Computers & Geosciences 88 (2016) 132–141

Contents lists available at ScienceDirect

Computers & Geosciences

journal homepage: www.elsevier.com/locate/cageo

Case study

Using an improved BEMD method to analyse the characteristic scale of aeromagnetic data in the Gejiu region of Yunnan, China

Jie Zhao*, Pengda Zhao, Yongqing Chen

China University of Geosciences (Beijing), 29 Xueyuan Road, Haidian District, Beijing 100083, China

ARTICLE INFO

Article history:
Received 12 April 2015
Received in revised form
17 December 2015
Accepted 18 December 2015
Available online 7 January 2016

Keywords:
Intrinsic mode function
Empirical mode decomposition
Characteristic scale
Metallogenic prediction
HHT

ABSTRACT

The geological and metallogenic process is a typical non-stationary multifactor and multi-scale random process. Multiple measurement data assess the performance of the integrated process, and the combined data set is usually large and complex, among other characteristics. When different metallogenic prediction targets exist, the data must be decomposed on different scales in space. The study of the scale interval in which the object features are located can eliminate useless information and retrieve useful scale data that are needed for metallogenic prediction. Thus, the model that the specific deposit presents will be rapidly and accurately identified to enhance the efficiency of the prediction and analysis models. This paper employs an improved bidimensional empirical decomposition method to decompose aeromagnetic survey data and expresses and decomposes the spatial distribution of deposits with a mixed Gaussian model. By comparing the decomposition results on various sampling data scales with the distribution function for the deposit, the characteristic scale interval that contains the measurement information that exhibits the greatest similarity to the distribution of the deposits can be identified. This method was employed to analyse a Yunnan Gejiu tin–copper polymetallic deposit using aeromagnetic sampling data to calculate suitable decomposition-scale parameters. This approach provides valuable parameters for metallogenic prediction in other areas with aeromagnetic data.

1. Introduction

Geological processes are typical, non-stationary and complex temporal and spatially variable processes (Cheng, 2003; Turcotte, 1997; Bansal and Dimri, 2005a, 2005b, 2014). A basic task of geological research is modelling this process to quantitatively and accurately describe it. Due to the extreme complexity of the entire process, reproduction of the whole geological process by establishing a single or several mathematical models is not feasible. A feasible approach is to address subproblems and subprocesses in a fractionised domain and to build a concrete model (Zhao, 1982, 2002; Zhao et al., 1994).

Geological data consists of measurement signals and a non-stationary random process that are formed by the interaction of various modulation mechanisms (Diks, 1999; Huang et al., 1996; Widrow and Stearns, 1985). Scale decomposition and scale selection of geological signals are the primary focus of this paper.

The decomposition and reconstruction of signals are classic problems. Frequency analysis tools and scale analysis tools can completely expand the structures of a signal; depending on their application, they can be employed for selecting and re-organising information to obtain a new desired signal (Cheng, 2004). This study employs the HHT as the analysis tool. The basic principle of the HHT is to decompose a signal based on the symmetry of the signal at each scale. Compositions that are decomposed from each characteristic scale are referred to as the intrinsic mode function (IMF). For each IMF, the Hilbert transform was employed and the instantaneous frequency spectrum was calculated. Then, the instantaneous spectrum of each component was combined, and information about the signal frequency structure was obtained. Therefore, the HHT is an integrated decomposition tool in the space-wavenumber domain (Kantz and Schreiber, 1997).

In recent years, there have been many applications of the HHT transform in geosciences researches (Chen and Zhao, 2011, 2012; Hou et al., 2012; Huang et al., 2010; Jian et al. 2012); however, many researchers use it as a conventional decomposition tool for sample data processing, obtaining features from various scales, and using these characteristics as parameters or input in different mathematical models. This is only a basic application of the HHT. For the former applications of the HHT for geosciences data processing is a simple signal decomposition, neither the frequency decomposition mechanism of the BEMD is discussed, nor the self-adaption of the signal analysis is made by the mechanisms, a deeper understanding of the HHT, the analysis of its mathematical

* Corresponding author.
E-mail address: dayinggezhuang@163.com (J. Zhao).

http://dx.doi.org/10.1016/j.cageo.2015.12.016
0098-3004/

properties and how to more deeply and reasonably apply the characteristics of the HHT in the geosciences are problems that remain to be studied.

In this paper, the basic principles of the HHT (Han et al., 2002; Huang, 2005, 2006) are described and its scale decomposition function is analysed. The effective scale decomposition ability of the HHT is employed to reduce a signal to an integrated multi-scale channel and identify components that are closest to the specific metallogenic mode at each scale, which is the scale interval in which the specific metallogenic mechanism serves a role. Once a scale parameter has been determined, it can be employed as a reference for metallogenic prediction or ore-forming process analyses in other locations, which is performed by directly applying the data within the decomposed characteristic scale to predict the metallogenic process using pattern recognition or feature comparison. Calculation and analysis based on decomposition data in the characteristic scale interval produces not only higher accuracy but also greater efficiency (Bedrosian, 1963; Flandrin and Gonçalves, 2004; Flandrin et al., 2004, 2005; Huang, 2001; Huang et al., 1998; Bendat, 1990).

2. HHT Analysis

The HHT method is a time-domain decomposition and frequency spectrum analysis method that was proposed at the end of the 1990s. It has the characteristics of simple calculation, strong adaptability, instantaneous spectrum structure and strong robustness. It is a new signal processing mode of decomposition and a breakthrough of various types of analysis tools based on the Fourier transform.

The HHT method employs a multi-scale analysis framework that is similar to the wavelet analysis (Schlurmann, 2002); however, it does not consider the signal frequency as the basic object of calculation. Its scale decomposition object is the symmetry of the vibration on the time domain. This change abandons the convolution calculation, which is generally adopted by the Fourier transform and wavelet analysis; it prevents the poor adaptability that is caused by the pre-setting convolution kernel and is more flexible and stable. Correspondingly, the disadvantage of this decomposition is that does not have a clear boundary in the frequency domain. For many applications, the scale analysis is more important than the frequency analysis; thus, the HHT becomes very suitable for these cases. For example, a substantial amount of geological data that corresponds to the stochastic processes are very suitable via use of the HHT. The majority of geological signals are complex, are composed of a plurality of spectrum types, are non-stationary and nonlinear (Tong, 1990; Wu et al., 2007), and are combined with many local time variation processes. Use of a highly adaptive and multiscale analysis tool, such as the HHT, which is directly based on the spatial form, to decompose signals with these characteristics is more suitable.

Numerous studies have applied HHT analysis to geoscience problems (Coughlin and Tung, 2004; Coughlin and Tung, 2005; Datig and Schlurmann, 2004; Duffy, 2004; Han et al., 2002; Huang and Attoh-Okine, 2005; Huang and Shen, 2005; Huang et al., 2001; Komm et al., 2001; Nuttall, 1966; Schlurmann, 2002; Zhang, 2006).

The HHT involves two parts: the first part is an iterative sifting process, in which an original signal is decomposed into a trend term and a plurality of the component function, which is referred to as an intrinsic mode function (IMF) (Wu and Huang, 2005). Each IMF has the characteristics of an approximate zero mean and envelope symmetry. This decomposition process is known as empirical mode decomposition (EMD). The second part uses the Hilbert transform to decompose the instantaneous frequency for each IMF to obtain instantaneous spectra of the decomposed signal (Chen et al., 2006; Huang et al., 2008). As the EMD algorithm guarantees the symmetry of each IMF signal and the zero mean characteristics, the IMF is very suitable for the Hilbert transform; some common defects of the Hilbert transform would not appear. These two parts are combined to constitute the HHT, in which the function of the first part is the scale decomposition and the role of the second part is the frequency analysis.

2.1. EMD decomposition

First, we discuss empirical mode decomposition. The decomposition rules of EMD imply some assumptions and property requirements of the HHT for signals: (1) the signal should be a concussion signal, in which minimum and maximum values can be applied to the decomposition procedure, and signals usually have multiple concussion periods, so that the analysis has practical significance; (2) the time interval between the extreme points and the amplitude of the vibration are important scale characteristics of the signal; however, the IMF enables variation in the two characteristic scales. A relatively stable characteristic scale indicates the high probability of a single vibration source; and (3) if the signal data have no extreme points but have inflection points, an extreme point can be obtained by differentiating the signal data and integrating it to obtain the decomposition results (Xu et al., 2006).

