# A Blind Cloud/Shadow Removal Strategy for Multi-Temporal Remote Sensing Images

**Jie Lin**[1],

Ting-Zhu Huang[1], Xi-Le Zhao[1], Meng Ding[1], Yong Chen[2], Tai-Xiang Jiang[3]

1. University of Electronic Science and Technology of China, China
2. Jiangxi Normal University, China
3. Southwestern University of Finance and Economics, China

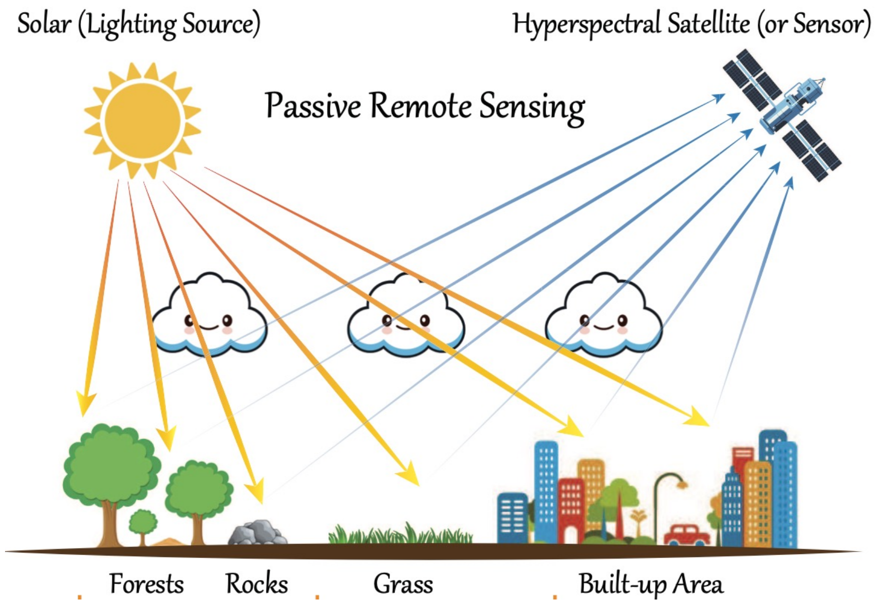
**IGARSS 2021**

# Outline

- Background
- Methodology
- Experiment
- Conclusion

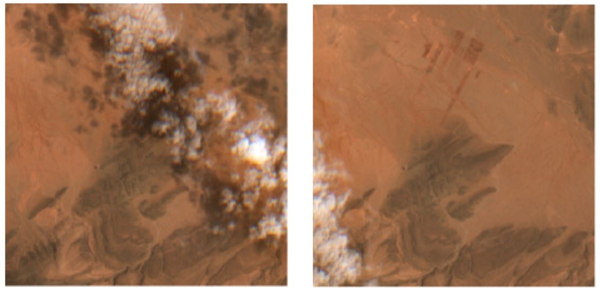
# Background

- Imaging Process[1]

