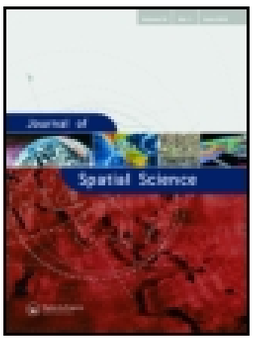


Journal of Spatial Science

ISSN: 1449-8596 (Print) 1836-5655 (Online) Journal homepage: https://www.tandfonline.com/loi/tjss20

Application of unsupervised learning of finite mixture models in ASTER VNIR data-driven land use classification

Bo Zhao, Fan Yang, Rongzhen Zhang, Junping Shen, Jürgen Pilz & Dehui Zhang

To cite this article: Bo Zhao, Fan Yang, Rongzhen Zhang, Junping Shen, Jürgen Pilz & Dehui Zhang (2019): Application of unsupervised learning of finite mixture models in ASTER VNIR data-driven land use classification, Journal of Spatial Science, DOI: 10.1080/14498596.2019.1570478

To link to this article: https://doi.org/10.1080/14498596.2019.1570478

Published online: 01 Mar 2019.

JOURNAL OF SPATIAL SCIENCE
https://doi.org/10.1080/14498596.2019.1570478

Taylor & Francis
Taylor & Francis Group

Application of unsupervised learning of finite mixture models in ASTER VNIR data-driven land use classification

Bo Zhao[a], Fan Yang[b,c,d], Rongzhen Zhang[e], Junping Shen[a], Jürgen Pilz[f] and Dehui Zhang[g]

[a]Advanced Algorithm Research Division, Beijing PIESAT Information Technology Co., Ltd, Beijing, China; [b]Beijing Institute of Geology for Mineral Resources, Beijing, China; [c]Key Laboratory of Geochemical Cycling of Carbon and Mercury in the Earth's Critical Zone, Institute of Geophysical and Geochemical Exploration, Chinese Academy of Geological Sciences, Langfang, China; [d]State Key Laboratory of Ore Deposit Geochemistry, Institute of Geochemistry, Chinese Academy of Sciences, Guiyang, China; [e]Henan Institute of Geological Survey, Zhengzhou, China; [f]Institute of Statistics, Alpen-Adria-Universität Klagenfurt, Klagenfurt, Austria; [g]School of Earth Sciences and Resources, China University of Geosciences, Beijing, China

ABSTRACT

Based on an ASTER VNIR image, we studied the applicability of the MML-EM (Minimum Message Length Criterion-Expectation Maximization) algorithm for land-use classification in southern Austria. Firstly, the RVI (ratio vegetation index) and PC_1 (first principal component) bands have been utilized to enhance the targeted information; secondly, the MML-EM algorithm and the terrain analysis-based imagery clipping were jointly used for surface type discrimination. Findings showed that the MML-EM method can provide refined imagery classification results and this is the first time it has been applied in this realm.

KEYWORDS

Mixture Gaussian distribution; land use; topographical analysis; remote sensing

1. Introduction

Remotely sensed (RS) data have been increasingly applied to characterize features on the Earth's surface for different purposes. Image classification is the most common analysis of multispectral remote sensing data in order to produce thematic maps that provide representation of the spatial distribution of a particular theme (Schowengerdt 2006). In this article we will focus on land-cover classification. Its main task is to classify the pixel-values into components, such as forest, urban, agriculture and other classes, and to establish the relationships between them. Unsupervised and supervised image classification techniques are the two most common methods in use (Camps-Valls *et al.* 2011), including supervised maximum likelihood, neural net, support vector machine (SVM), etc., and unsupervised K-means, ISODATA (iterative self-organizing data), etc. (Comber *et al.* 2012); for high- and ultrahigh-resolution data, the literature clearly points to object-based image analysis methods (Geneletti and Gorte 2003). The emphasis in this study, rather, is on the statistical rules behind multispectral image classification, and to survey how to model the joint histogram by a mixture of 2D Gaussians, corresponding to different land

CONTACT Fan Yang yangfan@igge.cn

cover classes. According to Sridharan and Qiu (2013), the probability distributions of image object-level reflectance values are often not normal, but multi-peaked due to the physical structure of the object being analysed and the lighting condition at image-acquisition time. This implies that some summary statistics (particularly the commonly used 'mean' and 'standard deviation' values) can be a misrepresentation of the spectral reflectance pattern of an object. Thus, the relevant experiments have been done only in the pixel-oriented framework.

Thresholding is a fundamental task in image classification. The grey-level values below or equal to the selected threshold in an enhanced image are usually classified as a class or a real-world feature, and the values above this threshold are classified as another class. The grey-level histogram provides readers with an appreciation of the quality of an RS scene, i.e. whether it is low/high in contrast, or multi-modal in nature (Swarnajyoti *et al*. 2011). This, in a sense, can be the theoretical basis for land-use classification, and what really matters is setting appropriate thresholds for the histogram using appropriate statistical analysis methods. For example, Gaussian distribution-based statistical methods, assuming that the brightness values recorded in a given scene are normally distributed (sometimes outliers cause the data to become skewed), especially when the study area is extensive (Liu *et al*. 2017), have been widely employed in most GIS packages for assigning thresholds for land classification – the maximum likelihood method is an example. Panahi and Cheng (2004) noticed that none of the thresholding methods under normality assumptions has taken into account the spatial properties of real-world features, and spatial fractal algorithms, characterizing not only pixel-value frequency distributions, but also the geometrical or spatial properties of the features reflected by pixel zones in the image, were highly recommended. Fractal theory has been successfully used to characterize spatial patterns defined on the basis of RS data, but acting merely as a thresholding proxy it cannot provide more information about the content of an image (Sun *et al*. 2006).

A threshold is a criterion that allows us to classify an image. Ju *et al*. (2003) and the references therein believed that pixel-values reflecting multiple land cover types in an enhanced image can be modelled as a mixture of subclasses, with each of them represented by an (approximately) Gaussian/normal distribution, and naturally the thresholds amongst them are used for image classification (Tobias and Seara 2002). This is because a single Gaussian distribution at regional scale may not represent various types of land cover well, and thus it is more realistic to treat the image as a mixture of Gaussian distributions (Kurita and Abdelmalek 1992, Richards and Richards 1999). It is well known that when the outcome is produced by many small effects acting additively and independently, its distribution will be normal or close to normal. This is supposed to be the case with a physically or spectrally homogeneous object/feature in an image, and multiple objects usually produce a multimode distribution, sometimes seeming like a skewed distribution, in which a single population (mass) may partially overlap with another because of the spectral similarities (Zhao *et al*. 2016). The implication of the Gaussian mixture model is actually beyond the thresholding issue, and it can be linked to the scale (weight), statistical independence (separability), spectral sensitivity (the mean value of a distribution), inhomogeneity (standard deviation), the source of false anomalies and the probabilistic pattern of a ground feature (Zhao *et al*. 1994), thus providing an unbiased approach to knowledge discovery during image classification.

Fortunately, accurate separation of the mixed distributions has been made possible with the help of the Minimum Message Length Criterion-Expectation Maximization (MML-EM) algorithm, namely the unsupervised learning of a finite mixture model (Figueiredo and Jain 2002), which is believed to among the state-of-the-art algorithms for separating originally mixed Gaussian distributions on the histogram (Liu *et al*. 2011). It can provide a holistic insight into the probability distribution patterns of the original RS data, while it avoids several drawbacks of the classic EM (Expectation Maximization) algorithm and can learn finite mixtures from data by selecting the number of components in an unsupervised way (Biernacki *et al*. 2000). However, MML-EM is still unfamiliar to most imagery interpreters, and there is no report validating its availability in RS products. Around the land-use issue, what we are most concerned about is whether MML-EM can be applied to separate various features by screening the mixed distributions of enhanced RS data. This is a question because the spectral similarities tend to make the probability histogram featureless (Han *et al*. 2017a). Using ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer) multispectral data and taking the Klagenfurt district of Austria as a case study, the objective of this research is to map, for the first time in this kind of study, the present land-use status by introducing the MML-EM model and analysis. The rangeland/agricultural land, urban or built-up land and water are our primary concerns.

2. Material and preprocessing

2.1 Overview of the study area

The land use status constitutes a major global change driver (Guo and Gifford 2010). The main body of the study area (Figure 1(a)) is located in the southern part of the Eastern Alps (Nemes *et al*. 1997), the Klagenfurt district (State of Kärnten), southern Austria. Its spatial range is 46°23′N – 46°40′N and 14°02′E – 14°42′E. According to the 30 m resolution GDEM (Global Digital Elevation Model) data (downloaded from http://www.gscloud.cn), this area covers over 1500 km^2 of land with an elevation of 362 to 2231 m (810.39 m on average) and slope of 0.00° to 83.30° (25.00° on average). As is seen in Figure 1(b), this terrain is heavily forested, and it represents the heterogeneous characteristics of land use status in Central Europe, which consists of free-standing waters, residential areas, rangelands, bare rocks, snow-capped mountains, shrub lands, wetlands, forest ecosystems, as well as a variety of agricultural systems (such as fallow land, irrigated or rain-fed field crops, cornfields, and orchards).

