The Paradox of Stationarity in Hydrology: Toward a Deterministic Solution[1]

Messele Zewdie Ejeta

Abstract

A retrospective analysis of over a century of observed full natural flow data of four rivers in California shows that the average water production of each river could have been predicted with about 96% precision on a climate time scale. Paradoxically, the water production of the watersheds for shorter periods could not have been predicted with the same level of precision. The realization of this level of stationarity and its counter intuitive paradox for a shorter period avail us a unique problem in which the moving average of a continuous subset of any *n* data points from a total set of *N* data points is predictable with a very high degree of precision whereas the natural variability of the components in each subset remains to be meaningfully characterized. This paper: 1) provides a formulation for this uniquely paradoxical problem of stationarity, and 2) presents empirical evidences for possible forcing factors of stationarity, as observed in tide level trend shifts in the Pacific Ocean, long-term natural streamflow extremes in California, and the shift in the phases of the Pacific Decadal Oscillation (PDO). The implication of this formulation and the presentation of these empirical evidences appear to point to underlying forcing factors that can be generally characterized beyond what has been understood so far. The characterization and solution of the problem will have immense importance for a robust water resources planning as well as for the prediction of extreme events such as droughts and floods, including under projected climate change scenarios.

Introduction

Stationarity has been qualitatively described as the idea that natural systems fluctuate within an unchanging envelope of variability (Milly, et al. 2008). While the hydrologic cycle theory attempts to explain the path for the transfer of water from one state to another and its distribution within a given state, it lacks a robust metric that is based on a deterministic process to measure the transfer rate of water from one store to another. The efforts of hydrologists to close earth's water balance by measuring where and in what quantities earth stores water and how water moves between those stores are yet to provide convincing results (Lettenmaier and Famiglietti, 2006). To the extent that quantifiable water amounts in earth's stores can be linked to earth's climate

---

[1] This article was submitted on January 9, 2010, to the Journal of Hydrology for review and possible publication. A letter of rejection was received on August 28, 2010, nearly 9 months later.

or vice versa, a better understanding of the natural mechanisms that cause the transfers of water from one store to other stores is essential for the analysis of both the natural and possibly induced variability in hydrological data. A closer study of the forcing factors for these transfers and their periodicities is likely to hold the keys to solve the unique problem of stationarity that is formulated herein and supported by over a century of observed full natural flow data of four major rivers in the San Joaquin Valley of California. Thus, the first effort of this paper is to present this formulation as a quantitative representation for the idea of stationarity as an added value to its qualitative description that was noted earlier. The second effort of this paper is to present observed empirical evidences relating the shifts in the trends of tide levels in the Pacific Ocean, long-range natural streamflow extremes in California, and the shift from cool to warm phases in the Pacific Decadal Oscillation (PDO).

A full natural flow, which is also called unimpaired runoff, is defined by the California Department of Water Resources as the flow that represents the natural water production of a river basin, unaltered by upstream diversions, storage, or by export or import of water to or from other watersheds (CA DWR, 2009). For a given river location, this flow is estimated by adjusting the recorded or gage flow for upstream operations for the built environment. The four river locations in California's San Joaquin Valley where over a century of estimated full natural flow data is used for this study include: 1) Stanislaus River at New Melones Dam, 2) Tuolumne River at New Don Pedro Dam, 3) Merced River at Lake McClure, and 4) San Joaquin River at Friant Dam, which are all in the western foothills of the Sierra-Nevada mountain range.

The nature of this unique problem of stationarity, as realized by the observed full natural flow data of these rivers according to the stationarity formula presented in this paper, is likely to be a function of measureable forcing factors and their periodicities that can be characterized. If done successfully, this characterization will enable us to predict over an extended period of time the transfer rate of water from one store to another, which will have immense significance for a robust water resources planning as well as the management of extreme events, such as droughts and floods.

Several researchers have been making efforts to characterize discernible forcing factors such as solar forcing (Willett, 1974) and multi-decadal variations due to natural factors such as the North American Oscillation (McCabe, et al., 2004, McCabe, et al., 2007). However, long-term prediction of quantifiable river flow that is not predicated on the statistics of observed data is currently practically non-existent. Willett (1974) made an important effort to link solar-climatic cycles to climatic trend forecasting.

