



High-Resolution Source Estimation of Volcanic Sulfur Dioxide Emissions Using Large-Scale Transport Simulations

Mingzhao Liu¹, Yaopeng Huang¹, Lars Hoffmann⁵, Chunyan Huang⁶, Pin Chen^{2,3},
and Yi Heng^{2,3,4}(✉)

¹ School of Chemical Engineering and Technology, Sun Yat-sen University, Guangzhou, China

² School of Data and Computer Science, Sun Yat-sen University, Guangzhou, China
hengyi@mail.sysu.edu.cn

³ National Supercomputing Center in Guangzhou (NSCC-GZ), Guangzhou, China

⁴ Guangdong Province Key Laboratory of Computational Science, Guangzhou, China

⁵ Jülich Supercomputing Centre, Forschungszentrum Jülich, Jülich, Germany

⁶ School of Statistics and Mathematics, Central University of Finance and Economics,
Beijing, China

Abstract. High-resolution reconstruction of emission rates from different sources is essential to achieve accurate simulations of atmospheric transport processes. How to achieve real-time forecasts of atmospheric transport is still a great challenge, in particular due to the large computational demands of this problem. Considering a case study of volcanic sulfur dioxide emissions, the codes of the Lagrangian particle dispersion model MPTRAC and an inversion algorithm for emission rate estimation based on sequential importance resampling are deployed on the Tianhe-2 supercomputer. The high-throughput based parallel computing strategy shows excellent scalability and computational efficiency. Therefore, the spatial-temporal resolution of the emission reconstruction can be improved by increasing the parallel scale. In our study, the largest parallel scale is up to 1.446 million compute processes, which allows us to obtain emission rates with a resolution of 30 min in time and 100 m in altitude. By applying massive-parallel computing systems such as Tianhe-2, real-time source estimation and forecasts of atmospheric transport are becoming feasible.

Keywords: Source estimation · High-throughput computing · Transport simulations · Volcanic emissions

1 Introduction

Model simulations and forecasts of volcanic aerosol transport are of great importance in many fields, e.g., aviation safety [1], studies of global climate change [2, 3] and atmospheric dynamics [4]. However, existing observation techniques, e.g., satellite measurements, cannot provide detailed and complete spatial-temporal information due to their own limitations. With appropriate initial conditions, numerical simulations can provide

relatively complete and high-resolution information in time and space. Model predictions can help to provide early warning information for air traffic control or input to studies of complex global or regional atmospheric transport processes.

In order to achieve accurate atmospheric transport simulations, it is necessary to first combine a series of numerical techniques with limited observational data to achieve high-resolution estimates of the emission sources. These techniques include backward-trajectory methods [5], empirical estimates [6] and inverse approaches. Among them, the inverse approaches are universal and systematic in the identification of atmospheric emission sources due to their mathematical rigor.

For instance, Stohl et al. [7] used an inversion scheme to estimate the volcanic ash emissions related to the volcanic eruptions of Eyjafjallajökull in 2010 and Kelut in 2014. They utilized Tikhonov regularization to deal with the ill-posedness of the inverse problem. Flemming and Inness [8] applied the Monitoring Atmospheric Composition and Climate (MACC) system to estimate sulfur dioxide (SO_2) emissions by Eyjafjallajökull in 2010 and Grimsvötn in 2011, in which the resolution of the emission rates is about 2–3 km in altitude and more than 6 h in time. Due to limitations in computational power and algorithms, the spatial-temporal resolution of the reconstructed source obtained in previous studies is relatively low.

The main limitations of real-time atmospheric transport forecasts are the great computational effort and data I/O issues. Some researchers tried to employ graphics processing units to reduce the computational time and got impressive results [9–11]. Lagrangian particle dispersion models are particularly well suited to distributed-memory parallelization, as each trajectory is calculated independently of each other. To reduce the computational cost, Larson et al. [12] applied a shared- and distributed-memory parallelization to a Lagrangian particle dispersion model and achieved nearly linear scaling in execution time with the distributed-memory version and a speed-up factor of about 1.4 with the shared-memory version. In the study of Müller et al. [11], the parallelization of the Lagrangian particle model was implemented in the OpenMP shared memory framework and good strong scalability up to 12 cores was achieved.

In this work, we implement the Lagrangian particle dispersion model Massive-Parallel Trajectory Calculations (MPTRAC) [5] on the Tianhe-2 supercomputer, along with an inverse modeling algorithm based on the concept of sequential importance resampling [13] to estimate time- and altitude-dependent volcanic emission rates. In order to realize large-scale SO_2 transport simulations on a global scale, high-resolution emission reconstructions and real-time forecasts, the implementation is based on state-of-the-art techniques of supercomputing and big-data processing. The computing performance is assessed in the form of strong and weak scalability tests. Good scalability and computational efficiency of our codes make it possible to reconstruct emission rates with unprecedented resolution both in time and altitude and enable real-time forecasts.

The remainder of this manuscript is organized as follows: Sect. 2 introduces the forward model, the inverse modeling algorithm and the parallelization strategies. Section 3 presents the parallel performance of the forward and inverse code on the Tianhe-2 supercomputer. In Sect. 4, the results of the emission reconstruction and forward simulation are presented for a case study. Discussions and conclusions are provided in Sect. 5.

