



**AgEcon** SEARCH  
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

*The World's Largest Open Access Agricultural & Applied Economics Digital Library*

**This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.**

**Help ensure our sustainability.**

Give to AgEcon Search

AgEcon Search  
<http://ageconsearch.umn.edu>  
[aesearch@umn.edu](mailto:aesearch@umn.edu)

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*



**Paper prepared for the 122<sup>nd</sup> EAAE Seminar**  
**"EVIDENCE-BASED AGRICULTURAL AND RURAL POLICY MAKING:  
METHODOLOGICAL AND EMPIRICAL CHALLENGES OF POLICY  
EVALUATION"**

**Ancona, February 17-18, 2011**



**Sub-vector Efficiency analysis in Chance Constrained  
Stochastic DEA: An application to irrigation water use in the  
Krishna river basin, India**

Veetil P.C.<sup>1</sup>, Ashok A.<sup>2</sup>, Speelman S.<sup>1</sup>, Buysse J.<sup>1</sup> and van Huylenbroeck G.<sup>1</sup>

1 Department of Agricultural Economics, Ghent University, Gent, Belgium

2 EIS Division, Central Institute of Fisheries Technology, Kochi, India

Prakashan.ChellattanVeetil@Ugent.be

*Copyright 2011 by Prakashan Chellattan Veetil, Arathy Ashok, Stijn Speelman, Jeroen Buysse and Guido van Huylenbroeck. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.*



# **Sub-vector Efficiency Analysis in Chance Constrained Stochastic DEA: An Application to Irrigation Water Use in the Krishna River Basin, India**

Chellattan Veetil P. , Arathy Ashok, Stijn Speelman, Jeroen Buysse and Guido van Huylenbroeck

## *Abstract*

*All deviations from the frontier is inefficiency in deterministic DEA (DDEA); thus making the DDEA unable to accommodate the measurement and specification errors. But, most of the production relationships are stochastic in nature with some inputs fixed in the short run. This paper addressed the above two issues by formulating a sub-vector efficiency model in a Stochastic DEA (SDEA) framework to analyze the efficiency of sub vector of inputs. The results illustrate that there is a wide scope for stochastic efficiency analysis. The overall efficiency in SDEA is higher than DDEA under both Constant and Variable Return to Scale frameworks. SDEA revealed that some efficient producers are not sub-vector efficient in our case study. Thus, overall efficiency oriented policy may not be sufficient for optimizing water use. The proposed model has limitations in terms of the degree of stochastic variability and the level of tolerance that the model can accommodate.*

*Keywords: stochastic DEA, sub-vector efficiency, chance constrained programming, irrigation water use efficiency*

*JEL classification: Enter JEL codes.*

## **1. INTRODUCTION**

Estimating the performance of productive units (Decision Making Unit or DMU) requires an appropriate methodology. Stochastic Frontier Analysis (SFA) and Data Envelopment Analysis (DEA) are the two dominant methods respectively in the most widely used parametric and non-parametric approaches for efficiency analysis. Non-parametric approaches on efficiency analysis has gained greater momentum after the pioneer work by Charnes et al in 1978 (CCR model) for a constant return to scale (CRS) version of DEA, which was later extended by Banker et. al. (1984) to variable return to scale (VRS) DEA framework (BCC model). In CCR-BCC models and other deterministic DEA (DDEA) methods observations at the frontiers are assigned to have an efficiency of unity and all those behind this frontier envelopment are given a value less than unity. That implies all deviations from the frontier are considered as inefficiency; thus making the DDEA unable to accommodate measurement and specification errors. As most production relationships are stochastic in nature, recently researchers started paying attention to incorporate stochastic considerations into DEA models. This paper attempts to extend the concept of stochastic DEA (SDEA) to analyze the sub-vector efficiencies in the context of irrigation water use in Krishna river basin, India.

The importance of sub-vector efficiency is illustrated here by taking the efficiency of irrigation water in an agricultural production perspective. Irrigation water is the limiting factor in most arid and semi-arid regions. In this situation the agricultural production relationships will be sub-optimal if we do not take into account the sub-vector efficiencies of irrigation water (Water use efficiency – WUE). The scope for improving water resource allocation is high if we know the farms/sectors with high WUE and very low WUE. Theoretically best allocation can be achieved between farms or sectors when the marginal productivity of water is equal in all farms or sectors. Additionally, we want to know whether farmers with efficient overall production are also efficient water users.

