



Supplement of

The first application of a numerically exact, higher-order sensitivity analysis approach for atmospheric modelling: implementation of the hyperdual-step method in the Community Multiscale Air Quality Model (CMAQ) version 5.3.2

Jiachen Liu et al.

Correspondence to: Shannon L. Capps (shannon.capps@drexel.edu)

The copyright of individual parts of the supplement might differ from the article licence.

1. Operations with hyperdual numbers

Examples of operations including addition, multiplication, exponential, natural log, power, and sine functions for hyperdual numbers $H_a = a_0 + a_1\epsilon_1 + a_2\epsilon_2 + a_{12}\epsilon_{12}$ and $H_b = b_0 + b_1\epsilon_1 + b_2\epsilon_2 + b_{12}\epsilon_{12}$ are listed in Eqs. S1-S6. These operations are derivations based on Fike and Alonso (2011).

$$H_a + H_b = a_0 + b_0 + (a_1 + b_1)\epsilon_1 + (a_2 + b_2)\epsilon_2 + (a_{12} + b_{12})\epsilon_{12} \quad (S1)$$

$$H_a H_b = a_0 b_0 + (a_1 b_0 + a_0 b_1)\epsilon_1 + (a_2 b_0 + a_0 b_2)\epsilon_2 + (a_0 b_{12} + a_1 b_2 + a_2 b_1 + a_{12} b_0)\epsilon_{12} \quad (S2)$$

$$\exp(H_a) = \exp(a_0) + [a_1 \exp(a_0)]\epsilon_1 + [a_2 \exp(a_0)]\epsilon_2 + [(a_1 a_2 + a_{12})\exp(a_0)]\epsilon_{12} \quad (S3)$$

$$\ln(H_a) = \ln(a_0) + \left(\frac{a_1}{a_0}\right)\epsilon_1 + \left(\frac{a_2}{a_0}\right)\epsilon_2 + \left(\frac{a_{12}}{a_0} - a_1 a_2\right)\epsilon_{12} \quad (S4)$$

$$H_a^{H_b} = \exp[H_b \ln(H_a)] \quad (S5)$$

$$\sin(H_a) = \sin(a_0) + [a_1 \cos(a_0)]\epsilon_1 + [a_2 \cos(a_0)]\epsilon_2 + [a_{12} \cos(a_0) - a_1 a_2 \sin(a_0)]\epsilon_{12} \quad (S6)$$

- 5 The power function (Eq. S5) is defined with both the natural log and exponential to avoid explicitly writing out the operations, which could be very complicated.

The Taylor series expansion of the multiplicative hyperdual perturbation to the real value, x , (Eq. 5) is shown in Eq. S7.

$$\begin{aligned} f(xH_h) &= f(x * (1.0 + h_1\epsilon_1 + h_2\epsilon_2)) \\ &= f(x + xh_1\epsilon_1 + xh_2\epsilon_2) \\ &= f(x) + (xh_1\epsilon_1 + xh_2\epsilon_2)f'(x) + \frac{1}{2!}(xh_1\epsilon_1 + xh_2\epsilon_2)^2 f''(x) \\ &\quad + \frac{1}{3!}(xh_1\epsilon_1 + xh_2\epsilon_2)^3 f'''(x) + \dots \\ &= f(x) + (xh_1\epsilon_1 + xh_2\epsilon_2)f'(x) + \frac{1}{2}(x^2 h_1^2 \epsilon_1^2 + 2x^2 h_1 \epsilon_1 h_2 \epsilon_2 + x^2 h_2^2 \epsilon_2^2) f''(x) \\ &= f(x) + xh_1\epsilon_1 f'(x) + xh_2\epsilon_2 f'(x) + x^2 h_1 h_2 \epsilon_{12} f''(x) \end{aligned} \quad (S7)$$

2. Slopes and R² values of additional one-to-one plots

- 10 The slopes of additional first-order (Table S1) and second-order (Table S2) FDM and HYD sensitivities comparisons are shown.

