



به کارگیری مدل رگرسیون ماشین بردار پشتیبان به منظور تخمین میزان اشباع شدگی آب سازند یکی از میدان های نفتی بزرگ جنوب غرب ایران

رضا احمدی^{۱*} و محمد صادق امیری بختیار^۲

۱- استادیار، دانشکده مهندسی معدن، دانشگاه صنعتی اراک

۲- دانشجوی کارشناسی ارشد، دانشگاه صنعت نفت آبادان

دریافت مقاله: ۱۳۹۶/۰۱/۱۲؛ پذیرش مقاله: ۱۳۹۶/۰۷/۲۵

* نویسنده مسئول مکاتبات: rezahmadi@gmail.com

واژگان کلیدی

چکیده

اشباع شدگی آب (S_w) سنگ مخزن یکی از پارامترهای پتروفیزیکی مهم است؛ که تأثیر زیادی بر دقت تخمین میزان نفت اولیه مخزن دارد. به دلیل اهمیت زیاد این پارامتر در محاسبات اقتصادی توسعه مخزن، تعیین دقیق آن اجتناب ناپذیر است. در پژوهش حاضر برای تخمین این پارامتر، مدل رگرسیون ماشین بردار پشتیبان شامل ۵ متغیر ورودی یعنی داده های چاهنگاری پرتو گامای طبیعی، تخلخل نوترونی، چگالی کیه ای سازند، زمان گذر امواج صوتی و مقاومت ویژه الکتریکی حقیقی و پارامتر Sw به عنوان تک خروجی برای سه حلقه چاه در یکی از میدان های نفتی بزرگ سازند آسماری واقع در جنوب غرب کشور ایران مورد استفاده قرار گرفته است. به منظور مقایسه نتایج تخمین با واقعیت به طور بصری، ستون چینه شناسی و اشباع شدگی آب و هیدروکربور سازند نیز توسط نرم افزار Geolog برای چاه های مورد مطالعه ترسیم شده است. از تعداد کل ۱۲۱۱ داده نقطه ای موجود برای سه حلقه چاه، حدود ۸۰ درصد به عنوان داده های آموزشی و حدود ۲۰ درصد به عنوان داده های آزمون انتخاب شدند. عملکرد الگوریتم از طریق اعتبارسنجی متقابل بر اساس معیارهای مختلف همانند ترسیم نمودار پراکندگی مقادیر اندازه گیری های آزمایشگاهی S_w توسط مغزه ها در مقابل مقادیر تخمینی با استفاده از داده های چاهنگاری سه حلقه چاه توسط مدل SVR و محاسبه پارامترهای آماری معرف خطا، نیز اعتبارسنجی شده است. نتایج تحقیق نشان می دهد که مدل مذکور از قابلیت بالایی برای تخمین میزان S_w سنگ مخزن با استفاده از داده های چاهنگاری برخوردار است. به گونه ای که داده های آموزشی را با ضریب تعیین همبستگی عالی بیش از ۸۷ درصد و داده های آزمون را با ضریب تعیین همبستگی مطلوب بیش از ۷۶ درصد تخمین زده است.

اشباع شدگی آب (S_w)
رگرسیون ماشین بردار پشتیبان (SVR)
سازند آسماری
داده های چاهنگاری

۱- مقدمه

اشباع‌شدگی آب سازند به کسری از فضاها خالی سنگ گفته می‌شود؛ که توسط آب اشغال شده باشد. این پارامتر یکی از کاربردی‌ترین پارامترهای پتروفیزیکی مخازن نفتی است؛ که درجه موفقیت بسیاری از فعالیت‌های حفاری، تکمیل و تولید چاه‌های نفت و گاز، به دقت روش‌های استفاده شده در اندازه‌گیری آن وابسته است. بنابراین تخمین دقیق آن در اکتشاف، توسعه و بهره‌برداری از مخازن نفت و گاز ضروری است. در صنعت نفت پارامترهای مخزن توسط دو روش متداول و کاربردی تحلیل مغزه به عنوان روش مستقیم و ارزیابی داده‌های چاه‌نگاری به عنوان روش غیرمستقیم تعیین می‌شوند. روش تحلیل مغزه که در آزمایشگاه و با روش آزمایش بر روی مغزه صورت می‌گیرد، اطلاعات مفیدی درباره میزان S_w سنگ مخزن ارائه می‌کند. به عنوان نمونه مرادزاده و همکاران (۱۳۹۲) را در یکی از مخازن هیدروکربوری کربناته با استفاده از داده‌های آزمایشگاهی حاصل از آزمایش‌های مغزه محاسبه کردند. محاسبات آن‌ها از طریق تلفیق داده‌های آزمایشگاهی فشار مؤینگی، به روش ترریق جیوه و طبقه‌بندی سنگ‌های کربناته با استفاده از مفهوم واحدهای جریان هیدرولیکی، استفاده از تابع لورت و ارتباط آن با اشباع آب نرمال شده، انجام شد.

استفاده از روش آزمایشگاهی برای اندازه‌گیری این پارامتر، گران قیمت و زمان‌بر است. بنابراین استفاده از این روش تنها در تعداد محدودی از چاه‌های یک میدان، امکان‌پذیر است. از این رو ارائه روشی توانا به منظور تخمین این پارامتر در میدان‌های نفتی، با استفاده از داده‌های چاه‌نگاری به‌ویژه در چاه‌های فاقد مغزه، ضروری است.

Archie (1942) با بررسی تغییرات مقاومت ویژه واقعی سازند (R_t : True resistivity) و آب سازندی (R_w : Water resistivity) رابطه اساسی زیر را برای محاسبه S_w در مخازن هیدروکربوری ارائه کرد:

$$S_w = \sqrt[n]{\frac{a}{\phi^m} \times \frac{R_w}{R_t}} \quad (1)$$

که در این رابطه n توان اشباعی، a یک عدد ثابت و تابع جنس سنگ (لیتولوژی)، ϕ تخلخل و m ضریب سیمان‌شدگی سازند است. معمولاً در مخازن هیدروکربوری، S_w با استفاده از داده‌های چاه‌نگاری و به کارگیری این رابطه، تخمین زده می‌شود؛ اما رابطه (۱) فقط در مورد سازندهای ماسه‌سنگی تمیز (فاقد رس) برقرار است.

بررسی‌های انجام شده نشان می‌دهد که با حضور درصد کمی از کانی‌های رسی در سازند، به دلیل کاهش مقاومت ویژه سازند، درصد اشباع آب به کمک فرمول آرچی، بیش از مقدار واقعی به دست می‌آید. بنابراین به کارگیری رابطه (۱) در این گونه موارد با خطای زیادی همراه بود (Luthi, 1941). پارامترهای n ، a و m ثابت‌های مجهول و ناشناخته‌ای هستند، که مقادیر آن‌ها برای

سازندهای مختلف، با بررسی‌های پتروفیزیکی و تجزیه و تحلیل مغزه در آن سازند به دست می‌آید. به همین دلیل پژوهشگران به روش‌های مختلف سعی کرده‌اند تا با در نظر گرفتن مجموعه‌ای از فرضیات و لحاظ کردن بعضی عوامل تأثیرگذار، روابطی را برای محاسبه اشباع آب در سازندهای رس‌دار ارائه کنند. تاکنون مدل‌های تجربی زیادی برای اندازه‌گیری S_w سنگ مخزن از داده‌های چاه‌نگاری ارائه شده است؛ که از بزرگ‌ترین معایب این مدل‌ها وابستگی شدید آن‌ها به سازند خاص مورد مطالعه و عدم جامعیت کاربرد آن‌ها برای انواع مختلف سازندهای دیگر است.