For the original signal $f(x, y)$, the specific sifting process of the EMD method is as follows: first, initialise the residual function $res(x, y)$ and the current step signal $f_0(x, y)$. Let $res(x, y) = f(x, y) = f_0(x, y)$. Second, calculate all extreme points of the current signal $f_0(x, y)$: employ a double three-spline interpolation function to calculate the upper envelope $up(x, y)$ of all local maximum points and employ this same method to obtain the lower envelope $low(x, y)$ of all local minimum points. Third, obtain the average surface $mean(x, y) = (up + low)/2$ of the two envelopes. The mean represents the development trend of the data. Fourth, the trend from the original signal. Let the remaining part be the signal of the next step, that is, $f_1(x, y) = f_0(x, y) - mean(x, y)$. Last, determine the relationship between the current step signal f_1 and the previous step signal f_0 by placing them into the stop condition. If the stop condition is not satisfied, repeat this fitting and subtracting of mean processes to obtain $f_2, f_{3,\ldots}$ until the stop condition is satisfied. The stop criterion serves a critical role in BEMD; it determines the scale, which resembles a frequency band, for which the current pass of decomposition generates relatively stable components. Additional levels are obtained for smaller thresholds. Let the current step signal be f_n. Then, f_n is the minimum scale IMF of the first layer that is sifted, which is denoted as imf_1. Remove the decomposed intrinsic mode function imf_1 from the original signal and re-initialise the remaining part as a new residual function and a current step signal, i.e., $res(x, y) = f_0(x, y) = f(x, y) - imf_1$. Repeat this sifting process for IMF2, IMF3,…, and the final residual function $res(x, y) = f_0(x, y) = f(x, y) - imf_1$ is obtained. Their relationship is

$$f(x, y) = \sum_{i=1}^{n} imf_i(x, y) + res(x, y).$$

A flow chart of the EMD decomposition process is shown in Fig. 1.

2.2. Hilbert Transform

Consider a one-dimensional function as an example to introduce the Hilbert transform, which analyses the instantaneous frequency of the decomposed IMFs. Let the original signal be $x(t)$. The Hilbert transform is the convolution of the signal and $(\pi t)^{-1}$:

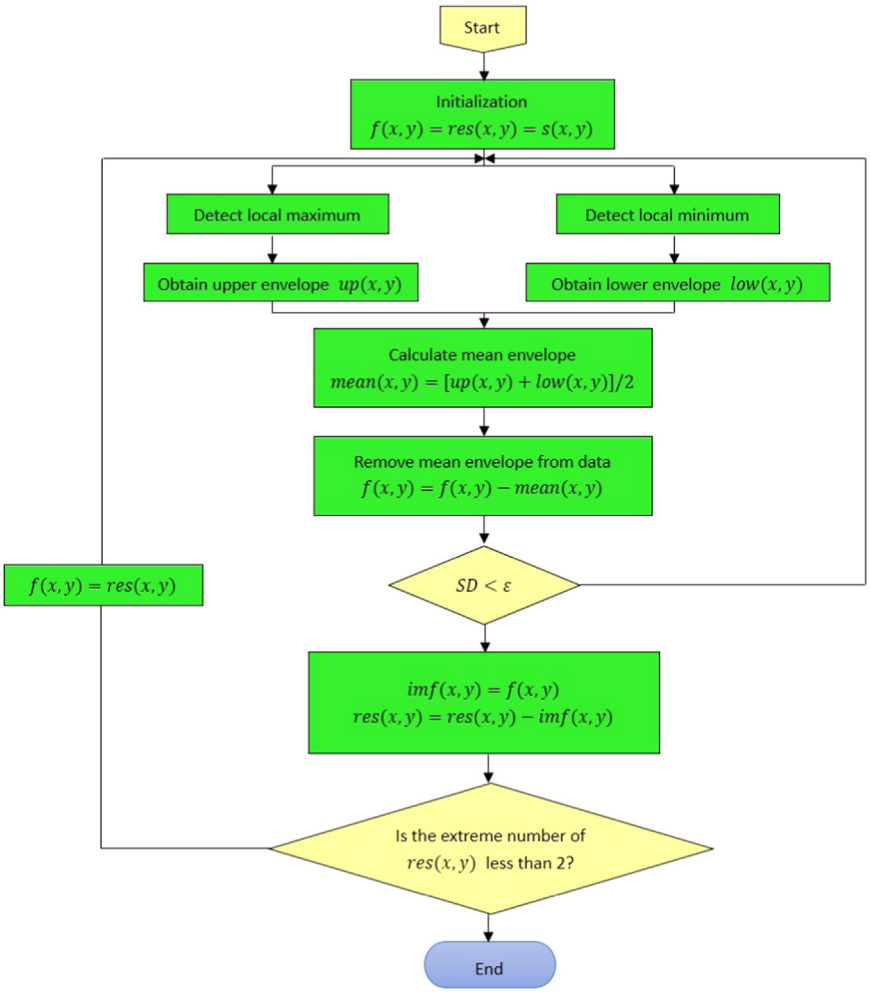

Fig. 1. EMD decomposition process.

$$H(x) = x(t)*\left(\frac{1}{\pi t}\right) = \frac{1}{\pi} p.v. \int_{-\infty}^{+\infty} \frac{x(\tau)}{t-\tau} d\tau.$$

This is the Cauchy principal value integral.

As $x(t)$ is a real signal, the real part of the frequency in the frequency domain is evenly symmetric, and the imaginary part of the frequency is oddly symmetric in the frequency domain. The frequency characteristic of the signal satisfies

$$F(\omega) = F(x(t)) = \begin{cases} \mathrm{Re}(\omega) + j\mathrm{Im}(\omega) & \omega > 0 \\ \mathrm{Re}(\omega) - j\mathrm{Im}(\omega) & \omega < 0 \end{cases},$$

where denotes the real part of its argument and Im represents the image part of its argument, is the complex number, and

$$j = \sqrt{-1}.$$

As the frequency characteristic of the Hilbert transform is

$$H(\omega) = F[H] = -i\,\mathrm{sgn}(\omega),$$

the Hilbert transform is an all-phase shift $\pi/2$ transform. Therefore, the Hilbert transform $x(t)$ becomes

$$F\{H\}(\omega) = F(H[x(t)]) = \begin{cases} -j\,\mathrm{Re}(\omega) - \mathrm{Im}(\omega) & \omega > 0 \\ j\,\mathrm{Re}(\omega) + \mathrm{Im}(\omega) & \omega < 0 \end{cases}.$$

Therefore,

$$F(\omega) + jF\{H\}(\omega) = F\{1 + jH\}(\omega) = \begin{cases} 2[\mathrm{Re}(\omega) + j\mathrm{Im}(\omega)] & \omega > 0 \\ 0 & \omega < 0 \end{cases}.$$

If $y(t) = H(x(t))$, the frequency spectrum of $x(t) + jy(t)$ is positive. Thus, based on the Hilbert transform, let $\varphi(t) = \arg\tan(y(t)/x(t))$; then, the circular frequency $\omega(t) = \varphi'(t)$. Thus, we can calculate the instantaneous frequency of vibration

$$f(t) = \frac{\omega(t)}{2\pi}.$$

while the amplitude is

$$a(t) = \sqrt{x^2(t) + y^2(t)},$$

the original signal is

$$x(t) = \mathrm{Re}\left[a(t)e^{j\int 2\pi f(t)dt}\right],$$

where e is the base of the natural logarithm.

