WEB-BASED FORECASTING OF POTENTIAL EVAPOTRANSPIRATION FOR IMPROVED WATER RESOURCE MANAGEMENT IN CALIFORNIA

David Yates, Scientist, National Center for Atmospheric Research, 3450 Mitchell Ln., Boulder, CO 80301, (303)-497-8394, Yates@ucar.edu; Michael Tansey, Special Projects Coordinator, Division of Planning, Mid Pacific Regional Office, 2800 Cottage Way, Sacramento, CA 95825, (916) 978-5197, mtansey@usbr.gov; Conrad Roesch, Software Engineer, Digital Dynamics, Boulder CO, (720)-347-7373, conrad006@live.com

Abstract: What to plant, how much and when to deliver irrigation water among the many decisions agricultural, water, and irrigation managers must make throughout California's Central Valley. If these planners and managers could have forecasts of water demand in the near (8-day) and long (90-day) term, the strong possibility exists for improved resource management decisions. To this end, a web-based tool that generates both 8-day forecasts and 90-outlooks of potential evapotranspiration (ETo) for use in physically based water management models has been developed. The forecasts and outlooks are made for California Irrigation Management Information System (CIMIS) sites throughout California, with the 8-day forecasts based on NCEP High Resolution Global Forecast System (1 degree GFS) model output with site specific bias-correction. The 90-day outlook is based on the International Research Institute for Climate Prediction's (IRI) monthly forecasts of global temperature and precipitation out 3-months (e.g. 90-days) in advance. A web-site, http://californiaPET.org, allows the user to select specific stations of interest, and save the forecasts in either text or DSS format for direct use in water management models such as the United States Bureau of Reclamation's (USBR) Land Atmosphere Water Simulator (LAWS) model. The tool easily accommodates automation through custom Universal Resource Locator (URL) tags. In addition to ETo, the site also generates daily precipitation estimates in units of either mm/day or inches/day.

INTRODUCTION

There is increasing pressure to better manage California's water resources into the future. For example, on February 28, 2008 Governor Schwarzenegger wrote to the California State Senate, asking for "a plan to achieve a 20 percent reduction in per capita water use statewide by 2020", affectionately known as "20 by 2020". Since agriculture is the biggest user of water in California, innovative strategies to reduce water consumption in this sector would go a long ways in achieving this goal. Most likely, these reductions will not come about through single decisions that lead to large reductions, rather they will come from many small-scale decisions informed through better science and improved management. This paper presents a methodology and the implementation of a methodology within a web-based framework for forecasting potential evapotranspiration (ETo) throughout California for use in water management models. ETo is the basis for estimating crop water use and we contend that an improved estimation and forecasting of ETo could help water and agricultural managers alike (Landeras and Oritz-Barredo 2009).

What to plant, how much and when to deliver irrigation water are among the many decisions agricultural, water, and irrigation managers must make throughout California's Central Valley. If these planners and managers could have forecasts of crop water requirements in the near (8-day) and long (90-day) term, the strong possibility exists for improved resource management decisions (Acra 2004). A web-based tool that generates both 8-day forecasts and 90-outlooks of potential evapotranspiration for use in physically based water management models is being developed for use by California agriculturalists. The forecasts and outlooks are made for California Irrigation Management Information System (CIMIS) sites throughout California, with the 8-day forecasts based on NCEP High Resolution Global Forecast System (1 degree GFS) model output with site specific bias-correction. The 90-day outlook is based on the International Research Institute for Climate Prediction's (IRI) monthly forecasts of regional temperature and precipitation, out 3-months (e.g. 90-days) in advance. The 90-day outlook includes the generation of ensemble members consistent with the IRI outlook, which should allow for a risk-based approach to water management decisions.

A web-site, http://californiaPET.org, allows users to select either the 8-day forecast or 90-day outlook methodology, select specific stations of interest, and save the forecasts in either text or DSS format for direct use in water management models such as the United States Bureau of Reclamation's (USBR) Land Atmosphere Water Simulator (LAWS) model. The tool easily accommodates automation through custom Universal Resource Locator (URL) tags, so new forecasts can be entered into agricultural and water management models on a daily basis. In addition to ETo, the site also generates daily precipitation estimates in units of either mm/day or inches/day.

