GlobEmission

NOx Emission Downscaling for South Africa

Version	Author	Date	Comments
1.0	Bino Maiheu, Nele Veldeman (VITO)	12/12/2013	First experimental version of algorithm & methodology

Algorithm Description

Input data

- ***Regional emission inventory for South Africa, provided by KNMI:*** KNMI will provide emission estimates of NOx on a 0.25 degree resolution, derived from satellite observations and inverse modelling, for South Africa.

- ***Sector split provided by SAWS/DEA:*** The regional emission inventory for South Africa, will be provided by KNMI as total emissions, i.e. without any breakdown over different economical sectors. As further downscaling is highly dependent on the breakdown over different sectors, information on the sectoral split is required. In the presentation 'National Emission Inventories Developments in South Africa: Current Status and Needs', presented by Patience Gwaze (SAWS) at the GlobEmission User Workshop on 26-27 Nov 2009 in Frascati, a sectoral split was presented, see **Error! Reference source not found.**.

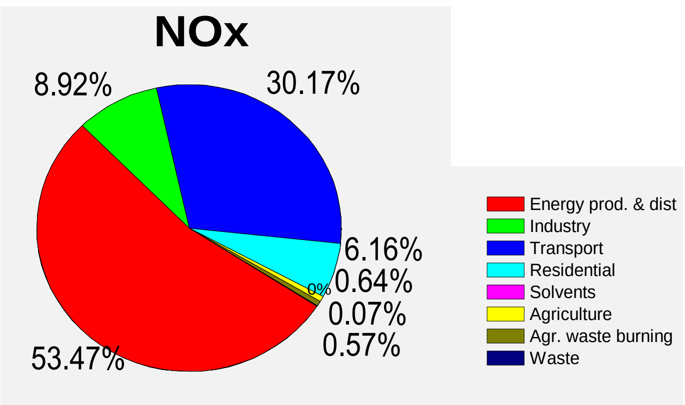


Figure 1 : Sector split information for anthropogenic emissions. Source: Lamarque J.-F., et al., (2009) Gridded emissions in support of IPCC AR5. IGACtivities, (41), 12-18.

The sector information contained in the above pie chart, only reflects anthropogenic emissions. A large unknown is the ratio between anthropogenic and biogenic emissions.

Ideally additional information on the ratio between anthropogenic and biogenic emissions will be provided.

- ***Proxy data provided by SAW/DEA:*** According to the national emission split over economical sectors for South Africa, 99% of anthropogenic NOx emissions in South Africa originate from the following four sectors: Energy production and distribution, Road Transport, Industry, and Residential Sources. Consequently, good proxy data for each of these four sectors are required: locations and emissions of large point sources, road map and traffic volumes, land use map, population density map, ... Apart from these, also good proxies to spatially allocate biogenic emissions are required.

Algorithm

The proposed methodology for downscaling the regional inventory to a 0.05x0.05 degrees resolution consists of the following steps:

1. The gridded regional (0.25 degree resolution) emission estimates from KNMI will be added up to get ***a national emission total for South Africa***, based on satellite observations;

2. ***Spatial patterns*** will be generated both, on a 0.25x0.25 degrees grid (step 2a) and on a 0.05x0.05 degrees grid (step 2b) using the available geographical proxy data, i.e. roads, population density, land use data, … in order to
   a. “reconstruct” ***the low resolution*** 0.25x0.25° KNMI emissions map (see step 3),
   b. allow further downscaling ***to high resolution*** (0.05x0.05°, see step 4);

3. ***Reconstruction of the low resolution emissions map*** from
   a. the national emissions total for South Africa (step 1) and
   b. the spatial patterns on the original 0.25x0.25 degrees grid (step 2a),

   making use of an ‘optimization procedure’. An optimized linear combination of the different spatial patterns will be determined. Hereto, validation indicators between the original map derived from inverse modelling and the reconstructed map are used as cost function for the optimization (minimization):

$$\zeta(\mathbf{a}) = 1 - r_{xy(\mathbf{a})}^2 = 1 - \frac{\left[\sum_{i=1}^{n}(x_i - \bar{x})(y_i(\mathbf{a}) - \overline{y(\mathbf{a})})\right]^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i(\mathbf{a}) - \overline{y(\mathbf{a})})}$$

   where $r_{xy(a)}$ is the spatial correlation coefficient between the original grid values x_i (as derived by the inverse modelling) and the reconstructed grid values $y_i(\mathbf{a})$ (as calculated using **a** as a weight vector between de different proxy parameters).

   Via maximization of the spatial correlation, the methodology tries to explain as much as possible of the spatial variance of the original KNMI emission dataset through a linear combination of the proxy-patterns.

Key to the methodology is therefore:

i. determination of relative importance of each sector
ii. at national level
iii. through optimization of weights

4. ***Validation and optimization of the minimization procedure***. Different minimization techniques, i.e. based on three constraining modes, will be applied and compared against each other. The three constraining modes that will be taken into account can be summarized as follows:
   a. Free: where the only constraint that applies is that the sum of the weights should equal 1 and the weights should all be non-negative.
   b. Hard constraint: where, next to being non-negative, 4 additional constraints guaranteeing that the explicit sector split known for anthropogenic emissions is maintained, are built in. These constraints are:
      i. The traffic proxiy weight should equal 30.17 % of (1 – the sum of the biogenic emission weights)
      ii. The residential proxy weight should equal 6.16% of (1 – the sum of the biogenic emission weights)
      iii. The industry proxy weight should equal 8.92 % of (1 – the sum of the biogenic emission weights)
      iv. The powerplant weight should equal 5..47 + 1.28 % of (1 – the sum of the biogenic emission weights). The remaining 1.28 % of the above pie-chart was added in order to make the different anthropogenic proxies add up to a total of 100 % for anthropogenic emissions.
   c. Soft constraint: where the same methodology as in the hard constraints mode is adapted, however, allowing the different contributions of the anthropogenic emissions to vary within +/- 5 % of the split given above.

5. ***Computation of high resolution emission maps from***
   a. the national emissions total for South Africa (step 1)
   b. the spatial patterns on the high resolution 0.05x0.05 degrees grid (step 2b)
   c. the optimized weights, i.e. vector a in the above formula) (step 3)

6. Note to the algorithm: national totals (as derived from original KNMI data) are conserved in the algorithm. However, locally there might be differences due to the nature of the approach. Hence, ***a local residue correction scheme*** might be applied as a last step.

Analysis

Derivation of proxy data & methodological aspects

The downscaling methodology tries to stay true to the sectorsplit for anthropogenic emissions we found in literature. This split is depicted below in Figure 2.

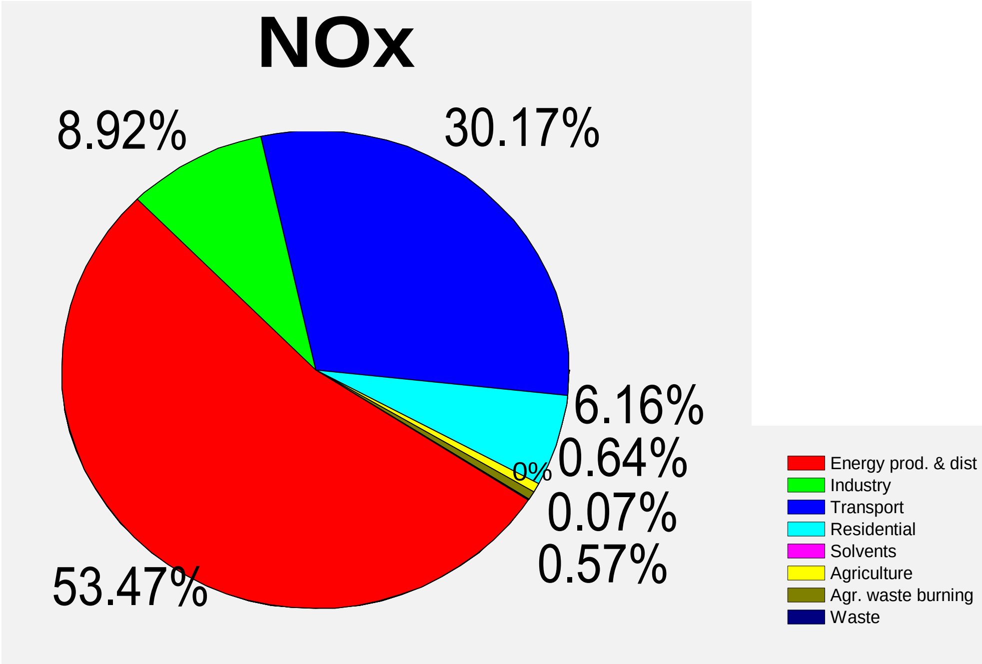


Figure 2 : Sector split information for antrhopogenic emissions. Source : Source: Lamarque J.-F., et al., (2009) Gridded emissions in support of IPCC AR5. IGACtivities, (41), 12-18.

