

REINFORCEMENT LEARNING-BASED RESERVOIR OPERATIONAL MODEL DUE TO CLIMATE CHANGE AND LAND USE: CASE STUDY IN HERITAGE SOLO VALLEY DEVELOPMENT, INDONESIA

MODELE D'EXPLOITATION DE RESERVOIR BASE SUR L'APPRENTISSAGE PAR RENFORCEMENT FACE AU CHANGEMENT CLIMATIQUE ET A L'USAGE DES SOLS: ETUDE DE CAS DU PROJET DE DÉVELOPPEMENT DE LA VALLEE DE HERITAGE SOLO (INDONESIE)

Y. Indriani^{1, 2}, S.I. Wahyudi¹, T. Andersen², and S. Soedarsono³

ABSTRACT

The Heritage Solo Valley irrigation and flood-control system is a strategic water facility in the Bengawan Solo basin, Indonesia. It supplies irrigation to approximately 62,000 ha of agricultural land through a main canal of about 191 km and 23 branch canals totaling about 117 km. Climate change, upstream land-use change and non-stationary inflow have made fixed reservoir and canal operating rules less reliable for balancing irrigation supply and flood control. This paper develops a reinforcement learning-based reservoir operation framework for the Heritage Solo Valley system and the Karangnongko watershed. The framework integrates GR4J inflow simulation, RCP4.5 and RCP8.5 climate scenarios, Monte Carlo-based stochastic evaluation, satellite-derived land-use analysis for 2017-2025, and comparative testing of STATIC, DYNAMIC, PPO and PPO-RISK operation policies. Land-use analysis shows settlement expansion from 29.025% to 33.554%, shrub/grassland decline from 12.562% to 6.260%, and an increase in runoff coefficient proxy from 0.424 to 0.443. However, land-use change has limited influence on annual mean GR4J inflow compared with rainfall and potential evapotranspiration. The climate-only model already explains annual mean inflow with $R^2 = 0.998$, while climate + LULC increases it slightly to $R^2 = 0.999$. LULC adds more information for maximum inflow, where R^2 increases from 0.105 to 0.254. Under RCP4.5, PPO reduces shortage by 92.3%, spill by 79.7% and the Climate Stress Index by 86.0% compared with STATIC operation. Under RCP8.5, PPO-RISK increases reliability to 0.988 and reduces shortage to 61.72 MCM, although spill remains 6,618.28 MCM and CSI remains 1.045. These findings show that operational risk shifts from deficit toward surplus and spill under extreme climate conditions. Land-use change is therefore better used as a risk-aware state and reward modifier in PPO-RISK than as a direct inflow correction factor.

Keywords: climate change; land-use change; PPO-RISK; reinforcement learning; reservoir operation; Solo Valley. 8/6.7.2026 FINAL.

RESUME

Le système d'irrigation et de la maîtrise des inondations de la vallée de Heritage Solo (Heritage Solo Valley) constitue une infrastructure hydraulique stratégique dans le bassin du Bengawan Solo, en Indonésie. Il fournit l'irrigation à environ 62 000 ha de terres agricoles grâce à un canal principal d'environ 191 km et 23 canaux secondaires totalisant environ 117 km. Le changement climatique, le changement d'usage des sols en amont et le débit non stationnaire de l'influx d'eau ont réduit la fiabilité des règlements d'exploitation fixes (réservoir et canaux) pour équilibrer l'approvisionnement en irrigation et la maîtrise des crues. Cet article présente un cadre d'exploitation de réservoir basé sur l'apprentissage par renforcement, appliqué au système de la vallée de Heritage Solo et au bassin

¹ Universitas Islam Sultan Agung, Semarang, Indonesia, E-mail: wahyudi@unissula.ac.id

