



Development of an Inverse Reinforcement Learning Method Integrated with the ICSS Hydrodynamic Model in the Dez canal

Amirhossein Tavana¹ | Kazem Shahverdi² | Hesam Ghodousi³

1. Department of Water Engineering, Faculty of Agriculture, University of Zanjan, Zanjan, Iran. E-mail:

amirhossein.tavana@znu.ac.ir

2. Corresponding Author, Department of Water Science Engineering, Faculty of Agriculture, Bu-Ali Sina University, Hamedan, Iran. E-mail: k.shahverdi@basu.ac.ir

3. Department of Water Engineering, Faculty of Agriculture, University of Zanjan, Zanjan, Iran. E-mail:

Ghodousi@znu.ac.ir

Article Info	ABSTRACT
Article type: Research Article	Optimal water resource management in irrigation networks has been recognized as a key strategy to address water scarcity challenges. This research developed a novel hybrid method based on inverse reinforcement learning (IRL) and hydrodynamic modeling with ICSS to optimize the operation of the E1-R1 canal in the Dez network. In this method, the reward function was automatically extracted from the experiences of expert operators (simulated using ICSS), and then, by combining it with the hydrodynamic simulation model, optimal control policies for water control structures and intakes were determined. Operational data under various scenarios, including random inflow rates (1.1 to 3.1 cubic meters per second) and different water withdrawal patterns, were analyzed. The results indicated high efficiency of the proposed system, with an average water delivery efficiency of 0.97, a supply adequacy of 0.95, and a maximum depth control error of 14.3%. The cumulative depth error also remained below 10% in all scenarios, indicating long-term system stability. These findings confirm that the IRL approach, by learning the implicit knowledge of operators and converting it into policies, can effectively reduce water losses and improve the performance of irrigation networks.
Article history:	
Received: March. 28, 2025	
Revised: May. 3, 2025	
Accepted: May. 6, 2025	
Published online: July. 2025	
Keywords: Control Structures, Intelligent Method, Hydrodynamic Model, Learning.	
Cite this article: Tavana, A. H., Shahverdi, K., & Ghodousi, H. (2025). Development of an Inverse Reinforcement Learning Method Integrated with a Hydrodynamic Model, <i>Iranian Journal of Soil and Water Research</i> , 56 (5), 1339-1353. https://doi.org/10.22059/ijswr.2025.392613.669911	
© The Author(s).	Publisher: The University of Tehran Press.
DOI: https://doi.org/10.22059/ijswr.2025.392613.669911	

EXTENDED ABSTRACT

Introduction:

Many people worldwide experience water scarcity in their daily lives. Therefore, water resource management is a serious challenge for communities. A significant portion of water resource consumption in a basin is used in agriculture and irrigation and drainage networks, which act like the human body's veins, serving as the main arteries for water distribution and conveyance in an agricultural area. Improper operation of these networks leads to significant water losses, making their efficient management crucial for substantially reducing these losses.

Purpose:

This research develops an inverse reinforcement learning algorithm integrated with a hydrodynamic canal simulation model to address this issue.

Research method:

To achieve the objective, a reward function is learned based on the different actions of the learning agent (water control structures such as weirs and intakes) in various states (obtained based on an expert's experience) in the inverse reinforcement learning algorithm. Then, using this reward function, the actions corresponding to different states are extracted. Using operational data and considering various scenarios, operational patterns in the E1-R1 canal of the Dez network were extracted.

Results:

The findings of this research showed that in most scenarios, the calculated values for the delivery efficiency index are close to one. The key parameters affecting the calculation of this index include the number of off-takes, the required (or requested) discharge, and the actual delivered discharge. According to the standard range of Molden and Gates performance evaluation indices for delivery efficiency, if the values fall within the range of 0.85 to 1, the water is delivered to consumers with higher efficiency, effectively minimizing water losses or excess delivery. In this study, based on the obtained results, the minimum and maximum delivery efficiencies were 0.97 and 1, respectively, which fall within the standard range and indicate a favorable performance. Overall, the efficiency values for Off-takes No. 1, 2, 5, and 6 were in their ideal state across all scenarios. Additionally, the average delivery efficiency for Off-takes No. 1 to 6 was 1, 1, 0.98, 0.98, 1, and 1, respectively.

The cumulative absolute error values were all below 10%. The minimum and maximum cumulative absolute errors were 0.5% and 2.4%, respectively, with an average error of 1.65% across all scenarios - all indicating good performance. The cumulative absolute error values also remained below 10%. The minimum and maximum errors were 1% and 9.4%, respectively, with an average error of 4.34% across all scenarios. Overall, these results similarly fall within the acceptable performance range.

Conclusion:

Based on the various defined scenarios and the results of simulations and learning, all evaluation indices fell within the "good" performance class. This confirms the successful performance of the inverse reinforcement learning algorithm and validates the learned reward function.

The evaluation results for water depth under the inverse reinforcement learning method demonstrated that the model consistently maintained water level variations within the permissible depth range across all scenarios. In all simulated scenarios, the depth remained within the allowable limits, and the average depth fluctuations were centered around the target depth.

Author Contributions

All authors contributed equally to the conceptualization of the article and writing of the original and subsequent drafts.

Data Availability Statement

Data available on request from the corresponding author.

Conflict of interest

The author declares no conflict of interest.

توسعه روش یادگیری تقویتی معکوس به صورت تلفیق با مدل هیدرودینامیک ICSS در کانال دز

امیرحسین توانا^۱ | کاظم شاهوردی^۲ | حسام قدوسی^۳^۱. گروه مهندسی آب، دانشکده کشاورزی، دانشگاه زنجان، زنجان، ایران. رایانامه: amirhossein.tavana@znu.ac.ir^۲. نویسنده مسئول، گروه علوم و مهندسی آب، دانشکده کشاورزی، دانشگاه بوعلی سینا، همدان، ایران. رایانامه: k.shahverdi@basu.ac.ir^۳. گروه مهندسی آب، دانشکده کشاورزی، دانشگاه زنجان، زنجان، ایران. رایانامه: Ghodousi@znu.ac.ir

اطلاعات مقاله	چکیده
نوع مقاله: مقاله پژوهشی	
تاریخ دریافت: ۱۴۰۴/۱/۸	مدیریت بهینه منابع آب در شبکه‌های آبیاری به عنوان راهکاری کلیدی برای مقابله با چالش کم‌آبی است. در این پژوهش، یک روش ترکیبی نوین مبتنی بر یادگیری تقویتی معکوس (IRL) و مدلسازی هیدرودینامیک با ICSS برای بهینه‌سازی بهره‌برداری از کانال E1-R1 شبکه دز توسعه داده شد. در این روش، تابع پاداش به صورت خودکار از تجربیات بهره‌برداران خبره (شبیه‌سازی شده با ICSS) استخراج گردید. با تلفیق آن با مدل شبیه‌ساز هیدرودینامیک، سیاست‌های بهینه کنترل سازه‌های آب‌بند و آبگیر تعیین شد. داده‌های بهره‌برداری تحت سناریوهای متنوعی شامل دبی‌های ورودی تصادفی (۱/۱ تا ۱/۳ مترمکعب بر ثانیه) و الگوهای مختلف آبیگری تحلیل شد. نتایج حاکی از کارایی بالایی سیستم پیشنهادی بود، به طوری که میانگین راندمان تحویل آب ۰/۹۷، کفایت تأمین نیاز ۰/۹۵ و حداکثر خطای کنترل عمق ۱۴/۳ درصد شد. خطای تجمعی عمق آب نیز در تمامی سناریوها کمتر از ۱۰ درصد باقی ماند که نشان‌دهنده پایداری سیستم در بلندمدت است. این یافته‌ها مؤید آن است که رویکرد IRL با یادگیری دانش ضمنی بهره‌برداران و تبدیل آن به سیاست‌ها، می‌تواند به صورت مؤثری تلفات آب را کاهش داده و عملکرد شبکه‌های آبیاری را ارتقا بخشد.
واژه‌های کلیدی: سازه‌های کنترل و تنظیم، روش هوشمند، مدل هیدرودینامیک، یادگیری.	
استناد: توانا، امیرحسین، شاهوردی، کاظم، قدوسی، حسام (۱۴۰۴). توسعه روش یادگیری تقویتی معکوس به صورت تلفیق با مدل هیدرودینامیک، مجله تحقیقات آب و خاک ایران، ۵۶ (۵)، ۱۳۵۳-۱۳۳۹. https://doi.org/10.22059/ijswr.2025.392613.669911	
ناشر: مؤسسه انتشارات دانشگاه تهران.	© نویسندگان.
DOI: https://doi.org/10.22059/ijswr.2025.392613.669911	



مقدمه

امروزه کمبود آب به عنوان یک چالش اساسی در سطح جهان و به ویژه در ایران توسط مردم و کارشناسان احساس می شود. این بحران نه تنها زندگی روزمره افراد را تحت تأثیر قرار داده است، بلکه ضرورت مدیریت و حفاظت از منابع آب را به اولویاتی حیاتی تبدیل کرده است. در پژوهشی، راندمان کل آبیاری در ایران به طور متوسط حدود ۵۷٪ برآورد شد. براساس آمار ذکر شده در این پژوهش، بخش قابل توجهی از منابع آبی کشور به دلیل ناکارآمدی در توزیع و مصرف از دست می رود ارتقاء بهره‌وری و استفاده از فناوری‌های نوین ضرورت است (عباسی و عباسی، ۱۴۰۲).

