

Statistical toolkit for climate analysis of hydrologic time series

Jesse Dickinson¹, Randy Hanson²

Problem: The potential consequences of climate variability and climate change were identified as major issues for the sustainability and availability of the water resources of the United States (Gleick and Adams, 2000). Unlike global climate change, climate variability represents reversible and periodic changes in the global weather systems that occur over periods of a few years to several decades. A systematic approach to analyzing hydrologic time series for climatic variability was established (Hanson and others, 2004) as part of the Southwest Ground Water Resource Project. Relations between climatic variability and hydrologic time series (groundwater levels, streamflow, precipitation, and tree-ring data) in the Southwestern United States (Hanson and others, 2006) established the first step to estimating time-varying ground-water recharge based on inferred climatic variability. The ability to estimate periodic groundwater recharge from single and multiple climate cycles was then established from these climate cycles (Dickinson and others, 2004).

The combination of data preparation and statistical methods recently developed by the USGS as part of the Ground Water Resources Program is relatively unavailable to hydrologists and engineers that may benefit from the estimates of climate variability and its effects on periodic recharge and water resource availability. A statistical toolkit needs to be created for hydrologists and engineers to facilitate investigations of the effects of climate variability within their unique study areas or types of hydrologic time series.

Objective: The objective of this project is to develop a software toolkit that 1) identifies responses of hydrologic systems to climate variability, 2) quantifies statistical relations between multiple time series and known climate indices, and 3) projects hydrologic time series using time series models and spectral analysis for assigning groundwater model boundary conditions and calibration targets.

Relevance and benefits: The methods of studying climatic variations within hydrologic time series has been established but is not available to a wide variety of users that can potentially benefit from this analysis. The USGS already maintains a network of about 150 wells as part of the Ground Water Resources Program to monitor the effects of droughts and other climate variability on groundwater levels. While this network is a valuable resource for groundwater statistics, it lacks a systematic and composite analysis tool. Groundwater data on the Climate Response Network can be incorporated with streamflow, precipitation, and tree-

¹ Arizona Water Science Center

² California Water Science Center

ring data that can all be analyzed individually and in combination with respect to climate variability. These analyses provide a basis for understanding climate variability, trends, and estimates of periodic recharge needed for groundwater flow modeling.

Approach: The toolkit will be released as a Geographic Information System (GIS) application (plug-in) composed of a dynamic link library and a Visual Basic application file. Inputs are long-term time series data of groundwater levels, streamflow, precipitation, tree ring data, and climate indices. Input can comprise single or multiple time-series for batch analysis from a specific region of study. Output can be interpreted within the GIS or exported to spreadsheet software.

The toolkit includes the following features:

- Data acquisition and preprocessing
 - Retrieval from NWIS or manual entry
 - Formatting, windowing, grouping, and transformation to generate standardized time series for statistical analyses.
- Analyses
 - Removal of non-stationary elements
 - Frequency analysis with Singular and Multiple Spectral Analysis
 - Principal Component Reconstructions including inverse transformations
 - Statistical Estimation—Correlation, lag, covariance, and other multi-series relations
- Projections
 - Estimation of periodic ground-water recharge using analytical groundwater flow models from single or multiple frequencies (Fig. 5)
 - Estimation of projected (future) hydrologic time series with an ARMA time series model (Fig. 6)
 - Observations and simulated equivalents for groundwater model calibration that are based on climate-related responses

Program Design: The toolkit will comprise a graphical user interface (GUI), a database system, and programs for preprocessing, analysis, and projections (Figure 1). The GUI can be run on any Windows operated personal computer. The GUI will be accessible as a stand-alone application and within GIS software. The GUI guides the user in selecting data input types, choices of statistical calculations for selected time series, and output options. Steps done by the toolkit are listed below.

1. Input hydrologic time series: The toolkit acquires time-series data through manual user input, by a batch load from locally-stored files, or by retrieval from NWIS. The database stores these single or multiple time series that are accessed by other computational processes of the toolkit. Having a local database expedites multiple statistical calculations across multiple data types.