² Bojonegoro University, Bojonegoro, Indonesia, E-mail: admin@unigoro.ac.id

³ VIA University College, Horsens, Denmark, E-mail: THRA@via.dk

versant de Karangnongko. Le cadre intègre la simulation de l'influx GR4J, les scénarios climatiques RCP4.5 et RCP8.5, l'évaluation stochastique basée sur Monte Carlo, l'analyse de l'usage des sols par satellite pour 2017-2025, ainsi que les tests comparatifs des politiques d'exploitation STATIC, DYNAMIC, PPO et PPO-RISK. L'analyse de l'usage des sols montre une expansion des zones de peuplement de 29,025% à 33,554%, un déclin des arbustes/prairies de 12,562% à 6,260% et une augmentation de l'indicateur du coefficient ruissellement de 0,424 à 0,443. Cependant, le changement d'usage des sols exerce une influence limitée sur l'influx annuel moyen de GR4J par rapport aux précipitations et à l'évapotranspiration potentielle. Le modèle basé uniquement sur les données climatique explique déjà l'influx annuel moyen avec $R^2 = 0,998$, tandis que l'intégration du climat + de l'usage des sols/de la couverture terrestre (LULC) augmente légèrement cette valeur à $R^2 = 0,999$. Le modèle LULC ajoute plus d'informations concernant l'influx maximal, où R^2 augmente de 0,105 à 0,254. Dans le scénario RCP4.5, le PPO réduit le déficit de 92,3%, les déversements de 79,7% et l'indice de stress climatique de 86,0% par rapport à l'exploitation STATIC. Dans le scénario RCP8.5, le PPO-RISK augmente la fiabilité à 0,988 et réduit le déficit à 61,72 MCM, bien que le déversement reste 6 618,28 MCM et le CSI reste 1,045. Ces résultats montrent que, dans des conditions climatiques extrêmes, le risque opérationnel passe d'une situation de déficit vers une situation d'excédent et de déversement. Par conséquent, il est plus pertinent d'intégrer le changement d'usage des sols dans PPO-RISK comme un état conscient du danger et modificateur de la récompense, plutôt que comme un facteur de correction direct d'influx.

Mots-clés : Changement climatique; Changement d'usage des sols; PPO-RISK; apprentissage par renforcement; Exploitation du réservoir; Vallée du Solo.

INTRODUCTION

The Heritage Solo Valley system is an irrigation and flood-control system in the Bengawan Solo basin. It is important for crop production, drainage and flood mitigation because it links long storage, canals and downstream irrigation areas. The system is expected to serve approximately 62,000 ha of crops and to help reduce flood risk through coordinated canal and reservoir operation. However, most operating rules still rely on expected inflow and fixed seasonal assumptions, which become weaker when hydroclimate and land-use conditions change (Chow and Maidment 1998; Loucks and van Beek 2017; De Girolamo et al. 2022; Hossein-Panahi et al. 2025).

Climate change affects rainfall timing, drought frequency and river flow regimes. The resulting uncertainty creates a serious challenge for water-system management, especially in tropical basins with high rainfall variability and frequent extreme events. At the same time, upstream land-use conversion to settlement, industry and other economic activities can increase impervious surfaces, produce faster runoff, raise flood peaks and reduce dry-season baseflow (Di Prima et al. 2020; Li et al. 2021; Fritzie et al. 2025). Reservoir operators therefore need procedures that can remain reliable under both water deficit and excess water.

Recent reservoir operation studies use stochastic optimization, interval programming, evolutionary algorithms, aggregation-decomposition methods and artificial intelligence for water allocation and flood control (Aljanabi et al. 2018; Bekri et al. 2015; Dariane and Farahmandfar 2013; Farriansyah et al. 2018; Bachtiar et al. 2023; Min et al. 2024; Wahyuni et al. 2024; Zha et al. 2021). Reinforcement learning (RL) is increasingly relevant because it learns a policy from interaction between system state, action, reward and response (Watkins and Dayan 1992; Sutton and Barto 2020; Mnih et al. 2015; Mnih et al. 2016; Schulman et al. 2017). PPO is useful for continuous operational decisions, but standard reward structures may not adequately penalize simultaneous shortage and spill.

This paper focuses on practical operation. The contribution is not a new theoretical RL algorithm. The contribution is an operational framework that combines climate scenarios, land-use signals, GR4J inflow and PPO-RISK for decision support. Land use is not treated as a stand-alone correction to annual inflow. It is translated into risk-aware states and reward modifiers that help the agent identify when to conserve water and when to release or transfer surplus water.

MATERIALS AND METHODS

Study system and data

The study focuses on the Heritage Solo Valley development and the Karangnongko watershed in Indonesia. Solo Valley is treated as a multi-reservoir or long-storage system for irrigation supply and surplus-water control. The Karangnongko watershed is used as the unit for inflow simulation and land-use analysis because it represents an important upstream hydrological signal for the operational framework.

