EFFECTS OF DOWNSCALED HIGH-RESOLUTION METEOROLOGICAL DATA ON THE PSCF IDENTIFICATION OF EMISSION SOURCES

Meng-Dawn Cheng^{1*} and Erik D. Kabela²

¹ Environmental Sciences Division

² Nuclear Security and Isotope Technology Division

Oak Ridge National Laboratory, Oak Ridge, Tennessee, U.S.A.

ABSTRACT

The Potential Source Contribution Function (PSCF) model has been successfully used for identifying regions of emission source at a long distance in this study, the PSCF model relies on backward trajectories calculated by the Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT) model. In this study, we investigated the impacts of grid resolution and Planetary Boundary Layer (PBL) parameterization (e.g., turbulent transport of pollutants) on the PSCF analysis. The Mellor-Yamada-Janjic (MYJ) and Yonsei University (YUS) parameterization schemes were selected to model the turbulent transport in the PBL within the Weather Research and Forecasting (WRF version 3.6) model. Two separate domain grid sizes (83 and 27 km) were chosen in the WRF downscaling in generating the wind data for driving the HYSPLIT calculation. The effects of grid size and PBL parameterization are important in incorporating the influence of regional and local meteorological processes such as jet streaks, blocking patterns, Rossby waves, and terrain-induced convection on the transport of pollutants by a wind trajectory. We found high resolution PSCF did discover and locate source areas more precisely than that with lower resolution meteorological inputs. The lack of anticipated improvement could also be because a PBL scheme chosen to produce the WRF data was only a local parameterization and unable to faithfully duplicate the real atmosphere on a global scale. The MYJ scheme was able to replicate PSCF source identification by those using the Reanalysis and discover additional source areas that was not identified by the Reanalysis data. A potential benefit for using high-resolution wind data in the PSCF modeling is that it could discover new source location in addition to those identified by using the Reanalysis data input.

Notice: This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the US Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (<http://energy.gov/downloads/doe-public-access-plan>).

1. INTRODUCTION

Identifying emission sources is of great interest to international treaty for mitigation of climate change in the Arctic. There are instances where an emission source inventory does not exist and the source location is impractical to know *a priori*. Inversion technique is therefore undertaken for emission source identification and apportionment. The complexity of solving the inverse problem of source identification and attribution has been well-known to the field of environmental chemistry and air quality management (Hopke, 1985; 1991; Fleming et al., 2012) and nuclear forensics (Anderson-Cook, 2015), for example. In general, there is no single solution to the inversed problem.

To estimate source emission factors by an inversion technique, a large solution uncertainty could exist due to the condition imposed on the solution. For example, the number of sources (the unknowns) may be more than the number of measurement data (the knowns) that results in multiple possible solutions and large uncertainty about each solution. There may be cases where the number of sources is more than the number of available material signatures leading to an under-determined system of equations. In other words, there are more than one solution that can meet the needs of the equation system.

Alternatively, for identifying the geographic location of a source or sources (of material signatures) using an inverse atmospheric modeling approach, one may use backward wind trajectory. Among which the Potential Source Contribution Function (PSCF) model (Cheng 2014; Cheng and Lin, 2001; Lin et al., 2001) is one such approach, and the PSCF has been successfully used over the past two decades for air quality resources management at spatial scales ranging from the metropolitan scale (Gao et al., 1994; Cheng et al., 1993a) to continental (Cheng and Lin, 2001) and hemispherical (Cheng et al., 1993b; Cheng, 2014). The PSCF model derives a geographical map showing the probability of grid cells of known latitude and longitude coordinates. A grid cell of high PSCF value (> 0.8 for example) indicates the cell or area is very likely to be the source of the emitted material. At its present formulation, PSCF does not provide a quantitative measure of the emission factor of that identified grid cell.

It has been reported that the accuracy of a backward wind trajectory is dependent on the accuracy of available meteorological data (e.g., wind direction, wind speed, humidity, etc.) (Kahl and Samson, 1986 & 1988). The mechanics of computing the backward trajectory is straightforward, but the uncertainty at each endpoint can accumulate through each time step backward in time into an error so large that it becomes unrealistic at the end of a long (e.g., 10-day) backward trajectory. The accumulation of numerical errors in the model enumeration could be from factors such as input data errors due to instrument malfunction or out of calibration, lack of key meteorological data over a large area such as the Arctic or over the Atlantic and Pacific Ocean, representativeness of the model process parameterization and model grid resolution. Since a backward trajectory has to travel a large distance in a hemispheric PSCF analysis, investigation on the effects of the trajectory caused by uncertainty in the wind data, which is coupled to Planetary Boundary Layer (PBL) processes as well as the representation of them over a long distance (several thousand kilometers), is required.