- The main contribution of this paper is to present a stochastic DEA model for sub-vector efficiency analysis and to compare the results obtained from this model with that of DDEA. Further, the proposed model is illustrated in the context of irrigation WUE in Krishna river basin where water is one of the scarce and limiting resource for agricultural production.
- The remaining sections of the paper are organized as follows. The next section briefly discusses the basic concepts of deterministic and stochastic efficiency. We here define the condition for an  $\alpha$ -stochastically efficient DMU. This is followed by a brief discussion about the family of deterministic DEA and a formulation of the basic CCR-BCC model under VRS and CRS framework. The concept of stochastic DEA model is introduced in a chance constrained framework. Later on these models are further extended to incorporate the sub-vector efficiency concept introduced by Fare et al (1994). The third section provides an empirical illustration of the proposed models in the context of irrigation water use efficiency of the agricultural production system in Krishna river basin, India. A comparison is made between the efficiency calculations under the two approaches. Finally the paper concludes with highlighting the scope and limitations of the model.

## **2. NON-PARAMETRIC FRONTIER EFFICIENCY ANALYSIS**

In the group of non-parametric efficiency analysis, here we focus on DEA and its variants which can accommodate the stochastic data for inputs and outputs. Farrell (1957) introduced the relative efficiency concept in his seminal work on technical efficiency. He defined technical efficiency as the ability of a farm to produce the maximum feasible output from a given bundle of inputs (output-oriented efficiency) or to use minimum feasible amounts of inputs to produce a given level of outputs (input-oriented efficiency). Extending the relative efficiency concept of Farrell, Charnes et al (1978) developed the first DDEA model (CCR model). The DDEA uses linear programming to calculate the efficient or best practice frontier through the piecewise linear envelopment of observed input-output combinations with the assumptions concerning scaling and disposability of inputs and outputs (Olesen & Petersen, 1995). The DMUs on this technical efficiency frontier are assigned with an efficiency score of unity and others behind this frontier get an efficiency score less than unity, treating all the deviations from this frontier as inefficiency. The CRS assumption in CCR model is further extended to VRS specification by

Banker et al (1984) famously known as BCC model. But, unfortunately these DDEA approaches cannot be used for applications with errors and random noises in data which often occur in reality. In response to this criticism, efforts have been made to extend these DDEA to accommodate stochasticity of inputs and outputs (Bruni, Conforti, Beraldi, & Tundis, 2009; Desai, Ratick, & Schinnar, 2005; Kenneth, Lovell, & Sten, 1993; Olesen & Petersen, 1995). Here we extend the stochastic DEA model to analyze the sub-vector efficiencies of inputs. First, we discuss the basic concepts of deterministic and stochastic efficiency followed by a brief discussion of the CCR-BCC model and then formulate the chance constrained DEA (CCDEA). Finally we extend the CCDEA to accommodate the sub-vector efficiency.

### ***2.1. Deterministic and Stochastic Efficiency : The concept***

Let  $j = 1, \dots, n$  be the collection of DMUs,  $X = (X_1, X_2, \dots, X_M) \in \mathfrak{R}_+^M$  denotes quantity vectors of  $m$  productive inputs and  $Y = (Y_1, Y_2, \dots, Y_K) \in \mathfrak{R}_+^K$  denotes the vector of  $k$  outputs. The production technology  $V(y)$  can be characterized by the production possibility set (PPS) which consists of all combinations of  $(x_j, y_j)$ ,  $j = 1, \dots, n$  which can be formulated as (Bruni, et al., 2009):

$$PPS = \{(x, y) : x = X'\lambda, y = Y'\lambda, I'\lambda = 1, \lambda \geq 0\} \quad (1)$$

Where  $Y$  is the  $(N \times K)$  matrix of observed outputs;  $y_j$  is the vector of outputs of current DMU;  $X$  is the  $(N \times M)$  matrix of observed productive inputs;  $x_j$  is the vector of productive inputs of current DMU;  $\lambda$  is a  $(N \times 1)$  vector of intensity variable representing the influence of each DMU in determining the technical efficiency of the current DMU;  $I'\lambda$  is a convexity constraint which specifies the VRS specification without which, the DEA model will be a CCR model describing a CRS situation. In a deterministic DEA model, the  $DMU_j$  is efficient if it is impossible to find a feasible solution for the following problem (Bruni, et al., 2009):

$$\begin{cases} X'\lambda \leq x_{mj} \\ Y'\lambda \geq y_{kj} \end{cases} \quad (2)$$

with  $\lambda \geq 0$  satisfying  $I'\lambda = 1$  and strict inequality holding for at least one constraint. The concept of efficiency can be extended to stochastic DEA by jointly comparing the outputs and inputs of DMU under study. Following Bruni et al (2009) the  $DMU_j$  is  $\alpha$ -stochastically efficient if and only if for any  $\lambda \geq 0$  satisfying  $I'\lambda = 1$ ,

$$\text{Prob} \left\{ \begin{array}{l} \tilde{X}(\omega)' \lambda \leq \tilde{x}_{mj}(\omega), \quad m = 1, \dots, M \\ \tilde{Y}(\omega)' \lambda \geq \tilde{y}_{kj}(\omega), \quad k = 1, \dots, K \end{array} \right\} \leq \alpha \quad (3)$$

with strict inequality holding for at least one constraint;  $\tilde{x}_j(\omega)$  and  $\tilde{y}_j(\omega)$  representing the

random input and output vectors respectively with  $\tilde{X}(\omega) = \sum_{j=1}^n \tilde{x}_{mj}(\omega)$  and

$\tilde{Y}(\omega) = \sum_{j=1}^n \tilde{y}_{kj}(\omega)$ . We assume that the distribution function of  $(\tilde{x}_j(\omega), \tilde{y}_j(\omega))$  is known.

