1 **Improving the computation efficiency of a source-oriented chemical mechanism for the**

2 **simultaneous source apportionment of ozone and secondary particulate pollutants**

3 Qixiang Xu[1, 2], Fangcheng Su[1, 2*], Ke Wang[1, 2], Ruiqing Zhang[1, 2], Qi Ying[3*], Michael J.

4 Kleeman[4]

5 1   School of Ecology and Environment, Zhengzhou University, Zhengzhou, 450001, China

6 2   Institute of Environmental Sciences, Zhengzhou University, Zhengzhou 450001, China

7 3   Zachry Department of Civil and Environmental Engineering, Texas A&M University,

8   College Station, TX 77845-3136

9 4   Department of Civil and Environmental Engineering, University of California, Davis, CA

10   95616

11 **Co-Correspondence:** Fangcheng Su (sufangcheng@zzu.edu.cn)

12 **Correspondence:** Qi Ying (qying@ust.hk) (Current Unit: Division of Environment and

13 Sustainability, Hong Kong University of Science and Technology, Hong Kong SAR, China)

14 **Abstract:**

15   Source-oriented chemical mechanisms enable direct source apportionment of air

16 pollutants by explicitly representing precursor emissions and their reaction products in

17 atmospheric models. These mechanisms use source-tagged species to track emissions and their

18 evolution. However, scalability was previously limited by the large number of reactions

19 required for interactions between two tagged species, such as the NOx-NOx or VOC-NOx

20 reactions. This study improves computational efficiency and scalability with a new method

21 that tracks the total concentration of tagged species, reducing the n^2 second-order reactions for

22 n sources into 2n pseudo first-order reactions. The overall production and removal rate of

23 individual species remain unchanged in the new approach. The number of reactions and

24 number of model species increase linearly with the number of source types, thus greatly

25 improved the computation efficiency. In addition, a source-oriented Euler Backward Iterative

26 (EBI) solver was implemented to replace the Gear solver used in previous applications of the

27 source-oriented mechanism. The source-oriented EBI solver has been assessed by comparing

28 predicted results with the Gear solver. Good agreement between those two methods has been

29 achieved, as the results from the EBI scheme are linearly correlated to Gear and average of

30 absolute relative error is below 5%. In the timing assessment, the proposed EBI scheme can

31 effectively reduce the total chemistry time by 73% to 90% for grids with different resolutions,

32 which leads to the reduction of total simulation time by 46% to 74%. The proposed source-

33 oriented scheme is efficient enough for practical long-term source apportionment applications
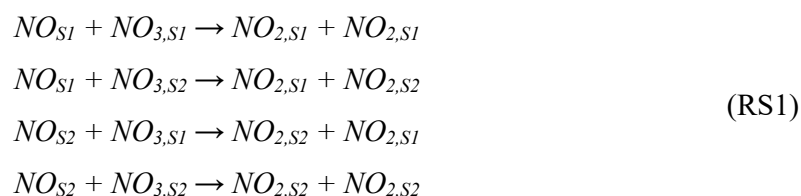
34 on nested domains.

35 **Keywords:** Atmospheric chemical solver, Euler Backward Iterative, Source apportionment

36

37 **1 Introduction**

38 **1.1 Source-oriented chemical mechanisms**

39 Source-oriented air quality models have been used extensively in source apportionment

40 modeling studies to determine source (or source region) contributions to NO_x (Zhang & Ying,

41 2011), VOCs (Ying & Krishnan, 2010), secondary inorganic (Ying & Kleeman, 2006) and

42 organic aerosols (Wang et al., 2018), and ozone (Wang et al., 2019a; Wang et al., 2020). In

43 these models, source-tagged species and their reactions are introduced in the gas phase

44 chemical mechanisms to track primary emissions and their reaction products from different

45 sources. For the source apportionment of secondary aerosol products from gas-to-particle

46 partitioning, aerosol and cloud processes are also modified to include additional model species

47 to represent the semi-volatile products from different sources.

48 While this is conceptually simple, the source-oriented mechanisms are computationally

49 expensive because the number of reactions increases almost quadratically with the number of

50 source types due to reactions that involve two source-tagged species. For example, consider

51 the simple reaction of $NO + NO_3 \rightarrow 2NO_2$, if the source-oriented mechanism is designed to

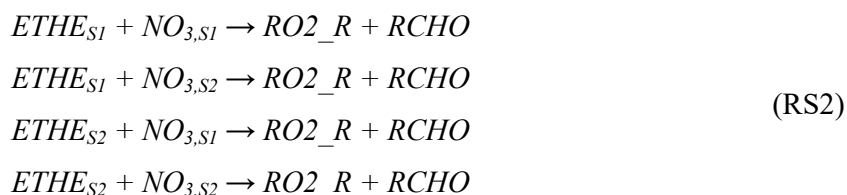52 track two explicit sources, four reactions in reaction set 1 (RS1) are needed:

$$NO_{S1} + NO_{3,S1} \rightarrow NO_{2,S1} + NO_{2,S1}$$
$$NO_{S1} + NO_{3,S2} \rightarrow NO_{2,S1} + NO_{2,S2}$$
$$NO_{S2} + NO_{3,S1} \rightarrow NO_{2,S2} + NO_{2,S1} \qquad \text{(RS1)}$$
$$NO_{S2} + NO_{3,S2} \rightarrow NO_{2,S2} + NO_{2,S2}$$

53 where the subscript $_{Sn}$ are tags attached to the name of the species to differentiate their

54 source-origin. For a total of N_s sources of NO_x to be tracked explicitly, N_s^2 reactions are needed

55 instead of one reaction. As there are quite a number of such $NO_x + NO_x$ reactions in the gas

56 phase inorganic chemistry, the number of reactions needed for the chemical mechanism grows

57 quickly. The necessity to deal with N_2O_5, which can be generated from NO_2 and NO_3 from

58   different sources, is handled with double-source-tagged species $N_2O_{5,Sij}$. The number of N_2O_5

59   species also increases quadratically with the number of explicit sources, leading to near-

60   quadratic growth of the overall number of species when the number of types to track gets

61   higher.

62   In ozone source apportionment calculations, it is also necessary to track the sources of

63   primary emitted VOCs as well as some of their reaction products in addition to the sources of

64   NO_x (Wang et al., 2020). Some of the unsaturated VOCs such as olefins can react with the

65   NO_3 radical. In the source-oriented mechanism, the number of reactions needed for these

66   VOC+NO_3 reactions also increases quadratically, as shown in RS2 below, using the ethene

67   (ETHE) + NO_3 reaction from the SAPRC-07 mechanism (Carter, 2010) as an example for two

68   sources. For accurate VOC source apportionment calculations that involve reactions between

69   two source-oriented organic radical species, such quadratic dependence of source types and

70   reaction numbers also arises   (Ying & Krishnan, 2010).

$$ETHE_{S1} + NO_{3,S1} \rightarrow RO2\_R + RCHO$$
$$ETHE_{S1} + NO_{3,S2} \rightarrow RO2\_R + RCHO$$
$$ETHE_{S2} + NO_{3,S1} \rightarrow RO2\_R + RCHO \quad \text{(RS2)}$$
$$ETHE_{S2} + NO_{3,S2} \rightarrow RO2\_R + RCHO$$

71   Due to the necessity of explicitly handling some or all of these reactions in source-oriented

72   mechanisms, the source-oriented modeling approach is computationally intensive so that

73   previous applications were limited to up to 9 explicit sources for secondary nitrate in a single

74   run (Kleeman & Cass, 2001; Ying et al., 2004; Ying et al., 2014). In some previous work for

75   VOC and secondary organic aerosol source apportionment, only one explicit source was

76   tracked at a time to simplify the reactions and to reduce the computation burden (Ying &

77   Krishnan, 2010; Wang et al., 2018). However, multiple model runs are needed to determine

78   the contributions from all sources. To make the source-oriented approach practical for a larger

79   number of source types, it is necessary to improve the computation efficiency of the source-

80   oriented approach.

