A statistical meta model of DNDC to estimate nitrogen fate and the water cycle at 1x1 km grid at Pan-European scale

Wolfgang Britz[1] & Adrian Leip[2]

[1] Rural, Water & Ecosystem Resource Unit
[2] Climate Change Unit
IES, JRC/ISPRA

Context of the study

- Develop and validate downscaling algorithms for agricultural drivers of environmental impacts from administrative regions:
  - To allow for spatial analysis of water abstraction, nutrient balances, greenhouse emissions, bio-diversity, landscape ..
  - To allow for ex-ante spatial CAP impact analysis based on results of agricultural sector models
- Specifically here:
  - Estimate a meta model from the bio-physical model DNDC,
  - In order to estimate nitrogen and water balance elements for clusters of 1x1 km grid cells

Overview on algorithm

- Three step procedure:
  1. Estimation of non-linear regression models for result variables of DNDC
  2. Calibrate the potential yields used in in the meta model in order to let the meta model reproduced to yield and farming practise estimated at 1x1 km grid
  3. Forecasting the results from DNDC at 1x1 km grid

Data for the Estimation Step

- Data sources are:
  - About 100.000 DNDC runs
  - Based on combinations of crops (soft wheat, durum, barley, maize, rye, oats, paddy rice, potatoes, sugar beet, pulses, rape seed, sunflower seed, soya seed, tomatoes, other vegetables, annual fodder) and HSMU (homogenous soil mapping units) representative for NUTS II regions
  - For each crop-HSMU (climate/soil) combination, 8 runs of 100 years were performed, covering all possible combinations of:
    - No – high mineral fertilisation dose
    - No – high manure fertilisation dose
    - No – automated irrigation
  - The last year from the runs was used as to let carbon and nitrogen content level out to continuous arable cropping

Variables in the Estimation

- Explanatory variables:
  - Soil parameters (clay content, packing density, soil organic carbon, PH)
  - Average annual climatic variables (rainfall, maximum and minimum temperature, E0, ET0)
  - Mineral and manure nitrogen applied, realized yield, potential yield
- All explanatory variables are possibly included:
  - in linear and quadratic form, as square roots and logs
  - multiplied with each other variable
  - multiplied with the log of organic carbon
  - multiplied with the square of organic carbon
  - which leads to possibly 140 explanatory variables

Estimation procedure

- Regression models were run per crop, differentiating irrigated and non-irrigated runs for:
  - Different N losses (leaching, gaseous emissions)
  - Mineralization
  - Water balance elements (transpiration, evaporation, leaching)
  - Crop yields (measured in removed nitrogen)
- Due to low number of observation or other reasons, some crops were grouped (durum and soft wheat; tomatoes and other vegetables; oats, rye and rice)
- Backwards elimination of insignificant variables based on OLS:
  - Remove step-wise least insignificant variable, as long:
    - The adjusted R squared is still increasing
    - The number of regressors exceeds 1/5 of the observations
    - There are variables with Prob=0 > 0.1%

Results

- R^2 are typically for 40% of the estimation >=95%, for 72% >=85% and only for 10% below 75%
- Some problems in selected cases with water balance elements and gaseous emissions (R^2 around 65%)
- Possible causes:
  - DNDC runs were using the same weather each year, which may in some cases far away from the average climate data used as explanatory variables (new runs healing that problem are planned)
  - Mixing rice (where only a few observations are available) and other cereals, and tomatoes and other vegetables
- But:
  - Remember: all variables are significant at the 0.1% level
  - Additional security limits introduced in forecasting, discussed later