2.2 Data source

The ASTER data (Level 1) consist of three separate subsystems with a total of 14 spectral bands (0.52–11.65 μm). Only the visible-near infrared (VNIR) subsystem containing three bands (B1 (Green), B2 (Red) and B3N (NIR), from 0.52 to 0.86 μm) with 15 m spatial resolution was involved in this research. The ASTER data (Path/Row: 191/28) were downloaded from http://glovis.usgs.gov/ without cost, and the acquisition date was 24 August 2011, the sun azimuth 154.769566°, the sun elevation 52.476347° and cloud coverage <1.0%. Moreover, the matched ASTER GDEM (Global Digital Elevation Model)

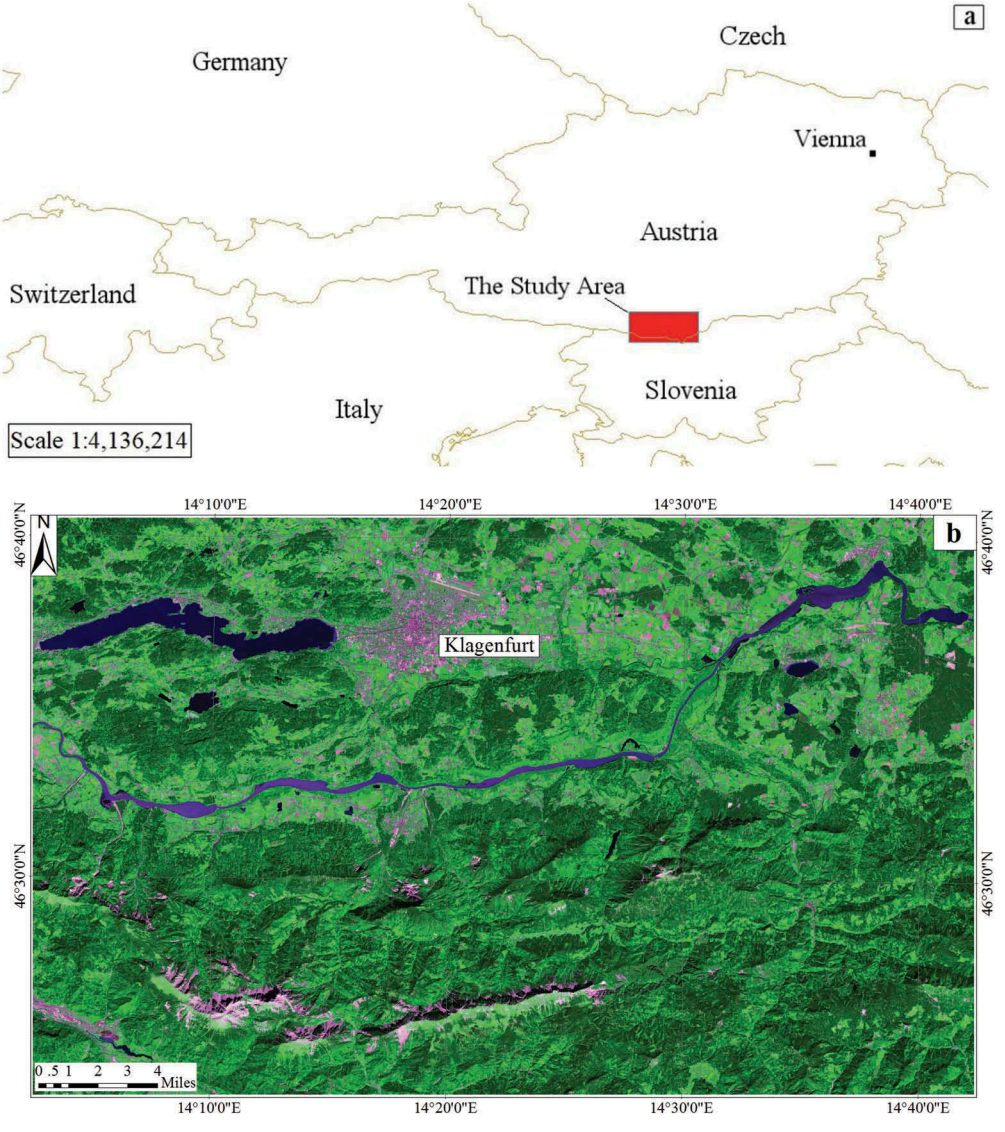


Figure 1. The location of the study area on the topographic map (a) of Central Europe and on the ASTER image (b) – bands 2, 3N and 1 in RGB.

data (30 m resolution) were acquired freely from www.gscloud.cn/, the relevant Google Earth images were downloaded using the 'Locaspace Viewer' software and site-specific ground validation was also conducted, using a high-definition camera and a portable GPS receiver with 5 m accuracy.

People may question whether the three bands of ASTER are sufficient for classification. This will be shown by the fact that even three bands would do a decent job of classification if the spectral information is fully utilized. In addition, high-resolution satellite/aerial digital imagery is typically used for identification of ground objects, but

the restriction of its bands to the visible wavelengths often limits the relevant application. Thus, our study may have some reference significance in this regard as well.

2.3 Image preprocessing

Based on the PIE v4.0 (Pixel Information Expert, a professional software product for RS image processing developed by Beijing PIESAT Information Technology Co., Ltd) and ENVI v5.3 software, pre-processing of the above data involves geometric rectification, image cropping, reprojection and orthorectification. The images were georeferenced to a UTM zone 33 (North Hemisphere) projection using the WGS-84 datum. At the suggestion of Lin *et al*. (2015), atmospheric correction is dispensable for land use classification, especially when no standard spectra were involved in the classification process, and may produce no improvement in (unsupervised) classification accuracy. However, in this article atmospheric correction was still executed because vegetation is preferentially more susceptible to the atmosphere and is a background feature that we cannot avoid (Lillesand *et al*. 2014).

3. Analysis methods

3.1 Image enhancement

The ASTER image improves the ability to identify various land cover types with remotely sensed data, but extracting band combinations or band operations that give the best results for image segmentation in most cases is difficult. There are only three bands in the first subsystem of ASTER data, and intuitively the exploitable spectral information for classification should be band 3N (NIR, near-infrared). This band is of diagnostic significance for many ground objects: e.g. in comparison to a tree object (including the grass), the peak values of a building object are often located at the lower end of the reflectance spectrum, which is understandable as buildings reflect less in the NIR region than trees. A similar scenario may apply to water features.

The real probability distribution pattern of an object cannot be approached if there are more pixel values badly distorted by terrain-induced lighting intensity differences. Here the Ratio Vegetation Index (RVI), namely the band ratio 3N/2 or NIR/R, was employed to compensate for variations caused by the topographic shading. Findings show that band 3N has the maximum standard deviation (i.e. information content), 2,985,484.47, and the coefficient of correlation (i.e. information redundancy) between band 3N and 2 is at a minimum, 0.305, and their ratio (RVI) was successfully used in identifying the vegetated areas and different reflectance levels they create. However, it may not perform well for areas with <50% vegetation cover (Huete *et al*. 2000).

The PCT (principal component transformed) image was also considered here. It is assumed that the first component of a PCT image (PC_1), usually having the highest factor loading of each input component, is connected to the textural, brightness or geomorphological aspects, and these can capture the essential qualities of a feature (Mars and Rowan 2010). PC_2 and PC_3 were not taken into account because their histograms follow a monomodal normal distribution. In addition, it must be cautioned that sometimes land

cover classes with very dissimilar colour but similar brightness in the PC bands might not be well separated.

3.2 Multimodal distribution and its applicability

The probabilistic mixture model assumes that all the data points are generated from a mixed distribution, where the confidence that a data point belongs to a certain cluster has a strict probability meaning (Jiang *et al*. 2010). Vistelius and Andrew (1960) first proposed the 'fundamental law of the geochemical processes' describing the probability distributions of concentrations of chemical elements in the earth's crust: distributions of the fixed stage of a geochemical process are normal, and distributions of mixed products of several stages are (positively) skewed, forming a joint distribution in many cases (Zhang *et al*. 2014). The RS data, as consistent with the geochemical data, are structured in *X-Y-Z* format, where *X-Y* sets the spatial position (longitude/latitude) of a pixel and *Z* sets the reflectance value (or DN values), so a similar scenario applies to RS image processing (Zhao *et al*. 2016). Commonly, the histogram of one feature's DN values is unimodal and may be Gaussian, although skewing is likely (Kerroum *et al*. 2010). For the mixture model, multivariate Gaussian distribution is most commonly used in statistical analysis due to its complete theory and analytical tractability, and it is also commonly used in image classification and object detection because of the Law of Large Numbers on which the characteristics of RS data are based (Xu and Wunsh 2005).

Han *et al*. (2017b) posited that the enhanced greyscale images help in formulating the thresholding problem as a binary classification problem with a mixture of Gaussian distributions, if an image contains two or more dominant ground features with distinctly different (often narrow) ranges of reflectance. As seen in am), a multimodal model usually gives a clear separation of various ground features, but two adjacent populations can be strongly overlapped in most cases, meanwhile violating the bell-shaped normal curve of a population (class). Figure 2(c) is such an example, showing a common situation in the remote sensing field when trying to classify classes very similar to each other. Muñoz-Marí *et al*. (2007) further reported that in some extreme cases a mixture of two or more unimodal distributions with differing means is not necessarily multi-peaked, and Figure 2(d) is an example. Obviously, a mixture of overlapped Gaussian populations is a difficult problem for histogram thresholding and analysis, and the MML-EM model and its extensions may provide an efficient solution.

3.3 The MML-EM based unsupervised classification

According to Liu *et al*. (2011) and references therein, the general expression of a mixed distribution is assumed to be as follows:

$$p(x;\theta)=\sum_{i=1}^{k} \alpha_i \cdot f(x,\theta_i)$$

where k is the number of the mixture populations; $\theta = (k, \theta_1, \ldots, \theta_k, \alpha_1, \ldots, \alpha_k)$, θ_i is the statistical parameters weight (α), mean (m) and covariance (Cov), α_i is the weight (percentage) of a population i, and $\sum_{i=1}^{k} \alpha_i = 1$ and $\alpha_i > 0$; $f(x, \theta_i)$ is the probability density function of population i. In brief, the MML-EM algorithm was put forward to

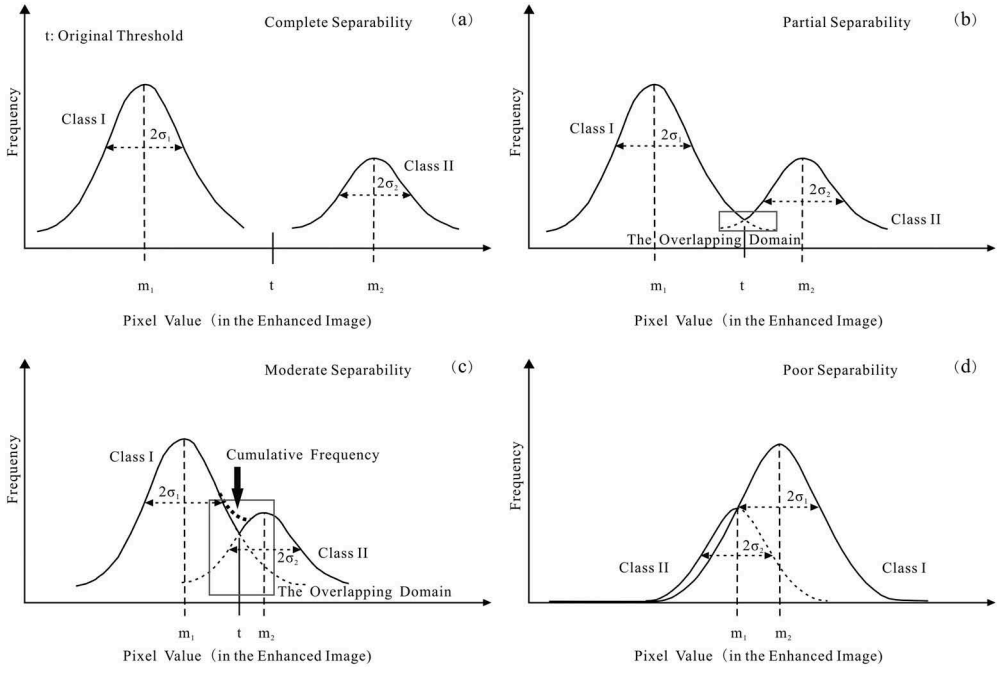


Figure 2. (a) to (d) are the mixture Gaussian schemas (bimodal) demonstrating the concept of separating populations from background (modified after Sinclair (1991)), and *t* represents the original threshold between two neighbouring populations.

estimate k and θ_i in a mixed Gaussian distribution, and it is based on 'a MML-like criterion which is directly implemented by a modified EM algorithm'. We have no intention of giving the detailed algorithm process here, and readers are referred to Figueiredo and Jain (2002). The related computer program is available from their open MATLAB code.