Here we restrict the probability of the existence of dominating DMU to be  $\leq \alpha$ . Hence, the stochastic efficiency of the  $DMU_j$  can be measured by solving the following model:

$$\alpha^* = \max_{\lambda} \text{Prob} \left\{ \begin{array}{l} \tilde{X}(\omega)' \lambda \leq \tilde{x}_{mj}(\omega), \quad m = 1, \dots, M \\ \tilde{Y}(\omega)' \lambda \geq \tilde{y}_{kj}(\omega), \quad k = 1, \dots, K \end{array} \right\} \quad (4)$$

with  $\lambda \geq 0$  satisfying  $I'\lambda = 1$  and strict inequality holding for at least one constraint.  $\hat{\alpha}^*$  is the risk of incorrectly identifying  $DMU_j$  as non-dominated stochastically in its efficiency (see Bruni, et al., 2009 for more details). Cooper et al (1998) and Huang and Li (2001) have suggested separate chance constraints as the necessary and sufficient condition for  $DMU_j$  is  $\alpha$ -stochastically efficient. That is,  $DMU_j$  is  $\alpha$ -stochastically efficient if the following condition is satisfied:

$$\text{Prob} \left\{ \sum_{m=1}^M (\tilde{X}(\omega)' \lambda - \tilde{x}_{mj}(\omega)) - \sum_{k=1}^K (\tilde{Y}(\omega)' \lambda - \tilde{y}_{kj}(\omega)) < 0 \right\} \leq \alpha \quad (5)$$

## 2.2. Deterministic and stochastic DEA model for efficiency analysis

The non-parametric representation of the underlying production technology  $V(y)$  as described in the above section, is given below (Lansink & Silva, 2004):

$$V(y) = \{(x, y) : Y'\lambda \geq y_j, X'\lambda \leq x_j, I'\lambda = 1, \lambda \geq 0\} \quad (6)$$

The following dual formulation of the above production technology in a mathematical programming formulation can be written in an input oriented BCC framework as follows:

$$\begin{aligned} & \text{Min}_{\theta, \lambda} \theta \\ & \text{s.t.} \quad Y'\lambda \geq y_j, \\ & \quad X'\lambda \leq \theta x_j, \end{aligned}$$

$$I'\lambda = 1,$$

$$\theta \in \mathfrak{R}, \lambda \in \mathfrak{R}_+^N \quad (7)$$

Where  $\theta$  is the radial input contraction factor representing the technical efficiency of the above input-oriented programming formulation.

The above model corresponds to deterministic DEA (Schmidt, 1985) where we assume that there is no uncertainty affecting input-output vectors. This implicit assumption of no random noise in data is overcome by the following stochastic DEA approach.

### 2.1.1 Chance constrained formulation of DEA

The LLT (Land, Lovell and Thore) model formulates (Kenneth, et al., 1993) a basic chance constrained programming approaches for incorporating stochasticity in input and output vectors. The LLT model imposed the probabilistic constraints individually on each output and input, and does not account for the intra-DMU correlations. Oppositely, the OP (Olesen and Pertersen) model is formulated to introduce intra-DMU correlations (Olesen & Petersen, 1995), but overlooks the inter-DMU dependencies. By using joint probability, the stochastic DEA model allows to simultaneously handle inter and intra-DMU dependencies (LLT and OP models) as shown in model (5). Since in our present case, we focus on inter-DMU dependencies rather than intra-DMU correlations, we nevertheless built on the LLT model in our analysis. The programming model can be formulated as:

$$\begin{aligned} & \text{Min}_{\theta, \lambda} \theta \\ & \text{s.t. Prob} \left( \begin{array}{l} \tilde{X}(\omega)' \lambda \leq \theta \tilde{x}_{mj}(\omega) \\ \tilde{Y}(\omega)' \lambda \geq \tilde{y}_{kj}(\omega) \end{array} \right) \leq \alpha \\ & I'\lambda = 1 \\ & \theta \in \mathfrak{R}, \lambda \in \mathfrak{R}_+^N \end{aligned} \quad (8)$$

