Front. Earth Sci. 2015, 9(3): 437–448
DOI 10.1007/s11707-014-0476-1

RESEARCH ARTICLE

# Identifying AMSR-E radio-frequency interference over winter land

**Sibo ZHANG, Li GUAN (✉)**

Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters, Key Laboratory for Aerosol-Cloud-Precipitation of China Meteorological Administration, Nanjing University of Information Science & Technology, Nanjing 210044, China



**Abstract** Satellite microwave emission mixed with signals from active sensors is referred to as radio-frequency interference (RFI). RFI affects greatly the quality of data and retrieval products from space-borne microwave radiometry. An accurate RFI detection will not only enhance geophysical retrievals over land but also provide evidence of the much-needed protection of the microwave frequency band for satellite remote sensing technologies. It is difficult to detect RFI from space-borne microwave radiometer data over winter land, because RFI signals are usually mixed with snow in mid-high latitudes. A modified principal component analysis (PCA) method is proposed in this paper for detecting microwave low frequency RFI signals. Only three original variables, one RFI index (sensitive to RFI signal) and two scattering indices (sensitive to snow scattering), are included in the vector for principal component analysis in this modified method instead of the nine or seven RFI index original variables used in a normal PCA algorithm. The principal component with higher correlation and contribution to the original RFI index is the RFI-related principal component. In the absence of a reliable validation data set of the "true" RFI, the consistency in the identified RFI distribution obtained from this method compared to other independent methods, such as the spectral difference method, the normalized PCA method, and the double PCA method, give confidence to the RFI signals' identification over land. The simple and reliable modified PCA method could successfully detect RFI not only in summer but also in winter AMSR-E data.

**Keywords** microwave remote sensing, radio-frequency interference (RFI), the Advanced Microwave Scanning Radiometer for Earth Observing System (AMSR-E), principal component analysis (PCA)

Received December 5, 2013; accepted May 29, 2014

E-mail: liguan.nuist@gmail.com

## 1 Introduction

X-band and C-Band have been occupied and used by both passive and active microwave devices in recent years (Zou, 2012). Satellite passive microwave radiometer-measured natural radiation mixed with signals from active microwave transmitters is referred to as radio-frequency interference (RFI) (Ruf et al., 2006; Guner et al., 2007; Misra and Ruf, 2008; Piepmeier et al., 2008). RFI emerges because artificial emissions from active microwave transmitters easily contaminate the remote measurements of the Earth's relatively weak scattered and emitted radiation by adding unpredictable spurious noise. The interference signals come dominantly from low-frequency active microwave transmitters, including radar, air traffic control, cell phone, garage door remote control, GPS signals on highways, defense tracking, and vehicle speed detection for law enforcement, etc. (Zhao et al., 2013). RFI seriously affects the quality of the data and retrieval products from space-borne microwave radiometers (Wentz et al., 2000; Weng et al., 2001; Njoku et al., 2004; Yang and Weng, 2011).

The Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) was on board the Aqua satellite in 2002 (Kawanishi et al., 2003). AMSR-E provides dual-polarized passive microwave measurements at six frequencies, 6.9 GHz, 10.7 GHz, 18.7 GHz, 23.8 GHz, 36.5 GHz, and 89 GHz. However, the observations from its C-band (6.9 GHz) and X-band (10.7 GHz) revealed strong and widespread radio-frequency interference (RFI), which further impacted soil moisture estimation (Njoku and Li, 1999; Njoku et al., 2003; Chaurasia et al., 2012). Besides, RFI was also found in the WindSat which was on board the Department of Defense Coriolis satellite (Ellingson and Johnson, 2006; Li et al., 2006) and in the Microwave Radiation Imager (MWRI) X-band, which was on board the Chinese second-generation polar-orbiting satellite (FY-3B) (Yang et al., 2011; Zou et al.,

2012). The low-frequency observations of the latest AMSR-2, which was the payload of the Global Change Observation Mission Water (GCOM-W) satellite launched in May 2012, might also be contaminated by RFI. All of these sensors have similar channel frequencies near X-band and C-Band.

To properly identify and correct the RFI contamination, a spectral difference technique has been developed to quantify the RFI magnitude and scope over land areas (Li et al., 2004). Based on examining the spatial and temporal characteristics of the RFI by the use of spectral indices (differences between brightness temperatures at two different frequencies for a given polarization), Njoku et al. (2005) demonstrated that using means and standard deviations of the spectral indices was effective in identifying strong RFI. Surface often produces smooth and ultrawideband microwave radiation which is different from RFI signatures. The multichannel correlations of radiometer data from natural radiations are usually high. The spectral difference technique did not explicitly utilize this channel correlation information from natural radiations. Li et al. (2006) proposed a principal component analysis (PCA) method with several consecutive RFI indices. Their PCA method integrated the spectral difference technique with the multivariate correlation of radiometer data. The spectral difference method and the PCA method both worked well for data over land in summer, but failed to detect RFI reliably over frozen ground and scattering surface.

Scattering from natural targets (e.g., snow and ice) can decrease substantially the observed brightness temperatures at high frequencies (Kunzi et al, 1976). The reversed (positive) spectral difference can also be derived from surface snow and ice. To avoid taking snow and ice as false RFI signals, Zou et al. (2012) added a scattering-sensitive spectral difference criterion $TB_{89}-TB_{18}\geqq -10$ K for detecting RFI signals over continental Asia and Europe in winter, based upon the X-band observations of FY-3B MWRI (MicroWave Radiation imager). Although this method was simple and convenient to understand, it did not work when snow was mixed together with RFI signals in high- latitude winter conditions. It removed the snow areas directly and might include some real RFI regions. To improve this deficiency they proposed a normalized PCA (NPCA) method to detect winter RFI signals. The RFI indices must be normalized by the mean and standard deviation before principal component analysis. Using the NPCA method, the snow scattering effects were removed without altering the PCA-detected RFI results. The latest method used to identify WindSat RFI over Greenland and Antarctic regions was a double PCA (DPCA) method developed by Zhao et al. (2013). The DPCA method took advantage of the multi-channel correlation for natural land and ice surface radiations, as well as the de-correlation between different frequencies for RFI detection. This method increased the stability and the reliability of the RFI identification better than the NPCA method; meanwhile it was effective in detecting C- and X-band RFI signals under permanent sea ice-covered conditions and other snow-free or snow-rich regions.

Previous research has proven that RFI detection for satellite microwave imagery of low-frequency radiances over land is extremely important before the data can be used for either geophysical retrievals or data assimilation in numerical weather prediction models. Detection of RFI in space-borne microwave radiometer data is difficult under snow/ice-covered underlying surface where snow/ice scattering signal mixes with RFI signal. A software solution must be derived to identify and remove RFI. It is necessary to develop a simplified method to quantify the RFI magnitude and extent. A modified PCA (MPCA) method to detect the RFI-contaminated AMSR-E data over snow-covered land was developed with this research, which performed well not only in summer but also in winter. A combination of one RFI Index (RI) and two Scattering Indices (SI) replaced the original variables (consecutive several RFI Indices) in the previous PCA methods.

This paper is organized as follows: Section 2 provides a brief description of AMSR-E data. Section 3 introduces some RFI detection methods and describes the modified RFI identification method in detail. Numerical examples are presented in Section 4. Section 5 gives some conclusions.