EM is a widely known iterative algorithm for estimating the maximum probability for a set of parameters of a statistical model. As its name indicates, the EM algorithm produces a sequence of estimates $\{\theta_i\}$ by alternatingly applying two steps: computing the conditional expectation of the complete log-likelihood (the expectation (E) step) and updating the parameter estimates according to the maximization (M) step; EM iterates between the 'E' step and the 'M' step until some convergence criterion is satisfied (see details in Zhang 2004). MML gives the optimal model selection criterion to choose one among a set of candidate modes. Considering a c-dimensional dataset y, known to have been generated according to $p\ (y|\theta)$, MML leads to the following criterion (where the minimization with respect to θ can be understood as simultaneously in θ and c, namely the dimension of θ):

$$\hat{\theta} = \arg\min_{\theta}\left\{-\log p(\theta) - \log \mathrm{p}(y|\theta) + \frac{1}{2}\log|I(\theta)| + \frac{c}{2}\left(1 + \log\frac{1}{12}\right)\right\}$$

where $\mathrm{I}(\theta) \equiv -\mathrm{E}[\mathrm{D}_{\theta}^{2}\log \mathrm{p}(\mathrm{y}|\theta)]$ is the expected Fisher information matrix and $|\mathrm{I}(\theta)|$ denotes its determinant (Christakos *et al*. 2012); $\hat{\theta}$ is a finite precision version of θ.

Putting all the elements together, MML-EM was designed to seamlessly integrate estimation and model selection in a single algorithm. It avoids two major drawbacks of classic EM: sensitivity to initialization and possible convergence to the boundary of the parameter space. Many experimental results have indicated that it has good performance in learning mixtures of Gaussian and mixture of factor analysers. This approach was successfully applied for processing of ore geochemical data (Liu *et al*. 2011), even when there is no mathematical or physical reason for the original data to have a mixed-distribution pattern (Tobias and Seara 2002). The implementation of MML-EM is under the assumption that pixel values in the input image follow a mixture normal distribution; notwithstanding, it can process and yield new insights for 'non-Gaussian' data or arbitrarily complex probability density functions, which is a common and possible scenario in the histogram.

3.4 Extensions of the MML-EM model

People may doubt the merit of thresholding the one-dimensional histogram consisting of highly overlapping Gaussian distributions (Figure 2), because no matter how one defines the thresholds, there are always false-positive or false-negative errors (Richards and Richards 1999). It is beneficial to understand the histograms in a multidimensional space formed by the three spectral bands and relevant spectral indices, and there are existing classification methods (e.g. the support vector machine) that are designed to tackle this problem. These, however, are beyond the scope of the current study.

Here we will introduce a necessary extension of MML-EM. As displayed in Figure 3, supposing there are several populations interfering with each other in the histogram of enhanced imagery, we must define an appropriate decision threshold, so that the grey-level values below or equal to the selected threshold can be classified as one ground object, while the values above it can be classified as another. Nussbaum *et al*. (2006) formulized the thresholding problem as a binary classification problem with a mixture of Gaussian distributions, and the computational formula is shown below:

$$t = \frac{m_2\sigma_1^2 - m_1\sigma_2^2 + \sigma_1\sigma_2\sqrt{(m_1 - m_2)^2 + 2A(\sigma_1^2 - \sigma_1^2)}}{(\sigma_1^2 - \sigma_1^2)} \quad \text{and } A = \log\left(\frac{\sigma_1}{\sigma_2} \times \frac{m_2}{m_1}\right)$$

where m_1 and m_2 are the mean values of two adjacent populations (classes), respectively, and $m_1 < m_2$; σ_1 and σ_2 are their standard deviations. These parameters are by-products of the MML-EM analysis, for the (unimodal) normal probability-density curve in Figures 2 and 3 cannot be separated and diagrammed without them.

However, there will be some misclassifications when using t for classification, because t is located at or close to the midpoint of the overlapping domain, namely the inflection point where two neighbouring probability curves intersect (Gao *et al*. 2011). In order to avoid the overlapping domain shown in Figure 2(b–d), the adjustment threshold (T) is highly desired, while t is regarded as the original threshold. The determination of T, according to Nussbaum and coworkers, is only possible with the help of another parameter – separability (J) – and its formula is given below:

$$J = 2\left(1 - e^{-B}\right), \text{ where } J \in [0,\ 2], \text{and}$$

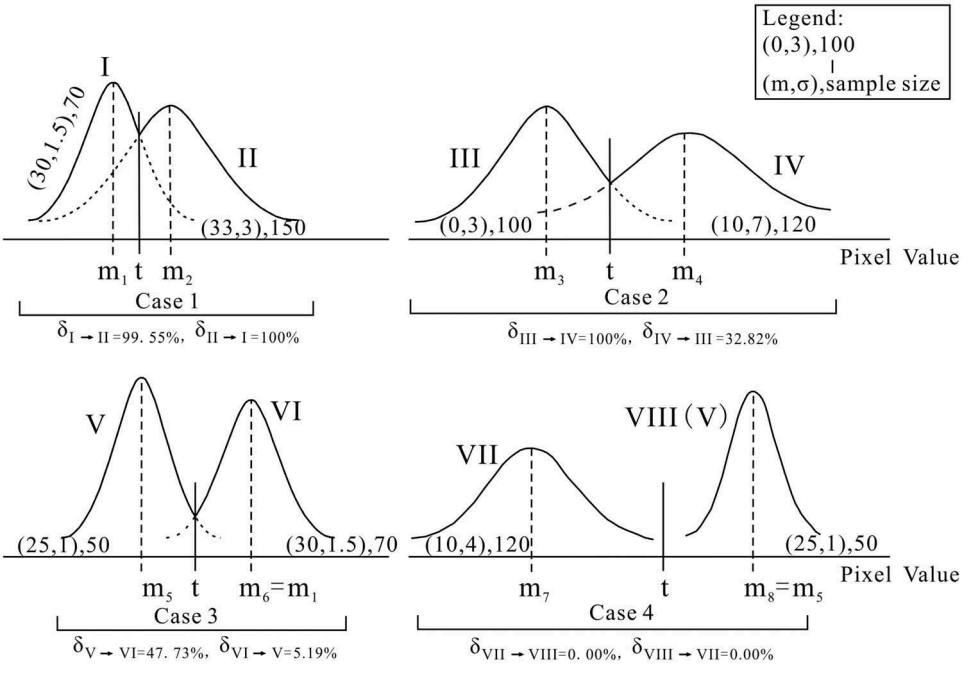


Figure 3. Different cases for the pairwise separability (*J*) measure. Here, I and II, III and IV, V and VI … represent any two neighbouring populations that are approximately normally distributed; see the text for explanations of other abbreviations.

$$B = \frac{1}{8}(m_1 - m_2)^2 \frac{2}{\sigma_1^2 + \sigma_1^2} + \frac{1}{2}\ln\left[\frac{\sigma_1^2 + \sigma_1^2}{2\sigma_1\sigma_2}\right]$$

We can empirically specify that: when $J < 0.5$, namely Case 1 shown in Figure 3, there is no available decision threshold; when $0.5 \leq J < 1.25$, namely Case 2 in Figure 3, the adjustment threshold $T = m_2$ (if the target is located in bright areas) or m_1 (if in dark areas), indicating a moderate separability; when $1.25 \leq J < 1.75$, namely Case 3, $T = (t + m_2)/2$ (bright) or $m_1 + (t - m_1)/2$ (dark), implying a good pairwise separability; and when $J \geq 1.75$, namely Case 4, completely separated, $T = t$ (Wang *et al.* 2013). As in a normal distribution, 99.70% of the samples will fall within $\pm 3\sigma$ of the mean (Gao *et al.* 2011), so the concept of the overlapping rate (δ) between any two neighbouring populations (noted by I and II) can be derived in the following:

$$\delta_{I \to II} = [(m_1 + 3\sigma_1) - (m_2 - 3\sigma_2)]/[(m_1 + 3\sigma_1) - (m_1 - 3\sigma_1)]$$

$$\delta_{II \to I} = [(m_1 + 3\sigma_1) - (m_2 - 3\sigma_2)]/[(m_2 + 3\sigma_2) - (m_2 - 3\sigma_2)]$$

Apparently, δ quantifies the proportion of how many samples in I or II fall into the overlapping domain where the main confusion occurs (Bazi and Melgani 2010); it also provides a deep insight into the distribution of RS data. For example, when two distributions vary both in mean and variance, $\delta_{I \to II}$ (the overlapping rate from population I to II) and $\delta_{II \to I}$ (the rate from II to I) can be very different, as seen in row 12 of Table 1, Population I is almost completely submerged in II, while only 30% of the elements in II fall

into the overlapping domain, and $\delta_{I\to II} >> \delta_{II\to I}$. The simulation experiments of Table 1 and Figure 3 also give some examples of how to understand the pairwise separability measure (J): keeping the standard deviation (σ) unchanged, and with increasing mean (m) value, the separability (J) keeps on increasing while the overlapping rate (δ) keeps on decreasing; on the other hand, keeping m unchanged, and with increasing σ, the separability generally decreases while δ increases.

3.5 The post-classification process

It is not possible to eliminate false anomalies in terms of spectrum alone, and in order to improve the correct interpretation ratio, here the digital terrain feature that is independent of the spectral nature was investigated. From the topographical perspective, water, built-up land, agricultural lands and so on are usually located on flat terrain, while forests and large areas of rangeland are found to be present mainly on hilly terrain. During the post-classification process, it is necessary to conduct a GDEM-based topographical analysis, so that those level terrains with slopes below 15° and relief amplitudes less than 3 m can be retained or removed, and correspondingly the rugged terrains be removed or retained. Note that the topographical feature map is a by-product easily extractable from GDEM data based on PIE v4.0, and the relevant thresholds (15° and 3 m) were determined experientially.

In addition, the 'sieve' module in ENVI v5.3 can remove the isolated pixels (random noise) based on a size threshold (kernel size), and 'clump' is run to add spatial coherence to existing anomaly pixels by combining adjacent similarly classified areas. As the case may be, these two methods can be used either jointly or separately.