2 Data and Methods

2.1 Lagrangian Particle Dispersion Model

In this work, the forward simulations are conducted with the Lagrangian particle dispersion model MPTRAC, which has been successfully applied for volcanic eruption cases of Grímsvötn, Puyehue-Cordón Caulle and Nabro [5]. Meteorological fields of the ERA-Interim reanalysis [14] provided by the European Centre for Medium-Range Weather Forecasts (ECMWF) are used as input data for the transport simulations. The trajectory of an individual air parcel is calculated by

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{v}(\mathbf{x}(t), t), \quad (1)$$

where $\mathbf{x} = (x, y, z)$ denotes the spatial position and $\mathbf{v} = (u, v, w)$ denotes the velocity of the air parcel at time t . Here, x and y coordinates refer to latitude and longitude whereas the z coordinate refers to pressure. The horizontal wind components u and v and the vertical velocity $w = dp/dt$ are obtained by 4-D linear interpolation from the meteorology data, which is common in Lagrangian particle dispersion models [15]. Small-scale diffusion and subgrid-scale wind fluctuations are simulated based on a Markov model following Stohl et al. [16].

In our previous work [17], truncation errors of different numerical integration schemes of MPTRAC have been analyzed in order to obtain an optimal numerical solution strategy with accurate results and minimum computational cost. The accuracy of the MPTRAC trajectory calculations has been analyzed in different studies, including [18], which compared trajectory calculations to superpressure balloon tracks.

2.2 Evaluation of Goodness-of-Fit of Forward Simulation Results

Atmospheric InfraRed Sounder (AIRS) satellite observations are used to detect volcanic SO_2 based on a brightness temperature differences (BTD) algorithm [19]. To evaluate the goodness-of-fit of the forward simulation results obtained by MPTRAC, the critical success index (CSI) [20] is calculated by

$$\text{CSI} = C_x / (C_x + C_y + C_z). \quad (2)$$

Here, the number of positive forecasts with positive observations is C_x , the number of negative forecasts with positive observations is C_y , and the number of positive forecasts with negative observations is C_z . The CSI, representing the ratio of successful predicts to the total number of predicts that were either made ($C_x + C_z$) or needed (C_y), is commonly used for the assessment of the simulation results of volcanic eruptions and other large-scale SO_2 transport problems. Basically, it provides a measure of the overlap of a simulated volcanic SO_2 plume from the model with the real plume as found in the satellite observations. CSI time series are calculated using the AIRS satellite observations and MPTRAC simulation results mapped on a discrete grid, which are essential to the inverse modeling algorithm presented in the next section.

2.3 Inverse Source Estimation Algorithm

The strategy for the inverse estimation of time- and altitude-dependent emission rates is shown in Fig. 1 and Algorithm 1. The time- and altitude-dependent emissions are considered for the domain $E := [t_0, t_f] \times \Omega$, which is discretized with n_t and n_h uniform intervals into $N = n_t \cdot n_h$ subdomains. For each subdomain, a forward calculation of a set of air parcel trajectories is conducted with MPTRAC, which is referred to here as ‘unit simulation’ for a given time and altitude. Each unit simulation is assigned a certain amount of SO_2 , where we assume that the total SO_2 mass over all unit simulations is known a-priori. During the inversion, a set of importance weights $w_i (i = 1, \dots, N)$, which satisfy $\sum_{i=1}^N w_i = 1$, are estimated to represent the relative posterior probabilities of the occurrence of SO_2 emission mass.

At first, the subdomains are populated with SO_2 emissions (air parcels) according to an equal-probability strategy. N parallel unit simulations with a certain amount of air parcels are performed in an iterative process and the corresponding time series (CSI_k^i) with $k = 1, \dots, n_k$ at different times t_k and $i = 1, \dots, N$ are calculated to evaluate the agreement of the simulations with the satellite observations. Then, the importance weights are updated according to the following formulas:

$$w_i = m_i / \sum_{a=1}^N m_a, \quad (3)$$

$$m_i = \left(\sum_1^{n_k} \text{CSI}_k^i \right) / n_k. \quad (4)$$

During the iteration, the m_i represents the probability of emitted source air parcels that fall in the i^{th} temporal and spatial subdomain. Finally, after the termination criterion is satisfied, the emission source is obtained based on the final importance weight distribution. To define the stopping criterion, we calculate the relative difference d by

$$d(\mathbf{W}^{r+1}, \mathbf{W}^r) = \frac{\|\mathbf{W}^{r+1} - \mathbf{W}^r\|}{\max(\|\mathbf{W}^{r+1}\|, \|\mathbf{W}^r\|)}, r \geq 1, \quad (5)$$

where r denotes the iterative step and the norm is defined by

$$\|\mathbf{W}^r\| = \sqrt{\sum_{i=1}^N |w_i^r|^2}, \mathbf{W}^r = (w_i^r)_{i=1 \dots N}. \quad (6)$$

As a stopping criterion, threshold of the relative difference d is chosen to be 1%.

