Evaluation of precipitation forecasts from NOAA global forecast system in hydropower operation

Huicheng Zhou, Guolei Tang, Ningning Li, Feng Wang, Yajun Wang and Deping Jian

ABSTRACT

Forecasts of 10-day average inflow into the Ertan hydropower station of the Yalong river basin are needed for seasonal hydropower operation. Medium-range inflow forecasts have usually been obtained by Auto-Regressive-Moving-Average (ARMA) models, which do not utilize any precipitation forecasts. This paper presents a simple GFS-QPFs-based rainfall–runoff model (GRR) using the 10-day accumulated Quantitative Precipitation Forecasts from the Global Forecast System (GFS-QPFs) run at the American National Oceanic and Atmospheric Administration (NOAA). In this study, 10-day accumulated GFS-QPFs over the Yalong river basin are verified by first using a three-category contingency table. Then this paper presents the results from a proposed hydrological model using 10-day accumulated GFS-QPFs. Results show that inflow forecast errors can be reduced considerably, compared with those from the currently used ARMA model by both quantitative and qualitative analysis. Finally, simulations of medium-range hydropower operation are also presented using the historical data and forecasts of 10-day average inflows into the Ertan dam during May to September 2006 to evaluate the efficiency of the proposed hydrological model using the GFS-QPFs. The simulations demonstrate that the use of GFS-QPFs has improved reservoir inflow predictions and hydropower operation of the Ertan hydropower station in the Yalong river basin during the wet season.

Key words | global forecast system, hydropower operation, inflow forecasting, quantitative precipitation forecasts

INTRODUCTION

Forecasts of 10-day average inflow into the hydropower stations of the Yalong river basin are needed for the medium-range hydropower operation, to bring benefits from a reduction in flood damage, increased dam safety and greater efficiency in power generation. The currently-used forecasting models for reservoir inflow do not include any information on weather or climate forecasts, but are based on Periodic Auto-Regressive (PAR) or Auto-Regressive Moving Average (ARMA) models (Maceira et al. 1999; Han et al. 2007). Recently, medium-range Quantitative Precipitation Forecasts (QPFs) over the forthcoming 10-day periods are addressed as input variables to improve inflow forecasts coupling hydrological models for reservoir inflow predictions (Collischonn et al. 2007). An increasing number of results suggest that progress is being achieved in bringing QPFs to the stage of operational usefulness for hydrological applications (Collier & Krzysztofowicz 2000; Damrath et al. 2000; Golding 2000; Mao et al. 2000; McBride & Ebert 2000; Hollingsworth 2003; Habets et al. 2004; Collischonn et al. 2005, 2007), while other authors highlight the limitations of these forecasts in the case of extreme events (Meneguzzo et al. 2004; Bartholmes & Todini 2005) as the QPFs are highly uncertain.
The use of QPFs obtained by numerical weather prediction models as input data to run hydrological rainfall–runoff models, thereby obtaining extended streamflow forecasts, has been explored by several authors who in general conclude that QPFs are useful, although their usefulness was limited by their large uncertainty (Yu et al. 1999; Ibbitt et al. 2000; Anderson et al. 2002; Jasper et al. 2002; Koussis et al. 2003; Habets et al. 2004; Collischonn et al. 2005, 2007). Collischonn et al. (2007) present the results of medium-range reservoir inflow predictions by the use of a large-scale hydrological model applied to a part of the Paraiba river basin using precipitation forecasts from the regional Era model run by the Brazilian Center for Weather Prediction, and the results show forecast errors can be reduced considerably during both wet and dry seasons, compared with those from the ARMA model. There have been recent attempts to consider the uncertainty in forecasts, using ensemble rainfall forecasts (Bartholmes & Todini 2005; Goweleeuw et al. 2005; Zhang et al. 2006) and to combine the inherent uncertainty of hydrological models with ensemble forecasts (Pappenberger et al. 2005).

Krzyżtofowicz & Henry (2001) present a hydrologic uncertainty processor (HUP), which produces a short-term probabilistic river stage forecast based on a probabilistic quantitative precipitation forecast. However, most of these results are from work that is still at the research stage, since operational forecasting systems still rely more on radar estimates and telemetry of measured rainfall or short-range nowcasting (Sivapragasam et al. 2001; Tsanis & Gad 2003; Li & Laim 2004; Qiu et al. 2004; Yu et al. 2004; Moore et al. 2005; Wang et al. 2005; Yuan et al. 2008). Nevertheless, QPFs are gradually being introduced into operational streamflow forecasting systems in an attempt to extend the range of forecasts (Moore et al. 2005; Bremicker et al. 2006), but whether the extended streamflow forecasts that are obtained could serve for hydropower dispatching needs to be further explored on a case-by-case basis (Hamlet & Lettenmaier 2000; Hettiarachchi et al. 2005).

