University of Zagreb

Faculty of Mining, Geology and Petroleum Engineering

Ana Kamenski

IMPROVEMENT OF THE DEEP-GEOLOGICAL CHARACTERIZATION OF THE EASTERN PART OF THE DRAVA DEPRESSION – SPATIAL PREDICTION OF LITHOLOGICAL PROPERTIES BASED ON SEISMIC AND WELL DATA

DISSERTATION

Zagreb, 2025

University of Zagreb

Faculty of Mining, Geology and Petroleum Engineering

Ana Kamenski

IMPROVEMENT OF THE DEEP-GEOLOGICAL CHARACTERIZATION OF THE EASTERN PART OF THE DRAVA DEPRESSION – SPATIAL PREDICTION OF LITHOLOGICAL PROPERTIES BASED ON SEISMIC AND WELL DATA

DISSERTATION

Zagreb, 2025

Sveučilište u Zagrebu

Rudarsko-geološko-naftni fakultet

Ana Kamenski

UNAPREĐENJE DUBINSKO-GEOLOŠKE KARAKTERIZACIJE ISTOČNE DRAVSKE DEPRESIJE – PROSTORNO PROGNOZIRANJE LITOLOŠKIH KARAKTERISTIKA NA TEMELJU SEIZMIČKIH I BUŠOTINSKIH PODATAKA

DOKTORSKI RAD

Mentor:
dr. sc. Marko Cvetković, izvanredni profesor

Zagreb, 2025

Supervisor:

Assoc. Prof. Marko Cvetković, PhD

University of Zagreb

Faculty of Mining, Geology and Petroleum Engineering

Department of Geology and Geological Engineering

ZAHVALE

Prvenstveno se zahvaljujem mentoru dr. sc. Marku Cvetkoviću koji me prihvatio, dao ideju i pomogao ju provesti u djelo i što smo uspjeli postići velike rezultate bez klasičnog financiranja doktorskog istraživanja. Zahvaljujem se i na svevremenskoj podršci i vjeri u dobre ishode tijekom cijelog ovog procesa, a najviše na beskonačnom strpljenju tijekom mog razvoja.

Želim se posebice zahvaliti dr. sc. Ivi Kolenković Močilac, koja je srčano navijala za mene od samog početka, koja je odradila i dobar dio posla tijekom provedbe ovog istraživanja. Hvala joj na svim razgovorima koji su sezali mnogo dalje od interpretacije karotažnih dijagrama i geologije općenito.

Od srca se zahvaljujem članovima Povjerenstva za ocjenu i obranu, dr. sc. Marku Špeliću, dr. sc. Josipi Kapuralić i dr. sc. Ivi Kolenković Močilac, čije su konstruktivne sugestije značajno doprinijele kvaliteti završne verzije ovog doktorata.

Također se zahvaljujem svim recenzentima znanstvenih radova koji čine glavni sadržaj ovog doktorata.

Iskreno zahvaljujem Rudarsko-geološko-naftnom fakultetu na podršci u svim mojim ostvarenjima tijekom studija, a posebice na dodijeljenoj Stipendiji koja je pokrila troškove školarine Poslijediplomskog doktorskog studija.

Zahvaljujem svim članovima projekta GEODEP, čiji sam dio bila od samog početka.

Zahvaljujem se Agenciji za ugljikovodike, bez čijeg dopuštenja ne bih imala pristup podacima na kojima se temelje svi znanstveni doprinosi stečeni tijekom ovog istraživanja.

Veliko hvala svim kolegama s Hrvatskog geološkog instituta na podršci, a posebno dragom prijatelju i kolegi dr. sc. Matiji Vukovskom za sve znanstvene i životne rasprave te na nesebičnoj podršci u oba aspekta. Svakako se zahvaljujem i dr. sc. Tvrtku Korbaru na bezuvjetnom razumijevanju i slobodi u organizaciji mojih obveza na njegovim projektima i ukazanom povjerenju, na čemu sam iznimno zahvalna. Posebno se zahvaljujem i dr. sc. Aniti Grizelj na beskrajnom povjerenju i razumijevanju. Veliko hvala svim "HGI-babama" koje mi svaki radni dan čine ugodnijim i ljepšim.

Od srca zahvaljujem "Malim Garavima" koje mi pružaju utjehu u svim vremenima, nose moje poraze i pobjede zajedno samnom, neke i većinu mog života.

Na kraju, najviše od svega želim zahvaliti Mami, Nevenu i Javoru, koji su mi bili i ostali bezuvjetna podrška kad god sam ju trebala, i kad sam ju zaslužila i u trenucima kad ju nisam zaslužila: Hvala im što su me još u najranijim danima naučili da je ulaganje u znanje jedino ulaganje bez kojeg nikada ne mogu ostati. Kosti sam vječno zahvalna na bezuvjetnoj ljubavi, podršci i razumijevanju…

ABSTRACT

This dissertation presents the development of an AI-supported methodology that integrates artificial neural networks (ANNs) to enhance subsurface characterization of the Pannonian age clastic interval in the Drava Basin. The aim was to improve the accuracy of the geological model and reduce economic and technical risks in mature basins with incomplete datasets by combining machine learning with geological expertise. The research was conducted in three phases. In the first phase, ANNs were combined with conventional geostatistical methods to improve lithology prediction. Although the approach showed potential, its limitations became evident in the overestimation of certain lithological categories, particularly sandstone, as it was predicted with greater certainty during the process preceding the definition of lithological values in the model (upscaling). Nevertheless, the integration of ANN models with fundamental geological knowledge proved successful in modeling lithological composition. The second phase introduced an ANN-supported approach for time-to-depth relationship prediction, significantly outperforming traditional techniques. Models achieved correlation coefficients above 0.99, attributed to their ability to capture complex geological patterns. Results also highlight the importance of developing locally specific ANN models and provide recommendations for potential adaptations for broader application. The third phase focused on lithology distribution using well log and 3D seismic attributes. Data preprocessing proved to be crucial, as it was found that standardized input data produced the most geologically consistent results. Normalized and raw data led to less reliable predictions. Application to a 2D seismic profile confirmed the methodology's suitability for regional-scale studies. Beyond lithology prediction and time-to-depth conversion, the developed models support subsurface characterization for geoenergy applications, including hydrocarbon exploration, geothermal assessment, and CO_2 storage. By capturing complex geological relationships, ANN-supported workflows offer improved prediction accuracy and form a replicable framework adaptable to other basins. This research establishes a foundation for advanced modelling approaches in future resource management.

PROŠIRENI SAŽETAK

Ova doktorska disertacija predstavlja razvoj i vrednovanje metodologije osnovane na umjetnoj inteligenciji (*Artificial Intelligence*, AI) koja koristi umjetne neuronske mreže (*Artificial Neural Networks*, ANN) za poboljšanje geološke karakterizacije podzemlja panonskog stratigrafskog intervala u Dravskom bazenu. Primarni cilj istraživanja bio je unaprijediti preciznost i pouzdanost geoloških modela korištenjem strojnog učenja u područjima s visokim stupnjem istraženosti koja sadrže brojne, ali često nepotpune geološke informacije.

Početni skup podataka obuhvaćao je četiri bušotine i 3D seizmički volumen od 36,25 km^3. Cilj ovog prostorno ograničenog skupa podataka bio je obuhvatiti velik broj karotažnih dijagrama, kao i seizmički volumen visoke rezolucije kako bi se osiguralo da utjecaj kvalitete podataka na sposobnost predviđanja bude znatno smanjen. Unutar naslaga panonske starosti karakterizirane su tri litologije, pješčenjaci, lapori i ugljen na podacima iz četiri bušotine. Podinska i krovinska granica panonskih naslaga određene su na temelju bušotinskih podataka te su iste interpretirane unutar seizmičkog volumena. Navedene stratigrafske granice korištene su za definiranje granica modela. Za unapređenje ulaznih podataka za ANN analizu izračunato je sljedećih 16 seizmičkih atributa: *Original Amplitude*, *Reflection Intensity*, *Root Mean Square (RMS) Amplitude*, *Sweetness*, *Amplitude Contrast*, *Chaos*, *Edge Evidence*, *Iso-frequency Component*, *RMS Iterative*, *RMS Time Gain*, *t*Attenuation*, *Variance (Edge Method)*, *Instantaneous Frequency*, *Instantaneous Phase*, *Cosine of Phase* i *Relative Acoustic Impedance (RAI)*.

Litološki podaci su preneseni na rezoluciju modela procesom *upscale*, a za ANN analizu izdvojeno je 800 ulaznih vrijednosti. Prosječna visina ćelija modela iznosila je 25 m, u skladu s vertikalnom rezolucijom seizmičkih podataka prikupljenih 1990-ih. ANN analiza načinjena je s ukupno 15.480 podataka, uključujući vrijednosti seizmičkih atributa. Fiksni parametri mreže sastojali su se od 16 neurona (16 seizmičkih atributa) u ulaznom sloju i tri neurona u izlaznom sloju (odgovarajući trima litološkim kategorijama). Korištene su dvije stohastičke metode DAANN (*diferent architecture artificial neural network*) i SAANN (*same architecture artificial neural network*) kako bi se procijenio utjecaj arhitekture neuronske mreže i distribucije vrijednosti na uspješnost predikcije. DAANN pristup koristio je različite arhitekture neuronskih mreža s različitim brojem neurona u skrivenom sloju i aktivacijskim funkcijama, dok je SAANN koristio istu arhitekturu, ali s varijacijama u razdiobi ulaznog seta podataka u skupu podataka za uvježbavanje, testiranje i validaciju ANN-a. Za predviđanje

litološkog sastava korišteno je ukupno 2.000 ANN mreža (1.000 DAANN i 1.000 SAANN), a odabrano je 200 najuspješnijih za daljnu predikciju (100 po svakom pristupu). ANN analiza proizvela je veliki broj podataka za variogramsku analizu, pri čemu se broj podataka smanjivao kako je rasla vjerojatnost podudaranja (P50, P75, P90). Variogramska analiza pokazala je da prostorna korelacija litoloških kategorija značajno varira ovisno o tipu podataka i uspješnosti predviđanja. Zbog dominacije pješčenjaka u P75 i P90 modelima, rezultati zahtijevaju opreznu interpretaciju jer veća preciznost u ovom slučaju nije značila i nužnu geološku vjerodostojnost.

Prva faza istraživanja istaknula je izazove u razvoju metodologije za karakterizaciju distribucije litološkog sastava u podzemlju pomoću ANN-a, čak i kada su primijenjeni sveobuhvatni skupovi podataka visoke rezolucije. Glavni izazovi proizlaze iz neadekvatne obrade podataka (*upscale*-a) te prilikom koreliranja podataka u dubinskoj domeni (bušotinski podaci) s onima iz vremenske domene (seizmički podaci). Mnoge bušotine nemaju podatke o vremensko-dubinskom odnosu (*Time-to-Depth Relationship*, TDR), koji se obično određuje na temelju vertikalnog seizmičkog profiliranja (VSP) ili zvučnim karotažama.

Kako bi se svladale opisane poteškoće, druga faza istraživanja fokusirala se na razvoj nove metodologije za precizniju odredbu vremensko-dubinskog odnosa podataka. Ova metodologija posebno je značajna za bušotine bez VSP-a ili zvučne katorože, gdje su prethodne metode određivanja odnosa vrijeme-dubina (TDR) uglavnom uključivale interpolaciju između bušotina s i bez uspostavljenih TDR-a. Novi pristup uzima u obzir i bušotinske i seizmičke podatke, koristeći ANN za povećanje preciznosti i smanjenje pogrešaka uzrokovanih ekstrapolacijom funkcije brzina iz susjednih bušotina. Ovaj pristup pruža ekonomično i učinkovito rješenje za istraživanje podzemlja, nadmašujući tradicionalne metode.

Predložena metodologija koristi ANN za predviđanje odnosa vrijeme-dubina na temelju interpretacije karotažnih dijagrama i stratigrafskih intervala. Na taj način rješava probleme povezane s definiranjem odnosa vremensko-dubinskih domena i značajno poboljšava točnost predviđanja, čak i za bušotine s ograničenim geofizičkim podacima. Nadalje, integracija ANN-a s litološkim podacima iz osnovnih karotažnih dijagrama – uključujući one iz starijih bušotina – omogućila je smanjenje pogrešaka, odnosno povećanje točnosti u uspostavljanju vremensko-dubinskih odnosa, što rezultira preciznijim modeliranjem podzemlja.

Ključan čimbenik u uspješnosti primjene ANN-a bio je odabir reprezentativnog skupa podataka za uvježbavanje. Odabrano je 18 bušotina, od kojih je 14 korišteno za uvježbavanje modela, dok su četiri bušotine korištene za testiranje uspješnosti izgrađenih neuronskih mreža. Testne bušotine odabrane su kako bi predstavljale različite scenarije taloženja litostratigrafskih jedinica (različite debljine), odnosno slučajeve s prisutnošću svih četiriju interpretiranih stratigrafskih intervala te slučajeve gdje su bušotine zahvatile samo prva dva intervala. Prilikom odabira uzet je u obzir i njihov geografski položaj u donosu na bušotine za uvježbavanje kako bi se omogućila procjena utjecaja udaljenosti na uspješnost predviđanja. Interpretacija bušotinskih podataka bila je usmjerena na izdvajanje propusnih i nepropusnih jedinica korištenjem osnovnih karotažnih dijagrama. Njihova interpretacija omogućila je definiranje vertikalne distribucije litološkog sastava, kao i odredbu triju regionalnih markera α (granica pliocen-miocen), Rs_7 (granica srednji – kasni miocen) i PNg (podloga neogena). Na temelju ovih glavnih granica, za svaku bušotinu interpretirana su do četiri stratigrafska intervala, grupirajući naslage prema starosti, litološkom sastavu i okolišu taloženja. Ovi intervali definiraju četiri stratigrafske jedinice koje se međusobno razlikuju po petrofizičkim svojstvima, pri čemu ta svojstva variraju i s dubinom zalijeganja. Prvi interval čine slabo vezane naslage pliocensko-kvartarne starosti, drugi uključuje gornjomiocenske pješčenjake i lapore, treći obuhvaća donjomiocenske i srednjemiocenske heterogene naslage, dok četvrti interval uključuje starije predneogenske stijene koje čine podlogu bazenske ispune. Najizraženiji kontrast seizmičke brzine očekivan je na prijelazu neogena u podlogu bazena.

ANN analiza provedena je pomoću softvera TIBCO Statistica, koristeći analizu *Time series (regression)* kako bi se osiguralo predviđanje u skladu s geološkim i geofizičkim načelima. Umjetne neuronske mreže višeslojno su konfigurirane (*Multi-layer percepton*, MLP) kao arhitekture s minimalno 3 i maksimalno 17 neurona u skrivenom sloju. Kako bi se izbjeglo pre-uvježbavanje modela, korištena je tehnika *weight decay*, čime se potaknula jednostavnija i generaliziranija mreža. Za treniranje modela korišteno je više od 27.000 slučajeva iz 14 bušotina, a 10 najuspješnijih neuronskih mreža združeno je u jedan model za analizu. ANN model postigao je korelacijski koeficijent veći od 0,99 za skupove podataka korištene za uvježbavanje, testiranje i validaciju, s prosječnom apsolutnom pogreškom (*mean absolute error*, MAE) od oko 25 ms i korijenom srednje kvadratne pogreške (*root mean square error*, RMSE) od približno 34 ms. Uspješnost skupa mreža testirana je na četiri bušotine koje nisu bile uključene u predviđanje, pokazujući visoku uspješnost predviđanja. Vrednovanje točnosti predviđanja otkrila je da ANN pristup nadmašuje tradicionalne metode ekstrapolacije dvostrukog vremena putovanja vala (*two way travel time*, TWT-a) u više od 75% slučajeva, pri čemu je pet bušotina pokazalo superiorne rezultate u odnosu na ekstrapolaciju, postižući u

100% slučajeva najuspješnije rezultate. Testne bušotine također su pokazale visoku točnost, čak i u usporedbi s predviđanjima temeljenima na ekstrapolaciji iz susjednih bušotina. Analiza je pokazala da blizina bušotina ne jamči nužno uspješnost ekstrapolacije, naglašavajući složen odnos između prostorne blizine i varijabilnosti litološkog sastava. Rezultati sugeriraju da geološki odnosi i uvjeti u podzemlju imaju značajniju ulogu u preciznosti predviđanja nego sama udaljenost između bušotina.

Iako je dokazano da umjetne neuronske mreže (ANN) mogu biti uspješno primijenjene za rješavanje poteškoća prilikom određivanja odnosa vrijeme-dubina, odnosno za predviđanje dvostrukog vremena putovanja vala (TWT) iz stratigrafskih i petrofizikalnih parametara u slučajevima kada konvencionalni podaci nisu dostupni, određeni izazovi i dalje su prisutni. Među njima se posebno ističe problem neadekvatnog skaliranja (*upscale*) bušotinskih zapisa. U trećoj fazi istraživanja primijenjeni su inovativni procesi pripreme podataka s ciljem poboljšanja performansi ANN modela te predviđanja raspodjele litologije korištenjem seizmičkih atributa. Stoga je razvijen ANN-model temeljen na analizi 3D seizmike koja pokriva površinu od 4365 km², s fokusom na panonske naslage.

Ključna inovacija ove metodologije bila je uvođenje volumena šejla (V_{sh}) kao kontinuirane varijable, umjesto dosadašnjeg pristupa gdje se litološki sastav tretirao kao diskretna kategorijska varijabla. V_{sh} izračunat je na temelju karotaže spontanog potencijala (SP). Kako bi se optimizirali procesi uvježbavanja i predviđanja ANN modela te smanjila subjektivnost interpretacije, razvijena su četiri modela podzemlja s različitim brojem slojeva (20, 50, 100 i 200). Model s 200 slojeva, prosječne visine ćelije od 6,5 metara, pokazao se najuspješnijim jer je omogućio smanjenje pogrešaka precjenjivanja dominantnog litološkog sastava uz očuvanje geološke rezolucije i statističke vjerodostojnosti. ANN modeli uvježbani na podacima sa 100 ili manje slojeva pokazali su slabije performanse ili potpuni neuspjeh, što naglašava osjetljivost ANN modela na rezoluciju ulaznih podataka. Za treniranje ANN modela kreirano je 747.800 ćelija s pripadajućim seizmičkim atributima. Kao ključni ulazni parametri odabrano je 12 seizmičkih atributa: *Sweetness*, *3D Curvature*, *Variance*, *Original Amplitude*, *Instantaneous Frequency*, *Envelope*, *Instantaneous Phase*, *Generalized Spectral Decomposition*, *Apparent Polarity*, *Reflection Intensity*, *RMS Amplitude* i *Relative Acoustic Impedance.*

Standardizirani ulazni podaci rezultirali su najboljim predviđanjima, dok su sirovi (*raw*) podaci pokazali najslabije performanse. Predviđene V_{sh} vrijednosti podijeljene su u tri litološke klase (pješčenjak, pješčenjak-lapor i lapor) te su ugrađene u litološki model. Model izveden iz standardiziranih podataka sadržavao je 52,17% lapora, 29,89% pješčenjaka-lapora i 17,94% pješčenjaka, dok su modeli temeljeni na normaliziranim i sirovim podacima pokazali značajna odstupanja u raspodjeli litološkog sastava.

Za testiranje hipoteze o primjenjivosti razvijene metodologije na 2D seizmičkim podacima odabran je jedan seizmički profil. Odabir profila temeljen je na mogućnosti procjene uspješnosti ANN-a prilikom predviđanja te da djelomično pokriva 3D seizmički volumen i jednu bušotinu. Prvi korak bila je interpretacija gornje i donje granice panona (α i Rs_7). Zatim su izračunati seizmički atributi. Izdvojeno je 12 seizmičkih atributa koji su prethodno prepoznati kao visoko učinkoviti za predviđanje litološkog sastava. Podaci su obrađeni u tri skupa podataka: sirovi, normalizirani i standardizirani. Svaki skup podataka kasnije je korišten kao ulazni podatak za ANN predviđanje. Razvijene neuronske mreže korištene su za predviđanje volumena šejla (V_{sh}), koji je zatim podijeljen u tri litološke klase: pješčenjak (≤0,5), pješčenjak-lapor (0,5-0,7) i lapor (≥0,7).

Konačni rezultati ponovo su pokazali ključnu važnost obrade podataka za predviđanje raspodjele litološkog sastava u podzemlju. Sirovi ulazni podaci doveli su do značajnog precjenjivanja pješčenjaka, dosegnuvši maksimalnih 100%, čineći rezultate potpuno neupotrebljivima za bilo kakvu interpretaciju. S druge strane, normalizirani podaci značajno su podcijenili kategoriju pješčenjak-lapor (1,67%), međutim, predviđena raspodjela pješčenjaka (43,36%) i lapora (64,97%) odgovarala je geološki očekivanim rezultatima. Standardizirani podaci, iako su malo precijenili lapor, proizveli su geološki najsmislenije rezultate, sa sljedećim raspodjelama pješčenjaka (18,92%), pješčenjaka-lapora (19,91%) i lapora (61,17%). Rezulati potvrđuju standardizaciju, kao alat obrade podataka, koji značajno poboljšava geološku vjerodostojnost dobivenu na temelju predviđanja raspodjele litološkog sastava podzemlja. Uz pomoć uspješnog predviđanja litološkog sastava i odnosa vrijeme-dubina, razvijeni ANN modeli značajno doprinose karakterizaciji podzemlja u geoenergetskim primjenama, uključujući istraživanje ugljikovodika, geotermalne energije i skladištenje ugljikovog dioksida.

KEYWORDS

Pannonian Basin

Lithology

Well logs

Volume of shale

Seismic 3D and 2D data

Seismic attributes

Time-to-depth relationship

Artificial neural networks

Lithology distribution

Geoenergy exploration

KLJUČNE RIJEČI

Panonski bazen

Litologija

Karotažni podaci

Volumen šejla

3D i 2D seizmički podaci

Seizmički atributi

Odnos vrijeme-dubina

Umjetne neuronske mreže

Distribucija litologije

Geoenergetska istraživanja

TABLE OF CONTENTS

TABLE OF FIGURES

1. INTRODUCTION

Structural-stratigraphic relationships are essential for qualitative subsurface characterization, which hold scientific and economic significance. Lithology distribution plays a critical role in estimating the location and volume of geoenergy resources. Traditionally, subsurface lithology is inferred from well data using deterministic mapping (**Forgotson, 1960; Bianchi et al., 2015**) or stochastic algorithms (**Dubrule, 1993; Feng et al., 2018**). However, these methods carry substantial uncertainty, especially in regional studies where well data is sparse. Even in developed basins with dense well coverage, interpretations still depend heavily on the interpreter's expertise (**Hohn, 1999**). Subsurface exploration typically focuses on age, structure, and lithology, which directly influence interpretation and have substantial economic implications when applied to resource estimation (**Selley & Sonnenberg, 2015**). Unlike surface outcrops, subsurface characterization often lacks directly measured data (well cores) and relies on indirect methods as well logs. In many cases, conventional techniques (e.g. extrapolation) and mathematical algorithms are used to infer lithology between wells, but their accuracy is limited – particularly where acoustic logs or vertical seismic profiling are lacking (**Hart et al., 2000; Cao et al., 2017; Sun et al., 2023**). Given these challenges, the aim of this dissertation is to develop a methodology that reduces uncertainty in subsurface lithology distribution by integrating more reliable predictive techniques.

For this purpose, the Drava Basin in North Croatia was chosen as research area since it exemplifies this challenge. As a prominent hydrocarbon exploration area within the Pannonian Basin System, it has undergone extensive exploration for over 70 years, yielding considerable geological and geophysical data (**Velić et al., 2012a, 2012b**). Nonetheless, many older wells lack acoustic logs or vertical seismic profiling, complicating accurate time-to-depth conversion. Conventional solutions often involve extrapolation using velocity functions derived from neighboring wells (**Aker et al., 2020**), but this approach risks inaccuracies that could compromise subsequent analyses and interpretations of structures and lithology. This particularly poses a significant challenge when assessing the geoenergy potential of an area, as accurate depth information is crucial for reliable evaluations. Another major challenge is the inherent subjectivity of geological interpretations, mainly due to insufficient data that makes building reliable models with traditional methods nearly impossible, or at least only good enough for preliminary research. To tackle this problem, machine learning techniques have emerged and proven to be helpful tools, making geological modeling more accurate and thorough (**Smirnoff et al., 2008; Zhou et al., 2019; Feng et al., 2024; Zhou & Liu, 2024**). In particular, artificial neural networks (ANNs) have demonstrated the capacity to enhance predictive accuracy by analyzing complex, multidimensional datasets that are too complex to be processed using traditional methods. These findings encourage this research to make an assumption that by training prediction on well log and seismic data, ANNs can effectively model lithological distribution and time-to-depth relationships, even in areas with limited direct measurements, i.e. in Drava Basin.

This transition toward AI-supported methodologies in geoscience field aligns with a broader trend of digitalization and automation (**Karpatne et al., 2019; Sagi et al., 2020; Irrgang et al., 2021; Sun et al., 2022; Zhao et al., 2024; Khosravi & Ashkpour, 2024**). The integration of ANN models not only reduces subjectivity of geological and geophysical models, but also provides data-supported, probabilistic assessments that can improve geological interpretation and resource management. Such advancements are essential for enhancing predictive accuracy in hydrocarbon exploration, geothermal energy projects, and gas storages – domains where precise subsurface characterization is vital. The motivation for adopting ANNs in geological modeling stems from their ability to recognize complex patterns within datasets that conventional methods often overlook. For instance, in the Drava Basin, where incomplete datasets cause difficulties in obtaining accurate interpretations, ANNs are expected to demonstrate superior performance in predicting lithological variations.

Despite their potential, ANN-supported methods remain underutilized in some geoscience applications, even within mature basins. This research aimed to bridge this gap by applying ANN models to subsurface characterization and evaluating their effectiveness compared to traditional approaches. This dissertation presents the development of an AI-supported methodology designed to assist geoscientists in enhancing the accuracy and reliability of their geological models, with a particular focus on litology distribution and time-to-depth relationship extrapolation. While the framework is broadly applicable, its successful implementation requires adaptation to the specific geological parameters of each study area.

2. PREVIOUS INVESTIGATIONS OF THE RESEARCH AREA

The study area is located in the eastern part of the Drava Basin in northern Croatia (Figure 1). It forms a significant portion of the southwestern marginal part of the Pannonian Basin System (PBS). Within the Croatian part of the PBS, the Drava Basin contains the thickest Neogene-Quaternary sedimentary succession. The pre-Neogene basement is composed of Paleozoic and Mesozoic crystalline and sedimentary rocks. In the Drava Basin, the top of this basement reaches depths of up to approximately seven kilometers (**Saftić et al., 2003; Velić, 2007; Cvetković et al., 2019**).

The study area lies within the eastern part of the Drava Basin, which is part of the North Croatian Basin (NCB). Like the broader PBS, the NCB is characterized by syn-rift and post-rift tectonic phases (**Royden, 1988; Tari et al., 1992; Horváth et al., 2006, 2015**). During the early rifting phase, spanning from Early to Middle Miocene, deposits consisted predominantly of diverse clastics deposits formed in alluvial and lacustrine environments, with sporadically volcanic activity contributing pyroclastic material (**Saftić et al., 2003; Brlek et al., 2024**). The end of the syn-rift phase marked the onset of the post-rift thermal subsidence started in the late Badenian (**Lučić et al., 2001; Balazs et al., 2018; Rukavina et al., 2023**). This transition is accompanied by a regional transgression (**Balazs et al., 2016; Pavelić & Kovačić, 2018**). Transgression phase is followed by the deposition of clastics, limestones, and eventually deep-water marls. A subsequent regression phase cause a transition from turbidites to deltaic environments, driven by high sediment influx (**Pavelić & Kovačić, 2018**).

In the study area, the Neogene-Quaternary infill of the North Croatian Basin overlies a complex pre-Neogene basement composed of Paleozoic crystalline rocks and their Mesozoic sedimentary cover. These lithological units are part of the European continental Tisza block and consists primarily of Paleozoic magmatic and metamorphic rocks, as well as Paleozoic and Mesozoic sediments (**Pamić, 1984; Pamić, 1998; Pamić & Jurković, 2002**).

This dissertation primarily focuses on Pannonian-age clastic deposits from the regressive stage, and specifically adopts the Pannonian interval for analyzing Late Miocene sedimentary sequences (**Sebe et al., 2020**). This interval began at the end of the Sarmatian when the isolation from the Central Paratethys started, which led to freshening of the environment. During this stage, clay-rich marls were deposited in a large Pannonian lake. This was followed by the deposition of a thick turbidite sequence composed of marls, sandstones, and siltstones, driven by the development of regional fluvial systems. This systems transported sediment from the Alpine region, as confirmed by **Matošević et al. (2024)**. Subsequent basin shallowing and eventual infill were marked by a succession of depositional environments, transitioning from turbidites of a morphological shelf (alternating marls and sandstones), to deltaic systems (alternating coarse sandstones, marls, and coal) (**Saftić et al., 2003, Magyar et al., 2013; Sztanó et al., 2015; Matošević et al., 2024**).

Pliocene, Pleistocene and Holocene sedimentation was influenced by neotectonic compression and reverse faulting (**Pavelić & Kovačić, 2018**), and characterized by coarse clastics, clay, loess, and aeolian sands associated with glacial periods (**Wacha et al., 2013**).

As one of the primary depositional areas within the Pannonian Basin System (PBS), the Drava Basin has been a focal point of petroleum-geological exploration for over 70 years. During this time, extensive seismic surveys, including both 2D and 3D, have been conducted, along with the drilling of a substantial number of wells. These efforts have confirmed the presence of an active petroleum-geological system, resulting in the discovery of several relatively large oil and gas fields (**Hernitz, 1980; Velić, 2010**). Furthermore, the region shows promising potential for CO_2 geological storage within depleted or nearly depleted reservoirs and regional saline aquifers (**Vulin et al., 2023; Saftić et al., 2024; Rukavina, 2021**).

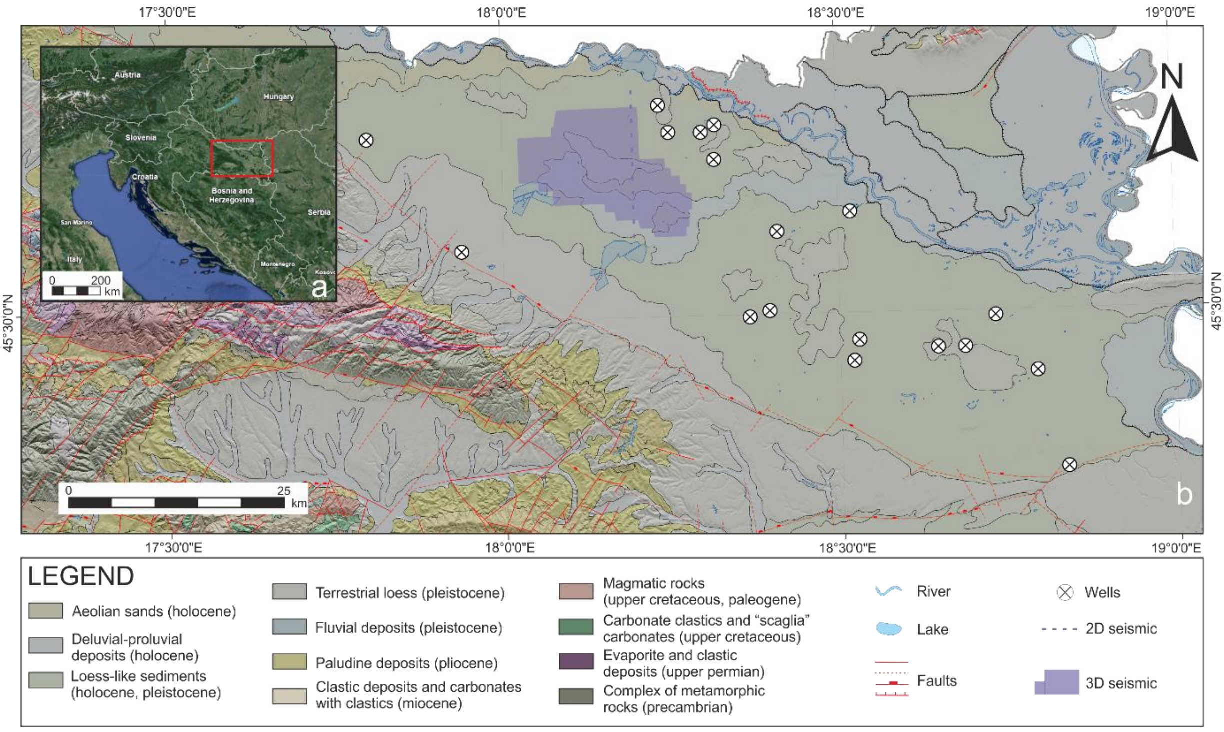


Figure 1. (a) Geographical position of the study area (modified map from Google Earth); (b) Map of the research area indicating the locations of data used in this doctoral thesis, including 3D seismic data, 2D seismic data, and wells (modified after Croatian Geological Survey, 2009).

Artificial intelligence, particularly artificial neural networks (ANNs) have so far seen limited application for lithology prediction within Croatia, yet it holds significant potential for enhancing geological modeling of the North Croatian Basin (NCB) deposits, primarily due to the availability of extensive high-resolution well log data (**Malvić & Cvetković, 2013**). These data provide a robust foundation for training and validating ANN models, enabling more precise subsurface interpretations. Early regional studies, such as **Brcković et al. (2017)**, demonstrated the benefits of integrating seismic and well data into ANN frameworks, showing notable improvements in lithological distribution modeling.

Globally, ANNs have been applied to a range of subsurface characterization purposes, including porosity and permeability prediction (**Iturrarán-Viveros & Parra, 2014**), seismic reservoir characterization (**Othman et al., 2021**), and shale volume estimation (**Taheri et al., 2021; Mohammadinia et al., 2023**). Among earth sciences, hydrology represents the field with the most widespread and long-standing applications of artificial intelligence, as evidenced by numerous studies focused on runoff prediction, groundwater modeling, and water quality forecasting (**Bonafe et al., 1994; Johnson & Rogers, 1995; Maier & Dandy, 1996; Muttiah et al., 1997; Valizadeh et al., 2017; Hatampour et al., 2018**).

3. OBJECTIVES AND HYPOTHESES

The general hypothesis is that ANN-supported approaches will outperform traditional extrapolation methods in terms of accuracy and reliability, providing a more robust framework for geological modeling. Moreover, it is assumed that machine learning techniques can reveal previously undetected correlations within geological datasets, enabling more precise subsurface predictions.

Objectives of this dissertation are:

1. Reinterpret the lithological composition within the study area using available well data
2. Perform neural network analysis on 3D seismic and well data
3. Make a geological model with distribution of the Pannonian sediments in the subsurface of Donji Miholjac area
4. Investigate the applicability of the method on 2D seismic data.

These objectives were based on the following hypotheses:

1. It is expected that the use of artificial neural networks will advance geological characterization of the subsurface
2. It is assumed that the incorporation of 3D seismic and well data will greatly improve accuracy by removing the bias in the interpretation
3. It is speculated that research results will also be applicable on limited 2D seismic data, which are far more common in regional surveys.

Thus, this doctoral thesis aims to enhance geological modeling by implementing artificial neural networks, demonstrating their potential to improve traditional approaches. By bridging the gap between conventional methods and AI-supported modeling, this research contributes to the advancement of machine learning applications in geosciences.

4. METHODOLOGICAL APPROACH

The methodological approach was structured into three distinct stages, each representing a higher level of development within the newly proposed geological modeling framework using machine learning. This chapter outlines all methodological stages encompassed within the three research phases, described in the chronological order proposed for their implementation by this dissertation.

The study focused on a subsurface volume of the Pannonian-age clastic interval, extracted from seismic data and well logs. The geological and geophysical data analysis followed a comprehensive workflow, integrating well log and seismic interpretation, seismic attribute extraction, artificial neural network modeling, and spatial distribution analysis. This dissertation targeted Pannonian clastic interval and utilized data from 4, 18, and 11 wells – corresponding to the three stages of methodological approach development. The wells were drilled by the INA company during their oil and gas exploration campaigns between the 1980s and 2010s. Permission to use the necessary data for the planned research was granted from the Croatian Hydrocarbon Agency. The selected wells, ranging in depth from 1,300 m to 4,500 m, were chosen based on the availability of acoustic well logs, allowing for precise control over time-to-depth conversion.

It was necessary to convert well data to the time domain since seismic and well log data belong to different domains. Check-shot data available from well reports were utilized for this purpose. The data preparation process (Figure 2) involved calibrating acoustic logs using check-shot data to correct for acoustic log drift, which may arise from equipment calibration issues, well conditions, or environmental factors (**Mari et al., 2020**). To enhance the quality of the acoustic logs, despiking was performed to eliminate outliers and anomalies (**Rider, 2002**). The despiked acoustic logs were calibrated based on vertical seismic profiling (VSP) data which enabled transformation from measured depth to two-way travel time (TWT) for validation purposes, employing time-to-depth relationships (TDRs) derived from smooth check-shot times.

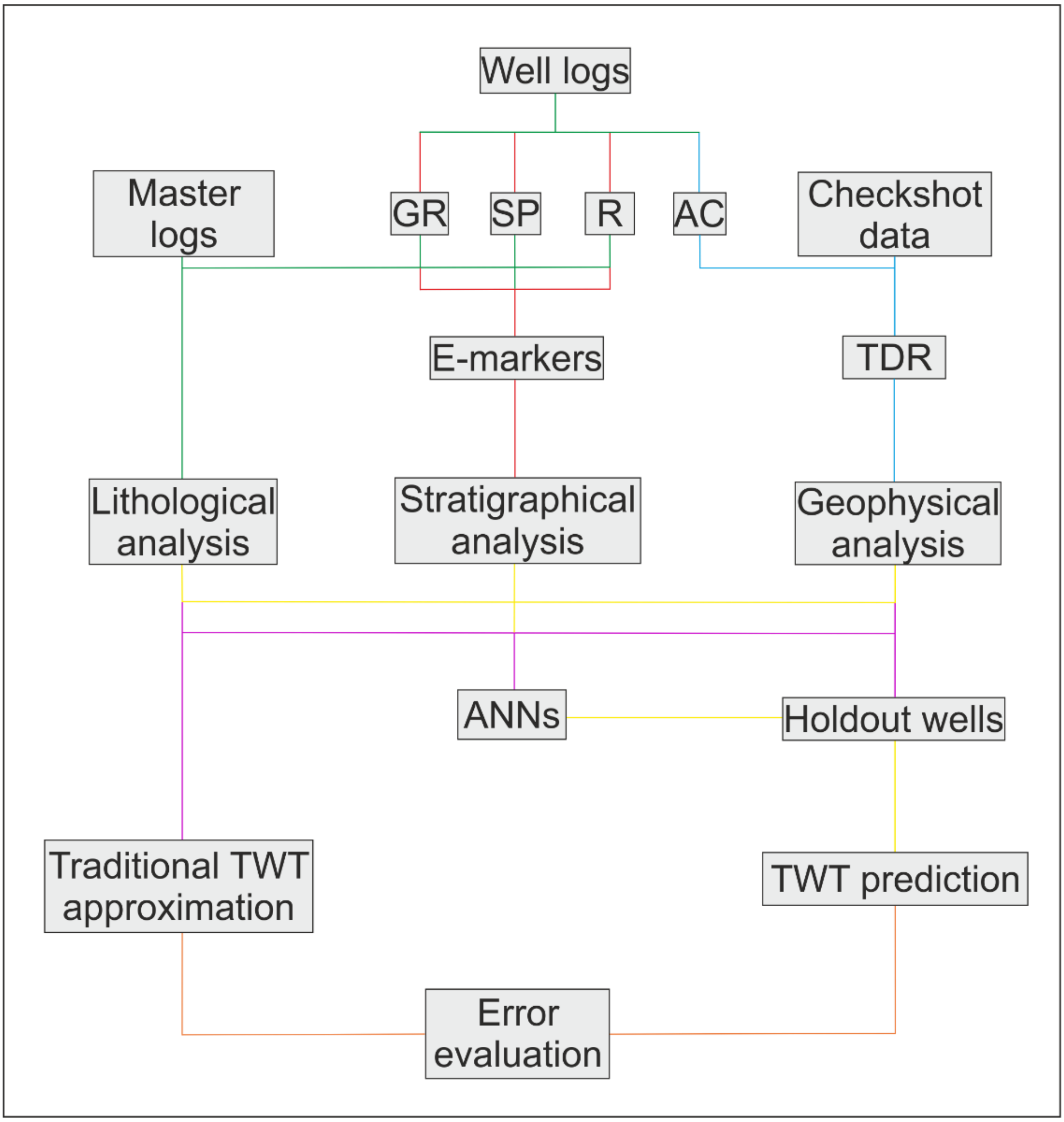


Figure 2. Workflow chart of the second phase of methodology development. TDR stands for time-to-depth relationship (after Kamenski et al., 2024).

This conversion facilitated the delineation of Pannonian deposits boundaries, corresponding to the model boundaries. These boundaries, identified by well tops (E-log horizons), were interpreted on seismic sections to accurately map the bottom and top boundaries throughout the entire seismic volume.

Interpretation of well logs was performed in Interactive Petrophysics software. Before interpretation, all well logs underwent pre-processing, which includes curve rescaling (this step was applied only to SP logs to correct curve inversion occurring in the shallowest sections, where the presence of freshwater in the pores caused the readings to be reversed), interactive curve splicing, and baseline shifting.

Availability of the legacy data played a crucial role in delineating model boundaries and mapping the top and base of the Pannonian interval across the seismic volume. The initial lithology interpretation utilized master logs and available well logs, including spontaneous potential (SP), shallow and deep resistivity (R16, R64), and gamma ray (GR) measurements. The interpretation of resistivity and SP curves, supplemented with additional well data, served as the base for lithology interpretation and unit correlation throughout the study area (**Bassiouni, 2013; Rider, 2002**). These well log data were used for differentiating permeable and impermeable units, establishing boundaries, and estimating the volume of shale. Lithological units were classified as categorical variables representing distinct rock types, following standard interpretation practices.

During the initial phase of methodological approach development, the lithological classification was simplified to focus on sandstone and marl as the prevailing lithological categories within the Pannonian interval. Additionally, coal was incorporated as a third category due to its significantly different acoustic properties; omitting coal would have introduced errors during artificial neural network (ANN) analysis (**Walton et al., 2000**). As the methodological approach evolved, the analysis of lithological variability was further refined with delineation of permeable and impermeable units in meter resolution – as an essential aspect of reservoir characterization.

To obtain continuous input data, the volume of shale (V_{sh}) was introduced. The V_{sh} analysis was performed using the SP log, as it effectively estimates shale volume by interpreting the SP deflection between the static SP value in a quartz-rich sandstone and the shale baseline, representing 100% shale content (**Serra, 1984; Asquith & Krygowski, 2004; Rider, 2002**).

Lithology distribution estimates were upscaled within the geological model to provide initial data points for ANN analysis. While thinner model layers are generally more successful in heterogeneous environments since they contain more input data for ANN analysis, this research in its initial phase, employed coarser layering to preserve seismic data quality, as the building of the ANN predominantly relies on seismic rather than well data. Lithological categories were treated as categorical variables, each identified by a distinct value to ensure consistency and accuracy in the modeling process.

To establish connections between well data (lithology) and the broader study area, a large number of seismic attributes were calculated from the seismic volume. These attributes capture information that is not readily apparent through conventional seismic visualization methods. Seismic attributes serve as either physical or geometric indicators (**Taner et al., 1994**). Physical attributes provide insights into lithology and rock properties, while geometric attributes assist in stratigraphic and structural interpretation. In this study, physical attributes were prioritized, as they are more relevant for assessing lithological spatial distribution. The primary purpose of generating seismic attributes in this research was to increase the input variables for the neural network analysis. Several seismic attributes have enhanced the identification of model boundaries, especially the deeper regional E-log horizon (Rs_7), which is difficult to map on classic seismic section.

Seismic attributes are derived by representing the seismic signal as a complex function with both real and imaginary components. The real component corresponds to the recorded seismic trace (kinetic energy), while the imaginary component (potential energy) is calculated using the Hilbert transformation (**Taner et al., 1976; Khan & Akhter, 2015**). This transformation produces basic instantaneous attributes that form the foundation for calculating additional attributes through mathematical operations. The complex seismic signal is obtained by treating the recorded seismic trace as the real part and initially setting the imaginary part to zero. Through Fourier transformation, the signal is transformed into the frequency domain, where positive transformation components are doubled, and negative components are disregarded. The inverse Fourier transformation then reconstructs a function with the real part preserved and the imaginary part equal to the Hilbert transformation of the input signal. This process aims to achieve a polar representation of the seismic trace, effectively separating amplitude from phase and frequency information while preserving the original spectrum (**Taner, 2001**).

To enhance lithology distribution predictions using ANNs, seismic attribute extraction focused on capturing key lithological and morphological features. Attributes such as Sweetness, Root-Mean-Square (RMS) Amplitude, Variance, Reflection Intensity, Apparent Polarity, Instantaneous Frequency, 3D Curvature, and Generalized Spectral Decomposition were selected for their potential to highlight lithological and morphological characteristics

(**Chopra & Marfurt, 2006; Liu & Marfurt, 2006; Chopra & Marfurt, 2008; Brcković et al., 2017; Ker et al., 2014; Li et al., 2019; Oumarou et al., 2021**). These attributes help reveal complex relationships within seismic data, significantly aiding structural, stratigraphic, and petrophysical interpretation (**Taner et al., 1976; Taner, 2001; Djeddi, 2016**).