A large unknown off course is the ration between anthropogenic and biogenic emissions. For this split, various information sources have been identified, however, not all of them are directly usable as they refer more to "Southern Africa" instead of south Africa as such, or refer to Africa as a whole. Given the significant heterogeneity of the land use, it is unclear to what extent numbers are directly transferrable to the South African domain. Suffice it to say that we do expect a quite sizeable fraction of the NOx emissions to have biogenic sources (where we count biomass burning also as a biogenic source). From the table below (REF), we can infer that the biogenic part (biogenic activity + biomass burning) is roughly 2.7-2.8 times larger than anthropogenic NO_x emissions for southern Africa.

Table 2. Emission totals for the contributions to NO_x, CO, isoprene and organic carbon from (in listed order) anthropogenic activity, biomass burning and biogenic activity for each latitudinal zone. For the biogenic component integrated totals are shown for both the BASE and LATH simulations, respectively, along with resulting percentage difference for each of the regions. For biogenic CO the same inventory is used in both runs. For the NOSOIL simulations the emissions denoted with (#) are removed and for the NOBIO simulation the emissions denoted with (*) are removed. Summing these segregated emission totals gives 6.91 (6.91) $TgN\,yr^{-1}$, 250.15 $TgCO\,yr^{-1}$, 112.07 (137.62) $TgC\,yr^{-1}$ (isoprene) and 20.34 (11.27) $TgC\,yr^{-1}$ (BVOCs), where the BASE sums are given in parenthesis except for CO.

Trace species	Emission Source	Sahara (20–40° N)	Sahel (10—20° N)	Guinea (0–10° N)	Southern Africa (40° S–0° N)
NO_x (TgN/yr)	Anth	1.12	1.13×10^{-1}	2.30×10^{-1}	6.96×10^{-1}
	BB	2.81×10^{-3}	2.40×10^{-1}	1.01	1.19
	Bio (BASE)	2.18×10^{-1}	6.12×10^{-1}	7.74×10^{-1}	7.02×10^{-1}
	Bio (LATH)	$2.01\times10^{-1(\#)}$	$5.75\times10^{-1(\#)}$	$7.64\times10^{-1(\#)}$	$7.64\times10^{-1(\#)}$
	%diff Bio	(7.5%)	(6.3%)	(1.3%)	(+8.0%)
CO (TgCO/yr)	Anth	13.83	14.90	23.10	18.46
	BB	1.64×10^{-1}	13.77	65.37	70.43
	Bio (ALL)	3.09	4.43	10.59	12.02
	%diff Bio	(N/A)	(N/A)	(N/A)	(N/A)
Isoprene (TgC/yr)	Bio (BASE)	3.46	16.10	51.45	41.06
	Bio (LATH)	3.58	20.54	56.99	56.51
	%diff Bio	(3.4%)	(27.6%)	(9.7%)	(27.3%)
Organic Carbon (TgC/yr)	Anth	2.48	1.07	1.97	1.40
	BB	4.24×10^{-3}	3.57×10^{-1}	1.87	7.83×10^{-1}
	Bio (BASE)	1.47×10^{-1}	1.46×10^{-1}	3.59×10^{-1}	6.86×10^{-1}
	Bio (LATH)	$5.05\times10^{-1(*)}$	$1.25^{(*)}$	$4.84^{(*)}$	$3.81^{(*)}$
	%diff Bio	(+343%)	(+858%)	(+1348%)	(+556%)

In order to do the downscaling, we have selected a number of proxy parameters, to which we will assign individual weights :

- For anthropogenic emissions we make the explicit split as in the above pie-chart. Note that we also split up the traffic emissions in a highway, primary, secondary and tertiary road part. As we do not have local information w.r.t. traffic intensities, we will simply derive the split in the optimization (see further)
- For biogenic/biomass burning we will use the entire GlobCover dataset (except for those classes that are non-natural and that a priori do not have NOx emissions). I.e. we do not use **190** (Artificial surfaces and associated areas (Urban areas >50%)), **210** (Water bodies) and **220** (Permanent snow and ice)
-

Nr	Origin	Sector	Subsector	Proxy	Proxy_ID
1	Anthropogenic	Transport	Highways	transport_highway	trans_high
2			Primary	transport_primary	trans_prim
3			Secondary	transport_secondary	trans_sec
4			Tertiary	transport_tertiary	trans_ter
5					
6		Residential		residential	residential
7		Industrial		industrial	industrial
8		Powerplants			pplants

10 - 21	Biogenic/Biomass burning	GlobCOVER V2.3

This amounts to a total of 21 proxy-maps. Each of the proxy data was normalized to a national total (total surface or total length of line-segments for line sources (traffic-proxies)), where each gridcell contains the fraction of the national total for that particular proxy. The different sector totals can therefore be mapped spatially by simple multiplication. Key to the methodology is that we now derive the weights of the sector (or split over the different proxy parameters) at a national level, each weight will therefore emphase the matching spatial proxy in the reconstruction of the total low resolution map. Validation indicators between the original map derived from inverse modelling (KNMI) and the reconstructed map are used as cost function for the optimization (minimization).

After careful analysis, it appears that defining a cost function based upon the spatial correlation yields the most satisfactory results :

$$\zeta(\mathbf{a}) = 1 - r^2_{xy(\mathbf{a})} = 1 - \frac{\left[\sum_{i=1}^{n}(x_i - \bar{x})(y_i(\mathbf{a}) - \overline{y(\mathbf{a})})\right]^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i(\mathbf{a}) - \overline{y(\mathbf{a})})}$$

Where $r_{xy(\mathbf{a})}$ is the spatial correlation coefficient between the original gridvalues x_i, as derived by the inverse modelling and the reconstructed grid values ($y_i(\mathbf{a})$), as calculated using a as a weight vector between de different proxy parameters. Via maximization of the spatial correlation, the methodology tries to explain as much as possible of the spatial variance of the original KNMI emission dataset through a linear combination of the proxy-patterns.

Clearly a number of constraints have to be built in here in order to stay true to the original sector split in the anthropogenic emissions. We have defined 3 constraining modes :

- **Free** : where the only constraint that applies is that the sum of the weights should equal 1 and the weights should all be non-negative.
- **Hard constraint** : where, next to being non-negative, 4 additional constraints are built in, requiring that :
  - o The sum of the 4 traffic proxies should equal 30.17 % of 1 – the sum of the biogenic emission weights
  - o The residential proxy weight should equal 6.16% of 1 – the sum of the biogenic emission weights
  - o The industry proxy weight should equal 8.92 % of 1 – the sum of the biogenic emission weights
  - o The powerplant weight should equal 5..47 + 1.28 % of 1 – the sum of the biogenic emission weights. Here we have added the remaining 1.28 % of the above pie-chart in order to make the different anthropogenic proxies add up to a total of 100 % for antropogenic emissions as such.
- **Soft constraint** : were we adopted the exact same methodology as in the hard constraints, however, here we allow the different contributions of the anthropogenic emissions to vary within +/- 5 % of the split given above. This means that we e.g. require the sum of the 4

traffic proxies to be withing 25.17 and 35.17 % of the total anthropogenic part (being: 1 – the sum of the biogenic emission weights).

