Error estimate

Beijing 2012-2013

Once obtained the solution

- strong: Cl, K, Ca, Ti, Fe, SO24, NO3, NH4, OC, and EC.
- weak: Na, Mg, Al, Si, V, Cr, Mn, Co, Ni, Cu, Zn, As,Br, Pb. PM2.5 was set as the total variable
- 10% extra uncertainty
- 6 factors
- Seed: 20

Error estimate – Base run

- BS (random errors)
- DISP (rotational errors)
- BS-DISP (random+rotational)

- Paatero, P., Eberly, S., Brown, S.G., Norris, G.A. Methods for estimating uncertainty in factor analytic solutions (2014) Atmospheric Measurement Techniques, 7 (3), pp. 781-797.

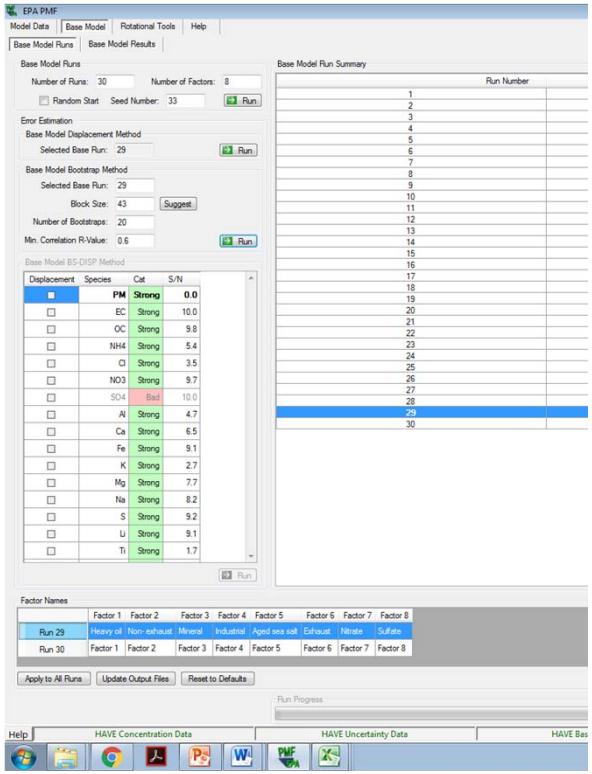
- Brown, S.G., Eberly, S., Paatero, P., Norris, G.A. Methods for estimating uncertainty in PMF solutions: Examples with ambient air and water quality data and guidance on reporting PMF results (2015) Science of the Total Environment, 518-519, pp. 626-635.

Output error estimate

1. Random errors (Bootstrapping).
2. Rotational errors (Displacement)
3. Both (BS-DISP)

Modelling errors difficult to estimate, and they are not minimized by large datasets

If uncertainties
are not well specified, use only
BS-DISP or BS



Bootstrapping diagnostic

- Random errors, partially evaluates the rotational ambiguity
- not influenced by the user-specified sample uncertainties.
- >80% mapped factors

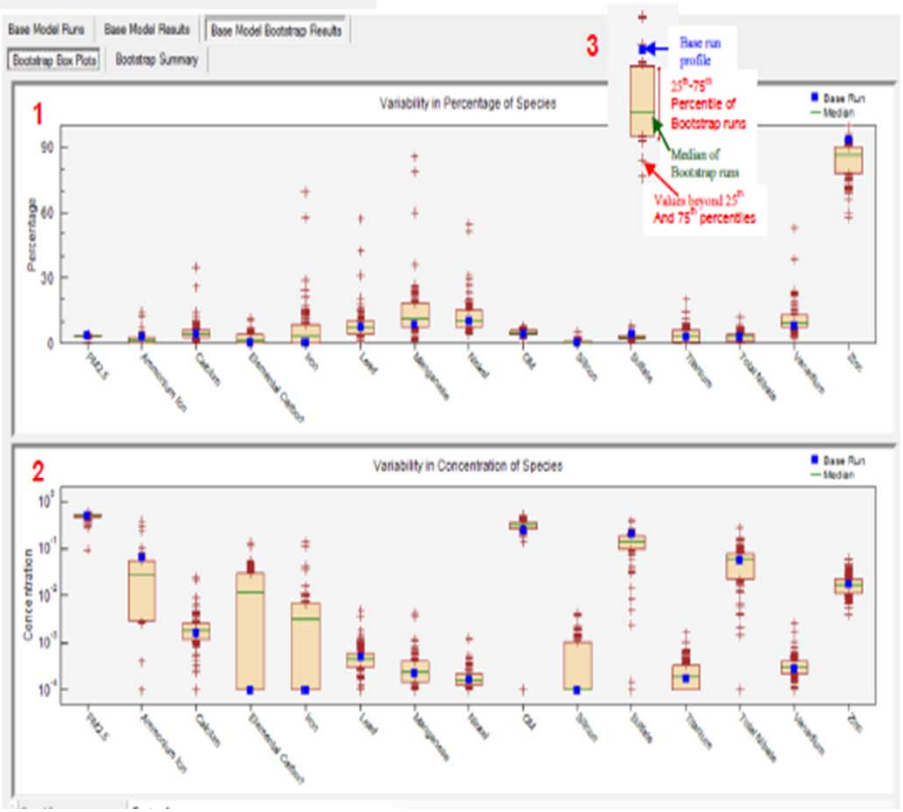
EPA PMF

Model Data | Base Model | Rotational Tools | Help

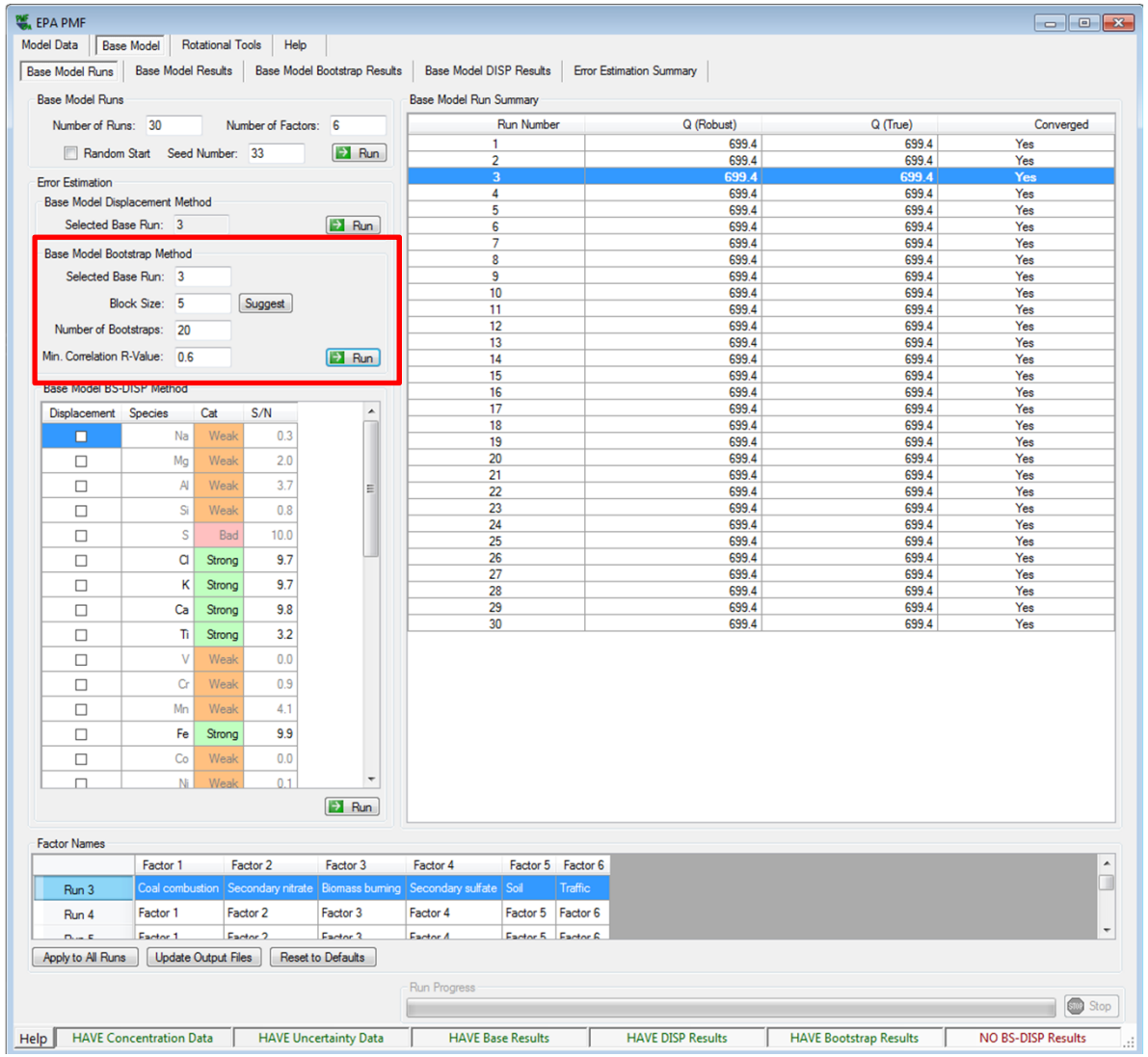
Base Model Runs | Base Model Results | Base Model DISP Results | Base Model Bootstrap Results | Error Estimation Summary

