This document describes detailed instructions for evaluating the accuracy of the RS-WRF coupled model to make the manuscript more clear and readable.

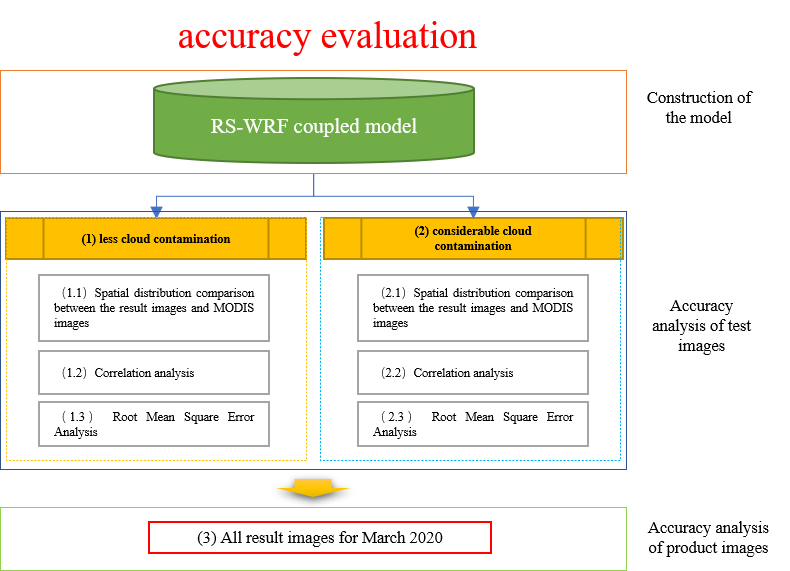


Fig. 1 Accuracy evaluation flowchart

**1. less cloud contamination**

**1.1 Spatial distribution comparison between the result images and MODIS images**

The LST distribution fitted by random forest model (Rf) is the most similar to the spatial distribution of the MODIS LST (Fig. 2). The correlation coefficient (Cor) is the highest, while the root mean square error (RMSE) is the smallest (Fig. 3). Because the images feature some salt and pepper noise, it is necessary to use median filtering (MF) to denoising to achieve optimization effect.

