Available online at www.HEFJournal.org

ISSN: 2785-2997

Journal of
Human, Earth, and Future

Vol. 6, No. 2, June, 2025

A Predictive Model for Estimating Reservoir Impounding Time Based on the Net Inflow and Storage Volume

Harimukti Rosita R. [1*], Pitojo T. Juwono [1], Lily M. Limantara [1*], Emma Yuliani [1]

[1] Department of Water Resources Engineering, Faculty of Engineering, University of Brawijaya, Malang, Indonesia.

Received 07 March 2025; Revised 22 May 2025; Accepted 27 May 2025; Published 01 June 2025

Abstract

This research aims to develop a duration model for predicting the rise in water level during the initial impounding phase of a reservoir (from bed elevation to the normal storage volume elevation). The initial impounding of a reservoir represents the first test following the completion of dam construction, during which the reservoir begins to function as intended. Therefore, the expected outcome of this research is to formulate a model that estimates the duration required to fill the reservoir to its designated storage level. The study is conducted on 10 reservoirs. The methodology involves analysing existing conditions (based on observed data from initial impounding in several reservoirs), identifying influential variables, and developing a predictive model for reservoir impounding duration. The research begins with a review of elevation data recorded during the impounding process across selected reservoirs. Field data are filtered to extract continuous daily elevation records from reservoirs that underwent a single-stage impounding process. A linear regression model is employed to predict the duration of reservoir impounding, as it provides clear, interpretable results and supports accurate decision-making during the implementation phase. By using the base equation of linear regression: $\mathrm{Ln(D)} = \mathrm{m} \times \mathrm{Ln(S)} + \mathrm{n} \times \mathrm{Ln(I_{net})}$ and based on the analysis result of 28 combinations, there is selected the combination with the best determination coefficient (R^2) that is 0.99 with m = 1.047 and n = -1.08807. After being carried out the verification to 4 locations of dams that are processing the reservoir impounding, it is produced good determination coefficient and it is near to 1, so there is obtained the linear regression equation for analyzing the impounding time by using data of reservoir storage volume (S) and net inflow (net inflow = inflow to reservoir – outflow from reservoir during the impounding process) as follows: $Ln\left(D_{(rl)}\right) = 1.047 \times Ln(S) + (-1.08807) \times \mathrm{Ln(I_{net})}$.

Keywords: Duration Model; Impounding; Water Level; Reservoir; Formulation.

1. Introduction

The initial impounding of a reservoir serves as the first test to determine whether the dam will function as planned. Once dam construction is completed, the diversion tunnel is closed, allowing river flow to enter the dam site and begin filling the reservoir with water [1]. This first impounding can be defined as the rise in water level from the riverbed up to the desired operational head. The duration and rate of initial reservoir impounding can vary depending on factors such as the dam's location, type, size, and purpose. During the impounding phase, certain construction activities must be completed within a limited timeframe, as water level rises due to the closure of the diversion tunnel, hydro-mechanical operations [2], and ongoing monitoring of instruments and the dam body.

At the start of the impounding process, the diversion tunnel gate is closed, causing the reservoir's water level to gradually rise. The rate of this rise is influenced by several factors, including the season in which impounding occurs—whether dry or rainy [3, 4]. For safety reasons, particularly concerning the integrity of the dam structure, it is generally recommended that the daily increase in water level not exceed 1.0 meter per day [5]. Therefore, a predictive formulation is essential to estimate the impounding duration. This study aims to model the process by analysing observed impounding data and identifying the influencing variables.

* Corresponding author: lilymont@ub.ac.id; alwaysrosita@gmail.com

http://dx.doi.org/10.28991/HEF-2025-06-02-06

According to Tschernutter & Kainrath [6], every reservoir impounding requires an understanding of the load impact and deformation behaviour of the dam body and core. Wang et al. [7] emphasized that determining the initial impounding period and its procedure is a research priority, which should be based on a comprehensive understanding of natural conditions, including the timing of normal discharge and flood events. Meanwhile, Liu et al. [8] suggested that reservoir impounding should follow a well-designed scheme and accurate algorithm, ideally initiated after the end of the rainy season. Their findings, when applied to reservoirs in Indonesia, indicate that water level elevation increases can be more accurately predicted under normal discharge conditions.

Researchers observing the impounding process at the Three Gorges Reservoir in Chongqing Municipality, Hubei Province, China, reported that changes in water level elevation affect both downstream reservoir operations and hydrological conditions. Their study analysed the relationship between hydraulic gradient and discharge ratio in the downstream river [9]. In this context, variables such as rainfall, inflow and outflow discharge, the ratio between storage and inflow volume, watershed area, and evaporation will be examined to develop the reservoir impounding equation. A hydrological statistical approach is used in the model analysis, evaluating data from 10 reservoirs, from bed elevation to normal storage elevation. The experience of small-scale reservoir impounding (309,000 m^3) in Qionglin Reservoir on Kinmen Island, Taiwan, as reported by Hung et al. [10], concluded that rainfall plays a crucial role in impounding, especially in climates where evaporation significantly exceeds rainfall depth.

According to He et al. [11], reservoir impounding operations can be accurately modelled when initiated from either the Minimum Operating Level (MOL) or the Full Supply Level (FSL). In Figure 1, the upper limit elevation is referred to as the "annual top of buffer pool." However, in this study, the impounding model begins from the riverbed elevation. Since the initial impounding of a reservoir is a critical phase in dam development, it is essential for dam engineers to maintain control over this process. This ensures strategic oversight and the implementation of appropriate work methods, including precise supervision, observation, and analysis of instrument data.

To improve the accuracy of reservoir water level predictions, a statistical analysis approach is employed. Seasonal variations, changing rainfall patterns due to climate change, and fluctuating water demands present significant challenges in optimizing reservoir impounding [12]. Nonetheless, accurate prediction of the impounding process is vital for effective water resource design and management. Linear regression analysis is used to model the relationship between total reservoir volume, average inflow, and impounding time in order to identify a reliable predictive model. Using data from 10 different reservoirs, this study develops a series of linear regression models [13] with varying levels of complexity and evaluates their performance based on Root Mean Square Error (RMSE) [14].

This research aims to develop a model for predicting the duration of water level rise during reservoir impounding (from bed elevation to normal storage volume elevation) by analysing existing conditions (based on observed impounding data from various reservoirs) and the influencing variables. The expected outcome is a predictive model that helps minimize the time required to raise the water level from the bed to the normal storage elevation.

2. Material and Methods

2.1. Research Location

Model formulation of initial impounding is built by using the observed data elevation during the initial impounding of reservoir that are taken from some dams as presented in Tables 1 and 2, Figures 1 and 2.