Bootstrap Box Plots | Bootstrap Summary

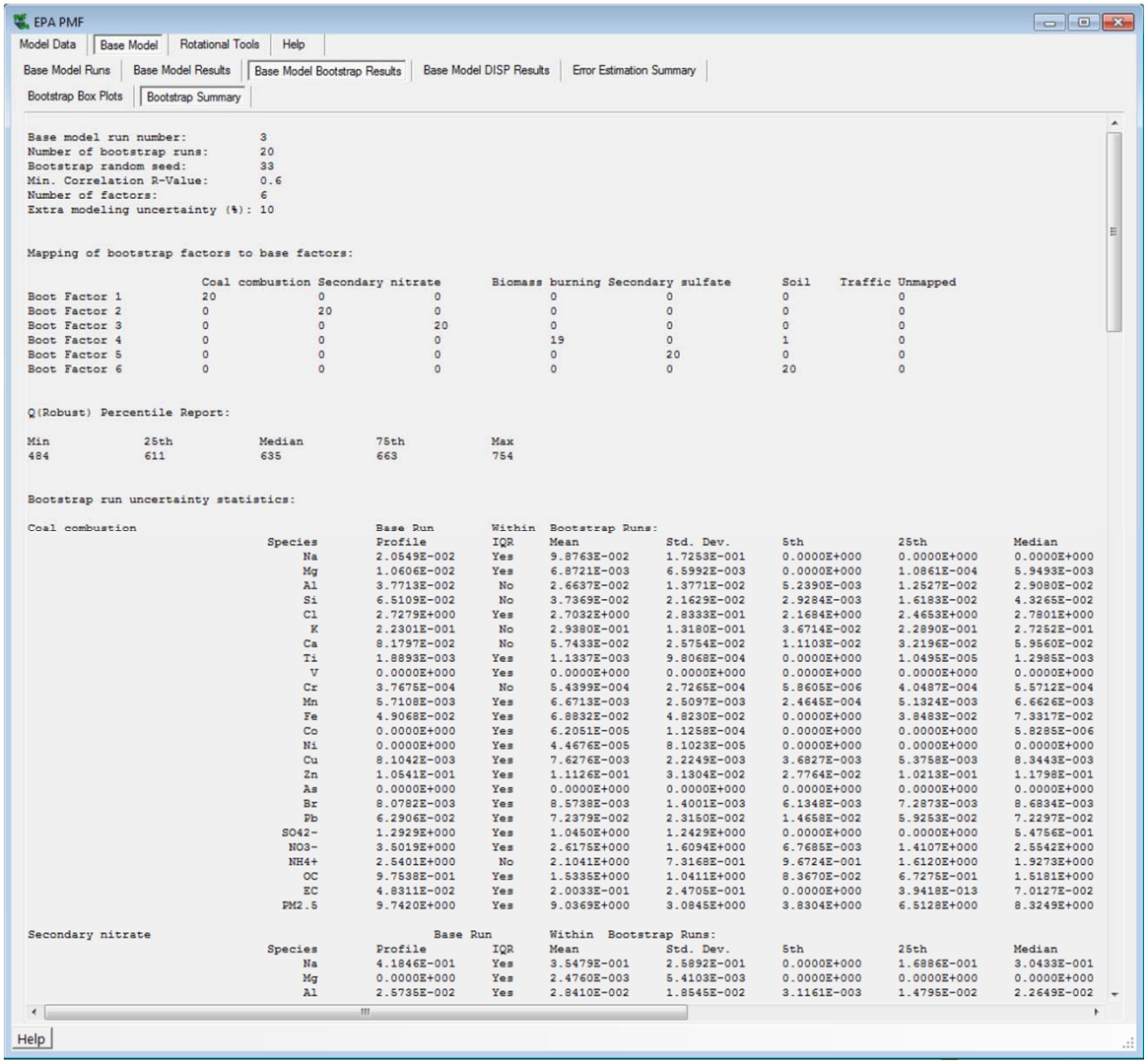
```
Base model run number:          29
Number of bootstrap runs:       20
Bootstrap random seed:          33
Min. Correlation R-Value:       0.6
Number of factors:              8
Extra modeling uncertainty (%): 7


Mapping of bootstrap factors to base factors:

                    Heavy oil       Non- exhaust    Mineral Industrial      Aged sea salt   Exhaust Nitrate Sulfate Unmapped
Boot Factor 1           20              0               0               0               0               0               0               0               0
Boot Factor 2           0               20              0               0               0               0               0               0               0
Boot Factor 3           0               0               20              0               0               0               0               0               0
Boot Factor 4           0               0               0               20              0               0               0               0               0
Boot Factor 5           0               0               0               0               20              0               0               0               0
Boot Factor 6           0               1               0               0               0               19              0               0               0
Boot Factor 7           0               0               0               0               0               0               20              0               0
Boot Factor 8           0               0               0               0               1               0               0               19              0


Q(Robust) Percentile Report:

Min             25th            Median          75th            Max
5611            6214            6464            6663            6978
```