Table 1. A numerical simulation experiment for understanding the distribution of RS data.

(m, σ), N	J	t	T	$\delta_{I\to II}$	$\delta_{II\to I}$
(0, 1),100 vs. (1,1), 100	0.00	None	None	100%*	99.50%*
(0,1), 100 vs. (2,1), 100	0.00	None	None	99.50%*	99.50%*
(0,1), 100 vs. (3,1), 100	1.3798	1.5373	2.2624	40.21%	56.73%
(0,1), 100 vs. (5,1), 100	1.8759	2.5352	2.5352	3.09%	2.88%
(0,1), 100 vs. (7,1), 100	1.9932	4.0042	4.0042	0.00%	0.00%
(0, 1),100 vs. (5,1.2), 100	1.8256	2.3157	2.3157	7.29%	7.62%
(0,1), 100 vs. (5,1.4), 100	1.6773	2.0549	3.4650	35.79%	16.98%
(0,1), 100 vs. (5,1.6), 100	1.5991	2.2205	3.5534	43.00%	13.86%
(0,1), 100 vs. (5,1.8), 100	1.6235	1.9467	3.3918	47.31%	8.33%
(0,1), 100 vs. (5,2.0), 100	1.5480	2.0480	3.5210	64.95%	15.39%
Case 1 in Figure 3 (30, 1.5), 70 vs. (33, 3), 150	0.00	None	None	99.55%*	100%*
Case 2 in Figure 3 (0, 3), 100 vs. (10, 7), 120	0.9925	3.1931	9.4660	100%*	32.82%
Case 3 in Figure 3 (25, 1), 50 vs. (30, 1.5), 70	1.6575	26.3601	27.9576	47.73%	5.19%
Case 4 in Figure 3 (10, 4), 120 vs. (25, 1), 50	1.9716	22.0055	22.0055	0.00%	0.00%

Note: N, sample size; m, mean value; σ, standard deviation; δ, overlap rate; I is an independent normal population while II is the other one right-adjacent to I. $\delta_{I\to II}$ ($\delta_{II\to I}$) measures the proportion of samples in Population I (II) falling into the overlap domain ($m_1 + 3\sigma_1$)–(m_2-$3\sigma_2$). The * means that the relevant populations may not be used as a class for object recognition because their δ values exceed 90%. The Populations (distributions) I and II used as input were randomly generated based on MATLAB, and the resulting argument values (e.g. t, T, δ, etc.) are not fixed during the simulation runs.

4. Processing results

4.1 Separation of cropland and water

Using bands 1, 2 and 3N in the ASTER VNIR imagery as input to PCA, the histogram of the PC_1 band is given in Figure 4(a–c)g, giving the resulting MML-EM schema. Note that the original PC_1 band (having 7,093,170 pixels) was resized (resampled) using a proportionality factor of 0.35, namely the pixel size was resampled from 15 m to about 42 m (there are 869,430 pixels left for classification). This process helped to eliminate the salt and pepper noise in the classification results. The grey levels indicated at least three populations (in fact four) that are normally and independently distributed, and each of them might correspond to a real-world feature on the ground.

(1) Population (Class) I turns out to be rangelands and croplands (cornfields) and may contain shrubs or groves. The separability between Population I and II is about 0.7220, $\delta_{I\to II} = 54.92\%$, and $\delta_{II\to I} = 97.37\%$, and hence the threshold $T = m_1 = -691.50$ was used here. (2) Pixels in Population III and IV have formed an integrated and independent normal distribution, representing lakes, ponds, rivers and other free-standing water features. The J value between II and III is about 1.75, $\delta_{I\to II} = 30\%$ and $\delta_{II\to I} = 37.63\%$, so the recommended threshold is $T = t = 2395$. (3) Due to spectral similarities, quite a few unwanted anomaly pixels have been mixed into III, most of which reflect mountainous shadows (Figure 5). The weight of I ($\boldsymbol{\alpha}_1$) is 60.50% while $\boldsymbol{\alpha}_{3+4}$ is only 3.42%. The agricultural lands are distinguishable because they usually exhibit lower reflectance (brightness) in the PC_1 imagery; however, there is probably no solution to spectrally discriminate between 'Agricultural lands' and 'Rangelands' subclasses in I, for their separability (J) is close to 0. Pixels reflecting urban or built-up areas are also inseparable within Population II (Figure 4(b)).

4.2 Elimination of false anomalies

In order to eliminate as many false anomalies as possible, while retaining the real ones, topographic analysis-based image clipping was conducted. As seen in Figure 6(a) regarding the Water class, false anomalies pertinent to mountainous shadows were removed completely. During the field survey, we also noticed that in sloping terrain the main land cover types include forests, rangelands and sporadic orchards (Figure 7). Croplands (cornfields) are mainly distributed on flat ground, and, unlike in Asia, people in Kärnten in Austria are not in the habit of building terraces, so the application of ASTER GDEM-derived slope masks helps to give the best results for cropland classification (Figure 6(b)). However, this inevitably comes at the price of introducing many unwanted 'grassland' or 'rangeland' pixels. Finally, we come to the conclusion that the water-area rate of the whole study area is at least 2.79% (namely 44.4771 km^2) and the upper limit of the cropland area is <18.57% (296.34 km^2).

4.3 Extraction of the urban or built-up areas

The PCA bands become featureless in classifying the urban or built-up areas, and instead the RVI image was considered. Its distribution histogram is given in Figure 8, which

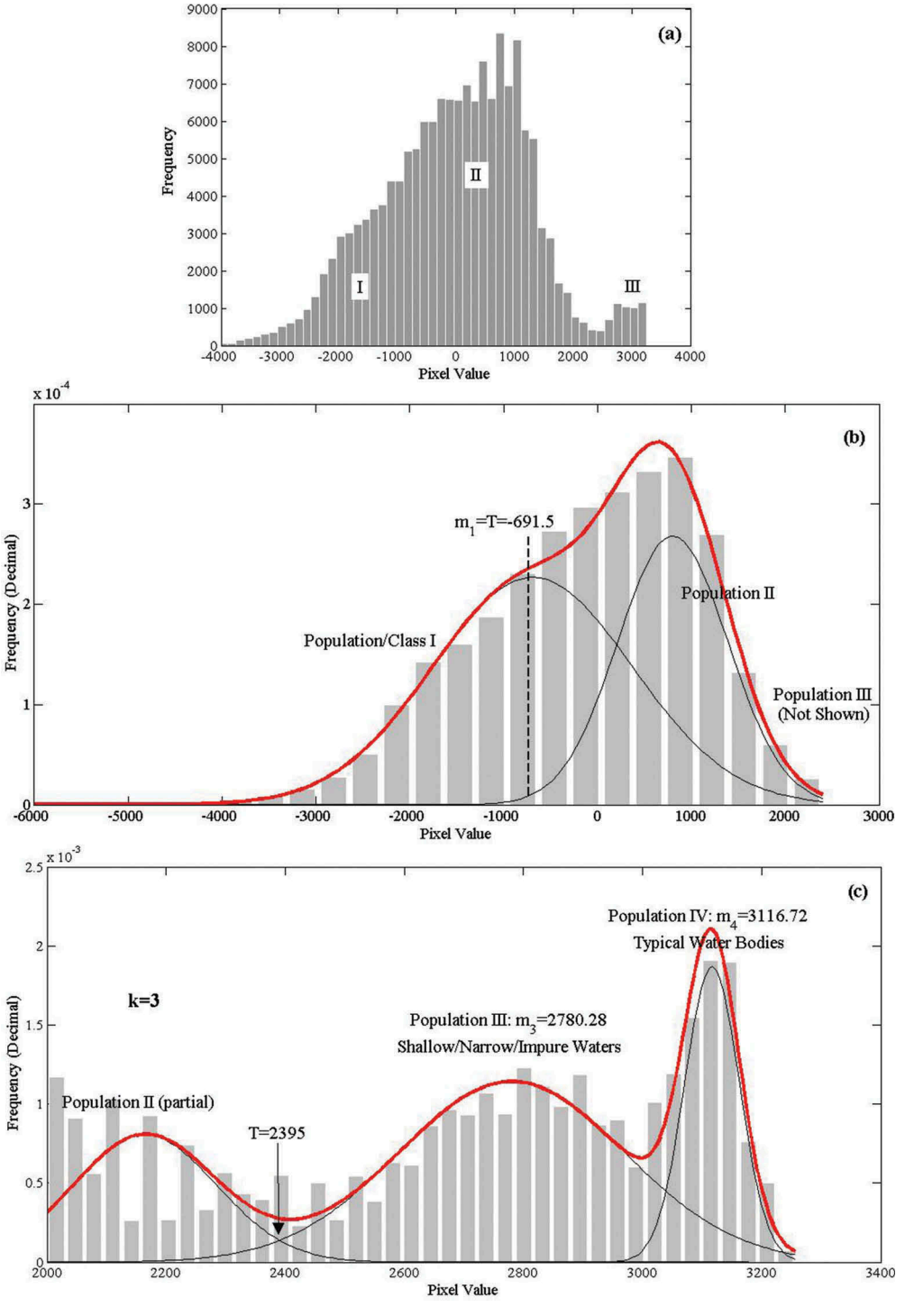


Figure 4. Grey histogram of the PC_1 imagery (a), as well as its MML-EM schema: (b) for Populations I and II, and (c) for IV and V. Note: (1) the computing program (in the MATLAB language) was quoted directly from Figueiredo and Jain (2002). (2) In order to conduct the screening algorithm in MATLAB, the original imagery in ENVI standard format was resampled (resized) using a proportionality factor of 0.35 and saved as ASCII format. (3) *k* represents the optimal number of populations (branches); *Tn* (n = 1, 2…) is the threshold between two neighbouring populations.

