Fusion of VNIR and Thermal Infrared Remote Sensing Data Based on GA-SOFM Neural Network

WANG Chongchang  YANG Guijun  MA Zhen li  XING Zhurong

Abstract  The multi-source data fusion methods are rarely involved in VNIR and thermal infrared remote sensing at present. Therefore, the potential advantages of the two kinds of data have not yet been adequately tapped, which results in low calculation precision of parameters related with land surface temperature. A new fusion method is put forward where the characteristics of the high spatial resolution of VNIR (visible and near infrared) data and the high temporal resolution of thermal infrared data are fully explored in this paper. Non-linear fusion is implemented to obtain the land surface temperature in high spatial resolution and the high temporal resolution between the land surface parameters estimated from VNIR data and the thermal infrared data by means of GA-SOFM (genetic algorithms & self-organizing feature maps)-ANN (artificial neural network). Finally, the method is verified by ASTER satellite data. The result shows that the method is simple and convenient and can rapidly capture land surface temperature distribution of higher resolution with high precision.

Keywords  fusion; VNIR data; thermal infrared; land surface parameter; GA-SOFM mapping; ANN

CLC number  P237

Introduction

For the calculation of the land surface energy balance in local areas, the thermal infrared data of high spatial resolution feature is usually needed. However, the spatial resolution of satellite-borne thermal infrared data is relatively low in the general case that it is an option to fuse the visible data of high spatial resolution and the thermal infrared data of high temporal resolution in the closed time phases. At present, a few methods for the fusion of visible and thermal infrared data emphasize mostly the image fusing visual effect but always neglect the physical significance of fused thermal infrared data resulting in incapable calculation for the remote sensing physical model[1].

It is put forward in this paper that the non-linear neural networks mapping relationship between the parameters and the surface temperature in the same spatial resolution is established by reducing the spatial resolution of visible and near infrared (VNIR) data to inverse the relevant parameters of land surface. The relationship is applied to the visible data and thermal infrared data in original high resolution to obtain the land surface temperature of sub-pixels, and to realize the fusion of the two kinds of data. This

Received on April 26, 2009

Supported by the Key Laboratory of Mapping from Space of State Bureau of Surveying and Mapping(No.200815), the Natural Science Foundation of China (NSFC 40371087, 40701119), the Major State Basic Research Development Program of China (973 Program) (No. 2007CB714401), the National High Technology Research and Development Program of China (863 Program) (No. 2007AA102201 ).

WANG Chongchang, School of Geomatics, Liaoning Technical University, 47 Zhonghua Road, Fuxin 123000, China.
E-mail: wcc_clxwang@yahoo.com.cn
method is not a simple combination of two types of data, but belongs to a semi-experimental model realized by their transitional parameters. Though the estimated temperature of the sub-pixel has a little difference from the actual reversed temperature or measured temperature, the estimated precision is much higher than that of simple vegetation index estimation\(^{[1, 2]}\), so it can satisfy the precision requirements in the computation for local evapotranspiration and surface fluxes.

1 Study site and the data

The testing area is located in Shunyi District of Beijing City (about 22 km×15 km), and the types of surface being covered mainly are vegetation, bare soil, towns (including artificial buildings) and water. Vegetations generally are winter wheat and some low bushes. From April to May 2001, the Institute of Remote Sensing Applications, Chinese Academy of Science and some survey departments carried out a large scale remote sensing ground experiment and satellite synchrony observation, including the conventional observation and continuous sampling in Shunyi meteorological station. The meteorological data includes air pressure, air temperature, relative humidity and general cloud amount of multi-periods, visibility, speed and direction of wind, ground temperature and total sunshine hours and evaporation.

The ASTER 1B data passing through on May 19, 2001 was selected, the wavebands of 1\(^{st}\) (0.52-0.60 µm), 2\(^{nd}\) (0.63-0.69 µm) and 3\(^{rd}\) (0.76-0.86µm) were used, and the spatial resolution ratio is 15 m as the VNIR data. The temperature product (spatial resolution 90 m) is taken as true temperature of the earth’s surface (Ts), and the atmospheric correction to the above data is implemented by means of MODTRAN 4, according to air sounding meteorological data of satellites passing through. Meanwhile, the geometric registration for both is referenced to ASTER VNIR waveband. The supervision typing is achieved with VNIR data to estimate the sub-pixel temperatures under the different coverage. The different input parameters are selected for the different cover types to have neural networks train during the estimation, so that the estimating precision is improved.

