Article

Research on Formation Pressure Prediction Method for Ultra-Deep Tight Sandstone Based on Collocated Cokriging

Qiang Wei [1], Yaoting Lin [2,*], Gang Gao [1], Zhixian Gui [1], Zhendong Wu [3] and Jiaqi Liu [1]

[1]   Key Laboratory of Exploration Technologies for Oil and Gas Resources, Yangtze University, Ministry of Education, Wuhan 430100, China
[2]   Department of Mathematics and Statistics, University of Calgary, 2500 University Dr. NW, Calgary, AB T2N 1N4, Canada
[3]   Dagang Branch, GRI, BGP Inc., CNPC, Tianjin 300280, China
[*]   Correspondence: yaoting.lin@ucalgary.ca

Abstract: Compared to conventional reservoirs, the prediction of pressure in ultra-deep tight sandstone formations is difficult. The prediction of seismic pressure is more challenging than well-logging pressure prediction. The main methods for seismic pressure prediction include the equivalent depth method, Eaton method, Fillippone formula, and modified versions. Among them, the Eaton method is widely used and has good effectiveness. However, this method relies on difficult-to-obtain normal compaction trend lines, which leads to low prediction accuracy in space. To address this issue, a method combining the Eaton method and collocated cokriging is proposed. Herein, the Eaton formula is used to predict formation pressure at the well, with compressional wave velocity as the covariate for predicting the main variable—formation pressure. By simulating the shear wave velocity based on seismic compressional wave velocity, the influence of various parameters on the prediction results is analyzed, and the accuracy of this method is verified by comparing it with other methods. The proposed method is then applied to predict formation pressure in the ultra-deep formations of the Junggar Basin. The simulation results show that the collocated cokriging method achieves higher planar accuracy and better matches the experimental expectations in terms of prediction results. The application results also demonstrate the scientific effectiveness of the combined method, which has achieved good results in practical production applications.

1. Introduction

In the process of petroleum exploration and development, accurate prediction of formation pressure has always been an extremely important aspect. In the petroleum industry, there are many methods used to predict formation pressure, which can be broadly divided into two categories: pre-drilling prediction based on seismic exploration data and post-drilling detection using well-logging data.

Currently, in China, the seismic layer velocity obtained from seismic exploration and the logging data obtained after drilling are commonly used as the basis for establishing a system that can quickly understand the geological environment at the drilling site [1]. However, research on how to predict formation pressure more accurately has been conducted both domestically and internationally. In terms of well logging, Zhao et al. [2] improved the prediction accuracy by using cluster analysis to separate shale from other lithologies and obtain normal trend line velocities. Xie et al. [3] effectively calculated formation pore pressure by combining dipole sonic and other conventional well-logging data, achieving good application results. Li et al. [4] found that the Dc index method and the time difference of sonic waves are suitable for calculating formation pressure

in Block A of the Bohai Sea, with a maximum error of 5%, which meets engineering requirements and is of practical significance for on-site drilling operations. Fan et al. [5] proposed and modified a formation pressure prediction model based on the Eaton method, taking into account the geological influences of the actual area. They established three pressure profiles for the formations in that area, providing important support for subsequent drilling. In terms of seismic exploration, pre-drilling pressure prediction mainly involves calculating formation pore pressure using seismic layer velocity data. Wang et al. [6] introduced the effective stress principle in porous media to develop a new method for calculating pore fluid pressure in unconventional reservoirs, achieving good application results. Zhang et al. [7] used the DIX formula to obtain layer velocities based on velocity spectrum data and predicted formation pressure using seismic layer velocities. Yang et al. [8] combined seismic and well-logging data to establish pressure prediction profiles for deep formations and conducted in-depth analysis of the drilling geological conditions in the study area. Ma [9] used an improved Fillippone method to predict formation pressure ahead of the drill bit, and the results showed that the method is scientifically effective and meets the requirements of actual field applications. Qian et al. [10] applied the equivalent medium theory to determine the upper and lower limits of rock velocities in the Fillippone formula, obtaining more reasonable estimation results.

So far, the methods for predicting formation pressure based on seismic data can be broadly classified into two categories. One category relies on seismic velocity spectrum data and uses the Fillippone formula and its modified versions to predict formation pressure. The advantage is the ability to establish the spatial distribution of formation pressure in the underground three-dimensional space. However, this method relies on seismic layer velocities, which are often difficult to obtain accurately. Moreover, when predicting formation pressure, it requires the establishment of an empirical equation between velocity and pressure, resulting in insufficient prediction accuracy of formation pressure in the entire three-dimensional space. The other category involves predicting formation pressure using seismic data under well constraints, including the equivalent depth method and the Eaton formula [11].

In recent years, the introduction of geostatistical stochastic simulation methods has provided effective tools for reservoir prediction. In particular, the collocated cokriging has strong advantages in integrating well-logging information and effectively utilizing precise well location data. Gao [12] used three different kriging methods to interpolate the sandstone formation in Yuncheng City. The fitting effect of the exponential model in the variogram fitting was the best, and the results showed that the ordinary kriging method could more accurately reflect the spatial distribution of sandstone formation elevation in Yuncheng. Ma et al. [13] combined functional data analysis and kriging interpolation techniques to improve the accuracy and reliability of oil and gas well productivity prediction. Li et al. [14] addressed the "banding effect" that occurs during kriging interpolation by using distance constraints to correct the weights obtained from kriging estimation, making the kriging interpolation method more stable.

However, whether it is simple kriging or ordinary kriging, the interpolation effect decreases rapidly when the data points are sparse, and the interpolated results do not meet expectations. Cokriging technology has wider applications. Under the constraint of secondary data that has a certain correlation with the primary data, it conducts collaborative simulation on a small amount of sparse and irregularly distributed primary data, and the simulation results are similar to the spatial distribution pattern of the secondary data. Wang et al. [15] used cokriging interpolation to predict the spatial distribution trend of low-pressure pipeline networks in suburban and rural areas. The results showed that the prediction model performed well, with average error and root mean square error within the acceptable range. Du et al. [16] used cokriging to predict coal seam thickness, and the results showed that the method was effective, with small errors and improved accuracy. Geng et al. [17] applied cokriging to a three-dimensional inversion of gravity gradient tensor data, reducing the ambiguity of inversion and improving the resolution of inversion

results. Yu et al. [18] proposed a cokriging porosity prediction method under facies control based on cokriging. Through comparative analysis, the estimated results had smaller errors and the predicted results were more realistic.

