Pan-European Estimation of crop specific irrigation shares at 1x1 km grid

Wolfgang Britz & Gunter Wriedt

Rural, Water & Ecosystem Resource 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 irrigation shares,
  - In order to estimate yields, water abstraction, fertilizer application rates and other crop specific parameters for clusters of 1x1 km grid cells
  - European land use map with irrigation shares

Irrigation shares at 1x1 km grid

Overview on algorithm

- Two step procedure:

1. Estimation of missing irrigation shares at the level of administrative regions
2. Forecasting at 1x1 km grid, using a Highest Posterior Density estimator

Data for the Estimation Step

- Data sources are:
  - the Farm Structure Survey 1999 (agricultural census in December), available from Eurostat at Pan-European level for NUTS II/III regions (ca. 500 regions for EU25), reporting areas irrigated once a year (but no crop specific data)
  - And irrigated areas for selected crops, reported for selected European NUTS II regions (France, Italy, Spain, Greece), again for 1999

Variables in the Estimation

- Irrigation shares are estimated for durum wheat, maize, potatoes, sugar beet, sunflower, soya, vineyards, forage crops, citrus, other fruit & berry plantations, average of all other crops
- Explanatory variables:
  - Mean slope and altitude
  - Average rain fall
  - Temperature sum over 8 months
  - Vegetation days over 8 months
- All explanatory variables are possibly included in linear and quadratic form, and as products

Estimation procedure

- Three step procedure:
  1. Backwards elimination of insignificant variables based on OLS:
     - Remove step-wise least insignificant variables, 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 > 10%
  2. Non-linear Ordinary Least Squares with the remaining variables, allowing for an exponential transformation between 0.75 and 1.25, to provide a starting point for
  3. Final non-linear tobit estimation, using the remaining variables

OLS/tobit regression function

Error term

Observed irrigation share

Explanatory variables

$$e_{c,r} = Y_{c,r} - B_{c,const} - \sum_i B_{c,i} X_{i,r}^{\exp_{c,i}}$$

Constant term

Regression coefficients

Exponential transformation
$0.75 < exp < 1.25$

c	crops
i	regressors
r	NUTS II regions

Irrigation shares at 1x1 km grid

The tobit Estimator

- tobit estimator:
  - Respects the truncated nature of the variables to estimate, which are by definition non-negative
  - Thus prevents estimation bias resulting from the truncated nature of the variable
  - Maximum Likelihood Estimator, here assumption of normally distributed error term is maintained as in OLS

Tobit Maximum Likelihood estimator

Log of Likehood

squared Errors

Estimated variance of error terms

$$\max_{\hat{\sigma},B,e,\exp} \log(L) = \sum_{c,r \neg Y_{c,r}>0} -\frac{1}{2} e_{c,r}{}^{2} \Big/ \hat{\sigma}^{2} - \log(\hat{\sigma})$$

Log Prob. Density of error terms for non-zero obs.

$$+ \sum_{c,r \neg Y_{c,r}=0} \log\left[1 - errorf\left(e_{c,r} / \hat{\sigma}\right)\right]$$

Log Prob. Density of error terms for zero obs.

Irrigation shares at 1x1 km grid

Results: Discussion

- Average R2 in the range of 50%
- But remember: all regression coefficients are significantly different from zero at last at the 90% level
- And final estimation step will ensure consistency to total irrigated area (which in most regions without reported irrigation shares per crop will be zero)

Irrigation shares at 1x1 km grid

Forecasting at 1x1 km grid I

- Find the most probable irrigation shares, s.t.
  - Total irrigation area from FSS is met
  - Deviation between irrigated share at 1x1 km grid and share from FAO map becomes small
  - The weighted average of irrigated (potential yield) and non-irrigated (potential water limited yields) yields at 1x1 km grid level, aggregated to NUTS II regions, comes close to observed yields at NUTS II level

Forecasting at 1x1 km grid II

- Application of Highest Posterior Density estimator:
  - Maximize joint probability density of forecasts for the crop specific irrigation shares
  - Using the statistically observed or estimated irrigation shares as means of the a priori distribution
- Motivation:
  - Ensure mutual compatibility of downscaled results and statistical data
  - Exploit the informational content of the different data sources (FSS irrigation shares, FAO map, potential yields from MARS)

HDP estimator - constraints

Exhaustion of FSS irrigated area

$$\overline{Y}_{FSS,"UAA"}\overline{ha}_{FSS,"UAA"} = \sum_{HSMU \in FSS} Y_{HSMU,"UAA"}\overline{ha}_{HSMU,"UAA"}$$

Definition of total irrigated share

$$Y_{HSMU,"UAA"}\overline{ha}_{HSMU,"UAA"} = \sum_{c} Y_{c,a}\overline{ha}_{HSMUc}$$

Definition of relative error of NUTS II yield

$$YieldCorr_{NUTSII,c} = \frac{\sum_{HSMU}\left(Y_{HSMU,c}\overline{Yield}_{HSMU,pot} + \left[\left(1 - Y_{HSMU,c}\right)\overline{Yield}_{HSMU,watLimited}\right]\right)\overline{ha}_{HSMUc}}{\overline{Yield}_{NUTSII,c}\overline{ha}_{NUTSII,c}}$$

Irrigation shares at 1x1 km grid

HDP estimator - objective

Log of joint prob.
Density, weighing observations with crop areas

Manual weights per crops
(irrigated share given, residual, estimated, estimated residual)

$$\max_{Y} \log(L) =$$

$$\sum_{HSMU,c}\left(-\frac{1}{2}\overline{w}_{NUTSII,c}\left[\frac{Y_{HSMU,c}-\overline{Y}_{NUTSII,c}}{\overline{Y}_{NUTSII,c}}\right]^2\overline{ha}_{HSMU,c}\right)$$

Deviation of HSMU share from FAO data

$$+\sum_{HSMU,c}\left(-\frac{1}{2}\left[\frac{Y_{HSMU,"UAA"}-\overline{Y}_{HSMU,"UAA"}}{\overline{Y}_{HSMU,"UAA"}}\right]^2\overline{ha}_{HSMU,"UAA"}\right)$$

Deviation of HSMU share for crops from NUTS II one

$$+\sum_{HSMU,c}\left(-\frac{1}{2}\left[YieldCorr_{NUTSII,c}-1\right]^2\overline{ha}_{NUTSII,c}\right)$$

Deviation from NUTSII yields

Irrigation shares at 1x1 km grid

Summary

- Methodology allows:
  - to estimate probable irrigation share per crop and for total area at 1x1 km grid cells, exploiting information from different sources
  - while consistently down-scaling from administrative regions
- Allows to improve yield estimation for 1x1 km grid cells, and to improve environmental impact analysis
- .. especially, as meta model from DNDC (N compartments, water abstraction) differentiates between irrigated and non-irrigated farming practise