پژوهش نفت شماره ۱۴۲، مرداد و شهریور ۱۴۰۴، صفحه ۵۷-۴۴

شبیه‌سازی داده‌محور یک میدان نفتی فرضی و مقایسه نتایج با شبیه‌سازی عددی

مهدی رشنو و سعید جمشیدی*

دانشکده مهندسی شیمی و نفت، دانشگاه صنعتی شریف، تهران، ایران

تاریخ دریافت: ۱۴۰۳/۰۶/۱۱ تاریخ پذیرش: ۱۴۰۳/۱۲/۲۱

چکیده

شبیه‌سازی داده‌محور مخزن روشی نوین در امر مدل‌سازی مخازن است که مکمل و یا حتی جایگزین روش‌های عددی مدل‌سازی مخازن است. این روش که به مدل‌سازی بالا به پایین هم شناخته می‌شود با استفاده از هوش مصنوعی و به شکل خاص از یادگیری عمیق استفاده می‌کند و با توجه به ماهیت این ابزارها نیازمند به داده‌هایی است که در صنعت نفت و گاز از اندازه‌گیری‌های میدانی چه برای چاه و چه برای مخزن به‌دست می‌آیند. روش‌های عددی مرسوم با استفاده از مدل‌سازی‌های عددی و فهم کنونی که از فیزیک حاکم بر جریان سیال در محیط متخلخل وجود دارد، فرآیند شبیه‌سازی را انجام می‌دهند. در این تحقیق تلاش شده است که برای یک مورد میدان فرضی در نرم‌افزار پترل مدل‌سازی انجام شود و سپس از داده‌های خروجی از این نرم‌افزار به‌عنوان ورودی برای شبیه‌سازی مدل بالا به پایین مدنظر با استفاده از زبان برنامه‌نویسی پایتون استفاده گردیده و پس از ساخت سه مدل داده‌محور به‌ترتیب برای فشار متوسط مخزن، اشباع متوسط آب مخزن و تولید گاز هر چاه در نهایت عملکرد مدل ساخته‌شده با استفاده از یادگیری ماشین سنجیده شود. خروجی‌های مدل داده‌محور می‌توانند تمام آن چیزهایی باشند که در مدل‌های عددی انتظار می‌رود اما در این تحقیق به خروجی‌های مذکور بسنده شده است. به‌عبارت دیگر تلاش خواهد شد که مدل داده‌محور، فیزیک حاکم بر جریان سیال در محیط متخلخل را از طریق داده‌های اندازه‌گیری شده متوجه شود و در طول این فرآیند با استفاده از اطلاعات دو سال ابتدایی تولید از میدان، تطبیق تاریخچه انجام شده است و برای مدل فشار متوسط مخزن، درصد اشباع آب متوسط میدان و تولید گاز، ضریب تعیین به‌ترتیب ۰/۹۸، ۰/۹۷ و ۰/۹۸ محاسبه شده است. همچنین برای سال سوم پیش‌بینی انجام خواهد شد و نتیجه به‌دست آمده از مدل‌های داده‌محور با نتایج به‌دست آمده توسط شبیه‌ساز عددی مقایسه می‌گردد.

کلمات کلیدی: مدل‌سازی بالا به پایین، داده‌محور، روش‌های عددی، یادگیری عمیق، مخزن

مقدمه

در صنعت نفت و گاز به‌دلیل پیچیدگی‌های موجود و هزینه‌های بسیار بالای توسعه مخازن داشتن یک مدل که دقت مناسب و ردپای محاسباتی کوچکی داشته باشد اهمیت بالایی دارد. هر چقدر هم که یک مدل دقت بالایی داشته باشد چون مدل مبنای فعالیت‌های مهندسین مخزن برای بررسی حساسیت‌سنجی، عدم قطعیت‌ها و بهینه‌سازی

*مسؤول مکاتبات
آدرس الکترونیکی jamshidi@sharif.edu
شناسه دیجیتال: (DOI: 10.22078/pr.2025.5524.3456)

موارد مدنظر در امور مدیریت مخزن قرار می‌گیرد باید ردپای محاسباتی کوچکی داشته باشد. رویکردهای متعددی برای شبیه‌سازی مخزن وجود دارد: ۱. رویکرد کلاسیک[1] ۲. شبیه‌ساز مخزن عددی ۳. رویکردهای مبتنی بر یادگیری ماشین [۱]. رویکرد اول ایجاد یک مدل ریاضی ساده و حل برای یافتن برخی ضرایب است که اساسا یک مسئله برازش منحنی است روش‌های مطرح در این زمینه تحلیل منحنی نزول[2] [۲] و مدل‌سازی مقاومت خازنی[3] [۳] است که اطلاعات محدودیت‌های عملیاتی، طراحی چاه‌ها و ... را در نظر نمی‌گیرند. روش دوم ایجاد یک مدل ریاضی با پیچیدگی‌های بیشتر و گسسته‌کردن زمان و مکان و حل معادلات به‌صورت عددی است. رویکرد سوم ساخت یک مدل با استفاده از اندازه‌گیری‌های واقعی از میدانی است [۱]. در مدل‌سازی عددی روابط مورد استفاده شامل قانون بقای جرم، قانون دارسی، ترمودینامیک و بقای انرژی و... است. اعتقاد بر این است که این روابط عملکردی درست ، قطعی و غیرقابل تغییر هستند. بنابراین اگر تولید حاصل از مدل‌سازی عددی با اندازه‌گیری های ما از میدان مطابقت نداشته باشد نتیجه می‌گیریم که ویژگی‌های مخزن (مدل ثابت[4]) ممکن است به‌طور ایده‌آل اندازه‌گیری و تفسیر نشوند و باید اصلاح شوند [۴]. مدل‌های عددی سریع نیستند و نیازمند سرمایه‌گذاری زیاد، زمان و نیروی انسانی هستند. مدل‌های پروکسی هوشمند[5] جایگزینی برای شبیه‌سازی‌های عددی هستند که به مهندسان مخزن در توسعه مخزن و مدیریت تولید کمک می‌کنند [۵]. در روش‌های عددی برای رسیدن به نتایج مطلوب مدل ثابت دچار تغییر می‌شود و برای ایجاد امکان حل مسئله فرض‌های ساده‌شده‌ای در نظر گرفته می‌شود که مسئله را از حالت واقعی خود دور می‌کند. باید انتظار داشت همان‌طور که در طول زمان درک موجود نسبت به این مسئله توسعه یافته است در زمان‌های آینده نیز به حد کمال خود برسد و شاید درک کنونی دارای نواقصی باشد. در صنعت نفت و گاز یکی از محبوب‌ترین شیوه‌ها، یادگیری نظارت‌شده است [۶]. درک شبکه‌های عصبی نیاز به درک علوم اعصاب دارد [۷]. و شبکه عصبی مصنوعی به‌عنوان یکی از فن‌آوری‌های هوش مصنوعی طبقه‌بندی می‌شود [۸]. در فرآیند آموزش از فرآیند تکرای پسا انتشار[6] تا زمان رسیدن به همگرایی وزن بین نورون‌ها[7] اصلاح می‌شود [۹]. و برای جلوگیری از بیش‌برازش[8]، داده‌های واسنجی[9] برای تایید قابلیت‌های تعمیم مدل استفاده می‌شوند [۱۰]. در این صنعت از هوش مصنوعی برای حل مسائل مربوط به تجزیه و تحلیل فشار گذرا، حفاری، شناسایی ویژگی‌ها مخزن و تفسیر نمودار چاه استفاده شده است [۱۱]. شهاب محقق یکی از پیشگامان به کارگیری هوش مصنوعی در مهندسی نفت، مدل‌سازی داده محور را به‌عنوان فرآیند به کارگیری هوش مصنوعی و روش‌های داده‌کاوی مانند یادگیری ماشین و تشخیص الگو به‌منظور کشف الگوهای پنهان در مجموعه داده‌های بزرگ که نشان‌دهنده جریان سیال در محیط های متخلخل هستند تعریف کرد. سپس الگوهای کشف‌شده در یک مدل مخزنی جامع و منسجم با قابلیت پیش‌بینی مورد تأیید گنجانده شده‌اند که می‌توانند برای مدیریت میادین و تولید آن‌ها مورد استفاده قرار گیرد. فن‌آوری داده‌محور کاملا مبتنی بر اندازه‌گیری‌های میدانی است. سه فاز در مدل‌سازی داده‌محور مخزن گنجانده شده است، فاز اول ماهیت تجزیه و تحلیل اکتشافی است که مربوط به داده‌کاوی است. فاز دوم شامل توسعه یک مدل مخزن پیش‌بینی‌کننده برای کل میدان با استفاده از هوش مصنوعی است.

1. Classic Approaches
2. Decline Curve Analysis
3. Capacitance Resistance Modeling
4. Static Model
5. Smart Proxy Model
6. Backpropagation
7. Neuron
8. Overfitting
9. Calibration

