**Interactive comment on** “Estimating surface CO\textsubscript{2} fluxes from space-borne CO\textsubscript{2} dry air mole fraction observations using an ensemble Kalman Filter” *by* L. Feng et al.

**Anonymous Referee #2**

Received and published: 9 January 2009

This paper presents a new scheme to infer the surface fluxes of CO2 from atmospheric concentration measurements. It is based on the ensemble Kalman filter (EnKF) principle. The authors apply it to conduct several observation system simulation experiments with OCO-like observations of XCO\textsubscript{2}. These experiments differ from each other for the observation density, for the probability density function of the observation uncertainty and for some internal parameters of the ensemble filter. The paper is well-written and interesting. It describes a system that may significantly contribute to the interpretation of the OCO and GOSAT observations. However, a number of major issues should be addressed before publication in Atmos. Chem. Phys. I list them hereafter.
• In order to save computing time, the authors have reduced the dimension of the inference problem so much that one may wonder why the ensemble gain matrix (Eq. 6) is used in place of the true gain matrix (Eq. 3). The classical analytical computation would be more accurate and faster. An ensemble approach could be justified if the concentration error statistics were cycled from one window to the next, but there is no cycle here. Instead overlapping windows are used.

• It is said in lines 229 and 351 that the paper evaluates the EnKF approach, but this is hardly done. Given the small size of the problem, a proper evaluation would involve comparing the reduced-rank statistics with the full-rank ones. Doing it is simple, as explained in point 1, and would greatly benefit the paper.

• The state vector is conveniently small in the presented EnKF, but no physical arguments are given in favor of this choice. It is just shown (section 4.5) that it has a strong impact. Can the prior error statistics be safely represented over the globe with just 144 regions?

In addition, I wish to list a series of minor issues that can easily be fixed:

• l.25: the adjective ‘quantitative’ is too optimistic to be used in the statement.

• l.28: most of the ‘in situ measurements’ have actually been made in laboratories from air samples. ‘Surface measurements’ would be more appropriate.

• l.89: the spatial resolution should be given with the latitude and the longitude increments given explicitly.

• l.99-100: the authors should comment their choice.

• l.129-135: this paragraph should be linked somehow with the information given in line 156 about the inclusion of model errors.
• l.161: One may think that correlations between successive observations would not give a block structure to the R matrix, but rather a more general band one. Or are the correlations truncated?

• l.163: it may be relevant to mention the influence of the initial state of CO2 concentrations that is not seem to be part of the state vector, even though it is uncertain.

• l.171: it should be mentioned that the length of the window is discussed later in Section 4.6 (actually the results of Section 4.6 could be more detailed).

• Eq. 4: the formula actually depends on the flux unit and on some space-time correlation structure. Both should be given.

• l.184: how does this statement relate to Eq. 14, that calculates H?

• l.190: 'advantage' with respect to what?

• l.211: I understand that the n-dimensional pdf is approximated by an ensemble of size n. The authors seem not to be hit by the curse of dimensionality.

• l.254-255: is the estimated error reduction realistic?

• l.264: 'small' a priori errors compared to what?

• l.282: 'greater' is excessive.

• l.284: 'obscured AODs' are obscure to me.

• l.298: 'geological' is awkward

• l.313: this result may depend on the state vector resolution.

• l.315: 'error' should be inserted after 'measurements'.

S10358
• l.326: 'weaker' than what?

• Section 4.6: what about increasing the size of the ensemble?

• l.357: ... but Eq. (14) is a linearized model

• l.360-275: lengthy summary

• l.376-383: these results were obvious beforehand, they are not brought by the paper.

• l.389, end of sentence: it may be relevant to add that other biases may have more impact.

• l.475 and 484: typos

• l.489: it is said before that Ne = Nx

• Fig. 6: do the vertical lines represent prior errors or posterior statistics? Also the meaning of 'error' is not obvious (it does not refer to the error statistics but to an error realization).

Interactive comment on Atmos. Chem. Phys. Discuss., 8, 19917, 2008.