Table 1. Locations for Calibration of Model Simulation

1)	Jatigede Reservoir	-6.856393° 108.096502°	Sumedang Regency, Jawa Barat
2)	Kuningan Reservoir	-7.063538° 108.706065°	Kuningan Regency, Jawa Barat
3)	Gongseng Reservoir	-7.362705° 111.901293°	Bojonegoro Regency, Jawa Timur
4)	Kuningan Reservoir	-7.063538° 108.706065°	Kuningan Regency, Jawa Barat
5)	Pidekso Reservoir	-8.036884° 110.997033°	Wonogiri Regency, Jawa Tengah
6)	Gondang Reservoir	-7.567428° 111.078189°	Karanganyar Regency, Jawa Tengah
7)	Rotiklot Reservoir	-9.067193°, 124.836406°	Sumbawa Regency, Provinsi Nusa Tenggara Timur
8)	Titab Uaran Reservoir	8° 14.664'S, 114° 56.579'E	Desa Busungbiu, Kecamatan Busungbiu, Buleleng Regency
9)	Tukul Reservoir	-8.059114° 111.138851°	Pacitan Regency, Jawa Timur
10)	Karian Reservoir	-6.413716° 106.283824°	Rangkasbitung, Lebak Regency, Provinsi Banten

Table 2. Locations for Validation of Model

a)	Jlantah Reservoir	7°42'43.36"S 111° 4'50.55"E	Karanganyar Regency, Jawa Tengah
b)	Pamukkulu Reservoir	5°23'59.14"S 119°35'39.11"E	Takalar Regency, Provinsi Sulawesi Selatan
c)	Sidan Reservoir	8°18'58.67"S 115°14'54.06"E	Badung Regency, Provinsi Bali
d)	Cipanas Reservoir	6°39'51.54"S 108° 1'40.41"E	Sumedang Regency, Provinsi Jawa Barat

Figure 1. Location Map of Indonesia Country and Reservoirs Location

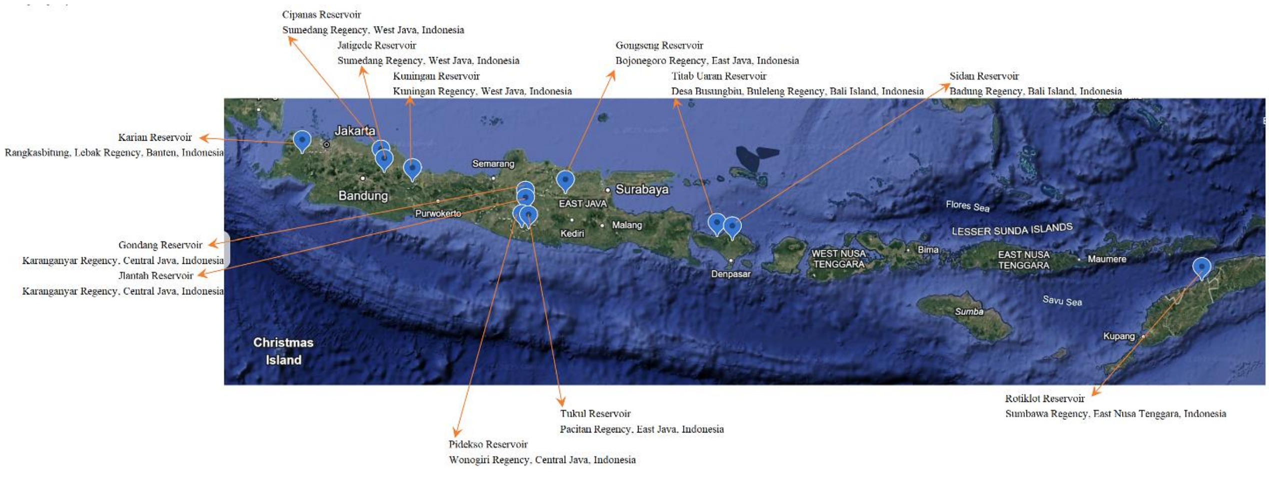

Figure 2. Location Map of The Study

2.2. Initial Impounding of Reservoir

The initial step in the reservoir impounding process is to close the diversion tunnel gate, allowing the water level to gradually rise. The rate of water level increase in the reservoir is influenced by the timing of the impounding—whether it occurs during the rainy or dry season. For safety reasons related to the dam structure, the daily rise is regulated to not exceed 1.0 meter per day. Based on observations of reservoir water level elevation during the initial impounding phase [10, 15], it has been found that a parabolic equation, specifically a first-degree polynomial, best represents the relationship between elevation and time during this process.

According to Bhadoriya et al. [16], accurately modelling reservoir impounding operations typically starts from the Minimum Operating Level (MOL) or lower limit elevation and proceeds to the Full Supply Level (FSL) or upper limit elevation. In the Figures 3 and 4, the upper limit elevation is labeled as the "annual top of buffer pool." In contrast, this study develops a model beginning from the riverbed elevation.

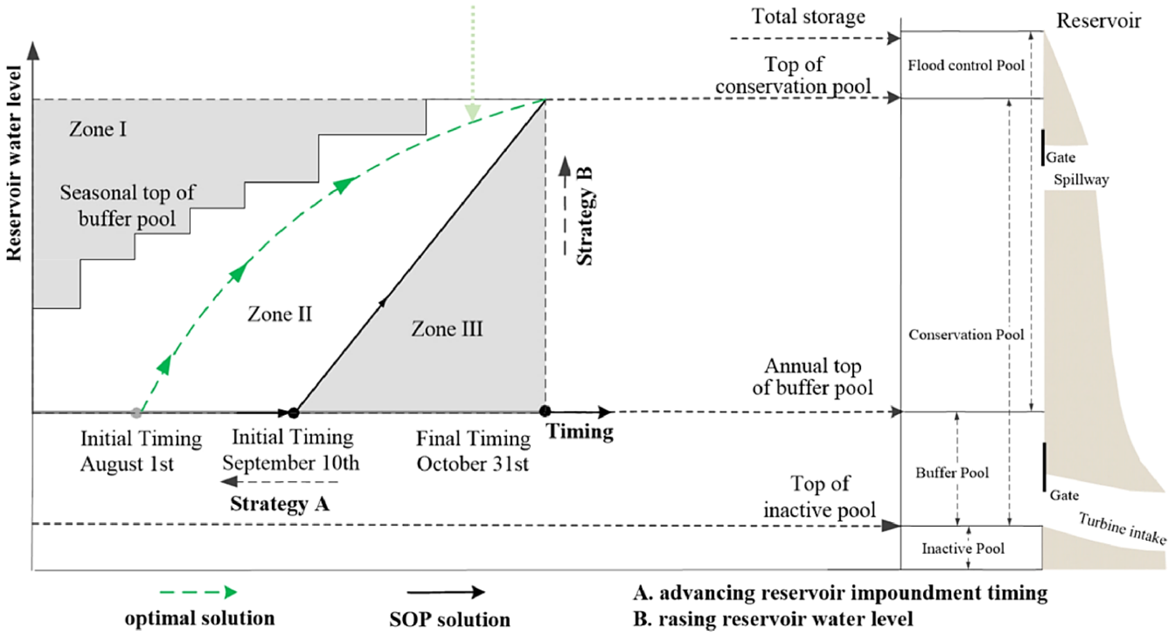

Figure 3. Process of Reservoir Impounding Operation [11]

