ORIGINAL PAPER - EXPLORATION ENGINEERING

Data-driven model for shear wave transit time prediction for formation evaluation

David Onalo[1,2] · Sunday Adedigba[3] · Olalere Oloruntobi[1] · Faisal Khan[3] · Lesley A. James[2] · Stephen Butt[1]

© The Author(s) 2020

Abstract
Sonic well logs provide a cost-effective and efficient non-destructive tool for continuous dynamic evaluation of reservoir formations. In the exploration and production of oil and gas reservoirs, these sonic logs contain crucial information about the formation. However, shear sonic logs are not acquired in all oil and gas exploration wells. More so, many offset wells are not run with the most recent sonic logging tools capable of measuring both shear and compressional sonic transit times due to the relatively high costs of running such equipment. And in wells where they are deployed, they are run only over limited intervals of the formation. Such wells lack continuous shear wave transit time measurements along the formation. In this study, an exponential Gaussian process model is presented. The model accurately predicts the shear wave transit times in the formations which lack reliable shear wave transit time measurements. The proposed model is developed using an array of well logs, namely depth, density, porosity, gamma ray, and compressional transit time. A Monte Carlo simulation is used to quantify the proposed model uncertainty. The shear sonic transit time predictions are used to estimate some formation deformation properties, namely Young's modulus and Poisson's ratio of a reservoir formation. The results suggest that shear transit time can be represented and predicted by Gaussian-based process model with RMSE, R^2, and MSE of 11.147, 0.99, and 124.6, respectively. The proposed model provides a reliable and cost-effective tool for oil and gas dynamic formation evaluation. The findings from this study can help for better understanding of shear transit times in formations which do not have multipole sonic logs or where data have been corrupted while logging in the Niger Delta.

Keywords Gaussian process · Exploration · Sonic log · Well logging · Velocity · Reservoir formation

List of symbols

RHOB	Bulk density log (g/cm^3)
DTCO	Compressional wave travel time (µs/ft)
Δt_c	Compressional wave travel time (µs/ft)
RESD	Deep resistivity log (ohm m)
PHIE	Effective porosity log (m^3/m^3)
ϕ_{den}	Electron density porosity (g/cc)
ρ_{fl}	Fluid density (g/cc)
ρ, ρ_b	Formation density (g/cc)
Δt_{fl}	Formation fluid compressional wave travel time (µs/ft)
Δt_{ma}	Formation matrix compressional wave travel time (µs/ft)
GR$_{min}$	Gamma ray in clean sandstone
GR$_{min}$	Gamma ray in shale
GR	Gamma ray log (gAPI)
GR	Gamma ray log reading
ρ_{ma}	Matrix density (g/cc)
MAE	Mean absolute error
MPE	Mean percentage error
Δt_{log}	Measured compressional wave travel time (µs/ft)
ϕ_N	Neutron porosity
$\phi_{N,fl}$	Neutron response of the fluid

✉ Faisal Khan
fikhan@mun.ca

1   Drilling Technology Laboratory (DTL), Faculty of Engineering and Applied Science, Memorial University of Newfoundland, St. John's, NL A1B 3X5, Canada

2   Hibernia Enhanced Oil Recovery (EOR) Group, Faculty of Engineering and Applied Science, Memorial University of Newfoundland, St. John's, NL A1B 3X5, Canada

3   Centre for Risk, Integrity and Safety Engineering (CRISE), Faculty of Engineering and Applied Science, Memorial University of Newfoundland, St. John's, NL A1B 3X5, Canada

مدينة الملك عبدالعزيز
KACST للعلوم والتقنية

🖄 Springer

$\phi_{N,\text{ma}}$	Neutron response of the matrix
ϑ	Poisson's Ratio
ϕ	Rock porosity
Vsh	Shale volume
DTSM	Shear wave travel time (µs/ft)
Δt_s	Shear wave travel time (µs/ft)
ϕ_s	Sonic porosity
PHIT	Total porosity log (m³/m³)
E	Young's modulus (MPa)

Introduction

Sonic well logs have been around since the 1900s in the petroleum industry (Alford et al. 2012; Doh and Alger 1958). Over the years, geologists, petrophysicists, and petroleum engineers have come to see the reliability and usefulness of sonic well logs in the exploration and production of hydrocarbon reservoirs. Onalo et al. (2018a) developed an artificial neural network to predict the compressional and shear wave sonic logs along a wellbore from a producing well. Drilling engineers use sonic data to improve drilling efficiency and reduce target offset margins (Alford et al. 2012). The transmission of sonic waves through the formation, "sonic well logging", provides valuable data such as compressional transit time and shear transit time that is used in formation evaluation (Minear and Fletcher 1983). Sonic logging was the first tool that provided the industry with a means to estimate formation porosity without knowledge of the fluid saturation (Raymer et al. 1980). As far back as 1958, researchers like Doh and Alger (1958) perceived formation porosity estimation to be the major advantage of sonic logs. The transit arrival times of the sonic waves have evolved and now being used for formation, porosity determination, lithology identification, fluid saturation indication, formation strength characterization, hydrocarbon indication, and much more (Khazanehdari and Mccann 2005; Williams 1990). This is due to the fact that the sonic transit times are affected by reservoir properties that include compaction, porosity, anisotropy, density, lithology, cementation, consolidation, overburden stress and pore pressure (Khazanehdari and Mccann 2005; Krief et al. 1990; Thomsen 1986; Toksöz et al. 1976; Williams 1990). A good understanding of how these properties change over the life of the reservoir is essential for proper reservoir planning, development and management (Dakhelpour-Ghoveifel et al. 2019; Khazanehdari and Mccann 2005; Saboorian-Jooybari et al. 2015).

Well-calibrated and reliable sonic logging tools are necessary to acquire accurate measurements of compressional and shear wave transit time, otherwise, the formation evaluations and estimation become false and misleading (Onalo et al. 2018a, 2019). Typical anomalies observed in well logs have been presented in the literature (Saboorian-Jooybari

et al. 2016, 2015). This may result in the development of non-potential reservoirs and the abandonment of potential reservoir formations. Sonic logging tools have also evolved over the years, from single transmitters and receivers to two receivers to compensate for discrepancies from the transmission source due to borehole and mud. This is known as the borehole effect (Doh and Alger 1958). The spacing between the receivers is usually about one feet to ensure a proper description of the medium. To correct the errors generated as a result of the irregularities of the borehole, borehole-compensated sonic tools were developed (Kokesh et al. 1965). To further improve the quality of the sonic measurements, array sonic logging tools were adopted that contains an array of transmitters and receivers (Hsu et al. 1987). The above-mentioned sonic logging tools are mainly monopole sonic logging as they do not provide measurements of the shear wave, especially in fast formations (Alford et al. 2012; Harrison et al. 1990). Fast formations are formations in which the shear wave response of the formation arrives at the receivers before the compressional wave response of the wellbore fluid. In situations where the compressional wave response of the borehole fluid arrives before the shear wave response of the formation, the formation is known as a slow formation. More modern sonic logging tools include dipole sonic and multipole sonic logging tools which are capable of measuring both compressional and shear wave properties directly or indirectly by generating flexural waves (Alford et al. 2012; Market and Canady 2006).

Shear wave transit time is vital for many geophysical and engineering analyses including seismic interpretations and bright spot analysis (Greenberg and Castagna 1992a; Onalo et al. 2018b). The lack of shear wave transit time data limits the number of valuable relationships and correlations that can be derived from sonic logging, especially for lithology identification, fluid saturation identification and porosity estimation (Domenico 1984; Onalo et al. 2018a, b). Shear wave transit time alone is not sufficient to provide a full description of the diversity across the reservoir formation (Greenberg and Castagna 1992a).

Empirical relationships have been developed to estimate the shear wave velocity from compressional wave velocity in situations where the shear wave data were missing (Bailey 2012; Castagna et al. 1985; Domenico 1984; Eberhart-Phillips et al. 1989; Esene et al. 2018; Gardner et al. 1974; Greenberg and Castagna 1992b; Hamada 2004; Han et al. 1986; Jorstad et al. 1999; Krief et al. 1990; Lee 2006; Miller and Stewart 1974, 1990; Oloruntobi et al. 2019; Oloruntobi and Butt 2019; Ramcharitar and Hosein 2016; Raymer et al. 1980; Takahashi et al. 2000; Vernik et al. 2002). Though these estimations provide simple correlation for quick estimations, they are not as robust as modern-day machine learning techniques that have been applied in several engineering applications (Kumar et al. 2014; Nourafkan and

Kadkhodaie-Ilkhchi 2015; Onalo et al. 2018a, 2019; Ramcharitar and Hosein 2016; Reichel et al. 2012).

Gaussian process (GP) is a powerful technique for predicting and modeling complex mathematical and engineering data-driven problems. GP involves defining a finite vector space function of infinite dimension over a Gaussian distribution. The GP has been used in many engineering applications due to its flexibility to model nonlinear complex patterns between dataset variables (MacKay 2005). The GP has been adopted in solving many engineering and real-life problems because of its ability to handle data in various forms and sizes (Akin et al. 2008; Ali Ahmadi and Golshadi 2012; Asadisaghandi and Tahmasebi 2011; Ashoori et al. 2010; Babakhani et al. 2015; Derakhshanfard and Mehralizadeh 2018; Ebden 2008; Huang et al. 2003; Iturrarán-Viveros and Molero 2013; Kelechukwu et al. 2013; Riazi et al. 2014; Sheremetov et al. 2014; Vaferi et al. 2014). A general sketch of the problem is presented in Fig. 1 after Yu et al. (2016).

Some examples of Gaussian-based processes that have been developed to solve problems in the industry are presented in Table 1.

Considering the success that Gaussian-based processes have had in several petroleum engineering applications, the objective of the paper is to develop a reliable model that can reproduce shear wave sonic logs using a Gaussian-based process from the available array of well log data. The importance of such a model to the industry is invaluable for offset wells that have been drilled and logged without dipole or multipole sonic logging tools and therefore do not have the corresponding shear wave sonic logs. Also, in formations where the inaccurate log data have been obtained due to damaged equipment or calibration (human) error. Sonic logs are essential components for drilling, exploration and reservoir management. The shear wave sonic logs provide accurate continuous predictions of the reservoir properties for better reservoir planning and management.

To the best of the authors' knowledge, the current work presents the first Gaussian distribution of shear transit time from other well logs located in West Africa. The current work will help push for the development of similar models in the region without the need for costly well interventions.