The joint probability constraint of the model (8) can be simplified using the following assumptions (for more details see Kenneth et al, 1993):

$$\begin{aligned} & Ey_{kj} = y_{kj} \quad \text{for all } j \text{ and } k \\ & Cov(y_{kj}, y_{ki}) = 0 \quad \text{for all } k \text{ and for } i \neq j \\ & Var y_{kj} = \sigma, \text{ constant for each } j \end{aligned}$$

as independent probability constraints and the resulting model is LLT specification:

$$\begin{aligned} & \text{Prob}(\tilde{X}(\omega)' \lambda \leq \theta \tilde{x}_{mj}(\omega)) \leq \alpha \\ & \text{Prob}(\tilde{Y}(\omega)' \lambda \geq \tilde{y}_{kj}(\omega)) \leq \alpha \end{aligned}$$



### 2.1.2 Sub-vector efficiency

Färe et al (1994) described the need for a notion of sub-vector efficiency as follows “in the short run some inputs might be fixed or uncontrollable and therefore it may be possible to contract only a sub-vector of inputs. Alternatively, some outputs may be produced under a fixed contract while others may be adjustable”. Sub-vector efficiency measures efficiency for a the sub-vector of inputs and outputs rather than for the entire vector of inputs and outputs (Färe, et al., 1994; Lansink & Silva, 2004; Speelman, D’Haese, Buysse, & D’Haese, 2008). Sub-vector efficiency of the DEA model (1) can be formulated as follows:

$$\begin{aligned}
 & \text{Min}_{\theta, \lambda} \theta^s \\
 & \text{s.t.} \quad Y' \lambda \leq y_{kj}, \\
 & \quad \quad X'_s \lambda \leq \theta^s x_{s,j}, \\
 & \quad \quad X'_{-s} \lambda \leq x_{-s,j}, \\
 & \quad \quad I' \lambda = 1, \\
 & \quad \quad \theta \in \mathfrak{R}, \lambda \in \mathfrak{R}_+^N
 \end{aligned} \tag{9}$$

Where  $\theta^s$  is the sub-vector efficiency;  $X_s$  is the sub-vector of the inputs contracted for the production of outputs,  $X_{-s}$  is the vector of all other inputs.

### 2.3. Chance constrained formulation of sub-vector efficiency

Using independent probability constrained the sub-vector efficiency model can be formulated as

$$\begin{aligned}
 & \text{Min}_{\theta, \lambda} \theta^s \\
 & \text{s.t. Prob} \left\{ \begin{aligned} & \left( \tilde{Y}(\omega)' \lambda \leq \tilde{y}_{kj}(\omega) \right) \\ & \left( \tilde{X}_{-s}(\omega)' \lambda \geq \tilde{x}_{-s,j}(\omega) \right) \\ & \left( \tilde{X}_s(\omega)' \lambda \geq \theta^s \tilde{x}_{sj}(\omega) \right) \end{aligned} \right\} \leq \alpha \\
 & \quad \quad I' \lambda = 1 \\
 & \quad \quad \theta \in \mathfrak{R}, \lambda \in \mathfrak{R}_+^N
 \end{aligned} \tag{10}$$

### 3. EMPIRICAL ILLUSTRATION

#### 3.1. The Data

Data on agricultural production systems were collected from the farmers of the Krishna river basin area of the northern Karnataka state in India from December to March 2008 by face-to-face interview method using a structured questionnaire. The Krishna river basin constitutes 8% of the total geographical area of India and flows through three Southern Indian states: Maharashtra, Karnataka and Andhrapradesh. This part includes four sub-basins, namely Lower Krishna, Ghataprabha, Malaprabha and Tungabhadra. Fig. 1 shows the map of the Krishna river basin. About 77% of the total basin area is cultivable (203,000 Km<sup>2</sup>) with an irrigation potential of 47,200 km<sup>2</sup> (IWMI, 2007). The majority of the basin area is arid or semi arid and faces high water scarcity. The per capita total renewable water resources availability of the basin is estimated to be 1,133 m<sup>3</sup> (Amarasinghe, et al., 2005). More than 90% of total water in Krishna River is used for irrigation. The cropping pattern in this basin is very diverse with field crops constituting the principal share. In this study, the villages and farmers within villages were selected randomly. The production details of 120 farms were collected and used in the DEA analysis.

