On possibilities of assimilation of near-real-time pollen data by atmospheric composition models

Mikhail Sofiev

Received: 25 January 2019 / Accepted: 27 March 2019 / Published online: 9 April 2019
© The Author(s) 2019

Abstract Assimilation of pollen observations for increasing accuracy of the deterministic model forecast and reanalysis faces several roadblocks. The most-evident problem today is the long delay of the data because of the manual character of the observations. Automatic monitors are about to eliminate this issue but more are on the way. This paper shows that the classical assimilation of the model state, i.e. pollen concentrations, has very little effect on the forecasts. Due to short relaxation time of the system, the updates generated by the assimilation are forgotten within a few hours. In a search of approaches with a longer-lasting impact, a numerical experiment is conducted assimilating the total seasonal pollen emission, which controls the overall season severity. It turned out to be a prominent example of parameters affecting the model predictions over the long period—in the retrospective simulations. It remains to be demonstrated that this parameter can substantially benefit from assimilation of the near-real-time data that are becoming available from the automatic pollen monitors.

Keywords Pollen forecasting · Data assimilation · Seasonal pollen index · SILAM

1 Introduction

Possibility to improve predictions of atmospheric composition models via assimilation of the observed concentrations has been extensively discussed during last 15 years (Bocquet et al. 2015; Elbern et al. 2007, 2000; Gaubert et al. 2014; Schwinger and Elbern 2010; Vira et al. 2017; Vira and Sofiev 2012). The data assimilation (DA) technology, in its classical form, adjusts the model concentrations bringing them closer to the observations. There are many mathematical procedures of achieving this goal, but they all solve some optimization problems trying to find the concentration field, which agrees with observations but does not deviate too much from the initial model state. Upon completing the assimilation step, the model is let to run freely through the forecasting period starting from this optimized field, which is referred as the analysis state.

Unlike the meteorological problems, which are primarily controlled by the model initial state (for regional applications, also by boundary conditions), the atmospheric composition problems are known to be driven by the external forcing, first of all, by emission of atmospheric tracers, boundary conditions for regional tasks and by meteorological forcing for offline applications (see, e.g. (Seinfeld and Pandis 2006), chapter 25).

With a closer look, the problem appears somewhat more complicated. At the beginning of the

M. Sofiev (✉)
Finnish Meteorological Institute, Erik Palmenin Aukio 1,
Helsinki, Finland
e-mail: mikhail.sofiev@fmi.fi

simulations, the air quality forecast is fully determined by the initial state, i.e. it can be successfully improved by the classical DA algorithms developed for the meteorological problems decades ago, e.g. (Le Dimet and Talagrand 1986). However, this “free run” will eventually drift to the “forced run” dictated by the external forcing. It happens with all systems and the only question is whether the length of the free run is sufficiently long to cover the desired forecast length. For weather forecasts, the free run lasts long enough, except for the seasonal forecasts and the climate-scale simulations. Therefore, assimilation of the model initial state is usually sufficient for good weather prediction. For the atmospheric composition problems, this relaxation period depends on species, meteorological conditions, region and can be too short even for basic AQ prediction tasks (Elbern et al. 2007; Virá and Sofiev 2012).

To date, there are no practical examples of data assimilation applied by pollen or spores dispersion models. An experimental research performed with the SILAM model (System for Integrated modelling of Atmospheric composition, <http://silam.fmi.fi>) of Finnish Meteorological Institute (FMI) was outlined in the report of the MACC project (Rouïl et al. 2011), but the broad agenda of the report did not allow for the in-depth discussion of the subject.

The objective of this paper is to highlight the promising directions of development of the data assimilation procedures suitable for pollen and spores forecasting. The paper describes a series of numerical experiments with the SILAM model to reveal the pros and cons of the DA algorithms. The next section presents the motivation for these tasks and formulates the scientific questions for each experiment. Section 3 then describes the set-up of the model runs, whose outcome is presented in Sect. 4 and discussed in the Sect. 5.

2 Motivation and general problem formulation

The DA technology can be useful for two tasks: (1) retrospective analysis of the past pollen season and (2) forecasting the pollen concentrations during the course of the season in real time. The roles of the model and the observations in these tasks are different: for the retrospective analysis, the observational data are already at hands, i.e. the model is needed as a smart

interpolation/extrapolation tool for the areas with insufficient density, frequency and coverage of the observational network. Since the observations are expensive, the model-based reanalysis is an efficient way to evaluate the distribution pattern over large areas using the (few) available observations to “anchor” it to reality. For the forecasting task, the model is the only available instrument to predict the evolution of the regional/continental pattern several days ahead using the observations available in the past. Evidently, timeliness of the observations becomes critical: the shorter the free-run forecasting interval the stronger the assimilation impact will be.