Assume that an original signal $X(t)$ that is decomposed by EMD splits into n IMF functions $x_i(t)$ and into the residual function $res(t)$:

$$X(t) = res(t) + \sum_i x_i(t).$$

Then, the Hilbert transform is performed on each IMF function to obtain the final synthetic signal expression as follows:

$$X(t) = \mathrm{Re}\sum_i \left[a_i(t)e^{j\int 2\pi f_i(t)dt}\right] + res(t).$$

Although the Hilbert transform can analyse the instantaneous frequency, its applicability is restricted by the form of the original signal (Datig and Schlurmann 2004; Huang et al., 2003; Nuttall, 1966; Sharpley and Vatchev, 2006). For poorly formed signals, such as very irregular signals, use of the Hilbert transform produces results that significantly differ from the actual physical status. After EMD decomposition, the IMF functions have superior characteristics with a small bandwidth, reasonable symmetry and smooth envelope changes, which are suitable for the Hilbert transform. Thus, the two parts are highly connected.

3. Improved HHT analysis: optimisation of the stop condition

In the EMD sifting process, a very important control parameter is the threshold of the stop condition (Damerval et al., 2005). Assume that it currently has a split m-1 intrinsic mode function, in which imf_1, imf_2,..., imf_{m-1}, imf_m are calculated. The current processing step is $f_{i-1}(x, y)$. The maximum of $f_{i-1}(x, y)$ will fit an up envelope, whereas the minimum of $f_{i-1}(x, y)$ will fit a down envelope. The average of the two envelopes are the mean function. If the mean is removed, then the signal $f_i(x, y)$ is generated. To clearly mark the level of decomposition for each operation signal, we respectively denote the signals as $f_{m,i-1}(x, y)$ and $f_{m,i}(x, y)$. Then, $f_{m,i}(x, y)$ can be employed as a stop condition for imf,

$$\frac{\iint_{x,y}\left|f_{m,i}(x, y) - f_{m,i-1}(x, y)\right|^2 dx\, dy}{\iint_{x,y}\left|f_{m,i-1}(x, y)\right|^2 dx\, dy} < \omega_0.$$

Consequently, the relative change rate ω is defined as the symmetry factor; its physical meaning is a measure of the approximate symmetry degree of the current function, which varies from the symmetry axis zero. ω_0 is referred to as the permission symmetric threshold and indicates the acceptable degree of symmetry of each IMF in the EMD process. In a practical computation, the value of ω_0 falls in the interval [0.01, 0.3]. The smaller the ω_0 value, the greater the number of IMFs that will become decomposed.

We consider the one-dimensional IMF calculation process as an example to estimate the value range of the symmetry factor. The previous explanation of the HHT is derived on the time domain for simplicity. However, the space domain variable s is adopted in the formulas instead of t in the following sections due to the fact that magnetic data and synthesised functions are based on the space domain. Assume that the current signal $x(s)$ is close to an IMF function, which may be equivalent to an approximate irregular circular motion. The slowly varying radius is $a(s)$, and the slowly varying phase is $\varphi(s)$. Thus, the signal is $a(s)\cos(\varphi(s))$. The accurate envelopes of the maxima and minima of this signal should be $a(s)$ and $-a(s)$. Due to the limitations in practical calculations, the obtained envelopes have errors. For example, using the cubic spline interpolation to approximate the envelope, the true envelope cannot be approximated with orders greater than two.

Assume that in a practical calculation, the upper envelope is $a(s) + e_1(s)$ and the lower envelope is $-a(s) + e_2(s)$, where $e_1(s)$and $e_2(s)$ are the error functions. Then, the mean value of the envelope is as follows:

$$e(s) = \frac{\left[e_1(s) + e_2(s)\right]}{2}.$$

As

$$f_{m,i} = f_{m,i-1} - c_{m,i-1} = f_{m,i-1} - e,$$

where $c_{m,i-1}$ are the $(i-1)$-th central function for the m-th pass. By substituting it into the stop condition, we obtain

$$\frac{\int_D\left|f_{m,i}(s) - f_{m,i-1}(s)\right|^2 ds}{\int_D\left|f_{m,i-1}(s)\right|^2 ds} < \omega_0$$

$$\Leftrightarrow \frac{\int_D\left|f_{m,i-1}(s) - e - f_{m,i-1}(s)\right|^2 ds}{\int_D\left|f_{m,i-1}(s)\right|^2 ds} < \omega_0$$

$$\Leftrightarrow \frac{\int_D e^2(s)ds}{\int_D a^2(s)\cos^2(\varphi(s))ds} < \omega_0$$

$$\Rightarrow \frac{\int_D e^2(s)ds}{\int_D a^2(s)ds} < \omega_0.$$

as $e_1(s)$ and $e_2(s)$ are the results from the cubic spline interpolation, $e_1(s) = O(\Delta s)$ and $e_2(s) = O(\Delta s)$. Then,

$$M_s = M_{cs}e^{-n_b\Psi} \quad \text{and} \quad M_{pt} = M_{cs}e^{n_d\Psi}$$

is the transverse spacing between the maxima or minima. In the interval of Δs, we obtain

$$\int_D e^2(s)ds = \int_D [O(\Delta s)]^2 ds = O\left[(\Delta s)^3\right].$$

when $a(s)$ slightly varies in the interval, $a(s) \approx a$ can be considered; therefore, in the interval, we obtain

$$\int_D a^2(s)ds \approx \int_D a^2\, ds = a^2 O(\Delta s).$$

Therefore,

$$\frac{\int_D e^2(s)ds}{\int_D a^2(s)ds} = \frac{O\left[(\Delta s)^3\right]}{a^2 O(\Delta s)} = a^{-2}O\left[(\Delta s)^2\right].$$

Let

$$\omega = \frac{\int_D e^2(s)ds}{\int_D a^2(s)ds} < \omega_0,$$

We define ω as the symmetry factor, where the agreed threshold is denoted as ω_0.

As $\omega = O(\Delta s/a)^2$ in our experiment, $K = \lambda\Delta s/a$ is applied as a dynamic adjustment factor to adjust the stop threshold ω_0, which ensures that the stop threshold is sufficiently small when the wavenumber is high and the amplitude is large; when the wavenumber is low and the amplitude is small, the stop threshold is sufficiently large. In the two-dimensional case, it only needs to replace Δs with the characteristic scale Δx, Δy. Therefore, the dynamic adjustment factor is $K = \lambda\ \min(\Delta x, \Delta y)/a$. The improved stop condition becomes

$$\frac{\iint_{x,y}\left|f_{m,i}(x, y) - f_{m,i-1}(x, y)\right|^2 dx\, dy}{\iint_{x,y}\left|f_{m,i-1}(x, y)\right|^2 dx\, dy} < K\omega_0.$$

In the experiment, the general value of ω_0 is [0.1, 0.3], and the value of λ is [0.5, 2].

4. Characteristic scale and local HHT

We define $(\Delta L(x, y), a(x, y))$ of an IMF that is decomposed from a signal using the HHT as the characteristic scale at the point $P(x, y)$, where ΔL is the average spacing between extreme points in the vicinity of point P and $a(x, y)$ is the amplitude near point P. For the entire signal or a local interval of the signal, if the variation in ΔL and a is relatively small, we can approximately define the mean value of ΔL and a in this range as the characteristic scale of the IMF.

The characteristic scale can serve as a reasonable expression of the scale characteristics of a signal for use as an index to design a

targeted optimisation decomposition method and as a valuable research direction. For example, as described in the previous section, dynamically adjusting the symmetry factor by the ratio of two characteristic scales can produce excellent decomposition results. Similarly, the recognition of a complex signal, using a characteristic scale that is based on the initial decomposition to perform local optimisation decomposition in geological studies, may have many applications (Wang et al., 2006).

Assume that the fluctuation of $(\Delta L, a)$ at the local range is relatively small. We define this fluctuation as the characteristic domain. The emergence of a characteristic domain frequently implies the existence of a relatively ideal vibration component. First, we employ the dynamic symmetry factor threshold to preliminarily decompose the global signal and identify the characteristic domain of each decomposition function. For the common characteristic domain of multiple levels, we consider this local part to be a distinct multi-scale signal region. Second, we re-decompose the original signal in the region to obtain better decomposition results. This type of transformation process is defined as the local HHT. We can calculate accurate characteristic scale parameters using the re-decomposed signal.

