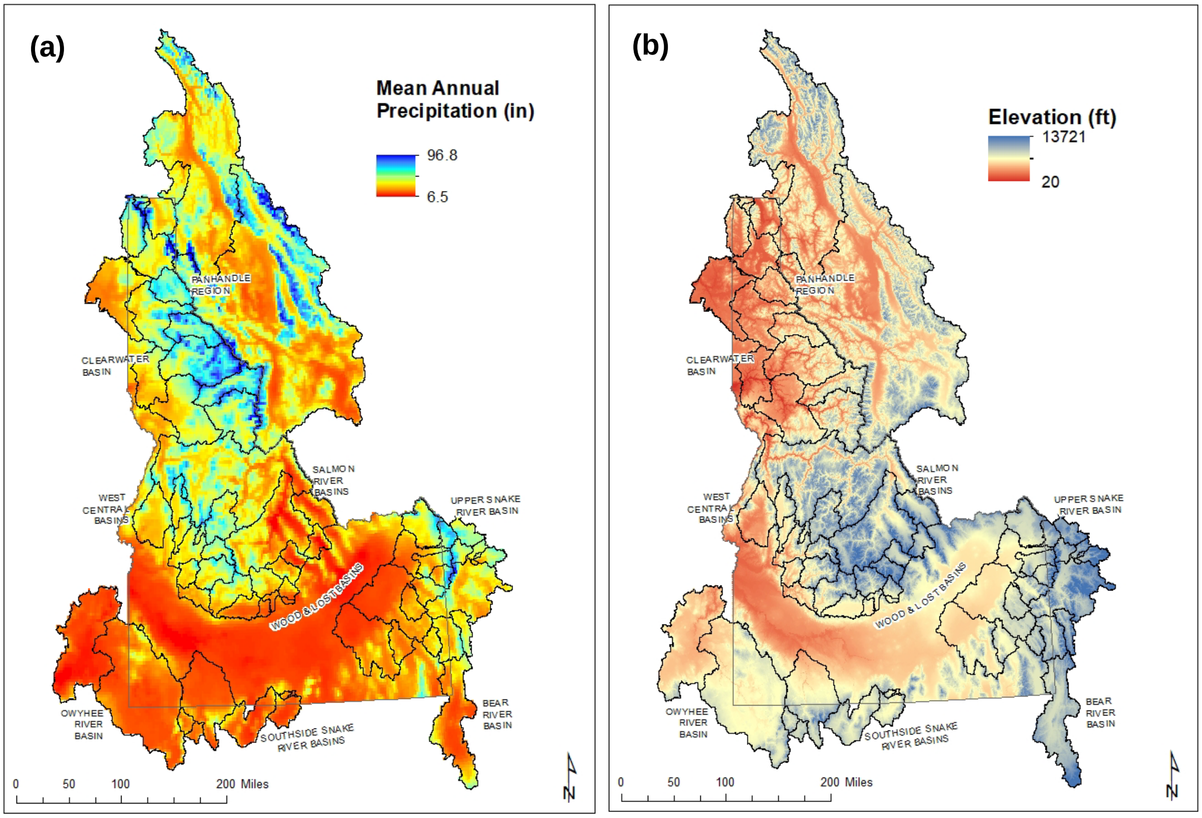


Figure 2: (a) Average annual precipitation in Idaho (PRISM Climate Group, 2004), (b) elevation map of Idaho

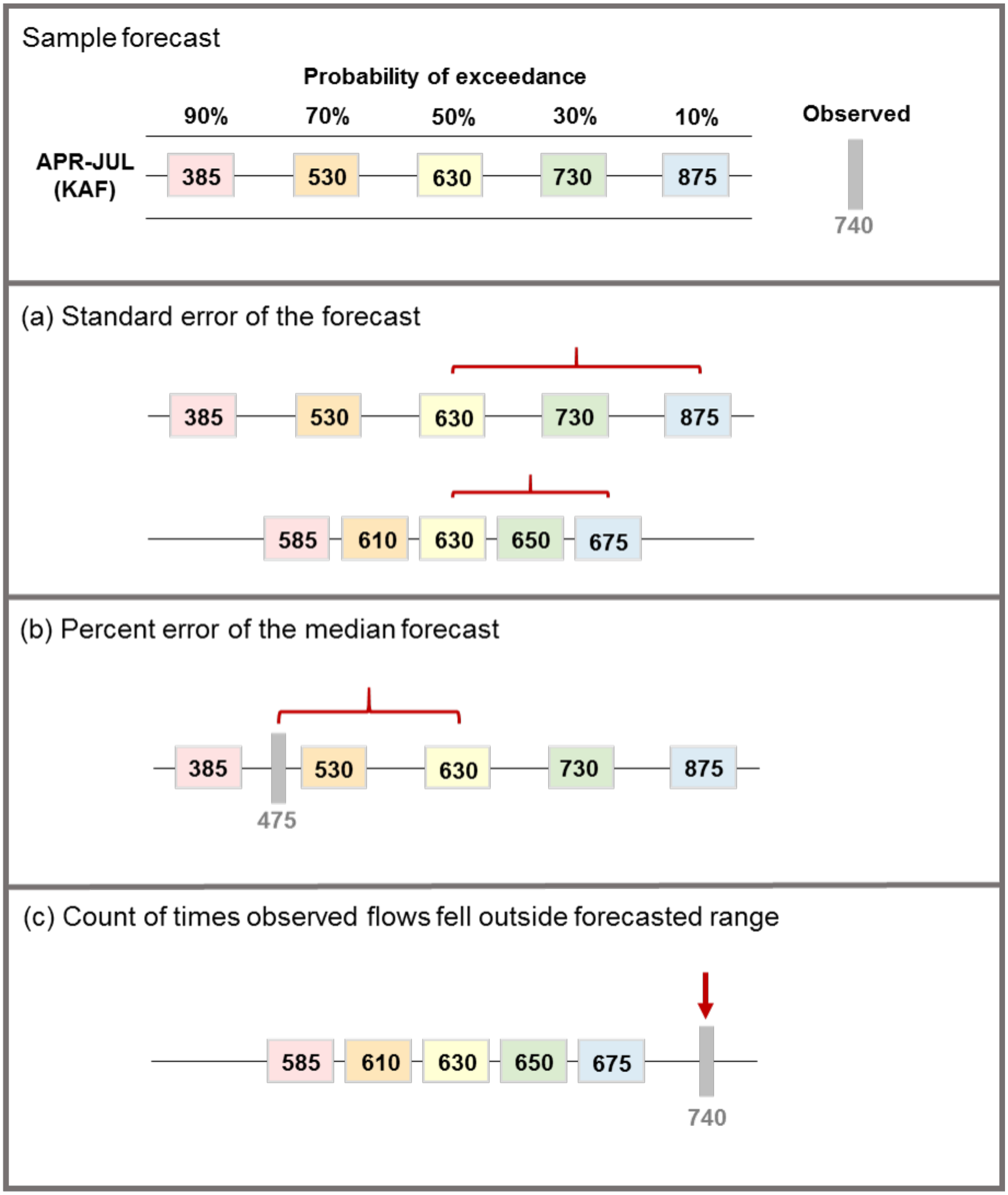


Figure 3: Illustration of error metrics used to quantify forecast performance, including (a) the standard error of the forecast equation, (b) the percent error of the median forecast, and (c) the count of times observed flows fell outside the forecasted range of volumes.

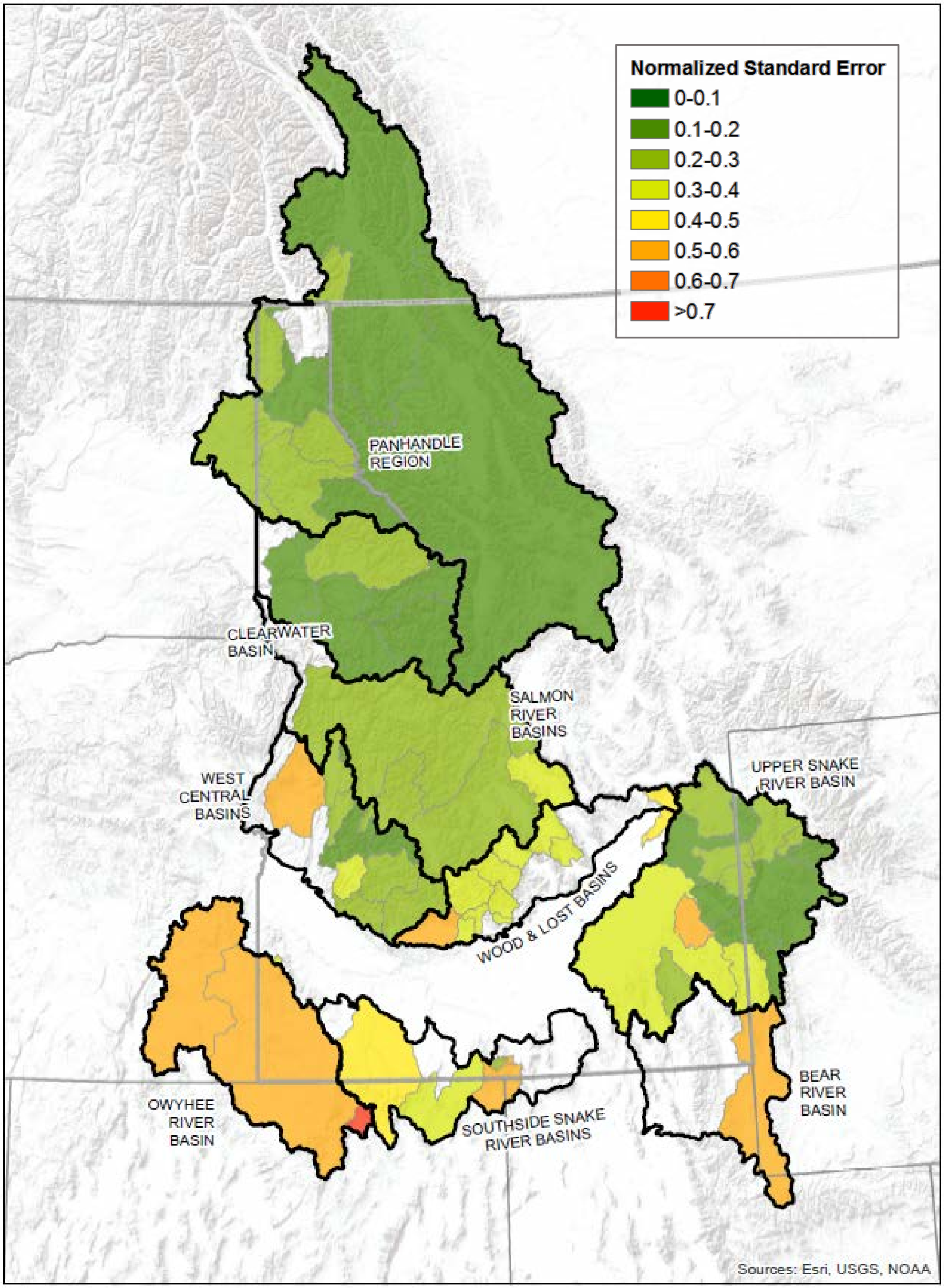
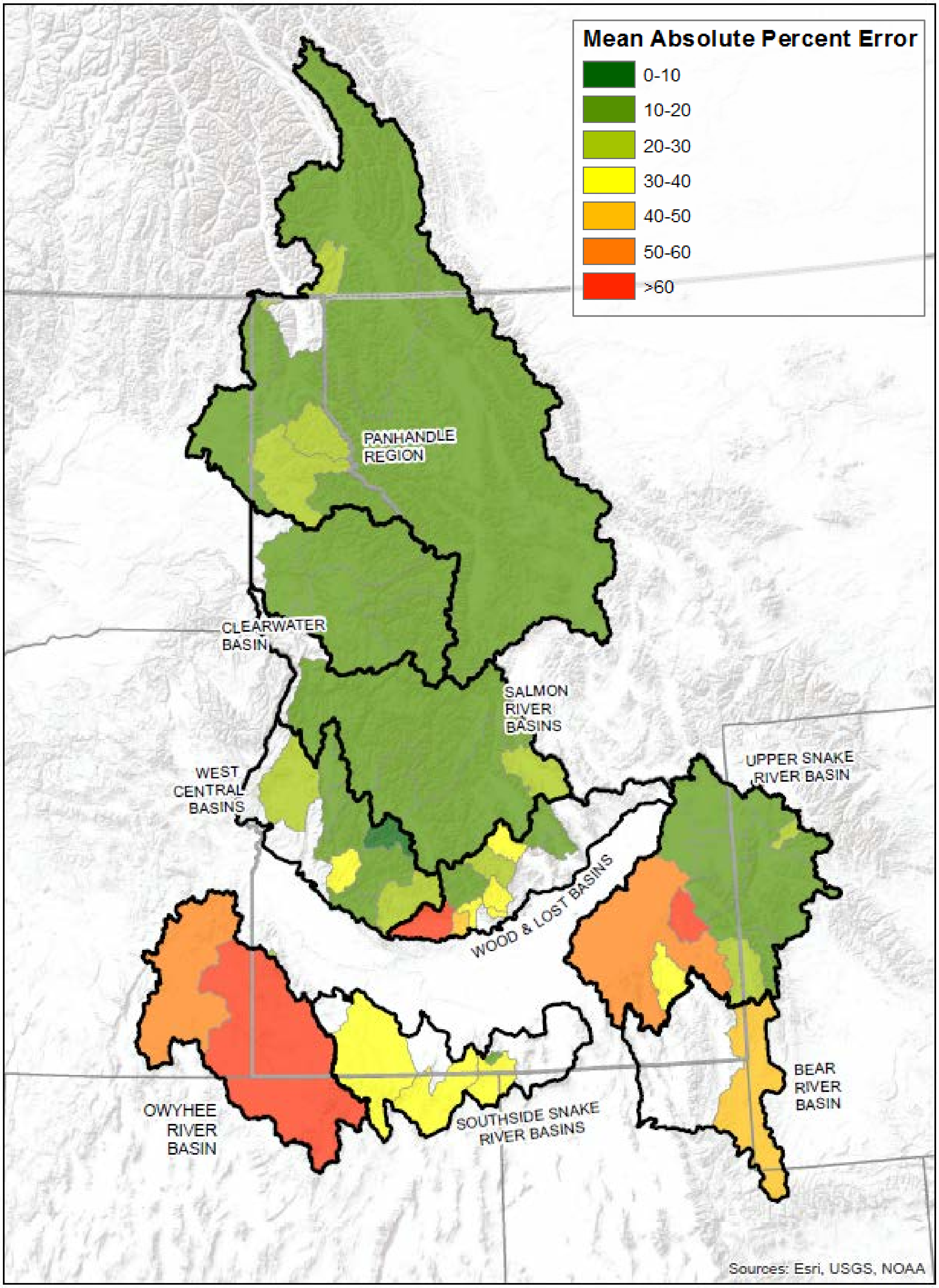


Figure 4: Map of 10-year mean normalized standard errors in the April 1 forecasts for the April-July period. Basins with the highest uncertainty in their forecasts are in the southern portion of the state and line the Snake River Plain.



