

A Graph-Based Deep Learning Framework for Forecasting the Dynamics of Socio-Ecological Systems: A Case Study of the Sistan Watershed

Babak Masoudi^{1*}, Safoora Bazzi², Mohammad SalKhordeh Haghighi³

¹Department of Information Technology, Payame Noor University (PNU), Tehran, Iran

²Department of Biology, Payame Noor University (PNU), Tehran, Iran

³Department of civil engineering, University of Zabol, Zabol, I.R of Iran

Corresponding Author's E-mail: b.masoudi@pnu.ac.ir

Received: 30 Oct. 2025
Revised: 25 Nov. 2025

Accepted: 19 Feb. 2026
Published: 21 June 2024



Homepage: ijwer.uoz.ac.ir

Abstract: Arid socio-ecological systems (SEs), such as the Sistan Watershed, face critical crises from wetland desiccation and dust storms. Modeling their non-linear dynamics is often hindered by data scarcity and the opacity of conventional AI. This study introduces a "gray-box" Graph Neural Network (GNN) framework for predictive modeling and scenario simulation. By leveraging a reconstructed 21-year aerosol index time series, we developed a process-informed, decomposition-based GNN that embeds the system's causal structure. The model demonstrated robust performance on unseen data (R^2 of 0.708 for water presence and 0.810 for aerosol index), significantly outperforming several baselines. Simulations reveal that wetland restoration could improve regional air quality by 16%, reducing estimated $PM_{2.5}$ concentrations by over $160 \mu g/m^3$. This framework serves as an interpretable "digital twin," providing a powerful quantitative decision-support tool for managing complex environmental systems..

Keywords: Socio-Ecological System, Deep Learning, Decomposition Modeling, Graph Neural Network, Sistan Watershed, Hydrology Indices, Dust Pollution

Introduction: Socio-ecological systems (SEs) in arid regions are increasingly vulnerable to climate change and anthropogenic pressures, often approaching critical "tipping points" that threaten ecosystem services and human well-being (Scheffer et al., 2001). The Sistan Basin, a transboundary watershed on the Iran-Afghanistan border, epitomizes this crisis. The desiccation of its Hamoun Wetlands has triggered severe dust storms (Yousefi-Kebriya et al. 2025), impacting public health and regional stability (Karimi et al. 2025; Zhou et al. 2025). While advanced predictive models are urgently needed, traditional approaches face limitations. Previous studies have successfully established statistical correlations between wetland shrinkage and dust pollution (e.g., (Yousefi-Kebriya et al. 2025)), but they often provide a retrospective analysis with limited forecasting capability. On the other hand, data-driven AI models often act as "black boxes," learning spurious correlations and lacking the interpretability required for high-stakes policy-making (Reichstein et al. 2019).

This research addresses this methodological gap by developing a predictive, process-informed, "gray-box" deep learning framework. Moving beyond correlational analysis, our primary objective is to create a dynamic and robust "digital twin" of the Sistan SES. This model is designed not only to forecast the system's future dynamics with high accuracy but also to serve as a quantitative tool for evaluating the impact of different management interventions, thereby providing a robust decision-support system.

Methodology: This study employed a multi-stage framework. First, we conducted data acquisition and enrichment. Time-series data (2001-2021) for six key environmental variables were extracted from multi-source satellite data using Google Earth Engine (details in Table 1). To address a significant gap in the aerosol time series, a machine learning-based data fusion technique was implemented. Subsequently, an extensive feature engineering

step was performed, creating multiple lag and seasonal features to enrich the dataset.

Second, we developed a novel decomposition-based, process-informed Graph Neural Network (GNN). This approach trains the model to predict the residuals of a simple persistence baseline, focusing the model's capacity on learning complex, non-linear patterns that simpler models cannot capture. The model architecture embeds a predefined causal graph (Fig. 3), which consists of node-specific Gated Recurrent Unit (GRU) layers to capture individual temporal patterns for each variable, and a Graph Convolutional Network (GCN) layer to fuse this information and learn systemic interdependencies (Fig. 4). The model's optimal hyperparameters were determined via systematic grid search, and it was trained using an Adam optimizer and an MSE loss function. Finally, a series of post-hoc analyses were performed, including a comprehensive and comparative performance evaluation against baseline models (SARIMA, LSTM, Persistence), and the simulation of future management scenarios.

Results and Discussion: The trained model demonstrated excellent and robust predictive performance on the unseen test set (January 2019 - December 2021). As detailed in Table 2, our final GNN model achieved an R^2 of 0.708 for water presence and a remarkable R^2 of 0.810 for the aerosol index. This performance significantly surpassed that of classical statistical (SARIMA) and simple persistence models and also outperformed a standard deep learning baseline (LSTM) in forecasting the key aerosol index variable. This confirms the value of our advanced decomposition approach and process-informed architecture. A comprehensive evaluation using multiple metrics is provided in Table 3. The comparison plots in Fig. 5 visually validate the model's ability to capture the core seasonal and long-term dynamics of the system.

Scenario simulations yielded critical, actionable insights for policymaking. We simulated two scenarios over a 5-year horizon: a 'Baseline' and a 'Wetland Restoration' scenario, incorporating uncertainty analysis via Monte Carlo Dropout. The restoration scenario, which involves maintaining water presence above its historical 75th percentile, projected a powerful positive cascade across the ecosystem. As detailed quantitatively in Table 4 and visualized in Fig. 6, this intervention led to a significant 16.1% improvement (reduction) in the aerosol index compared to the baseline.

To translate this finding into a more tangible, health-oriented metric, we employed a validated conversion model from the literature (Hassaan et al., 2023) to estimate particulate matter concentrations. The analysis revealed that the 16.1% improvement in the aerosol index corresponds to a substantial estimated reduction of over 166 $\mu\text{g}/\text{m}^3$ in average PM_{2.5} concentrations. This result quantitatively demonstrates that proactive, internal management interventions can be a highly effective lever for mitigating a key regional environmental hazard. This finding is consistent with studies in similar regions of Iran (e.g., Yousefi-Kebriya et al., 2025), but provides a forward-looking, quantitative forecast of a specific policy's impact. The findings strongly suggest that hydrological restoration of the Hamoun Wetlands is a powerful local strategy for combating the regional dust crisis.

Conclusion: This study successfully developed and validated a novel, decomposition-based, process-informed deep learning framework for modeling complex, data-scarce socio-ecological systems. The model demonstrated superior predictive performance compared to several standard baselines, and through robust scenario simulations, provided critical, actionable insights. Our findings quantitatively confirm that the hydrological restoration of the Hamoun Wetlands is a highly effective strategy for mitigating local dust pollution, projected to reduce PM_{2.5} concentrations by a significant margin. This framework serves as a powerful and interpretable "digital twin," offering a quantitative decision-support tool to guide sustainable management and enhance resilience in the Sistan Basin and other vulnerable arid regions. Future work should focus on external validation of the reconstructed data and incorporating the direct modeling of upstream water management policies.



