

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

مهدی رشنو، سعید جمشیدی*

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

jamshidi@sharif.edu

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^۱ 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 به دست می‌آیند. روش‌های عددی مرسوم با استفاده از مدل‌سازی‌های عددی و فهم کنونی که از فیزیک حاکم بر جریان سیال در محیط متخلخل وجود دارد، فرایند شبیه‌سازی را انجام می‌دهند.

در این تحقیق تلاش شده است که برای یک مورد میدان فرضی در نرم‌افزار پترل^۲ مدل‌سازی انجام شود و سپس از داده‌های خروجی از این نرم‌افزار به عنوان ورودی برای شبیه‌سازی مدل بالا به پایین مدنظر با استفاده از زبان برنامه‌نویسی پایتون^۳ استفاده گردیده و پس از ساخت سه مدل داده‌محور به ترتیب 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. به عبارت 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 فرایند با استفاده از اطلاعات دو سال ابتدایی تولید از میدان، تطبیق تاریخیچه انجام شده است و برای مدل فشار متوسط مخزن، درصد اشباع آب متوسط میدان و تولید گاز، ضریب تعیین به ترتیب ۰/۹۸۰۲، ۰/۹۷ و ۰/۹۸۷ محاسبه شده است. همچنین برای سال سوم پیش‌بینی انجام خواهد شد و نتیجه‌ی به دست آمده از مدل‌های داده‌محور با نتایج به دست آمده توسط شبیه‌ساز عددی مقایسه می‌گردد.

^۱ TDM: Top-Down Reservoir Modeling

^۲ PETREL Software

^۳ Python

۱- مقدمه

در صنعت نفت و گاز به دلیل پیچیدگی های موجود و هزینه های 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 فعالیت های مهندسی مخزن برای بررسی حساسیت سنجی، عدم قطعیت ها و بهینه سازی موارد هدف در امور مدیریت مخزن قرار می گیرد باید ردپای محاسباتی کوچکی داشته باشد. رویکردهای متعددی برای شبیه سازی مخزن وجود دارد: ۱. رویکرد کلاسیک^۱ ۲. شبیه ساز مخزن عددی^۳. رویکردهای مبتنی بر یادگیری ماشین^۴ ۱. رویکرد اول ایجاد یک مدل ریاضی ساده و حل برای یافتن برخی ضرایب است که اساسا یک مسئله برازش منحنی است روش های مطرح در این زمینه تحلیل منحنی نزول^۲ [۲] و مدل سازی مقاومت خازنی^۳ [۳] است که اطلاعات محدودیت های عملیاتی، طراحی چاه ها و ... را در نظر نمی گیرند. روش دوم ایجاد یک مدل ریاضی با پیچیدگی های بیشتر و گسسته کردن زمان و مکان و حل معادلات به صورت عددی است. رویکرد سوم ساخت یک مدل با استفاده از اندازه گیری های واقعی از میدانی است. [۱] در مدل سازی 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 غیر قابل تغییر هستند. بنابراین اگر تولید حاصل از مدل سازی عددی با اندازه گیری های ما از میدان مطابقت نداشته باشد نتیجه می گیریم که ویژگی های مخزن (مدل ثابت^۴) ممکن است به طور ایده آل اندازه گیری و تفسیر نشوند و باید اصلاح شوند. [۴] مدل های عددی سریع نیستند و نیازمند سرمایه گذاری زیاد، زمان و نیروی انسانی هستند. مدل های پروکسی هوشمند^۵ جایگزینی برای شبیه سازی های عددی هستند که به مهندسان مخزن در توسعه مخزن و مدیریت تولید کمک می کنند. [۵] در روش های عددی برای رسیدن به نتایج مطلوب مدل ثابت دچار تغییر می شود و برای ایجاد امکان حل مسئله فرض های ساده شده ای در نظر گرفته می شود که مسئله را از حالت واقعی خود دور می کند. باید انتظار داشت همانطور که در طول زمان درک موجود نسبت به این مسئله توسعه یافته است در زمان های آینده نیز به حد کمال خود برسد و شاید درک کنونی دارای نواقصی باشد. در صنعت نفت و گاز یکی از محبوب ترین شیوه ها، یادگیری نظارت شده است. [۶] درک شبکه های عصبی نیاز به درک علوم اعصاب دارد. [۷] و شبکه عصبی مصنوعی به عنوان یکی از فناوری های هوش مصنوعی طبقه بندی می شود. [۸] در فرایند آموزش از فرایند تکرار پسانتشار^۶ تا زمان رسیدن به همگرایی وزن بین نورون ها^۷ اصلاح می شود [۹] و برای جلوگیری از بیش برآزش^۸، داده های

^۱ Classic Approaches

^۲ Decline Curve Analysis

^۳ Capacitance Resistance Modeling

^۴ Static Model

^۵ Smart Proxy Model

^۶ Backpropagation

^۷ Neuron

^۸ Overfitting

کالیبراسیون برای تایید قابلیت‌های تعمیم مدل استفاده می‌شوند. [۱۰] در این صنعت از هوش مصنوعی برای حل مسائل مربوط به تجزیه و تحلیل فشار گذرا^۱، حفاری، شناسایی ویژگی‌ها مخزن^۲ و تفسیر نمودار چاه استفاده شده است. [۱۱] شهاب محقق یکی از پیشگامان به کارگیری هوش مصنوعی در مهندسی نفت، مدل‌سازی داده محور را به عنوان فرایند به کارگیری هوش مصنوعی و روش‌های داده‌کاوی مانند یادگیری ماشین و تشخیص الگو به منظور کشف الگوهای پنهان در مجموعه داده‌های بزرگ که نشان دهنده جریان سیال در محیط‌های متخلخل هستند تعریف کرد. سپس الگوهای کشف‌شده در یک مدل مخزنی جامع و منسجم با قابلیت پیش‌بینی مورد تأیید گنجانده شده‌اند که می‌تواند برای مدیریت میادین و تولید آنها مورد استفاده قرار گیرد. فناوری داده‌محور کاملاً 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. در حالی که فاز سوم مدل‌سازی مخزن داده‌محور تحلیل پس داده^۳ 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 فناوری 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 است. (۲) فرض نمی‌شود که به طور 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 فناوری پیشگام است که تعداد زیادی از رشته‌ها مانند مدل‌سازی مخزن، مهندسی مخزن، تجزیه و تحلیل داده محور 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 فناوری نوین است و پتانسیل این فناوری 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 کاربردی بودن 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^۴ در حوضه رودخانه پودر وایومینگ^۵ [۱۷] پیاده‌سازی کردند و علاوه بر تطبیق تاریخیچه با استفاده از مدل‌های اعتبارسنجی‌شده برای چند سال پیش‌بینی را در مقایسه با روش‌های عددی انجام داده‌اند که نتایج خوبی حاصل شده است. یارا آل‌اتراک^۶ مطالعه‌ای را بر روی یک مخزن کربناته در ابوظبی انجام داد و توانست با استفاده از داده‌های تاریخیچه برای بهینه‌سازی تولید و تزریق برای به حداکثر رساندن تولید کل میدان، مدل را آموزش دهد. [۱۸] زرگری و محقق در مدل‌سازی داده‌محور مخزن را بر روی بخشی از سازند شیل باکن در حوضه ویلیستون داکوتای شمالی^۷ پیاده‌سازی کرده‌اند. از مدل آنها می‌توان برای پیشنهاد استراتژی‌های توسعه، شناسایی ذخایر باقیمانده و در نتیجه مکان چاه‌های جدید استفاده کرد. علاوه بر این یک مدل

