The examination of reproducibility in hydro-ecological characteristics by daily synthetic flow models

Gene Jiing-Yun You *, Bing-Hao Thum, Fu-Hua Lin

Department of Civil Engineering, National Taiwan University, Taipei 10617, Taiwan

ARTICLE INFO

Article history:
Received 22 October 2013
Received in revised form 6 February 2014
Accepted 12 February 2014
Available online 24 February 2014
This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Peter F. Rasmussen, Associate Editor

SUMMARY

Research into synthetic streamflow generation previously focused primarily on engineering purposes, while disregarding hydro-ecological implications with regard to daily flow. This study investigated the applicability of daily synthetic flow models in the representation of hydro-ecological characteristics. This study applied flow duration curves and three performance indices as well as a method of statistical testing to identify whether Indicators of Hydrologic Alteration (IHA) derived from historical and synthetic flow data have the same distribution. This enabled us to evaluate the capacity of the synthetic flow models in capturing other important hydrological characteristics. This study examined various representative methods, most of which proved effective in simulating the magnitude of monthly flow and high flow within short durations. However, low flow conditions remain problematic, particularly over short durations. Among the various models, annual flow simulation using annual–daily disaggregation demonstrated good reproducibility, but its applicability with regard to within year fluctuations may be limited by the relatively low degree of freedom. Direct simulation models (shot noise and Markov-based) proved suitable in dealing with engineering problems based on monthly averages and annual daily maximum flow. Linear regression models, PARMA, and nonparametric models, modified k-NN, were shown to be well-suited to the representation of monthly stream flow. In conjunction with nonparametric bootstrap disaggregation, these methods proved effective in representing daily flow patterns. Coupling these methods with the shot noise disaggregation model was the least recommended due to bias resulting from numerical errors associated with repeated convolutions.

Available online 24 February 2014
Accepted 12 February 2014
Received in revised form 6 February 2014
Received 22 October 2013

1. Introduction

Reliable hydrologic records are fundamental to the planning and utilization of water resources; however, records of sufficient length are seldom available. As a result, synthetic streamflow models are commonly used to generate representative data. Retaining the statistical characteristics of historical flow time series make it possible for synthetic hydrologic sequences to reproduce hydrologic features and generate variability sufficient to meet the needs of water resource engineers. Conventional synthetic hydrologic sequences are generated using linear and parametric models, such as autoregressive moving average models (ARMA) and lag-1 periodic autoregression models (PAR) [Salas, 1985]. These stochastic hydrological models provide hydrological sequence simulations over relatively long timescales, such as monthly, seasonally or annually; however, they tend to be somewhat limited in the simulation of daily discharge.

Beard (1967) and Green (1973) were the first to attempt daily synthetic flow simulation. Beard used a second-order Markov model and Log-Pearson Type III distribution to generate daily hydrological sequences; however, these efforts proved insufficient in the reproduction of storm hydrographs and baseflow recession. Green developed a method based on linear interpolation using five-day averages for the synthesis of daily flow; however, this approach failed to provide sufficient accuracy for the simulation of high-flow conditions. Both of these approaches were only able to reproduce a limited number of hydrological characteristics of discharge [Treiber and Plate, 1977]. Weiss (1977) took into account the characteristics of rainfall runoff and adopted a filtered Poisson method to develop a shot noise (SN) model for synthetic flow simulation. That study provided a comprehensive discussion of the fundamental concepts underlying the model as well as the corresponding processes for the calibration of various parameters. A disaggregation model based on the same conceptual approach has also been proposed. Treiber and Plate (1977) furthered the shot noise model by utilizing pulse processes and system transform functions to simulate runoff associated with daily synthetic

* Corresponding author. Tel/fax: +886 2 33664238.
E-mail address: genejyu@ntu.edu.tw (G.J.-Y. You).

http://dx.doi.org/10.1016/j.jhydrol.2014.02.047
0022-1694/© 2014 Published by Elsevier B.V.
streamflow. They employed a Markov chain and autoregressive processes to produce correlated pulses. Variations on this approach using various techniques for the estimation of pulse and transform function parameters were also developed by Hino and Hasebe (1981, 1984), Battaglia (1986), and Wang and Vandewiele (1994). Schneider and Schultz (1982) and Kron et al. (1990) modified Treiber and Plate’s (1977) approach and applied it to various regions. Subsequent researchers have focused on overcoming the limitations of the Treiber and Plate model and shot noise models to extend their applicability (Claps and Murren, 1994; Claps et al., 1993; Murren et al., 1997). More recently, Claps et al. (2005) discovered that using discharge increment pulses (DIP) to derive corresponding response functions can lead to the overestimation of flow events, with small pulses generating a considerably greater number of events than actually occur. Thus, the filtered peaks over threshold (FPTO) approach was proposed to filter pulses, such that only significant peak values are retained. A case study conducted on seven catchment areas in northern Italy demonstrated the superior simulation results of the FPTO method.

Another daily flow model utilizes Markov models to determine the probability of excess rainfall (wet/dry), prior to estimating the magnitude of runoff in accordance with hydrologic and statistical characteristics (Xu et al., 1997, 2002, 2003). Aksoy and Bayazit (2000a, 2000b) applied second-state (2000a) and third-state (2000b) Markov models to conduct similar hydrologic synthetic flow simulations. Claps et al. (2005) compared these Markov-based methods with a shot noise model according to their complexity and the number of associated parameters, concluding that shot noise models benefit from model parsimony when taking both the quality and efficiency of the simulations into account. However, all of these models are based primarily on parametric methods, which are affected by assumptions related to probability distribution, thereby leading to inconsistencies between simulation and historical data.

In contrast, non-parametric methods do not require assumptions related to the distribution of simulated flow, making them more flexible in their applicability. A number of studies have investigated non-parametric flow synthesis (Yakowitz, 1973, 1979, 1985, 1993; Schuster and Yakowitz, 1979; Karlsson and Yakowitz, 1987a,b; Smith, 1991; Smith et al., 1992). The bootstrap approach is a widely accepted non-parametric technique. Lall and Sharma (1996) developed a monthly flow generation model based on a k-nearest neighbor (k-NN) bootstrap method using historical data. In their model, the k-NN method takes the events of the previous period to obtain the k number of historical data nearest the simulation values. By assigning a weighted factor with respect to distance, the model resamples from the data set of the k number to formulate events in the current period. Lall and Sharma (1996) demonstrated that the k-NN approach outperformed other parametric methods in fitting the trends and peak values of monthly streamflow. Variations of the k-NN approach have also been applied to the simulation of univariate and multivariate weather data, such as rainfall, radiation, wind speed, dew point, humidity, and daily maximum and minimum temperatures (Rajagopalan and Lall, 1999; Yates et al., 2003; Harrold et al., 2003; Prairie et al., 2007).

The bootstrap approach cannot be used to simulate observed values that have not previously occurred, which severely limits its capacity to reproduce variability in data. Prairie et al. (2005, 2006) improved on the original method with the development of the modified k-NN approach, which uses residuals as sampling standards to enhance the variability of simulated data. The applicability of the modified k-NN method has been verified in a number of subsequent studies (Grantz et al., 2005; Singhhatna et al., 2005).