The normalization and standardization of all data, including seismic attributes and V_{sh} values, were performed to ensure consistent scaling. This process employed standard equations for normalization and standardization, guaranteeing uniform data representation and compatibility throughout the analysis.

The obtained attribute and lithology data were subjected to artificial neural network (ANN) analysis, designed to solve specific tasks. Artificial neural networks, first conceptualized in the 1940s (**McCulloch & Pitts, 1943**) and further refined over time (**Anderson & Rosenfeld, 1988; Rosenblatt, 1958**), aim to replicate human cognitive processes. The fundamental structure includes an input layer, a hidden layer, and an output layer, where each input variable is represented by a neuron. The input layer represents input variables, a hidden layer consists of an arbitrary number of neurons, and an output layer contains the number of neurons that matches the predicted variables or categories (**Agatonovic-Kustrin & Beresford, 2000; Lean et al., 2007**). During the training process, the network iteratively adjusts weight factors to minimize prediction errors. This structure enables the ANN model to capture complex relationships and patterns within the data. One of the primary challenges in ANN training is preventing overfitting – when the model achieves near-perfect accuracy on training data but fails to generalize to new cases (**James et al., 2013**). To address this issue, the data is divided into training, testing, and validation subsets, enabling continuous error monitoring and adjustment throughout the process.

Initial ANN training employed a quasi-stochastic approach to enhance prediction accuracy and minimize uncertainties. This stochastic approach was applied in two distinct ways (Figure 3). The first was a quasi-stochastic method, involving a wide range of network architectures with varying numbers of neurons in the hidden layer and different activation functions (DAANN). The second approach utilized a consistent network architecture but varied the distribution of cases among the three datasets (training, testing, and validation) and the initial starting point (case) for the analysis (SAANN). The methodology is conceptually similar to Gaussian simulation with seed values (**Nowak & Verly, 2005; Paulsen et al., 2018**). The ANN approach calculates probability values for each prediction point within the model, allowing for the generation of multiple geological models through numerous iterations. Each iteration step integrates different combinations of input data, while different workflows apply ANN analysis to diverse data types.

The artificial neural networks were used to analyze well and seismic data, setting it apart from the traditional deterministic approach (**Agatonovic-Kustrin & Beresford, 2000; Brcković et al., 2017**), and from the stochastic components directly embedded within the neural network algorithm itself (**Kappen, 2001**). The ANN analysis was performed using Tibco Statistica (Neural Nets module), where network architecture optimization was achieved through iterative training while minimizing prediction errors.

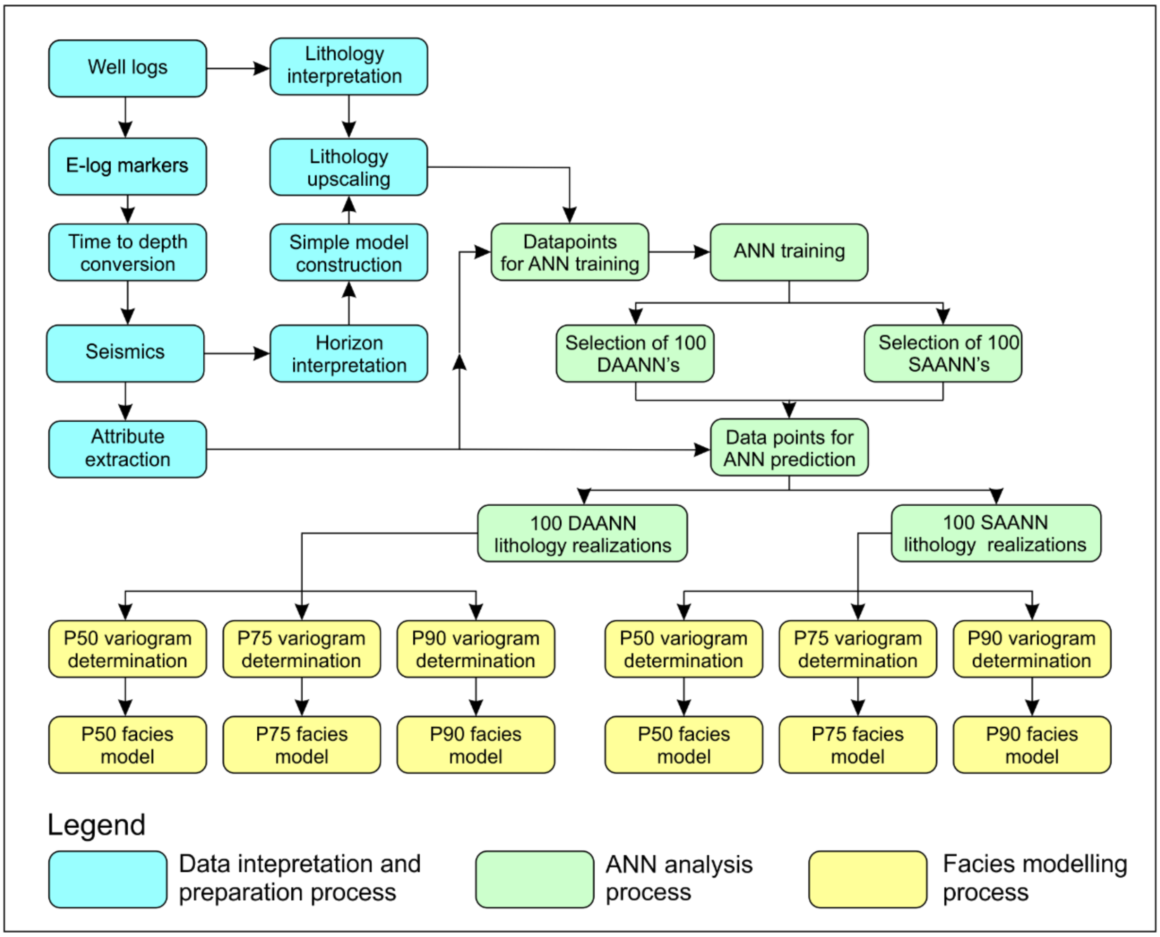


Figure 3. Workflow chart of the first phase of methodology development. DAANN stands for different architecture of ANN, while SAANN stands for same architecture ANN (after Kamenski et al., 2020).

The learning process was structured with data distributed into training, testing, and validation sets to ensure model reliability. Three possible stopping conditions were defined: achieving optimal parameters, encountering prediction failures, or detecting overtraining, indicated by increasing error values during validation (**James et al., 2013; Anderson &**

Rosenfeld, 1988). The ANN model validated predictions through multiple error metrics to ensure accuracy and consistency.

Following the ANN analysis, the results were incorporated into the geological model to map the distribution of lithological categories. The subsequent geostatistical analysis included building of variograms to determine the spatial correlations and kriging to estimate lithology at points where the lithology was not predicted by the ANN. Variogram parameters, including sill, range, and nugget effect, were carefully optimized to enhance the accuracy of spatial interpolation (**Cressie, 1990; Hohn, 1999**). Kriging as a method contains several variations, here indicator kriging was utilized for categorical data, enabling the estimation of probabilities for specific lithological classes based on spatial continuity and indicator variograms (**Journel, 1983; Nikravesh & Aminzadeh, 2003**).

Variograms were constructed to evaluate estimation probabilities for each lithological category obtained from ANN modeling, providing insights into spatial correlations and determining whether the data exhibit isotropy or anisotropy. Geostatistical analysis using variograms is essential for spatial modeling, as it correlates variance and distance to develop spatial dependence (**Nikravesh & Aminzadeh, 2003**). A typical variogram has three main parameters: sill, range, and nugget effect. Sill represents the maximum semivariance, indicating the plateau where data no longer correlate with distance. Range defines the distance at which spatial correlation becomes negligible. Nugget effect reflects variability at minimal distances or experimental errors (**Cressie, 1990**).

Various theoretical models, such as spherical and exponential, were employed to approximate experimental variograms (**Hohn, 1999**). Spatial variability often displays mixed anisotropy (**Nikravesh & Aminzadeh, 2003**), and final variogram optimization was achieved using residual concepts and the kriging method. Variogram analysis of categorical variables, such as facies and lithologies, led to smoother category representations and improved spatial continuity, with enhanced nugget effects compared to more common continuous variables (**Falivene et al., 2007; Hengl et al., 2007**).

To address spatial uncertainty, indicator kriging was selected as the primary method for estimating spatial lithology distribution. Kriging, a statistical estimation method, predicts values at unknown points by incorporating neighboring data points (**Krige, 1951; Matheron et al., 1965**). While kriging excels in localized accuracy, it inherently produces deterministic outcomes (**Nikravesh & Aminzadeh, 2003**). In contrast, indicator kriging, employed in this research, estimates the local probability distribution function without presuming a specific model (**Journel, 1983**). This non-parametric technique is particularly advantageous for reconstructing probability functions and evaluating the likelihood of specific lithological events, especially when dealing with categorical variables transformed into binary values (**Kanevski & Dumolard, 2008**). Consequently, it enables robust probability and risk mapping.

The fundamental concept of indicator kriging lies in selecting possible states for classification, dividing the data into classes of nearly equal sample numbers. Each indicator variable requires a corresponding variogram model to ensure accurate probability estimations. This approach proves especially beneficial in estimating the likelihood of class membership, forming the basis for lithology assessment in this research.

The estimation of lithological composition initially targeted localized areas with 3D seismic and well data. In advanced research phases, the analysis expanded to include larger seismic (3D and 2D) and well datasets, enabling regional subsurface mapping of both categorical and continuous variables. Improved lithology categorization, along with denser spacing of well data, is expected to deliver even more reliable results while meeting both statistical and geological standards.

Data preparation and interpretation were facilitated through a combination of Schlumberger’s Petrel software, Interactive Petrophysics (IP 2021), and Microsoft Excel, ensuring comprehensive workflow for data integration, analysis, and validation. The ANN analysis was performed using TIBCO Software Inc. Statistica 13, leveraging its Neural Nets module for optimizing neural network architecture and minimizing prediction errors. Final output – modeling was conducted using Schlumberger’s Petrel E&P software platform (academic licence), which served as the primary tool for building and visualizing the geological model.

5. ORIGINAL SCIENTIFIC PAPERS

5.1 Lithology prediction in the subsurface by artificial neural networks on well and 3D seismic data in clastic sediments: a stochastic approach to a deterministic method

By

Ana Kamenski, Marko Cvetković, Iva Kolenković Močilac & Bruno Saftić

Published in GEM – International Journal on Geomathematics

DOI: https://doi.org/10.1007/s13137-020-0145-3

5.2 From traditional extrapolation to neural networks: time-depth relationship innovations in the subsurface characterization of Drava basin, Pannonian Super Basin

By

Ana Kamenski, Marko Cvetković, Josipa Kapuralić, Iva Kolenković Močilac & Ana Brcković

Published in Advances in Geo-Energy Research

DOI: https://doi.org/10.46690/ager.2024.10.05

Advances in Geo-Energy Research

Vol. 14, No. 1, p. 25-33, 2024

Original article

From traditional extrapolation to neural networks: Time-depth relationship innovations in the subsurface characterization of Drava Basin, Pannonian Super Basin

Ana Kamenski[1], Marko Cvetković[2]*, Josipa Kapuralić[2], Iva Kolenković Močilac[2], Ana Brcković[2]

[1] *Department of Geology, Croatian Geological Survey, Zagreb 10000, Croatia*

[2] *Faculty of Mining, Geology and Petroleum Engineering, University of Zagreb, Zagreb 10000, Croatia*

Keywords:
Well logs
time-to-depth relationship
seismic interpretation
artificial neural networks
Pannonian Super Basin

Cited as:
Kamenski, A., Cvetković, M., Kapuralić, J., Kolenković Močilac, I., Brcković, A. From traditional extrapolation to neural networks: Time-depth relationship innovations in the subsurface characterization of Drava Basin, Pannonian Super Basin. Advances in Geo-Energy Research, 2024, 14(1): 25-33.
https://doi.org/10.46690/ager.2024.10.05

Abstract:
The estimation of time-to-depth relationships can prove challenging in regions with rare acoustic logs. This study focuses on the eastern part of the Drava Basin in north Croatia, chosen as a mature hydrocarbon exploration area with abundant geophysical and well data. As only a small portion of wells have well log measurements or seismic profiling performed, a time-to-depth extrapolation is often performed, which potentially results in the erroneous placement of well log markers in the time domain and affects the interpretation of seismic sections or volumes. This study proposes a novel methodology for predicting two-way travel time values in wells without vertical seismic profiling or acoustic logging. This research evaluates the parameters for the characterization of the velocity distribution in the subsurface and the efficiency of artificial neural networks versus conventional methods for this task. The constructed artificial neural network model has a correlation coefficient above 0.99 for the training, testing, and validation datasets, with a mean absolute error of approximately 25 milliseconds for each network. Artificial neural networks proved to have a lesser error in predicting the two-way time and are not sensitive to outlier values.

1. Introduction

The subsurface of the Drava Basin in North Croatia proved to be a rich hydrocarbon exploration area at a Pannonian Super Basin scale (Velić et al., 2012a, 2012b). This has led to extensive exploration activities in the last 70 years, resulting in a substantial amount of geophysical and geological data. The data collected outside the active hydrocarbon exploitation blocks is available for research purposes. However, the lack of acoustic logs or vertical seismic profiling in many older wells presents a significant challenge for accurate time-to-depth conversion, rendering this region relatively underexplored in this aspect, like many other mature basins around the world (Hart et al., 2000; Cao et al., 2017; Sun et al., 2023).

Traditionally, solving time-to-depth relationship (TDR) gaps within wells without acoustic logs or vertical seismic profiling relies on the implementation of velocity functions (hereafter referred to as extrapolation). These velocity functions contain information about the depth in the two-way travel time domain (hereafter referred to as TWT) derived from neighboring wells (Aker et al., 2020; Inichinbia and Saule, 2021; Al-Khazraji, 2023). Like any other assumption, this extrapolation could be erroneous which results in inaccuracies in time-to-depth conversion. This would influence every subsequent analysis and cause misinterpretation of the structures and lithofacies distribution in the subsurface which in turn is a crucial step in assessing hydrocarbon and/or geothermal reservoirs, as well as underground energy storage

Yandy Scientific Press

*Corresponding author.
E-mail address: akamenski@hgi-cgs.hr (A. Kamenski); marko.cvetkovic@rgn.unizg.hr (M. Cvetković); josipa.kapuralic@rgn.unizg.hr (J. Kapuralić); iva.kolenkovic@rgn.unizg.hr (I. Kolenković Močilac); ana.brckovic@rgn.unizg.hr (A. Brcković).

Received June 27, 2024; revised July 20, 2024; accepted August 12, 2024; available online August 17, 2024.

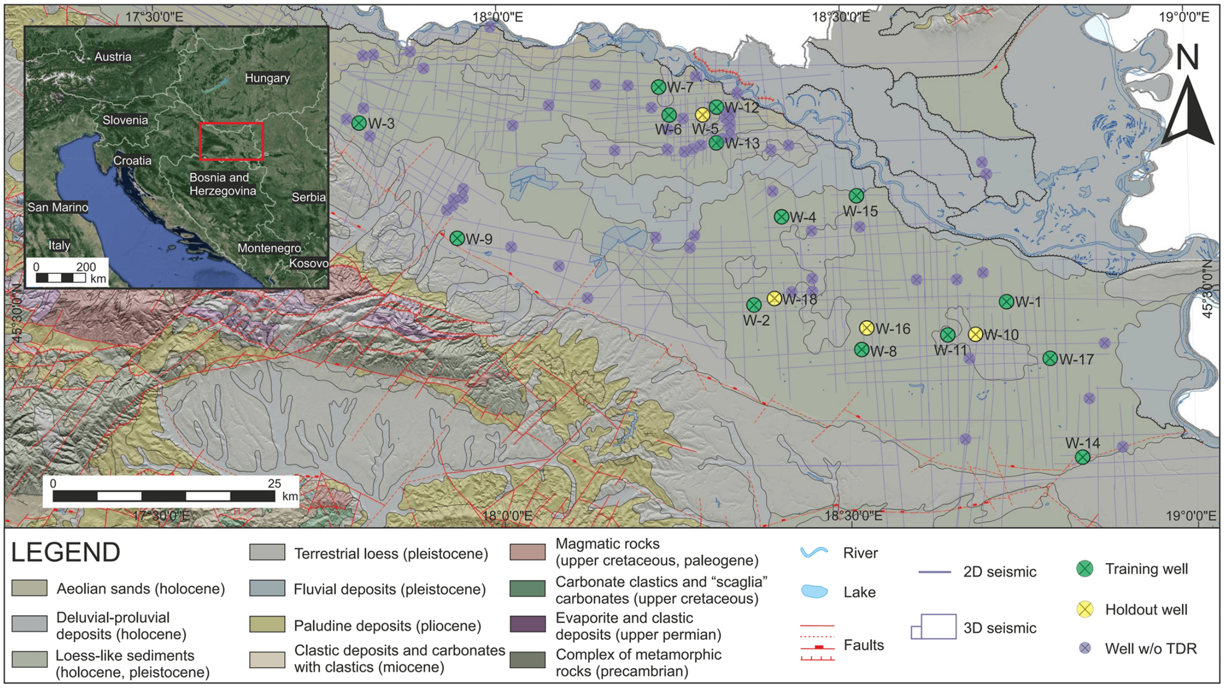


Fig. 1. Study area with well and seismic data displayed alongside basic surface geology adopted from the basic geological map of the Republic of Croatia 1:300,000 (HGI-CGS, 2009).

and CO_2 storage objects.

This research aims to investigate the feasibility of developing a more precise methodology for solving the TDR in cases with limited or no geophysical data. Initially, the accuracy of the conventional methodology was assessed in a way that TWT values were extrapolated from a nearby well. Subsequently, a methodology was developed for deriving TWT data for wells lacking acoustic logs or vertical seismic profiling, using artificial neural networks (ANNs) in combination with interpretation of basic well logs which were obtained even on older wells. Finally, results obtained using these two approaches were compared to show the average difference in the time-to-depth domain and the distribution of error from both methods.

The study aims to present a novel approach and assess its efficiency compared to traditional methods, potentially leading to significant cost savings in future subsurface explorations. Additionally, its application may unveil previously unknown correlations and dependencies within existing data, which would be beneficial for geo-energy exploration, not only in the Drava Basin but also in similar regions undergoing exploration.

2. Geological settings

Data prediction and analysis were performed on an area with a complex geological setting, situated in the eastern part of the Drava Basin which is in the southwestern part of the Pannonian Super Basin (Fig. 1).

Within the broader area of interest, three distinctive geological units can be distinguished. The first unit consists of the crystalline basement, primarily composed of partially metamorphosed Paleozoic magmatic rocks with the presence of metamorphosed sediments (Pamić and Lanphere, 1991; Pamić, 1998). The second unit is characterized by Mesozoic carbonates (Velić, 2007; Malvić and Cvetković, 2013), often referred to as "Base Tertiary" (Velić, 2007). The third unit includes Neogene and Quaternary sediments representing the basin infill (Saftić et al., 2003; Malvić and Cvetković, 2013).

The area experienced continental rifting from the Ottnangian to the Badenian periods, accompanied by a shift in stress orientation that led to sinistral transcurrent faults and the formation of narrow asymmetrical half-grabens (Lučić et al., 2001; Pavelić, 2001; Saftić et al., 2003; Pavelić and Kovačić, 2018). During the Otnnangian and Carpathian, sedimentation was predominantly characterized by coarse-grained clastic sediments deposited in alluvial to lacustrine environments. Meanwhile, sporadic occurrences of pyroclastics in the Drava Basin (Fig. 1) are related to the volcanic activity associated with rifting (Saftić et al., 2003).

A significant change in the depositional environment took place during Middle Badenian, due to marine transgression (Ćorić et al., 2009), which caused a shift from lacustrine to marine deposition. This transition resulted in sedimentation of thick marl layers with occurrences of coarse-grained clastic sediments, reflecting the occasional activity of gravity flows (Ćorić et al., 2009). Sarmatian is marked by the end of syn-rift extension and local compression (Saftić et al., 2003; Pavelić and Kovačić, 2018), as well as the isolation of Paratethys and accompanying salinity fluctuations (Pavelić and Kovačić, 2018). In these circumstances, deposition of coarse-grained clastic sediments, calcarenites, and limestones took place, whereas, in deeper parts of the basin, fine-grained sediments deposited, occasionally with sandstone occurrences resulting from sediment gravity flow (Pavelić and Kovačić, 2018).

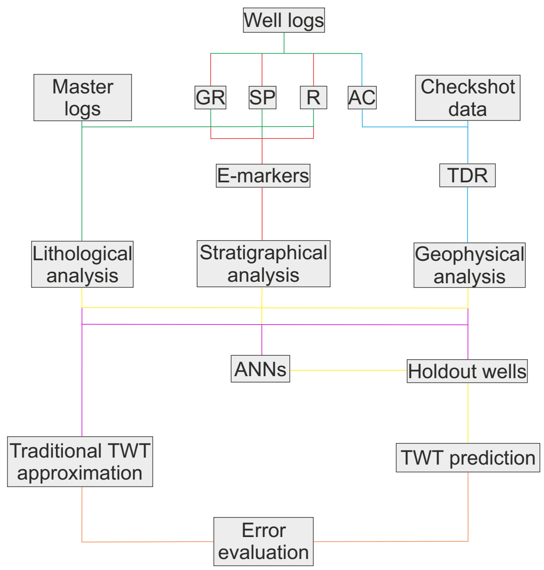


Fig. 2. Schematic representation of the workflow in which each line color represents the sequential process steps necessary for conducting each analysis from the provided input data.

During the Pannonian period, post-rift thermal subsidence took place (Lučić et al., 2001), due to the isostatic sinking of the crust thinned by mantle diapirism (Stegena et al., 1975), initially causing the deepening of the lake and deposition of deep-water marls (Pavelić and Kovačić, 2018). However, significant sediment supply without a corresponding increase in accommodation space resulted in the transition of sedimentation from lacustrine to deltaic environments, leading to the basin's infilling (Saftić et al., 2003; Pavelić and Kovačić, 2018). Neotectonics activity during the Pliocene and Quaternary resulted in compression and dextral transcurrent displacements, filling the remnants of Lake Pannon with coarse clastic sediments and clay. The Pleistocene glacial periods are marked by the deposition of loess sediments and aeolian sands (Wacha et al., 2013), while interglacial periods were marked by lacustrine and marsh sedimentation (Pavelić and Kovačić, 2018).

3. Data and methodology

This study aimed to develop a method for defining TDR that would be more reliable compared to extrapolation of the velocity functions from a neighboring well. This implies evaluating the data which has been almost always available within the well and which would have an impact on the change of velocities in the subsurface. All the historical well data had the basic Electrical well log curves which were employed to differentiate permeable and impermeable units and establish their boundaries, as well as to estimate fluid saturations. Units were correlated throughout the study area based on the interpretation of the resistivity and spontaneous potential curves, supplemented with other well data. These interpretations provided the basis for stratigraphic analysis. Another key part was to evaluate how many distinctive units with different lithological compositions could be defined in the studied area. This was evaluated from well reports based on the descriptions of lithology and selection of interpreted well markers.

This research utilized data from 18 wells (Fig. 1) drilled by the INA company as a part of their oil and gas exploration campaign between the 1980s and 2010s. The wells have an average depth of around 2,300 m, with Well-7 being the shallowest at 1,300 m depth and Well-3 being the deepest well with 4,110 m depth. These wells were selected based on the availability of acoustic well logs through the entire or the majority of the drilled section. The summarized workflow covering the entire process, from data preparation to the neural network deployment, is illustrated in Fig. 2.

Initially, a TDR was applied to all wells to enable transformation from measured depth to TWT domain for validation purposes. These steps are highlighted in blue in Fig. 2. The wells in which seismic velocity measurements or vertical seismic profiling data had been recorded were selected as input for the presented study (wells which are labeled in green and yellow in Fig. 1). The check-shot data were used to calibrate the acoustic logs with the check-shot travel times to create the TDRs for these 18 wells. Acoustic log calibration provides more accurate time-to-depth relationships and corrects for acoustic log drift due to equipment calibration issues, well conditions, or environmental factors (Mari et al., 2020). The calibration adjusts the acoustic log cumulative travel times to match the smoothed check-shot times. Prior to calibration, the acoustic logs were despiked. Despiking improves the overall quality of the acoustic logs by eliminating outliers and anomalies that do not represent the true properties of the subsurface (Rider, 2002).

Well logs, including gamma ray, spontaneous potential (SP), and short- and long-normal resistivity (R16, R64) were analyzed primarily to distinguish permeable and impermeable layers. Data preparation of these logs was performed in the Interactive Petrophysics software. Well logs are typically recorded in multiple intervals at various depths, necessitating the merging of these intervals. In our case, nearly every well log type required some form of conversion, rescaling, or normalization since well conditions change with every technical column due to the difference in temperature and mud properties (Bassiouni, 2013). Given the significance of the SP log and the common occurrence of SP inversion in certain sections of the log, the initial step involved analyzing the resistivity and gamma ray logs to identify inversed SP intervals. The occurrence of inversed intervals is due to the extremely low mineralization of the formation water, which can be below 5 g/l at several thousand meters of depth in the Drava Basin (Pavlin, 2022). Once these intervals were determined, they were assigned appropriate SP values for shale and clean formation, to ensure their reliability in further analyses.

Creation of TDRs, as well as lithological and stratigraphic interpretation were conducted in the Schlumberger Petrel software. This was followed by the differentiation of permeable from impermeable intervals per meter resolution. Regional well log markers were identified on the resistivity curve, defining boundaries of stratigraphic units (1-4). This characterization provided input parameters for ANN analysis.

ANNs are computational models inspired by the architecture and functionality of biological neural networks in the human brain. They comprise interconnected nodes, or "neurons", that process and transmit data (Prieto et al. (2016) and references therein).

The input parameters for the ANNs learning process consisted of three variables. Firstly, the true vertical depth subsea and the ratio of cumulative "sandstone to shale" ratio (S_s/S_h) both serve as continuous input variables. Here, S_s denotes the cumulative thickness of permeable layers up to a certain depth, while S_h represents the cumulative thickness of impermeable layers as one of the controlling factors of the velocity distribution in the subsurface. The third input variable denotes the categorical variable type, namely the stratigraphic category of the unit situated at the observed depth/case. These categories represent assigned values ranging from 1 to 4 which were defined by the stratigraphic differentiation conducted using interpreted well log markers. Finally, these input parameters were the basis for the ANNs prediction of TWT.

The analysis was conducted in TIBCO Statistica software using a regression approach for the time series. Networks were configured as multi-layer perceptron (MLP). MLPs are a type of artificial neural network with multiple layers of nodes, which process and transmit information (Thimm and Fiesler, 1997). Out of 18 wells, 14 were utilized for training the neural networks, while four were reserved for testing ANN on holdout data (Fig. 1). Neural network architecture was restricted to a minimum of three and a maximum of 17 neurons in the hidden layer. To mitigate overfitting during neural network training, a smaller value of weight decay was applied to enhance the network training. Weight decay penalizes large weights, thereby promoting the development of a simpler and more generalized model structure and ultimately enhancing performance on new, unseen data (Thimm and Fiesler, 1997).

Neural networks went through training, testing, and validation using a dedicated dataset consisting of four variables with a resolution of one meter. In total, there were more than 27,000 cases from 14 wells that were used for the training process of the neural networks. Ten neural networks with the most successful performance in training and testing were selected and used as an ensemble for the analysis (Hansen and Salamon, 1990). In a subsequent phase of ANNs analysis, trained neural networks were applied to predict target TWT values for four test wells (holdout wells) not included in the training process consisting of more than 8,000 cases.

The performance of the model was evaluated by calculating the error relative to the measured TWT values for each data point per resolution of one meter and calculating the absolute mean value for all cases within a well (orange steps in Fig. 2). For better visual comparison of the magnitude and distribution of the error, box and whiskers plots as well as maps illustrating error distribution were used.

4. Results

Well log interpretation, following the workflow shown in Fig. 2, enabled the differentiation of permeable and impermeable units. Moreover, the combination of these results with lithology data from Master logs enabled the definition of vertical lithology distribution which is displayed in the "Lithology" column in Fig. 3. Additionally, a maximum of four stratigraphic intervals were interpreted for each well. These intervals represent sediments of an age interval, similar in lithology and, if applicable, sedimentary environment as factors that can influence the change of petrophysical properties of rocks with burial. The first interval represents Pliocene-Quaternary unconsolidated sands, clays, gravels, and occasional coals. The second interval consists of Upper Miocene-Pannonian, which are predominantly sandstones and marls. The third interval represents a lithologically heterogeneous sediment of Lower and Middle Miocene breccias, conglomerates, sandstones, marls, and limestones, with sporadic occurrences of effusive. The fourth interval encompasses all rocks older than Miocene, including older pre-Neogene sediments as well as the metamorphic and magmatic complex comprising the Basement. One of the factors which was considered when the division was performed, was the impact of the present-day burial depth and its accompanying compaction and diagenetic processes influencing units' petrophysical properties. Consequently, it was presumed that the largest change of interval velocities would be from Neogene infill to Basement Neogene rocks (from intervals three to four). This transition represents the change from the lithologies in which petrophysical properties are governed by compaction to lithologies where fractures play a crucial role in controlling petrophysical properties, including crystalline rocks or Triassic dolomites. Most of the selected wells have been drilled through the first three intervals and finished in the fourth as hydrocarbon accumulations at some locations were expected even in the Basement rocks.

The selection of a representative training dataset is paramount for successful ANNs prediction. As was already established, a comprehensive training dataset was defined, reflecting the complexities of geology and petrophysics in the subsurface. The dataset comprised 18 wells, partitioned into a training set of 14 wells and four holdout wells. Holdout wells were used to evaluate the ANN model on previously unseen data. Wells W-5, W-10, and W-16 were selected as holdout wells because they were comprised of all four stratigraphic intervals, while well W-18 was chosen to represent the case where only the first two stratigraphic intervals were developed. The holdout wells have significant geographical positions, being close to one or more training wells. This arrangement allowed for examining whether proximity is always the best criterion for selecting the source for data extrapolation.

Given the continuous nature of the input dataset, time series regression was selected to ensure sequence prediction, aligning with geological and geophysical principles. The neural network architecture was optimized using various parameters and validated through repeated training, testing, and validation processes. The resulting correlation coefficients per each ANN are presented in Table 1, encompassing training, test, and validation process phases.

The correlation coefficient of the training set exhibited an accuracy range between 0.995711 and 0.995877. The testing data displayed a high accuracy range of 0.997796 to 0.997986. Similarly, the validation dataset yielded comparable performance, with an accuracy range of 0.997811 to 0.997982.

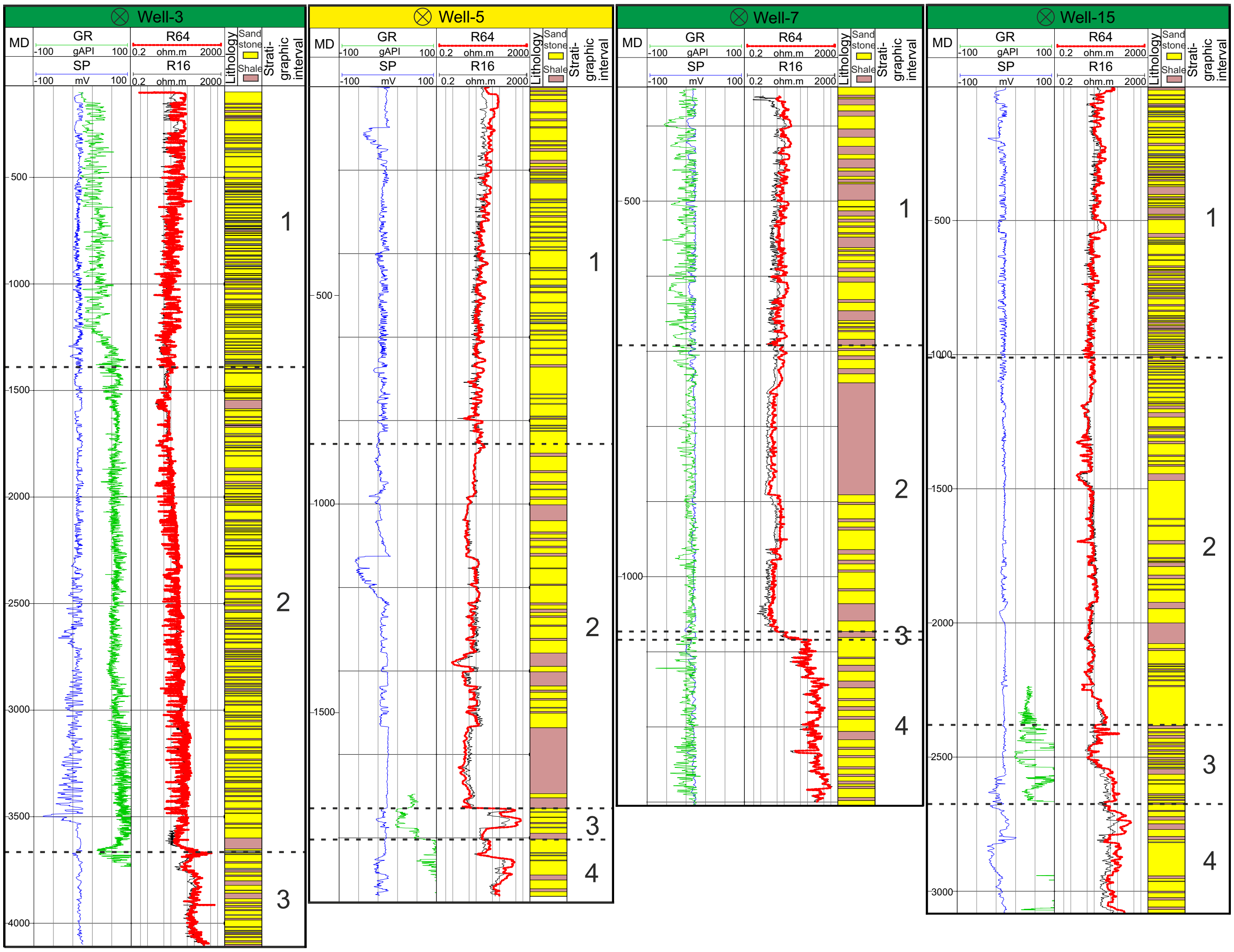


Fig. 3. Representation of the available well logs, interpretation of S_s and S_h as general lithology and stratigraphic intervals for four wells.

Table 1. Correlation coefficients, RMSE, MAE and RSQ of observed and predicted data for ten neural networks for training, test, and validation datasets from 14 wells.

Architecture	TWT train (ms)	TWT test (ms)	TWT validation (ms)	RMSE	MAE	RSQ
MLP 6-17-1	0.995770	0.997845	0.997878	34.008562	25.699704	0.995660
MLP 6-11-1	0.995823	0.997940	0.997967	33.426744	24.911437	0.995816
MLP 6-4-1	0.995716	0.997812	0.997829	34.364610	25.690295	0.995575
MLP 6-9-1	0.995831	0.997959	0.997937	33.399862	25.004618	0.995812
MLP 6-14-1	0.995852	0.997881	0.997911	33.568955	25.978290	0.995774
MLP 6-16-1	0.995877	0.997986	0.997982	33.111525	24.540709	0.995890
MLP 6-15-1	0.995743	0.997837	0.997848	34.237448	25.675535	0.995602
MLP 6-4-1	0.995845	0.997963	0.997971	33.256294	25.232209	0.995851
MLP 6-11-1	0.995804	0.997906	0.997895	33.734191	25.044579	0.995732
MLP 6-10-1	0.995711	0.997796	0.997811	34.499635	25.997086	0.995535

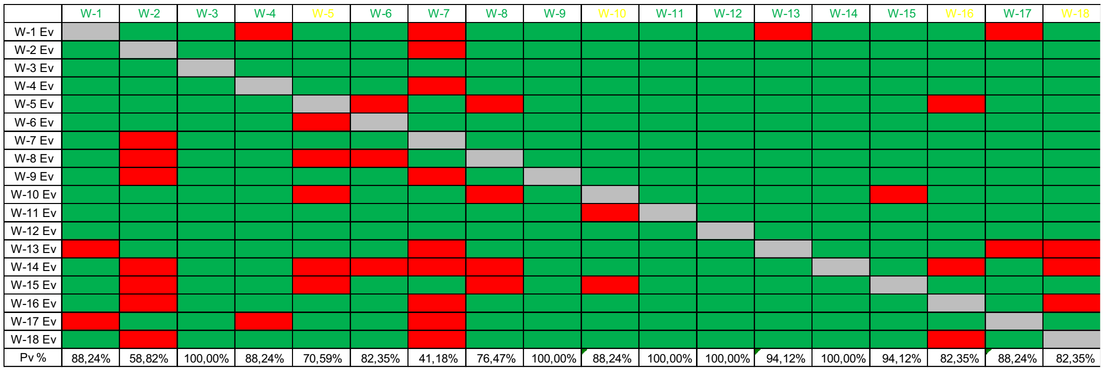
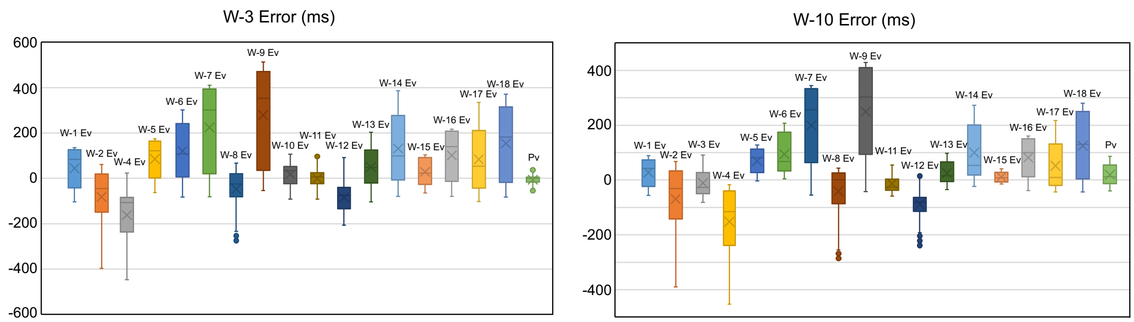


Fig. 4. A visual representation of the successfulness of the ANN prediction over the extrapolation approach. Well names used in the ANN training process are in green color, while holdout wells are highlighted in yellow. Cases with better Ev are labeled in red, while wells with better Pv are in green. Pv% represents a percentage of wells for which prediction via ANN gave more successful results while the green field indicates better results for Pv and the red field with the Ev approach.



(a) (b)

Fig. 5. Box and whisker plots representing the distribution of error in milliseconds (ms) for wells W-3 and W-10.

The root mean square error (RMSE) indicates that the average difference between values predicted by the ANN model and the actual values is approximately 34 ms. The mean absolute error (MAE) is even smaller, around 25 ms. The coefficient of determination (RSQ) value which is a measure of the goodness of the fit, demonstrates that the data fits the regression model very well. For prediction purposes, all ten neural networks were combined into an ensemble.

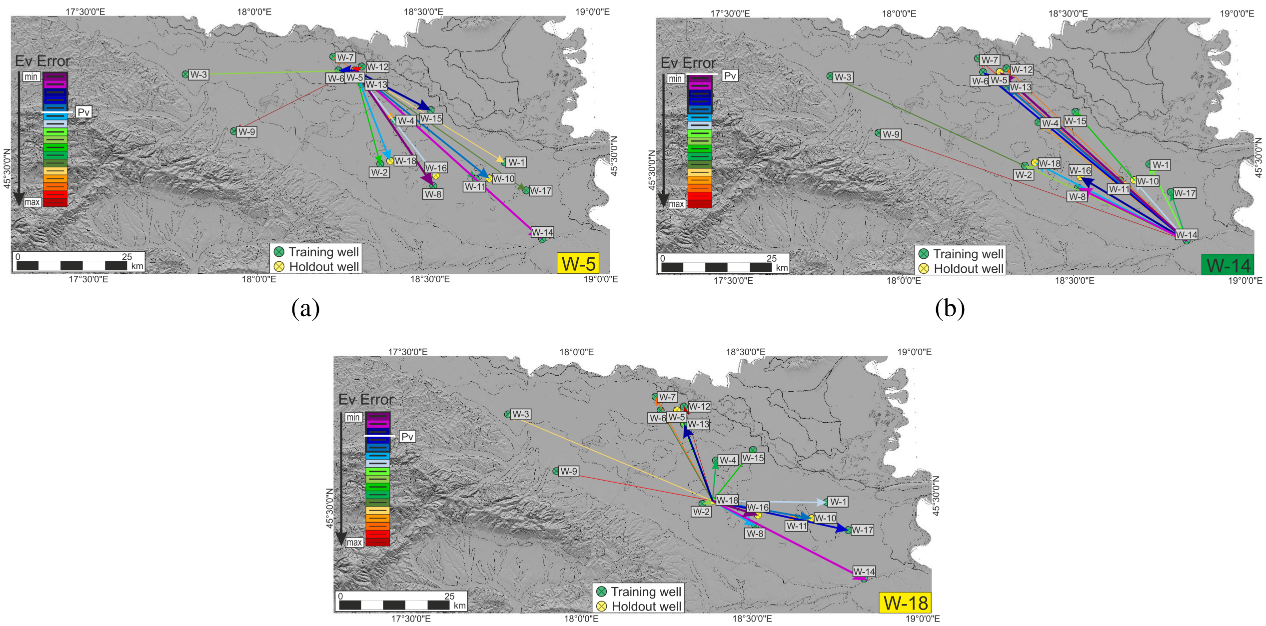
To evaluate the successfulness of the ANN deployment for the task, the TWT prediction results were compared to the TWT extrapolation values from nearby wells based on the average error relative to measured values from the well data. Error values were calculated for both the extrapolated TWT values versus measured values (Ev) and the ANN-predicted values versus measured TWT values (Pv). A matrix table was generated for each well to summarize these comparisons (Appendix 1).

When focusing solely on the success of the ANN prediction results, it is evident that 15 out of 18 wells show a closer fit to the measured values compared to those extrapolated from surrounding wells, i.e., they are more successful in over 75% of cases. It was found that for wells W-3, W-9, W-11, W-12, and W-14 prediction of TWT was outperforming the extrapolation method in 100% of cases (Fig. 4). Notably, four of these wells serve as holdout wells, indicated with yellow labels in Fig. 4 and Appendix 1. All of them achieved successful predicted values (Pv), with three demonstrating more than 80% more accurate outcomes than the extrapolation method (Fig. 4). Only a few instances showed smaller errors than those when TWT was predicted by the ANN analysis, highlighted in red in Appendix 1. The extrapolation approach achieved the best results for well W-7, with values from nine wells showing lower average errors.

Results are even better illustrated through the box and whisker plot (Fig. 5). The plot shows two representative wells: W-3, which was used in the ANN training process, and W-10, a holdout well. For W-3, ANN-predicted values exhibit the smallest average error compared to all errors calculated from extrapolating values from neighboring wells. In the case of W-10, there are only two instances (that correspond to extrapolations from W-11 and W-15) where the extrapolated values have a smaller error than the ANN-predicted values. Appendix 2 illustrates all case scenarios.

5. Discussion

The results of the ensemble ANN model demonstrate its ability to generate synthetic values for time-to-depth conversion using lithological parameters that can be interpreted from the most basic well logs and stratigraphic interval delineation. The trend observed in the predicted values closely aligns with



(a) (b) (c)

Fig. 6. Distribution of average errors obtained through extrapolation of TWT values from wells with available velocity information. The five best results, indicated by the smallest errors, are highlighted with thick arrows ranging in color from dark blue to dark purple (colors within the thick black rectangle in Ev Error Legend).

the measured values, as evidenced by the small average error, which ranges from a minimum of 11 ms to a maximum of 52 ms, with an average of 26 ms. Outlier and spread of the error are significantly lower when applying the ANN approach instead of an extrapolation of the TDR. Appendix 1 quantitatively presents the average errors, selected as the decisive parameter to evaluate the accuracy of ANN predictions.

The distribution of these errors and their relationship are illustrated in Fig. 6 for three representative cases and in Appendix 3 for all 18 cases. Upon thorough examination, several observations can be drawn. When extrapolating time-to-depth relationships, wells W-7, W-9 and W-12 consistently provide the largest errors (Appendix 2), indicating the poorest results for estimating TWT values in wells lacking velocity information. For Well-9, this is understandable given its greater distance from other wells in the area of exploration.

The analysis presents that proximity between wells does not guarantee accurate extrapolation of time-to-depth relations. Instead, the ANN predictions offer consistently more reliable results, even for the holdout wells which were never included in the building of the ANN model. This can be observed in Fig. 6 and Appendix 3. Despite the expectation that closely located wells would provide dependable time-to-depth information for extrapolation, the findings of this study disprove this assumption. For instance, in Fig. 6, case W-5 shows that wells W-7 and W-12, though remarkably close to well W-5 which is treated as a no velocity data well in this case scenario, yield extremely poor results with the highest errors compared to more distant wells.

This pattern is not uniform across all close pairs of wells. For example, extrapolating TWT values from W-6 to W-5 results in a satisfactory small average error of 21 ms. On the other hand, well W-2 despite being the closest, displays a poor correlation when extrapolating TWT information for W-18.

It is evident that, in most instances, extrapolating time-to-depth relations from nearby wells leads to significantly poorer outcomes compared to extrapolation from distant wells. This discrepancy could be attributed to substantial variations in subsurface lithology distribution and/or general orientation of structures.