درحالی‌که فاز سوم مدل‌سازی مخزن داده‌محور، تحلیل پس داده است که ترکیبی از داده‌کاوی و هوش مصنوعی است [۱۲]. در مدل‌سازی مخزن داده‌محور نقش فیزیک از معمار معادلات حاکم، به چراغ راهنما و طرح اولیه‌ای که چارچوبی را برای فرآیند توسعه مدل فراهم می‌کند تغییر می‌کند [۱۳]. ویژگی‌های متمایز این فن‌آوری از شبیه‌سازی مخزن سنتی این است که: ۱) فرض بر این نیست که تمام اطلاعات لازم برای ساخت یک مدل زمین‌شناسی که نمایندگی کاملی از میدان است در دسترس می‌باشد . ۲) فرض نمی‌شود که به‌طور کامل فیزیک حاکم درک می‌شود و می‌توان تمام پیچیدگی‌ها و تفاوت‌های ظریف جریان سیال را در محیط متخلخل برای میدان در حال مدل‌سازی فرمول‌بندی کرد [۱۴]. این مدل‌سازی یک فن‌آوری پیشگام است که تعداد زیادی از رشته‌ها مانند مدل‌سازی مخزن، مهندسی مخزن، تجزیه و تحلیل داده‌محور پیشرفته و تجزیه و تحلیل آماری را با استفاده از یادگیری ماشین و هوش مصنوعی ادغام می‌کند[۱۳]. مدل‌سازی مخزن مبتنی بر هوش مصنوعی یک فن‌آوری نوین است و پتانسیل این فن‌آوری بالا است زیرا مدل‌های مخزن را می‌توان در کسری از زمان و بودجه مدل‌های سنتی تولید کند و در عین حال بیشتر قابلیت‌های آن را ارائه می‌کند. چندین مقاله در سال‌های اخیر منتشر شده است که نشان‌دهنده کاربردی‌بودن و عملیاتی بودن مدل‌سازی مخزن داده محور را ارائه می‌دهند [۱۵]. حقیقت و همکاران مدل‌سازی داده‌محور را برای میدان نیوبرارا [۱۶]، میثمی و همکاران برای میدان هایلایت در حوضه رودخانه پودر وایومینگ [۱۷] پیاده‌سازی کردند و علاوه‌بر تطبیق تاریخچه با استفاده از مدل‌های اعتبارسنجی‌شده، برای چند سال پیش‌بینی را درمقایسه با روش‌های عددی انجام داده‌اند که نتایج خوبی حاصل شده است. یارا آلاتراک مطالعه‌ای را برروی یک مخزن کربناته در ابوظبی انجام داد و توانست با استفاده از داده‌های تاریخچه برای بهینه‌سازی تولید و تزریق برای به حداکثر رساندن تولید کل میدان، مدل را آموزش دهد [۱۸]. زرگری و محقق مدل‌سازی داده‌محور مخزن را بر روی بخشی از سازند شیل باکن در حوضه ویلیستون داکوتای شمالی پیاده‌سازی کرده‌اند. از مدل آن‌ها می‌توان برای پیشنهاد راهبردهای توسعه، شناسایی ذخایر باقی‌مانده و در نتیجه مکان چاه‌های جدید استفاده کرد. علاوه‌بر این یک مدل پیش‌بینی‌کننده تولید شده و تاریخچه موجود را تطبیق داده و تحلیل اقتصادی برای برخی از چاه‌های جدید پیشنهادی انجام شده است [۱۹]. محقق و همکاران مدل‌های داده‌محور را برای سه مورد منابع شیلی (نفتی و گازی) که تطابق خوبی با تاریخچه دارد ایجاد کرده و پس از اعتبارسنجی در نهایت برای پیش‌بینی و کمک به برنامه‌ریزی راهبردهای توسعه میدان از آنها استفاده کردند [۲۰]. حسن الحیفی نتایج تطبیق تاریخچه و پیش‌بینی مدل با داده‌های اصلی تولید شده از شبیه‌سازی عددی را برای یک مدل فرضی مقایسه نمود. از طرفی او بر برخی از چالش‌های استفاده از هوش مصنوعی و یادگیری ماشین تأکید کرد زیرا این فن‌آوری‌ها به حجم زیادی از داده‌های با کیفیت نیاز دارند تا بتوانند یک مدل پیش‌بینی دقیق به‌دست آورند [۱۴] همچنین مارتینز برای یک میدان فرضی عملکرد مدل‌های داده‌محور آموزش دیده را تحت راهبردهای مختلف مورد بررسی قرار داد [۲۱]. در این مقاله، برای اولین بار در ایران شبیه‌سازی داده‌محور مخزن به صورت عملی انجام می‌گیرد. همچنین در این تحقیق بنابر شرایط و پیچیدگی‌های مورد تحت مطالعه، علی‌رغم اینکه در تحقیقات قبلی از روش منطق فازی برای تشخیص اثرگذارترین پارامترها برروی خروجی مدنظر استفاده شده است، از دو روش مرسوم در یادگیری ماشین (نقشه حرارتی و روش اسپیرمن[1]) برای یک میدان نفتی فرضی استفاده شده است.

1. Spearman Method

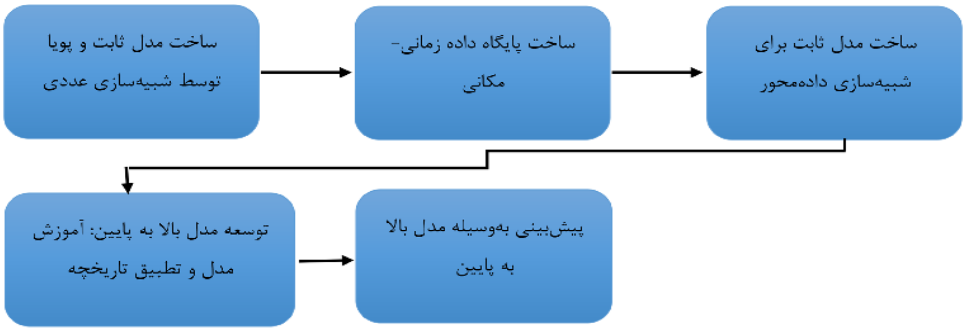
روش‌شناسی

شکل ۱ روند انجام این مدل‌سازی را طی پنج مرحله نشان داده است که با استفاده از زبان برنامه‌نویسی پایتون انجام شده است.

ساخت مدل ثابت و پویا توسط شبیه‌سازی عددی

شبیه‌سازی با استفاده از نرم‌افزار پترل برروی یک میدان با ابعاد مربعی به طول ضلع ۴۸۷۶/۸ m واقع در عمق ۹۱۴/۴ m از سطح زمین و ضخامت کلی ۱۸۲/۸۸ m با اندازه شبکه‌بندی‌های مربعی با طول ضلع ۶۰/۹۶ m انجام شده است. این میدان دو لایه با خط جداکننده‌ای در عمق ۱۰۳۶/۳۲ m دارد. همان‌طور که از جدول ۱ برمی‌آید ویژگی‌های تخلخل و تراوایی تعریف شده در این لایه‌ها به‌شرح زیر است:

در رابطه با خواص سیال، نفت سبک و گاز در نظر گرفته شده است. جدول ۲ بیانگر مهم‌ترین خواص تعریف‌شده است: جنس سنگ مخزن از نوع سنگ آهک تثبیت‌شده است. خواص سنگ و سیال همچون تراوایی نسبی (با استفاده از رابطه کوری) و فشار مویینگی هم مقدار پیش‌فرض انتخاب شده است. برنامه توسعه میدان به شرح زیر است: ۱) در هر سه ماه می‌توان پنج چاه را حفر نمود و چاه‌ها به محض حفرشدن شروع به تولید/تزریق می‌کنند.



شکل ۱ روند انجام شبیه‌سازی داده‌محور

جدول ۱ ویژگی‌های لایه‌های میدان

ویژگی	توزیع	میانگین	انحراف معیار استاندارد	سبد اعداد تصادفی	کمینه	بیشینه
تخلخل لایه اول	نرمال	۰/۲۵	۰/۰۶۲۵	۲۸۱۰۶	۰	۰/۵
تخلخل لایه دوم	نرمال	۰/۱۵	۰/۰۶۲۵	۲۲۴۰۸	۰	۰/۵
تراوایی لایه اول (m^2)	لاگ‌نرمال	$۷/۸۲۸۹\times۱۰^{-۱۴}$	$۲/۴۶۱۲\times۱۰^{-۱۵}$	۲۳۹۴	$۵/۹۲۱۵\times۱۰^{-۱۴}$	$۷/۸۹۵۴\times۱۰^{-۱۴}$
تراوایی لایه دوم (m^2)	لاگ‌نرمال	$۶/۸۴۲۰\times۱۰^{-۱۴}$	$۲/۴۶۱۲\times۱۰^{-۱۵}$	۱۲۱۰۴	$۵/۹۲۱۵\times۱۰^{-۱۴}$	$۷/۸۹۵۴\times۱۰^{-۱۴}$

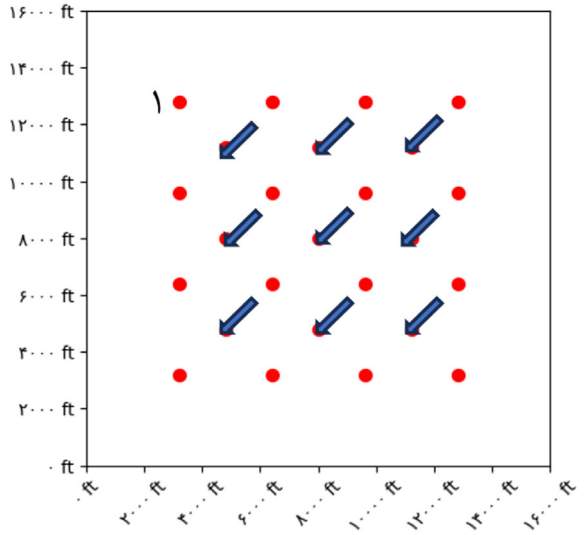
جدول ۲ خواص سیال میدان

وزن مخصوص گاز	۰/۶۵
ای پی آی نفت	۴۵
فشار نقطه حباب (Pa)	۳×۱۰^{۷}
شوری آب (ppm)	۳۵۰۰۰
فشار اولیه در عمق مرجع (Pa)	۳۴۰۴۶۳۱۱/۵

۲) فاصله یک چاه با چاه کناری را ۱۵۲/۴ m فرض شده که با توجه به در نظر گرفتن ابعاد میدان حداکثر می‌توان ۲۵ چاه در آن حفر نمود. ۳) حداقل فشار ته‌چاهی را ۱۳۷۸۹۵۱۴/۵۸ Pa و حداقل تولید نفت از یک چاه ۷/۹۴۹۳۶۵ m^3/day فرض شده است. ۴) حداکثر ظرفیت خطوط لوله انتقال سیال (نفت و آب) را ۷۹۴/۹۳۶۵ m^3/day و همچنین حداکثر ظرفیت تزریق آب نیز همین مقدار فرض شده است. ۵) فشار تزریق ته‌چاهی برای جلوگیری از شکاف‌خوردن ۴۱۳۶۸۵۴۳/۷۴ Pa و مقدار کسر جایگزین‌شده ۰/۵ در نظر گرفته شده است. میدان دارای ۲۵ چاه است که ۱۶ چاه تولیدی و ۹ چاه تزریقی با الگوی تزریق پنج نقطه حفر شده است و همگی بدون لوله جداری هستند و تمامی چاه‌ها از هر دو لایه تولید می‌کنند. محل حفر چاه‌ها و نوع آن‌ها در شکل ۲ نمایش داده شده است. حفر چاه‌ها از ۰۱/۱۰/۲۰۲۳ شروع شده و در تاریخ ۰۱/۰۱/۲۰۲۴ چهار چاه تولیدی شروع به تولید کرده و یک چاه تزریقی شروع به تزریق می‌کند و در نهایت در تاریخ ۰۱/۰۱/۲۰۲۵ تمامی ۲۵ چاه شروع به فعالیت تا تاریخ ۰۱/۰۱/۲۰۲۶ می‌کنند. نحوه گزارش داده‌ها در این دو سال به‌صورت روزانه است.