5. Application of the local HHT in mineral information extraction and prediction of mineral deposits

5.1. Matching parameter calculation

In previous geoscience studies, the HHT is frequently and directly applied as a decomposition tool for analysing some geological datasets and obtaining decomposition images. However, this type of study resembles an interpretation instead of a prediction. A local HHT and a method with a characteristic scale that can establish the relationship between geological mineral deposit models and some measurement signals are proposed in this paper. After the deposit information is located in the measurement signals using strategies such as pattern recognition, metallogenic prediction can be achieved.

The basic approach is divided into several steps:

1) Establish a spatial model using the geographical distribution information of a deposit, which will be treated as a two-dimensional signal. Then, apply the HHT to the signal to achieve a complex multi-scale decomposition signal. Remove the trend that has no major effect on the characteristics, and calculate the characteristic scale of the IMF signal at each level as characteristic variables, which can help to establish the corresponding relationship.
2) Perform BEMD decomposition of some geological measurement signals (such as aeromagnetic data) as the location of the deposit is known; the local HHT in the range of the deposit needs to be applied. Using the characteristic scale of the deposit decomposition signal, which is calculated as a guide, each IMF of the measurement signal and each IMF of the deposit signal are compared to find the closest signals. After all matching is completed, each characteristic scale of the measurement decomposition signal as a parameter of an identifier is recorded.
3) In the future, if we obtain the same measurement signal in another area, we recognise the characteristic region of the signal, apply the local HHT, and compare each IMF parameter that has already been calculated with the IMF parameter that was obtained using the local HHT (which can be calculated by a variety of pattern recognition algorithms) to obtain the mineralisation prediction results.

5.2. Deposit distribution model

Generally, a deposit is a three-dimensional distribution with a fuzzy boundary and an internal distribution that is not always continuous. It is neither realistic nor necessary to establish a space model that is strictly in accordance with the real physical situation. In this paper, we utilise a two-dimensional mixed Gaussian model to express the distribution of ore deposits.

Select $_n$ main deposit locations for the ore field and represent their positions as $\{P_1, P_2, \dots, P_n\}$. Let the mineral reserves of each position be $\{Q_1, Q_2, \dots, Q_n\}$, respectively, and let $\{q_1, q_2, \dots, q_n\}$ be their normalised positions, where

$$q_i = \frac{Q_i}{\max_i(Q_i)}.$$

Let the spatial parameters of each represented ore deposit be

$$\left\{ (L_1, W_1, \theta_1)_1, (L_2, W_2, \theta_2), \dots, (L_n, W_n, \theta_n) \right\},$$

where (L, W, θ) represents the length, the width and the axis angle, respectively, of the Gaussian model of the deposit. Then, the mixed Gaussian functions of the distribution deposits are expressed as

$$\sum_{i=1}^{n} q_i \cdot \mathbf{R}^{\theta_i} \left[G\left(P_i, \left(\frac{L_i}{\sqrt{2}}, \frac{W_i}{\sqrt{2}} \right) \right) \right],$$

in which $\mathbf{R}^{\theta_i}$ is the rotation transformation and $G(P, (\sigma_x, \sigma_y))$ is the Gaussian function. The model can represent the metal distribution of the deposits, which is considered to be a signal, and the relationship between the target measurement signal and the actual distribution of mineral deposits can be established.

The HHT trend is similar to the DC component of the Fourier transform: it only reflects the total trend in the characteristics and has no effect on the alternating characteristics of the signal. Therefore, we continue the deposit distribution decomposition function and only analyse the individual IMF components.

6. Case study

We consider the aeromagnetic data from the Gejiu area as the analysis object, apply this method, and discuss and analyse the dynamic control parameters and characteristic scale of the deposit prediction model, based on an HHT.

6.1. Background introduction

Gejiu is the most important tin–copper polymetallic deposit and mining area in China. Its geologic structure is typical and distinct, with very distinct characteristics of controlled metallogenic tectonics. It not only has a long mining history but also has produced very detailed studies with abundant data. Thus, Gejiu is a suitable location for a case study of the development of digital and geological metallogenic prediction models.

The geotectonics of Gejiu's tin and copper polymetallic deposits are located at the joint of the Yangtze block, in the Southern China fold system and the Indochina block and on the western margin of the Youjiang geosynclinal fold belt in southern China's geosynclinal fold area.

The main types of ore deposits are divided into an interlayer sulphide tin polymetallic ore deposit, a granite contact zone skarn-type tin polymetallic ore deposit and an ore vein. The main ore sections are as follows: Damoshan, Malage, Songshujiao, Lutangba, Laochang, Shuangzhu, Jinguangpo, and Kafang.

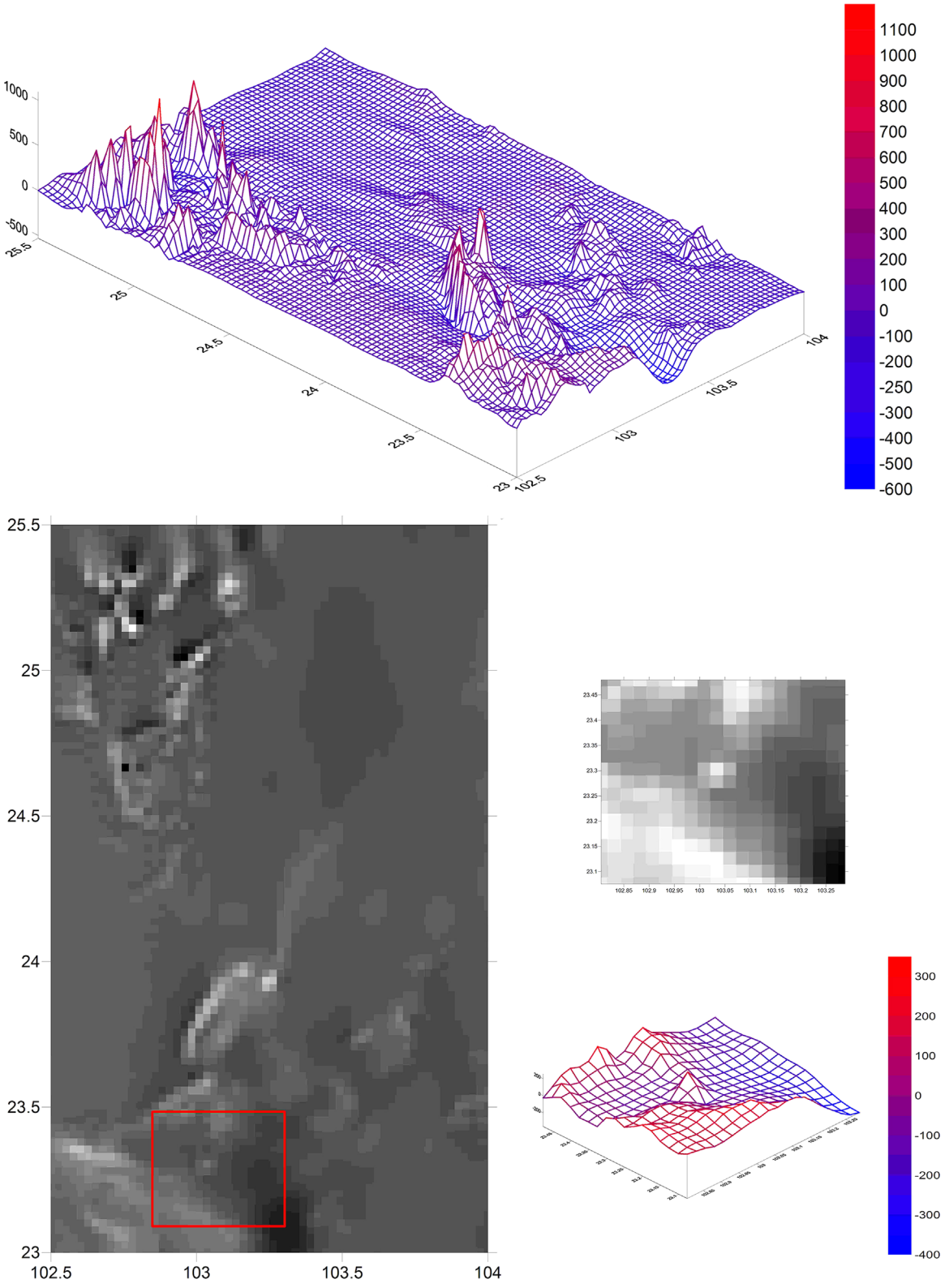

Fig. 2. Aeromagnetic data in the southeast Yunnan region.