Cokriging is relatively complex in computation, with a large amount of calculation, and the actual values of primary data at the same location cannot be consistent, which limits the development environment of cokriging. Collocated cokriging technique simplifies the complexity of cokriging in computation and matches well with the known data points and actual parameters. Chen et al. [19] integrated seismic, well-logging, geological, and other information to predict the distribution and variation of channel sand bodies using collocated cokriging, achieving significant improvement in prediction accuracy compared to traditional methods. Wang et al. [20] improved the prediction accuracy of seismic inversion by combining collocated cokriging to effectively utilize information from horizontal sections. Zhang et al. [21] applied sequential Gaussian collocated cokriging to predict reservoirs by incorporating other seismic parameter data, demonstrating high prediction accuracy and effective reduction in drilling risks. Niu et al. [22] estimated the variogram function by filtering the expected values of covariates using primary variable observations, and proposed and derived a new collocated cokriging method.

2. Methods

2.1. Principle of Formation Pressure Prediction

The Eaton method predicts formation pressure by establishing a normal compaction trend line. The calculation formula is as follows:

$$P_p = P_{ob} - (P_{ob} - P_h)\left(\frac{V}{V_n}\right)^N \tag{1}$$

In the above, P_p represents formation pressure; P_{ob} represents overburden pressure; P_h represents normal hydrostatic pressure; V represents actual formation velocity, which is obtained from sonic log data or stacking velocity; V_n represents the normal trend line velocity, mainly obtained by fitting the velocity of mudstone; and N represents the Eaton index, which is a coefficient related to the formation. The value of the Eaton index varies for different geological periods and regions.

In the formula, P_h hydrostatic pressure increases with the increase of depth; at the same depth, its value increases with the increase in formation water density:

$$P_h = \int_0^{D_0} \rho(z) g \, dD \tag{2}$$

In the formula, P_{ob} overburden pressure is an important factor in generating underground pressure. Its driving force is mainly a combination of sedimentary and compaction effects of the formation. This pressure is a fundamental parameter in the process of predicting formation pressure. In the detection of formation pressure, its value is determined first.

$$G_{oi} = \frac{\rho_\omega h_\omega + \rho_0 h_0 + \sum_{i=1}^{n} \rho_{bi} \Delta h}{h_\omega + h_0 + \sum_{i=0}^{n} \Delta h_i} \tag{3}$$

In this formula, G_{oi} represents the gradient of overburden pressure at a certain depth; ρ_ω represents seawater density; h_ω represents seawater depth; ρ_0 represents the average density of the upper density-free logging formation section; h_0 represents the thickness of the upper density-free logging formation section; ρ_{bi} represents density scatter data at a certain depth; and h_i represents depth interval. The calculation of overburden pressure involves multiple data and is a tedious process.

2.2. Cokriging Principle

When there are two or more characteristic parameters in the interpolation area and there is a significant correlation between the main variable and covariates to be interpolated within that area, the cokriging method can be used. This method belongs to multivariate geostatistics and involves analyzing multiple parameters in the study area, studying the linear and nonlinear relationships between these parameters, and understanding the spatial differences of different parameters to achieve a certain level of precision in estimating the main variable.

The cokriging estimation is formulated as follows:

$$Z^*(\mathbf{u}_0) = \sum_{i=1}^{n} \lambda_{z_i} Z(\mathbf{u}_{z_i}) + \sum_{k=1}^{m} \lambda_{y_k} Y(\mathbf{u}_{y_k}) \tag{4}$$

In this equation, $Z^*(\mathbf{u}_0)$ represents the estimated value at the estimation point $\mathbf{u}_0$; $Z(\mathbf{u}_{z_i})$ represents the actual attribute value of the main variable at point $\mathbf{u}_{z_i}$; $Y(\mathbf{u}_{y_k})$ represents the actual attribute value of the covariate at point $\mathbf{u}_{y_k}$; and λ_{z_i} and λ_{y_k} are the weight coefficients corresponding to the main variable and covariate, respectively.

Cokriging estimation is defined as a linear combination of available samples. Similar to ordinary Kriging, cokriging estimation requires unbiasedness and minimum error variance.

$$E[Z(\mathbf{u}_0) - Z^*(\mathbf{u}_0)] = 0 \tag{5}$$

$$Var[Z(\mathbf{u}_0) - Z^*(\mathbf{u}_0)] = \min \tag{6}$$

By incorporating the minimum variance condition with weight constraints, each considered random variable introduces Lagrange multipliers during the minimization process. By taking partial derivatives of each weight λ_{z_i}, λ_{x_j}, and λ_{y_k} and setting the results to zero, the minimum variance can be determined. After expanding and processing the variance, derivative calculations yield the cokriging equations:

$$\sum_{i=1}^{n} \lambda_{z_i} C_{zz}(Z(\mathbf{u}_{z_i}), Z(\mathbf{u}_I)) + + \sum_{k=1}^{m} \lambda_{y_k} C_{zy}(Y(\mathbf{u}_{y_k}), Z(\mathbf{u}_I))$$
$$+\xi_1 = C_{zz}(Z(\mathbf{u}_0) - Z(\mathbf{u}_I)) \quad I = 1, 2, \ldots n \tag{7}$$

$$\sum_{i=1}^{n} \lambda_{z_i} C_{zy}(Z(\mathbf{u}_{x_i}), Y(\mathbf{u}_K)) + \sum_{k=1}^{m} \lambda_{y_k} C_{yy}(Y(\mathbf{u}_{y_k}), Y(\mathbf{u}_K))$$
$$+\xi_2 = C_{zy}(Z(\mathbf{u}_0) - Y(\mathbf{u}_K)) \quad K = 1, 2, \ldots m \tag{8}$$

$$\sum_{i=1}^{n} \lambda_{z_i} = 1; \sum_{k=1}^{m} \lambda_{y_k} = 0 \tag{9}$$

In the equations, C_{zz} represents the autocovariance of the main variable; C_{yy} represents the autocovariance of the covariate; C_{zy} represents the cross-covariance between the main variable and covariate; and ξ_1 and ξ_2 are Lagrange factors.