ساخت پایگاه داده زمانی – مکانی

تهیه پایگاه داده مهم‌ترین کار در مدل‌سازی داده‌محور است. در این پایگاه داده دو دسته داده ثابت و پویا وجود دارد که دسته ثابت به دو دسته واقعا ثابت و ثابت اصلاح‌شده تقسیم‌بندی می‌شوند. داده‌های ثابت اصلاح‌شده مربوط به داده‌های یک ناحیه زهکشی هستند. با حفر چاه‌های جدید در میدان ناحیه زهکشی برای یک چاه کمتر خواهد شد و به‌عنوان مثال تخلخل متوسط در ناحیه زه‌کشی آن چاه که میانگینی از تخلخل شبکه‌بندی‌های موجود در آن است متفاوت خواهد شد. لیست این داده‌ها در جدول ۳ آورده شده است. در بین داده‌های ورودی، علی‌رغم اهمیت تراوایی در مهندسی نفت، اطلاعات آن استفاده نشده است. علت این امر آن است که در رابطه با تراوایی متوسط در ناحیه زه‌کشی یک چاه، به‌علت خواص مخزنی انتخاب شده بعد از محاسبه تراوایی متوسط برای هر دولایه و بعد از گرفتن میانگین وزنی برای تمام ناحیه‌های زه‌کشی، مقداری یکسان به‌دست آمد که به همین خاطر از ورودی‌ها حذف شده است برای تحقیقات آینده توصیه می‌شود که از داده‌های واقعی استفاده شود که تفاوت تراوایی در مناطق مختلف به خوبی نمایان شود. چون چاه‌های مجاور بر تولید یک چاه مرکزی اثرگذار هستند در این تحقیق با توجه به الگوی حفر چاه‌ها، ۴ چاه اطراف یک چاه به‌عنوان چاه‌های مجاور چاه مرکزی در نظر گرفته می‌شوند.



شکل ۲ محل چاه‌ها و نوع آن‌ها

جدول ۳ داده‌های مورد استفاده در توسعه مدل بالا به پایین

نوع داده	داده
واقعا ثابت	طول جغرافیایی، عرض جغرافیایی و تخلخل در محل چاه
ثابت اصلاح‌شده	تخلخل متوسط در ناحیه زه‌کشی چاه
پویا	فشار ته‌چاهی، تولید نفت، تزریق آب، نسبت گاز به نفت، فشار متوسط میدان، درصد اشباع متوسط آب میدان و تولید گاز

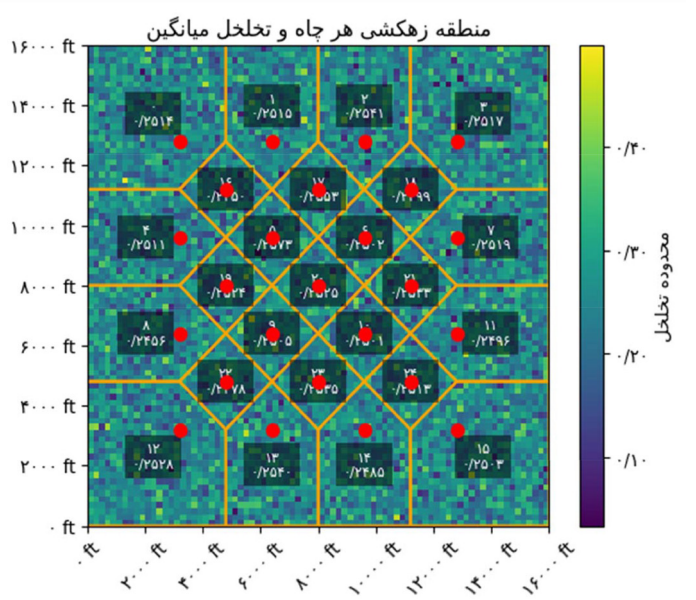
و علاوه‌بر اطلاعات یک چاه مرکزی، در ستون‌های مجاور اطلاعات چاه‌های مجاور آن گنجانده می‌شود. این باعث می‌شود که یک چاه هم به‌عنوان چاه مرکزی و هم به‌عنوان چاه مجاور چندین بار مورد مطالعه توسط مدل‌های داده‌محور قرار گیرد.

ساخت مدل ثابت برای شبیه‌سازی داده‌محور

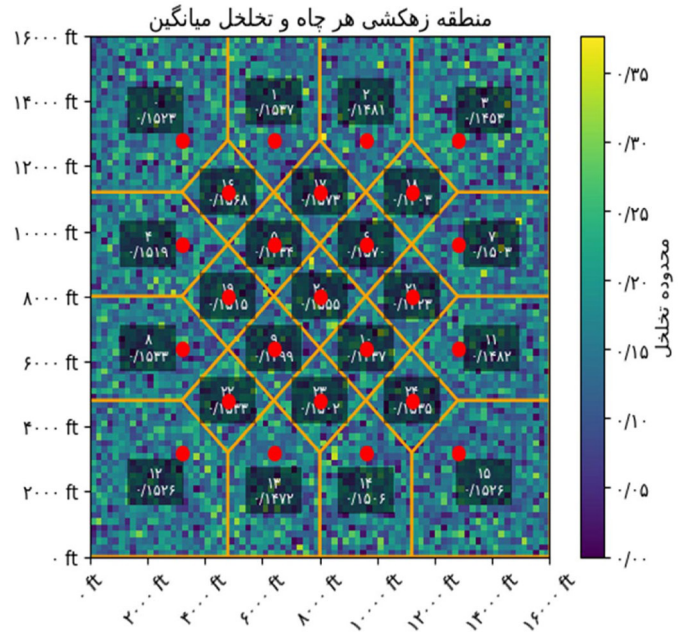
در مدل داده‌محور، میدان به شبکه‌بندی‌هایی با ضلع ۶۰/۹۶ m تقسیم‌بندی شده و بر طبق ویژگی‌های جدول ۳-۱ خواص تخلخل تعریف می‌شود و میانگین مقدار تخلخل شبکه‌بندی‌های داخل یک ناحیه زه‌کشی محاسبه می‌گردد. به‌علت تغییر مساحت ناحیه زه‌کشی یک چاه بعد از حفر چاه‌های جدید، این کار برای هر پنج مرحله حفر چاه و هم برای هر دو لایه انجام می‌گردد. برای تعیین ناحیه زه‌کشی هر چاه از نظریه وورونوی استفاده شده است. در شکل‌های ۳ و ۴ تخلخل میانگین در هر ناحیه زه‌کشی برای لایه بالا و پایین پس از حفر تمام چاه‌ها نشان داده شده است. چون میدان دو لایه دارد برای پارامترهای تخلخل و تخلخل میانگین از داده‌های لایه اول و دوم براساس ضخامت لایه‌ها میانگین وزنی گرفته شده و در پایگاه داده قرار داده شده است.

توسعه مدل بالا به پایین: آموزش مدل و تطبیق تاریخچه

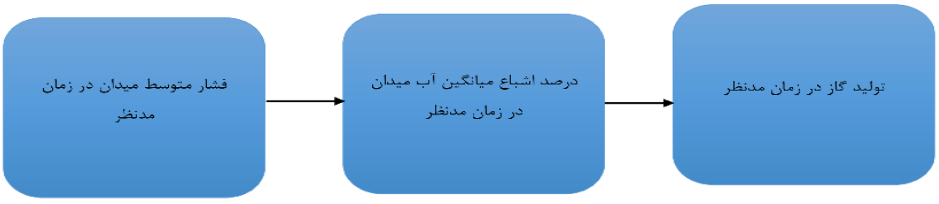
برای تعیین تعداد مدل‌ها از آنجا که در ابتدا میدان فشار بالایی دارد نرخ تولید نفت و تزریق آب ثابت است پس نیازی به آن‌ها نیست و مدل‌هایی که هدف مدل‌سازی هستند طبق شکل ۵ سه مدل خواهند بود: در شکل ۵ برای پیش‌بینی خروجی فشار متوسط میدان در گام زمانی مشخص، علاوه‌بر تمام داده‌های دیگر از اطلاعات فشار متوسط میدان در گام زمانی قبلی هم استفاده می‌شود. بعد از ساخت مدل اول، فشار متوسط در گام زمانی مدنظر، درصد اشباع میانگین در گام زمانی قبلی و تمام داده‌های دیگر برای پیش‌بینی درصد اشباع میانگین در گام زمانی مدنظر استفاده می‌شود. سپس فشار متوسط در گام زمانی مدنظر، درصد اشباع میانگین در گام زمانی مدنظر، تولید گاز در گام زمانی قبلی و تمام داده‌های دیگر برای پیش‌بینی تولید گاز در گام زمانی مدنظر استفاده خواهد شد. جدول ۴ تعداد داده‌های استفاده‌شده در ساخت یک مدل داده‌محور را تحت قسمت‌بندی تصادفی نشان می‌دهد. همچنین برای آموزش یکی از مدل‌های داده‌محور مدل بالا به پایین از شبکه عصبی مصنوعی که یک لایه پنهان و یک گره خروجی دارد استفاده شده و تمامی گره‌ها دو لایه‌ی مجاور، به یکدیگر متصل هستند. به‌عبارت دیگر از شبکه‌ی عصبی با مدل پرسپترون چند لایه (با یک لایه پنهان) استفاده شده است و الگوریتم مورد استفاده برای بهبود وزن‌های شبکه عصبی پس‌انتشار می‌باشد. طبق شکل ۶ برای کاهش تعداد ورودی‌ها، ارتباط بین داده‌های ورودی با استفاده از روش‌های مرسوم در یادگیری ماشین بررسی شده است. همان‌طور که از شکل ۶ پیدا می‌باشد ستون داده‌های ورودی مربوط به نوع تولیدی/تزریقی‌بودن و نسبت گاز به نفت تولیدی را می‌توان حذف نمود (نام چاه‌ها و تاریخ هم به‌علت حضور طول و عرض جغرافیایی و ماهیت داده‌های پویا حذف می‌شوند).