81   **1.2 Numerical solution of stiff ODEs for gas phase reaction kinetics**

82   The gas phase chemical reaction kinetics are described by of a non-linear system of stiff

83   ordinary differential equations (ODEs) which must be solved to predict the transient evolution

84   of the concentrations of gas species. One of the most widely used schemes is the Gear method

85   (Ralph, 1973)

$$(\boldsymbol{I} - \hbar\beta_s \boldsymbol{J})\Delta \vec{C}_t^{m+1} = -\vec{C}_t^m + \sum_{j=1}^{s} \alpha_j \vec{C}_{t-jh} + \hbar\beta_s \frac{d\vec{C}_t^m}{dt} \tag{1}$$

86   where h is the integration time step; $\vec{C}_t^m$ is the vector of species concentrations to be solved

87   for time t during the mth iteration; $\Delta\vec{C}_t^{m+1}$ is the correction vector to estimate $\vec{C}_t^{m+1}$; $\boldsymbol{J}$ is the

88   Jacobian matrix of partial derivatives for all species that $J_{i,k} = \frac{\partial^2 C_{i,t}}{\partial C_{k,t} \partial t}$. It is calculated either

89   analytically or numerically initially based on $\vec{C}_{t-h}$ and updated when necessary using the most

90   recent values of $\vec{C}_t$; $\boldsymbol{I}$ is the identity matrix; s is the order of the accuracy; β_s and α_j are

91   scalar multipliers that depend on the order of the method. For each iteration, the new

92   concentrations for the next time step t is evaluated as $\vec{C}_t^{m+1} = \vec{C}_t^m + \Delta\vec{C}_t^{m+1}$. The iteration stops

93   when $\Delta\vec{C}_t^{m+1}$ becomes less than a provided error. A practical solver also needs to automatically

94   adjust the time step size h, the order of accuracy s and recalculate the Jacobin matrix when

95   necessary to ensure that the estimated error in one time step is less than a prescribed criteria

96   (Jacobson, 2006).

97   The advantage of using the Gear solver is that it is a general stiff solver so that no special

98   modifications are needed for a specific chemical mechanism. However, it is computationally

99   intensive as it involves evaluating the Jacobian matrix and performing LU factorization for the

100   left-hand side matrix. A Sparse-Matrix Vectorized Gear (SMVGEAR) solver was developed

101   by Jacobson and Turco (1994) and has been included in a number of atmospheric chemical

102   transport models (Zhang et al., 2011; Hu et al., 2012; Hu et al., 2014). The SMVGEAR solver

103   was also used previously to solve the gas phase reaction kinetics in the source oriented CTM

104   (Shi et al., 2017; Li et al., 2021). The gas chemistry is the most time-consuming step that

105   normally takes more than half of total simulation time, as shown in Figure 1.
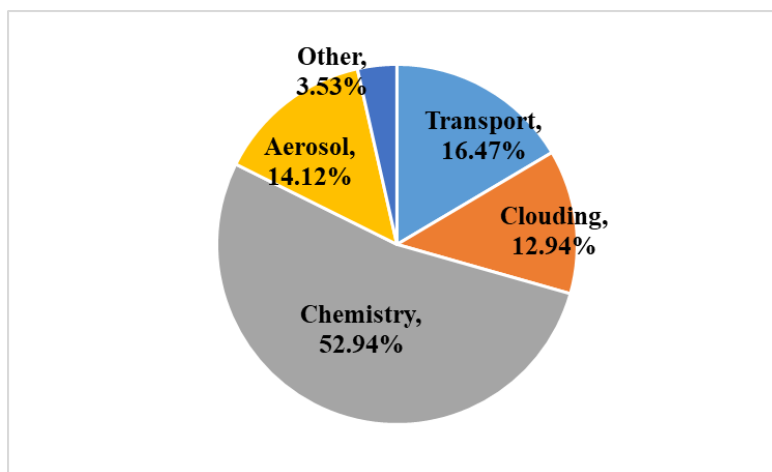
106
107 **Fig. 1.** Typical fraction of time spent in the scientific processes in the source-oriented CMAQ

108 model. This is based on a 36-km resolution domain (160 rows*130 columns*44 layers), 6

109 source types and solved using the SMVGEAR solver.

110 The Euler Backward Iterative (EBI) (Hertel et al., 1993) is a faster method to solve the stiff

111 ODE systems arising from a gas-phase photochemical mechanism. The basis of this method is

112 the backward Euler method as shown in equation (2),

$$C_{i,t} = \frac{C_{i,t-h} + hP_{i,t}}{1 + hL_{i,t}} \tag{2}$$

113 where $C_{i,t}$ and $C_{i,t-h}$ are the concentrations of species i at time t and t-h, respectively; h is

114 the integration time step; $P_{i,t}$ and $L_{i,t}$ are the chemical production and loss terms, respectively,

115 evaluated using the concentrations of the species at time t. Equation (2) represents a set of

116 coupled non-linear equations and a solution can be obtained by first evaluating the production

117 and loss terms using the concentrations from the previous time step *t-h* to calculate an initial

118 estimation of the species concentrations for the time step *t* based on equation (2). These

119 concentrations are applied to update the *P* and *L* terms so that and updated estimation of the

120 species concentrations for time step *t* are obtained. This procedure is repeated until the changes

121 in *C* are less than a prescribed value.

122 For atmospheric photochemical reactions, there are several families of species whose

123 concentrations are strongly coupled in reversible reactions. The general backward Euler

124 method described above has a slow rate of convergence or even fail to converge. In the EBI

125 solver, four family groups of species, are excluded from the general equation (2): (1) NO, NO_2,

126 O_3 and $O(^3P)$, (2) OH, HO_2, HONO and HNO_4, (3) peroxyacetyl radical (CH3C(O)OO·, or

127 C_2O_3) and peroxyacetyl nitrate (PAN), and (4) NO_3 and N_2O_5. For these species, explicit

128   solutions are applied to determine their concentrations at time t instead of using the P and L

129   terms with equation (2).

130   The accuracy of the EBI solver has been evaluated against more accurate solvers (Hertel

131   et al., 1993) and is the default choice in the CMAQ model for a number of build-in chemical

132   mechanisms. However, it cannot be directly used to solve the source-oriented chemical

133   mechanisms. In these mechanisms, groups of tagged reactive nitrogen species are included so

134   that the explicit solutions used for a non-typed regular chemical mechanism do not apply. An

135   EBI solver that can handle the chemical mechanisms with source tagged species and their

136   reactions will greatly improve the efficiency and applicability of the source-oriented air quality

137   model.

138   The objective of this study is to develop a computationally efficient source-oriented gas

139   phase chemical mechanism for the simultaneous source apportionment of O_3 and other gaseous

140   pollutants such as CO, primary VOCs, NO, NO_2, SO_2 and NH_3. The mechanism, when linked

141   with a proper source-oriented aerosol mechanism, can be used to determine the sources

142   contributions to nitrate, sulfate and ammonium ion. The following sections, the method to

143   improve the efficiency of the source-oriented mechanism by simplifying the representation of

144   the reactions and modifying the EBI ODE solver for source-oriented nitrogen species are

145   described in Section 2 and the testing of the improved mechanism and the source-oriented EBI

146   solver is described in Section 3.

2 Method

2.1 Reduce the number of reactions and source-tagged species

149   In the original source-oriented model, a general reaction set that involves two source-

150   tagged species as reactants for N_s source types can be written in the following form:

$$A_i + B_j \xrightarrow{k_{AB}} C_i + D_j + E, \quad i = 1,2,\ldots,N_s; j=1,2,\ldots,N_s \qquad (RS3)$$
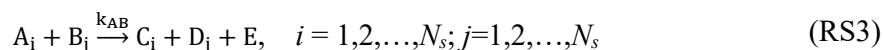
151   where A, B, C and D are source-tagged model species and the subscripts denote source

152   origin index of these species. For simplicity, assuming that C and D are the reaction products

153   from A and B, respectively. E represents a general product whose source-origin is not tracked

154   in the model simulation. k_{AB} is the second order reaction rate coefficient, which is the same for

155   all the reactions in this reaction set. Reaction sets RS1 and RS2, as shown in the examples in

156   Section 1, can both be expressed in this form. The loss rate of A_i is calculated using equations

157   (3a) and (3b):

$$\frac{d[A_i]}{dt} = -k_{AB}[B_1][A_i] \dots - k_{AB}[B_{N_s}][A_i] = -k_{A,effB}[A_i] \tag{3a}$$

$$k_{A,effB} = k_{AB}[B_{tot}] = k_{AB} \sum_{j=1}^{N_s}[B_i] \tag{3b}$$

158   where $k_{A,effB}$ is the pseudo-first order reaction rate coefficient for $[A_i]$ based on the total

159   concentration of B, $[B_{tot}]$, as defined in equation (3b). A similar set of equations can be derived

160   for the loss rate of individual tagged species B_i. Thus, the N_s^2 second-order reactions

161   represented by RS3 can be equivalently described by the following $2N_S$ pseudo first-order

162   reactions,

$$A_i \xrightarrow{k_{A,effB}} C_i + E, \quad i=1,2,\dots,N_s$$
$$B_i \xrightarrow{k_{B,effA}} D_i, \quad\quad i=1,2,\dots,N_s \tag{RS4}$$
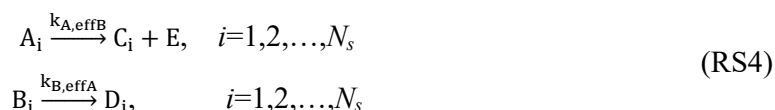
163   For the non-typed product E, it can appear in either the A_i reactions or the B_i reactions and

164   it is easy to show that the formation rate of E based on RS4 is the same as these from RS3.