The data used in this study include rainfall, potential evapotranspiration (PET), GR4J inflow for the Karangnongko watershed, annual inflow summaries for 2006-2079, land-use data for 2017-2025, and reservoir-operation results under RCP4.5 and RCP8.5. Operation performance is evaluated using reliability, service ratio, shortage, spill, the Climate Stress Index (CSI), balanced score and transfer volume. All volumes are reported in million cubic metres (MCM).

Hydrological, land-use and counterfactual analysis

The GR4J rainfall-runoff model is used to generate inflow because it is parsimonious and suitable for long time-series simulation (Perrin et al. 2003). Bias-corrected and downscaled climate scenarios are used to represent RCP4.5 and RCP8.5 conditions (Teutschbein and Seibert 2012). Monte Carlo-based stochastic evaluation is used to represent a wide range of inflow conditions and to test operation policies under dry, normal and wet situations (Metropolis and Ulam 1949; Kroese et al. 2014; Rubinstein and Kroese 2017).

Land-use analysis calculates the annual percentage of settlement, forest, agriculture, shrub/grassland, water body, wetland and bare land. A runoff coefficient proxy and a vegetation index are derived from this composition. These variables are not measured runoff coefficients; they are practical indicators of watershed response. The relationship between land use and GR4J inflow is examined over the 2017-2025 overlap period using correlation, climate-only regression, climate + LULC regression and counterfactual testing, following the need to link hydrological data quality, watershed response and operation decisions (Chow and Maidment 1998; Statistical Methods in the Atmospheric Sciences 2019; Ketut Nuraga et al. 2025).

The counterfactual test compares two predictions. In the actual-LULC case, annual land-use indicators follow mapped land-use conditions. In the fixed-LULC case, the 2017 land-use condition is repeated for all years. The difference between the two predictions is interpreted as an approximate land-use signal. Because the overlap period contains only nine years, the results are treated as indicative, not causal.

PPO-RISK formulation

PPO-RISK strengthens PPO by adding risk-sensitive operation features. The best policy is selected using hold-out validation rather than training reward alone. The validation objective combines balanced score, CSI penalty and service-ratio deviation. Uncertainty-aware action adjusts release and intake according to low-flow p05 and high-flow p95 signals. Risk-based transfer redistributes water when surplus and deficit conditions occur in different parts of the system. This design follows earlier work on reinforcement learning, hydropower/reservoir optimization and forecast-informed reservoir operation (Dariane and Moradi 2016; Hooshyar et al. 2020; Castro-Freibott et al. 2025; Turner et al. 2017).

Reliability represents the frequency of satisfactory supply. Service ratio compares total supply with total demand. Shortage is unmet water demand. Spill is surplus water that cannot be stored or used. CSI is a combined stress indicator reflecting shortage, spill and excess operation pressure. A low CSI indicates a more controlled operating condition.

Table 1. Risk-aware variables used in PPO-RISK

Variable	Operational role in PPO-RISK
Settlement percentage	Indicator of impervious-surface pressure
Forest percentage	Indicator of vegetation retention capacity
Agricultural percentage	Indicator of productive land and irrigation demand
Vegetation index	Indicator of general vegetative condition
Runoff coefficient proxy	Indicator of potential surface runoff
Spill risk index	Signal of surplus and spill risk
Shortage risk index	Signal of deficit and shortage risk

RESULTS AND DISCUSSION

Land-use change in the Karangnongko watershed

This subsection evaluates how land-use composition in the Karangnongko watershed changed between 2017 and 2025 and how these changes may influence reservoir-operation risk. The analysis focuses on major land-use classes that are relevant to hydrological response, especially settlement, forest, agriculture and shrub/grassland. The comparison between 2017 and 2025 is used to identify the direction of change, while the annual series supports interpretation of temporal fluctuation.

The results in Table 2 should be read as a watershed-risk screening result rather than as a direct causal estimate of inflow change. Settlement expansion and the increase in the runoff coefficient proxy indicate a tendency toward faster runoff generation. This signal is important for PPO-RISK because it can help the operation policy anticipate peak inflow and spill risk, even when the statistical effect of land-use change on annual mean GR4J inflow remains limited.