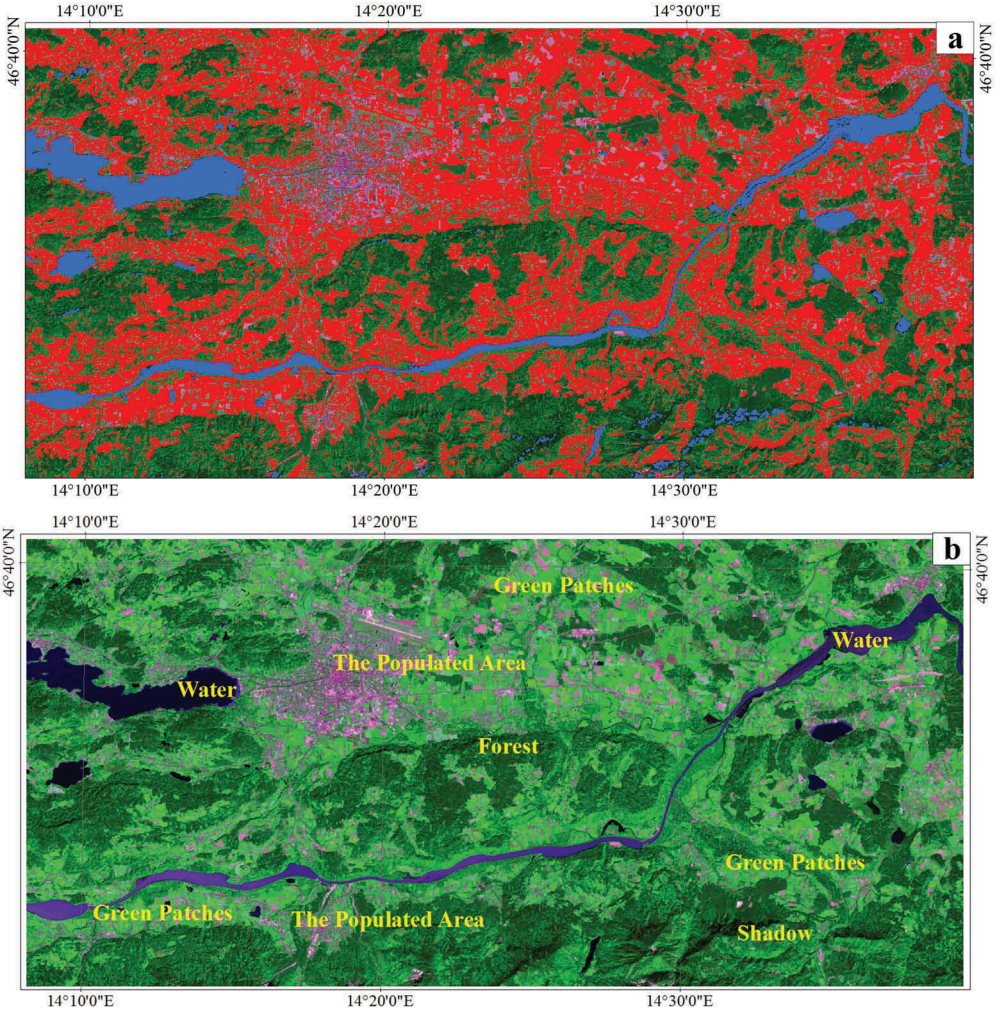


Figure 5. (a) The classification result derived from the MML-EM algorithm (base map: the PC_1 image (local)); (b) the cross-reference ASTER imagery (local) – bands 2, 3N and 1 in RGB. Only the northern study region where the farmlands and natural waters concentrate is exhibited in this diagram. The visual interpretation signs in (b) are as follows: forest: dark green; the populated area: light purple; water: light to dark blue; agricultural land/rangeland: bright green patches; and mountain shadow: dark tone.

shows the preliminary MML-EM analysis result. Again the raw RVI image was resampled using a proportionality factor of 0.350 (42 m pixel size), and there are 869,430 pixels left. As exhibited in Figure 8(a), two main populations (classes) are discriminated, and their separability is $J = 1.1204$, $\delta_{I\to II} = 44.43\%$, and $\delta_{II\to I} = 83.27\%$. Population I has turned out to reflect the non- and sparsely vegetated areas mainly including waters, bare lands and the built-up areas, and its weight $\boldsymbol{\alpha}_1 = 35.75\%$. Here we recommend using $T = m_1 = 4.0082$ as the classification threshold. Population II can discriminate vegetation (including forests, grasslands and croplands) from the background. Evidently,

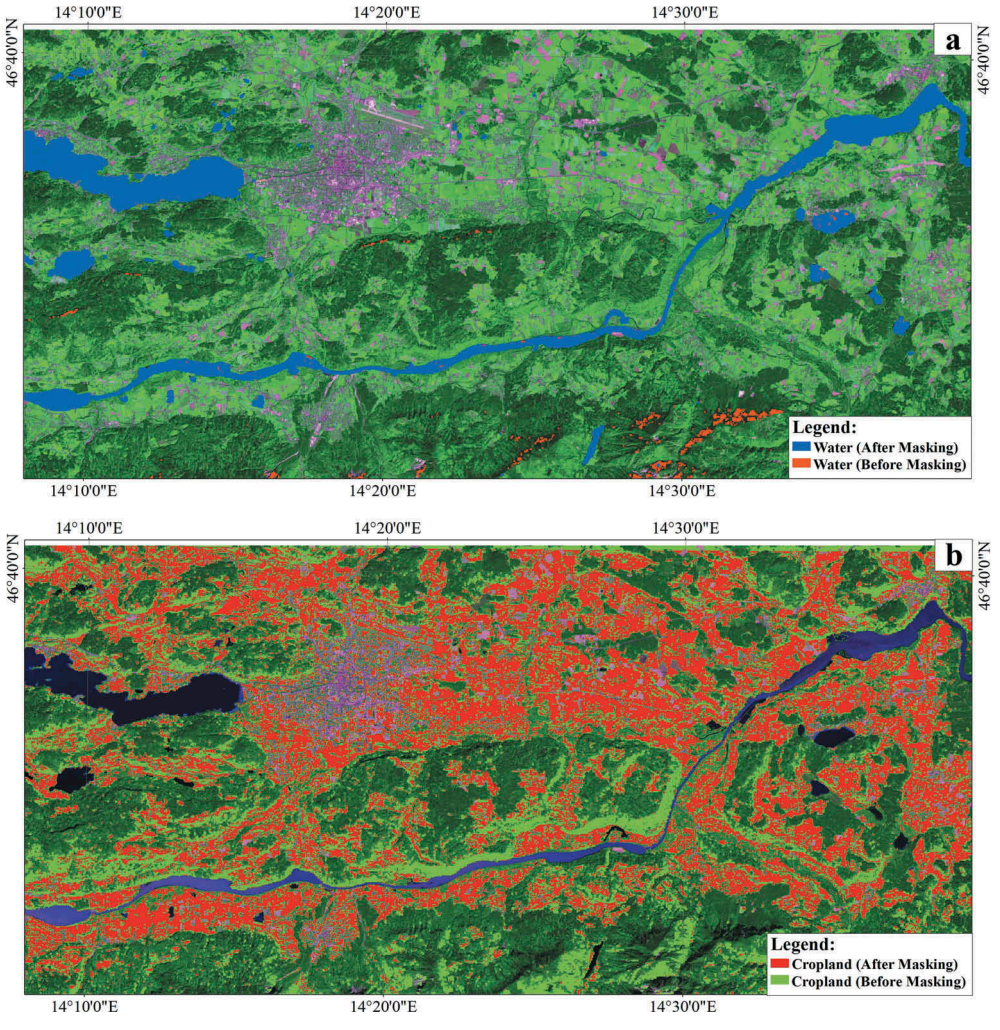


Figure 6. The spatial distribution of natural waters (a) and croplands (including grasslands or rangelands) (b). Note that (1) the anomalous patches are overlaid on the ASTER image (2, 3N and 1 in RGB). (2) The original anomalies of water/cropland (before masking) were spatially clumped using an 11 × 11 kernel size because of their sparse density after the terrain analysis-based imagery clipping/masking.

discrimination of built-up areas is only possible when there are separable subpopulations coexisting within Population I.

Figure 8(b) gives the mixed distribution schema of pixels only within Population I in Figure 8(a), which is a second MML-EM analysis. Based on visual interpretation, the Subpopulation (Subclass) I-1 turns out to be free-standing waters exhibiting stable spectral profiles; due to mixed reflectance, however, pixels in shallow and narrow waters having an unstable spectral expression have been excluded (Jiang *et al*. 2014). Its weight ($a_{1\text{-}1}$) and average ($m_{1\text{-}1}$) are 12.690% and 0.5289, respectively. As shown in Figure 9, the Subpopulations I-2 and I-3 are a good reflection of built-up areas, and $a_{1\text{-}2}$ = 41.84%, $m_{1\text{-}2}$ = 1.7,733, $a_{1\text{-}3}$ = 33.68% and $m_{1\text{-}3}$ = 3.0930. The separability (J)

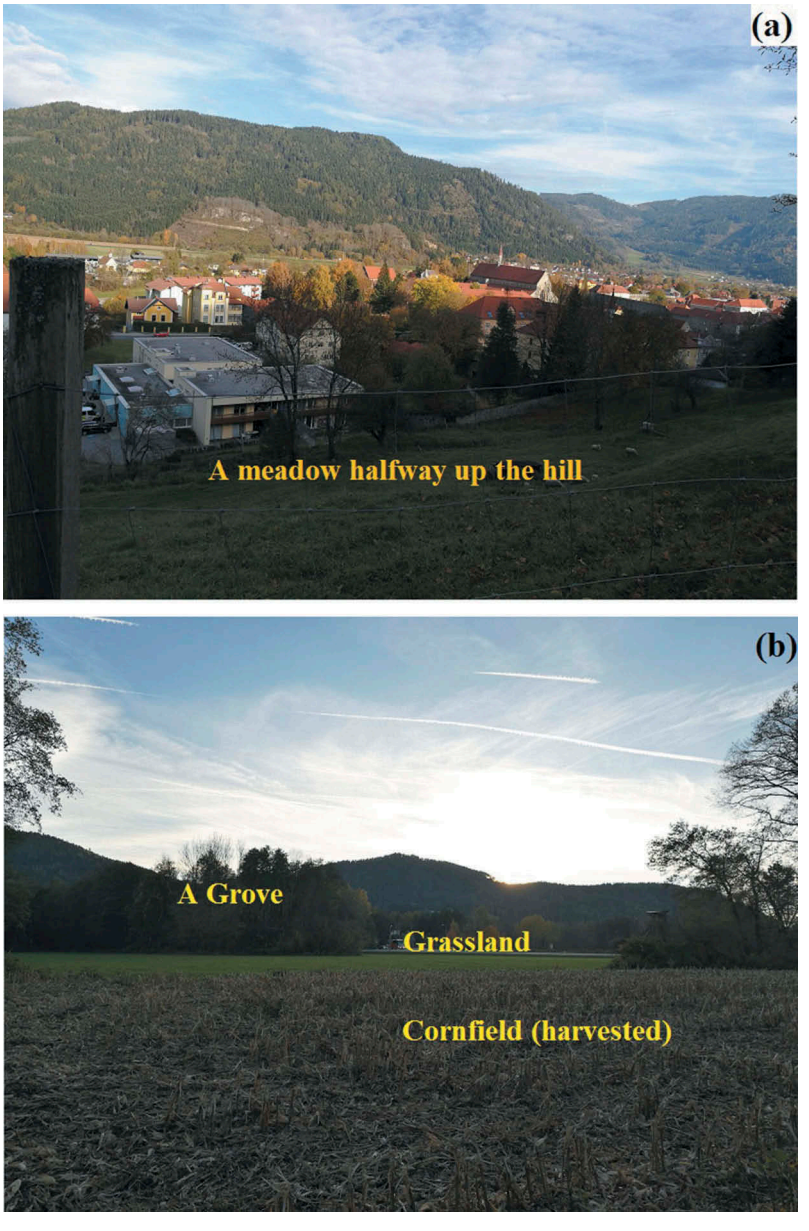


Figure 7. (a) The rangelands halfway up the hill, and these are very common in the study area. Time and location of the photo: 21 October 2017 in Friesach, Austria (46°57′13.3 9″N, 14°23′ 58.91′ ′E). (b) indicates that the grove, grassland and cornfield often coexist in one place (mixed classes). Time and location of the photo: 28 October 2017 in Klagenfurt, Austria (46°36′45.92″N, 14°15′ 42.73″E).