از طرفی پیش‌بینی میزان S_w در مخازن کربناته به دلیل ناهمگنی پارامترهای پتروفیزیکی، تغییرات گسترده اندازه حفره‌های موجود و پیچیده و مبهم بودن ارتباط ضرایب آرچی با کانی‌شناسی و حفره‌های موجود در مخزن، امری دشوار و پیچیده بوده و همین مسئله باعث می‌شود تا مقادیر S_w در سنگ‌های کربناته در گستره وسیعی از توان اشباع دچار نوسان شود.

علاوه بر روش‌های تجربی، از روش‌های رگرسیون خطی متعددی نیز برای تخمین میزان S_w سنگ مخزن، از داده‌های چاه‌نگاری استفاده شده است که به دلیل پیچیدگی ویژگی‌های این پارامتر، روابط این مدل‌ها نیز قادر به پیش‌بینی صحیح آن نمی‌باشند. در مقاله حاضر از روش هوشمند مبتنی بر آمار رگرسیون ماشین بردار پشتیبان (SVR: Support Vector Regression) برای تخمین S_w سنگ مخزن یکی از میدان‌های نفتی بزرگ واقع در جنوب غربی کشور از طریق داده‌های چاه‌نگاری مناسب استفاده شده است.

۲- مرور پیشینه تحقیق

اخیراً روش‌های هوشمند توسط بسیاری از محققین در زمینه‌های مختلف صنعت نفت به‌ویژه جهت تخمین پارامترهای مختلف پتروفیزیکی مخزن استفاده شده است. به عنوان نمونه علی‌مرادی (۱۳۹۰) از داده‌های لرزه‌نگاری سه‌بعدی به منظور بهبود روش‌های تخمین پارامتر اشباع‌شدگی از هیدروکربور در مخازن تیپ کربناته استفاده نمود. برای این منظور ابتدا با استفاده از روش بهینه‌سازی الگوریتم ژنتیکی، معادله سرعت گسمن بهبود یافت. سپس از دو نشانگر فاز و فرکانس پوش وزنی برای مدل‌سازی معکوس دو پارامتر تخلخل و ابعاد حفرات با دو روش هوشمند شبکه‌های عصبی مصنوعی و ماشین بردار پشتیبان استفاده نمود؛ که نتایج بیانگر عملکرد بهتر روش ماشین بردار پشتیبان در تخمین مقادیر این پارامترها بود. در نهایت از سه پارامتر تخلخل، بعد حفرات و سرعت موج به منظور پیش‌بینی مقادیر اشباع‌شدگی از آب سازند استفاده کرد.

Cheng et al. (1993) از روش شبکه عصبی مصنوعی برای تخمین تخلخل در یکی از میدان‌های نفتی کانادا بر اساس داده‌های

رشته کوه زاگرس واقع شده؛ که در ناحیه فروافتادگی دزفول دارای بیشترین گسترش است (مطیعی، ۱۳۸۲). این میدان به صورت طاق‌دیزی ملایم با روند شمال غربی - جنوب شرقی و در امتداد روند عمومی چین‌خوردگی‌های زاگرس قرار دارد. میدان مورد مطالعه فاقد عوارض و شواهد سطحی چین‌خورده است؛ که این مسئله با استفاده از داده‌های لرزه‌نگاری نیز به اثبات رسیده است. طول این میدان بر روی آخرین منحنی تراز بسته حدود ۶۳ کیلومتر و عرض آن حدود ۵ کیلومتر است.

مخزن آسماری در محدوده میدان‌های نفتی مناطق نفت‌خیز جنوب همچون اهواز، منصوری و مارون، از جمله مخازن کربناته ترشیری ایران محسوب می‌شود. سازند آسماری در میدان نفتی مورد مطالعه شامل سازند آسماری بالایی از جنس دولومیت، آهک دولومیتی به همراه تکه‌هایی از انیدرید که در زیر بخش یک سازند گچساران (پوش سنگ) قرار دارد، آسماری میانی شامل تناوبی از ماسه‌سنگ، آهک، آهک ماسه‌ای، دولومیت و ماسه شیلی و آسماری زیرین از جنس آهک، آهک دولومیتی و ماسه‌سنگ است. ماسه‌سنگ‌های این زیرلایه به صورت غیرسیمانی و یا سیمان شده هستند؛ که سیمان آن‌ها از جنس آهک، دولومیت و گاهی نیز انیدرید است. ضخامت این سازند در میدان مطالعه شده بین ۲۹۰ تا ۳۵۰ متر است. متراژ حفاری سه حلقه چاه A، B و C در سازند آسماری میدان نفتی مورد مطالعه به ترتیب، ۳۵۰، ۳۸۵ و ۴۷۵ متر است؛ که به‌طور کلی از توالی سنگ‌های کربناته (سنگ آهک و دولومیت)، ماسه‌سنگ و شیل تشکیل شده است. در این میدان نفتی فرآیند دولومیتی شدن در کل مخزن رخ داده و بیشترین درصد دولومیت در بخش بالایی مخزن مشاهده می‌شود (مطیعی، ۱۳۸۲؛ محمدی، ۱۳۸۳).

تخلخل متوسط مناطق نفتی دولومیتی و ماسه‌سنگی به ترتیب ۱۳ و ۱۶ درصد و نفوذپذیری این مناطق به ترتیب ۱۰ و ۶۰۰ میلی‌داریسی است (اداره مطالعات زمین‌شناسی شرکت ملی مناطق نفت‌خیز جنوب، ۱۳۸۷).

۴- معرفی روش استفاده شده (SVR)

همان‌گونه که از رابطه (۱) ملاحظه می‌شود، نگارهای مقاومت ویژه و تخلخل از مهم‌ترین نگارهای مورد استفاده در تعیین اشباع‌شدگی آب در رابطه اساسی تجربی آرچی هستند. تخلخل سازند را می‌توان با دقت قابل قبولی از روی داده‌های چاه‌نگاری صوتی، چگالی و نوترون به دست آورد؛ اما نگارهای مقاومت ویژه به شدت به حضور شیل و ناخالصی‌های رسی سازندها حساس هستند؛ که با استفاده از نمودار GR می‌توان تأثیر میزان رس سازند را نیز اعمال نمود. بنابراین به‌کارگیری یک روش هوشمند با استفاده از داده‌های چاه‌نگاری مناسب، برای تخمین S_w سازند سودمند خواهد بود. در پژوهش حاضر برای این منظور روش هوشمند SVR مورد استفاده قرار گرفته