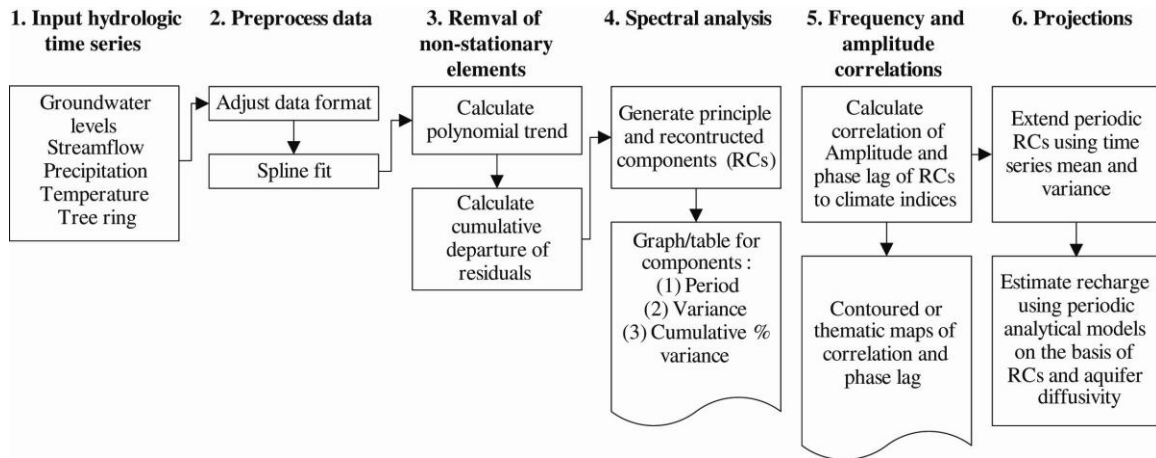


Figure 1 Workflow by the climate variability toolkit for climate analysis in hydrologic time series.

2. Preprocess Data: Data preprocessing steps are done to standardize data formats required by the toolkit. Necessary preprocessing steps often include (1) changes to calendar dates to a decimal date format, and (2) spline fits that smooth time periods with no data and that create an evenly-sampled data set.

3. Removal of non-stationary elements: Data evaluation includes identification and removal of non-stationary elements and any low-frequency cycles. Non-stationary elements are trends in the mean or variance that are caused (1) by human activity or changes in watershed characteristics, or (2) by changes in climate that available information cannot identify as a repeating cycle. Non-stationary trends can follow a linear pattern, curvilinear pattern, a step-change pattern, or other pattern. Low-frequency cycles are removed in this step because the spectral analysis or time-series analysis in step two cannot identify low-frequency cycles that are longer than half the period of record. The non-stationary trend or low-frequency cycle are removed by fitting a regression relation or other fitting technique to the trend or cycle. Spectral analysis is done on the residuals from the fitted relation.

4. Spectral Analysis: Time-series frequency analysis will be done using singular-spectrum analysis (SSA), which is a form of principle component analysis in lag-time domain (Vautard and others, 1992). SSA detects periodic signals in short and noisy time series. The toolkit will embed computational routines implemented by Dettinger and others (1995) to isolate and reconstruct significant periodic components (RCs) in the time series. The RCs are nearly-sinusoidal oscillations (Figure 2) typically with periods between 2–50 years that often correlate to known climate indices such as ENSO and PDO (Hanson and others, 2004). The sum of all RCs may account for up to 60% of the total variance for a time series. The remaining variance is complex noise that is preserved in order to generate time series projections (see step 6).

5. Frequency and Amplitude Correlations: Frequency and amplitude correlations are statistical estimates of how closely related two time series are as they change in frequency and amplitude domains. Frequency and amplitude lags are the phase shift (in months) of the

dependent time series for the best correlation. Identical time series will have a frequency correlation of 1 and a lag of 0. Users can compute correlation statistics of the RCs of any time series. The toolkit will generate graphs, tables, and maps showing correlation statistics for display within GIS or other plotting software.

6. Projections: Time series projections can be generated using a combination of random-number synthesis and extrapolation of the cyclical reconstructed components as demonstrated by Hanson and others (2002). The projections are statistically-realistic extensions of previous climatic conditions. RCs which include most of the total time series variance can be projected using autoregressive models (e.g. an ARMA time-series model). Because the RCs do not include all variance of the original time series, additional random-looking residuals that are based on the remaining variance are added to the projection.

Time-varying recharge related to climate variability can be estimated on the basis of RCs of groundwater levels (Dickinson, 2004). Recharge rates obtained from the toolkit can provide users with initial estimates of the magnitude of variable recharge rates, which may assist in developing conceptual and numerical groundwater flow models. Recharge is estimated by the inverse application of an analytical one-dimensional groundwater flow model for periodic boundary conditions (Townley, 1995). The inverse application requires the following data: groundwater level RCs, distances from likely recharge and discharge areas, and some knowledge of aquifer diffusivity.