The covariance function can be obtained from the variogram function, but calculating the cross-covariance or cross-variogram requires a significant amount of computation and involves complex derivation, which severely reduces the interpolation efficiency of cokriging. Even if the cross-variogram function is obtained, solving the equations can lead to a singular, resulting in situations where the estimation point has no solution.

2.3. Collocated Cokriging Principle

Collocated cokriging is a simplified form of cokriging that greatly reduces the computational burden on the equation system. In cokriging, a considerable number of covariates need to be selected for calculation. However, collocated cokriging only requires the covariates at the same positions as the estimation point. The covariates around the estimation

point will be masked by the covariance in the same position. This also requires that there are corresponding covariates for each estimation location.

In collocated cokriging, only three functions are needed: the autocovariance function of the main variable, the cross-covariance function between the main variable and covariates. The latter can be derived from the Markov model, significantly improving the computational speed of collocated cokriging.

The estimation value in collocated cokriging is given by the equation:

$$Z^*(\mathbf{u}_0) = \sum_{i=1}^{n} \lambda_{x_i} Z(\mathbf{u}_{x_i}) + \beta_1 X(\mathbf{u}_0) \tag{10}$$

In this equation, $Z^*(\mathbf{u}_0)$ represents the estimated value at the estimation point $\mathbf{u}_0$ and $Z(\mathbf{u}_{z_i})$ represents the actual attribute value of the main variable at point $\mathbf{u}_{z_i}$. Since there is only one secondary datum used for calculation, the weight for the secondary data has only one value.

Then, by incorporating the condition of unbiasedness and minimum error variance:

$$E[Z(\mathbf{u}_0) - Z^*(\mathbf{u}_0)] = 0 \tag{11}$$

$$Var[Z(\mathbf{u}_0) - Z^*(\mathbf{u}_0)] = \min \tag{12}$$

A series of mathematical operations leads to the collocated cokriging equation matrix:

$$\begin{bmatrix} C_{zz}(\mathbf{u}_{x_1}, \mathbf{u}_{x_1}) & \cdots & C_{zz}(\mathbf{u}_{x_1}, \mathbf{u}_{x_n}) & C_{zx}(\mathbf{u}_{x_1}, \mathbf{u}_0) & 1 \\ \vdots & & \vdots & \vdots & 1 \\ C_{zz}(\mathbf{u}_{x_n}, \mathbf{u}_{x_1}) & \cdots & C_{zz}(\mathbf{u}_{x_n}, \mathbf{u}_{x_n}) & C_{zx}(\mathbf{u}_{x_n}, \mathbf{u}_0) & 1 \\ C_{xz}(\mathbf{u}_{u_0}, \mathbf{u}_{x_1}) & \cdots & C_{xz}(\mathbf{u}_0, \mathbf{u}_{x_n}) & C_{xx}(\mathbf{u}_0, \mathbf{u}_0) & 1 \\ 1 & \cdots & 1 & 1 & 0 \end{bmatrix} \times \begin{bmatrix} \lambda_{x_1} \\ \vdots \\ \lambda_{x_n} \\ \beta_1 \\ \mu \end{bmatrix} = \begin{bmatrix} C_{zz}(\mu_0, \mu_{x_1}) \\ \vdots \\ C_{zz}(\mu_0, \mu_{x_n}) \\ C_{zx}(\mu_0, \mu_0) \\ 1 \end{bmatrix} \tag{13}$$

2.4. Workflow of Collocated Cokriging Method for Prediction

The Eaton formula method is widely used and has high accuracy. However, its predictive accuracy is not high over the entire spatial domain. Geostatistical stochastic simulation methods, including the collocated cokriging method, serve as effective tools to improve accuracy by integrating well-logging information in spatial analysis. Combining the Eaton formula method with the isochronous co-simulation kriging method is expected to enhance the accuracy of formation pressure prediction in the plane. Numerical simulations of seismic P-wave and S-wave velocities are conducted to analyze the experimental variogram of P-wave velocity and fit it using a spherical model. Anisotropy of the formation is considered, and an elliptical anisotropic model is established to make the simulation results better match the actual formation. By comparing with simple kriging and cokriging methods, the accuracy of the collocated cokriging method is validated. The effects of range, azimuth, and number of reference points on the simulation results are analyzed, and appropriate parameters are selected. Finally, the Eaton formula method combined with the collocated cokriging method is applied to predict formation pressure in the ultra-deep rock formations of the Junggar Basin. The operation process is shown in Figure 1.

The Eaton method is used to calculate the formation pressure at well locations. In collocated cokriging, the assumption is that the variogram function of the secondary variable is the same as that of the primary variable. Therefore, fitting only the secondary variable is required. A spherical model is used for fitting, and the parameters are obtained using the least squares method to determine the range. An elliptical model is established to improve the interpolation accuracy. The covariance and cross-covariance functions of the primary and secondary variables are calculated using the variogram function. A certain number of primary data points are selected around the estimation point, and their weights, along with the weights of the corresponding secondary data points, are obtained by solving the equation matrix for estimation.
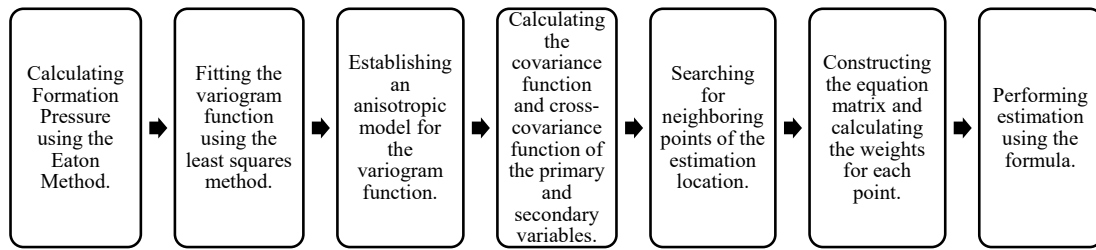
Figure 1. Collocated cokriging prediction process diagram.