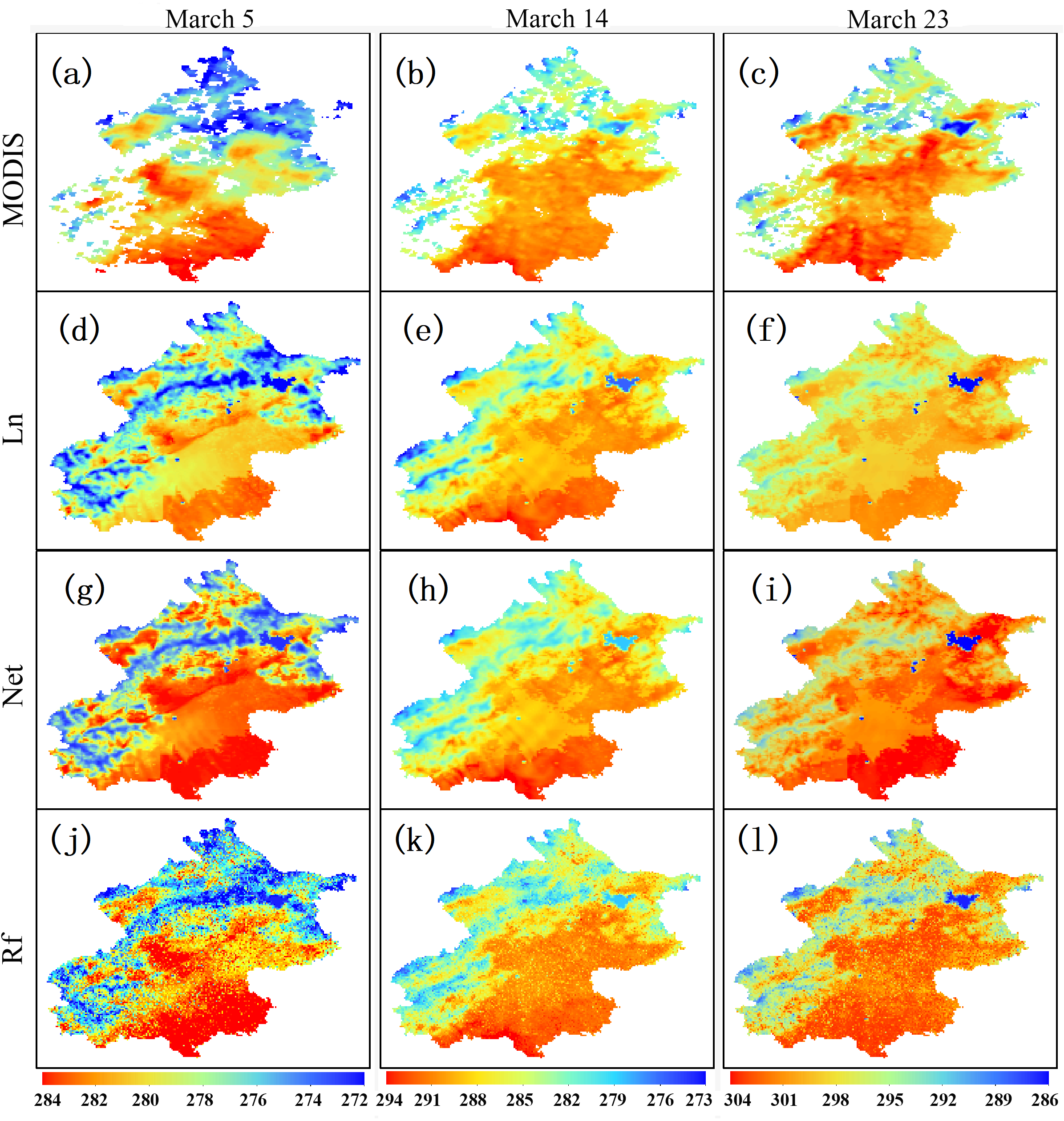


Fig. 2 MODIS LSTs on March 5 (a), March 14 (b) and March 23, 2020 (c). LST images simulated by scheme 6 on March 5 (d), March 14 (e) and March 23 (f) fitted by Ln, March 5 (g), March 14 (h) and March 23 (I) fitted by Net, and on March 5 (j), March 14 (k) and March 23 (l) fitted by Rf. The pixels in the MODIS images (a, b, c) are obtained by masking according to MODIS QA products, and the unit of LST is degrees Celsius (K).

