

Selection of optimum wavelet in CWT analysis of geophysical downhole data

Amrita Singh¹, Saumen Maiti¹ and R.K.Tiwari²

¹Department of Applied Geophysics, Indian Institute of Technology (Indian School of Mines), Dhanbad-826004, India

²CSIR-National Geophysical Research Institute, Hyderabad-500007, India

*Corresponding Author: saumen_maiti2002@yahoo.co.in

ABSTRACT

Continuous wavelet transform (CWT)-based scalogram analysis is appropriate for the modelling of discontinuous well log signal. The choice of appropriate mother wavelet in scalogram analysis is crucial to model the geophysical well log data precisely. Here, we examine some of the key points related to the choice of optimum wavelet for the analysis of well log signal. We used the scalogram analysis to detect the formation interface and explored the impact of each mother wavelet. Following three key indicators, e.g. (i) histogram analysis of CWT coefficient at even number scale (ii) statistical significance test of each sub-signal at even number scale and (iii) Principal component analysis (PCA) of those sub-signals are used. These key steps help to identify the precise localization of formation tops detection problem of well log data of KTB borehole, Germany. Here, five mostly used wavelet functions (Haar, Gaussian wavelet of order 1 (Gaus1), Gaussian wavelet of order 3 (Gaus3), Morlet and Daubechies wavelet of order 2 (Db2)) are employed on two sets of data: (i) spectral gamma ray (SGR) log in pilot hole and (ii) density (RHOB) log in main hole data. The comparative results suggest that Gaus1 is better among all mother wavelets rendering maximum number of occurrences of CWT coefficients in the total well log signal modelling. Statistical significance test shows that CWT coefficients and their respective scales are statistically important. The results based on PCA analysis further suggest that Gaus1 wavelet is in good agreement with the gross structure of the signal.

Key words: Continuous wavelet transform, PCA analysis, Histogram analysis, Significance test, Formation interface, KTB.

INTRODUCTION

Wavelet transform has become a very popular tool for data analysis in almost all domains of applied sciences (Sang et al., 2016; Shoaib et al., 2014). The wavelet transform-based analysis is appropriate for the investigation of the non-stationary signals, which are common in real geophysical time/space series data. Wavelet (Morlet et al., 1982), has been broadly used in time series analysis (e.g., identifying periods and duration of signal band, time series modelling/forecasting, de-noising of time series, and identification of true components, characterization of subsurface geology/geological rock types, litho-logical boundary identification/modelling discontinuity etc) (Kumar and Foufoula-Georgiou, 1997; Kisi, 2009; Percival and Walden, 2000; Adamowski and Chan, 2011; Doveton, 1986; Farge, 1992; Daubechies, 1992; Chandrashekhar and Rao, 2012; Labat, 2008; Perez-Munoz et al., 2013; Liu et al., 2014; Maraun et al., 2007; Pan et al., 2008; Prokoph and Agterberg, 2000; Goupillaud et al., 1985).

However, selection of optimum mother wavelet and their time-based scale distribution of the wavelet coefficients are two important issues in any applications (Percival and Walden, 2000). Particularly, the choice of suitable mother wavelet in scalogram analysis is crucial for the precise analysis of geophysical well log data

analysis (Chandrashekhar and Rao, 2012). Several studies have discussed the procedure for appropriate selection of optimum mother wavelet function. For example, Morlet wavelet is found suitable for the identification of key sedimentary cycles (Prokoph and Agterberg, 2000). The Haar is found realistic for the identification of formation interface in well logging (Pan et al., 2008). To identify fluids in wells and pay zones, Chandrashekhar and Rao (2012) experimented with various mother wavelets and performed mainly histogram analysis of CWT coefficient and found that Gaussian wavelet is the most appropriate. Therefore, there is no widely accepted consensus on the choice of the best method and selection of optimum mother wavelet and temporal scale.

Recently, several researchers have carried out experiment to search for suitable optimum wavelets for modelling (Nourani et al., 2011; Maheswaran and Khosa, 2012; Singh, 2011; Shoaib et al., 2014). Torrence and Compo (1998) recommended for choosing a non-orthogonal wavelet. They considered both 'width' and 'shape' of the signal/wavelet and performed similarity measures between wavelet and original series for finalizing better mother wavelet. Schaeffli et al., (2007) reported that wavelet should be chosen such that it shows good time-frequency localization by compromising between time and scale resolution. Nourani et al., (2014) advocated a setting criterion by utilizing

similarity measures between wavelet and raw series. In their work, it is explained how wavelet match with malleable depth-scale window that contracts to view small scale structures and broadens to view large-scale structures, akin to a zoom lens (Walker, 1999; Percival and Walden, 2000; Singh et al., 2016; Kumar and Foufoula-Georgiou, 1997; Javid and Tokhmechi, 2012). Particularly, Sang et al., (2016) presented a succinct way of various technical issues in wavelet analysis. They focus mainly on DWT-based analysis and its usefulness is credited by the fact that categorical distributions of DWTs are approximately uncorrelated. However, their recommendation for mother wavelet selection and scale has not resolved the ongoing debate completely. In fact, their approach may not be appropriate for CWT-based analysis where each sub-signal is intimately coupled. Recently, Chandrashekhar and Rao (2012) have done scalogram analysis and efficient histogram analysis of different CWT coefficients to select the optimum mother wavelet. The above scheme is found to be strong for particular downhole data and has been used to model the space localization of formation tops by choosing the appropriate mother wavelet. However, they did not deal with the quantitative statistical assessment and significance of the results at various scales exclusively in choosing the best mother wavelet and a more convincing careful quantitative assessment is required. Therefore, we carried out an experiment to select optimum mother wavelet and see how the associated CWT coefficients are statistically significant to explain the original signal in different scales. Moreover, our combined approach is different in a sense that we have (i) considered and compared performance of both popular orthogonal (Haar, Daubechies wavelets) and non-orthogonal (Morlet, Gaussian wavelets) wavelet in scalogram analysis for good localization in space domain, (ii) taken care of even number scales of CWT coefficients for histogram analysis, (iii) included statistical significance test analysis of the wavelet coefficients to show how the coefficients are significant in matching signal characteristics (iv) used PCA analysis to examine relative influence of each sub-signal at various even number scale using various mother wavelet on over all signal matching characteristics.

Before going to the data analysis, we briefly present the KTB data and wavelet transform theory for the completeness of the article.

MATERIALS AND METHODS

KTB borehole data

The KTB is located in the north-eastern Bavaria, southern Germany. The KTB crust consists of paragneisses, metabasites and alternations of gneiss-amphibolites, with minor occurrence of marbles, calcsilicates, orthogneisses,

lamprophyres and diorites (Pechinig et al., 1997). The maximum depths of pilot and main hole are 4 km and 9.1 km respectively. The data is digitized at 0.1524 m interval. We have used gamma ray log of pilot hole and density log of KTB main hole to model boundary via optimum wavelet selection.

Wavelet analysis via CWT

Goupillaud et al., (1985) defined a wavelet function such as

$$\psi_{u,s}(z) = \frac{1}{\sqrt{|s|}} \psi\left(\frac{z-u}{s}\right), u > 0, s \in R \quad (1)$$

where the ψ is mother wavelet, s is scale factor which depends on wavelength, u is shift parameter and z is a variable. In this framework, the transformation is performed for different sections of the convolved signal by changing s and u . Accordingly, CWT of signal $f(z)$ takes the form of

$$CWT_f(u,s) = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{\infty} f(z) \psi^*\left(\frac{z-u}{s}\right) dz \quad (2)$$

where $*$ denotes the complex conjugate. CWT was performed on different downhole data sets, like density, neutron porosity, the gamma-ray intensity, and seismic p-wave velocity and resistivity log. But in the present study, the results of gamma ray log of pilot hole and density log of main hole data in KTB are presented. Experiments were performed to select optimum mother wavelet from a pool consisting of mother wavelet namely Haar, Gaus1, Gaus3, Morlet and Db2 for the application of modelling well log data (Figure 1).

We have employed the CWT-based histogram analysis and examined their statistical significance to emulate the signal characteristics at different scales. Finally, PCA analysis of CWT coefficient was used to examine relative influence of each mother wavelet to model the well log signal boundary detection.

Optimum mother wavelet selection in CWT-based wavelet analysis and boundary detection

A complete computational procedure is presented as a flow diagram in Figure 2. Accordingly after reading the well log data CWT has been applied. Since edge effect is common problem in high value of the CWT coefficients at the edge of the scalograms (contour map of CWT coefficient), which might obscure the signal of interest, we have followed symmetric half-point method of Strang and Nguyen (1995) and Chandrasekhar and Rao (2012). The CWT coefficient was obtained by using CWT analysis with various mother wavelets.