نگارهای صوتی، چگالی نوترون و پرتو گاما استفاده کردند. در این پژوهش نتایج تخمین توسط شبکه‌های عصبی، با تخمین مبتنی بر روش رگرسیون خطی، مقایسه شد و بر اساس نتایج حاصل به وسیله شبکه عصبی تخمین بهتری از تخلخل صورت گرفت. Cuddy (2000) با استفاده از منطق فازی و بر اساس قوانین تجربی، ستون چینه‌شناسی سنگ مخزن میدان گازی وایکینگ (Viking) واقع در قسمت جنوبی دریای شمال را ترسیم نمود. با مشخص بودن ستون چینه‌شناسی، مقادیر نفوذپذیری سنگ مخزن را نیز بر اساس قوانین تجربی تخمین زد. Singh (2005) در پژوهشی که در میدان یوتا (Utah) در خلیج یواینتا (Uintah) انجام داد، شبکه‌های عصبی مصنوعی را برای تخمین نفوذپذیری با استفاده از داده‌های نگارهای پرتو گاما، نوترون و چگالی بکار گرفت. Lim et al. (2006) در یکی از میدان‌های نفتی کره از تلفیق دو روش شبکه عصبی و الگوریتم ژنتیک برای تخمین نفوذپذیری سنگ مخزن در یکی از چاه‌های این میدان نفتی بر اساس نگارهای چاه و با هدف تخمین تخلخل بر اساس توابع چندجمله‌ای استفاده کردند. در این پژوهش هدف، بهینه‌سازی ضرایب چندجمله‌ای تخمین بود؛ که بر اساس روش تلفیقی مذکور، نتایج بهتری نسبت به زمانی که از شبکه‌های عصبی مصنوعی به‌تنهایی استفاده می‌شد، به دست آمد. Aminzadeh and Brouwer (2006) از شبکه‌های عصبی - فازی برای بهبود پیش‌بینی خصوصیات پتروفیزیکی مخزن در یکی از میدان‌های نفتی آمریکا استفاده کردند. Adeniran et al. (2009) پارامتر پتروفیزیکی تخلخل را با استفاده از داده‌های شش نمودار و پارامتر S_w سازند را با داده‌های پنج نمودار چاه‌نگاری با استفاده از روش شبکه‌های تابعی (حالت تعمیم یافته‌ای از شبکه‌های عصبی با تغذیه پیشرو) تخمین زدند. Jafari Kenari and Mashohor (2013) از ترکیب روش‌های مختلف یادگیری ماشین (Machine learning) برای تخمین S_w مخزن از روی داده‌های چاه‌نگاری استفاده کردند. Bagheripour and Asoodeh (2014) از روش منطق فازی بهبودیافته (ترکیب ژنتیک الگوریتم با جستجوی الگو جهت تنظیم قواعد فازی و استخراج پارامترهای بهینه دخیل در ساختار محاسباتی مدل فازی) برای تخمین S_w سازند بر اساس داده‌های چاه‌نگاری استفاده نمودند.

هر چند تحقیقات مبتنی بر روش‌های هوشمند قبلی در خصوص تخمین پارامترهای مخزن، موفقیت‌آمیز بوده‌اند؛ برای تخمین پارامتر مهمی چون S_w مخزن، همچنان ارائه مدل‌های توانا تر به‌ویژه روش‌های هوشمند مبتنی بر آمار پیشرفته با استفاده از داده‌های چاه‌نگاری، مورد نیاز است.

۳- زمین‌شناسی و خصوصیات مخزنی منطقه مورد مطالعه

میدان نفتی مورد مطالعه درون سازند آسماری در جنوب غربی

حال بیشینه شدن اندازه حاشیه جدایش است. تابع L_ε تابع وینیک است؛ که به صورت زیر تعریف می‌شود:

$$|y - f(x, w)|_\varepsilon = \begin{cases} 0 & \text{if } |y - f(x, w)| \leq \varepsilon \\ |y - f(x, w)| - \varepsilon & \text{Otherwise} \end{cases} \quad (4)$$

مسئله بالا به صورت بیشینه شدن معادله زیر بازنویسی می‌شود:

$$L_p(\alpha_i, \alpha_i^*) = -\frac{1}{2} \sum_{i,j=1}^l (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) x_i^T x_j - \varepsilon \sum_{i=1}^l (\alpha_i + \alpha_i^*) + \sum_{i=1}^l (\alpha_i - \alpha_i^*) y_i \quad (5)$$

که شرایط آن به صورت زیر است:

$$\begin{cases} \sum_{i=1}^l (\alpha_i - \alpha_i^*) = 0 \\ 0 \leq \alpha_i \leq C, \quad i = 1, \dots, l \\ 0 \leq \alpha_i^* \leq C, \quad i = 1, \dots, l \end{cases} \quad (6)$$

با حل معادله بالا می‌توان تابع SVR یعنی f را با استفاده از تابع کرنل به صورت زیر محاسبه نمود:

$$f(x, w) = w_0^T x + b = \sum_{i=1}^l (\alpha_i - \alpha_i^*) x_i^T x + b \quad (7)$$

۴-۲- جمع‌آوری، آماده‌سازی و پردازش داده‌های مورد نیاز

در این پژوهش از داده‌های چاه‌نگاری مناسب شامل GR، تخلخل نوترونی، چگالی کپ‌های سازند، زمان گذر امواج صوتی و مقاومت ویژه الکتریکی حقیقی به دست آمده از سوند القایی عمیق برداشت شده درون سه حلقه چاه، در محدوده مورد مطالعه استفاده شده است. مقادیر S_w اندازه‌گیری شده در آزمایشگاه از طریق مغزه نیز برای تمام طول این سه حلقه چاه، در دسترس می‌باشند.

از تعداد کل ۱۲۱۱ داده نقطه‌ای (شامل داده‌های ۵ سوند چاه‌نگاری و مقدار S_w اندازه‌گیری شده از طریق مغزه)، موجود برای سه حلقه چاه، حدود ۸۰ درصد (۹۸۸ عدد) به عنوان داده‌های آموزشی و حدود ۲۰ درصد (۲۲۳ عدد) به عنوان داده‌های آزمون (Test) انتخاب شدند.

از آنجایی که لیتولوژی تأثیر بسزایی بر روی پارامترهای پتروفیزیکی مخازن دارد. بنابراین در پژوهش حاضر به منظور ارائه تصویری از وضعیت واقعی ساختار مخزن، امکان تجسم و درک ملموس‌تری از آن و مقایسه نتایج تخمین مدل‌سازی با واقعیت به‌طور بصری، با استفاده از نرم‌افزار Geolog مطالعه لیتولوژی سازند محدوده مخزن نفتی مورد نظر بر اساس مطالعه مغزه‌ها و تفسیر همزمان پنج نگار چاه، برداشت شده در سه حلقه چاه میدان نفتی مورد مطالعه نیز صورت گرفت و ستون چینه‌شناسی هر سه حلقه چاه ترسیم گردید. به عنوان نمونه در شکل ۲ ستون چینه‌شناسی همراه با جزئیات پتروفیزیکی سازند آسماری در حلقه چاه A (با شماره ۰۶۸) ترسیم شده است. در این شکل از چپ به راست به

۴-۱- مبانی نظری روش SVR

الگوریتم SVR یکی از الگوریتم‌های یادگیری ماشین است؛ که بر اساس تئوری آموزش آماری استوار است. این روش که جزو روش‌های آموزش نظارت‌شده (Supervised) است؛ بین داده‌های ورودی و مقدار پارامتر وابسته، بر اساس کمینه‌سازی ریسک ساختاری ارتباط برقرار می‌کند. کمینه‌سازی ریسک ساختاری، نسبت به روش متداول کمینه‌سازی ریسک تجربی مورد استفاده در الگوریتم‌هایی مانند شبکه‌های عصبی و روش‌های کلاسیک آماری، دارای برتری است و برخلاف روش‌هایی همانند شبکه‌های عصبی، به جواب‌های موضعی همگرا نمی‌شود. SVR روشی برای تخمین یک تابع است که بر اساس داده‌های آموزشی از یک شیء ورودی به یک عدد واقعی نگاشت می‌شود. در مسائل رگرسیون، بردارهای ورودی به یک فضای چندبعدی، نگاشت می‌شوند؛ سپس یک ابرصفحه ساخته می‌شود که بردارهای ورودی را با بیشترین فاصله ممکن، از یکدیگر جدا می‌کند. برای رفع مشکل انجام عملیات در فضای با ابعاد زیاد، از یک تابع کرنل استفاده می‌شود. در این صورت عملیات می‌تواند با همان سرعت فضای داده‌های ورودی انجام شود. در واقع با استفاده از تابع کرنل، مشکل چندبعدی و غیرخطی بودن نگاشت، مرتفع می‌شود. فرآیند بهینه‌سازی باید با یک تابع افت اصلاح شده همراه شود؛ تا سنجش فاصله را دربر گیرد. در حقیقت هدف SVR، تخمین پارامترهای وزن‌ها و ارباب تابعی است که بهترین انطباق را بر داده‌ها داشته باشد (Yu and Kim, 2012). تابع SVR می‌تواند خطی (شکل ۱-الف) یا غیرخطی باشد (شکل ۱-ب) و در حالت غیرخطی، محاسبه تابع رگرسیون در یک فضای ویژگی با ابعاد بالا صورت می‌گیرد؛ که داده‌های ورودی از طریق یک تابع غیرخطی، نگاشت می‌شوند.

