Water Resour Manage
DOI 10.1007/s11269-014-0896-4

Global Storage-Reliability-Yield Relationships for Water Supply Reservoirs

Faith W. Kuria · Richard M. Vogel

Received: 17 January 2014 / Accepted: 30 November 2014

Abstract Storage-Reliability-Yield (SRY) relationships are used to determine the reservoir storage capacity for delivery of a specified yield with a given reliability, or to compute the yield and/or reliability of an existing reservoir system. Several studies have developed generalized SRY relations using synthetic inflows arising from a variety of theoretical streamflow models. Fewer studies have used actual streamflow datasets to develop generalized SRY relationships and most of those studies were for small geographic regions. This study uses a global dataset of monthly streamflows combined with robust regression methods to develop improved generalized SRY models suitable for use anywhere in the world. Comparisons are provided between the models developed here and other studies documenting a number of innovations over previous relationships. In cross validation experiments our global reservoir yield model exhibited extremely high goodness-of-fit with values of Nash Sutcliffe Efficiency and adjusted R^2 values both always in excess of 0.99 and negligible bias. The resulting SRY model should prove useful in screening studies which seek to evaluate the benefits of constructing reservoirs for surface water supply.

Keywords Surface water · Infrastructure · Dams · Municipal · Commercial · Hydropower · Irrigation · Demand · Water use

1 Introduction

In the design of water supply reservoirs, the Storage-Reliability-Yield (SRY) relationship is the tool that has traditionally been used to determine the reservoir storage capacity required for delivery of a water supply yield with a specified reliability or yield that can be supplied from a reservoir with known storage capacity. Here reliability is the steady-state time based value that indicates the probability that the reservoir yields are met. Behaviour analysis (BA) is the simulation method that has been widely used in many parts of the world to develop such a relationship (see McMahon and Adeloye 2005, pg 86). Using BA, the minimum reservoir

F. W. Kuria (✉) · R. M. Vogel
Department of Civil and Environmental Engineering, Tufts University, Medford, MA 02155, USA
e-mail: fwmwang@gmail.com

R. M. Vogel
e-mail: Richard.Vogel@tufts.edu

Published online: 11 December 2014

storage capacity required for delivery of a specified yield with a given reliability is determined by trial and error. The appropriate storage capacity S is typically one that will provide the specified yield Y at the specified reliability R. A prespecified operating policy of the reservoir is needed to determine the appropriate storage capacity. Pretto et al. (1997) showed that very long streamflow sequences of at least 1,000 years are required to obtain a stable and steady state solution to the SRY relationship for reservoirs dominated by year-to-year fluctuations in storage, termed 'over-year' systems. The classification of reservoirs into within-year and over-year reservoirs is given by Vogel et al. (1999) and Vogel and Bolognese (1995). An over-year reservoir is one which does not normally refill at the end of each year. A variant of the BA approach that has been widely used in the USA and elsewhere is the mass curve (Rippl 1883) or its automated equivalent sequent peak algorithm (SPA) introduced by Thomas and Burden (1963). The SPA assumes failure free (100 % reliable) reservoir operations over a prespecified planning horizon which is often based on a historical record of inflows. During future planning periods, inflows into the reservoir are likely to be wetter or drier than the historical record indicates, thus the actual reliability of reservoirs designed using the SPA approach remain unknown. Other limitations of this approach are highlighted in Fiering (1963, pg. 7) and Vogel and Stedinger (1987).

1.1 Literature Review on Generalized SRY Relationships

Use of simulation procedures to derive the steady-state SRY relationship is computationally intensive because a stochastic streamflow model and a reservoir simulation model must be combined and implemented repeatedly, using thousands of Monte-Carlo experiments. Attempts have been made to develop generalized SRY relations that can be used to mimic the results of such Monte-Carlo experiments based on the use of a stochastic streamflow models combined with a simple reservoir simulation model. Depending on the type of streamflow traces considered in the analysis, two types of generalized SRY relations exist: for reservoirs fed by theoretical inflows and for reservoirs fed by actual streamflow traces. Examples of 'theoretical' generalized SRY models include the tabulations and graphical relations developed by Pegram (1980) for reservoirs fed by lognormal flows, Bayazit and Bulu (1991) for over year reservoirs fed by normal and lognormal autoregressive and autoregressive moving average (AR(1)) and (AR(1,1)), and by Bayazit and Önöz (2000) for over year reservoirs fed by normal flows, Bayazit (1982) for reservoirs fed by first-order autoregressive (AR(1)) model. Silva and Portela (2012) developed SRY relations using synthetic streamflows based on the hydrological characteristics of 54 rivers located in various parts of Portugal. There are many examples of studies which sought to generalize the SRY relationship for a particular river, however such studies are too numerous to summarize here. A simple model known as the Dincer model (see McMahon and Mein 1986; McMahon and Adeloye 2005; and McMahon et al. 2007a) is a simple analytical SRY relation based on the water balance of inputs and outputs of a storage reservoir for overyear reservoirs fed by normal flows. Gould (1964) extended for the Dincer model for Gamma flows. McMahon et al. (2007a) provide other modifications of the Dincer and Gould method for autocorrelated lognormal inflows.

Others have used multivariate regression to develop analytical generalized SRY relations using synthetic streamflows. For example Vogel and Stedinger (1987) developed relations for over year reservoirs fed by lag one autoregressive lognormal (AR(1)LN) streamflows with Cv <0.5. Similarly Phien (1993) developed relations for overyear reservoirs fed by Gamma autoregressive (GAR) streamflows. These tabular, graphical and analytical 'theoretical' generalized SRY models and others of this type are limited for use with over-year systems where the stochastic structure of annual streamflows is well approximated by the particular theoretical model of inflows assumed in the analysis. Using actual streamflows, McMahon et al. (2007b) showed that the Vogel and Stedinger (1987) SRY relations performed satisfactorily when applied to streamflows with similar statistical characteristics as were considered in the development of their model though biases were observed when such models were used outside the range of conditions considered in the development of the model.

