

# Interstate Comparison of Environmental Performance using Stochastic Frontier Analysis: The United States Case Study

Alexander Y. Vaninsky

**Abstract**—Environmental performance of the U.S. States is investigated for the period of 1990 – 2007 using Stochastic Frontier Analysis (SFA). The SFA accounts for both efficiency measure and stochastic noise affecting a frontier. The frontier is formed using indicators of GDP, energy consumption, population, and CO<sub>2</sub> emissions. For comparability, all indicators are expressed as ratios to total. Statistical information of the Energy Information Agency of the United States is used. Obtained results reveal the bell - shaped dynamics of environmental efficiency scores. The average efficiency scores rise from 97.6% in 1990 to 99.6% in 1999, and then fall to 98.4% in 2007. The main factor is insufficient decrease in the rate of growth of CO<sub>2</sub> emissions with regards to the growth of GDP, population and energy consumption. Data for 2008 following the research period allow for an assumption that the environmental performance of the U.S. States has improved in the last years.

**Keywords**—Stochastic Frontier Analysis, Environmental Performance, Interstate Comparisons.

## I. INTRODUCTION

WITH environmental issues being among major concerns of contemporary society, decrease in the emissions of greenhouse gases is amid main issues of the Department of Energy of the United States (DOE) activity. Report [1] issued by the Energy Information Administration (EIA) states that total U.S. greenhouse gas emissions in 2008 were 2.2% below the 2007 total, declining from 7,209.8 million metric tons carbon dioxide equivalent (MMTCO<sub>2</sub>e) in 2007 to 7,052.6 MMTCO<sub>2</sub>e in 2008. The drop was largely the result of the decrease in carbon dioxide (CO<sub>2</sub>) emissions. Even in the presence of small percentage increases in emissions of other greenhouse gases, their absolute contributions to the change in total emissions were relatively small and were more than offset by the drop in CO<sub>2</sub> emissions. The decrease in U.S. CO<sub>2</sub> emissions in 2008 resulted primarily from economic contraction, lower demand for electricity along with lower carbon intensity of electricity supply, and higher energy prices. In this section below we follow [1] to describe the situation in more details and to stress the necessity of mathematical modeling of environmental performance aimed at its further improvement. We use statistical data provided by the EIA available on the web site <http://www.eia.doe.gov> and focus on the States of the U.S., the main level of governmental environmental legislation.

Carbon dioxide (CO<sub>2</sub>) is the most abundant human-caused greenhouse gas in the atmosphere. It constituted 82.7% of the total energy-related total U.S. greenhouse gas emissions in 2008. Petroleum is the largest fossil fuel source constituting 42% of total. Coal is the second-largest fossil fuel contributor at 37%. Natural gas has a carbon intensity of 55% of that for coal and 75% of the carbon intensity for petroleum. In 2008 it accounted for 28.5% of U.S. fossil energy use but only 21% of total energy-related CO<sub>2</sub> emissions. Together with the observation that the U.S. can provide about 70% of its demand in natural gas on its own, relatively low carbon intensity of natural gas allows for its consideration as primary source of energy from fossil fuels in future.

The largest source of all energy-related CO<sub>2</sub> emissions is the electric power sector. Its share in total is 40.6%. The transportation sector is the second-largest source, at 33.1% of the total. Combustion of motor gasoline, diesel fuel, and jet fuel are main sources of CO<sub>2</sub> emissions in this sector. In the residential and commercial sectors, the main sources of CO<sub>2</sub> emissions are heating processes that involve direct fuel use. These sectors account for 26.3% of total CO<sub>2</sub> emissions.

As stated in [1], the greenhouse gas intensity of the U.S. economy—measured as metric tons carbon dioxide equivalent (MTCO<sub>2</sub>e) emitted per million dollars of real gross domestic product (GDP)—fell by 2.6% from 2007 to 2008. It states also that based on 0.4% of economic growth and 2.2% decrease in total greenhouse gas emissions, the U.S. greenhouse gas intensity decreased in the period from 2007 to 2008.