شکل ۳ ناحیه زهکشی چاه‌ها به‌همراه تخلخل میانگین هر ناحیه برای لایه اول مخزن پس از حفر تمامی چاه‌ها



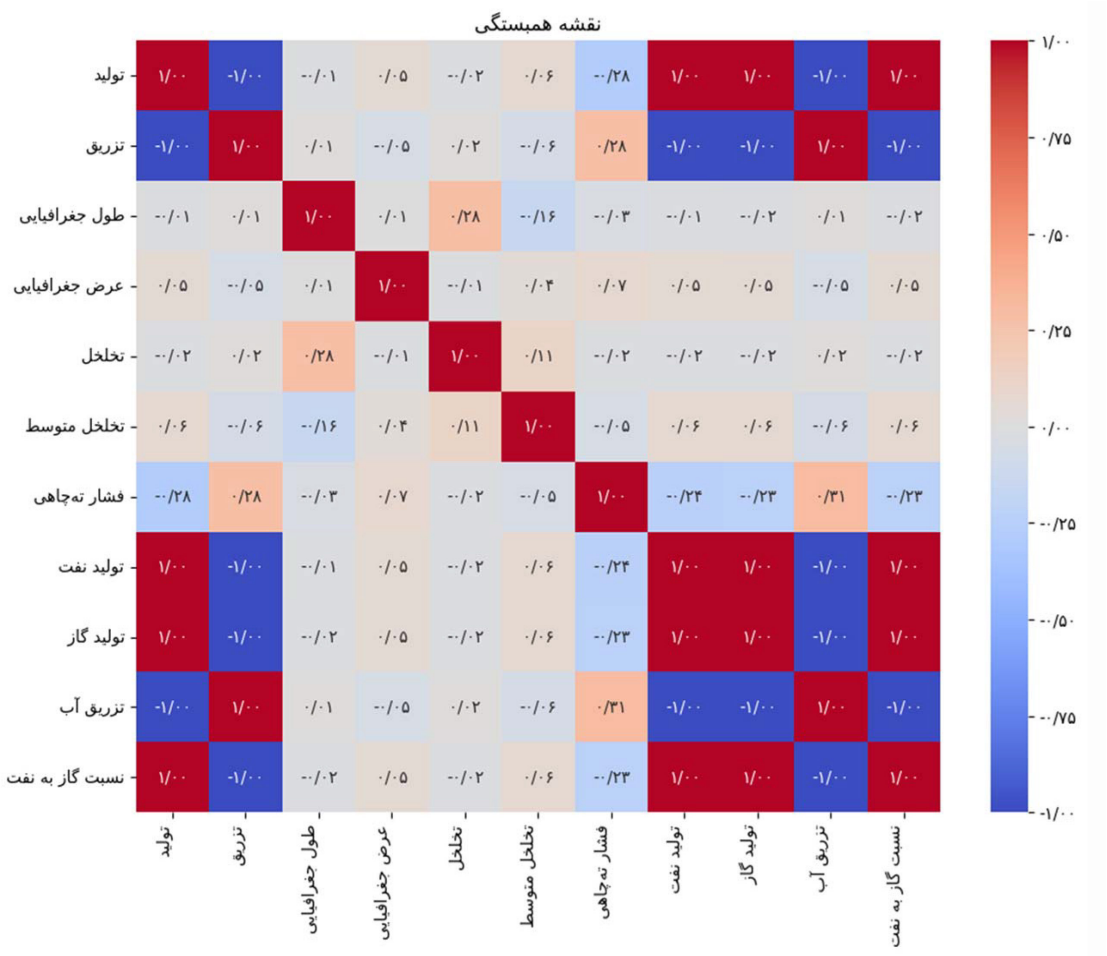
شکل ۴ ناحیه زهکشی چاه‌ها در مراحل مختلف به‌همراه تخلخل میانگین هر ناحیه برای لایه دوم مخزن



شکل ۵ عنوان و ترتیب مدل‌های داده‌محور مورد نیاز برای توسعه مدل بالا به پایین

جدول ۴ تعداد داده‌های استفاده‌شده در ساخت یک مدل داده‌محور

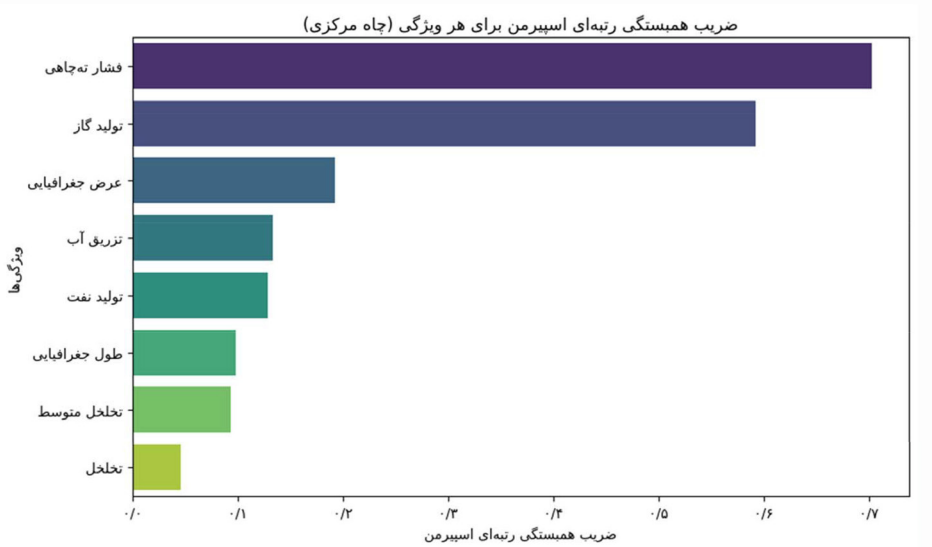
کل داده‌ها	داده‌های آموزشی	داده‌های اعتبارسنجی	داده‌های آزمون
۱۳۷۳۵	۸۷۹۰	۲۱۹۸	۲۷۴۷



شکل ۶ ارتباط بین پارامترهای ورودی با استفاده از روش‌های مرسوم در یادگیری ماشین

علاوه‌بر این طبق شکل ۷ از روش اسپیرمن به‌منظور مشخص‌سازی ورودی‌هایی با بیشترین اثرگذاری برروی خروجی مدل برای چاه مرکزی استفاده شده و در نهایت سه پارامتر با بیشترین اثرگذاری انتخاب گردیده است (این کار برای چهار چاه مجاور آن نیز باید انجام شود). با وجود کم‌اثربودن طول جغرافیایی حذف نشده زیرا این پارامتر اگر حذف شود مسئله فیزیک معناداری نخواهد داشت. پارامترهای تولید نفت و تزریق آب هم چون اعداد ثابت بودند کمترین اثرگذاری را داشتند. برای نرمال‌سازی مقادیر داده‌ها روش کمینه و بیشینه[1]، برای تابع فعال‌سازی بین لایه ورودی و لایه پنهان تابع رلو، برای تابع فعالسازی بین لایه پنهان و لایه خروجی تابع خطی[2]، برای معیار کاهش در طول آموزش مدل، میانگین مربع خطا[3] و به‌عنوان معیار بررسی عملکرد، میانگین خطای مطلق[4] استفاده شده است. در بحث بهینه‌سازی پارامترها هم با استفاده از روش هایپربند (با معیار میانگین خطای مطلق) تلاش شده است که تعداد گره‌های لایه میانی و نرخ یادگیری بهینه شود. از سوی دیگر برای بهینه‌سازی تعداد مشاهده‌ی داده‌های ورودی به‌صورت دستی اقدام شده است. در مدل فشار متوسط میدان بعد از پیدا کردن بهترین ساختار مدل و پیداکردن مناسب‌ترین پارامترها آموزش مدل انجام خواهد شد در نهایت برای ارزیابی مدل بر روی داده‌های آزمون از معیار ضریب تعیین استفاده می‌شود.

1. MinMaxScaler
2. Linear Function
3. Mean Squared error
4. Mean Absolute error



شکل ۷ نتایج بررسی اثرگذارترین ورودی‌ها برروی مدل فشار متوسط میدان با استفاده از روش اسپیرمن برای چاه مرکزی

