UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

Surrogate models, physics-informed neural networks and climate change

PhD Co-tutored Thesis with
Università degli Studi di Parma

Submitted by
Daniele Secci

Advisors:
Maria Giovanna Tanda
J. Jaime Gómez-Hernández
Valeria Todaro

April 2024

Surrogate models, physics-informed neural networks and climate change

PhD Thesis submitted by
Daniele Secci

Advisors:
Maria Giovanna Tanda
J. Jaime Gómez-Hernández
Valeria Todaro

Departments:
Ingeniería Hidráulica y Medio Ambiente
Universitat Politècnica de València

Ingegneria Civile e Architettura
Università degli Studi di Parma

April 2024, Parma, Italy

Acknowledgements

Foremost, heartfelt thanks go to my esteemed academic mentors, Prof. Tanda and Prof. Gómez-Hernández.

To Prof. Tanda, your mentorship has been a perfect blend of academic prowess and genuine care and your wealth of knowledge and sharp insights have been indispensable.

To Prof. Gómez-Hernández, your knack for making complex concepts accessible has been invaluable in shaping my understanding and your commitment to excellence has left a lasting impression on my academic journey.

A special mention goes to my co-tutor and friend, Dr. Todaro. Your enthusiasm for always tackling the most complex methodologies pushed me to give my best. Despite the occasional playful banter, your presence has played a crucial role in guiding my path.

My gratitude also extends to my academic dream team of all the professors and colleagues from Parma and Valencia who have been part of this journey. Your collaborative spirit has made the academic environment both enjoyable and motivating.

To my partner-in-crime, Sara, you've been the Batman to my Robin. Your infinite patience, unwavering love and understanding have saved me from countless "thesis-induced" meltdowns. And, of course, to my furball of joy, Zoe, you were the ultimate stress-buster, wagging your tail and providing comic relief.

Heartfelt appreciation to my Mom and Dad, cheering me on, and setting the bar pretty high as role models. To my Sister and Brother, you've been my strength, always believing in me. To my nephews, Gabriele and Vittoria, your energy and your smiles could light up the Parma sky on a usual foggy day. To Monica and Francesco, your kindness and positivity have added an extra layer of warmth to this journey.

A special mention goes to my grandparents, your enduring support and timeless wisdom have left an indelible mark on my character.

To my squad of friends, you are my lifeline. Your support turned bad moments into comedy clubs and padel matches.

Last but not least, gratitude to the family of my fiancé for adopting me into their clan. Your warmth and hospitality made me feel like part of the tribe.

I can't wait to celebrate this achievement with each one of you.

With sincere thanks,
Daniele

Abstract

This research contributes to the advancement of surrogate modelling as a powerful technique in the field of computational simulation that offers numerous advantages for solving complex problems efficiently. In particular, this study emphasizes the pivotal role of surrogate modeling in groundwater management. By integrating key factors like climate change and leveraging machine learning, particularly neural networks, the research facilitates more informed decision-making, significantly reducing the computational cost of complex numerical models.

The impact of climate change is a central focus and the first study aims to construct surrogate data-driven models for evaluating climate change effects on groundwater resources, also in the future. The study involves a comparison between statistical methods and different types of artificial neural networks (ANNs). The effectiveness of surrogate models was demonstrated in Northern Tuscany (Italy) but can easily extend to any area of interest. The adopted statistical method involves analyzing historical precipitation and temperature data along with groundwater levels recorded in monitoring wells. Initially, the study explores potential correlations between meteorological and groundwater indices; if a correlation is identified, a linear regression analysis is employed to establish relationships between them. These established relationships are then used to estimate future groundwater levels based on projected precipitation and temperature obtained from an ensemble of Regional Climate Models, under two Representative Concentration Pathways,

namely RCP4.5 and RCP8.5.

Then, three distinct Artificial Intelligence (AI) models, Nonlinear AutoRegressive with eXogenous inputs (NARX), Long-Short Term Memory (LSTM) and Convolutional Neural Network (CNN) were implemented to evaluate the impact of climate change on groundwater resources for the same case study. Specifically, these models were trained using directly historical precipitation and temperature data as input to provide groundwater levels as output. Following the training phase, the developed AI models were utilized to forecast future groundwater levels using the same precipitation and temperature projections and climate scenarios described above. The results highlighted different outputs among the models used in this work. However, most of them predict a decrease in groundwater levels as a result of future variations in precipitation and temperature. The study also presents the strengths and weaknesses of each model. Notably, the LSTM model emerges as the most promising approach to predict future groundwater levels.

Within the same field, an ANN was developed with the capability to simulate groundwater conditions in the Konya closed basin, Turkey, one of the pilot sites investigated as part of the InTheMED project. This model serves as a tool for examining the potential impacts of climate change and agricultural policies on groundwater resources within the region. The final goal of this application, is to provide a user-friendly tool, based on the trained neural network. The inherent simplicity of the surrogate model, with a straightforward interface and results that are simple to understand, plays a crucial role in decision-making processes.

Shifting to pollutant transport, an ANN was implemented to solve different direct and inverse problems. The direct problem deals with the evaluation of concentrations in monitoring wells, while the inverse problem involves the identification of contaminant sources and their release history. It demonstrated efficiency in addressing both direct and inverse transport problems, offering reliable results with reduced computational burden.

The study also addresses the interpretability challenge of ANNs and the so called "generalization problem" through Physics-Informed Neural Networks (PINNs). By incorporating physics-based constraints, PINNs bridge the gap between data-driven modeling and physics-based interpretations, offering a promising approach for groundwater numerical simulations. In this study, a PINN is developed to simulate flow in an unconfined aquifer.

Finally, two extra content are presented. First, an ANN is used to solve an inverse problem in the field of sewer systems. Then, an easily interpretable example of numerical groundwater flow modeling using spreadsheets, from a didactic perspective, is described.

In conclusion, this research underscores the importance of surrogate modeling, machine learning, climate change analysis, and physics-informed approaches in advancing groundwater management strategies and beyond, providing valuable tools for decision-makers to address complex groundwater flow problems in changing environmental conditions.

Sommario

Questa ricerca propone nuovi avanzamenti nella modellazione surrogata che, nella simulazione di problemi complessi, offre vantaggi rilevanti a supporto della modellazione numerica usuale. In particolare, questo studio sottolinea il ruolo cruciale della modellazione surrogata nella gestione delle risorse idriche sotterranee. Integrando fattori chiave come il cambiamento climatico e sfruttando l'apprendimento automatico, in particolare le reti neurali, lo studio concorre a rendere più facile il processo decisionale informato, riducendo significativamente il costo computazionale dei complessi modelli numerici.

L'impatto del cambiamento climatico è al centro dell'attenzione e il primo studio mira a costruire modelli surrogati del tipo "data-driven" per valutare gli effetti del cambiamento climatico sulle risorse idriche sotterranee nel futuro. Esso confronta un metodo statistico e diversi tipi di reti neurali artificiali (ANN) per migliorare la comprensione e facilitare le decisioni nella gestione delle acque sotterranee. L'efficacia dei modelli surrogati è stata dimostrata in una applicazione nella Toscana settentrionale, ma può facilmente estendersi a qualsiasi area di interesse. Il metodo statistico adottato coinvolge l'analisi di dati storici sulle precipitazioni e sulla temperatura insieme ai livelli freatici registrati nei pozzi di monitoraggio. Inizialmente, lo studio esplora correlazioni potenziali tra indici meteorologici e indici delle acque sotterranee. Se viene individuata una valida correlazione tra questi, si costruisce una regressione lineare che stabilisce una relazione tra di essi. Queste

relazioni vengono poi utilizzate per stimare futuri livelli di falda sulla base delle proiezioni di precipitazione e temperatura ottenute da un insieme di Modelli Climatici Regionali, considerando due Scenari di Emissione Rappresentativi, ovvero RCP4.5 e RCP8.5.

Successivamente, sono stati implementati tre distinti modelli di Intelligenza Artificiale (AI), Rete neurale Autoregressiva con Ingressi Eterogenei (NARX), Rete Neurale con Memoria a Lungo e Breve Termine (LSTM) e Rete Neurale Convoluzionale (CNN), per valutare l'impatto del cambiamento climatico sui livelli di falda per lo stesso caso di studio. In particolare, questi modelli sono stati addestrati utilizzando direttamente dati storici di precipitazioni e di temperatura come input e per fornire i livelli freatici come output. Dopo la fase di addestramento, i modelli di AI sviluppati sono stati utilizzati per prevedere i livelli delle acque sotterranee utilizzando le stesse proiezioni di precipitazioni e temperatura e gli scenari climatici descritti in precedenza. I risultati hanno evidenziato diversi output tra i modelli utilizzati in questo studio. Tuttavia, la maggior parte di essi prevede una diminuzione dei livelli di falda a seguito di future variazioni di precipitazione e temperatura. Lo studio presenta anche i punti di forza e debolezza di ciascun modello. In particolare, il modello LSTM emerge come l'approccio più promettente per prevedere i futuri livelli di falda.

Nello stesso campo, è stata sviluppata una rete neurale artificiale con la capacità di simulare lo stato dell'acquifero nel bacino di Konya, Turchia, uno dei siti pilota indagati nell'ambito del progetto InTheMED. Questo modello si comporta da strumento per esaminare gli impatti potenziali del cambiamento climatico e delle politiche agricole sulle risorse idriche sotterranee. L'obiettivo finale di questa applicazione è fornire uno strumento "user-friendly" basato sulla rete neurale addestrata. La semplicità intrinseca del modello surrogato, sviluppato con un'interfaccia chiara e risultati di facile comprensione, svolge un ruolo cruciale nei processi decisionali.

Passando al trasporto di inquinanti, è stata implementata una rete neurale artificiale per risolvere diversi problemi diretti e inversi. Il problema diretto riguarda la valutazione delle concentrazioni nei pozzi di monitoraggio, mentre il problema inverso comporta l'identificazione delle fonti di contaminazione e la loro storia di rilascio. La tecnica ha dimostrato efficienza nell'affrontare sia problemi diretti che inversi di trasporto, offrendo risultati affidabili con un ridotto onere computazionale.

Lo studio affronta anche la sfida dell'interpretabilità fisica delle reti neurali artificiali e del cosiddetto "problema della generalizzazione" attraverso le Reti Neurali Fisicamente Basate (PINN). Integrando vincoli basati sulla fisica, le PINN colmano il divario tra la modellazione basata sui dati e i modelli numerici costruiti sulle equazioni differenziali dedotte dalla fisica, offrendo un approccio promettente per le simulazioni numeriche delle acque sotterranee. In questo studio, una PINN è stata sviluppata per simulare il flusso in un acquifero non confinato.

Infine, vengono presentati due contenuti aggiuntivi. Nel primo, una rete neurale artificiale è utilizzata per risolvere un problema inverso nel campo dei sistemi fognari. In secondo, un esempio di modellazione numerica del flusso delle acque sotterranee mediante fogli di calcolo con una ottima prospettiva didattica.

In conclusione, questa ricerca sottolinea l'importanza della modellazione surrogata, dell'apprendimento automatico, dell'analisi del cambiamento climatico e degli approcci basati sulla fisica per progredire nelle strategie di gestione delle acque sotterranee e affrontare sfide complesse, offrendo strumenti preziosi ai decisori.

Resumen

Esta investigación contribuye al avance de la modelación sustitutiva como una técnica poderosa en el campo de la simulación computacional que ofrece numerosas ventajas para resolver eficientemente problemas complejos. En particular, este estudio destaca el papel crucial de la modelación sustitutiva en la gestión de aguas subterráneas. Integrando factores clave como el cambio climático y aprovechando el aprendizaje automático, especialmente las redes neuronales, la investigación facilita la toma de decisiones más informada, reduciendo significativamente el costo computacional de modelos numéricos complejos.

El impacto del cambio climático es un enfoque central, y el primer estudio tiene como objetivo construir modelos de datos sustitutivos para evaluar los efectos del cambio climático en los recursos de aguas subterráneas, también en el futuro. El estudio implica la comparación entre métodos estadísticos y diferentes tipos de Redes Neuronales Artificiales (ANN). La eficacia de los modelos sustitutivos se demostró en el norte de la Toscana (Italia), pero puede extenderse fácilmente a cualquier área de interés. El método estadístico adoptado implica analizar datos históricos de precipitación y temperatura junto con niveles de agua registrados en pozos de monitoreo. Inicialmente, el estudio explora posibles correlaciones entre índices meteorológicos e índices de agua subterránea; si se identifica una correlación, se emplea un análisis de regresión lineal para establecer relaciones entre ellos. Estas relaciones establecidas se utilizan luego para estimar los futuros niveles de

agua subterránea en función de las proyecciones de precipitación y temperatura obtenidas de un conjunto de Modelos Climáticos Regionales, bajo dos Trayectorias de Concentración Representativa, conocidas como RCP4.5 y RCP8.5.

Posteriormente, se implementaron tres modelos distintos de Inteligencia Artificial (AI), AutoRegressive No Lineal con Entradas Exógenas (NARX), Memoria a Largo y Corto Plazo (LSTM) y Red Neuronal Convolucional (CNN) para evaluar el impacto del cambio climático en los recursos de aguas subterráneas para el mismo caso de estudio. Específicamente, estos modelos fueron entrenados utilizando directamente datos históricos de precipitación y temperatura como entrada para proporcionar niveles de agua subterránea como salida. Después de la fase de entrenamiento, los modelos de IA desarrollados se utilizaron para prever los futuros niveles de agua subterránea utilizando las mismas proyecciones de precipitación y temperatura y escenarios climáticos descritos anteriormente. Los resultados resaltaron diferentes salidas entre los modelos utilizados en este trabajo. Sin embargo, la mayoría de ellos predice una disminución en los niveles de agua subterránea como resultado de futuras variaciones en la precipitación y temperatura. El estudio también presenta las fortalezas y debilidades de cada modelo. Notablemente, el modelo LSTM emerge como el enfoque más prometedor para predecir futuros niveles de agua subterránea.

Dentro del mismo campo, se desarrolló una ANN con la capacidad de simular las condiciones de agua subterránea en la cuenca cerrada de Konya, Turquía, uno de los sitios piloto investigados como parte del proyecto InTheMED. Este modelo sirve como herramienta para examinar los impactos potenciales del cambio climático y las políticas agrícolas en los recursos de agua subterránea dentro de la región. El objetivo final de esta aplicación es proporcionar una herramienta fácil de usar, basada en la red neuronal entrenada. La simplicidad inherente del modelo sustitutivo, con una interfaz directa y resultados fáciles de entender, juega un papel crucial en los procesos de toma de decisiones.

En cuanto al transporte de contaminantes, se implementó una ANN para resolver diferentes problemas directos e inversos. El problema directo trata sobre la evaluación de concentraciones en pozos de monitoreo, mientras que el problema inverso implica la identificación de fuentes de contaminantes y su historial de liberación. Demostró eficiencia al abordar problemas de transporte tanto directos como inversos, ofreciendo resultados confiables con una carga computacional reducida.

El estudio también aborda el desafío de la interpretabilidad de las ANNs y el llamado "problema de generalización" a través de las Redes Neuronales Informadas por la Física (PINNs). Al incorporar restricciones basadas en la física, las PINNs llenan la brecha entre la modelación basada en datos y las interpretaciones basadas en la física, ofreciendo un enfoque prometedor para las simulaciones numéricas de aguas subterráneas. En este estudio, se desarrolló una PINN para simular el flujo en un acuífero no confinado.

Finalmente, se presentan dos contenidos adicionales. Primero, se utiliza una ANN para resolver un problema inverso en el campo de los sistemas de alcantarillado. Luego, se describe un ejemplo fácilmente interpretable de modelado numérico del flujo de aguas subterráneas utilizando hojas de cálculo, desde una perspectiva didáctica.

En conclusión, esta investigación subraya la importancia de la modelación sustitutiva, el aprendizaje automático, el análisis del cambio climático y los enfoques informados por la física en el avance de las estrategias de gestión de aguas subterráneas y más allá, proporcionando herramientas valiosas para que los tomadores de decisiones aborden problemas complejos de flujo de aguas subterráneas en condiciones ambientales cambiantes.

Resum

Aquesta investigació contribueix al progrés de la modelació substitutiva com una tècnica poderosa en el camp de la simulació computacional que ofereix nombrosos avantatges per resoldre eficientment problemes complexos. En particular, aquest estudi destaca el paper crucial de la modelació substitutiva en la gestió d'aigües subterrànies. Integrant factors clau com el canvi climàtic i aprofitant l'aprenentatge automàtic, especialment les xarxes neuronals, la investigació facilita la presa de decisions més informada, reduint significativament el cost computacional de models numèrics complexes.

L'impacte del canvi climàtic és un enfocament central, i el primer estudi té com a objectiu construir models de dades substitutius per avaluar els efectes del canvi climàtic en els recursos d'aigües subterrànies, també en el futur. L'estudi implica la comparació entre mètodes estadístics i diferents tipus de Xarxes Neuronals Artificials (ANN). L'eficàcia dels models substitutius es va demostrar al nord de la Toscana (Itàlia), però pot estendre's fàcilment a qualsevol àrea d'interès. El mètode estadístic adoptat implica analitzar dades històriques de precipitació i temperatura juntament amb nivells d'aigua registrats en pous de monitoratge. Inicialment, l'estudi explora possibles correlacions entre índexs meteorològics i índexs d'aigua subterrània; si s'identifica una correlació, s'emplea una anàlisi de regressió lineal per establir relacions entre ells. Aquestes relacions establertes s'utilitzen després per estimar els futurs nivells d'aigua subterrània en funció de les projec-

cions de precipitació i temperatura obtingudes d'un conjunt de Models Climàtics Regionals, sota dues Trajectòries de Concentració Representativa, conegudes com RCP4.5 i RCP8.5.

Posteriorment, es van implementar tres models diferents d'Intel·ligència Artificial (IA), AutoRegressive No Lineal amb Entrades Exògenes (NARX), Memòria a Llarg i Curt Terminis (LSTM) i Xarxa Neuronal Convolucional (CNN) per avaluar l'impacte del canvi climàtic en els recursos d'aigües subterrànies pel mateix cas d'estudi. Específicament, aquests models van ser entrenats utilitzant directament dades històriques de precipitació i temperatura com a entrada per proporcionar nivells d'aigua subterrània com a sortida. Després de la fase d'entrenament, els models d'IA desenvolupats es van utilitzar per preveure els futurs nivells d'aigua subterrània utilitzant les mateixes projeccions de precipitació i temperatura i escenaris climàtics descrits anteriorment. Els resultats van destacar diferents sortides entre els models utilitzats en aquest treball. No obstant això, la majoria d'ells prediu una disminució en els nivells d'aigua subterrània com a resultat de futures variacions en la precipitació i temperatura. L'estudi també presenta les forces i les debilitats de cada model. Notablement, el model LSTM emergeix com l'enfocament més prometedor per predir futurs nivells d'aigua subterrània.

Dins del mateix camp, es va desenvolupar una ANN amb la capacitat de simular les condicions d'aigua subterrània a la conca tancada de Konya, Turquia, un dels llocs pilot investigats com a part del projecte InTheMED. Aquest model serveix com a eina per examinar els impactes potencials del canvi climàtic i les polítiques agrícoles en els recursos d'aigua subterrània dins de la regió. L'objectiu final d'aquesta aplicació és proporcionar una eina fàcil d'utilitzar, basada en la xarxa neuronal entrenada. La simplicitat inherent del model substitutiu, amb una interfície directa i resultats fàcils d'entendre, juga un paper crucial en els processos de presa de decisions.

Pel que fa al transport de contaminants, es va implementar una ANN per resoldre diferents problemes directes i inversos. El problema directe tracta sobre l'avaluació de concentracions en pous de monitoratge, mentre que el problema invers implica la identificació de fonts de contaminants i el seu historial de lliberació. Va demostrar eficiència en abordar problemes de transport tant directes com inversos, oferint resultats fiables amb una càrrega computacional reduïda.

L'estudi també aborda el repte de la interpretabilitat de les ANNs i la anomenada "problema de generalització" a través de les Xarxes Neuronals Informades per la Física (PINNs). Al incorporar restriccions basades en la física, les PINNs omplen la bretxa entre la modelació basada en dades i les interpretacions basades en la física, oferint un enfocament prometedor per a les simulacions numèriques d'aigües subterrànies. En aquest estudi, es va desenvolupar una PINN per simular el flux en un aqüífer no confinat.

Finalment, es presenten dos continguts addicionals. Primer, s'utilitza una ANN per resoldre un problema invers en el camp dels sistemes d'aigües residuals. Després, es descriu un exemple fàcilment interpretable de modelat numèric del flux d'aigües subterrànies utilitzant fulls de càlcul, des d'una perspectiva didàctica.

En conclusió, aquesta investigació subratlla la importància de la modelació substitutiva, l'aprenentatge automàtic, l'anàlisi del canvi climàtic i els enfocaments informats per la física en el progrés de les estratègies de gestió d'aigües subterrànies i més enllà, proporcionant eines valioses perquè els prescriptors abordin problemes complexos de flux d'aigües subterrànies en condicions ambientals canviant.

Contents

Introduction **1**

1. Surrogate modeling **5**

1.1. Introduction and state of the art 5
1.2. Statistical approach based on linear regression 9
1.3. Artificial neural networks 12
  1.3.1. Artificial neural network design and training process 14
  1.3.2. Multilayer perceptron networks 20
  1.3.3. Nonlinear autoregressive networks with exogenous inputs . 21
  1.3.4. Long-short term memory networks 24
  1.3.5. Convolutional neural networks 28
  1.3.6. Physics-informed neural networks 31

2. Climate change **37**

2.1. Introduction 37
2.2. Historical data processing 40
  2.2.1. Gap filling and interpolation procedures 40
  2.2.2. Meteorological and groundwater drought indices 41
2.3. Future climate projections 46
  2.3.1. Downscaling and bias correction 47

2.3.2. Future meteorological drought indices 49

3. Applications 51

3.1. Impact of climate change on groundwater resources 52

3.1.1. State of the art 53

3.1.2. Study area and available data 57

3.1.3. Future climate projections 63

3.1.4. Statistical approach 64

3.1.5. Comparison of three different artificial neural network techniques 80

3.1.6. Discussion and conclusions 98

3.2. An artificial neural network as a quick tool to assess the effects of climate change and agricultural policies on groundwater 101

3.2.1. Case study 102

3.2.2. Surrogate model: ANN 103

3.2.3. Conclusion and discussion 109

3.3. Artificial neural networks for solving forward and inverse transport problems 110

3.3.1. State of the art 111

3.3.2. Case study 112

3.3.3. Groundwater flow and transport 115

3.3.4. Set up of the ANN 117

3.3.5. Results 125

3.3.6. Discussion and conclusions 135

3.4. Physics-informed neural networks for solving transient unconfined groundwater flow 137

3.4.1. Transient unconfined flow 137

3.4.2. Case study 140

3.4.3. PINN set up 142
3.4.4. Evaluation of the performance 149
3.4.5. Results 150
3.4.6. Discussion and conclusions 161

4. Extra contents **165**
4.1. Identification of the inflow source in a foul sewer system using an artificial neural network as inverse model 166
4.1.1. Case study 167
4.1.2. Set up of the ANN 168
4.1.3. Results 171
4.1.4. Discussion and conclusions 171
4.2. Enhancing user-friendliness: a comprehensive and accessible example of numerical groundwater flow modeling using spreadsheets . . 173
4.2.1. Unconfined two-dimensional groundwater flow in a horizontal plane 174
4.2.2. Unconfined two-dimensional groundwater flow in a vertical cross-section 198
4.2.3. Conclusions and discussion 206

Conclusions **209**
Limitations and future directions 211

List of Figures

1.1. Left: experimental cumulative distribution function of the variable of interest and fitted cumulative distribution function that best approximates experimental data. Right: transformation into a standard normal distribution. 11
1.2. Artificial Neural Network sketch. 14
1.3. MLP sketch. 22
1.4. NARX closed-loop scheme. 24
1.5. LSTM cell. 28
1.6. CNN sketch. 31
1.7. PINN sketch. 35

2.1. Distribution Mapping method (Gupta et al., 2019). 49

3.1. Location of the study area with indication of the climate stations, monitoring wells and river basins. 59
3.2. Total annual precipitation (a) and annual average of the mean daily temperature (b) in terms of 10-year moving average observed and forecasted by the RCMs under the RCP4.5 and RCP8.5 scenarios. Average values over the entire study area. 64

3.3. SPIs (a) and SPEIs (b) for the four analyzed basins and time windows of 6, 9 and 12 months. 69
3.4. SGI values for the 15 monitoring wells used in this study. The white color indicates missing data, the grey color indicates positive values, while the color scale classifies the negative SGIs. 70
3.5. SGI-SPI Pearson correlation coefficients. 71
3.6. Heat maps of the SGI-SPI correlation coefficients (R) for different time windows and lags. The black box represents the highest correlation coefficient. 72
3.7. SGIs versus SPI6; the points represent the data, the red line indicates the regression line and the black line denotes the identity line. For each well, the correlation coefficient (R) and the regression equation is reported. 73
3.8. SGI-SPEI Pearson correlation coefficients. 74
3.9. SGIs versus SPEI9; the points represent the data, the red line indicates the regression line and the black line denotes the identity line. For each well, the correlation coefficient (R) and the regression equation is reported. 75
3.10. Cumulative probability distributions according to the whole RCM ensemble obtained for the Paganico monitoring well through the SGI-SPI6 (a) and the SGI-SPEI9 (b) regression equations for the historical period and at short- (ST), medium- (MT) and long-term (LT) under the RCP8.5 scenario. Envelope curves obtained by the 13 RCM models. 76

3.11. Box-plots of the SGIs obtained for the Paganico monitoring well. The SGIs were calculated according to the whole RCM, through the SGI-SPI6 and SGI-SPEI9 regression equations for the historical period and at short-term (ST), medium-term (MT), and long-term (LT) under the two RCP scenarios. The boxplot marks points as outliers if they are greater than the mean ± 2.7σ, where σ is the standard deviation. 77
3.12. Schematic view of the CNN (top) and sample of the standardized input data (bottom). 86
3.13. Observed and predicted groundwater levels for the testing phase (period 2019-2020) for the well Paganico. 89
3.14. Predicted groundwater levels with the NARX in April for the Paganico well in terms of 10-year moving average under the RCP4.5 and RCP8.5 scenarios. 90
3.15. Cumulative distribution probability functions according to the whole RCM ensemble obtained with the NARX for the Paganico well for the historical period and at short- (ST), medium- (MT), and long-term (LT) under the RCP4.5 (left) and RCP8.5 (right) scenarios, along with the envelope curves provided by the 13 RCMs. 91
3.16. Predicted groundwater levels with the LSTM in April for the Paganico well in terms of 10-year moving average under the RCP4.5 and RCP8.5 scenarios. 92
3.17. Cumulative distribution probability functions according to the whole RCM ensemble obtained with the LSTM for the Paganico well for the historical period and at short- (ST), medium- (MT), and long-term (LT) under the RCP4.5 (left) and RCP8.5 (right) scenarios, along with the envelope curves provided by the 13 RCMs. 93

3.18. Predicted groundwater levels with the CNN in April for the Paganico well in terms of 10-year moving average under the RCP4.5 and RCP8.5 scenarios. 93

3.19. Cumulative distribution probability functions according to the whole RCM ensemble obtained with the CNN for the Paganico well for the historical period and at short- (ST), medium- (MT), and long-term (LT) under the RCP4.5 (left) and RCP8.5 (right) scenarios, along with the envelope curves provided by the 13 RCMs. 94

3.20. Konya closed basin (Turkey) and the 30 control points. 103

3.21. Training, validation and test performances of the neural network. . 106

3.22. ANN tool interface developed using MATLAB. 107

3.23. ANN tool result. 108

3.24. Discretization grid of the two-dimensional aquifers. 116

3.25. Observed and estimated concentration at 7 monitoring wells for 5 years of simulation recorded one time per year, forward simulation with two release sources (FWD 1). 127

3.26. Observed and estimated release obtained as average of the results of 10 neural networks at known source, inverse simulation with one release source and different error level, error-free data ($\alpha = 0$) and corrupted data ($\alpha = 0.1$) (INV 1). 128

3.27. Observed and estimated release fluxes obtained as average of the results of 10 neural networks at estimated unknown source, inverse simulation with one release and different error level, error-free data ($\alpha = 0$) and corrupted data ($\alpha = 0.1$) (INV 2). 130

3.28. Observed and estimated release fluxes described by four stress period and obtained as average of the results of 10 neural networks results, inverse simulation with two release sources and different error level, error-free data ($\alpha = 0$) and corrupted data ($\alpha = 0.1$) (INV 3). 132
3.29. Estimated release fluxes in reference works for corrupted data ($\alpha = 0.1$) and, for any time step, the error bars related to one time the standard deviation (INV 3). 132
3.30. Observed and estimated release fluxes described by four stress periods at two known sources obtained as the average of the results of 10 neural networks at a known source. Inverse simulation with two release sources under different estimated error levels: error-free data ($\hat{\alpha} \to 0$) and perturbed data ($\hat{\alpha} \approx 0.10$) (INV 4). 134
3.31. Synthetic domain. 142
3.32. Sketch of the implemented neural networks. 144
3.33. Sketch of ANN1. 147
3.34. Sketch of ANN2. 149
3.35. Scenario SC1. Training loss for ANN1 alone (left), ANN2 alone with ANN1 frozen (center) and ANN1 jointly with ANN2 (right). The iteration axis in the right plot starts at the number of iterations already performed in the previous training. 151
3.36. Scenario SC2. Training loss for ANN1 alone (left), ANN2 alone with ANN1 frozen (center) and ANN1 jointly with ANN2 (right). The iteration axis in the right plot starts at the number of iterations already performed in the previous training. 152
3.37. SC1: Error plot of the estimated piezometric field (PINN minus MODFLOW), using the active cells in the numerical model. 153

3.38. SC1: Estimated piezometric field by the numerical model (left) and PINN (right) for the selected time $t = 0.01$. 154
3.39. SC1: Estimated piezometric field by the numerical model (top) and PINN (bottom) for the selected time $t = 0.25$, $t = 0.5$, $t = 1$. . . . 155
3.40. Heterogeneous aquifer: Error plot of the estimated piezometric field (PINN minus MODFLOW), using the active cells in the numerical model. 156
3.41. Heterogeneous aquifer: Estimated piezometric field by the numerical model (left) and PINN (right) for the selected time $t = 0.01$. . . 157
3.42. Heterogeneous aquifer: Estimated piezometric field by the numerical model (top) and PINN (bottom) for the selected time $t = 0.25$, $t = 0.5$, $t = 1$. 158
3.43. Heterogeneous aquifer. Left: numerical solution of the piezometric field with respect to the active cells at time $t = 0$, $t = 0.25$, $t = 0.5$ and $t = 1$. Right: conventional ANN prediction of the piezometric field with respect to the active cells at time $t = 0$, $t = 0.25$, $t = 0.5$ and $t = 1$. 160

4.1. The sewer network plan for Polesine Parmense. The 122 junctions are denoted by black points, while the red squares mark the 20 junctions where the basins are connected, representing potential inflow locations. The green star indicates the downstream end of the network, where the flow hydrograph is observed. 168
4.2. Flow hydrographs, observed at the network outlet (bottom), in relation to a precipitation event (top) and the unauthorized inflow from various nodes. 169

4.3. Confusion matrix for the test phase results of the ANN. The distance of each node from the network outlet is provided at the top of the confusion matrix. 172
4.4. Confined and unconfined aquifer sketch 175
4.5. Horizontal plane aquifer sketch 176
4.6. Active cells 181
4.7. Prescribed heads 182
4.8. Bottom elevations. Promontory located at the center of the lower half, which is expected to remain dry 183
4.9. W-E directional hydraulic conductivities 184
4.10. S-N directional hydraulic conductivities 184
4.11. Wells 185
4.12. Piezometric heads solution 195
4.13. Piezometric heads solution of different iterations 196
4.14. Aquifer saturated thickness after the completion of iterations. . . . 197
4.15. River inflow to the aquifer. 197
4.16. Vertical cross-section sketch of the aquifer. 198
4.17. Active cells in the vertical cross-section of the aquifer 200
4.18. Saturated thickness in the vertical cross-section of the aquifer . . . 203
4.19. Piezometric heads solution in the vertical cross-section of the aquifer 204
4.20. Earth dam sketch of a section with core and embankment 205
4.21. Piezometric heads solution of the earth dam section with core and embankment 206

List of Tables

2.1. Combination of GCMs and RCMs from the EURO-CORDEX project used in this study. 47

3.1. Annual mean temperature and annual precipitation over the basins: average, maximum and minimum values in the period 1934-2020. . 58

3.2. Type of data and elevation of the precipitation and temperature gauges. 61

3.3. ID, name, reference groundwater body, percentage of available data and ground elevation of the monitoring wells. 62

3.4. Differences of the median, 25th and 75th percentiles of the future SGIs with those evaluated in the historical period. The SGIs are defined through the SGI-SPI6 regression relationships for the historical period and at short- (ST), medium- (MT) and long-term (LT), under the RCP4.5 and RCP8.5. 78

3.5. Differences of the median, 25th and 75th percentiles of the future SGIs with those evaluated in the historical period. The SGIs are defined through the SGI-SPEI9 regression relationships for the historical period and at short- (ST), medium- (MT) and long-term (LT), under the RCP 4.5 and RCP 8.5 79

3.6. ID, name, reference groundwater body, percentage of available data and ground elevation of the monitoring wells considered. 81
3.7. MSE (m^2) between the output of the neural networks (NARX, LSTM, CNN) and the observed groundwater levels. 87
3.8. RMSE (m) between the output of the neural networks (NARX, LSTM, CNN) and the observed groundwater levels in the learning and testing period 2005-2018. 88
3.9. Differences between the 25th, 50th and 75th percentiles of the future standardized GWLs at short- (ST), medium- (MT), and long-term (LT) and the historical ones under the RCP4.5 scenario. Results obtained with the AI models proposed in this study and the regression model presented in Section 3.1.2. 96
3.10. Differences between the 25th, 50th and 75th percentiles of the future standardized GWLs at short- (ST), medium- (MT) and long-term (LT) and the historical ones under the RCP8.5 scenario. Results obtained with the AI models proposed in this study and the regression model presented in Section 3.1.2. 97
3.11. MSE between the output of the neural network and the simulated groundwater levels. 105
3.12. Hydraulic and geometry characteristics of the study domain. . . . 114
3.13. Summary of the input-target data for the investigated scenarios. . 115
3.14. Summary of the input-target data for the investigated scenarios. . 120
3.15. ME, MAE, RMSE, and NRMSE computed on concentrations (mg/l) and related to the 35 average concentration values of the 7 monitoring wells (FWD 1). 126

3.16. Observed and estimated source release fluxes (g/s) obtained as the average of the results of 10 neural networks with related metrics PAEE, SD for different error levels, error-free data ($\alpha = 0$) and corrupted data ($\alpha = 0.1$) (INV 1). 128

3.17. ME, MAE, RMSE, NRMSE and NE computed on source fluxes (g/s) described by four stress period and obtained as average of the results of 10 neural networks for different error level (INV 1). . . . 129

3.18. Actual and estimated source location (ζ, η) obtained as the average of the results of 10 neural networks with different data error levels, error-free data ($\alpha = 0$) and corrupted data ($\alpha = 0.1$) (INV2). . . . 129

3.19. Observed and estimated source release fluxes (g/s) obtained as the average of the results of 10 neural networks with related metrics PAEE, SD for different error levels, error-free data ($\alpha = 0$), and corrupted data ($\alpha = 0.1$) (INV2). 130

3.20. ME, MAE, RMSE, NRMSE and NE computed on source fluxes (g/s) described by four stress period and obtained as average of the results of 10 neural networks for different error level (INV 2). . . . 131

3.21. Comparison of the estimated and actual source release fluxes described by four stress periods at two known sources obtained as the average of the results of 10 neural networks with Ayvaz (2010), Jamshidi et al. (2020), and the present work, with level error $\alpha = 0.1$ and related statistical metrics (INV 3). 131

3.22. Comparison of statistical metrics with Ayvaz (2010) and Jamshidi et al. (2020) for different error levels, error-free data ($\alpha = 0$), and corrupted data ($\alpha = 0.1$) (INV 3). 133

3.23. Actual and estimated order of magnitude obtained as average of the results of 10 neural networks of the error on concentrations (INV 4). 133

3.24. Observed and estimated source release fluxes (g/s) described by four stress periods at two known sources obtained as the average of the results of 10 neural networks with related metrics PAEE, SD for different estimated error levels, error-free data ($\hat{\alpha} \to 0$) and perturbed data ($\hat{\alpha} \approx 0.10$) (INV 4)). 134
3.25. ME, MAE, RMSE, NRMSE and NE computed on source fluxes (g/s) described by four stress period at two known sources obtained as average of the results of 10 neural networks for different estimated error level (INV 4). 135
3.26. Hydraulic and geometric characteristics of the study domain 141
3.27. SC1: RMSE, MAE and NSE of the estimated solution by the PINN compared to the one obtained by the numerical model 154
3.28. SC2: RMSE, MAE and NSE of the estimated solution by the PINN compared to the one obtained by the numerical model 157

Introduction

Surrogate modeling is a powerful technique in the field of computational simulation that offers numerous advantages for solving complex problems efficiently. At its core, surrogate modeling involves creating a simplified approximation of a more complex system: the surrogate model emulates the behavior of the original system but it is computationally less demanding, making it a valuable tool in various domains, from engineering and physics to finance and environmental science. By providing an effective bridge between computational efficiency and accuracy, surrogate modeling empowers researchers and practitioners to explore and solve complex problems with confidence, making it an indispensable asset in the toolbox of scientific and engineering disciplines. Furthermore, surrogate models enable sensitivity analysis and optimization, facilitating better decision-making and design improvements. They can also be integrated into real-time systems, enhancing their predictive capabilities.

This PhD thesis aims to explore the application of surrogate modeling in groundwater. Investigating alternative modeling techniques for groundwater arises from the need that groundwater serves as a critical lifeline for approximately 2 billion people worldwide, acting as their primary source of drinking water. However, the global scenario reveals alarming statistics. Annually, humans withdraw a staggering 1000 cubic kilometers of groundwater (U.S. Geological Survey, 2018), equivalent to the volume of North America's Lake Erie. This massive consumption is primarily driven by agriculture, which accounts for 70% (United Nations, 2022) of total withdrawals, emphasizing its indispensable role in sustaining food production. Unsustainable practices have taken a toll on groundwater resources. The overexploitation of aquifers has led to depletion in 21 out of the world's 37 largest aquifers (Earth Security Group, 2016). Furthermore, the distribution of groundwater resources is highly uneven globally, with countries like India, China, the United States, Pakistan, and Iran being major consumers due to population density, agricultural demands, and water scarcity challenges (United Nations, 2022).

Climate change exacerbates these issues (Intergovernmental Panel on Climate Change, 2021), altering global temperatures and precipitation patterns, significantly impacting groundwater resources. Changing precipitation patterns limit water infiltration and recharge, reducing groundwater availability. Increasing temperatures lead to increasing water demand and excessive pumping further deplete stressed aquifers. Rising sea levels lead to increasing saltwater intrusion, rendering coastal groundwater unusable. Within this context, there is an urgent need to consider the climate change as a key factor that influences groundwater resources.

Aquifer contamination poses another grave concern. Industrial activities, agricultural practices, and improper waste disposal contribute to the pollution of groundwater, posing risks to both human health and the environment. Moreover, the recharge time of groundwater is of extended duration, ranging from decades to centuries, impeding its natural replenishment in some regions. Using surrogate modeling, researchers and practitioners have the capability to simulate and forecast the movement and distribution of contaminants, evaluate potential hazards, and formulate efficient strategies for pollution mitigation and safeguarding groundwater resources within practical timeframes. Understanding these impacts is crucial to ensure a sustainable future, by implementing effective groundwater management practices.

Surrogate modeling, as applied to groundwater, is a valuable approach to address the challenges and complexities associated with modeling subsurface water flow and transport (Asher et al., 2015). In fact, groundwater systems are inherently intricate, often characterized by a multitude of variables and parameters, and simulating their behavior with high-fidelity numerical models can be computationally expensive or, in some instances, impossible. These surrogate models can capture the essential relationships between inputs (such as geological properties, boundary conditions, and pumping rates) and outputs (such as water table elevations, contaminant concentrations, or groundwater flow rates), allowing for rapid simulations and decision support.

This work incorporates the goals provided by the European project InTheMED (https://inthemedprima.com/) of achieving integrated water resources management tools, by implementing innovative surrogate modeling approaches. In particular, this study focuses on developing surrogate models based on statistical approaches, conventional neural networks, deep learning techniques and physics-informed neural networks.

The thesis also addresses the challenges and limitations associated with surrogate modeling, including the selection of appropriate surrogate methods, the incorporation of uncertainty, and the lack of physical interpretability. In particular, the limitation in providing a physical interpretation has sparked debate within the scientific community. In this work, physics informed neural networks are proposed as a promising solution to handle groundwater problems by integrating the core concept of conventional neural networks with the fundamental laws of physics, allowing for a more comprehensive and interpretable modeling approach.

The structure of the thesis is as follows: In Chapter 1, the focus is on surrogate modeling by an overview of different types of surrogate models and an in-depth description of the models employed across various applications within the thesis. Chapter 2 focuses on climate change, offering an outline of the required methodologies for obtaining future climate projections at a local scale. In Chapter 3, different applications of surrogate modeling are presented. In particular it involves statistical approaches using standardized indices and regional climate models, and three different artificial neural network techniques to asses the impact of climate change on groundwater resources; neural networks for solving forward and inverse transport problems; neural networks for building user-friendly tools to assess different scenarios of climate change and agricultural polices; physics-informed neural network for solving transient unconfined groundwater flow. Chapter 4 presents some extra contents that have sparkled during the research work performed: a neural network is used to identify the inflow source in a foul sewer system and a teaching application is presented to explain how to model numerical groundwater flow with spreadsheets. Finally, the last chapter summarizes the thesis conclusions and offers suggestions for future research directions.

1

Surrogate modeling

1.1. Introduction and state of the art

Surrogate modeling, also known as metamodeling or smart modeling, is a technique used to create simplified and computationally efficient models that mimic the behavior of complex, computationally expensive models or systems. The primary goal of surrogate modeling is to reduce the computational burden associated with running complex simulations or models while still providing reasonably accurate results.

Surrogate modeling finds applications in optimization, sensitivity analysis, uncertainty quantification, and decision-making processes, offering researchers and engineers an efficient way to explore scenarios and make informed decisions without sacrificing result accuracy.

Surrogate models can be categorize into three broad groups (Asher et al., 2015):

data-driven approaches, projection-based methods and hierarchical or multifidelity methods.

Data-driven models

Data-driven surrogate modeling techniques have gained prominence in recent years due to their flexibility and adaptability. These methods rely on empirical models, often incorporating machine learning algorithms, to approximate the behavior of complex systems. The key feature of data-driven approaches is their reliance on input-output datasets, which consist of pairs of input parameters and corresponding system outputs. These datasets can either be generated by the complex model itself through simulations or experiments, or they can be available from observed data.

One of the strengths of data-driven approaches is their ability to capture complex and nonlinear relationships between inputs and outputs, making them suitable for a wide range of applications. They excel in cases where the underlying physics of the system is not fully understood or when the complexity of the model makes traditional analytical methods impractical. However, challenges such as overfitting and the need for high-quality training data remain areas of active research in this field.

Data-driven techniques have evolved over time, starting with Response Surface Methodology (RSM) in the early 1950s. Pioneered by Box and Wilson (1951), RSM aimed to approximate the relationship between input factors and the response of interest using low-degree polynomials. This early work laid the foundation for surrogate modeling and experimental design.

In the 1970s, kriging emerged as a powerful surrogate modeling approach. It was initially developed by Krige (1951) and later formalized by Matheron et al. (1962). Kriging incorporated spatial correlation and provided accurate predictions even with limited data. Its applications expanded into fields such as geology, hydrology, and environmental modeling.

Gaussian Processes (GPs) gained prominence in the late 1990s and early 2000s as a flexible and non-parametric surrogate modeling technique (Rasmussen, 2004). By allowing the modeling of complex relationships without assuming a specific functional form, GPs became effective in handling noisy or sparse data. They found applications in machine learning, optimization, and computer experiments.

Furthermore, the growth in computational ability of computers has played a significant role in the popularity of machine learning methods as surrogate models. Introduced in the 1990s, Support Vector Machines (SVMs) became powerful machine learning algorithms for surrogate modeling (Cortes and Vapnik, 1995). SVMs constructed hyperplanes in high-dimensional feature spaces to separate and classify data points, finding success in pattern recognition, regression, and anomaly detection.

Artificial Neural Networks (ANNs) recently have been employed as surrogate models. They excelled in approximating complex nonlinear relationships between inputs and outputs (McCulloch and Pitts, 1943).

Projection-based methods

Projection-based surrogate modeling techniques focus on reducing the dimensionality of parameter spaces. These methods project the governing equations of the system onto a basis of orthonormal vectors. By virtue of this practice, they transform the problem from a high-dimensional space to a lower-dimensional space, which can significantly simplify the modeling process.

Projection-based methods are particularly useful when dealing with systems that exhibit a high degree of parameter dependence or when seeking to identify dominant modes of behavior.

Principal Component Analysis (PCA) and Proper Orthogonal Decomposition (POD) are examples of projection techniques commonly used in surrogate modeling. While effective in reducing computational complexity, these methods assume linearity and may not capture highly nonlinear interactions present in some systems.

PCA, originally conceived in 1901 by Karl Pearson (Pearson, 1901) as an analog of the principal axis theorem in mechanics, later received independent development and nomenclature from Harold Hotelling in the 1930s (Hotelling, 1936). Depending on its specific application domain, PCA assumes various names, such as: the discrete Karhunen–Loève transform (KLT) within the realm of signal processing; the Hotelling transform within multivariate quality control; Proper Orthogonal Decomposition (POD) in the field of mechanical engineering; Singular Value Decomposition (SVD) of X (originating in the latter part of the 19th century); Eigenvalue Decomposition (EVD) of $X^T X$ within the field of linear algebra or factor analysis (for a comprehensive discussion on the distinctions between PCA and factor analysis, refer to Chapter 7 of Jolliffe and Cadima (2016).

Multifidelty methods

Hierarchical or multifidelity surrogate modeling approaches take a different route by simplifying the representation of the physical system itself. This simplification can involve ignoring certain processes and/or reducing numerical resolution. These methods are particularly valuable in cases where computational resources are limited or when rapid approximations are required.

One of the key advantages of hierarchical or multifidelity methods is their ability to balance accuracy and computational cost. By employing surrogates with varying levels of detail, practitioners can choose the appropriate level of fidelity depending on the specific task or stage of the analysis. However, determining the optimal hierarchy and ensuring smooth transitions between surrogate models are ongoing challenges in this field.

The Multiscale Finite Volume Method (MsFVM, Jenny et al., 2004) and the

Multiscale Finite Element Method (MsFEM, Hou and Wu, 1997) are the two mostly used techinques to upscale properties from the scale of measurements or a fine grid to a coarse grid for rapid computation.

This work mainly focus on data-driven surrogate modeling techniques. Specifically, the exploration includes statistical approaches involving simple linear regressions, traditional neural networks, deep learning neural networks, and physics-informed neural networks. The following subsections present the methodological aspects of surrogate modeling techniques employed to carry out the various applications developed in this thesis.

1.2. Statistical approach based on linear regression

Linear regression, a fundamental statistical approach, offers valuable insights into relationships between variables. It excels in modeling the linear associations between a dependent variable and one or more independent variables, providing interpretability through coefficient analysis. This technique, employed extensively in diverse fields, has both simple and multiple regression variations. Researchers often leverage it to perform data analysis, assess trends, build simple models and make predictions, aided by metrics such as the coefficient of determination (R-squared) and the Root Mean Square Error (RMSE). Linear regression is not only easy to implement but also straightforward to interpret, making it a widely favored choice in statistical modeling and data analysis. Its simplicity and transparency empower analysts and researchers to quickly gain insights from their data. Additionally, linear regression is computationally efficient, which means it can handle relatively large datasets and complex models without significant computational burden. However, it is essential to acknowledge its limitations, such as the assumption of linearity, sensitivity to outliers, and its inability to capture complex, non-linear relationships. Understanding its strengths, limitations, and contemporary extensions is essential for leveraging linear regression effectively in modern data-driven research.

Standardized indices

The computation of standardized indices involves the utilization of extensive historical time series of observed data of the variable in question. This process typically begins with the fitting of the observed variable values to a suitable probability distribution. Then, these fitted values are transformed into a standard normal distribution. Once the probability distribution is fitted along a representative historical period, it can be used to compute standardized indices for other periods. Standardized values close to zero suggest that the variable closely resembles the mean of the historical period, whereas positive or negative values indicate deviations from the mean of the reference period. The conceptual procedure to obtain standardized values of a variable of interest is reported in Figure 1.1.

Correlation analysis

The objective is to identify potential relationships among different standardized indices. In this study, the Pearson correlation coefficient is used to assess the correlation between different investigated indices. In fact, the Pearson correlation coefficient, denoted as r, is a statistical measure of the strength and direction of a linear relationship between two variables. The formula for the Pearson correlation coefficient is given by:

$$r = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum (X_i - \bar{X})^2 \sum (Y_i - \bar{Y})^2}}. \tag{1.1}$$

In this formula, X_i and Y_i represent individual data points from the two datasets,

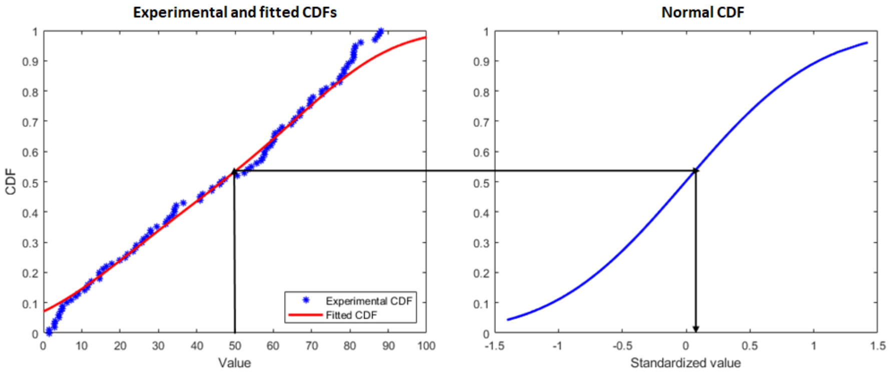


Figure 1.1. Left: experimental cumulative distribution function of the variable of interest and fitted cumulative distribution function that best approximates experimental data. Right: transformation into a standard normal distribution.

$\bar{X}$ and $\bar{Y}$ denote the means of the respective datasets. The Pearson correlation coefficient r ranges between -1 and 1, where $r = 1$ indicates a perfect positive linear correlation; $r = -1$ indicates a perfect negative linear correlation, and $r = 0$ suggests no linear correlation between the datasets. This coefficient provides valuable insights into the linear relationship between variables and is a fundamental tool in statistics and data analysis.

In this work, when the correlation coefficient between the investigated variables exceeds the predefined threshold of 0.6 (Evans, 1996), a linear regression model is employed to capture this relationship.

Linear regression model

A linear regression model is a statistical approach used to model the linear relationship between a dependent variable (y, target) and one or more independent variables (x, predictors or features). The general form of a simple linear regression model with one independent variable can be expressed as:

$$y = \beta_0 + \beta_1 x + \epsilon \ , \tag{1.2}$$

where y is the dependent variable; x is the independent variable; β_0 is the intercept; β_1 is the slope, and ϵ represents a random error term. The values of β_0 and β_1, that best fit the data points, are estimated by minimizing the sum of the squared differences between the observed and the target values. Linear regression can be extended to multiple linear regression when there are multiple independent variables:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p + \epsilon \ , \tag{1.3}$$

where $x_1, x_2, \ldots, x_p$ are the independent variables; $\beta_0, \beta_1, \ldots, \beta_p$ are the corresponding coefficients; ϵ represents the random error term.

1.3. Artificial neural networks

Machine learning models, such as neural networks, extend the applicability of simple linear regression to capture intricate, nonlinear relationships. Artificial Neural Networks (ANNs) are computational models inspired by the structure and functioning of the human brain. They are a subset of Artificial Intelligence (AI) learning algorithms that aim to mimic the way neurons interact and learn from data. ANNs have gained significant attention and achieved remarkable advancements across a wide range of domains, revolutionizing fields such as computer vision, natural language processing, robotics, and more.

The story of ANNs spans several decades and encompasses various milestones. It begins with the inspiration drawn from biological neurons, as Warren McCulloch and Walter Pitts introduced the perceptron in 1943 (McCulloch and Pitts, 1943) laying the foundation for ANNs. In the late 1950s, Frank Rosenblatt expanded upon the perceptron (Rosenblatt, 1958), developing an early form of ANNs capable of learning and decision-making. However, the limitations of early ANNs and the high computational costs led to a decline in interest during the 1970s, known as the "AI Winter" (Minsky and Papert, 1969). The field experienced reduced funding and a slowdown in AI and ANN research.

The resurgence came in the 1980s with the discovery of the backpropagation algorithm (Werbos, 1990), enabling the efficient training of multilayer perceptrons. This breakthrough allowed ANNs to learn complex patterns and gave rise to deep neural networks. Additionally, Yann LeCun's introduction of convolutional neural networks (CNNs) in the late 1980s (LeCun et al., 1990) revolutionized computer vision, demonstrating impressive results in image recognition tasks with LeNet-5.

In the 1990s, researchers developed recurrent neural networks (RNNs) tailored for sequential data processing (Rumelhart et al., 1986). RNNs, with their recurrent connections, excelled in tasks like speech recognition, language modeling, and machine translation.

The true resurgence of ANNs came around 2012, driven by advancements in computational power, availability of large-scale labeled datasets, and improved network architectures.

In recent years, ANNs have consistently achieved state-of-the-art performance across various domains. They have excelled in natural language processing, speech recognition, machine translation, and highlighting their versatility and potential for real-world applications. Additionally, ANNs provide valuable solutions for engineers, enabling advancements in fields such as image and signal processing, image recognition, object detection, denoising, image segmentation, audio processing, and time series analysis.

The field of ANNs continues to evolve rapidly, with ongoing research exploring novel architectures, optimization algorithms, regularization techniques, and interpretability methods. Transfer learning, reinforcement learning, and generative models like GANs further extend the capabilities of ANNs, ensuring they remain at the forefront of artificial intelligence and machine learning research.

1.3.1. Artificial neural network design and training process

At their core, ANNs consist of interconnected nodes, called artificial neurons or units, organized into layers. The most common type of ANN design is the feedforward neural network, which follows a unidirectional flow of information from the input layer through one or more hidden layers to the output layer (Figure 1.2).

In a feedforward neural network, the input layer receives the initial data or features. Each artificial neuron in the input layer is associated with a specific input value or feature. The information from the input layer is then passed to the neurons in the subsequent layers.

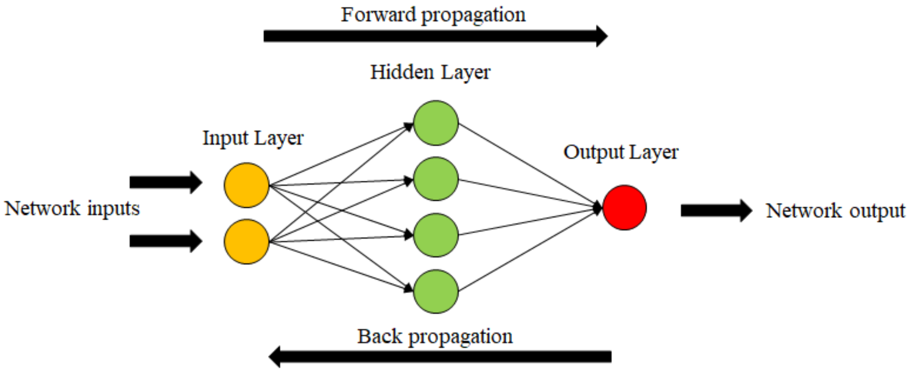


Figure 1.2. Artificial Neural Network sketch.

The hidden layers in an ANN provide intermediate processing stages between the input and output layers. Each hidden layer consists of multiple artificial neurons, which receive inputs from the previous layer and apply an activation function to produce an output.

The output layer is the final layer of the network and produces the desired output or prediction. The number of neurons in the output layer depends on the nature of the problem being addressed. For example, in a binary classification problem, there might be a single neuron in the output layer, while a multi-class classification problem may require multiple neurons corresponding to each class.

Activation functions play a crucial role in ANN design. They introduce non-linearity into the network, enabling the network to learn complex patterns and relationships. Commonly used activation functions include sigmoid, tanh (hyperbolic tangent), ReLU (Rectified Linear Unit) and softmax (for multi-class classification).

The design of an ANN involves choices regarding the number of hidden layers, the number of neurons in each layer, and the specific activation functions to be used. These choices depend on factors such as the complexity of the problem, the available data, and the computational resources.

The connections between neurons in different layers are defined by specific parameters of the network called weights and biases. Each connection has an associated weight and bias that determines the strength or importance of the connection. During the training process, these weights are adjusted iteratively to minimize the difference between the predicted output and the target, using techniques like backpropagation. The main steps for the traning procedure are described in the following,

Epochs and batch training

Training an ANN typically involves multiple passes over the entire training dataset, referred to as epochs. In each epoch, the training dataset are presented to the network, and the weights and biases are updated based on the calculated gradients. In large datasets, it is common to divide the data into smaller subsets called batches. The weights and biases are updated after each batch has been processed, rather than after each individual training dataset. This approach, known as mini-batch training, helps in achieving a balance between computational efficiency and convergence accuracy.

Normalization procedures

For arbitrary-sized and parameterized cases, appropriate normalization procedures can be applied to transform them into the unit domain $[0,1] \times [0,1]$. Even when dealing with intricate geometries, the scaling of coordinates and parameters to this specific range is feasible. When employing ANNs, it is common practice to normalize input variables to a standardized range, typically between 0 and 1. This normalization plays a pivotal role in stabilizing neural network training, mitigating issues such as vanishing gradients, and accelerating model convergence. Additionally, by employing this normalized range, the model gains the capacity to recognize patterns and relationships that transcend the specific magnitudes and units of input variables. This adaptability is essential when applying the model to scenarios of varying sizes and parameters. Crucially, the success of this approach hinges on the reversibility of the normalization process. By meticulously recording the scaling factors and means used for normalization, it becomes straightforward to apply an inverse transformation to the model's predictions, effectively reintroducing them into the original variable space. This ensures that the results remain interpretable and relevant within the context of the original problem domain. The utilization of normalization and transformation procedures not only enhances model stability and generalization but also equips the model to address a wide spectrum of scenarios, thereby increasing its versatility and adaptability.

Forward propagation

During forward propagation, the input data is fed into the network, and the information flows through the layers from the input layer to the output layer. The inputs are multiplied by the corresponding weights and biases and passed through the activation function of each neuron to produce an output. The outputs of one layer serve as inputs to the next layer until the final output is generated.

Calculating the loss

Once the output is generated, the prediction of the ANN is compared to the actual target. The discrepancy between the predicted output and the target is quantified using a loss function. The choice of the loss function depends on the specific problem being addressed. Common loss functions include mean squared error (MSE) for regression problems and categorical cross-entropy for classification problems. In this work, regression problems are presented and, in this regard, the Loss equation $L(\Theta)$ is reported:

$$L(\boldsymbol{\Theta}) = \frac{1}{N \cdot d_3} \sum_{i=1}^{N} \sum_{j=1}^{d_3} (\boldsymbol{h}_{\boldsymbol{\Theta}}(\boldsymbol{x}^{(\boldsymbol{i})}))_j - \boldsymbol{y}_{\boldsymbol{j}}^{(\boldsymbol{i})})^2 \quad . \tag{1.4}$$

Here, N represents the number of training input-target pairs. The term d_3 denotes the number of existing outputs, $\boldsymbol{h}_{\boldsymbol{\Theta}}(\boldsymbol{x}^{(\boldsymbol{i})})_{\boldsymbol{j}}$ is the network's output for the i-th input data $\boldsymbol{x}^{(\boldsymbol{i})}$, dependent on all network parameters $\boldsymbol{\Theta}$, and $\boldsymbol{y}_{\boldsymbol{j}}^{(\boldsymbol{i})}$ represents the target vector.