The Global Forecast System (GFS), run by the American National Oceanic and Atmospheric Administration (NOAA), has made quantitative precipitation forecasts (QPFs) up to 16 days at each data assimilation cycle (00, 06, 12 and 18 UTC), and the QPFs are made available for free over the internet. However, the use of QPFs obtained by the GFS (termed GFS-QPFs) as input into rainfall–runoff models is relatively undeveloped for medium-range reservoir inflow predictions. So the GFS-QPFs are evaluated in this study to explore the potential improvements in medium-range reservoir inflow predictions and hydropower production comprehensively. In this paper, taking the Ertan hydropower station located in the Yalong river basin as an example, forecasts of 10-day accumulated precipitation are verified objectively by comparing them with the corresponding observed rainfall, at first with a three-category contingency table showing the frequency of forecasts and observations in the various bins. Then the use of the 10-day accumulated GFS-QPFs as input data into a simple hydrological model is presented for forecasting the inflow into the reservoir, and the results from the proposed hydrological model will be compared with forecasts obtained by the currently used Auto-Regressive (AR) model both quantitatively and qualitatively. Finally, simulations for medium-range hydropower operation will be presented to evaluate the efficiency of reservoir inflow forecasting model using the GFS-QPFs.

**STUDY SITE AND FORECASTING METHODOLOGY**

The Ertan hydropower station

The case study reported in this paper is concentrated on the Ertan hydropower station located in the lower reaches of the Yalong river basin in the Sichuan province of southern China. Figure 1 shows the location of the Ertan dam, the main features of the basin and gauging stations. As seen in Figure 1, this part is relatively well covered with rain gauges and stream gauges, which transmit rainfall and flow at 6-h intervals in real time. The Yalong river basin lies on the eastern edge of the Tibetan plateau that covers 26°32’–33°58’N and 96°52’–102°48’E. The mean annual rainfall in catchment ranges is between 500 and 2470 mm and it is higher in the south and east. The area has two distinct seasons: dry and wet season, as its climate is mainly influenced by high-altitude westerly circulation and the southwest monsoon. Generally, this basin has plenty of rain, and about 90–95% of the annual rainfall in the wet season (from May to September), so there is a large possibility for hydropower...
improvement in that season. However, there is less rainfall with only 5–10% of the annual rainfall, so the runoff is relatively stable in the dry season (from October to April of the following year).

The Ertan hydropower station is one of the key power sources for the Sichuan electric network with an annual hydropower generation of 16,880 GWh. It has a dead storage capacity of 2,430 Mm$^3$ with a dead storage level (DSL) of 1,155 m, a gross storage capacity of 5,800 Mm$^3$ with a normal pool level (NPL) of 1,200 m and an active storage capacity of 3,370 Mm$^3$. The installed capacity of power generation is 3,500 MW through its power generation system composed of six hydrogenerators with an installed capacity of 550 MW each. The firm capacity committed for the Ertan reservoir is 1,028 MW. Statistical analyses of historical hydropower operation data show that the wet season can be roughly divided into two parts: the delivery period from early May to mid-June and the storage period from late June to late September. During the delivery period, the Ertan is running mainly for water supply, so the operating water level of the Ertan should be lower; while during the storage period, the Ertan will run primarily for water storage and the corresponding water level should be kept higher. More features of the Ertan station include the mass balance equation for reservoir storage and inflow, releases from the reservoir, plant capability, turbine capacity and power production; details can be found in Wang (2008).

Forecasts of 10-day average inflow into the Ertan reservoir are needed for seasonal operational hydropower operation and obtained from the model in current operational use, known as PREVIVAZ (Maceira et al. 1999; CEPEL 2004). PREVIVAZ is a program that uses several different configurations of ARMA$(p, q)$ and periodic ARMA models, with values for $p$ in the range 1–4 and $q$ not larger than 1. Different configurations of the model are tested every 10 days before the forecast is issued, so as to identify the best amongst all possible model configurations. The procedure is repeated every 10 days, but the QPFs are not used. The average absolute error (ABE), the average relative error (ARE), the root mean square error (RMSE) and Nash–Sutcliffe efficiency (NSE) are the statistical parameters (Sahai et al. 2000) used to describe the accuracy of inflow forecasting models. From the statistics of forecasts, the ARMA model can be considered as an acceptable one during the dry season, but not during the wet season. So the main objective of this study is to obtain forecasts of 10-day average inflows into the Ertan dam every 10 days from May to September during the wet season.

Quantitative precipitation forecasts

The Global Forecast System run by the NOAA is a global Numerical Weather Prediction (NWP) computer model, which produces QPFs up to 16 days ahead at each data
assimilation cycle (00, 06, 12 and 18 UTC), but with decreasing spatial and temporal resolution over time. The NWP model is run in two parts: the first part has a higher resolution and goes out to 180 h in the future; the second part runs from 180–384 h at a lower resolution. The resolution of the model varies in each part of the model: horizontally, it divides the surface of the Earth into 35 or 70 km grid squares; vertically, it divides the atmosphere into 64 layers and temporally it produces a forecast for every third hour for the first 180 h; after that, they are produced for every twelfth hour. The GFS is the only global NWP model for which all output including QPFs in the GRIB1 format (which has been updated to GRIB2 since 12 February 2008) is available on the NOAA FTP servers for free over the internet, and as such is the basis for non-state weather companies, e.g. Wunderground.com, Weatheronline.co.uk, Weather.com.au and t7online.com (GRIB 2008).