Table 2. Land-use change in the Karangnongko watershed, 2017-2025

Variable	2017	2025	Change
Settlement	29.025%	33.554%	+4.529 pp
Forest	18.010%	20.920%	+2.910 pp
Agriculture	39.335%	38.065%	-1.271 pp
Shrub/grassland	12.562%	6.260%	-6.302 pp
Runoff coefficient proxy	0.424	0.443	+0.019

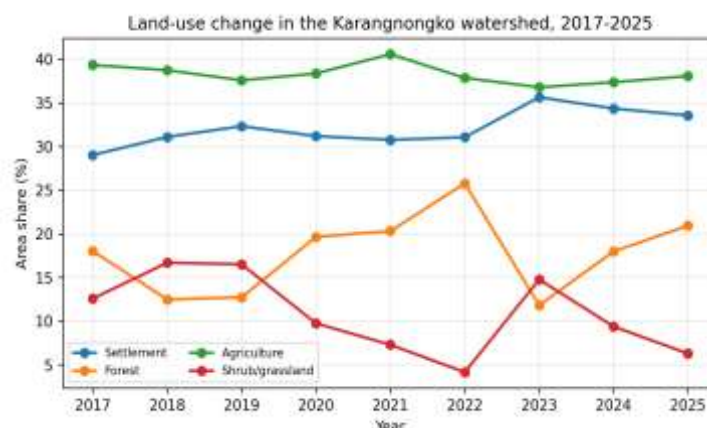


Figure 1. Annual percentage of major land-use classes in the Karangnongko watershed, 2017-2025.

The dominant land-use signals are settlement expansion and shrub/grassland reduction. The runoff coefficient proxy increases from 0.424 to 0.443, indicating a more responsive watershed. This is operationally important because a more responsive watershed can increase fast runoff and spill risk even when the effect on annual mean inflow is limited.

The increase in forest percentage should be interpreted carefully because land-cover classification may fluctuate between forest and shrub classes. For operation, the key message is the change in watershed risk. Future work should use a spatial transition matrix to confirm whether shrub/grassland was converted to settlement, agriculture, forest or other land uses.

Climate-only model versus climate + LULC model

The comparison between climate-only and climate + LULC models shows that annual mean inflow is mainly controlled by rainfall and PET. Adding LULC increases R2 for mean inflow only from 0.998 to 0.999. For maximum inflow, however, R2 increases from 0.105 to 0.254 and RMSE decreases from 3,127.257 to 2,855.638. This indicates that LULC is more useful for identifying peak-flow and spill risk than for explaining annual mean inflow.

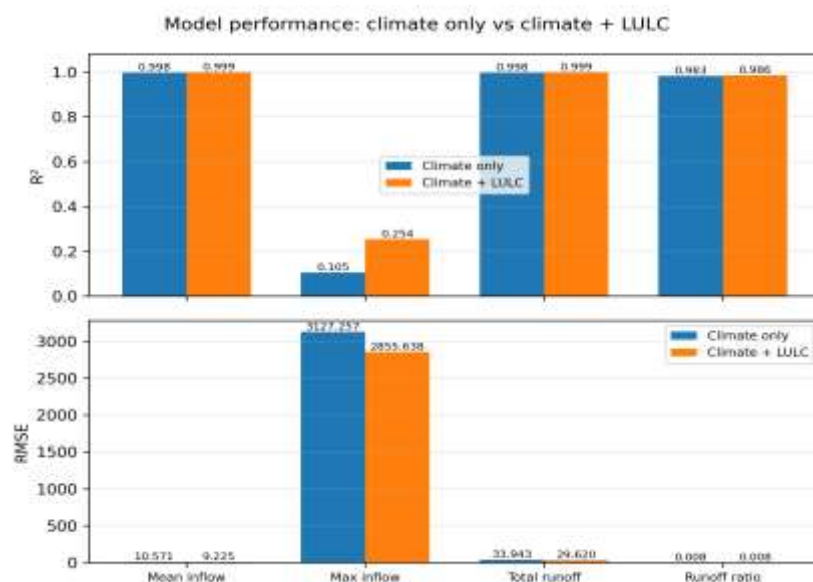


Figure 2. Performance comparison between climate-only and climate + LULC models using R2 and RMSE.

Counterfactual interpretation

The counterfactual analysis shows nearly overlapping curves between actual LULC and fixed-2017 LULC. The percentage effect of actual LULC ranges from +0.48% to -2.58%. This confirms that land-use change is not a dominant driver of annual mean GR4J inflow in the overlap period. The result also prevents over-interpretation of a short statistical relationship.