165

166   The double-tagged N_2O_5 species and their reactions can be simplified as well. For $N_2O_{5,ij}$

167   which represents N_2O_5 based on NO_2 from source i and NO_3 from source j, it can be

168   equivalently written as $0.5N_2O_{5,i} + 0.5N_2O_{5,j}$ as shown in reaction set RS5, in terms of

169   preserving the source contributions to NO_2 and NO_3,

$$NO_{2,i} + NO_{3,j} \xrightarrow{k_{NO2NO3}} 0.5N_2O_{5,i} + 0.5N_2O_{5,j} \ , \ i=1,2,\dots,N_s; j=1,2,\dots,N_s \tag{RS5}$$
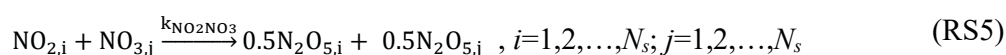
170   With this simplification, as well as the pseudo-first order reaction technique described

171   above, the reactions of N_2O_5 formation from N_s types of NO_x can be expanded into $2N_s$

172   reactions with N_s tagged N_2O_5 species as shown in the general reaction set RS4, where A is

173   NO_2, B is NO_3 and C and D are N_2O_5 species that have the same source tag as A and B,

174   respectively. This represents a significant reduction in terms of the number of species as well

175   as the number of reactions for the source-oriented mechanism.

176   The total concentration of tagged species needed in equation (3b) for the pseudo first order

177   reaction rate coefficient need not be tracked separately. Instead, they are calculated on-the-fly

178   and then used to calculate the pseudo first-order rate coefficients for the reactions of the tagged

179   species shown above. The function that calculates the reaction rates to be used in the stiff ODE

180   solvers needs to be modified to recognize these special pseudo first-order reactions. The

181   CMAQ model is capable of dealing with these special pseudo-first order reaction natively with

182   its included mechanism preprocessor (CHEMMECH). An example input to the CHEMMECH

183   on how the reactions are constructed for $NO + NO_3 \rightarrow 2NO_2$ for 10 source types is illustrated

184   in List 1.

2.2 Source-oriented Euler Backward Iteration (EBI) scheme

2.2.1 Solution for family of source-tagged species in coupled reversible reactions

187   For source apportionment of ozone and secondary inorganic aerosols, the reactive nitrogen

188   species NO, NO_2, HONO, HNO_4, PAN, NO_3 and N_2O_5 are source-tagged. For PAN, its source

189   is determined by the source of NO_2 while the peroxyacetyl group is not source-tagged.

190   The standard solution procedure of the source-tagged species in the source-oriented EBI

191   method includes two major steps: (1) evaluating of total concentrations of these tagged species

192   using equation sets (9), (10) and (11) and (12) in Hertel et al. (1993); and (2) predicting the

193   concentrations for each tagged species based on the total concentrations; In the following,

194   equations for step (1) are summarized first, followed by equations for step (2). Equations are

195   separately listed for each family.

196   The first set of equations (4a-4d) are used to solve the total concentrations of NO, NO_2, O_3

197   and $O(^3P)$. These equations are based on the corresponding ones in Hertel et al. (1993).

$$[NO_{tot}]_t = [NO_{tot}]_{t-h} + hP_1' + h(r_{1,2} + J_1)[NO_{2,tot}]_t - hk_{1,3}[O_3]_t[NO_{tot}]_t$$
$$- hr_{2,1}[NO_{tot}]_t - hL_1'[NO_{tot}]_t \tag{4a}$$

$$[NO_{2,tot}]_t = [NO_{2,tot}]_{t-h} + hP_2' + hr_{2,1}[NO_{tot}]_t + hk_{1,3}[O_3]_t[NO_{tot}]_t$$
$$- h(r_{1,2} + J_1)[NO_{2,tot}]_t - hL_2'[NO_{2,tot}]_t \tag{4b}$$

$$[O_3]_t = [O_3]_{t-h} + hJ_2[O(^3P)]_t - hk_{1,3}[NO_{tot}]_t[O_3]_t - hL_3'[O_3]_t \tag{4c}$$

$$[O(^3P)]_t = [O(^3P)]_{t-h} + hP_{12}' + hJ_1[NO_{2,tot}]_t - hJ_2[O(^3P)]_t - hL_{12}'[O(^3P)]_t \tag{4d}$$

198   In the above equations, species with a subscript *tot* represent the total concentration of a

199   set of tagged species, which is calculated by adding the concentrations of the individual tagged

200   species; h is the size of the current time step; $r_{1,2}$ is the production rate coefficient of NO from

201   NO_2, excluding photo-dissociation and J_1 is the photolysis rate of NO_2 to form NO; $k_{1,3}$ is the

202   second-order rate coefficient for $NO+O_3$ reaction to form NO_2; $r_{2,1}$ is the pseudo-fist order

203   rate constant for the production of NO_2 from NO from all other pathways, excluding the

204    $NO+O_3$ reaction. J_2 is the first-order reaction rate constant for $O(^3P)+O_2$ to form O_3. The terms

205   P_1', P_2', and P_{12}' account for all the remaining production for the total concentrations of NO,

206   NO_2 and $O(^3P)$ in the mechanism, and the terms L_1', L_2', , L_3', and L_{12}' are the losses of the

207   total concentrations of NO, NO$_2$, O$_3$ and O(^{3}P), respectively. Analytical solutions for the total

208   concentrations based on 4a-4d were derived and described in detail in the Appendix A of Hertel

209   et al. (1993) and are not repeated here.

210   Once the total concentrations of NO, NO$_2$, O$_3$ and O(^{3}P) are solved, concentrations of the

211   source-tagged NO and NO$_2$ are solved from the following two equations,

$$[NO_i]_t = [NO_i]_{t-h} + hP'_{1,i} + h(r_{1,2} + J_1)[NO_{2,i}]_t - hk_{1,3}[O_3]_t[NO_i]_t - hr_{2,1}[NO_i]_t$$
$$- hL'_1[NO_i]_t \tag{5a}$$

$$[NO_{2,i}]_t = [NO_{2,i}]_{t-h} + hP'_{2,i} + hr_{2,1}[NO_i]_t + hk_{1,3}[O_3]_t[NO_i]_t - h(r_{1,2} + J_1)[NO_{2,i}]_t$$
$$- hL'_2[NO_{2,i}]_t \tag{5b}$$

212   where i is the source index for the tagged species. For each i, the two unknows $[NO_i]_t$ and

213    $[NO_{2,i}]_t$ are solved analytically using the following equations,

$$[NO_i]_t = \frac{1}{\det(A)}(A_{22}b_1 - A_{12}b_2) \tag{6a}$$

$$[NO_{2,i}]_t = \frac{1}{\det(A)}(-A_{21}b_1 + A_{11}b_2) \tag{6b}$$

$$A = \begin{bmatrix} 1 + hk_{1,3}[O_3]_t + hr_{1,2} + hL'_1 & -h(r_{2,1} + J_1) \\ -h(r_{2,1} + k_{1,3}[O_3]_t) & 1 + hr_{1,2} + hJ_1 + hL'_2 \end{bmatrix} \tag{6c}$$

$$b = \begin{bmatrix} [NO_i]_{t-h} + hP'_{1,i} \\ [NO_{2,i}]_{t-h} + hP'_{2,i} \end{bmatrix} \tag{6d}$$

214   In equations 6a and 6b, det(**A**) is the determinant of the 2 x 2 matrix **A**, as defined in

215   equation 6c.