For example, topography, type of surface elements (water body, vegetation, sand, ice, etc.), and man-made structures can all affect the three-dimensional movement of a trajectory backward in time in the atmosphere. These land attributes affect surface-atmosphere exchange of heat (sensible and latent) and momentum (mass like CO₂, water vapor, particulate matter and trace gases), which play an important role in the genesis and life cycle of boundary-layer turbulence controlling the vertical transport and mixing of pollutants. All these boundary layer processes are typically parameterized in so-called PBL schemes for a modern weather model. There are more than a dozen PBL schemes available in the Weather Research and Forecasting (WRF) model (Skamarock, 2008) ARW core.

Trajectory calculations using meteorological data are constrained by spatial and temporal resolution. Kahl and Samson (1986; 1988) have investigated the impacts on the trajectory of the uncertainty of meteorological data used. They suggested accuracy of the trajectory operating at a mesoscale can be improved by an increased spatial resolution. Regularization of the meteorological data is typically done through a process called Reanalysis such that meteorological data are in uniform format and quality controlled for further use. The Reanalysis data were provided in a spatial resolution on a model domain grid size of 250 km by 250 km.

Many micro- and meso-scale atmospheric processes; e.g., hurricane and vertical up-draft, that can significantly affect the trajectory movement but cannot be adequately simulated when a domain grid size is coarse. Thus, it is also likely that the trajectory endpoint can be effected by the domain grid size.

Since WRF-ARW was used to derive high-resolution wind data for driving the Hybrid Single-Particle Lagrangian Integrated Trajectory model (HYSPLIT; Draxler, 1999) backward trajectory calculation, we investigated the effects of model domain grid size and choice of PBL scheme on the accuracy of the computed backward trajectory. We compared the PSCF results that were based on the trajectories calculated from the WRF meteorological data, against the one performed by NCEP/NCAR Reanalysis used earlier in Cheng (2014) and showed how the effects eventually propagated and impacted the PSCF outcome and, therefore, the end results of source identification.

2. DESCRIPTION OF AMBIENT SAMPLING SITES AND AMBIENT DATA

Black carbon (BC) data, obtained through the Canadian NAtChem portal and used in the PSCF analysis, were from ambient air filter samples collected at Alert, Nunavut, Canada ([82.5°N, 62.3°W] at a height about 200 m above mean sea level) and analyzed by Environment Canada. Hirdman et al. (2010) described extensively the long-term trends of black carbon in the Arctic, and change in the atmospheric transport. **Figure 1** shows the location of the Alert sampling site in the Arctic Circle. The Alert station has a long, remarkable history contributing to the understanding of Arctic haze and of air chemistry in cold climates since the 60s (Barrie et al., 1967; Sharma, et al., 2004). The details of sample collection, analysis, and quality control of the NAtChem data can be found in Sharma et al. (2006) and will not be repeated here.

Figure 1. Map of the Arctic showing the location of the Alert Sampling Station (to the northwest tip of the Greenland). The blue dotted circle shows the Arctic circle at latitude of 60-degree N. The red curve enclosed the 10°C isotherm in July. Source: <http://www.arctic.noaa.gov/>

BC concentration from the year 2000 was measured by an aethalometer (McGee model AE31) on an hourly interval for this analysis. The BC data are displayed in a box-whisker plot for each

month in 2000 in **Figure 2**. The 75th percentile lines with the top of the box, the middle line is the 50th percentile (or median), and the bottom of the box is the 25th percentile of the hourly BC data. Data larger than the 95th percentile values are displayed as points outside the whisker. The seasonality of the BC data is clearly shown in **Figure 2**; winter or dark season has higher BC concentration, while summer or light season has lower BC concentration. This seasonal pattern is consistent with those for other Arctic Haze species reported in the past (e.g., Shaw, 1995; Barrie, 1986; Cheng et al., 1993b).