The standard pollen observations today use the Hirst-type traps (Hirst 1952), which require tedious manual analysis of the slides collected during the previous week. Such data cannot be used for assimilation in the forecasting mode. With the delay of 1–2 weeks, the season might already be over by the moment the model gets information about its beginning. Because of that, all numerical pollen forecasts today are performed without any usage of the observations.

Direct usage of the pollen data for the model-based forecasting becomes feasible with introduction of online pollen monitors, which can deliver the information within few hours from the observation moment (Crouzy et al. 2016; Oteros et al. 2015). The operational maturity of the technology is yet to be achieved (Šauliėne et al. 2019), but the pollen DA will also take time to become practically usable. Outlining the possible directions of its development and presenting the results of the first experiments are the objectives of this paper.

3 The modelling experiments

Two modelling experiments were constructed as problems with the known answers. For both cases, the “observational” dataset was generated by the model using actual locations of the stations of European Aeroallergen Network (EAN, <https://ean.polleninfo.eu/Ean/>, totally 687 places) but not involving the actual observed concentrations. The dispersion model used in all experiments was SILAM v.5 (Sofiev et al. 2012; Sofiev et al. 2015a, b).

3.1 Adjusting the concentrations in the forecasting regime

As pointed out above, the short relaxation time of the system can make the classical assimilation schemes useless. For pollen, the problem is potentially severe because this coarse aerosol has comparatively short atmospheric lifetime (Sofiev et al. 2006). The aim of the first experiment was to quantify this effect and to verify if any practically useful results can be obtained with the classical DA algorithms.

The experiment included three SILAM runs. The first one was the simulation closely following the template of (Sofiev et al. 2012, 2015). Namely, the pollen emission was computed by the dual-threshold thermal-type model (Linkosalo et al. 2010; Sofiev et al. 2012). The domain covered the whole Europe, computations were made for birch pollen in 2018, resolution was $0.1 \times 0.1^\circ$ with 8 uneven vertical layers reaching up to ~ 7 km. The model output consisted of the hourly concentration and deposition fields.

The second and the third runs were forked from the first one at arbitrary moments of time: at midnight, 00:00 20 April 2018 and mid-day, 12:00 20 April 2018—hereinafter referred as “night-perturbed” and “day-perturbed”, respectively. At these moments, the 3-D concentration maps were multiplied with a factor

of 10 leaving the other parameters intact (see screenshots in Fig. 1). The model then continued the simulations, and the time period needed for the perturbed runs to relax back to the reference one was calculated.

The measure of the distance between the runs was taken to be the L_2 -metric (also referred as Frobenius norm) between the hourly near-surface concentration fields:

$$D_{L2}(t) = \sqrt{\sum_{i,j} (C_{\text{pert}}(i,j,1,t) - C_{\text{ref}}(i,j,1,t))^2}$$

where C_{pert} and C_{ref} are perturbed and reference concentrations, respectively, t is time, i and j are the grid cell indices and D_{L2} is distance between the fields.

3.2 Assimilation of seasonal pollen emission: efficient for retrospective analysis

The modelling experiment mimicking the retrospective analysis was constructed assimilating the map of the total pollen release over the whole season. It is the static parameter of the model and stays unchanged during the season. The aim of the experiment was to demonstrate that (some) static model parameters can be successfully used as control variables for the data assimilation.

Fig. 1 Birch pollen concentrations at 13:00 20 April 2018 predicted by three SILAM runs: standard one (left), night-perturbed (middle, 12 h after the perturbation) and day-perturbed (right, 1 h after the perturbation). [pollen m^{-3}]

Fig. 2 Scheme of the retrospective assimilation experiment

The experiment was conducted in two steps (Fig. 2).