Another approach to daily streamflow simulation involves simulating total flow over longer time steps, such as monthly or annually (ex. ARMA, k-NN, and modified k-NN), and then applying a disaggregation model to divide the total flow over the period into days. Valencia and Schakke (1973) began work into disaggregation models, followed by numerous other researchers (Mejia and Roussele, 1976; Tao and Delleur, 1976; Stedinger and Vogel, 1984; Maheepala and Perera, 1996; Koutsoudiannis and Manetas, 1996). Most disaggregation methods have been based on parametric statistical theory and various non-parametric methods, such as the kernel-based approach, have proven inefficient in high-dimensional problems resulting from calculation complexity (Lall et al., 1996; Sharma et al., 1997; Tarboton et al., 1998). To overcome the limitations of the kernel-based approach, Prairie et al. (2007) developed the k-NN based disaggregation approach, which replaced the kernel density function with the k-NN bootstrap technique. This modification simplified computation and avoided the prerequisite of Gaussian assumptions and boundary issues. However, the application of this approach to daily flow disaggregation remains problematic. Kumar et al. (2000) adopted a k-NN bootstrap technique in conjunction with an optimization scheme for the spatial and temporal disaggregation of monthly streamflow to daily flow. This optimization scheme significantly increases the number of parameters required. In the estimation of discharge at five stations, over 1500 parameters were required for calibration. This demonstrates the fact that complexity associated with daily flow characteristics is exceedingly difficult to describe using parametric methods, suggesting that non-parametric methods might be more suitable. Using the San Juan River as an example, Nowak et al. (2010) demonstrated a k-NN-based disaggregation model from annual to daily streamflow. This model resampled historical daily flow proportion vectors to maintain the continuity of daily flow behavior in the results. Unfortunately, parameters relevant to synthetic flow data, such as the mean, skewness, maximum, and minimum, did not differ significantly from previous trends. As a result, flow simulation may suffer from a lack of variability.

Computer software for the analysis and simulation of stochastic hydrology was first introduced in the 1970s, in conjunction with theoretical developments in this field. The US Bureau of Reclamation (USBR) completed Lane’s applied stochastic techniques (LAST) in the late 1970s. LAST was used to generate flow conditions for multiple stations within a given watershed for the purpose of engineering (Lane, 1979; Lane and Frevert, 1988). LAST has also been widely adopted by other governmental agencies and academic/research institutions. Another software package, SPIGOT (Grygier and Stedinger, 1990), made a number of improvements to the algorithms employed by LAST, and applied univariate/multivariate statistical generation methods with disaggregation models for the simulation of streamflow. Stochastic analysis modeling and simulation (SAMS) is a more comprehensive model recently developed by Colorado State University (Salas et al., 2006; Sveinsson et al., 2007). The model provides comprehensive functionality for the analysis of hydrologic time series using single or multiple stations. This program is applicable to monthly, seasonal, and annual flow simulation; however, it still lacks complete functionality with regard to daily flow analysis.

Until recently, synthetic streamflow generation has focused primarily on engineering purposes, largely disregarding the ecological implications. In the last few decades however, ecological issues have been receiving increased attention. The integrity of ecosystems is closely bound to hydrologic flow conditions. Effective ecological management requires that the existing regime be characterized using ecologically relevant hydrological parameters. The Indicators of Hydrologic Alteration (IHA) is a hydrological analysis tool commonly used for ecological evaluation (Richter et al., 1996, 1997). By processing daily hydrologic records, the IHA characterizes within-year variations in streamflow on the basis of a series
of hydrologic attributes. These hydrologic parameters assess river flow regimes with respect to the magnitude, frequency, duration, timing, and rate of change associated with flow. These hydrological indicators are widely applied in the study of water resources related to ecosystems from a quantitative perspective, helping to bridge the gap between hydrology and ecology (Koel and Sparks, 2002; Magilligan and Nislow, 2005; Suen and Eheart, 2006; Shiau and Wu, 2006; Hu et al., 2008; Chen et al., 2010; Lian et al., 2011). Including IHA, currently over 170 hydrologic indicators have been developed to describe various components of flow regimes. The intercorrelation of these indicators has resulted in a degree of redundancy. Olden and Poff (2003) demonstrated the efficacy of IHA in representing the entire ordination space and capturing most of the information provided by 171 indices. However, their findings indicated a high degree of intercorrelation among a number of IHA indices. Gao et al. (2009) further evaluated the redundancy of IHA indices and proposed metrics (termed ecodeficit and ecosurplus) comprising a total of ten parameters. Poff and Zimmerman (2010) published a comprehensive review of the literature related to quantitative relationships between flow alteration and ecological response, thereby providing a highly detailed discussion on the ecological consequences of changes in the flow regime.

Based on conventional experience, most synthetic flow techniques using longer time intervals are able to meet the demands of water resource management and flood prevention. However, it remains uncertain whether synthetic daily flow data can be applied to engineering problems, particularly those related to ecology, and little research has been conducted to address this issue. Black et al. (2002) applied micro-low flows (MLF) to investigate ecological requirements in the UK, and then used IHA to evaluate the MLF program. They also addressed the need for the means with which to evaluate the application of synthetic techniques for ecological purpose (Black et al., 2005; Arthington et al., 2006; Richter et al., 2006).

Motivated by this concern, this study investigated whether daily synthetic hydrologic flow sequences are capable of representing ecological characteristics. At present, no commonly accepted standards have been devised for the generation of daily flow analysis. Therefore, we examined several representative methods found in the literature and evaluated their representation of ecological characteristics to provide a reference for future planning.

2. Methodology

This study selected representative methods from the literature and compared the synthetic flow data with historical records as they pertain to hydrologic, statistical, and ecological characteristics. In this way, we were able to determine the reproducibility of various methods. The methods were divided into two categories: (I) stochastic synthetic flow models that generate daily flows directly and (II) models that generate flows at higher resolutions, which are then disaggregated into daily flows. The latter category includes annual to daily and monthly to daily disaggregation models. Based on this framework, we selected a variety of methods and applied flow duration curves, hydrological indices and IHA indicators with statistical tests for comparison. The overall framework is outlined in Fig. 1.

2.1. Daily flow simulation

Category (I) comprises models capable of deriving daily flows directly, including the shot noise model and the Markov-based model. These methods are introduced in the following.

2.1.1. Shot noise model

The shot noise model is used to reproduce the basic mechanisms of rainfall–runoff based on the filtered Poisson process. Pulses are employed to reproduce the occurrence of rainfall events, in conjunction with a response function, which is interpreted as the unit hydrograph describing the response to a unit volume of effective rainfall. This type of simulation of daily streamflow series is better able to reproduce the presence of peaks and recessions. In addition, this approach provides a means to clearly differentiate the estimation of the stochastic component (effective rainfall) from that of the deterministic component (response function). Thus, flows can be expressed as

\[
q(t) = \sum_{i=1}^{N} u_i \cdot h(t - \tau_i)
\]

Fig. 1. A conceptual framework to examine the reproducibility by daily synthetic flow models in this study.
where \( t \) is the time; \( i \) represents the pulse indicator; \( u_i \) is the corresponding intensity of the \( i \)th pulse; \( N \) is the number of pulses occurring up to time \( t \); \( h(\cdot) \) denotes the response function; \( \tau_i \) is random times when the \( i \)th pulse occurs, and \( t - \tau_i \) is the time difference between the occurrence of the \( i \)th pulse and simulation time \( t \), which is also the delay time of recession. Based on the concept of unit hydrographs and the principle of superposition, Eq. (1) calculates the input and output relationships from the linear conceptual scheme. As outlined by Weiss (1977), the Poisson process can be used to determine whether pulse \( u_i \) occurs and its time \( \tau_i \), as displayed in Eq. (2).

\[
 f(k, \lambda) = \frac{\lambda^k}{k!} e^{-\lambda}
\]

where \( \lambda \) denotes the Poisson rate (the density of occurrences within a given time interval), reflecting hydrologic conditions varying month by month; and \( f(k, \lambda) \) represents the probability of \( k \) occurring under these conditions.

Shot-noise models require the assignment of probability distribution, \( f(\cdot; u) \) for random variable \( u \), the intensity of the pulse. Most previous studies assumed exponential distribution, in which the probability density function can be expressed as

\[
 f(\cdot; u) = \frac{1}{\mu_u} e^{u/\mu_u}
\]

where \( \mu_u \) is the mean of pulse intensity. Eqs. (2) and (3) are generally referred to as Poisson-exponential models.