Gaussian process (GP)

Modeling complex engineering problems present a real challenge in the petroleum industry. The GP is a probabilistic modeling technique that is nonparametric, meaning that the prior is placed in space and the actual distribution that fits the data is not known before the initialization (Huang et al. 2017; Kuss and Rasmussen 2006). GP has been recognized as a promising data mining technique in machine learning due to its ability to handle large amounts of data (Han and Kamber 2010). GP is generally classified into supervised and unsupervised. Simply put, supervised GP involves establishing functions of input datasets used for the training to predict the corresponding output dataset (Rostami et al. 2013). In unsupervised, there is no prediction as there is no target output dataset or prior history from which to establish functions. Nonetheless, this is very useful functionality for classifying large datasets. When the GP is used for prediction, it is referred to as a GP. On the other hand, if the GP is used for classification, it is referred to as GP classification (Rostami et al. 2013). GP captures set finite random variables and attempts to represent them by a joint Gaussian distribution (Rasmussen 2004). GP is defined fully by its mean and covariance functions (Seeger 2004). Gaussian process-based models are highly capable of establishing nonlinear relationships from nonparametric data and deriving algorithms for future predictions (Abdollahzadeh et al. 2012). GP is highly universal and can be adapted to various problems presented; however, care must be taken to select the best covariance, kernel, and hyperparameters describing the multidimensional distribution (Kuss and Rasmussen 2006).
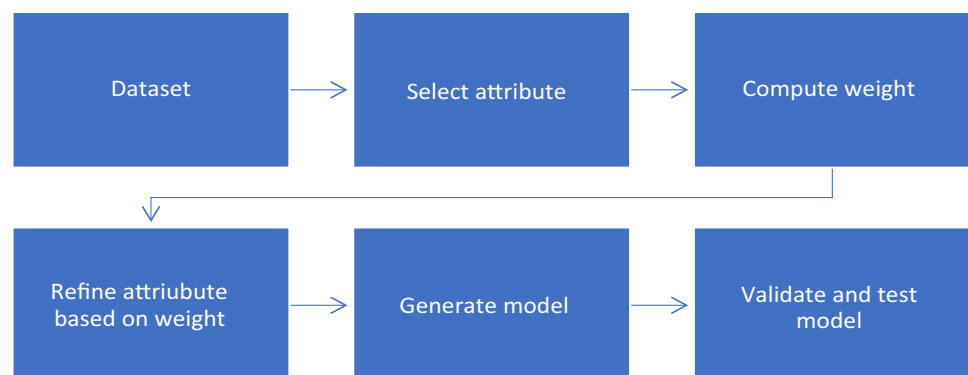
Fig. 1 A general sketch of the problem

Springer

Table 1 Examples of Gaussian-based processes used in the petroleum industry

S/no.	Reference	Intelligent system	Input data	Output data	Application
1	Williams (1990)	Gaussian distribution	Compressional and shear velocity-vs	Distribution of a random variable $e = vp/vs$ observed $-vp/vs$ predicted	Statistical differentiation of hydrocarbon bearing formations
2	Su (2009)	Gaussian process machine learning	Day and measured displacement	Predicted displacement	Nonlinear deformation time series
3	Abdollahzadeh et al. (2012)	Gaussian-based estimation of distribution algorithms	Facies description, porosity, and permeability	History match	Optimized solutions to continuous history matching for numerical reservoir models.
4	Iglesias et al. (2013)	Gaussian approximations	Production rate, pressure, and time interval	Log permeability and well locations	Data assimilation in reservoir models
5	Rostami et al. (2013)	Gaussian process regression	Temperature, pressure, molecular weight, specific gravity, and acentric factor	Crude oil density	Prediction of undersaturated crude oil density
6	Rostami and Khaksar Manshad (2013)	Gaussian process regression	Subsets of asphaltene content, C_1-C_{7+} composition, $H_2S+N_2+CO_2$, specific gravity, molecular weight, pressure, temperature and weight percent	Asphaltene precipitation	Prediction of asphaltene precipitation
7	Izadi and Elsworth (2014)	Gaussian distribution	Stresses, pressures, temperature, and peak strength	Permeability evolution and induced seismicity	Reservoir stimulation and induced seismicity:
8	Kumar et al. (2014)	Gaussian process regression	Joint frequency (Jn), joint inclination parameter (n), joint roughness parameter (r), confining pressure (s3), and elastic modulus (ei) of intact rock	Elastic modulus jointed rock	Rock mechanics
9	Yu et al. (2016)	Gaussian process regression	Gamma ray, density, sonic, neutron, resistivity logs, and photoelectric absorption factor	Total organic carbon (toc)	Toc estimation
10	Silva and Deutsch (2016)	Gaussian mixture models	Measurements of nickel in iron (Fe), silica (SiO_2), and magnesium (Mgo)	Missing measurements in dataset	Optimization of computational efficiency in the geostatistical analysis
11	Adeli et al. (2017)	Multivariate Gaussian simulation	Grades of iron (Fe), silica (SiO_2), phosphorus (P), alumina (Al_2O_3), manganese (Mn), loss on ignition (loi), and the granulometric fraction	Rock or ore type	Interpretation of geological model,
12	Huang et al. (2017)	Gaussian process regression	Secondary stresses (σ_2 and σ_3)	Principal stress (σ_1)	Determine intact rock strength
13	Current work	Gaussian process regression	Depth, porosity, bulk density, gamma ray, and compressional travel time logs	Shear transit time	Formation evaluation

Methodology

The framework of the proposed model development and testing is presented in Fig. 2.

Data collection and preparation

The data required for the proposed methodology are actual well logs. The data should contain the relevant logs required for the proposed model, namely depth (ft), density—RHOB (g/cc), total porosity—PHIT, gamma ray—GR (GAPI), DTCO (μs/ft), and DTSM (μs/ft). This methodology is applicable to any combination of logs. Various combinations of log data have been used; however, the proposed combination gives a more accurate prediction. Reducing the set of well logs combination reduces the accuracy of the prediction. Accuracy was one of the objectives of the study; therefore, this work uses the combination of gamma ray, porosity, density, and compressional transit time logs along with their corresponding depth to predict the shear transit log.

Quality assurance and quality checks (QAQC)

Quality assurance and quality control (QAQC) were performed on the suite of well logs to ensure the reliability of the data. Firstly, the logs were analyzed to identify null

Fig. 2 Framework for the proposed model development of the Gaussian process
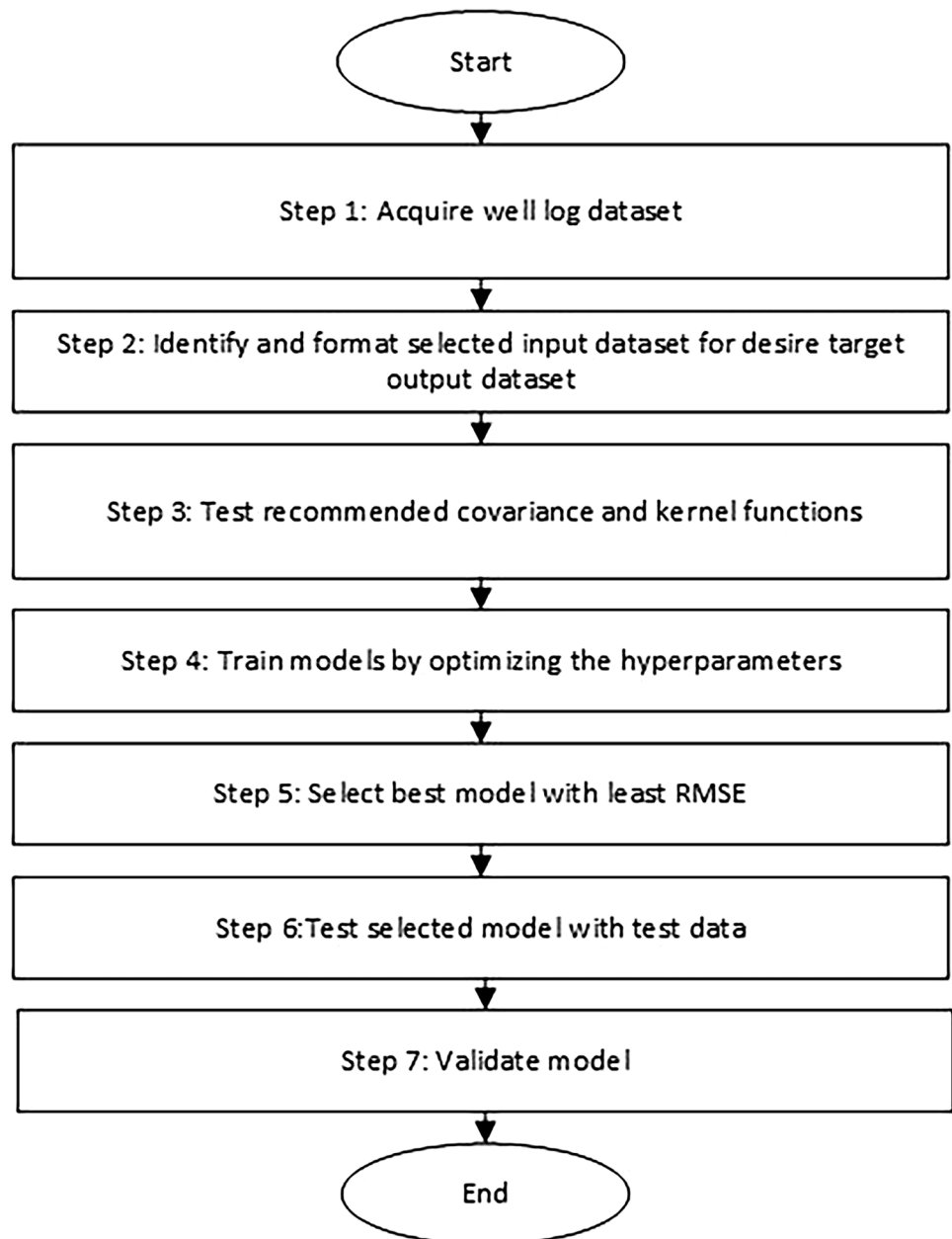
readings where the logging tools failed to accurately record the corresponding measurements. Secondly, the sections where washout and key-seat (define) sections were observed were removed for the model development by referencing with the caliper logs for adjacent formation sections. Poisson's ratio calculations were used to ensure only valid sections were represented in the dataset. These tests calculate Poisson's ratio from the measured compressional and shear sonic transit times to ensure reasonable values, and are an industry standard logging quality check.

Gaussian process model development

A shaley sandstone formation located in West Africa was used in this study. For the model development, well log data from 2850 to 6000 ft and from 8000 to 12,500 ft of a sandstone reservoir were used to build and train the model. To test the model, the section from 6000 to 8000 ft was used. The prepared data were formatted to match the initial model set up with five predictors and one response (name). The actual target response is presented in Fig. 3. Shear transit time on the y-axis and record number represents the number of reference data points by the model on the x-axis.

The GP distribution is applied to the dataset; however, the kernel function that best represents the distribution function is not known. Therefore, a set of kernel functions were applied to the dataset to ascertain which kernel function was able to best smoothen the dataset and provide the least errors. The squared exponential kernel, exponential kernel, Matern 5/2 kernel, and the rational quadratic kernel

were applied to the dataset. Each GP model and kernel function were trained by constantly updating the hyperparameters until the best match describing the well log correlation was reached by the respective models. GP generally does not suffer from overfitting like other intelligent systems like neural networks (Adedigba et al. 2017; Onalo et al. 2018a). Nevertheless, overfitting can arise from the marginal likelihood optimization, especially with many hyperparameters (Mohammed and Cawley 2017; Rasmussen and Williams 2006). To solve this problem, cross-validation was used. First, the dataset is divided into five disjoint sets (folds). For each fold, the out-of-fold data points are used to train the model, and then the performance of the model is assessed using the in fold data points. The average error is then calculated across all the folds (Matlab Documentation 2018).