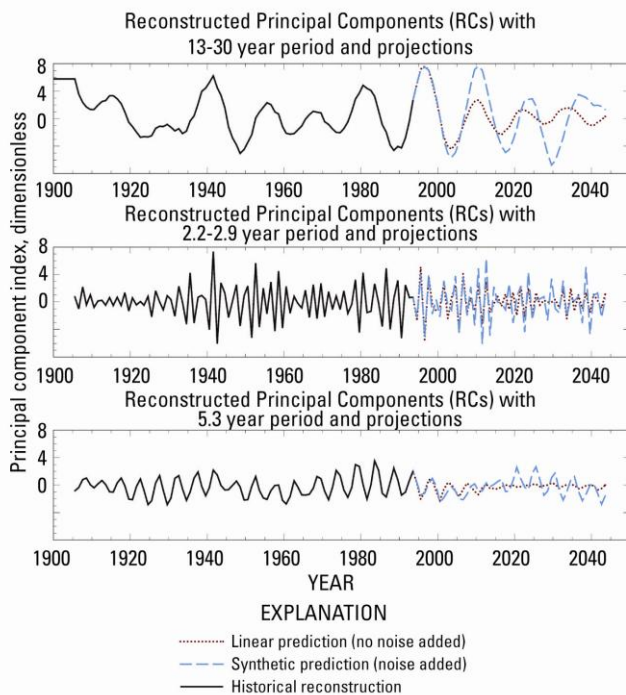


Figure 2. Example of precipitation projections based on ARMA time series models built from RCs (Modified from Hanson and others, 2002).

Output options: Outputs of the SSA include graphs and tables that show period (years) of the RCs and the percent of total variance associated with the RCs. Outputs from the frequency and amplitude correlation calculations include graphs of RCs and phase-lag adjusted RCs with reference climate index time series, contoured maps of correlation and phase lag. Time-series projections include a series of extrapolations that can be used to assign groundwater model boundary conditions as demonstrated by Hanson and others (2002) (Figure 2).

Products: The toolkit will be documented in a USGS Techniques and Methods or Open-File Report series report. An executable program file and the computer code will be available for public download. Example datasets in the documentation will be selected to support ongoing water availability studies.

Response to Challenge Area 3: Monitoring Effects of Climate Change on Groundwater

Arizona and California Water Science Centers

Personnel: The staff required for the project will include one mid-level hydrologist and one senior-level research hydrologist. The staff will include Jesse Dickinson (900 hours in both FY2010 and FY2011; 400 hours in FY2012), Randy Hanson (300 hours in both FY2010 and FY2011; 150 hours in FY2012) and a GIS specialist (80 hours in both FY2010 and FY2011).

Budget:

| Net Costs | | | | |
|--------------------|-----------|-----------|----------|-----------|
| Category | FY 2010 | FY 2011 | FY 2012 | Total |
| Salary | \$64,000 | \$65,000 | \$30,000 | \$159,000 |
| Training | \$1,000 | \$0 | \$0 | \$1,000 |
| Travel | \$2,600 | \$2,600 | \$1,900 | \$7,100 |
| Postage | \$100 | \$100 | \$100 | \$300 |
| Printing | \$0 | 0 | \$8,000 | \$8,000 |
| Supplies | \$500 | \$500 | \$500 | \$1,500 |
| Contingency | \$500 | \$500 | \$500 | \$1,500 |
| Total | \$68,700 | \$68,700 | \$41,000 | \$178,400 |
| Assessments | | | | |
| Distributed Direct | \$8,300 | \$8,300 | \$5,000 | \$21,600 |
| DOTSC | \$23,000 | \$23,000 | \$14,000 | \$60,000 |
| Gross Costs | | | | |
| Total | \$100,000 | \$100,000 | \$60,000 | \$260,000 |

References:

Dettinger, M. D., M. Ghil, C. M. Strong, W. Weibel, and P. Yiou, 1995, Software expedites singular-spectrum analysis of noisy time series: *Eos Trans. AGU*, 76(2), 12, 14, 21

Dickinson, J.E., R.T. Hanson, T.P.A. Ferré, S.A. Leake, (2004) Inferring time-varying recharge from inverse analysis of long-term water levels, *Water Resources Research*, 40, W07403.

Gleick, P.H. and Adams, D.B., 2000, Water: the potential consequences of climate variability and change for the water resources of the United States: the report of the Water Sector Team of the National Assessment of the Potential Consequences of Climate Variability and Change for the U.S. Global Research Program, 151 pp, Pac. Inst. for Stud. In Dev., Environ., and Security, Oakland, Calif.

Hanson, R.T., Dettinger, M.D., and Newhouse, M.W., 2006, Relations between climatic variability and hydrologic time series from four alluvial basins across the southwestern United States: *Journal of Hydrogeology*, vol.14, no. 7 p. 1122-1146, doi:10.1007/s10040-006-0067-7.

Hanson, R. T., P. Martin, and K. M. Koczot, 2002, Simulation of ground-water/surface-water flow in the Santa Clara-Calleguas Basin, Ventura County, California: *U.S. Geol. Surv. Water Resour. Invest. Rep.*, 02-4136, 214 pp.

Hanson, R.T., Newhouse, M.W., and Dettinger, M.D., 2004, A methodology to assess relations between climatic variability and variations in hydrologic time series in the southwestern United States: *Journal of Hydrology*, 287, p. 252-269.

Townley, L.R., 1995, The response of aquifers to periodic forcing, *Adv. Water Resour.*, 18, 125-146.

Vautard, R., P. Yiou, and M. Ghil, 1992, Singular-spectrum analysis: A toolkit for short, noisy chaotic signals, *Physica D*, 58, 95–126