A number of formulas have been proposed for the response function \( h(\cdot) \). The features of recession are generally described using exponential functions. This study adopted the response function suggested by Claps et al. (2005), as follows:

\[
 h(t) = c_0 \delta(t) + \frac{c_1}{k_1} e^{-t/k_1} + \frac{c_2}{k_2} e^{-t/k_2}
\]

where \( \delta(t) \) is the Dirac delta function; \( c_0, c_1, c_2 \) are dimensionless constants subject to \( c_0 + c_1 + c_2 = 1 \) for mass conservation, and \( k_1 \) and \( k_2 \) are storage coefficients, in which \( k_2 > k_1 \). The first term in the response function of Eq. (4) presents the instant response that reaches the basin outlet directly with sub-daily response times; the second term shows the direct runoff recession over the short-term. The third is the long-term recession flow regime capable of providing an appropriate representation of baseflow.

Murrone et al. (1997) proposed a well-designed method for the estimation of model parameters. First, we identify pulses (effective rainfall events) corresponding to the days when streamflow increases, known as discharge increments pulses (DIPs). The magnitude of these pulses is initially assumed to be equal to the discharge increment. With the DIP providing initial pulse conditions, an optimization technique is used to determine the parameters of the response function \( h(\cdot) \) by minimizing the sum of the quadratic distances between observed and synthetic flow data. Using the newly generated response function, a new pulse series is inversely estimated by performing deconvolution with the original discharge data. These steps are repeated until convergence to derive the optimal pulse sequence and response function. Readers are referred to Murrone et al. (1997) and Claps et al. (2005) for more details related to this parameter estimation procedure.

### 2.1.2. Markov-chain method

Another branch of daily flow simulation includes methods based on the Markov-chain. The framework is very similar to the daily rainfall generation model proposed by Richardson (1981), Aksoy and Bayazit (2000a, 2000b) and Aksoy (2000, 2003, 2004) made improvements to the same basic model and simulated increases and decreases in discharge using the Markov process. Their model treats the magnitude of increase as a random variable, generated using a given probability distribution function, whereas the magnitude of a decrease is described by exponential recession. The model for flow generation proceeds through the following steps: (1) determining whether the flow on a given day increases or decreases; and (2) deriving the ascension and recession curves of the hydrograph. In this way, daily flows are generated sequentially.

For the purpose of analysis, this study used two-state (W–D) Markov chains to determine the probability of wet and dry days. According to the definitions given by Aksoy and Bayazit (2000a), days with or without increases in flow are wet days, whereas days on which the flow decreased are considered dry. Probability matrix \( P \) reflects the hydrologic characteristics in a given month:

\[
 P = \begin{bmatrix} P_{WW} & P_{WD} \\ P_{DW} & P_{DD} \end{bmatrix}
\]

The total of each row in matrix \( P \) equals 1. \( P_{WW} \) indicates the probability of a wet day remaining a wet day; \( P_{DW} \) denotes the probability of a wet day becoming a dry day; \( P_{WD} \) signifies the probability of a dry day becoming a wet day, and \( P_{DD} \) is the probability of a dry day remaining a dry day. Matrix \( P \) is produced using historical observation records and varies according to month.

In accordance with the probability of being wet or dry from matrix \( P \), the model simulates the hydrologic states of each day. A wet day identified by flow greater than the flow of the previous day is located on the ascension curve of the hydrograph; a dry day with flow that is lower than the flow of the previous day is located on the recession curve. Once ascension or recession has been determined, the magnitude of the increase or decrease in flow is simulated. For ascension, this approach treats discharge increments as random variables \( (q_i) \), rather than the flow during ascension. Statistical frequency analysis is used to obtain the probability distribution of \( q_i \). Previous researchers have suggested that discharge increment \( q_i \) follows a Gamma distribution, with the probability density function as follows:

\[
 f(q_i) = \frac{1}{\Gamma(q_0) \cdot \beta_q} q_i^{q_0 - 1} e^{-q_i/\beta_q}
\]

The moment method is applied to obtain parameters \( \mu_q \) and \( \beta_q \) for each month, resulting in a total of 24 parameters. In the simulations, the flow is stochastically generated according to distribution. Several successive increments are treated as a single event and flows are ranked from small to large to present the characteristics of ascension.

The recession curve is split into two stages. The first stage is the upper recession, containing curves with a peak flow value greater than the observed monthly mean. The second stage is the lower recession, containing curves with a peak flow value smaller than the observed monthly mean. The upper recession corresponds to the fast component of the flow and the lower recession corresponds to the slow component of the flow feeding the stream in a rainless period. Recession \( q_i \) is assumed to take the following form:

\[
 q_i = \begin{cases} q_0 \cdot e^{-h_1 t_1} & I \\ q_0 \cdot e^{-h_2 (t_i - t_{i-1})} & II \end{cases}
\]

where \( h_1 \) and \( h_2 \) are the recession coefficients for the upper and lower parts of the recession curve, respectively. In stage I, \( t_1 \) is the number of days after the peak, \( q_0 \) is the preceding peak flow value. In stage II, \( t_i \) is the time from the start of the lower recession, \( q_i \) is the initial flow in the lower part of the recession. A recession curve decays with the stage I equation until the flow takes a value smaller than the value of the observed monthly mean flow. If the peak flow of a recession or flow value for a given day on the recession curve is smaller than the observed monthly mean, the equation for stage II is used until the end of the recession.
2.2. Monthly and annual flow simulation

In Part (II), the models were used to generate simulated flows at large resolutions (months and years) in conjunction with disaggregation, which divided the larger flows into daily flows. We adopted the common linear parametric model for annual and monthly streamflow generation and added a non-parametric model for monthly flows to enable comparison.

2.2.1. Linear parametric model for generation of monthly and annual flow

In stochastic hydrology, the development of flow generation models based on linear regression for monthly and annual simulation are relatively advanced, and many relevant software packages are available. This study employed the 2007 version of Stochastic Analysis, Modeling, and Simulation (SAMS). SAMS deals with the stochastic analysis, modeling, and the simulation of hydrologic time series, such as annual and monthly streamflows. This program provides several parametric modeling approaches and data analysis capabilities, such as multivariate autoregressive modeling and disaggregation linear models. This study used the commonly applied ARMA model for annual flow and the periodic autoregressive moving average (PARMA) model for monthly flow. Both of these models have been widely applied in the generation of streamflow time series in real-world situations, demonstrating reasonable effectiveness in the simulation of complex hydrologic processes and water resources systems (Yevjevich, 1972; Stedinger and Taylor, 1982; Bras and Iturbe, 1985; Salas, 1985).

2.2.2. Modified k-NN monthly method

A Parametric framework restricts data to Gaussian distribution and is limited in its ability to capture structures of nonlinear dependence. Recent developments in nonparametric time series modeling (Lall and Sharma, 1996; Tarboton, 1997; Rajagopalan and Lall, 1999) have alleviated some of the shortcomings of parametric frameworks. In addition to parametric methods, this study employed the non-parametric modified k-NN method for monthly flow generation. The kk-NN bootstrap technique takes the generated distances and historical streamflow of the previous month as references (known as feature vectors). Generally, historical data with a neighborhood (k) of specific size are considered. The data are designated using various weights before being selected at random using a stochastic model. From the selected historical data, the streamflow for the following month can be read and assigned as a simulated or forecast value (known as the successor). Thus, the flows in the month following selected historical events constitute the monthly flow data to be generated. This procedure is then repeated until the end of simulation procedure. In this approach, k represents the size of the sampling candidates. Previous studies have suggested that k = \sqrt{n}, where n is the number of historical records. The corresponding weighting function (ex, kth sample) is based on the distance with the feature vector and reference written as follows:

\[
w(k') = \left( \frac{1}{k} \right) \left[ \sum_{i=1}^{k} \frac{1}{k} \right] \quad \forall \ k' = 1, 2, \ldots k
\]

where i is the value closest to 1. This weight function gives more weight to the nearer neighbors and less weight to farther neighbors. The weights are normalized to create a probability mass function or weight metric. Prairie et al. (2006) developed a modified version of the k-NN bootstrap approach to overcome the shortcomings of previous versions. As a bootstrap method, values not observed in the historic data are not generated in the simulations. Prairie et al. (2006) obtained the main trends in the previous and subsequent months using polynomial regression. In their model, the streamflow time series is partially integrated with a periodic autoregressive model, a local polynomial fitting for each month, dependent on the previous month. The k-NN approach was used only to resample residuals from the partially prewhitened streamflow. Readers are referred to Prairie et al. (2006) for more details of the modified k-NN streamflow simulation model.