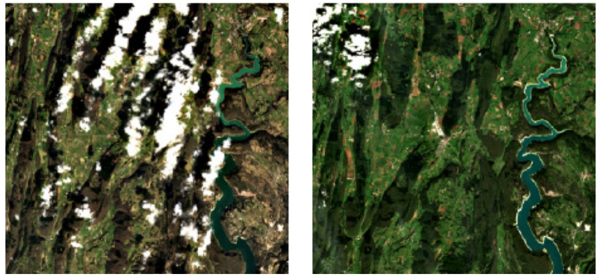



- Observed Remote Sensing (RS) Images

Sentinel-2 MSI



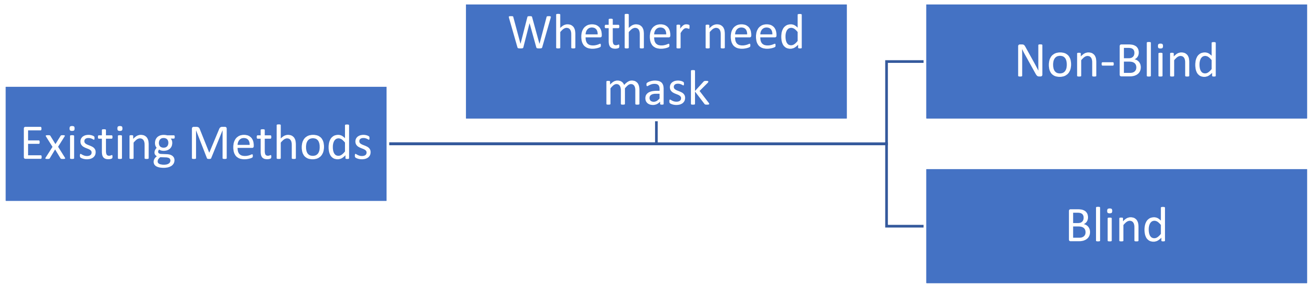
Landsat-8 MSI



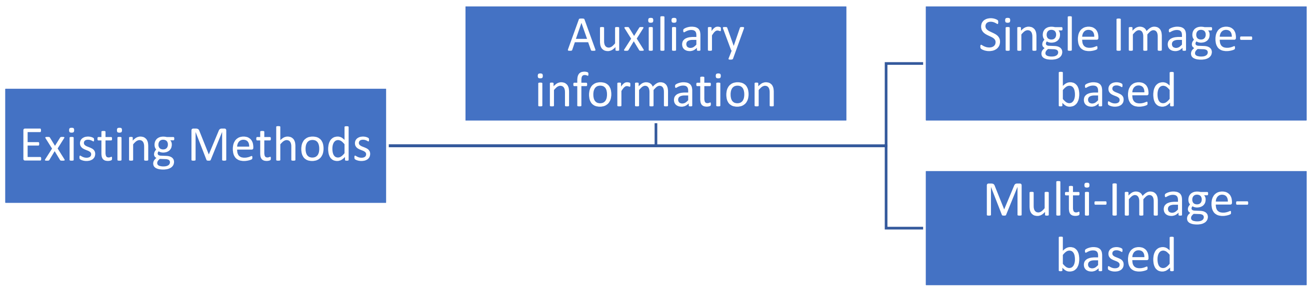
[1] D. Hong *et al.*, "Interpretable Hyperspectral Artificial Intelligence: When nonconvex modeling meets hyperspectral remote sensing," *IEEE Geoscience and Remote Sensing Magazine*, doi: 10.1109/MGRS.2021.3064051.

# Background

- Existing Thick Cloud Removal Methods




- Can achieve a more ideal cloud removal when the mask is inaccurate




- Can handle the large thick cloud contamination

# Outline

- Background
- Methodology
- Experiment
- Conclusion

# Methodology

➢ **Aim:** Thick cloud removal for remote sensing image

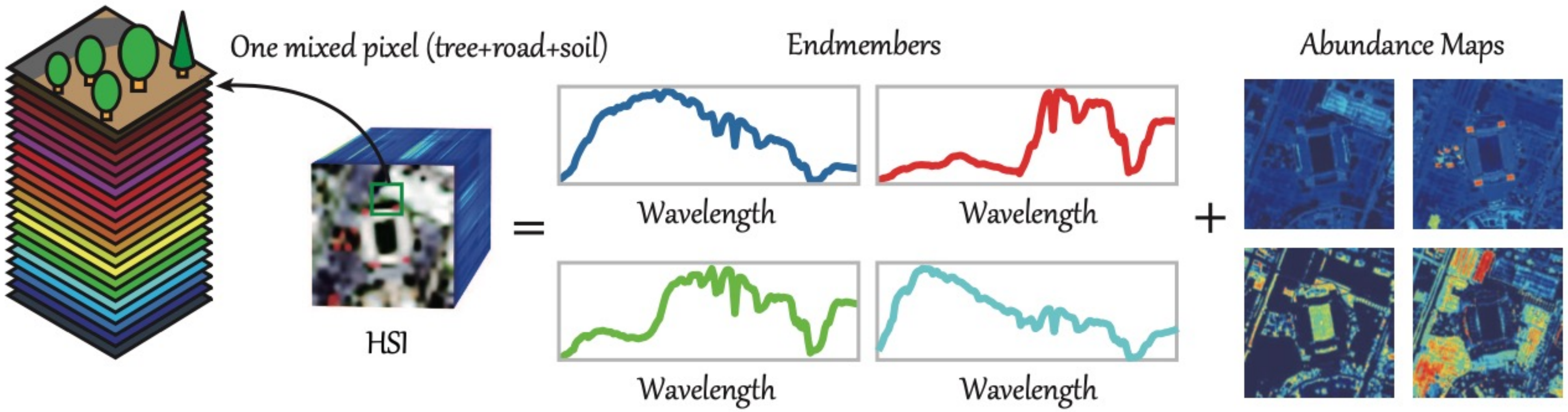
**Blind + Multi-Image-based method**

➢ **Question**

- How to reduce the higher computational complexity due to the introduction of more auxiliary images?
- Is there any latent relationships between the multi-temporal RS images, which can be exploited to finely reconstruct the multi-temporal information?