با فرض آنکه l داده آموزشی وجود داشته باشد، اگر هر ورودی X دارای تعداد D ویژگی بوده (به عبارت دیگر متعلق به فضایی با بعد D باشد) و هر نقطه دارای مقدار Y نظیر ویژگی‌های باشد -مانند تمام روش‌های رگرسیون- هدف پیدا کردن تابعی است که بین ورودی و خروجی رابطه برقرار کند.

$$f(x, w) = w^T x + b \quad (2)$$

برای به دست آوردن تابع f محاسبه مقادیر w و b ضروری است. برای محاسبه مقادیر w و b رابطه آتی باید کمینه شود (Vapnik, 1995).

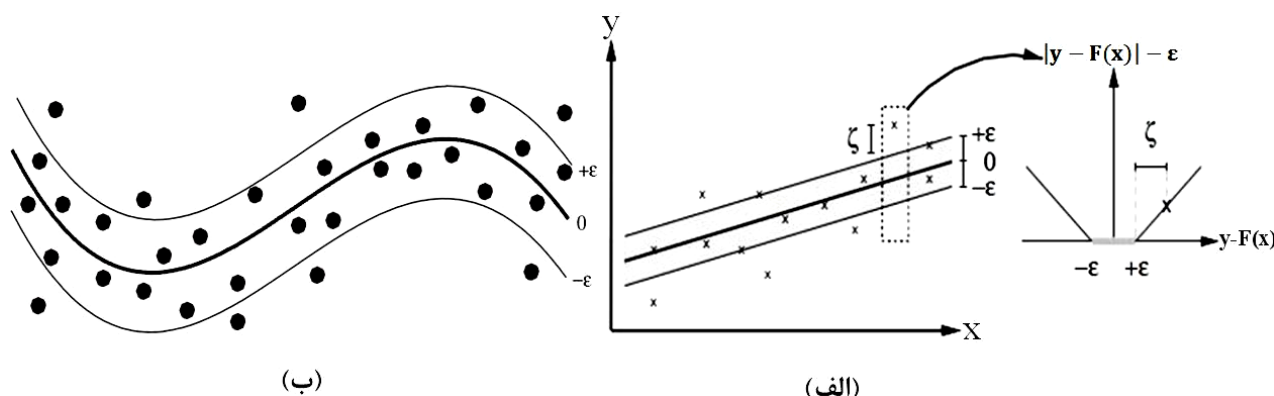
$$R(C) = \frac{1}{2} \|w\|^2 + C \frac{1}{l} \sum_{i=1}^l L_\varepsilon(y_i, f_i(x, w)) \quad (3)$$

که در آن C پارامتری ثابت بوده و مقدار آن باید به وسیله کاربر مشخص شود. در واقع وظیفه پارامتر ثابت C ایجاد تعادل و تغییر وزن‌های میزان جریمه ناشی از اغماض (متغیر ε) و در عین

راست و حجم اشغال شده تخلخل مفید با آب (VUWA) با رنگ طوسی روشن (و افزایش از چپ به راست) نمایش داده شده است. بدیهی است بقیه حجم اشغال شده تخلخل مؤثر (رنگ سبز) به هیدروکربور اختصاص می‌یابد.

مقایسه و انطباق نتایج تخمین میزان اشباع‌شدگی سازند در مرحله بعدی، با اشکال نمایشی ستون چینه‌شناسی و ستون اشباع‌شدگی از آب و هیدروکربور سازند ترسیم شده برای سه حلقه چاه، هم از طریق مشاهده چشمی و هم از نظر دقت عددی، به ارزیابی سریع عملکرد الگوریتم مورد استفاده کمک می‌کند.

ترتیب ستون اول، نمودار پرتو گاما (GR)، ستون دوم نمودار قطرسنجی چاه (Caliper)، ستون سوم شامل سه نمودار زمان گذر امواج صوتی (DT)، تخلخل نوترونی (NPHI) و چگالی کپه‌ای سازند (RHOB) و ستون چهارم نمودار مقاومت ویژه حقیقی سوند القایی عمیق (ILD) است. در ستون پنجم نیز لیتولوژی سازند شامل انیدرید (VANH)، دولومیت (VDOL)، کلسیت (VLST)، کوارتز (VSST) و شیل (VSH) به صورت درصد حجمی با رنگ و الگوهای نمایشی مربوطه، نشان داده شده است. در ستون ششم درجه اشباع‌شدگی سازند از آب با نمودار آبی و با افزایش از راست به چپ، میزان تخلخل مؤثر (PHIE) با نمودار مشکی و افزایش از چپ به



شکل ۱: الف - نمایش یک تابع SVR خطی (Ivanciuc, 2007)، ب - نمایش تابع SVR غیرخطی که یک تیوب به شعاع ϵ را بر داده‌ها برازش می‌دهد (Smola and Schölkopf, 2003).

مورد استفاده قرار می‌گیرند که تابع کرنل خطی ساده‌ترین نوع آن‌ها است. امروزه تابع پایه شعاعی نسبت به توابع دیگر، اهمیت بیشتری پیدا کرده و مطابق رابطه (۸) اغلب به صورت گوسی بکار می‌رود. هر چند که این تابع به صورت نمایی هم کاربرد دارد. علت استفاده از تابع پایه شعاعی در این پژوهش آن است که این نوع تابع کرنل در موارد زیادی در عمل مورد استفاده قرار گرفته و در موارد استفاده شده به نتایج مطلوبی رسیده است. هرچند توابع کرنل دیگر نیز به صورت سعی و خطا آزمایش شدند؛ ولی بهترین نتیجه مطابق پیش‌بینی، با استفاده از تابع پایه شعاعی به دست آمد. این تابع در حالتی که داده‌ها در بازه -1 و 1 باشند، بهترین عملکرد را از خود نشان می‌دهد. بنابراین ابتدا داده‌ها با استفاده از رابطه (۹) بین این دو مقدار نرمالیزه شدند.

$$y_i = 2 \times \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} - 1 \quad (9)$$

در واقع با این تبدیل، کمترین مقدار داده‌ها به عدد -1 و بیشترین مقدار، به عدد 1 نگاشت می‌شود. مقادیر بهینه پارامترهای به دست آمده نیز در جدول ۱ آمده است. در این جدول ثابت مثبت C (پارامتر تنظیم برازش) یا پارامتر توازن (Trade off) بین اندازه