## 2 AMSR-E data

The AMSR-E is onboard the National Aeronautics and Space Administration (NASA) Earth Observing System (EOS) Aqua satellite which was launched on 4 May 2002 with equator crossings at 1:30 pm and 1:30 am local time. AMSR-E has vertical and horizontal polarization channels at six frequencies, i.e., 6.925 GHz, 10.65 GHz, 18.7 GHz, 23.8 GHz, 36.5 GHz, and 89.0 GHz, for a total of 12 channels. The AMSR-E provides a fixed earth incidence angle of 55°. The antenna beams scan conically about the nadir axis with an observation swath-width of 1,445 km. The spatial resolution of the individual measurement decreases from 5.4 km at 89 GHz to 56 km at 6.9 GHz.

The AMSR-E Level 1A instrument data are provided by the Japan Aerospace Exploration Agency (JAXA). Then Level 2A brightness temperatures are generated by reconstructing the AMSR-E antenna gain patterns at each channel to five footprints at the Remote Sensing Systems (RSS) facility in Santa Rosa, CA. Hence, there are five different resolutions for the re-sampled data, from 56 km to 5.4 km. The resolution of 56 km was chosen for each channel in this study. For the Level 2A data if the measured brightness temperatures of one channel exceed 330 K the value is set to zero. The same Field of View (FOV) observations of other channels are all set to zero. The

FOVs with zero value observations are represented by some small white areas in the displayed figures of brightness temperature.

## 3 RFI detection methods

### 3.1 Spectral difference technique

The principle of the spectral difference RFI detection technique is based on the distinct signature differences between natural radiation and RFI induced by human activities. Because RFI has narrower bandwidth relative to natural radiation, its emission/scattering characteristics are very different. The brightness temperature of natural surface tends to increase with frequency at frequencies below 30 GHz, while RFI can only increase brightness temperature significantly at a particular frequency and generate a negative spectral gradient (Li et al., 2004). Thus, Li proposed an RFI Index (RI) to identify the location of RFI and quantify its intensity:

$$\mathrm{RI}_{f_1 p} = \mathrm{TB}_{f_1 p} - \mathrm{TB}_{f_2 p}, \tag{1}$$

where TB denotes brightness temperature, the subscript $p$ stands for horizontal or vertical polarization, and $f_1$ and $f_2$ represent the two neighborhood frequencies ($f_1 < f_2$). For example, land RFI is identified for the 6-GHz horizontal polarization channel if the RI exceeds a threshold beyond its natural variation (5 K) (Li et al., 2004; Wu and Weng, 2011).

$$\mathrm{RI}_{6\mathrm{H}} = \mathrm{TB}_{6\mathrm{H}} - \mathrm{TB}_{10\mathrm{H}}. \tag{2}$$

### 3.2 Principal Component Analysis (PCA) method

In principle, RFI identification relies on two kinds of information. The first one is the emission/scattering characteristics represented by RFI indices, which are the physical information determined by the Earth materials and structures of the natural targets. The second one is the correlations between different channels or indices, which is the statistical information introduced by the natural variability of the targets. The multichannel correlations of radiometer data from natural radiations are often high. While the channel correlations between RFI contaminated channels and the remaining AMSR-E channels are low since RFI only increases brightness temperatures significantly at a particular frequency. The spectral difference technique does not explicitly utilize multichannel correlations of radiometer data, which are key information in separating RFI from natural radiations.

Li et al. (2006) advised the principal component analysis (PCA) of RFI Indices over land. This method integrates statistics of target emission/scattering characteristics (through RFI Indices) and multivariate correlation of radiometer data into a single statistical framework for PCA. It generates an RFI-related principal component that is orthogonal to multivariate natural radiations. For 6 GHz RFI identification, the RFI indices used in the PCA are

$$\overrightarrow{\mathbf{RFI}}_{\text{indicies}} = \begin{pmatrix} \mathrm{RI}_{6\mathrm{H(V)}} \\ \mathrm{RI}_{10\mathrm{V}} \\ \mathrm{RI}_{10\mathrm{H}} \\ \mathrm{RI}_{18\mathrm{V}} \\ \mathrm{RI}_{18\mathrm{H}} \\ \mathrm{RI}_{23\mathrm{V}} \\ \mathrm{RI}_{23\mathrm{H}} \\ \mathrm{RI}_{36\mathrm{V}} \\ \mathrm{RI}_{36\mathrm{H}} \end{pmatrix} = \begin{pmatrix} \mathrm{TB}_{6\mathrm{H(V)}} - \mathrm{TB}_{10\mathrm{H(V)}} \\ \mathrm{TB}_{10\mathrm{V}} - \mathrm{TB}_{18\mathrm{V}} \\ \mathrm{TB}_{10\mathrm{H}} - \mathrm{TB}_{18\mathrm{H}} \\ \mathrm{TB}_{18\mathrm{V}} - \mathrm{TB}_{23\mathrm{V}} \\ \mathrm{TB}_{18\mathrm{H}} - \mathrm{TB}_{23\mathrm{H}} \\ \mathrm{TB}_{23\mathrm{V}} - \mathrm{TB}_{36\mathrm{V}} \\ \mathrm{TB}_{23\mathrm{H}} - \mathrm{TB}_{36\mathrm{H}} \\ \mathrm{TB}_{36\mathrm{V}} - \mathrm{TB}_{89\mathrm{V}} \\ \mathrm{TB}_{36\mathrm{H}} - \mathrm{TB}_{89\mathrm{H}} \end{pmatrix}. \tag{3}$$

To avoid the coupling of RFI signals between vertical and horizontal polarizations, $\mathrm{RI}_{6\mathrm{V}}$ and $\mathrm{RI}_{6\mathrm{H}}$ are grouped into separate datasets with the rest of the RFI Indices.

In the area where data have a higher projection onto the first PC mode, RFI is presented, in general (Li et al., 2006). The principal component scores of RFI components are then used to identify RFI-contaminated observations. RFI is identified for scores with a value greater than zero. Higher score values indicate stronger RFI.

### 3.3 Normalized Principal Component Analysis (NPCA) method

However, neither the spectral difference nor PCA methods detect RFI reliably over frozen grounds and scattering surfaces. The scattering from natural targets (e.g., snow and ice) can substantially decrease brightness temperatures at high frequencies (Zou et al., 2012). The reversed spectral differences can also be related to surface snow and ice. This presents a problem in simply implementing the spectral difference method or PCA method that works for data in summer to winter cases. To avoid taking snow and ice as false RFI signals from the aforementioned indices in

winter, Zou suggested a normalized PCA (NPCA) method (Zou et al., 2012). The RFI Indices need to be normalized before the principal component decomposition. For example, for 6 GHz RFI identification the RFI indices used in the NPCA method are

$$\overrightarrow{\mathbf{RFI}}^{\text{normalized}}_{\text{indicies}} = \begin{pmatrix} \frac{RI_{6H(V)}-\mu}{\sigma} \\ \frac{RI_{10V}-\mu}{\sigma} \\ \frac{RI_{10H}-\mu}{\sigma} \\ \frac{RI_{18V}-\mu}{\sigma} \\ \frac{RI_{18H}-\mu}{\sigma} \\ \frac{RI_{23V}-\mu}{\sigma} \\ \frac{RI_{23H}-\mu}{\sigma} \end{pmatrix} = \begin{pmatrix} \frac{TB_{6H(V)}-TB_{10H(V)}-\mu}{\sigma} \\ \frac{TB_{10V}-TB_{18V}-\mu}{\sigma} \\ \frac{TB_{10H}-TB_{18H}-\mu}{\sigma} \\ \frac{TB_{18V}-TB_{23V}-\mu}{\sigma} \\ \frac{TB_{18H}-TB_{23H}-\mu}{\sigma} \\ \frac{TB_{23V}-TB_{36V}-\mu}{\sigma} \\ \frac{TB_{23H}-TB_{36H}-\mu}{\sigma} \end{pmatrix} \quad (4)$$

where $\mu$ and $\sigma$ are the mean and standard deviation of the RFI indices. The snow scattering effects are removed by normalizing the RFI indices. This works more effectively than PCA for identifying RFI signals in the presence of snow.

