TROPOMI NO_2: a new method to identify unresolved features in DOAS residuals

Jos van Geffen[1]*, Henk Eskes[1], Maarten Sneep[1], Pepijn Veefkind[1,2]

with thanks to Andreas Richter (University of Bremen)

The problem

In NO_2 slant column (SCD) retrieval DOAS minimises χ^2 by adjusting fit parameters for known reference spectra: NO_2, O_3, O_2-O_2, H_2O_{vap}, H_2O_{liq}, Ring, plus polynomial (cf. ATBD [3]). This works well in most cases worldwide, but not in cases where other absorption or scattering effects occur. In those cases DOAS uses avialable fit parameters to compensate for missing reference spectra, thus possibly giving incorrect NO_2 SCDs, which leads to incorrect vertical columns (VCDs).

The main case — Lake Siling, Tibet [4,5]

Two ground pixels over 'land' and 'water' (*Fig. A*):

- some 75 km apart $\Longrightarrow$ $\mathrm{VCD}^{\mathrm{land}}_{\mathrm{strat}} \approx \mathrm{VCD}^{\mathrm{water}}_{\mathrm{strat}}$
- if no NO_2 sources $\Longrightarrow$ $\mathrm{VCD}^{\mathrm{land}}_{\mathrm{trop}} \approx \mathrm{VCD}^{\mathrm{water}}_{\mathrm{trop}}$
- albedo (*Fig. B*): $\mathrm{A}^{\mathrm{land}}_{\mathrm{surf}} > \mathrm{A}^{\mathrm{water}}_{\mathrm{surf}} \Longrightarrow \mathrm{GCD}^{\mathrm{land}}_{\mathrm{total}} > \mathrm{GCD}^{\mathrm{water}}_{\mathrm{total}}$

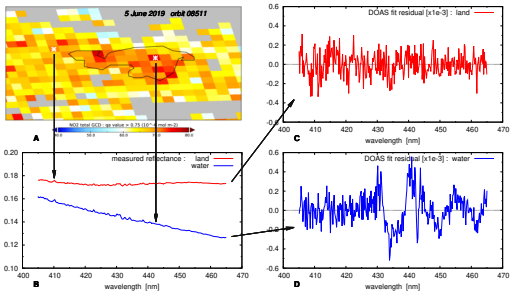
[GCD = SCD/AMFgeo = geometric column]

Observations (*Fig. A*) show: $\mathrm{GCD}^{\mathrm{land}}_{\mathrm{total}} < \mathrm{GCD}^{\mathrm{water}}_{\mathrm{total}}$
$\Longrightarrow$ either NO_2 source in the lake
or incorrect $\mathrm{GCD}^{\mathrm{water}}_{\mathrm{total}}$, i.e. incorrect SCD

Fit residual = measured − modelled reflectance:

- 'land' (*Fig. C*) looks OK $\Longrightarrow$ SCD OK
- 'water' (*Fig. D*) shows strong low-frequency structures $\Longrightarrow$ reference spectrum missing
  $\Longrightarrow$ NO_2 SCD possibly incorrect !

Same issue occurs over mulitple lakes in and around Tibet, but not e.g. over lakes in Mongolia



The test

Can we detect low-frequency structures in fit residuals ?

Wald-Wolfowitz runs test [6,7] or **runs test** for short:

- checks a randomness hypothesis based on the number of positive and negative fit residual values, where a 'run' is a sequence of same-signed values
- determines the deviation between actual and expected number of runs in terms of standard deviation σ

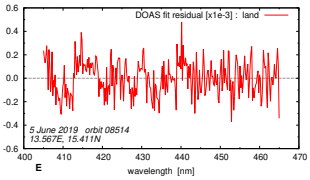
We added [3]:

- sign to the deviation R_D:
  $R_D < 0$ for fewer | $R_D > 0$ for more actual than expected runs
- length of the longest run R_L

