Remote Sensing of Environment 265 (2021) 112665

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

A method for land surface temperature retrieval based on model-data-knowledge-driven and deep learning

Han Wang [a,1], Kebiao Mao [a,b,1,*], Zijin Yuan [a], Jiancheng Shi [c], Mengmeng Cao [a], Zhihao Qin [a], Sibo Duan [a], Bohui Tang [d]

[a] Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, China
[b] School of Physics and Electronic-Engineering, Ningxia University, Yinchuan 750021, China
[c] National Space Science Center, Chinese Academy of Sciences, Beijing 100190, China
[d] Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

ARTICLE INFO

Editor: Jing M. Chen

Keywords:
Land surface temperature (LST)
Model-data-knowledge-driven
Deep learning
Geophysical logical reasoning
Expert knowledge

ABSTRACT

Most algorithms for land surface temperature (LST) retrieval depend on acquiring prior knowledge. To overcome this drawback, we propose a novel LST retrieval method based on model-data-knowledge-driven and deep learning, called the MDK-DL method. Based on the expert knowledge and radiation transfer model, we deduce LST retrieval mechanism and determine the best combination of the thermal infrared (TIR) bands of the sensor. Then, we use the radiation transfer model simulation and reliable satellite-ground data to establish a training and test database, and finally use the deep learning neural network for optimal computation. Three typical high-, medium- and low-spatial-resolution TIR remote sensing datasets (from Gaofen, the Moderate Resolution Imaging Spectroradiometer (MODIS), and Fengyun) are used for theoretical simulation and application analysis. The simulation shows that the minimum mean absolute error (MAE) is less than 0.1 K (standard deviation: 0.04 K; correlation coefficient: 1.000) at a small viewing direction (<7.5°) and less than 0.8 K at a large viewing direction (<65°). The *in situ* validation shows that the minimum MAE obtained by the optimal band combination is approximately 1 K (root mean square error (RMSE) = 1.12 K; coefficient of determination (R^2) = 0.902). The retrieval accuracy is improved by increasing the number of TIR bands in the atmospheric window, and adding accurate atmospheric water vapor information produces better results. In general, four TIR bands in the atmospheric window bands are sufficient to retrieve the LST with high accuracy. Likewise, three TIR bands plus atmospheric water vapor information are sufficient for the retrieval requirements. All analyses indicate that our method is feasible and reliably accurate and can also be used to help design the instrument band to retrieve the LST with high precision.

1. Introduction

Land surface temperature (LST) is an important parameter in the study of the physical processes of regional and global earth-atmosphere system interactions. LST is used in many research fields, including surface energy balance, global climate change, hydrology, agricultural production, and urban thermal environments (Huang et al., 2018; Justice et al., 1998; Price, 1990; Pu et al., 2006; Schmugge et al., 2002b). Space-based remote sensing technology using measurements in thermal infrared (TIR) spectral regions provides an effective way to obtain LSTs on a regional and global scale. Many studies have been conducted based on the thermal radiative transfer equation (RTE). Different algorithms have been proposed to derive LSTs from TIR data with different band settings. These algorithms can be roughly classified into four types: single-channel, split-window, multichannel, and machine learning methods.

The single-channel method calculates LST based on a single RTE. Specifically, the method mainly needs to obtain the land surface emissivity (LSE), atmospheric transmittance (or atmospheric water vapor content, WVC) and the effective mean atmospheric temperature as prior

* Corresponding author at: Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, China.
E-mail addresses: 82101182280@caas.cn (H. Wang), maokebiao@caas.cn (K. Mao), yuanzijin@caas.cn (Z. Yuan), shijiancheng@nssc.ac.cn (J. Shi), 82101182280@caas.cn (M. Cao), qinzhihao@caas.cn (Z. Qin), duansibo@caas.cn (S. Duan), tangbh@igsnrr.ac.cn (B. Tang).
[1] These authors contributed equally to this work and should be considered co-first authors.

https://doi.org/10.1016/j.rse.2021.112665
Received 12 January 2021; Received in revised form 3 August 2021; Accepted 18 August 2021
Available online 26 August 2021

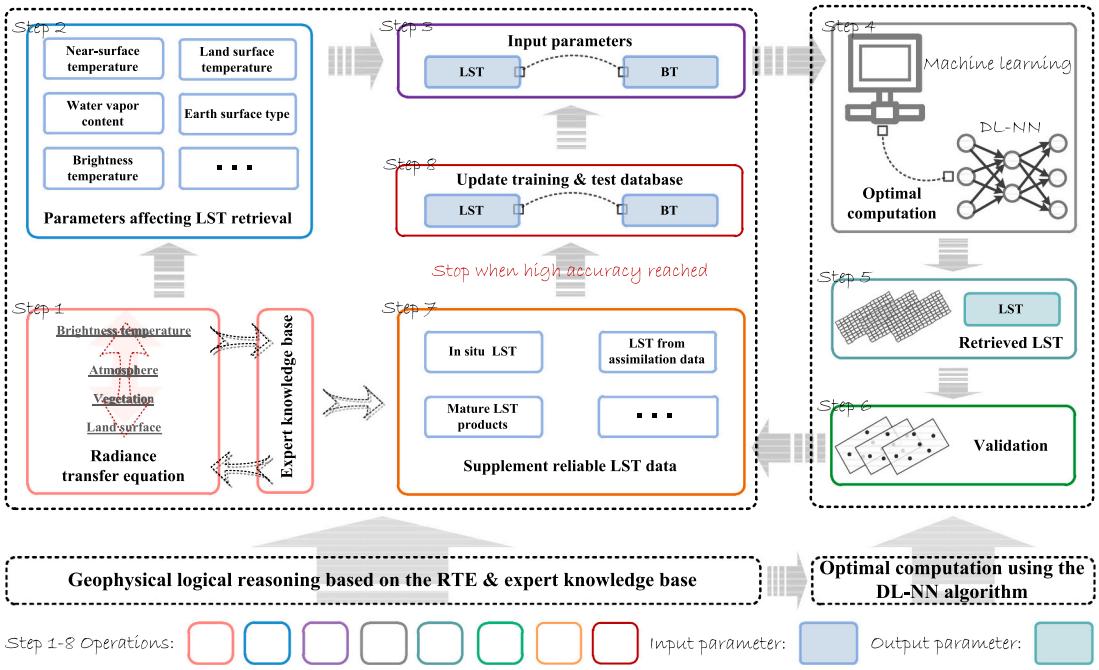

Fig. 1. The MDK-DL method consists of two parts: (a) GLR based on RTE and expert knowledge and (b) optimal computation using the DL-NN algorithm.

knowledge (Cristóbal et al., 2009; Hook et al., 1992; Price, 1983; Sobrino et al., 2004). Qin et al. (2001b) proposed a mono-window algorithm to estimate the LST specifically from Landsat 5 by using only the near-surface air temperature and WVC as empirical knowledge. Jiménez-Muñoz et al. (2009) updated a generalized single-channel algorithm developed by Jiménez-Muñoz and Sobrino (2003), which retrieved the LST from any TIR channel with an FWHM (full width at half maximum) of approximately 1 μm. The algorithm assumed that the LSE and the total WVC were known. Coll et al. (2012) used the single-channel method to retrieve LST based on different sources of atmospheric water vapor and temperature profiles, and applied the method to sensors with a single TIR band. These sensors are the Enhanced Thematic Mapper Plus (ETM+), the Advanced Spaceborne Thermal Emission and Reflection radiometer (ASTER), the Moderate Resolution Imaging Spectroradiometer (MODIS), and the Advanced Along Track Scanning Radiometer (AATSR)). Although the theoretical accuracy of the single-channel method is relatively high, the actual accuracy of this method still needs to be further improved. The main reason is that the accuracy of this method is determined mainly by the accuracy of land cover classification, and the accuracy of transmittance which determined by WVC. However, thermal infrared remote sensing data with high spatial resolution generally lacks water vapor bands, and WVC mainly comes from weather station observation data.

The split-window method establishes two RTEs to calculate LST. The method initially utilized the differential absorption in two TIR bands to correct most atmospheric effects, represents the LST as semiempirical regression equations of the two bands' brightness temperatures (BTs) and usually depended on the LSE. Subsequently, this method further considered treating the information about transmittance, WVC, the view zenith angle (VZA) and other atmospheric parameters to improve the accuracy (Mao et al., 2005; Peres et al., 2010; Price, 1984; Sobrino et al., 1991; Wan and Dozier, 1989). Wan and Dozier (1996) developed a generalized split-window algorithm to retrieve LSTs from the MODIS TIR channels which considered the dependence of the VZA, WVC and atmospheric lower boundary temperature, and the LSEs were estimated from land cover types. The algorithm has achieved very good accuracy and has been continuously improved to help the optimization of MODIS LST products (Wan, 2008; Wan et al., 2002). Qin et al. (2001a) proposed a split-window algorithm to retrieve LST from the Advanced Very High Resolution Radiometer (AVHRR) data, and the algorithm required only two essential parameters (LSE and transmittance). Jiang and Li (2008) constructed a non-linear split-window algorithm for the Spinning Enhanced Visible and Infrared Imager (SEVIRI) on board the first geostationary satellite Meteosat Second Generation (MSG1), which still required two essential parameters (LSE and WVC). LSEs were acquired from global land cover products, and WVC was obtained by the modified split-window covariance and variance ratio (MSWCVR) method from the TIR data. In general, although different researchers have proposed different split-window methods, these methods are basically similar in principle. Most of these methods need to obtain two key parameters, LSE and atmospheric transmittance or WVC. The accuracy validation of the split-window method by using simulation data is approximately 1 K or better than 1 K, but the validation accuracy in actual application still depends on the accuracy of the LSE and transmittance obtained through prior knowledge. The advantage of split-window method over single-channel method is that there is one less unknown than single-channel method, which reduces the difficulty and uncertainty of LST retrieval.

The multichannel method mainly establishes three or more RTEs to calculate LST (Gillespie et al., 1998; Liang, 2001; Mushkin et al., 2005; Schmugge et al., 2002a). Two algorisms are widely used in application. Li and Becker (1993) proposed a method to estimate both LSE and LST using pairs of day/night co-registered AVHRR images. Wan and Li (1997) proposed a day-and-night (D/N) algorithm to retrieve LSEs and LSTs from day/night pairs of MODIS data in seven TIR bands and considered air temperature at the surface level, WVC, and angular form factor. These two algorithms assumed LSEs do not significantly change from day to night. The air temperature at the surface level and WVC were used to modify the initial atmospheric profiles. Finally, 14 nonlinear equations were calculated using the statistical regression method and the least-squares fit method. This algorithm has been tested and obtained great accuracy from simulated MODIS data in wide ranges of atmospheric and surface conditions. Unlike the former two methods, the multichannel method usually takes LSE as an unknown, assuming that LSE does not change significantly during the day and night. However, although the simulation accuracy of the multichannel method is very high (better than 1 K), sometimes the retrieval accuracy of this method cannot be guaranteed when the LSE changes significantly between day and night. In addition, due to the influence of clouds, it is

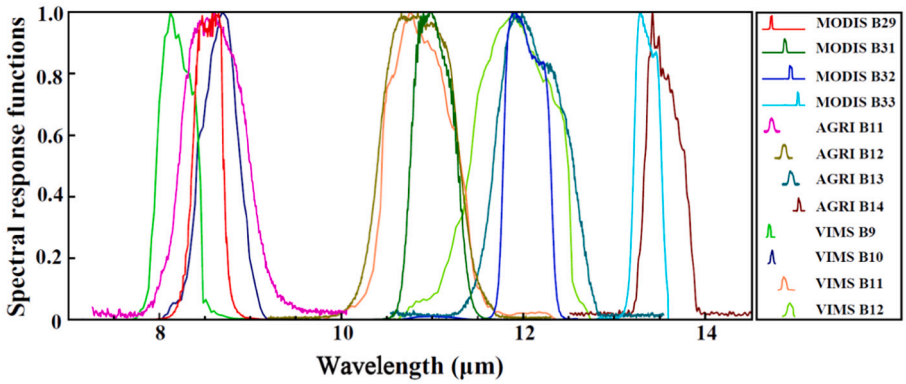

Fig. 2. Spectral response functions of VIMS TIR bands, MODIS TIR bands and AGRI TIR bands, where "B" is an abbreviation of band. For example, "B9" denotes band 9.

difficult to accurately register the data during the day and night, which will also reduce the practicality of the algorithm, such as reducing the resolution.

In the past few decades, machine learning has been applied to the retrieval of geophysical parameters (Frate and Solimini, 2004; Khoob, 2008; Shen et al., 2020; Taylor et al., 2014). Aires et al. (2002) developed a regularized neural network to retrieve atmospheric profiles and LSTs from the high-resolution infrared atmospheric sounding interferometer (IASI) spaceborne instrument. The neural network was trained and tested using real atmospheric situations as measured by radiosondes and taken from the Thermodynamic Initial-Guess Retrieval (TIGR) database. Similarly, Jang et al. (2004), Blackwell (2005), and Mao et al. (2008) also achieved good results in atmospheric profiles/LST/LSE retrieval by using neural network algorithms. Over the past decade, machine learning methods are gradually applied to the retrieval of geophysical parameters, which has considerably outperformed traditional models (Yuan et al., 2020; Zhang et al., 2016; Zhu et al., 2017). Recently, some of the related studies focused on reconstructing the LST of satellite datasets, forecasting daily LST from time series data, and fusing multisource data to estimate subpixel LST (Jia et al., 2021; Wu et al., 2019; Yang et al., 2010; Zhang et al., 2018).

Although machine learning methods have been used in LST retrieval, they have not yet been widely used. There are two main reasons that limit the application of deep learning neural networks (DL-NN). The first reason is that the accuracy of DL-NN depends on the training and test database. The second reason is that most of the physical mechanisms of DL-NN algorithms used in the past are not clear. To make full use of the advantages of traditional physical models and machine learning, we construct a novel LST retrieval method based on model-data-knowledge-driven and deep learning, called the MDK-DL method. First, we perform geophysical logical reasoning (GLR) based on RTE with the help of expert knowledge. Then, we use a machine learning (DL-NN) algorithm to perform optimal computation to approximate the solution (LST) of the established RTEs. Finally, we apply the MDK-DL method to the newly available Gaofen-5 (GF-5)/Visual and Infrared Multispectral Sensor (VIMS) data with high spatial resolution, the reliable Earth Observation System (EOS)/MODIS data with medium spatial resolution, and the newly available Fengyun-4A (FY-4A)/Advanced Geostationary Radiation Imager (AGRI) data with low spatial resolution.