Backward propagation

Backward propagation is the key step in the training process. It involves calculating the gradients of the loss function with respect to the weights and biases in the network. This step determines how much each weight and bias contribute to the overall error of the network. Starting from the output layer, the gradients are computed using the chain rule of calculus (Griewank, 1988). The gradients are then propagated backward through the layers, updating the weights and biases based on the calculated gradients. This process is repeated iteratively for the whole dataset, adjusting weights and biases of the network to minimize the loss function.

To update the weights and biases, an optimization algorithm is used, most commonly gradient descent. Gradient descent adjusts the parameters in the direction of steepest descent of the loss function, gradually moving the network towards a configuration that minimizes the error. The learning rate is a crucial parameter in gradient descent, as it determines the step size taken in each iteration. A high learning rate may cause the network to overshoot the optimal solution, while a low learning rate can result in slow convergence. Different variations of gradient descent, such as stochastic gradient descent (SGD), mini-batch gradient descent, the Levenberg-Marquardt (LM) and the Adam optimization algorithms, are commonly used to balance computational efficiency and convergence speed. The Levenberg-Marquardt algorithm is particularly effective for nonlinear least squares problems and neural network training tasks with well-behaved loss function. However, it may encounter challenges when dealing with highly nonlinear or ill-conditioned problems. On the other hand, the Adam optimization algorithm is widely used in deep learning tasks due to its robustness and efficiency. It adapts the learning rates for individual parameters, allowing it to handle different scales of gradients effectively. The adaptive nature of this algorithm makes it well-suited for scenarios with sparse gradients, noisy gradients, or non-stationary objectives.

Regularization and overfitting

During the training process, overfitting can occur, where the network becomes too specialized to the training data and performs poorly on unseen data. Regularization techniques can be employed to mitigate overfitting. Common regularization methods include L1 and L2 regularization, dropout, and early stopping. L1 and L2 regularization add penalty terms to the loss function, discouraging large weight values and promoting simpler models. Dropout randomly deactivates a fraction of neurons during training, preventing the network from relying too heavily on specific neurons. Early stopping involves monitoring the validation error and stopping the training process when the validation error starts to increase.

Convergence and evaluation

The training process continues until a stopping criterion is met, such as a predefined number of epochs or reaching a desired level of accuracy. After training, the performance of the trained network is evaluated using separate validation and test datasets to assess its generalization ability. The validation dataset is used to fine-tune the hyperparameters and make decisions during the training process, while the test dataset provides an unbiased evaluation of the network's performance on unseen data. Evaluation metrics such as accuracy and precision are used to measure the performance of classification models. For regression tasks, metrics like Mean Squared Error (MSE), Mean Absolute Error (MAE), and R-squared are commonly used to assess the prediction accuracy.

In recent years, advancements in ANN design have led to the development of specialized architectures. For example, Convolutional Neural Networks (CNNs) have proven highly effective in image and signal processing tasks, leveraging convolutional layers to capture spatial features. Recurrent Neural Networks (RNNs) are designed for sequential data, incorporating recurrent connections that enable memory of past information. Moreover, researchers have explored innovative architectures, such as Long Short-Term Memory (LSTM) networks. LSTMs are designed to tackle the vanishing gradient problem encountered in traditional RNNs, and they exhibit exceptional performance in tasks related to natural language processing.

In the following, the implemented methodology for the applications are presented.

1.3.2. Multilayer perceptron networks

The MultiLayer Perceptron (MLP, Figure 1.3) is a type of artificial neural network that was invented in the 1960s as an extension of the Perceptron model (Rosenblatt, 1958). While the Perceptron model was a single-layer network capable of linearly separating data, the MLP introduced multiple layers of artificial neurons, enabling it to learn and model non-linear relationships.

The basic mathematical formula of an artificial neuron in an MLP involves three main components: the weighted sum of inputs, a bias term, and an activation function. Let's consider a neuron in a hidden layer or the output layer.

Weighted sum

The neuron receives inputs $(x_1, x_2, \ldots, x_n)$ from the previous layer, and each input is multiplied by a corresponding weight $(w_1, w_2, \ldots, w_n)$. The weighted inputs are summed together, along with a bias term (b), to produce the weighted sum (z):

$$z = w_1 \cdot x_1 + w_2 \cdot x_2 + \ldots + w_n \cdot x_n + b. \tag{1.5}$$

Activation function

The weighted sum is then passed through an activation function (g) to introduce non-linearity and produce the output (a) of the neuron:

$$a = g(z). \tag{1.6}$$

Commonly used activation functions include the sigmoid function, which squashes the output between 0 and 1; the hyperbolic tangent function, which maps the output between -1 and 1; and the rectified linear unit (ReLU) function, which returns the input value if positive, or zero otherwise.

To better understand the mathematical concept from a rigorous point of view, let's consider a simple MLP composed by three layers: input, hidden and output. Each layer is a Euclidean vector space. Considering d_1, d_2, d_3 the dimensions of such vector spaces, it means that input data are real vectors with d_1 components, output data are real vectors with d_3 components and, in the hidden layer, d_2 neurons are considered. The network, interpreted as a differentiable system, is simply a composition of multivariable vector-valued functions: affine transformations and linear or nonlinear activation functions (g), from $\mathbb{R}^{d_1}$ in $\mathbb{R}^{d_3}$. The signal moves from one layer to the next following (1.5) applied to all neurons in the layer:

$$\boldsymbol{a_d} = g(\boldsymbol{W} \cdot \boldsymbol{a_u} + \boldsymbol{b}), \tag{1.7}$$

where subscripts d and u refer to downstream and upstream, respectively, $\boldsymbol{a_u}$ is the input vector from a given layer, $\boldsymbol{a_d}$ is the output vector of a given layer, $\boldsymbol{W} \in \mathbb{R}^{d_d \times d_u}$, contains all the weights applying to the current layer and $\boldsymbol{b} \in \mathbb{R}^{d_d \times 1}$, contains all the bias terms.

1.3.3. Nonlinear autoregressive networks with exogenous inputs

The Nonlinear AutoRegressive network with eXogenous input (NARX) is a type of artificial neural network that is commonly used for modeling and predicting time series data. It is particularly effective in capturing non-linear dependencies and incorporating exogenous inputs, which are external factors that can influence the variable of interest.

The NARX model extends the basic AutoRegressive (AR) model (Pandit and Wu, 1983) by incorporating additional terms that capture the influence of past observations and exogenous inputs on the current output. This allows the NARX

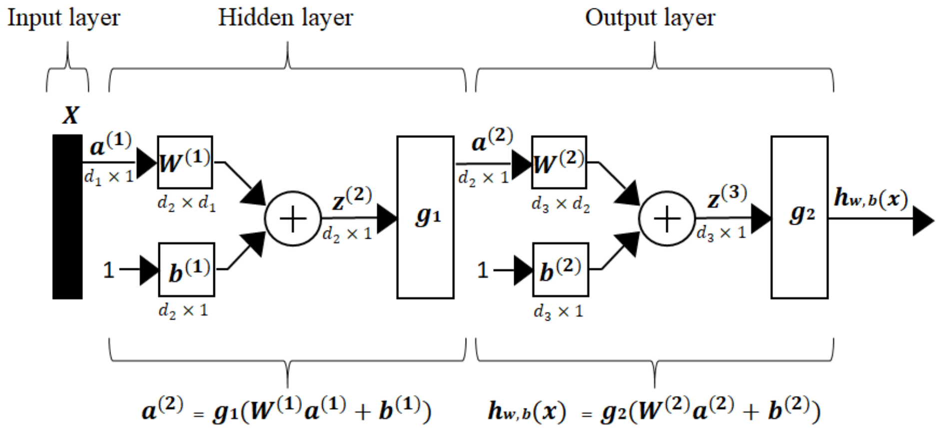


Figure 1.3. MLP sketch.

model to capture more complex and non-linear relationships compared to traditional AR models.

The specific architecture of the NARX network can vary. There are two types of architectures in the NARX network: open-loop (Series-Parallel Architecture) and closed-loop (Parallel Architecture). During the training phase, the open-loop network is utilized, incorporating available observations (y) and exogenous inputs (x) as input information. In contrast, the closed-loop scheme is employed during the prediction phase when observations are no longer accessible, resulting in the network's output also being used as an input.

Mathematically, the NARX model for the open- and closed-loop can be represented as follows, respectively:

$$h(t) = f(p(t-1), p(t-2), \ldots, p(t-m), x(t), x(t-1), \ldots, x(t-n)). \tag{1.8}$$

In the equations above, $h(t)$ represents the network output of the variable of interest at time t, and p is a generic term that should be replaced with the observed values y in the open-loop scheme or with the previously predicted output h of the network in the closed-loop scheme. Consequently, $h(t)$ is influenced by its past values of the observed values of the variable of interest $y(t-1)$, $y(t-2)$, $\dots$, $y(t-m)$, for the open-loop scheme, or is influenced by the previously predicted output of the network $h(t-1)$, $h(t-2)$, $\dots$, $h(t-m)$, for the closed-loop scheme as well as the exogenous inputs $x(t)$, $x(t-1)$, $\dots$, $x(t-n)$. The values of m and n determine the number of lagged terms of the output and exogenous inputs, respectively, that are considered in the model.

The function f represents the non-linear mapping between the inputs and the output. It is typically modeled using artificial neurons and activation functions within the NARX network. The NARX design generally consists of multiple layers of neurons (MLP neural network), including input, hidden, and output layers. The input layer receives the lagged values of the output and exogenous inputs, and the hidden layers perform computations to capture complex patterns. The output layer produces the predicted value of the target variable. Let's consider a NARX with three classic layers. The output $\boldsymbol{a^{(2)}}$ produced by the input and hidden layers, following the open-loop scheme, is:

$$\boldsymbol{a^{(2)}} = g_1(\boldsymbol{IW^{(1)}}\boldsymbol{a^{(1)}} + \boldsymbol{b^{(1)}} + \boldsymbol{LW^{(1)}}\boldsymbol{p^{(1)}}), \tag{1.9}$$

where the superscript indicates the reference layer of the network, g_1 represents the non-linear activation function, $\boldsymbol{IW^{(1)}} \in \mathbb{R}^{d_2 \times (d_1 \cdot n)}$ represents the weight matrix related to exogenous inputs, $\boldsymbol{a^{(1)}} \in \mathbb{R}^{d_1 \cdot n}$ is the exogenous input vector, $\boldsymbol{b^{(1)}} \in \mathbb{R}^{d_2}$ is the bias term vector, $\boldsymbol{LW^{(1)}} \in \mathbb{R}^{d_2 \times (d_3 \cdot m)}$ is the matrix of the input weights relative to the observations, and $\boldsymbol{p^{(1)}} \in \mathbb{R}^{d_3 \cdot m}$ is the given input. For the open-loop the scheme lagged observed values are used as $\boldsymbol{p^{(1)}}$, for the closed-loop scheme the term $\boldsymbol{h^{(1)}}$, that represents the lagged output predicted by the network used as new input is used for $\boldsymbol{p}^{(\mathbf{1})}$. The output is then transferred from the hidden layer to the output layer. This process is described by:

$$\boldsymbol{h}(\boldsymbol{x},\boldsymbol{y})^{(\mathbf{3})} = g_2(\boldsymbol{OW}^{(\mathbf{2})}\boldsymbol{a}^{(\mathbf{2})} + \boldsymbol{b}^{(\mathbf{2})}), \tag{1.10}$$

where g_2 represents the linear transfer function, $\boldsymbol{OW}^{(\mathbf{2})} \in \mathbb{R}^{d_3 \times d_2}$ is the matrix of the weights of the outputs of the hidden layer $\boldsymbol{a}^{(\mathbf{2})} \in \mathbb{R}^{d_2}$, $\boldsymbol{b}^{(\mathbf{2})} \in \mathbb{R}^{d_3}$ is the bias term vector. The predicted value of the network is $\boldsymbol{h}(\boldsymbol{x},\boldsymbol{y})^{(\mathbf{3})}$ for the open-loop scheme; for the closed-loop is $\boldsymbol{h}(\boldsymbol{x},\boldsymbol{h})^{(\mathbf{3})}$. For the sake of brevity, is shown only the closed-loop in Figure 1.4.

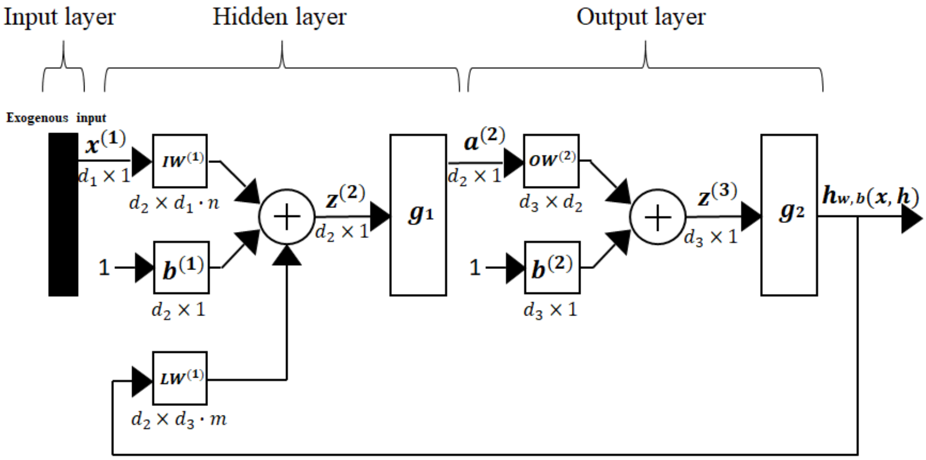


Figure 1.4. NARX closed-loop scheme.

1.3.4. Long-short term memory networks

Long Short-Term Memory (LSTM) is a type of recurrent neural network (RNN) architecture that was first introduced by Sepp Hochreiter and Jürgen Schmidhuber in 1997 (Hochreiter and Schmidhuber, 1997). LSTM was designed to address the vanishing gradient problem that commonly occurs in traditional RNNs, which hinders their ability to capture and retain long-term dependencies in sequential data (Hochreiter, 1991). The key innovation of LSTM is the introduction of memory cells (see Fig. 1.5), which are responsible for storing and updating information over time. Each memory cell has three main components: a forget gate, an input gate, and an output gate. These gates regulate the flow of information into, out of, and within the memory cell, allowing LSTM to selectively retain or discard information as needed. The functioning of an LSTM can be explained as follows.

Input processing

The LSTM layer processes each time step of the given time series sequentially by receiving an input vector and the previous hidden state as inputs. It is worth noting that the C state (the upper line in Fig. 1.5) represents the long-term memory, while the hidden state (the lower line in Fig. 1.5) represents the short-term memory. These inputs are utilized to calculate the forget gate, input gate, and output gate, in addition to determining the candidate value that will be stored in the memory cell.

Forget gate

The forget gate decides which information from the long-term memory line should be discarded from the memory cell. It takes the input x and the previous hidden state h as inputs, and calculates a sigmoid activation value

$$\boldsymbol{f_t} = \sigma(\boldsymbol{W}^{(\boldsymbol{f_{gate}})}[\boldsymbol{x_t}, \boldsymbol{h_{(t-1)}}] + \boldsymbol{b}^{(\boldsymbol{f_{gate}})}). \tag{1.11}$$

This value (between 0 and 1) multiplies the previous memory cell value and therefore retains only a fraction of its value. This allows the LSTM to selectively forget information that is no longer relevant.

Input gate

The next step is to select the new information to be provided to the chain. This procedure is divided into two parts: first, the input gate layer (i_t) is activated via a sigmoid function to define which inputs will be considered for updating the Cell state. Then, another layer, defined as the "new candidate gate layer" (nc$_t$), produces a new set of possible candidates through a hyperbolic tangent function.

$$\boldsymbol{i_t} = \sigma(\boldsymbol{W}^{\boldsymbol{(i_{gate})}}[\boldsymbol{x_t}, \boldsymbol{h_{(t-1)}}] + \boldsymbol{b}^{\boldsymbol{(i_{gate})}}). \tag{1.12}$$

and the new set of possible candidates is given by:

$$\boldsymbol{nc_t} = \tanh(\boldsymbol{W}^{\boldsymbol{(nc_{gate})}}[\boldsymbol{x_t}, \boldsymbol{h_{(t-1)}}] + \boldsymbol{b}^{\boldsymbol{(nc_{gate})}}), \tag{1.13}$$

Memory cell update

The input gate and forget gate outputs are combined to update the memory cell state, C. The previous memory cell state is multiplied by the forget gate output to discard unnecessary information. The result is added to the product of the input gate by the candidate value:

$$\boldsymbol{C_t} = \boldsymbol{f_t} \cdot \boldsymbol{C_{t-1}} + \boldsymbol{i_t} \cdot \boldsymbol{nc_t}. \tag{1.14}$$

Output gate

The output gate controls the flow of information from the memory cell to the output. First, it computes a potential hidden state value by applying the hyperbolic tangent to the updated memory cell value, and only a fraction of this value is retained; fraction that is computed from the input data and the hidden state as:

$$\boldsymbol{o_t} = \sigma(\boldsymbol{W^{(o_{gate})}}[\boldsymbol{x_t}, \boldsymbol{h_{(t-1)}}] + \boldsymbol{b^{(o_{gate})}}). \tag{1.15}$$

The output h for the next step is given by:

$$\boldsymbol{h_t} = \boldsymbol{o_t} \cdot \tanh\boldsymbol{(C_t)}. \tag{1.16}$$

Then, the procedure continues by utilizing the hidden state obtained at the final time step. To mitigate overfitting, the vector undergoes processing through the Dropout Layer, which stochastically sets certain input elements to zero with a specified probability (0.5 in the context of this application). The output generated by the Dropout Layer is then directed to the Fully Connected Layer, which establishes the synaptic connections required to link the dropout-affected hidden state with the Regression Output Layer, as a conventional neural network. The resulting weight matrix assumes dimensions of $R^{d_3 \times d_2}$, where d_3 represents the number of output neurons and d_2 is the number of hidden neurons. When the described weight matrix is multiplied by the hidden state vector, yields the final output vector of dimensions R^{d_3}. Finally, the Regression Output Layer computes the half-mean-squared-error loss between the generated output and the target output.

The ability of LSTMs to selectively retain and forget information over long sequences makes them particularly effective in tasks that require modeling complex temporal dependencies.

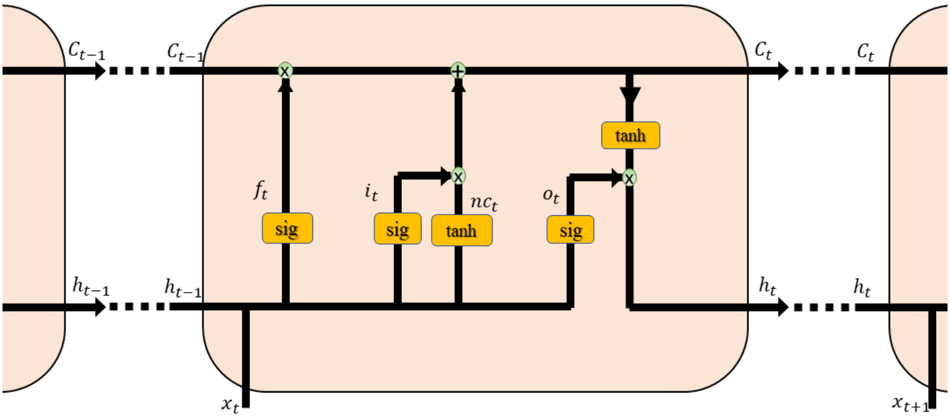


Figure 1.5. LSTM cell.

1.3.5. Convolutional neural networks

Convolutional Neural Networks (CNNs) are a type of deep learning model specifically designed for processing grid-like data, such as images. CNNs were first introduced by Yann LeCun in the late 1980s (LeCun et al., 1990). The functioning of a CNN can be summarized as follows (Figure 1.6).

Convolutional layers

The core building blocks of a CNN are the convolutional layers. Each layer consists of multiple learnable filters (also called kernels) that slide across the input image in a convolution operation. The filters extract spatial features by computing the dot product between their weights and the "pixel values" of the input. This operation captures different patterns such as edges, textures, and shapes at different scales. Let's consider an input image with dimensions $h \times w \times c$, where h represents the height, w represents the width, and c represents the number of channels. This input image undergoes processing in a convolutional layer that consists of

nf filters with dimensions $n \times m \times nf$, where n represents the height and m represents the width of each filter. The number of filters determines the number of channels in the output. Moreover, two additional parameters, namely stride (s) and padding (p), need to be specified. The stride determines the step size at which the convolutional filter moves vertically and horizontally across the input image, while the padding defines the number of extra border pixels added around the input image. In most cases, the stride is set to 1, meaning the filter moves one pixel at a time. The padding is adjusted so that the output produced by the convolutional layer maintains the same size as the initial input. The size of the output feature maps (FM) produced by the convolutional layer can be calculated using the following formula:

$$\mathrm{FM}_{\mathrm{row,column,channel}} = \left[\left(\frac{h-n+2p}{s}+1\right);\left(\frac{w-m+2p}{s}+1\right);\mathrm{nf}\right]. \tag{1.17}$$

Nonlinear activation

After the convolutional operation, a nonlinear activation function, typically ReLU (Rectified Linear Unit), is applied element-wise to introduce nonlinearity into the network. This activation function helps CNNs learn complex relationships between features.

Pooling layers

Pooling layers are used to downsample the spatial dimensions of the feature maps obtained from the convolutional layers. Common pooling operations include max pooling, which selects the maximum value within each local region, and average pooling, which calculates the average value. Pooling reduces the spatial resolution while retaining the essential features, allowing the network to be more robust to variations in the input, making the model translation-invariant and reducing the number of parameters.

Dropout layer

It is a regularization technique used to reduce overfitting and improve the generalization ability of the model. It works by randomly deactivating a fraction of neurons or units within the layer during each training iteration. This means that during training, some of the neurons do not contribute to the forward or backward pass, effectively making the network more robust and less reliant on any specific neuron.

Fully-connected layers

After several convolutional and pooling layers, the resulting feature maps are flattened and fed into fully connected layers. These layers connect every neuron in one layer to every neuron in the next layer (as in a feedforward neural network), allowing the network to learn high-level representations and make predictions based on the extracted features.

Softmax or regression layers

In classification tasks, a softmax layer is typically used at the output to produce a probability distribution over different classes. The softmax function normalizes the outputs and ensures that their sum adds up to 1, providing class probabilities for the input. In contrast, in regression tasks, the output layer is typically structured differently. Instead of using a softmax layer, a regression layer is employed. The regression layer directly outputs a continuous value, without the need for normalization or probability distribution interpretation. This setup is common in tasks such as predicting a numeric value based on a set of input features.

The strength of CNNs lies in their ability to automatically learn hierarchical representations from raw data, capturing local patterns in lower layers and gradually building more abstract features in higher layers. This hierarchical feature extraction enables CNNs to excel in tasks such as image classification, object detection, image segmentation, and even tasks beyond computer vision, such as natural language processing and speech recognition when combined with appropriate architectures.

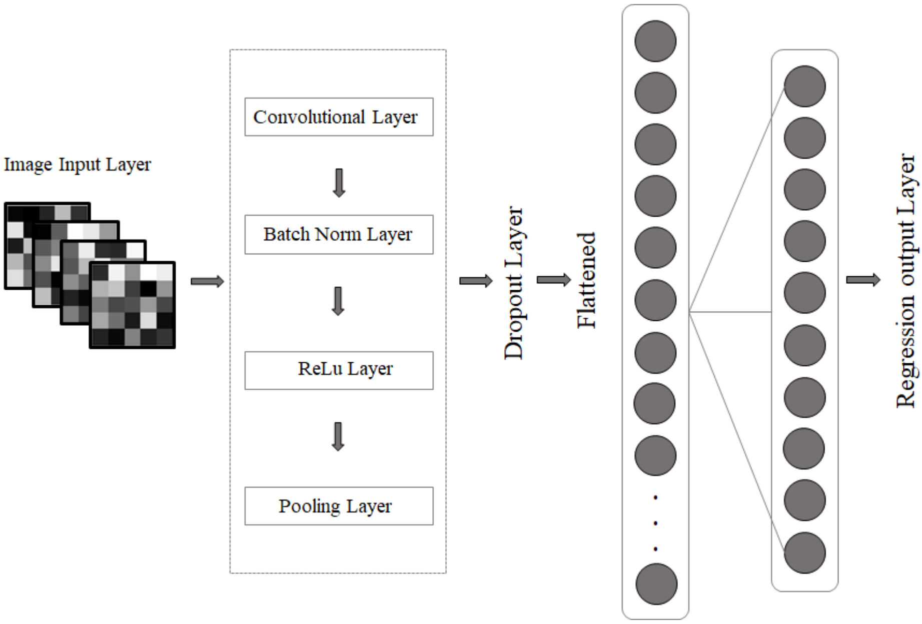


Figure 1.6. CNN sketch.

1.3.6. Physics-informed neural networks

Neural networks have proven to be highly effective in various machine learning applications. However, they face limitations when it comes to providing a clear understanding of their decision-making processes and adhering to the physical constraints essential for many scientific and engineering problems. These limitations can lead to inaccuracies and unreliability in their predictions, particularly when dealing with complex physical systems.

To address these challenges, there is a class of machine learning models known as Physics-Informed Neural Networks (PINNs). These models combine the computational power of neural networks with the governing physical laws that describe natural phenomena. Raissi et al. (2019) recently introduced PINNs to augment the physical interpretability of conventional ANNs and enhance their predictive capabilities. PINNs offer several advantages over traditional physics-based models. Through the incorporation of physics-based constraints into the ANN architecture, PINNs empower the model to more accurately capture the underlying physical principles governing the system under consideration. This integration can result in more precise and dependable predictions, especially in scenarios where traditional models may struggle due to high complexity or data scarcity. PINNs offer flexibility and generalization, as they can be trained on limited or noisy data, and they can accommodate intricate geometries and boundary conditions without requiring a specific mesh. This versatility positions PINNs as powerful tools for addressing a wide array of physical problems. Computational efficiency is a defining attribute of PINNs; once trained, they can be used swiftly, rendering them efficient for real-time decision-making or optimization tasks.

This physics-informed methodology has found application in various domains, spanning fluid dynamics, materials science, geophysics, and others (Waheed et al., 2021; Bajracharya and Jain, 2022; Cai et al., 2021; Mao et al., 2020; Lv et al., 2021; Zheng and Wu, 2023). For a comprehensive overview of the current state of PINNs, interested readers are directed to the recent publication by Lawal et al. (2022).

The core of the definition of a PINN typically involves incorporating physical principles or governing equations into the loss function. These constraints are essential for preserving critical principles, such as mass conservation, and have proven highly effective in solving complex PDEs within meshless domains, often surpassing the performance of traditional numerical methods (Raissi et al., 2019; Yang et al., 2021; He and Tartakovsky, 2021; Rezaei et al., 2022; Zhang et al., 2022). The primary idea behind a PINN is to train a neural network to approximate a solution to a partial differential equation (PDE) or a set of physical equations while simultaneously fitting the data (Figure 1.7). This hybrid approach allows PINNs to capture complex relationships between variables while upholding physical constraints, making them suitable for challenging problems, including those with limited data availability (He et al., 2020). Moreover, PINNs have demonstrated remarkable extrapolation capabilities, both spatially and temporally, a task that often challenges conventional ANNs, owing to the incorporation of physics principles into their training (Rezaei et al., 2022; Almajid and Abu-Al-Saud, 2022).

Here is how the loss function for a PINN is commonly structured:

Data-fitting term

This term measures the error between the prediction of the network and the target data. It encourages the model to fit the available data accurately. It represents the same term already presented in Section 1.3.1 for conventional ANNs.

Physics-informed term

The Physics-Informed Term works by quantifying the agreement between the predictions of the neural network and the expected values based on the governing physics. This agreement, when optimized, results in a network that respects and adheres to the fundamental physical constraints, allowing for accurate predictions while preserving the integrity of the system's underlying physics.

Conceptually, collocation points randomly generated within the domain of interest are used as given input for the neural network, aiming to generate corresponding outputs that represent the physical quantity. The PDE governing the investigated phenomenon is computed utilizing automatic differentiation (Griewank, 1988) at the collocation points. Then, by rearranging all the terms in this PDE to achieve equilibrium (i.e., making them equal zero), the loss associated with errors in replicating the PDE can be determined through the use of residuals. In essence, this involves calculating the average of the squared sums of the PDE values, which are derived from the network's output at the collocation points, and this average represents the corresponding error.

$$L_{\text{Physics-Informed}} = \sum_{\Omega} \left(f(\boldsymbol{x}, t, o, \frac{\partial o}{\partial \boldsymbol{x}}, \frac{\partial o}{\partial t}) \right)^2 . \tag{1.18}$$

In this expression: Ω represents the spatial and temporal domain over which the PDE is solved; o is the predicted solution by the neural network, that corresponds to the physical quantity, and $f(\boldsymbol{x}, t, o, \frac{\partial o}{\partial \boldsymbol{x}}, \frac{\partial o}{\partial t})$ is the PDE residual, which depends on the predicted solution o and its derivatives with respect to $\boldsymbol{x}$, and t.

Loss function

The overall loss function is a combination of the Data-Fitting and the Physics-Informed terms, typically weighted by hyperparameters that control the trade-off between fitting the data and adhering to the physics:

$$\text{Total Loss} = L_{\text{Data-Fitting}} + \lambda \cdot L_{\text{Physics-Informed}} . \tag{1.19}$$

λ is a hyperparameter that controls the importance of the physics-informed term relative to the data-fitting term. It determines the balance between fitting data and respecting physics.

The core challenge in defining the loss function for a PINN is to ensure that the neural network learns both the underlying physics and the data while maintaining a balance between these aspects. By optimizing this combined loss function, the PINN can provide solutions that are consistent with the physics governing the system while also fitting available data accurately.

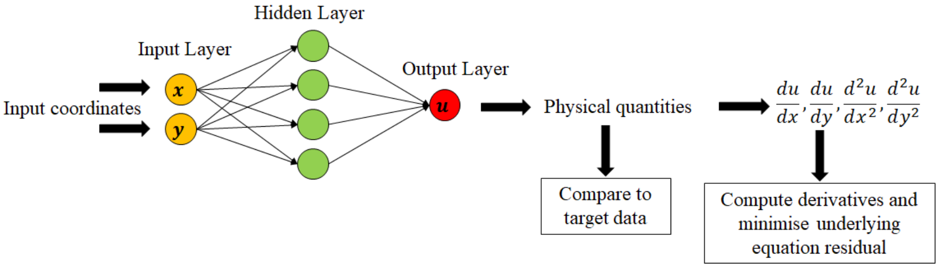


Figure 1.7. PINN sketch.

2

Climate change

2.1. Introduction

One of the primary objectives of this thesis is to evaluate the potential effects of climate change on groundwater resources, contributing to informed decision-making for sustainable water resource management and adaptation strategies.

Climate change exerts a profound and widespread impact on our planet. It affects ecosystems, weather patterns, sea levels, and poses significant risks to biodiversity, agriculture, and human health. This phenomenon leads to more frequent and severe extreme weather events, including hurricanes, droughts, heatwaves, and floods, with substantial economic and societal consequences.

Scientists widely agree that human activities, especially the release of greenhouse gases like carbon dioxide (CO2) from burning fossil fuels, are the primary drivers of climate change (Intergovernmental Panel on Climate Change, 2021).

Efforts to address climate change involve two main approaches: mitigation, which focuses on reducing emissions, and adaptation, which involves adapting to the changes that are already happening. These efforts span various domains, including policy, technology, and individual behavior changes. Transitioning to renewable energy sources, such as solar and wind power, plays a crucial role in reducing greenhouse gas emissions. International agreements like the Paris Agreement (UNFCCC, 2015) aim to unite countries in their fight against climate change by setting targets for emissions reduction and temperature limits.

Assessing the effectiveness of climate policies and fostering interdisciplinary collaborations across various fields are crucial steps in addressing the multifaceted challenges of climate change. With this aim, the Intergovernmental Panel on Climate Change (IPCC) is a globally recognized institution established by the United Nations to assess and communicate the scientific knowledge surrounding climate change. Comprising thousands of scientists and experts from diverse backgrounds, the IPCC conducts rigorous assessments of the latest climate research. Its comprehensive reports, published periodically, are considered authoritative references for climate science and its impacts (Stocker et al., 2013). Policymakers, governments, and organizations worldwide rely on these findings to make informed decisions and formulate strategies to combat climate change.

In the last Assessment Report (AR6, Intergovernmental Panel on Climate Change, 2021), the IPCC states that future climate conditions are closely tied to different future projections of human societies and their interactions with the environment. In this regard, Shared Socioeconomic Pathways (SSPs) are constructed on the basis of assumptions concerning socio-economic progress, demographic shifts, technological advancements, energy usage, land utilization, and other determinants influencing GreenHouse Gas (GHG) emissions, which are the primary drivers of global warming. In this context, the Coupled Model Intercomparison Project, Phase 6 (CMIP6) encompasses historical climate reanalysis and future projections providing Global Climate Models (GCMs) that are employed to simulate long-term climate forecasts spanning from 2015 to 2100 across the various SSP scenarios. The resolution of GCMs, typically ranging from 100 to 500 kilometers, may not be sufficient for reliable regional scale projections. To address this limitation, dynamic downscaling techniques have been employed to generate Regional Climate Models (RCMs) with higher resolutions, ranging from 10 to 50 kilometers, crucial for evaluating vulnerabilities, assessing impacts, and formulating adaptation strategies at the regional and local levels. In the context of CMIP6, RCM experiments are still under development and are not yet available. For this reason, this reaseach work uses the RCMs developed within the Coupled Model Intercomparison Project, Phase 5 (CMIP5), as part of the fifth Assessment Report of the Intergovernmental Panel on Climate Change (AR5, Pachauri et al., 2014). Differently from AR6, the climate projections provided by the AR5 relies on four Representative Concentration Pathways (RCPs) or scenarios, namely RCP2.6, RCP4.5, RCP6, and RCP8.5, to provide insights into GHG emissions (Stocker et al., 2013; Moss et al., 2010) and the climate projections span from 2006 to 2100 across the RCPs.

Nevertheless, raw outputs from RCMs require a bias correction process when used in medium-small scale basins for climate change impact studies (Teutschbein and Seibert, 2012). Additionally, to account for uncertainty in the results, it is recommended to use an ensemble of climate models, involving different combinations of GCMs and RCMs (D'Oria et al., 2018b; Todaro et al., 2022a). This ensemble approach helps capture the range of possible outcomes and provides a more comprehensive assessment of the uncertainties associated with climate change impacts.

In this chapter, an overview of the methodologies to be applied to obtain climate projections of precipitation and temperature at station scale and meteorological drought indices are presented. These features will be used in the Applications Chapter to build some of the presented surrogate modeling applications.

2.2. Historical data processing

2.2.1. Gap filling and interpolation procedures

In accordance with the guidelines established by the World Meteorological Organization (1987) to perform bias correction, a minimum of 30 years of observed data is required. Typically, in the AR5 framework, a common period for implementing bias corrections spans from 1976 to 2005, aligning with the projection period of climate models, commencing in 2006. The continuity of time series data is imperative for effective bias correction. Consequently, the need arises to address data gaps within the historical time series.

One frequently adopted method is the FAO procedure (Allen et al., 1998). According to this method, the gaps are filled using a linear relationship between the data at the target location and a twin location where the missing data are available. The data available in the two locations must exhibit a satisfactory correlation. This method was applied to fill gaps in groundwater level, precipitation, and temperature datasets. As recommended by Allen et al. (1998), a threshold correlation coefficient was adopted (0.7) to select twin stations. When the FAO method cannot be applied, another procedure is adopted. The gap is filled with a random value from a normal distribution (or a lognormal distribution for precipitation) which has been calculated on the available data for the considered month.

In addition, to compute meteorological indices that requires both precipitation and temperature (standardized precipitation-evaporation index, i.e. SPEI), it is crucial to have contemporaneous records of temperature and precipitation at the same location. In this work, some of the climate gauging stations considered lacked of temperature data. Consequently, once the gaps in the available time series were filled, the temperature data were interpolated to the precipitation station locations. In accordance with Moisello (1998), a reduction in temperature with increasing ground elevation was considered. Thus, during the recorded period and on a monthly scale, the coefficients "q" and "m" were determined in the following linear equation using the Ordinary Least Squares (OLS) method applied to the N locations with known temperature, T^o_j, and elevation, E_j:

$$T^o_j = q - m \cdot E_j \quad (j = 1, \ldots, N), \tag{2.1}$$

once the coefficients "q" and "m" were estimated, if Eq. 2.1 is applied to the sites where the temperature records exist, deviations (residuals) can be observed due to local peculiarities not accounted for by the linear regression. Subsequently, when estimating the final temperature, T_i, at any elevation, E_i, the residuals, weighted using an inverse square distance method, were incorporated to the result of Eq. 2.1, yielding the following relationship:

$$T_i = q - m \cdot E_i + \sum_{j=1}^{N} \lambda_{i,j} \cdot \epsilon_j, \tag{2.2}$$

where $\lambda_{i,j}$ is the weight of the ϵ_j residual of the temperature values in the j location with known temperature.

2.2.2. Meteorological and groundwater drought indices

Meteorological drought indices serve as pivotal components in the evaluation and quantification of the occurance, severity and extent of drought events. These indices rely on meteorological data, including precipitation, temperature, and evapotranspiration. Among the widely used indices, in this work, the Standardized Precipitation Index (SPI) and the Standardized Precipitation Evapotranspiration Index (SPEI) are used. The SPI quantifies precipitation anomalies across various time scales, aiding in the identification of drought conditions based solely on precipitation data. On the other hand, the SPEI is a drought index that integrates both precipitation and temperature, in terms of potential evapotranspiration, to assess and quantify meteorological drought conditions. Similarly, the Standardized Groundwater Index (SGI) serves as a statistical indicator to assess the severity of groundwater drought, and it shares a conceptual similarity with SPI and SPEI.

2.2.2.1. Standardized precipitation index

The SPI was developed by Mckee et al. (1993) as a statistical tool to assess the severity of meteorological droughts. The calculation of SPI, as for the bias correction, requires a long-term series of monthly precipitation data (typically 30 years or more, as recommended by the World Meteorological Organization (1987)) accumulated over different time periods of interest (e.g., 1, 3, 6, 9, 12, 24 months). The precipitation values for a specific month and time period are first fitted to an appropriate probability distribution, which is then transformed into a standard normal distribution. The methodology related to this procedure has been already explained in Section 1.2. SPI values close to zero indicate precipitation close to the long-term average, positive or negative values indicate above-normal or below-normal rainfall, and negative values less than -1 indicate the occurrence of meteorological drought.

The gamma distribution is commonly used to fit the cumulative precipitation data in the probability distribution function (PDF) for SPI calculation (Mckee et al., 1993; Soľáková et al., 2014; Stagge et al., 2015b), and it was also employed in this work.

$$p(x;k,\theta) = \frac{1}{\theta^k \Gamma(k)} x^{k-1} e^{-\frac{x}{\theta}}, \tag{2.3}$$

where:

- $p(x;k,\theta)$ represents the probability density function of the gamma distribution;

- x is the random variable for which you want to calculate the probability density;
- k is the shape parameter of the gamma distribution;
- θ is the scale parameter of the gamma distribution;
- $\Gamma(k)$ represents the gamma function.

In addition, it is important to consider the "zero precipitation problem" when dealing with seasons of low precipitation, where the accumulated precipitation over short time periods (typically 1 or 3 months) can be zero. However, the gamma distribution can only handle positive values. Therefore, following the approach suggested by Stagge et al. (2015b), the cumulative gamma distribution function was transformed into a piecewise probability distribution:

$$p(x) = \begin{cases} p_0 + (1 - p_0)G(x_{p>0}, \boldsymbol{\gamma}) & \text{for } x > 0 \\ p_0 = \frac{n_{p=0}+1}{2(n+1)} & \text{for } x = 0 \end{cases}, \tag{2.4}$$

where p is the probability distribution, p_0 is the zero precipitation probability, $n_{(p=0)}$ is the number of zeros occurring in the total data set of n values, $G(x_{p>0}, \boldsymbol{\gamma})$ denotes the Gamma distribution with parameters $\boldsymbol{\gamma}$, and x is one element in the series.

2.2.2.2. Standardized precipitation-evapotranspiration index

Temperature plays a significant role in hydrological processes, which is why, in addition to the SPI, the SPEI was considered in this study. The computation of the SPEI (Vicente-Serrano et al., 2010) follows a similar procedure to that of the SPI. However, in this case, the reference meteorological variable is the useful precipitation, i.e., the difference between precipitation and potential evapotranspiration

(PET). Due to the availability of only mean temperature data for the study area, PET was evaluated using the Thornthwaite method (Thornthwaite, 1948).

The potential evapotranspiration according to the Thornthwaite formulation (Thornthwaite, 1948) PET_{Thorn} [mm/day] is calculated as follows:

$$PET_{Thorn} = \begin{cases} 0, & \text{if } T < 0 \\ 16 \cdot \frac{N}{360} \cdot \left(10 \cdot \frac{T}{I}\right)^a & \text{if } T > 0 \end{cases}, \tag{2.5}$$

where N is the duration of sunlight in hours, varying with season and latitude, T is the average daily air temperature [°C], and I is a heat index calculated as follows:

$$I = \sum_{\text{Jan}}^{\text{Dec}} \max[0, T_m]^5 \cdot 1.514, \tag{2.6}$$

where T_m is the monthly mean temperature [°C]. The exponent a is calculated as follows:

$$a = (6.75 \times 10^{-7} \cdot I^3) - (7.71 \times 10^{-5} \cdot I^2) + (0.01792 \cdot I) + 0.49239. \tag{2.7}$$

The gamma distribution used for the SPI no longer accommodates the inclusion of useful precipitation data when negative values arise due to the influence of potential evapotranspiration. To address this, the log-logistic distribution, as suggested by Stagge et al. (2015b), was employed.

$$p(x; \alpha, \beta, \gamma) = \sum \left(\log\left(\frac{\beta}{\alpha} \left(\frac{x-\gamma}{\alpha} \right)^{\beta-1} \right) - \log\left(1 + \left(\frac{x-\gamma}{\alpha} \right)^{\beta} \right)^2 \right). \tag{2.8}$$

In this formula:

- x is the random variable for which you want to calculate the probability density;
- α is the shape parameter of the log-logistic distribution;
- β is the scale parameter of the log-logistic distribution;
- γ is the location parameter of the log-logistic distribution.

2.2.2.3. Standardized groundwater index

The SGI presents conceptual similarity with SPI and SPEI. However, there are notable distinctions. Unlike SPI and SPEI, SGI does not involve the accumulation of values over a specific period. Additionally, the distribution of monthly groundwater levels observed does not conform to the probability distribution functions (PDFs) previously analyzed. To address this, different distributions have been employed by various when analyzing groundwater data. These include the plotting position method (Osti et al., 2008) and the kernel non-parametric distribution (Vidal et al., 2010; Bloomfield and Marchant, 2013; Soleimani Motlagh et al., 2017).

The plotting position method is sensitive to the sample size, particularly when dealing with limited data. Hence, the preferred approach in this study is the kernel non-parametric method. The PDF of the model adopted in this study follows the formulation presented by Horová et al. Horová et al. (2012):

$$p(x) = \frac{1}{Nh} \sum_{k=1}^{N} K\left(\frac{x - x_k}{h}\right). \tag{2.9}$$

The formula suggests that the probability density at a given point x is computed as the sum of K kernel functions centered at each data point x_k. The kernel function is $\frac{1}{h}$ times a base function, typically a Gaussian or Epanechnikov kernel, which determines the shape of the kernel. The bandwidth h controls the spread of these kernels.

In simpler terms, the PDF is calculated by summing up contributions from individual data points, with each contribution being a scaled and shifted kernel function centered around that data point. The $\frac{1}{Nh}$ term normalizes the probability density function, where N is the sample size.

$$K(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}. \tag{2.10}$$

This equation represents the Gaussian kernel function, where the value of $K(x)$ is computed using the square root of 2π as the normalization factor and the exponential function with a coefficient of $-\frac{1}{2}$ multiplied by the square of x.

Once the distribution is established, the normalization procedure for obtaining the SGIs follows the same process as described for the meteorological indices.

2.3. Future climate projections

In this work, the future climate estimates for daily precipitation and daily mean temperature were obtained from a collection of 13 RCMs as part of the EURO-CORDEX initiative (Jacob et al., 2014). The combinations of GCMs and RCMs used in this study are listed in Table 2.1. The RCM data encompass a historical control period (1950/1970-2005) and a projection period from 2006 to 2100, considering various Representative Concentration Pathways (RCPs). Specifically, this study focuses on RCP4.5 and RCP8.5 scenarios.

Table 2.1. Combination of GCMs and RCMs from the EURO-CORDEX project used in this study.

		GCM				
		CNRM-CM5	**EC-EARTH**	**HadGEM2-ES**	**MPI-ESM-LR**	**IPSL-CM5A-MR**
RCM	**CCLM4-8-17**	x	x	x	x	
	HIRHAM5		x			
	WRF331F					x
	RACMO22E		x	x		
	RCA4	x	x	x	x	x

2.3.1. Downscaling and bias correction

To obtain climate information at local scale, the climate model data were downscaled from a regional scale, to the climate station locations. It consist of an interpolation of climate data from the nearest 9 nodes on the model grid to the designated station. This procedure was accomplished using the Inverse Weighted Distance method:

$$\hat{Z}(P_0) = \frac{\sum_{i=1}^{9} \frac{Z(P_i)}{d(P_i,P_0)^p}}{\sum_{i=1}^{9} \frac{1}{d(P_i,P_0)^p}}, \tag{2.11}$$

where:

- $\hat{Z}(P_0)$ represents the interpolated value at the target location P_0.
- $Z(P_i)$ is the RCM value at the grid node at location P_i.
- $d(P_i, P_0)$ represents the distance between the target location P_0 and the data point P_i.
- p is the power parameter that can be adjusted to control the influence of distance on the interpolation.

The value of p to control the weighting of data points based on their distance from the target location can be adjusted. Typically, a value of 2 is used for p, as in this study, depending on the desired weighting scheme.

After this interpolation, RCMs data underwent bias correction (Teutschbein and Seibert, 2012; D'Oria et al., 2017), since they are affected by systematic errors (biases) in their output, due to the need to simplify intricate climate processes, parameterize small-scale phenomena, and make assumptions about initial conditions and external forcing factors. The bias correction is based on the reference historical period (1976-2005). In this study, the Distribution Mapping method (Teutschbein and Seibert, 2012) was employed to adjust the Cumulative Distribution Functions (CDFs) of the climate model data (daily precipitation and temperature) at a monthly scale, i.e., processing each month separately, to align with the observed data distribution during the selected historical period. The same correction derived from the historical period was then applied to the future projections. The Gamma distribution to represent wet-day rainfall was employed (Piani et al., 2010) and the Gaussian distribution for the temperature time series (Teutschbein and Seibert, 2012) was adopted. For instance, following Teutschbein and Seibert (2012) and with reference to the day d of the month m of the control period, the corrected climate variable V^*_{contr} is derived from the RCM climate data V_{contr} using the following equation:

$$V^*_{\mathrm{contr}}(d) = F^{-1}\left(F\left(V_{\mathrm{contr}}(d)\middle|\beta_{\mathrm{contr},m}, \beta_{\mathrm{obs},m}\right)\right), \tag{2.12}$$

here, F represents the cumulative distribution function that approximates the data (e.g. Gamma and Gaussian distributions for precipitation and temperature, respectively), and F^{-1} is its inverse. $\beta_{\mathrm{contr},m}$ and $\beta_{\mathrm{obs},m}$ are the parameter vectors of the fitted distribution for the daily climate variable simulated by the RCM and observed for the month m in the control period, respectively.

The corrected daily variable for the scenario period V_{scen} is then evaluated from the RCM V_{scen} as:

$$V^*_{\mathrm{scen}}(d) = F^{-1}\left(F\left(V_{\mathrm{scen}}(d)\middle|\beta_{\mathrm{contr},m}, \beta_{\mathrm{obs},m}\right)\right). \tag{2.13}$$

Furthermore, before implementing the Distribution Mapping correction, the wet-day frequencies was adjusted. This procedure, applied only for precipitation, involved the identification of an RCM-specific threshold, as outlined by Teutschbein and Seibert (2012). This threshold was determined to ensure that the count of rainy days of the climate model aligned with observations. Consequently, any days with precipitation falling below this threshold were categorized as dry. By way of explanation, Figure 2.1 shows the concept of the bias correction.

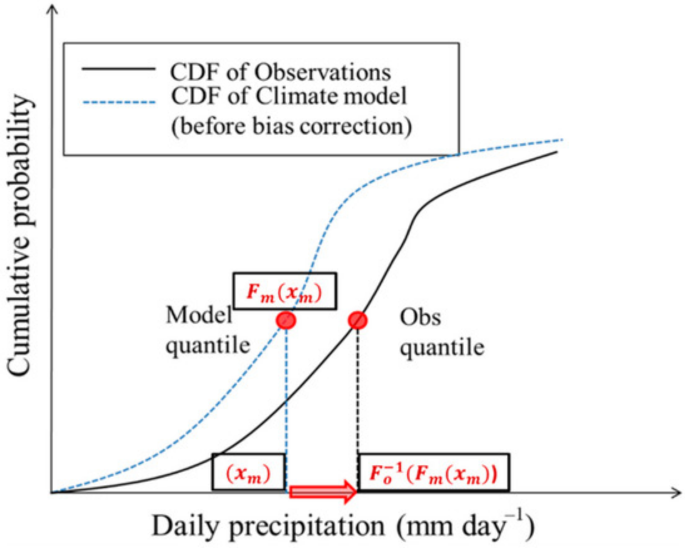


Figure 2.1. Distribution Mapping method (Gupta et al., 2019).

2.3.2. Future meteorological drought indices

Using climate model outputs, SPIs and SPEIs can be computed for selected future periods. These indices were calculated based on the probability distributions fitted to the historical observations. It should be noted that the computation of SPEIs outside the historical reference period may involve extrapolation, due to increasing temperature, which can lead to unreasonable values. In such cases, the SPEIs were constrained to fall within the range of extreme values allowed by the historical distribution.

Another critical question arises regarding the accuracy of climate models in representing standardized indices, such as SPIs and SPEIs. To this aim, as proposed by Stagge et al. (2015b) and Osuch et al. (2016), it is necessary to verify the alignment between the observed and estimated meteorological indices in a common historical period. If the results align well, it provides confidence that climate models can reliably predict meteorological indices in the future. To investigate this matter, a congruence check was conducted by comparing the probability distributions of observed SPIs and SPEIs with those derived from climate models for the historical period. The two-sample Kolmogorov-Smirnov test, determining whether two samples are drawn from the same distribution, was employed. This test was applied individually to the SPIs and SPEIs calculated using observed data and RCM data for each station and specific time windows. Each climate model underwent this analysis.

3

Applications

The simulation and prediction of aquifer conditions require the use of a conceptual model, which presents several challenges. Developing comprehensive conceptual models often requires a vast amount of data, including information about the geological structure, hydraulic properties of aquifers, boundary conditions, and pumping rates. Obtaining this level of detail can be difficult and expensive. In some cases, crucial data may be entirely unavailable. Even when sufficient data is available, constructing and analyzing complex conceptual models can be computationally demanding. This is because these models often involve numerous interconnected parameters, and simulating their behavior over different scenarios can require extensive computer processing time. This prevents the application of models in real-time decision support scenarios.

Surrogate models provide an efficient and expedient means to approximate complex, computationally intensive models. By capturing the essential features of the underlying system, surrogate models enable feasible simulations with increased processing speed.

In Chapter 3, the spotlight is on the application of different surrogate modeling techniques in different fields of groundwater. The impact of climate change on groundwater resources is analyzed using a statistical approach and three different artificial intelligence techniques. The aim is modeling and predicting groundwater behavior using historical and future climate information. In addition, surrogates can function as educational tools, offering valuable insights into the process of simplifying real-world physical models and demonstrating how they streamline complex realities. For this purpose, an ANN is used to build a user-friendly tool for aquifer management. Then, a neural network is implemented to solve direct and inverse problems in the field of groundwater transport and contamination. Finally, physics-informed neural networks are used for solving unconfined groundwater flow.

3.1. Impact of climate change on groundwater resources

The impacts of climate change extend to all corners of the globe, emphasizing the urgent need to evaluate its effects on various environmental, societal, and economic aspects (Jiménez Cisneros et al., 2015). One critical concern revolves around the future availability of high-quality freshwater resources.

This section of the thesis explores the potential implications of climate change on one of our most vital natural resources: groundwater. With surface water sources already experiencing challenges in terms of both quantity and quality, the reliance on groundwater as a primary water supply becomes increasingly common. Unfortunately, this reliance, coupled with the direct influence of climate change on aquifer recharge, raises serious concerns about the sustainability of groundwater in terms of availability and quality. By understanding these impacts, valuable insights into the challenges ahead can be obtained, and effective strategies for mitigating the effects of climate change on groundwater resources can be explored.

Two different methodologies have been applied to data collected in northern Tuscany (Italy) in an area served by a water company, but the procedure can be easily applied to different areas of interest. This research provides a comprehensive overview of the study area and the available data. The main results are then presented and thoroughly discussed, leading to the formulation of appropriate conclusions.

3.1.1. State of the art

Numerous studies have been conducted to quantify the effects of climate change on groundwater resources. However, this task presents significant challenges, particularly when employing traditional approaches that rely on complex hydrological models driven by future climate projections. In fact, as already highlighted in Chapter 2, climate change analyses typically involve considering an ensemble of climate models and various emission scenarios. This approach, however, often imposes substantial computational demands, potentially restricting the analysis to short time periods and a constrained number of scenarios.

However, although coupling complex hydrological models with climate change scenarios can pose a computationally demanding task, literature provides several studies addressing this challenge. Malcolm and Soulsby (2000) conducted an assessment of climate change impacts on a shallow coastal aquifer in northern Scotland utilizing a numerical MODFLOW model. They examined various climate change scenarios to investigate future changes in climate variables, such as precipitation and temperature, and their influence on groundwater levels (GWLs). Croley and Luukkonen (2003) applied a similar methodology in the Lansing area of Michigan. Brouyère et al. (2004) employed an integrated hydrological model

(MOHISE) driven by future climate projections from three General Circulation Models (GCMs) to evaluate the effects of climate change on groundwater availability in the Geer basin, Belgium. Citrini et al. (2020) assessed the impacts of climate change on karst spring discharge using the GR4J model and data from three climate models under different scenarios (RCP2.6, RCP4.5, and RCP8.5). Azizi et al. (2021) developed a MODFLOW model for the Varamin plain in Iran to evaluate changes in GWLs until 2050, utilizing climate data from ten Atmosphere-Ocean GCMs under the RCP2.6, RCP4.5, and RCP8.5 scenarios.

In the groundwater environment, surrogate models have proven useful in various groundwater management problems (Razavi et al., 2012; Asher et al., 2015; Rajaee et al., 2019). For instance, surrogate models can be employed to forecast groundwater levels using a set of drivers or proxy variables such as precipitation, temperature, and withdrawals. These models undergo a preliminary training process using historical data of the drivers (e.g., precipitation and temperature) and corresponding groundwater level responses. In this field, surrogate models typically fall into two categories: statistical approaches and machine learning algorithms, which have gained popularity due to advancements in computing capabilities.

Numerous studies (Khan et al., 2008; Bloomfield and Marchant, 2013; Kumar et al., 2016; Leelaruban et al., 2017; Soleimani Motlagh et al., 2017; Van Loon et al., 2017; Uddameri et al., 2019; Guo et al., 2021) have employed statistical approaches to explore the relationships between groundwater levels and various proxy variables.

Standardized indices are frequently employed to represent key variables of interest. These indices, including the Standardized Precipitation Index (SPI), the Standardized Precipitation-Evapotranspiration Index (SPEI), and the Standardized Groundwater Index (SGI), offer a standardized framework for quantifying and comparing these variables.

Notable findings from these studies include the research of Khan et al. (2008) in the Murra-Darling Basin, Australia, where they observed a strong SPI correlation with groundwater fluctuations, though diminished in areas with intensive irrigation and complex recharge patterns.

Bloomfield and Marchant (2013) extended this understanding to the United Kingdom, highlighting the SPI-SGI relationship. Specifically, the analysis revealed that the computation of the SGI introduces novel complexities in defining an appropriate statistical distribution for monthly groundwater levels, influenced by distinct local characteristics. Despite these challenges, a robust and conspicuous association between SPIs and SGIs was discerned. It is worth noting, however, that the authors emphasized the spatial variability of hydrological processes, which hinge on a multitude of driving factors beyond mere meteorological conditions.

Kumar et al. (2016) conducted research in Germany and the Netherlands, emphasizing the need for variable precipitation accumulation periods (3-24 months) to fit SPI and SGI, indicative of groundwater's delayed response to precipitation.

Leelaruban et al. (2017) found SPI with accumulation period of 24 months to be optimal for groundwater level assessment in Central US, underlining the region-specific nature of drought relationships.

Soleimani Motlagh et al. (2017) adopted clustering techniques to investigate groundwater drought in Iran, discovering varying precipitation accumulation periods and correlation coefficients among different clusters.

Van Loon et al. (2017) emphasized the importance of spatially variable accumulation periods when reconstructing groundwater droughts over extensive regions.

Uddameri et al. (2019) explored the use of SPI as a surrogate for groundwater drought in the Edwards Aquifer, Texas. While SPI and SGI showed statistical correlation, they cautioned that SPI might only provide qualitative predictions of groundwater drought

Guo et al. (2021) explored groundwater drought variations across US regions, attributing differences to agricultural and human activities and analyzing the temporal lag between meteorological and groundwater droughts.

Furthermore, climate models give the opportunity to project SPI and SPEI trends in future scenarios, with Stagge et al. (2015b) noting increased drought severity in the Mediterranean region and reduced occurrences in northern Europe due to climate change, aligning with historical data. These studies collectively shed light on the intricate relationships between meteorological indices and groundwater levels, underlining regional changes and the importance of customized approaches in assessing groundwater droughts, while also highlighting the potential implications of climate change on future drought patterns.

Shifting from statistical approaches to machine learning techniques, these algorithms have been employed to establish the relationships between groundwater levels and meteorological data, as discussed in the extensive reviews by Rajaee et al. (2019) and Tao et al. (2022). Several studies have utilized different machine learning techniques to assess piezometric levels over a historical period using a subset of available observations. These techniques include artificial neural networks (Coppola et al., 2003; Lallahem et al., 2005; Mohanty et al., 2010; Trichakis et al., 2011; Karthikeyan et al., 2013; Taormina et al., 2012; Sahoo and Jha, 2013; Shiri et al., 2013; Emamgholizadeh et al., 2014), and support vector machines (Yoon et al., 2011; Shiri et al., 2013; Suryanarayana et al., 2014). Some studies consider precipitation as the sole input feature, while others incorporate additional variables such as temperature, humidity, runoff, and evapotranspiration. Coastal aquifers have used tidal levels as input data, while other studies have included antecedent groundwater levels, pumping rates, and water demand.

Recent research has also investigated the impact of climate change on groundwater levels by combining machine learning techniques with future climate projections (Chen et al., 2010; Chang et al., 2015; Jeihouni et al., 2019; Idrizovic et al., 2020; Javadinejad et al., 2020; Ghazi et al., 2021; Gonzalez and Arsanjani, 2021).

Afrifa et al. (2022) provide a comprehensive review of mathematical and machine learning models used to examine the effects of climate change on groundwater level fluctuations.

Chen et al. (2010) investigated the effects of climate change and human activities on shallow groundwater levels in Wuqiao, North China Plain, using projections from 20 GCMs. They trained a Back-Propagation Artificial Neural Network (BP-ANN) with observed meteorological and pumping rate data to replicate groundwater levels. The ensemble mean of the climate models was used to project future climate variables and simulate deviations in groundwater levels using the trained network.

Chang et al. (2015) developed two ANNs to simulate and predict suprapermafrost groundwater levels, considering different scenarios of precipitation and temperature increases.