در پژوهشی دیگر نیز به معرفی دو مدل هیدرودینامیک مدل‌سازی و مدیریت کانال Rootcanal و ICSS^۱ پرداخته شد. همچنین، کارایی این نرم‌افزارها با استفاده از اطلاعات واقعی در کانال درجه دوم شبکه آبیاری و زهکشی قوریچای مورد بررسی و مقایسه قرار گرفت. نتایج نشان داد که مدل‌های هیدرودینامیک به خوبی می‌توانند به شبیه‌سازی جریان در شبکه‌های آبیاری بپردازند و در مواردی که تفاوت‌هایی با مقادیر واقعی مورد اندازه‌گیری وجود دارد، مدل Rootcanal نتایج بهتری را ارائه می‌دهد (قدوسی و داداشی، ۱۳۹۵). در تحقیقی، عملکرد دو روش یادگیری سارسای فازی^۲ و روش سنتی، در بهره‌برداری از کانال عقلی شرقی مورد مقایسه قرار گرفت. نتایج این تحقیق نشان داد که روش سارسای فازی نه تنها دقت بیشتری دارد، بلکه در شرایط کم‌آبی توانایی توزیع عادلانه آب بین مصرف‌کنندگان را نیز داراست (Savari et al., 2016). در مطالعه ای با استفاده از نرم افزار مدل هیدرودینامیکی ICSS بهره‌برداری کانال اصلی شبکه آبیاری رودشت شبیه‌سازی شد. سناریو کاهش مقدار آب مورد نیاز به عنوان یکی از سناریوهای سازه‌ای برای بهبود بهره‌برداری کانال مورد مطالعه، آزموده شد. نتایج ارزیابی عملکرد بهره‌برداری از شبکه با کاربرد شاخص‌های چهارگانه مولدن و گیتس^۳ نشان دهنده عملکرد قابل قبول جهت بهبود این شاخص‌ها بوده است (Shahdany et al., 2016).

در پژوهشی به بررسی و مقایسه استفاده از روش‌های هوشمند، از جمله یادگیری تقویتی (FSL)، برای بهبود بهره‌برداری از کانال‌های آبیاری پرداخته شد. نتایج تحقیق نشان داد که استفاده از این روش‌ها منجر به بهبود راندمان، کفایت، پایداری و عدالت در توزیع آب شده است. همچنین، این روش‌ها توانسته‌اند نوسانات سطح آب را کاهش داده و به بهینه‌سازی مصرف آب در سیستم‌های آبیاری کمک کنند (بیات و همکاران، ۱۴۰۱). در پژوهش دیگری نیز به استخراج الگوهای بهره‌برداری با در نظر گرفتن خطای میراب در تنظیم سازه‌ها در شبکه‌های آبیاری با استفاده از کاربرد الگوریتم یادگیری تقویتی پرداخته شد. در فرآیند تنظیم، یکی از خطاهای رایج، خطای ناشی از عملکرد میراب‌ها است. در پژوهشی، بررسی توانایی الگوریتم یادگیری تقویتی در استخراج الگوهای بهینه بهره‌برداری با در نظر گرفتن خطای احتمالی میراب‌ها (حداکثر ۱۰ درصد) انجام شده است که در آن، برای شبیه‌سازی بخشی از کانال دز، از یک مدل غیرخطی استفاده شد. یافته‌های تحقیق حاکی از دقت بالای الگوهای استخراج شده است، به طوری که ارزش کنش‌های انتخاب شده در بیشتر موارد به ۰/۹ می‌رسد. همچنین، الگوریتم موردنظر پس از ۶۵۰ تکرار به جواب نهایی همگرا شده است (شاهوردی، ۱۴۰۱).

در تحقیق دیگری نیز روش‌های بهینه‌سازی جامعه مورچگان و یادگیری تقویتی در بهره‌برداری از کانال عقلی شرقی با سناریوهای مختلف مورد بررسی و مقایسه قرار گرفت. نتایج این تحقیق نشان داد که در حالت شبیه‌سازی با روش یادگیری تقویتی شاخص‌های ارزیابی عملکرد بهتر و در نتیجه راندمان بهره‌برداری بیشتر می‌باشد (Omidzade et al., 2020). در پژوهشی به بهینه‌سازی آبیاری با استفاده از مدل یادگیری تقویتی عمیق پرداخته شد. در این پژوهش یک روش کنترل هوشمند برای آبیاری کشاورزی بر اساس یادگیری تقویتی پیشنهاد شد. با ساخت یک شبکه یادگیری عمیق برای استخراج ویژگی‌ها از داده‌های حسگر خام، مطالعه انجام شد. با استفاده از قابلیت‌های یادگیری و داده‌های یادگیری تقویتی عمیق، می‌توان دقت کنترل آبیاری کشاورزی را به طور موثر بهبود بخشید. نتایج پژوهش نشان داد مدل آموزش دیده بهره‌وری را تا ۱۱ درصد افزایش داده و مصرف آب، ۲۰ تا ۳۰ درصد کاهش می‌یابد (Alibabaei et al., 2022). بررسی کنترل خودکار کانال‌های آبیاری با استفاده از روش یادگیری تقویتی فازی بهبود یافته با هدف بهبود مدیریت و بهره‌برداری از کانال‌های آبیاری، انجام شد و الگوریتمی توسعه داده شد که قادر است نوسانات سطح آب را کاهش دهد و پایداری و کارایی سیستم‌های توزیع آب را افزایش دهد. نتایج این پژوهش نشان داده است که استفاده از این روش می‌تواند عملکرد کانال‌های آبیاری را بهبود بخشیده و توزیع آب را در شرایط مختلف عملیاتی بهینه‌سازی کند (Shahverdi and Monem, 2022).

1 Irrigation Conveyance System Simulation

2 Fuzzy SARSA Learning

3 Molden and Gates

در تحقیقی، بهینه‌سازی چندهدفه توزیع آب در سیستم کانال با در نظر گرفتن پوشش آبیاری مورد بررسی قرار گرفته است. این تحقیق از الگوریتم NSGA-III با استراتژی تصحیح مؤثر و روش وزن‌دهی آنتروپی به منظور بهینه‌سازی استفاده کرد و نتایج آن نشان داد که این روش زمان انتقال آب را ۳ ساعت و تلفات نشت را $\frac{1}{3}$ درصد کاهش می‌دهد (Zhou et al., 2024). در مطالعه دیگری به بهینه‌سازی راندمان آبیاری با استفاده از یادگیری تقویتی عمیق پرداخته شد. از شبکه‌های عصبی و الگوریتم‌های یادگیری تقویتی برای طراحی یک سیستم آبیاری هوشمند استفاده شد که قادر می‌باشد با توجه به شرایط مختلف محیطی و منابع آبی، تصمیمات بهینه‌تری در خصوص زمان‌بندی و مقدار آبیاری اتخاذ کند. نتایج تحقیق نشان می‌دهد که این سیستم توانسته است مصرف آب را تا $\frac{9}{52}$ درصد کاهش دهد (Ding and Du, 2024). در تحقیقی، تأثیر پارامترهای طراحی شبکه توزیع آبیاری بر مدیریت پایداری مورد بررسی قرار گرفت. هدف از این تحقیق تحلیل تأثیر این پارامترها بر بهره‌وری منابع آب و پایداری سیستم‌های آبیاری می‌باشد. در این مطالعه از مدل‌های شبیه‌سازی و تحلیلی برای ارزیابی تأثیرات طراحی شبکه استفاده شده است. نتایج نشان داد که بهینه‌سازی طراحی شبکه می‌تواند، علاوه بر بهبود راندمان، به کاهش اثرات منفی زیست‌محیطی کمک کند (Garcia-Espinal et al., 2024).

تأثیر تغییرات اقلیمی بر عملکرد شبکه آبیاری عقلی مورد بررسی قرار گرفت. این مطالعه با هدف ارزیابی تأثیر تغییرات دمایی و بارشی بر کارایی تحویل آب در این کانال انجام گرفته است. برای تحلیل روند داده‌های اقلیمی از آزمون من-کندال^۱ استفاده شد. نتایج نشان داد که دمای حداکثر، تبخیر- تعرق و جریان ورودی به ترتیب $\frac{3}{7}$ درجه سانتی‌گراد، $\frac{1}{45}$ میلی‌متر در روز و $\frac{39}{1}$ لیتر بر ثانیه افزایش یافته‌اند (Mohebbi et al., 2024). در تحقیقی، تأثیر عدم قطعیت در عملکرد کانال‌های آبیاری در شبکه آبیاری شرقی عقلی مورد بررسی قرار گرفت. هدف این مطالعه ارزیابی تغییرات در باز شدن دریچه‌های انشعاب و تأثیر آن بر کارایی سیستم آبیاری بود. نتایج این تحقیق نشان داد که با افزایش باز شدن دریچه‌ها، نوسانات در جریان آب کاهش، کفایت تأمین آب و کاهش قابلیت اطمینان سیستم افزایش می‌یابد (Aghayee et al., 2024). توزیع آب و کارایی آن در شبکه‌های آبیاری با استفاده از مدل‌سازی ترکیبی هیدرولیکی- هیدرولوژیکی مورد بررسی قرار گرفت. هدف از این مطالعه ارزیابی توزیع آب در بخش‌های مختلف شبکه آبیاری بود. نتایج نشان داد که کانال‌های بالادست با کارایی آبیاری بیش از ۶۰ درصد و کانال‌های پایین‌دست با کارایی کمتر از ۵۰ درصد مواجه هستند (Shin et al., 2024).

در تحقیقات پیشین، بهره‌برداری از شبکه‌های آبیاری و بهینه‌سازی توزیع و تحویل آب با استفاده از روش‌های مختلفی مورد بررسی قرار گرفته است که در برخی از این مطالعات از روش یادگیری تقویتی نیز برای بهبود عملکرد سیستم‌های آبیاری استفاده شده است. با این حال، علیرغم پیشرفت‌های حاصل شده در این زمینه، یک خلأ تحقیقاتی قابل توجه در به کارگیری روش یادگیری تقویتی معکوس (IRL)^۲ برای استخراج و مدل‌سازی تجربیات بهره‌برداران و استفاده از این دانش در بهبود عملکرد شبکه‌های آبیاری وجود دارد. نوآوری اصلی این پژوهش در به کارگیری روش یادگیری تقویتی معکوس برای تحلیل و مدل‌سازی رفتار بهره‌برداران در شبکه‌های آبیاری است. در این روش، با استخراج دانش ضمنی و تجربیات بهره‌برداران و تبدیل آن به یک تابع پاداش بهینه، امکان تصمیم‌گیری هوشمند در توزیع و تحویل آب فراهم می‌شود. در این پژوهش، از روش یادگیری تقویتی معکوس برای اولین بار در حوزه مدیریت شبکه‌های آبیاری استفاده شد.

