Modelling irrigation water consumption at microdata level in the Census of Agriculture 2009 (Spain)

Saralegui, Jorge
INE
Paseo Castellana 183
28046 Madrid (Spain)
jorge.saralegui.gil @ine.es

1. Introduction

During the first half of 2010, INE (Spain) carried out the second phase of the Agricultural Census 2009, by collecting additional data on production methods through the statistical operation Survey of Production Methods (SPM 2009).

The survey was addressed to a sample of 60,000 holdings drawn from the census frame (1,400,000 holdings, approx.). One of the goals of SPM 2009, is to estimate at the microdata level for each of the sampling units, the volume of irrigation water used for crop year 2009. The INE must submit these estimates to Eurostat in December 2011, as required by FSS EU Regulation.

Preliminary studies came the conclusion on the non-feasibility of including in the questionnaires direct questions about amounts of water consumed by the holding , mainly because of the risk of incurring high observation and measurement errors. Consequently the decision was to launch a project of model assisted estimation in several stages, combining direct data with additional information available from external sources.

The paper first formulates the target variables and the primary information available to build the model. Next, the general model and the process of preparing the input data bases are described , to go on with the formulation of the different estimation stages. Finally a summary of real results of the model for some crops in a particular region of the south of Spain are presented together with a comparison with external administrative data –'irrigations quantities recommended'- referred to the same crops.

2. Statistical information available as input to the model

2.1 Data from the 2009 Census of Agriculture

In the 2009 Agricultural Census information is requested on the irrigated area per crops for each holding, according to a standard classification adapted to the requirements of the EU Regulation. At this point it is important to keep in mind the fact that the census files contain the same elementary identifiers as those used in the EMP2009, as the latter is a sample of holdings drawn from the census, considered as sampling frame.

2.2 Complementary information collected in SPM 2009

Although the SPM 2009 questionnaire does not provide direct information on water consumption by the holding, important data is collected about certain characteristics of irrigation: a) distribution (%) of the volume of water by source (surface, underground) and b) distribution (%) of the irrigated area per irrigation technique applied (drip, localised, surface irrig., sprinkling).

The latter distribution will be used, as shown below, in the second stage of the model, (adjustment according to the efficiency of irrigation techniques).

2.3 Agro-climatic data

The Ministry of Environment and Rural and Marine Affairs (MARM) manages a network of 419 agro-climatic stations geo-referenced (through UTM coordinates) that register the weather-related and other parameters that condition the local growth of crops (rain, wind, sun, soil moisture). In particular, each agro-climatic station provides average evapotranspiration data (ETP) which corresponds to that occurring on the surface of a standard plot completely covered with (common) grass.

It should be noted that a municipality and therefore the irrigated area of holdings included within them, are assigned to one, and only one, agro-climatic station.

2.4 Areas of irrigation and crop coefficients

The MARM, as part of its preparatory technical studies for the elaboration of the National Hydrological Plan 2008 (NHP) conducted a subdivision of Spain in 811 irrigation areas , with its limits defined by UTM coordinates, defined on the basis of agronomic characteristics and water environment conditions (shared river basins, common water supply for the area, type of crops, among others). Within the irrigation area , each crop is assigned a crop coefficient Kc, adimensional, which is the correction factor to be applied to the standard evapotranspiration ETP as defined above. The Kc for each crop within an irrigation area is identical, regardless of the situation of the crop. However, Kc coefficients of the same crop may vary from one irrigation area to another. MARM databases also allow to determine, for every crop present in the irrigation area, an associated ***q*** coefficient representing the weight of the crop acreage in respect of the total of the irrigation area.

3. Target variables

The main target variable, which must be calculated at the microdata level is the total volume of water used for irrigation in each of SPM 2009 holdings (about 60,000 units, of which 35,000 are expected to declare irrigated crops).

The information to be produced by the model is the 'theoretical water needs' of each crop grown in a given municipality. The municipality is the basic geographical unit in the context of the survey to which every holding(and its crops) can be assigned. In addition to its official code, the municipalities perimeter is also geo-referenced (UTM coordinates).

The water needs of each crop within a municipality is thus applied to the irrigated surface of the crop declared by the holding. For a given month it can be formulated as:

$$H = (Kc \cdot ETP) - Pe$$

Where:

- *H* is the theoretical crop water need for the month (m3).
- *Kc* is the crop coefficient, within each irrigation area (without dimension).
- *ETP* is the standard evapo_transpiration. ETP units are in mm, which are converted to m3/ha for introduction in the model.
- *Pe* is effective precipitation for the month, mm, which are converted to m3/ha for introduction in the model.

The theoretical crop water needs in a year (in this case, 2009) are given by the sum of the theoretical water needs calculated for each month.

4. Preparatory work of the auxiliary information for the model

4.1 Assigning agro-climatic station to municipalities

In this stage, the correspondence between municipalities (about 8000 in Spain) and agro-climatic stations is defined. At this end, the cartography unit within INE , by using UTM coordinates, assigns the appropriate station to each municipality of the SPM 2009 frame, based on the geographical proximity to its centroid.