APPROACH

8-DAY FORECAST

An eight-day ETo forecast is generated for select CIMIS sites from the CaliforniaPET.org website, for a 3-hour or a daily accumulation period using either the Penman-Montieth or the CIMIS method in mm or inches (Penman 1948 and Sharma 1985). The meteorological forcing data comes straight from the output of NOAA's Global Forecast System (GFS) spectral model. The GFS model output is posted to a 1x1 degree equally spaced longitude/latitude grid with 3-h forecast interval out to 180-hours, cycled 4X/day, and referred as the "GFS high resolution" data set. Custom data requests are made available through a National Operational Model Archive and Distribution System (NOMADS), whose website at the time of writing can be found at http://nomads.ncep.noaa.gov/. Available meteorological variables include the full complement of those needed to compute ETo including surface temperature, humidity, radiation and wind. However, some required ETo variables, most notably net radiation (Rnet) and windspeed require post-processing of the GFS output. Server side, CaliforniaPET.org uses the down and up welling short and long wave radiation to compute Rnet in units of Watts per square meters (W/m^2) and the U and V surface wind components to compute a surface windspeed in units of meters per second (m/s).

Since the GFS surface temperature, T_m are output on a coarse, 1x1 degree grid, site specific, sub-grid temperature variability has been estimated using a simple lapse rate adjustment of 6.5°C per kilometer. The grid average elevation, E_g has been computed for each of the 1x1 degree, cells over California. The station temperature, T_s is then estimated as the difference between this grid average elevation and the station elevation, E_s, e.g. $T_s = T_m*(E_s - E_g)*6.5$.

Figure 1 shows a screen-shot of the 8-day ETo, CaliforniaPET.org website. The interactive map on the left side of the image is where the user selects specific CIMIS sites for returning the ETo forecasts. The stations are color coded according to ETo zones (Jones 1999). On the server side, a forecast request triggers the update of the GFS data and computation of the variables needed to compute ETo, which are returned either in html, text, or in Hydrologic Engineering Center Data Storage System (DSS) format. Since the site has only been operational in the late fall of 2009, forecast verification has been very limited.

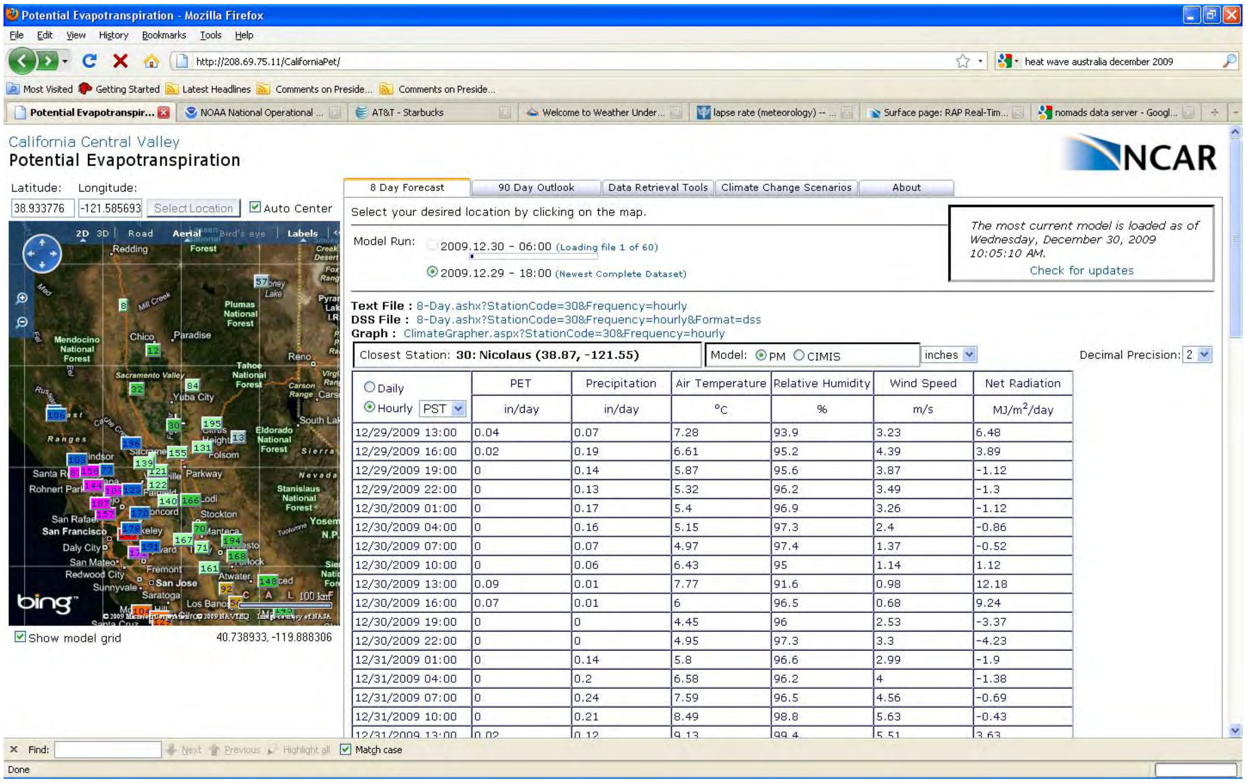

Figure 1. Screenshot of the CaliforniaPET.org's 8-day forecast web page.