2. Methodology

Some interdisciplinary subjects related to earth science, especially big data and artificial intelligence technology, are on a great upsurge (Zhou, 2019). This situation provides a new perspective to solve geophysical parameter retrieval problems using an interdisciplinary approach, especially by using the advantages of deep learning. Based on these, we propose the MDK-DL method, as shown in Fig. 1. The MDK-DL method described in this section is a general method for LST retrieval. In Section 3, we apply the method to three specific datasets and made detailed analysis.

We use the relevant theories and research results of previous experts for LST retrieval as an expert knowledge base as a knowledge-driven, which condenses human practical experience and wisdom. In Steps 1–3, the mathematical derivation of the retrieval mechanism according to the radiative transfer equation (RTE) is carried out as GLR, which is model-driven and makes full use of expert knowledge. Then, the input parameters of the machine learning process are determined. That is, we study the relationship between the main parameters affecting LST retrieval and demonstrated that the BTs of TIR bands reflect the main information, such as atmospheric WVC and LSE. This process considers data availability, greatly reduces redundant information, and directly uses BTs as the main input nodes for machine learning.

Shown in Fig. 1, based on big-data thinking, we analyze the optimal LST retrieval band combination for specific sensors and use machine learning technology (DL-NN algorithm) as an optimal computation method (data-driven) to obtain the retrieved LST (Step 4 and Step 5). In Step 6, if the validation shows the results to be unsatisfactory, the input is revised by adding more reliable data, where knowledge and data-driven also work (Step 7). These data can be *in situ* ground truth measurements of LST, high-precision LST products (such as MODIS LST products) or other reliable LST data (such as LST from assimilating data). Then, the training and test databases are updated (Step 8). Step 7 and Step 8 are the LST abduction processes. Abduction, learned from artificial intelligence (Zhou, 2019), can also be explained as retro-production, refers to the process of selectively inferring certain facts that explain phenomena and observations based on background knowledge. The MDK-DL method implements knowledge exploiting by means of GLR and evidence/facts exploiting by means of deep learning.

3. Materials and methods

3.1. Data

3.1.1. Remote sensing data

To verify the suitability of the MDK-DL method, we selected representative TIR remote sensing data with high, medium and low spatial resolution for analysis. Specifically, three kinds of remote data were collected: GF-5/VIMS data, EOS/MODIS data, and FY-4A/AGRI data. Fig. 2 shows the spectral response functions of the TIR bands of these three remote sensing datasets.

- As a polar-orbiting satellite launched in May 2018 as part of a series of China High-resolution Earth Observation System (CHEOS) satellites operated by the China National Space Administration (CNSA), GF-5 is a comprehensive hyperspectral satellite that detects the

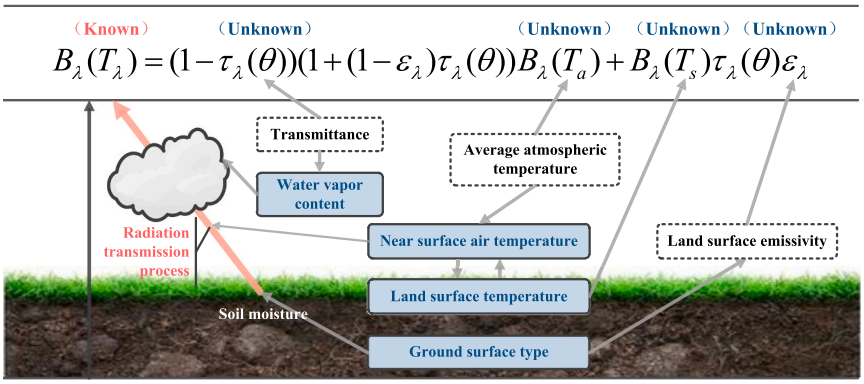

Fig. 3. A simplified diagram of the thermal RTE (unknowns: WVC, near-surface air temperature, LST, and surface type).

atmosphere and land. The VIMS contains 12 channels, has spatial resolutions of 20 m (0.45–2.35 μm) and 40 m (3.5–12.5 μm), and has a ground coverage width of 60 km. VIMS data have application potential in environmental monitoring, resource surveys and surface ecological environment quality surveys (Zhang et al., 2017).

- MODIS, aboard the two polar orbit satellites Terra and Aqua, is an important sensor for observing global biological and physical processes as part of the US Earth Observing System (EOS) program and was first launched in 1999. The sensor has 36 channels with spatial resolutions of 250 m (0.62–0.876 μm), 500 m (0.459–2.155 μm), and 1 km (0.405–14.385 μm), and the scan width is 1330 km. In addition, MODIS has a wide range of applications in terms of urban heat islands, air temperature estimation/mapping, soil moisture, evapotranspiration estimation, and drought monitoring/estimation (Phan and Kappas, 2018) and has mature LST products (Wan et al., 2004).
- FY-4A is a second-generation geostationary meteorological satellite in the Fengyun series of Chinese meteorological satellites that was put into use in September 2017. The AGRI aboard FY-4A has 14 channels in total, with spatial resolutions of 0.5–1 km (0.45–0.90 μm), 2–4 km (1.36–4.0 μm), and 4 km (5.8–13.8 μm) and an area scanning time of 1 min (1000 km × 1000 km). The AGRI can be used to improve applications in a wide range of ocean, land, and atmosphere monitoring and in forecasting extreme weather, especially typhoons and thunderstorms (Yang et al., 2017).

Different satellite data have their own advantages. We choose MODIS data mainly because the it has stable performance, multiple TIR bands, and representative validation. The most important thing is that there are already mature LST products after more than 20 years of development and application since the 1990s. MODIS data can be used as a supplement to the reliable data in the abduction process to compensate for the deficiencies of other data in LST retrieval. After selecting the representative medium spatial resolution data, we considered the newly available high (VIMS) and low spatial resolution (AGRI) remote sensing data. The calibrations of the two new sensors onboard are unstable, and there are deviations in the center wavelength, and there are no corresponding mature LST products. To promote the research and development of the LST products of the new sensors as soon as possible, we deliberately made the corresponding analysis. The spatial resolution of VIMS data is high, but the time resolution is low, and some deviation in BTs has been found through research. The spatial observation of AGRI data has high temporal resolution, but low spatial resolution, and there are problems of unstable calibration on satellite. Although the data quality of these two sensors is not very high, they are also two very important data sources. These three datasets can be used to test the MDK-DL method. The application of MODIS data can achieve a comprehensive and representative analysis, and the application of AGRI and VIMS data can confirm that our method can achieve good results through avoiding deviations due to data quality.

Some TIR bands of AGRI and VIMS are not very stable in terms of satellite calibration. Therefore, we first use the ground station and MODIS LST products of the adjacent time as the corresponding ground data through calibration when the satellite transits. Then, for each dataset, we resample the high spatial resolution images to match the low spatial resolution images, thereby establishing a correspondence relationship of images information to increase the DL-NN training and test samples. The problem of further on-satellite correction can be avoided, and the retrieval calculation can be directly performed.

3.1.2. Auxiliary data

The auxiliary data mainly refer to the reliable LSTs that can be added in the DL-NN and the LSTs used in the validation. The data sources of the former can be obtained from a) reliable *in situ* ground truth measurements of LST, b) well-validated products, and c) assimilating data with high accuracy. The data sources of the latter can be obtained from a) simulation data from the MODerate spectral resolution atmospheric TRANsmittance mode (MODTRAN), b) MODIS LST products and c) *in situ* ground truth measurements of LST from meteorological observatory stations. In practical applications, abnormal and unrepresentative data should be eliminated to ensure that the selected data reflect all the physical conditions.

3.2. Geophysical logical reasoning

3.2.1. Reasoning process

A good geophysical parameter retrieval method should satisfy three conditions. First, the retrieval model should have physical meaning. Second, the retrieval equation should be mathematically solvable in theory. Third, the retrieval accuracy is high and meets application requirements. LST retrieval is based on the thermal radiance of the ground and the transmission of thermal radiance from the ground through the atmosphere to the remote sensor. The ground is not a blackbody. Thus, LSE must be considered when computing the thermal radiance emitted by the ground. The atmosphere has important effects on the received radiance at the level of the remote sensor (Dozier, 1981). Considering all these impacts, the general radiance transfer equation after simplification for remote sensing of LST can be formulated as follows (Qin et al., 2001b):

$$B_i(T_i)=\varepsilon_i\tau_i(\theta)B_i(T_s)+[1-\tau_i(\theta)](1-\varepsilon_i)\tau_i(\theta)B_i\left(T_a^{\downarrow}\right)+[1-\tau_i(\theta)]B_i(T_a), \tag{1}$$

where T_s is the LST, T_i is the BT in channel i, $\tau_i(\theta)$ is the atmospheric transmittance in channel i in viewing direction θ (zenith angle from nadir), ε_i is the LSE, $B_i(T_i)$ is the ground radiance, T_a is the effective mean atmospheric temperature, and $T_a^{\downarrow}$ is the downward effective mean atmospheric temperature. Qin et al. (2001b) made a reasonable simplification and analysis and concluded that using T_a instead of $T_a^{\downarrow}$ has little influence, so the equation can be depicted as Eq. (2). Thus, we simplified

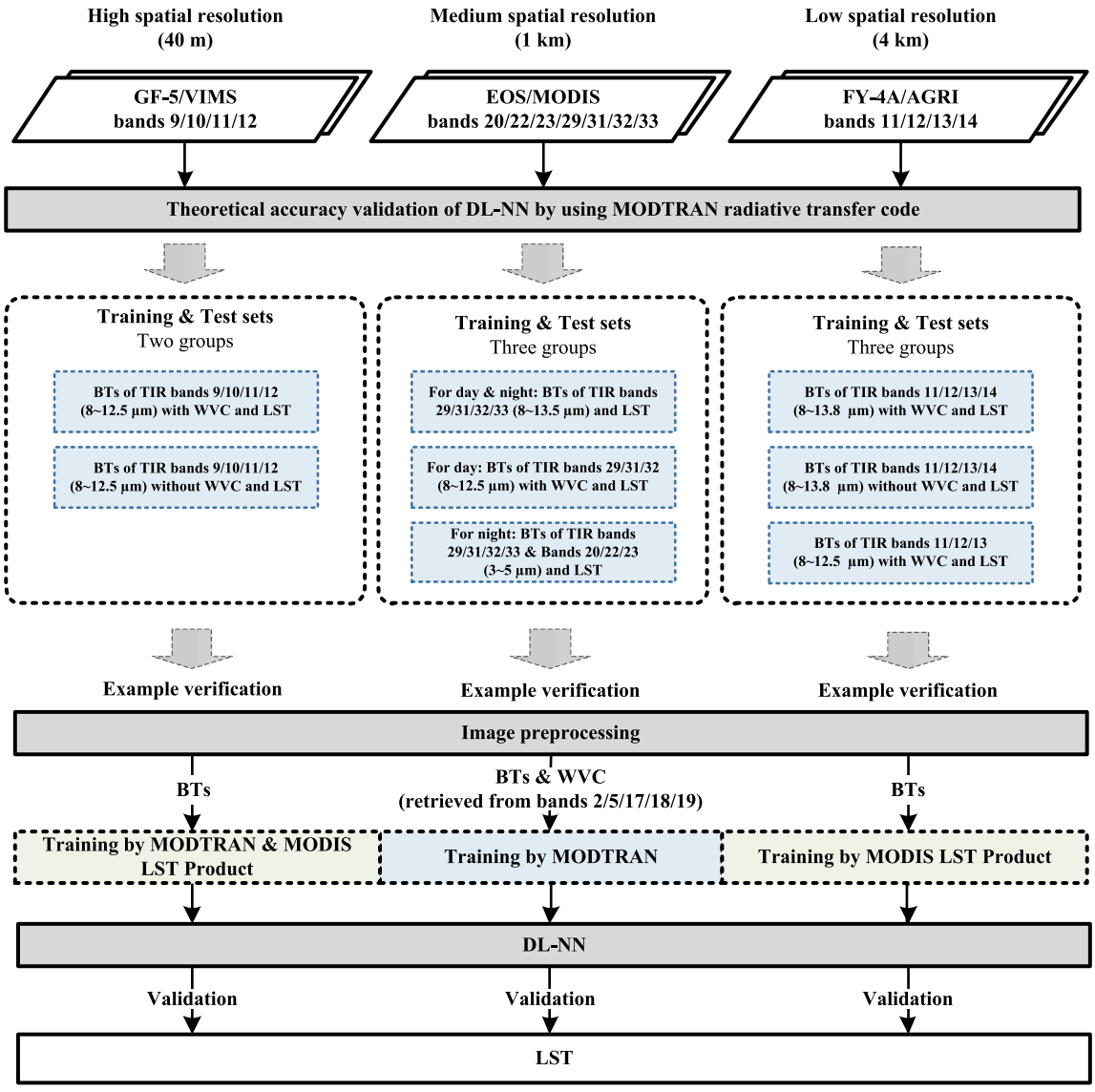

Fig. 4. The workflows. The overall process of MDK-DL method is applied to LST retrieval from three typical high-, medium- and low-spatial-resolution remote sensing datasets.

the radiation transmission process, which is shown in Fig. 3.

$$B_i(T_i) = \varepsilon_i\tau_i(\theta)B_i(T_s) + [1-\tau_i(\theta)](1-\varepsilon_i)\tau_i(\theta)B_i(T_a) + [1-\tau_i(\theta)]B_i(T_a) \tag{2}$$

The LSE (ε_i) of each band is unknown. If each type of ground surface is known, then the LSE of each band can be determined. Therefore, all unknown LSEs of different bands can be unified into a single unknown parameter (surface type).

$$\varepsilon_i = f(ground\_surface\_type) \tag{3}$$

High-precision surface type datasets and a stable relationship between LSEs at different wavelengths are critical factors. Moreover, LST retrieval depends on the accuracy of the classification algorithm.