6.2. *Aeromagnetic data pre-processing and feature domain determination*

In this paper, we employ aeromagnetic data that were collected from the southeast Yunnan region at longitude [102.5°E, 104°E] and latitude [23°N, 25.5°N] as the experimental data source (as shown in Fig. 2). As southeastern Yunnan is located at a low latitude, the aeromagnetic data need to be pole pre-treated. We employed a standardised polar transform algorithm to calculate the data by a declination value of 22.5° and an inclination value of 49°.

The larger images on the left depict the entire aeromagnetic sampling data in mesh form and image form. The two smaller images on the right show the study area by mesh and image.

Fig. 2 shows that the magnetic anomalies in the lower left

corner of the region are caused by metal deposits in Gejiu, which is the focus of this paper. Therefore, a characteristic region near Gejiu is selected as the research domain. According to the geographical coordinates of Gejiu and a radius, we constrained the general reference region. Then, we decomposed the aeromagnetic data for the entire southeast Yunnan region using HHT decomposition. We discovered a square characteristic area around the reference area with a relatively stable characteristic scale $(\Delta L, a)$ using a sampling method that is based on the previously mentioned principles. The final calculation generates a geographical spatial window [E102.81667, E103.31667] [N23.0834, N23.51667] as the research domain. The calculation processes five levels of HHT decomposition on the magnetic field data; consequently, the variation range of the characteristic scale $(\Delta L, a)$ of low wavenumber IMF1–IMF3 in this region is within 15%.

6.3. Spatial deposit distribution model

As previously mentioned, we utilise a Gaussian mixture model (GMM) to construct the spatial distribution of a magnetic anomaly that is caused by deposits or tin-containing rock point factors. For the convenience of the calculations in the test algorithm, the anisotropic Gaussian function is not applied with the same direction as the basis function; instead, the standard Gaussian function is employed to construct the basic unit of the GMM. The parameters of the Gaussian base are completely consistent; the shapes of different Gaussian functions are not adjusted by controlling the variance σ and the vertical scaling factor k. As long as an ore occurrence position is based on mineral reserves and the deposit spatial shape proportionately interpolates additional ore occurrences, it can also predict the directional control and control of the height and breadth of the distribution function; this approach is simpler. The only requirement is the need for more accurate data. The Gaussian model that is employed in this paper is expressed as

$$G(x, y) = \frac{k}{\sqrt{2\pi}\sigma} \exp\left[\frac{(x-\bar{x})^2 + (y-\bar{y})^2}{2\sigma^2}\right],$$

where $\sigma = 0.01$ and $k = -1.6$.

According to previous exploration data from Gejiu, deposit extraction distribution information from Gejiu, the ore occurrences and ore deposit reserve and grade data, a discrete approximation sample of the distribution of tin–copper polymetallic ore deposits in Gejiu can be generally established.

Specific ore occurrence information is shown in Table 1 and Fig. 3A.

Assuming that the magnetic anomaly amplitude caused by tin deposits is proportionate to the ore reserves and using the data in Table 1 and the GMM, we can approximately simulate the spatial distribution of the magnetic signals of tin deposits in Gejiu, as shown in Figs. 3B and 4.

Fig. 3A shows the distribution of tin ore occurrences in Gejiu, and Fig. 3B shows the distribution of the magnetic anomaly generated by the GMM, which is caused by tin deposit factors.

Fig. 4A shows the generated negative anomaly signals, Fig. 4B shows the grid form of the GMM, and Fig. 4C shows the contour distribution of the GMM.

Fig. 5 shows the contrast between the actual aeromagnetic sampling data and the algorithm-simulated magnetic data for the tin deposits in Gejiu. It depicts a plot of the relative error, which belongs to the [0,1] interval, in our sample window. By removing the trend, simulated magnetic data can satisfy the actual sampling aeromagnetic data. In the southwest corner of this area, the magnetic field shows a positive anomaly, which is primarily attributed to metal mineral distribution chaos or other conditions that affect the magnetic field intensity in west Gejiu. However, the eastern region is relatively flat, and dense mineral deposits form a significant negative abnormal distribution of a geomagnetic field. Its distribution of fluctuation characteristics basically fits with our model, and the only large deviation is observed in the southeastern region, which was probably caused by incomplete data.

Table 1
Presumed positions of tin deposits in the Gejiu area.

Deposit	Latitude	Longitude	Deposit	Latitude	Longitude
1	23.339138	103.144698	23	23.293472	103.209088
2	23.345509	103.158496	24	23.288161	103.203339
3	23.332767	103.161945	25	23.277539	103.215987
4	23.325334	103.149297	26	23.279663	103.230935
5	23.326396	103.164245	27	23.350818	103.218287
6	23.326396	103.118252	28	23.344978	103.213113
7	23.310466	103.102154	29	23.339138	103.211963
8	23.313652	103.167694	30	23.335953	103.218287
9	23.301969	103.118252	31	23.329581	103.213688
10	23.312590	103.184942	32	23.325865	103.222886
11	23.310466	103.160795	33	23.334360	103.221162
12	23.291348	103.142398	34	23.358249	103.225761
13	23.292410	103.156196	35	23.363027	103.244158
14	23.279663	103.171144	36	23.384788	103.225186
15	23.271165	103.195290	37	23.384788	103.233235
16	23.254166	103.175743	38	23.327989	103.061910
17	23.245665	103.171144	39	23.334360	103.032589
18	23.240352	103.164245	40	23.340731	103.041788
19	23.248853	103.150447	41	23.347632	103.051561
20	23.239290	103.150447	42	23.358249	103.010742
21	23.230789	103.157346	43	23.348694	102.986021
22	23.297721	103.219437			

6.4. HHT dynamic decomposition and the selection of a characteristic scale

The GMM approximately represents the magnetic signals caused by tin deposits, whereas aeromagnetic acquisition comprises the magnetic field signals that are actually formed by the comprehensive geology. The main purpose of this paper was to search for the pattern that was contained in the magnetic signals caused by the tin deposits from the actual signals. First, we decomposed the GMM. After the removal of the trend, we determined the scale at which the pattern of the magnetic field of tin deposit exists, typically in the decomposed IMFs. Then, we decomposed the original magnetic field signals using HHT at various levels, searching for a characteristic scale that resembled the pattern of the tin deposit signals and finding the most similar scale combination. The characteristic scale range obtained from the combination is considered to be the proposed scale when HHT decomposition is employed to process aeromagnetic signals to search for the existence of deposits.

Assuming that the signal channel obtained from the GMM, which is applied by the HHT, is G_i, all IMF sets are denoted as S_G. Let M_i denote the signal component of the aeromagnetic sampling data decomposed by the HHT and the denote the entire set as S_H, where s_g, s_h are the subset of S_G and the subset of S_H, respectively. Thus, the search characteristic scale problem can be summarised as an optimisation problem as follows:

$$\arg\min_{s_g \subset S_G, s_h \subset S_H} \left(\frac{\| \sum_{G_i \in s_g} G_i - \sum_{M_j \in s_h} M_j \|^2}{\| \sum_{G_i \in s_g} G_i \| \| \sum_{M_j \in s_h} M_j \|} \right),$$

i.e., searching for the IMF subsets decomposed by the GMM and IMFs, which are decomposed by aeromagnetic data with a minimum relative error. To reduce the large number of combinations, we only computed the subsets that are composed of continuous IMF layers.