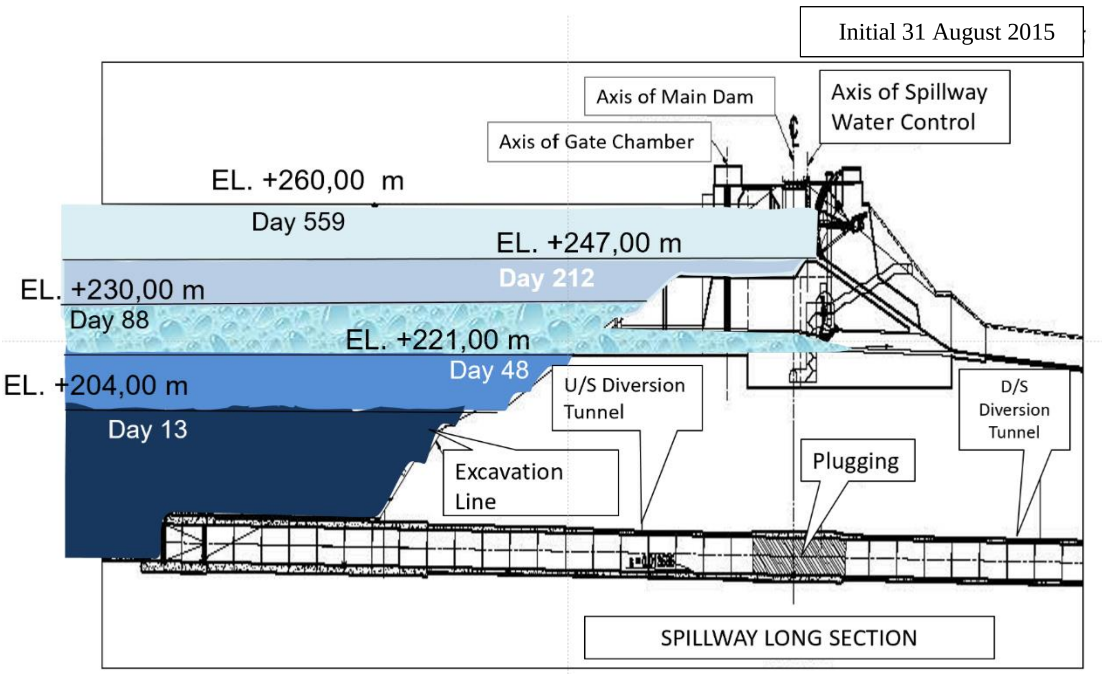

Figure 4. Illustration of Impounding Simulation

2.3. Reservoir Simulation Theory

Jain et al. [17] stated "A reservoir is operated according to a set of rules or guidelines for storing and releasing water, depending on the purposes it is required to serve. Decisions regarding water releases are made for different time periods in accordance with the demands. One commonly used method for reservoir management is based on rule curves. The following formula is used." Rule curves are frequently employed to simulate reservoir operations, and the corresponding formulation is as follows:

$$\text{Storage volume}_{\mathrm{i}} = \text{Storage volume}_{\mathrm{i-1}} + \text{Inflow} - \text{Demand} - \text{Evaporation} - \text{Spillout} \quad (1)$$

The increase in reservoir water level during the impounding period is analysed using this equation, assuming that the initial storage $V_0 = 0$. In some cases, the outflow demand is also zero, or outflow does not occur until the reservoir water level reaches the full elevation as specified in the design.

Meanwhile, according to Linsley & Franzini [18], the change in water level elevation in a natural channel is based on the principle of continuity, which is applied to river systems. The formulation is as follows:

$$\hat{y} = b + a \cdot \log X \tag{2}$$

where: $\bar{I}$ and $\bar{O}$ = average inflow and outflow for the time period of Δt. Δs = water volume change in channel that is located between the sections of inflow and outflow during the period of Δt.

If $\bar{O}$ can be measured using a flow meter structure or a staff gauge (peilschaal) and is referred to as outflow, then $\bar{I}$ can be calculated based on the analysis of Δs. The value of Δs can be approximated using the reservoir capacity curve. If the average discharge during a given period is equal to the average discharge at the beginning and end of that period, then Equation 1 simplifies to:

$$\frac{I_1+I_2}{2}\Delta t - \frac{O_1+O_2}{2}\Delta t = s_2 - s_1 \tag{3}$$

2.4. Analysis of Reservoir Impounding Model

According to Yin et al. [2], optimal initial impounding and the impounding process are influenced by the stochastic characteristics of natural river flow, and flood events must be considered during the impounding phase. Wang et al. [7] conducted a study comparing water availability (accumulated inflow) and the required design volume to estimate the reservoir impounding duration. In their model, Wang et al. [7] introduced an additional index that illustrates the potential for the highest interannual fluctuation. By developing such a model, the regulation of reservoir impounding operations can be formulated through an algorithmic approach, enabling the initial impounding to begin at the end of the flood season [3]. Based on several studies, in a stochastic context, the rise in reservoir water level elevation is influenced by both inflow into the reservoir and outflow from it.

The linear regression model is selected due to its simplicity, ease of interpretation, and effectiveness in modelling linear relationships between independent and dependent variables. Additionally, linear regression is widely applied in hydrology because of its reliability when dealing with limited datasets and its relatively stable distribution properties. Masselot et al. [19] introduced the functional linear model and applied it to hydrological forecasting, enabling the prediction of the entire flow curve rather than just specific daily or hourly points. River flow, as a natural phenomenon, exhibits continuous behaviour over time, similar to meteorological variables that influence its variability. The functional linear model operates on curves instead of discrete values, allowing it to capture the full process rather than just isolated time-based observations.

A linear regression approach was employed to develop this model due to its simplicity, interpretability, and effectiveness in capturing linear relationships. Its proven reliability in hydrology, especially when working with limited datasets, highlights its practical value. In 2016, Masselot et al. [19] introduced the functional linear model for hydrological forecasting. Unlike traditional models that predict isolated data points, this method forecasts the entire flow curve. Given that both river flow and the meteorological variables influencing it are continuous phenomena, a curve-based model allows for a more holistic analysis of the hydrological process. By accurately capturing the full temporal dynamics, it enhances forecasting accuracy and provides deeper insight into water behaviour.