یک چارچوب یادگیری عمیق مبتنی بر گراف برای پیش‌بینی دینامیک سیستم‌های اجتماعی-اکولوژیکی: مطالعه موردی حوضه آبریز سیستان

بابک مسعودی^{۱*}، صفورا بزی^۲، محمد سالخورده حقیقی^۳

^۱استادیار، گروه مهندسی فناوری اطلاعات، دانشگاه پیام نور، تهران، ایران

^۲استادیار، گروه زیست‌شناسی، دانشگاه پیام نور، تهران، ایران

^۳استادیار، گروه مهندسی عمران، دانشگاه زابل، زابل، ایران

پست الکترونیکی نگارنده مسئول: b.masoudi@pnu.ac.ir

وبگاه نشریه: ijwer.uoz.ac.ir



تاریخ دریافت: ۱۴۰۴/۰۷/۰۸ تاریخ پذیرش: ۱۴۰۴/۱۱/۳۰

تاریخ ویرایش: ۱۴۰۴/۰۹/۰۴ تاریخ چاپ: ۱۴۰۳/۰۴/۰۱

چکیده: سیستم‌های اجتماعی-اکولوژیکی (SES) در مناطق خشک، با بحران‌های شدیدی از جمله خشکیدگی تالاب‌ها و طوفان‌های گرد و غبار روبرو هستند. مدل‌سازی دینامیک پیچیده و غیرخطی این سیستم‌ها به دلیل کمبود داده و ماهیت "جعبه سیاه" مدل‌های هوش مصنوعی رایج، یک چالش بزرگ محسوب می‌شود. این پژوهش یک چارچوب نوآورانه "جعبه خاکستری" مبتنی بر یادگیری عمیق و شبکه‌های عصبی گراف (GNN) را برای مدل‌سازی پیش‌بینی‌کننده و شبیه‌سازی سناریو در این سیستم‌ها ارائه می‌دهد. ابتدا، با استفاده از مهندسی ویژگی و تلفیق داده مبتنی بر یادگیری ماشین، یک مجموعه داده غنی‌شده و کامل در دوره‌ای ۲۱ ساله برای متغیرهای کلیدی سیستم ایجاد شد. سپس، یک معماری پیشرفته GNN مبتنی بر تجزیه (Decomposition-GNN) که ساختار علی سیستم در آن تعبیه شده بود، برای پیش‌بینی دینامیک‌های پیچیده آموزش داده شد. نتایج اعتبارسنجی بر روی داده‌های دیده‌نشده، عملکرد برتر مدل را در مقایسه با مدل‌های پایه (شامل SARIMA و LSTM) به اثبات رساند و به ضریب تعیین (R^2) برابر با ۰.۷۰۸ برای وسعت آب و ۰.۸۱۰ برای شاخص آئروسول دست یافت. شبیه‌سازی سناریوهای مدیریتی نشان داد که مداخله فعال برای احیای تالاب می‌تواند به طور قابل توجهی (بهبود ۱۶ درصدی در شاخص آئروسول) منجر به کاهش گرد و غبار محلی شود، که معادل کاهش تخمینی بیش از ۱۶۰ میکروگرم بر متر مکعب در غلظت PM2.5 است. این چارچوب، یک "دوقلوی دیجیتال" قدرتمند و قابل تفسیر برای پشتیبانی از تصمیم‌گیری و مدیریت پایدار در سیستم‌های پیچیده محیطی ارائه می‌دهد.

کلیدواژه‌ها: سیستم اجتماعی-اکولوژیکی، یادگیری عمیق، شبکه عصبی گراف، مدل‌سازی مبتنی بر تجزیه، حوضه آبریز سیستان، شاخص‌های هیدرولوژی، آلودگی گرد و غبار.

۱- مقدمه

سیستم‌های اجتماعی-اکولوژیکی^۱ (SES) که بیانگر برهم‌کنش‌های پیچیده بین جوامع انسانی و محیط زیست هستند، در مناطق خشک و نیمه‌خشک با آسیب‌پذیری فزاینده‌ای در برابر فشارهای اقلیمی و انسانی مواجه‌اند (Amani et al. 2025; Kanan and Giupponi 2024). این سیستم‌ها اغلب در نزدیکی آستانه‌های بحرانی یا "نقاط عطف"^۲ قرار دارند، بطوری که تغییرات کوچک می‌توانند منجر به دگرگونی‌های ناگهانی و اغلب بازگشت‌ناپذیر، مانند بیابان‌زایی یا فروپاشی اکوسیستم، شوند (Scheffer et al. 2001). از این رو، توسعه ابزارهای تحلیلی پیشرفته برای درک دینامیک و پیش‌بینی آینده این سیستم‌ها برای مدیریت سازگار و توسعه پایدار امری ضروری است.

حوضه سیستان، یک حوضه آبخیز فرامرزی در مرز ایران و افغانستان، نمونه‌ای برجسته از یک SES تحت بحران است. این منطقه که در گذشته به واسطه تالاب‌های بین‌المللی هامون از تنوع زیستی و پویایی اقتصادی بالایی برخوردار بود، در دهه‌های اخیر به دلیل خشکسالی‌های مداوم، تغییرات اقلیمی و به ویژه مدیریت منابع آب در بالادست رودخانه هیرمند، با خشکیدگی گسترده تالاب‌ها روبرو شده است. این فروپاشی اکولوژیکی، منجر به شکل‌گیری یکی از بزرگترین کانون‌های گرد و غبار در جهان شده (Yousefi-Kebriya et al. 2025) که پیامدهای شدیدی بر سلامت عمومی، کشاورزی و پایداری اجتماعی-اقتصادی منطقه داشته است. در حالی که بخش قابل توجهی از بحران گرد و غبار ایران، منشأ فرامرزی دارد (Alaiejad et al. 2025; Mojgan et al. 2022). این چالش‌ها، نیاز فوری به مدل‌های پیش‌بینی‌کننده‌ای را نشان می‌دهند که بتوانند به سیاست‌گذاران در ارزیابی مداخلات مدیریتی و افزایش تاب‌آوری سیستم کمک کنند (Karimi et al. 2025; Zhou et al. 2025).

