Climate data: the ins and outs and where to find what

By Zack Guido

This article is the second in a two-part series. Part One, featured last month, discussed the National Weather Service's Cooperative Observer Program and the related Historical Climate Network. This article describes data from Remote Automated Weather Stations (RAWS) and the Arizona Meteorological Network (AZMET), and data generated by the Parameter-elevation Regressions on Independent Slopes Model (PRISM) statistical technique.

Not enough data is bad. Too much data is overwhelming. But not knowing what data exists and where to find it is worse.

Hundreds of weather stations in the Southwest dot the landscape, piping measurements to many different users. The National Weather Service (NWS), for example, intertwines the information in models that help forecast tomorrow's weather, while a coordinated group of federal wildfire agencies eyes data from different stations to monitor fire risk. The Arizona Cooperative Extension uses data from yet another network to derive "degree days" from temperature measurements, which allow farmers to estimate an outbreak of the infamous pink bollworm.

While climate and weather data support many actions, it is often difficult for users outside each data network's administration to track down and understand the data. Three networks—Remote Automated Weather Stations (RAWS), the Arizona Meteorological Network (AZMET), and data generated by a sophisticated algorithm called the Parameter-elevation Regressions on Independent Slopes Model (PRISM)—offer detailed data that may have gone unnoticed to some.

While RAWS and AZMET capture extreme conditions and weather representative of agricultural areas, PRISM meshes observations from several networks into a mathematical model that estimates climate for small grid-boxes that span the entire U.S. All three, along with the Cooperative Observer Programs (Coop) and the Historical Climate Network (HCN), which were discussed in the March *Southwest Climate Outlook*, can help researchers understand climate change, businesses relate product demand to climate, and resource managers dole water to irrigation districts, among other uses.

Remote Automated Weather Stations

The RAWS network was established principally to help fire managers predict fire behavior and monitor the conditions of fuels, such as standing and fallen trees. As a result, the stations have been systematically located in remote areas that capture extreme conditions, including windy areas and sites that receive a hefty dose of sunlight—areas that are the most susceptible to fire. RAWS are generally not sited on northern facing slopes, which receive less sunlight than southern aspects. While RAWS are predominantly used for fire-risk assessments, the data also assist in air quality monitoring and research.

Nearly 2,200 RAWS are strategically located throughout the United States. There are 130 stations in Arizona and New Mexico, and the oldest stations have been active since the mid-1980s. Most RAWS are operated by the wildland fire agencies, such as the National Forest Service, the Bureau of Land Management, and the U.S. Fish and Wildlife Service.

RAWS record weather conditions every minute to every hour, depending on the variable being measured, and transmit the information via satellite to the National Interagency Fire Center and the Western Regional Climate Center (WRCC). This allows users to obtain real-time information. Most RAWS record temperature, precipitation, wind speed and direction, barometric pressure, and relative humidity; some stations also record the moisture and temperature of fire fuels. The data are free and most easily accessed at the WRCC.

The RAWS data have some limitations. First, the data available to the public are not quality controlled. The raw values recorded at the stations are the same as those archived at the WRCC. Second, not all stations in the western U.S. continuously collect data—some stations sleep in the winter when fire risk is low, particularly those at higher elevations. In addition, some stations are portable and are moved during the year and between years. Because micro climate can impact weather conditions, data from portable stations are not useful for long-term analysis without careful inspection. Furthermore, some RAWS data are not well annotated, making it difficult to decipher which stations moved and the site characteristics of the new and old locations.

Like all networks, RAWS have a specific purpose, which influences how data is recorded. To monitor fire risk, for example, wind speeds are measured at a height of 20 feet and are averaged over 10 minutes. Weather stations at airports, in contrast, measure wind speeds at 33 feet and the values are averaged over two-minutes. Knowing these and other RAWS data issues can help make this detailed dataset useful.