The GFS-QPFs over the continental United States (CONUS) have been evaluated by different measures including Equitable Threat Score (ETS), True Skill Statistics (TSS) and Bias (BI). The statistics show that the precipitation forecast has more skill in winter (December, January and February) when comparing to summer (June, July and August), and the skills have improved gradually for the past several years when the model increases the resolution and improves the analysis system and physical processes (Zhu 2007). The ETS seems to be a good estimate for overall forecast skill and has a range of –1/3 to 1. The higher the value of ETS, the better the forecast model skill is for that particular threshold. The TSS measures the ability of the forecasts to discriminate between “Yes” and “No” observations based on contingency tables (Doswell et al. 1990). It ranges from –1.0 (no correct forecasts) to 1.0 (perfectly correct). The Bias measures the relative frequency of predicted and observed rainfall. The best model is generally the one that remains near the 1.0 line. If the model verifies over 1.0, it is over-predicting precipitation, and if it is below 1.0 it is under-predicting precipitation. Figure 2 shows the objective scores of GFS precipitation skills over CONUS from Zhu (2007) for 00 UTC of August 15, 2006–18 UTC of November 6, 2006 and 00 UTC of November 1, 2006–18 UTC of February 5, 2007. Three different forecasts (12–36 h, 36–60 h and 60–84 h, leading time) have been verified by calculating ETS, TSS and BI. The x axis is 24-h threshold precipitation amounts in mm. The numbers above the x axis are total observed grids/boxes in the verified period for that threshold. The BIs (Figure 2(a)) are very similar at a range of 0.9 (10% under-forecast) to 2.0 (100% over-forecast) for all lead time forecasts, while the TSS and ETS are reasonably decreasing their skills when increasing lead time. There is the same future for Figure 2(b) except the BI are less, and the TSS and ETS are larger. A more comprehensive description of GFS, including model parameters and its implementation changes, precipitation maps and evaluation documentation, is given on the official website of NOAA.

The GFS-QPFs used in this study are the modeled precipitation forecasts over the entire East Asia region. A precipitation map published at 00 UTC September 15, 2005 by t7online.com is given in Figure 3, showing the modeled precipitation in mm for the next 6 h from GFS. The precipitation areas are encircled by isohyets—lines with equal amounts of precipitation. For research purposes, the GRIB1 datasets are collected from the NOAA server explained above during the wet season for the periods firstly from 2002–2006. Then the GFS-QPFs for the next 10 days made at the first data assimilation cycle of 0000 UTC every day have been extracted from these GRIB1 datasets for all gauging stations of the Yalong river basin shown in Figure 1, by a GRIB1 encoder/decoder program, which also can be downloaded from GRIB (2008). Finally, the observed and forecast 10-day accumulated mean areal precipitation of the Yalong river basin (termed 10-day accumulated precipitation) is estimated by the Thiessen polygon method (McCuen 1998), which assigns an area called a Thiessen polygon to each station in the Yalong river basin shown in Figure 1. The verification and practical use of the collected 10-day accumulated GFS-QPFs over the Yalong river basin will be discussed in the following sections in detail.

**Verification of 10-day accumulated GFS-QPFs**

In Zhu’s (2007) verification procedure, nine threshold precipitations in mm are selected above, covering 0.2, 2, 5, 10, 15, 25, 35, 50 and 75, as seen in Figure 2. It is obvious that these threshold precipitations can only reflect the variation law of rainfall itself, without considering the nature of the reservoir hydropower operation. So we
promote a data-mining approach using the decision tree algorithm (Breiman et al. 1984; Quinlan 1986; Fayyad & Irani 1992) to extract knowledge from historical generation data and to learn how the 10-day accumulated precipitation (denoted by \( P \)) is relevant to power generation decisions (denoted by \( N \)) under what conditions or criteria. In the study, the historical hydropower operation data are obtained from the Ertan hydropower station, including the observed 10-day accumulated precipitations in mm, the initial storage in m (denoted by \( Z \)) and the power generation decisions in MW on the 10-day timescale during May to September from 1998 to 2006. A sampling of information including hydrological data and reservoir power generation decisions is listed in Table 1 for the Ertan reservoir. The derived decision tree rules for the month of September are plotted in Figure 4 and show that the Ertan reservoir should run with the power generation decision in September: 3,300 MW, when the 10-day accumulated precipitation

Figure 2 | The verification scores over CONUS including ETS, TSS and BIAS are for (a) 00 UTC of August 15, 2006–18 UTC of November 6, 2006 and (b) 00 UTC of November 1, 2006–18 UTC of February 5, 2007.
reaches 60 mm or more; 3,000 MW or more, when the precipitation varies from 30 to 60 mm; and below 3,000 MW, when the precipitation is less than 30 mm. Then the 10-day accumulated precipitation can be among three possible categories: 1, 2 or 3 by these threshold precipitations: 30 and 60 mm in September. Now the derived threshold precipitations can respond to the operational characteristics of hydropower generation for the Ertan hydropower station to a certain extent.

