# Response to Referees

We would like to thank the referees for their valuable feedback. The original referee comments are included below with our responses in red.

Based on the feedback, several improvements have been made to the feature clustering and ranking (FCR) method, the evaluation, and the layout of the manuscript. The FCR method was modified to use affinity propagation clustering rather than hierarchical clustering. Affinity propagation produces a more stable grouping of features than hierarchical clustering under different data samplings, and does not require input parameters to determine the number of clusters. The result of this change is that FCR can operate in a fully automatic mode without the need for a user selected clustering threshold.

The maximum-relevance minimum-redundancy (mRMR) criterion was added to the methods used in the comparison and results were re-generated. The accuracy and stability performance of the modified FCR method was improved relative to the other methods.

Changes to the previous version of the manuscript have been marked in the revised version. Insertions are marked in red and deletions in strikethrough red, while layout moves were left unmarked for readability.

Referee 1  
This paper presents a feature selection method for high dimensional remotely sensed data. It mainly uses the average-linkage hierarchical clustering to group the corrected features, and the importance is measured based on the median of a relevance criterion. The manuscript is not written well. The structure is not clear which makes it hard to understand. Some important problems should be considered as follows.

1.The logic of the introduction is not clear, and there are too many different categories according to various classifying criteria.

The introduction has been restructured and shortened where appropriate. Related methods have been separated into section 2.1.

The introduction begins by describing why there is a need for feature selection and then gives a basic background on approaches to feature selection to explain where FCR fits in. Problems commonly encountered when selecting features from high dimensional redundant data are then described, to provide background on the motivation for the FCR method. Lastly, the proposed FCR method is introduced and its contributions are summarised.

The introduction is not comprehensive as well. Some related works should be added as follows.  
“Dual Clustering Based Hyperspectral Band Selection by Contextual Analysis,” IEEE T-GRS, 2016. “Salient Band Selection for Hyperspectral Image Classification via Manifold Ranking,” IEEE T-NNLS, 2016. “Hyperspectral Band Selection by Multi-task Sparsity Pursuit,” IEEE T-GRS, 2015. “Hyperspectral Image Band Selection via Global Optimal Clustering,” IGARSS, 2017.

While the articles listed above represent interesting feature selection methods, we do not believe they are relevant to the problem we are addressing as they are unsupervised, hyperspectral feature selection methods that do not consider the problem of stability. In contrast, FCR is a supervised method for generic remote sensing data that addresses the problem of instability resulting from redundant data. We hope that the referees will agree.

2.The structure of the whole manuscript is not clear. It is really hard to understand the motivations of the proposed method.

Some structural changes have been made to address this comment and motivations for the proposed method have been clarified at the end of the introduction.

It is better to reorganize the structure, and put the experimental figures and tables under the corresponding manuscript description.

Unfortunately, the ScholarOne online system for uploading manuscripts does not allow figures and tables to be inserted in the relevant sections. One has to upload these separately and have them included at the end of the manuscript.

3.One important motivation of the proposed is taking the computation time and measurement cost into consideration when selecting features. In the detailed algorithm implementation, how to select bands based on the computation time?

The proposed method can function automatically where the most relevant features are selected from the top ranked clusters without user intervention. Optionally, factors like computation time and measurement cost can be considered in a manual procedure. In this scheme, the user would choose features from the top-ranked clusters that achieve the best trade-off between the feature relevance and measurement cost (or other criterion) values. Means of automating this multiple-criterion mode will be investigated in future. Details of these options have been added to section 2.2 and are referred to again in the conclusion.

4.The dissimilarity threshold is an important parameter which is determined by visual inspection dendrogram for each data set. This parameter selection is subjective and not stable. Therefore, how does this parameter affect the stability of the whole method? Is the defined stability meaningful when the proposed method involves an unstable parameter?

With the change from hierarchical clustering to affinity propagation, the dissimilarity threshold is no longer necessary, and the clustering process requires no user input.

5.Some grammar mistakes should be corrected. The reviewer just illustrates some examples. Please recheck the whole manuscript.   
On page 4 of 33, Line 7, some criterion...  
On page 4 of 33, Line 12, a variety of search methods exists...  
On page 4 of 33, Line 14, Of these, only the branch and bound method is globally optimal, the rest achieve reduced computation at the price of optimality.

The manuscript has been re-checked. Thank you for bringing this to our attention.

Referee 2  
Comments to the Author  
This paper presents a new feature selection method based on feature clustering and ranking for high-dimensional remote sensing data classification. The description of the proposed method is not sufficient. The contribution of this paper is not enough. The specific comments are as follow:  
1) In the Introduction, the review of existing feature selection methods are poorly organized. Moreover, the main contribution of this paper should be added and clearer.

The introduction has been restructured and edited for clarity. Related methods have now been moved to section 2.1. The motivation for, and contribution of the method is described at the end of the introduction. See also the response to comment 1 from referee 1 above.

2) In 2.1, the introduction of the proposed method is too short. Some more detailed procedure of the proposed method should be added.

Detailed descriptions of the clustering procedure, cluster importance measure and feature selection options have been added to section 2.2.

3) In the experiment, the analysis of selected features for different methods may be investigated.

A forward selection search with maximum-relevance minimum-redundancy (mRMR) criterion has been added to the experiment for comparison to FCR. The mRMR criterion is widely used and addresses the problem of redundancy and resultant instability. We believe that the combination of methods included in the comparison provides a competitive benchmark against which to compare FCR.