McCabe, et al. (2004) and McCabe, et al. (2007) attempted to link drought occurrences in the conterminous United States and sea surface temperature variability in both the tropical Pacific and North Atlantic oceans on a decadal to multi-decadal time scales. Wu, et al. (2009) presented an empirical model to predict the East Asian Summer Monsoon strength using the El Niño Southern Oscillation (ENSO) and the spring North Atlantic Oscillation (NAO).

These empirical evidences appear to show correlations between the momentum of the trends of these natural variables and the unique problem of stationarity that was realized in the full natural flow data of the four rivers. Even though emerging studies suggest that stationarity is dead (Milly, et al., 2008), the discernible factors behind the above variables and stationarity are yet to be definitely characterized. In the likely scenario that these factors contribute to the realization of stationarity, they may continue to play a dominant role even under presumed dead stationarity, or projected climate change. In other words, even in the projected altered state of the earth that has been attributed to elevated green house gas concentration in the atmosphere, natural variability will most likely continue to live as the undercurrent beneath dead stationarity. Therefore, a meaningful characterization of natural variability of the past that maintained stationarity will help in the attribution of future variability due to nature and the influence of humans.

Stationarity in Observed Data

The long-range observed full natural flow data by water year is used herein to provide an overview of stationarity. In California, a water year runs from the beginning of October of the previous calendar year to the end of September of the current calendar year. This characterization of a water year was recommended by Loewe and Radok (1948) to avoid splitting the Southern Hemisphere summer wet season, or equivalently, the Northern Hemisphere winter wet season. Although precipitation data would be preferable for such an analysis, since it doesn't contain possible bias due to the abstraction, surface evaporation, and infiltration of water, long-range and distributed historical precipitation data is not as widely available as the flow data. Thus, the flow data is used as a proxy for the precipitation data.

The analysis of the full natural flow data of the four major rivers in California using a 30-year climate period's moving average flow traces for these rivers shows that the average water year flow for any given climate period since 1901 is nearly deterministic, as shown in Figure 1. The long-range average water productions of the Stanislaus, Tuolumne, Merced, and San Joaquin rivers that drain to the stated locations are 1.2, 1.9, 1.0, and 1.8 million acre-feet, respectively. Note that the graph shows pronounced peaks for the climate periods ending in water years 1938 and 1998 and a pronounced trough for the climate period ending in water year 1968. Note also that the trough is situated between the peaks, thus suggesting conceivable periodic characteristics in the data. This will become more apparent that looks at the time series residuals of stationarity (Figure 2). Although these pronouncements don't appear to be particularly large for a 30-year average data, they are caused by single data points that stand out from among the data points used in the averaging.