2.3. Disaggregation model

The use of disaggregation models for the distribution of flows of larger time resolution into flows with smaller time interval was first proposed by Harms and Campbell (1967). Later, Valencia and Schakke (1973) used mathematical models to divide annual flows into seasonal flows. This work has been followed by many other studies (Meja and Rousselle, 1976; Tao and Delleur, 1976; Srikanthan, 1978; Lane, 1979; Salas et al., 1980). Other disaggregation models include those proposed by Stedinger and Vogel (1984); Stedinger et al. (1985); Grygier and Stedinger (1988); Santos and Salas (1992); Bartolini and Salas (1993); Koutsoyiannis (1992), and Koutsoyiannis and Manetas (1996). This study distributed annual or monthly streamflows into daily flows using a shot noise disaggregation model (Weiss, 1977), a k-NN disaggregation model (Prairie et al., 2007), and an annual-to-daily, k-NN disaggregation model (Nowak et al., 2010).

2.3.1. Shot noise disaggregation model

The shot noise disaggregation approach was presented by Weiss (1977), based on the concept of shot noise models for the distribution of simulated monthly streamflows into daily flows. This scheme simulates the total number of pulses, s, based on Poisson distribution with parameters estimated from historical data, in which s number of pulses are assumed to occur at time t1, t2, \ldots, ts. Next, the model generates s – 1 values at interval (0, y_sim), assuming uniform distribution, and ranked according to magnitude as y1, y2, \ldots, y_s-1. Let

\[ y_n = \sum_{i=1}^{n} u_i \quad \forall \ n = 1, 2, \ldots s - 1 \]

\[ y_{\text{sim}} = \sum_{i=1}^{s} u_i \]  

(9)

The intensities of the pulses, u1, u2, \ldots, us at times t1, t2, \ldots, ts, can be determined according to Eq. (9). By applying Eq. (1) to the pulses and the previously mentioned response function (Eq. (4)), we can obtain the newly generated daily flow values.

2.3.2. k-NN disaggregation model

Prairie et al. (2007) replaced the kernel function in the method proposed by Tarboton et al. (1998) with a k-nearest bootstrap sampling scheme. Their model is employed mainly in the disaggregation of annual flows into monthly flows. This study examined the effectiveness of this method for monthly–daily flow disaggregation. The k-NN disaggregation Model uses the conditional probability function \( f(X|Y) \), in which \( X \) is originally assumed to be a matrix with dimensions of \( d \) comprising monthly flow vectors, and \( Y \) is the total of the monthly flows. In our simulation, dimensions were replaced with the number of days is each month for the purpose of monthly–daily flow disaggregation. This function is written as

\[
f(X|Y) = \frac{f(X, Y)}{\int f(X, Y) dX} \quad X_1 + X_2 + \cdots + X_d = Y
\]

X and Y are dependent. As a result, they must undergo orthogonal and unitized transformation to enable subsequent analysis. For
details on this process, please refer to Prairies et al. (2007) and Tarboton et al. (1998). The concept of disaggregation involves devising the means to divide the simulated $y_{sim}$ (annual or monthly flow) to $X$ vectors. Using the kernel-based method, high-dimensional calculations are complex and inefficient. Thus, Prairies et al. (2007) replaced the kernel-based method with the $K$-NN approach. Following a similar bootstrapping scheme, the new method compares differences between monthly flows $y_{sim}$ and historical records $y$. Taking the same weighting function as Eq. (7), it provides greater weight to the nearer neighbors and less weight to the farther neighbors during resampling. The number of samples, $k$, are suggested to be the square root of the total number of observed years in historical records. After weighting and random sampling, the total flows are converted using vectors and $y_{sim}$ is assigned to each time step as a series of simulated realizations.

2.3.3. $k$-NN disaggregation model

The annual-to-daily disaggregation model adopted for this study is also a non-parametric $K$-NN disaggregation model, similar to the monthly-to-daily $k$-NN model. It is based on the concept in Eq. (10) for the disaggregation of annual flows into daily flows. This study replaced monthly flows $Y$ with annual flows $Z$, and the condition probability is written as $f(X|Z)$, in which $Z$ denotes the annual flow vectors and $X$ is the daily flow matrix. Unlike the monthly-to-daily $K$-NN disaggregation model, the annual-to-daily model does not convert the $X$ matrix orthogonally, but merely unites the rows into vector $X^*$ with a sum of 1. In application, the observed daily flows of each year in the historical records are converted into daily flow vectors based on the proportions in the annual flow. We thus obtained matrix $P_{n \times 365}$, in which $n$ is the number of observed years in the historical records. $z_{sim}$ is the annual flow generated from linear parametric model for disaggregation. The $K$-NN approach utilizes the differences between the observed annual flows and $z_{sim}$, and similarly takes $k = \sqrt{n}$ number of nearest records as sampling candidates. The weight function in Eq. (8) is also used for the sampling of proportional vector $X^*$ of one year in the historical records. Multiplying annual simulation flow $Z$ provides the daily streamflow sequence $q$ within a year.

\[
q = z_{sim} \cdot X^* 
\]

Nowak et al. (2010) provided numerical examples as well as detailed explanations on the application of this method.

3. Model evaluation

This study compared the results of the models with historical data to evaluate their performance in the generation of synthetic flow sequences. The methods include flow duration curves (FDC), daily flow indices of simulation performance, and statistical tests between Indicators of Hydrologic Alteration (IHA) derived from historical and simulated streamflow data. FDCs provide a graphical illustration of the flow characteristics between historical and simulated streamflow data. The indices of performance associated with daily flow simulation can be used to quantify the ability of the model to reproduce the statistical features of the observed record in the generated time series. These indices assess performance of the model by deriving the difference between the model results and historical records by determining whether the hydrologic characteristics are similar. The IHAAs were employed in conjunction with statistical test methods to determine whether the IHAAs derived from simulated streamflow using different models originate from the same distribution group as the historical records and to inspect the ecological characteristics.

3.1. Flow duration curves

FDCs are a cumulative curve showing the percentage of time that streamflow is likely to equal or exceed a specified value. They combine in a single curve the flow characteristics of a stream throughout the range of discharge, disregarding the sequence of occurrence. The vertical axis of FDC measures flow discharge by day, week, month, or other duration, and the horizontal axis measures the probability that a given flow will be equaled or exceeded. Such curves indicate the percentage of time that river discharge can be expected to exceed a pre-determined flow, or likewise to show the river discharge that occurs or is exceeded for given percent of the time. As a result, when plotting an FDC, the data must be ranked and calculated according to the probabilities corresponding to each flow, before being plotted on the graph. Daily FDCs are well-suited to the illustration of river flow characteristics. In addition, FDCs present a visual representation of the relationship between the flow of natural rivers and the corresponding probability of being exceeded, thereby providing valuable information related to the flow regime. Applications using FDC are applicable to many hydrologic problems related to hydropower generation, river and reservoir sedimentation, water quality assessment, water-use assessment, water allocation, and habitat suitability.