Table 1. Required and optional parameters for a URL tag that facilitates automated download of station data.

<table>
<tr><td rowspan="8">8-Day forecast</td><td>Required Parameters:</td><td>StationCode
Numeric Codes, see map</td></tr>
<tr><td rowspan="6">Optional Parameters:</td><td>Format
Accepted Format values: txt, dss (text is default)</td></tr>
<tr><td>Model
Accepted Model format values: PM, CIMIS (PM is default)</td></tr>
<tr><td>Units
Accepted Units format values: mm, in (inches is default)</td></tr>
<tr><td>Frequency
Accepted Frequency format values: daily, hourly (daily is default)</td></tr>
<tr><td>TimeZone
Accepted Time Zone format values: utc, pst (UTC is default)</td></tr>
<tr><td>DecimalPrecision:
Accepted Decimal Precision Values: non-negative int (2 is default)</td></tr>
<tr><td colspan="2"><u>Example URL</u>: Davis Station, Text format, hourly, utc , 2 decimal points
http://208.69.75.11/CaliforniaPet/8-Day.ashx?StationCode=6&Frequency=hourly&Format=text&TimeZone=utc&DecimalPrecision=2</td></tr>
</table>

In addition, a user can make a forecast request for a specific station using a url tag, which facilitates automated downloads. In this way, multiple stations can be selected for download and use in hydrologic and agricultural models. Table 1 shows the required and optional values in a URL tag and a complete formatted tag at the bottom of the table.

90-DAY OUTLOOK

In contrast to the 8-day forecast, which makes use of the explicit output from a numerical weather prediction model, the 90-day seasonal outlook uses a non-parametric, statistical technique to generate total, daily ETo based on a historical climate archive and a seasonal outlook. The historical archive is from Maurer et al. 2002, and includes a daily time series of the key meteorological variables for the period 1950 through 2002. We will consider using the CIMIS archive in the near future.

These seasonal outlooks can come from organizations such as the National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center or from Columbia University's International Research Institute (IRI). Beyond about 10-days, the climate system is basically chaotic and unpredictable from a physical or Global Climate Modeling perspective, and so climatologists must rely on the development of statistical models (some quite clever) that produce these seasonal forecasts. Seasonal outlooks are typically generated through fairly sophisticated combinations of global climate model output and statistical techniques such as canonical correlation analysis (Landman and Mason 1999; Mason and Weigel 2009; Thomson et al. 2006).