In practice, in order to deal with the complexity of the SO_2 air parcel transport, a so-called ‘product rule’ is utilized in the resampling process, in which the average CSI time series is replaced by the product of two average CSI time series in subsequent and separate time periods:

$$m_i = \left(\sum_{k=1}^{n'_k} \text{CSI}_k^i / n'_k \right) \cdot \left(\sum_{k=n'_k+1}^{n_k} \text{CSI}_k^i / (n_k - n'_k) \right), 1 \leq n'_k < n_k, \quad (7)$$

where n'_k is a “split point” of the time series. This strategy can better eliminate some low-probability local emissions when reconstructing source terms, thus leading to accurate final forward simulation results locally and globally. A detailed description of the inverse algorithm and the improvements due to applying the product rule can be found in [21].

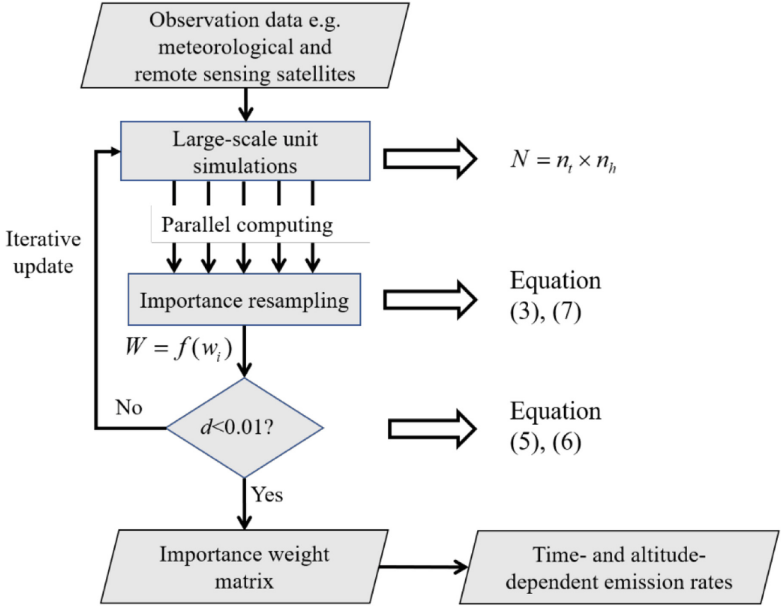


Fig. 1. Flow chart of the inverse modeling strategy

Algorithm 1 Importance sampling algorithm in the inverse modeling

Initialize: Set $r = 1$, and initially set all importance weights $w_i^0 = 1/N$;

repeat

for $i = 1 : N$ **do** in parallel

 Perform forward simulation with w_i^{r-1} ;

 Calculate $CSI_k^i = C_x / (C_x + C_y + C_z)$;

 Update m_i^r, w_i^r with Equation (3), (7);

end

 Calculate $d(\mathbf{W}^r, \mathbf{W}^{r-1})$ with Equation (5), (6);

 Set $r=r+1$;

until $d \leq 0.01$

Output: \mathbf{W}^r .

2.4 Parallel Implementation

The Tianhe-2 supercomputer at the National Supercomputing Center of Guangzhou (NSCC-GZ) consists of 16000 compute nodes, with each node containing two 12-core Intel Xeon E5-2692 CPUs with 64 GB memory [22]. The advanced computing performance and massive computing resources of Tianhe-2 provide the possibility to conduct more complex mathematical research and simulations on much larger scales than before. Based on off-line simulations in previous work, we expect that Tianhe-2 will facilitate applications of real-time forecasts for larger-scale problems. The computational efficiency of the high-precision inverse reconstruction of emission source will directly determine whether the atmospheric SO₂ transport process can be predicted in real time.

To our best knowledge, few studies focus on both, direct inverse source estimation and forecasts, at near-real-time. Fu et al. [23] conducted a near-real-time prediction study on volcanic eruptions based on the LOTOS-EUROS model and an ensemble Kalman filter. Santos et al. [10] developed a GPU-based code to process the calculations in near real time. In this work, we attempt to further develop the parallel inverse algorithm for reconstruction of volcanic SO₂ emission rates based on sequential importance resampling methods, utilizing the computational power of the Tianhe-2 supercomputer to achieve large-scale SO₂ transport simulations and real-time or near-real-time predictions.

The parallelization of MPTRAC and the inverse algorithm is realized by means of a hybrid scheme based on the Message Passing Interface (MPI) and Open Multi-Processing (OpenMP). Since each trajectory can be computed independently, the ensembles of unit simulations are distributed to different compute nodes using the MPI distributed memory parallelization. On a particular compute node, the trajectory calculations of the individual unit simulations are distributed using the OpenMP shared memory parallelization. Theoretically, the calculation time will decrease near linearly with an increasing number of compute processes. Therefore, sufficient computational performance can greatly reduce the computational costs and enable simulations of hundreds of millions of air parcels on the supercomputer system.

