International Environmental Modelling and Software Society (iEMSs)
7th Intl. Congress on Env. Modelling and Software, San Diego, CA, USA,
Daniel P. Ames, Nigel W.T. Quinn and Andrea E. Rizzoli (Eds.)
http://www.iemss.org/society/index.php/iemss-2014-proceedings

The Value of Seasonal Productivity Forecasting in Biodiesel Plans

G. Fiorese[a], G. Guariso[b]

[a]*European Commission, Joint Research Centre, Institute for Environment and Sustainability, Via E. Fermi 2749, I-21027 Ispra (VA), Italy, giulia.fiorese@jrc.ec.europa.eu*
[b]*Dept. of Electronics Information and Bioengineering, Politecnico di Milano, Italy, giorgio.guariso@polimi.it*

Abstract: Crop productivity is commonly assumed as a deterministic function when developing agricultural plans. Actual data prove however that, even for the same soil at the same location, crop productivity can be better interpreted as a random variable due to the meteorological conditions of the specific year. For the production of biodiesel, crops are easily substitutable and the farmer can chose every year between various alternatives. Without information on the seasonal meteorology, the farmers select the crop to cultivate mainly on the basis of the expected productivity. However, changes in the meteorological situation may result in a reduction in crop profitability. As a result, a crop, that on average is less interesting, may become the best choice in a specific year. Given that seasonal forecasts based on long range climatic variables, such as ENSO, are becoming available, the paper examines their effectiveness in biodiesel production plans, with reference to an area in Mato Grosso, Brazil. We formulate and solve a mathematical programming problem to determine the most efficient crop plan under different scenarios: (i) no information about the seasonal meteorology, (ii) perfect information and (iii) meteorological forecasts with different precision. This allows us to quantitatively evaluate how important the availability of seasonal productivity forecasting might be and shows that even a rough seasonal forecast, if systematically applied, may improve the average production and reduce the risk of traditional agricultural decisions.

Keywords: biodiesel; agricultural planning; mathematical programming.

1. INTRODUCTION

When developing agricultural plans, the productivity of each crop is commonly assumed to be a deterministic function of known variables (soil characteristics, local climate, farming practices,…) (see for instance, Dury et al., 2012 for an extended review). Actual data prove however that, even for the same soil at the same location, crop productivity is highly unstable. It can be better interpreted as the sum of a long term trend plus a random variation due to the meteorological conditions of the specific year (Chambers, 2013). In most areas of the world, the past trend has been positive due to a progressive expansion of mechanization, improved agricultural practices, and the availability of more productive varieties.

For the production of biodiesel, crops are easily substitutable and farmers can chose every year between various alternatives. Decision support systems (e.g. Yield Prophet, 2014, in use in Australia), may help them take into account various factors, but basically, without information on the seasonal meteorology, their choice is guided by the highest profitability (see again Dury et al., 2012), which in turn is mainly related to the expected productivity. However, changes in the seasonal meteorological situation may result in an inversion of crop profitability. As a result, for instance a crop, that on average is less productive, may become the best choice in a specific year.

Recent developments of meteorological forecasting based on long term phenomena such as El Niño Southern Oscillation (ENSO) (e.g. Solis and Latson, 2013) provide useful information for the agricultural decisions with a reasonable reliability. In Brazil, for instance, the meteorological institutes of some coastal states already issue regular seasonal forecasting. In Europe, Meteo France publishes a forecast for the next three months, while the AGRI4CAST system (JRC, 2014) monitors crop vegetation growth, including the short-term effects of meteorological events on crop production, and provides seasonal yield forecasts of key European crops. The systematic adoption of these forecasts is hindered by agricultural traditions and risk aversion, but also by the difficulty of forwarding the relevant information to farmers and the lack of the spatial and temporal details that would be needed to support agricultural activities at local scale (Hansen et al., 2006; Calanca et al., 2011). Several studies have been recently published on the topic: for instance, Dorneles Machado and Porfírio da Rocha (2011) discuss the accuracy of a seasonal forecast model in Brazil, while Kumar (2010) surveys some of the problem of practical adoption of seasonal forecast in general.