The volume of reservoir impounding, in the context of initial reservoir filling, refers to the volume calculated over the duration from the river bed elevation up to the reservoir's normal full condition—also known as the Full Supply Level (FSL) or Normal Water Level (NWL). This does not include the High-Water Level (HWL), which is associated with flood conditions and overflow through the spillway. Figure 5 illustrates the volume boundary used in the model: the lower limit is the river bed elevation, and the upper limit is the full reservoir elevation, marked by red vertical lines. The height of the reservoir impounding, therefore, represents the elevation difference between these two boundaries.

2.5. Boundary Condition of Reservoir Impounding

The volume of reservoir impounding refers to the volume accumulated over a defined period, starting from the bed elevation up to the reservoir's normal full condition—also known as the Full Supply Level (FSL) or Normal Water Level (NWL). It does not include the High-Water Level (HWL), which corresponds to flood conditions when overflow occurs through the spillway. Figure 5 illustrates the volume boundaries considered in the model, marked by red vertical lines, where the lower limit represents the river bed elevation and the upper limit corresponds to the full reservoir elevation. The height of the reservoir impounding, therefore, is defined as the elevation difference between these two limits.

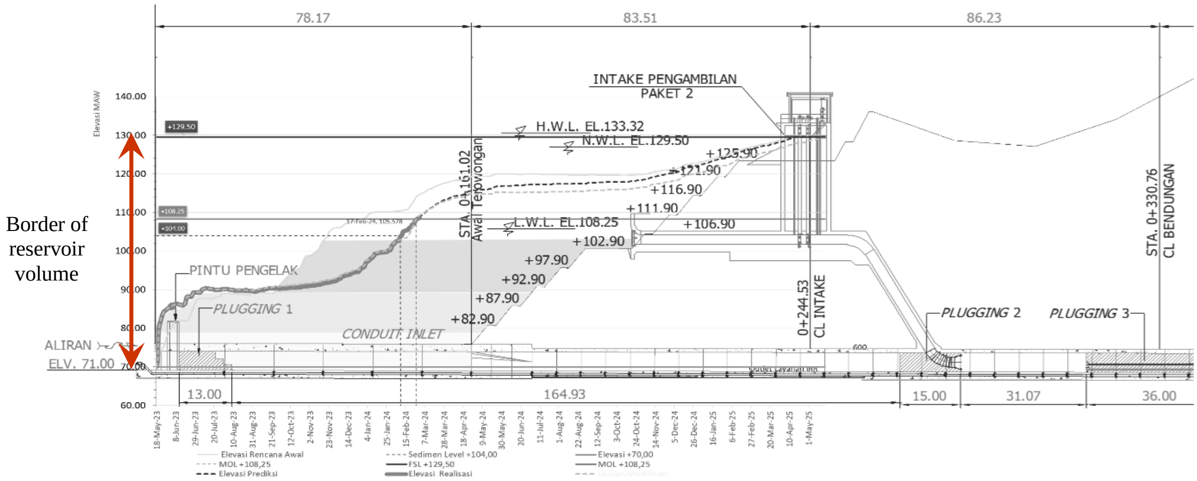

Figure 5. Illustration of Reservoir Volume Boundary that is Modelled

3. Results and Discussion

3.1. Evaluation of Initial Variable for Model

The variables and their explanations are as follows:

- **Independent variables** include storage depth, the reservoir's capacity curve, reservoir volume, and watershed area. The reservoir volume refers to the target volume at the first-stage impounding.
- **Intermediate variables** (which also serve as independent variables) include daily inflow and outflow (analysed based on demand or as a percentage of inflow), average inflow and outflow, and rainfall observed during the impounding process.
- **Dependent variables** include the design elevation (y_{2t}) at time *t* and the duration (y_1) of the water level rise from the reservoir bed elevation to the design storage volume.

Data related to the initial impounding used in the model is presented in Table 3.

Table 3. Initial Assessment of Data to Variable

No.	Name of reservoir	Watershed area (A)	Total volume (V)	Height (H)	Average inflow (I)	Average outflow (O)	Rainfall (P)	Ratio Vi/Vt	Evapo-transpiration	Duration
		x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	Y
		(km^2)	(Million m^3)	(m)	(m^3/s)	(m^3/s)	(mm)			(day)
1	Jatigede	1460.00	727.08	88,85	72.67	110.07	2559.10	3.16	1533.00	472.00
2	Kuningan	23.07	24.65	28.10	1.03	0.09	942.64	1.62	1447.75	287.00
3	Gongseng	51.21	22.43	22.09	2.32	0.00	251.77	3.13	1443.47	114.00
4	Bendo	120.63	43.46	67.82	2.52	0.00	867.20	4.18	1635.81	203.00
5	Pidekso	55.00	24.55	25.55	0.93	0.78	880.44	2.85	1262.98	123.00
6	Gondang	19.18	7.09	36.12	0.79	0.26	2806.00	6.23	1410.23	157.00
7	Lolak	73.11	16.23	58.00	3.31	0.00	1022.50	6.43	1669.66	213.00
8	Titab	69.54	12.80	46.00	8.13	6.21	674.80	13.53	1861.86	66.00
9	Tukul	47.10	8.63	63.71	1.03	0.67	2773.00	6.50	1638.55	371.00
10	Karian +57.50	288.00	126.49	37.50	11.96	1.84	1209.92	3.12	1200.00	145.00
11	Sindang Heula	73.47	9.99	20.48	1.21	0.24	2430.59	4.17	1277.50	115.00

According to Chen et al. [3], through model development, the formulation for regulating reservoir storage operations can be constructed using an algorithmic approach, and the initial impounding time can be further refined. Based on the discussion above, the rise in water level elevation is influenced by reservoir inflow (in a stochastic context) and outflow. Table 4 presents the correlation between daily elevation and variables x_4, x_5, and x_6. Conversely, the ratio of Vi/Vt (x_7) and evapotranspiration (x_8) do not show a significant correlation. The linear regression model for impounding does not explicitly incorporate the sedimentation variable, which is indirectly represented through reservoir volume (x_2) based on its capacity curve. This implies that only one type of reservoir capacity curve is used per reservoir to determine the storage volume variable (x_2).

Table 4. Initial Assessment of Correlation between Daily Elevation and Variables of x_4, x_5, x_6

No.	Name of reservoir	Watershed area (A)	Total volume (V)	Tinggi (H)	Average inflow (I)	Average outflow (O)	Rainfall (P)	Ratio Vi/Vt	Evapo-transpiration	Duration
		x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	Y
Correlation to Y		0.676	0.688	0.500	0.657	0.693	0.507	-0.301	0.181	

However, the approach by using daily elevation is presented in Table 5:

Table 5. Correlation between Daily Elevation and x_4, x_5, x_6

No.	Name of dam	Correlation to		
		Daily inflow (I) (x_4)	Daily outflow (O) (x_5)	Daily rainfall (P) (x_6)
1	Jatigede	0.58	0.69	0.23
2	Kuningan	0.55	-0.02	-0.02
3	Gongseng	0.46	0.00	0.22
4	Bendo	0.52	0.00	0.32
5	Pidekso	0.60	0.96	-0.02
6	Gondang	0.45	-0.38	0.40
7	Rotiklot	-0.81	0.00	0.12
8	Tukul	0.40	0.51	0.22

Based on Tables 4 and 5, the variables of storage volume (x_2), inflow (x_4), and outflow (x_5) are more significant than daily rainfall (x_6) to the reservoir elevation change.