Figure 2. BC concentration in the year of 2000

3. DESCRIPTION OF SELECTED PBL SCHEMES

Turbulent mixing within the lower troposphere is needed to transport pollutants from emission sources to a distance far away. Turbulence is mostly a subgrid-scale process, but its presence in the PBL directly modulates a simulation's depiction of mass fields relevant for transport problems. Each scheme represents mixing on a local and/or nonlocal basis. Local schemes only

consider immediately adjacent vertical levels in the model, whereas nonlocal schemes can consider a deeper layer covering multiple levels in representing the effects of vertical mixing through the PBL. To investigate the effect of grid size and PBL on the computed trajectories, we chose two schemes: the MYJ (the Mellor-Yamada-Janjic) scheme (Janjic, 1994) and YSU (the Yonsei University) scheme (Hong et al., 2006). Both of these are available in the Weather Forecast and Research (WRF) model. The MYJ scheme is a 1.5-order closure scheme for prognosis of turbulent kinetic energy, while YSU is a first-order closure and represents entrainment at the top of the PBL explicitly. MYJ under-mixes PBL for locations up-stream convection (Coniglio et al., 2013), while YSU can more accurately simulate deeper vertical mixing in buoyancy-driven PBLs with shallower mixing in strong-wind regimes, and has been found to over-deepen the PBL for deep convective environments, resulting in too much dry air near the surface (Cohen et al., 2015).

Figure 3 shows two examples of the influence of PBL scheme on the resultant trajectories calculated using the same input meteorological data processed through WRF runs. All trajectories in **Figure 3** arrived at Alert on 1200 UTC on January 31, 2000. At either 27- or 82-km resolution, trajectories calculated using the WRF-generated wind field with different PBL scheme were significantly different. For instance, at 82-km resolution, the trajectory calculated with WRF MYJ data extended farther than that calculated with WRF YSU data. Difference was also found in the trajectories calculated at 27-km resolution using the two PBL schemes. Thus, it is of interest to PSCF modeling to study the impact of these effects on the final PSCF mapping results. In comparison, the trajectories calculated from the Reanalysis data set were at 250-km resolution.

5. DESCRIPTION OF HYSPLIT TRAJECTORY DATA

Our prior experience with PSCF analysis (Cheng et al., 1993a, 1993b; Cheng et al., 1996; Cheng and Schroeder, 2000; Cheng and Lin, 2001) suggests that for a trajectory to be able to traverse the distance enabling source identification at this large spatial scale from Alert, Canada, a large number of 5- to 10-day back trajectories would be needed. Thus, 10-day three-dimensional backward reference trajectories were chosen and calculated using the HYSPLIT model in a terrain-following model-vertical-motion mode using reanalyzed meteorological data. The trajectories were calculated at a 6-h interval (00, 06, 12, and 18Z), four times a day on a 2.5° by 2.5° latitude and longitude grid. The arrival and sampling inlet heights at Alert are set at 210 m above ground level.

Previous research has found that the further back in time and the longer a trajectory traveled, the higher the endpoint uncertainty (Kahl and Samson, 1986 and 1988). For instance, a boundary layer trajectory could have a spatial displacement of 350 km after 72 hours of travel time. Increased spatial resolution in the meteorological data used in trajectory calculations would improve the accuracy of trajectory.

6. DESCRIPTION OF PSCF ANALYSIS TECHNIQUE

PSCF analysis yields a two-dimensional map that shows a probability field describing the source strength of a geographical area (i.e., a grid cell), called the “potential source contribution.” PSCF modeling is briefly described as follows: let N equal the total number of trajectory endpoints that are counted in the entire modeling domain, and n_{ij} equal the total number of endpoints that fall in the ij^{th} grid cell during time interval T . The “cumulative probability” that a selected air parcel resides in the ij^{th} cell is $P[A_0] = n_{ij}/N$. This probability represents the chance that BC was transported to the receptor site when trajectories passed through the ij^{th} cell. An event would be measured as “dirty” if the event signal received at the receptor were above a “threshold” level.