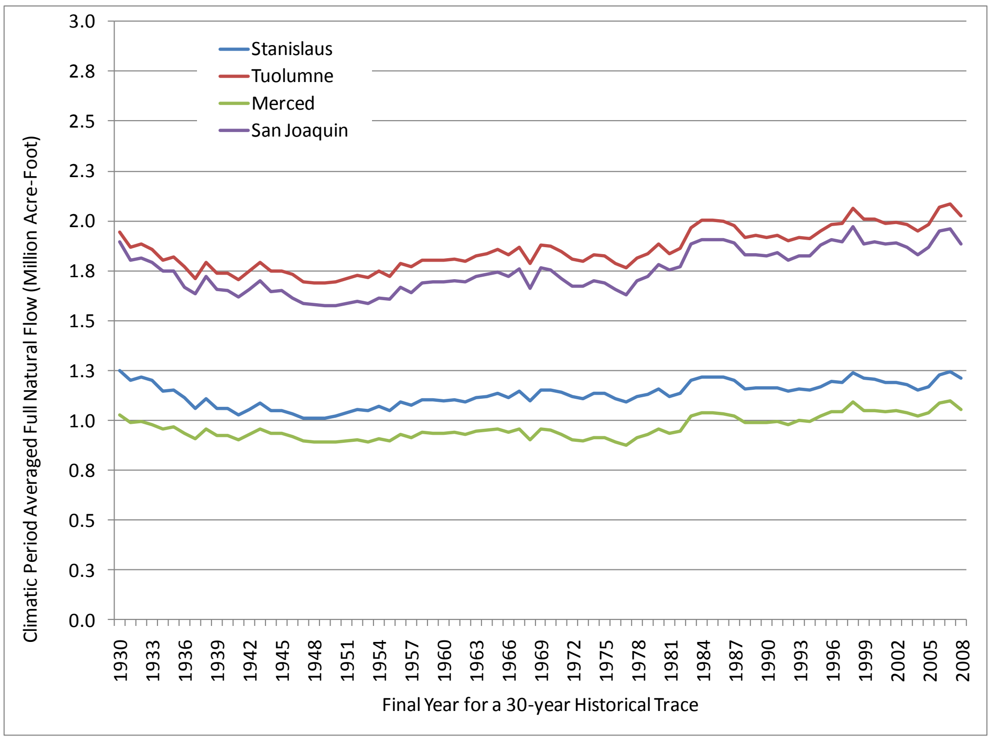

Figure 1. Observed 30-year moving average trace of full natural flow data of the four major rivers in the San Joaquin Valley of California

Formulation of the Hydrological Stationarity Problem

As the 30-year moving average traces in Figure 1 show, the trend lines of the graphs for all the four rivers have nominal slopes and practically identical patterns over time throughout the period of record. For any given period of consecutive 30 years, the average water production of any of the watersheds can be well estimated by the long-term average water production of the respective watershed. Thus, if $Q_1, Q_2, Q_3, \ldots, Q_N$ are the water year full natural flows of a given river for N years where $N \geq 30$, we can define the average water production of the watershed using a simple arithmetic mean formulation given by Equation 1.

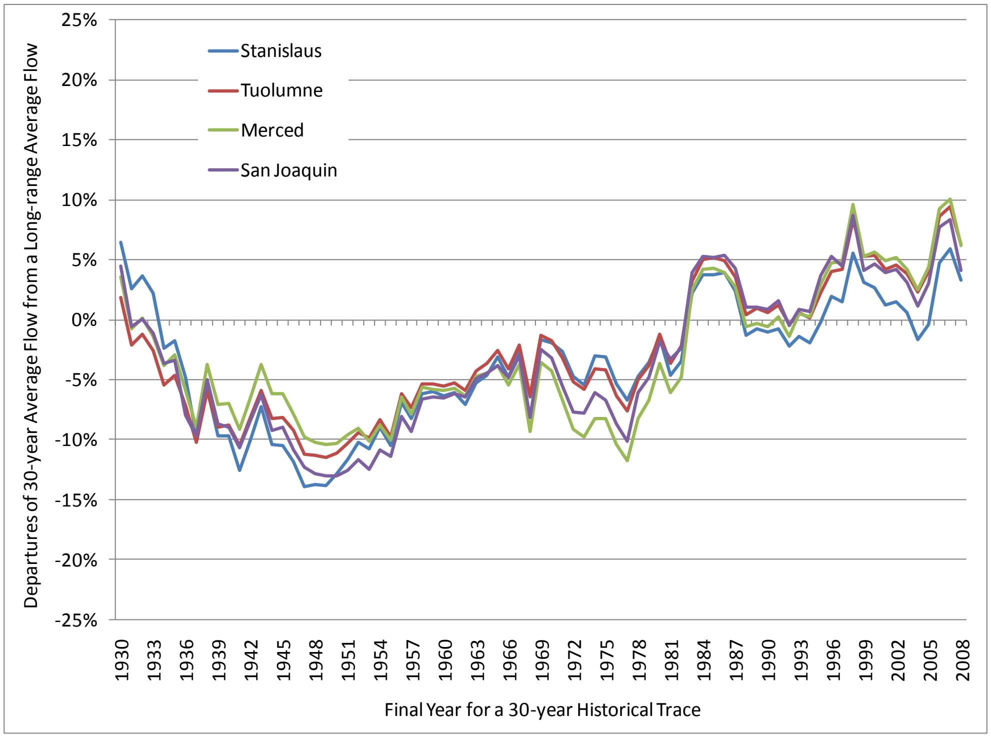

Figure 2. Departures from 30-year moving average trace of full natural flow data of the four major rivers in the San Joaquin Valley of California

$$Q_y = \frac{\sum_{i=1}^{N} Q_i}{N} \tag{1}$$

where Q_y is the average water production of the watershed.

