

Use of generalized extreme value covariates to improve estimation of trends and return frequencies for lake levels

Shayne Paynter and Mahmood Nachabe

ABSTRACT

One of the most important tools in water management is the accurate forecast of long-term and short-term extreme values for flood and drought conditions. Traditional methods of trend detection are not suited for hydrologic systems while traditional methods of predicting extreme frequencies may be highly inaccurate in lakes. Traditional frequency estimates assume independence from trend or initial stage. However, due to autocorrelation of lake levels, initial stage can greatly influence the severity of an event. This research utilizes the generalized extreme value (GEV) distribution with time and starting stage covariates to more accurately identify trend direction and magnitude and provide improved predictions of flood and drought stages. Traditional methods of predicting flood or drought stages significantly overpredict or underpredict stages depending on the initial stage. Prediction differences can exceed one meter, a substantial amount in regions with flat topography; these differences could result in significant alterations in evacuation plans or other management decisions such as how much lake water to release in preparation for an approaching hurricane, appropriate lake levels to maintain, minimum structure floor elevations and more accurate forecasting of future water supply or impacts to tourism. The methods utilized in this research can be applied globally.

Key words | drought, flood frequency, generalized extreme value, lake stage, return period, trend

Shayne Paynter (corresponding author)
Mahmood Nachabe
Department of Civil and Environmental
Engineering,
University of South Florida,
Tampa, Florida 33620,
USA
E-mail: pshayne21@gmail.com

INTRODUCTION

One of the most important tools in effective water management is the accurate forecast of both long-term and short-term extreme values for both flood and drought conditions. High water stages associated with flood can cause extensive erosion or property damage while low stages associated with drought affect wildlife, ecology, recreation, and water supply. Accurately identifying trends in lake levels can affect long-term decision making such as forecasting water supply, while improving the prediction of near-term frequency return periods can affect short-term planning such as the determination of evacuation zones in the face of an approaching hurricane or giving resource managers adequate tools to set annual lake levels. Changes

in the general trends of lake, stream and other surface water bodies have been observed in many parts of the world. These trends may be due to factors such as watershed urbanization, water supply pumping and morphological changes to the water body itself or climatic changes. Traditional methods of trend detection, such as ordinary least squares (OLS) or the Mann–Kendall test, are not aptly suited for hydrologic systems since these systems often exhibit time scale issues, non-normal distributions, seasonality, autocorrelation, inconsistent data collection, missing data other complications that render these traditional methods unreliable (Hirsch *et al.* 1982; Katz *et al.* 2002). Zhang *et al.* (2004) utilized Monte Carlo simulations to

compare OLS, the nonparametric Kendall test, and allowing the parameters of the GEV distribution to vary with time. According to the study, while the nonparametric test is more effective at identifying trends than OLS, allowing a GEV parameter covariate significantly outperforms both OLS and the Kendall test in hydrologic trend detection. We investigated lake level trends in both flood and drought in this research utilizing the GEV distribution with a time parameter covariate.

In a similar fashion to trends, traditional methods of predicting extreme flood and drought frequencies, such as distribution fitting without parameter covariates, may be highly inaccurate in lake-type systems, especially in the short-term. In the case of lakes, traditional frequency return estimates assume extremes are independent of trend or starting lake stages. However, due to the significant autocorrelation of lake levels, the initial stage can have a significant influence on the severity of a given event. The GEV distribution has been widely applied to recent hydrologic studies (Morrison & Smith 2002; Nadarajah & Shiau 2005; Garcia *et al.* 2007). Several studies have analyzed the relation between initial stages, antecedent conditions and flood return periods in various hydrologic systems. Buchberger (1995) developed near-term flood risk estimates for Lake Erie, United States, based on an autoregressive time series model and the joint occurrence of a normally distributed storm surge and found that conventional frequency analysis underestimates flood risk when starting lake stages are high and overestimates flood risk when starting lake stages are low. Several studies have developed flood return periods dependent upon thresholds of evaporation, rainfall frequency, catchment response time, field capacity storage, catchment storage capacity or lake storage (Struthers & Sivaplan 2007; Kusumastuti *et al.* 2007, 2008). Kusumastuti (2008) found that the antecedent lake storage was the dominant control on flood frequency and magnitude. We modeled lake flood and drought stages with the GEV distribution utilizing covariates of starting lake stage and time. If the addition of time or lake stage covariates offered a significant improvement of the fit, frequency return period curves were developed for these cases. Lakes studied are located in Florida, United States, and have at least 50 years of data that are not significantly anthropogenically altered.

Trend identification in lake levels utilizing the GEV distribution as well as the development of variable return periods based on starting lake stages are a practical application of GEV distribution theory that has not yet been applied to lakes. More accurate flood and drought frequencies based on starting water level and trend will be of significant use in water resource management in terms of hurricane evacuation decisions, lake management decisions including letting an adequate amount of water out of a lake to minimize flooding impacts from an approaching hurricane or tropical storm, development of appropriate average water levels to maintain throughout the year based upon return curves that can be adjusted to the average water levels selected and preparation for increases or decreases in future flooding or drought.

The objectives of this research in regards to lake levels were to 1) accurately identify the direction and magnitude of trends in flood and drought stages and 2) provide more accurate predictions of both long-term and short-term flood and drought stage return frequencies utilizing GEV with time and starting stage covariates.

MATERIALS AND METHODS

Lake information and data

We selected lakes with at least 50 years of data across the southwestern portion of Florida that were mostly anthropogenically unaltered, i.e. from significant dredging, placement of berms, pumping, installation of major control structures, etc. in such a way that would significantly change the time series signature and, hence, the underlying distribution. Given the degree of urbanization across Florida, it is not possible to find completely unaltered lakes with sufficient data. However, four lakes, including Lake Arbuckle, Lake Carroll, Lake Trafford and Lake Weohyakapka (Figure 1) that are relatively unaltered were utilized.