Idrizovic et al. (2020) studied the potential influence of climate change on the Toplica River catchment, where a calibrated hydrological model (HBV-light) was used to simulate runoff based on historical precipitation, temperature, and potential evapotranspiration data. An ANN was then trained using runoff values to reproduce groundwater levels, and future projections of precipitation and temperature from an ensemble of seven RCMs were used to estimate future groundwater levels under different climate scenarios (RCP4.5 and RCP8.5).

The following sections of this thesis present the implementation of two innovative methodologies for the application of surrogate models. These methodologies will encompass statistical approaches and machine learning techniques, providing a comprehensive and detailed exploration of their integration and utilization.

3.1.2. Study area and available data

The geographical focus of this study, as depicted in Figure 3.1, encompasses the northern part of Tuscany in Italy, covering an approximate area of 3000 km^2. The study area is of particular interest to an Italian water company, which seeks to assess the impact of climate change on water resources within this region. Over the years, there have been significant transformations in the anthropic occupation of this area. While agriculture had been the predominant activity in the past century, it has experienced a decline, with tourism emerging as the primary source of income (Pranzini et al., 2019). Currently, agricultural land accounts for approximately 14-16% of the total area, resulting in a relatively modest water demand. Natural forests, on the other hand, cover a significant portion, ranging from 55% to 70% of the total area (PTA, 2005).

Previous studies have already conducted investigations within this area (D'Oria et al., 2017, 2019), and accordingly, it has been divided into four watersheds based on water divides (Fig. 3.1): Magra, Serchio, Coastal Basins, and a portion of the Arno River basin known as Arno Portion. This division was necessary due to the distinct characteristics exhibited by each basin. Table 3.1 provides a summary of the annual precipitation and mean temperature recorded in these four basins during the period of 1934-2020.

Table 3.1. Annual mean temperature and annual precipitation over the basins: average, maximum and minimum values in the period 1934-2020.

Annual mean temperature (°C)	**MAGRA**	**COASTAL BASINS**	**SERCHIO**	**ARNO PORTION**
Average	13.2	13.2	12.9	14.8
Max	14.8	14.8	14.3	16
Min	11.3	11.8	11.4	13.3
Annual precipitation (mm)	**MAGRA**	**COASTAL BASINS**	**SERCHIO**	**ARNO PORTION**
Average	1539	1578	1536	1205
Max	2608	2579	2650	2039
Min	810	803	825	444

The basin of the Magra River covers three distinct areas: coastal, hilly, and mountain portions, with a total area of 938 km^2. The coastal part of the basin is not included in the study. The hilly area exhibits high spatial variability in temperature due to the influence of the coastal climate. The inner mountain

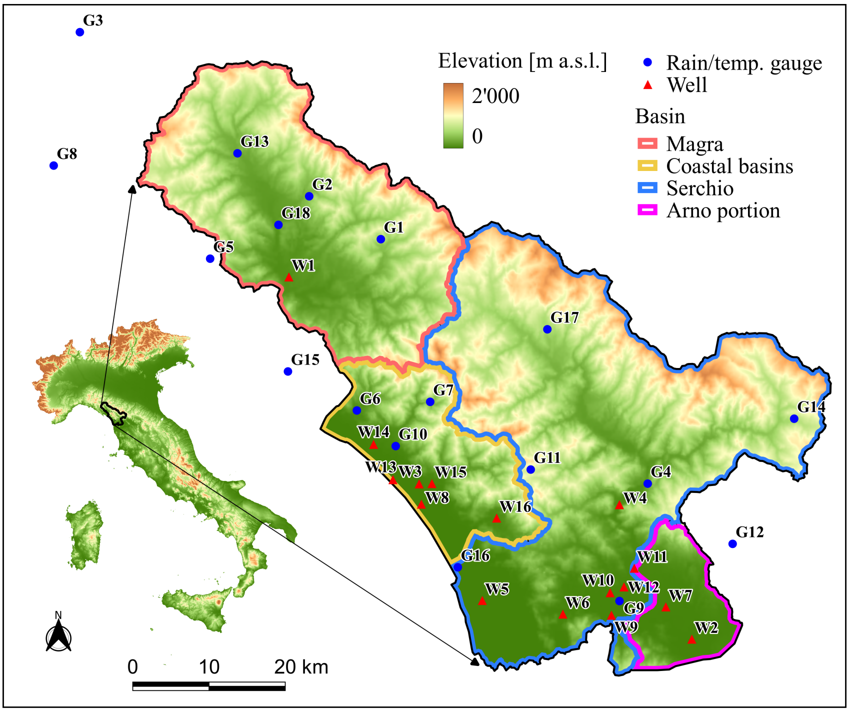


Figure 3.1. Location of the study area with indication of the climate stations, monitoring wells and river basins.

area experiences average winter temperatures close to zero and moderate snow accumulation, with high precipitation occurring in the internal regions.

The Coastal basins, spanning an area of 383 km^2, are situated between the Apuan Alps and the Tirreno Sea. These basins receive high levels of precipitation due to the proximity of the Apuan Alps, with the most intense rainfall occurring in late spring, late summer, and autumn. The winter season is characterized by sporadic and short-duration snowfall.

The Serchio River basin, covering 1545 km^2, originates in the Apennine area north of the Province of Lucca and flows into the Tirreno Sea. The longitudinal orientation of the basin in relation to the sea makes it one of the wettest areas in Italy, with annual total precipitation exceeding 2500 mm per year on the Apuan hills.

The Arno portion, with an area of 186 km^2, was historically occupied by swamps and a lake with irregular drainage to the Serchio River or Arno River based on seasonal variations. The area was later reclaimed through an artificial channel, redirecting the water to the Arno River. Precipitation in this area is distributed throughout the year, with two distinct periods: abundant and regular rainfall from January to May, and significant but irregularly distributed precipitation from October to December.

For this study, data from 18 gauging stations (Table 3.2) and 16 wells (Table 3.3) were considered. The precipitation and temperature data extend to neighboring regions (Liguria and Emilia Romagna regions) and were obtained from the Environmental Agency of each region. The historical daily precipitation and temperature database used in previous studies was updated until 2020. The selected stations represent the historical climate due to their long records. The wells provide groundwater level measurements from 2005 to 2020, except for the S. Pietro a Vico well, which had limited data and was excluded from the analysis.

All the wells included in the study have been classified according to the European Directive 2008/105/CE and subsequent national laws as belonging to underground water bodies. In the Magra basin, only one monitoring well (Bandita 7) is available, located in the city of Aulla in the bed aquifer of the Magra River. The Magra groundwater body (21MA010) extends to a depth of a few tens of meters, resting on the impermeable sediments of the Rusciniano-Villafranchiano substratum.

The Coastal basin comprises seven monitoring wells, which are part of the Versilia and Apuan Riviera groundwater body (33TN010). It is a multilayer system with silt or clayed-silt lenses, and direct contact between the aquifer horizons is

Table 3.2. Type of data and elevation of the precipitation and temperature gauges.

ID	**Name**	**Rain gauge**	**Temp. gauge**	**Elevation [m a.s.l.]**
G1	Arlia	x	x	460
G2	Bagnone	x	x	195
G3	Bedonia	x	x	500
G4	Borgo a Mozzano	x		100
G5	Calice al Cornoviglio	x	x	402
G6	Carrara	x	x	55
G7	Casania	x		845
G8	Cembrano	x	x	410
G9	Lucca	x	x	16
G10	Massa	x	x	150
G11	Palagnana	x		861
G12	Pescia	x	x	78
G13	Pontremoli	x	x	340
G14	S. Marcello Pistoiese	x	x	618
G15	Sarzana	x	x	26
G16	Viareggio	x	x	0
G17	Villacollemandina	x		502
G18	Villafranca Lunigiana	x	x	156

prevalent. The main groundwater flow is supplied by the upstream basins and alluvial fans of the streams.

The Serchio basin includes six monitoring wells. The Decimo well is situated in the upper-medium valley of the Serchio River groundwater body (12SE020), with a depth of 20-30 meters resting on the impermeable sediments of the Pliocene substratum. This phreatic aquifer exhibits lateral continuity along the course of the Serchio River and its tributaries. The other wells are located in the Lucca plain groundwater body, characterized as a phreatic aquifer within the Serchio zone (12SE011).

Two wells are located in the Arno portion basin, belonging to the Lucca plain -

Bientina area groundwater body (11AR028). The aquifer in this area is primarily phreatic, with a shallow confining layer recognized only in the southern region.

Table 3.3. ID, name, reference groundwater body, percentage of available data and ground elevation of the monitoring wells.

ID	Name	Groundwater body	% data	Elevation m a.s.l.
W1	Bandita 7	21MA010	73.4	68.00
W2	Corte Spagni	11AR028	83.8	9.07
W3	Cugnia	33TN010	91.7	4.00
W4	Diecimo	12SE020	60.9	65.00
W5	Flor Export	12SE011	64.6	1.67
W6	Nozzano	12SE011	78.6	16.43
W7	Paganico	11AR028	72.4	13.00
W8	Percorso vita	33TN010	78.1	1.56
W9	Ronco	12SE020	79.7	11.67
W10	Salicchi	12SE011	83.3	27.12
W11	S.Alessio	12SE011	71.9	18.87
W12	S.Pietro a Vico	12SE011	12.0	30.69
W13	Sat 1	33TN010	75.5	1.50
W14	Unim	33TN010	91.7	19.91
W15	Via Barsanti	33TN010	91.7	20.00
W16	Via Romboni	33TN010	88.0	37.92

During data collection, gaps were present within the time series. To fill these blanks and ensure a continuous set of observations, the FAO method (Allen et al., 1998) was adopted, as presented in Section 2.2. This method was applied to fill gaps in groundwater level, precipitation, and temperature datasets. Notably, the Bandita7 well, even after the FAO filling process, still contains missing data due to its inadequate correlation with the other wells.

In addition, out of the 18 climate gauging stations, four lack temperature data. Since this study requires both precipitation and potential evapotranspiration data, it is crucial to have contemporaneous records of temperature and precipitation at the same location. Consequently, once the gaps in the time series were filled, the temperature data were interpolated to the precipitation station locations following the procedure described in Section 2.2.

3.1.3. Future climate projections

The future climate estimates for daily precipitation and daily mean temperature were obtained from a collection of 13 RCMs (Table 2.1) as part of the EURO-CORDEX initiative (Jacob et al., 2014). Then, as highlighted in Section 2.3, the climate model data were downscaled to 18 climate stations and underwent bias correction based on the reference historical period of 1976-2005. The Distribution Mapping method (Teutschbein and Seibert, 2012) was employed to adjust the Cumulative Distribution Functions (CDFs) of the climate model data (daily precipitation and temperature) at a monthly scale to align with the observed data distribution during the selected historical period (see Section 2.3.1). The future analysis was conducted for short-term (2006-2035), medium-term (2036-2065), and long-term (2066-2095) periods. For the sake of conciseness and to provide an overarching view of predicted climate changes, Figure 3.2 presents annual precipitation and annual mean temperature values averaged across the entire study area for both historical and future periods. The data are displayed in the form of a 10-year moving average to reduce the influence of natural climate variability and emphasize climate change components.

Analysis based on both the RCP4.5 and RCP8.5 scenarios, as well as median values, reveals that annual precipitation is not expected to undergo substantial modifications in the future under both scenarios (Fig. 3.2a). However, it is important to note that there is considerable variability among different climate models, indicating a significant level of uncertainty in future precipitation estimations. Regarding temperature trends (Fig. 3.2b), a clear and consistent upward trajectory is observed for the future across both RCP scenarios. Both historical climate data and climate models indicate that temperature has been increasing since the 1990s, and this upward trend is projected to continue until around 2040 for both RCP scenarios. After this period, RCP8.5 suggests a more pronounced warming of the study area.

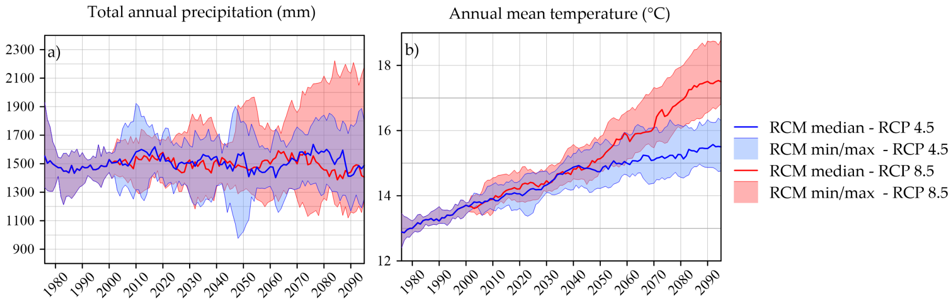


Figure 3.2. Total annual precipitation (a) and annual average of the mean daily temperature (b) in terms of 10-year moving average observed and forecasted by the RCMs under the RCP4.5 and RCP8.5 scenarios. Average values over the entire study area.

3.1.4. Statistical approach

In this part of the thesis a useful methodology to evaluate the impacts of climate change on groundwater levels is presented. Linear regression models were used to establish predictive relationships between groundwater index (SGI) and meteorological indices (SPI and SPEI). The relationships evaluated according to the historical data have been applied to future SPI and SPEI values, estimated by means of an ensemble of regional climate models (RCMs), to infer future SGI indices under different climate scenarios (RCP 4.5 and RCP 8.5).

3.1.4.1. Set up of the surrogate model

This section offers a brief description of the steps involved in setting up the surrogate model. It outlines the key procedures such as indices calculation, correlation analysis, and regression, as well as considerations for future levels.

Standardized Precipitation Index (SPI)

The SPI was evaluated at the station scale using long-term precipitation records from the reference period of 1934-1993. The gamma distribution was used to fit the cumulative precipitation data in the probability distribution function (PDF) for SPI calculation (see Section 2.2.2.1). The distribution function fitted over the reference period was utilized to calculate the SPIs for more recent years (2005-2020) when groundwater level data were available. The decision to not extend the reference period until 2020 was based on the detection of climate change effects in the study area since the '90s (D'Oria et al., 2017).

After computing the SPIs at each gauging station, they were processed to derive an average value using the Thiessen polygon method. Specifically, the average SPIs for each basin as well as for the entire study area were obtained.

Standardized Precipitation-Evapotranspiration Index (SPEI)

The SPEI was evaluated at the station scale using long-term temperature records from the reference period of 1934-1993. The log-logistic distribution was employed to fit the cumulative useful precipitation data, i.e., the difference between precipitation and potential evapotranspiration, in the probability distribution function (PDF) for SPEI computation (see Section 2.2.2.2). Similar to the SPI, the distribution function fitted over the reference period was utilized to calculate the SPEIs for more recent years (2005-2020) when groundwater level data were available; Thiessen polygons were employed to compute areal averages of the SPEIs at the station level for each basin as well as for the entire study area.

Standardized Groundwater Index (SGI)

The SGI was evaluated for each well using groundwater level records from the period of 2005-2020. The kernel distribution was used to fit the data in the probability distribution function (PDF) for SGI calculation (see Section 2.2.2.3). After adjusting the distribution to fit the data, the method outlined in Section 1.2 was employed to derive the standardized index.

Correlation between meteorological and groundwater indices

To investigate potential relationships between meteorological and groundwater indices, the Pearson correlation coefficient between meteorological and groundwater indices was computed for each monitoring well. This involved weighing meteorological indices based on the corresponding basin and correlating them with the SGI in the period 2005-2020. The analysis spanned eight time windows (1, 3, 6, 9, 12, 18, 24, and 36 months). Additionally, as suggested by Bloomfield and Marchant (2013), investigations were conducted to explore whether introducing potential delays between meteorological and groundwater indices (e.g., shifting the SPI or SPEI backward) could enhance their correlation. A threshold correlation coefficient of 0.6 was adopted, following (Evans, 1996), to identify a satisfactory linkage between the two indices. For correlations meeting this criterion, the relationship between these two indices was further examined through linear regression analysis (see Section 1.2).

Future SPIs and SPEIs

Using the climate model outputs, the SPIs and SPEIs were computed for both the historical period and selected future periods (see Section 2.3.2). These indices were calculated based on the probability distributions fitted to the historical observations from 1934 to 1993.

The question arises as to whether the downscaled and bias-corrected results obtained from the climate models accurately represent the observed standardized indices, such as SPIs and SPEIs, during a common historical period, as emphasized by Stagge et al. (2015a) and Osuch et al. (2016). If the results align well, it provides confidence that the climate models can reliably predict the meteorological indices in the future.

To investigate this matter, a congruence check was conducted by comparing the probability distributions of the observed SPIs and SPEIs with those derived from the climate models for the historical period of 1976-2005. The two-sample Kolmogorov-Smirnov test, which determines whether two samples are drawn from the same distribution, was employed. In fact, the application of the distribution mapping procedure as a bias correction method (Teutschbein and Seibert, 2012) guarantees congruence at the single-month scale, although it may not be assured for longer time windows. This test was applied individually to the SPIs and SPEIs calculated using both the observed data and the RCM data for each station and specific time windows. Each climate model was subjected to this analysis.

Future SGIs

Assuming that the regression equations determined for the historical period (2005-2020) remain valid in the future, they were utilized to estimate the SGIs based on the projected meteorological indices (SPIs or SPEIs). As already mentioned for the future climate projections, the future analysis was conducted for short-term (2006-2035), medium-term (2036-2065), and long-term (2066-2095) periods.

3.1.4.2. Results

This section provides a summary of the main findings. Firstly, the SPIs, SPEIs, and SGIs calculated during the historical period (2005-2020) are presented. Then, the correlations between meteorological and groundwater indices are examined, identifying their relationships. Finally, the future projections of SGIs are presented.

Historical SPIs, SPEIs and SGIs

Although SPIs and SPEIs were originally calculated at the station scale, for the sake of brevity, Fig. 3.3 illustrates their averages over the basins of interest and during the period when groundwater level data were available (2005-2020). The selection of time windows of 6, 9, and 12 months is based on the highest observed correlations between meteorological and groundwater indices within these accumulation periods.

Regarding the variability of SPIs among the basins, it appears to be statistically non-significant. Focusing on the 12-month time window (Fig. 3.3a, SPI-12), all basins exhibit a drought period that initiates in 2005 and concludes in 2009. Another significant drought event is identified from 2012 to 2013, with the Magra basin experiencing relatively milder conditions. Furthermore, within the context of the 12-month time window, the Arno portion basin records the lowest SPI values in 2008, while the Serchio basin exhibits the highest values during 2012-2013.

SPEI values (Fig. 3.3b) reveal drought periods that closely align with those identified by SPIs. On average, focusing on the 12-month time window, SPEIs tend to exhibit lower negative values and moderately higher positive values. Specifically, the lowest SPEI value is observed in the Arno portion basin in 2012, while the highest value is noted in the Serchio basin in 2014.

In addition, as previously indicated, the two-sample Kolmogorov-Smirnov test was conducted to compare meteorological indices obtained using historical and RCM information. For SPIs, at a significance level of 5%, nearly all samples passed the test, with only a few exceptions (1%) resulting in a *p*-value slightly below the threshold. For SPEIs, the percentage of no-passing samples increases to 20% but remains relatively low. The results of the Kolmogorov-Smirnov test confirm the reliability of SPIs and SPEIs evaluated using climate model data.

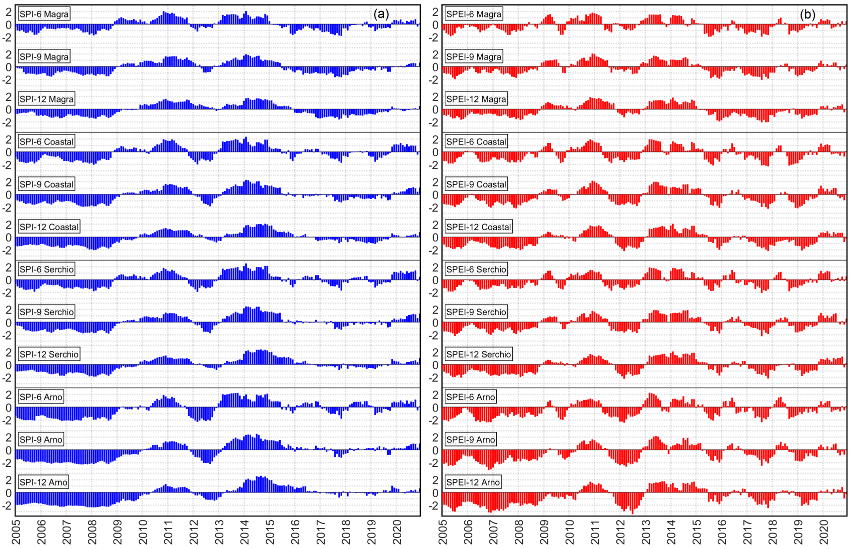


Figure 3.3. SPIs (a) and SPEIs (b) for the four analyzed basins and time windows of 6, 9 and 12 months.

The SGIs, as depicted in Fig. 3.4, detect drought periods that bear similarities to those identified by SPIs and SPEIs for nearly all wells. For selected wells, such as Bandita7 in the Magra basin, Unim in the Coastal basin, and Corte Spagni in the Arno basin, certain periods exhibit positive or slightly negative SGI values, potentially attributed to external factors influencing groundwater dynamics. For instance, the proximity of the Magra River to the Bandita7 well may impact groundwater levels, while Unim and Corte Spagni may be influenced by withdrawals from nearby well fields.

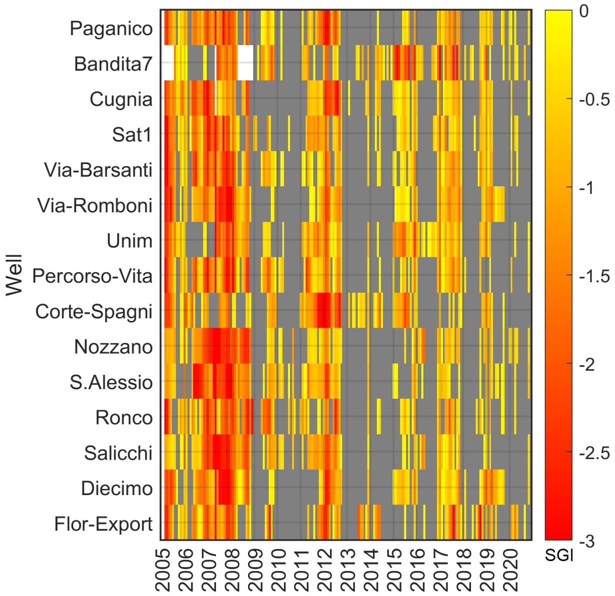
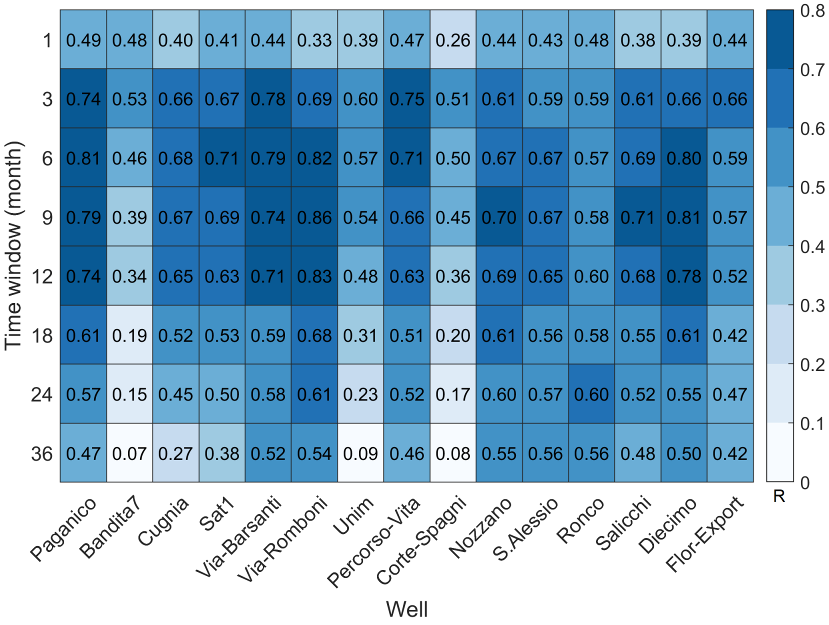


Figure 3.4. SGI values for the 15 monitoring wells used in this study. The white color indicates missing data, the grey color indicates positive values, while the color scale classifies the negative SGIs.

Meteorological and groundwater index relationships

A notable aspect to highlight is that the correlations obtained using basin-weighted SPIs and SPEIs tend to be higher compared to those calculated using indices weighted over the entire study area. This enhanced correlation aligns with findings in existing literature, which have highlighted the influence of both climate and basin characteristics on the propagation of precipitation signals to groundwater (Kumar et al., 2016).

The results in terms of SPI-SGI relationships are presented in Figure 3.5. The correlation coefficients exceeding the chosen threshold of 0.6 exhibit consistently stronger correlations with time windows of six, nine, and twelve months. This observation aligns with the expectations, given that the wells are located in shallow aquifers with moderate distances from the ground surface (Kumar et al., 2016). However, certain wells display low correlation values across all considered time windows, particularly evident for Bandita7, Unim, and Corte Spagni wells, consistent with the results reported in the previous section.



Time window (month) \ Well	Paganico	Bandita7	Cugnia	Sat1	Via-Barsanti	Via-Romboni	Unim	Percorso-Vita	Corte-Spagni	Nozzano	S.Alessio	Ronco	Salicchi	Diecimo	Flor-Export
1	0.49	0.48	0.40	0.41	0.44	0.33	0.39	0.47	0.26	0.44	0.43	0.48	0.38	0.39	0.44
3	0.74	0.53	0.66	0.67	0.78	0.69	0.60	0.75	0.51	0.61	0.59	0.59	0.61	0.66	0.66
6	0.81	0.46	0.68	0.71	0.79	0.82	0.57	0.71	0.50	0.67	0.67	0.57	0.69	0.80	0.59
9	0.79	0.39	0.67	0.69	0.74	0.86	0.54	0.66	0.45	0.70	0.67	0.58	0.71	0.81	0.57
12	0.74	0.34	0.65	0.63	0.71	0.83	0.48	0.63	0.36	0.69	0.65	0.60	0.68	0.78	0.52
18	0.61	0.19	0.52	0.53	0.59	0.68	0.31	0.51	0.20	0.61	0.56	0.58	0.55	0.61	0.42
24	0.57	0.15	0.45	0.50	0.58	0.61	0.23	0.52	0.17	0.60	0.57	0.60	0.52	0.55	0.47
36	0.47	0.07	0.27	0.38	0.52	0.54	0.09	0.46	0.08	0.55	0.56	0.56	0.48	0.50	0.42

Figure 3.5. SGI-SPI Pearson correlation coefficients.

Then, for the following analysis, only the wells with correlation coefficients exceeding the selected threshold of 0.6 within the 6-, 9-, and 12-month time windows were consider. Ten wells meet this criterion, distributed across the Arno portion (1 well), Coastal (5 wells), and Serchio (4 wells) basins.

The heat maps in Figure 3.6 summarize the analysis carried out to investigate the presence of delay (lag) between meteorological and groundwater indices. The results reveal that the highest correlation coefficient is observed with zero-lag.

This indicates that, for the study area, meteorological droughts closely align with groundwater droughts.

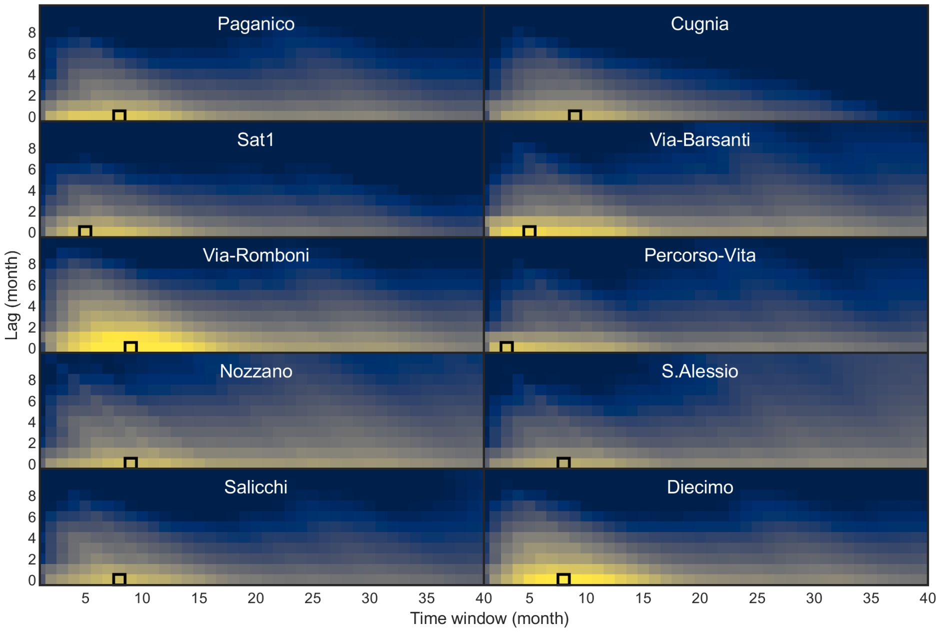


Figure 3.6. Heat maps of the SGI-SPI correlation coefficients (R) for different time windows and lags. The black box represents the highest correlation coefficient.

In the study area and for the 10 selected wells, the precipitation accumulation periods yielding the highest correlations do not exhibit significant spatial variability. For all these wells, except one, the maximum correlations are consistently observed within the six- and nine-month time windows, with correlation coefficients showing little variation within these accumulation periods. Consequently, for clarity and uniformity, the SPI with a six-month time window (hereafter referred to as SPI6) was selected for all 10 wells in the subsequent analyses.

Having established the correlation between SPIs and SGIs, the relationship between these two indices using linear regression analysis was analyzed (Fig. 3.7).

Across all wells, the slope of the regression line consistently falls below one, indicating that, in the study area, a damping mechanism is at play during the propagation process from meteorological to groundwater droughts, smoothing out negative anomalies (Van Loon et al., 2017). The dispersion around the regression line (Fig. 3.7) suggests that, as expected, factors beyond precipitation, such as lateral inflow/outflow and human activities, contribute to the drought propagation process (Wang et al., 2016). Nevertheless, the high correlation between SPIs and SGIs justifies considering this straightforward relationship for subsequent analyses.

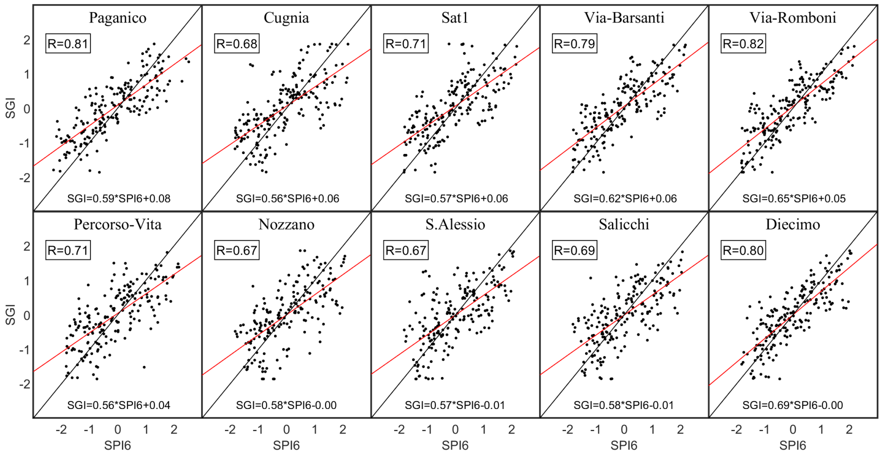
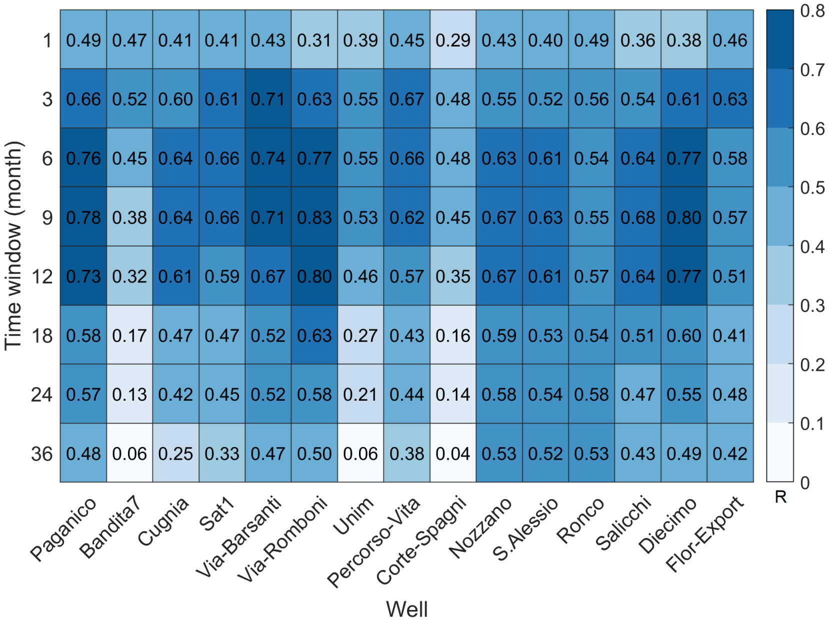


Figure 3.7. SGIs versus SPI6; the points represent the data, the red line indicates the regression line and the black line denotes the identity line. For each well, the correlation coefficient (R) and the regression equation is reported.

The same approach as described above was applied to investigate the correlations and relationships between SPEIs and SGIs. Also in this case, correlations tend to be higher for accumulation periods of 6, 9, and 12 months (Fig. 3.8). The same set of 10 wells that meet the threshold criterion for the SPI-SGI correlation analysis also fulfills the threshold condition for the SPEI-SGI relation. Generally, the correlations between SPEIs and SGIs are somewhat lower than those observed for SPIs and SGIs. In most cases, the 9-month time window yields the best results, with correlation coefficients similar to those of adjacent accumulation periods. Therefore, for clarity and consistency, further analyses were conducted using the SPEI with a 9-month time window (hereafter referred to as SPEI9), weighted according to the four basins.



Time window (month)	Paganico	Bandita7	Cugnia	Sat1	Via-Barsanti	Via-Romboni	Unim	Percorso-Vita	Corte-Spagni	Nozzano	S.Alessio	Ronco	Salicchi	Diecimo	Flor-Export
1	0.49	0.47	0.41	0.41	0.43	0.31	0.39	0.45	0.29	0.43	0.40	0.49	0.36	0.38	0.46
3	0.66	0.52	0.60	0.61	0.71	0.63	0.55	0.67	0.48	0.55	0.52	0.56	0.54	0.61	0.63
6	0.76	0.45	0.64	0.66	0.74	0.77	0.55	0.66	0.48	0.63	0.61	0.54	0.64	0.77	0.58
9	0.78	0.38	0.64	0.66	0.71	0.83	0.53	0.62	0.45	0.67	0.63	0.55	0.68	0.80	0.57
12	0.73	0.32	0.61	0.59	0.67	0.80	0.46	0.57	0.35	0.67	0.61	0.57	0.64	0.77	0.51
18	0.58	0.17	0.47	0.47	0.52	0.63	0.27	0.43	0.16	0.59	0.53	0.54	0.51	0.60	0.41
24	0.57	0.13	0.42	0.45	0.52	0.58	0.21	0.44	0.14	0.58	0.54	0.58	0.47	0.55	0.48
36	0.48	0.06	0.25	0.33	0.47	0.50	0.06	0.38	0.04	0.53	0.52	0.53	0.43	0.49	0.42

Figure 3.8. SGI-SPEI Pearson correlation coefficients.

An examination of the influence of time delays between SPEIs and SGIs showed that, once again, the maximum correlations are achieved with zero-lag for all 10 wells (figure not shown for brevity). Across all wells, the slopes of the regression lines are lower than those evaluated using SPIs, indicating a greater attenuation in drought propagation processes in the study area when considering evapotranspiration processes. Similarly, the spread around the regression line (Fig. 3.9) underscores that factors beyond precipitation also influence groundwater levels. Nevertheless, the strong correlation between SPEIs and SGIs justifies the use of this straightforward relationship for subsequent analyses.

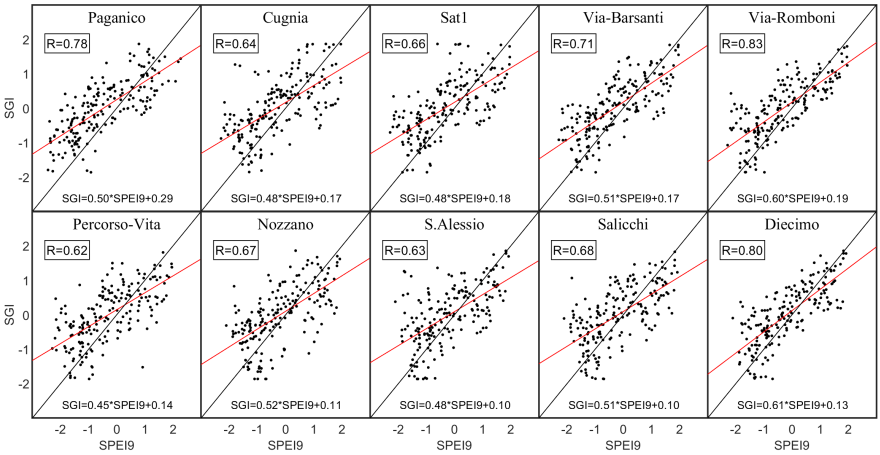


Figure 3.9. SGIs versus SPEI9; the points represent the data, the red line indicates the regression line and the black line denotes the identity line. For each well, the correlation coefficient (R) and the regression equation is reported.

Future groundwater indices

The SPI6 and SPEI9 values, derived from climate models at each station location, were subjected to averaging within each basin, under the two different RCP scenarios. These values were then employed to estimate SGIs for both historical and future periods, using the relationships depicted in Figures 3.7 and 3.9. To perform this estimation, the time series generated by the 13 RCMs were assembled to create a single dataset, assuming equal reliability among these 13 model realizations as they represent statistical outcomes of the same stochastic process. Hereafter, this dataset is referred as the "whole RCM ensemble."

As an illustrative instance, Figure 3.10 presents empirical cumulative distribution functions of SGIs for the Paganico well (located within the Arno portion basin) under the RCP8.5 scenario. The envelope curves encompassing various CDFs, corresponding to each climate model separately, demonstrate substantial uncertainty attributable to differences among individual models, with this uncertainty being more pronounced in the long-term.

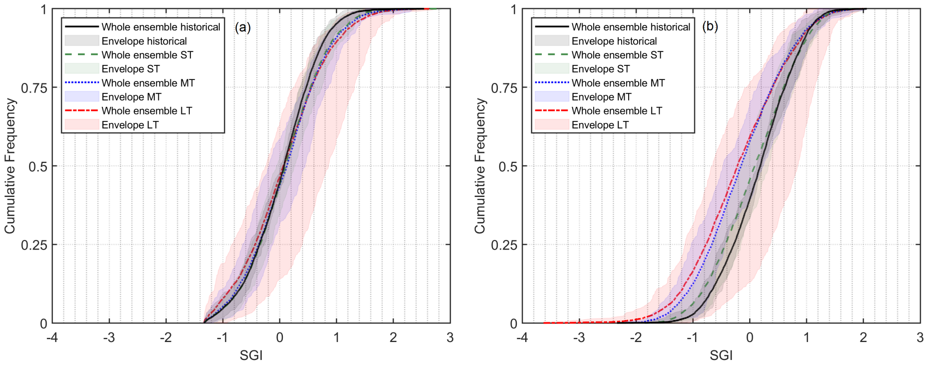


Figure 3.10. Cumulative probability distributions according to the whole RCM ensemble obtained for the Paganico monitoring well through the SGI-SPI6 (a) and the SGI-SPEI9 (b) regression equations for the historical period and at short- (ST), medium- (MT) and long-term (LT) under the RCP8.5 scenario. Envelope curves obtained by the 13 RCM models.

The results for the Paganico well are summarized in Figure 3.11 using box-whisker plots. When applying the SGI-SPI6 regression relationships, no significant alterations are apparent between the historical and future periods, with the median value remaining close to zero throughout all periods. Notably, there are positive outliers due to the outcomes of a specific model that predicts abundant future precipitation, in contrast to the other RCMs. Conversely, employing the SGI-SPEI9 regression relationships reveals a systematic reduction in SGIs, particularly at medium- and long-term projections. Accounting for temperature, the effects of the model projecting abundant precipitation are mitigated, and there is an increase in negative outliers.

To quantify these outcomes across all observation wells, characteristic SGI values derived from the SGI-SPI6 and SGI-SPEI9 regression relationships are presented in Tables 3.4 and 3.5. Under the SGI-SPI6 relationships, a minor decrease in SGIs is evident in the medium-term for RCP4.5 and in the long-term for RCP8.5 when examining the 25th percentile and median values. Conversely, when utilizing the SGI-SPEI9 relationships, substantial reductions in future SGIs are observed across most wells. For RCP4.5, the medium-term period experiences the most significant declines, while for RCP8.5, the most critical groundwater level conditions are anticipated in the long-term. These detected changes exhibit similar characteristics across all wells, particularly within the same basin.

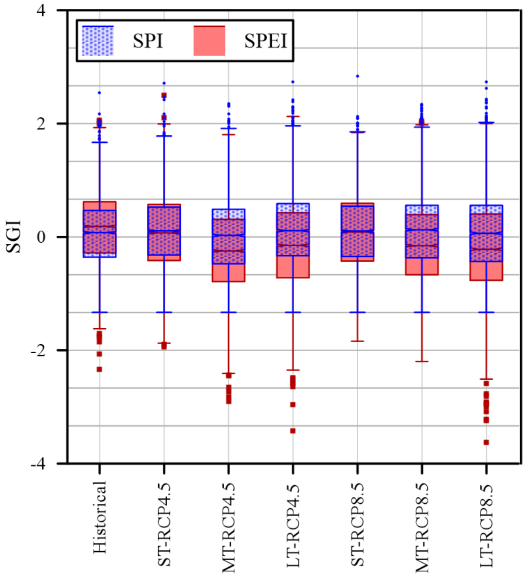


Figure 3.11. Box-plots of the SGIs obtained for the Paganico monitoring well. The SGIs were calculated according to the whole RCM, through the SGI-SPI6 and SGI-SPEI9 regression equations for the historical period and at short-term (ST), medium-term (MT), and long-term (LT) under the two RCP scenarios. The boxplot marks points as outliers if they are greater than the mean ± 2.7σ, where σ is the standard deviation.

Table 3.4. Differences of the median, 25th and 75th percentiles of the future SGIs with those evaluated in the historical period. The SGIs are defined through the SGI-SPI6 regression relationships for the historical period and at short- (ST), medium- (MT) and long-term (LT), under the RCP4.5 and RCP8.5.

		SGI-SPI6						
		Historical	Differences with the historical period					
			RCP4.5			RCP8.5		
			ST	MT	LT	ST	MT	LT
Paganico	25th	-0.36	0.04	-0.12	0.03	0.01	-0.01	-0.07
	50th	0.08	0.03	-0.05	0.03	0.03	0.05	-0.01
	75th	0.47	0.06	0.02	0.12	0.07	0.09	0.09
Cugnia	25th	-0.33	-0.08	-0.21	-0.12	-0.09	-0.12	-0.17
	50th	0.01	0.03	-0.04	0.04	0.01	0.02	-0.01
	75th	0.33	0.15	0.13	0.20	0.15	0.15	0.17
SAT 1	25th	-0.41	0.01	-0.12	-0.03	-0.01	-0.04	-0.08
	50th	0.01	0.04	-0.03	0.05	0.02	0.02	-0.01
	75th	0.41	0.07	0.06	0.13	0.08	0.08	0.10
Via Barsanti	25th	-0.46	0.01	-0.13	-0.04	-0.01	-0.04	-0.09
	50th	0.00	0.04	-0.03	0.05	0.02	0.02	-0.01
	75th	0.44	0.08	0.06	0.14	0.09	0.09	0.11
Via Romboni	25th	-0.49	-0.15	-0.05	-0.01	-0.06	-0.11	-0.06
	50th	0.03	-0.08	0.01	0.02	-0.02	-0.05	-0.04
	75th	0.53	-0.02	0.07	0.04	0.01	0.03	0.01
Percorso Vita	25th	-0.43	0.01	-0.12	-0.03	-0.01	-0.04	-0.08
	50th	-0.01	0.04	-0.03	0.05	0.02	0.02	-0.01
	75th	0.39	0.07	0.06	0.13	0.08	0.08	0.10
Nozzano	25th	-0.40	0.01	-0.14	-0.04	-0.01	-0.04	-0.12
	50th	-0.01	0.02	-0.05	0.01	0.01	-0.01	-0.05
	75th	0.35	0.04	0.02	0.07	0.06	0.04	0.03
S. Alessio	25th	-0.38	0.01	-0.14	-0.04	-0.02	-0.04	-0.12
	50th	0.03	-0.01	-0.08	-0.02	-0.03	-0.04	-0.08
	75th	0.40	-0.02	-0.03	0.02	0.00	-0.01	-0.02
Salicchi	25th	-0.39	0.01	-0.14	-0.04	-0.01	-0.04	-0.12
	50th	0.00	0.02	-0.05	0.01	0.00	-0.01	-0.05
	75th	0.36	0.04	0.02	0.07	0.06	0.04	0.03
Diecimo	25th	-0.47	0.01	-0.16	-0.04	-0.02	-0.05	-0.14
	50th	-0.01	0.03	-0.06	0.01	0.01	-0.01	-0.06
	75th	0.42	0.04	0.03	0.09	0.07	0.05	0.04

Table 3.5. Differences of the median, 25th and 75th percentiles of the future SGIs with those evaluated in the historical period. The SGIs are defined through the SGI-SPEI9 regression relationships for the historical period and at short- (ST), medium- (MT) and long-term (LT), under the RCP 4.5 and RCP 8.5

		SGI-SPEI9						
		Historical	Differences with the historical period					
			RCP4.5			RCP8.5		
			ST	MT	LT	ST	MT	LT
Paganico	25th	-0.28	-0.13	-0.51	-0.44	-0.15	-0.39	-0.49
	50th	0.19	-0.12	-0.43	-0.34	-0.10	-0.34	-0.40
	75th	0.62	-0.05	-0.31	-0.20	-0.03	-0.23	-0.22
Cugnia	25th	-0.32	-0.08	-0.29	-0.25	-0.08	-0.21	-0.27
	50th	0.02	-0.04	-0.21	-0.16	-0.03	-0.17	-0.19
	75th	0.38	0.01	-0.14	-0.07	0.01	-0.13	-0.10
SAT 1	25th	-0.30	-0.08	-0.28	-0.24	-0.08	-0.21	-0.27
	50th	0.02	-0.04	-0.21	-0.16	-0.03	-0.17	-0.19
	75th	0.37	0.00	-0.14	-0.07	0.00	-0.13	-0.10
Via Barsanti	25th	-0.10	-0.08	-0.29	-0.25	-0.08	-0.22	-0.28
	50th	0.25	-0.04	-0.22	-0.17	-0.03	-0.17	-0.20
	75th	0.62	0.00	-0.14	-0.07	0.00	-0.13	-0.11
Via Romboni	25th	-0.13	-0.10	-0.35	-0.29	-0.10	-0.25	-0.33
	50th	0.28	-0.05	-0.26	-0.20	-0.04	-0.21	-0.23
	75th	0.71	0.01	-0.17	-0.09	0.01	-0.16	-0.13
Percorso Vita	25th	-0.17	-0.07	-0.27	-0.23	-0.07	-0.20	-0.25
	50th	0.14	-0.04	-0.20	-0.15	-0.03	-0.16	-0.18
	75th	0.47	0.00	-0.13	-0.07	0.00	-0.12	-0.10
Nozzano	25th	-0.18	-0.09	-0.37	-0.30	-0.11	-0.27	-0.36
	50th	0.18	-0.06	-0.26	-0.20	-0.05	-0.22	-0.26
	75th	0.53	-0.03	-0.17	-0.10	0.00	-0.15	-0.15
S. Alessio	25th	-0.17	-0.08	-0.33	-0.27	-0.10	-0.24	-0.32
	50th	0.15	-0.06	-0.23	-0.18	-0.04	-0.19	-0.23
	75th	0.47	-0.02	-0.15	-0.09	0.00	-0.13	-0.13
Salicchi	25th	-0.06	-0.21	-0.47	-0.41	-0.23	-0.39	-0.47
	50th	0.28	-0.17	-0.36	-0.31	-0.15	-0.32	-0.36
	75th	0.36	0.60	-0.12	-0.26	-0.20	-0.10	-0.24
Diecimo	25th	-0.21	-0.11	-0.42	-0.35	-0.13	-0.32	-0.42
	50th	0.20	-0.07	-0.30	-0.23	-0.05	-0.25	-0.30
	75th	0.61	-0.03	-0.19	-0.12	0.00	-0.17	-0.17

3.1.5. Comparison of three different artificial neural network techniques

In this part of the thesis, the primary objective is to bridge an existing gap in research by investigating the application of deep learning models to evaluate the effects of climate change on future groundwater levels. The study compares three distinct Artificial Intelligence (AI) techniques 1.3: a machine learning approach referred to as Non-Linear Autoregressive Neural Network (NARX), alongside two deep learning methods, the Long Short-Term Memory neural network (LSTM) and Convolutional Neural Network (CNN). MATLAB is the environment for building and training these models.

The ultimate outputs of all these models are predictions of groundwater levels. The training and validation processes involve historical meteorological and groundwater level time series data. Once trained, these models utilize climate model projections to estimate future groundwater levels.

One of the primary contributions of this study lies in the comparative analysis of various AI techniques for assessing the influence of climate change on groundwater resources. Particular emphasis is placed on the implementation of the LSTM neural network, a sequential deep learning method, within this specific context. The study focuses into the use of neural networks for long-term forecasting, enabling the models to make predictions beyond the range of data encountered during the training phase. Another noteworthy aspect is the extensive ensemble of regional climate projections under different scenarios, which offers a comprehensive understanding of the associated uncertainty in the outcomes of the study.

3.1.5.1. Set up of the surrogate model

To establish a benchmark, the 10 wells (Table 3.6) that exhibited strong correlations with the meteorological indices, as detailed in the previous work, were chosen.

Table 3.6. ID, name, reference groundwater body, percentage of available data and ground elevation of the monitoring wells considered.

ID	**Name**	**Groundwater body**	**% data**	**Elevation m a.s.l.**
W3	Cugnia	33TN010	91.7	4.00
W4	Diecimo	12SE020	60.9	65.00
W6	Nozzano	12SE011	78.6	16.43
W7	Paganico	11AR028	72.4	13.00
W8	Percorso vita	33TN010	78.1	1.56
W10	Salicchi	12SE011	83.3	27.12
W11	S.Alessio	12SE011	71.9	18.87
W13	Sat 1	33TN010	75.5	1.50
W15	Via Barsanti	33TN010	91.7	20.00
W16	Via Romboni	33TN010	88.0	37.92

For all neural networks utilized in this study, the exogenous input data comprise monthly precipitation and mean monthly temperature measurements from the 18 climate stations, as listed in Table 3.2. This totals to 36 features considered in the input data. The output dataset corresponds to monthly groundwater levels at the 10 wells detailed in Table 3.6, resulting in 10 responses. The target data used for constructing the models encompass the period from March 2005 to December 2020, spanning 190 months during which observed monthly groundwater levels are available. The precipitation and temperature data, used as meteorological variables, extend from July 2004 to December 2020, covering a total of 198 months. This extended time frame accounts for potential delayed responses of groundwater level data to meteorological variables.

To facilitate the parameter update process (involving weights and biases), both the input and target datasets underwent standardization using the following formula:

$$Z_i = \frac{z_i - \bar{z}}{\sigma_z} \quad . \tag{3.1}$$

Here, z_i represents an individual data point within the input or target vector, $\bar{z}$ is the arithmetic mean, and σ_z is the standard deviation.

AI models: training, validation, test and evaluation metrics

The training and validation phases of the neural networks were based on a dataset that encompasses the years from 2005 to 2018, partitioned into a training set (90%) and a validation set (10%). The subsequent test phase covered the years 2019 and 2020.

During the training process, the network aimed to minimize the Loss function (Eq. 1.4). Then, the trained AI models were employed to predict groundwater levels for the period spanning from 1976 to 2095, using precipitation and temperature data supplied by the 13 climate models as input.

The Mean Squared Error (MSE) was employed to assess the performance of the developed machine learning models during the training, validation, and test phases:

$$MSE = \frac{\sum_{i=1}^{N_d}(W_i - \hat{W}_i)^2}{N_d} \tag{3.2}$$

Where W_i represents the actual value, $\hat{W}_i$ denotes the corresponding value estimated by the neural networks, and N_d stands for the number of observations within the training, validation, testing, or the entire dataset.

Furthermore, for each well, the Root Mean-Square Error (RMSE) between predicted and observed groundwater levels in the test period was calculated.

For each neural network, hyperparameters and activation functions were manually adjusted to identify the architecture with the minimum Loss function value while managing computational costs during training.

NARX

The operational framework of the NARX model was previously introduced in Section 1.3.3. In this study, the neural network training parameters were configured as follows: 50 training epochs, a mini-batch size of 18, and a learning rate of 0.001. The training phase employed historical precipitation and temperature data alongside observed groundwater levels in an open-loop mode. Then, the NARX model utilized precipitation and temperature projections from the ensemble of climate models as input for estimating future groundwater levels. During the prediction phase, the network adopted a closed-loop scheme.

The input delay parameter, denoted as n was set to 9, consistent with the research findings of Section 3.1.2, where the maximum correlation between SGI and SPEI occurred for an accumulation period of 9 months. Consequently, to predict groundwater levels at time $t + 1$, the model considered precipitation and temperature data spanning from time $t - 8$ to time t as exogenous inputs.

Feedback delays, represented by m were configured as 2, aligning with the rapid response time of the aquifer. This setting assumed that the predicted levels at time $t + 1$ were related to the groundwater levels at times t and $t - 1$.

At each time step, the input matrix had dimensions of 36×9, where 9 accounted for the monthly values due to the input delay, and 36 represented the features corresponding to the monthly values of precipitation and temperature obtained from the 18 climate stations.

The objective function is minimized using the Levenberg–Marquardt algorithm (Hagan and Menhaj, 1994). The neural network incorporated a single hidden layer comprising 10 neurons, while the output vector size was set to 10 to match the number of wells. The total number of time steps considered amounted to 190, aligning with the length of the groundwater level time series.

LSTM

The deep learning structure of the Long Short-Term Memory networks (LSTMs) starts with the initial layer, known as the Sequence Input Layer. It serves as the entry point for sequential data into the neural network. Then, the LSTM layer, positioned as a central component within the deep learning architecture, assumes a pivotal role in the sequence processing chain, embodied by the LSTM cell unit (Fig. 1.5). For a comprehensive understanding of the LSTM cell intricacies, the readers are encouraged to refer to Section 1.3.4 for more exhaustive information and insights.

In this study, several hyperparameters were configured: the number of training epochs was set to 100, the mini-batch size was established as 18, the initial learning rate was fixed at 0.005, the learning rate was reduced every 40 epochs with a reduction factor of 0.1. The dimension of the sequence matrix was defined as 190×1 of data type cell. Each row in the matrix corresponds to an input block comprised of a 36×9 matrix, where rows denote the features (such as precipitation and temperature from 18 climate stations), and columns represent the length of the input sequence. The number of hidden units was set at 100. At each time step, an output vector of dimension 10, corresponding to the number of wells, is generated. The objective function is minimized using a widely recognized backpropagation algorithm known as "Adam," as extensively documented in the literature (Kingma and Ba, 2015).

CNN

The operational framework of the CNN model was previously introduced in Section 1.3.5. A schematic representation of the CNN for a single image is depicted in Figure 1.6.

The initial layer is designated as the Image Input Layer, serving as the entry point for input image data into the neural network. Then, this data flows into a structural unit composed of multiple layers, including the Convolutional Layer, Batch Normalization Layer, ReLu Layer, and Average Pooling 2D Layer. Notably, three of these distinct structural units are sequentially arranged in the implemented model.

Each input image is characterized by dimensions of 36×9×1, where 36 represents the key features derived from monthly precipitation and temperature data gathered from 18 climate stations. The value 9 corresponds to the temporal window selected to curate the input data, while 1 signifies the number of channels in the image. For example, the initial image, utilized to predict groundwater levels for March 2005, encompasses precipitation and temperature data spanning from June 2004 to February 2005. Subsequently, the second image, associated with groundwater levels in April 2005, includes climate variables recorded from July 2004 to March 2005.

The Convolutional Layers maintain a consistent filter height and width of 5 across all three layers. The number of filters is distributed as 8, 16, and 32 for the first, second, and third blocks, respectively. The Batch Normalization Layer operates by independently normalizing a mini-batch of extracted features for each channel, an approach designed to expedite the training process of the CNN and reduce its sensitivity to initial parameter values.

The ReLu Layer employs the Rectified Linear Unit Function to activate the signal within the network. Following this, the normalized and activated feature maps in the Average Pooling 2D Layer are partitioned into non-overlapping zones. These zones have a height and width of 2, and a stride of 2 is employed to extract the average value from each zone.

At each temporal step, the network generates an output vector with dimensions of 10, corresponding to the number of wells under consideration.

For the sake of explanation, Figure 3.12 includes a depiction of the raw map, which comprises standardized input data and serves as the foundation for generating all input images. Precipitation is represented by values on the Y-axis ranging from position 1 to 18, while temperature is depicted by values from position 19 to 36. This map notably illustrates the significant variability in precipitation values, reflecting natural fluctuations in weather patterns. In contrast, temperature values, exhibit a prominent seasonal pattern. The red rectangle demarcates the dimensions of the image, and a sequential one-step scrolling process is executed to create a total of 190 images.

Once fully trained, the neural network is deployed to estimate groundwater levels based on the input of precipitation and temperature data provided by the 13 distinct climate models. The Adam algorithm is used to minimize the loss function.

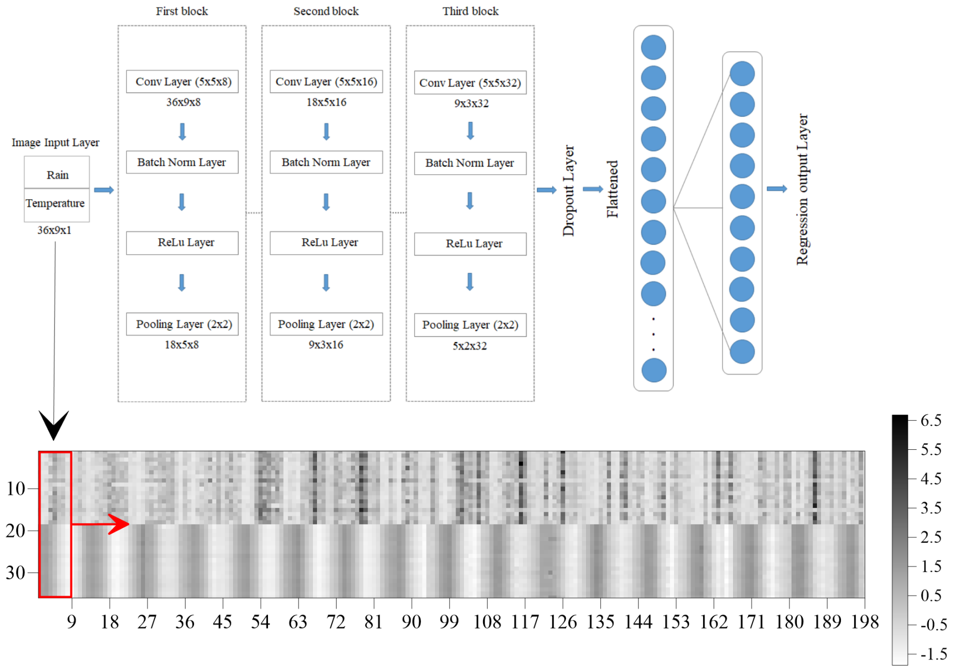


Figure 3.12. Schematic view of the CNN (top) and sample of the standardized input data (bottom).