If a criterion value is chosen above which a measurement at the receptor is considered “significant,” then the cumulative probability of these exceeding value events, B_0 , can be

determined as $P[B_0] = m_{ij} / N$, where m_{ij} is the total “polluted” counts for the ij^{th} cell. Given $P[A_0]$ and $P[B_0]$, PSCF is calculated as a ratio statistic as follows:

$$PSCF = \frac{P[B_0]}{P[A_0]} = \frac{\left(\frac{m_{ij}}{N}\right)}{\left(\frac{n_{ij}}{N}\right)} = \frac{m_{ij}}{n_{ij}} . \quad (1)$$

The final gridded PSCF values were weighted using the approach that has been applied to previous studies (e.g., Lin et al, 2001; Cheng and Lin, 2001). The interpretation of the PSCF is that the computed value indicates the likelihood that a grid cell is an emission source. Thus, a grid cell having a value of 0.8, for example, is more likely than one having a value of less than 0.2 to be an emission source for the chemical species under investigation. In our analysis, a 6-h averaged BC concentration above the 75th percentile for each month was considered a signal for Alert, Canada.

7. RESULTS AND DISCUSSION

7.1. Effects of Domain Grid Size and chosen PBL Scheme on the Calculated Trajectory

For all the trajectories calculated using different PBL schemes and grid domain sizes, there were significant differences in terms of direction and range covered in a 10-day journey. As discussed earlier, we arbitrarily selected two cases to exemplify our observation in **Figure 3** that shows two panels of six 10-day backward trajectories each calculated by HYSPLIT using different wind data inputs. The input data include those obtained from NCAR/NCEP Reanalysis (for 2.5° by 2.5° resolution) or by WRF downscaling for 82- and 27-km grid domains (**Figure 4**). In the left panel (**Figure 3[a]**), the PBL scheme for the 82- and 27-km domains was MYJ, while the YSU scheme trajectories are displayed in the right panel (**Figure 3[b]**). The trajectory obtained with Reanalysis data for January 31, 2000 at 1200 UTC is displayed in green, the 82 km domain results are shown in blue, and the last trajectory (27-km domain) is displayed in red.

It is clear that the computed trajectories were strongly influenced by the grid size as well as the PBL scheme chosen in the wind field reconstruction. For example, with the 82-km resolution runs, the backward trajectory calculated based on the wind data generated by invoking the MYJ

parameterization lifted off from the receptor site almost immediately and subsequently traversed a substantially larger area than the trajectory calculated based on the YSU scheme, which stayed below 500 m AGL for the first 3 days. The traverse distance covered by the two trajectories constructed from the 27-km grid WRF domain wind data are different in the vertical direction based on the two PBL schemes. The trajectory based on the MYJ scheme circulated on the North American side of Greenland for the first 10-day period, while the trajectory based on the YSU scheme followed a different path from day 1, and was able to travel directly toward Russia.

Our interest was to learn the difference in the final PSCF outcomes that could be caused by the difference in the PBL schemes and grid resolutions over a long period of time, it might be of interest to investigate the causes of differences in a separate paper. Furthermore, using the same PBL parameterization, we also noted significant differences in the calculated trajectories (see **Figures 3[a] for MYJ and 3[b] for YSU**). The single contributing factor causing the differences was the grid resolution used in the WRF downscaling to generate wind data for the HYSPLIT model. Interestingly, none of the trajectories in **Figure 3** replicated the trajectory calculated based on the Reanalysis data, which had a lower grid resolution compared to the two downscaled data sets. Also, no PBL parameterization was invoked in producing the Reanalysis data set because the sub-grid processes could not be reproduced effectively at the 250-km scale.

In the following Sections, we show the PSCF results obtained using the HYSPLIT trajectories calculated by the WRF meteorological data obtained from different domain grid resolution and choice of PBL scheme. We denote the PSCF result obtained from using the MYJ scheme at 82-km resolution as “MYJ at 82-km”, the same scheme at 27-km as “MYJ at 27-km”. The results for YSU are then “YSU at 82-km” and “YSU at 27-km” for those obtained using the HYSPLIT trajectories calculated from the WRF meteorological data under the YSU scheme and at the two resolutions. The effects of WRF model domain grid resolution on using MYJ are shown in **Figure 4** and discussed in the Section 7.2, while the effects on using YSU are shown in **Figure 5** and discussed in the Section 7.3. The effects of PBL scheme given the same domain grid resolution on the outcome of PSCF model are discussed in the Section 7.4.