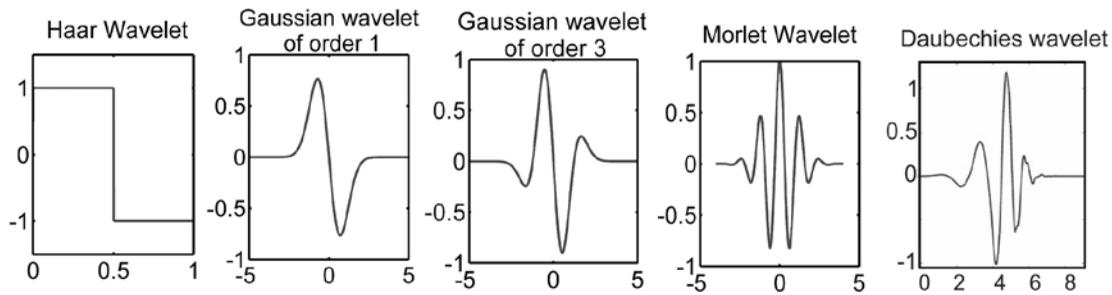


Figure 1. Shape and nature of the mother wavelet used in the present study (a) Haar, Gaus 1, Gaus 3, Morlet and Db2.

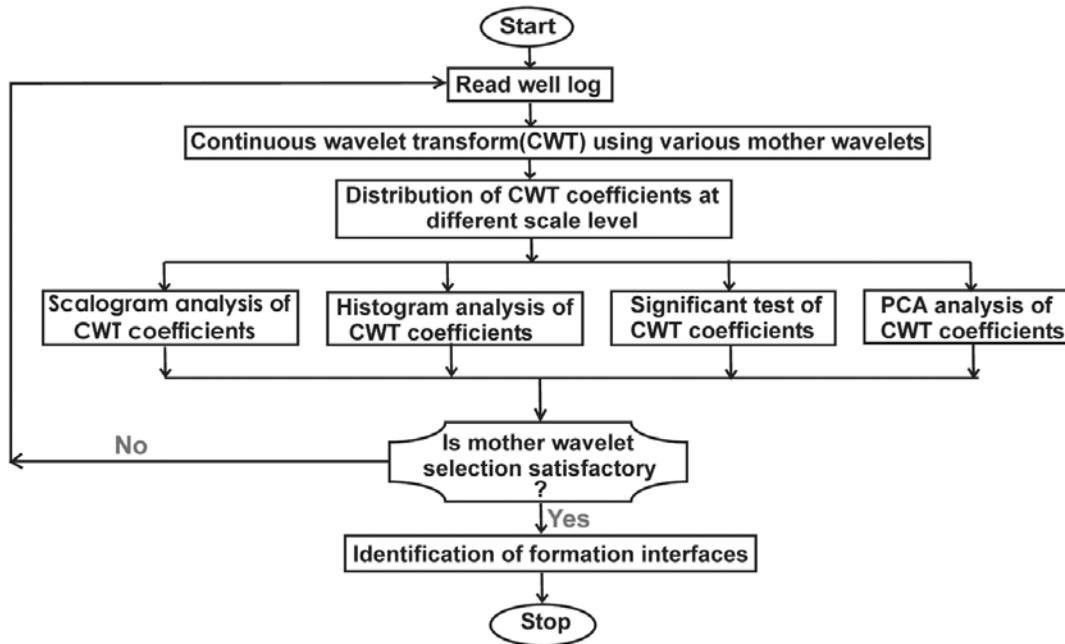


Figure 2. Basic flow diagram of CWT-based boundary detection and mother wavelet selection.

Then histogram analysis, statistical t-test by function t-test2, and PCA-based analysis were conducted on the obtained CWT coefficient distribution at even number scale. The best mother wavelet was chosen with highest histogram peak, t-test results ($h=0$ and p -value less than 1.0) and central characteristics of PCA results in PC1 and PC2 because, in most of the cases using PC1 and PC2, the output data variance is explained to the extent of 90%. At the end of the boundaries, interfaces are detected by using the best mother wavelet and are given in the table 2 and 3.

RESULTS AND DISCUSSIONS

Histogram analysis

Histograms of absolute normalised value of CWT coefficients were prepared for entire well log signal using various mother wavelets. The histogram analysis plot of gamma ray log of KTB pilot hole is presented in figure 3b.

Figure 3b demonstrates that maximum number of occurrence of CWT coefficients is close to 1300 for Haar wavelet, which is slightly lower than the results obtained by the Gaus1. Figure 3b implies that maximum number of occurrence of CWT coefficient for Gaus 3, Db2 and Morlet are around 1400, 1100 and 1000 respectively. Figure 4a shows the scalogram of density log of KTB main hole.

Figure 4b reveals that maximum number of occurrence of CWT coefficients is almost equal to 2000 whereas for Gaus1 it is close to 2100. The maximum number of occurrence of CWT coefficient for Gaus 3, Db2 and Morlet are around 1900, 1750 and 1150 respectively (Figure 4b).

Therefore, histogram-based CWT coefficient suggests that Gaus1 mother wavelet is better among the other wavelets being currently used for matching the gross signal characteristics. Conversely, it can be said that Gaus1 shows a better similarity between wavelet and original signal. It may be noted that CWT-based scalogram analysis gives the CWT coefficient (Amplitude) distribution against depth

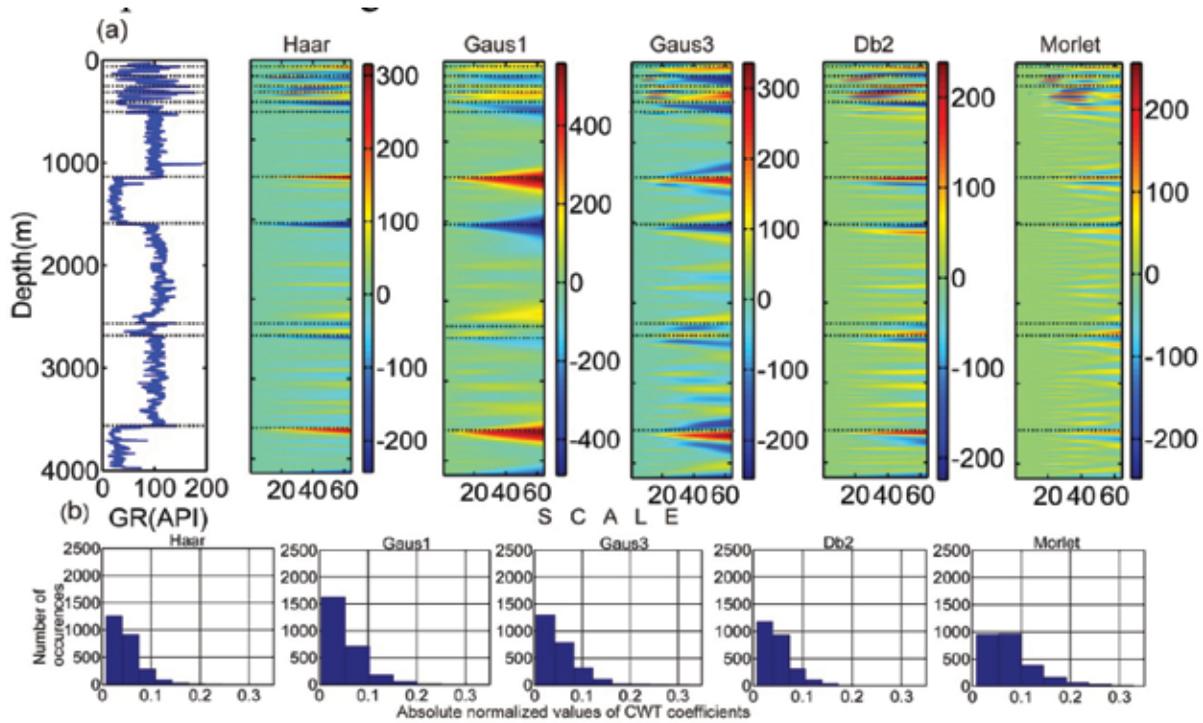


Figure 3. (a) Scalogram plots of gamma ray log data obtained from depth range 28 m to 4000 m at KTB pilot hole corresponding to Haar, Gaus1, Gaus3, Db2, and Morlet wavelets (b) Histogram analysis of CWT coefficients corresponding to Haar, Gaus1, Gaus3, Db2 and Morlet mother wavelet applied on gamma ray log.