The predicted values generated by the ANN analysis (indicated by the thick white line and "Pv" mark in the color legend in Fig. 5) demonstrate exceptional accuracy for both the training wells and, most importantly, the holdout wells (marked with green and yellow labels, respectively). With a high precision percentage observed, TWT values obtained through neural network analysis prove to be more reliable, especially for holdout wells (highlighted in yellow in Appendix 1, Fig. 6 and Appendix 3).

The distribution trend is also evident in Fig. 6, particularly highlighting the best five approximations. The results obtained from extrapolating TWT values reveal a predominant NNW-SSE strike, as observed in the top five approximations marked with the thickest arrows ranging from dark blue to dark purple. These findings are consistent with the strike of geological structures and sediment paleotransport orientation documented in recent investigations of the area (Rukavina et al., 2023; Špelić et al., 2023; Matošević et al., 2024).

6. Conclusion

Accurate determination of time-to-depth parameters plays a crucial role in various applications, including drilling optimization and reservoir characterization. However, this process often entails significant economic and technical risks. To address these challenges and mitigate associated risks, a novel approach has been developed and is presented herein.

Building an ANN model to solve time-to-depth relationships has proven highly effective. The ANN model exhibits a high correlation coefficient for the training, testing, and validation set, all above 0.99, with root mean square errors under 35 ms and mean absolute errors around 25 ms. This level of accuracy surpasses any method applied so far on Pannonian Super Basin data, including the common approach of extrapolating values from nearby wells. The ANN model not only has smaller absolute errors but is also significantly less sensitive to outliers. Since model predictions depend on the local geological characteristics of the training data, a separate ANN model for solving TDRs should be developed for each basin or super-basin to account for their unique geological features.

Overall, this study presents the effectiveness of the ANN framework in conventional, dominantly clastic environments, tailored to the specific objectives of parameter prediction. Depending on the desired objective, the focus can vary from precise values to understanding broader trends and variations within the wells. This approach enhances efficiency and adaptability, improving the accuracy of subsurface models. The methodology remains open to further refinement through activities such as spatial information integration, hyper-parameter fine-tuning, and the development of tailored models for specific geological settings.

The results highlight the effectiveness of the proposed methodology in deriving TWT values from depth, lithological, and stratigraphical parameters by ANN analysis. This ANN-driven solution proves to be an effective approach for obtaining time-to-depth relations in mature basins with a large number of historical well data often lacking acoustic well log measurements and vertical seismic profiling.

Acknowledgements

This research was conducted as part of the research project founded by the Croatian Science Foundation (No. HRZZ UIP-2019-04-3846): "GEOlogical characterization of the Eastern part of the Drava depression subsurface intended for the evaluation of Energy Potentials (GEODEP)". The authors would like to extend their gratitude to Schlumberger company for generously donating the Petrel academic software license. Additionally, we acknowledge the Croatian Hydrocarbon Agency for granting permission to access the data, without which the analysis presented in this paper would not have been possible.

Supplementary file

https://doi.org/10.46690/ager.2024.10.05

Conflict of interest

The authors declare no competing interest.

References

Aker, E., Tveiten, O. G., Wynn, T. Seismic pore pressure prediction at the Halten Terrace in the Norwegian Sea. Petroleum Geoscience, 2020, 26(2): 346-354.

Al-Khazraji, O. N. A. Cross-validation of time-depth conversion and evaluation of different approaches in the Mesopotamian Basin, Iraq. Exploration Geophysics, 2023, 54(3): 288-315.

Bassiouni, Z. Theory, Measurement, and Interpretation of Well Logs (4th edition). Texas, USA, Society of Petroleum Engineers, 2013.

Cao, J., Shi, Y., Wang, D., et al. Acoustic log prediction on the basis of kernel extreme learning machine for wells in GJH survey, Erdos Basin. Journal of Electrical and Computer Engineering, 2017, 2017: 3824086.

Ćorić, S., Pavelić, D., Rögl, F., et al. Revised Middle Miocene datum for initial marine flooding of North Croatian Basins (Pannonian Basin System, Central Paratethys). Geologia Croatica, 2009, 62(1): 31-43.

Hansen, L. K., Salamon, P. Neural network ensembles. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1990, 12(10): 993-1001.

Hart, D. M., Balch, R. S., Weiss, W. W., et al. Time-to-depth conversion of nash draw "L" seismic horizon using seismic attributes and neural networks. Paper SPE 59555 Presented at SPE Permian Basin Oil and Gas Recovery Conference, Midland, Texas, 21-23 March, 2000.

HGI-CGS. Basic Geological Map of Croatia 1:300,000. Zagreb, Croatia, HGI-CGS, 2009.

Inichinbia, S., Saule, P. O. Well-to-seismic tie of a field onshore of the nigerian delta. Journal of Applied Sciences and Environmental Management, 2021, 25(1): 53-58.

Lučić, D., Saftić, B., Krizmanić, K., et al. The Neogene evolution and hydrocarbon potential of the Pannonian Basin in Croatia. Marine and Petroleum Geology, 2001, 18(1): 133-147.

Malvić, T., Cvetković, M. Lithostratigraphic units in the Drava Depression (Croatian and Hungarian parts)-a correlation. Nafta, 2013, 64(1): 27-33.

Mari, J. L., Vergniault, C., Coppens, F. Acoustic logging, in Well seismic surveying and acoustic logging, edited by J. L. Mari and C. Verginault, EDP Sciences, Paris, pp. 77-102, 2020.

Matošević, M., Garzanti, E., Šuica, S., et al. The Alps as the main source of sand for the Late Miocene Lake Pannon (Pannonian Basin, Croatia). Geologia Croatica, 2024, 77(2): 69-83.

Pamić, J. Crystalline basement of the South Pannonian Basin based on surface and subsurface data. Nafta, 1998, 49(12): 371-390.

Pamić, J., Lanphere, M. Hercynian granites and metamorphic rocks from the Mts. Papuk, Psunj, Krndija, and the surrounding basement of the Pannonian Basin in Slavonija (Northern Croatia, Yugoslavia). Geologija, 1991, 34(1): 81-253.

Pavelić, D. Tectonostratigraphic model for the North Croatian and North Bosnian sector of the Miocene Pannonian Basin System. Basin Research, 2001, 13(3): 359-376.

Pavelić, D., Kovačić, M. Sedimentology and stratigraphy of the Neogene rift-type North Croatian Basin (Pannonian Basin System, Croatia): A review. Marine and Petroleum Geology, 2018, 91: 455-469.

Pavlin, I. Formation water salinity in deep permeable layers in Eastern part of Drava Basin. Zagreb, University of Zagreb, 2022. (in Croatian)

Prieto, A., Prieto, B., Ortigosa, E. M., et al. Neural networks: An overview of early research, current frameworks and new challenges. Neurocomputing, 2016: 214: 242-268.

Rider, M. The Geological Interpretation of Well Logs (2nd edition). Sutherland, Scotland, Rider-French Consulting Ltd., 2002.

Rukavina, D., Saftić, B., Matoš, B., et al. Tectonostratigraphic analysis of the syn-rift infill in the Drava Basin, south-western Pannonian Basin System. Marine and Petroleum Geology, 2023, 152: 106235.

Saftić, B., Velić, J., Sztanó, O., et al. Tertiary subsurface facies, source rocks and hydrocarbon reservoirs in the SW part of the Pannonian Basin (Northern Croatia and south-western Hungary). Geologia Croatica, 2003, 56(1): 101-122.

Špelić, M., Kovács, Á., Saftić, B., et al. Competition of deltaic feeder systems reflected by slope progradation: A high-resolution example from the Late Miocene-Pliocene, Drava Basin, Croatia. International Journal of Earth Sciences, 2023, 112: 1023-1041.

Stegena, L., Géczy, B., Horváth, F. Late Cenozoic evolution of the Pannonian basin. Tectonophysics, 1975, 26(1-2): 71-90.

Sun, Z., Yang, S., Zhang, F., et al. A reconstructed method of acoustic logging data and its application in seismic lithological inversion for uranium reservoir. Remote Sensing, 2023, 15(5): 1260.

Thimm, G., Fiesler, E. High-order and multilayer perceptron initialization. IEEE Transactions Neural Networks, 1997, 8(2): 349-359.

Velić, J. Petroleum geology. Zagreb, University of Zagreb, 2007. (in Croatian)

Velić, J., Krasić, D., Kovačević, I. Exploitation, reserves and transport of natural gas in the Republic of Croatia. Tehnički Vjesnik-Technical Gazette, 2012a, 19(3): 633-641.

Velić, J., Malvić, T., Cvetković, M., et al. Reservoir geology, hydrocarbon reserves and production in the Croatian part of the Pannonian Basin System. Geologia Croatica, 2012b, 65(1), 91-101.

Wacha, L., Galović, L., Koloszár, L., et al. The chronology of the Šarengrad II loess-palaeosol section (Eastern Croatia). Geologia Croatica, 2013, 66(3): 191-203.

Supplementary file

From traditional extrapolation to neural networks: Time-depth relationship innovations in the subsurface characterization of Drava Basin, Pannonian Super Basin

Ana Kamenski[1], Marko Cvetković[2,*], Josipa Kapuralić[2], Iva Kolenković Močilac[2], Ana Brcković[2]

[1] *Department of Geology, Croatian Geological Survey, Zagreb 10000, Croatia*

[2] *Faculty of Mining, Geology and Petroleum Engineering, University of Zagreb, Zagreb 10000, Croatia*

E-mail address: akamenski@hgi-cgs.hr (A. Kamenski); marko.cvetkovic@rgn.unizg.hr (M. Cvetković); josipa.kapuralic@rgn.unizg.hr (J. Kapuralić); iva.kolenkovic@rgn.unizg.hr (I. Kolenković Močilac); ana.brckovic@rgn.unizg.hr (A. Brcković).

[*] Corresponding author (ORCID: 0000-0002-4555-6083)

Kamenski, A., Cvetković, M., Kapuralić, J., Kolenković Močilac, I., Brcković, A. From traditional extrapolation to neural networks: Time-depth relationship innovations in the subsurface characterization of Drava Basin, Pannonian Super Basin. Advances in Geo-Energy Research, 2024, 14(1): 25-33.

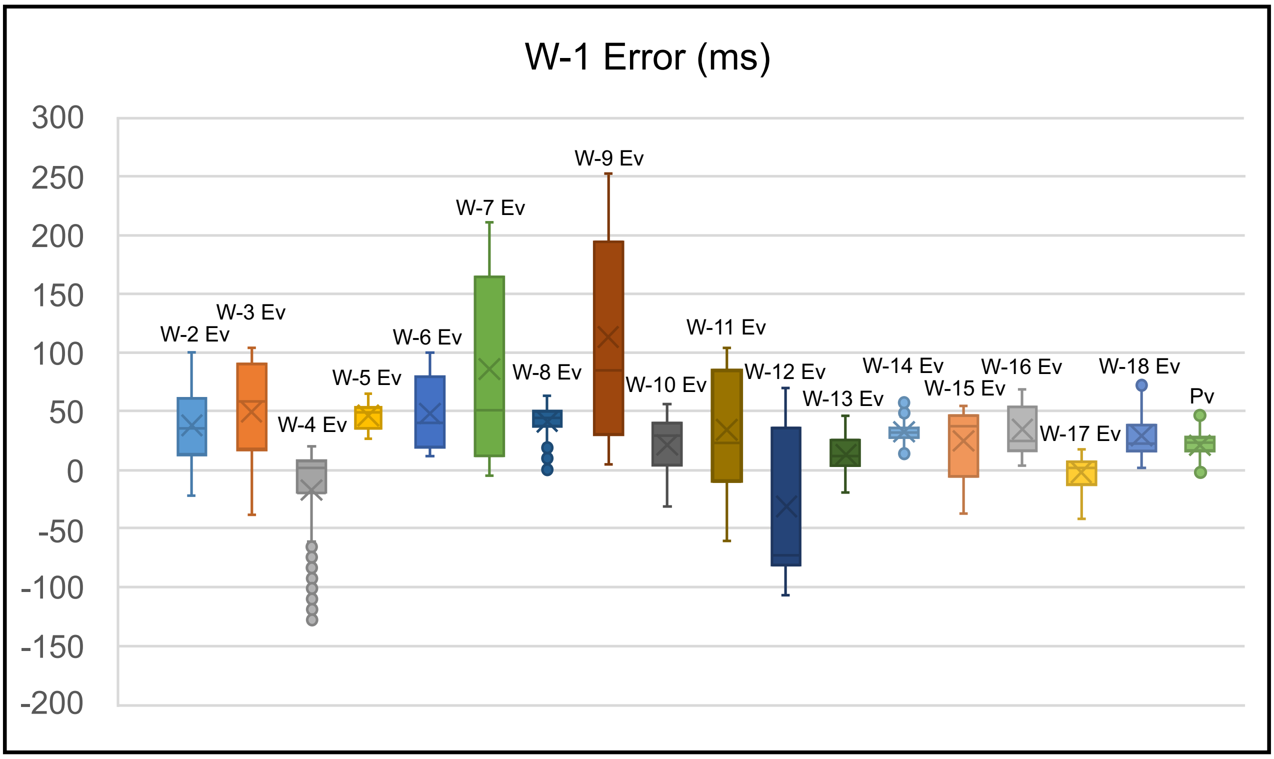
The link to this file is: https://doi.org/10.46690/ager.2024.10.05

Appendix 1

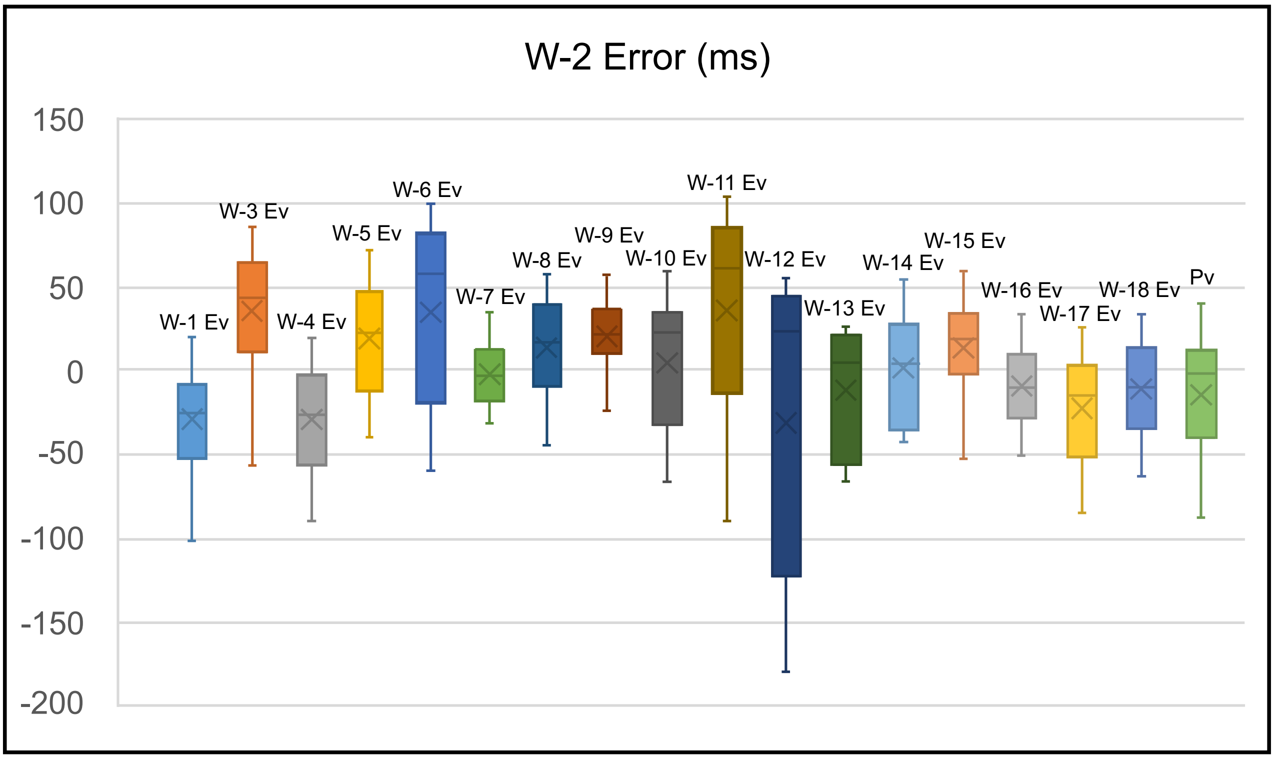
Average absolute error in milliseconds calculated for results TWT obtained through both the extrapolation method (Ev) and ANN analysis (Pv). Wells used in the ANN training process are marked with green labels, while holdout wells are highlighted in yellow. Numbers in red indicate instances where the extrapolation method yielded better results, i.e., smaller errors than those produced by the ANN analysis. Pv% represents a percentage of wells for which prediction via ANN gave more successful results.

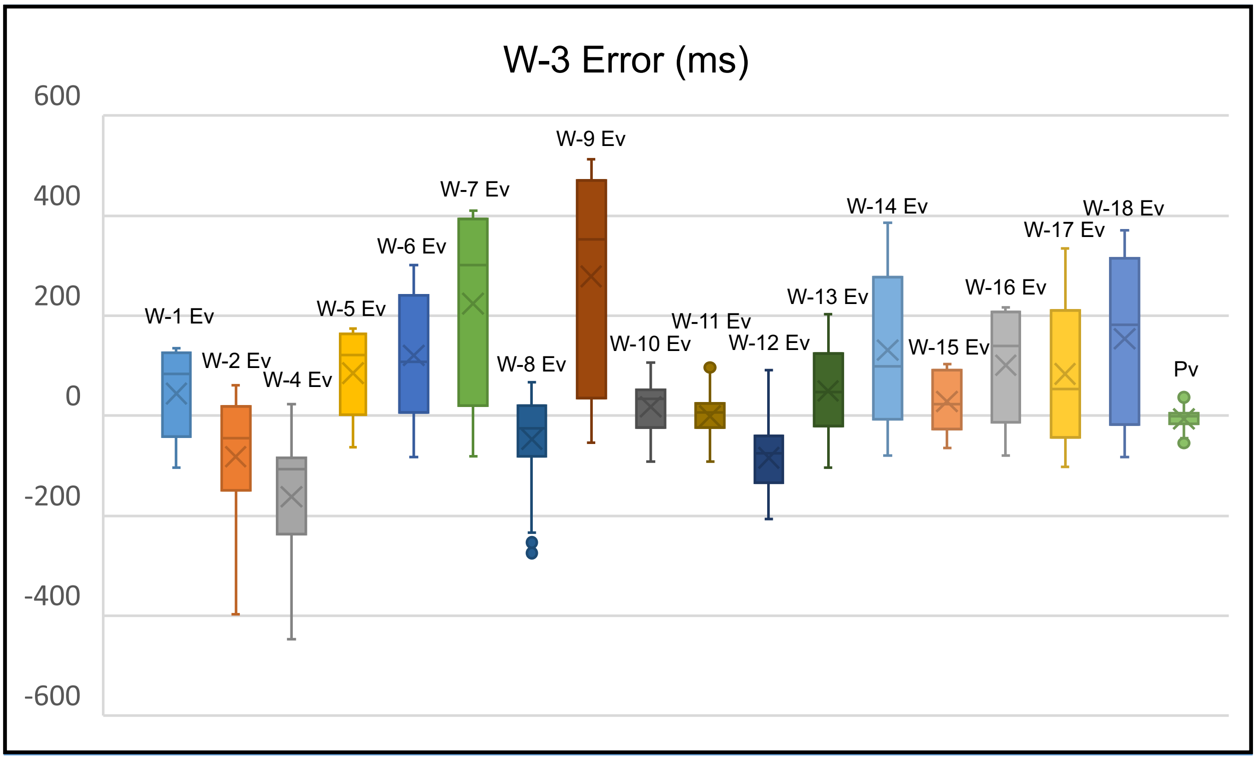
	W-1	W-2	W-3	W-4	W-5	W-6	W-7	W-8	W-9	W-10	W-11	W-12	W-13	W-14	W-15	W-16	W-17	W-18
W-1 Ev		34.5	88.9	28.2	45.0	44.7	18.1	45.2	46.5	49.6	73.1	59.7	22.5	29.5	45.9	38.2	14.6	25.3
W-2 Ev	42.8		101.5	69.5	31.5	52.6	14.8	29.4	24.9	101.5	64.5	95.1	88.6	31.8	55.6	36.2	66.3	27.4
W-3 Ev	55.8	45.1		73.8	31.1	52.3	47.2	37.3	27.3	43.3	35.2	111.5	64.8	45.6	37.7	62.4	65.8	44.2
W-4 Ev	25.7	34.7	162.7		64.8	93.3	19.5	46.1	48.4	152.1	91.7	59.3	138.4	82.3	102.1	86.2	93.4	36.9
W-5 Ev	44.8	34.0	101.1	69.3		19.8	38.7	5.7	32.5	69.2	84.4	74.9	47.9	15.0	54.5	25.2	47.1	25.4
W-6 Ev	48.9	58.7	139.3	72.0	21.2		63.5	29.4	57.4	94.5	78.6	77.1	63.1	24.5	64.6	40.4	51.1	37.3
W-7 Ev	83.6	15.7	246.6	116.3	74.7	114.3		36.6	25.0	212.0	210.5	101.0	170.1	97.6	174.0	74.1	142.5	53.8
W-8 Ev	41.1	29.0	67.4	66.9	7.0	30.8	33.6		27.1	64.8	42.9	74.2	65.0	17.7	32.7	36.2	55.7	22.8
W-9 Ev	111.9	25.9	289.5	145.3	82.5	129.9	25.6	30.0		254.7	240.2	113.3	211.9	116.9	202.8	100.0	179.8	71.6
W-10 Ev	27.5	36.4	47.5	44.8	23.4	36.0	34.5	13.7	35.7		31.7	57.9	29.2	28.9	13.3	47.2	33.9	22.7
W-11 Ev	48.8	63.6	32.3	55.4	38.5	35.9	66.8	34.6	60.8	32.5		73.4	60.2	58.1	34.3	75.9	59.9	50.7
W-12 Ev	65.4	76.4	91.9	48.5	78.5	89.7	63.5	56.1	74.0	89.9	56.6		99.6	86.8	74.1	102.9	85.1	74.9
W-13 Ev	16.8	32.0	84.7	39.7	30.2	39.4	25.2	29.5	42.8	36.3	55.3	56.9		27.9	24.6	39.7	20.8	17.6
W-14 Ev	33.1	26.6	153.6	59.0	12.6	25.1	26.4	12.7	26.3	103.5	97.2	69.1	63.2		62.3	22.3	42.0	13.3
W-15 Ev	34.1	27.7	54.0	47.5	20.0	39.1	31.7	5.6	25.2	17.9	34.7	72.9	25.9	32.2		50.8	36.3	27.4
W-16 Ev	34.2	21.6	128.9	62.2	26.8	40.6	13.6	25.1	29.0	91.5	107.7	72.1	60.4	22.8	74.3		48.9	7.5
W-17 Ev	10.7	30.0	127.0	24.6	44.5	50.9	15.6	38.2	40.2	72.1	68.3	58.8	31.9	36.3	40.4	44.4		22.6
W-18 Ev	29.4	24.6	182.8	56.0	25.6	42.8	15.5	26.0	29.0	134.5	130.8	67.6	92.0	28.7	97.2	11.8	65.9	
Pv	19.9	29.9	15.0	35.4	24.8	31.4	30.4	16.4	20.3	34.7	28.9	51.9	24.3	10.8	15.5	26.8	26.1	20.1
Pv%	88.24	58.8	100.0	88.24	70.59	82.35	41.18	76.47	100	88.24	100	100	94.12	100	94.12	82.35	88.24	82.35

Appendix 2

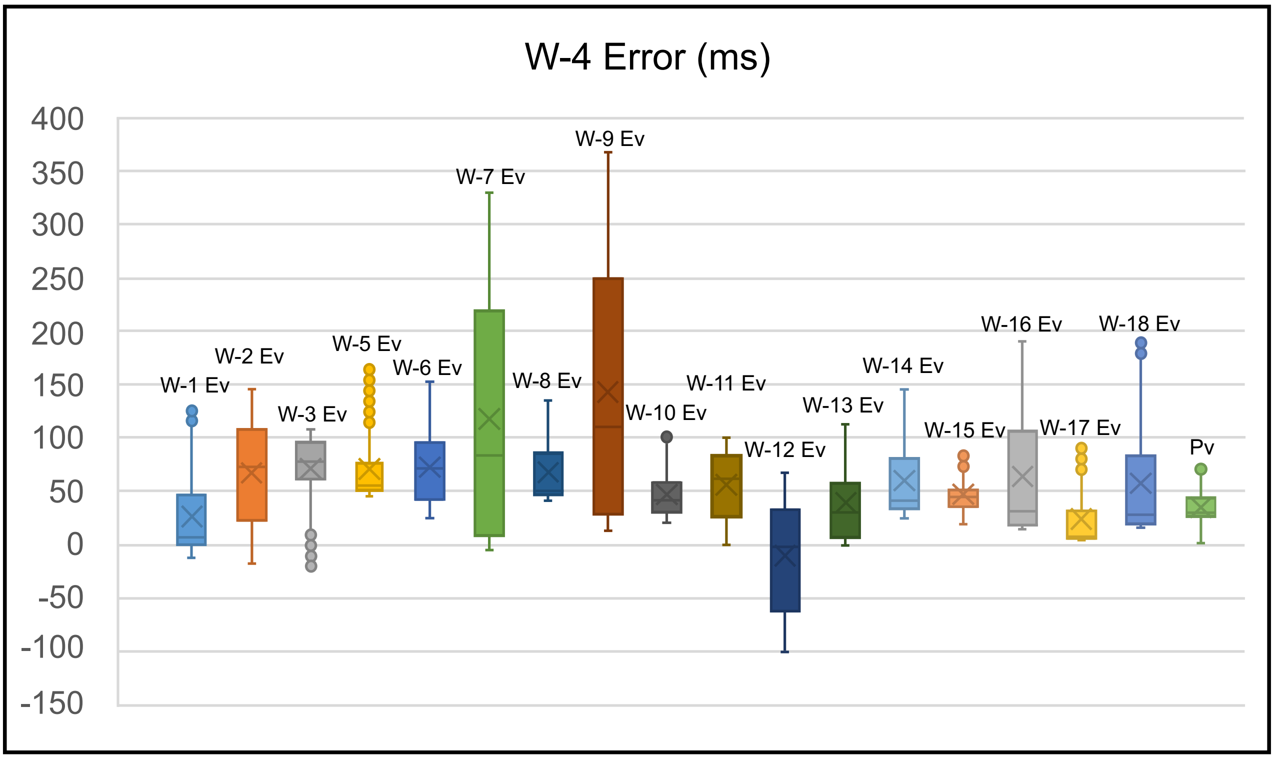


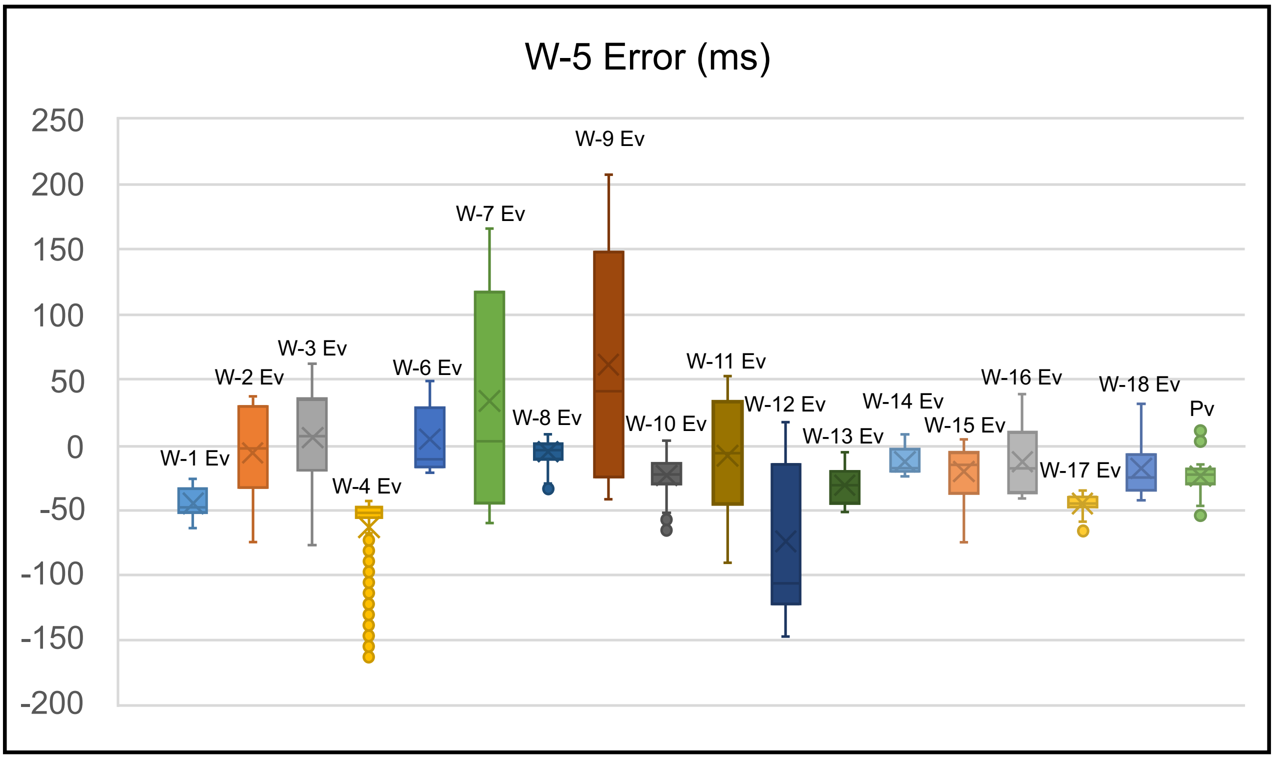
W-1 Error (ms)
300
250
200
150
100
50
0
-50
-100
-150
-200
W-2 Ev
W-3 Ev
W-4 Ev
W-5 Ev
W-6 Ev
W-7 Ev
W-8 Ev
W-9 Ev
W-10 Ev
W-11 Ev
W-12 Ev
W-13 Ev
W-14 Ev
W-15 Ev
W-16 Ev
W-17 Ev
W-18 Ev
Pv


W-2 Error (ms)
150
100
50
0
-50
-100
-150
-200
W-1 Ev
W-3 Ev
W-4 Ev
W-5 Ev
W-6 Ev
W-7 Ev
W-8 Ev
W-9 Ev
W-10 Ev
W-11 Ev
W-12 Ev
W-13 Ev
W-14 Ev
W-15 Ev
W-16 Ev
W-17 Ev
W-18 Ev
Pv

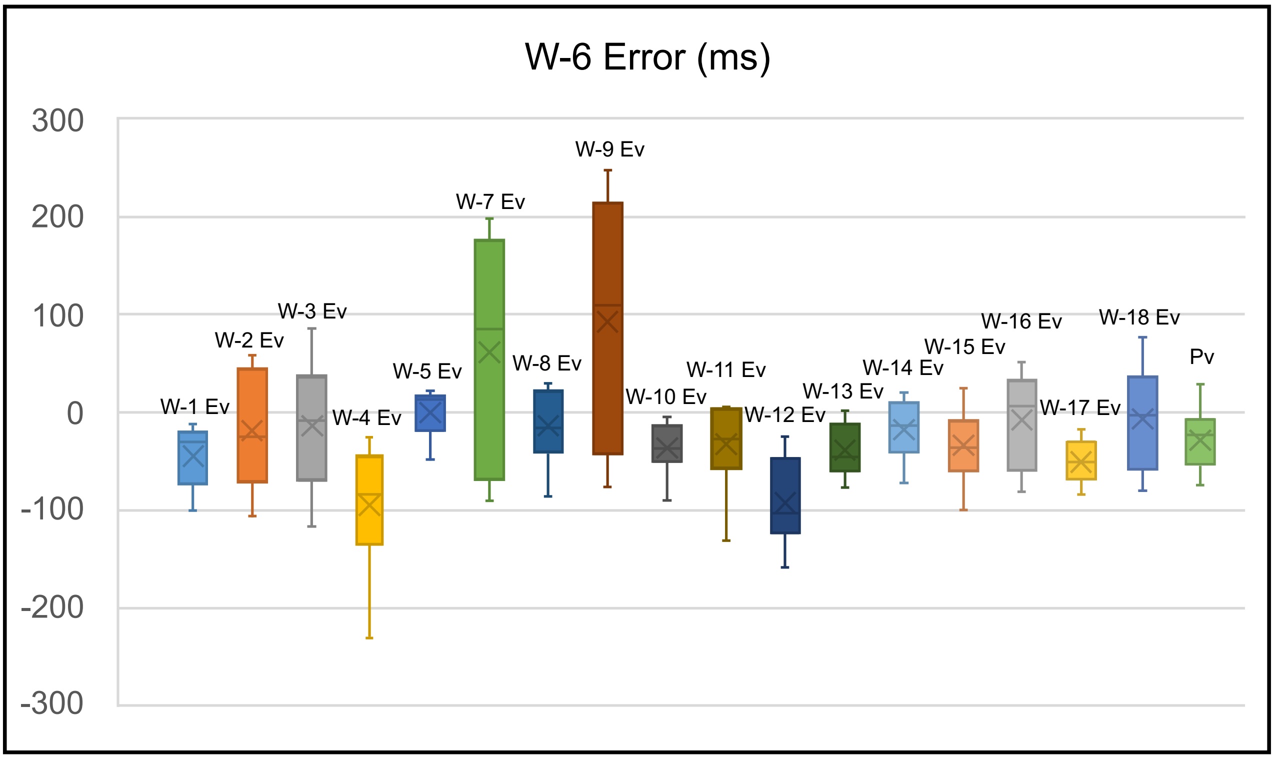


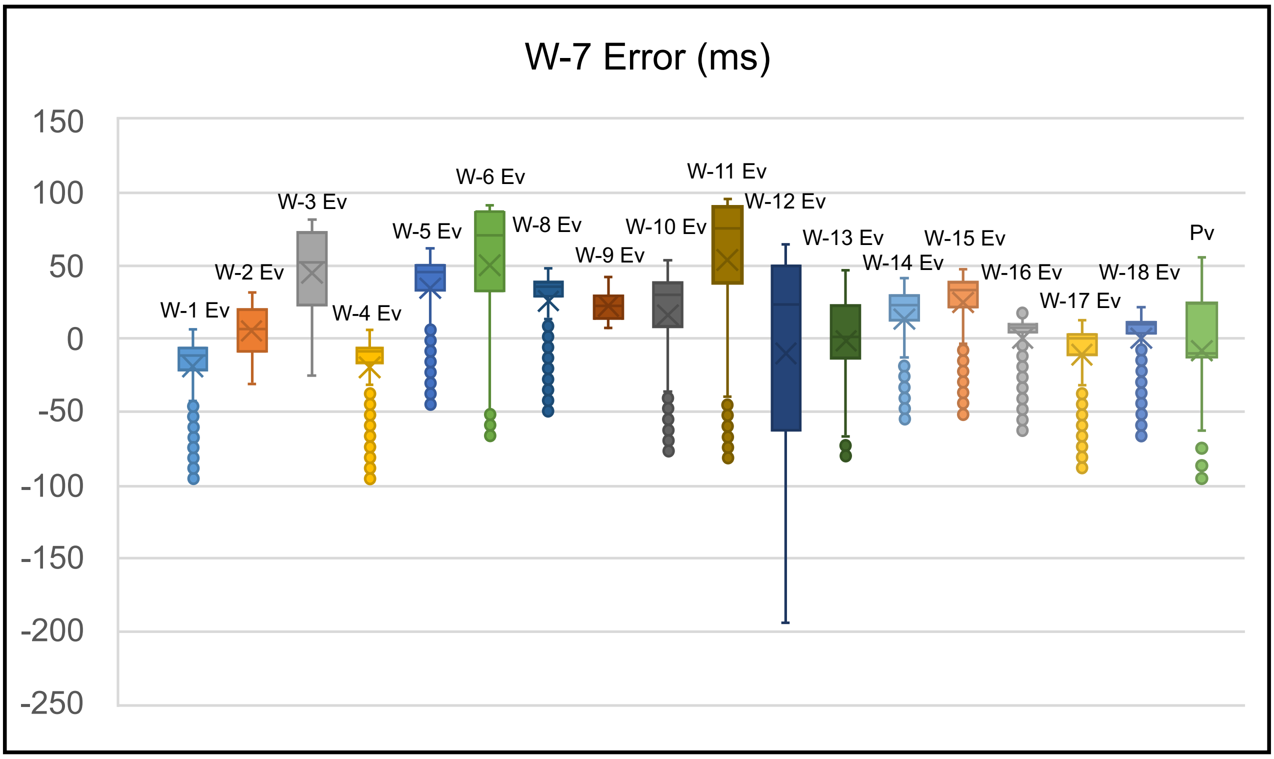
W-3 Error (ms)
600
400
200
0
-200
-400
-600
W-1 Ev
W-2 Ev
W-4 Ev
W-5 Ev
W-6 Ev
W-7 Ev
W-8 Ev
W-9 Ev
W-10 Ev
W-11 Ev
W-12 Ev
W-13 Ev
W-14 Ev
W-15 Ev
W-16 Ev
W-17 Ev
W-18 Ev
Pv


W-4 Error (ms)
400
350
300
250
200
150
100
50
0
-50
-100
-150
W-1 Ev
W-2 Ev
W-3 Ev
W-5 Ev
W-6 Ev
W-7 Ev
W-8 Ev
W-9 Ev
W-10 Ev
W-11 Ev
W-12 Ev
W-13 Ev
W-14 Ev
W-15 Ev
W-16 Ev
W-17 Ev
W-18 Ev
Pv

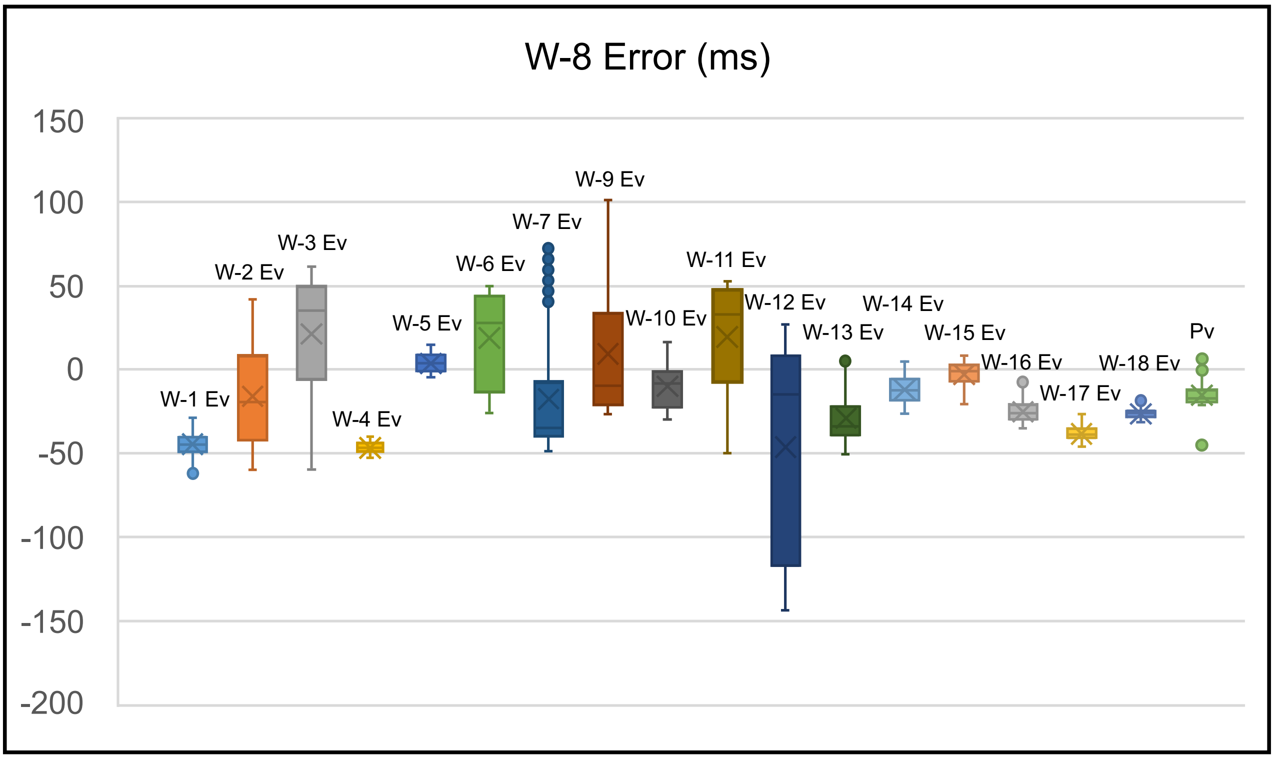


W-5 Error (ms)
250
200
150
100
50
0
-50
-100
-150
-200
W-1 Ev
W-2 Ev
W-3 Ev
W-4 Ev
W-6 Ev
W-7 Ev
W-8 Ev
W-9 Ev
W-10 Ev
W-11 Ev
W-12 Ev
W-13 Ev
W-14 Ev
W-15 Ev
W-16 Ev
W-17 Ev
W-18 Ev
Pv

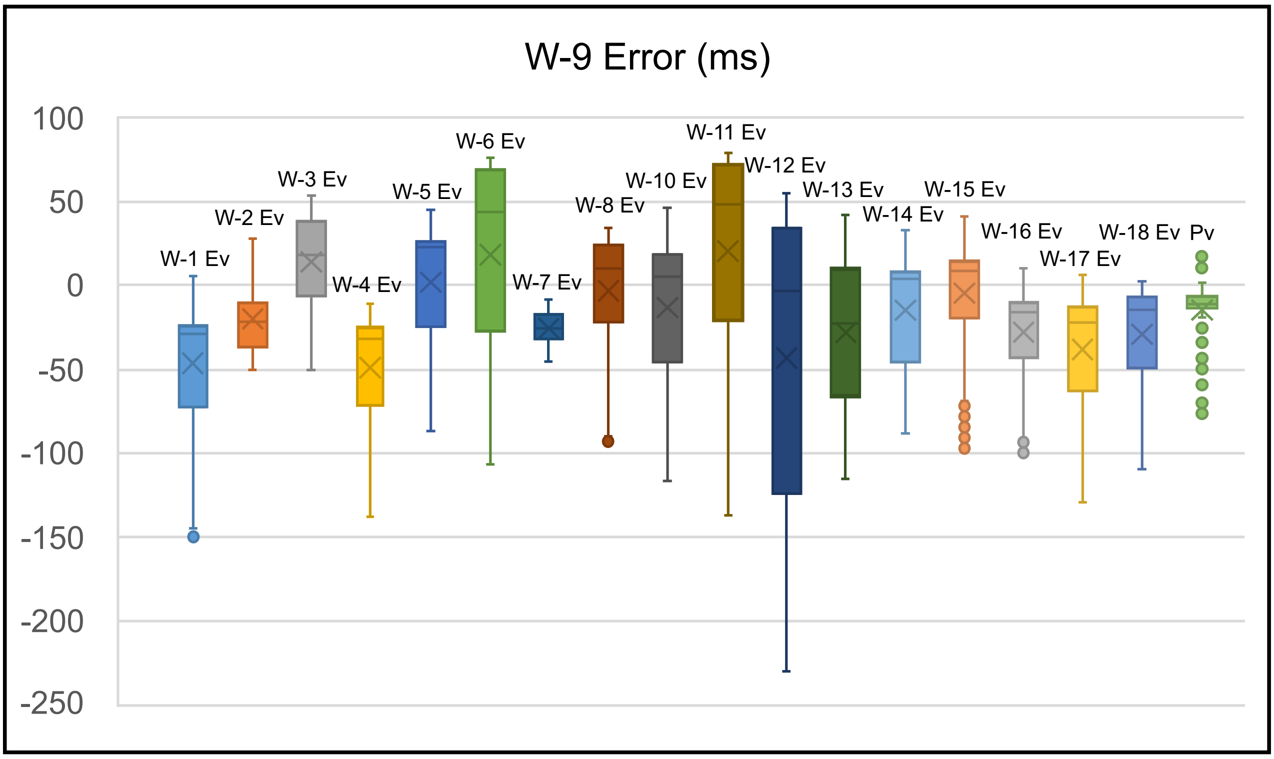

W-6 Error (ms)
300
200
100
0
-100
-200
-300
W-1 Ev
W-2 Ev
W-3 Ev
W-4 Ev
W-5 Ev
W-7 Ev
W-8 Ev
W-9 Ev
W-10 Ev
W-11 Ev
W-12 Ev
W-13 Ev
W-14 Ev
W-15 Ev
W-16 Ev
W-17 Ev
W-18 Ev
Pv