3.2. Indices of performance

To evaluate the ability of stochastic synthetic flow models in the reproduction of statistical features associated with observed records in a generated time series, this study proposed three indices respectively based on comparisons of real and generated flow duration curves, annual maxima and annual 7-day minima statistics. Based on the distance between flow duration curves of historical records $q_i$ and synthetic flow sequences $q$, Claps et al. (2005) proposed that two indices be used to measure the reproducibility of daily flow models. One index is used to evaluate the model performance according to the ratio of the mean squared distance between the two flow duration curves to variance in the observed discharge. The other evaluates the adequacy of the model with regard to the reproduction of annual maxima statistics. This study adopted an approach similar to that of Claps et al. (2005), in which the third index was used to evaluate the reproduction of the annual 7-day minima of models, which is of greater importance in ecological applications.

\[
I_1 = 1 - s^2/\sigma^2, \quad s^2 = \frac{\sum_{i=1}^{n_f} q_0(F_i) - q(F_i))^2}{n_f} 
\]

\[
I_2 = 1 - S_{2AM}/\sigma_{2AM}, \quad S_{2AM} = \frac{\sum_{i=1}^{n_f} [q_0^{2AM}(F_i) - q^{2AM}(F_i)]^2}{k_f} 
\]

\[
I_3 = 1 - s_{7-min}^2/\sigma_{7-min}^2, \quad s_{7-min}^2 = \frac{\sum_{i=1}^{n_f} [q_0^{7-min}(F_i) - q^{7-min}(F_i)]^2}{k_f} 
\]

Assuming that the historical and generated data comprise $n_f$ number of data sets, we ranked the two data groups by magnitude. In this way, the frequency of occurrence $F_i$ of the $i$th value is \( i/n_f + 1 \), the corresponding values of historical flow are $q_{0(i)} = q(F_i)$, and simulated flow $q_{0(i)} = q(F_i)$. Eq. (12) is used to calculate the mean square distances between historical flow $q_{0(i)}$ and simulated flow $q_{i}$ under various probabilities, $s^2$. This is then compared with the variance in the original historical records, $\sigma^2$, to derive index $I_1$. The other two indices, $I_2$ and $I_3$, identify the distance between annual maxima and 7-day minima in each year, where $k_f$ denotes the
number of annual data points, \( s^2_{AM} \) is the average squared distance for the maximum annual daily flow in the historical data \( (q^2_{AM}) \) and synthetic data \( (q^2_{AM}) \), and \( s^2_{min} \) is the average squared distance for the 7-day minimum in the historical records \( (q^2_{min}) \) and synthetic data \( (q^2_{min}) \). Comparing \( s^2_{AM} \) with \( s^2_{AM} \) provides the variance in observed annual maximum flow, which can be used to calculated \( I_2 \) according to Eq. (13). Similarly, \( I_3 \) is obtained using Eq. (14), based on the variance in observed 7-day minimum flow \( s^2_{min} \) and \( s^2_{min} \). Values for \( I_1, I_2, \) and \( I_3 \) that are closer to 1 represent smaller differences between the flow duration curves and give an indication of the adequacy of the model.

3.3. IHA and statistical tests

FDCs and performance indices, as discussed in previous sections, determine the reproducibility between simulated and recorded sequences. However, FDCs examine daily flow data from a frequency analysis perspective without regard for the chronologically significant events. In this study, Indicators of Hydrologic Alterations (IHA version 7.1) (Richter et al., 1996, 1997) were used to evaluate the capacity of the synthetic flow model to capture other important hydrological characteristics. IHA is a statistical package that computes 33 parameters or indices from historic flow or stage records. It was developed by Richter et al. (1996, 1998) under the support of the Nature Conservancy. These parameters can be placed in five groups relevant to ecological and hydrological conditions: (1) magnitude of monthly water conditions, (2) magnitude and duration of annual extreme conditions, (3) timing of annual extreme conditions, (4) frequency and duration of high and low pluses, and (5) rate and frequency of changes in conditions (Richter et al., 1996). A detailed description of the influences of each IHA on the riverine ecosystem can be found in Richter et al. (1996, 1998).

To evaluate the reproducibility of these hydrological characteristics in synthetic flow models, we compared each IHA parameter series derived from historical records for the generation of simulated flow data, and applied statistical tests to determine whether they came from the same population. If a significant difference exists between two instances of an IHA parameter (e.g. monthly average), the model is recognized as unable to capture the corresponding features, and vice versa. We applied the Wilcoxon rank-sum test and the Kolmogorov–Smirnov test to evaluate the performance. Both tests are nonparametric approaches comparing differences between two samples in order to determine whether the two data sets differ significantly. These tests are more robust and require no assumptions regarding the distribution of data, unlike a t-test.

4. Case study

A case study was performed using flow records from the Tamsui River watershed in northern Taiwan. The basin is approximately 2726 square kilometers in area with a subtropical marine monsoon climate. It encompasses three major tributaries, the Dahan River, the Hsindian River, and the Keelung River. The average annual rainfall ranges between 2200 mm and 2800 mm, concentrated between May and August. In Taiwan, the major rainy seasons are generally divided into the mei-yu season (mid-May to mid-June) and the typhoon season (mid-July to August). During typhoon season, tropical storms frequently strike Taiwan with very heavy resulting rainfall. The high intensity of this precipitation is the primary cause of weather-related disasters in Taiwan.

This region is the politico-economic center of Taiwan and therefore has relatively complete records of streamflow discharge. This study selected six streamflow-gauging stations with at least 15 years of daily streamflow records in different locations through-out the Tamsui River watershed. These stations included 1140H041 (Hsiuluan), 1140H058 (Wudu), 1140H066 (Hsiulang), 1140H067 (Sanyinciao), 1140H078 (Jieshouciao), and 1140H082 (Baociao). The purpose of this study was to test the capacity of models to synthesize a time series; therefore, station records were considered independently for the sake of simplification. Table 1 and Fig. 2 present relevant information related to the selected stations and their geographical locations.

In accordance with the framework proposed in Fig. 1, we analyzed the hydrological characteristics based on historical records and generated simulated flow data using daily flow data with each of the models for the selected stations. Seven synthetic sequence models were adopted, as outlined in Section 2 and Table 2. To avoid the statistical complexity resulting from differences in population size, this study used historical observations representing the same length of time for each of the hydrological sequence models. Despite these precautions, we generated the sequences for each station ten times to ensure stochastic variation within the model. Taking the Hsiulang Station as an example, Table 1 lists the records spanning 34 years from 1957 to 2003. Thus, we generated 34-year flow sequences using each model through ten iterations for comparison. Using the comparison methods outlined in Section 3, the statistical and hydrological characteristics of the generated and historical daily flow rates were compared with regard to various aspects of reproducibility.

5. Results and discussion

We first compared the FDCs of observation and simulation data to evaluate the reproducibility of different models in different stations. Overall, these FDCs demonstrated the ability of most of the models to simulate the relationship between the frequency and magnitude of streamflow. Plotting the FDCs on a linear scale would reveal no obvious differences; therefore, we plotted the FDCs along the horizontal axis in probability-scale to illustrate the occurrences of high and low probability. Along the vertical axis (discharge), a linear scale was used for flow exceeding 20 cms, while flow below 20 cms was converted to a logarithmic scale for better resolution of low flow conditions. These measures were taken to illustrate differences among the models.

Fig. 3 illustrates the FDCs for six streamflow-gauging stations. Among the various approaches, it was observed that models incorporating SAMS for the generation of annual flow and k-NN annual to daily disaggregation (SAMS & annual kNN) consistently outperformed the other models. At most of the sites, the FDCs derived through simulation coincided closely with those obtained through historical observation. At the stations with the major portion of FDCs below 20 cms (e.g. 1140H041 and 1140H082), the difference seems more distinct visually for low flow condition, but actually as pronounced as they appeared on the graph. We can conclude that this model shows very high reproducibility in the simulation of discharge flow across a range of frequencies. However, it should be noted that for this model, the randomness was only generated.

