This paper presents a Robust and Reduced Rank Spatial Kriging Model (R3-SKM), which is resilient to the impact of outliers and supports fast spatial prediction. Differing from other work, the R3-SKM method used a knot-based technique to select a number of knot locations whose size is much less than the size of all observations. The application of the reduced rank technique improves the time efficiency in a linear order. The key ideas of the paper include reduced rank spatial kriging model, and the parameter estimation using Baysian framework. In my opinion, the novel ideas include the application of knot-based technique, and Moran's I-statistic to capture the spatial dependency of the observations.

The strong points are in the following. First, the method in the paper uses a reduced rank method to reduce the number of observations which are needed in the spatial kriging model. The application of this reduced rank technique significantly improves the efficiency of the SKM method in a linear time complexity, especially when the dataset has a large size. Second, it designs an approximate algorithm for robust parameter estimation. The application of the iterated re-weighted least squares algorithm is a contribution of the paper, because the convergence rate has been proved theoretically in linear time, which is more suitable than other approximation method, such as expectation propagation (EM). Finally, the measurement error is modeled using heavy tailed distributions which lead to the robustness of the proposed method (R3-SKM).

The weak points are as follows. First, the type of outliers addressed is not presented in more details in this paper, such as clusters of outliers or dispersed distributed outlier points. These two distinct types of outliers may induce different results. For example, in the Yang’s paper [1], the experiment result shows that the method performs well for dispersed distributed outlier points, but not well for outlier clusters. From the context in this paper, I guess that they only handle the second situation which is the dispersed distributed outlier points. Second, although the paper uses a kind of reduced rank technology, named as knot-based model, to improve the running efficiency of their method, they did not discuss how much the accuracy changes or degrades compared to the same method which does not reduce the number of observations. Third, the paper did not prove the optimal number of knots theoretically. I guess, in fact, the number of knots may have relationships with the distribution and other characteristics of the observations. Finally, when the paper explains why the SKM method (a baseline in the comparison experiment) did not perform as good as the regression model, the reason is that the hidden spatial-temporal process captures spatial association which is assumed to be a multivariate Gaussian process. But as far as I am concerned, the hidden spatial-temporal process in the proposed method also follows a multivariate Gaussian distribution. Then the true reason of SKM model should be that the measurement error is modeled by a heavy tailed distribution.

As a research discussion point, my opinion is that first, the spatial kriging model can be used together with some temporal model, such as Kalman filtering to make the robust prediction and outlier detection for spatial-temporal datasets. Second, I do not understand that, in the experimental dataset which has a high spatial dependency, why the SKM performs still no better than the linear regression method. I think the SKM should be better because it captures the spatial dependency in the spatial dataset and the advantage of this characteristic should be obvious in the spatial dataset with high spatial dependencies.

[1] Student-t based Robust Spatio-Temporal Prediction, ICDM 2012