The IMF layer that is decomposed by HHT has many possibilities, and different stop factors may result in a different number of

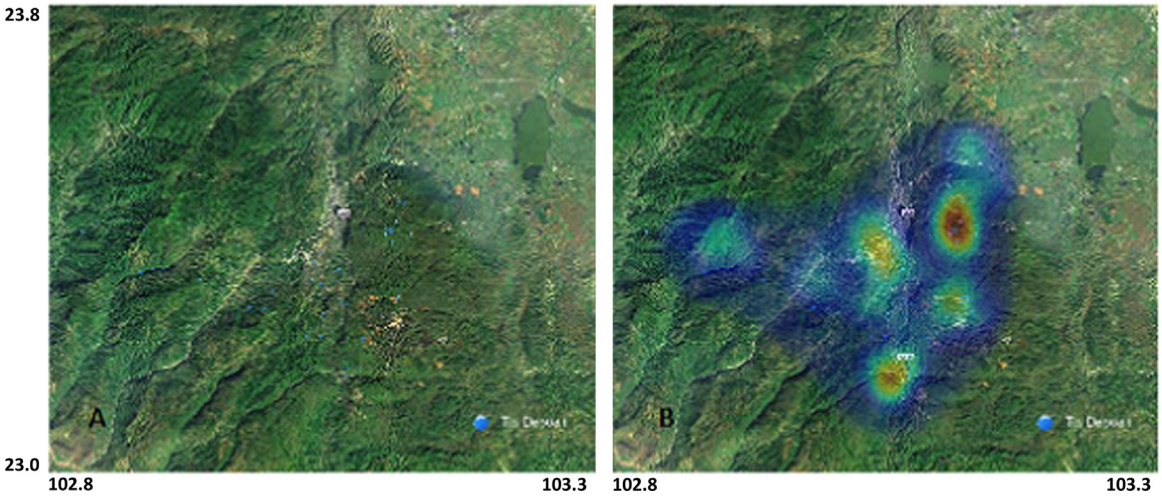

Fig. 3. Tin ore occurrence distribution and simulating magnetic field distribution in Gejiu.

layers and a specific IMF at each layer. Simultaneously, the subset of s_g, s_h with minimum relative error must be determined. Therefore, the computational complexity of the optimisation problem for two variables is difficult to imagine when searching the entire computational domain; the calculation cannot be completed. Therefore, we employed a single variable optimisation iterative algorithm with two circular levels and an approximated reduced search domain to obtain a better solution. For the convenience of description, we agreed on a form DATA [(a, b), C] to represent an s_g or an s_h. When DATA is GMM, the GMM model is decomposed, and when DATA is MEG, the aeromagnetic sampling data are decomposed. *C* denotes the number of layers of the decomposed HHT, namely, the C IMF functions and a residual function (a, b), which indicates a subset formed by the a–b layers.

We processed a 2–5 layer HHT decomposition with a dynamic stop condition on the GMM and a 3–8 layer HHT decomposition on the aeromagnetic sampling function. Then, we calculated the distance between a randomly selected s_g, s_h subset with a total of 100 tests. The pre-test calculation was employed to determine the

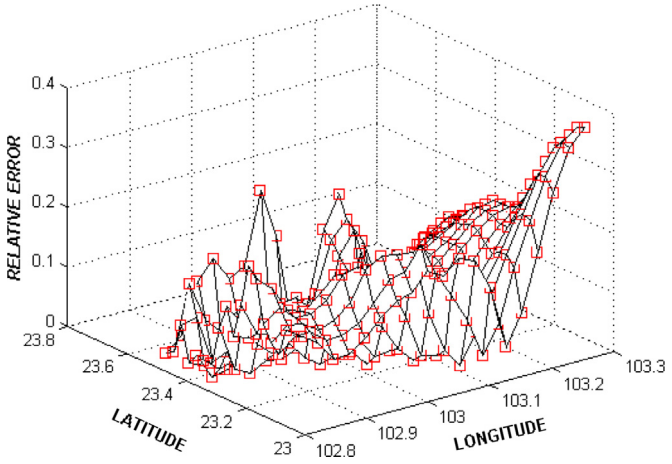

Fig. 5. Relative error of aeromagnetic sampling data and GMM-simulated data of tin deposits in GEJIU.

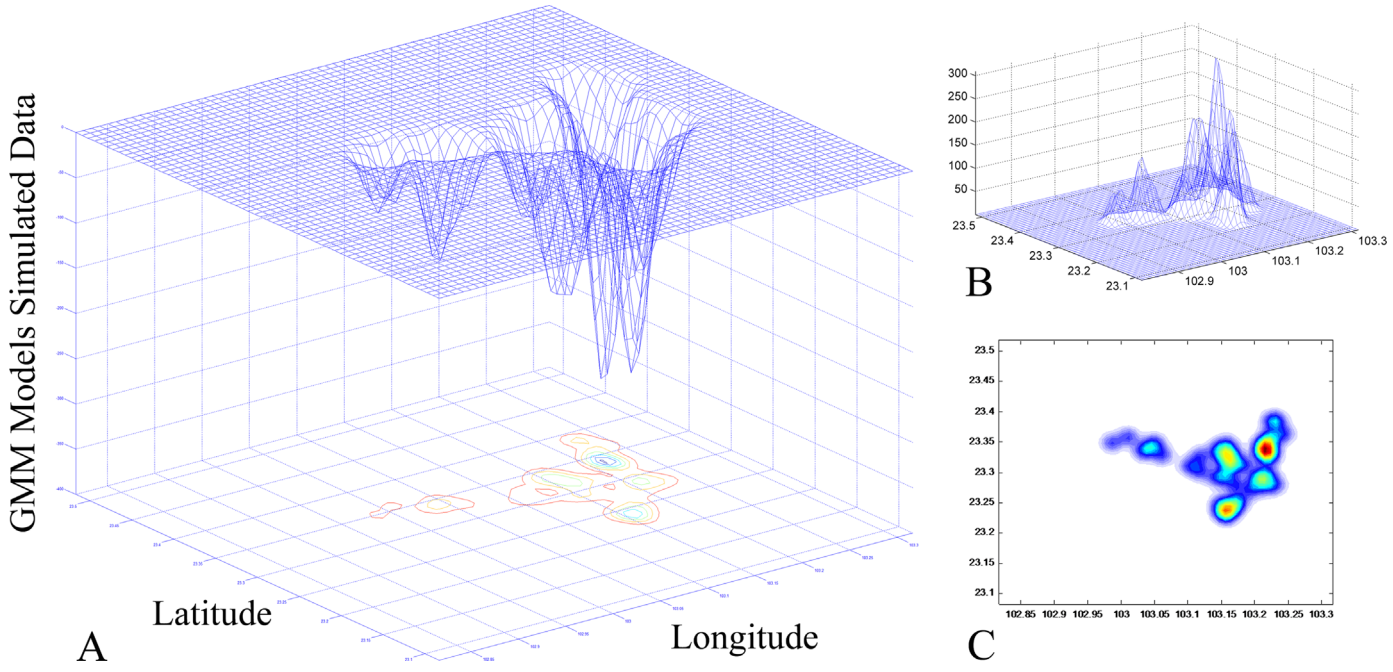

Fig. 4. Magnetic field distribution established by the GMM.