Sources: Esri, USGS, NOAA

Figure 5: Map of 10-year (2008-2017) mean absolute percent error of the April 1 median forecast (50% exceedance) for the April-July forecast period, in relation to the cumulative observed flow volume.

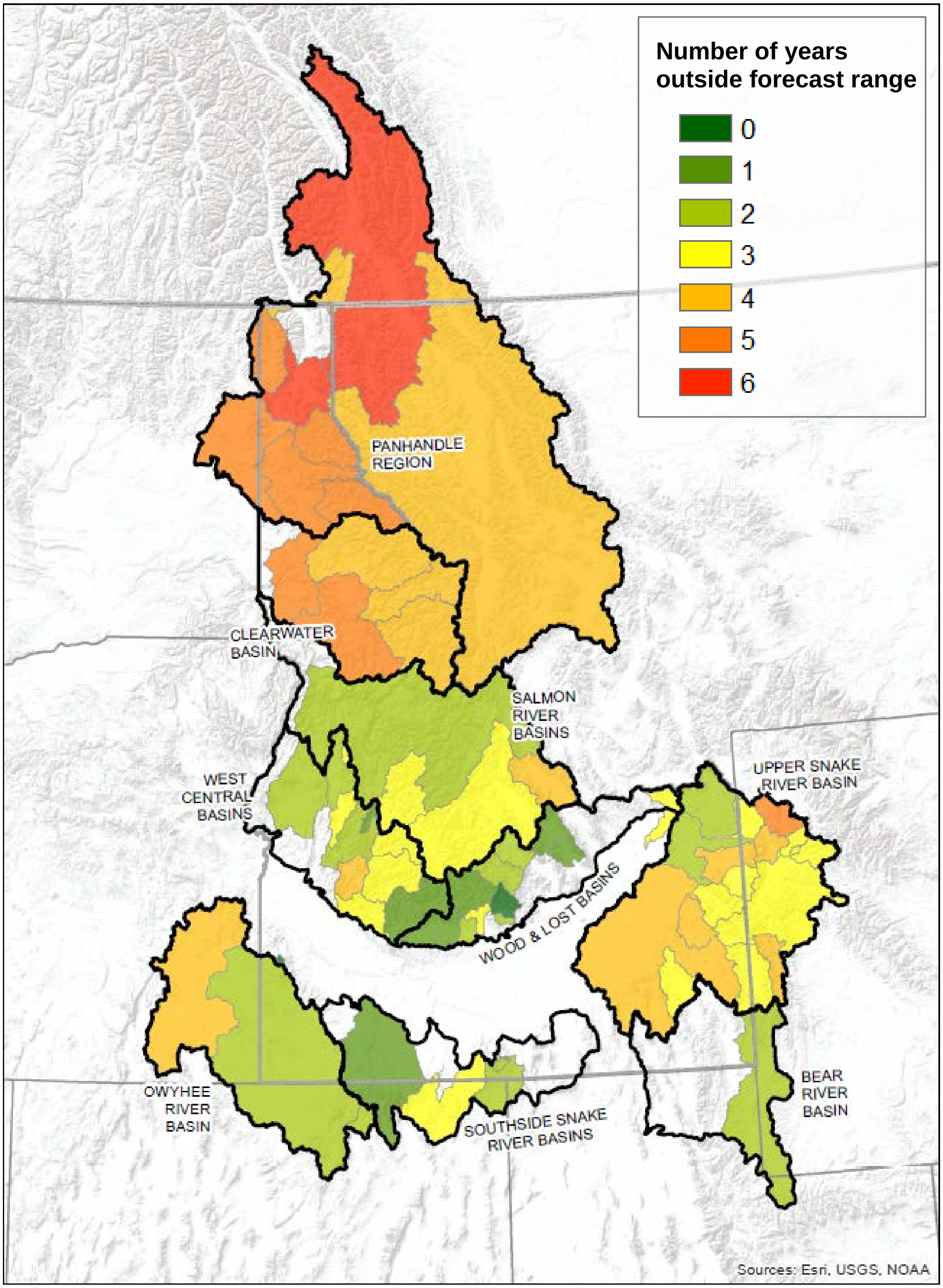


Figure 6: Map of 10-year out-of-forecast-range counts for each forecast basin. A "count" represents any year that the observed flows fell below the minimum forecasted volume (90% exceedance) or above the maximum forecasted volume (10% exceedance).

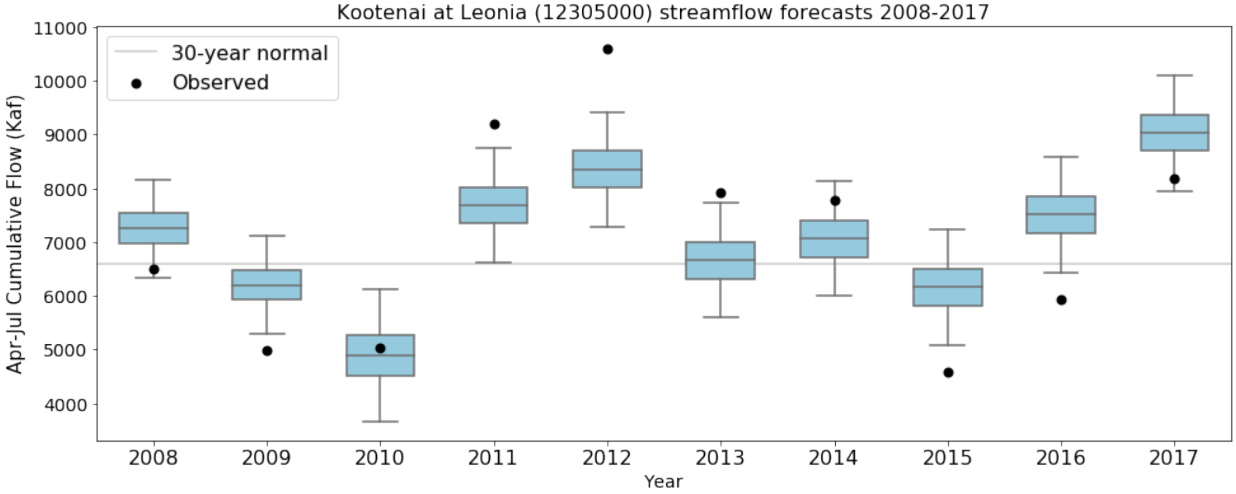


Figure 7: Forecasts and observations for Kootenai at Leonia (12035000), located in the Panhandle region. Lower and upper whiskers represent the 90% exceedance and 10% exceedance forecast volumes, respectively. Lower and upper bounds of boxes represent the 70% and 30% exceedance volumes, and the center line of box represents the 50% exceedance volume.

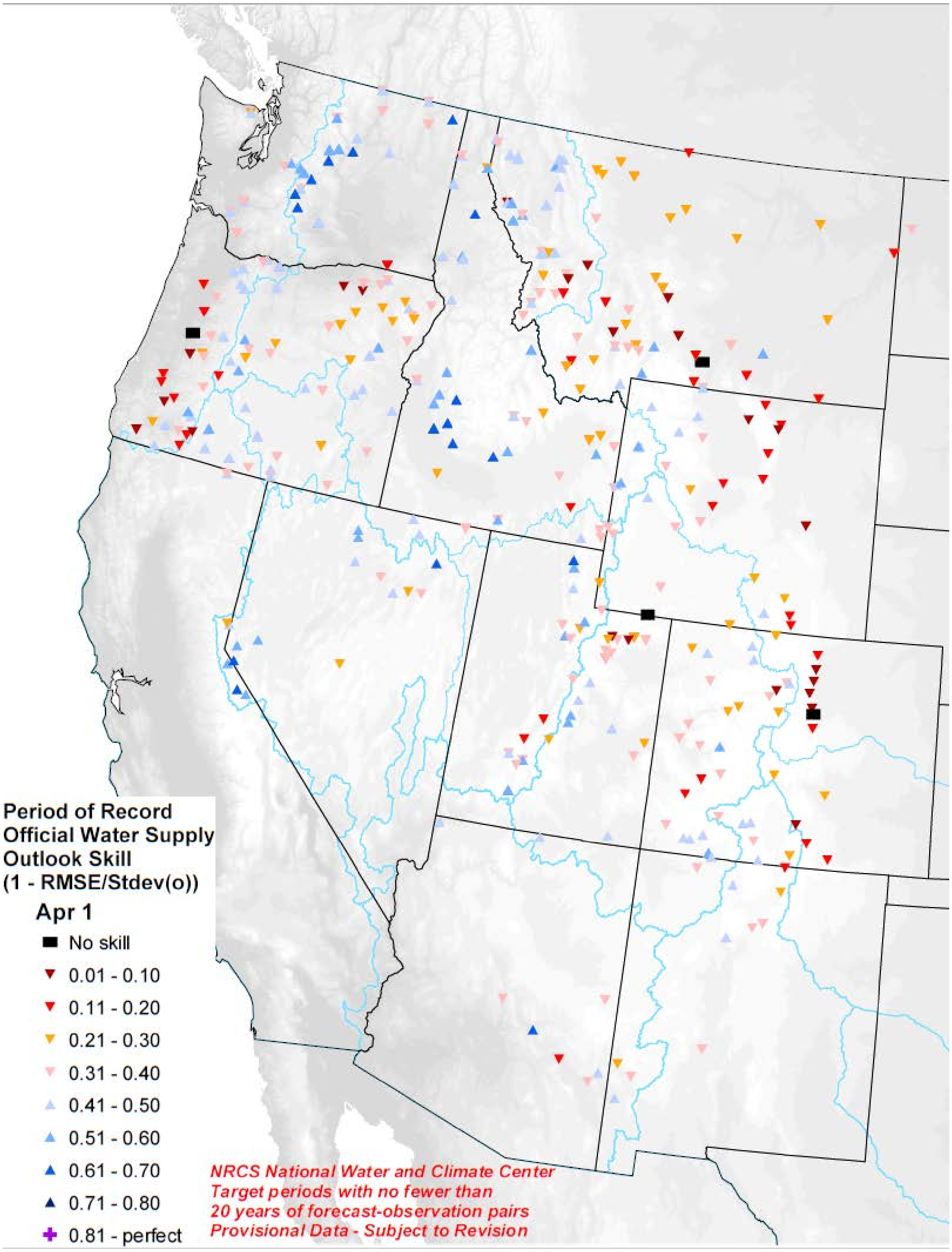


Figure 8: Map of NS skill score values for western U.S. NRCS streamflow forecasts, period of record through 2005, produced by Pagano (2005).