216   The second set of equations are for the total concentrations of OH, HO$_2$, HONO and HNO$_4$,

217   as shown in equations (7a) - (7d),

$$[OH]_t = [OH]_{t-h} + hP'_4 + hr_{4,5}[HO_2]_t + hr_{4,19}[HONO_{tot}]_t - hL_4[OH]_t \tag{7a}$$

$$[HO_2]_t = [HO_2]_{t-h} + hP'_5 + hr_{5,4}[OH]_t + hr_{5,21}[HNO_{4,tot}]_t - 2hk_{5,5}[HO_2]_t^2$$
$$- hL'_5[HO_2]_t \tag{7b}$$

$$[HONO_{tot}]_t = [HONO_{tot}]_{t-h} + hr_{19,4}[OH]_t - hL_{19}[HONO_{tot}]_t \tag{7c}$$

$$[HNO_{4,tot}]_t = [HNO_{4,tot}]_{t-h} + hr_{21,5}[HO_2]_t - hL_{21}[HNO_{4,tot}]_t \tag{7d}$$

218   The notations and symbols used in the above equations are similar to those used in

219   equations (5a-5d). $r_{4,5}$ and $r_{4,19}$ are pseudo first order production rate coefficients of OH

220   from HO$_2$ and HONO, respectively. $r_{5,4}$ and $r_{5,21}$ are pseudo first order production rate

221   coefficients of HO$_2$ from OH and HNO$_4$, respectively. $k_{5,5}$ is the HO$_2$+HO$_2$ self-reaction rate

222 coefficient. $r_{19,4}$ and $r_{21,5}$ are pseudo first order rate coefficients for the production of HONO

223 and HNO_4 from OH+NO and HO_2+NO_2, respectively. The terms P_4' and P_5' account for all

224 the remaining production of OH and HO_2 and the terms L_4', L_5', L_{19} and L_{21} account for all

225 the other losses of OH, HO_2, HONO and HNO_4, respectively. Analytical solutions for (7a)-(7d)

226 were also derived and described in detail in Appendix A of Hertel et al. (1993). As the OH and

227 HO_2 concentrations are determined, concentrations of individual HONO and HNO_4 from

228 different sources are solved using equations 8a and 8b, respectively.

$$[HONO_i]_t = \frac{[HONO_i]_{t-h} + hr_{19,4}^i[OH]_t}{1 + hL_{19}} \tag{8a}$$

$$\left[HNO_{4,i}\right]_t = \frac{\left[HNO_{4,i}\right]_{t-h} + hr_{21,5}^i[HO_2]_t}{1 + hL_{21}} \tag{8b}$$

229 where $r_{19,4}^i$ and $r_{21,5}^i$ are pseudo first order production rate coefficients of HONO and

230 HNO_4 from source i due to NO and NO_2 from the same source with OH and HO_2, respectively.

231 Concentrations of NO_i and $NO_{2,i}$ for the current timestep has already been determined using

232 (6a) and (6b) and these concentrations will be applied to calculate $r_{19,4}^i$ and $r_{21,5}^i$ used in the

233 above two equations.

234 The third set of equations are for C_2O_3 and PAN:

$$[C_2O_3]_t = [C_2O_3]_{t-h} + hP_8' + hr_{8,9}[PAN_{tot}]_t - 2hk_{8,8}[C_2O_3]_t^2 - hL_8'[C_2O_3]_t \tag{9a}$$

$$[PAN_{tot}]_t = [PAN_{tot}]_{t-h} + hr_{9,8}[C_2O_3]_t - hL_9[PAN_{tot}]_t \tag{9b}$$

235 In the equation, it is assumed that in the source-oriented mechanism, PAN is a source-

236 tagged species and its source is based on the source of NO_2 but C_2O_3 is not source-tagged. This

237 is sufficient for the source apportionment of ozone, as described in the Section 2.1. A quadradic

238 equation for C_2O_3 can be obtained (see Appendix A of Hertel et al. 1993). In $r_{9,8}$, total NO_2

239 concentration at the current timestep t has already been determined by solving equations 5a-

240 5d. Once the concentrations of C_2O_3 at timestep t is solved, the concentrations of each of the

241 tagged PAN species can be used by:

$$[PAN_i]_t = \frac{[PAN_i]_{t-h} + hr_{9,8}^i[C_2O_3]_t}{1 + hL_{21}} \tag{9c}$$

242 where the $r_{9,8}^i$ includes the concentration of $NO_{2,i}$ (NO_2 attributed to source i) at the

243 current timestep t.

244 The last set of equations treated specially in the source-oriented EBI solver is for NO_3 and

245   N_2O_5, as shown in equations (10),

$$[NO_{3,tot}]_t = [NO_{3,tot}]_{t-h} + hP_{15}' + hr_{15,16}[N_2O_{5,tot}]_t - hL_{15}[NO_{3,tot}]_t \tag{10a}$$

$$[N_2O_{5,tot}]_t = [N_2O_{5,tot}]_{t-h} + hr_{16,15}[NO_{3,tot}]_t - hL_{16}[N_2O_{5,tot}]_t \tag{10b}$$

246   The two equations are linear equations and can be solved easily for $NO_{3,tot}$ and $N_2O_{5,tot}$.

247   The $NO_{3,i}$ and $N_2O_{5,i}$ (as discussed in section 2.1, sources of N_2O_5 can be tracked with a single

248   type instead of double typed) can be solved explicitly based on the two equations as well, as

249   shown in (10c) and (10d),

$$[NO_{3,i}]_t = \frac{(1 + hL_{16})\left([NO_{3,i}]_{t-h} + hP_{15}^{i,'}\right) + hr_{15,16}[N_2O_{5,tot}]_{t-h}}{(1 + hL_{15})(1 + hL_{16}) - h^2 r_{15,16} r_{16,15}} \tag{10c}$$

$$[N_2O_{5,i}]_t = \frac{(1 + hL_{15})[N_2O_{5,tot}]_{t-h} + hr_{16,15}\left([NO_{3,i}]_{t-h} + hP_{15}^{i,'}\right)}{(1 + hL_{15})(1 + hL_{16}) - h^2 r_{15,16} r_{16,15}} \tag{10d}$$

250   Note that $r_{16,15}$ includes the total concentration of NO_2 thus is the same as that used in

251   equation (10b). However, the production of NO_3 from other reactions do have to be source

252   specific, thus the P15's used in equation (10c) and (10a) are different.

253   To avoid round-off errors introduced in the calculation for the source-tagged species

254   involved in these four groups so that the sum of the source-tagged species always exactly

255   matches the total concentrations, their concentrations are readjusted by the total concentrations

256   solved independently.

2.2.2 Successive under-relaxation

257
258   During the testing of the above algorithm, non-converging oscillations were sometimes

259   observed, mostly due to the low concentration source-tagged species. The iterative process

260   used in the EBI solver can be modified to include a relaxation factor α so that the final

261   concentrations at the current timestep t is modified by a weighted average of the results from

262   the previous time step *t-h* and the present estimated values for the current timestep *t*, as shown

263   in equation (11),

$$C_{i,t}^{final} = (1 - \alpha)C_{i,t-h} + \alpha C_{i,t} \tag{11}$$

264   Generally, larger α values lead to faster convergence, but have a higher chance of fall into

265   oscillation. Based on the testing, α=0.8 appears to be a conservative choice that always lead to

266   convergence. A dynamic under-relaxation scheme using a set of varying α values between 0.79

267   and 1.0, based on the number of iterations in the EBI scheme, is shown to lead to faster

268   convergence. This is further discussed in the Results section.

2.3 Test mechanism and model set up

269   To evaluate how much improvement in computation efficiency can be achieved by using
270 the simplified reaction representation and source-oriented EBI solver, a series of source-
271 oriented mechanisms for simultaneous attribution of ozone and secondary inorganic aerosol
272 were constructed based on the SAPRC-07 photochemical mechanism and implemented in
273 CMAQv5.2. The SAPRC-07 mechanism is chosen instead of the more recent versions of
274 SAPRC because it is faster with fewer species and reactions, and thus is more suitable for
275 simulations requiring rapid response, such as operational air quality forecasting and for source
276 apportionment of ozone and secondary inorganic aerosols. The source-oriented mechanism
277 based on this will be applied in a future air quality forecasting model that also forecasts source-
278 tagged species concentrations and source-region contributions to air pollution. Since the main
279 purpose of this paper is to test the efficiency of the gas phase algorithm, although aerosol and
280 cloud processes were enabled in the simulations, aerosol results are not compared in the
281 analyses described below.

