

Tracking the Uncertainty Propagation Process Between Hydrological Forecasting and Reservoir Real-Time Optimal Operation

Feilin Zhu

Hohai University

Ping-an Zhong (✉ zhongpinganhhu@163.com)

Hohai University

Bin Xu

Hohai University

Yufei Ma

Hohai University

Qingwen Lu

Hohai University

Han Wang

Hohai University

Research Article

Keywords: uncertainty propagation, hydrological forecasting, reservoir real-time optimal operation, temporal correlation, stochastic optimization

Posted Date: May 25th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-513460/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

1 **Tracking the uncertainty propagation process between hydrological forecasting**
2 **and reservoir real-time optimal operation**

3

4 Feilin Zhu^a; Ping-an Zhong^{a, b*}; Bin Xu^a; Yufei Ma^a; Qingwen Lu^a; Han Wang^a

5

6 ^a College of Hydrology and Water Resources, Hohai University, No.1 Xikang Road,
7 Nanjing 210098, China

8 ^b National Engineering Research Center of Water Resources Efficient Utilization and
9 Engineering Safety, Hohai University, No.1 Xikang Road, Nanjing 210098, China

10

Name	E-mail address
Feilin Zhu	zhufeilin@hhu.edu.cn
Ping-an Zhong	pazhong@hhu.edu.cn
Bin Xu	xubin_hhu@hhu.edu.cn
Yufei Ma	mayufei@hhu.edu.cn
Qingwen Lu	luqw@hhu.edu.cn
Han Wang	wanghan_hhu@163.com

11

12 **Corresponding author:** Ping-an Zhong (pazhong@hhu.edu.cn)

13 **Abstract**

14 The inherent uncertainty in hydrological forecasting poses a challenge for reservoir
15 real-time optimal operation. In this paper, a stochastic framework is proposed to track
16 the uncertainty propagation process between hydrological forecasting and reservoir
17 operation. The framework simulates the comprehensive uncertainty of hydrological
18 forecasts in the form of ensemble forecasts and scenario trees. Based on the derived
19 analytic relationship between the performance metric Nash-Sutcliffe efficiency
20 coefficient (NSE) and forecast uncertainty probability distribution, we use three
21 methods (two are commonly used classical methods and one is the Gaussian copula
22 method) simultaneously to generate inflow forecast ensembles. Compared with the two
23 classical methods, the Gaussian copula method additionally takes into account the
24 temporal correlation of reservoir inflows. Then, the neural gas method is employed to
25 transform the generated ensembles into a scenario tree, which is further used as an input
26 for reservoir stochastic optimization. To improve the adaptability to uncertainties in
27 inflow forecasts, we establish a stochastic optimization model that optimizes the
28 expectation of objective values over all scenarios. Meanwhile, we propose a parallel
29 differential evolution (DE) algorithm based on parallel computing techniques for
30 solving the stochastic optimization model efficiently. Risk assessment is performed to
31 capture the uncertainty and corresponding risk associated with the reservoir optimal
32 decision. The proposed framework is demonstrated in a flood control reservoir system
33 in China. Furthermore, we conduct several numerical experiments to explore the effect
34 of forecast uncertainty level and temporal correlation on reservoir real-time optimal

35 operation. The results indicate that the temporal correlation of inflows must be
36 considered in inflow stochastic simulation and reservoir stochastic optimization,
37 otherwise the operational risk is likely to be overestimated or underestimated, thus
38 leading to operation failures. Based on the risk simulation surface, reservoir operators
39 can examine the robustness of operational decisions and thus make more reliable final
40 decisions.

41 **Keywords:** uncertainty propagation; hydrological forecasting; reservoir real-time
42 optimal operation; temporal correlation; stochastic optimization

43 **1. Introduction**

44 Reservoirs play an important role in the integrated development and management of
45 water resources globally, and the real-time optimal operation of reservoirs is an
46 important and hot research problem in the field of water resources management (Ahmad
47 et al., 2014). The main task of reservoir real-time optimal operation is to make optimal
48 decisions (e.g., reservoir release process or water level trajectory) for a future period
49 using optimization techniques, thus maximizing benefits of the reservoir. Hydrological
50 forecasts provide predictive information about reservoir inflows for a certain lead time,
51 which is essential to guide optimal reservoir operation.

52 However, hydrological forecasts on different time scales (short-, medium-, and
53 long-term) are subject to varying degrees of uncertainty, and this inherent uncertainty
54 is still the most serious challenge faced by reservoir operation (Ding et al., 2015;
55 Kujawa et al., 2020; Li et al., 2018). The occurrence and development of hydrological
56 processes in a basin are closely related to unstable weather systems and complex
57 subsurface conditions, and is a very complex natural process, which contains both
58 deterministic dynamic patterns and uncertain statistical patterns. The complexity of
59 hydrological processes and the limitations of human understanding make the three
60 aspects of hydrological model input, model structure and parameter estimation
61 uncertain, which leads to uncertainty in hydrological forecast results. Despite the great
62 progress and breakthroughs in research on water information telemetry, radar rainfall
63 measurement technology and hydrological models, hydrological forecasting has always
64 been accompanied by inherent uncertainty (Moges et al., 2020). In order to quantify the

65 uncertainty of hydrological forecasts, a large amount of research work has been carried
66 out by scholars worldwide in the last 20 years, and many effective theoretical methods
67 have been proposed (Ajami et al., 2007; Beven and Freer, 2001; Bourdin et al., 2014;
68 Ghaith and Li, 2020). According to the form of forecast release, hydrological forecasts
69 can be divided into three categories: deterministic forecasts, probabilistic forecasts and
70 ensemble forecasts (Cloke and Pappenberger, 2009).

71 Uncertainty in reservoir inflow forecasts often leads to deviations in the actual
72 implementation of previously planned optimal decisions to some extent, resulting in a
73 variety of risk events. Therefore, uncertainty will propagate from hydrological forecasts
74 to reservoir operation. Accurate tracking of the uncertainty propagation process is
75 essential to capture the uncertainty and risk in reservoir operation. Generally, risk
76 analysis methods for reservoir operation can be classified into two types: analytical
77 methods and numerical methods, which are still hot research topics in the field of
78 reservoir operation. Chen et al. (2014) derived the risk calculation formula based on the
79 stochastic differential equation of reservoir flood regulation. Liu et al. (2015) proposed
80 a two-stage risk analysis method by using the forecast horizon point to divide the future
81 time into two stages, i.e., the forecast lead-time and the unpredicted time. Zhang et al.
82 (2019) extended the two-stage risk analysis method to multi-reservoir systems. Huang
83 et al. (2018) developed a risk analysis method for flood control reservoir operation
84 considering the uncertainties of flood forecasting, hydrograph shape, streamflow
85 simulation, reservoir storage, water level, and discharge outflow. Chen et al. (2017)
86 proposed a decomposition-integration risk analysis method by decomposing the

87 original complex flood control system into a number of independent subsystems. Chen
88 and Zhong (2019) presented a multi-time-scale risk analysis approach for hydropower
89 reservoirs based on dynamic Bayesian network theory, considering the uncertainties of
90 reservoir inflow, electricity price, and hydropower consumption rate.

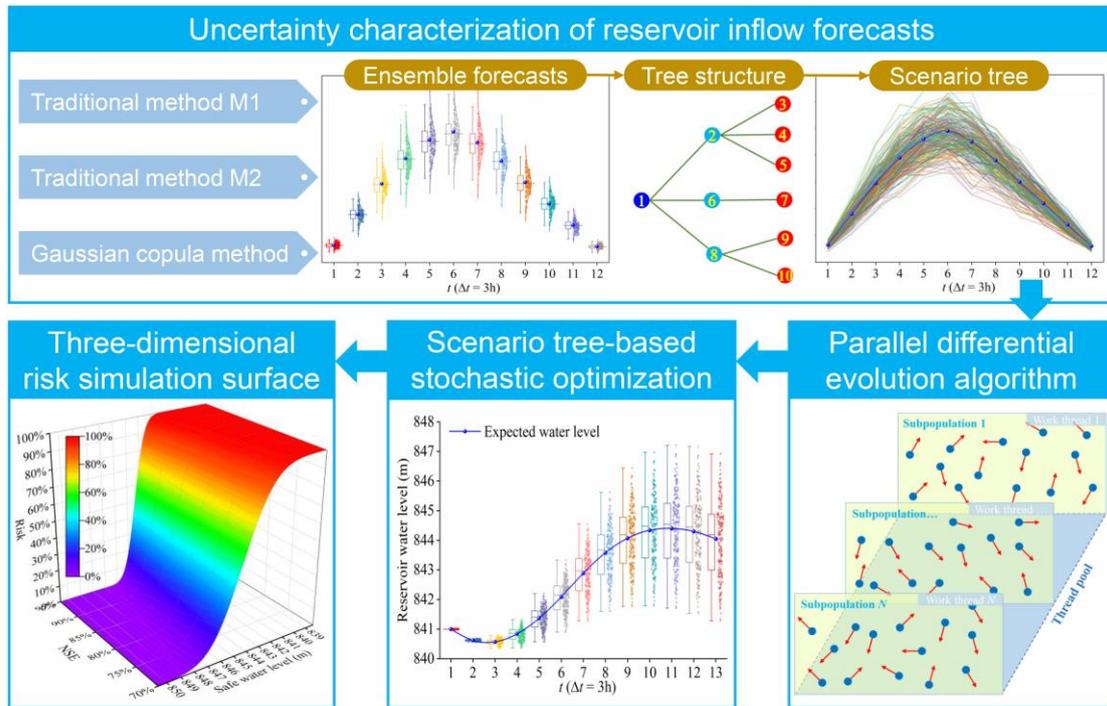
91 Although a number of methods are available for assessing reservoir operation risks,
92 most of the existing studies use fixed operation rules without considering optimization
93 under uncertainty. Zhu et al. (2017) pointed out that real-time optimization models
94 enable more efficient and flexible use of forecast information than operation rules, and
95 can be easily integrated into real-time operational systems. Therefore, it is necessary to
96 use an optimal operation model in risk analysis studies and to explicitly incorporate the
97 uncertainty of inflow forecasts in the optimization process. In addition, stochastic
98 simulation of hydrological forecast uncertainty is the first step in performing risk
99 analysis of reservoir scheduling. Traditional stochastic simulation methods tend to
100 focus only on the uncertainty of each time period itself, ignoring the temporal
101 correlation of inflows between adjacent time periods. In fact, the formation and
102 development of reservoir inflows are closely related to the development of weather
103 systems (especially rainfall), and the inflow between adjacent periods are often strongly
104 correlated. Hence, it is unrealistic to assume that inflow to the reservoir is uncorrelated
105 in time. Serial correlation is important and must be considered in reservoir optimization.
106 For example, in stochastic dynamic programming, inflow is modeled by a first order
107 Markov Chain and the transition probability matrix can be estimated from historical
108 records.

109 In this study, we propose a stochastic framework for tracking the uncertainty
110 propagation process between hydrological forecasting and reservoir real-time optimal
111 operation. First, based on the analytic relationship between the probability distribution
112 of forecast uncertainty and the performance metric Nash-Sutcliffe efficiency coefficient
113 (NSE), three methods (two are commonly used classical methods and one is the
114 Gaussian copula method) are used to generate inflow forecast ensembles. The
115 difference between these three methods is whether the temporal correlation of inflows
116 is taken into account. Then, the neural gas method is employed to transform the
117 generated ensembles into a scenario tree. To explicitly consider inflow forecast
118 uncertainty, we establish a stochastic optimization model for reservoir real-time
119 operation. A parallel differential evolution (DE) algorithm is designed, based on parallel
120 computing strategy, to solve the stochastic optimization model efficiently. We carry out
121 risk assessment to quantify the uncertainty in stochastic optimization results. Finally,
122 we design several numerical experiments to explore the effect of forecast uncertainty
123 level and temporal correlation on reservoir real-time optimal operation.

124 The remainder of this paper is organized as follows. Section 2.1 presents methods
125 for characterizing the uncertainty of reservoir inflow forecasts. Mathematical formulas
126 of the stochastic optimization model and the parallel DE algorithm are presented in
127 Section 2.2. Section 2.3 presents an overview of the risk assessment method. Section 3
128 introduces the case study and three numerical experiments. Finally, discussion and
129 conclusions are summarized in Section 4.

130 **2. Methodology**

131 Fig. 1 illustrates the logical relationships between each part of this study, and details of
132 the methodology are presented in the following subsections.



133

134 **Fig. 1. Flowchart of the proposed framework for tracking the uncertainty propagation**

135 **process between hydrological forecasting and reservoir real-time optimal operation**

136 **2.1. Uncertainty characterization of reservoir inflow forecasts**

137 Based on the derived analytic relationship between the performance metric Nash-
138 Sutcliffe efficiency coefficient (NSE) and forecast uncertainty probability distribution,
139 three methods (two are commonly used classical methods and one is the Gaussian
140 copula method) are employed to generate inflow forecast ensembles. Finally, the neural
141 gas method is used to transform these ensembles into a scenario tree as the input to a
142 reservoir stochastic optimization model.

143 **2.1.1. Relationship between NSE and standard deviation of forecast errors**

144 Nash-Sutcliffe efficiency coefficient (NSE), also referred to as the coefficient of
 145 determination (R^2), is a widely used performance metric to evaluate the forecast skills
 146 of hydrological models. In this paper, the probability distribution of inflow forecast
 147 uncertainty is derived via NSE, i.e., quantifying the analytic relationship between NSE
 148 and standard deviation of forecast errors (Yan et al., 2014).

149 NSE reflects the goodness-of-fit between the observed and forecasted inflows,
 150 which can be defined as follows:

$$151 \quad NSE = 1 - \frac{\sum_{t=1}^T (Q_{fct,t} - Q_{obs,t})^2}{\sum_{t=1}^T (Q_{obs,t} - \overline{Q_{obs}})^2} \quad (1)$$

152 where $Q_{obs,t}$ and $Q_{fct,t}$ represent the observed and forecasted reservoir inflow at time
 153 t , respectively; $\overline{Q_{obs}}$ is the average value of the observed inflow; T is the forecast
 154 horizon. The relative forecast error ε_t is assumed to follow a normal distribution with
 155 mean of zero and standard deviation of σ , i.e., $\varepsilon_t \sim N(0, \sigma^2)$. The relative forecast
 156 error at time t can be expressed as:

$$157 \quad \varepsilon_t = \frac{Q_{fct,t} - Q_{obs,t}}{Q_{obs,t}} \quad (2)$$

158 Substituting Eq. (2) into Eq. (1) yields:

$$159 \quad \sum_{t=1}^T \varepsilon_t^2 \cdot Q_{obs,t}^2 = (1 - NSE) \cdot \sum_{t=1}^T (Q_{obs,t} - \overline{Q_{obs}})^2 \quad (3)$$

160 Taking the mean value for both sides of Eq. (3) gives:

$$161 \quad \sum_{t=1}^T Q_{obs,t}^2 \cdot E(\varepsilon_t^2) = E(1 - NSE) \cdot \sum_{t=1}^T (Q_{obs,t} - \overline{Q_{obs}})^2 \quad (4)$$

162 As mentioned previously, $\varepsilon_t \sim N(0, \sigma^2)$, so $E(\varepsilon_t^2) = D(\varepsilon_t) + [E(\varepsilon_t)]^2 = \sigma^2$, where

163 $D(\varepsilon_t)$ is the variance of ε_t . Eq. (4) can be further written as:

$$164 \quad \sigma^2 \cdot \sum_{t=1}^T Q_{obs,t}^2 = (1 - NSE) \cdot \sum_{t=1}^T (Q_{obs,t} - \overline{Q_{obs}})^2 \quad (5)$$

165 Finally, the relationship between NSE and the standard deviation σ can be obtained as
166 follows:

$$167 \quad \sigma = \sqrt{(1 - NSE) \cdot \sum_{t=1}^T (Q_{obs,t} - \overline{Q_{obs}})^2 / \sum_{t=1}^T Q_{obs,t}^2} \quad (6)$$

168 Based on Eq. (6), the probability distribution of ε_t can be easily determined,
169 because σ is the only parameter of the distribution. For a specific region, NSE can be
170 estimated by quantitative analysis of historical inflow forecast records and observations
171 or qualitative analysis of forecast accuracy grade levels.

172 **2.1.2. Two traditional methods for generating forecast ensembles**

173 In recent years, a number of authors have focused on the problem of inflow forecast
174 uncertainty in reservoir operation (Li et al., 2010; Lu et al., 2020; Zhou et al., 2015). In
175 previous studies, reservoir inflow was often described as a stochastic process. After
176 deriving the probability distribution of inflow forecast uncertainty using Eq. (6), several
177 Monte Carlo sampling-based methods are available to synthesize ensemble inflow
178 forecast, which serves as a discrete approximation of the original stochastic process. In
179 this paper, we consider two representative methods (i.e., M1 and M2). In essence, both
180 of the M1 and M2 use Monte Carlo simulations to generate random numbers, while the
181 difference is how the random numbers are used for generating ensembles. The formulae
182 of the two methods are expressed as follows:

183
$$Q_{syn,t}^i = \begin{cases} Q_{fct,t} \cdot (1 + \sigma \cdot e^i) & \text{method M1} \\ Q_{fct,t} \cdot (1 + \sigma \cdot e_t^i) & \text{method M2} \end{cases} \quad (7)$$

184 where $Q_{syn,t}^i$ ($t = 1, 2, \dots, T$; $i = 1, 2, \dots, I$) represents the i th synthetic inflow sample
 185 at time t ; I is total number of the generated samples; $Q_{fct,t}$ is the deterministic reservoir
 186 inflow forecast; e is a random variable with standard normal distribution $N(0, 1)$.