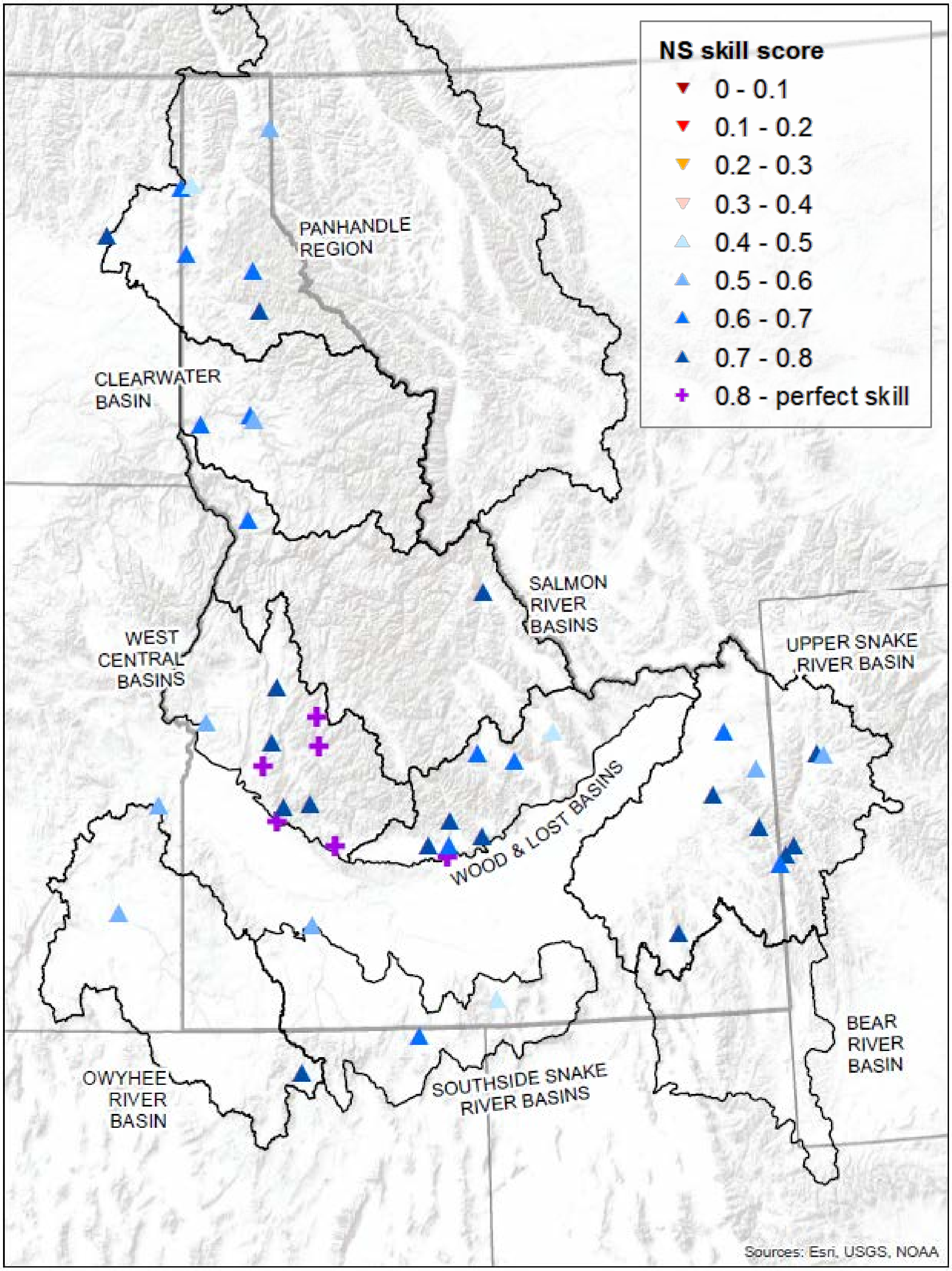


Figure 9: Nash-Sutcliffe skill score values for 46 forecast points.

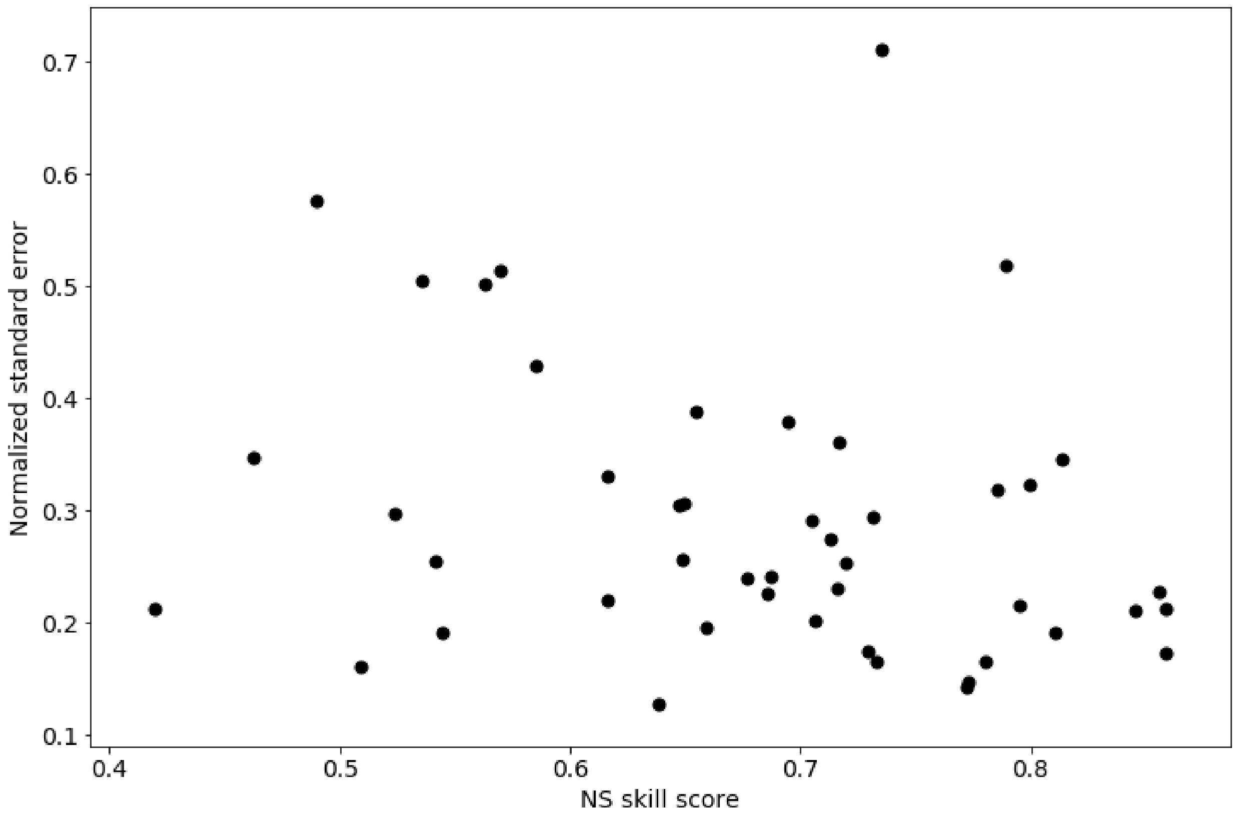


Figure 10: Plot of normalized standard error vs NS skill score.

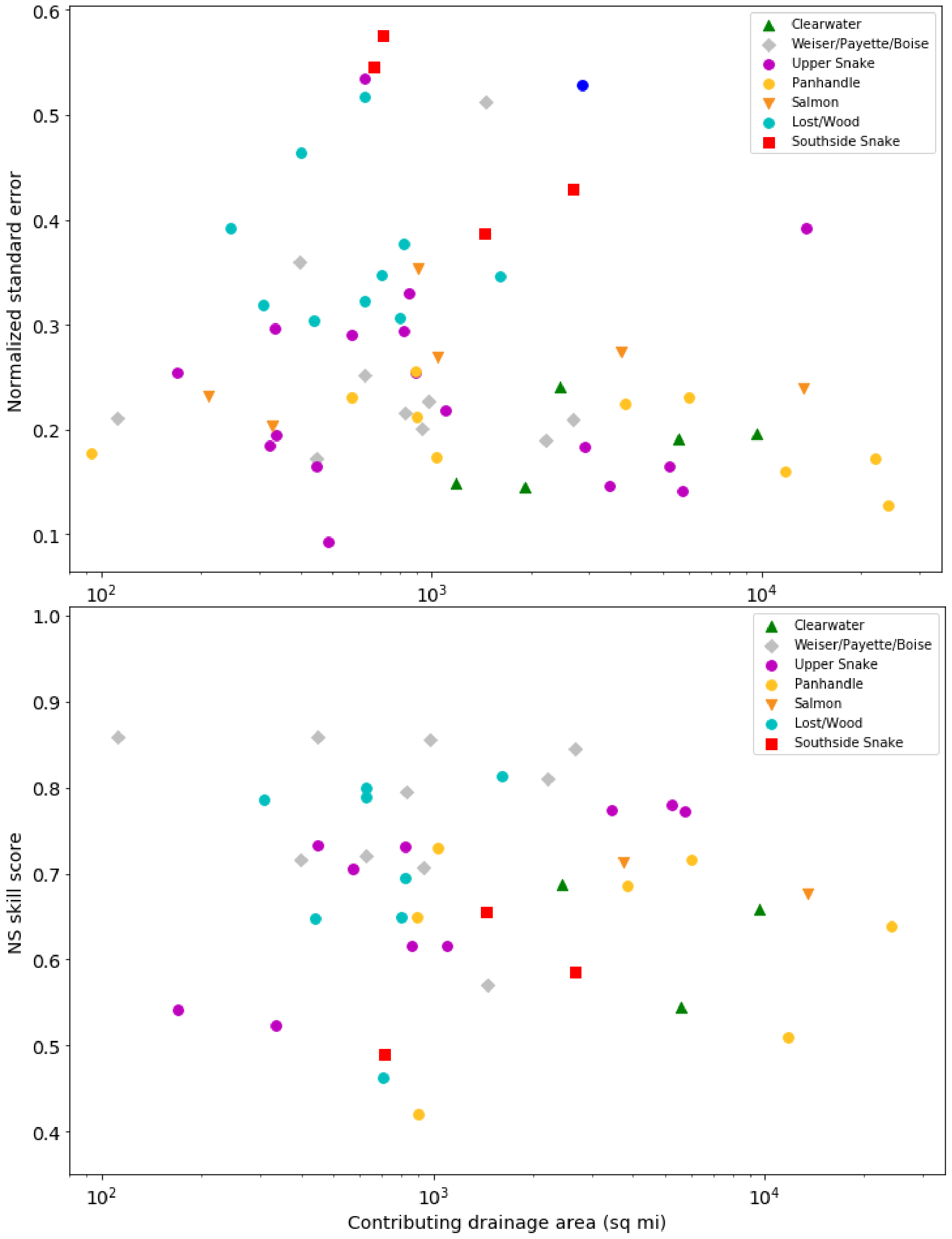


Figure 11: Contributing drainage area plotted against (a) normalized standard error and (b) NS skill score.

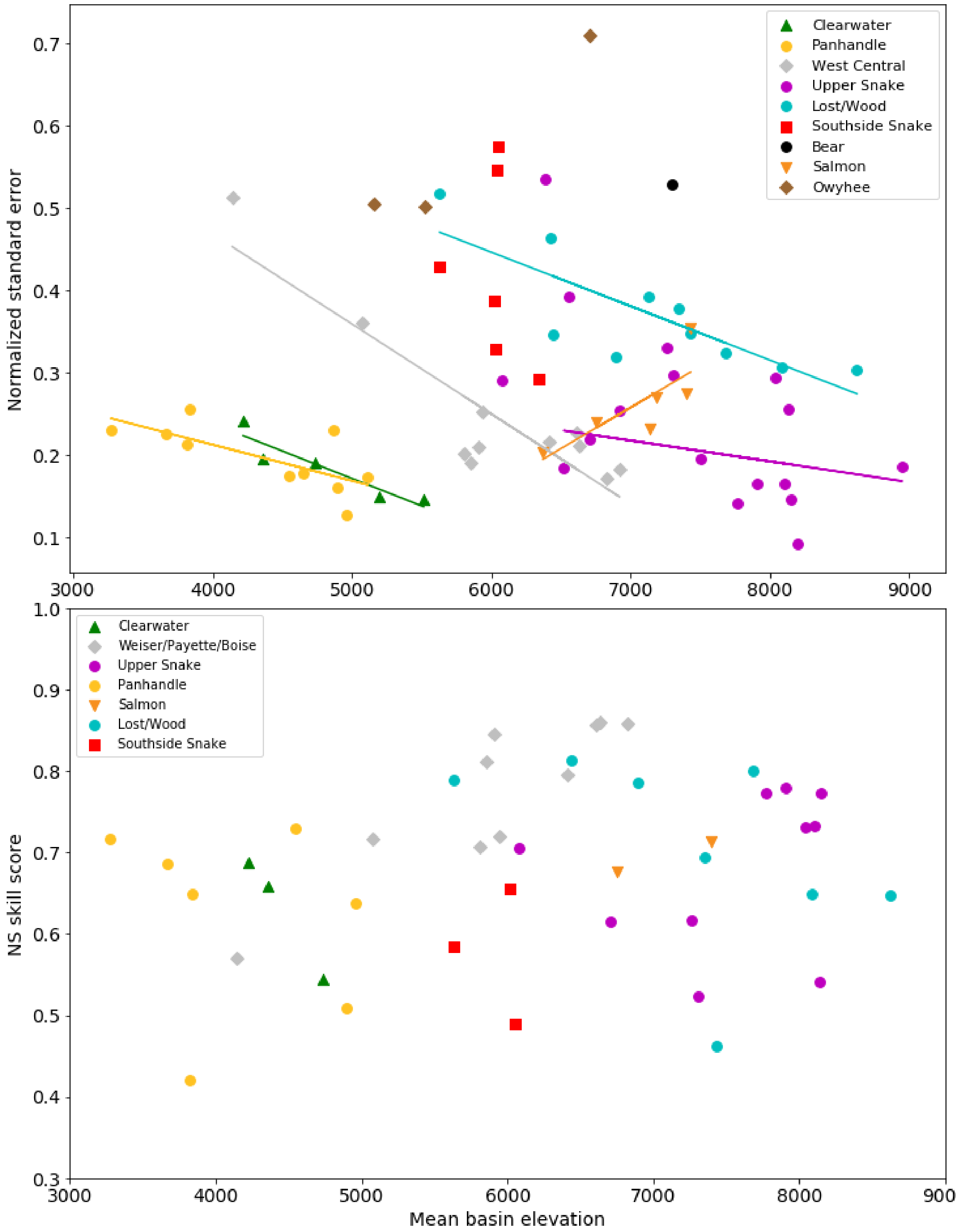


Figure 12: Mean basin elevation plotted against (a) normalized 10-year (2008-2017) mean standard error of April 1 forecasts and (b) 20-year (1998-2017) NS skill score for each forecast point. Each point shows the mean elevation of the contributing drainage area to a forecast point and the associated error parameter. Points are colored by greater forecast area (Figure 1). (a) also includes fitted regression lines for each of the greater forecast area groupings.