با این حال، مدل‌سازی SESها با محدودیت‌های روش‌شناختی قابل توجهی همراه است. مطالعات پیشین

عمدتاً بر روی تحلیل‌های آماری و گذشته‌نگر برای نشان دادن همبستگی بین متغیرها متمرکز بوده‌اند (به عنوان مثال، (Yousefi-Kebriya et al. 2025) اگرچه این رویکردها در شناسایی عوامل موثر بسیار ارزشمند هستند، اما توانایی محدودی در پیش‌بینی دینامیک‌های پیچیده و غیرخطی آینده سیستم و ارزیابی کمی پیامدهای سناریوهای مدیریتی مختلف دارند. مدل‌های سنتی مبتنی بر فرآیند نیز اغلب در مواجهه با پیچیدگی و ابعاد بالای این سیستم‌ها دچار مشکل می‌شوند (Aleahmad et al. 2025). از سوی دیگر، مدل‌های هوش مصنوعی و یادگیری عمیق، علی‌رغم قدرت پیش‌بینی بالا، معمولاً به عنوان مدل‌های "جعبه سیاه" عمل می‌کنند. این مدل‌ها ممکن است همبستگی‌های جعلی^۳ را بیاموزند که فاقد معنای فیزیکی هستند و این امر اعتبار آن‌ها را برای تصمیم‌گیری‌های حساس کاهش می‌دهد (Reichstein et al. 2019). برای پر کردن این شکاف روش‌شناختی، پژوهش حاضر یک چارچوب مبتنی بر یادگیری عمیق را معرفی می‌کند که از تحلیل همبستگی فراتر رفته و یک "دوقلوی دیجیتال" پیش‌بینی‌کننده برای سیستم سیستان ارائه می‌دهد. این چارچوب بر پایه شبکه‌های عصبی گراف آگاه از فرآیند^۴ استوار است. در این رویکرد "جعبه خاکستری، دانش علمی موجود در مورد روابط علی بین اجزای مختلف سیستم سیستان به صورت یک گراف علی در معماری یک شبکه عصبی گراف (GNN) تعبیه می‌شود. این کار به مدل اجازه می‌دهد تا دینامیک‌های پیچیده سیستم را با تکیه بر داده‌ها و با هدایت دانش فرآیندی بیاموزد و از این طریق، شفافیت و سازگاری فیزیکی آن افزایش یابد (Bhasme et al. 2022; Mu et al. 2025). علاوه بر این، برای غلبه بر چالش کمبود داده، که یک مانع بزرگ در مدل‌سازی محیطی است، از یک تکنیک پیشرفته تلفیق داده مبتنی بر یادگیری ماشین برای بازسازی یک سری زمانی کامل و ۲۱ ساله از شاخص گرد و غبار استفاده شده است. نوآوری اصلی این پژوهش، ارائه یک ابزار پشتیبان تصمیم کمی است که قادر به شبیه‌سازی و ارزیابی پیامدهای سناریوهای مدیریتی و اقلیمی آینده بر سلامت کل اکوسیستم می‌باشد.

³ Spurious Correlations⁴ Process-Informed GNN¹ Socio-Ecological Systems² Tipping Points

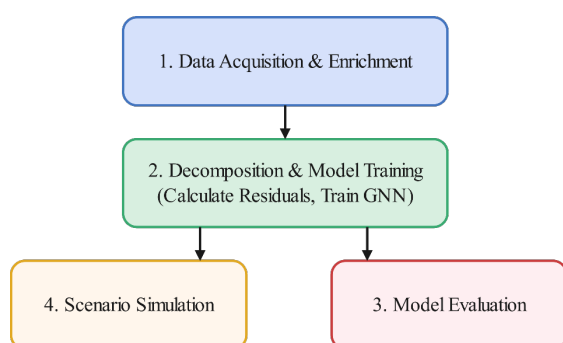


Fig. 1 Methodological workflow of the research.

شکل ۱ گردش کار روش‌شناختی پژوهش.

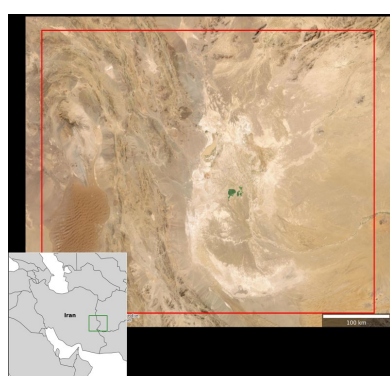


Fig. 2 Map of the Sistan Basin study area.

شکل ۲ نقشه منطقه مورد مطالعه حوضه سیستان.

برای تعبیه دانش علمی موجود در مدل، یک گراف علی اولیه (شکل ۳) بر اساس مرور جامع پژوهش‌ها در زمینه هیدرولوژی و علوم جوی مناطق خشک تدوین شد. این گراف، فرضیات ما در مورد قوی‌ترین روابط مستقیم در سیستم را نشان می‌دهد. یال‌ها در این گراف جهت‌دار و بدون وزن هستند و صرفاً وجود یک مسیر بالقوه برای جریان اطلاعات در معماری شبکه عصبی گراف (GNN) را مشخص می‌کنند؛ وزن واقعی و میزان تاثیرگذاری این روابط در طول فرآیند آموزش مدل، توسط خود شبکه عصبی آموخته خواهد شد. در این گراف، گره‌ها همان شش متغیر سیستم هستند و یال‌ها نشان‌دهنده روابط تاثیرگذاری اولیه هستند. برای مثال، یک یال از گره "بارش (pr)" به گره "وسعت آب" نشان‌دهنده تاثیر مستقیم بارش بر میزان آب موجود در تالاب است. به طور مشابه، یال‌های خروجی از گره‌های پوشش گیاهی و به سمت گره "شاخص آئروسول" یا نگر این

در ادامه این مقاله، ابتدا مواد و روش‌های مورد استفاده شامل منطقه مورد مطالعه، فرآیند جمع‌آوری و بازسازی داده‌ها، و معماری مدل پیشنهادی تشریح می‌شود (بخش ۲). سپس، نتایج عملکرد مدل و یافته‌های حاصل از مقایسه با مدل‌های پایه، تحلیل حساسیت و شبیه‌سازی سناریوها ارائه می‌گردد (بخش ۳). در نهایت، مقاله با یک نتیجه‌گیری در مورد اهمیت یافته‌ها و پیشنهاد برای کارهای آتی به پایان می‌رسد (بخش ۴).

۲- مواد و روش‌ها

چارچوب روش‌شناختی این پژوهش از سه مرحله اصلی و یکپارچه تشکیل شده است: (۱) جمع‌آوری داده‌های ماهواره‌ای چندمنظوره و بازسازی سری‌های زمانی کامل با استفاده از روش‌های یادگیری ماشین، (۲) طراحی، آموزش و اعتبارسنجی یک مدل یادگیری عمیق مبتنی بر گراف و آگاه از فرآیند، و (۳) اجرای تحلیل‌های پیشرفته پس از مدل‌سازی شامل ارزیابی عملکرد، تحلیل حساسیت سیستم و شبیه‌سازی سناریوهای مدیریتی و اقلیمی. فلوچارت کلی این فرآیند در شکل ۱ نمایش داده شده است.