187 For the i th inflow sample, method M1 only generates one random number e^i (I
 188 random numbers in total), and this random number is used to generate the inflow
 189 process over the entire forecast horizon. The final generated ensemble inflow forecasts
 190 have the characteristic of non-overlapping. However, for each time period t , method
 191 M2 generates I random numbers e_t^i ($I \times T$ random numbers in total), and the random
 192 numbers at each time are combined in sequence to form the ensemble inflow forecasts.

193 **2.1.3. Gaussian copula method for generating forecast ensembles**

194 The two commonly used methods mentioned above can be easily implemented in
 195 practice by Monte Carlo sampling techniques. However, both methods only focus on
 196 the uncertainty of the inflow forecast value at a single moment, without considering the
 197 correlation of inflows between adjacent time periods from the perspective of the whole
 198 inflow process. In fact, the formation and development of reservoir inflows are closely
 199 related to the development of weather systems (especially rainfall), and the inflow
 200 between adjacent periods are often strongly correlated. Hence, it is unrealistic to assume
 201 that inflow to the reservoir is uncorrelated in time. Serial correlation is important and
 202 must be considered in reservoir optimization. As a comparison with methods M1 and
 203 M2, we propose the third method, called Gaussian copula method, to describe the
 204 temporal correlation of inflows and further simulate forecast ensembles.

205 As described in Section 2.1.1, for a given lead time h , we can obtain the probability
 206 density function $f_{t+h|t}$ and the cumulative distribution function $F_{t+h|t}$ of the forecast
 207 relative error based on the performance metric NSE, which quantitatively describes the
 208 forecast uncertainty at each time, but cannot take into account the temporal correlation.
 209 In this paper, we use the Gaussian copula function to describe the interdependence
 210 structure of the multivariate stochastic process. By employing a multivariate Gaussian
 211 random number generator, we can issue at time t a number I of realizations
 212 $\{x_{t+1}^i, x_{t+2}^i, \dots, x_{t+T}^i\}$, $i = 1, 2, \dots, I$, from a multivariate Gaussian variable, for a modelled
 213 or estimated covariance structure. Based on the inverse probability function ϕ , as well
 214 as the cumulative distribution function $F_{t+h|t}$, these multivariate Gaussian realizations
 215 can be transformed into samples of inflow forecast errors with the same marginal
 216 distributions derived from NSE,

$$217 \quad y_{t+h|t}^i = F_{t+h|t}^{-1} \left[\phi \left(x_{t+h|t}^i \right) \right] \quad (8)$$

218 An exponential covariance function, which actually proved realistic in view of the
 219 empirical correlations observed, is used to model the covariance structures. Denoting
 220 by X_{t+h} the Gaussian random variable for lead time h , the exponential covariance
 221 function can be expressed as:

$$222 \quad \text{cov} \left(X_{t+h_1}, X_{t+h_2} \right) = \exp \left(- \frac{|h_1 - h_2|}{\varphi} \right) \quad (9)$$

223 where φ represents the parameter controlling the correlation strength of random
 224 variables among lead times. The φ value can be estimated by statistical analysis of
 225 historical forecasts and observed data. Eq. (9) shows that the strength of temporal
 226 correlation between different time period is related to the length of the time interval

227 $|h_1 - h_2|$, the closer the time interval, the stronger the correlation.

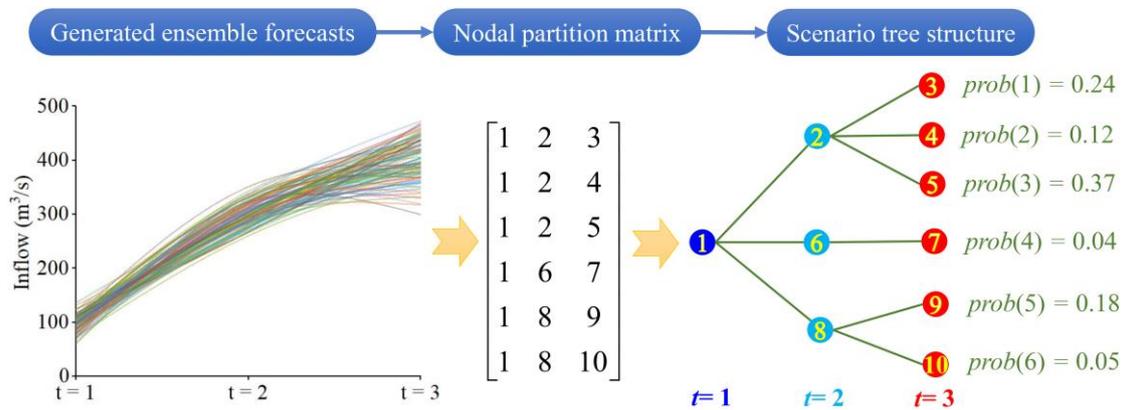
228 By coupling the samples of inflow forecast errors generated by Eq. (8) with the
229 deterministic forecast results, we can obtain the simulated forecast ensembles, which
230 have the same marginal distributions derived from NSE, and having different temporal
231 dynamics.

232 **2.1.4. Inflow scenario tree generation utilizing the neural gas method**

233 Ensemble forecasts are usually transformed into scenario trees in order to use the
234 recourse strategy in optimization, i.e., to adjust operational decisions in real time as
235 uncertainty is reduced through time (Raso et al., 2014). A stochastic optimization model
236 that considers all members in ensemble forecasts is computationally expensive. Hence,
237 an appropriate scenario reduction technique must be used to trim down the ensemble
238 forecasts. The ensemble forecast is actually a collection of trajectories representing all
239 possible future outputs, and the scenario tree determines when these trajectories branch
240 from each other, which can be considered as a simplified form of the ensemble forecast.
241 The scenario tree divides these ensembles into a set of branches with their respective
242 probabilities of occurrence. The difference between scenario tree and ensemble
243 forecasts is that scenario tree has a bifurcation structure that can be identified by
244 stochastic optimization models, and each scenario has its own occurrence probability,
245 while ensemble members are assumed to be equally likely to occur.

246 Fig. 2 shows the schematic of a scenario tree with three stages. In this paper, a
247 scenario tree is defined as $\{\omega_{s,t}\}_{t=1}^T$, $s \in [1, S]$, in which S is the size of the scenario
248 tree and T is operation horizon. The tree starts from the root node in the first period and

249 gradually branches to the node in the next period. Each root-to-leaf path represents an
 250 inflow scenario, representing a possible realization of the stochastic process over the
 251 entire prediction horizon. In the first time period, all scenarios share the same node. As
 252 time progresses, the tree gradually starts to branch. When $t = T$, each scenario has its
 253 own exclusive node. Each scenario has a probability $prob(s) > 0$, and the sum of $prob(s)$
 254 for all scenarios $\sum_{s=1}^S prob(s) = 1$.



255
256 **Fig. 2. Schematic of a scenario tree with three stages**

257 The neural gas method is a well-known algorithm that has been widely applied to
 258 simulate representative vectors from an existing set of vectors. Previous studies have
 259 proved that the neural gas method is superior to other clustering methods in generating
 260 scenario trees for streamflow (Latorre et al., 2007; Xu et al., 2015). In this paper, this
 261 method is used to generate scenario trees for reservoir inflow based on the generated
 262 ensemble forecasts. The main procedures of the neural gas method are briefly described
 263 as follows.

264 **Step 1. Initialization of nodal values**

265 The initialization is achieved by randomly selecting the trajectory from ensemble
 266 forecasts, initializing each node value of the scenario tree:

267
$$\omega_{s,t} = IN_{rand(\cdot),t} \quad s \in [1, S], t \in [1, T] \quad (10)$$

268 where IN denotes the simulated ensemble forecasts; $rand(\cdot)$ is a random number in
 269 the range $[1, K]$; and K is the size of ensemble forecasts.

270 Step 2. Selecting a new series and determining distance-orders

271 Before iteration, a complete series IN_k is randomly selected from IN . Euclidean
 272 distance between IN_k and the s th scenario is expressed below:

273
$$d_{s,k} = \|\omega_s - IN_k\| \quad s \in [1, S] \quad (11)$$

274 where the distance array d is sorted in increasing order and recorded as a new sequence
 275 in the array O .

276 Step 3. Updating nodal values based on iteration

277 The position of the featured vector is shifted towards the selected vector. Then the
 278 value of each node is updated according to its iteration time and distance-order:

279
$$\Delta\omega_{s,t} = \varepsilon(j) \cdot \sum_{s'=1,2,K,S | \omega_{s',t} \in \omega_s} h(O_{s'}, \lambda(j)) \cdot (IN_{k,t} - \omega_{s',t}) / \sum_{s'=1,2,K,S | \omega_{s',t} \in \omega_s} 1, j \in [1, jm] \quad (12)$$

280 where $\varepsilon(j) = \varepsilon_0 \cdot (\varepsilon_f / \varepsilon_0)^{j/jm}$ is the step size which decreases from ε_0 to ε_f as
 281 iteration time j increases from 0 to jm ; $h(O_{s'}, \lambda(j)) = \exp(-O_{s'} / \lambda(j))$ is an
 282 exponential function that is used to calculate the adaptation value of each node to its
 283 scenario distance-order; and $\lambda(j) = \lambda_0 \cdot (\lambda_f / \lambda_0)^{j/jm}$ changes from λ_0 to λ_f as
 284 iteration proceeds. When iteration progresses to $j+1$, the node value $\omega_{s,t}$ is updated
 285 according to $\Delta\omega_{s,t}$, i.e., $\omega_{s,t}^{j+1} = \omega_{s,t}^j + \Delta\omega_{s,t}$. Thus, featured vectors move towards the
 286 selected series IN_k .

287 Step 4. Calculating probabilities for each scenario

288 Based on the number of series with the minimum distance from the scenario, the

289 probability of each scenario is calculated:

$$290 \quad \text{prob}(s) = \text{Count} \left\{ k' \in [1, K], k' \mid d_{s,k'} = \min_{s' \in [1, S]} \{ d_{s',k'} \} \right\} / K \quad (13)$$

291 where $\text{prob}(s)$ represents the probability of the s th scenario and $\text{Count} \{ \cdot \}$ is a function
292 that counts numbers. Finally, a reduced number of reservoir inflow scenarios with their
293 respective probabilities can be obtained and used as an input to the stochastic
294 optimization model for reservoir real-time operation.

295 **2.2. Stochastic optimization model for reservoir real-time operation**

296 In reservoir real-time operation, decision makers need to make operational decisions
297 over an operation horizon with uncertain forecast information. Stochastic optimization
298 is a technique for solving optimization problems where partial information is uncertain,
299 and related uncertainties can be described with estimated probability distributions (Côté
300 and Leconte, 2016). Stochastic optimization has been successfully used in many areas,
301 and proved to be superior to traditional methods in providing better insights of
302 optimization results (Bayesteh and Azari, 2021; Xu et al., 2017; Zhu et al., 2020).

303 To handle inflow forecast uncertainty and improve the adaptability to uncertainties
304 in inflow forecasts, a stochastic optimization model is established for reservoir real-
305 time operation. The stochastic model ensures that the optimal value of the objective
306 function is reached in an average sense, i.e., the final optimal decision is derived on the
307 basis of optimizing the expected value of the objective function for all scenarios. In
308 addition, we design a parallel differential evolution (DE) algorithm, based on parallel
309 computing strategy, to solve the stochastic optimization model more efficiently. The
310 difference between our model and previous risk analysis papers lies in that the use of

311 stochastic optimization technique which enables inflow forecast uncertainty to be
 312 explicitly incorporated into the optimization process. However, previous studies mostly
 313 used fixed operation rules without optimization or developed deterministic
 314 optimization models (Chen et al., 2014; Zhang et al., 2019). Therefore, the established
 315 stochastic model in this paper can more precisely describe the propagation process of
 316 uncertainty from hydrological forecast to reservoir optimal operation.

317 **2.2.1. Objective function**

318 This paper considers the problem of reservoir real-time operation during floods. The
 319 essence of flood control operation is the reallocation of flood control capacity in time
 320 and space to ensure the safety of the dam. For a reservoir flood control system, the
 321 considered objective is to maximize the expected value of the reservoir flood control
 322 safety degree (FCSD), which is equal to the weighted sum of the safety degree under
 323 each inflow scenario according to the scenario probability:

$$324 \quad \max F = R\left(\left\{\omega_{s,t}\right\}_{t=1}^T\right) = \sum_{s=1}^S \text{prob}(s) \cdot R_s \quad (14)$$

$$325 \quad R_s = \left(V_s^{des} - Vu_s\right) / V_s^{des} \quad (15)$$

326 where R represents the flood control safety degree of a reservoir; R_s is the FCSD under
 327 the s th inflow scenario, which is defined in Eq. (15) as the ratio of the remaining flood
 328 control capacity divided by the design flood control capacity during the entire operation
 329 horizon; V_s^{des} is the design flood control capacity of a reservoir; Vu_s represents the
 330 occupied flood control storage used for flood routing in the s th inflow scenario, i.e., the
 331 difference between the highest water level and the flood limited water level. It should
 332 be mentioned that the objective function value is related to the inflow scenario tree, and

333 the reservoir water level is a stochastic process closely related to each scenario and its
334 occurrence probability.

335 2.2.2. Constraints

336 In the stochastic optimization model, the following constraints are considered.

337 (1) Mass balance equation:

$$338 \quad V_{s,t+1} = V_{s,t} + (I_{s,t} - O_{s,t}) \cdot \Delta t \quad (16)$$

339 where $V_{s,t}$ and $V_{s,t+1}$ are reservoir storage at the beginning and ending of time period
340 t in the s th scenario; $O_{s,t}$ and $I_{s,t}$ are reservoir outflow and inflow at time t in
341 scenario s , respectively; Δt is time span.

342 (2) Reservoir water level limits:

$$343 \quad \underline{Z}_t \leq Z_{s,t} \leq \overline{Z}_t \quad (17)$$

344 where $Z_{s,t}$ represents reservoir water level at time t in scenario s ; \overline{Z}_t and \underline{Z}_t
345 represent the maximum and minimum limits of reservoir water level at time t .

346 (3) Reservoir outflow limits:

$$347 \quad \underline{O}_t \leq O_{s,t} \leq \overline{O}_t \quad (18)$$

348 where \underline{O}_t and \overline{O}_t represent lower and upper bounds of reservoir outflow at time t ,
349 respectively.

350 (4) Reservoir outflow capacity limits:

$$351 \quad O_{s,t} \leq O_{\max}(Z_{s,t}) \quad (19)$$

352 where $O_{\max}(Z_{s,t})$ represents a function of the reservoir outflow capacity, which
353 determines the maximum reservoir outflow when all spillways reach their release
354 capacities.

355 (5) Reservoir outflow fluctuation limits:

$$356 \quad |O_{s,t} - O_{s,t-1}| \leq \Delta O_m \quad (20)$$

357 where ΔO_m is the permitted maximum outflow fluctuation. Reservoir outflow is not
358 allowed to fluctuate sharply, the constraint is set to avoid the sediment erosion of levee
359 systems and ensure the safety of downstream navigation.

360 (6) Initial and boundary conditions:

$$361 \quad Z_{s,t} = Z_{INI}; Z_{s,T} \rightarrow Z_{END} \quad (21)$$

362 where Z_{INI} and Z_{END} represent the initial and target end water level. Reservoir operators
363 usually expect the water level to fall back to the target water level. However, for real-
364 time flood control operation, the constraint is usually loosened, since the end water
365 level cannot always reach Z_{END} . This is especially true if the initial water level is high,
366 or if the flood peak is large and occurs late.