In order to verify the applicability of the method, a group of data sets in different spatial resolution was structured based on the existing data. Since only ASTER temperature product data of 90 m resolution can be used to verify, two estimation modes are set-up, which are VNIR(90 m)-Ts(360 m) and VNIR(90 m)-Ts (1 080 m) (defined as Mode 1 and Mode 2). VNIR is the visible & near infrared data, and Ts is the earth’s surface temperature data. Both modes of data were obtained from ASTER 15 VNIR data and 90 m temperature data after aggregate resampling, and the surface temperature of sub-pixel 90 m resolution was achieved by fusing both modes and the estimated results were compared and analyzed.

2 Fusion methods

Taking Mode 1 for example, the surface temperature data of resolution 90 m is obtained by fusing the data of spatial resolution 360 m. As shown in Fig.1, the main steps of fusion are: (1) Using the actual measurement data and existing spectrum to build the training area, the supervised classification being in progress to VNIR 90 m data, make the earth’s surface be allotted by four kinds which are vegetation, bare soil, towns and water; (2) Separating the corresponding reflective data of the vegetation, bare soil, towns and water surface from VNIR 90 m. Theoretically, there is no mix-pixel question in the process of space resolution lessening if classification is very accurate. It is hard to assure 100% accuracy, because noise still exists partly. The spatial resolution of the reflectivity data and the classification result reduces to 360 m in proper order, respectively for VNIR\(_{90}\), CLASS\(_{90}\), VNIR\(_{360}\), CLASS\(_{360}\) and Ts\(_{360}\) (CLASS is the earth’s surface classification chart, the subscript is resolution); (3) According to the different coverage, the main parameters that effect the earth’s surface temperature are ascertained. All above data are used to determine the earth’s surface parameters on two scales; (4) Under the spatial resolution of 360 m, build the mapping relation of the earth’s surface parameters of the different covers and Ts by using the GA-SOFM artificial neural network; (5) Applying the
nonlinearity mapping relation to VNIR90 data with high resolution, and then the pixel earth’s surface temperature of 90m resolution can be acquired. Finally, the fusion is realized with two kinds of data.

The characteristics of the method are: (1) Compared with the statistics in linearity or multinomial relationship between NDVI(normalized difference vegetation index) and Ts, the reliability of nonlinearity trained by ANN is higher than the former, the precision of sub-pixel temperature calculation is also improved; (2) Avoiding the weakness to adopt the remote sensing physical model as radiation transformation which needs to input a large amount of parameters; (3) It can build the knowledge base of ANN in different space-time dimensions, and is beneficial to gain the earth’s surface temperature of the sub-pixel rapidly according to VNIR data.

Fig.1  Flowchart of fusing VNIR and TIR data based on land surface parameters and ANN

2.1 Parameter capture from network input

In order to realize the data fusion in different surface coverage, the key parameters that affect the surface temperature should be chosen as the network inputs to keep network training optimization. The network input parameters are gained by reversing from VNIR, Ts and relevant arithmetic in both modes.

(1) Vegetation area. When using the CUPID model to simulate the canopy temperature of vegetation, the main factors that affect canopy temperature are air temperature, soil surface moisture, LAI (leaf area index), near-infrared reflectivity, etc. When considering only the surface parameter reversed from VNIR, the network input parameters in the covered area are: NDVI, TVDI (temperature-vegetation dryness index), LAI, VWC (vegetation water content), SWC (soil water content), near-infrared reflectivity (RefNIR), red wave band reflectivity (RefR), degree of vegetation coverage (fv) and RVI (ratio vegetation index). The reason why TVDI is chosen as the earth’s surface parameter is that: (1) The surface temperature information of 360m spatial resolution can be introduced into SOFM artificial neural network by using VNIR360 and Ts360 to implicitly implicate the priori knowledge of surface temperature under low resolution in the mapping relationship training, which will play a key role in establishing the nonlinearity mapping relationship; (2) It has been a consensus in the world to adopt TVDI as the key parameter for evaluation of the surface moisture or drought condition, and besides TVDI is also a helpful supplement to apply VWC and SWC parameter.