Actual streamflow observations have also been used to develop generalized SRY relationships. For example Adeloye et al. (2003) used monthly streamflow data for 12 international rivers to develop generalized SRY relations for within year and total storage capacity requirements. Total storage is the sum of within-year and over-year storage requirements. Adeloye (2009a, b) used these same 12 international rivers together with three more rivers in Italy to develop regression equations for over year storage capacity requirements. Portela and Quintela (2012) used historical streamflows to develop generalized SRY for rivers in Portugal. The application of the models by Adeloye et al. (2003), Adeloye (2009a, b), and Silva and Portela (2012) is limited to the geographical areas that were considered in the analysis. McMahon et al. (2007c) used a global database of 729 unregulated rivers all of which had at least 25 years of monthly streamflow data to develop generalized SRY models based on storage estimates derived from both SPA and BA. Their resulting multivariate regression equations predict the storage capacity, S, as a function of numerous independent variables including: the mean μ, standard deviation σ, and skew γ of the annual inflows and parameters describing the system yield Y and reliability R. Their empirical regression equation for storage capacity based on BA summarized in Table 4 of their paper is:

$$S = 1.932\ \mu^{-3.254}\sigma^{1.599}\gamma^{-0.074}Y^{2.67}{Z_R}^{1.445} \quad \left(R^2 = 0.899\right) \tag{1}$$

where S is storage capacity, μ, σ and γ are the mean, standard deviation and skewness coefficient of the annual inflows, Y is the yield, all of which have units in millions of m^3. Here Z_R is the standardized normal variate with R equal to the reliability (For example, a system with reliability R=0.95 corresponds to a value of Z_R=1.645). We employ the same global database of rivers here as was used by McMahon et al. (2007c).

1.2 Previous Applications of Generalized SRY Relationships

Such generalized theoretical SRY relations are useful for improving our understanding of the general behavior of reservoirs, because they can represent the behavior of an extremely wide class of reservoir systems. Here we describe numerous examples of the how such SRY relationships have proved useful in previous studies: Vogel and Stedinger (1988) used generalized SRY relations developed by Vogel and Stedinger (1987) to illustrate the value of stochastic streamflow models in the design of water supply reservoirs as well as to document the variability of estimates of reservoir storage capacity based on short streamflow records. Since they employed generalized SRY relations their results apply to an extremely wide class of reservoir behavior, unlike most reservoir studies which sought only to understand the behavior of a particular reservoir system. Vogel et al. (1999) used generalized SRY relations to explore the behavior of thousands of actual storage reservoir systems across the continental United States and Vogel et al. (1995) performed a similar analysis for a few complex multiple reservoir systems in the Northeastern USA. Similarly, Vogel et al. (1997), Lane et al. (1999) and Brown et al. (2012) used generalized SRY relationships to explore the impact of climate change on reservoir system behavior. Vogel and Bolognese (1995) and Vogel and McMahon (1996) used generalized SRY to validate analytical relationships for reliability, resilience and vulnerability indices used to understand the behaviour of over year reservoirs.

An interesting application and extension of SRY relationships was the work of Hanson and Vogel (2014) who developed analogous generalized SRY relationships for rain water harvesting systems. Kuria (2014) used generalized SRY relationships to document the variability and to estimate confidence intervals for estimates of water supply reservoir yields. We anticipate an even broader set of applications which may result from the generalized SRY relationships developed in this study, because they are based on a global dataset of actual monthly streamflow records and thus mimic actual hydrologic and associated SRY relationships across the globe.

1.3 Study Goals

Our literature review reveals that although there are several studies that have developed generalized SRY relations, there is only one study by McMahon et al. (2007c) that has developed an analytical empirical model (Eq. 1) based on a global dataset of actual streamflow observations that can be used as a preliminary tool to design new reservoirs or evaluate the reliability and/or yield of existing reservoirs anywhere in the world. Since the models developed by McMahon et al. (2007c) are based on actual monthly streamflow series, they account for the actual complex stochastic structure of the monthly streamflows considered, and they are applicable for both within year and over year reservoirs. The relations also provide preliminary estimates of existing system reliability, something which is often unavailable when only an SPA algorithm was applied in the design of the reservoir system. Thus we conclude that Eq. (1) developed by McMahon et al. (2007c) is an extremely useful tool for both water supply engineering research and practice, yet the model is buried within Table 4 along with five other models in a paper which deals with many other findings, and thus is not likely to receive either the attention and/or scrutiny that it deserves. One goal of this study is therefore to more fully evaluate the adequacy of the SRY model given in Eq. (1). We use Iteratively Reweighted Least Squares (IRLS) method (Beaton and Tukey 1974) to develop improved models of Eq. (1) using the same database as was used by McMahon et al. (2007c). We also use influence statistics (Helsel and Hirsch 2002) in an effort to remove particular rivers which exert unrealistic influence over model coefficient estimates.

Also none of the previous studies developed models with reservoir yield as the dependent variable and few of the previous models incorporate reliability of the specified yields in the resulting model. Relationships which consider reservoir yield as the dependent variable are increasingly important as the focus in water resources development is changing from constructing new reservoirs to managing/improving the performance of existing ones especially in developed countries. Previous SRY relationships could be solved for yield, such as in Eq. (1), however such estimates of yield will not necessarily be less than the mean annual flow, thus estimates of yield could be unrealistic, particularly for systems with very large storage ratios S/μ. Thus, when one solves an SRY relationship (such as that given in Eq. 1) for Y, one can easily obtain yield ratios (Y/μ) which exceed unity. Thus a central goal of this study is to develop a global SRY model for estimation of water supply reservoir yields as a function of S, R, and the hydrologic characteristics of reservoir inflows which cannot exceed the mean annual inflow.. Reservoir yield is often treated as a constant, yet due to natural variability of streamflows and sampling errors resulting from limited streamflow records, the yields from the reservoirs are expected to be variable during the actual performance of the reservoirs. Such a multivariate equation can be very useful for determining the variability of future yields as is shown by Kuria (2014).