The implementation of the inverse algorithm is designed based on a high-throughput computing strategy. At each iterative step, the time- and altitude-dependent domain $E := [t_0, t_f] \times \Omega$ is discretized with n_t and n_h uniform intervals $N = n_t \cdot n_h$, which leads to $N = n_t \cdot n_h$ unit simulations that are calculated in parallel as shown in Fig. 1. Theoretically, the high-throughput parallel computing strategy can greatly improve the resolution of the inversion in time and altitude through increasing the value of N . Only little communication overhead is needed to distribute the tasks and gather the results. With more computing resources, it is possible to operate on more detailed spatial-temporal grids and to obtain more accurate results. In this work, we have achieved a resolution of 30 min in time and 100 m in altitude for the first time with our modeling system.

In summary, the goal of this work is to develop an inverse modeling system using parallel computing on a scale of millions of cores, including high-throughput submission, monitoring, error tolerance management and analysis of results. A multi-level task scheduling strategy has been employed, i.e., the computational performance of each sub-task was analyzed to maximize load balancing. During the calculation, every task is monitored by a daemon with an error tolerance mechanism being established to avoid accidental interruptions and invalid calculations.

3 Parallel Performance Analysis

In this section, we evaluate the model parallel performance on the Tianhe-2 supercomputer based on the single-node performance of MPTRAC for the unit simulations and the multi-node performance of the sequential importance resampling algorithm.

Since our parallel strategy is based on high-throughput computing to avoid communication across the compute nodes, the single-node computing performance is essential in determining the global computing efficiency. To test the single-node computing performance, we employ the Paratune Application Runtime Characterization Analyzer to measure the floating-point speed. An ensemble of 100 million air parcels was simulated on a single node and the gigaflops per second (Gflops) turned out to be 13.16. The strong scalability test on a single node is conducted by simulating an ensemble of 1 million air parcels. The results on strong scaling are listed in Table 1 and the results on weak scaling are listed in Table 2. Referred to a single process calculation, the strong and weak scaling efficiency using 16 computing processes reach 84.25% and 85.63%, respectively.

Table 1. Strong scaling of a single-node MPTRAC simulation

Problem size	Number of processes N_{process}	Clock time/s	Strong speed-up ratio R_s	Efficiency (R_s/N_{process})
1 million parcels	1	1496	1x	100%
	4	399	3.75x	93.75%
	16	111	13.48x	84.25%

Table 2. Weak scaling of a single-node MPTRAC simulation

Problem size	Number of processes N_{process}	Clock time/s	Weak speed-up ratio R_w	Efficiency (R_w/N_{process})
1 million parcels	1	1496	1x	100%
4 million parcels	4	1597	3.75x	93.75%
16 million parcels	16	1753	13.7x	85.63%

Since the ensemble simulations with MPTRAC covering multiple unit simulations are conducted independently on each node, the scaling efficiency of the MPI parallelization is mostly limited by I/O issues rather than communication or computation. Nevertheless, we tested the strong and weak scalability of the high-throughput based inverse calculation process, with a maximum computing scale of up to 38400 computing processes. The results are shown in Tables 3 and 4. The scaling is nearly linear with respect to the number of compute nodes. Especially for weak scaling, the efficiency is close to ideal conditions, except for little extra costs related to the calculation of the CSI and I/O issues.

Table 3. Strong scaling of the multi-node inverse algorithm (Each unit simulation cover 1 million air parcels. The same goes for Table 4)

Problem size	Number of processes N_{process}	Clock time/s	Strong speed-up ratio R_s	Efficiency (R_s/N_{process})
1600 unit simulations	2400	16775	1x	100%
	9600	4452	3.77x	94.25%
	38400	1138	14.74x	92.13%

Table 4. Weak scaling of multi-node inverse algorithm

Problem size	Number of processes N_{process}	Clock time/s	Weak speed-up ratio R_w	Efficiency (R_w/N_{process})
100 unit simulations	2400	1123	1x	100%
400 unit simulations	9600	1132	3.97x	99.25%
1600 unit simulations	38400	1138	15.79x	98.69%

In summary, the high-throughput based hybrid MPI/OpenMP parallel strategy of MPTRAC and the inverse algorithm show good strong and excellent weak scalability on the Tianhe-2 supercomputer. That means the inverse modeling system has high potential in massive parallel applications, meeting the requirements of real-time forecasts. However, the forward calculation still has some potential for optimization. In future work, we will investigate the possibility of cross-node computing with MPTRAC and try to further improve the single-node computing performance by using hyper threading. Besides, some further improvements may also be possible for the multi-node parallelization, in particular for the I/O issues and the efficiency of temporary file storage.

4 Case Study of the Nabro Volcanic Eruption

Following Heng et al. [21], we choose an eruption of the Nabro volcano, Eritrea, as a case study to test the inverse modeling system on the Tianhe-2 supercomputer. The Nabro volcano erupted at about 20:30 UTC on 12 June 2011, causing a release of about 1.5×10^9 kg of volcanic SO_2 into the troposphere and lower stratosphere. The volcanic activity lasted over 5 days with varying plume altitudes. The simulation results obtained for the Nabro volcanic eruption are of particular interest for studies of the Asian monsoon circulation [4, 24].