4.2 Crops correspondence

The crop coefficients Kc to be introduced in the model, are available for 300 categories of crops that do not correspond with the 50 categories used in the agricultural census.

Crop categories used in the technical studies for the NHP (MARM) are aimed at classifying crops based on their water requirements and, particularly, to reflect its seasonality, while crop categories in the Agricultural Census follow an harmonized list, according to EU regulation on the Farm Structure Survey. It is therefore necessary to built a correspondence between crop categories used in the NHP 2008 (MARM crops from now on) and crop categories used in the Agricultural Census 2009 (INE crops in the following).

4.3 Delimitation of polygons intersection municipality-irrigation area

When it comes to the spatial integration of the information, it may occur or not intersections between the polygons that form the municipalities and the polygons that form the irrigated areas, which poses further methodological problems. The following table summarizes the distribution of the number of polygons generated by the intersection of municipalities and irrigation areas:

Table 1. Number of polygons resulting from the intersection Municipality –irrigation area

Number of irrigation areas intersected	**Nr of municipalities**	**Nr resulting polygons**
1	3.017	3.017
2	2.963	5.926
3	1.293	3.879
4 ó más	907	44128
Total	**8.180**	**17.252**

The intersection of an irrigation area with one or more municipalities and vice versa, makes it necessary to define coefficients w, whose use will be shown in the following, representing the weight in respect of the total irrigation area of the surface intersected with a municipality. The general idea is to assign weighted means of Kc values for the same crop among all irrigation areas intersecting the municipality.

Preparatory works described in 4.1, 4.2 and 4.3, allow the generation of databases to be used in subsequent phases of the project, containing ETP, Kc and Pe values per crop within every polygon of the intercession.

ETP data are extremely variable, and greater than zero in every month of the year. For region C chosen here as an example to present some data, ETP values rang from 404-1517 m3 / h, the lowest in spring-autumn and the highest in the summer months , when Pe (rainfall) values are minimal . For seasonal crops, Kc values only take nonzero values for the months with the presence of the crop, while in perennial crops (i.e. citrus) they have a very stable value along the twelve months of the year . For region C, the Kc for citrus vary from 0.500 to 0.650.

5. Final formulation of the theoretical model of water needs

M: Municipality;
a: irrigation area intersecting with municipality M;
n: INE crop;
j: MARM crop;
i: irrigation month ;
H: water needs;

K_{aijn}: coefficient of irrigation in the area (a) of MARM crop (j) corresponding to INE crop (n) , in month (i) ;

Let q be the relative weights:

$$\sum_{j \in a} q_{aj} = 1$$

corresponding to fraction of the acreage covered by each crop MARM (j) in respect of the total irrigated area (a).

Let be the adjusted weights of the MARM crops (a) corresponding to the same INE crop (n) such as:

$\sum_{j\in a,n} \overline{q}_{ajn} = 1$; Es decir, $\overline{q}_{ajn} = \dfrac{q_{ajn}}{\sum_{j \in a,i,n} q_{ajn}}$

Let ETO_i^M and P_i^M be respectively the evapotranspiration and monthly rainfall for the month i in the municipality M (agro-climatic station assigned by the criterion of proximity UTM to its centroid).

Then, the annual water needs, H, for an INE crop (n) in the irrigation area (a) that intersects with the municipality are given by the expression:

$$H_{an}^M = \sum_{j\in a,n} \overline{q}_{ajn} \sum_{i}^{12} (K_{aijn}.ETO_i^M - P_i^M);$$

with $H_{aijn}^M = 0$ if $K_{aijn} \cdot ETO_i^M - P_i^M \le 0$, where $H_{aijn}^M = K_{aijn} \cdot ETO_i^M - P_i^M$

Let w be the weights:

$\sum_{a\in M} w_a^M = 1$; corresponding to the intersection of each irrigation area with a municipality M, adjusted as

And

$\overline{w}_{an}^M = \dfrac{w_a^M \delta_{an}}{\sum w_a^M \delta_{an}}$; wtih $\delta_{an} =$ 0 if and only if $\sum_{j\in n} q_{ajn} = 0$; $\delta_{an} = 1$ if not;

Then: $H_n^M = \sum_{a\in M} \overline{w}_{an}^M H_{an}^M$

(annual theoretical water needs for an INE crop n in a municipality M).

6. Adjustments according to the efficiency of irrigation techniques

At this stage of the model the effective consumption of irrigation water per ha to be allocated to each crop produced by the holding is estimated. At this end, a correction to the water needs as estimated above, is applied based on coefficients provided by MARM agricultural expert services, related to the irrigation technique in use by the holding. It is collected through SPM 2009 questionnaire as % of irrigated area under each irrigation technique. Efficiency coefficiens related to the irrigation management are also available.

The coefficients currently vary at the national level in the following ranges:
Surface irrigation (g): 0.60 to 0.70; Sprinkling, rangers (a): 0.70 to 0.90; Localised, drip (l): 0.90 to 0.95
Management efficiency coefficients (e): 0.90 to 0.95.

