Dear Olivier,

Thank you very much for your valuable comments and suggestions. We take them into account in the revised version. Below are our responses to each item.

1. P. 15, l. 3, ... t1 = t 0 + Δt. Does it mean the assimilation was performed over only one timestep Δt = 5 minutes of the model ? If so, that reduces somewhat the interest of the experiment. It means that there is no propagation of information between grid-points, and that the adjustment of heat fluxes to observed temperatures is purely local. In addition, no significant convection can occur over the assimilation window, thus rendering inappropriate the explanation given p. 15, ll. 16-18 for the larger sensitivities seen in shallow areas on Fig. 1. Clarification of these points is desirable.

We performed the assimilation for different time intervals, and then $t1=t0 + M \Delta t$, where M is the number of subintervals we use for approximation. The number M could be sufficiently large, such that $M\Delta t$ is one day, or several months. However, in Fig.1, we present the results of calculations just for M = 1, to demonstrate the work of the algorithm to construct the gradient G' for one time subinterval. Due to the splitting method, for each subinterval we assimilate T_{obs} to reconstruct the heat flux Q (on the sea surface) which depends on x,y,t. It is not an initialization problem, but the parameter estimation problem, and in this case the model itself plays a role of an interpolant and propagates the information between grid points in x,y,z. Therefore, the adjustment of heat fluxes to the observed temperature is done in each grid point x,y, and the assimilation takes into account interactions in z-direction, because we use for assimilation an iterative process, and all the model parameters are consistent with the full INM RAS model (where temperature, velocities, salinity, sea level etc. are calculated at each time step). Therefore, even for a short time interval, the convection term plays a role in assimilation, and effects the result in z-direction (depth). The splitting procedure for assimilation and the iterative process used are described in detail in our former papers [1]-[2].

2. Since numerical values are given without units (and without elements for comparison) in the numerical experiments (Figure 1, parameter α on p. 15, l. 8), they are almost meaningless (and would not allow comparison with other experiments). It would be desirable to say more. In particular were SI units used in the numerical experiments ?

We use here the SI units, namely, K (kelvin) is used for temperature, ms^{-1} for velocities, mKs^{-1} for the heat flux Q. The parameter α is defined as s^2m^{-2} to give the both terms in (6.5) the same dimension. It is easily seen that in this case, the units of the gradient G' from (6.18) are defined as $m^{-2}s^{-1}$. We introduce these details in the revised version.

3. P. 9, l. 5. Since λ is defined in Eq. (5.2) as the minimizer of the function J, J(v) would be more appropriate here (check for possible similar corrections elsewhere).

Corrected.

4. P. 4, l. 12, ... it contains all the available information. That is vague. Can you be more precise?

Corrected.
5. P. 15, l. 7, *a regularization parameter* α, *which appears near the term involving* $\boldsymbol{Q}$ *and* $\boldsymbol{Q}(0)$ → … *parameter* α, *which weights the squared difference* $\left|\boldsymbol{Q} - \boldsymbol{Q}(0)\right|^2$.

Corrected.

6. P. 14, l. 17, … *with zero initial condition* … I understand this means that the initial velocity was zero. But what about temperature ? What is uniform ?

Zero initial conditions (including temperature, velocities, salinity) are taken to overclock the system, and this is done at the preliminary stage for the INM RAS model. Climatic data from the atmosphere are used as boundary conditions and partly in the right-hand side. Therefore, after 20 years of calculating the model, a solution close to the climate for the region is obtained. After this, the nonzero solution obtained is taken as the initial condition for the experiment described in this paper.

New explanations are introduced in the text of the revised version, and they are marked red.

We are greatly thankful to you for very useful critical remarks and comments which helped us to improve the paper.

Sincerely,

Victor Shutyaev, Francois-Xavier Le Dimet, and Eugene Parmuzin