At the first step, the standard birch season simulations were performed for 2014 following the same template as the first experiment and essentially repeating the SILAM computations of (Sofiev et al. 2015a) but with coarser resolution $0.2^\circ \times 0.2^\circ$ to reduce the computational demand. Meteorological input was taken from the operational archives of European Centre for Medium-Range Weather forecast (ECMWF). The emission fluxes of birch pollen were stored hour by hour as a time-resolving emission dataset. Its integral over the whole season is shown in Fig. 3a. The output of the reference run was also used for setting up pseudo-observations, which were just

the model predictions at the EAN stations locations scaled by a constant factor of 0.5.

The second step was the data assimilation run, which used the stored hourly pollen emission of the reference run as a prescribed first guess, $E_{fg}(i,j,t)$. The outcome of the assimilation was the map of correction factors $\alpha(i,j)$, constant-in-time multipliers for the first-guess emission fields, which turn them into the a posteriori solution of the data assimilation problem $E_{assim}(i,j,t)$:

$$E_{assim}(i,j,t) = E_{fg}(i,j,t) \alpha(i,j)$$

The meteorological dataset was taken from the ERA-Interim archive (Dee et al. 2011) instead of the operational ECMWF forecasts used in the reference

Fig. 3 Total seasonal emission of birch pollen (a) and seasonal pollen index (b) obtained from the SILAM reference run for the season of 2014

run. The “measurements” were the pseudo-observations from the reference run, i.e. the halved reference model predictions.

This set-up brings about two challenges for the model: different meteorological driver and the two times lower observed concentrations compared to what one would expect from the stored emission. The objective of the DA was to correct the emission overestimation disregarding the noise due to the different meteorological driver.

The assimilation run used four-dimensional variational technique (4D-Var) closely following the template of Vira and Sofiev (2012) with the exception for the horizontal correlation distance, which was set to 250 km. In order to match the birch season duration in Europe and average out the fluctuations due to meteorology, a very long assimilation window was set from 15 March till 15 June 2014. The initial value of the control variable (output of the assimilation, the correction factor α) was set to 1 for the whole domain.

4 Results

4.1 The model relaxation experiment

The qualitative effect of the massive perturbation of the concentration field is shown in Figs. 1 and 4. As

one can see, the difference from the reference run is visible after 12 h (Fig. 1, middle panel for night-perturbed run, and Fig. 4, right-hand panel for day-perturbed run) but mostly fades out within 24 h (Fig. 4, middle panel for night-perturbed run).

The quantitative measure of the model memory ($L2$ distance between the reference run and the perturbed ones) is presented in Fig. 5. As the graphs show, a half of the initial perturbation is gone within ~ 5 h while 90% vanishes within 20 and 24 h for night- and daytime perturbations, respectively.

Exceptions, such as some Mediterranean regions (compare the left and middle panels of Fig. 4) that maintained the tenfold difference after 24 h, illustrate the complexity highlighted in the introduction. The effect of the perturbation disappears because the external forcing takes over. But in the areas that have no external forcing the effect of perturbation will last. And this is what happened in the Mediterranean: (1) there are no birches there, i.e. the emission forcing is zero; (2) the region is far from the domain borders, i.e. the boundary conditions are irrelevant, (3) there was no perturbation of meteorology, i.e. the transport was the same in all runs. Under these conditions, the plumes released in Central Europe and perturbed during their flow towards the south, continued without further disturbance until get deposited or brought out of the domain. This is in contrast with Central Europe,

Fig. 4 Birch pollen concentrations at 00:00 21 April 2018 predicted by three SILAM runs: reference one (left), night-perturbed (middle, 24 h after the perturbation) and day-perturbed (right, 12 h after the perturbation). [pollen m^{-3}]

Fig. 5 Fraction of the initial D_{L2} distance between the reference and perturbed runs as a function of time passed since the perturbation moment. [relative unit]

where the emission forcing quickly overwhelms the perturbation.

The relaxation time of 10–20 h, already too short for practical applications, has to be considered as the upper-most estimate since the update of the whole 3D domain cannot be done in real life: one cannot measure in every grid cell at all heights. As a result, the actual relaxation time in any practically relevant case of the near-surface perturbation will be even shorter. We will not have the “exceptional” plumes with long memory because the unperturbed upper layers will quickly mix out the effect. Therefore, the assimilation of the pollen concentrations can hardly bring a significant added value in the operational forecasting context.