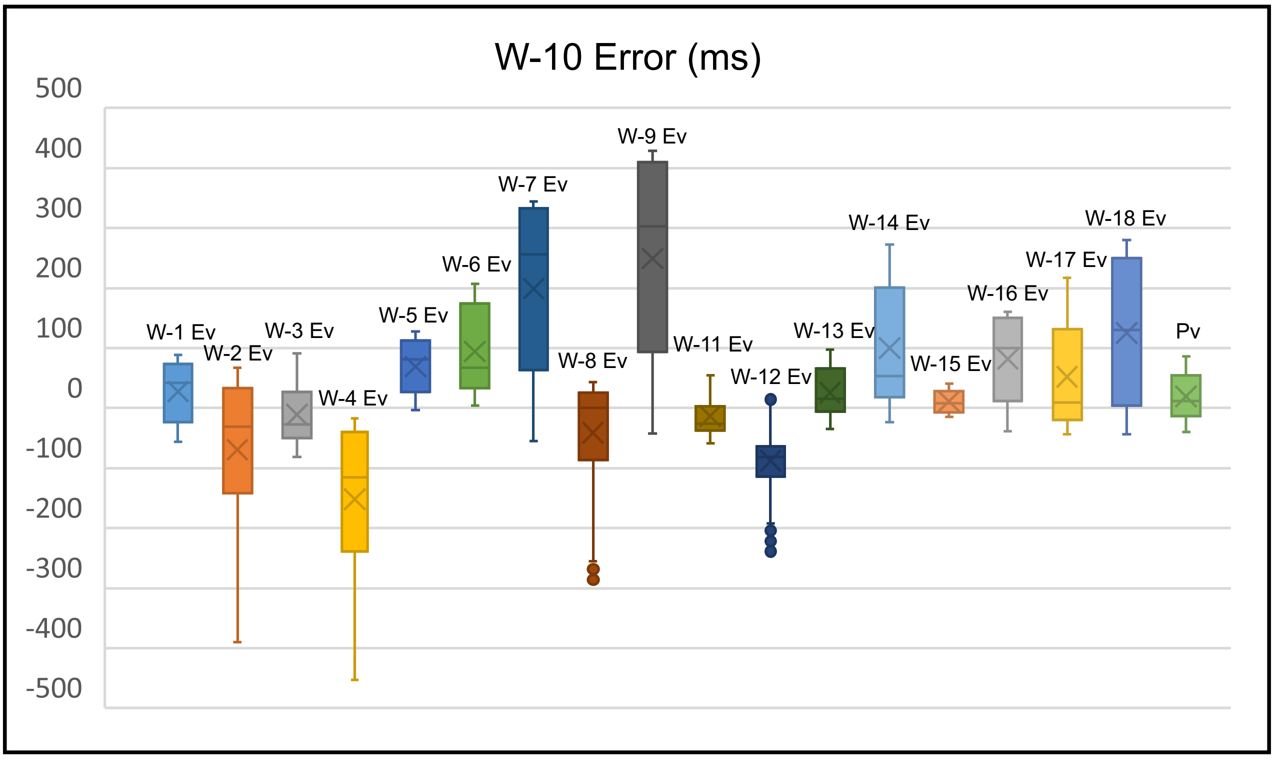


W-7 Error (ms)
150
100
50
0
-50
-100
-150
-200
-250
W-1 Ev
W-2 Ev
W-3 Ev
W-4 Ev
W-5 Ev
W-6 Ev
W-8 Ev
W-9 Ev
W-10 Ev
W-11 Ev
W-12 Ev
W-13 Ev
W-14 Ev
W-15 Ev
W-16 Ev
W-17 Ev
W-18 Ev
Pv

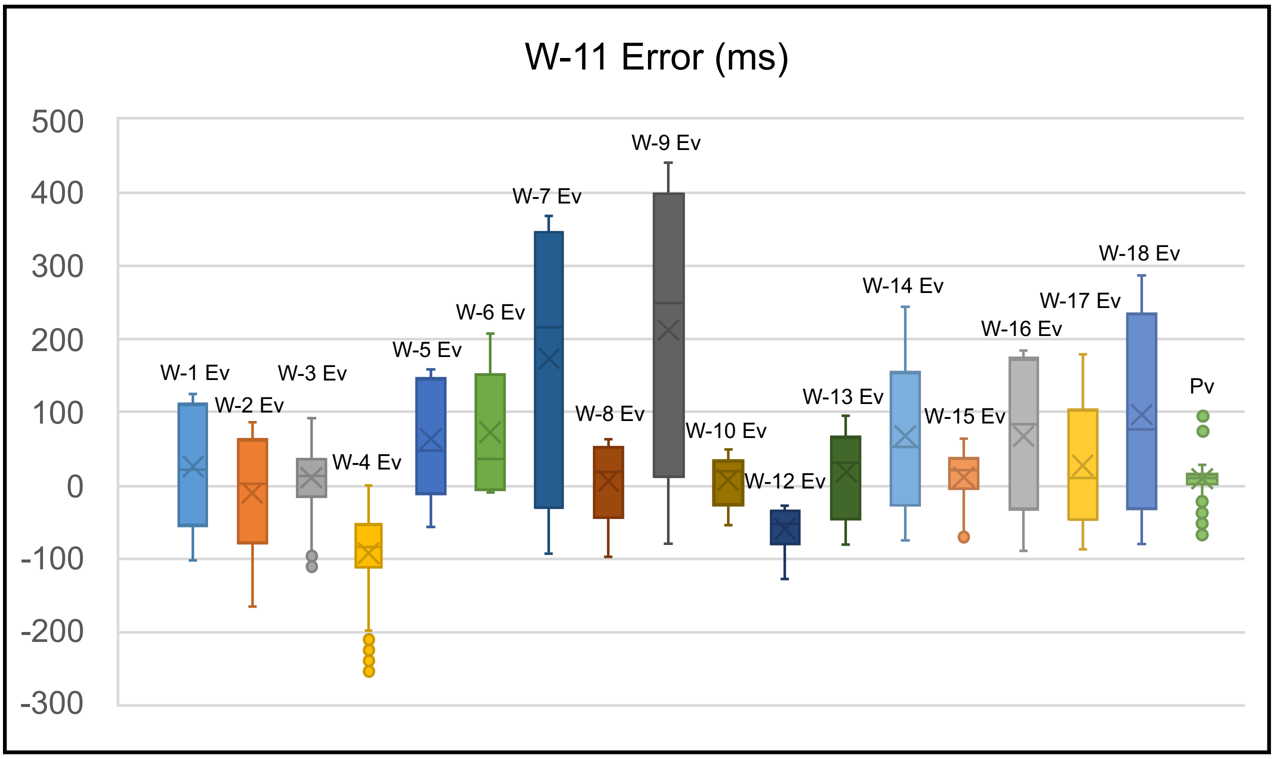


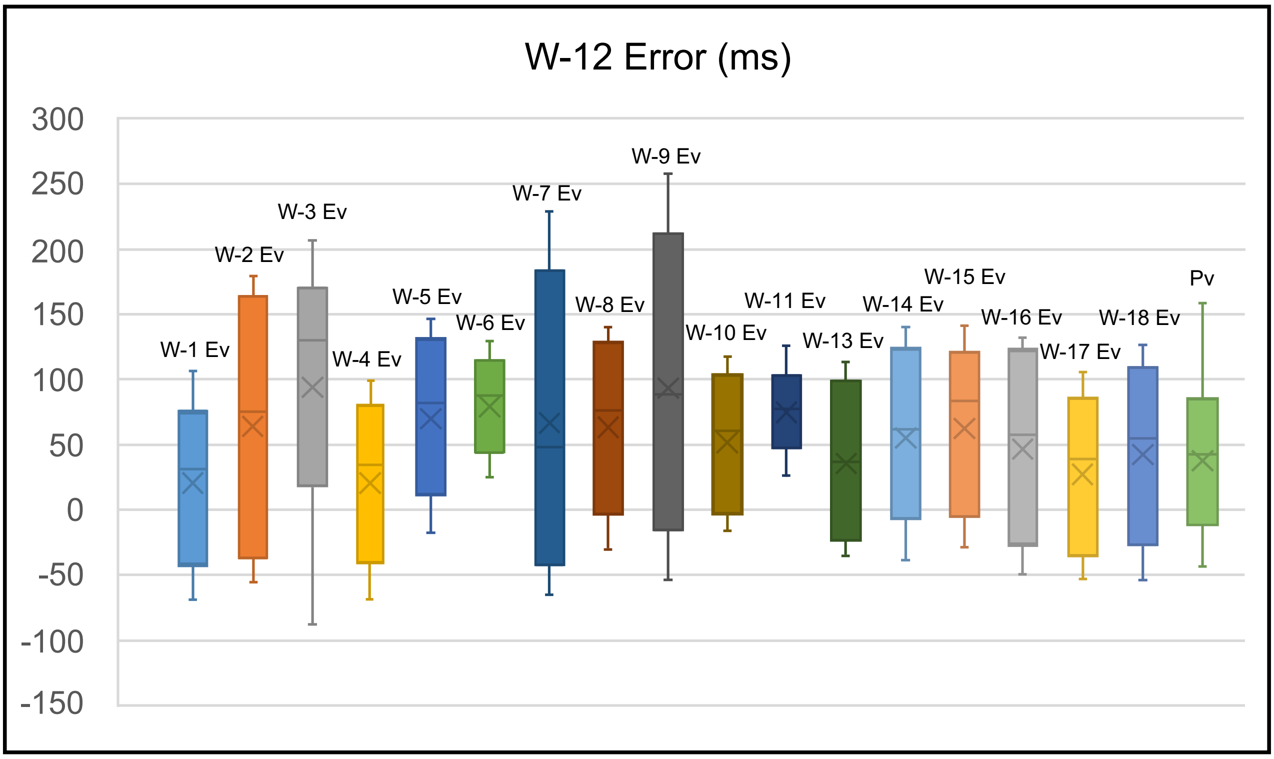
W-8 Error (ms)
150
100
50
0
-50
-100
-150
-200
W-1 Ev
W-2 Ev
W-3 Ev
W-4 Ev
W-5 Ev
W-6 Ev
W-7 Ev
W-9 Ev
W-10 Ev
W-11 Ev
W-12 Ev
W-13 Ev
W-14 Ev
W-15 Ev
W-16 Ev
W-17 Ev
W-18 Ev
Pv


W-9 Error (ms)
100
50
0
-50
-100
-150
-200
-250
W-1 Ev
W-2 Ev
W-3 Ev
W-4 Ev
W-5 Ev
W-6 Ev
W-7 Ev
W-8 Ev
W-10 Ev
W-11 Ev
W-12 Ev
W-13 Ev
W-14 Ev
W-15 Ev
W-16 Ev
W-17 Ev
W-18 Ev
Pv

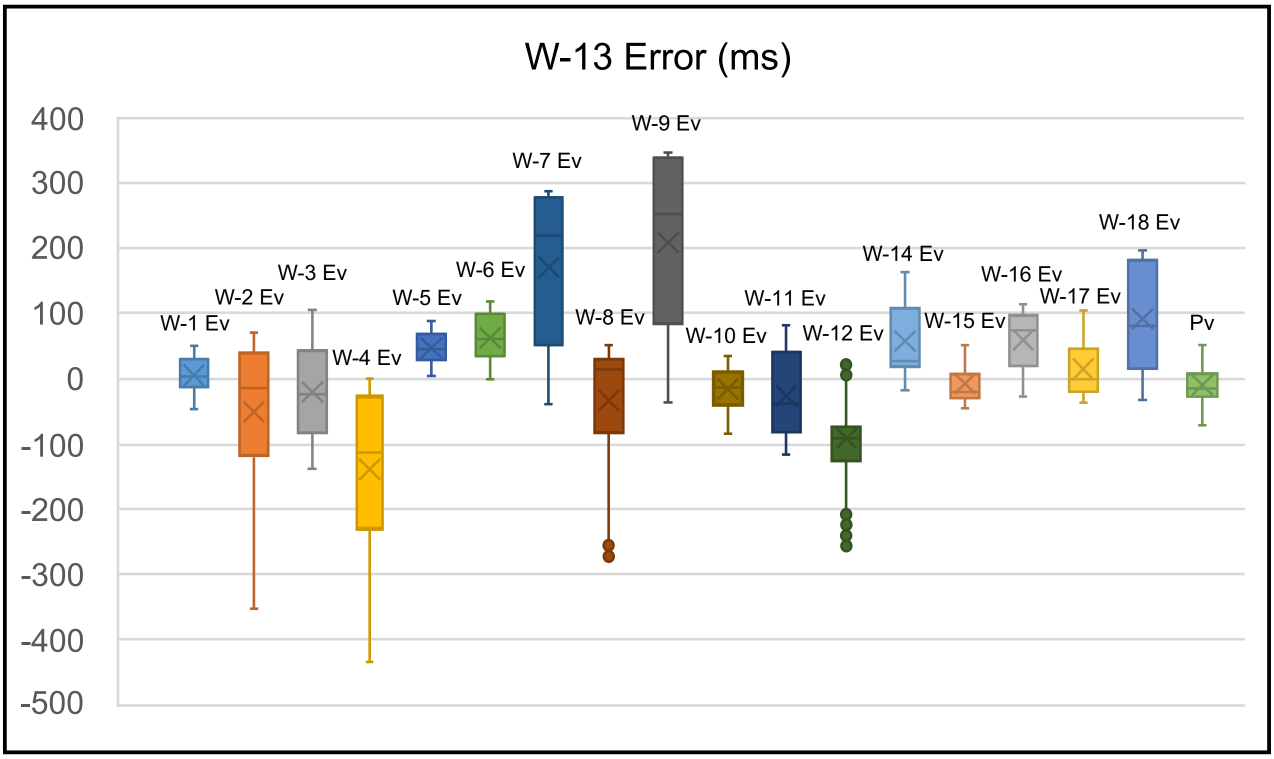


W-10 Error (ms)
500
400
300
200
100
0
-100
-200
-300
-400
-500
W-1 Ev
W-2 Ev
W-3 Ev
W-4 Ev
W-5 Ev
W-6 Ev
W-7 Ev
W-8 Ev
W-9 Ev
W-11 Ev
W-12 Ev
W-13 Ev
W-14 Ev
W-15 Ev
W-16 Ev
W-17 Ev
W-18 Ev
Pv

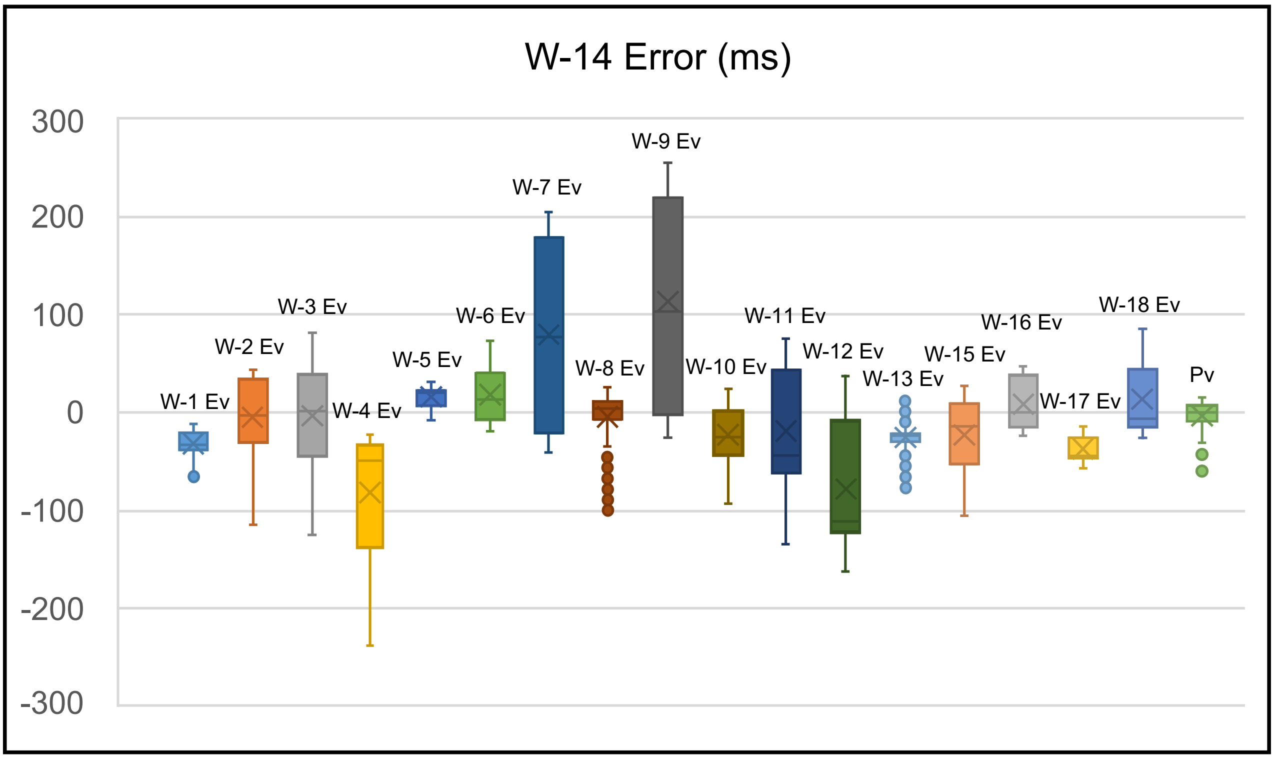

W-11 Error (ms)
500
400
300
200
100
0
-100
-200
-300
W-1 Ev
W-2 Ev
W-3 Ev
W-4 Ev
W-5 Ev
W-6 Ev
W-7 Ev
W-8 Ev
W-9 Ev
W-10 Ev
W-12 Ev
W-13 Ev
W-14 Ev
W-15 Ev
W-16 Ev
W-17 Ev
W-18 Ev
Pv



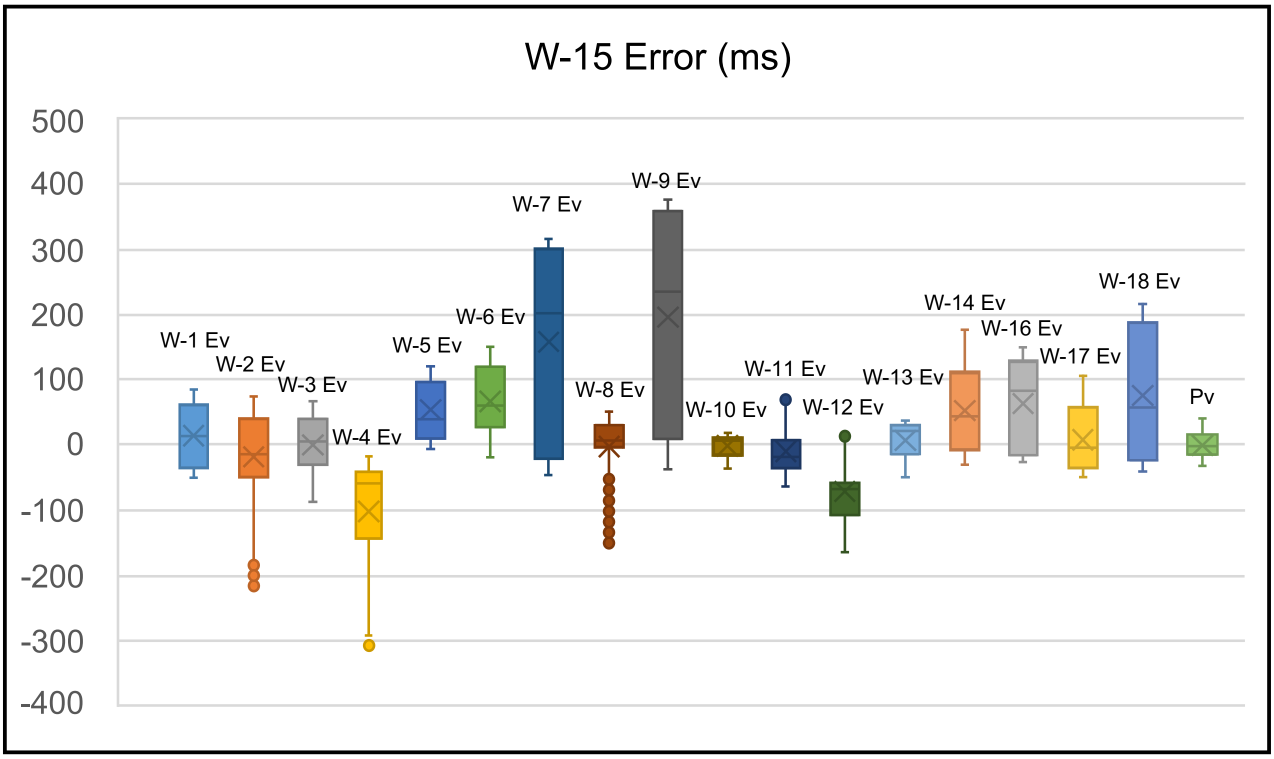
W-12 Error (ms)
300
250
200
150
100
50
0
-50
-100
-150
W-1 Ev
W-2 Ev
W-3 Ev
W-4 Ev
W-5 Ev
W-6 Ev
W-7 Ev
W-8 Ev
W-9 Ev
W-10 Ev
W-11 Ev
W-13 Ev
W-14 Ev
W-15 Ev
W-16 Ev
W-17 Ev
W-18 Ev
Pv



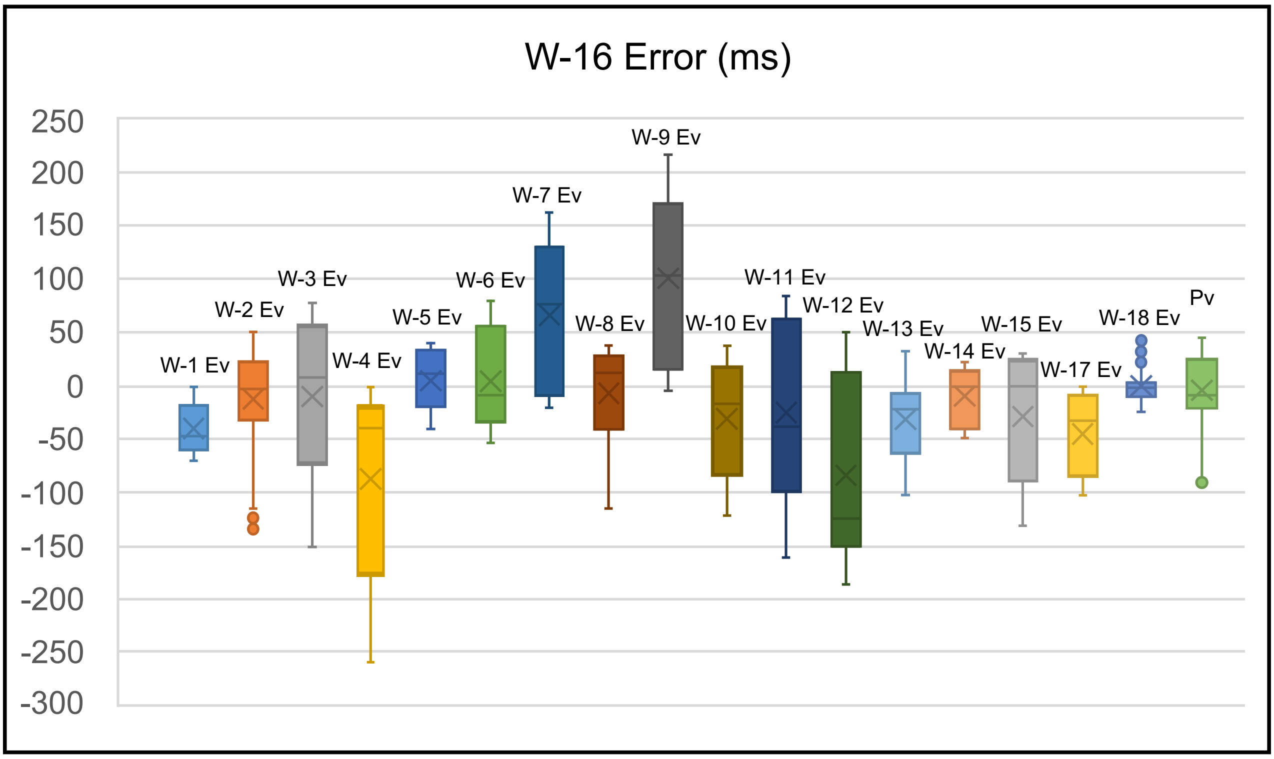
W-13 Error (ms)
400
300
200
100
0
-100
-200
-300
-400
-500
W-1 Ev
W-2 Ev
W-3 Ev
W-4 Ev
W-5 Ev
W-6 Ev
W-7 Ev
W-8 Ev
W-9 Ev
W-10 Ev
W-11 Ev
W-12 Ev
W-14 Ev
W-15 Ev
W-16 Ev
W-17 Ev
W-18 Ev
Pv



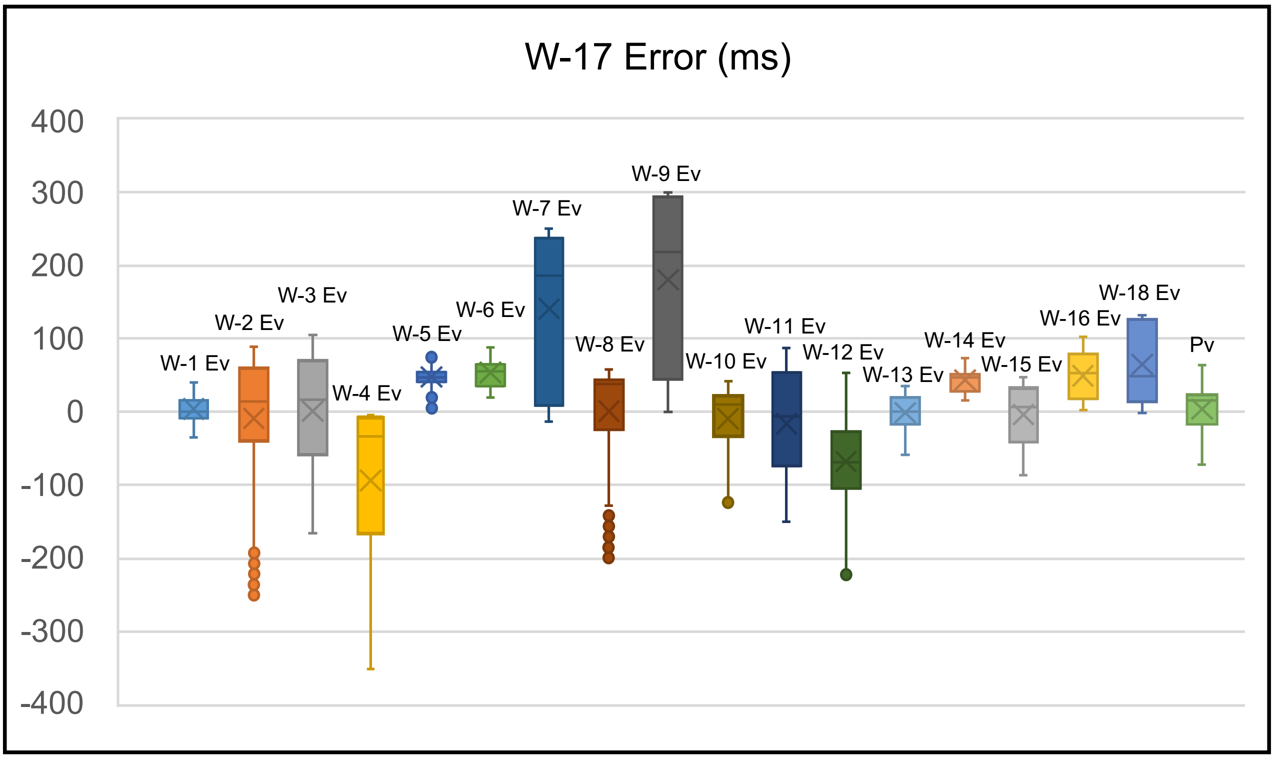
W-14 Error (ms)
300
200
100
0
-100
-200
-300
W-1 Ev
W-2 Ev
W-3 Ev
W-4 Ev
W-5 Ev
W-6 Ev
W-7 Ev
W-8 Ev
W-9 Ev
W-10 Ev
W-11 Ev
W-12 Ev
W-13 Ev
W-15 Ev
W-16 Ev
W-17 Ev
W-18 Ev
Pv



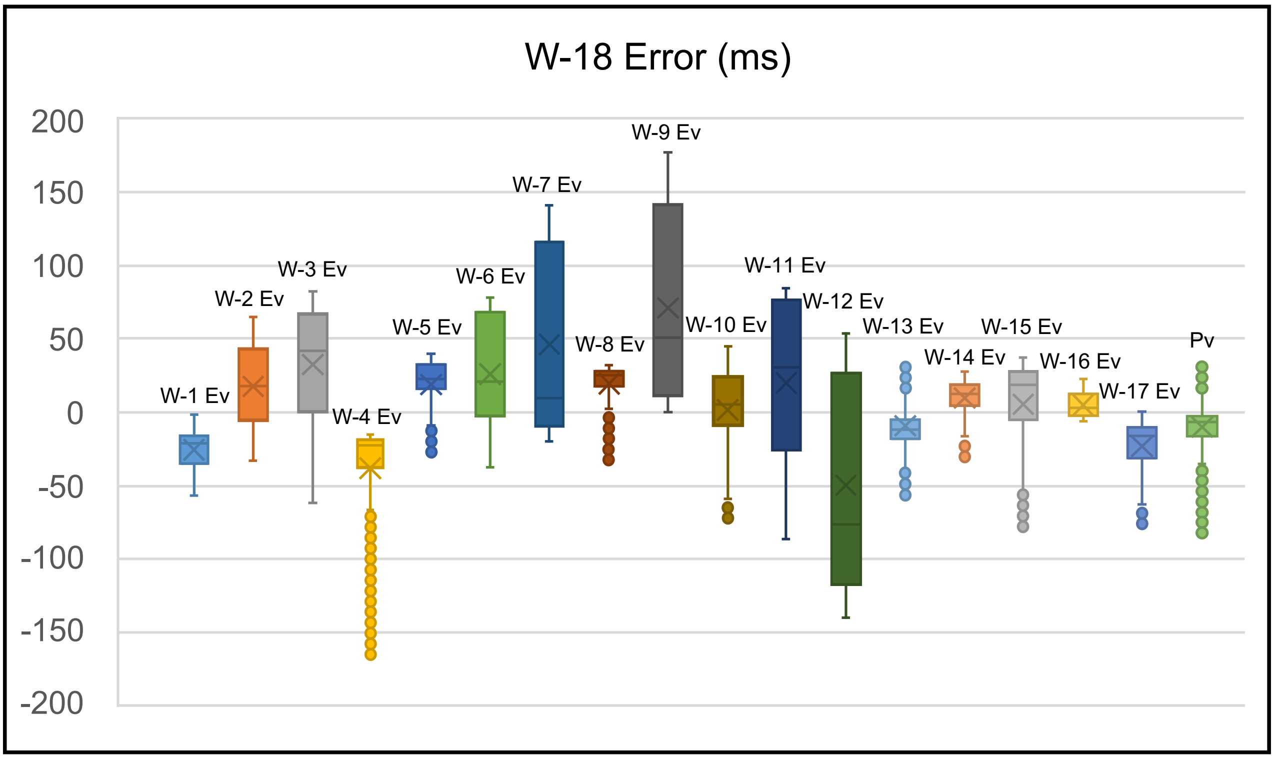
W-15 Error (ms)
500
400
300
200
100
0
-100
-200
-300
-400
W-1 Ev
W-2 Ev
W-3 Ev
W-4 Ev
W-5 Ev
W-6 Ev
W-7 Ev
W-8 Ev
W-9 Ev
W-10 Ev
W-11 Ev
W-12 Ev
W-13 Ev
W-14 Ev
W-16 Ev
W-17 Ev
W-18 Ev
Pv



W-16 Error (ms)
250
200
150
100
50
0
-50
-100
-150
-200
-250
-300
W-1 Ev
W-2 Ev
W-3 Ev
W-4 Ev
W-5 Ev
W-6 Ev
W-7 Ev
W-8 Ev
W-9 Ev
W-10 Ev
W-11 Ev
W-12 Ev
W-13 Ev
W-14 Ev
W-15 Ev
W-17 Ev
W-18 Ev
Pv

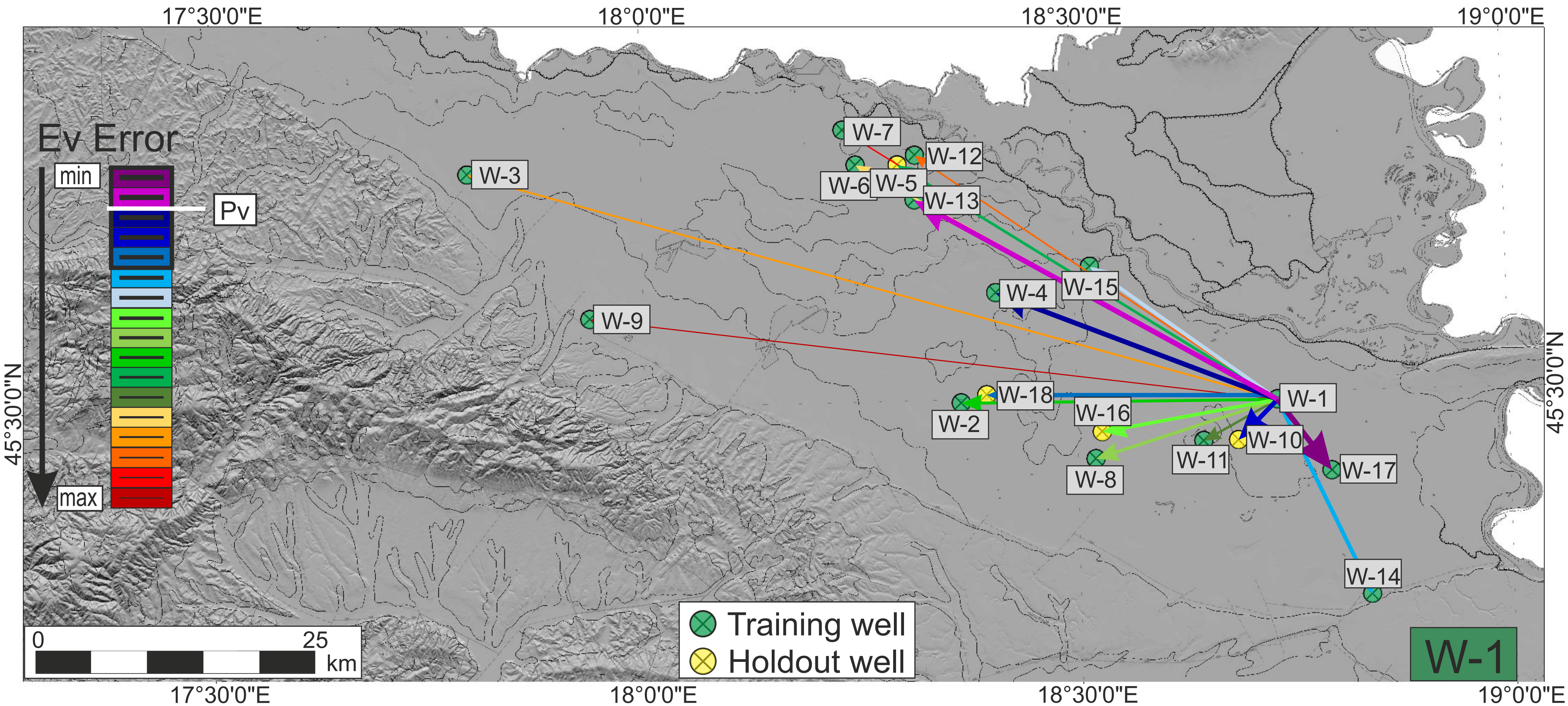


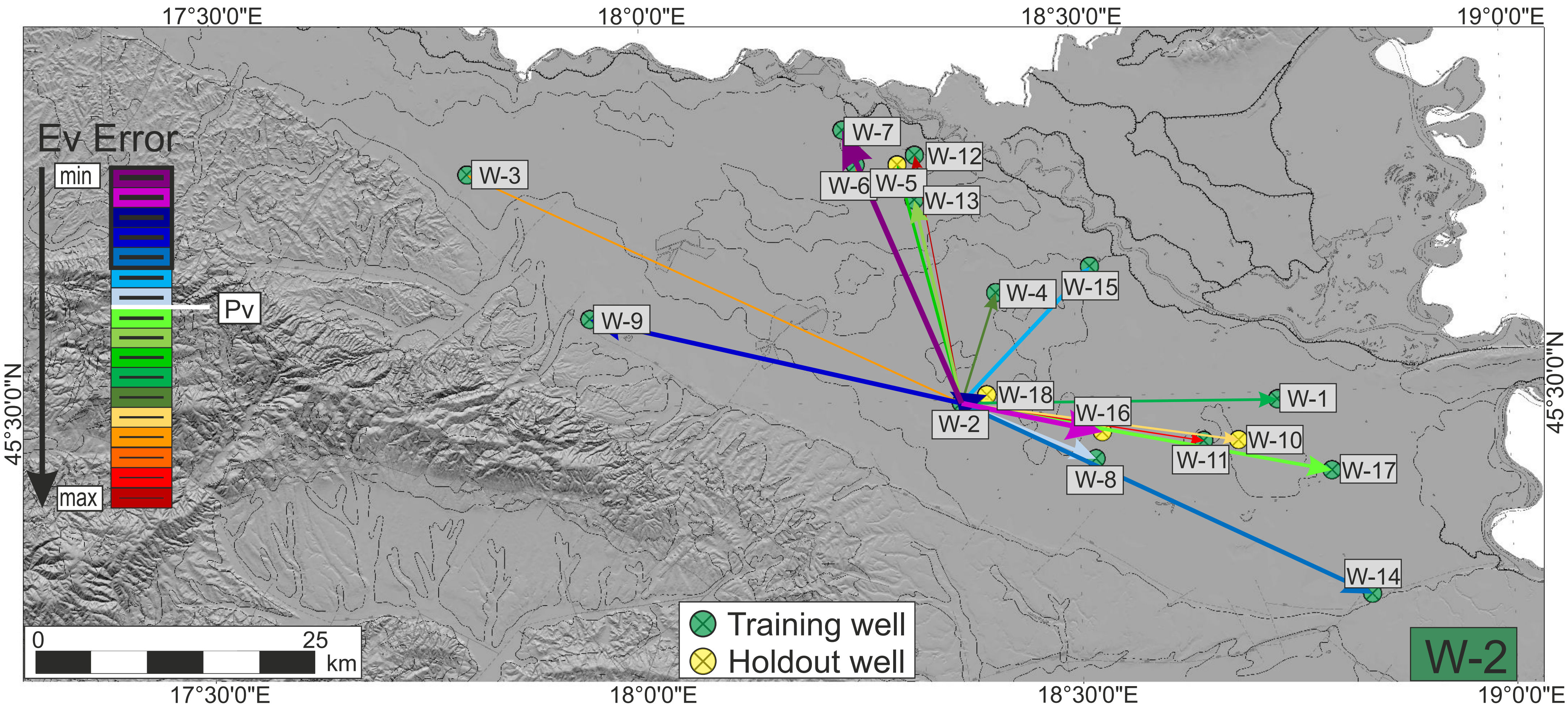
W-17 Error (ms)
400
300
200
100
0
-100
-200
-300
-400
W-1 Ev
W-2 Ev
W-3 Ev
W-4 Ev
W-5 Ev
W-6 Ev
W-7 Ev
W-8 Ev
W-9 Ev
W-10 Ev
W-11 Ev
W-12 Ev
W-13 Ev
W-14 Ev
W-15 Ev
W-16 Ev
W-18 Ev
Pv



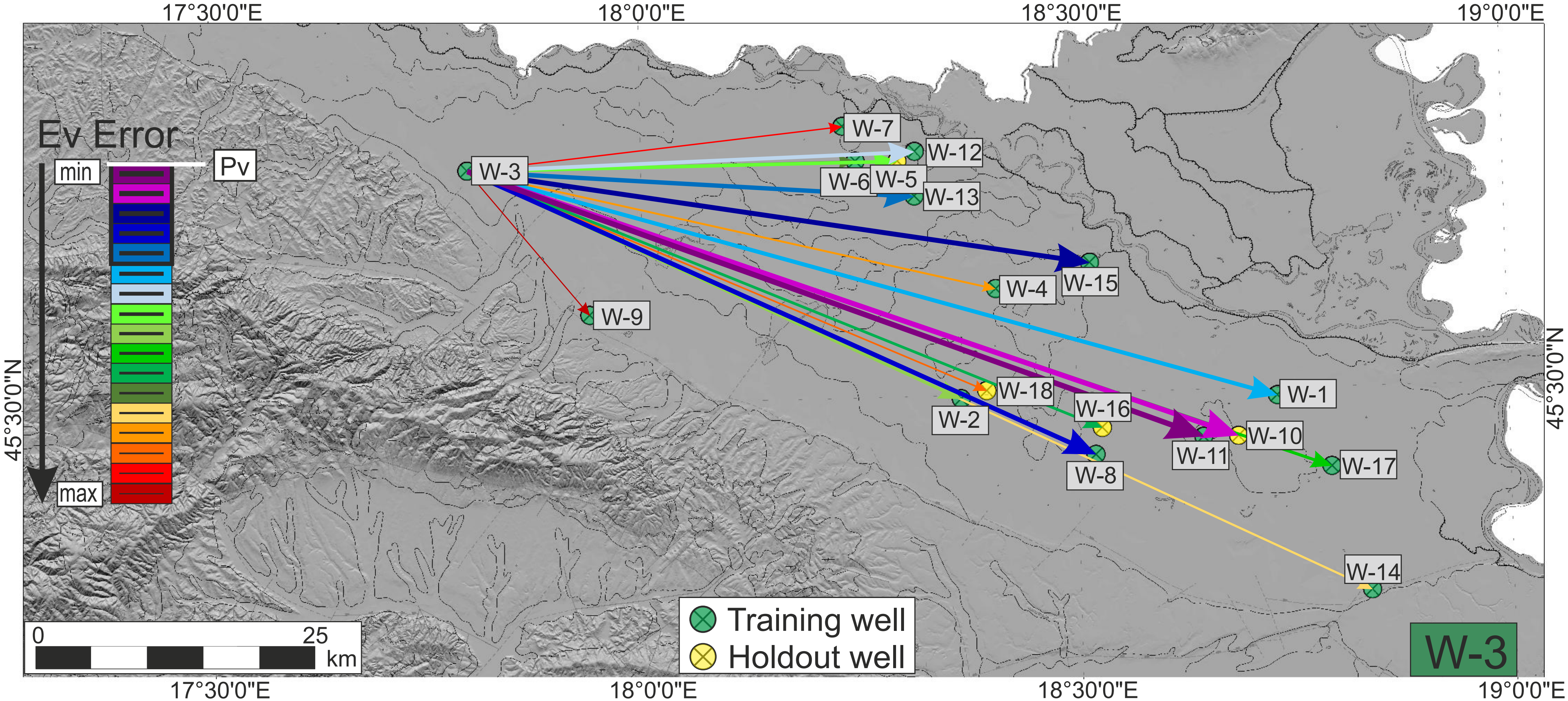
W-18 Error (ms)
200
150
100
50
0
-50
-100
-150
-200
W-1 Ev
W-2 Ev
W-3 Ev
W-4 Ev
W-5 Ev
W-6 Ev
W-7 Ev
W-8 Ev
W-9 Ev
W-10 Ev
W-11 Ev
W-12 Ev
W-13 Ev
W-14 Ev
W-15 Ev
W-16 Ev
W-17 Ev
Pv

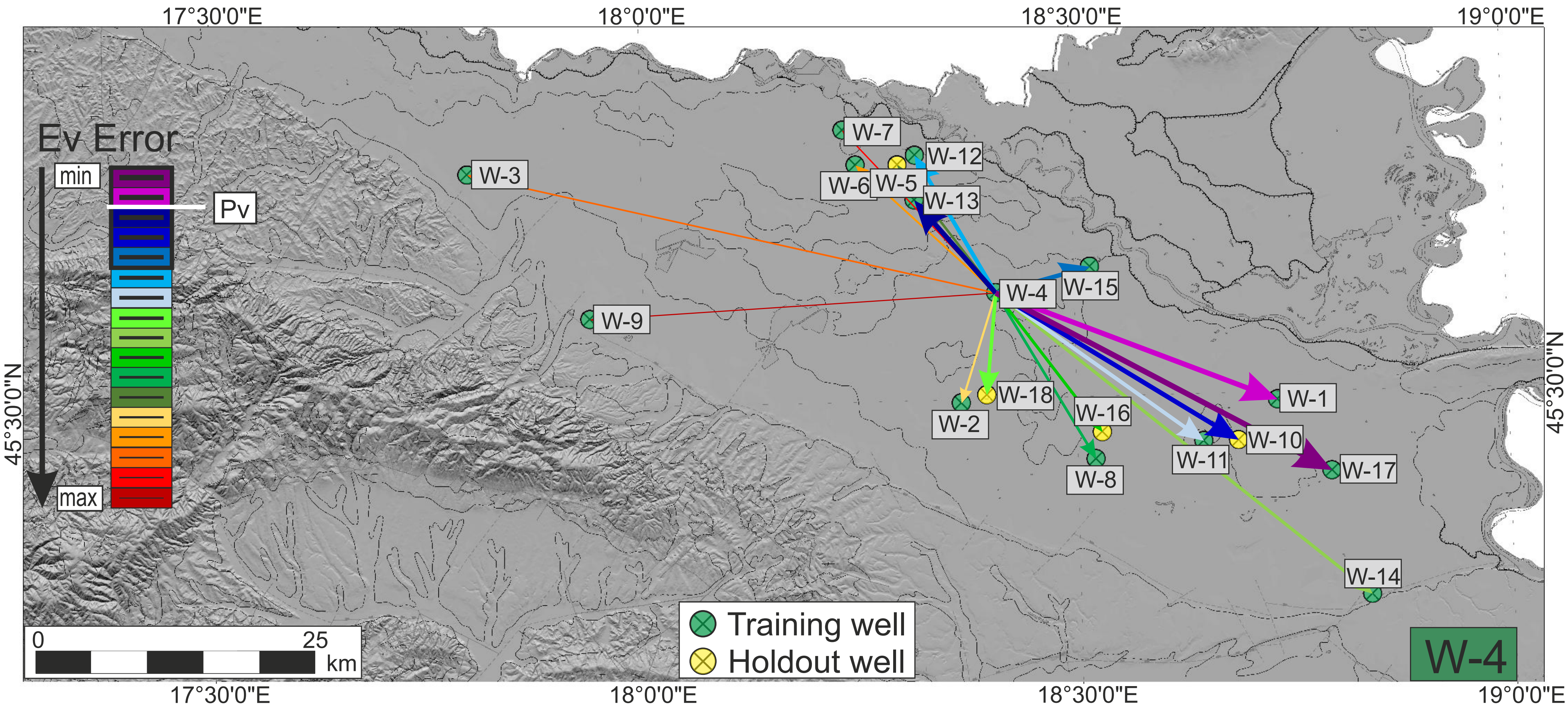
Appendix 3


17°30'0"E
18°0'0"E
18°30'0"E
19°0'0"E
45°30'0"N
Ev Error
min
Pv
max
W-7
W-12
W-3
W-6
W-5
W-13
W-15
W-4
W-9
W-18
W-1
W-2
W-16
W-10
W-11
W-17
W-8
W-14
Training well
Holdout well
0
25
km
W-1