Gaussian process selection

The test results of each model are presented in Table 2. In Fig. 4, the prediction response of each model is presented. All models performed well and were able to represent the data with a relatively high degree of accuracy; however, the exponential GP model seemed to better follow the actual plotted response (Fig. 3) more closely. The predicted responses of the models are plotted against the actual responses in a cross-plot in Fig. 5 to validate the model prediction accuracy. All the tested models have coefficients of determination (R^2) of 0.99. This proved that GP was highly capable of predicting the responses and would thus provide relatively sufficient models. Nonetheless, the coefficient of

Fig. 3 Actual target response of the shear sonic log
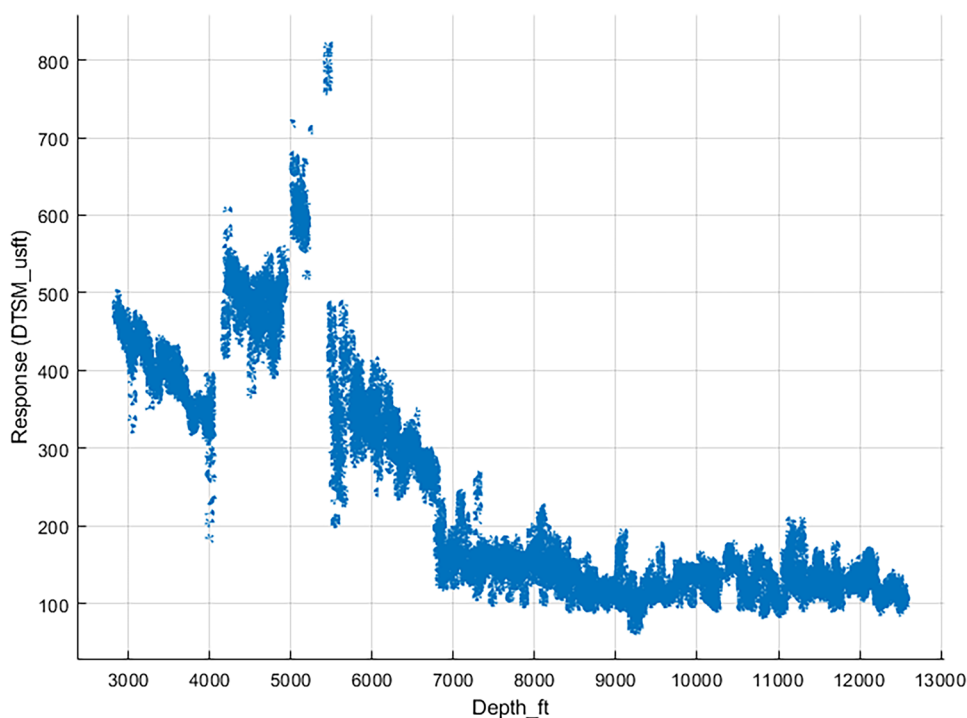
determination could not be used as a basis for the selection of the optimum GP model. The residuals of the model predictions are plotted in Fig. 6. The exponential GP model presented the least error through the dataset.

The exponential GP model had the least root mean square (RMSE) with a value of 11.147. This was followed by the Matern 5/2 GP model with an RMSE = 12.718, and then the rational quadratic Gaussian model with an RMSE = 12.786. The squared exponential Gaussian model, though popular, presented the highest error with an RMSE = 13.774. The mean square error (MSE) and mean average error (MAE) follow the same trend as the root mean square error. The squared exponential kernel is the most popular GP; however, the exponential GP outperformed the squared GP. It is difficult to say with utmost certainty why the exponential GP outperformed the other models. The margin of difference between the models is within 5%; however, this may be due

to the exponential relationship between density, porosity and sonic logs along the depths of a formation as suggested by several researchers (Dey and Stewart 1997; Gardner et al. 1974; Miller and Stewart 1974).

A summary of the results of the tested models is presented in Table 2. The exponential Gaussian process model was selected as the best of the four models tested.

The results of the models are compared with a multilinear regression model with the same predictors (input well logs) and output to ensure that the proposed GP model is not redundant. The result of the multilinear regression is presented in Table 3. The result alludes to the improvement and accuracy of the new model prediction by the use of the GP model. The RMSE, R^2, and MSE reduced to 37.486, 0.93, and 24.692 in the multilinear model, respectively.

Uncertainty analysis

The authors acknowledge that even though the boundary limits of the input well logs and output shear transit time for a sandstone formation can be defined to a reasonable extent, tackling the presence of uncertainty in the development of the model presents a real challenge. To address the uncertainty in the developed model, a Monte Carlo simulation based on a $\mp 10\%$ uncertainty in the well log array used in the model development. The results of the simulated models

Table 2 Summary of the results of the GP model selection

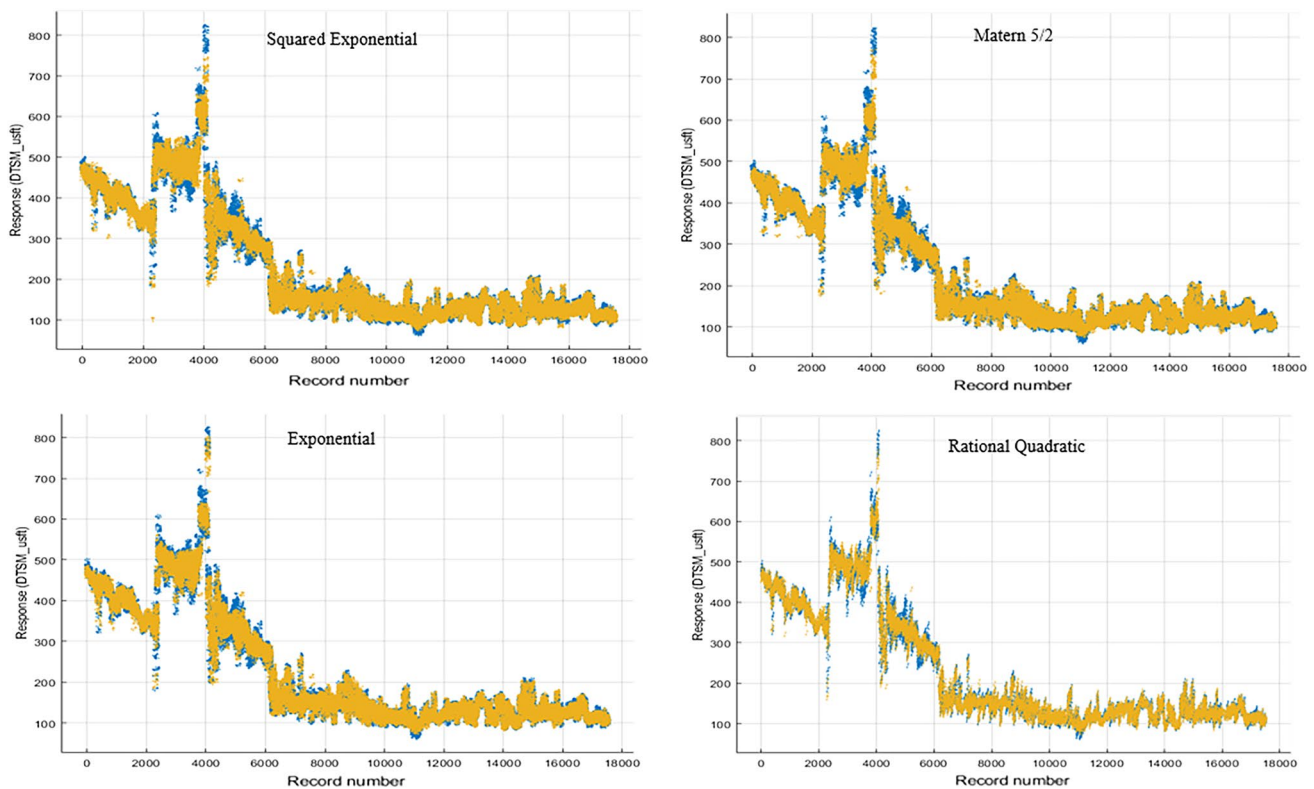| Model type	RMSE	R^2	MSE	MAE
Exponential GP	11.147	0.99	124.26	6.6046
Squared exponential GP	13.774	0.99	189.71	8.0162
Matern 5/2 GP	12.718	0.99	161.74	7.4618
Rational quadratic GP	12.786	0.99	163.48	7.5401

Fig. 4 The response of the tested GP models (actual in blue and predicted in yellow)
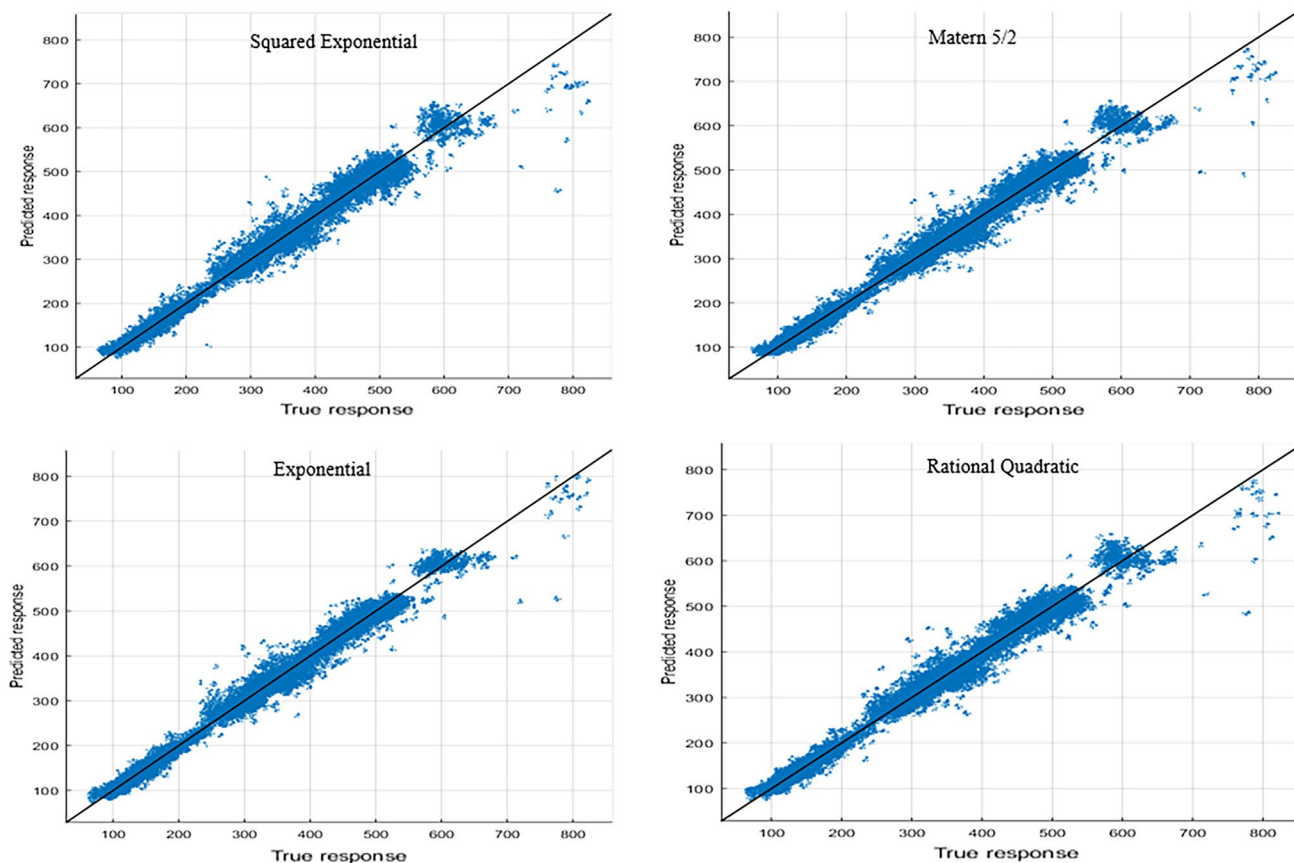
Fig. 5 Cross-plot of the tested GP models

are presented in Table 4. Table 4 suggests that the exponential GP model outperforms all other GP models with an RMSE of 20.11 and MAE of 12.11. The squared exponential model has the least performance with an RMSE of 21.87 and MAE of 13.13. This is consistent with the previous results in the initial model development. However, unlike in the initial model set up, the rationale quadratic GP model outperformed the Matern 5/2 GP with an RMSE of 20.92 and MAE of 12.52. The coefficient of determination, although lower than the initial model, was 0.98 for all Monte Carlo simulated models. The uncertainty simulation portrays a significant increase in the RSME from 11.18 to 20.92 and an increase in MAE from 6.6 to 12.11 for the exponential GP. Nevertheless, the model is robust as the coefficient of determination was only reduced from 0.99 to 0.98.