۳-۴- تخمین اشباع‌شدگی آب سازند با مدل SVR

به منظور به کارگیری روش SVR برای تخمین مقادیر S_w لازم است که مدل با استفاده از داده‌های ورودی و خروجی مناسب در محیط MATLAB آموزش داده شود. در این پژوهش ورودی‌ها شامل ۵ متغیر (داده‌های چاه‌نگاری) و خروجی نیز تک پارامتر S_w است. برای ساخت بهترین مدل SVR باید پارامترهای تأثیرگذار آن بهینه شود. به منظور بهینه‌سازی پارامترهای مدل می‌توان از روش جستجو در فضای استفاده نمود که در این پژوهش از روش جستجو در فضای شبکه‌بندی شده پارامترهای بردار پشتیبان با تابع کرنل پایه شعاعی (RBF: Radial Basis Function) گوسی استفاده شد. این تابع کرنل برای دو متغیر a و b مطابق رابطه (۸) تعریف می‌شود (Yu and Kim, 2012):

$$K(a, b) = \exp(-\gamma \|a - b\|^2) = \exp\left(-\frac{\|a - b\|^2}{2\sigma^2}\right) \quad (8)$$

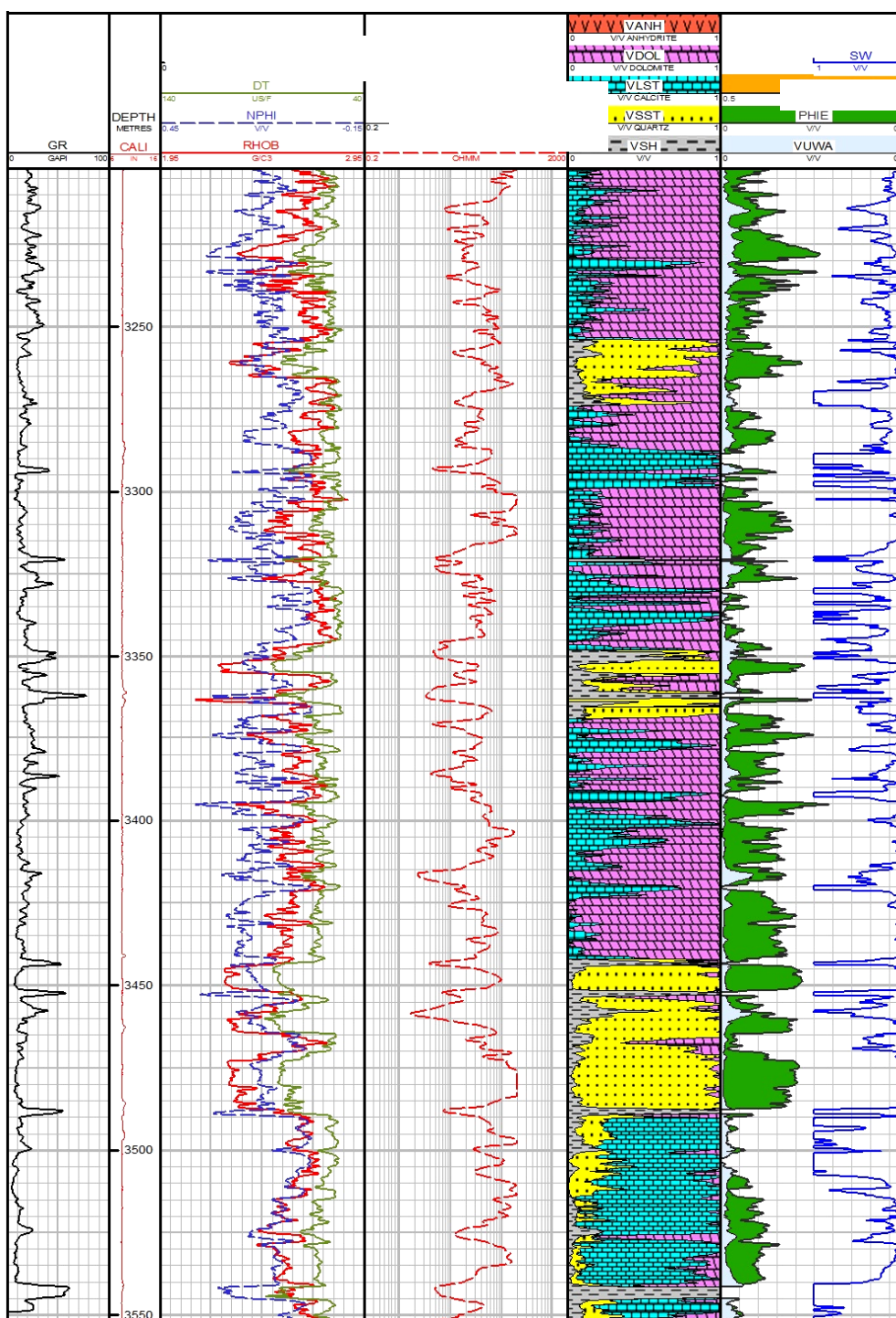
که در این رابطه پارامتر γ با عکس مجذور عرض (پهنای) کرنل گوسی (σ) متناسب است.

به طور کلی در مسائل SVM چهار نوع تابع کرنل خطی، تابع چندجمله‌ای از درجه n ، تابع پایه شعاعی و تابع تانژانت سیگموئید،

احمدی و امیری بختیار، به کارگیری مدل رگرسیون ماشین بردار پشتیبان به منظور تخمین میزان اشباع شکی آب سازند یکی از میدان‌های نفتی ...، صفحات ۲۱۰-۱۹۹.

راستای عمود برهم مشابه سه ضلع. یک مکعب سه بعدی ایجاد کرده و این مکعب، شبکه بندی می شود. سپس با استفاده از سه حلقه For در محیط نرم افزار MATLAB مقدار یک تابع هدف (خطا) کمینه می شود. تابع هدف نیز میانگین قدر مطلق اختلاف مقادیر واقعی پارامتر اندازه گیری (S_w) با مقادیر تخمینی مدل SVR بر اساس مقادیر C ، ϵ و γ در هر گره از این شبکه، تعریف شد. مقدار اولیه پارامتر ϵ معمولاً در ابتدا ۰/۱ انتخاب می شود (Ivanciuc, 2007).

حاشیه (Margin) و میزان خطاها و پارامتر ϵ معرف دقت تخمین یعنی اختلاف بین مقادیر واقعی و مقادیر تخمینی است (Yu and Kim, 2012). هرچه مقدار C بزرگتر انتخاب شود، اندازه حاشیه کوچکتر خواهد بود و برعکس (Hastie et al., 2001; Parrella, 2007). مقادیر بهینه پارامترهای C ، ϵ و γ با استفاده از روش جستجو در یک فضای شبکه بندی شده به دست آمدند. در این روش از طریق تعریف حدود مقادیر بالا و پایین برای این سه پارامتر، در سه



شکل ۲: نمایش ستون چینه شناسی همراه با جزئیات پتروفیزیکی سازند آسماری در حلقه چاه A (با شماره ۰۶۸).

جدول ۱: مقادیر بهینه پارامترهای مدل SVR استفاده شده برای تخمین مقادیر S_w .