<table>
<thead>
<tr>
<th>Station name</th>
<th>Station no.</th>
<th>Record period</th>
<th>Record length</th>
<th>Tributary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hsiuluan</td>
<td>1140H041</td>
<td>1957–2003</td>
<td>47</td>
<td>Dahan River</td>
</tr>
<tr>
<td>Wudu</td>
<td>1140H058</td>
<td>1966–1999</td>
<td>34</td>
<td>Keelung River</td>
</tr>
<tr>
<td>Hsiulang</td>
<td>1140H066</td>
<td>1966–1999</td>
<td>34</td>
<td>Hsindian River</td>
</tr>
<tr>
<td>Jieshouciao</td>
<td>1140H078</td>
<td>1981–2000</td>
<td>20</td>
<td>Keelung River</td>
</tr>
<tr>
<td>Baociao</td>
<td>1140H082</td>
<td>1987–2002</td>
<td>16</td>
<td>Jingmei River</td>
</tr>
</tbody>
</table>
Table 2
A summary of daily streamflow synthetic models and their abbreviations in the application.

<table>
<thead>
<tr>
<th>No.</th>
<th>Abbreviations</th>
<th>Model description</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Shot noise mode</td>
<td>Shot noise daily streamflow simulation model</td>
</tr>
<tr>
<td>II</td>
<td>Markov-based model</td>
<td>Markov-based daily streamflow simulation model</td>
</tr>
<tr>
<td>III</td>
<td>SAMS &amp; SN</td>
<td>Linear parametric monthly streamflow simulation model using PARMA by SAMS with shot noise monthly–daily disaggregation model</td>
</tr>
<tr>
<td>IV</td>
<td>SAMS &amp; kNN</td>
<td>Linear parametric monthly streamflow simulation model using PARMA by SAMS with k-NN monthly–daily disaggregation model</td>
</tr>
<tr>
<td>V</td>
<td>M-kNN &amp; SN</td>
<td>Modified k-NN monthly flow simulation model with shot noise monthly–daily disaggregation model</td>
</tr>
<tr>
<td>VI</td>
<td>M-kNN &amp; kNN</td>
<td>Modified k-NN monthly flow simulation model with k-NN monthly–daily disaggregation model</td>
</tr>
<tr>
<td>VII</td>
<td>SAM &amp; Annual-kNN</td>
<td>Linear parametric annual streamflow simulation model using ARMA by SAMS with k-NN annual–daily disaggregation model</td>
</tr>
</tbody>
</table>

Fig. 2. Location of 6 streamflow stations used in this study in the Tamsui watershed, Northern Taiwan.
from total annual discharge, then the sequence resampled exactly from unitized daily streamflow vector in the historical records. With relatively low degree of freedom, the excellent reproducibility of this model in the simulation of streamflow is not surprising. The range of freedom can be expanded through the use of models simulating monthly streamflow and then disaggregating to daily flow. This study adopted two monthly flow simulation models (linear parametric model PARMA by SAMS and non-parametric modified k-NN), with two monthly–daily disaggregation models (shot noise disaggregation and k-NN disaggregation), resulting in the following four combinations: SAMS & SN, SAMS & kNN, M-kNN & SN, and M-kNN & kNN. In a comparison of the FDCs of these models, no significant differences were observed in the results from the monthly flow simulation. The FDCs from SAMS & SN and M-kNN & SN were similar, as were SAMS & kNN and M-kNN & kNN. It could be inferred that the PARMA and modified k-NN models are equivalent in the simulation of monthly discharge for disaggregation.

Unlike monthly flow models, the choice of disaggregation model could alter the characteristics of the simulation. Under extremely high-flow conditions (1% exceedence probability or less), a shot noise model generated less flow than did a k-NN-based disaggregation model in all cases except 1140H067. Under medium and low flow conditions, streamflow simulated by the k-NN disaggregation model was relatively high, compared to shot noise disaggregation. According to FDC, it was observed that either of the monthly flow simulation models coupled with k-NN disaggregation (SAMS & kNN and M-kNN & kNN) could be the next-best option comparing to SAMS & annual kNN. However, when coupled with shot-noise disaggregation, streamflow could be underestimated. The processes of these two approaches reveal that k-NN disaggregation is able to maintain mass conservation only from the magnitude of the pulse. Following transformation by the response function, the recession might extend for the subsequent month or longer. As a result, the flow amount within a monthly interval cannot only be exactly maintained. In addition, the discretization of the response function to daily units and numerical errors in the computation of convolution may also lead to difficulties in maintaining the principle of mass conservation. In the shot noise model, the low flow or baseflow is a composite of flow recession from numerous pulses in advance. As a result, even small approximation errors can accumulate/propagate monotonically with respect to the number of pulses in the convolution. As a result, shot noise disaggregation always produces simulations with relatively low flow rates.

Direct daily simulation approaches, shot noise and Markov-based synthetic models each have specific concerns. We can observe from the results that shot noise simulation models provide relatively high low-flow, compared to other methods. Comparing to shot noise disaggregation, the difference between them mainly lies in the schemes of pulse generation. While shot noise disaggregation model is based on the assumption of uniform distribution to determine the magnitude of pulses, direct shot noise simulation model applies exponential distribution for the same task. As a result, even with the similar statistics patterns (ex, approximative mean of monthly flow), shot noise simulation model with exponential distribution assumption produces more pulses of low magnitude so that it results in relatively higher low-flow conditions. Markov-based models tend to overestimate under high-flow conditions because they determine whether the day is dry or wet before determining the magnitude. For sequence of wet days, Markov-based models assume that the increase in flow each day is independent and calculates the magnitude according to distribution fitted from historical data. However, for periods of high-flow, the correlation in flow rates from one day to the next can be disregarded. For example, when a very large increase in flow is observed on one day, the probability that the following day will present another increase is diminished. In this method, this situation is referred to as “wet conditions”. Markov-based models are limited...
to describing this type of relationship and therefore tend to overestimate high-flow discharge to a certain degree.

Table 3 lists the results of $I_1$ for all stations with each of the various methods. An index was used to quantify actual and generated FDCs. As observed in a comparison of FDCs, the use of SAMS to generate annual flow in conjunction with k-NN for annual to daily disaggregation (SAMS & annual kNN) provides results superior to all other models. It has the highest the averaged $I_1$ for SAMS & annual kNN of six stations reaches among all models. Regarding to monthly simulation model with monthly to daily disaggregation, k-NN disaggregation model seems have better performance comparing to shot noise disaggregation model, either based on flows generated from linear regression or modified k-NN monthly models. Integrating monthly simulation with shot noise disaggregation might be the least recommended among all methods proposed in this study, particularly due to the accumulation of numerical errors resulting from convolution using this framework. Despite the fact that no significant differences were observed in the FDCs between linear regression and modified k-NN monthly model, the results of $I_1$ show that the modified k-NN approach may be slightly better in reproducing the magnitude and frequency of average daily streamflows. Nonetheless, the difference in $I_1$ was very small (0.01–0.02). For the direct simulation of daily flow, both Markov-based and shot-noise models demonstrate good reproducibility in FDCs, as indicated by $I_1$, producing both over- and under-estimation for low- and high-flow conditions. Nonetheless, the probability of exceedance under high-flow conditions and the magnitude of low-flow are very small. As a result, estimation flaws cannot be appropriately reflected in $I_1$. In a comparison of different stations, it was found that 1140H058 and 1140H066 had relatively higher average $I_1$, probably due to their location further downstream. The flow discharge in these areas is greater and conditions tend to be more tranquil, making it easier for simulation. In contrast, 1140H041 is located in the high upstream portion of the watershed; therefore, the average flow in the area was the smallest among all of the stations.