Generalization of the GP

GP machine learning is adaptive in nature and can be generalized for datasets in a similar format as the original training data. To validate and ensure that the proposed model is generalized, it is applied to the section of the well log data that were omitted in the development of the model in Sect. 3.

Application of the developed model

The proposed shear wave transit time model is validated by applying the proposed model to actual well logs. The geological setting of the formation used for calibrating and validating the model is a shaley sandstone formation located in West Africa. The formation is normally pressured producing oil reservoir. The well log data presented in this study cover a 2000-ft section, from 6000 to 8000 ft of an actual oil and gas sandstone reservoir. This is an improvement to most studies conducted on a section of only several hundred feet. The location and details of the well log data have been withheld in this study to protect the privacy and confidentiality of the logging company. Nevertheless, the first 100 ft of the data is presented in Table 5 for interested users. A plot of the available well log data for the study is presented in Fig. 7.
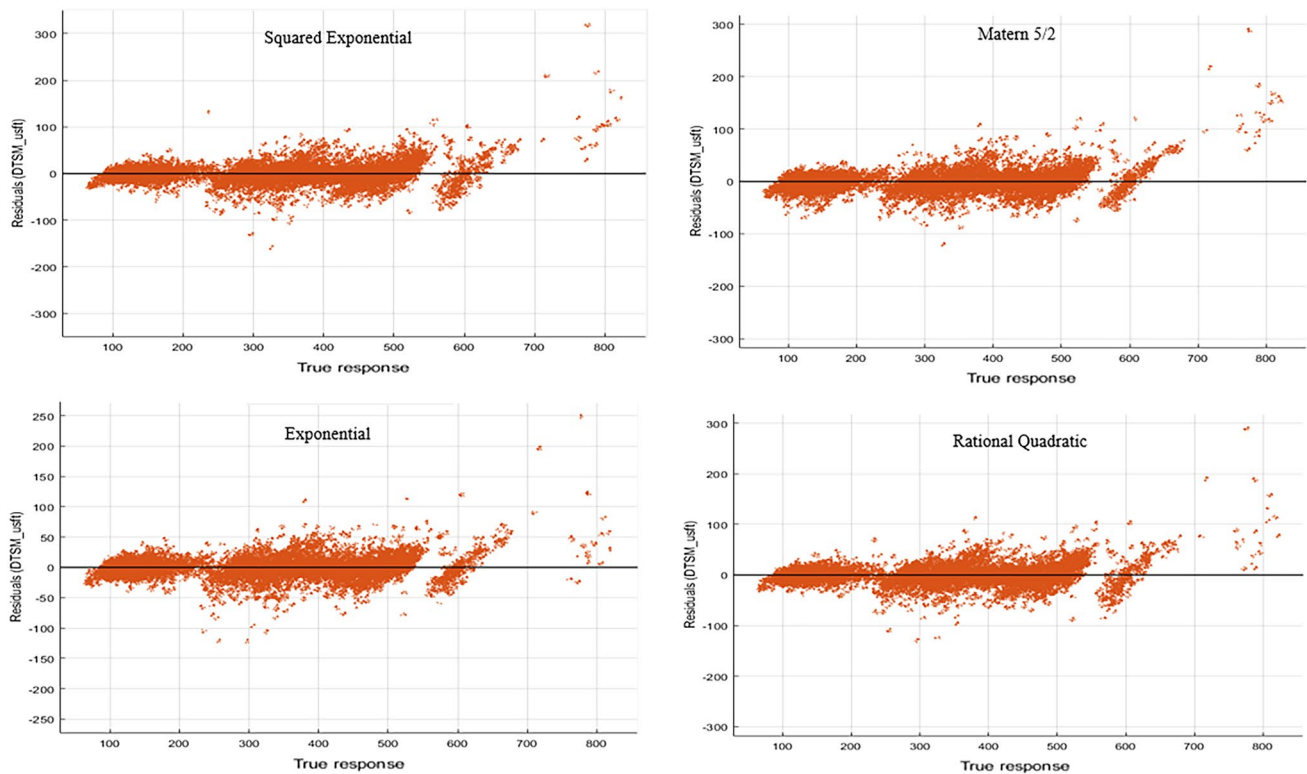
Fig. 6 Residuals of the tested GP models

Table 3 Summary of the results of the multilinear regression model selection

Model type	RMSE	R^2	MSE	MAE
Multilinear regression	37.486	0.93	1405.2	24.692

Table 4 Summary of the results of the Monte Carlo simulated GP models

Model type	RMSE	R^2	MSE	MAE
Exponential GP	20.11	0.98	404.46	12.11
Squared exponential GP	21.87	0.98	478.35	13.13
Matern 5/2 GP	21.18	0.98	448.55	12.68
Rational quadratic GP	20.92	0.98	437.49	12.52

Results and discussion

Shear Sonic transit time log estimation

The primary objective of developing a Gaussian-based process model from well log data is to provide a tool adequate enough to furnish reliable shear sonic transit time logs in wells from offset wells run with monopole sonic logging tools. More so, in wells with corrupted datasets or erroneous readings from faulty equipment. In Fig. 8, the measured shear sonic transit log is plotted against the depth profile of the wells used for the case study from 6000 to 800 ft. This is followed by a plot of the shear sonic transit log along the same depth in Fig. 8. The predicted shear sonic log closely matches the measured shear sonic log values. The most disparity is seen from 6010 to 6050 ft with less than a 5% difference in value. What is very intriguing is that the proposed model is relatively conservative in the sense that it tries to follow the measured shear sonic log trend, without going out of the measured shear sonic log boundaries in the well. This ensures that analysis conducted using the models are reliable and safe as they do not venture away from or to the extreme boundary scenarios of the formation. To further depict the success of the model in predicting the shear sonic transit log from the well logs proposed in the previous section, a cross-validation plot of the predicted shear sonic transit time versus the measured shear sonic transit time log is presented in Fig. 9. The proposed model does a good job of almost matching the measured shear sonic logs with a coefficient of determination of 0.9923. The trend line in Fig. 9 also falls on the perfect unity slope line, the figure thereby portraying a non-bias in the predictions of the proposed model. The results show that the proposed model achieves the desired objective of the proposed model by accurately predicting the shear sonic transit log of the well.

مدينة الملك عبدالعزيز
للعلوم والتقنية KACST
Springer

	Depth (ft)	DTCO (µs/ft)	DTSM (µs/ft)	GR (gAPI)	PHIT (m³/m³)	RHOB (g/cc)
Table 5 50 ft of available well log data for the study	6000.00	131.82	324.03	87.14	0.19	2.32
	6000.50	133.00	349.97	87.44	0.20	2.31
	6001.00	134.20	363.23	86.52	0.20	2.30
	6001.50	133.26	342.61	86.24	0.20	2.31
	6002.00	131.83	312.74	86.66	0.19	2.32
	6002.50	130.86	304.08	85.50	0.18	2.33
	6003.00	130.94	304.68	84.60	0.18	2.33
	6003.50	129.35	315.31	85.13	0.18	2.34
	6004.00	129.85	338.09	85.31	0.18	2.33
	6004.50	131.38	363.53	87.60	0.19	2.32
	6005.00	133.30	379.96	88.31	0.19	2.32
	6005.50	133.26	379.00	91.86	0.19	2.32
	6006.00	132.83	373.47	92.08	0.19	2.32
	6006.50	132.59	361.99	89.04	0.19	2.32
	6007.00	131.24	358.47	86.43	0.18	2.34
	6007.50	131.50	361.11	88.95	0.17	2.35
	6008.00	132.92	371.02	94.94	0.17	2.36
	6008.50	132.07	374.90	97.52	0.17	2.36
	6009.00	133.00	378.18	95.00	0.17	2.35
	6009.50	133.95	379.63	89.53	0.18	2.34
	6010.00	133.47	350.07	85.95	0.18	2.33
	6010.50	133.49	332.87	84.92	0.19	2.32
	6011.00	134.39	346.93	85.52	0.19	2.31
	6011.50	133.68	367.18	84.63	0.19	2.32
	6012.00	133.95	363.88	82.08	0.18	2.33
	6012.50	132.86	356.71	86.91	0.18	2.34
	6013.00	132.95	363.06	93.85	0.18	2.34
	6013.50	131.60	342.72	94.98	0.19	2.32
	6014.00	131.81	316.72	91.16	0.19	2.31
	6014.50	131.79	311.10	85.44	0.20	2.31
	6015.00	131.07	307.31	82.28	0.19	2.31
	6015.50	132.31	310.24	77.20	0.19	2.31
	6016.00	131.97	322.17	77.37	0.19	2.32
	6016.50	130.74	364.35	83.84	0.19	2.31
	6017.00	132.42	385.76	86.86	0.20	2.30
	6017.50	133.16	387.82	86.94	0.21	2.29
	6018.00	133.07	372.00	82.35	0.20	2.30
	6018.50	132.27	335.59	81.35	0.20	2.30
	6019.00	131.38	323.30	80.80	0.20	2.30
	6019.50	131.01	318.53	80.37	0.20	2.31
	6020.00	130.82	319.37	82.46	0.19	2.32
	6020.50	130.81	330.04	80.39	0.19	2.32
	6021.00	130.69	366.92	82.67	0.19	2.32
	6021.50	131.93	373.65	81.17	0.19	2.31
	6022.00	131.37	366.63	87.45	0.19	2.31
	6022.50	130.85	326.51	88.73	0.19	2.31
	6023.00	131.18	315.40	92.32	0.19	2.31
	6023.50	128.95	315.07	91.10	0.19	2.31
	6024.00	130.31	319.97	91.63	0.19	2.32
	6024.50	130.17	323.64	89.42	0.19	2.32
	6025.00	129.66	317.45	86.47	0.19	2.32

Table 5 (continued)