# Methodology

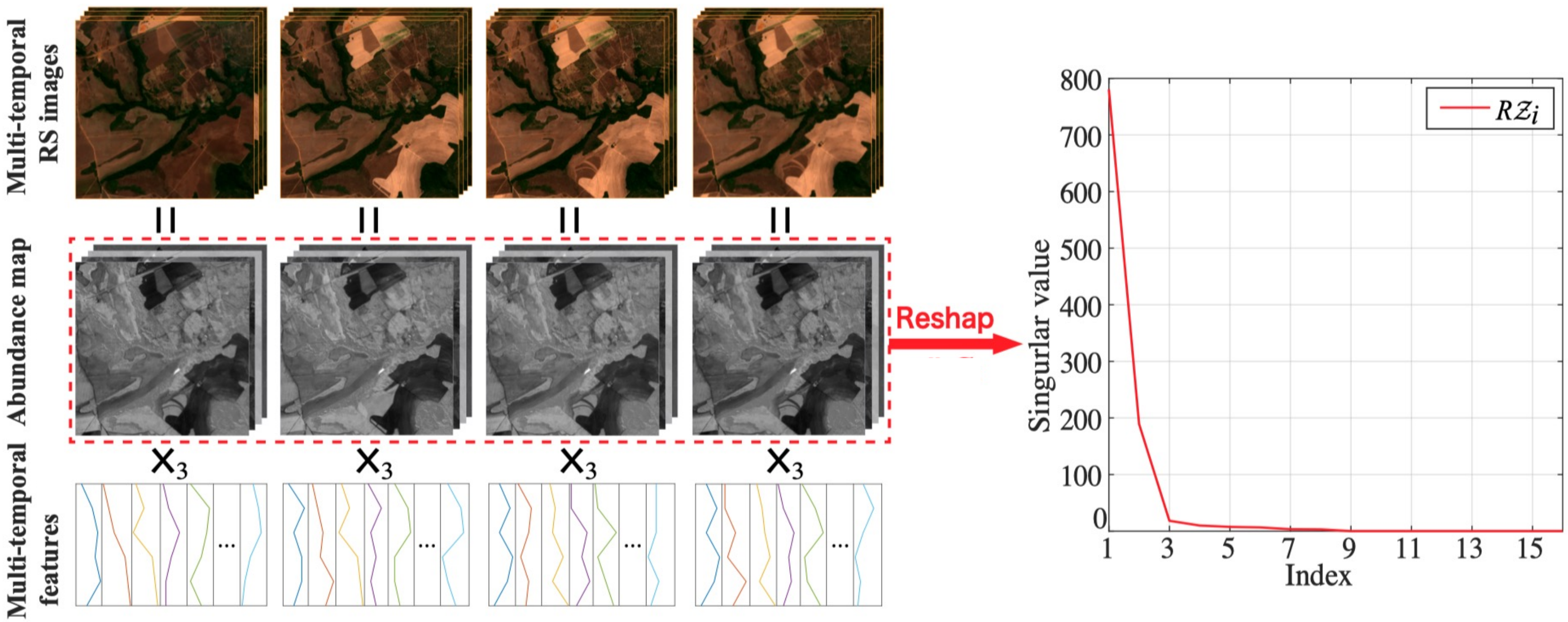
## ➢ Inspiration[1]




Inspired by unmixing , as the distribution of surface material is constant over a period and the same material shows different spectral signatures at different time nodes, the multi-temporal images in the same scene share the same abundances.

[1] D. Hong *et al.*, "Interpretable Hyperspectral Artificial Intelligence: When nonconvex modeling meets hyperspectral remote sensing," in *IEEE Geoscience and Remote Sensing Magazine*, doi: 10.1109/MGRS.2021.3064051.

➢ **Key Observation**




Then, we use a coupled tensor factorization to explore this relationship, which decomposes the image at each time node into an abundance tensor that implies material distribution and orthogonal endmembers. There is a strong similarity between abundance tensors over a period.