The report [1] provides economic analysis of the factors of the decrease using factorial decomposition of the resulting indicator. The factorial decomposition methodology was invented in [2] and adapted to environmental issues in [3] referred to now in literature as *Kaya identity*. Publication [4] applies the *Kaya identity* to analysis of the U.S. energy sector. This identity is based on the following factorial decomposition:

$$F = P \cdot (G/P) \cdot (E/G) \cdot (F/E) = P \cdot g \cdot e \cdot f, \quad (1)$$

where  $F$  stands for CO<sub>2</sub> emissions,  $P$  is population,  $G$  is GDP,  $E$  is primary energy consumption,  $g = G/P$  is the per-capita GDP,  $e = E/G$  is the energy intensity of GDP, and  $f = F/E$  is the carbon intensity of energy. Upper- and lowercase symbols distinguish extensive and intensive variables. For analytical purposes, this identity may be compressed to three terms only by combining the product  $e \cdot f$  into a new factor  $h$ :

$$h = e \cdot f = F/G, \quad (2)$$

carbon intensity of GDP. By doing so, the basic identity (1) takes a form

$$F = P \cdot (G/P) \cdot (F/G) = P \cdot g \cdot h. \quad (3)$$

A. Y. Vaninsky is with Hostos Community College of The City University of New York, Bronx, NY 10451 USA (phone: +1-718-319-7930; fax: +1-718-518-6607; e-mail: [avaninsky@hostos.cuny.edu](mailto:avaninsky@hostos.cuny.edu)).

The last identity is used in [1] for the following approximation:

$$\% \Delta F \approx \% \Delta P + \% \Delta g + \% \Delta h, \quad (4)$$

where combination of symbols  $\% \Delta$  represents percentage change. The last equation may be interpreted as follows, [1]:

$$\% \Delta CO_2 \approx \% \Delta GDP + \% \Delta (Energy/GDP) + \% \Delta (CO_2/Energy). \quad (5)$$

When applied to the data of 2000 – 2008, this equation reveals that energy intensity of GDP (Energy/GDP) has gone down in every year since 2000. The carbon intensity of the energy supply (CO<sub>2</sub>/Energy) has gone down in some years and up in others. The GDP growth was positive in all years from 2000 through 2008, but has varied. As a result, in 2008, economic growth was low (0.4%) coupled with decreases in both energy intensity (-2.5%) and carbon intensity of energy (-0.8%). This interplay of the factors has led to a 2.9% decline in energy-related CO<sub>2</sub> emissions from 2007 to 2008.

Publication [1] mentions findings that the six key greenhouse gases (GHGs), with CO<sub>2</sub> being the main, pose a threat to public health and welfare for current and future generations. It was stated also that GHG emissions from new motor vehicles and motor vehicle engines contribute to climate change. As the result, the U.S. Environmental Protection Agency (EPA) was authorized to request the mandatory reporting the GHGs emissions from 31 different source categories. It is expected that 80% - 85% of total U.S. GHG emissions from 10,000 facilities will be covered in the reports. Expectedly, they will allow for the development of operational and strategic measures aimed at the decrease in the negative effect of GHG emissions. The monitoring begins in January 2010 with first reports due in 2011.

The EPA drafted the Prevention of Significant Deterioration / Title V Greenhouse Gas Tailoring Rule that limits the applicability of CO<sub>2</sub> emissions standards to new and modified stationary sources only those emitting more than 25,000 MTCO<sub>2</sub>e annually. By doing so, the EPA expects provision of operating permits to about 14,000 large industrial sources, which are responsible for nearly 70 percent of U.S. GHG emissions: power plants, refineries, and other large industrial operations. The EPA and the U.S. Department of Transportation (DOT) jointly proposed new nationwide standards for corporate average fuel economy (CAFE) and GHG emissions standards for new light- and medium-duty vehicles. The proposal imposes nationwide the GHG standards sought by California.

In 2009, the EPA published a Standard (RFS2) that requires the establishment of new standards for cellulosic biofuel, biomass-based diesel, advanced biofuels, and total renewable fuel for use in transportation. The proposed standard includes definitions and criteria for both renewable fuels and the feedstocks used to produce them. It includes, in particular, the guidelines on how life-cycle emissions from each type of renewable fuel would be calculated. It proposes a method of life-cycle accounting including GHG emissions from production and transport, land use change, production, distribution, and blending of the renewable fuel.