GEV distribution

We identified trends in lake levels and return level frequencies utilizing extreme value models. The main

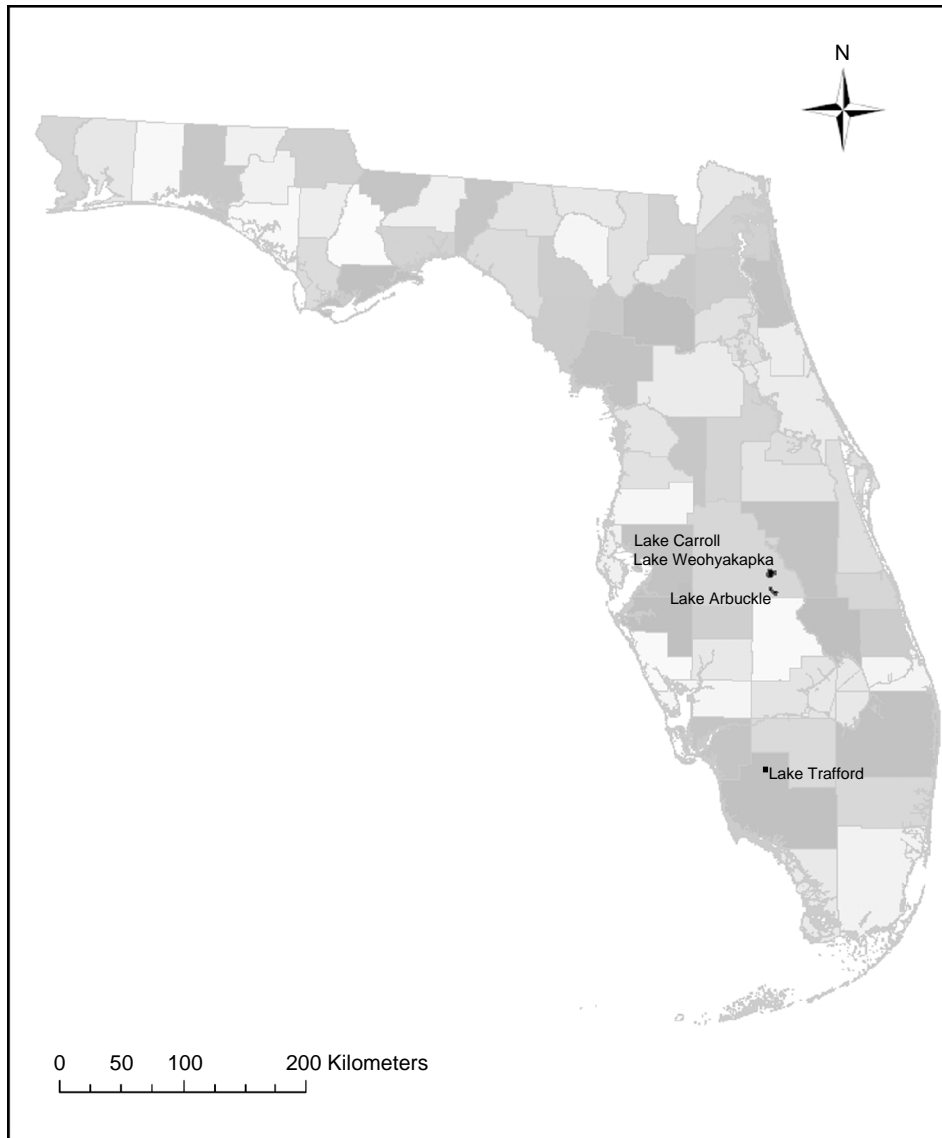


Figure 1 | Location map of study lakes.

variables modeled were the annual maximum and minimum lake levels, the flood and drought stages. In order to analyze any trends, distribution parameters were allowed to vary with time. Because lake levels exhibit substantial autocorrelation, it is surmised that annual starting lake levels have a significant impact on the distribution of annual extremes; therefore, the GEV parameters also were allowed to vary with initial stage. The starting lake stage was taken as the water level on

January 1st of any given year. We compared the time and starting lake stage covariate models to the original distribution model to determine if a statistically significant better fit was achieved. If covariates do significantly improve the fit, the distribution itself is potentially changing as these covariates change. Changing distribution parameters with time or starting stage allows for the distribution to be non-stationary and also gives an estimate on the rate of change.

The GEV is the generalized form of three commonly applied extreme value distributions: the Gumbel, the Frechet and the Weibull. The GEV is applicable to variables of block maxima, where the blocks are equal divisions of time. The GEV cumulative distribution function is given by:

$$F(x) = \exp\left\{-\left[1 + \xi\left(\frac{x - \mu}{\sigma}\right)\right]^{-1/\xi}\right\} \quad (1)$$

where x is the random variable, μ is the location parameter, σ is the scale parameter and ξ is the shape parameter and $1 + \xi(x - \mu)/\sigma > 0$. It readily follows that the sub-distributions are:

$$\text{Gumbel: } F(x) = \exp\left\{-\exp\left[-\left(\frac{x - \mu}{\sigma}\right)\right]\right\}, -\infty < x < \infty \quad (2)$$

$$\text{Frechet: } F(x) = \begin{cases} 0 & x \leq \mu \\ \exp\left\{-\left(\frac{x - \mu}{\sigma}\right)^{-1/\xi}\right\} & x > \mu \end{cases} \quad (3)$$

$$\text{Weibull: } F(x) = \begin{cases} \exp\left\{-\left(\frac{x - \mu}{\sigma}\right)^{-1/\xi}\right\} & x < \mu \\ 1 & x \geq \mu \end{cases} \quad (4)$$