۲-۱- منطقه مورد مطالعه و ساختار سیستم

منطقه مورد مطالعه، حوضه سیستان، یک حوضه آبخیز فرامرزی به وسعت تقریبی ۳۹۰۰۰ کیلومتر مربع در مرز ایران و افغانستان است (شکل ۲).

این منطقه شامل تالاب‌های بین‌المللی هامون است که در دهه‌های اخیر به دلیل خشکسالی و مدیریت منابع آب بالادست به یکی از کانون‌های اصلی گرد و غبار در شرق ایران تبدیل شده است. دینامیک این سیستم اجتماعی-اکولوژیکی از طریق شش متغیر کلیدی هیدرولوژی که به عنوان گره‌های یک سیستم پویا در نظر گرفته شده‌اند، مدل‌سازی می‌شود: (۱) بارش (Pr)، (۲) پتانسیل تبخیر و تعرق (PET)، (۳) وسعت آب^۱ تالاب هامون، (۴) شاخص گیاهی تالاب^۲، (۵) شاخص گیاهی کل حوضه^۳، و (۶) شاخص آئروسول^۴ به عنوان نماینده گرد و غبار.

³ Basin_NDVI

⁴ Aerosol_Index

¹ Water_Presence_Score

² Wetland_NDVI

بین عمق نوری آئروسول^۳ از سنجنده MODIS به عنوان ورودی اصلی از سال ۲۰۰۱ و شاخص آئروسول (AI) از Sentinel-5P (به عنوان خروجی هدف) را در دوره همپوشانی داده‌ها (۲۰۱۹-۲۰۲۱) بیاموزد. برای افزایش دقت، سایر متغیرهای محیطی نیز به عنوان ورودی‌های کمکی به مدل اضافه شدند. مدل آموزش داده شده با ضریب تعیین (R^2) برابر با ۰.۶۱ بر روی داده‌های تست داخلی، عملکرد قابل قبولی از خود نشان داد. از آنجا که اعتبارسنجی کامل با داده‌های زمینی به دلیل نبود ایستگاه‌های AERONET در منطقه امکان پذیر نبود، این سطح از دقت برای ادامه تحلیل‌ها کافی ارزیابی شد. این مدل سپس برای بازسازی مقادیر شاخص آئروسول در دوره ۲۰۱۸-۲۰۰۱ به کار گرفته شد. پس از تلفیق داده‌ها، یک مرحله مهندسی ویژگی برای غنی‌سازی دیتاست انجام شد. برای هر یک از شش متغیر اصلی، مقادیر تاخیری ۱، ۲، ۳ و ۶ ماه قبل به عنوان ویژگی‌های جدید اضافه شدند. علاوه بر این، برای کمک به مدل در یادگیری الگوهای فصلی، دو ویژگی مبتنی بر زمان (سینوس و کسینوس ماه سال) نیز به مجموعه داده افزوده شد. این فرآیند منجر به ایجاد یک دیتاست جامع و غنی شده با ۳۲ ویژگی و ۲۴۶ مشاهده ماهانه (پس از حذف رکوردهای دارای NaN ناشی از ایجاد lag شد که مبنای تمام تحلیل‌های بعدی قرار گرفت.

۳-۲- معماری و آموزش مدل

برای مدل‌سازی دینامیک‌های پیچیده سیستم، یک معماری نوآورانه مبتنی بر تجزیه^۴ و شبکه‌های عصبی گراف توسعه داده شد. این رویکرد اذعان دارد که بخش قابل توجهی از سیگنال متغیرهای محیطی (به ویژه آئروسول) دارای خودهمبستگی بالایی است. بنابراین، به جای پیش‌بینی مستقیم مقادیر، مدل برای پیش‌بینی باقیمانده‌ها (residuals) آموزش داده شد. باقیمانده (R) برای هر متغیر (V) در زمان t ، طبق رابطه (۱) محاسبه می‌شود:

$$R(t) = V_{actual}(t) - V_{actual}(t-1) \quad (1)$$

فرضیه است که افزایش پوشش گیاهی منجر به تثبیت خاک و کاهش انتشار گرد و غبار محلی می‌شود. این گراف، دانش فرآیندی را به عنوان یک قید ساختاری به مدل تزریق کرده و به یادگیری روابط معنادار کمک می‌کند.

۲-۲- جمع‌آوری و تلفیق داده‌ها

داده‌های سری زمانی ماهانه برای تمام متغیرها در بازه زمانی ۲۰۰۱ تا ۲۰۲۱ با استفاده از پلتفرم Google Earth Engine (GEE) استخراج شدند. برای اطمینان از شفافیت و تکرارپذیری، جزئیات کامل منابع داده، باندهای مورد استفاده، رزولوشن مکانی و زمانی اولیه و روش تجمع ماهانه^۱ در جدول ۱ خلاصه شده است. یک چالش کلیدی، وجود گپ زمانی در داده‌های باکیفیت شاخص آئروسول (AI) بود که توسط سنجنده Sentinel-5P از اواسط سال ۲۰۱۸ به بعد در دسترس است. برای حل این مشکل داده‌های ناقص از نوع MNAR^۲، یک روش تلفیق داده

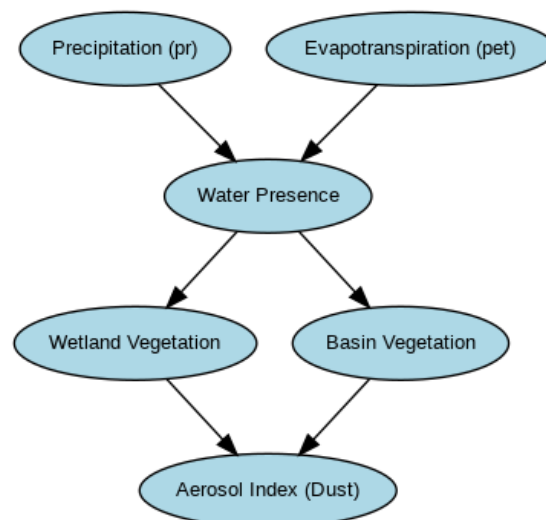


Fig. 3 The predefined causal graph for the system.

شکل ۳ گراف علی تعریف شده برای سیستم.

مبتنی بر یادگیری ماشین پیاده‌سازی شد.

در این روش، یک مدل Random Forest Regressor با ۱۰۰ درخت تصمیم، آموزش داده شد تا رابطه غیرخطی

³ Aerosol Optical Depth - AOD

⁴ Decomposition-based

¹ Aggregation method

² Missing Not at Random

جدول ۱ خلاصه منابع داده و فرآیند پیش‌پردازش

Table 1 Summary of data sources and preprocessing steps.