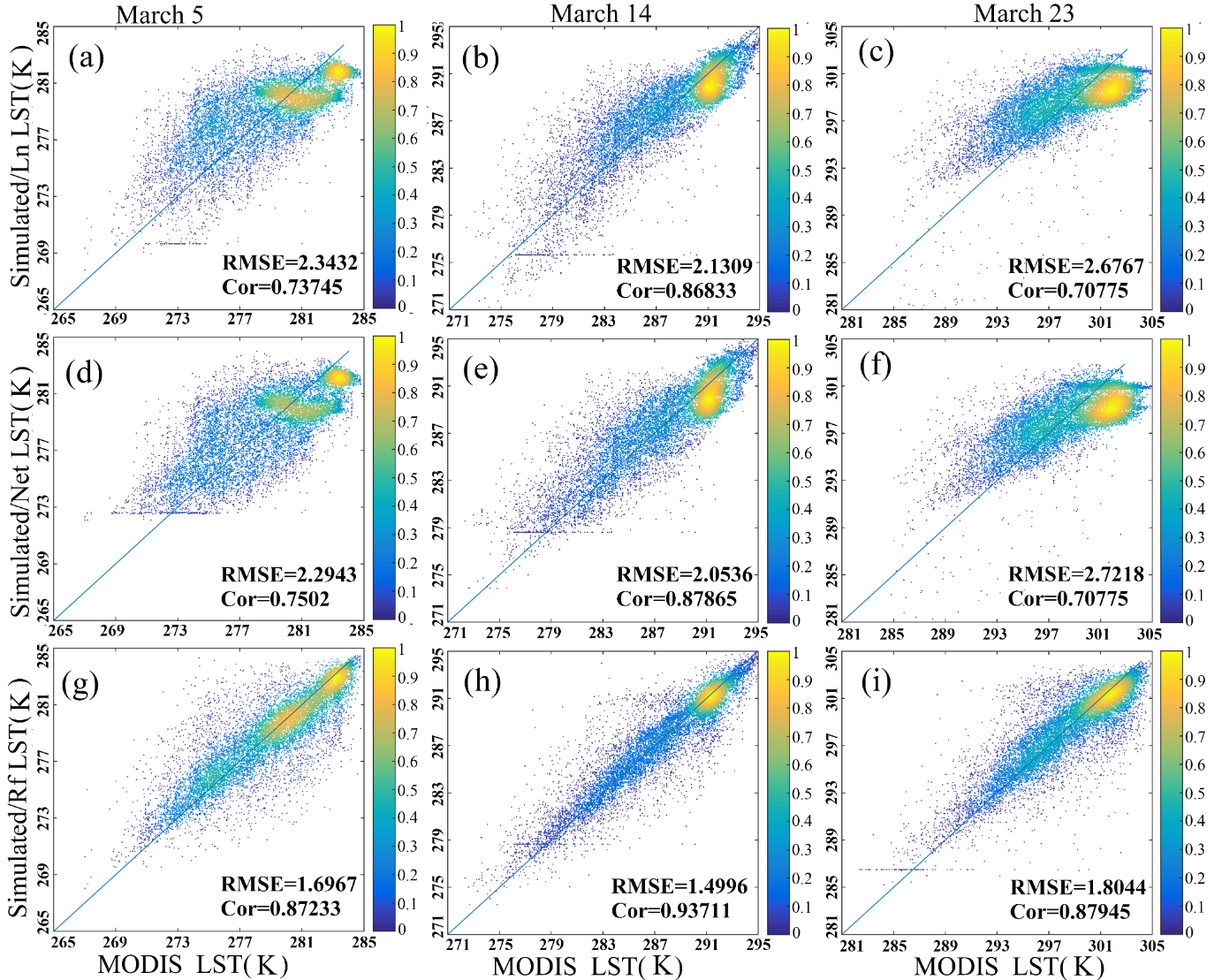


Fig. 3 RMSE and Cor between the MODIS and machine learning LSTs obtained through Ln, Net, and Rf fitting on March 5, 2020 (a, d, g, respectively), March 14, 2020 (b, e, h), and March 23, 2020 (c, f, i).

**1.2 Correlation analysis**

Correlation heat maps (Fig. 4 g, h, i) show that Cor after optimization is always greater than that before optimization. The overall correlation coefficient between the LST images constructed from the RS-WRF coupled model and the MODIS LST images exceeds 0.9.