Arizona Meteorological Network

AZMET—a service of the Cooperative Extension at The University of Arizona— provides meteorological data and weather-based information to agricultural and horticultural interests operating in southern and central Arizona. Each hour, AZMET stations record numerous climate and weather variables that have been useful for irrigation districts, golf courses, cotton and citrus growers, fertilizer and pesticide companies, researchers, and others.

continued on page 4

Climate data, continued

The earliest AZMET stations began operating in 1987, and 28 stations are currently active. The stations are located in both rural and urban agricultural areas and are often positioned in open spaces over grass and away from buildings. As a result, the data is not as affected by urban heat island effects, which can amplify temperature and alter other climate variables. One asset of AZMET is that it measures many climate and weather variables, including air temperature, soil temperature at two depths, precipitation, wind speed and direction, solar radiation, and humidity. From those measurements, AZMET calculates heat units and chill hours, which help characterize the life stages of plants, and evapotranspiration. Because the data are recorded hourly, the dataset is rich and detailed. Furthermore, AZMET stations measure climate and weather variables not collected by other networks, including evapotranspiration. A census of data collection organizations indicates that AZMET is the only network to monitor evapotranspiration continuously in Arizona.

Another positive feature of the AZMET network is that stations are well maintained, which helps create consistent data. A technician visits each site at least every three months and erects a temporary station with laboratory-calibrated sensors. A comparison of the results between the official and temporary stations helps AZMET evaluate the reliability of the data and ensure accurate measurements. In addition, AZMET changes the wind speed and solar sensors each year and changes the temperature and humidity sensor every two years to prevent sensor failure or measurement drift. Many other networks change equipment only after problems occur, often making it difficult to locate in the data when values became inaccurate.

AZMET data also are quality controlled, although not as rigorously as the HCN network. Most quality control is performed by computer statistical analysis, in which measurements are cross-checked with nearby stations to make sure that one station is not recording artificial conditions. Additional computer programs comb the data for negative values or uncharacteristically extreme values. The presence of these anomalies tells technicians to review the data manually. Each morning at about one a.m., the data is transferred onto a Web server where it is free and available to the public.

Like RAWS, however, AZMET data have limitations. First, the period of record is relatively short: a maximum of 22 years, and only 12 stations span this period. AZMET data is therefore not as useful for deriving long-term climate trends as other networks such as the

Network	Data Source	Climate Variables	Recording Intervals	Record Length	Primary Application	Quality Control
Coop	12,000 active Coop stations; ~170 in AZ and ~180 in NM	1. Maximum temp. 2. Minimum temp. 3. Daily total precip. 4. Daily total snow 5. Others	Once a day	1880 – present; varies by station	Support public services with near real-time data	Some quality control after data acquisition
HCN	1,221 stations selected from Coop network	1. Maximum temp. 2. Minimum temp. 3. Daily total precip. 4. Daily total snow	Once a day	Most stations have data for 80 years or more	Detect and monitor changes in regional climate	Extensive quality control after data acquisition
RAWS	2,200 remote automated stations; 130 in AZ and NM	1. Temperature 2. Precipitation 3. Wind speed 4. Relative humidity 5. Others	Minute to hourly	Many stations became active in the mid-1980s	Monitor fire-risk	No quality control
AZMET	28 automated stations in rural and urban areas in AZ	1. Temperature 2. Precipitation 3. Evapotranspiration 4. Others	Hourly	1986 – present; varies by station	Support agriculture and horticulture in southern and central Arizona	Some quality control after data acquisition; routine station maintenance
PRISM	Coop, SNOTEL, local stations, and statistically generated data	1. Maximum temp. 2. Minimum temp. 3. Average temp. 4. Precipitation	Monthly	1895 – present	Produce detailed, high-quality spatial climate datasets	Depends on data source

Coop: Cooperative Observer Program; HCN:Historical Climate Network; AZMET: Arizona Meteorological Network; PRISM: Parameter-elevation Regressions on Independent Slopes Model; RAWS: Remote Automated Weather Stations; SNOTEL: snow telemetry

Table 1. Characteristics of common sources of climate and weather data.

continued on page 5

Climate data, continued

HCN. Also, the station density is sparse, except in the Phoenix area, and the data is predominantly limited to southern and central Arizona. Finally, the data is representative of agricultural locations, providing information that is suitable for aiding agricultural decisions but not as appropriate for understanding the climate of ranchlands or assessing the urban heat island as other data sources.