R2 of estimations

		Nitrogen								Water				Yield
		Leached	Mineral.	Losses	NH3	N2/N2O/NO	N2	N2O	NO	Total	Evapor.	Leached	Transpir.	
Barley	no Irri	0.95	1.00	0.98	0.97	0.82	0.87	0.64	0.89	0.86	0.83	0.86	0.94	0.96
	Irri	0.94	1.00	0.99	0.99	0.94	0.98	0.82	0.95	0.85	0.85	0.85	0.93	0.98
Maize	no Irri	0.96	1.00	0.96	0.95	0.72	0.66	0.76	0.89	0.85	0.94	0.82	0.96	0.91
	Irri	0.83	1.00	0.93	0.95	0.72	0.71	0.65	0.89	0.82	0.88	0.75	0.89	0.91
Potatoes	no Irri	0.80	1.00	0.97	0.98	0.87	0.87	0.84	0.95	0.84	0.92	0.83	0.80	1.00
	Irri	0.80	1.00	0.97	0.98	0.90	0.88	0.87	0.95	0.87	0.93	0.84	0.80	1.00
Sugar beet	no Irri	0.91	0.99	0.98	0.99	0.81	0.78	0.81	0.93	0.84	0.95	0.78	0.95	0.99
	Irri	0.80	0.99	0.99	0.99	0.88	0.84	0.87	0.95	0.92	0.94	0.88	0.88	0.99
Pulses	no Irri	0.98	1.00	0.99	0.99	0.94	0.94	0.77	0.93	0.85	0.92	0.81	1.00	0.95
	Irri	1.00	1.00	1.00	1.00	0.99	0.99	0.99	0.99	1.00	1.00	0.99	1.00	0.97
Rapeseed	no Irri	0.93	1.00	0.99	0.98	0.90	0.89	0.84	0.93	0.75	0.87	0.72	0.97	0.96
	Irri	0.98	1.00	1.00	0.99	0.91	0.85	0.90	0.98	0.95	0.92	0.93	0.98	0.98
Sunflower seed	no Irri	0.96	1.00	1.00	0.99	0.93	0.93	0.88	0.94	0.83	0.93	0.77	1.00	0.98
	Irri	0.96	1.00	1.00	1.00	0.93	0.90	0.91	0.96	0.86	0.96	0.79	1.00	0.99
Annual fodder	no Irri	0.98	0.98	1.00	0.98	0.86	0.89	0.56	0.91	0.69	0.89	0.76	0.88	0.90
	Irri	0.98	0.99	1.00	0.98	0.77	0.79	0.60	0.91	0.74	0.83	0.77	0.83	0.90
Wheat	no Irri	0.93	0.98	0.98	0.96	0.75	0.74	0.67	0.84	0.73	0.72	0.71	0.76	0.89
	Irri	0.94	0.98	0.99	0.98	0.80	0.71	0.79	0.89	0.86	0.72	0.78	0.74	0.91
Other cereals	no Irri	0.86	0.99	0.97	0.96	0.68	0.63	0.63	0.86	0.76	0.74	0.73	0.95	0.90
	Irri	0.78	1.00	0.99	0.99	0.87	0.87	0.81	0.94	0.87	0.93	0.84	0.97	0.97
Vegetables	no Irri	0.87	0.99	0.96	0.97	0.90	0.89	0.81	0.92	0.69	0.92	0.66	0.73	0.95
	Irri	0.86	0.99	0.96	0.97	0.88	0.87	0.76	0.93	0.71	0.91	0.67	0.72	0.95

Min		0.78	0.98	0.93	0.95	0.68	0.63	0.56	0.84	0.69	0.72	0.66	0.72	0.89
Max		1.00	1.00	1.00	1.00	0.99	0.99	0.99	0.99	1.00	1.00	0.99	1.00	1.00
Mean		0.91	0.99	0.98	0.98	0.85	0.84	0.78	0.92	0.82	0.89	0.80	0.89	0.95

DNDC meta model estimation

Re-calibrating the statistical model

- Yields in the spatial down-scaling are derived from MARS non-water limited and water limited potential yields, estimated irrigation shares, and scaled as too recover average yields at NUTS II
- Unfortunately, nothing guarantees that DNDC will simulated for soil/climate/fertilising management combination that estimated yields
- Rather simple and drastic remedy: DNDC potential yields – the main driver of realized yields in DNDC – are re-estimated from the yield estimate in CAPRI

$$Yield^{*}_{HSMU,c} = f(\overline{potYield}_{HSMU,}\mathbf{x}_{HSMU})$$

$$\Leftrightarrow$$

$$potYield^{*}_{HSMU,c} = f(\overline{Yield}_{HSMU,}\mathbf{x}_{HSMU})$$

Forecasting at 1x1 km grid I

- After re-calibration of the potential yield, result variables are estimated from the regression results
- Results were bounded to the min and max of the result variables in the observation sample (crude outlayer control)
- Specific treatment of results with adding up conditions, as discussed below

Forecasting at 1x1 km grid I (cntd.)

- Observation that e.g. total nitrogen losses have typically a higher R2 as the individual loss positions
- In such cases, a correction is introduced, which distribution the difference between the estimated total (e.g. total N loss) and the estimated sum over the single positions (e.g. leaching, NH3, N2, N2O, NO)

$$Err = Y_{tot}^{*} - \sum_{pos} Y_{pos}^{*} \qquad CorrWeights = \sum_{pos} Y_{pos}^{*} \sqrt{1 - R_{pos}^{2}} + Y_{tot}^{*} \sqrt{1 - R_{tot}^{2}}$$

$$Y_{pos}^{corr} = \max\left(\min\left(10, Y_{pos}^{*}/2\right), Y_{pos}^{*} + \frac{Err}{CorrWeights} Y_{pos}^{*} \sqrt{1 - R_{pos}^{2}} \right)$$

$$Y_{tot}^{corr} = \min\left(2\left(Y_{tot}^{*} + 1\right), \max\left(Y_{tot}^{*}/2, Y_{tot}^{*} - \frac{Err}{CorrWeights} Y_{tot}^{*} \sqrt{1 - R_{tot}^{2}} \right)\right)$$

$$Y_{pos}^{corrf} = Y_{pos}^{corr} \frac{\sum_{pos} Y_{pos}^{corr}}{Y_{tot}^{corr}}$$

DNDC meta model estimation

Summary

- Methodology allows:
  - to estimate rather efficiently major result variables from DNDC for crops under climate-soil-farming practise combination
  - removes bias resulting from non-calibrated crop growth model by re-calibrating the potential yields in the meta model
- Major remaining problems:
  - No observations for permanent grass land and tree cultures
  - Runs used same weather in every year