(2) Bare soil area. The earth’s surface temperature of bare soil area is mainly affected by the air temperature, soil moisture, atmosphere, terrain and the soil physical characteristics (such as soil thermal capacity and thermal conductivity). In addition, the difference in reflectivity between visible light and near-infrared wavebands for bare soil may be diminished greatly. For bare soil or low rates coverage soil, the surface temperature and the earth’s surface moisture content go hand in hand. Choose SWC, SSC (soil sand content), TVDI, RefSNI, RefR and RVI as the parameter of bare soil by referring the parameter of vegetation cover area. Except for SWC and SSC, the other algorithms of the parameter are the same for vegetation. The parametric calculation of SSC is...
based on the soil spectrum of different silt content and moisture content that is gained by laboratory reality measures, which count out the exponent experience relation between silt concentration and moisture content, dry silt reflectivity and dry ground reflectivity, and then make use of the VNIR data to oppose silt concentration and moisture content[9].

(3) Cities and towns area. Estimate the earth’s surface temperature of cities and towns area, consulting the method for estimating city and town surface exchange amounts based on the balancing theory of energy. The main parameters affecting cities and towns energy balance are building geometry structure, super crust emissive and reflectivity (roof, road and wall), atmosphere and solar radiation[10]. Because it is difficult to estimate these parameters in the existing data of VNIR and Ts, we have chosen RVI, near-infrared reflectivity (RefNIR), red wave band reflectivity (RefR), DVI and TVDI as parameters. The main foundation for choosing is the characteristic of man-power material in visible light and the near-infrared wave band reflectivity. Besides, the surface of cities and towns is a comixture of building, vegetation, water and so on, so TVDI may be the key parameter among them.

(4) Water area. In the entire situation testing where water area takes up only about 8%, the water temperature is mainly decided by atmosphere, solar radiation, matter distribution in the water and the physical and chemical state. Considering that it is difficult to invert wave component using the data provided by the paper, RVI, near-infrared reflectivity (RefNIR), red wave band reflectivity (RefR) and DVI are chosen only as water surface parameters.

According to the above-mentioned method, calculate the earth’s surface parameters under two kinds of resolutions (90 m and 360 m) respectively, providing GA-SOFM with neural networks as input parameter after standardization.

2.2 Structure and training of GA-SOFM artificial neural network

Self-Organizing Feature Maps (SOFM) are obtained by exerting a feedback dimension restraining to output different layer artificial neural network for clustering the topological property of multidimensional input data to the weighted artificial neural network. Not only can it update the weighted value of winning artificial neural value in the process of training, and update all neighbor of winning artificial neural value, but also the input data is changed into the non-intersect category by partition. In the process of network training, the genetic algorithm (GA) is introduced simultaneously to handle the network parameter continuously until the training is over. After network training, carry on the learning vector quantization (LVQ), tune the network values carefully, and make the training result ultimately reach higher precision[11]. The network is trained respectively according to the four kinds of earth’s surface covers. Finally, four network training results are obtained.

2.2.1 The network structure of SOFM

The sample of input parameters was selected randomly, the sample choosing accounts for 30% of the experimental area, and the other 20% is detection data. Because entering parameters have a certain correlativity, the value’s range of the input parameters that corresponds to the earth’s surface temperature can be divided into different numerical values, that is, the relationship between earth’s surface parameter and temperature can be classified and the network training accuracy can be improved. SOFM is the best neural network, and can execute the classification according to the organization characteristic. The paper has designed a SOFM network with three implications; the network parameters are shown in Table 1.