Parameter-elevation Regressions on Independent Slopes Model

All monitoring stations, including Coop, HCN, RAWS, and AZMET, measure weather and climate conditions at a location. But climate can vary dramatically across short distances and over small elevation changes. Even Coop, which has 170 active stations in Arizona, cannot adequately cover the entire state. What about the weather in areas between the stations?

To fill in data gaps between stations, Oregon State University developed PRISM, an observation-based statistical algorithm that uses measurements made at monitoring stations from several data networks. PRISM generates climate data for a 2.5 by 2.5 mile (or four-kilometer) grid that covers the continental United States.

The PRISM model computes climate values in a sophisticated way. Essentially, the model overlays a grid on a three-dimensional relief map of the U.S. and marks the grid-boxes containing monitoring stations. It then assigns the observed values for precipitation, temperature, and other variables to each box with an established station. After this, boxes remain that do not have stations. PRISM populates these grids with climate values, for each box, derived from the unique relationship between climate and elevation, coastal proximity, topography, distance to known observations, and aspect. The PRISM algorithm is specifically designed to generate realistic climate data for areas prone to complex weather, such as mountainous regions, places in rain shadows, and regions near water.

PRISM has been used to create a continuous monthly climate data for 1895 to the present. The length of record and the fine spatial resolution make PRISM data unique, meeting the needs of resource managers, land-use planners, researchers, and many other stakeholders.

PRISM data, however, have some drawbacks. Monitoring stations at higher elevations are few and far between, and therefore some people believe that PRISM data for higher elevations is less reliable. Also, any statistical procedure introduces additional sources of error. In addition, only monthly data are available.

Until recently, PRISM data were not easily analyzed without specialized software. However, the need for more accessible, fine-scale climate datasets spawned the Western Climate Mapping Initiative (WestMap), a collaborative effort between The University of Arizona, The Desert Research Institute, and Oregon State University. CLIMAS also played a role, helping identify demand for Web-based PRISM data.

WestMap has developed a Web-based climate analysis and mapping tool that enables users to download and graphically display PRISM data for the western U.S. The tool allows users to query data for different time periods and regions, download the data in a common format, and create maps and charts. For example, users can obtain monthly data for any period between 1895 and the present for a user-defined area, such as a single location, an entire state, or a watershed. Users may also create custom maps to suit their needs.

Conclusion

Weather and climate data come from many sources and possess unique qualities. While stations in the RAWS network are in remote, sun-baked areas, PRISM sites are virtual. While many HCN stations span more than 80 years, AZMET stations have made measurements since 1987. And while Coop stations and RAWS have minimal quality control, HCN and AZMET are processed with a finer-tooth comb.

Regardless of which networks are used, however, knowing the ins and outs of each will help match the proper dataset to the question at hand and can help enable businesses, farmers, researchers, natural resource managers, and others to more effectively make decisions.

For questions or comments, please contact Zack Guido, CLIMAS Associate Staff Scientist, at zguido@email.arizona.edu or (520) 882-0879.

Related Links

Arizona Meteorological Network

1. Access all AZMET data and encounter more information: http://ag.arizona.edu/AZMET/

Remote Automated Weather Stations

1. Access data through a map interface, hosted by Western Regional Climate Center: http://www.raws.dri.edu/index.html

2. RAWS home page provides overview of RAWS program: http://www.fs.fed.us/raws/

PRISM

1. User-friendly graphical interface for accessing PRISM data for Western U.S., developed by WestMap: http://www.cefa.dri.edu/Westmap/Westmap_home.php

2. Access datasets for entire U.S. via Oregon State University: http://www.prism.oregonstate.edu/

Evaluating forecasts with the RPSS

By Zack Guido

In response to user feedback, the Southwest Climate Outlook has changed its temperature and precipitation forecast verification highlights to incorporate a more accurate evaluation method, the Rank Probability Skill Score (RPSS). To the mathematically wary, this name likely causes anxiety. Indeed, the RPSS is an equation and is complicated. But it helps answer a critical question: have the forecasts been accurate? Knowing this helps users incorporate the forecasts into decisions, such as when to purchase hay to avoid high costs or how much water to dole to irrigation districts.