7.2. Effects of Domain Grid Resolution on PSCF Results Using MYJ Scheme

Give a PBL scheme, the problem of grid resolution on the accuracy of PSCF results is investigated by comparing the PSCF map using the NCEP/NCAR Reanalysis data with the two maps produced using the WRF meteorological data modeled at the 82km and 27km. Three maps displayed in **Figure 4** are [a] NCEP/NCAR, [b] MYJ at 82-km, and [c] MYJ at 27-km. Since the map shown in **Figure 4[a]** does not reflect any possible atmospheric transport processes occurring within a 2.5° by 2.5° latitude by longitude grid, we expect the results shown in **Figure 4[b]** and or **4[c]** to be different. Also note that the crosses “+” in **Figure 4** mark the “known” locations of power plants.

Figure 4[a] shows the PSCF map produced using the NCAR/NCEP Reanalyzed data and the 75th percentile cut in the BC data displayed in **Figure 2**. Again, the **Figure 4[a]** map shows the “PSCF source potential” of a grid cell (250km x 250km in this case) where the PSCF value greater than 0.8 indicating the grid cell is highly likely to be an emission source of the measured black carbon. Based on this map, the region surrounding Moscow (particularly to the south) was identified as medium to high potential in **Figure 4[a]**, which is consistent with that found by Cheng (2014) and Huang et al. (2014). Also, note that area south of Murmansk was identified as of medium source potential and another big city, St. Petersburg, was found to be of low source potential based on the Reanalysis meteorological data input.

There is a larger number of grid cells with PSCF values greater than 0.8 from the Central Siberia to the Far East. Also note that all but a couple of these high PSCF grid cells did not cover known power plants marked on the map. Since **Figure 4[a]** used only 1-year worth of trajectory data shorter than 7 years used in the previous paper (Cheng, 2014), the distribution of high PSCF-value cells in **Figure 4[a]** suggests that some of the “hot” grid cells might be annually transient. In other words, the PSCF value of these cells could be down-averaged when a source was not emitting persistently over time. When combined with other source signature markers such as potassium as did in Cheng (2014) and Huang et al. (2014), we suggested that these Siberian BC sources could have been forest fires and biomass burning occurring over a short period of time.

When comparing **Figure 4[b]** and **4[c]**, distinct features from the map in **Figure 4[a]** can be found. The first feature is the extent of trajectory reach; i.e., the PSCF colors in **Figures 4[b]** and **4[c]** were further south than that in **Figure 4[a]** indicating that the 10-day trajectories calculated using high-resolution WRF data could reach areas located further south of those that reached by using the Reanalysis data. Those in the 27-km map (**Figure 4[c]**) reached close at 30°N, a 10 degrees further south of those in the Reanalysis map (**Figure 4[a]**). Given all factors (including the PBL scheme) but the grid resolution fixed in the WRF downscaling exercises, we concluded that trajectories could reach farther away from the receptor as the grid resolution for meteorological data increased.

As the grid resolution increased from 250 km (**Figure 4[a]**) to 82 km (**Figure 4[b]**), many Central Siberian cells' potentials were weakening and the corresponding PSCF values decreased from 0.8 and above category to that of less than 0.2. Similar PSCF pattern was observed in the 27-km map. However, many grid cells in the European Russia region remain highly probably with the PSCF values greater than 0.8 for all three grid resolutions. The grid cells at the southeastern of Moscow were found to be in the highest PSCF value category. Murmansk and St. Petersburg were identified as emission sources with high PSCF value using 82-km meteorological data, which was not identified at all by using the Reanalysis data. Note that the grid cell potential (i.e., PSCF value) for these three big cities in Russia decreased as the meteorological data resolution increased again from 82-km to 27-km. The only exception is the area in Murmansk. As stated previously, we think that it is likely those grid cells in Central Siberian were temporary; e.g., forecast fire, agricultural burning, and etc. So when the grid resolution was increased as WRF modeling down-scaled, the PSCF model lost the signal to identify these transient emission sources.

Figure 4. PSCF Results using the Reanalysis data [a], WRF downscaled to 82-km [b], and WRF downscaled to 27-km [c]. Both WRF data sets were generated using the MYJ PBL scheme in the computer runs. The crosses “+” dotted in the maps are the power plant locations.