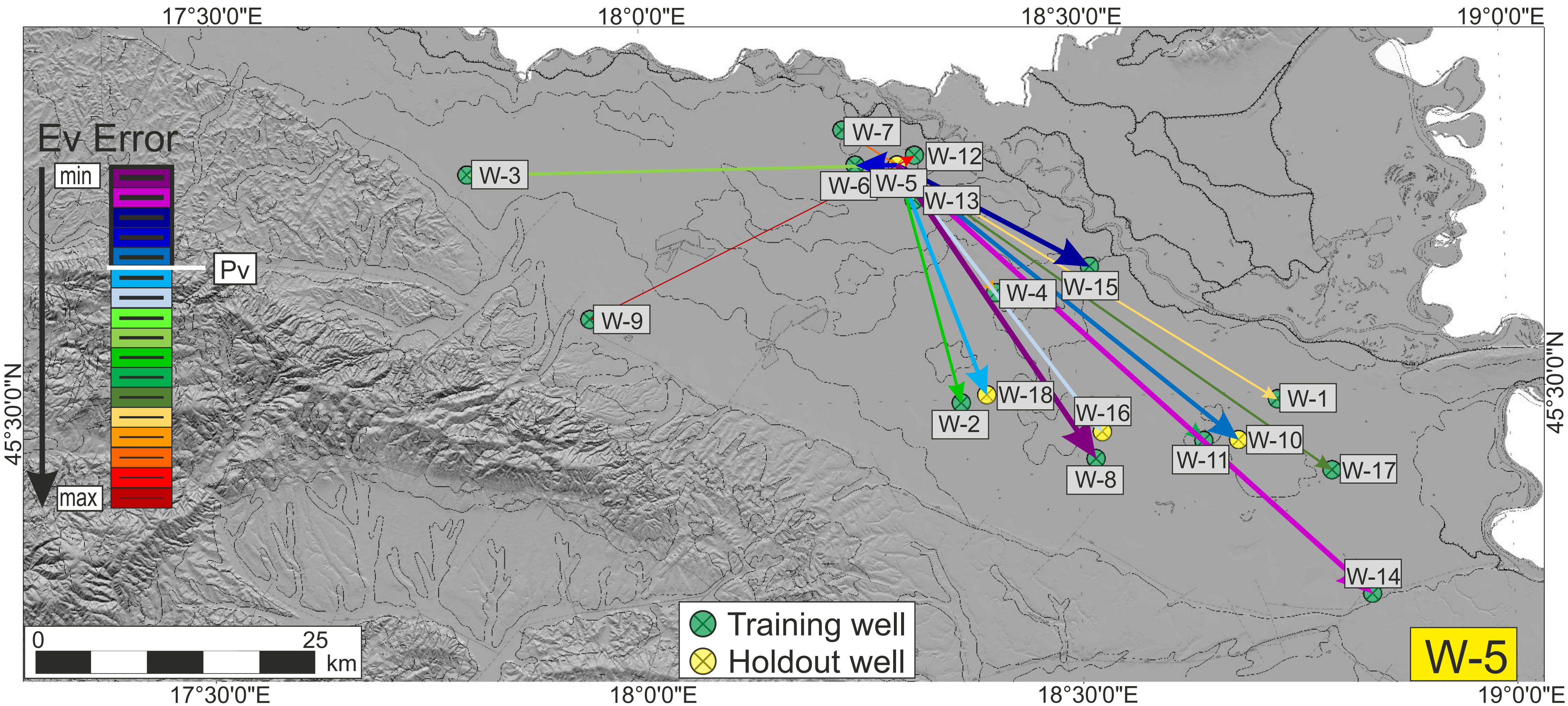


17°30'0"E
18°0'0"E
18°30'0"E
19°0'0"E
45°30'0"N
Ev Error
min
max
Pv
W-1
W-2
W-3
W-4
W-5
W-6
W-7
W-8
W-9
W-10
W-11
W-12
W-13
W-14
W-15
W-16
W-17
W-18
Training well
Holdout well
0
25
km
W-2

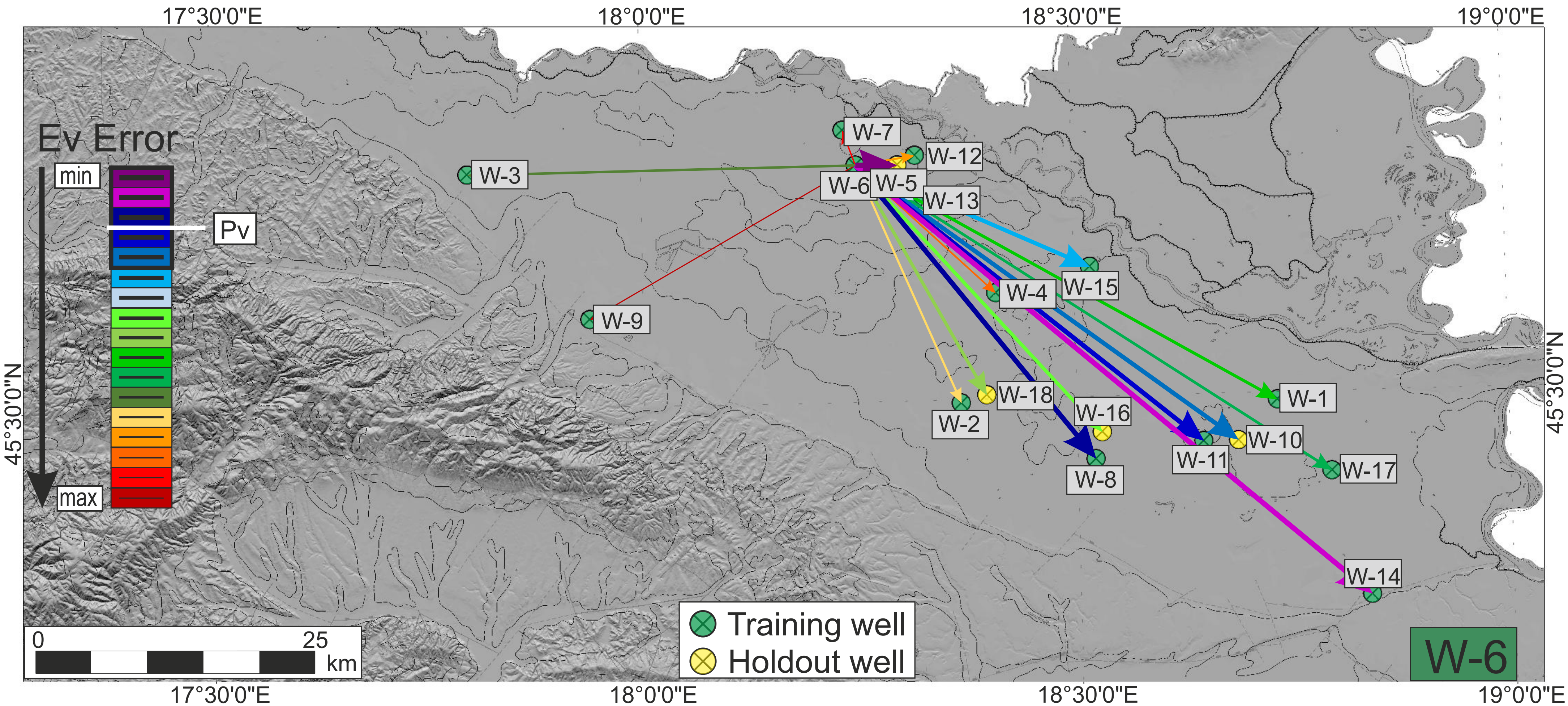

17°30'0"E
18°0'0"E
18°30'0"E
19°0'0"E
45°30'0"N
Ev Error
min
max
Pv
W-1
W-2
W-3
W-4
W-5
W-6
W-7
W-8
W-9
W-10
W-11
W-12
W-13
W-14
W-15
W-16
W-17
W-18
Training well
Holdout well
0
25
km
W-3

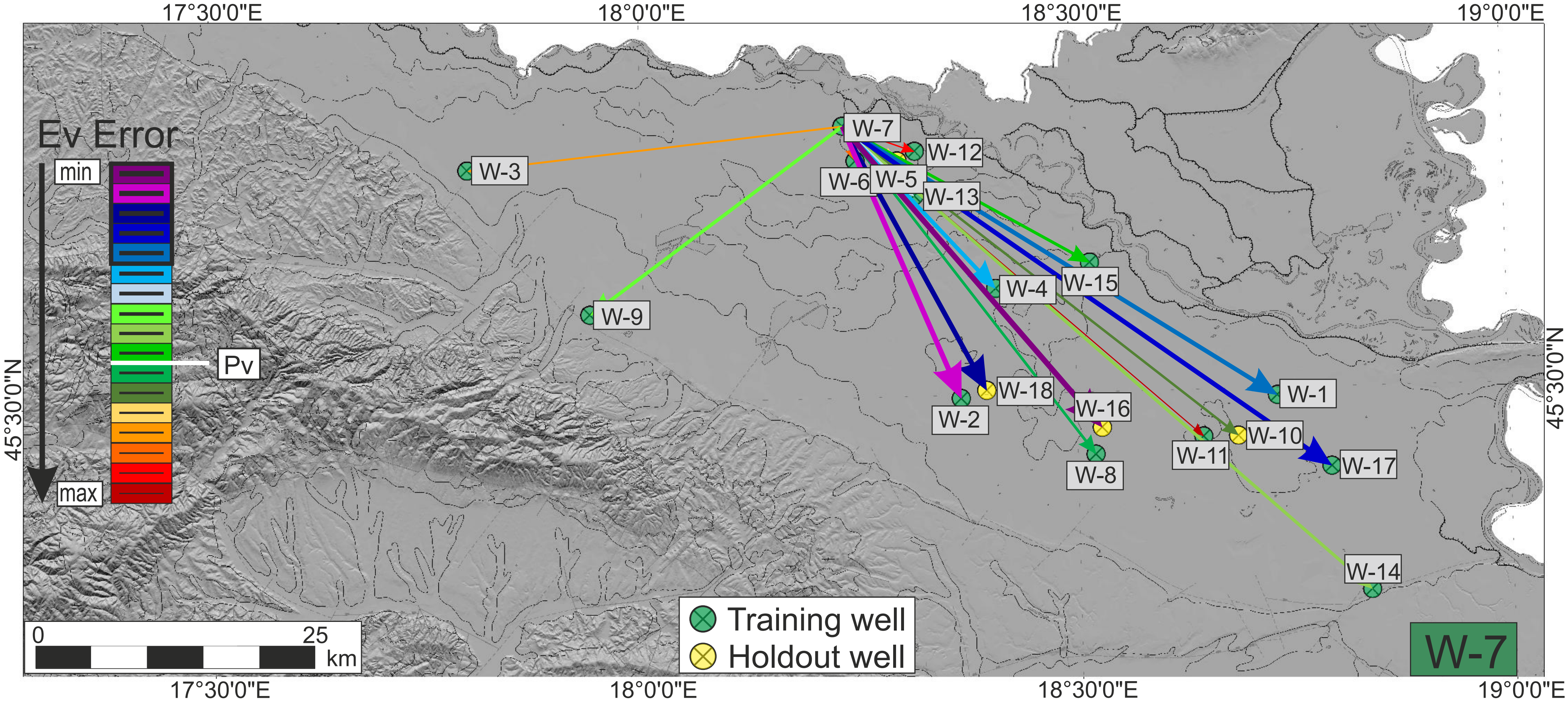


17°30'0"E
18°0'0"E
18°30'0"E
19°0'0"E
45°30'0"N
Ev Error
min
max
Pv
W-1
W-2
W-3
W-4
W-5
W-6
W-7
W-8
W-9
W-10
W-11
W-12
W-13
W-14
W-15
W-16
W-17
W-18
Training well
Holdout well
0
25
km
W-4

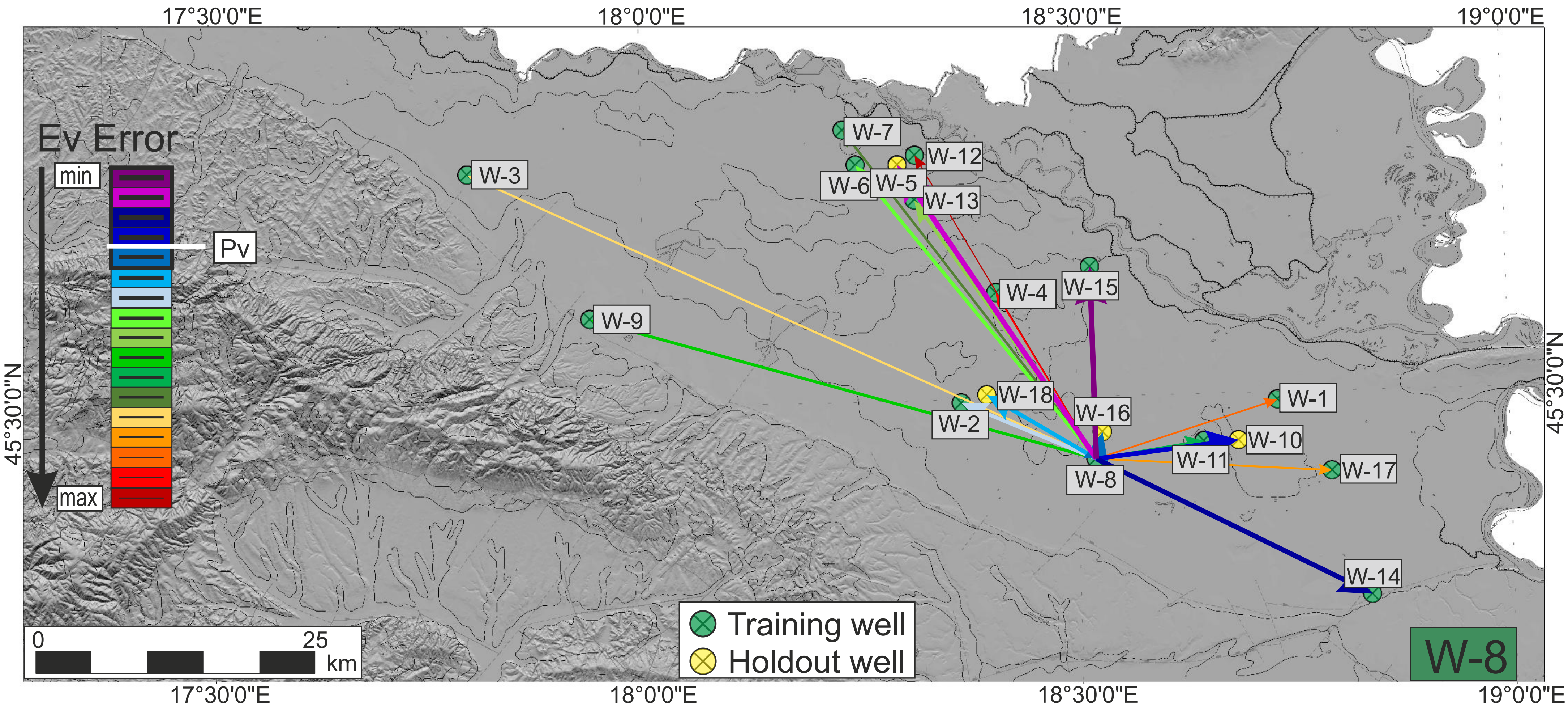


17°30'0"E
18°0'0"E
18°30'0"E
19°0'0"E
45°30'0"N
Ev Error
min
Pv
max
W-1
W-2
W-3
W-4
W-5
W-6
W-7
W-8
W-9
W-10
W-11
W-12
W-13
W-14
W-15
W-16
W-17
W-18
Training well
Holdout well
0
25
km
W-5

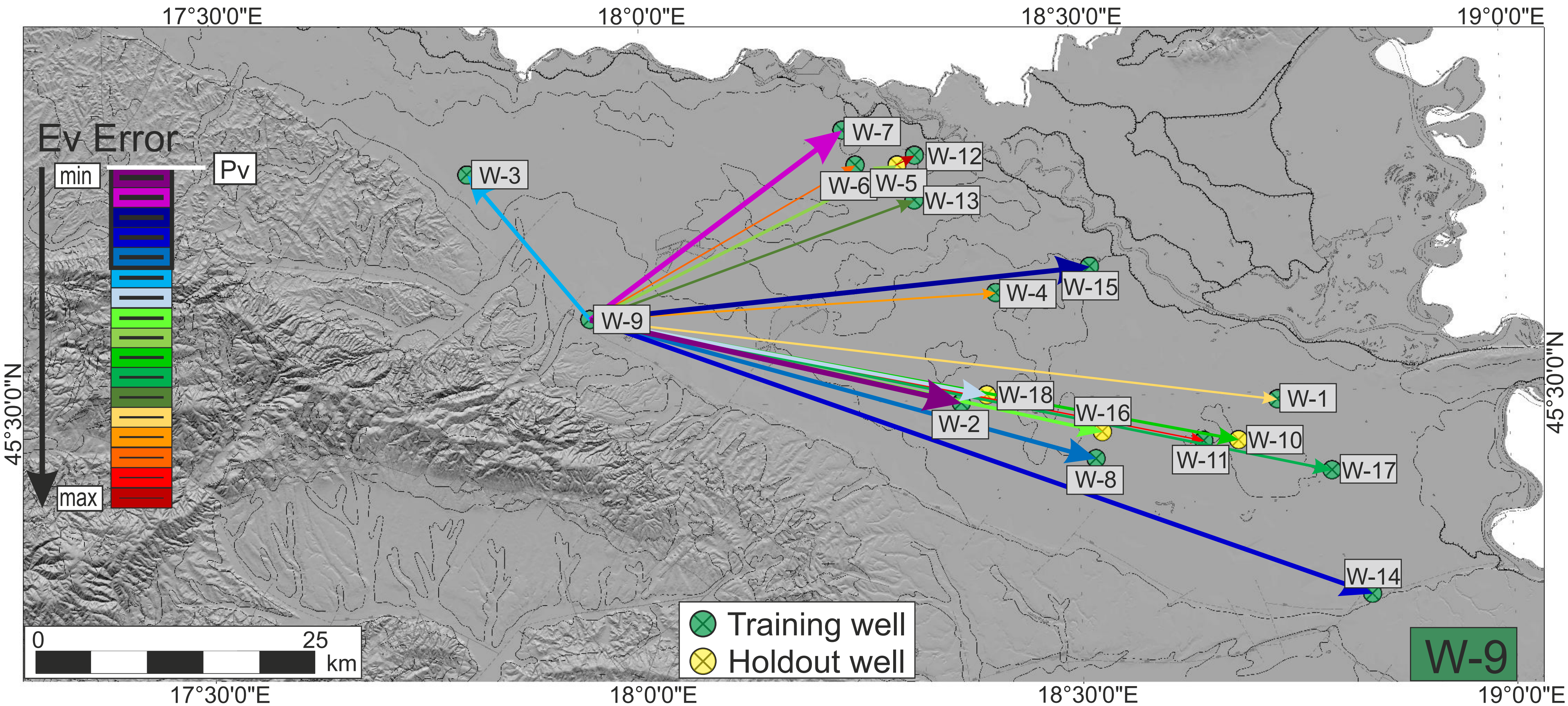

17°30'0"E
18°0'0"E
18°30'0"E
19°0'0"E
45°30'0"N
Ev Error
min
Pv
max
W-7
W-12
W-3
W-6
W-5
W-13
W-15
W-4
W-9
W-1
W-18
W-2
W-16
W-10
W-11
W-17
W-8
W-14
0
25
km
Training well
Holdout well
W-6



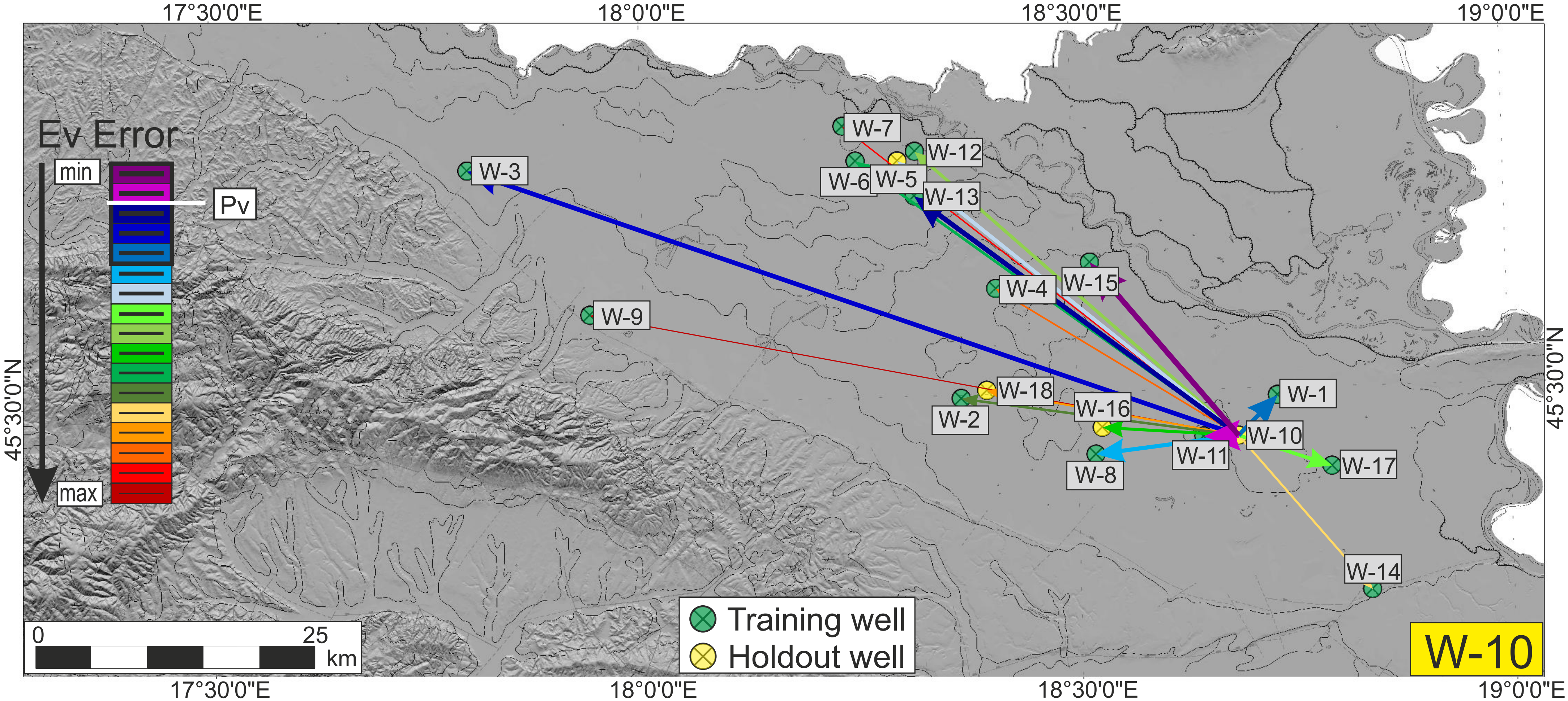
17°30'0"E
18°0'0"E
18°30'0"E
19°0'0"E
45°30'0"N
Ev Error
min
max
Pv
W-7
W-12
W-3
W-6
W-5
W-13
W-15
W-4
W-9
W-1
W-18
W-2
W-16
W-10
W-11
W-17
W-8
W-14
Training well
Holdout well
0
25
km
W-7



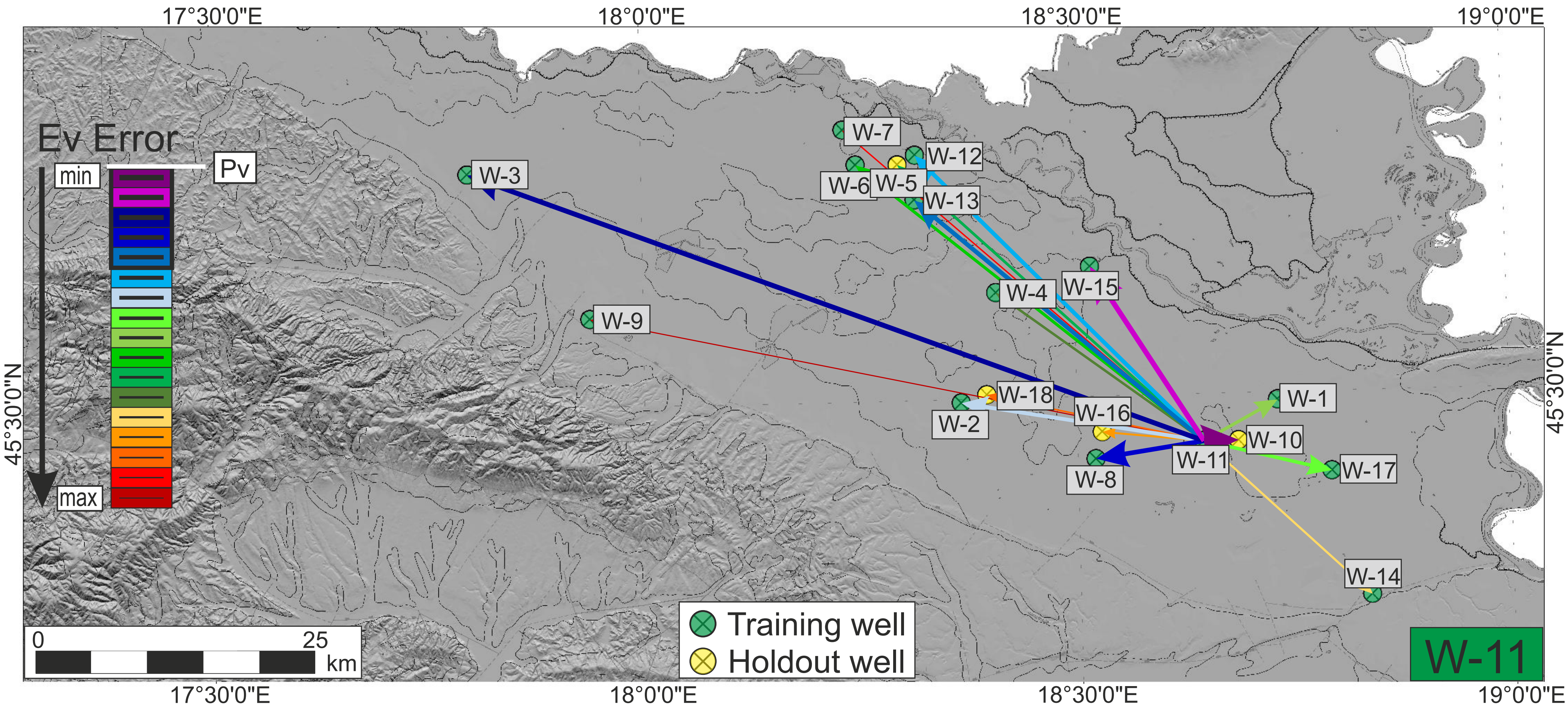
17°30'0"E
18°0'0"E
18°30'0"E
19°0'0"E
45°30'0"N
Ev Error
min
Pv
max
W-1
W-2
W-3
W-4
W-5
W-6
W-7
W-8
W-9
W-10
W-11
W-12
W-13
W-14
W-15
W-16
W-17
W-18
0
25
km
Training well
Holdout well
W-8



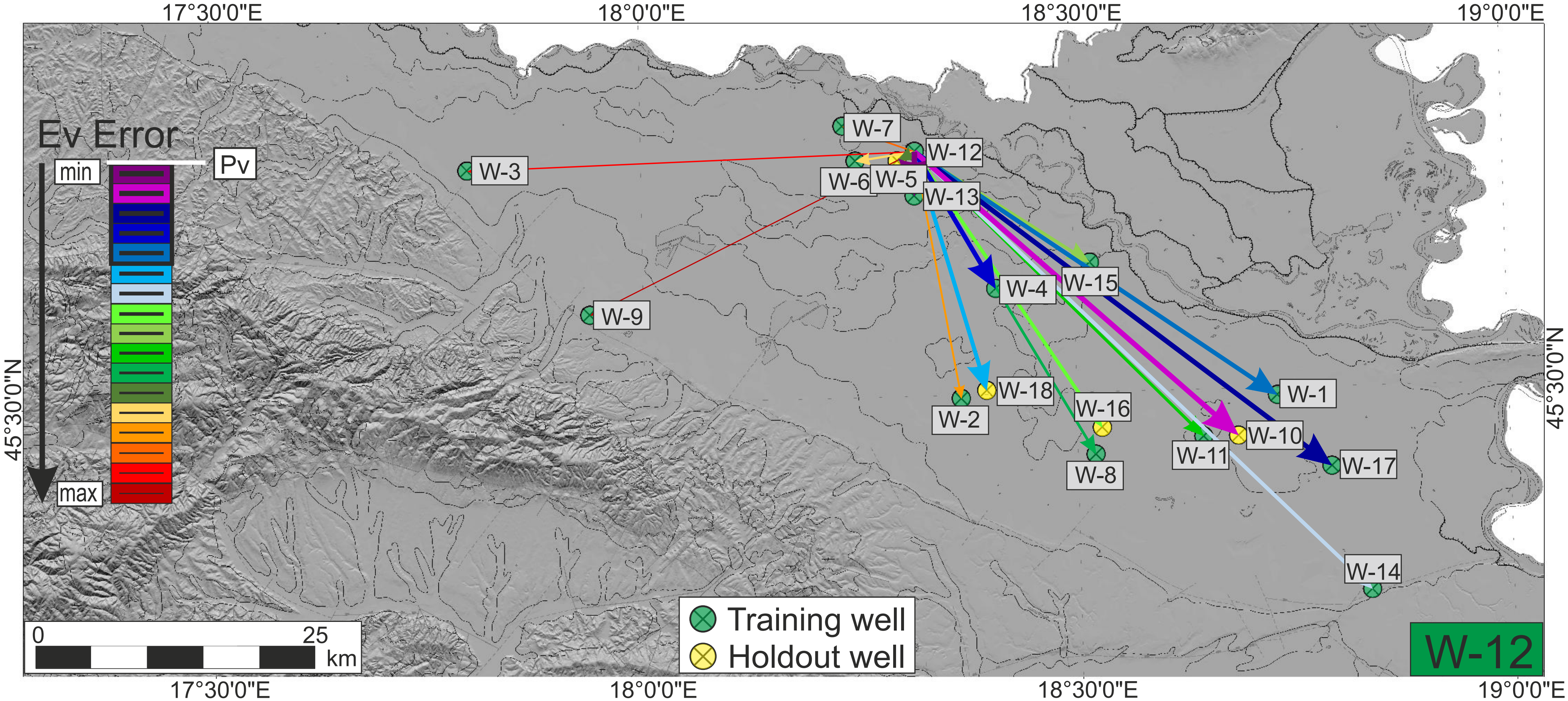
17°30'0"E
18°0'0"E
18°30'0"E
19°0'0"E
45°30'0"N
Ev Error
min
max
Pv
W-1
W-2
W-3
W-4
W-5
W-6
W-7
W-8
W-9
W-10
W-11
W-12
W-13
W-14
W-15
W-16
W-17
W-18
Training well
Holdout well
0
25
km
W-9



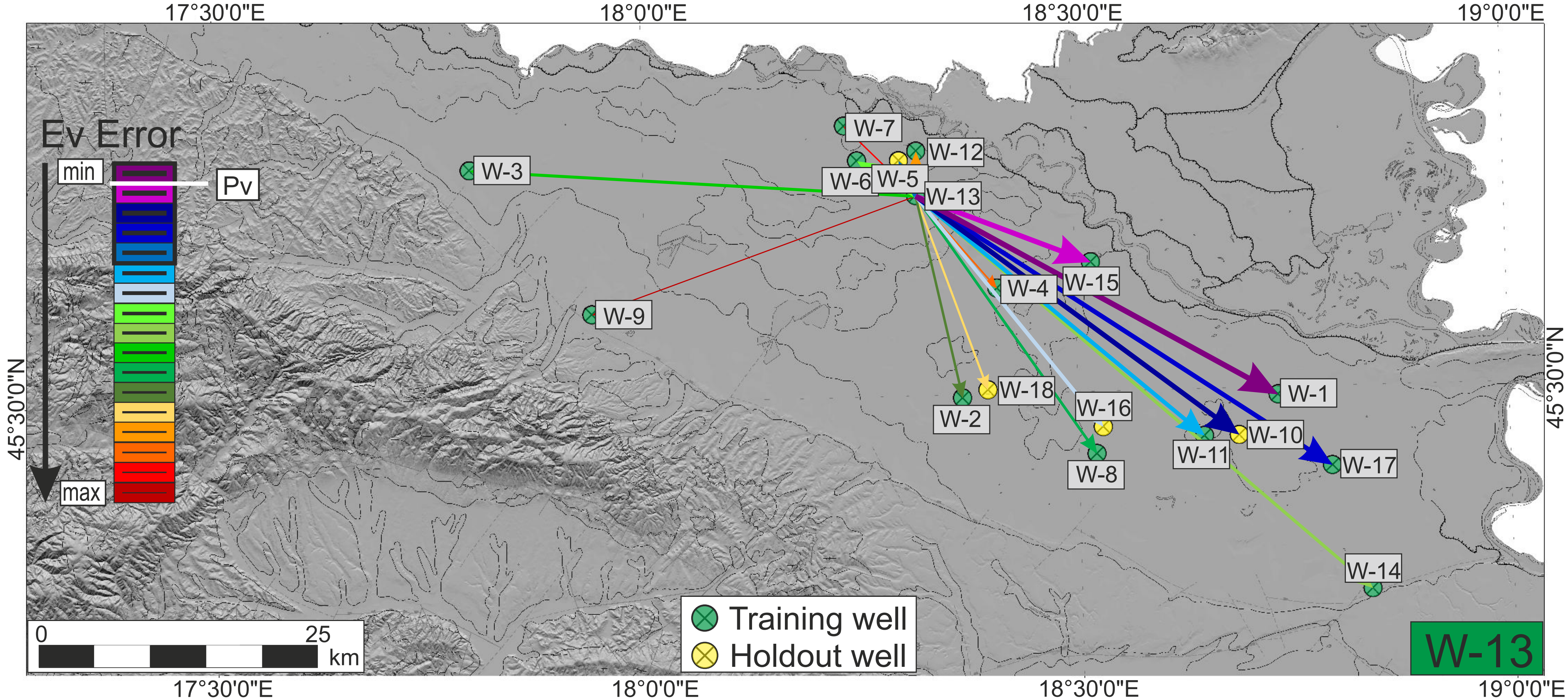
17°30'0"E
18°0'0"E
18°30'0"E
19°0'0"E
45°30'0"N
Ev Error
min
max
Pv
W-1
W-2
W-3
W-4
W-5
W-6
W-7
W-8
W-9
W-10
W-11
W-12
W-13
W-14
W-15
W-16
W-17
W-18
Training well
Holdout well
0
25
km
W-10



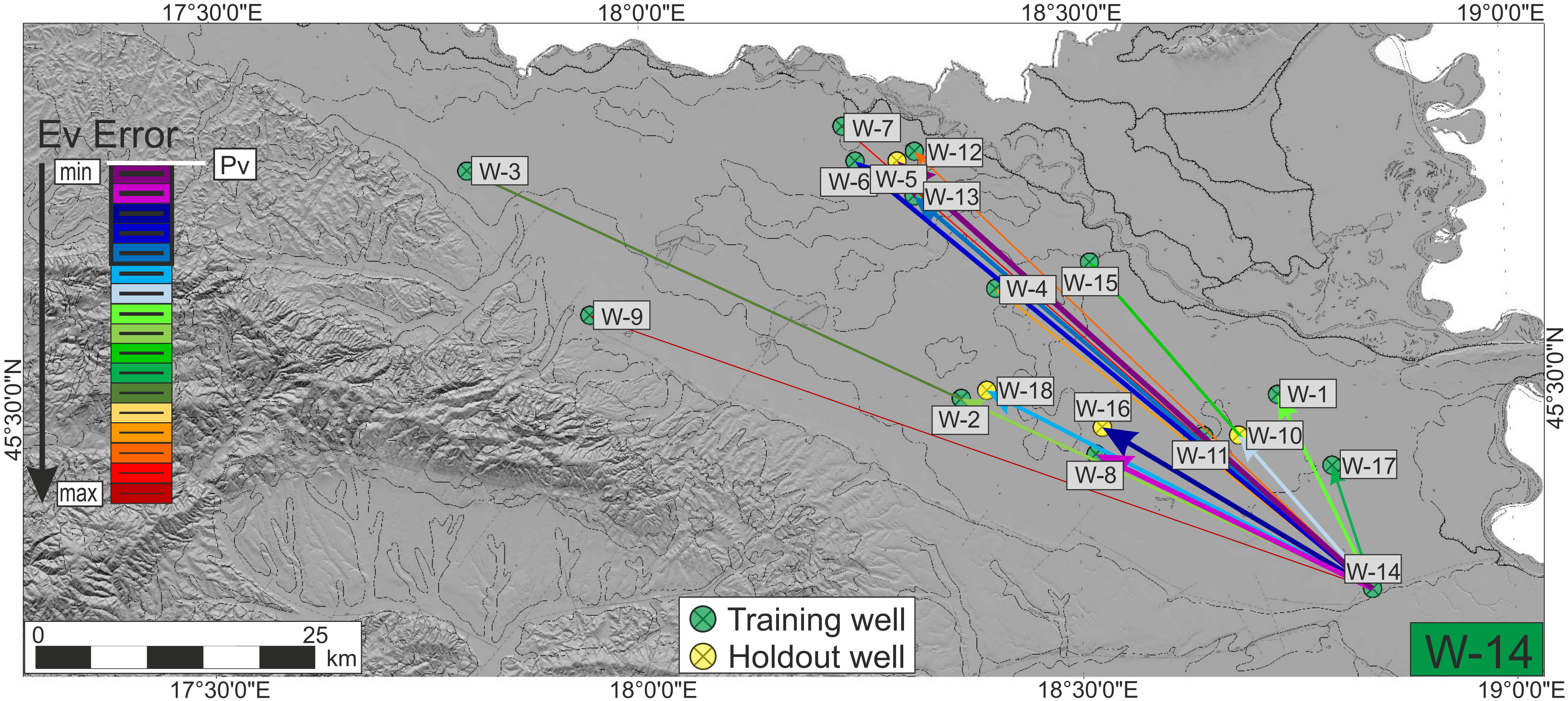
17°30'0"E
18°0'0"E
18°30'0"E
19°0'0"E
45°30'0"N
Ev Error
min
max
Pv
W-1
W-2
W-3
W-4
W-5
W-6
W-7
W-8
W-9
W-10
W-11
W-12
W-13
W-14
W-15
W-16
W-17
W-18
Training well
Holdout well
0
25
km
W-11

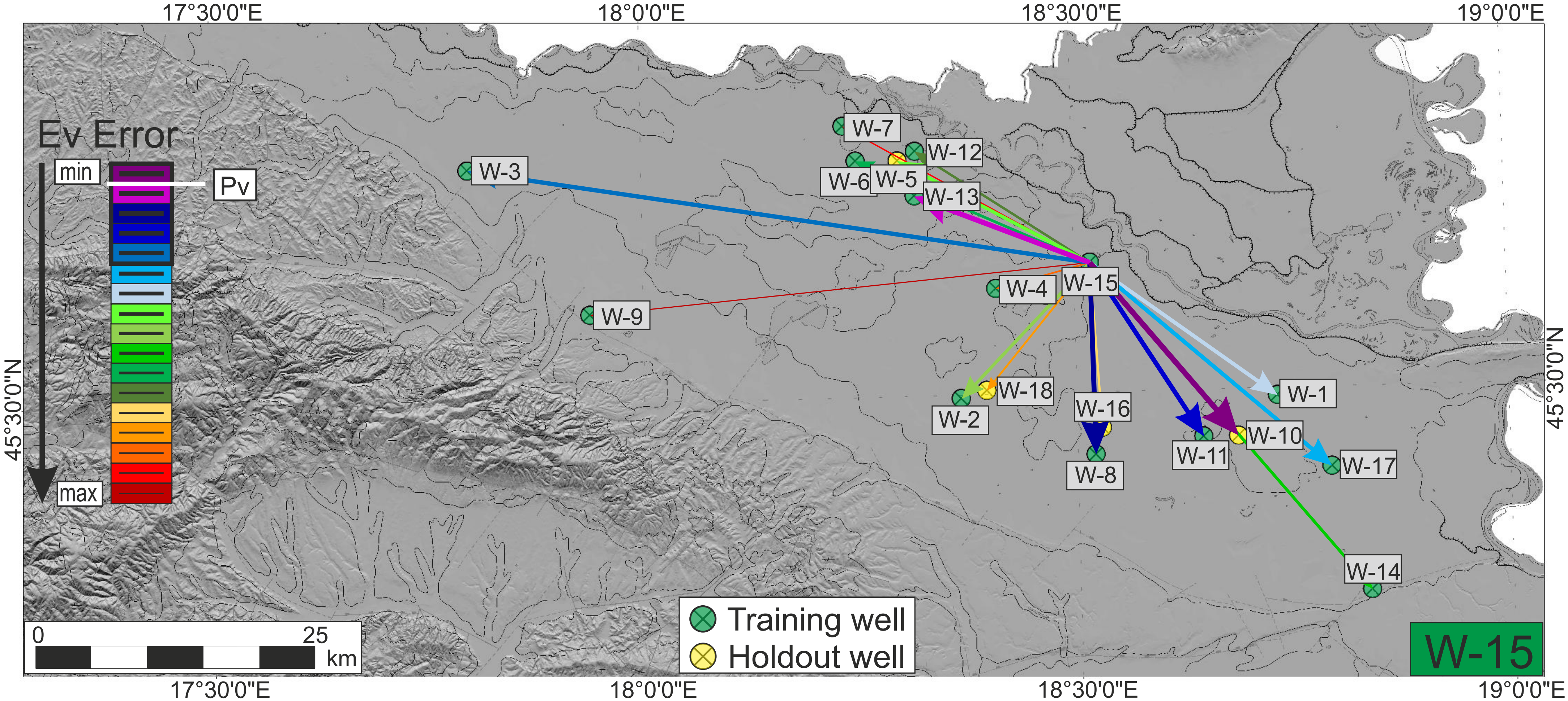


17°30'0"E
18°0'0"E
18°30'0"E
19°0'0"E
45°30'0"N
Ev Error
min
max
Pv
W-1
W-2
W-3
W-4
W-5
W-6
W-7
W-8
W-9
W-10
W-11
W-12
W-13
W-14
W-15
W-16
W-17
W-18
Training well
Holdout well
0
25
km
W-12

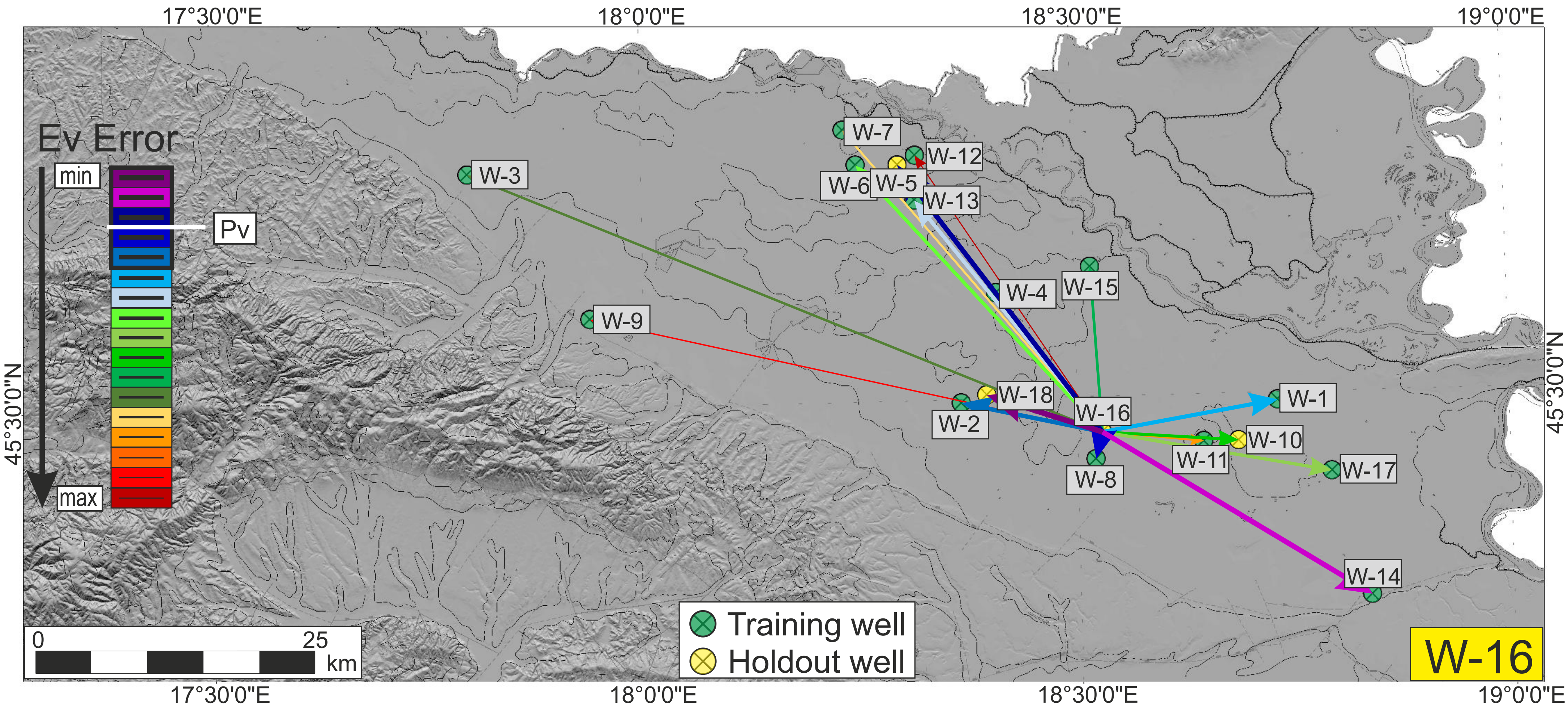


17°30'0"E
18°0'0"E
18°30'0"E
19°0'0"E
45°30'0"N
Ev Error
min
max
Pv
W-1
W-2
W-3
W-4
W-5
W-6
W-7
W-8
W-9
W-10
W-11
W-12
W-13
W-14
W-15
W-16
W-17
W-18
Training well
Holdout well
0
25
km
W-13