The U.S. House of Representatives passed The American Clean Energy and Security Act of 2009 (ACESA). It includes a “cap-and-trade” program effective 2012. The Act requires that by 2020 total GHG emissions be 17% and by 2050 83%

below the 2005 levels. It also includes provisions for funding for carbon capture and sequestration projects, standards and programs designed to increase in energy efficiency, and energy efficiency standards for 2020.

At the State – level, the nation – wide initiatives of the U.S. Federal Government were supported and further developed. California imposed a fee on carbon emissions in 2009. Five auctions of CO<sub>2</sub> emissions allowances were held by 2009, with approximately 90 million allowances being sold. They generated more than \$432.8 million aimed to be used by the participating States for energy efficiency and renewable energy programs. The Western Climate Initiative (WCI) continues development of the comprehensive regional market - based cap - and-trade program that seeks to reduce emissions across participating States. Its objective is to launch the emissions trading system in 2012. Participating U.S. States include Arizona, California, Montana, New Mexico, Oregon, Utah, and Washington. The Midwestern Greenhouse Gas Reduction Accord (MGGRA) released draft recommendations for a regional cap – and - trade program. Sectors for which emission caps are proposed include electricity generation and imports, industrial combustion and process sources, transportation fuels, and residential, commercial, and industrial fuels. It is proposed that about 33% of allowances be auctioned initially, with the remainder sold for a small fee. The program would transition to full auction over time. Member States include Iowa, Illinois, Kansas, Michigan, Minnesota, and Wisconsin.

An approach suggested in this paper below allows for the estimation of the combined effect of all these undertakings. It models the objectively needed amounts of the CO<sub>2</sub> emissions (the frontier) based on GPD, energy consumption, and population of the states, and weights the frontier amounts of CO<sub>2</sub> emissions against their actual amounts as their ratio. As a result, it produces an environmental efficiency score that measures the environmental performance of the U.S. States.

The obtained results reveal the dynamics of the States’ environmental efficiency scores and allow for the evaluation of the trends in the environmental performance.

## II. METHODOLOGY

Methodology of the research in this paper is Stochastic Frontier Analysis (SFA). The SFA was developed in [5, 6] and is designed to separate actual inefficiency from the advantageous or disadvantageous impact of stochastic noise, considered an exogenous factor. The initial version of SFA was restricted to finding average efficiency in a group only. Later, publications [7, 8] revealed the opportunities for individualizing of the stochastic efficiency scores. At present, SFA is a well-established and widely used research tool. Monograph [9] provides a comprehensive review of its contemporary state.

SFA considers either production or cost function as a frontier subject to stochastic noise. It was a breakthrough for the SFA to assume that these functions pertain to frontier performance rather than describe a fully efficient mode of operation of actual objects. Deviation from the frontier is due to two factors: inefficiency and stochastic noise. The former is manageable; the latter is exogenous. Separation of the two

factors allows for more precise estimation of inefficiency, because the exogenous factor may be eliminated by using the maximum likelihood method described in [10]. In fact, this feature provides the main potential advantage of SFA over the Data Envelopment Analysis (DEA), [11]. The DEA measures efficiency by using deterministic frontier only and thus, ignores stochastic noise.

Below, we focus only on the cost function and follow the description of SFA provided by [9] and [12]. SFA considers objects producing outputs by using a vector of inputs  $\mathbf{x}$ . The best possible practice is determined by a deterministic cost frontier  $g(\boldsymbol{\beta}, \mathbf{x})$ , where  $\boldsymbol{\beta}$  represents a vector of unknown parameters. The frontier is subject to stochastic fluctuations, and the actual objects are not fully efficient in the sense that actual cost of their outputs is greater than that given by the stochastic frontier. Usually, it is assumed that the function  $g(\boldsymbol{\beta}, \mathbf{x})$  is the Cobb–Douglas function or the transcendental logarithmic (translog) function. We use the latter below. It is also typical for SFA to use data in logarithmic form. In light of these practices, the SFA cost frontier model may be presented as