The atmospheric transmittance (τ_i) of each band is unknown which can be determined by the WVC:

$$\tau_i = f(WVC) \tag{4}$$

Consequently, all unknown transmittances of different bands can be unified into a single unknown parameter (WVC). Obtaining a high-precision WVC dataset and ensuring a stable relationship between WVC and transmittance at different wavelengths are other critical tasks in LST retrieval algorithms. The algorithm also depends on the design of the WVC band. The atmospheric WVC can also be retrieved from the TIR channels (Eq. (5)) and is related to the channel BT differences (Jedlovec, 1990; Kleespies and McMillin, 1990). In addition, the effective mean atmospheric temperature (T_a) is mainly determined by the near-surface air temperature:

$$WVC = f(\Delta BT) \tag{5}$$

$$T_a = A_1 + B_1T_0 \tag{6}$$

where ΔBT represents a difference in observed BTs due to changing skin temperature, T_0 is the near-surface air temperature at two meters above ground, A_1 is a constant, and B_1 is a coefficient. Clearly, for different regions and seasons, the coefficient of eq. (6) is different. According to reciprocity theory, the effective mean atmospheric temperature T_a can also be determined by the BT (T_i) at the satellite. The expression can be depicted as Eq. (7) (Mao et al., 2007).

$$T_a \approx A_2 + B_2T_i \tag{7}$$

where T_i is the BT (which is known) at the satellite, A_2 is a constant, and B_2 is a coefficient. There is a strong constraint relationship between the

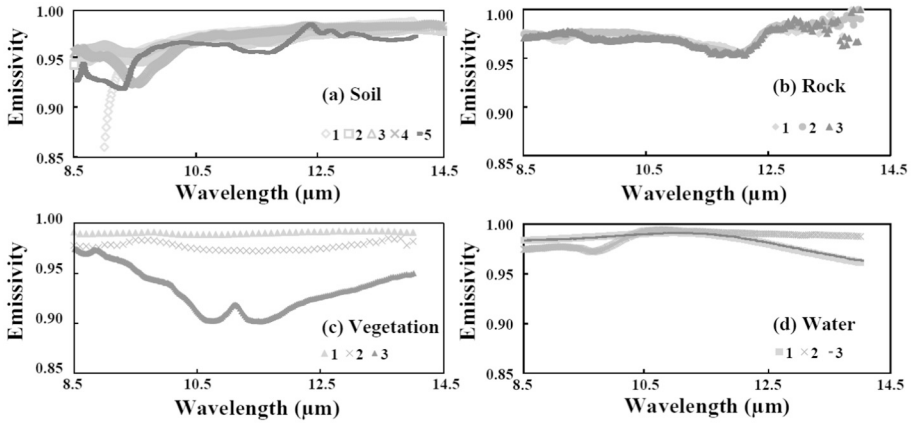

Fig. 5. (a) The LSE of the five kinds of soil. (b) The LSE of the three kinds of rocks. (c) The LSE of the three kinds of vegetation. (d) The LSE of the three kinds of water.

effective mean atmospheric temperature, the LST, near surface air temperature, the atmospheric WVC, and the BT at the satellite.

According to the above analysis, there are only four unknowns (LST, ground surface type, WVC and effective mean atmospheric temperature or near-surface air temperature) in the RTE. In theory, without any prior knowledge, the LST can be retrieved from four RTEs established from four TIR bands. If the ground surface type is obtained from visible light and near-infrared (NIR) bands, only three RTEs established from three TIR bands are needed to retrieve LST. If WVC is obtained from NIR bands, three RTEs established from three TIR bands are also sufficient to retrieve LST. If ground surface type and WVC are obtained from visible light and NIR bands, only two RTEs established from two TIR bands are needed to retrieve LST, and this technique is the split-window method. The geophysical parameters depend on each other, but we did not fully use the connections among these parameters in the previous algorithms. When we solve mathematical equations, the relationships cannot be depicted in a strict mathematical equation, so the error will be amplified, and the simplification of the Planck function also leads to some errors. To improve the accuracy, optimal computation methods (such as neural networks) are one of the best choices (Mao et al., 2008). In this study, we use the DL-NN algorithm to overcome the ill-conditioned difficulty and improve the accuracy in LST retrieval.

3.2.2. *Research approach*

Based on GLR, the overall process of the MDK-DL method applied to three typical high-, medium- and low- spatial-resolution remote sensing datasets is depicted in Fig. 4. First, we verify the theoretical accuracy of LST retrieval on the three datasets. MODTRAN is used to simulate TIR radiation transmission with input atmospheric profile data to simulate atmospheric parameters. Different satellite sensor technical parameters should be considered. Second, we conduct instance verification. Preprocessing, such as radiometric calibration, geometric correction, atmospheric correction, and reflectance calibration, is performed to obtain more accurate spectral information. The BTs of the three kinds of data are obtained. Third, when the training database of the simulation data (obtained by MODTRAN) matches the real situation in the example application, the database remains unchanged. Otherwise, reliable LST data are increased according to the principle of abduction (Fig. 1) to update the training database with the LSTs and the corresponding BTs of images. Then, the DL-NN for each dataset used for LST retrieval is well trained and can process the input parameters and output the results.

Different sensor channels have different characteristics, so the settings of parameters such as LST, WVC, LSE, and field of view (FOV) in MODTRAN are also different. Our study selected four land surface types (soil, vegetation, water and rocks) for the convenience of unified processing of different spatial resolution data. The LSEs of the different TIR bands of VIMS, MODIS and AGRI, as measured by the Jet Propulsion Laboratory, are taken as the input parameters of MODTRAN (Fig. 5).

Through preliminary research on these three datasets, we determined suitable and representative band combinations based on specific sensor performance. In the DL-NN training stage (which is described in Section 3.3), the training and test sample groups of these three datasets are as follows. For VIMS data, there are two groups: 1) the BTs of TIR bands 9, 10, 11, and 12 (8–12.5 μm) with WVC and 2) the BTs of TIR

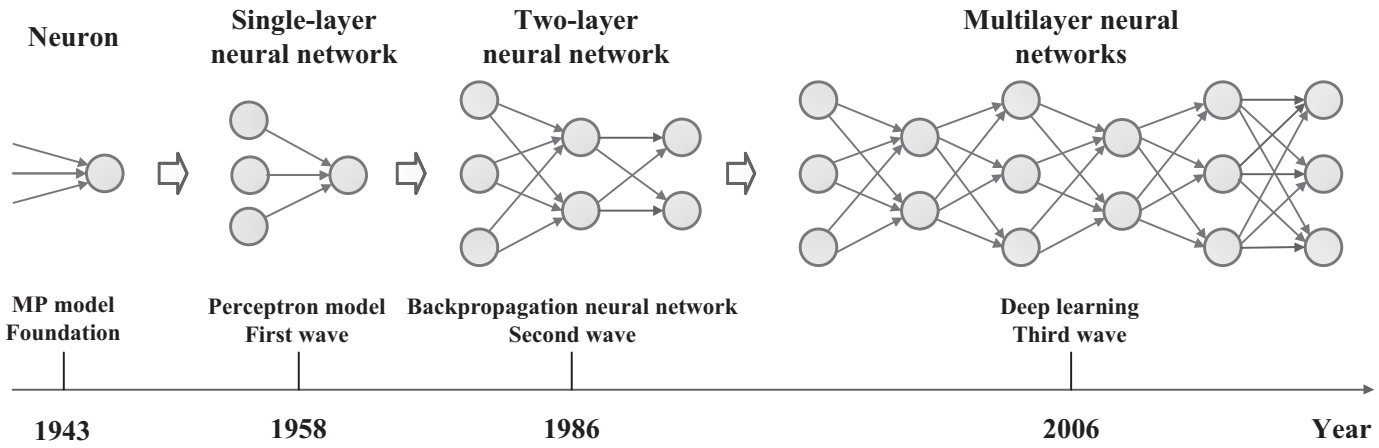

Fig. 6. Neural network development history. The neural network originated in the 1840s (McCulloch and Pitts, 1943), and the introductions of the perceptron (Rosenblatt, 1958), BP neural network algorithms (Rumelhart et al., 1986) and DL (Hinton, 2005) were the three climaxes in the development of neural networks.

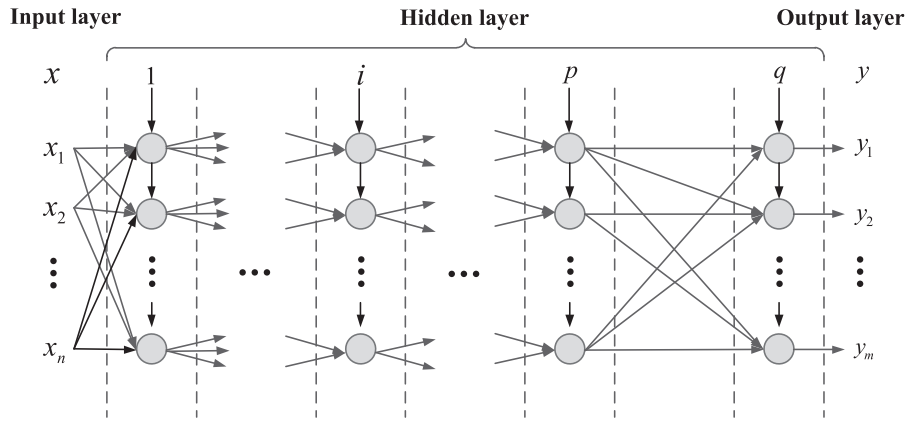

Fig. 7. Fully connected layer DL-NN.

bands 9, 10, 11, and 12 (8–12.5 μm) without WVC. For MODIS data, there are three groups: 1) the BTs of TIR bands 29, 31, 32, and 33 (8–13.5 μm) for day and night; 2) the BTs of TIR bands 29, 31, and 32 (8–12.5 μm) with WVC for day; and 3) the BTs of TIR bands 29, 31, 32, and 33, bands 20, 22, and 23 (3–5 μm) for night. For AGRI data, there are three groups: 1) the BTs of TIR bands 11, 12, 13, and 14 (8–13.8 μm) with WVC; 2) the BTs of TIR bands 11, 12, 13, and 14 without WVC; and 3) the BTs of TIR bands 11, 12, and 13 (8–12.5 μm) with WVC.

3.3. Optimal computation using the DL-NN algorithm

To optimize the calculation of the MDK-DL method, we consider machine learning technology, specifically the neural network algorithm. Neural networks are constantly evolving (deep learning is the third wave) (Fig. 6) and can be classified based on two factors: the network structure and learning style. In our study, the selected DL-NN is a feedback neural network with supervised learning. In general, DL-NN does not need to accurately grasp the relationship between input parameters and output parameters. It extracts information from training data, simplifies the complex physical model, and transforms the feature extraction process into automatic feature learning and application-dependent feature exploration (LeCun et al., 2015).

3.3.1. Establishment of the DL-NN

The deep learning convolutional neural network (DL-CNN), a feedback neural network with supervised learning, is one of the most popular pattern recognition methods. The convolution and pooling operation extract features from the input and hidden layers and use local connections and shared weights to effectively extract spatial and spectral information from an image. The greatest advantage of the DL-CNN is the ability to reduce the number of parameters and the complexity of the operation and provide strong fault tolerance, parallel computing power and self-assessment capabilities. Therefore, this method has high optimal computation or classification ability in target identification (Chen et al., 2014; Mao et al., 2007; Zhao and Du, 2016).

To reduce the number of calculations and accelerate the training speed, the DL-CNN model connects the pooling process after the convolution operation to reduce the amount of data. Each pooling layer corresponds to the previous convolution layer. Pooling can ensure invariance by reducing the resolution of feature maps (Zuo et al., 2016). Convolution and pooling operations are suitable mainly for feature extraction from two-dimensional images. Due to the low resolution of remote sensing images, it is unrealistic to identify features (except for extremely large object types, such as roads and rivers) by the geometric features in remote sensing images. Typically, we use the spectral information of remote sensing images to identify ground surface types. Hu et al. (2015) used deep convolutional neural networks to classify hyperspectral images directly in the spectral domain. The number of spectral segments must to be at least 10 or more if convolution and pooling are to be effective. In theory, the greater the number of spectral segments is, the better the effect. Considering that the number of bands we studied is small, we choose a fully connected layer neural network for optimal computation. As shown by the analysis in Section 3.2, function approximation and optimization are needed for the LST retrieval algorithm. Thus, we construct a fully connected layer DL-NN, as shown in Fig. 7.

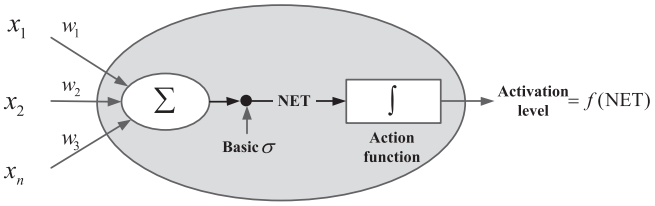

Fig. 8. A single neuron.

A fully connected layer DL-NN consists of one input layer, one output layer, and two or more hidden layers between the input and output layers. Fig. 8 shows the arrangement of an individual neuron, where the sum $(x \cdot w + \sigma)$ is commonly termed the net of the neuron. The input to the neuron can be either the actual input to system x or the output from other neurons in preceding layers.

Although many types of activation functions exist, the most common is the nonlinear sigmoid function shown in Eq. (8).

$$f(Net) = \frac{1}{1 + e^{-Net}} = \frac{1}{1 + e^{-(w \cdot x + \sigma)}} \tag{8}$$

The representation of the function by the neural network is accomplished by a set of individual neurons that have learned the appropriate response to an input. During the training phase, training patterns are sequentially presented to the network. After all patterns have been presented, the interconnecting synaptic strength (weight) of each neuron is adjusted. The DL-NN constructed in our study utilizes the Kalman filtering technique, which is a recursive minimum mean square estimation procedure. This technique greatly improves the accuracy, training speed and stability of the DL-NN. More details can be found in Tzeng et al. (1994).

3.3.2. Training and testing of the DL-NN

The training and testing of the DL-NN algorithm is not complex and can be divided into four basic steps. The operation is applicable to both theoretical accuracy analysis and practical application.

- Reliable training and test databases (of LSTs and their corresponding BTs with/without WVC) were obtained for theoretical accuracy

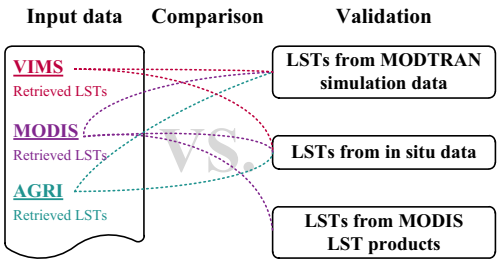

Fig. 9. Validation strategy of three datasets (only MODIS data are used for cross-validation to ensure the independence of validation).

analysis and practical application. The acquisition time of the supplementary data must be matched with the imaging time of the research data, and the time values can be corrected based on the value measured at meteorological stations. Moreover, if necessary, resampling is performed to ensure the matching of different images.

