Development of C-TRAIL v1.0 model for Investigating Long-range Transport of Pollutants

Arman Pouyaei, Yunsoo Choi*, Jia Jung, Bavand Sadeghi
Department of Earth and Atmospheric Sciences, University of Houston, Houston, TX, USA

and Chul Han Song
School of Environmental Science and Engineering, Gwangju Institute of Science and Technology (GIST), Gwangju, South Korea

1. INTRODUCTION

Determining the long-range transport (LRT) of pollutants has been a challenge for air quality researchers. As the chemical composition of outflow over a region or continent can significantly affect air quality downwind, information about LRT must be reliable. Several studies have applied a number of methods to examine the role that LRT plays in the concentrations of particulate matter (PM), ozone, trace gases, and biomass burning tracers over target regions (Choi et al., 2014; Lee et al., 2019; Oh et al., 2015; Pu et al., 2015; Stohl, 2002). Spatial concentration patterns of pollutants incorporated with other models (i.e., back-trajectory models) or satellite data enhance our understanding of the impact of LRT and other related processes such as the formation of aerosols, emissions, and dry deposition in various regions (Xu et al., 2019; Zhang et al., 2019).

The conventional way of estimating potential source regions of air-mass transport is to use back-trajectory modeling. Frequently used for source-receptor linkage, such models combine their output with measurements of pollutant concentrations. As this source-receptor linkage approach uses meteorology-based models for back trajectories, it is not fully accepted because it is unable to directly determine whether an originated air mass is polluted or non-polluted (Lee et al., 2019). Thus, back-trajectory modeling sometimes provides unreliable information from which to assess the variation of pollutants at a receptor point, raising concern about its use for interpreting the contribution of the effect of LRT on concentrations of a target pollutant. In addition, other factors such as emissions and the local production of air pollutants contribute to variation in a target pollutant.

In this study, we implement a Lagrangian advection scheme that we refer to as the trajectory grid (TG) (Chock et al., 2005), into the Eulerian CMAQ v5.2 model. We introduce a new type of output from the Concentration Trajectory Route of Air pollution with the Integrated Lagrangian (C-TRAIL v1.0) stand-alone model in addition to CMAQ v5.2 output to simultaneously accomplish two objectives: (1) to provide a direct link between polluted air masses from sources and a receptor and (2) to provide the spatial concentration distribution of several pollutants that explains relevant physical processes. Chock et al. (2005) incorporated the TG into an air quality model to study the accuracy of this Lagrangian advection method over the Bott advection scheme applied in the Eulerian domain. One significant outcome of the TG model applied to CTMs is its ability to account for the concentrations of pollutants in air masses in its investigation of trajectories. This outcome addresses the unreliability of meteorology-based Lagrangian models when the pollutedness or cleanliness of an originated air mass becomes an issue. For this study, we have selected CO as our trace gas target. We begin by introducing the methodology behind TG and the implementation of TG into CMAQ. Then, we present a simple case and our interpretation of the C-TRAIL output. Finally, we present a case study of C-TRAIL for Korea and the United States Air Quality (KORUS-AQ) campaign over South Korea. More detailed results and illustrations regarding this study can be found in (Pouyaei et al., 2020).

2. Methodology

2.1 Description of the TG approach

The TG method rewrites the advection equation for concentration as follows:

$$\frac{dC}{dt} = \frac{\partial C}{\partial t} + \mathrm{v} \cdot \nabla C = -(\nabla \cdot \mathrm{v})C, \qquad (1)$$

where C is the concentration of species in velocity field v. Following this approach, the TG automatically and accurately conserves the mass,

**Corresponding author:* Yunsoo Choi, Department of Earth and Atmospheric Sciences, University of Houston, Houston, TX, USA; e-mail: ychoi6@uh.edu

sign, and shape of the concentration profile. As interpreted from the equation, the concentration profile of the species along trajectories can be described. Since we can use the TG method to calculate the concentration from an ordinary differential equation, it is mass conserving, monotonic, and accurate. Although interpolation errors occur during the diffusion step, but they are typically considerably smaller than Eulerian advection errors (Chock et al., 2005). In addition, the trajectory will be three-dimensional and as accurate as the input for wind velocity and direction. In particular, for large-scale vertical winds, in which CTMs typically modify the scheme to address the mass-conservation issue, TG removes numerical diffusion from upwind vertical advection schemes and generates more physical vertical winds (Chock et al., 2005). In this study, the Lagrangian points will be called "packets".