7.3. Effects of Domain Grid Resolution on PSCF Results Using YSU Scheme

We investigated the effects of grid resolution on PSCF results using the YSU scheme in this Section. The three maps displayed in **Figure 5** are [a] the PSCF map using the Reanalysis meteorological data, [b] YSU-82km, and [c] YSU-27km. In other words, the only variable in this study was the grid resolution of the WRF meteorological data used in the HYSPLIT trajectory calculations. It is important to note that **Figure 5[a]** is identical to **Figure 4[a]**. It is displayed here again only for easing comparison with **Figure 5[b]** and **5[c]**, no further discussion of **Figure 5[a]** will be made nor repeated here. The discussion was found in the Section 7.2 for **Figure 4[a]**.

Again, we saw the grid cells of high PSCF values in the Central Siberia disappeared from the maps created as the WRF modeling was down-scaled (i.e., **Figure 5[b]** and **5[c]**). This observation of high PSCF-valued Siberia grid cells disappeared in **Figure 5[b]** and **5[c]** is consistent with those shown in **Figure 4**. Some of the grid cells in the European Russia retain their high PSCF values using downscaled WRF meteorological input data. For example, at the 82-km resolution, many grid cells over Europe were found to be of high potential (**Figure 5[b]**). These areas were not identified in **Figure 5[a]** when the Reanalysis data were used. Also, the Murmansk and St. Petersburg areas were not identified in **Figure 5[b]** as the grid cells of high emission potential. The greater Moscow area was identified as of high PSCF value. As WRF further down-scaled from the 82-km resolution to 27-km, all the European grid cells lost their high PSCF values (i.e., white cells). This indicates that the trajectories of YSU at 27-km simply did not traverse these areas, which were different from any of the cases present so far. A feature in **Figure 5[c]** that is distinct from any other maps is the grid cells of high PSCF values East of the Ural Mountain Range in Russia.

7.4. Comparison of PBL Schemes

The results using the two PBL schemes on the PSCF modeling using the same WRF-downscaled data at the 27-km resolution are compared with **Figures 4[c]** and **5[c]**. **Figure 4[c]** is MYJ-27km and **Figure 5[c]** is YSU-27km. Most of the high PSCF grid cells are located in the

379 western end of the maps. The MYJ-27km scheme produced high potential grid cells in two big
380 regions: (1) the northern Scandinavia including Murmansk area and (2) the Urals Mountain area,
381 while the YSU-27km's high potential grid cells are located at (1) the St. Petersburg, (2) the
382 Moscow area, and (3) the Urals Mountain. Thus, the comparison shows only the Urals
383 Mountains area were identified by both PBL scheme at the 27-km scale. If the results using the
384 Reanalysis data are compared, we found that MYJ-27 km show more consistent identification
385 because the Murmansk area, southern part of Scandinavia, Moscow, and the Urals Mountains
386 were all of high PSCF values in the Reanalysis.

387
388 At the 82-km scale, MYJ-82km found (1) Murmansk, (2) St. Petersburg, (3) southern
389 Scandinavia, (4) northern central Europe, (5) north of Moscow, and (6) the Urals Mountains of
390 high PSCF values (i.e., greater than 0.6). YSU-82km found (1) southern Scandinavia, (2) most
391 of the mainland Europe, and (3) west of Moscow were the areas of high PSCF values. Again,
392 when comparing with the Reanalysis results, MYJ-82km appear to produce more consistent
393 identification than YSU-82km.

Figure 5. PSCF Results using the Reanalysis data [a], WRF downscaled to 82-km [b], and WRF downscaled to 27-km [c]. Both WRF data sets were generated using the YSU PBL scheme in the computer runs. The crosses “+” dotted in the maps are the power plant locations.

CONCLUSIONS AND DISCUSSION

Our original hypothesis was that increased model grid resolution and inclusion of PBL scheme in the HYSPLIT trajectory calculation could improve the PSCF analysis and hopefully lead to a more accurate source identification. We found that hypothesis was questionable. A PSCF map is a constructed ensemble result of many trajectories integrated over a long time scale; i.e., 365 days in this study. The final map is a direct convolution of the backward transport over a long time span and variation in the source emission time profile. Thus, increasing the grid resolution alone for the WRF downscaled data did not linearly improve the PSCF identification. However, high resolution PSCF did discover and locate source areas more precisely than that with lower resolution meteorological inputs. The lack of anticipated improvement could also be because a PBL scheme chosen to produce the WRF data was only a local parameterization and unable to faithfully duplicate the real atmosphere on a global scale. The MYJ scheme was able to replicate PSCF source identification by those using the Reanalysis and discover additional source areas that was not identified by the Reanalysis data. A potential benefit for using high-resolution wind data in the PSCF modeling is that it could discover new source location in addition to those identified by using the Reanalysis data input.