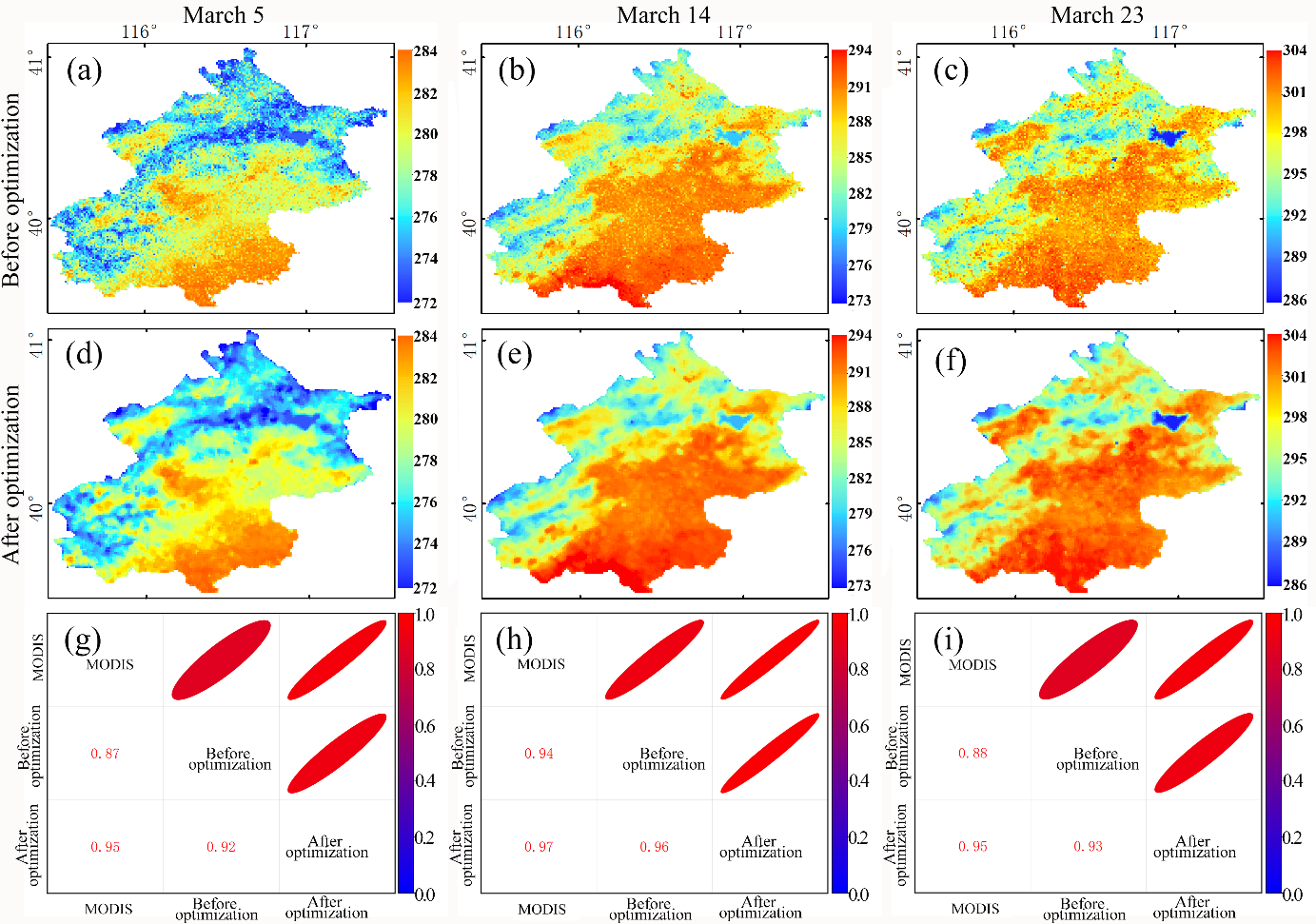


Fig. 4 LST images before optimization (a, b, c) and after optimization (d, e, f) on March 5, 14, and 23, 2020, respectively, and correlation heat maps of the LST before and after optimization (g, h, i)

**1.3 Root Mean Square Error Analysis**

Fig. 5 a shows the overall decrease in the RMSE after optimization. The overall RMSE between the LST images constructed from the RS-WRF coupled model and the MODIS LST images is approximately 1.2 K, and the correlation exceeds 0.9. Hence, when the MODIS images are contaminated by little clouds, the LST can be well constructed.