GEV distribution parameters are determined using maximum likelihood estimation. The log-likelihood function, for $\xi \neq 0$, is given by (Coles 2004):

$$l(\mu, \sigma, \xi) = -m \log \sigma - (1 + 1/\xi) \sum_{i=1}^m \log \left[1 + \xi \left(\frac{x_i - \mu}{\sigma} \right) \right] - \sum_{i=1}^m \left[1 + \xi \left(\frac{x_i - \mu}{\sigma} \right) \right]^{-1/\xi} \quad (5)$$

given that $1 + \xi \left(\frac{x_i - \mu}{\sigma} \right) > 0$ for $i = 1, \dots, m$

The log-likelihood for the GEV distribution with parameters that are a function of time t or starting lake stage s is given by (Coles 2004):

$$l(\mu, \sigma, \xi) = - \sum_{i=1}^m \left\{ \log \sigma(t, s) + (1 + 1/\xi(t, s)) \times \log \left[1 + \xi(t, s) \left(\frac{x_{t,s} - \mu(t, s)}{\sigma(t, s)} \right) \right] + \left[1 + \xi(t, s) \left(\frac{x_{t,s} - \mu(t, s)}{\sigma(t, s)} \right) \right]^{-1/\xi(t, s)} \right\} \quad (6)$$

given that $1 + \xi(t, s) \left(\frac{x_{t,s} - \mu(t, s)}{\sigma(t, s)} \right) > 0$
for all $t = 1, \dots, m$

For purposes of this research, model 1 is the GEV distribution with parameters μ , σ and ξ held constant. We estimated the distribution parameters for model 1 for each lake and evaluated the goodness-of-fit with the Kolmogorov–Smirnov test statistic at the 95-percent significance level. For model 2, the location parameter of model 1 was allowed to vary with time or starting stage or both to investigate the presence of trends and determine if model 1 could be improved. It should be noted that covariates could be assigned to any of the three GEV parameters, however the location parameter is generally most sensitive to nonstationarities and given the difficulty in estimating the shape parameter, it is impractical to model this parameter as unstationary (Coles 2004). Therefore, the location parameter was the only parameter explicitly modeled with covariates, however the scale parameter was varied for comparison purposes. Model 2 is therefore a submodel of model 1 with,

$$\mu = a + by \quad (7)$$

where y is either the time in years or the starting lake stage and a and b are constants. Model 3 is a submodel of model 2 with,

$$\mu = c + dt + es \quad (8)$$

where t is the time in years, s is the starting lake stage and c , d and e are constants. Once parameters were estimated for all three cases, the models were compared to determine if the time and/or starting lake stage covariate give a statistically significant better fit. In order to test one model against another, the likelihood ratio test was utilized. If l_1 and l_2 represent the maximized log-likelihoods of the models to be compared, then a deviance statistic

is given by:

$$D = 2\{l_2 - l_1\} \quad (9)$$

Assuming a chi-square distribution, a quantile, c_{α} , at significance α can be determined and if $D > c_{\alpha}$, the submodel explains significantly more of the variation in the data (Coles 2004). Model 2 will be compared to model 1 while model 3 will be compared to both model 1 and model 2. In cases where a model with parameter covariates demonstrated a significantly better fit, fits were further investigated by examining standard quantile plots for visual confirmation of the fit improvement. However, because models 2 and 3 are non-stationary and parameters are varying at each observation, the random variable X should be transformed to a new variable Z for the quantile plot. A transform to the standard Gumbel distribution is given by (Coles 2004):

$$Z_t = \frac{1}{\xi(t, s)} \log \left\{ 1 + \xi(t, s) \left(\frac{X_t - \mu(t, s)}{\sigma(t, s)} \right) \right\} \quad (10)$$

Quantile-quantile plots were developed for these transformed standardized variables. If models 2 or 3 demonstrate an improved fit, it means estimated frequency return periods are changing with time or starting lake stage. Although the maximum likelihood ratio test is given more weight than the quantile-quantile plots, the test compares the fit of all actual data points to the model and gives even weight to all frequency events. Because low frequency events are of main interest, quantile-quantile plots were utilized to focus on the fit in the extreme end of the distribution. If an adequate fit in this region was not confirmed via the plots, the simplest model that adequately predicted extremes was selected. Return level plots for the most appropriate model for both flood and drought

were developed at each lake. Estimates of quantiles for the return level plots are given by:

$$q_{x,p} = \begin{cases} \mu(t, s) + \frac{\sigma}{\xi} [(-\ln(1-p))^{-\xi}], & \gamma \neq 0 \\ \mu(t, s) - \sigma \ln(-\ln(1-p)), & \gamma = 0 \end{cases} \quad (11)$$

where q is the quantile estimate for lake stage x at frequency p (Beirlant *et al.* 2004).

RESULTS AND DISCUSSION

A summary of the data utilized is given in Table 1. Specifically, the number of years of record, maximum, minimum and average stages and variance are provided. Plots of the maximum, minimum and starting stage for Lakes Carroll and Weohyakapka are provided in Figures 2 and 3. From the table, the standard deviation in lake levels is consistently near 0.3 m. The average difference between the maximum and minimum for the lakes analyzed is 2.13 m. Given the flat topography of west-central Florida and other similar regions, relatively small differences in water level fluctuations can inundate large areas and impact structures that are routinely set as low as 0.3 m above expected high water marks. From inspection of the figures, it appears likely that annual starting stage is correlated with both annual maximum flood and minimum drought stages as the starting stage approximately parallels both the flood and drought stages. The fits of the lake stages for flood and drought are given in Tables 2 and 3, respectively, for all GEV models. The Kolmogorov–Smirnov values were well within the 95-percent test statistic for the no-covariate fits, indicating the fits are acceptable.

Table 1 | Lake data summary

Lake	Period of record	Average (m) (NGVD [*])	Maximum (m) (NGVD [*])	Minimum (m) (NGVD [*])	Standard deviation (m)
Arbuckle	1942–2008	16.35	17.79	15.59	0.36
Carroll	1946–2003	10.76	12.10	9.41	0.34
Trafford	1941–2007	5.98	6.95	4.85	0.29
Weohyakapka	1958–2008	18.64	19.46	17.95	0.24

*National Geodetic Vertical Datum.