$$y_{it} = g(\boldsymbol{\beta}, \mathbf{x}_{it}) \cdot e^{U_{it} + V_{it}}, \quad (6)$$

or in logarithmic form as

$$\ln y_{it} = \ln g(\boldsymbol{\beta}, \mathbf{x}_{it}) + (V_{it} + U_{it}), \quad (7)$$

where  $i = 1, \dots, I$  stands for the object, and  $t = 1, \dots, T$ , for the time period;  $y_{it}$  is the cost of production;  $V_{it}$  are random variables corresponding to stochastic noise; and  $U_{it}$ , nonnegative random variables associated with the inefficiency of an object  $i$  at time period  $t$ . In this paper, the random variables  $V_{it}$  and  $U_{it}$  are considered to be independent of each other. Each of them is also assumed to be identically distributed and independent for different objects or periods of time. Random variables  $V_{it}$ , stochastic noise, are usually considered to be normally distributed with zero mean and unknown dispersion  $N(0, \sigma_v^2)$ . Random variables  $U_{it}$ , associated with nonnegative inefficiency, are usually considered to have a half-normal, truncated normal, exponential, or gamma distribution with unknown parameters. Dispersions may or may not depend on a specific object or time (heteroskedastic or homoskedastic models, respectively). In this paper, we assume that the dispersions  $\sigma_v^2$  and  $\sigma_u^2$  are the same for all objects and time periods. Following [12], we consider a model with the mean value of  $U_{it}$  depending on time.

SFA models may be estimated for one period of time (cross-sectional models) or for a series of periods of time (panel data). The SFA literature shows that panel data models have better statistical properties. Analytically, parameters of the SFA models are estimated in three steps. At the first step, parameters of the deterministic component of the stochastic frontier are evaluated by using an appropriate modification of the least-squares method. Residuals obtained at the first step represent observed values of the random variable  $\varepsilon_{it}$ , the sum of the two random components  $U$  and  $V$ :

$$\varepsilon_{it} = V_{it} + U_{it} \quad (8)$$

At the second step, parameters of the distribution of  $U$  and  $V$  are found analytically or numerically by using the maximum

likelihood method; see [10] for details. The main idea of the maximum likelihood method is to find the values of the parameters that maximize the probability density function conditioned on observed data. Different statistical hypotheses may be tested, and the values of the previously found parameters may be corrected. At the third step, parameters of the conditional distribution of the inefficiency component  $U_{it}$  are calculated, conditioned on the observed values of the corresponding random variable  $\varepsilon_{it}$ , as suggested in [7] and [8]:

$$f(U_{it} | \varepsilon_{it}) = \frac{f(U_{it}, \varepsilon_{it})}{f(\varepsilon_{it})}, \quad (9)$$

where functions  $f(\cdot)$  stand for corresponding probability density functions. The term  $\varepsilon_{it}$  may or may not depend on the object or the time. Efficiency scores are usually defined either as an exponent of the conditional expectation of the random variable  $U_{it}$ ,

$$TE_{it} = e^{E(-U_{it} | \varepsilon_{it})}, \quad (10)$$

or as an expectation of the conditional exponent,

$$TE_{it} = E\left(e^{-U_{it}} | \varepsilon_{it}\right). \quad (11)$$

Monograph [9] argues that the latter is preferable; we have used it in this paper.

Different modifications of this procedure are known in the literature, but the main component is always a solution of an optimization problem related to the maximum likelihood method. In some cases, an analytical solution may be obtained, but it is more typical for SFA applications to use computer software. Several software products are available for this purpose. Article [13] furnishes a review of the most popular products – FRONTIER [14], and LIMDEP [15]. We have used the former for the calculations in this paper.

The application of SFA to environmental performance is considered in [15, 16, 17]. It may be noted that no insuperable barrier separates DEA and SFA. Publications [18, 19, 20] show that the SFA technique may be incorporated into DEA. To do this, the DEA frontier should be considered as an estimate for the deterministic component of the SFA frontier. We do not use this approach in this paper.

