The authors would like to thank the Referee for reviewing our manuscript and for providing the authors with the constructive remarks and recommendations, which we have found to be very enlightening. We address the comments in the revised version of our manuscript. Here is a list of our preliminary responses to the Referee comments:

1. It is not clear whether the current formulation of the data assimilation system may not be simply incorporated into the previous case of initial condition estimation through a state augmentation procedure for joint state and parameter estimation, see for example, Dee (2005), Smith et al. (2013). As such, the authors should clearly state the need for the re-derivation of several equations presented here and whether the current context may not by reduced to a previously developed theory through an appropriate change in notation.

The main difference between Dee (2005), Smith et al. (2013) and our paper is that we consider a dynamic formulation of variational data assimilation problem in a continuous form. Therefore, the presented equations can not be derived directly from Dee (2005), Smith et al. (2013).
The presented sensitivity analysis formulas do not follow from our previous results for the initial condition problem (Shutyaev et al., 2017). Of course, the initial condition function may be considered as a parameter, however, in our dynamic formulation we have 2 equations for the model: one equation for describing an evolution of the model operator (involving parameters such as right-hand sides, coefficients, boundary conditions etc.), and another equation is considered as an initial condition.
We also can consider joint state and parameter estimation problem, and it will be some generalization of this paper and the previous one (Shutyaev et al., 2017). In this case we need to introduce an additional term related to the initial condition into the cost function (2.2) to find simultaneously u and λ. The optimality system (2.8)-(2.10) will be supplemented by an additional equation related to the gradient of the cost function with respect to u. In this case, the Hessian is a 2x2 operator-matrix, and all the derivations are more complicated, cumbersome and lengthy. Of course, we can do this, but we decided to present here only parameter estimation case, because this is the case we deal with at our Institute for numerical experiments to find the heat fluxes for the Baltic Sea thermodynamic model. In our experiments, the initial condition is supposed to be known and taken from the run of the model on the previous time step.

2. The significance of the numerical results is only briefly discussed and it appears that the sole purpose of the experiments is to illustrate the practical ability to evaluate the observation sensitivity in a non-trivial application. Little can be learned from these results and, in particular, important practical issues need further clarification. For example, the observation sensitivity calculations are derived from the first order optimality system however, in practice, only an approximate solution to the minimization problem is obtained through an iterative procedure. As such, solving the continuous sensitivity equations may result in inconsistencies between the optimization process and the observation sensitivity calculations. It is not clear what approach has been adopted here: discretize-then-optimize or optimize-then-discretize? Some practical issues regarding the accuracy of the sensitivity estimates should be discussed in the manuscript.

In the revised version of the paper we give more details concerning the numerical experiments. We use the discretize-then-optimize approach, and for numerical experiments all the presented equations are understood in a discrete form, as finite-dimensional analogues of the corresponding problems, obtained after approximation. This allows us to consider model equations as a perfect model, with no approximation errors. Therefore, the accuracy of the sensitivity estimates given by the algorithm (5.16)-(5.18) are determined by the accuracy of solving the Hessian equation $\mathcal{H}\chi = \Phi$ (step 2 of the Algorithm). Due to (5.9)-(5.11), this equation is equivalent to a linear data assimilation problem, and an approximate solution to the minimization problem is obtained by an iterative procedure.

3. In my opinion, the manuscript will benefit from the insertion of a proof-of-concept with a simple model and numerical results using an easily reproducible assimilation setup where several practical aspects can be investigated and illustrated.

In the revised version of the paper we include a proof-of-concept analytic example with a simple model to demonstrate how the sensitivity analysis algorithm (5.16)-(5.18) works. Numerical analysis of the algorithm is given for a non-trivial application for the Baltic Sea dynamics model.

Most of the explanations presented here are introduced in the text of the revised version.

We are greatly thankful to the Referee for general appreciation of our work and for very useful remarks and comments which helped us to improve the paper.

Sincerely,

On behalf of the authors,
Victor Shutyaev