These optimization schemes were implemented in MATLAB, via the *fmincon* routine using the interior-point algorithm (Byrd et al, 1999; Byrd et al, 2000; Waltz et al, 2006).

Optimization of weights

Robustness of weight determination

In a first step, we assessed the robustness of the optimization with respect to the initial values for the split between the proxy parameters. Clearly, we want to investigate whether we are in a global or local minimum in the 21-dimensional parameter space. We have performed a rough assessment via initial weights (which add up to 1) drawn from a random distribution or simply using all the same weights. Example results are given in the table below, where we used the emission total for the year 2011, provided by the KNMI as original emission map together with the hard constraints. We show an example starting from initial uniform weight, or 3 different random sets.

	initial uniform	***optim***	**initial random 1**	***optim***	**initial random 2**	***optim***	**initial random 3**	***optim***
trans_high :	0.0476	*0.0277*	0.0404	*0.0277*	0.0302	*0.0277*	0.0306	*0.0278*
trans_prim :	0.0476	*0.0450*	0.0628	*0.0450*	0.0173	*0.0450*	0.0512	*0.0451*
trans_sec :	0.0476	*0.0164*	0.0328	*0.0164*	0.0632	*0.0164*	0.0086	*0.0164*
trans_ter :	0.0476	*0.0007*	0.0672	*0.0007*	0.0759	*0.0007*	0.0107	*0.0004*
residential :	0.0476	*0.0183*	0.036	*0.0183*	0.0262	*0.0183*	0.0131	*0.0183*
industrial :	0.0476	*0.0265*	0.0624	*0.0265*	0.0537	*0.0265*	0.0651	*0.0265*
pplants :	0.0476	*0.1628*	0.0643	*0.1628*	0.0351	*0.1628*	0.0475	*0.1627*
gc14 :	0.0476	*0.0000*	0.0404	*0.0000*	0.0667	*0.0000*	0.0182	*0.0000*
gc20 :	0.0476	*0.0024*	0.0018	*0.0024*	0.0615	*0.0024*	0.0475	*0.0024*
gc30 :	0.0476	*0.3622*	0.0302	*0.3622*	0.0134	*0.3622*	0.0142	*0.3623*
gc40 :	0.0476	*0.0000*	0.0387	*0.0000*	0.0689	*0.0000*	0.0053	*0.0000*
gc50 :	0.0476	*0.0106*	0.0247	*0.0106*	0.0792	*0.0106*	0.0816	*0.0106*
gc60 :	0.0476	*0.0000*	0.018	*0.0000*	0.0411	*0.0000*	0.0538	*0.0000*
gc90 :	0.0476	*0.0000*	0.075	*0.0000*	0.0707	*0.0000*	0.0892	*0.0000*
gc100 :	0.0476	*0.0000*	0.0393	*0.0000*	0.047	*0.0000*	0.0669	*0.0000*
gc110 :	0.0476	*0.0250*	0.0811	*0.0250*	0.0124	*0.0250*	0.0559	*0.0251*
gc120 :	0.0476	*0.0001*	0.0357	*0.0001*	0.016	*0.0001*	0.0782	*0.0000*
gc130 :	0.0476	*0.0881*	0.0702	*0.0881*	0.0325	*0.0881*	0.0843	*0.0881*
gc140 :	0.0476	*0.1889*	0.0362	*0.1889*	0.0599	*0.1889*	0.0949	*0.1890*
gc150 :	0.0476	*0.0174*	0.0738	*0.0174*	0.066	*0.0174*	0.0001	*0.0174*

gc200 :	0.0476	*0.0079*	0.0689	*0.0079*	0.0632	*0.0079*	0.083	*0.0079*
SUM OF WEIGHTS	1.0000	**1.0000**	1.0000	**1.0000**	1.0000	**1.0000**	1.0000	**1.0000**
RMSE	0.7888	**0.5413**	0.7272	**0.5413**	0.8711	**0.5413**	0.8381	**0.5412**
SPATIAL CORR	0.2102	**0.6367**	0.3235	**0.6367**	0.096	**0.6367**	0.1456	**0.6367**

Each time, the same final weights are retrieved with an optimized spatial correlation coefficient. The same exercise was repeated with soft constraints (allowing up to +/- 10 % deviation of the original split between anthropogenic sector) with identical results; albeit yielding different weights off course.

We clearly see the significant increase in R2 and decrease in RMSE in the optimized weights.

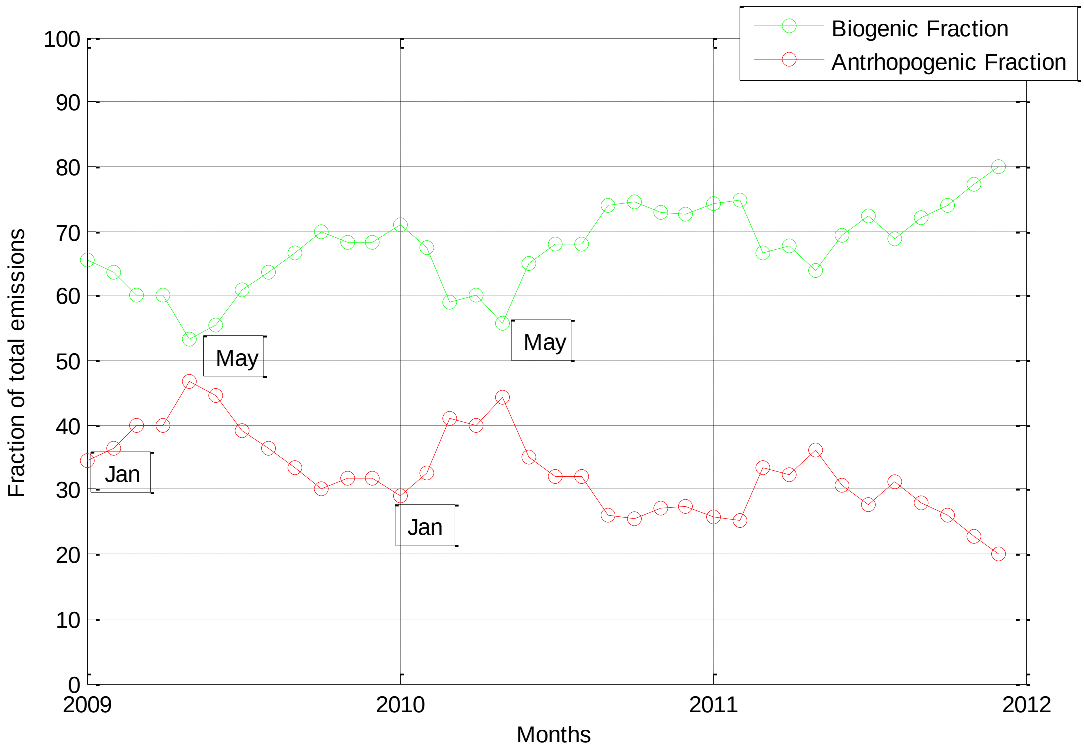
Optimization analysis

The table on the next page shows the results of the weights for each proxy from a free, hard & soft constrained optimization. In the hard optimization, the original split for anthropogenic sources is indicated in red. A number of observations :