3. Numerical Simulation of Collocated Cokriging Method

3.1. Fitting Experimental Model Variogram Functions

The numerical simulation uses compressional wave velocity (Vp) as the covariate to constrain the interpolation of the main variable—shear wave velocity (Vs). To simulate the distribution of actual well points, six randomly sampled data points of Vs are chosen as the primary data, as shown in Figure 2.
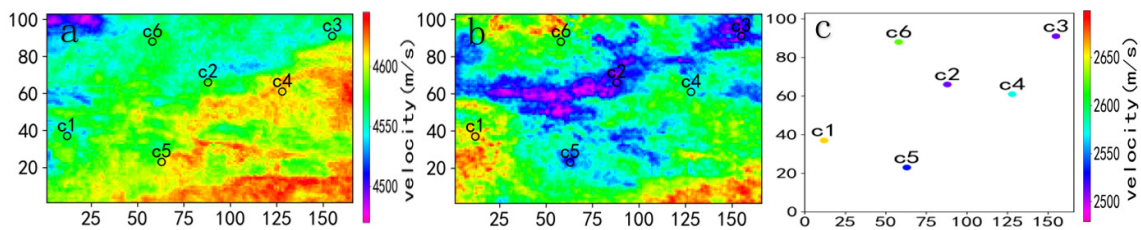

Figure 2. Distribution map of data Vp (**a**), Vs (**b**), and primary data (**c**) in the same area.

The Markov MM1 model assumes that the covariance function of the covariate is the same as that of the primary data, so only the experimental variogram function of the covariate needs to be discussed. Variogram functions exhibit directional characteristics in anisotropic media, with the azimuth of the variogram function being consistent with its opposite direction. An angle tolerance of 30° is chosen, and the variogram functions of the covariate in the six directions on the plane are analyzed within 30° intervals, as shown in Figure 3. The length of the range is defined in terms of the number of grid cells, and this definition size is used for all ranges in the experimental model.
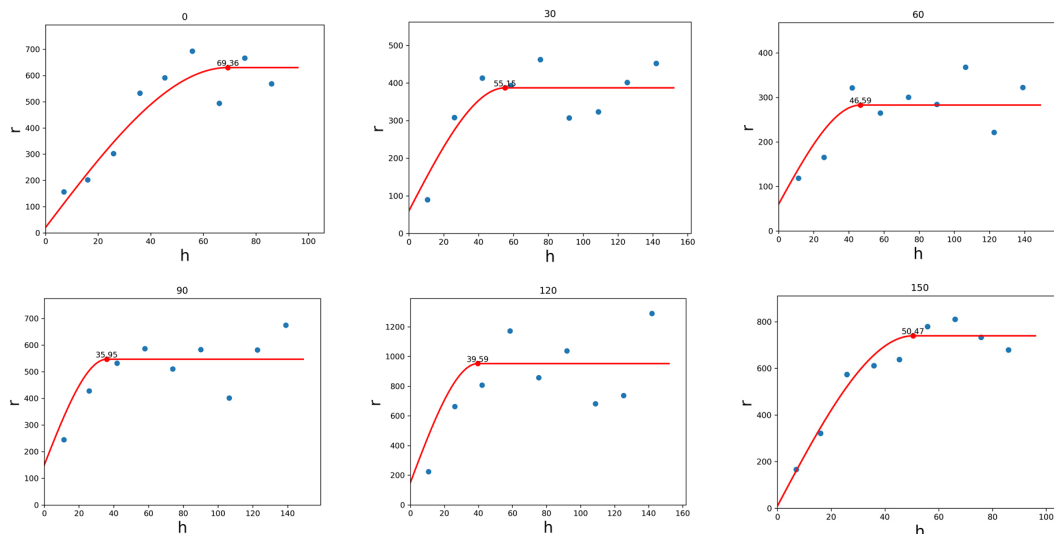

Figure 3. Experimental variogram of Vp in six directions of plane (Blue dots are Experimental variation function value, and red lines are fitting lines.).

From the experimental variogram function graph (Figure 3) and the fitted range ellipse graph (Figure 4), it can be observed that the maximum range occurs in the northeast–southwest direction (21°) and is equal to 71.704 grid units in length. This direction represents the maximum extension of the compressional wave velocity, indicating good continuity of the compressional wave velocity in that direction.
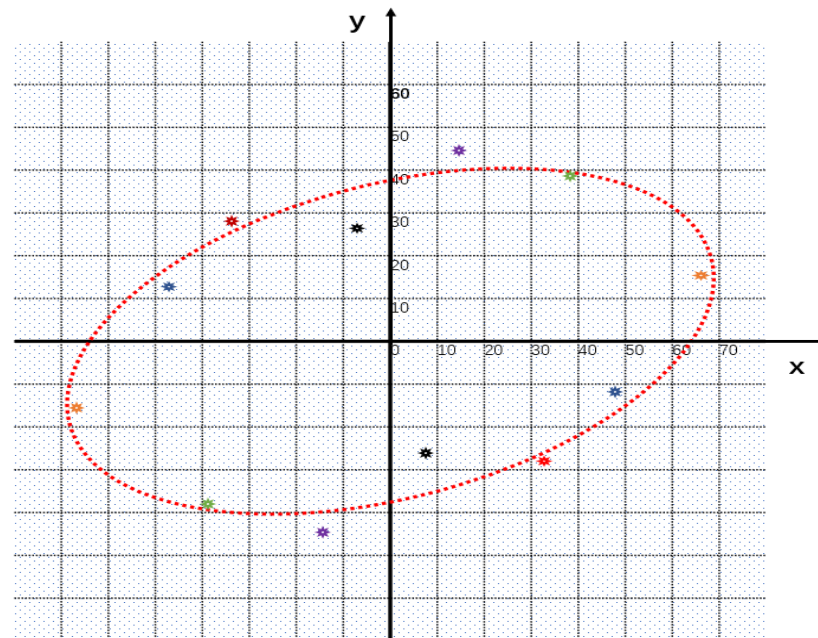
Figure 4. Elliptic variogram (different color points correspond to the range change in different directions).

Figure 4 is the ellipse variogram function graph on the plane of the compressional wave velocity. From this function, it can be observed that the variogram function structure type within the work area is an ellipse model. The maximum range in all directions is 71.704 grid units in length, with an azimuth of 21°. The minimum range is 36.872 grid units in length, with an azimuth of 111°. The ratio of the short axis to the long axis of the range ellipse is approximately 0.514, indicating a strong directional distribution of compressional wave velocity Vp within the work area, and its distribution on the plane is not uniform.