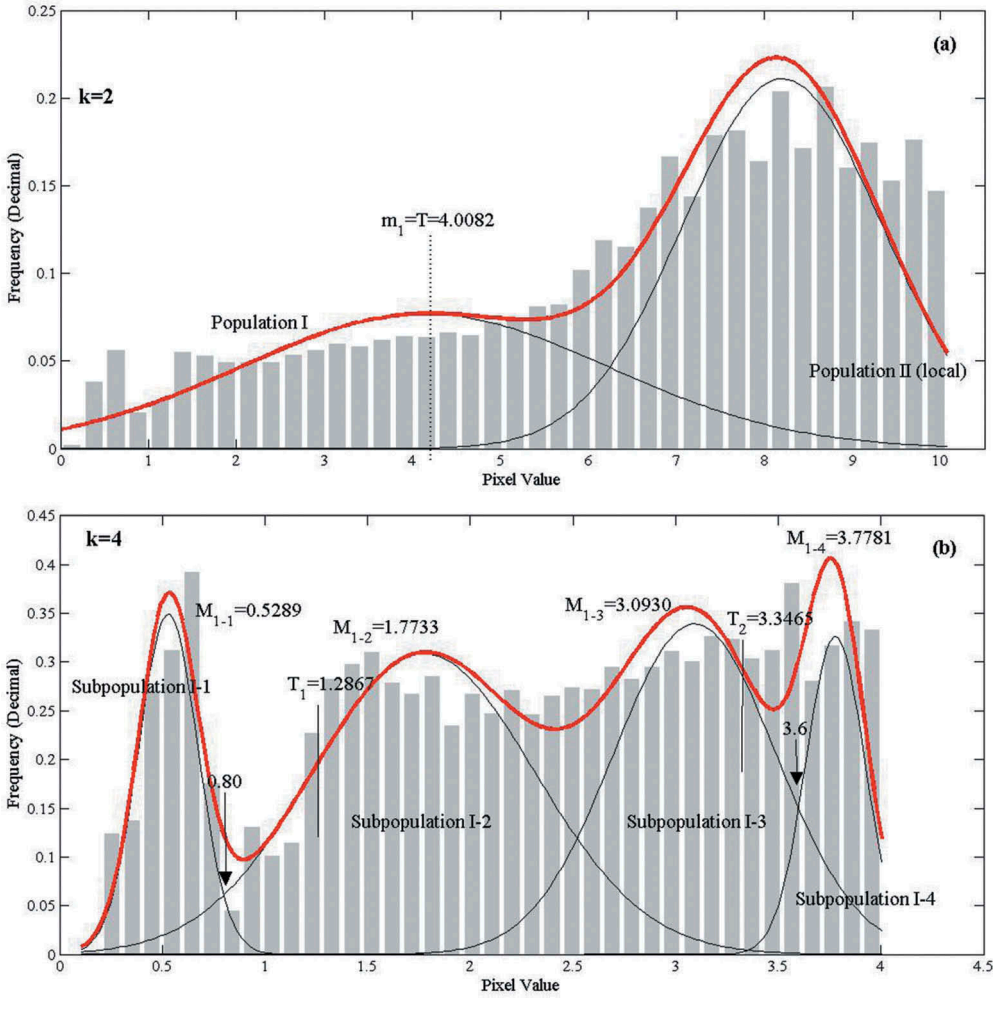


Figure 8. (a) the MML-EM schema of the RVI image and (b) the secondary MML-EM schema for pixels only within Population I in (a). Note that the raw RVI image was resampled using a proportionality factor of 0.35. Although there are fewer pixels in the input image, their standard deviation is only 1.30, implying that the resampling process has no obvious effect on the threshold value. The meanings of the parameters marked in the diagram are the same as those in Figure 4.

between I-1 and I-2 is 1.7045, between I-2 and I-3 it is 1.4303, and between I-3 and I-4 it is 1.307. According to the aforementioned thresholding rules, the pixel-value range of the built-up area class should be within [T_1(1.2,867), T_2(3. 3,465)]. In addition, Subpopulation I-4 is more like a transitional subpopulation between the city and the forest, reflecting clearings, urban green space, tree-lined paths, etc.

A closer scrutiny of Figure 9(a) confirmed that the shallow/narrow/impure waters incorporated in Subpopulation I-2 are greatly reduced by applying the adjustment threshold T_1, and so are the vegetation-dominant pixels occurring in I-3. However, there are some mixed bareland-dependent pixels which mostly come in the form of fallow fields, exposed rocks, snow-capped mountains, etc. This is because the dark-

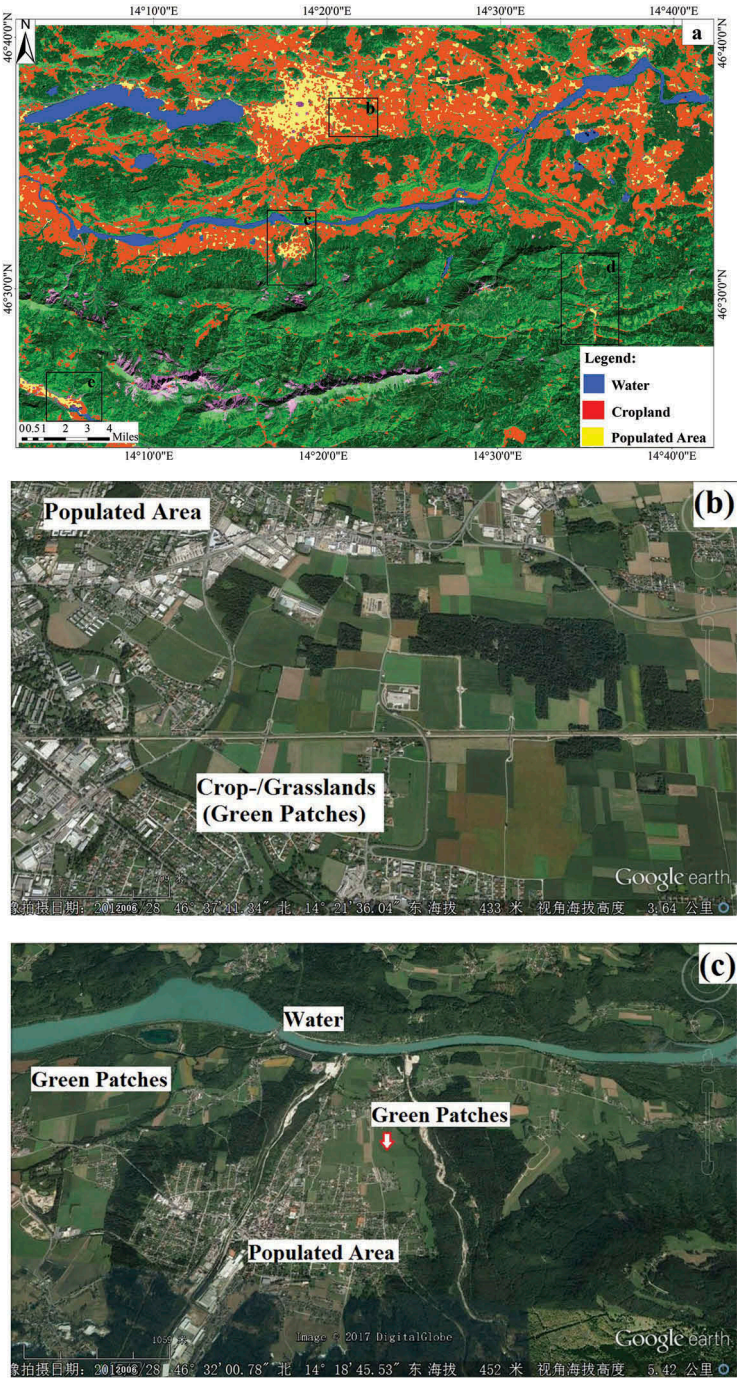
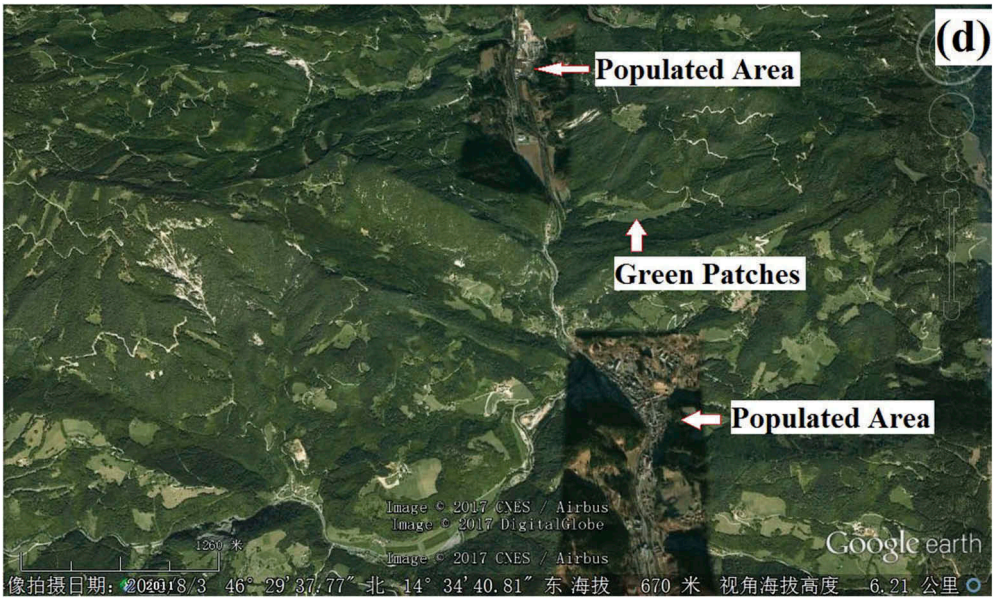


Figure 9. (a) The resulting land cover classification; (b), (c), (d) and (e) are several representative sub-areas for cross-reference. Note: (1) anomaly patches are overlaid on the ASTER imagery (2, 3N and 1 in RGB). (2) The original anomaly pixels of water and built-up areas were empirically clumped using a 10 × 10 kernel size, and those of agricultural areas were clumped using a 5 × 5 size. The 'sieve' process was also conducted to remove the isolated patches or random noise. (3) The item 'crop land' can be confused with grasslands/lawns. (4) (b)–(e) are true-colour Google Earth scenes.