2 Dataset and Development of the Regression Models

We use the same global dataset that was used by McMahon et al. (2007c) based on 729 rivers each of which had at least 25 years of monthly streamflows (see their Fig. 1 for the spatial coverage and location of river gages). Required storage capacities of hypothetical reservoirs for each of the 729 unregulated rivers were determined by McMahon et al. (2007a) using BA for delivery of yield ratios (yield ratio=Y/μ) in the range of 0.3–0.8 in increments of 0.1, with monthly reliabilities of 0.9, 0.95 and 0.98. The Standard Operating Policy was used, which assumes the desired yield will be satisfied if there is sufficient water, otherwise whatever is available is supplied until the reservoir runs empty (McMahon and Adeloye 2005). A total of 12,412 estimates of reservoir storage capacity were generated and form the database employed in the development of our global SRY model.

This study employs multiple linear regression as a method for generalizing the SRY relationship using the global dataset. We begin by developing a regression model for yield using the

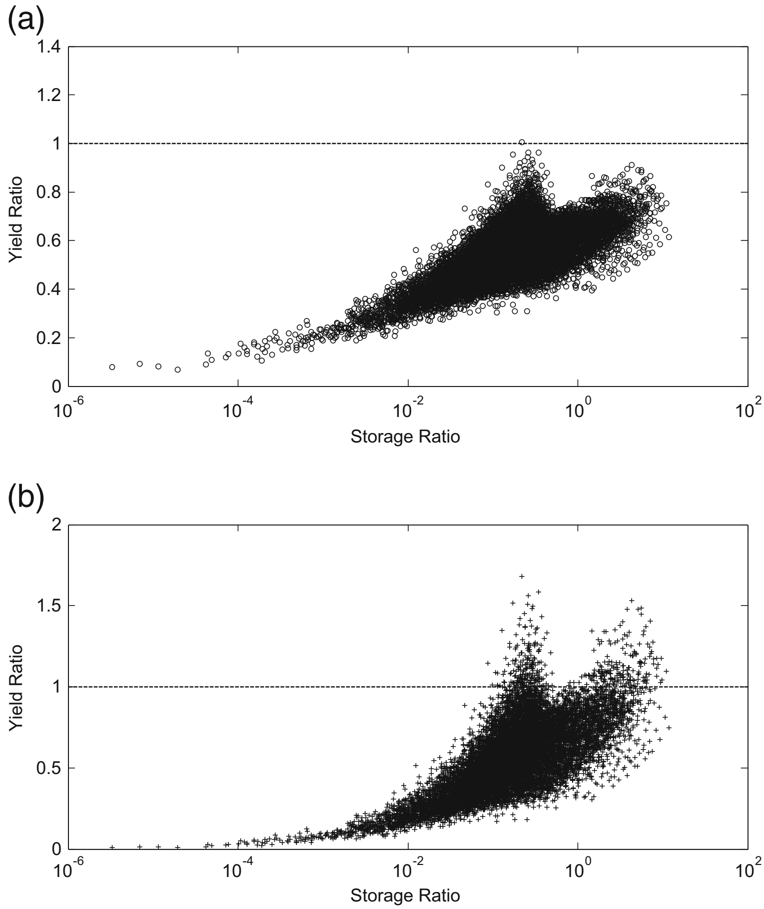

Fig. 1 **a** The range of predicted values of the Storage ratio and the Yield ratio using our yield model summarized in Eq. 2 and Table 2. The dashed line (at unity) indicates the physical limit of the yield ratio. **b** The range of predicted values of the Storage ratio and the Yield ratio by solving Eq. 1 developed by McMahon et al. (2007c) for the yield ratio. The dashed line (at unity) indicates the physical limit of the yield ratio

same predictor variables as were used in Eq. (1) i.e. mean μ, standard deviation σ, and skewness coefficient of annual streamflows, storage capacity S and standardized normal reliability Z_R. In addition we consider the autocorrelation ratio, $\phi = \frac{1+\rho}{1-\rho}$ introduced using completely independent derivations by Phatarfod (1977) and Vogel and McMahon (1996). The variable ϕ was found to be a significant explanatory variable in the SRY studies by Vogel and Stedinger (1987), Phien (1993) and others. We used stepwise regression to evaluate alternative multivariate models using the goodness-of-fit metric termed 'prediction R^2' described by Helsel and Hirsch (2002). The prediction R^2 is based on the use of 'delete one residuals' which are used to compute the prediction sum of squares (PRESS) metric. PRESS and prediction R^2 arevalidation goodness of fit metrics which which are equivalent to performing a jackknife, or leave one out model validation (Helsel and Hirsch 2002). The prediction R^2, unlike R^2, will generally not increase when adding additional explanatory variables unless their addition leads to improved model predictions under validation (leave-one out) conditions. Weighted least squares (WLS) regression analysis with the weights based on length of the available historical record was used, thus rivers with longer flow records will be weighted more heavily than those with shorter records. Normal probability plots together with the probability plot correlation coefficient (PPCC) test of normality (Heo et al. 2008) were used to evaluate normality of the model residuals. Variance Inflation Factors (VIF) were used to detect multicollinearity among the independent or explanatory variables. Traditionally VIF>10 indicates the presence of multicolinearity (Helsel and Hirsch 2002) resulting in potentially unstable model parameter estimates and misleading model selection. However Kroll and Song (2013) found that high values of VIF should not necessarily be of concern when sample sizes are large and overall model goodness-of-fit is high which is the case here. Finally K-fold cross validation experiments were implemented to evaluate the ability of the model to predict water supply reservoir yields in practice.

3 Results

In the following sections we describe the SRY models which were developed using the multiple regression methods outlined in the previous section.