### III. STATISTICAL DATA AND MATHEMATICAL MODEL

In this section, we present statistical data and SFA model of environmental efficiency used in this paper. Our main source of statistical information was the EIA web site [www.doe.eia.gov](http://www.doe.eia.gov). The site provides data for GDP, population, total primary energy consumption, and carbon dioxide emissions for the period 1990–2007. For comparability, we have transformed the data into the ratios to the U.S. total. Ratios corresponding to CO<sub>2</sub> emissions for 2006 and 2007 were obtained by parabolic extrapolation. The objects under investigation in our study are the States of the U.S. The States are very different in all of the parameters: GDP, population, energy consumption, and CO<sub>2</sub> emissions. Table 1 provides an example of the data for 2007 for selected States: California(CA), Connecticut (CT), New York (NY), and Texas (TX). As follows from this table, big States like CA,

TABLE I  
DATA FOR SELECTED STATES, 2007, % OF TOTAL

State	CO <sub>2</sub> <sup>a</sup>	Energy Consumption	GDP	Population
CA	6.49	8.44	13.20	12.11
CT	0.80	0.85	1.54	1.17
NY	3.78	4.01	8.06	6.45
TX	8.63	11.66	8.37	7.91

a.Extrapolated

and TX, coexist with much smaller ones, like CT. Table 2 presents data on carbonization ratios for these states (value of 1.00 corresponds to the U.S. average).

TABLE II  
CARBONIZATION RATIOS  
FOR SELECTED STATES, 2007

State	CO <sub>2</sub> /GDP	CO <sub>2</sub> /Energy Consumption	CO <sub>2</sub> /Population
CA	0.49	0.77	0.54
CT	0.52	0.94	0.68
NY	0.47	0.94	0.59
TX	1.03	0.74	1.09

As may be seen from this table, CA is a very “clear” state, while TX is above the average in two indicators out of three. Energy carbonization ratio is independent from other two: TX has better level of 0.74 than CA’s 0.77. Both CT and NY have higher levels of energy carbonization (0.94 for both) though they are also below the U.S. average.

It is worth noting that large GDP, population, or even energy consumption, may or may not lead to correspondingly large carbon dioxide emissions ratios. Such interplay of the factors demonstrates the need for mathematical modeling of the environmental performance of the States’ economies aimed at determining those that are not sufficiently environmentally efficient. These states should increase their environmental performance in the overall perspective.

For our application of SFA, we used a model suggested in [12] and well adapted to the objectives of this paper. The resulting indicator is CO<sub>2</sub> emissions. The frontier is formed by the indicators of GDP, energy consumption, and population and total energy consumption, all expressed as ratios to the U.S. total. The rationale underlying the choice of the indicators is an assumption that a State that has population greater than the U.S. average, is economically more active (as measured by the GDP ratio), or is objectively energy intensive may be allowed to emit proportionally greater share of CO<sub>2</sub>. For our calculations, we used the cost function translog model:

$$\ln Z_{it} = \beta_0 + \beta_1 \ln r_{1it} + \beta_2 \ln r_{2it} + \beta_3 \ln r_{3it} + \beta_4 (\ln r_{1it})^2 + \beta_5 (\ln r_{2it})^2 + \beta_6 (\ln r_{3it})^2 + \beta_7 \ln r_{1it} \cdot \ln r_{2it} + \beta_8 \ln r_{1it} \cdot \ln r_{3it} + \beta_9 \ln r_{2it} \cdot \ln r_{3it} + (V_{it} + U_{it}), \quad (12)$$

where *i* stands for a state; *t*, for a period of time; *r*<sub>1</sub>, *r*<sub>2</sub>, and *r*<sub>3</sub> for the ratios of energy consumption, GDP, and population, respectively, all expressed as ratios to the U.S. totals; *V*<sub>it</sub> = *N*(0, σ<sup>2</sup>) is a normally distributed random variable; and *U*<sub>it</sub> = |*N*(*z*<sub>it</sub>, σ<sub>it</sub><sup>2</sup>)| is a nonnegative, half-normally distributed random variable intended to account for technical inefficiency. Random variables *U*<sub>it</sub> and *V*<sub>it</sub> corresponding to different states or periods of time are assumed to be independent and to have

parameters that are neither object nor time specific. The mathematical expectation of *U*<sub>it</sub> is assumed to be a quadratic function of time:

$$E(U_{it}) = z_{it} = \beta_i + \delta_{1i}t + \delta_{2i}t^2, \quad (13)$$

where *E* stands for mathematical expectation. For our calculations, we used the FRONTIER 4.1 software developed in [14].