We propose a method to measure the potential impact of seasonal forecasting on a regional biodiesel exploitation plan. First, we formulate a linear programming model with the objective of maximizing the net energy production, i.e. the potential energy from biodiesel, minus the energy necessary to grow, transport and process the feedstock. Second, we model different scenarios with an increasing accuracy (and adoption) of seasonal weather forecast: (i) no information about the seasonal meteorology, (ii) perfect information and (iii) meteorological forecasts with different precision. Specifically, the availability of seasonal forecasts is simulated by adding a white noise with a given variance to the value of crop productivity. Finally, we analyze the impact of forecasting the productivity on the problem solution. This method allows to quantitatively evaluate the importance of seasonal productivity forecasting. We show that even a rough seasonal forecast, if systematically applied, may improve the average production and reduce the risk of traditional agricultural decisions. The focus is on an area in Mato Grosso, Brazil.

2. THE STUDY AREA

The study area lies within a region with very specific climatic and ecological characteristics called Cerrado in southeastern Mato Grosso. The area extends for 15,031 km^2 and includes four municipalities: Rondonopolis, Pedra Preta, Alto Garças and Alto Araguaia.

In 2011, the cultivation of soybean was predominant in the region, exceeding 60% of the cultivated area. Other common crops were cotton (18% of the surface), corn (13%) and sorghum (3%) (IBGE, 2013). The tropical savannah of the Cerrado is rich in plant and animal biodiversity and is currently threatened by the progressive expansion of extensive soybean plantations. Further growth of the agricultural areas should thus be avoided (Janssen and Rutz, 2011). There are two large biodiesel production plants, one in Rondonopolis with capacity of 245 thousand m^3 per year and one in Goias (118,000 m^3) plus other small plants for an additional 13 thousand m^3.

The Brazilian Government is fostering the development of biodiesel through a specific National Program for Production and Use of Biodiesel (PNPB), approved in 2005. It called for an increase of biodiesel in the car fuel mix and is also aimed at decreasing GHG emissions, at promoting the use of renewable sources of energies and at fostering social development.

3. THE VARIABILITY OF CROP PRODUCTIVITY

In this paper we briefly describe the premises of the allocation problem, and refer the reader to a previous work (Baglivi et al., 2014) for details and numerical values. A set of oleaginous crops that can be successfully cultivated in the study area was selected. Specifically, these crops are: soybeans, cotton, castor, peanuts and sunflower. Each of these crops has specific needs in terms of temperature and precipitation patterns. A suitability map was derived for each crop based on local climatic, geographic and pedological characteristics, using GIS data and software, following the approach described in Fiorese and Guariso (2010). The suitable area was then divided into a number of cells, that are small enough to be considered having the same characteristics. These cells are the basis of the linear programming model, described in Section 4. In our formulation, square cells with size equal to 2,500 ha each have been adopted and resulted in a total of 105 cells available to the allocation of the crops. In previous works (e.g. Fisher et al., 2010; Cesar Leão et al. 2011), productivity was always assumed to vary with land suitability and with climatic variables that describe the past temperature and precipitation patterns. Here a meteorological variable for the current season is also introduced. Therefore, the decisions about the cultivation of all the considered crops are modelled under several scenarios assuming different degrees of knowledge of future seasonal weather conditions.

We first analyse actual productivity data of the candidate oleaginous crops in the study area. Yields are subject to significant variations over the years that can be explained by the changes in agricultural practices, the specific weather of a particular year and several other factors.

Figure 1 shows how the yield of soybean and peanuts has changed from 1990 up to recent years in Brazil and in Mato Grosso (IBGE, 2013). It is possible to observe both an increasing trend over the timespan and the sensible annual oscillations which are related to that specific year climate. Similar behaviors can be observed for other crops (see Table 1).