Table 4 presents the results for $I_3$, the index used to evaluate the simulation of annual maxima. The results show that the use of annual flow in conjunction with k-NN annual to daily disaggregation (SAMS & annual kNN) performed exceedingly well in the simulation of annual maxima. In the use of monthly simulation models, it appears that the modified k-NN approach outperforms linear models in capturing maximum flow conditions, particularly when incorporated with k-NN disaggregation. When coupled with a shot noise aggregation model, the advantage did not hold. Neither the linear model nor the modified k-NN appeared able to capture the annual maximum flow conditions effectively and daily simulation models are unsuited to the reproduction of annual maxima. Nonetheless, the Markov-based model still appears to be somewhat more effective than the shot noise model. It appears that the annual maxima characteristics of 1140H058 and 1140H066 were the easiest to model, perhaps due to stable high-flow conditions and historical records dating back over a relatively long period of time. Nonetheless, despite a long record of historical data, station 1140H041 is located in a far upstream area of the watershed resulting in far greater variability in flow, such that simulating annual maximum flow conditions could be challenging under any circumstances. The results from downstream station 1140H067 also were not very good, perhaps due to the flood-control operations of the Shihmen reservoir interfering with the natural flow, thereby increasing the difficulty of simulating maximum flow conditions.
For the annual maxima, the Markov-based daily values are closer to 1. This implies the use of a stochastic model on historical and synthetic flow data did not present significant differences. As a result, this relatively large p-value is less meaningful. Thus, for each IHA parameter used to evaluate the reproducibility of the models, this study used the acceptance ratio obtained from 60 simulation series (10 sequences x 6 stations) that did not present a significant difference from the historical data. The p-value of each IHA is provided in the appendix.

Table 5 presents the acceptance ratio that failed to reject the null hypothesis for group 1 IHA parameters related to the magnitude of monthly water conditions with a subtotal of 12 parameters. From the results, it was observed that the modified k-NN monthly flow model with k-NN monthly–daily disaggregation model (M-kNN & kNN) resulted in the highest acceptance ratio. The SAMS model with annual disaggregation (SAMS & Annual-kNN) or monthly k-NN disaggregation (SAMS & k-KNN) provided the second best results with regard to the magnitude of monthly water conditions. Both daily simulation approaches reproduced patterns in the monthly flow conditions reasonably well. However, regardless of which monthly simulation models were used, the coupling shot noise disaggregation model tended to bias the magnitude of monthly flow.

Both of the statistical tests used in this study, Wilcoxon Rank-Sum test (WR) and Kolmogorov–Smirnov test (KS), provided similar results with regard to model reproducibility. These two tests had the same assumptions and the same null hypothesis; however, their alternative hypotheses were different. The WR test is a non-parametric test that did not assume a specific distribution of the data. The KS test is a more general instrument designed to detect differences in both the location and shape of the distributions, based on the maximum absolute difference between the cumulative distribution functions.

For regression appears to perform far better than the shot noise aggregation model, from the perspective of 7-day low-flow regime. Un-aggregation appears to perform far better than the shot noise aggregation model. In a comparison of these approaches, the models that adopt exponential recession curves (SAMS & SN, M-kNN & SN, Markov-based and Shot noise daily simulation models) are less able to reproduce 7-day minimum flow regimes. Three of the downstream stations, 1140H041, 1140H058, and 1140H066, produced higher I3 than the other upstream stations, 1140H041, 1140H078, and 1140H082. A close examination of the data revealed that the 7-day minimum flow in the downstream stations included greater annual variation due to flow concentration, as well as a corresponding high variance in observed annual 7-day minimum, which resulted in I3 values closer to 1. This implies the use of a stochastic framework is not necessarily sufficient for the simulation of low-flow regimes due to a relatively small variation/uncertainty.

This study also compared each IHA parameter derived from historical records and simulated flow series using the Wilcoxon rank-sum test (WR) and Kolmogorov–Smirnov test (KS). The null hypothesis is that these two populations (IHA from historical and synthetic flow data) have the same distribution. This comparison helped to reveal the reproducibility of various flow patterns. In both WR and KS tests, this study adopted a significance level of 0.05 to reject the null hypothesis. Most of the IHA parameters based on historical and synthetic flow data did not present significant differences. As a result, this relatively large p-value is less meaningful. Thus, for each IHA parameter used to evaluate the reproducibility of the models, this study used the acceptance ratio obtained from 60 simulation series (10 sequences x 6 stations) that did not present a significant difference from the historical data. The p-value of each IHA is provided in the appendix.

Table 6 presents the acceptance ratio that failed to reject the null hypothesis for group 1 IHA parameters related to the magnitude of monthly water conditions with a subtotal of 12 parameters. From the results, it was observed that the modified k-NN monthly flow model with k-NN monthly–daily disaggregation model (M-kNN & kNN) resulted in the highest acceptance ratio. The SAMS model with annual disaggregation (SAMS & Annual-kNN) or monthly k-NN disaggregation (SAMS & k-KNN) provided the second best results with regard to the magnitude of monthly water conditions. Both daily simulation approaches (Markov and shot noise) reproduced patterns in the monthly flow conditions reasonably well. However, regardless of which monthly simulation models were used, the coupling shot noise disaggregation model tended to bias the magnitude of monthly flow.

Both of the statistical tests used in this study, Wilcoxon Rank-Sum test (WR) and Kolmogorov–Smirnov test (KS), provided similar results with regard to model reproducibility. These two tests had the same assumptions and the same null hypothesis; however, their alternative hypotheses were different. The WR test is a popular two-independent-samples test used to examine whether the same distribution is present but has been shifted to some degree. Observations are combined from two groups to assess whether the mean ranking of their populations differs. The KS test is a more general instrument designed to detect differences in both the location and shape of the distributions, based on the maximum absolute difference between the cumulative distribution functions.
observed in both samples. When this difference is large, the two distributions are considered different. Under IHA group 1 parameters, only a slight disagreement was observed between SAMS & kNN and SAMS & Annu-l-kNN models with regard to the goodness of simulation. Compared to SAMS & kNN, SAMS & Annu-l-kNN provided a smaller difference in maximum magnitude in monthly flow but larger differences in distribution.

In the case study, the nonparametric monthly simulation model appeared to perform more effectively in dealing with monthly patterns. According to a comparison of FDC curves and related indices, SAMS & Annual-kNN proved to be the most powerful scheme with regard to the reproduction of flow conditions. However, the performance of this approach was not as strong under group 1 IHA parameters (0.90 for WR and 0.91 for KS). One possible reason may be the fact that the flow series pattern utilized by SAMS & Annual-kNN may be strongly correlated to the magnitude of annual flow. For example, years with high annual flow may have relatively high flow conditions during flooding season but a negligible increase in discharge during the dry season. Considering these two factors independently may result in a bias in monthly flow simulation. FDC is unable to illustrate this concern because it lumps everything together; however, a comparison of a series of IHA parameters reflects this problem effectively.

Markov-based models appear to be slightly better than the shot noise model with a higher acceptance ratio in the KS test (0.89) compared to the shot noise model (0.84) with no obvious difference in the WR test (0.89 vs. 0.88). From these results, we can infer that Markov-based models may possess better distributional similarity in monthly flow. Although the monthly simulation models with k-NN monthly–daily disaggregation outperform daily simulation models, shot noise disaggregation does not produce good results. Theoretically, using the same monthly flow simulation model (SAMS and modified k-NN in this study) should produce the same results under monthly flow conditions from IHA group 1. However, the nonparametric modified k-NN model presents a higher acceptance ratio in monthly flow, but only when combined with the k-NN disaggregation model. The SN disaggregation model raises concerns regarding mass conservation. As a result, the acceptance ratio (0.79 for WR, KS) is relatively low among the various approaches. In the use of monthly simulation models, it appears that the modified k-NN approach is better able to capture maximum flow conditions, compared to linear models incorporated with the k-NN disaggregation model.