Bootstraap base run

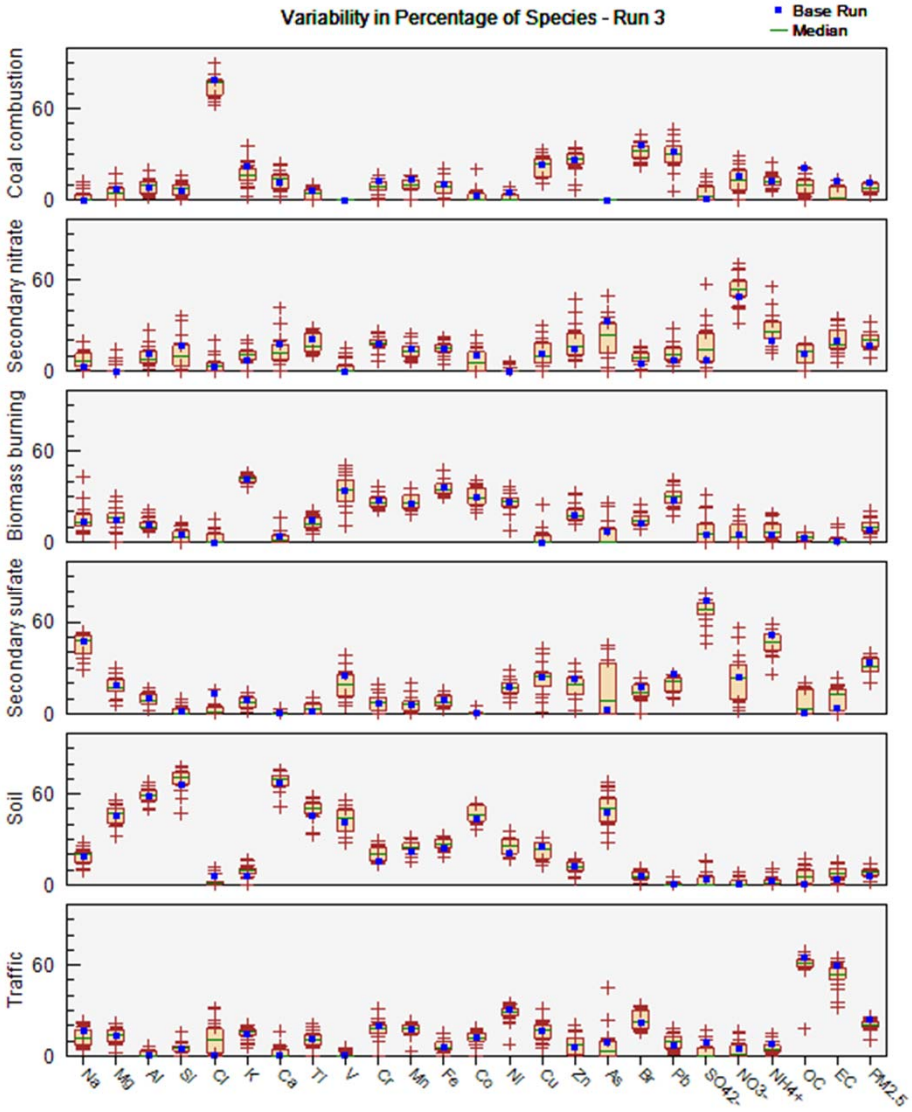


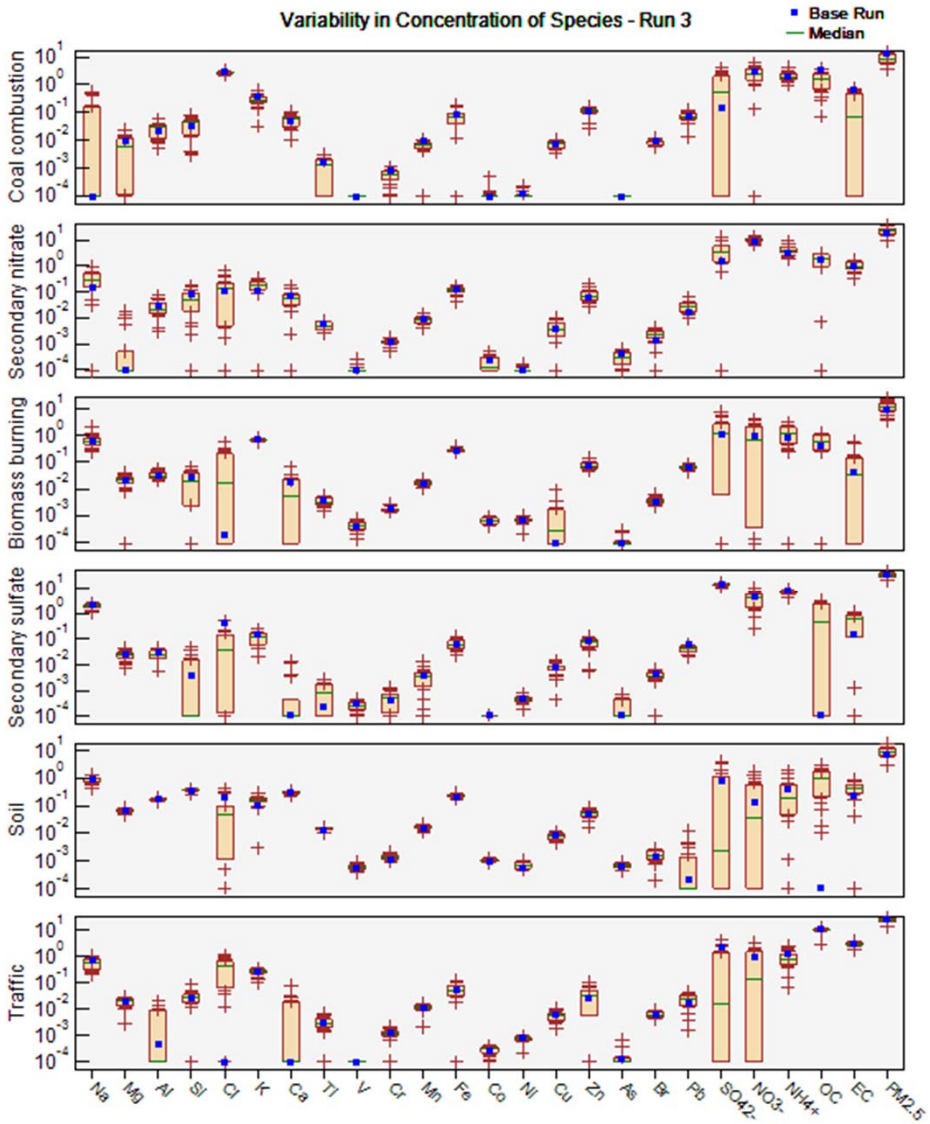
BS summary


EPA PMF
Model Data
Base Model
Rotational Tools
Help
Base Model Runs
Base Model Results
Base Model Bootstrap Results
Base Model DISP Results
Error Estimation Summary
Bootstrap Box Plots
Bootstrap Summary
Base model run number: 3
Number of bootstrap runs: 20
Bootstrap random seed: 33
Min. Correlation R-Value: 0.6
Number of factors: 6
Extra modeling uncertainty (%): 10
Mapping of bootstrap factors to base factors:
Coal combustion Secondary nitrate Biomass burning Secondary sulfate Soil Traffic Unmapped
Boot Factor 1 20 0 0 0 0 0 0
Boot Factor 2 0 20 0 0 0 0 0
Boot Factor 3 0 0 20 0 0 0 0
Boot Factor 4 0 0 0 19 0 1 0
Boot Factor 5 0 0 0 0 20 0 0
Boot Factor 6 0 0 0 0 0 20 0
Q(Robust) Percentile Report:
Min 25th Median 75th Max
484 611 635 663 754
Bootstrap run uncertainty statistics:
Coal combustion
Species Base Run Profile Within IQR Bootstrap Runs: Mean Std. Dev. 5th 25th Median
Na 2.0549E-002 Yes 9.8763E-002 1.7253E-001 0.0000E+000 0.0000E+000 0.0000E+000
Mg 1.0606E-002 Yes 6.8721E-003 6.5992E-003 0.0000E+000 1.0861E-004 5.9493E-003
Al 3.7713E-002 No 2.6637E-002 1.3771E-002 5.2390E-003 1.2527E-002 2.9080E-002
Si 6.5109E-002 No 3.7369E-002 2.1629E-002 2.9284E-003 1.6183E-002 4.3265E-002
Cl 2.7279E+000 Yes 2.7032E+000 2.8333E-001 2.1684E+000 2.4653E+000 2.7801E+000
K 2.2301E-001 No 2.9380E-001 1.3180E-001 3.6714E-002 2.2890E-001 2.7252E-001
Ca 8.1797E-002 No 5.7433E-002 2.5754E-002 1.1103E-002 3.2196E-002 5.9560E-002
Ti 1.8893E-003 Yes 1.1337E-003 9.8068E-004 0.0000E+000 1.0495E-005 1.2985E-003
V 0.0000E+000 Yes 0.0000E+000 0.0000E+000 0.0000E+000 0.0000E+000 0.0000E+000
Cr 3.7675E-004 No 5.4399E-004 2.7265E-004 5.8605E-006 4.0487E-004 5.5712E-004
Mn 5.7108E-003 Yes 6.6713E-003 2.5097E-003 2.4645E-004 5.1324E-003 6.6626E-003
Fe 4.9068E-002 Yes 6.8832E-002 4.8230E-002 0.0000E+000 3.8483E-002 7.3317E-002