Table S1. Slope and R² of the comparisons of sensitivities of ground layer species concentrations to domain-wide perturbations

Sensitivities: slope, R ²			
NO _x	SO ₂	2 nd -order NO _x	2 nd -order SO ₂
$S_{NO_x}^{ANH_4}$: 0.99, 0.94	$S_{SO_2}^{ANH_4}$: 1.00, 0.97	$S_{NO_x}^{(2)ANH_4}$: 0.74, 0.67	$S_{SO_2}^{(2)ANH_4}$: 0.04, 0.04
$S_{NO_x}^{ANO_3}$: 1.00, 0.99	$S_{SO_2}^{ANO_3}$: 0.10, 0.30	$S_{NO_x}^{(2)ANO_3}$: 0.61, 0.38	$S_{SO_2}^{(2)ANO_3}$: 0.01, 0.00

$s_{NO_x}^{ASO_4}$: 1.00, 0.94	$s_{SO_2}^{ASO_4}$: 1.04, 1.00	$s_{NO_x}^{(2)ASO_4}$: 0.98, 0.96	$s_{SO_2}^{(2)ASO_4}$: 0.04, 0.06
---------------------------------	---------------------------------	------------------------------------	------------------------------------

15 **Table S2.** Slope and R^2 of the comparisons of ground layer species concentrations to domain-wide perturbations

Sensitivities: slope, R^2			
TERP	APIN	2 nd -order TERP	2 nd -order APIN
s_{TERP}^{AMT1} : 1.00, 1.00	s_{APIN}^{AMT1} : 1.00, 1.00	$s_{TERP}^{(2)AMT1}$: 0.50, 0.49	$s_{APIN}^{(2)AMT1}$: 0.51, 0.50
s_{TERP}^{AMT2} : 1.00, 1.00	s_{APIN}^{AMT2} : 1.00, 1.00	$s_{TERP}^{(2)AMT2}$: 0.38, 0.38	$s_{APIN}^{(2)AMT2}$: 0.40, 0.39
s_{TERP}^{AMT3} : 1.01, 1.00	s_{APIN}^{AMT3} : 1.01, 1.00	$s_{TERP}^{(2)AMT3}$: 0.84, 0.72	$s_{APIN}^{(2)AMT3}$: 0.85, 0.74
s_{TERP}^{AMT4} : 1.01, 1.00	s_{APIN}^{AMT4} : 1.02, 1.00	$s_{TERP}^{(2)AMT4}$: 1.02, 0.94	$s_{APIN}^{(2)AMT4}$: 1.06, 0.94
s_{TERP}^{AMT5} : 1.02, 1.00	s_{APIN}^{AMT5} : 1.03, 1.00	$s_{TERP}^{(2)AMT5}$: 1.03, 0.95	$s_{APIN}^{(2)AMT5}$: 1.07, 0.95
s_{TERP}^{AMT6} : 1.02, 1.00	s_{APIN}^{AMT6} : 1.03, 1.00	$s_{TERP}^{(2)AMT6}$: 1.03, 0.95	$s_{APIN}^{(2)AMT6}$: 1.06, 0.95

Table S3. Individual module wall time of regular CMAQ (REG) and CMAQ-hyd (HYD) runs.

	HYD, 1 (s)	HYD, 2 (s)	HYD, 4 (s)	HYD, 8 (s)	REG, 1 (s)	REG, 2 (s)	REG, 4 (s)	REG, 8 (s)
Chem	5810	2930	1490	725	3180	1610	794	403
Aero	3840	1890	960	487	793	385	206	104
Vdiff	3860	1870	924	487	1390	724	393	217
Hadv	1270	757	474	320	589	334	216	160
Phot	462	205	124	64.0	235	111	53.4	28.9
Cldproc	301	173	94.8	56.9	149	81.4	41.3	27.4
Hdiff	221	141	83.0	56.9	70.6	47.1	29.1	23.1
Zadv	422	238	124	56.9	94.2	51.4	26.7	13.0
MPI_Barrier	2630	1830	1160	853	1160	796	561	368
Other	1290	778	492	448	188	137	107	99.6