For any n consecutive water years that are a subset of N water years, the paradoxical stationarity problem, using observed full natural flow data as a case example, is formulated as follows.

$$Q_y = \frac{\sum_{i=1}^{n} Q_i}{n} - \varepsilon_1 = \frac{\sum_{i=2}^{n+1} Q_i}{n} - \varepsilon_2 = \frac{\sum_{i=3}^{n+2} Q_i}{n} - \varepsilon_3 = \cdots = \frac{\sum_{i=(N-n+1)}^{N} Q_i}{n} - \varepsilon_{(N-n+1)} \tag{2}$$

where $\varepsilon_1, \varepsilon_2, \varepsilon_3, \ldots \varepsilon_{(N-n+1)}$ are residual departures from stationarity. For practically insignificant residuals, Equation (2) may be rewritten as follows:

$$Q_y \approx \frac{\sum_{i=1}^{n} Q_i}{n} \approx \frac{\sum_{i=2}^{n+1} Q_i}{n} \approx \frac{\sum_{i=3}^{n+2} Q_i}{n} \approx \cdots \approx \frac{\sum_{i=(N-n+1)}^{N} Q_i}{n} \quad (3)$$

Both equations (2) and (3) can be extended to all other variables that have the characteristics of stationarity. The period *n* that is used for averaging to establish stationarity may be called the stationarity period. The residual departures, ε_1, ε_2, ε_3, …, $\varepsilon_{(N-n+1)}$, may be called the departures from stationarity. Under theoretical stationarity:

$$\lim_{p \to n} \varepsilon_1 = \lim_{p \to n} \varepsilon_2 = \lim_{p \to n} \varepsilon_3 = \cdots = \lim_{p \to n} \varepsilon_{(N-n+1)} = 0 \quad (4)$$

where *p* is the test case period for the stationarity period *n*, which could vary from an instantaneous time, such as for the measurement of the mass of a stable atom, to infinity, such as for the known state of our universe. Under practical stationarity, equation (4) may be modified as follows:

$$\frac{\sum_{i=1}^{N-n+1} \varepsilon_i}{N-n+1} \leq \varepsilon_T \quad (5)$$

where ε_T may be called the average tolerance limit for departures from stationarity.

As the data of the graphs in Figure 2 indicate, the average of the departures from stationarity of all the successive 30-year periods from the long-term average of the full natural flows of all the four rivers is about -3%. The biggest observed departure is about -14% for the Stanislaus River above New Melones Dam during the 1925 – 1954 climate period. In hindsight, at any time in the past since 1901, the average water productions of any of these watersheds for the subsequent 30 years could have been predicted with about 96% precision on average and 86% precision in the worst case scenario. Therefore, the paradox of this problem is that if the water production of a watershed for the next thirty years, starting any year in the last century, could be predicted with this high level of precision, why can't the water production of the same watershed for the subsequent 5 or 10 years be predicted with the same level of precision? This is a counter-intuitive problem that demands characterizing the causes. Further investigation of observed empirical evidences that relate extreme data points of interest in the stationarity with the trend shifts in the Pacific Ocean tide levels and the phase shifts in the PDO may provide an area of focus in continuous endeavors to solve this unique problem.