In a comparison of IHA group 1 parameters, it was found that the average values for monthly flow were relatively low during the dry season (Dec. to May). The low acceptance ratio can be attributed more to shot-noise disaggregation approach, and also partly to Markov-based and shot-noise models. All of these models adopt exponential recession curves, which do not necessarily reflect long-term low flow conditions such as baseflow. Perhaps this disadvantage only exists in the study area only. Overall, the simulation of monthly magnitude for high flow season poses fewer challenges. In the low flow season, a nonparametric approach may outperform other methods.

Table 7 lists the results under IHA group 2 parameters, representing extreme water conditions. These parameters include 1-, 3-, 7-, 30-, and 90-day minimums and maximums from moving averages of an appropriate length calculated within the water year. Different IHA group 1, for this group, SAMS & Annual-kNN has best simulation as demonstrated by high acceptance ratio (0.99). Monthly models with k-NN disaggregation (M-kNN & kNN and SAMS & kNN) produce the next best results, followed by Markov-based models. Shot noise approaches, including the shot noise model, SAMS & SN, and M-kNN & SN, did not perform as well as the other approaches. The WR test and KS test both produced the same ranking in the model acceptance ratio, but with small differences in values. These results suggest that the evaluations are the same, when viewed from the perspective of either maximum difference or distributional similarity.

The non-parametric models (M-kNN & kNN, SAMS & kNN, and SAMS & Annual-kNN) did not differ greatly in the acceptance ratio with respect to duration, for either minimum or maximum flow. Using the other four models, the acceptance ratio of minimum flow increased with its duration. In contrast, with longer duration, models’ simulation in maximum flow resulted in a lower acceptance ratio. This finding implies that longer durations are easier to simulate under low flow conditions than under high flow conditions. All models provide a reasonable illustration of 1-day maximum flow, with a sufficiently high acceptance ratio (0.85–1.00). These results are consistent with those in the $L_2$ index, which also exhibits the annual maxima of daily flow. The baseflow parameter is a 7-day minimum flow over mean annual flow. Although the annual mean is not a serious concern, this parameter is similar to the
Table 8
The acceptance ratio of IHA group 3–5 parameters from historical and synthetic flow data recognized as the same distribution.

<table>
<thead>
<tr>
<th>IHA parameters</th>
<th>Group 3</th>
<th>Group 4</th>
<th>Group 5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Date min</td>
<td>Date max</td>
<td>Average</td>
</tr>
<tr>
<td>(a) Wilcoxon rank-sum test</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M-4NN &amp; kNN</td>
<td>0.90</td>
<td>1.00</td>
<td>0.95</td>
</tr>
<tr>
<td>SAMS &amp; kNN</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
</tr>
<tr>
<td>SAMS &amp; Annual-4NN</td>
<td>0.92</td>
<td>0.93</td>
<td>0.93</td>
</tr>
<tr>
<td>Markov-based model</td>
<td>0.98</td>
<td>0.97</td>
<td>0.98</td>
</tr>
<tr>
<td>Shot noise model</td>
<td>0.95</td>
<td>0.95</td>
<td>0.93</td>
</tr>
<tr>
<td>SAMS &amp; SN</td>
<td>0.95</td>
<td>0.98</td>
<td>0.97</td>
</tr>
<tr>
<td>M-4NN &amp; SN</td>
<td>0.92</td>
<td>0.95</td>
<td>0.93</td>
</tr>
<tr>
<td>Average</td>
<td>0.94</td>
<td>0.97</td>
<td>0.95</td>
</tr>
<tr>
<td>(b) Kolmogorov-Smirnov test</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M-4NN &amp; kNN</td>
<td>0.83</td>
<td>1.00</td>
<td>0.92</td>
</tr>
<tr>
<td>SAMS &amp; kNN</td>
<td>0.92</td>
<td>1.00</td>
<td>0.96</td>
</tr>
<tr>
<td>SAMS &amp; Annual-4NN</td>
<td>0.92</td>
<td>0.90</td>
<td>0.91</td>
</tr>
<tr>
<td>Markov-based model</td>
<td>0.95</td>
<td>1.00</td>
<td>0.98</td>
</tr>
<tr>
<td>Shot noise model</td>
<td>0.75</td>
<td>0.90</td>
<td>0.82</td>
</tr>
<tr>
<td>SAMS &amp; SN</td>
<td>0.73</td>
<td>1.00</td>
<td>0.87</td>
</tr>
<tr>
<td>M-4NN &amp; SN</td>
<td>0.73</td>
<td>0.90</td>
<td>0.82</td>
</tr>
<tr>
<td>Average</td>
<td>0.82</td>
<td>0.96</td>
<td>0.89</td>
</tr>
</tbody>
</table>

6. Conclusions

This study investigated the reproducibility of daily synthetic flow models, including daily synthetic flow simulation models and monthly/annual simulation models coupled with disaggregation. Graphical FDCs, quantified indices, and statistical tests based on IHA parameters were used to compare differences between daily synthetic flow simulations and historical observations. Despite the convenience of FDCs as a graphical summary that illustrates river flow (under both high and low flow conditions), quantified indices (based on FDCs) tend to be more robust in the evaluation of simulation quality. However, neither FDCs nor indices provide information regarding the sequential relationships of flow, which may be important with regard to ecological concerns. This study also examines IHA parameters to determine the simulation capacity of models with respect to various flow patterns. Most of the models were effective in simulating the magnitude of monthly flow and high flow within short durations. However, low flow conditions remain problematic, particularly over short durations. Non-parametric disaggregation models based on historical flow patterns proved more effective in capturing fluctuations in flow patterns (high/pulse numbers, rise/fall rates, reversals); however, these patterns are still difficult to illustrate using simulation models.

Among the various models tested in this study, the approach involving annual flow simulation with annual–daily disaggregation demonstrated the good reproducibility with high value in most indices. From the perspective of duplicating patterns in flow regions, this may be the most powerful option. However, one purpose of synthetic flow modeling is to represent realistic future scenarios. In this model, only the total annual flow provides room for flexibility, which may be fine for long-term planning with a focus on inter-year variation (e.g. climate change simulation over decades); however, its applicability with regard to within year fluctuations (e.g. water management or ecological engineering) is doubtlessly limited. Linear regression models (such as PARMA) and nonparametric models (modified k-NN) are well suited to the representation of monthly streamflow. In conjunction with nonparametric bootstrap disaggregation schemes, they provide a good option for the generation of synthetic flow series capable of representing daily flow patterns with reasonable fidelity. This study determined that both of the daily simulation models tested...
data obtained through case studies, such that under different hydrological conditions, the performance of these models may differ from that observed in this study. This issue would be worthy of further investigation through the application of data from various regions around the world. Nonetheless, this is beyond the scope of our present work; therefore, we leave it for further study.

Acknowledgement

The authors wish to acknowledge the financial support for this paper provided by the National Science Council, Taiwan (NSC 101-2221-E-002-152- and NSC 102-2221-E-002-142-MY2).

Appendix A

For each IHA parameter used to evaluate the reproducibility of the models, this study used the acceptance ratio obtained from 60
simulation series that did not present a significant difference from the historical data. The p-value of the Wilcoxon rank-sum test (WR) and Kolmogorov–Smirnov test (KS) for IHA of group 1, 2 and 3–5 is provided in Tables A1–A3 respectively. The p-value is the probability of obtaining a test statistic at least as extreme as the one that was actually observed, assuming that the null hypothesis is true. However, it should be noticed the p-value is not the probability that the null hypothesis is true.

References