Each vector of the network computing unit corresponds to a group of the input parameters (8 vegetation areas, 7 bare soil areas, 5 urban areas, 4 water districts), the number of unit vector represents the number of categories mapping relation. The network training is divided into two steps: First, carry out a self-organizing map based on the input parameters then divide the input parameters into categories. Finally, re-train each category and access a group of training set including different mapping. In the self-organizing map learning process the parameter settings are: start learning rate from 0.1 to 0.001 termi-
nation of learning in the learning process, and when the change in the value of connection contingency is less than 0.0001, terminate the study. The largest number of study times is 20,000. Finally, in the re-training process carry on the parameter configuration: the condition for the termination of RMS error is less than 0.0001, and to a minimum; the largest number of study times is 30,000.

| Table 1 | Initial configuration of SOFM neural network |
|---------|-----------------------------------------------|
| **Competitive layer** | **Connotative layer 1** | **Connotative layer 2** | **Connotative layer 3** | **Deferent layer** |
| Neighborhood dimension | 5×5 | — | — | — | — |
| Neighborhood shape | Kohonens | — | — | — | — |
| Vector operation unit | — | * 31 | * 15 | * 10 | — |
| Activation function | — | Tangential hyperbola | S function | Tangential hyperbola | S function |
| Learning rule | — | Grads and contingency approach | Grads and contingency approach | Grads and contingency approach | Grads and contingency approach |
| Learning rate | — | * 1.00 | * 0.80 | * 0.60 | * 0.001 |
| Factor of momentum | — | * 0.70 | * 0.60 | * 0.40 | * 0.400 |

Note: * GA-based optimizing

2.2.2 SOFM network training

Network training is essentially a network for each pixel to obtain a nonlinear mapping relationship between the surface temperature and surface parameters. In fact, even the same type of surface coverage was affected by factors such as topography, the surrounding environment and cover degree differences. This non-linear mapping relation is not single, but has a number of similar mapping relations.

Where $w_j$ is connectivity weight vector located from the output layer neurons $j$ (surface temperature) to the input layer nodes $i$ (surface parameters), $x = \{x_1, x_2, \cdots, x_n\}$ is surface parameters input vector, and $n$ is surface parameters dimension. The network training is divided into two parts: course to fine-tune reconcile. Rough-tune is self-organization’s competitive learning process for unsupervised learning; the process is as follows:

**Step 1** Weight initialization with the real numbers between 0.0-1.0 to each random neurons;

**Step 2** Surface parameters for each input vector $x$, using Euclidean distance as a similarity measure. The calculation of the output layer neuron activation value (distance) is

$$a_j = (\sum_{i=1}^{n} (x_i - w_{ij})^2)^{1/2} = ||x - w_j||$$ (1)

**Step 3** Find the corresponding input vector $x$, which is the smallest value of the activation of neurons, and then follow formula (6) updating weights:

$$w_j(t+1) = w_j(t) + \alpha(t)[x(t) - w_j(t)], \text{if } j \in N_j(t)$$

$$w_j(t+1) = w_j(t), \text{ if } j \notin N_j(t)$$ (2)

where $N_j(t)$ is the moment of victory modules $t$ side neighborhood, $j$ is the output layer neurons, $\alpha(t)$ is the learning rate, which decreases according to the $t$. Its initial value is set general in the range of 0.0 to 1.0. $c$ is the different regional mapping arising from self-organization competitive learning;

**Step 4** Enter the new vector, repeat step 2 to step 3, reach the reset cycles or conditions of termination;

**Step 5** Recalculate each input vector in accordance with the contingency after self-organization competitive learning, divide the correspondence relationship of the input surface parameters and the output surface temperature into a number of mapping center. This is divided into a number of regional mapping and eventually acquire more than a mapping relationship, and not the same type of surface cover only having one mapping relationship. Only in this way can we meet the actual inversion of the surface parameters under abnormal circumstances of the real existence of noise and surface temperature.

The learning rate $\alpha(t)$ decreases with the passage of time slowly: $\alpha(t) = \alpha(t-1) \cdot \sigma$, the initial values of $\alpha$ is chosen between 0.5 and 0.9.