# Methodology

## Proposed Method

- Degradation model

$$\mathcal{Y}_i = \mathcal{X}_i + \mathcal{C}_i + \mathcal{E}_i,$$

- Decomposition model

$$\mathcal{X}_i = \mathcal{A}_i \times_3 \mathbf{F}_i,$$

- Proposed Model

$$\min_{\mathcal{X}_i,\mathcal{C}_i,\mathcal{A}_i,\mathbf{F}_i} \sum_i \left\{ \frac{1}{2}\|\mathcal{Y}_i - \mathcal{C}_i - \mathcal{X}_i\|_F^2 + \beta\|\mathcal{C}_i\|_1 \right\} + \alpha\|\mathscr{R}\mathcal{A}_i\|_*,$$
$$\text{s.t.} \quad \mathcal{X}_i = \mathcal{A}_i \times_3 \mathbf{F}_i,\ \mathbf{F}_i^T\mathbf{F}_i = \mathbf{I},\ (\mathcal{X}_i)_{\overline{\Omega}} = (\mathcal{Y}_i)_{\overline{\Omega}}.$$

- Developed ALM algorithm

**Algorithm 1** ALM-based Algorithm for the Proposed Model

**Input:** Observed RS images and Reference RS images $\{\mathcal{Y}_i\}$.

1: Initialize: $\mathcal{X}_i = \mathcal{Y}_i$, $\mathcal{C}_i = \mathcal{P}_i = \mathcal{O}$, and $\mathbf{W}_i = \mathbf{Q}_i = \mathbf{0}$.
2: **while** not converged **do**
3: Update $\{\mathbf{F}_i\}$, $\{\mathcal{A}_i\}$ by (4) and (5);
4: Update $\mathbf{W}$, $\{\mathcal{C}_i\}$, $\{\mathcal{X}_i\}$ by (6), (7), and (8);
5: Update $\{\mathcal{P}_i\}$ and $\mathbf{Q}$ by (9);
6: Refine mask $\Omega$ by Algorithm 2.
7: **end while**

**Output:** Estimated RS images $\{\mathcal{X}_i\}$.

# Methodology

## ➢ Proposed Method

- Mask refinement strategy

**Algorithm 2** Cloud/Shadow Detection

**Input:** Sparse component $\mathcal{C}_i$, threshold $\epsilon$.

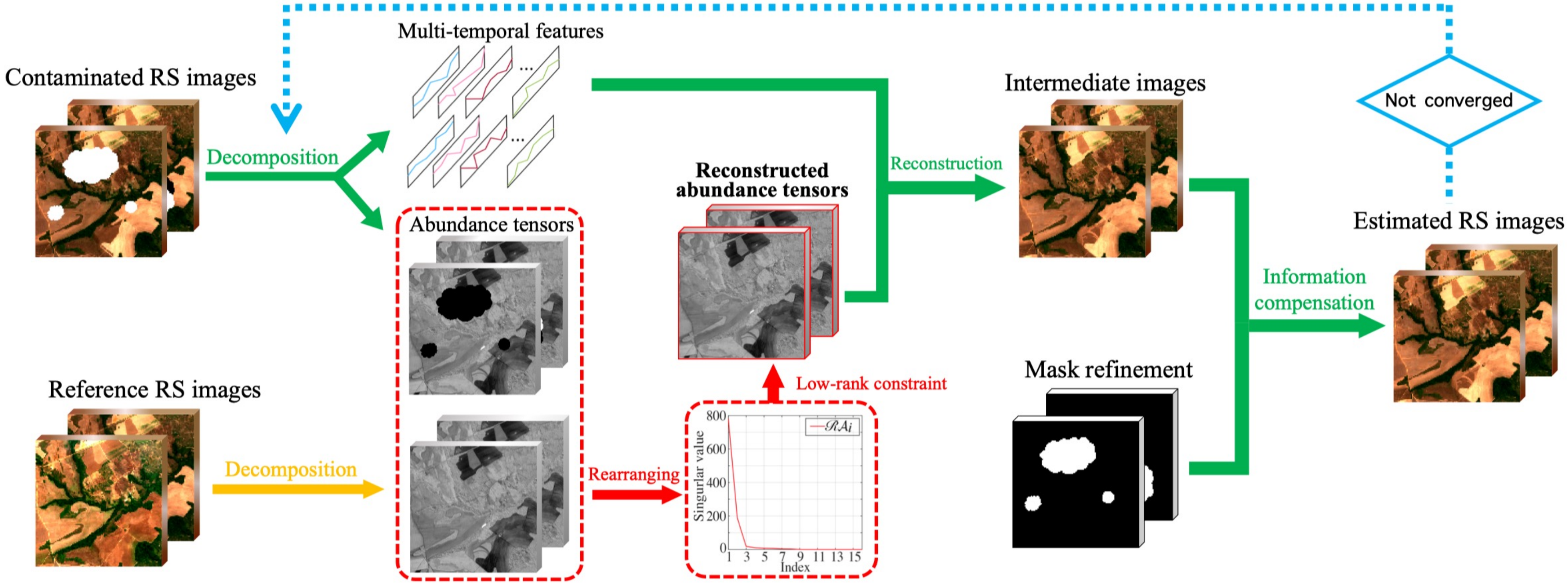
Initialize: $\Omega = \varnothing$.<br>
**for** $p_1 = 1 : m$ **do**<br>
    **for** $p_2 = 1 : n$ **do**<br>
        Compute $a = \text{mean}\left(\mathcal{C}_i(p_1, p_2, :)\right)$;<br>
        $\Omega = \Omega \cup (p_1, p_2, :)$, if $|a| > \epsilon$;<br>
    **end for**<br>
**end for**