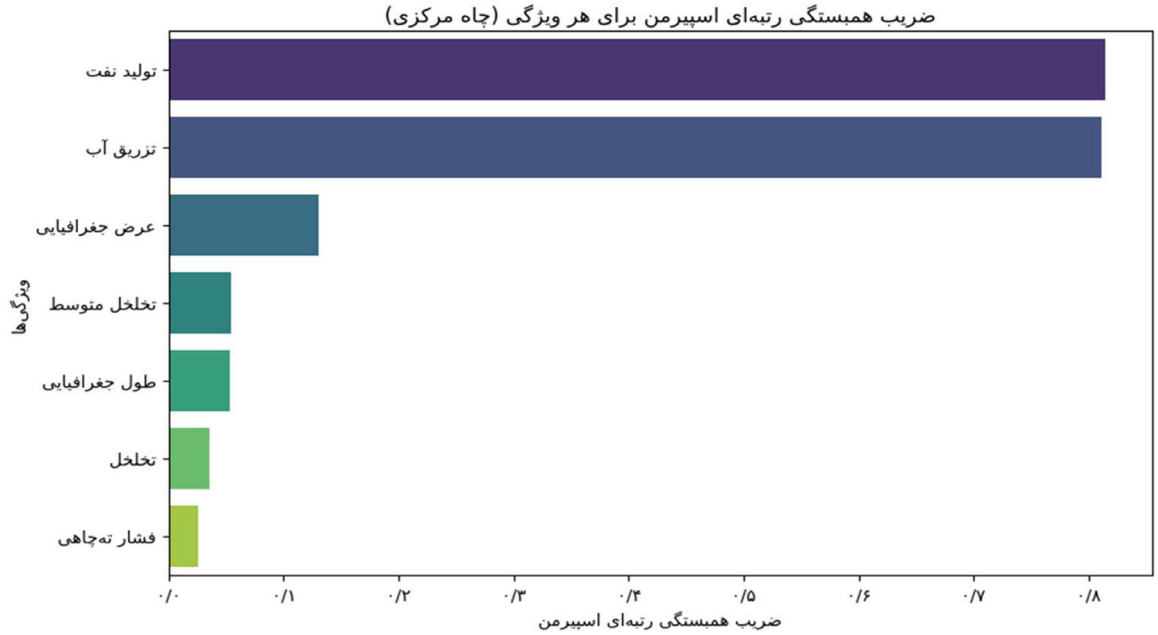
در آخر با استفاده از داده‌های یک چاه فشار متوسط میدان پیش‌بینی شده و در کنار فشار متوسط واقعی میدان رسم می‌شود. در ساخت مدل درصد اشباع میانگین نیز روند بدین شکل است و اثرگذارترین پارامترها همانند مدل فشار متوسط میدان هستند. در ساخت مدل تولید گاز نیز همانند طبق شکل ۸ اثرگذارترین پارامترها برروی مدل تولید گاز برای چاه مرکزی مشخص شده‌اند (این کار برای چهار چاه مرکزی نیز انجام شده است) و تعداد گره‌های ورودی ۲۲ گره خواهد بود زیرا ورودی‌های بیشتری (تولید گاز و نسبت گاز به نفت تولیدی) حذف شده‌اند. برای حفظ فیزیک مسئله طول و عرض جغرافیایی چاه مرکزی حذف نشده است ولی در چاه‌های مجاور به‌علت اثرگذاری کم آنها حذف شده‌اند. برای تمامی مدل‌ها و به‌منظور جلوگیری از بیش‌برازش این مدل‌ها از روش حذف موقت تعدادی از گره‌های لایه پنهان در حین آموزش استفاده شده است که مقدار بهینه به‌صورت حدس و خطا و در بازه ۱۰ تا ۵۰٪ انتخاب شده است.

پیش‌بینی به‌وسیله مدل بالا به پایین

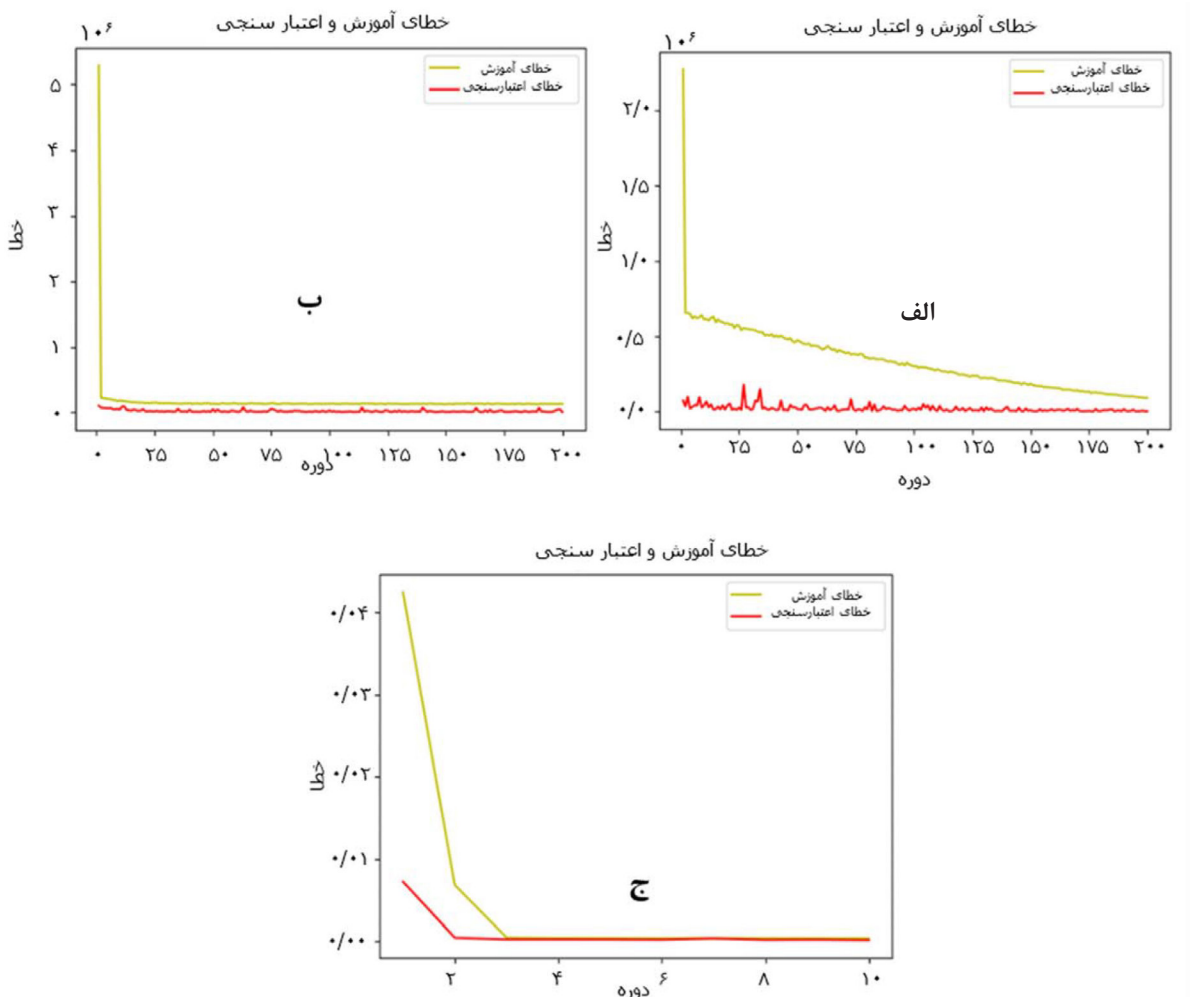
پس از تطبیق تاریخچه در مراحل قبل از مدل‌های ساخته‌شده استفاده شده تا با استفاده از داده‌هایی که هرگز مشاهده نکرده است عملکرد میدان پیش‌بینی شود.

نتایج و بحث

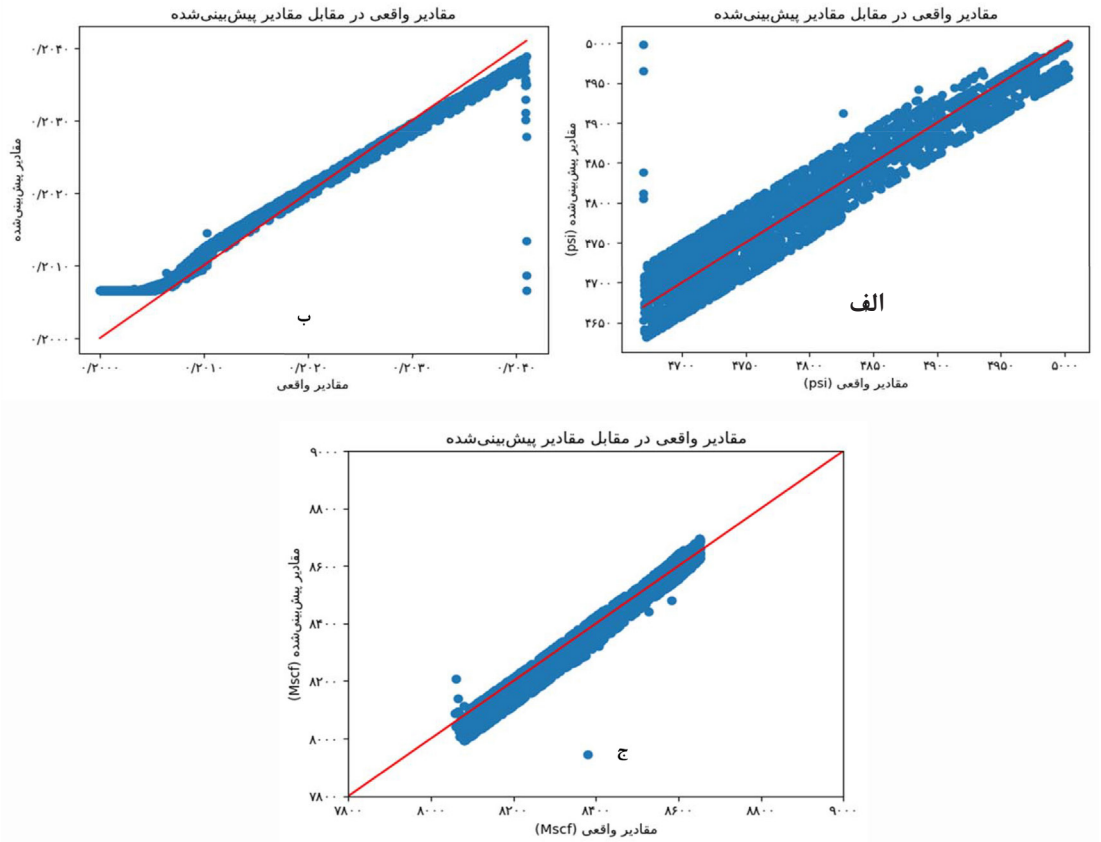
در شکل ۹ روند یادگیری مدل‌های مختلف براساس معیار میانگین مطلق خطا برای درصد اشباع آب میانگین و میانگین مربع خطا برای مدل‌های فشار متوسط مخزن و تولید گاز نمایش داده شده است. در شکل ۹ برای مدل درصد میانگین اشباع آب تنها از معیار میانگین مطلق خطا استفاده شده است زیرا مقادیر اشباع کمتر از یک هستند و با محاسبه مربع خطاها اصلاح وزن‌ها با تاخیر انجام خواهد شد. همچنین همان‌طور که در شکل ۹ مشخص است خطای مدل پس از دیدن چندین باره داده‌ها به‌خصوص در بخش داده‌های آموزشی کاهش یافته است که همین روند به‌صورت کلی در داده‌های اعتبارسنجی نیز دیده می‌شود که به‌علت جلوگیری از بیش‌برازش، از دیدن بسیار زیاد داده‌ها توسط مدل جلوگیری شده است. در شکل ۱۰ مقادیر پیش‌بینی‌شده توسط مدل‌های داده‌محور براساس مقادیر محاسبه‌شده توسط مدل عددی رسم شده است که نشان‌دهنده دقت مناسب مدل داده‌محور و تطابق مناسب بین نتایج مدل‌های داده‌محور و مدل شبیه‌ساز عددی و یادگیری فیزیک جریان در محیط متخلخل می‌باشد و نشان می‌دهد درصورتی‌که از مدل‌های داده‌محور برای پیش‌بینی استفاده شود نتایج مناسبی را به‌عنوان خروجی محاسبه می‌کند.