2.1 Implementation of TG into CMAQ v5.2

The C-TRAIL v1.0 model requires the same meteorology, initial conditions (ICs), boundary conditions (BCs), and emissions as CMAQ (Figure 1). All CMAQ modules and parameters are associated with cells of the Eulerian grid in the model domain. Since TG is based on CMAQ in this study and some of the CMAQ processes cannot be satisfactorily carried out by Lagrangian models (e.g., eddy diffusion) at this time, grid cells are the primary structure for initiating and listing packets. By grouping the packets into grid cells, keeping track of which packets are close to each other is easier.

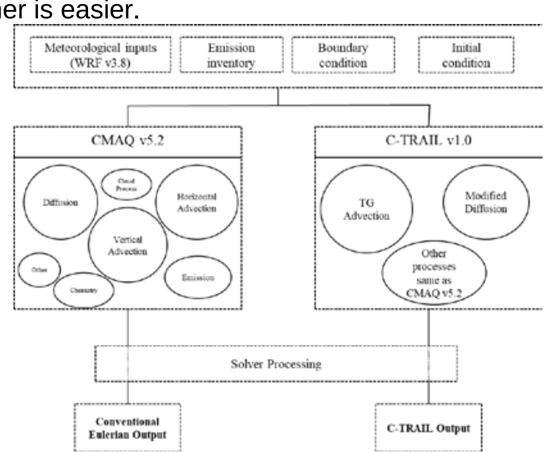

Figure 1: Schematic of conventional CMAQ versus C-TRAIL

When the initiated packets in the domain follow the Lagrangian equation, they land in different grid cells after each time step. To balance the density of packets in grid cells, we apply a simple packet management technique that includes spawning (filling) and pruning (emptying) processes. In the spawning process, every step entails the creation of a group of new packets in each cell with insufficient packets. The initial composition of a spawned packet is estimated from nearby packets. The pruning process entails the removal of extra packets from cells that have become overpopulated. During this process, the packets closest to the cell center are retained. Such packet management with favorable options contributes to reducing the computational costs of the C-TRAIL model. The limitation of this packet management approach, however, is that it violates mass conservation. These errors are caused by sub-grid interpolations of packets in the spawning or pruning process. The underlying algorithms for both vertical and horizontal diffusion, emissions, and other processes are the same as those in standard CMAQ with some minor modifications (Pouyaei et al., 2019).

3. Setup and Validation of the Model

In this study, we implement TG in the CMAQ model version 5.2. Shown in Figure 2, the model domain, with a horizontal grid resolution of 27-km over East Asia, covers the eastern parts of China, the Korean Peninsula, and Japan. Jung et al. (2019) validated the air quality model set up by comparing simulated and observed aerosol optical depths; they showed a correlation of 0.64 for the entire KORUS-AQ campaign period. Their comparison of various gaseous and particulate species also showed close agreement with observations.