3.2. Model Analysis of Reservoir Impounding Time with Linear Regression

This analysis represents a stage in developing the reservoir impounding time model using a linear regression approach. Consistent with the work of Masselot et al. [19], this model incorporates a formulation that linearly relates time to inflow, outflow, and reservoir volume variables. However, when the model is used to determine the elevation at a specific time, a parabolic relationship emerges—an aspect that will be explored in future research.

A linear regression model may require modifications for reservoirs located in different climate zones or geographical regions, as factors such as precipitation, temperature, and geological conditions significantly influence water flow dynamics. Studies have shown that hydrological modelling must be location-specific to produce more accurate results [20, 21]. Research by Schrunner et al. [22] introduces the Gaussian sliding window regression model, which offers an alternative approach to hydrological inference that is more adaptable to various geographical and climatic conditions.

To solve the linear regression model (i.e., to analyse parameters *m* and *n*), six data sets are required, including duration (y), storage volume (x_2), reservoir inflow (x_4), reservoir outflow (x_6), and net inflow (x_4 net), which is calculated based on the inflow and outflow values. The foundational equations for developing the model with these variables include the exact equation, regression equation, and the linear regression equation, as follows:

$$D_{eksak} = \frac{S}{I_{net}} \tag{4}$$

where: D_{eksak} is duration based on the exact equation (day); S is volume of storage (1000 m^3); I_{net} is net inflow = (I – O) (1000 m^3/ day); I is inflow to reservoir (1000 m^3/ day); O is outflow from reservoir (1000 m^3/ day).

$$D_{regresi} = (S)^m \times (I_{net})^n \tag{5}$$

where: $D_{regresi}$ is duration based on the regression equation (day); m is parameter (to be calculated); n is parameter (to be calculated).

Both Equations 4 and 5 are exact equations, however the linear regression equation is as follows:

$$Ln\left(D_{(regresi\ linier)}\right) = m \times Ln(S) + \mathrm{n} \times Ln(I_{net}) \tag{6}$$

where: $D_{regresi\ linier}$ = duration based on the linear regression equation (day).

The steps for applying the combined data set linear regression method include: first, combining the data sets; then conducting the linear regression analysis; evaluating the coefficient of determination (R^2); selecting the best combination; and finally, determining the values of *n* and *m* based on the highest R^2 value that is closest to 1.

3.3. Combination of Data Set

To solve the linear regression model (for calculating the parameters *m* and *n*), six sets of data (D, S, I, and O) are required. This model uses eight data sets, which results in 28 possible combinations based on the binomial coefficient C(8,6) = 28. Therefore, from the eight data sets (D, S, I, and O), 28 alternative combinations are generated. Tables 6 and 7 present the data used for the linear regression application.

Tabel 6. Data for Application of Linear Regression-1 Model

No.	No. of selected raw	Name of reservoir	D	S	I	O
			Duration	Volume of reservoir storage	Average inflow	Average outflow
			[day]	[million m³]	[m³/s]	[m³/s]
1	2	Kuningan	287.00	24.65	1.03	0.09
2	3	Gongseng	114.00	22.43	2.32	0.00
3	4	Bendo	203.00	43.46	2.52	0.00
4	5	Pidekso	123.00	24.55	0.93	0.78
5	6	Gondang	157.00	7.09	0.79	0.26
6	8	Lolak	213.00	16.23	3.31	0.00
7	10	Tukul	371.00	8.63	1.03	0.67
8	12	Karian +57.50	145.00	126.49	11.96	1.84

Table 7. Tabel of Data for Application of Linear-Regression-2 Model

D [day]	S [million m³]	I_net [million m³/day]
287.00	24650.0	80.96
114.00	22430.0	200.10
203.00	43458.0	217.73
123.00	24550.0	12.69
157.00	7090.0	45.79
213.00	16230.0	285.80
371.00	8630.0	31.10
145.00	126490.0	874.35

3.4. Analysis of Linear Regression

For every alternative of combination, there is carried out the analysis of linear regression (with the function of LINEST) that produces the coefficient of *n* and *m* as presented in Table 8.

Table 8. Analysis of Linear Regression Due to the Function of LINEST

Ln(D)	Ln(S)	Ln(I_net)
5.31321	10.67955	5.38325
4.81218	10.10847	2.54051
5.05625	8.86644	3.82411
5.36129	9.69462	5.65528
5.91620	9.06300	3.43734
4.97673	11.74792	6.77348

	n	m
	-0.1550	0.5880
	0.30	0.14
R^2	0.98144	0.88
	105.76	4.00
	162.39	3.07

From the analysis of each combination alternative, the following are obtained: (1) the coefficient of determination (R^2); (2) parameter *m*; and (3) parameter *n*. The combination alternative with the highest coefficient of determination—i.e., the value closest to 1—is considered the best. Table 9 presents the analysis results of R^2, *n*, and *m* for each combination.

Table 9. Analysis of R^2, n, and m from Each Combination

	R^2
Average	0.98783
Max.	0.99966
Min.	0.98024

No.	Calculated Parameter		
	R^2	m	n
1	0.99454	0.47394	0.09857
2	0.98692	0.56107	-0.06790
3	0.99070	0.57213	-0.17071
4	0.98747	0.53242	-0.00421
5	0.99049	0.51273	-0.04192
6	0.98404	0.62854	-0.27278
7	0.99437	0.82299	-0.57631
8	0.99308	0.80779	-0.58775
9	0.99966	1.04700	-1.08807
10	0.99226	0.91533	-0.78588
11	0.98783	0.52832	0.03345
12	0.98783	0.53694	-0.07494
13	0.98228	0.64128	-0.28682
14	0.98120	0.59289	-0.17170
15	0.99249	0.91229	-0.77309
16	0.98999	0.50968	0.08435
17	0.98870	0.53249	-0.05780
18	0.98311	0.63444	-0.26478
19	0.98234	0.58774	-0.15399
20	0.99287	0.90142	-0.75040
21	0.98133	0.59667	-0.15418
22	0.98675	0.53344	-0.00590
23	0.98999	0.51054	-0.03976
24	0.98324	0.62824	-0.27357
25	0.98207	0.57583	-0.15342
26	0.99209	0.92909	-0.80968
27	0.98024	0.59293	-0.17248
28	0.98144	0.58801	-0.15501

3.5. Evaluation of Determination Coefficient (R^2)

The combination alternatives of data sets with the highest coefficient of determination (R^2)—that is, the value closest to 1—are selected as the best combinations. These combinations provide the most accurate linear regression model in illustrating the relationship among the variables D, S, I, and O. Figure 6 presents a ranking of the combinations from highest to lowest based on their R^2 values, identifying the best composition accordingly. Additionally, the ranking of the determination coefficients is illustrated in a graph.