Variable	Data Source	ID GEE	Selected Band	Original Resolution	Monthly Aggregation Method
pr	TerraClimate	IDAHO_EPSCOR/TERRACLIMATE	'pr'	~۴ km	Sum
pet	TerraClimate	IDAHO_EPSCOR/TERRACLIMATE	'pet'	~۴ km	Sum
Water Extent	JRC GSW	JRC/GSW1_4/YearlyHistory	'waterClass'	۳۰ m	Mean
Wetland NDVI	MODIS	MODIS/061/MOD13A2	'NDVI'	۱ km	Mean
AOD	MODIS	MODIS/061/MCD19A2_GRANULES	'Optical_Depth_055'	۱ km	Mean
AI	Sentinel-5P	COPERNICUS/S5P/OFFL/L3_AER_AI	'absorbing_aerosol_index'	~۷ x ۳/۵ km	Mean

اختصاصی همان گره (شامل مقادیر تاخیری و ویژگی‌های زمانی) را پردازش می‌کند تا یک نمایش پنهان غنی و ویژه برای همان متغیر را استخراج نماید.

۲. لایه همجوشی گراف^۱: خروجی‌های تمام لایه‌های GRU (که هر کدام وضعیت یک گره را نشان می‌دهند) به یک لایه شبکه کانولوشن گراف (GCN) وارد می‌شوند. این لایه، با استفاده از ساختار گراف علی از پیش تعریف شده، اطلاعات را بین گره‌های همسایه منتشر کرده و به مدل اجازه می‌دهد تا برهم‌کنش‌های کل سیستم و تأثیرات متقابل بین متغیرها را بیاموزد. یک لایه Dropout نیز برای جلوگیری از بیش‌برازش پس از لایه GCN اعمال می‌شود.

پیش‌بینی نهایی مدل ($V_{predicted}(t)$) از ترکیب پیش‌بینی پایه (که همان $V_{actual}(t-1)$ است) و پیش‌بینی باقیمانده توسط GNN به دست می‌آید. این روش به مدل اجازه می‌دهد تا بر روی یادگیری بخش‌های پیچیده و غیرخطی سیگنال که توسط مدل ساده قابل درک نیستند، تمرکز کند.

معماری مدل پیشنهادی (شکل ۴) از اجزای زیر تشکیل شده است:

۱. لایه‌های ورودی و پردازش زمانی مجزا: برخلاف معماری‌های رایج، برای هر یک از شش گره سیستم (متغیر)، یک واحد بازگشتی درجه‌دار (GRU) مستقل در نظر گرفته شده است. هر GRU، سری زمانی ۱۲ ماه گذشته ویژگی‌های

¹ Graph Fusion Layer

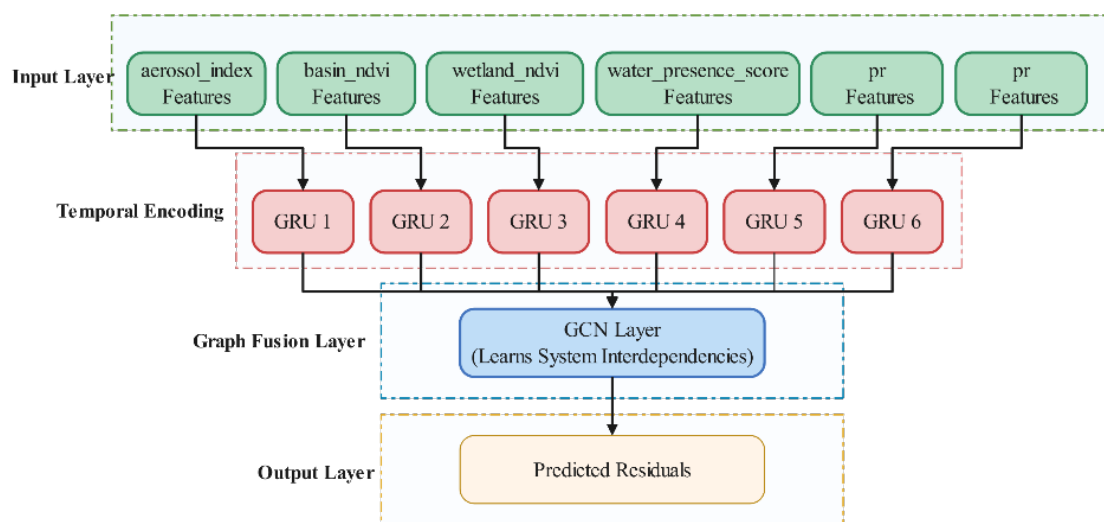


Fig. 4 The architecture of the proposed node-specific Decomposition-GNN model

شکل ۴ معماری مدل پیشنهادی Decomposition-GNN با پردازشگرهای مجزا برای هر گره.

۳- نتایج و بحث

در این بخش، نتایج کلیدی حاصل از چارچوب مدل‌سازی، ارائه و به تفصیل مورد بحث و تفسیر قرار می‌گیرد. این تحلیل‌ها بر سه حوزه اصلی متمرکز هستند: (۱) اعتبارسنجی جامع عملکرد مدل پیش‌بینی‌کننده و مقایسه آن با مدل‌های پایه، (۲) شبیه‌سازی سناریوهای مدیریتی آینده و تحلیل عدم قطعیت آن‌ها، و (۳) تفسیر پیامدهای سیاستی یافته‌ها.

۳-۱- اعتبارسنجی و ارزیابی مقایسه‌ای عملکرد مدل

به منظور ارزیابی دقیق و همه‌جانبه مدل پیشنهادی (Decomposition-GNN)، عملکرد آن بر روی مجموعه داده تست (دوره زمانی ژانویه ۲۰۱۹ تا دسامبر ۲۰۲۱) که در فرآیند آموزش مشاهده نشده بود، با سه مدل پایه قدرتمند مقایسه شد: یک مدل آماری کلاسیک (SARIMA)، یک مدل یادگیری عمیق استاندارد در حوزه سری‌های زمانی (LSTM) و یک مدل ساده تداوم (Persistence). ارزیابی با استفاده از مجموعه‌ای از معیارهای استاندارد انجام شد تا پاسخی جامع به نیازهای تحلیلی ارائه دهد.

اجازه می‌دهد تا برهم‌کنش‌های کل سیستم و تاثیرات متقابل بین متغیرها را بیاموزد. یک لایه Dropout نیز برای جلوگیری از بیش‌برازش پس از لایه GCN اعمال می‌شود. ۳. لایه خروجی: یک لایه خطی در نهایت، نمایش غنی و یکپارچه شده از لایه GCN را به پیش‌بینی باقیمانده هر شش متغیر در ماه آینده تبدیل می‌کند.