### 3.4 Double Principal Component Analysis (DPCA) method

Zhao proposed the DPCA method, which proved to be effective in snow-free or snow-rich regions, such as winter data over high latitude (Zhao et al., 2013). The DPCA method developed for global RFI detection consists of two PCA steps. In the first PCA step a vector of ten-variable brightness temperature (except for the 89 GHz) is defined for PCA analysis. The data matrix can finally be reconstructed using the PC coefficients by two parts. The sharp gradients across the edge of ice sheets or snow are captured by the correlated matrix $\boldsymbol{A}_1$, which is reconstructed by first number PC modes due to strong channel by channel correlations for these natural surface conditions. RFI signals are contained in the residual matrix $\boldsymbol{A}_2$. The second step of the DPCA method is to apply a normalized PCA to the residual matrix $\boldsymbol{A}_2$. The DPCA works with both non-scattering and scattering surfaces.

### 3.5 Modified Principal Component Analysis (MPCA) method

The PCA method is modified for RFI identification of AMSR-E data over high latitude snow areas in winter. The MPCA method takes advantage of the de-correlation between RFI frequencies with other channels and the multi-channel correlation for natural land including snow surface radiations.

Surface brightness temperatures tend to increase with frequency at frequencies below 30 GHz, so RI values should be negative over natural land surfaces. RFI is the possible cause of positive RI values (Li et al., 2004). Moderate and strong RFI can be detected if the RI exceeds a threshold (5 K) beyond its natural variation. The higher positive RI value means stronger RFI. But the scattering effect of the snow-rich or ice regions cannot be ignored in winter. When volume scattering effects dominate, brightness temperatures are reduced with increasing frequencies, thereby snow areas also make RI positive. To describe the low-frequency snow characteristics, the 18 GHz and 36 GHz frequencies were often used as a Scattering Index (SI $= TB_{18}-TB_{36}$) in numerous previous studies (Grody, 1991; Grody and Basist, 1996; Kelly et al., 2003; Qiu et al., 2010). The TBs at 18.7−36.5GHz (SI) are more reliable for snow evaluation than those at high frequency. Since the scattering effects from snow and ice are much stronger at the high-frequency channels (36 GHz) than at the low-frequency channels (10.65 GHz or 18.7 GHz), high positive SI values mean radiation from scattering surfaces. So we add the variable SI in the vector for principal components analysis. Since the aim of this paper is to identify RFI signal instead of snow detection, we just display qualitatively the snow distribution from the high-frequency 89 GHz observation (Fig. 1(f)) and Scattering Index (Figs. 2(b) and 2(c)).

In this research, we found that the objective of identifying RFI in winter still can be reached by retaining part channels of AMSR-E and using PCA calculation only once. The Scattering Index was introduced in our RFI analysis vector together with RI to identify the RFI signals mixed with snow. There are three variables constructing the vector for principal component analysis in our modified PCA method, one RFI index (RI) and two Scattering Indices (SI). The first RI variable is sensitive to RFI contamination and the SI is snow-sensitive. Hence, we believe that replacing the consecutive RFI indices with RI and SI will help to identify RFI distribution and exclude snow disturbance.

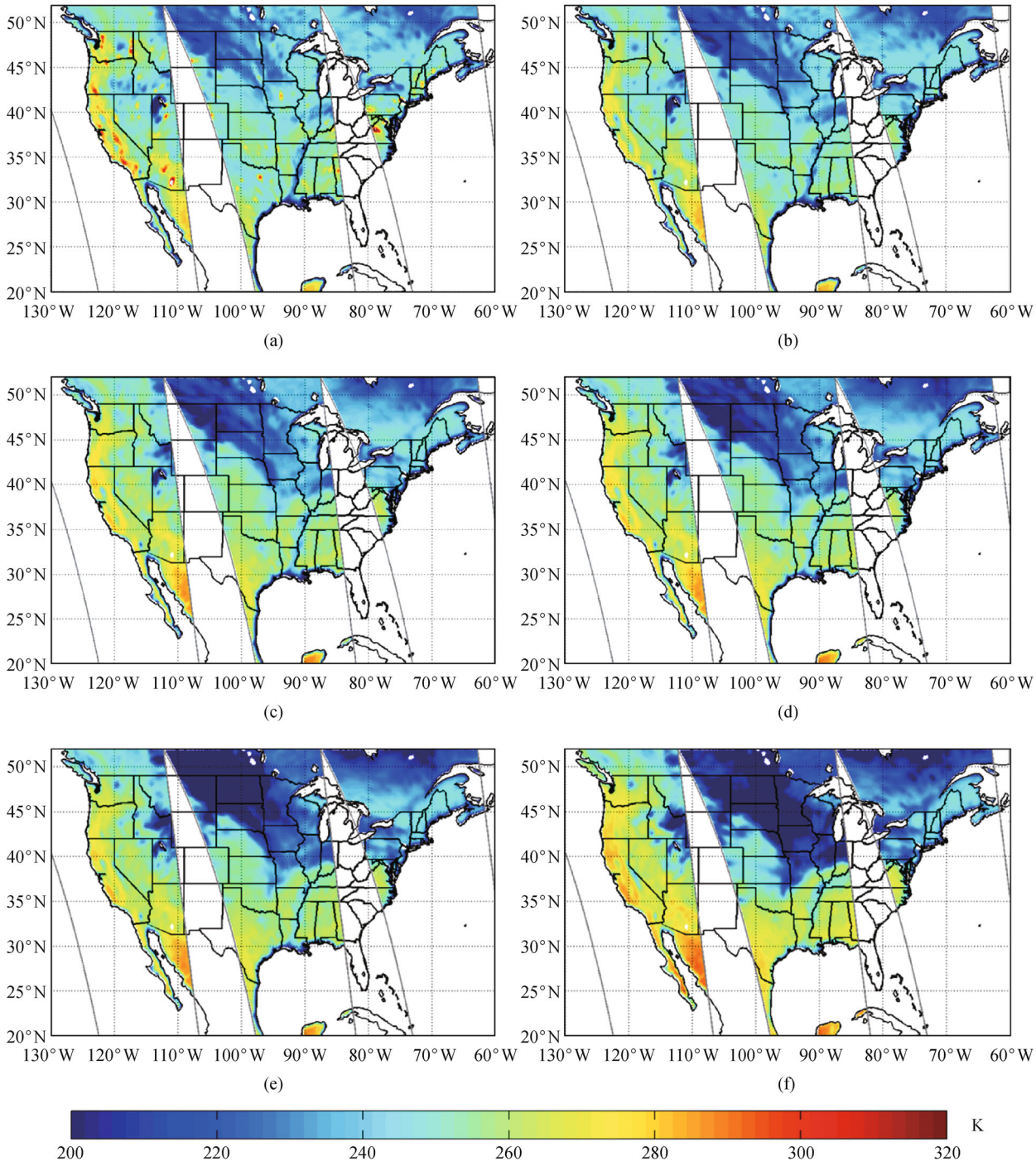


**Fig. 1** AMSR-E brightness temperatures at (a) 6.9 GHz, (b) 10.7 GHz, (c) 18.7 GHz, (d) 23.8 GHz, (e) 36.5 GHz and (f) 89 GHz with horizontal polarizations for ascending passes over the United States on February 10, 2011.