Figure 1. The map of Krishna river basin showing the study area



Three outputs and eight productive inputs (water, land, labor, capital, manure, fertilizer, seed and chemicals) are distinguished in the production process. The outputs measured were

$y_1$  : quantity of rice grain in 100kg

$y_2$  : quantity of corn grain in 100kg

$y_3$  : quantity of sugarcane in tons

and the inputs were

$x_1$  : total amount of irrigation water used in acre - inches (1 acre - inch = 102.8 m<sup>3</sup>)

$x_2$  : total area under crops in acres (2.5 acre = 1 hectare)

$x_3$  : total family and hired labor used for cropping in man - days

$x_4$  : capital invested on machinery

$x_5$  : quantity of organic manure applied in tons

$x_6$  : quantity of fertilizer applied in 100kg

$x_7$  : seed in kg

$x_8$  : chemicals in kg

The outputs consist of rice, corn and sugarcane. Water is the total amount of irrigation water used in acre-inches (1 acre-inch = 102.8 cubic meter); Land represents the total area cultivated and is measured in acres (2.5 acre = 1 hectare); labor is measured in man-days and includes family as well as hired labor; capital consists of the capital invested in machinery; manure represents the organic manure which consists of natural products comprising farmyard manure, green manures, compost prepared from crop residues and other farm wastes, oil cakes and other manures from decaying plant/animal matter, and is measured in tons; fertilizer is measured in 100 kg; seed and chemicals are measured in kg. Table A in appendix provides the summary characteristics of the production factors used in the analysis.

The agricultural production faces uncertainties especially in irrigation water use, hence it is an ideal case to illustrate the advantage stochastic DEA in assessing sub-vector efficiency. We assume that within a farm, outputs are approximately normally distributed and the observed outputs serves as an unbiased estimate of the true outputs of the farm. Additionally we assume that all farms are stochastically independent, implying that the agricultural production of one farm is independent of other farms. It is quite reasonable to assume this independence as the agricultural production depends on the productive inputs and since the farms are randomly selected from different villages belonging to same agro-climatic conditions, the dependencies are minimal. Both linear (DDEA) and non-linear (SDEA) programming models are solved using GAMS (General Algebraic Modeling System). We use an input-oriented SDEA model because farmers have more control over the inputs than they have on output (Tuna & Oren, 2006).

### 3.2. Agricultural production efficiency of the farming system

Table 1 lists the results of efficiency scores of agricultural production system of the first 40 farms. The table can be divided into two parts: the left side presenting the DDEA efficiency scores and the right side the SDEA efficiency scores. In each method, both overall and sub-vector efficiencies under CRS and VRS specifications are reported. Chance constrained DEA efficiency scores are higher than deterministic DEA efficiency scores. The soft frontier in chance constrained DEA in contrast with the hard frontier in deterministic DEA allow output observations crossing the frontier (Kenneth, et al., 1993), but not too often.

A close look at the Table 1 reveals that the efficiency score for SDEA is higher than that of DDEA. This is quite logical as for DDEA the efficiency is bounded to a maximum efficiency ratio score of 1 where as in SDEA this hard frontier is relaxed to a soft frontier (Kenneth, et al., 1993). The greater be the stochasticity of outputs, the greater is the band of this frontier that can be crossed in SDEA. In our present study the tolerance limit is set to 5%. Almost half of the farmers lie on the output frontier in DDEA under the VRS framework where as it is slightly lower (44%) in CRS framework (see Table 3 for detailed distribution of the efficiency scores in different bins). In both cases, the fraction of farmers on the stochastic output frontier is higher (about 15%). Stochasticity of inputs and outputs allow more farmers obtaining higher efficiency scores (or lie on the ‘band of soft’ frontier). The average overall efficiency for DDEA is 0.870 and 0.856 under VRS and CRS framework respectively (Table 2). Similar deterministic

Table 1 Overall and Sub-vector Efficiency scores for DDEA and SDEA (for first 38 farms)

Farm	Overall Efficiency				Sub-Vector (Water Use) Efficiency			
	Deterministic DEA		stochastic DEA		Deterministic DEA		stochastic DEA	
	VRS	CRS	VRS	CRS	VRS	CRS	VRS	CRS
Farm001	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Farm002	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Farm003	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Farm004	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Farm005	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Farm006	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Farm007	0.902	0.902	1.000	1.000				
Farm008	0.819	0.819	0.977	1.000				
Farm009	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Farm010	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Farm011	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Farm012	0.686	0.686	0.720	1.000	0.200	0.200	0.200	0.200
Farm013	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Farm014	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Farm015	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Farm016	0.699	0.699	1.000	0.699	0.250	0.250	0.250	0.250

*Ancona - 122<sup>nd</sup> EAAE Seminar*  
*"Evidence-Based Agricultural and Rural Policy Making"*