367 **2.2.3. Parallel differential evolution algorithm**

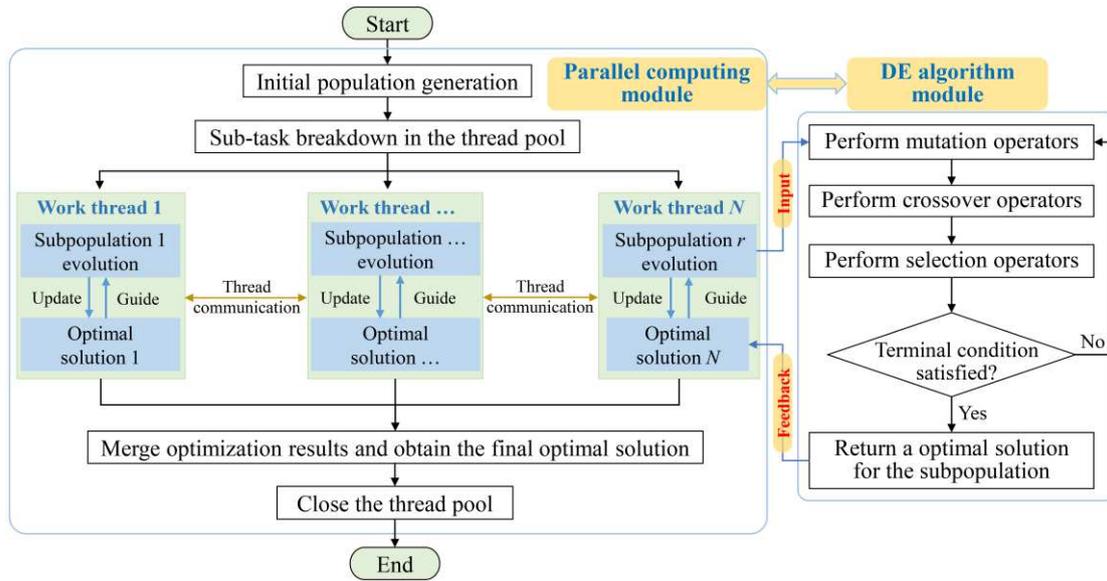
368 Differential evolution (DE) algorithm is a heuristic search algorithm based on
369 population intelligence theory. DE algorithm is similar to well-known genetic algorithm
370 in terms of process structure, which implements population evolution via mutation,
371 crossover and selection operators (Ahmadianfar et al., 2021). The basic steps of the DE
372 algorithm include: first, starting from a randomly generated initial population, two
373 individuals are selected from parent population to calculate the difference vector;
374 second, another individual is selected to sum with the difference vector to generate
375 experimental individuals; third, a crossover operator is performed between parent
376 individuals and experimental individuals to generate new offspring individuals; finally,

377 a selection operator is performed between parent individuals and the new offspring
378 individuals, and the search is guided to gradually approach the optimal solution by
379 continuously evolving to retain good individuals and eliminate inferior ones. DE
380 algorithm has the advantages of simple structure, fast convergence and robustness, and
381 has been widely used to solve complicated nonlinear constrained optimization problems
382 (Arora et al., 2021; Yan et al., 2020).

383 Deterministic reservoir operation models are generally solved using reservoir
384 water level or outflow as decision variables. Unlike deterministic optimization models,
385 the stochastic optimization model constructed in this paper considers all inflow
386 scenarios, which ensures the expected value of the objective function is maximized for
387 all scenarios. Therefore, the reservoir water level process under each scenario is chosen
388 as the decision variable in this paper, and the length of individuals in the population
389 generated by DE algorithm is S times larger than that of deterministic models (S is the
390 scenario tree size). As described in Section 2.2.2, some complex nonlinear constraints
391 exist in the stochastic optimization model, and we use the dynamic penalty function
392 (Zhu et al., 2020) method to deal with these constraints and transform the constrained
393 optimization problem into an unconstrained optimization problem. Specifically, the
394 violation of each constraint is dynamically calculated during each iteration and added
395 to Eq. (14) as a penalty term to further construct the fitness function. The inclusion of
396 these penalty terms can effectively identify the feasibility of candidate solutions, thus
397 improving the computational efficiency and ensuring the quality of the final optimal
398 solution.

399 In recent decades, with the rapid development of computer technology, almost all
400 computers (including laptops) are equipped with multi-core processors, which provides
401 a solid hardware foundation for parallel computing. Compared with the traditional
402 serial computing model, the execution efficiency of the same algorithm can be greatly
403 improved in a multi-core parallel computing environment. As a result, parallel
404 computing techniques have attracted a great deal of attention from researchers and are
405 becoming increasingly popular in the fields of reservoir operation and many other
406 engineering fields. The stochastic optimization model constructed in this paper has a
407 higher dimensionality than the traditional deterministic optimization model since the
408 values of the fitness function need to be calculated for all inflow scenarios in each
409 iteration. The computational effort of the stochastic optimization model grows
410 exponentially with the size of the scenario tree, and numerous scenarios often render
411 the model computationally intractable. In addition, the standard DE algorithm uses only
412 one core in serial computing mode, thus greatly wasting the computer's free multi-core
413 resources. In fact, the DE algorithm is naturally parallelizable due to the fact that each
414 iteration needs to evaluate the fitness of each individual in the population at the same
415 time. To this end, we design a parallel DE algorithm to solve the stochastic optimization
416 model more efficiently, and the flowchart of the algorithm is shown in Fig. 3. The
417 parallel DE algorithm divides the initial population into several smaller subpopulations
418 according to the number of computer threads. Each thread simultaneously executes the
419 standard DE algorithm to implement the computational tasks (i.e., mutation, crossover,
420 and selection) of its corresponding subpopulation until the terminal condition is

421 satisfied. After all threads have completed their computations, the thread manager
 422 collects the optimization results of all subtasks and obtains the final optimal solution.
 423 In this paper, we use the *multiprocessing* library in Python to implement the above
 424 parallel computing procedure.



425
 426

Fig. 3. Flow chart of DE algorithm using parallel computing strategy

427 **2.3. Risk assessment for reservoir real-time optimal operation**

428 Selecting an appropriate quantitative indicator is important to accurately track the
 429 uncertainty propagation process from hydrological forecasting to reservoir operation.
 430 For this reason, we carry out risk assessment of the results of stochastic optimization.
 431 Most of previous risk analysis studies were based on fixed operation rules and without
 432 using optimization models, i.e., directly using a large number of randomly simulated
 433 inflow samples as input and obtaining risk values by statistical analysis of flood
 434 regulation results (Chen et al., 2014). The difference between this paper and previous
 435 studies is that we explicitly consider hydrological forecast uncertainty in the stochastic
 436 optimization process. For stochastic optimal operation, the risk actually arises from the

437 variability between the optimal release decisions under different inflow scenarios,
438 which is closely related to each scenario and its probability of occurrence.

439 Strictly speaking, risk should be defined as a function of the probability of a risk
440 event occurring and the consequences it produces. However, for reservoir real-time
441 flood control operation, assessing the consequences of risk events is often difficult to
442 achieve. As a result, researchers have typically simplified the definition of risk to the
443 probability of a risk event occurring, and the simplified definition has been broadly
444 accepted in the field of reservoir operations (Chen et al., 2017). This paper also adopts
445 this simplified definition of risk, which is expressed as the probability that the
446 optimized highest water level exceeds the safe water level:

$$447 \quad Risk = \int_{Z_s}^{+\infty} f(Z_{\max}) dZ_{\max} \quad (22)$$

448 where $f(Z_{\max})$ denotes the probability density function of Z_{\max} , which can be
449 estimated quantitatively using entropy theory; Z_s represents the safe water level set by
450 decision makers.

451 3. Case study

452 3.1. Numerical experiment setting

453 The Pubugou reservoir system located in the Dadu River basin in China is selected as
454 the case study. The Dadu River is one of the major tributaries of the Yangtze River, with
455 a basin area of 90,000 square kilometers. According to the parameter settings shown in
456 Table 1, we designed three numerical experiments. Experiment 1 is used to demonstrate
457 the presented framework, where NSE and the parameter φ controlling the strength of
458 correlation in the Gaussian copula method is obtained by statistical analysis of forecasts
459 and observations of 389 historical floods. We perform experiment 2 to examine the
460 impact of forecast uncertainty level on reservoir real-time optimal operation by varying
461 NSE from 0.70 to 0.95 while fixing φ at its base value 4.68. Experiment 3 is
462 conducted to investigate the influence of forecast uncertainty temporal correlation on
463 reservoir real-time optimal operation, φ is varied from 0.50 to 150 while NSE is fixed
464 at its base value 0.91.

465 **Table 1 Parameter setting of three numerical experiments**

Experiment No.	Parameter symbol	Type	Value range	Base value
Experiment 1	NSE	Constant	/	0.91
	φ	Constant	/	4.68
Experiment 2	NSE	Variable	[0.70, 0.95]	/
	φ	Constant	/	4.68
Experiment 3	NSE	Constant	/	0.91
	φ	Variable	[0.50, 150]	/

466

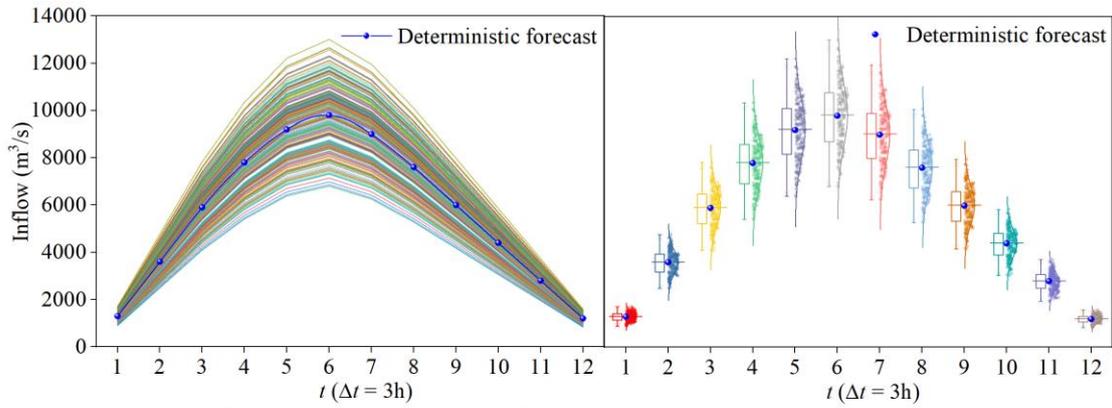
467 3.2. Numerical experiment 1: demonstration of the proposed framework

468 3.2.1. Generation of ensemble forecasts and scenario trees

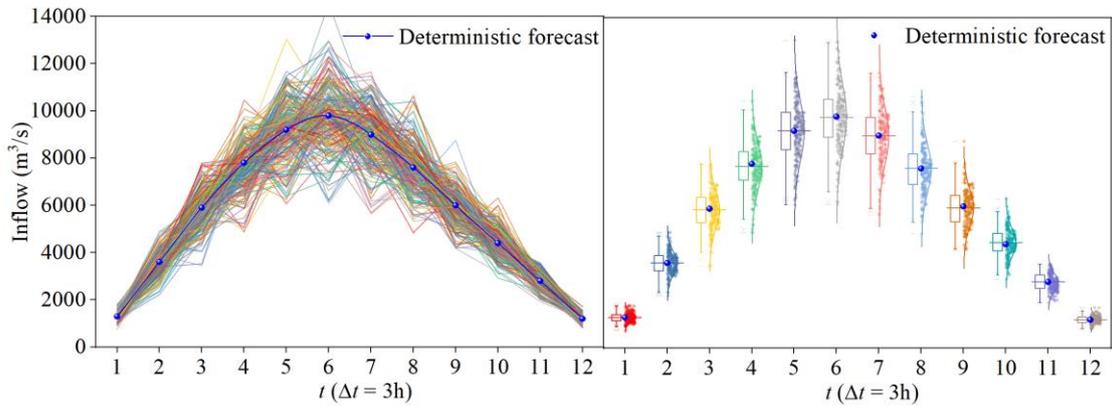
469 According to the parameter settings in Table 1, we use three methods (i.e., the
470 traditional methods M1, M2, and the Gaussian copula method) to generate ensemble
471 forecasts, which are then converted into scenario trees via the neural gas method (tree
472 size is set as 200). The methods M1 and M2 are only influenced by the parameter NSE,
473 while the Gaussian copula method incorporates an additional parameter φ to measure
474 the correlation structure between moments.

475 Fig. 4 compares inflow scenario trees generated by the three different methods. As
476 can be seen from Fig. 4, all three methods are capable of simulating the inflow
477 uncertainty since the generated scenario trees cover the original deterministic
478 forecasting process symmetrically. The box plot on the right side in Fig. 4 indicates that
479 the range of uncertainties simulated by the three methods at a single moment is basically
480 the same, and the mean values of their probability distributions are all approximately
481 equal to the deterministic forecast values. However, the difference between the three
482 methods lies mainly in the significant variability of the generated scenario trees in terms
483 of process shape. The inflow scenarios generated by method M1 are parallel to each
484 other without overlapping, while the inflow scenarios generated by method M2
485 fluctuate very sharply. The fluctuation of the inflow scenario process generated by the
486 Gaussian copula method lies between these two extremes. In general, the Gaussian
487 copula method can more realistically simulate the correlation of inflow between
488 adjacent time periods, and this correlation is reflected in the fact that the generated
489 inflow scenarios do not fluctuate drastically at random, but show a certain regularity in
490 the fluctuation direction and magnitude. The main reason is that the Gaussian copula

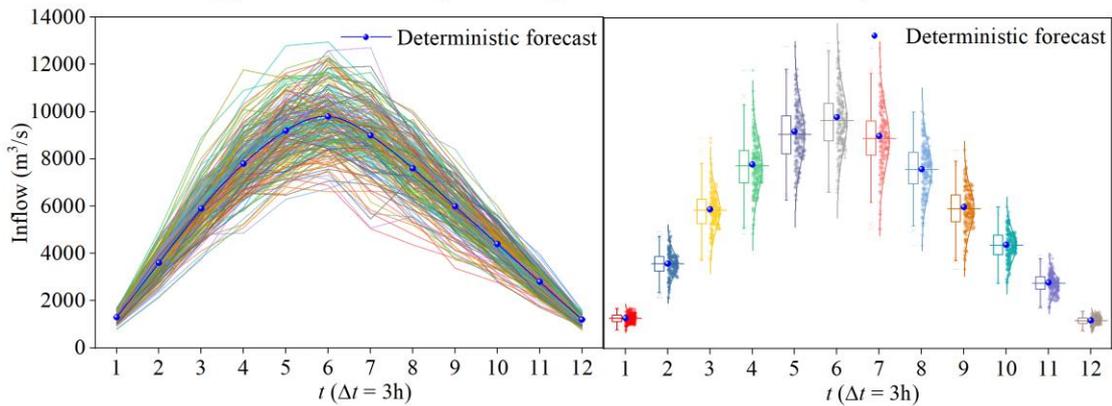
491 method takes into account the temporal dynamics of inflow processes, thus ensuring
 492 that the inflow scenarios follow both the marginal distribution at a given moment and
 493 the joint distribution throughout the entire forecast horizon.



(a) Line chart and box plot of the generated inflow forecasts using method M1



(b) Line chart and box plot of the generated inflow forecasts using method M2

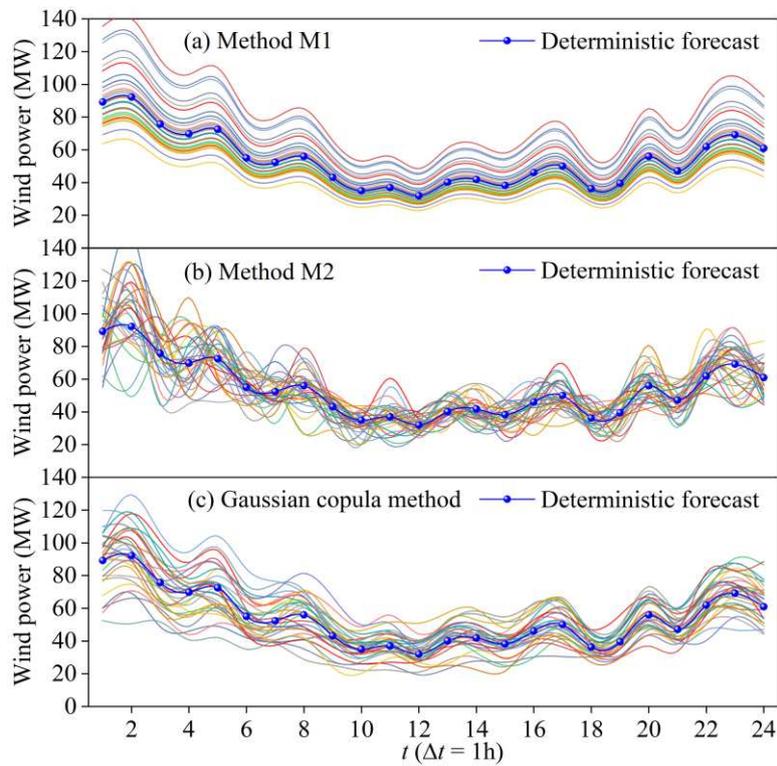


(c) Line chart and box plot of the generated inflow forecasts using Gaussian copula method

494
 495 **Fig. 4. Line charts and box plots of the generated inflow scenario trees using three different**
 496 **methods**

497 In addition, we simulate the hourly wind power process of a wind farm in the

498 eastern coastal region of China using the three methods. The wind power process is
 499 similar to reservoir inflows, and both are generally considered to be stochastic processes
 500 with strong uncertainties. Fig. 5 shows the generated wind power outputs using three
 501 different methods, and the scenario tree size is set as 30. As with the results of the inflow
 502 simulation, the Gaussian copula method yields results for wind power fluctuations that
 503 are closer to the real situation.