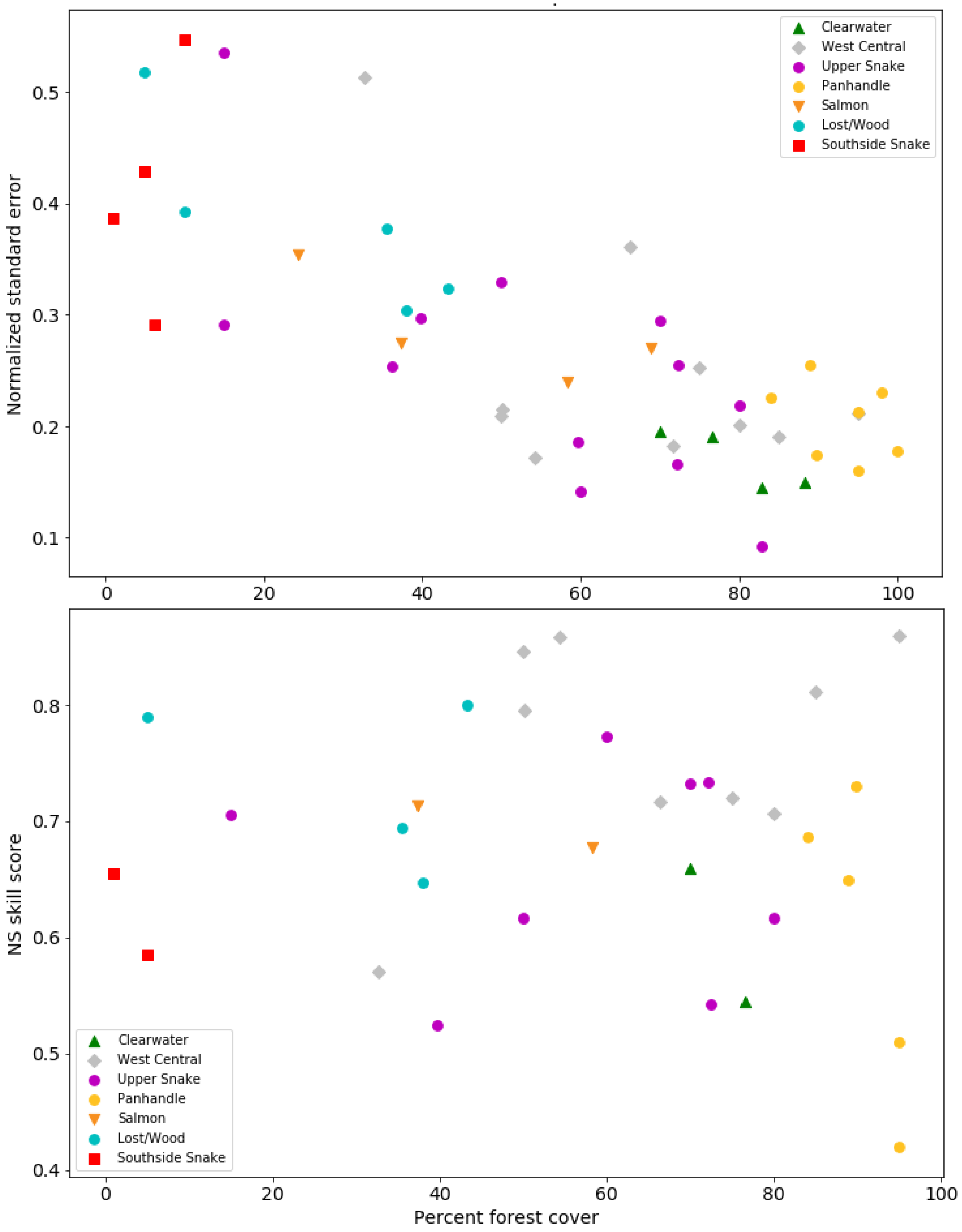


Figure 13: (a) Normalized 10-year (2008-2017) mean standard error of the April 1 forecast for each forecast point vs percent of basin area covered by forest, and (b) 20-year (1998-2017) April 1 NS skill score vs percent forest cover.

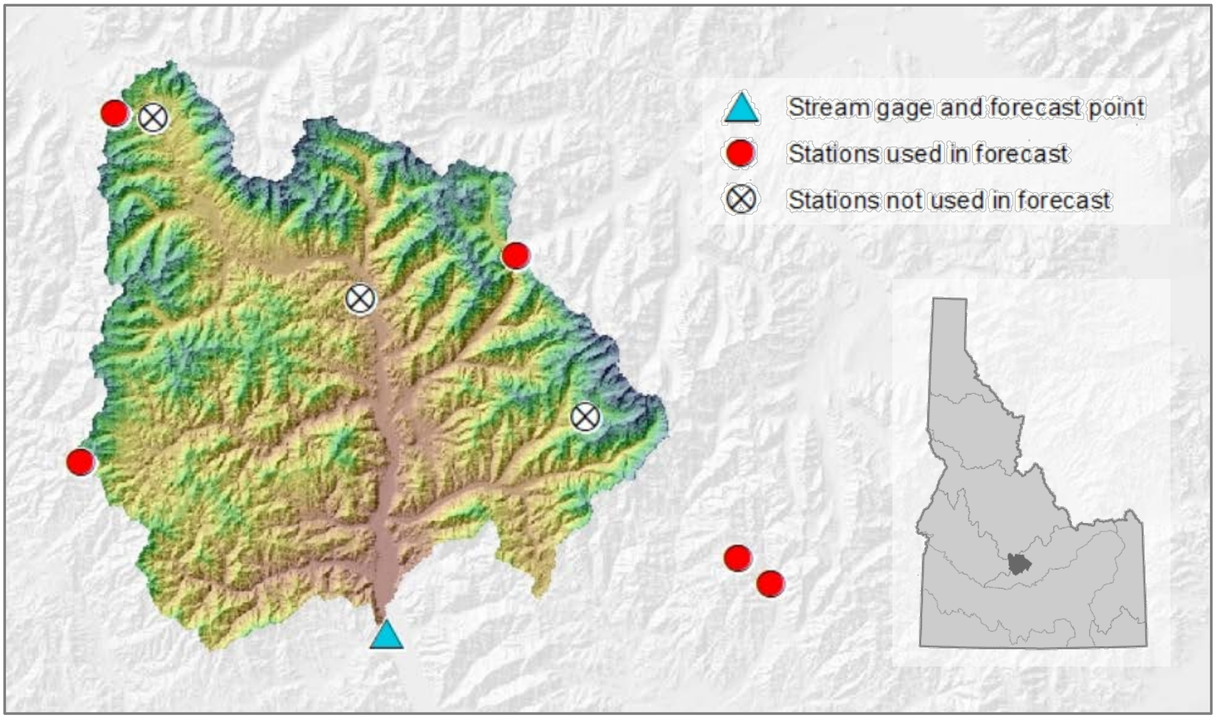
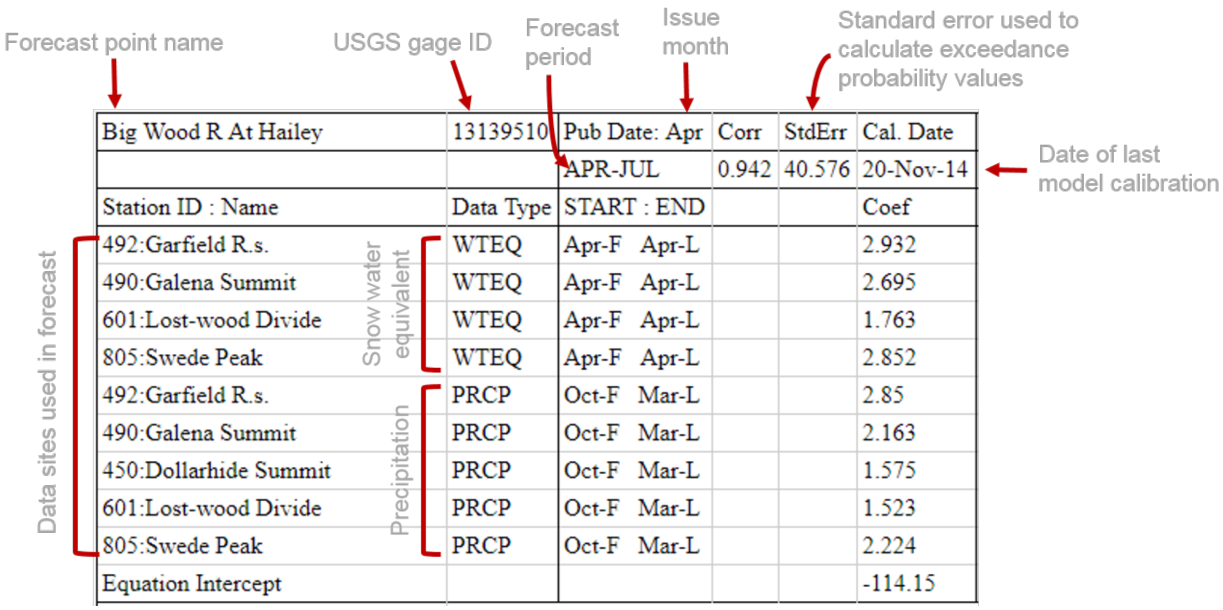


Figure 14: Map of the Big Wood River at Hailey forecast basin, displaying SNOTEL sites used in the forecast, SNOTEL sites available in the basin but not used in the forecast, and the stream gage that corresponds to the forecast point. Colors correspond to relative elevation (red = lowest, blue = highest). Inset shows location of contributing drainage area.



Big Wood R At Hailey	13139510	Pub Date: Apr	Corr	StdErr	Cal. Date
		APR-JUL	0.942	40.576	20-Nov-14
Station ID : Name	Data Type	START : END			Coef
492:Garfield R.s.	WTEQ	Apr-F Apr-L			2.932
490:Galena Summit	WTEQ	Apr-F Apr-L			2.695
601:Lost-wood Divide	WTEQ	Apr-F Apr-L			1.763
805:Swede Peak	WTEQ	Apr-F Apr-L			2.852
492:Garfield R.s.	PRCP	Oct-F Mar-L			2.85
490:Galena Summit	PRCP	Oct-F Mar-L			2.163
450:Dollarhide Summit	PRCP	Oct-F Mar-L			1.575
601:Lost-wood Divide	PRCP	Oct-F Mar-L			1.523
805:Swede Peak	PRCP	Oct-F Mar-L			2.224
Equation Intercept					-114.15

Figure 15: Sample forecast model summary for Big Wood R at Hailey detailing the components of each forecast model equation. The location of this forecast point can be found on an inset map in Figure 14.