3.1.5.2. Results

Evaluation of the performance

Table 3.7 presents the MSEs computed on the entire dataset, as well as for the training, validation, and testing subsets for all proposed AI methods. For the NARX, the overall performance of the network is good, with an MSE of 0.17 m^2 on the entire dataset. The training phase shows an MSE of 0.04 m^2, and the validation phase has an MSE of 0.54 m^2, while the test phase has an MSE of 0.82 m^2. These results suggest a potential issue with overfitting, where the network performs well on the training data but struggles to generalize to data beyond the training range.

Regarding the LSTM, the overall MSE performance is 0.14 m^2. The training MSE is 0.12 m^2, the validation phase yields an MSE of 0.23 m^2, and the test phase reports an MSE of 0.30 m^2, indicating good performance across all phases. Unlike the NARX, the Dropout Layer in the LSTM allows for slightly less learning during training, promoting better generalization.

The CNN achieves an overall performance of 0.10 m^2, with a training MSE of 0.05 m^2. The validation phase results in an MSE of 0.20 m^2, and the test phase has an MSE of 0.31 m^2, comparable to the LSTM's performance.

Table 3.7. MSE (m^2) between the output of the neural networks (NARX, LSTM, CNN) and the observed groundwater levels.

	NARX	**LSTM**	**CNN**
Training	0.04	0.12	0.05
Validation	0.54	0.23	0.20
Test	0.82	0.30	0.31
Whole dataset	0.17	0.14	0.10

The Table 3.8 provides a comparison between observed and predicted groundwater levels, in terms of RMSE values, for the 10 wells studied, during both the learning, which includes training and validation, and testing phases, using the three implemented AI models. Overall, the models demonstrate satisfactory performance metrics (RMSE less than 0.4 m) for both learning and testing phases, except for the Via Romboni well, which exhibits higher errors (greater than 0.8 m) across all three networks during testing. In the testing phase, the well Nozzano performs well only with the LSTM network, while the wells S. Alessio and Diecimo show unsatisfactory metrics for the NARX. In general, the LSTM appears to be more suitable for predicting groundwater levels. As an illustrative example, Figure 3.13 compares predicted and observed groundwater levels in the test period 2019-2020 for the Paganico well, highlighting the superior performance of the LSTM.

Table 3.8. RMSE (m) between the output of the neural networks (NARX, LSTM, CNN) and the observed groundwater levels in the learning and testing period 2005-2018.

Well	**NARX**		**LSTM**		**CNN**	
	Learning	Testing	Learning	Testing	Learning	Testing
Paganico	0.10	0.37	0.11	0.17	0.15	0.30
Cugnia	0.07	0.35	0.08	0.10	0.16	0.25
Sat 1	0.11	0.36	0.09	0.11	0.14	0.26
Via Barsanti	0.10	0.27	0.11	0.14	0.10	0.29
Via Romboni	0.30	1.46	0.42	0.83	0.48	0.98
Percorso Vita	0.05	0.17	0.07	0.12	0.06	0.23
Nozzano	0.18	0.93	0.23	0.36	0.23	0.66
S. Alessio	0.06	0.59	0.11	0.18	0.17	0.27
Salicchi	0.07	0.43	0.10	0.17	0.11	0.38
Diecimo	0.13	0.57	0.21	0.28	0.16	0.39

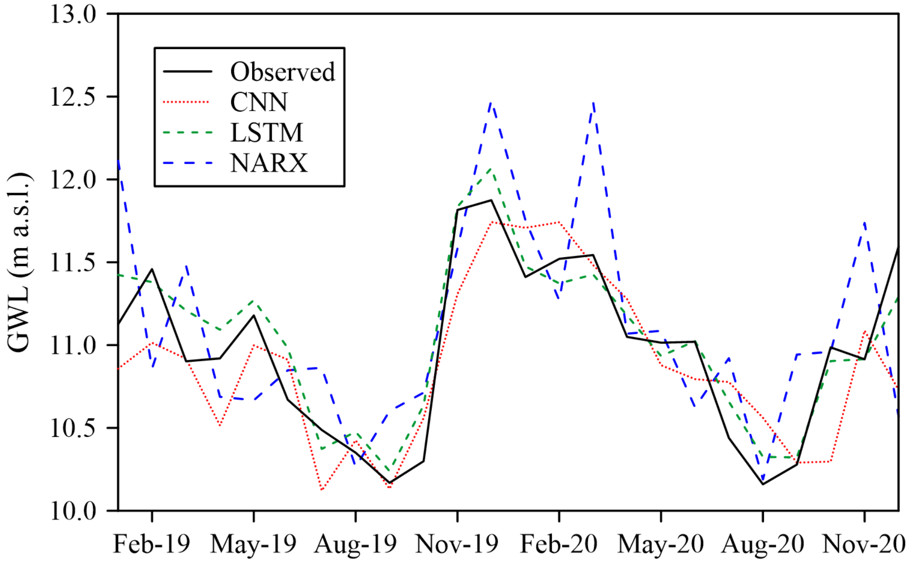


Figure 3.13. Observed and predicted groundwater levels for the testing phase (period 2019-2020) for the well Paganico.

Future projections

The trained neural networks were applied in conjunction with climate model projections to forecast groundwater levels under the two distinct scenarios, RCP4.5 and RCP8.5, spanning from 1976 to the end of the current century.

To emphasize the evolving trends in piezometric levels over time, the outcomes for the Paganico well in April are presented. April was chosen due to its usual minimal anthropogenic influences, occurring before the commencement of the irrigation withdrawal period. Predicted groundwater levels for the entire simulated period (1976-2095) are illustrated using a 10-year moving average, which helps elucidate the impacts of climate change on natural variability. The outcomes are reported in terms of the median value and minimum and maximum range derived from the ensemble of climate models.

Figure 3.14 showcases the results obtained with the NARX model. For both

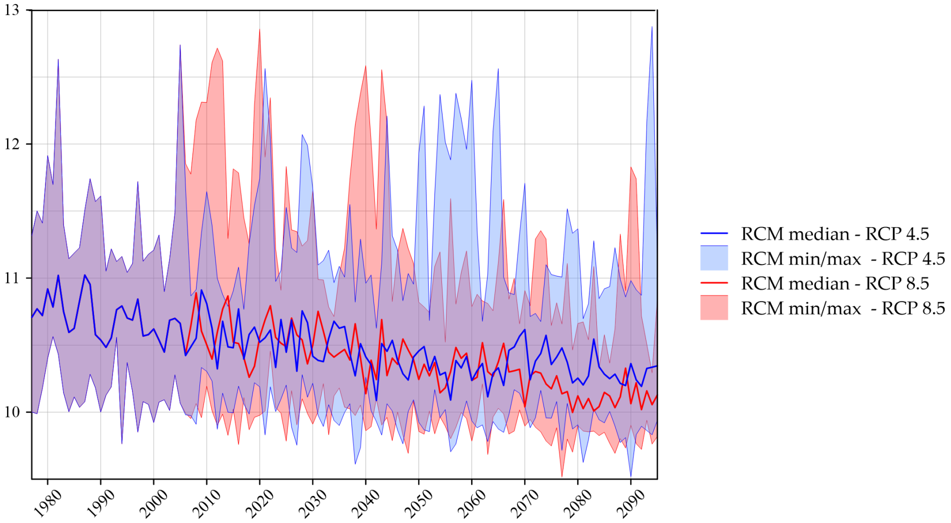


Figure 3.14. Predicted groundwater levels with the NARX in April for the Paganico well in terms of 10-year moving average under the RCP4.5 and RCP8.5 scenarios.

RCP scenarios, a discernible reduction in groundwater levels over time is evident based on the median values. This decline is projected to be more pronounced under RCP 8.5 in the long term. This phenomenon can be attributed to the higher greenhouse gas concentrations associated with RCP 8.5, leading to more substantial temperature increases and subsequent elevated evapotranspiration rates, which in turn impact recharge processes. Conversely, RCP4.5 anticipates the implementation of mitigation measures to curb greenhouse gas emissions. The max/min range, which considers different regional climate models, underscores the uncertainties inherent in future predictions.

To elucidate temporal changes, we analyzed results across four distinct climate model periods: historical (1976-2005), short-term (2006-2035), medium-term (2036-2065), and long-term (2066-2095).

Figure 3.15 presents empirical CDFs of standardized groundwater levels pre-

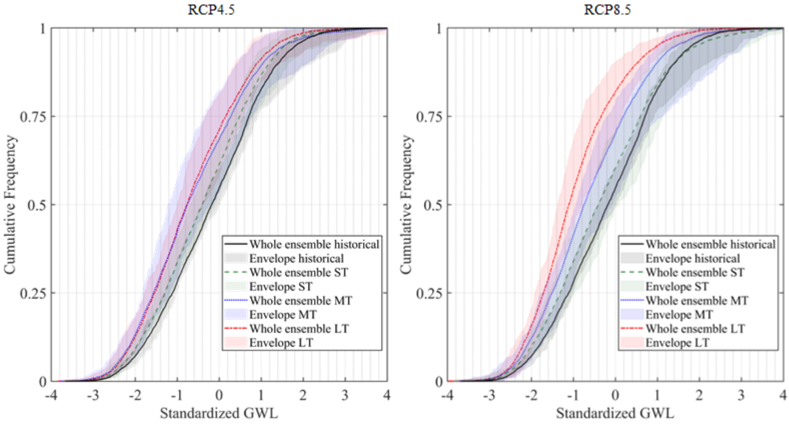


Figure 3.15. Cumulative distribution probability functions according to the whole RCM ensemble obtained with the NARX for the Paganico well for the historical period and at short- (ST), medium- (MT), and long-term (LT) under the RCP4.5 (left) and RCP8.5 (right) scenarios, along with the envelope curves provided by the 13 RCMs.

dicted over these four projection periods for the Paganico well, under both RCP4.5 and RCP8.5 scenarios. The envelope curves represent individual climate models, emphasizing prediction uncertainties. Additionally, data from 13 climate models were aggregated into a single dataset referred to as the "whole RCM ensemble". The leftward shift of the CDFs from the historical one highlights increased likelihood or frequency of lower values. While RCP4.5 shows similar results in the medium and long term, RCP8.5 indicates a more substantial decrease in negative standardized groundwater levels over the long term. Concerning this, for both scenarios, a clear decline in standardized groundwater levels in the future compared to the historical period is discernible, especially in the medium and long term.

Figures 3.16 and 3.17 illustrate results obtained with the LSTM model. Looking at the 10-year moving average of predicted groundwater levels in April (Figure 3.16), both RCP scenarios exhibit median values indicating a moderate decline in piezometric levels over time, with a slight pronounced trend for RCP8.5 toward

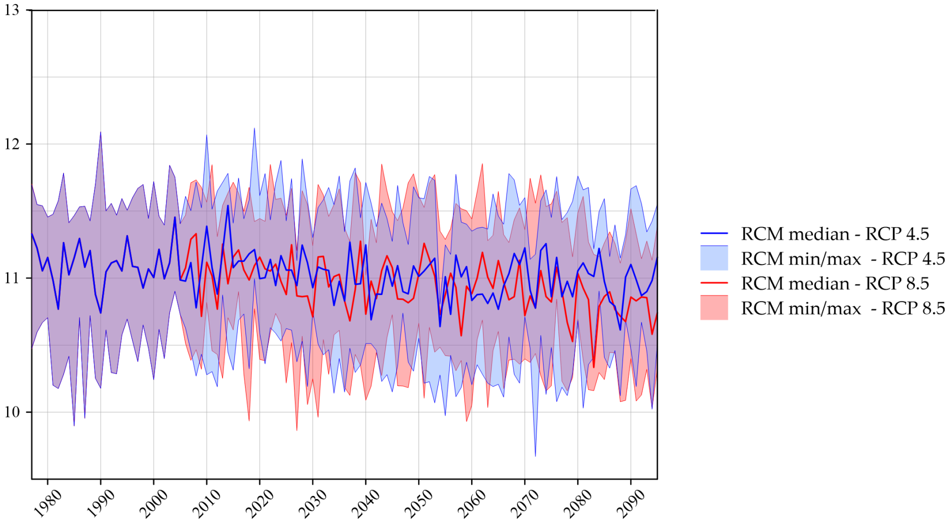


Figure 3.16. Predicted groundwater levels with the LSTM in April for the Paganico well in terms of 10-year moving average under the RCP4.5 and RCP8.5 scenarios.

the end of the simulation period. The variability, represented by the maximum and minimum range between climate models, is similar to that observed with the NARX.

CDFs of predicted standardized GWLs (Figure 3.17) indicate similar trends in the medium and long term under RCP4.5, with slight deviations when compared to the historical period. These deviations are minimal in the short term. Conversely, RCP8.5 shows a progressive reduction in standardized GWLs over time, particularly in the long term. Furthermore, envelope curves reveal lower uncertainty in future GWL estimates using LSTM compared to NARX. Overall, LSTM predicts a smaller reduction in standardized groundwater levels compared to NARX.

Figures 3.18 and 3.19 present results obtained with the CNN model. Figure 3.18 reveals no systematic trends in April groundwater levels under both RCP scenarios, with lower variability between RCMs compared to NARX and LSTM.

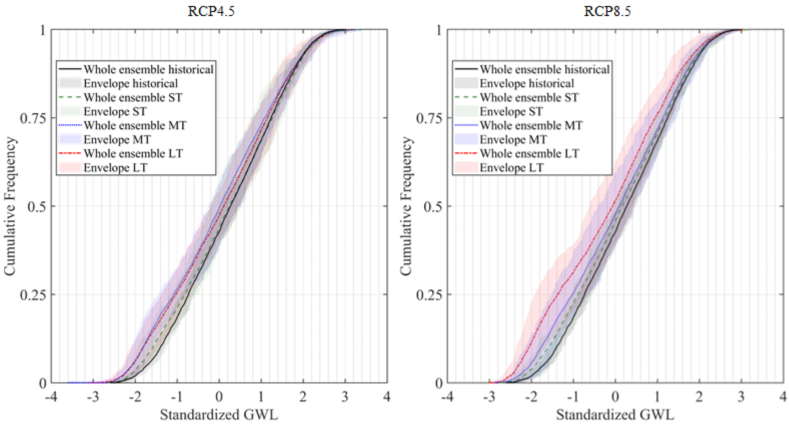


Figure 3.17. Cumulative distribution probability functions according to the whole RCM ensemble obtained with the LSTM for the Paganico well for the historical period and at short- (ST), medium- (MT), and long-term (LT) under the RCP4.5 (left) and RCP8.5 (right) scenarios, along with the envelope curves provided by the 13 RCMs.

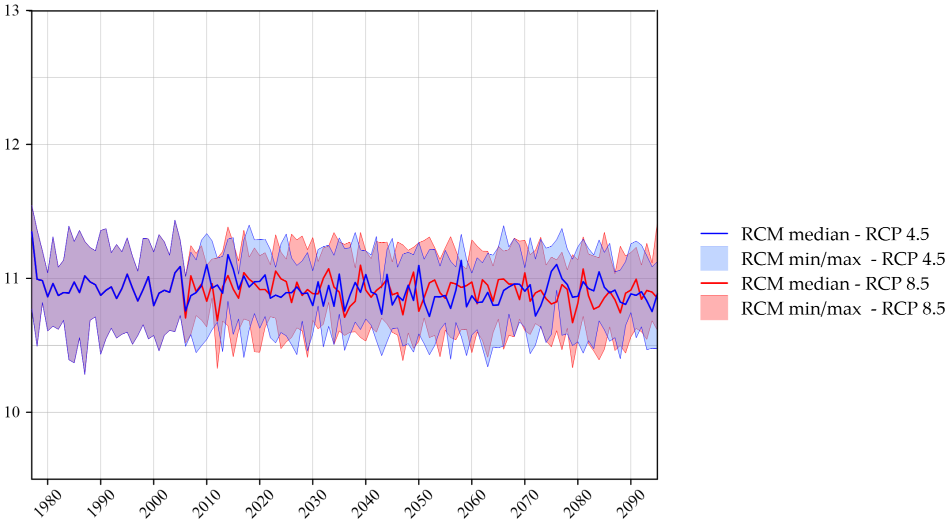


Figure 3.18. Predicted groundwater levels with the CNN in April for the Paganico well in terms of 10-year moving average under the RCP4.5 and RCP8.5 scenarios.

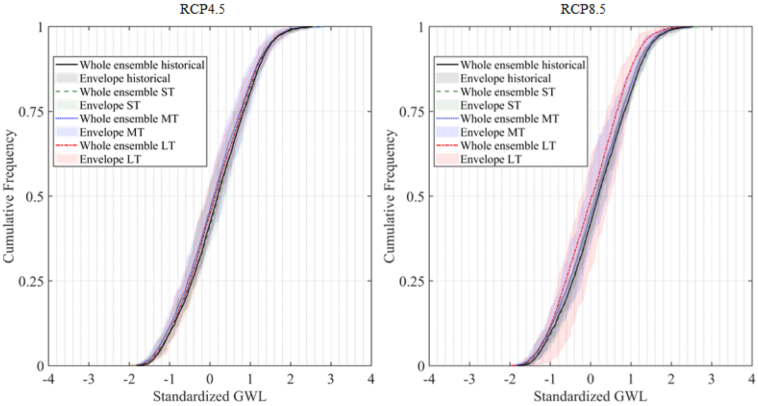


Figure 3.19. Cumulative distribution probability functions according to the whole RCM ensemble obtained with the CNN for the Paganico well for the historical period and at short- (ST), medium- (MT), and long-term (LT) under the RCP4.5 (left) and RCP8.5 (right) scenarios, along with the envelope curves provided by the 13 RCMs.

According to Figure 3.19, CDFs of future GWLs for the Paganico well do not exhibit significant changes compared to the historical period for both RCP scenarios. Only a slight decrease in standardized GWLs is detectable in the long term under RCP8.5. This indicates a distinct behavior of the CNN compared to the other two AI models.

To quantify results for all wells and make comparisons with the outcomes of Section 3.1.4, Tables 3.9 and 3.10 depict differences in standardized GWLs, evaluated by means of the statistical approach and the three ANNs, between future periods (short-term, medium-term and long-term) under RCP4.5 and RCP8.5 and the historical period. The 25th, 50th, and 75th percentiles of the whole RCM ensemble were used for these calculations. Color intensity highlights the magnitude of differences in standardized GWLs between future and historical periods. Overall, CNN exhibits smaller differences compared to the other two AI models. For instance, for the Paganico well under RCP8.5, the difference in standardized GWL median values between the long term and the historical period is -0.93 for NARX, -0.34 for LSTM, and only -0.07 for CNN (Table 3.10). The regression model previously presented in Section 3.1.2 generally predicts slightly greater reductions in groundwater levels compared to CNN. Conversely, LSTM shows more pronounced declines in future levels, with few exceptions. NARX predicts the most severe declines in future groundwater levels but has the highest MSEs poorest performance in the test phase (Table 3.8).

Table 3.9. Differences between the 25th, 50th and 75th percentiles of the future standardized GWLs at short- (ST), medium- (MT), and long-term (LT) and the historical ones under the RCP4.5 scenario. Results obtained with the AI models proposed in this study and the regression model presented in Section 3.1.2.

		REGRESSION			NARX			LSTM			CNN		
		ST	MT	LT	ST	MT	LT	ST	MT	LT	ST	MT	LT
Paganico	25th	-0.04	-0.28	-0.23	-0.19	-0.46	-0.43	-0.08	-0.35	-0.30	0.05	-0.01	0.05
	50th	-0.08	-0.38	-0.32	-0.22	-0.58	-0.59	-0.03	-0.24	-0.16	0.03	-0.08	-0.02
	75th	-0.08	-0.36	-0.29	-0.22	-0.43	-0.55	0.02	-0.16	-0.10	0.01	-0.06	-0.04
Cugnia	25th	0.01	-0.13	-0.11	0.14	0.01	0.03	-0.08	-0.32	-0.21	0.01	-0.10	-0.04
	50th	-0.03	-0.20	-0.17	0.15	0.05	0.06	-0.04	-0.19	-0.12	0.00	-0.11	-0.07
	75th	-0.07	-0.22	-0.19	0.25	0.19	0.22	0.02	-0.14	-0.05	-0.03	-0.11	-0.08
SAT 1	25th	0.01	-0.13	-0.11	-0.27	-0.50	-0.46	-0.23	-0.57	-0.61	0.04	-0.08	-0.02
	50th	-0.03	-0.20	-0.17	-0.18	-0.40	-0.46	-0.11	-0.34	-0.33	0.02	-0.08	-0.04
	75th	-0.07	-0.22	-0.19	-0.12	-0.29	-0.37	-0.05	-0.24	-0.21	-0.04	-0.10	-0.08
Via Barsanti	25th	0.03	-0.16	-0.11	-0.18	-0.45	-0.41	-0.12	-0.38	-0.33	0.00	-0.10	-0.05
	50th	-0.02	-0.21	-0.16	-0.09	-0.27	-0.27	-0.05	-0.27	-0.19	0.02	-0.10	-0.06
	75th	-0.05	-0.19	-0.16	-0.02	-0.12	-0.12	0.01	-0.18	-0.12	0.00	-0.09	-0.05
Via Romboni	25th	0.02	-0.20	-0.16	-0.17	-0.57	-0.61	-0.11	-0.38	-0.30	-0.04	-0.20	-0.17
	50th	-0.04	-0.25	-0.20	-0.09	-0.51	-0.51	-0.04	-0.21	-0.17	-0.02	-0.18	-0.14
	75th	-0.06	-0.20	-0.17	0.04	-0.25	-0.25	-0.02	-0.12	-0.10	-0.04	-0.10	-0.07
Percorso Vita	25th	0.01	-0.13	-0.11	-0.19	-0.51	-0.45	-0.10	-0.39	-0.35	0.03	-0.04	0.02
	50th	-0.02	-0.17	-0.14	-0.13	-0.36	-0.35	-0.06	-0.26	-0.19	0.02	-0.09	-0.03
	75th	-0.07	-0.20	-0.19	-0.14	-0.25	-0.32	0.02	-0.19	-0.12	-0.04	-0.11	-0.07
Nozzano	25th	0.01	-0.18	-0.14	-0.45	-0.90	-0.94	-0.01	-0.12	-0.02	-0.06	-0.12	-0.13
	50th	-0.04	-0.25	-0.20	-0.41	-0.82	-0.91	0.01	-0.16	-0.01	-0.05	-0.18	-0.16
	75th	-0.07	-0.24	-0.20	-0.30	-0.66	-0.75	0.10	-0.02	0.06	-0.05	-0.12	-0.12
S. Alessio	25th	-0.01	-0.16	-0.13	-0.39	-0.74	-0.67	-0.01	-0.15	-0.05	0.01	0.00	0.05
	50th	-0.04	-0.21	-0.18	-0.43	-0.74	-0.73	0.03	-0.14	0.01	0.02	-0.06	-0.01
	75th	-0.08	-0.23	-0.20	-0.49	-0.81	-0.78	0.06	0.01	0.07	0.03	0.02	0.05
Salicchi	25th	0.02	-0.16	-0.12	-0.85	-1.39	-1.45	0.01	-0.14	-0.04	-0.05	-0.11	-0.10
	50th	-0.04	-0.22	-0.18	-0.84	-1.50	-1.65	0.00	-0.18	-0.05	-0.03	-0.17	-0.14
	75th	-0.06	-0.20	-0.16	-0.82	-1.38	-1.59	0.08	-0.02	0.05	-0.03	-0.11	-0.10
Diecimo	25th	0.12	-0.10	-0.05	-0.29	-0.69	-0.77	-0.10	-0.33	-0.22	-0.05	-0.16	-0.13
	50th	0.07	-0.18	-0.12	-0.22	-0.64	-0.71	-0.02	-0.20	-0.12	-0.01	-0.14	-0.09
	75th	0.05	-0.13	-0.09	-0.02	-0.30	-0.36	0.00	-0.11	-0.07	0.00	-0.02	0.01

Table 3.10. Differences between the 25th, 50th and 75th percentiles of the future standardized GWLs at short- (ST), medium- (MT) and long-term (LT) and the historical ones under the RCP8.5 scenario. Results obtained with the AI models proposed in this study and the regression model presented in Section 3.1.2.

		REGRESSION			NARX			LSTM			CNN		
		ST	MT	LT	ST	MT	LT	ST	MT	LT	ST	MT	LT
Paganico	25th	-0.04	-0.22	-0.29	-0.20	-0.40	-0.60	-0.16	-0.28	-0.64	0.05	0.03	0.06
	50th	-0.11	-0.32	-0.35	-0.23	-0.57	-0.93	-0.12	-0.18	-0.34	0.03	-0.02	-0.07
	75th	-0.06	-0.27	-0.29	-0.11	-0.52	-1.04	-0.08	-0.11	-0.30	0.02	-0.05	-0.14
Cugnia	25th	-0.01	-0.11	-0.13	0.15	0.04	-0.35	-0.13	-0.21	-0.40	0.01	-0.05	-0.16
	50th	-0.06	-0.17	-0.18	0.20	0.08	-0.38	-0.07	-0.11	-0.24	0.00	-0.06	-0.14
	75th	-0.08	-0.20	-0.21	0.36	0.20	-0.30	-0.06	-0.07	-0.19	-0.02	-0.08	-0.14
SAT 1	25th	-0.01	-0.11	-0.13	-0.23	-0.46	-0.52	-0.31	-0.61	-1.11	0.04	-0.01	-0.11
	50th	-0.06	-0.16	-0.17	-0.15	-0.43	-0.64	-0.17	-0.33	-0.63	0.03	-0.03	-0.13
	75th	-0.08	-0.20	-0.21	-0.06	-0.33	-0.61	-0.13	-0.24	-0.47	-0.01	-0.09	-0.19
Via Barsanti	25th	0.00	-0.12	-0.15	-0.17	-0.37	-0.51	-0.17	-0.36	-0.68	0.00	-0.07	-0.17
	50th	-0.06	-0.14	-0.17	-0.07	-0.25	-0.35	-0.13	-0.23	-0.43	0.00	-0.05	-0.16
	75th	-0.05	-0.15	-0.17	0.02	-0.13	-0.23	-0.07	-0.14	-0.34	0.00	-0.06	-0.18
Via Romboni	25th	-0.01	-0.16	-0.21	-0.26	-0.54	-0.93	-0.15	-0.31	-0.51	-0.05	-0.20	-0.34
	50th	-0.07	-0.19	-0.22	-0.15	-0.50	-0.97	-0.08	-0.19	-0.37	-0.06	-0.14	-0.27
	75th	-0.06	-0.15	-0.18	-0.04	-0.32	-0.77	-0.08	-0.14	-0.22	-0.01	-0.10	-0.19
Percorso Vita	25th	0.00	-0.10	-0.14	-0.20	-0.44	-0.68	-0.16	-0.34	-0.70	0.05	0.01	-0.03
	50th	-0.05	-0.12	-0.16	-0.09	-0.31	-0.53	-0.14	-0.22	-0.39	0.00	-0.03	-0.08
	75th	-0.08	-0.19	-0.21	-0.03	-0.32	-0.61	-0.07	-0.13	-0.32	-0.02	-0.06	-0.16
Nozzano	25th	-0.01	-0.14	-0.20	-0.45	-0.92	-1.28	-0.04	-0.02	-0.10	-0.03	-0.16	-0.21
	50th	-0.07	-0.19	-0.24	-0.38	-0.90	-1.32	-0.02	-0.07	-0.17	-0.07	-0.17	-0.27
	75th	-0.06	-0.21	-0.24	-0.20	-0.75	-1.16	0.05	0.03	-0.06	-0.02	-0.12	-0.22
S. Alessio	25th	-0.02	-0.13	-0.18	-0.49	-0.61	-0.74	-0.04	-0.04	-0.12	0.06	0.02	0.05
	50th	-0.06	-0.16	-0.21	-0.50	-0.70	-0.70	0.00	-0.06	-0.14	0.02	-0.01	-0.05
	75th	-0.08	-0.20	-0.25	-0.52	-0.81	-0.81	0.04	0.02	-0.01	0.05	0.04	0.04
Salicchi	25th	-0.01	-0.12	-0.19	-0.87	-1.43	-1.70	-0.02	-0.03	-0.13	-0.03	-0.14	-0.14
	50th	-0.06	-0.17	-0.22	-0.87	-1.64	-2.15	-0.04	-0.10	-0.22	-0.05	-0.16	-0.23
	75th	-0.06	-0.17	-0.21	-0.73	-1.53	-2.34	0.04	0.02	-0.09	-0.02	-0.11	-0.19
Diecimo	25th	0.10	-0.06	-0.11	-0.33	-0.73	-1.16	-0.12	-0.25	-0.39	-0.05	-0.13	-0.24
	50th	0.04	-0.12	-0.16	-0.25	-0.69	-1.14	-0.06	-0.15	-0.31	-0.02	-0.08	-0.17
	75th	0.07	-0.10	-0.13	-0.11	-0.40	-0.74	-0.07	-0.10	-0.19	0.01	-0.04	-0.04

3.1.6. Discussion and conclusions

An element of novelty in this work lies in the application of expeditious methods to assess potential climate change impacts on groundwater quantity. The methods are applied in northern Tuscany, Italy, but they are adaptable to various regions, requiring only basic observed data on groundwater levels, precipitation, and temperature. These approaches serve as valuable tools for promptly assessing aquifer conditions under climate change scenarios, essential for developing integrated mitigation and adaptation strategies. In general, the findings suggest a potential reduction in future groundwater availability for the studied region, particularly under the RCP8.5 scenario in the long term.

The statistical approach, employing linear relationships, provides a straightforward methodology for easily assessing the impacts of climate change on groundwater quantity. Certainly, this can be an advantage if the objective is to implement a user-friendly tool that is easy to interpret. However, the AI models outperform in capturing non-linear relationships, which are prevalent in water infiltration and aquifer recharge, rendering them more suitable than linear functions. Additionally, varying methods for calculating potential evapotranspiration may influence SPEI evaluations, leading to uncertainty in the evaluation of the SGI-SPEI relationship when using the statistical approach. In contrast, the AI models, directly incorporating climate variables like temperature and precipitation without relying on meteorological indices, remain unaffected by this source of uncertainty.

Another point worthy of discussion is to consider the validity of the relationships between meteorological and groundwater indices/variables, evaluated in the historical period, for future predictions. The reliability of these relationships for future projections has sparked some debates, as many factors may change. For instance, the dynamics of evapotranspiration may undergo changes as atmospheric CO_2 concentrations rise. As posited by Vicente-Serrano et al. (2020), increased atmospheric evaporative demand, driven by elevated radiation and temperature, may not necessarily intensify droughts. The outcome may differ based on whether the region has a humid or arid climate, affecting meteorological, hydrological, and agricultural droughts differently. Nonetheless, the authors concur that while higher CO_2 levels may lead to reduced plant water consumption due to optimized functions, increased temperatures could escalate evaporation from water surfaces and soil. Bloomfield et al. (2019) provide evidence from long-term observations in certain UK locations, suggesting that groundwater droughts are becoming more frequent due to anthropogenic warming, even in the absence of decreased precipitation and increased withdrawals. Elevated temperatures can also alter root systems, with plants in Mediterranean regions adapting to warming climates by extending and deepening their roots. Other studies (Teuling et al., 2013; Vicente-Serrano et al., 2014; Diffenbaugh et al., 2015; Dierauer and Zhu, 2020) stress the significance of considering temperature in drought indices, as it substantially heightens drought severity. Consequently, assessing the effects of climate change solely based on precipitation variations is intrinsically unreliable. It is imperative to incorporate thermal effects when discerning future climate and hydrological trends. In several regions, future precipitation alterations are negligible, while temperature increases are conspicuous, as illustrated in Figure 3.2, which is particularly relevant to this case study. In this context, for the statistical approach, while SPIs and SPEIs yield similar results for the historical period in this application, this may not necessarily hold true for the future. In alignment with the perspective of other researchers (Kumar et al., 2016), the relationships between SGIs and SPEIs may be more suitable for drought studies in the context of global warming than SGI-SPI relationships. In general, and also in the context of the AI approach, the consideration of temperature together with precipitation is significantly important when assessing the impact of climate change on groundwater resources. Even under the assumption that other factors influencing hydrological processes remain constant in the future, the models can capture the effect of changes in pumping rates, which correlate with precipitation and temperature changes. Higher temperatures often lead to increased irrigation demand, driving greater groundwater extraction, while reduced precipitation and prolonged droughts decrease recharge and increase extraction due to limited surface water supplies. Thus, despite not explicitly including these factors, the models indirectly account for them since they depend on climate variables.

Another interesting aspect, is the importance to acknowledge the substantial variability introduced by different climate models. Consequently, it is prudent to encompass an ensemble of models in the analysis (Jackson et al., 2015; Mascaro et al., 2018; D'Oria et al., 2018a), which helps visualize result uncertainties. In this study, a downscaling/bias correction technique was employed aimed at aligning raw climate model outputs with the statistical distribution of observed precipitation and temperature data at a monthly scale. While this approach adequately replicates the historical period, disparities between models persist in future projections, contributing significantly to result uncertainty. As evidenced in Figure 3.10, the CDF envelope of SGIs derived from climate models in future periods exhibits wide dispersion. A specific model plays a substantial role in introducing uncertainty, offering projections of abundant precipitation and consequently elevated groundwater levels compared to other models. However, estimates derived from the entire RCM ensemble align well with median and mean CDFs, validating this approach to consider model projections collectively as realizations of the same stochastic process.

Consideration must also be given to the selection of accumulation periods, as it varies depending on the aquifer's characteristics. For the statistical approach this regard the computation of meteorological indices, while for the AI approach this influence the sequence length to be considered. The type of natural recharge (precipitation or inflow from contiguous aquifers, lakes, or streams) and its conditions (e.g., distance between the ground level and the water table) contribute to this variability.

To conclude, comparisons between AI models and the statistical approach reveal nuances in outputs, highlighting the importance of considering model characteristics and design. LSTMs, known for their ability to capture temporal dependencies, outperform other models, suggesting their suitability for time series prediction. Both applications and their respective methodologies and results have been published in international journals (Secci et al., 2021, 2023).

3.2. An artificial neural network as a quick tool to assess the effects of climate change and agricultural policies on groundwater

Complex models, particularly those concerning aquifer dynamics, can often pose significant challenges for a wide array of stakeholders, including farmers, governmental authorities, and other involved parties. The intricate nature of these models, with their complex equations and extensive datasets, can be a deterrent to effective understanding and application. This complexity may hinder timely decision-making processes and the efficient management of aquifer resources, which are crucial for ensuring sustainable water usage and environmental preservation. In light of these challenges, the development and implementation of user-friendly tools, facilitated by surrogate models, emerge as a paramount solution. Innovative methodologies, as neural networks, with user-friendly interfaces and easily interpretable results, can bridge the gap between intricate scientific models and practical, real-world decision-making processes, ultimately benefiting the aquifer ecosystem and the communities that depend on it.

The objective of this part of the thesis is to develop a surrogate model, based on an artificial neural network, with the capability to simulate groundwater flow in the Konya closed basin, Turkey, one of the pilot sites investigated as part of the InTheMED project. This model serves as a tool for examining the potential impacts of climate change and agricultural policies on groundwater resources within the region.

3.2.1. Case study

The Konya closed Basin (Figure 3.20) is an important agricultural area, with over 50% of its flat terrain dedicated to farming. However, owing to the absence of rivers and streams in the region, farmers heavily depend on groundwater resources, especially during arid periods, to sustain their agricultural activities. Consequently, the groundwater reserves within the basin are under substantial stress. The last decade highlighted a transition from traditional wheat cultivation to more profitable crops, which demand increased water usage (Yılmaz et al., 2021). Furthermore, the anticipated impacts of climate change are expected to exacerbate the situation, potentially jeopardizing the sustainability of the basin due to excessive aquifer exploitation.

In pursuit of promoting a sustainable use of these groundwater resources, the Boğaziçi University of Istanbul, a partner of the InTheMED project, has developed a numerical hydrogeological model for the Konya closed basin. The full model simulates vertical water flow in the vadose zone as well as horizontal flow in the underlying aquifer system. The model was based on the MODFLOW model coupled with the UZF1 module (Harbaugh, 2005) that simulates vertical flow through the vadose zone.It was calibrated using field data collected during 2000-2019.

The surrogate model should be capable of replacing the full model, simulating groundwater conditions in the Konya closed basin with minimal computational burden.

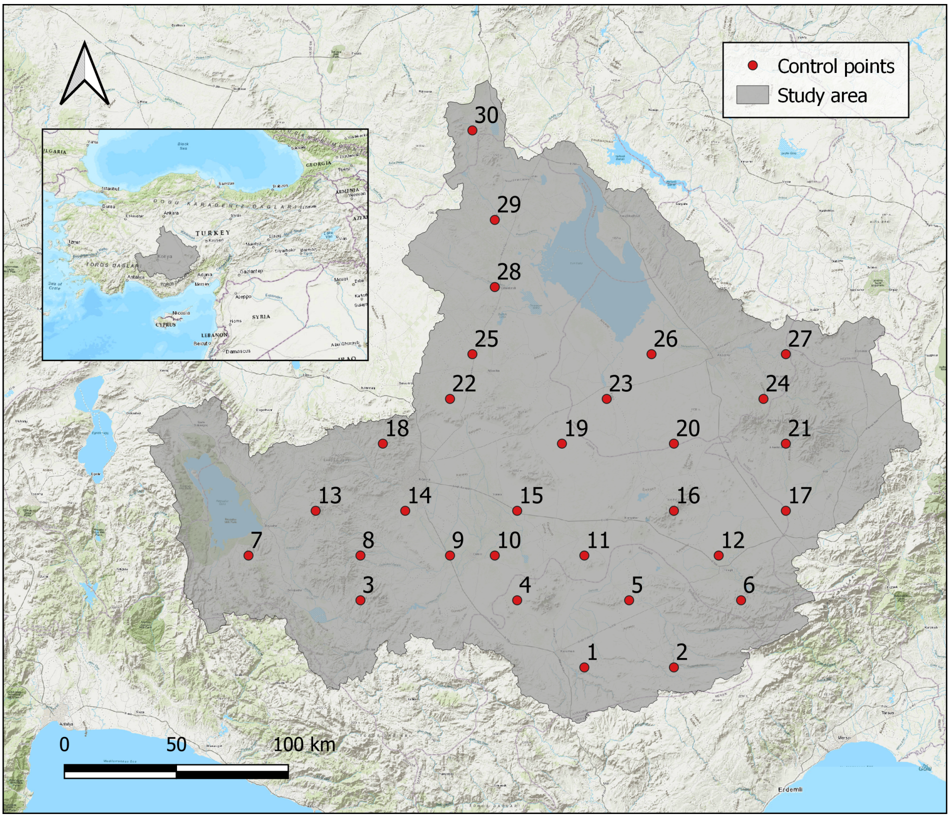


Figure 3.20. Konya closed basin (Turkey) and the 30 control points.

3.2.2. Surrogate model: ANN

A conventional neural network of the MLP type (Section 1.3.2) was trained to predict monthly groundwater levels, for the Konya basin, at 30 monitoring points (Figure 3.20) over the period from 2020 to 2039, considering a range of climate and agricultural scenarios. The climate scenarios are considered in terms of precipitation, while the agricultural scenarios are considered in terms of crop demand. Given the large number of unmetered wells in the Konya Plain, it was not possible to directly estimate groundwater extraction rates. For this reason, irrigation water extraction was estimated indirectly from water crop demand that, together with the precipitation, represent the field data used for model calibration.

3.2.2.1. Training dataset and set up of the ANN

The input features of the ANN are three and encompass two multiplicative coefficients, one applied to historical precipitation (prec coefficient) and the other to historical crop water demand (crop coefficient), and the time. The piezometric heads at the 30 monitoring points constitute the desired outputs.

A dataset comprising 100 combinations of precipitation and crop coefficients was generated using the Latin Hypercube Sampling method (McKay et al., 1979). The precipitation coefficient varies from 0.6 to 1.4, which means that precipitation ranges from -40% to +40% of that of the historical period. The crop coefficient varies from 0.75 to 1.25, which corresponds to a water demand varying from -25% to +25% of the historical one. For each combination, the full numerical model was executed starting from January 2020 until December 2039, to obtain monthly piezometric head data at the 30 monitoring points. The starting heads condition for this simulation is based on the December 2019 piezometric head data obtained from the calibrated model using historical information. The input data (100 combination of precipitation and crop along with the time) and the corresponding output (monthly piezometric head at the 30 control points) represent the dataset used to train the network. A single combination simulated by the full model produces 240 vectors of monthly piezometric head at the 30 control points(one for each month of simulation). Hence, the total size of the training dataset is 24,000 (100 combinations by 240 outputs each). The training dataset was divided into train (70%), validate (15%), and test (15%) sets.

The operational framework of the ANN model (MLP type) was previously introduced in Section 1.3.2. In this application, the neural network training parameters were configured as follows: 50 training epochs, a mini-batch size of 18, and a learning rate of 0.001. The objective function is minimized using the Levenberg–Marquardt algorithm (Hagan and Menhaj, 1994). The neural network incorporated an input layer composed of three neurons (prec coefficient, crop coefficient and time), a single hidden layer comprising 10 neurons, while the output vector size was set to 30 to match the number of control points. The total number of time steps considered amounted to 190, aligning with the length of the groundwater level time series.

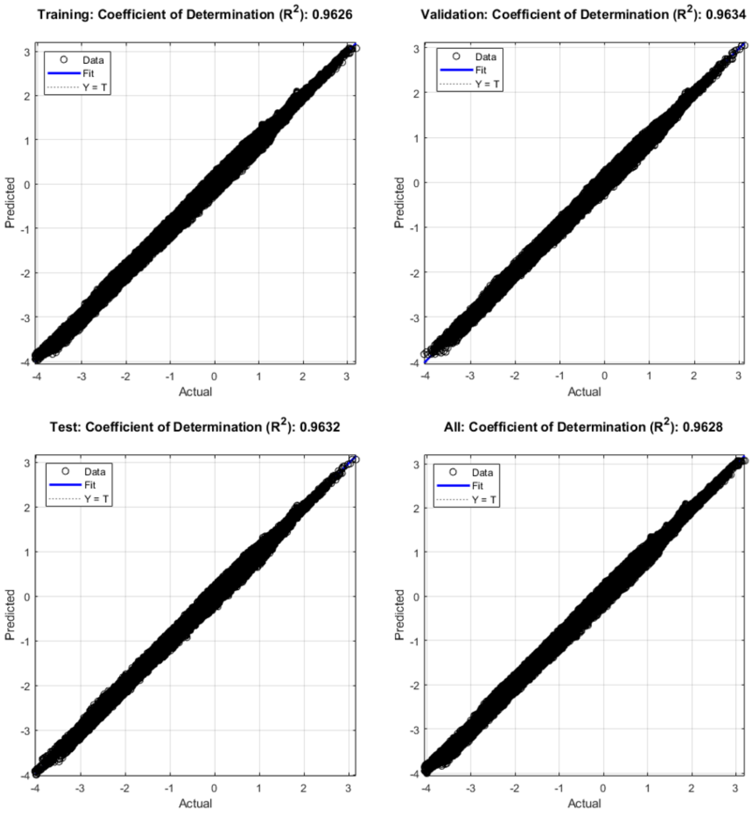
3.2.2.2. Evaluation of the performance

Table 3.11 displays the Mean Squared Errors (MSEs) calculated across the complete dataset, including evaluations for the training, validation, and testing subsets. At iteration one, the overall performance of the network is about 20.74 in terms of MSE. At the final step of the training, the overall performance for the entire dataset is good, with an MSE of 0.0061 on the entire dataset. This signifies a decrease of at least three orders of magnitude compared to the initial loss calculated with randomly initialized weights and biases. The training phase shows an MSE of 0.0060, and the validation phase has an MSE of 0.0061, while the test phase has an MSE of 0.0062. These findings might imply a concern related to overfitting, due to the good performance of the training phase; however, the consistent loss value associated with the validation dataset confirms effective training. This is further supported by the strong ability of the neural network to generalize when confronted with unseen data during the testing phase.

Table 3.11. MSE between the output of the neural network and the simulated groundwater levels.

Training	**Validation**	**Test**	**All**
0.0060	0.0061	0.0062	0.0061

To better gauge the quality of the implemented model, another crucial metric in assessing the quality of a regression model is explored: the Coefficient of Determination (R^2). It quantifies the proportion of the variance in the dependent variable that is predictable from the independent variable(s). As observed in Figure 3.21, alongside the presented MSE metrics, the R^2 values reaffirm the exceptional performance of the neural network across all phases. Notably, the R^2 values approach unity, indicating a remarkably high level of predictability in capturing the variance within the data.


Training: Coefficient of Determination (R^2): 0.9626

Validation: Coefficient of Determination (R^2): 0.9634

Test: Coefficient of Determination (R^2): 0.9632

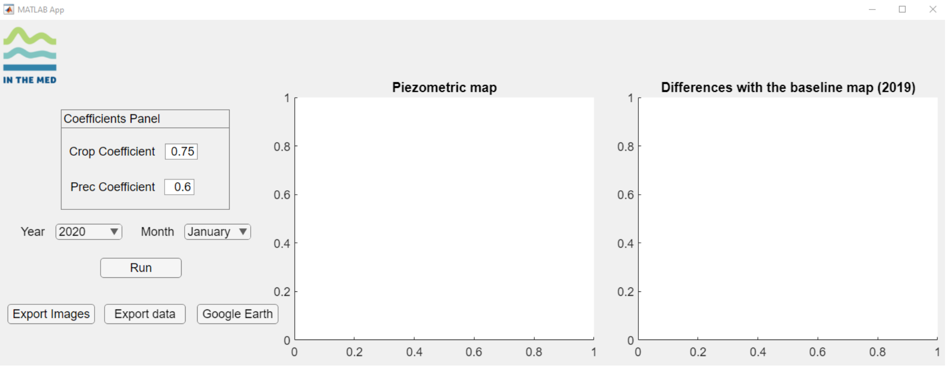
All: Coefficient of Determination (R^2): 0.9628

Figure 3.21. Training, validation and test performances of the neural network.

3.2.2.3. ANN tool

In this section the ANN tool is presented. The general idea is to use the fully trained ANN, to create a user-friendly application that users can use effortlessly, featuring intuitive interfaces and results that are easily understood. MATLAB (MathWorks, 2022) was selected as programming language and environment for codes implementation.

The tool interface exhibits an aspect as illustrated in Figure 3.22. Looking the panel on the left, the user only needs to select the values of the two coefficients and specify the desired investigation time for which the results are required.



Coefficients Panel

Crop Coefficient 0.75

Prec Coefficient 0.6

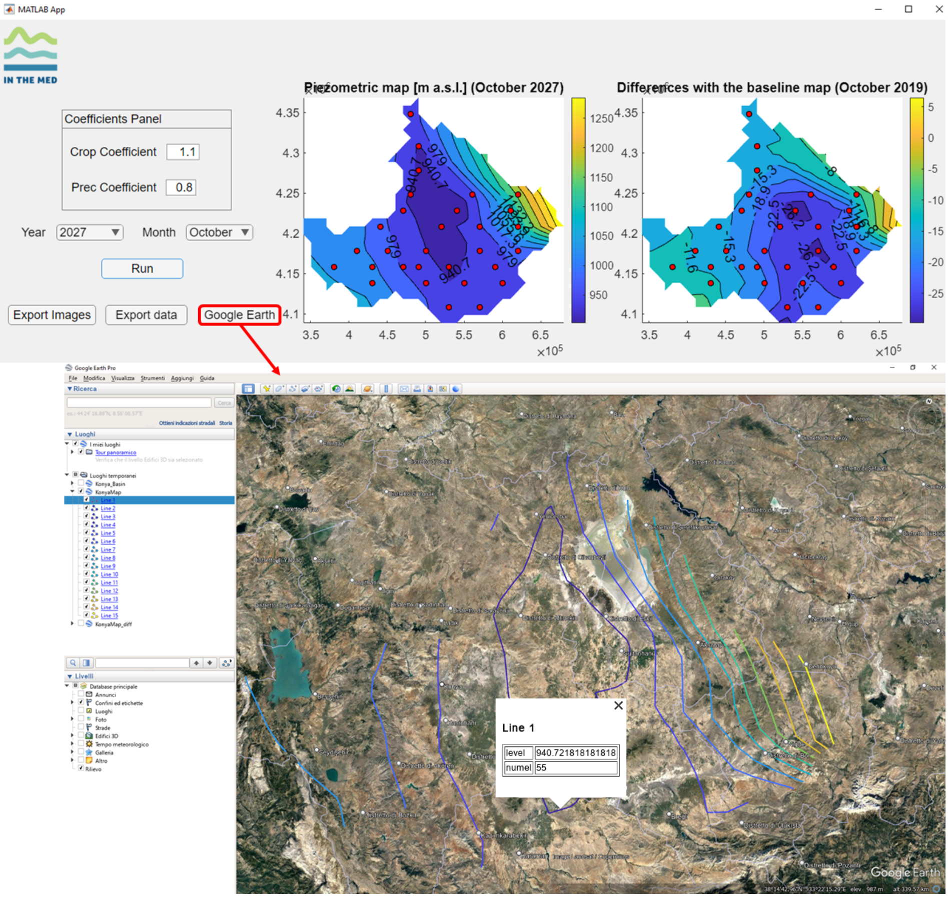
Year 2020

Month January

Figure 3.22. ANN tool interface developed using MATLAB.

By clicking the "Run" button, the fully trained ANN is invoked, utilizing the input features provided by the user to generate the desired result (Figure 3.23). It is worth noting that the results are not only presented as 30 monthly groundwater levels at the 30 control points but are also displayed in the form of a piezometric map. This is due to the implementation of an interpolation procedure that transforms the output of the network into maps. Specifically, interpolation utilizing the Delaunay triangulation of the scattered sample points was employed, as described in (Amidror, 2002). This interpolation method is readily available in MATLAB, ensuring ease and efficiency in its application. The interpolation method creates an interpolant that fits a surface of the form $v = F(x, y)$. Vectors x and y specify the (x, y) coordinates of the sample points. v is a vector that contains the sample values associated with the points (x, y).

Additionally, a map of differences is also generated, illustrating the disparities between the obtained piezometric map for the user selected time and the baseline map from the year 2019.



Crop Coefficient 1.1

Prec Coefficient 0.8

Year 2027 Month October

Line 1

level 940.721818181818

numel 55

Figure 3.23. ANN tool result.

The user can save the generated graphs as images (in PNG format) by clicking the "Export images" button and export the data associated with the 30 control points and the interpolated values in a .txt file using the "Export data" button. Finally, the resulting maps can be visualized, in the form of contour lines, in one of the most well-known free geographic information systems available, such as Google Earth Pro (Google LLC, 2023), as shown in Figure 3.23.

3.2.3. Conclusion and discussion

The importance of enabling people to understand complex systems practically cannot be overstated. When individuals, especially decision-makers, are equipped with user-friendly tools and easily interpretable results, they gain a practical grasp of intricate models. By bridging the gap between complexity and practicality, such tools facilitate communication, foster collaboration, and ultimately lead to more efficient problem-solving and decision-making processes. As in this case, the surrogate models allow the users to effortlessly access and visualize the results, which are presented as piezometric maps.

Future work in this field could explore the development of web applications or smartphone applications that incorporate these user-friendly systems, making them widely accessible and readily available to a broader audience. By creating intuitive platforms that can be accessed via the web or mobile devices, the reach of these tools can be extended, allowing a wider range of users, including the general public, to benefit from easy-to-use systems for understanding and managing complex models.

3.3. Artificial neural networks for solving forward and inverse transport problems

The demands of modern society have driven the development of economies reliant on intensive agriculture and industrial production. This poses a challenge in ensuring water sustainable utilization without compromising environmental integrity and jeopardizing groundwater quality.

A sustainable approach to address this challenge involves safeguarding groundwater quality, thereby obviating the need for costly treatment systems (Katsanou and Karapanagioti, 2019). Furthermore, the identification of the location of the contaminant source with its release history has attracted great attention within the scientific community called upon to provide theoretical methods to identify and limit the spread of the contaminant. Indeed, source identification of contaminants has emerged as a focal point of research over the last four decades, as underscored in recent reviews by Gómez-Hernández and Xu (2022) and Barati Moghaddam et al. (2021).

To identify remediation strategies immediately is essential to have accurate results in real time. With this aim, surrogate models can become the conceptual models of primary choice being able to study forward and inverse transport problem reducing in this way the computational cost compared with more complex models.

In the present study, ANNs were harnessed to solve direct groundwater contamination problems and to estimate the source location and its temporal release dynamics, starting from a sparse concentration dataset observed at monitoring wells. To assess and compare the proposed methodology, a complex literature-based case study was adopted (Ayvaz, 2010). The benchmark case entails a heterogeneous aquifer housing multiple contaminant sources and featuring seven monitoring wells. Additionally, for the first time, ANNs were deployed to simultaneously estimate the release history and quantify observation errors.

The investigated test case is examined across diverse scenarios and objectives: 1. Estimation of pollutant concentrations in monitoring wells in the presence of a known release source (direct problem). 2. Estimation of the release history at a single contaminant source with a known location. 3. Estimation of the release history at two contaminant sources with known locations. 4. Simultaneous estimation of the release history and location of a contaminant source with an unknown location. 5. Simultaneous estimation of the release history of two sources with known locations and the quantification of observation errors.

3.3.1. State of the art

The identification of contaminant sources from sparse concentration data is illustrative of an inverse problem within hydrology. The literature provides diverse methodologies to address such problems, categorizing inverse techniques for contaminant source identification into three principal domains (Barati Moghaddam et al., 2021): mathematics-based, stochastic-based, and optimization-based. Mathematics-based methods directly engage inverse source problems employing numerical or analytical methodologies. Mitigating stability concerns is typically managed through regularization and stabilization techniques (Skaggs and Kabala, 1994; Liu and Ball, 1999). Stochastic approaches (Woodbury and Ulrych, 1996; Butera et al., 2013; Cupola et al., 2015; Gzyl et al., 2014; Zanini and Woodbury, 2016; Xu and Gómez-Hernández, 2016; Todaro et al., 2021; Wang et al., 2021, 2022) contextualize the problem within a stochastic framework, treating the parameters for estimation as random variables. Optimization-based methodologies involve integrating simulation and optimization models. The simulation model resolves the flow and transport equations under specified initial and boundary conditions, subsequently minimizing the disparities between simulated and observed data through an optimization algorithm (Ayvaz, 2010; Jamshidi et al., 2020). Interested readers are referred to Gómez-Hernández and Xu (2022) and Barati Moghaddam et al. (2021) for exhaustive reviews on source reconstruction in groundwater hydrology and groundwater-surface hydrology, respectively.

Over the last decade, the rapid advancement of ANNs has been propelled by enhanced computational capabilities and technological innovations. The pioneering application of ANNs to contaminant source reconstruction was undertaken by Singh and Datta (2004) and Singh et al. (2004). They employed concentrations observed at monitoring points as network inputs and the release history at the contaminant source as the output. Singh and Datta (2004) devised an ANN-based approach to concurrently tackle groundwater pollution source identification and aquifer hydro-dispersive parameter estimation. Singh et al. (2004) explored the efficacy of ANNs in scenarios involving multiple sources and noise in observations. Chaubey and Srivastava (2022) introduced an ANN-based method for estimating source location and release concentrations within a simple 1D case study. Ayaz (2021) proposed an ANN for estimating the release history of groundwater pollution sources without prior knowledge of the release start time. Most recently, Pan et al. (2022) introduced a deep residual neural network as a forward surrogate model, combined with an ensemble smoother particle filter, to estimate groundwater contamination sources and aquifer hydraulic conductivity.

The black box nature of ANNs enables consideration of diverse scenarios by approximating various functions, including highly nonlinear ones, without a physical perspective on the underlying phenomena, while concurrently reducing computational costs and the requisite number of observations for model implementation.

3.3.2. Case study

To assess the reliability of the proposed methodology, a case study from the existing literature, initially presented by Ayvaz (2010) and subsequently adopted by other studies (Xing et al., 2019; Jamshidi et al., 2020; Todaro et al., 2022b), has been employed as a benchmark. In the first step, the model was established using the approach outlined in Ayvaz (2010).

The modeling domain was represented using a grid with block-centered cells (Figure 3.24). Table 3.12 provides a summary of the hydraulic and geometric characteristics of the domain. Regarding boundary conditions, specified head boundary conditions were applied at the upper-left (A-B) and lower-right (C-D) boundaries, while considering no-flow boundary conditions for the remaining parts of the domain. The aquifer was divided into five zones (as depicted in Figure 3.24), each characterized by distinct hydraulic conductivity values: $HK_1 = 0.0004\,\mathrm{m/s}$, $HK_2 = 0.0002\,\mathrm{m/s}$, $HK_3 = 0.0001\,\mathrm{m/s}$, $HK_4 = 0.0003\,\mathrm{m/s}$, and $HK_5 = 0.0007\,\mathrm{m/s}$. Hydraulic conductivity remained consistent within each zone, resulting in a scenario of steady-state and non-uniform flow conditions. Within the aquifer, we introduced two active sources and positioned seven monitoring locations strategically. The simulation covered a total period of five years, split into ten stress periods, each spanning six months. The sources remained active for the initial two years of the simulation, releasing a conservative contaminant, commonly referred to as the "golden-test", following the approach proposed by Ayvaz (2010). Consequently, the contaminant transport process exhibited a transient behavior.

The longitudinal and transverse dispersivity coefficients, denoted as α_L and α_T, respectively, were linked to the longitudinal and transversal dispersion components D_L and D_T of the dispersion tensor D through the relationships $D_L = \alpha_L u$ and $D_T = \alpha_T u$, with "u" representing the effective velocity of the flow field.

The data-driven model underwent training and validation to address two distinct study scenarios: forward and inverse transport problems (Table 3.13). In the forward approach, a single scenario was considered, focusing on the estimation of contaminant concentration values resulting from the influence of two sources at known positions (referred to as FWD 1). In contrast, the inverse approach encompassed a range of scenarios, including the estimation of the release history of a single source with a known position (INV 1), the estimation of both the release history and the location of a single source with an unknown position (INV 2), the estimation of the release histories of two sources with known positions (INV 3), and the estimation of the release histories of two sources with known positions, along with the determination of the order of magnitude of observation errors (INV 4).

Table 3.12. Hydraulic and geometry characteristics of the study domain.

Parameters	**Values**
Effective porosity, ϕ	0.3
Longitudinal dispersivity, α_L (m)	40
Transverse dispersivity, α_T (m)	4
Saturated thickness, b (m)	30
Grid spacing in the ζ direction, $\Delta\zeta$ (m)	100
Grid spacing in the η direction, $\Delta\eta$ (m)	100
Length of the stress periods, Δt (months)	6
Initial concentration (ppm)	0

Table 3.13. Summary of the input-target data for the investigated scenarios.