Depth (ft)	DTCO (µs/ft)	DTSM (µs/ft)	GR (gAPI)	PHIT (m³/m³)	RHOB (g/cc)
6025.50	129.85	321.53	82.90	0.19	2.31
6026.00	130.95	333.50	81.50	0.20	2.30
6026.50	130.85	341.21	82.07	0.20	2.31
6027.00	131.47	336.61	83.38	0.19	2.31
6027.50	132.42	332.92	85.78	0.19	2.32
6028.00	132.36	324.53	87.77	0.19	2.32
6028.50	132.19	320.65	91.52	0.19	2.32
6029.00	132.28	315.04	91.68	0.19	2.32
6029.50	132.71	313.44	87.51	0.19	2.32
6030.00	132.55	316.99	85.33	0.19	2.32
6030.50	132.35	324.50	87.06	0.19	2.32
6031.00	131.76	357.34	92.97	0.19	2.33
6031.50	132.15	396.80	95.18	0.19	2.33
6032.00	131.76	401.90	95.50	0.19	2.33
6032.50	132.21	402.31	95.32	0.19	2.32
6033.00	133.53	406.33	93.26	0.20	2.31
6033.50	134.37	413.76	89.90	0.20	2.29
6034.00	135.44	404.15	89.06	0.21	2.28
6034.50	135.84	398.00	91.33	0.21	2.28
6035.00	131.67	392.08	96.32	0.21	2.28
6035.50	130.85	387.02	97.00	0.20	2.29
6036.00	130.95	354.82	96.31	0.20	2.30
6036.50	129.68	339.62	88.86	0.20	2.31
6037.00	129.51	329.55	83.93	0.19	2.32
6037.50	125.87	307.18	77.83	0.19	2.32
6038.00	126.05	292.21	78.55	0.19	2.31
6038.50	122.92	284.44	79.17	0.19	2.31
6039.00	125.14	290.37	81.78	0.19	2.32
6039.50	128.86	307.30	84.62	0.19	2.33
6040.00	128.77	322.01	82.73	0.18	2.34
6040.50	129.55	365.22	83.52	0.18	2.34
6041.00	130.52	354.40	84.58	0.18	2.33
6041.50	130.84	315.02	89.45	0.19	2.32
6042.00	131.80	356.62	93.99	0.19	2.32
6042.50	132.12	322.97	92.13	0.19	2.31
6043.00	130.62	317.84	89.23	0.19	2.32
6043.50	128.64	310.26	84.82	0.18	2.33
6044.00	126.87	309.84	84.66	0.18	2.34
6044.50	126.66	305.22	82.88	0.18	2.34
6045.00	127.37	299.49	80.86	0.18	2.33
6045.50	128.29	305.08	81.14	0.19	2.32
6046.00	129.59	311.22	82.66	0.19	2.32
6046.50	129.64	320.78	85.12	0.19	2.33
6047.00	129.81	327.75	82.91	0.18	2.33
6047.50	129.74	352.66	82.84	0.18	2.34
6048.00	130.00	358.57	81.03	0.18	2.34
6048.50	130.38	324.24	81.19	0.18	2.33
6049.00	131.37	323.73	81.31	0.19	2.32
6049.50	132.51	377.43	84.68	0.20	2.31
6050.00	132.55	386.24	91.19	0.20	2.31
6050.50	132.55	393.98	90.16	0.19	2.31

Table 5 (continued)

Depth (ft)	DTCO (µs/ft)	DTSM (µs/ft)	GR (gAPI)	PHIT (m³/m³)	RHOB (g/cc)
6051.00	132.75	399.55	88.44	0.19	2.32

Fig. 7 Available well log data array
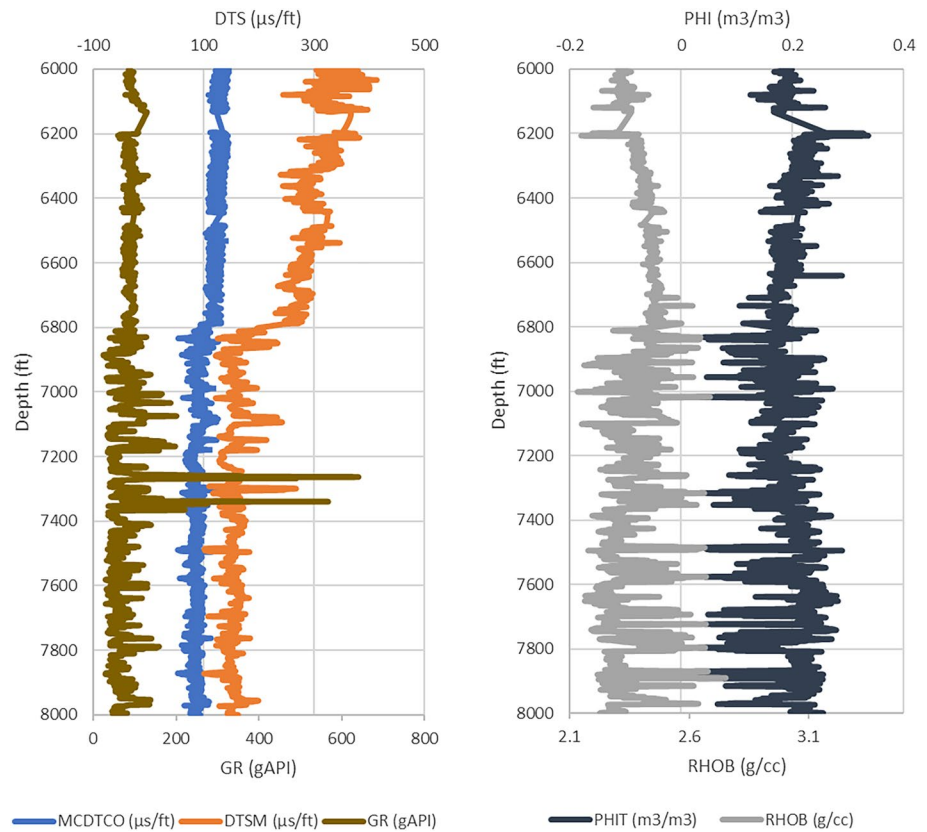
Although GP is not a new technique for predicting and representing data in the oil and gas industry, GP has not been applied to estimate or predict shear wave transit time with the array of well logs proposed in this study. The model proves that shear sonic travel time can be adequately represented with a Gaussian distribution; thus, GP can serve as a reliable tool in the prediction and reproduction of shear transit time from offset wells with no shear sonic logs for formation evaluation.

The model was developed using a reservoir located in the sand and shale sequence. Nonetheless, the methods could be used for any reservoir type, but the response would not be the same. This eludes to the limitations of any correlation, linear or multidimensional, where its applicability is useful within the range of physical data used to derive it.

Predicting dynamic geomechanical properties

This section illustrates the common uses of sonic logs in the evaluation of formation mechanical properties. To illustrate

these uses, the dynamic Young's modulus and Poisson's ratio are estimated from the measured sonic logs and compared the dynamic Young's modulus and Poisson's ratio estimated from the proposed model sonic log predictions. Rock elastic properties, particularly Young's modulus and Poisson's ratio, tell a lot about the formation because they are deformation properties (Ma et al. 2013). Poisson's ratio is used as a calibration tool in the industry to determine the accuracy of well logs (Oloruntobi et al. 2018; Onalo et al. 2018a, 2019). In most cases, if a sonic log model is able to predict Poisson's ratio accurately, then, it can be said that the model is robust and reliable (Onalo et al. 2018a).

Dynamic Young's modulus

Young's modulus is commonly known as the modulus of elasticity because it is a measure of the stiffness of the formation and can be estimated using Eq. (1) (Mullen et al. 2007).
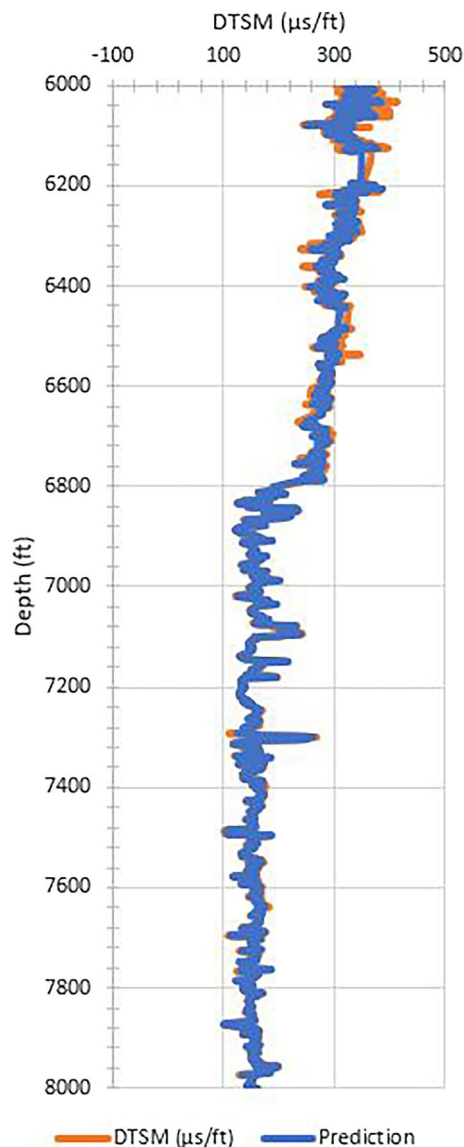
DTSM (μs/ft)

Depth (ft)

— DTSM (μs/ft) — Prediction

Fig. 8 Shear wave transit time versus depth

$$E = \frac{\rho}{\Delta t_s^2} \times \left(\frac{3\Delta t_s^2 - 4\Delta t_c^2}{\Delta t_s^2 - \Delta t_c^2} \right) \times 1.34 \times 10^{10} \qquad (1)$$

The results of the estimation of dynamic Young's modulus from the measured sonic logs and proposed model predicted sonic logs are presented and compared in Fig. 10. The cross-validation of the Young's modulus from the predicted and measured sonic logs presented in Fig. 10 shows very good agreement with a coefficient of determination of 0.9953. As the dynamic Young's modulus increases, the deviation from the perfect slop increases. The estimation from the proposed model sonic logs slightly underpredicts Young's modulus from approximately 40–70 GPa. The highest deviation is observed at a depth of 7480–7490 ft with the average dynamics Young's

modulus of 55 GPa and 50 GPa which both signify a good consolidation at such depths for measured and predicted, respectively.

Poisson's ratio (PR)

Poisson's ratio is another rock mechanical property that is estimated during formation evaluation. It is literally the ratio of the lateral to the vertical strain of a specimen and is estimated from sonic logs as follows (Mullen et al. 2007).

$$\vartheta = 0.5 \times \left(\frac{\Delta t_s^2 - 2\Delta t_c^2}{\Delta t_s^2 - \Delta t_c^2} \right) \qquad (2)$$

The results of the estimation of Poisson's ratio from the measured sonic logs and proposed model predicted sonic logs are presented and compared in Fig. 11. The cross-validation of the Poisson's ratio from the predicted and measured sonic logs presented in Fig. 11 portrays a good match with a coefficient of determination of 0.9413. As the Poisson's ratio increases, the deviation from the perfect slope decreases. The estimations from the proposed model overpredict points of Poisson's ratio values below 0.25. The accuracy of the estimations from the predicted model is increased as the formation weakens.

The main reason why the Young's modulus and Poisson's ratio predictions are reasonably accurate is because of the accuracy of the Shear transit time predictions which are then used in the theoretical and empirical relationships given in Eqs. 15 and 16. A major disadvantage of such model would be that poor predictions of the shear transit time would result in poor estimations of the geomechanical properties of the formation. Sample data of the measured and predicted geomechanical models are presented in Table 6. In general, both estimation of Young's modulus and Poisson's ratio from the measured and predicted sonic logs allude to a good agreement. Therefore, the model can be used in place of actual sonic logs with a high confidence level. Access to real-life data was one of the limiting factors in this research. To further improve this work, the authors recommend that well log data from different regions of the world in different formations can be used to develop a more robust model. Another advancement to this work would be to compare GP model to other models like artificial neural networks, recurrent neural networks and support vector machines with similar data from the same formation.