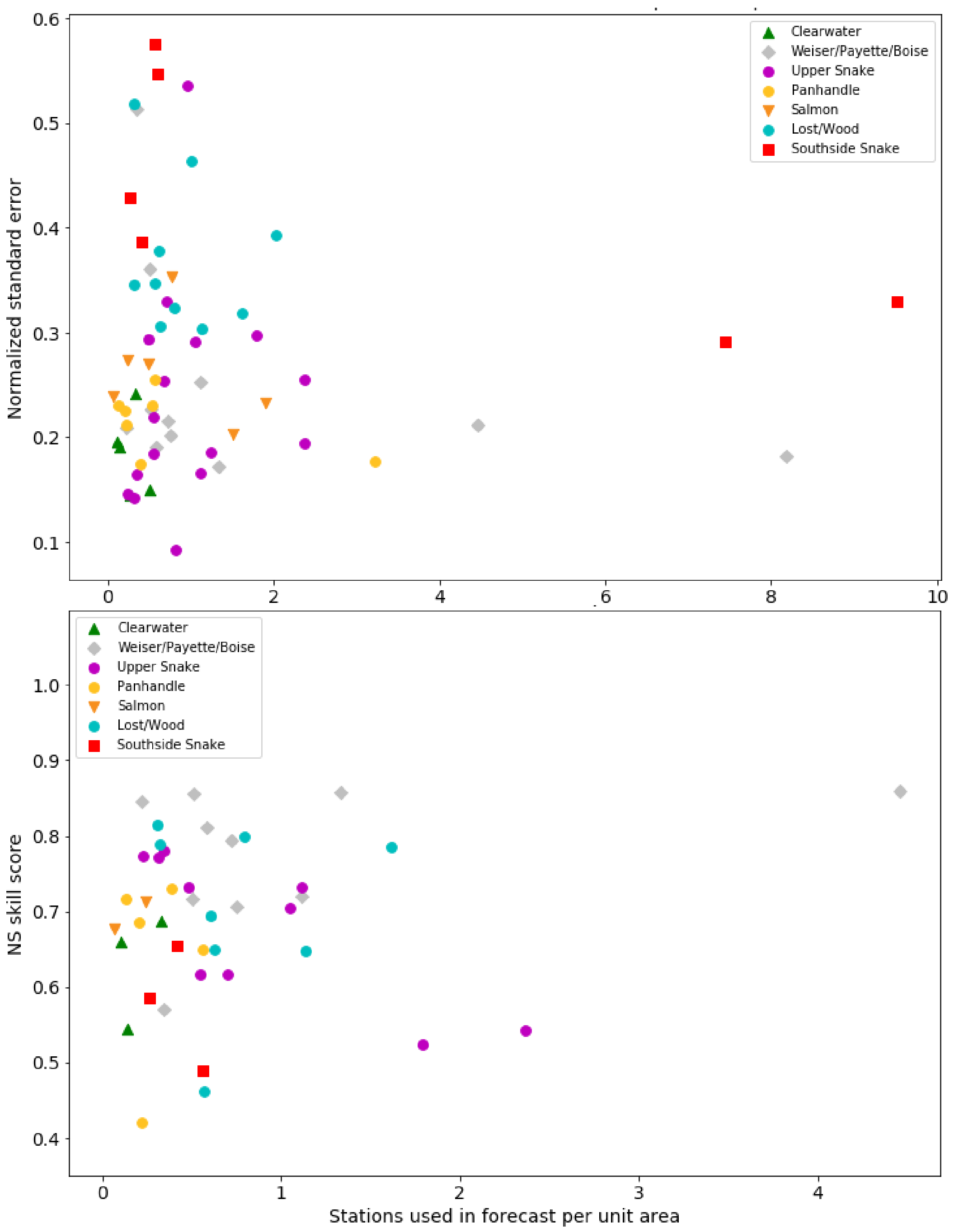


Figure 16: Number of data sites used in each forecast per 100 square miles vs. (a) normalized 10-year mean standard error of the April 1 forecast, and (b) 20-year April 1 NS skill score for each forecast point.

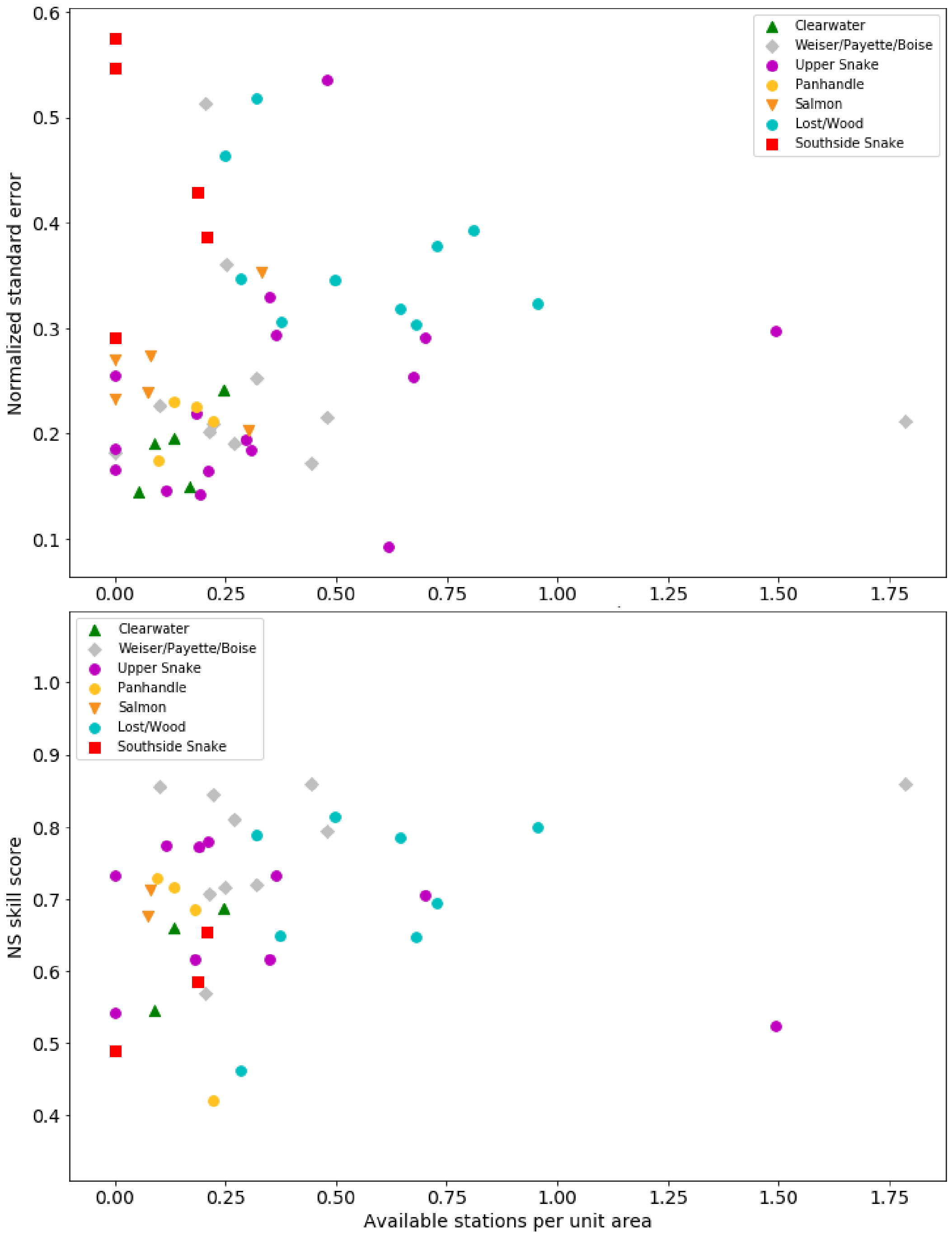


Figure 17: Number of data sites per 100 square miles available in each forecast basin vs. (a) normalized 10-year mean standard error of the April 1 forecast, and (b) 20-year April 1 NS skill score for each forecast point.

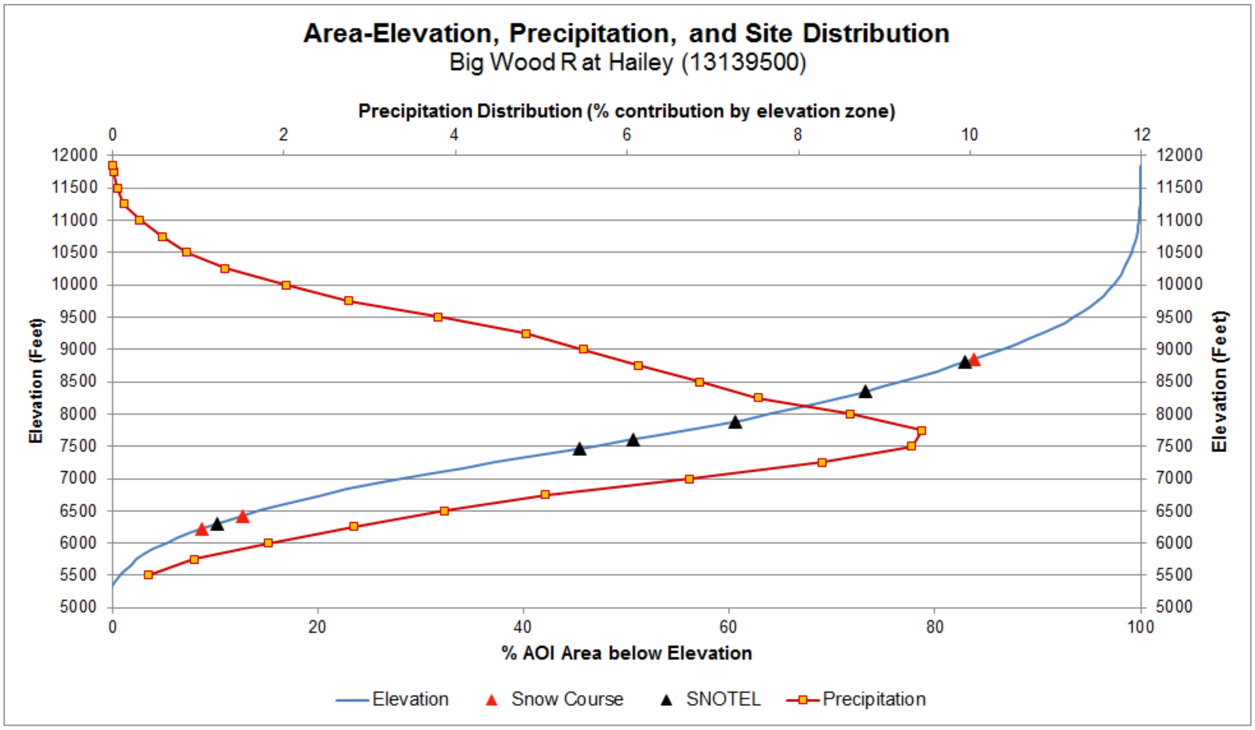


Figure 18: Area-elevation, precipitation distribution, and site distribution for the 'Big Wood R at Hailey' forecast point.

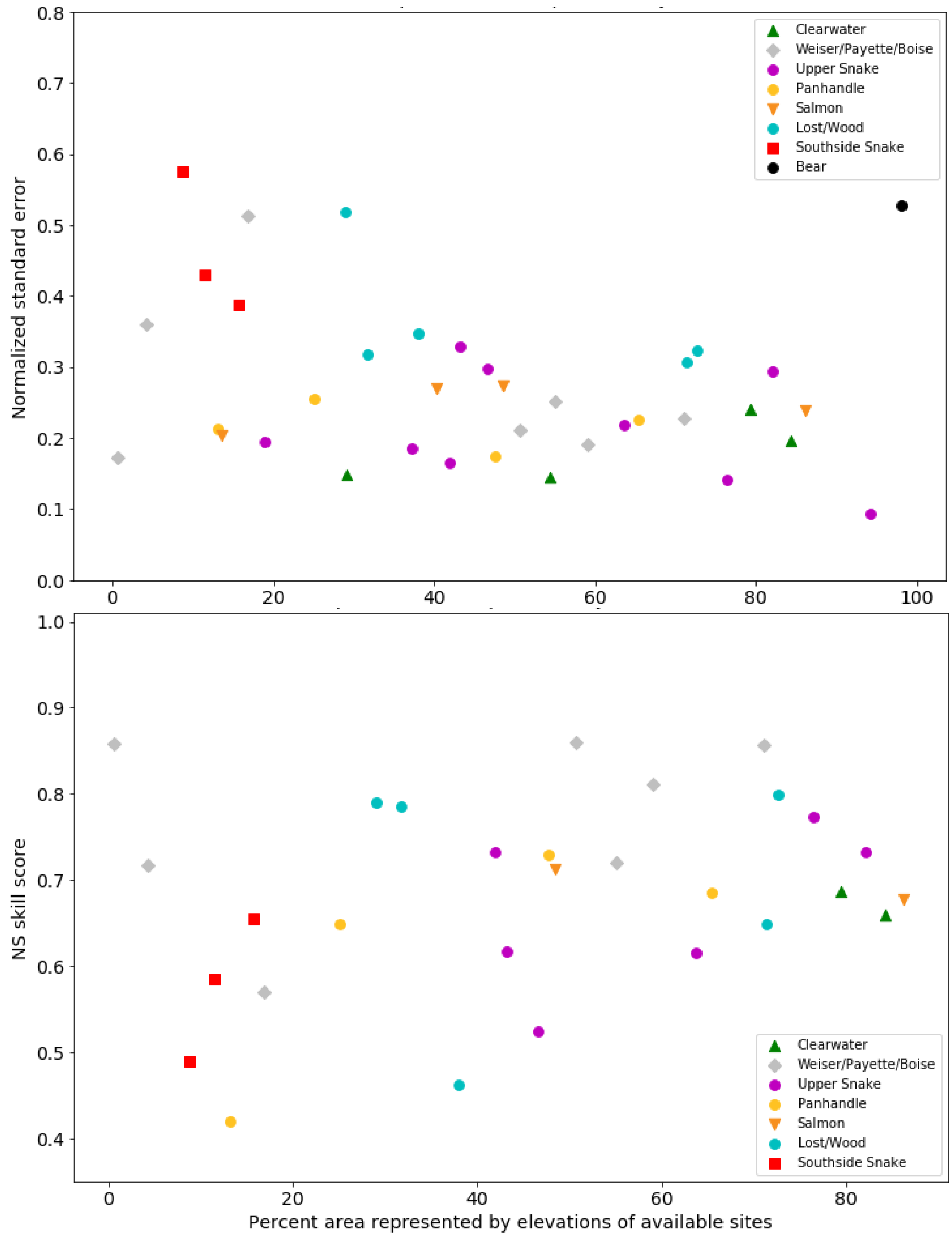


Figure 19: (a) Normalized standard error vs percent area represented by the elevations of available sites in a basin (n = 37), and (b) NS skill score vs percent area represented by the elevations of available sites in a basin (n = 29).