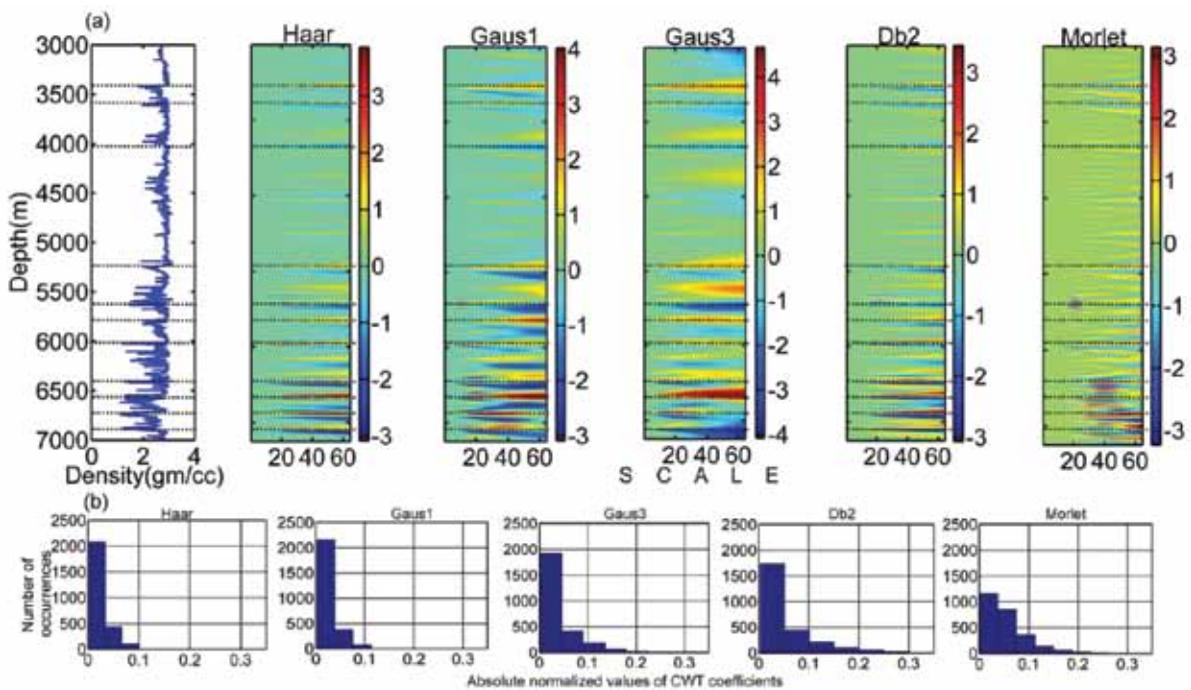


Figure 4. (a) Scalogram plots of density log data obtained from depth range 3000 m to 7000m at KTB main hole corresponding to Haar, Gaus1, Gaus3, Db2, and Morlet wavelets (b) Histogram analysis of CWT coefficients corresponding to Haar, Gaus1, Gaus3, Db2 and Morlet mother wavelet applied on density log.

Table 1. Significance test of the gamma ray log of pilot hole (P) and density log of main hole (M) data using various wavelets of Haar, Gaus1, Gaus 3, Db2 and Morlet. Here $Ph_{Haar,Gaus1}$ means h value using Haar and Gaus1 wavelets in pilot hole data and $Mh_{Haar,Gaus1}$ means h value using Haar and Gaus1 wavelets in main hole data. Others are denoted similarly.

Parameter of Hypothesis t- test	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6
$Ph_{Haar,Gaus1}$ $Pp_{Haar,Gaus1}$	0 0.97	0 0.60	0 0.81	0 0.37	0 0.47	0 0.41
$Ph_{Haar,Gaus3}$ $Pp_{Haar,Gaus3}$	0 0.36	0 0.34	0 0.70	0 0.53	0 0.81	0 0.95
$Ph_{Haar,Db2}$ $Pp_{Haar,Db2}$	0 0.87	0 0.78	0 0.26	0 0.17	0 0.46	0 0.44
$Ph_{Haar,Morlet}$ $Pp_{Haar,Morlet}$	0 0.89	0 0.94	0 0.71	0 0.36	0 0.96	0 0.50
$Mh_{Haar,Gaus1}$ $Mp_{Haar,Gaus1}$	0 0.41	0 0.30	0 0.13	0 0.25	0 0.51	0 0.55
$Mh_{Haar,Gaus3}$ $Mp_{Haar,Gaus3}$	0 0.55	0 0.01	0 0.32	0 0.61	0 0.92	0 0.99
$Mh_{Haar,Db2}$ $Mp_{Haar,Db2}$	0 0.23	0 0.33	0 0.06	0 0.13	0 0.31	0 0.63
$Mh_{Haar,Morlet}$ $Mp_{Haar,Morlet}$	0 0.80	0 0.14	0 0.02	0 0.29	0 0.86	0 0.18

of the well log. From the scalogram figure (3a & 4a), one can identify the litho-facies boundary zones by utilizing the property of space localization which is particularly important to identify the boundary of the litho-facies/ and/ or pay zones in well log analysis. This analysis is particularly significant because the maximum occurrence of CWT coefficient is related to the resolution of the scalogram, which in turn, helps identifying formation boundary. Therefore, it may be noted that optimum mother wavelet selection is critical for similarity measurement between wavelet and original log data in the space-localization of CWT based well log analysis.

Statistical significance test for CWT coefficient at various scales

We performed t-test by the function ttest2 between CWT coefficient distributions using different mother wavelets at various even numbers of scales with 5% significance level. Table 1 demonstrates the statistical significance test results between CWT coefficient of Haar and Gaus1 mother wavelet of scale 2, 4, 8, 16, 32 and 64 of the gamma ray log of KTB pilot hole data. The h-value gives 0.0 for all scales (even number considered here) and p-value ranges from 0.37 to 0.97. Here, h is equal to zero, which implies that null hypothesis cannot be rejected at this point with 5% significance level. A statistical significant value ($p < 1.0$) implies that CWT coefficient of Haar and Gaus 1 is pertinent to model the data at all scales. The result of hypothesis test between CWT coefficient of Haar and Gaus 3 mother wavelet of scale 2, 4, 8, 16, 32 and 64 of the gamma ray log of KTB pilot hole data is shown in table1.

The h-value gives 0 for all scales (even number considered) and p-value ranges from 0.34 to 0.95. Moreover, statistical significant value ($p < 1.0$) implies that CWT distribution coefficients between Haar and Gaus 3 are relevant and closely associated with each other. Similarly, the result of statistical significance test between CWT coefficient of Haar and Db2 mother wavelet of scale 2, 4, 8, 16, 32 and 64 of the gamma ray log of KTB pilot hole data is documented (Table 1). The h-value gives 0 for all scales (even number considered), while the p-value varies from 0.17 to 0.87 suggesting that CWT coefficients between Haar and Db2 are closely tied to their relevancy in current analysis. The statistical significance test result between CWT coefficient of Haar and Morlet mother wavelet of scale 2, 4, 8, 16, 32 and 64 of the gamma ray log of KTB pilot hole data is demonstrated (table 1). The h-value gives 0 for all scales (even number considered) and the p-value ranges from 0.36 to 0.94. The p-values are also found statistically significant ($p < 1.0$).

Likewise, Table 1 also demonstrates the hypothesis test result between CWT coefficient of Haar and Gaus1 mother wavelet of scale 2, 4, 8, 16, 32 and 64 of the density log of KTB main hole data. The h-value gives 0 for all scales (even number considered) and the p-value ranges from 0.13 to 0.55. Here also, null hypothesis cannot be rejected with 5% significance level because the h value is found to be zero. All the p-values are found to be less than 1.0 implying that CWT coefficients of Haar and Gaus 1 corresponding to all even number scales are statistically significant to model the original gamma log variation of KTB pilot hole.

The hypothesis test result between CWT coefficient of Haar and Gaus 3 mother wavelet of scale 2, 4, 8, 16,

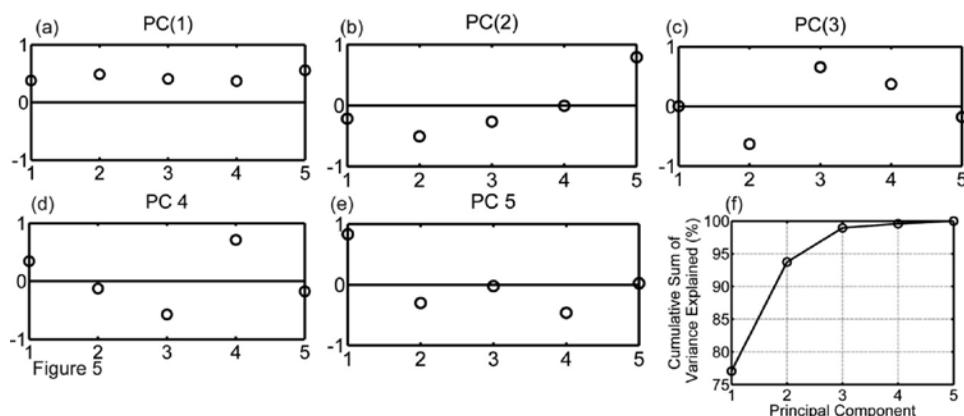


Figure 5. (a-e) PCA plot displays the impact of each variable (CWT coefficient series at scale 2 of various mother wavelets of gamma ray log of pilot hole) for modelling of well log signal. In the x-axis, position of Haar, Gaus 1, Gaus 3, Db2 and Morlet wavelet is represented by 1, 2, 3, 4, and 5 respectively. Y-axis denotes the impact of each wavelet and the value is normalized between -1 and +1. (f) The PCA analysis of well log data from cumulative sum of variance is explained by the different PCs.