After the completion of self-organization competitive learning, because of the many-to-many relationship between the surface parameters and the surface temperature, the supervised learning will be on the
back of each mapping of the region once more. That is, apply learning vector quantization (LVQ) algorithm to the network vector which has been rough-tuned, and carry out fine-tuning according to the equation below to gain a number of precise mapping relations:

1) Selecting training input vector \( x \) randomly, and find the value \( c \) which makes \( ||x - w_c|| \) smallest.

2) Using LVQ algorithm, if \( x \) and \( w_c \) lie in the same regional mapping, update \( w_c \) according to the next formula:

\[
    w_c(t + 1) = w_c(t) + \alpha(t)[x(t) - w_c(t)]
\]

Otherwise, the value of \( w_c \) is

\[
    w_c(t + 1) = w_c(t) - \alpha(t)[x(t) - w_c(t)]
\]

When \( i \) does not equal to \( c \), the weight vector: \( w_i(t + 1) = w_i(t) \). Learning rate \( \alpha \) is the small normal number, and decreases by each iteration, until reducing to 0.001.

3) If it reaches the maximum number of iterations or termination conditions, stop and put out the power matrix on behalf of the mapping relationship; otherwise, return to Step 1.

2.2.3 Genetic optimization of network parameters

SOFM

SOFM network uses gradient and contingency, but its biggest flaw is the slow convergence and lacks effective evaluation mechanisms to terminate computing. Also, the larger search space, multi-peak and non-differentiable function easily reach local maximum points, which will affect the forecasting accuracy. Another neural network initial weight and threshold as well as the choice of network structure lack sufficient prior knowledge and certain randomness. Therefore, it is difficult to choose the initial point which has the overall situation, so there is a low probability to obtain the global optimum. GA algorithms have adaptive, global optimization and implicit parallelism-features and strong problem-solving abilities. However, genetic algorithms also have shortcomings, especially in searching a near optimal solution. It is unable to pinpoint the location of the optimal solution, which means the fine-tuning capacity of the local search space is relatively poor.

In order to take the advantages of GA and SOFM, we combined the two algorithms, in the process of rough-tuning with fine-tuning, using GA to optimize the value of the number of vector computing unit. The learning steps and learning rate listed in Table 1 are shown in Fig.2.

The general steps of parameter optimization are as follows:

1) Coding method and the initial formation of groups. As the training data in the pilot have been normalized within the range of \([0.0, 1.0]\), floating-point code can be applied. The initial formation’s production is used to construct a group of initial chromosomes for the parameters waiting for optimization in each level of the SOFM network. As there are three types of SOFM network parameters to be optimized, the chromosome contains only three genes. The chromosome structure can be defined as follows:

Chromosome \( j(j =1,2,\cdots,n) \);

Learning rate: 0.4;

Momentum factor: 0.7;

The number of vector computing unit: 15.

These genes in the chromosome are in the scope given by SOFM network, in accordance with uniform distribution of the random distribution function in generating.

2) Groups selection, cross and mutation genetic manipulation. Choose one operator by using the roulette selection method; cross operator by using floating-point cross-linear; mutant operator by choosing uniform mutation. During the mutation, when a unit vector computing variation is deleted, the corresponding code is zero, and when the variability computing unit is increased, a vector operation randomly generates the initialization code.

3) According to SOFM the neural network function is defined as:

\[
    F(x) = \frac{1}{\sqrt{\frac{1}{m} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}}. \quad \text{Training sample collection: } \phi = \{(x_i, y_i)\}, \quad i = 1, 2, \cdots, m (m < n). \quad \text{Where } m \text{ is the number of training samples; } n \text{ is the total number of samples; } x_i \text{ and } y_i \text{ are the input vector and output of expectations, respectively.}
\]

Suppose the training network is \( \text{net} \), \( \text{net}(x) = \hat{y}_i \) is
the actual output. By calculating adaptation of all individuals of the current groups, first string individual floating-point digital decoders will learn the rate and momentum factors and vector computing unit number of neural networks. Then enter the training sample and calculate each individual’s genetic adaptation degree in accordance with the fitness function.