Thus, the effective consumption of a crop (n) for a particular holding in a municipality M becomes (the index for the holding is omitted):

$$\overline{H}_n^M = H_n^M . \frac{1}{(k(g)p(g) + k(a)p(a) + k(l)p(l))k(e)}$$

Where a simplification is used to apply to the theoretical water needs the mean of the central values in

each interval of the coefficients K of irrigation technique (g, a, l), weighted by the % (p) collected in the SPM 2009 questionnaire of the holding for each applied technique. This average is then adjusted too by the central value of the efficiency coefficients k (e). It allows to assign to each crop of the holding its effective water consumption and , through aggregation , the total water consumption to be imputed to the holding.

7. Estimation and Calibration

All steps described above have been already completed at the time of drafting the paper. Next step under development is the process of assignment to the sampling units (holdings) of the effective sample the design weights referred to the census frame, corrected for non-response. The EMP2009 sample follows an stratified sampling design with strata defined by the technical-economic typology and a measure of size of the holdings. Subsequently, the census area under different crops, for a set of crop aggregations by region, are used as calibration marginals.

Once available EMP2009 weights it will be possible to estimate actual consumption for particular crop aggregates (cereals, pastures, olive, vineyards, other fruits) for which external sources on water consumption are regionally available. The main one is the Survey of Water Use in the Agricultural Sector (SWUA) carried out by INE every year. The survey frame consists of lists of Irrigation Entities (IE) within a river basin. IE are professional associations of long tradition in Spain, which provide data of water supply to associated holdings for the above mentioned aggregates. Primary data are expanded to total irrigated land estimated by MARM through its annual area survey.

The availability of SWUA 2009 facilitates the contrast with the estimates of effective water consumption obtained from the model. As a result of the analysis within each region (NUT II) a decision will be made on whether or not the calibration of weights in order to estimate aggregates of water consumption for specific subpopulations of holdings would be necessary.

The decision process will take into account the results of the sensitivity analysis performed making the values of the coefficients K (g, a, l, e) dealt with in (5.2) vary along its intervals.

As part of the analysis prior to making decisions on the final calibration, comparisons can also be made to other external sources, such as the 'Irrigation recommendations of the National Hydrological Plan (NHP) 2008. These recommendations are linked to objectives of official policies on saving water resources, especially in the Spanish regions most exposed to negative rainfalls trends.

Table 2 follows, showing for region C sample mean among municipalities of the real values of the theoretical water needs for three of the most important crops in the region, as well as the external 'administrative' estimates (recommendations) NHP 2008.

The model estimates fall systematically near the top of NHP data. Given that they are sample data, referred to a different year (the annual cycles of rainfall are showing a great variability in the last decade), and that region C is one with scarce water resources (thus water use savings are officially strongly recommended in PHN), it is considered that estimates of the theoretical water needs in the region used as example here –as in all other regions where the model has been already implemented- behave as expected.

Table 2. H crop*Municipality vs External source (NHP 2008) in region C 2009 (m3/Ha)

3506 CITRUS	3590 GRAPES QUALITY WINES	3339 VEGETABLES UNDER COVER	3066 CORN	Municipality
6571,60256	4637,46897	5593,81167	6451,89416	Mean
609,157592	317,398533	500,03625	235,519498	Standard Dev.
4950-6100	2850-3800	2750-5500	4300-6250	NHP m3/Ha

REFERENCES

Regulation (EC) Nº 1166/2008 of the European Parliament and of the Council of 19 November 2008 on farm structure and the survey on agricultural production methods.

Carcelén , Vicente. (1999) “Estimations of the water demand in the irrigation planning”. In: XVII National Congress of Irrigation. Murcia (Spain).

Spanish National Plan of Irrigation. Royal Decree 329/2002 , April 5th.

INE (2008). “Survey on Water Use in Agriculture “ http://www.ine.es/

ABSTRACT

During the first half of 2010, INE carried out the second phase of the Agricultural Census 2009 by collecting additional data through the Survey of Production Methods (EMP2009). One of the goals of the survey (also required by EU regulation) was to estimate at the microdata level (holding) the volume of irrigation water used during the agricultural year 2009.

Preliminary studies recommended not to include in the survey forms the quantity of water consumed as an specific question , mainly due to the risk of high measurement errors. Therefore it was decided to launch a project of model assisted estimation in several stages, combining direct observations with exogenous data from diverse administrative and statistical sources.

The document describes the different stages of the modeling process:

I) Theoretical water needs. Administrative sources provide agro-meteorological data on actual and net rainfall and evapo-transpiration for every crop within a particular area belonging to a national network of 800 irrigation areas. After several treatments, the theoretical water requirements per crop in a given month and a municipality is estimated.

II) Models of irrigation efficiency. In this stage the actual consumption of irrigation water per crop is imputed and aggregated within each holding. At this end, detailed coefficients have been provided to INE by the Ministry of Environment, Rural and Marine Affaires (MARM) according to the irrigation technique (as collected in the EMP2009) and the management efficiency of irrigation.

III) Final estimation and calibration. In this last stage external data from the annual environmental Survey on Water Use (carried out by INE) are used for control and/or calibration of the sample design weights.