Summary and conclusions

The present study has demonstrated that in the absence of shear sonic transit logs, a GP model can be used to model the shear sonic logs from the depth, density, gamma ray,

Fig. 9 Predicted shear wave transit time versus measured shear wave transit time
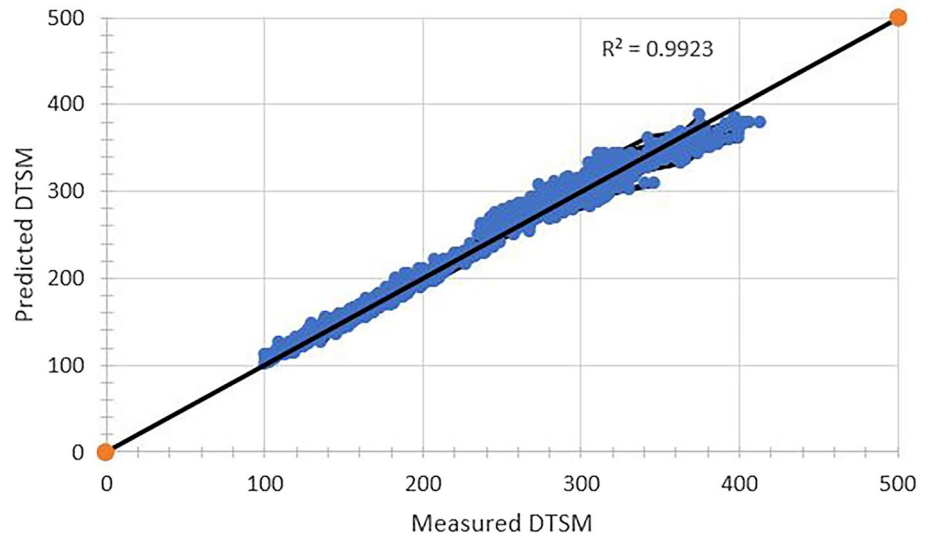

Fig. 10 Cross-validation plot of predicted and measured dynamic Young's modulus
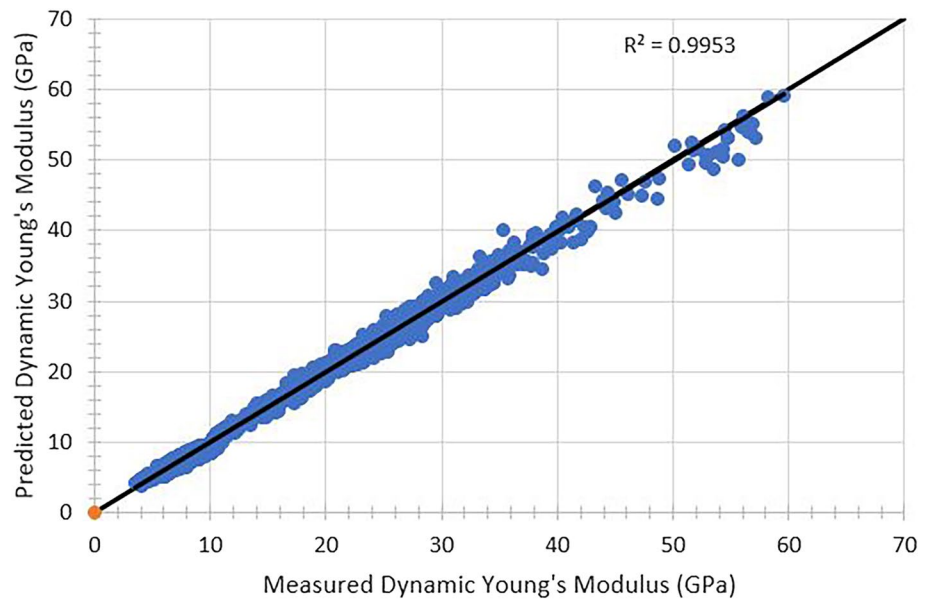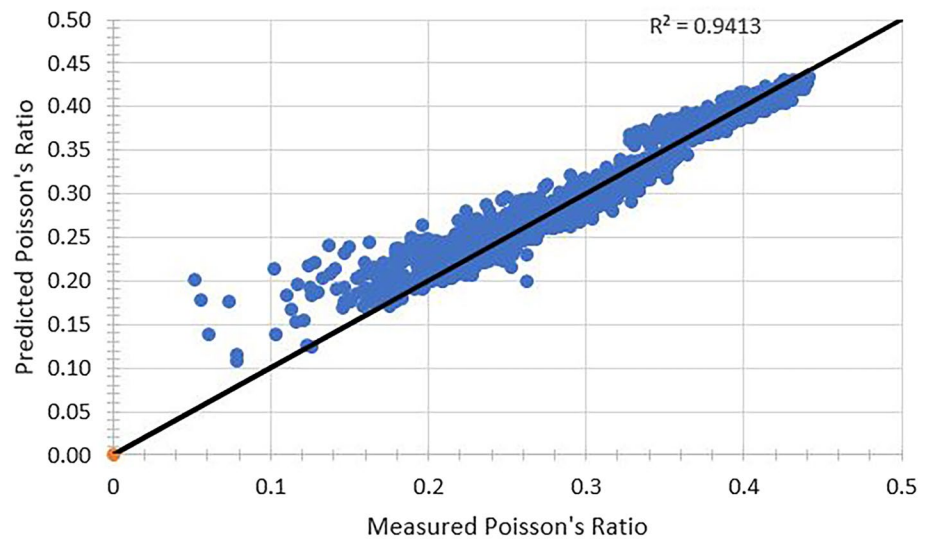

Fig. 11 Cross-validation plot of predicted and measured Poisson's ratio

porosity, and compressional transit time logs. The new model can be particularly useful in wells where accurate or continuous shear wave transit times are missing or where borehole-compensated sonic tool has not been used.

The proposed GP model development offers the following benefits to the oil and gas industry:

- The GP model offers operators with offset wells that only contain compressional sonic wells a reliable tool to predict the shear sonic log for better formation evaluation analysis.
- The GP model provides a cost-effective and safe tool to operators by offering a reliable means of predicting shear transit time in a field instead of carrying out more expensive dipole and multipole sonic logging on several wells in the field. This leads to cost savings and human (work hours) reduction leading to higher days without accidents (days since last accident or hazard exposure) on projects.
- The Gaussian model provides a cheap method of establishing mechanical rock formation property tables for several geographical regions and geological settings.
- The GP model provides a calibration and validation tool for cross-checking already measured or acquired sonic shear logs from sonic loggers that may be faulty or run in complicated hole sections.

The GP model accurately predicts shear sonic time log for the case study with an R^2 of 0.99. The model is also used to estimate some mechanical formation properties, namely Young's modulus and Poisson's ratio. The results are compared to the same mechanical rock properties using the measured sonic logs. The coefficients of determination between the measured and predicted sonic logs used for the estimations of Young's modulus and Poisson's ratio are 0.99 and 0.94, respectively.

Generally, the GP models are highly efficient in recognizing nonlinear patterns with the complex dataset including well logs used in the oil and gas industry as is evident in this study. GP models are recommended for developing nonparametric correlations between other well log dataset of interest.

The present study provides the oil and gas industry with a roadmap for estimating shear sonic well logs and also validating measured shear sonic transit time logs. Future work can be done to estimate both compressional and shear sonic transit logs from a Gaussian model, thereby eliminating the need to run countless expensive sonic logging tools in the formation. The significance of such a future model will be highly valuable in terms of cost savings and man-hours resources that could potentially be saved.

Appendix

Gaussian process theory

A description of the GP is presented below; however, a more detailed explanation can be found in Rasmussen (2004) and Williams and Rasmussen (2006).

Assume a set of data is provided in the following format (Kumar et al. 2014):

$$D = \left\{ x_i, y_i \right\}_{i=1}^n, x_i \in R^d \text{ and } y_i \in R \tag{3}$$

x_i = input data, y_i = output data, n = number of data points, R = 1-dimensional vector space, R^d = d-dimensional vector space. In this study, the input data are depth (ft), RHOB (g/cc), PHIT, GR (GAPI), and DTCO (μs/ft). The output is DTSM (μs/ft). Mathematically: x = [depth, RHOB, PHIT, GR, and DTCO] and y = [DSTM].

The GP can be defined by the mean ($m(x)$) and the covariance function ($k(x, x')$) for the function ($f(x)$) (Rostami and Khaksar Manshad 2013).

$$m(x) = E\big[f(x)\big] \tag{4}$$

$$k(x, x') = E\big[(f(x) - m(x))(f(x') - m(x'))\big], x \text{ and } x' \in R^d \tag{5}$$

Thus, the GP is written as follows (Rostami et al. 2013)

$$f(x) = GP\big(m(x), k(x, x')\big) \tag{6}$$

The GP regression is then expressed similarly to linear regression with the main function and Gaussian noise (ε) function as follows (Yu et al. 2016):

$$y = f(x) + \varepsilon \tag{7}$$

Table 6 Sensitivity analysis data (dynamic Young's modulus and Poisson's ratio)

Depth (ft)	Measured Ed (Gpa)	Predicted Ed (Gpa)	Measured PR	Measured PR
6000.00	5.71	5.55	0.40	0.40
6000.50	4.92	4.91	0.42	0.42
6001.00	4.58	4.76	0.42	0.42
6001.50	5.13	5.06	0.41	0.41
6002.00	6.10	5.69	0.39	0.40
6002.50	6.45	5.87	0.39	0.40
6003.00	6.44	5.54	0.39	0.41
6003.50	6.08	5.67	0.40	0.41
6004.00	5.33	5.62	0.41	0.41
6004.50	4.63	5.37	0.42	0.41
6005.00	4.24	4.94	0.43	0.42
6005.50	4.26	4.59	0.43	0.42
6006.00	4.38	4.70	0.43	0.42
6006.50	4.66	4.87	0.42	0.42
6007.00	4.78	5.16	0.42	0.41
6007.50	4.74	4.93	0.42	0.42
6008.00	4.52	4.62	0.43	0.42
6008.50	4.43	4.74	0.43	0.42
6009.00	4.33	4.52	0.43	0.43
6009.50	4.28	4.45	0.43	0.43
6010.00	4.97	5.01	0.41	0.41
6010.50	5.44	5.24	0.40	0.41
6011.00	5.01	4.91	0.41	0.41
6011.50	4.52	4.94	0.42	0.41
6012.00	4.63	4.99	0.42	0.41
6012.50	4.83	4.81	0.42	0.42
6013.00	4.66	4.54	0.42	0.43
6013.50	5.16	4.65	0.41	0.42
6014.00	5.94	5.30	0.40	0.41
6014.50	6.12	5.16	0.39	0.41
6015.00	6.28	5.46	0.39	0.41
6015.50	6.17	5.57	0.39	0.40
6016.00	5.78	5.66	0.40	0.40
6016.50	4.58	5.07	0.43	0.42
6017.00	4.09	4.67	0.43	0.42
6017.50	4.03	4.66	0.43	0.42
6018.00	4.37	4.96	0.43	0.41
6018.50	5.32	5.19	0.41	0.41
6019.00	5.70	5.42	0.40	0.41
6019.50	5.88	5.60	0.40	0.40
6020.00	5.87	5.56	0.40	0.41
6020.50	5.54	5.86	0.41	0.40
6021.00	4.54	5.41	0.43	0.41
6021.50	4.37	5.14	0.43	0.41
6022.00	4.53	5.16	0.43	0.41
6022.50	5.63	5.48	0.40	0.41
6023.00	5.99	5.44	0.40	0.41
6023.50	6.02	5.88	0.40	0.40