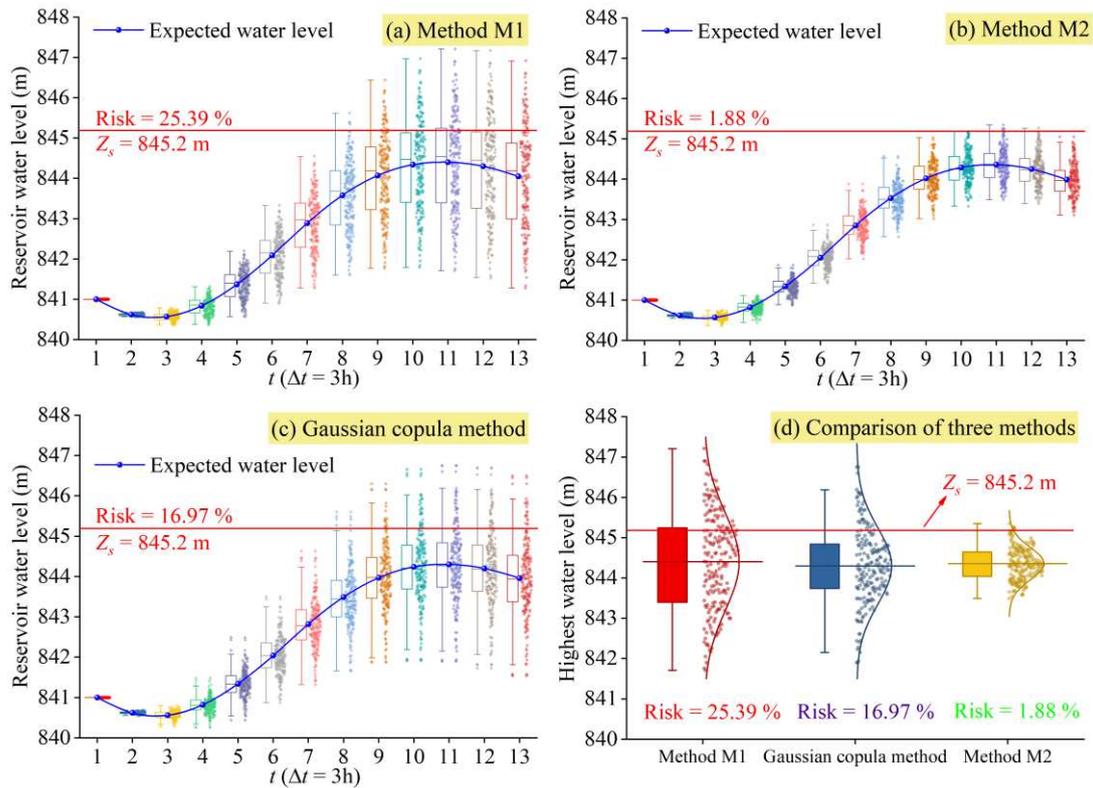
504

505 **Fig. 5. Comparison of the generated wind power outputs using three different methods**

506 **3.2.2. Stochastic optimal operation and risk assessment**

507 The inflow scenario trees generated by the three methods (shown in Fig. 4) are used as
 508 inputs to the stochastic optimization model respectively, and the parallel DE algorithm
 509 is implemented to solve the model to obtain the reservoir scheduling process under
 510 consideration of forecast uncertainty. The risk assessment of the stochastic optimal
 511 operation results is performed according to Eq. (22), and the safe water level set by the

512 decision maker is 845.2m. Fig. 6 presents the results of stochastic optimal operation
513 and risk calculation, where the water level process becomes a stochastic process under
514 the influence of hydrological forecast uncertainty. However, the stochastic processes of
515 reservoir water level obtained by the three methods are obviously different. The
516 uncertainty of the water level at each moment obtained by method M1 is the largest,
517 which is reflected in the most divergent distribution range of the box plot; the Gaussian
518 copula method has the second highest uncertainty, and method M2 has the smallest
519 uncertainty. As shown in Fig. 6(d), the risks corresponding to the three methods follow
520 the same pattern as above, with the methods M1 and M2 yielding a risk of 25.39% and
521 1.88% (the maximum and minimum, respectively), and the Gaussian copula method
522 yielding a risk of 16.97%. As shown in Fig. 4(b), the fluctuation of inflow scenario
523 generated by method M2 is the sharpest among the three methods, and its overall inflow
524 uncertainty seems to be the largest. However, a very interesting result can be found in
525 Fig. 6, after reservoir regulation, the uncertainty and risk of the water level process for
526 method M2 is much smaller than the other two methods.



527

528

Fig. 6. Results of reservoir stochastic optimal operation and risk assessment

529

3.2.3. Mechanism analysis of the risk differences between the three methods

530

This section carries out a mechanistic analysis to explore the reasons for the differences

531

in risks calculated by the three methods. For reservoir flood control operation, water

532

level variations are mainly influenced by the flood volume. For the method M2, since

533

the correlation between time periods is not considered in the inflow scenario generation,

534

the overestimated (or underestimated) inflow in the current time period will be offset

535

by the underestimated (or overestimated) inflow in the next time period, and this

536

offsetting effect of forecast errors in adjacent time periods (see Fig. 7(b)) reduces the

537

uncertainty in the cumulative flood volume and final operation results (i.e., reservoir

538

water level processes in Fig. 6(b)), thus significantly underestimates the risk. For the

539

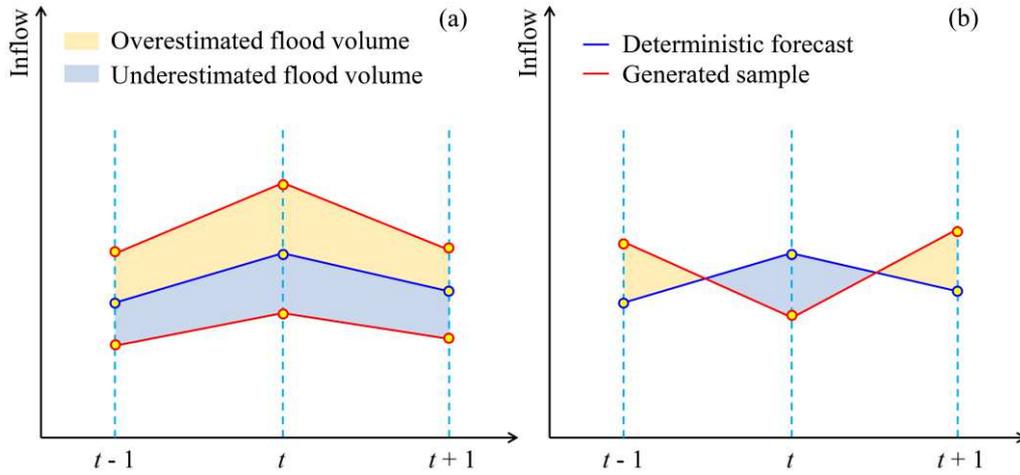
method M1, since the same random number is used to generate the inflow scenario in

540

each time period, the generated inflow processes are parallel to each other. Therefore,

541 the overestimated (or underestimated) inflow in the current time period will be
542 cumulated in the next time period, and this cumulative effect (see Fig. 7(a)) increases
543 the uncertainty of the cumulative flood volume and thus overestimates the flood control
544 risk. The Gaussian copula method estimates the temporal correlation of inflow based
545 on a large number of historical forecasts and observation samples, and takes them into
546 account when generating the inflow scenario tree. Therefore, the results of the Gaussian
547 copula method lie between the two extremes of the method M1 and M2, where the
548 stronger the correlation of historical samples, the stronger the cumulative effect and the
549 weaker the offsetting effect, and vice versa.

550 Fig. 8. compares the flood volume and cumulative flood volume forecasting errors
551 simulated by three different methods. The correlation has little effect on the inflow
552 value and flood volume for a single time period, but has a significant effect on the shape
553 of inflow processes and cumulative flood volume errors. As can be seen from Fig. 8,
554 the cumulative flood volume error simulated by method M1 is greatly enlarged due to
555 the cumulative effect over time, while the cumulative flood volume error simulated by
556 method M2 remains basically unchanged from the 6th time period due to the offsetting
557 effect. Hence, methods M1 and M2 actually overestimate and underestimate the
558 operation uncertainty and corresponding risk, respectively. In summary, for stochastic
559 optimal reservoir operation, the stochastic simulation of inflow processes cannot ignore
560 the correlation between time periods, otherwise the risk is likely to be overestimated or
561 underestimated, and the true risk value cannot be obtained, which will lead to operation
562 failures.



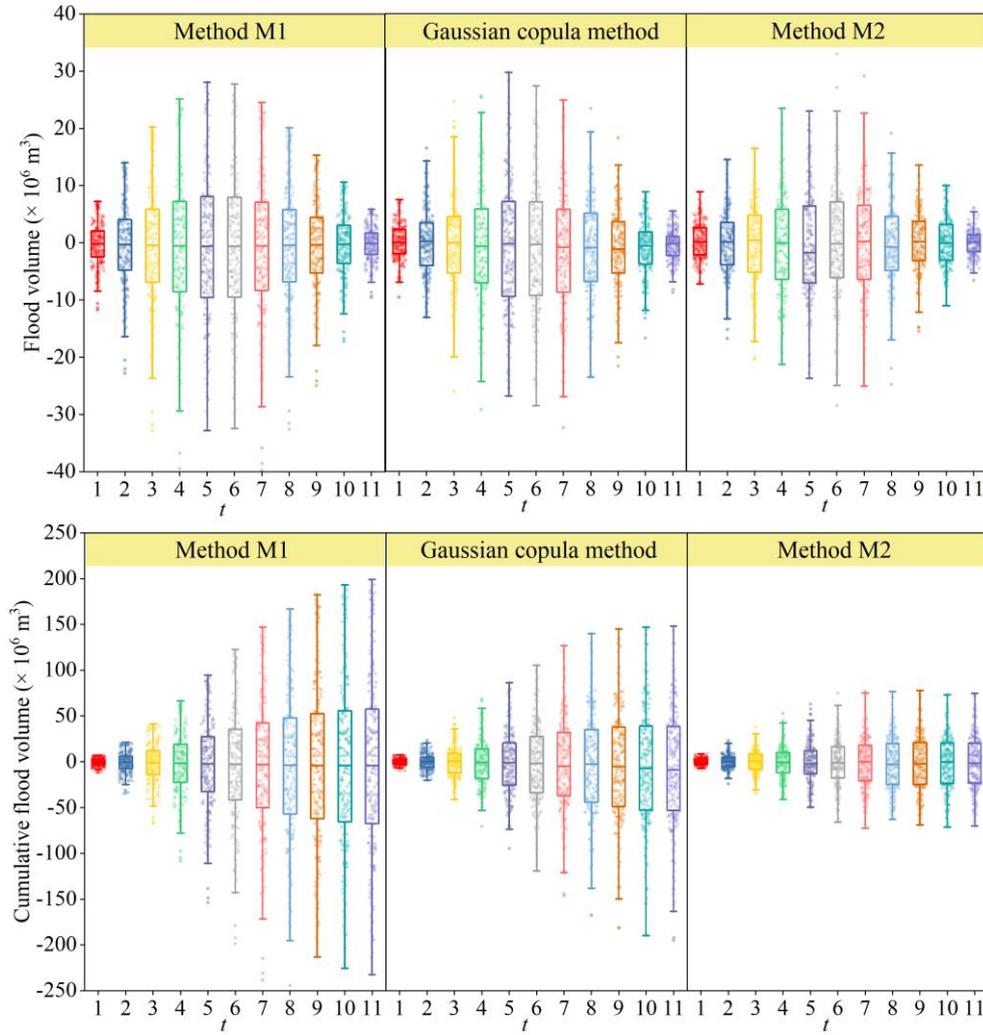
563

564

Fig. 7. Diagrammatic sketch of the cumulative effect and offsetting effect of flood volume

565

forecasting errors



566

567

Fig. 8. Comparison of flood volume and cumulative flood volume forecasting errors

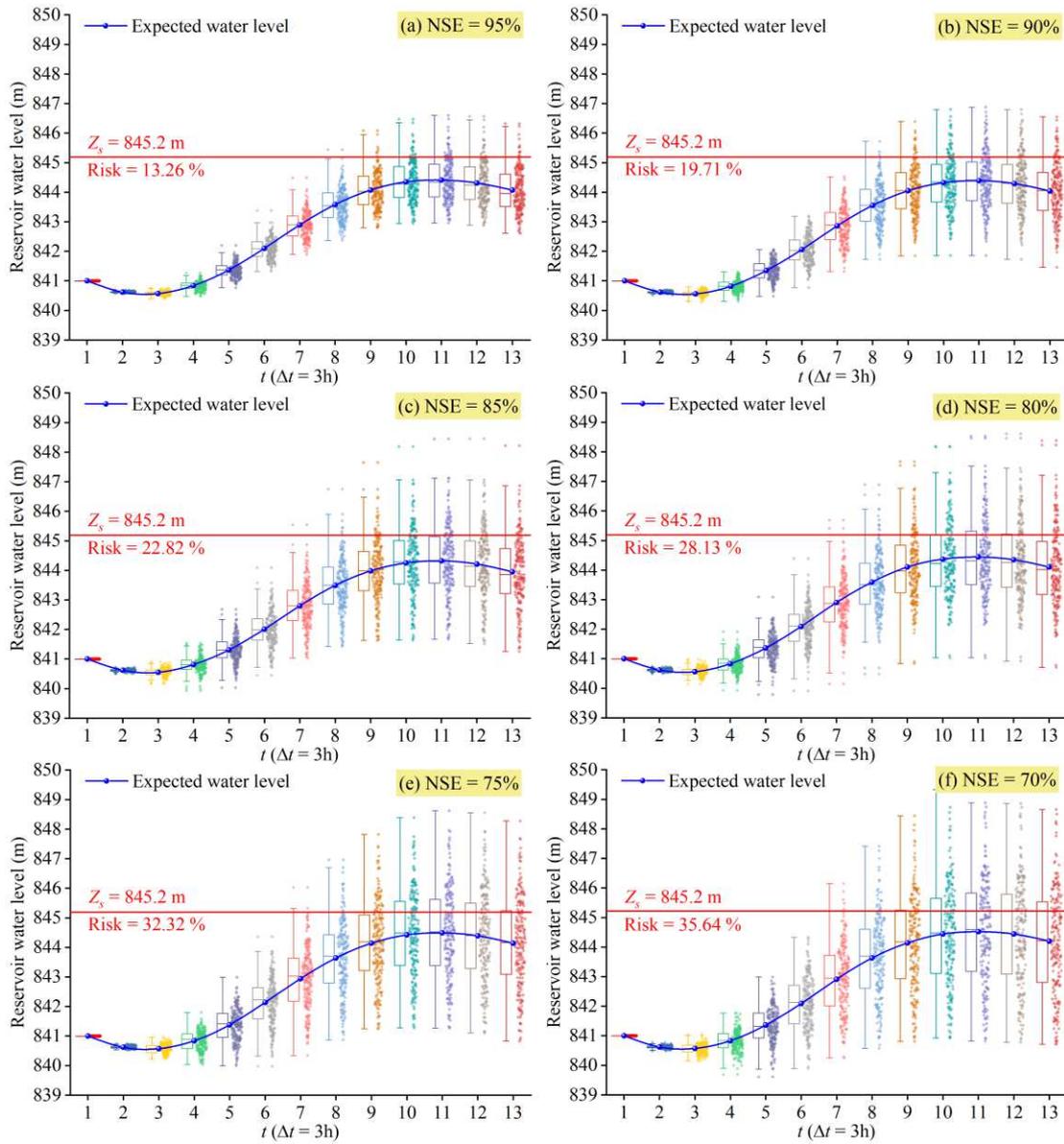
568 **simulated by three different methods**

569 **3.3. Numerical experiment 2: effect of forecast uncertainty level on reservoir real-**
570 **time optimal operation**

571 In this section, we perform numerical experiment 2 to examine the influence of forecast
572 uncertainty level on reservoir real-time optimal operation. Numerical experiment 2 uses
573 the control variables method, i.e., fixing the parameter φ equal to its base value 4.68
574 and varying NSE from 0.70 to 0.95. Fig. 9 illustrates the box plot of stochastic water
575 levels and corresponding risks under different forecast uncertainty levels (represented
576 by NSE). As can be seen from Fig. 9, when the NSE gradually decreases from 0.95 to
577 0.70, the uncertainty of the optimized water level process also increases, as shown by
578 the increasingly divergent distribution of water level values at each moment (especially
579 the 11th moment when the highest water level occurs), and the corresponding risk
580 increases from 13.26% to 35.64% as well. The above results suggest that the uncertainty
581 of reservoir optimal operation increases with the uncertainty of hydrological forecasts.

582 According to the risk formula Eq. (22), the risk is associated with the safe water
583 level set by decision makers, in addition to the probability density function of Z_{\max}
584 (depending on the NSE). To this end, we carry out a perturbation analysis, i.e., we
585 calculate risk values at a series of safe water levels and obtain a three-dimensional risk
586 simulation surface and its two-dimensional projection curves, as shown in Fig. 10.
587 Figure 10 is very useful in real-world applications. On the one hand, reservoir operators
588 can quickly determine the corresponding risk value according to a certain safe water
589 level, and on the other hand, we can determine the safe water level values under

590 different NSE conditions according to the acceptable risk of decision makers, thus
 591 providing an important reference for the boundary condition setting of the stochastic
 592 optimization model.

