Agricultural Systems 103 (2010) 316–326

Contents lists available at ScienceDirect

Agricultural Systems

journal homepage: www.elsevier.com/locate/agsy

The yield gap of global grain production: A spatial analysis

Kathleen Neumann [a,*], Peter H. Verburg [b], Elke Stehfest [c], Christoph Müller [c,d]

[a] *Land Dynamics Group, Wageningen University, P.O. Box 47, 6700 AA Wageningen, The Netherlands*
[b] *Institute for Environmental Studies, VU University Amsterdam, De Boelelaan 1087, 1081 HV Amsterdam, The Netherlands*
[c] *Netherlands Environmental Assessment Agency (PBL), P.O. Box 303, 3720 AH Bilthoven, The Netherlands*
[d] *Potsdam Institute for Climate Impact Research (PIK), Telegrafenberg, P.O. Box 601203, 14412 Potsdam, Germany*

ARTICLE INFO

Article history:
Received 14 April 2009
Received in revised form 29 January 2010
Accepted 22 February 2010
Available online 26 March 2010

Keywords:
Grain production
Yield gap
Land management
Intensification
Inefficiency
Frontier analysis

ABSTRACT

Global grain production has increased dramatically during the past 50 years, mainly as a consequence of intensified land management and introduction of new technologies. For the future, a strong increase in grain demand is expected, which may be fulfilled by further agricultural intensification rather than expansion of agricultural area. Little is known, however, about the global potential for intensification and its constraints. In the presented study, we analyze to what extent the available spatially explicit global biophysical and land management-related data are able to explain the yield gap of global grain production. We combined an econometric approach with spatial analysis to explore the maximum attainable yield, yield gap, and efficiencies of wheat, maize, and rice production. Results show that the actual grain yield in some regions is already approximating its maximum possible yields while other regions show large yield gaps and therefore tentative larger potential for intensification. Differences in grain production efficiencies are significantly correlated with irrigation, accessibility, market influence, agricultural labor, and slope. Results of regional analysis show, however, that the individual contribution of these factors to explaining production efficiencies strongly varies between world-regions.

1. Introduction

Human diets strongly rely on wheat (*Triticum aestivum* L.), maize (*Zea mays* L.), and rice (*Oryza sativa* L.). Their production has increased dramatically during the past 50 years, partly due to area extension and new varieties but mainly as a consequence of intensified land management and introduction of new technologies (Cassman, 1999; Wood et al., 2000; FAO, 2002a; Foley et al., 2005). For the future, a continuous strong increase in the demand for agricultural products is expected (Rosegrant and Cline, 2003). It is highly unlikely that this increasing demand will be satisfied by area expansion because productive land is scarce and also increasingly demanded by non-agricultural uses (Rosegrant et al., 2001; DeFries et al., 2004). The role of agricultural intensification as key to increasing actual crop yields and food supply has been discussed in several studies (Ruttan, 2002; Tilman et al., 2002; Barbier, 2003; Keys and McConnell, 2005). However, in many regions, increases in grain yields have been declining (Cassman, 1999; Rosegrant and Cline, 2003; Trostle, 2008). Inefficient management of agricultural land may cause deviations of actual from potential crop yields: the yield gap. At the global scale little information is available on the spatial distribution of agricultural yield gaps and the potential for agricultural intensification. There are three main reasons for this lack of information.

First of all, little consistent information of the drivers of agricultural intensification is available at the global scale. Keys and McConnell (2005) have analyzed 91 published studies of intensification of agriculture in the tropics to identify factors important for agricultural intensification. They emphasize that a plentitude of factors drive changes in agricultural systems. The relative contribution of them varies greatly between regions. This problem was confirmed by a number of studies that have investigated grain yields, and tried to identify factors that either support or hamper grain production at different scales (Kaufmann and Snell, 1997; Timsina and Connor, 2001; FAO, 2002a; Reidsma et al., 2007). These studies also indicate that most of these factors are locally or regionally specific, which makes it difficult to derive a generalized set of factors that apply to all countries. A second reason for the absence of reliable information on the global yield gap is the limited availability of consistent data at the global scale. Especially land management data are lacking. When it comes to quantifying potential changes in crop yields often only biophysical factors, such as climate are considered while constraints for increasing actual crop yields are often neglected or captured by a simple management factor that is supposed to include all factors that cause

* Corresponding author. Tel.: +31 317 482430; fax: +31 317 419000.
E-mail addresses: kathleen.neumann@wur.nl (K. Neumann), Peter.Verburg@ivm.vu.nl (P.H. Verburg), elke.stehfest@pbl.nl (E. Stehfest), christoph.mueller@pik-potsdam.de (C. Müller).

0308-521X/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.agsy.2010.02.004

a deviation from potential yields (Alcamo et al., 1998; Harris and Kennedy, 1999; Ewert et al., 2005; Long et al., 2006). Finally, lack of data also leads to another difficulty. Many yield gap analyses have in common that they apply crop models for simulating potential crop yields which are compared to actual yields (Casanova et al., 1999; Rockstroem and Falkenmark, 2000; van Ittersum et al., 2003). Potential yields, however, are a concept describing crop yields in absence of any limitations. This concept requires assumptions on crop varieties and cropping periods. While such information is easily attainable at the field scale it is not available at the global scale. Moreover, different simplifications of crop growth processes exist between the models. This may result in uncertainties of globally simulated potential yields, and makes an appropriate model calibration essential for global applications. Comparing simulated global crop yields to actual yields therefore bears the risk of dealing with error ranges and uncertainties of different data sources (i.e., observations and simulation results) which might even outrange the yield gap itself.