3.1 Development of Generalized SRY Models for Estimating Yield

Using the global dataset of monthly streamflows to develop an SRY model with yield as the dependent variable led to the model:

$$Y = aS^b {Z_R}^c \mu^d \sigma^e \gamma^f \tag{2}$$

where all the variables have the same meaning and units as described in Eq. 1 and the model coefficients are summarized in Table 1. Table 1 illustrates very different values for the estimates of the model coefficients associated with our yield model in (2) and the implied yield model given by solving Eq. (1) for yield Y. Also included in Table 1 are t-ratios, standard error of the estimates (SEE), adjusted R squared, prediction R squared, Nash Sutcliffe efficiency (NSE) (Nash and Sutcliffe 1970) and statistical tests for normality of the model residuals using the PPCC test (Heo et al. 2008) and the VIF test statistic for evaluating multicollinearity among the independent variables (see Helsel and Hirsch 2002; and Kroll and Song 2013). The signs of the model coefficients are stable and consistent with our theoretical expectations, for example: reservoir yields increase with increases in mean annual flow and decrease with increases in the standard deviation of the annual flows.

Table 1 Estimated coefficients for reservoir yields implied by McMahon SRY model and our yield model

Dependent variable	Model parameter	McMahon yield model	Our yield model	VIF
Constant (10^6)	exp(a)	0.780	0.651 (−54.69)	
Ln (μ)	b	1.225	1.135 (287.24)	23.5
Ln (σ)	c	−0.643	−0.342 (−70.03)	28.9
LN ( )	d	0.0253	0.017 (17.01)	1.46
LN (Z_R)	e	−0.566	−0.306 (−29.06)	1.05
LN (S)	f	0.411	0.203 (121.43)	4.48
NSE (%)		98.1	99.2	
Adj R^2			99.2	
Pred R^2			99.17	
SEE (%)			22.1	
Normality test		0.9246	0.9833	

For all predictors the *p*-values <0.00001. Included in parenthesis are the t-ratios associated with model coefficients. At 5 % significance level, yields a critical value of PPCC=0.9998 using Heo et al. (2008) thus we must reject assumption of normality of model residuals

Analogous to the study by McMahon et al. (2007c), inclusion of autocorrelation of the annual streamflows did not lead to improvement in model goodness-of-fit. Both predicted R^2 and adjusted R^2 are extremely high indicating very high explanatory and prediction power of our yield model. The VIFs for μ and σ are high indicating multicollinearity between these two variables, however, given the extremely high goodness-of-fit associated with this model, combined with the large sample size used to create the model (sample size=12,413), concerns over the high values of VIF are not warranted here (see Kroll and Song 2013).

The very high NSE values show that the goodness-of-fit of the model predictions is excellent. All of yield ratios (ratio of yield to mean annual streamflow) obtained using our yield model and the implied McMahon et al. (2007a) yield model are compared in Fig. 1a and b, respectively. Figure 1b indicates that when the McMahon et al. (2007c) model in (1) is used to estimate yield, the yield ratio is greater than unity for 406 of the 12,412 cases (3.3 %) whereas when our model summarized in (2) and Table 1 is used all the values of the yield ratio illustrated in Fig. 1a remain in the expected range [0 1].

Cross validation experiments were performed to evaluate the ability of the McMahon et al. (2007a) regression model in (1) (solved for yield) as well as our yield model in (2) to predict water supply reservoir yields in practice. The K-fold validation method was used where for each of the fold (or split of the data), K-1 folds were used as the training dataset while the other one fold was used for validation. In this study a 10-fold was considered which implies that validation was carried out ten times. Regression was done for each of the training sets using the predictor variables in the water supply model. This training regression model was then used to estimate the yields for all the cases in the validation dataset. The advantage of using the K-fold validation method is that each of the observation in the dataset is used both in the training and a validation set. The Nash-Sutcliffe Efficiency, NSE (Nash and Sutcliffe 1970), and percent bias were then used to evaluate the goodness-of-fit of these cross validation estimates derived from the training regression model. The NSE statistic is a normalized value of Mean Square Error, thus in combination with the percent bias, these two statistic provide a full understanding of the overall goodness of fit of the various model outputs. Our cross validation analyses considers both our yield model and the McMahon et al. (2007a) model in (1) solved for yield. The NSE values obtained for our training regression model and the

McMahon Model are then compared. The values of each of the NSE and the percentage bias estimates obtained from each of the 10-fold validations are summarized in Table 2 along with their mean values at the bottom of the table. The mean NSE from our yield model is 99.1 % while for the McMahon model it is slightly lower at 98.1 %. From Table 2, the percent biases for all the 10-fold validations were all less than 1 % indicating negligible bias.

Figure 2a compares predicted and observed yield values using our yield model and similarly Fig. 2b compares predicted and observed yields using the McMahon et al. (2007a) implied yield model for one of the $K=10$ fold validations. As expected water supply yields predicted from our yield model have higher NSE values than those predicted by the McMahon yield model. Furthermore, for this particular validation fold, our yield model results in negligible biased estimates of yield (Table 2 and Fig. 2a) whereas the mean percent bias in estimates of yield from the McMahon yield model for this particular validation fold was equal to 1.5 % of the sum observed yield (Table 2 and Fig. 2b).

As it was shown in Fig. 1a, the McMahon et al. (2007a) model solved for yield occasionally results in yield values that are greater than mean annual flow which is not physically realistic. Therefore our yield model is generally preferable over McMahon yield model for obtaining preliminary estimates of water supply reservoir yields corresponding to a particular combination of storage and reliability. Overall the high value of NSE, combined with the small percent bias indicates that the both yield models have extremely high predictive power.

3.2 Generalized SRY Models for Estimation of Storage Capacity

In this section we evaluate the global SRY model for estimating storage capacity developed by McMahon et al. (2007a) summarized in Eq. (1) and we develop a new model. Since their study dealt with many other issues, their analysis of the goodness-of-fit of that model was limited to reporting only a few statistics such as adjusted R squared and the predicted R squared. Here we considered several other goodness-of-fit measures and influence statistics to more fully evaluate their resulting model.