DISCLAIMERS

The view and opinions expressed in this manuscript are solely of the authors and are not of the Oak Ridge National Laboratory nor the Department of Energy. Any mention of instrument brand names, chemicals, software, or commercial merchandises does not represent endorsement by the authors or the organization the authors are associated with.

ACKNOWLEDGEMENTS

This work was unfunded and performed by the authors at their own time, there were no conflicts of interest. The authors acknowledge the Canadian Environment Canada for making the Alert

data available on the NAtChem web site (<https://www.ec.gc.ca/natchem/>) and the Air Resources Laboratory of the National Oceanic and Atmospheric Administration for making the HYSPLIT model available online at <http://ready.arl.noaa.gov/HYSPLIT.php>. Oak Ridge National Laboratory is managed by UT-Battelle, LLC, for the U. S. Department of Energy under Contract No. DE-AC05-00OR22725.

REFERENCES

- Anderson-Cook, C.M., T. Burr, M.S. Hamada, and E.V. Thomas (2015) Statistical Analysis for Nuclear Forensics Experiments, *Stat. Anal. Data Min., The ASA Data Sci. J.*, 8(5-6): 364-377.
- Barrie, L.A. (1985) Arctic Air Pollution: An Overview of Current Knowledge, *Atmos. Environ.*, 20(4): 643-663.
- Cheng, M.-D., P. K. Hopke, and Y. Zeng (1993a) A Receptor-Oriented Methodology for Determining Source Regions of Particulate Sulfate Observed at Dorset, Ontario, *J. Geophys. Res.-Atmos.*, 98(D9): 16,839.
- Cheng, M.-D., P. K. Hopke, A. Rippe, L. Barrie, M. Olson, and S. Landsberger (1993b) Qualitative Determination of Source Regions of Aerosol in Canadian High Arctic, *Environ. Sci. Technol.*, 27(10): 2,063.
- Cheng, M.-D. and P. K. Hopke (1986a) Investigation on the Use of Chemical Mass Balance Receptor Model: Numerical Computations, *Chemo. Intel. Lab. Sys.*, 1: 33.
- Cheng, M.-D. and P. K. Hopke (1986b) Investigation on the Use of Chemical Mass Balance Receptor Model: Numerical Computations - Responses, *Chemo. Intel. Lab. Sys.*, 1(1): 48.
- Cheng, M.-D., N. Gao, and P. K. Hopke (1996) Source Apportionment Study of Nitrogen Species Measured in Southern California in 1987, *J. Environ. Engr./ASCE*, 122(3): 183.
- Cheng, M.-D. and W. H. Schroeder (2000) Potential Atmospheric Transport Pathway for Mercury Measured in the Canadian High Arctic, *J. Atmos. Chem.*, 35(1): 101-107.
- Cheng, M.-D. and C. J. Lin (2001) Receptor Modeling for Smoke of 1998 Biomass Burning in Central America, *J. Geophys. Res. - Atmos.*, 106(D19): 22,871-22,886.
- Cheng, M.-D. (2014) Geolocating Russian Sources for Arctic Black Carbon, *Atmos. Environ.*, 92(August): 398-410. <http://dx.doi.org/10.1016/j.atmosenv.2014.04.031>
- Cohen, A.E., S.M. Cavallo, M.C. Coniglio, and H.E. Brooks (2015) A Review of Boundary Layer Parameterization Schemes and Their Sensitivities in Simulating Southeastern U.S. Cold Season Severe Weather Environments, *Wea. Forecasting*, 30(6): 591-612, DOI: 10.1175/WAF-D-14-00105.1.