Table 6 (continued)

Depth (ft)	Measured Ed (Gpa)	Predicted Ed (Gpa)	Measured PR	Measured PR
6024.00	5.86	5.70	0.40	0.40
6024.50	5.75	5.78	0.40	0.40
6025.00	5.95	5.92	0.40	0.40
6025.50	5.79	5.56	0.40	0.41
6026.00	5.39	5.39	0.41	0.41
6026.50	5.17	5.32	0.41	0.41
6027.00	5.32	5.22	0.41	0.41
6027.50	5.43	5.43	0.41	0.41
6028.00	5.71	5.17	0.40	0.41
6028.50	5.83	5.12	0.40	0.41
6029.00	6.01	5.16	0.39	0.41
6029.50	6.06	5.27	0.39	0.41
6030.00	5.95	5.44	0.39	0.41
6030.50	5.71	5.37	0.40	0.41
6031.00	4.78	4.75	0.42	0.42
6031.50	3.92	4.25	0.44	0.43
6032.00	3.83	4.34	0.44	0.43
6032.50	3.81	4.30	0.44	0.43
6033.00	3.72	4.24	0.44	0.43
6033.50	3.57	4.22	0.44	0.43
6034.00	3.70	4.19	0.44	0.43
6034.50	3.81	4.08	0.43	0.43
6035.00	3.94	4.32	0.44	0.43
6035.50	4.06	4.52	0.44	0.43
6036.00	4.80	4.76	0.42	0.42
6036.50	5.23	5.43	0.41	0.41
6037.00	5.55	5.76	0.41	0.40
6037.50	6.36	6.56	0.40	0.39
6038.00	6.94	6.65	0.39	0.39
6038.50	7.31	7.25	0.39	0.39
6039.00	7.04	6.93	0.39	0.39
6039.50	6.34	6.15	0.39	0.40
6040.00	5.85	5.67	0.40	0.41
6040.50	4.63	5.22	0.43	0.42
6041.00	4.88	5.41	0.42	0.41
6041.50	6.04	5.56	0.40	0.41
6042.00	4.78	4.75	0.42	0.42
6042.50	5.74	5.14	0.40	0.41
6043.00	5.93	5.66	0.40	0.40
6043.50	6.25	6.02	0.40	0.40
6044.00	6.30	6.28	0.40	0.40
6044.50	6.48	6.31	0.40	0.40
6045.00	6.66	6.48	0.39	0.39
6045.50	6.42	6.35	0.39	0.39
6046.00	6.17	5.94	0.40	0.40
6046.50	5.86	5.95	0.40	0.40
6047.00	5.64	5.77	0.41	0.40
6047.50	4.94	5.25	0.42	0.42
6048.00	4.79	5.25	0.42	0.42

Table 6 (continued)

Depth (ft)	Measured Ed (Gpa)	Predicted Ed (Gpa)	Measured PR	Measured PR
6048.50	5.75	5.65	0.40	0.41
6049.00	5.73	5.58	0.40	0.40
6049.50	4.28	4.79	0.43	0.42
6050.00	4.10	4.68	0.43	0.42

The Gaussian noise has a mean of 0 and a variance of σ.

$$\varepsilon \sim N(0, \sigma) \tag{8}$$

For a new input dataset (x^*) and output dataset (y^*), the GP prior distribution is given as follows (Kumar et al. 2014):

$$\left(\frac{y}{y^*}\right) \sim N(0, k^*) \tag{9}$$

$$k^* = \begin{bmatrix} k & k(x, x^*) \\ k(x, x^*)^T & k(x^*, x^*) \end{bmatrix} \tag{10}$$

$k(x, x^*)$ = the covariance between the training input data and test input data; $k(x, x^*)^T$ = the transpose of $k(x, x^*)$; $k(x^*, x^*)$ = the covariance of the test data.

Thus, the mean and variance of the posterior Gaussian distribution of (y^*) can be written as follows, respectively (Yu et al. 2016):

$$m^* = k(x, x^*)^T k^{-1} y \tag{11}$$

$$\sigma^* = k(x^*, x^*) - k(x, x^*)^T k^{-1} y k(x, x^*) \tag{12}$$

Covariance and kernel Function

The covariance function can be defined by the kernel functions in order to provide better response across the dataset to which they are similar (Ebden 2008). A set of kernel functions or hyperparameters ($\theta = \{\sigma_f, \sigma_l\}$) parameterizes the covariance function. The kernel functions are needed to reduce the error and improve the accuracy by smoothing the dataset predictions. The dependency of the covariance function is written as $k(x, x'|\theta)$. Most problems can be presented as GP distribution; however, accuracy and efficiency are improved by the kernel and hyperparameter functions. Therefore, to ensure an adequate model is attributed to a problem, the most suitable kernel function that describes the nonlinear relationship should be chosen. Although Ebden (2008) suggests that squared exponential kernel is the popular choice, in this study the following set

of kernel functions are explored to provide a justification for the model selection (Matlab Documentation 2018).

Exponential kernel:

$$k(x_i, x_j|\theta) = \sigma_f^2 \exp\left[-\frac{(x_i - x_j)^T (x_i - x_j)}{\sigma_l}\right] \tag{13}$$

σ_f = the standard deviation, σ_l = the characteristic length and a = positive scale mixture parameter

Squared exponential kernel:

$$k(x_i, x_j|\theta) = \sigma_f^2 \exp\left[-\frac{1}{2}\left(\frac{(x_i - x_j)^T (x_i - x_j)}{\sigma_l^2}\right)\right] \tag{14}$$

Matern 5/2 kernel:

$$k(x_i, x_j|\theta) = \sigma_f^2 \left(1 + \frac{\sqrt{5}(x_i - x_j)^T (x_i - x_j)}{\sigma_l} \right.$$
$$\left. + \left(\frac{\sqrt{5}(x_i - x_j)^T (x_i - x_j)}{\sqrt{3}\sigma_l}\right)^2\right)$$
$$\times \exp\left[-\left(\frac{\sqrt{5}(x_i - x_j)^T (x_i - x_j)}{\sigma_l}\right)\right] \tag{15}$$

Rational quadratic kernel:

$$k(x_i, x_j|\theta) = \sigma_f^2 \left[1 + \left(\frac{(x_i - x_j)^T (x_i - x_j)}{2a\sigma_l^2}\right)^{-a}\right] \tag{16}$$

The process of finding the most suitable values of the hyperparameters is called the GP learning that illustrates how the GP trains the model to define the problem with the least errors (Huang et al. 2017). The GP is developed using MATLAB simulation software which initializes and finds the hyperparameters that minimize cross-validation loss by using automatic hyperparameter optimization (Matlab Documentation 2018).

References

Abdollahzadeh A, Christie MA, Corne D (2012) Gaussian-based estimation of distribution algorithms for history matching. In: Abu Dhabi international petroleum conference and exhibition. https://doi.org/10.2118/161951-MS

Adedigba SA, Khan F, Yang M (2017) Dynamic failure analysis of process systems using neural networks. Process Saf Environ Prot 111:529–543. https://doi.org/10.1016/J.PSEP.2017.08.005

Adeli A, Emery X, Dowd P (2017) Geological Modelling and Validation of Geological Interpretations via Simulation and

Classification of Quantitative Covariates. Minerals. https://doi.org/10.3390/min8010007

Akin S, Ross CM, Kovscek AR (2008) Combination of well log and pore-scale data to predict petrophysical properties of diatomite. J Pet Sci Eng 60:133–149. https://doi.org/10.1016/J.PETROL.2007.05.013

Alford J, Blyth M, Tollefsen E, Crowe J, Loreto J, Mohammed S, Pistre V, Rodriguez-Herrera A (2012) Sonic logging while drilling-shear answers. Oilf Rev 24:4–15

Ali Ahmadi M, Golshadi M (2012) Neural network based swarm concept for prediction asphaltene precipitation due to natural depletion. J Pet Sci Eng 98–99:40–49. https://doi.org/10.1016/J.PETROL.2012.08.011

Asadisaghandi J, Tahmasebi P (2011) Comparative evaluation of back-propagation neural network learning algorithms and empirical correlations for prediction of oil PVT properties in Iran oilfields. J Pet Sci Eng 78:464–475. https://doi.org/10.1016/J.PETROL.2011.06.024

Ashoori S, Abedini A, Abedini R, Nasheghi KQ (2010) Comparison of scaling equation with neural network model for prediction of asphaltene precipitation. J Pet Sci Eng 72:186–194. https://doi.org/10.1016/J.PETROL.2010.03.016

Babakhani SM, Bahmani M, Shariati J, Badr K, Balouchi Y (2015) Comparing the capability of artificial neural network (ANN) and CSMHYD program for predicting of hydrate formation pressure in binary mixtures. J Pet Sci Eng 136:78–87. https://doi.org/10.1016/J.PETROL.2015.11.002

Bailey T (2012) An empirical Vp/Vs shale trend for the Kimmeridge Clay of the Central North Sea. In: 74th EAGE conference & exhibition incorporating SPE EUROPEC 2012. Copenhagen, Denmark, pp 4–7

Castagna JP, Batzle ML, Eastwood RL (1985) Relationships between compressional-wave and shear-wave velocities in clastic silicate rocks. Geophysics 50:571–581. https://doi.org/10.1190/1.1441933

Dakhelpour-Ghoveifel J, Shegeftfard M, Dejam M (2019) Capillary-based method for rock typing in transition zone of carbonate reservoirs. J Pet Explor Prod Technol 9:1–10

Derakhshanfard F, Mehralizadeh A (2018) Application of artificial neural networks for viscosity of crude oil-based nanofluids containing oxides nanoparticles. J Pet Sci Eng 168:263–272. https://doi.org/10.1016/J.PETROL.2018.05.018

Dey AK, Stewart RR (1997) Predicting density using Vs and Gardner's relationship. CREWES Res Rep 9:1–9

Doh CA, Alger RP (1958) Sonic logging, a new petrophysical tool

Domenico SN (1984) Rock lithology and porosity determination from shear and compressional wave velocity. Geophysics 49:1188–1195. https://doi.org/10.1190/1.1441748

Ebden M (2008) Gaussian processes for regression: a quick introduction. Website Robot Res Gr Dep Eng Sci Univ, Oxford

Eberhart-Phillips D, Han D-H, Zoback MD (1989) Empirical relationships among seismic velocity, effective pressure, porosity, and clay content in sandstone. Geophysics 10(1190/1):1442580

Esene C, Onalo D, Zendehboudi S, James L, Aborig A, Butt S (2018) Modeling investigation of low salinity water injection in sandstones and carbonates: effect of Na+ and SO42−. Fuel 232:362–373. https://doi.org/10.1016/j.fuel.2018.05.161

Gardner GHF, Gardner LW, Gregory AR (1974) Formation velocity and density—the diagnostic basics for stratigraphic traps. Geophysics 39:770–780. https://doi.org/10.1190/1.1440465

Greenberg ML, Castagna JP (1992) Shear-wave velocity estimation in porous rocks: theoretical formulation, preliminary verification and applications1. Geophys Prospect 40:195–209. https://doi.org/10.1111/j.1365-2478.1992.tb00371.x