کل دیتاست به سه بخش آموزش (۷۰٪)، اعتبارسنجی (۱۵٪) و تست (۱۵٪) به صورت زمانی تقسیم شد. برای یافتن هایپرپارامترهای بهینه، یک جستجوی جامع شبکه‌ای^۱ انجام شد. بهترین مدل با استفاده از بهینه‌ساز Adam، تابع زیان میانگین مربعات (MSE) و مکانیزم توقف زودهنگام (با صبر ۵۰ دوره) بر روی خطای اعتبارسنجی، آموزش داده شد. هایپرپارامترهای بهینه نهایی شامل: ابعاد پنهان GRU برابر با ۳۲، ابعاد خروجی GCN برابر با ۳۲، نرخ یادگیری ۰/۰۰۱، نرخ Dropout برابر با ۰/۳، و L2 regularization (weight_decay) برابر با ۰/۰۰۰۱ بودند.

^۱ Grid Search

مدل در دنبال کردن روندهای فصلی و بلندمدت داده‌های واقعی را تایید می‌کند. علاوه بر این، تحلیل باقیمانده‌ها که در شکل ۶ ارائه شده، تایید می‌کند که اریب سیستماتیک قابل توجهی در پیش‌بینی‌های مدل وجود ندارد؛ زیرا خطاها به صورت تصادفی در اطراف خط صفر پراکنده شده‌اند.

۲-۳- شبیه‌سازی سناریوهای مدیریتی و پیامدهای سیاستی

نوآوری اصلی این پژوهش، فراتر رفتن از پیش‌بینی صرف و ارائه ابزاری برای "شبیه‌سازی آینده" است. در حالی که مطالعات پیشین در ایران (مانند (Yousefi-Kebriya et al. 2025)) همبستگی قوی بین کاهش سطح تالاب و روزهای گرد و غباری را در گذشته نشان داده‌اند، مدل ما قادر است دینامیک‌های آینده این رابطه را تحت سناریوهای مختلف شبیه‌سازی کند.

جدول ۲ مقایسه عملکرد مدل GNN پیشنهادی با مدل‌های پایه بر روی داده تست.

Table 2 Comparative performance evaluation of the proposed GNN model against baseline models on the test set

Model	R ² (Water Presence)	R ² (Aerosol Index)
Suggested Model	0.708	0.810
LSTM	0.667	0.802
Persistence	0.046	0.713
SARIMA	0.035	0.015

نتایج کمی این ارزیابی مقایسه‌ای در جدول ۲ خلاصه شده است. این نتایج چندین یافته کلیدی را آشکار می‌سازند. اولاً، مدل آماری SARIMA عملکرد بسیار ضعیفی در پیش‌بینی هر دو متغیر کلیدی از خود نشان داد که بیانگر ماهیت شدیداً غیرخطی و پیچیده دینامیک‌های سیستم سیستم است. ثانیاً، مدل ساده Persistence عملکرد قابل توجهی در پیش‌بینی شاخص آئروسول ($R^2=0.713$) داشت که نشان‌دهنده خودهمبستگی زمانی بسیار بالای این پدیده است. ثالثاً و از همه مهم‌تر، مدل پیشنهادی ما در هر دو متغیر کلیدی از مدل‌های کلاسیک و پایه بهتر عمل کرد. به طور مشخص، مدل GNN به ضریب تعیین (R^2) برابر با ۰.۷۰۸ برای وساب و ۰.۸۱۰ برای شاخص آئروسول دست یافت. این موفقیت، به ویژه در پیشی گرفتن از معیار قدرتمند Persistence برای شاخص آئروسول، نشان‌دهنده ارزش افزوده رویکرد ماست. با تجزیه سیگنال و تمرکز بر روی پیش‌بینی باقیمانده‌ها، و همچنین با تعبیه ساختار علی سیستم در معماری گراف، مدل ما توانسته است الگوهای پیچیده‌ای را فراتر از خودهمبستگی ساده یاد بگیرد.

جدول ۳، ارزیابی جامع عملکرد مدل نهایی را برای تمام متغیرهای سیستم با استفاده از معیارهای مختلف (MAE, RMSE, NSE, sMAPE) نشان می‌دهد. همانطور که مشاهده می‌شود، مدل در پیش‌بینی وضعیت پوشش گیاهی حوضه ($R^2=0.767$) و تبخیر و تعرق ($R^2=0.943$) نیز بسیار موفق بوده است که این امر قابلیت آن را به عنوان یک مدل سیستم جامع تایید می‌کند.

نمودارهای مقایسه‌ای در شکل ۵ به صورت بصری توانایی

جدول ۳ ارزیابی جامع عملکرد مدل نهایی بر روی مجموعه داده تست برای تمام متغیرها.

Table 3 Comprehensive performance evaluation of the final Decomposition-GNN model on the test set for all system variables

Parameter	R ²	MAE	RMSE	NSE	sMAPE (%)
pr	0.443	8.35	11.54	0.443	116.98
pet	0.943	1.51e+02	1.99e+02	0.943	14.80
water_presence_score	0.708	1.31e+06	1.82e+06	0.708	42.37
wetland_ndvi	0.049	0.008	0.010	0.049	8.89
basin_ndvi	0.767	0.007	0.009	0.767	6.84
aerosol_index	0.810	0.161	0.264	0.810	41.00

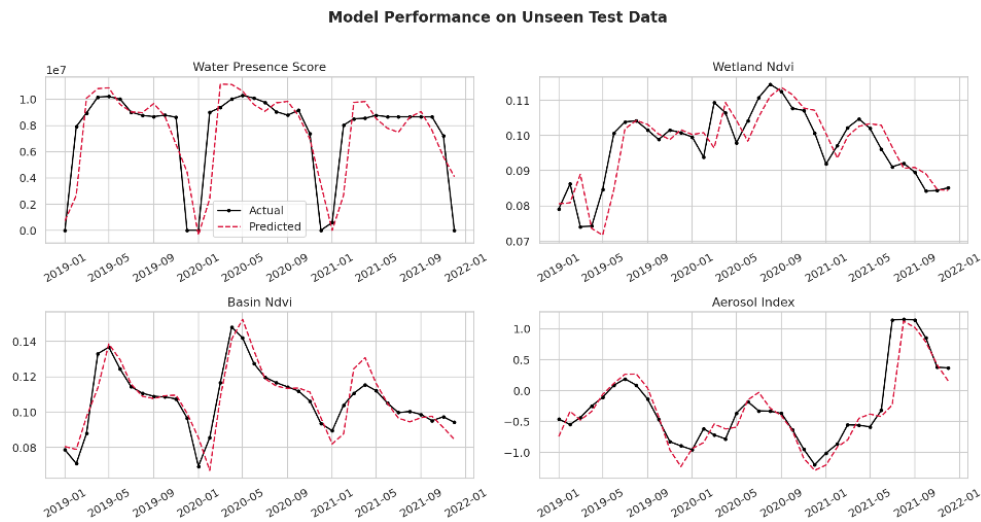


Fig. 5 Comparison of model performance on the test dataset for four key variables.