For identifying RFI at 6.9 GHz a vector of three-variable RFI indices is defined as

$$\overrightarrow{\mathbf{RFI}}_{\text{indicies}} = \begin{pmatrix} \mathrm{RI}_{6\mathrm{H(V)}} \\ \mathrm{SI_V} \\ \mathrm{SI_H} \end{pmatrix} = \begin{pmatrix} \mathrm{TB}_{6\mathrm{H(V)}} - \mathrm{TB}_{10\mathrm{H(V)}} \\ \mathrm{TB}_{18\mathrm{V}} - \mathrm{TB}_{36\mathrm{V}} \\ \mathrm{TB}_{18\mathrm{H}} - \mathrm{TB}_{36\mathrm{H}} \end{pmatrix}. \tag{5}$$

For at 10.7 GHz is

$$\overrightarrow{\mathbf{RFI}}_{\text{indicies}} = \begin{pmatrix} \mathrm{RI}_{10\mathrm{H(V)}} \\ \mathrm{SI_V} \\ \mathrm{SI_H} \end{pmatrix} = \begin{pmatrix} \mathrm{TB}_{10\mathrm{H(V)}} - \mathrm{TB}_{18\mathrm{H(V)}} \\ \mathrm{TB}_{18\mathrm{V}} - \mathrm{TB}_{36\mathrm{V}} \\ \mathrm{TB}_{18\mathrm{H}} - \mathrm{TB}_{36\mathrm{H}} \end{pmatrix}. \tag{6}$$

The three-dimensional data matrix for identifying RFI at

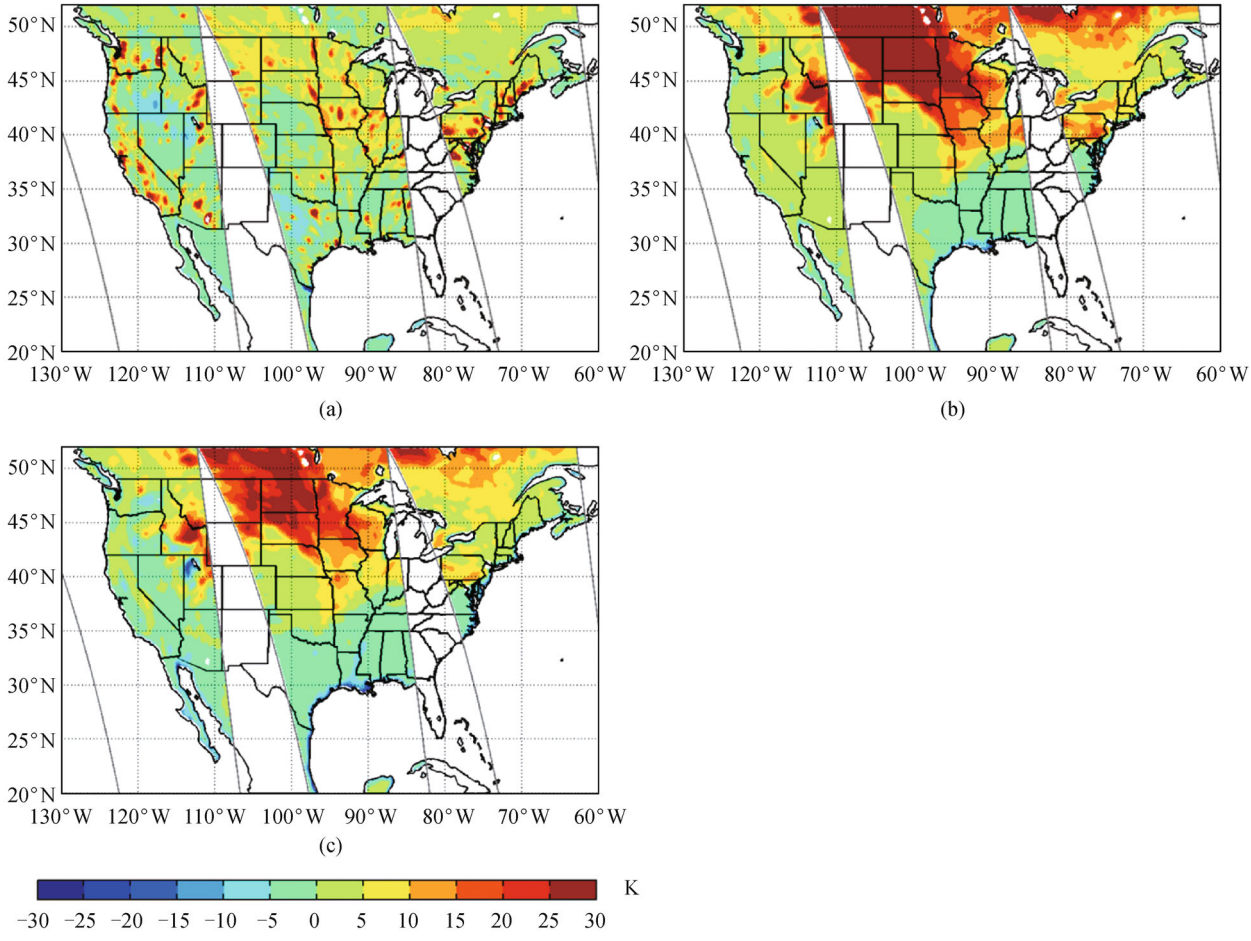


**Fig. 2** (a) $RI_{6H}$, (b) $SI_V$, and (c) $SI_H$ distribution over the United States on 10 February, 2011 from AMSR-E ascending measurements.

6.9 GHz horizontal polarization using PCA is defined as,

$$\boldsymbol{A}_{3\times N}=\begin{pmatrix}(\mathrm{TB_{6H}}-\mathrm{TB_{10H}})_1 & (\mathrm{TB_{6H}}-\mathrm{TB_{10H}})_2 & \cdots\cdots & (\mathrm{TB_{6H}}-\mathrm{TB_{10H}})_N\\ (\mathrm{TB_{18V}}-\mathrm{TB_{36V}})_1 & (\mathrm{TB_{18V}}-\mathrm{TB_{36V}})_2 & \cdots\cdots & (\mathrm{TB_{18V}}-\mathrm{TB_{36V}})_N\\ (\mathrm{TB_{18H}}-\mathrm{TB_{36H}})_1 & (\mathrm{TB_{18H}}-\mathrm{TB_{36H}})_2 & \cdots\cdots & (\mathrm{TB_{18H}}-\mathrm{TB_{36H}})_N\end{pmatrix}, \tag{7}$$

where $N$ is the total number of data points over a specified region.

Next, construct the covariance matrix $\boldsymbol{S}_{5\times5}=\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}$, whose eigenvalues ($\lambda=[\lambda_1,\lambda_2,\lambda_3]$) and eigenvectors ($u=[u_1,u_2,u_3]$) are found to satisfy the following equation:

$$\boldsymbol{S}\boldsymbol{u}=\boldsymbol{\lambda}\boldsymbol{u}. \tag{8}$$

Then, the data matrix $\boldsymbol{A}$ can be projected onto a new orthonormal data space as

$$Z_{3\times N}=\boldsymbol{u}^{\mathrm{T}}\boldsymbol{A}=\begin{pmatrix}\mathrm{Z_1}\\ \mathrm{Z_2}\\ \mathrm{Z_3}\end{pmatrix}. \tag{9}$$

The elements of Z are often defined as the principal components, which are uncorrelated with each other (Rothrock et al., 1988). Each principal component (e.g., $Z_1$, $Z_2$) is an exact linear combination (i.e., weighted sum) of the original variables (e.g., RI, SI) (Lattin et al., 2003). The first principal component ($Z_1$) exhibits maximum variance, and the second principal component ($Z_2$) accounts for the remaining maximum variance not yet accounted for by $Z_1$. PCA is a method for re-expressing multivariate data to highlight their similarities and differences. In our data matrix, the first principal component tends to describe the similarities of the original data set, while the second and third ones present the uncorrelation.