مواد و روش‌ها

در شکل ۱، فلوچارت انجام این تحقیق نشان داده شده است. ابتدا منطقه مورد مطالعه بررسی و انتخاب شد و سپس داده‌های هیدرولیکی و فیزیکی آن جمع‌آوری شد. با شبیه‌سازی حالت‌های مختلف مقادیر دبی، تغییرات عمق و میزان بازشدگی مشخص شد. این داده‌ها به عنوان ورودی به الگوریتم یادگیری تقویتی معکوس داده شد و تابع پاداش مربوطه استخراج گردید. نتایج حاصل با استفاده از شاخص‌های ارزیابی بررسی شد. در ادامه جزئیات مربوط به هر مرحله تشریح می‌شود.

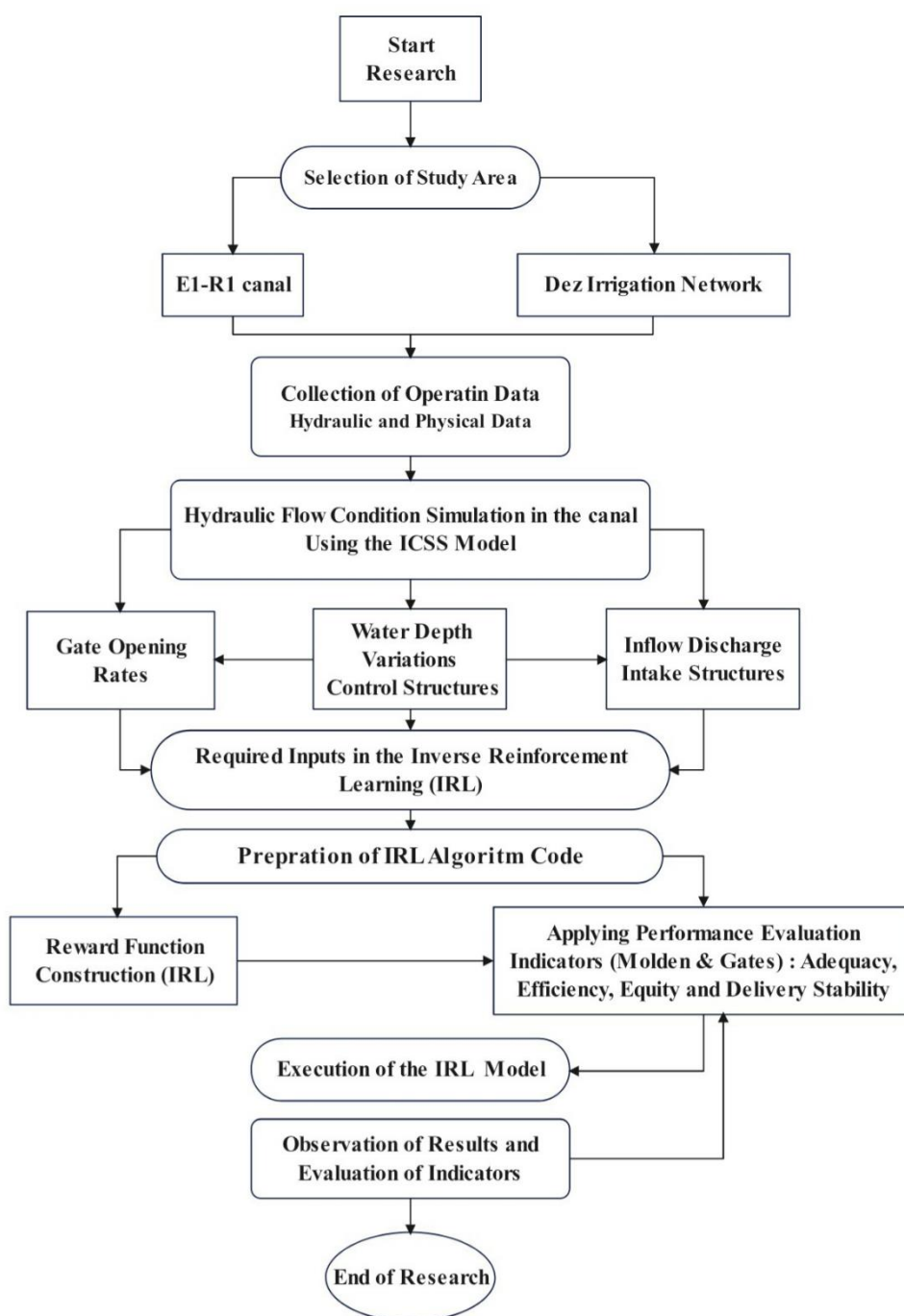
منطقه مورد مطالعه

شبکه آبیاری و زهکشی دز، یکی از بزرگترین شبکه‌های آبیاری ثقلی کشور می‌باشد که در استان خوزستان در ایران واقع شده است. اراضی تحت پوشش شبکه آبیاری دز از سمت شمال با تپه ماهورهای شمالی دزفول، از سمت جنوب به اراضی هفت تپه و رودخانه شاوور، از شرق با رودخانه شور (گللال کهنک) و از غرب با رودخانه کرخه هم‌مرز است. در این منطقه، رودخانه‌های شاوور و دز، راه آهن سراسری، جاده تهران - خرمشهر، دو مسیل بالارود و سیاه منصور از شبکه عبور می‌کنند. علاوه بر این شهرهای دزفول، اندیمشک و شوش نیز در محدوده

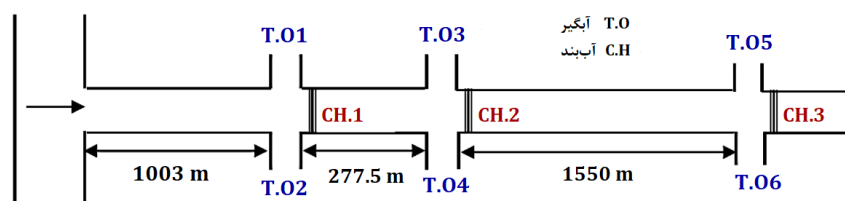
1 Mann-Kendall

2 Inverse Reinforcement Learning

این اراضی قرار دارند. سطح زیر کشت خالص این شبکه برابر ۹۳ هزار هکتار می‌باشد (توانا، ۱۴۰۳). در این تحقیق، بخشی از کانال E1-R1 از شبکه آبیاری دز مورد مطالعه قرار گرفته است. در این کانال از تعداد ۱۵ سازه تقاطعی از نوع سیفون، چک سیفون و کالورت استفاده شده است. وسعت اراضی تحت پوشش آبیاری این کانال برابر ۱۰۲۱۸ هزار هکتار می‌باشد که توزیع آب بین واحدهای کشت و صنعت و بهره‌برداران محلی توسط ۲۱ کانال فرعی (انشعابات) و ۴۹ سازه آبرگیر مزرعه مستقر در طول کانال انجام می‌شود (توانا، ۱۴۰۳). حداکثر ظرفیت این کانال ۲/۴۷ مترمکعب بر ثانیه می‌باشد که آب مورد نیاز را توسط شش آبرگیر ثقیلی با دریچه‌های کشویی مستطیلی در مسیری به طول ۲۸۳۰/۵ متر تأمین می‌کند. در طول مسیر کانال، شیب از ۰/۰۰۱۲ تا ۰/۰۰۱۵ متغیر است. از ابتدای کانال تا فاصله ۱۰۰۳ متری که آب‌بند یک قرار دارد، عرض کف برابر ۱/۵ متر و از نقطه مذکور در فاصله ۲۸۳۰/۵ متری تا محل آب‌بند ۳، معادل ۱ متر می‌باشد. شیب جانبی کانال در طول مسیر به نسبت ۱/۵ افقی به ۱ عمودی است. ضریب زبری مانینگ متوسط در طول مسیر نیز ۰/۰۱۷ می‌باشد. طرح شماتیک کانال E1-R1 و موقعیت سازه‌های آبرگیر و آب‌بند در شکل ۲ نشان داده شده است.



شکل ۱- فلوچارت روند کلی تحقیق حاضر.



شکل ۲- طرح شماتیک کانال E1-R1 در شبکه آبیاری دز (شاهوردی و همکاران، ۱۴۰۲)

یادگیری تقویتی معکوس

یادگیری تقویتی معکوس (IRL)، شاخه‌ای از یادگیری تقویتی است که هدف آن استخراج تابع پاداش^۱ از رفتار یک عامل یا کارشناس انسانی است. در این زمینه، یادگیری تقویتی به دنبال یافتن سیاستی است که بالاترین میزان پاداش را برای یک عامل به ارمغان می‌آورد، در حالی که در IRL، توجه اصلی بر روی شناسایی تابع پاداش است که رفتار مشاهده‌شده از عامل را توضیح دهد. این تابع پاداش نه تنها به شبیه‌سازی رفتار تصمیم‌گیرنده کمک می‌کند بلکه می‌تواند به عنوان معیاری برای نمایان‌سازی اهداف و اولویت‌های عامل نیز عمل کند. استفاده از یادگیری تقویتی معکوس به خاطر چندین ویژگی خاص آن نسبت به روش‌های سنتی و دیگر روش‌های هوشمند مورد توجه قرار گرفته است. یکی از مهم‌ترین دلایل، قابلیت استخراج تابع پاداش از رفتار انسانی است. در بسیاری از سناریوها، نمی‌توان تابع پاداش واقعی را به راحتی مشخص کرد، اما یادگیری تقویتی معکوس این امکان می‌دهد تا با مشاهده رفتار یک کارشناس یا عامل خبره، این تابع را به دست آورد. این ویژگی کمک می‌کند تا رفتارهای پیچیده انسانی بهتر درک و شبیه‌سازی و به طور موثری محیط‌های مختلف مدل‌سازی شود. همچنین، یادگیری تقویتی معکوس به سیستم‌ها این امکان را می‌دهد که با تغییرات محیطی سازگار شوند و به طور خودکار خود را بهبود بخشند. مراحل یادگیری تقویتی این الگوریتم به شرح ادامه است (Arora and Doshi, 2021).