^۹ Pressure transient analysis

^{۱۰} Reservoir characterization

^۱ Post data analysis

^۲ Hilight Field

^۳ Powder River Basin Wyoming

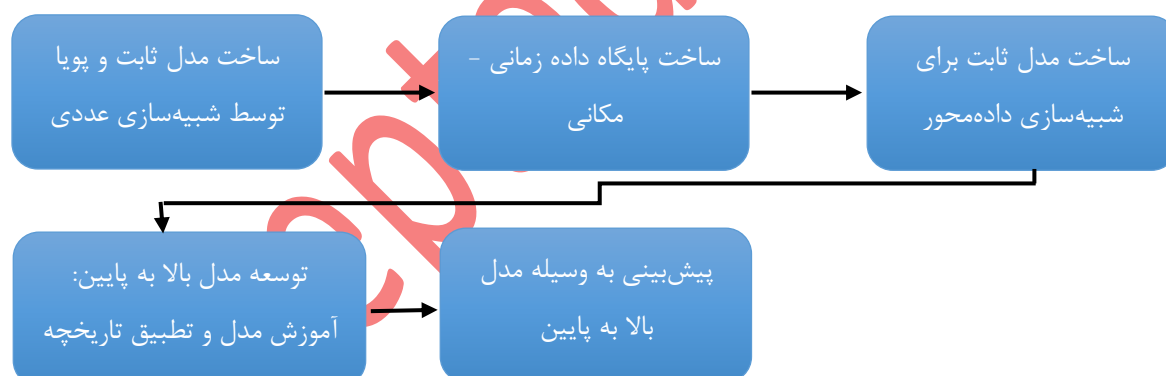
^۴ Yara Alatrach

^۵ Williston basin of North Dakota

پیش بینی کننده تولید شده و تاریخچه‌ی 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 استراتژی‌های توسعه میدان از آنها استفاده کردند. [۲۰] حسن الحیفی نتایج تطبیق تاریخچه و پیش‌بینی مدل پیش‌بینی 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 فناوری‌ها به حجم زیادی از داده‌های با کیفیت نیاز دارند تا بتوانند یک مدل پیش‌بینی دقیق به دست آورند. [۱۴] همچنین مارتینز برای یک میدان فرضی عملکرد مدل‌های داده‌محور آموزش دیده را تحت سناریوهای مختلف مورد بررسی قرار داد. [۲۱] در این مقاله برای اولین بار در ایران است که 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^۱ برای تشخیص اثرگذارترین پارامترها بر روی خروجی مدنظر استفاده شده است در این تحقیق از دو روش مرسوم در یادگیری ماشین (نقشه حرارتی^۲ و روش اسپیرمن^۳) برای یک میدان نفتی فرضی استفاده شده است.

۲- روش‌شناسی

شکل ۱ روند انجام این مدل‌سازی را طی پنج مرحله نشان داده است که با استفاده از زبان برنامه‌نویسی پایتون^۴ انجام شده است.



شکل ۱. روند انجام شبیه‌سازی داده‌محور

۲-۱ ساخت مدل ثابت و پویا توسط شبیه‌سازی عددی

^۱ Fuzzy Logic

^۲ Heatmap Correlation

^۳ Spearman method

^۴ Python

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

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 (مترمربع)	لاگ نرمال	$۷/۸۲۸۹ \times ۱۰^{-۱۴}$	$۲/۴۶۱۲ \times ۱۰^{-۱۵}$	۲۳۹۴	$۵/۹۲۱۵ \times ۱۰^{-۱۴}$	$۷/۸۹۵۴ \times ۱۰^{-۱۴}$
تراوایی لایه دوم (مترمربع)	لاگ نرمال	$۶/۸۴۲۰ \times ۱۰^{-۱۴}$	$۲/۴۶۱۲ \times ۱۰^{-۱۵}$	۱۲۱۰۴	$۵/۹۲۱۵ \times ۱۰^{-۱۴}$	$۷/۸۹۵۴ \times ۱۰^{-۱۴}$

در رابطه با خواص سیال، نفت سبک و گاز در نظر گرفته شده است. جدول ۲ بیانگر مهم‌ترین خواص تعریف‌شده است:

جدول ۲. خواص سیال میدان

وزن مخصوص گاز	۰/۶۵
ای پی آی نفت	۴۵
فشار نقطه حباب (پاسکال ^۳)	۲۹۹۹۹۹۹۹/۷۷
شوری آب	۳۵۰۰۰ قسمت در میلیون
فشار اولیه در عمق مرجع (پاسکال)	۳۴۰۴۶۳۱۱/۵

جنس سنگ مخزن از نوع سنگ آهک تثبیت‌شده است. خواص سنگ و سیال همچون تراوایی نسبی (با استفاده از رابطه کوری^۴) و فشار موئینگی هم مقدار پیش فرض 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 حفر شدن شروع به تولید/تزریق می‌کنند. (۲) فاصله یک چاه با چاه کناری را ۱۵۲/۴ متر فرض شده که با توجه به در نظر گرفتن ابعاد میدان حداکثر می‌توان ۲۵ چاه در آن حفر نمود. (۳) حداقل فشار ته چاهی را ۱۳۷۸۹۵۱۴/۵۸ پاسکال و حداقل تولید نفت از یک چاه را ۷/۹۴۹۳۶۵ مترمکعب در روز فرض شده است. (۴) حداکثر ظرفیت