Scenario	**Input data**	**Output data**
FWD1	Mass release at 2 sources for 2 years every 6 months	Concentrations at 7 monitoring points observed for 5 years, one time per year
INV1	Concentrations at 7 monitoring points observed at time 5 years after the release	Mass release at one source for 2 years every 6 months
INV2	Concentrations at 7 monitoring points observed at time 5 years after the release	Mass release at one source for 2 years every 6 months. Planar coordinates of the source
INV3	Concentrations at 7 monitoring points observed for 5 years, one time per year	Mass release at 2 sources for 2 years every 6 months
INV4	Concentrations at 7 monitoring points observed for 5 years, one time per year	Mass release at 2 sources for 2 years every 6 months and error on observations

3.3.3. Groundwater flow and transport

To comprehensively investigate issues concerning the transport of contaminants within an aquifer, a thorough understanding of the flow characteristics is imperative. The analysis focuses on a confined aquifer with well-known hydraulic parameters, characterized by a two-dimensional flow equation (Eq. 3.3). Including the Darcy law, Eq. 3.3 expresses (in Cartesian coordinates $\boldsymbol{\xi} = (\zeta, \eta)$) the mass balance within a heterogeneous and anisotropic confined aquifer (under the ssumption to align with the principal directions of the symmetric tensor representative of transmissivity):

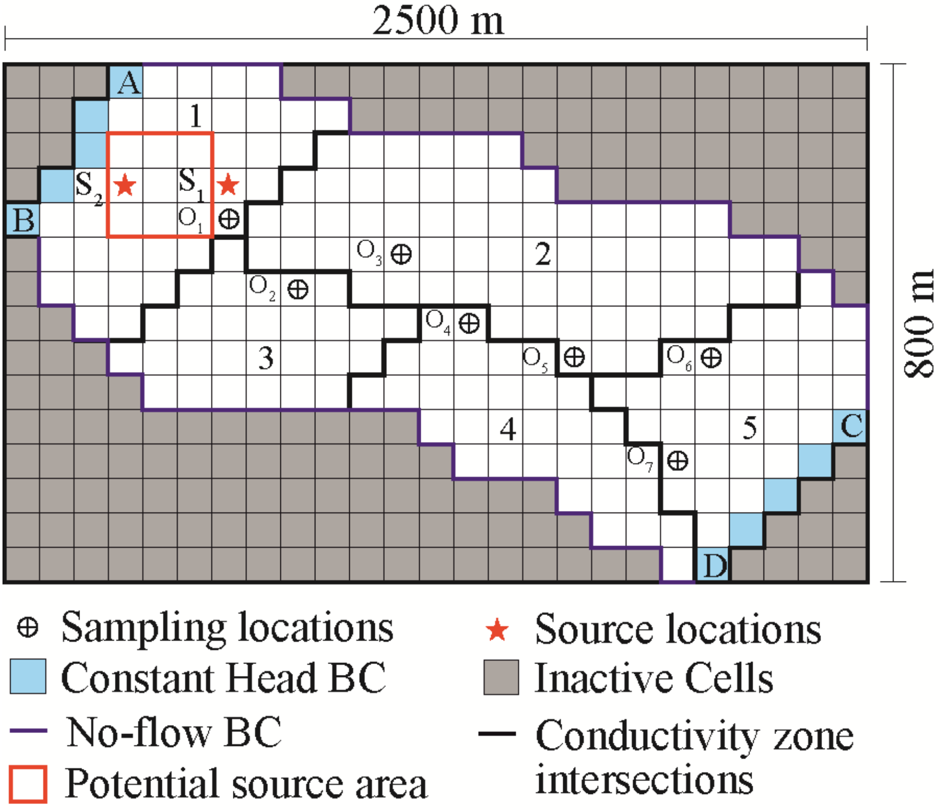


Figure 3.24. Discretization grid of the two-dimensional aquifers.

$$\frac{\partial}{\partial \zeta}\left(T_{\zeta\zeta}(\zeta,\eta)\frac{\partial h}{\partial \zeta}(\zeta,\eta,t)\right)+\frac{\partial}{\partial \eta}\left(T_{\eta\eta}(\zeta,\eta)\frac{\partial h}{\partial \eta}(\zeta,\eta,t)\right)=S\frac{\partial h}{\partial t}(\zeta,\eta,t)+Q(\zeta,\eta,t). \tag{3.3}$$

Here, $T_{\zeta\zeta}$ and $T_{\eta\eta}$ $[L^2T^{-1}]$ represent the principal transmissivity values along the ζ and η directions, t $[T]$ is time, h is the piezometric head $[L]$, Q $[LT^{-1}]$ is the flow rate per unit area entering or leaving the aquifer (positive when entering), and S [-] is the storativity of the porous medium.

Firstly, the flow problem was solved; then, the transport problem was addressed, utilizing the known flow field, while specifying particular boundary and initial conditions. Eq. 3.4 defines the transport process for an injection of a non-reactive and non-sorbing solute at a point source:

$$\frac{\partial(\phi C(\boldsymbol{\xi},t))}{\partial t} = \nabla\cdot[\phi D(\boldsymbol{\xi})\nabla C(\boldsymbol{\xi},t)] - \nabla\cdot[\phi u(\boldsymbol{\xi},t)C(\boldsymbol{\xi},t)] + s(\boldsymbol{\xi}_0,t)\delta(\boldsymbol{\xi}-\boldsymbol{\xi}_0). \quad (3.4)$$

In this equation, $\boldsymbol{\xi}$ represents the position vector of a specific point within the two-dimensional aquifer, $\boldsymbol{\xi}_0$ is the source location, $C(\boldsymbol{\xi},t)$ $[ML^{-3}]$ is the concentration at the location $\boldsymbol{\xi}$ and time t $[T]$, ϕ [-] is the effective porosity, $u(\boldsymbol{\xi},t)$ $[LT^{-1}]$ is the effective velocity vector field at the location $\boldsymbol{\xi}$ and time t $[T]$, $D(\boldsymbol{\xi})$ $[L^2T^{-1}]$ is the dispersion tensor, ∇ represents the spatial differential operator Nabla in the spatial coordinates $\boldsymbol{\xi}$, $s(\boldsymbol{\xi}_0,t)$ $[MT^{-1}]$ signifies the rate of mass release of the contaminant per unit time injected into the aquifer through the source, and δ $[L^{-3}]$ denotes the Dirac delta function.

To solve these equations a numerical model was developed using MODFLOW (Harbaugh, 2005) and MT3D (Zheng, 1999).

3.3.4. Set up of the ANN

The implemented artificial neural network (ANN) takes the form of a Multilayer Perceptron Network (Section 1.3.2). As previously introduced, the MLP model operates with three layers: the input, hidden, and output layers. The dimensions of the input and output layers are variable, depending on the specific scenario under consideration, while the hidden layer consistently comprises 10 neurons. The hidden layer is activated by the hyperbolic tangent function, while for the output layer the activation function employed is the identity function. The neural network Loss Function, denoted as $L(\Theta)$, where $\Theta = (\theta_1,\ldots,\theta_n)$ represents all the network parameters (weights and biases), is defined as reported in Eq. 1.4. To train the network, the Levenberg-Marquardt numerical optimization technique was selected. Building a data-driven model with an ANN requires both a training and validation dataset, comprising input and output values of a process. In this work, two distinct approaches have been considered: the forward ANN, which takes released mass fluxes at the sources as input and concentrations observed at monitoring points as output, and the inverse ANN, which considers concentrations observed at monitoring points as input and released mass fluxes at the sources as output. These approaches serve different purposes: the forward approach predicts concentrations at monitoring points based on a known mass flux release at the source, while the inverse approach estimates the released mass flux at the source based on known concentrations observed at the monitoring points.

Typically, training and validation datasets for data-driven models can be generated from field data or by utilizing results from a numerical model. In this case, a synthetic example from Ayvaz (2010) was adopted, and the dataset was generated using the numerical model. The dataset generation procedure involved the following steps:

1. Creation of groundwater flow and transport numerical models to replicate the studied aquifer. 2. Specification of the mass released at the source. 3. Iterative execution of forward flow and transport models to compute concentrations at monitoring points.

The concentrations monitored at various locations $\boldsymbol{\xi}$ are then associated with varying levels of error

$$C_{\text{error}}(\xi, t) = C_{\text{real}}(\xi, t) + \alpha\varepsilon C_{\text{real}}(\xi, t), \tag{3.5}$$

here, ε represents a random value drawn from a standard Gaussian distribution, and α signifies the magnitude of the error. In particular, this study explores free of errors and normal random errors equivalent to 0.1%, 1%, and 10% of the standard deviation.

The key consideration in this process is defining the size of the dataset and the extreme values of the mass released. The size (corresponding to the number of forward simulations in the numerical model) is determined based on the complexity of the scenario. Moreover, it is designed to minimize computational costs while ensuring effective training and validation. To reduce the input dataset size, Latin Hypercube Sampling (LHS) was employed. LHS generates variables that are uniformly distributed and uncorrelated, thereby optimizing the dataset efficiency.

For forward ANNs, the range of mass release should encompass available mass release data. For inverse ANNs, given the concentrations observed at monitoring wells (C_{true}), the dataset's boundaries can be established through a preliminary run of the numerical flow and transport models. This preliminary run involves injecting a constant mass rate (M_0) at the source, observing the maximum concentrations at monitoring wells ($C_{\max}$), and computing the ratio ($R = C_{\text{true}}/C_{\max}$) between C_{true} and the maximum concentrations obtained from the numerical model by injecting M_0. Using a linear relationship between mass release and concentrations at monitoring wells, it is possible to define the upper limit of the input dataset as a value greater than $M_0 \cdot R$, with the lower limit set to 0.

Before training the network, a preprocessing phase was performed on the dataset using the 'mapminmax' function (MathWorks, 2022). This phase involved transforming the input dataset $\{x^{(i)}\}_{i=1,\dots,N}$ and the target dataset $\{y^{(i)}\}_{i=1,\dots,N}$ into $\{\widetilde{x}^{(i)}\}_{i=1,\dots,N}$ and $\{\widetilde{y}^{(i)}\}_{i=1,\dots,N}$, scaling the values to the range [-1, 1]. This transformation was applied independently to each component, where $j \in \{1, \dots, d_1\}$ and $h \in \{1, \dots, d_3\}$, using the following formulas:

$$\widetilde{x}_j^{(i)} = 2\left(\frac{x_j^{(i)} - \min_{k=1,\dots,N}\{x_j^{(k)}\}}{\max_{k=1,\dots,N}\{x_j^{(k)}\} - \min_{k=1,\dots,N}\{x_j^{(k)}\}}\right) - 1 \tag{3.6}$$

$$\widetilde{y}_h^{(i)} = 2\left(\frac{y_h^{(i)} - \min_{k=1,\dots,N}\{y_h^{(k)}\}}{\max_{k=1,\dots,N}\{y_h^{(k)}\} - \min_{k=1,\dots,N}\{y_h^{(k)}\}}\right) - 1 \tag{3.7}$$

Then, during the training and validation phases of the network, certain "control" criteria were established to evaluate the performance. Firstly, the maximum number of training epochs was set to 1000. Secondly, to prevent overfitting, a limit of 6 validation checks was imposed. For each scenario, 70% of the dataset was allocated for network training, the remaining 30% was reserved for validation, while the golden-test was used to compare the ANN solution with Ayvaz (2010) and Jamshidi et al. (2020).

3.3.4.1. Set up of each scenario

Table 3.14 outlines the dimensions of the input and target datasets employed across all investigated scenarios.

Table 3.14. Summary of the input-target data for the investigated scenarios.

Scenario	Size of the dataset	Input layer dimension	Output layer dimension
FWD1	500	8 data	35 data
INV1	256	7 data	4 data
INV2	2304	7 data	6 data
INV3	500	26 data	8 data
INV4	500	26 data	9 data

Two sources with known positions (FWD 1)

In this particular scenario, the objective was to obtain concentration values at the observation points, totaling 35 values (concentrations at 7 monitoring points observed for 5 years, one time per year). To accomplish this, a dataset consisting of 500 input data vectors (mass release at 2 sources for 2 years every 6 months: 8 data) and 500 output data vectors (concentration values for each year) was utilized for both training and validation, as presented in Table 3.14.

During the testing phase, reference was made to concentration values derived from the original Ayvaz (2010) case study. In this network architecture, the input layer represented an eight-dimensional real linear space, featuring eight neurons that corresponded to the six-month release history values of the two sources. The output layer was a thirty-five-dimensional real linear space, consisting of 35 neurons representing concentration values observed at monitoring locations for each simulation year.

One source with known position (INV 1)

The dataset included 256 input data vectors (concentrations at 7 monitoring points observed at the last time step after the release: 7 data) and 256 output data vectors (mass release at one source for 2 years every 6 months: 4 data), as outlined in Table 3.14. The golden-test was used to assess the performance of the training and validation phases.

The network architecture consisted of an input layer representing a seven-dimensional real linear space with seven neurons corresponding to concentration values at monitoring locations at t=5 years. The output layer was a four-dimensional real linear space with four neurons representing the four six-month release history values.

One source with unknown position (INV 2)

In this scenario, the objective is to determine both the release history of Source 2 and its corresponding location, represented by the coordinates ζ and η. To achieve this, nine potential cells within the study domain have been identified, each characterized by its unique ζ and η coordinates. These coordinates, along with the mass release rate values for the six-month release history, will be provided as output by the network.

The dataset employed for training and validation purposes consists of 2304 input data vectors (concentrations at 7 monitoring points observed at the last time step after the release: 7 data) and 2304 output data vectors, encompassing both the six-month release history and the source coordinates (6 data). The dataset size of 2304 results from combining nine separate datasets, each containing 256 data points, corresponding to hypothetical source locations within the domain.

The network's type and structure remain consistent with that described for INV1, with the only distinction being in the output layer (3.14). In this case, the output layer is a real linear space with a dimension of six, comprising six neurons that represent the four six-month release history values in addition to the ζ and η coordinates.

Two sources with known positions (INV 3)

In this particular application, a dataset comprising 500 input data vectors (concentrations at 7 monitoring points observed for 5 years, one time per year: 35 data, reduced to 26) and 500 output data vectors (mass release at 2 sources for 2 years every 6 months: 8 data) has been employed to facilitate the training and validation of the network, as indicated in Table 3.14. During the testing phase, the six-month release history values, originally examined in the Ayvaz (2010) case study, have been adopted for evaluation.

It is noteworthy that the input data no longer correspond to the concentration values exclusively at the monitoring locations at time $t = 5$ years. Instead, they encompass the concentration values recorded annually at the observation points over a span of 5 years, resulting in a total of 35 values (representing data from 7 monitoring points for the entire simulation period).

As a practical consideration, among these 35 concentration values recorded annually at the observation points, 9 consistently exhibit values close to zero, irrespective of the released mass. Consequently, the input data has been effectively reduced to 26 values for the purpose of conducting 500 synthetic simulations. Consequently, the input layer now comprises a real linear space with a dimension of 26, constituted by 26 neurons. Since there are two sources under consideration, the output layer has evolved to a real linear space of dimension eight, consisting of eight neurons. These neurons serve the dual role of representing the four six-month release history values for the two sources within the domain.

Two sources with unknown observation error (INV 4)

The final application is focused on the concurrent estimation of the release history for the two sources with known positions, while also aiming to ascertain the order of magnitude associated with observation errors (α value). For this particular scenario, the training dataset remains consistent with that utilized in INV3. However, a notable modification has been introduced in the output dataset, which now incorporates an additional component representing the magnitude of the observation error (9 data, as detailed in Table 3.14).

As a result of this augmentation in the output dataset's complexity, the output layer has undergone a transformation, expanding to a real linear space characterized by a dimension of nine. Within this nine-dimensional output space, eight neurons serve the purpose of representing the four six-month release history values associated with the two sources within the domain. Simultaneously, one neuron is allocated to convey the crucial information pertaining to the order of magnitude concerning observation errors.

3.3.4.2. Evaluation of the performance

To facilitate a comparison between the outcomes produced by the data-driven model and those yielded by the physical model, a set of metrics has been established following the methodologies employed in Ayvaz (2010) and Jamshidi et al. (2020). These metrics encompass the following:

Normalized Error (NE):

$$NE(\%) = \frac{\sum_{i=1}^{M} |Z_i^{\wedge} - Z_i|}{\sum_{i=1}^{M} Z_i} \cdot 100 \tag{3.8}$$

Percent Average Estimation Error (PAEE):

$$\mathrm{PAEE}_i(\%) = \frac{|Z_i^{\wedge} - Z_i|}{Z_i} \cdot 100 \tag{3.9}$$

Standard Deviation (SDt):

$$\mathrm{SD}_t = \sqrt{\frac{\sum_{r=1}^{N_R} (Z_{t,r}^{\wedge} - \overline{Z}_t)^2}{N_R - 1}} \tag{3.10}$$

Mean Error (ME):

$$ME = \frac{\sum_{i=1}^{M} (Z_i^{\wedge} - Z_i)}{M} \tag{3.11}$$

Mean Absolute Error (MAE):

$$MAE = \frac{\sum_{i=1}^{M} |Z_i^{\wedge} - Z_i|}{M} \tag{3.12}$$

Root Mean Squared Error (RMSE):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{M} (Z_i^{\wedge} - Z_i)^2}{M}} \tag{3.13}$$

Normalized Root Mean Squared Error (NRMSE):

$$NRMSE(\%) = \frac{RMSE}{Z_{\max} - Z_{\min}} \cdot 100 \tag{3.14}$$

in these expressions, M signifies the count of unknowns, Z_i stands for the actual observed value (for the forward problem, this represents the concentration observed at monitoring points, and for the inverse problem, it signifies the mass flux released at the source), $Z_i^{\wedge}$ denotes the estimated value, $Z_{t,r}^{\wedge}$ represents the estimated value at time t and realization r, $\overline{Z}_t$ corresponds to the estimated value at time t, averaged over N_R realizations, $Z_{\max}$ and $Z_{\min}$ respectively denote the maximum and minimum actual observed values. These metrics provide a comprehensive assessment of model performance.

To ensure robust and reliable results, the metrics previously outlined were applied by averaging the outputs generated by 10 distinct neural networks, all trained using the same dataset.

It is important to note that the training and validation process involves a degree of randomness. This randomness encompasses the division of the dataset into training and validation sets, as well as the initialization of weights within the neural networks. Consequently, the 10 neural networks yield slightly varying results. The decision to employ this ensemble of 10 networks is motivated by the objective of assessing not only the desired output but also the inherent uncertainty associated with the results.

In this regard, Eq. 3.10 plays a pivotal role as it defines the confidence interval within which the processed output, derived as the average of these 10 realizations, falls. Furthermore, this approach proves especially effective in scenarios where the number of network parameters, typically contingent on the chosen architecture, is on par with the total data points in the training set or when dealing with noisy data, as in this study. In such cases, employing multiple neural networks and subsequently averaging their outputs serves as a valuable strategy to mitigate the risk of overfitting.

3.3.5. Results

This section provides an overview of the outcomes achieved through the data-driven models in both the forward and inverse approaches.

3.3.5.1. FWD 1

The results generated by the data-driven model for the FWD 1 scenario closely align with those obtained through the numerical model. Notably, the FWD 1 scenario represents a unique case where the data-driven model replaces the numerical model as a surrogate. The estimated concentrations exhibit a high level of agreement with the actual values, as visually demonstrated in Figure 3.25. Table 3.15 provides additional insights, showcasing that the disparities between the estimated and actual values are minimal, thereby underscoring the effectiveness of the neural network.

Furthermore, it is worth noting that the artificial neural network (ANN) exhibits exceptional computational efficiency, with response times as low as 0.2 seconds. In contrast, the MODFLOW + MT3DMS numerical model necessitates approximately 6 seconds to complete its computations. This significant disparity in computational speed underscores another advantage of employing the ANN in this context.

Table 3.15. ME, MAE, RMSE, and NRMSE computed on concentrations (mg/l) and related to the 35 average concentration values of the 7 monitoring wells (FWD 1).

ME (g/L)	0.0028
MAE (g/L)	0.0096
RMSE (g/L)	0.0153
NRMSE	0.52%

3.3.5.2. INV 1

INV 1 deals with an inverse simulation featuring a single release source with a known position. For the sake of brevity, in Figure 3.26 and Tables 3.16 and 3.17, only the results obtained for the error-free data and data corrupted with the highest magnitude of the error $\alpha = 0.1$ are presented. Remarkably, the results

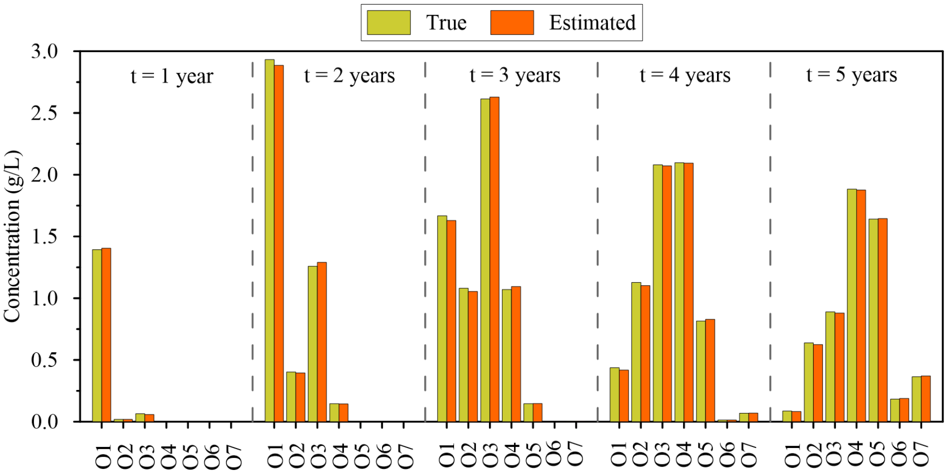


Figure 3.25. Observed and estimated concentration at 7 monitoring wells for 5 years of simulation recorded one time per year, forward simulation with two release sources (FWD 1).

for both error scenarios closely resemble each other and align well with the true release history (golden-test). The ANN exhibits a strong capability to accurately predict the desired output for this inverse application. Furthermore, Table 3.16 provides a comprehensive overview of the observed and estimated source release histories, along with computed metrics for the varying error levels. For a more detailed analysis of the error levels, Table 3.17 presents the corresponding metrics.

3.3.5.3. INV 2

In the case of the inverse application INV 2, the problem complexity increases compared to the previous scenario, but the data-driven model yields quite satisfactory results. In INV 2, we utilize the same observations as in INV 1. However, the unknowns have expanded from 4 (representing 4 releases in INV 1) to 6 (encompassing 4 releases and the source coordinates in INV 2). As a result, the output generated by the artificial neural network (ANN) in the INV 2 scenario may not

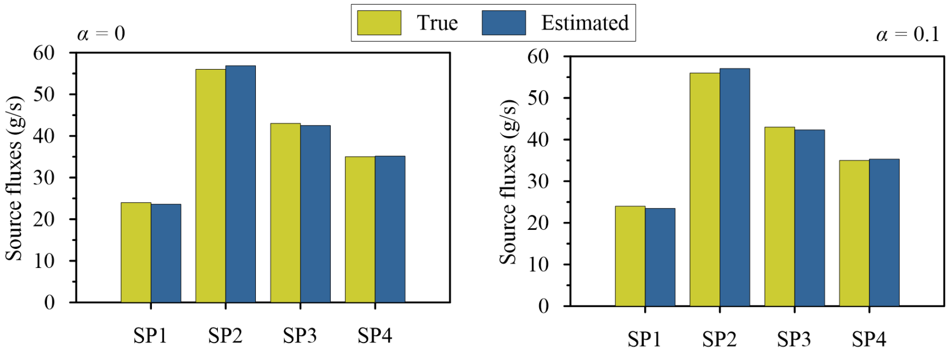


Figure 3.26. Observed and estimated release obtained as average of the results of 10 neural networks at known source, inverse simulation with one release source and different error level, error-free data ($\alpha = 0$) and corrupted data ($\alpha = 0.1$) (INV 1).

Table 3.16. Observed and estimated source release fluxes (g/s) obtained as the average of the results of 10 neural networks with related metrics PAEE, SD for different error levels, error-free data ($\alpha = 0$) and corrupted data ($\alpha = 0.1$) (INV 1).

Source	Stress period	Actual source fluxes (g/s)	Average estimated source fluxes (g/s)			Average estimated source fluxes (g/s)		
			$\alpha = 0$	PAEE (%)	SDt (g/s)	$\alpha = 0.10$	PAEE (%)	SDt (g/s)
S2	1	24	23.61	1.65	0.39	23.48	2.18	0.32
	2	56	56.88	1.58	0.75	57.07	1.92	0.92
	3	43	42.52	1.12	0.65	42.33	1.56	0.78
	4	35	35.16	0.47	0.37	35.30	0.86	0.37

be as accurate as in INV 1. Nevertheless, with an increased number of observations and consequently, more information being fed into the neural network, the network has the potential to perform better in the INV 2 scenario.

Despite these complexities, the ANN manages to estimate not only the release history but also the source location effectively, even for varying error levels. Table 3.18 provides the actual and estimated coordinates of the source, demonstrating the network's capability to make accurate predictions across different error sce-

Table 3.17. ME, MAE, RMSE, NRMSE and NE computed on source fluxes (g/s) described by four stress period and obtained as average of the results of 10 neural networks for different error level (INV 1).

	$\alpha = 0$	$\alpha = 0.1$
ME (g/s)	-0.04	-0.04
MAE (g/s)	0.48	0.64
RMSE (g/s)	0.55	0.70
NRMSE	1.71%	2.19%
NE	1.22%	1.63%

narios. This visual agreement can be observed in Figure 3.27, where the estimated release mass rates are compared to the real values, showcasing a strong agreement between them. Additionally, Tables 3.19 and 3.20 present the computed metrics for various error levels.

Table 3.18. Actual and estimated source location (ζ, η) obtained as the average of the results of 10 neural networks with different data error levels, error-free data ($\alpha = 0$) and corrupted data ($\alpha = 0.1$) (INV2).

		$\alpha = 0$		$\alpha = 0.1$	
Source	Actual location	Average estimated location	SDt (-)	Average estimated location	SDt (-)
S_2	$\zeta = 4$	$\zeta = 4.02$	0.199	$\zeta = 4.16$	0.212
	$\eta = 4$	$\eta = 3.83$	0.178	$\eta = 3.85$	0.271

3.3.5.4. INV 3

Scenario INV3 involves the estimation of the release history of two sources with known positions using concentration data from monitoring points within the domain. Different levels of error are considered in this analysis. The results obtained from the data-driven model are compared to two literature study cases: Example 2 from Ayvaz (2010) and the first case from Jamshidi et al. (2020). This comparison aims to assess the reliability of the neural network model.

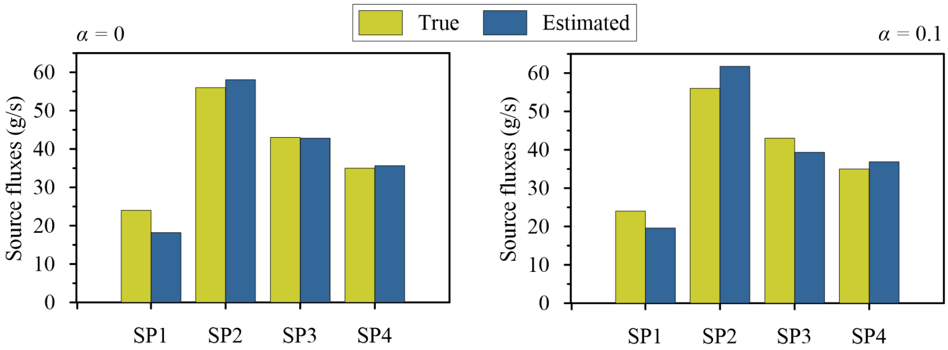


Figure 3.27. Observed and estimated release fluxes obtained as average of the results of 10 neural networks at estimated unknown source, inverse simulation with one release and different error level, error-free data ($\alpha = 0$) and corrupted data ($\alpha = 0.1$) (INV 2).

Table 3.19. Observed and estimated source release fluxes (g/s) obtained as the average of the results of 10 neural networks with related metrics PAEE, SD for different error levels, error-free data ($\alpha = 0$), and corrupted data ($\alpha = 0.1$) (INV2).

Source	Stress period	Actual source fluxes (g/s)	Average estimated source fluxes (g/s)			Average estimated source fluxes (g/s)		
			$\alpha = 0$	PAEE (%)	SDt (g/s)	$\alpha = 0.10$	PAEE (%)	SDt (g/s)
S2	1	24	18.21	24.13	3.98	19.59	18.41	6.40
	2	56	58.09	3.74	8.31	61.73	10.24	10.67
	3	43	42.80	0.47	7.08	39.33	8.54	9.99
	4	35	35.67	1.91	8.53	36.85	5.27	7.43

While the results from the two literature studies are valid, the neural network demonstrates superior precision in estimating release histories (Table 3.21). Furthermore, when considering the NE (%) value for a corrupted error level of $\alpha = 0.1$ (as shown in Table 3.22), this work achieves an NE (%) of only 1.23%, significantly lower than the values obtained in Ayvaz (8.06%) and Jamshidi et al. (18.06%) for the same error level.

In Figure 3.28, we present the results of estimated release histories for both

Table 3.20. ME, MAE, RMSE, NRMSE and NE computed on source fluxes (g/s) described by four stress period and obtained as average of the results of 10 neural networks for different error level (INV 2).

	$\alpha = 0$	$\alpha = 0.1$
ME (g/s)	0.81	0.13
MAE (g/s)	2.19	3.92
RMSE (g/s)	3.10	4.16
NRMSE	9.69%	13.01%
NE	5.54%	9.92%

error-free conditions ($\alpha = 0$) and the corrupted error level ($\alpha = 0.1$). These figures highlight the excellent agreement between the ANN predictions and the actual values. Additionally, Figure 3.29 provides a visual comparison, clearly indicating that the release histories generated by the neural network align more closely with the real data. Error bars, representing one standard deviation, are included to emphasize result reliability.

Table 3.22 offers a comparative analysis of statistical metrics between the literature cases and our work, considering various error levels.

Table 3.21. Comparison of the estimated and actual source release fluxes described by four stress periods at two known sources obtained as the average of the results of 10 neural networks with Ayvaz (2010), Jamshidi et al. (2020), and the present work, with level error $\alpha = 0.1$ and related statistical metrics (INV 3).

	Source		S_1				S_2			
	Stress period		1	2	3	4	1	2	3	4
	Actual source fluxes	(g/s)	35	90	65	47	24	56	43	35
Ayvaz (2010)	Average estimated source fluxes	(g/s)	35.4	87.5	62.9	53.4	31.5	48.5	46.49	33.6
	PAEE	(%)	1.23	2.8	3.27	13.7	31.1	13.4	9.14	4.13
	SDt	(g/s)	3.1	6.56	15.5	9.6	7.97	10.9	13.5	6.07
Jamshidi et al. (2020)	Average estimated source fluxes	(g/s)	41.6	63.3	77.7	43.6	22.2	48.5	47.7	27
	PAEE	(%)	18.9	29.6	19.5	7.15	7.6	13.4	11	22.8
	SDt	(g/s)	8	29.9	42.1	23.5	11.8	35.2	42	16.9
Present work	Average estimated source fluxes	(g/s)	35	89.2	64.9	47.3	23.6	58.3	42.1	35
	PAEE	(%)	0.05	0.9	0.15	0.69	1.76	4.09	2.06	0.01
	SDt	(g/s)	0.17	0.43	0.34	0.34	0.29	0.81	0.79	0.27

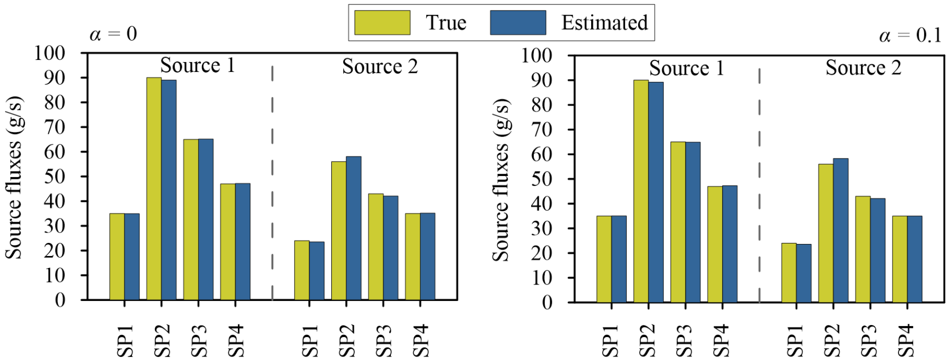


Figure 3.28. Observed and estimated release fluxes described by four stress period and obtained as average of the results of 10 neural networks results, inverse simulation with two release sources and different error level, error-free data ($\alpha = 0$) and corrupted data ($\alpha = 0.1$) (INV 3).

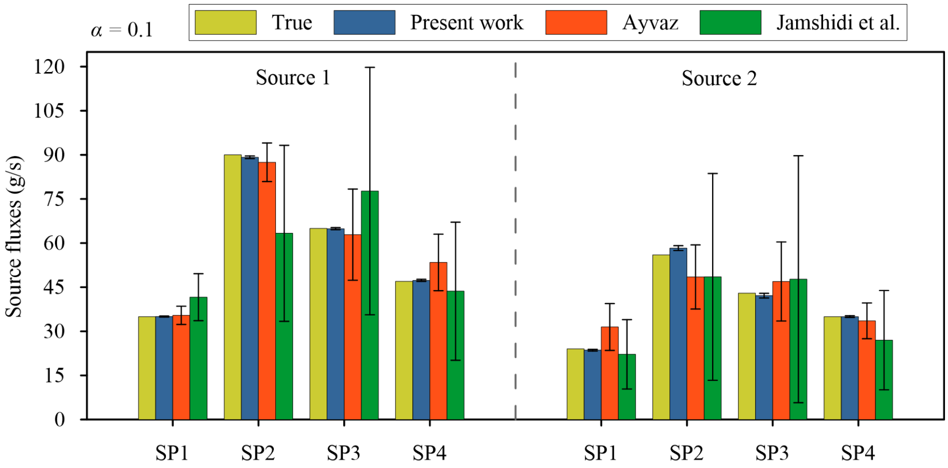


Figure 3.29. Estimated release fluxes in reference works for corrupted data ($\alpha = 0.1$) and, for any time step, the error bars related to one time the standard deviation (INV 3).

Table 3.22. Comparison of statistical metrics with Ayvaz (2010) and Jamshidi et al. (2020) for different error levels, error-free data ($\alpha = 0$), and corrupted data ($\alpha = 0.1$) (INV 3).

	$\alpha = 0$			$\alpha = 0.1$		
	Ayvaz (2010)	Jamshidi et al. (2020)	Present Work	Ayvaz (2010)	Jamshidi et al. (2020)	Present Work
ME (g/s)	0.00	-2.92	-0.02	0.58	-2.91	-0.05
MAE (g/s)	0.85	5.65	0.63	3.98	8.92	0.61
RMSE (g/s)	1.06	7.34	0.90	4.77	11.58	0.93
NRMSE	1.6%	11.1%	1.4%	7.2%	17.5%	1.4%
NE	1.72%	8.06%	1.17%	8.06%	18.06%	1.23%

3.3.5.5. INV 4

Scenario INV 4 introduces a novel aspect in the literature, as it involves estimating not only the release histories but also the order of magnitude of the error in the observations. The neural network demonstrates its ability to accurately recognize the observation errors, as evidenced in Table 3.23, which presents the actual and estimated values of the observation errors.

Additionally, as depicted in Figure 3.30 and summarized in Table 3.24, there is a strong agreement between the actual and estimated source fluxes for both $\alpha = 0$ and $\alpha = 0.1$. For further insights, Table 3.25 provides the statistical metrics calculated for the source fluxes at different estimated error levels.

Table 3.23. Actual and estimated order of magnitude obtained as average of the results of 10 neural networks of the error on concentrations (INV 4).

Actual Value	Average estimated value	SDt(-)
$\alpha = 0$	$\hat{\alpha} \to 0$	$\to 0$
$\alpha = 0.001$	$\hat{\alpha} = 0.0008$	0.0001
$\alpha = 0.01$	$\hat{\alpha} = 0.0099$	0.0008
$\alpha = 0.1$	$\hat{\alpha} = 0.1006$	0.0053

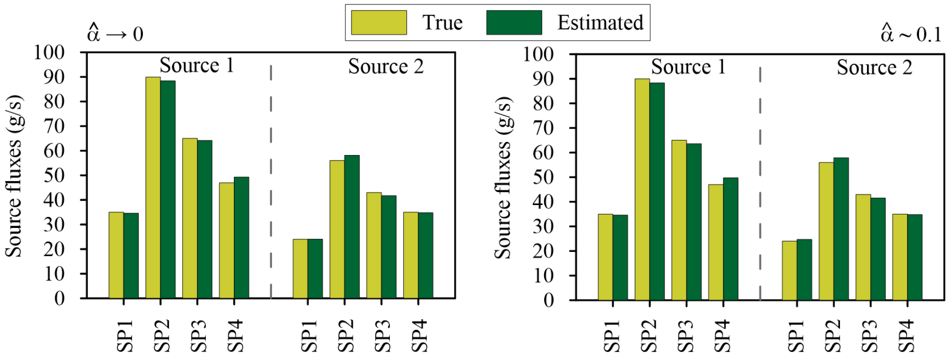


Figure 3.30. Observed and estimated release fluxes described by four stress periods at two known sources obtained as the average of the results of 10 neural networks at a known source. Inverse simulation with two release sources under different estimated error levels: error-free data ($\hat{\alpha} \to 0$) and perturbed data ($\hat{\alpha} \approx 0.10$) (INV 4).

Table 3.24. Observed and estimated source release fluxes (g/s) described by four stress periods at two known sources obtained as the average of the results of 10 neural networks with related metrics PAEE, SD for different estimated error levels, error-free data ($\hat{\alpha} \to 0$) and perturbed data ($\hat{\alpha} \approx 0.10$) (INV 4)).

Source	Stress period	Actual source fluxes (g/s)	Average estimated source fluxes (g/s)			Average estimated source fluxes (g/s)		
			$\alpha = 0$	PAEE (%)	SDt (g/s)	$\alpha = 0.10$	PAEE (%)	SDt (g/s)
S1	1	35	34.59	1.17	0.69	34.58	1.19	0.57
	2	90	88.40	1.78	0.83	88.33	1.86	0.76
	3	65	64.17	1.28	1.42	63.63	2.11	1.16
	4	47	49.29	4.86	0.67	49.69	5.72	0.54
S2	1	24	24.1	0.38	1.15	24.71	2.95	1.32
	2	56	58.14	3.84	1.93	57.88	3.37	2.02
	3	43	41.71	2.99	1.19	41.50	3.48	1.26
	4	35	34.80	0.58	0.67	34.79	0.59	0.81

Table 3.25. ME, MAE, RMSE, NRMSE and NE computed on source fluxes (g/s) described by four stress period at two known sources obtained as average of the results of 10 neural networks for different estimated error level (INV 4).

	$\hat{\alpha} \to 0$	$\hat{\alpha} \approx 0.001$	$\hat{\alpha} \approx 0.01$	$\hat{\alpha} \approx 0.10$
ME (g/s)	-0.02	-0.03	0.00	-0.01
MAE (g/s)	1.11	1.11	1.27	1.31
RMSE (g/s)	1.37	1.36	1.57	1.52
NRMSE	2.10%	2.10%	2.40%	2.30%
NE	5.54%	3.10%	2.71%	2.65%

3.3.6. Discussion and conclusions

In this study, neural networks have been employed as data-driven models to tackle various challenges associated with both forward and inverse transport problems. These challenges involve the utilization of concentration data from different monitoring points and release history information during the training phase. The outcomes of this investigations indicate that this data-driven approach is highly effective in providing solutions with minimal computational overhead, offering valuable insights for aquifer management and expediting the development of remediation strategies.

One key advantage of our work lies in the use of LHS, which significantly reduces the number of forward simulations required for network training, thus alleviating the computational burden. Notably, Ayvaz (2010) employed an approach based on minimizing an objective function derived from the forward model, necessitating a substantial number of simulations to achieve convergence (32,859 simulations). Conversely, Jamshidi et al. (2020) adopted a transfer function theory-based approach, allowing them to run the simulation model only once. The transfer matrices were optimized using an algorithm requiring fewer iterations, which is comparable to the number of simulations performed in our study. However, the results obtained in our work surpass those of the other approaches.

Furthermore, both Ayvaz (2010) and Jamshidi et al. (2020) utilized a dataset comprising 140 concentration values during the optimization process. This led to a situation where the number of unknown parameters was much smaller than the number of measurements. In contrast, our proposed method proved highly efficient, substantially diminishing the volume of a priori information (only 26 observations) required to tackle the inverse problem.

Upon evaluating the performance using computed metrics (Tables 3.22 and 3.25), the implemented ANN consistently demonstrated excellent results, outperforming the other two existing literature cases. For example, the SDt values calculated for the INV 3 application showcased a narrower confidence interval, signifying lower uncertainty compared to the SDt values from the other studies (Figure 3.29). Additionally, the SDt values calculated for the INV 4 application underscored the ANN's capability to handle observation error estimation, addressing a previously unexplored aspect in ANN-based research.

One of the most significant advantages of employing a data-driven model like ANN is its ability to provide computationally efficient solutions once trained. This eliminates the need for complex numerical models for each evaluation. Neural networks rely on the numerical model only during the training phase to generate a training dataset within specified ranges. After successful training, the network can swiftly provide the desired output.

In practice, applying these procedures has been met with challenges, primarily due to the assumption of known aquifer parameters when they are often poorly characterized and highly heterogeneous (Chen et al., 2021). Building on the promising results of this study, future research endeavors will focus on simultaneously estimating network parameters while identifying release histories, adapting the methodology also for unconfined and three-dimensional aquifers. Furthermore, addressing the issue of generalization is of paramount importance. Neural networks can generalize within their training range but struggle to do so outside of it. To mitigate this challenge, future work may explore the application of "Physically Informed Neural Networks" (Raissi et al., 2019).

In conclusion, the methodology and the results obtained were submitted to an international journal and were published (Secci et al., 2022).

3.4. Physics-informed neural networks for solving transient unconfined groundwater flow

In Chapter 3.4, the spotlight is on the intricacies of training PINNs and investigates the use of PINNs for solving transient unconfined groundwater flow.

The capability of PINNs to assimilate physical constraints into neural networks, as discussed in Section 1.3.6, navigate intricate geometries and boundary conditions, and operate effectively with limited data renders them suitable for simulating unconfined groundwater flow. Nonetheless, their utilization for this specific purpose has been relatively unexplored (Shadab et al., 2021; Zhang et al., 2022).

The application of PINNs is demonstrated in two scenarios: an isotropic and homogeneous aquifer, and an anisotropic and heterogeneous aquifer. The results are compared with the finite difference numerical solution obtained through a numerical model implemented with MODFLOW 2005 (Harbaugh, 2005). In addition, the result of the PINN are compared to the results obtained by a conventional neural network MLP (Section 1.3.2).

3.4.1. Transient unconfined flow

Understanding unconfined groundwater flow is pivotal in the context of water resource management, water quality preservation, and the mitigation of environmental ramifications like land subsidence and saltwater intrusion in coastal aquifers. However, tackling the complexities of unconfined groundwater flow is far from straightforward, given the necessity to account for spatially and temporally fluctuating boundary conditions. In an effort to simplify this intricate problem, initial assumptions were introduced by Dupuit and Forchheimer (Bear, 2012, Eq. 4-64), and later generalized by Boussinesq (Bear, 2012, Eq. 5-76). Although the Boussinesq equation provides a useful yet simplified framework, it fails to encompass certain intricate physical phenomena inherent in unconfined groundwater flow, such as vertical flows and steep hydraulic gradients. Consequently, its accuracy may be compromised in specific scenarios, such as those in proximity to pumping wells, discharge zones in coastal aquifers, or areas characterized by rugged topography. Over time, researchers have actively sought optimal solutions for the equation of simplified unconfined groundwater flow (Meenal and Eldho, 2011; Pulido-Velazquez et al., 2007; Taigbenu and Nyirenda, 2010).

Addressing the spatiotemporal-variant nature of the phreatic surface as a boundary condition presents challenges due to its intricate characteristics and substantial computational demands (Guo, 1997). ANNs have gained increasing traction in environmental and water resource studies due to their adeptness at swiftly and accurately processing extensive datasets (Sit et al., 2020; Tahmasebi and Sahimi, 2021; Mariethoz and Gómez-Hernández, 2021). While ANNs are data-driven models that offer cost-efficiency compared to process-based models, they demand substantial data volumes to yield precise outcomes and lack the intrinsic physical interpretability characteristic of process-based models. While some literature exists on the utilization of ANNs as surrogate models for groundwater flow equations (Asher et al., 2015), their application in the context of unconfined flow, particularly for predictive purposes, remains relatively limited.

In this study, PINNs are used to compute the phreatic surface and piezometric heads within a synthetic unconfined aquifer. Unlike prior investigations (Shadab et al., 2021; Zhang et al., 2022), the phreatic surface is considered as a boundary condition with a spatiotemporal-varying nature and an unknown geometry, subject to the requirement that the piezometric head equals the elevation at the phreatic surface. The groundwater flow partial differential equation (PDE) for a transient unconfined aquifer is employed as the foundational model, without simplifications.

The PDE governing unconfined groundwater flow in a two-dimensional heterogeneous aquifer in the vertical plane (XZ) under transient conditions is expressed as follows:

$$\begin{aligned}\frac{\partial}{\partial x}\left(K_{xx}(x,z)\frac{\partial h}{\partial x}(x,z,t)\right)+\frac{\partial}{\partial z}\left(K_{zz}(x,z)\frac{\partial h}{\partial z}(x,z,t)\right)\Bigg) \\ = S(x,z)\frac{\partial h}{\partial t}(x,z,t)+W(x,z,t)\end{aligned} \tag{3.15}$$

In this equation, $K_{xx}(x,z)$ and $K_{zz}(x,z)$ $[LT^{-1}]$ are the principal values of the hydraulic conductivity tensor, assuming that the principal directions are parallel to axes x and z, t $[T]$ is time, $h(x,z,t)$ $[L]$ is piezometric head, W $[T^{-1}]$ is an external flow extraction rate per unit volume, and S $[L^{-1}]$ is specific storage. Additionally, S_y is the specific yield, a dimensionless quantity. The initial condition is specified as $h(x,z,0) = h_0(x,z)$, where h_0 is a known function. The problem is further constrained by boundary conditions, which can be of standard types, such as prescribed head, prescribed flow, or prescribed linear combinations of head and flow. However, in the case of unconfined aquifers, a distinctive boundary condition applies to the phreatic surface, where the piezometric head h must equal the elevation z along the phreatic surface for all points (x,z). Notably, the position of the phreatic surface is unknown in advance and varies over time, adding complexity to the solution process.

$$h(x,z) = z \quad \forall(x,z) \quad \text{along the phreatic surface.} \tag{3.16}$$

3.4.2. Case study

In this study, a two-dimensional unconfined aquifer defined in the domain $(x, z) \in [0, 1] \times [0, 1]$ is considered. Being a synthetic application and aligned with the normalization procedures discussed in Section 1.3.1, to ensure the generality of the results, all parameters and simulation outcomes are presented without units, with validity retained as long as consistent units are employed.

The transient groundwater flow is simulated over a time period (t) ranging from 0 to 1, with four specific time instances analyzed at $t = 0.01$, $t = 0.25$, $t = 0.5$, and $t = 1$. The bottom boundary is impermeable throughout the simulation. The model captures the transient behavior that ensues from a specific set of boundary conditions. At time 0, the left and right boundaries correspond to reservoirs prescribing heads equal to 1, with the initial heads conforming to the steady-state solution for these conditions, specifically $h(x, z) = 1$, with the phreatic surface coinciding with the top boundary. Suddenly, at time 0, the left reservoir's level drops to 0.4, and the right reservoir's level drops to 0.6. These new boundary conditions translate to $h(0, z) = 0.4$ for $z \in [0, 0.4]$ and undefined for $z > 0.4$ on the left boundary, and $h(1, z) = 0.6$ for $z \in [0, 0.6]$ and undefined for $z > 0.6$ on the right boundary. This abrupt change induces transient behavior, which we aim to model using PINNs.

Two synthetic aquifers are investigated in this study: a homogeneous and isotropic unconfined aquifer (SC1) and a heterogeneous and anisotropic unconfined aquifer (SC2). Table 3.26 outlines the geometric and hydraulic characteristics of these cases. To facilitate benchmarking, the MODFLOW solution is computed. While the PINN solution operates in a meshless manner (as detailed below), MODFLOW necessitates domain discretization. For this purpose, a discretization of 20 by 20 cells in space and 1440 steps in time is chosen. Figure 3.31 illustrates the utilized discretization, the spatial distribution of conductivity for the heteroge-

Table 3.26. Hydraulic and geometric characteristics of the study domain

Parameters	**SC1**	**SC2**
Specific yield	10^{-3}	10^{-3}
Horizontal hydraulic conductivity, K_{xx}	10^{-3}	$K_1 = 4 \cdot 10^{-3}, K_2 = 10^{-3}, K_3 = 2 \cdot 10^{-3}, K_4 = 3 \cdot 10^{-3}$
Vertical hydraulic conductivity, K_{zz}	10^{-3}	$K_1 = 4 \cdot 10^{-4}, K_2 = 10^{-4}, K_3 = 2 \cdot 10^{-4}, K_4 = 3 \cdot 10^{-4}$
Grid spacing in the x direction, Δx	0.05	0.05
Grid spacing in the z direction, Δz	0.05	0.05
Simulation time, Δt	1	1
Total time steps	1440	1440

neous case, and the boundary conditions implemented in MODFLOW after time 0.

For training the artificial network, a specific number of observations within the saturated aquifer zone is considered. Particularly, 10% of active cells from the in the (XZ) plane are sampled from the MODFLOW solution randomly for each of the four time steps, excluding locations above the phreatic surface. This choice of observations may seem arbitrary but aims to establish control points inside the domain for improved network training. To ensure that the network satisfies the PDE (3.15) with its initial and boundary conditions, specific points where PDE verification is required were identified. Accordingly, 1000 points was selected along the left boundary, randomly chosen within the segment $[0, 0] \times [0, 0.4]$, and similarly, 1000 points along the right boundary within the segment $[1, 0] \times [1, 0.6]$. These points enforce adherence to the prescribed head boundary conditions. Additionally, 1000 points are randomly selected along the bottom boundary within the segment $[0, 0] \times [1, 0]$, ensuring compliance with the bottom impermeable condition. For initial conditions, 500 point locations are randomly chosen from the simulation domain at time 0, serving to enforce the specified initial conditions. Lastly, 25,000 points for SC1 and 35,000 points for SC2 are randomly selected within the simulation domain $[0, 0] \times [1, 1]$ and used as collocation points to ensure the accurate reproduction of the PDE.

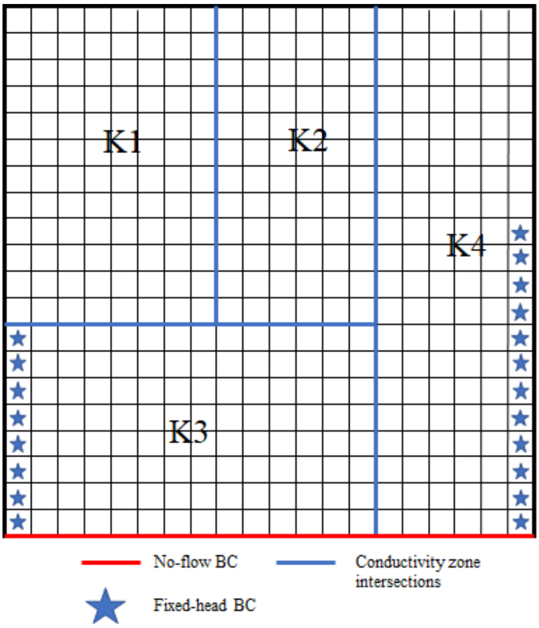


Figure 3.31. Synthetic domain.

3.4.3. PINN set up

The primary objective of the implemented PINN is to integrate the supervised loss function derived from the data while simultaneously adhering to the governing PDE, boundary conditions and initial conditions. This is accomplished by formulating a comprehensive loss function that incorporates four essential components: the prediction residual with the observed values, PDE residual, boundary residual, and initial conditions residual:

$$\text{Loss} = \text{Loss OBS} + \text{Loss PDE} + \text{Loss BC} + \text{Loss IC} \tag{3.17}$$

In this study, two structurally identical neural networks (Figure 3.32), of the MLP type (Section 1.3.2), are employed, differing only in their input and output layers. The first network, denoted as ANN1, is designed to calculate the piezometric head value (output) using the spatial coordinates (x, z) and time (t) as inputs. The second network, known as ANN2, takes the x coordinate and time as input values and predicts the z_s coordinate value (output), indicating the position of the free surface at a specific time. While both networks could be trained simultaneously with a single loss function, our approach involves a preliminary iteration where ANN1 is trained first, followed by the training of ANN2 (with ANN1 parameters fixed). The weights and biases obtained in this preliminary iteration serve as starting values for the joint training of the two networks.

Each artificial network comprises an input layer, seven hidden layers (each containing 20 neurons), and an output layer. Functionally, the ANN can be viewed as a differentiable system consisting of a sequence of multivariable vector-valued functions, encompassing affine transformations and linear or nonlinear functions (activation functions), mapping from $\mathbb{R}^{d_1}$ to $\mathbb{R}^{d_3}$ as illustrated in Eq. (3.18):

$$\mathbb{R}^{d_1} \Rightarrow \mathbb{R}^{d_2} \Rightarrow \mathbb{R}^{d_2} \Rightarrow \mathbb{R}^{d_2} \Rightarrow \mathbb{R}^{d_2} \Rightarrow \mathbb{R}^{d_2} \Rightarrow \mathbb{R}^{d_2} \Rightarrow \mathbb{R}^{d_2} \Rightarrow \mathbb{R}^{d_3}. \tag{3.18}$$

Here, d_1 and d_3 represent the dimensions of the input and output layers, respectively. In this study, d_1 is three for ANN1 and two for ANN2, d_2 is the number of neurons in the hidden layer (20 for both ANN1 and ANN2), and d_3 is one for both networks.

The selection of the number of hidden layers, evaluation points for network performance assessment, activation function choice, and other hyperparameters necessary for network definition were determined through initial manual testing. These tests aimed to identify configurations that optimize the ANN's performance, minimizing errors while ensuring efficient processing times. Specifically, the activation function used is the hyperbolic tangent (tanh), the number of training

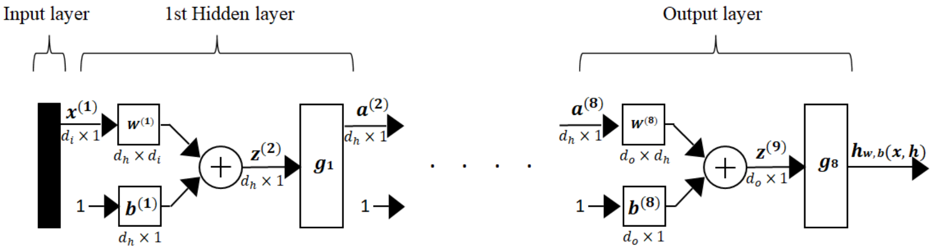


Figure 3.32. Sketch of the implemented neural networks.

epochs is set to 200, the mini-batch size is 128, the initial learning rate is 0.01, and the decay rate is 0.005.

3.4.3.1. First neural network (ANN1)

ANN1 (Figure 3.33) is responsible for predicting the piezometric head based on three input parameters: spatial coordinates (x, z) and temporal variable (t). These spatiotemporal coordinates do not require a mesh and, once trained, provide a meshless solution detached from any space-time discretization.

The training of ANN1 revolves around minimizing a loss function with the following components:

1. Loss associated with the error in reproducing the observed values, which equals the average of the squared differences between observations and predictions at the chosen locations and times.

$$L_{OBS} = \sum_{\Gamma} \left(h_{\text{predicted}} - h_{\text{observed}}\right)^2 \tag{3.19}$$

where:

- the summation symbol $\sum_{\Gamma}$ represents the summation over spatial and temporal regions corresponding to the observed values.

- $h_{\text{predicted}}$ is the predicted hydraulic head.

- h_{observed} is the observed hydraulic head as a function of space and time.

2. Loss associated with the error in reproducing the initial conditions, which equals the sum of the squared differences between the network prediction at time zero and the known initial values at the sampled locations

$$L_{IC} = \sum_{\Gamma} (h_{\text{predicted}} - h_{\text{initial}})^2 \tag{3.20}$$

where:

- the summation symbol $\sum_{\Gamma}$ represents the summation over over space corresponding to the initial conditions.

- $h_{\text{predicted}}$ is the predicted hydraulic head.

- h_{initial} is the initial hydraulic head as a function of space and time.

3. Loss associated with the error in reproducing the known heads at the prescribed head boundaries, which equals the average of the squared differences between the network prediction and the known prescribed heads at the four chosen time steps

$$L_{BC} = \sum_{\Gamma} (h_{\text{predicted}} - h_{\text{specified}})^2 \tag{3.21}$$

where:

- the summation symbol $\sum_{\Gamma}$ represents the summation over spatial and temporal regions corresponding to the boundary conditions.

- $h_{\text{predicted}}$ is the predicted hydraulic head.

- $h_{\text{specified}}$ is the specified hydraulic head as a function of space and time.

4. Loss associated with the error in reproducing the no flow boundary. Using automatic differentiation (Griewank, 1998), it is possible to evaluate any partial differential of the artificial network output (h) with respect to the input variables (x, z, t); therefore, $\frac{\partial h}{\partial z}$ can be evaluated at the selected points along the bottom boundary, and the average squared difference with respect to its known value of zero computed for each of the chosen time steps.

   $$L_{\text{noflow}} = \sum_{\Omega} \left(f(x, z, t, h, \frac{\partial h}{\partial x}, \frac{\partial h}{\partial z}) \right)^2 \tag{3.22}$$

   In this expression:

   - Ω represents the spatial and temporal domain over which the no flow boundary conditions is solved.

   - h is the predicted solution by the neural network.

   - $f(x, z, t, h, \frac{\partial h}{\partial x}, \frac{\partial h}{\partial z})$ is the no flow residual, which depends on the predicted solution h and its derivatives with respect to x and z.

5. Loss associated with the error in reproducing the PDE. Again, thanks to automatic differentiation, and as displayed in Figure 3.33, the partial derivatives involved in (3.15) can be computed at the collocation points selected. After rearranging all terms in (3.15) so that they equal zero, the average squared sum of the PDE values computed with the heads provided as output from the neural network at the collocation points and the four selected times will represent the associated error.

   $$L_{\text{PDE}} = \sum_{\Omega} \left(f(x, z, t, h, \frac{\partial h}{\partial x}, \frac{\partial h}{\partial z}, \frac{\partial h}{\partial t}) \right)^2 \tag{3.23}$$

   In this expression:

   - Ω represents the spatial and temporal domain over which the PDE is solved.

- h is the predicted solution by the neural network.

- $f(x, z, t, h, \frac{\partial h}{\partial x}, \frac{\partial h}{\partial z}, \frac{\partial h}{\partial t})$ is the PDE residual, which depends on the predicted solution h and its derivatives with respect to x, z, and t.

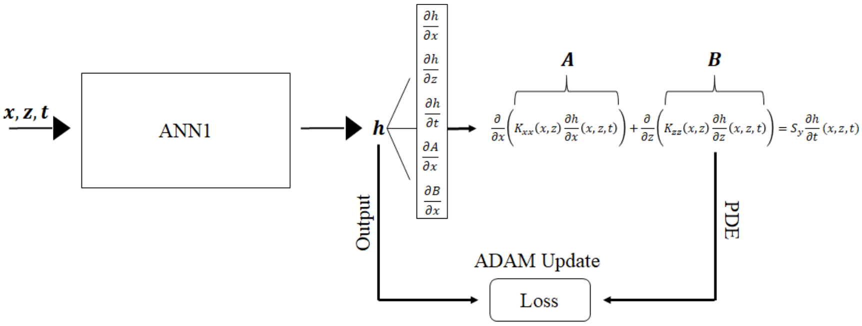


Figure 3.33. Sketch of ANN1.

3.4.3.2. Second neural network (ANN2)

The objective of ANN2 (Fiure 3.34) is to predict the spatial coordinate (z_s) corresponding to the phreatic surface, using (x, t) as input variables. The loss function is redefined as the summation of squared discrepancies between the predicted elevation z_s and the piezometric head forecasted at that specific location by ANN1, which is expected to match the actual elevation. This summation is calculated over 2500 points (SC1) and 3500 points (SC2), randomly distributed within the range [0, 1] and across the four selected time intervals. The loss function for ANN2 comprises two components:

1. Loss associated with the error in reproducing the phreatic surface, which equals the sum of the squared difference between the ANN2 predictions and the piezometric head predicted by ANN1.

$$L_{BC} = \sum_{\Gamma} \left(z_{s_{\text{predicted}}} - h_{\text{predicted}} \right)^2 \tag{3.24}$$

where:

- the summation symbol $\sum_{\Gamma}$ represents the summation over spatial and temporal regions corresponding to the prheatic surface prediction.

- $z_{s_{\text{predicted}}}$ is the predicted elevation of the phreatic surface by ANN2.

- $h_{\text{predicted}}$ is the predicted hydraulic head by ANN1.

2. Loss associated with the error in reproducing the initial conditions, which equals the sum of the squared differences between the network predictions and the known initial values for the phreatic surface.

$$L_{IC} = \sum_{\Gamma} \left(z_{s_{\text{predicted}}} - z_{s_{\text{initial}}} \right)^2 \tag{3.25}$$

where:

- the summation symbol $\sum_{\Gamma}$ represents the summation over space corresponding to the initial position of the phreatic surface.