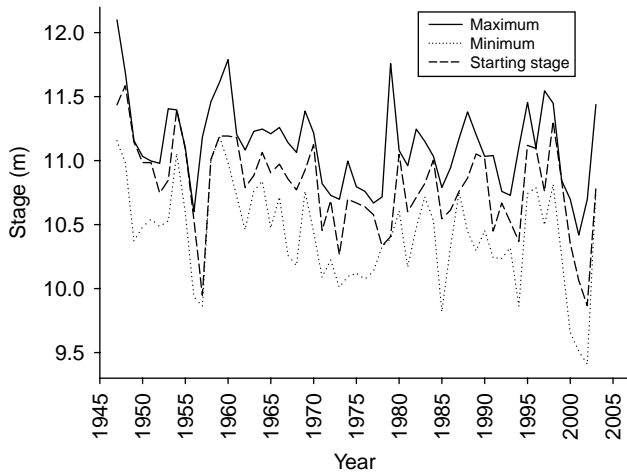


Figure 2 | Lake Carroll stage data.

Trend analysis

In regards to modeling both lake flood and drought stages with the GEV distribution and a time covariate, only Lake Carroll demonstrated a statistically significant improvement in the model 2 fit over the model 1 fit with the GEV distribution alone. Lake Arbuckle exhibited a trend in annual flood stages but not drought stages. For Lake Carroll, the model 2 location parameter for flood stages, which yields an estimate of the relation between lake stage and time t in years, indexed as 1,2,3... etc. years into the future, is given by:

$$11.176 - 0.006t$$

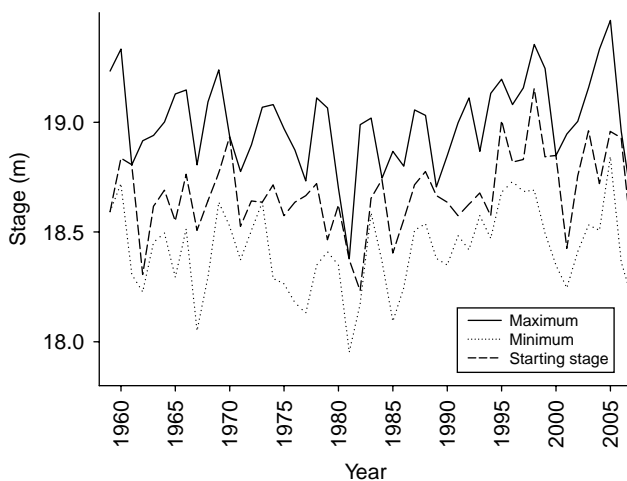


Figure 3 | Lake Weohyakapka stage data.

And the model 2 location parameter for drought stages is given by:

$$10.561 - 0.008t$$

Although the maximum likelihood ratio for both trends is substantially larger than the 95-percent confidence limit threshold, the actual change in flood or drought stage is relatively small, 0.006 m and 0.008 m of decrease per year that the trend is extended into the future. This slight trend is visually confirmed in Figure 2. For Lake Arbuckle, the model 2 location parameter for flood stages is given by:

$$17.085 - 0.007t$$

The trend is again downward and of similar order, a decrease of 0.007 m per year. The lakes studied are relatively unaltered in regards to excessive pumping, dredging, management or other mechanisms that may induce dramatic trends. Furthermore, Paynter & Nachabe (2009) determined that the rainfall patterns in the southwest Florida region do not exhibit significant trends that would correlate to changes in lake levels. Lakes Arbuckle, Trafford and Weohyakapka are fairly undeveloped when compared to Lake Carroll, which is highly urbanized. Although many lakes in Florida have demonstrated significant trends due to pumping or anthropogenic change, it appears lakes left in a fairly natural state such as the four studied for this research exhibit slight but statistically significant trends in the case of Lakes Carroll and Arbuckle or no trends in the cases of Lakes Trafford and Weohyakapka.

Starting stage analysis

Flood return period

According to the maximum likelihood ratios, model 2, with a starting stage covariate, is most appropriate for Lakes Carroll, Trafford and Weohyakapka while model 3 is most appropriate for Lake Arbuckle. It should be noted that the magnitude of the likelihood ratio is proportional to the degree of improvement of the fit; the model 2 ratios are generally high. Only Lake Arbuckle demonstrated a statistically significant improvement in fit when covariates for both time and starting stage are included. However,

Table 2 | GEV flood parameter summary

Lake	Location μ	μ -Time covariate*	μ -Starting stage covariate*	Scale σ	Shape ξ	Likelihood ratio [†]
No covariate (model 1)						
Arbuckle	16.831	N/A	N/A	0.347	-0.286	
Carroll	10.984	N/A	N/A	0.300	-0.147	
Trafford	6.285	N/A	N/A	0.193	-0.191	
Weohyakapka	18.928	N/A	N/A	0.214	-0.335	
Time covariate (model 2)						
Arbuckle	17.085	-0.007 (0.002)	N/A	0.336	-0.385	Model 2/model 1 10.301
Carroll	11.176	-0.006 (0.002)	N/A	0.282	-0.149	6.567
Trafford	6.301	-0.000 (0.001)	N/A	0.193	-0.192	0.090
Weohyakapka	18.876	0.002 (0.000)	N/A	0.216	-0.386	1.204
Starting stage covariate (model 2)						
Arbuckle	8.954	N/A	0.484 (0.191)	0.343	-0.387	Model 2/model 1 8.203
Carroll	3.299	N/A	0.712 (0.074)	0.165	0.194	45.895
Trafford	4.015	N/A	0.376 (0.198)	0.179	-0.115	4.457
Weohyakapka	5.691	N/A	0.710 (0.015)	0.192	-0.464	16.431
Time and starting stage covariates (model 3)						
Arbuckle	13.178	-0.007 (0.001)	0.240 (0.019)	0.335	-0.488	Model 3/model 1: model 3/model 2 16.626/8.422
Carroll	3.470	-0.000 (0.001)	0.697 (0.078)	0.169	0.169	45.782/0.113
Trafford	3.858	-0.001 (0.001)	0.407 (0.204)	0.177	-0.109	4.898/0.441
Weohyakapka	6.341	-0.000 (0.000)	0.676 (0.000)	0.190	-0.449	17.081/0.649

*Standard Errors of slope parameters are noted in parenthesis.