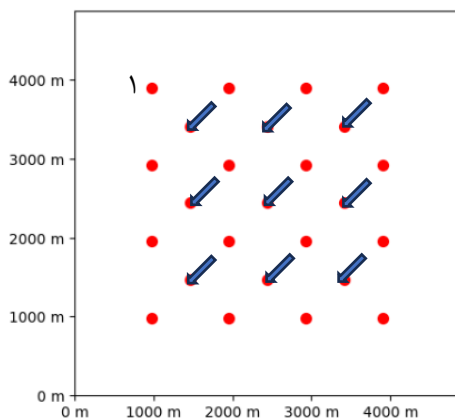
^۵ PETREL Software

^۶ Grid

^۱ Pascal

^۲ Corey's Correlation

خطوط لوله انتقال سیال (نفت و آب) را ۷۹۴/۹۳۶۵ مترمکعب در روز و همچنین حداکثر ظرفیت تزریق آب را نیز همین مقدار فرض شده است. ۵) فشار تزریق ته‌چاهی برای جلوگیری از شکاف‌خوردن ۴۱۳۶۸۵۴۳/۷۴ پاسکال و مقدار کسر جایگزین‌شده^۱ 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^۲ حفر شده است و همگی بدون لوله‌ی جداری هستند و تمامی چاه‌ها از هر دو لایه تولید می‌کنند. محل حفر چاه‌ها و نوع آنها در شکل ۲ نمایش داده شده است.



شکل ۲. محل چاه‌ها و نوع آنها

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 درنهایت در تاریخ ۰۱/۰۱/۲۰۲۵ تمامی ۲۵ چاه شروع به فعالیت تا تاریخ ۰۱/۰۱/۲۰۲۶ می‌کنند. نحوه‌ی گزارش داده‌ها در این دو سال به صورت روزانه است.

۲-۲ ساخت پایگاه داده زمانی - مکانی

تهیه پایگاه داده مهم‌ترین کار در مدل‌سازی داده‌محور است. در این پایگاه داده دو دسته داده‌ی ثابت و پویا وجود دارد که دسته ثابت به دو دسته واقعا ثابت و ثابت اصلاح‌شده^۳ 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 به عنوان مثال تخلخل متوسط در ناحیه زهکشی آن چاه که میانگینی از تخلخل گریدهای موجود در آن است متفاوت خواهد شد. لیست این داده‌ها در جدول ۳ آورده شده است. در بین داده‌ها تراوایی در محل چاه و در ناحیه زهکشی یک چاه به چشم نمی‌خورد. علت این امر آن است که در رابطه با تراوایی متوسط در ناحیه زهکشی یک چاه، به علت خواص مخزنی انتخاب شده بعد از محاسبه تراوایی متوسط برای هر دولایه و بعد از گرفتن میانگین وزنی برای تمام ناحیه‌های زهکشی مقداری یکسان به دست آمد که به همین خاطر از ورودی‌ها حذف شده است. چون چاه‌های مجاور بر تولید یک چاه مرکزی اثرگذار هستند در این تحقیق با توجه به الگوی حفر چاه‌ها، ۴ چاه اطراف یک چاه به عنوان چاه‌های مجاور چاه مرکزی در نظر گرفته می‌شوند و علاوه بر اطلاعات یک چاه مرکزی

^۳ Group voidage replacement

^۴ Five Spot Pattern

^۱ Dynamically modified Static data

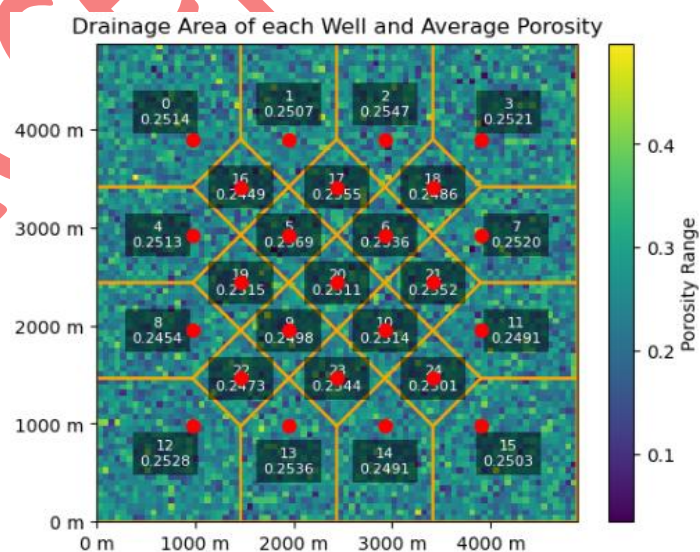
در ستون‌های مجاور اطلاعات چاه‌های مجاور آن گنجانده می‌شود. این باعث می‌شود که یک چاه هم به عنوان چاه مرکزی و هم به عنوان چاه مجاور چندین بار مورد مطالعه توسط مدل‌های داده‌محور قرار گیرد.

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 ته چاهی، تولید نفت، تزریق آب، نسبت گاز به نفت، فشار متوسط میدان، درصد اشباع متوسط آب میدان و تولید گاز

۲-۳ ساخت مدل ثابت برای شبیه‌سازی داده‌محور

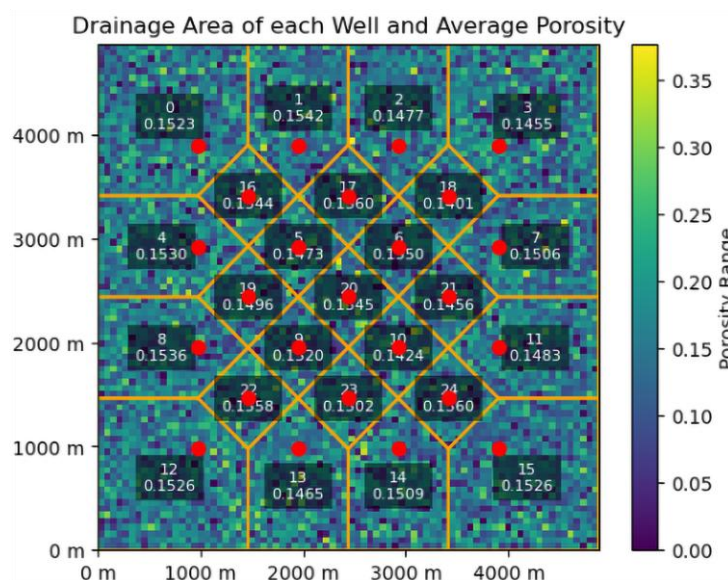
در مدل داده‌محور، میدان به گریدهایی با ضلع ۶۰.۹۶ متری تقسیم‌بندی شده و بر طبق ویژگی‌های جدول شماره ۳-۱ خواص تخلخل تعریف می‌شود و میانگین مقدار تخلخل گریدهای داخل یک ناحیه زهکشی محاسبه می‌گردد. به علت تغییر مساحت ناحیه زهکشی یک چاه بعد از حفر چاه‌های جدید، این کار برای هر پنج مرحله حفر چاه و هم برای هر دو لایه انجام می‌گردد. برای تعیین ناحیه زهکشی هر چاه از نظریه وورونوی^۲ استفاده شده است. در شکل‌های ۳ و ۴ تخلخل میانگین در هر ناحیه زهکشی برای لایه‌ی بالا و پایین پس از حفر تمام چاه‌ها نشان داده شده است.