282   The tested SAPRC-07 mechanism used in this study contains a total of 134 species and
283 341 reactions. Among these species, 15 species are reactive nitrogen species. For each of these
284 species, tagged species are used to track their source origins. Reactions involving these species
285 are expanded in the source-oriented mechanism. In addition, CO, SO_2 and sulfuric acid (SULF)
286 were also expanded in the source-oriented mechanism. To evaluate source contributions to
287 ozone, 14 primary VOC species were also treated as source-oriented species in addition to
288 source-tagged non-reactive O_3 species to track contributions from different sources of NOx
289 and VOCs to O_3 formation. As HCHO is an important oxidation product from several parent
290 VOCs, sources of secondary HCHO from the first-generation oxidation of parent VOCs are
291 also tracked. Details for the source apportionment of O_3 has been described by Wang et al.
292 (Wang et al., 2019a; Wang et al., 2019b) and are not repeated here. Two versions of the source-
293 oriented SAPRC-07 mechanisms are prepared. The first version (V1) uses double-tagged N_2O_5
294 and fully expanded reactions without the pseudo first order reactions described in Section 2.
295 The second version (V2) is based on single-tagged N_2O_5 and a simplified treatment of reactions
296 involving two source-tagged reactants, as descried in section 2.1. Both mechanisms are
297 constructed to track emissions from ten source types. The number of reactions and species in
298 each mechanism is listed in Table 1.

Table 1. Computation time needed for a one-day simulation using two different versions of the source-oriented chemical mechanism and two versions of the ODE solvers. Both versions are capable of tracking 10 different source types in a single simulation.

	V1 SMVGEAR	V2 SMVGEAR			V2 EBI*		
Domain resolution	36km	36km	12km	4km	36km	12km	4km
# of species							
# of reactions							
Total Chem. Time^ (min)	133	49.9	172	325	13.4	15.1	32.4
Total Sim. Time^ (hr)	2.63	1.32	3.59	6.27	0.707	1.03	1.63
Chem %	84%	63%	80%	86%	32%	24%	33%
Chem time reduction wrt. V1	-	62%	-	-	90%	-	-
Chem time reduction wrt. V2-S	-	-	-	-	73%	91%	90%
Total time reduction wrt. V1	-	50%	-	-	73%	-	-
Total time reduction wrt. V2-S	-	-	-	-	46%	71%	74%

* Simulation time based on dynamic under-relaxation coefficient.

^ Wall-clock time

The testing is performed for a one-day (July 1, 2020) simulation using three-level nested domains with 36, 12 and 4 km resolutions that cover eastern Asia, central and eastern China and Henan province in central China, respectively. The meteorology inputs were based on the Weather Research and Forecasting (WRF) model v4.1.4. The anthropogenic emissions are based on the Multi-scale Emission Inventory for China (MEIC, for 36- and 12-km domains, available from http://www.meicmodel.org/) and a local emission inventory (for the 4-km domain) (Lu et al., 2023). Emissions are grouped into six source categories, including five anthropogenic source sectors (power, industrial, residential, transportation and agriculture) and one biogenic sector, whose emission is based on MEGAN (Model for Emissions of Gases and Aerosol from Nature) v2.1(Guenther et al., 2006). The initial concentrations of the species are based on a 7-day non-source-oriented simulation. The boundary conditions for the 36-km domain is based on the clean continental vertical profiles included in the CMAQ model. The boundary conditions of the 12- and 4-km domains are based on results from the parent 36- and 12-km domains, respectively.

Three sets of simulations using the source-oriented mechanisms were conducted: (1) V1, solved using the SMVGEAR solver (V1-GEAR), (2) V2, solved with SMVGEAR (V2-GEAR), and (3) V2, solved with source-oriented EBI (V2-EBI). For the SMVGEAR a relative tolerance (RTOL) of 1×10^{-3} and an absolute tolerance (ATOL) of 1×10^{-9} were used. For the EBI solver, only a relative tolerance RTOL was used in the convergence check. For most of the species, a relative tolerance of 1×10^{-3} is used. Integrated reaction rate analysis (IRR) was

326 used in all three simulations as it is needed for the ozone source apportionment algorithm

327 (Wang et al., 2019b). In addition to the source-oriented model simulations, two base case

328 simulations (V0) were conducted using the unmodified SAPRC-07 solved with the

329 SMVGEAR (V0-SMVGEAR) and the EBI (V0-EBI) solvers, respectively.

330   All the simulations were conducted on a Dell Precision-Tower 7810 working station (2XE-

331 2660-v4, 28/56 cores/threads and 256G of DDR4 RAM), and the run was in parallel in a

332 configuration of 8x6 domain decomposition.

3 Results

3.1 Timing results

335   The wall-clock time for the gas-phase mechanism as well as the total run time were

336 recorded using the time function (MPI_WTIME) in the Massage Passing Interface. An

337 MPI_BARRIER call was issued before each MPI_TIME call to make sure that the measured

338 wall-clock time represents the actual time for a chemistry time step in a static domain

339 decomposition setting with imbalanced loads.

340   The wall-clock time for the 1-day simulation using the source-oriented EBI solver was

341 compared with the SMVGEAR solver and presented in Table 1. The fully expanded source-

342 oriented mechanism with SMVGEAR (V1-SMVGEAR) is the slowest and only the 36-km

343 resolution simulation was conducted. The simplified reaction representation alone (V2-

344 SMVGEAR, see section 2.1) leads to a reduction of chemistry time by ~62% when compared

345 with V1-SMVGEAR (133 min to 49.9 min) and total computation time by ~50% (2.63 hr to

346 1.32 hr). Using the source-oriented EBI on V2 (V2-EBI) further reduces the total chemistry

347 time to ~13.4 min. Compared to the V1-SMVGEAR, V2-EBI reduced the chemistry time by

348 ~90% and the total time by ~75% for a one-day simulation in the 36 km domain.

349   The EBI solver represents a significant reduction in both chemistry time and total

350 computation time comparing to the SMVGEAR solver. When V2-SMVGEAR and V2-EBI

351 are compared, the total chemistry time saving of EBI scheme increases with grid resolution,

352 from 70% for 36 km to 90% for 4 km grid. As the result, the total simulation time saving

353 increases from 46% for coarse grid to 74% for fine grid. The increase in time saving of EBI

354 solver with grid resolution can be attributed to the smaller time step size which is determined

355 from flow Courant stability criterion. For fixed simulation duration (1 day in this study), the

356 required total number of chemistry steps increase dramatically with grid resolution, this results

357   in significant reduction in total chemistry time with faster EBI scheme for finer grid. Therefore,

358   for time consuming applications such as long-term source apportionment simulation of nested

359   domains, the time efficiency can be improved for more than three folds.

360   The V2-EBI results shown in Table 1 are based on the dynamic under-relaxation using an

361   iteration count dependent under-relaxation factor (α) as shown in Table 2. In this scheme, the

362   α is initially set to 1.0 and gradually decreases to smaller values. If the solution does not

363   converge in 15 iterations, a constant α of 0.79 is used. Using this dynamic α scheme is

364   demonstrated to be more efficient than using a constant α, as shown in Figure 2. The optimal

365   value for fixed α is 0.8, at which the total chemistry time could be reduced by ~62%, for the

366   36-km domain simulations, which is approximately 10% less than the dynamic α scheme.
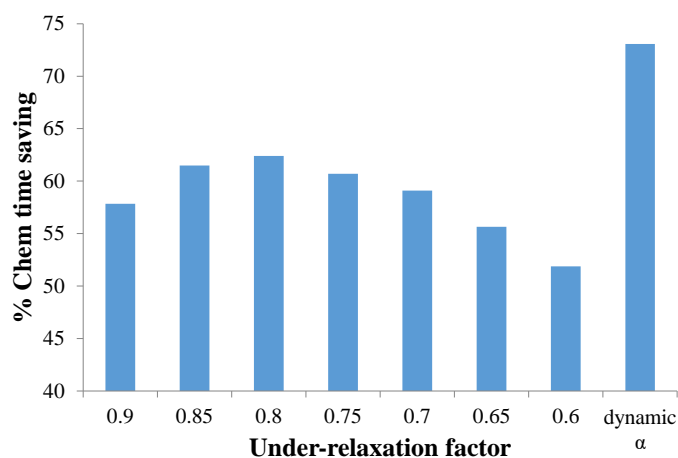

367

368   **Fig. 2**. Percentage chemistry time saving with respect to the SMVGEAR solver for the

369   identical source-oriented chemical mechanism for the 36-km domain. See text and Table 2 for

370   the details of the dynamic under-relaxation scheme.

371   **Table 2.** Dynamic section of the under-relaxation factor (α) based on iteration count

Iteration #	α	Iteration #	α
1	1.00	6-7	0.79
2	0.90	8-10	0.78
3	0.85	11-15	0.80
4-5	0.81	≥16	0.79

372

373   **3.2 Accuracy assessment of the source oriented EBI solver**

374   The general accuracy of the EBI solver has been tested by Hertel et al. (1993). The

375   accuracy of proposed source apportionment EBI scheme was evaluated by comparing

376   predicted results with those from the SMVGEAR and with the non-source-oriented simulation

377   V0-EBI.

378   Hourly average concentrations of NO, NO_2, PAN, HONO, OH and HO_2 at each grid cell

379   in the surface layer of the 36-km domain were selected as sample species for accuracy

380   evaluation because these are the important species treated specially by the source oriented EBI

381   solver. The maximum and mean values of the normalized error of these species are listed in

382   the Table 3. For all these species, the maximum normalized error among all grid cells is less

383   than 15% and the mean normalized error does not exceed 4%.