- The picture where biogenic/biomass burning takes about 65 % of the total emissions and the antropogenic about 35 % is fairly constant over all years of analysis/optimization modes.
- The spatial correlation (R^2) does not depend hugely on the optimization mode or the application of constrants, more so, it appears the hard constrained method even yields overall better results, which is a little surprising and probably would deserve some further attention. The slight difference may be related to the random selection of initial weights, which do not necessarily conform to the constraints imposed (both hard/soft). Never the less, the differences in R^2 between the hard and soft constrained versions are really minimal, so further investigation was not performed here.
- The free optimization completely destroys the original sector split in the anthropogenic emissions and if probably therefore not really an option as we wanted to stay as true as possible to the original information.
    - The soft constrained method seems to want to set the traffic emissions to 0 (consistently yielding the lower boundary of the constraint interval), whereas it seems to want to increase drastically the residential emissions (consistently yielding the upper interval boundary). Industrial & powerplants are better preserved in the split.
- Overall emission of certain sectors does seem to be fairly constant over the different years / optimization modes , e.g. : powerplants remains at ~20 % of total, GlobCOVER 200, 150, 140, 130.
- Certain classes are put to 0 : tertiary roads, GlobCOVER 14, 40, 60,90,100 and 120 in 2011, and tiny in 2009-2010.
-

	2011				2010				2009			
	free	constr_hard	constr_soft 5 % dev	constr_soft 10 %	free	constr_hard	constr_soft 5 % dev	constr_soft 10 %	free	constr_hard	constr_soft 5 % dev	constr_soft 10 %
Transport Highway	0.0001	0.0277	0.0244	0.0204	0.0000	0.0122	0.0101	0.0065	0.0026	0.0279	0.0242	0.0193
Transport Primary Road	0.0001	0.0450	0.0402	0.0337	0.0098	0.0641	0.0581	0.0496	0.0104	0.0526	0.0478	0.0399
Transport Secundary Road	0.0000	0.0164	0.0172	0.0151	0.0000	0.0206	0.0207	0.0180	0.0053	0.0227	0.0231	0.0204
Transport Tertiary Road	0.0000	0.0007	0.0008	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Residential	0.1133	0.0183	0.0366	0.0556	0.1083	0.0198	0.0394	0.0594	0.1072	0.0211	0.0422	0.0638
Industrial	0.0019	0.0265	0.0129	0.0053	0.0033	0.0287	0.0139	0.0080	0.0063	0.0305	0.0148	0.0084
Powerplants/Energy	0.2007	0.1628	0.1961	0.2136	0.2588	0.1759	0.2111	0.2262	0.2742	0.1872	0.2259	0.2430
GlobCOVER 14	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
GlobCOVER 20	0.0024	0.0024	0.0021	0.0019	0.0000	0.0000	0.0000	0.0000	0.0037	0.0036	0.0032	0.0031
GlobCOVER 30	0.3376	0.3622	0.3425	0.3319	0.3074	0.3483	0.3287	0.3187	0.3241	0.3721	0.3485	0.3365
GlobCOVER 40	0.0000	0.0000	0.0000	0.0000	0.0000	0.0009	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000
GlobCOVER 50	0.0123	0.0106	0.0104	0.0101	0.0226	0.0230	0.0228	0.0222	0.0103	0.0124	0.0112	0.0104
GlobCOVER 60	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
GlobCOVER 90	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
GlobCOVER 100	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
GlobCOVER 110	0.0292	0.0250	0.0248	0.0251	0.0278	0.0264	0.0260	0.0263	0.0238	0.0212	0.0211	0.0216
GlobCOVER 120	0.0000	0.0001	0.0001	0.0001	0.0018	0.0018	0.0016	0.0016	0.0013	0.0016	0.0015	0.0014
GlobCOVER 130	0.0795	0.0881	0.0829	0.0800	0.0622	0.0737	0.0687	0.0662	0.0544	0.0641	0.0595	0.0570
GlobCOVER 140	0.1971	0.1889	0.1840	0.1824	0.1744	0.1799	0.1746	0.1733	0.1579	0.1638	0.1581	0.1564
GlobCOVER 150	0.0176	0.0174	0.0171	0.0169	0.0165	0.0173	0.0168	0.0167	0.0129	0.0136	0.0133	0.0131
GlobCOVER 200	0.0082	0.0079	0.0078	0.0078	0.0071	0.0074	0.0072	0.0072	0.0057	0.0057	0.0056	0.0056
SUM	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
RMSE	0.60	0.54	0.59	0.63	0.62	0.48	0.52	0.54	0.60	0.51	0.53	0.55
R2	0.64	0.64	0.63	0.63	0.68	0.68	0.68	0.68	0.71	0.72	0.71	0.71
Biogenic/Biomass burning	68.39%	70.26%	67.16%	65.62%	61.97%	67.87%	64.66%	63.23%	59.40%	65.81%	62.19%	60.51%
Antropogenic Fraction	31.61%	29.74%	32.84%	34.38%	38.03%	32.13%	35.34%	36.77%	40.60%	34.19%	37.81%	39.49%
Traffic	*0.05%*	***30.17%***	*25.17%*	*20.17%*	*2.58%*	***30.17%***	*25.17%*	*20.17%*	*4.49%*	***30.17%***	*25.17%*	*20.17%*
Residential	*35.83%*	***6.16%***	*11.16%*	*16.16%*	*28.49%*	***6.16%***	*11.16%*	*16.16%*	*26.41%*	***6.16%***	*11.16%*	*16.16%*
Industrial	*0.61%*	***8.92%***	*0.04%*	*1.54%*	*0.87%*	***8.92%***	*3.93%*	*2.17%*	*1.55%*	***8.92%***	*3.92%*	*2.14%*
Powerplants	*63.51%*	***54.75%***	*59.73%*	*62.13%*	*68.06%*	***54.75%***	*59.74%*	*61.50%*	*67.55%*	***54.75%***	*59.75%*	*61.53%*

When performing the analysis on the monthly data, using the hard constrained version, we do find some interesting patterns emerging in the biogenic/anthropogenic split (as derived from the optimization weights) :



We clearly see the biogenic fraction decreasing in the South African winter (May-Sept) and increasing in the South African summer (one might suspect wildfires/agriculture to be a cause of this ?). This appears consistent with the following seasonal cycle, as identified by (REF, plot below), showing the biogenic component for 2 simulations as a function of month for southern Africa (in red). Here too, the biogenic component appears to reach it's minimum near May-June.

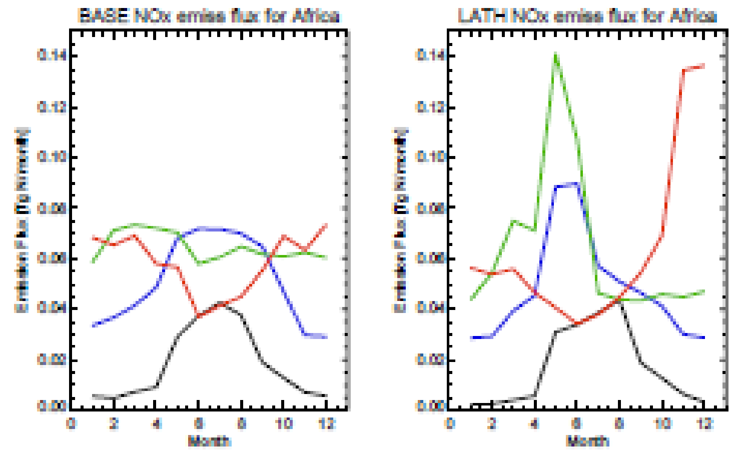


Fig. 1. A comparison of the monthly emission flux of NO from soils in TgN as provided in **(a)** the POET inventory (from Yienger and Levy, 1995) and **(b)** the RETRO inventory (from Lathiere et al., 2006). The key for the latitudinal regions is: (black) Saharan, (blue) Sahel, (green) Guinea and (red) southern Africa.