Tide Level Trend Shifts

Figure 3 shows the Pacific Ocean's tide level for the period of 1900 to 2007 as observed at the Golden Gate Bridge, off the coast of California near San Francisco. A 19-year moving average of this tide level data is also shown in Figure 3 for three time segments, which show apparent shifts in trend. These segments are the 19-year periods ending between 1918-1933, 1934-1998, and 1999-2007. Using the slope of the trend line for the 19-year moving average that ends between 1918 and 1933 as the baseline, the trend line slopes for the subsequent two periods change by 5.4 and -2.6 times, respectively. From 1924 to 1998, there was a marked increasing trend in the tide levels; note that 1924 is the mid-point for the 19-year moving average that ends in 1933. Since 1998, there is a marked falling trend in the tide levels. Of particular importance here are the years 1924 and 1998, the first one being one of the two driest years on record in California since 1901 and the second one being one of the wettest years on record in California. Whether the marked falling trend since 1998 can be fully explained by arguments made in favor of global cooling since 1998 (Easterbrook, 2009) or that future warming will be strongly modulated by natural climate variations especially those driven by the slowly varying oceans on a time scale of decades (Hurrell, 2008) will be a subject of further investigation.

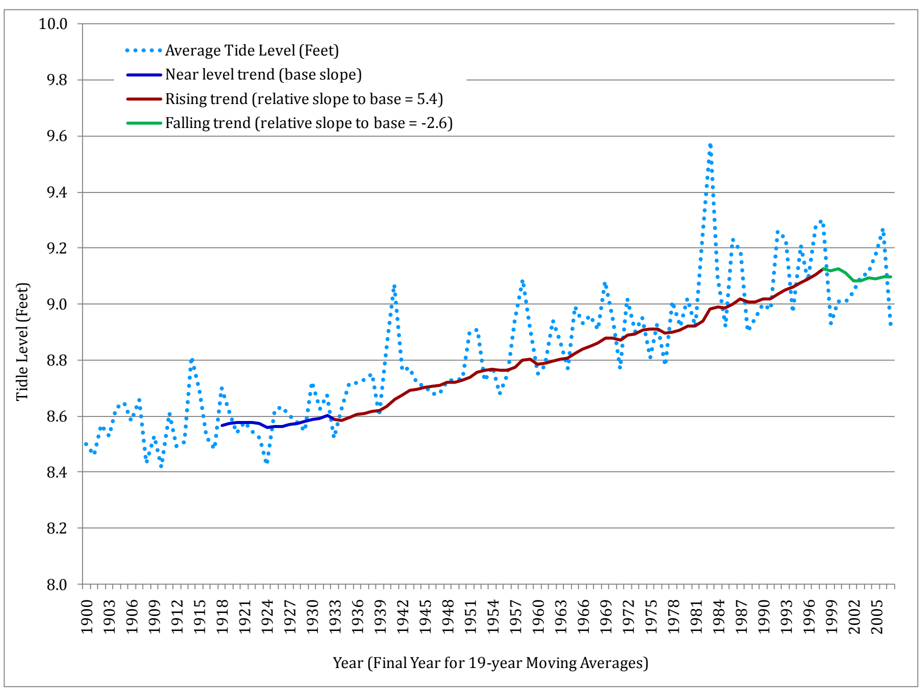

Figure 3. Pacific Ocean's tide levels and their 19-year moving averages, as observed at the Golden Gate Bridge near San Francisco

Figure 4 shows the tide level changes for the three time segments noted above. The biggest fluctuations in the tide levels were observed between 1942-1943 and 1983-1984, which were after the Dust Bowl period in the U.S. and the wettest year on record in California, respectively. Coincidentally, the time between 1983 and 1984 saw the highest tide level fluctuation on record at this location. In addition, since this highest tide level fluctuation, the positive peaks in the tide level show a dampening trend at this location.