After determining the range size and direction, with the maximum range being 71.704 grid units, the minimum range being 36.872, and the direction being 21°, the variogram function of the covariate is established using the range ellipse model. According to the assumption of the Markov MM1 model:

$$\rho_y(\mathrm{h}) = \rho_z(\mathrm{h})$$
$$\rho_{zy}(\mathrm{h}) = \rho_{zy}(0) \cdot \rho_z(\mathrm{h}) \tag{14}$$

where $\rho_z(\mathrm{h})$ and $\rho_y(\mathrm{h})$ are the standardized autocovariance functions or autocorrelation plots of the main and covariate variables, $\rho_{zy}(\mathrm{h})$ is the cross-covariance function or cross-correlation plot between the main and covariate variables, and $\rho_{zy}(0)$ is the linear correlation coefficient derived from the collocated main and covariate variables.

The MMI model does not require the covariance function of the covariate variable because of the following assumptions.

It is assumed that $\rho_{zy}(\mathrm{h})$ and $\rho_z(\mathrm{h})$ have the same shape and continuity, and by adjusting the covariate variable y(u) with the main variable z(u) located at the same position, the influence of any more distant main variables can be eliminated. If z(u) is defined on the same or larger volume than y(u), then MM1 is a reasonable model.

3.2. Results

In the case of limited primary data points, the simple kriging, cokriging, and collocated cokriging methods are sequentially applied to simulate and predict the S-wave velocity, as shown in Figure 5.
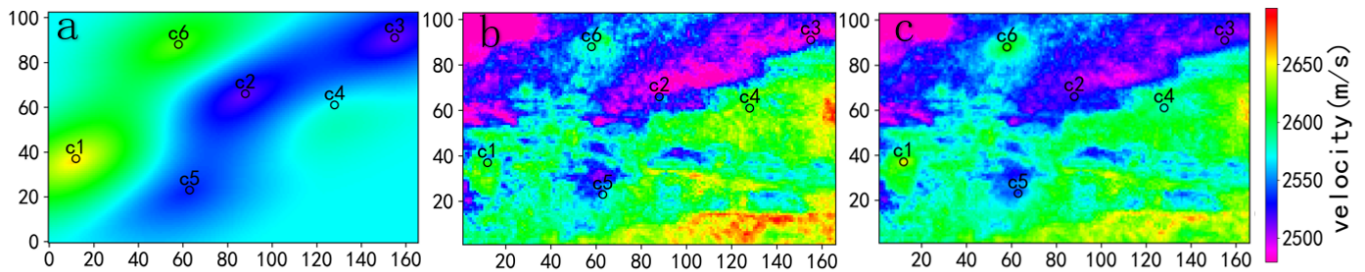
Figure 5. Simulated prediction maps using simple kriging (**a**), cokriging (**b**), and collocated cokriging (**c**).

In comparing the three kriging methods, it is found that simple kriging has limited practical reference value when interpolating with few data points. When the number of primary data points is small and the secondary data is densely distributed, cokriging relies more on the secondary data, with less influence from the primary data during interpolation. Collocated cokriging uses only the secondary data located at the same position as the estimated point, and the calculated weights are dependent on the correlation coefficients. Cokriging requires a significant amount of computation in calculating the cross-covariance or cross-variogram functions, and its computational efficiency is significantly slower compared to collocated cokriging.

3.3. Discussion

The simple kriging method uses only the primary variable in the interpolation process, and since no secondary data are used, the predicted distribution map generated by simple kriging has limited practical reference value in the current situation. When introducing secondary data with certain correlation coefficients, the predicted distribution map simulated by cokriging approximately aligns with the spatial distribution of the secondary data. The influence of the primary data on the estimated values in the surrounding area is limited, and the predicted distribution map is mostly determined by the secondary data. On one hand, the scarcity of primary data points results in a limited influence range. On the other hand, because the prediction grid is based on the distribution of secondary data, when searching for data points around the estimated point, the secondary data points are closer and more densely distributed. According to the variogram function model, the closer the lag distance, the larger the covariance value. Therefore, in the process of solving the equation system, the weights corresponding to the secondary data are larger, which leads to the dominance of the secondary data in the prediction when the primary data points are relatively sparse. The predicted results of collocated cokriging show a spatial distribution that is roughly consistent with the distribution of the secondary data. However, compared to cokriging, the weight of the secondary data in collocated cokriging depends on the correlation coefficient between the primary and secondary data, and the influence of the primary data on the estimated values in the surrounding area is enhanced, while still preserving the information from the primary data.

3.4. Parameter Optimization for Collocated Cokriging

The variogram function is an important input parameter in the collocated cokriging method, and thus the characteristic parameters of the variogram function will inevitably have certain effects on the results.

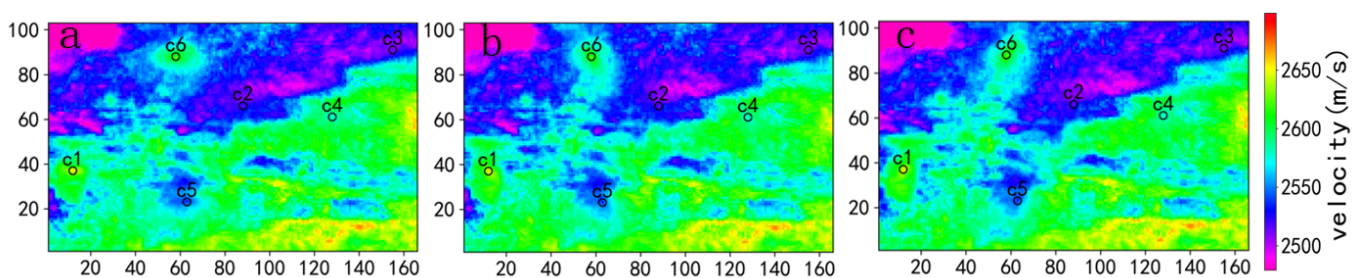### 3.4.1. Influence of Range on Experimental Results