Based on the fact that multichannel correlations of radiometer data are often high from natural radiations but low between an RFI low-frequency channel (6.8 GHz or 10.7 GHz) with the remaining AMSR-E channels, the RFI-related principal component is extracted for RFI detection. The principal component which has the highest correlation

and contribution to the original RFI index is the RFI-related principal component. Higher values of the RFI-related principal component indicate stronger RFI. The similarities and correlations of original variables will be described in $Z_1$ when RFI mixes together with snow because snow and RFI areas both own larger variance. In winter, the variance formed by widespread snow data is usually larger than RFI data. $Z_2$ is the RFI-related principal component and denotes the uncorrelated information. Higher values of the second principal component ($Z_2$) would suggest greater probabilities of RFI. In summer, only the RFI areas own larger variance under non-scattering surfaces or shallow snow regions, so the RFI signals will be presented by $Z_1$. Therefore this modified PCA method works not only in winter, but also in the summer season. In winter, the RFI-related principal component is $Z_2$, while in summer it is $Z_1$. Based on previous research, the threshold to identify the RFI signal was also set to 5 K (Li et al., 2004; Wu and Weng, 2011). Those pixels with the RFI-related principal component greater than 5 K are identified as RFI effect.

Compared to PCA and NPCA methods, computer time is saved by our MPCA method. The vector for analysis is simplified to only three variables, while nine or seven variables are used in the PCA or NPCA methods. The vector in the first decomposition of the DPCA method includes ten variables; the second PCA step has seven variables. The dimensions in the data matrix are decreased greatly, so the calculation of the covariance matrix, eigenvalues, eigenvectors, and the principal components are faster. The NPCA method needs to normalize the RFI indices before decomposition, and DPCA has to do PCA decomposition twice, which are all time consuming. It takes just several seconds for one AMSR-E orbit to do the MPCA analysis. Another advantage over other PCA methods is that the principal component with higher correlation and contribution to the original RFI index is used as the RFI-related principal component for RFI identification, whereas the PCA-identified RFI signal always appears in the first principal component by previous PCA methods.

## 4 MPCA algorithm application

To avoid coupling of RFI signals between vertical and horizontal polarizations, $RI_{6h}$ and $RI_{6v}$ were grouped into separate datasets from the rest of the indices and used in two independent PCA studies, which have very similar results. Therefore, only the case of horizontal polarization will be presented here. For the purposes of this research, a snow case over the northern United States on 10 February 2011 was taken as an example. Figure 1 shows AMSR-E brightness temperature observations at 6.9 GHz (Fig. 1(a)), 10.7 GHz (Fig. 1(b)), 18.7 GHz (Fig. 1(c)), 23.8 GHz (Fig. 1(d)), 36.5 GHz (Fig. 1(e)), and 89 GHz (Fig. 1(f)) with horizontal polarization for ascending passes over the United States on 10 February, 2011. In all the figures presented, gray lines indicate the swath width bound of each orbit, and the blank zones between two intersected gray lines are the observation gap areas. There are some visible scattered "hot spots" with red color in the 6.9 GHz measurements (Fig. 1(a)) (e.g., along the United States West and East Coasts) where the brightness temperatures are abnormally higher than the naturally-emitted radiation. These "hot spots" are potential locations of RFI. Since the scattering effects from snow and ice are much stronger at the high-frequency channels, the dark-blue areas over 40 N with brightness temperature below 240 K from the observation of channel 36.5 GHz (Fig. 1(e)) and 89 GHz (Fig. 1(f)) denote areas covered by thick snow. Table 1 shows the correlation matrix for horizontal polarization channels from land data over the U.S. on February 10, 2011. The 6.9 GHz channel has significantly decreased correlations against all other channels. The correlation decreases from 0.81 to 0.67 as the frequency increases. Except for the 6.9 GHz channels, correlations between other channels remain quite high, around 0.9. Therefore, the RFI signal in 6.9 GHz is indeed de-correlated with natural radiation in the contaminated areas.

From Fig. 1(a) and Fig. 1(b) we can see that the RFI at 6.9 GHz distributes widely over the United States and is less populated at 10.7 GHz. So the first variable in the data matrix of our modified PCA method used $RI_6$. The brightness temperature difference, 6.9 minus 10.7 GHz with horizontal polarization ($RI_{6H}$), 18.7 minus 36.5 GHz with vertical polarization ($SI_V$), and horizontal polarization ($SI_H$) are displayed in Fig. 2(a), Fig. 2(b) and Fig. 2(c), respectively. The moderate and strong RFI areas are indicated when $RI_{6H}$ is larger than 5 K (with red color) in Fig. 2(a). Snow also contributes to the positive $RI_{6H}$ (e.g., the yellow regions in Montana and North Dakota). This

**Table 1** Correlation matrix for horizontal polarization channels from land data over the U.S. on February 10, 2011

| | 6H | 10H | 18H | 23H | 36H |
|---|---|---|---|---|---|
| 6H | 1.0000 | 0.81412 | 0.77312 | 0.74112 | 0.67152 |
| 10H | | 1.0000 | 0.97091 | 0.92380 | 0.90323 |
| 18H | | | 1.0000 | 0.98585 | 0.94499 |
| 23H | | | | 1.0000 | 0.97769 |
| 36H | | | | | 1.0000 |

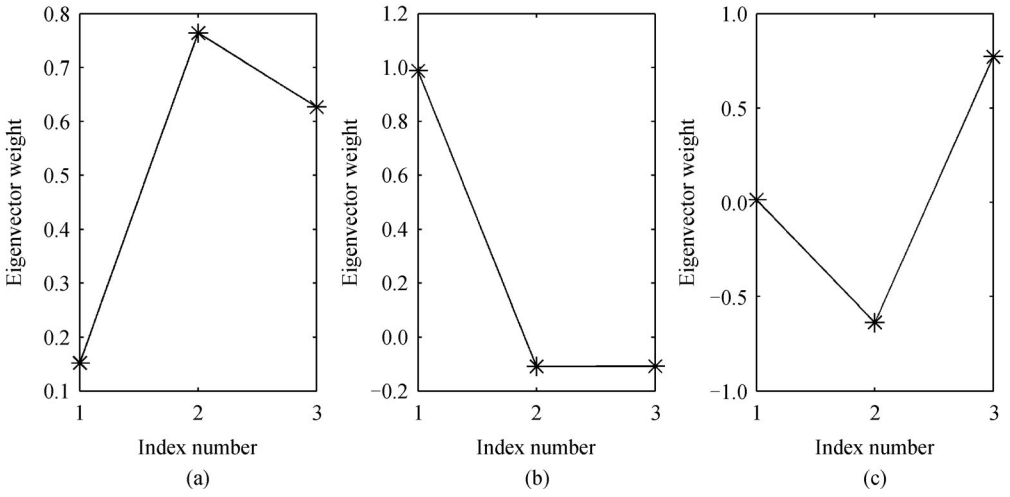


**Fig. 3** Three eigenvectors of MPCA indices: (a) the first eigenvector; (b) the second eigenvector; (c) the third eigenvector.

presents a problem when simply implementing the spectral difference method, that works for data in summer, to winter cases. Snow areas are depicted by SI index in Figs. 2(b) and 2(c) with values greater than 10 K (brown and red color). The spectral difference method only with the positive spectral gradient criterion $RI_{6H} > 5$ K cannot separate RFI from the mixed snow surfaces.