The threshold precipitations for other periods can also be inferred by the decision tree algorithm (Breiman et al. 1984; Quinlan 1986; Fayyad & Irani 1992) as done for September in Figure 4. Table 2 details the threshold precipitations in mm of the 10-day accumulated precipitation for the Ertan hydropower operation during May to September. Then the forecast and observed 10-day accumulated precipitations are discretized into three categories—category 1, 2 and 3—by these threshold precipitations in Table 2 during the wet season from 2002 to 2006. Finally, a three-category contingency table for different thresholds of precipitation amount is evaluated with datasets of forecast and observed precipitation listed in Table 3, which shows the frequency of forecasts and observations in the various bins.

From the frequencies $p(F, O)$ in Table 3, it comes to the preliminary conclusion that: when the 10-day accumulated precipitation forecast is in category 1, the frequency of observed precipitation being “below Category 3” is higher than 86% on average, so the heavier precipitation corresponding to Category 3 is less likely to fall over the Yalong river basin; When the precipitation forecast is in Category 2 or 3, the frequency of observed precipitation being “in Category 2 or higher” reaches 80%, and even up to 95% on average, so the lighter precipitation corresponding to Category 1 is less likely to fall over this river basin. Also different scores including Accuracy and Heidke skill score (HSS) can be derived from the numbers in the contingency Table 3 (Murphy & Winkler 1987). The Accuracy shows what fraction of the forecasts is in the correct category, and has a range of 0–1, with 1 being the perfect score. The HSS is the answer of “What is the accuracy of the forecast in predicting the correct category, relative to that of random chance?”. It has a range of $-\infty$ to 1 (perfect score), and 0 indicates no skill. The Accuracy and HSS are obtained as

![Figure 3](https://iwaponline.com/jh/article-pdf/13/1/81/386529/81.pdf) | Precipitation map obtained from t7onle.com, showing a 6-h precipitation forecast in mm from GFS[0].

![Figure 4](https://iwaponline.com/jh/article-pdf/13/1/81/386529/81.pdf) | The derived decision tree rules for September by a decision tree algorithm.

### Table 1 | A sampling of information including hydrological data and power generation decisions for the Ertan reservoir

<table>
<thead>
<tr>
<th>Attributes</th>
<th>Decision variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Month</td>
<td>$P$ (mm)</td>
</tr>
<tr>
<td>Sept.</td>
<td>35</td>
</tr>
<tr>
<td>Sept.</td>
<td>50</td>
</tr>
<tr>
<td>Sept.</td>
<td>65</td>
</tr>
<tr>
<td>... More training data</td>
<td></td>
</tr>
</tbody>
</table>

### Table 2 | Threshold precipitation in mm for the Ertan hydropower operation

<table>
<thead>
<tr>
<th>Month</th>
<th>Attributes</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$P$ (mm)</td>
<td>$Z$ (m)</td>
</tr>
<tr>
<td>May</td>
<td>(0,10)</td>
<td>(10,55)</td>
</tr>
<tr>
<td>June (Jun.)</td>
<td>(0,40)</td>
<td>(40,80)</td>
</tr>
<tr>
<td>July (Jul.)</td>
<td>(0,50)</td>
<td>(50,100)</td>
</tr>
<tr>
<td>August (Aug.)</td>
<td>(0,40)</td>
<td>(40,80)</td>
</tr>
<tr>
<td>September (Sept.)</td>
<td>(0,30)</td>
<td>(30,60)</td>
</tr>
</tbody>
</table>
Table 3 | Three-category contingency table showing the frequency of 10-day accumulated precipitation forecasts obtained by GFS and observations in the various bins from May to September in the Yalong river basin

<table>
<thead>
<tr>
<th>Month</th>
<th>Forecast, ( F_i )</th>
<th>Indicators</th>
<th>Observed, ( O_j )</th>
<th>( n(F_i, O_j) )</th>
<th>( p(F_i, O_j) ) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( F_i )</td>
<td>( O_j )</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>May</td>
<td>( n(F_i, O_j) )</td>
<td>11</td>
<td>3</td>
<td>0</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>( p(F_i, O_j) ) (%)</td>
<td>78.6</td>
<td>21.4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Jun.</td>
<td>( n(F_i, O_j) )</td>
<td>18.2</td>
<td>72.7</td>
<td>9.1</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>( p(F_i, O_j) ) (%)</td>
<td>18.2</td>
<td>72.7</td>
<td>9.1</td>
<td>11</td>
</tr>
<tr>
<td>Jul.</td>
<td>( n(F_i, O_j) )</td>
<td>12</td>
<td>23</td>
<td>7</td>
<td>42</td>
</tr>
<tr>
<td></td>
<td>( p(F_i, O_j) ) (%)</td>
<td>12</td>
<td>23</td>
<td>7</td>
<td>42</td>
</tr>
<tr>
<td>Aug.</td>
<td>( n(F_i, O_j) )</td>
<td>11</td>
<td>5</td>
<td>5</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>( p(F_i, O_j) ) (%)</td>
<td>11</td>
<td>5</td>
<td>5</td>
<td>21</td>
</tr>
<tr>
<td>Sept.</td>
<td>( n(F_i, O_j) )</td>
<td>8</td>
<td>15</td>
<td>6</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>( p(F_i, O_j) ) (%)</td>
<td>8</td>
<td>15</td>
<td>6</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>( n(O_j) )</td>
<td>27</td>
<td>65</td>
<td>32</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>( N(O_j) )</td>
<td>4</td>
<td>22</td>
<td>28</td>
<td>50</td>
</tr>
</tbody>
</table>

In this table \( n(F_i, O_j) \) denotes the number of forecasts in category \( F_i \) that had observations in category \( O_j \); \( p(F_i, O_j) \) denotes the frequency of forecasts in category \( F_i \) that had observations in category \( O_j \); \( n(F_i) \) denotes the accumulated number of forecasts in category \( F_i \); \( n(O_j) \) denotes the accumulated number of observations in category \( O_j \); and \( N \) is the accumulated number of forecasts.