(4) Sort the individual value according to adaptation value and carry out a SOFM operator to the highest individual after decoding, and acquire a new generation of the population $p(t + 1)$, then increase the number of generations.

(5) Determine that the optimization criterion is in line or not, then return to (2), otherwise end the computation.

The parameters that need to be chosen in genetic algorithms mainly include group size $n$, choose probability $p_s$, cross probability $p_c$ as well as mutation probability $p_m$. In this paper, $n$ is 60; $p_s$ is 0.06; $p_c$ is 0.5; $p_m$ is 0.1. When the pre-set maximum algebra is reached or the average fitness’s difference of the continuous evolution of the three-generation groups is less than 0.01, end up the iteration.

3 Results and discussion

3.1 Analysis on the sensitivity of surface parameters

Choose different parameters for different land cover types to carry out neural network training. In order to effectively evaluate the degree of the contribution of SOFM network parameters, we carry out an analysis of these parameters’ sensitivity, respectively. The sensitivity analysis steps are: (1) First, carry out network training for the various types of coverage to obtain network mapping and surface temperature output under normal circumstances. (2) Re-structure the input parameter samples. Individually select the input parameters to be analyzed according to the parameter sample’s original mean and variance, re-construct arithmetic series of samples, and replace this parameter samples. (3) Completely take the corresponding parameter sample mean regarding other parameter samples of network input parameters. (4) Put the input parameters into the trained network and gain the surface temperature output corresponding with waiting analysis parameters. (5) Compare the output of this surface temperature with the output of the surface temperature under normal circumstances, and analyze the surface temperature’s change resulting from the input parameter’s change. If the surface temperature’s changing scope is large, it shows that the sensitivity of parameters is high, otherwise the sensitivity is low.

The results show that in the vegetation cover, the area of sensitivity of the input parameters is divided into three levels: parameters SWC directly relating to soil moisture have the greatest impact on the estimated results. In addition to the indirect parameter TVDI, red and near-infrared bands reflectivity are indirect parameters about the soil, vegetation, moisture and surface status. Other parameters make little contribution to the network output. In terms of the bare soil area input parameters in the RVI, red and near-infrared bands reflectivity are the main factors, which shows that the difference of spectrum space in the bare soil areas is more important than the inversion soil moisture and silt content. TVDI and NDVI are even more important than SWC and SSC, which also can confirm the importance of the spectral information. Urban areas mainly rely on the red-band and near-infrared bands reflectivity, while other parameters account for only 1/3 of all contributions. The conclusion is consistent with that town is the mixture...
of buildings, bare soil, vegetation and water. It is
difficult for the index parameter to play a role. The
reason why water surface mainly relies on the red
band reflectivity and associated RVI index is that the
distribution of chlorophyll in water has a certain
impact on the water temperature on the surface, and
the near infrared reflectance and related index have
less affect.

3.2 Result of fusion and the precision analysis

Use the trained GA-SOFM network to carry on the
fusion computation under the two models, and com-
pare the 90 m special resolution surface temperature
after fusion with the ASTER temperature product
(90m), and count the variance, the average value and
the correlation coefficient separately and so on. The
result and analysis of fusion are shown in Fig.3 and
Table 2, respectively.

The training results shows that in Model 1, looking
from the root mean square error (MSE) of normali-
Zation, the average absolute error and the correlation
coefficient, the fusion precision in the bare soil area
was highest. This proved that the surface temperature
in the bare soil area was less affected by other factors,
which was associated with the network input pa-
rameters. In the town coverage areas the estimation
accuracy was worst. The root-mean-square error was
up to 1.58 and the correlation coefficient was large;
there was still some noise pixels. Thus, at present no
matter whether the town surface temperature was es-

timated by using the physical model or experiential
model, it would be very difficult to obtain satisfactory
results. Because the type of vegetation cover is simi-
lar in the pilot area, the vegetation cover of the vege-
tation coverage area is relatively high, just below the
bare soil. The estimation precision in the water sur-
face is slightly higher than that in the urban area. The
reason is that optical remote sensing cannot reflect all
of the internal-water to a certain extent. In addition,
although the purity of the pixel in the surface area is
high, the water temperature is susceptible to tem-
perature and the internal heat transfer process.