Co 0.0000E+000 Yes 6.2051E-005 1.1258E-004 0.0000E+000 0.0000E+000 5.8285E-006
Ni 0.0000E+000 Yes 4.4676E-005 8.1023E-005 0.0000E+000 0.0000E+000 0.0000E+000
Cu 8.1042E-003 Yes 7.6276E-003 2.2249E-003 3.6827E-003 5.3758E-003 8.3443E-003
Zn 1.0541E-001 Yes 1.1126E-001 3.1304E-002 2.7764E-002 1.0213E-001 1.1798E-001
As 0.0000E+000 Yes 0.0000E+000 0.0000E+000 0.0000E+000 0.0000E+000 0.0000E+000
Br 8.0782E-003 Yes 8.5738E-003 1.4001E-003 6.1348E-003 7.2873E-003 8.6834E-003
Pb 6.2906E-002 Yes 7.2379E-002 2.3150E-002 1.4658E-002 5.9253E-002 7.2297E-002
SO42- 1.2929E+000 Yes 1.0450E+000 1.2429E+000 0.0000E+000 0.0000E+000 5.4756E-001
NO3- 3.5019E+000 Yes 2.6175E+000 1.6094E+000 6.7685E-003 1.4107E+000 2.5542E+000
NH4+ 2.5401E+000 No 2.1041E+000 7.3168E-001 9.6724E-001 1.6120E+000 1.9273E+000
OC 9.7538E-001 Yes 1.5335E+000 1.0411E+000 8.3670E-002 6.7275E-001 1.5181E+000
EC 4.8311E-002 Yes 2.0033E-001 2.4705E-001 0.0000E+000 3.9418E-013 7.0127E-002
PM2.5 9.7420E+000 Yes 9.0369E+000 3.0845E+000 3.8304E+000 6.5128E+000 8.3249E+000
Secondary nitrate
Species Base Run Profile Within IQR Bootstrap Runs: Mean Std. Dev. 5th 25th Median
Na 4.1846E-001 Yes 3.5479E-001 2.5892E-001 0.0000E+000 1.6886E-001 3.0433E-001
Mg 0.0000E+000 Yes 2.4760E-003 5.4103E-003 0.0000E+000 0.0000E+000 0.0000E+000
Al 2.5735E-002 Yes 2.8410E-002 1.8545E-002 3.1161E-003 1.4795E-002 2.2649E-002
Help

BS box plots

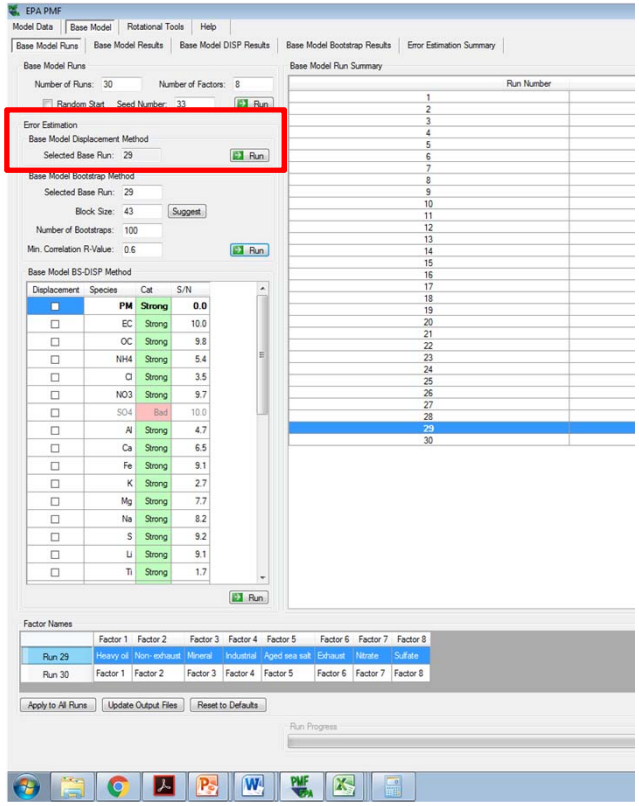
Species with base run outside the IQR need evaluation





DISP

Only on strong species.
Valid only if data uncertainties are reliable
Invalid if Q/Qexp>10

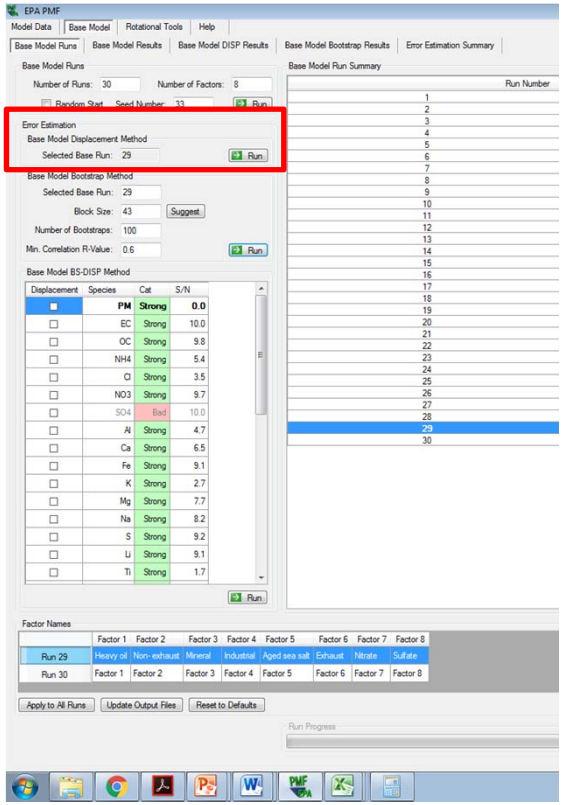


- Diagnostics:
- Error code: 0
- Second value <1%
- Number of swaps

Case 1. Small Errors	Case 2. Realistic Errors	Case 3. Simple Errors + 1 extra factor modeled
0 -0.111	0 0.000	0 -2.557
0 0 0 0	0 0 0 0	6 5 3 0 14
0 0 0 0	0 0 0 0	8 6 4 1 19
0 0 0 0	0 0 0 0	8 7 4 1 20
0 0 0 0	0 0 0 0	8 8 5 1 22