شکل ۸ نتایج بررسی اثرگذارترین ورودی‌ها برروی مدل تولید گاز با استفاده از روش اسپیرمن برای چاه مرکزی



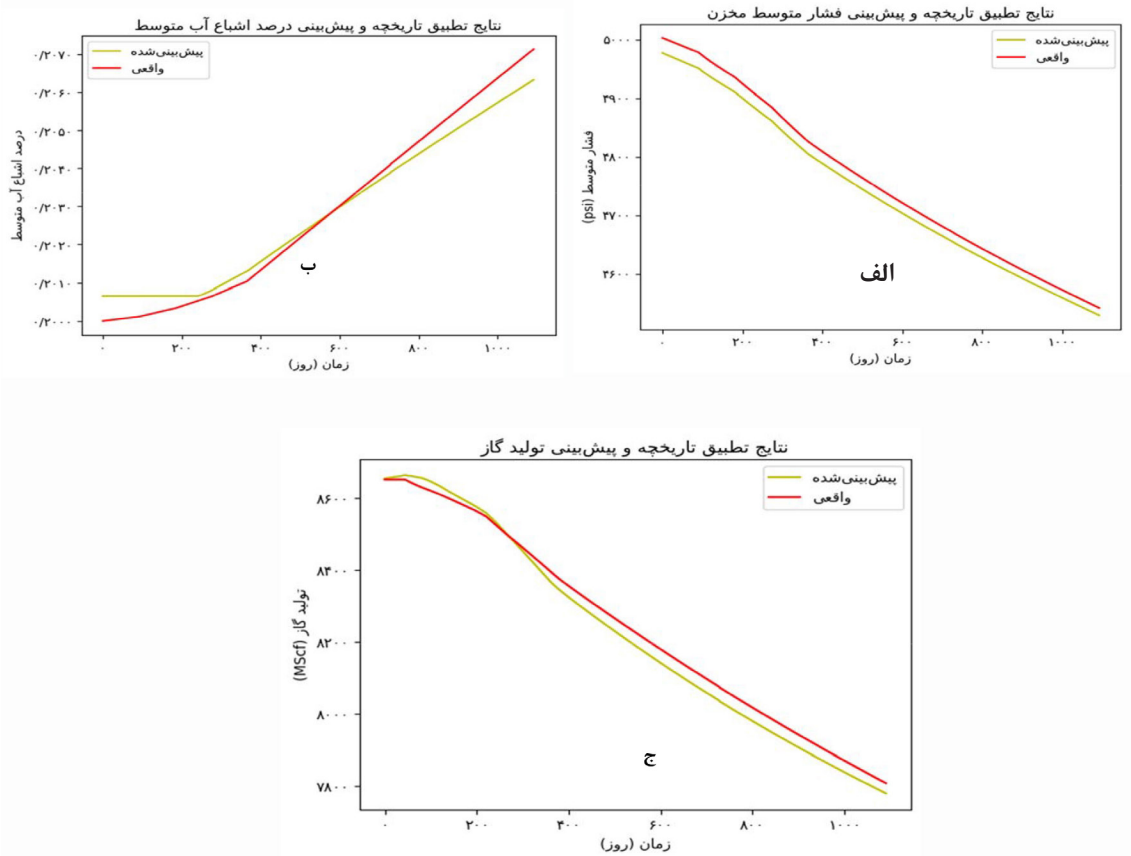
شکل ۹ روند یادگیری در مدل‌های مختلف بر اساس میانگین مربع خطا برای الف) مدل فشار متوسط میدان، و ب) مدل تولید گاز و براساس میانگین مطلق خطا برای ج) مدل درصد میانگین اشباع آب



شکل ۱۰ رسم مقادیر پیش‌بینی‌شده برحسب مقادیر واقعی در مدل‌های داده‌محور مختلف: الف) فشار متوسط میدان، ب) درصد اشباع میانگین آب و ج) تولید گاز

در شکل ۱۰ ب مدل داده‌محور درصد میانگین اشباع آب برای دامنه‌ای از مقادیر ابتدایی و کوچک یک عدد ثابت را پیش‌بینی کرده است. پس از بررسی داده‌های موجود در پایگاه داده مشخص شده که اعداد مربوطه در روزهای ابتدایی تا پنج رقم بعد از اعشار هم یکسان هستند درحالی‌که ما از اطلاعات روزانه استفاده کرده‌ایم و همچنین در سال‌های ابتدایی توسعه میدان جمع‌آوری شده است. همان‌طور که از شکل ۱۱ برمی‌آید مدل‌های داده‌محور توانسته‌اند روند تغییرات متغیرهای مدنظر را به‌خوبی درک کنند. همچنین تغییرات شیب مشاهده شده در نمودارهای مربوط به نتایج شبیه‌ساز عددی به‌علت حفر چاه‌های جدید توسط مدل‌های داده‌محور فهمیده شده است و در نمودارها قابل مشاهده است. در جدول ۵ نتایج بهینه‌سازی برای پارامترهای مدنظر به‌منظور آموزش مدل‌ها گنجانده شده است و تلاش شده است که با کوچک‌ترین شبکه عصبی ممکن و کمترین حجم محاسبات به نتایج مناسبی رسید. در جدول ۶ نیز نتایج معیار ضریب تعیین[1] و سایر معیارهای اندازه‌گیری دقت برای مدل‌های داده‌محور آموزش دیده برروی داده‌های دسته آزمون برای بررسی عملکرد آنها آورده شده است که ضریب تعیین همگی بالاتر از ۰/۹ هستند و سایر معیارها نیز دقت خوبی را برای مدل‌ها نشان می‌دهند. از این رو می‌توان فهمید که مدل‌های داده‌محور به‌خوبی توانسته‌اند فیزیک حاکم بر جریان سیال در محیط متخلخل که توسط شبیه‌ساز عددی در دسترس آن‌ها قرار گرفته است به‌خوبی درک کنند. اکنون که مدل‌ها اعتبارسنجی شدند و از عملکرد آن‌ها در فرآیند تطبیق تاریخچه اطمینان حاصل شده از آن‌ها برای پیش‌بینی براساس اطلاعات چاه شماره ۱ استفاده شده است.

1. R2 Score



شکل ۱۱ نتایج تطبیق تاریخچه و پیش‌بینی در سه سال متوالی در کنار هم توسط مدل‌های مختلف در مقایسه با روش عددی با استفاده از اطلاعات چاه شماره ۱: الف) فشار متوسط میدان، ب) درصد اشباع میانگین آب و ج) تولید گاز

جدول ۵ نتایج پارامترهای بهینه‌شده در مدل‌های داده‌محور مذکور برای توسعه مدل بالا به پایین

مدل	ورودی‌ها	نرخ یادگیری	شمار گره‌های استفاده شده در لایه پنهان	دفعات مشاهده داده‌ها	اندازه دسته	حذف موقت
فشار متوسط میدان	۲۱	۰/۱	۹۰	۲۰۰	۳۲	۰/۵
درصد اشباع آب متوسط میدان	۲۳	۰/۰۰۱	۱۰	۱۰	۳۲	۰/۲
تولید گاز	۲۲	۰/۱	۱۰۰	۲۰۰	۳۲	۰/۱

جدول ۶ نتایج معیارهای اندازه‌گیری دقت برای مدل‌های داده‌محور به‌منظور توسعه مدل بالا به پایین برای داده‌های آزمون

مدل	ضریب تعیین	میانگین مربع خطا	میانگین مطلق خطا
فشار متوسط میدان (Pa)	۰/۹۰۷	۱۲۹۴۹۲۶۳۵۱/۴	۱۴۶۸۵۱/۴۵
درصد اشباع آب متوسط میدان	۰/۹۵۹	۰/۰۰۰۲۴۵۶۷۹۵۵	۰/۰۰۰۱۷۴۸۷۵۴۳
تولید گاز (m^3)	۰/۹۹۷	۵۴۵۲/۷	۱۱۴۱/۶۸

شکل ۱۱ نتایج تطبیق تاریخچه و پیش‌بینی (دو سال تطبیق تاریخچه و یک سال پیش‌بینی) را برای فشار متوسط میدان، درصد اشباع میانگین آب میدان و تولید گاز چاه شماره ۱ را نشان می‌دهند که همان‌طور که از شکل‌ها مشخص می‌باشد تطابق مناسبی بین نتایج شبیه‌ساز عددی و شبیه‌ساز داده‌محور وجود دارد چه برای اطلاعات دو سال ابتدایی که به‌منظور تطبیق تاریخچه استفاده شده است و چه برای سال سوم که مدل‌های داده‌محور اطلاعات سال سوم را مشاهده نکرده‌اند و برای آن پیش‌بینی انجام داده‌اند.

همچنین در جدول ۷ نیز نتایج معیارهای دقت مدل‌های داده‌محور آموزش دیده براساس اطلاعات سه سال، مورد بررسی قرار گرفته است. همان‌طور که از نتایج این معیارها مشخص است مدل‌های داده‌محور به‌خوبی توانسته‌اند رفتار مخزن در رابطه با مدل فشار متوسط مخزن، مدل درصد اشباع میانگین آب مخزن و مدل تولید گاز چاه به‌خوبی چه در مرحله تطبیق تاریخچه و چه در مرحله پیش‌بینی شبیه‌سازی نمایند.