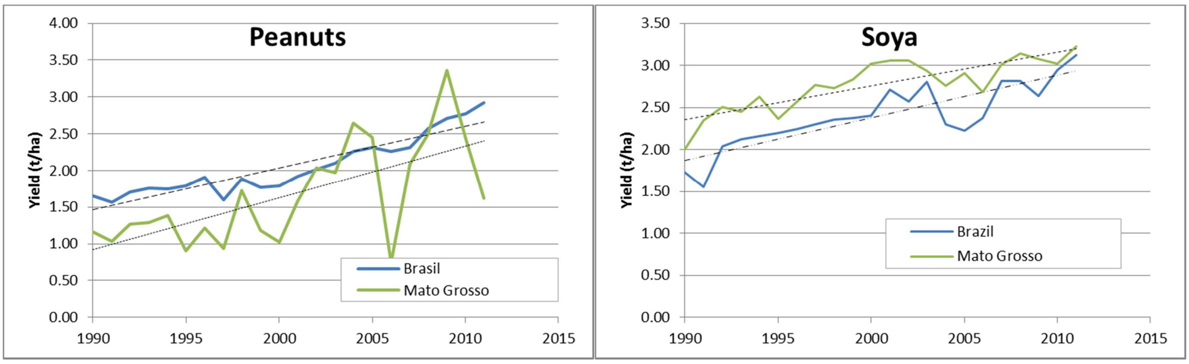

Figure 1. Yield of soybeans and peanuts in Brazil and in Mato Grosso in the period 1990-2011 (IBGE, 2013). Dashed lines are the respective trends.

Table 1. Yields and standard deviations for the selected crops (years 1990-2011; IBGE, 2013).

	Avg yield (t/ha)	Standard deviation of data	Standard deviation (after removing the trend)
Soybeans	2.78	0.31	0.16
Castor	1.08	0.39	0.38
Sunflower*	1.33	0.17	0.13
Peanuts	1.66	0.69	0.52
Cotton seeds	3.68	1.16	1.08

* data are available only for 2005-2011

For the purpose of our analysis, we consider only annual variations of the yield after removing the trends from the time series (Table 1). With a linear interpolation, we estimated the residuals (i.e. the differences between actual values and values predicted by the linear trend) of each crop. A Kolmogorov-Smirnov goodness-of-fit test implemented in MATLAB proved that all the considered time series of the residuals have a normal distribution at the standard 5% significance level. We can thus conclude that the productivity of the considered oleaginous crops can be described by an increasing trend plus a random stochastic variable with zero mean and a Gaussian probability distribution.

4. PROBLEM FORMULATION

Assuming the productivity of each cell for the cropping season to come is known, the optimal energy production plan can be formulated as a (linear) mathematical programming problem aiming to maximize the net energy output from the system. The decision variables z_{ijs} are the fractions of area cultivated in cell i, with crop s and shipped to plant j for processing. Considering 105 cells times five possible crops times two existing plants (small ones have been accounted for with the largest one, since they are very close) makes a total of 1050 decision variables. The maximization of the net energy output can be written as (Fiorese et al., 2013):

$$\max_{\{z\}} J = \sum_{i=1}^{N_c}\sum_{j=1}^{N_p}\sum_{s=1}^{N_s} A\left[E_b w_s p_s s_{is} z_{ijs} - e_{tr} p_s s_{is} d_{ij} z_{ijs} - (e_s^{agr} + e_s^{pr} p_s s_{is}) z_{ijs}\right] \quad (1)$$

where the first term represents the yearly energy output (in terms of GJ of biodiesel at the conversion plants), the second is the energy spent to transport the feedstock to the plants and the third is the energy employed for the cultivation and the conversion process. More precisely:

N_c, N_p, and N_s represent, respectively, the number of cells, of plants, and of crops considered;
E_b is the energy content of a unit mass of biodiesel;
p_s is the crop productivity in terms of mass of seeds per unit area per year;
s_{is} represents the land suitability of cell i for crop s: following the approach suggested in Fischer et al. (2010), it is equal to 1 for very suitable conditions or to 0.8 for less suitable (but still acceptable) ones;
w_s is the amount of biodiesel (kg) that can be extracted from a unit weight of seeds type s. This parameter depends on the oil content of the seeds and on the efficiency of the extraction operations;
A is the land surface (ha) available in each cell which is constant since we have assumed a regular grid.

As to the second term:
e_{tr} is the energy necessary to transport a unit of biomass over a unit distance;
d_{ij} is the distance between cell i and plant j. Since there is only one main road crossing the region, connecting the plant at the western border with that on the eastern side, the distance has been computed using GIS tools (the geometric distance from the cell center to the existing roads is summed to the distance along the road to the plant).