The practical interpretation is that LULC should be used in the risk layer of PPO-RISK rather than forced into direct annual inflow correction. When rainfall forecasts indicate wet conditions and the watershed has a higher runoff proxy, LULC helps the agent anticipate fast response, peak inflow and possible spill.

Reservoir-operation performance under RCP4.5 and RCP8.5

Under RCP4.5, PPO strongly improves system performance compared with STATIC. Shortage decreases from 4,230.92 MCM to 327.65 MCM, spill decreases from 13,976.08 MCM to 2,830.68 MCM

and CSI decreases from 3.050 to 0.428. This shows that PPO improves supply-demand balance and reduces surplus pressure at the same time.

Under RCP8.5, PPO-RISK produces the highest reliability and the lowest shortage. Reliability increases to 0.988 and shortage decreases to 61.72 MCM. However, spill remains 6,618.28 MCM and CSI remains 1.045. The RCP8.5 result indicates that the main operational problem shifts from deficit to surplus-water control. Improving reliability alone is therefore not sufficient. Wet-season release rules, transfer routes and storage targets need to respond to surplus forecasts.

Table 3. Performance of reservoir-operation models under RCP4.5 and RCP8.5

Scenario	Model	Reliability	Shortage (MCM)	Spill (MCM)	CSI
RCP4.5	STATIC	0.144	4,230.92	13,976.08	3.050
RCP4.5	DYNAMIC	0.144	4,230.92	13,976.08	3.050
RCP4.5	PPO	0.226	327.65	2,830.68	0.428
RCP8.5	STATIC	0.493	374.94	8,509.55	1.286
RCP8.5	DYNAMIC	0.629	308.38	7,368.45	1.121
RCP8.5	PPO-RISK	0.988	61.72	6,618.28	1.045

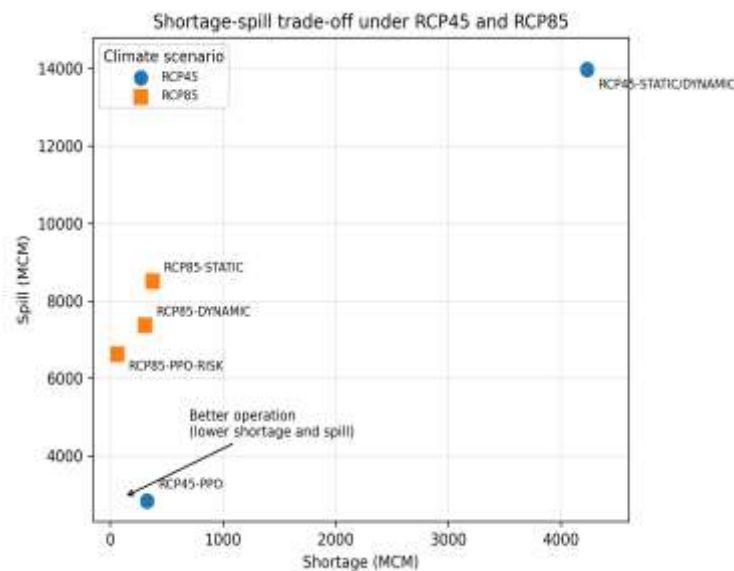


Figure 3. Trade-off between shortage and spill under RCP4.5 and RCP8.5. Points closer to the lower-left corner indicate better operation.

Risk index and operational implications

The risk-index analysis shows that the system does not face a uniform risk over time. Many simulated years are dominated by shortage risk, while several extreme years show increased spill risk. This supports multi-mode operation. During deficit years, the policy should conserve storage and limit non-priority releases. During surplus years, the policy should increase early release, lower target storage or activate risk-based transfer.

Operationally, the Heritage Solo Valley system should be treated as an active climate buffer rather than a passive storage system. A single rule curve is not sufficient. Practical policy should include deficit, normal, surplus, peak-risk and storage-conservation modes. Operators can use PPO-RISK to identify when to conserve water, when to release earlier and when to activate transfer. The final decision should still consider gate capacity, downstream safety, irrigation schedules, maintenance constraints and institutional operating procedures, as emphasized in practical reservoir operation and water-allocation studies (Loucks and van Beek 2017; Jeferson Bale et al. 2023; Ketut Nuraga et al. 2025).