محیط یادگیری تقویتی معکوس به صورت یک فرایند تصمیم‌گیری مارکوف^۲ بدون تابع پاداش مطابق رابطه ۱ تعریف می‌شود.

$$MDP = (s, a, p(s'|s, a), \gamma) \quad \text{رابطه ۱}$$

که در آن، s مجموعه‌ای از حالت‌ها (State)، a مجموعه‌ای از اعمال یا اقدام‌ها (Actions)، $p(s'|s, a)$ تابع توزیع انتقال حالت s به s' وقتی که عمل a انجام می‌شود. γ عامل تخفیف، به عنوان یک ضریب که با آزمون و خطا تعیین می‌شود، که میزان اهمیت آینده در مقابل حال را مشخص می‌کند. برای یادگیری، مشاهده دنباله‌ای از جفت‌های حالت - عمل^۳ از عامل خبره بسیار مهم است. این جفت‌ها به صورت رابطه ۲ نمایش داده می‌شوند.

$$D = \{(s_1, a_1), (s_2, a_2), \dots, (s_T, a_T)\} \quad \text{رابطه ۲}$$

که در آن $s_1, \dots, s_T$ برابر با حالت‌های ۱ تا T و $a_1, \dots, a_T$ برابر با عمل‌های متناظر ۱ تا T است. در هر MDP، دو تابع ارزش مهم وجود دارد که شامل تابع ارزش^۴ و تابع ارزش عمل می‌باشد.

تابع ارزش $V^\pi(s)$ که امیدریاضی مجموع پاداش‌های تنزیل‌یافته از حالت s با پیروی از سیاست π است که به صورت رابطه ۳ تعریف می‌شود.

$$V^\pi(s) = \mathbb{E} \left[\sum_{t=0}^{\infty} \gamma^t R(s_t) \mid s_0 = s \right] \quad \text{رابطه ۳}$$

تابع ارزش عمل $Q^\pi(s, a)$ که امیدریاضی مجموع پاداش‌های تنزیل‌یافته از حالت s و عمل a با پیروی از سیاست π است که به صورت رابطه ۴ است.

$$Q^\pi(s, a) = \mathbb{E} \left[\sum_{t=0}^{\infty} \gamma^t R(s_t) \mid s_0 = s, a_0 = a \right] \quad \text{رابطه ۴}$$

هدف اصلی در یادگیری تقویتی معکوس، یافتن تابع پاداش $R(s_t)$ با مشاهده رفتار بهینه (یا رفتارهای نمونه) است. معمولاً، رفتار بهینه تحت سیاست بهینه π^* انجام می‌شود که در آن علامت * نشان دهنده عمل بهینه است. یکی از متداول‌ترین روش‌ها جهت تنظیم

1 Reward

2 Markov Decision Process

3 State-action

4 Value

پارامتر عامل تخفیف، که در این تحقیق نیز استفاده شد، تنظیم دستی است که شامل آزمون و خطا و بررسی آثار هر تغییر بر عملکرد الگوریتم از جهت زمان رسیدن به همگرایی و کیفیت پاسخ ارائه شده توسط الگوریتم می‌باشد.

مدل شبیه‌سازی ICSS

مدل‌سازی کانالهای آبیاری، یکی از ابزارهای کلیدی در بهینه‌سازی مدیریت منابع آب و افزایش بهره‌وری در بخش کشاورزی می‌باشد. این فرایند به شبیه‌سازی جریان آب و بررسی پارامترهای موثر در طراحی و بهره‌برداری کمک می‌نماید. ضرورت مدل‌سازی در این حوزه به دلیل کاهش منابع آب، نیاز به توزیع عادلانه و کارآمد آب و پیشگیری از هدررفت آن به خصوص در مناطق کم آب می‌باشد. همچنین، استفاده از مدل‌سازی امکان ارزیابی سناریوهای مختلف و برنامه‌ریزی بهتر برای پاسخگویی به نیازهای آینده را فراهم می‌کند. مدل ICSS در سال ۱۹۸۵ توسعه یافت. این مدل، یک مدل هیدرودینامیک یک بعدی است که به منظور شبیه‌سازی هیدرولیکی، هیدرولوژیکی و بهره‌برداری از سیستم‌های انتقال و توزیع آب که در طول زمان دارای تغییرات جریانات ورودی و خروجی و تغییرات شرایط فیزیکی سیستم هستند، طراحی شده است. این مدل توانایی شبیه‌سازی جریان‌های ماندگار و غیرماندگار را در شرایط مختلف بهره‌برداری از کانال‌های آبیاری داراست و می‌تواند انواع اشکال مقاطع کانال، طیف گسترده‌ای از سازه‌های هیدرولیکی، و همچنین جریان‌های ورودی و خروجی متنوع را پوشش دهد. به منظور ارزیابی دقت مدل، خروجی‌های حاصل از اجرای مدل توسط کمیته تخصصی مدل‌های ریاضی در جامعه مهندسين عمران کانادا (CSCE) در چهارچوب برنامه ارزیابی مدل‌های ریاضی هیدرودینامیک به طور دقیق مورد بررسی قرار گرفت و نتایج مورد تایید قرارگرفت. قبل از شروع به شبیه‌سازی با مدل، سه فایل اطلاعات ورودی شامل فایل‌های RCHN، ICSS4-CFG و ICSS4-REP بایستی تکمیل شوند. فایل RCHN یکی از فایل‌های ورودی کلیدی است که در آن اطلاعات فیزیکی و هیدرولیکی لازم جهت شبیه‌سازی وارد می‌شود. این اطلاعات شامل تعداد کانال‌ها، تعداد بازه‌ها در هر کانال، گام زمانی برای محاسبات هیدرولیکی و تعداد گره‌های هر بازه می‌باشد. در فایل ورودی ICSS4-CFG، ضرائب وزنی مرتبط با الگوی چهارنقطه‌ای در روش تفاضل محدود وارد می‌شود (Shahverdi, 2023). این مدل جهت شبیه‌سازی جریان غیرماندگار از معادلات سنت و نانت^۱ که یکی از کامل‌ترین مدل‌های شبیه‌سازی هیدرولیکی حرکت آب در مجاری روباز (کانال‌ها و زهکش‌ها) است، استفاده می‌کند. جمع‌آوری داده‌ها در قالب سه سناریوی اصلی با توجه به مقادیر دبی ورودی به کانال و دبی‌های آبیگرها، به عنوان معیارهای اصلی، با استفاده از ICSS انجام شد که در بخش بعدی به تفصیل تشریح شده است.

سناریوهای شبیه‌سازی با مدل ICSS

در این تحقیق به طور کلی سه سناریو شامل ۱۰۱ اجرا در نظر گرفته شد که با استفاده از مدل هیدرودینامیک ICSS شبیه‌سازی شد. اطلاعات ورودی مدل‌سازی شامل (دبی ورودی در بالادست کانال و دبی ورودی به آبیگرها) می‌باشند. همانگونه که قبلاً بیان شد شبیه‌سازیها در کانال E1-R1 از شبکه آبیاری دز با شش بازه صورت گرفت.

سناریو اول: مقدار دبی (Inflow)، که از بالادست وارد کانال می‌شود، برای اجراهای یک تا ۱۰۱ به صورت مشخص به ترتیب ۱، ۱/۰۵، ۱/۱، ۱/۱۵، ۱/۲۵ و ۱/۳ مترمکعب برثانیه در نظر گرفته شد. در این سناریو مقدار دبی (Q) ورودی به سازه‌های آبیگر براساس الگوی مشخص برای هر کدام از اجراها مشابه می‌باشد. به طور نمونه برای شروع، در اجرای شماره اول مقدار دبی ورودی از بالادست به کانال ۱ مترمکعب برثانیه و دبی ورودی به سازه‌ها به دلیل بسته بودن آبیگرها صفر می‌باشد. به ترتیب در اجراهای بعدی، در حالت کلی مقدار دبی ورودی به آبیگرها مقدار ۰/۰۰۲ مترمکعب برثانیه (۲ لیتر برثانیه) می‌باشد. زمان شبیه‌سازی نیز برابر ۲۴ ساعت در نظر گرفته شد.

سناریو دوم: باتوجه به موارد گفته شده در سناریو اول، در هنگام انجام شبیه‌سازی به دلیل تغییرات زیاد جریان و ناپایدار شدن شرایط کانال، به زمان شبیه‌سازی در صورت لزوم افزوده می‌شود. این کار تا زمانی که شرایط تاحد قابل قبولی پایدار شود تکرار می‌شود. به طور نمونه، در اجرای شماره ۴۴، دبی ورودی از بالادست به کانال که ۱/۱ مترمکعب برثانیه و مقدار دبی که به ازای هر یک از آبیگرها که عبور می‌کند برابر ۰/۰۸۶ مترمکعب برثانیه است، از این اجرا به بعد، زمان شبیه‌سازی با ۲۵ ساعت انجام شده است.