Hamada GM (2004) Reservoir fluids identification using Vp/Vs ratio. Oil Gas Sci Technol Rev IFP 59:649–654

Han I, Kamber M (2010) Data mining: concepts and techniques. Morgan Kaufinann, Burlington

Han D, Nur A, Morgan D (1986) Effects of porosity and clay content on wave velocities in sandstones. Geophysics 51:2093–2107. https://doi.org/10.1190/1.1442062

Harrison AR, Randall CJ, Aron JB, Morris CF, Wignall AH, Dworak RA, Rutledge LL, Perkins JL (1990) Acquisition and analysis of sonic waveforms from a borehole monopole and dipole source for the determination of compressional and shear speeds and their relation to rock mechanical properties and surface seismic data. In: SPE annual technical conference and exhibition. society of petroleum engineers. https://doi.org/10.2118/20557-MS

Hsu K, Brie A, Plumb RA (1987) A new method for fracture identification using array sonic tools. J Pet Technol. https://doi.org/10.2118/14397-PA

Huang YF, Huang GH, Dong MZ, Feng GM (2003) Development of an artificial neural network model for predicting minimum miscibility pressure in CO2 flooding. J Pet Sci Eng 37:83–95. https://doi.org/10.1016/S0920-4105(02)00312-1

Huang XB, Zhang Q, Zhu HH, Zhang LY (2017) An estimated method of intact rock strength using gaussian process regression. In: 51st US rock mechanics/geomechanics symposium

Iglesias MA, Law KJH, Stuart AM (2013) Evaluation of Gaussian approximations for data assimilation in reservoir models. Comput Geosci 17:851–885. https://doi.org/10.1007/s10596-013-9359-x

Iturrarán-Viveros U, Molero M (2013) Simulation of sonic waves along a borehole in a heterogeneous formation: accelerating 2.5-D finite differences using [Py]OpenCL. Comput Geosci 56:161–169. https://doi.org/10.1016/J.CAGEO.2013.03.014

Izadi G, Elsworth D (2014) Reservoir stimulation and induced seismicity: roles of fluid pressure and thermal transients on reactivated fractured networks. Geothermics 51:368–379. https://doi.org/10.1016/J.GEOTHERMICS.2014.01.014

Jorstad A, Mukerji T, Mavko G (1999) Model-based shear-wave velocity estimation versus empirical regressions. Geophys Prospect 47:785–797. https://doi.org/10.1046/j.1365-2478.1999.00154.x

Kelechukwu EM, Said Al-Salim H, Saadi A (2013) Prediction of wax deposition problems of hydrocarbon production system. J Pet Sci Eng 108:128–136. https://doi.org/10.1016/j.petrol.2012.11.008

Khazanehdari J, Mccann C (2005) Acoustic and petrophysical relationships in low-shale sandstone reservoir rocks. Geophys Prospect 53(4):447–461

Kokesh FP, Schwartz RJ, Wall WB, Morris RL (1965) A new approach to sonic logging and other acoustic measurements. J pet Technol. https://doi.org/10.2118/991-PA

Krief M, Garat J, Stellingwerf J, Ventre J (1990) A petrophysical interpretation using the velocities of P and S waves (full-waveform Sonic). Log Anal 31:355–369

Kumar M, Bhatt MR, Samui P (2014) Modeling of elastic modulus of jointed rock mass: Gaussian process regression approach. Int J Geomech 14:06014001. https://doi.org/10.1061/(ASCE)GM.1943-5622.0000318

Kuss M, Rasmussen CE (2006) Assessing approximations for Gaussian process classification. In: Advances in neural information processing systems 18 proceedings of 2005 conference, pp 699–706

Lee MW (2006) A simple method of predicting S-wave velocity. Geophysics 71:F161–F164. https://doi.org/10.1190/1.2357833

Ma L, Liu X, Wang M, Xu H, Hua R, Fan P, Jiang S, Wang G, Yi Q (2013) Experimental investigation of the mechanical properties of rock salt under triaxial cyclic loading. Int J Rock Mech Min Sci 62:34–41

MacKay DJC (2005) Information theory, inference, and learning algorithms david J.C. MacKay, Learning. https://doi.org/10.1198/jasa.2005.s54

Market J, Canady WJ (2006) Dispersion corrections are not just for LWD dipole sonic tools. In: SPE Annual Technical Conference and Exhibition. https://doi.org/10.2118/102345-MS

Matlab Documentation (2018) The MathWorks Inc

Miller SLM, Stewart RR (1974) The relationship between elastic-wave velocities and density in sedimentary rocks: a proposal. Crewes Res Rep:260–273

Miller S, Stewart R (1990) Effects of lithology, porosity and shaliness on the P and S-wave velocities from sonic logs. J Can Soc Explor Geophys 26:94–103

Minear JW, Fletcher CR (1983) Full-wave acoustic logging. In: SPWLA 24th annual logging symposium

Mohammed RO, Cawley GC (2017) Over-fitting in model selection with gaussian process regression. In: Petrou M, Perner P (eds) Machine learning and data mining in pattern recognition. Springer, Cham, pp 192–205. https://doi.org/10.1007/978-3-319-62416-7_14

Mullen MJ, Roundtree R, Turk GA (2007) A composite determination of mechanical rock properties for stimulation design (What to do when you don't have a sonic log). In: Rocky mountain oil & gas technology symposium. https://doi.org/10.2118/108139-MS

Nourafkan A, Kadkhodaie-Ilkhchi A (2015) Shear wave velocity estimation from conventional well log data by using a hybrid ant colony-fuzzy inference system: a case study from Cheshmeh-Khosh oilfield. J Pet Sci Eng 127:459–468. https://doi.org/10.1016/j.petrol.2015.02.001

Oloruntobi O, Butt S (2019) The shear-wave velocity prediction for sedimentary rocks. J Nat Gas Sci, Eng, p 103084

Oloruntobi O, Adedigba S, Khan F, Chunduru R, Butt S (2018) Overpressure prediction using the hydro-rotary specific energy concept. J Nat Gas Sci Eng 55:243–253

Oloruntobi O, Onalo D, Adedigba S, James L, Chunduru R, Butt S (2019) Data-driven shear wave velocity prediction model for siliciclastic rocks. J Pet Sci Eng 183:106293

Onalo D, Adedigba S, Khan F, James LA, Butt SD (2018a) Data Driven model for sonic well log prediction. J Pet Sci Eng 170:1022–1037. https://doi.org/10.1016/j.petrol.2018.06.072

Onalo D, Oloruntobi O, Adedigba S, Khan F, James L, Butt S (2018b) Static Young's modulus model for drilling operation planning. J Pet Sci Eng 171:394–402

Onalo D, Oloruntobi O, Adedigba S, Khan F, James L, Butt S (2019) Dynamic data driven sonic well log model for formation evaluation. J Pet Sci, Eng

Ramcharitar K, Hosein R (2016) Rock mechanical properties of shallow unconsolidated sandstone. In: SPE trinidad and tobago section energy resources conference https://doi.org/10.2118/180803-MS

Rasmussen CE (2004) Gaussian processes in machine learning. In: Advanced lectures on machine learning. Springer, pp 63–71

Rasmussen CE, Williams CKI (2006) Model selection and adaptation of hyperparameters. Gaussian Process. Mach. Learn. (Adaptive Comput. Mach. Learn. Ser. 105–128. https://doi.org/10.1142/S0129065704001899

Raymer LLL, Hunt ERR, Gardner JS (1980) An improved sonic transit time to porosity-to-porosity transform. In: 21st SPWLA logging symp. trans, pp 1–13

Reichel N, Evans M, Allioli F, Mauborgne M-L, Nicoletti L, Haranger F, Stoller C, Schlumberger VC, El E, Sipetrol H (2012) Neutron-gamma density (Ngd): principles, field test results and log quality control of a radioisotope-free bulk density measurement. In: SPWLA 53rd annual logging symposium. society of petrophysicists and well-log analysts, pp 1–15

Riazi SH, Heydari H, Ahmadpour E, Gholami A, Parvizi S (2014) Development of novel correlation for prediction of hydrate formation temperature based on intelligent optimization algorithms. J Nat Gas Sci Eng 18:377–384. https://doi.org/10.1016/J.JNGSE.2014.03.012

Rostami H, Khaksar Manshad A (2013) Prediction of asphaltene precipitation in live and tank crude oil using gaussian process regression. Pet Sci Technol 31:913–922. https://doi.org/10.1080/10916466.2010.531349

Rostami H, Azin R, Dianat R (2013) Prediction of undersaturated crude oil density using gaussian process regression. Pet Sci Technol 31:418–427. https://doi.org/10.1080/10916466.2010.531346

Saboorian-Jooybari H, Dejam M, Chen ZJ, Pourafshary P (2015) Fracture identification and comprehensive evaluation of the parameters by dual laterolog data. In: SPE middle east unconventional resources conference and exhibition. Society of petroleum engineers

Saboorian-Jooybari H, Dejam M, Chen Z, Pourafshary P (2016) Comprehensive evaluation of fracture parameters by dual laterolog data. J Appl Geophys 131:214–221

Seeger M (2004) Gaussian processes for machine learning. Int J Neural Syst 14:69–106

Sheremetov L, Cosultchi A, Martínez-Muñoz J, Gonzalez-Sánchez A, Jiménez-Aquino MA (2014) Data-driven forecasting of naturally fractured reservoirs based on nonlinear autoregressive neural networks with exogenous input. J Pet Sci Eng 123:106–119. https://doi.org/10.1016/J.PETROL.2014.07.013

Silva DSF, Deutsch CV (2016) Multivariate data imputation using Gaussian mixture models. Spat Stat 27:74–90. https://doi.org/10.1016/J.SPASTA.2016.11.002

Su G (2009) Modeling non-linear deformation time series of tunnel using Gaussian process machine learning. In: ISRM international symposium on rock mechanics-SINOROCK 2009

Takahashi I, Mukerji T, Mavko G (2000) Vp–Vs relations of sandstones and carbonates: implications about the pore structure. In: SEG technical program expanded abstracts 2000. Society of exploration geophysicists, pp 1838–1841. https://doi.org/10.1190/1.1815786

Thomsen L (1986) Weak elastic anisotropy. Geophysics 51:1954–1966. https://doi.org/10.1190/1.1442051

Toksöz MN, Cheng CH, Timur A (1976) Velocities of seismic waves in porous rocks. Geophysics 41:621–645. https://doi.org/10.1190/1.1440639

Vaferi B, Gitifar V, Darvishi P, Mowla D (2014) Modeling and analysis of effective thermal conductivity of sandstone at high pressure and temperature using optimal artificial neural networks. J Pet Sci Eng 119:69–78. https://doi.org/10.1016/J.PETROL.2014.04.013

Vernik L, Fisher D, Bahret S (2002) Estimation of net-to-gross from P and S impedance in deep-water turbidites. Lead Edge 21:380–387. https://doi.org/10.1190/1.1471602

Williams DM (1990) The acoustic log hydrocarbon indicator. Soc. Petrophysicists Well-Log Anal

Williams CKI, Rasmussen CE (2006) Gaussian processes for machine learning, vol 2. MIT Press, Cambridge, p 4

Yu H, Wang Z, Rezaee R, Zhang Y, Xiao L, Luo X, Wang X, Zhang L (2016) The gaussian process regression for TOC Estimation using wireline logs in shale gas reservoirs. Pet Technol Conf, Int. https://doi.org/10.2523/IPTC-18636-MS