^۱ Grid

^۲ Voronoi Graph theorem

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

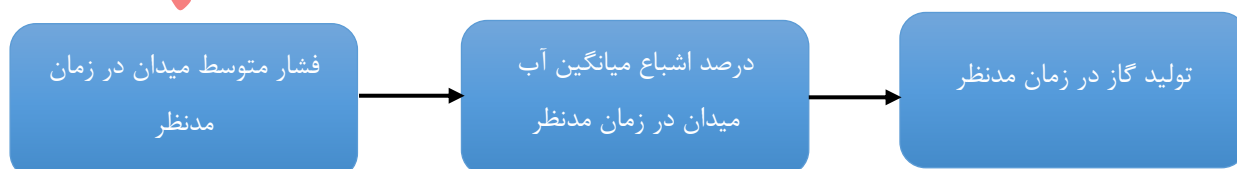


شکل ۴. ناحیه زهکشی چاه‌ها در مراحل مختلف به همراه تخلخل میانگین هر ناحیه برای لایه دوم مخزن

چون میدان دو لایه دارد برای پارامترهای تخلخل و تخلخل میانگین از داده‌های لایه اول و دوم بر اساس ضخامت لایه‌ها میانگین وزنی گرفته شده و در پایگاه داده قرار داده شده است.

۲-۴ 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:



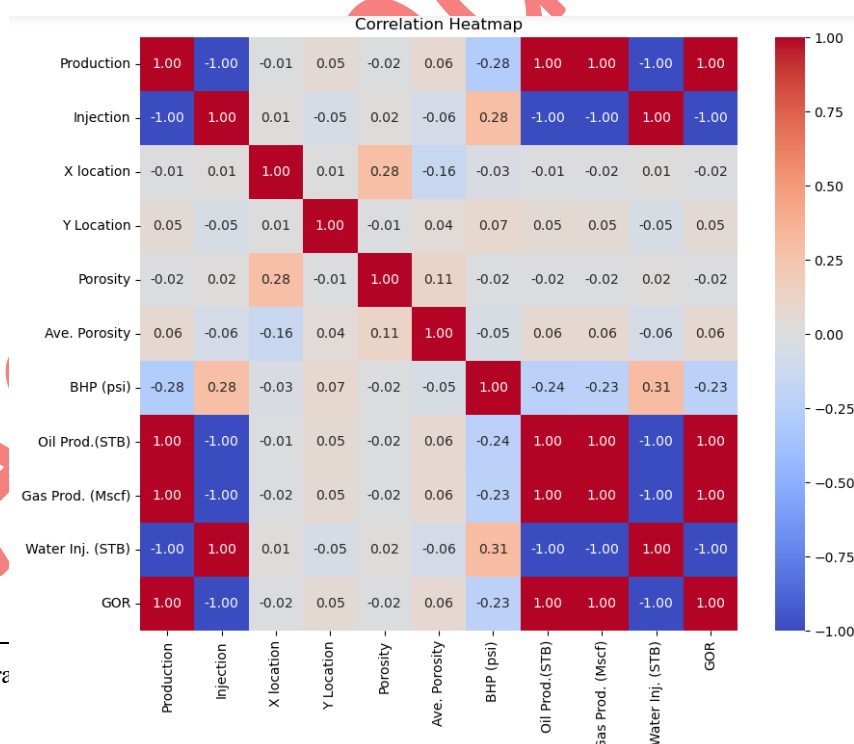
شکل ۵. عنوان و ترتیب مدل‌های داده‌محور مورد نیاز برای توسعه مدل بالا به پایین

در شکل ۵ برای پیش‌بینی خروجی فشار متوسط میدان در گام زمانی مشخص، علاوه بر 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.

جدول ۴. تعداد داده‌های استفاده‌شده در ساخت یک مدل داده‌محور

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

همچنین برای آموزش یکی از مدل‌های داده‌محور مدل بالا به پایین از شبکه‌ی عصبی مصنوعی^۱ که یک لایه‌ی پنهان^۲ و یک گره^۳ خروجی دارد استفاده شده و تمامی گره‌ها دو لایه‌ی مجاور، به یکدیگر متصل^۴ هستند. به عبارت دیگر از شبکه‌ی عصبی با مدل پرسپترون^۵ چند لایه (با یک لایه‌ی پنهان) استفاده شده است و الگوریتم مورد استفاده برای بهبود وزن‌های شبکه‌ی عصبی پس‌انتشار^۶ می‌باشد. طبق شکل ۶ برای کاهش تعداد ورودی‌ها، ارتباط بین داده‌های ورودی با استفاده از روش‌های مرسوم^۷ در یادگیری ماشین بررسی شده است.



^۱ Artificial Neural Network

^۲ Hidden Layer

^۳ Node/Neuron

^۴ Fully interconnected network

^۵ Perceptron

^۶ Backpropagation

^۷ Correlation Heatmap

شکل ۶. ارتباط بین پارامترهای ورودی با استفاده از روش‌های مرسوم در یادگیری ماشین همانطور که از شکل ۶ پیدا می‌باشد ستون داده‌های ورودی مربوط به نوع تولیدی/تزریقی بودن و نسبت گاز به نفت تولیدی را می‌توان حذف نمود (نام چاه‌ها و تاریخ هم به علت حضور طول و عرض جغرافیایی و ماهیت داده‌های پویا حذف می‌شوند). علاوه بر این طبق شکل ۷ از روش اسپیرمن^۱ به منظور مشخص‌سازی ورودی‌هایی با بیشترین اثرگذاری بر روی 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 کم‌اثر بودن 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^۲ مقادیر داده‌ها روش کمینه و بیشینه^۳، برای تابع فعال‌سازی^۴ بین لایه ورودی و لایه پنهان تابع رلو^۵، برای تابع فعال‌سازی بین لایه پنهان و لایه خروجی تابع خطی^۶، برای معیار کاهش^۷ در طول آموزش مدل، میانگین مربع خطا^۸ و به عنوان معیار بررسی عملکرد، میانگین خطای مطلق^۹ استفاده شده است. در بحث بهینه‌سازی پارامترها هم با استفاده از روش هایپربرد^{۱۰} (با معیار میانگین خطای مطلق) تلاش شده است که تعداد گره‌های لایه میانی و نرخ یادگیری^{۱۱} بهینه شود. از سوی دیگر برای بهینه‌سازی تعداد مشاهده‌ی داده‌های ورودی^{۱۲} به صورت دستی اقدام شده است. در مدل فشار متوسط میدان بعد از پیدا کردن بهترین ساختار مدل و پیدا کردن مناسب‌ترین پارامترها آموزش مدل انجام خواهد شد و در نهایت برای ارزیابی مدل توسط معیار ضریب تعیین^{۱۳} بر روی داده‌های آزمون سنجیده می‌شود و برای بصری‌سازی مقادیر پیش‌بینی شده بر حسب مقادیر واقعی رسم می‌شوند. 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 بر روی 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