Figure 6 shows the impact of range variation on simulated predictions. Since an elliptical variogram model is established, two ranges in different directions are needed to correspond to the major and minor axes of the ellipse, reflecting the elliptical influence of the primary data points on the surrounding data in the prediction. Decreasing the range reduces the affected area by the primary data, while increasing the range expands its influence. The magnitude of the range represents the extent of the influence of the primary data points. When there is a strong correlation between the primary and secondary data, the strong correlation weakens the influence of the primary data on the surrounding estimated points. Strong correlation plays an important role in collocated cokriging, and it can weaken the influence of fewer primary data points.
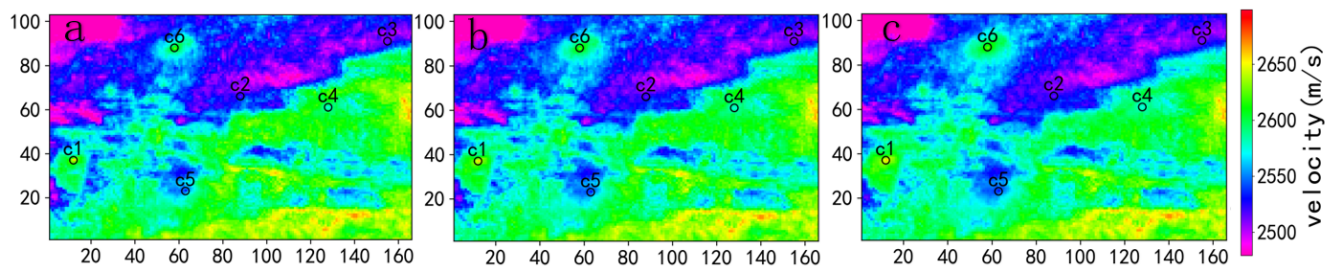
Figure 6. The maximum range and minimum range are 40 and 20 (**a**); 60 and 30 (**b**); 80 and 40 (**c**).

3.4.2. Influence of Anisotropy on Experimental Results

Anisotropy includes the azimuth and anisotropy ratio. The azimuth is based on the positive direction of the *x*-axis, rotating clockwise as the positive direction. Changing the azimuth to $0°$, $60°$, and $120°$ (Figure 7), the collocated cokriging is simulated for the primary data. When different azimuths are set, the simulated prediction distribution map shows evident directional characteristics around the existing primary data points, and the data outside the range, i.e., beyond the influence of the primary data, is not significantly affected by the primary data. Similarly to the range, when there is a strong correlation between the primary and secondary data, the strong correlation weakens the influence of the primary data on the estimated points within the range.


Figure 7. The angles are $0°$ (**a**), $60°$ (**b**), and $120°$ (**c**).

3.4.3. Influence of Number of Conditioning Points on Experimental Results

In the simulation experiment, if there are too many conditioning points involved in the estimation, it will result in complex calculations for the equation matrix, severely affecting the execution speed of the algorithm. On the other hand, a certain number of conditioning points are needed for the estimation, as an insufficient number of conditioning points will lead to interpolated results that do not reflect the true distribution of reservoir parameters. Therefore, the selection of the number of conditioning points is crucial.

In the simulation experiment, the primary and secondary data have a strong correlation. To better illustrate the influence of the number of conditioning points on the results, the same set of data is used and the correlation coefficient between the primary and secondary data is artificially changed. The collocated cokriging simulation experiments

under different numbers of conditioning points are shown in Figure 8. When performing interpolation on the estimated points using collocated cokriging, only the surrounding primary data points are considered. From the images, it can be observed that when there are too few primary data points involved in the estimation, the interpolated images have distinct abrupt boundaries. On one hand, this is due to the scarcity and uneven distribution of primary data points; on the other hand, the significant differences in parameter values among different primary data points lead to large variations between different regions. Therefore, the number of conditioning points should be reasonably selected based on the actual situation.
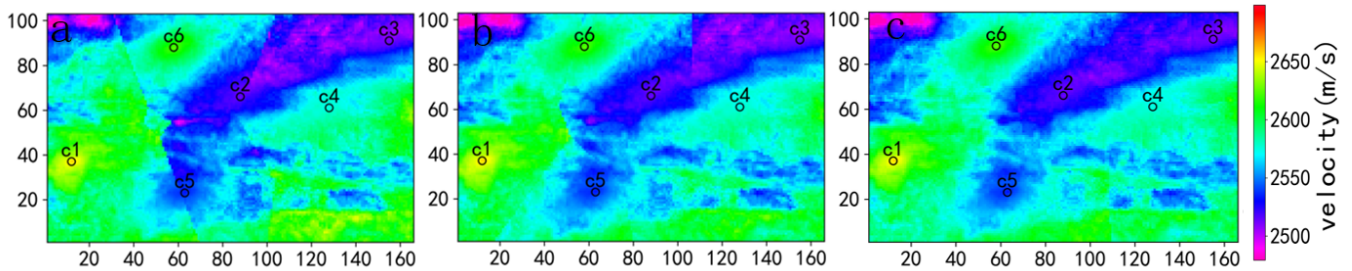

Figure 8. Number of conditioning points are 2 (**a**), 3 (**b**) and 4 (**c**).