Farm017	0.952	0.952	1.000	1.000	0.938	0.938	0.937	0.938
Farm018	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Farm019	0.927	0.927	0.927	0.927	0.625	0.625	0.625	0.625
Farm020	0.760	0.760	1.000	0.760	0.700	0.700	0.700	0.700
Farm021	0.774	0.774	0.834	0.834	0.153	0.153	0.204	0.191
Farm022	1.000	0.952	1.000	1.000	1.000	0.771	1.000	1.000
Farm023	0.692	0.692	0.692	0.756	0.380	0.380	0.446	0.789
Farm024	0.596	0.596	0.754	0.754	0.400	0.400	0.577	0.570
Farm025	0.544	0.544	0.581	0.581	0.404	0.404	0.451	0.441
Farm026	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Farm027	0.832	0.832	0.928	1.074	0.452	0.452	0.632	0.452
Farm028	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Farm029	0.891	0.891	1.000	0.891	0.572	0.572	1.000	1.000
Farm030	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Farm031	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Farm032	0.898	0.898	0.898	0.898				
Farm033	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Farm034	0.742	0.742	1.000	0.742	0.250	0.250	0.250	0.292
Farm035	0.991	0.991	1.000	1.000	0.357	0.357	0.357	0.506
Farm036	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Farm037	0.719	0.719	0.719	0.801	0.460	0.460	0.535	0.460
Farm038	0.945	0.815	1.000	1.000	0.546	0.459	1.000	0.459

technical efficiency scores were estimated for the wheat based cropping system (Tuna & Oren, 2006). The overall efficiency in SDEA is higher than DDEA under both frameworks and they are 0.912 and 0.894 respectively for VRS and CRS. Table 2 also reveals that the minimum and maximum efficiency ratios are also higher for SDEA.

Table 2. summary statistics of the efficiency ratios in two approaches

Statistic	Overall Efficiency				Sub-Vector (Water Use) Efficiency			
	Deterministic DEA		stochastic DEA		Deterministic DEA		stochastic DEA	
	VRS	CRS	VRS	CRS	VRS	CRS	VRS	CRS
Mean	0.870	0.856	0.912	0.894	0.681	0.627	0.708	0.682
Minimum	0.265	0.265	0.329	0.265	0.050	0.050	0.050	0.058
Maximum	1.000	1.000	1.198	1.074	1.000	1.000	1.0001	1.00001
Std. dev.	0.180	0.182	0.158	0.167	0.362	0.370	0.345	0.348

The distribution of the efficiency scores for DDEA and SDEA models are given in table 3. The percentage of farmers in the lower efficiency score bin is lower for SDEA compared with DDEA. The sub-vector efficiency also shows a similar trend. On an average the percentage of farmers in the lower efficiency bins are found higher for CRS model. The figures 2 and 3

showed the difference between efficiency scores of SDEA and DDEA for overall and sub-vector efficiency respectively. There are no difference between both efficiency scores for 60% and 54% of farmers under VRS and CRS overall efficiencies. 76.79% and 66.07% of farmers show no differences for sub-vector efficiencies (WUE) in both models. The average differences between SDEA and DDEA are 4.28, 3.82, 3.32 and 5.84 percentage point respectively for overall VRS, overall CRS, sub-vector VRS and sub-vector CRS models. In the case of sub-vector efficiencies this difference is found higher for CRS model compared to VRS model.

Table 3. distribution of efficiency scores (%) in two approaches

Efficiency score	Overall Efficiency				Sub-Vector Efficiency			
	Deterministic DEA		stochastic DEA		Deterministic DEA		stochastic DEA	
	VRS	CRS	VRS	CRS	VRS	CRS	VRS	CRS
<0.100	0.00	0.00	0.00	0.00	2.68	3.57	1.79	2.68
0.100-0.200	0.00	0.00	0.00	0.00	8.93	12.50	7.14	8.04
0.200-0.400	1.67	1.67	0.83	1.67	15.18	16.07	14.29	16.07
0.400-0.600	10.83	12.50	6.67	7.50	13.39	13.39	11.61	10.71
0.600-0.800	15.83	15.00	11.67	13.33	2.68	4.46	4.46	5.36
0.800-0.900	10.00	15.00	8.33	11.67	0.89	0.89	0.89	0.89
0.900-0.950	5.83	5.00	4.17	5.00	0.89	1.79	0.89	1.79
0.950-0.999	3.33	6.67	3.33	3.33	0.00	0.00	0.00	0.00
>0.999	52.50	44.17	65.00	57.50	55.36	47.32	58.93	54.46