^۱ Spearman's rank coefficient

^۲ Normalization

^۳ MinMaxScaler

^۴ Activation Function

^۵ ReLU Activation Function

^۶ Linear Function

^۷ Loss Function

^۸ Mean Squared error

^۹ Mean Absolute error

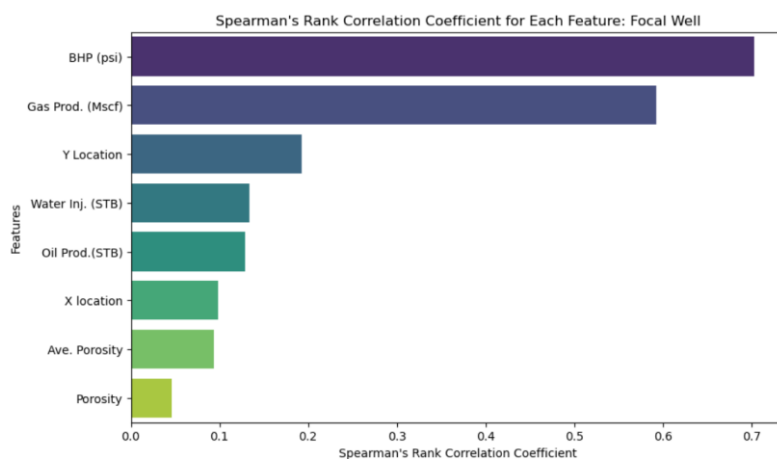
^{۱۰} Hyperband

^{۱۱} Learning rate

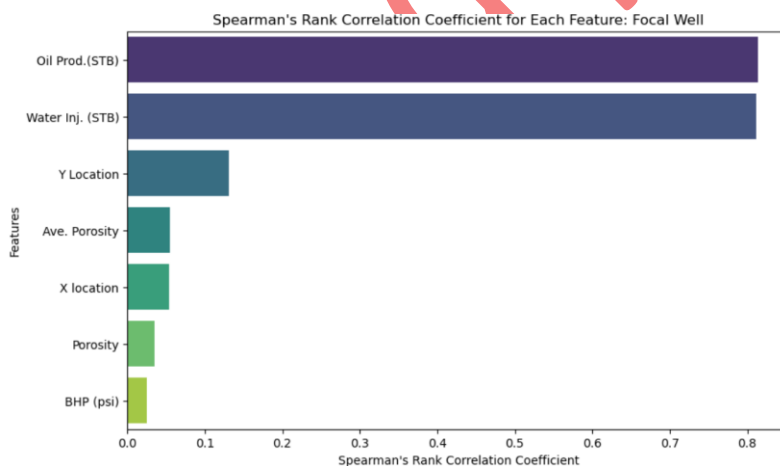
^{۱۲} Epoch

^{۱۳} R² Score

نسبت گاز به نفت تولیدی) حذف شده‌اند. برای حفظ فیزیک مسئله طول و عرض جغرافیایی چاه مرکزی حذف نشده است ولی در چاه‌های مجاور به علت اثرگذاری کم آنها حذف شده‌اند. برای تمامی مدل‌ها و به منظور جلوگیری از بیش‌برازش این مدل‌ها از روش حذف موقت^۱ تعدادی از گره‌های لایه پنهان در حین آموزش استفاده شده است که مقدار بهینه به صورت حدس و خطا و در بازه‌ی ۱۰ تا ۵۰ درصد انتخاب شده است.



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



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

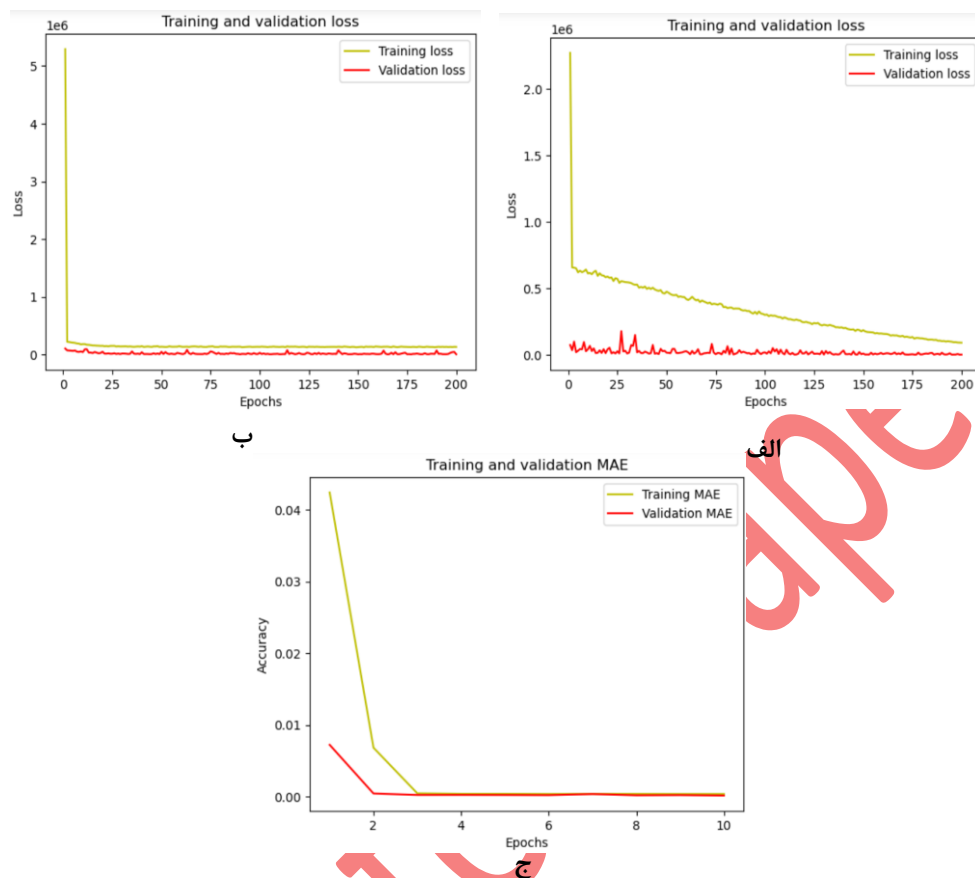
۲-۵ پیش‌بینی به وسیله مدل بالا به پایین

پس از تطبیق تاریخچه در مراحل قبل از مدل‌های ساخته‌شده استفاده شده تا با استفاده از داده‌هایی که هرگز مشاهده نکرده است عملکرد میدان از سال ۲۰۲۶ تا ۲۰۲۷ پیش‌بینی شود.