We begin by evaluating whether or not there were any potentially influential sets of independent variables which could cause difficulty when fitting a regression model. We employ the DFFIT statistic (Belsley et al. 1980) to evaluate whether or not any influential observations exist in the dataset used by McMahon et al. (2007c) to develop Eq. (1). Influential

Table 2 NSE and percent bias results for the yield models

	Our yield model NSE	McMahon model NSE	Our yield model percent bias	McMahon model percent bias
	99.1	98.1	1.40	0.278
	99.1	98.1	1.42	0.392
	99.2	98.1	1.44	−0.267
	99.0	98.0	1.38	−0.645
	99.1	98.1	1.57	−0.323
	99.1	98.1	1.53	0.232
	99.1	98.1	1.53	0.879
	99.1	98.1	1.68	−0.153
	99.1	98.1	1.59	−0.076
	99.2	98.7	1.59	−0.321
Mean	99.1	98.1	1.50	0.035

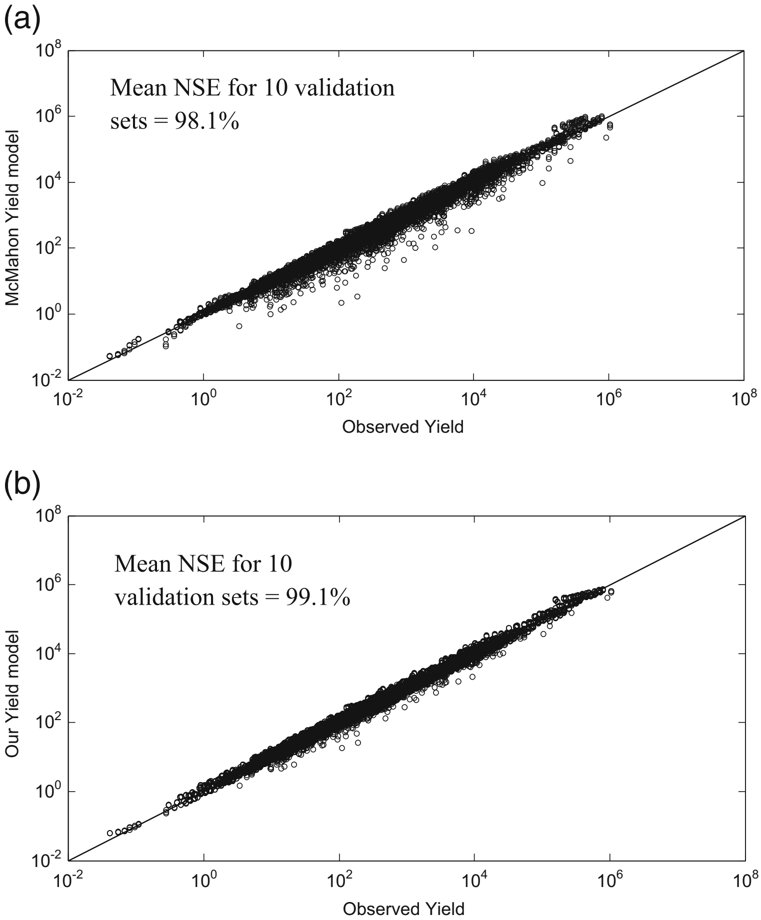

Fig. 2 **a** Comparison of predicted yield estimates using McMahon et al. (2007c) model in (1) solved for yield with observed yields for one of the validation datasets. **b** Comparison of predicted yield estimates using our yield model summarized in (2) and Table 2 with observed yields for one of the validation datasets

observations are defined as observations whose inclusion would lead to very different model coefficient values then if they were omitted. DFFIT(i) is an influence statistic for the ith observation, that is determined from the prediction residuals which are obtained from a model developed when the ith observation is left out. Using DFFIT(i), an observation is said to have unusually large influence when $|DFFIT(i)| > 2\sqrt{p/n}$ (see Belsley et al. 1980; and Helsel and Hirsch 2002) where p is the number of parameters in the model including the constant and n is the total number of observations in the dataset.

Figure 3a and b illustrate a plot of reservoir storage capacity estimates versus the influence statistic DFFIT for the storage model given in Eq. (1) developed by McMahon et al. (2007c) using weighted least squares (WLS) and the storage model developed here using Iteratively Reweighted Least Squares (IRLS), respectively. Both models are also summarized in Table 3. For a model with p=6 and n=12,413, the critical value for DFFIT is 0.044. For the McMahon et al. (2007c) model developed using WLS (Fig. 1a) there are about 700 data points with |DFFITs(i)|>0.044 and these are considered to have unusual influence on the estimates of

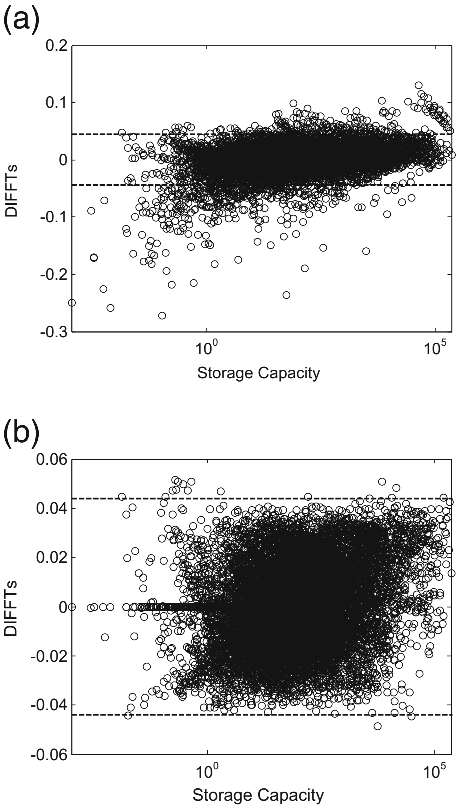

Fig. 3 **a** A plot of reservoir storage estimates (in cubic meters) and DFFIT for the SRY model in Eq. (1) developed by McMahon et al. (2007c) using Weighted Least Squares. The horizontal lines represent the critical value for DFFIT. **b** A plot of reservoir storage estimates (in cubic meters) versus the influence statistic DFFITS for models developed in this study using Iterative Reweighted Least Squares (IRLS). The horizontal lines represent the critical value for DFFIT

model coefficients for the regression model given in Eq. (1). Surprisingly, there are several observations with DFFIT as large as five times the critical value. The model residuals corresponding to the WLS model in (1) are reasonably well approximated by a normal distribution as evaluated using the PPCC value (=0.9975) which is only very slightly less than the critical value (PPCC=0.9998) corresponding to a 5 % level significance, for this sample size.