4. Pressure Prediction of Actual Strata

4.1. Data Introduction

The dataset used in this paper comes from the Jurassic Badaowan Formation in Junggar Basin, mainly composed of sandstone and mudstone. The reservoir in the area of deep burial, low porosity, low permeability and complex pore structure is a typical unconventional tight hydrocarbon reservoir. There are eight wells in the area, of which four wells contain parameters required by the Eaton method, which are marked ca1, ca2, ca3, and ca4, respectively(Figure 9). The exponent N in the Eaton formula varies in different regions, and multiple trial calculations have shown that an N value of 5.0 is more suitable for the work area. The normal acoustic time difference is calculated using the fitted trend of normal acoustic time difference from wells ca1, ca2, ca3, and ca4.
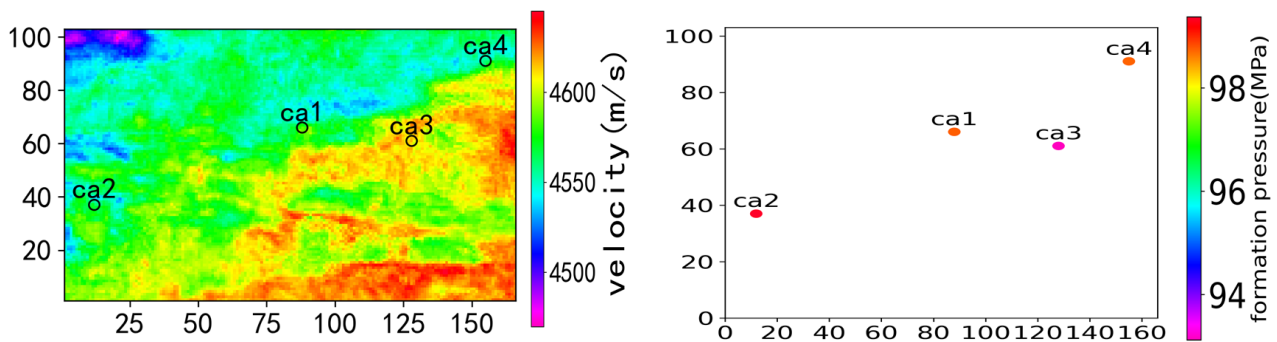

Figure 9. Vp distribution map (**left**) and well point distribution(**right**) in the same area.
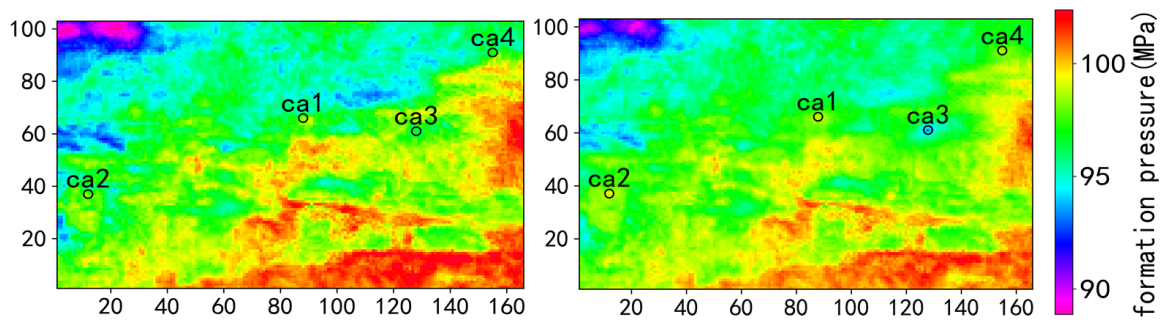
4.2. Results

The formation pressures at the same depth calculated by the Eaton formula for these wells are 98.761 MPa, 99.372 MPa, 93.121 MPa, and 98.714 MPa, respectively. Figure 10 shows the overburden pressure, hydrostatic pressure, and formation pressure at the same depth for well ca2 calculated using the Eaton formula.

In the actual calculation process, calculating the formation pressure for individual wells is cumbersome, and obtaining the normal compaction velocity is not straightforward. Analyzing the Eaton formula reveals that there is a certain relationship between formation pressure and measured velocity, and the calculated formation pressure from the Eaton formula and the measured compressional wave velocity exhibit a linear correlation

with a correlation coefficient of 0.8034. Therefore, cokriging and collocated cokriging are simultaneously used to predict formation pressure (Figure 11).
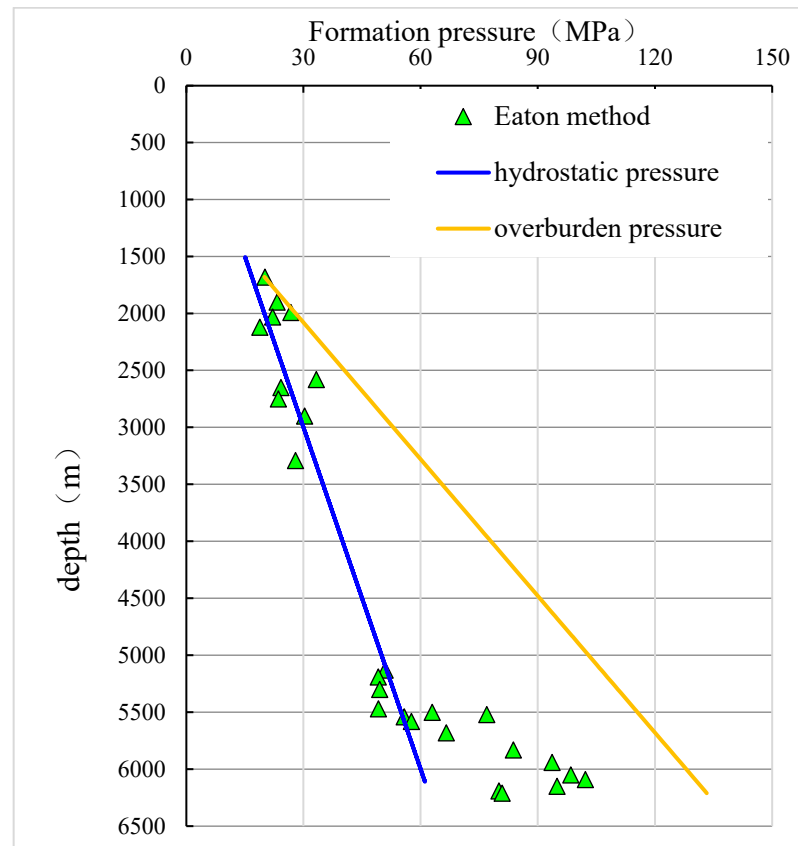
Figure 10. Formation pressure calculated by the Eaton formula for well ca2.

Figure 11. Formation pressure map is obtained by cokriging (**left**) and collocated cokriging (**right**).

4.3. Discussion

　　Consistent with the results of simulation testing, using fewer primary data points for cokriging interpolation aligns more with the distribution trend of secondary data, but it fails to match the existing data at known point locations, resulting in the loss of existing data. On the other hand, collocated cokriging, while maintaining the existing data, assigns appropriate weights to surrounding data points, resulting in predicted results that better align with expectations. Compared to the Eaton formula for predicting formation pressure, which inevitably introduces errors in the process of obtaining normal compaction velocity, the collocated cokriging method predicts formation pressure using measured layer velocities that have a certain correlation with formation pressure, thereby avoiding the difficulty of obtaining normal compaction velocity.