**Output:** The location of the cloud/shadow $\Omega$.

We embed the cloud/shadow detection (Algorithm 2) in each iteration of Algorithm 1 to refine the mask. The refined mask will help to introduce true information from observed images for multi-temporal feature learning.

## ➢ Proposed Method

- Flowchart of the Proposed Method

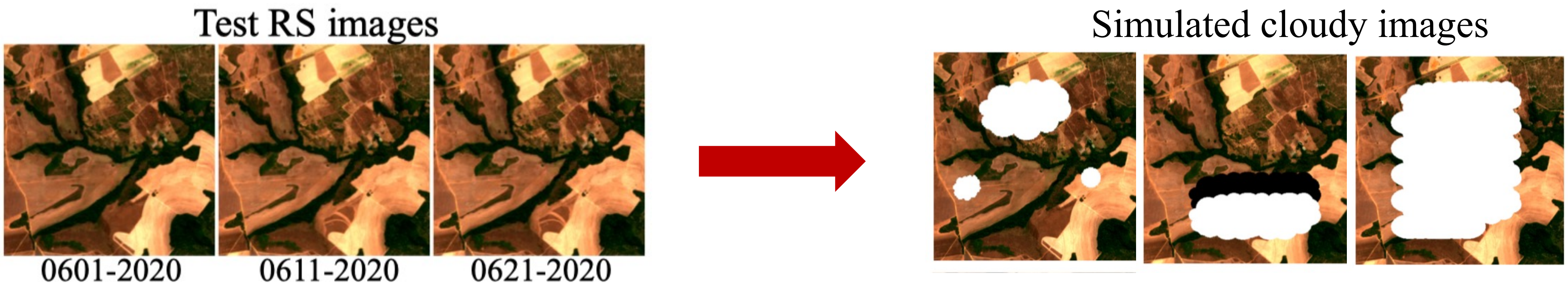


# Outline

- Background
- Methodology
- Experiment
- Conclusion

# Experiment

- Dataset: Brazil Dataset[2]




This dataset is taken over Brazil by Sentinel-2 and each temporal data contains four spectral bands (B2, B3, B4, and B8). The sub-images of size 400 × 400 of 6 different temporal data are used in our experiments.

- Comparison methods

  Non-Blind temporal-based methods: ALM-IPG and WLR
  Blind temporal-based method: TVLRSDC