4.1 Reconstructed Emission Results with Different Resolutions

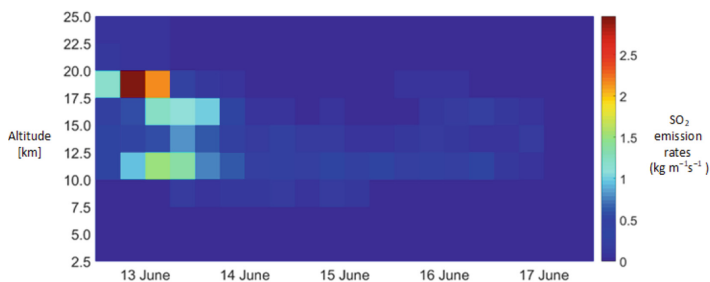
In general, with increasing resolution of the initial emissions, the forward simulation results are expected to become more accurate, but the calculation cost will also be much larger. In this work, the resolution of the volcanic SO₂ emission rates has been raised to 30 min of time and 100 m of altitude for the first time. The largest computing scale employs 60250 compute nodes on Tianhe-2 simultaneously. Each node calculated the kinematic trajectories of 1 million air parcels using a total of 24 cores. On such a computing scale, the inverse reconstruction and final forward simulation take about 22 min and require about 530,000 core hours in total.

Based on the inverse algorithm and parallel strategy described in Sects. 2 and 3, the SO₂ emission rates are reconstructed at different temporal and spatial resolutions, as shown in Fig. 2. The resolutions are (a) 6 h of time, 2.5 km of altitude, (b) 3 h of time, 1 km of altitude, (c) 1 h of time, 250 m of altitude, and (d) 30 min of time, 100 m of altitude. More fine structures in the emission rates become visible at higher resolution in Figs. 2a to 2c. However, the overall result in Fig. 2d at the highest resolution appears to be unstable with oscillations occurring between 12 to 16 km of altitude. The reason of this is not clear and will require further study, e.g., in terms of regularization of the inverse problem. For the time being, we employ the results in the Fig. 2c for the final forward simulation.

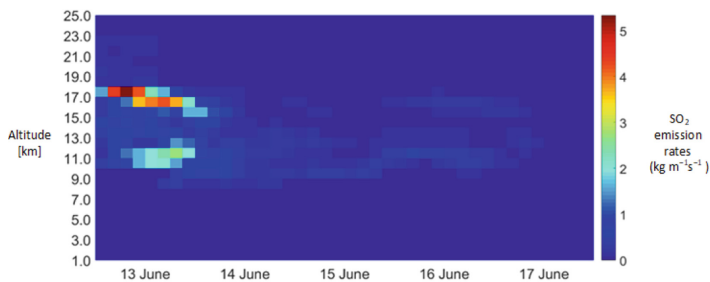
Compared with our previous work performed on the JuRoPA supercomputer at the Jülich Supercomputing Centre [21], the simulation results of this work performed on Tianhe-2 are rather similar. The reconstructed emissions show that the Nabro volcano had three strong eruptions on June 13, 14 and 16. For validation, Table 5 compares altitude and time of the major eruptions obtained with observations from different satellite sensors, which shows that the emission data constructed by the inverse modeling approach qualitatively agree with the measurements. Here, we also refer to the time series of the 2011 Nabro eruption based on Meteosat Visible and InfraRed Imager (MIVIRI) infrared imagery (IR) and water-vapor (WV) measurements, which were used as validation data sets in [21] as shown in Fig. 3.

Table 5. Major eruption altitudes of the Nabro volcano on different days

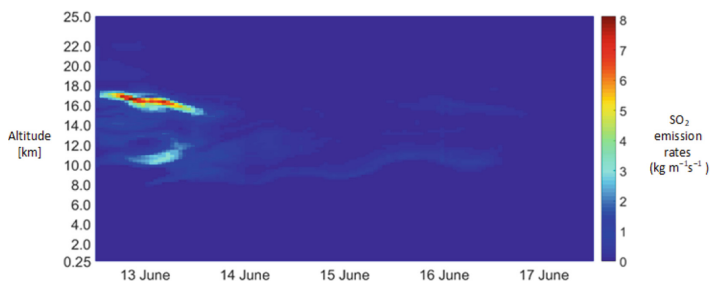
	June 13	June 14	June 16
CALIOP and MIPAS data	19 km	9–13 km	–
Fromm et al. (2013) [25]	15–19 km	–	–
Fromm et al. (2014) [26]	–	–	17.4 km
Result from this work	15–17 km	9–13 km	17 km



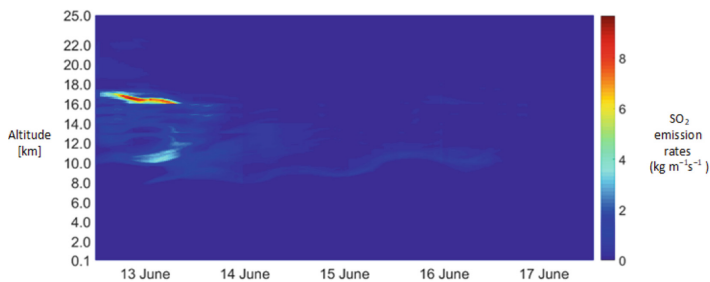
(a) Simulation on 210 compute nodes (5040 processes)