[†]At the 95-percent confidence interval, a maximum likelihood ratio of greater than 3.842 for model 2/model 1 or model 3/model 2 and 5.992 for model 3/model 1 indicates a significantly better fit. The model 2/model 1 and model 3/model 1 ratios were also compared to determine if the additional degree of freedom improves the fit. Selected models are bolded.

as the trend component of model 3 is negligible, the simpler model 2 was selected. The Lake Arbuckle (Figure 4) and Lake Carroll (Figure 5) quantile-quantile plots demonstrate an adequate fit for model 2. Quantile-quantile plots for

Lakes Trafford and Weohyakapka also demonstrated adequate model 2 fits. With the exception of Lake Carroll, in most of the quantile-quantile plots for flood, the fit breaks down at the extreme end for all models. This is partly due to

Table 3 | GEV drought parameter summary

Lake	Location μ	μ -Time covariate*	μ -Starting stage covariate*	Scale σ	Shape ξ	Likelihood ratio [†]
No covariate (model 1)						
Arbuckle	15.844	N/A	N/A	0.182	-0.150	
Carroll	10.304	N/A	N/A	0.409	-0.395	
Trafford	5.554	N/A	N/A	0.301	-0.504	
Weohyakapka	18.354	N/A	N/A	0.194	-0.326	
Time covariate (model 2)						
Arbuckle	15.915	-0.002 (0.001)	N/A	0.181	-0.169	Model 2/model 1 2.484
Carroll	10.561	-0.008 (0.002)	N/A	0.380	-0.455	11.005
Trafford	5.498	0.001 (0.002)	N/A	0.295	-0.465	0.549
Weohyakapka	18.293	0.0026 (0.001)	N/A	0.193	-0.363	2.023
Starting stage covariate (model 2)						
Arbuckle	10.414	N/A	0.335 (0.000)	0.154	-0.383	Model 2/model 1 7.536
Carroll	0.117	N/A	0.947 (0.076)	0.200	-0.264	71.923
Trafford	-0.804	N/A	1.055 (0.140)	0.200	-0.323	39.937
Weohyakapka	3.855	N/A	0.778 (0.007)	0.149	-0.445	32.044
Time and starting stage covariates (model 3)						
Arbuckle	9.441	-0.001 (0.000)	0.396 (0.334)	0.166	-0.276	Model 3/model 1: model 3/model 2 21.753/14.217
Carroll	0.439	-0.001 (0.002)	0.921 (0.085)	0.198	-0.259	72.413/0.489
Trafford	-1.519	0.000 (0.000)	1.177 (0.096)	0.177	-0.329	34.432/5.505
Weohyakapka	15.839	0.002 (0.000)	0.134 (0.000)	0.176	-0.430	8.160/23.884

*Standard errors of parameters are noted in parenthesis.

[†]At the 95-percent confidence interval, a maximum likelihood ratio of greater than 3.842 for model 2/model 1 or model 3/model 2 and 5.992 for model 3/model 1 indicates a significantly better fit. The model 2/model 1 and model 3/model 1 ratios were also compared to determine if the additional degree of freedom improves the fit. Selected models are bolded.

these points representing hurricanes or tropical storms that are not part of the same distribution as normal rainfall events and partly due to extrapolating extreme events with 50 years of data. After evaluating both the maximum

likelihood ratios and the quantile-quantile plots, we selected model 2 for all four lakes in terms of flood stage. The location parameter, which yields an estimate of the relation between starting stage and flood stage, is given by

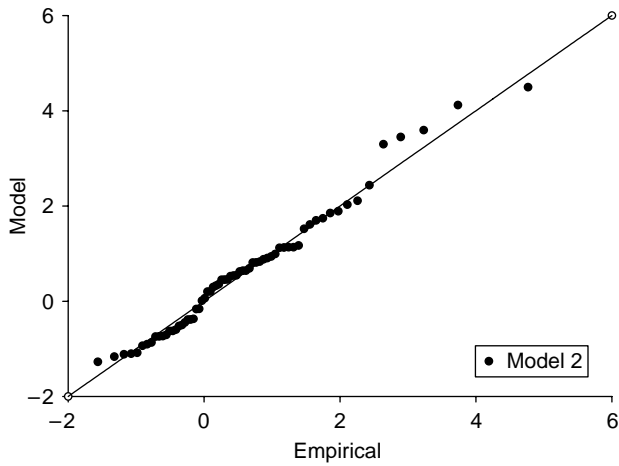


Figure 4 | Lake Arbuckle flood stage standardized residual quantiles.

the following for Lakes Arbuckle, Carroll, Trafford and Weohyakapka, respectively, for starting stage s :

$$8.954 + 0.484s$$

$$3.299 + 0.712s$$

$$4.015 + 0.376s$$

$$5.691 + 0.710s$$

For every unit change in starting stage, there is a substantial change ranging from 0.376 to 0.712 m in the flood stage for a given year, indicating a very high degree of correlation. Figures 6 and 7 give the model 2 flood return period for Lakes Arbuckle and Trafford associated with the maximum, minimum and average starting stage as well as