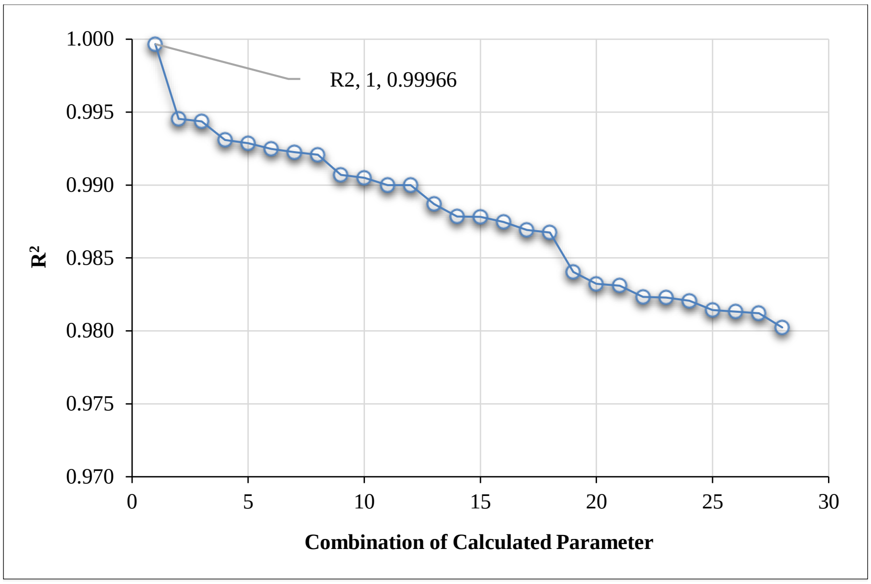

Figure 6. Graphic of 28 Combinations of n and m and R^2 (Determination Coefficient)

3.6. Selection of the Best Combination

The combination alternatives of data sets with the highest coefficient of determination (R^2)—i.e., those closest to 1—are selected as the best combinations. These provide the most accurate linear regression model for illustrating the relationship among the variables D, S, I, and O. This process enables the identification of the most reliable model for representing the factors influencing the reservoir's initial impounding time. By using the optimal combination of data sets (Table 10), the accuracy of predicting the variables that affect the impounding duration can be significantly improved.

Table 10. Parameter of the Best Model

The best	Parameter of the best model			No.
	R^2	m	n	Composition
1	0.99966	1.04700	-1.08807	9

Based on the analysis result of linear regression, there is obtained the best determination coefficient (R^2) is 0.99966 or number 9 of compositions, with the value of m = 1.047 and n = -1.08807.

The comparison results between observed reservoir impounding time in each dam with the exact method, and by applying n and m from the linear regression, is presented in Table 11.

Table 11. Application Result Comparison of Linear Regression Coefficient between Observed Data and Exact Formulation

No.	Composition	Reservoir	Comparation of duration		
			$D_{\text{linear regression}}$ (day)		
			Observed data	Regression model	Formulation (S)/(I-O)
1	1	Kuningan	287.00	332.59	304.48
2	2	Gongseng	114.00	112.56	112.09
3	3	Bendo	203.00	205.23	199.60
4	5	Gondang	157.00	167.71	154.83
5	7	Tukul	371.00	313.84	277.46
6	8	Karian +57.50	145.00	138.38	144.67

Based on the analysis, the difference between the observed data and the linear regression model application is 0.43%, while the exact calculation yields a difference of -4.01%. Therefore, the linear regression model with parameters *m* = 1.047 and *n* = -1.08807 can be effectively used for modelling the reservoir impounding time, as it results in a minimal average error. Table 12 presents the comparison between the observed data, the regression model, and the exact formulation.

Table 12. The Difference between Observed Data, Regression Model, and Exact Formula

No.	Composition	Reservoir	Comparison of duration D (day): Recorded data	Regression model	Formula (S)/(I-O)	Percentage error of model	Percentage error of formula
1	1	Kuningan	287.00	332.59	304.48	15.89%	6.09%
2	2	Gongseng	114.00	112.56	112.09	-1.26%	-1.67%
3	3	Bendo	203.00	205.23	199.60	1.10%	-1.68%
4	5	Gondang	157.00	167.71	154.83	6.82%	-1.38%
5	7	Tukul	371.00	313.84	277.46	-15.41%	-25.21%
6	8	Karian +57.50	145.00	138.38	144.67	-4.56%	-0.23%
						0.43%	-4.01%

Based on the analysis above, the equation of exact regression model that can be used is as follows:

$$D_{\text{regression}} = (S)^{1.047} \times (I_{net})^{n\text{-}1.08807} \quad (7)$$

where: $D_{\text{regression}}$ is duration based on the regression equation (day); m is parameter =1.047; n is parameter = -1.08807.

Based on the equation-7, the linear regression is as follows:

$$Ln\left(D_{(linear\ regression)}\right) = 1.047 \times Ln(S) + (-1.08807) \times Ln(I_{net}) \quad (8)$$

where: $D_{linear\ regression}$ is duration based on the linear regression equation (day).

The analysis results indicate that the linear regression model can predict inflow with reasonable accuracy under normal conditions. However, we recognize that this model may not fully account for substantial inflow fluctuations caused by extreme weather events. Linear regression relies on the assumption of a linear relationship and homogeneity of variance, both of which can be affected by outliers or extreme data points. To address this limitation, we plan to investigate alternative approaches—such as robust regression or non-parametric models—and incorporate variables that represent extreme weather conditions in future studies.

3.7. Model Validation

Cross-validation of the model was performed to evaluate its generalizability beyond the initially selected reservoirs. This validation involved observing four additional reservoirs in Indonesia that were undergoing the impounding process between 2023 and 2024. The assessment included daily monitoring of water level, outflow, and technical data to calculate daily inflow. By applying the derived equation Ln(D) = 1.047 × Ln(S) + (−1.08807) × Ln (Inet), the estimated reservoir impounding time closely matched the observed duration.

Moriasi et al. [23] recommend using three quantitative metrics—Nash-Sutcliffe Efficiency (NSE), Percent Bias (PBIAS), and the Ratio of the Root Mean Square Error to the Standard Deviation of measured data (RSR)—along with graphical techniques for model evaluation. The graphical results demonstrated a correlation exceeding 98% between the observed data and the model-predicted impounding duration. The model yielded the highest R^2, confirming its suitability as the best-performing model for the dataset used in this study. Table 13 presents the validation results from the four additional reservoirs, and Figure 7 illustrates the graphical comparison between the observed and calculated durations.