5. Conclusions

1. Compared with different kriging methods, the prediction results of the collocated cokriging method are more in line with experimental expectations. The calculation and fitting of the variogram function in the process of prediction directly affect the prediction results. When performing calculations, geological knowledge should be considered, appropriate models should be selected, and relevant parameters should be obtained reasonably to avoid using variogram functions that do not match the reality, thereby improving the prediction accuracy.
2. Combining the Eaton formula with the collocated cokriging method for predicting formation pressure demonstrates feasibility. This method can be combined with seismic data, and the predicted results align with expectations, providing guidance for exploration and development in actual work areas.

Author Contributions: Writing—original draft, Q.W.; Writing—review & editing, Y.L., G.G., Z.G., Z.W. and J.L. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: Data is unavailable due to ethical restrictions.

References

1. Zu, F.; Zhang, Z.L.; Feng, Q.H.; Fan, H. Abnormal formation pressure monitoring and predicting method. *Oil Drill. Prod. Technol.* **2004**, *26*, 35–38. [CrossRef]
2. Zhao, M.; Wei, C.J.; Wang, S.H. Algorithm research and realization of formation pressure prediction based on eaton method. *J. Qingdao Univ.* **2017**, *30*, 86–88+92.
3. Xie, Y.; Li, X.Y.; Zhai, Y. Establishing formation pressure profiles and guiding adjacent wells balanced drilling in new area by logging data. *Pet. Instrum.* **2011**, *25*, 46–48+103.
4. Li, H.; Yu, Z.T.; Yuan, H.S.; He, P.F. Application of formation pressure prediction technology in drilling in Bozhong A block. *Petrochem. Ind. Appl.* **2018**, *37*, 33–36.
5. Fan, X.Z.; Li, X.; Wu, Y.C.; Zhang, G.; Lou, Y.; Liu, S.; Zhu, Y. A formation pressure prediction method for well drilling in the Arctic permafrost region. *Nat. Gas Ind.* **2022**, *42*, 99–105.
6. Wang, B.; Yong, X.S.; Pan, J.G.; Yin, L.; Xu, D.N.; Kong, X.; Qu, J.H.; Tan, K.J.; Huang, Y. A new method to predecit pore pressure in low permeability reservoirs and its application. *Prog. Geophys.* **2015**, *30*, 695–699. [CrossRef]
7. Zhang, Q.; Zhi, Y.P. Applying seismic velocity to predicting formation pressure. *Chin. J. Eng. Geophys.* **2015**, *30*, 695–699.
8. Yang, H.; Zhou, P.G.; Sun, W.G.; Shi, J.G.; Chen, W.F. Using seismic data to predict formation pressure in piedmont structures at the southern margin of Junggar Basin. *Xinjiang Pet. Geol.* **2017**, *38*, 347–351.
9. Ma, H. Pore pressure prediction with seismic interval velocity by the modified Fillippone Method. *Pet. Drill. Tech.* **2012**, *40*, 56–61. [CrossRef]
10. Qian, L.P.; Wang, X.; Li, F.; Li, J.H.; Wang, J.; Qu, Y.W. Formation pore pressure prediction using Fillipone formula combined with equivalent medium theory. *Oil Geophys. Prospect.* **2018**, *53*, 224–229.
11. Shi, M.X.; Liu, Z.D.; Yang, X.F.; Yang, J.R.; Chen, X.J.; Liu, H.Z.; Cao, J.L. Review and prospect prediction technology for formation pore pressure by geophysical well logging. *Prog. Geophys.* **2020**, *35*, 1845–1853. [CrossRef]
12. Gao, R.N. Study on Sandstone Distribution in Yuncheng Based on Geostatistics. *Shanxi Sci. Technol.* **2020**, *35*, 37–41.
13. Ma, W.L.; Deng, S.Z. Application of Functional Kriging on the Well Productivity Prediction. *Inn. Mong. Petrochem. Ind.* **2021**, *47*, 92–96.
14. Li, S.H.; Wang, Y.B.; Li, J.; Liu, X.T.; Wang, J.; Gong, W.Q. Correcting the string effect of Kriging. *OGP* **2012**, *47*, 978–983.
15. Wang, C.Y.; Qu, Y.; Shuai, Y.M. Prediction of the Spatial Distribution of Gas Pipe Network Using Cokriging Method Based on ArcGIS Technologies. *Geomat. Spat. Inf. Technol.* **2022**, *45*, 28–32.
16. Du, W.F.; Peng, S.P. Coal seam Thickness Prediction With Geostatistics. *Chin. J. Rock Mech. Eng.* **2010**, *29*, 2762–2767.
17. Geng, M.X.; Huang, D.N.; Yu, P.; Yang, Q.J. Three-dimensional constrained inversion of full tensor gradiometer data based on cokriging method. *Chin. J. Geophys.* **2016**, *59*, 1849–1860. (In Chinese) [CrossRef]
18. Yu, Z.J.; Dong, D.D.; Song, W.Q.; Gao, Y.K.; Wu, S.G. Porosity prediction with co-Kriging method controlled by sedimentary facies. *Prog. Geophys.* **2012**, *4*, 1581–1587. [CrossRef]
19. Chen, J.Y.; Yu, X.H.; Li, S.L.; Hou, G.W. Application and method of predicting reservoir by collocated seismic attributes. *Inn. Mong. Petrochem. Ind.* **2007**, *121*, 1–3.

20. Wang, D.D.; Yang, X.F.; Zhou, Y.B.; Ma, Z.Z.; Liu, Y.M. Improving the accuracy of seismic inversion using the horizontal well data. *Prog. Geophys.* **2019**, *34*, 185–190. [CrossRef]
21. Zhang, S.P.; Yu, X.H. The application of collocated co-kriging stochastic modeling in reservoir description. *Nat. Gas Geosci.* **2006**, *17*, 378–381.
22. Niu, W.J.; Meng, X.H.; Li, J.G. A new estimation method of collocated CoKing combined with soft data. *Coal Geol. Explor.* **2011**, *39*, 13–17. [CrossRef]