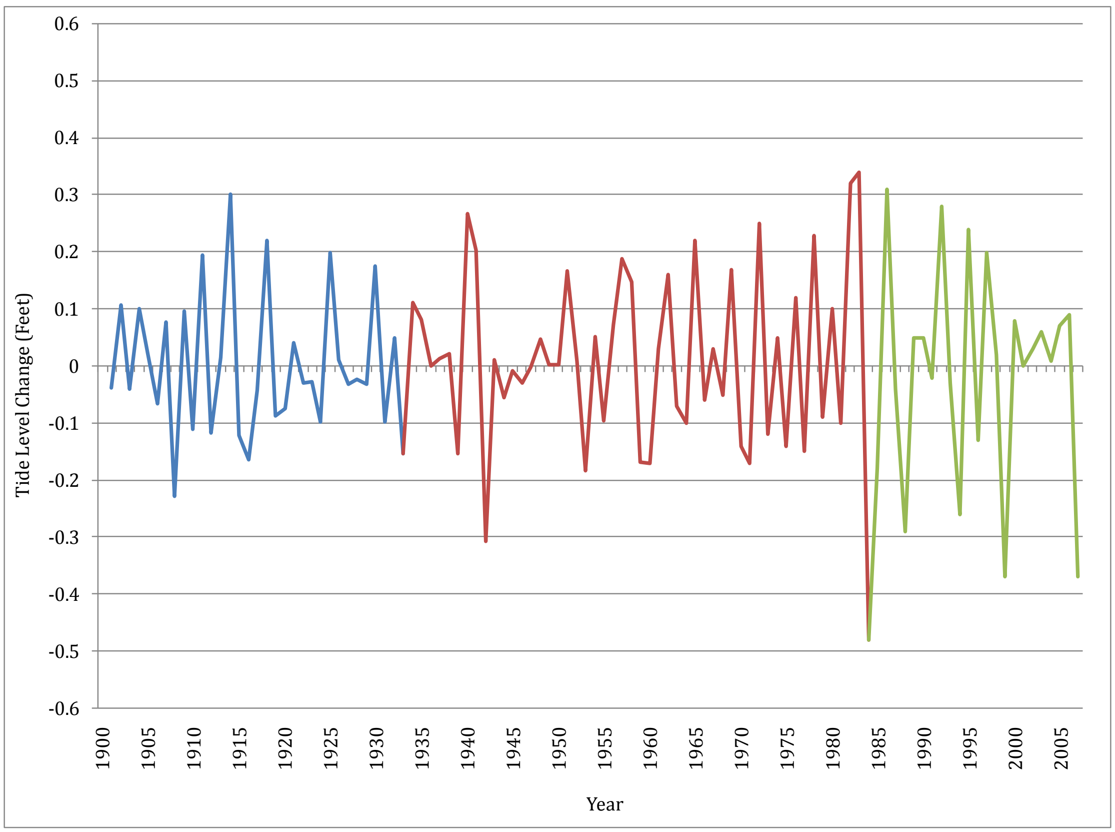

Figure 4. Pacific Ocean's tide level changes as observed at the Golden Gate Bridge

Pacific Decadal Oscillation (PDO) Phase Shifts

The Pacific Decadal Oscillation (PDO) is a long-lived El-Niño like pattern of Pacific climate variability. The PDO Index is defined as the leading principal component of North Pacific monthly sea surface temperature variability (pole ward of 20º north for the 1900–1993 period). "Cool" PDO regimes prevailed from 1890 – 1924 and again from 1947 – 1976, while “warm” PDO regimes dominated from 1925 – 1946 and from 1977 through at least the mid 1990’s (University of Washington, 2009).

The analysis of the full natural flows of the four major rivers in California in relation to the PDO Index showed a strong correlation between the two “cool” to “warm” PDO regime shifts on record and the two lowest natural streamflows on record since 1901. For all these rivers, the two lowest natural flows occurred in 1924 and 1977 whereas the two “cool” to “warm” PDO regime shifts are noted to have occurred between 1924 and 1925 and 1976 and 1977.

Discussion

Going beyond a qualitative description of the idea of stationarity, over a century of observed full natural flow data of four major rivers in the San Joaquin Valley of California have been used to formulate and quantitatively illustrate the idea of stationarity in hydrological data. The data shows, in hindsight, that starting any year since 1901 the water production of the watersheds of these four rivers could be predicted, on average, with 96% precision on a climate time scale of 30-years. Based on this analysis, a paradoxical question has been raised that if we could have predicted, starting from any year since 1901, the average water production of a given watershed in the subsequent 30 years with such a very high level of precision, why couldn’t have we predicted the water production of the same watershed for a shorter time period, such as 5 or ten years, with the same level of precision?