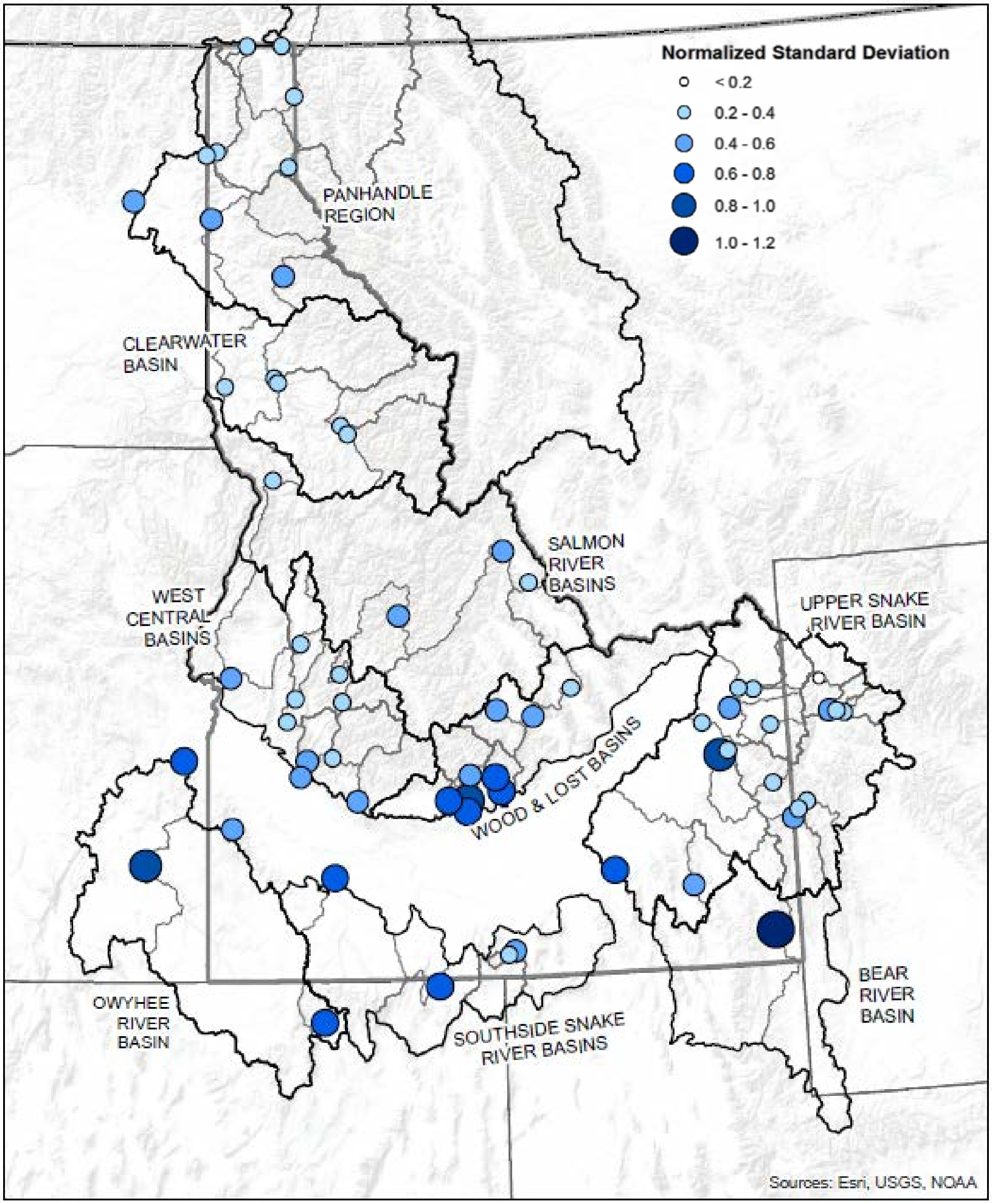


Figure 20: Standard deviation of observed April-July streamflow volumes (2008-2017) at each forecast point, normalized to the 30-year average April-July streamflow volume.

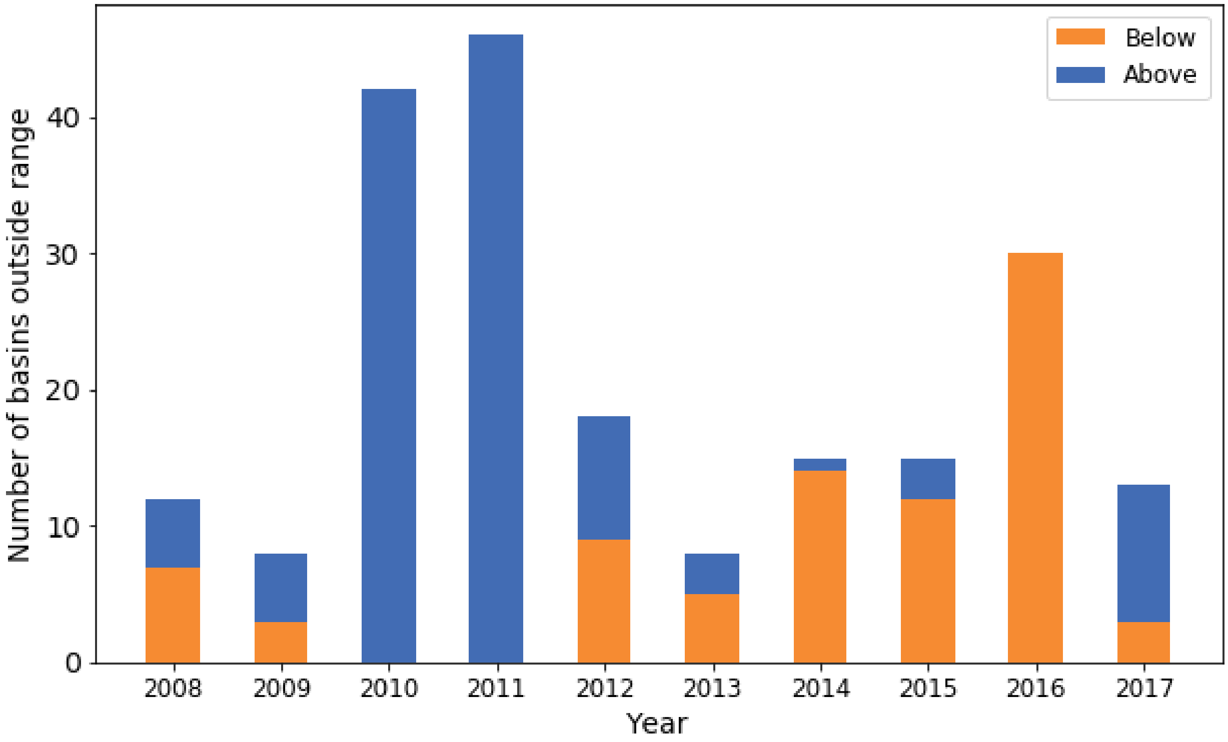


Figure 21: Count of basins in which observed flows fell either below the minimum forecast or above the maximum forecast, per year (total number of forecast basins = 69).

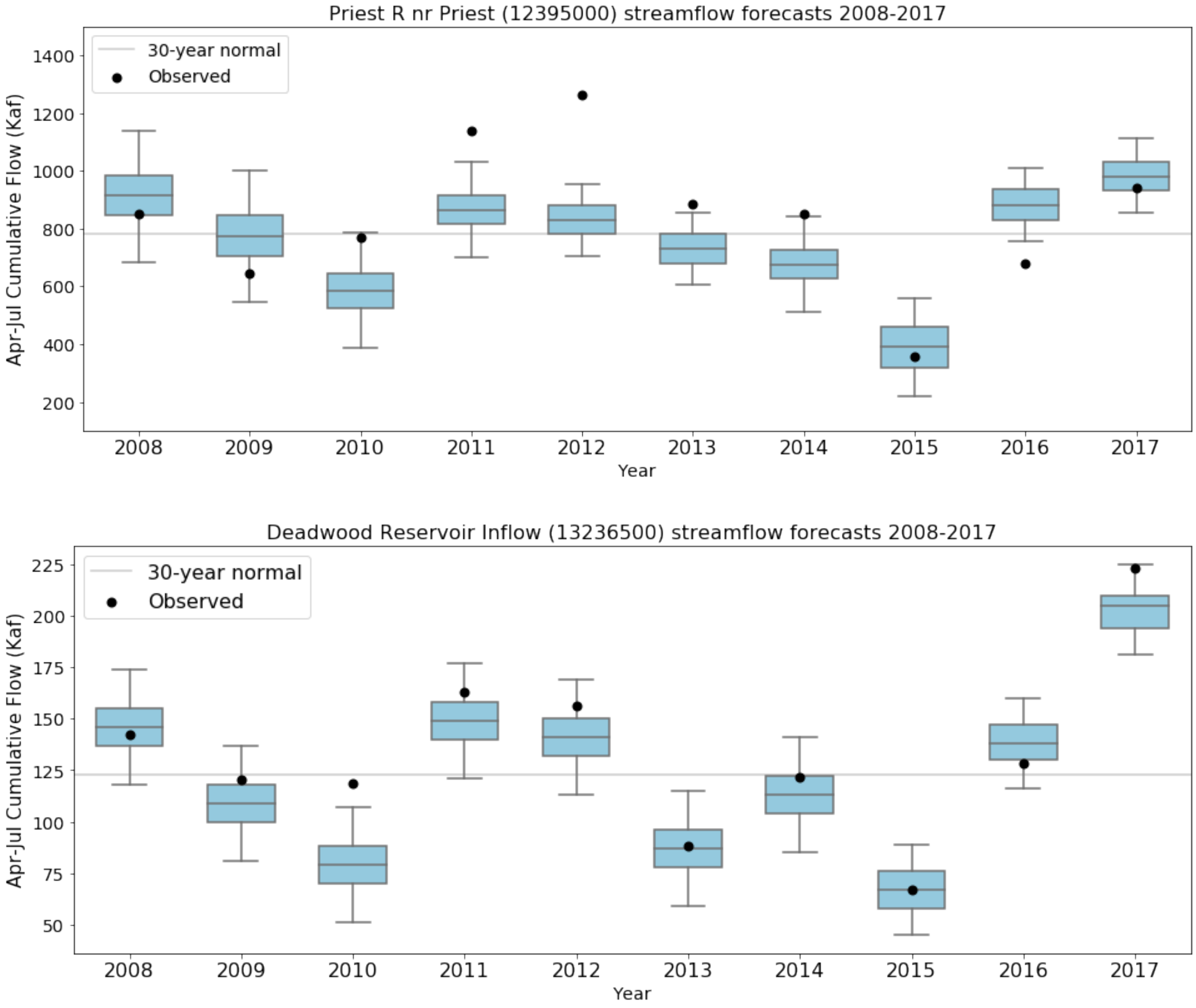
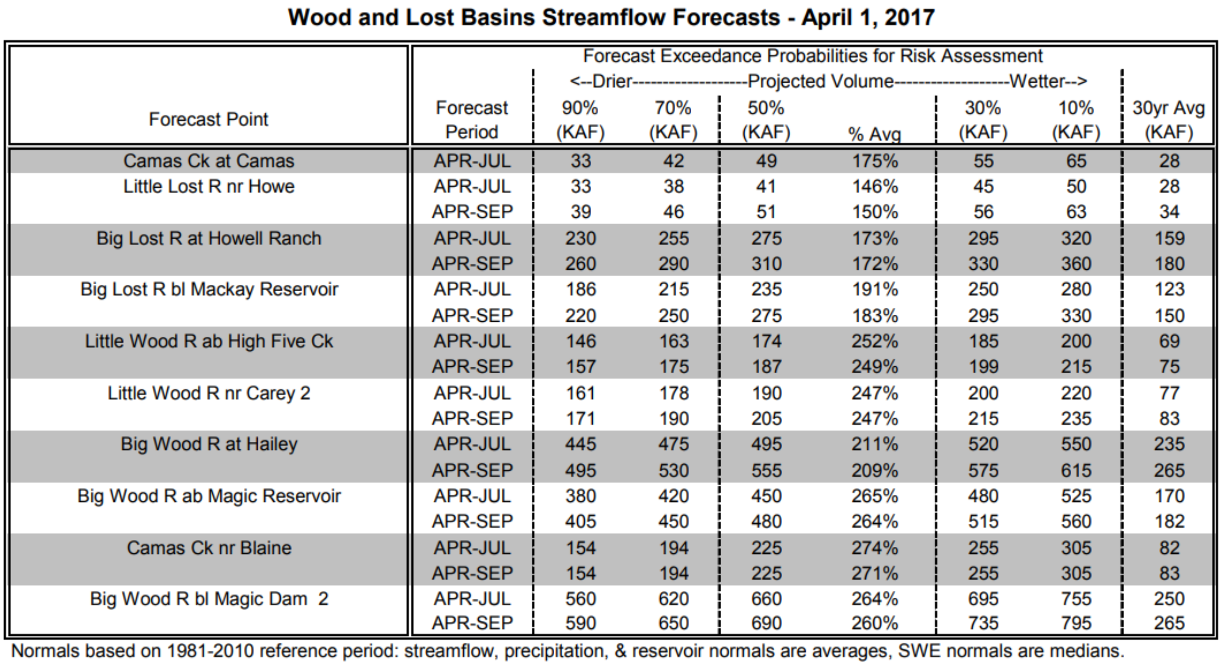


Figure 22: Exceedance probability forecasts and observations for (a) Priest R nr Priest, the forecast point with the lowest NS skill score, and (b) Deadwood Reservoir Inflow, the forecast point with the highest NS skill score.

