**Geoscientific
Model Development
Discussions**

Interactive comment on "Technical Note: Improving computational efficiency in large linear inverse problems: an example from carbon dioxide flux estimation" by V. Yadav and A. M. Michalak

Anonymous Referee #1

Received and published: 14 December 2012

The technical note "Improving Computational Efficiency in Large Linear Inverse Problems: an Example from Carbon Dioxide Flux Estimation" is a well written paper that does fall perfectly within the scope of the GMD. It proposes two new algorithms for fast computation of large linear inverse problems.

The abstract is well written and it summarises the paper, focusing on the conclusions. The overall presentation is well structured, with a clear separation between sections 2 and 3. The method in section 3 can use the multiplication method described in section 2, allowing the reader to understand the methods in sequence. The English is excellent, this reviewer didn't find any typos of grammar problems. The formulae is well

presented, and the symbols are coherent and explained in the text. References are appropriate in both number and quality. The authors produced a good literature review and gave the appropriate credit to the cited authors. The supplementary material is also very good, and it allows the reader to play with the implementation of the algorithm in Matlab.

The first algorithm is based on the fact that one of the matrices involved in the problem can be factored as a Kronecker product, so the matrices can be computed in blocks. The second algorithm relies on the fact that there is no need for the explicit calculation of the a posteriori covariance matrix for this particular problem. The methods are described very clearly, and the results are sufficient for a technical note, and enough information is given to allow readers to re-implement them.

The claim in section 2 that the savings regarding Gourdji et al. (2012) would be 99.9 % needs caution. It would be better to phrase it "... a savings of over 99.9% in the number of floating point operations..." rather than "... a savings of over 99.9% in the computational cost...". Runtime will often depend on other factors such as hardware and memory usage. Some phenomena related to the operating systems can also play an important role in the runtime, e.g., memory contention. It would be nice to have empirical evaluation on at least one particular hardware set, in order to compare the proposed approach with the traditional one. The same idea is also relevant to section 3, where the claim is a savings of over 99.9999%.

The two algorithms can yield a substantial improvement in the runtime for large problem sizes, with a relatively simple implementation. For more practical purposes, it would be very useful to have an additional section or an appendix with some real runtime measurements. Moreover, it would be very interesting to see an empirical comparison between the proposed approach implemented in parallel systems and some well established libraries used in parallel computation of matrices, using the same problem size used as an example in the article. Perhaps this could be part of future work.

Summing up, this paper makes an original contribution to the efficiency improvement of very large inverse problems.