Scientists often evaluate a forecast by calculating its skill, which is the accuracy of a forecast in relation to another, reference forecast. A "skillful" forecast shows improvement over the reference forecast. For example, a poker player may say he or she can beat the house more often than losing. If the game played has 50:50 odds, the poker player must win more than 50 percent of the games to show skill over the odds (the reference forecasts).

The National Oceanic and Atmospheric Administration's Climate Prediction Center (NOAA-CPC) began forecasting successive three-month periods in 1994, and these forecasts spanned two weeks to 13 months into the future. But the usefulness of these forecasts depends on their accuracy. If the forecasts have been historically worse than simply using a coin to predict the weather, than what value do they have?

To help address this question for readers, the Southwest Climate Outlook verification pages will present the average RPSS calculated for all the temperature and precipitation forecasts issued since 1994 for four different lead times. The RPSS is calculated by the Forecast Evaluation Tool, which was developed by The University of Arizona in partnership with NOAA, NASA, the National Science Foundation, and the University of California-Irvine.

In essence, the RPSS communicates how much more or less accurate the CPC forecasts have been than the reference forecast. The reference forecast for the CPC forecasts is equal probabilities that temperatures or precipitation will be one of three categories—"above," "below," or "neutral"—or a 33 percent chance for each category. These forecasts give probabilities, for example, that temperature will be similar to the 10 warmest, coolest, or normal temperatures observed during the period 1971–2000. This equal probability is often referred to as a climatology forecast.

The actual formula of the RPSS is complicated and is beyond the scope of this article. The two important characteristics of the RPSS, however, are easily articulated. First, the higher the RPSS value, the better the forecast; the RPSS value is the percent improvement the forecast exhibits over the reference forecast. Positive values also give an indication that the forecasts and the actual weather conditions are similar—the higher the RPSS, the more similar the forecast and the actual conditions. Negative values, on the other hand, mean that the forecast is less accurate than the climatology forecast.

Second, the value of the RPSS incorporates the degree of correctness or incorrectness. This "ranked" scoring system values correct forecasts and incorrect forecasts differently—some inaccurate forecasts are worse than others. For example, if a forecast indicated a 90 percent chance for "above" temperatures but temperatures were actually "below," the RPSS would be lower than if the forecast stated a 40 percent chance for "above" temperatures.

The usefulness of forecast verifications such as the RPSS becomes apparent in the example of an early forecaster. In 1884, Sergeant John Finley began forecasting tornado occurrences east of the Rocky Mountains. Shortly thereafter, he reported a 95.6–98.6 percent forecast accuracy. Other scientists, however, pointed out that the accuracy could have been 98.2 percent had he simply always forecasted no tornados. Although Finley's forecasts seemed accurate, they were not the best forecasts. Had an RPSS been calculated, it would have been negative.

While forecasts will continue to be made—each additional year helps make the RPSS more robust—knowing the accuracy of past forecasts will help evaluate the usefulness of the current forecast.

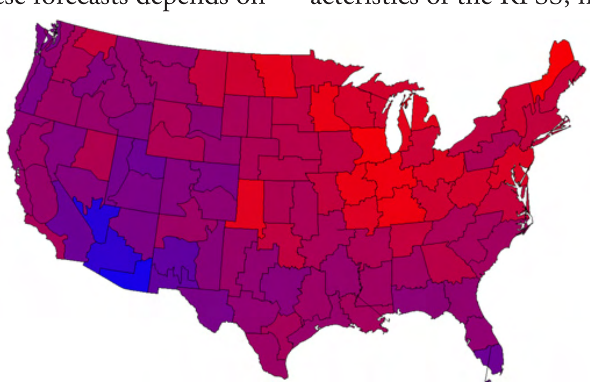

Figure 1. The new verification highlights incorporate a more sophisticated measure of forecast performance than the highlights featured in the past. The new color maps like this one that help readers visualize the historical accuracy of the forecasts.

For questions or comments, please contact Zack Guido, CLIMAS Associate Staff Scientist, at zguido@email.arizona.edu or (520) 882-0879.