Effective downscaling & generation of high-resolution emissions

- We're using the hard constrained version for all years, example year 2009 (72 % of spatial variance explained by downscaling model)

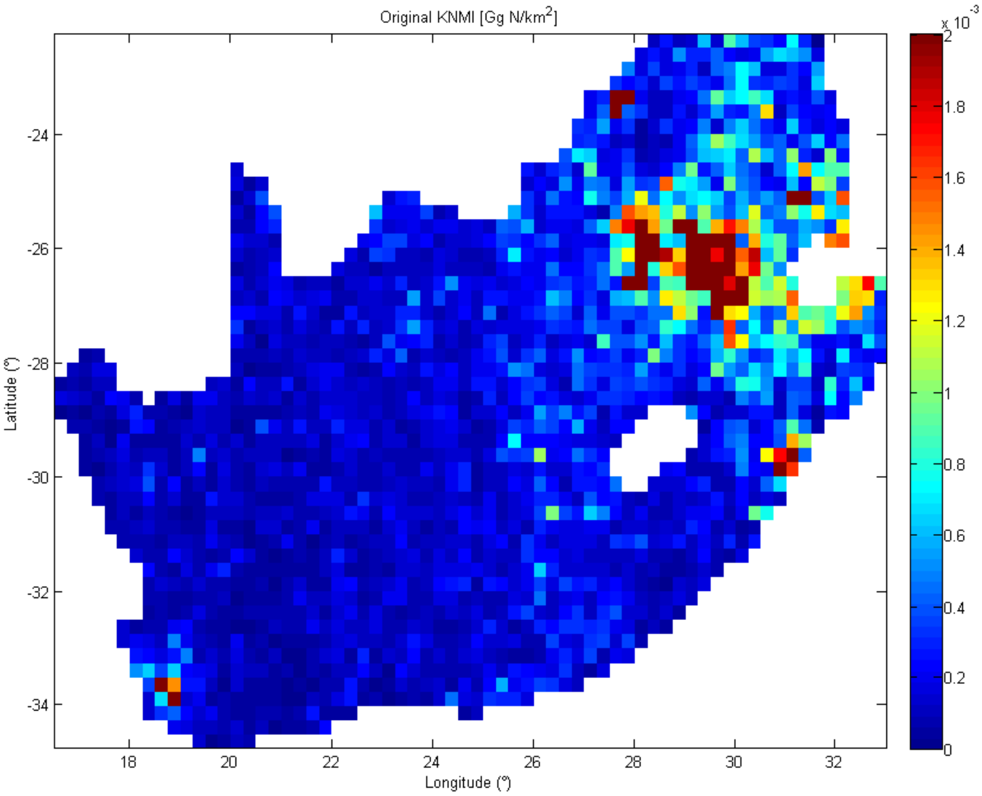


Figure 3 : Original KNMI low resolution emissions, normalized by surface area (2009)

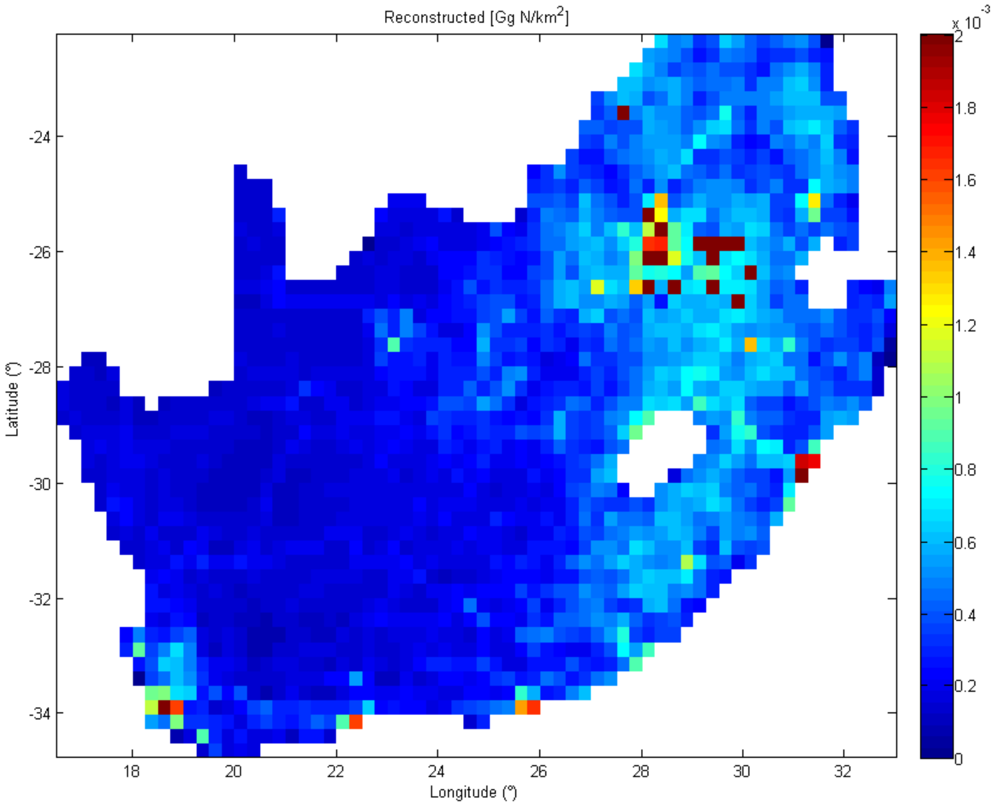


Figure 4 : Reconstructed low resolution emissions, normalized by surface area (2009)

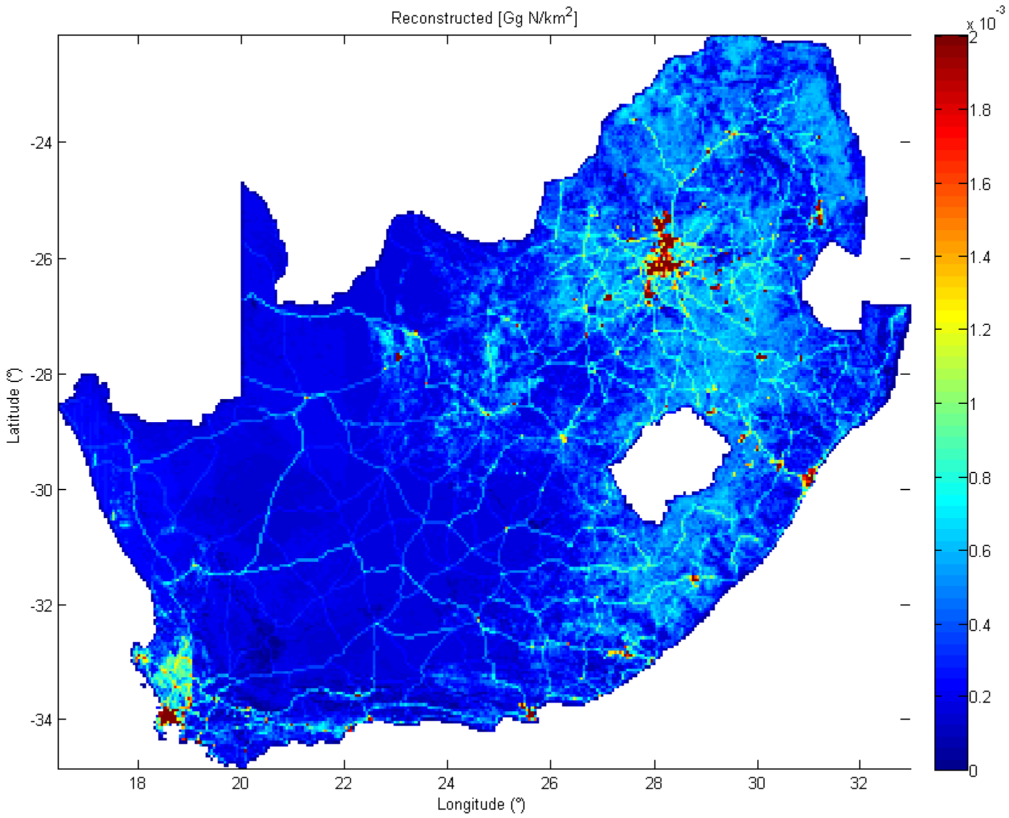


Figure 5 : High resolution emissions, normalised by area (2009)