(b) Simulation on 1025 compute nodes (24600 processes)



(c) Simulation on 12100 compute nodes (290,400 processes)



(d) Simulation on 60250 compute nodes (1,446,000 processes)

Fig. 2. Reconstructed volcanic SO_2 emission rates of the Nabro eruption in June 2011. The x-axis refers to time, the y-axis refers to altitude (km), and the color bar refers to the emission rate ($\text{kg m}^{-1} \text{s}^{-1}$). (Color figure online)

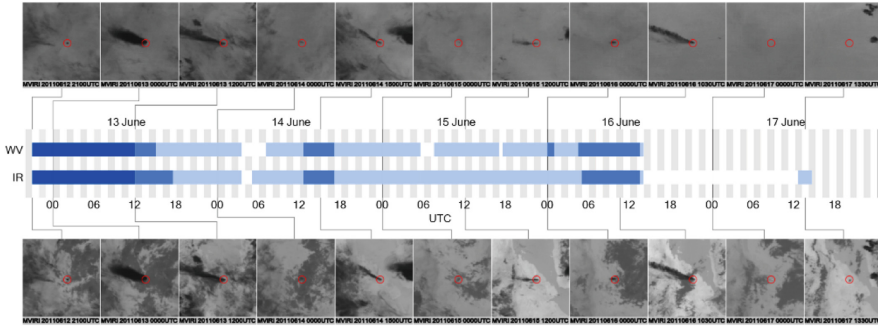


Fig. 3. Time line of the 2011 Nabro eruption based on MVIRI IR and WV measurements. Here white is none, light blue is low level, blue is medium level, dark blue is high level [21] (Color figure online)

4.2 Final Forward Simulation Results

Based on the reconstructed emission data with 1 h in time and 250 m in altitude resolution, the final simulation applying the product rule was conducted on the Tianhe-2 supercomputer for further evaluation. Figure 5 illustrates the simulated SO₂ transport, providing information on both, altitude and concentration, which are comparable to the AIRS observation maps shown in Fig. 4, suggesting the results are stable and accurate.

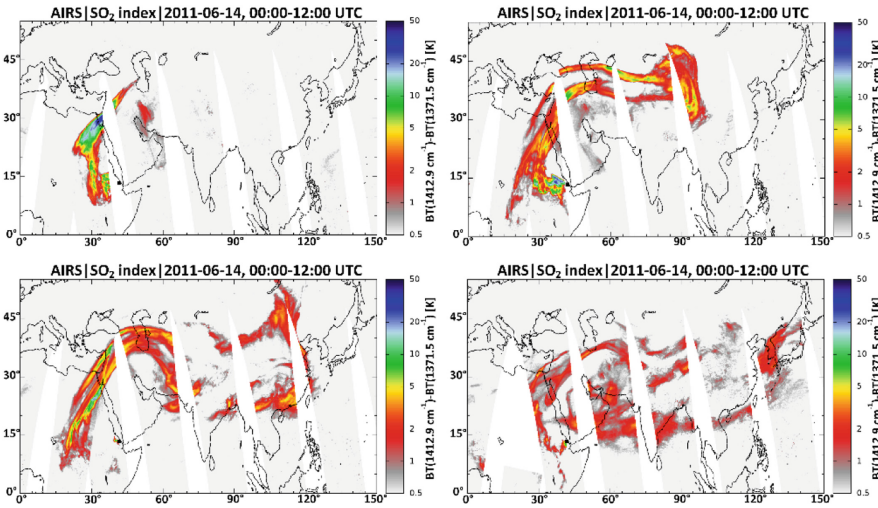
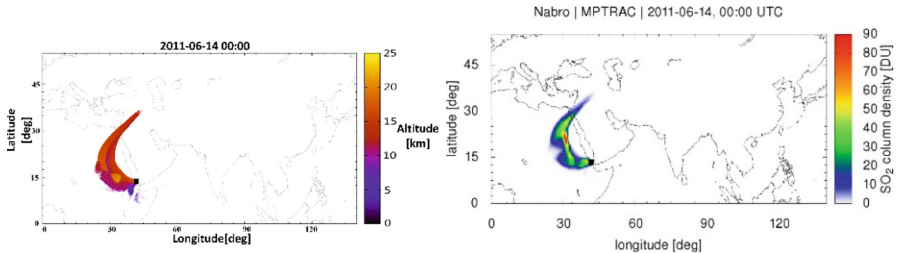
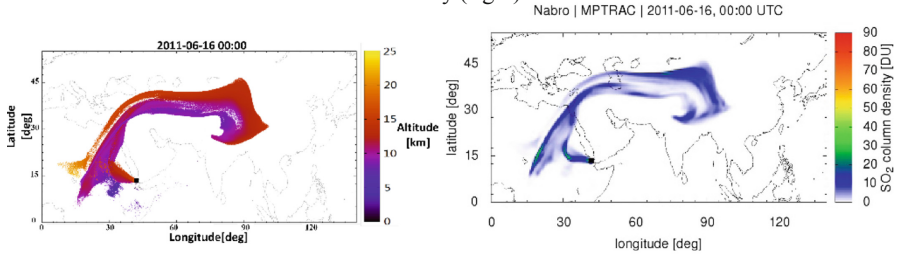