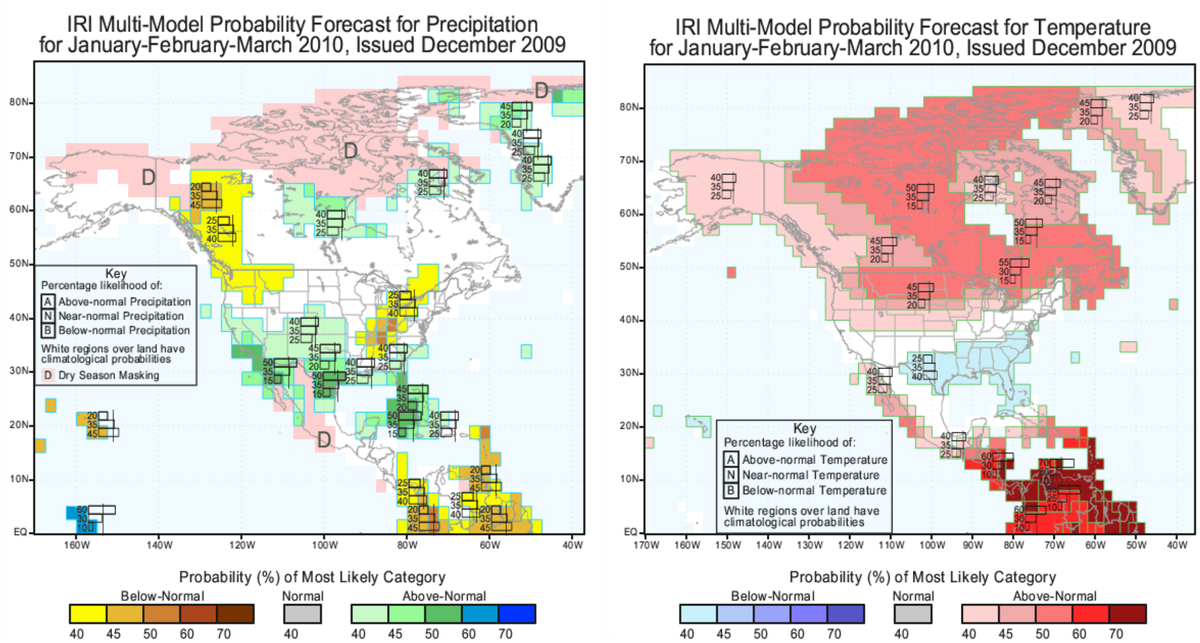

Figure 2. IRI's seasonal outlook of temperature (left) and precipitation (right). Use by permission. For details, see http://iri.columbia.edu/climate/forecast/net_asmt/

The seasonal outlooks are typically updated in the middle of each month. For example, the IRI 90-day outlook for January, February, and March of 2010 was issued on 17 December, 2009. The outlooks suggested above normal precipitation and near normal temperature for California. A non-rigorous validation of the forecast, now that is March 1 of 2010 suggests strong agreement, as most of California experienced above normal precipitation in January and February of 2010.

An ensemble of 90-day, daily ETo outlooks is generated using a non-parametric technique known as K-Nearest Neighbor or *K-NN (*Yates et al. 2003; Clark et al. 2004), which are guided by the IRI outlooks. The K-NN resamples from an observational archive of climate data in such a way that new, synthetic time series possess the biases associated with the seasonal outlook probabilities such as those shown in Figure 2. The *K-NN* scheme resamples from a historic dataset, generating alternative weather sequences of daily data. The *K-NN* simply creates a new, synthetic weather sequence that differs from the one-and-only observed historic weather sequence, with the goal of preserving the spatial correlation and most statistical properties of the observed weather data. If there is interest in generating climate data for multiple stations within a similar climate zone, then a regional mean can be computed from the stations of interest and that regional mean time series used in the resampling procedure. The output from the K-NN algorithm is simply a sequence of historical dates, and once a new sequences of dates is generated, the climate data are simply associated with that date. The steps in K-NN are summarized below.

In Step 1, the algorithm selects a random starting calendar day, say January 1st, from all *Y* available January 1sts, with the steps resulting in the selection of the next day's weather on January 2nd. Step 2 defines a window of days, *w* to the left and right of the starting day selected in Step 1. With a window with of *w*=7 and a dataset with *Y*=24 years, the total population of days would be *2w x Y* or *N*=336 days. A sub-set of candidate days from the population, similar to the day chosen in Step 1, is defined based on a Mahalanobis distance measure (see Yates et al. 2003 for details). The size of this *K* subset of days is given as, $K=\sqrt{\left[(2w+1)\times Y\right]-1}$ or *K*=19 (Rajagopalan and Lall, 1999) in this example. One of the *K* similar days is then randomly selected (Step 3), and represents a day similar to the January 1st selected in Step 1. The subsequent day to this selected day is used as the successor, leading to a new January 2nd (Step 4). The window of days is then shifted one day to the right, and Steps 2 through 4 are repeated. Successive iterations of these steps then generates a new, unique daily time series that similar statistical properties as the historical data such as autocorrelations, cross-covariance, and mean values as the original data (Yates et al., 2003). If all 336 days are given equally likelihood of being selected, then the statistical attributes of the resulting time series should be similar to "climatology".

If a warmer sequence were desired (e.g. a "biased resample"), then the population of candidate days could simply be limited to a subset of warmer days. Biased resampling requires a conditioning criterion to select this $n \in N$ subset of days, which will be used in the same 4-step process described above. Our sub-setting approach uses the IRI probabilistic outlook in the selection of the *n* subset, which is used in the biased resampling. The pool of “similar days” from which the next day is selected can be biased according to precipitation and temperature outlooks based on a binning of the data into their respective terciles. Precipitation and temperature are the two climate variables available from these seasonal outlooks. Precipitation is not directly used in the ETo estimate, but can be used to improve the generation of synthetic weather sequences, which would mean a better ETo forecast.