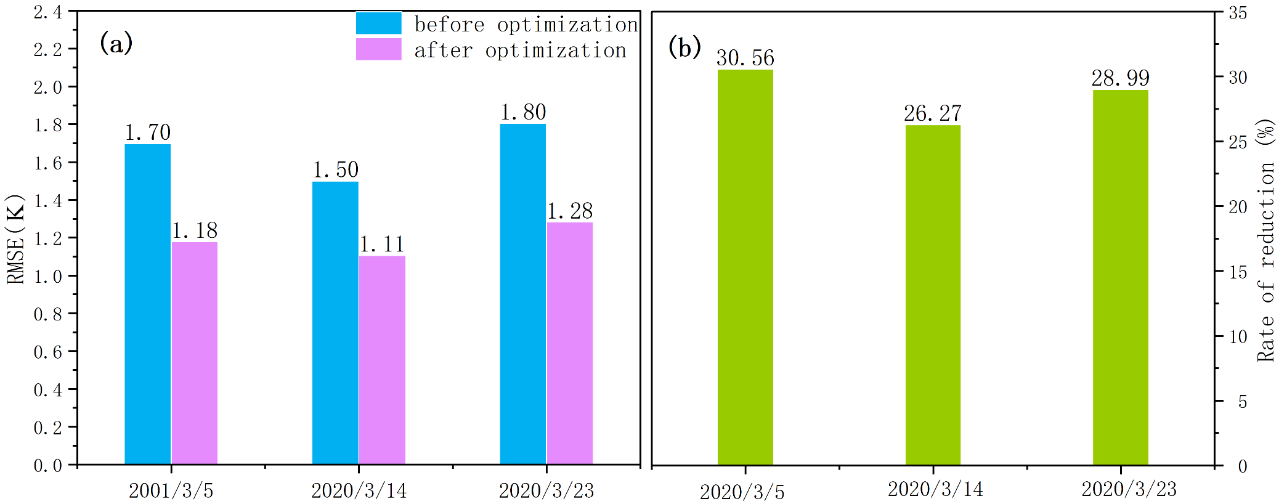


Fig. 5 RMSE values (a) and rates of change (b) before and after LST optimization on March 5, 14 and 23, 2020.

**2. considerable cloud contamination**

**2.1 Spatial distribution comparison between the result images and MODIS images**

Visually, the spatial distributions of the LST in the images after optimization (Fig. 6 e, f) are very similar to those in the MODIS images on the corresponding dates.