32 and 64 of the gamma ray log of KTB pilot hole data is presented in Table 1. The h-value gives 0 for all scales (even number considered) and the p-value ranges from 0.01 to 0.99. The statistical significance ($p < 1.0$) is closely linked to the authenticity and reliability of relevance. The result of hypothesis test between CWT coefficient of Haar and Db2 mother wavelet of scale 2, 4, 8, 16, 32 and 64 of the density log of KTB main hole data are documented in Table 1. The h-value shows 0 for all scales (even number considered) and the p-value varies from 0.13 to 0.63. The relevance between the CWT coefficient distribution of Haar and Db2 mother wavelet is distinct and noteworthy ($p < 1.0$). The hypothesis test result between CWT coefficient of Haar and Morlet mother wavelet of scale 2, 4, 8, 16, 32 and 64 of the density log of KTB main hole data is provided (Table 1). The h-value gives 0 for all scales (even number considered) and the p-value ranges from 0.02 to 0.80. All the p-values are found to be less than 1.0 meaning that CWT coefficients of Haar and Morlet corresponding to all even number scales are statistically significant to model the original density log variation of KTB main hole.

PCA analysis

The Principal component analysis (PCA) is an effective tool to remove linear dependencies among variables (Jolliffe 1972). Ultimately it helps to reduce the data dimension for visualization (Ojha and Maiti 2016). Total data variance is distributed among PCs such as PC(1), PC(2), PC(3) etc. In order to represent the data characteristics satisfactorily, it is suggested to keep PCs whose eigen values are larger than 0.7 (Jolliffe, 1972).

We carried out PCA analysis of CWT coefficients at different scales obtained from CWT of gamma ray log of KTB pilot hole and density log of KTB main hole. Here, the

total variables are five (i) Haar wavelet (ii) Gaus 1 wavelet (iii) Gaus 3 wavelet (iv) Db2 wavelet (v) Morlet wavelet. The results of PC on the five mother wavelets for the depth 28-4000 m interval are shown in figures 5-10 for pilot hole and figures 11-16 for main hole.

Pilot hole data

Figures 5a-e demonstrate the role of each PC corresponding to each mother wavelet for modeling well log at scale 2 and suggests that the Gaus 1 shares relatively principal role in modeling gamma ray in panel PC(1), although, mother wavelets show a significant positive involvement in PC(1) and PC(2), Gaus 1 and Morlet play key role in opposite direction. In PC(3), Gaus 1 and Gaus 3 have main role in opposite direction. In PC(4), Gaus 3 and Db2 have foremost role in opposite direction.

In PC(5), Haar has a key role opposite to Gaus 1 and Db2. Figure 5f suggests that 94% variance of the downhole data described by first two PCs. Thus, involvement of CWT coefficient data can in principle be described by first two PCs which are particularly significant in analyzing CWT coefficient of voluminous industrial data. Thus, the multivariable data set can be described using only two coordinate axes. The PCA based leading role amongst Haar, Gaus 1, Gaus 3, Db 2 and Morlet at scale 2 for the depth 28-4000 m interval of gamma ray log is demonstrated in figures 5a-e.

Figures 6(a-e) reveal the role of each PC corresponding to each mother wavelet for modeling well log at scale 4. Figures 6a-e suggest that the Gaus 3 shares relatively central role in modeling gamma ray in panel PC(1), although, mother wavelets have significant positive involvement in PC(1). In PC(2), Gaus 1 and Morlet have leading role in opposite direction. In PC(3), Gaus 1 and Gaus 3 have governing role in opposite direction. In PC (4),

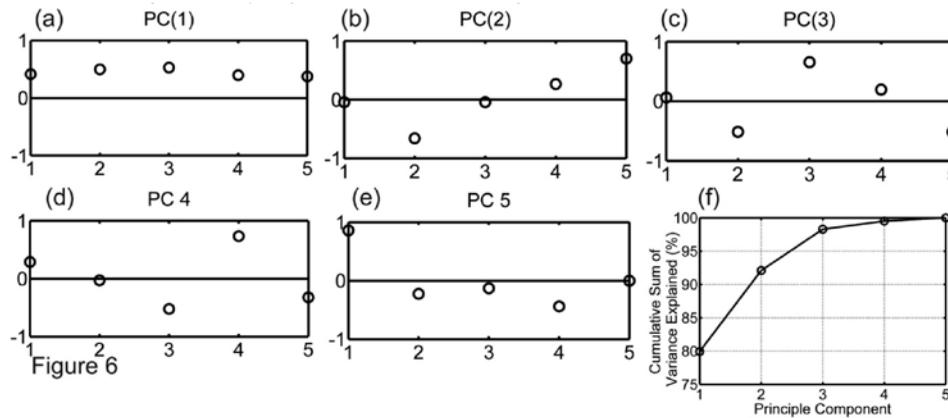


Figure 6. (a-e) PCA plot displays the impact of each variable (CWT coefficient series at scale 4 of various mother wavelets of gamma ray log of pilot hole) for modelling of well log signal. In the x-axis, position of Haar, Gaus 1, Gaus 3, Db2 and Morlet wavelet is represented by 1, 2, 3, 4, and 5 respectively. Y-axis denotes the impact of each wavelet and the value is normalized between -1 and +1. (f) The PCA analysis of well log data from cumulative sum of variance is explained by the different PCs.

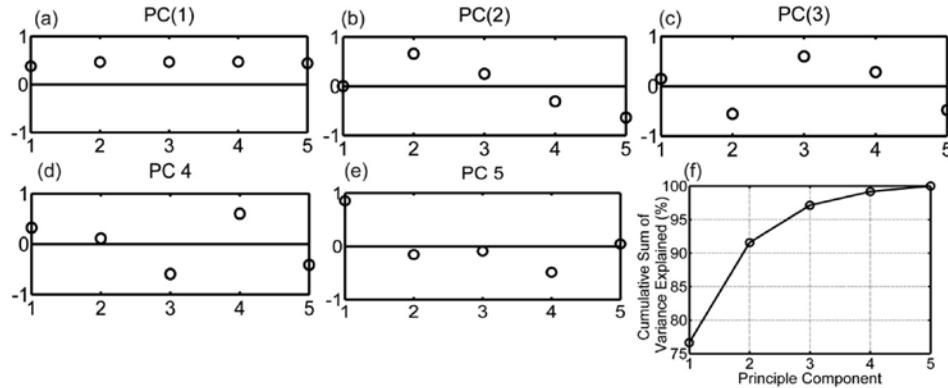


Figure 7. (a-e) PCA plot displays the impact of each variable (CWT coefficient series at scale 8 of various mother wavelets of gamma ray log of pilot hole) for modelling of well log signal. In the x-axis, position of Haar, Gaus 1, Gaus 3, Db2 and Morlet wavelet is represented by 1, 2, 3, 4, and 5 respectively. Y-axis denotes the impact of each wavelet and the value is normalized between -1 and +1. (f) The PCA analysis of well log data from cumulative sum of variance is explained by the different PCs.

Gaus 3 and Db2 have chief role but in opposite direction. In PC(5), Haar has major role opposite to Gaus 1 and Db 2. Figure 6f submits that 92% variance of the downhole data are elucidated by first two PCs.

Therefore, involvement of CWT coefficient data can in principle be described by first two PCs which are particularly significant in analyzing voluminous CWT coefficient of industrial data. Thus the multivariable data set can be described using only two coordinate axes. The PCA-based leading role among Haar, Gaus 1, Gaus 3, Db2 and Morlet at scale 4 for the depth 28-4000 m interval of gamma ray log is demonstrated in figures 6a-e.

Figure 6f suggests that 91% variance of the downhole data could be explained by first two PCs. So involvement of CWT coefficient data could in principle be described by first two PCs without losing much information, which is particularly significant in analyzing voluminous CWT coefficient of industrial data. Thus, the multivariable data set can be described using only two coordinate axes. The

PCA-based significant role among Haar, Gaus 1, Gaus 3, Db2 and Morlet at scale 8 for the depth 28-4000 m interval of gamma ray log is demonstrated in figures 7a-e.

Figures 7(a-e) display the role of each PC corresponding to each mother wavelet for modeling well log at scale 8. Figure 7a-e suggest that the Db2 shares slightly prominent role in modeling gamma ray in panel PC(1), although, mother wavelets have significant positive involvement in PC(1). In PC(2), Gaus 1 and Morlet have prominent role in opposite direction. In PC (3), Gaus 1 and Gaus 3 have important role in opposite direction. In PC(4), Gaus 3 and Db 2 have significant role in opposite direction. In PC(5), Haar and Db 2 have major role in opposite direction.

Figures 8(a-e) exhibit the role of each PC corresponding to each mother wavelet for modeling well log at scale 16. Figures 8a-e suggest that the Gaus 1 shares slightly significant role in modeling gamma ray in panel PC(1), although, mother wavelets have significant positive involvement in PC(1).