Table 13. Validation Result with Four Additional Reservoirs

No.	Reservoir Name	Initial Impounding date	Impounding Duration: D (observed) (days)	D (model) (days)	R^2
1	Jlantah Dam	20-Dec-24	46.00	44.32	98.43%
2	Pamukkulu Dam	19-Apr-24	257.00	257.86	
3	Sidan Dam	18-Nov-24	64.00	61.97	
4	Cipanas Dam	19-May-23	296.00	257.26	

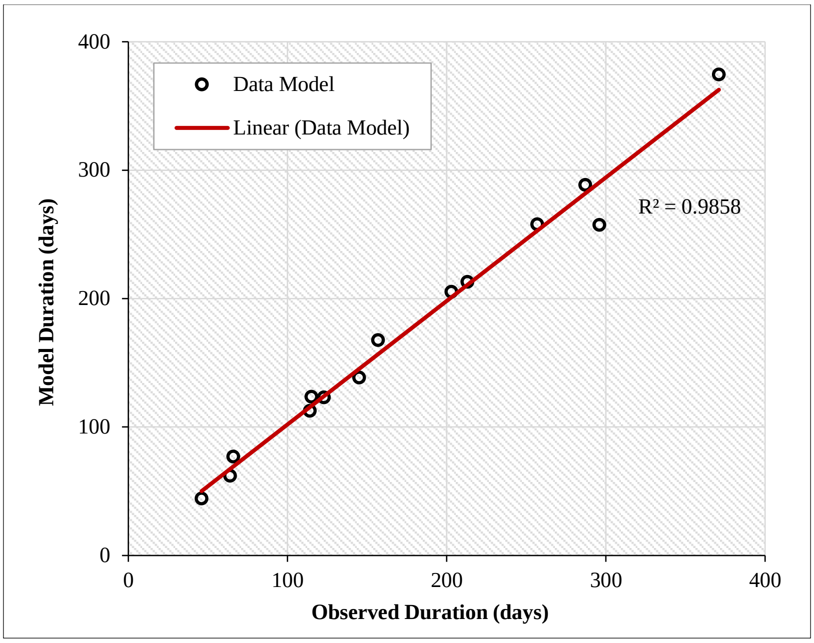

Figure 7. Graphical of Comparison between Observed and Calculated Duration

The performance of the calibrated model was evaluated using the Nash–Sutcliffe Model Efficiency (NSE). Skhakhfa & Ouerdachi [24] applied the NSE to assess the level of agreement between observed and simulated values. The NSE is commonly used to measure how well the model predictions match the observed data. Mathematically, it is expressed as:

$$NSE = 1 - \frac{\sum_{i=1}^{n}(OBS_I - SIM_i)^2}{\sum_{i=1}^{n}(OBS_I - \overline{OBS})^2} \tag{9}$$

Where: NS is Nash-Sutcliffe Efficiency (NS); SIM_i represents the predicted values (model outputs); OBS_I represents the observed values (data points); $\overline{OBS}$ is the mean of the observed values (m^3/s); n is the total number of observations.

The model's performance was evaluated using three statistical measures: the Nash-Sutcliffe Efficiency (NSE), Percent Bias (PBIAS), and the Ratio of the Root Mean Square Error to the Standard Deviation of Observed Data (RSR). In general, model simulation can be judged as satisfactory if NSE > 0.50 and RSR < 0.70, and if PBIAS + 25% for streamflow, PBIAS + 55% for sediment, and PBIAS + 70% for N and P [23]. By using the Equation 9, $\sum_{i=1}^{n}(OBS_I - SIM_i)^2$ is1879.18, $\sum_{i=1}^{n}(OBS_I - \overline{OBS})^2$ is found of 126641.50, NSE = 0.985, Nash–Sutcliffe efficiencies can range from ∞ to 1. An NSE value of 0.985 indicates that the model possesses excellent predictive capability, as it is very close to 1. This suggests that the simulated values closely align with the observed data. The PBIAS value of -0.50% reflects minimal bias, implying that the model slightly underestimates the observed values. Additionally, the RSR value of 0.117, which is well below the threshold of 0.50, confirms the model's high accuracy and strong predictive reliability. These results demonstrate that the regression-based model is highly effective in estimating reservoir impounding time with minimal error.

While a simple approach to estimate reservoir filling duration involves dividing the reservoir volume by the inflow discharge, this method is less accurate. During the impounding process, it is essential to regulate outflow discharge to control the rate of water level rise, ensuring it does not exceed 1.0 meter per day for the safety of embankment dams. Therefore, the validated model equation offers a more accurate estimate of the required duration.

4. Conclusion

The initial impounding of a reservoir is conducted after the completion of dam construction, beginning with the closure of the diversion tunnel gate, after which the water level gradually rises. The duration and rate of increase in reservoir water level are influenced by whether the initial impounding occurs during the dry or rainy season. For safety reasons, the rate of water level rise is regulated so that it does not exceed 1.0 meter per day to protect the dam structure. This study was conducted on 10 reservoirs that underwent single-stage impounding. The research began by analysing observed data on water level elevations during the initial impounding process in several dams. To develop the model, field data were first selected, focusing on continuous elevation records from the riverbed up to the spillway (full supply level). Verification was also performed using data from four additional dams currently undergoing initial impounding. Following the analysis using linear regression and selection based on the highest determination coefficient (R^2), the resulting linear regression equation for predicting reservoir impounding time is as follows: $\mathrm{Ln}\left(\mathrm{D}_{(\text{linear regression})}\right) = 1.047 \times Ln(S) + (-1.08807) \times \mathrm{Ln}(\mathrm{I}_{\text{net}})$. The result of model evaluation method consists of NSE > 0.50 and RSR = 0.117 < 0.70, and if PBIAS = -0.50% < 25% for streamflow, demonstrate that the regression-based model is highly effective in estimating reservoir impounding time with minimal error. For further application, the linear regression model can be integrated into real-time decision-making tools for dam operators when combined with automated monitoring systems and hydrological sensors. The study by Skhakhfa & Ouerdachi [24] demonstrated the use of a real-time decision support system (DSS) for power generation and flood control, which consists of three key components: real-time operation and visualization for dam operators, a forecasting model, and optimized decision-making modules.

5. Declarations

5.1. Author Contributions

Conceptualization, H.R.R. and L.M.L.; methodology, P.T.J.; software, H.R.R.; validation, E.Y., L.M.L., and E.Y.; formal analysis, H.R.R.; investigation, H.R.R.; resources, E.Y.; data curation, H.R.R.; writing—original draft preparation, L.M.L.; writing—review and editing, E.Y.; visualization, P.T.J.; supervision, P.T.J., project administration, H.R.R.; funding acquisition, H.R.R. All authors have read and agreed to the published version of the manuscript.

5.2. Data Availability Statement

The data presented in this study are available in the article.

5.3. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

5.4. Institutional Review Board Statement

Not applicable.