جدول ۷ نتایج معیارهای دقت برای مدل‌های داده‌محور آموزش‌دیده با استفاده از اطلاعات مربوط به سه سال

مدل	ضریب تعیین	جذر میانگین مربع خطا	میانگین خطای مطلق	خطای میانگین درصدی مطلق	میانگین مربع خطا	بیشینه خطا
فشار متوسط میدان (Pa)	۰/۹۸۰۲	۵۵/۹۳۶۰۱۶۸۸۱	۱۳۲۰۳۴/۶۰	۰/۰۰۴۰	۱۸۴۲۹۴۰۷۰۴۱/۱۶	۱۸۶۳۴۰/۲۴
درصد اشباع میانگین آب میدان	۰/۹۷	۰/۰۰۰۳۹	۰/۰۰۰۳	۰/۰۰۱۶۱	۰/۰۰۰۰۰۱۵۳۹	۰/۰۰۰۸۰۳۴
تولید گاز چاه (m^3)	۰/۹۸۷	۲۴۲۶۸/۳۰۶	۷۹۸/۹۴	۰/۰۰۳۵	۷۳۴۴۸۸/۸۶	۱۰۹۱/۸۴

نتیجه‌گیری

در شبیه‌سازی داده‌محور نگرانی‌ای که وجود دارد این است که چون ابتدا از فیزیک مسئله استفاده نمی‌شود ممکن است این مدل حتی اگر تطبیق تاریخچه‌ی خوبی داشته باشد نتواند به‌علت عدم درک فیزیک جریان سیال در محیط متخلخل، پیش‌بینی خوبی داشته باشد اما مشاهده شد که مدل بالا به پایین که از سه مدل داده‌محور تحت عناوین فشار متوسط میدان، درصد اشباع میانگین آب و تولید گاز ساخته شده بود توانست برای این موارد خواسته‌شده فیزیک حاکم که توسط داده‌های شبیه‌ساز عددی ارائه می‌شد را درک کند و ضریب تعیین برای مدل‌های فشار متوسط مخزن، درصد متوسط اشباع آب میدان و تولید گاز چاه‌ها به‌ترتیب ۰/۹۸، ۰/۹۷ و ۰/۹۸ در مرحله تطبیق تاریخچه دیده شد که نتایج قابل قبولی هستند. علاوه‌براین مدل‌ها عملکرد مناسبی هم در زمینه پیش‌بینی سال سوم از خود نشان دادند. بنابراین در این تحقیق شبیه‌سازی داده‌محور به‌عنوان یک روش مکمل شبیه‌سازی‌های عددی معرفی می‌شود که می‌تواند هم دقت و هم سرعت قابل قبولی را ارائه کند. در آینده که انتظار می‌رود میادین موجود با پیشرفت تکنولوژی به‌سمت میادین هوشمند پیش روند و با گذشت زمان به میادین بالغ‌تری نیز تبدیل گردند، حجم بسیار زیادی از داده‌ها تولید می‌گردد که شبیه‌سازهای عددی توانایی استفاده از تمامی این داده‌ها را ندارند پس راه‌حلی که می‌تواند به مهندسین مخزن در این امر کمک کند استفاده از شبیه‌سازی داده محور و با استفاده از هوش مصنوعی است. در نهایت بعد از ساخت مدل می‌توان از مدل بالا به پایین برای مدیریت میدان استفاده نمود و صدها راهبردی مختلف را در زمان کمی اجرا و با یکدیگر مقایسه کرد و همچنین به‌دلیل سرعت بالای اجرای این مدل‌ها می‌توان برای بررسی عدم‌قطعیت‌ها و بهینه‌سازی مورد استفاده قرار گیرند. همچنین در این مقاله محدودیت‌هایی وجود داشت که به موارد زیر می‌توان اشاره نمود:

- عدم دسترسی به اطلاعات واقعی مخازن موجود در ایران که به‌علت این محدودیت مطالعه برروی یک مدل میدان فرضی انجام شده است و تمامی خواص سنگ و سیالات مورد استفاده به‌صورت پیش‌فرض از نرم‌افزار پترل مورد استفاده قرار گرفته‌اند.
- به‌علت عدم دسترسی به اطلاعات پتروفیزیکی چاه‌های میادین موجود در ایران، خواص همچون

تخلخل و تراوایی میدان مدل‌سازی‌شده در این تحقیق توسط توابع توزیع پخش شده‌اند که در واقعیت با استفاده اطلاعات نمونه‌گیری شده از چاه‌های مختلف چنین اطلاعاتی استخراج می‌شود.

مراجع

[1]. Ansari, A. (2023). Reservoir simulation of the volve oil field using AI-based top-down modeling approach (Doctoral dissertation, West Virginia University). doi.org/10.33915/etd.11970.

[2]. Sayarpour, M., Zuluaga, E., Kabir, C., & Lake, L. W. (2009). The use of capacitance–resistance models for rapid estimation of waterflood performance and optimization. Journal of Petroleum Science and Engineering, 69(3–4), 227–238. https://doi.org/10.1016/j.petrol.2009.09.006.

[3]. "Equinor, Disclosing all Volve data," 2018. [Online]. Available: https://www.equinor.com/news/archive/14jun2018-disclosing-volve-data.

[4]. Mohaghegh, S. D. (2011). Reservoir simulation and modeling based on pattern recognition. All Days. doi.org/10.2118/143179-ms.

[5]. Faisal, A., & Mohaghegh, S. D. (2016). A data-driven smart proxy model for a comprehensive reservoir simulation.

[6]. Schuld, M., & Petruccione, F. (2018). Supervised learning with quantum computers. In Quantum science and technology.doi.org/10.1007/978-3-319-96424-9.

[7]. Donald J. Ford, P. (2011). Training industry. Retrieved from How the Brain Learns: trainingindustry.com/articles/content-development/how-the-brain-learns/.

[8]. Shahkarami, A., Mohaghegh, S. D., Gholami, V., & Haghighat, S. A. (2014). Artificial intelligence (AI) assiste d history matching. All Days. doi.org/10.2118/169507-ms.

[9]. Haykin, S. S. (2010). Neural networks and learning machines. ci.nii.ac.jp/ncid/BB00465945.

[10]. Hernandez, A. (2016). Model calibration with neural networks. SSRN Electronic Journal. doi.org/10.2139/ssrn.2812140.

[11]. Mohaghegh, S. D. (2017b). Shale Analytics. In Springer eBooks. doi.org/10.1007/978-3-319-48753-3.

[12]. Mohaghegh, S. D. (2017). Data-Driven reservoir modeling. In Society of Petroleum EngineersRichardson, Texas, USA eBooks. doi.org/10.2118/9781613995600.

[13]. Mohaghegh, S. D., Gaskari, R. ., Maysami, M. ., & Khazaeni, Y. (2014). Data-Driven reservoir management of a giant mature oilfield in the Middle East. All Days. doi.org/10.2118/170660-ms.

[14]. Haifi, A. H. M. A. (2019). Confirmation of data-driven reservoir modeling using numerical reservoir simulation. In Partial Fulfillment of the Requirements for the Degree of Master of Science in Petroleum and Natural Gas Engineering doi.org/10.33915/etd.3835.

[15]. Gomez, Y., Khazaeni, Y., Mohaghegh, S. D., & Gaskari, R. (2009). Top-Down intelligent reservoir modeling (TDIRM). All Days. doi.org/10.2118/124204-ms.

[16]. Haghighat, S. A., Mohaghegh, S. D., Gholami, V., & Moreno, D. (2014). Production analysis of a Niobrara field using intelligent Top-Down modeling. All Days. doi.org/10.2118/169573-ms.

[17]. Maysami, M., Gaskari, R., & Mohaghegh, S. D. (2013, September). Data driven analytics in powder river basin, WY. In SPE Annual Technical Conference and Exhibition? (p. D031S033R002). SPE.

[18]. Alatrach, Y., Saputelli, L., Narayanan, R., Mohan, R., Alklih, M. Y., & Rubio, E. (2019, November). Data-driven vs. traditional reservoir numerical models: A case study comparison of applicability, practicality and performance. In Abu Dhabi International Petroleum Exhibition and Conference (p. D021S030R002). SPE.

[19]. Zargari, S. & Mohaghegh, S. D. (2010). Field development Strategies for bakken shale Formation. All Days. doi.org/10.2118/139032-ms.

[20]. Mohaghegh, S. D., Gruic, O., Zargari, S., Dahaghi, A. K., & Bromhal, G. S. (2012). Top-down, intelligent reservoir modelling of oil and gas producing shale reservoirs: case studies. International Journal of Oil Gas and Coal Technology, 5(1), 3. doi.org/10.1504/ijogct.2012.044175.

[21]. Martinez, Y. A. (2020). Top-Down model development using data generated from a complex numerical reservoir simulation with water injection. In Partial Fulfillment of the Requirements for the Degree of Master of Science in Petroleum and Natural Gas Engineerin doi.org/10.33915/etd.7571.

Petroleum Research
Petroleum Research, 2025(August-September), Vol. 35, No. 142, 7-10
DOI: 10.22078/pr.2025.5524.3456

Data-driven Simulation of a Hypothetical Oil Field and Comparison of Results with Numerical Simulation

Mahdi Rashnoo and Saeid Jamshidi*
Faculty of Chemical and Petroleum Engineering, Sharif university of Technology, Tehran, Iran
jamshidi@sharif.edu
DOI: 10.22078/pr.2025.5524.3456

Received: September 01, 2024 Accepted: March 11, 2025

Introduction

In the oil and gas industry, due to existing complexities and the very high costs of reservoir development, having a model with appropriate accuracy and a small computational footprint is of great importance. Also, the conventional approach is to use a numerical reservoir simulator, but the novel method is simulation based on machine learning. This approach is carried out using real field measurements [1]. In numerical modeling, it is believed that these functional relationships are correct, deterministic, and unchangeable. Therefore, if the output from numerical modeling does not match field measurements, it is concluded that the reservoir properties (static model) may not be ideally measured and interpreted and require modification [2]. Numerical models are not fast and require significant investment, time, and manpower. AI-based reservoir modeling is a novel technology with high potential because it can produce reservoir models in a fraction of the time and budget of traditional models while offering most of their capabilities and can be a substitute or complement to the conventional method. In this research, attempt has been made to create models for a hypothetical oil field using existing data with artificial intelligence for the desired outputs, and finally, the results are compared with a numerical simulator.