پارامتر مدل	ضریب تنظیم برازش (c)	دقت تخمین (ε)	گاما (γ)
مقدار بهینه	۱/۲۰۸	۰/۱۲۳	۲۰/۰۹۴

برای داده‌های آموزش و آزمون در جدول ۲ خلاصه شده است. پارامترهای آماری R^2 ، MSE، MRE و AMRE برای متغیری مانند X به ترتیب مطابق روابط (۱۰) تا (۱۳) تعریف می‌شوند.

$$R^2 = 1 - \frac{\sum_{i=1}^n (X_i - \hat{X}_i)^2}{\sum_{i=1}^n (\bar{X} - \hat{X}_i)^2} \quad (10)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (X_i - \hat{X}_i)^2 \quad (11)$$

$$MRE = \frac{1}{n} \sum_{i=1}^n \left(\frac{X_i - \hat{X}_i}{X_i} \right) \quad (12)$$

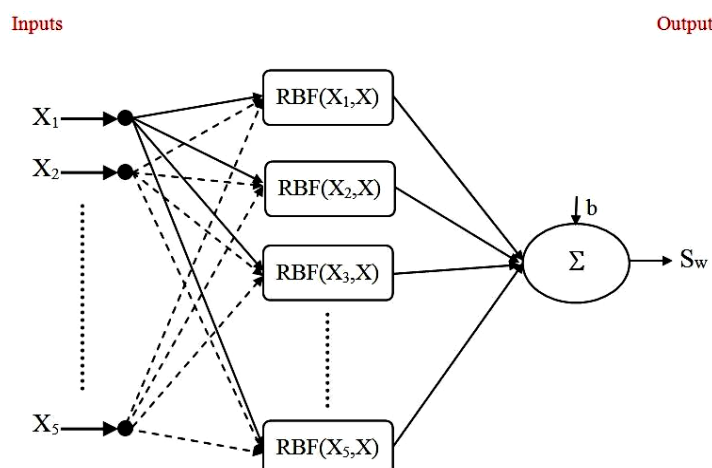
$$AMRE = \frac{1}{n} \sum_{i=1}^n \left| \left(\frac{X_i - \hat{X}_i}{X_i} \right) \right| \quad (13)$$

که در این روابط X متغیر مشاهده‌ای یا اندازه‌گیری شده، $\bar{X}$ میانگین متغیر X و $\hat{X}$ مقدار تخمینی آن متغیر است. با توجه به این روابط ملاحظه می‌شود که پارامتر MSE معرف میزان خطای مطلق است؛ ولی دو پارامتر دیگر (MRE و AMRE) بیانگر میزان خطای نسبی می‌باشند.

با توجه به داده‌های جدول ۲ ملاحظه می‌شود که میزان ضریب تعیین مدل SVR هم برای داده‌های آموزشی و هم برای داده‌های آزمون دارای مقدار بیشتر از ۰/۷ است که مطلوب ارزیابی می‌شود. همچنین مقادیر پارامترهای خطا هم برای داده‌های آموزشی و هم برای داده‌های آزمون، اندک است. بنابراین رویکرد پیشنهادی پارامتر S_w سنگ مخزن را با دقت مطلوبی پیش‌بینی نموده است.

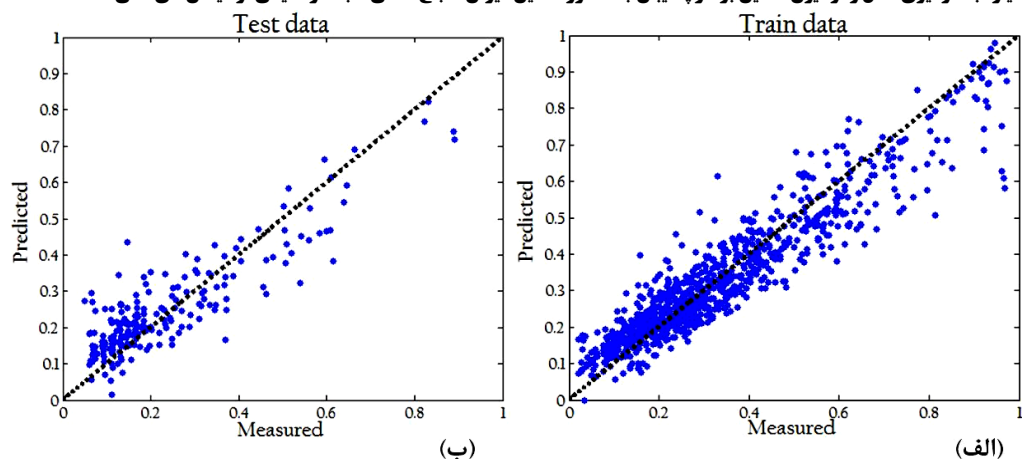
در شکل ۳ شمای کلی از ساختار (معماری) مدل SVR استفاده شده در این مقاله نشان داده شده است. عملکرد الگوریتم از طریق اعتبارسنجی متقابل بر اساس معیارهای مختلف همانند ترسیم نمودار پراکندگی مقادیر اندازه‌گیری‌های آزمایشگاهی S_w از روی مغزه‌ها در مقابل مقادیر تخمینی، با استفاده از داده‌های چاه‌نگاری سه حلقه چاه، توسط مدل SVR و محاسبه پارامترهای آماری معرف خطا، اعتبارسنجی شده است. شکل ۴ نمودار پراکندگی مقادیر اندازه‌گیری شده S_w در مقابل مقادیر تخمینی توسط مدل SVR را برای داده‌های آموزشی (شکل الف) و آزمون (شکل ب) نشان می‌دهد. با توجه به این شکل مشاهده می‌شود که پراکندگی داده‌ها عمدتاً حول خط نیمساز ربع اول است. در شکل‌های ۵ و ۶ نیز مقادیر اندازه‌گیری شده S_w در مقابل مقادیر تخمینی توسط مدل SVR به ترتیب برای داده‌های آموزشی و آزمون، مقایسه شده است. با دقت در این دو شکل و مشاهده چشمی نیز می‌توان میزان انطباق داده‌های مشاهده‌ای با مقادیر تخمینی را مطلوب ارزیابی نمود. بنابراین درمجموع با توجه به این شکل‌ها می‌توان نتیجه‌گیری نمود که مدل SVR از توانایی لازم برای مدل‌سازی S_w سنگ مخزن با استفاده از داده‌های مناسب چاه‌نگاری برخوردار است.

ارزیابی عملکرد رویکرد پیشنهادی (مدل SVR) بر اساس پارامترهای آماری مختلف شامل ضریب تعیین (Coefficient of determination) داده‌های مشاهده‌ای و تخمینی (R^2)، میانگین مربعات خطای تخمین (MSE: Mean Squared Error)، میانگین خطای نسبی (MRE: Mean Relative Error) و قدر مطلق میانگین خطای نسبی (AMRE: Absolute of Mean Relative Error) که همگی معرف میزان دقت (خطای) تخمین می‌باشند،

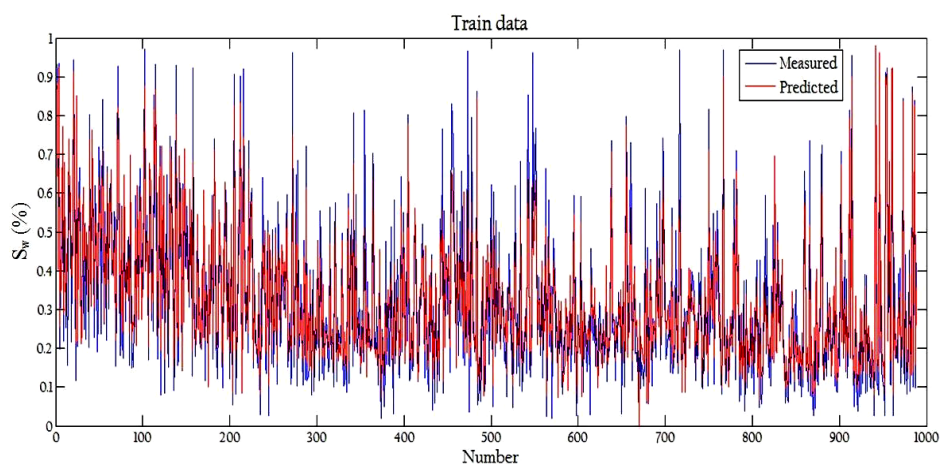


شکل ۳: شمایی از ساختار SVR استفاده شده در پژوهش حاضر با ۵ ورودی.