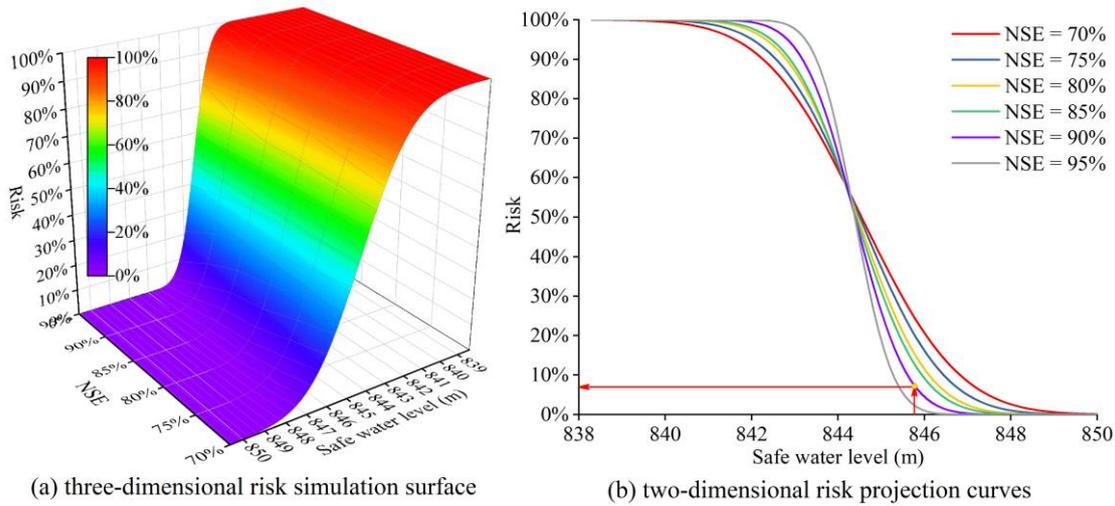


593

594 **Fig. 9. Box plot of reservoir water levels and related risks under different forecast**

595

uncertainty levels



596

(a) three-dimensional risk simulation surface

(b) two-dimensional risk projection curves

597 **Fig. 10. Relationship between safe water level and risk under different forecast uncertainty**

598 **levels**

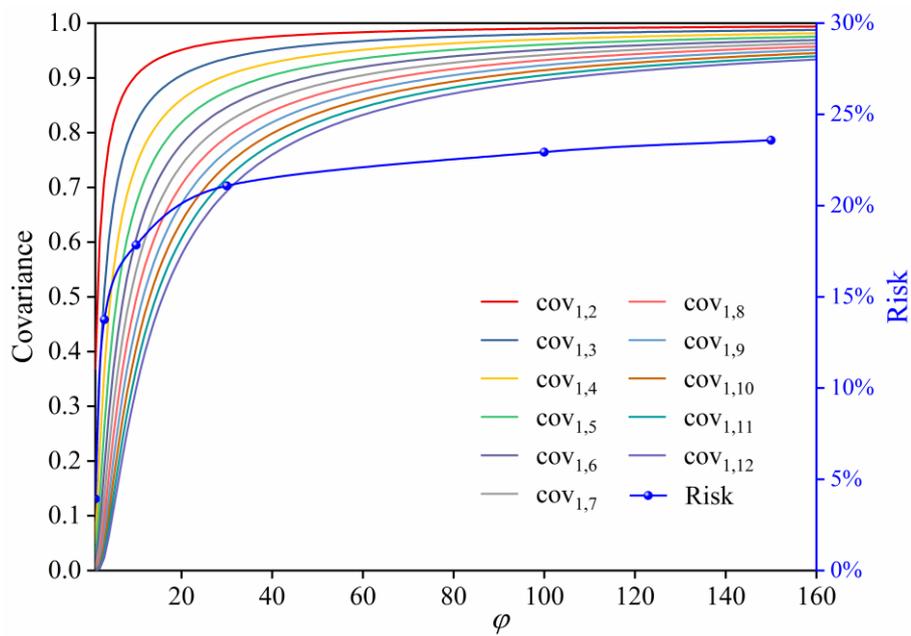
599 **3.4. Numerical experiment 3: effect of forecast uncertainty temporal correlation**

600 **on reservoir real-time optimal operation**

601 We further perform numerical experiment 3 to explore the impact of forecast
 602 uncertainty temporal correlation on reservoir real-time optimal operation. Unlike the
 603 previous section, we fix NSE at its base value 0.91, and vary the parameter φ from
 604 0.50 to 150. Fig. 11 shows covariance curves between time period 1 and other time
 605 periods under different temporal correlations, as well as the risk curves. Fig. 11
 606 indicates that the closer the time interval, the stronger the temporal correlation of inflow.
 607 The covariance and operational risk show an obvious nonlinear relationship with the
 608 parameter φ . When the parameter φ increases from 0.5 to about 40, the covariance
 609 and risk show a rapid increase, and then the variation of both tends to stabilize, i.e.,
 610 remains basically constant with the increase of the parameter φ .

611 Fig. 12 presents the relationship between safe water level and risk under different
 612 temporal correlations of inflow forecast uncertainty. From Fig. 12, it can be seen that

613 temporal correlations have a significant impact on reservoir operation risk. Therefore,
 614 when quantifying the uncertainty in hydrological forecasts, it is important to consider
 615 not only the probability distribution of each single time period, but also to estimate the
 616 correlation between time periods, which has important influences on reservoir
 617 stochastic optimization and risk assessment. However, this has usually been ignored in
 618 previous studies, thus yielding unrealistic (over- or underestimated) risk calculations.

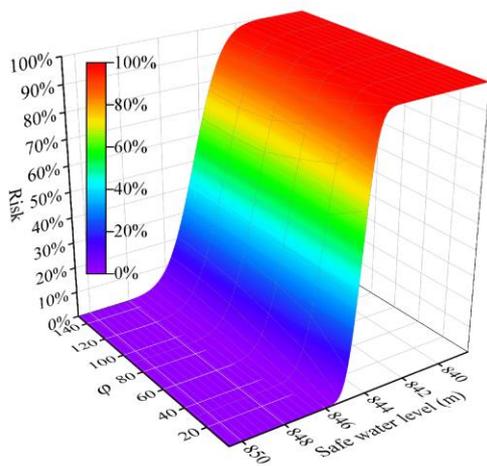


619

620 **Fig. 11. Covariance and risk curves under different temporal correlations of inflow forecast**

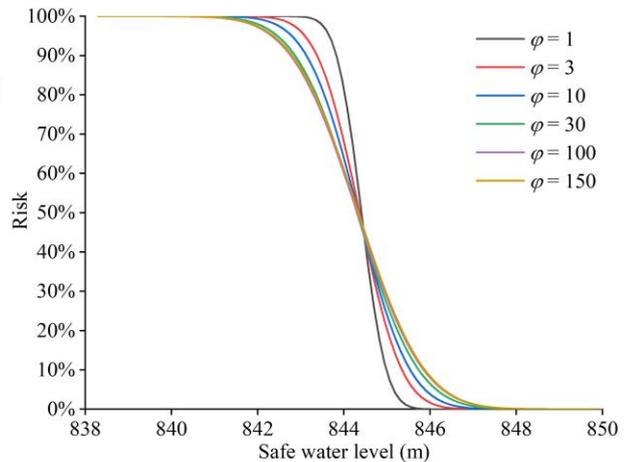
621

uncertainty



622

(a) three-dimensional risk simulation surface



(b) two-dimensional risk projection curves

623 **Fig. 12. Relationship between safe water level and risk under different temporal correlations**
624 **of inflow forecast uncertainty**

625 **4 Discussion and Conclusions**

626 Hydrological forecasts are inherently uncertain because of errors in hydrological model
627 inputs, complicated model structures and inexact parameter estimations. Hydrological
628 forecast uncertainty is an important challenge for reservoir real-time optimal operation.
629 In this paper, a stochastic framework was proposed to track the uncertainty propagation
630 process between hydrological forecasting and reservoir operation. First, three methods
631 were used to generate inflow forecast ensembles based on the derived analytic
632 relationship between the performance metric NSE and forecast uncertainty probability
633 distribution, respectively. The first two were the commonly used traditional methods
634 M1 and M2, and the third was the Gaussian copula method, which additionally
635 considered the temporal correlation of inflows between time periods. Then, the neural
636 gas method was used to transform the generated ensembles into a scenario tree, which
637 was further used as an input for reservoir stochastic optimization. To handle inflow
638 forecast uncertainty and improve the adaptability to uncertainties in inflow forecasts, a
639 stochastic optimization model was established, where the final optimal decision was
640 derived on the basis of optimizing the expected value of the objective function for all
641 scenarios. Based on parallel computing strategy, we designed a parallel differential
642 evolution (DE) algorithm to solve the stochastic optimization model more efficiently.
643 We performed risk assessment to capture the uncertainty and corresponding risk
644 associated with the reservoir optimal decision. Finally, three numerical experiments
645 were performed to demonstrate the effectiveness of the proposed framework, and to
646 investigate the effect of forecast uncertainty level and temporal correlation on reservoir

647 real-time optimal operation.

648 The Pubugou reservoir system located in the Dadu River basin in China was
649 selected as the case study. The main conclusions were outlined as follows.

650 (1) Compared with methods M1 and M2, the Gaussian copula method can more
651 realistically simulate the inflow correlation between adjacent time periods by
652 considering the temporal dynamics of inflow processes. Thus, the generated inflow
653 scenarios do not fluctuate dramatically at random, but show a certain regularity in the
654 fluctuation direction and magnitude.

655 (2) The optimized water level process becomes a stochastic process under the impact
656 of hydrological forecast uncertainties. However, the stochastic processes of reservoir
657 water level obtained by the three methods are obviously different, and the ranking
658 according to the uncertainty degree is $M1 > \text{the Gaussian copula method} > M2$.
659 Similarly, the obtained risks corresponding to the three methods also follow the same
660 pattern. It is interesting that inflow scenarios generated by method M2 fluctuate
661 sharpest and seem to be the most uncertain among all three methods, while after
662 reservoir regulation the uncertainty of the optimized water level process as well as the
663 risk is much smaller than the other two methods.

664 (3) The difference in the results of three methods mainly results from the cumulative
665 and offsetting effect of forecast errors. Method M1 has the strongest cumulative effect
666 of forecast error, method M2 has the strongest offsetting effect, and the Gaussian copula
667 method lies between these two extremes. The stronger the correlation of historical
668 samples, the stronger the cumulative effect and the weaker the offsetting effect, and

669 vice versa. The cumulative effect results in the cumulative flood volume error being
670 greatly enlarged over time, thus leading to an overestimation of risk; while the offsetting
671 effect tends to underestimate the corresponding risk. Therefore, the temporal correlation
672 between time periods must be considered when simulating stochastic inflow process,
673 otherwise it is likely to overestimate or underestimate the risk, thus leading to operation
674 failures.

675 (4) The uncertainty of reservoir stochastic optimization tends to increase with the
676 uncertainty of hydrological forecasts. The obtained three-dimensional risk simulation
677 surface is very useful in real-world applications. On the one hand, reservoir operators
678 can quickly determine the corresponding risk value according to a certain safe water
679 level, and on the other hand, we can determine the safe water level values under
680 different NSE conditions according to the acceptable risk of decision makers, thus
681 providing guidance for setting boundary conditions of the stochastic optimization
682 model. The risk simulation surface enables reservoir operators to examine the
683 robustness of operational decisions, and make risk-informed decisions with higher
684 reliabilities.

685 (5) The covariance and operational risk show a clear nonlinear relationship with the
686 parameter φ . The closer the time interval, the stronger the temporal correlation of
687 inflow. The covariance and risk show a rapid increase as the parameter φ increases
688 from 0.5 to about 40, and then the variations of both stabilize (parameter φ is greater
689 than 40). Temporal correlation has a significant impact on reservoir operation risk. The
690 above findings highlight again the importance of considering temporal correlation in

691 inflow stochastic simulation and reservoir stochastic optimization.

692 This paper assumes that the relative forecast errors follow a normal distribution,
693 and most of previous studies have been based on this assumption. However, this
694 assumption does not always work in practice; for instance, Zhang et al. (2015) found
695 that the probability distribution of forecast errors for the Three Gorges Reservoir in
696 China is skewed. When the assumption of normal distribution is violated, we can
697 transform the original data by the method of normal quantile transformation, and then
698 use the framework proposed in this paper to trace the uncertainty propagation process.
699 In addition, numerical ensemble forecasting techniques have been greatly developed in
700 recent years. The increasing availability and accuracy of ensemble forecasts will
701 provide a new way for reservoir real-time optimal operation under uncertainty.

702 **Acknowledgments**

703 This work was supported by the National Natural Science Foundation of China (Grant
704 No. 52009029, 52079037); the Fundamental Research Funds for the Central
705 Universities (Grant No. B200202032); the China Postdoctoral Science Foundation
706 (Grant No. 2020T130169, 2019M661715). Data will be made available on contacting
707 the first author.

708 **References**

- 709 Ahmad, A., El-Shafie, A., Razali, S.F.M., Mohamad, Z.S., 2014. Reservoir
710 optimization in water resources: a review. *Water Resour. Manag.* 28, 3391-3405.
711 <https://doi.org/10.1007/s11269-014-0700-5>
- 712 Ahmadianfar, I., Kheyrandish, A., Jamei, M., Gharabaghi, B., 2021. Optimizing
713 operating rules for multi-reservoir hydropower generation systems: An adaptive
714 hybrid differential evolution algorithm. *Renew. Energ.* 167, 774-790.
715 <https://doi.org/10.1016/j.renene.2020.11.152>
- 716 Ajami, N.K., Duan, Q.Y., Sorooshian, S., 2007. An integrated hydrologic Bayesian
717 multimodel combination framework: Confronting input, parameter, and model
718 structural uncertainty in hydrologic prediction. *Water Resour. Res.* 43(1).
719 <https://doi.org/10.1029/2005WR004745>
- 720 Arora, A., Arabameri, A., Pandey, M., Siddiqui, M.A., Shukla, U.K., Bui, D.T.,
721 Mishra, V.N., Bhardwaj, A., 2021. Optimization of state-of-the-art fuzzy-
722 metaheuristic ANFIS-based machine learning models for flood susceptibility
723 prediction mapping in the Middle Ganga Plain, India. *Sci. Total Environ.* 750,
724 141565. <https://doi.org/10.1016/j.scitotenv.2020.141565>
- 725 Bayesteh, M., Azari, A., 2021. Stochastic optimization of reservoir operation by
726 applying hedging rules. *J. Water Res. Plan. Man.* 147(2), 04020099.
727 [https://doi.org/10.1061/\(ASCE\)WR.1943-5452.0001312](https://doi.org/10.1061/(ASCE)WR.1943-5452.0001312)
- 728 Beven, K., Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation
729 in mechanistic modelling of complex environmental systems using the GLUE

730 methodology. J. Hydrol. 249, 11-29. <https://doi.org/10.1016/S0022->
731 [1694\(01\)00421-8](https://doi.org/10.1016/S0022-1694(01)00421-8)

732 Bourdin, D.R., Nipen, T.N., Stull, R.B., 2014. Reliable probabilistic forecasts from an
733 ensemble reservoir inflow forecasting system. Water Resour. Res. 50(4), 3108-
734 3130. <https://doi.org/10.1002/2014WR015462>

735 Chen, J., Zhong, P.A., 2019. A multi-time-scale power prediction model of
736 hydropower station considering multiple uncertainties. Sci. Total Environ. 677,
737 612-625. <https://doi.org/10.1016/j.scitotenv.2019.04.430>

738 Chen, J., Zhong, P.A., Xu, B., Zhao, Y.F., 2014. Risk Analysis for Real-Time Flood
739 Control Operation of a Reservoir. J. Water Res. Plan. Man. 141(8), 04014092.
740 [https://doi.org/10.1061/\(ASCE\)WR.1943-5452.0000490](https://doi.org/10.1061/(ASCE)WR.1943-5452.0000490)

741 Chen, J., Zhong, P.A., Zhang, Y., Navar, D., Yeh, W.W.G., 2017. A decomposition-
742 integration risk analysis method for real-time operation of a complex flood
743 control system. Water Resour. Res. 53, 2490-2506.
744 <https://doi.org/10.1002/2016WR019842>

745 Cloke, H.L., Pappenberger, F., 2009. Ensemble flood forecasting: A review. J. Hydrol.
746 375, 613-626. <https://doi.org/10.1016/j.jhydrol.2009.06.005>

747 Côté, P., Leconte, R., 2016. Comparison of stochastic optimization algorithms for
748 hydropower reservoir operation with ensemble streamflow prediction. J. Water
749 Res. Plan. Man. 142(2), 04015046. [https://doi.org/10.1061/\(ASCE\)WR.1943-](https://doi.org/10.1061/(ASCE)WR.1943-)
750 [5452.0000575](https://doi.org/10.1061/(ASCE)WR.1943-5452.0000575)