Our biased resampling uses a binning approach, where candidate days of both precipitation and temperature are binned into terciles. For example, Figure 3 shows the distribution of historic temperature of an arbitrary day for all years, where values below 4^{o}C are in the lowest tercile, those 4^{o}C and 18^{o}C are in the 2nd tercile, and values greater than 18^{o}C are in the 3rd tercile. If we wanted to generate a new sequence of days that retained the climatological attributes, then we would resample equally from each bin. In the example in Figure 2, 100 samples each draw equally from the three terciles (left) would result in a climate sequence that should be similar to climatology, while 100 samples with 25 drawn from tercile 1, 35 from tercile 2 and 40 from tercile 3 would result in a sequence that should be warmer than climatology.

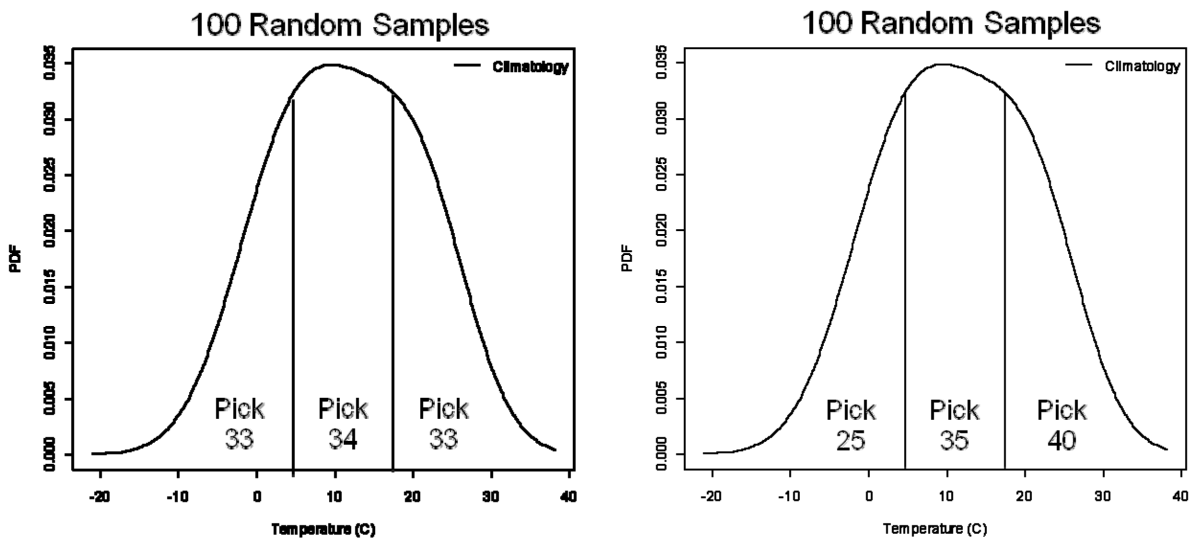

Figure 3. All climate data for an arbitrary day (e.g. all January 1sts from the historical archive) are put into one of tercile bins. Resampling equally from all 3 bins (33%) will return a climate sequence similar to climatology (left), while biased resampling that selects more from the warmer tercile will yield a climate sequence warmer than climatology (right).

The example above assumes that precipitation and temperature are treated independently, but the approach can be extended to reflect their joint-distribution. Figure 4 shows a 3x3 matrix that is used to categorize and distribute the historical data for resampling. For a given window of days and all years, days are assessed according to their relative attribute and binned accordingly, e.g. warm and dry days are placed in the top-left bin, while cool wet days reside in the bottom-right bin. Of course, if the K-NN resampling algorithm only resampled from the top-left bin, then the climate sequence would be warm and dry.

Warm	$T_r > T_{2/3}$ & $P_r < P_{1/3}$	$T_r > T_{2/3}$ & $P_{2/3} < P_r < P_{1/3}$	$T_r > T_{2/3}$ & $P_r < P_{2/3}$
	$T_{1/3} < T_r < T_{2/3}$ & $P_r < P_{1/3}$	$T_{1/3} < T_r < T_{2/3}$ $P_{2/3} < P_r < P_{1/3}$	$T_{1/3} < T_r < T_{2/3}$ & $P_r > P_{2/3}$
Cool	$T_r < T_{1/3}$ & $P_r < P_{1/3}$	$T_r < T_{1/3}$ & $P_{2/3} < P_r < P_{1/3}$	$T_r < T_{1/3}$ & $P_r > P_{2/3}$
	Dry		Wet

Figure 4. A 3x3 matrix used to illustrate the joint-distribution between precipitation and temperature. The historic climate data are placed into each of nine bins.