Tables

Table 1: Sample NRCS April 1 streamflow forecast for the Wood and Lost River Basins. The ten forecast points correspond to USGS streamflow gages distributed throughout the Wood and Lost Basins. Cumulative flows are forecasted for two periods: April through July, and April through September. Units are in thousand acre-feet (KAF). The forecast table indicates the five forecast exceedance probability volumes (90%, 70%, 50%, 30%, and 10%), as well as the median forecast percent of average and the 30-year average (Idaho April 1 WSOR: *https://www.wcc.nrcs.usda.gov/ftpref/states/id/webftp/wsor/2017/ borid417.pdf).*



Wood and Lost Basins Streamflow Forecasts - April 1, 2017

		Forecast Exceedance Probabilities for Risk Assessment						
		<--Drier------------------Projected Volume------------------Wetter-->						
Forecast Point	Forecast Period	90% (KAF)	70% (KAF)	50% (KAF)	% Avg	30% (KAF)	10% (KAF)	30yr Avg (KAF)
Camas Ck at Camas	APR-JUL	33	42	49	175%	55	65	28
Little Lost R nr Howe	APR-JUL	33	38	41	146%	45	50	28
	APR-SEP	39	46	51	150%	56	63	34
Big Lost R at Howell Ranch	APR-JUL	230	255	275	173%	295	320	159
	APR-SEP	260	290	310	172%	330	360	180
Big Lost R bl Mackay Reservoir	APR-JUL	186	215	235	191%	250	280	123
	APR-SEP	220	250	275	183%	295	330	150
Little Wood R ab High Five Ck	APR-JUL	146	163	174	252%	185	200	69
	APR-SEP	157	175	187	249%	199	215	75
Little Wood R nr Carey 2	APR-JUL	161	178	190	247%	200	220	77
	APR-SEP	171	190	205	247%	215	235	83
Big Wood R at Hailey	APR-JUL	445	475	495	211%	520	550	235
	APR-SEP	495	530	555	209%	575	615	265
Big Wood R ab Magic Reservoir	APR-JUL	380	420	450	265%	480	525	170
	APR-SEP	405	450	480	264%	515	560	182
Camas Ck nr Blaine	APR-JUL	154	194	225	274%	255	305	82
	APR-SEP	154	194	225	271%	255	305	83
Big Wood R bl Magic Dam 2	APR-JUL	560	620	660	264%	695	755	250
	APR-SEP	590	650	690	260%	735	795	265

Normals based on 1981-2010 reference period: streamflow, precipitation, & reservoir normals are averages, SWE normals are medians.

Table 2: Example of the range of April 1 forecasted flows, observed flow volumes, and bin classification for 'Kootenai at Leonia', a forecast point in the Panhandle region with one of the highest out-of-range counts (Figure 1). Units are in thousand acre-feet (KAF). Bin values express the range the observed flows fall within: <90 for below the minimum (90% exceedance) forecast, 90-70 for between the minimum (90%) and the low (70%) forecast, and so on.

Year	**Min (90%)**	**Low (70%)**	**Mid (50%)**	**Upp (30%)**	**Max (10%)**	**Observed**	**Bin**
2008	6340	6970	7250	7530	8160	6497	90-70
2009	5290	5920	6200	6480	7110	4977	**<90**
2010	3660	4500	4890	5280	6120	5043	50-30
2011	6610	7340	7680	8020	8750	9187	**>10**
2012	7280	8010	8350	8690	9420	10589	**>10**
2013	5590	6320	6660	7000	7730	7929	**>10**
2014	5990	6720	7060	7400	8130	7787	30-10
2015	5090	5820	6160	6490	7230	4585	**<90**
2016	6430	7170	7510	7840	8580	5930	**<90**
2017	7950	8690	9030	9360	10100	8171	90-70

Appendices

Appendix A: The NRCS streamflow forecasting procedure

This overview summarizes Chapter 7 of the National Engineering Handbook, Part 622, on Water Supply Forecasting (Garen, 2011).

Forecast points

Forecast points are established based on the needs of end users, sufficient data availability, and feasibility of producing a skillful forecast. Forecast points are located at active streamgaging stations, typically in basins where most runoff originates as snowmelt. In basins where water management activities such as reservoirs and diversions significantly alter streamflows, observed flows are adjusted based on available information to reflect natural conditions and to allow for comparison with the forecasted volumes.

NRCS forecast hydrologists are assigned a group of forecast areas that each correspond to a major river basin. Several forecast points are located within each forecast area, and are located on the major rivers and tributaries of importance.

Data used in forecasts

The primary sources of data used in streamflow forecasts are the NRCS SNOTEL network and the USGS streamgage network. Other supplementary data sources less frequently used include the NRCS manual snow course measurements, weather stations managed by other agencies, and larger scale climate models. The most commonly used predictor variables in the forecast model equations include snow water equivalent and precipitation obtained from data sites in and near the forecast basin. Other predictor variables less commonly used include antecedent streamflow and climate teleconnection indices (e.g. ENSO). At least 10 years of continuous data (from both the predictor variables and streamflow, the predictand) is used to develop the model equation, but most streamflow forecast models use 30 years of data.

Statistical forecasting methods

The procedure currently used by all NRCS forecasters was developed in-house by David Garen in the early 1990s and is fundamentally based on multiple linear regression. Specifically, the NRCS procedure primarily employs a principal components regression technique. The main purpose of the principal components regression is to address potential multicollinearity issues and minimize the number of highly intercorrelated predictor variables used in the model equation. For example, if the value of snow water equivalent at Station A can be readily predicted based on its relationship to the snow water equivalent at Station B, the utilization of both predictor variables in the model equation can produce unreliable or inconsistent forecasts. Instead, the principal components method summarizes a set of intercorrelated predictor variables into a single “component,” which can then be used as a predictor variable in the model equation. Ideally, the predictor variables (whether summarized "components" or raw input data) are uncorrelated and provide unique information about the hydrologic system of interest.

An alternative statistical technique is used in cases of missing or incomplete data records. This technique is called Z-score regression, and though it does not account for intercorrelations between predictor variables, it produces similar results to the principal components technique in cases of complete data records.

Forecast production

NRCS personnel have also developed software to support both of the above forecast techniques in a visual-based interface. The software is called VIPER (Visual Interactive Prediction and Estimation Routines) and is described as an "application with data retrieval, visualization, and forecast calibration and execution functions" (NWCC Technical Note: Statistical Techniques in VIPER).

An automated search algorithm in VIPER identifies a set of combinations of predictor variables with the lowest standard errors of the resulting regression equations. The forecaster then chooses the model from this list that he or she determines is the most parsimonious and physically meaningful. A separate model equation is developed for each forecast point and for each month that forecasts are issued. The output of the model is used as the median value of the streamflow forecast.

Once a model is chosen, the forecaster performs a cross-validation procedure called the jackknife test. The jackknife test iterates over the calibration dataset, leaves one year of data out, and uses the forecast model to predict the streamflow from the excluded year. A new "jackknifed" standard error is computed from the cross-validation procedure and is typically larger than the standard error computed from the original model equation. The jackknifed standard error is then used to calculate the minimum, lower, upper, and maximum exceedance probability values.

Appendix B: List of forecast points, their associated forecast areas, and basin metrics.

FORECAST POINT NAME	FORECAST AREA	CONTRIBUTING DRAINAGE AREA (SQ MI)	MEAN BASIN ELEVATION (FT)
BEAR RIVER BELOW STEWART DAM NEAR MONTPELIER, ID	BEAR	2853	7295
CLEARWATER RIVER AT OROFINO ID	CLEARWATER	5580	4736
CLEARWATER RIVER AT SPALDING ID	CLEARWATER	9640	4360
DWORSHAK RES NR AHSAHKA ID	CLEARWATER	2440	4220
LOCHSA RIVER NR LOWELL ID	CLEARWATER	1180	5197
SELWAY RIVER NR LOWELL ID	CLEARWATER	1910	5515
BIG LOST RIVER AT HOWELL RANCH NR CHILLY ID	LOST/WOOD	440	8626
BIG LOST RIVER BL MACKAY RES NR MACKAY ID	LOST/WOOD	801	8090
BIG WOOD R AB MAGIC RESERVOIR	LOST/WOOD	824	7347
BIG WOOD RIVER AT HAILEY ID	LOST/WOOD	628	7686
BIG WOOD RIVER BL MAGIC DAM NR RICHFIELD ID	LOST/WOOD	1611	6440
CAMAS CK AT CAMAS	LOST/WOOD	400	6429
CAMAS CREEK NR BLAINE ID	LOST/WOOD	626	5630
LITTLE LOST RIVER NR HOWE ID	LOST/WOOD	703	7430
LITTLE WOOD RIVER AB HIGH FIVE CREEK NR CAREY, ID	LOST/WOOD	246	7130
LITTLE WOOD RIVER NR CAREY ID	LOST/WOOD	310	6890
OWYHEE R BL OWYHEE DAM 2	OWYHEE	11160	5160
OWYHEE R NR GOLD CK 2	OWYHEE	209	6710
OWYHEE R NR ROME	OWYHEE	7690	5520
BOUNDARY CREEK NR PORTHILL ID	PANHANDLE	93	4650
CLARK FORK AT WHITEHORSE RAPIDS NR CABINET ID	PANHANDLE	22073	5115
COEUR D ALENE RIVER ABOVE N FORK NEAR ENAVILLE ID	PANHANDLE	895	3836
KOOTENAI RIVER AT LEONIA ID	PANHANDLE	11740	4900
MOYIE RIVER AT EASTPORT ID	PANHANDLE	570	4870
PEND OREILLE RIVER AT NEWPORT ID	PANHANDLE	24200	4960
PRIEST RIVER NR PRIEST RIVER ID	PANHANDLE	902	3820
SPOKANE RIVER AT LONG LAKE	PANHANDLE	6020	3270
SPOKANE RIVER NR POST FALLS ID	PANHANDLE	3840	3670
ST JOE RIVER AT CALDER ID	PANHANDLE	1030	4546
JOHNSON CK AT YELLOW PINE	SALMON	211	7135
LEMHI RIVER NR LEMHI, ID	SALMON	907	7431
MF SALMON RIVER AT MF LODGE NR YELLOWPINE, ID	SALMON	1038	7190
SALMON RIVER AT SALMON ID	SALMON	3760	7398
SALMON RIVER AT WHITE BIRD ID	SALMON	13421	6754
SF SALMON R NR KRASSEL RANGER STATION	SALMON	330	6372
BRUNEAU RIVER NR HOT SPRING ID	SOUTHSIDE	2687	5630

GOOSE CREEK AB TRAPPER CREEK NR OAKLEY, ID	SOUTHSIDE	670	6040
OAKLEY RESERVOIR NR OAKLEY ID	SOUTHSIDE	711	6050
REYNOLDS CK AT TOLLGATE	SOUTHSIDE	21	6030
SALMON FALLS CK NR SAN JACINTO	SOUTHSIDE	1450	6020
TRAPPER CREEK NR OAKLEY, ID	SOUTHSIDE	54	6339
BUFFALO FK AB LAVA CK NR MORAN	UPPER SNAKE	323	8951
FALLS RIVER NR ASHTON	UPPER SNAKE	337	7500
GREYS R AB RESERVOIR NR ALPINE	UPPER SNAKE	448	8105
HENRYS FORK NR ASHTON ID	UPPER SNAKE	1096	6710
HENRYS FORK NR REXBURG ID	UPPER SNAKE	2920	6520
PACIFIC CK AT MORAN	UPPER SNAKE	169	8135
PORTNEUF RIVER AT TOPAZ ID	UPPER SNAKE	570	6080
SALT R AB RESERVOIR NR ETNA	UPPER SNAKE	857	7260
SNAKE R AB RESERVOIR NR ALPINE	UPPER SNAKE	3465	8150
SNAKE R NR MORAN	UPPER SNAKE	824	8040
SNAKE RIVER AB JACKSON LAKE AT FLAGG RANCH WY	UPPER SNAKE	486	8199
SNAKE RIVER AT NEELEY ID	UPPER SNAKE	13600	6560
SNAKE RIVER NR HEISE ID	UPPER SNAKE	5752	7770
SNAKE RIVER NR IRWIN ID	UPPER SNAKE	5225	7910
TETON RIVER AB SOUTH LEIGH CREEK NR DRIGGS ID	UPPER SNAKE	335	7303
TETON RIVER NR ST ANTHONY ID	UPPER SNAKE	890	6921
WILLOW CREEK NR RIRIE ID	UPPER SNAKE	627	6390
BOISE RIVER NR BOISE ID	WEST CENTRAL	2686	5910
BOISE RIVER NR TWIN SPRINGS ID	WEST CENTRAL	833	6413
DEADWOOD RIVER BL DEADWOOD RES NR LOWMAN ID	WEST CENTRAL	112	6630
LAKE FORK PAYETTE RIVER AB JUMBO CR NR MCCALL ID	WEST CENTRAL	49	6922
MORES CREEK NR ARROWROCK DAM ID	WEST CENTRAL	399	5071
NF PAYETTE RIVER AT CASCADE ID	WEST CENTRAL	626	5940
NF PAYETTE RIVER NR BANKS ID	WEST CENTRAL	933	5810
PAYETTE RIVER NR HORSESHOE BEND ID	WEST CENTRAL	2220	5850
SF BOISE RIVER AT ANDERSON RANCH DAM ID	WEST CENTRAL	978	6610
SF PAYETTE RIVER AT LOWMAN ID	WEST CENTRAL	449	6825
WEISER RIVER NR WEISER ID	WEST CENTRAL	1460	4141

Appendix C: 10-year mean normalized standard error for each forecast basin, sorted in ascending order of the error. Forecast area codes are as follows: US – Upper Snake; P – Panhandle; C – Clearwater; WC – West Central; SM – Salmon; SS – Southside Snake; LW – Lost/Wood; B – Bear; O – Owyhee.