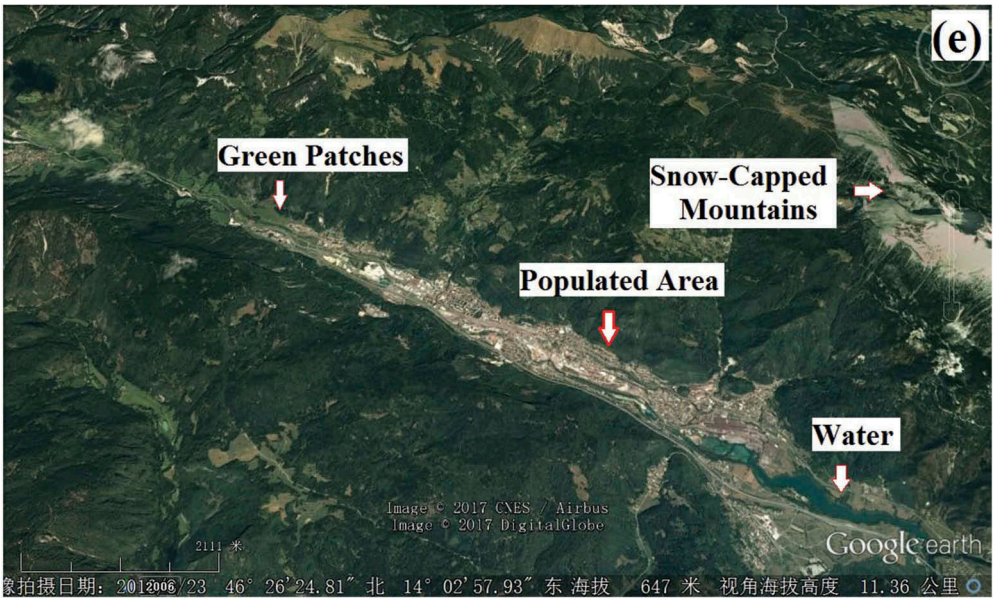


Figure 9. (Continued).

coloured patches in an RVI image are not specific only to built-up areas in the real world. Based on PIE v4.0, here the postprocessing procedure is to overlay the vector shapefiles of postprocessed water + agriculture land classes and the non-processed built-up area class, so that they are intersected to eliminate the relevant false anomalies, and then use the GDEM-derived slope masks to remove residual false anomalies occurring in mountains. The final detected built-up areas are given in Figure 9(b), and are 3.41% (about 54.455 km^2) of the study region.

5. Precision validation

The extracted water features, built-up areas and crop/grass lands were combined to produce the land cover classification result (Figure 9(a)). Generally it indicates a satisfactory result that can be well validated by the Google Earth image (Figure 9(b–e)) and implies the usefulness of MML-EM. As on-site validation is not feasible for land-use mapping on a local or regional scale, here a (digitized) reference base map is utilized to give the 'true' land cover, and it is acquired by artificial vectorization on the basis of existing classification shapefiles, which involves the ACUD (addition, change, update and deletion) operations of these .shpfiles. On that basis, calculating the relevant confusion matrix, Table 2 gives the exact Overall Classification Accuracy quantities of Figure 9(a), which are close to unity in general, implying very good observed agreement (Camps-Valls *et al*. 2011). For comparative purposes, we conducted traditional unsupervised classification methods (K-means and ISODATA) as well, and, as expected, their classification results are eclipsed by the proposed method's. As seen in Table 2, even without topographic masks, the new method still outperforms the traditional ones.

This precision, although satisfactory for the work reported here, is not as good as can be obtained by this method. Actually, the real-world land use status can be very complicated: e.g. a piece of cropland may be bare soil, or cultivated with crops at various growth stages, or with different crops, and it might even be covered by water due to rain or irrigation. For the same reason, a settlement area may consist of roads, grass, shrubs, water and trees in between buildings. Therefore, 'omission errors' are statistically undetectable in many cases, and the relevant quantities may have been overestimated more or less (that is why we use the symbol '<' in Table 2). In Figure 9, the main cause for misclassification should be the incomplete image enhancement and mixed classes, rather than the proposed method itself.

6. Discussion

Quantitatively characterizing the probability distribution pattern of an object is the foundation for performing imagery segmentation (Swarnajyoti *et al*. 2011). In this contribution, the MML-EM algorithm, in combination with terrain analysis, was introduced and utilized to map the distribution of different land use classes in the study region. A homogeneously textured or coloured region in an enhanced image can be well modelled by a Gaussian distribution, whereas an image with complex scenes usually consists

Table 2. Table of the overall classification accuracy.

Ground feature	MML-EM	K-means*	K-means**	ISODATA*	ISODATA**
Surface water	99.7325%	12.8682%	95.8838%	10.8921%	94.7279%
Built-up areas	94.3537%	78.9804%	90.9232%	75.3251%	88.1117%
Crop/grass lands	<99.9970%	83.4078%	94.4032%	81.9697%	94.6028%

Note: * without terrain mask; ** with terrain mask. The overall accuracy is calculated by dividing the total correct pixels by the total number of pixels in the error matrix. Here, the Google Earth images were used as a cross-reference to inspect the existing classification results and to help distinguish confusing feature pixels from background noise, and by using PIE v4.0 the standard or reference base map used for accuracy assessment is digitized by modifying existing classification shapefiles. In addition, the K-means and ISODATA classifications were conducted based on the RVI and PC_1 bands, as is routine in MML-EM analysis.

of multiple intertwined target regions (Bazi *et al*. 2007); so MML-EM is utilized to model the distribution of the features as a mixture of Gaussian distributions. On the other hand, the real-world land use status can be very complicated, and our objective is to classify those well-defined land cover classes, e.g. water, agricultural lands/rangeland, built-up area, and so on; a further subdivision is not possible.

Another meaningful result we achieved concerns the presence of false anomalies: there is typically an overlapping or cumulative frequency domain between two adjacent populations (classes) which are assumed to be normally and independently distributed; this domain, due to its transitional nature, is a possible source of the spurious anomalies pertinent to spectral similarities. This argument is supported by several observations in Figure 10: in (a), urban areas and water features coexist within the overlapping domain between I-1 and I-2 in Figure 8(b); regarding the settlements, water features act as the false anomalies, and vice versa. In (b), the domain between I-3 and I-4 in Figure 8(b) generally represents the 'buildings in wooded grounds'. In this sense, no thresholding method can achieve 100% accuracy as there is overlap between the classes. There are many reasons to believe that the resulting accuracy is at least over 95%, or close to 100% in local terrain – like the case in Figures 9(b–e) and 10. Nevertheless, closer scrutiny indicated that in the classified output, some misclassified, isolated pixels (noise) or small regions of pixels still exist. This will not have any appreciable impact on the overall interpretation accuracy, but gives the output a 'salt and pepper' appearance. The object-based classification that can group pixels into a small number of representative shapes and sizes was developed to cope with the problem; it offers the MML-EM program fewer input data, thus being useful in improving the computation efficiency and stability. We will report on this issue in a separate study.

7. Conclusions

Using the ASTER VNIR image as data source, an attempt has been made to conduct land use classification in southern Austria. We obtained a satisfactory classification result (the kappa coefficients are close to unity in general). Given the validity of the normal approximation assumption, the key findings from our work are as follows.

(1) The PC_1 band has greatly enhanced the targeted information of water features and green patches in the imagery, while there is probably no solution to spectrally discriminate between 'cropland' and 'grassland/lawn' classes. The RVI band is more applicable for discriminating water features and populated areas.
(2) We obtained a satisfactory result of land cover classification that is well validated by the Google Earth image, and it has demonstrated the capability of MML-EM for extracting useful information from processed remotely sensed products. The overall accuracies of agricultural land/rangeland, water and built-up area are always above 94%, and the kappa coefficients above 0.9 (except the built-up area class).
(3) The water-area ratio of this region is expected to be 2.79%, the farmland area ceiling is more than 290 km^2 and the built-up areas are 3.41% (54.455 km^2) of the study region. To balance against the interpretation accuracy required, here terrain analysis-based image clipping was introduced to eliminate rugged or steep

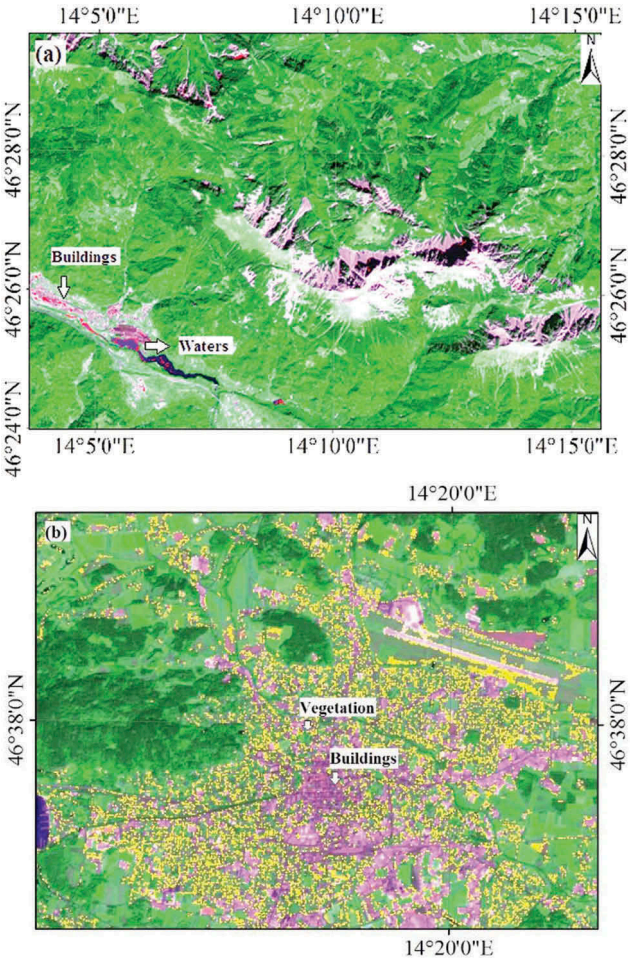


Figure 10. (a) The anomaly pixels (shown in red) corresponding to the overlapping domain between I-1 and I-2 in Figure 8(b) (the RVI range: 0.7–1.1, centred at 0.9); (b) the anomalies (shown in yellow) corresponding to the domain between I-2 and I-3 and 4 in Figure 8(b) (the RVI range: 2.31–2.71, centred at 2.51). Note that these anomaly pixels are shown in a partial base map (ASTER 2, 3N and 1 in RGB).

terrain where false anomalies pertinent to spectral similarities or random noise appear most frequently, and retain the areas of level terrain within which the features of interest occur.