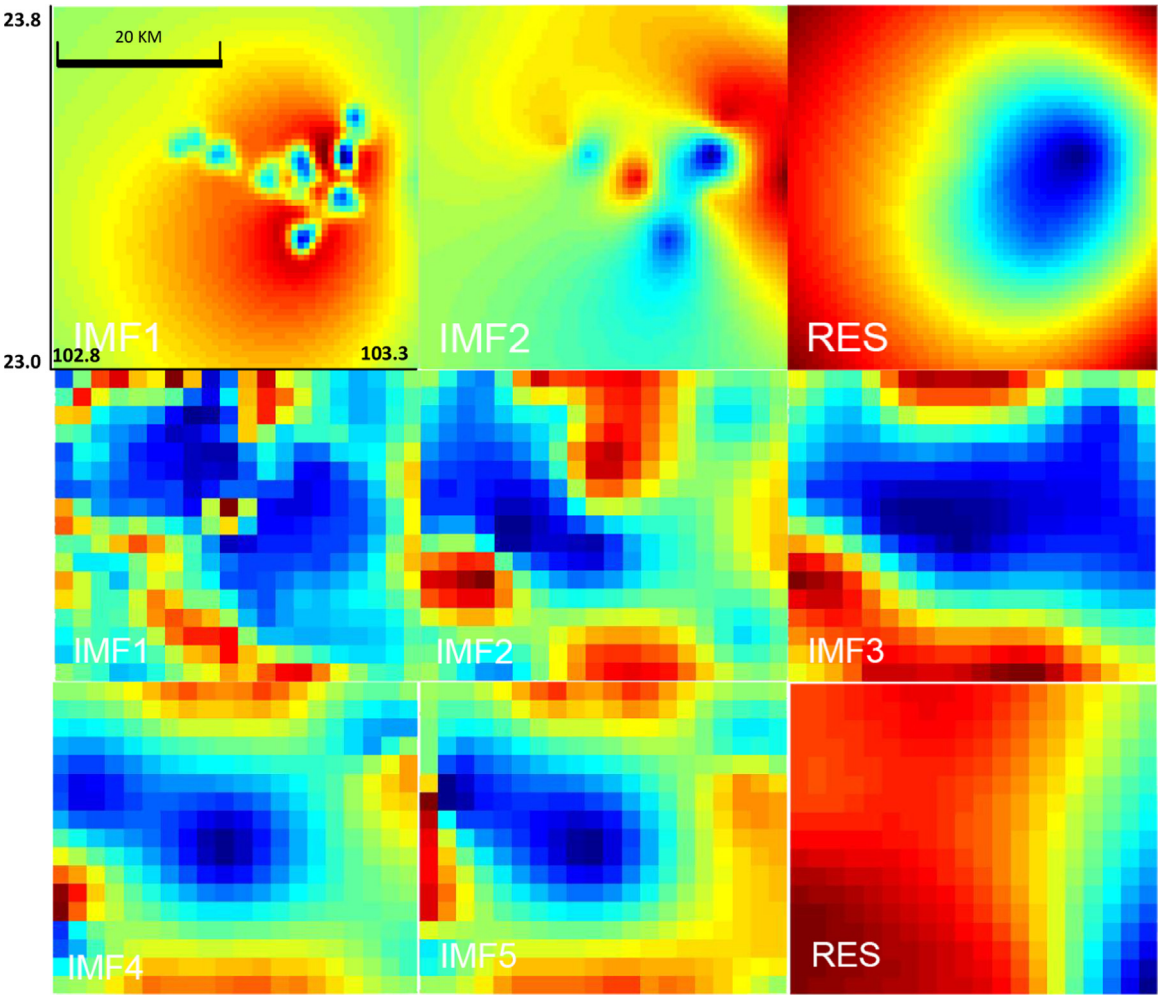

Fig. 6. Matching results of HHT decomposition of the GMM and aeromagnetic data.

initial decomposition layer number of the GMM function, the HHT decomposition and the aeromagnetic sampling HHT decomposition function. The test results are GMM [(2,3),3] and MEG [(3),4].

Then, a two-level iterative optimisation is employed to solve the problem. First, fix the selected s_g; then, search the optimal s_g within the majority of the eight layers of aeromagnetic data decomposition. Second, fix s_h and search for the optimal s_h within the majority of the six layers of the GMM decomposition function. Iterate until the optimum s_g, s_h pair is obtained.

Lastly, the results for the best matching signals are GMM [(2), 2] and MEG [(3,4), 5]; these results offer a reasonable decomposition scale parameter for predicting the magnetic field pattern of the tin deposit model of aeromagnetic data. This parameter can be used to predict other problems and to improve the prediction accuracy and efficiency. The results are shown in Fig. 6.

The first row shows the HHT decomposition results of the GMM, and the second and third rows show the HHT decomposition results of the aeromagnetic data, in which GMM [(2),2] and MEG [(3,4),5] yield the best matches. GMM [(2), 2] indicates that the simulation data of the GMM model is decomposed into two layers of IMFs and that the subset is formed by the second layer; MEG [(3,4) 5] indicates that the aeromagnetic sampling data are decomposed into five layers of IMFs and that the subset is formed by the third and fourth layers.

7. Conclusions and discussion

This paper discusses the mechanism of the HHT signal decomposition tool and proposes that the characteristic scale of signal decomposition is an important feature of the signal components. A self-adaptive HHT decomposition method and a dynamic adjustment stop condition are proposed.

As actual geological sampling data are usually representative of complicated geological factors, a particular factor, such as a signal that is induced from the signal source, is usually only distributed at a characteristic scale interval in the original signal. From the point of view of the HHT, it falls in the signal pathway that is constructed by some of the IMF functions. Therefore, the parameters of the signal pathway can be obtained with analysis and calculations, which help to improve the accuracy and efficiency of a metallogenic prediction. We used the aeromagnetic sampling data from Gejiu in a numerical test to validate the method. The tin deposit magnetic field distribution is approximately expressed by a GMM model and addresses the characteristic search problem, in which the target signal is a two-parameter optimisation problem; the calculation is performed using a heuristic iterative algorithm to obtain the best matching signal channel.

The matching results provides not only effective technical parameters for searching magnetic field data for other tin deposits but also the framework of the method, which can also be employed for other metallogenic prediction problems.

However, the algorithm has many problems. First, the training data for the channel parameter calculations are very important; gaps in the deposit data may cause a substantial difference between the distribution of the GMM and the actual situation, which may affect the validity of the algorithms. Second, the dynamic HHT decomposition that is proposed in this paper substantially relies on the calculation and estimation of the characteristic parameters (ΔL, a). However, a simple estimation sampling method is performed in the paper; a more accurate estimation method requires further study. The calculations of the characteristic signal channel parameters in this paper are only based on the aeromagnetic data from Gejiu. This approach is equivalent to the training process of machine learning technology. When applied to actual metallogenic predictions, we need to test the prediction results using this method for other deposits. Lastly, the method must be practical for real geological prospecting problems. We also need to study the characteristics matching problem in multi-data source conditions. These considerations are important for future research.

Acknowledgements

This research is supported by the project of Yunnan Gejiu large-super large Sn-Cu-polymetallic deposit metallogenic geodynamics background, process and quantitative evaluation, the National Natural Science Foundation of China (41272365, 40972232 and 40772197) and the National High-Tech Research & Development (No: 2006AA06Z113), as well as the China Geological Survey Project (No: 1212011220922). We thank the Yanna Geological Survey for providing the original aeromagnetic data for this study. I am also grateful to my tutor Professor Pengda Zhao, Professor Yongqing Chen and my friend Yuanting Gu.