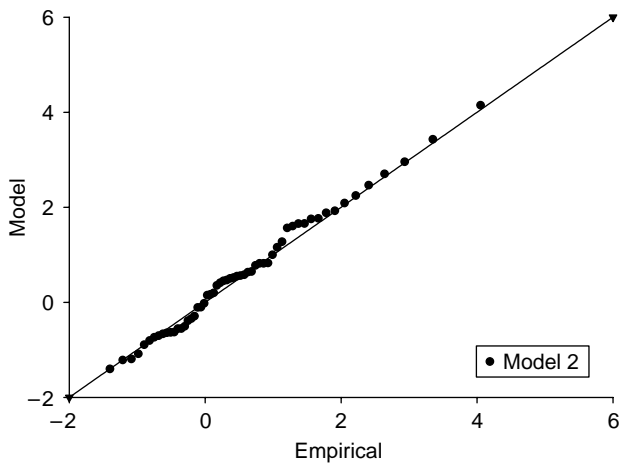


Figure 5 | Lake Carroll flood stage standardized residual quantiles.

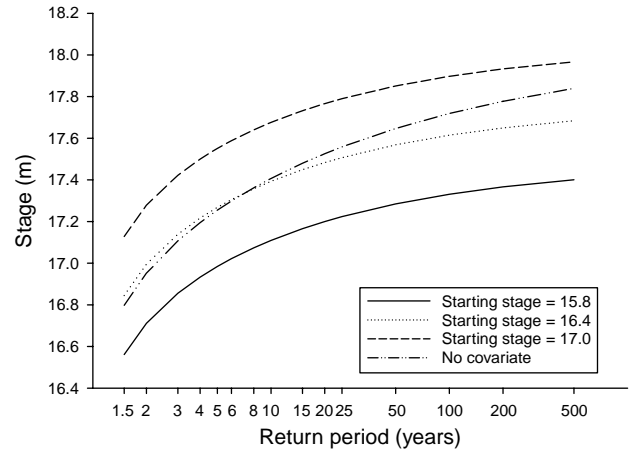


Figure 6 | Lake Arbuckle flood return frequencies with and without starting lake stage covariates.

the return period associated with no covariate. For each lake, the return period associated with the average starting stage is fairly close to the return period associated with no covariate. For Lakes Arbuckle, Carroll and Weohyakapka, there is some divergence between these two curves towards the larger return periods. At Lakes Arbuckle and Carroll, this is likely due to the fact that these lakes exhibit some trends and since the starting stage should correlate to any trends, the inclusion of the starting stage covariate improves the fit and causes divergence from the fit without a covariate. The flood return period associated with no covariate is bounded by that associated with the maximum and minimum starting stage. In years with a low starting stage, traditional frequency analysis overpredicts the

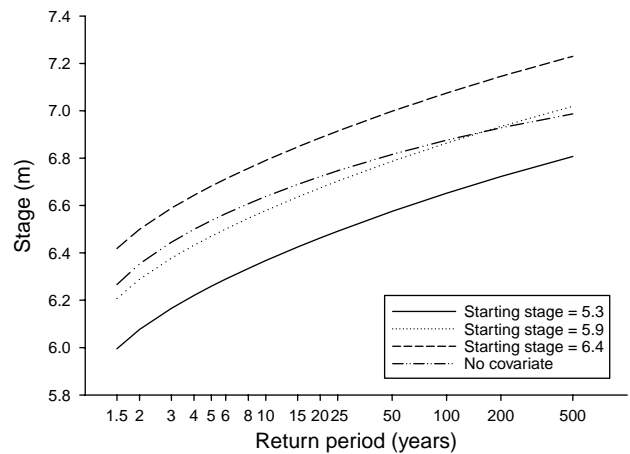


Figure 7 | Lake Trafford flood return frequencies with and without starting lake stage covariates.

100-year flood by 108.3, 129.4, 75.9 and 179.2 percent of standard deviation for Lakes Arbuckle, Carroll, Trafford and Weohyakapka, respectively. In years with a high starting stage, traditional frequency analysis underpredicts the 100-year flood by 50, 232.4, 69.0, and 91.7 percent of standard deviation for the same lakes. As such there is a 0.57 m, 1.22 m, 0.42 m and 0.65 m difference, respectively, between the 100-year return period stage for the maximum and minimum starting lake stage covariate. Given the flat topography in Florida and other similar regions, a difference of as much as 1.22 m can mean a substantial increase in the extent of flooding and potential number of structures flooded.

Since more area is available at consistently higher elevations of a lake, it takes more runoff or baseflow volume to cause a unit rise in stage at higher lake elevations. Because of this it would be expected that in a lake left in its natural stage, return period curves would flatten out at more extreme frequencies. However, once a lake basin is urbanized, the watershed infilled with construction and management structures installed, it is difficult to consistently predict the shape of these curves in a general sense. Lakes Arbuckle, Trafford and Weohyakapka are relatively undeveloped and they demonstrate the expected flattening of the return period curves at higher frequencies. Lake Carroll is the most urbanized and it shows some steepening of the return period curves at extreme events.