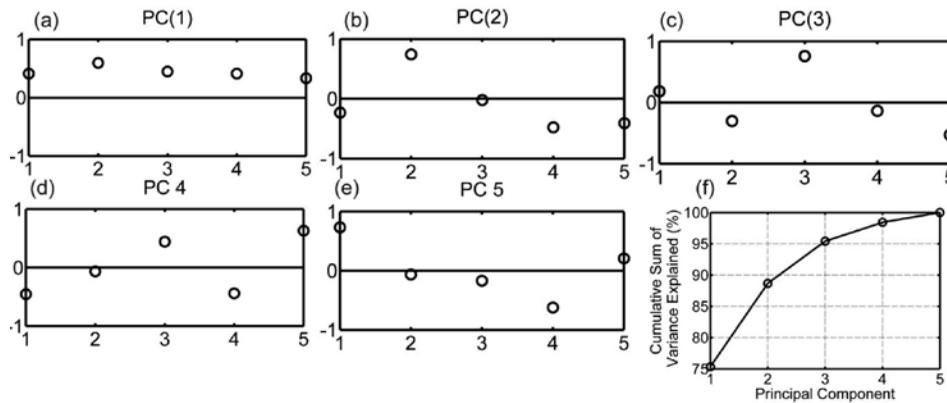


Figure 8. (a-e) PCA plot displays the impact of each variable (CWT coefficient series at scale 16 of various mother wavelets of gamma ray log of pilot hole) for modelling of well log signal. In the x-axis, position of Haar, Gaus 1, Gaus 3, Db2 and Morlet wavelet is represented by 1, 2, 3, 4, and 5 respectively. Y-axis denotes the impact of each wavelet and the value is normalized between -1 and +1. (f) The PCA analysis of well log data from cumulative sum of variance is explained by the different PCs.

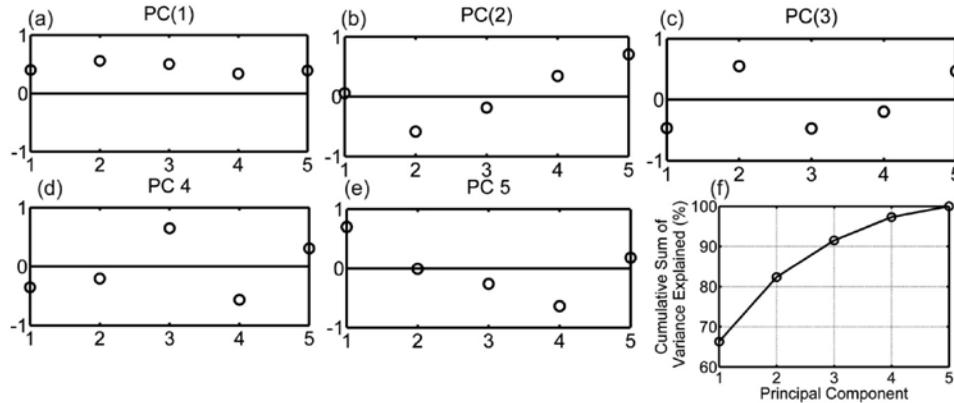


Figure 9. (a-e) PCA plot displays the impact of each variable (CWT coefficient series at scale 32 of various mother wavelets of gamma ray log of pilot hole) for modelling of well log signal. In the x-axis, position of Haar, Gaus 1, Gaus 3, Db2 and Morlet wavelet is represented by 1, 2, 3, 4, and 5 respectively. Y-axis denotes the impact of each wavelet and the value is normalized between -1 and +1. (f) The PCA analysis of well log data from cumulative sum of variance is explained by the different PCs.

As seen in PC (2) panel, Gaus1 and Morlet have significant roles but they are in opposition to each other. PC(3) demonstrates the dominance of Gaussian wavelets (order 1 and 2) but their role is in opposite direction. On the other hand, PC(4) shows that the impact and direction of influence of Db2 and Morlet are distinct. Figure 8f recommends that 89% variance of the downhole data are explained by first two PCs. Thus, the multivariate data set can be described using only two coordinate axes. The PCA based significant role among Haar, Gaus 1, Gaus 3, Db2 and Morlet at scale 16 for the depth 28-4000 m interval of gamma ray log is demonstrated in figures 8a-e.

Figures 9(a-e) show the role of each PC corresponding to each mother wavelet for modeling well log at scale 32. Figures 9a-e suggest that the Gaus 1 plays a better role in modeling gamma ray in panel PC (1), although the mother wavelets have significant positive role in PC (1). In PC(2), Gaus1 and Morlet have central role in opposite direction.

In PC(3), Gaus1 and Gaus3 have significant role in opposite direction. In PC (4), Gaus3 and Db2 have significant role in opposite direction. In PC(5) Haar and Morlet play key roles in opposite direction. According to figure 9f, 82% variance of the downhole data are explained by the first two PCs. Thus, the multivariable data set can be described using only two coordinate axes. The PCA-based key role among Haar, Gaus 1, Gaus 3, Db2 and Morlet at scale 32 for the depth 28-4000 m interval of gamma ray log is demonstrated in figures 9a-e.

Figures 10(a-e) demonstrate the role of PC corresponding to each mother wavelet for modeling well log at scale 64 suggesting that the Gaus 1 shares slightly persuasive role in modeling gamma ray in panel PC(1), although, mother wavelets have significant positive involvement in PC(1). Gaus 1 and Morlet have persuasive role in PC(2) in opposite direction. Similarly in PC(3), Gauss 1 and Gaus 3 have credible roles in opposite direction. Further, one can

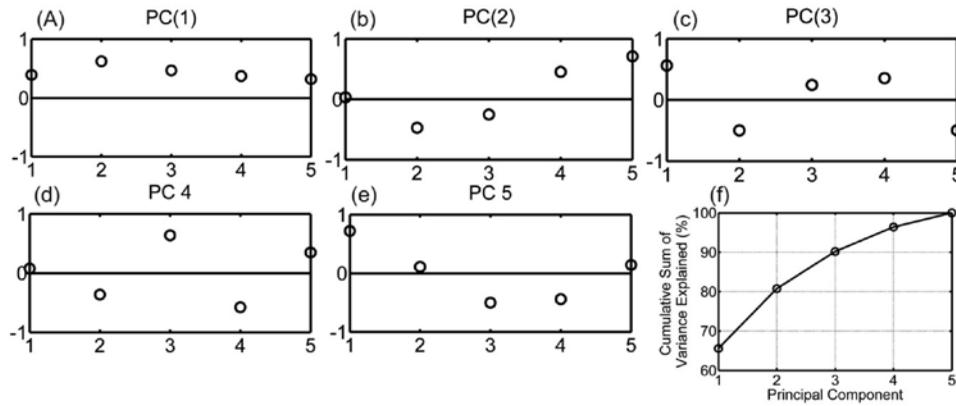


Figure 10. (a-e) PCA plot displays the impact of each variable (CWT coefficient series at scale 64 of various mother wavelets of gamma ray log of pilot hole) for modelling of well log signal. In the x-axis, position of Haar, Gaus 1, Gaus 3, Db2 and Morlet wavelet is represented by 1, 2, 3, 4, and 5 respectively. Y-axis denotes the impact of each wavelet and the value is normalized between -1 and +1. (f) The PCA analysis of well log data from cumulative sum of variance is explained by the different PCs.

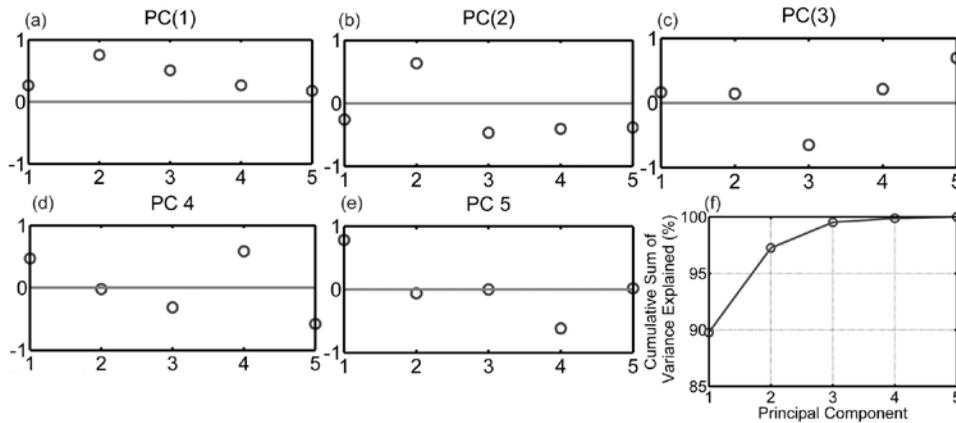


Figure 11. (a-e) PCA plot demonstrates the impact of each variable (CWT coefficient series at scale 2 of various mother wavelets of density log of main hole) for modelling of well log signal. In the x-axis, position of Haar, Gaus 1, Gaus 3, Db2 and Morlet wavelet are represented by 1, 2, 3, 4, and 5 respectively. Y-axis denotes the impact of each wavelet and the value is normalized between -1 and +1. (f) The PCA analysis of well log data from cumulative sum of variance is explained by the different PCs.

see that in PC(4), Gauss 3 and Db 2 have credible roles in opposite direction. In PC (5), Haar and Morlet have credible roles in opposite direction. Figure 10f proposes that 80% variance of the downhole data are explained by first two PCs.