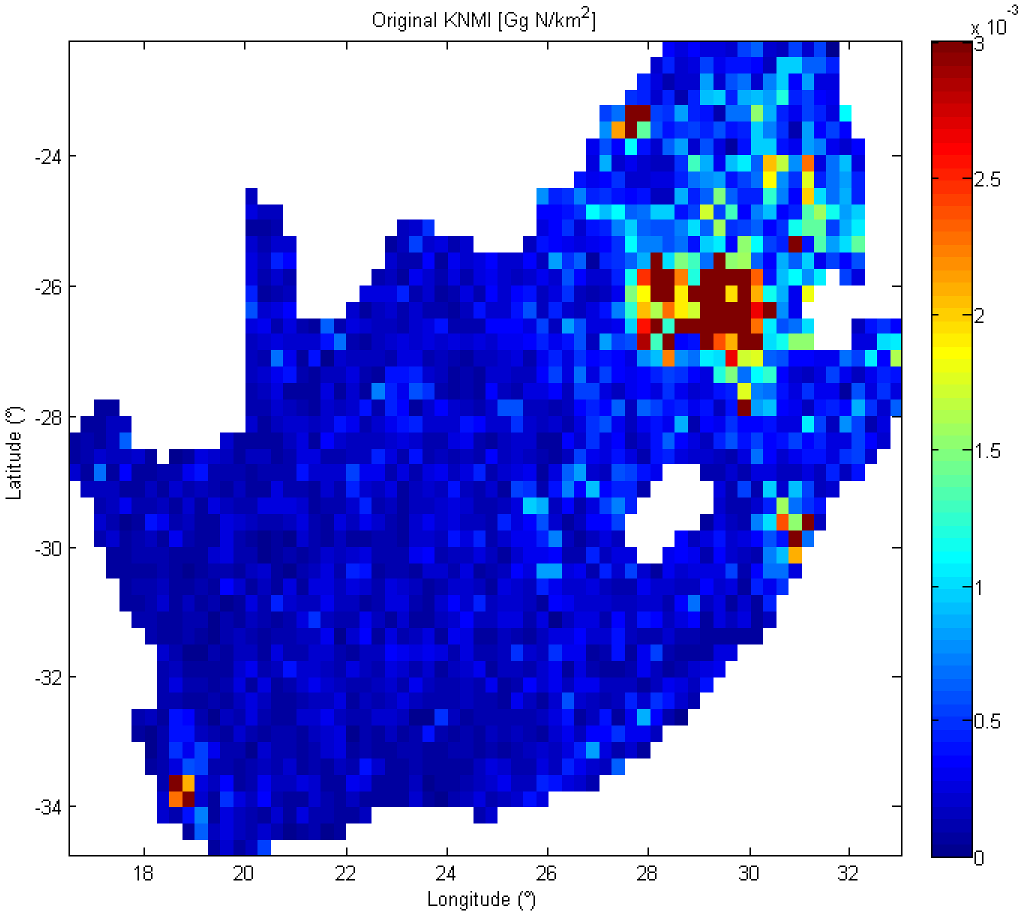


Figure 6 : Low resolution original emissions, normalized by surface area (2011)

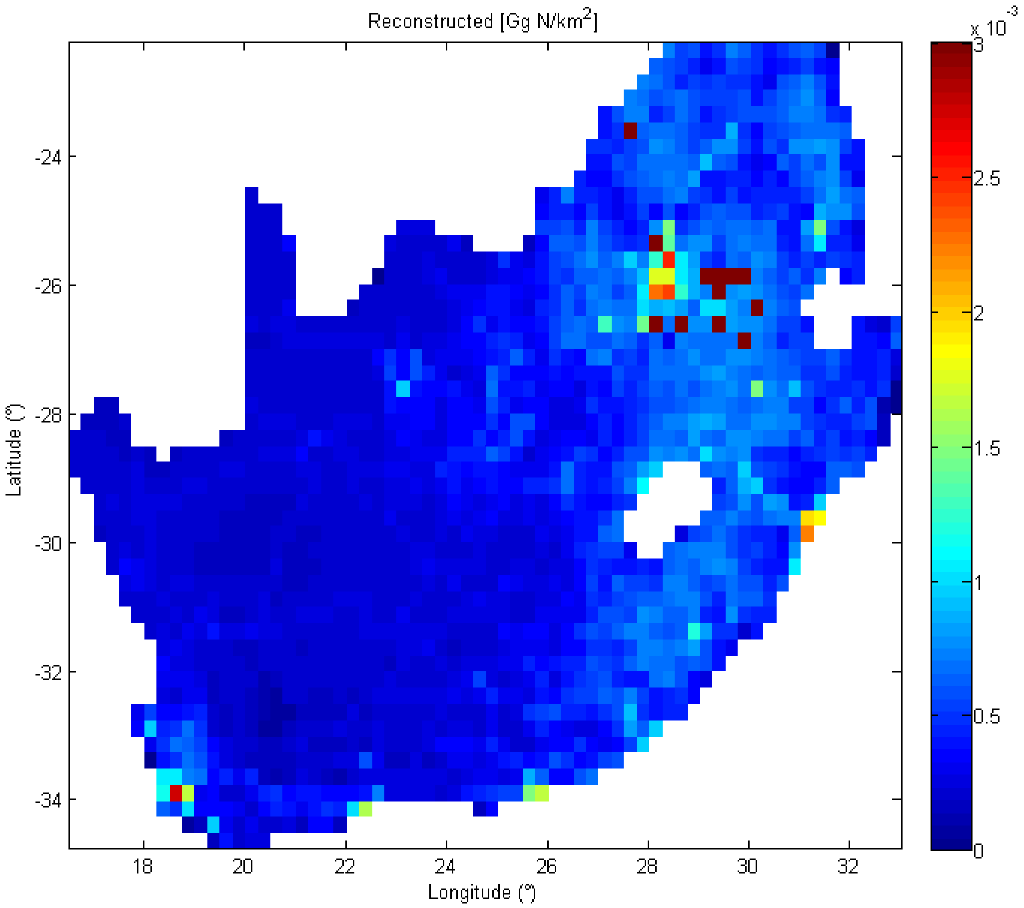


Figure 7 : Low resolution reconstructed emissions, normalized by surface area (2011)