0.50 and 0.46, respectively, for the 10-day accumulated precipitation over the Yalong river basin. Figure 5 compares the observed and predicted forecasts for 10-day accumulated precipitation in May and July 2006, taking an extremely dry year as an example. It is obviously concluded that the predicted categories display similar change trends to the observed ones. In May, the GFS-QPFs categories are equal to the observed ones in most cases but for May 1, May 2, May 25 and May 27, the maximum difference between observed and predicted categories is only 1, as illustrated in Figure 5(a). In July, observed rainfalls display a descending trend, i.e. there is more rainfall in the first few days, and then the rainfall is on the decrease on the following days. As seen in Figure 5(b), the descending trend can also be predicted accurately by the NOAA GFS. As mentioned above, the season of heavy or torrential rain is between July and August for the Ertan reservoir. However, there is less rainfall than usual in 2006, and that the observed rainfall is in the majority with precipitation being in Categories 1 and 2 in July, Figure 5(b). Fortunately, the GFS model can predict the facts successfully as illustrated in Figure 5(b), which will play a key role in hydropower operation for the Ertan hydropower station during the extremely dry years. In all, it has been found that the Global Forecast System has a considerable capability of forecasting for all categories of 10-day accumulated precipitation, especially for the highest category of precipitation considering the Ertan hydropower operation.

**FORECASTS OF 10-DAY AVERAGE INFLOW USING QPFs**

**GFS-QPFs-based rainfall–runoff model**

Many hydrological models can be used to make reservoir inflow predictions based on quantitative precipitation forecasts, and the comparative study includes lumped rainfall–runoff models (Reed et al. 2004) and more complex distributed hydrological models (Collischonn et al. 2005, 2007). However, the input data for the distributed hydrological models including land use, topography, vegetation cover and soil types (Beven 2001) have not been collected in this study for guiding the calibration of parameter values.
So the results presented in this paper are all obtained using a simple GFS-QPFs-based rainfall–runoff (denoted by GRR) model. The GRR model is incorporated with the estimated forecast of 10-day accumulated precipitation by the Thiessen polygon method (McCuen 1998), at 00 UTC the first day of every 10 days, e.g. May 1, May 11 and May 21 during the month of May, as an input variable.

The GRR model is a two-segment multi-factor inflow forecasting model. It is configured with multiple predictors, such as the 10-day accumulated GFS-QPFs of the current period $t$, $P_f^t$, the observed 10-day accumulated precipitation of the previous period $t - 1$, $P_{t-1}$ and the observed 10-day average inflow into the Ertan dam of the previous period $t - 1$, $Q_{t-1}$. In this study it is simply assumed that the values of antecedent soil moisture differ somewhat for the delivery period and the storage period, as the data are not available for inferring the relationships of antecedent soil moisture and runoff. Then the forecasting model has to be organized into two parts: one for the delivery period and the other for the storage period. The parameters of this model are estimated using the stepwise regression algorithm (Sun et al. 1998). The model is calibrated by the least-squares technique, using observed and forecast precipitation and 10-day average inflow into the Ertan dam on the 10-day timescales during the wet season from 2002 to 2005 (about 60 ten-days), and verified using data of 2006 (about 15 ten-days). It is given as a piecewise function in Equation (1):

$$Q_t = \begin{cases} 
79 + 4.32P_f^t + 11.94P_{t-1} + 0.75Q_{t-1} & \text{for } t \text{ belonging to the delivery period} \\
-274 + 21.33P_f^t + 4.42P_{t-1} + 0.62Q_{t-1} & \text{for } t \text{ belonging to the storage period}
\end{cases} \tag{1}$$

such as $Q_t$ is the modeled 10-day inflow into the Ertan dam for period $t$ using the model; $P_f^t$ is the estimated 10-day accumulated GFS-QPFs during period $t$ by the Thiessen polygon method (McCuen 1998) as explained in the subsection on quantitative precipitation forecasts; $P_{t-1}$ is the observed 10-day accumulated precipitation during the previous period $t - 1$, also estimated by the Thiessen polygon method and $Q_{t-1}$ is the observed 10-day average inflow into the Ertan dam of the previous period $t - 1$.
Ten-day average inflows into the Ertan dam obtained by the GRR model using input from GFS-QPFs are compared with forecasts obtained by the PREVIVAZ ARMA-type model currently in use. Figure 6 presents the hydrographs of observed and predicted 10-day average inflows during the wet season from 2002 to 2006. It can be seen that forecasts obtained by the GRR model are closer to the observed inflow in most cases, especially for the first five periods each year, which correspond to the end of the recession of the dry season, and for the rising parts of the hydrograph. The AR forecasts show a pattern of a one 10-day delay with maximum and minimum values postponed by one 10-day delay, which is a consequence of the model structure. The value of including new information given by rainfall forecasts can also be seen during periods when the hydrograph is rising and during sharp changes in inflow when the hydrograph may increase or decrease, which is consistent with that from Collischonn et al. (2007).