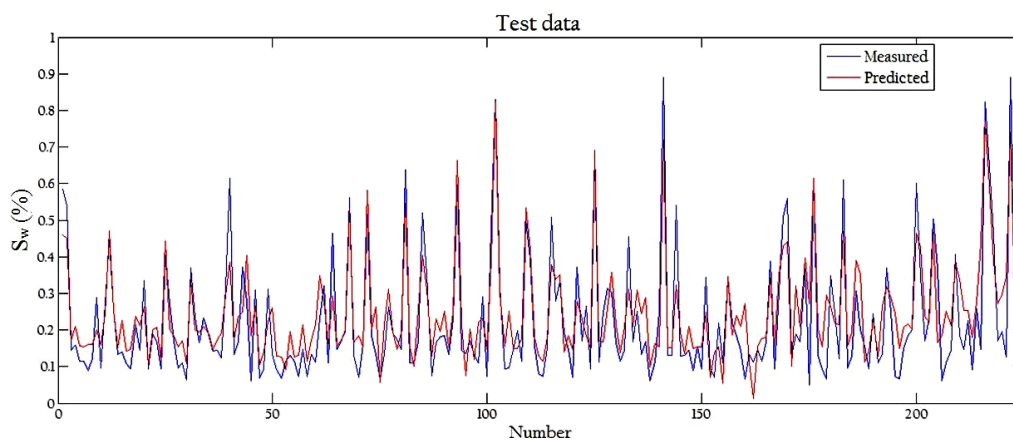
احمدی و امیری بختیار، به کارگیری مدل رگرسیون ماشین بردار پشتیبان به منظور تخمین میزان اشباع‌شدگی آب سازند یکی از میدان‌های نفتی ...، صفحات ۲۱۰-۱۹۹.



شکل ۴: مقادیر اندازه‌گیری شده S_w در مقابل مقادیر تخمینی مدل SVR برای الف- داده‌های آموزشی و ب- داده‌های آزمون.



شکل ۵: مقایسه مقادیر اندازه‌گیری شده S_w با مقادیر تخمینی مدل SVR برای داده‌های آموزشی.



شکل ۶: مقایسه مقادیر اندازه‌گیری شده S_w با مقادیر تخمینی مدل SVR برای داده‌های آزمون.

جدول ۲: پارامترهای آماری مدل SVR استفاده شده برای تخمین مقادیر S_w .

روش	نوع داده	R^2	MSE	MRE	AMRE
SVR	آموزشی	۰/۸۷۳۷	۰/۰۰۵۲	-۰/۱۴۹۲	۰/۲۸۱۱
	آزمون	۰/۷۶۴۱	۰/۰۰۶۵	-۰/۳۰۱۷	۰/۴۳۰۱

در پژوهش حاضر از روش هوشمند مبتنی بر آمار رگرسیون ماشین بردار پشتیبان با استفاده از داده‌های مناسب چاه‌نگاری پرتو گاما، تخلخل نوترونی، چگالی کپه‌ای سازند، زمان گذر امواج صوتی و مقاومت ویژه حقیقی سازند برای تخمین S_w سنگ مخزن یکی از میدان‌های بزرگ نفتی جنوب غربی کشور استفاده شد. نتایج پژوهش بر اساس ترسیم نمودار پراکندگی مقادیر اندازه‌گیری شده S_w در مقابل مقادیر تخمینی توسط مدل SVR و محاسبه پارامترهای آماری مختلف معرف دقت، نشان می‌دهند که مدل SVR پیشنهادی می‌تواند این پارامتر پتروفیزیکی مهم را به روش غیرمستقیم و با دقت بالایی تخمین زند، به‌گونه‌ای که میزان ضریب تعیین همبستگی بین داده‌های اندازه‌گیری شده در آزمایشگاه از طریق مغزه، با داده‌های تخمینی در مرحله آموزشی، بیش از ۸۷ درصد و برای مرحله آزمون، بیش از ۷۶ درصد است. این میزان همبستگی برای داده‌ها به‌ویژه در مرحله آزمون، مطلوب ارزیابی می‌گردد. مقایسه و مطابقت نتایج حاصل از تخمین، با اشکال نمایشی ستون لیتولوژی و ستون اشباع‌شدگی آب و هیدروکربور سازند، همراه با جزئیات پتروفیزیکی سازند آسماری در سه حلقه چاه محدوده مورد مطالعه، ترسیم شده با استفاده از نرم‌افزار Geolog، هم به روش مشاهده چشمی و هم از نظر دقت عددی، اعتبار عملکرد الگوریتم مورد استفاده را تأیید می‌نماید. بنابراین روش پیشنهادی (SVR) یک روش دقیق، سریع و ارزان برای ارزیابی پارامتر پتروفیزیکی مهم S_w سنگ مخزن است.

۶- منابع

اداره مطالعات زمین‌شناسی شرکت ملی مناطق نفت‌خیز جنوب، ۱۳۸۷، گزارش زمین‌شناسی گسترشی (از مجموعه گزارش‌های داخلی مناطق نفت‌خیز جنوب).

علی‌مرادی، ا.، ۱۳۹۰، ارائه روشی برای تخمین اشباع هیدروکربور مخازن کربناته به کمک نشانگرهای لرزه‌ای و داده‌های چاه، رساله دکتری، دانشکده مهندسی معدن، نفت و ژئوفیزیک، دانشگاه صنعتی شاهرود.

محمدی، ا.، ۱۳۸۳، تعیین گروه‌های سنگی مخزنی سازند آسماری بر مبنای لاگ‌های پتروفیزیکی با استفاده از روش منطق فازی، پایان‌نامه کارشناسی ارشد، دانشکده زمین‌شناسی، دانشگاه تهران.

مرادزاده، ع.، سلیمی دلشاد، ی. و کاظم‌زاده، ع. ا.، ۱۳۹۲، بهبود محاسبه اشباع آب در مخازن هیدروکربوری کربناته به کمک داده‌های چاه‌نگاری و آزمایشگاهی، مجله فیزیک زمین و فضا، ۳۹ (۱)، ۱۰۷-۱۲۱.

مطیعی، ه.، ۱۳۸۲، زمین‌شناسی ایران، چینه‌شناسی زاگرس، سازمان

Adeniran, A., Elshafei, M. and Hamada, G., 2009, Functional network soft sensor for formation porosity and water saturation in oil wells, Instrumentation and Measurement Technology Conference, I2MTC '09. IEEE, Singapore, 113-1143.

Aminzadeh, F. and Brouwer, F., 2006, Integrating neural networks and fuzzy logic for improved reservoir property prediction and prospect ranking, 2006, SEG Annual Meeting, 1-6 October, New Orleans, Louisiana, Society of Exploration Geophysicists.

Archie, G.E., 1942, The electrical resistivity log as an aid in determining some reservoir characteristics, Petroleum Transactions of AIME, 146, 54-62.

Bagheripour, P. and Asoodeh, M., 2014, Genetic implanted fuzzy model for water saturation determination, Journal of Applied Geophysics, 103, 232-236.