751 Ding, W., Zhang, C., Peng, Y., Zeng, R.J., Zhou, H.C., Cai, X.M., 2015. An analytical

752 framework for flood water conservation considering forecast uncertainty and
753 acceptable risk. *Water Resour. Res.* 51, 4702-4726.
754 <https://doi.org/10.1002/2015WR017127>

755 Ghaith, M., Li, Z., 2020. Propagation of parameter uncertainty in SWAT: A
756 probabilistic forecasting method based on polynomial chaos expansion and
757 machine learning. *J. Hydrol.* 586, 124854.
758 <https://doi.org/10.1016/j.jhydrol.2020.124854>

759 Huang, K.D., Ye, L., Chen, L., Wang, Q.S., Dai, L., Zhou, J.Z., Singh, V.P., Huang,
760 M.T., Zhang, J.H., 2018. Risk analysis of flood control reservoir operation
761 considering multiple uncertainties. *J. Hydrol.* 565, 672-684.
762 <https://doi.org/10.1016/j.jhydrol.2018.08.040>

763 Kujawa, H., Kalcic, M., Martin, J., Aloysius, N., et al., 2020. The hydrologic model as
764 a source of nutrient loading uncertainty in a future climate. *Sci. Total Environ.*
765 724, 138004. <https://doi.org/10.1016/j.scitotenv.2020.138004>

766 Latorre, J.M., Cerisola, S., Ramos, A., 2007. Clustering algorithms for scenario tree
767 generation: Application to natural hydro inflows. *Eur. J. Oper. Res.* 181, 1339-
768 1353. <https://doi.org/10.1016/j.ejor.2005.11.045>

769 Li, H., Liu, P., Guo, S.L., Ming, B., Cheng, L., Zhou, Y.L., 2018. Hybrid Two-stage
770 stochastic methods using scenario-based forecasts for reservoir refill operations.
771 *J. Water Res. Plan. Man.* 144(12), 04018080.
772 [https://doi.org/10.1061/\(ASCE\)WR.1943-5452.0001013](https://doi.org/10.1061/(ASCE)WR.1943-5452.0001013)

773 Li, X., Guo, S.L., Liu, P., Chen, G.Y., 2010. Dynamic control of flood limited water

774 level for reservoir operation by considering inflow uncertainty. J. Hydrol. 391,
775 124-132. <https://doi.org/10.1016/j.jhydrol.2010.07.011>

776 Liu, P., Lin, K.R., Wei, X.J., 2015. A two-stage method of quantitative flood risk
777 analysis for reservoir real-time operation using ensemble-based hydrologic
778 forecasts. Stoch. Environ. Res. Risk Assess. 29, 803-813.
779 <https://doi.org/10.1007/s00477-014-0986-0>

780 Lu, Q.W., Zhong, P.A., Xu, B., Zhu, F.L., Ma, Y.F., Wang, H., Xu, S.Y., 2020. Risk
781 analysis for reservoir flood control operation considering two-dimensional
782 uncertainties based on Bayesian network. J. Hydrol. 589, 125353.
783 <https://doi.org/10.1016/j.jhydrol.2020.125353>

784 Moges, E., Demissie, Y., Li, H.Y., 2020. Uncertainty propagation in coupled
785 hydrological models using winding stairs and null-space Monte Carlo methods.
786 J. Hydrol. 589, 125341. <https://doi.org/10.1016/j.jhydrol.2020.125341>

787 Raso, L., Schwanenberg, D., van de Giesen, N.C., van Overloop, P.J., 2014. Short-
788 term optimal operation of water systems using ensemble forecasts. Adv. Water
789 Resour. 71, 200-208. <https://doi.org/10.1016/j.advwatres.2014.06.009>

790 Xu, B., Zhong, P.A., Wu, Y.N., Fu, F.M., Chen, Y.T., Zhao, Y.F., 2017. A
791 multiobjective stochastic programming model for hydropower hedging
792 operations under inexact information. Water Resour. Manag. 31, 4649-4667.
793 <https://doi.org/10.1007/s11269-017-1771-x>

794 Xu, B., Zhong, P.A., Zambon, R.C., Zhao, Y.F., Yeh, W.W.G., 2015. Scenario tree
795 reduction in stochastic programming with recourse for hydropower operations.

796 Water Resour. Res. 51, 6359-6380. <https://doi.org/10.1002/2014WR016828>

797 Yan, B.W., Guo, S.L., Chen, L., 2014. Estimation of reservoir flood control operation
798 risks with considering inflow forecasting errors. Stoch. Environ. Res. Risk
799 Assess. 28, 359-368. <https://doi.org/10.1007/s00477-013-0756-4>

800 Yan, H., Zhang, J.X., Rahman, S.S., Zhou, N., Suo, Y., 2020. Predicting permeability
801 changes with injecting CO₂ in coal seams during CO₂ geological sequestration: A
802 comparative study among six SVM-based hybrid models. Sci. Total Environ.
803 705, 135941. <https://doi.org/10.1016/j.scitotenv.2019.135941>

804 Zhang, J.H., Chen, L., Singh, V.P., Cao, H.W., Wang, D.W., 2015. Determination of
805 the distribution of flood forecasting error. Nat. Hazards. 75, 1389-1402.
806 <https://doi.org/10.1007/s11069-014-1385-z>

807 Zhang, X.Q., Liu, P., Xu, C.Y., Gong, Y., Cheng, L., He, S.K., 2019. Real-time
808 reservoir flood control operation for cascade reservoirs using a two-stage flood
809 risk analysis method. J. Hydrol. 577, 123954.
810 <https://doi.org/10.1016/j.jhydrol.2019.123954>

811 Zhou, Y.L., Guo, S.L., Xu, J.J., Zhao, X.F., Zhai, L.N., 2015. Risk analysis for
812 seasonal flood-limited water level under uncertainties. J. Hydro-Environ. Res. 9,
813 569-581. <https://doi.org/10.1016/j.jher.2015.05.001>

814 Zhu, F.L., Zhong, P.A., Sun, Y.M., Xu, B., Ma, Y.F., Liu, W.F., Zhang, D.C., Dawa,
815 J.M., 2020. A coordinated optimization framework for long-term complementary
816 operation of a large-scale hydro-photovoltaic hybrid system: Nonlinear
817 modeling, multi-objective optimization and robust decision-making. Energy

818 Convers. Manage. 226, 113543. <https://doi.org/10.1016/j.enconman.2020.113543>

819 Zhu, F.L., Zhong, P.A., Sun, Y.M., Yeh, W.W.G., 2017. Real-time optimal flood
820 control decision making and risk propagation under multiple uncertainties. Water
821 Resour. Res. 53, 10635-10654. <https://doi.org/10.1002/2017WR021480>

822 Zhu, F.L., Zhong, P.A., Xu, B., Liu, W.F., Wang, W.Z., Sun, Y.M., Chen, J., Li, J.Y.,
823 2020. Short-term stochastic optimization of a hydro-wind-photovoltaic hybrid
824 system under multiple uncertainties. Energy Convers. Manage. 214, 112902.
825 <https://doi.org/10.1016/j.enconman.2020.112902>

Figures

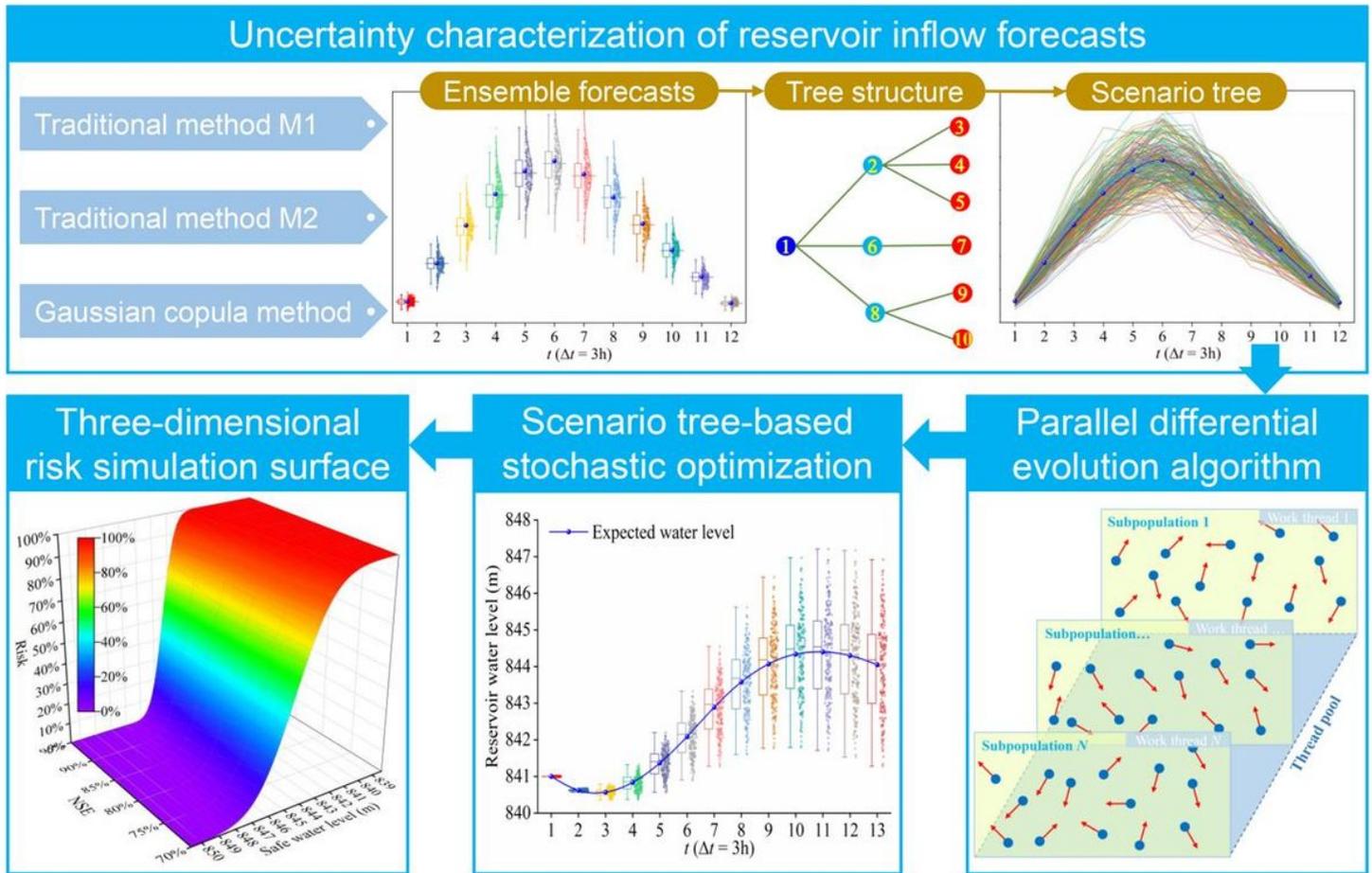


Figure 1

Flowchart of the proposed framework for tracking the uncertainty propagation process between hydrological forecasting and reservoir real-time optimal operation