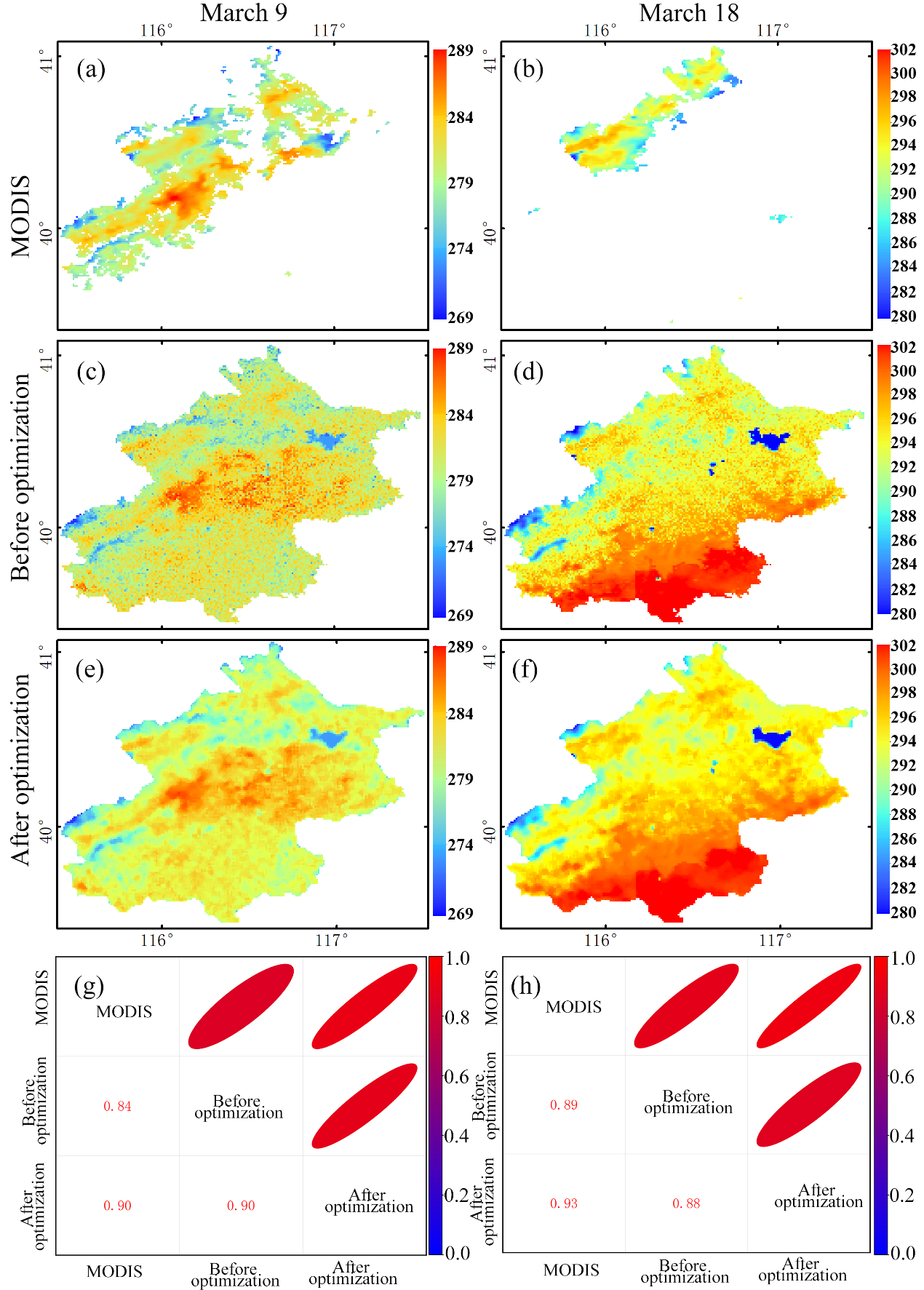


Fig. 6 MODIS images on March 9 (a) and 18 (b), 2020, images before optimization (c, d) and after optimization (e, f) on March 9 and 18, 2020

**2.2 Correlation analysis**

On March 9, 2020, the correlation coefficient increases from 0.84 to 0.90. On March 18, 2020, the correlation coefficient increases from 0.89 to 0.93. Hence, the Cor increases after optimization.

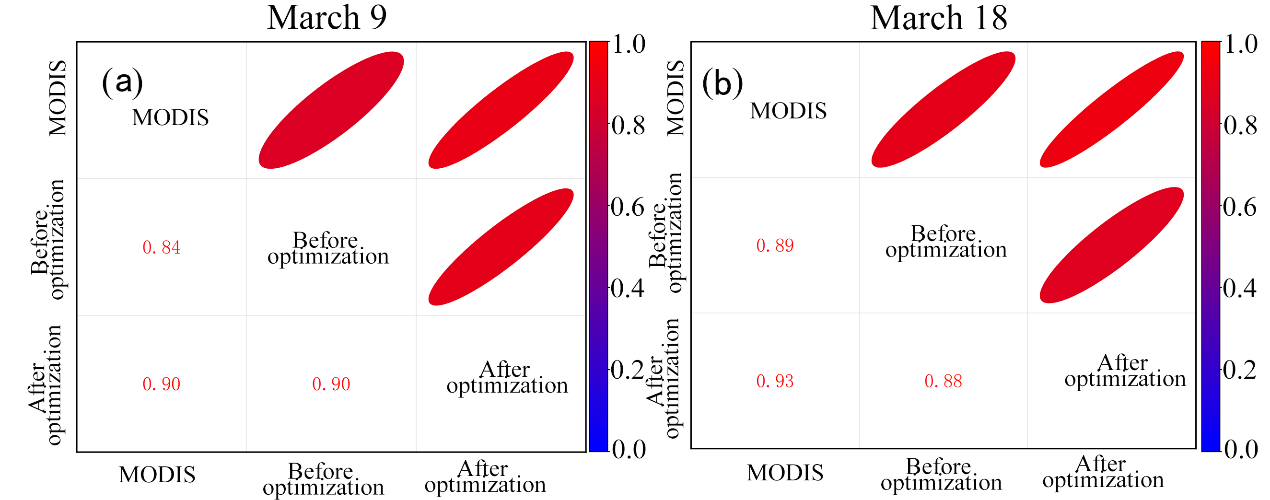


Fig. 7 Correlation heat maps of the LSTs before and after optimization on March 9 (a) and 18 (b).

**3.2 Correlation analysis**

Fig. 8 a shows the overall decrease in the RMSE after optimization. The overall RMSE between the LST images constructed from the RS-WRF coupled model and the MODIS LST images is approximately 1.8 K, and the correlation exceeds 0.9. Hence, when the MODIS images are contaminated by considerable clouds, the RS-WRF coupled model can still construct the LST with high precision.