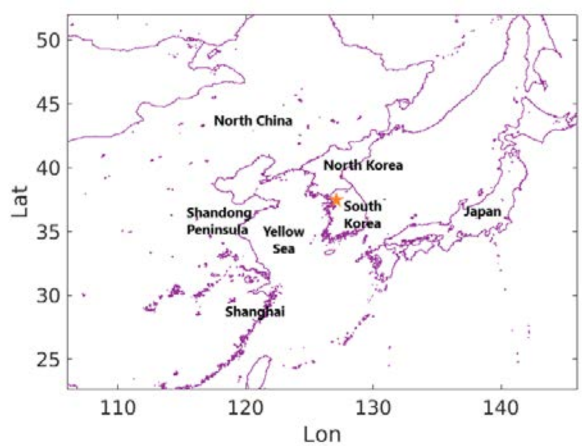

Figure 2: Domain of the study

We run C-TRAIL simulations for May 2016 during the KORUS-AQ campaign. Studies pertaining to this campaign have separated the time frame into three periods (Table 1) based on meteorological conditions: 1) the dynamic weather period (DWP), a rapid cycle of clear and rainy days in the Korean Peninsula (May 10-16); 2) the stagnant period (SP), in which the area was under the influence of a high-pressure system (May 17-22) and which showed the influence of local emissions; and 3) the extreme pollution period (EPP) with high peaks of pollutants that showed strong direct transport from China (May 25-28). The overall accuracy of the CMAQ CO simulation compared to aircraft measurements during all periods is presented in Table 1.

Table 1: Comparison of the statistical parameters of CMAQ CO concentrations to aircraft measurements

	Abbreviation	COR	IOA	RMSE	MAE
Entire month of May 2016		0.71	0.72	91.3	66.7
Dynamic Weather Period	DWP	0.72	0.62	81.5	66.2
Stagnant Period	SP	0.65	0.58	98.4	83.3
Extreme Pollution Period	EPP	0.89	0.88	68.7	47.7

The Eulerian output from CMAQ, including CO concentrations and surface wind fields, is displayed in Figure 3. High peaks of CO concentrations appeared in south-eastern China, including the Shanghai region and the Shandong Peninsula, because of high anthropogenic emissions in these areas. The impact on pollution from LRT was greater in this region because the dominant wind over East Asia in May was westerly, which explains our observations of high CO concentrations over the Yellow Sea.

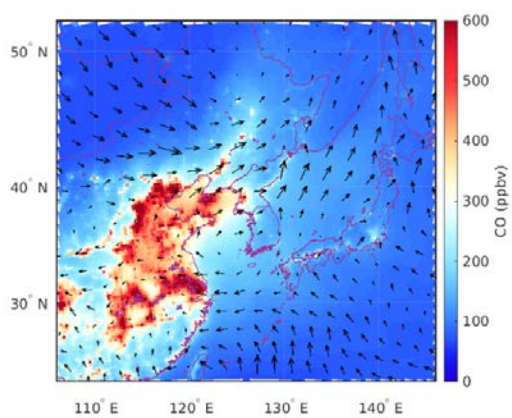

Figure 3: : CMAQ CO concentrations and wind patterns over the surface during the entire month of May 2016

The raw hypothesis from Eulerian outputs is that a high CO concentration at a receptor during a specific period is due to LRT from a source because the direction of the wind is typically toward the receptor during the period of simulation. This hypothesis is based on the average wind speed and direction and the average CO concentration, which do not constitute a reliable source of this assumption. We will briefly explain why we require merged output with simultaneous changes in trajectories and concentrations. To determine the source of LRT, researchers should include one major parameter in their investigations: the trajectory of the air mass. Once the location of the source and the trajectory of the air mass are known, the air mass is assumed to be polluted. If the air mass is not polluted, then that source is not responsible for high concentrations in the receptor location. Therefore, linking the source to the receptor based on only mean wind patterns and concentrations is not a reliable approach. The following section will discuss how we combine concentrations and trajectories into one set of outputs to explain the trajectories more clearly.

4. Case Study for the C-TRAIL Analysis: The May 2016 KORUS-AQ Period

For our case study, covering May 2016, we selected Seoul, South Korea, over East Asia as the receptor. We plotted C-TRAIL outputs according to variations in the packet concentrations and their distances from the receptor. Figure 4(a-b) presents a general and zoomed in path of all packet trajectories reaching the Seoul area at various altitudes at 9:00 AM local time throughout May 2016. The color bar represents the altitude at which the packets were traveling. Generally, packets at low altitudes traveled from local areas to Seoul, and those at high altitudes traveled from more distant regions. One exception was packets that originated in the Shandong Peninsula; Some traveled at high altitudes and some at low altitudes. Figure 4(c) displays a C-TRAIL that represented a unique type of packet that followed the concentrations of trajectories. In this case, each packet at each location (or hour of the trajectory) had a specific CO concentration that depended on its altitude (high altitude/ surface), its location (land/ sea /urban/ forest), and the hour of the day (traffic hours/ non-traffic hours). To more clearly explain the location of packets and the variability in their trajectory paths before reaching Seoul, we created a boxplot of packet distances in kilometers from the receptor at each hour before the packets reached Seoul, shown in Figure 4(d).