Drought return period

According to the maximum likelihood ratios, model 2 (with a starting stage covariate) is most appropriate for Lakes Carroll, Trafford and Weohyakapka while model 3 is most appropriate for Lake Arbuckle. As in the flood analysis, the trend component is quite small and the simpler model 2 was deemed appropriate. Similar to the flooding case, the likelihood ratios for model 2 are quite high, indicating that model 2 explains substantially more of the variation. The quantile-quantile plots for Lakes Arbuckle, Carroll, Trafford (Figure 8) and Weohyakapka (Figure 9) indicate an adequate fit for model 2. As with the flood quantiles, there is divergence between the model and empirical data at the extremes. This is likely due to longer time-scale cycles, such as La Nina, that cause excessively dry years and are not

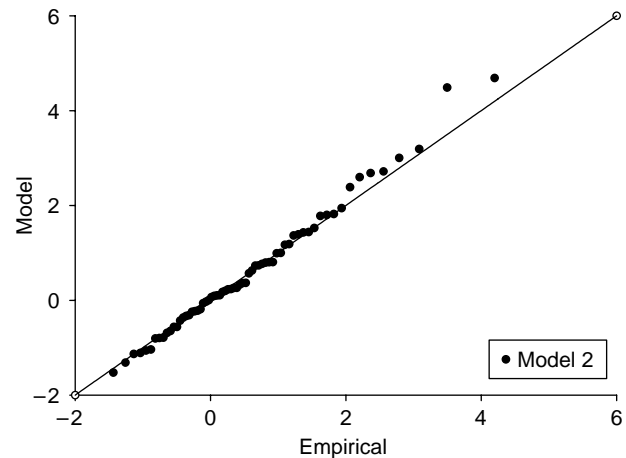


Figure 8 | Lake Trafford drought stage standardized residual quantiles.

explicitly included in the models; model 2 should capture some, but not all, of these longer cycles with the inclusion of starting stage. Some of the fit breakdown is also due to extrapolating events greater than the 50-year from 50 years of data. After evaluating both the maximum likelihood ratios and the quantile-quantile plots, we selected model 2 for all four lakes. The location parameter associated with the most appropriate model for each lake is given by the following for Lakes Arbuckle, Carroll, Trafford and Weohyakapka, respectively, for starting stage s :

$$10.414 + 0.335s$$

$$0.117 + 0.947s$$

$$-0.804 + 1.055s$$

$$3.855 + 0.778s$$

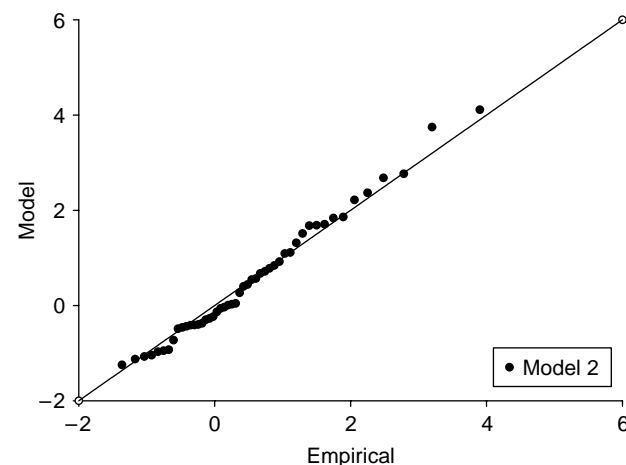


Figure 9 | Lake Weohyakapka drought stage standardized residual quantiles.

As in the flood case, for every unit change in starting stage, there is a substantial change in the drought stage for that year, in this case ranging from 0.335 m to 1.055 m. Figures 10 and 11 give the drought return period for Lakes Arbuckle and Carroll associated with the maximum, minimum and average starting stage as well as the return period associated with no covariate. Similar to the flood return period case, the return period associated with no covariate is bounded by that associated with the maximum and minimum starting stage and nearly parallels the return period associated with the average starting stage covariate. One exception is Lake Carroll where the no-covariate return period curves deviate significantly from the average starting stage covariate curves towards the extreme end. Lake Carroll was the only lake to exhibit a significant drought trend and, as in the flood case, including the starting stage as a covariate captures some of this trend and provides a better fit. In years with a low starting stage, traditional frequency analysis overpredicts the 100-year drought by 41.4, 105.9, 158.6 and 104.2 percent of standard deviation for Lakes Arbuckle, Carroll, Trafford and Weohyakapka, respectively. In years with a high starting stage, traditional frequency analysis underpredicts the 100-year drought by 66.7, 373.5, 251.7 and 191.7 percent of standard deviation for the same lakes. As such there is a 0.39 m, 1.63 m, 1.19 m and 0.72 m difference, respectively, between the 100-year return period stage for the maximum and minimum starting lake stage covariate. In similar fashion to

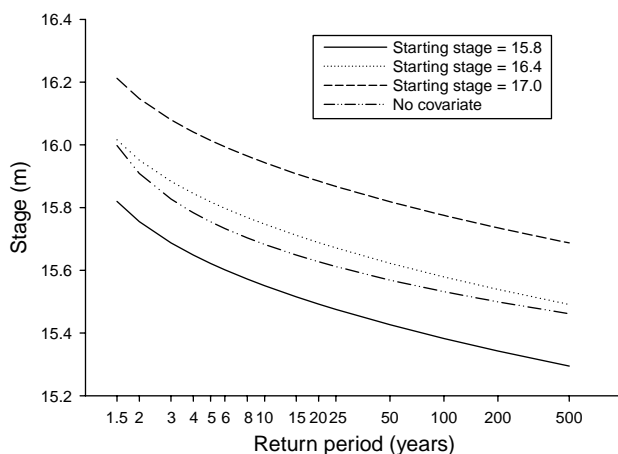


Figure 10 | Lake Arbuckle drought return frequencies with and without starting lake stage covariates.

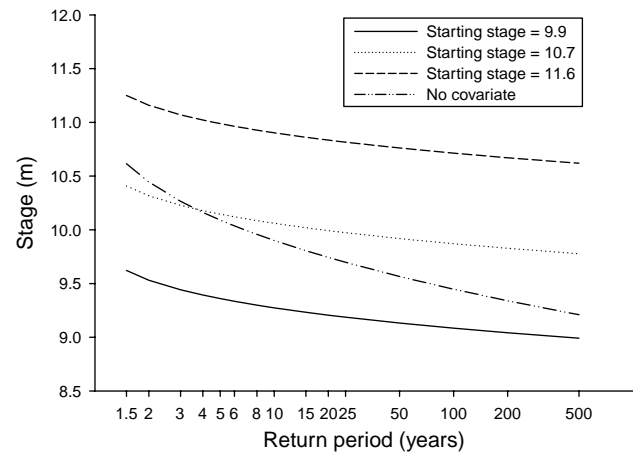


Figure 11 | Lake Carroll drought return frequencies with and without starting lake stage covariates.

flood stages, it is expected that drought return period curves would flatten at more extreme return periods since there are more water loss mechanisms at higher lake stages. At lower stages, the only method of water loss may be evapotranspiration or recharge to the ground. All four lake drought return curves follow this general pattern.