References

Bansal, A.R., Dimri, V.P., 2005a. Self-affine gravity covariance model for the Bay of Bengal. Geophys. J. Int. 161 (1), 21–30.
Bansal, A.R., Dimri, V.P., 2005b. Depth determination from a non-stationary magnetic profile for scaling geology. Geophys. Prospect. 53 (3), 399–410.
Bansal, A.R., Dimri, V.P., 2014. Modelling of magnetic data for scaling geology. Geophys. Prospect. 62 (2), 385–396.
Bedrosian, E., 1963. On the quadrature approximation to the Hilbert transform of modulated signals. Proc. IEEE 51, 868–869.
Bendat, J.S., 1990. Nonlinear System Analysis and Identification From Random Data. Wiley Interscience, New York, p. 267.
Chen, Q., Huang, N.E., Riemenschneider, S., Xu, Y., 2006. A B-spline approach for empirical mode decomposition. Adv. Comput. Math. 24, 171–195. http://dx.doi.org/10.1007/s10444-004-7614-3.
Chen, Y.Q., Zhao, B.B., 2011, Aplication of BEMD in extraction of regional and local gravity anomalies associated with Au mineralization within Western Shandong Uplift Block, Eastern China. In: Proceedings of the Eighth International Conference on Fuzzy Systerms and Knowledges Discovery (FSKD).
Chen, Y.Q., Zhao, B.B., 2012. Extraction of gravity anomalies associated with gold mineralization: a comparison of singular value decomposition and bi-dimensional empirical mode decomposition. Adv. Mater. Res. 455–456, 1567–1577.
Cheng, Q.M., 2003. Non-linear mineralization model and information processing methods for prediction of unconventional mineral resources. Earth Sci.—J. China Univ. Geosci. 28 (4), 1–10 (in Chinese with English abstract).
Cheng, Q.M., 2004. A new model for quantifying an isotropic scale invariance and for decomposition of mixing patterns. Math. Geol. 36 (3), 345–360.
Coughlin, K., Tung, K.-K., 2004. 11-year solar cycle in the stratosphere extracted by the empirical mode decomposition method. Adv. Space Res. 34, 323–329.
Coughlin, K., Tung, K.-K., 2005. Empirical mode decomposition and climate variability. In: Huang, N.E., Shen, S.S.P. (Eds.), Hilbert-Huang Transform: Introduction and Applications. World Sci, Singapore, pp. 149–165.
Damerval, C., Meignen, S., Perrier, V., 2005. A fast algorithm for bidimensional EMD. IEEE Signal Process. Lett. 12, 701–704. http://dx.doi.org/10.1109/LSP.2005.855548.
Datig, M., Schlurmann, T., 2004. Performance and limitations of the Hilbert-Huang transformation (HHT) with an application to irregular water waves. Ocean. Eng. 31, 1783–1834. http://dx.doi.org/10.1016/j.oceaneng.2004.03.007.
Diks, C., 1999. Nonlinear Time Series Analysis. World Sci, Singapore, p. 209.
Duffy, D.G., 2004. The application of Hilbert–Huang transforms to meteorological datasets. J. Atmos. Ocean. Technol. 21, 599–611. http://dx.doi.org/10.1175/1520-0426(2004)021 < 0599: TAOHTT > 2.0.CO;2.
Flandrin, P., Gonçalves, P., 2004. Empirical mode decompositions as data-driven wavelet- like expansions. Int. J. Wavelets Multiresolut. Inf. Process. 2, 477–496. http://dx.doi.org/10.1142/S0219691304000561.
Flandrin, P., Rilling, G., Gonçalves, P., 2004. Empirical mode decomposition as a filter bank. IEEE Signal Process. Lett. 11, 112–114. http://dx.doi.org/10.1109/LSP.2003.821662.
Flandrin, P., Gonçalves, P., Rilling, G., 2005. EMD equivalent filter bank, from interpretation to applications. In: Huang, N.E., Shen, S. (Eds.), Hilbert-Huang Transform and its Applications. World Sci, Singapore, pp. 57–74.
Han, C.M., Guo, D.D., Wang, C.L., Fan, D., 2002. A novel method to reduce speckle in SAR images. Int. J. Remote. Sens. 23, 5095–5101. http://dx.doi.org/10.1080/01431160210153110.
Hou, W.S., Yang, Z.J., Zhou, Y.Z., Zhang, L.P., Wu, W.L., 2012. Extracting magnetic anomalies based on an improved BEMD method: a case study in the Pangxidong Area, South China. Comput. Geosci. 48, 1–8.
Huang, J.N., Zhao, B.B., Chen, Y.Q., Zhao, P.D., 2010. Bidimensional empirical mode decomposition (BEMD) for extraction of gravity anomalies associated with gold mineralization in the Tongshi gold field, Western Shandong Uplifted Block, Eastern China. Comput. Geosci. 36, 987–995.
Huang, N.E., 2001. Computer Implemented Empirical Mode Decomposition Apparatus, Method and Article of Manufacture for Two-dimensional Signals, Patent 6311130. U.S. Patent and Trademark Off., Washington, D.C.
Huang, N.E., 2005. Computing Instantaneous Frequency by Normalizing Hilbert Transform, Patent 6901353. U.S. Patent and Trademark Off., Washington, D.C.
Huang, N.E., 2006. Computing Frequency by Using Generalized Zero-crossing Applied to Intrinsic Mode Functions, Patent 6990436. U.S. Patent and Trademark Off., Washington, D. C.
Huang, N.E., Attoh-Okine, N.O. (Eds.), 2005. Hilbert–Huang Transforms in Engineering. CRC Press, Boca Raton, Fla, p. 313.
Huang, N.E., Shen, S.S.P. (Eds.), 2005. Hilbert-Huang Transform and its Applications. World Science, Singapore, p. 311.
Huang, N.E., Long, S.R., Shen, Z., 1996. The mechanism for frequency downshift in nonlinear wave evolution. Adv. Appl. Mech. 32, 59–111.
Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.-C., Tung, C. C., Liu, H.H., 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis. Proc. R. Soc. Lond., Ser. A 454, 903–993.
Huang, N.E., Chern, C.C., Huang, K., Salvino, L.W., Long, S.R., Fan, K.L., 2001. A new spectral representation of earthquake data: Hilbert spectral analysis of station TCU129. Chi-Chi.
Huang, N.E., Wu, M.L., Long, S.R., Shen, S.S., Qu, W.D., Gloersen, P., Fan, K.L., 2003. A confidence limit for the position empirical mode decomposition and Hilbert spectral analysis. Proc. R. Soc. Lond., Ser. A 459, 2317–2345.
Huang, N.E., Wu, Z., Long, S.R., Arnold, K.C., Blank, K., Liu, T.W., 2008. On instantaneous frequency. Adv. Adapt. Data Anal.
Jian, Z.Z., Zhao, B.B., Chen, Y.Q., 2012. Application of bi-dimensional empirical mode decomposition (BEMD) in extraction of Platinum and Palladium anomalies features. Adv. Adapt. Data Anal. 4 (1 & 2).
Kantz, H., Schreiber, T., 1997. Nonlinear Time Series Analysis. Cambridge University Press, Cambridge, U.K., p. 320.
Komm, R.W., Hill, F., Howe, R., 2001. Empirical mode decomposition and Hilbert analysis applied to rotation residuals of the solar convection zone. Astrophys. J. 558, 428–441.
Nuttall, A.H., 1966. On the quadrature approximation to the Hilbert transform of modulated signals. Proc. IEEE 54, 1458–1459.
Schlurmann, T., 2002. Spectral analysis of nonlinear water waves based on the Hilbert–Huang transformation. J. Offshore Mech. Arch. Eng. 124, 22–27. http://dx.doi.org/10.1115/1.1423911.
Sharpley, R.C., Vatchev, V., 2006. Analysis of the intrinsic mode functions. Constr. Approx. 24, 17–47. http://dx.doi.org/10.1007/s00365-005-0603-z.
Tong, H., 1990. Nonlinear Time Series Analysis. Oxford University Press, Oxford, U.K, p. 584.
Turcotte, D.L., 1997. Fractals and Chaos in Geology and Geophysics, Second edition. Cambridge University Press.
Wang, F.T., Chang, S.H., Lee, C.Y., 2006. Signal detection in underwater sound using the empirical mode decomposition. IEICE Trans. Fundam. Electron. Commun. Comput. Sci., E89-A, 2415–2421.
Widrow, B., Stearns, S.D., 1985. Adaptive Signal Processing. Prentice Hall, Upper Saddle River, N.J., p. 474.
Wu, Z., Huang, N.E., 2005. Huang, N.E., Shen, S.S.P. (Eds.), Statistical Significant Test of Intrinsic Mode Functions, in Hilbert–Huang Transform: Introduction and Applications. World Science, Singapore, pp. 125–148.
Wu, Z., Huang, N.E., Long, S.R., Peng, C.-K., 2007. On the trend, detrending, and the variability of nonlinear and nonstationary time series. Proc. Natl. Acad. Sci. U.S. A. 104 (14), 14889–14894.
Xu, Y., Liu, B., Liu, J., Riemenschneider, S., 2006. Two dimensional empirical mode decomposition by finite elements. Proc. R. Soc. Ser. 462, 3081–3096.
Zhang, R.R., 2006. Characterizing and quantifying earthquake induced site nonlinearity. Soil Dyn. Earthq. Eng. 26, 799–812. http://dx.doi.org/10.1016/j.soildyn.2005.03.004.
Zhao, P.D., 1982w. On the mathematical characteristics of geological bodies. Earth Sci. – J. China Univ. Geosci. 1, 145–155 (in Chinese with English abstract).
Zhao, P.D., 2002. Three-Component quantitative resource prediction and assessments: theory and practice of digital mineral prospecting. Earth Sci. – J. China Univ. Geosci. 27 (5), 482–489 (in Chinese with English abstract).
Zhao, P.D., Hu, W.L., Li, Z.J., 1994. Statistical Prediction of Mineral Deposit, Beijng. Geological Publishing House.