4.2 Season intensity assimilation experiment

The outcome of the assimilation experiment—the correction factor α for the emission field of Fig. 3a—is shown in Fig. 6a. The colour scale is made so that the deviation of this correction from the known answer—a constant factor of 0.5—is easy to decipher. Figure 6b shows the so-called L-curve, which depicts the RMSE of each iteration as a function of the deviation of the control variable (the correction factor map) from its initial value of unity. As the L-curve shows, the assimilation achieves $\sim$ tenfold reduction of the RMSE, i.e. eliminates $\sim$ 90% of the initial L_2 discrepancy between the model and the observations.

5 Discussion

The fact that the efficiency of the data assimilation in the chemistry transport problems is hampered by the limited impact period was recognized more than a decade ago—e.g. (Elbern et al. 2007)—and also noticed in experiments with SILAM (Vira and Sofiev 2012, 2015). But it was also pointed out that the effect strongly varies between the species. In particular, it is affected by the ratio of the formation (or emission) and removal fluxes and the total amount of the species in the air: the larger the reservoir and the smaller the exchange fluxes the longer will the assimilation effect last. The first experiment described above suggested that pollen has very fast exchange and relatively small amount in the air. Therefore, the effect of the classical DA is bound to be short.

Looking for the longer-lasting improvement of the forecasts, Sofiev et al. (2017) applied data fusion procedure to an ensemble of models and showed that the model weighting coefficients are quite stable: the five-day-long forecast showed very little degradation. However, data fusion is a post-processing, whose power lies on the assumption that the model quality has some systematic features, which are stable in time. In such a case, one can construct a correction procedure, which does not affect the model itself but improves the final forecasts. This is different from the data assimilation goals: we would like to “teach” the model rather than tweak its output.

From the above flux-vs-reservoir model, one can see that an assimilation procedure constructed so that it improves fluxes rather than concentrations, will have a longer-lasting effect. This suggests assimilating the pollen emission and/or deposition parameters. This problem is noticeably more complicated than the classical assimilation and much less studied. Choosing among the possible options, constraining the total seasonal emission should have the most-profound impact lasting throughout the whole season. This hypothesis has been confirmed by the second assimilation experiment.

The second experiment, despite bringing about a very stable and pronounced effect (tenfold reduction of the error—see Fig. 6b), has several features that require further work.

The final scaling map (Fig. 6a) shows by no means a constant map of $\sim$ 0.5. Only the major sources of birch pollen in the Northern Europe are constrained,

Fig. 6 Outcome of assimilation: Panel **a** correction factor to the total seasonal emission, iteration 10, **b** L-curve describing the assimilation trajectory

whereas the areas in Southern Europe where the birch SPI is low are practically not touched. This is despite the fact that the 83 station points were in Spain (out of 687 EAN locations all over Europe). The reason is that RMSE, as a sum of squared model-measurement differences, is not sensitive to small values: if both the model prediction and the pseudo-observation are small their contribution to the cost function is negligible. The effect can be corrected by a more accurate selection of the observation covariance matrix with a smaller uncertainty attributed to small observed values. However, such details are beyond the scope of the current paper, which aims at illustrating just the main problems and directions.

One can also notice certain overregulation: the scaling is ~ 0.4 over the bulk of the continent. It reflects the impact of the meteorological disturbance, which partly de-correlates the time series and makes the limited underestimation of the concentrations a plausible way to reduce RMSE.

Finally, the procedure is not yet a solution for the forecasting problem. Indeed, the procedure relies on availability of the data throughout the whole season. The technology needs long assimilation window in order to reduce the noise and accumulate the signal from as many observations as possible. Lifting this

requirement, one can apply it daily for the past days following the season propagation. Three examples of such sequential applications are depicted in Fig. 7, where the correction is determined using the observations from (a) the early season in Southern Europe 15.3–1.4, (b) the start of the Central-European season 15.3–15.4, (c) the main season in Central Europe 15.3–1.5. Comparison of these maps with the whole-season-based correction (Fig. 6) shows that the assimilation can follow the season features gradually correcting the emission intensity when information appears. One still have to keep in mind that the sequential application is retrospective: it cannot correct the emission before seeing it going off. Experience of the birch season of 2013 (Sofiev et al. 2015a) showed that spatial extrapolation has to be done with care. For instance, the moderate season in Central Europe in 2013 was accompanied with one of the lowest-ever seasons in the north. Therefore, the added value of the sequential assimilation of the season severity in the forecasting context still needs to be demonstrated.