Main hole data

Figures 11a-e reveals the role of each PC corresponding to each mother wavelet for modeling well log at scale 2. Figure 11 suggests that the Gaus1 shares relatively cogent role in modeling gamma ray in panel PC(1), although, mother wavelets have significant positive involvement in PC(1). In PC(2), Gaus 1 and Gaus 3 have cogent role in opposite direction. In PC(3), Gaus 3 and Morlet have cogent roles in opposite direction. In PC(4), Morlet and Db(2) have main role in opposite direction. In PC(5) Haar has a sound role opposed to Db 2

Figure 11f reveals that 97% variance of the downhole data are explained by first two PCs. The PCA-based dominant role among Haar, Gaus 1, Gaus 3, Db2 and Morlet at scale 2 for the depth 3000-7000 m interval of density log is demonstrated in figures 11a-e. Figures 12a-e demonstrate the role of each PC corresponding to each mother wavelet for modeling well log at scale 4.

Figures 12(a-e) suggest that the Gaus 1 shares relatively sound role in modeling density log in panel PC(1), although, mother wavelets have significant positive involvement in PC(1). PC(2) panel shows the influence of Gaus 1 and Morlet are distinct but they are opposite in direction. PC(3) exhibits convincing role of Gaus 3 and Morlet wavelet although they are found in opposite direction. In PC (4), Morlet and Db2 have principal role in opposite direction. In PC(5) Haar has central role opposite to Db2. Figure 12f advocates that 95% variance of the downhole data are explained by first two PCs. The PCA-based convincing role among Haar, Gaus 1,

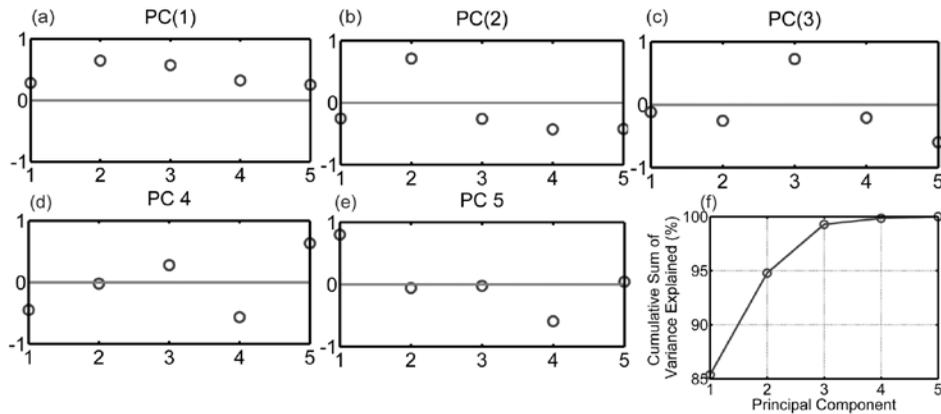


Figure 12. (a-e) PCA plot demonstrates the impact of each variable (CWT coefficient series at scale 4 of various mother wavelets of density log of main hole) for modelling of well log signal. In the x-axis, position of Haar, Gaus 1, Gaus 3, Db2 and Morlet wavelet are represented by 1, 2, 3, 4, and 5 respectively. Y-axis denotes the impact of each wavelet and the value is normalized between -1 and +1. (f) The PCA analysis of well log data from cumulative sum of variance is explained by the different PCs.

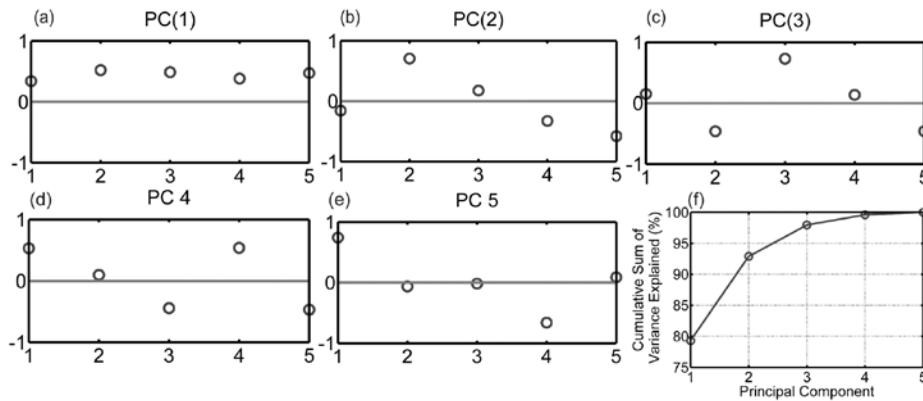


Figure 13: (a-e) PCA plot demonstrates the impact of each variable (CWT coefficient series at scale 8 of various mother wavelets of density log of main hole) for modelling of well log signal. In the x-axis, position of Haar, Gaus 1, Gaus 3, Db2 and Morlet wavelet are represented by 1, 2, 3, 4, and 5 respectively. Y-axis denotes the impact of each wavelet and the value is normalized between -1 and +1. (f) The PCA analysis of well log data from cumulative sum of variance is explained by the different PCs.

Gaus 3, Db 2 and Morlet at scale 4 for the depth 3000-7000 m interval of density log is demonstrated in figures 12a-e.

Figures 13(a-e) exhibit the role of each PC corresponding to each mother wavelet for modeling well log at scale 8. Figures 13a-e suggest that the Gaus 1 shares slightly convincing role in modeling density in panel PC(1), although, mother wavelets have significant positive involvement in PC(1). In PC(2), Gaus 1 and Morlet have strong role in opposite direction. In PC (3), Gaus 1 and Gaus 3 have strong role in opposite direction. In PC (4), Gaus 3 and Db 2 have strong role in opposite direction. In PC(5), Haar and Db2 have strong role in opposite direction. Figure 13f submits that 92% variance of the downhole data are explained by first two PCs. The PCA based strong role among Haar, Gaus1, Gaus3, Db2 and Morlet at scale 8 for the depth 3000-7000 m interval of gamma ray log is demonstrated in figures 13a-e.

Figures 14(a-e) divulge the role of each PC corresponding to each mother wavelet for modeling well log at scale 16. Figures 14a-e suggest that the Gaus 1 shares slightly strong role in modeling density in panel PC(1), although, mother wavelets have significant positive involvement in PC(1). As seen in PC(2) panel, Gaus 1 and Morlet play major role to characterize the signal but their influence is found in opposite direction. PC(3) demonstrates dominance of Gaussian wavelets although they are acting in opposite direction, to match with the signal. It is evident from panel PC(4) that Db2 dominates over Morlet wavelet but the role of action is in opposite direction. PC(5) guides that Haar and Db2 can be more influential than the rest of the wavelets used.

Figure 14f advocates that 90% variance of the downhole data are explained by first two PCs. The PCA-based dominant role among Haar, Gaus 1, Gaus 3, Db2

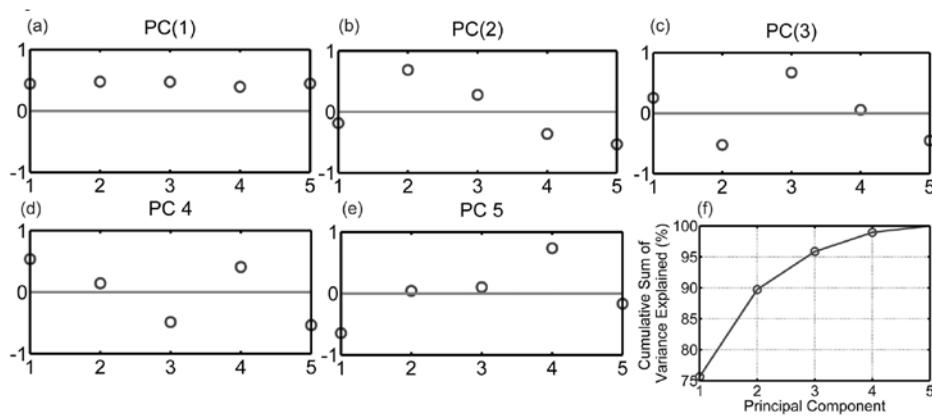


Figure 14: (a-e) PCA plot demonstrates the impact of each variable (CWT coefficient series at scale 16 of various mother wavelets of density log of main hole) for modelling of well log signal. In the x-axis, position of Haar, Gaus 1, Gaus 3, Db2 and Morlet wavelet are represented by 1, 2, 3, 4, and 5 respectively. Y-axis denotes the impact of each wavelet and the value is normalized between -1 and +1. (f) The PCA analysis of well log data from cumulative sum of variance is explained by the different PCs.