- $z_{s_{\text{predicted}}}$ is the predicted phreatic surface elevation at time $t = 0$.

- $z_{s_{\text{initial}}}$ is the initial phreatic surface elevation as a function of space.

As previously mentioned, the training process begins with ANN1, followed by training ANN2 using the output from ANN1. Then, both networks are concurrently trained, utilizing a loss function that combines the components described for each network. We have found this sequential training approach to be more efficient than attempting to train both networks simultaneously from the outset. Given the inherent relationship between the elevation of the free surface, denoted as the output of ANN2, and the hydraulic head, represented as the output of ANN1, within

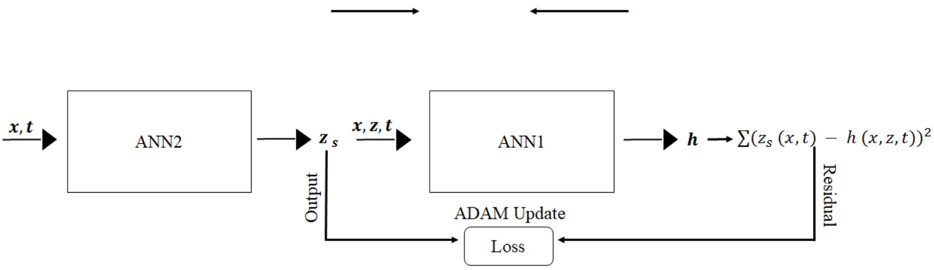


Figure 3.34. Sketch of ANN2.

an unconfined aquifer, it proves more effective to predict the phreatic surface elevation using ANN2 following the training of ANN1. This methodology optimizes computational processes by harnessing the knowledge acquired from the initial neural network, thereby enhancing predictive accuracy and efficiency, leading to smaller values of the final loss function.

3.4.4. Evaluation of the performance

The solution provided by MODFLOW will serve as the benchmark for evaluating the performance of the implemented PINN. To quantify the agreement between the fully trained PINN and the numerical model, several evaluation metrics are employed.

First, the Root Mean Squared Error (RMSE) is utilized to assess the overall fit:

$$RMSE = \sqrt{\sum_{i=1}^{N} \frac{(\hat{h_i} - h_i)^2}{N}}, \tag{3.26}$$

where N denotes the total number of verification points across both spatial and temporal dimensions. In this context, h_i represents the predicted values obtained from MODFLOW, while $\hat{h_i}$ corresponds to the predicted values produced by the

PINN.

Additionally, the Mean Absolute Error (MAE) is chosen as a reference metric due to its resilience to outliers:

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |\hat{h}_i - h_i|. \tag{3.27}$$

Furthermore, the Nash-Sutcliffe Efficiency (NSE) is employed as a goodness-of-fit measure:

$$NSE = 1 - \frac{\sum_{i=1}^{n} (h_i - \hat{h}i)^2}{\sum i = 1^n (h_i - \bar{h})^2}, \tag{3.28}$$

where $\bar{h}$ represents the mean of the MODFLOW-predicted values.

It is essential to bear in mind that the numerical solution provided by MODFLOW is itself an approximation of the underlying PDE and may not be exact. This is particularly evident in the representation of the phreatic surface, where MODFLOW employs a discretized approach, resulting in a less smooth representation compared to the meshless approximation offered by the PINN. Therefore, while these metrics serve as valuable tools for comparison, they should be interpreted with the understanding that both approaches are approximations, each with its unique characteristics and limitations.

3.4.5. Results

3.4.5.1. Training phase

As previously mentioned, the most efficient training strategy for both neural networks involves a sequential approach. Initially, ANN1 is trained independently for 200 epochs, followed by training ANN2 for 200 epochs while keeping ANN1's weights fixed. Subsequently, a joint training phase is carried out for an additional 200 epochs. The progression of the loss functions for both case studies is depicted in Figures 3.35 and 3.36, showcasing the loss function values and corresponding elapsed times on an Intel(R) Core(TM) i9-10920X CPU 3.50GHz with 32GB of RAM. In both scenarios, the final loss converges to approximately 10^{-5}, signifying a reduction of at least three orders of magnitude compared to the initial loss computed with random weight and bias initialization.

It is crucial to underscore that the available dataset predominantly comprises physics-based information, offering only a limited amount of prior knowledge for model training. Consequently, the risk of encountering overfitting issues, where the PINN memorizes the training data rather than generalizing effectively, is substantially mitigated. Furthermore, by assigning a significant portion of the loss function's emphasis to enforcing physical constraints and reserving a smaller fraction for data-driven training, we prioritize the model's capacity to capture the underlying physics while maintaining robust and dependable performance on unseen data, thus avoiding underfitting. This approach aligns seamlessly with the inherent characteristics of PINNs, showcasing their effectiveness in addressing intricate, physics-driven problems. Consequently, we have chosen to integrate all available prior information into the training data, forgoing the creation of a separate validation dataset.

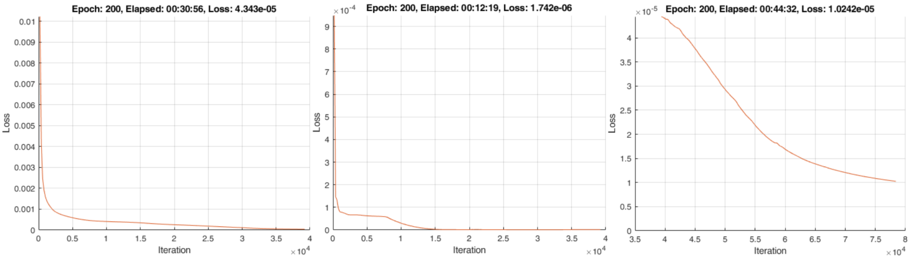


Figure 3.35. Scenario SC1. Training loss for ANN1 alone (left), ANN2 alone with ANN1 frozen (center) and ANN1 jointly with ANN2 (right). The iteration axis in the right plot starts at the number of iterations already performed in the previous training.

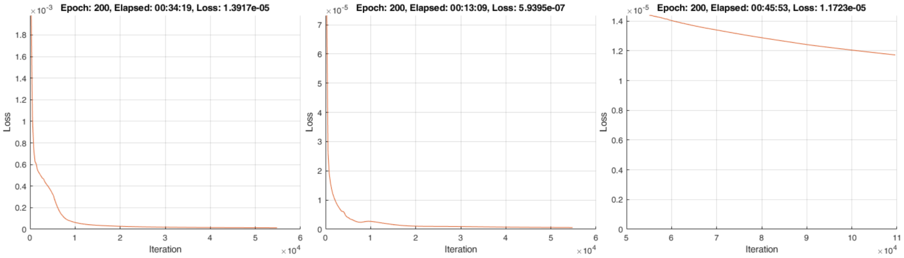


Figure 3.36. Scenario SC2. Training loss for ANN1 alone (left), ANN2 alone with ANN1 frozen (center) and ANN1 jointly with ANN2 (right). The iteration axis in the right plot starts at the number of iterations already performed in the previous training.

3.4.5.2. Testing phase

After the completion of network training, the validation process entails a thorough comparison between the network predictions and the outcomes derived from MODFLOW. Specifically, the Root Mean Squared Error (RMSE), the Mean Absolute Error (MAE) and the Nash-Sutcliffe Efficiency (NSE) are calculated for the piezometric heads at the central points of the discretization grid, as well as for the elevation of the phreatic surface at four distinct time instances: 0.01, 0.25, 0.5, and 1. To obtain predictions from the neural networks, coordinates (x, z, t) into ANN1 and (x, t) into ANN2 are considered. Additionally, a visual assessment is performed by comparing the piezometric head maps and representations of the phreatic surface.

Unconfined homogeneous isotropic aquifer (SC1)

Figure 3.37 visually illustrates the disparities between the predictions generated by the neural networks and those produced by MODFLOW at the four designated time points. Meanwhile, Table 3.27 provides quantitative insights by presenting the RMSE and MAE values calculated at the central locations of the active discretization cells.

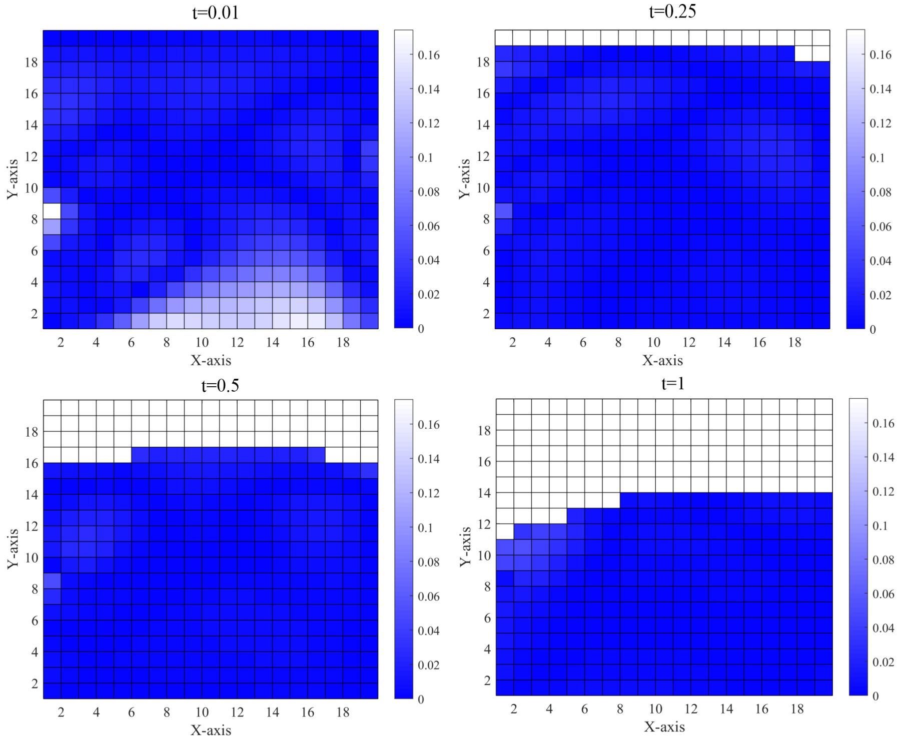


Figure 3.37. SC1: Error plot of the estimated piezometric field (PINN minus MODFLOW), using the active cells in the numerical model.

In general, the errors observed are minimal, with the most significant discrepancies arising at $t = 0.01$ when both models are simulating the abrupt reduction in prescribed head values along the boundaries. Subsequently, as the simulation progresses towards a steady-state condition, the errors diminish, as evidenced by the declining RMSE and MAE values. Table 3.27 further illustrates this strong agreement by presenting the Nash-Sutcliffe Efficiency (NSE) values, affirming the excellent alignment between the PINN predictions and the numerical solution.

Figure 3.38 illustrates the piezometric head maps at the specified time $t = 0.01$ as obtained from both the MODFLOW simulation and the PINN prediction. No-

Table 3.27. SC1: RMSE, MAE and NSE of the estimated solution by the PINN compared to the one obtained by the numerical model

Time	**RMSE**	**MAE**	**NSE**
0.01	0.0423	0.0239	0.90
0.25	0.0098	0.0070	0.99
0.50	0.0093	0.0064	0.99
1.00	0.0099	0.0053	0.98

tably, the MODFLOW maps are pixel-based representations aligned with the discretization scheme used for solving the equation, while the PINN maps exhibit a finer resolution, capitalizing on the meshless nature of the neural network. This distinction becomes particularly pronounced in Figure 3.39, where the delineation of the phreatic surface at the other three examined time points appears jagged in the MODFLOW solution but exhibits a smoother profile in the PINN results.

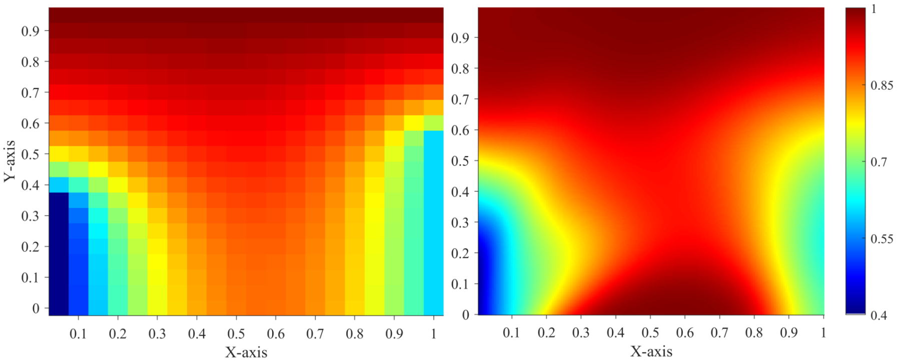


Figure 3.38. SC1: Estimated piezometric field by the numerical model (left) and PINN (right) for the selected time $t = 0.01$.

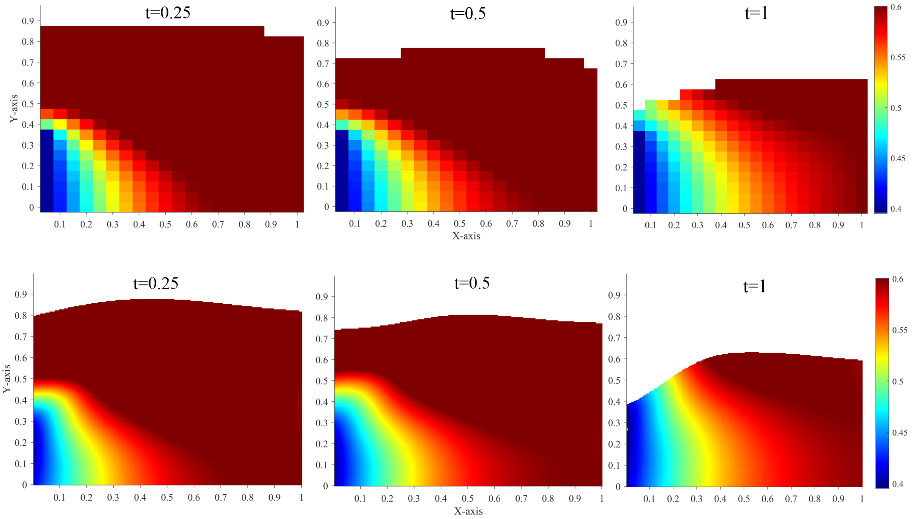


Figure 3.39. SC1: Estimated piezometric field by the numerical model (top) and PINN (bottom) for the selected time $t = 0.25$, $t = 0.5$, $t = 1$.

Unconfined heterogeneous anisotropic aquifer (SC2)

Figure 3.40 visually represents the disparities between the network predictions and the MODFLOW predictions at the four designated time points. For a quantitative evaluation of these differences, Table 3.28 present the computed RMSE and MAE values specifically calculated at the central positions of the active discretization cells.

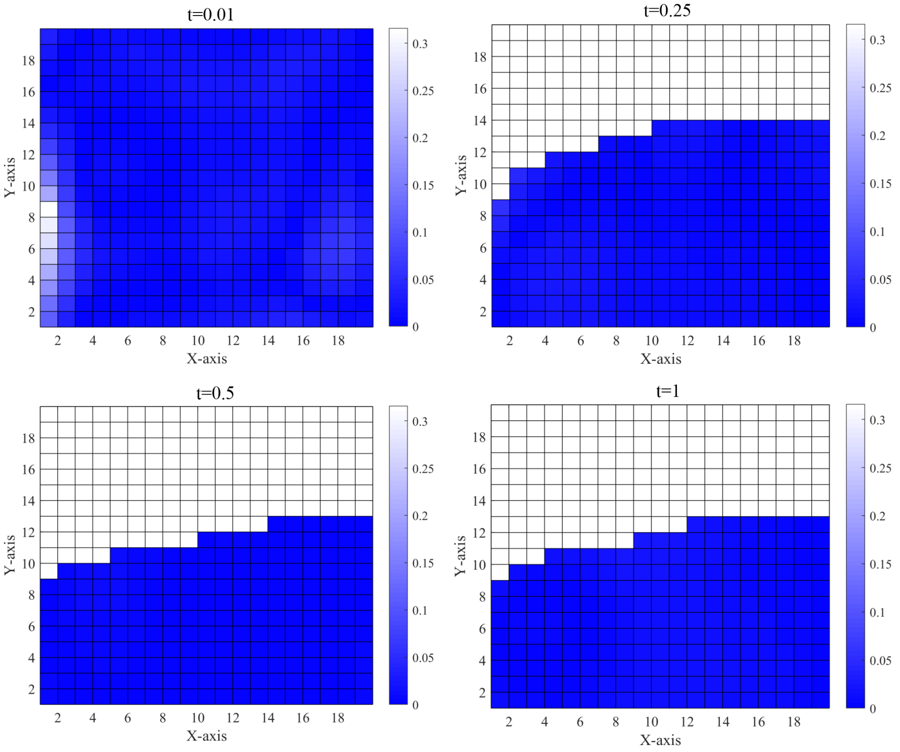


Figure 3.40. Heterogeneous aquifer: Error plot of the estimated piezometric field (PINN minus MODFLOW), using the active cells in the numerical model.

Generally, these errors are small, with the most significant deviations occurring at the initial starting time. This is expected since both models are simulating the abrupt changes along the boundaries at time zero. However, as the simulation progresses towards the steady-state condition, the errors gradually diminish. Additionally, Table 3.28 provides NSE values, emphasizing the strong correspondence between the PINN predictions and the numerical solution.

Table 3.28. SC2: RMSE, MAE and NSE of the estimated solution by the PINN compared to the one obtained by the numerical model

Time	**RMSE**	**MAE**	**NSE**
0.01	0.0437	0.0231	0.87
0.25	0.0136	0.0109	0.95
0.50	0.0122	0.0103	0.96
1.00	0.0047	0.0039	0.99

In Figure 3.41 and Figure 3.42, the piezometric head maps are presented at the selected times $t = 0.01$, $t = 0.25$, $t = 0.5$, and $t = 1$, generated from both the MODFLOW simulation and the PINN prediction. Just like in the homogeneous case, the piezometric heads obtained from the PINN and the numerical model exhibit comparable overall behavior and patterns.

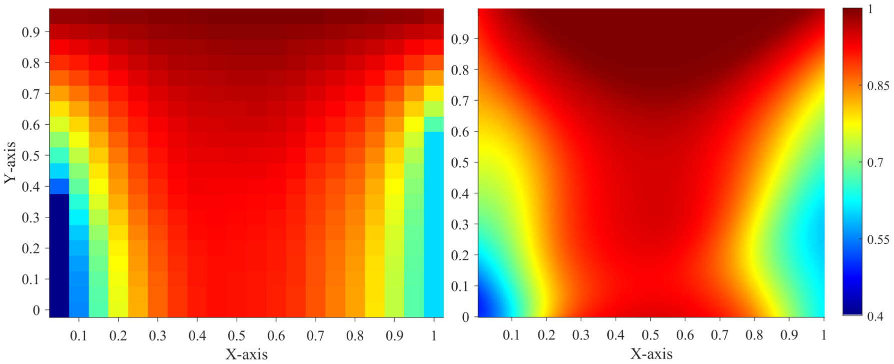


Figure 3.41. Heterogeneous aquifer: Estimated piezometric field by the numerical model (left) and PINN (right) for the selected time $t = 0.01$.

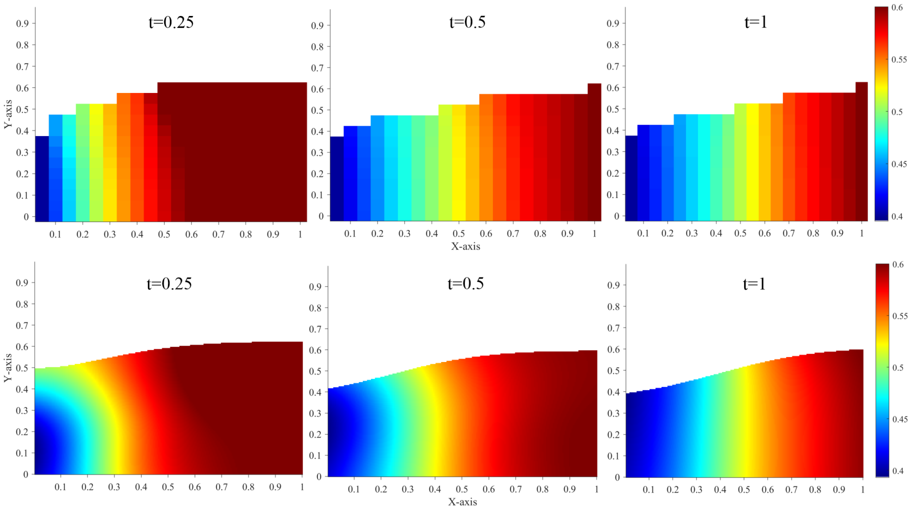


Figure 3.42. Heterogeneous aquifer: Estimated piezometric field by the numerical model (top) and PINN (bottom) for the selected time $t = 0.25$, $t = 0.5$, $t = 1$.

Heterogeneous aquifer (SC2): PINN vs conventional ANN

Drawing a more extensive comparison between the performance of PINNs and traditional neural networks is of particular interest. For this reason, in this section, the results obtained for SC2 using a MLP neural network (Section 1.3.2) are presented (Figure 3.43). Traditional neural networks exclusively depend on a priori information, specifically, the availability of known data. In the context of this specific case study, training a conventional neural network effectively would require a substantial volume of groundwater level data collected over time to serve as target data during network training. However, in this particular case, the known data is limited to prescribed boundary head conditions, initial conditions, and only a sparse dataset within the domain (comprising merely 10% of active cells from the MODFLOW simulation and only for four specific time steps). Even the impermeable boundary condition is incorporated into the loss function as physics constrain through automatic differentiation.

It is vital to emphasize that, despite being synthetic, the case study is physically complex. The abrupt alteration of boundary conditions initiates a transient flow behavior that is far from trivial. With the limited known data used for the training, by excluding the underlying physics the network lacks the necessary information for effective training and achieving high-performance results, as observed for the PINN in the metrics reported in Tables 3.28. Furthermore, training the network with sparse data within the domain is insufficient for establishing an appropriate training range that accurately represents the spatial extent of the domain.

Numerous studies in the field have shown that conventional neural networks are fundamentally incapable of extrapolating beyond their training range. Consequently, the utilization of PINNs, which leverage the underlying physics, is essential. This approach enables the creation of a network that can extrapolate based on physical principles, facilitating the development of a meshless model that delivers reliable and functional results in contrast to a conventional neural network.

The stark contrast is evident in Figure 3.43, where the output of the conventional neural network, trained solely with available a priori information, clearly fails to accurately reproduce the flow field.

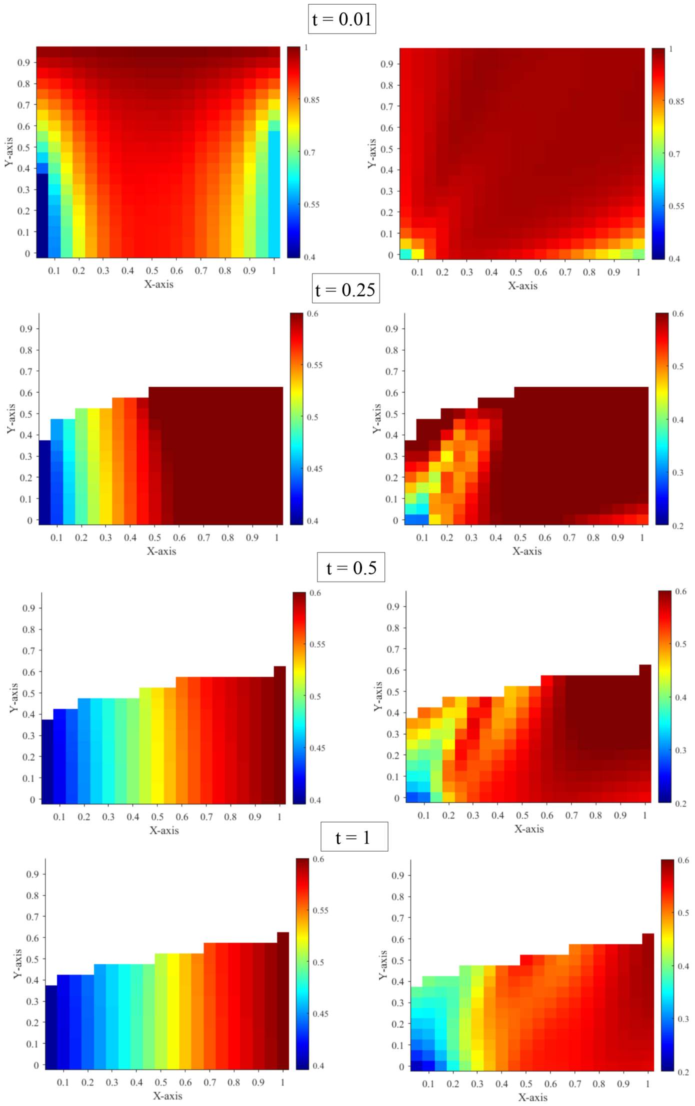


Figure 3.43. Heterogeneous aquifer. Left: numerical solution of the piezometric field with respect to the active cells at time $t = 0$, $t = 0.25$, $t = 0.5$ and $t = 1$. Right: conventional ANN prediction of the piezometric field with respect to the active cells at time $t = 0$, $t = 0.25$, $t = 0.5$ and $t = 1$.

3.4.6. Discussion and conclusions

This study presents a successful application of Physics-Informed Neural Networks (PINNs) to address forward groundwater flow problems in unconfined aquifers. It is important to note that the governing partial differential equation (PDE) for unconfined aquifer flow includes a space-time-varying boundary condition related to the phreatic surface's position, making it a particularly challenging problem. The following discussion highlights the main conclusions and implications of this research:

Firstly, our results affirm the capability of PINNs to accurately predict piezometric head values in unconfined aquifer systems. This accuracy is evident from the small errors observed between the PINN predictions and the numerical model results, especially as the simulation progresses towards a stationary condition (refer to Tables 3.27 and 3.28). These findings align with previous research that has showcased the effectiveness of PINNs in tackling complex problems in fields like fluid mechanics and geosciences (Waheed et al., 2021; Bajracharya and Jain, 2022; Cai et al., 2021; Mao et al., 2020; Lv et al., 2021; Zheng and Wu, 2023).

Secondly, it was highlighted that incorporating physical constraints into the training process can significantly reduce the need for extensive observations when using a conventional Artificial Neural Network (ANN). This underscores the potential of PINNs to operate effectively in scenarios with limited data availability. In contrast, traditional ANNs struggle to perform well under data-limited conditions, while PINNs can efficiently harness existing data and physics-based constraints to overcome data scarcity. This quality makes PINNs particularly valuable for applications in remote or data-scarce regions, as well as in situations where data collection is resource-intensive.

Furthermore, our findings suggest that PINNs have the potential to complement or even replace traditional numerical models in simulating unconfined aquifer flow problems. With their faster computation times and capacity to handle complex datasets, PINNs offer a promising alternative for modeling and simulating hydrogeological systems.

However, it is essential to acknowledge certain limitations. The accuracy of PINN results can be influenced by factors such as the neural network architecture and the quality of training data. Hyperparameters like the number of layers, neurons, and learning rates can exert a significant impact on results. While this study manually calibrated hyperparameters to evaluate PINN effectiveness in solving forward flow problems in unconfined aquifers, future research could explore sensitivity analysis or automated tools for optimizing these parameters. Additionally, PINNs require a substantial number of collocation points for evaluating physical constraints, which can extend the training phase, particularly when observational data is scarce.

Another crucial consideration is the potential impact of data errors on AI model outcomes. In the context of neural networks, this issue has been extensively studied, demonstrating that neural networks can be trained to accommodate measurement inaccuracies (Coppola et al., 2003; Secci et al., 2022). This approach involves introducing a range of measurement errors and perturbing individual data points accordingly. While this process demands more computational resources, it enhances the model's ability to handle and adapt to measurement errors, a feature that can be explored in future research to enhance model reliability.

Furthermore, future investigations could explore three-dimensional (3D) unconfined aquifer problems to assess whether the introduction of an additional input dimension (the y-coordinate) imposes limitations on performance or necessitates a significant increase in data and collocation points for effective training.

In conclusion, this study demonstrates the effectiveness of using PINNs to address unconfined aquifer flow problems, providing accurate estimates of time-varying phreatic surface and piezometric head values. PINNs offer an efficient and alternative approach to solving complex groundwater flow problems, with potential applications spanning environmental management, civil engineering, and hydrogeology. Future research can further explore the potential of PINNs in addressing other groundwater challenges, including contaminant transport, heterogeneity characterization, and anisotropy.

The methodology and the results obtained were submitted to an international journal and have been published (Secci et al., 2024).

4

Extra contents

This section of the thesis introduces additional contents. It begins with the application of a conventional neural network to solve an inverse problem in the field of sewer systems. Then, an easily interpretable example of numerical groundwater flow modeling using spreadsheets from a didactic perspective is presented.

The inclusion of additional materials in this work aim to extend the exploration of ANNs beyond the confines of the initial research domain and was motivated by the interest to uncover novel insights and potential applications that can enhance the overall depth and applicability of the study. Furthermore, the incorporation of diverse perspectives aimed to present alternative approaches to addressing the flow equation, ultimately contributing to a more holistic comprehension of the complexities inherent in solving the flow equation.

4.1. Identification of the inflow source in a foul sewer system using an artificial neural network as inverse model

Sewer infiltration, often stemming from pipe cracks or defective joints, alongside unauthorized inflow into sanitary sewer systems, can lead to a spectrum of issues. These include reduced treatment plant efficiency, more frequent overflows due to surcharged pipelines, and the necessity to expand pipeline capacities (Staufer et al., 2012; Beheshti and Sægrov, 2018). Integrated water service operators are increasingly tasked with addressing these challenges by mitigating infiltrations and pinpointing the origins of illegal inflows (of Massachusetts, 1993; , EPA). However, the conventional techniques employed in these scenarios tend to be costly, time-consuming, and often yield only partial results (Wittenberg and Aksoy, 2010; Rezaee and Tabesh, 2022). Typical approaches involve methods such as CCTV inspections, smoke tests, and the installation of extensive sensor networks for continuous data collection, encompassing parameters such as flow rates, water levels, temperatures (DTS), and pollutant concentrations (Beheshti and Sægrov, 2018; Vosse et al., 2013; Lepot et al., 2017; Beheshti and Sægrov, 2019; Panasiuk et al., 2019).

The primary objective of this study is to employ an inverse mathematical technique, by means of an artificial neural network, to identify the source location of illegal inflows into wastewater systems. This identification is based on information collected at the outlet of the drained basin, coupled with a well-calibrated numerical model of the sewer network. In this research, the numerical model is developed through the utilization of the Storm Water Management Model (SWMM) software, which is distributed by the Environmental Protection Agency in the United States (Rossman and Huber, 2016). A realistic foul sewer system, characterized by known dimensions and hydrological features, is considered as the basis for this investigation.

To test the inverse methodology, a synthetic case study is established. Assuming a hypothetical rainfall event and an unauthorized inflow introduced at a specified location within the sewer system, the numerical model is run to obtain the flow hydrograph at the network outlet. This simulated hydrograph data serves as observed measurements for the subsequent inverse modeling process.

4.1.1. Case study

In this study, the problem was addressed by creating synthetic scenarios applied to a real sewer network. The sewer system depicted in Figure 4.1 is the infrastructure serving the municipality of Polesine Parmense in northern Italy. Originally designed as a combined sewer, it was later modified to transport only sewage, necessitating the separation of rainwater from wastewater. This network is situated in a flat area within the Parma plain and comprises 122 junctions and 20 catchments connected to the network when operating in a mixed mode. The geometric information about the network were obtained through surveys and video inspections. Information regarding the basins potentially linked to the network was collected from inspections and aerial photo analysis.

For generating synthetic observations, such as flow hydrographs at the downstream end of the system in Figure 4.1, one-hour storms were considered and discretized into 15-minute intervals, with a maximum depth of 20 mm. In addition, potential inflows from the 20 basins previously connected to the network were factored, excluding simultaneous unauthorized entries from multiple basins (one basin at a time). The known hydrological characteristics of these basins was utilized. The daily wastewater discharges were determined through previous on-site campaigns. The hourly foul water pattern was assumed to align with the standard hourly aqueduct demand for the municipality in question.

As an example, Figure 4.2 depicts the flow hydrographs observed during the

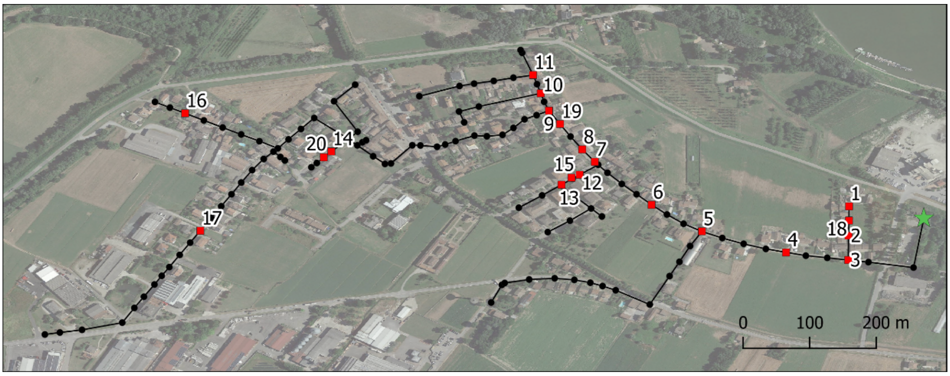


Figure 4.1. The sewer network plan for Polesine Parmense. The 122 junctions are denoted by black points, while the red squares mark the 20 junctions where the basins are connected, representing potential inflow locations. The green star indicates the downstream end of the network, where the flow hydrograph is observed.

first five hours at the network outlet during a storm event, considering unauthorized inflow from three possible basins. Notably, node 17 is the farthest from the outlet, resulting in a minimal peak discharge at the downstream end due to damping effects. Conversely, the unauthorized inflow from the closest node (3) leads to the highest peak discharge.

The calculations were carried out in a MATLAB environment (MathWorks, 2022), coupled with SWMM for simulating the sewer network. For the inverse method, the identification of inflow considering all 20 nodes shown in Figure 4.1 was examined, connected to their respective sub-basins, but without simultaneous entries.

4.1.2. Set up of the ANN

The implemented neural network takes the form of an MLP (Section 1.3.2). This MLP model comprises four layers: input, two hidden, and output layers. The input layer contains 2,881 neurons, representing the 24-hour downstream flow

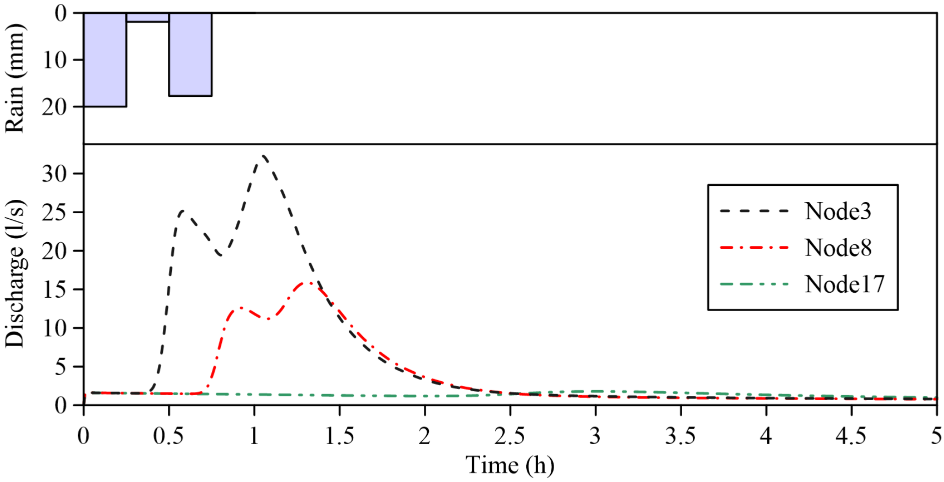


Figure 4.2. Flow hydrographs, observed at the network outlet (bottom), in relation to a precipitation event (top) and the unauthorized inflow from various nodes.

hydrograph discretized at 30-second intervals. The two hidden layers have 50 and 70 neurons, and the output layer is composed of 20 neurons. These 20 neurons correspond to the number of junctions within the sewer system that are under consideration for potential illegal inflow locations. An hyperbolic tangent was selected as activation function for the input and hidden layers, while, since it is a classification problem, the softmax activation function was used between the hidden and output layers. The softmax activation function is commonly used in neural networks for multi-class classification problems. It is used to transform the raw output scores of a neural network into a probability distribution over multiple classes. The softmax function takes a vector of real numbers as input and converts these numbers into values between 0 and 1, where the sum of all the values is equal to 1. Mathematically, the softmax function is defined as follows for an input vector Z:

$$\text{softmax}(Z)_i = \frac{e^{Z_i}}{\sum_{j=1}^{K} e^{Z_j}} \tag{4.1}$$

where Z_i represents the raw score for class i and K is the total number of classes.

The loss function for the neural network, denoted as $L(\Theta)$, where $\Theta = (\theta_1, \ldots, \theta_n)$ encompasses all network parameters, including weights and biases. It is defined as presented in Eq. 1.4.

The neural network underwent training, validation, and testing using a synthetic dataset created from 4,000 forward simulations of SWMM. The flow hydrographs, which served as inputs for the neural network, were derived from 200 storms, each lasting one hour and divided into four intervals of 15 minutes, with rainfall depths ranging from 0 to 20 mm. These hyetographs were generated through a stochastic technique known as Latin Hypercube Sampling which generates random variables from a multidimensional distribution while minimizing correlations between them.

The runoff resulting from these 200 synthetic storms was introduced individually into the 20 potential inflow locations, resulting in a dataset of 4,000 flow hydrographs. Each flow hydrograph was associated with an output vector of 20 values. The actual inflow location was marked with a value of 1, while the remaining 19 nodes were assigned a value of 0.

The flow hydrographs produced by SWMM were intentionally corrupted by adding random errors to simulate real-world conditions. These errors were sampled from a normal distribution with a mean of zero and a variance of $1 \times 10^{-3}\ \left(m^3/s\right)^2$, which was chosen to mimic plausible instrument errors. A dataset of 120,000 elements was created by artificially corrupting each of the 4,000 flow hydrographs. Although computationally intensive, this method enables the network to learn how errors affect individual flow hydrographs and helps to accurately classify the inflow location. The final dataset was partitioned into training (70%), validation (15%), and test (15%) subsets.

4.1.3. Results

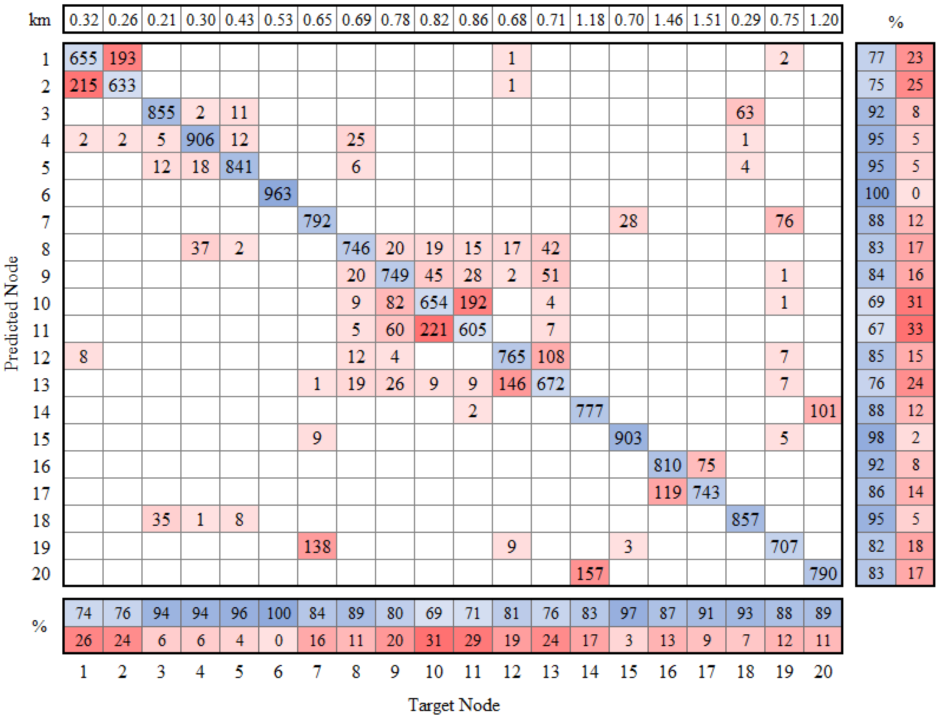
Figure 4.3 illustrates the confusion matrix obtained during the testing phase of the network after completing the training and validation processes. In Figure 4.3, the values outside the diagonal in the rows represent false positives (e.g., for junction 1, the network correctly recognized the inflow location 655 times, but 193 times junction 1 was mistakenly identified instead of node 2, and occasionally substituted nodes 12 and 19 once or twice, respectively). The values outside the diagonal in the columns represent false negatives (e.g., 215 times junction 2 was mistakenly identified instead of node 1).

The network correctly classified the inflow location in 86% of the test simulations. The classification error during training is as low as 0.06%, and the error percentages for the validation and test phases are both 0.14%. Examining Figure 4.3, it is evident that when the network does misclassify the desired output, it often selects a node that is in close proximity to the correct one. However, in a few cases, the prediction deviates significantly from the actual location.

4.1.4. Discussion and conclusions

In this study, a machine learning procedure based on an artificial neural network is applied to solve the inverse problem of identifying the source location illegal inflow of rainwater into a foul sewer system using flow observations recorded at the downstream of the sewer network. Overall, the artificial intelligence model has demonstrated good results even in the presence of errors in the observations.

The main advantage of ANN is its adaptability to address diverse problem configurations. Once adequately trained, it serves as a data-driven model that no longer requires the complete model. However, the training, validation, and



km	0.32	0.26	0.21	0.30	0.43	0.53	0.65	0.69	0.78	0.82	0.86	0.68	0.71	1.18	0.70	1.46	1.51	0.29	0.75	1.20	%	
1	655	193										1							2		77	23
2	215	633										1									75	25
3			855	2	11													63			92	8
4	2	2	5	906	12			25										1			95	5
5			12	18	841			6										4			95	5
6						963															100	0
7							792								28				76		88	12
8				37	2			746	20	19	15	17	42								83	17
9								20	749	45	28	2	51						1		84	16
10								9	82	654	192		4						1		69	31
11								5	60	221	605		7								67	33
12	8							12	4			765	108						7		85	15
13							1	19	26	9	9	146	672						7		76	24
14											2			777						101	88	12
15							9								903				5		98	2
16																810	75				92	8
17																119	743				86	14
18			35	1	8													857			95	5
19							138					9			3				707		82	18
20														157						790	83	17
%	74	76	94	94	96	100	84	89	80	69	71	81	76	83	97	87	91	93	88	89		
	26	24	6	6	4	0	16	11	20	31	29	19	24	17	3	13	9	7	12	11		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20		

Predicted Node

Target Node

Figure 4.3. Confusion matrix for the test phase results of the ANN. The distance of each node from the network outlet is provided at the top of the confusion matrix.

testing phases are computationally intensive. In this case, achieving the correct calibration of the neural network required 4,000 simulations of the SWMM model.

The method may encounter the problem of multiple solutions, a common challenge in inverse problems, where several solutions can provide a good fit to the observations used to address the inverse problem.

It is important to note that this study specifically focuses on point inflows among the various undesirable water sources in sewer systems. Future research will explore the application of these approaches to real-world scenarios, extending the scope beyond the identification of entry points to consider the extent of the basin area responsible for unauthorized inflow, which is often unknown.

The implemented methodology and results was published in Journal of Physics: Conference Series (Tanda et al., 2023)

4.2. Enhancing user-friendliness: a comprehensive and accessible example of numerical groundwater flow modeling using spreadsheets

The use of spreadsheets to solve groundwater flow equations is not new, dating back to the pioneering work of Olsthoorn (1985). Since then, many research papers have been published using spreadsheets for both steady-state and transient groundwater modeling (Akhter et al., 2006; Ankor and Tyler, 2019; Bair and Lahm, 2006; Bhattacharjya, 2011; Elfeki and Bahrawi, 2015; Fox, 1996; Karahan and Ayvaz, 2005a,b; Molano, 2014; Niazkar and Afzali, 2015; Olsthoorn, 1999). However, even if spreadsheet-based numerical groundwater flow modeling is a well-known approach, its potential in education has not been fully realized. In this part of the thesis, a new supplementary educational tool complements a previous one (Gómez-Hernández, 2022). Gómez-Hernández (2022) presented a spreadsheet-based solution for the partial differential equation describing steady-state flow in a single-layer confined aquifer. Readers are encouraged to review that paper to understand the foundational structure of the workbooks, which remains consistent in this work.

This study extends the applicability from solving numerical problems for single-layer confined aquifers (Gómez-Hernández, 2022) to encompass unconfined aquifers: horizontal layers and multi-layered vertical cross-sections. The updated spreadsheets can also handle hydraulic conductivity anisotropy, irregular aquifer shapes, all three types of boundary conditions, and pumping and recharge in steady-state flow regimes. This seemingly minor conceptual change has a significant impact on the numerical solution of the flow problem. The work describes how this change is addressed, and the findings are extended to analyze seepage through an earth dam. The research concludes with a discussion of possible extensions and limitations.

These advanced features make the spreadsheet model a versatile tool for simulating various steady-state groundwater flow scenarios. This user-friendly platform allows students to experiment with diverse configurations, explore the underlying mathematical expressions, and gain a deeper understanding of more sophisticated numerical codes such as MODFLOW.

The work includes examples of the three common types of boundary conditions, including the challenging phreatic surface boundary condition in unconfined flow scenarios. All cases consider steady-state flow. These examples were prepared using Excel spreadsheets and tested in Google spreadsheets, with Excel being significantly faster.

4.2.1. Unconfined two-dimensional groundwater flow in a horizontal plane

A quasi-rectangular unconfined aquifer (Figure 4.4), crossed by a river and connected to a lake, is studied. The aquifer has three distinct hydraulic conductivity zones, three pumping wells, and uniform infiltration. The aquifer is assumed to be in steady state.

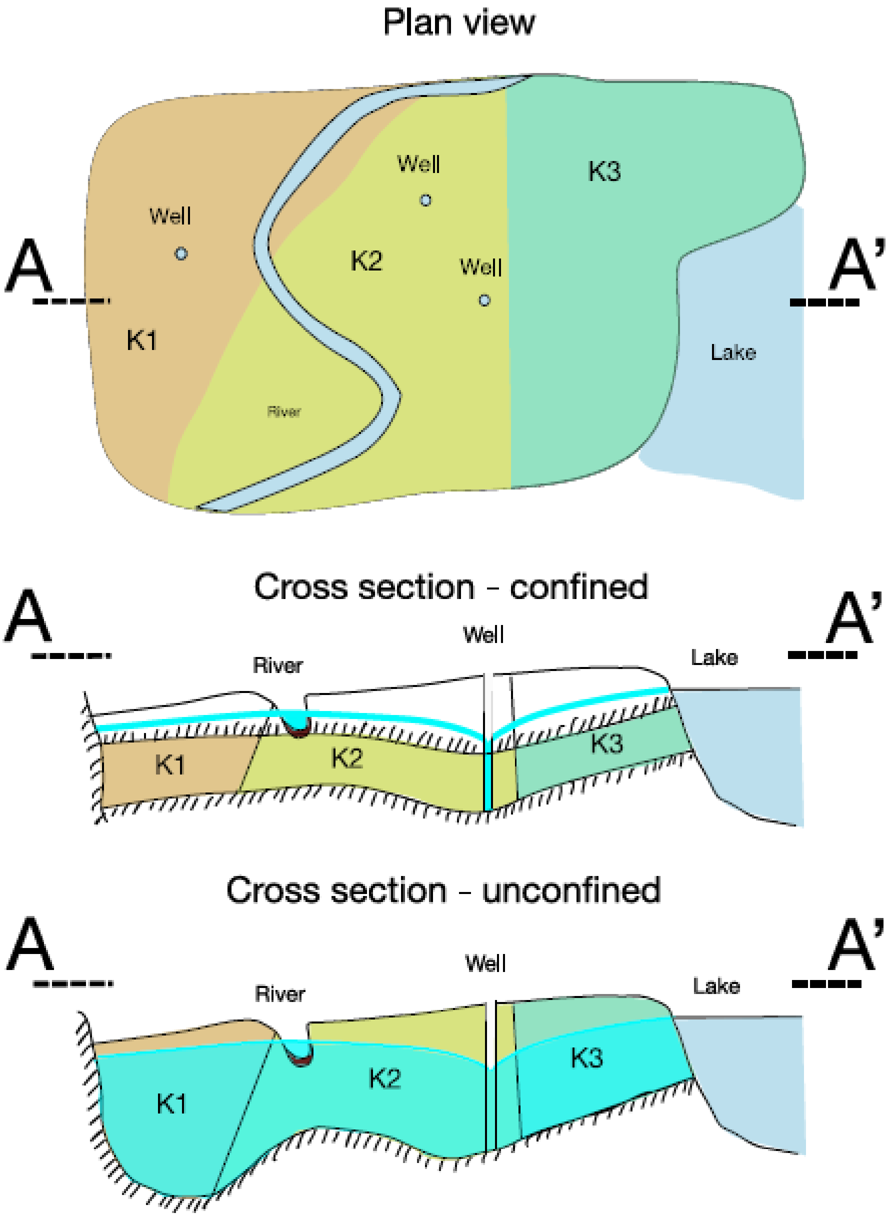


Figure 4.4. Confined and unconfined aquifer sketch

For the sake of brevity, the equation governing the hydraulic head for a generic cell and its derivation can be found in Section 2 of Gómez-Hernández (2022) and is not reported here. In the confined aquifer application (Gómez-Hernández, 2022), looking at Figure 4.4 the saturated thickness of the confined aquifer is independent of the piezometric head. Consequently, the transmissivity, which is equal to the product of conductivity and saturated thickness, can be determined beforehand and was indeed part of the input data set. However, the saturated thickness of the unconfined aquifer is unknown and is a function of the piezometric head. The saturated thickness in the unconfined aquifer is equal to the difference between the piezometric head and the aquifer bottom elevation. Fig. 4.5 shows the domain discretization: 19 rows by 33 columns of 100 m by 100 m cells.

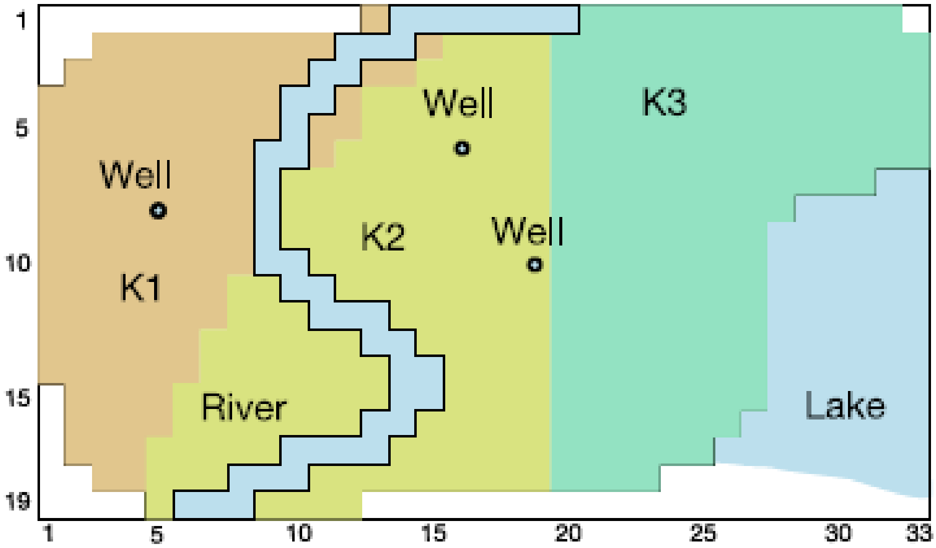


Figure 4.5. Horizontal plane aquifer sketch

The dependency of transmissivity on the saturated thickness, which is unknown a priori and depends on the piezometric head, yields the problem highly nonlinear.

This aspect complicates the iterative solution process and can lead to convergence instability. This is particularly challenging because some cells may dry out during iterations, necessitating a mechanism for rewetting them in subsequent iterations if their surrounding cells remain wet. In this regard, there are some factors to consider. First, the input data for the unconfined aquifer problem should be hydraulic conductivity, not transmissivity. At each iteration of the numerical solution, the saturated thickness of each cell must be computed as a function of the piezometric head from the previous iteration. The transmissivities must then be recomputed as the product of hydraulic conductivity and saturated thickness for each cell. This sensitivity to the initial heads makes the solution particularly challenging to converge. If the initial head values are far from the actual solution, unrealistic transmissivities will result, leading to unrealistic head estimates and potentially preventing convergence. In addition, a mechanism must be implemented to deactivate cells when their saturated thickness becomes zero, effectively treating them as dry cells. Conversely, a mechanism must also be implemented to reactivate or rewet inactive cells when the hydraulic head estimates in the surrounding cells are above the aquifer bottom of the inactive cell. One purely technical yet crucial consideration for achieving the final solution is related to the fact that the iterative nature of the numerical solution necessitates the use of the recalculate option of Microsoft Excel to iterate on cells containing circular references. Practical endeavors led to the realization that simply adding new spreadsheets for saturated thickness and transmissivity calculations was not sufficient. Instead, the order of calculations must be enforced, with saturated thickness being calculated first, followed by transmissivity. However, determining the order of spreadsheet calculations was not immediately apparent.It was observed that computations are performed in alphabetical order of spreadsheet names (Williams, 2001). Below are summarized all the modifications required for the unconfined problem compared to the confined one (Gómez-Hernández, 2022).

- Spreadsheet names should be altered to ensure their alphabetical order aligns with the intended calculation sequence;
- A spreadsheet containing cell bottom elevations should be incorporated;
- Two spreadsheets should be added for conductivity data:
    - One spreadsheet containing conductivity tensor values along the x-axis, used for calculating transmissivities TW and TE;
    - Another spreadsheet containing conductivity tensor values along the y-axis, used for calculating transmissivities TN and TS;
- A spreadsheet should be added for computing saturated thickness;
- The transmissivity spreadsheets should be converted to calculated spreadsheets, computed after determining the saturated thickness;
- During the iterative process, calculated heads may fall below the cell bottom, rendering the cell dry. However, if surrounding cell heads are above the cell bottom in the next iteration, it indicates that the cell should be partially saturated. Therefore, a mechanism is needed to enable the drying and rewetting of any cell. This requires two new spreadsheets:
    - One spreadsheet containing a flag indicating whether a cell should be wet based on the heads in surrounding cells;
    - Another spreadsheet containing the head value that should be assigned to cells that were dry but should be wet because the surrounding cells are above the cell bottom;
- An intermediary binary spreadsheet containing active cells after the last iteration is also needed due to the possibility of cells drying out and becoming inactive during the iteration process;

- Infiltration due to recharge should only be calculated for wet cells;

Following these modifications, the spreadsheet structure walkthrough proceeds as follows:

Input parameters (9 sheets):

- A_i: active cells;
- A_Bot: bottom elevation $[L]$;
- A_hfix: prescribed heads $[L]$;
- A_Kx: conductivities along the x-axis $[LT^{-1}]$;
- A_Ky: conductivities along the y-axis $[LT^{-1}]$;
- A_W: pumping well extraction $[L^3T^{-1}]$ (negative if injection);
- A_hR: river stage $[L]$;
- A_hB: elevation of riverbed bottom $[L]$;
- A_R: riverbed conductance $[L^2T^{-1}]$;

Intermediate variables (14 sheets):

- C_i: active cells after the last iteration;
- C_Sat: saturated thickness $[L]$;
- D_wet: Boolean variable indicating whether a cell should be wet or not as a function of the hydraulic heads in the surrounding cells;
- D_Rewet: value of the hydraulic head to be used only for those cells that became dry but should rewet because D_wet is True;
- D_Tx: transmissivity along the W-E direction $[L^2T^{-1}]$;

- D_Ty: transmissivity along the S-N direction $[L^2T^{-1}]$;
- D_Qriv: river recharge flow $[L^3T^{-1}]$;
- D_Qn: infiltration flow $[L^3T^{-1}]$;
- E_TN: transmissivity at the interface with the north cell $[L^2T^{-1}]$, computed as the harmonic average of adjacent y transmissivities;
- E_TS: transmissivity at the interface with the south cell $[L^2T^{-1}]$, computed as the harmonic average of adjacent y transmissivities;
- E_TW: transmissivity at the interface with the west cell $[L^2T^{-1}]$, computed as the harmonic average of adjacent x transmissivities;
- E_TE: transmissivity at the interface with the east cell $[L^2T^{-1}]$, computed as the harmonic average of adjacent x transmissivities;
- E_TZ: sum of E_{TN}, E_{TS}, E_{TW}, and E_{TE}. (Note that the name has been changed from TT to TZ to ensure that it is computed after the four summands have been updated.);

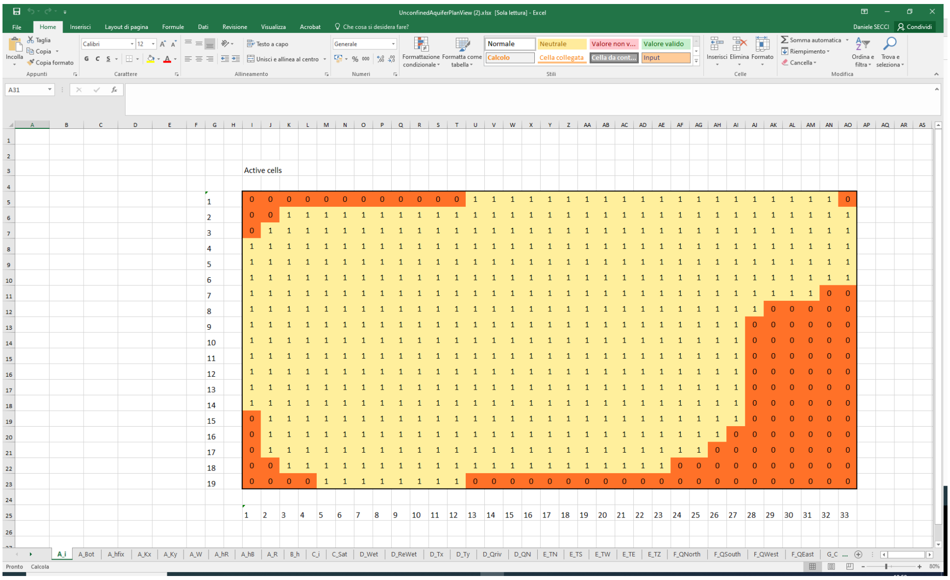
Final results (6 sheets):

- B_h: final results with the hydraulic heads in the aquifer $[L]$;
- F_QNorth: flows through cell north boundaries $[L^3T^{-1}]$;
- F_QSouth: flows through cell south boundaries $[L^3T^{-1}]$;
- F_QWest: flows through cell west boundaries $[L^3T^{-1}]$;
- F_QEast: flows through cell east boundaries $[L^3T^{-1}]$;
- G_CellBal: residual flow in the active cells (it should be zero if convergence has been achieved) $[L^3T^{-1}]$;

The presented example utilizes meters and days as the reference units for all variables. The step-by-step process is outlined below and is available in the GitHub repository https://github.com/jaumegomez/GroundwaterFlowByExcel that contains all the spreadsheets discussed:

Step 1: Define the input parameters

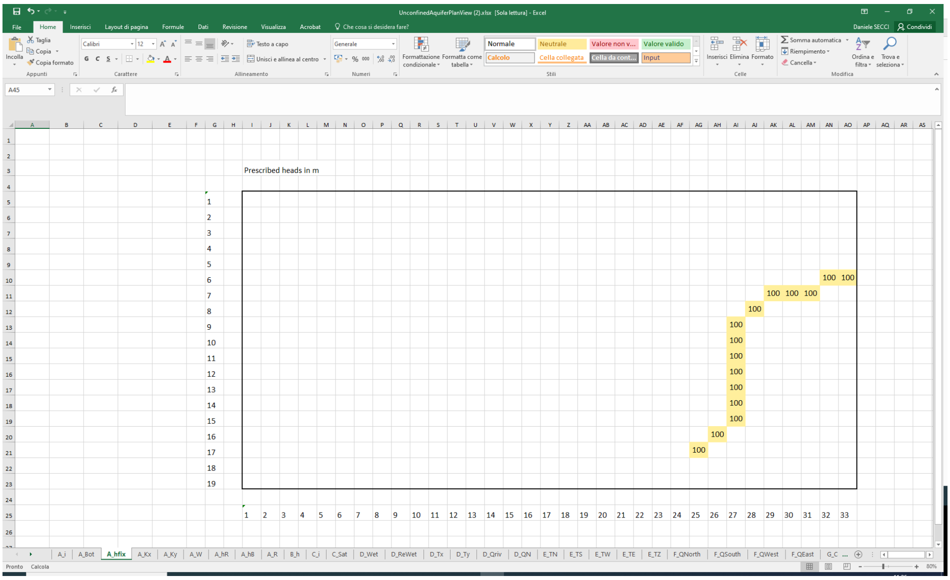
The active cells and the prescribed cells are shown in Figures 4.6 and 4.7.