There appears to be no correlation in the difference between flood stages and drought stages for the minimum and maximum starting lake stages within each lake, i.e. a small difference in the Lake Carroll flood stage associated with the maximum and minimum starting stage does not indicate a small difference in the drought stage associated with the maximum and minimum starting stage. This is likely due to different physical dynamics operating in the flood and drought cases. Flood stages are generally controlled by some management mechanism, i.e. a weir, culvert, gate, etc. while drought stages are largely uncontrolled other than natural losses such as evapotranspiration or seepage to the ground. Furthermore, at extreme flood stages, the basin morphology may change relative to the lake at lower stages, i.e. higher stages may be flatter than at lower stages, a basin popoff to another basin may be reached or housing construction may have significantly altered the historic basin by infill.

In all cases for both flood and drought, adding covariates for both trend and starting stage offer little improvement over a starting stage alone. This is likely because any monotonic trend in time should be captured in the starting stage variable and because the trends identified were very

small. Potential scenarios for the inclusion of both time and starting stage covariates improving a fit include situations in which overall trends may not be reflected in the January 1st stage. One possibility may be seasonal trends such as an increase in summer floods due to hurricanes or tropical storms followed by periods of low rainfall whereby the annual starting stage returns to normality.

CONCLUSIONS

The lakes studied were relatively unaltered in terms of extensive pumping, dredging, filling or other measures that would significantly alter the underlying lake level distribution. All of the lakes researched evidenced either no trend or very small trends unlikely to significantly alter prediction of future flood or drought return levels. However, for all of the lakes, significant improvement in the fits was obtained with the inclusion of starting lake stage as a covariate. This is likely because any monotonic trends are captured in the starting stage itself and the trends identified were negligible. Traditional methods of estimating flood or drought stages significantly overpredict stages when starting lake stages are low and underpredict stages when starting stages are high. The difference between these predictions can be substantially more than one meter, a significant amount in urbanized watersheds in areas of the world with flat topography. Flood differences of over one meter can mean significant alterations in evacuation or other water management decisions. In addition to improving prediction of extreme events, utilizing GEV with time or starting stage covariates can provide guidance in lake management decisions in regards to how much water to release from a lake in preparation for an approaching hurricane, appropriate lake levels to maintain throughout the year or determining minimum structure flood elevations in the watershed. Although there is less that can be done from a management standpoint in regards to drought, utilizing GEV with covariates provides a more accurate estimate of expected drought return periods, which can be useful in forecasting future water supply or impacts to tourism.

The methodology employed in this research provides a means to estimate the direction and magnitude of lake trends that is robust despite the inherent difficulties in determining trends in hydrologic data. The methodology also allows for more accurate prediction of flood and drought return frequencies that can be applied to nearly any region globally.

REFERENCES

- Beirlant, J., Goegebeur, Y., Segers, J. & Teugels, J. 2004 *Statistics of Extremes*. John Wiley & Sons Limited, New York.
- Buchberger, S. 1995 Conditional frequency analysis of autocorrelated lake levels. *J. Water Resour. Plan. Manage.* **121**, 158–170.
- Coles, S. 2004 *An Introduction to Statistical Modeling of Extreme Values*. Springer-Verlag London Limited, London.
- Garcia, J., Gallego, M., Serrano, A. & Vaquero, J. 2007 Trend in block-seasonal extreme rainfall over the Iberian peninsula in the second half of the twentieth century. *J. Clim.* **20**, 113–130.
- Hirsch, R., Slack, J. & Smith, R. 1982 Techniques of trend analysis for monthly water quality data. *Water Resour. Res.* **18**, 107–121.
- Katz, R., Parlange, M. & Naveau, P. 2002 Statistics of extremes in hydrology. *Adv. Water Resour.* **25**, 1287–1304.
- Kusumastuti, D., Struthers, I., Sivapalan, M. & Reynolds, D. 2007 Threshold effects in catchment storm response and the occurrence and magnitude of flood events: implications for flood frequency. *Hydrol. Earth Syst. Sci.* **11**, 1515–1528.
- Kusumastuti, D. I., Sivapalan, M., Struthers, I., Reynolds, D. A., Murray, K. & Turlach, B. A. 2008 Thresholds in the storm response of a catchment-lake system and the occurrence and magnitude of lake overflows: implications for flood frequency. *Water Resour. Res.* **44**, W02438.
- Morrison, J. & Smith, J. 2002 Stochastic modeling of flood peaks using the generalized extreme value distribution. *Water Resour. Res.* **38**, 41-1–41-12.
- Nadarajah, S. & Shiau, J. 2005 Analysis of extreme flood events for the Pachang River, Taiwan. *Water Resour. Manage.* **19**, 363–374.
- Paynter, S. & Nachabe, M. 2009 Regional scale spatio-temporal consistency of precipitation variables related to water resource management and planning. *Meteorol. Appl.* **16**, 413–423.
- Struthers, I. & Sivapalan, M. 2007 A conceptual investigation of process control upon flood frequency: role of thresholds. *Hydrol. Earth Syst. Sci.* **11**, 1405–1416.
- Zhang, X., Zwiers, F. & Li, G. 2004 Monte Carlo experiments on the detection of trends in extreme values. *J. Clim.* **17**, 1945–1952.

First received 26 October 2008; accepted in revised form 1 September 2009. Available online 31 March 2010