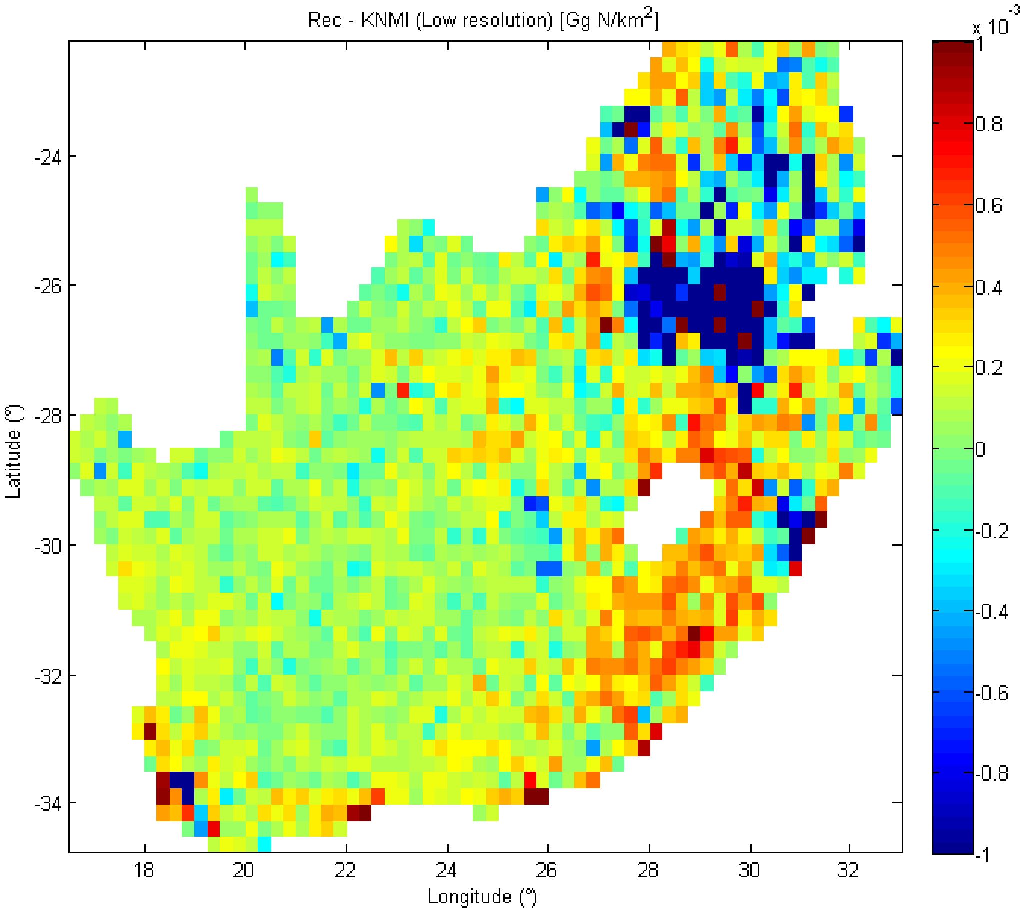


Figure 8 : Residue between reconstructed and original KNMI emissions, (2011)

The local residue can be quite sizeable… see remarks below

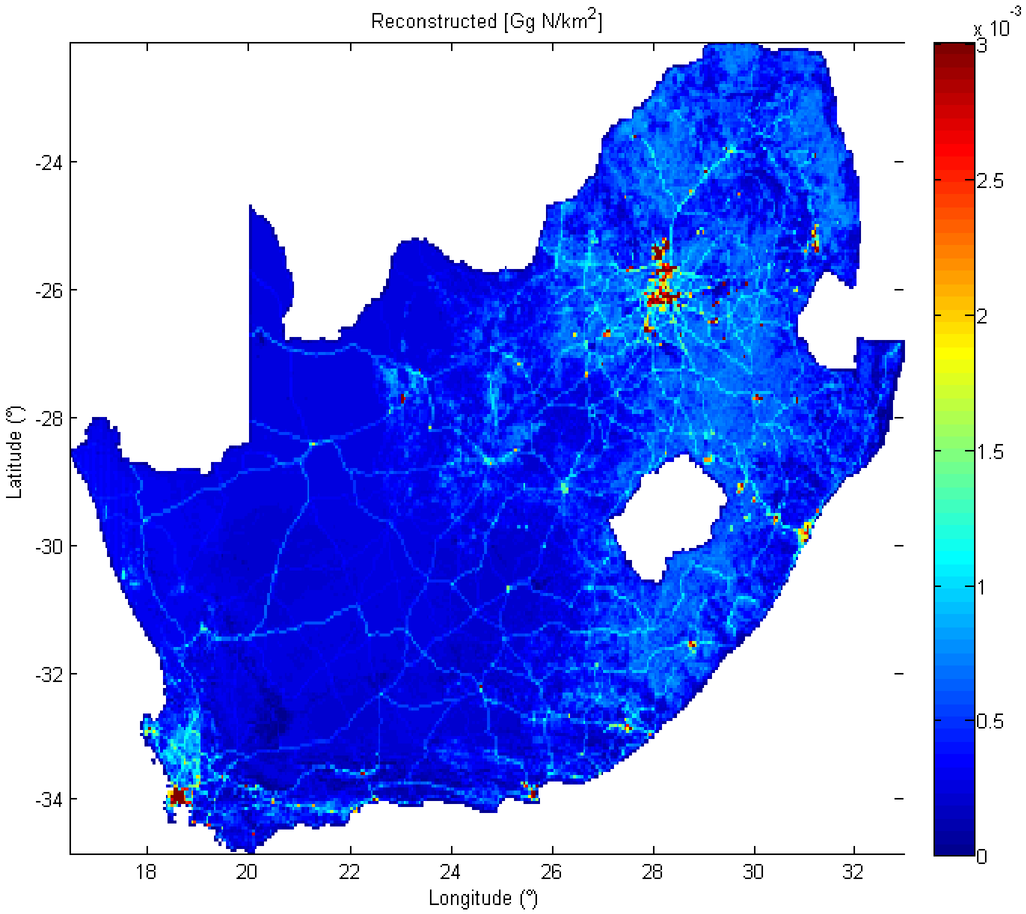


Figure 9 : High resolution downscalied emissioned, normalized by surface area (2011)

Note : **national totals** (as derived from original KNMI data) are **conserved i**n these maps, however , locally there might be differences due to the nature of our approach. Hence, we propose a local residue correction scheme…

Spatial view on the seasonality : May 2009 vs. January 2010

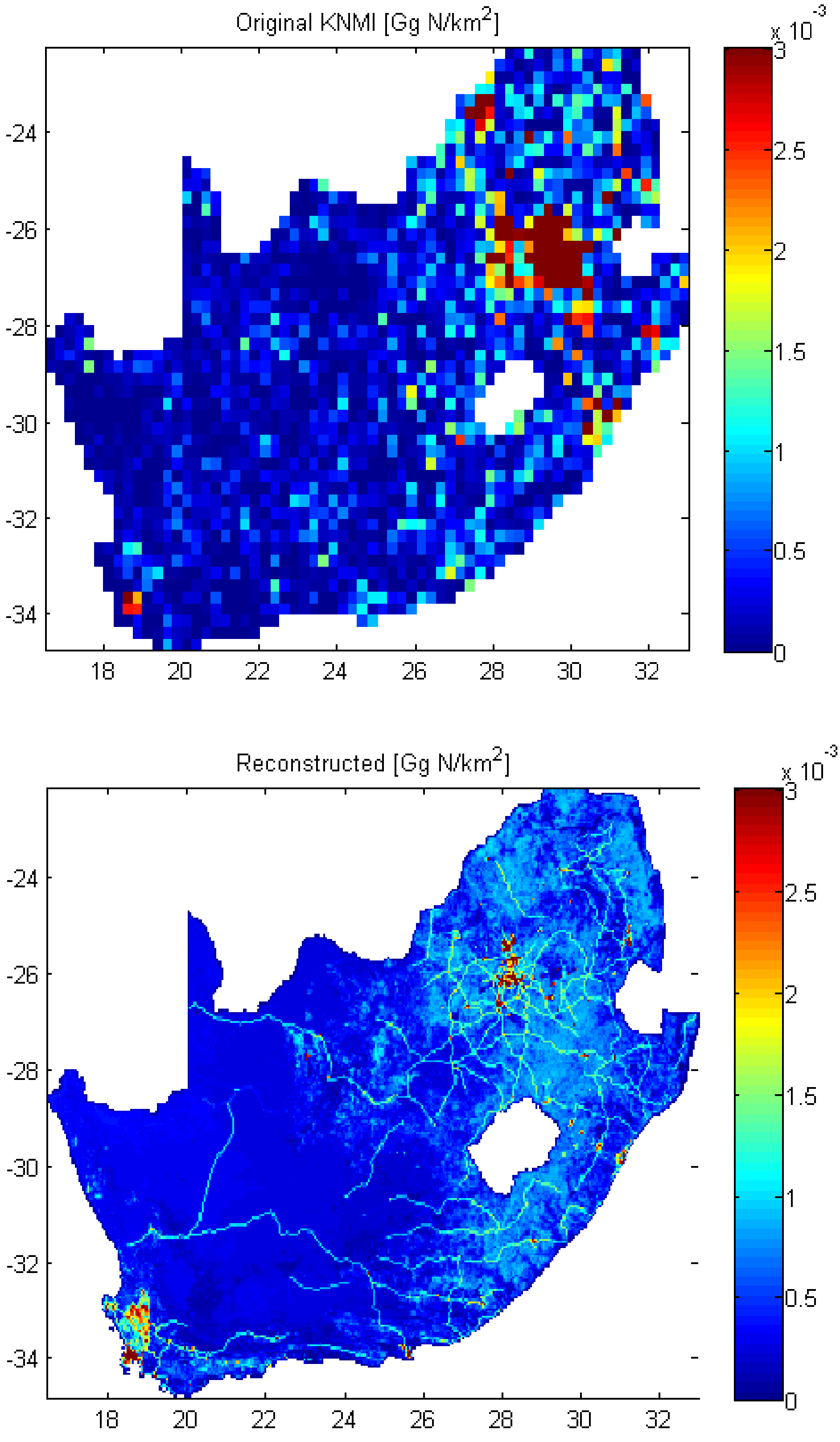


Figure 10 : High resolutioni emission, january 2010 → summer : more biogenic sources, top : original KNMI image, bottom : downscaled