- Coniglio, M.C., J. Correia, P.T. Marsh, and F. Kong (2013) Verification of Convection-Allowing WRF Model Forecasts of the Planetary Boundary Layer Using Sounding Observations. *Wea. Forecasting*, 28, 842–862, doi:[10.1175/WAF-D-12-00103.1](https://doi.org/10.1175/WAF-D-12-00103.1).
- Draxler, R. R. (1999) *HYSPLIT4 User's Guide*, NOAA Tech. Memo. ERL ARL-230, NOAA Air Resources Laboratory, Silver Spring, MD.
- Fleming, Z.L., P.S. Monks, and A.J. Manning (2012) Review: Untangling the influence of air-mass history in interpreting observed atmospheric composition, *Atmos. Res.*, 104-105, 1-39, doi: [10.1016/j.atmosres.2011.09.009](https://doi.org/10.1016/j.atmosres.2011.09.009)
- Gao, N., M.-D. Cheng, and P. K. Hopke (1994) Receptor Modeling of Airborne Ionic Species Collected in SCAQS, *Atmos. Environ.*, 28(8): 1,447-1,470.
- Hines, K.M. and D.H. Bromwich (2008) Development and Testing of Polar Weather Research and Forecasting (WRF) Model. Part I: Greenland Ice Sheet Meteorology, *Monthly Weather Revw.*, 136, 1971-1989, doi: [10.1175/2007MWR2112.1](https://doi.org/10.1175/2007MWR2112.1)
- Hirdman, D., J. F. Burkhardt, H. Sodemann, S. Eckhardt, A. Jefferson, P. K. Quinn, S. Sharma, J. Ström, and A. Stohl (2010) Long-Term Trends of Black Carbon and Sulphate Aerosol in the Arctic: Changes in Atmospheric Transport and Source Region Emissions, *Atmos. Chem. Phys.*, **10**: 9,351-9,386.
- Hong, S.-Y., Y. Noh, and J. Dudhia (2006) A New Vertical Diffusion Package with an Explicit Treatment of Entrainment Processes, *Mon. Wea. Rev.*, **134**, 2318-2341.
- Hopke, P.K. (1985) *Receptor Modeling in Environmental Chemistry*, John Wiley & Sons, New York, NY.
- Hopke, P.K. (1991) *Receptor Modeling for Air Quality Management*, Elsevier Pub., New York, NY.
- Huang, K., J.S. Fu, E.L. Hodson, X. Dong, J. Cresko, V.Y. Prikhodko, J.M. Storey, and M.-D. Cheng (2014) Identification of Missing Anthropogenic Emission Sources in Russia: Implication for Modeling Arctic Haze, *J. Aerosol Air Qual. Res.*, 14 (12): 1799-1811, doi: [10.4209/aaqr.2014.08.0165](https://doi.org/10.4209/aaqr.2014.08.0165)
- Janjic, Z. I. (1994) The Step-Mountain Eta Coordinate Model: Further Developments of the Convection, Viscous Sublayer, and Turbulence Closure, *Mon. Wea. Rev.*, **122**, 927-945.
- Kahl, J.D. and P.J. Samson (1986) Uncertainty in Trajectory Calculations Due to Low Resolution Meteorological Data, *J. Climate Appl. Meteo.*, 25(12): 1816-1831.
- Kahl, J.D. and P.J. Samson (1988) Trajectory Sensitivity to Rawinsonde Data Resolution, *Atmos. Environ.*, 22(7): 1291-1299.

- 523 Lin, C. J., M.-D. Cheng, and W. H. Schroeder (2001) Transport Patterns and Potential Sources of
524 Total Gaseous Mercury Measured in the Arctic in 1995, *Atmos. Environ.*, 35(6): 1141-1154.
525
- 526 Sharma, S., Lavoue, D., Cachier, H., Barrie, L., and Gong, S.: Long-Term Trends of the Black
527 Carbon Concentrations in the Canadian Arctic, *J. Geophys. Res.*, **109**, D15203-D15213,
528 doi:10.1029/2003JD004331, 2004.
- 529 Shaw, G.E. (1995) The Arctic Haze Phenomenon, *Bull Amer. Met. Assoc.*, 76(12): 2403-2413.
530
- 531 Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, M. G. Duda, X.-Y. Huang,
532 W. Wang, and J. G. Powers (2008) *A Description of the Advanced Research WRF Version 3*,
533 NCAR Technical Note, NCAR/TN-475+STR, June 2008.