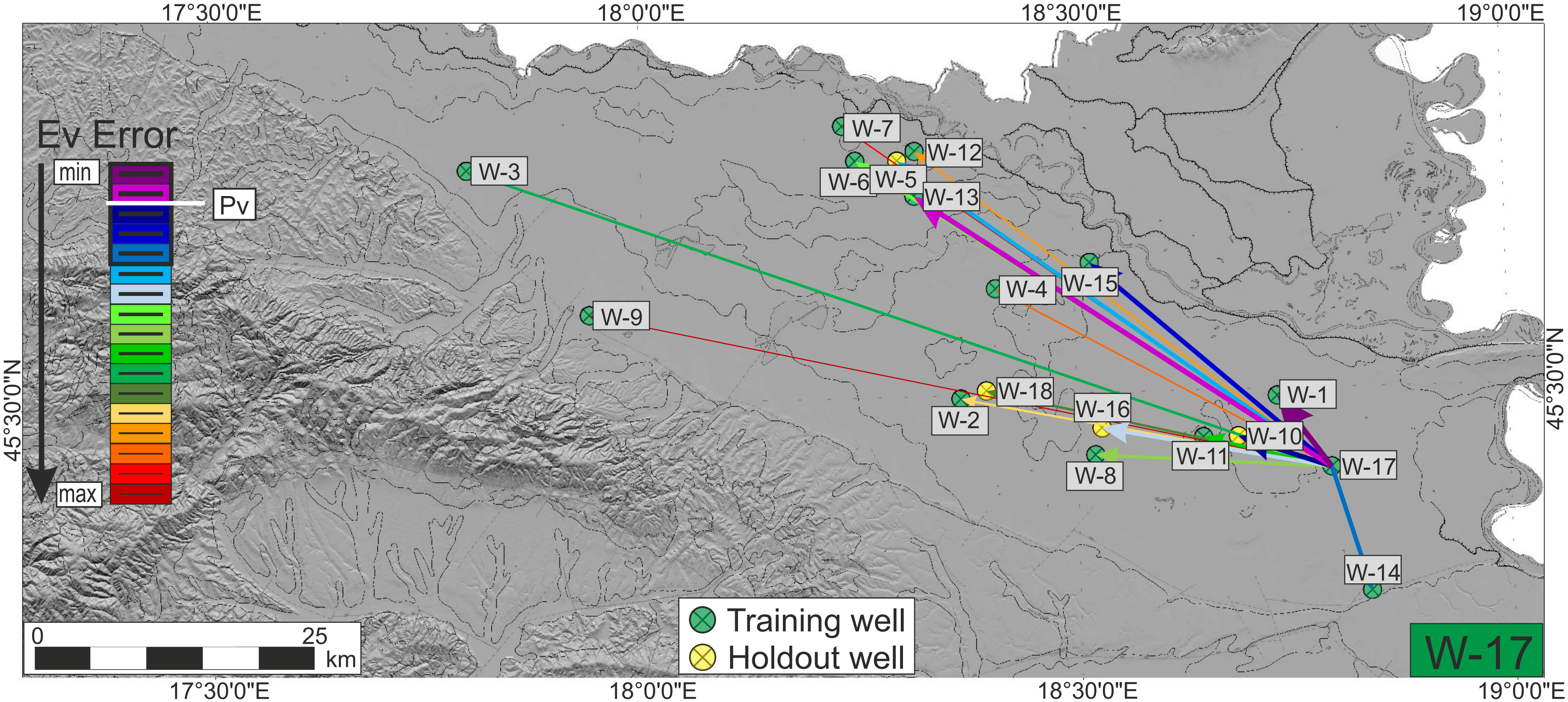

17°30'0"E
18°0'0"E
18°30'0"E
19°0'0"E
45°30'0"N
Ev Error
min
Pv
max
W-1
W-2
W-3
W-4
W-5
W-6
W-7
W-8
W-9
W-10
W-11
W-12
W-13
W-14
W-15
W-16
W-17
W-18
Training well
Holdout well
0
25
km
W-14



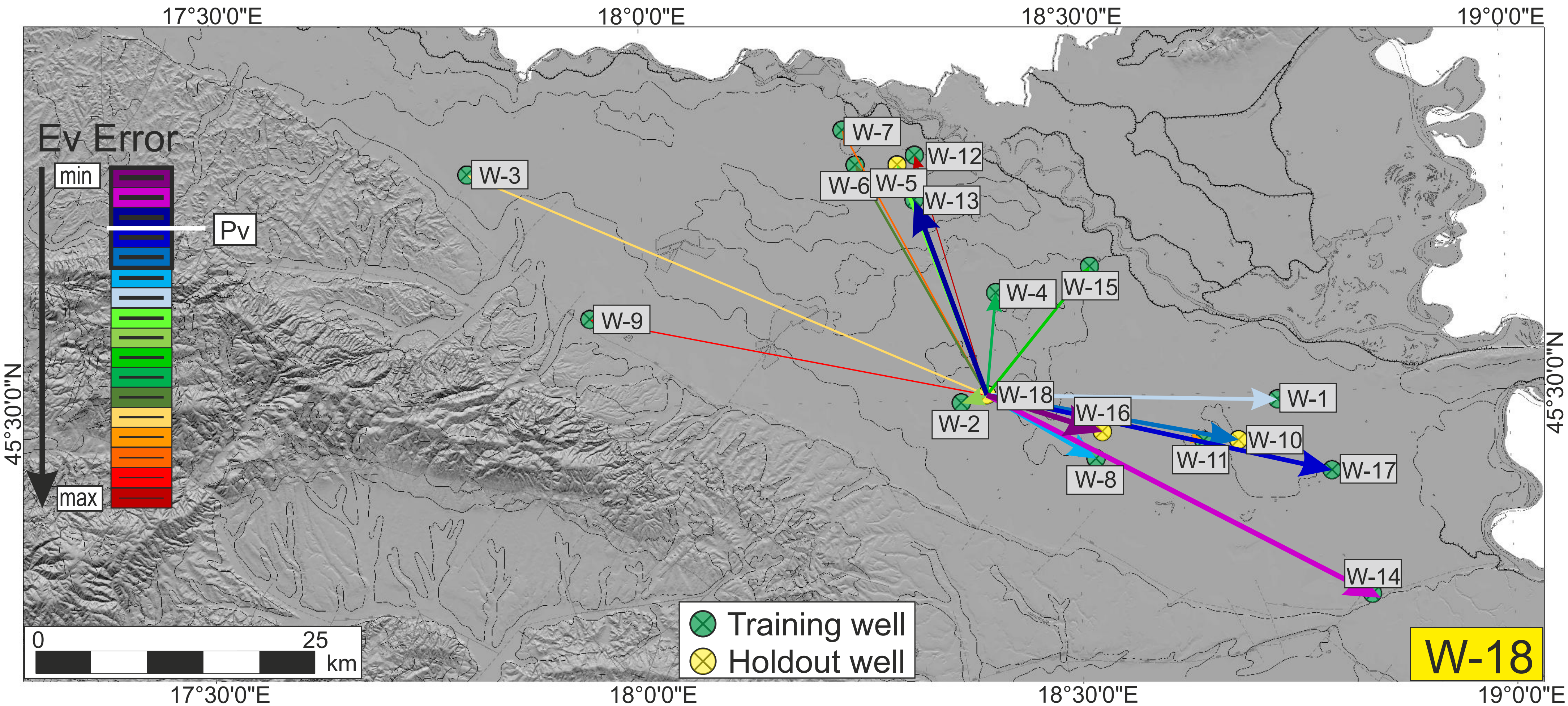
17°30'0"E
18°0'0"E
18°30'0"E
19°0'0"E
45°30'0"N
Ev Error
min
Pv
max
W-1
W-2
W-3
W-4
W-5
W-6
W-7
W-8
W-9
W-10
W-11
W-12
W-13
W-14
W-15
W-16
W-17
W-18
0
25
km
Training well
Holdout well
W-15



17°30'0"E
18°0'0"E
18°30'0"E
19°0'0"E
45°30'0"N
Ev Error
min
max
Pv
W-1
W-2
W-3
W-4
W-5
W-6
W-7
W-8
W-9
W-10
W-11
W-12
W-13
W-14
W-15
W-16
W-17
W-18
Training well
Holdout well
0
25
km
W-16



17°30'0"E
18°0'0"E
18°30'0"E
19°0'0"E
45°30'0"N
Ev Error
min
Pv
max
W-1
W-2
W-3
W-4
W-5
W-6
W-7
W-8
W-9
W-10
W-11
W-12
W-13
W-14
W-15
W-16
W-17
W-18
Training well
Holdout well
0
25
km
W-17


17°30'0"E
18°0'0"E
18°30'0"E
19°0'0"E
45°30'0"N
Ev Error
min
max
Pv
W-1
W-2
W-3
W-4
W-5
W-6
W-7
W-8
W-9
W-10
W-11
W-12
W-13
W-14
W-15
W-16
W-17
W-18
Training well
Holdout well
0
25
km
W-18

5.3 Shale Volume, Seismic Attributes, and Proper Data Preparation: Critical Components for Modeling Subsurface Lithology Distribution

By

Ana Kamenski, Iva Kolenković Močilac & Marko Cvetković

Published as online first in Geologia Croatica

DOI: https://doi.org/10.4154/gc.2025.12

Geologia Croatica
Journal of the Croatian Geological Survey
and the Croatian Geological Society

2025 | 78/2 | XXX– XXX | 5 Figs. | 1 Tab. | www.geologia-croatica.hr

Shale Volume, Seismic Attributes, and Proper Data Preparation: Critical Components for Modelling Subsurface Lithology Distribution

Ana Kamenski[1], Iva Kolenković Močilac[2,*], Marko Cvetković[2]

[1] Croatian Geological Survey, Sachsova 2, 10000 Zagreb Croatia; (akamenski@hgi-cgs.hr)
[2] Faculty of Mining, Geology and Petroleum Engineering, University of Zagreb, Pierottijeva 6, 10000 Zagreb, Croatia;
(*corresponding author: iva.kolenkovic@rgn.unizg.hr; marko.cvetkovic@rgn.unizg.hr)

doi: 10.4154/gc.2025.12

Crossref Similarity Check

Article history:
Manuscript received: February 10, 2025
Revised manuscript accepted: March 25, 2025
Available online: April 28, 2025

Abstract

The scarcity of well data and the inherent subjectivity of geological interpretations often leads to imprecise or oversimplified subsurface models. Traditional interpretation methods struggle with sparse datasets, necessitating the application of advanced machine learning techniques to enhance subsurface characterization. This study leverages artificial neural networks to predict lithology distribution using seismic attributes in the northern Croatian part of the Pannonian Basin System, an area with numerous exploratory wells. Seismic data, long employed as a supplementary interpretation tool, was used to generate a predictive lithological model, overcoming data limitations inherent to well-based methods. A key focus was the volume of shale, a lithological indicator, which was estimated using an extensive set of seismic attributes and processed through innovative data preparation techniques for artificial neural network analysis. A comprehensive artificial neural network based modelling approach was implemented over a 4365 km² 3D seismic dataset, targeting Pannonian (Late Miocene–Early Pliocene) sediments deposited in deltaic, turbiditic, and lacustrine environments. Results show that standardization of input data significantly improved model accuracy, particularly in capturing key geological features such as meandering sandstone-filled channels. In contrast, normalization led to unreliable predictions, while raw data substantially underestimated sandstone volumes. Despite its advantages, the method's limitations stem from the inherent uncertainty in the volume of shale estimation and interpreter subjectivity. The approach is well-suited for geological settings with two or three dominant lithologies distinguishable on geophysical well logs. While applicable to coal-bearing strata and shale-rich carbonates, its effectiveness in more complex geological settings requires further refinement. The findings highlight the untapped potential of legacy seismic data for geo-energy applications, including hydrocarbon exploration, geothermal studies, and carbon storage.

Keywords: volume of shale, artificial neural networks, lithology mapping, geo-energy exploration

1. INTRODUCTION

The lack of subsurface data and the inherent subjectivity of interpreters have been major contributors to imprecise or overly simplified interpretations and resulting geological models. Well data are often too sparse to enable meaningful interpretations using traditional methods, so the implementation of seismic data with advanced machine learning techniques is expected to yield better modelling results (SMIRNOFF et al., 2008; ZHOU et al., 2019; FENG et al., 2024; ZHOU & LIU, 2024). Machine learning techniques enable the combination of well and seismic derived properties in a time efficient way which can greatly improve the accuracy and the resolution of the resulting geological model.

Visualizing subsurface lithology, including the volume and spatial distribution of rock types, typically starts with well data. This generally refers to well logs and information about the lithology of well cuttings and core samples if available. However, in many regions worldwide, the number of wells is limited, they are either sparsely distributed, or the available data is old. The Croatian part of the Pannonian Basin System is characterized by a several hundred, old exploratory wells, some of which were drilled over 50 years ago. Therefore, seismic data has long been used as a supplementary tool to aid the geological interpretation process, offering valuable insights into subsurface characterization (NOVAK ZELENIKA et al., 2018; VUKADIN, 2022; XU & HAQ, 2022). In recent years, this type of study has increasingly relied on the application of artificial neural networks for various purposes: lithology prediction (BRCKOVIĆ et al., 2017; KAMENSKI et al., 2020), estimation of porosity and permeability (ITURRARÁN-VIVEROS & PARRA, 2014), seismic reservoir characterization (OTHMAN et al., 2021), shale volume prediction (TAHERI et al., 2021; MOHAMMADINIA et al., 2023), two-way-time prediction (KAMENSKI et al., 2024).

Recent studies in northern Croatia, specifically within the southwestern part of the Pannonian Basin System, have explored the application of neural networks to enhance the independence and accuracy of subsurface interpretations based on various geophysical datasets (seismics, well logs), in combination with well data (BRCKOVIĆ et al., 2017; KAMENSKI et al., 2020; KAMENSKI et al., 2024). These studies demonstrated the effectiveness of neural networks, particularly in scenarios with limited data. For instance, the inevitable challenges of time-to-depth conversions were successfully surpassed using neural networks, which predicted two-way-time data from stratigraphical and petrophysical parameters in the depth domain, particularly in scenarios where conventional time-to-depth conversion data were unavailable (KAMENSKI et al., 2024). Efforts have also been made to predict lithology based on sparse well and seismic data (BRCKOVIĆ et al., 2017; KAMENSKI et al., 2020). These investigations provided valuable new insights into subsurface characterization in northern Croatia, while also highlighting challenges of lithology prediction, such as the inadequate upscaling of well logs (KAMENSKI et al., 2020).

To determine the lithology distribution in areas where only seismic data is available and well data is sparse, this study utilized an extensive set of seismic attributes to predict the volume of shale, a parameter that serves as an indicator of lithology distribution. Furthermore, this approach was tested with legacy data, as pre-stack seismic data, which can be utilized for lithology prediction, is often unavailable (ADEOTI et al., 2017; BORNARD et al., 2005). To achieve this, innovative data preparation processes were implemented.

This study highlights the successful application of available legacy data, revealing its substantial untapped potential that has yet to be fully utilized when processed by artificial neural network algorithms. Neural networks were chosen for lithology distribution prediction over other machine learning approaches for several key reasons. Primarily, neural networks represent an excellent tool for capturing complex, non-linear relationships inherent in geological data, which makes them well-suited for modelling intricate subsurface patterns. Additionally, they have proven successful in similar applications, such as predicting porosity and permeability (e.g., ITURRARÁN-VIVEROS & PARRA, 2014), further validating their effectiveness in geoscientific tasks. To demonstrate this approach, a study area with available 3D seismic data, covering 4365 km^2 in the northern part of Croatia was selected (Fig. 1) to investigate the possibility of predicting the general lithology distribution within the Pannonian stratigraphic interval based on seismic attributes.

Results from this study can have application in geo-energy characterization for various purposes, from hydrocarbon exploration, geothermal investigations, carbon capture, utilization and storage, etc.

2. GEOLOGICAL OVERVIEW

The study area, located in the North Croatian Basin (NCB), lies within the southwestern part of the Pannonian Basin System (PBS). Base of the Neogene-Quaternary infill is represented by Palaeozoic crystalline and partially metamorphosed rocks, which are in places overlain by Mesozoic carbonates (PAMIĆ & LANPHERE, 1991; PAMIĆ, 1998; PAVELIĆ, 2001; VELIĆ, 2007; MALVIĆ & CVETKOVIĆ, 2013; PAVELIĆ & KOVAČIĆ, 2018). Basin evolution is associated with rifting, and syn-rift and post-rift sediments can be distinguished (LUČIĆ et al., 2001; SAFTIĆ et al., 2003; PAVELIĆ & KOVAČIĆ, 2018; RUKAVINA et al., 2023). The extension began during the Ottnangian and Carpathian and is believed to have been driven by the eastward extrusion of the Alps (FODOR et al., 1999). In these conditions, deposition of coarse-grained sediments (rock-fall breccias and conglomerates) interlayered with sandy and silty layers took place (PAVELIĆ & KOVAČIĆ, 2018). The extension was accompanied by a later marine transgression and volcanic activity in the Badenian (LUČIĆ et al., 2001; SAFTIĆ et al., 2003; ĆORIĆ et al., 2009; MARKOVIĆ et al., 2021). Depositional environments during the Badenian were very diverse. Locally, marsh-type fine-grained sediments can be found, which are overlain by carbonate deposits (BAKRAČ et al., 2010). Deepening of the depositional environment resulted in the lacustrine sedimentation of shales, and siltstones, together with thin, sandy turbidites and occasional conglomerates (PAVELIĆ & KOVAČIĆ, 2018), with sporadic occurrences of pyroclastics resulting from accompanying volcanism (PAVELIĆ, 2001; SAFTIĆ et al., 2003). A change of depositional environment from lacustrine to marine occurred during the Middle Badenian with deposition of shales interlayered with coarse grained clastics (PAVELIĆ & KOVAČIĆ, 2018). The Late Badenian is characterized by the end of the syn-rift phase and the beginning of the post-rift phase (PAVELIĆ, 2001; PAVELIĆ & KOVAČIĆ, 2018). During the Late Badenian, carbonate sedimentation on small carbonate platforms, formed around islands, was followed by marl deposition in the deeper parts of the sea (VRSALJKO et al., 2006; PAVELIĆ & KOVAČIĆ, 2018). By the latest Badenian, a general shallowing occurred, marked by deposition of biocalcarenites and conglomerates, reduced volcanic activity, and localized emersions. The breakup of central Paratethys started in the latest Badenian when it lost connection to the Indo-Pacific Ocean and the palaeo-Mediterranean Sea (RÖGL, 1999). The isolation of the basin and subsequent salinity fluctuations at the Badenian/Sarmatian boundary led to the extinction of most stenohaline marine organisms, while unique associations adapted to the new conditions, emerged or migrated from Eastern Paratethys, marking the Sarmatian age of the deposits (PAVELIĆ & KOVAČIĆ, 2018 and references therein). The final separation of the Pannonian Basin System from other surrounding marine environments took place at the end of the Middle Miocene (ĆORIĆ et al., 2009).

The post-rift phase of PBS development was characterized by a thermal subsidence due to lithospheric cooling (PAVELIĆ & KOVAČIĆ, 2018), resulting in the creation of significant accommodation space. This phase was characterized by the deposition of thick sand and marl sequences in brackish conditions within Lake Pannon (LUČIĆ et al., 2001; SAFTIĆ et al., 2003; PAVELIĆ & KOVAČIĆ, 2018). During the Pliocene and Quaternary, the stress regime turned to compressional, which resulted in activation of reverse faults and reactivation of normal faults with reverse displacement characteristics (HORVÁTH & CLOETINGH, 1996). Sedimentary environments were shallower than in the Pannonian with

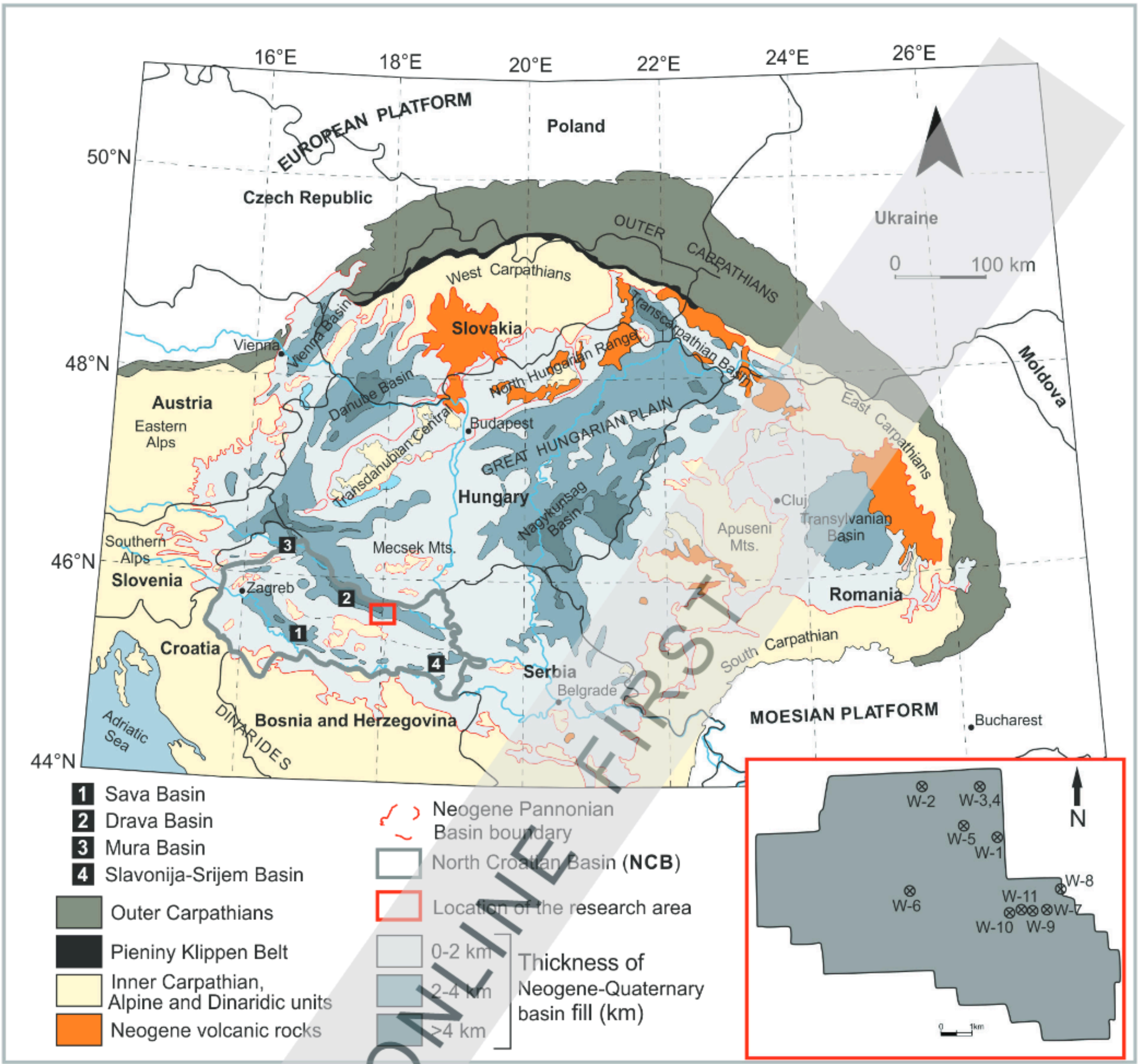


Figure 1. The Pannonian Basin System with the outline of the North Croatian Basin and the study area location. The red rectangle represents the extent of 3D seismic coverage of the "Donji Miholjac" 3D seismic block with the locations of 11 wells highlighted, enlarged in the inset at lower right corner (modified from Cvetković et al. (2019), after Dolton (2006) and Schmid et al. (2008)).

fluctuations from lacustrine and marsh to terrestrial (CVETKOVIĆ, 2013). Lithologically, they are represented by sands, clays and gravel with occasional coal seams or layers. As the tectonic regime shifted from extensional to compressional, basin inversion occurred which led to the formation of structural traps within the older Pannonian sediments.

The NCB's marginal position within the PBS results in generally thinner Neogene sedimentary sequences compared to the central part of the PBS, with the exception of the Drava Basin where the Neogene sedimentary sequence reaches a thickness of almost seven kilometres (SAFTIĆ et al., 2003; VELIĆ, 2007; CVETKOVIĆ et al., 2019).

The focus of this research is the sediments of the Pannonian (Late Miocene–Early Pliocene), that were deposited after the Central Paratethys Sea transitioned into the brackish Lake Pannon around 11.6 million years ago. Pannonian sediments, primarily derived from the Eastern Alps and the Western Carpathians, were deposited in a variety of environments, including deltaic, turbiditic, and lacustrine settings (KOVAČIĆ & GRIZELJ, 2006; PAVELIĆ & KOVAČIĆ, 2018; MATOŠEVIĆ et al., 2024a). The Pannonian deposits include lacustrine marls and limestones in the early stages, followed by sands and siltstones from deltaic environments as the lake progressively filled. These deposits serve as significant source rocks, reservoirs, and caprocks in the Croatian part of the PBS (LUČIĆ et al., 2001; SAFTIĆ et al., 2003).

3. METHODOLOGY

For the task of determining the spatial distribution of lithology throughout the study volume, a comprehensive workflow was implemented. This began with the definition of the model boundaries. The top and bottom of the model were delineated based on regional well tops identified using resistivity well logs. For this study, focusing on the sediments of the Pannonian stratigraphic interval, the regional marker "α" was chosen as the model's top surface and "Rs_7" as its bottom, i.e. Top Pannonian and Base Pannonian surfaces respectively. Lateral boundaries were defined with the 3D seismic volume coverage.

Exploratory wells within the study area were very scarce, so all the wells which at least partially intersected the chosen interval were taken into the analysis. A total of 11 wells were included: eight that drilled trough both the Top Pannonian ("α") and Base Pannonian ("Rs_7") boundaries, and three which were terminated before reaching the Base of the Pannonian, i.e. intersecting only part of the interval of interest. The Top Pannonian well top ("α") is not identified by distinct patterns in the apparent electrical resistivity curve. Instead, it is defined as the transition point where the resistivity curve shifts from a shallower zone with high variability in resistivity to a deeper zone characterized by more stable resistivity values. This transition is the result of change in depositional conditions, specifically the shift from deeper-water lacustrine sedimentation during the Pannonian to shallow lacustrine and alluvial sedimentation in the Pliocene of the Drava Basin and also in the western part of the Sava Basin, which is characterized by more frequent vertical and lateral lithological variations, reflected on the resistivity curve (CVETKOVIĆ, 2017). Well top "Rs_7" was defined by an emphasized increase in resistivity values at the transition from the Lower Pannonian limey shales to the Middle Miocene limestone, due to pronounced resistivity differences between the limestones and shales (PALACKY, 1988). A well to seismic tie was performed either on the basis of available vertical seismic profiling measurements in the wells or with the artificial neural network approach as in KAMENSKI et al. (2024). These horizons were mapped across the "Donji Miholjac" 3D seismic block (Fig. 2), producing interpreted surfaces that defined a study area of 4,365 km² with a total volume of 11,660 km³.

A volume of shale (V_{sh}) analysis was performed by interpreting the Spontaneous Potential (SP) log (Fig. 3), as Neutron, Spectral Radioactivity, Resistivity, and Gamma ray logs were excluded from the analysis due to technological and/or geological constraints specific to the study area. This widely used procedure (SERRA, 1984; ASQUITH & KRYGOWSKI, 2004) is based on the assumption that the SP deflection between the static value of SP (SSP) in a clean sandstone and shale baseline (representing 100% shale) is proportional to the volume of shale (RIDER, 2002), i.e. it assumes that the volume of shale at any given point can be estimated by linear interpolation between the SP value having 0% of shale (SP_{clean}) and the shale baseline value (SP_{shale}) having 100% of shale (Equation 1):

$$V_{sh} = \frac{SP - SP_{clean}}{SP_{shale} - SP_{clean}} \cdot 100 \qquad (1)$$

where SP represents reading of SP value in any point of interest.

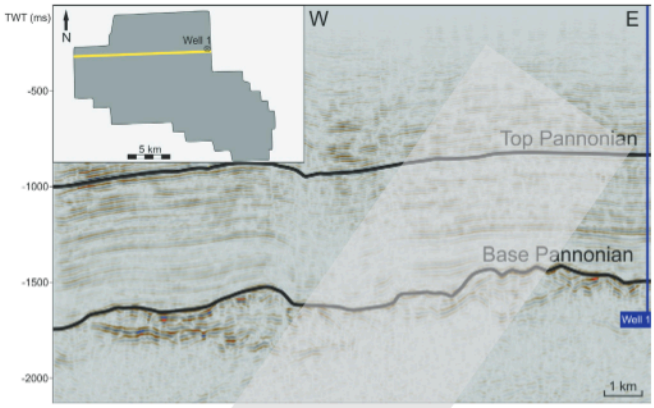


Figure 2. Inline seismic section highlighting the interpreted Top and Base Pannonian surfaces of the subsurface model. The inset indicates the location of the seismic section within the study area.

Once the volume of shale (V_{sh}) values were obtained for all 11 wells, upscaling was performed to firstly average the V_{sh} value within the corresponding model cell and to enable integration with the seismic data. This was performed for models with 20, 50, 100 and 200 layers. Based on the results and to mitigate overestimation of the predominant lithology, the selected model was the one stratified into 200 layers, maintaining an average cell height of 6.5 metres. This approach effectively prevented excessive layering thickness, ensuring that the upscaling process did not introduce biases in lithology distribution predictions by the artificial neural networks (ANNs). Overestimation of the predominant lithology had been a significant challenge in previous studies (KAMENSKI et al., 2020), but this refined layering strategy minimized such distortions.

Following the upscaling of the volume of shale (V_{sh}) values, selected seismic attributes were extracted at the upscaled data points, forming a comprehensive dataset for ANN training. This step ensured that the model retained both the geological resolution necessary for more realistic lithology prediction and the statistically valid dataset required for effective machine learning applications.

Seismic attributes contain a huge amount of data which holds significant relationships between the physical characteristics that remain undetectable through conventional seismic visualization techniques (TANER et al., 1976; TANER, 2001). These attributes, derived from seismic data, capture kinematic, dynamic, geometric, and statistical characteristics, play a fundamental role in structural, stratigraphic, and petrophysical interpretation. Their application significantly enhances subsurface analysis and reservoir characterization (DJEDDI, 2016).

In this study, the selection of seismic attributes was guided by their ability to emphasize lithological and morphological features, thereby aiding artificial neural networks in producing geologically coherent lithology distribution predictions. Seismic attributes that capture key lithological and morphological features were generated based on comprehensive reviews of seismic attribute application (CHOPRA & MARFURT, 2006; LIU & MARFURT, 2006; KER et al., 2014; BRCKOVIĆ et al., 2017; LI et al., 2019; OUMAROU et al., 2021). Twelve attributes were constructed: Sweetness, 3D Curvature, Vari-

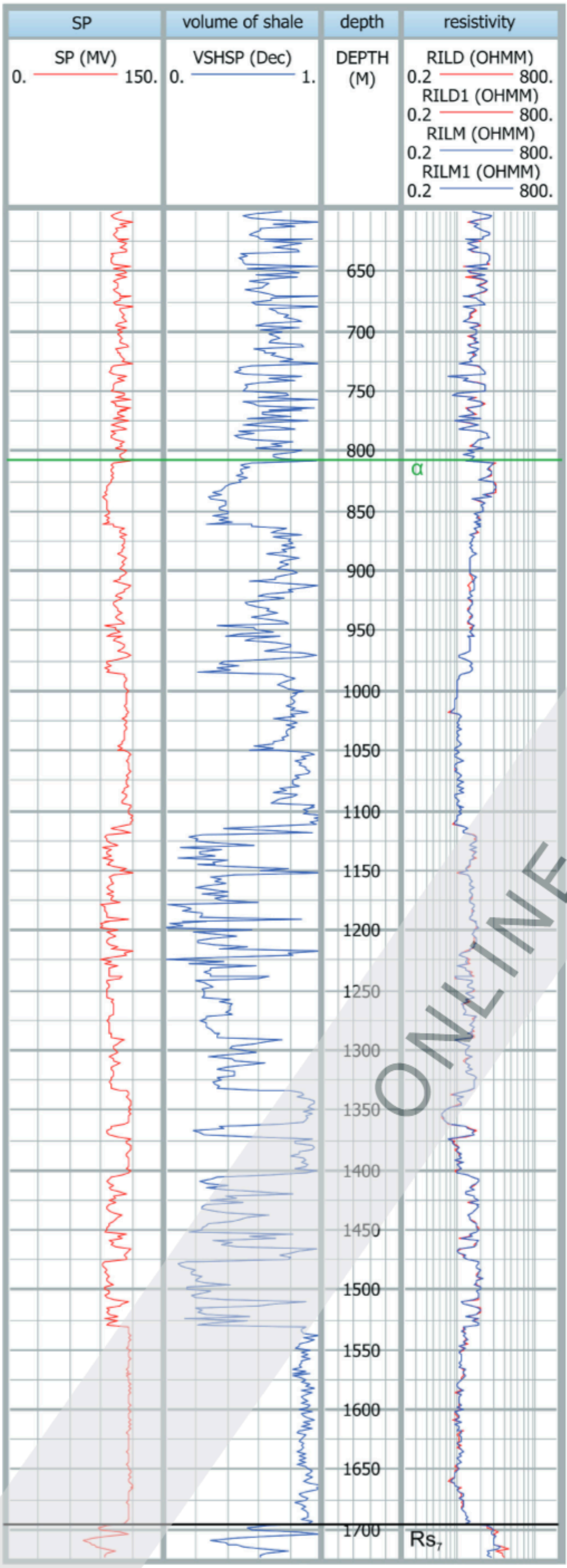


Figure 3. Well logs for Well-1: the SP log is shown in red, the calculated volume of shale is represented in blue, and resistivity logs in both red and blue. Green and black horizontal lines indicate the regional markers "α" and "Rs_7".

ance, Original Amplitude, Instantaneous Frequency, Envelope, Instantaneous Phase, Generalized Spectral Decomposition, Apparent Polarity, Reflection Intensity, Root-Mean-Square (RMS) Amplitude and Relative Acoustic Impedance.

Sweetness and RMS Amplitude effectively detect and display coarse-grained intervals and compaction features (SUBRAHMANYAM & RAO, 2008; KOSON et al., 2014). Variance (edge) and Reflection Intensity serve as reliable indicators of lithology variations (PIGOTT et al., 2013; KOSON et al., 2014), while Original Amplitude provides a clear representation of sediment continuity and discontinuity (BRCKOVIĆ et al., 2017 and references therein). Additionally, Apparent Polarity is mostly related to the useful detection of gas-charged layers (KER et al., 2014).

OUMAROU et al. (2021, and references therein) demonstrated that Instantaneous Frequency aids in seismic facies recognition, while 3D Curvature is crucial for identifying structural features such as channels, faults, anticlines, synclines, and salt domes. Instantaneous Phase delineates subsurface layering, whereas Instantaneous Frequency and Generalized Spectral Decomposition assist in layer thickness estimation and seismic geomorphology analysis (LIU & MARFURT, 2006; LI et al., 2019). Furthermore, Envelope and Relative Acoustic Impedance (RAI) provide insights into lithology, thickness estimation, and sequence delineation, and offer valuable information regarding porosity and permeability (OUMAROU et al., 2021; PIGOTT et al., 2013; KOSON et al., 2014).

Input data for ANN analysis consisted of data points created along well paths, each containing X, Y, Z coordinates, 12 seismic attributes and shale volume values (Fig. 4a, b). Data was statistically processed before the ANN training process. Feature scaling was applied to mitigate the significant scale differences among seismic attributes. Three input versions were prepared: Raw data, Normalized data (rescaled between 0 and 1), and Standardized data (centered at a mean of 0 with a standard deviation of 1). Normalization and standardization processes ensured uniform scaling across all features, including seismic attributes and volume of shale values, which inherently range from 0 to 1.

Data normalization was performed using following Equation 2:

$$x_{normalized} = \frac{(x - x_{\min})}{(x_{max} - x_{\min})} \tag{2}$$

Data standardization was performed using following Equation 3:

$$x_{standardized} = \frac{(x - \text{mean } of\ range)}{\text{standard deviation of range}} \tag{3}$$

ANN analysis was performed in Tibco Statistica within the Statistica (Neural Nets) module. The process consists of the general selection of the ANN architecture constraints (minimum and maximum number of neurons in the hidden layer, activation functions and number of networks to be trained and retained), and the distribution of data into training, test and validation datasets. In this study, 80% of the cases were used for the training dataset, while the remaining 20%

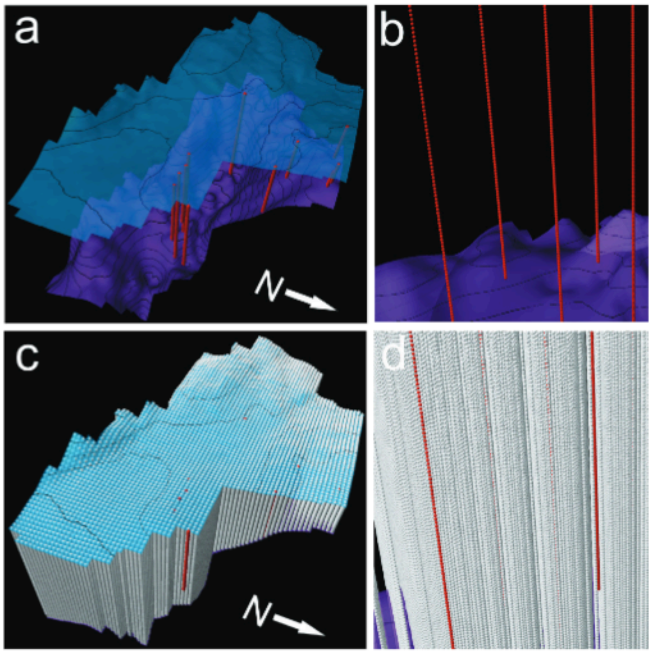


Figure 4. a, b) The blue surface represents the Top Pannonian surface, while the purple surface denotes the Base Pannonian surface. Red points indicate cells containing values for 12 seismic attributes and volume of shale (V_{sh}), which served as training input data for the artificial neural networks (ANNs); **c, d)** White points represent the data points where ANN predicted the volume of shale (V_{sh}) based on assigned seismic attribute values.

was evenly split between testing and validation. The number of iterations of the learning process is not strictly defined but is in the function of the prediction error decay. The learning process stops when the error on the training dataset does not significantly change or the error value is extremely variable for 20 consecutive iterations. The workflow follows three possible end conditions. The first occurs when optimal parameters are achieved, ensuring accurate predictions. The second arises when the ANN algorithm fails to predict the target variable, resulting in random output values. The third end condition is triggered when error values in the test and validation dataset begin to rise, indicating overtraining of the ANN. In all cases, the final network parameters are determined based on their optimal performance across the training, test, and validation datasets.

For prediction, 747,800 cells were generated within the geological model (Fig. 4c, d), with appended seismic attribute values. The trained ANNs (using Raw, Standardized, and Normalized datasets) were applied to predict shale volume for each cell.

Well-log interpretation was performed using Interactive Petrophysics (IP 2021) software, Petrel Schlumberger software package was used for seismic interpretation, attribute extraction, and model construction, and Statistica Tibco for ANN analysis. Additional calculations were performed in Excel.

4. RESULTS

Four subsurface models were constructed to determine the optimal layer thickness for the training and prediction process. These models varied only in the number of layers, which were set to 20, 100 and 200. The corresponding vertical point spacings for these models were an average of 79 m, 16 m and 8 m, respectively. The layering itself has an impact on the analysis in two ways. First, the layer height has a direct effect on the upscaling of the V_{sh} and seismic attribute values. As the layer thickness increases, more values will be averaged representing one data point in the well trajectory (Fig. 4b). Secondly, the number of cases for the ANN analysis significantly decreases from 1888 for the 200-layer case to 193 for the 20-layer case. This has a significant impact on the degree of success of the ANNs training process as it is sensitive to the number of cases for analysis (ALWOSHEEL et al., 2018).

Numerous iterations of the ANN parameters were tested to obtain the best output, as determined by correlation coefficients. The optimal neural network architecture was achieved with a learning rate of 80%, using a neural architecture search as the optimizer and the correlation coefficient as the error metric. Regarding predictions based on standardized and normalized data, the logistic sigmoid function proved to be the best activation function for both the hidden and output layers. In contrast, for raw data, best performance was obtained using the logistic sigmoid function for the hidden layer and the sinusoidal function for the output layer. The predictive performance of ANNs for the volume of shale values was evaluated using correlation coefficients of the target (interpreted V_{sh} value) and the predicted value of V_{sh} (Table 1), with the correlation coefficient serving as the primary metric for distinguishing between high- and low-efficient neural networks. The highest predictive accuracy was achieved in the model with most layers (Table 1, Model 200 layers). Among the tested models, the ANN trained on input data from the 200-layer model demonstrated to be the most successful and was therefore selected for further investigation (Table 1). ANNs trained on models with less than 100 layers had poorer performance (20-layer model) or were completely unable to predict the V_{sh} data (50-layer model which was omitted from the study).

To predict the volume of shale across the entire 3D seismic coverage, a model with 747,800 cells was generated (Fig. 4c, d), each containing values from 12 selected seismic attributes. These input data were processed using the same methodology as the training dataset. The most effective ANN architecture was then deployed to predict the volume of shale values across

Table 1. Correlation coefficients representing the performance of artificial neural networks for the three differently layered models. Bold results represent the model which is selected for artificial neural network analysis to predict the volume of shale values throughout the investigated area. The network architecture is represented by the number of neurons in the input, hidden and output layer, while the number in brackets indicates the number of learning iterations.

Input data	Correlation coefficients		
	Model 20 layers	Model 100 layers	**Model 200 layers**
Raw	0.34	0.49	**0.68**
Normalized (Norm)	0.31	0.60	**0.70**
Standardized (STD)	0.33	0.43	**0.72**
	Network architecture		
Raw	12-123-1 (23)	12-161-1 (83)	**12-128-1 (270)**
Normalized (Norm)	12-181-1 (22)	12-138-1 (124)	**12-145-1 (208)**
Standardized (STD)	12-132-1 (21)	12-168-1 (92)	**12-80-1 (209)**

the target dataset, which consisted solely of 3D seismic attribute values. Table 1 demonstrates that standardized data produced the most efficient ANN predictions, whereas raw datasets yielded the least accurate predictions (Table 1, Model 200 layers).

Predicted shale volume values were categorized as follows: ≤0.5 as Sandstone, 0.5–0.7 as Sandstone-Shale, and ≥0.7 as Shale. These classifications were upscaled for lithology modelling, resulting in three distinct models corresponding to Raw, Standardized, and Normalized input data (Fig. 5).

The predicted lithology distribution for the model constructed from standardized (STD) input data suggests 52.17% Shale, 29.89% Sandstone-Shale and 17.94% Sandstone. For the model based on normalized (Norm) input data, the distribution was 55.14% Shale, 42.59% Sandstone-Shale, and 2.27% Sandstone. Predictions from raw (Raw) input data resulted in 38.12% Shale, 48.57% Sandstone-Shale, 13.31% Sandstone.