Active cells

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33
1	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
2	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0
9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0
10	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0
11	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0
12	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0
13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0
14	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0
15	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0
16	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0
17	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
18	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
19	0	0	0	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 4.6. Active cells

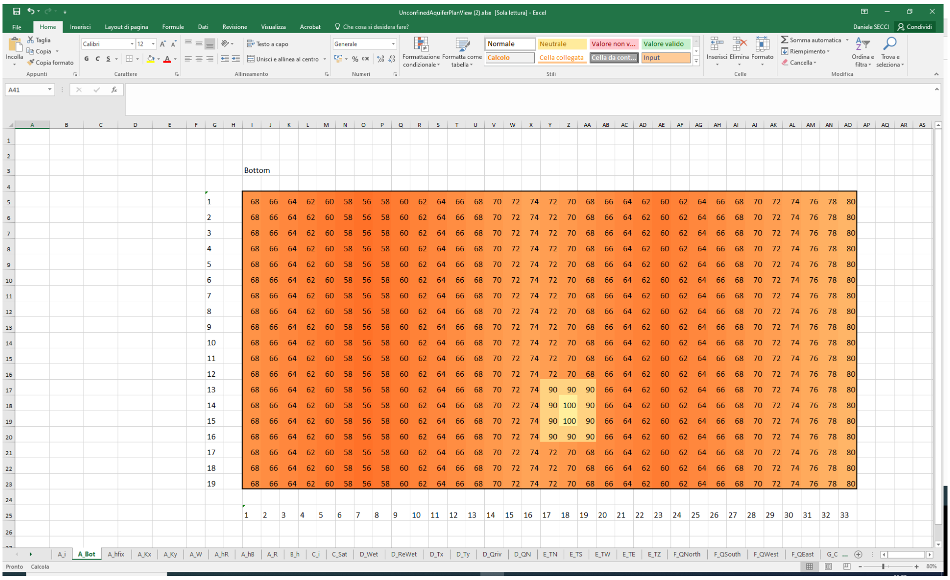


Prescribed heads in m

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33
1																																	
2																																	
3																																	
4																																	
5																																	
6																																100	100
7																													100	100	100		
8																												100					
9																											100						
10																											100						
11																											100						
12																											100						
13																											100						
14																											100						
15																											100						
16																										100							
17																									100								
18																																	
19																																	

Figure 4.7. Prescribed heads

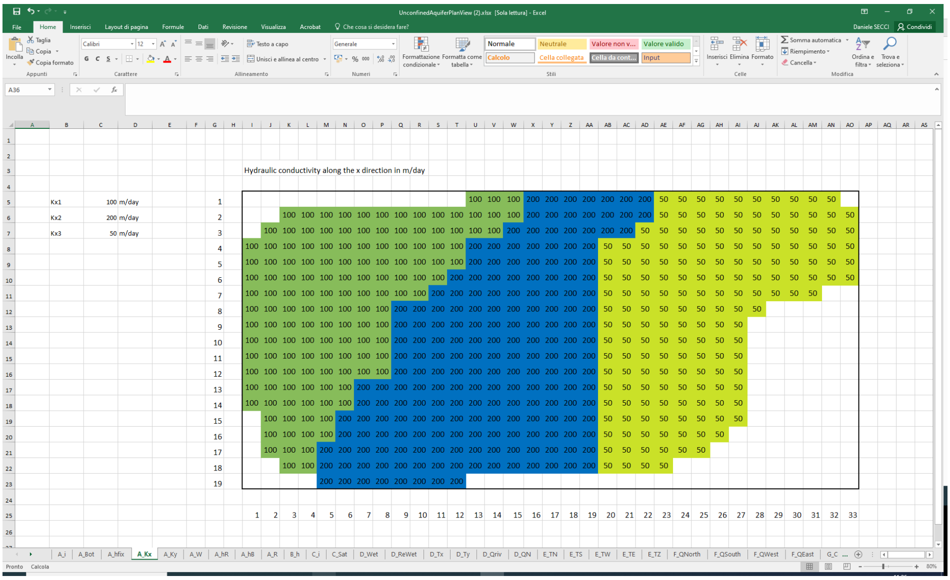
In contrast to the confined aquifer application, this model takes into account cell bottom elevation (Figure 4.8) and hydraulic conductivities (Figures 4.9 and 4.10), which exhibit anisotropic behavior, with conductivity along the W-E direction being ten times higher than that along the S-N direction. The cell bottom elevation features an undulating profile with elevations ranging from 68 m to 80 m, including a bump reaching 100 m towards the center of the aquifer to demonstrate the methodology's new drying capabilities. While the hydraulic conductivities follow the same zonation as in the confined case, their values have been adjusted to yield calculated transmissivities of similar magnitude.



Bottom

1	68	66	64	62	60	58	56	58	60	62	64	66	68	70	72	74	72	70	68	66	64	62	60	62	64	66	68	70	72	74	76	78	80
2	68	66	64	62	60	58	56	58	60	62	64	66	68	70	72	74	72	70	68	66	64	62	60	62	64	66	68	70	72	74	76	78	80
3	68	66	64	62	60	58	56	58	60	62	64	66	68	70	72	74	72	70	68	66	64	62	60	62	64	66	68	70	72	74	76	78	80
4	68	66	64	62	60	58	56	58	60	62	64	66	68	70	72	74	72	70	68	66	64	62	60	62	64	66	68	70	72	74	76	78	80
5	68	66	64	62	60	58	56	58	60	62	64	66	68	70	72	74	72	70	68	66	64	62	60	62	64	66	68	70	72	74	76	78	80
6	68	66	64	62	60	58	56	58	60	62	64	66	68	70	72	74	72	70	68	66	64	62	60	62	64	66	68	70	72	74	76	78	80
7	68	66	64	62	60	58	56	58	60	62	64	66	68	70	72	74	72	70	68	66	64	62	60	62	64	66	68	70	72	74	76	78	80
8	68	66	64	62	60	58	56	58	60	62	64	66	68	70	72	74	72	70	68	66	64	62	60	62	64	66	68	70	72	74	76	78	80
9	68	66	64	62	60	58	56	58	60	62	64	66	68	70	72	74	72	70	68	66	64	62	60	62	64	66	68	70	72	74	76	78	80
10	68	66	64	62	60	58	56	58	60	62	64	66	68	70	72	74	72	70	68	66	64	62	60	62	64	66	68	70	72	74	76	78	80
11	68	66	64	62	60	58	56	58	60	62	64	66	68	70	72	74	72	70	68	66	64	62	60	62	64	66	68	70	72	74	76	78	80
12	68	66	64	62	60	58	56	58	60	62	64	66	68	70	72	74	72	70	68	66	64	62	60	62	64	66	68	70	72	74	76	78	80
13	68	66	64	62	60	58	56	58	60	62	64	66	68	70	72	74	90	90	90	66	64	62	60	62	64	66	68	70	72	74	76	78	80
14	68	66	64	62	60	58	56	58	60	62	64	66	68	70	72	74	90	100	90	66	64	62	60	62	64	66	68	70	72	74	76	78	80
15	68	66	64	62	60	58	56	58	60	62	64	66	68	70	72	74	90	100	90	66	64	62	60	62	64	66	68	70	72	74	76	78	80
16	68	66	64	62	60	58	56	58	60	62	64	66	68	70	72	74	90	90	90	66	64	62	60	62	64	66	68	70	72	74	76	78	80
17	68	66	64	62	60	58	56	58	60	62	64	66	68	70	72	74	72	70	68	66	64	62	60	62	64	66	68	70	72	74	76	78	80
18	68	66	64	62	60	58	56	58	60	62	64	66	68	70	72	74	72	70	68	66	64	62	60	62	64	66	68	70	72	74	76	78	80
19	68	66	64	62	60	58	56	58	60	62	64	66	68	70	72	74	72	70	68	66	64	62	60	62	64	66	68	70	72	74	76	78	80
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33

Figure 4.8. Bottom elevations. Promontory located at the center of the lower half, which is expected to remain dry

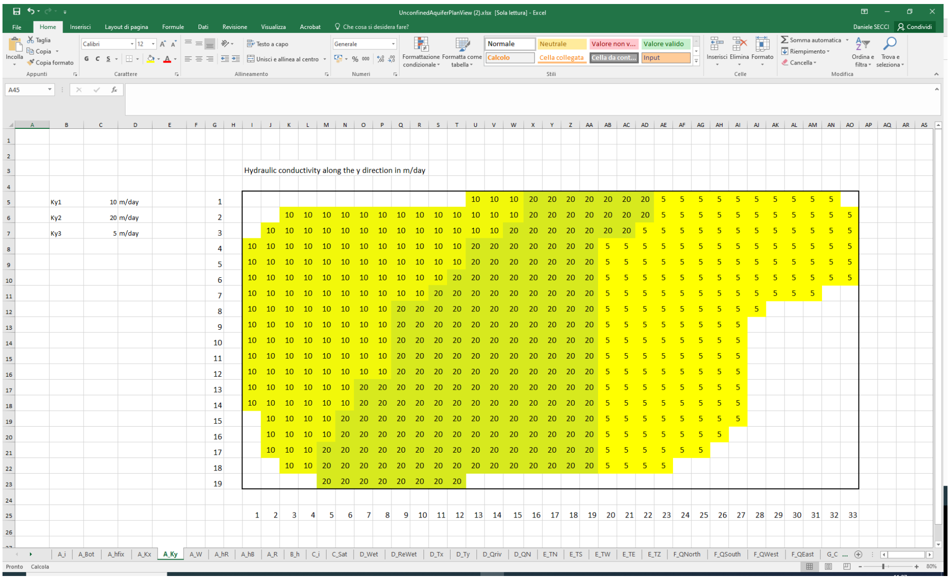


Hydraulic conductivity along the x direction in m/day

Kx1 100 m/day
Kx2 200 m/day
Kx3 50 m/day

1													100	100	100	200	200	200	200	200	200	200	50	50	50	50	50	50	50	50	50	50	
2			100	100	100	100	100	100	100	100	100	100	100	100	100	200	200	200	200	200	200	200	50	50	50	50	50	50	50	50	50	50	50
3		100	100	100	100	100	100	100	100	100	100	100	100	100	200	200	200	200	200	200	200	50	50	50	50	50	50	50	50	50	50	50	50
4	100	100	100	100	100	100	100	100	100	100	100	100	200	200	200	200	200	200	200	50	50	50	50	50	50	50	50	50	50	50	50	50	50
5	100	100	100	100	100	100	100	100	100	100	100	100	200	200	200	200	200	200	200	50	50	50	50	50	50	50	50	50	50	50	50	50	50
6	100	100	100	100	100	100	100	100	100	100	100	200	200	200	200	200	200	200	200	50	50	50	50	50	50	50	50	50	50	50	50	50	50
7	100	100	100	100	100	100	100	100	100	100	200	200	200	200	200	200	200	200	200	50	50	50	50	50	50	50	50	50	50	50	50		
8	100	100	100	100	100	100	100	100	200	200	200	200	200	200	200	200	200	200	200	50	50	50	50	50	50	50	50	50					
9	100	100	100	100	100	100	100	100	200	200	200	200	200	200	200	200	200	200	200	50	50	50	50	50	50	50	50						
10	100	100	100	100	100	100	100	100	200	200	200	200	200	200	200	200	200	200	200	50	50	50	50	50	50	50	50						
11	100	100	100	100	100	100	100	100	200	200	200	200	200	200	200	200	200	200	200	50	50	50	50	50	50	50	50						
12	100	100	100	100	100	100	100	100	200	200	200	200	200	200	200	200	200	200	200	50	50	50	50	50	50	50	50						
13	100	100	100	100	100	100	200	200	200	200	200	200	200	200	200	200	200	200	200	50	50	50	50	50	50	50	50						
14	100	100	100	100	100	100	200	200	200	200	200	200	200	200	200	200	200	200	200	50	50	50	50	50	50	50	50						
15		100	100	100	100	200	200	200	200	200	200	200	200	200	200	200	200	200	200	50	50	50	50	50	50	50	50						
16		100	100	100	100	200	200	200	200	200	200	200	200	200	200	200	200	200	200	50	50	50	50	50	50	50							
17		100	100	100	200	200	200	200	200	200	200	200	200	200	200	200	200	200	200	50	50	50	50	50	50								
18			100	100	200	200	200	200	200	200	200	200	200	200	200	200	200	200	200	50	50	50	50										
19					200	200	200	200	200	200	200	200																					
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33

Figure 4.9. W-E directional hydraulic conductivities



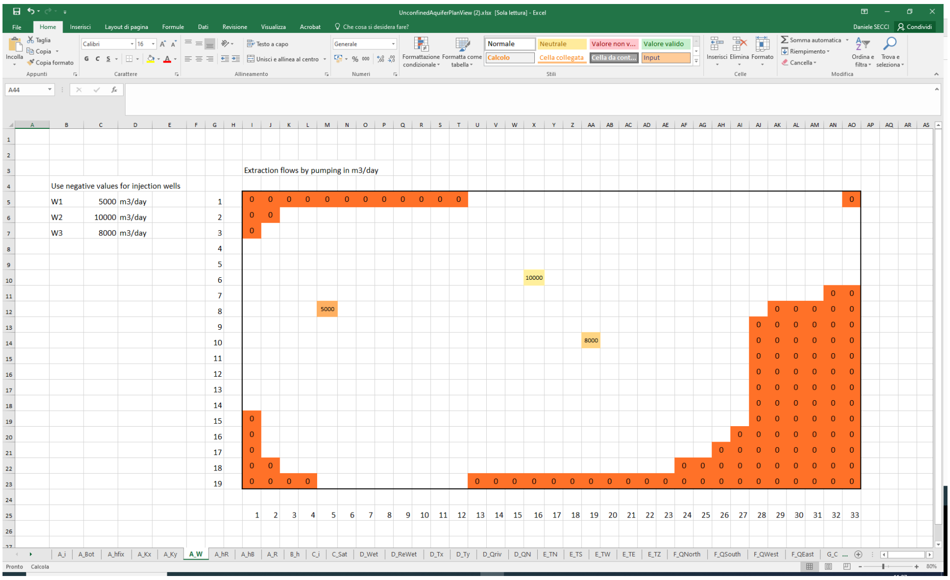
Hydraulic conductivity along the y direction in m/day

Ky1 10 m/day
Ky2 20 m/day
Ky3 5 m/day

1													10	10	10	20	20	20	20	20	20	20	5	5	5	5	5	5	5	5	5	5	
2			10	10	10	10	10	10	10	10	10	10	10	10	10	20	20	20	20	20	20	20	5	5	5	5	5	5	5	5	5	5	5
3		10	10	10	10	10	10	10	10	10	10	10	10	10	20	20	20	20	20	20	20	5	5	5	5	5	5	5	5	5	5	5	5
4	10	10	10	10	10	10	10	10	10	10	10	10	20	20	20	20	20	20	20	5	5	5	5	5	5	5	5	5	5	5	5	5	5
5	10	10	10	10	10	10	10	10	10	10	10	10	20	20	20	20	20	20	20	5	5	5	5	5	5	5	5	5	5	5	5	5	5
6	10	10	10	10	10	10	10	10	10	10	10	20	20	20	20	20	20	20	20	5	5	5	5	5	5	5	5	5	5	5	5	5	5
7	10	10	10	10	10	10	10	10	10	10	20	20	20	20	20	20	20	20	20	5	5	5	5	5	5	5	5	5	5	5	5		
8	10	10	10	10	10	10	10	10	20	20	20	20	20	20	20	20	20	20	20	5	5	5	5	5	5	5	5	5					
9	10	10	10	10	10	10	10	10	20	20	20	20	20	20	20	20	20	20	20	5	5	5	5	5	5	5	5						
10	10	10	10	10	10	10	10	10	20	20	20	20	20	20	20	20	20	20	20	5	5	5	5	5	5	5	5						
11	10	10	10	10	10	10	10	10	20	20	20	20	20	20	20	20	20	20	20	5	5	5	5	5	5	5	5						
12	10	10	10	10	10	10	10	10	20	20	20	20	20	20	20	20	20	20	20	5	5	5	5	5	5	5	5						
13	10	10	10	10	10	10	20	20	20	20	20	20	20	20	20	20	20	20	20	5	5	5	5	5	5	5	5						
14	10	10	10	10	10	10	20	20	20	20	20	20	20	20	20	20	20	20	20	5	5	5	5	5	5	5	5						
15		10	10	10	10	20	20	20	20	20	20	20	20	20	20	20	20	20	20	5	5	5	5	5	5	5	5						
16		10	10	10	10	20	20	20	20	20	20	20	20	20	20	20	20	20	20	5	5	5	5	5	5	5							
17		10	10	10	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	5	5	5	5	5	5								
18			10	10	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	5	5	5	5										
19					20	20	20	20	20	20	20	20																					
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33

Figure 4.10. S-N directional hydraulic conductivities

Wells are located in the same cells as in the previous paper; however, the pumping rates have been lowered from a total of 35,000 to 23,000 m^{3}d^{-1} distributed in three wells pumping at 5,000, 10,000, and 8,000 m^{3}d^{-1} as shown in Figure 4.11. The reason for this reduction is that with the previous values, the aquifer would go dry at the well locations.



Use negative values for injection wells

W1	5000	m3/day
W2	10000	m3/day
W3	8000	m3/day

Extraction flows by pumping in m3/day

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33
1	0	0	0	0	0	0	0	0	0	0	0	0																					0
2	0	0																															
3	0																																
4																																	
5																																	
6																10000																	
7																																0	0
8					5000																								0	0	0	0	0
9																												0	0	0	0	0	0
10																			8000									0	0	0	0	0	0
11																												0	0	0	0	0	0
12																												0	0	0	0	0	0
13																												0	0	0	0	0	0
14																												0	0	0	0	0	0
15	0																											0	0	0	0	0	0
16	0																										0	0	0	0	0	0	0
17	0																									0	0	0	0	0	0	0	0
18	0	0																						0	0	0	0	0	0	0	0	0	0
19	0	0	0	0									0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Figure 4.11. Wells

The remaining input data for river stage, elevation of riverbed bottom, and riverbed conductance remain unchanged from the previous work (Gómez-Hernández, 2022).

Step 2: Compute intermediate variables

Based on the heads from the previous iteration, several intermediary variables must be calculated before proceeding to the next iteration. Firstly, determine whether any active cell has dried out. This occurs when the calculated head falls below the cell bottom elevation. Spreadsheet C_i performs this check and the expression, referred to the cell I5 in Eq. 4.2, is the same for all the cells of the domain. From this point forward, only the expression for cell I5 will be presented:

$$\begin{aligned} &= \mathrm{IF}(\mathrm{Restart} > 0, \mathrm{A\_i!I5}, \\ &\quad \mathrm{IF}(\mathrm{B\_h!I5} = \mathrm{hundef}, \\ &\quad\quad 0, \\ &\quad\quad 1)) \end{aligned} \tag{4.2}$$

To enhance clarity, this expression can be understood as follows:

- During initialization or upon restarting calculations due to input data changes, utilize the provided input information to identify active cells.
- Otherwise:
  - If the current cell's head value is "undef" (the default value for initially inactive cells or cells with zero transmissivity due to zero saturated thickness), set the cell as inactive.
  - Otherwise, set the cell as active.

The resulting spreadsheet contains the active cells after the last iteration, encompassing all active cells from the input spreadsheet except those that have dried out.

Then, the saturated thickness at the end of the iteration must be calculated (C_sat represents the utilized spreadsheet). For each cell, the saturated thickness is set to zero if the cell is designated as inactive in the input data or if the calculated head from the previous iteration is lower than the cell bottom elevation. Otherwise, the saturated thickness is determined by subtracting the cell bottom elevation from the piezometric head using the following expression:

$$
\begin{aligned}
&= \mathrm{IF}(\mathrm{OR}(\mathrm{A\_i!I5} = 0, \mathrm{B\_h!I5} < \mathrm{A\_Bot!I5}),\\
&\quad 0,\\
&\quad \mathrm{B\_h!I5} - \mathrm{A\_Bot!I5}))
\end{aligned}
\tag{4.3}
$$

At this point, is essential to define a spreadsheet that "decide" if a cell should be rewetted in the next iteration. When any of the active neighboring cells possess a piezometric head exceeding the bottom elevation of the current cell, it should be rewetted in the subsequent iteration. This rewetting is facilitated by a Boolean flag that is determined using the following expression:

$$
\begin{aligned}
&= \mathrm{IF}(\mathrm{A\_i!I5} =,\\
&\quad \mathrm{FALSE},\\
&\quad \mathrm{OR}(\mathrm{AND}(\mathrm{A\_i!I6!B\_h!I6} > \mathrm{A\_Bot!I5}),\\
&\qquad \mathrm{AND}(\mathrm{A\_i!I4!B\_h!I4} > \mathrm{A\_Bot!I5}),\\
&\qquad \mathrm{AND}(\mathrm{A\_i!J5!B\_h!J5} > \mathrm{A\_Bot!I5}),\\
&\qquad \mathrm{AND}(\mathrm{A\_i!H5!B\_h!H5} > \mathrm{A\_Bot!I5}))
\end{aligned}
\tag{4.4}
$$

To aid comprehension, this expression can be understood as follows: if the cell is classified as inactive in the input data, it should remain dry in the next iteration. Otherwise, if any of the four adjacent cells are active and their piezometric heads exceed the current cell's bottom elevation, the current cell should be wetted in the next iteration. For the sake of simplicity, although this calculation is performed for all aquifer cells, it will only be utilized for those cells that were dry in the preceding iteration, as will be further explained.

At this point, the head value to be used in those cells that should rewet is computed. Upon transitioning from a dry to a wet state, a cell's head value is calculated using a function of the heads of its surrounding cells. This method is similar to the one implemented in the MODFLOW program (Harbaugh, 2005). A WetFactor, defined in spreadsheet D_ReWet, is utilized to compute the rewetting head. If all surrounding cells are active and their piezometric heads surpass the cell bottom elevation, the new head is set to the cell bottom plus the average of the differences between the heads in the surrounding cells and the cell bottom, multiplied by the WetFactor. Otherwise, the surrounding heads are sequentially evaluated in the order top, bottom, left, and right. The first active cell encountered with a head value above the cell bottom elevation is selected, and the rewetting head is determined as the cell bottom plus the difference between the adjacent cell head and the cell bottom, multiplied by the WetFactor. This procedure is expressed by the following equation:

$$
\begin{aligned}
&= \mathrm{IF}(\mathrm{A\_i!I5} = 0,\\
&\quad \mathrm{FALSE},\\
&\quad \mathrm{IF}(\mathrm{AND}(\mathrm{B\_h!I4} \neq \mathrm{hundef}, \mathrm{B\_h!I6} \neq \mathrm{hundef}, \mathrm{B\_h!H5} \neq \mathrm{hundef}, \mathrm{B\_h!J5} \neq \mathrm{hundef}),\\
&\quad \mathrm{A\_Bot!I5} + \left(\frac{\mathrm{B\_h!I4} + \mathrm{B\_h!J5} + \mathrm{B\_h!I6} + \mathrm{B\_h!H5}}{4} - \mathrm{A\_Bot!I5}\right) \cdot \mathrm{WetFactor},\\
&\quad\quad \mathrm{IF}(\mathrm{B\_h!I4} \neq \mathrm{hundef},\\
&\quad\quad\quad \mathrm{A\_Bot!I5} + (\mathrm{B\_h!I4} - \mathrm{A\_Bot!I5}) \cdot \mathrm{WetFactor},\\
&\quad\quad\quad \mathrm{IF}(\mathrm{B\_h!I6} \neq \mathrm{hundef},\\
&\quad\quad\quad\quad \mathrm{A\_Bot!I5} + (\mathrm{B\_h!I6} - \mathrm{A\_Bot!I5}) \cdot \mathrm{WetFactor},\\
&\quad\quad\quad\quad \mathrm{IF}(\mathrm{B\_h!H5} \neq \mathrm{hundef},\\
&\quad\quad\quad\quad\quad \mathrm{A\_Bot!I5} + (\mathrm{B\_h!H5} - \mathrm{A\_Bot!I5}) \cdot \mathrm{WetFactor},\\
&\quad\quad\quad\quad\quad \mathrm{IF}(\mathrm{B\_h!J5} \neq \mathrm{hundef},\\
&\quad\quad\quad\quad\quad\quad \mathrm{A\_Bot!I5} + (\mathrm{B\_h!J5} - \mathrm{A\_Bot!I5}) \cdot \mathrm{WetFactor}))))
\end{aligned}
\tag{4.5}
$$

This calculation aims to establish the rewetted cell's head value above its bottom elevation by a specific fraction of the difference between the bottom and the surrounding cell heads. This allows cells that had dried out during an iteration to regain wet status if the surrounding cells possess piezometric heads that justify such a transition.

Subsequent to calculating the saturated thickness, transmissivities should be determined as the product of the hydraulic conductivity and the saturated thickness. This process is carried out for both the D_Tx and D_Ty components of the transmissivity tensor. For example, the transmissivity in the W-E direction is given by:

$$\begin{aligned} = \mathrm{IF}(&\mathrm{C\_i!I5} <> 1, \\ &0, \\ &A\_\mathrm{Kx!I5} \cdot \mathrm{C\_Sat!I5}) \end{aligned} \tag{4.6}$$

It is important to note that the transmissivity is set to zero for any cell that was inactive in the input data or became inactive during the previous iteration due to drying out.

Then, to compute the flow entering the aquifer from the river or exiting to the river the same equation of the confined application can be used, but limited to the cells that are active in the current iteration. Also for the infiltration, a small modification with the confined application is implemented because of the rewetting issue. To accurately calculate the global balance in subsequent steps, the infiltration should be assessed solely on the active cells in the last iteration. For inactive cells, the recharge remains undefined; for cells with a prescribed head, the recharge is set to zero. Otherwise, the recharge is calculated by multiplying the infiltration rate N1 and the cell area Delta*Delta:

$$
\begin{aligned}
&= \text{IF(C\_i!I5 <> 1,}\\
&\quad \text{””,}\\
&\quad \text{IF(A\_hfix!I5 <> ””,}\\
&\qquad \text{0,}\\
&\qquad\quad \text{N1*Delta*Delta}
\end{aligned}
\tag{4.7}
$$

The interblock transmissivities are calculated similarly to the confined application, using the harmonic average of adjacent cell transmissivities. However, for the horizontal interblock transmissivities in E_TE and E_TW, the transmissivities along the W-E direction are employed, while the transmissivities along the S-N direction are used for the interblock transmissivities in E_TN and E_TS. Corresponding expressions are derived for the remaining three interblock transmissivities. Additionally, sheet E_TZ is constructed with the sum of the four interblock transmissivities to expedite calculations.

Step 3: Compute the iterative solution for the piezometric heads

The head calculations are conducted in sheet B_h, leveraging Excel's iterative capabilities to handle circular dependencies. The equation governing the head at each active cell is reported here for reference:

$$
h = \frac{T'_N h_N + T'_S h_S + T'_W h_W + T'_E h_E - W + N\mathit{\Delta}^2 + Q_{Riv}}{T'_N + T'_S + T'_W + T'_E}. \tag{4.8}
$$

where h represents the head at the cell being calculated, T'_N, T'_S, T'_W, T'_E are the interblock transmissivities to the North, South, West, and East of the cell, h'_N, h'_S, h'_W, h'_E are the heads at the cells to the North, South, West, and East of the cell, W is the extraction flow at the cell, N is the net infiltration rate, $\mathit{\Delta}^2$ is the cell area (assuming square cells) and Q_{Riv} is the flux entering from the river.

Two variables in spreadsheet B_h govern the initialization and commencement of calculations. The first, "hini", assigns the initial head value to all active cells during the initialization phase. The second, "Restart", initializes the heads to the hini value. When Restart is positive, all active cells are set to hini; when Restart is zero, computations begin. A third variable, hundef, is assigned to inactive cells within the rectangular area encompassing the aquifer. In this case, the chosen value is -99, which is easily identifiable.

In unconfined aquifers, the piezometric head calculation at cell I5 is more complex than in confined aquifers due to the potential for cells to become dry or rewet during the simulation. To account for this, the expression for computing the head value at cell I5 involves additional considerations:

$$
\begin{aligned}
&= \mathrm{IF}(\mathrm{A\_i!I5} <> 1,\\
&\quad \mathrm{hundef},\\
&\quad \mathrm{IF}(\mathrm{A\_hfix!I5} <> \text{""},\\
&\qquad \mathrm{A\_hfix!I5},\\
&\qquad \mathrm{IF}(\mathrm{Restart} > 0,\\
&\qquad\quad \mathrm{hini},\\
&\qquad\quad \mathrm{IF}(\mathrm{AND}(\mathrm{D\_Wet!I5}, \mathrm{B\_h!I5} = \mathrm{hundef}),\\
&\qquad\qquad \mathrm{D\_ReWet!I5},\\
&\qquad\qquad \mathrm{IFERROR}((\mathrm{E\_TN!I5} * \mathrm{B\_h!I4} + \mathrm{E\_TS!I5} * \mathrm{B\_h!I6}\\
&\qquad\qquad + \mathrm{E\_TE!I5} * \mathrm{B\_h!J5} + \mathrm{E\_TW!I5} * \mathrm{B\_h!H5}\\
&\qquad\qquad - \mathrm{A\_W!I5} + \mathrm{D\_Qriv!I5} + \mathrm{D\_QN!I5})/(\mathrm{E\_TZ!I5}), \mathrm{hundef})))))
\end{aligned}
\tag{4.9}
$$

This expression determines the piezometric head at cell I5. It follows a series of conditional steps and can be read in this terms:

- Inactive Cell Check: If the current cell (I5) is inactive in the input data set, the piezometric head is set to a predefined value (hundef);
- Prescribed Head Check: If the cell (I5) corresponds to a prescribed head location, the piezometric head is set to the specified prescribed head value;
- Initialization Check: If the Restart flag is greater than zero, indicating a simulation restart, the cell's piezometric head is initialized to the initial head value (hini);
- Rewetting Check: If the cell's head is undefined (due to dryness in the previous iteration) and the D_Wet flag is True, implying that the cell should be wet (surrounding cell heads are above the cell bottom), the piezometric head is set to the rewetting value computed in D_ReWet;
- Piezometric Head Calculation: If none of the above conditions apply, the piezometric head is calculated using Eq. 4.8. The IFERROR function is employed to handle the case where the denominator in Eq. 4.8 is zero, which occurs for cells with zero saturated thickness. In such cases, the head is set to hundef (undefined), otherwise, it is set to the calculated value;

Step 4: Checking both global and local balances

The global balance check is conducted on the same sheet where the heads are calculated. It involves evaluating six variables: the flux extracted through wells, the recharge from infiltration, the recharge from the river, and the flow entering the aquifer from prescribed head cells. The global error is determined by summing these four variables. Additionally, a relative error is calculated by dividing the absolute error by the total flux entering from the prescribed head boundary.

The local balance check involves computing the fluxes entering each cell and ensuring that they sum to zero. This process utilizes sheets F_QNorth, F_QSouth, F_QEast, and F_QWest. For example, F_QNorth contains the fluxes entering each cell through its northern boundary. It is calculated using the expression:

$$\begin{aligned}
&= \mathrm{IF}(\mathrm{B\_h!I5} = \mathrm{hundef},\\
&\quad \text{""},\\
&\quad \mathrm{IF}(\mathrm{AND}(\mathrm{C\_i!I5} = 1, \mathrm{C\_i!I4} = 0,\\
&\qquad 0,\\
&\qquad \mathrm{E\_TN!I5} * (\mathrm{B\_h!I4} - \mathrm{B\_h!I5})))
\end{aligned} \tag{4.10}$$

This expression means that before calculating the flux entering each cell through its northern boundary (QNorth), a check is done to verify if the head value is undefined. If it is, leave QNorth as an empty string. Otherwise, it is necessary to proceed with the following steps:

- Active Cell Check: Determine if the current cell is active after the last iteration. If it is inactive, set QNorth to zero;
- Adjacent Cell Check: Check if the adjacent cell to the north is active. If it is inactive, set QNorth to zero;
- Darcy's Law Calculation: If both the current cell and the adjacent cell to the north are active, proceed to calculate QNorth using Darcy's law. Darcy's law relates the flow rate of a fluid through a porous medium to the pressure difference between two points and the hydraulic properties of the medium;

By following these steps, the QNorth value for each cell can be accurately determined, ensuring that the local balance check is performed effectively.

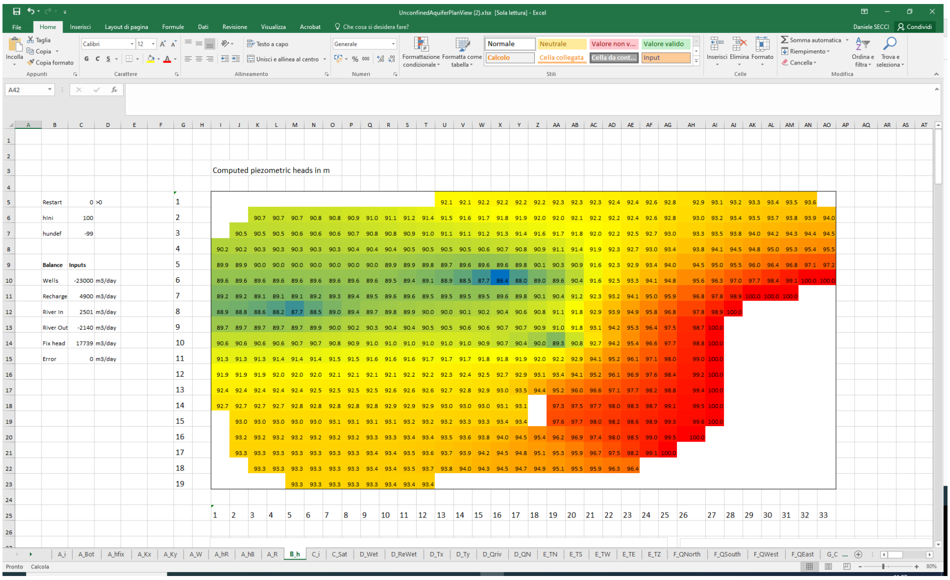
The remaining flow components are calculated using similar expressions to QNorth. Once all flow components are determined, sheet G_CellBalance is created, which summarizes the sum of all incoming flow components for each cell.

Upon convergence, this sheet should contain zero values for all cells except those corresponding to prescribed head boundaries. These cells will exhibit values representing the flow exchange between the aquifer and the boundary.

Iterations continue until the Error value in sheet B_h is close to zero and all active cells in sheet G_CellBal show values close to zero. Figure 4.12 illustrates the resulting piezometric heads obtained after 8000 iterations for the given input parameters. A heat map is generated, filtering out the hundef values (-99) and ranging from the minimum computed head to the maximum one.

To create this heat map, follow these steps:

- Select the rectangular area covering the aquifer (I5:AO23);
- Access the Format menu, choose Conditional Formatting, and establish two conditions;
- For the first condition, apply a custom formatting style to cells containing the value -99. Set the font color to white;
- For the second condition, use a Three-color scale. In the Formula field, enter =MIN(IF($I$5:$AO$23<>-99,$I$5:$AO$23)), which calculates the minimum head value while excluding cells with a value of -99;
- For the midpoint, set the 50th percentile;
- For the maximum, enter the highest value;
- Assign the desired colors for the minimum (dark blue), midpoint (yellow), and maximum (red);



Restart	0	>0
hIni	100	
hundef	-99	
Balance	**Inputs**	
Wells	-23000	m3/day
Recharge	4900	m3/day
River In	2501	m3/day
River Out	-2140	m3/day
Fix head	17739	m3/day
Error	0	m3/day

Computed piezometric heads in m

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33
1													92.1	92.1	92.2	92.2	92.2	92.2	92.3	92.3	92.3	92.4	92.4	92.6	92.8	92.9	93.1	93.2	93.3	93.4	93.5	93.6	
2			90.7	90.7	90.7	90.8	90.8	90.9	91.0	91.1	91.2	91.4	91.5	91.6	91.7	91.8	91.9	92.0	92.0	92.1	92.2	92.2	92.4	92.6	92.8	93.0	93.2	93.4	93.5	93.7	93.8	93.9	94.0
3		90.5	90.5	90.5	90.6	90.6	90.6	90.7	90.8	90.8	90.9	91.0	91.1	91.1	91.2	91.3	91.4	91.6	91.7	91.8	92.0	92.2	92.5	92.7	93.0	93.3	93.5	93.8	94.0	94.2	94.3	94.4	94.5
4	90.2	90.2	90.3	90.3	90.3	90.3	90.3	90.4	90.4	90.4	90.5	90.5	90.5	90.5	90.6	90.7	90.8	90.9	91.1	91.4	91.9	92.3	92.7	93.0	93.4	93.8	94.1	94.5	94.8	95.0	95.3	95.4	95.5
5	89.9	89.9	90.0	90.0	90.0	90.0	90.0	90.0	90.0	89.9	89.9	89.8	89.7	89.6	89.6	89.6	89.8	90.1	90.3	90.9	91.6	92.3	92.9	93.4	94.0	94.5	95.0	95.5	96.0	96.4	96.8	97.1	97.2
6	89.6	89.6	89.6	89.6	89.6	89.6	89.6	89.6	89.6	89.5	89.4	89.1	88.9	88.5	87.7	86.4	88.0	89.0	89.6	90.4	91.6	92.5	93.3	94.1	94.8	95.6	96.3	97.0	97.7	98.4	99.1	100.0	100.0
7	89.2	89.2	89.1	89.1	89.1	89.2	89.3	89.4	89.5	89.6	89.6	89.5	89.5	89.5	89.5	89.6	89.8	90.1	90.4	91.2	92.3	93.2	94.1	95.0	95.9	96.8	97.8	98.9	100.0	100.0	100.0		
8	88.9	88.8	88.6	88.2	87.7	88.5	89.0	89.4	89.7	89.8	89.9	90.0	90.0	90.1	90.2	90.4	90.6	90.8	91.1	91.8	92.9	93.9	94.9	95.8	96.8	97.8	98.9	100.0					
9	89.7	89.7	89.7	89.7	89.7	89.9	90.0	90.2	90.3	90.4	90.4	90.5	90.5	90.6	90.6	90.7	90.7	90.9	91.0	91.8	93.1	94.2	95.3	96.4	97.5	98.7	100.0						
10	90.6	90.6	90.6	90.6	90.7	90.7	90.8	90.9	91.0	91.0	91.0	91.0	91.0	91.0	90.9	90.7	90.4	90.0	89.3	90.8	92.7	94.2	95.4	96.6	97.7	98.8	100.0						
11	91.3	91.3	91.3	91.4	91.4	91.4	91.5	91.5	91.6	91.6	91.6	91.7	91.7	91.7	91.8	91.8	91.9	92.0	92.2	92.9	94.1	95.2	96.1	97.1	98.0	99.0	100.0						
12	91.9	91.9	91.9	92.0	92.0	92.0	92.1	92.1	92.1	92.1	92.2	92.2	92.3	92.4	92.5	92.7	92.9	93.1	93.4	94.1	95.2	96.1	96.9	97.6	98.4	99.2	100.0						
13	92.4	92.4	92.4	92.4	92.4	92.5	92.5	92.5	92.5	92.6	92.6	92.6	92.7	92.8	92.9	93.0	93.5	94.4	95.2	96.0	96.6	97.1	97.7	98.2	98.8	99.4	100.0						
14	92.7	92.7	92.7	92.7	92.8	92.8	92.8	92.8	92.8	92.9	92.9	92.9	93.0	93.0	93.0	93.1	93.1		97.3	97.5	97.7	98.0	98.3	98.7	99.1	99.5	100.0						
15		93.0	93.0	93.0	93.0	93.0	93.1	93.1	93.1	93.1	93.2	93.2	93.2	93.3	93.3	93.4	93.4		97.6	97.7	98.0	98.2	98.6	98.9	99.3	99.6	100.0						
16		93.2	93.2	93.2	93.2	93.2	93.2	93.2	93.3	93.3	93.4	93.4	93.5	93.6	93.8	94.0	94.5	95.4	96.2	96.9	97.4	98.0	98.5	99.0	99.5	100.0							
17		93.3	93.3	93.3	93.3	93.3	93.3	93.3	93.4	93.4	93.5	93.6	93.7	93.9	94.2	94.5	94.8	95.1	95.3	95.9	96.7	97.5	98.2	99.1	100.0								
18			93.3	93.3	93.3	93.3	93.3	93.3	93.4	93.4	93.5	93.7	93.8	94.0	94.3	94.5	94.7	94.9	95.1	95.5	95.9	96.3	96.4										
19					93.3	93.3	93.3	93.3	93.3	93.4	93.4	93.4																					

Figure 4.12. Piezometric heads solution

The iterative process converges to an almost exact solution, as indicated by an absolute error of 0 m^3/d after 8000 iterations. Figure 4.13 illustrates the evolution of piezometric heads throughout the iterations.

After 1000 iterations, the solution is far from convergence, exhibiting a large absolute error, limited total recharge, and substantial discharges through wells, the river, and the boundary. At 2000 iterations, the head pattern closely resembles the final solution, with the anisotropy becoming evident. The two cells with exceptionally high bottom values have already dried up, but the absolute error remains high, with the river appearing to draw water from the aquifer. After 3000 iterations, the error is reduced, but convergence is not achieved until 8000 iterations, resulting in the solution shown in Figure 4.12. The final water balance indicates 23,000 m^3/d extracted by wells, 4,900 m^3/d recharged by infiltration, 361 m^3/d net recharge from the river, and 17,739 m^3/d entering the aquifer from the lake.

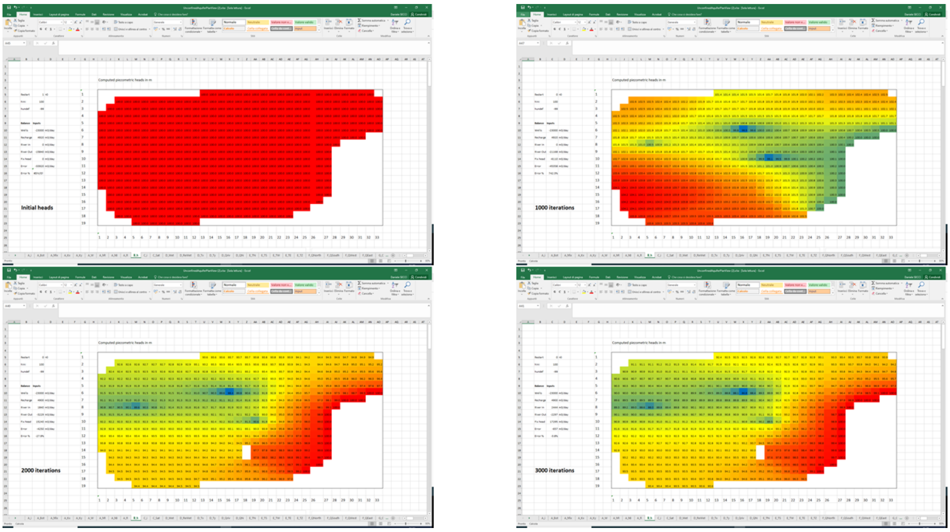
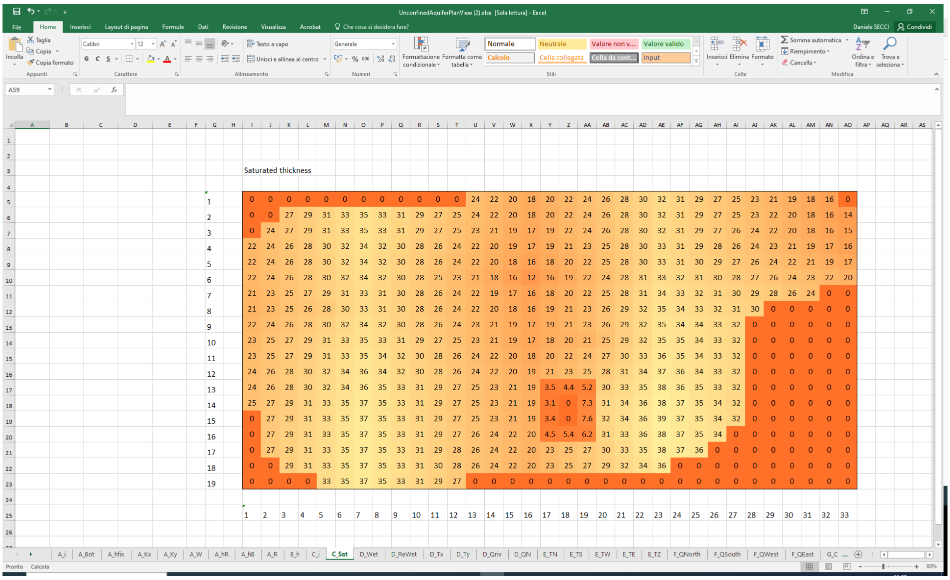


Figure 4.13. Piezometric heads solution of different iterations

Figure 4.14 depicts the aquifer's saturated thickness, highlighting the dry cells and the surrounding cells with low saturated thickness. Figure 4.15 illustrates the flow entering the aquifer from the river, revealing that the river is a losing river in the northern section, disconnected for most of its upper stretch, and becomes a gaining river in the lower part.

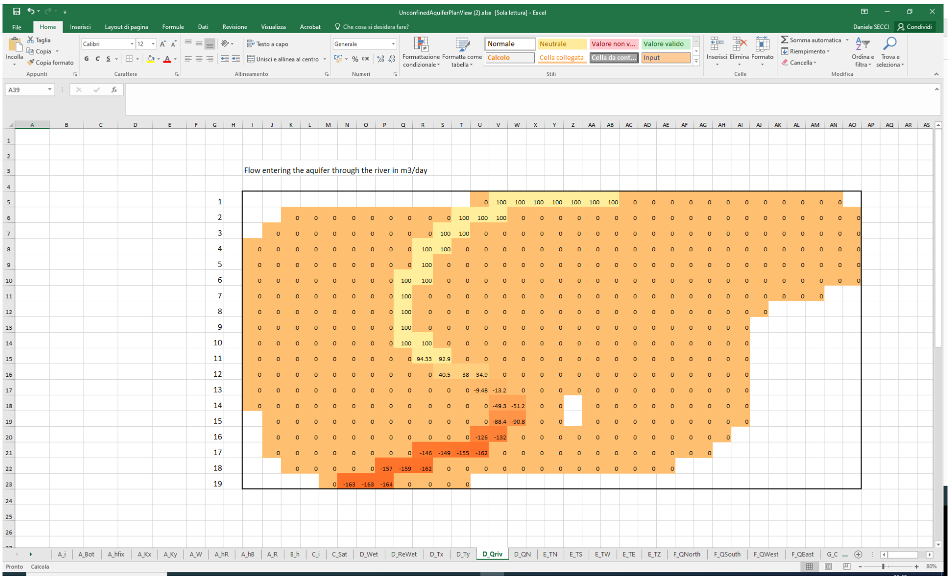
The elliptical cones of depression formed around the three pumping wells are visible in Figure 4.12. The elliptical shape is a result of the strong hydraulic conductivity anisotropy in the W-E direction.



Saturated thickness

1	0	0	0	0	0	0	0	0	0	0	0	0	24	22	20	18	20	22	24	26	28	30	32	31	29	27	25	23	21	19	18	16	0
2	0	0	27	29	31	33	35	33	31	29	27	25	24	22	20	18	20	22	24	26	28	30	32	31	29	27	25	23	22	20	18	16	14
3	0	24	27	29	31	33	35	33	31	29	27	25	23	21	19	17	19	22	24	26	28	30	32	31	29	27	26	24	22	20	18	16	15
4	22	24	26	28	30	32	34	32	30	28	26	24	22	20	19	17	19	21	23	25	28	30	33	31	29	28	26	24	23	21	19	17	16
5	22	24	26	28	30	32	34	32	30	28	26	24	22	20	18	16	18	20	22	25	28	30	33	31	30	29	27	26	24	22	21	19	17
6	22	24	26	28	30	32	34	32	30	28	25	23	21	18	16	12	16	19	22	24	28	31	33	32	31	30	28	27	26	24	23	22	20
7	21	23	25	27	29	31	33	31	30	28	26	24	22	19	17	16	18	20	22	25	28	31	34	33	32	31	30	29	28	26	24	0	0
8	21	23	25	26	28	30	33	31	30	28	26	24	22	20	18	16	19	21	23	26	29	32	35	34	33	32	31	30	0	0	0	0	0
9	22	24	26	28	30	32	34	32	30	28	26	24	23	21	19	17	19	21	23	26	29	32	35	34	34	33	32	0	0	0	0	0	0
10	23	25	27	29	31	33	35	33	31	29	27	25	23	21	19	17	18	20	21	25	29	32	35	35	34	33	32	0	0	0	0	0	0
11	23	25	27	29	31	33	35	34	32	30	28	26	24	22	20	18	20	22	24	27	30	33	36	35	34	33	32	0	0	0	0	0	0
12	24	26	28	30	32	34	36	34	32	30	28	26	24	22	20	19	21	23	25	28	31	34	37	36	34	33	32	0	0	0	0	0	0
13	24	26	28	30	32	34	36	35	33	31	29	27	25	23	21	19	3.5	4.4	5.2	30	33	35	38	36	35	33	32	0	0	0	0	0	0
14	25	27	29	31	33	35	37	35	33	31	29	27	25	23	21	19	3.1	0	7.3	31	34	36	38	37	35	34	32	0	0	0	0	0	0
15	0	27	29	31	33	35	37	35	33	31	29	27	25	23	21	19	3.4	0	7.6	32	34	36	39	37	35	34	32	0	0	0	0	0	0
16	0	27	29	31	33	35	37	35	33	31	29	27	26	24	22	20	4.5	5.4	6.2	31	33	36	38	37	35	34	0	0	0	0	0	0	0
17	0	27	29	31	33	35	37	35	33	31	29	28	26	24	22	20	23	25	27	30	33	35	38	37	36	0	0	0	0	0	0	0	0
18	0	0	29	31	33	35	37	35	33	31	30	28	26	24	22	20	23	25	27	29	32	34	36	0	0	0	0	0	0	0	0	0	0
19	0	0	0	0	33	35	37	35	33	31	29	27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33

Figure 4.14. Aquifer saturated thickness after the completion of iterations.



Flow entering the aquifer through the river in m3/day

1													0	100	100	100	100	100	100	100	0	0	0	0	0	0	0	0	0	0	0	0	
2			0	0	0	0	0	0	0	0	0	100	100	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3		0	0	0	0	0	0	0	0	0	100	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	100	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	100	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		
8	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0					
9	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0						
10	0	0	0	0	0	0	0	0	100	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0						
11	0	0	0	0	0	0	0	0	0	94.33	92.9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0						
12	0	0	0	0	0	0	0	0	0	0	40.5	38	34.9	0	0	0	0	0	0	0	0	0	0	0	0	0	0						
13	0	0	0	0	0	0	0	0	0	0	0	0	-9.48	-13.2	0	0	0	0	0	0	0	0	0	0	0	0	0						
14	0	0	0	0	0	0	0	0	0	0	0	0	0	-49.3	-51.2	0	0		0	0	0	0	0	0	0	0	0						
15		0	0	0	0	0	0	0	0	0	0	0	0	-88.4	-90.8	0	0		0	0	0	0	0	0	0	0	0						
16		0	0	0	0	0	0	0	0	0	0	0	-126	-132	0	0	0	0	0	0	0	0	0	0	0	0							
17		0	0	0	0	0	0	0	0	-146	-149	-155	-162	0	0	0	0	0	0	0	0	0	0	0	0								
18			0	0	0	0	0	-157	-159	-162	0	0	0	0	0	0	0	0	0	0	0	0	0										
19					0	-163	-163	-164	0	0	0	0																					

Figure 4.15. River inflow to the aquifer.

4.2.2. Unconfined two-dimensional groundwater flow in a vertical cross-section

With the functioning model for an unconfined aquifer in plan view, it is straightforward to adapt it to simulate a vertical cross-section of an unconfined aquifer. This modification aims to determine the phreatic surface, an unknown boundary condition that makes the solution of the groundwater flow problem highly nonlinear. The model setup represents a vertical cross-section of a formation situated on an inclined impermeable bottom with a river crossing, a pumping well, and recharge through infiltration. Figure 4.16 depicts the sketch of the considered aquifer domain, including three distinct hydraulic conductivity zones, a well that operates only in the lower part of the aquifer, and key elevation data required for defining the boundary conditions. The shaded area represents the saturated zone, where recharge from the surface infiltrates and replenishes the cells that intersect the free surface.

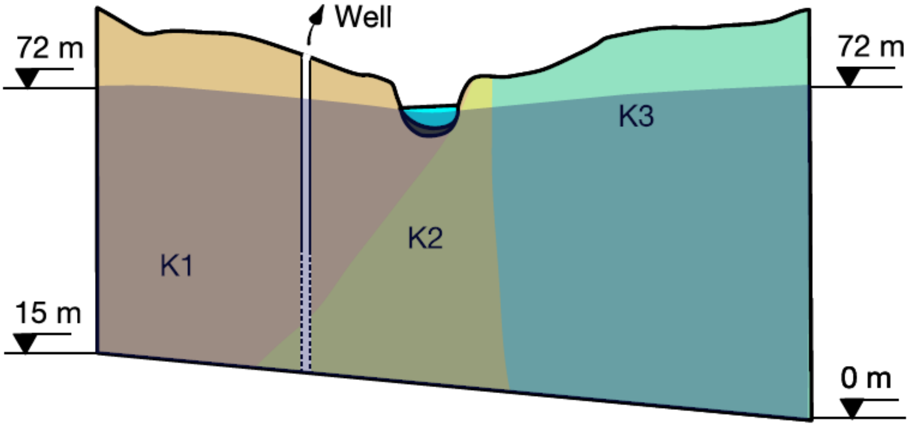


Figure 4.16. Vertical cross-section sketch of the aquifer.

To achieve the final result, some modifications are necessary compared to the horizontal plane application.

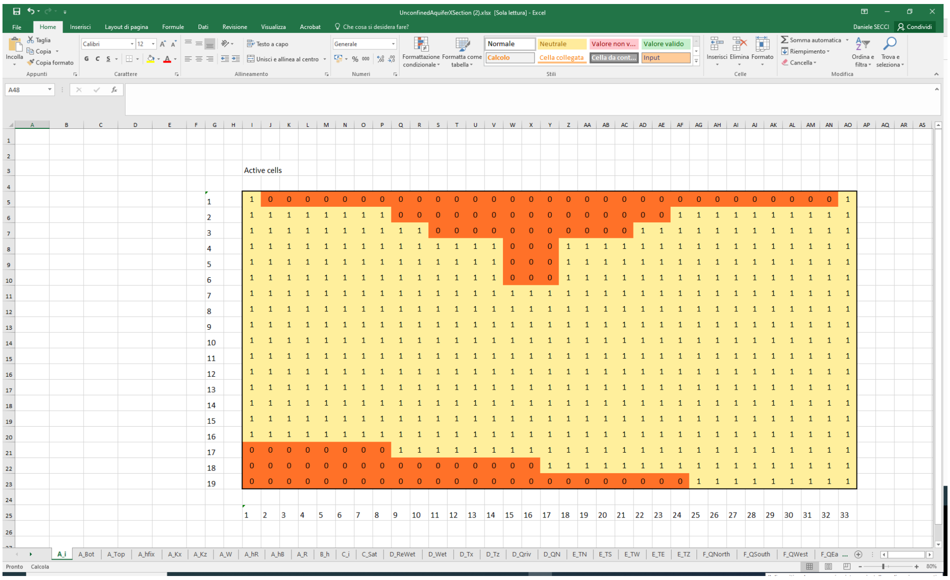
Step1: Input parameters

In this application, the active cells represent the aquifer below the ground surface and up to the impermeable base. Figure 4.17 shows the active cells in the cross-section. The water table at both vertical boundaries is at an elevation of 72 m, and the base slopes from 15 m to 0 m as shown in Figure 4.16. It is assumed that the vertical piezometric head gradient is zero along the vertical boundaries. Therefore, the boundary conditions are prescribed heads of 72 m at the first and last columns, an impermeable bottom, and an unknown phreatic boundary condition at the top.

The prescribed head cell that intersects the phreatic surface must have a head value that coincides with the elevation of the cell. There cannot be any active cells above the phreatic surface. In this case, since the prescribed head is 72 m, the prescribed head cells that intersect the phreatic surface are in row 5, which has a bottom elevation of 70 m and a top elevation of 75 m.

The cross-section is in the XZ plane and is larger along the horizontal direction than along the vertical direction (Figure 4.16). This means that in a realistic scenario, it is impractical to use square cells of equal dimensions, such as 5 by 5, when modeling a section of 90 m by 3,300 m, as in this case. This approach results in increased computational costs without a proportional improvement in solution accuracy. There are two main implications of this: the horizontal (Δx) and vertical (Δz) cell spacings must be specified. The cell aspect ratio (Δx/Δz) plays a role in the calculation of flows into and out of each cell, and in the expression for computing the piezometric head as a function of surrounding heads (Eq. 4.8). The cell aspect ratio is 1 for square cells, which is why it does not appear in Eq. 4.8. The bottom and top of each cell are fully defined by the elevation of the lower left corner of the rectangle that encloses the aquifer and the value of Δz. Therefore, it is only necessary to input the elevation of the lower left corner in spreadsheet

A_Bot. The bottom and top values for all cells are automatically calculated using Δz, which is given in spreadsheet B_h. In addition, in order to properly model recharge by infiltration and river-aquifer flow exchange, there is also the necessity to specify the width of the aquifer perpendicular to the cross-section (Δy).



Active cells

1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
2	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1
3	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1
4	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
5	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
7	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
8	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
10	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
11	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
12	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
14	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
15	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
17	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33

Figure 4.17. Active cells in the vertical cross-section of the aquifer

The recharge and river inflow/outflow will vary depending on the width of the aquifer section. For this reason, spreadsheets B_h and A_R are used to provide values for the cell spacings and riverbed properties. In spreadsheet A_R, the riverbed conductance is computed assuming that the river covers the entire cell. However, river infiltration is not considered when the head is below the cell bottom. It is possible to develop a mechanism to compute infiltration from a disconnected river and carry it down to the phreatic surface, but this would make the workbook too complex. The recharge is computed in spreadsheet D_QN by identifying the active cells after the last iteration that intersect the phreatic surface. These cells are assigned a flow entering by recharge equal to $QN = N\Delta x\Delta y$, where N is the infiltration rate. The expression below explains the procedure to evaluate the net infiltration:

$$
\begin{aligned}
&= \mathrm{IF}(\mathrm{C\_i!I5} = 0,\\
&\quad \text{""},\\
&\quad \mathrm{IF}(\mathrm{A\_hfix!I5} <> \text{""},\\
&\qquad 0,\\
&\qquad \mathrm{IF}(\mathrm{C\_i!I4} = 0, \mathrm{N1} \cdot \mathrm{Deltax} \cdot \mathrm{Deltay})))
\end{aligned}
\tag{4.11}
$$

This expression means that if the cell is inactive, set its value to empty cell. Otherwise, if the cell has a prescribed head, set the recharge to zero. Otherwise, if the cell above is inactive, assign a recharge equal to the infiltration recharge (N1) times the cell size in the x direction (deltax) times the cell size in the y direction (deltay). Otherwise, set it to zero.