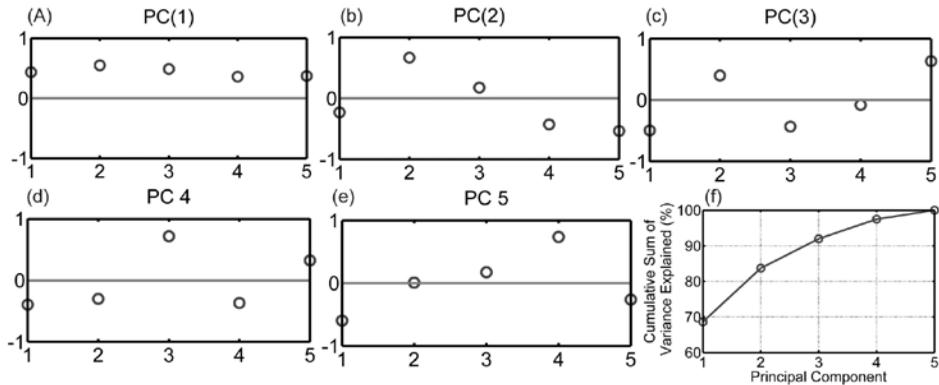


Figure 15. (a-e) PCA plot demonstrates the impact of each variable (CWT coefficient series at scale 32 of various mother wavelets of density log of main hole) for modelling of well log signal. In the x-axis, position of Haar, Gaus 1, Gaus 3, Db2 and Morlet wavelet are represented by 1, 2, 3, 4, and 5 respectively. Y-axis denotes the impact of each wavelet and the value is normalized between -1 and +1. (f) The PCA analysis of well log data from cumulative sum of variance is explained by the different PCs.

and Morlet at scale 16 for the depth 3000-7000 m interval of density log is demonstrated in figures 14a-e.

Figures 15(a-e) suggest the role of each PC corresponding to each mother wavelet for modeling well log at scale 32. Figures 15a-e imply that the Gaus1 shares slightly forceful role in modeling gamma ray in panel PC(1), although, mother wavelets have significant positive contribution in PC(1). In PC(2), Gaus1 and Morlet have powerful role in opposite direction. In PC(3), Morlet and Gaus 3 have clear role in opposite direction. In PC (4), Gaus 3 and Db2 have clear role in opposite direction. In PC (5) Haar and Morlet have key role in opposite direction. Figure 15f proposes that 82% variance of the downhole data are explained by first two PCs. Thus the multivariate data set can be described using only two coordinate axes. The PCA based central role among Haar, Gaus 1, Gaus 3, Db2 and Morlet at scale 32 for the depth 3000-7000 m interval of density log is demonstrated in figures 15a-e.

Figures 16a-e demonstrate the role of each PC corresponding to each mother wavelet for modeling well log at scale 64. Figures 16a-e suggest that the Gaus 1 shares slightly central role in modeling gamma ray in panel PC(1), although, mother wavelets have significant positive involvement in PC(1). In PC(2), Gaus1 and Morlet have key role but in opposite direction. In PC(3), Gaus 3 and Morlet have key role in opposite direction. In PC (4), Gaus 3 and Db2 have key role in opposite direction. In PC (5) Haar and Morlet have key role in opposite direction. Figure 16f recommends that 74% variance of the downhole data are explained by first two PCs. The PCA-based key role among Haar, Gaus 1, Gaus 3, Db2 and Morlet at scale 64 for the depth 3000-7000 m interval of density log is demonstrated in figures 16a-e.

The scalogram plots corresponding to mother wavelet Haar, Gaus 1, Gaus 3, Db2 and Morlet of gamma ray log of KTB pilot hole shown in figure 3 demonstrate that

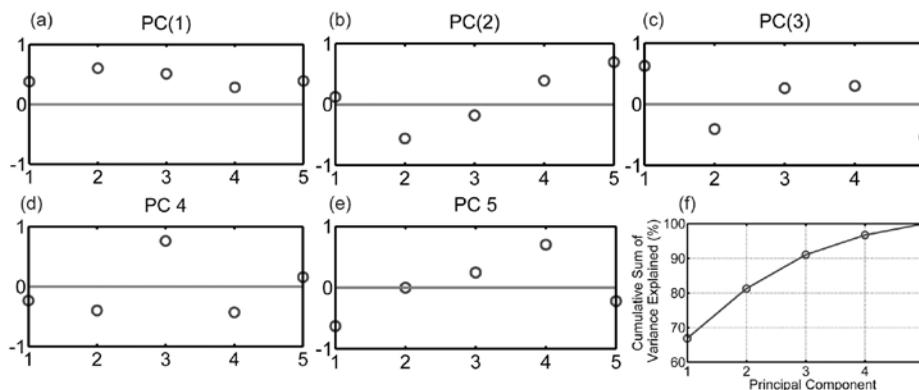


Figure 16. (a-e) PCA plot demonstrates the impact of each variable [CWT coefficient series at scale 64 of various mother wavelets of density log of main hole) for modelling of well log signal. In the x-axis, position of Haar, Gaus 1, Gaus 3, Db2 and Morlet wavelet are represented by 1, 2, 3, 4, and 5 respectively. Y-axis denotes the impact of each wavelet and the value is normalized between -1 and +1. (f) The PCA analysis of well log data from cumulative sum of variance is explained by the different PCs.

Table 2. Calculating depths of formation tops from the scalogram plots of the gamma ray intensity log in pilot hole data by using Haar, Gaus1, Gaus3, Morlet and Db2 wavelets.

Actual depths of the formation top obtained from log data (m)	Depths to the formation top by using different wavelets from scalogram plots (m)				
	Haar	Gaus1	Gaus3	Db2	Morlet
60.04	60.04	59.74	59.89	59.89	58.82
112.77	112.31	112	110.49	111.55	Not clear
196.44	196.13	Not clear	Not clear	195.98	Not clear
303.58	303.58	302.51	302.97	303.12	Not clear
336.65	336.65	335.73	334.67	336.04	Not clear
511.45	511.45	510.84	Not clear	510.08	Not clear
1150.01	1150.01	1150.01	1149.85	1150.01	Not clear
1597	1597	1597	1596.84	1596.84	1596.69
2683.61	2683.30	2683.15	2683	2683.30	Not clear
2733.29	2733.14	2732.83	2732.53	2783.83	2732.68
3564.02	3564.02	3564.02	3563.72	3563.87	3563.56

the formation boundary are well resolved grossly against the log response changes. The detailed comparison of identified formation boundary for KTB pilot hole and main hole by using CWT-based scalograms of various mother wavelets is presented in the tables 2 and 3 respectively. Overall, the boundary picked by the different scalograms of corresponding mother wavelet is found to be satisfactory and is well tuned with the boundary detected from the changes of the well log at the KTB site (Table 2 and 3). Looking at the scalogram figures, Gaus 1 gives better resolution (spectral bands are more distinct) in comparison to the rest of the mother wavelets used, however, it also provides thicker spectral distribution at the sharp boundary of the bed. In the case of sharp bed boundary picking, Haar-based scalogram gives spectral distribution, which is relatively sharp at the boundary (Figure 3 and 4). The blocky kernel functions of the Haar wavelet, which helps

to find the depth of the boundary with less smearing of spectral bands, is appropriate to model blocky signal at the boundary.

Conversely, Gaus 1 provides attractive structures for the interpreter for the overall signal matching for pay zone/litho-facies zone because it captures over all changes of the spectral band and or/distribution of CWT coefficients at different scales. However, comparatively poor resolution of non-Gaussian wavelets is evident because there is more miss-match of shape between mother wavelet and well log gross signal. Moreover, shifts and dilations parameters and their averaging at the boundary could be the cause for poor resolution. Therefore, in many of the cases, there is no clarity of boundary detection (table 2 and 3). The distribution and changes from positive to negative and vice versa of CWT coefficients over the colour panel clearly demonstrates the lithology/litho-facies changes at the KTB

Table 3. Calculating depths of formation tops from the scalogram plots of the density log in main hole data by using Haar, Gaus1, Gaus3, Db2 and Morlet.

Actual depths of the formation top obtained from log data (m)	Depths to the formation top by using different wavelets from scalogram plots (m)				
	Haar	Gaus1	Gaus3	Db2	Morlet
3409.79	3409.79	3409.64	3409.49	3409.18	3410.10
3546.34	3546.34	3546.19	3545.89	3546.19	3545.12
4000.51	4007.81	4007.51	4007.05	4008.12	4008.12
5204.61	5204.61	5204.61	5204.91	5204.15	5204.15
5773.21	5773.21	5773.21	5773.52	5773.82	Not clear
6007.30	6007.30	6007.30	Not clear	6007.30	6006.69
6436.46	6436.46	6436.31	Not clear	6436.31	6435.85
6566.30	6566.45	6566.30	6566.76	6566.45	6566.00
6709.10	6709.10	6709.10	6708.95	6709.25	6708.80
6882.38	6882.38	6882.38	6882.23	6882.53	6882.07

site. The boundary detected by the scalogram analysis is closely matching with the boundary obtained by the KTB research team using multiple methods (Pechnig et al., 1997; Maiti and Tiwari, 2009, 2010a-b; Singh et al., 2016).