Table 2  Comparing between fusion of land surface temperature and ASTER products under two models

| Covering type         | Estimation model | RMS error | Normalization RMS error | Average absolute error | Minimum absolute error | Maximum absolute error | Linear correlation coefficient |
|-----------------------|------------------|-----------|-------------------------|------------------------|------------------------|------------------------|------------------------------|
| Vegetation coverage area | Model 1         | 0.000 176 | 0.916 008               | 0.010 300              | 6.29E-06               | 0.063 289               | 0.585 000                    |
| Bare soil covered area | Model 2         | 0.000 169 | 0.913 224               | 0.010 219              | 5.91E-07               | 0.061 898               | 0.560 031                    |
| Model 1               | 0.000 159       | 1.250 386 |                        | 0.009 915              | 2.38E-06               | 0.059 348               | 0.623 734                    |
| Model 2               | 6.41E-05        | 0.504 555 |                        | 0.005 866              | 7.01E-07               | 0.059 836               | 0.716 934                    |
| Towns coverage area   | Model 1         | 0.000 195 | 1.581 621               | 0.011 096              | 1.30E-07               | 0.068 638               | 0.428 686                    |
| Water area coverage   | Model 2         | 0.000 336 | 2.719 702               | 0.014 877              | 2.09E-06               | 0.065 066               | 0.225 407                    |
In Model 2, when the number of iterations of the network reached 10,000 times, the four kinds of cover types converged very quickly. The largest MSE of the town coverage area is 0.000336, the root-mean-square error is up to 2.71, the average absolute error and root mean square error have some reduction and the correlation coefficient increased to 0.72. The main reason was that the lower resolution pixel value is the weighted average of the original pixel value. It can suppress the noise pixel contribution, which weaken the complexity of network mapping. All these are beneficial to the fast convergence of the network and the improvement of the accuracy. The fusion precision of the vegetation coverage area is a little lower, but that of the water surface area is increased slightly.

Compared the result integrated through the GA-SOFM neural network method with the results estimated through the Ref. [1] method, the accuracy of this method is increased by 3-4 times, as shown in Fig.4.

![Fig.4 The comparison of fusing accuracy between GA-SOFM and vegetation index method[1]](image)

Overall, for the two modes, the fusion precision of vegetation and town coverage areas is closer, but the fusion accuracy of the town coverage areas in Model 2 dropped significantly. On the other hand, that of the bare soil surface and water surface coverage area is the same generally. The reason why the accuracy of Model 2 for vegetation coverage area is lower than the accuracy of Model 1 is that multiple coverage exists in the vegetation region. That is, there exist a variety of bare soil, vegetation mixed pixel in different proportion, surface parameters inverse-performed from these pixel correspond to the input parameters of high sensitivity, such as SWC, and so on. When using low-resolution surface temperature as the known condition, the information of the input network is mainly provided by surface parameters inverse-performed from mixed pixel of middle and high coverage. If the information provided by the mixed pixel of low coverage is in the low spatial resolution, it has been substantially curbed or abated, so the mapping relations trained by the network is more suitable for the pixel of high coverage, and it responds less to the low coverage, which causes the estimation accuracy of Model 2 to decrease. Furthermore, in order to test the impact on the estimation caused by different vegetation coverage, separately calculate the sub-pixel surface temperature estimated under the two models when the coverage is more than 0.5 and less than 0.2. The results showed that for high vegetation coverage conditions, the estimation accuracy of the two models is high and vice versa. However, under two different vegetation coverage conditions, the accuracy of Model 1 clearly is the best. The urban areas covered by the same laws also exist, but also because the difference of different buildings “cover” results in a large number of mixed pixels in different proportions of the buildings bare soil, vegetation and water surface. Bare soil and surface coverage area are coincidently the contrary of the first two, mainly due to a little impact of the mixed pixel problem; pixel value is roughly similar. Using Model 1, network access to the importation of the information is contained in the more “abnormal temperature noise” element, which is a less precise estimation network. While using Model 2, the lower-resolution process which is the low-pass filtering of the spatial domain, the network with the input parameters with a certain noise, the estimation accuracy of the bare soil and water surface coverage area have improved, but the rate of increase is limited. Experiments show that the spatial resolution difference of VNIR and Ts data which was used for estimating cannot be too large, and should be limited between 3-5 times.
4 Conclusion

