Interactive comment on “Regularization methods for the combination of heterogeneous observations using spherical radial basis functions” by Qing Liu et al.

Anonymous Referee #2

Received and published: 15 May 2019

General comments: The authors present a regularization method that is used in the combination of heterogeneous data. The novelty is the combination of two regularization methods, namely by combining VCE estimation and the L-curve criteria. Various combinations are discussed and compared to existing methods, primarily to VCE estimation or L-curve regularization alone. However, the applied methodology is questionable as the combination of the two criteria is essentially equivalent to a double regularization. The comparison to the calculation "VCE based on CM2" reveals that the same result can be achieved by ordinary approaches. Further, the usage of the Shannon function for analysis and Blackman/CuP function results in additional smoothing which has not been further explained or described. Thus, I consider this paper inconclusive.

Specific comments: Section 1: The motivation of the regularization is unclear. Why are new methods needed? What are the limitations of existing methods? Why is the specific approach of the authors chosen and what benefits do the authors expect from their approach?

page 3, line 5: The authors argue to find the best-performing method (in what sense?) for regularization. However, they do not consider other methods than VCE and L-curve, e.g. GCV. Further, the method will be best-performing for their specific problem as no general criteria is derived which allows to conclude that the proposed method is best-performing.

Section 2.3: The authors present three different SRBFs with various smoothing features. Why is the approach of Eicker (2008) not considered? By including gravity field information into Bn, a considerable improvement can be achieved.

Section 2.3: If I understood the author's approach correctly, they use the Shannon function for the analysis of the simulated data but apply the estimated coefficients using either the Blackman or CuP function in the synthesis step. This approach is at least odd and inconsistent if not wrong from the beginning. In fact, the approach introduces an additional smoothing. The authors state correctly that the latter two have smoothing features. Thus, the approach is unsuitable for the conducted research as it masks the effects of the regularization. It is another implicit regularization and thus the results cannot unambiguously assigned to the performance of the chosen methods. The only correct approach is therefore to use the same function for the analysis and synthesis step. The approach is even more questionable as Bentel2013 showed that differences between SRBFs matter (as also stated by the authors).

Section 3.1 provides no new information. The content can be reduced to the most significant equations and appropriate referencing.

Section 3.2: CM1 can obviously be removed as the assumption $\sigma_1^2 = \sigma_2^2 = ...$ is hardly valid in any case (except for simulated data with exactly
this assumption). Furthermore, applying VCE is the proper tool to consider data with varying variance factors. Thus, the results of CM1 are superfluous and the results prove the invalidity of the assumption.

Section 4.3: The regularization is essentially a double differentiation as the estimated variance factors during the VCE will reflect the regularization parameters. Practically the \( \lambda \) of equation 30 is split in \( \lambda_1 + \lambda_2 \) where one is estimated by VCE and the other by the L-curve criterion or vice-versa. Due to the double regularization, the results will be further smoothed than in case of applying just one of the methods alone. A better fit is therefore expected as the inherent effects due to ill-posedness is dominating. Also, the authors do not motivate the need for a second regularization and also do not discuss the effect of the second regularization step.

Section 5.3: The authors present two study cases: A and F; why not naming them A and B as you only present results of those two. The reader will have no information on cases B to E. Further, the results of CuP function can also be removed as they do not introduce any new insight.