Finally, the last term represents the energy costs of biodiesel production. This term is given by the sum of the energy necessary for the cultivation and for the conversion into biodiesel, so:
e_s^{agr} is the energy of all agricultural operations to grow a unit area of crop s;

e_s^{pr} is the energy needed to process a unit weight of seeds of crop s to extract biodiesel.

The constraints of the problem represent:
The use of land in each cell cannot exceed the land available

$$\sum_{j=1}^{N_p}\sum_{s=1}^{N_s} z_{ijs} \le 1 \qquad \forall i \quad (2)$$

The biomass shipped to each plant cannot exceed the plant capacity C_j

$$\sum_{i=1}^{N_c}\sum_{s=1}^{N_s} p_s s_{is} z_{ijs} \le C_j \qquad \forall j \quad (3)$$

The non-negativity of decision variables

$$z_{ijs} \ge 0 \qquad \forall i,j,s \quad (4)$$

Cotton requires some additional consideration, since it is already grown in the area as a supply for the textile industry. The net balance of energy production of oil from cottonseeds is negative because its energy output, derived from oil, is less than the energy used for the cultivation and processing of its seeds. However, oil seeds are a by-product of the production chain of textile fibers: this means that the net energy balance should not include the (high) energy spent for the production of seeds and the overall balance of biodiesel production may return a positive value. The positivity and convenience of the production still depend on the energy spent for processing and transportation. Assuming a very conservative approach, this is computed as if the cottonseeds should be transported from the fields to the plant, while they would be probably already available at some textile plants. If cotton seeds are to be considered as a “free” by-product of other activities, their area must be limited at the current level.

The solution of problem (1-4) above gives the optimal value of the decision variables z_{ijs}^o, which mean what to grow in each parcel and where to process the seeds, that corresponds to a net energy output

$J°$. However, this will not coincide to the real production, since this will be determined by the actual productivity $\underline{p}_s$, determined by the specific meteorological condition and possibly quite different, as seen before (Figures 1), from the expected one. This means that the actual energy output J will be a value:

$$J = \sum_{i=1}^{N_c}\sum_{j=1}^{N_p}\sum_{s=1}^{N_s} A\left[E_b w_s \underline{p}_s s_{is} z_{ijs}^o - e_{tr} \underline{p}_s s_{is} d_{ij} z_{ijs}^o - (e_s^{agr} + e_s^{pr} \underline{p}_s s_{is}) z_{ijs}^o\right] \quad (5)$$

Clearly, the more the forecasted productivity p_s will approach the actual one $\underline{p}_s$, the more J will approach $J°$. Using the method described above, it is possible to measure the benefit that can be obtained by a systematic use of a seasonal forecast as a function of the quality of the forecast itself.

Since actual experience in using seasonal weather forecasts is limited, we simulate their availability by adding a white noise with a given variance to the value of productivity. In case a forecasting model would indeed be developed, it should in fact be unbiased (the average error must be zero) and the residual (i.e. the difference between actual and forecasted values) must be a white noise, which means it should have a Gaussian distribution. By sorting a random number and inverting the Gaussian distribution, we "simulate" an equivalent forecast and, by varying the assumed variance, we can regulate its precision. These values can then be examined according with the standard indexes used to evaluate model performances (see Bennett et al., 2013). As an example, in the following, the precision of the forecast is evaluated in terms of correlation coefficient, which, as it is well known, is also related to the variance explained by the model.

5. RESULTS AND DISCUSSION

Results for the selected area of Mato Grosso are summarized in Figure 2, where the net energy production is plotted as a function of the actual-predicted correlation of the hypothesized forecasting model. It clearly appears that, as soon as the forecasting model reaches a sufficient precision (correlation larger than 0.5), there is an almost linear increase of the production that may reach about 25% when predictive performances are very good (correlation equal to 0.9). It is interesting to note that the presence of the forecasting model also reduces the standard deviation of the production distribution (represented by the dashed lines in Figure 2). When taking decisions only looking at the average productivity, the standard deviation of the production values is entirely due to the natural variation illustrated above, while it reduces to about half (when compared with the average production value) in case of perfect prediction.