۳- نتایج و بحث

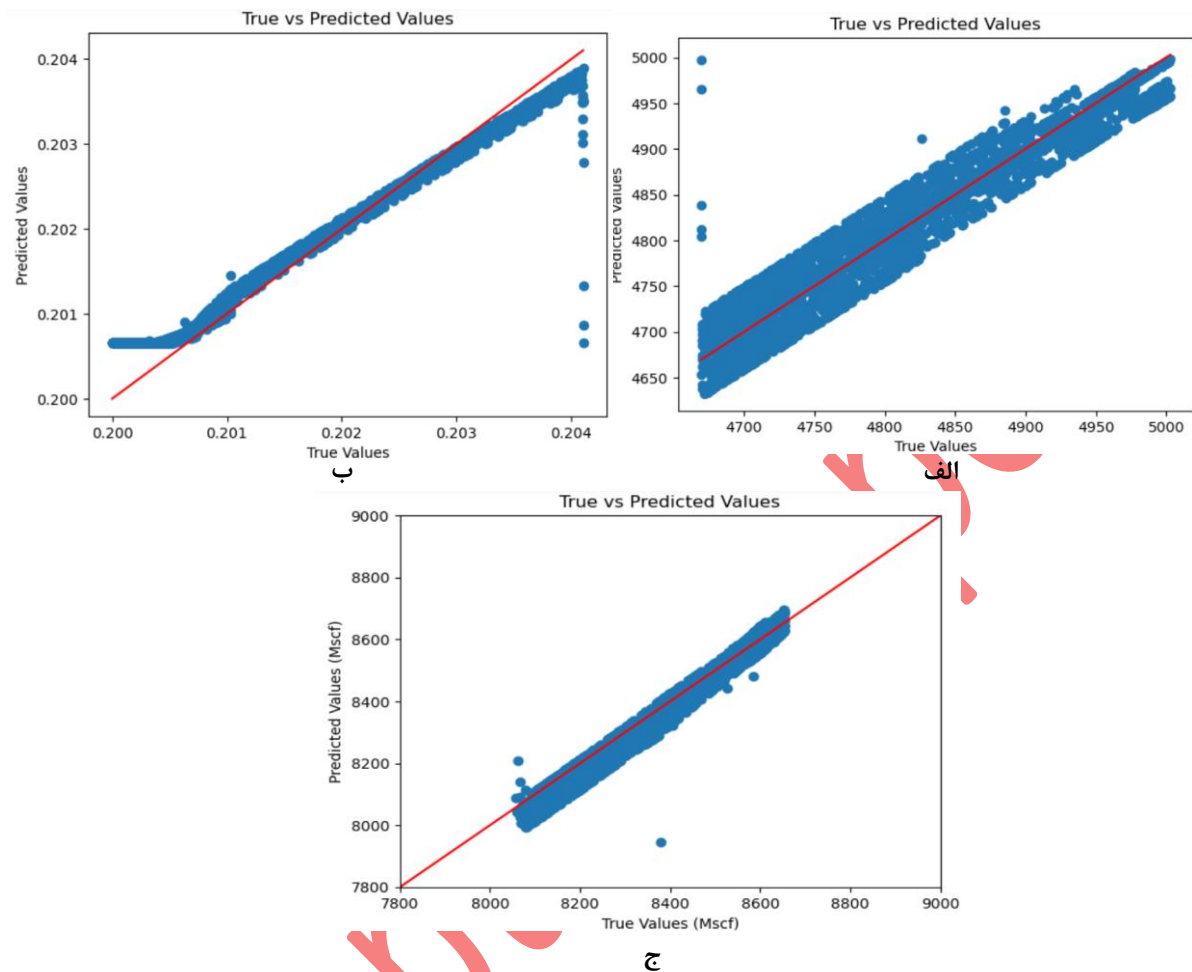
^۱ Dropout

در شکل ۹ روند یادگیری مدل‌های مختلف بر اساس 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.



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

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 باره‌ی داده‌ها به خصوص در بخش داده‌های آموزشی کاهش یافته است که همین روند به صورت کلی در داده‌های اعتبارسنجی نیز دیده می‌شود که به علت جلوگیری از بیش‌برازش، از دیدن بسیار زیاد داده‌ها توسط مدل جلوگیری شده است. در شکل ۱۰ مقادیر پیش‌بینی‌شده توسط مدل‌های داده‌محور بر اساس مقادیر محاسبه‌شده توسط مدل عددی رسم شده است که نشان‌دهنده‌ی 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 در صورتی که از مدل‌های داده‌محور برای پیش‌بینی استفاده شود نتایج مناسبی را به عنوان خروجی محاسبه می‌کند. 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 در حالی که ما از اطلاعات روزانه استفاده کرده‌ایم و همچنین در سال‌های ابتدایی توسعه میدان جمع‌آوری شده است.



شکل ۱۰. رسم مقادیر پیش‌بینی شده بر حسب مقادیر واقعی در مدل‌های داده‌محور مختلف: الف) فشار متوسط میدان، ب) درصد اشباع میانگین آب و ج) تولید گاز

همانطور که از شکل ۱۱ برمی‌آید مدل‌های داده‌محور توانسته‌اند روند تغییرات متغیرهای مدنظر را به خوبی درک کنند. همچنین تغییرات شیب مشاهده شده در نمودارهای مربوط به نتایج شبیه‌ساز عددی به علت حفر چاه‌های جدید توسط مدل‌های داده‌محور فهمیده شده است و در نمودارها قابل مشاهده است. در جدول ۵ نتایج بهینه‌سازی برای پارامترهای مدنظر به منظور آموزش مدل‌ها گنجانده شده است و تلاش شده است که با کوچکترین شبکه‌ی عصبی ممکن و کمترین حجم محاسبات به نتایج مناسبی رسید. در جدول ۶ نیز نتایج معیار ضریب تعیین^۱ و سایر معیارهای اندازه‌گیری دقت برای مدل‌های داده‌محور آموزش دیده بر روی داده‌های دسته‌ی 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 به خوبی توانسته‌اند فیزیک حاکم بر جریان سیال در محیط متخلخل که توسط شبیه‌ساز عددی در دسترس آنها قرار گرفته است به خوبی درک کنند.