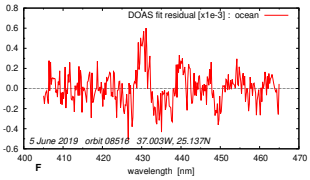
	Lake Siling land	Lake Siling water	Saraha desert	Atlantic ocean	unit
figure of fit residual	C	D	E	F	
NO_2 SCD	170.19	187.41	128.97	146.12	μmol/m^2
NO_2 SCD error	8.08	13.07	6.93	9.52	μmol/m^2
NO_2 GCD	70.55	75.67	58.01	60.65	μmol/m^2
NO_2 GCD error	3.35	5.28	3.12	3.95	μmol/m^2
RMS of the fit	1.21	1.63	1.45	1.66	$\times 10^{-4}$
χ^2 of the fit	3.60	6.55	4.06	5.32	$\times 10^{+2}$
no. of spectral points	305	305	305	305	
no. of positive values	155	152	152	144	
no. of negative values	150	153	153	161	
no. of observed runs	138	110	100	98	
no. of expected runs	153.46	153.50	153.50	153.03	
sigma of that	8.72	8.72	8.72	8.69	
R_D = deviation from no. of expected runs	−1.77	−4.99	−6.14	−6.33	sigma
R_L = length of longest observed run	10	35	33	33	

Two other cases

Fit results of several pixels over land (*Fig. E: over Sahara; similar seen over land elsewhere*): clear broad but not high peak at 415-420 nm; no indication that the NO_2 SCD is incorrect.



Fit results of many pixels over ocean (*Fig. F: Atlantic Ocean*): clear high peak around 430 nm (due to VRS in water?) as well as some indication of small structures similar to *Fig. D* $\Longrightarrow$ NO_2 SCD possibly incorrect – *see below for some remarks.*



Recipe for the NO_2 data user

if persistent unexpected tropospheric column values
then check the GCD, i.e. the actual measurements

if unexpected GCD (& maybe other fit parameter) values
then check results of the runs test

if R_D and/or R_L show large values
then DOAS fit residual has remaining structures
$\Longrightarrow$ the NO_2 SCD may be incorrect
$\Longrightarrow$ contact data product leads to look at fit residual

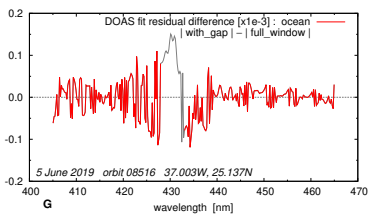
Note: large R_D and/or R_L values do not necessarily mean that the NO_2 SCD is incorrect

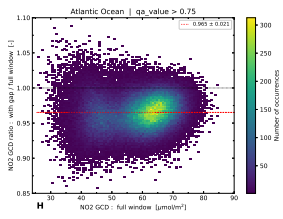
Concluding remarks

- DOAS uses reference spectra to minimise χ^2 of fit residuals
- Missing reference spectra lead to structures in the fit residuals
  $\Longrightarrow$ retrieved NO_2 SCD may be incorrect
- The runs test provides additional independent information on the quality of DOAS fit residuals
- As of v2.7.1 (Sept. 2024) the NO_2 data product contains:
  `runs_deviation` $= R_D$: deviation in terms of σ
  `runs_longest` $= R_L$: length of longest run
- Negative (positive) R_D values indicate low (high) frequency structures left in the fit residual
- We cannot use runs test results to adjust the `qa_value`

Bias in NO_2 over oceans

Tropospheric NO_2 VCDs over oceans are small but are known to have a high bias in some areas. The fit residual of such ground pixels (*Fig. F*) shows a clear high peak around 430 nm and the NO_2 SCDs may thus be incorrect.
Performing a DOAS fit omitting spectral pixels between 428 and 433 nm (*Fig. G: difference between absolute fit residuals*) reduces χ^2, RMS error and SCD error, and leads to a 5.6% lower SCD for this specific pixel.
For cloud-free pixels over the Atlantic ocean (*Fig. H: ratio for scanlines with nadir latitude [10°:40°], 5 June 2019, orbit 08516*) the GCD is reduced by about 4% on average, which in turn leads to lower VCDtrop values. Does this approach solve the bias, without causing problems elsewhere? To be investigated further.





[1] KNMI – Royal Netherlands Meteorological Institute, De Bilt, NL
[2] TU/Delft – Delft University of Technology, Delft, NL
[3] van Geffen et al., TROPOMI NO_2 ATBD v2.7.0, 2024
[4] Kong et al., 2023: *Nat. Geosci.* 16, 474-477, doi: 10.1038/s41561-023-01200-8
[5] Labzovskii et al., 2024: *Nat. Geosci.*, in press = "arising matter" response to ref. 4
[6] Barlow, 1989: "Statistics: a guide to the use of statistical methods in the physical sciences," Sect. 8.3.2
[7] https://en.wikipedia.org/wiki/Wald-Wolfowitz_runs_test

*** Contact:** jos.van.geffen@knmi.nl