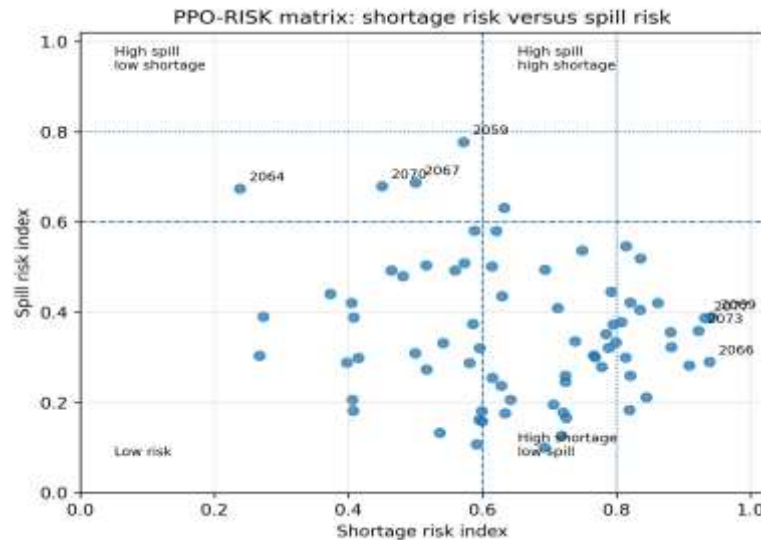


Figure 4. PPO-RISK matrix based on shortage risk index and spill risk index.

LIMITATIONS

The LULC-inflow analysis uses only nine overlapping years from 2017 to 2025, so the statistical relationship remains indicative. A transition matrix between land-use classes is still needed to confirm spatial conversion pathways. The PPO-RISK indices are annual indicators and should be developed further at seasonal or daily resolution for real-time operation. Economic losses, ecological flow, sediment flushing, gate operation reliability and social acceptance of release changes are not explicitly represented. These factors should be added before the framework is used for formal operating rules.

CONCLUSIONS

The results lead to several operational conclusions for the Heritage Solo Valley system. They show that future reservoir and canal operation must respond to both climate-driven inflow uncertainty and watershed-condition signals from land-use change. The main practical implication is that PPO-RISK should be used as a decision-support framework that distinguishes deficit years from surplus years, rather than as a fixed operating rule. The following conclusions summarize the main findings from the hydrological analysis, land-use assessment and reservoir-operation evaluation.

1. The Heritage Solo Valley system requires adaptive operation because climate change and land-use change create both deficit and surplus risks.
2. The Karangnongko watershed experienced clear land-use change in 2017-2025, mainly settlement expansion and shrub/grassland reduction.
3. Land-use change has limited influence on annual mean GR4J inflow compared with rainfall and PET, but it adds information for maximum inflow and spill-risk interpretation.
4. PPO improves RCP4.5 operation by reducing shortage, spill and CSI. PPO-RISK improves RCP8.5 reliability and shortage, although spill remains the dominant challenge.
5. Land-use change should be integrated into PPO-RISK as a risk-aware state and reward modifier, not as a single inflow correction factor.

RECOMMENDATIONS

The following recommendations translate the main findings into practical steps for data improvement, model development and field implementation. They are intended to strengthen PPO-RISK before it is used in reservoir and canal operation. The first recommendations improve the reliability of land-use and hydrological inputs, while the later recommendations guide risk-aware policy design, stress testing and operational use. The aim is to keep PPO-RISK as an explainable decision-support tool that supports, rather than replaces, reservoir and canal operators.

1. Validate land-use transitions using a spatial transition matrix and extend the LULC series with consistent image-classification methods.
2. Include runoff coefficient proxy, vegetation index, spill risk index and shortage risk index in PPO-RISK state variables.
3. Use adaptive rewards with stronger spill penalties during high spill-risk years and stronger shortage penalties during deficit years.
4. Use extreme years such as 2059 for spill risk and 2069 for shortage risk as stress tests before operational implementation.
5. Apply PPO-RISK as a decision-support tool, not as a full replacement for reservoir and canal operators.