Table 2 gives us the statistical characteristics of these three original variables for the studied winter snow case. The variance of $RI_{6H}$ (Fig. 2(a)), $SI_V$ (Fig. 2(b)), and $SI_H$ (Fig. 2(c)) are 44.07, 126.59, and 86.4, respectively. We also can see from Fig. 2 that the snow index contributes more variance because it is widespread with big values. After principal components analysis Fig. 3 shows the enginevectors of the three principal components derived from the MPCA method. Each principal component (e.g., $Z_1$, $Z_2$) is an exact linear combination (i.e., weighted sum) of the original variables (e.g., RI, SI). In the second principal component $Z_2$, SI owns small weight (about −0.1) while RFI-contaminated $RI_{6H}$ owns a 1.0 weight that is much larger than that of SI (Fig. 3(b)). In the first principal component $Z_1$, more weight contribution comes from SI (Fig. 3(a)). The second principal component $Z_2$ is a probable RFI-related component. According to Eq. (9), the principal components $Z_1$, $Z_2$, and $Z_3$ are computed. The values are shown in Table 3. The variance contribution of the first principal component $Z_1$ to accumulative variances is 83.08%, and the second principal component $Z_2$ is 15.58%. However, the correlation coefficient of $Z_2$ with $RI_{6H}$ (0.9311) is far higher than $Z_1$ or $Z_3$ with $RI_{6H}$ (0.051 or 0.0001). Based on the weight contribution and correlation to the RFI information, the principal component $Z_2$ is the RFI-related principal component in winter. So the principal component $Z_2$ is displayed in Fig. 4(a). The magnitude and scope of RFI can be identified reasonability when the $Z_2$ of the MPCA method is greater than 5 K. Compared with Fig. 2(a), the false RFI identification in high latitudes (yellow color with $RI_{6H} >$ 5K) derived by snow influence are dispelled at the same time the RFI areas are saved.

**Table 2** Characteristics of each original variable

| Original variables | Variance | Percentage of variance/% |
|---|---|---|
| $RI_{6H}$ | 44.07 | 17.14 |
| $SI_V$ | 126.59 | 49.23 |
| $SI_H$ | 86.46 | 33.63 |

To compare our MPCA method with other PCA methods, we computed the first principal component of the PCA method, the NPCA method, and the DPCA method separately. The results are shown in Fig. 4(c), Fig. 4(d), and Fig. 4(e). The pattern of our MPCA method is very similar to that of the DPCA method. In the DPCA method a coastline mask has been applied to remove ocean and large inland water regions. However, brightness temperatures within land pixels near the coastline are affected by ocean emission viewed in the antenna sidelobes (Njoku et al., 2005). So the red lines along the coastline are not RFI regions in Fig. 4(e). In Fig. 4(c) the

**Table 3** Characteristics of each principal component

| Principal component | Variance | Percentage of variance/% | Correlation coefficient with $RI_{6H}$ | Correlation coefficient with $SI_V$ | Correlation coefficient with $SI_H$ |
|---|---|---|---|---|---|
| $Z_1$ | 213.62 | 83.08 | 0.0510 | 0.7585 | 0.6177 |
| $Z_2$ | 40.06 | 15.58 | 0.9311 | 0.0066 | 0.0079 |
| $Z_3$ | 3.44 | 1.34 | 0.0001 | 0.0667 | 0.1188 |

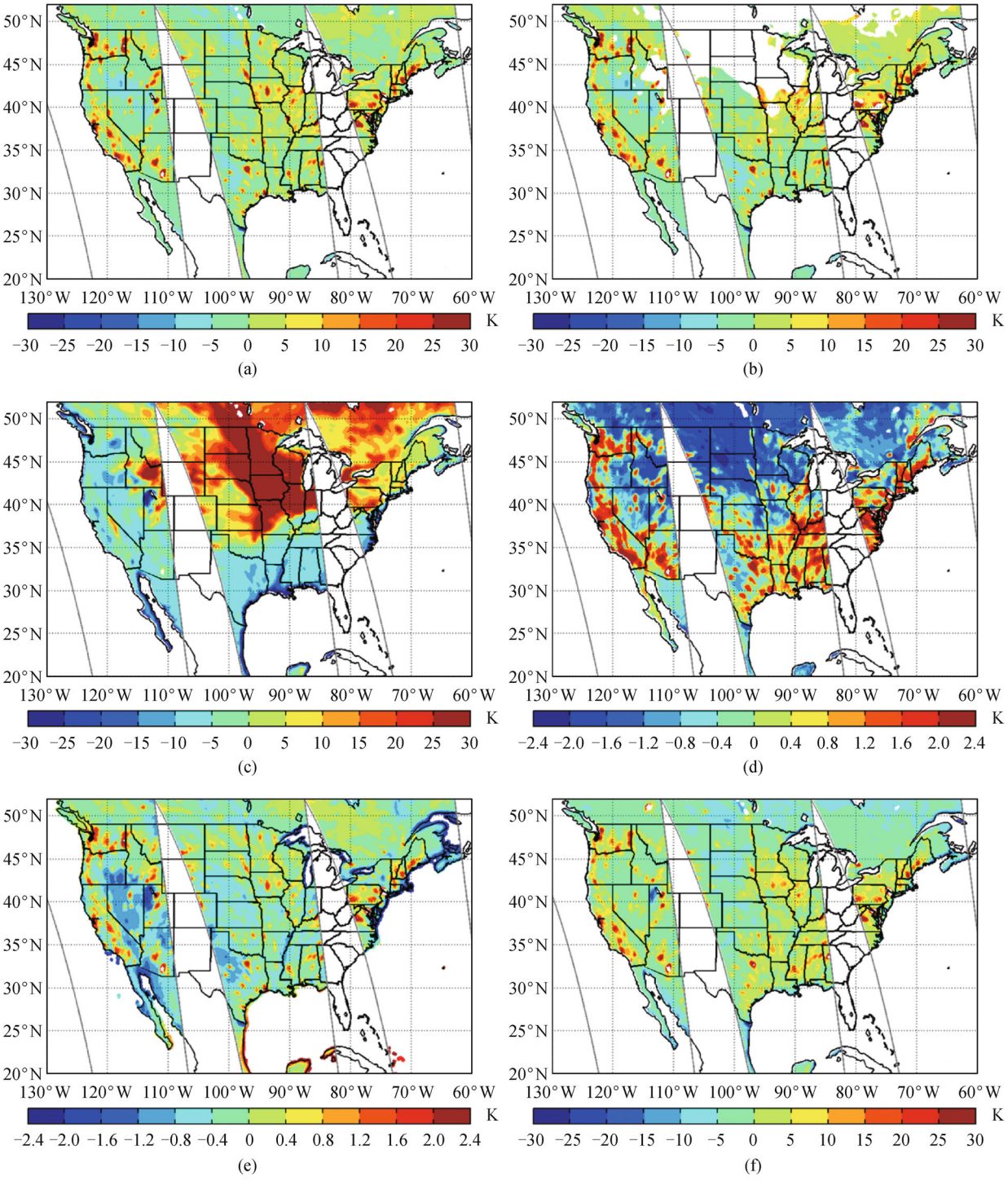


**Fig. 4** RFI-related principal component of 6.9 GHz with horizontal polarization over the U.S. on 10 February, 2011 calculated by (a) our modified PCA method, (b) spectral difference method (only show the data that meet $TB_{18h}-TB_{89h} < 10$ K), (c) PCA method, (d) NPCA method, (e) DPCA method and (f) our MPCA method on 4 July, 2011.

values of snow area are much larger than real RFI signals. The spectral difference of 6.9 GHz and 10.7 GHz measurements is shown in Fig. 4(b) with pixels $TB_{18h}-TB_{89h} > 10$ K removed. This scatter criterion $TB_{18h}-TB_{89h} > 10$ K can identify snow areas in high latitudes, but at the same time the RFI regions mixed with snow are removed too.