Consequently, available knowledge about the yield gap is rather inconsistent and regional and global levels of agricultural production have hardly been studied together.

The aim of this paper is to overcome some of the mentioned shortcomings by analyzing actual yields of wheat, maize, and rice production at both regional and global scale accounting for biophysical and land management-related factors. We propose a methodology to explain the spatial variation of the potential for intensification and identifying the nature of the constraints for further intensification. We estimated a stochastic frontier production function to calculate global datasets of maximum attainable grain yields, yield gaps, and efficiencies of grain production at a spatial resolution of 5 arc min (approximately 9.2 × 9.2 km on the equator). Applying a stochastic frontier production function facilitates estimating the yield gap based on the actual grain yield data only, instead of using actual and potential grain yield data from different sources. Therefore, the method allows for a robust and consistent analysis of the yield gap. The factors determining the yield gap are quantified at both global and regional scales.

2. Methodology

2.1. The stochastic frontier production function

Stochastic frontier production functions originate from economics where they were developed for calculating efficiencies of firms (Aigner et al., 1977; Meeusen and Broeck, 1977). Since agricultural farms are a special form of economic units this econometric methodology can also be used to calculate farm efficiencies and efficiencies of agricultural production in particular. In our global analysis, the agricultural production within one grid cell (5 arc min resolution) is considered as one uniform economic unit. The stochastic frontier production function represents the maximum attainable output for a given set of inputs. Hence, it describes the relationship between inputs and outputs. The frontier production function is thus "a regression that is fit with the recognition of the theoretical constraint that all actual productions lie below it" (Pesaran and Schmidt, 1999). In case of agricultural production the frontier function represents the highest observed yield for the specified inputs. Inefficiency of production causes the actual observations to lie below the frontier production function. The stochastic frontier accounts for statistical noise caused by data errors, data uncertainties, and incomplete specification of functions. Hence, observed deviations from the frontier production function are not necessarily caused by the inefficiency alone but may also be caused by statistical noise (Coelli et al., 2005).

The frontier production function to be estimated is a Cobb-Douglas function as proposed by Coelli et al. (2005). Cobb-Douglas functions are extensively used in agricultural production studies to explain returns to scale (Bravo-Ureta and Pinheiro, 1993; Bravo-Ureta and Evenson, 1994; Battese and Coelli, 1995; Reidsma et al., 2009b). If the output increases by the same proportional change in input then returns to scale are constant. If output increases by less than the proportional change in input the returns decrease. The main advantage of Cobb-Douglas functions is that returns to scale can be increasing, decreasing or constant, depending of the sum of its exponent terms. In agricultural production decreasing returns to scale are common. The Cobb-Douglas function is specified as following:

$$\ln(q_i) = \beta_1 x_i + v_i - u_i \tag{1}$$

where $\ln(q_i)$ is the logarithm of the production of the ith grid cell ($i = 1, 2, \ldots, N$), x_i is a $(1 \times k)$ vector of the logarithm of the production inputs associated with the ith grid cell, β is a $(k \times 1)$ vector of unknown parameters to be estimated and v_i is a random (i.e., stochastic) error to account for statistical noise. Statistical noise is an inherit property of the data used in our study resulting from reporting errors and inconsistencies in reporting systems. The error can be positive or negative with a mean zero. The non-negative variable u_i represents inefficiency effects of production and is independent of v_i. Fig. 1 illustrates the frontier production function.

Stochastic frontier analyses are widely used for calculating efficiencies of firms and production systems. The most common measure of efficiency is the ratio of the observed output to the corresponding frontier output (Coelli et al., 2005):

$$E_i = \frac{q_i}{\exp(x_i'\beta + v_i)} = \frac{\exp(x_i'\beta + v_i - u_i)}{\exp(x_i'\beta + v_i)} = \exp(-u_i) \tag{2}$$

where E_i is the efficiency in the ith grid cell. The efficiency is an index without a unit of measurement. The observed output at the ith grid cell is represented by q_i while $x_i'\beta$ is the frontier output. The efficiency E_i determines the output of the ith grid cell relative to the output that could be produced if production would be fully efficient given the same input and production conditions. The efficiency ranges between zero (no efficiency) and one (fully efficient).

Kudaligama and Yanagida (2000) applied stochastic frontier production functions to study inter-country agricultural yield differences at the global scale. However, that study disregards spatial variability within countries, which can be very large. To our knowl-

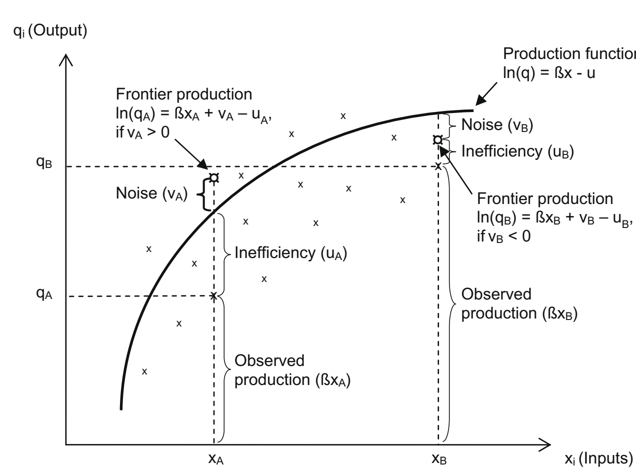

Fig. 1. The stochastic production frontier (after Coelli et al., 2005). Observed productions are indicated with × while frontier productions are indicated with ¤. The frontier function is based on the highest observed outputs under the inputs accounting for random noise (v_i). Further deviations of the observations are due to inefficiencies (u_i). The frontier production q_i can lie above or below the frontier production function, depending on the noise effect (v_i).