REFERENCES

- Ahmadianfar, I., B. Halder, S. Heddam, L. Goliatt, M.L. Tan, Z. Sa'adi, Z. Al-Khafaji, R.Z. Homod, T.A. Rashid and Z.M. Yaseen. 2023. An enhanced multioperator Runge-Kutta algorithm for optimizing complex water engineering problems. *Sustainability* 15(3). <https://doi.org/10.3390/su15031825>.
- Aljanabi, A.A., L.W. Mays and P. Fox. 2018. Optimization model for agricultural reclaimed water allocation using mixed-integer nonlinear programming. *Water* 10(10). <https://doi.org/10.3390/w10101291>.
- Bachtiar, S., L.M. Limantara, M. Sholichin and W. Soetopo. 2023. Optimization of integrated reservoir for supporting the raw water supply. *Civil Engineering Journal* 9(4): 860-872. <https://doi.org/10.28991/CEJ-2023-09-04-07>.
- Bekri, E., M. Disse and P. Yannopoulos. 2015. Optimizing water allocation under uncertain system conditions for water and agriculture future scenarios in Alfeios River basin (Greece). *Water* 7(11): 6427-6466. <https://doi.org/10.3390/w7116427>.
- Castro-Freibott, R., A. Garcia-Sanchez, F. Espiga-Fernandez and G. Gonzalez-Santander de la Cruz. 2025. Deep reinforcement learning for intraday multireservoir hydropower management. *Mathematics* 13(1). <https://doi.org/10.3390/math13010151>.
- Chow, V.T. and D.R. Maidment. 1998. *Applied hydrology*. McGraw-Hill, New York.
- Darlane, A.B. and Z. Farahmandfar. 2013. A comparative study of marriage in honey bees optimisation algorithm in multi-reservoir system optimisation. *Water SA* 39(2): 327-334. <https://doi.org/10.4314/wsa.v39i2.17>.
- Darlane, A.B. and A.M. Moradi. 2016. Comparative analysis of evolving artificial neural network and reinforcement learning in stochastic optimization of multireservoir systems. *Hydrological Sciences Journal* 61(6): 1141-1156. <https://doi.org/10.1080/02626667.2014.986485>.
- De Girolamo, A.M., E. Barca, M. Leone and A. Lo Porto. 2022. Impact of long-term climate change on flow regime in a Mediterranean basin. *Journal of Hydrology: Regional Studies* 41. <https://doi.org/10.1016/j.ejrh.2022.101061>.
- Di Prima, S., M. Castellini, J. Rodrigo-Comino and A. Cerda. 2020. Soil hydrology for sustainable land management: theory and practice. *Water* 12(4). <https://doi.org/10.3390/W12041109>.
- Farriarsyah, A.M., P.T. Juwono, E. Suhartanto and V. Dermawan. 2018. Water allocation computation model for river and multi-reservoir system with sustainability-efficiency-equity criteria. *Water* 10(11). <https://doi.org/10.3390/w10111537>.
- Fritzie, W., P. Qatrinnada, E. Hidayah and G. Halik. 2025. Determination of rainfall thresholds as flood triggers in the Klopogowok watershed, Jember. *Jurnal Teknik Pengairan* 16(2). <https://doi.org/10.21776/ub.pengairan.2025.016.02.1>.