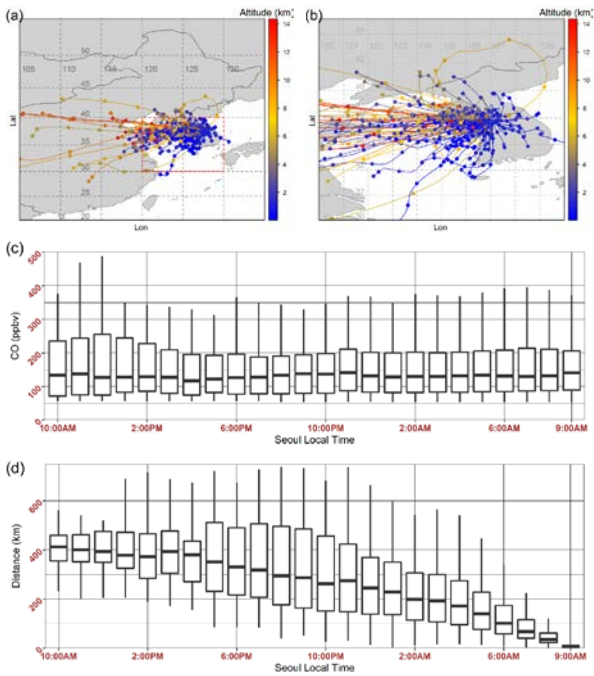

Figure 4: C-TRAIL output for the entire month of May 2016 for Seoul as the receptor: (a) 24-hour trajectories of packets for the entire domain, (b) 24-hour trajectories of packets for the zoomed area in South Korea, (c) the boxplot of the CO concentrations of all packets at each hour before they reached Seoul, and (d) the boxplot of packet distances from Seoul at each hour before the packets reached Seoul

In a study of C-TRAIL outputs, it is better to account for trajectories, concentrations, and distances simultaneously. As a result, the concentrations and distances of packets in early hours (10:00 AM to 2:00 PM of local time) in Figure 4 show high variability in concentrations with a median of around 150 ppbv and a maximum as high as 500 ppbv. Most of these packets originated far from the receptor (i.e., eastern, northern, and south-eastern China). The median concentration, shown in the boxplot, rose slightly between 6:00 PM and 10:00 PM. Distances also showed more variation during this time, which could be explained by the different paths of the trajectories (i.e., local trajectories with shorter distances and LRT trajectories with longer distances). As the packets approached Seoul (6:00 AM to 9:00 AM), the upper whisker of concentration values increased to as high as 400 ppbv, and the distances approached zero, indicating higher concentrations of CO of local trajectories resulting from surface on-road emissions and other emission sources.

Because of variable weather and wind (i.e., cloudy, rainy, or clear) during the DWP, C-TRAIL showed a mixed response of trajectories from both local and long-range transport, shown in Figure 5(a-b). A wide interquartile range and a median of close to the 25th percentile at 11:00 AM and 12:00 PM indicate that a few packets contained high concentrations of CO (close to 300 ppbv), but the majority consisted of low concentrations (around 100 ppbv). The distance output of low-concentration packets showed distances as long as 500 km (over the Shandong Peninsula). As the packets approached Seoul, the median concentration values were as high as 150 ppbv. Thus, from Figure 5, we conclude that most of the long trajectories followed a path at high altitudes (higher than 7 km), and the polluted trajectories, which originated in the Shandong Peninsula, were from the near-surface, shown in Figure 5(a).