(4) Acknowledging the uncertainties, this is the first time that MML-EM has been applied for ASTER VNIR data-driven image classification in Europe, and based on it an extended separability-threshold model was used to avoid the issue of generalized and inappropriate product thresholds as well as offering more subtle insights into how different ground features are related to and distinguished from each other. In addition, terrain and derived terrain variables (such as slope) may have major implications for improving the correct classification rate during imagery classification.

Acknowledgements

Helpful review comments on the manuscript were provided by several anonymous reviewers, and we would like to thank them here.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was financially supported by the NSFC (Natural Science Foundation of China) projects 'formation mechanism of snowball quartz and relationship between snowball quartz and raremetal mineralization (Authorization No. 41773030)' and 'mechanism of five-story vertical morphological zonation in vein-type wolframite deposits (No. 41373048)'.

ORCID

Dehui Zhang http://orcid.org/0000-0002-7229-0970

References

Bazi, Y., Bruzzone, L., and Melgani, F., 2007. Image thresholding based on the EM algorithm and the generalized Gaussian distribution. *Pattern Recognition*, 40 (2), 619–634. doi:10.1016/j.patcog.2006.05.006

Bazi, Y. and Melgani, F., 2010. Gaussian process approach to remote sensing image classification. *IEEE Transactions on Geoscience & Remote Sensing*, 48 (1), 186–197. doi:10.1109/TGRS.2009.2023983

Biernacki, C., Celeux, G., and Govaert, G., 2000. Assessing a mixture model for clustering with the integrated completed likelihood. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22 (7), 719–725. doi:10.1109/34.865189

Camps-Valls, G., *et al*., 2011. Introduction to the issue on advances in remote sensing image processing. *IEEE Journal of Selected Topics in Signal Processing*, 5 (3), 365–369. doi:10.1109/JSTSP.2011.2142490

Christakos, G., Bogaert, P., and Serre, M., 2012. *Temporal GIS: advanced functions for field-based applications*. Heidelberg: Springer Science & Business Media.

Comber, A., *et al*., 2012. Spatial analysis of remote sensing image classification accuracy. *Remote Sensing of Environment*, 127 (140), 237–246. doi:10.1016/j.rse.2012.09.005

Feizi, F., Mansouri, E., and Ramezanali, A.K., 2016. Prospecting of Au by remote sensing and geochemical data processing using fractal modelling in Shishe-Botagh area (NW Iran). *Journal of the Indian Society of Remote Sensing*, 44 (4), 1–14. doi:10.1007/s12524-015-0510-0

Figueiredo, M.A.T. and Jain, A.K., 2002. Unsupervised learning of finite mixture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24 (3), 381–396. doi:10.1109/34.990138

Gao, Y., *et al*., 2011. Object-based classification with features extracted by a semi-automatic feature extraction algorithm–sEaTH. *Geocarto International*, 26 (3), 211–226. doi:10.1080/10106049.2011.556754

Geneletti, D. and Gorte, B.G.H., 2003. A method for object-oriented land cover classification combining Landsat TM data and aerial photographs. *International Journal of Remote Sensing*, 24 (6), 1273–1286. doi:10.1080/01431160210144499

Guo, L.B. and Gifford, R.M., 2010. Soil carbon stocks and land use change: a meta analysis. *Global Change Biology*, 8 (4), 345–360. doi:10.1046/j.1354-1013.2002.00486.x

Han, L., Zhao, B., and Jürgen, P., 2017a. *Extraction of the ore-caused anomaly and study of the metallogenic efficiency in (heavily-covered) western Zhejiang province of China*. Saarbrücken: Golden Light Academic Publishing. in Chinese with English Abstracts.

Han, L., *et al*., 2017b. An integrated approach for extraction of lithology information using the SPOT 6 imagery in a heavily Quaternary- covered region — north Baoji district of China. *Geological Journal*, 53, 352–363. doi:10.1002/gj.3061

Huete, A., *et al*., 2000. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 83 (1–2), 195–213. doi:10.1016/S0034-4257(02)00096-2

Jiang, H., *et al*., 2014. An automated method for extracting rivers and lakes from Landsat imagery. *Remote Sensing*, 6 (6), 5067–5089. doi:10.3390/rs6065067

Jiang, M., *et al*., 2010. Towards hypothesis testing and lossy minimum description length: a unified segmentation framework. *In*: *Asian Conference on Computer Vision*, November 2010 Queenstown, Berlin, Heidelberg: Springer, 343–354.

Ju, J., Kolaczyk, E.D., and Gopal, S., 2003. Gaussian mixture discriminant analysis and sub-pixel land cover characterization in remote sensing. *Remote Sensing of Environment*, 84 (4), 550–560. doi:10.1016/S0034-4257(02)00172-4

Kerroum, M.A., Hammouch, A., and Aboutajdine, D., 2010. Textural feature selection by joint mutual information based on Gaussian mixture model for multispectral image classification. *Pattern Recognition Letters*, 31 (10), 1168–1174. doi:10.1016/j.patrec.2009.11.010

Kurita, T. and Abdelmalek, O.N., 1992. Maximum likelihood thresholding based on population of mixture models. *Pattern Recognition*, 25, 1231–1240. doi:10.1016/0031-3203(92)90024-D

Lillesand, T., Kiefer, R.W., and Chipman, J., 2014. *Remote sensing and image interpretation*. Hoboken: John Wiley & Sons.

Lin, C., *et al*., 2015. Effects of atmospheric correction and Pansharpening on LULC classification accuracy using WorldView-2 imagery. *Information Processing in Agriculture*, 2 (1), 25–36. doi:10.1016/j.inpa.2015.01.003

Liu, L., *et al*., 2017. Mineral mapping and ore prospecting using Landsat TM and Hyperion data, Wushitala, Xinjiang, northwestern China. *Ore Geology Reviews*, 81, 280–295. doi:10.1016/j.oregeorev.2016.10.007

Liu, X.C., *et al*., 2011. MML-EM algorithm and its application on mixed distributions of geochemical data. *Earth Science — Journal of China University of Geosciences*, 36 (2), 355–359. [in Chinese with English abstract].

Mars, J.C. and Rowan, L.C., 2010. Spectral assessment of new ASTER SWIR surface reflectance data products for spectroscopic mapping of rocks and minerals. *Remote Sensing of Environment*, 114 (9), 2011–2025. doi:10.1016/j.rse.2010.04.008

Muñoz-Marí, J., Bruzzone, L., and Camps-Valls, G., 2007. A support vector domain description approach to supervised classification of remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 45 (8), 2683–2692. doi:10.1109/TGRS.2007.897425

Nemes, F., *et al*., 1997. The klagenfurt basin in the eastern alps: an intra-orogenic decoupled flexural basin? *Tectonophysics*, 282 (1–4), 189–203. doi:10.1016/S0040-1951(97)00219-9

Nussbaum, S., Niemeyer, I., and Canty, M.J., 2006. SEATH – a new tool for automated feature extraction in the context of object-based image analysis. *In*: *1st International Conference on Object-based Image Analysis (OBIA)*, July 2006 Austria: Salzburg.

Panahi, A. and Cheng, Q., 2004. Multifractality as a measure of spatial distribution of geochemical patterns. *Mathematical Geology*, 36 (7), 827–846. doi:10.1023/B:MATG.0000041181.32596.5d

Richards, J.A. and Richards, J.A., 1999. *Remote sensing digital image analysis*. Vol. 3, Berlin, Heidelberg: Springer.

Schowengerdt, R.A., 2006. *Remote sensing: models and methods for image processing*. Amsterdam: Elsevier.

Sinclair, A.J., 1991. A fundamental approach to threshold estimation in exploration geochemistry: probability plots revisited. *Journal of Geochemical Exploration*, 41 (1–2), 0–22. doi:10.1016/0375-6742(91)90071-2

Sridharan, H. and Qiu, F., 2013. A spatially disaggregated areal interpolation model using light detection and ranging-derived building volumes. *Geographical Analysis*, 45 (3), 238–258. doi:10.1111/gean.2013.45.issue-3

Sun, W., *et al*., 2006. Fractal analysis of remotely sensed images: a review of methods and applications. *International Journal of Remote Sensing*, 27 (22), 4963–4990. doi:10.1080/01431160600676695

Swarnajyoti, P., Susmita, G., and Ashish, G., 2011. Histogram thresholding for unsupervised change detection of remote sensing images. *International Journal of Remote Sensing*, 32 (21), 6071–6089. doi:10.1080/01431161.2010.507793

Tobias, O.J. and Seara, R., 2002. Image segmentation by histogram thresholding using fuzzy sets. *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, 11 (12), 1457–1465. doi:10.1109/TIP.2002.806231

Vistelius, A. and Andrew, B., 1960. The skew frequency distributions and the fundamental law of the geochemical processes. *The Journal of Geology*, 68 (1), 1–22. doi:10.1086/626634

Wang, H., Chen, J.S., and Yu, X.M., 2013. Feature selection and its application in object- oriented classification. *Journal of Remote Sensing*, 17 (4), 816–829. [in Chinese with English abstract].

Xu, R. and Wunsh, D., II, 2005. Survey of clustering algorithms. *IEEE Transactions on Neural Networks*, 16 (3), 645–678. doi:10.1109/TNN.2005.845141

Zhang, S.F., 2004. EM algorithm and its application in parameter estimation for Gaussian mixture. *Journal of Spacecraft TT and C Technology*, 23 (4), 47–52. [in Chinese with English abstract].

Zhang, Y.J., *et al*., 2014. *Practical techniques for extraction of ore exploration information from multispectral remote sensing data*. Beijing: Geological Publishing House.

Zhao, B., *et al*., 2016. A spatial Gaussian mixture model for optical remote sensing image clustering. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9 (12), 5748–5759. doi:10.1109/JSTARS.2016.2546918

Zhao, P.D., Hu, W.L., and Li, Z.J., 1994. *Statistical prediction of mineral deposit*. 2nd. Beijing: Geological Publishing House.