CONCLUSIONS

In this paper, CWT based wavelet analysis has been carried out and scalogram was prepared using various mother wavelets to identify the very complex formation interfaces using the KTB well log data. For this histogram analysis, statistical significance test and PCA-based analysis were performed to understand how each CWT coefficient plays role in overall signal modeling and characterizing the KTB well log data. These results will be useful to provide future guidelines for selecting appropriate wavelet functions for such complex well log signal modeling.

From the above study, we reached at the following conclusions.

Histogram analysis shows that the CWT coefficients of Gaus 1 occur maximum times in the entire well log signals of KTB pilot and main hole. Hence, Gaus 1 is found to be the relatively best mother wavelet for CWT- based scalogram analysis of well log signal. Results after significance test demonstrate that spatial series at even number scale are statistically significant among all mother wavelets. PCA- based analysis suggests that CWT coefficient of Gaus 1 plays a major role in most of the cases at PC (1) and PC (2) for both pilot and main hole well log data.

ACKNOWLEDGEMENTS

AS and SM are gratified to Director, IIT (ISM), Dhanbad for permitting to publish the work. AS is thankful to IIT(ISM)

for SRF fellowship. Financial support from the Ministry of Earth Sciences, Govt. of India, New Delhi (Grant No: MoES/P.O. (Geosci)/44/2015) is acknowledged. RKT is grateful to DAE for award of RRR. We are thankful to Dr. Gautam Gupta, Indian Institute of Geomagnetism for meticulous review, which has greatly improved the paper. Authors are also thankful to Dr.M.R.K.Prabhakar Rao for constructive editing of the manuscript. We thank the Chief Editor for his support.

Compliance with Ethical Standards

The authors declare that they have no conflict of interest and adhere to copyright norms.

REFERENCES

Adamowski, J., and Chan, H.F., 2011. A wavelet neural network conjunction model for groundwater level forecasting, *J. Hydrol.*, v.407, pp: 28-40.

Chandrashekhar, E., and Rao, V.E., 2012. Wavelet analysis of Geophysical Well- log Data of Bombay offshore Basin, *India Math. Geosci.*, v.44, pp: 901-928.

Daubechies, I., 1992. Ten lectures on wavelets, Soc. for Ind. and Appl. Math. Philadelphia, Pa. pp: 357.

Doveton, J.H., 1986. Log Analysis of Subsurface Geology Concepts and Computer Methods Wiley, New York, NY., pp: 273.

Farge, M., 1992. Wavelet transforms and their applications to turbulence, *Annu. Rev. Fluid Mech.*, v.24, pp: 395-457.

Goupillaud, R.A., Grossmann, A., and Morlet, J., 1985. Cycle-octave and related transform in seismic signal analysis, *Georex.*, v.23 pp: 85-102.

Javid, M., and Tokhmechi, B., 2012. Formation interface detection using Gamma Ray log: A novel approach, *J. Mining & Env.*, v.3, no.1, pp: 41-50.

- Jolliffe, I.T., 1972. Discarding variables in a principle components analysis 1: Artificial Data, *Appl. Stat.*, v.21, pp: 160-173.
- Kisi, O., 2009. Wavelet regression model as an alternative to neural networks for monthly streamflow forecasting, *Hydrol. Process.*, v.23, no.25, pp: 3583-3597.
- Kumar, P., and Foufoula-Georgiou, E., 1997. Wavelet analysis for geophysical applications Review. *Geophys.*, v.35, no.4, pp: 385-412.
- Labat, D., 2008. Wavelet analysis of the annual discharge records of the world's largest rivers, *Adv. Water Resour.*, v.31, no.1, pp: 109-117.
- Liu, Z., Zhou, P., Chen, G., and Guo, L., 2014. Evaluating a coupled discrete wavelet transform and support vector regression for daily and monthly stream flow forecasting, *J. Hydrol.*, v.519, pp: 2822-2831.
- Maheswaran, R., and Khosa, R., 2012. Comparative study of different wavelets for hydrologic forecasting, *Comput. Geosci.*, v.46, pp: 284-295.
- Maiti, S., and Tiwari, R. K., 2009. A hybrid Monte Carlo method based artificial neural networks approach for rock boundaries identification: A case study from the KTB bore hole, *Pure and Appl. Geophys.*, v.166, pp: 2059-2090.
- Maiti, S., and Tiwari, R. K., 2010a. Automatic discriminations among geophysical signals via the Bayesian neural networks approach, *Geophys.*, v.75, no.1, pp: E67-E78.
- Maiti, S., and Tiwari, R. K., 2010b. Neural network modeling and an uncertainty analysis in Bayesian framework: A case study from the KTB borehole site, *J. Geophys. Res.*, B10208, v.115.
- Maraun, D., Kurths, J., and Holschneider, M., 2007. Nonstationary Gaussian processes in wavelet domain: synthesis, estimation, and significance testing. *Physical Rev. E*, v.75, pp: 016707.
- Morlet, J., Arens, G., Fourgeau, E., and Giard, D., 1982. Wave propagation and sampling theory-part I:Complex signal and scattering in multilayered media *Geophysics.*, v.47, no.2, pp: 203-221.
- Nourani, V., Baghanam, A.H., Adamowski, J., and Kisi, O., 2014. Applications of hybrid wavelet-artificial intelligence models in hydrology: A review *J. Hydrol.*, v.514, pp: 358-377.
- Nourani, V., Kisi, O., and Komasi, M., 2011. Two hybrid artificial intelligence approaches for modelling rainfall-runoff process, *J. Hydrol.*, v.402, no.1-2, pp: 41-59.
- Ojha, M., and Maiti, S., 2016. Sediment classification using neural networks: an example from the site-U1344A of IODP Expedition 323 in the Bering Sea Deep-Sea Research Part II: Topical Studies in Oceanography, v.125-126, pp: 202-213.
- Pan, S., Hsieh, B., Lu, M., and Lin, Z., 2008. Identification of stratigraphic formation interfaces using Wavelet and Fourier transform *Comput. Geosci.*, v.34, pp: 77-92.
- Pechinig, P., Haverkamp, S., Wohlenberg, J., Zimmermann, G., and Burkhardt, H., 1997. Integrated interpretation in the German Continental Deep Drilling Program: Lithology, porosity, and fracture zones. *J. Geophys. Res.*, v.102, pp: 18363-18390.
- Percival, D., and Walden, A. T., 2000. Wavelet methods for time series analysis Cambridge University Press, Cambridge, U.K.
- Perez-Munoz, T., Velasco-Hernandez, J., and Hernandez-Martinez, E., 2013. Wavelet transform analysis for lithological characteristics identification in siliciclastic oil fields, *J. Appl. Geophys.*, v.98, pp: 298-308.
- Prokoph, A., and Agterberg, F. P., 2000. Wavelet analysis of well-logging data from oil source rock, Egret Member offshore eastern Canada *Am Assoc. Pet. Geol. Bull.*, v.84, pp: 1617-1632.
- Sang, Y.F., Singh, V.P., Sun, F., Chen, Y., Liu, Y., and Yang, M., 2016. Wavelet-based hydrological time series forecasting *J. Hydrol. Eng.*, v.21, no.5, pp: 06016001.
- Schaeffli, B., Maraun, D., and Holschneider, M., 2007. What drives high flow events in the Swiss Alps? Recent developments in wavelet spectral analysis and their application to hydrology *Adv. Water Resour.*, v.30, no.12, pp: 2511-2525.
- Shoaib, M., Shamseldin, A.Y., and Melville, B.W., 2014. Comparative study of different wavelet based neural network models for rainfall runoff modelling, *J. Hydrol.*, v.515, pp: 47-58.
- Singh, A., Maiti, S., and Tiwari, R.K., 2016. Modelling discontinuous well log signal to identify lithological boundaries via wavelet analysis: An example from KTB borehole data, *J. Earth Sys. Sci.*, v.125, no.4, pp: 761-776.
- Singh, R.M., 2011. Wavelet-ANN model for flood events *Adv. Intel. Software Comput.*, v.131, pp: 165-175.
- Strang, G., and Nguyen, T., 1995. Wavelets and filter banks. Wellesley-Cambridge Press, Prentice Hall.
- Torrence, C., and Compo, G.P., 1998. A practical guide to wavelet analysis *B. Am. Meteorol. Soc.*, v.79, no.1, pp: 61-78.
- Walker, J.S., 1999. A primer on wavelets and their scientific applications, Chapman and Hall, CRC, New York.