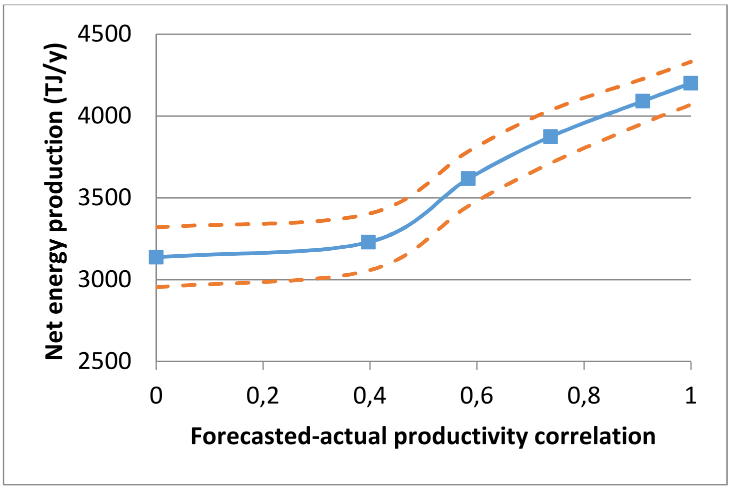

Figure 2. Average and ±standard deviation of net energy production as a function of forecasting accuracy.

Another interesting comparison is shown in Table 2, which illustrates the frequency with which the energy production may drop below that with no forecast, due to possible errors in the forecasting model. The table shows the percentage of years in which the energy production is less than that can be obtained without any forecast as a function of the model performance (measured by the forecasted – actual productivity correlation). For instance, using only mean values (i.e. without any model), obviously 50% of the time the production is less than the average, 16% of the time there is a loss of more than 5%, and rarely (1% of the time) a loss of more than 15%.

Table 2. Risk of producing less energy than without forecast as a function of forecasting accuracy.

	Forecast model performance		
Decrease of energy production with respect to no forecast	*0 (mean values)*	*0.4*	*0.6*
0%	50	26.5	0
-5%	16	7.5	0
-15%	1	0	0
-25%	0	0	0

It clearly appears that even a poorly performing forecast model can strongly reduce the risk of low production, which, according to Meza et al. (2008), constitutes one of the most critical points in rain-fed and not highly industrialized agriculture. For instance, even with a forecasted-actual productivity correlation of 0.6, the risk of performing worse than without forecasting is negligible (see Table 2). Indeed, a sequence, and sometimes even a single, low year may disrupt the economic balance of farmers at the point that they may even be unable to exploit better performances in the following years. A more complex model would however be needed to take this longer planning horizon into account.

The adoption of forecasting models of different accuracy results in different land allocation between crops, which is the decision that farmers have to take every year. If such a decision is based on the average productivity, which approximately means the average income, only the most productive crops are used. This means for instance that soybean is completely abandoned. On the contrary, if information about actual productivity is available in advance, sometimes very productive crops such as peanuts must be suspended since they are more sensitive to environmental conditions. With adverse meteorology and perfect prediction, soybean surface may be again 2-3% of the total. Additionally, castor, that without forecast should use 15% of the surface, may grow up to 26% in some years, when perfect forecast is assumed. Finally, cotton uses 15% of the area in all conditions, since its seeds are a byproduct of the textile industry and thus by far the most convenient choice, with all the meteo conditions assumed in this study.

6. CONCLUSION

The approach proposed in this paper tries to explicitly account for the variability of agricultural productivity due to meteorological conditions. This is often assumed to be constant, in some cases also with the products and input prices, which, on the contrary, are also subject to relevant fluctuations due to the evolution of the economy.

The formulation and solution of few hundreds mathematical programming problems allows to quantitatively evaluate the improvement that the systematic use of forecasts of given accuracy may have in reducing the effects of meteorological variation in agriculture. This should be compared with the effort that local governments and agricultural agencies should deploy to reach that level of seasonal forecast performances.

As underlined by many authors, however, the problem in not purely technical and several limitations exist in the practical adoption of these forecasts: from how the pertinent information is distributed, to the obvious tendency of farmers to be strongly risk-averse, to the distrust in adopting new, and at times unreliable, technologies. From a theoretical viewpoint, a relevant change in the type of results presented above could be the introduction of new crops such as jatropha (*Jatropha curcas*) or Chinese pistache (*Pistacia chinensis*) that are permanent and thus occupy the land for a number of consecutive years.