It has been argued that without some inherently underlying forcing factors that can conceivably be characterized, a subject of an ongoing investigation, there wouldn’t be such a paradoxical stationarity in hydrology. The nature of the historical departures from stationarity is far from being random. In fact, at any given time, the pattern of these residuals may provide us clues about a natural inertia of the direction of the trend of river flows on multi-decadal time scales. The inertia of these departures suggests that natural variability is likely to continue to exist beneath what is believed to be dead stationarity. Establishing a baseline for historical natural stationarity will provide a reference that can be used to estimate the shift from this baseline.

Empirical evidences relating observed full natural flow data extremes to significant trend shifts in the Pacific Ocean’s tide levels and phase shifts in the PDO that have been presented appear to provide useful information about the forcing factors for stationarity. The coincidences of these shifts and the occurrences of extreme full natural flow data points of the four rivers in the San Joaquin Valley of California appear to suggest that some relationship exist between these natural variables. To the extent that natural variability lives beneath dead stationarity, a better understanding and characterization of the natural variability component of hydrology is likely to go a long way in overcoming the uncertainties in the use of such data that is based on the results of General Circulation Modeling. The tide level trend shifts on a multi-decadal time scale appears poised to lead us to revisit the universal law of gravity that assumes the earth as a monolithic mass, instead of an integral mass with its various components that may have their own centers of masses that may shift marginally. In fact, ocean level data shows the gravitational influence of the moon and the sun on the ocean level (Rahmstorf, 2002). The current assumption is that the gravitational pull of the ocean mass by the sun and the moon leaves the ocean tide's wave front in its integrity, which is a subject of focus in ongoing investigations as part of an effort to move toward a deterministic solution for the paradoxical hydrological stationarity problem.

References

California Department of Water Resources (2009,) http://cdec.water.ca.gov/cgi-progs/stages/FNF

Easterbrook, D. (2009,) Global Cooling is Here! Evidence for Predicting Global Cooling for the Next Three Decades**,** http://icecap.us/index.php/go/joes-blog/global_cooling_is_here_evidence_for_predicting_global_cooling_for_the_next_ (accessed November 30, 2009)

Hanson, R. T., M. W. Newhouse, M. D. Dettinger (2004), A Methodology to Assess Relations between Climatic Variability and Variations in Hydrologic Time Series in the Southwestern United States, Journal of Hydrology, 287, 252-269.

Hurrell, J. W. (2008), Decadal climate prediction: challenges and opportunities, Journal of Physics: Conference Series, doi:10.1088/1742-6596/125/1/012018.

Lettenmaier, D. P. and J. S. Famiglietti (2006), Water from on high, *Nature*, 444(30).

Loewe, F. and Radok, U. (1948), Variability and periodicity of meteorological elements in the Southern Hemisphere with particular reference to Australia, Canberra, Commonwealth of Australia.

McCabe, G.J., M. A. Palecki, J. L. Betancourt (2004), Pacific and Atlantic ocean influences on multidecadal drought frequency in the United States, Proceedings of the National Academy of Sciences, 101 (12), 4136-4141.

McCabe, G. J., J. L. Betancourt, S. T. Gray, M. A. Palecki, H. G. Hidalgo (2008), Associations of multi-decadal sea-surface temperature variability with US drought, Science Direct Quaternary International 188, 31-40.

Milly, P. C. D., J. Betancourt, M. Falkenmark, R. M. Hirsch, Z. W. Kundzewicz, D. P. Lettenmaier, R. J. Stouffer (2008), Stationarity is dead: wither water management?, Science 319, 573-574.

Rahmstorf, S. (2002), Ocean circulation and climate during the past 120,000 years, Nature, 419, 207-214

University of Washington (2009), The Pacific Decadal Oscillation (PDO), http://jisao.washington.edu/pdo/ (accessed November 30, 2009)

Willett, H. C. (1974), Recent statistical evidence in support of the predictive significance of solar-climate cycles, Monthly Weather Review, 102, 679-686.

WMO (1988), General Meteorological Standards and Recommended Practices, Technical Regulations, Volume I, WMO Publication No. 49, Geneva, Switzerland.

Wu, Z., B. Wang, J. Li, and F.-F. Jin (2009), An empirical seasonal prediction model of the east Asian summer monsoon using ENSO and NAO, J. Geophys. Res., 114, D18120, doi:10.1029/2009JD011733