- Training and testing of the DL-NN.
- Computing the BTs from VIMS, MODIS, and AGRI TIR data. In addition, the WVC data in MODIS data were retrieved from MODIS bands 2, 5, 17, 18, and 19.
- Retrieving the LST by using the DL-NN.

The specific settings are provided below. The grouping information of each dataset is shown in Section 3.2. This operation is applicable to both theoretical accuracy analysis and practical application. For VIMS simulation data (from MODTRAN), we specified the following settings for a total of 25,088 sets of training data and 6720 sets of test data: LST (280–325 K), WVC (0.1–3.5 g/cm^2), angle of view (0–7.5°), and LSE (vegetation, water, soil, and rocks). The small angle of the VIMS, of which the FOV is approximately 15°, can be ignored without losing noticeable accuracy in example applications, and similar conclusions were reached in Tang and Li (2018). For the MODIS simulation data, we specified the following settings for the training data (158,400 sets) and testing data (72,600 sets): LST (280–330 K), WVC (0.1–4.5 g/cm^2), and LSE (vegetation, water, soil, and rocks). The viewing angle affects the retrieval accuracy. According to Masuoka et al. (1998), the footprint of MODIS pixels in TIR bands is approximately 1 km by 1 km at nadir, 1.3 km (along track) by 1.6 km (across track) at a viewing zenith angle near 40°, and 1.7 km by 3.3 km at a viewing angle of 60°. Because the maximum MODIS viewing angle is 65° from nadir, pixels with viewing angles larger than 45° account for nearly 30% of the total pixels, or almost 50% of the total coverage area within each swath. To cover all ranges, the analog angle setting range is 0–65°. According to the study and comparison of MODIS and AGRI data, the simulation analysis of MODIS data is representative of large FOV observations. The theoretical accuracy of AGRI data is similar to that of MODIS. In addition, due to the inaccurate calibration of the AGRI data at satellite, the actual application does not depend on the training database simulated by MODTRAN, Therefore, it is not necessary to present a simulation analysis similar to that of MODIS, but we did an evaluation when evaluating the theoretical

Table 1
Retrieval errors of MODTRAN simulated data from VIMS TIR bands 9, 10, 11, and 12 with WVC (SD: standard deviation of the fit). For all columns (validation dataset), the minimum MAE with maximum R and minimum SD is indicated in bold; colors denote the level of performance: red denotes the minimum MAE with maximum R and minimum SD, and blue denotes the maximum MAE with minimum R and maximum SD. These colors are also used in Table 2-Table 5.

Hidden layers, nodes (0~7.5°), and retrieval errors (from VIMS bands 9, 10, 11, and 12 with WVC)

Hidden	400			500			600			700			800		
layers	MAE	SD	R	MAE	SD	R	MAE	SD	R	MAE	SD	R	MAE	SD	R
3	0.13	0.11	1.000	0.12	0.10	1.000	0.10	0.11	1.000	0.09	0.09	1.000	0.09	0.07	1.000
4	0.28	0.33	1.000	0.07	0.09	1.000	0.07	0.06	1.000	**0.04**	**0.04**	**1.000**	0.06	0.08	1.000
5	0.08	0.07	1.000	0.39	0.32	0.999	0.09	0.26	1.000	0.07	0.14	1.000	0.16	1.60	0.993
6	0.21	0.85	0.998	0.21	0.44	0.999	0.09	0.14	1.000	0.08	0.27	1.000	0.08	0.15	1.000
7	0.39	2.93	0.978	0.10	0.43	0.999	0.14	0.33	1.000	0.25	2.86	0.979	0.19	0.26	1.000
8	0.08	0.13	1.000	0.16	0.22	1.000	0.10	0.25	1.000	0.66	1.74	0.991	0.28	0.87	0.998
9	0.16	0.27	1.000	0.44	1.31	0.995	0.36	0.72	0.998	0.28	1.14	0.996	0.37	0.40	0.999
10	0.65	2.19	0.987	0.20	0.53	0.999	0.25	0.27	1.000	0.25	0.47	0.999	0.38	0.92	0.997

Table 2
Retrieval errors of MODTRAN simulated data from VIMS TIR bands 9, 10, 11, and 12 without WVC.

Hidden layers, nodes (0~7.5°), and retrieval errors (from VIMS bands 9, 10, 11, and 12 without WVC)															
Hidden	400			500			600			700			800		
layers	MAE	SD	R	MAE	SD	R	MAE	SD	R	MAE	SD	R	MAE	SD	R
3	0.42	0.37	0.999	0.45	0.39	0.999	0.35	0.37	0.999	0.34	0.35	0.999	**0.31**	**0.34**	**0.999**
4	0.32	0.34	0.999	0.35	0.37	0.999	0.32	0.37	0.999	0.36	0.55	0.999	0.32	0.37	0.999
5	0.57	0.64	0.998	0.35	0.49	0.999	0.39	0.66	0.998	0.34	0.65	0.999	0.43	1.29	0.995
6	0.38	0.64	0.999	0.54	3.47	0.970	0.40	0.50	0.999	0.93	7.94	0.864	0.71	2.17	0.987
7	0.40	0.46	0.999	0.64	1.22	0.995	0.73	2.57	0.982	0.99	1.83	0.989	0.92	3.28	0.970
8	0.47	0.71	0.998	0.64	1.91	0.989	2.35	28.97	0.467	0.97	1.56	0.991	1.14	2.41	0.982
9	1.12	6.34	0.905	1.25	7.14	0.885	0.73	0.98	0.996	1.03	2.06	0.986	0.94	1.09	0.995
10	0.91	2.32	0.984	0.78	2.16	0.987	0.75	0.90	0.996	1.11	2.38	0.982	1.22	6.24	0.909

accuracy.

3.4. Validation approach

The validation of the retrieved LST involves evaluating the accuracy and representativeness of the true values of the satellite pixels and the ground. The universality of the LST retrieval algorithm is closely related to the quality of the verification results. The land surface is generally a nonhomogeneous, nonisothermal surface, resulting in a very complicated validation process. It is extremely difficult to obtain *in situ* ground truth measurements comparable to the pixel size of VIMS, MODIS and AGRI data at the satellite pass. An alternative practical approach is to use simulated data generated by atmospheric simulation programs such as MODTRAN. This program can simulate the thermal radiance reaching

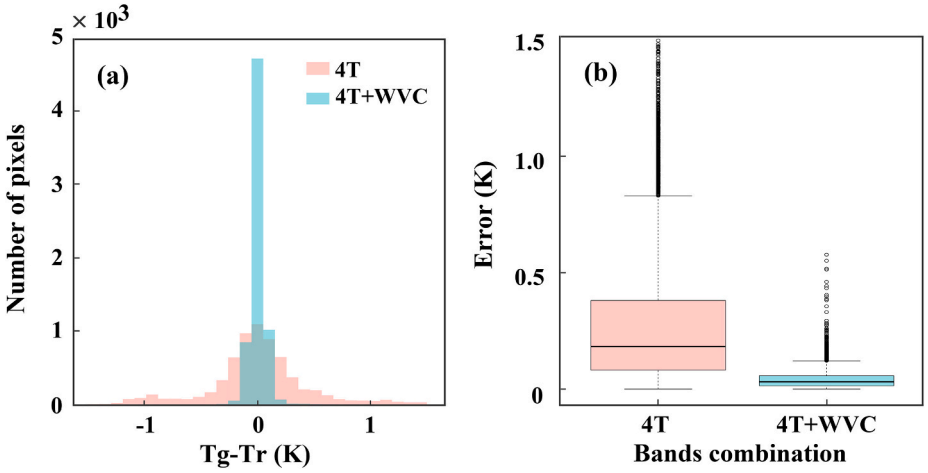

Fig. 10. (a) Histogram of the difference between ground temperature (Tg) and the retrieved LST (Tr). (b) A boxplot of the absolute difference between Tg and Tr; the x-axis is the band combination, where "4 T" represents VIMS TIR bands 9, 10, 11, and 12 without WVC and "4 T + WVC" represents VIMS TIR bands 9, 10, 11, and 12 with WVC.

Table 3
Retrieval errors of MODTRAN simulated data from MODIS TIR bands 29, 31, 32, and 33.

Hidden layers, nodes (0~65°), and retrieval errors (from MODIS bands 29, 31, 32, and 33)															
Hidden layers	400			500			600			700			800		
	MAE	SD	R	MAE	SD	R	MAE	SD	R	MAE	SD	R	MAE	SD	R
3	1.04	0.91	0.994	0.95	0.87	0.995	0.97	0.87	0.995	0.93	0.87	0.995	0.91	0.86	0.995
4	0.92	0.85	0.995	0.93	0.87	0.995	0.90	0.84	0.996	0.93	0.85	0.995	0.89	0.85	0.996
5	0.94	0.86	0.995	0.97	0.89	0.995	0.91	0.84	0.995	0.90	0.85	0.995	0.90	0.86	0.995
6	0.94	0.88	0.995	0.92	0.85	0.995	0.90	0.84	0.996	0.92	0.87	0.995	**0.87**	**0.89**	**0.995**
7	0.96	0.91	0.995	0.98	0.87	0.995	0.91	0.85	0.995	0.92	0.86	0.995	0.91	0.94	0.995
8	0.92	0.85	0.995	0.95	0.90	0.995	0.95	0.87	0.995	0.92	0.85	0.995	0.90	0.85	0.995
9	0.91	0.85	0.995	0.94	0.87	0.995	0.90	0.85	0.995	0.91	0.85	0.995	0.92	0.85	0.995
10	1.08	9.81	0.793	1.07	0.93	0.994	0.92	0.85	0.995	0.93	1.41	0.992	0.90	0.85	0.995

a remote sensor at the satellite level for input profile data with known ground thermal properties. In addition, cross-validation between the retrieved results and verified LST products is a useful approach. Specifically, our validation is divided into three parts (Fig. 9). The first part is a validation based on MODTRAN simulation data, which demonstrates the theoretical accuracy of our method. The second part is a cross-validation based on LST products. The third part is a validation based on *in situ* data, confirming the actual application accuracy of our method. Our validation is conducted based on four statistical metrics: the coefficient of determination (R^2), RMSE, mean absolute error (MAE), and correlation coefficient (R).

4. Results and validation

4.1. Theoretical accuracy analysis and validation

4.1.1. Theoretical accuracy from small viewing directions

The theoretical accuracy analysis of VIMS data is used to verify the LST retrieval effect of a small viewing direction. We divide the test data into two groups to test the accuracy of the DL-NN. Specifically, the first group uses bands 9, 10, 11, and 12 with WVC, and the second group uses bands 9, 10, 11, and 12 without WVC. Table 1 and Table 2 summarize the retrieval errors for the first and second groups, respectively. Table 1 shows that the accuracy is the highest when the number of hidden layers is 4 and the number of hidden nodes is 700 (MAE = 0.04 K, SD = 0.04, and $R = 1.000$). Table 2 shows that the accuracy is the highest when the number of hidden layers is 3 and the number of hidden nodes is 800 (MAE = 0.31 K, SD = 0.34, and $R = 0.999$). Therefore, according to the preliminary analysis, we can consider combining the four TIR channels to retrieve the LST in practical applications. In addition, this analysis verifies the feasibility of our previous analysis (GLR driven by expert knowledge).

In addition, we made statistics on the errors of the two groups with the highest precision in Table 1 and Table 2 (as shown in Fig. 10). **Fig. 10**a clearly shows that the retrieval accuracy of the four TIR bands can be improved by increasing WVC, where most of the error distribution accumulates near zero. The same conclusion can be obtained from the analysis in **Fig. 10**b. This figure shows that although the average level of the two band combinations is within 0.5 K, the error fluctuation after increasing WVC is greatly reduced, and the abnormal value is also reduced.

4.1.2. Theoretical accuracy from large viewing directions

Through verification, we found that the results of the theoretical accuracies of the AGRI and MODIS data are similar for similar combinations. In addition, due to the instability of AGRI on-board calibration and center wavelength deviation and to avoid redundancy, the analysis results of MODIS are mainly displayed here. We use its theoretical accuracy as the representative of different band combinations with large viewing directions. We divide the test MODIS data into three groups to assess the accuracy of the MDK-DL method. For the first group, the combination of bands 29, 31, 32, and 33 is suitable for day and night. Table 3 presents a summary of the retrieval errors for the first group and shows that the accuracy is highest (MAE = 0.87 K, SD = 0.89, and $R = 0.995$) when the number of hidden layers is 6 and the number of hidden nodes is 800.

Table 4
Retrieval errors of MODTRAN simulated data from MODIS TIR bands 29, 31, and 32 with WVC.

Hidden layers, nodes (0~65°), and retrieval errors (from MODIS bands 29, 31, and 32 with WVC)															
Hidden layers	400			500			600			700			800		
	MAE	SD	R	MAE	SD	R	MAE	SD	R	MAE	SD	R	MAE	SD	R
3	0.71	0.76	0.997	0.71	0.76	0.997	0.70	0.76	0.997	0.70	0.78	0.997	0.68	0.75	0.997
4	0.70	0.76	0.997	0.70	0.76	0.997	0.71	0.76	0.997	0.72	0.78	0.997	0.77	1.17	0.994
5	0.69	0.77	0.997	0.75	0.80	0.996	0.92	1.61	0.990	0.89	2.37	0.982	0.95	1.95	0.987
6	**0.67**	**0.74**	**0.997**	0.72	0.84	0.996	0.73	0.76	0.997	1.05	2.64	0.977	0.72	1.11	0.995
7	0.75	1.12	0.995	0.70	0.88	0.996	0.88	1.66	0.990	0.93	1.49	0.991	0.74	1.79	0.989
8	0.88	2.20	0.984	0.91	2.70	0.977	0.72	0.95	0.996	1.04	1.53	0.990	0.98	1.40	0.991
9	0.80	1.62	0.991	0.78	0.90	0.996	0.73	2.86	0.975	0.94	1.39	0.992	0.98	1.10	0.994
10	0.81	1.10	0.994	0.74	0.87	0.996	0.79	1.02	0.995	1.12	1.97	0.986	0.71	0.81	0.997

For the second group (the combination of bands 29, 31, and 32 with WVC), the TIR band is affected mainly by atmospheric water vapor absorption, so we consider atmospheric water vapor as an input parameter. The ratios of atmospheric water vapor-influenced channels (bands 2, 5, 17, 18, and 19) can be used to estimate the WVC from the MODIS data, which are suitable mainly for daytime. Table 4 provides a summary of retrieval errors for the second group, and shows that the accuracy is highest (MAE = 0.67 K, SD = 0.74 K, and R = 0.997) when the number of hidden layers is 6 and the number of hidden nodes is 400. Comparison of the data from this table with that in Table 3 shows a difference between the two test groups: adding atmospheric water vapor information significantly improves the accuracy of LST retrieval. Therefore, in practical applications, as much accurate atmospheric water vapor information as possible should be added to improve the LST retrieval accuracy.