Figure 7 compares observed and predicted 10-day average inflows into the Ertan dam considering both the GRR hydrological model with input from GFS-QPFs and the currently used PREVIVAZ ARMA-type model during the wet season from 2002 to 2006. It can be seen that, for low inflows, both forecasting models perform relatively well, with points representing the GRR forecasts rather closer to the line of perfect forecasts. For inflows larger than 1,500 m$^3$/s, both forecasting models do not perform well, though points are considerably more dispersed with a lower pattern of larger dispersion for the AR model. As seen from Figure 6, that often occurs in the main flood season from late June to late September, when there is a dramatic change in the Ertan reservoir inflows. For example, the average inflows in August vary greatly over the last 49 years from 1958 to 2006: the minimum inflow is about 1,200 m$^3$/s, while the maximum inflow is approaching 7,200 m$^3$/s, which makes it hard to get perfect forecasts of inflow.
In some cases, however, it is not necessary to have very accurate forecasts, since relatively rough estimates can improve the operation of hydraulic structures, or can yield estimates of the risk that rivers will exceed specified discharge thresholds (Rabuffetti & Barbero 2005), especially for forecasts of inflows into reservoirs in the medium range. So qualitative analysis of 10-day inflow forecasts using the GRR model will be further carried out for evaluation tasks in the following subsections.

Several error analyses compare forecasts obtained by the GRR model with QPF as inputs with forecasts obtained by the AR model for both calibration (from 2002 to 2005) and verification (2006). The results are given in Table 4, showing that the GRR model performs better in all cases, no matter whether for calibration or verification. The reduction of average absolute errors (ABE) is of the order of 15% and the improvement in other statistics is similar. It is not possible to assert at present whether this improvement results in better decisions in reservoir operation. However, we expect that better forecasts will probably lead to better decisions.

### Qualitative analysis of forecast models

Statistical analysis of historical generation data of the Ertan reservoir shows that the reservoir can run safely with corresponding power outputs in early August: 3,300 MW, when the 10-day average inflow is above 2,700 m$^3$/s and the initial storage is above 1,180 m$^3$/s; 3,000 MW or larger, when the inflow is lower than 1,800 m$^3$/s. So the 10-day average inflow can be among the three possible categories: low, median or large by the specified inflow thresholds 1,800 m$^3$/s and 2,700 m$^3$/s for early August. The threshold inflows for other periods can also be inferred by a Decision Tree Algorithm (Breiman et al. 1984; Quinlan 1986; Fayyad & Irani 1992) as done in the third section. Threshold inflows in m$^3$/s for 10-day average inflow into the Ertan reservoir are detailed in Table 5.

### Table 5: Threshold inflows in m$^3$/s for 10-day average inflow into the Ertan reservoir

<table>
<thead>
<tr>
<th>Month</th>
<th>Category</th>
<th>Calib.</th>
<th>Median</th>
<th>Large</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early/mid/late May</td>
<td>Low</td>
<td>(0, 800)</td>
<td>(800, 1,300)</td>
<td>&gt;1,300</td>
</tr>
<tr>
<td></td>
<td>Median</td>
<td>(920, 2,000)</td>
<td>&gt;2,000</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Large</td>
<td>(1,200, 2,500)</td>
<td>&gt;2,500</td>
<td></td>
</tr>
<tr>
<td>Mid June</td>
<td>Low</td>
<td>(2,200)</td>
<td>(2,200, 3,600)</td>
<td>&gt;3,600</td>
</tr>
<tr>
<td></td>
<td>Median</td>
<td>(1,800, 3,200)</td>
<td>&gt;3,200</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Large</td>
<td>(1,800, 2,700)</td>
<td>&gt;2,700</td>
<td></td>
</tr>
<tr>
<td>Late June</td>
<td>Low</td>
<td>(1,000)</td>
<td>(1,000, 2,000)</td>
<td>&gt;2,000</td>
</tr>
<tr>
<td></td>
<td>Median</td>
<td>(1,000)</td>
<td>(1,000, 1,950)</td>
<td>&gt;1,950</td>
</tr>
</tbody>
</table>

Table 6 shows a categorical statistics of the forecasts of 10-day average inflow into the Ertan dam obtained by both GRR and AR models with observations from May to September during the wet season. The hit rate (HTR) and the false alarm rate (FAR) are used to describe the accuracy of inflow forecasting models. The HTR is defined as the probability of the occurrence that the predicted category is correct, while the FAR is defined as the probability of the occurrence that the predicted category is incorrect.