#### IV. RESULTS AND DISCUSSION

For the period of investigation from 1990 through 2007, the efficiency scores ranged from 97.35% to 99.62% with standard deviation of 0.54%. Table 3 and Fig. 1 present average efficiency scores calculated by years.

TABLE III  
AVERAGE EFFICIENCY  
SCORES, %

Year	Score	Year	Score
1990	97.56	1999	99.60
1991	98.55	2000	99.59
1992	99.03	2001	99.57
1993	99.28	2002	99.54
1994	99.42	2003	99.48
1995	99.50	2004	99.39
1996	99.55	2005	99.23
1997	99.58	2006	98.94
1998	99.59	2007	98.36
Min	97.56	Max	99.59

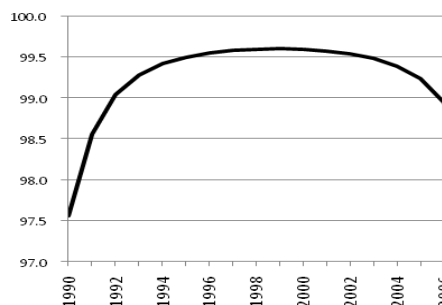


Fig. 1. Average efficiency scores, %

In spite of the differences in both the size of economies and their specialization, all States revealed very similar dynamics of efficiency scores. In our opinion, it is the result of similarity in the market forces and States’ and Federal environmental policies and regulations.

As follows from the presented data, efficiency scores follow the bell – shaped pattern. They are on the rise from 1990 (97.56%) through 1999 (99.60%), and then, on the decline to 98.36% in 2007. To provide explanations to this pattern of change, we analyzed the rates of growth of all of the ratios included in the model. Obtained results are shown in table 4. As follows from the table, the average rates of growth of all factors of the CO<sub>2</sub> emissions are lower in the period of 2000 – 2007, together with the average rate of growth of the CO<sub>2</sub> emissions itself. However, the decrease in the rate of growth of the CO<sub>2</sub> emissions was insufficient to offset the growth in GDP, population, and energy consumption and thus, to provide increase in the environmental efficiency. This observation stresses the importance of setting restrictions on

TABLE IV  
AVERAGE RATES OF GROWTH, %

	1990 - 1999	2000 - 2007
GDP	5.52	5.12
Population	1.25	0.96
Energy consumption	1.50	0.60
CO <sub>2</sub> emissions	2.69	1.87

the absolute amounts of the greenhouse gas emissions. Restrictions set on their rates of growth may or may not be sufficient. Mathematical models, like the one presented in this paper, can provide a solution to the problem in cases when decrease in the quantity of the emissions is not feasible. In such situations, the models can be used to compute different scenarios of restrictions on the rates of growth that result in the increase of the environmental performance.

#### V. CONCLUSIONS

An SFA translog model is suggested as a means of comparative evaluation of environmental performance of the U.S. States. A ratio – based model is used with mathematical expectation of the inefficiency term being a quadratic function of time. Inclusion of time in the model allows for the investigation of efficiency dynamics. Environmental efficiency of the U.S. States was analyzed using suggested approach for the period of 1990 – 2007. Obtained results reveal similarity in the dynamics of States' efficiency scores and the bell - shaped change in time. After the steady rise of the average scores from 97.6% in 1990 to 99.6% in 1999, they fell to 98.4% in 2007. Comparative analysis of the rates of growth allows for the assumption that the main factor of the decrease in environmental efficiency is insufficient decrease in the rate of growth of CO<sub>2</sub> emissions needed offset the rates of growth of GDP, population and energy consumption. Data for 2008, the year following the research period, allow for assumption that environmental performance of the U.S. States has improved in 2008 and the last years. The suggested model may be used for environment – related research and policy – making as a means for investigation or strategic regulation of CO<sub>2</sub> emissions. In particular, it may be used as a tool for the cap-and-trade regulatory system.