Due to the existence of influential observations and because the null formal hypothesis of normality of the model residuals must be rejected, robust regression is needed to develop the regression model. We use Iteratively Reweighted Least Squares (IRLS) also known as Iteratively Weighted Least Squares. IRLS is an attractive algorithm for fitting regression models because it is resistant to influential data points and non-normal model residuals as is the case here. Using IRLS, weights are derived from the observations, as follows. An ordinary least squares (OLS) regression is first fit to the data with the weights initially set equal to one. Data points near the OLS model are given weights near unity, while points further away have lower weights. A weighted least squares regression is then computed and the process repeated. After about two iterations the weights become stabilized, and a final IRLS line results (Helsel

Table 3 Summary of model coefficients and goodness-of-fit statistics for model shown in Eq. (1) developed using WLS (McMahon et al. 2007c) and IRLS (this study)

Dependent variable (1)	Model parameter (2)	Model parameter estimate using WLS (3)	Model parameter estimates using IRLS (5)
Constant (10^6)	exp(a)	1.932	1.828 (26.8)
Ln (μ)	b	−3.254	−2.977 (−155)
Ln (σ)	c	1.599	1.563 (144)
LN ( )	d	−0.074	−0.061 (−23.9)
LN (D)	e	2.670	2.431 (153)
LN (Z_R)	f	1.445	1.376 (51.0)
Adj R^2 (%)		89.8	94.2
Pred R^2 (%)		89.8	94.1
NSE_{Real} (%)		89.66	89.39
NSE_{Log} (%)		65.21	66.26
SEE (%)		94.9	57.8
PPCC		0.9975	0.9319
BCF		0.8009	1.187

For all predictors the *p*-values <0.00001. Included in the parenthesis are the t-ratios. At 5 % significance level, PPCC<PPCC critical=0.9998 using Heo et al. (2008) thus reject normality assumption for residuals in both cases

and Hirsch 2002). There are several weight functions which have been used to compute weights. A common and useful one is the bisquare weight function (Mosteller and Tukey 1977; Helsel and Hirsch 2002) which is used here. The results of the analysis using IRLS is shown in Fig. 3b. There is a great reduction in the number of influential observations with now only about 20 such observations. The maximum DFFIT is now less than 1.5 times the cut-off value of 0.044. Based on the results of Fig. 4a and b the robust regression method IRLS appears more suitable for developing the regression models for reservoir storage capacity than WLS for the problem considered here.

Therefore we used the IRLS to develop a slight improvement over the SRY storage model developed by McMahon et al. (2007a) in Eq. 1. Columns 3 and 4 in Table 3 compare the storage models developed using the WLS and IRLS respectively. In general, the model coefficients based on IRLS are always lower than when WLS is used. There are also significant improvements in the adjusted and prediction R squared values of 89.9 and 89.7 respectively when WLS is used compared to the values of 94.2 and 94.1 respectively when IRLS is used as well as associated reductions in the standard error of the model residuals (SEE). Similar to the McMahon et al. (2007c) study, the model coefficient for the independent autocorrelation variable φ was not found to be significantly different from zero. Using the probability plot correlation coefficient (PPCC) test (Heo et al. 2008), we could not accept the normality hypothesis of the model residuals for either the WLS or the IRLS model, and the model residuals of the WLS model are much more nearly normally distributed, thus we would be reluctant to construct confidence and/or prediction intervals associated with model coefficients or predictions corresponding to the IRLS model due to this fact. Also included in Table 2 are the Nash-Sutcliffe Efficiency (NSE) (normalized mean square error) values for both methods computed in both real and log space. These NSE values are high showing that the fit of both models to the observations can be considered satisfactory with little difference