3. Additional Figures

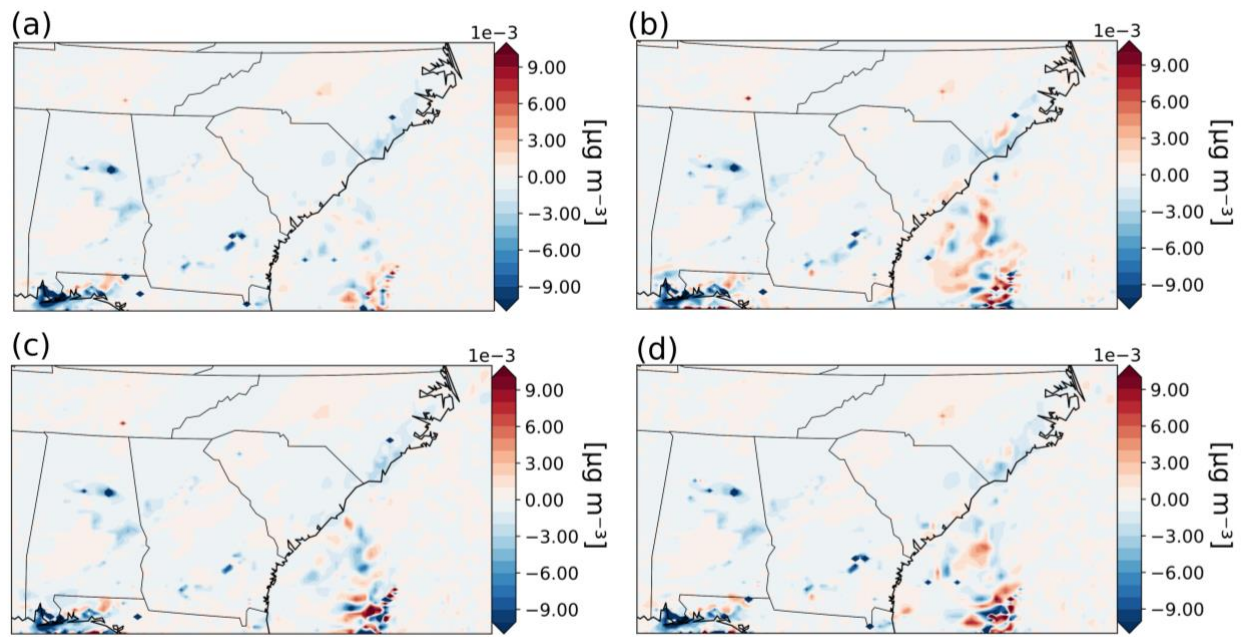


Figure S1. The first-order sensitivities of ground layer aerosol nitrate (ANO₃) concentration with respect to domain-wide perturbation of SO₂ emission calculated by FDM with different perturbation sizes: (a) 125%, 75%; (b) 110%, 90%; (c) 125%, 100%; (d) 100%, 75%. The sensitivities are apparently noisy, especially over the coast of South Carolina and Georgia.

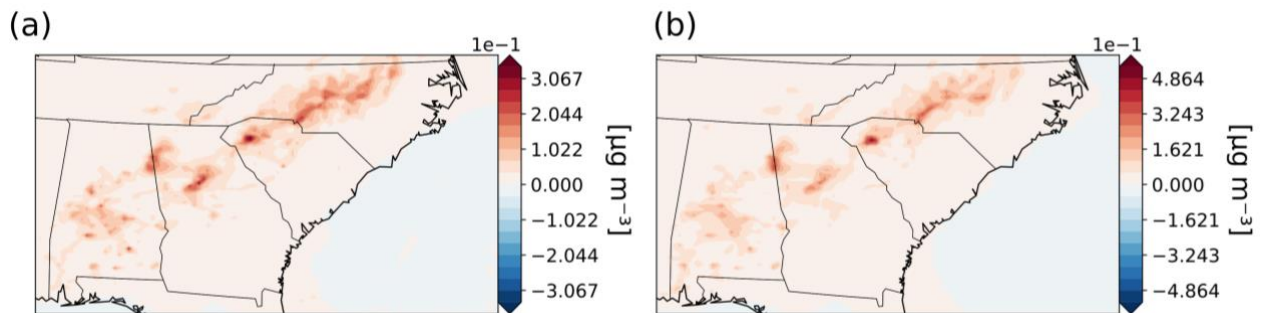


Figure S2. The first order sensitivities of (a): AMTNO₃ with respect to NO_x, $s_{NO_x}^{AMTNO_3}$ and (b): AMTNO₃ with respect to monoterpenes, $s_{TERP}^{AMTNO_3}$