Assimilation of the whole-season pollen emission is just one possible direction, which is highly promising for birch. But for other taxa, the effect durability can be associated with other variables, so that their

Fig. 7 Sequential assimilations of the total seasonal emission performed at different stages of the season

assimilation can be more efficient. As a result, the DA formulations can be taxon-specific.

The results of this paper are all obtained with the single model—SILAM—and, strictly speaking, a possibility of their generalization is yet to be shown. Indeed, the specific concentration levels and the algorithm details are model-dependent. However, as shown by (Sofiev et al. 2015a) and (Sofiev et al. 2017), the difference between the different models is usually much less than the effect demonstrated above. For instance, RMSE for birch (Sofiev et al. 2015a) differs by less than a factor of two between the seven CAMS models, whereas the above assimilation reduced it almost tenfold (Fig. 6). Similar consideration holds for the model memory experiment: even twofold increase in the pollen lifetime in the air would be still too short to last over the desired forecasting period.

6 Conclusions

A series of the SILAM-based modelling experiments showed that assimilation of pollen data, even if they are available in near-real-time, is a challenge. In particular, the effect of the update of the pollen concentrations—a standard approach of the classical data assimilation—vanishes within a few hours. Even full update of the 3D concentration field lasts about 20 h, which is then the upper estimate of the longevity of the concentration assimilation effect. Therefore, this approach hardly has any practical value for pollen.

A lasting improvement can be obtained if assimilation targets the emission parameters, which are static or nearly static. The second model experiment confirmed that correction of the total seasonal pollen emission field is (1) possible, (2) efficient, (3) has

positive impact over the whole season. The experiment, however, used the whole-season data and thus cannot serve as a prototype for the operational forecasting. Its area of applications is rather reanalysis.

The same assimilation set-up applied sequentially at different stages of the season showed good results successfully reducing the model error at least tenfold in all runs. However, the adjustment was all but entirely concentrated on the areas of active pollen release during the assimilation window. Possibility of the spatial extrapolation beyond the area with ongoing pollination needs further investigation.

Acknowledgements The study was performed within the scope of the projects REALTIME (received funding from European Social Fund project No. 09.3.3-LMT-K-712-010066 under grant agreement with LMTLT) and PS4A (No 318194) of Academy of Finland. Support of Copernicus CAMS50 service and EUMETNET AutoPollen programme is kindly appreciated.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Bocquet, M., Elbern, H., Eskes, H., Hirtl, M., Aabkar, R., Carmichael, G. R., et al. (2015). Data assimilation in atmospheric chemistry models: Current status and future prospects for coupled chemistry meteorology models. *Atmospheric Chemistry and Physics*, 15, 5325–5328. <https://doi.org/10.5194/acp-15-5325-2015>.