^۱ R2 Score

جدول ۵. نتایج پارامترهای بهینه شده در مدل های داده محور مذکور برای توسعه مدل بالا به پایین

مدل	ورودی ها	نرخ یادگیری ^۱	شمار گره های استفاده شده در لایه پنهان	دفعات مشاهده داده ها ^۲	اندازه دسته ^۳	حذف موقت ^۴
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 اندازه گیری دقت برای مدل های داده محور به منظور توسعه مدل بالا به پایین برای داده های آزمون

مدل	ضریب تعیین	میانگین مربع خطا	میانگین مطلق خطا
فشار متوسط میدان	۰/۹۰۷	۱۲۹۴۹۲۶۳۵۱/۴ مربع پاسکال	۱۴۶۸۵۱/۴۵ پاسکال
درصد اشباع آب متوسط میدان	۰/۹۵۹	۰/۰۰۰۲۴۵۶۷۹۵۵	۰/۰۰۰۱۷۴۸۷۵۴۳
تولید گاز	۰/۹۹۷	۵۴۵۲/۷ مترمکعب	۱۱۴۱/۶۸ مترمکعب

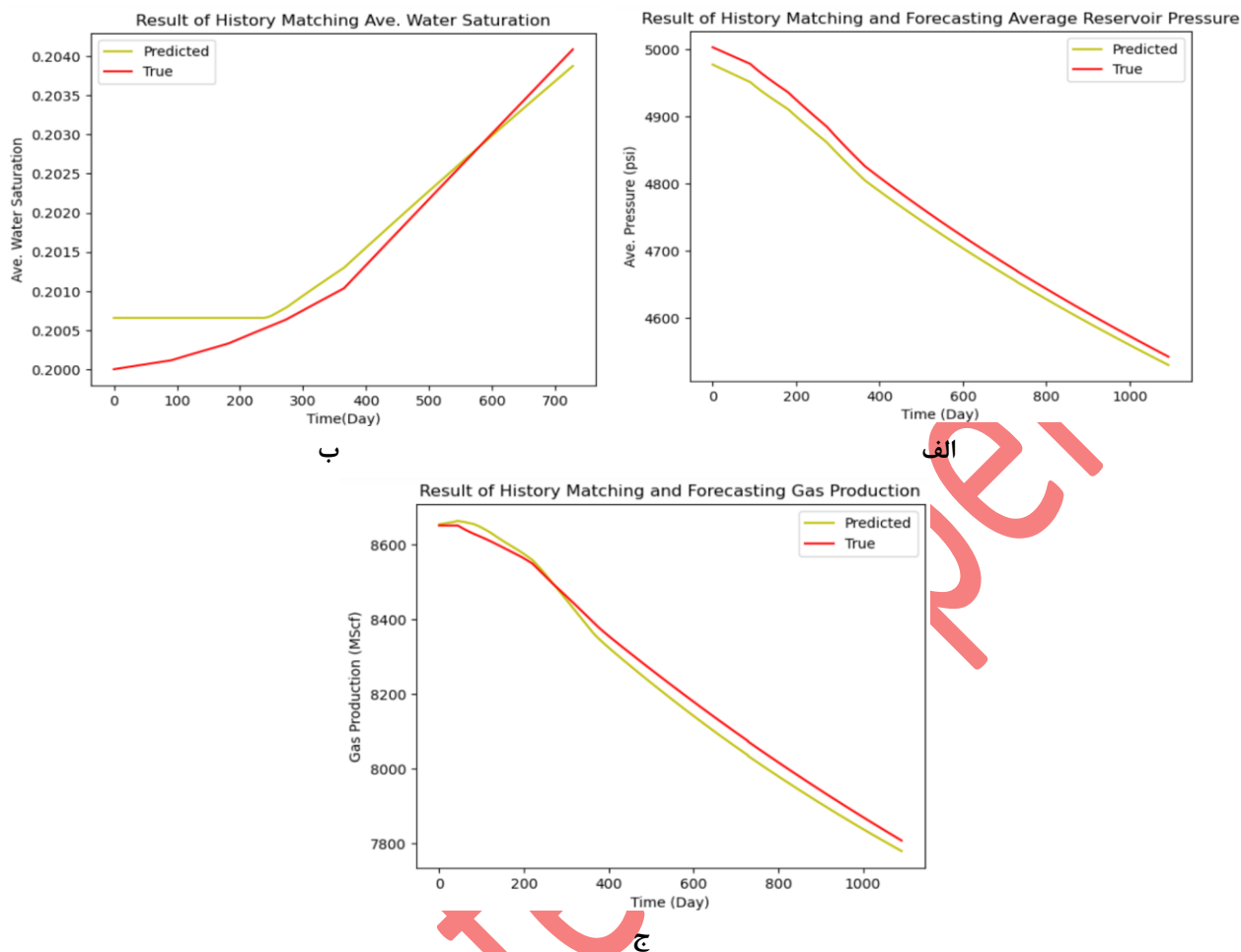
اکنون که مدل ها اعتبارسنجی شدند و از عملکرد آنها در فرایند تطبیق تاریخچه اطمینان حاصل شده از آنها برای پیش بینی از سال ۲۰۲۶ تا ۲۰۲۷ بر اساس اطلاعات چاه شماره ۱ استفاده شده است. شکل ۱۱ نتایج تطبیق تاریخچه و پیش بینی (دو سال تطبیق تاریخچه و یک سال پیش بینی) را برای فشار متوسط میدان، درصد اشباع میانگین آب میدان و تولید گاز چاه شماره ۱ را نشان می دهند که همانطور که از شکل ها مشخص می باشد تطابق مناسبی بین نتایج شبیه ساز عددی و شبیه ساز داده محور وجود دارد چه برای اطلاعات دو سال ابتدایی که به منظور تطبیق تاریخچه استفاده شده است و چه برای سال سوم که مدل های داده محور اطلاعات سال سوم را مشاهده نکرده اند و برای آن پیش بینی انجام داده اند.