A new method was developed where the surface sub-pixel temperature is obtained by fusing GA-SOFM and neural network, which is studied based on the data of the high spatial resolution visible-near-infrared and the low spatial resolution surface temperature. The experiment and error analyses are carried through ASTER data to prove that the fusing precision is 3 to 4 times higher than that of the conventional method, so that it is able to meet the needs of practical application.

The method does not need magnitude measuring data and is fused by two remote sensing data directly, so its technology has strong feasibility. Meanwhile, it is simple and feasible to obtain the sub-pixel surface temperature by using the nonlinear mapping relationship between the surface parameters from neural network parameters and the surface temperatures of high or low resolutions linking with surface parameters. This makes it ideal in maneuverability and repeatability to fit the estimation of surface temperature of high resolution under the condition of the lack of prior knowledge for the surface.

The method can also be used in the estimation of surface temperature field in continuous time phases considering the correction factors of environment and surface in a short period with the establishing of relevant error correcting models. For instance, the surface temperature of 90m resolution can be fused in the local area in the satellite transiting period to provide a new path for the computation of surface evapotranspiration and energy balance in a minor area by using remote sensing data.

A big influence of sorting error on the fusing precision may exist since the method is based on the different surface types. Therefore, a selection of a more suitable and accurate sorting method is needed to improve the sorting precision and reduce the influence of noisy pixels on the neural network mapping relationship. The future work is to perfect the fusing method by using existing satellite data, to finish a magnitude of verification, and to improve the fusing precision by means of the introduction of error correction models.

References

[1] Ling Han, Wu Hanning (2005) Study of multi-source remote sensing data fusion method in the earth science[J]. Earth Science and Environment Journal, 27(3): 77-80
[2] Kustas William P, Norman John M, Anderson Martha C, et al. (2003) Estimating subpixel surface temperatures and energy fluxes from the vegetation index-radiometric temperature relationship[J]. Remote Sensing of Environment, 85(4): 429-440
[3] Norman J M, Divakarla M G (1995) Algorithms for extracting information from remote thermal-IR observations of the Earth’s surface[J]. Remote Sensing of Environment, 51(1): 157-168
[4] Fisher (1994) A model for the seasonal variations of vegetation indices in coarse resolution data and its inversion to extract crop parameters[J]. Remote sensing and environment, 4: 220-230
[5] Carlson T N, Gillies R R, Perry E M (1994) A method to make use of thermal infrared temperature and NDVI measurements to infer soil water content and fractional vegetation cover[J]. Remote Sensing Reviews, 52: 45-59
[6] Liu Peijun, Zhang Lin (1997) A method of estimating soil moisture with remote sensing satellite[J]. Journal of Remote Sensing, 1(2): 135-138
[7] Thomas J Jackson, Chen Daoyi, Michael Cosh, et al. (2004) Vegetation water content mapping using landsat data derived normalized difference water index for corn and soybeans[J]. Remote Sensing of Environment, 92: 475-482
[8] Wang Changyao, Qi Shuhua, Niu Zheng, et al. (2004) Evaluating soil moisture status in China using the temperature-vegetation dryness index (TVDI)[J]. Can.J.Remote Sensing, 30(5): 671-679
[9] Qing Xiao (2002) The quantitative research of airborne remote sensing data[D]. Beijing: Institute of Remote Sensing Applications, Chinese Academy of Sciences
[10] Noilhan J, Mahfouf J.-F (1996) The ISBA land surface parameterization scheme[J]. Global Planet Change, 13: 145-159
[11] Hasibagan, Ma Jianwen, Li qijing (2003) The self-organizing neural network classification study of ASTER data[J]. Earth Science Progress, 18 (3): 345-350