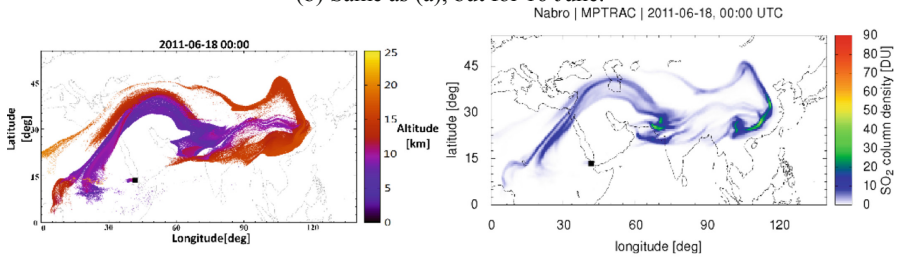
Fig. 4. The AIRS satellite observations on 14, 16, 18, 20 June 2011, 06:00 UTC (SO₂ index is a function of column density obtained from radiative transfer calculations. Here we refer to [19] for more detailed description of detection of volcanic emissions based on brightness temperature differences (BTDs) technique.)



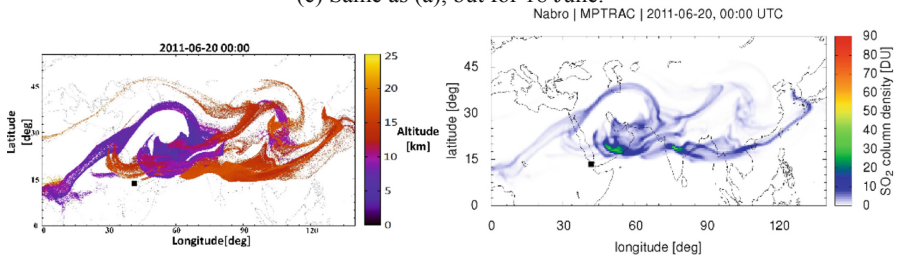
(a) Simulation results for 14 June: air parcels altitude distribution (left), SO_2 vertical column density (right).



(b) Same as (a), but for 16 June.



(c) Same as (a), but for 18 June.



(d) Same as (a), but for 20 June.

Fig. 5. Final forward simulation results of volcanic SO_2 released by the Nabro eruption. The black square indicates the location of the Nabro volcano.

5 Conclusions and Outlook

The high-resolution reconstruction of source information is critical to obtain precise atmospheric aerosol and trace gas transport simulations. The work we present in this

paper has potential applications for studying the effects of large-scale industrial emissions, nuclear leaks and other pollutions of the atmosphere and environment. The computational costs and efficiency of the inverse model will directly determine whether the atmospheric pollutant transport process can be predicted in real time or near real time. For this purpose, we implemented and assessed a high-throughput based inverse algorithm using the MPTRAC model on the Tianhe-2 supercomputer. The good scalability demonstrates that the algorithm is well suited for large-scale parallel computing. In our case study, the computational costs for the inverse reconstruction and final forward simulation at unprecedented resolution satisfy the requirements of real-time forecasts.

In the future work, we will study further improvements of the computational efficiency, e.g., multi-node parallel usage of MPTRAC, mitigation of remaining I/O issues, post-processing overhead, efficient storage of temporary files, etc. Also, some stability problems at the highest resolution problems need to be addressed, e.g., by means of regularization techniques. Nevertheless, we think that the inverse modeling system in its present form is ready to be tested in further applications.

Acknowledgements. The corresponding author Yi Heng acknowledges support provided by the “Young overseas high-level talents introduction plan” funding of China, Zhujiang Talent Program of Guangdong Province (No.2017GC010576) and the Natural Science Foundation of Guangdong (China) under grant no. 2018A030313288. Chunyan Huang is supported by the National Natural Science Foundation of China (No. 11971503), the Young Talents Program (No. QYP1809) and the disciplinary funding of Central University of Finance and Economics. Pin Chen is supported by the Key-Area Research and Development Program of Guangdong Province (2019B010940001). Part of this work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Projektnummer 410579391.

References

1. Brenot, H., Theys, N., Clarisse, L., et al.: Support to Aviation Control Service (SACS): an online service for near-real-time satellite monitoring of volcanic plumes. *Nat. Hazards Earth Syst. Sci.* **14**(5), 1099–1123 (2014)
2. Sigl, M., Winstrup, M., McConnell, J.R., et al.: Timing and climate forcing of volcanic eruptions for the past 2,500 years. *Nature* **523**(7562), 543–549 (2015)
3. Solomon, S., Daniel, J.S., Neely, R.R., et al.: The persistently variable “background” stratospheric aerosol layer and global climate change. *Science* **333**(6044), 866–870 (2011)
4. Bourassa, A.E., Robock, A., Randel, W.J., et al.: Large volcanic aerosol load in the stratosphere linked to Asian monsoon transport. *Science* **337**(6090), 78–81 (2012)
5. Hoffmann, L., Rößler, T., Griessbach, S., et al.: Lagrangian transport simulations of volcanic sulfur dioxide emissions: impact of meteorological data products. *J. Geophys. Res.-Atmos.* **121**(9), 4651–4673 (2016)
6. Lacasse, C., Karlsdottir, S., Larsen, G., et al.: Weather radar observations of the Hekla 2000 eruption cloud, Iceland. *Bull. Volcanol.* **66**(5), 457–473 (2004). <https://doi.org/10.1007/s00445-003-0329-3>
7. Stohl, A., Prata, A.J., Eckhardt, S., et al.: Determination of time- and height-resolved volcanic ash emissions and their use for quantitative ash dispersion modeling: the 2010 Eyjafjallajökull eruption. *Atmos. Chem. Phys.* **11**(9), 4333–4351 (2011)