- Crouzy, B., Stella, M., Konzelmann, T., Calpini, B., & Clot, B. (2016). All-optical automatic pollen identification: Towards an operational system. *Atmospheric Environment*, 140, 202–212. <https://doi.org/10.1016/j.atmosenv.2016.05.062>.
- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137, 553–597. <https://doi.org/10.1002/qj.828>.
- Elbern, H., Schmidt, H., Talagrand, O., & Ebel, A. (2000). 4D-variational data assimilation with an adjoint air quality model for emission analysis. *Environmental Modelling and Software*, 15, 539–548. [https://doi.org/10.1016/S1364-8152\(00\)00049-9](https://doi.org/10.1016/S1364-8152(00)00049-9).
- Elbern, H., Strunk, A., Schmidt, H., & Talagrand, O. (2007). Emission rate and chemical state estimation by 4-dimensional variational inversion. *Atmospheric Chemistry and Physics*, 7, 3749–3769. <https://doi.org/10.5194/acpd-7-1725-2007>.
- Gaubert, B., Coman, A., Foret, G., Meleux, F., Ung, A., Rouil, L., et al. (2014). Regional scale ozone data assimilation using an ensemble Kalman filter and the CHIMERE chemical transport model. *Geoscientific Model Development Discussions*, 7, 283–302. <https://doi.org/10.5194/gmd-7-283-2014>.
- Hirst, J. M. (1952). An automatic volumetric spore trap. *Annals of Applied Biology*, 39, 257–265. <https://doi.org/10.1111/j.1744-7348.1952.tb00904.x>.
- Le Dimet, F.-X., & Talagrand, O. (1986). Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. *Tellus A*, 38A, 97–110. <https://doi.org/10.1111/j.1600-0870.1986.tb00459.x>.
- Linkosalo, T., Ranta, H., Oksanen, A., Siljamo, P., Luomajoki, A., Kukkonen, J., et al. (2010). A double-threshold temperature sum model for predicting the flowering duration and relative intensity of *Betula pendula* and *B. pubescens*. *Agricultural and Forest Meteorology*, 150, 1579–1584. <https://doi.org/10.1016/j.agrformet.2010.08.007>.
- Oteros, J., Pusch, G., Weichenmeier, I., Heimann, U., Möller, R., Röseler, S., et al. (2015). Automatic and online pollen monitoring. *International Archives of Allergy and Immunology*, 167, 158–166. <https://doi.org/10.1159/000436968>.
- Rouil, L., Beekman, M., Foret, G., Sofiev, M., & Vira, J. (2011). Assessment report: Air quality in Europe in 2009. Reading.
- Šaulienė, I., Šukienė, L., Daunys, G., Valiulis, G., Vaitkevičius, L., Matavulj, P., et al. (2019). Automatic pollen recognition with the rapid-E particle counter: The first-level procedure, experience and next steps. *Atmospheric Measurement Techniques Discussion*, 2, 1–33. <https://doi.org/10.5194/amt-2018-432>.
- Schwinger, J., & Elbern, H. (2010). Chemical state estimation for the middle atmosphere by four-dimensional variational data assimilation: A posteriori validation of error statistics in observation space. *Journal of Geophysical Research*. <https://doi.org/10.1029/2009jd013115>.
- Seinfeld, J. H., & Pandis, S. N. (2006). *Atmospheric chemistry and physics. From air pollution to climate change* (2nd ed.). Hoboken: Wiley.
- Sofiev, M., Berger, U., Prank, M., Vira, J., Arteta, J., Belmonte, J., et al. (2015a). MACC regional multi-model ensemble simulations of birch pollen dispersion in Europe. *Atmospheric Chemistry and Physics*, 15, 8115–8130. <https://doi.org/10.5194/acp-15-8115-2015>.
- Sofiev, M., Ritenberga, O., Albertini, R., Arteta, J., Belmonte, J., Bonini, M., et al. (2017). Multi-model ensemble simulations of olive pollen distribution in Europe in 2014. *Atmospheric Chemistry and Physics Discussion*, 17, 1–32. <https://doi.org/10.5194/acp-2016-1189>.
- Sofiev, M., Siljamo, P., Ranta, H., Linkosalo, T., Jaeger, S., Rasmussen, A., et al. (2012). A numerical model of birch pollen emission and dispersion in the atmosphere. Description of the emission module. *International Journal of Biometeorology*, 57, 54–58. <https://doi.org/10.1007/s00484-012-0532-z>.
- Sofiev, M., Siljamo, P., Ranta, H., & Rantio-Lehtimäki, A. (2006). Towards numerical forecasting of long-range air transport of birch pollen: Theoretical considerations and a feasibility study. *International Journal of Biometeorology*. <https://doi.org/10.1007/s00484-006-0027-x>.
- Sofiev, M., Vira, J., Kouznetsov, R., Prank, M., Soares, J., & Genikhovich, E. (2015b). Construction of an Eulerian atmospheric dispersion model based on the advection algorithm of M. Galperin: dynamic cores vol 4 and 5 of SILAM v.5.5. *Geoscientific Model Development Discussions*, 8, 3497–3522. <https://doi.org/10.5194/gmd-8-3497-2015>.
- Vira, J., Carboni, E., Grainger, R. G., & Sofiev, M. (2017). Variational assimilation of IASI SO₂ plume height and total column retrievals in the 2010 eruption of Eyjafjallajökull using the SILAM v5.3 chemistry transport model. *Geoscientific Model Development*, 10, 1985–2008. <https://doi.org/10.5194/gmd-10-1985-2017>.
- Vira, J., & Sofiev, M. (2012). On variational data assimilation for estimating the model initial conditions and emission fluxes for short-term forecasting of SO_x concentrations. *Atmospheric Environment*, 46, 318–328. <https://doi.org/10.1016/j.atmosenv.2011.09.066>.
- Vira, J., & Sofiev, M. (2015). Assimilation of surface NO <math><inf>2</inf></math> and O <math><inf>3</inf></math> observations into the SILAM chemistry transport model. *Geoscientific Model Development Discussions*. <https://doi.org/10.5194/gmd-8-191-2015>.