Displacement

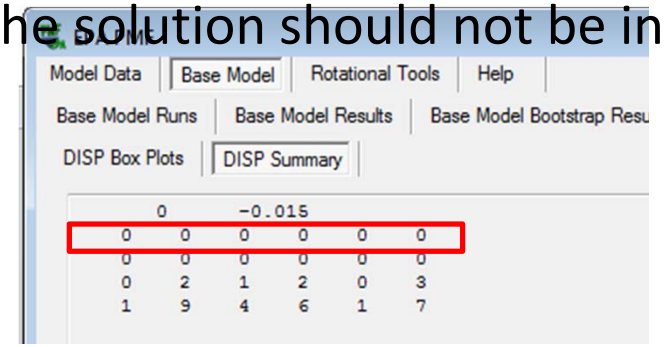
- Assesses the largest range of source profile values without an appreciable increase in the Q-value
- For example, if 20 species are in a data set and a 7-factor model has been fitted, then the DISP method will estimate 20 x 7 = 140 intervals for each dQmax value.
- The model generates results for the following dQMax values: 4, 8, 15, and 25. Plots are only shown for dQmax of 4



DISP results

- 4 txt files generated (one for each dQmax)

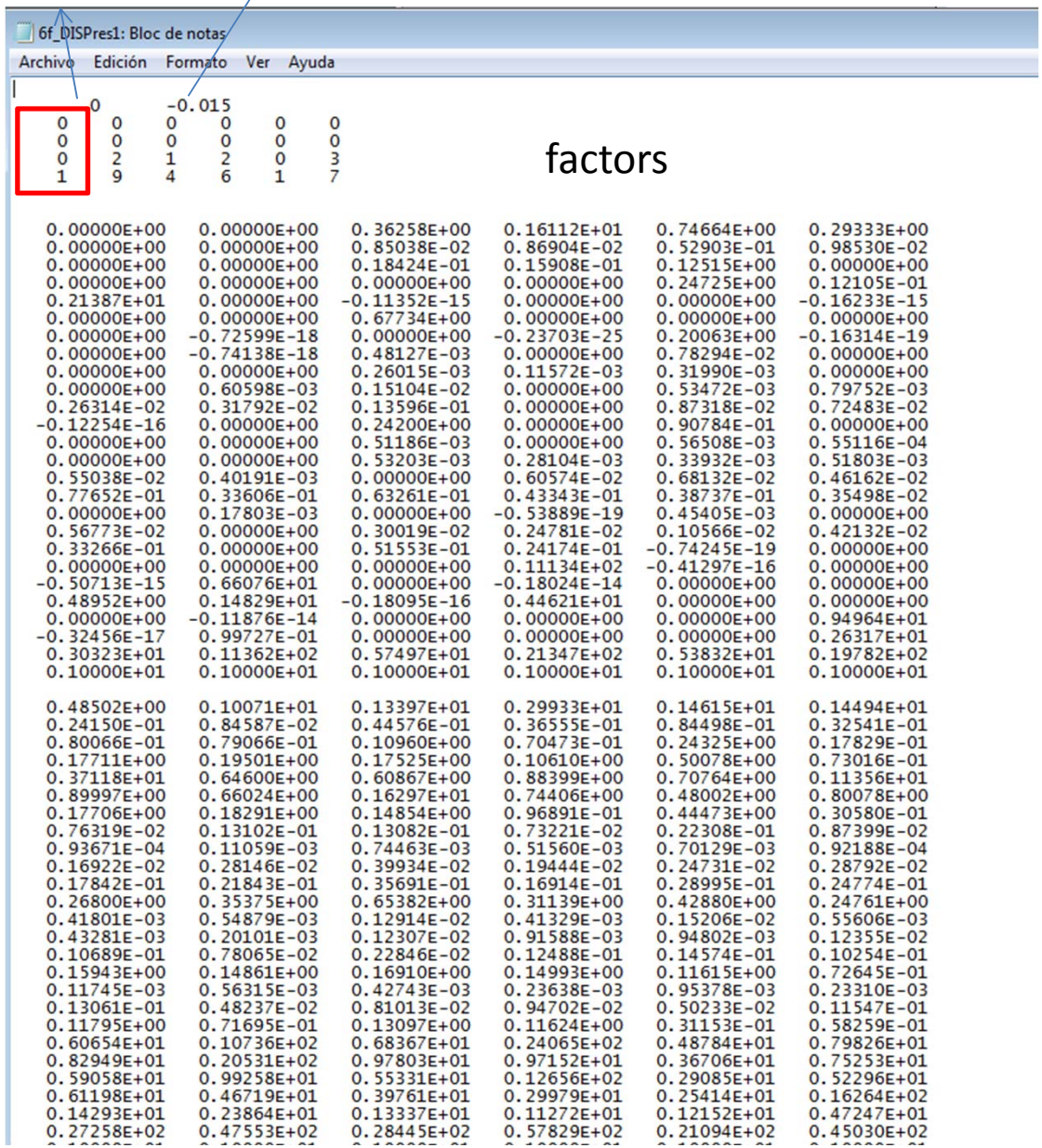
- In each file, there is a line with two numbers, followed by four lines of data. In the first line, the first value is an error code: 0 means no error; 6 or 9 indicates that the run was aborted; 6 indicates there were too many swaps, and 9 indicates Q decreased too much. If this first value is non-zero, the DISP analysis results are considered invalid. The second value is the largest observed drop of Q during DISP.
- Below the first line is a four-line table that contains swap counts for factors (columns) for each dQmax level (rows). The first row is for dQmax = 4, the second row dQmax=8, the third dQmax=15 and the fourth dQmax=25. The swap counts are a key indicator of the stability of a PMF solution and swaps at dQmax = 4 or the first row in the table indicate that the solution should not be interpreted.



DISPres1.txt

<1% (otherwise it was not a global minimum)