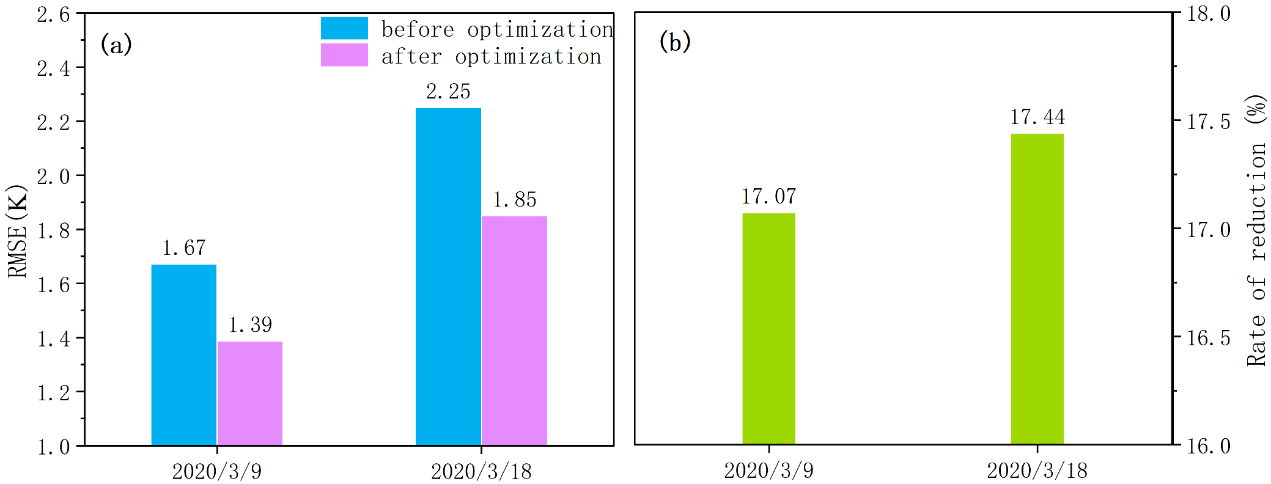


Fig. 8 RMSE values (a) and rates of change (b) before and after optimization on March 9 and 18, 2020.

**3. All result images for March 2020**

All LSTs in March 2020 are constructed, and according to the Cor and RMSE plotted in Fig. 9, the RMSE is less than 2 K, the overall average RMSE is 1.23 K, and the average Cor is 0.93; therefore, our research results are reliable.

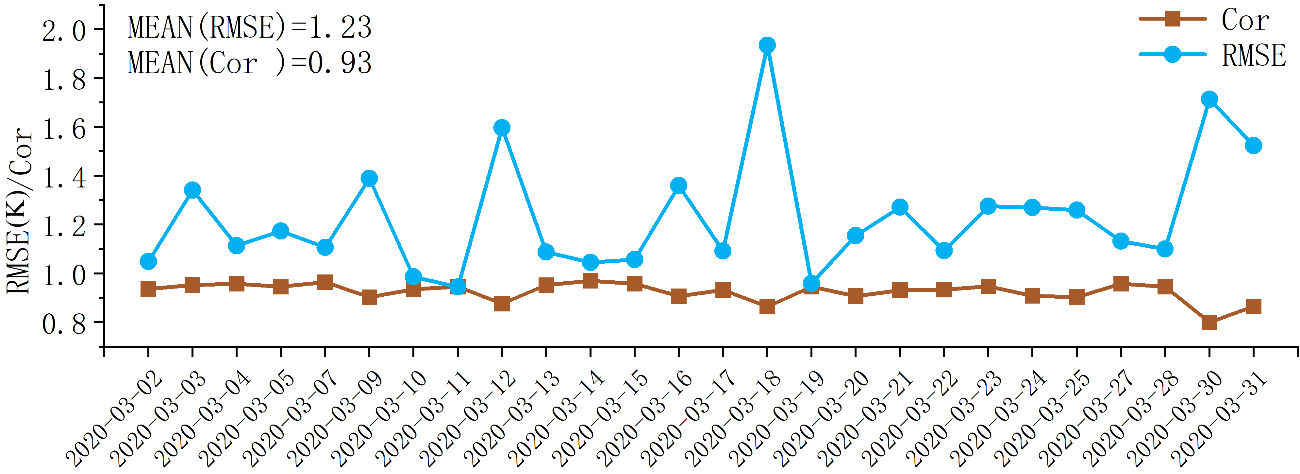


Fig. 9 RMSE and Cor of the LST in March 2020 constructed by the RS-WRF coupled model. Because the data of the WRF model in its early operation phase are unstable, the data from March 1 are excluded. On March 6, 8, 26, and 29, MODIS was completely contaminated by clouds, so the RMSE and Cor could not be calculated, and thus, these 5 days are not included in the figure.