384   **Table 3.** Max and mean values of the normalized error* for selected species in the 36-km

385   domain.

Species	Max Normalized error (%)	Mean Normalized Error (%)	Species	Max Normalized error (%)	Mean Normalized Error (%)
NO_X0	14.45	3.01	NO₂_X0	12.69	3.99
NO_X1	5.92	2.50	NO₂_X1	11.89	3.37
NO_X2	2.30	0.0059	NO₂_X2	9.59	3.14
NO_X3	6.29	1.35	NO₂_X3	8.26	2.19
NO_X4	5.64	0.59	NO₂_X4	11.34	1.17
NO_X5	0.99	0.57	NO₂_X5	5.47	1.46
NO_X6	1.35	0.41	NO₂_X6	4.7	0.83
PAN_X0	9.52	2.98	HONO_X0	4.55	1.52
PAN_X1	6.42	1.09	HONO_X1	5.49	1.13
PAN_X2	8.12	1.23	HONO_X2	4.36	0.65
PAN_X3	6.22	0.96	HONO_X3	13.89	1.56
PAN_X4	5.99	0.75	HONO_X4	8.56	1.61
PAN_X5	2.08	0.014	HONO_X5	7.41	1.32
PAN_X6	4.97	0.672	HONO_X6	5.73	0.41
OH	3.01	0.19	HO₂	5.15	0.35

386   * Normalized error is calculated as $|C_{V2-EBI}-C_{V2-SMVGEAR}|/C_{V2\_SMVGEAR}$. This is calculated for hourly

387   concentrations for all the grid cells in the entire day. The mean normalized error is calculated by averaging the

388   normalized error for all the grid cells.

389   For OH and HO_2, the EBI solver agrees with the SMVGEAR solver very well, with mean

390   differences of ~0.19% and 0.35%, respectively. Figure 3 shows the comparison of all the

391   hourly concentrations of OH and HO_2 at hour 6, which represent average concentrations

392   between 1400-1500 local time. The EBI results agree with SMVGEAR solver results across

393   all concentration ranges that span more than three orders of magnitude. Figure 4 shows the

394   comparison of hourly concentrations of NO_2, NO, NO_3, HONO and PAN for hour 24. For the

395 total concentrations these species (i.e. sum of the concentrations of the source-tagged species),

396 the source-oriented EBI predictions agree very well with the observations. For the individual

397 source-tagged species, differences between the EBI and SMVGEAR results are highest among

398 the NO_2, NO and HONO species that are used to track the initial and boundary contributions

399 (IC/BC type, first column of Figure 3). The results shown in the plots are essentially initial

400 concentrations because boundary conditions for NO and NO_2 are quite low, and won't

401 contribute to such high concentrations (see Figure 5 – the regional NO2 and NO plots). The

402 source-oriented EBI predictions for IC/BC species are biased high than those predicted using

403 the Gear solver. These over-predictions are balanced by the general under-predictions of EBI

404 for species associated with emissions.
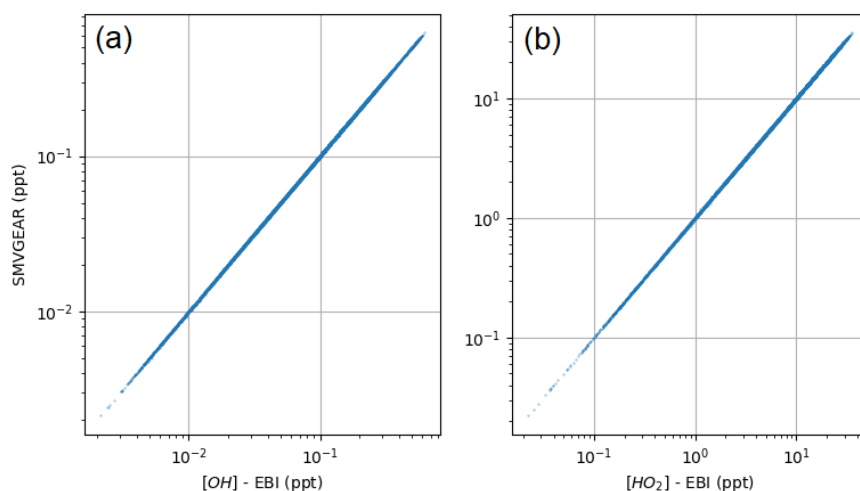
406 **Fig. 3.** Comparison of predicted OH (a) and HO_2 (b) radicals in the surface layer of the 36-

407 km resolution domain using the source-oriented EBI and the SMVGEAR solvers.

408 Concentrations at hour 6 (local time 1400-1500) are shown.
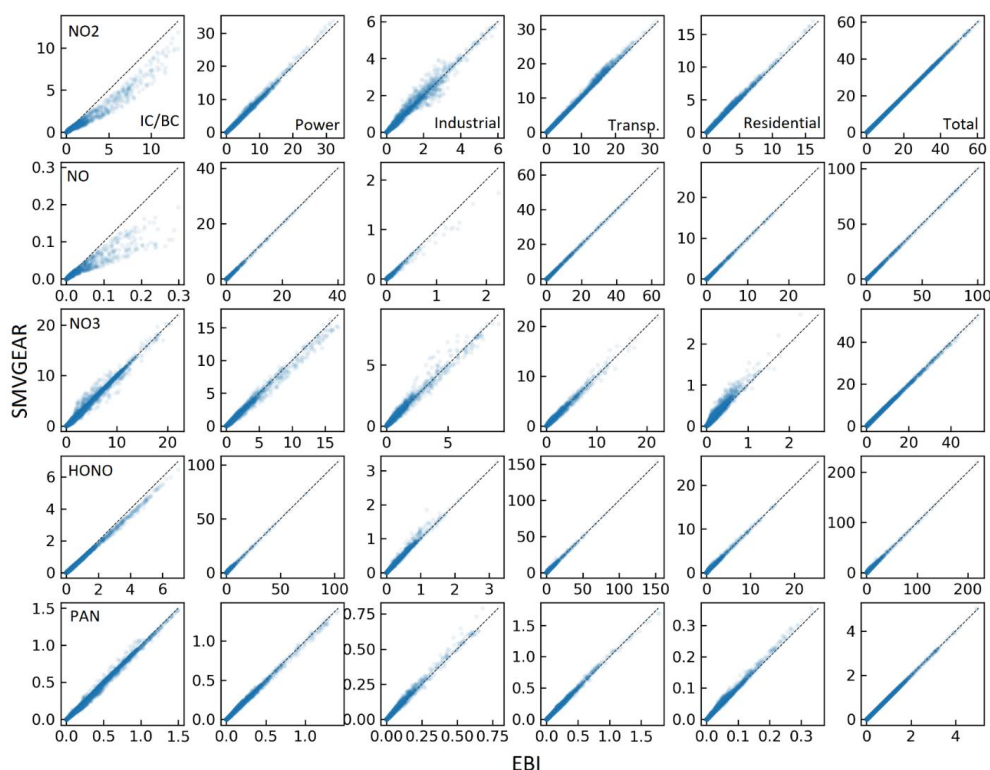
Fig. 4. Predicted hourly-averaged NO_2, NO, NO_3, HONO and PAN concentrations for different source types in the last hour of the 1-day simulation using source-oriented EBI and the SMVGEAR solver. Concentrations of all grid cells in the surface layer are included in the plot. Concentrations are in units of ppb for NO, NO_2 and PAN and in units of ppt for NO_3 and HONO.