=0




```
                0       -0.015
    0    0    0    0    0    0
    0    0    0    0    0    0
    0    2    1    2    0    3
    1    9    4    6    1    7

 0.00000E+00  0.00000E+00  0.36258E+00  0.16112E+01  0.74664E+00  0.29333E+00
 0.00000E+00  0.00000E+00  0.85038E-02  0.86904E-02  0.52903E-01  0.98530E-02
 0.00000E+00  0.00000E+00  0.18424E-01  0.15908E-01  0.12515E+00  0.00000E+00
 0.00000E+00  0.00000E+00  0.00000E+00  0.00000E+00  0.24725E+00  0.12105E-01
 0.21387E+01  0.00000E+00 -0.11352E-15  0.00000E+00  0.00000E+00 -0.16233E-15
 0.00000E+00  0.00000E+00  0.67734E+00  0.00000E+00  0.00000E+00  0.00000E+00
 0.00000E+00 -0.72599E-18  0.00000E+00 -0.23703E-25  0.20063E+00 -0.16314E-19
 0.00000E+00 -0.74138E-18  0.48127E-03  0.00000E+00  0.78294E-02  0.00000E+00
 0.00000E+00  0.00000E+00  0.26015E-03  0.11572E-03  0.31990E-03  0.00000E+00
 0.00000E+00  0.60598E-03  0.15104E-02  0.00000E+00  0.53472E-03  0.79752E-03
 0.26314E-02  0.31792E-02  0.13596E-01  0.00000E+00  0.87318E-02  0.72483E-02
-0.12254E-16  0.00000E+00  0.24200E+00  0.00000E+00  0.90784E-01  0.00000E+00
 0.00000E+00  0.00000E+00  0.51186E-03  0.00000E+00  0.56508E-03  0.55116E-04
 0.00000E+00  0.00000E+00  0.53203E-03  0.28104E-03  0.33932E-03  0.51803E-03
 0.55038E-02  0.40191E-03  0.00000E+00  0.60574E-02  0.68132E-02  0.46162E-02
 0.77652E-01  0.33606E-01  0.63261E-01  0.43343E-01  0.38737E-01  0.35498E-02
 0.00000E+00  0.17803E-03  0.00000E+00 -0.53889E-19  0.45405E-03  0.00000E+00
 0.56773E-02  0.00000E+00  0.30019E-02  0.24781E-02  0.10566E-02  0.42132E-02
 0.33266E-01  0.00000E+00  0.51553E-01  0.24174E-01 -0.74245E-19  0.00000E+00
 0.00000E+00  0.00000E+00  0.00000E+00  0.11134E+02 -0.41297E-16  0.00000E+00
-0.50713E-15  0.66076E+01  0.00000E+00 -0.18024E-14  0.00000E+00  0.00000E+00
 0.48952E+00  0.14829E+01 -0.18095E-16  0.44621E+01  0.00000E+00  0.00000E+00
 0.00000E+00 -0.11876E-14  0.00000E+00  0.00000E+00  0.00000E+00  0.94964E+01
-0.32456E-17  0.99727E-01  0.00000E+00  0.00000E+00  0.00000E+00  0.26317E+01
 0.30323E+01  0.11362E+02  0.57497E+01  0.21347E+02  0.53832E+01  0.19782E+02
 0.10000E+01  0.10000E+01  0.10000E+01  0.10000E+01  0.10000E+01  0.10000E+01

 0.48502E+00  0.10071E+01  0.13397E+01  0.29933E+01  0.14615E+01  0.14494E+01
 0.24150E-01  0.84587E-02  0.44576E-01  0.36555E-01  0.84498E-01  0.32541E-01
 0.80066E-01  0.79066E-01  0.10960E+00  0.70473E-01  0.24325E+00  0.17829E-01
 0.17711E+00  0.19501E+00  0.17525E+00  0.10610E+00  0.50078E+00  0.73016E-01
 0.37118E+01  0.64600E+00  0.60867E+00  0.88399E+00  0.70764E+00  0.11356E+01
 0.89997E+00  0.66024E+00  0.16297E+01  0.74406E+00  0.48002E+00  0.80078E+00
 0.17706E+00  0.18291E+00  0.14854E+00  0.96891E-01  0.44473E+00  0.30580E-01
 0.76319E-02  0.13102E-01  0.13082E-01  0.73221E-02  0.22308E-01  0.87399E-02
 0.93671E-04  0.11059E-03  0.74463E-03  0.51560E-03  0.70129E-03  0.92188E-04
 0.16922E-02  0.28146E-02  0.39934E-02  0.19444E-02  0.24731E-02  0.28792E-02
 0.17842E-01  0.21843E-01  0.35691E-01  0.16914E-01  0.28995E-01  0.24774E-01
 0.26800E+00  0.35375E+00  0.65382E+00  0.31139E+00  0.42880E+00  0.24761E+00
 0.41801E-03  0.54879E-03  0.12914E-02  0.41329E-03  0.15206E-02  0.55606E-03
 0.43281E-03  0.20101E-03  0.12307E-02  0.91588E-03  0.94802E-03  0.12355E-02
 0.10689E-01  0.78065E-02  0.22846E-02  0.12488E-01  0.14574E-01  0.10254E-01
 0.15943E+00  0.14861E+00  0.16910E+00  0.14993E+00  0.11615E+00  0.72645E-01
 0.11745E-03  0.56315E-03  0.42743E-03  0.23638E-03  0.95378E-03  0.23310E-03
 0.13061E-01  0.48237E-02  0.81013E-02  0.94702E-02  0.50233E-02  0.11547E-01
 0.11795E+00  0.71695E-01  0.13097E+00  0.11624E+00  0.31153E-01  0.58259E-01
 0.60654E+01  0.10736E+02  0.68367E+01  0.24065E+02  0.48784E+01  0.79826E+01
 0.82949E+01  0.20531E+02  0.97803E+01  0.97152E+01  0.36706E+01  0.75253E+01
 0.59058E+01  0.99258E+01  0.55331E+01  0.12656E+02  0.29085E+01  0.52296E+01
 0.61198E+01  0.46719E+01  0.39761E+01  0.29979E+01  0.25414E+01  0.16264E+02
 0.14293E+01  0.23864E+01  0.13337E+01  0.11272E+01  0.12152E+01  0.47247E+01
 0.27258E+02  0.47553E+02  0.28445E+02  0.57829E+02  0.21094E+02  0.45030E+02
```

4 matrixes

(1) the profile matrix upper bound, in concentration units;

(2) the profile matrix lower bound, in concentration units;

(3) the profile matrix upper bound, in % species units;

(4) the profile matrix lower bound, in % species

DISP box plots

DISP intervals

Small DISP range for tracers

Down to zero for:

- NO3 in Sulfate
- Ions and C in Soil

BB likely contributes more (K)

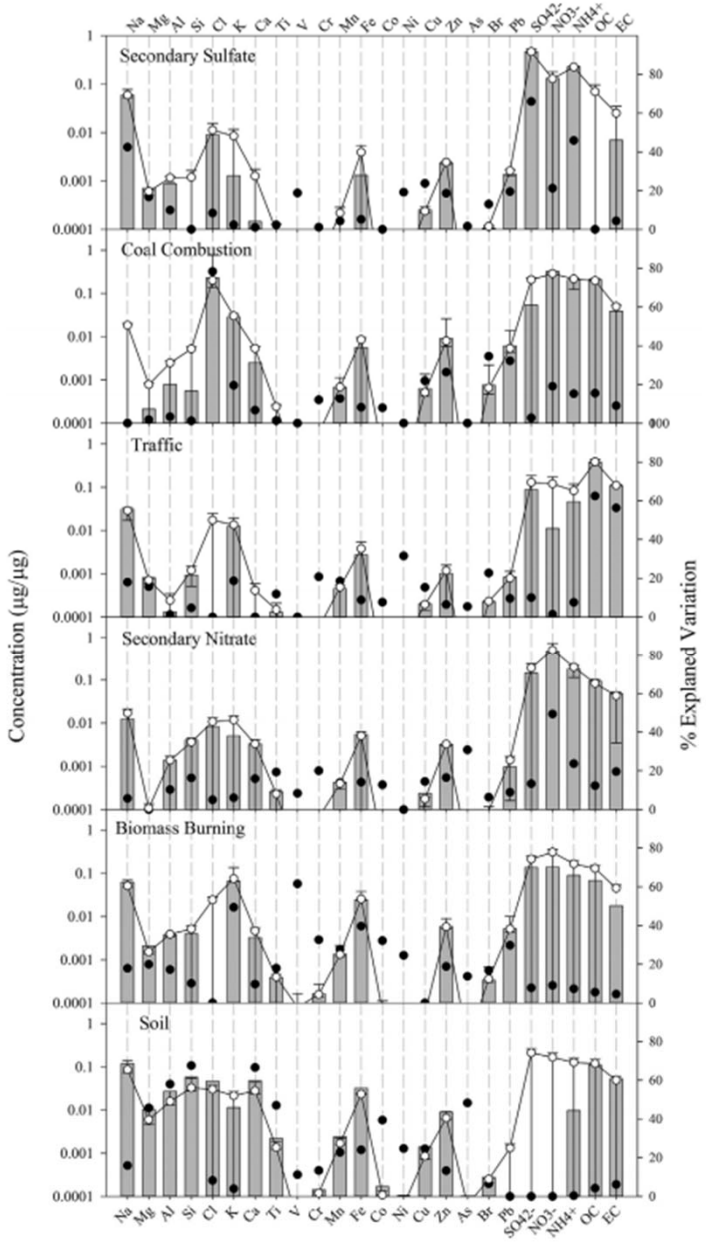
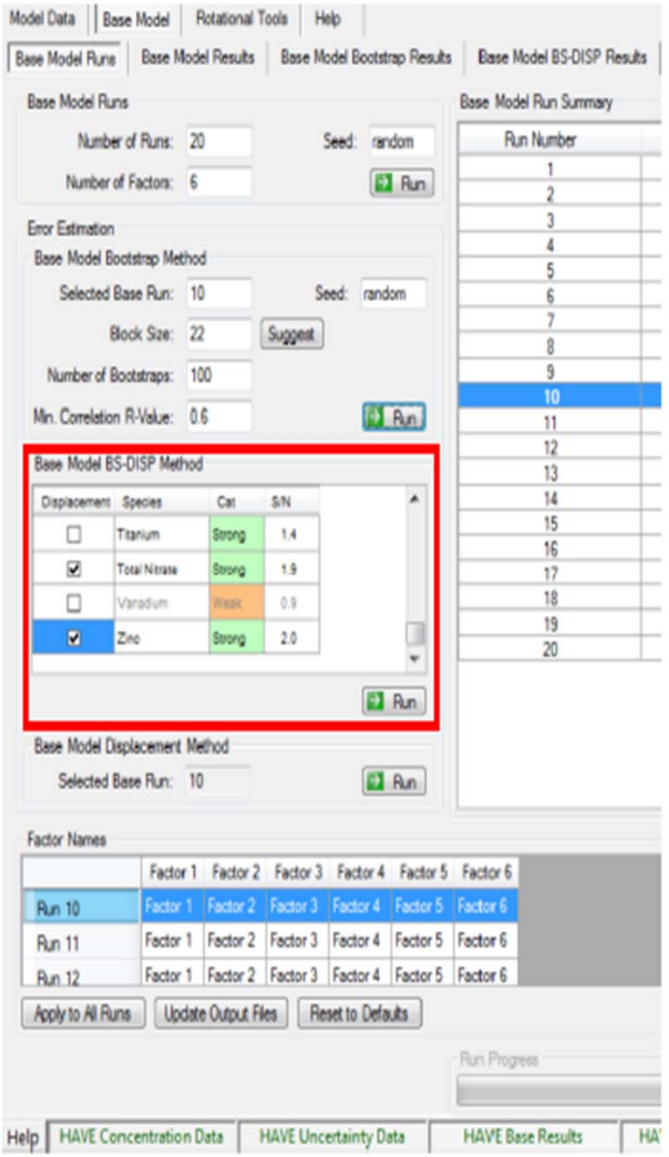