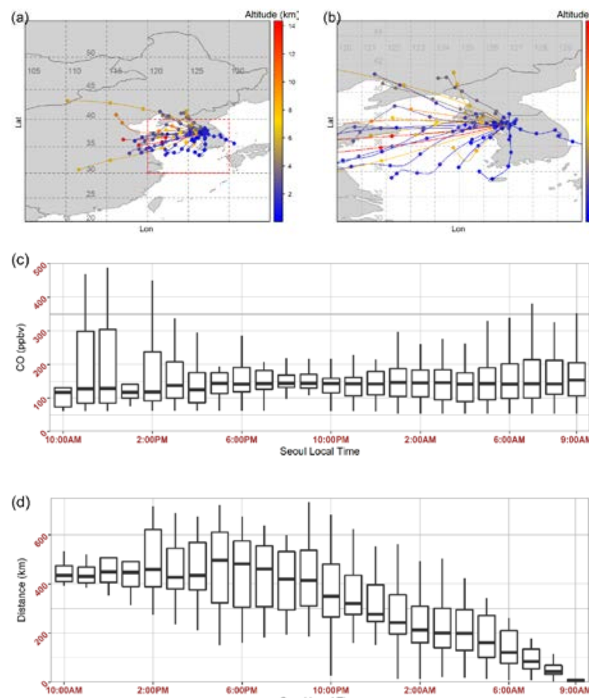

Figure 5: C-TRAIL output for the dynamic weather period (DWP) for Seoul as the receptor.

Unlike the DWP, the SP showed a more vivid display of trajectories, nearly all of which could be considered local trajectories. Long-range trajectories could not be considered responsible for the CO concentration values of Seoul. After all, from 10:00 AM to 4:00 PM (Figure 6), nearly all of the long-distance packets had concentrations of less than 100 ppbv. The local origination of highly polluted trajectories can be explained by a high-pressure system over the Korean Peninsula during this time period, which was responsible for very low wind speeds. The poor emission inventory over East Asia, however, provided extreme under-predictions of high concentration values during this time period. Therefore, when studying model outputs, we should account for various aspects of the model (e.g., the transport, diffusion, formation, deposition, and convention), in which diffusion, in this case, played a significant role in CO concentration values at the receptor location.

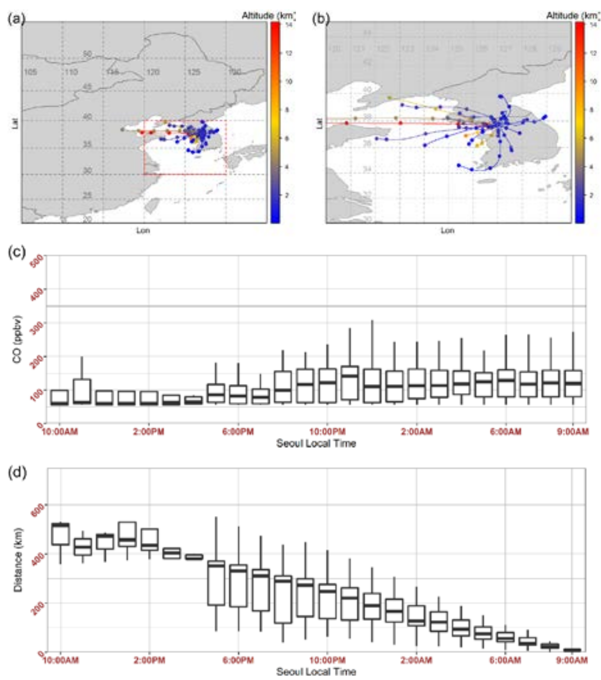

Figure 6: C-TRAIL output for the stagnant period (SP) for Seoul as the receptor.