سناریو سوم: با افزایش مقدار دبی ورودی به سازه‌های آبیگر در هر اجرا و با در نظر گرفتن مقدار دبی ورودی از بالادست کانال، جهت انجام شبیه‌سازی و تامین دبی ورودی به آبیگرها، از اجرا شماره ۷۰ تا ۱۰۱، سازه آبیگر شماره دوم به صورت بسته در نظر گرفته شد تا جریان

از آن عبور نکنند. لازم به ذکر است که عمق آب در بالادست سازه‌های تنظیم کننده ۱/۲ متر معادل عمق هدف بوده که بایستی در کلیه اجراها عمق ذکر شده تثبیت شود. در جدول ۱ اطلاعات تعدادی از اجراهای حاصل از شبیه‌سازی با مدل هیدرودینامیکی ICSS قابل مشاهده است. در جدول مذکور inflow مقدار دبی ورودی، Qch1-Qch3 دبی عبوری از سازه‌های تنظیم کننده ۱-۳، Y1-Y3 عمق آب در بالادست سازه‌های تنظیم کننده ۱-۳، Goch1- Goch3 بازشدگی سازه‌های تنظیم کننده ۱-۳، GoTO1- GoTO6 بازشدگی آبگیرهای ۱-۶ و QTO1- QTO6 مقادیر دبی آبگیرهای ۱-۶ است.

جدول ۱- بخشی از اطلاعات خروجی شبیه‌سازی با مدل هیدرودینامیکی ICSS

اجرا	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰
inflow	۱/۰۰	۱/۰۰	۱/۰۰	۱/۰۰	۱/۰۰	۱/۰۰	۱/۰۰	۱/۰۰	۱/۰۰	۱/۰۰
Qch1	۱/۰۰	۰/۹۶۴	۰/۹۲۴	۰/۹۳۴	۰/۹۴۴	۰/۹۰۴	۰/۸۶۴	۱/۱۱۱	۱/۰۹۴	۱/۱۲۴
Qch2	۱/۰۰	۰/۹۲۸	۰/۸۴۸	۰/۸۱۸	۰/۷۸۸	۰/۷۰۷	۰/۶۲۸	۰/۸۴۲	۰/۷۸۲	۰/۷۷۲
Qch3	۰/۹۹۷	۰/۸۸۹	۰/۷۶۹	۰/۶۹۹	۰/۶۲۹	۰/۵۰۸	۰/۳۸۹	۰/۵۶۷	۰/۴۶۷	۰/۴۱۷
Y1	۱/۲۰	۱/۲۰	۱/۲۰	۱/۲۰	۱/۲۰	۱/۲۰	۱/۲۰	۱/۲۰	۱/۲۰	۱/۲۰
Y2	۱/۲۰	۱/۲۰	۱/۲۰	۱/۲۰	۱/۲۰	۱/۲۰	۱/۲۰	۱/۲۰	۱/۲۰	۱/۲۰
Y3	۱/۲۰	۱/۲۰	۱/۲۰	۱/۲۰	۱/۲۰	۱/۲۰	۱/۲۰	۱/۲۰	۱/۲۰	۱/۲۰
Goch1	۰/۳۸۵	۰/۳۷۱	۰/۳۵۳	۰/۳۵۸	۰/۳۶۳	۰/۳۴۶	۰/۳۳۰	۰/۴۳۲	۰/۴۲۴	۰/۴۳۶
Goch2	۰/۰۰	۰/۳۷۹	۰/۳۴۰	۰/۳۳۶	۰/۳۱۲	۰/۲۷۵	۰/۲۴۰	۰/۳۳۷	۰/۳۰۹	۰/۳۰۵
Goch3	۰/۰۰	۰/۳۱۷	۰/۲۷	۰/۲۴۳	۰/۲۱۷	۰/۱۷۲	۰/۱۲۹	۰/۱۹۳	۰/۱۵۷	۰/۱۳۹
GoTO1	۰/۰۰	۰/۰۱۲	۰/۰۲۶	۰/۰۴۰	۰/۰۵۴	۰/۰۶۸	۰/۰۸۲	۰/۰۹۵	۰/۱۱۰	۰/۱۲۴
GoTO2	۰/۰۰	۰/۰۰۹	۰/۰۲	۰/۰۳	۰/۰۴۱	۰/۰۵۱	۰/۰۶۲	۰/۰۰	۰/۰۰	۰/۰۰
GoTO3	۰/۰۰	۰/۰۰۹	۰/۰۲	۰/۰۳	۰/۰۴۱	۰/۰۵۱	۰/۰۶۲	۰/۰۷۱	۰/۰۸۲	۰/۰۹۳
GoTO4	۰/۰۰	۰/۰۱۲	۰/۰۲۶	۰/۰۴۰	۰/۰۵۴	۰/۰۶۹	۰/۰۸۳	۰/۰۹۶	۰/۱۱۰	۰/۱۲۴
GoTO5	۰/۰۰	۰/۰۱۲	۰/۰۲۶	۰/۰۴۰	۰/۰۵۴	۰/۰۶۸	۰/۰۸۳	۰/۰۹۶	۰/۱۱۰	۰/۱۲۴
GoTO6	۰/۰۰	۰/۰۰۹	۰/۰۲	۰/۰۳	۰/۰۴۱	۰/۰۵۱	۰/۰۶۲	۰/۰۷۱	۰/۰۸۲	۰/۰۹۳
QTO1	۰/۰۰	۰/۰۱۸	۰/۰۳۸	۰/۰۵۸	۰/۰۷۸	۰/۰۹۸	۰/۱۱۸	۰/۱۳۶	۰/۱۵۶	۰/۱۷۶
QTO2	۰/۰۰	۰/۰۱۸	۰/۰۳۸	۰/۰۵۸	۰/۰۷۸	۰/۰۹۸	۰/۱۱۸	۰/۰۰	۰/۰۰	۰/۰۰
QTO3	۰/۰۰	۰/۰۱۸	۰/۰۳۸	۰/۰۵۸	۰/۰۷۸	۰/۰۹۸	۰/۱۱۸	۰/۱۳۶	۰/۱۵۶	۰/۱۷۶
QTO4	۰/۰۰	۰/۰۱۸	۰/۰۳۸	۰/۰۵۸	۰/۰۷۸	۰/۰۹۸	۰/۱۱۸	۰/۱۳۶	۰/۱۵۶	۰/۱۷۶
QTO5	۰/۰۰	۰/۰۱۸	۰/۰۳۸	۰/۰۵۸	۰/۰۷۸	۰/۰۹۸	۰/۱۱۸	۰/۱۳۶	۰/۱۵۶	۰/۱۷۶
QTO6	۰/۰۰	۰/۰۱۸	۰/۰۳۸	۰/۰۵۸	۰/۰۷۸	۰/۰۹۸	۰/۱۱۸	۰/۱۳۶	۰/۱۵۶	۰/۱۷۶

با بهره‌گیری نتایج حاصل از خروجی مدل ICSS، می‌توان شاخص‌های ارزیابی عملکرد سیستم شامل کفایت، راندمان، عدالت و پایداری در تحویل را براساس نیاز آبی و الگوی توزیع زمانی آن در آبگیرهای مختلف محاسبه نمود.

شاخص‌های ارزیابی عملکرد

شاخص‌های استاندارد ارزیابی عملکرد، تعاریف کمی برای بهره‌برداری از کانال ارائه می‌دهند که به تفکیک مسائل مدیریتی از سیستم فیزیکی انتقال و توزیع کمک می‌کنند. مهم‌ترین ویژگی‌های این شاخص‌ها، بدون بُعد بودن و سهولت در اندازه‌گیری می‌باشند. در تحقیق حاضر جهت ارزیابی عملکرد توزیع و تحویل آب در شبکه‌های آبیاری از شاخص‌های کفایت (MPA)^۱، راندمان (MPF)^۲ تحویل ارائه شده توسط (Molden and Gates, 1990) و شاخص‌های خطای حداکثر مطلق (MAE)^۳ و خطای مطلق تجمعی (IAE)^۴ برای ارزیابی کنترل سطح آب که توسط ASCE (Clemmens et al., 1998) ارائه شده، استفاده می‌شود. شاخص کفایت تحویل معیاری است که بیان می‌کند آیا دبی به مقدار کافی توسط آبگیرها به مصرف کنندگان تحویل شده یا خیر. شاخص راندمان تحویل معیاری برای ارزیابی میزان حفظ منابع آبی است و نشان می‌دهد که تحویل آب مازاد بر مصرف وجود دارد یا خیر (Shahverdi and Maestre 2022). شاخص‌های

1 Measure of performance relative to adequacy

2 Measure of performance relative to efficiency

3 Maximum absolute error

4 Integral of absolute magnitude of error

کفایت و راندمان به ترتیب در روابط ۵ و ۶ ارائه شده‌اند.

$$\begin{aligned} \text{MPA} &= \frac{1}{N} \sum_N \frac{1}{M} \sum_M (\text{PA}), & \text{PA} &= \frac{QD}{QR} \quad \text{IF} \quad QR > QD \\ & & \text{PA} &= 1 \quad \text{IF} \quad 1 < \frac{QD}{QR} \\ \text{MPF} &= \frac{1}{N} \sum_N \frac{1}{M} \sum_M (\text{PF}), & \text{PF} &= \frac{QR}{QD} \quad \text{IF} \quad QD > QR \\ & & \text{PF} &= 1 \quad \text{IF} \quad 1 < \frac{QR}{QD} \end{aligned}$$