Fig. 2. The source profiles (bars represent mass contribution, open circles represent the mean DISP values with the error bars providing the confidence intervals, the lines are drawn to guide the eye, y-axis on the left; filled circles represent contribution in percentage, y-axis on the right).

BS-DISP

- Only key and major species
- Long computation times

1. Each DISP defines the span of rotationally accessible space.
2. Each BS resample moves this space around, randomly in different directions.
3. Taken together, all the replications of the rotationally accessible space, in random locations, represent both the random uncertainty and the rotational uncertainty.



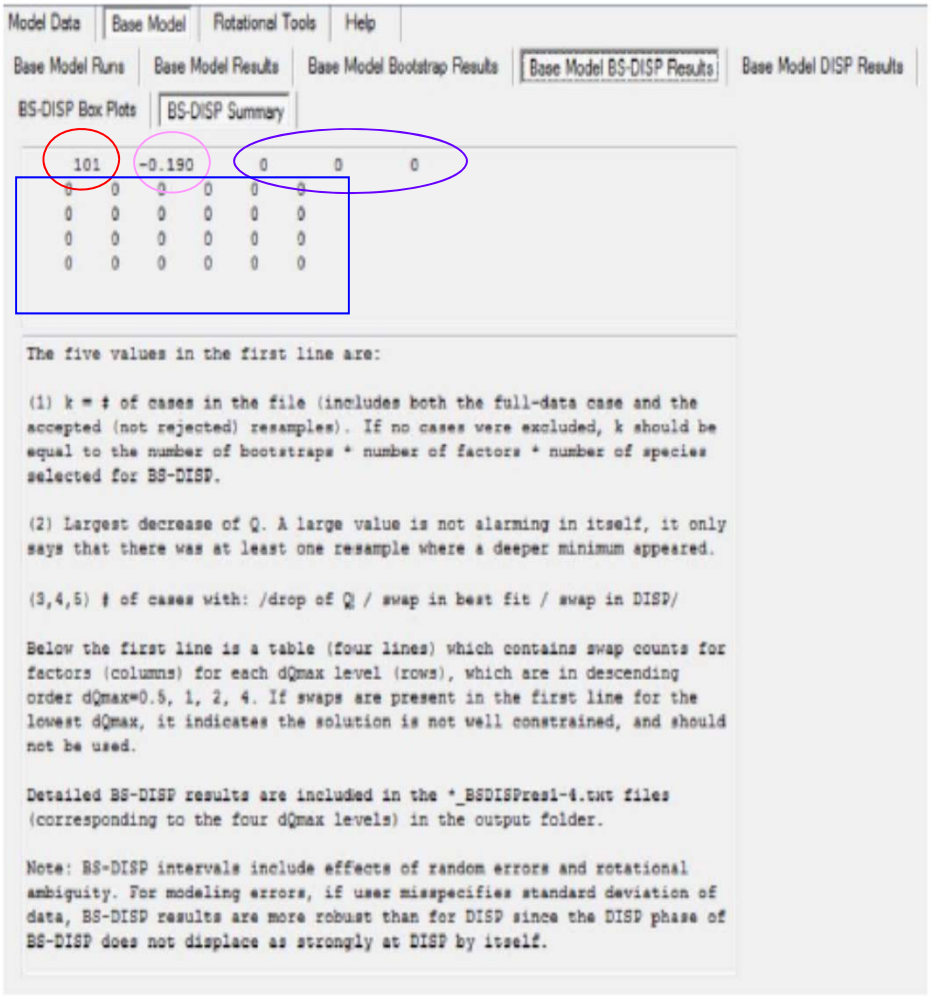
LIFE11
ENV/ES/000584

AIRUSE

BS-DISP diagnostic

Ideally

- Nº of runs+1
- <1%
- Zero values
- 0 in first row



CSIC

LIFE11
ENV/ES/000584

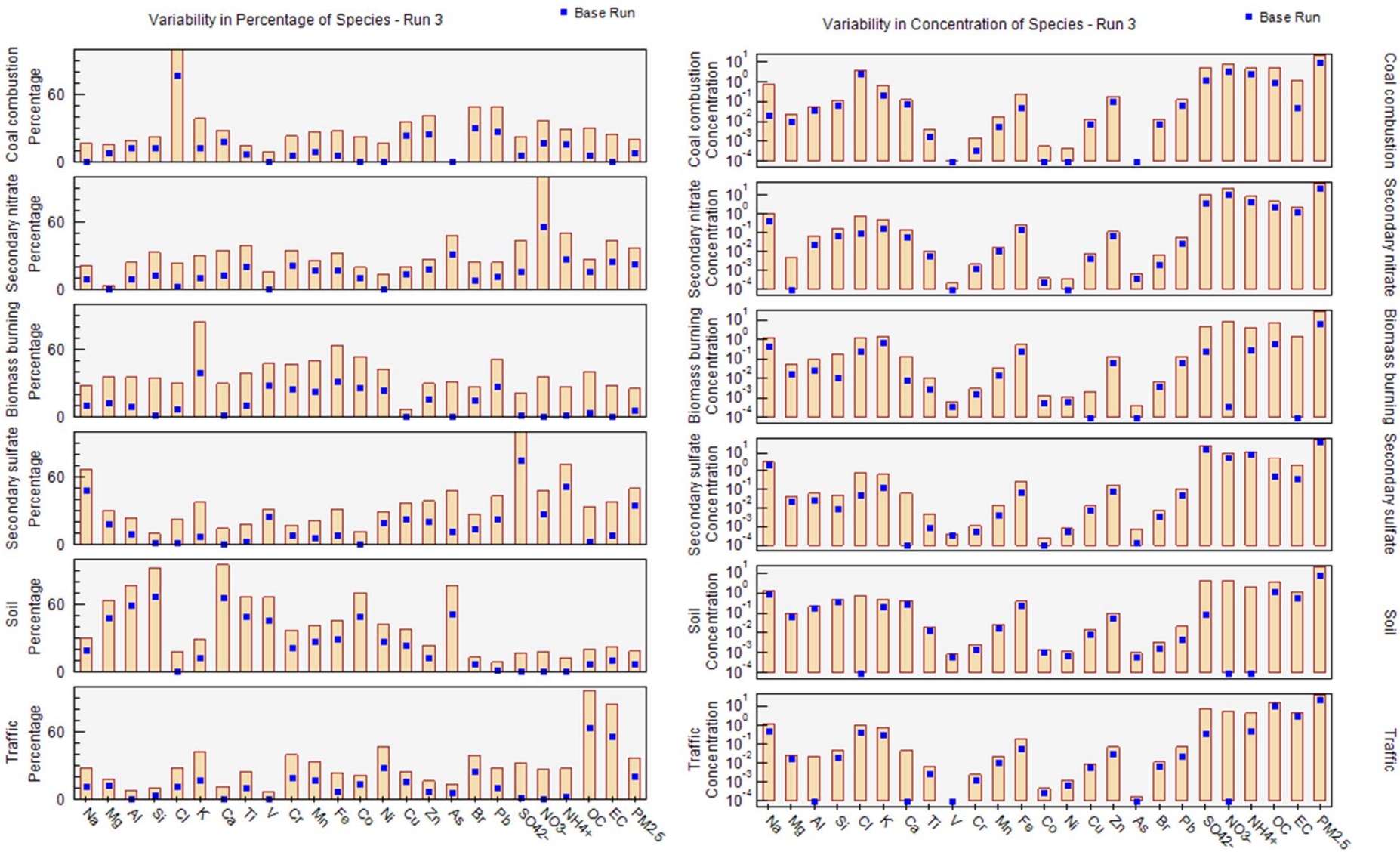
AIRUSE

BS-DISP: three examples