MODIS bands 20, 22, and 23 are in the atmospheric window in the 3.5–4.2 μm medium wavelength range. The midinfrared band is affected by the sun and is therefore suitable mainly for night use. The third group is the combination of bands 20, 22, and 23 (3–5 μm) and bands 29, 30, 31, 32, and 33 (8–13.5 μm), and this combination is suitable for LST retrieval during the night. We analyzed different combinations: (1) bands 20, 29, 31, and 32; (2) bands 20, 22, 31, and 32; (3) bands 20, 22, 29, 31, and 32; (4) bands 20, 22, 29, 31, 32, and 33; (5) bands 20, 22, 23, 29, 31, and 32; and (6) bands 20, 22, 23, 29, 31, 32, and 33. The accuracy of the combination of bands 20, 22, 23, 29, 31, and 32 and the combination of bands 20, 22, 23, 29, 31, 32, and 33 is similar and better than that of the other combinations, but band 33 is substantially affected by carbon dioxide. From the perspective of reducing the quantity of data and the number of calculations, we choose the combination of bands 20, 22, 23, 29, 31, and 32. Here, we provide detailed information for only this combination (above), which is presented in Table 5. This table clearly shows that the accuracy is highest when the number of hidden layers is 6 and the number of hidden nodes is 700 (MAE = 0.55 K, SD = 0.56 K, and R = 0.998).

In addition, we compare the Tg and Tr of the optimal layer and the node settings of the three combinations, as shown in Fig. 11. **Fig. 11**a clearly shows that the retrieval accuracy of the three TIR bands can be improved by increasing WVC or the TIR bands in the atmospheric window, where most of the error distribution accumulates near zero. The same conclusion can be obtained from the analysis in **Fig. 11**b. The figure also shows that although the average level of the three combinations is within 1 K, the error fluctuation after increasing the WVC or TIR bands in the atmospheric window is greatly reduced. However, compared to the first combination, the latter two combinations have more outliers. The reason for this may be that the accuracy of the obtained atmospheric water vapor information is not sufficiently high, and the input of multiple TIR bands also adds some redundant information.

In summary, band combinations with high theoretical accuracy can be considered for application. However, in fact, the availability, accuracy, redundancy, *etc.*, of the data should also be considered, and the application should consider the actual circumstances.

4.1.3. Summary of the validations based on MODTRAN simulation data

One of our validation approaches (Section 3.4) uses MODTRAN simulation data to prove the theoretical accuracy of our method. To summarize the validations above based on MODTRAN simulation data,

Table 5
Retrieval errors of MODTRAN simulated data from MODIS bands 20, 22, 23, 29, 31, and 32.

Hidden layers, nodes (0~65°), and retrieval errors (from MODIS bands 20, 22, 23, 29, 31, and 32)															
Hidden layers	400			500			600			700			800		
	MAE	SD	R	MAE	SD	R	MAE	SD	R	MAE	SD	R	MAE	SD	R
3	0.81	0.65	0.997	0.80	0.63	0.997	0.73	0.61	0.997	0.61	0.64	0.998	0.68	0.59	0.997
4	0.80	0.64	0.997	0.73	0.99	0.995	0.67	0.70	0.997	0.62	0.60	0.998	0.76	2.20	0.984
5	0.91	0.67	0.996	0.93	0.64	0.996	0.62	0.55	0.998	0.58	0.56	0.998	0.66	1.43	0.993
6	0.74	0.63	0.997	0.56	0.53	0.998	0.57	0.54	0.998	**0.55**	**0.56**	**0.998**	0.57	0.53	0.998
7	0.61	0.56	0.998	0.68	3.30	0.968	0.70	0.61	0.997	0.63	0.65	0.998	0.65	0.64	0.998
8	0.86	1.03	0.995	0.68	1.31	0.994	0.65	0.61	0.998	0.61	0.75	0.997	0.61	1.02	0.996
9	0.75	0.66	0.997	0.69	0.60	0.998	0.59	0.61	0.998	0.60	0.59	0.998	0.63	0.60	0.998
10	0.71	0.65	0.997	0.76	0.74	0.997	0.64	0.74	0.997	0.70	1.18	0.994	0.64	1.50	0.992

we performed scatter diagram analysis (Fig. 12) of the error when each data point is in the optimal band combination and the optimal structure of its DL-NN. For VIMS data, the best band combination is four TIR bands and atmospheric water vapor information (namely, bands 9, 10, 11, and 12 with WVC), and the accuracy is highest when the number of hidden layers is 4 and the number of hidden nodes is 700 (MAE = 0.04 K, SD = 0.04, and $R = 1.000$). For MODIS data, the best band combination is six bands (namely, bands 20, 22, 23, 29, 31, and 32), and the accuracy is highest when the number of hidden layers is 6 and the number of hidden nodes is 700 (MAE = 0.55 K, SD = 0.56 K, and $R = 0.998$). For AGRI data, the best band combination is four TIR bands (namely, bands 11, 12, 13 and 14), and the accuracy is highest when the number of hidden layers is 10 and the number of hidden nodes is 700 (MAE = 0.70 K, SD = 1.19 K, and $R = 0.995$).

In **Fig. 12**a, due to the high accuracy of the simulation validation of VIMS, the data points overlap substantially. Shown from **Fig. 12**b and **Fig. 12**c, the retrieval error is relatively large at both ends, and there are two main reasons. First, the training and test data at both ends are relatively small. Second, when the observation angle is greater than 55° and the WVC is higher than 2.5 g/cm^2, the retrieval error will increase. In the simulation, we set the WVC on the vertical path. Shown in Fig. 13, when the WVC on the vertical path is 2.5 g/cm^2 and the observation angle is greater than 60°, which means that the WVC on the slant path exceeds 5 g/cm^2 and the transmittance is already lower than 0.4 for some TIR bands (Mao et al., 2005). It is no longer suitable for LST retrieval when the transmittance is relatively low. When the observation angle is relatively large and the WVC is high, the retrieval can still maintain a certain accuracy to prove the robust of this method. Clearly, in simulation and practical applications, we can also choose to remove this part of the retrieval result to improve accuracy. In general, the simulation validation accuracy of the three datasets is high, which proves the feasibility of our proposed method.

4.2. Practical analysis and validation

The LST abduction we propose is universal, and we selected the Bohai Rim region to demonstrate the application. This region is located around the Bohai Sea in the Jing-Jin-Ji area (namely, Beijing, Tianjin and Hebei Provinces), Liaodong Peninsula and Shandong Peninsula in northern China (Fig. 14). As one of China's agriculturally developed regions, the Bohai Rim region has a temperate monsoon climate. More importantly, there are many ground observation sites, and we have a wealth of prior knowledge about the area. Therefore, the region is conducive to verification analysis. Considering that the quality of MODIS data is the best, we take the LST retrieval process of MODIS data as the research object, and the main results of the other two datasets are briefly discussed.

4.2.1. Practical analysis, cross-validation, and in situ validation of MODIS data

The DL-NN trained above is used to retrieve the LST from MODIS image data. Specifically, the midinfrared data (bands 20, 22, and 23; 3–5 μm), TIR data (bands 29, 31, 32, and 33; 8–13.5 μm) and atmospheric water vapor data (retrieved from bands 2, 5, 17, 18, and 19; 0.8–1.25 μm) of EOS/MODIS are selected and used as the input parameters of the corresponding DL-NN. The inputs of the DL-NN are the

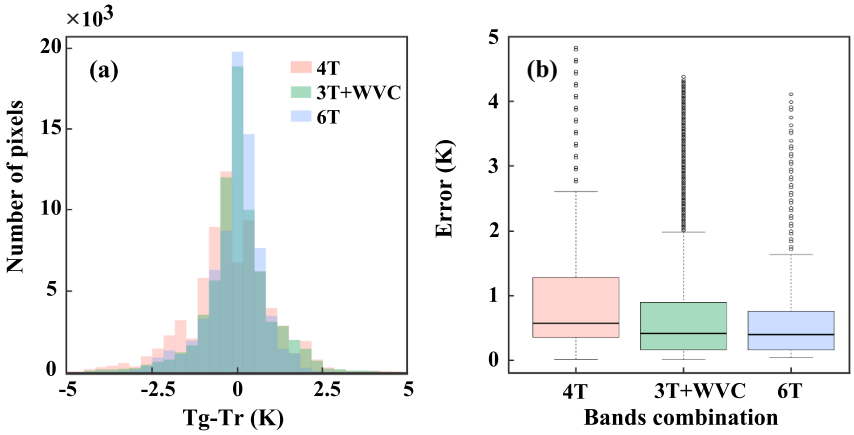

Fig. 11. (a) Histogram of the difference between Tg and Tr. (b) A boxplot of the absolute difference between Tg and Tr; the x-axis is the band combination, where "4 T" represents MODIS TIR bands 29, 31, 32, and 33; "3 T + WVC" represents MODIS TIR bands 29, 31, and 32 with WVC; and "6 T" represents MODIS bands 20, 22, 23, 29, 31, and 32.

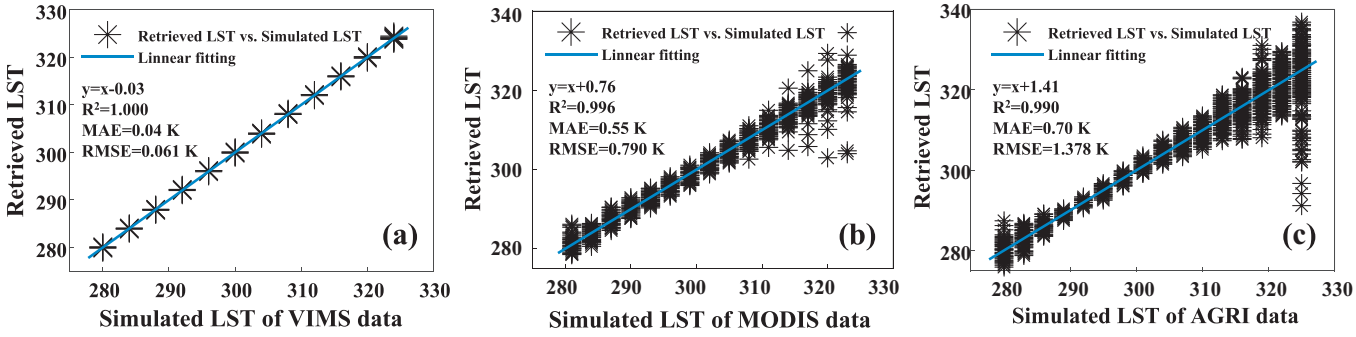

Fig. 12. Validations based on MODTRAN simulation data of (a) VIMS data, (b) MODIS data and (c) AGRI data when each data point has the optimal band combination and the optimal hidden layer and nodes of the well-trained DL-NN.

BT (BT_i, i=20, 22, 23, 29, 31, 32, 33) and the WVC retrieved from MODIS bands 2, 5, 17, 18, and 19, and the output is the LST. We select the Terra/MODIS and Aqua/MODIS images from September 9, 2005 (for daytime), and October 4, 2018 (for nighttime) for examples. **Fig. 15**b and **Fig. 16**a show the MODIS LST products MOD11A1 and MYD11A1, respectively. The LSTs extracted in Kelvin with the generalized split-window algorithm (Wan and Dozier, 1996) represent the differences in LSTs between the DL-NN retrieved results and MODIS LST products.

As Fig. 15 and Fig. 16 show that the spatial distribution of the retrieved LSTs is similar to the MODIS products. Moreover, Fig. 17 indicates that most of the difference values between MODIS LST products and the retrieved LSTs are within −1–1 K. The retrieved LSTs of a few areas were slightly higher than the MODIS LST products in the main agricultural production area of China. In September and October, straw burning incidents were common in North China after the autumn harvest. Ground surveys support this point, and we can also see the slight aerosol caused by burning straw from the RGB image (**Fig. 15**a). Thus, we can infer that the MODIS LST product algorithm underestimates the true LST because the LSEs are changed due to the burning of straw on the ground. Fortunately, the LST retrieval method based on the combination of DL-NN and RTE can overcome this shortcoming. The four band combinations for retrieving LSTs are effective and reasonable, and the band combination with added water vapor information is closer to the actual situation. In conclusion, the generalized split-window LST algorithm (Wan and Dozier, 1996) achieves a high accuracy of MODIS LST products, but the retrieval process is relatively complex and requires more parameters. Our method achieves an accuracy close to that of this algorithm and can simplify the process.

Moreover, we sampled stations under clear skies and with flat terrain, representative and single feature types for the *in situ* validation of the LSTs retrieved from MODIS data. Fig. 18 displays the validation of the best band combination, namely, MODIS TIR bands 29, 31, and 32 with WVC. The figure shows that *in situ* validation has a high-precision fitting accuracy (R^2 = 0.902) with an MAE of 1.12 K and an RMSE of 1.12 K.

4.2.2. *Practical analysis and in situ validation of AGRI and VIMS data*

We have analyzed the AGRI and VIMS datasets and found that the on-board calibrations of AGRI and VIMS instruments are unstable and the center wavelengths are offset. Therefore, the deviation from the model simulation data is relatively large, which is difficult. Therefore, the deviation between actual data and model simulation data is relatively large, and it is difficult to directly perform retrieval calculation and analysis with simulated data. Another advantage of the DL-NN algorithm does not require the relationship between the input parameters and the output parameters to be specified and can perform optimal computation. Therefore, based on the principle of LST abduction, we have established a new training database by supplementing reliable LST data and the BT of the images. The abduction was verified to have a good internal correction effect, which effectively avoided the BT deviation.