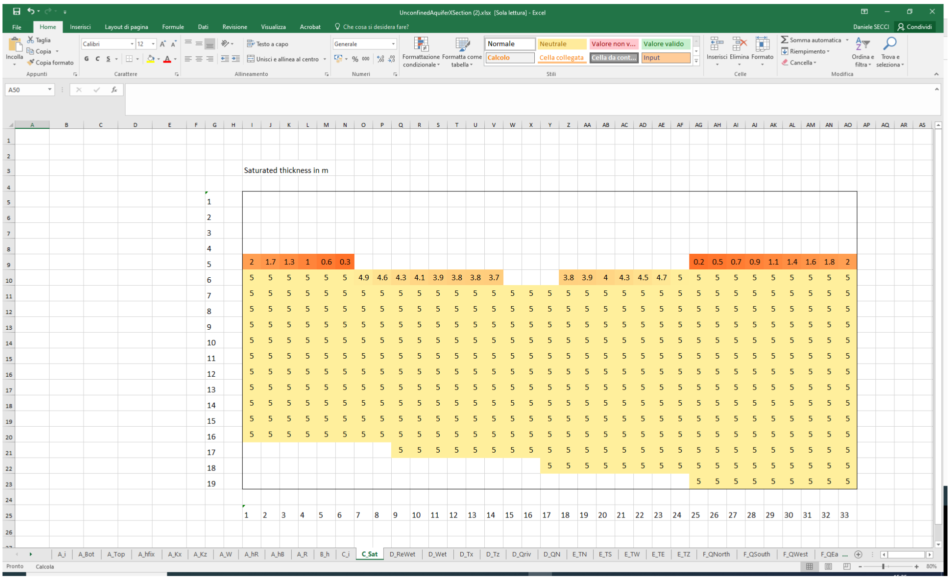
Step 2: Calculate intermediate variables

The saturated thickness intermediate variable refers to the saturated thickness of the aquifer on the plane of the cross-section, not the saturated thickness orthogonal to the plane view. It is equal to the cell thickness (Δz) for all saturated cells, and equal to the difference between the calculated piezometric head and the cell bottom elevation for the cells that intersect the phreatic surface.

It is necessary to account for the aspect ratio (Δx/Δz) when calculating the flows into and out of each cell, and to modify the calculation of the piezometric head accordingly. Eq. 4.8 can still be used for calculating the piezometric head if the cell aspect ratio is incorporated into the calculation of the terms T'_N, T'_S, T'_W, T'_E, which transform them from interface transmissivities to conductances.

Conductances are values that, when multiplied by the head difference between two adjacent cells, provide the flow crossing their interface. This value is equal to the interface transmissivity multiplied by the interface area and divided by the distance between cell centers.

To account for the cell aspect ratio, the spreadsheets D_Tx, D_Tz, E_TN, E_TS, E_TW, and E_TE are modified as follows: D_Tx and D_Tz contain the transmissivity, which is defined as the conductivity multiplied by the saturated thickness measured orthogonal to the cross-section. Since the saturated thickness orthogonal to the cross-section is $\mathit{\Delta}$y, D_Tx and D_Tz are equal to the values in A_Kx and A_Kz multiplied by $\mathit{\Delta}$y. To obtain the conductances, the interface transmissivity is first computed as the harmonic average of adjacent transmissivities. Then, this value is multiplied by the cell width and divided by the distance between cell centers. The calculation of conductances is different for the vertical and horizontal axes. For the vertical axis, the cell width is constant ($\mathit{\Delta}$x) and the distance between cell centers is constant ($\mathit{\Delta}$z). For the horizontal axis, the cell height is variable (equal to the saturated thickness) and the distance between cell centers is constant ($\mathit{\Delta}$x). Figure 4.18 shows the saturated thickness at the end of the iterations.



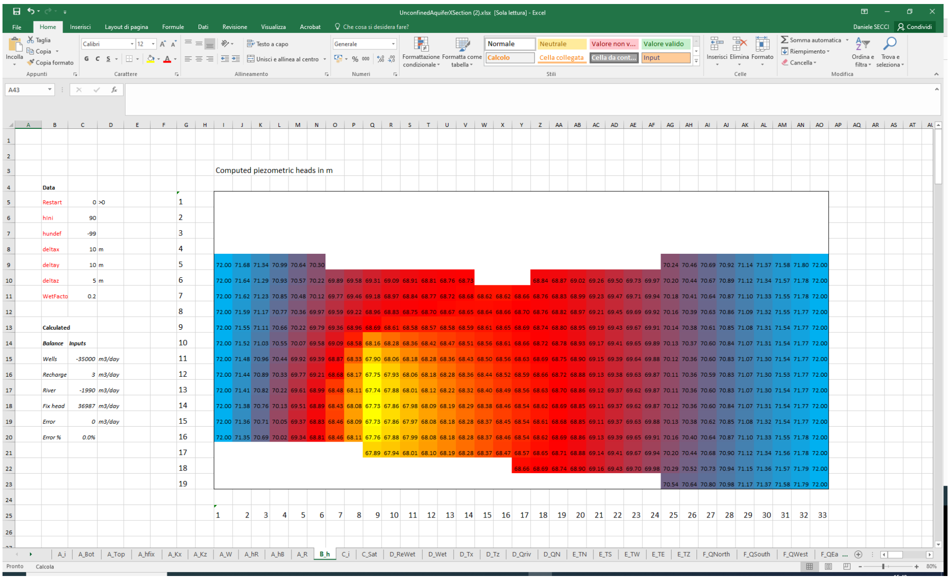
Saturated thickness in m

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33
1																																	
2																																	
3																																	
4																																	
5	2	1.7	1.3	1	0.6	0.3																			0.2	0.5	0.7	0.9	1.1	1.4	1.6	1.8	2
6	5	5	5	5	5	5	4.9	4.6	4.3	4.1	3.9	3.8	3.8	3.7				3.8	3.9	4	4.3	4.5	4.7	5	5	5	5	5	5	5	5	5	5
7	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
8	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
9	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
10	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
11	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
12	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
13	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
14	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
15	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
16	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
17									5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
18																	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
19																									5	5	5	5	5	5	5	5	5

Figure 4.18. Saturated thickness in the vertical cross-section of the aquifer

Step 3: Evaluation of the iterative solution of the piezometric heads

Following the same workflow of the horizontal plane application, the iterative calculation of the piezometric head has not changed. However, it is important to note that convergence slows down when the aspect ratio (Δz/Δz) is very large. Figure 4.19 shows the hydraulic head distribution for an aquifer vertical cross-section discretized in cells that are 10 meters wide and 5 meters tall, with a well that pumps 35,000 m^3d^{-1}. The resulting distribution shows a clear cone of depression around the cells where the well is open. Water flows from the prescribed head boundaries towards the river and towards the well. The total infiltration is very small because the exposed surface through which recharge would enter the aquifer is only 31 × 10 × 10 × 0.001 (number of cells crossing the phreatic surface × Δx × Δy × infiltration rate). The river drains 1,990 m^3d^{-1} from the aquifer.



Data		
Restart	0	>0
hIni	90	
hundef	-99	
deltax	10	m
deltay	10	m
deltaz	5	m
WetFacto	0.2	
Calculated		
Balance	***Inputs***	
Wells	*-35000*	*m3/day*
Recharge	*3*	*m3/day*
River	*-1990*	*m3/day*
Fix head	*36987*	*m3/day*
Error	*0*	*m3/day*
Error %	*0.0%*	

Computed piezometric heads in m

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33
1																																	
2																																	
3																																	
4																																	
5	72.00	71.68	71.34	70.99	70.64	70.30																			70.24	70.46	70.69	70.92	71.14	71.37	71.58	71.80	72.00
6	72.00	71.64	71.29	70.93	70.57	70.22	69.89	69.58	69.31	69.09	68.91	68.81	68.76	68.73				68.84	68.87	69.02	69.26	69.50	69.73	69.97	70.20	70.44	70.67	70.89	71.12	71.34	71.57	71.78	72.00
7	72.00	71.62	71.23	70.85	70.48	70.12	69.77	69.46	69.18	68.97	68.84	68.77	68.72	68.68	68.62	68.62	68.66	68.76	68.83	68.99	69.23	69.47	69.71	69.94	70.18	70.41	70.64	70.87	71.10	71.33	71.55	71.78	72.00
8	72.00	71.59	71.17	70.77	70.36	69.97	69.59	69.22	68.96	68.83	68.75	68.70	68.67	68.65	68.64	68.66	68.70	68.76	68.82	68.97	69.21	69.45	69.69	69.92	70.16	70.39	70.63	70.86	71.09	71.32	71.55	71.77	72.00
9	72.00	71.55	71.11	70.66	70.22	69.79	69.36	68.96	68.69	68.61	68.58	68.57	68.58	68.59	68.61	68.65	68.69	68.74	68.80	68.95	69.19	69.43	69.67	69.91	70.14	70.38	70.61	70.85	71.08	71.31	71.54	71.77	72.00
10	72.00	71.52	71.03	70.55	70.07	69.58	69.09	68.58	68.16	68.28	68.36	68.42	68.47	68.51	68.56	68.61	68.66	68.72	68.78	68.93	69.17	69.41	69.65	69.89	70.13	70.37	70.60	70.84	71.07	71.31	71.54	71.77	72.00
11	72.00	71.48	70.96	70.44	69.92	69.39	68.87	68.33	67.90	68.06	68.18	68.28	68.36	68.43	68.50	68.56	68.63	68.69	68.75	68.90	69.15	69.39	69.64	69.88	70.12	70.36	70.60	70.83	71.07	71.30	71.54	71.77	72.00
12	72.00	71.44	70.89	70.33	69.77	69.21	68.68	68.17	67.75	67.93	68.06	68.18	68.28	68.36	68.44	68.52	68.59	68.66	68.72	68.88	69.13	69.38	69.63	69.87	70.11	70.36	70.59	70.83	71.07	71.30	71.53	71.77	72.00
13	72.00	71.41	70.82	70.22	69.61	68.99	68.48	68.11	67.74	67.88	68.01	68.12	68.22	68.32	68.40	68.49	68.56	68.63	68.70	68.86	69.12	69.37	69.62	69.87	70.11	70.36	70.60	70.83	71.07	71.30	71.54	71.77	72.00
14	72.00	71.38	70.76	70.13	69.51	68.89	68.43	68.08	67.73	67.86	67.98	68.09	68.19	68.29	68.38	68.46	68.54	68.62	68.69	68.85	69.11	69.37	69.62	69.87	70.12	70.36	70.60	70.84	71.07	71.31	71.54	71.77	72.00
15	72.00	71.36	70.71	70.05	69.37	68.83	68.46	68.09	67.73	67.86	67.97	68.08	68.18	68.28	68.37	68.45	68.54	68.61	68.68	68.85	69.11	69.37	69.63	69.88	70.13	70.38	70.62	70.85	71.08	71.32	71.54	71.77	72.00
16	72.00	71.35	70.69	70.02	69.34	68.81	68.46	68.11	67.76	67.88	67.99	68.08	68.18	68.28	68.37	68.46	68.54	68.62	68.69	68.86	69.13	69.39	69.65	69.91	70.16	70.40	70.64	70.87	71.10	71.33	71.55	71.78	72.00
17									67.89	67.94	68.01	68.10	68.19	68.28	68.37	68.47	68.57	68.65	68.71	68.88	69.14	69.41	69.67	69.94	70.20	70.44	70.68	70.90	71.12	71.34	71.56	71.78	72.00
18																	68.66	68.69	68.74	68.90	69.16	69.43	69.70	69.98	70.29	70.52	70.73	70.94	71.15	71.36	71.57	71.79	72.00
19																									70.54	70.64	70.80	70.98	71.17	71.37	71.58	71.79	72.00

Figure 4.19. Piezometric heads solution in the vertical cross-section of the aquifer

4.2.2.1. Modeling flow through an earth dam: a practical application

Groundwater flow through an earth dam can be modeled using the same equations as groundwater flow in an unconfined aquifer. Consider a dam with a core with a conductivity of 0.01 m/d and an embankment with a conductivity of 0.1 m/d, as shown in Figure 4.20.

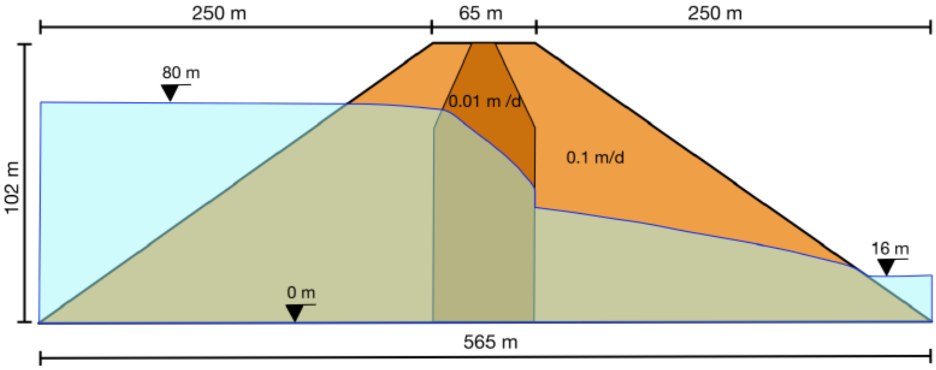


Figure 4.20. Earth dam sketch of a section with core and embankment

A vertical cross-section of the dam is modeled using a grid of 51 rows and 113 columns, with each cell being 5 meters wide and 2 meters tall. The water level on the upstream side of the dam is 80 meters and on the downstream side is 16 meters. The bottom of the dam is horizontal and impermeable at a depth of 0 meters. Using these parameters, a spreadsheet model is developed to simulate groundwater flow through the dam. The final solution is shown in Figure 4.21, which shows that high hydraulic gradients occur within the nucleus, while there are virtually no gradients on the upstream side of the embankment and some gradients on the downstream side. Assuming that the dam is 1000 meters long perpendicular to the cross-section shown, the total seepage through the dam is calculated to be 105.4 m^3d^{-1}.

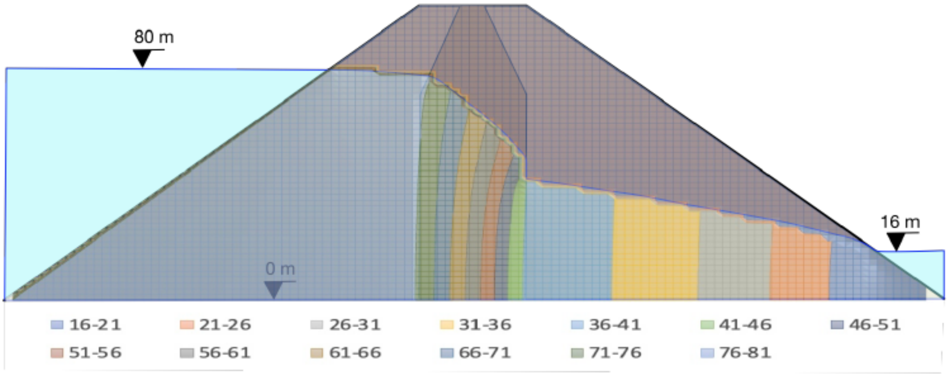


Figure 4.21. Piezometric heads solution of the earth dam section with core and embankment

4.2.3. Conclusions and discussion

This analysis extends the work of Gómez-Hernández (2022) to solve the unconfined groundwater flow equation in both the horizontal plane and in a vertical cross-section. The unconfined groundwater flow equation is nonlinear because the transmissivities depend on the piezometric heads, and also because cells may dry out.

To allow cells to dry out and rewet, a mechanism is introduced that monitors the piezometric heads in nearby cells. If the head of a cell goes below its bottom elevation, the cell is marked as dry. If, later during the iterations, the heads in the nearby cells suggest that the cell should be wet, the cell is marked as wet.

In the case of simulating a cross-section, the phreatic surface is a boundary condition that is not known in advance, and that depends on the solution of the equation. This introduces additional challenges that must be addressed during the implementation.

Despite these challenges, the spreadsheet approach yields very good results, which have been compared to the solution of the groundwater flow equation with

MODFLOW (Harbaugh, 2005).

Future work should prioritize modeling changes in groundwater flow over time, emphasizing transient modeling. Although calculating the evolution of piezometric time would not pose a significant challenge, effectively managing the results to generate time-dependent graphs requires careful consideration. Furthermore, addressing local grid refinements in specific areas could be a focal point for future research.

Moreover, constructing a three-dimensional model is feasible by incorporating multiple horizontal layers into each spreadsheet. However, it's worth noting that while implementing this approach may not be overly complex, the management of such a large spreadsheet could diminish its appeal as a tool for teaching or quick demonstrations.

The implemented methodology and results was published in the journal Mathematical Geosciences (Gómez-Hernández and Secci, 2023).

Conclusions

The objective of this thesis is to implement innovative surrogate modeling approaches for a sustainable use and protection of groundwater resources, also in the context of climate change. The study focuses on statistical approaches, conventional neural networks, deep learning techniques, and physics-informed neural networks. These diverse techniques offer a comprehensive toolkit for addressing various challenges including complexity, computational expense, and lack of physical interpretability, making surrogate modeling a valuable tool for groundwater management and sustainability.

In particular, the thesis recognizes climate change as a key factor influencing groundwater resources. Changing precipitation patterns and increasing temperatures significantly impact groundwater availability. Surrogate modeling techniques have been demonstrated as effective tools for facilitating the analysis of multiple climate change scenarios. As a result of the study, the strengths and weaknesses of different models, including a linear regression approach and three machine learning techniques, were analyzed. The linearity of the regression model, the lack of deep learning capability of the NARX, and the difficulty of extrapolating beyond the training range of the CNN make the LSTM the most promising approach for analyzing time-series data to assess the impacts of climate change on groundwater resources. The insights gained through these adaptive surrogate models are crucial for ensuring sustainable groundwater management in response to changing environmental conditions.

In addition, ANNs emerge as a promising technique for effectively addressing contamination issues. In fact, the implemented methodology can accurately assess and predict contamination scenarios (direct and inverse problems). The learning capabilities of ANNs enhance their ability to pinpoint potential sources of contamination and predict the spread of pollutants over time. Therefore, ANNs contribute to more sustainable and resilient environmental practices by offering robust solutions for the precise and proactive management of groundwater contamination.

Furthermore, the thesis discussed the lack of physics interpretability of surrogate models. A promising solution was introduced in the form of PINNs, a new approach that combines physics and machine learning. PINNs demonstrated their reliability in addressing complex groundwater problems. The comparison of results obtained by the PINN for solving unconfined groundwater flow with those obtained by a conventional ANN underscored their remarkable ability to yield accurate outcomes, particularly in scenarios with limited data.

Lastly, the incorporation of supplementary materials in this study expanded the exploration of ANNs beyond the original research field. The neural networks perform well in the classification task of identifying the illicit inflow source in a sewage system, resulting in satisfactory results. Moreover, the integration of a diverse perspective aimed to introduce alternative approaches to address the flow equation, provide a more comprehensive understanding of the intricacies involved in solving this equation. The use of spreadsheets, as illustrated in this thesis, empowers students to experiment with various configurations, delve into the underlying mathematical expressions, and acquire a profound understanding of more advanced numerical codes like MODFLOW.

To conclude, this thesis explores diverse surrogate modeling techniques to address groundwater challenges, emphasizing the significance of integrating physics, tackling climate change impacts, and proposing avenues for future research in this field.

Limitations and future directions

This section of the thesis provides an overview of limitations, future potential improvements and new lines of research.

The first challenge and limitation associated with surrogate modeling is the selection of appropriate methods able to map the investigated phenomena. Furthermore, there are other challenges that need to be addressed. Overfitting and the need for high-quality training data are two of the most important challenges in data-driven surrogate modeling. Overfitting occurs when the model learns the training data too well and fails to generalize to new data. This can lead to inaccurate predictions when the model is used to forecast on unseen data. Researchers are developing new methods to address these challenges. Early stopping techniques are one way to prevent overfitting. Early stopping works by monitoring the performance of the model on a held-out validation set and stopping the training when the performance on the validation set starts to decline.

Moreover, large amounts of high-quality training data are necessary to build the surrogate model. This can be expensive and time-consuming to collect, especially for complex problems. Data augmentation is a technique that can be used to improve the performance of data-driven models in data-limited conditions by artificially generating new training data from existing data. This can help to make the model more robust to noise and variability in the data.

In addition, in the realm of machine learning surrogate modeling, two common limitations are frequently encountered: the challenge of comprehending the inner workings of models and the reasons behind their predictions, and the incapacity to generate accurate predictions beyond their training range. Concerning the first limitation, despite the user-friendly nature of these models, their black-box characteristics often create issues when used in applications requiring a high level of confidence. This thesis has already explored future directions by integrating physics into the modeling process, as exemplified in the study utilizing PINNs for groundwater numerical simulations. This integration aims to enhance the understanding and robustness of the trained model, particularly in situations with limited data. Furthermore, PINNs present a prospective solution for the second limitation. While the inability to generate accurate predictions beyond the training range was a significant obstacle until recently, the introduction of PINNs enables the incorporation of physics to construct more reliable models that consider both data-driven insights and the fundamental principles of physics. However, intrinsically, this approach can represent a limitation, as it necessitates a thorough understanding of the physical phenomena being modeled, including the associated partial differential equations describing the problem. Additionally, the underlying partial differential equation inherently serves as an approximation of reality, and consequently, PINNs might enforce adherence to an imperfect model.

As a final point, it is worth discussing the approach employed in identifying hyperparameters in neural networks (such as number of layers and neurons, learning rate, decay rate, etc.). The current approach involves a manual process for determining hyperparameters, which could be time-consuming and could introduce potential of subjective judgments or personal biases into the process. Exploring and implementing automated methods for hyperparameter tuning in neural networks can not only streamline the model development process but also contribute to improved efficiency, ensuring the model to be well-optimized for the specific tasks at hand.

Future works could extend the application of surrogate models for simulating and forecasting contaminant movement in aquifers by investigating additional strategies, such as the application of Generative Adversarial Networks (GAN).

This recently implemented technique can address both direct and inverse problems related to contaminant transport. In fact, by leveraging GANs, it becomes possible to generate diverse and realistic scenarios of contaminant behavior, thereby enhancing the robustness of surrogate models in capturing the complexity of contaminant movement in aquifers. Furthermore, GANs offer a unique capability for addressing inverse problems by generating plausible contaminant source scenarios based on observed data. This can aid in identifying potential sources of contamination and refining the predictive accuracy of surrogate models.

Finally, to explore how different socioeconomic developments may influence future greenhouse gas emissions, upcoming research efforts will leverage the new climate scenarios known as Shared Socioeconomic Pathways (SSPs). However, at the moment RCM experiments, useful for vulnerability, impact and adaptation studies at regional and local scales, are still under development. A potential area for future exploration in this direction would be the use of artificial intelligence for downscaling and rectifying systematic errors in GCMs.

Bibliography

Afrifa, S., Zhang, T., Appiahene, P., and Varadarajan, V. (2022). Mathematical and Machine Learning Models for Groundwater Level Changes: A Systematic Review and Bibliographic Analysis. *Future Internet*, 14(9):259.

Akhter, M., Ahmad, Z., and Khan, K. (2006). Excel-based finite difference modeling of groundwater flow. *Journal of Himalayan Earth Sciences*, 39:49–53.

Allen, R., Pereira, L., Raes, D., and Smith, M. (1998). *FAO Irrigation and Drainage Paper No. 56*, volume 56. Food and Agriculture Organization of the United Nations, Rome, Italy.

Almajid, M. M. and Abu-Al-Saud, M. O. (2022). Prediction of porous media fluid flow using physics informed neural networks. *Journal of Petroleum Science and Engineering*, 208:109205.

Amidror, I. (2002). Scattered data interpolation methods for electronic imaging systems: a survey. *Journal of Electronic Imaging*, 11(2):157–176.

Ankor, M. J. and Tyler, J. J. (2019). Development of a spreadsheet-based model for transient groundwater modelling. *Hydrogeology Journal*, 27(5):1865–1878.

Asher, M. J., Croke, B. F., Jakeman, A. J., and Peeters, L. J. (2015). A review of surrogate models and their application to groundwater modeling. *Water Resources Research*, 51(8):5957–5973.

Ayaz, M. (2021). Estimation of release history of groundwater pollution source using ANN model. *Modeling Earth Systems and Environment*, (0123456789).

Ayvaz, M. T. (2010). A linked simulation-optimization model for solving the unknown groundwater pollution source identification problems. *Journal of Contaminant Hydrology*, 117(1-4):46–59.

Azizi, H., Ebrahimi, H., Samani, H. M. V., and Khaki, V. (2021). Evaluating the effects of climate change on groundwater level in the Varamin plain. *Water Science and Technology: Water Supply*, 21(3):1372–1384.

Bair, E. and Lahm, T. (2006). *Practical Problems in Groundwater Hydrology*. Prentice Hall, Hoboken.

Bajracharya, P. and Jain, S. (2022). Hydrologic similarity based on width function and hypsometry: An unsupervised learning approach. *Computers & Geosciences*, 163:105097.

Barati Moghaddam, M., Mazaheri, M., and Mohammad Vali Samani, J. (2021). Inverse modeling of contaminant transport for pollution source identification in surface and groundwaters: a review. *Groundwater for Sustainable Development*, 15(August):100651.

Bear, J. (2012). *Hydraulics of groundwater*. Dover Publications.

Beheshti, M. and Sægrov, S. (2018). Quantification assessment of extraneous water infiltration and inflow by analysis of the thermal behavior of the sewer network. *Water Switzerland*, 10.

Beheshti, M. and Sægrov, S. (2019). Detection of extraneous water ingress into the sewer system using tandem methods-a case study in trondheim city. *Water Science and Technology*, 79:231–239.

Bhattacharjya, R. K. (2011). Solving Groundwater Flow Inverse Problem Using Spreadsheet Solver. *Journal of Hydrologic Engineering*, 16(5):472–477.

Bloomfield, J. P. and Marchant, B. P. (2013). Analysis of groundwater drought building on the standardised precipitation index approach. *Hydrology and Earth System Sciences*, 17(12):4769–4787.

Bloomfield, J. P., Marchant, B. P., and McKenzie, A. A. (2019). Changes in groundwater drought associated with anthropogenic warming. *Hydrology and Earth System Sciences*, 23(3):1393–1408.

Box, G. and Wilson, K. B. (1951). On the Experimental Attainment of Optimum Conditions. *Journal of the Royal Statistical Society: Series B (Methodological)*, 13(1):1–38.

Brouyère, S., Carabin, G., and Dassargues, A. (2004). Climate change impacts on groundwater resources: Modelled deficits in a chalky aquifer, Geer basin, Belgium. *Hydrogeology Journal*, 12(2):123–134.

Butera, I., Tanda, M. G., and Zanini, A. (2013). Simultaneous identification of the pollutant release history and the source location in groundwater by means of a geostatistical approach. *Stochastic Environmental Research and Risk Assessment*, 27(5):1269–1280.

Cai, S., Wang, Z., Wang, S., Perdikaris, P., and Karniadakis, G. E. (2021). Physics-informed neural networks for heat transfer problems. *Journal of Heat Transfer*, 143(6).

Chang, J., Wang, G., and Mao, T. (2015). Simulation and prediction of suprapermafrost groundwater level variation in response to climate change using a neural network model. *Journal of Hydrology*, 529:1211–1220.

Chaubey, J. and Srivastava, R. (2022). Simultaneous identification of groundwater pollution source location and release concentration using Artificial Neural Network. *Environmental Forensics*, 23(3-4):293–300.

Chen, H., Wang, S., Gao, Z., and Hu, Y. (2010). Artificial neural network approach for quantifying climate change and human activities impacts on shallow groundwater level - A case study of Wuqiao in north China plain. *2010 18th International Conference on Geoinformatics, Geoinformatics 2010*.

Chen, Z., Xu, T., Gómez-Hernández, J. J., and Zanini, A. (2021). Contaminant Spill in a Sandbox with Non-Gaussian Conductivities: Simultaneous Identification by the Restart Normal-Score Ensemble Kalman Filter. *Mathematical Geosciences*, 53(7):1587–1615.

Citrini, A., Camera, C., and Beretta, G. P. (2020). Nossana Spring (Northern Italy) under Climate Change: Projections of future discharge rates and water availability. *Water (Switzerland)*, 12(2).

Coppola, E., Szidarovszky, F., Poulton, M., and Charles, E. (2003). Artificial neural network approach for predicting transient water levels in a multilayered groundwater system under variable state, pumping, and climate conditions. *Journal of Hydrologic Engineering*, 8(6):348–360.

Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3):273–297.

Croley, T. E. and Luukkonen, C. L. (2003). Potential effects of climate change on ground water in Lansing, Michigan. *Journal of the American Water Resources Association*, 39(1):149–163.

Cupola, F., Tanda, M. G., and Zanini, A. (2015). Contaminant release history identification in 2-D heterogeneous aquifers through a minimum relative entropy approach. *SpringerPlus*, 4(1):1–19.

Dierauer, J. R. and Zhu, C. (2020). Drought in the twenty-first century in awater-rich region: Modeling study of the Wabash River Watershed, USA. *Water (Switzerland)*, 12(1).

Diffenbaugh, N. S., Swain, D. L., Touma, D., and Lubchenco, J. (2015). Anthropogenic warming has increased drought risk in California. *Proceedings of the National Academy of Sciences of the United States of America*, 112(13):3931–3936.

D'Oria, M., Cozzi, C., and Tanda, M. G. (2018a). Future precipitation and temperature changes over the Taro, Parma and Enza River basins in Northern Italy. *Italian Journal of Engineering Geology and Environment*, (Special Issue):49–63.

D'Oria, M., Ferraresi, M., and Tanda, M. G. (2017). Historical trends and high-resolution future climate projections in northern Tuscany (Italy). *Journal of Hydrology*, 555:708–723.

D'Oria, M., Ferraresi, M., and Tanda, M. G. (2019). Quantifying the impacts of climate change on water resources in northern Tuscany, Italy, using high-resolution regional projections. *Hydrological Processes*, 33(6):978–993.

D'Oria, M., Tanda, M. G., and Todaro, V. (2018b). Assessment of local climate change: Historical trends and rcm multi-model projections over the salento area (italy). *Water (Switzerland)*, 10(8):978.

Earth Security Group (2016). CEO Briefing: Global Depletion of Aquifers. Global companies must take an active role in groundwater governance to avoid existential risks. page 20.

Elfeki, A. M. and Bahrawi, J. (2015). A fully distributed spreadsheet modeling as a tool for analyzing groundwater level rise problem in Jeddah city. *Arabian Journal of Geosciences*, 8(4):2313–2325.

Emamgholizadeh, S., Moslemi, K., and Karami, G. (2014). Prediction the Groundwater Level of Bastam Plain (Iran) by Artificial Neural Network (ANN) and Adaptive Neuro-Fuzzy Inference System (ANFIS). *Water Resources Management*, 28(15):5433–5446.

(EPA), U. E. P. A. (2014). *Guide for Estimating Infiltration and Inflow, June 2014.*

Evans, J. (1996). *Straightforward Statistics for the Behavioral Sciences.* Pacific gr edition.

Fox, P. (1996). Spreadsheet solution method for groundwater flow problems. In *Subsurface Fluid-Flow (Ground-Water and Vadose Zone) Modeling.* ASTM International.

Ghazi, B., Jeihouni, E., and Kalantari, Z. (2021). Predicting groundwater level fluctuations under climate change scenarios for Tasuj plain, Iran. *Arabian Journal of Geosciences*, 14(2).

Gómez-Hernández, J. J. (2022). Teaching Numerical Groundwater Flow Modeling with Spreadsheets. *Mathematical Geosciences*, 54(6):1121–1138.

Gómez-Hernández, J. J. and Secci, D. (2023). Teaching Numerical Groundwater Flow Modeling with Spreadsheets: Unconfined Aquifers and Multilayered Vertical Cross-Sections. *Mathematical Geosciences.*

Gómez-Hernández, J. J. and Xu, T. (2022). Contaminant Source Identification in Aquifers: A Critical View. *Mathematical Geosciences*, 54(2):437–458.

Gonzalez, R. Q. and Arsanjani, J. J. (2021). Prediction of groundwater level variations in a changing climate: A danish case study. *ISPRS International Journal of Geo-Information*, 10(11).

Google LLC (2023). Google earth.

Griewank, A. (1988). On automatic differentiation. Preprint ANL/MCS-P10-1088, Argonne National Laboratory, 9700 South Cass Avenue, Argonne, Illinois 60439.

Guo, M., Yue, W., Wang, T., Zheng, N., and Wu, L. (2021). Assessing the use of standardized groundwater index for quantifying groundwater drought over the conterminous US. *Journal of Hydrology*, 598(March):126227.

Guo, W. (1997). Transient groundwater flow between reservoirs and water-table aquifers. *Journal of Hydrology*, 195(1-4):370–384.

Gupta, R., Bhattarai, R., and Mishra, A. (2019). Development of climate data bias corrector (CDBC) tool and its application over the agro-ecological zones of India. *Water (Switzerland)*, 11(5).

Gzyl, G., Zanini, A., Frączek, R., and Kura, K. (2014). Contaminant source and release history identification in groundwater: A multi-step approach. *Journal of Contaminant Hydrology*, 157:59–72.

Hagan, M. T. and Menhaj, T. M. (1994). Training feed-forward networks with the marquardt algorithm. *IEEE Transactions on Neural Networks*, 5(6):989–993.

Harbaugh, A. W. (2005). *MODFLOW-2005, the US Geological Survey modular ground-water model: the ground-water flow process*, volume 6. US Department of the Interior, US Geological Survey Reston, VA, USA.

He, Q., Barajas-Solano, D., Tartakovsky, G., and Tartakovsky, A. M. (2020). Physics-informed neural networks for multiphysics data assimilation with application to subsurface transport. *Advances in Water Resources*, 141:103610.

He, Q. and Tartakovsky, A. M. (2021). Physics-informed neural network method for forward and backward advection-dispersion equations. *Water Resources Research*, 57(7):e2020WR029479.

Hochreiter, S. (1991). Untersuchungen zu dynamischen neuronalen Netzen. *Institut für Informatik, Technische Universität, Munchen*, pages 1–71.

Hochreiter, S. and Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8):1735–1780.

Horová, I., Koláček, J., and Zelinka, J. (2012). *Kernel smoothing in MATLAB: Theory and practice of kernel smoothing.* World Scientific Publishing Co.

Hotelling, H. (1936). Relations Between Two Sets of Variates. *Biometrika*, 28(3/4):321–377.

Hou, T. Y. and Wu, X. H. (1997). A multiscale finite element method for elliptic problems in composite materials and porous media. *Journal of Computational Physics*, 134(1):169–189.

Idrizovic, D., Pocuca, V., Vujadinovic Mandic, M., Djurovic, N., Matovic, G., and Gregoric, E. (2020). Impact of climate change on water resource availability in a mountainous catchment: A case study of the Toplica River catchment, Serbia. *Journal of Hydrology*, 587(April 2019):124992.

Intergovernmental Panel on Climate Change (2021). *Climate Change 2021 The Physical Science Basis Summary for Policymakers Working Group I Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.*

Jackson, C. R., Bloomfield, J. P., and Mackay, J. D. (2015). Evidence for changes in historic and future groundwater levels in the UK. *Progress in Physical Geography*, 39(1):49–67.

Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O. B., Bouwer, L. M., Braun, A., Colette, A., Déqué, M., Georgievski, G., Georgopoulou, E., Gobiet, A., Menut, L., Nikulin, G., Haensler, A., Hempelmann, N., Jones, C., Keuler, K., Kovats, S., Kröner, N., Kotlarski, S., Kriegsmann, A., Martin, E., van Meijgaard, E., Moseley, C., Pfeifer, S., Preuschmann, S., Radermacher, C., Radtke, K., Rechid, D., Rounsevell, M., Samuelsson, P., Somot, S., Soussana, J. F., Teichmann, C., Valentini, R., Vautard, R., Weber, B., and Yiou, P. (2014). EURO-CORDEX: New high-resolution climate change projections for European impact research. *Regional Environmental Change*, 14(2):563–578.

Jamshidi, A., Samani, J. M. V., Samani, H. M. V., Zanini, A., Tanda, M. G., and Mazaheri, M. (2020). Solving inverse problems of unknown contaminant source in groundwater-river integrated systems using a surrogate transport model based optimization. *Water (Switzerland)*, 12(9).

Javadinejad, S., Dara, R., and Jafary, F. (2020). How groundwater level can predict under the effect of climate change by using artificial neural networks of NARX. *Resources Environment and Information Engineering*, 2(1):90–99.

Jeihouni, E., Mohammadi, M., Eslamian, S., and Zareian, M. J. (2019). Potential impacts of climate change on groundwater level through hybrid soft-computing methods: a case study—Shabestar Plain, Iran. *Environmental Monitoring and Assessment*, 191(10).

Jenny, P., Lee, S. H., and Tchelepi, H. A. (2004). Adaptive multiscale finite-volume method for multiphase flow and transport in porous media. *Multiscale Modeling and Simulation*, 3(1):50–64.

Jiménez Cisneros, B. E., Oki, T., Arnell, N. W., Benito, G., Cogley, J. G., Döll, P., Jiang, T., Mwakalila, S. S., Kundzewicz, Z., and Nishijima, A. (2015). Freshwater resources. *Climate Change 2014 Impacts, Adaptation and Vulnerability: Part A: Global and Sectoral Aspects*, pages 229–270.

Jolliffe, I. T. and Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, 374(2065):20150202.

Karahan, H. and Ayvaz, M. T. (2005a). Time-dependent groundwater modeling using spreadsheet. *Computer Applications in Engineering Education*, 13(3):192–199.

Karahan, H. and Ayvaz, M. T. (2005b). Transient groundwater modeling using spreadsheets. *Advances in Engineering Software*, 36(6):374–384.

Karthikeyan, L., Kumar, D. N., Graillot, D., and Gaur, S. (2013). Prediction of Ground Water Levels in the Uplands of a Tropical Coastal Riparian Wetland using Artificial Neural Networks. *Water Resources Management*, 27(3):871–883.

Katsanou, K. and Karapanagioti, H. K. (2019). Surface water and groundwater sources for drinking water. *Handbook of Environmental Chemistry*, 67(April 2022):1–19.

Khan, S., Gabriel, H. F., and Rana, T. (2008). Standard precipitation index to track drought and assess impact of rainfall on watertables in irrigation areas. *Irrigation and Drainage Systems*, 22(2):159–177.

Kingma, D. P. and Ba, J. L. (2015). Adam: A method for stochastic optimization. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, pages 1–15.

Krige, D. (1951). A statistical approach to some basic mine valuation problems on the witwatersrand. *Journal of the Chemical, Metallurgical and Mining Society of South Africa*, 52:119–139.

Kumar, R., Musuuza, J. L., Van Loon, A. F., Teuling, A. J., Barthel, R., Ten Broek, J., Mai, J., Samaniego, L., and Attinger, S. (2016). Multiscale evaluation of the Standardized Precipitation Index as a groundwater drought indicator. *Hydrology and Earth System Sciences*, 20(3):1117–1131.

Lallahem, S., Mania, J., Hani, A., and Najjar, Y. (2005). On the use of neural networks to evaluate groundwater levels in fractured media. *Journal of Hydrology*, 307(1-4):92–111.

Lawal, Z. K., Yassin, H., Lai, D. T. C., and Che Idris, A. (2022). Physics-informed neural network (pinn) evolution and beyond: A systematic literature review and bibliometric analysis. *Big Data and Cognitive Computing*, 6(4).

LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., and Jackel, L. D. (1990). Backpropagation applied to handwritten zip code recognition. *AT&T Bell Laboratories.*

Leelaruban, N., Padmanabhan, G., and Oduor, P. (2017). Examining the relationship between drought indices and groundwater levels. *Water (Switzerland)*, 9(2).

Lepot, M., Makris, K. F., and Clemens, F. H. L. R. (2017). Detection and quantification of lateral, illicit connections and infiltration in sewers with infra-red camera: Conclusions after a wide experimental plan. *Water Research*, 122:678–691.

Liu, C. and Ball, W. P. (1999). Application of inverse methods to contaminant source identification from aquitard diffusion profiles at Dover AFB, Delaware. *Water Resources Research*, 35(7):1975–1985.

Lv, A., Cheng, L., Aghighi, M. A., Masoumi, H., and Roshan, H. (2021). A novel workflow based on physics-informed machine learning to determine the permeability profile of fractured coal seams using downhole geophysical logs. *Marine and Petroleum Geology*, 131:105171.

Malcolm, R. and Soulsby, C. (2000). Modelling the potential impact of climate change on a shallow coastal aquifer in northern Scotland. *Geological Society Special Publication*, 182:191–204.

Mao, Z., Jagtap, A. D., and Karniadakis, G. E. (2020). Physics-informed neural networks for high-speed flows. *Computer Methods in Applied Mechanics and Engineering*, 360:112789.

Mariethoz, G. and Gómez-Hernández, J. J. (2021). Machine learning for water resources. *Frontiers in Artificial Intelligence*, 4:63.

Mascaro, G., Viola, F., and Deidda, R. (2018). Evaluation of Precipitation From EURO-CORDEX Regional Climate Simulations in a Small-Scale Mediterranean Site. *Journal of Geophysical Research: Atmospheres*, 123(3):1604–1625.

Matheron, G., Blondel, F., and *de recherches géologiques et minières (France)., Bureau* (1962). Traité de géostatistique appliquée. Tome I. Technip Paris, Paris SE - 1 vol. (333 p.) : ill. ; 28 cm.

MathWorks (2022). Matlab version: 9.13.0 (r2022b).

McCulloch, W. S. and Pitts, W. (1943). A logical calculus of the ideas imminent in nervous activity. *Bull. Math. Biophys.*, 5(4):115–133.

McKay, M. D., Beckman, R. J., and Conover, W. J. (1979). A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. *Technometrics*, 21(2):239–245.

Mckee, T. B., Doesken, N. J., and Kleist, J. (1993). THE RELATIONSHIP OF DROUGHT FREQUENCY AND DURATION TO TIME SCALES. Technical report.

Meenal, M. and Eldho, T. (2011). Simulation of groundwater flow in unconfined aquifer using meshfree point collocation method. *Engineering Analysis with Boundary Elements*, 35(4):700–707.

Minsky, M. and Papert, S. (1969). *Perceptrons.* M.I.T. Press.

Mohanty, S., Jha, M. K., Kumar, A., and Sudheer, K. P. (2010). Artificial neural network modeling for groundwater level forecasting in a river island of eastern India. *Water Resources Management*, 24(9):1845–1865.

Moisello, U. (1998). Idrologia tecnica.

Molano, C. (2014). Métodos de evaluación de flujo y contaminación de acuíferos utilizando hojas de cálculo. In *2014 NGWA Groundwater Expo and Annual Meeting.* NGWA.

Moss, R. H., Edmonds, J. A., Hibbard, K. A., Manning, M. R., Rose, S. K., Van Vuuren, D. P., Carter, T. R., Emori, S., Kainuma, M., Kram, T., Meehl, G. A., Mitchell, J. F., Nakicenovic, N., Riahi, K., Smith, S. J., Stouffer, R. J., Thomson, A. M., Weyant, J. P., and Wilbanks, T. J. (2010). The next generation of scenarios for climate change research and assessment. *Nature*, 463(7282):747–756.

Niazkar, M. and Afzali, S. (2015). Application of excel spreadsheet in engineering education. In *First International and Fourth National Conference on Engineering Education*, pages 10–12. Shiraz University.

of Massachusetts, C. (1993). *Guidelines for Performing Infiltration/Inflow Analyses And Sewer System Evaluation Survey.*

Olsthoorn, T. N. (1985). THE POWER OF THE ELECTRONIC WORKSHEET: MODELING WITHOUT SPECIAL PROGRAMS. *Groundwater*, 23(3):381–390.

Olsthoorn, T. N. (1999). *Groundwater Modelling: Calibration and the Use of Spreadsheets*. PhD thesis, Ph.D. thesis.

Osti, A. L., Lambert, M. F., and Metcalfe, A. V. (2008). On spatiotemporal drought classification in New South Wales: Development and evaluation of alternative techniques. *Aust. J. Water Resour.*, 12:21–34.

Osuch, M., Romanowicz, R. J., Lawrence, D., and Wong, W. K. (2016). Trends in projections of standardized precipitation indices in a future climate in Poland. *Hydrology and Earth System Sciences*, 20(5):1947–1969.

Pachauri, R. K., Meyer, L., Hallegatte France, S., Bank, W., Hegerl, G., Brinkman, S., van Kesteren, L., Leprince-Ringuet, N., and van Boxmeer, F. (2014). AR5 Synthesis Report: Climate Change 2014 — IPCC. Technical report.

Pan, Z., Lu, W., and Bai, Y. (2022). Groundwater contamination source estimation based on a refined particle filter associated with a deep residual neural network surrogate. *Hydrogeology Journal*, 30(3):881–897.

Panasiuk, O., Hedström, A., Langeveld, J., de Haan, C., Liefting, E., Schilperoort, R., and Viklander, M. (2019). Using distributed temperature sensing (dts) for locating and characterising infiltration and inflow into foul sewers before, during and after the snowmelt period. *Water Switzerland*, 11:1–12.

Pandit, S. M. and Wu, S.-M. (1983). *Time Series and System Analysis with Applications*. John Wiley & Sons.

Pearson, K. (1901). LIII. On lines and planes of closest fit to systems of points in space . *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 2(11):559–572.

Piani, C., Haerter, J. O., and Coppola, E. (2010). Statistical bias correction for daily precipitation in regional climate models over europe. *Theoretical and Applied Climatology*, 99:187–192.

Pranzini, G., Martino, F., Fanti, R., and Fontanelli, K. (2019). Map of the vulnerabiliy to pollution of the Apuo-Versilia aquifer (Tuscany - Italy). *Acque Sotterranee - Italian Journal of Groundwater*, 8.

PTA (2005). Piano di tutela delle acque della toscana n.6 del 25 gennaio 2005. Accessed 2.18.21.

Pulido-Velazquez, D., Sahuquillo, A., Andreu, J., and Pulido-Velazquez, M. (2007). A general methodology to simulate groundwater flow of unconfined aquifers with a reduced computational cost. *Journal of Hydrology*, 338(1-2):42–56.

Raissi, M., Perdikaris, P., and Karniadakis, G. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707.

Rajaee, T., Ebrahimi, H., and Nourani, V. (2019). A review of the artificial intelligence methods in groundwater level modeling. *Journal of Hydrology*, 572(December 2018):336–351.

Rasmussen, C. E. (2004). Gaussian Processes in machine learning. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 3176:63–71.

Razavi, S., Tolson, B. A., and Burn, D. H. (2012). Review of surrogate modeling in water resources. *Water Resources Research*, 48(7).

Rezaee, M. and Tabesh, M. (2022). Effects of inflow, infiltration, and exfiltration on water footprint increase of a sewer system: A case study of tehran. *Sustainable Cities and Society*, 79:103707.

Rezaei, S., Harandi, A., Moeineddin, A., Xu, B.-X., and Reese, S. (2022). A mixed formulation for physics-informed neural networks as a potential solver for engineering problems in heterogeneous domains: Comparison with finite element method. *Computer Methods in Applied Mechanics and Engineering*, 401:115616.

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6):386–408.

Rossman, L. A. and Huber, W. C. (2016). *Storm Water Management Model Reference Manual*. Number Res Dev III 231.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088):533–536.

Sahoo, S. and Jha, M. K. (2013). Groundwater-level prediction using multiple linear regression and artificial neural network techniques: A comparative assessment. *Hydrogeology Journal*, 21(8):1865–1887.

Secci, D., A. Godoy, V., and Gómez-Hernández, J. J. (2024). Physics-Informed Neural Networks for solving transient unconfined groundwater flow. *Computers & Geosciences*, 182(November 2023):105494.

Secci, D., Molino, L., and Zanini, A. (2022). Contaminant source identification in groundwater by means of artificial neural network. *Journal of Hydrology*, 611.

Secci, D., Tanda, M. G., D'Oria, M., and Todaro, V. (2023). Artificial intelligence models to evaluate the impact of climate change on groundwater resources. *Journal of Hydrology*, 627(PB):130359.

Secci, D., Tanda, M. G., D'Oria, M., Todaro, V., and Fagandini, C. (2021). Impacts of climate change on groundwater droughts by means of standardized indices and regional climate models. *Journal of Hydrology*, 603(PD):127154.

Shadab, M. A., Luo, D., Shen, Y., Hiatt, E., and Hesse, M. A. (2021). Investigating steady unconfined groundwater flow using physics informed neural networks. *arXiv preprint arXiv:2112.13792*.

Shiri, J., Kisi, O., Yoon, H., Lee, K. K., and Hossein Nazemi, A. (2013). Predicting groundwater level fluctuations with meteorological effect implications-A comparative study among soft computing techniques. *Computers and Geosciences*, 56:32–44.

Singh, R. M. and Datta, B. (2004). Groundwater pollution source identification and simultaneous parameter estimation using pattern matching by artificial neural network. *Environmental Forensics*, 5(3):143–153.

Singh, R. M., Datta, B., and Jain, A. (2004). Identification of Unknown Groundwater Pollution Sources Using Artificial Neural Networks. *Journal of Water Resources Planning and Management*, 130(6):506–514.

Sit, M., Demiray, B. Z., Xiang, Z., Ewing, G. J., Sermet, Y., and Demir, I. (2020). A comprehensive review of deep learning applications in hydrology and water resources. *Water Science and Technology*, 82(12):2635–2670.

Skaggs, T. H. and Kabala, Z. J. (1994). Recovering the release history of a groundwater contaminant. *Water Resources Research*, 30(1):71–79.

Soleimani Motlagh, M., Ghasemieh, H., Talebi, A., and Abdollahi, K. (2017). Identification and Analysis of Drought Propagation of Groundwater During Past and Future Periods. *Water Resources Management*, 31(1):109–125.

Soľáková, T., De Michele, C., and Vezzoli, R. (2014). Comparison between Parametric and Nonparametric Approaches for the Calculation of Two Drought Indices: SPI and SSI. *Journal of Hydrologic Engineering*, 19(9):1–11.

Stagge, J., Tallaksen, L., and Rizzi, J. (2015a). Future meteorological drought: projections of regional climate models for Europe. *Geophysical Research Abstracts*, 17(25):2015–7749.

Stagge, J. H., Tallaksen, L. M., Gudmundsson, L., Van Loon, A. F., and Stahl, K. (2015b). Candidate Distributions for Climatological Drought Indices (SPI and SPEI). *International Journal of Climatology*, 35(13):4027–4040.

Staufer, P., Scheidegger, A., and Rieckermann, J. (2012). Assessing the performance of sewer rehabilitation on the reduction of infiltration and inflow. *Water Research*, 46:5185–5196.

Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M. (2013). *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.* Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Suryanarayana, C., Sudheer, C., Mahammood, V., and Panigrahi, B. K. (2014). An integrated wavelet-support vector machine for groundwater level prediction in Visakhapatnam, India. *Neurocomputing*, 145:324–335.

Tahmasebi, P. and Sahimi, M. (2021). Special issue on machine learning for water resources and subsurface systems.

Taigbenu, A. and Nyirenda, E. (2010). Revisiting the stream-aquifer flow problem with a flux-based Green element model. *Water SA*, 36:287 – 294.

Tanda, M. G., D'Oria, M., Secci, D., and Todaro, V. (2023). Identification of the inflow source in a foul sewer system through techniques of inverse modelling. *Journal of Physics: Conference Series*, 2444(1).

Tao, H., Hameed, M. M., Marhoon, H. A., Zounemat-Kermani, M., Heddam, S., Sungwon, K., Sulaiman, S. O., Tan, M. L., Sa'adi, Z., Mehr, A. D., Allawi, M. F., Abba, S. I., Zain, J. M., Falah, M. W., Jamei, M., Bokde, N. D., Bayatvarkeshi, M., Al-Mukhtar, M., Bhagat, S. K., Tiyasha, T., Khedher, K. M., Al-Ansari, N., Shahid, S., and Yaseen, Z. M. (2022). Groundwater level prediction using machine learning models: A comprehensive review. *Neurocomputing*, 489:271–308.

Taormina, R., Chau, K. W., and Sethi, R. (2012). Artificial neural network simulation of hourly groundwater levels in a coastal aquifer system of the Venice lagoon. *Engineering Applications of Artificial Intelligence*, 25(8):1670–1676.

Teuling, A. J., Van Loon, A. F., Seneviratne, S. I., Lehner, I., Aubinet, M., Heinesch, B., Bernhofer, C., Grünwald, T., Prasse, H., and Spank, U. (2013). Evapotranspiration amplifies European summer drought. *Geophysical Research Letters*, 40(10):2071–2075.

Teutschbein, C. and Seibert, J. (2012). Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. *Journal of Hydrology*, 456-457:12–29.

Thornthwaite, C. W. (1948). An Approach toward a Rational Classification of Climate. *Geographical Review*, 38(1):55.

Todaro, V., D'Oria, M., Secci, D., Zanini, A., and Tanda, M. G. (2022a). Climate Change over the Mediterranean Region: Local Temperature and Precipitation Variations at Five Pilot Sites. *Water (Switzerland)*, 14(16).

Todaro, V., D'Oria, M., Tanda, M. G., and Gómez-Hernández, J. J. (2021). Ensemble smoother with multiple data assimilation to simultaneously estimate the source location and the release history of a contaminant spill in an aquifer. *Journal of Hydrology*, 598(January).

Todaro, V., D'Oria, M., Tanda, M. G., and Gómez-Hernández, J. J. (2022b). genES-MDA: A generic open-source software package to solve inverse problems via the Ensemble Smoother with Multiple Data Assimilation. *Computers and Geosciences*, 167(July).

Trichakis, I. C., Nikolos, I. K., and Karatzas, G. P. (2011). Artificial Neural Network (ANN) Based Modeling for Karstic Groundwater Level Simulation. *Water Resources Management*, 25(4):1143–1152.

Uddameri, V., Singaraju, S., and Hernandez, E. A. (2019). Is Standardized Precipitation Index (SPI) a Useful Indicator to Forecast Groundwater Droughts? — Insights from a Karst Aquifer. *Journal of the American Water Resources Association*, 55(1):70–88.

UNFCCC (2015). The Paris Agreement. *Conference: Paris Climate Change Conference - November 2015.*

United Nations (2022). *2022. The United Nations World Water Development Report 2022. Groundwater: Making the Invisible Visible. Paris, UNESCO. https://unesdoc. unesco.org/ark:/48223/pf0000380721.*

U.S. Geological Survey (2018). Groundwater decline and depletion. Report, U.S. Geological Survey.

Van Loon, A. F., Kumar, R., and Mishra, V. (2017). Testing the use of standardised indices and GRACE satellite data to estimate the European 2015 groundwater drought in near-real time. *Hydrology and Earth System Sciences*, 21(4):1947–1971.

Vicente-Serrano, S. M., Beguería, S., and López-Moreno, J. I. (2010). A multiscalar drought index sensitive to global warming: The standardized precipitation evapotranspiration index. *Journal of Climate*, 23(7):1696–1718.

Vicente-Serrano, S. M., Lopez-Moreno, J. I., Beguería, S., Lorenzo-Lacruz, J., Sanchez-Lorenzo, A., García-Ruiz, J. M., Azorin-Molina, C., Morán-Tejeda, E., Revuelto, J., Trigo, R., Coelho, F., and Espejo, F. (2014). Evidence of increasing drought severity caused by temperature rise in southern Europe. *Environmental Research Letters*, 9(4).

Vicente-Serrano, S. M., McVicar, T. R., Miralles, D. G., Yang, Y., and Tomas-Burguera, M. (2020). Unraveling the influence of atmospheric evaporative demand on drought and its response to climate change. *Wiley Interdisciplinary Reviews: Climate Change*, 11(2).

Vidal, J. P., Martin, E., Franchistéguy, L., Habets, F., Soubeyroux, J. M., Blanchard, M., and Baillon, M. (2010). Multilevel and multiscale drought reanalysis over France with the Safran-Isba-Modcou hydrometeorological suite. *Hydrology and Earth System Sciences*, 14(3):459–478.

Vosse, M., Schilperoort, R., de Haan, C., Nienhuis, J., Tirion, M., and Langeveld, J. (2013). Processing of dts monitoring results: Automated detection of illicit connections. *Water Practice and Technology*, 8:375–381.

Waheed, U., Haghighat, E., Alkhalifah, T., Song, C., and Hao, Q. (2021). Pinneik: Eikonal solution using physics-informed neural networks. *Computers & Geosciences*, 155:104833.

Wang, H., Lu, W., and Chang, Z. (2021). An iterative updating heuristic search strategy for groundwater contamination source identification based on an ACPSO–ELM surrogate system. *Stochastic Environmental Research and Risk Assessment*, 2.

Wang, W., Ertsen, M. W., Svoboda, M. D., and Hafeez, M. (2016). Propagation of drought: From meteorological drought to agricultural and hydrological drought. *Advances in Meteorology*, 2016(4).

Wang, Z., Lu, W., Chang, Z., and Wang, H. (2022). Simultaneous identification of groundwater contaminant source and simulation model parameters based on an ensemble Kalman filter – Adaptive step length ant colony optimization algorithm. *Journal of Hydrology*, 605(December 2021):127352.

Werbos, P. (1990). Backpropagation through time: what it does and how to do it. *Proceedings of the IEEE*, 78(10):1550–1560.

Williams, C. (2001). Excel calculation process.

Wittenberg, H. and Aksoy, H. (2010). Groundwater intrusion into leaky sewer systems. *Water Science and Technology*, 62:92–98.

Woodbury, A. D. and Ulrych, T. J. (1996). Minimum Relative Entropy Inversion: Theory and Application to Recovering the Release History of a Groundwater Contaminant. *Water Resources Research*, 32(9):2671–2681.

World Meteorological Organization (1987). Standardized Precipitation Index User Guide. *Journal of Applied Bacteriology*, 63(3):197–200.

Xing, Z., Qu, R., Zhao, Y., Fu, Q., Ji, Y., and Lu, W. (2019). Identifying the release history of a groundwater contaminant source based on an ensemble surrogate model. *Journal of Hydrology*, 572(February):501–516.

Xu, T. and Gómez-Hernández, J. J. (2016). Joint identification of contaminant source location, initial release time, and initial solute concentration in an aquifer via ensemble Kalman filtering. *Water Resources Research*, 52(8):6587–6595.

Yang, L., Meng, X., and Karniadakis, G. E. (2021). B-pinns: Bayesian physics-informed neural networks for forward and inverse pde problems with noisy data. *Journal of Computational Physics*, 425:109913.

Yılmaz, G., Çolak, M. A., Özgencil, İ. K., Metin, M., Korkmaz, M., Ertuğrul, S., Soyluer, M., Bucak, T., Ülkü Nihan Tavşanoğlu, Özkan, K., Akyürek, Z., Beklioğlu, M., and Jeppesen, E. (2021). Decadal changes in size, salinity, waterbirds, and fish in lakes of the konya closed basin, turkey, associated with climate change and increasing water abstraction for agriculture. *Inland Waters*, 11(4):538–555.

Yoon, H., Jun, S. C., Hyun, Y., Bae, G. O., and Lee, K. K. (2011). A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer. *Journal of Hydrology*, 396(1-2):128–138.

Zanini, A. and Woodbury, A. D. (2016). Contaminant source reconstruction by empirical Bayes and Akaike's Bayesian Information Criterion. *Journal of contaminant hydrology*, 185-186:74–86.

Zhang, X., Zhu, Y., Wang, J., Ju, L., Qian, Y., Ye, M., and Yang, J. (2022). GW-PINN: A deep learning algorithm for solving groundwater flow equations. *Advances in Water Resources*, 165(September 2021):104243.

Zheng, C. (1999). MT3DMS : A Modular Three-Dimensional Multispecies Transport Model for Simulation of Advection , Dispersion , and Chemical Reactions of Contaminants in Groundwater Systems ; Documentation and User ' s Guide by. *US Army Corps of Engineers*, (December):220.

Zheng, Y. and Wu, Z. (2023). Physics-informed online machine learning and predictive control of nonlinear processes with parameter uncertainty. *Industrial & Engineering Chemistry Research*, 62(6):2804–2818.