Materials and Methods

Using PETREL software, a square field at a depth of 914.4 meters from the surface and a total thickness of 182.88 meters is considered, which it has two layers with different thicknesses. After determining the porosity and permeability characteristics, the fluid is light oil and gas, which has a bubble point pressure of 2999999.77 Pascal and API 45, and the reservoir rock is consolidated limestone. Moreover, the relative permeability and capillary pressure are also selected as the software's default values. In addition, the field has 16 production wells and 9 injection wells without casing in a five-spot pattern, and data collection is done daily. Then, a database is created using X and Y Location, porosity at the well location, average porosity in the well drainage area, bottom-hole pressure, oil production, water injection, gas-oil ratio, average field pressure, average water saturation, and gas production. Furthermore, for each well, 4 offset wells are selected based on the drilling pattern. Using the Python programming language, a static model similar to a numerical simulator is designed, and the average porosity of networks in a drainage area is calculated (using Voronoi graph theory). In the next step, three data-driven models are considered, which are the average field pressure, average water saturation, and gas production, respectively.

For training the data-driven models, an artificial neural network with one hidden layer and one output node has been used. Furthermore, to reduce the number of inputs, a Correlation Heatmap and the Spearman method have been employed. For normalizing the data values, the Min-Max method is used. The mean squared error is chosen as the loss function during model training, and the mean absolute error is selected as the performance evaluation metric. In parameter optimization, efforts are made to optimize the number

of nodes in the hidden layer, the learning rate, and epochs. To prevent overfitting, the dropout method along with its optimization is used. After finding the most suitable parameters, the model training will be carried out, and then the model will be evaluated using the R2 method on the test data and for visualization, the predicted values will be plotted against the actual values. Finally, using data from production well 1, the target value will be predicted and plotted alongside the value obtained from the numerical simulator. Ultimately, for all models and after history matching in previous steps, the constructed models will be used to predict the reservoir performance from 2026 to 2027 using data that has never been seen before.

Results and Discussion

In Fig. 1, the values predicted by the data-driven models are plotted against the values calculated by the numerical model, indicating the appropriate accuracy of the data-driven model and its learning of fluid flow physics in porous media. In Fig. 1b, the data-driven model predicts a constant value for the average water saturation percentage for a range of initial and small values. After reviewing the data in the database, it was found that the relevant numbers in the early days are identical up to five decimal places; in addition, we used daily information that it was collected in the early years of field development.

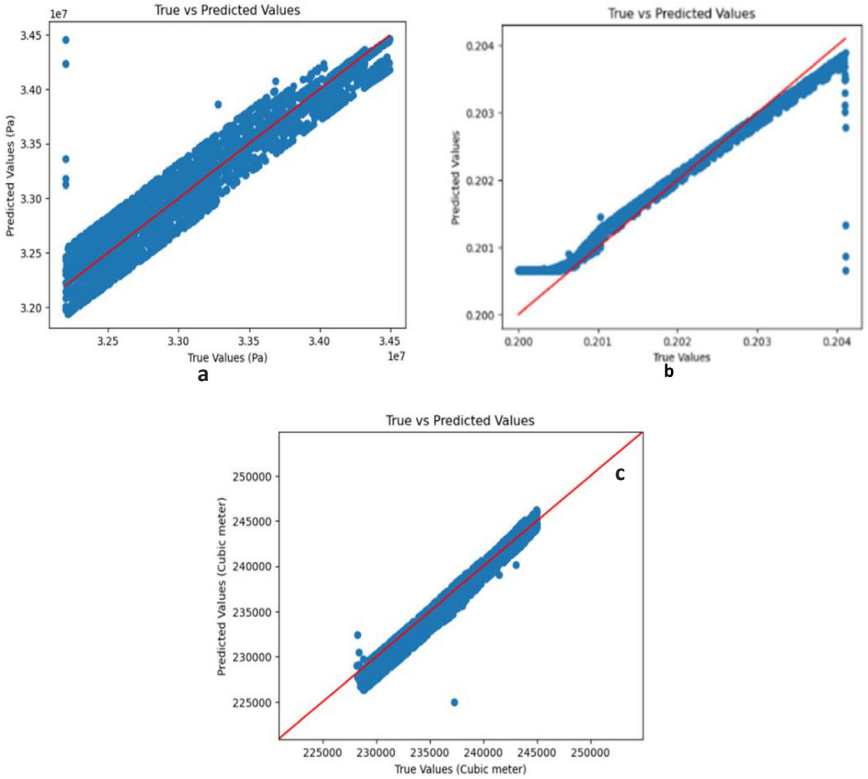


Fig. 1 Plot of predicted values versus actual values in different data-driven models: a) average reservoir pressure, b) average water saturation and c) gas production.

In Fig. 2, the history matching results by the data-driven models and the numerical model based on the information from well number 1 are displayed. Now, that the models have been validated and their performance in the history matching process has been confirmed, they have been used for predictions from 2026 to 2027 based on the information from well number 1. Fig. 3 shows the results of history matching and prediction (two years of history matching and one year of prediction) for the average field pressure, average water saturation of the field, and gas production of well number 1.

Conclusions

In data-driven simulation, there is a concern that since the physics of the problem is not used initially, the model may not be able to make good predictions, even if it has a good history fit, due to a lack of understanding of fluid flow physics in porous media. However, it was observed that the top-down model, which was built from three data-driven models under the titles of average reservoir pressure, average water saturation and gas production, was able to understand the governing physics provided by the numerical simulator data for these requested outputs. In addition to showing good history matching, it also demonstrated appropriate performance in prediction.

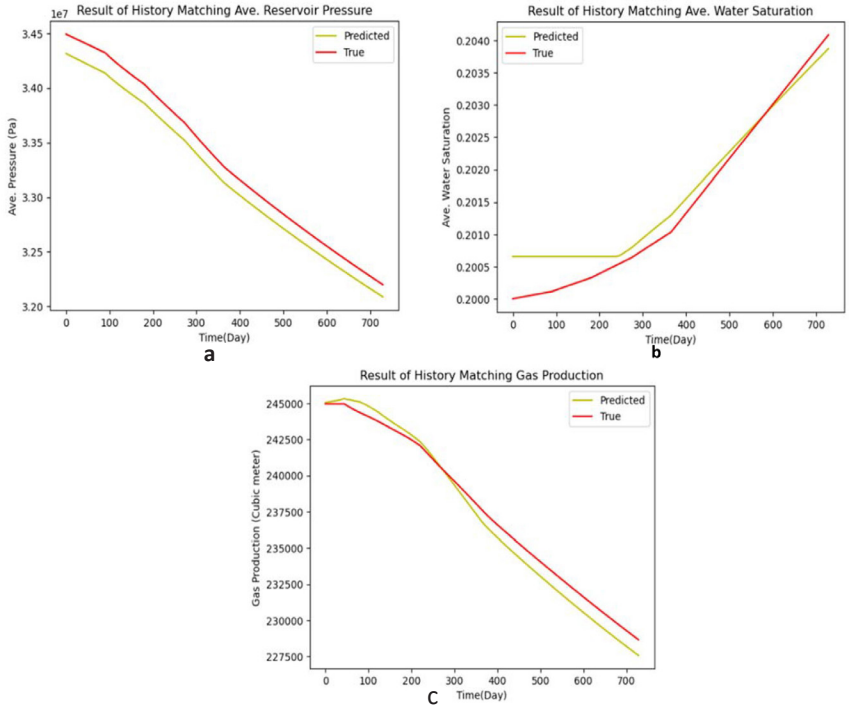


Fig. 2 Plot of history matching data for different models using data from production well number 1 and comparison with numerical values: a) average reservoir pressure, b) average water saturation and c) gas production.

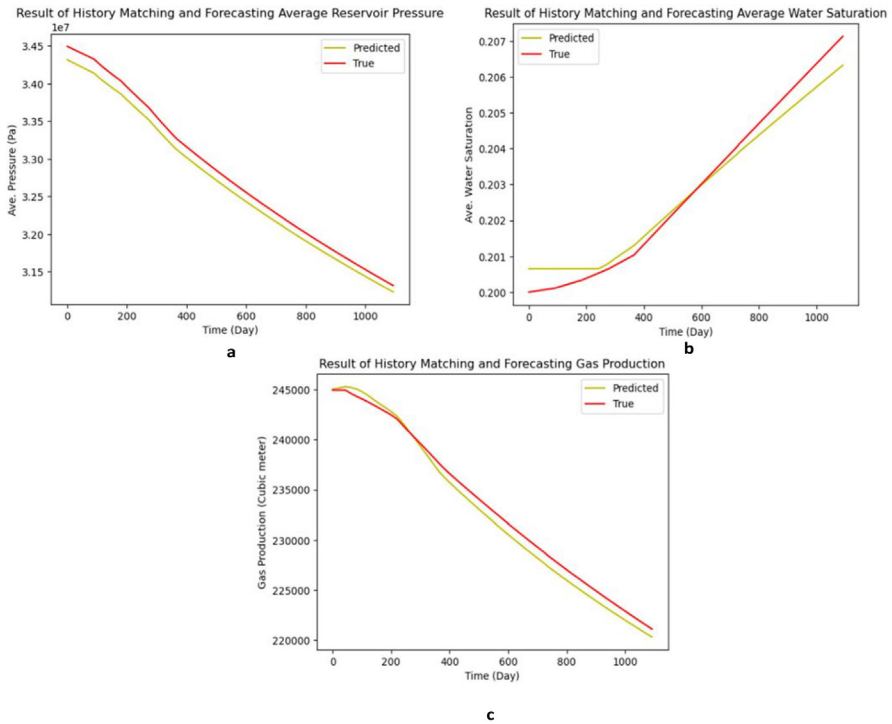


Fig. 3 The results of history matching and prediction by different models compared to the numerical method using information from well number 1: a) average field pressure, b) average water saturation and c) gas production.

Therefore, in this research, data-driven simulation is introduced as a complementary method to numerical simulations, which it can offer both acceptable accuracy and speed. In the future, as existing fields are expected to advance towards smart fields with technological progress and over time become more mature fields, a large volume of data will be generated that numerical simulators may not be able to utilize fully. Thus, a solution that can assist reservoir engineers in this matter is the use of data-driven simulation with artificial intelligence. Ultimately, after building the model the top-down model can be used

for field management, executing hundreds of different scenarios in a short time and comparing them with each other. Additionally, due to the high speed of executing these models, they can be used for uncertainty analysis and optimization.

References

1. Ansari, A. (2023). Reservoir Simulation of the Volve Oil field using AI-based Top-Down Modeling Approach. https://doi.org/10.33915/etd.11970.
2. Mohaghegh, S. D. (2011). Reservoir simulation and modeling based on pattern recognition. All Days. https://doi.org/10.2118/143179-ms.