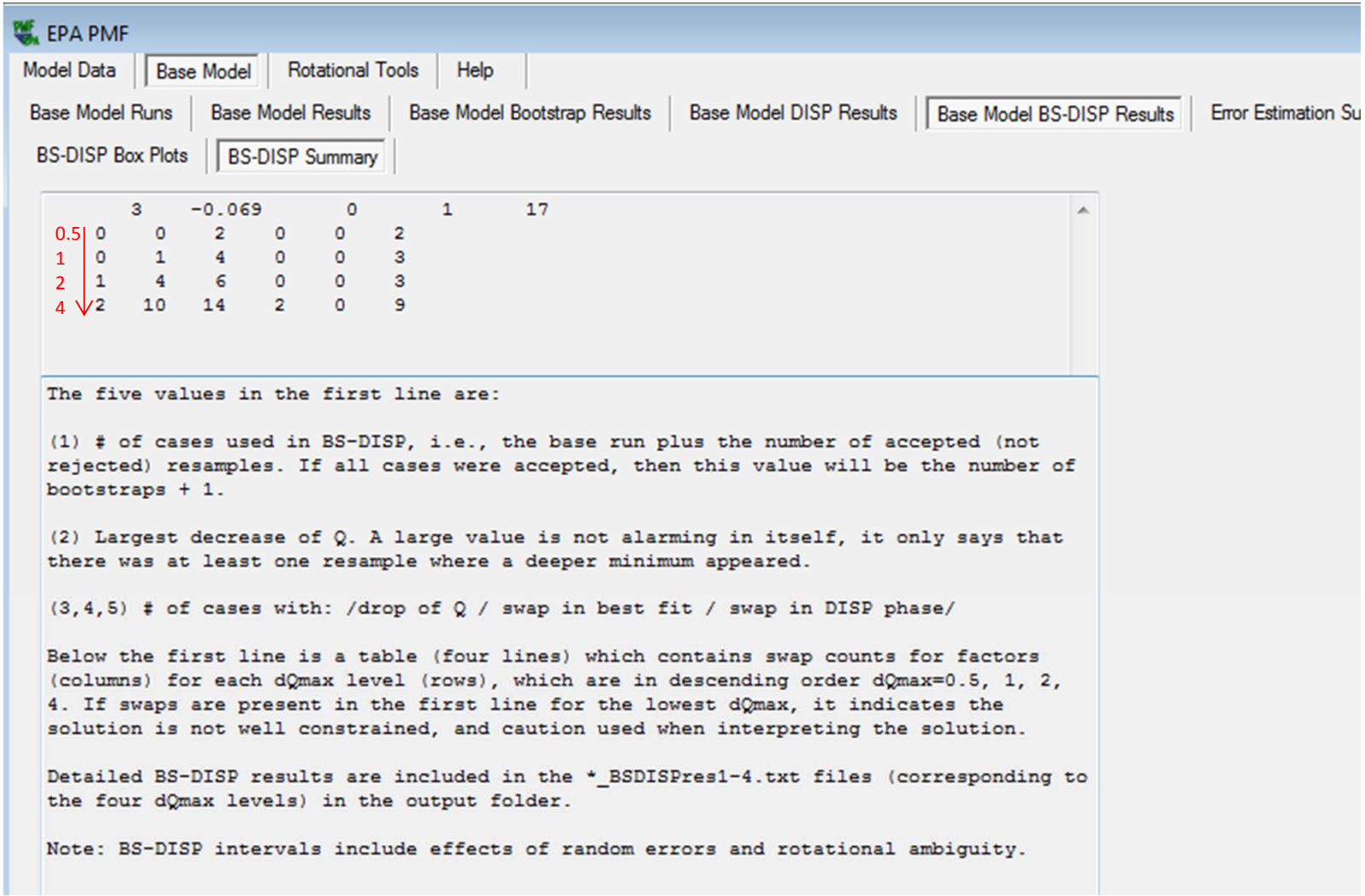
51 -0.209 0 0 0	39 -21.149 8 0 4	11 -0.461 0 0 135
0 0 0 0	0 0 0 0	7 40 27 1 54
0 0 0 0	1 0 2 1	9 52 44 1 71
0 0 0 0	1 0 2 1	18 59 53 1 93
0 0 0 0	2 1 4 1	26 67 66 1 113

CSIC

BS-DISP


Variability in Percentage of Species - Run 3
Base Run
Coal combustion Percentage
Secondary nitrate Percentage
Biomass burning Percentage
Secondary sulfate Percentage
Soil Percentage
Traffic Percentage
0
60
Na Mg Al Si Cl K Ca Ti V Cr Mn Fe Co Ni Cu Zn As Br Pb SO42- NO3- NH4+ OC EC PM2.5
Variability in Concentration of Species - Run 3
Base Run
Coal combustion Concentration
Secondary nitrate Concentration
Biomass burning Concentration
Secondary sulfate Concentration
Soil Concentration
Traffic Concentration
Coal combustion
Secondary nitrate
Biomass burning
Secondary sulfate
Soil
Traffic
Na Mg Al Si Cl K Ca Ti V Cr Mn Fe Co Ni Cu Zn As Br Pb SO42- NO3- NH4+ OC EC PM2.5

BS-DISP summary



Error Estimation summary