FORECAST POINT NAME	NORMALIZED STANDARD ERROR	FORECAST AREA
SNAKE RIVER AB JACKSON LAKE AT FLAGG RANCH WY	0.093	US
PEND OREILLE RIVER AT NEWPORT ID	0.127	P
SNAKE RIVER NR HEISE ID	0.142	US
SELWAY RIVER NR LOWELL ID	0.145	C
SNAKE R AB RESERVOIR NR ALPINE	0.147	US
LOCHSA RIVER NR LOWELL ID	0.149	C
KOOTENAI RIVER AT LEONIA ID	0.160	P
SNAKE RIVER NR IRWIN ID	0.164	US
GREYS R AB RESERVOIR NR ALPINE	0.166	US
SF PAYETTE RIVER AT LOWMAN ID	0.172	WC
CLARK FORK AT WHITEHORSE RAPIDS NR CABINET ID	0.172	P
ST JOE RIVER AT CALDER ID	0.174	P
BOUNDARY CREEK NR PORTHILL ID	0.177	P
LAKE FORK PAYETTE RIVER AB JUMBO CR NR MCCALL ID	0.182	WC
HENRYS FORK NR REXBURG ID	0.184	US
BUFFALO FK AB LAVA CK NR MORAN	0.185	US
PAYETTE RIVER NR HORSESHOE BEND ID	0.190	WC
CLEARWATER RIVER AT OROFINO ID	0.191	C
FALLS RIVER NR ASHTON	0.195	US
CLEARWATER RIVER AT SPALDING ID	0.196	C
NF PAYETTE RIVER NR BANKS ID	0.201	WC
SF SALMON R NR KRASSEL RANGER STATION	0.204	SM
BOISE RIVER NR BOISE ID	0.210	WC
DEADWOOD RIVER BL DEADWOOD RES NR LOWMAN ID	0.211	WC
PRIEST RIVER NR PRIEST RIVER ID	0.212	P
BOISE RIVER NR TWIN SPRINGS ID	0.215	WC
HENRYS FORK NR ASHTON ID	0.219	US
SPOKANE RIVER NR POST FALLS ID	0.225	P
SF BOISE RIVER AT ANDERSON RANCH DAM ID	0.227	WC
SPOKANE RIVER AT LONG LAKE	0.230	P
MOYIE RIVER AT EASTPORT ID	0.231	P
JOHNSON CK AT YELLOW PINE	0.232	SM
SALMON RIVER AT WHITE BIRD ID	0.240	SM
DWORSHAK RES NR AHSAHKA ID	0.241	C
NF PAYETTE RIVER AT CASCADE ID	0.252	WC
TETON RIVER NR ST ANTHONY ID	0.254	US

PACIFIC CK AT MORAN	0.255	US
COEUR D ALENE RIVER ABOVE NORTH FORK NEAR ENAVILLE ID	0.255	P
MF SALMON RIVER AT MF LODGE NR YELLOWPINE, ID	0.270	SM
SALMON RIVER AT SALMON ID	0.274	SM
PORTNEUF RIVER AT TOPAZ ID	0.291	US
TRAPPER CREEK NR OAKLEY, ID	0.291	SS
SNAKE R NR MORAN	0.294	US
TETON RIVER AB SOUTH LEIGH CREEK NR DRIGGS ID	0.297	US
BIG LOST RIVER AT HOWELL RANCH NR CHILLY ID	0.304	LW
BIG LOST RIVER BL MACKAY RES NR MACKAY ID	0.306	LW
LITTLE WOOD RIVER NR CAREY ID	0.318	LW
REYNOLDS CK AT TOLLGATE	0.329	SS
SALT R AB RESERVOIR NR ETNA	0.330	US
BIG WOOD RIVER BL MAGIC DAM NR RICHFIELD ID	0.346	LW
LITTLE LOST RIVER NR HOWE ID	0.347	LW
LEMHI RIVER NR LEMHI, ID	0.354	SM
MORES CREEK NR ARROWROCK DAM ID	0.360	WC
BIG WOOD R AB MAGIC RESERVOIR	0.378	LW
SALMON FALLS CK NR SAN JACINTO	0.387	SS
BIG WOOD RIVER AT HAILEY ID	0.388	LW
SNAKE RIVER AT NEELEY ID	0.392	US
LITTLE WOOD RIVER AB HIGH FIVE CREEK NR CAREY, ID	0.393	LW
BRUNEAU RIVER NR HOT SPRING ID	0.429	SS
CAMAS CK AT CAMAS	0.464	LW
OWYHEE R NR ROME	0.501	O
OWYHEE R BL OWYHEE DAM 2	0.504	O
WEISER RIVER NR WEISER ID	0.513	WC
CAMAS CREEK NR BLAINE ID	0.518	LW
BEAR RIVER BELOW STEWART DAM NEAR MONTPELIER, ID	0.528	B
WILLOW CREEK NR RIRIE ID	0.535	US
GOOSE CREEK AB TRAPPER CREEK NR OAKLEY, ID	0.546	SS
OAKLEY RESERVOIR NR OAKLEY ID	0.575	SS
OWYHEE R NR GOLD CK 2	0.710	O

Appendix D: Mean percent error of each forecast basin.

FORECAST POINT NAME	10-YEAR MEAN ABSOLUTE PERCENT ERROR
Deadwood Reservoir Inflow	8.73
SF Payette R at Lowman	9.66
Lake Fork Payette R nr McCall	11.89
Snake R Nr Alpine	12.06
Greys R Nr Alpine	12.15
Selway R nr Lowell	12.36
Falls R nr Ashton	12.45
Snake R nr Heise	12.95
Snake R nr Irwin	13.42
Buffalo Fork ab Lava nr WY	13.44
CLARK FK at Whitehorse Rpds	13.53
Salmon R at White Bird	13.56
Lochsa R nr Lowell	13.56
MF Salmon R at MF Lodge	13.70
Payette R nr Horseshoe Bend	13.77
Boise R nr Twin Springs	14.20
PEND OREILLE Lake Inflow	14.22
SF Salmon R nr Krassel RS	14.73
Snake R Nr Moran	15.29
Henrys Fork nr Ashton	15.51
Clearwater R at Spalding	15.62
Snake River At Flagg Ranch	15.68
Boise R nr Boise	15.84
Henrys Fork nr Rexburg	15.89
NF Payette R at Cascade	15.97
Big Wood River at Hailey	16.88
Johnson Ck at Yellow Pine	17.09
KOOTENAI at Leonia	17.36
ST. JOE at Calder	17.70
Salmon R at Salmon	17.84
Teton R nr St. Anthony	17.93
NF Payette R nr Banks	18.06
Clearwater R at Orofino	18.11
SPOKANE at Long Lake	18.71
Little Lost R nr Howe	18.85
PRIEST near Priest River	19.08
Trapper Ck nr Oakley	19.41
Dworshak Reservoir Inflow	19.50

Teton R nr Driggs	19.55
Big Lost R at Howell Ranch	20.56
BOUNDARY CREEK	21.45
Lemhi R nr Lemhi	21.48
MOYIE RIVER at Eastport	21.54
Pacific Ck At Moran	22.29
SPOKANE near Post Falls	22.85
NF COEUR D'ALENE RIVER at Enaville	23.18
SF Boise R at Anderson Ranch Dam	23.36
Reynolds Ck at Tollgate	23.71
Salt R Nr Etna	23.73
Weiser R nr Weiser	26.47
SNAKE RIVER near Murphy	29.22
Bruneau R nr Hot Springs	30.77
Little Wood R ab High Five Creek	31.19
Mores Ck nr Arrowrock Dam	31.21
Big Lost R bl Mackay Res	31.84
Portneuf R at Topaz	32.66
Salmon Falls Ck nr San Jacinto	33.34
Little Wood R nr Carey	34.76
Oakley Reservoir Inflow	36.91
Big Wood R ab Magic Reservoir	38.71
Bear River at Stewart Dam	42.61
Big Wood R bl Magic Dam	44.78
Owyhee R blw Owyhee Dam	51.12
Snake River at Neeley	53.08
Willow Ck nr Ririe	73.35
Owyhee R nr Rome	79.29
Camas Ck nr Blaine	86.52
Owyhee R nr Gold Creek	180.88
Camas Ck at Camas	1949.97

Appendix E: Nash-Sutcliffe skill score of points with complete forecast-observation pairs for the 1998-2017 period. Forecast area codes are as follows: US – Upper Snake; P – Panhandle; C – Clearwater; WC – West Central; SM – Salmon; SS – Southside Snake; LW – Lost/Wood; B – Bear; O – Owyhee.

FORECAST POINT NAME	FORECAST AREA	NS SKILL SCORE
PRIEST RIVER NR PRIEST RIVER ID	P	0.420
LITTLE LOST RIVER NR HOWE ID	LW	0.462
OAKLEY RESERVOIR NR OAKLEY ID	SS	0.490
KOOTENAI RIVER AT LEONIA ID	P	0.509
TETON RIVER AB SOUTH LEIGH CREEK NR DRIGGS ID	US	0.524
OWYHEE R BL OWYHEE DAM 2	O	0.536
PACIFIC CK AT MORAN	US	0.542
CLEARWATER RIVER AT OROFINO ID	C	0.545
OWYHEE R NR ROME	O	0.563
WEISER RIVER NR WEISER ID	WPB	0.570
BRUNEAU RIVER NR HOT SPRING ID	SS	0.585
HENRYS FORK NR ASHTON ID	US	0.616
SALT R AB RESERVOIR NR ETNA	US	0.616
PEND OREILLE RIVER AT NEWPORT ID	P	0.638
BIG LOST RIVER AT HOWELL RANCH NR CHILLY ID	LW	0.647
COEUR D ALENE RIVER ABOVE NORTH FORK NEAR ENAVILLE ID	P	0.649
BIG LOST RIVER BL MACKAY RES NR MACKAY ID	LW	0.649
SALMON FALLS CK NR SAN JACINTO	SS	0.655
CLEARWATER RIVER AT SPALDING ID	C	0.659
SALMON RIVER AT WHITE BIRD ID	SM	0.677
SPOKANE RIVER NR POST FALLS ID	P	0.686
DWORSHAK RES NR AHSAHKA ID	C	0.687
BIG WOOD R AB MAGIC RESERVOIR	LW	0.694
PORTNEUF RIVER AT TOPAZ ID	US	0.705
NF PAYETTE RIVER NR BANKS ID	WPB	0.707
SALMON RIVER AT SALMON ID	SM	0.713
SPOKANE RIVER AT LONG LAKE	P	0.716
MORES CREEK NR ARROWROCK DAM ID	WPB	0.717
NF PAYETTE RIVER AT CASCADE ID	WPB	0.720
ST JOE RIVER AT CALDER ID	P	0.729
SNAKE R NR MORAN	US	0.732
GREYS R AB RESERVOIR NR ALPINE	US	0.733
OWYHEE R NR GOLD CK 2	O	0.735
SNAKE RIVER NR HEISE ID	US	0.772

SNAKE R AB RESERVOIR NR ALPINE	US	0.773
SNAKE RIVER NR IRWIN ID	US	0.780
LITTLE WOOD RIVER NR CAREY ID	LW	0.785
CAMAS CREEK NR BLAINE ID	LW	0.789
BOISE RIVER NR TWIN SPRINGS ID	WPB	0.795
BIG WOOD RIVER AT HAILEY ID	LW	0.799
PAYETTE RIVER NR HORSESHOE BEND ID	WPB	0.811
BIG WOOD RIVER BL MAGIC DAM NR RICHFIELD ID	LW	0.814
BOISE RIVER NR BOISE ID	WPB	0.846
SF BOISE RIVER AT ANDERSON RANCH DAM ID	WPB	0.856
SF PAYETTE RIVER AT LOWMAN ID	WPB	0.859
DEADWOOD RIVER BL DEADWOOD RES NR LOWMAN ID	WPB	0.859