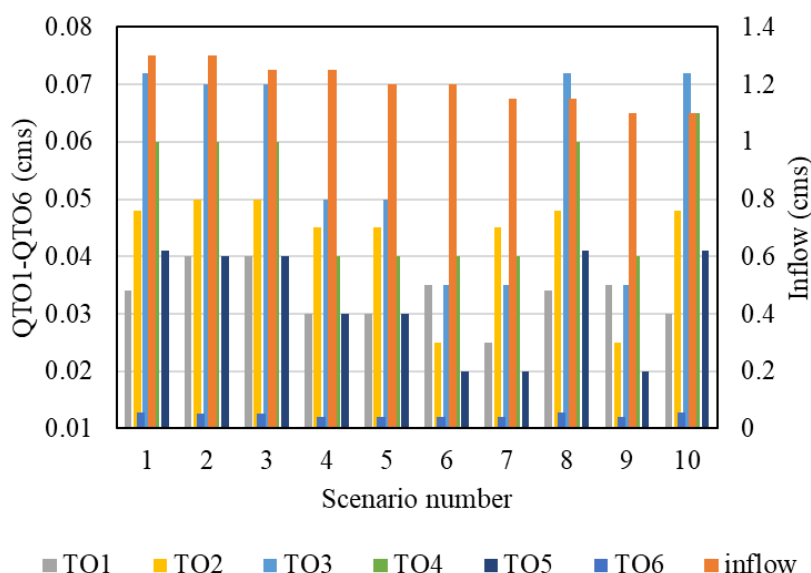
(رابطه ۵)

(رابطه ۶)

در روابط فوق، QR دبی مورد نیاز درخواست شده یا مورد تقاضای اراضی پائین هر دریاچه آبگیر، QD دبی واقعی تحویلی به هر دریاچه آبگیر، N تعداد دریاچه‌های آبگیر و T تعداد گام‌های زمانی مناسب در یک دوره تحویل است که برابر با نسبت طول دوره بهره‌برداری به طول گام زمانی محاسبات هیدرولیکی می‌باشد. لازم به ذکر است که مقدار ایده‌آل شاخص‌های کفایت و راندمان تحویل برابر با یک می‌باشد.

نتایج و بحث

با توجه به اینکه تعداد حالت‌های ممکن و قابل بهره‌برداری زیاد می‌باشد جهت انجام شبیه‌سازی در حالت کلی تعداد ۱۰ سناریو به صورت کامل تصادفی در نظر گرفته شد. مقادیر دبی ورودی به کانال در سناریوهای در نظر گرفته شده به ترتیب برابر با ۱/۱، ۱/۱۵، ۱/۲، ۱/۲۵ و ۱/۳ مترمکعب بر ثانیه است. مقادیر دبی آبگیرها نیز به صورت تصادفی بوده و محدوده وسیعی را شامل می‌شود. داده‌های مورد استفاده در این پژوهش با استفاده از مدل هیدرودینامیک کانال استخراج گردید. سناریوهای مختلف بهره‌برداری (با دبی‌های ورودی ۱، ۱، ۳، ۱ تا ۳، ۱ مترمکعب بر ثانیه و الگوهای تصادفی آبگیری) در این مدل شبیه‌سازی شدند به‌طوری‌که با دادن دبی ورودی به کانال و نیاز آبگیرها، بازسازی سازه‌ها به‌طوری‌که تنظیم گردید که شاخص‌های ارزیابی به مقادیر ایده‌آل خود نزدیک شوند. شکل ۳، با وارد کردن مقادیر دبی هر سناریو به مدل یادگیری تقویتی معکوس و با استفاده از پاداش یادگیری شده، مقادیر خروجی استخراج شد و سپس شاخص‌های ارزیابی محاسبه گردید.



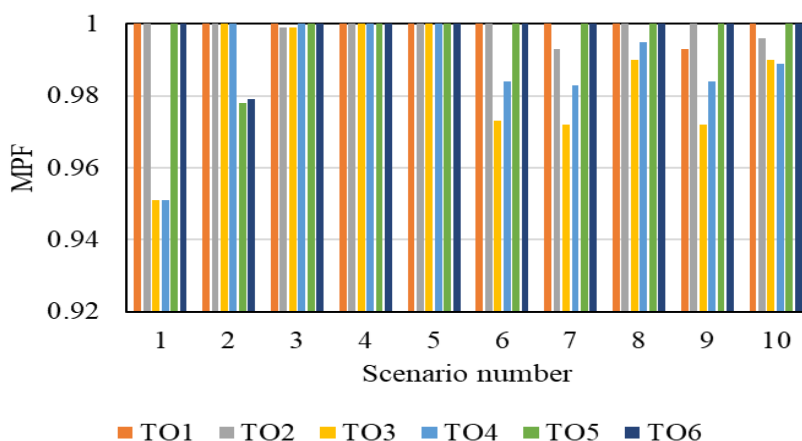
شکل ۳- سناریوهای تصادفی انتخاب شده برای ارزیابی نتایج الگوریتم یادگیری تقویتی معکوس (TO1-TO6: آبگیرهای ۱ تا ۶ و QTO1-QTO6: دبی آبگیرهای ۱ تا ۶)

با توجه به شکل ۳، در حالت کلی مقادیر حداقل و حداکثر دبی ورودی به آبگیر شماره یک به ترتیب ۲۵ و ۴۰ لیتر بر ثانیه، آبگیر شماره دو به ترتیب ۲۵ و ۵۰ لیتر بر ثانیه، آبگیر شماره سه به ترتیب ۳۵ و ۷۲ لیتر بر ثانیه، آبگیر شماره چهار به ترتیب ۴۰ و ۶۵ لیتر بر ثانیه، آبگیر شماره پنج ۲۰ و ۴۱ لیتر بر ثانیه و آبگیر شماره شش ۴۰ و ۵۵ لیتر بر ثانیه می‌باشد. با توجه سناریوهای تصادفی مطرح شده که تمامی شرایط

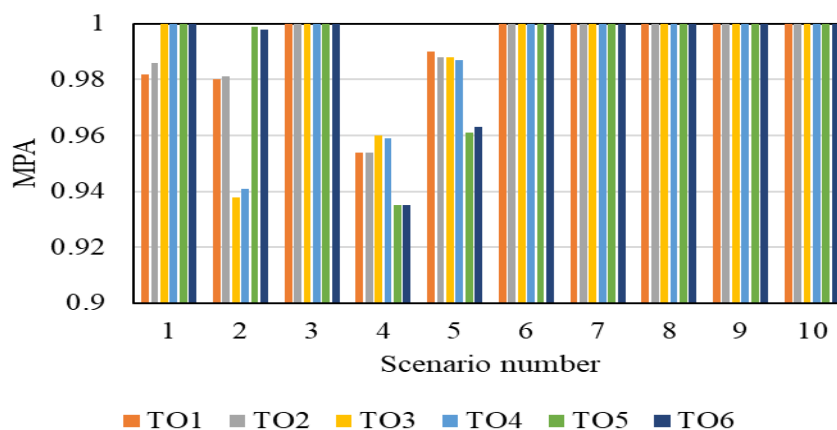
کانال را پوشش می‌دهند.

در شکل ۴ نتایج مقادیر شاخص راندمان در سناریوهای مختلف ارائه شده است. همان‌طور که مشاهده می‌شود، در اغلب سناریوها مقادیر محاسبه شده برای شاخص راندمان تحویل، نزدیک به یک می‌باشند. پارامترهای موثر در محاسبه مقادیر این شاخص شامل تعداد آگیرها، دبی موردنیاز یا دبی درخواست شده و دبی واقعی تحویلی می‌باشند. بر اساس محدوده استاندارد شاخص‌های ارزیابی عملکرد مولدن و گیتس برای شاخص راندمان، اگر مقادیر در محدوده ۰/۸۵ تا ۱ قرار گرفته باشند دبی با راندمان بالاتری به مصرف‌کنندگان تحویل می‌شود و عملاً تلفات آب یا مازاد تحویل به حداقل خود می‌رسد. در این تحقیق، براساس مقادیر بدست آمده، حداقل و حداکثر راندمان تحویل به ترتیب ۰/۹۷ و ۱ می‌باشد که در محدوده استاندارد قرار گرفته و در وضعیت خوبی قرار دارند. در حالت کلی، مقادیر راندمان آگیرهای شماره یک، دو، پنج و شش در تمامی سناریوها در حالت ایده آل خود قرار دارند. همچنین، به طور میانگین راندمان تحویل در آگیرهای شماره یک تا شش به ترتیب ۱، ۱، ۰/۹۸، ۰/۹۸، ۱ و ۱ بدست می‌آید.

در شکل ۵، مقادیر شاخص کفایت نشان داده شده است. همان‌طور که مشاهده می‌شود، در اکثریت سناریوها مقادیر محاسبه شده برای شاخص کفایت تحویل، نزدیک به یک می‌باشند. پارامترهای موثر در محاسبه مقادیر این شاخص شامل تعداد آگیرها، دبی موردنیاز یا دبی درخواست شده و دبی واقعی تحویلی می‌باشند. براساس محدوده استاندارد شاخص‌های ارزیابی عملکرد مولدن و گیتس برای شاخص کفایت اگر مقادیر در محدوده ۰/۹ تا ۱ قرار گرفته باشند، دبی به مقدار کافی توسط آگیرها به مصرف‌کنندگان تحویل داده می‌شود. در این تحقیق، بر اساس مقادیر بدست آمده، حداقل و حداکثر کفایت تحویل به ترتیب ۰/۹۵ و ۱ می‌باشد که در محدوده استاندارد قرار گرفته و در وضعیت خوبی قرار دارند. در حالت کلی، مقادیر کفایت تحویل در آگیرهای شماره سه، چهار، پنج و شش در تمامی سناریوها در حالت ایده آل خود قرار دارند. ضمن اینکه به طور میانگین کفایت تحویل در آگیرهای شماره یک و دو ۰/۹۹ و در آگیرهای شماره سه تا شش ۱ حاصل شده است.