- Hartel, P. and M. Korpas. 2017. Aggregation methods for modelling hydropower and implications for a highly decarbonised energy system in Europe. *Energies* 10(11). <https://doi.org/10.3390/en10111841>.
- Hooshyar, M., S.J. Mousavi, M. Mahootchi and K. Ponnambalam. 2020. Aggregation-decomposition-based multi-agent reinforcement learning for multi-reservoir operations optimization. *Water* 12(10): 1-20. <https://doi.org/10.3390/w12102688>.
- Hosseini-Panahi, B., S.M. Samani, A.R. Sadeghi, M. Shahi, S.M. Hosseini and E. Parizi. 2025. River baseflow in supplying reservoirs inflows of Tehran metropolis: a machine learning modeling based on influencing factors. *Journal of Hydrology: Regional Studies* 60. <https://doi.org/10.1016/j.ejrh.2025.102528>.
- Jain, S., D. Ramesh, M.C. Trivedi and D.R. Edla. 2023. Evaluation of metaheuristic optimization algorithms for optimal allocation of surface water and groundwater resources for crop production. *Agricultural Water Management* 279. <https://doi.org/10.1016/j.agwat.2023.108181>.
- Jeferson Bale, F., M. Bisri and S. Wahyuni. 2023. Optimization utilization study storage of Raknamo Dam, Amabi Oefeto District, Kupang Region. *Jurnal Teknik Pengairan* 14(02). <https://doi.org/10.21776/ub.pengairan.2023.014.02.4>.
- Ketut Nuraga, I., P. Indah and D. Putri. 2025. Review of Gelar Reservoir operation simulation in Jembrana Regency. *Jurnal Teknik Pengairan* 16(2). <https://doi.org/10.21776/ub.pengairan.2025.016.02.3>.
- Kroese, D.P., T. Brereton, T. Taimre and Z.I. Botev. 2014. Why the Monte Carlo method is so important today. *WIREs Computational Statistics* 6(6): 386-392. <https://doi.org/10.1002/wics.1314>.
- Li, Z., B. Huang, Z. Yang, J. Qiu, B. Zhao and Y. Cai. 2021. Mitigating drought conditions under climate and land-use changes by applying hedging rules for the multi-reservoir system. *Water* 13(21). <https://doi.org/10.3390/w13213095>.
- Loucks, D.P. and E. van Beek. 2017. *Water resource systems planning and management: an introduction to methods, models, and applications*. Springer, Cham.
- Lv, Y., G.H. Huang, Y.P. Li, Z.F. Yang, Y. Liu and G.H. Cheng. 2010. Planning regional water resources system using an interval fuzzy bi-level programming method. *Journal of Environmental Informatics* 16(2): 43-56. <https://doi.org/10.3808/jei.201000177>.
- Metropolis, N. and S. Ulam. 1949. The Monte Carlo method. *Journal of the American Statistical Association* 44(247): 335-341. <https://doi.org/10.1080/01621459.1949.10483310>.
- Min, I., N. Lee, S. Kim, Y. Bang, J. Jang, K. Jung and D. Park. 2024. An improved aggregation-decomposition optimization approach for ecological flow supply in parallel reservoir systems. *Sustainability* 16(17). <https://doi.org/10.3390/su16177475>.
- Mnih, V., K. Kavukcuoglu, D. Silver, A.A. Rusu, J. Veness, M.G. Bellemare, A. Graves, M. Riedmiller, A.K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg and D. Hassabis. 2015. Human-level control through deep reinforcement learning. *Nature* 518(7540): 529-533. <https://doi.org/10.1038/nature14236>.
- Mnih, V., A.P. Badia, M. Mirza, A. Graves, T. Lillicrap, T. Harley, D. Silver and K. Kavukcuoglu. 2016. Asynchronous methods for deep reinforcement learning. *arXiv:1602.01783*.
- Perrin, C., C. Michel and V. Andreassian. 2003. Improvement of a parsimonious model for streamflow simulation. *Journal of Hydrology* 279(1-4): 275-289.
- Rubinstein, R.Y. and D.P. Kroese. 2017. *Simulation and the Monte Carlo method*. Wiley, Hoboken.
- Schulman, J., F. Wolski, P. Dhariwal, A. Radford and O. Klimov. 2017. Proximal policy optimization algorithms. *arXiv:1707.06347*.
- Statistical Methods in the Atmospheric Sciences. 2019. Elsevier. <https://doi.org/10.1016/C2017-0-03921-6>.
- Sutton, R.S. and A.G. Barto. 2020. *Reinforcement learning: an introduction*. MIT Press, Cambridge.
- Teutschbein, C. and J. Seibert. 2012. Bias correction of regional climate model simulations for hydrological climate-change impact studies. *Journal of Hydrology* 456-457: 12-29. <https://doi.org/10.1016/j.jhydrol.2012.05.052>.
- Turner, S.W.D., J.C. Bennett, D.E. Robertson and A. Galelli. 2017. Complex relationship between seasonal streamflow forecast skill and value in reservoir operations. *Hydrology and Earth System Sciences* 21(9): 4841-4859.
- Wahyuni, S., D. Sisinggih, I. Elhuda, K. Souma and I.Z. Dasyilva. 2024. Optimization of reservoir release operation using genetic algorithm method. *Results in Engineering* 23. <https://doi.org/10.1016/j.rineng.2024.102610>.
- Watkins, C.J.C.H. and P. Dayan. 1992. Q-learning. *Machine Learning* 8: 279-292.
- Zha, G., J. Zhou, X. Yang, W. Fang, L. Dai, Q. Wang and X. Ding. 2021. Modeling and solving of joint flood control operation of large-scale reservoirs. *Water* 13(1). <https://doi.org/10.3390/w13010041>.