RFI signals typically originate from a wide variety of coherent point target sources and are often isolated. So the locations of the RFI distribution tend to be persistent in time, though their magnitudes exhibit temporal and directional variability (Li et al., 2004). The first principal component $Z_1$ of our MPCA method for a summer case on 4 July, 2011 is shown in Fig. 4(f). Figure 4(f) is very close to Fig. 4(a) and Fig. 4(e). The location of RFI in summer

and winter is nearly same, which validates the effect of the modified PCA method.

The RFI distributions at 6.9 GHz with horizontal polarization over the U.S. on 10 February, 2011 identified by the modified PCA method are shown in Fig. 5(a). To illustrate the robustness of the proposed modified RFI detection methods, the RFI distributions are displayed in Figs. 5(b)−5(e), which are the results from the spectral difference method, the PCA method, the NPCA method, and the DPCA method, respectively. RFI-free locations are

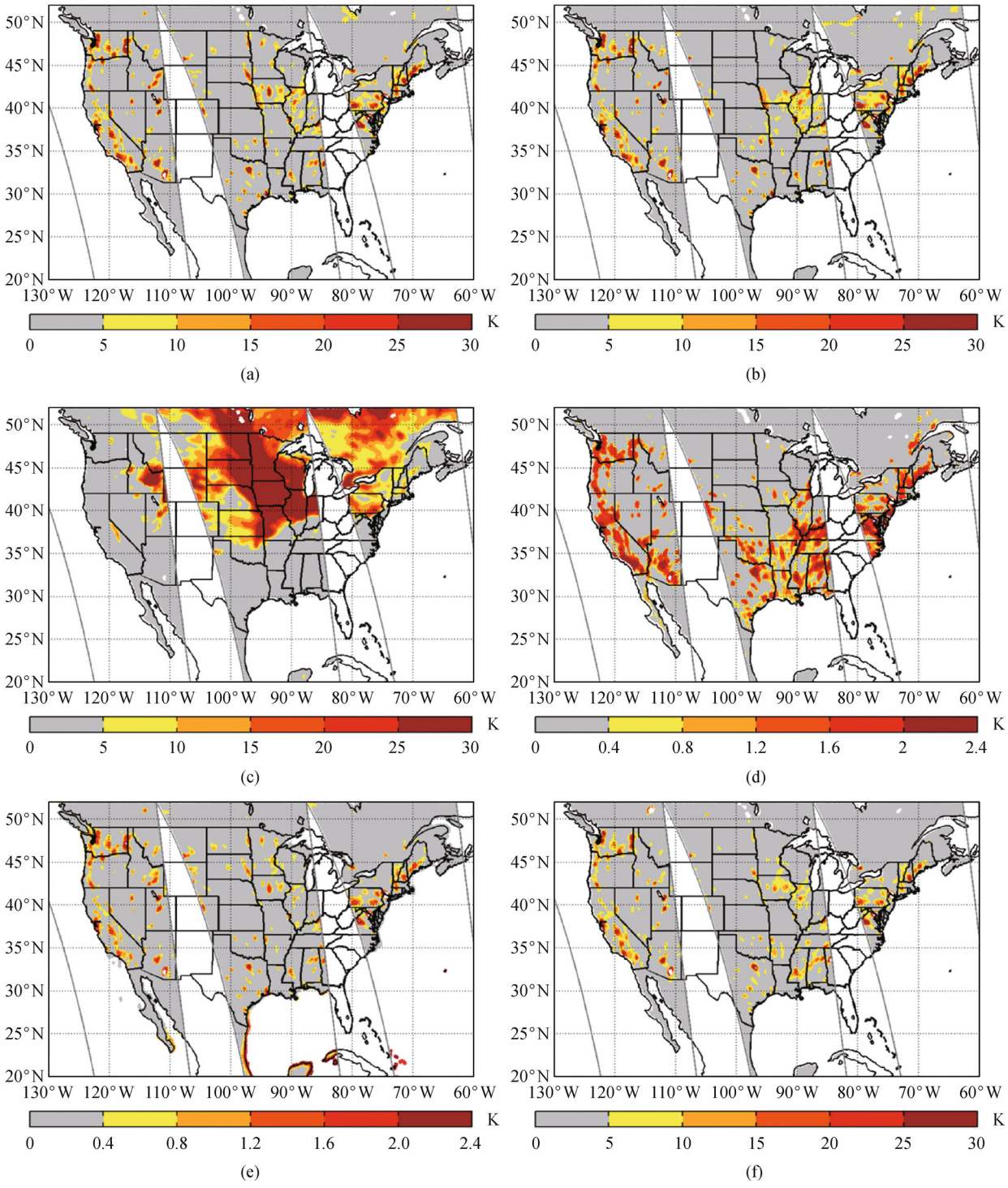


**Fig. 5** RFI distributions at 6.9 GHz with horizontal polarization over the U.S. on 10 February, 2011 identified by (a) our modified PCA method, (b) spectral difference method, (c) PCA method, (d) NPCA method, (e) DPCA method, and (f) our MPCA method on 4 July, 2011.

in gray color. The magnitude and scope of RFI from the MPCA method in winter (Fig. 5(a)), summer (Fig. 5(f)), and from the DPCA method (Fig. 4(e)) are nearly same. The spectral difference method (Fig. 5(b)) excludes directly the RFI sources covered with snow, especially in the states of Iowa, Illinois, and partly Pennsylvania. The RFI identified by the PCA method is from snow, instead of real RFI signals. So the PCA method does not perform well at detecting RFI in winter. The NPCA method uses normalized RFI indices as the original variables to identify RFI. In Fig. 5(d) the scope of the RFI distribution seems to be extensive, and the magnitude is also a little stronger by this method. The feasibility of the NPCA method for detecting RFI mixed with snow is weaker than MPCA and DPCA methods.

The statistical characteristics of the RFI-related principal component from four different PCA methods are shown in Table 4. The RFI-related PC from PCA, NPCA, and DPCA methods is the first PC, but in the MPCA method it is the second PC. The variance and $r^2$ are computed for the RFI-related PC. The variance of the PCA method is the largest with very small $r^2$ (0.0814), because the PCA method identifies the snow as RFI signals. The variance of DPCA is smaller than that of NPCA because the RFI scope detected by DPCA is smaller than that using NPCA (Fig. 5(d) and Fig. 5(e)). The MPCA method owns larger variance and the largest *r*-square (0.8567).

**Table 4** Characteristics of the RFI-related principal component from four different PCA methods

| Principal component | Variance | $r^2$ |
|---|---|---|
| PCA: PC1 | 244.88 | 0.0814 |
| NPCA: PC1 | 0.91 | 0.3072 |
| DPCA: PC1 | 0.46 | 0.2651 |
| MPCA: PC2 | 40.06 | 0.8567 |

Using the proposed MPCA method to identify RFI is also feasible for vertical polarization channels, and the RFI distribution results are very similar (The figures were omitted). The encouraging research results indicate that the RFI identification algorithm is effective. This method can successfully detect RFI not only in AMSR-E summer data but also in winter data.