Basak, D., Srimanta Pal, S. and Patranabis, D.C., 2007, Support vector regression, neural information processing.

Cheng, D.Z., Xi Ling, W. and Ju-An, C., 1993, Determining reservoir properties in reservoir studies using a fuzzy neural network, SPE Annual Technical Conference and Exhibition, 3-6 October, Houston, Texas, Society of Petroleum Engineers, 141-150.

Cuddy, S.J., 2000, Litho-facies and permeability prediction from electrical logs using fuzzy logic, SPE Reservoir Evaluation & Engineering, Society of Petroleum Engineers, 3(4), 319-324.

Gunn, S.R., 1998, Support vector machines for classification and regression, Technical Report, Faculty of engineering, science and mathematics, University of Southampton.

Hastie, T., Tibshirani, R. and Friedman, J., 2001, The elements of statistical learning: data mining, Inference, and Prediction, Springer, New York.

Ivanciuc, O., 2007, Applications of support vector machines in chemistry, Reviews in Computational Chemistry, 23, 291.

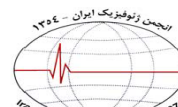
Jafari Kenari, S.A. and Mashohor, S., 2013, Robust committee machine for water saturation prediction, Journal of Petroleum Science and Engineering, 104, 1-10.

Joachims, T., 2003, Support Vector and Kernel Methods, Cornell University Computer Science Department, <http://www.Joachims.org>

Li, H.D., Liang, Y. and Xu, Q., 2009, Support vector machines and its applications in chemistry, Chemometrics and Intelligent Laboratory Systems,

احمدی و امیری بختیار، به کارگیری مدل رگرسیون ماشین بردار پشتیبان به منظور تخمین میزان اشباع شدگی آب سازند یکی از میدان‌های نفتی ...، صفحات ۱۹۹-۲۱۰.
RBF SVM parameters, <http://scikit-learn.org>.

- Singh, S., 2005, Permeability prediction using artificial neural network (ANN): A case study of Uinta basin, SPE Annual Technical Conference and Exhibition, 9-12 October, Dallas, Texas, Society of Petroleum Engineers.
- Smola, A.J. and Schölkopf, B., 2003, A tutorial on support vector regression, *Statistics and Computing*, 14 (3), 199-222.
- Vapnik, V., 1995, *The nature of statistical learning theory*, Springer, New York.
- Yu, H. and Kim, S., 2012, SVM tutorial: classification, regression, and ranking, *Handbook of Natural Computing*, Springer Berlin Heidelberg, 479-506.
- Zisserman, A., 2015, Lecture 3: SVM dual, kernels and regression, C19 Machine Learning, Hilary, <http://www.robots.ox.ac.uk/~az/lectures/ml/>.
- Lim, J.S., Park, H.J. and Kim, J., 2006, A new neural network approach to reservoir permeability estimation from well logs, SPE Asia Pacific Oil & Gas Conference and Exhibition, 11-13 September, Adelaide, Australia, Society of Petroleum Engineers.
- Luthi, E.M., 2001, *Geological well logs, their use in reservoir modeling*, Springer-Verlag Berlin Heidelberg, Germany.
- Parrella, F., 2007, Online support vector regression, A thesis presented for the degree of information science, Department of Information Science, University of Genoa, Italy.
- Parviainen, M., 2012, Radial Basis Function (RBF) and Support Vector Machines (SVM) networks, Tampere University of Technology, www.cs.tut.fi.
- Scikit-learn developers (BSD License), 2010-2014,



Application of support vector regression to estimate the formation water saturation in one of the largest oil fields located in the southwest of Iran

Reza Ahmadi^{1*} and Mohammad-Sadegh Amiri Bakhtiar²

1- Assistant professor, Mining Engineering Department, Arak University of Technology, Arak, Iran

2- M.Sc. Student, Abadan Petroleum University of Technology, Abadan, Iran

Received: 1 April 2017; Accepted: 17 October 2017

* Corresponding author: rezahmadi@gmail.com

Keywords

Water Saturation (S_w)

Support Vector Regression (SVR)

Asmari Formation

Well Logs Data

Extended Abstract

Summary

Water saturation (S_w) of a hydrocarbon reservoir is an important petrophysical parameter having a great impact on the accuracy of primitive estimation of the reservoir. Due to highly importance of this parameter dealing with the economic calculations of the reservoir, it must be estimated precisely.

Although experimental analysis of core samples taken from a reservoir leads to very useful information about S_w of the reservoir, this experimental method is highly expensive and time consuming; and therefore, this method is applicable only for a small number of wells in a field. To overcome this problem, an intelligent pattern recognition method, known as support vector regression (SVR), has been employed in the current research to estimate S_w from well logs data of 3 wells in one of the largest oil fields of Iran. The performance of the algorithm has also been validated through different criteria. The results of this research indicate that the SVR model can estimate S_w from well logs data accurately, in which the determination coefficients of 87 and 76 percent have been obtained from the training and test steps, respectively.

Introduction

Generally in most commonly hydrocarbon reservoirs, S_w is estimated using well logs data through applying Archie's fundamental empirical relation. However, this relation is just satisfied for clean sandstone formations (without clay minerals). So far several empirical models have been proposed to measure S_w using well logs data. The main disadvantage of these models is their formation dependency, which makes the models specific and not comprehensive to be applied in a variety of other formations. In addition to empirical methods, several linear regression techniques have also been applied to estimate this parameter using well logs data. These techniques cannot estimate S_w appropriately due to the complexity of the parameter features. Resistivity and porosity logs are the most important well logs used to estimate S_w by Archie's relation. The porosity of a formation can be very accurately determined through sonic, density and neutron logs. However, resistivity logs are very sensitive to the presence of shale and other clayey impurities in formations. Their effects can be adjusted by means of gamma ray (GR) log. Therefore, to estimate S_w , employing an intelligent method using appropriate well logs data will be useful. The oil reservoir, studied in this research, is located in Asmari formation in southwest of Zagros Mountain. Overall this formation in the investigated region has been formed from a sequence comprising of carbonate rocks (limestone and dolomite), sandstone and shale.

Methodology and Approaches

In the current research, to estimate S_w , SVR method has been applied to well logs data from 3 wells in one of the largest oil fields of Iran. In this study, appropriate well logs data comprising of GR, neutron porosity, formation bulk density, sonic transit time and true resistivity from deep induction log (ILD) have been used. Moreover, S_w values measured from cores in the laboratory are available for whole depth of the wells. In order to employ SVR to estimate S_w , the model needs to be trained using appropriate input and output data in MATLAB environment. In the current research, the input consists of 5 variables (well logs data) while the output is only the S_w parameter. From 1211 data points (containing 5 variables of well logs data and S_w parameter measured by core) available from the 3 wells, about 80 percent (i.e. 988 samples) were selected for training and the remaining 20 percent (i.e. 223 samples) were chosen for test. To compare the estimated values with the measured ones for the reservoir in the study region, visually, chart of lithology, water and hydrocarbon saturations of the formation were also depicted for the 3 wells by means of Geolog software.

Results and Conclusions

The performance of the algorithm has been validated through different criteria such as scatter plot of S_w values from cores versus the estimated S_w values from well logs data of 3 study wells by means of SVR model as well as computing statistical parameters indicating the accuracy of the results. Furthermore, the results of the research revealed that the SVR model can estimate S_w using well logs data accurately so that it has estimated the training and test data with the determination coefficients of 87 and 76 percent, respectively. As a result, the proposed method, i.e. SVR, is an accurate, fast and cost-effective method to evaluate the petrophysical parameter S_w .