Considering that the two data-processing methods are similar, we take the AGRI as an example to illustrate the LST abduction-based correction process. MODIS LST products have been widely recognized, applied, and validated (Wan, 2008; Wan et al., 2002). AGRI is a sensor mounted on a geostationary satellite, and its time resolution is very high (a local area is higher than 35 min), so the MODIS LST products

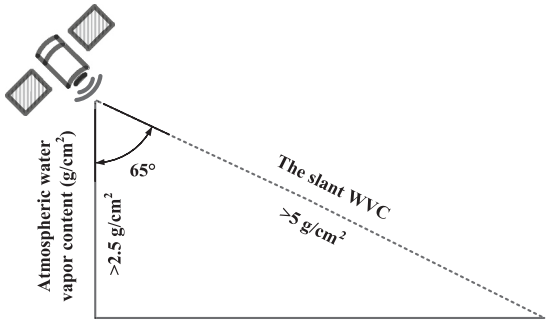

Fig. 13. Relationship of WVC between in vertical path and in slant path.

Fig. 14. The location of the Bohai Rim region.

corresponding to the time are very good ground data. We resampled different data and unified the resolution, and then used latitude, longitude and time as control conditions, and selected about 30 sceneries in eastern China to build a training and testing database. Especially for the VIMS data with a high spatial resolution, we have added images around the Bohai Sea region to provide the retrieved LSTs corresponding to sufficient and reliable validation station data. After training and testing, the best band combination (AGRI TIR bands 11, 12 13, and 14) is used to make the retrieval application. As shown in Fig. 19, the LSTs are 290.02–329.54 K. This figure provides a wealth of information about the spatial distribution of retrieved LSTs. That is, developed urban areas have an urban heat island effect due to high urbanization and have LSTs that are higher than those of surrounding areas.

Similarly, we did the same processing and analysis for VIMS data. The best band combination (TIR bands 9, 10, 11, and 12) is used to make the retrieval application. As shown in Fig. 20, the LST range is 280–330 K, and the geometric information in the resulting picture is rich. We can clearly identify the built-up area of Dongying City, Shandong Province (whose LSTs are higher than those of the surrounding area). Geometric information on the surrounding aquaculture areas, farmland, and various bodies of water (including reservoirs, rivers and the Bohai Sea) is also clearly depicted.

we sampled stations under clear skies and with flat terrain, representative and single feature types for validation. For AGRI data, Fig. 21a displays the *in situ* validation of the best band combination, namely, AGRI TIR bands 11, 12, 13, and 14. The MAE between the retrieved LST and ground LST is 1.78 K, the R^2 is 0.938, and the RMSE is 1.76 K. Fig. 21b displays the validation of the best band combination of VIMS data, namely, TIR bands 9, 10, 11, and 12. The figure shows a good fitting accuracy ($R^2 = 0.959$) with an MAE of 1.61 K and an RMSE of 1.61 K.

In summary, for MODIS data, our method achieves high validation accuracy both in the *in situ* validation and the cross-validation to MODIS LST products. Since LST retrieval from both AGRI and VIMS data involves time correction based on ground stations and image resampling, the extracted information will inevitably contain certain errors. Moreover, the new AGRI and VIMS data are not very stable, and their onboard calibration needs to be further improved. This problem can be remedied by further increasing the quantity of reliable data used to retrain the DL-NN. If more reliable LSTs (such as MODIS LST products) can be added and if DL-NN training can be performed in different regions and seasons, the LST retrieval accuracy will be further improved.

5. Discussion and conclusions

5.1. Discussion

With the foundation of GLR, MDK-DL method does not need prior knowledge of the two parameters (LSE and transmittance). Our method does not need to obtain complex parameter information, uses only the BT on the satellite as input parameters. Clearly, if the satellite sensor has atmospheric water vapor band that can provide water vapor information, the retrieval accuracy can be greatly improved. Taking MODIS data as an example, we studied the relationship between the LST and BT at the satellite. The simulation results are displayed in Fig. 22. For different ground surface types (with different LSEs) and atmospheric conditions, the same LST corresponds to a plurality of BTs at the satellite, and the largest range of variation is more than 25 degrees. Clearly, if a single-band statistical regression retrieval method is used, the average error is more than 10 K. Therefore, the traditional method requires atmospheric parameters and LSE as *a priori* knowledge. In other words, an LST value corresponds to a vast number of measurable BTs, and the range fluctuates greatly, which will bring some uncertainty to LST retrieval.

Moreover, if the different LSTs are considered to be one type, LST

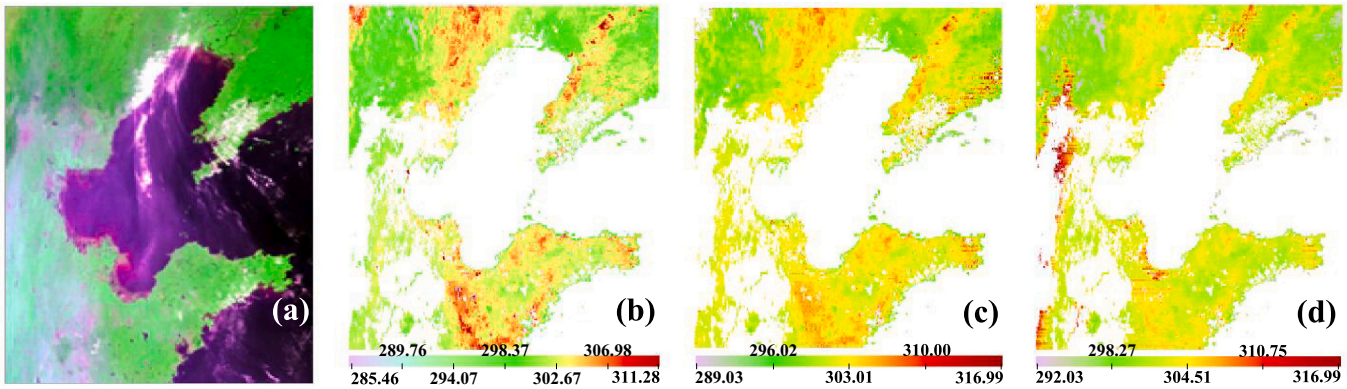

Fig. 15. (a) RGB image for the daytime of September 9, 2005. (b) MODIS product (MOD11A1). (c) LST retrieved from MODIS bands 29, 31, 32, and 33 using the DL-NN algorithm. (d) LST retrieved from bands 29, 31, and 32 with WVC using the DL-NN algorithm. The white areas denote invalid values.

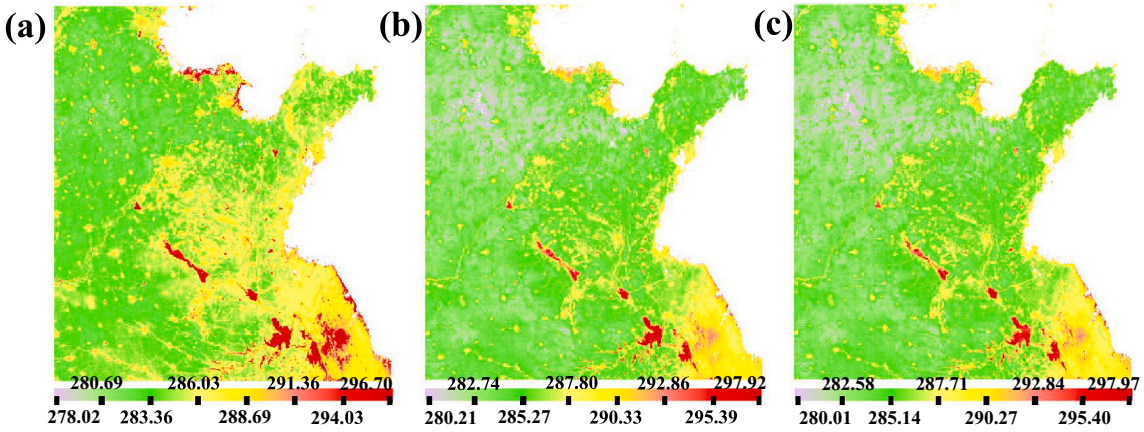

Fig. 16. Retrieved LSTs (K) for the nighttime of October 4, 2018. (a) LST product (MYD11A1). (b) LST retrieved from MODIS bands 29, 31, 32, and 33 using the DL-NN algorithm. (c) LST retrieved from bands 20, 22, 23, 29, 31, and 32 using the DL-NN algorithm. The white areas denote invalid values.

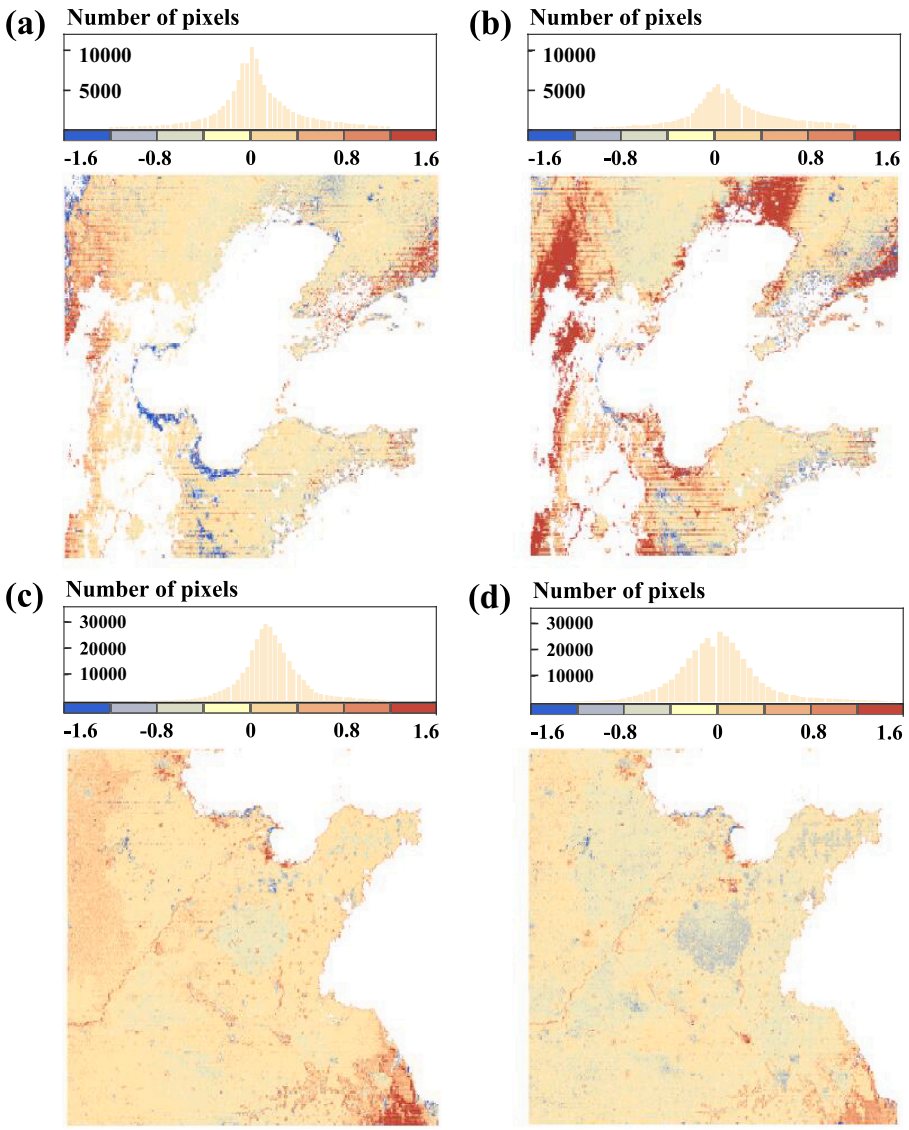

Fig. 17. Cross-validation. The difference (K) between MODIS LST products and the retrieved LSTs from MODIS data (retrieved LSTs minus LST products) using the DL-NN algorithm. (a) The difference between Fig. 15c and Fig. 15b. (b) The difference between Fig. 15d and Fig. 15b. (c) The difference between Fig. 16b and Fig. 16a. (d) The difference between Fig. 16c and Fig. 16a. The white areas denote invalid values.

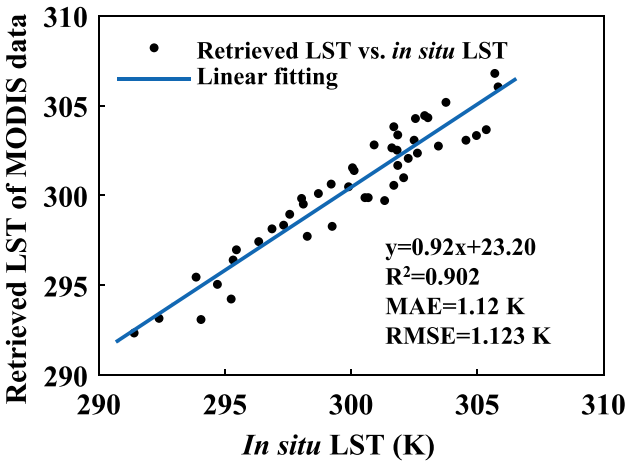

Fig. 18. *In situ* validation of MODIS data in practical analysis.

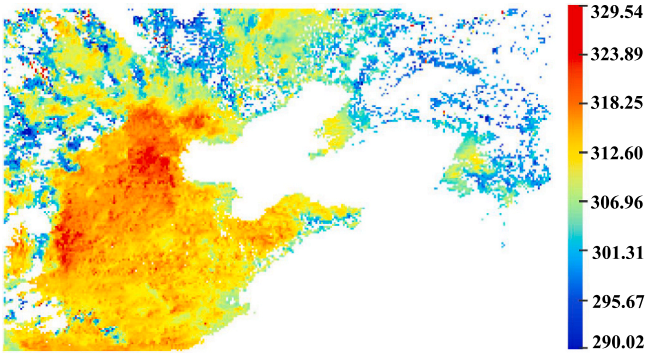

Fig. 19. Retrieved LSTs (K) from BT images of AGRI TIR bands 11, 12, 13, and 14 on July 1st, 2019, using the DL-NN algorithm; the white areas denote invalid values.