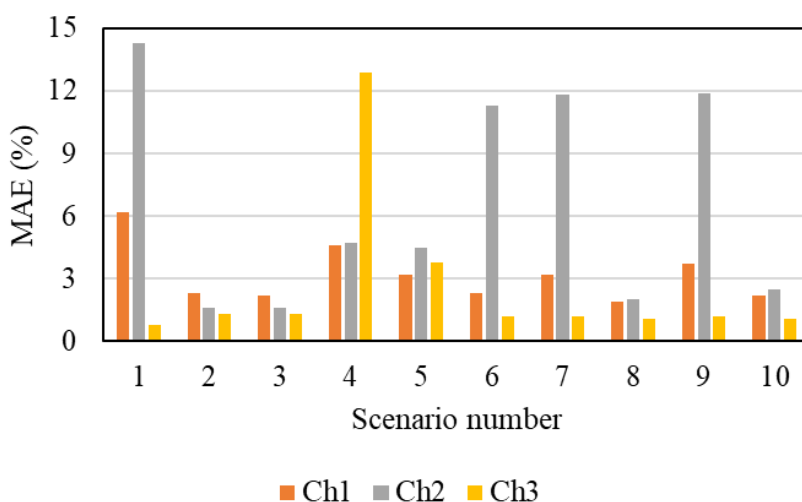


شکل ۴- مقادیر شاخص راندمان تحویل در فرایند یادگیری تقویتی معکوس (TO1-TO6: آگیرهای ۱ تا ۶)



شکل ۵- مقادیر شاخص کفایت تحویل در فرایند یادگیری تقویتی معکوس (TO1-TO6: آگیرهای ۱ تا ۶)

در شکل ۶ مقادیر شاخص حداکثر خطای عمق نشان داده شده است. در سازه آب‌بند شماره یک، مقادیر خطای حداکثر مطلق کمتر از مقدار ۱۰ درصد می‌باشد. حداقل و حداکثر خطا به ترتیب برابر با ۱/۹ و ۶/۹ درصد و میانگین مقدار این خطا برای تمامی سناریوها برابر با ۳/۱۸ درصد محاسبه شده که همگی در وضعیت خوبی قرار گرفته‌اند. در سازه آب‌بند شماره دو، مقادیر خطای حداکثر مطلق برای سناریوهای ۲، ۳، ۴، ۵، ۸ و ۱۰ درصد، در سناریوهای ۱، ۶، ۷ و ۹ کمتر از ۲۰ درصد و حداقل و حداکثر خطا به ترتیب برابر با ۱/۶ و ۱۴/۳ درصد و میانگین مقدار این خطا برای تمامی سناریوها برابر با ۶/۶۲ درصد محاسبه شده که در مجموع همگی در وضعیت قابل قبولی قرار گرفته‌اند. در سازه آب‌بند شماره سه، مقادیر خطای حداکثر مطلق برای تمامی سناریوها به جز سناریو شماره چهار کمتر از مقدار ۱۰ درصد و در سناریوی شماره چهار این مقدار ۱۲/۹ به‌دست آمده است که نسبت به بقیه سناریوها افزایش یافته است و در محدوده مجاز قرار گرفته است. لازم به ذکر است که در این سناریو با آنکه مقدار خطای بیشتر می‌باشد با این حال چون در محدوده مجاز قرار گرفته است؛ بنابراین، مدل حساسیتی به آن نشان نداده است. با این حال، می‌توان با کاهش محدوده مجاز و اجرای مجدد مدل به مقادیر کمتر نیز دست یافت. حداقل و حداکثر خطا به ترتیب ۰/۸ و ۱۲/۹ درصد و میانگین مقدار خطای حداکثر مطلق تمامی سناریوها برای سازه آب‌بند شماره سه ۲/۵۹ درصد می‌باشد. در حالت کلی، سازه آب‌بند شماره سه نسبت به دو سازه تنظیم‌کننده دیگر عملکرد بهتری داشته است.

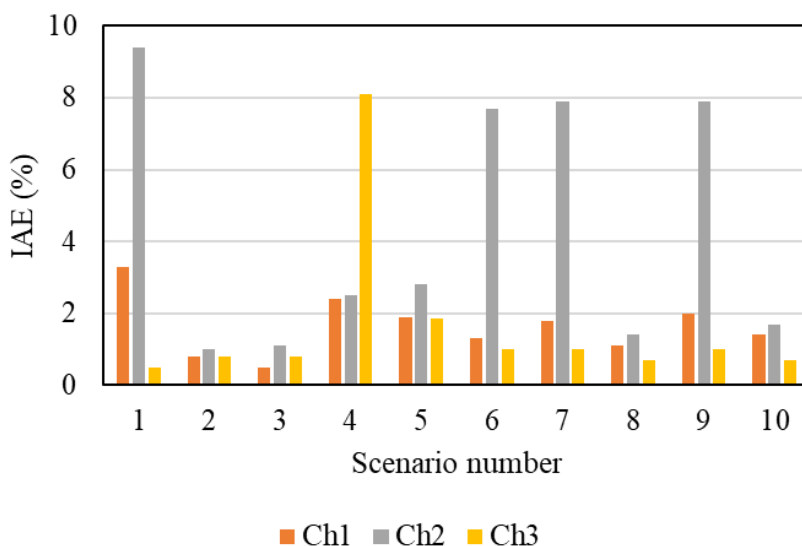


شکل ۶- مقادیر شاخص خطای حداکثر مطلق در فرایند یادگیری تقویتی معکوس (Ch1-Ch3: سازه‌های تنظیم‌کننده ۱ تا ۳)

مقادیر شاخص خطای تجمعی در فرایند یادگیری معکوس به صورت شکل ۷ ارائه شده است. در سازه آب‌بند شماره یک، مقادیر خطای مطلق تجمعی کمتر از مقدار ۱۰ درصد می‌باشد، حداقل و حداکثر خطا مطلق تجمعی به ترتیب ۰/۵ و ۲/۴ درصد و میانگین مقدار این خطا برای تمامی سناریوها برابر با ۱/۶۵ درصد محاسبه شده که همگی در وضعیت خوبی قرار گرفته‌اند. در سازه آب‌بند شماره دو، مقادیر خطای مطلق تجمعی کمتر از مقدار ۱۰ درصد می‌باشد، همچنین حداقل و حداکثر خطا مطلق تجمعی به ترتیب برابر با ۱ و ۹/۴ درصد و میانگین مقدار این خطا برای تمامی سناریوها ۴/۳۴ درصد محاسبه شده که در مجموع همگی در وضعیت خوبی قرار گرفته‌اند. در سازه آب‌بند شماره سه، مقادیر خطای مطلق تجمعی کمتر از مقدار ۱۰ درصد بوده به‌طوری که بیشتر سناریوها، مقادیر نزدیک به صفر می‌باشند. حداقل و حداکثر خطا به ترتیب ۰/۵ و ۸/۱ درصد و میانگین مقدار خطای حداکثر مطلق تمامی سناریوها برای سازه آب‌بند شماره سه ۱/۶۴ درصد می‌باشد. در حالت کلی نسبت به دو سازه تنظیم‌کننده قبلی عملکرد بهتری داشته است. لازم به ذکر است که شاخص خطای مطلق تجمعی، نشان‌دهنده میانگین انحرافات عمق آب از عمق هدف در طول بهره‌برداری است. هر چقدر این شاخص کمتر باشد، نوسانات حداکثر، میانگین سطح آب کمتر و در نهایت سیستم عملکرد بهتری خواهد داشت.

با توجه به مقادیر به‌دست آمده برای شاخص‌ها، مشاهده می‌شود تطابق مناسبی بین خروجی سیستم و نیاز واقعی شبکه وجود دارد، به‌طوری که حتی در شرایطی که دبی ورودی و خروجی سیستم به‌صورت تصادفی و در محدوده وسیعی تغییر کرده است، سیستم قادر به حفظ سطح مطلوبی از عملکرد است. این موفقیت را می‌توان به توانایی یادگیری تقویتی معکوس در یادگیری دقیق الگوهای بهره‌برداری و تبدیل آن به سیاست‌های بهینه، در شرایط عملیاتی واقعی که با عدم قطعیت‌های زیادی همراه است، نسبت داد. مقایسه نتایج این پژوهش با Savari et al. (2016) نشان می‌دهد که سیستم پیشنهادی از نظر شاخص‌های عملکردی مانند راندمان و کفایت تحویل آب، بهبود

محسوسی نسبت به روش‌های سنتی کنترل و بهره‌برداری دستی داشته است و حتی در مقایسه با روش یادگیری تقویتی، در تعداد حالات بیشتری به مقادیر ایده‌آل رسیده است. اگرچه نتایج این پژوهش امیدبخش است، اما برخی محدودیت‌ها نیز وجود دارد که می‌تواند زمینه‌ساز تحقیقات آینده باشد. از جمله این که مطالعه حاضر بر روی یک شبکه آبیاری با مشخصات خاص انجام شده و تعمیم‌پذیری آن به شبکه‌های با مقیاس بزرگتر نیاز به بررسی بیشتر دارد. همچنین، در این پژوهش فرض شده که داده‌های آموزشی از کیفیت مناسبی برخوردار هستند، در حالی که در شرایط واقعی ممکن است داده‌های نویزی یا ناقص وجود داشته باشد. استفاده از تجهیزات سخت‌افزاری مانند کامپیوتر با سرعت مناسب می‌تواند منجر به پیاده‌سازی راحت‌تر این مدل شود.