Figure 5 shows the predicted hourly time series of OH, HONO, NO_2 and PAN at five grid cells that represent no emission, and low, medium, high and intense emission conditions, respectively. For HONO, NO_2 and PAN, predicted concentrations are shown for IC/BC type and the sum of the tagged species for other emission-related types. This again demonstrates that predicted OH from EBI and SMVGEAR agree with each other at all times under all emission conditions. The fraction of initial concentrations to the total concentration decreases as the emission intensity increases and the difference between EBI and SMVEAR remains very small.
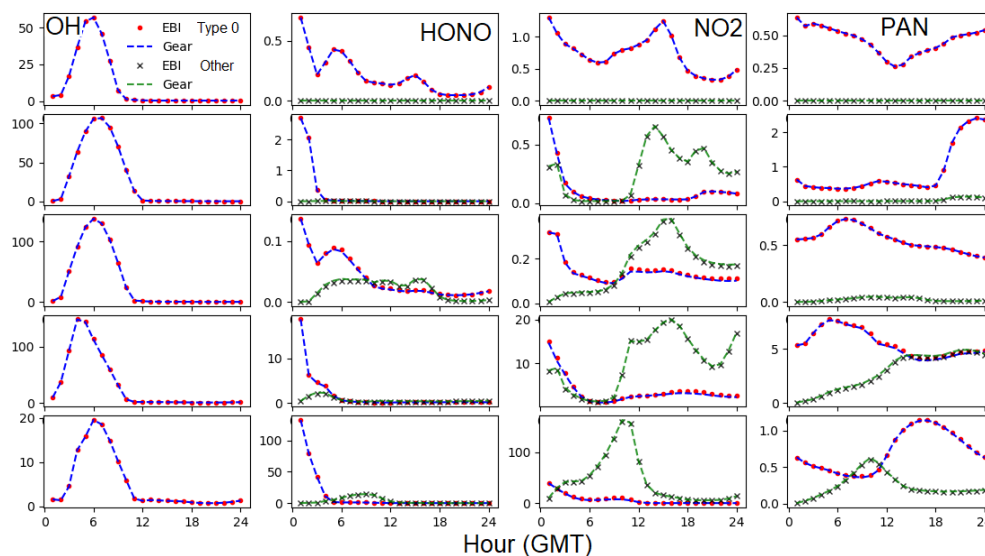
Fig. 5. Timeseries of OH, HONO, NO$_2$ and PAN at grid cells that represent different emission conditions (no emissions, low emission, medium emission, high emission and intense emission) for the one-day simulation. Units are ppq (parts per quadrillion) for OH, ppt for HONO and 0.1ppb for NO$_2$ and PAN. Type 0 is the concentration for the IC/BC source type and "other" represents the sum of the concentrations of all other tagged species.

Figures 6 and 7 illustrate the spatial distribution of daily average concentrations predicted from two methods. HO, HO$_2$ and two tagged NO, NO$_2$ and HONO of 36 km grid were selected as sample species, results from proposed EBI scheme are very close to the SMVGEAR results. The concentration of X0 species predicted by EBI scheme at high concentration zone are higher than the Gear, the reason of such slight deviation is mainly caused by the time lag in results of Euler backward scheme that has been explained in the previous section.
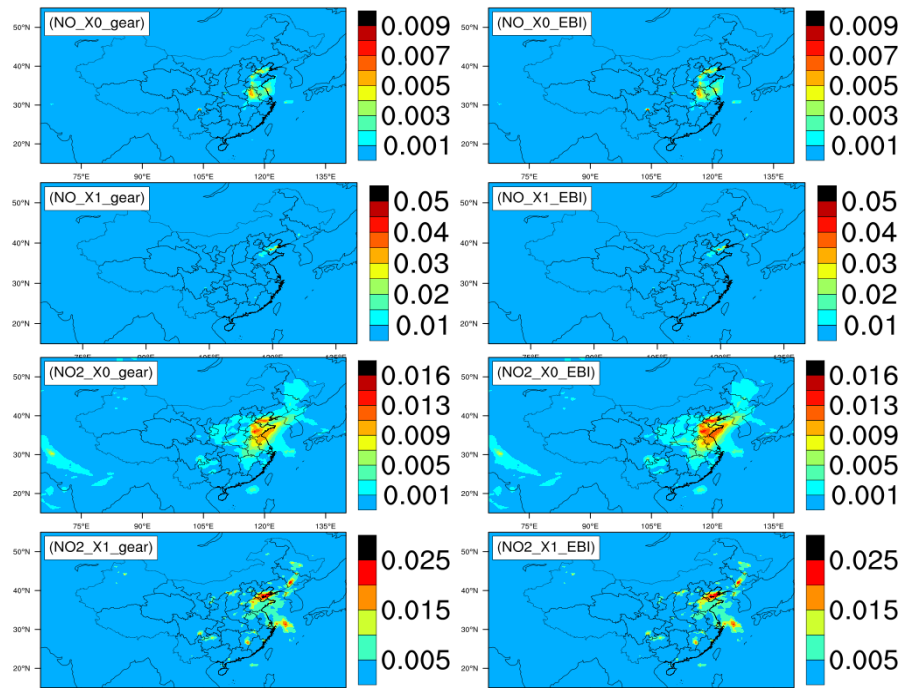
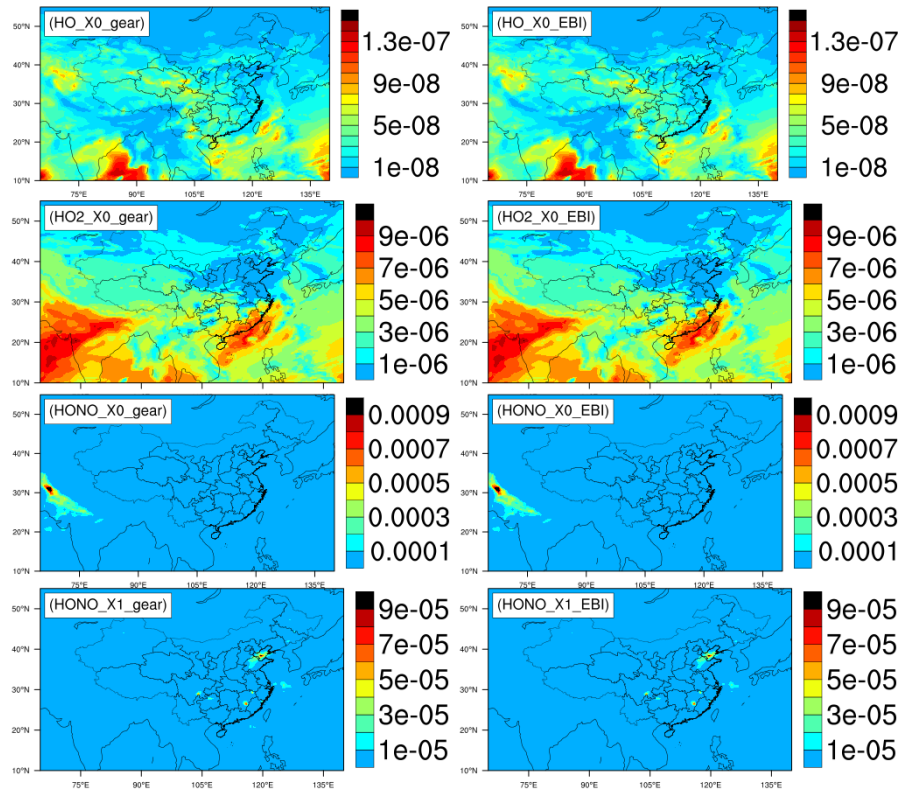**Fig. 6.** Regional distribution of NO and NO_2, units are ppm.

Fig. 7. Regional distribution of HO HO_2 and HONO, units are ppm.

4 Conclusions

In this study, the computation efficiency and thus scalability of the source-oriented approach is greatly improved with a new approach of dealing with these two-tagged-species reactions. The new approach is based on tracking the total concentration of the source-tagged species and reduce the n^2 number of second-order reactions for n sources into 2n pseudo first-order reactions. The production and removal rate of individual species remain unchanged in the new approach. The number of reactions and number of model species increase linearly with the number of source types, thus greatly improved the computation efficiency. Test cases based on the Texas Air Quality Study 2006 ozone episode showed that a source-oriented SAPRC-07 mechanism that simultaneously performs the source apportionment of NOx, SO_2, primary VOCs, HCHO, and ozone for 16 sources needs only 11 times of the computation time of the original non-source-oriented mechanism. While efficient source-oriented approach for primary particles are already available to track a large number of sources simultaneously, the efficient approach developed in this study has the potential to track a large number of sources to evaluate their impact on secondary pollutant formation, and has the potential to be applied in air quality forecasting models that provide source or source-region contribution information for policy makers for better emission regulations under adverse meteorological conditions.

Data availability. Data can be obtained upon request from the authors.

Authorship contributions. QX: Investigation, Methodology, Software, Visualization, Writing & editing. FS: Writing, Investigation, Conceptualization, Methodology, Software, Validation. KW: Supervision, Writing - review & editing, Visualization. RZ: Resources, Supervision, Visualization. QY: Software, Data curation, Validation, Visualization, Writing - review & editing. M. K: Software, Methodology.

Declaration of competing interest. The contact author has declared that neither they nor their co-authors have any competing interests.