شکل ۵ مقایسه عملکرد مدل بر روی مجموعه داده تست برای چهار متغیر کلیدی.

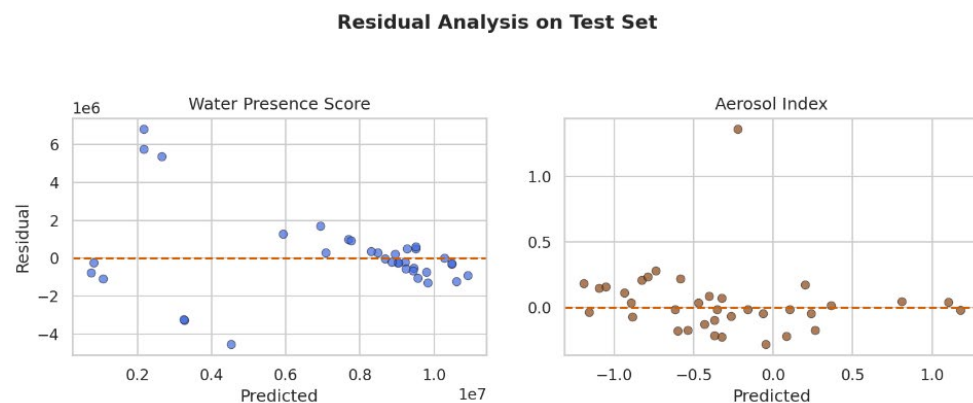


Fig. 6 Residual analysis for the two key variables on the test dataset.

شکل ۶ مقایسه عملکرد مدل بر روی مجموعه داده تست برای تحلیل باقیمانده‌ها برای دو متغیر کلیدی بر روی مجموعه داده تست.

گیاهی تالاب گردید. مهم‌ترین یافته این بخش، کاهش ۱۶/۰۸٪ در میانگین شاخص آئروسول (AI) نسبت به سناریوی پایه است.

برای تفسیر این یافته فنی با یک معیار ملموس برای سلامت عمومی، از رابطه کالیبره شده بین شاخص آئروسول ماهواره Sentinel-5P و غلظت زمینی ذرات معلق (Hassaan et al., 2023) استفاده شد. بر این اساس، سناریوی احیای تالاب منجر به کاهش تخمینی حدود ۱۶۶ میکروگرم بر متر مکعب ($\mu\text{g}/\text{m}^3$) در میانگین غلظت ذرات $\text{PM}_{2.5}$ در منطقه می‌شود. این نتیجه با یافته‌های

برای اطمینان از پایداری نتایج در افق زمانی ۵ ساله (۶۰ ماهه)، فرآیند شبیه‌سازی با استفاده از یک استراتژی ترکیبی و با اعمال عدم قطعیت (Monte Carlo Dropout) اجرا شد. دو سناریو مقایسه شدند: (۱) سناریوی پایه (ادامه روند فعلی)، و (۲) سناریوی احیای تالاب (حفظ سطح آب بالاتر از چارک سوم تاریخی).

نتایج شبیه‌سازی (جدول ۴ و شکل ۷) تاثیر چشمگیر مدیریت آب بر کیفیت هوا را آشکار می‌سازد. در سناریوی احیا، مداخله برای حفظ آب (که منجر به افزایش متوسط ۴/۰۵٪ در وسعت آب شد)، باعث بهبود ۳/۸۹٪ در پوشش

جدول ۴ مقایسه کمی نتایج میانگین ۵ ساله برای سناریوی پایه و سناریوی احیای تالاب.

Table 4 Quantitative comparison of the 5-year average outcomes for the Baseline and Wetland Restoration scenarios.

Variable (5 Year Ave.)	Baseline Scenario	Wetland scenario	Percentage Change
pr (mm)	10.84	12.41	+14.51%
pet (mm)	1397.78	1417.29	+1.40%
water_presence_score (km ²)	7.66e+06	7.97e+06	+4.05%
wetland_ndvi	0.09	0.09	+3.89%
basin_ndvi	0.10	0.11	+3.37%
aerosol_index	-0.21	-0.25	-16.08%

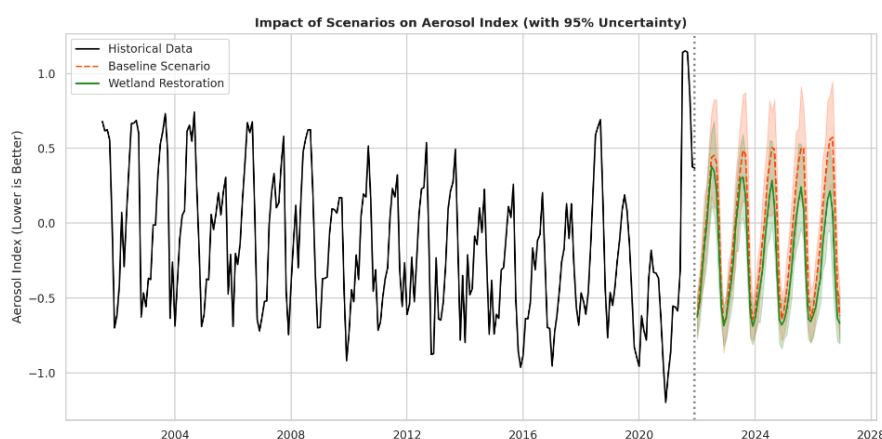


Fig. 7 Simulation of management scenarios' impact on the Aerosol Index over a 5-year horizon, including 95% uncertainty intervals derived from Monte Carlo Dropout.

شکل ۷ شبیه‌سازی تاثیر سناریوهای مدیریتی بر شاخص آئروسول در یک افق ۵ ساله، به همراه دامنه اطمینان ۹۵٪ استخراج شده از طریق Monte Carlo Dropout.