A draw-back of this resampling approach can be under-dispersive data. For example, if the historic data does not show much variability, then the resulting synthetic climate might not contain the desired attributes that are suggested by the climate outlook. For example, if there is strong correlation between temperature and precipitation, such that when it is wet it is also cool (e.g. precipitation and temperature are highly correlated), then there will be more candidate days placed in the cool-wet and warm-dry bins (bottom-right and top-left), and fewer days in the cool-dry and warm-wet bins (bottom-left and top-right). If the outlook is "warmer and wetter", the resulting climate sequences will likely under-represent historic variability, as fewer days are available from which to resample. Once the historic data are appropriately binned, then the K-NN algorithm is used to select a new, synthetic sequence.

The IRI forecast provides probablist outlooks for precipitation and temperature according to three qualtivative attributes: above-normal, near-normal and below-normal and assigns a probability that the coming season will fall into each of those categories. For example, the outlook for precipitation in Southern California in January, February and March of 2010 is for a 50%, 35%, and 15% likehihood of being above, near, or below normal, while for temperature, the outlook is for an equal likelihood of being above, near or below normal (e.g. 33.3%, 33.3%, 33.3%). If the preciptiation temperature outlooks are expressed as a vectors, $\boldsymbol{p}$ and, $\boldsymbol{t}$ (3x1), respectively, the resulting weighting matrix is, $\boldsymbol{w} = \boldsymbol{p} \times \boldsymbol{t}^T$. For this January, February, and March, 2010 outlook, the resulting joint weighting matrix is given in Table 2. If 100 samples were drawn according to the number in each bin, the resulting climate sequence would statistically reflect this IRI December 2009, 90-day outlook for Southern California.

Table 2. Resulting weighting matrix, *w* for the December outlook for January, February and March, 2010.

Warm

5	12	16
5	12	17
5	12	16

Wet

Figure 6 shows a plot of ETo generated from the CaliforniaPET.org website for the December 2009 outlook that extends through March 2010, for CIMIS Station 6, near Davis California. The seasonal outlook for this site shows a 40%, 35%, and 25% chance and equal change above, near, and below normal preciption and temperature, respecitively. We used the CaliforniaPET.org website to generate 10 ensemble members for this site. The strong, red line is the Davis CIMIS station's historical ETo mean estimate and the dashed lines are the +- 1-standard deviation from the mean. The ETo outlook suggests a wide range of ETo values, indictiative of the equal change of temperatures being above, near, or below normal. However, the fact that temperature and precipitation are strongly correlated (e.g. wetter weather corresponds to cooler temperatures in the Central Valley) appears to suggest a moderated range in the spread of the ETo estimate. There is also a shift in the ETo estimate around March 15, whose result still needs to be explored.

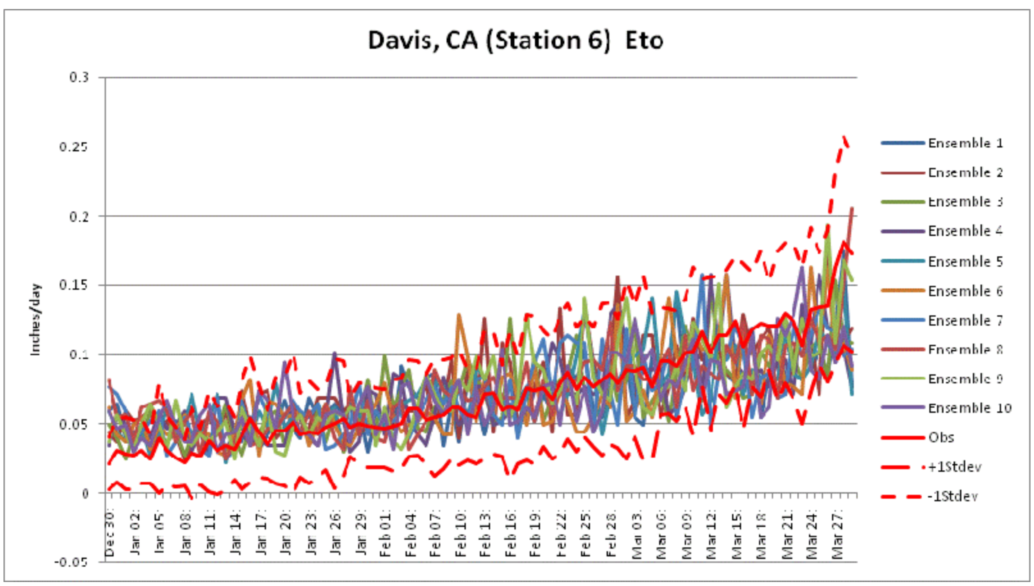

Figure 5. Plot ten ensemble members for the the 90-day ETo generated from the December 2009