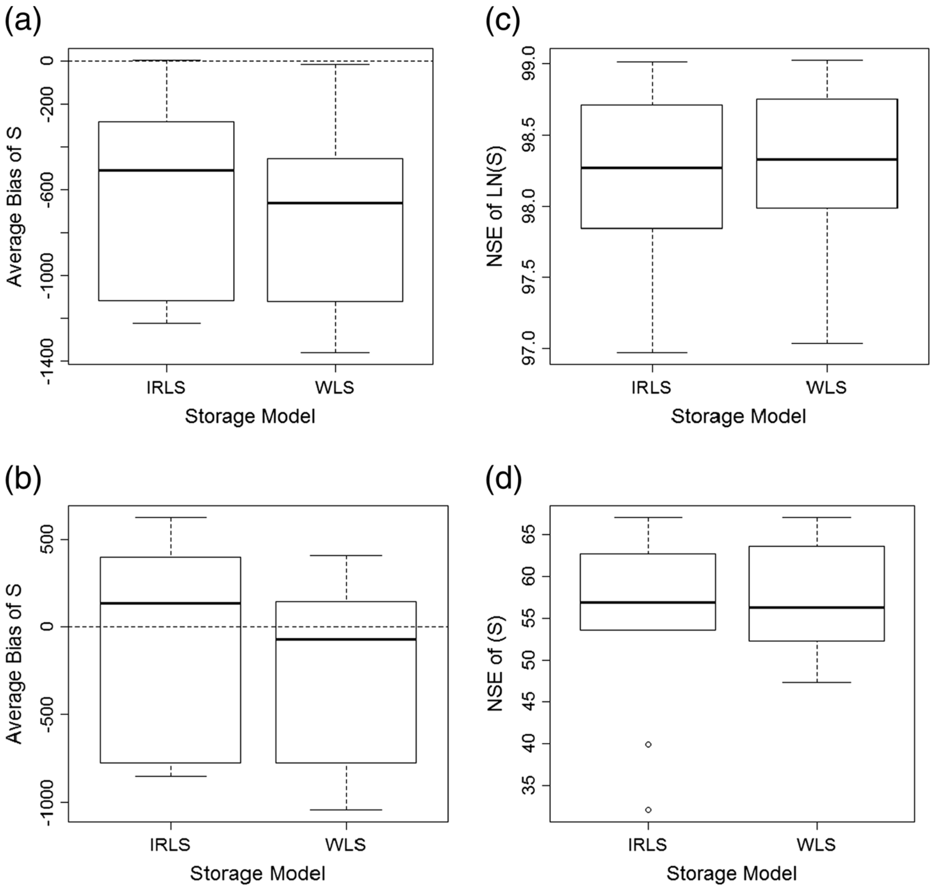

Fig. 4 **a** Boxplots of the average model residuals across the 10 fold validations corresponding to both the storage SRY regression model developed here (IRLS) and the McMahon et al. (2007a) storage model (WLS) given in Eq. 1 without a bias correction (BCF). **b** Boxplots of the average model residuals across the 10 fold validations corresponding to both the storage SRY regression model developed here (IRLS) and the McMahon et al. (2007a) storage model (WLS) given in Eq. 1, with a bias correction (BCF). **c** Boxplots of the average values of NSE across the 10 fold validations corresponding to both the storage SRY regression model developed here (IRLS) and the McMahon et al. (2007a) storage model (WLS) given in Eq. 1 without a bias correction. **d** Boxplots of the average values of NSE across the 10 fold validations corresponding to both the storage SRY regression model developed here (IRLS) and the McMahon et al. (2007a) storage model (WLS) given in Eq. 1 with a bias correction (BCF)

between the NSE values for both methods. A more rigorous goodness-of-fit evaluation is performed using k-fold cross validations.

The k-fold cross validation experiment described previously for our yield model was repeated to evaluate which of the two storage models perform better in validation mode, which more closely represents their ability to generate reliable predictions in practice than the summaries reported in Table 2. Model estimates of storage capacity, S based on IRLS and WLS regression were determined for each of the $k=10$ training and validation sets. We then computed the bias of the resulting storage estimate using both WLS and IRLS models. Due to bias introduced by the need for a retransformation of the natural logs of S back to real space, a bias correction factor, $BCF=\exp(s^2/2)$ was applied to predicted storage capacities (in real space, where s is the standard deviation of the model error in log space for each model). Figure 4a and b illustrate the average residuals associated with storage estimates corresponding to both storage models (with and without the bias correction factor) and Fig. 4c and d illustrate the average NSE values associated with estimates of Sacross the 10 fold validations. The residuals and NSE values illustrated in Fig. 4 are average values across the 10 fold cross validation experiments. Figure 4 illustrates documents that the BCF is effective for reducing the bias. Figure 4 illustrates that the bias associated with storage estimates from theWLS model is slightly less than for the IRLS model, and the NSE values for storage values based on IRLS is slightly better than for the WLS model. Therefore the IRLS model results in slightly higher (better) goodness-of-fit but at the slight expense of higher bias. These results are for average results across the 10 fold validation sets.

We conclude from the comparisons in Fig. 4 that the bias correction factor led to much more unbiased estimates of S for both the WLS and IRLS models, thus we strongly recommend the use of such a BCF when these models are applied in the future. The resulting values of the BCF for each model are summarized in Table 2 and we recommend that they be used in any application of either model. When a BCF is applied, both models perform similarly. We also note from a comparison of NSE of S in Fig.4c and d, that both WLS and IRLS perform similarly, thus there is no compelling reason to recommend either model, though the model based on IRLS has a very slight advantage in terms of NSE of S. We also note that in validation mode, neither IRLS nor WLS perform nearly as well as they appeared to perform in calibration mode, because the values of NSE of S reported in Fig. 4c and d, for the validation sets are generally much lower than the values reported in Table 2. This is to be expected and provides a much better reflection of the ability of these models to perform in practice.

4 Summary and Conclusion

A review of the literature revealed that although numerous analytical generalized relationships between storage S, yield Y and reliability R have been developed, nearly all were based on artificial streamflow traces or if based on actual flow records were limited in geographic scope. A primary goal of this study was to carry out a diagnostic analysis and extension to the one generalized global SRY model developed by McMahon et al. (2007c). Our review of the literature further reveals that generalized SRY relationships are useful in a variety of setting ranging from climate change evaluations (Vogel et al. 1997; Lane et al. 1999 and Brown et al. 2012), estimation of sampling properties of yields (Kuria 2014) and storage capacities (Vogel and Stedinger 1988), stochastic streamflow analysis for reservoir design (Vogel and Stedinger 1988), behavior analysis of water supply reservoir systems (Vogel and Bolognese 1995; Vogel and McMahon 1996; Vogel et al. 1999) and may be extended to rainwater harvesting systems (Hanson and Vogel, 201).. Our primary findings are:

First, previous studies have developed SRY models considering storage capacity as the dependent variable. We show that efforts to solve for yield in such models often leads to yield estimates which are larger than the mean annual flows which is physically unrealistic. We developed a regression model for determining water supply reservoir yield estimates which always leads to yield estimates which are less than the mean annual flow. The global regression model for yield developed here has high adjusted R^2 and prediction R^2 of 99.2 % and 99.12 % respectively, which indicates high explanatory power of the model. A 10-fold cross validation of the model also resulted in high NSE values with a mean value of 99.2, across the ten validation sets and a very small percent bias. As expected the yield model developed here leads to more accurate estimates of yield than those obtained from solving the McMahon et al. (2007c) reservoir storage capacity model given in Eq. (1) for yield.

Further evaluations of the weighted least squares (WLS) regression models developed by McMahon et al. (2007c) raised concerns due to some highly influential observations. A robust regression method termed Iteratively Reweighted Least Squares (IRLS) was used to develop a slightly improved model with both higher explanatory and predictive power and lower standard error of the model estimates. However, after detailed cross validation experiments, our IRLS storage model is only shown to yield a very marginal improvement over the SRY model introduced by McMahon et al. (2007c) summarized here in Eq. 1.

Acknowledgments We are indebted to Murray Peel, Thomas McMahon and Geoffrey Pegram for sharing the global dataset of monthly streamflow upon which this study is based, and for their encouragement and support of our efforts. We are further indebted to the Fulbright Foundation and Tufts University for their support of this research.