8. Flemming, J., Inness, A.: Volcanic sulfur dioxide plume forecasts based on UV satellite retrievals for the 2011 Grimsvotn and the 2010 Eyjafjallajökull eruption. *J. Geophys. Res.-Atmos.* **118**(17), 10172–10189 (2013)
9. Molnar, F., Szakaly, T., Meszaros, R., et al.: Air pollution modelling using a Graphics Processing Unit with CUDA. *Comput. Phys. Commun.* **181**(1), 105–112 (2010)
10. Santos, M.C., Pinheiro, A., Schirru, R., et al.: GPU-based implementation of a real-time model for atmospheric dispersion of radionuclides. *Prog. Nucl. Energy* **110**, 245–259 (2019)
11. Müller, E.H., Ford, R., Hort, M.C., et al.: Parallelisation of the Lagrangian atmospheric dispersion model NAME. *Comput. Phys. Commun.* **184**(12), 2734–2745 (2013)
12. Larson, D.J., Nasstrom, J.S.: Shared- and distributed-memory parallelization of a Lagrangian atmospheric dispersion model. *Atmos. Environ.* **36**(9), 1559–1564 (2002)
13. Gordon, N.J., Salmond, D.J., Smith, A.F.M.: Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proc. F (Radar Signal Process.)* **140**(2), 107–113 (1993)
14. Dee, D.P., Uppala, S.M., Simmons, A.J., et al.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. Roy. Meteorol. Soc.* **137**(656), 553–597 (2011)
15. Bowman, K.P., Lin, J.C., Stohl, A., et al.: Input data requirements for Lagrangian trajectory models. *Bull. Am. Meteorol. Soc.* **94**(7), 1051–1058 (2013)
16. Stohl, A., Forster, C., Frank, A., et al.: Technical note: the Lagrangian particle dispersion model FLEXPART version 6.2. *Atmos. Chem. Phys.* **5**, 2461–2474 (2005)
17. Rößler, T., Stein, O., Heng, Y., et al.: Trajectory errors of different numerical integration schemes diagnosed with the MPTRAC advection module driven by ECMWF operational analyses. *Geosci. Model Dev.* **11**(2), 575–592 (2018)
18. Hoffmann, L., Hertzog, A., Rößler, T., et al.: Intercomparison of meteorological analyses and trajectories in the Antarctic lower stratosphere with Concordiasi superpressure balloon observations. *Atmos. Chem. Phys.* **17**, 8045–8061 (2017)
19. Hoffmann, L., Griessbach, S., Meyer, C.I.: Volcanic emissions from AIRS observations: detection methods, case study, and statistical analysis. In: *Proceedings of SPIE*, vol. 9242 (2014)
20. Schaefer, J.: The critical success index as an indicator of warning skill. *Weather Forecast.* **5**, 570–575 (1990)
21. Heng, Y., Hoffmann, L., Griessbach, S., et al.: Inverse transport modeling of volcanic sulfur dioxide emissions using large-scale simulations. *Geosci. Model Dev.* **9**(4), 1627–1645 (2016)
22. Che, Y., Yang, M., Xu, C., et al.: Petascale scramjet combustion simulation on the Tianhe-2 heterogeneous supercomputer. *Parallel Comput.* **77**, 101–117 (2018)
23. Fu, G.L., Lin, H.X., Heemink, A., et al.: Accelerating volcanic ash data assimilation using a mask-state algorithm based on an ensemble Kalman filter: a case study with the LOTOS-EUROS model (version 1.10). *Geosci. Model Dev.* **10**(4), 1751–1766 (2017)
24. Wu, X., Griessbach, S., Hoffmann, L.: Equatorward dispersion of a high-latitude volcanic plume and its relation to the Asian summer monsoon: a case study of the Sarychev eruption in 2009. *Atmos. Chem. Phys.* **17**(21), 13439–13455 (2017)
25. Fromm, M., Nedoluha, G., Charvat, Z.: Comment on “large volcanic aerosol load in the stratosphere linked to Asian monsoon transport”. *Science* **339**(6120), 647 (2013)
26. Fromm, M., Kablick, G., Nedoluha, G., et al.: Correcting the record of volcanic stratospheric aerosol impact: Nabro and Sarychev Peak. *J. Geophys. Res.-Atmos.* **119**(17), 10343–10364 (2014)