During the EPP, several high concentrations of CO appeared at the early points of trajectories. These high concentrations, combined with high distance values, indicate that the LRT of polluted air masses was responsible for high concentrations of CO during this time period (Figure 7). Furthermore, the variability of CO concentrations from 10:00 PM to 9:00 AM at the receptor location stemmed from both the various paths of the trajectories and the distances. High concentration trajectories close to the surface, which originated in the Shandong Peninsula, passed over the Yellow Sea and landed in Seoul at 9:00 AM. When the surface packets reached urban areas, they presented maximum CO concentrations, depending on the time of day and the rush-hour traffic. An assumption made by studies that used Eulerian model outputs or meteorological-based Lagrangian models for this time period was that transport played an important role (Lee et al., 2019). The outputs from C-TRAIL also indicate that highly polluted air masses originated in China (the source) and landed in Seoul (the receptor). That is, the findings of this study regarding the trajectories and the origin of polluted air masses are similar to those of previous studies.

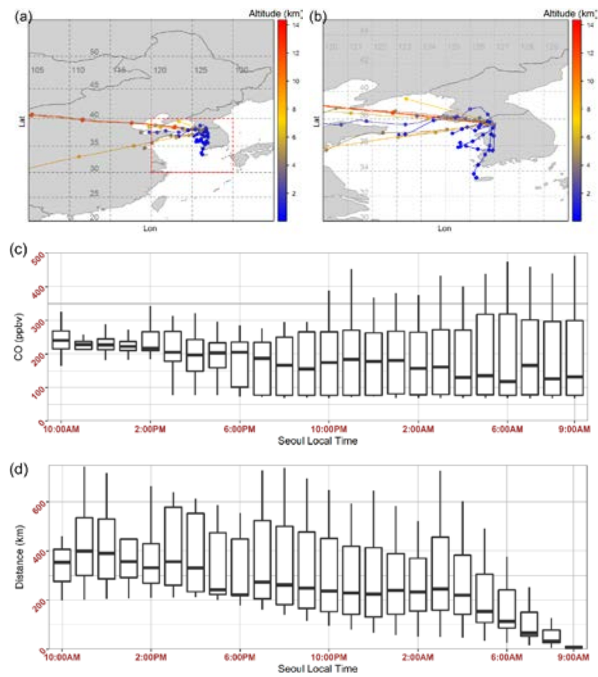

Figure 7: C-TRAIL output for the extreme pollution period (EPP) for Seoul as the receptor.

5. Conclusion

In this study, we introduced C-TRAIL Lagrangian output, extracted from the Eulerian CMAQ model. The comprehensive output of C-TRAIL directly linked the trajectories of pollution from the source to the receptor. We used concentration and trajectory values of C-TRAIL outputs to investigate the pollution status of originated air masses by classifying the outputs from May 2016 over East Asia into separate categories. Unlike the conventional Eulerian CO concentration plots for separate periods, which did not exhibit a clear relationship between the source and the receptor, the C-TRAIL outputs, which combined trajectories and concentrations, more vividly determined the impact of LRT on pollution during the EPP. Furthermore, during the dynamic weather period, C-TRAIL outputs showed that polluted packets from the Shandong Peninsula were responsible for high CO concentrations. The outputs for the SP revealed CO concentrations of less than 100 ppbv for distant packets, strong evidence supporting the link between local trajectories and CO concentrations over the SMA during this period. More comprehensive investigations on C-TRAIL outputs found that the Shandong Peninsula, local regions near the SMA, and the Pyongyang area were potentially strong sources of CO pollutants during the entire month of May 2016. Overall, by analyzing the trajectory paths of packets that reached specific locations, we were able to generalize that C-TRAIL represents a practical tool for ascertaining the impact of long-range transport on species concentrations over a receptor by simultaneously providing concentrations and trajectories. C-TRAIL can be applied to LRT-impacted regions such as East Asia, North America, and India. Owing to uncertainties inherent in emission inventories and immature diffusion modeling methods, however, C-TRAIL outputs may have limitations that we will address in future work. The objective of this study is to suggest an effective tool for establishing a link between real sources of pollution to a receptor via trajectory analysis. The results of this study over East Asia showed the reliability and various advantages of C-TRAIL output. Therefore, because of its capability to determine the trajectories of masses of CO concentrations, C-TRAIL output could prove to be a highly useful tool for those who model air quality over a specific region and investigate sources of polluted air masses.