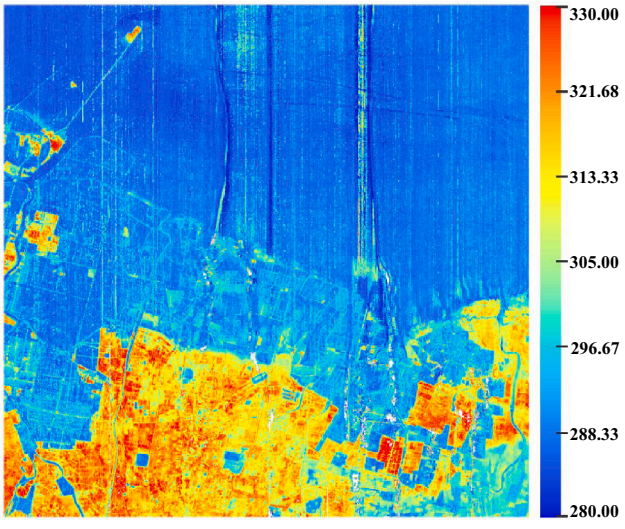

Fig. 20. LSTs (K) retrieved from VIMS TIR bands 9, 10, 11, and 12 of June 14, 2019, using the DL-NN algorithm; the white areas denote invalid values.

retrieval can be regarded as a classification problem. That is, LST retrieval is the same as high-dimensional classification. The traditional retrieval method classifies the ground features into different LST values according to the relevant radiation transmission information after determining each type of ground surface. If there is a method that can directly and accurately perform high-dimensional LST classification, the amplification error caused by step-by-step calculation in the conventional method can be overcome, and the target (LST) accuracy is judged to be the highest. Remote sensing images have problematic aspects with complex nonlinear characteristics and complementary, redundant spectral information. As a powerful machine learning methodology, deep learning can potentially extract more abstract and complex features at higher levels. This method has also proven to be efficient and powerful in extracting different ground surface types when processing hyperspectral and multispectral images (Geng et al., 2015). Mao et al. (2008) obtained good results, which concluded that neural networks are the best method for resolving the (LST and LSE) retrieval problem because these methods have function approximation, optimization computation and classification abilities.

In actual LST retrieval, for MODIS data, the best band combination is three TIR bands and atmospheric water vapor information (namely, MODIS TIR bands 29, 31, and 32 with WVC). For AGRI and VIMS data, the best band combination is four TIR bands (namely, AGRI TIR bands 11, 12, 13, and 14) and VIMS TIR bands (9, 10, 11, and 12), respectively. However, we also need to pay attention to the following two aspects. The first is the influence of TIR bands and atmospheric water vapor information. In general, most TIR bands in the atmospheric window are more stable than the atmospheric water vapor information, but instrument design is more difficult. For example, the TIR bands of the data from the new VIMS are unstable, and we can add the atmospheric water vapor information to improve retrieval accuracy and calculation efficiency. Therefore, an appropriate increase in the water vapor band is beneficial to improve the accuracy of LST retrieval. Second, the observation angle has an impact on LST retrieval. The biggest challenge of this method is that the retrieval accuracy decreases when the observation angle becomes larger and the WVC is high. To improve the retrieval accuracy, we can divide the observation angle into different intervals to retrain the DL-NN, and even divide it according to different seasons and regions.

5.2. Conclusions

To overcome the ill-conditioned problem of LST retrieval, a novel LST retrieval method is proposed based on model-data-knowledge-driven and deep learning, called the MDK-DL method. We discussed in detail the physical mechanism of parameter retrieval, the intrinsic relationship between geophysical parameters, and how to combine bands to achieve high retrieval accuracy. The analysis indicates that MDK-DL method consisting of model-data-knowledge-driven and deep learning is very suitable for resolving the ill-posed problem, which can use the interconnections between geophysical parameters.

Based on the validation using simulation data, for high-resolution TIR remote sensing data, the minimum MAE obtained from the optimal band combination is less than 0.1 K at a small viewing direction (<7.5°). For medium-resolution and low-resolution TIR remote sensing data at a large viewing direction (<65°), the minimum MAE is less than 0.8 K. The *in situ* validation shows that the minimum MAE obtained by the optimal band combination is approximately 1 K (RMSE = 1.123 K, and R^2 = 0.902). Additionally, the retrieved maps have abundant and credible spatial information. Increasing the number of TIR bands in the atmospheric window can improve the retrieval accuracy of the LST, while adding accurate atmospheric water vapor information obtains better results. In general, four TIR bands in the atmospheric window are sufficient to retrieve the LST with high precision. However, when atmospheric water vapor information and three TIR bands meeting the above conditions are available, adding additional TIR bands does not greatly improve retrieval accuracy.

In general, the MDK-DL method provides a new pattern, which is also a general framework to systematically retrieve LSTs and is suitable for complex mountainous areas and cities. We think that the use of deep learning as an optimized calculation requires sufficient and necessary

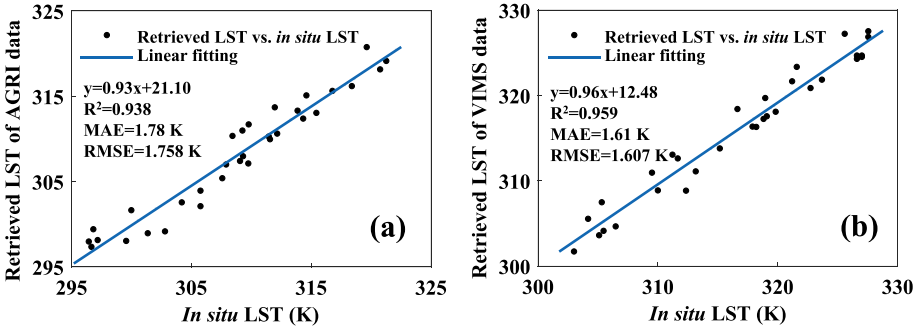

Fig. 21. *In situ* validation of (a) AGRI data and (b) VIMS data in practical analysis.

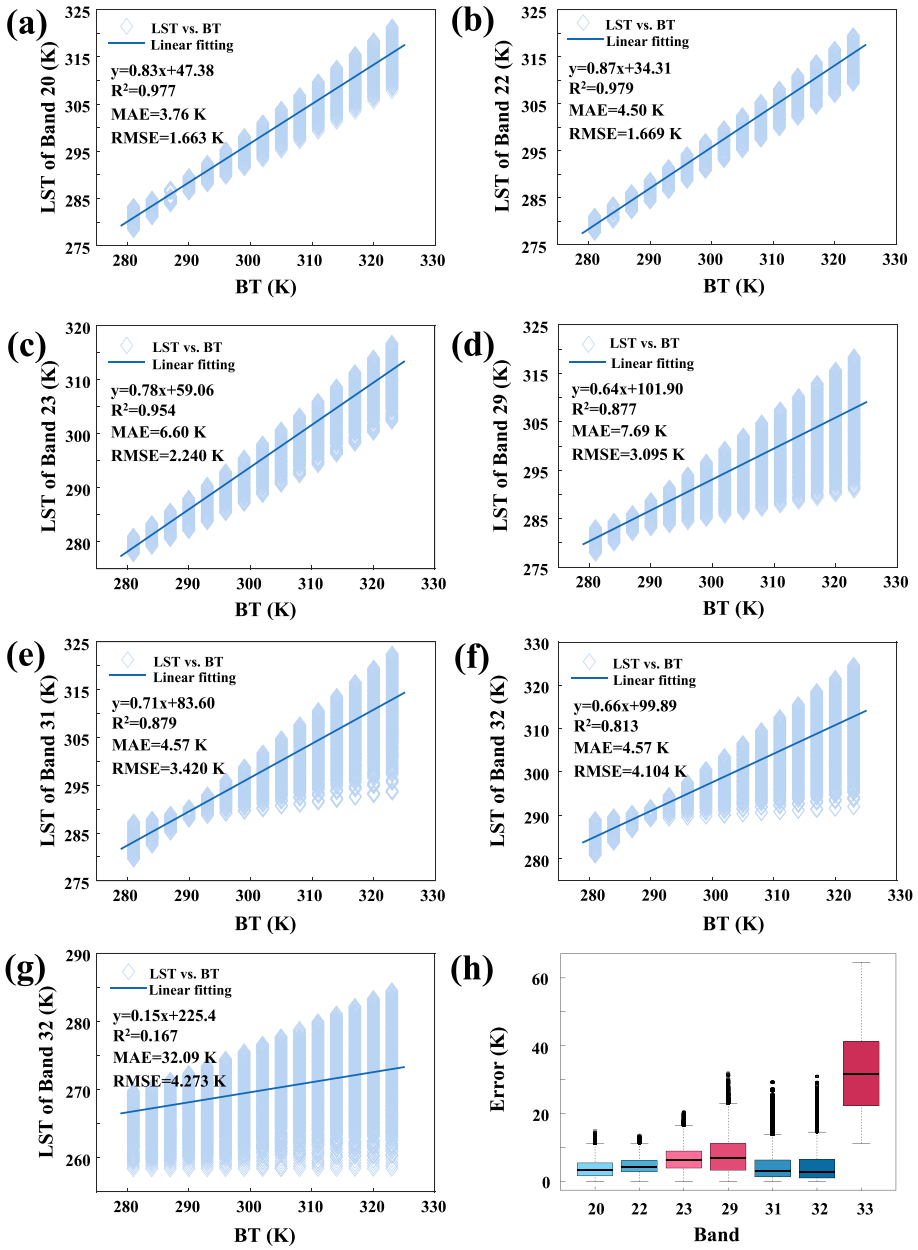

Fig. 22. The relationship between the LST and BT of (a) band 20, (b) band 22, (c) band 23, (d) band 29, (e) band 31, (f) band 32, and (g) band 33 at the satellite (in the case of MODIS data); (h) a boxplot of the absolute errors between the LSTs and BTs.

conditions. If we want to obtain high-precision results, the necessary condition is that the input parameters (nodes) must determine the output parameters (output nodes), and there is a causal relationship between input and output nodes. The sufficient condition is that the input parameters can theoretically construct enough equations and the output results can be calculated theoretically. After satisfying the necessary and sufficient conditions, the use of deep learning methods is no longer a black box theory, and it is an optimal computation. The training and test data is actually the solution of the theoretical equation. The use of deep learning is just to use the obtained representative solution to optimize the space curve of the simulation solution, so as to achieve the purpose of optimizing the solution equation. The MDK-DL method can help enrich the theoretical connotation of LST retrieval and make it suitable for interdisciplinary applications. In the future, we can train and test DL-NN by adding a large amount of reliable and high-precision training and test data to adapt to more situations.

Credit author statement

Han Wang: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing. **Kebiao Mao:** Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review, Editing, Project administration & Funding acquisition. **Zijin Yuan and Mengmeng Cao:** Software, Validation, Formal analysis, Investigation & Data Curation. **Jiancheng Shi and Zhihao Qin:** Resource, Formal analysis & Investigation. **Sibo Duan and Bohui Tang:** Resource & Editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank the China Meteorological Administration for providing the ground measurements, the China National Space Administration (CNSA) for providing the GF-5 data, the NASA Earth Observing System Data and Information System for providing the MODIS data, and the China National Satellite Meteorological Center (CNSMC) for providing the FY-4A data. This work was supported by the Second Tibetan Plateau Scientific Expedition and Research Program (STEP)-'Dynamic monitoring and simulation of water cycle in Asian water tower area (No. 2019QZKK0206), National Key Project of China (No. 2018YFC1506602), National Natural Science Foundation of China (No. 41921001), Fundamental Research Funds for Central Nonprofit Scientific Institution (No. 1610132020014), and the Open Fund of State Key Laboratory of Remote Sensing Science (Grant No. OFSLRSS201910).

References

Aires, F., Chédin, A., Scott, N.A., Rossow, W.B., 2002. A regularized neural net approach for retrieval of atmospheric and surface temperatures with the IASI instrument. J. Appl. Meteorol. 41, 144–159. https://doi.org/10.1175/1520-0450(2002)041<0144:ARNNAF>2.0.CO;2.

Blackwell, W.J., 2005. A neural-network technique for the retrieval of atmospheric temperature and moisture profiles from high spectral resolution sounding data. IEEE Trans. Geosci. Remote Sens. 43 (11), 2535–2546. https://doi.org/10.1109/TGRS.2005.855071.

Chen, Y., Lin, Z., Zhao, X., Wang, G., Gu, Y., 2014. Deep learning-based classification of hyperspectral data. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 7 (6), 2094–2107. https://doi.org/10.1109/JSTARS.2014.2329330.

Coll, C., Caselles, V., Valor, E., Niclòs, R., 2012. Comparison between different sources of atmospheric profiles for land surface temperature retrieval from single channel thermal infrared data. Remote Sens. Environ. 117, 199–210. https://doi.org/10.1016/J.RSE.2011.09.018.

Cristóbal, J., Jiménez-Muñoz, J.C., Sobrino, J.A., Ninyerola, M., Pons, X., 2009. Improvements in land surface temperature retrieval from the Landsat series thermal band using water vapor and air temperature. J. Geophys. Res. 114 https://doi.org/10.1029/2008JD010616.

Dozier, J., 1981. A method for satellite identification of surface temperature fields of subpixel resolution. Remote Sens. Environ. 11, 221–229. https://doi.org/10.1016/0034-4257(81)90021-3.

Frate, F.D., Solimini, D., 2004. On neural network algorithms for retrieving forest biomass from SAR data. IEEE Trans. Geosci. Remote Sens. 42 (1), 24–34. https://doi.org/10.1109/TGRS.2003.817220.

Geng, J., Fan, J., Wang, H., Ma, X., Li, B., Chen, F., 2015. High-resolution SAR image classification via deep convolutional autoencoders. IEEE Geosci. Remote Sens. Lett. 12 (11), 2351–2355. https://doi.org/10.1109/LGRS.2015.2478256.

Gillespie, A., Rokugawa, S., Matsunaga, T., Cothern, J.S., Hook, S., Kahle, A.B., 1998. A temperature and emissivity separation algorithm for advanced spaceborne thermal emission and reflection radiometer (ASTER) images. IEEE Trans. Geosci. Remote Sens. 36 (4), 1113–1126. https://doi.org/10.1109/36.700995.

Hinton, G.E., 2005. What kind of a graphical model is the brain. In: International Joint Conference on Artificial Intelligence (IJCAI), Edinburgh, Scotland, UK, pp. 1765–1775.

Hook, S.J., Gabell, A.R., Green, A.A., Kealy, P.S., 1992. A comparison of techniques for extracting emissivity information from thermal infrared data for geologic studies. Remote Sens. Environ. 42 (2), 123–135. https://doi.org/10.1016/0034-4257(92)90096-3.

Hu, W., Huang, Y., Wei, L., Zhang, F., Li, H., 2015. Deep convolutional neural networks for hyperspectral image classification. J. Sensors. 2015 (2015), 1–12. https://doi.org/10.1155/2015/258619.