^۱ Learning rate

^۲ Epoch

^۳ Batch size

^۴ Dropout



شکل ۱۲. 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

همچنین در جدول ۷ نیز نتایج معیارهای دقت مدل‌های داده‌محور آموزش دیده بر اساس اطلاعات سه سال، مورد بررسی قرار گرفته است. همانطور که از نتایج این معیارها مشخص است مدل‌های داده‌محور به خوبی توانسته‌اند رفتار مخزن در رابطه با مدل فشار متوسط مخزن، مدل درصد اشباع میانگین آب مخزن و مدل تولید گاز چاه به خوبی چه در مرحله‌ی تطبیق تاریخچه و چه در مرحله‌ی 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	۰/۹۸۰۲	۹۳۶۰۱۶۸۸۱/۵۵	۱۳۲۰۳۴/۶۰	۰/۰۰۴۰	۱۸۴۲۹۴۰۷۰۴۱/۱۶	۱۸۶۳۴۰/۲۴ پاسکال
درصد اشباع میانگین آب میدان	۰/۹۷	۰/۰۰۰۳۹	۰/۰۰۰۳	۰/۰۰۱۶۱	۰/۰۰۰۰۰۰۱۵۳۹	۰/۰۰۰۸۰۳۴
تولید گاز چاه	۰/۹۸۷	۲۴۲۶۸/۳۰۶	۷۹۸/۹۴	۰/۰۰۳۵	۷۳۴۴۸۸/۸۶	۱۰۹۱/۸۴ متر مکعب

۴- 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 تاریخی‌چهی خوبی داشته باشد نتواند به علت 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 به ترتیب ۰/۹۸۰۲، ۰/۹۷ و ۰/۹۸۷ در مرحله‌ی تطبیق تاریخی‌چه دیده شد که نتایج قابل قبولی هستند. علاوه بر این مدل‌ها عملکرد مناسبی هم در زمینه‌ی پیش‌بینی سال سوم از خود نشان دادند. بنابراین در این تحقیق شبیه‌سازی داده‌محور به عنوان 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 به سمت میادین هوشمند پیش روند و با گذشت زمان به میادین بالغ‌تری نیز تبدیل گردند، حجم بسیار زیادی از داده‌ها تولید می‌گردد که شبیه‌سازی‌های عددی توانایی استفاده از تمامی این داده‌ها را ندارند پس راه‌حلی که می‌تواند به مهندسين مخزن در این امر کمک کند استفاده از شبیه‌سازی داده‌محور 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 سناریوی مختلف را در زمان کمی اجرا و با یکدیگر مقایسه کرد و همچنین به دلیل سرعت بالای اجرای این مدل‌ها می‌توان برای بررسی عدم قطعیت‌ها و بهینه‌سازی مورد استفاده قرار گیرند. همچنین در این مقاله محدودیت‌هایی وجود داشت که از آنها می‌توان به مواردی همچون موارد زیر اشاره نمود:

- عدم دسترسی به اطلاعات واقعی مخازن موجود در ایران که به علت این محدودیت مطالعه بر روی یک مدل میدانی فرضی انجام شده است و تمامی خواص سنگ و سیالات مورد استفاده به صورت پیش‌فرض از نرم افزار پترل مورد استفاده قرار گرفته‌اند.
- به علت عدم دسترسی به اطلاعات پتروفیزیکی چاه‌های میدانی موجود در ایران، خواص ثابت همچون تخلخل و تراوایی میدان مدل‌سازی شده در این تحقیق توسط توابع توزیع پخش شده‌اند که در واقعیت با استفاده اطلاعات نمونه‌گیری شده از چاه‌های مختلف چنین اطلاعاتی استخراج می‌شود.

منابع و مراجع

- [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] 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*. <https://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*. <https://doi.org/10.1007/978-3-319-96424-9>
- [7] Donald J. Ford, P. (2011). Training Industry. Retrieved from How the Brain Learns: <https://trainingindustry.com/articles/content-development/how-the-brain-learns/>
- [8] Shahkarami, A., Mohaghegh, S. D., Gholami, V., & Haghighat, S. A. (2014). Artificial intelligence (AI) assisted history matching. *All Days*. <https://doi.org/10.2118/169507-ms>
- [9] Haykin, S. S. (2010). *Neural networks and learning machines*. <http://ci.nii.ac.jp/ncid/BB00465945>
- [10] Hernandez, A. (2016). Model Calibration with Neural Networks. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2812140>
- [11] Mohaghegh, S. D. (2017b). Shale Analytics. In *Springer eBooks*. <https://doi.org/10.1007/978-3-319-48753-3>
- [12] Mohaghegh, S. D. (2017). Data-Driven reservoir modeling. In *Society of Petroleum Engineers Richardson, Texas, USA eBooks*. <https://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*. <https://doi.org/10.2118/170660-ms>
- [14] Haifi, A. H. M. A. (2019). CONFIRMATION OF DATA-DRIVEN RESERVOIR MODELING USING NUMERICAL RESERVOIR SIMULATION. <https://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*. <https://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*. <https://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., Alkili, 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*. <https://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. <https://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*. <https://doi.org/10.33915/etd.7571>

Data-driven simulation of a hypothetical oil field and comparison of results with numerical simulation

Mahdi Rashnoo, Saeed Jamshidi

Faculty of Chemical and Petroleum Engineering, Sharif university of Technology, Tehran, Iran

jamshidi@sharif.edu

Abstract

Data-driven reservoir simulation is a novel approach in reservoir modeling that complements or even replaces traditional numerical modeling methods. This method, also known as top-down modeling, utilizes artificial intelligence, specifically deep learning, and due to the nature of these tools, it requires data obtained from field measurements in the oil and gas industry, both for wells and reservoirs. Traditional numerical methods perform the simulation process using numerical modeling and the current understanding of the physics governing fluid flow in porous media.

In this research, efforts have been made to model a hypothetical field using PETREL software, and then use the output data from this software as input for the desired top-down modeling using the Python programming language. After constructing three data-driven models for average reservoir pressure, average water saturation in the reservoir, and gas production from each well, the performance of the constructed models will finally be evaluated using machine learning. The outputs of the data-driven model can include all that is expected in numerical models; however, in this research, only the mentioned outputs have been considered. In other words, the goal is for the data-driven model to understand the physics governing fluid flow in the porous medium through the measured data. During this process, history matching has been performed using the production data from the first two years of the field. The R^2 score for the average reservoir pressure model, the average field water saturation, and gas production has been calculated as 0.9802, 0.97, and 0.987, respectively. A forecast for the third year will also be made and results obtained from the data-driven models will be compared with the results obtained by the numerical simulator.

Keywords: 1- Top-Down modeling 2- Data-driven 3- numerical methods 4- deep learning 5- reservoir