6. Acknowledgments

We wish to acknowledge Dr. Peter Percell for his technical support in the development of CMAQ-TG in this research. This study was funded by the National Strategic Project-Fine particle of the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (MSIT), the Ministry of Environment (ME), and the Ministry of Health and Welfare (MOHW) (NRF-2017M3D8A1092022).

7. References

Chock, D. P., Whalen, M. J., Winkler, S. L. and Sun, P.: Implementing the trajectory-grid transport algorithm in an air quality model, Atmos. Environ., 39(22), 4015–4023, doi:10.1016/j.atmosenv.2005.03.037, 2005.

Choi, S. H., Ghim, Y. S., Chang, Y. S. and Jung, K.: Behavior of particulate matter during high concentration episodes in Seoul, Environ. Sci. Pollut. Res., 21(9), 5972–5982, doi:10.1007/s11356-014-2555-y, 2014.

Jung, J., Souri, A. H., Wong, D. C., Lee, S., Jeon, W., Kim, J. and Choi, Y.: The Impact of the Direct Effect of Aerosols on Meteorology and Air Quality Using Aerosol Optical Depth Assimilation During the KORUS-AQ Campaign, J. Geophys. Res. Atmos., 124(14), 8303–8319, doi:10.1029/2019jd030641, 2019.

Lee, S., Kim, J., Choi, M., Hong, J., Lim, H., Eck, T. F., Holben, B. N., Ahn, J. Y., Kim, J. and Koo, J. H.: Analysis of long-range transboundary transport (LRTT) effect on Korean aerosol pollution during the KORUS-AQ campaign, Atmos. Environ., 204(September 2018), 53–67, doi:10.1016/j.atmosenv.2019.02.020, 2019.

Oh, H. R., Ho, C. H., Kim, J., Chen, D., Lee, S., Choi, Y. S., Chang, L. S. and Song, C. K.: Long-range transport of air pollutants originating in China: A possible major cause of multi-day high-PM10 episodes during cold season in Seoul, Korea, Atmos. Environ., 109, 23–30, doi:10.1016/j.atmosenv.2015.03.005, 2015.

Pouyaei, A., Choi, Y., Jung, J. and Sadeghi, B.: Trajectory Grid: A Lagrangian Advection Algorithm Implemented into CMAQ, in TACCSTER 2019 Proceedings, Austin, TX, USA., 2019.

Pouyaei, A., Choi, Y., Jung, J., Sadeghi, B. and Song, C. H.: Concentration Trajectory Route of Air pollution with an Integrated Lagrangian model (C-TRAIL Model v1.0) derived from the Community Multiscale Air Quality Model (CMAQ Model v5.2), Geosci. Model Dev., 13(8), 3489–3505, doi:10.5194/gmd-13-3489-2020, 2020.

Pu, W., Zhao, X., Shi, X., Ma, Z., Zhang, X. and Yu, B.: Impact of long-range transport on aerosol properties at a regional background station in Northern China, Atmos. Res., 153, 489–499, doi:10.1016/j.atmosres.2014.10.010, 2015.

Stohl, A.: Computation, accuracy and applications of trajectories- a review and bibliography, Dev. Environ. Sci., 1(C), 615–654, doi:10.1016/S1474-8177(02)80024-9, 2002.

Xu, S., Warner, N., Bohlin-Nizzetto, P., Durham, J. and McNett, D.: Long-range transport potential and atmospheric persistence of cyclic volatile methylsiloxanes based on global measurements, Chemosphere, 228, 460–468, doi:10.1016/j.chemosphere.2019.04.130, 2019.

Zhang, L., Zhang, M. and Yao, Y.: Multi-Time Scale Analysis of Regional Aerosol Optical Depth Changes in National-Level Urban Agglomerations in China Using Modis Collection 6.1 Datasets from 2001 to 2017, Remote Sens., 11(2), 201, doi:10.3390/rs11020201, 2019.