Figure 6 shows a screen-shot of the CaliforniaPET.org website. The left-hand map inset is where the user can select a specific station of interest, with the color coding used to group stations according to their zone. The user can select the number of ensemble members (currently limited to 5), the output format (html, text, or DSS), and the calculation method (Penman Montieth, PM; or the CIMIS method). Once a station selection is made, the outlook weighting is cells are automatically updated, or the user can manual enter there values into the cells labeled “Hot”, “Mild”, “Cool”; and “Wet”, “Avg”, “Dry.

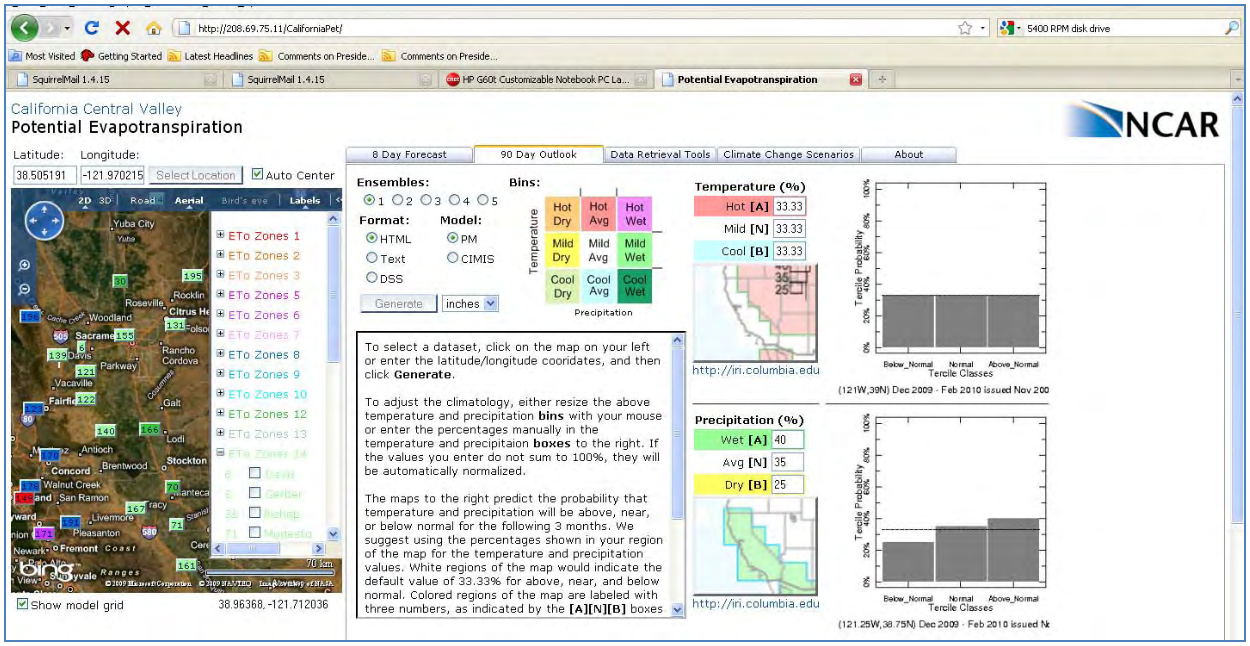

Figure 7. Screenshot of the CaliforniaPET.org’s 8-day forecast web page.

SUMMARY

This paper presents a methodology and its implementation via a web-based data portal for making and distributing 8-day forecasts and 90-day outlooks of ETo and precipitation for use in hydrologic and agricultural management models throughout California. This experimental project seeks to make ETo forecasts accessible and usable by resource managers by allowing for ease of use and the rapid ability to incorporate the ETo forecasts directly into models via url scripting, and text and DSS formats.