شکل ۷- مقادیر شاخص خطای مطلق تجمعی در فرایند یادگیری تقویتی معکوس (Ch1-Ch3: سازه‌های تنظیم‌کننده ۱ تا ۳)

نتیجه‌گیری

در این تحقیق برای اولین بار، روش یادگیری تقویتی معکوس به منظور استخراج دستورالعمل‌های بهره‌برداری معرفی و مدل مربوطه در نرم‌افزار متلب تهیه گردید. در این الگوریتم تابع پاداش با توجه به ورودی‌های مساله یادگیری می‌شود به‌طوری‌که ورودی و خروجی با مقدار پاداش حداکثر در آن ذخیره می‌شود. یکی از اجزاء اصلی الگوریتم یادگیری تقویتی معکوس، مدل محیط می‌باشد. که همواره با آن در تعامل است. مدل محیط در این تحقیق، مدل هیدرودینامیکی ICSS بوده که براساس تجربیات فرد (خبره) شبیه‌سازی صورت گرفته است. عامل یادگیری توسط الگوریتم در این بخش سازه‌های آبگیر است که در شرایط مختلف بهره‌برداری کنش‌های متفاوتی را انجام می‌دهد. محیط یادگیری در یادگیری تقویتی معکوس، کانال E1-R1 دز بود. با استفاده از این روش تجربیات بهره‌برداری در قالب یک تابع پاداش تعریف شده و در طول فرایند یادگیری، نتایج حاصل از شبیه‌سازی که با مدل ICSS انجام شده در هر تکرار در قالب مدل محیط به صورت گسسته در الگوریتم یادگیری تقویتی معکوس فراخوانی می‌شود. بر اساس سناریوهای مختلفی که تعریف شد و نتایج شبیه‌سازی و یادگیری، شاخص‌های ارزیابی همگی در کلاس عملکردی خوب قرار داشتند که نشانگر عملکرد موفقیت‌آمیز الگوریتم یادگیری تقویتی معکوس بود و نشانگر معتبر بودن تابع پاداش یادگیری شده می‌باشد. نتایج شاخص‌های ارزیابی عمق برای روش یادگیری تقویتی معکوس در همه سناریوها نشان داد که مدل همواره توانسته است تغییرات سطح آب را درون محدوده مجاز عمق تثبیت کند. عمق در تمامی سناریوهای شبیه‌سازی شده در محدوده مجاز عمق قرار گرفته است. همچنین، متوسط تغییرات عمق نیز در اطراف عمق هدف قرار گرفته است. یافته‌های این پژوهش به وضوح نشان می‌دهد که روش یادگیری تقویتی معکوس می‌تواند به عنوان یک ابزار قدرتمند در مدیریت بهینه شبکه‌های آبیاری مورد استفاده قرار گیرد. توانایی این روش در یادگیری از تجربیات بهره‌برداران و تبدیل آن به سیاست‌های بهینه، راه‌حلی نوین برای چالش‌های موجود در مدیریت شبکه‌های آبیاری ارائه می‌دهد. نتایج حاصل از ارزیابی‌های مختلف حاکی از آن است که سیستم پیشنهادی نه تنها از کارایی بالایی برخوردار است، بلکه قابلیت اطمینان مناسبی نیز در شرایط عملیاتی مختلف از خود نشان می‌دهد. با توجه

به اینکه در این تحقیق، داده‌های مورد استفاده، حاصل مدل شبیه‌سازی بود پیشنهاد می‌شود کارایی این روش در شبکه‌های بزرگ مقیاس با پیچیدگی بیشتر و با داده‌های واقعی مورد بررسی قرار گیرد. بررسی امکان تلفیق این روش با دیگر روش‌های هوش مصنوعی مانند یادگیری عمیق نیز می‌تواند به عنوان یک مسیر تحقیقاتی جذاب مورد توجه قرار گیرد.

"هیچ‌گونه تعارض منافع بین نویسندگان وجود ندارد"

منابع

- بیات، فاطمه، قدوسی، حسام، و شاهوردی، کاظم (۱۴۰۱). بهره‌برداری از کانال‌های آبیاری با استفاده از روش‌های هوشمند. *فصلنامه علوم محیطی*، ۲۰(۱): ۷۶-۵۷.
- توانا، امیر حسین (۱۴۰۳). کاربرد روش یادگیری تقویتی معکوس در شبکه‌های آبیاری (مطالعه موردی: شبکه آبیاری دز). پایان‌نامه کارشناسی ارشد. به‌راهنمایی حسام قدوسی. زنجان: دانشگاه زنجان، دانشکده کشاورزی.
- شاهوردی، کاظم (۱۴۰۱). استخراج الگوهای بهره‌برداری با درنظرگرفتن خطای میراب در تنظیم سازه‌ها در شبکه آبیاری. *مجله مدیریت آب و آبیاری*، ۱۲(۲): ۳۴۹-۳۵۸.
- شاهوردی، کاظم، قدوسی، حسام، و دوستی، فرشاد (۱۴۰۲). اثر بهره‌برداری به‌هنگام و دیر هنگام سازه‌های کانال بر توزیع عادلانه و پایدار آب. *مجله تحقیقات آب و خاک ایران*، ۵۵(۳): ۴۶۷-۴۸۱.
- عباسی، فریبرز، و عباسی، نادر (۱۴۰۲). تحلیلی بر راندمان‌های آبیاری ایران در بستر زمان. *نشریه آبیاری و زهکشی ایران*، ۱۷(۶): ۱۰۳۳-۱۰۲۵.
- قدوسی، حسام، و داداشی، زهرا (۱۳۹۵). ارزیابی و مقایسه مدل‌های Rootcanal و ICSS برای شبیه‌سازی جریان در کانال‌های آبیاری (مطالعه موردی شبکه آبیاری قوریچای). *مجله پژوهش آب ایران*، ۱۰(۲): ۱۱۶-۱۰۷.

REFERENCES

- Abbasi, F., & Abbasi, N. (2024). An analysis of irrigation efficiencies over time in Iran. *Iranian Journal of Irrigation & Drainage*, 17(6), 1025-1033. (In Persian)
- Aghayee, Z., Ghodousi, H., & Shahverdi, K. (2024). Uncertainty Analysis of Irrigation Canals Operation. *Iranian Journal of Science and Technology, Transactions of Civil Engineering*, 48(6), 4769-4779.
- Alibabaei, K., Gaspar, P. D., Assuncao, E., Alirezazadeh, S., & Lima, T. M. (2022). Irrigation optimization with a deep reinforcement learning model: Case study on a site in Portugal. *Agricultural Water Management*, 263: 0378-3774.
- Arora, S., & Doshi, P. (2021). A survey of inverse reinforcement learning: Challenges, methods and progress. *Artificial Intelligence*, 297, 103500.
- Bayat, F., Ghodousi, H., and Shahverdi, K. (2022). Operation of irrigation canals using intelligent methods. *Environmental Sciences*, 20(1), 57-76. (In Persian)
- Clemmens, A. J., Kacerek, T. K., Grawitz, B., & W, Schuurmans. (1998). Test cases for canal control algorithms. *Journal of irrigation and drainage engineering* 124(1): 23-30.
- Ding, X., & Du, W. (2024). Optimizing irrigation efficiency using deep reinforcement learning in the field. *ACM Transactions on Sensor Networks*, 20(4), 1-34.
- Garcia-Espinal, M. A., Pérez-Sánchez, M., Sánchez-Romero, F. J., & López-Jiménez, P. A. (2024). Irrigation Distribution Network Design Parameters and Their Influence on Sustainability Management. *Water*, 16(8), 1131.
- Ghodousi, H., and Dadashi, Z. (2016). Evaluation and comparison of ICSS and Rootcanal models for simulation of flow in irrigation canals (Case study: Ghoorichay irrigation network). *Iranian Water Researches Journal*, 10(2), 107-116. (In Persian)
- Mohebbi, H., Ghodousi, H., & Shahverdi, K. (2024). Evaluating the impact of climate change on irrigation canal performance. *Irrigation and Drainage*, 73(3), 961-973.
- Molden, D. J., & Gates, T. K. (1990). Performance measures for evaluation of irrigation-water-delivery systems. *Journal of irrigation and drainage engineering*, 116(6): 804-82.
- Omidzade, Fatemeh, Hesam Ghodousi, and Kazem Shahverdi. Comparing fuzzy SARSA learning and ant Colony optimization algorithms in water delivery scheduling under water shortage conditions. *Journal of Irrigation and Drainage Engineering* 146.9 (2020): 04020028.
- Savari, H., Monem, M. J. and Shahverdi, K. (2016). Comparing the performance of FSL and traditional operation methods for on-request water delivery in the Aghili network, Iran. *Journal of Irrigation and*

- Drainage Engineering*, 142(11): 04016055.
- Shahdany, S. H., Majd, E. A., & Firoozfar, A. (2016). Improving Operation of a Main Irrigation Canal Suffering from Inflow Fluctuation within a Centralized Model Predictive Control System: Case Study of Roodasht Canal, Iran. *Journal of Irrigation and Drainage Engineering*, 142(11): 05016007.
- Shahverdi, K., & Javad Monem, M. (2022). Irrigation canal control using enhanced Fuzzy SARSA learning. *Irrigation and Drainage*, 71(3), 766-775.
- Shahverdi, k. (2022). Determining operational patterns considering operator's error in structures settings in irrigation networks. *Water and Irrigation Management*, 12(2), 349-358. (In Persian)
- Shahverdi, K. (2023). AICSS: Automatic simulator-controller/optimizer model of open channels. *Irrigation and Drainage*, 72(4), 1124-1136.
- Shahverdi, K., & Maestre, J. M. (2022). Gray wolf optimization for scheduling irrigation water. *Journal of Irrigation and Drainage Engineering*, 148(7), 04022020.
- Shahverdi, K, Ghodousi, H, & Dosti, F. (2024). Effect of on time and late operation of canal structures for equitable and sustainable water distribution. *Iranian Journal of Soil and Water Research*, 55 (3), 467-481. (In Persian)
- Shin, J. H., Nam, W. H., Jeon, M. G., Hong, E. M., Zhang, X., Sharma, V., ... & Kisekka, I. (2024). Assessing water distribution and efficiency by coupled hydraulic-hydrological modeling for irrigation canal network. *Paddy and Water Environment*, 1-14.
- Tavana, A. H. (2024). *Application of Inverse Reinforcement Learning Method in Irrigation Networks (Case Study: Dez Irrigation Network)*. Master dissertation, Zanjan University, Zanjan, Iran.
- Zhou, K., Fan, Y., Gao, Z., Chen, H., & Kang, Y. (2024). Research progress on operation control and optimal scheduling of irrigation canal systems. *Irrigation and Drainage*.