### Table 6: Three-category contingency table showing the frequency of inflow forecasts obtained by the GRR model and AR model with observations in the various bins during the wet season

<table>
<thead>
<tr>
<th>Predicted category</th>
<th>Observed category</th>
<th>GRR model</th>
<th>AR model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
<td>Median</td>
<td>Large</td>
</tr>
<tr>
<td>Low</td>
<td>83.3</td>
<td>16.7</td>
<td>0</td>
</tr>
<tr>
<td>Median</td>
<td>26.1</td>
<td>65.2</td>
<td>8.7</td>
</tr>
<tr>
<td>Large</td>
<td>0</td>
<td>17.5</td>
<td>82.5</td>
</tr>
</tbody>
</table>

In this table, values are expressed in percentage terms.
equals the observed and the FAR denotes the probability of the occurrence of the predicted category totally opposed to the observed, e.g. the observed inflow belongs to the category Low but the predicted inflow reaches to the category Large and vice versa. For a good prediction, the HTR should be closer to 1 and the FAR should be near to zero. The statistics of qualitative forecasts are presented in Table 7 for both GRR and AR models.

The qualitative analyses in Tables 6 and 7 show that the GRR model performs relatively better in all cases than the AR model. The HTR of the GRR model is obtained as 80%, proving the precipitation forecasts obtained by the GFS can be applied to qualitative forecasts of 10-day average inflow. Moreover, the maximum difference between observed and predicted categories using the GRR model is only 1, with the FAR obtained as 0, so no false alarm will occur, which has little influence on the power generation decision-making of the Ertan hydropower station. In contrast, a false alarm occurs in the AR model, which may mislead decision-making of the reservoir operation. Therefore, the qualitative forecasts of 10-day average inflow using the GRR model could assist the decision-maker in selecting the better reservoir operating policy for the Ertan reservoir, combining with the initial storage of the current period and the Routine Generation Scheduling Chart (RGSC, as illustrated in Figure 8) policy or other policies.

**Table 7 | Statistics of qualitative forecasts using the GRR model and the AR model**

<table>
<thead>
<tr>
<th>GRR model</th>
<th>AR model</th>
</tr>
</thead>
<tbody>
<tr>
<td>HTR</td>
<td>FAR</td>
</tr>
<tr>
<td>80</td>
<td>0</td>
</tr>
</tbody>
</table>

In this table, the Hit rate (HTR) and the false alarm rate (FAR) are expressed in percentage terms.

**SIMULATIONS TO EVALUATE FORECASTING MODEL EFFICIENCY**

Hydropower operation for the Ertan reservoir is made by the RGSC policy in current operational use. As illustrated in Figure 8, the RGSC is divided into nine zones by different values of power generation decision \((N)\), e.g. 1,028, 2,200, 3,300 MW, etc. The RGSC policy is to release an amount of water to generate electrical energy equal to the demand of period \(t\), if possible. The demand here is determined by none other than the value of \(N\) obtained by which zone the initial storage of period \(t\) is located in in the RGSC. For example, when the hydropower operation calculation proceeds to mid-June and the initial storage is 1,155 m (point A of Figure 8), the value of \(N\) or the demand for this period will be 1,500 MW, as point A is located in the zone with \(N\) being 1,500 MW in the RGSC. When it is not possible to meet the demand, all the water in storage is released. If the

![Figure 8](https://iwaponline.com/jh/article-pdf/13/1/81/386529/81.pdf)
availability of water exceeds the sum of the demand and the capacity, the release is equal to the excess water available over the capacity. Obviously, the RGSC policy is made only by the initial storage of every period without taking full advantage of these available inflow forecasts. However, new forecasts of 10-day average flows become available at 00 UTC on the first day of every 10 days, e.g. 00 UTC of May 1, May 11 and May 21 as discussed in the fourth section. So the RGSC policy can be improved by considering the predicted inflows, and improved operational policies can also be derived by combining the RGSC policy with the predicted inflows.

To explore whether these predicted inflows result in better decisions in the Ertan hydropower operation, simulations for the improved policies by the predicted inflows are carried out using the historical inflow data and the associated forecasts using both the GRR and the AR models for the 10-day timescales during May to September 2006. The observed initial storage level of May 1 being 1,160 m is taken as the initial condition. In addition to release constraints and other physical and technical constraints, it has to be ensured that the end-of-period storage on September 30 reaches the NPL (1,200 m) to guarantee power generation for the following dry season. Table 6 lists observed and predicted categories of 10-day average inflows into the Ertan dam considering both the GRR model with input from GFS-QPFs and the currently used AR model in 2006. It is obviously shown that the GRR model performs better than the AR model. Unfortunately, false alarms occur in late July and late August for the AR model, which may mislead decision-making of the reservoir operation.

Figures 9 and 10 present the simulation results of the Ertan hydropower operation by the GRR, AR and RGSC policies from May to September 2006. The results in
Figure 9(a) show that the improved policies by the predicted inflows using GRR and AR models result in higher total hydropower generation and outperforming the RGSC policy, and the GRR policy performs better between these two improved policies. The increments of total power generation are about 476 GWh (6.3%) for GRR and 253 GWh (3%) for AR compared to that of RGSC.