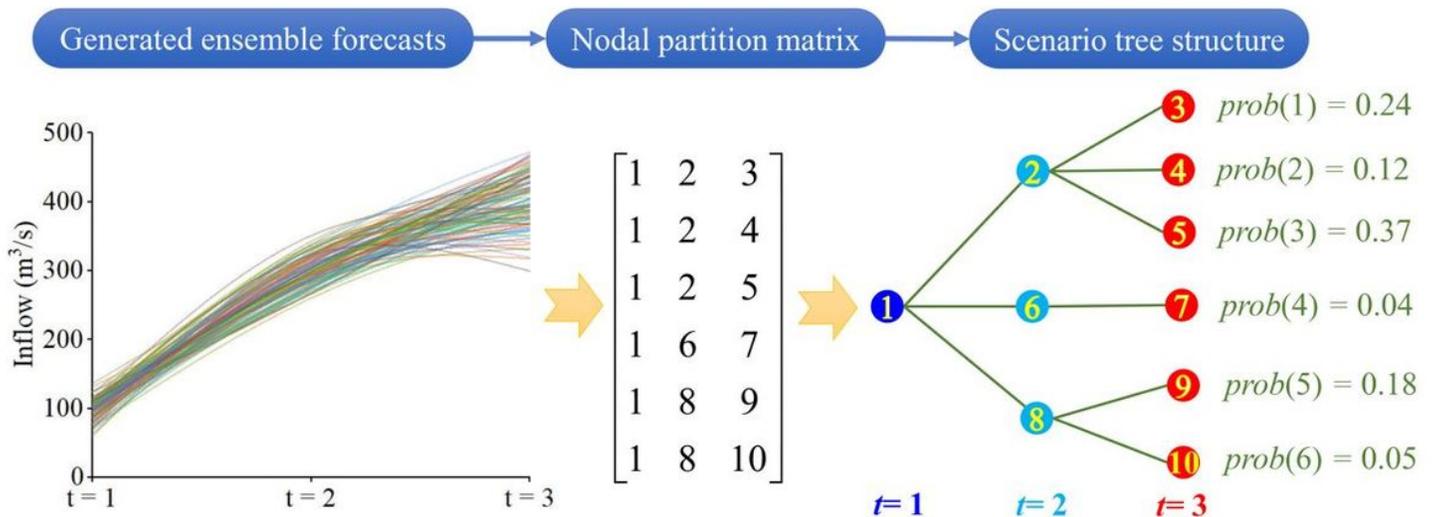


Figure 2

Schematic of a scenario tree with three stages

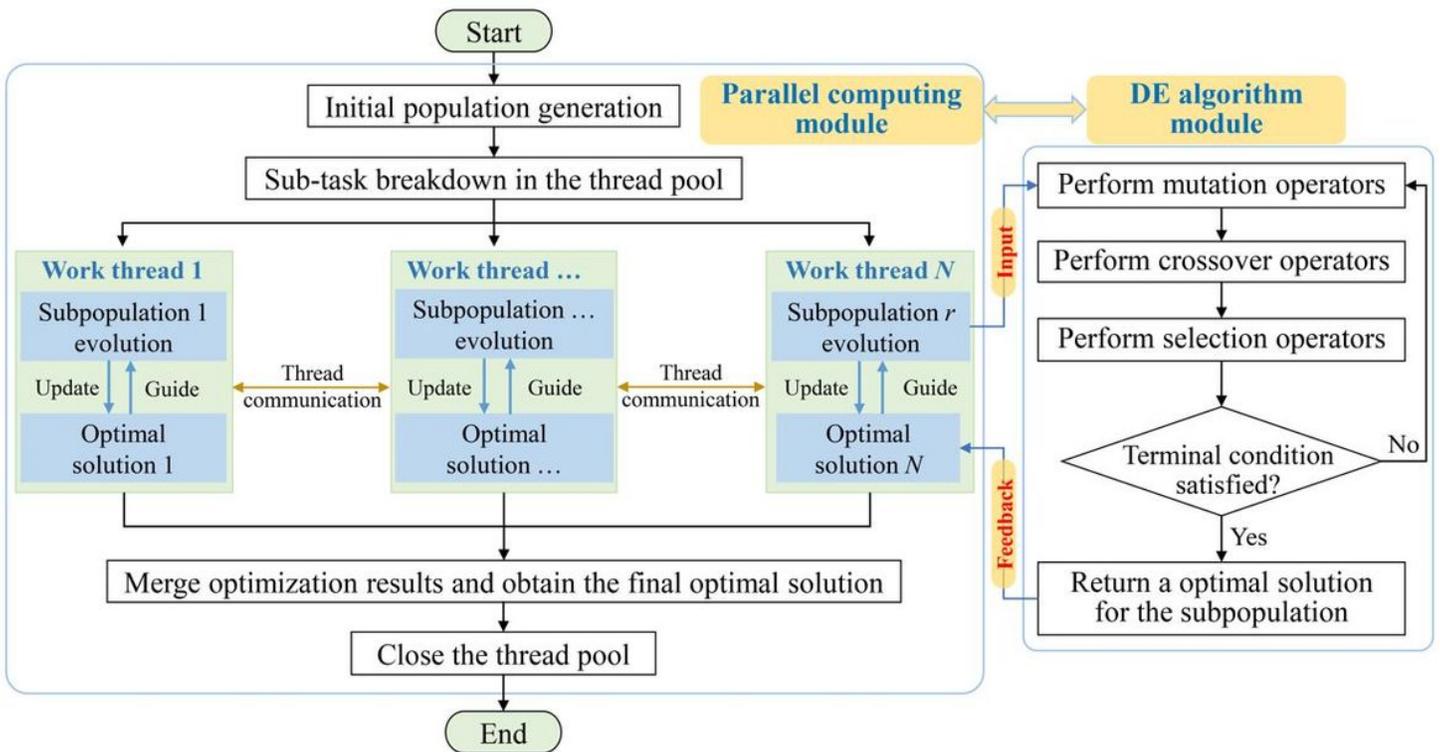
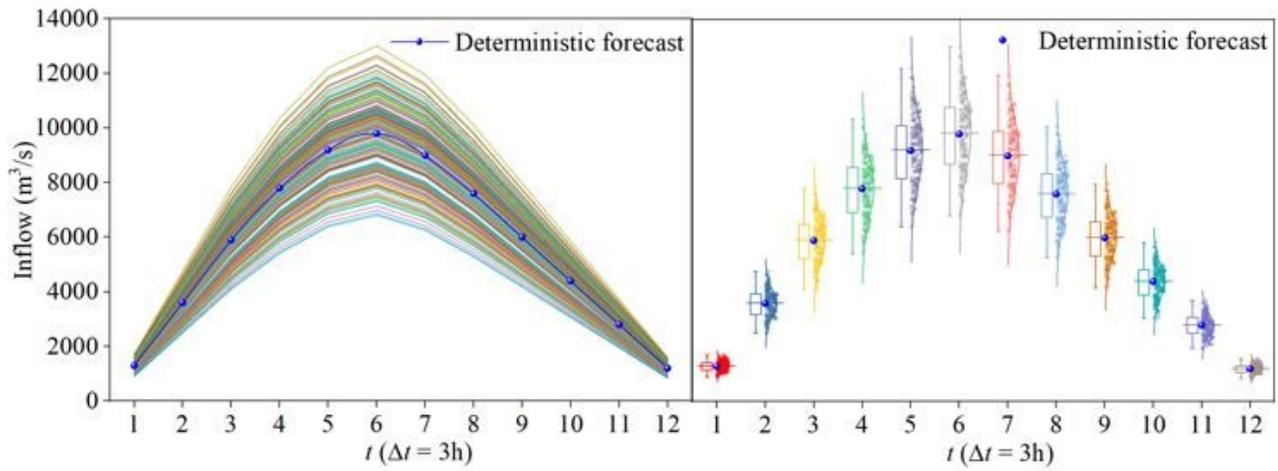
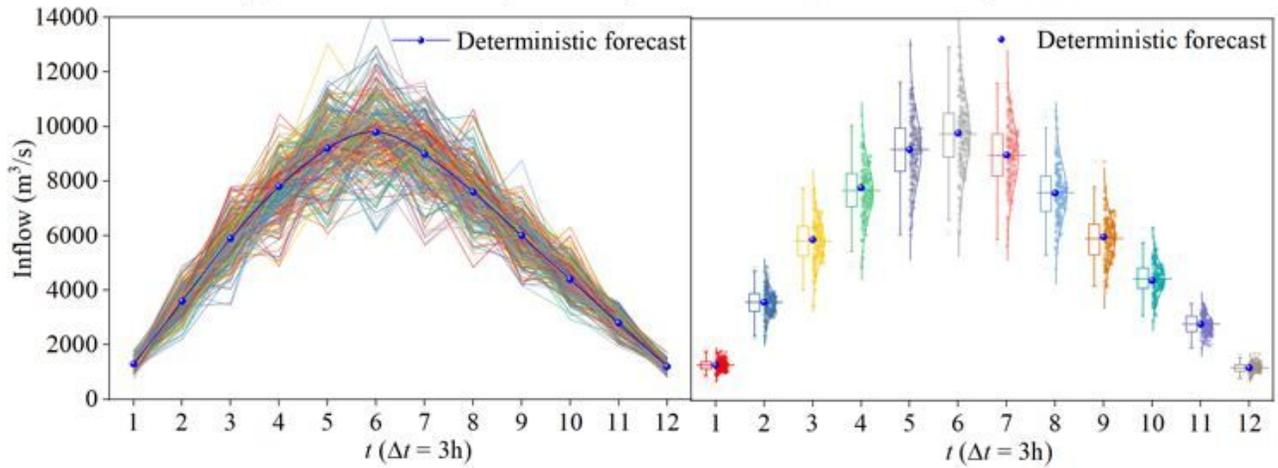


Figure 3

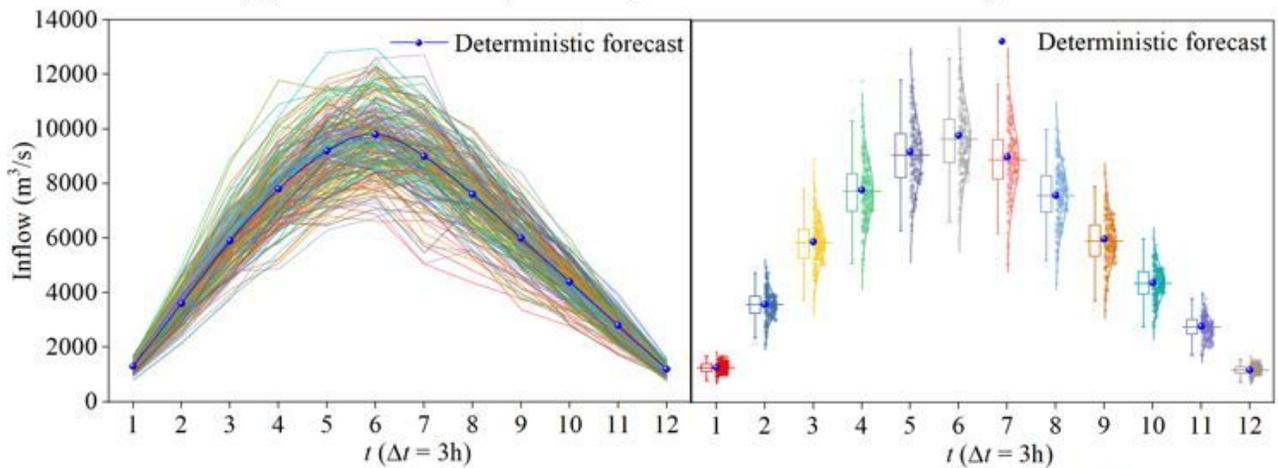
Flow chart of DE algorithm using parallel computing strategy



(a) Line chart and box plot of the generated inflow forecasts using method M1



(b) Line chart and box plot of the generated inflow forecasts using method M2



(c) Line chart and box plot of the generated inflow forecasts using Gaussian copula method

Figure 4

Line charts and box plots of the generated inflow scenario trees using three different methods

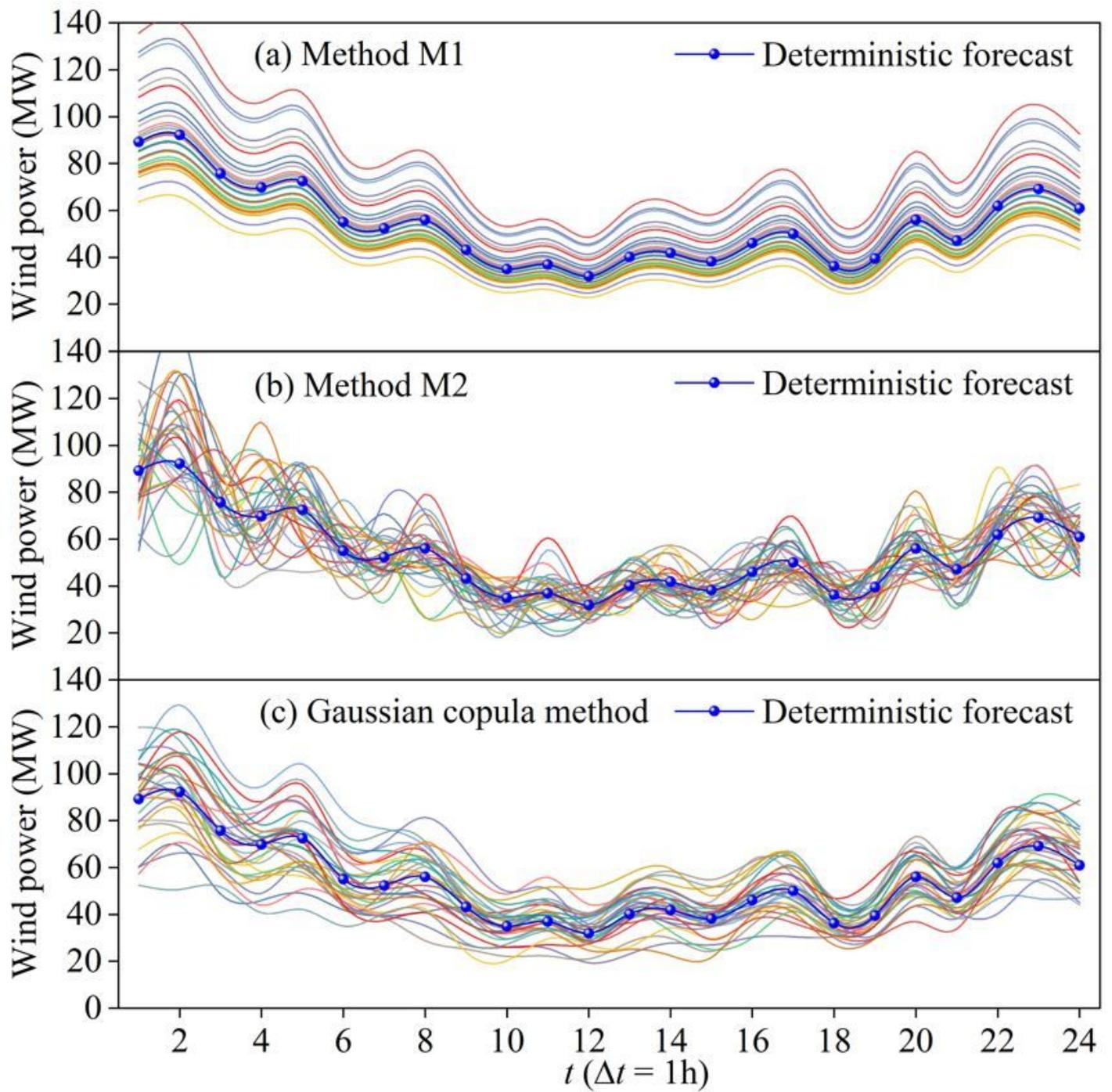


Figure 5

Comparison of the generated wind power outputs using three different methods

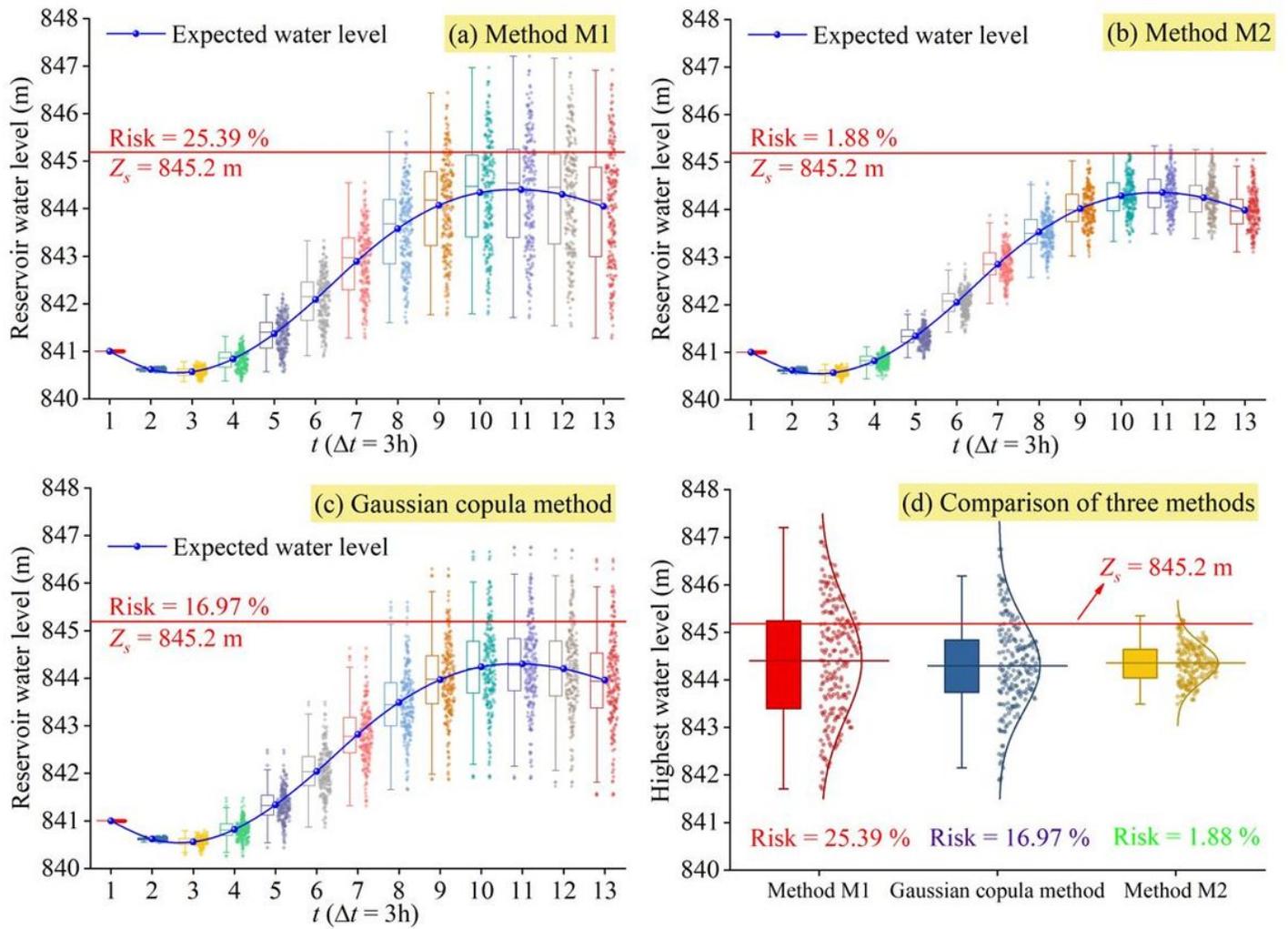


Figure 6

Results of reservoir stochastic optimal operation and risk assessment

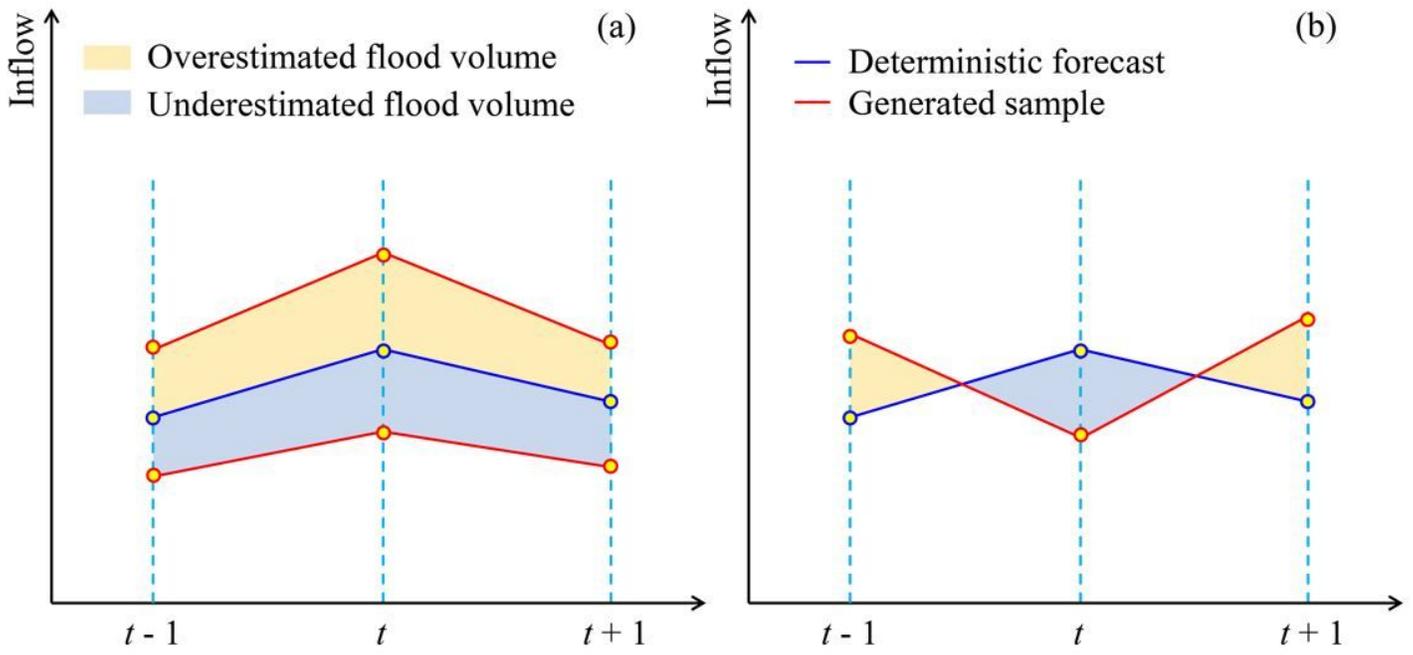


Figure 7

Diagrammatic sketch of the cumulative effect and offsetting effect of flood volume forecasting errors