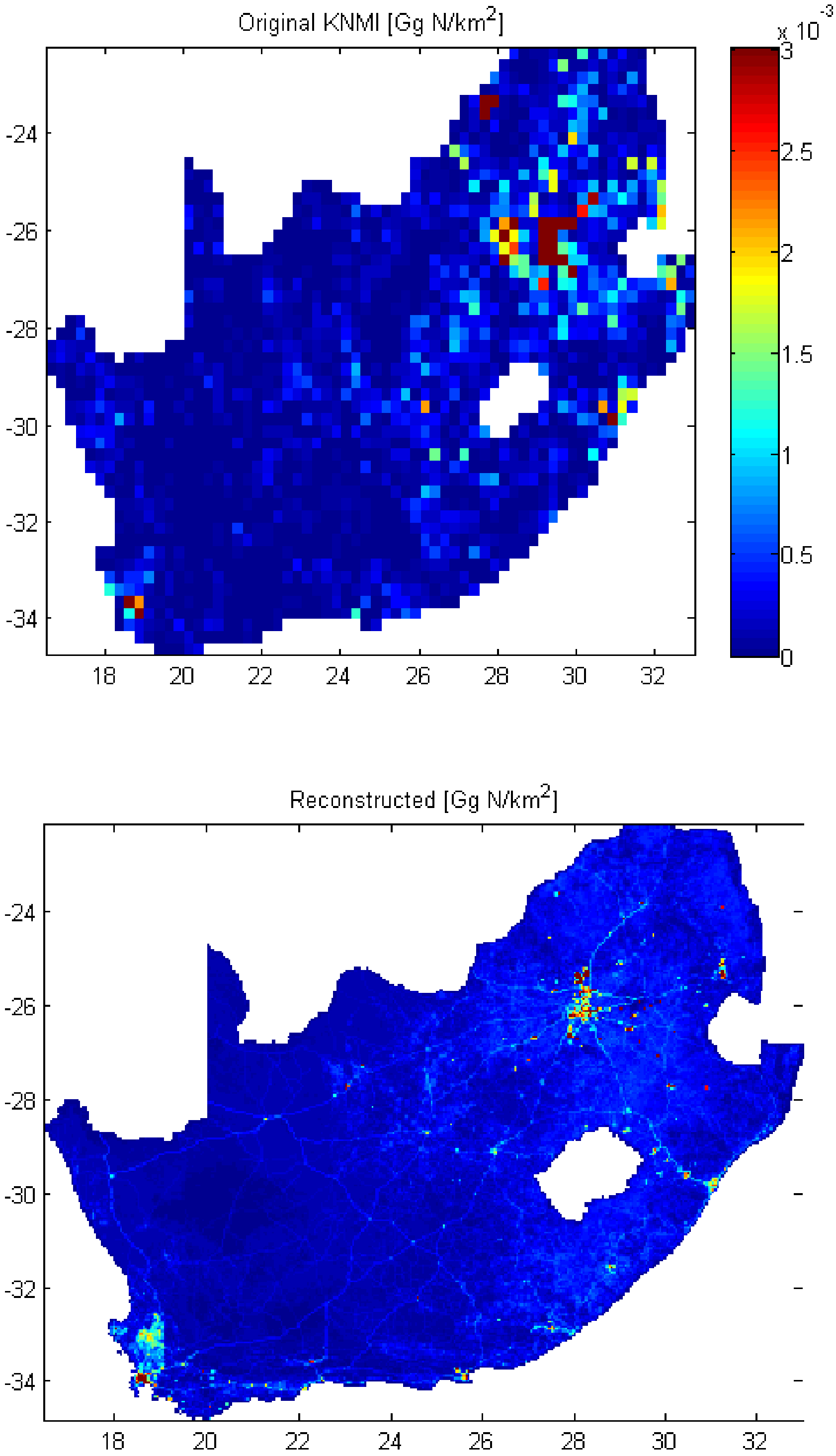


Figure 11 : high resolution emission, may 2010 → winter, fewer biogenic sources, top : original KNMI image, bottom : downscaled (same legend)

Local residue correction

- We tried, but so far not really satisfying results,
- It was noticed though that the optimization could suffer from the extreme values in the power plant pixels → some progress to be expected there

Additional remarks

- Original KNMI data has some smearing as well -> localized emissions (e.g. power plants) can be smeared out, which makes it hard for the downscaling (which used localized proxy parameters, e.g. coordinates of powerplants)
- Local proxy information, which contains local insights into the location of specific emissions could be a huge improvement, as opposed to relying on GlobCOVER.
- It was noted that some of the outlying emission values (e.g. coming from industry) had a significant impact on the optimization, so much can still be gained from making the optimization more robust against these outliers or leaving them out altogether. At present, there was not time to research this.
- In general : the methodology has very good potential, still some improvements to be made…

Appendix : GlobCOVER Classes

Value	Label
11	Post-flooding or irrigated croplands (or aquatic)
14	Rainfed croplands
20	Mosaic cropland (50-70%) / vegetation (grassland/shrubland/forest) (20-50%)
30	Mosaic vegetation (grassland/shrubland/forest) (50-70%) / cropland (20-50%)
40	Closed to open (>15%) broadleaved evergreen or semi-deciduous forest (>5m)
50	Closed (>40%) broadleaved deciduous forest (>5m)
60	Open (15-40%) broadleaved deciduous forest/woodland (>5m)
70	Closed (>40%) needleleaved evergreen forest (>5m)
90	Open (15-40%) needleleaved deciduous or evergreen forest (>5m)
100	Closed to open (>15%) mixed broadleaved and needleleaved forest (>5m)
110	Mosaic forest or shrubland (50-70%) / grassland (20-50%)
120	Mosaic grassland (50-70%) / forest or shrubland (20-50%)
130	Closed to open (>15%) (broadleaved or needleleaved, evergreen or deciduous) shrubland (<5m)
140	Closed to open (>15%) herbaceous vegetation (grassland, savannas or lichens/mosses)
150	Sparse (<15%) vegetation
160	Closed to open (>15%) broadleaved forest regularly flooded (semi-permanently or temporarily) - Fresh or brackish water
170	Closed (>40%) broadleaved forest or shrubland permanently flooded - Saline or brackish water
180	Closed to open (>15%) grassland or woody vegetation on regularly flooded or waterlogged soil - Fresh, brackish or saline water
190	Artificial surfaces and associated areas (Urban areas >50%)
200	Bare areas
210	Water bodies

220 Permanent snow and ice
230 No data (burnt areas, clouds,…)

References

Byrd, R.H., J. C. Gilbert, and J. Nocedal, "A Trust Region Method Based on Interior Point Techniques for Nonlinear Programming," Mathematical Programming, Vol 89, No. 1, pp. 149–185, 2000.

Byrd, R.H., Mary E. Hribar, and Jorge Nocedal, "An Interior Point Algorithm for Large-Scale Nonlinear Programming," SIAM Journal on Optimization, Vol 9, No. 4, pp. 877–900, 1999.

Waltz, R. A., J. L. Morales, J. Nocedal, and D. Orban, "An interior algorithm for nonlinear optimization that combines line search and trust region steps," Mathematical Programming, Vol 107, No. 3, pp. 391–408, 2006.