Huang, Y., Chen, Z., Yu, T., Huang, X., Gu, X., 2018. Agricultural remote sensing big data: management and applications. J. Integr. Agric. 17 (9), 1915–1931. https://doi.org/10.1016/S2095-3119(17)61859-8.

Jang, J.D., Viau, A.A., Anctil, F., 2004. Neural network estimation of air temperatures from AVHRR data. Int. J. Remote Sens. 25 (21), 4541–4554. https://doi.org/10.1080/01431160310001657533.

Jedlovec, G.J., 1990. Precipitable water estimation from high-resolution split window radiance measurements. J. Appl. Meteorol. 29 (9), 863–877. https://doi.org/10.1175/1520-0450(1990)029<0863:PWEFHR>2.0.CO;2.

Jia, H., Yang, D., Deng, W., Wei, Q., Jiang, W., 2021. Predicting land surface temperature with geographically weighed regression and deep learning. Wires. Data Min. Knowl 11 (1). https://doi.org/10.1002/WIDM.1396.

Jiang, G.M., Li, Z.L., 2008. Split-window algorithm for land surface temperature estimation from MSG1-SEVIRI data. Int. J. Remote Sens. 29 (20), 6067–6074. https://doi.org/10.1080/01431160802235860.

Jiménez-Muñoz, J.C., Sobrino, J.A., 2003. A generalized single-channel method for retrieving land surface temperature from remote sensing data. J. Geophys. Res. 108 https://doi.org/10.1029/2003JD003480.

Jiménez-Muñoz, J.C., Cristobal, J., Sobrino, J.A., Soria, G., Ninyerola, M., Pons, X., 2009. Revision of the single-channel algorithm for land surface temperature retrieval from Landsat thermal-infrared data. IEEE Trans. Geosci. Remote Sens. 47 (1), 339–349. https://doi.org/10.1109/TGRS.2008.2007125.

Justice, C.O., Vermote, E., Townshend, J.R.G., Defries, R., Roy, D.P., Hall, D.K., Salomonson, V.V., Privette, J.L., Riggs, G., Strahler, A., Lucht, W., Myneni, R.B., Knyazikhin, Y., Running, S.W., Nemani, R.R., Wan, Z., Huete, A.R., Leeuwen, W.v., Wolfe, R.E., Giglio, L., Muller, J., Lewis, P., Barnsley, M.J., 1998. The Moderate resolution imaging Spectroradiometer (MODIS): land remote sensing for global change research. IEEE Trans. Geosci. Remote Sens. 36 (4), 1228–1249. https://doi.org/10.1109/36.701075.

Khoob, A.R., 2008. Artificial neural network estimation of reference evapotranspiration from pan evaporation in a semi-arid environment. Irrig. Sci. 27 (1), 35–39. https://doi.org/10.1007/S00271-008-0119-Y.

Kleespies, T.J., McMillin, L.M., 1990. Retrieval of precipitable water from observations in the split window over varying surface temperatures. J. Appl. Meteorol. 29 (9), 851–862. https://doi.org/10.1175/1520-0450(1990)029<0851:ROPWFO>2.0.CO;2.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature. 521 (7553), 436–444. https://doi.org/10.1038/NATURE14539.

Li, Z., Becker, F., 1993. Feasibility of land surface temperature and emissivity determination from AVHRR data. Remote Sens. Environ. 43 (1), 67–85. https://doi.org/10.1016/0034-4257(93)90065-6.

Liang, S., 2001. An optimization algorithm for separating land surface temperature and emissivity from multispectral thermal infrared imagery. IEEE Trans. Geosci. Remote Sens. 39 (2), 264–274. https://doi.org/10.1109/36.905234.

Mao, K., Qin, Z., Shi, J., 2005. A practical split-window algorithm for retrieving land surface temperature from MODIS data. Int. J. Remote Sens. 26, 3181. https://doi.org/10.1080/01431160500044713.

Mao, K., Shi, J., Li, Z., Tang, H., 2007. An RM-NN algorithm for retrieving land surface temperature and emissivity from EOS/MODIS data. J. Geophys. Res. 112 (D21102) https://doi.org/10.1029/2007JD008428.

Mao, K., Shi, J., Tang, H., Li, Z., Wang, X., Chen, K., 2008. A neural network technique for separating land surface emissivity and temperature from ASTER imagery. IEEE Trans. Geosci. Remote Sens. 46 (1), 200–208. https://doi.org/10.1109/TGRS.2007.907333.

Masuoka, E., Fleig, A., Wolfe, R.E., Patt, F., 1998. Key characteristics of MODIS data products. IEEE Trans. Geosci. Remote Sens. 36 (4), 1313–1323. https://doi.org/10.1109/36.701081.

McCulloch, W.S., Pitts, W., 1943. A logical calculus of the ideas immanent in nervous activity. Bull. Math. Biol. 5 (4), 115–133. https://doi.org/10.1007/BF02478259.

Mushkin, A., Balick, L.K., Gillespie, A.R., 2005. Extending surface temperature and emissivity retrieval to the mid-infrared (3-5 μm) using the multispectral thermal

imager (MTI). Remote Sens. Environ. 98 (2), 141–151. https://doi.org/10.1016/J.RSE.2005.06.003.
Peres, L.F., Dacamara, C.C., Trigo, I.F., Freitas, S.C., 2010. Synergistic use of the two-temperature and split-window methods for land-surface temperature retrieval. Int. J. Remote Sens. 31 (16), 4387–4409. https://doi.org/10.1080/01431160903260973.
Phan, T.N., Kappas, M., 2018. Application of MODIS land surface temperature data: a systematic literature review and analysis. J. Appl. Remote. Sens. 12 (4), 041501.
Price, J.C., 1983. Estimating surface temperatures from satellite thermal infrared data—a simple formulation for the atmospheric effect. Remote Sens. Environ. 13 (4), 353–361. https://doi.org/10.1016/0034-4257(83)90036-6.
Price, J.C., 1984. Land surface temperature measurements from the split window channels of the NOAA 7 advanced very high resolution radiometer. J. Geophys. Res. 89, 7231–7237. https://doi.org/10.1029/JD089ID05P07231.
Price, J.C., 1990. Using spatial context in satellite data to infer regional scale evapotranspiration. IEEE Trans. Geosci. Remote Sens. 28 (5), 940–948. https://doi.org/10.1109/36.58983.
Pu, R., Gong, P., Michishita, R., Sasagawa, T., 2006. Assessment of multi-resolution and multi-sensor data for urban surface temperature retrieval. Remote Sens. Environ. 104 (2), 211–225. https://doi.org/10.1016/J.RSE.2005.09.022.
Qin, Z., Dall'Olmo, G., Karnieli, A., Berliner, P., 2001a. Derivation of split window algorithm and its sensitivity analysis for retrieving land surface temperature from NOAA-advanced very high resolution radiometer data. J. Geophys. Res. 106, 22655–22670. https://doi.org/10.1029/2000JD900452.
Qin, Z., Karnieli, A., Berliner, P., 2001b. A mono-window algorithm for retrieving land surface temperature from Landsat TM data and its application to the Israel-Egypt border region. Int. J. Remote Sens. 22 (18), 3719–3746. https://doi.org/10.1080/01431160010006971.
Rosenblatt, F., 1958. The perceptron: a probabilistic model for information storage and organization in the brain. Psychol. Rev. 65, 386–408. https://doi.org/10.1037/h0042519.
Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning representations by back-propagating errors. Nature. 323 (6088), 533–536. https://doi.org/10.1038/323533a0.
Schmugge, T., French, A., Ritchie, J.C., Rango, A., Pelgrum, H., 2002a. Temperature and emissivity separation from multispectral thermal infrared observations. Remote Sens. Environ. 79 (2), 189–198. https://doi.org/10.1016/S0034-4257(01)00272-3.
Schmugge, T.J., Kustas, W.P., Ritchie, J.C., Jackson, T.J., Rango, A., 2002b. Remote sensing in hydrology. Adv. Water Resour. 25 (8), 1367–1385. https://doi.org/10.1016/S0309-1708(02)00065-9.
Shen, H., Jiang, Y., Li, T., Cheng, Q., Zeng, C., Zhang, L., 2020. Deep learning-based air temperature mapping by fusing remote sensing, station, simulation and socioeconomic data. Remote Sens. Environ. 240 https://doi.org/10.1016/J.RSE.2020.111692.
Sobrino, J., Coll, C., Caselles, V., 1991. Atmospheric correction for land surface temperature using NOAA-11 AVHRR channels 4 and 5. Remote Sens. Environ. 38 (1), 19–34. https://doi.org/10.1016/0034-4257(91)90069-I.
Sobrino, J.A., Jimenez-Munoz, J.C., El-Kharraz, J., Gomez, M., Romaguera, M., Soria, G., 2004. Single-channel and two-channel methods for land surface temperature retrieval from DAIS data and its application to the Barrax site. Int. J. Remote Sens. 25 (1), 215–230. https://doi.org/10.1080/0143116031000115210.
Tang, B., Li, Z., 2018. Estimation of land surface temperature from Chinese Gaofen-5 satellite data. In: IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Valencia, Spain, pp. 2559–2562.
Taylor, M., Kazadzis, S., Tsekeri, A., Gkikas, A., Amiridis, V., 2014. Satellite retrieval of aerosol microphysical and optical parameters using neural networks: a new methodology applied to the Sahara desert dust peak. Atmos. Meas. Tech. 7 (9), 3151–3175. https://doi.org/10.5194/AMT-7-3151-2014.
Tzeng, Y.C., Chen, K.S., Kao, W.L., Fung, A.K., 1994. A dynamic learning neural network for remote sensing applications. IEEE Trans. Geosci. Remote Sens. 32 (5), 1096–1102. https://doi.org/10.1109/36.312898.
Wan, Z., 2008. New refinements and validation of the MODIS land-surface temperature/emissivity products. Remote Sens. Environ. 112 (1), 59–74. https://doi.org/10.1016/J.RSE.2006.06.026.
Wan, Z., Dozier, J., 1989. Land-surface temperature measurement from space: physical principles and inverse modeling. IEEE Trans. Geosci. Remote Sens. 27 (3), 268–278. https://doi.org/10.1109/36.17668.
Wan, Z., Dozier, J., 1996. A generalized split-window algorithm for retrieving land-surface temperature from space. IEEE Trans. Geosci. Remote Sens. 34 (4), 892–905. https://doi.org/10.1109/36.508406.
Wan, Z., Li, Z., 1997. A physics-based algorithm for retrieving land-surface emissivity and temperature from EOS/MODIS data. In: IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Singapore, pp. 980–996. https://doi.org/10.1109/36.602541.
Wan, Z., Zhang, Y., Zhang, Q., Li, Z., 2002. Validation of the land-surface temperature products retrieved from Terra Moderate resolution imaging Spectroradiometer data. Remote Sens. Environ. 83 (1), 163–180. https://doi.org/10.1016/S0034-4257(02)00093-7.
Wan, Z., Zhang, Y., Zhang, Q., Li, Z.L., 2004. Quality assessment and validation of the MODIS global land surface temperature. Int. J. Remote Sens. 25 (1), 261–274. https://doi.org/10.1080/0143116031000116417.
Wu, P., Yin, Z., Yang, H., Wu, Y., Ma, X., 2019. Reconstructing geostationary satellite land surface temperature imagery based on a multiscale feature connected convolutional neural network. Remote Sens. 11 (3) https://doi.org/10.3390/RS11030300.
Yang, G., Pu, R., Huang, W., Wang, J., Zhao, C., 2010. A novel method to estimate subpixel temperature by fusing solar-reflective and thermal-infrared remote-sensing data with an artificial neural network. IEEE Trans. Geosci. Remote Sens. 48 (4), 2170–2178. https://doi.org/10.1109/TGRS.2009.2033180.
Yang, J., Zhang, Z., Wei, C., Lu, F., Guo, Q., 2017. Introducing the new generation of Chinese geostationary weather satellites, Fengyun-4. Bull. Am. Meteorol. Soc. 98 (8), 1637–1658. https://doi.org/10.1175/BAMS-D-16-0065.1.
Yuan, Q., Shen, H., Li, T., Li, Z., Li, S., Jiang, Y., Xu, H., Tan, W., Yang, Q., Wang, J., Gao, J., Zhang, L., 2020. Deep learning in environmental remote sensing: achievements and challenges. Remote Sens. Environ. 241 https://doi.org/10.1016/J.RSE.2020.111716.
Zhang, L., Zhang, L., Du, B., 2016. Deep learning for remote sensing data: a technical tutorial on the state of the art. IEEE Geosci. Remote. Sens. Mag. 4 (2), 22–40. https://doi.org/10.1109/MGRS.2016.2540798.
Zhang, Y., Jiang, X., Wu, H., 2017. A generalized split-window algorithm for retrieving land surface temperature from GF-5 thermal infrared data. In: IEEE 2017 Progress in Electromagnetics Research Symposium - Spring (PIERS), St Petersburg, Russia, pp. 2766–2771. https://doi.org/10.1109/PIERS.2017.8262224.
Zhang, X., Zhang, Q., Zhang, G., Nie, Z., Gui, Z., Que, H., 2018. A novel hybrid data-driven model for daily land surface temperature forecasting using long short-term memory neural network based on ensemble empirical mode decomposition. Int. J. Environ. Res. Public Health 15, 1032. https://doi.org/10.3390/ijerph15051032.
Zhao, W., Du, S., 2016. Learning multiscale and deep representations for classifying remotely sensed imagery. ISPRS J. Photogramm. Remote Sens. 113, 155–165. https://doi.org/10.1016/J.ISPRSJPRS.2016.01.004.
Zhou, Z., 2019. Abductive learning: towards bridging machine learning and logical reasoning. SCIENCE CHINA Inf. Sci. 62 (07), 220–222. https://doi.org/10.1007/s11432-018-9801-4.
Zhu, X.X., Tuia, D., Mou, L., Xia, G., Zhang, L., Xu, F., Fraundorfer, F., 2017. Deep learning in remote sensing: a comprehensive review and list of resources. IEEE Geosci. Remote Sens. Mag. 5 (4), 8–36. https://doi.org/10.1109/MGRS.2017.2762307.
Zuo, Z., Shuai, B., Wang, G., Liu, X., Wang, X., Wang, B., Chen, Y., 2016. Learning contextual dependence with convolutional hierarchical recurrent neural networks. IEEE Trans. Image Process. 25 (7), 2983–2996. https://doi.org/10.1109/TIP.2016.2548241.