Figure 2. The difference between SDEA and DDEA overall efficiency scores

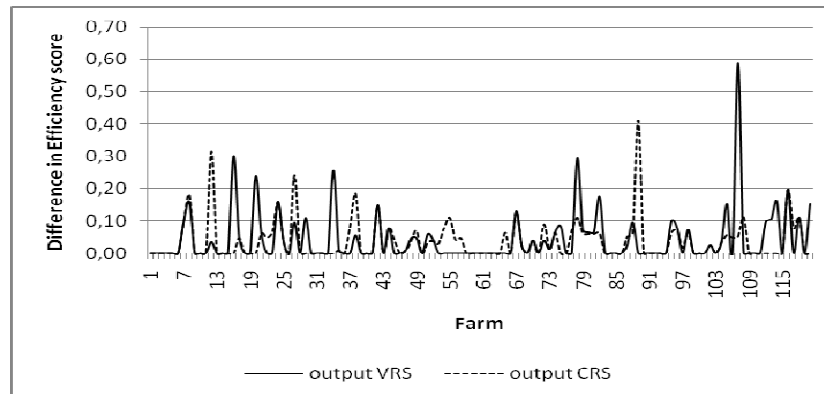
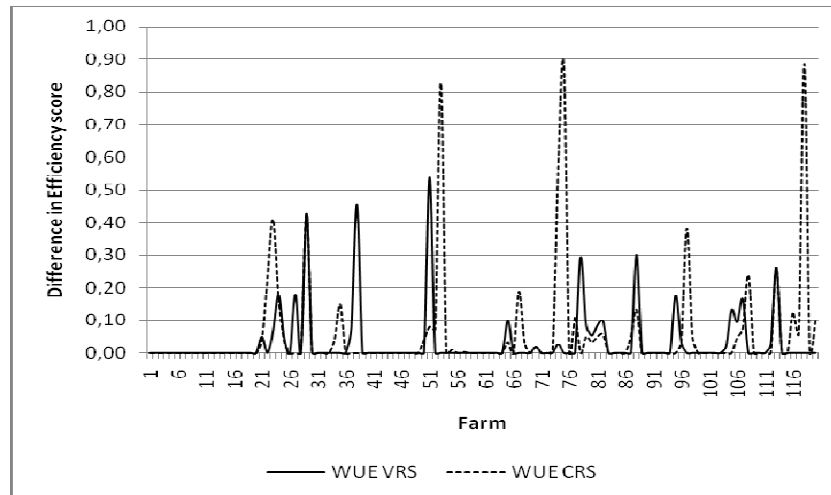


Fig 3. The difference between SDEA and DDEA subvector efficiency scores



### 3.3. Water use efficiency

Some of the sub-vector efficiency values are missing in table 1 because these farmers do not use irrigation (rainfed cropping) or use of extremely low amount of water. The sub-vector efficiency frontier shows a similar pattern as that of the overall efficiency frontier in terms of the number of farmers on the hard frontier where as the stochastic input efficiency shows a deviation from the stochastic overall efficiency frontier. 54% of the farmers lie on the stochastic input frontier for efficient water use under VRS framework where as it is 48% under CRS framework. According to the DDEA model outcome, those farmers who are efficient producers (overall efficiency) are also efficient in water use (sub-vector efficiency). This is not true in SDEA. Some efficient producers are not efficient in water use when we consider stochasticity in inputs and outputs. For example the Farm017 is an efficient producer but not efficient in water use under both the CRS and the VRS (WUE= 0.94) framework. The opposite can be true for both DDEA and SDEA. Though SDEA is more complex, the stochastic model is flexible and relies on fewer assumptions which can be violated than DDEA. 17% of the farms in the stochastic WUE frontier are not in the WUE frontier of DDEA under the CRS framework, but all farms in DDEA frontier are also in SDEA WUE frontier. Whereas under VRS framework it is only 7.5%.

In many decision problems, the uncertainties and existence of slacks are integral part of reality and the SDEA offers one format for this (Kenneth, et al., 1993). Due to the unpredictability of climate and resulting uncertainties in irrigation water availability, the water use decisions in agricultural production often demands due attention to the underlying uncertainties. Chance constrained formulation acknowledge this uncertainty by introducing the stochasticity in input use. High variations in the estimated efficiency scores for sub-vector reflects the wide variations in actual water use of farmers.

#### **4. CONCLUSION**

A vast literature of DEA discusses the efficiency based on the mathematical theory of production which is deterministic in nature (Farrel, 1957). The sub-vector efficiency analysis in SDEA provides ample opportunity to accommodate the stochasticity of inputs in the production relationships. Since, agricultural production often faces uncertainties due to changing climatic, physical, social and political conditions, forgoing of such random errors and noises are not appreciable. This paper provides an illustration of incorporating stochasticity in a sub-vector efficiency analysis in irrigation water use in agricultural production relationships in semi-arid farming systems.

The result of the stochastic DEA efficiency model has the advantage of the greater scope to accommodate noises and errors in data compared to deterministic DEA model. The theoretical consequence is that the possibility of a few number of farms, possibly outliers, dominating the frontier is lower. As illustrated in the case study, the SDEA is determined by more DMUs than the DDEA frontier. The SDEA frontier has as a results a more complex and possibly a more complete representation of the technology in the frontier. However, the case also shows that the ranking of the efficiency ratios in both cases are almost similar which explains that the ranking is robust against assumptions about the noise.