References

Adeloye A (2009a) Multiple linear regression and artificial neural networks models for generalized reservoir storage–yield–reliability function for reservoir planning. J Hydraul Eng-ASCE 14(7):731–738

Adeloye A (2009b) The relative utility of regression and artificial neural networks models for rapidly predicting the capacity of water supply reservoirs. Environ Model Softw 24(10):1233–1240

Adeloye A, Lallemand F, McMahon T (2003) Regression models for within-year capacity adjustment in reservoir planning. Hydrologic Sci J 48(4):539–552

Bayazit M (1982) Ideal reservoir capacity as a function of yield and risk. J Hydrol, (Amsterdam) 58(1–2):1–9

Bayazit M, Bulu A (1991) Generalized probability distribution of reservoir capacity. J Hydrol (Amsterdam) 126(3–4):195–205

Bayazit M, Önöz B (2000) Conditional distributions of ideal reservoir storage variables. J Hydrol Eng 5(1):52–58

Beaton AE, Tukey JW (1974) The fitting of power series, meaning polynomials, illustrated on band-spectroscopic data. Technometrics 16(2):147–185

Belsley DA, Kuh E, Welsch RE (1980) Regression diagnostics. Wiley, New York

Brown C, Ghile Y, Laverty M (2012) Decision scaling: linking bottom-up vulnerability analysis with climate projections in the water sector. Water Resour Res 48(9):1–12

Fiering MB (1963) Statistical analysis of the reservoir storage-yield relations. In: Thomas HA Jr, Burden RP (eds) Operations research in water quality management. Harv. Water Resour. Group, Cambridge

Gould BW (1964) Discussion of Alexander GN, Effect of variability of stream-flow on optimum storage capacity. In: Water resources use and management, proceedings of a symposium held in Canberra. Melbourne University Press, Melbourne, pp 161–164

Hanson LS, Vogel RM (2014) Generalized storage-reliability-yield relationships for rainwater harvesting systems. Environ Res Lett 9:075007. doi:10.1088/1748-9326/9/7/075007

Helsel DR, Hirsch RM (2002) Statistical methods in water resources. US Geological Survey, Techniques of Water-Resources Investigations. http://pubs.usgs.gov/twri/twri4a3/

Heo JH, Kho YW, Shin S, Kim S, Kim T (2008) Regression equations of probability plot correlation coefficient test statistics from several probability distributions. J Hydrol (Amsterdam) 355(1–4):1–15

Kroll CN, Song P (2013) Impact of multicollinearity on small sample hydrologic regression models. Water Resour Res 49(6):3756–3769

Kuria FW (2014) Uncertainty analysis of water supply reservoir yield estimates. PhD Dissertation, Tufts University, Medford

Lane ME, Kirshen PH, Vogel RM (1999) indicators of the impacts of global climate change on U.S. water resources. J Water Res Pl-ASCE 125(4):194–204

McMahon TA, Adeloye AJ (2005) Water resources yield. Water Resources Publications, LLC, Highlands Ranch, Chicago

McMahon TA, Mein RG (1986) Reservoir capacity and yield. Elsevier, New York
McMahon TA, Vogel RM, Pegram GGS, Peel MC (2007a) Review of Gould-Dincer reservoir storage-yield-reliability estimates. Adv Water Resour 30(9):1873–1882
McMahon TA, Pegram GGS, Vogel RM, Peel MC (2007b) Revisiting reservoir storage-yield relationships using a global streamflow database. Adv Water Resour 30(8):1858–1872
McMahon TA, Vogel RM, Pegram GGS, Peel MC, Etkin D (2007c) Global streamflows—Part 2 Reservoir storage-yield performance. J Hydrol (Amsterdam) 347:260–271
Mosteller F, Tukey JW (1977) Data analysis and regression. Addison-Wesley, Menlo Park
Nash JE, Sutcliffe JV (1970) River flow forecasting through conceptual models part I—a discussion of principles. J Hydrol (Amsterdam) 10(3):282–290
Pegram GGS (1980) On reservoir reliability. J Hydrol (Amsterdam) 47(3–4):269–296
Phatarfod RM (1977) The Sampling Error of Storage Size. Water Resour Res 13(6):967–969
Phien HN (1993) Reservoir storage capacity with gamma inflows. J Hydrol (Amsterdam) 146(1–4):383–389
Pretto PB, Chiew FHS, McMahon TA, Vogel RM, Stedinger JR (1997) The (mis)behavior of behavior analysis for storage estimates. Water Resour Res 33(4):703–709
Rippl W (1883) The capacity of storage-reservoirs for water-supply. Proc1nst Civ Eng 61:270–278
Silva AT, Portela NM (2012) Stochastic assessment of reservoir storage-yield relationships in Portugal. J Hydrol Eng-ASCE 18(5):567–575
Thomas HA Jr, Burden RP (1963) Operations research in water quality management. Harvard Water Resources Group, Cambridge
Vogel RM, Bolognese RA (1995) Storage-reliability-resilience-yield relations for overyear water supply systems. Water Resour Res 31(3):645–654
Vogel RM, McMahon TA (1996) Approximate reliability and resilience indices of over-year reservoirs fed by AR(1) gamma and normal flows. Hydrol Sci J 41(1):75–96
Vogel RM, Stedinger JR (1987) Generalized reservoir storage-reliability-yield relationships. J Hydrol(Amsterdam) 89(3–4):303–327
Vogel RM, Stedinger JR (1988) The value of stochastic streamflow models in overyear reservoir design applications. Water Resour Res 24(9):1483–1490
Vogel RM, Fennessey NM, Bolognese RA (1995) Storage-reliability-resilience-yield relations for Northeastern United States. J Water Res Pl-ASCE 121(5):365–374
Vogel RM, Bell CJ, Fennessey NM (1997) Climate, streamflow and water supply in Northeastern United States. J Hydrol(Amsterdam) 198(1–4):42–68
Vogel RM, Lane M, Ravindiran RS, Kirshen P (1999) Storage reservoir behavior in the United States. J Water Res Pl-ASCE 125(5):245–254