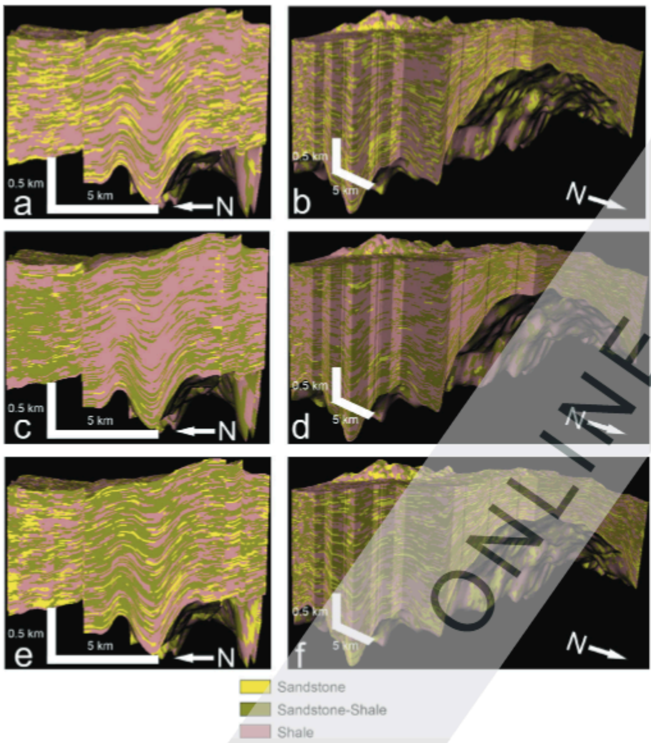


Figure 5. Lithology models developed using predicted volumes of shale values derived from: **a, b)** Standardized (STD) input data; **c, d)** Normalized (Norm) input data; and **e, f)** Raw input data.

5. DISCUSSION

The volume of shale (V_{sh}) was calculated from the Spontaneous Potential (SP) log. The Neutron log was not used due to the simple fact that only small intervals in the wells had a neutron log recorded, so in order to preserve consistency across the investigated interval, the electrolog was expected to be a better solution. Also, the Neutron log is, due to its shallow depth of penetration and consequential effect of wellbore conditions caused by the increased well diameter in clayey rocks, prone to misleading volume of shale estimates (KAMAYOU et al., 2021). The Resistivity log was not used because of the mineral composition of the Pannonian sandstones, which are lithic arenites with significant content of calcite, feldspars and micas (MATOŠEVIĆ et al., 2023, 2024a, 2024b). Namely, PIMIENTA et al. (2019) indicate significant differences in the resistivities of quartz-rich sandstones and feldspar rich sandstones, the latter showing significantly lower resistivities which could result in overestimated values of shale volume. Gamma ray (GR) cannot be used in the studied settings for the same reason – the significant content of feldspars and micas in the Upper Miocene sandstones, i.e. their potassium content, affects the GR readings (IMAM & TREWIN, 1991) leading to overestimated volume of clay/shale values (KAMEL & MABROUK, 2003). The thorium content from the Spectral Radioactivity log could be considered (RIDER, 2002) in a given setting, but the problem was the availability of Spectral Radioactivity logs that are very rare for the study area.

SP should not be used to estimate shale volume in areas where formation water resistivity (R_w) is not much different from mud filtrate resistivity (R_{mf}) (KAMEL & MABROUK, 2003), but that was not a limiting factor in the study area where the mud filtrate generally shows significantly different resistivity compared to the formation water. However, the shallower parts of the investigated unit/intervals could be affected by this limitation, due to the lower total dissolved solids (TDS) of the formation water, but short intervals of formation water salinity change mostly coincided with the upper boundary of the model – well top "α", Top Pannonian surface.

The upscaled volume of shale values served as the input variable for modelling, ensuring continuity rather than discrete categorization, as seen in previous studies (BRCKOVIĆ et al., 2017; KAMENSKI et al., 2020). This approach demonstrated the successful application of continuous input data in artificial neural networks, a methodology not commonly adopted in earlier research. Additionally, the upscaling process was carefully managed by implementing thin layering, effectively preventing overestimation of the predominant lithology. This was evident by the network performance increase with the increase of the number of layers within the geological model. Oversimplification in geological models with layers thicker than 16 metres led to poor performance or the complete inability of the ANN processing the input data to predict the V_{sh}.

Based on the performance of the ANN training (Table 1, Model 200 layers), the best results were achieved using Standardized input data, followed by Normalized data and, lastly, Raw data. It is anticipated that Normalized or Standardized input data would yield better predictions, as Raw data contains original seismic attribute values that show significant variations and difference in magnitude. For example, the range of 3D Curvature spans from -0.56 to 2.33, Original Amplitude units range from -63,668.92 to 52,227.58, and Envelope values range from 459.39 to 73,946.72, while the volume of shale (V_{sh}) values range from 0 to 1. This high inconsistency in scale poses substantial challenges for ANNs to accurately predict V_{sh} from seismic attributes. The normalization and standardization processes enable a more balanced representation of the features, improving model performance.

Better performance of standardization compared to normalization can be attributed to the original data being more accurately characterized by mean and standard deviation, resulting in a standardized value range of -10.53 to 14.13, as opposed to the constrained normalization range of 0 to 1. These findings strongly suggest that preprocessing data, either by standardization or normalization, is essential for accurately predicting the volume of shale values from seismic attributes.

However, the results from lithology modelling offer distinct insights into the impact of data preprocessing on model accuracy and geological consistency. While preprocessing input data generally improves model performance, this study confirms that the benefits are particularly pronounced when using Standardized input data. The lithological model generated with STD data produced results that align closely with geological expectations, exhibiting features such as well-developed meander channels filled with sandstones. These results are not only statistically reliable but also geologically meaningful, as they are supported by numerous independent geological studies (ŠPELIĆ et al., 2023; PAVELIĆ & KOVAČIĆ, 2018 and references therein). The enhanced accuracy of this model suggests that standardization effectively preserves and enhances key lithological trends within the dataset.

In contrast, models constructed using Normalized (Norm) input data produced less reliable predictions. In some cases, their performance was even inferior to models developed using Raw (unprocessed) data. This suggests that normalization may distort the original statistical relationships between variables, leading to a loss of critical geological information. Notably, the model derived from Raw input data exhibited a significant underestimation of the total volume of sandstones, highlighting the risks associated with inadequate preprocessing. The omission or misrepresentation of such key geological features could lead to inaccurate interpretations of reservoir quality, depositional environments, or resource potential.

These findings underscore the crucial role of data preprocessing in artificial neural network (ANN)-based lithology distribution modelling. Proper standardization of the input data has resulted in geologically accurate and statistically reliable models, reinforcing the importance of preprocessing in machine learning applications within geosciences. However, normalization appears to compromise model reliability, making it an unsuitable preprocessing method for lithological modelling. As a result, careful selection of preprocessing techniques is essential for ensuring reliable, geologically consistent predictions in lithology modelling. This study proved that combining machine learning with traditional geological methods holds great potential for improving predictive capabilities in the geosciences.

The approach presented here is applicable in geological settings characterized by the relative uniformity of lithological composition, i.e. in settings with two to three main distinctive lithologies that can be distinguished based on the geophysical exploration data from wells. This approach is therefore expected to be applicable in constructing the model of coal bearing strata, where the Density log could be used to identify different lithological categories. It could also be used for the characterization of shale-rich carbonates, i.e. the spatial zonation of carbonate reservoirs with respect to shale volume. Modelling the spatial distribution of lithological composition prior to porosity modelling can be beneficial, as each lithological category can be associated with a specific porosity range (KOLENKOVIĆ MOČILAC et al., 2022). This approach allows the lithological model to control the porosity model, therefore representing a significant improvement over a simple sandstone-shale system. Additionally, the methodology offers potential for various geo-energy explorations, including hydrocarbon and geothermal explorations, as well as CO_2 storage assessments. However, in more complex lithological settings, this methodology would require further refinement and development to achieve optimal accuracy.

The transfer of this methodology to other settings is conditioned not only by geological characteristics but also by differences in the input data set. The presented methodology is limited to 3D seismic data due to the usage of a large number of attributes, some of which are not at all suited for calculation on 2D seismic data. Furthermore, the effectiveness of standardization over normalization is linked to the specific numerical value ranges of the seismic attributes in this study. These ranges are not necessarily the same for all 3D volumes, meaning that in other cases, normalization may yield better results. Therefore, the transferability of direct ANN parameters to other regions with different input data is not expected to be as successful; however, the developed methodology concept has high potential to improve lithology prediction across diverse geological settings.

6. CONCLUSION

This study presents an ANN approach for determining the subsurface lithology distribution based on well log and 3D seismic data. The importance of data preprocessing in ANN-based lithology modelling is highlighted. Standardized input data produced geologically consistent results, accurately representing key features including fluvial meander channels filled with sandstones. In contrast, models using Normalized data were less reliable, sometimes even underperforming compared to those using Raw data, which significantly underestimated sandstone volume. These findings emphasize that standardization enhances model accuracy, while normalization should be avoided as it can compromise lithology predictions.

The limitations of the approach presented in this work are mainly related to the estimation of the volume of shale, since the volume of shale represents a key input parameter, and its estimate is influenced by the available dataset as well as specific geological settings. Furthermore, there is a certain amount of subjectivity of the interpreter affecting the final result, but these issues are inherent to all analyses which are influenced by the volume of shale parameter.

The approach is suitable for geological settings with two to three main lithologies, distinguishable through geophysical well data. It can be applied to coal-bearing strata using Density logs or to shale-rich carbonates for reservoir zonation. However, for geologically complex settings, it requires significant refinement.

ACKNOWLEDGEMENT

This research was conducted as part of the project funded by the Croatian Science Foundation (No. HRZZ UIP-2019-04-3846): "GEOlogical Characterization of the Eastern Part of the Drava Depression Subsurface Intended for the Evaluation of Energy Potentials (GEODEP)." The authors express their sincere gratitude to Schlumberger for generously providing the Petrel academic software license and Geoactive for providing academic software licences for Interactive Petrophysics. Additionally, we extend our appreciation to the Croatian Hydrocarbon Agency for granting access to the data, without which this research would not have been possible to conduct.

Authors would like to thank professor Marijan Kovačić as welll as an anonymous reviewer for their thorough and detailed review. Their insightful comments helped to identify and address shortcomings, ultimately improving the quality of the manuscript.

REFERENCES

ALWOSHEEL, A., VAN CRANENBURGH, S. & CHORUS, C.G. (2018): Is your dataset big enough? Sample size requirements when using artificial neural networks for discrete choice analysis.– Journal of Choice Modelling, 28, 167–182. https://doi.org/10.1016/j.jocm.2018.07.002

ASQUITH, G. & KRYGOWSKI, D. (2004): Basic Well Log Analysis.– AAPG Methods in Exploration 16, 21-30.

BAKRAČ, K., HAJEK-TADESSE, V., MIKNIĆ, M., GRIZELJ, A., HEĆIMOVIĆ, I. & KOVAČIĆ, M. (2010): Evidence for Badenian local sea level changes in the proximal area of the North Croatian Basin.– Geologia Croatica, 63/3, 259-269. https://doi.org/10.4154/GC.2010.21

BRCKOVIĆ, A., KOVAČEVIĆ, M., CVETKOVIĆ, M., KOLENKOVIĆ MOČILAC, I., RUKAVINA, D. & SAFTIĆ, B. (2017): Application of artificial neural networks for lithofacies determination based on limited well data.– Central European geology, 60/3; 299–315. https://doi.org/10.1556/24.60.2017.012

BORNARD, R., ALLO, F., COLÉOU, T., FREUDENREICH, Y., CALDWELL, D.H. & HAMMAN, J.G. (2005): Petrophysical seismic inversion to determine more accurate and precise reservoir properties. SPE, 94144.

CHOPRA, S. & MARFURT, K. (2006): Seismic attributes – a historical perspective.– Geophysics, 70/5, 3SO–28SO. http://doi.org/10.1190/1.2098670

CVETKOVIĆ, M. (2017): Possibilities for Well Log Correlation using Standard Deviation Trends in Neogene-Quaternary Sediments, Sava Depression, Pannonian Basin.– Geologia Croatica, 70/2, 79–85. https://doi.org/10.4154/gc.2017.07

CVETKOVIĆ, M., MATOŠ, B., RUKAVINA, D., KOLENKOVIĆ MOČILAC, I., SAFTIĆ, B., BAKETARIĆ, T., BAKETARIĆ, M., VUIĆ, I., STOPAR, A., JARIĆ, A. & PAŠKOV, T. (2019): Geoenergy potential of the Croatian part of Pannonian Basin: insights from the reconstruction of the pre-Neogene basement unconformity.– Journal of Maps, 15/2, 651–661. https://doi.org/10.1080/17445647.2019.1645052

ĆORIĆ, S., PAVELIĆ, D., RÖGL, F., MANDIC, O., VRABAC, S., AVANIĆ, R., JERKOVIĆ, L. & VRANJKOVIĆ, A. (2009): Revised Middle Miocene datum for initial marine flooding of North Croatian Basins (Pannonian Basin System, Central Paratethys).– Geologia Croatica, 62/1, 31–43. https://doi.org/10.4154/GC.2009.03

DJEDDI, M. (2016): Transformée de Hilbert, signal analytique et attributs sismiques.– Département de Géophysique, FHC-Université M'Hamed Bougara de Boumerdes, Algérie.

DOLTON, G.L. (2006): Pannonian Basin Province, Central Europe (Province 4808) – petroleum geology, total petroleum systems, and petroleum resource assessment.– Bulletin of the United States Geological Survey 2204, 47.

FENG, Y., WEN, G., SHANG, J., WEN S. & WU, B. (2024): Research on 3D geological modeling based on boosting integration strategy.– Ore Geology Reviews, 171, 106157. https://doi.org/10.1016/j.oregeorev.2024.106157

FODOR, L., CSONTOS, L., BADA, G., GYÖRFI, I. & BENKOVICS, L. (1999): Tertiary tectonic evolution of the Pannonian Basin system and neighbouring orogens: a new synthesis of palaeostress data.– Geological Society, London, Special Publications, 156/1, 295-334. https://doi.org/10.1144/GSL.SP.1999.156.01.15

HORVÁTH, F. & CLOETINGH, S. (1996): Stress-induced late-stage subsidence anomalies in the Pannonian Basin.– Tectonophysics, 266/1–4, 287–300. https://doi.org/10.1016/S0040-1951(96)00194-1

IMAM, Md.B. & TREWIN, N.H. (1991): Factors contributing to high gamma-ray levels in Upper Jurassic sandstone reservoirs of the Claymore Oilfield, North Sea.– Marine and Petroleum Geology, 8/4, 452–460. https://doi.org/10.1016/0264-8172(91)90067-B

ITURRARÁN-VIVEROS, U. & PARRA, J.O. (2014): Artificial Neural Networks applied to estimate permeability, porosity and intrinsic attenuation using seismic attributes and well-log data.– Journal of Applied Geophysics, 107, 45–54. https://doi.org/10.1016/j.jappgeo.2014.05.010

KAMAYOU, V., EHIRIM, C. & IKIENSIKIMAMA, S. (2021): Estimating Volume of Shale in a Clastic Niger Delta Reservoir from Well Logs: A Comparative Study.– International Journal of Geosciences, 12, 949–959. https://doi.org/10.4236/ijg.2021.1210049

KAMEL, M.H. & MABROUK, W.M. (2003): Estimation of shale volume using a combination of the three porosity logs.– Journal of Petroleum Science and Engineering, 40/3-4, 145–157. https://doi.org/10.1016/S0920-4105(03)00120-7

KAMENSKI, A., CVETKOVIĆ, M., KOLENKOVIĆ MOČILAC, I. & SAFTIĆ, B. (2020): Lithology prediction in the subsurface by artificial neural networks on well and 3D seismic data in clastic sediments: a stochastic approach to a deterministic method.– GEM - International journal on geomathematics, 11/8, 1–24. https://doi.org/10.1007/s13137-020-0145-3

KAMENSKI, A., CVETKOVIĆ, M., KAPURALIĆ, J., KOLENKOVIĆ MOČILAC, I. & BRCKOVIĆ, A. (2024): From Traditional Extrapolation to Neural Networks: Time-Depth Relationship Innovations in the Subsurface Characterization of Drava Basin, Pannonian Super Basin.– Advances in geo-energy research, 14/1, 25–33. https://doi.org/10.46690/ager.2024.10.05

KER, S., LE GONIDEC, Y., MARSSET, B., WESTBROOK, G.K., GIBERT, D. & MINSHULL, T.A. (2014): Fine-scale gas distribution in marine sediments assessed from deep-towed seismic data.– Geophysical Journal International, 196/3, 1466–1470. https://doi.org/10.1093/gji/ggt497

KOLENKOVIĆ MOČILAC, I., CVETKOVIĆ, M., SAFTIĆ, B. & RUKAVINA, D. (2022): Porosity and Permeability Model of a Regionally Extending Unit (Upper Miocene Sandstones of the Western Part of Sava Depression, Croatia) Based on Vintage Well Data.– Energies (Basel), 15/16, 1–18. https://doi.org/10.3390/en15166066

KOSON. S., CHENRAI, P. & CHOOWONG, M. (2014): Seismic Attributes and Their Applications in Seismic Geomorphology.– Bulletin of Earth Sciences of Thailand, 6/1, 1–9.

KOVAČIĆ, M. & GRIZELJ, A. (2006): Provenance of the Upper Miocene clastic material in the southwestern part of the Pannonian Basin.– Geologica Carpathica, 57, 495–510.

LI, W., YUE, D., WU, S., SHU, Q., WANG, W., LONG, T. & ZHANG, B. (2019): Thickness prediction for high-resolution stratigraphic interpretation by fusing seismic attributes of target and neighboring zones with an SVR algorithm.– Marine and Petroleum Geology, 113, 104153. https://doi.org/10.1016/j.marpetgeo.2019.104153

LIU, J. & MARFURT, K.J. (2006): Thin bed thickness prediction using peak instantaneous frequency.– SEG Technical Program Expanded, Abstracts Society of Exploration Geophysicists, 968–972. http://doi.org/10.1190/1.2370418

LUČIĆ, D., SAFTIĆ, B., KRIZMANIĆ, K., PRELOGOVIĆ, E., BRITVIĆ, V., MESIĆ, I. & TADEJ, J. (2001): The Neogene evolution and hydrocarbon potential of the Pannonian Basin in Croatia.– Marine and Petroleum Geology, 18, 133–147. https://doi.org/10.1016/S0264-8172(00)00038-6

MALVIĆ, T. & CVETKOVIĆ, M. (2013): Lithostratigraphic units in the Drava Depression (Croatian and Hungarian parts) – a correlation.– Nafta, 64/1, 27–33.

MARKOVIĆ, F., KUIPER, K., ĆORIĆ, S., HAJEK-TADESSE, V., HERNITZ KUČENJAK, M., BAKRAČ, K., PEZELJ, Đ. & KOVAČIĆ, M. (2021): Middle Miocene marine flooding: new 40Ar/39Ar age constraints with integrated biostratigraphy on tuffs from the North Croatian Basin.– Geologia Croatica, 74/3, 237-252. https://doi.org/10.4154/gc.2021.18

MATOŠEVIĆ, M., MARKOVIĆ, F., BIGUNAC, D., ŠUICA, S., KRIZMANIĆ, K., PERKOVIĆ, A., KOVAČIĆ, M. & PAVELIĆ, D. (2023): Petrography of the Upper Miocene sandstones from the North Croatian Basin: Understanding the genesis of the largest reservoirs in the southwestern part of the Pannonian Basin System.– Geologica Carpathica, 74/2, 155–179. https://doi.org/10.31577/GeolCarp.2023.06

MATOŠEVIĆ, M., ŠUICA, S., BERSANI, D., MARKOVIĆ, F., RAZUM, I., GRIZELJ, A., PETRINJAK, K., KOVAČIĆ, M. & PAVELIĆ, D. (2024a). The Alps as the main source of sand for the Late Miocene Lake Pannon (Pannonian Basin, Croatia).– Geologia Croatica, 77/2, 69-83. https://doi.org/10.4154/gc.2024.05

MATOŠEVIĆ, M., TOMAŠIĆ, N., PERKOVIĆ, A., KAMPIĆ, Š., KOVAČIĆ, M. & PAVELIĆ, D. (2024b): Reservoir Quality Evaluation: Unveiling Diagenetic Transformations through Mineralogical and Petrophysical Analyses of the Upper Miocene Lacustrine Sandstones in the Pannonian Basin System, Croatia.– Rudarsko-Geološko-Naftni Zbornik, 39/3, 153–172. https://doi.org/10.17794/rgn.2024.3.12

MOHAMMADINIA, F., RANJBAR, A., KAFI, M., SHAMS, M., HAGHIGHAT, F. & MALEKI, M. (2023): Shale volume estimation using ANN, SVR, and RF algorithms compared with conventional methods.– Journal of African Earth Sciences, 205, 104991. https://doi.org/10.1016/j.jafrearsci.2023.104991

NOVAK ZELENIKA, K., NOVAK MAVAR, K. & BRNADA, S. (2018): Comparison of the Sweetness Seismic Attribute and Porosity–Thickness Maps, Sava Depression, Croatia.– Geosciences, 8, 426. https://doi.org/10.3390/geosciences8110426

OTHMAN, A., FATHY, M. & MOHAMED, I.A. (2021): Application of Artificial Neural Network in seismic reservoir characterization: a case study from Offshore Nile Delta.– Earth Science Informatics, 14, 669–676. https://doi.org/10.1007/s12145-021-00573-x

OUMAROU, S., MABROUK, D., TABOD., T.C., MARCEL, J., NGOS III, S., ESSI, J.M.A. & KAMGUIA, J. (2021): Seismic attributes in reservoir characterization: an overview.– Arabian Journal of Geosciences, 14, 402. https://doi.org/10.1007/s12517-021-06626-1

PALACKY, G.V. (1988): Resistivity Characteristics of Geologic Targets.– In: Electromagnetic Methods in Applied Geophysics, 1, 53-129. Tulsa, OK: Society of Exploration Geophysicists. https://doi.org/10.1190/1.9781560802631.ch3

PAMIĆ, J. (1998): Crystalline basement of the South Pannonian Basin based on surface and subsurface data.– Nafta, 49, 371–390.

PAMIĆ, J. & LANPHERE, M. (1991): Hercynian Granites and Metamorphic Rocks from The Mts. Papuk, Psunj, Krndija, and The Surrounding Basement of The Pannonian Basin in Slavonija (Northern Croatia, Yugoslavia).– Geologija, 34/1, 81–253. https://doi.org/10.5474/geologija.1991.004

PAVELIĆ, D. (2001): Tectonostratigraphic model for the North Croatian and North Bosnian sector of the Miocene Pannonian Basin System.– Basin Research, 13, 359–376. https://doi.org/10.1046/j.0950-091x.2001.00155.x

PAVELIĆ, D. & KOVAČIĆ, M. (2018): Sedimentology and stratigraphy of the Neogene rift-type North Croatian Basin (Pannonian Basin System, Croatia): A review.– Marine Petroleum Geology, 91, 455–469. https://doi.org/10.1016/j.marpetgeo.2018.01.026

PIGOTT, J.D., KANG, M.-H. & HAN, H.-C. (2013): First order seismic attributes for clastic seismic facies interpretation: Examples from the East China Sea.– Journal of Asian Earth Sciences, 66, 34–54. https://doi.org/10.1016/j.jseaes.2012.11.043

PIMIENTA, L., ORELLANA, L.F. & VIOLAY, M. (2019): Variations in elastic and electrical properties of crustal rocks with varying degree of microfracturation.– Journal of Geophysical Research: Solid Earth, 124, 6376–6396. https://doi.org/10.1029/2019JB017339

RIDER, M.H. (2002): The geological interpretation of well logs (Second Edition).– Sutherland, Rider-French Consulting, 280 p.

RÖGL, F. (1999): Mediterranean and Paratethys – Facts and hypotheses of an Oligocene to Miocene paleogeography (short overview).– Geologica Carpathica, 50/4, 339–349.

RUKAVINA, D., SAFTIĆ, B., MATOŠ, B., KOLENKOVIĆ MOČILAC, I., PREMEC FUČEK, V. & CVETKOVIĆ, M. (2023): Tectonostratigraphic analysis of the syn-rift infill in the Drava Basin, southwestern Pannonian Basin System.– Marine and petroleum geology, 152, 106235, 23. https://doi.org/10.1016/j.marpetgeo.2023.106235

SAFTIĆ, B., VELIĆ, J., SZTANÓ, O., JUHÁSZ, G. & IVKOVIĆ, Ž. (2003): Tertiary subsurface facies, source rocks and hydrocarbon reservoirs in the SW part of the Pannonian Basin (Northern Croatia and south-western Hungary).– Geologia Croatica, 56, 101–122. https://doi.org/10.4154/232

SCHMID, S.M., BERNOULLI, D., FÜGENSCHUH, B., MATENCO, L., SCHEFER, S., SCHUSTER, R., TISCHLER, M. & USTASZEWSKI, K. (2008): The Alpine-Carpathian-Dinaridic orogenic system: Correlation and evolution of tectonic units.– Swiss Journal of Geosciences. 101, 139–183. https://doi.org/10.1007/s00015-008-1247-3

SERRA, O. (1984): 4. The Spontaneous Potential–SP.– In SERRA, O. (ed.): Developments in Petroleum Science, Elsevier, 15, Part A, 77–88. https://doi.org/10.1016/S0376-7361(08)70418-7

SMIRNOFF, A., BOISVERT, E. & PARADIS, S.J. (2008): Support vector machine for 3D modelling from sparse geological information of various origins.– Computers & Geosciences, 34/2, 127–143. https://doi.org/10.1016/j.cageo.2006.12.008

SUBRAHMANYAM, D. & RAO, P.H. (2008): Seismic attributes – A Review.– 7th International Conference & Exposition on Petroleum Geophysics, University of Hyderabad, Hyderabad, 398–403.

ŠPELIĆ, M., KOVÁCS, Á., SAFTIĆ, B. & SZTANÓ, O. (2023): Competition of deltaic feeder systems reflected by slope progradation: a high-resolution example from the Late Miocene-Pliocene, Drava Basin, Croatia.– International Journal of Earth Sciences, 113, 1–19. https://doi.org/10.1007/s00531-023-02290-w

TAHERI, M., CIABEGHODSI, A.A., NIKROUZ, R. & KADKHODAIE, A. (2021): Modeling of the Shale Volume in the Hendijan Oil Field Using Seismic Attributes and Artificial Neural Networks.– Acta Geologica Sinica (English Edition), 95/4, 1322–1331. https://doi.org/10.1111/1755-6724.14739

TANER, M.T. (2001): Seismic attributes.– Recorder, Canadian Society of Exploration Geophysicists, Houston, USA, 26/7, 26–48.

TANER, M.T., SHERIF, R.E., KOEHLER, E.D.F. (1976): Extraction and interpretation of the complex seismic trace.– In: 6th Annual Convention Proceedings, SEG, Huston, Texas, 2, 305–316.

VELIĆ, J. (2007): Petroleum Geology [*Geologija ležišta nafte i plina* – in Croatian].– in ZORIĆ, I. (ur.). University of Zagreb, Faculty of Mining, Geology and Petroleum Engineering, Zagreb.

VRSALJKO, D., PAVELIĆ, D., MIKNIĆ, M., BRKIĆ, M., KOVAČIĆ, M., HEĆIMOVIĆ, I., HAJEK-TADESSE, V., AVANIĆ, R. & KURTANJEK, N. (2006): Middle Miocene (Upper Badenian/Sarmatian) palaeoecology and evolution of the environments in the area of Medvednica Mt. (North Croatia).– Geologia Croatica, 59, 51–63.

VUKADIN, D. (2022): Development of seismic interpretation procedure based on seismic inversion for the Bjelovar depression reservoirs [*Razvoj postupka seizmičke interpretacije ležišta ugljikovodika Bjelovarske depresije na temelju inverzije seizmičkih brzina* – in Croatian].– PhD. University of Zagreb, Faculty of Mining, Geology and Petroleum Engineering, Zagreb.

XU, Q. & HAQ, U.B. (2022): Seismic facies analysis: Past, present and future.– Earth-Science Reviews, 224, 103876. https://doi.org/10.1016/j.earscirev.2021.103876

ZHOU, F. & LIU, L. (2024): Machine Learning Prediction of Deep Potential Ores and its Explanation Based on Integration of 3D Geological Model and Numerical Dynamics Simulation: An Example from Dongguashan Orefield, Tongling Copper District, China.– Natural Resources Research, 34, 121–147. https://doi.org/10.1007/s11053-024-10430-5

ZHOU, C., OUYANG, J., MING, W., ZHANG, G., DU, Z. & LIU, Z. (2019): A Stratigraphic Prediction Method Based on Machine Learning.– Applied Sciences, 9/17, 3553. https://doi.org/10.3390/app9173553

6. DISCUSSION

Objectives 1 and 2

(1) ***Reinterpret the lithological composition within the study area using available well data***

(2) Perform neural network analysis on 3D seismic and well data

Hypotheses 1 and 2

(1) ***It is expected that the use of artificial intelligence will advance geological characterization of the subsurface***

(2) It is assumed that the incorporation of 3D seismic and well data will greatly improve accuracy by removing bias in the interpretation

To assess the success of Objectives 1 and 2 a small area covering a depleted oil field was selected. This was also the case for testing Hypotheses 1 and 2 – which stated that applying artificial intelligence will enhance geological characterization of the subsurface and that integrating 3D seismic and well data will significantly improve accuracy by minimizing interpretational bias.

The initial dataset included only four wells and the 3D seismic volume of 36.26 km^3. This spatially limited dataset aimed to encompass the most comprehensive possible sets and sequences of well log measurements, as well as a high-quality and high-resolution seismic volume, which would ensure that the influence of data quality on model prediction capability is minimized. Pannonian sediments were characterized by identifying three primary lithologies (sandstone, marl, and coal) across four wells. Top and bottom model boundaries (top and bottom of Pannonian interval) were defined based on well data and mapped throughout the seismic volume. Seismic attributes such as RMS, Time Gain, Instantaneous Phase, Cosine of Phase, and RAI have demonstrated exceptional effectiveness in delineating boundaries that are particularly difficult to map using conventional seismic interpretation techniques (e.g. E-log horizon Rs_7). Sixteen seismic attributes were extracted to enhance input data for artificial neural network (ANN) analysis.

Lithology was upscaled within the model, and 800 data points were extracted for ANN analysis. The average layer thickness was 25 m, reflecting the vertical resolution of seismic data acquired in the 1990s (**Sheriff et al., 1992**). The ANN analysis involved extracting 15,480 data points with seismic attributes values. The fixed network parameters consisted of 16 neurons (seismic attributes) in the input layer and three neurons in the output layer (corresponding to three lithological categories). Two different stochastic approaches (DAANN and SAANN) were used to assess the impact of network architecture and data distribution on prediction accuracy. DAANN employed various architecture networks with differing hidden layers and activation functions, while SAANN used the same architecture but varied training, test, and selection datasets. To predict lithology, 2,000 ANNs were trained (1,000 DAANN and 1,000 SAANN), and the 100 most successful from each set (200 total) were selected for lithology prediction. ANN analysis produced numerous data points for variogram analysis. The number of these data points decreased as the probability of matching (P50, P75, P90) increased. After upscaling, variogram analysis was performed to assess the spatial relationships between lithological categories.

Results indicated that DAANN data is more susceptible to upscaling effects and tends to favor the dominant lithology (sandstone). The variogram analysis demonstrated that spatial correlation varies significantly based on data type and prediction accuracy. Due to the dominance of sandstone in P75 and P90 models, results should be interpreted cautiously, as the geological credibility of higher precision data may be compromised (**Cvetković et al., 2018; Kovács et al., 2019**).

Despite theoretical expectations that higher accuracy data yields more accurate results, the geologically sound results were achieved with 50% accuracy data. In contrast, higher accuracy results (P75 and P90) overemphasized sandstone lithology due to upscaling, experimental variogram calculation, and statistical analysis.

The first phase of the research revealed challenges in developing a novel methodological approach for characterizing lithology distribution within the subsurface using artificial neural networks (ANNs) even when applied to a comprehensive dataset of high-quality seismic and well data. This emphasizes the complexity of the problem with challenges primarily arising from inadequate data handling, particularly when synchronizing the depth domain (well data) with the time domain (seismic data), as many wells lack the time-to-depth relationship (TDR) data that could be obtained from vertical seismic profiling (VSP) or acoustic logs.

To overcome these issues, the second phase of research focused on developing a new methodological approach that ensures accurate conversion from depth-to-time domain. This is especially important for wells lacking VSP or acoustic logging, where previous conversions often involved interpolating between wells with and without established TDRs. The new approach accounts for both well and seismic data, employing ANNs to enhance accuracy and reduce errors caused by extrapolating velocity functions from nearby wells. This approach offers a cost-effective and efficient solution for subsurface exploration, surpassing traditional methods.

The innovative methodology presented here uses ANNs to predict TDRs based on well log interpretation and stratigraphic interval delineation. By doing so, it addresses issues related to time-to-depth domain conversion and significantly improves predictive accuracy, even for wells with limited geophysical data. Additionally, by integrating ANNs with lithological data from basic well logs – including those obtained from older wells – this research aimed to minimize errors and enhance accuracy in establishing time-to-depth relationships, which is expected to result in more accurate subsurface modeling.

The selection of a representative training dataset is crucial for the success of ANN predictions. A training dataset consisting of 18 wells was selected, which was split into 14 training wells and 4 holdout wells. The holdout wells (W-5, W-10, W-16, and W-18) were selected to represent diverse geological scenarios, including cases with all four stratigraphic intervals and cases where only the first two intervals were drilled trough by the well. Additionally, their geographical positions ensured that proximity to training wells could be assessed as a factor influencing prediction accuracy. The used dataset accurately reflects the geological and petrophysical complexities of the subsurface.

Well log interpretation focused on the differentiation between permeable and impermeable units using only basic well logs. Integrating these results with lithology data from master logs facilitated defining the vertical distribution of lithology. The interpretation of the well logs enabled the determination of three regional markers previously described by **Šimon (1980)** as α (Pliocene – Miocene boundary), Rs_7 (Middle – Late Miocene boundary) and PNg (Neogene basement). Based on three key boundaries, up to four stratigraphic intervals were interpreted for each well, grouping sediments by age, lithology, and depositional environment. These intervals define four stratigraphical units that influence petrophysical property variations with burial depth. The first consists of Pliocene-Quaternary unconsolidated sediments, the second of Upper Miocene sandstones and marls, the third of Lower and Middle Miocene heterogeneous deposits, and the fourth includes older pre-Neogene sediments and rocks forming the Basin Basement. The most significant velocity contrast is expected at the Neogene-

Basement transition, where compaction-governed lithologies are underlain by those where petrophysical properties are dominantly defined by secondary porosity.

The ANN analysis was conducted using TIBCO Statistica software, employing time series regression to ensure sequence prediction aligned with geological and geophysical principles. Neural networks were configured as multi-layer perceptron (MLP) architectures with a minimum of 3 and a maximum of 17 neurons in the hidden layer. To avoid overfitting, weight decay was applied, encouraging a simpler and more generalized model. In total, over 27,000 cases from 14 wells were used for training. Ten of the most successful neural networks were combined into an ensemble model for analysis. The ANN model achieved a correlation coefficient exceeding 0.99 for training, testing, and validation datasets, with a mean absolute error (MAE) of around 25 ms and a root mean square error (RMSE) of approximately 34 ms. To ensure the successfulness of the resulting ensemble it was tested on four holdout wells not included in the training process, demonstrating reliable performance with high accuracy.

Evaluation of prediction accuracy revealed that the ANN approach outperformed traditional TWT extrapolation methods in more than 75% of cases, with five wells demonstrating superior prediction success compared to extrapolation, achieving 100% better outcomes. Holdout wells also showed more accurate results, even when compared to predictions based on extrapolation from nearby wells. Interestingly, the analysis indicated that proximity between wells does not guarantee accurate extrapolation. This result highlights the complex relationship between spatial proximity and lithological variability, suggesting that geological relations and subsurface conditions play a more prominent role in prediction accuracy than mere distance.

The ANN model proved effective in generating TWT values, offering a promising alternative to conventional extrapolation methods. The methodology shows potential for application beyond the Drava Basin, especially in regions facing similar challenges in subsurface characterization. Additionally, the ANN-supported approach is capable of identifying previously unexplored correlations and dependencies within existing data, which could substantially enhance geoenergy exploration. Patterns proven in this research, although not directly recognizable within individual well or seismic datasets, align closely with previous studies documenting NNW-SSE structural trends and paleotransport sediment orientations (**Rukavina et al., 2023; Špelić et al., 2023; Matošević et al., 2024**). The results emphasize the reliability of ANN predictions compared to traditional methods, providing a valuable tool for future subsurface exploration efforts.

Objectives 1-3

(1) ***Reinterpret the lithological composition within the study area using available well data***

(2) Perform neural network analysis on 3D seismic and well data

(3) Make a geological model with distribution of the Pannonian sediments in the subsurface of Donji Miholjac area

Hypotheses 1 and 2

(1) ***It is expected that the use of artificial intelligence will advance geological characterization of the subsurface***

(2) It is assumed that the incorporation of 3D seismic and well data will greatly improve accuracy by removing bias in the interpretation

In the final phase of methodological development, Objectives 1–3 were successfully achieved: (1) reinterpret the lithological composition within the study area using available well data, (2) perform neural network analysis on 3D seismic and well data, and (3) construct a geological model illustrating the distribution of Pannonian sediments in the subsurface of the Donji Miholjac area. Additionally, Hypotheses 1 and 2 were confirmed, demonstrating that the application of artificial intelligence significantly enhances geological characterization of the subsurface and that the integration of 3D seismic and well data substantially improves accuracy by reducing interpretational bias.

Although it was demonstrated that ANNs can be successfully applied to address time-to-depth conversion challenges, predicting two-way time data from stratigraphic and petrophysical parameters in situations lacking conventional data (**Kamenski et al., 2024/Chapter 5.2**), some challenges remained, including the inadequate upscaling of well logs (**Kamenski et al., 2020/Chapter 5.1**). In response, this phase of the research applied innovative data preparation processes to improve ANN performance and predict lithological distribution using seismic attributes. Therefore, the study applied an ANN-supported modeling approach over a 4365 km² 3D seismic dataset targeting Pannonian (Late Miocene) sediments.

The volume of shale (V_{sh}), a continuous variable was introduced in opposition to regarding lithology directly as a categorical variable in **Chapter 5.1**. It was calculated from the Spontaneous Potential (SP) log. The usage of the Neutron log was excluded from analysis due to the limited intervals with recorded data in the wells, making it unsuitable for ensuring consistency across the investigated interval. Additionally, the Neutron log's shallow penetration depth and susceptibility to wellbore conditions, particularly in clayey rocks with increased well diameter, can result in misleading shale volume estimates (**Kamayou et al., 2021**). Similarly, the Resistivity log was not utilized because the mineral composition of the Pannonian sandstones (lithic arenites containing significant calcite, feldspar, and mica content) could produce misleading results. Feldspar-rich sandstones, such as the ones in the study area (**Matošević et al., 2023; Matošević et al., 2024)** exhibit notably lower resistivities compared to quartz-rich sandstones, potentially leading to overestimated shale volume (**Pimienta et al., 2019**). Gamma-ray (GR) logs were also unsuitable, as the potassium content of feldspar and mica within Upper Miocene sandstones can cause overestimated clay or shale volume values (**Imam & Trewin, 1991; Kamel & Mabrouk, 2003**). While thorium content from Spectral Radioactivity logs might theoretically serve as an alternative (**Rider, 2002**), their scarcity within the study area limited their practical application.

Despite certain limitations of the SP log – such as being unreliable in areas where formation water resistivity (Rw) closely matches mud filtrate resistivity (Rmf) (**Kamel & Mabrouk, 2003**) – this was not a significant issue in the research area, where formation water resistivity typically differed substantially from mud filtrate resistivity. Nevertheless, the shallow portions of the investigated interval were occasionally affected by lower total dissolved solids (TDS) in formation water, especially near the upper boundary of the model (Top Pannonian surface, well top "α"). Upscaled V_{sh} values were used as input variables for modeling, favoring continuity over discrete categorization, which is a step forward in the approach when compared to previous studies (**Brcković et al., 2017; Kamenski et al., 2020/Chapter 5.1**).

To address the inherent challenges of volume of shale estimation and interpreter subjectivity, four subsurface models with varying number of layers (20, 50, 100, 200) were constructed to optimize training and prediction processes. The model with 200 layers, averaging 6.5 meters in cell height, proved the most successful with minimizing the risk of overestimating dominant lithology by maintaining geological resolution and statistical validity. ANNs trained on models with fewer than 100 layers demonstrated poor performance or failed entirely, emphasizing the sensitivity of ANN training to input data resolution (**Alwosheel et al., 2018**).

For ANN training, 747,800 cells containing seismic attribute values were created, with the most effective ANN architecture deployed for shale volume prediction across the 3D seismic coverage. Twelve key seismic attributes – Sweetness, 3D Curvature, Variance, Original Amplitude, Instantaneous Frequency, Envelope, Instantaneous Phase, Generalized Spectral Decomposition, Apparent Polarity, Reflection Intensity, RMS Amplitude and Relative Acoustic Impedance – were selected based on their proven applicability from previous studies (**Chopra & Marfurt, 2006; Liu & Marfurt, 2006; Ker et al., 2014; Brcković et al., 2017; Li et al., 2019; Oumarou et al., 2021**). Utilization of standardized input data consistently produced the most accurate predictions, while directly using the raw data had the poorest performance. Predicted V_{sh} values were categorized into three lithological classes (Sandstone, Sandstone-Shale, and Shale) and upscaled for lithology modeling (Figure 4).

The lithological model derived from standardized input data contained 52.17% Shale, 29.89% Sandstone-Shale, and 17.94% Sandstone closely reflecting the lithological distribution observed in well data from the study area (**Kamenski et al., 2025**). In contrast, the model generated using normalized data yielded 55.14% Shale, 42.59% Sandstone-Shale, and 2.27% Sandstone – substantially underrepresenting the sandstone category. Similarly, the model based on raw input data produced 38.12% Shale, 48.57% Sandstone-Shale, and 13.3% Sandstone, which also deviates from lithological trends documented in both well and surface dataset observations (**Pavelić & Kovačić, 2018**).

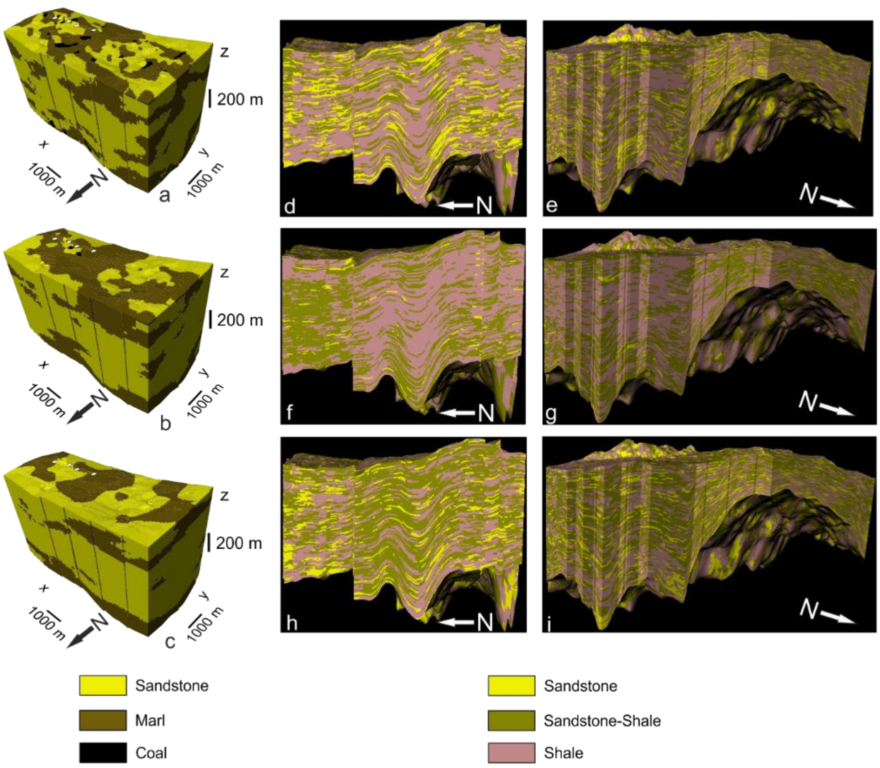


Figure 4. (a-c) Model results from the first phase of methodology development, showcasing the SAANN approach for cases P50 (a), P75 (b), and P90 (c). (d-i) Model results from the third phase of methodology development, illustrating outcomes using standardized input data (d, e), normalized data (f, g), and raw data (h, i).

This doctoral thesis demonstrated the successful application of continuous input data (V_{sh}) in artificial neural networks (ANNs), addressing a critical gap in previous research. A key aspect of the methodology was the upscaling process, which incorporated thin layering to avoid overestimation of dominant lithologies. Notably, models with layering thicker than 16 meters exhibited reduced performance or even complete failure in ANN predictions, highlighting the importance of fine-scale layering to maintain model accuracy.

Data preprocessing played a critical role in achieving accurate ANN predictions, with standardized input data proving most effective. Standardization maintained key numerical trends from input data that preserved the original lithology information and significantly improved the accuracy of predicted shale volume (V_{sh}) from seismic attributes. The resulting lithological model, derived from standardized data, demonstrated strong geological coherence that aligns with previous studies conserning deposition direction and facies of well-developed meander channels filled with sandstones (**Rukavina et al., 2023; Špelić et al., 2023; Matošević et al., 2024)**. In contrast, normalized data often distorted statistical relationships between variables, while raw data led to substantial underestimations of sandstone volumes, potentially compromising reservoir quality assessments and resource evaluations. These results emphasize the critical role of data standardization in ANN-supported lithological modeling and demonstrate the potential for the use of legacy seismic data in geoenergy explorations. These findings show the importance of selecting appropriate preprocessing techniques to achieve reliable and geologically accurate predictions in modeling lithology distribution.