5.5. Informed Consent Statement

Not applicable.

5.6. Declaration of Competing Interest

The authors declare that there are no conflicts of interest concerning the publication of this manuscript. Furthermore, all ethical considerations, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancies have been completely observed by the authors.

6. References

[1] Kaunda, R. B., & Asbury, B. (2016). Prediction of rock brittleness using nondestructive methods for hard rock tunneling. Journal of Rock Mechanics and Geotechnical Engineering, 8(4), 533–540. doi:10.1016/j.jrmge.2016.03.002.

[2] Yin, T., Li, Q., Hu, Y., Yu, S., & Liang, G. (2020). Coupled thermo-hydro-mechanical analysis of valley narrowing deformation of high arch dam: A case study of the Xiluodu project in China. Applied Sciences (Switzerland), 10(2), 524. doi:10.3390/app10020524.

[3] Chen, Y., Hong, J., Tang, S., & Zhou, C. (2016). Characterization of transient groundwater flow through a high arch dam foundation during reservoir impounding. Journal of Rock Mechanics and Geotechnical Engineering, 8(4), 462–471. doi:10.1016/j.jrmge.2016.03.004.

[4] Chuo, M., Ma, J., Liu, D., & Yang, Z. (2019). Effects of the impounding process during the flood season on algal blooms in Xiangxi Bay in the Three Gorges Reservoir, China. Ecological Modelling, 392, 236–249. doi:10.1016/j.ecolmodel.2018.11.017.

[5] Huang, W., Zhou, J., & Zhang, D. (2020). Exploring empirical linkage of water level-climate-vegetation across the Three Georges Dam Areas. Water (Switzerland), 12(4), 965. doi:10.3390/W12040965.

[6] Tschernutter, P., & Kainrath, A. (2016). Design considerations and behavior of reinforced concrete core dams during construction and impounding. Water Science and Engineering, 9(3), 212–218. doi:10.1016/j.wse.2016.11.006.

[7] Wang, X., Zhou, J., Ouyang, S., & Li, C. (2014). Research on Joint Impoundment Dispatching Model for Cascade Reservoir. Water Resources Management, 28(15), 5527–5542. doi:10.1007/s11269-014-0820-y.

[8] Liu, X., Guo, S., Liu, P., Chen, L., & Li, X. (2011). Deriving Optimal Refill Rules for Multi-Purpose Reservoir Operation. Water Resources Management, 25(2), 431–448. doi:10.1007/s11269-010-9707-8.

[9] Zhang, J., Feng, L., Chen, S., Huang, T., Chen, L., Wang, D., Dai, M., & Zhang, D. (2018). Impoundment impact of the three Gorge Reservoir on the hydrological regime in the lower Han River, China. Water (Switzerland), 10(11), 1670. doi:10.3390/w10111670.

[10] Hung, Y. C., Chen, T. T., Tsai, T. F., & Chen, H. X. (2021). A comprehensive investigation on abnormal impoundment of reservoirs—a case study of Qinglin reservoir in Kinmen Island. Water (Switzerland), 13(11), 1463. doi:10.3390/w13111463.

[11] He, S., Guo, S., Chen, K., Deng, L., Liao, Z., Xiong, F., & Yin, J. (2019). Optimal impoundment operation for cascade reservoirs coupling parallel dynamic programming with importance sampling and successive approximation. Advances in Water Resources, 131, 103375. doi:10.1016/j.advwatres.2019.07.005.

[12] Zheng, D., Yuan, B., & Moghanloo, R. G. (2017). Analytical modeling dynamic drainage volume for transient flow towards multi-stage fractured wells in composite shale reservoirs. Journal of Petroleum Science and Engineering, 149, 756–764. doi:10.1016/j.petrol.2016.11.023.

[13] Towfighi, S. (2019). pySRURGS - a python package for symbolic regression by uniform random global search. Journal of Open-Source Software, 4(41), 1675. doi:10.21105/joss.01675.

[14] Abeysingha, N. S., Singh, M., Sehgal, V. K., Khanna, M., Pathak, H., Jayakody, P., & Srinivasan, R. (2015). Assessment of water yield and evapotranspiration over 1985 to 2010 in the Gomti River basin in India using the SWAT model. Current Science, 108(12), 2202–2212.

[15] Amitaba, I. W., Juwono, P. T., Limantara, L. M., & Asmaranto, R. (2024). Real Time Operation Simulation Model with Early Release Reservoir Storage. Journal of Human, Earth, and Future, 5(4), 574-590. doi:10.28991/HEF-2024-05-04-03.

[16] Bhadoriya, U. P. S., Mishra, A., Singh, R., & Chatterjee, C. (2020). Implications of climate change on water storage and filling time of a multipurpose reservoir in India. Journal of Hydrology, 590, 125542. doi:10.1016/j.jhydrol.2020.125542.

[17] Jain, S. K., Shilpa, L. S., Rani, D., & Sudheer, K. P. (2023). State-of-the-art review: Operation of multi-purpose reservoirs during flood season. Journal of Hydrology, 618, 129165. doi:10.1016/j.jhydrol.2023.129165.

[18] Linsley, R.K, and Franzini, J. B. (2020). Water resources management. An International Journal published for the European Water Resources Association (EWRA), Athens, Greece.

[19] Masselot, P., Dabo-Niang, S., Chebana, F., & Ouarda, T. B. M. J. (2016). Streamflow forecasting using functional regression. Journal of Hydrology, 538, 754–766. doi:10.1016/j.jhydrol.2016.04.048.

[20] Ha, T. Van, Ha, T. N., & Nguyen, K. D. (2023). Real-time operating decision support system for reservoir operation using weather forecast and a hydrological model. Journal of Mining and Earth Sciences, 64(3), 41–50. doi:10.46326/jmes.2023.64(3).05.

[21] Nguyen, N.M. (2025). Empirical study of dynamic regret in online model predictive control for linear time-varying systems. Department of Computer Science University of Alberta, arXiv (Preprint), arXiv:2502.13418. doi:10.48550/arXiv.2502.13418.

[22] Schrunner, S., Pishrobat, P., Janssen, J., Jenul, A., Cao, J., Ameli, A. A., & Welch, W. J. (2025). A Gaussian sliding windows regression model for hydrological inference. Journal of the Royal Statistical Society Series C: Applied Statistics, qlaf009. doi:10.1093/jrsssc/qlaf009.

[23] D. N. Moriasi, J. G. Arnold, M. W. Van Liew, R. L. Bingner, R. D. Harmel, & T. L. Veith. (2007). Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. Transactions of the ASABE, 50(3), 885–900. doi:10.13031/2013.23153.

[24] Skhakhfa, I. D., & Ouerdachi, L. (2016). Hydrological modelling of Wadi ressoul watershed, Algeria, by HEC-HMS model. Journal of Water and Land Development, 31(1), 139–147. doi:10.1515/jwld-2016-0045.