#### ACKNOWLEDGMENT

Author is thankful to Perry Lindstrom (EIA) for productive discussion of the obtained results.

#### REFERENCES

[1] *Emissions of Greenhouse Gases in the United States 2008*. Energy Information Administration. <http://www.eia.doe.gov>, 2009.

- [2] Sheremet, A. D. *Kompleksnyi ekonomicheskii analiz deiatelnosti predpriiatiia : (voprosy metodologii)*. Moskva : Ekonomika, 1974. (In Russian).
- [3] K. Yamaji, R. Matsuhashi, Y. Nagata, Y. Kaya, An Integrated System for CO<sub>2</sub>/Energy/GNP Analysis: Case Studies on Economic Measures for CO<sub>2</sub> Reduction in Japan. *Workshop on CO<sub>2</sub> Reduction and Removal: Measures for the Next Century*, March 19, 1991, International Institute for Applied Systems Analysis, Laxenburg, Austria.
- [4] M. Raupach, G. Marland, P. Ciais, C. Le Quere, J. Canadell, G. Klepper, C. Field, Global and regional drivers of accelerating CO<sub>2</sub> emissions. *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 24, pp. 10288 – 10293. June 12, 2007, [www.pnas.org/cgi/doi/10.1073/pnas.0700609104](http://www.pnas.org/cgi/doi/10.1073/pnas.0700609104).
- [5] Aigner D, Lovell C, Schmidt P. Formulation and estimation of stochastic frontier production function models. *J Econometrics* 1977;6:21-37.
- [6] Meeseus W, van Den Broeck J. Efficiency estimation from Cobb–Douglas production functions with composed error. *International Economic Review* 1977;18:435-444.
- [7] Materov I. On full identification of the stochastic production frontier model. *Ekonomika i matematicheskie metody* 1981;17:784-788 (in Russian).
- [8] Jondrow J, Materov I, Lovell C, Schmidt P. On the estimation of technical inefficiency in the stochastic frontier production model. *J Econometrics* 1982;19:233-238.
- [9] Kumbhakar S, Lovell C. *Stochastic frontier analysis*. Cambridge, UK: Cambridge University Press, 2003.
- [10] Pawitan Y. *In all likelihood: Statistical modelling and inference using likelihood*. 1st ed. USA: Oxford University Press, 2001.
- [11] Cooper W, Seiford L, Zhu J. *Handbook on data envelopment analysis* (International Series in Operations Research & Management Science). New York: Springer, 2004.
- [12] Battese G, Coelli T. A model for technical efficiency effects in a stochastic frontier production function for panel data. *Empirical Economics* 1995;20:325-332.
- [13] Sena V. Stochastic frontier estimation: A review of the software options. *J Applied Econometrics* 1999;14:579-586.
- [14] Coelli T. A guide to FRONTIER Version 4.1: A computer program for stochastic frontier production and cost function estimation. Centre for Efficiency and Productivity Analysis (CEPA) Working Papers 1996; 7. Department of Econometrics, University of New England, Australia, <http://www.une.edu.au/econometrics/cepa.htm>.
- [15] Greene W. LIMDEP (Version 7): User's manual and reference guide. *Econometric Software*: New York, 1995.
- [16] Cuesta R, Lovell C, Zofio J. Environmental efficiency measurement with translog distance functions: A parametric approach. *Ecological Economics* 2009;68:2232-2242.
- [17] Hattori T. Relative performance of U.S. and Japanese electricity distribution: An application of stochastic frontier analysis. *J Productivity Analysis* 2002;18:269-284.
- [18] Reinhard S, Lovell C, Thijssen G. Environmental efficiency with multiple environmentally detrimental variables; estimated with SFA and DEA. *European J Operational Research*, 2000;121:287-303.
- [19] Avkiran N, Rowlands T. How to better identify the true managerial performance: State of the art using DEA. *Omega* 2008;36:317-324.
- [20] Fried H, Lovell C, Schmidt S, Yaisawarng S. Accounting for environmental effects and statistical noise in data envelopment analysis. *J Prod Anal* 2002;17(1-2):157-174.