By integrating machine learning with traditional geological methods, this dissertation demonstrated substantial improvements in predictive capabilities. The methodology proved particularly effective in settings with relatively uniform lithological compositions – typically two to three main lithologies distinguishable by geophysical well logs. Moreover, the proposed approach is well-suited for geoenergy applications such as hydrocarbon and geothermal exploration, as well as CO_2 storage assessments. However, in more geologically complex settings, further methodological refinement may be required to ensure optimal accuracy.

Objective 4

(4) Investigate the applicability of the method on 2D seismic data

Hypothesis 3

(3) It is speculated that research results will also be applicable on limited 2D seismic data, which are far more common in regional surveys

Finally, regarding Objective 4 and Hypothesis 3, a single 2D seismic section was selected to test the hypothesis concerning the applicability of the developed methodology to 2D seismic data. The selection of the section was based on the ability to compare the success of the ANN prediction. Thus, one seismic section which is partly covered by the 3D seismic volume and one well was selected. The location of this 2D section is presented in Figure 5, along with the modelled lithology prediction results.

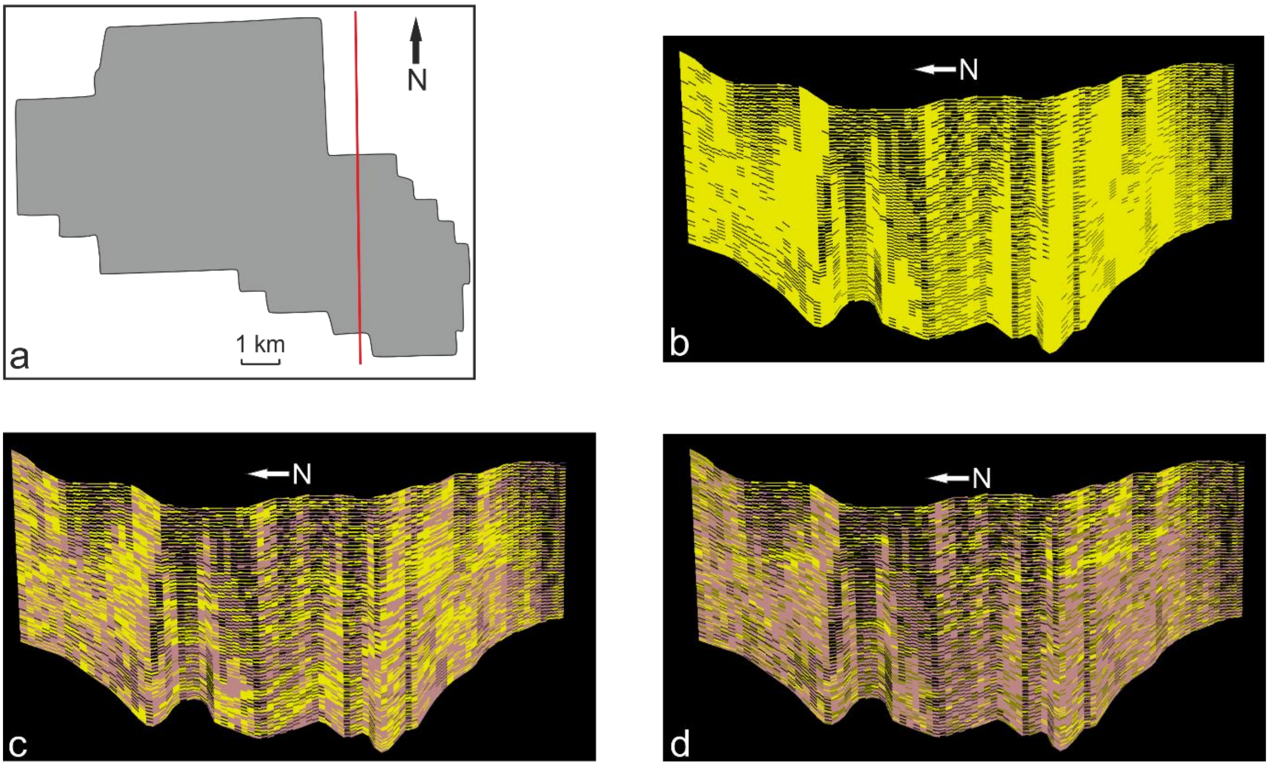


Figure 5. (a) Location of the 2D seismic section used for testing the developed methodology (red straight line). (b) Modelled lithology prediction results using raw data input. (c) Modelled lithology prediction results using normalized data input. (d) Modelled lithology prediction results using standardized data input.

The first step involved interpreting the top and bottom surfaces of the Pannonian interval (α and Rs_7). Following this, seismic attribute extraction was performed. Twelve seismic attributes, the same as in previous study, identified as highly effective for lithology prediction (see **Chapter 5.3**), were extracted from this section. These data were processed into three datasets: raw, normalized, and standardized. Each dataset was subsequently used as input data for ANN predictions. Previously developed, successful neural networks were employed to predict the volume of shale (V_{sh}) for each data type. The predicted shale volume was then categorized into three lithologies: Sandstone (≤0.5), Sandstone-Shale (0.5-0.7), and Shale (≥0.7).

Next, a model was developed following the parameters outlined in **Chapter 5.3**, and data from the three scenarios were upscaled. Once again, the final results proved the critical importance of data pre-processing in lithology distribution prediction.

The raw input data led to a significant overestimation of sandstone lithology, reaching a maximum of 100%, rendering it completely negligible for any interpretative considerations. On the other hand, normalized data considerably underestimated the Sandstone-Shale category (1.67%); however, the predicted distribution of Sandstone (43.36%) and Shale (64.97%) visually corresponded to a geologically reasonable setting, with distinguishable channel geometries. This outcome aligns well with the established geological framework (**Špelić et al., 2023; Pavelić & Kovačić, 2018**, and references therein).

Standardized data, while slightly overestimating Shale, produced the most geologically coherent results, with Sandstone (18.92%), Sandstone-Shale (19.91%), and Shale (61.17%) distribution ratios. These findings confirm that standardization significantly enhances the geological plausibility of predicted lithology distributions.

The developed methodology has been successfully applied to the 2D seismic section, demonstrating its feasibility. Results indicate a promising potential for reliable lithology distribution predictions, particularly for regional-scale studies. However, for investigations that require significantly more precise data than presented in this research, further refinement of the methodology will be necessary. This research demonstrates that applying ANN-supported methods to 2D seismic data can provide satisfactory subsurface imaging for broader regional assessments, paving the way for future advancements in subsurface characterization.

7. CONCLUSION

This dissertation presents the development of an AI-supported methodology that integrates artificial neural networks (ANNs) to improve subsurface characterization of the Pannonian-age clastic interval in the Drava Basin. The primary objective was to enhance the accuracy of geological models and mitigate economic and technical risks associated with mature basins, which often suffer from incomplete datasets. To achieve this, the approach involve integrating proven machine learning techniques with domain-specific geological knowledge. The research was structured into three phases: the first phase integrated ANNs with conventional geostatistical methods to enhance lithology prediction; the second phase developed an ANN-supported model for predicting time-to-depth relationships, demonstrating superior performance over traditional extrapolation techniques; and the third phase applied ANNs to lithology distribution modeling using well log data and 3D seismic attributes.

Hypothesis 1. It is expected that the use of artificial neural networks will advance geological characterization of the subsurface

Hypothesis 2. It is assumed that the incorporation of 3D seismic and well data will greatly improve accuracy by removing bias in the interpretation

An AI-supported methodology integrating artificial neural networks (ANNs) with well data and seismic attributes for subsurface characterization was developed and proven successful for lithology distribution modeling as well as for time-to-depth conversion.

ANN models successfully improved time-to-depth relationship predictions, incorporating stratigraphic and petrophysical parameters in situations with limited conventional data, achieving a correlation above 0.99 and outperforming traditional methods.

Lithology distribution modeling showed that standardized input data provided the most geologically consistent results, while raw and normalized data led to inconsistencies of model compared to observations in wells and regional studies.

ANN-supported models proved useful beyond lithological prediction, with potential applications in hydrocarbon exploration, geothermal assessment and CO_2 storage.

Hypothesis 3. It is speculated that research results will also be applicable on limited 2D seismic data, which are far more common in regional surveys

The methodology was successfully tested on 2D seismic data, which is of utmost significance for large areas in NCB covered only by 2D seismic data, but this approach is currently best suited for regional-scale analyses, i.e. level of detail of resulting model lacks high resolution compared to 3D seismic input data.

The ANN-supported models developed within this dissertation can be readily used in areas with similar geological settings. It should be emphasized that the main research object, the Pannonian sedimentary sequence in Drava Basin, is characterized by the relative simplicity of lithology distribution, so the lithology model would be best suited for analogues settings. Apart from the geological similarities, input data will have significant influence on model performance and other pre-processing methods may prove to be more effective for specific input dataset.

Tailoring ANN models to specific basins to ensure adaptability across different geological settings is expected to be needed. Future research directions should therefore aim to make the developed models suitable for different geological conditions:

- Expanding dataset diversity to minimize bias and improve model reliability, making more robust and generalizable predictions
- Integrating multiple AI-techniques to maximize the utilization of geological data, i.e. enhance geological insights to refine the model predictions
- Optimizing seismic attribute selection and neural network architectures to improve predictive accuracy for better adaptation to local specificities of research projects.

By actively improving current methods and refining modeling techniques, ANN-supported geological modeling is on track to become an indispensable tool for subsurface exploration and resource management. As methodologies evolve, applying AI-supported geological modeling to various sedimentary basins will further enhance prediction accuracy and support more informed decision-making processes in geoenergy research and resource management.

8. REFERENCES

Agatonovic-Kustrin, S. & Beresford, R. (2000) Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. *Journal of Pharmaceutical and Biomedical Analysis,* 22, 717–727. https://doi.org/10.1016/S0731 -7085(99)00272 -1

Aker, E., Tveiten, O.G. & Wynn, T. (2020). Seismic pore pressure prediction at the Halten Terrace in the Norwegian Sea. *Petroleum Geoscience*, 26(2), 346-354.

Alwosheel, A., Van Cranenburgh, S. & Chorus, C.G. (2018) Is your dataset big enough? Sample size requirements when using artificial neural networks for discrete choice analysis. *Journal of Choice Modelling,* 28, 167–182. https://doi.org/10.1016/j.jocm.2018.07.002

Anderson, J.A. & Rosenfeld, E. (1988) *Neurocomputing: foundations of research*. MIT Press, Cambridge.

Asquith, G. & Krygowski, D. (2004): Basic Well Log Analysis. *AAPG Methods in Exploration* 16, 21-30.

Balázs, A., Matenco, L., Magyar, I., Horváth, F & Cloetingh, S. (2016) The link between tectonics and sedimentation in back-arc basins: New genetic constraints from the analysis of the Pannonian Basin. *Tectonics*, 35, 1526-1559.

Balázs, A., Magyar, I., Matenco, L., Sztanó, O., Tőkés, L. & Horváth, F. (2018) Morphology of a large paleo-lake: Analysis of compaction in the Miocene-Quaternary Pannonian Basin. *Global and Planetary Change*, 171, 134-147.

Bassiouni, Z. (2013) *Theory, Measurement, and Interpretation of Well Logs* (4^{th} edition). Texas, USA, Society of Petroleum Engineers.

Bianchi, M., Kearsey, T. & Kingdon, A. (2015) Integrating deterministic lithostratigraphic models in stochastic realizations of subsurface heterogeneity. Impact on predictions of lithology, hydraulic heads and groundwater fluxes. *Journal of Hydrology*, 531(3), 557-573. https://doi.org/10.1016/j.jhydrol.2015.10.072

Bigdeli, A. & Delshad, M. (2024) The Evolving Landscape of Oil and Gas Chemicals: Convergence of Artificial Intelligence and Chemical-Enhanced Oil Recovery in the Energy Transition Toward Sustainable Energy Systems and Net-Zero Emissions. *Journal of Data Science and Intelligent Systems*, 2(2), 65-78. https://doi.org/10.47852/bonviewJDSIS42022111

Bonafe, A., Galeati, G. & Sforna, M. (1994) Neural networks for daily mean flow forecasting. *Hydrology and Engineering Software*, 1, 131-138.

Brcković, A., Kovačević, M., Cvetković, M., Kolenković Močilac, I., Rukavina, D. & Saftić, B. (2017) Application of artificial neural networks for lithofacies determination based on limited well data. *Central European Journal of Geosciences*, 60 (3), 299-315. https://doi.org/10.1556/24.60.2017.012

Brlek, M., Trinajstić, N., Gaynor, S.P., Kutterolf, S., Hauff, F., Schindlbeck-Belo, J., Šuica, S., Wang, K.-L., Lee, H.-Y., Watts, E., Georgiev, S.V., Brčić, V., Špelić, M., Mišur, I., Kukoč, D., Schoene, B. & Lukács, R. (2024) Spread and frequency of explosive silicic volcanism of the Carpathian-Pannonian Region during Early Miocene: Clues from the SW Pannonian Basin and the Dinarides. Journal of volcanology and geothermal research, 455, 108215, 22. doi: 10.1016/j.jvolgeores.2024.108215

Cao, J., Shi, Y., Wang, D. & Zhang, X. (2017) Acoustic Log Prediction on the Basis of Kernel Extreme Learning Machine for Wells in GJH Survey, Erdos Basin. *Journal of Electrical and Computer Engineering*, 2017, 3824086. https://doi.org/10.1155/2017/3824086

Chopra, S. & Marfurt, K.J. (2006) Seismic attributes — a historical perspective. *Geophysics*, 70/5, 3SO–28SO.

Chopra, S. & Marfurt, K.J. (2008) Emerging and future trends in seismic attributes. *The Leading Edge*, 27(3), 298–318. https://doi.org/10.1190/1.2896620

Cressie, N. (1990) The origins of kriging. *Mathematical Geosciences*, 22, 239–252. https://doi.org/10.1007/BF008 89887

Croatian Geological Survey (2009) Geological Map of Republic of Croatia, M 1:300.000. Croatian Geological Survey, Zagreb.

Cvetković, M., Emanović, I., Stopar, A. & Slavinić, P. (2018) Petroleum system modeling and assessment of the remaining hydrocarbon potential in the eastern part of Drava Depression. *Interpretation*, 6, 1, SB11-SB21. https://doi.org/10.1190/int-2017-0078.1

Cvetković, M., Matoš, B., Rukavina, D., Kolenković Močilac, I., Saftić, B., Baketarić, T., Baketarić, M., Vuić, I., Stopar, A., Jarić, A. & Paškov, T. (2019) Geoenergy potential of the Croatian part of Pannonian Basin: insights from the reconstruction of the pre-Neogene basement unconformity. *Journal of Maps*, 15/2, 651–661. https://doi.org/10.1080/17445647.2019.1645052

Djeddi, M. (2016) *Transformée de Hilbert, signal analytique et attributs sismiques*. Département de Géophysique, FHC-Université M'Hamed Bougara de Boumerdes, Algérie.

Dubrule, O. (1993) *Introducing more geology in stochastic reservoir modelling*. In: Soares, A. (ed.) Geostatistics Tróia '92, pp. 351–369. Springer, Dordrecht.

Falivene, O., Cabrera, L. & Sáez, A. (2007) Optimum and robust 3D facies interpolation strategies in a heterogeneous coal zone (Tertiary As Pontes basin, NW Spain). *International Journal of Coal Geology*, 71, 2-3, 185-208. https://doi.org/10.1016/j.coal.2006.08.008

Feng, R., Luthi, S.M., Gisolf, D. & Angerer, E. (2018) Reservoir lithology classification based on seismic inversion results by Hidden Markov Models: applying prior geological information. *Marine and Petroleum Geology*, 93, 218–229. https://doi.org/10.1016/j.marpe tgeo.2018.03.004

Feng, Y., Wen, G., Shang, J., Wen S. & Wu, B. (2024) Research on 3D geological modeling based on boosting integration strategy. *Ore Geology Reviews*, 171, 106157, https://doi.org/10.1016/j.oregeorev.2024.106157

Forgotson, J.M. (1960): Review and classification of quantitative mapping techniques. *American Association of Petroleum Geologists Bulletin*, 44, 83–100.

Gradstein, F.M, Ogg, J.G., Schmitz, M.D., Ogg, G.M., Agterberg, F.P., Anthonissen, D.E., Becker, T.R., Catt, J.A., Cooper, R.A., Davydov, V.I., Gradstein, S.R., Henderson, C.M., Hilgen, F.J., Hinnov, L.A., McArthur, J.M., Melchin, M.J., Narbonne, G.M., Paytan, A., Peng, S., Peucker Ehrenbrink, B., Pillans, B., Saltzman, M.R., Simmons, M.D., Shields, G.A., Tanaka, K.L., Vandenberghe, N., Van Kranendonk, M.J., Zalasiewicz, J., Altermann, W., Babcock, L.E., Beard, B.L., Beu, A.G., Boyes, A.F., Cramer, B.D., Crutzen, P.J., van Dam, J.A., Gehling, J.G., Gibbard, P.L., Gray, E.T., Hammer, O., Hartmann, W.K., Hill, A.C., Hoffman, P.F., Hollis, C.J., Hooker, J.J., Howarth, R.J., Huang, C., Johnson, C.M., Kasting, J.F., Kerp, H., Korn, D., Krijgsman, W., Lourens, L.J., MacGabhann, B.A., Maslin, M.A., Melezhik, V.A., Nutman, A.P., Papineau, D., Piller, W.E., Pirajno, F., Ravizza, G.E., Sadler, P.M., Speijer, R.P., Steffen, W., Thomas, E., Wardlaw, B.R., Wilson, D.S. & Xiao, S. (2012) *The Geologic Time Scale*. Elsevier, 1129 p., Boston.

Hart, D.M., Balch, R.S., Weiss, W.W. & Wo, S. (2000) Time-to-Depth Conversion of Nash Draw "L" Seismic Horizon using Seismic Attributes and Neural Networks. Paper Presented at SPE Permian Basin Oil and Gas Recovery Conference, Midland, Texas, 21-23 March, 2000. https://doi.org/10.2118/59555-MS

Hatampour, A., Schaffie, M. & Jafari, S. (2018) Hydraulic flow units' estimation from seismic data using artificial intelligence systems, an example from a gas reservoir in the Persian Gulf. *Journal of Petroleum Science and Engineering*, 170, 400-408. https://doi.org/10.1016/j.petrol.2018.06.086

Hengl, T., Toomanian, N., Reuter, H.I. & Malakouti, M.J. (2007) Methods to interpolate soil categorical variables from profile observations: Lessons from Iran. Geoderma, 140(4), 417-427. https ://doi.org/10.1016/j.geoderma.2007.04.022

Hernitz, Z. (1980) *Subsurface structure-tectonic relations in the area of eastern Slavonia* (In Croatian: Dubinski strukturno—tektonski odnosi u području istočne Slavonije).

Hohn, M.E. (1999) *Geostatistics and Petroleum Geology*. Springer, Dordrecht, 235 p. https://doi.org/10.1007/978-94-011-4425-4

Horváth, F., Bada, G., Szafian, P, Tari, G., Adam, A. & Cloetingh, S. (2006) Formation and deformation of the Pannonian basin: constrains from observational data. *Geological Society London Memoirs*, 32, 191-206.

Horváth, F., Mustiz, B., Balázs, A., Végh, A., Uhrin, A., Nádor, A., Koroknai B., Pap, N., Tóth, T. & Wórum, G. (2015) Evolution of the Pannonian basin and its geothermal resources. *Geothermics*, 53, 328-352.

Imam, Md.B. & Trewin, N.H. (1991) Factors contributing to high gamma-ray levels in Upper Jurassic sandstone reservoirs of the Claymore Oilfield, North Sea. *Marine and Petroleum Geology*, 8/4, 452–460. https://doi.org/10.1016/0264-8172(91)90067-B

Irrgang, C., Boers, N., Sonnewald, M., Barnes, E.A., Kadow, C., Staneva, J. & Saynisch-Wagner, J. (2021) Towards neural Earth system modelling by integrating artificial intelligence in Earth system science. *Nature Machine Intelligence*, 3, 667–674. https://doi.org/10.1038/s42256-021-00374-3

Iturrarán-Viveros, U. & Parra, J.O. (2014) Artificial Neural Networks applied to estimate permeability, porosity and intrinsic attenuation using seismic attributes and well-log data. *Journal of Applied Geophysics*, 107, 45–54. https://doi.org/10.1016/j.jappgeo.2014.05.010

James, G., Witten, D., Hastie, T. & Tibshirani, R. (2013) *An Introduction to Statistical Learning*. Springer, New York.

Johnson, V.M. & Rogers, L.L. (1995) Location analysis in ground-water remediation using neural networks. *Groundwater*, 33(5), 749-758.

Journel, A.G. (1983) Nonparametric estimation of spatial distributions. International Association for Mathematical Geosciences, 15, 445–468. https://doi.org/10.1007/BF010 31292

Kamayou, V., Ehirim, C. & Ikiensikimama, S. (2021) Estimating Volume of Shale in a Clastic Niger Delta Reservoir from Well Logs: A Comparative Study. International Journal of Geosciences, 12, 949–959. https://doi.org/10.4236/ijg.2021.1210049

Kamel, M.H. & Mabrouk, W.M. (2003) Estimation of shale volume using a combination of the three porosity logs. *Journal of Petroleum Science and Engineering*, 40/3-4, 145–157. https://doi.org/10.1016/S0920-4105(03)00120-7

Kamenski, A., Cvetković, M., Kolenković Močilac, I. & Saftić, B. (2020): Lithology prediction in the subsurface by artificial neural networks on well and 3D seismic data in clastic sediments: a stochastic approach to a deterministic method. *GEM - International journal on geomathematics*, 11/8, 1–24. https://doi.org/10.1007/s13137-020-0145-3

Kamenski, A., Cvetković, M., Kapuralić, J., Kolenković Močilac, I. & Brcković, A. (2024) From Traditional Extrapolation to Neural Networks: Time-Depth Relationship Innovations in the Subsurface Characterization of Drava Basin, Pannonian Super Basin. *Advances in geo-energy research*, 14/1, 25–33. https://doi.org/10.46690/ager.2024.10.05

Kanevski, M. & Dumolard, P. (2008) Advanced Mapping of Environmental Data. ISTE, London.

Kappen, H.J. (2001) *Chapter 13 An introduction to stochastic neural networks*. In: Moss, F., Gielen, S. (eds.): Neuro-Informatics and Neural Modelling, 517–552 p. Elsevier.

Karpatne, A., Ebert-Uphoff, I., Ravela, S., Babaie, H.A. & Kumar, V. (2019) *Machine Learning for the Geosciences: Challenges and Opportunities*. In: IEEE Transactions on Knowledge and Data Engineering, 31(8), 1544-1554. doi: 10.1109/TKDE.2018.2861006.

Ker, S., Le Gonidec, Y., Marsset, B., Westbrook, G.K., Gibert, D. & Minshull, T.A. (2014) Fine-scale gas distribution in marine sediments assessed from deep-towed seismic data. *Geophysical Journal International*, 196/3, 1466– 1470. https://doi.org/10.1093/gji/ggt497

Khan, K.A. & Akhter, G. (2015) Review of instantaneous, wavelet, and weighted seismic attributes along with a computational library. *Arabian Journal of Geosciences*, 2015. doi: 10.1007/s12517-015-2044-8

Khosravi, A. & Ashkpour, M. (2024) *Machine Learning and Digital Innovation for Managing and Monitoring Water Resources*. In: Moussaoui, M. & Rachid, A. (eds.): Emerging Trends and Technologies in Water Management and Conservation. IGI Global, 241-284. doi: 10.4018/979-8-3693-6920-3.ch008

Kovács, Á., Špelić, M., Balázs, A., Magyar, I. & Sztanó, O. (2019) *Late Miocene variation of accommodation in the Drava Trough: stratigraphic forward modeling and field observations*. In: Tari, G. & Sachsenhofer, R. (eds.) Paratethys Petroleum Systems Between Central Europe and the Caspian Region, 149 p. AAPG Europe, Vienna.

Krige, D.G. (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(6), 119–139.

Krijgsman, W., Stoica, M., Vasiliev, I. & Popov, V.V. (2010) Rise and fall of the Paratethys Sea during the Messinian Salinity Crisis. *Earth and Planetary Science Letters*, 290, 183–191. https://doi.org/10.1016/j.epsl.2009.12.020

Lean, Y., Shouyang, W. & Kin, K.L. (2007) *Basic learning principles of artificial neural networks*. In: Hillier, F. (ed.): Foreign-Exchange-Rate Forecasting with Artificial Neural Networks, pp. 27–37. Springer, New York.

Li, W., Yue, D., Wu, S., Shu, Q., Wang, W., Long, T. & Zhang, B. (2019) Thickness prediction for high-resolution stratigraphic interpretation by fusing seismic attributes of target and neighboring zones with an SVR algorithm. *Marine and Petroleum Geology*, 113, 104153. https://doi.org/10.1016/j.marpetgeo.2019.104153

Liu, J. & Marfurt, K.J. (2006): Thin bed thickness prediction using peak instantaneous frequency. *SEG Technical Program Expanded Abstracts*, Society of Exploration Geophysicists, 968–972.

Lučić, D., Saftić, B., Krizmanić, K., Prelogović, E., Britvić, V., Mesić, I. & Tadej, J. (2001) The Neogene Evolution and Hydrocarbon Potential of the Pannonian Basin in Croatia. *Marine and Petroleum Geology*, 18/1, 133-147. https://doi.org/10.1016/S0264-8172(00)00038-6

Magyar, I., Radivojević, D., Sztanó, O., Synak, R., Ujszaszi, K. & Pocsik, M. (2013) Progradation of the paleo-Danube shelf margin across the Pannonian Basin during the Late Miocene and Early Pliocene. *Global and Planetary Change*, 103, 168-173.

Maier, H.R. & Dandy, G.C. (1996) The use of artificial neural networks for the prediction of water quality parameters. *Water resources research*, 32(4), 1013-1022.

Malvić, T. & Cvetković, M. (2013) Lithostratigraphic units in the Drava Depression (Croatian and Hungarian parts) – a correlation. *Nafta*, 64/1, 27–33.

Mari, J. L., Vergniault, C. & Coppens, F. *Acoustic logging*. In: Well seismic surveying and acoustic logging, Mari, J.L. & Verginault, C. (eds), EDP Sciences, Paris, 77-102 p.

Matheron, G. (1965) Les variables régionalisées et leur estimation. Une application de la théorie des fonctions aléatoires aux sciences de la nature. Masson, Paris.

Matošević, M., Marković, F., Bigunac, D., Šuica, S., Krizmanić, K., Perković, A., Kovačić, M. & Pavelić, D. (2023) Petrography of the Upper Miocene sandstones from the North Croatian Basin: Understanding the genesis of the largest reservoirs in the southwestern part of the Pannonian Basin System. *Geologica Carpathica*, 74/2, 155–179. https://doi.org/10.31577/GeolCarp.2023.06

Matošević, M., Tomašić, N., Perković, A., Kampić, Š., Kovačić, M. & Pavelić, D. (2024) Reservoir Quality Evaluation: Unveiling Diagenetic Transformations through Mineralogical and Petrophysical Analyses of the Upper Miocene Lacustrine Sandstones in the Pannonian Basin System, Croatia. *Rudarsko-Geološko-Naftni Zbornik*, 39/3, 153–172. https://doi.org/10.17794/rgn.2024.3.12

McCulloch, W.S. & Pitts, W. (1943) A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biology*, 5, 115–133. https://doi.org/10.1007/BF02478259

Mohammadinia, F., Ranjbar, A., Kafi, M., Shams, M., Haghighat, F. & Maleki, M. (2023) Shale volume estimation using ANN, SVR, and RF algorithms compared with conventional methods. *Journal of African Earth Sciences*, 205, 104991. https://doi.org/10.1016/j.jafrearsci.2023.104991

Muttiah, R.S., Srinivasan, R. & Allen, P. M. (1997) Prediction of two-year peak stream-discharges using neural networks. *Journal of the American Water Resources Association*, 33(3), 625-630.

Nikravesh, M. & Aminzadeh, F. (2003) *Chapter 1 Soft computing for intelligent reservoir characterization and modeling*. In: Nikravesh, M., Aminzadeh, F., Zadeh, L.A. (eds.) Soft Computing and Intelligent Data Analysis in Oil Exploration, 3–32 p. Elsevier.

Nowak, M. & Verly, G. (2005) T*he practice of sequential Gaussian simulation*. In: Leuangthong, O., Deutsch, C.V. (eds.) Geostatistics Banff, 387–398 p. Springer, Dordrecht.

Othman, A., Fathy, M. & Mohamed, I.A. (2021) Application of Artificial Neural Network in seismic reservoir characterization: a case study from Offshore Nile Delta. *Earth Science Informatics*, 14, 669–676. https://doi.org/10.1007/s12145-021-00573-x

Oumarou, S., Mabrouk, D., Tabod., T.C., Marcel, J., Ngos Iii, S., Essi, J.M.A. & Kamguia, J. (2021) Seismic attributes in reservoir characterization: an overview. *Arabian Journal of Geosciences*, 14, 402. https://doi.org/10.1007/s12517-021-06626-1

Pamić, J. (1984) Triassic magmatism of the Dinarides in Yugoslavia. *Tectonophysics*, 109(3-4), 273-277, 281-307. https://doi.org/10.1016/0040-1951(84)90145-8

Pamić, J. (1998) Crystalline basement of the South Pannonian Basin based on surface and subsurface dana, *Nafta*, 49, 371–390.

Pamić, J. & Jurković, I. (2002) Paleozoic tectonostratigraphic units of the northwest and central Dinarides and the adjoining South Tisia. *International Journal of Earth Sciences*, 91, 538-554. https://doi.org/10.1007/s00531-001-0229-8

Paulsen, M., Rasmussen, T.K. & Nielsen, O.A. (2018) Output variability caused by random seeds in a multi-agent transport simulation model. *Procedia Computer Science*, 130, 850–857. https ://doi.org/10.1016/j.procs .2018.04.078

Pavelić, D. & Kovačić, M. (2018) Sedimentology and stratigraphy of the Neogene rift-type North Croatian Basin (Pannonian Basin System, Croatia): A review. *Marine and Petroleum Geology*, 91, 455–469. https://doi.org/10.1016/j.marpetgeo.2018.01.026

Pimienta, L., Orellana, L.F. & Violay, M. (2019) Variations in elastic and electrical properties of crustal rocks with varying degree of microfracturation. Journal of Geophysical Research: Solid Earth, 124, 6376–6396. https://doi.org/10.1029/2019JB017339

Rider, M.H. (2002): *The geological interpretation of well logs* (Second Edition). Sutherland, Rider-French Consulting, 280 p.

Rögl, F. & Steininger, F.F. (1984) Neogene paratethys, Mediterranean and Indo-pacific Seaways. U: Brenchley, P.J. (ur.): Fossils and Climate, London, Wiley.

Rosenblatt, F. (1958) The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review*, 65, 386–408. https://doi.org/10.1037/h0042 519

Royden, L. (1988) *Late Cenozoic tectonics of the Pannonian Basin system*. In: Royden, L.H., Horvath, F. (eds.): The Pannonian Basin: A Study in Basin Evolution. AAPG Memoir 45. AAPG, Tulsa.

Rukavina, D. (2021) Rift tectonostratigraphic sequences of Lower and Middle Miocene in eastern part of the Drava Depression: application to CO2 geological storage assessment (in Croatian). *Dissertation*, 329. University of Zagreb, Faculty of Mining, Geology and Petroleum Engineering, Zagreb, Croatia.

Rukavina, D., Saftić, B., Matoš, B., Kolenković Močilac, I., Premec Fuček, V. & Cvetković, M. (2023) Tectonostratigraphic analysis of the syn-rift infill in the Drava Basin, southwestern Pannonian Basin System. *Marine and petroleum geology*, 152, 106235, 23. https://doi.org/10.1016/j.marpetgeo.2023.106235

Saftić, B., Velić, J., Sztanó, O., Juhász, Gy. & Ivković, Ž. (2003) Tertiary subsurface facies, source rocks and hydrocarbon reservoirs in the SW part of the Pannonian Basin (northern Croatia and south-western Hungary). *Geologia Croatica*, 56/1, 102-122.

Saftić, B., Kolenković Močilac, I., Cvetković, M. & Vulin, D. (2024) CO2 geological storage potential in the Croatian part of the Pannonian Superbasin. *Special publication - Geological Society of London*, 555, 1; 555-24, 11. https://doi.org/10.1144/SP555-2024-23

Sagi, T., Lehahn, Y. & Bar, K (2020) Artificial intelligence for ocean science data integration: current state, gaps, and way forward. *Elementa: Science of the Anthropocene*, 8(21). https://doi.org/10.1525/elementa.418

Sebe, K., Kovačić, M., Magyar, I., Krizmanić, K., Špelić, M., Bigunac, D., Sütő-Szentai, M., Kovács, A., Szuromi-Korec, A., Bakrač, K., Hajek-Tadesse, V., Troskot-Čorbić, T. & Sztanó; O. (2020) Correlation of upper Miocene–Pliocene Lake Pannon deposits across the Drava Basin, Croatia and Hungary. *Geologia Croatica*, 73, 3; 177-195. doi: 10.4154/gc.2020.12

Selley, R.C. & Sonnenberg, S.A. (2015) Methods of Exploration. Elements of Petroleum Geology, 41–152 p. Elsevier, New York.

Serra, O. (1984) *4. The Spontaneous Potential—SP*.– In SERRA, O. (ed.): Developments in Petroleum Science, Elsevier, 15, Part A, 77–88. https://doi.org/10.1016/S0376-7361(08)70418-7

Sheriff, R.E. (1992): *Vertical and lateral seismic resolution and attenuation: part 7*. Geophysical methods. In: Morton-Thompson, D., Woods, A.M. (eds.) ME 10: Development Geology Reference Manual, 388–389 p. AAPG, Tulsa.

Smirnoff, A., Boisvert, E. & Paradis, S.J. (2008) Support vector machine for 3D modelling from sparse geological information of various origins. *Computers & Geosciences*, 34/2, 127–143. https://doi.org/10.1016/j.cageo.2006.12.008

Sztanó, O., Sebe, K., Csillag & Magyar, I. (2015) Turbidites as indicators of paleotopography, Upper Miocene Lake Pannon, Western Mecsek Mountains (Hungary). *Geologica Carpathica*, 66(4). doi: 10.1515/geoca-2015-0029

Sun, Z., Sandoval, L., Crystal-Ornelas, R., Mostafa Mousavi, S., Wang, J., Lin, C., Cristea, N., Tong, D., Hawley Carande, W., Ma, X., Rao, Y., Bednar, J.A., Tan, A., Wang, J., Purushotham, S., Gill, T.E., Chastang, J., Howard, D., Holt, B., Gangodagamage, C., Zhao, P., Rivas, P., Chester, Z., Orduz, J. & John, A. (2022) A review of Earth Artificial Intelligence. *Computers & Geosciences*, 159, 105034. https://doi.org/10.1016/j.cageo.2022.105034

Sun, Z., Yang, S., Zhang, F., Lu, J., Wang, R., Ou, X., Lei, A., Han, F., Cen, W., Wei, D. & Liu, M. (2023) A Reconstructed Method of Acoustic Logging Data and Its Application in Seismic Lithological Inversion for Uranium Reservoir. *Remote Sensing* (Basel), 15, 1260, https://doi.org/10.3390/rs15051260

Šimon, J. (1980): *Prilog stratigrafiji u taložnom sustavu pješčanih rezervoara Sava-grupe naslaga mlađeg tercijara u Panonskom bazenu sjeverne Hrvatske* (in Croatian). Disertacija, RGN fakultet Sveučilišta u Zagrebu, 66 str.

Špelić, M. (2023) Depositional dynamics of Upper Miocene and Pliocene sediments in the eastern part of Drava Depression (in Croatian). *Dissertation*, 313. University of Zagreb, Faculty of Mining, Geology and Petroleum Engineering, Zagreb, Croatia.

Špelić, M., Kovács, Á., Saftić, B. & Sztanó, O. (2023) Competition of deltaic feeder systems reflected by slope progradation: a high-resolution example from the Late Miocene-Pliocene, Drava Basin, Croatia. *International journal of earth sciences*, 113, 1–19. https://doi.org/10.1007/s00531-023-02290-w

Taheri, M., Ciabeghodsi, A.A., Nikrouz, R. & Kadkhodaie, A. (2021) Modeling of the Shale Volume in the Hendijan Oil Field Using Seismic Attributes and Artificial Neural Networks. *Acta Geologica Sinica* (English Edition), 95/4, 1322–1331. https://doi.org/10.1111/1755-6724.14739

Taner, M.T. (2001) *Seismic attributes*. Recorder, Canadian Society of Exploration Geophysicists, Houston, USA, 26/7, 26–48.

Taner, M.T., Sherif, R.E. & Koehler, E.D.F. (1976) *Extraction and interpretation of the complex seismic trace*. In: 6th Annual Convention Proceedings, SEG, Huston, Texas, 2, 305–316.

Taner, M.T., Schuelke, J.S., O’Doherty, R. & Baysal, E. (1994) *Seismic attributes revisited*. In: 1994 SEG Annual Meeting. 1104– 1106.

Tari, G., Horváth, F. & Rumpler, J. (1992) Styles of extension in the Pannonian Basin. *Tectonophysics*, 208, 203–219. https://doi.org/10.1016/0040-1951(92)90345-7

Valizadeh, N., Mirzaei, M., Allawi, M.F., Afan, H.A., Mohd, N.S., Hussain, A. & El-Shafie, A. (2017) Artificial intelligence and geo-statistical models for stream-flow forecasting in ungauged stations: state of the art. *Natural Hazards*, 86, 1377–1392. https://doi.org/10.1007/s11069-017-2740-7

Velić. J. (2007) *Geologija ležišta nafte i plina* (in Croatian). Sveučilište u Zagrebu, Rudarsko-geološko-naftni fakultet, 342 p.

Velić, J., Cvetković, M., Malvić, M. & Vrbanac, B. (2010) Characteristics of hydrocarbon fields in the Croatian part of the Pannonian basin. *Oil, gas*, 36(3), 146-147. Hamburg, Germany.

Velić, J., Krasić, D. & Kovačević, I. (2012a) Exploitation, reserves and transport of natural gas in the Republic of Croatia (in Croatian). *Tehnicki Vjesnik*, 13, 633–641.

Velić, J., Malvić, T., Cvetković, M. & Vrbanac, B. (2012b) Reservoir geology, hydrocarbon reserves and production in the Croatian part of the Pannonian Basin System. *Geologia Croatica*, 65. https://doi.org/10.4154/gc.2012.07

Vulin, D., Kolenković Močilac, I., Jukić, L., Arnaut, M., Vodopić, F., Saftić, B., Karasalihović Sedlar, D. & Cvetković, M. (2023) Development of CCUS clusters in Croatia. *International Journal of Greenhouse Gas Control*, 124, 103857, 13. https://doi.org/10.1016/j.ijggc.2023.103857

Wacha, L., Galović, L., Koloszár, L., Magyari, Á., Chikán, G. & Marsi, I. (2013) The chronology of the Šarengrad II loess-palaeosol section (Eastern Croatia). *Geologia Croatica*, 66, 191–203. https://doi.org/10.4154/GC.2013.18

Walton, C., Evans, B. & Urosevic, M. (2000) Imaging coal seam structure using 3-D seismic methods. *Exploration Geophysics*, 31, 509–514. https ://doi.org/10.1071/EG005 09

Zhao, T., Wang, S., Ouyang, C., Chen, M., Liu, C., Zhang, J., Yu, L., Wang, F., Xie, Y., Li, J., Wang, F., Grunwald, S., Wong, B.M. Zhang, F., Qian, Z., Xu, Y., Yu, C., Han, W., Sun, T., Shao, Z., Qian, T., Chen, Z., Zeng, J., Zhang, H., Letu, H., Zhang, B., Wang, L., Luo, L., Shi, C., Su, H., Zhang, H., Yin, S., Huang, N., Zhao, W., Li, N., Zheng, C., Zhou, Y., Huang, C., Feng, D., Xu, Q., Wu, Y., Hong, D., Wang, Z., Lin, Y., Zhang, T., Kumar, P., Plaza, A., Chanussot, J., Zhang, J., Shi, J. & Wang, L. (2024) Artificial intelligence for geoscience: Progress, challenges, and perspectives. *The Innovation*, 5(5), 100691. https://doi.org/10.1016/j.xinn.2024.100691

Zhou, F. & Liu, L. (2024) Machine Learning Prediction of Deep Potential Ores and its Explanation Based on Integration of 3D Geological Model and Numerical Dynamics Simulation: An Example from Dongguashan Orefield, Tongling Copper District, China. Natural Resources Research, 34, 121–147. https://doi.org/10.1007/s11053-024-10430-5

Zhou, C., Ouyang, J., Ming, W., Zhang, G., Du, Z. & Liu, Z. (2019) A Stratigraphic Prediction Method Based on Machine Learning. *Applied Sciences*, 9/17, 3553. https://doi.org/10.3390/app9173553

9. BIOGRAPHY OF THE AUTHOR

Ana Kamenski was born on November 6, 1994, in Zagreb. After graduating from V. Gymnasium in Zagreb, she enrolled in undergraduate program in Geological Engineering at the Faculty of Mininig, Geology and Petroleum Engineering in 2013. She completed undergaduate studies with thesis *Refraction exploration on Gomirje railway notch on Zagreb GK-Karlovac-Rijeka railway*. In 2016, she obtained her master's degree in Geology (Geology of the Mineral resources and Geophysical Exploration) with a thesis *Magnetic susceptibility in pit "Slovačka jama" and pit system "Kita Gaćešina-Draženova puhaljka" on Velebit*.

During studies, she actively engaged in the activities of the Student Council, the Student Section of the Croatian Geological Society, and the Ozren Lukić Speleological Club. Her achievements were recognized with the City of Zagreb scholarship in academic years 2015/2016 and 2017/2018, the recognition of Croatian Geological Society for contributions to the Society's work in 2018, the Rector's Award (2018), and the Special Dean's Award (2017). After completing studies, she obtained a certificate as a Project Preparation and Implementation Manager for EU-funded projects at Algebra.

In December 2018, she was awarded a doctoral tuition scholarship from Faculty of Mininig, Geology and Petroleum Engineering. In May 2019, she was employed as a expert associate at the Croatian Geological Survey, and in January 2024, she was promoted to the position of senior expert associate.

She has contributed as a collaborator to three projects funded by the Croatian Science Foundation, two EU-funded projects, and several others. As an author, she has published seven scientific papers.

Ana Kamenski is the recipient of the annual award from the Geomathematics Section of Croatian Geological Society for a scientific paper that had the greatest impact on the popularization of geomathematics through its quality and accessibility. Additionally, she currently holds the position of Technical Editor for the scientific journal *Geologia Croatica* and for *HGD Vijesti*.

List of published papers:

Kamenski, A., Kolenković Močilac, I. & Cvetković, M. (2025) Shale Volume, Seismic Attributes, and Proper Data Preparation: Critical Components for Modeling Subsurface Lithology Distribution. *Geologia Croatica*, 78(1).

Barudžija, U., Kamenski, A., Paar, D. & Malvić, T. (2024) Applicability of Magnetic Susceptibility Measurements on Cave Sediments in Karst Areas: Insight from Dinaric Karst (Velebit Mt., Croatia). *Applied sciences* (Basel), 14(16), 6973, 14. doi: https://doi.org/10.3390/app14166973

Kamenski, A., Cvetković, M., Kapuralić, J., Kolenković Močilac, I. & Brcković, A. (2024) From Traditional Extrapolation to Neural Networks: Time-Depth Relationship Innovations in the Subsurface Characterization of Drava Basin, Pannonian Super Basin. *Advances in geo-energy research*, 14(1), 25-33. doi: 10.46690/ager.2024.10.05

Brčić, V., Baranyi, V., Glumac, B., Špelić, M., Fuček, L., Kukoč, D., Petrinjak, K., Mišur, I., Budić, M., Palenik, D., Belić, N., Kamenski, A., Martinčević Lazar, J.; Šuica, S., Grizelj, A., Ciglenečki, I. & Dautović, J. (2024) Impact of the Jenkyns Event on shallow-marine carbonates and coeval emerged paleoenvironments: the Plitvice Lakes region, Croatia. *Palaeogeography, palaeoclimatology, palaeoecology*, 655, 112519-17. doi: 10.1016/j.palaeo.2024.112519

Kamenski, A. & Korbar, T. (2023) Platform-to-Basin Evolution of a Tectonically Indistinct Part of a Multiple Foreland—Analysis of a 3D Seismic Block in the Northern Adriatic Sea (Croatian Offshore). *Geosciences,* 13(11), 323, 22. doi: 10.3390/geosciences13110323

Cvetković, M., Kapuralić, J., Pejić, M., Kolenković Močilac, I., Rukavina, D., Smirčić, D., Kamenski, A., Matoš, B. & Špelić, M. (2021) Soil gas measurements of radon, CO2 and hydrocarbon concentrations as indicators of subsurface hydrocarbon accumulation and hydrocarbon seepage. *Sustainability*, 13(7), 3840, 16. doi: 10.3390/su13073840

Kamenski, A., Cvetković, M., Kolenković Močilac, I. & Saftić, B. (2020) Lithology prediction in the subsurface by artificial neural networks on well and 3D seismic data in clastic

sediments: a stochastic approach to a deterministic method. *GEM - International journal on geomathematics*, 11(8), 1-24. doi: 10.1007/s13137-020-0145-3