forecasting errors

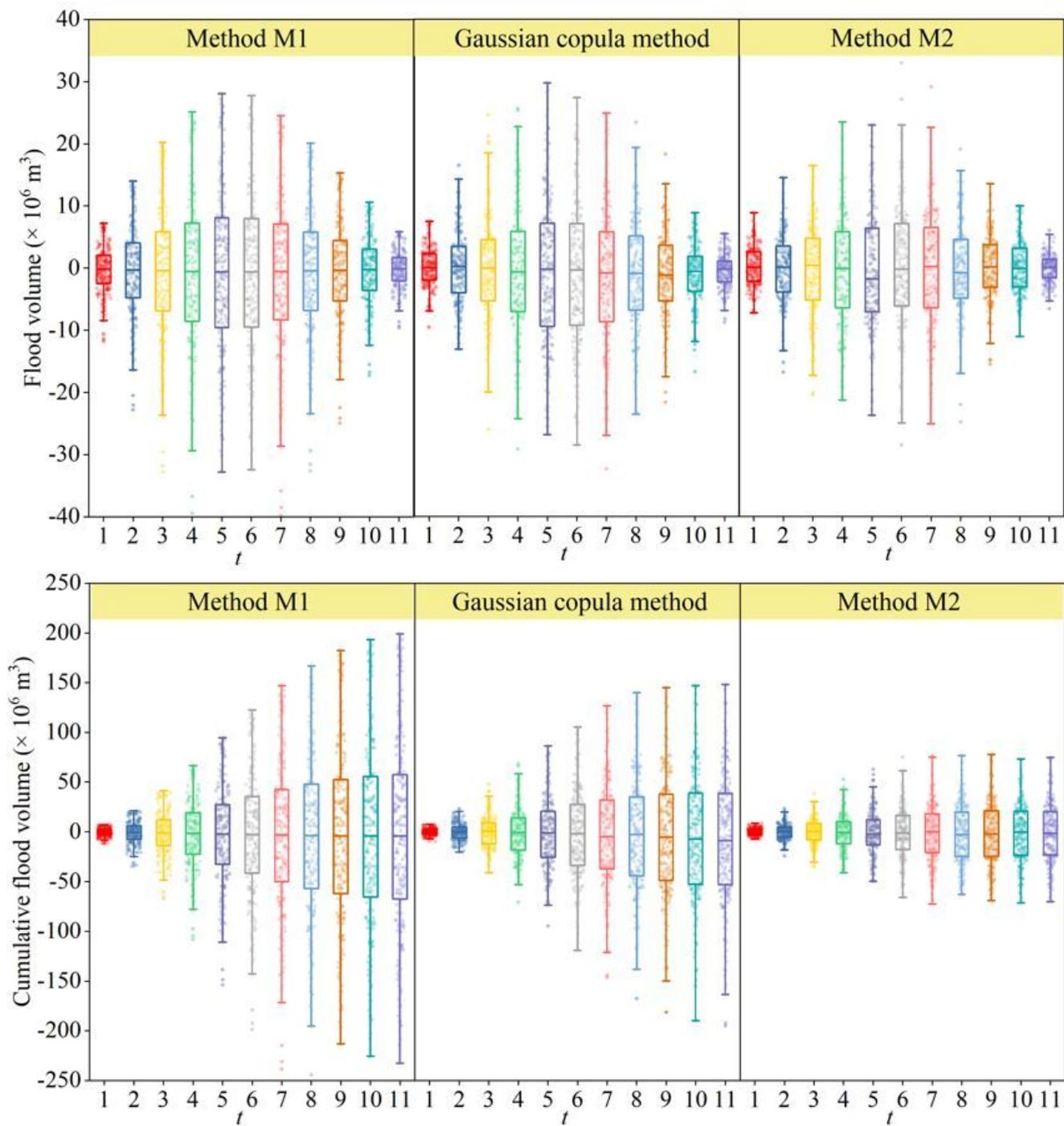


Figure 8

Comparison of flood volume and cumulative flood volume forecasting errors

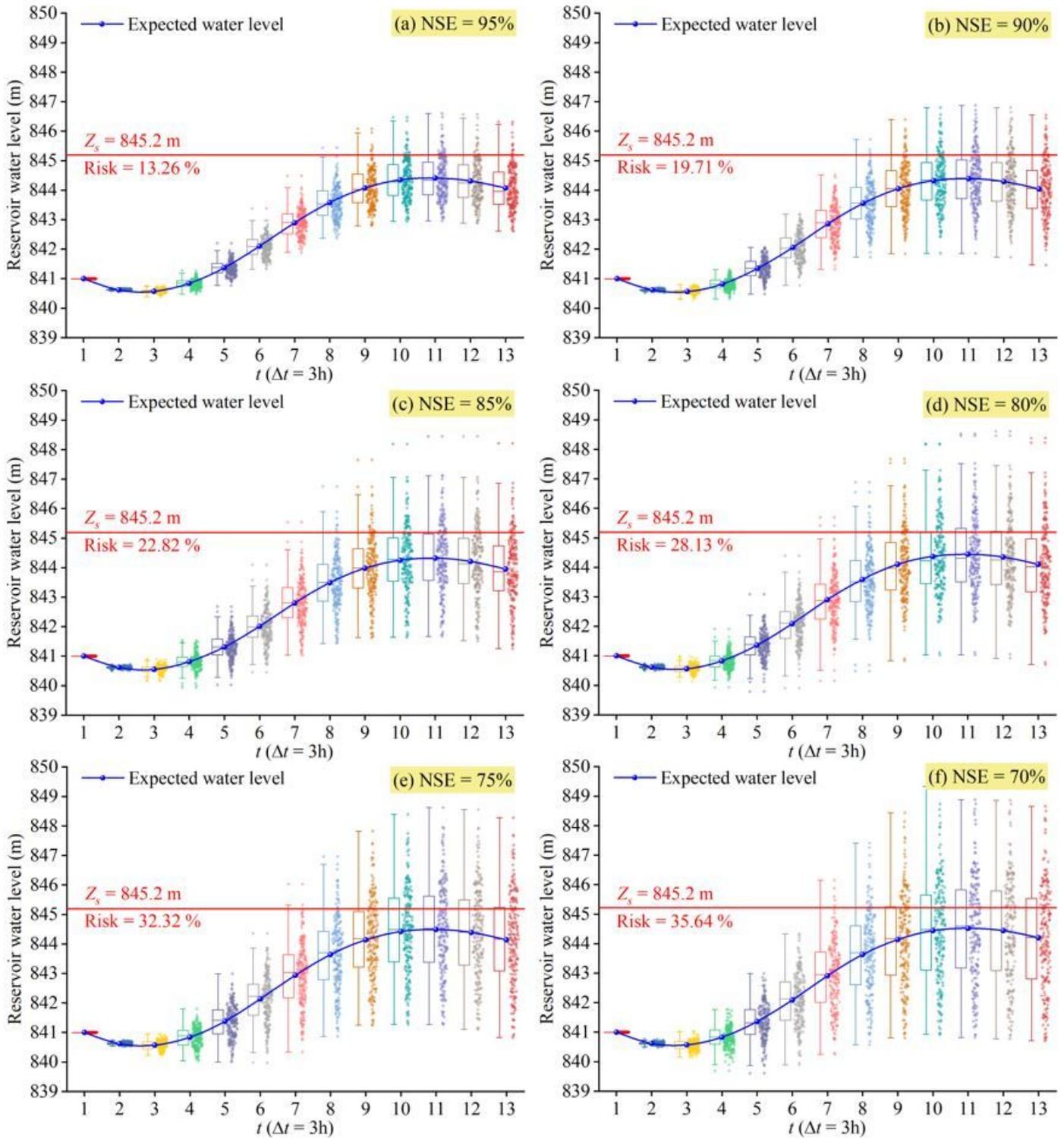
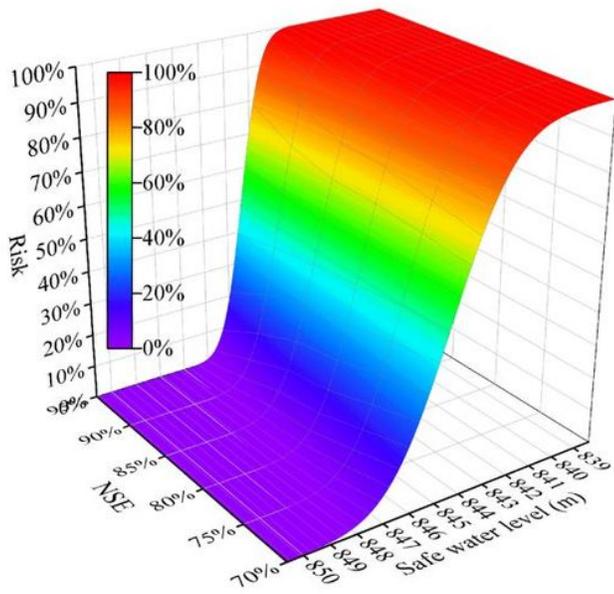
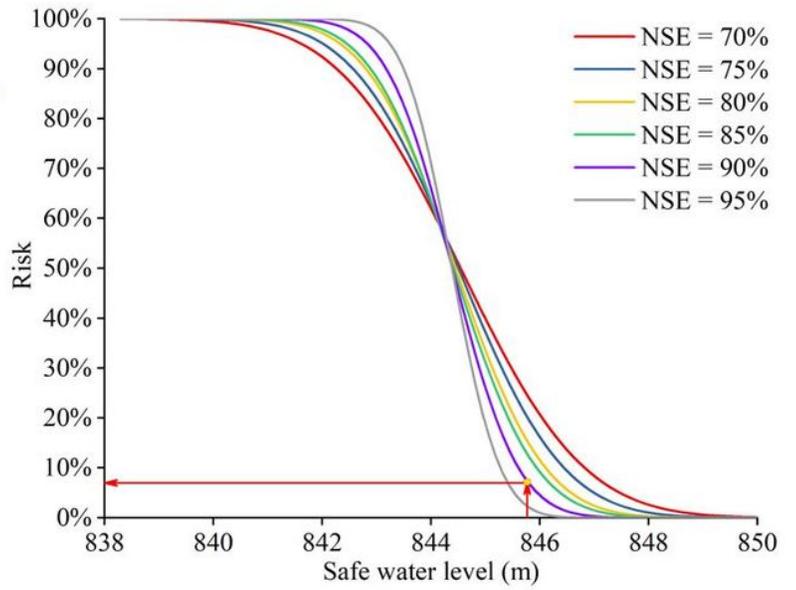


Figure 9

Box plot of reservoir water levels and related risks under different forecast uncertainty levels



(a) three-dimensional risk simulation surface



(b) two-dimensional risk projection curves

Figure 10

Relationship between safe water level and risk under different forecast uncertainty levels

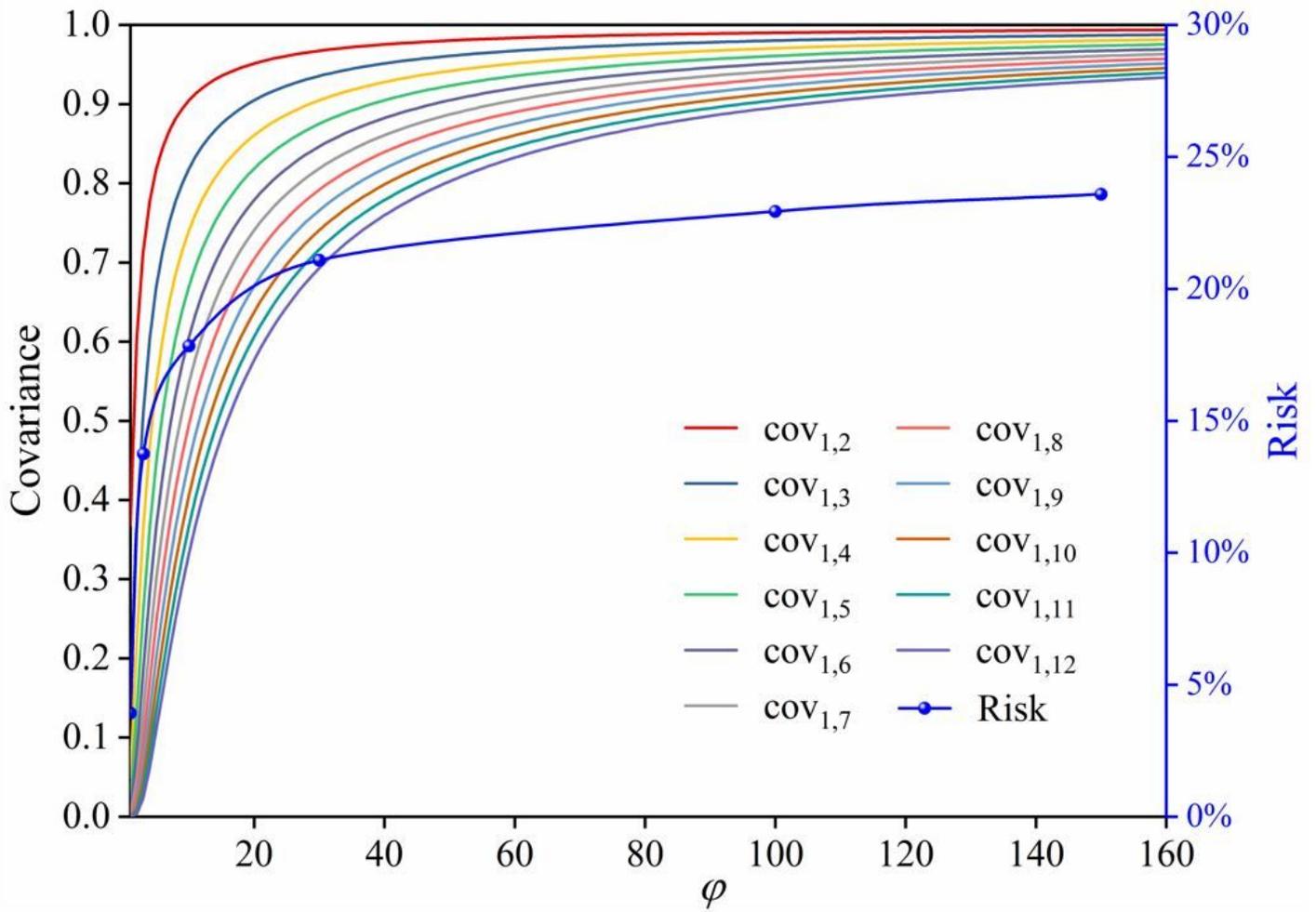
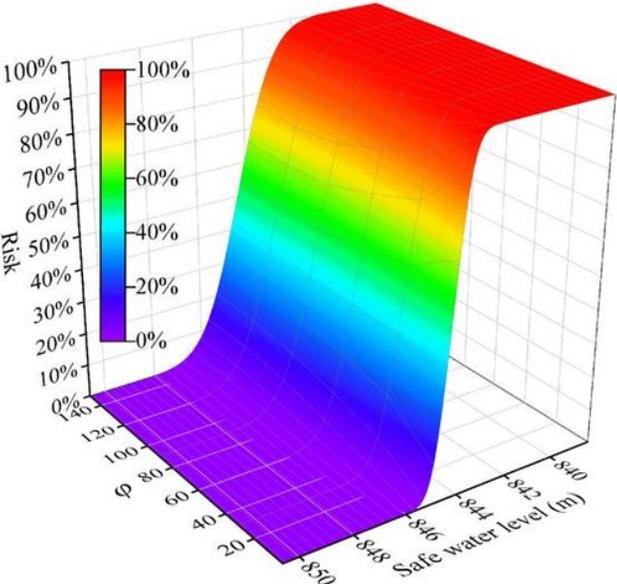


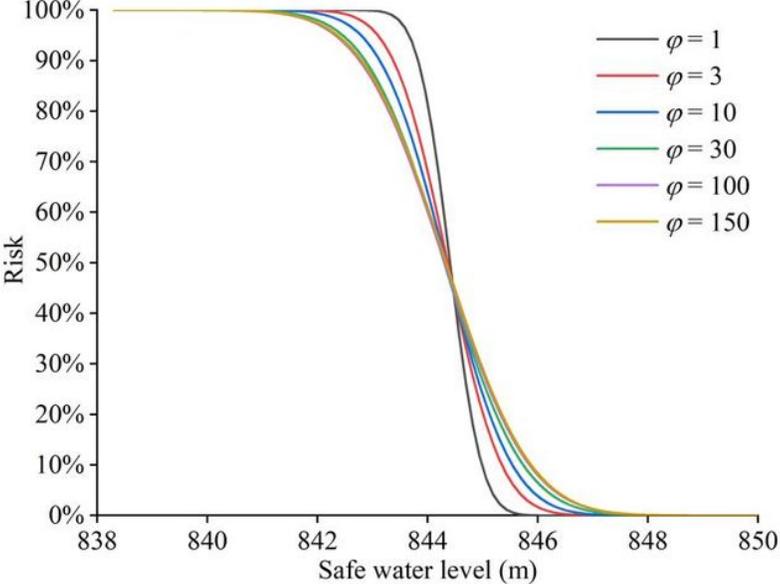
Figure 11

Covariance and risk curves under different temporal correlations of inflow forecast uncertainty

uncertainty



(a) three-dimensional risk simulation surface



(b) two-dimensional risk projection curves

Figure 12

Relationship between safe water level and risk under different temporal correlations of inflow forecast uncertainty