ERE445/645
Hydrologic Modeling
Assignment #4
Due Thursday March 10, 2016, 5 pm submitted to Blackboard

Sometimes one is forced to create a hydrologic model using only a limited amount of data. For instance, often precipitation and temperature are the only recorded climatic variables, though to model many phenomenon (such as evapotranspiration), ideally other climatic information would be available. In such instances, we are forced to use the information we have to the best of our ability.

Consider a crude estimator of streamflow, Q:

$$Q = P - E \qquad (1)$$

where P is precipitation and E is evapotranspiration. Such a model assumes basin storage is constant over time. Over short time periods this is obviously not true (i.e. soil moisture, groundwater, and surface water quantities vary with time), but for longer time periods (say a year), such an assumption might be adequate.

METEOROLOGIC DATA DOWNLOAD

For this assignment, you will obtain 20 years of daily precipitation and temperature estimates for the 1941 – 1960 water years for a weather station near our study site, the West Branch of the Delaware River in Hale Eddy, NY. This data will come from the National Climatic Data Center for a recording station located near our study watershed (in Binghamton). [While there are other gauges located in our watershed, most contain large gaps in the data record, especially for the record from 1940 to 1960.] To obtain this file do the following:

a) Go to http://www.ncdc.noaa.gov/cdo-web/
b) Scroll down and click on the box "Search Tool".
c) In the "Select Weather Observation Type/Dataset" box choose "Daily Summaries"; choose our date range (1940-10-01 to 1960-09-30); select "Stations" in the "Search For" box, and enter "Delhi, NY" in the "Enter a Search Term" box. Click SEARCH.
d) Click ADD in the box for the "DELHI 2 SE, NY US" result on the map screen.
e) In the top right corner of the page click on "Cart (Free Data) 1 item" and then "View All Items (1)".
f) Check the date range to make sure it's right (you may have to tell it 1940-10-01 to 1960-09-30 again) and choose "Custom GHCN-Daily CSV" for the output format and click the "Continue" box.
g) Under "Station Detail . . . " select the box for "Station Name" and under "Select data type . . . " select the boxes for "Precipitation", "Air Temperature", and "Weather Type" and click the "Continue" box..
h) Entire your e-mail address and click "Submit Order".
a) Wait a few minutes for the file to be available (you will be sent an e-mail). Note that if you try to access this file prior to the e-mail it may not be complete.

For this assignment we will use three data columns: daily precipitation (PRCP in tenths of mm), daily maximum temperature (TMAX in tenths of ˚C), and daily minimum temperature (TMIN in tenths of ˚C). You will employ this data to model streamflow on the West Branch of the Delaware at Hale Eddy, NY for the water years 1941 to 1960 (20 years, starting October 1st, 1940 and ending September 30th, 1960).

IMPORTANT NOTE: There are many issues with this data set. For instance:
a) Some temperatures are recorded as -9999 (see TMIN for 12/18/41). If these are single missing values, take the average of the preceding and folloing values (here the average of the TMINs on 12/17/41 and 12/19/41). If there are multiple days (7/28-31/51), then linearly interpolate across these values.
b) Some daily precipitation values are recorded as -9999 (see PRCP for 2/4/43). In these cases we will set the daily precipitation to zero.
c) In addition, 1 day is completely missing!
ERE445 Students: You can manually adjust this file in Excel to correct for these issues.
ERE645 Students: You are to write a separate script for this assignment to automate this correctly process.

<u>STREAMFLOW MODEL DEVELOPMENT</u>

Consider a daily time step for the streamflow model in equation (1). There are a number of temperature-based models of evapotranspiration (e.g. Hargreaves, Blaney-Criddle, etc.). Here we will use the method by Hamon (Estimating Potential Evapotranspiration, *Proceedings of the American Society of Civil Engineers, Journal of the Hydraulics Division,* 87, 107-120, 1961). In this model:

$$E = 0.021\, N^2\, e_{st}/(T + 273) \qquad \text{if } T > 0 \qquad (2)$$
$$E = 0 \qquad \text{if } T \le 0$$

where E = potential evapotranspiration (cm/day)
N = number of hours of daylight on day t
e_{st} = saturated water vapor pressure (mb) on day t
T = average daily air temperature (°C)
= (TMAX+TMIN)/2 (assumed)

Saturated vapor pressure can be approximated by:

$$e_{st} = 6.108 \exp(17.27\, T/(237.3 + T)) \qquad (3)$$

While the number of daylight hours could be estimated as a function of the sunset hour angle, in practice average monthly N values are often used (i.e. the daylight hours are assumed constant for the entire month). For our study site, the following information has been estimated:

April	N = 13.1	Oct.	N = 10.9
May	N = 14.3	Nov.	N = 9.7

June	N = 15.0	Dec.	N = 9.0
July	N = 14.6	Jan.	N = 9.3
Aug.	N = 13.6	Feb.	N = 10.4
Sept.	N = 12.3	Mar.	N = 11.7

Haith and Shoemaker (Generalized Watershed Loading Functions for Stream Flow Nutrients, *Water Resources Bulletin,* 23(3), 471-478, 1987) employed Hamon's methodology to estimate E with average monthly daylight values. You are to run your model on a daily time step for the water years 1941 - 1960, and calculate the statistics reported in Homework Assignments 1, 2, and 3. ***Note that when running your model, you should allow for negative streamflows (when E > P).***

NOTE: Make sure you keep track of your units! E in equation (2) is in cm/day, but the final units from your model should be in cfs (as in other homework assignments). The drainage area for our study site is 1541 km^2 (595 mi^2).

Deliverables: In this assignment, I suggest you add to your existing HW#3 program since you will need to calculate similar statistics as well as access the streamflow data you have previously analyzed. You are to perform this analysis by creating new functions to calculate evaporation and predict streamflow, as well as determine the bias and Nash Sutcliffe Efficiency for your model. You are to hand in the following:

a) The commented main script that was developed to create the model
b) A script of functions employed by your main script.
c) Your input file.
d) A scatter plot containing the predicted (from this model) and observed (from HW1) annual average daily streamflow versus water year.
e) A scatter plot of the predicted (from this model) and observed (from HW3) average daily streamflow for each month averaged across the 20 years (the month should be on the x-axis).
f) The value of the bias and Nash Sutcliffe Efficiency for annual average daily streamflow and monthly average streamflows, which can be calculated as follows:

Bias:

$$Bias(\hat{\theta}) = \frac{\sum_{i=1}^{N}\left(\hat{\theta}_i - \theta_i\right)}{N}$$

where:

θ_i	=	ith observation of the quantity to be estimated (here either annual average daily streamflow or monthly average streamflow.
$\hat{\theta}_i$	=	ith estimate of θ_i (from your model)
N	=	number of observations (N = 20 for annual average daily streamflow and N = 240 for monthly average streamflow)

We ideally want the bias to be close to 0.

The Nash-Sutcliffe Efficiency (NSE):

$$NSE(\hat{\theta}) = 1 - \frac{\sum_{i=1}^{N}\left(\hat{\theta}_i - \theta_i\right)^2}{\sum_{i=1}^{N}\left(\theta_i - \bar{\theta}_i\right)^2}$$

where: $\bar{\theta}_i$ = the average value of θ_i across the N observations

We ideally want the NSE to be close to 1. Note the NSE is similar to the coefficient of determination (R^2) employed in regression modeling.

g) A short discussion of the main problems with the model you created. This should include a list of 3 changes/additions you would make to improve your model. You should reflect on the plots produced in parts d) and e) to motivate your suggested changes/additions (i.e. where is the model doing well and where the model is doing poorly?).