The forecasts and outlooks are made for select sites throughout California based on the California Irrigation Management Information System (CIMIS). The 8-day forecasts are based on NCEP High Resolution Global Forecast System (1 degree GFS) model output with site specific bias-correction. The 90-day ensemble outlooks are based on the International Research Institute for Climate Prediction’s (IRI) monthly forecasts of global temperature and precipitation out 3-months (e.g. 90-days) in advance. A web-site, http://californiaPET.org, allows the user to select specific stations of interest, and save the forecasts in either text or DSS format for direct use in water allocation models such as the United State Bureau of Reclamation’s (USBR) new Land Atmosphere Water Simulator (LAWS) model. The tool easily accommodates automation through custom Universal Resource Locator (URL) tags. In addition to ETo, the site also generates daily precipitation estimates in units of either mm/day or inches/day.

The 90-day outlook makes use of a non-parametric resampling technique known as K-Nearest Neighbor. Its heuristic nature allows for the generation of multiple, independent climate sequence that are conditioned from probabilistic climate outputs that assign weights to the likelihood of precipitation and temperature being above, near or below normal. The method preserves their joint distribution, which adds value to outlooks that are inconsistent with their natural correlation. The CaliforniaPET.org web-site facilitates the creation of an ensemble of local ETo forecasts and outlooks that can be used by LAWS and other models for forecasting seasonal water demand, placed within an uncertainty analysis framework.

The proof-of-concept data portal was conceived and developed in 2009 and is being made quasi operational the beginning of 2010 in anticipation of the 2010 growing season. Current work includes forecast and outlook verification, improvement of the bias correction of the 8-day forecast and developing a method to consider net radiation in ETo estimates for the 90-day outlook. Although not as important since ETo outlooks are of little value during non-growing periods, early evaluation of the 90-day outlook suggests over-estimation of ETo in the late fall and winter, since the Penman Montieth model requires an estimation of radiation and is not available in the Maurer et al. (2002) historic data archive. We will consider using the CIMIS archive in the near future to rectify the problem.

REFERENCES

Arca, B., Duce, P., Snyder, R.L., Spano, D. and Fiori, M. 2004. Use of numerical weather forecast and time series models for predicting reference evapotranspiration,. Acta Hort. (ISHS) 664:39-46

Clark M. P., S. Gangopadhyay, D. Brandon, K. Werner, L. Hay, B. Rajagopalan, and D. Yates, 2004: A resampling procedure for generating conditioned daily weather sequences, Water Resour. Res., 40.

DehghaniSanij H. H.T. Yamamotoa and V. Rasiahb 2004, Assessment of evapotranspiration estimation models for use in semi-arid environments, Agricultural Water Management, 64, (2), pp. 91-106.

Gorka Landeras, M.D. and Amaia Ortiz-Barredo, 2009, Journal of Irrigation and Drainage Engineering, Vol. 135, No. 3, May/June 2009, pp. 323-334, (doi 10.1061/(ASCE)IR.1943-4774.0000008)

Jones D., 1999, California Irrigation Management Information System: Reference Evapotranspiration, State of California, Department of Water Resources, P.O. Box 942836, 901 P Street, Sacramento CA 94236

Landman, W.A. and S.J. Mason, 1999. “Operational long-lead prediction of South African rainfall using canonical correlation analysis”, International Journal of Climatology, 19(10), pp. 1073-1090.

Mason, S.J., and Weigel, A.P. (2009): A Generic Forecast Verification Framework for Administrative Purposes. Mon. Wea. Rev., 137, pp 331–349.

Maurer, E.P., A.W. Wood, J.C. Adam, D.P. Lettenmaier, and B. Nijssen (2002). “A long-term hydrologically-based data set of land surface fluxes and states for the conterminous United States.” J. Climate 15(22), pp. 3237-3251.

Penman, H. L. 1948. "Natural evaporation from open water, bare soil and grass." Proc. Roy. Soc. London, A193, 120-146.

Rajagopalan, B., and U. Lall, 1999, A k-nearest-neighbor simulator for daily precipitation and other variables, Water Resour. Res., 35(10), pp. 3089–3101.

Sharma, M. L. 1985. Estimating evapotranspiration. p. 213-281 in Adv. in Irrigation, Vol III, D. Hillel (Editor)., Academic Press, New York.

Thomson, M. C., F. J. Doblas-Reyes, S. J. Mason, R. Hagedorn, S. J. Connor, T. Phindela, A. P. Morse, and T. N. Palmer, 2006: Multi-model ensemble seasonal climate forecasts for malaria early warning. Nature, 439, 576–579.

Yates, D., S. Gangopadhyar, B. Rajagopalan, and K. Strzepek, 2003: A technique for generating regional climate scenarios using a nearest-neighbor algorithm. Water Resources Research, 39(7), 1199.