Acknowledgments

to thank Texas A&M High Performance Computing Center for providing the computation resources.

References

Carter, W. P. L.: Development of the SAPRC-07 chemical mechanism, Atmos. Environ., 44(40), 5324-5335. https://doi.org/10.1016/j.atmosenv.2010.01.026, 2010.

Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., & Geron, C.: Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature) [Review]. Atmos Chem Phys, https://doi.org/10.5194/acp-6-3181-2006, 2006.

Hertel, O., Berkowicz, R., Christensen, J., & Hov, Ø.: Test of two numerical schemes for use in atmospheric transport-chemistry models. Atmos. Environ. Part A. General Topics, 27(16), 2591-2611. https://doi.org/10.1016/0960-1686(93)90032-T, 1993.

Hu, J., Howard, C. J., Mitloehner, F., Green, P. G., & Kleeman, M. J.: Mobile Source and Livestock Feed Contributions to Regional Ozone Formation in Central California. Environ. Sci. Technol., 46(5), 2781-2789. https://doi.org/10.1021/es203369p, 2012.

Hu, J., Zhang, H., Chen, S., Ying, Q., Wiedinmyer, C., Vandenberghe, F., & Kleeman, M. J.: Identifying PM2.5 and PM0.1 Sources for Epidemiological Studies in California. Environ. Sci. Technol., 48(9), 4980-4990. https://doi.org/10.1021/es404810z, 2014.

Jacobson, M. Z: Fundamentals of Atmospheric Modeling (Second Edition ed.). Cambridge University Press., 2006.

Jacobson, M. Z., & Turco, R. P.: SMVGEAR: A sparse-matrix, vectorized gear code for atmospheric models. Atmos. Environ., 28(2), 273-284. https://doi.org/10.1016/1352-2310(94)90102-3, 1994.

Kleeman, M. J., & Cass, G. R.: A 3D Eulerian source-oriented model for an externally mixed aerosol. Environ. Sci. Technol., 35(24), 4834-4848. https://doi.org/10.1021/es010886m, 2001.

Li, L., Hu, J., Li, J., Gong, K., Wang, X., Ying, Q., Qin, M., Liao, H., Guo, S., Hu, M., & Zhang, Y.: Modelling air quality during the EXPLORE-YRD campaign – Part II. Regional source apportionment of ozone and PM2.5. Atmos. Environ., 247. https://doi.org/10.1016/j.atmosenv.2020.118063,2021.

Lu, X., Gao, D., Liu, Y., Wang, S., Lu, Q., Yin, S., Zhang, R., & Wang, S.: A recent high-resolution PM2.5 and VOCs speciated emission inventory from anthropogenic sources: A case study of central China. J. Cleaner Prod., 386. https://doi.org/10.1016/j.jclepro.2022.135795, 2023.

Shi, Z., Li, J., Huang, L., Wang, P., Wu, L., Ying, Q., Zhang, H., Lu, L., Liu, X., Liao, H., & Hu, J.: Source apportionment of fine particulate matter in China in 2013 using a source-oriented

502   chemical transport model. Sci. Total Environ., 601-602, 1476-1487. https://doi.org/10.1016/j.scitot

503   env.2017.06.019, 2017.

504   Ralph, A. W.: Numerical Initial Value Problems in Ordinary Differential Equations (C. William Gear). SIAM

505   Review., 15(3), 676-678. https://doi.org/10.1137/1015088, 1973.

506   Wang, P., Chen, Y., Hu, J., Zhang, H., & Ying, Q.: Attribution of Tropospheric Ozone to NOx and VOC

507   Emissions: Considering Ozone Formation in the Transition Regime. Environ. Sci. Technol., 53(3), 1404-

508   1412. https://doi.org/10.1021/acs.est.8b05981, 2019a.

509   Wang, P., Chen, Y., Hu, J., Zhang, H., & Ying, Q.: Source apportionment of summertime ozone in China using

510   a   source-oriented   chemical   transport   model.   Atmos.   Environ.,   211,   79-90.

511   https://doi.org/10.1016/j.atmosenv.2019.05.006, 2019b.

512   Wang, P., Wang, T., & Ying, Q.: Regional source apportionment of summertime ozone and its precursors in

513   the megacities of Beijing and Shanghai using a source-oriented chemical transport model. Atmos.

514   Environ., 224, 117337. https://doi.org/10.1016/j.atmosenv.2020.117337, 2020.

515   Wang, P., Ying, Q., Zhang, H., Hu, J., Lin, Y., & Mao, H.: Source Apportionment of Secondary Organic

516   Aerosol in China using a Regional Chemical Transport Model and Two Emission Inventories. Environ.

517   Pollut., 237, 756-766. https://doi.org/10.1016/j.envpol.2017.10.122, 2018.

518   Ying, Q., & Kleeman, M. J.: Source contributions to the regional distribution of secondary particula

519   te matter in California [Article]. Atmos. Environ., 40(4), 736-752. https://doi.org/10.1016/j.atmose

520   nv.2005.10.007, 2006.

521   Ying, Q., & Krishnan, A.: Source contributions of volatile organic compounds to ozone formation in southeast

522   Texas. J Geophys Res-atmos, 115(D17), D17306. https://doi.org/10.1029/2010jd013931, 2010.

523   Ying, Q., Mysliwiec, M., & Kleeman, M. J.: Source apportionment of visibility impairment using a three-

524   dimensional   source-oriented   air   quality   model.   Environ.   Sci.   Technol.,   38(4),   1089-1101.

525   https://doi.org/10.1021/es0349305, 2004.

526   Ying, Q., Wu, L., & Zhang, H.: Local and inter-regional contributions to PM 2.5 nitrate and sulfate in China.

527   Atmos. Environ., 94, 582-592. https://doi.org/10.1016/j.atmosenv.2014.05.078, 2014.

528   Zhang, H., Linford, J. C., Sandu, A., & Sander, R.: Chemical Mechanism Solvers in Air Quality Models.

529   Atmos. Environ., 2(3), 510-532. https://doi.org/10.3390/atmos2030510, 2011.

530   Zhang, H. L., & Ying, Q.: Contributions of local and regional sources of NOx to ozone concentrati

531   ons in Southeast Texas. Atmos. Environ., 45(17), 2877-2887. https://doi.org/10.1016/j.atmosenv.20

532   11.02.047, 2011.

533 **List 1.** The special reaction rate section (between the keywords SPECIAL and END SPECIAL)
534 and reactions used to implement the source-oriented NO2+NO3 reactions with 10 source types
535 using the chemical mechanism preprocessor CHEMMECH for the CMAQ model. Due to the
536 limitation of the current mechanism preprocessor, a dummy reaction (<10_dum>) is needed
537 so that the original reaction rate can be included in calculation of the special rate constants.
538 Using the special rate expression is signaled by including the symbol '?' in the reaction rate
539 coefficient expression.

```
SPECIAL =
RNO_NO3 = K<10_dum>*C<NO3> + K<10_dum>*C<NO3_X1> +
<10_dum>*C<NO3_X2>
        + K<10_dum>*C<NO3_X3> + K<10_dum>*C<NO3_X4> +
K<10_dum>*C<NO3_X5>
        + K<10_dum>*C<NO3_X6> + K<10_dum>*C<NO3_X7> +
K<10_dum>*C<NO3_X8>
        + K<10_dum>*C<NO3_X9>;
RNO3_NO = K<10_dum>*C<NO> + K<10_dum>*C<NO_X1> +
K<10_dum>*C<NO_X2> +
        K<10_dum>*C<NO_X3> + K<10_dum>*C<NO_X4> +
K<10_dum>*C<NO_X5> +
        K<10_dum>*C<NO_X6> + K<10_dum>*C<NO_X7> +
K<10_dum>*C<NO_X8> +
        K<10_dum>*C<NO_X9>;
...
END SPECIAL
...
<10_dum> dummy1 + dummy1 = dummy1 + dummy1 #1.80e-11@-
110;
<10_aX0> NO = NO2 #1.0?RNO_NO3;
<10_aX1> NO_X1 = NO2_X1 #1.0?RNO_NO3;
<10_aX2> NO_X2 = NO2_X2 #1.0?RNO_NO3;
...
<10_aX9> NO_X9 = NO2_X9 #1.0?RNO_NO3;
<10_bX0> NO3 = NO2 #1.0?RNO3_NO;
<10_bX1> NO3_X1 = NO2_X1 #1.0?RNO3_NO;
<10_bX2> NO3_X2 = NO2_X2 #1.0?RNO3_NO;
...
<10_bX9> NO3_X9 = NO2_X9 #1.0?RNO3_NO;
```

540