Hydrographs of power dispatch by these two improved policies will be analyzed next in detail. For example, when the calculation proceeds to June 1 and the initial storage is 1,155 m³, the power generation decision is 1,500 MW for mid-June according to the RGSC policy, point A in Figure 8. But the inflow of mid-June is large enough to reach to category Large (Table 8), so the improved power generation decision RGSC should be increased appropriately to provide as much electrical energy as possible by utilizing available water resources during mid-June, as illustrated in Figure 9(b). Likewise, when the computation proceeds to late August, the initial storages of August 21 are 1,181.5 m³ and 1,181.3 m³ using the GRR and AR policy, respectively, and then the station should run with a power generation of 3,300 MW by RGSC, point B in Figure 8. However, the predicted inflow into the reservoir using the GRR model is only category Low (equal to the observed category in Table 8), so the original power generation decision by RGSC should be decreased to keep the reservoir running with high head water for more benefits in the remaining periods after late August. For the GRR policy, the end-of-period storage of late August is 1,185 m³, and the remaining power production is 1,976 GWh (Figure 10). In contrast, the corresponding values drop to 1,173.6 m³ and 1,893 GWh obtained by the AR policy (Figure 10). This is due to the fact that the predicted inflow using the AR model unexpectedly reaches category Large totally opposed to the observation as seen in Table 8, and the occurrence of false alarms results in that the reservoir runs with relatively lower head water as more water in storage is released in late August. The comparison shows that the predicted inflows using the GRR model with rainfall forecasts from the GFS result in better decisions, and better forecasts indeed lead to better decisions for power generation scheduling of the Ertan hydropower station.

SUMMARY AND CONCLUSIONS

The forecasts of 10-day accumulated precipitation from the Global Forecast System run by the American National Oceanic and Atmospheric Administration are evaluated to explore the potential improvements in reservoir inflow predictions and hydropower production comprehensively for the Ertan hydropower station of the Yalong river basin. Forecasts of 10-day accumulated precipitation are verified objectively by comparing them firstly with observed rainfall. Then a methodology for forecasting 10-day average reservoir inflows based on quantitative precipitation forecasts obtained from the Global Forecast System (GFS-QPFs) has also been presented and tested qualitatively and quantitatively, using a simple lumped hydrological model to estimate inflows from QPFs. Finally, simulations for improved operating policies, obtained by the average reservoir inflow forecasts using both the GRR and AR model, have also been presented.

Forecasts of 10-day accumulated precipitation are assessed by comparing them with observed rainfall during the wet season from 2002 to 2006, first with a three-category contingency table showing the frequency of forecasts and observations in the various bins. From the
analysis carried out, it has been found that GFS-QPFs have a considerable capability of forecasting for all categories of 10-day accumulated precipitation, especially for the highest category considering the Ertan hydropower operation.

The most important part reported in this paper is the use of the 10-day accumulated GFS-QPFs as input data into a simple hydrological model (GRR) for forecasting the inflow into the reservoir. The results from the proposed hydrological GRR model are compared quantitatively and qualitatively, with forecasts obtained by the currently used PREVIVAZ ARMA-type forecasting model during the wet season from 2002 to 2006. The comparisons show that the GRR model performs better than the AR model, especially for low flows, both in terms of error statistics and of the visual inspection of hydrographs and scatter plots. The qualitative analysis is also presented in this study and the results show the GRR model performs relatively better than the AR model, in all cases with a hit rate as 80% and no occurrence of false alarms.

Simulations for medium-range hydropower operation have also been presented to evaluate the efficiency of the reservoir inflow forecasting model using the GFS-QPFs. As expected, the predicted inflows using the GRR model with the GFS-QPFs result in better decisions with an increment of power generation as 476 GWh, and better forecasts indeed lead to better decisions for hydropower operation of the Ertan hydropower station. In conclusion, the quantitative precipitation forecasts obtained by the Global Forecast System can be applied to 10-day average inflow predictions and hydropower productions of the Ertan hydropower station in the Yangtze river basin. Thus it can be concluded that the QPFs are quite promising in reservoir inflow predictions and very useful in deriving efficient operating polices for other hydropower station systems. Further improvement in precipitation prediction skill on the meteorological side is needed and further work has to be done to tackle the uncertainty issue on the hydrological side.

ACKNOWLEDGEMENTS

This research is supported by the National Natural Science Foundation of China (no. 50579095) and the Ertan Hydropower Development Company, Ltd.

REFERENCES


Golding, B. W. 2000 Quantitative precipitation forecasting in the UK. J. Hydrol. 239, 286–305.


GRIB 2008 NCEP WMO GRIB2 Documentation. Available at: http://www.nco.ncep.noaa.gov/pmb/docs/grib2/


Jasper, K., Gurtz, J. & Lang, H. 2002 Advanced flood forecasting in Alpine watersheds by coupling meteorological observations and forecasts with a distributed hydrological model. J. Hydro. 267, 40–52.


McBride, J. L. & Ebert, E. E. 2000 Verification of quantitative precipitation forecasts from operational numerical weather prediction models over Australia. Weather Forecast. 15, 103–121.


Pappenberger, F., Beven, K. J., Hunter, N. M., Bates, P. D., Gouwelleeuw, B. T., Thielen, J. & De Roo, A. P. J. 2005 Cascading model uncertainty from medium range weather forecasts (10 days) through a rainfall-runoff model to flood inundation predictions within the European Flood Forecasting System (EFFS). Hydrol. Earth Syst. Sci. 9 (4), 381–393.