[2] https://theia.cnes.fr/atdistrib/rocket/#/home

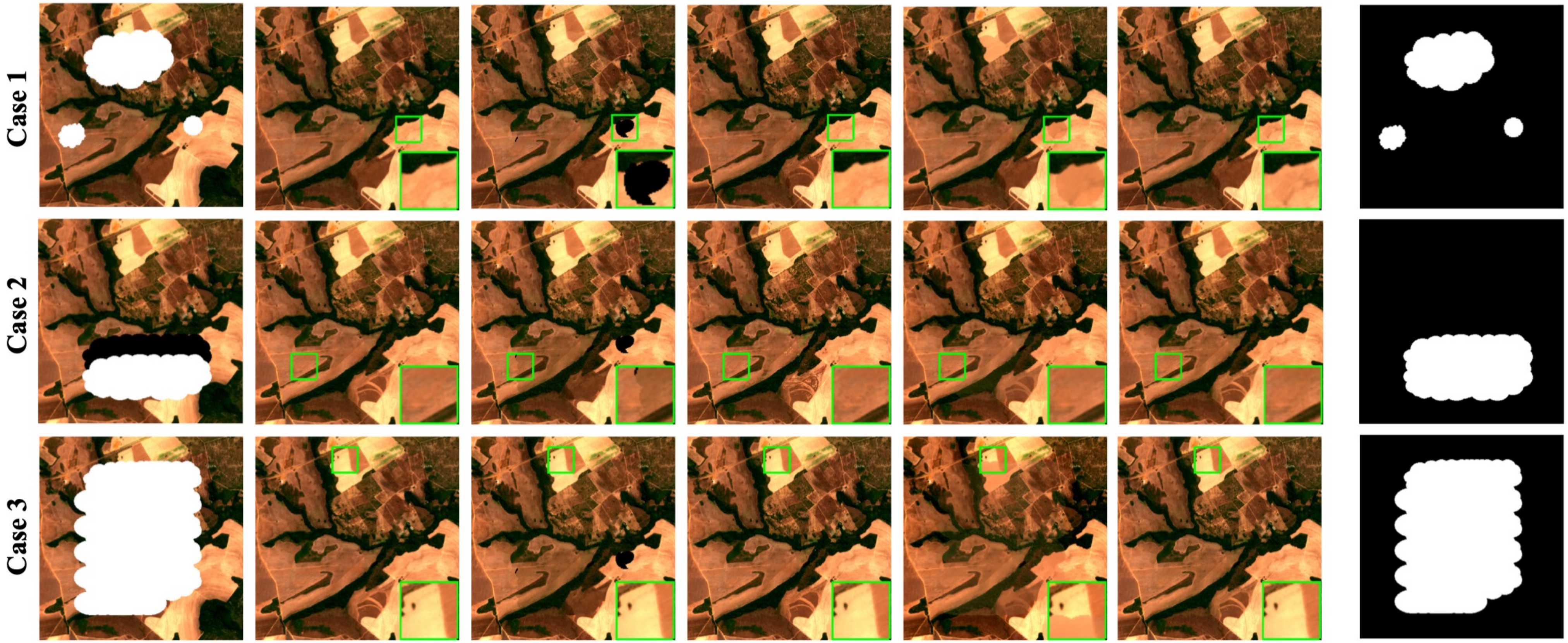
# Experiment

- Index Evaluation

| Case | Index | Observed | ALM-IPG | WLR | TVLRSDC | Proposed |
|---|---|---|---|---|---|---|
| Case 1 | PSNR | 11.921 | 32.164 | 41.420 | 40.578 | **49.264** |
| | SSIM | 0.8500 | 0.9840 | 0.9884 | 0.9947 | **0.9987** |
| | CC | 0.4905 | 0.9202 | 0.9901 | 0.9940 | **0.9987** |
| Case 2 | PSNR | 11.543 | 30.923 | 37.336 | 40.790 | **49.154** |
| | SSIM | 0.7955 | 0.9703 | 0.9714 | 0.9901 | **0.9982** |
| | CC | 0.4512 | 0.8959 | 0.9750 | 0.9902 | **0.9982** |
| Case 3 | PSNR | 6.2239 | 31.168 | 42.735 | 36.780 | **44.670** |
| | SSIM | 0.5580 | 0.9669 | 0.9933 | 0.9664 | **0.9952** |
| | CC | 0.1728 | 0.9010 | 0.9943 | 0.9542 | **0.9956** |
| Time | | — | 0.7528 | 2.0179 | 4.8299 | 1.0934 |

# Experiment

- Visual Evaluation




(a) Observed (b) Cloud-free (c) ALM-IPG (d) WLR (e) TVLRSDC (f) Proposed (g) Cloud/shadow mask

# Outline

- Background
- Methodology
- Experiment
- Conclusion

# Conclusion

- **Key Observation**: There is a strong similarity between the abundance tensors at different times since the material distribution is constant to time in multi-temporal RS images.

- **Contribution:**

  - Based on our key observation, we suggest a cloud/shadow removal model, which faithfully reconstructs the underlying multi-temporal information through reconstructing abundance tensors and learning multi-temporal features.

  - We design a refinement strategy to refine the detected cloud/shadow mask, which helps to learn more accurate multi-temporal features.

# Thanks!

**Jie Lin**
University of Electronic Science and Technology of China