

Figure S1. The cost function Z_2 with different block number N. Results are shown for the first day of February, April, August and November sampled every 6 hours. Shaded area denotes for 1 standard deviation.

1 2

3

4

5

6 7

8

9

10

GEOS-Chem species in 12 blocks IMO2H CH2I2 VINAO3 RP TOLU PPIMAE IPMN LBRO2H LBRO2N LBRO2N BRO2 R4M1LTRO2N BENZ CH2CI2 HOI CI2O2 I2O2 BrO OIO 12 11 HBr CINO3 CI OCIO BrNO2 I CIOO Br2 BrNO3 BrCI SO4H1 2 N IONITA INDIOL R4N2 RCO3 IONO CI2 HCI CIO BL HOCI 10 BrSALC ISALA INO ICI ALAINO ICI ISALC HI 1203 BrSALA 12 1204 AERI IBr CINO2 3 CH3CHOO OLNN PRPN ISOPNB MONITU RIO2 OLND CHB/3 RIPB HONIT CHC/3 CCC LISOPOH LISOPNOS HCFC22 ISOP RIPA PROPNN PRN1 MONITS HPALD CH2/C CH2/OO MONITA HC5/ISN1 HC5/OO INO2/ISN1 LIMO OH NO NO3 HO2 N2O5 HNO3 NO2 O3 H2O MPN MO2 CH2O PRPE SO4 ACET C2H6 ETO2 MAP ATOOH ALD2 DMS CO2 CO3 CO CH4 H ACTA ATO2 MSA O HNO4 N2C PAN O1D HNO2 MP G ISNOOAINPN MACROO ISN10G CCI DHPCARP MACRNO ISOPND CH3Br CH2IB* HCFC1414 ISOPNDO: NMAO MAOPO CH2Br PMNN* ISNP ISOPNBO HCFC142b GAOO MOBA ISOPNBO HCFC142b GAOO MOBA MVKOO HCFC ISNOHOO MAN MGLOO DIBOO MACRO HCCCI LYOC PYAC CFC DHDN LIMO MRP ISNOOB LVOCOA NPMN MACRN ISNOB LVOCOA NPMN IAP IEPOXD MOBAOO 7 IEPOXA DHMOB SOAGX ROHHCOOH XYLE MTPA SOAME MTPO GLYCXRO2 LXRO2N RCHO MGLY KO2 ALK4 MEK RB3P CH3CI R4P CH3I R4O2 PIO2 IEPOXOO MVK VRPHC187 DHDC IEPOXB

Figure S2. GEOS-Chem species in a 12-block system. The 228 species are split into 12 blocks based on the similarity of production and loss rates using the simulated annealing algorithm. More details can be found in text. The full names of these acronyms can be found at http://wiki.seas.harvard.edu/geos-chem/index.php/Species in GEOS-Chem.

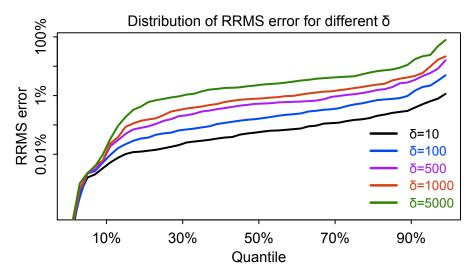


Figure S3. Distribution of RRMS error for different thresholds δ (N=12 and M=20). The y axis is on the logarithmic scale.

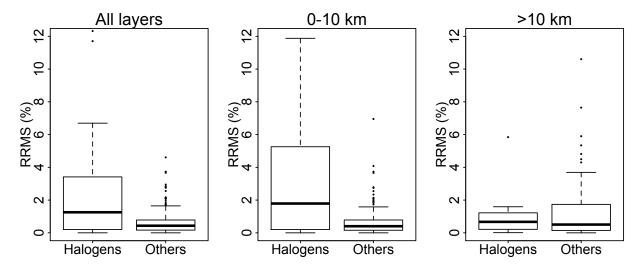


Figure S4. The RRMS error of halogen and non-halogen species for (a) all layers, (b) 0-10 km and (c) above 10 km. Results are shown for the daily-average concentrations of all species on the last day of the 2-year simulations, using 20 chemistry regimes and a threshold δ of 500 molecules cm⁻³ s⁻¹.

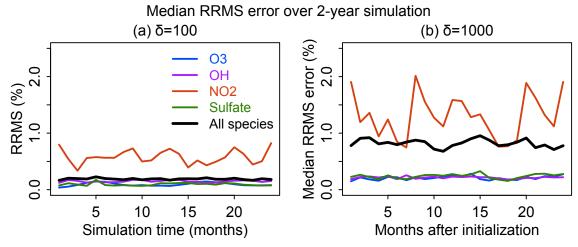


Figure S5. Same as Figure 4, but using a threshold of 100 and 1000 molecules cm⁻³ s⁻¹ to partition the fast and slow species.