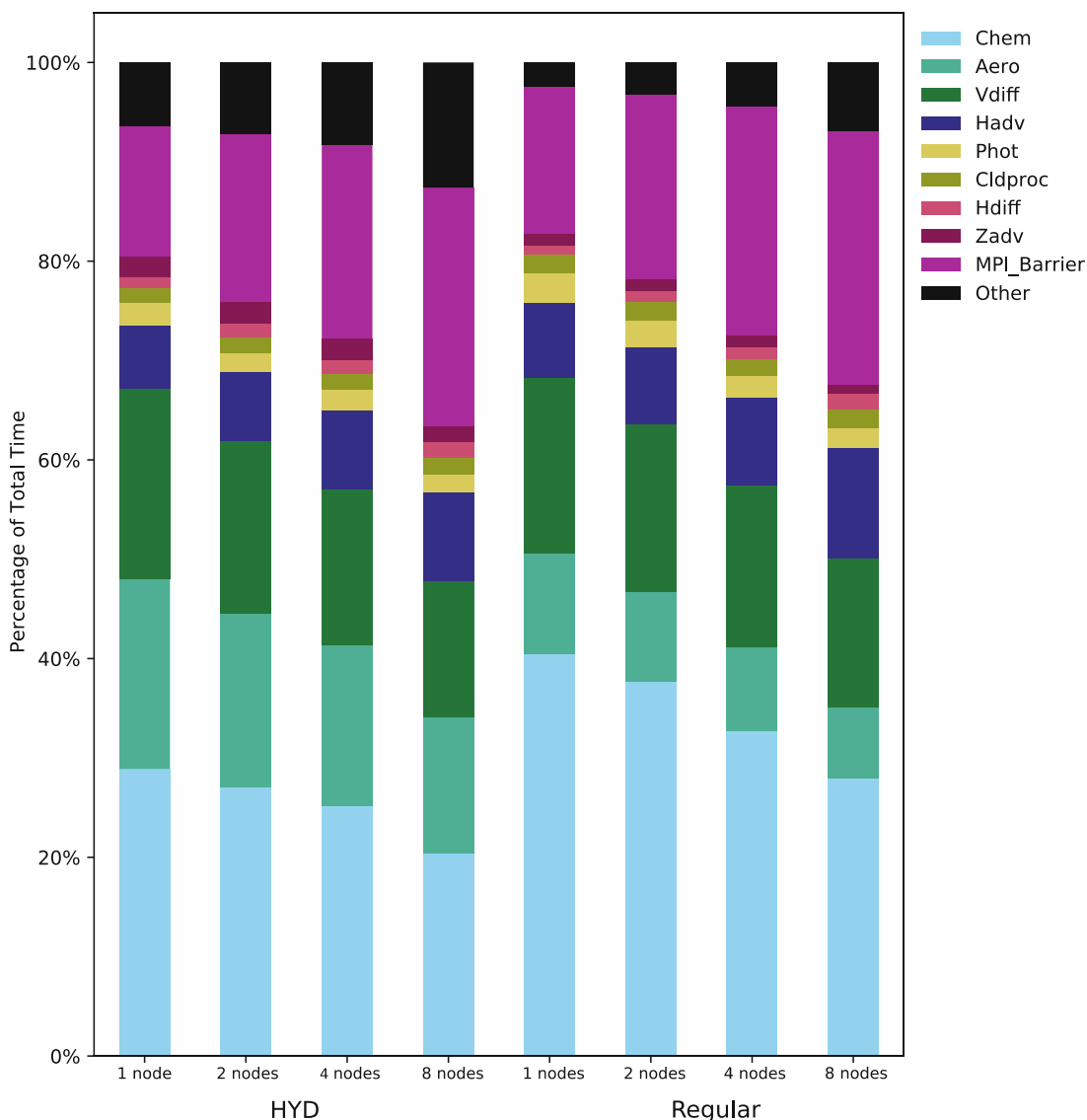


Figure S3. Computational cost of CMAQ-hyd and regular CMAQ separated by different modules. The CMAQ-hyd and regular CMAQ simulations with different number of 1, 2, 4, and 8 nodes and are shown on the x-axis. The percentage is shown on the y-axis.

4. High-Performance Computing Clusters

30 The simulations were run on NCAR's Cheyenne Supercomputing Clusters (Computational Information Systems
 Laboratory, 2017). Cheyenne has 145,152 processor cores of 2.3 GHz Intel Xeon E5-2697V4 processors. The
 total system memory is 313 TB, with 64 GB per node on 3,168 nodes and 128 GB per node on 864 nodes. All the
 simulations in this work were run with 64 GB of memory per node. The runs using one, two, four, and eight nodes
 use 36 cores, 72 cores, 144 cores, and 288 cores of processors, respectively. CMAQ delegates the computing job
 35 of the grid by partitioning the horizontal grid. Each processor is tasked with executing the scientific processes

before communicating the information on the boundaries. The specific grid allocation for each processor, in terms of columns and rows, are as follows: 6x6 for one node, 12x6 for two nodes, 12x12 for four nodes, and 24x12 for eight nodes in our domain. We applied the same computational configuration for the regular CMAQ and the CMAQ-hyd runs during code profiling processes.

40

References

- 45 Computational Information Systems Laboratory: Cheyenne: SGI ICE XA Cluster, 10.5065/D6RX99HX, 2017.
Fike, J. and Alonso, J.: The Development of Hyper-Dual Numbers for Exact Second-Derivative Calculations, 49th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition, 2011-01-04, 10.2514/6.2011-886,