7. ACKNOWLEDGMENTS

This work was partially supported by LEAP (Laboratorio Energia Ambiente Piacenza). The authors gratefully acknowledge the assistance of A. Baglivi and C. Uggè, who contributed to the development of an initial study on the area.

8. REFERENCES

Baglivi, A., Fiorese, G., Guariso, G., Uggè, C., 2014 in press. Energy and GHG emissions assessment of biodiesel production in Mato Grosso, Brazil. In: Ajay Kumar Bhardwaj, Terenzio Zenone, Jiquan Chen (Eds.), Sustainable Biofuels - An Ecological Assessment of Future Energy, De Gruyter.

Bennett, N.D., Croke, B.F.W., Guariso, G., Guillaume, J.H.A., Hamilton, S.H., Jakeman, A.J. et al., 2013. Characterizing performances of environmental models. Environ Modell Softw 40, 1–20.

Calanca, P., Bolius, D., Weigel, A.P., Liniger, M. A., 2011. Application of long-range weather forecasts to agricultural decision problems in Europe, J Agric Sc 149, 15–22.

Cesar Leão, R.R.C., Silvio Hamacher, S., Fabrício Oliveira, F., 2011.Optimization of biodiesel supply chains based on small farmers: A case study in Brazil, Biores. Tech. 102, 8958–8963.

Chambers, R.G., 2013. The Sources of Measured Agricultural Productivity Growth, International Agricultural Trade Research Consortium's (IATRC's) 2013 Symposium: Productivity and Its Impacts on Global Trade, June 2-4, Seville, Spain (available from http:// http://ageconsearch.umn.edu/).

Dury, J., Schaller, N., Garcia, F., Reynaud, A., Bergez, J. E., 2012. Models to support cropping plan and crop rotation decisions. A review. Agr. Sust. Develop. 32 (2), 567-580.

Dorneles Machado, R.,Porfírio da Rocha, R., 2011. Seasonal climatic forecast over Brazil: evaluation of the RegCM3 model nested to the CPTEC/COLA global model. Rev. Bras. Meteorol. 26(1), 121-136.

Fiorese, G., Gatto, M., Guariso, G., 2013. Optimisation of combustion bioenergy in a farming district under different localisation strategies. Biomass Bioener. 58, 20 – 30.

Fiorese, G., Guariso, G., 2010. A GIS-based approach to evaluate biomass potential from energy crops at regional scale. Environ. Model. Softw. 25(6), 702–711.

Fischer, G., Prieler, S., van Velthuizen, H., Lensink, S.M., Londo, M., de Wit, M., 2010. Biofuel production potentials in Europe: Sustainable use of cultivated land and pastures. Part I: Land productivity potentials. Biomass Bioenerg. 34(2), 159–172.

Hansen, J.W., Challinor, A., Inses, A., Wheeler, T., Moron, V., 2006. Translating climate forecasts into agricultural terms: advances and challenges, Clim Res 33, 27-41.

IBGE - Instituto Brasileiro de Geografia e Estatística, 2013. Sistema IBGE de Recuparação automática – SIDRA. http://www.sidra.ibge.gov.br/ (last accessed 15.04.13).

Janssen, R., Rutz, D.D., 2011. Sustainability of biofuels in Latin America: Risks and opportunities. Energy Policy 39, 5717–5725.

JRC – Joint Research Center, AGRI4CAST Action. http://mars.jrc.ec.europa.eu/mars/About-us/AGRI4CAST (last accessed 11.03.2014).

Kumar, A., 2010. On the assessment of the value of the seasonal forecast information. Meteorol Appl 17, 385–392.

Meza, J.F., Hansen, J.W., Osgood, D., 2008. Economic value of seasonal climate forecast for agriculture: Review of ex-ante assessments and recommendation for future research. JAMC 47, 1269-1286.

Solis, D., Letson, D., 2013. Assessing the value of climate information and forecasts for the agricultural sector in the Southeastern United States: multi-output stochastic frontier approach. Regional Environmental Change 13, 5-14.

Yield Prophet, 2014. http://www.yieldprophet.com.au/YP/wfLogin.aspx, visited May 2014.