## 5 Conclusions

The strong scattering signature of dry snow and ice leads to similar characteristics with positive spectral gradient as radio-frequency interference signals. This phenomenon makes the identification of RFI more difficult in the data of the AMSR-E and other similar imagery sensors in winter, especially in regions where RFI is mixed with scattering surfaces. The objective of this study is to find a way to identify RFI over snow covered land efficiently and consistently. We used one RFI Index (RI) and two Scattering Indices (SI) to construct the vector for principal components analysis in our proposed modified PCA method. It appears to be robust in separating RFI from scattering surfaces and time-saving in the computation process. AMSR-E PCA data analysis reveals widespread C-band land RFI in many regions of the U.S. To validate this new method we compare the identified RFI distribution among the spectral difference method, the PCA method, the NPCA method, and the DPCA method. Because the spectral difference method using snow criterion ($TB_{18h}$–$TB_{89h} > 10$ K) firstly excludes the scatter effect of snow, the RFI information mixed together with snow is also removed. The PCA method does not work for winter RFI identification. The magnitude and scope of RFI from our MPCA method are very similar to those from the DPCA method. The range of RFI seems to be extended, and the magnitude also enhanced a little by NPCA. The MPCA method is simple and reliable, which is suitable for both summer and winter seasons.

The new MPCA method is applicable over AMSR-E, but needs further tests before it is used with other satellite-based microwave imagers to detect RFI, such as the WindSat radiometer, MWRI (MicroWave Radiation Imager) uploaded on the Chinese FY-3 meteorological satellite series, and so on. The MPCA method can be applied to satellite radiometer data at the orbit-by-orbit or granule-by-granule data level, offering a real-time RFI detection method before C- and X-band data is delivered to users.

The MPCA method has been tested over U.S., Europe, and Asia, using AMSR-E observations. It performed well at detecting RFI over land. However, it is not optimal for ocean RFI due to the inherent large natural variability of the spectral difference over the ocean. It also does not work over regions with permanent snow and ice such as Greenland and the Antarctic.

**Acknowledgements** This work was jointly supported by the National Natural Science Foundation of China (Grant No. 41175034) and Key University Science Research Project of Jiangsu Province (No. 13KJA170003).

## References

Chaurasia S, Thapliyal P K, Pal P K (2012). Application of a time-series-based methodology for soil moisture estimation from AMSR-E observations over India. IEEE Geosci Remote Sens Lett, 9(5): 818–821

Ellingson S W, Johnson J T (2006). A polarimetric survey of radio-frequency interference in C- and X-Bands in the continental United States using windsat radiometry. IEEE Trans Geosci Rem Sens, 44(3): 540–548

Grody N C (1991). Classification of snow cover and precipitation using

the Special Sensor Microwave Imager. Journal of Geophysical Research: Atmospheres (1984–2012), 96(D4): 7423–7435

Grody N C, Basist A N (1996). Global identification of snowcover using SSM/I measurements. IEEE Trans Geosci Rem Sens, 34(1): 237–249

Guner B, Johnson J T, Niamsuwan N (2007). Time and frequency blanking for radio-frequency interference mitigation in microwave radiometry. IEEE Trans Geosci Rem Sens, 45(11): 3672–3679

Kawanishi T, Sezai T, Ito Y, Imaoka K, Takeshima T, Ishido Y, Shibata A, Miura M, Inahata H, Spencer R W (2003). The advanced microwave scanning radiometer for the earth observing system (AMSR-E), NASDA's contribution to the EOS for global energy and water cycle studies. IEEE Trans Geosci Rem Sens, 41(2): 184–194

Kelly R E, Chang A T, Tsang L, Foster J L (2003). A prototype AMSR-E global snow area and snow depth algorithm. IEEE Trans Geosci Rem Sens, 41(2): 230–242

Kunzi K F, Fisher A D, Staelin D H, Waters J W (1976). Snow and ice surfaces measured by the Nimbus 5 Microwave Spectrometer. J Geophys Res, 81(27): 4965–4980

Lattin J M, Carroll J D, Green P E (2003). Analyzing Multivariate Data. Beijing: China Machine Press, 83–123

Li L, Gaiser P W, Bettenhausen M H, Johnston W (2006). WindSat radio-frequency interference signature and its identification over land and ocean. IEEE Trans Geosci Rem Sens, 44(3): 530–539

Li L, Njoku E G, Im E, Chang P S, Germain K S (2004). A preliminary survey of radio-frequency interference over the U. S. in Aqua AMSR-E Data. IEEE Trans Geosci Rem Sens, 42(2): 380–390

Misra S, Ruf C S (2008). Detection of radio-frequency interference for the Aquarius Radiometer. IEEE Trans Geosci Rem Sens, 46(10): 3123–3128

Njoku E G, Ashcroft P, Chan T K, Li L (2005). Global survey and statistics of radio-frequency interference in AMSR-E land observations. IEEE Trans Geosci Rem Sens, 43(5): 938–947

Njoku E G, Chan T K, Crosson W, Limaye A (2004). Evaluation of the AMSR-E Data Calibration over Land. Italian Journal of Remote Sensing, 30(31): 19–37

Njoku E G, Jackson T J, Lakshmi V, Chan T K, Nghiem S V (2003). Soil moisture Retrieval from AMSR-E. IEEE Trans Geosci Rem Sens, 41 (2): 215–229

Njoku E G, Li L (1999). Retrieval of land surface parameters using passive microwave measurements at 6 –18 GHz. IEEE Trans Geosci Rem Sens, 37(1): 79–93

Piepmeier J R, Mohammed P N, Knuble J J (2008). A double detector for RFI mitigation in microwave radiometers. IEEE Trans Geosci Rem Sens, 46(2): 458–465

Qiu Y, Guo H, Shi J, Kang S, Wang J R, Lemmetyinen J, Jiang L (2010). Analysis between AMSR-E swath brightness temperature and ground snow depth data in winter time over Tibet Plateau, China. IEEE International Geoscience and Remote Sensing Symposium, 2367–2370

Rothrock D A, Thomas D R, Thorndke A S (1988). Principal component analysis of satellite passive microwave data over sea ice. Journal of Geophysical Research: Oceans (1978–2012), 93(C3): 2321–2332

Ruf C S, Gross S M, Misra S (2006). RFI detection and mitigation for microwave radiometry with an agile digital detector. IEEE Trans Geosci Rem Sens, 44(3): 694–706

Weng F, Yan B, Grody N C (2001). A microwave land emissivity model. Journal of Geophysical Research: Atmospheres (1984–2012), 106 (D17): 20115–20123

Wentz F J, Gentemann C, Smith D, Chelton D (2000). Satellite measurements of sea surface temperature through clouds. Science, 288(5467): 847–850

Wu Y, Weng F (2011). Detection and correction of AMSR-E radio-frequency interference. Acta Meteorologica Sinica, 25(5): 669–681

Yang H, Weng F (2011). Error sources in remote sensing of microwave land surface emissivity. IEEE Trans Geosci Rem Sens, 49(9): 3437–3442

Yang H, Weng F, Lv L, Lu N, Liu G, Bai M, Qian Q, He J, Xu H (2011). The FengYun-3 microwave radiation imager on-orbit verification. IEEE Trans Geosci Rem Sens, 49(11): 4552–4560

Zhao J, Zou X, Weng F (2013). WindSat radio-frequency interference signature and its identification over Greenland and Antarctic. IEEE Trans Geosci Rem Sens, 51(9): 4830–4839

Zou X (2012). Introduction to microwave imager radiance observations from polar-orbiting meteorological satellites. Advances in Meteorological Science and Technology, 2(3): 45–50

Zou X, Zhao J, Weng F, Qin Z (2012). Detection of radio-frequency interference signal over land from FY-3B Microwave Radiation Imager (MWRI). IEEE Trans Geosci Rem Sens, 50(12): 4994–5003