The advantage of flexibility of SDEA has also some consequences. In fact DDEA can be considered as a special case of SDEA where the tolerance limit of noise is set to zero. SDEA is less restrictive by incorporating noise but the trade-off is that one also has to make an assumption about the tolerance limit. The greater the stochastic variability, the greater would be the band of soft frontier which can be crossed. This makes the efficiency ratios of data with large uncertainties close to unity (Kenneth, et al., 1993). A greater tolerance limit could also lead to the fact that real inefficiencies are attributed to plain noise. Therefore, even if we keep the stochastic variability as constant, an increase in the tolerance level of chance constraints always increases the efficiency score.

#### **ACKNOWLEDGEMENT**

We would like to acknowledge Prof. Sten Thore for his valuable help during the development of GAMS code.

#### **REFERENCES**

- Amarasinghe, U. A., Sharma, B. R., Aloysius, N., Scott, N., Smakhtin, V., & de Fraiture, C. (2005). Spatial Variation in Water Supply and Demand across River Basins of India. In. Colombo, Srilanka: International Water Management Institute.
- Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. *Management Science*, 30, 1078-1092.
- Bruni, M. E., Conforti, D., Beraldi, P., & Tundis, E. (2009). Probabilistically constrained models for efficiency and dominance in DEA. *International Journal of Production Economics*, 117, 219-228.
- Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2, 429-444.



*Ancona - 122<sup>nd</sup> EAAE Seminar*  
*"Evidence-Based Agricultural and Rural Policy Making"*

- Cooper, W. W., Huang, Z., Lelas, V., Li, S. X., & Olesen, O. B. (1998). Chance Constrained Programming Formulations for Stochastic Characterizations of Efficiency and Dominance in DEA. *Journal of Productivity Analysis*, 9, 53-79.
- Desai, A., Ratick, S. J., & Schinnar, A. P. (2005). Data envelopment analysis with stochastic variations in data. *Socio-Economic Planning Sciences*, 39, 147-164.
- Färe, R., Grosskopf, S., & Lovell, C. A. K. (1994). *Production Frontiers*: Cambridge University Press, London.
- Farrel, M. J. (1957). The Measurement of Productive Efficiency. *Journal of Royal Statistical Society*, 120, 37.
- Huang, Z., & Li, S. X. (2001). Stochastic DEA Models With Different Types of Input-Output Disturbances. *Journal of Productivity Analysis*, 15, 95-113.
- IWMI. (2007). International Benchmark basin: Krishna. In. Colombo, Srilanka: International Water Management Institute (IWMI).
- Kenneth, C. L., Lovell, C. A. K., & Sten, T. (1993). Chance-Constrained Data Envelopment Analysis. *Managerial and Decision Economics*, 14, 541-554.
- Lansink, A. O., & Silva, E. (2004). Non-Parametric Production Analysis of Pesticides Use in the Netherlands. *Journal of Productivity Analysis*, 21, 49-65.
- Olesen, O. B., & Petersen, N. C. (1995). Chance Constrained Efficiency Evaluation. *Management Science*, 41, 442-457.
- Schmidt, P. (1985). Frontier production functions. *Econometric Reviews*, 4, 40.
- Speelman, S., D'Haese, M., Buysse, J., & D'Haese, L. (2008). A measure for the efficiency of water use and its determinants, a case study of small-scale irrigation schemes in North-West Province, South Africa. *Agricultural Systems*, 98, 31-39.
- Tuna, A., & Oren, M. N. (2006). Determinants of Technical Efficiency of Wheat Farming in Southeastern Anatolia, Turkey: A Nonparametric Technical Efficiency Analysis. *Journal of Applied Sciences*, 6, 827-830.

## APPENDIX

Table A. The summary characteristics of the production factors used in the model

Production factor	Mean	Std. dev.	Minimum	Maximum
<b>Inputs</b>				
Irrigation water ( $x_1$ )	48.89	52.24	0.00	360.00
Land area ( $x_2$ )	7.28	10.78	1.00	100.00
Labor ( $x_3$ )	87.84	73.48	11.00	519.00
Capital ( $x_4$ )	2.37	4.03	0.00	28.00
Organic manure ( $x_5$ )	8.49	11.07	0.00	80.00
Fertilizer ( $x_6$ )	4.63	5.05	0.00	40.00
Seed ( $x_7$ )	128.16	686.36	0.00	5000.00
Chemicals ( $x_8$ )	3.32	22.03	0.00	300.00
<b>outputs</b>				
Rice ( $y_1$ )	7.99	11.84	0.00	55.00
Corn ( $y_2$ )	3.07	6.12	0.00	24.00
Sugarcane ( $y_3$ )	29.82	38.05	0.00	360.00