ارزیابی مدل بر روی داده‌های تست، عملکرد برتر آن را با ضریب تعیین (R^2) برابر با ۰/۷۰۸ برای وسعت آب و ۰/۸۱۰ برای شاخص آئروسول به اثبات رساند. این مدل ضمن پیشی گرفتن از رقبای قدرتمندی نظیر Persistence، SARIMA و LSTM، ارزش افزوده ترکیب دانش علی با پردازش زمانی مجزا در ساختار گراف را تایید کرد. شبیه‌سازی سناریوها نشان داد که مداخله فعال برای احیای تالاب (افزایش ۴٪ در وسعت آب)، منجر به بهبود ۱۶٪ در شاخص آئروسول و کاهش تخمینی بیش از ۱۶۰ میکروگرم بر متر مکعب در غلظت ۲/۵ PPM می‌شود که تاثیری مستقیم بر سلامت عمومی دارد. این ابزار شفاف به سیاست‌گذاران امکان می‌دهد پیامد مداخلات مدیریتی را پیش از اجرا ارزیابی کرده و مسیرهای پایدارتری برای افزایش تاب‌آوری مناطق خشک انتخاب کنند. مطالعات آتی می‌تواند بر اعتبارسنجی میدانی و مدل‌سازی سیاست‌های مدیریت آب بالادست

(Yousefi-Kebriya et al. 2025) که نشان دادند هر کیلومتر مربع کاهش تالاب منجر به افزایش قابل توجه روزهای آلوده می‌شود، همخوانی دارد، با این تفاوت که مدل ما توانسته است این اثر را برای یک برنامه مدیریتی مشخص در آینده کمی‌سازی کند.

۴- نتیجه‌گیری

ا بر یادگیری عمیق و شبکه‌های عصبی گراف (GNN) را برای پیش‌بینی دینامیک سیستم‌های اجتماعی-اکولوژیکی در حوضه سیستان ارائه داد. برخلاف مدل‌های آماری گذشته‌نگر، این معماری قادر به شبیه‌سازی دقیق سناریوهای آینده در شرایط کمبود داده است. با بهره‌گیری از رویکرد Decomposition-GNN و تمرکز بر پیش‌بینی باقیمانده‌های سیگنال، مدل پیشنهادی توانست الگوهای غیرخطی پیچیده سیستم را با دقت بالا استخراج کند.

متمرکز شود.

5- List of Symbols

SES	Socio-Ecological System
GNN	Graph Neural Network
GEE	Google Earth Engine
MNAR	Missing Not At Random
AI	Aerosol Index
AOD	Aerosol Optical Depth
NDVI	Normalized Difference Vegetation Index
GRU	Gated Recurrent Unit
GCN	Graph Convolutional Network
MSE	Mean Squared Error
R ²	Coefficient of Determination
MAE	Mean Absolute Error
RMSE	Root Mean Squared Error
NSE	Nash-Sutcliffe Efficiency
LSTM	Long Short-Term Memory
SARIMA	Seasonal Auto-Regressive Integrated Moving Average

Acknowledgments

The authors are grateful to anonymous reviewers whose detailed comments substantially improved the clarity and scope of this manuscript.

Credit Authorship Contribution Statement

Conceptualization, B.m and M.S.H.; Methodology, B.m and M.S.H.; Software, B.M.; data analysis, M.S.H. and S.B; Writing-Original Draft Preparation, B.M and S.B.; Writing-Review and Editing, M.S.H. and S.B; Supervision, S.B.

All authors have read and agreed to the published version of the manuscript.

Funding

The authors state that no financial support was received from any organization for this research.

Data Availability Statement

The data that supports the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare no conflict of interest.

References

Alaiejad M, Goudarzi G, Yengejeh RJ, Davami AH, Hassan EM (2025) An investigation of long-term dust storms in western Iran: an insight into emission sources using HYSPLIT model and MERRA-2 data (2012–2018) *Natural Hazards*

121:9519-9542

doi:

<https://doi.org/10.1007/s11069-025-07177-4>

Aleahmad M, Karbassi A, Davami AH, Jalilzadeh Yengejeh R (2025) Air Pollution Modeling for the Optimal Location of Air Quality Monitoring Stations in the Petrochemical Special Economic Zone in MAHSHAHR (Iran) *Environmental Modeling & Assessment* doi: <https://doi.org/10.1007/s10666-025-10063-9>

Amani M, Omid Najafabadi M, Mirdamadi SM, Torkashvand AM (2025) Dynamic analysis of the water, food, and energy nexus in minab county: An integrated approach to sustainable water resource management in hormozgan province *Results in Engineering* 27:106185 doi: <https://doi.org/10.1016/j.rineng.2025.106185>

Bhasme P, Vagadiya J, Bhatia U (2022) Enhancing predictive skills in physically-consistent way: Physics Informed Machine Learning for hydrological processes *Journal of Hydrology* 615:128618 doi: <https://doi.org/10.1016/j.jhydrol.2022.128618>

Hassaan MA, Abdallah SM, Shalaby E-SA, Ibrahim AA (2023) Assessing vulnerability of densely populated areas to air pollution using Sentinel-5P imageries: a case study of the Nile Delta, Egypt *Scientific Reports* 13:17406 doi: <https://doi.org/10.1038/s41598-023-44186-4>

Kanan AH, Giupponi C (2024) Coastal Socio-Ecological Systems Adapting to Climate Change: A Global Overview *Sustainability* 16:10000 doi: <https://doi.org/10.3390/su162210000>

Karimi BA, Haziq MA, Hayat A (2025) Specific impacts of climate change on the hydrological patterns and land use dynamics in the Arghandab River Basin, Kandahar, Afghanistan *Natural Hazards Research* 5:380-390 doi: <https://doi.org/10.1016/j.nhres.2024.12.007>

Mojgan A, Gholamreza G, Reza J, Amir Hossein D, Elham Mobarak H (2022) Study on Origin of Dust and its Relationship With Climatic Elements in Ilam Province *Archives of Hygiene Sciences* 11:232-239

Mu T, Duan F, Ning B, Zhou B, Liu J, Huang M (2025) ST-GPINN: a spatio-temporal graph physics-informed neural network for enhanced water quality prediction in water distribution systems *npj Clean Water* 8:74 doi: <https://doi.org/10.1038/s41545-025-00499-7>

Reichstein M, Camps-Valls G, Stevens B, Jung M, Denzler J, Carvalhais N, Prabhat (2019) Deep learning and process understanding for data-driven Earth system science *Nature* 566:195-204 doi: <https://doi.org/10.1038/s41586-019-0912-1>

- Scheffer M, Carpenter S, Foley JA, Folke C, Walker B (2001) Catastrophic shifts in ecosystems Nature 413:591-596 doi:<https://doi.org/10.1038/35098000>
- Yousefi-Kebriya A, Nadi M, Afaridegan E, Sun Z (2025) Wetland shrinking and dust pollution in Khuzestan Iran: insights from sentinel-5 and MODIS satellites Scientific Reports 15:13626 doi: <https://doi.org/10.1038/s41598-025-96935-2>
- Zhou E, Zhou B, Zhang L, Chen H, Guo J, Zhang K, Luo B (2025) The effect and burden of sand-dust storms on asthma hospitalization: Evidence from cities with arid climate in China Environmental Research 264:120345 <https://doi.org/10.1016/j.envres.2024.120345>