METHODOLOGIES AND APPLICATION

The uncertain two-stage network DEA models

Bao Jiang[1] · Hao Chen[1] · Jian Li[1,2] · Waichon Lio[3]

Abstract

The two-stage network DEA models based on the framework that the efficiency of the whole stage is equal to the product of the efficiencies of two sub-stages can not only turn the 'black box' into the 'glass box' to identify the root causes of the inefficiency of the network system, but also consider the relationship between the two sub-stages within the whole stage. Nowadays, the two-stage network DEA models have been widely applied in the field of economy and management, such as green supply chain and reverse supply chain. Due to the novelty of evaluation indexes, these emerging research objects with network structure, such as green supply chain, involve not only traditional evaluation indexes such as cost and time, but also some novel evaluation indexes such as customer satisfaction and flexibility. However, these new evaluation indexes are difficult to quantify accurately, which will lead to the failure of the traditional two-stage network DEA models. Therefore, this paper attempts to extend the traditional two-stage network DEA models to the uncertain two-stage network DEA models with the application of uncertainty theory. In the new models, inputs, intermediates and outputs are considered to be uncertain variables to deal with the problem of inaccurate data. Finally, a numerical example of the uncertain two-stage network DEA models will be presented for illustration.

Keywords Two-stage network system · Data envelopment analysis · Uncertainty theory · Efficiency

1 Introduction

With the globalization of economy and the popularity of supply chain, the traditional 'black-box' DEA models such as CCR or BCC models Charnes et al. (1978) which can only recognize DMU as a whole rather than break down the whole system into sub-stages are no longer perfectly suitable for evaluating a system with complex network structure Cook and Zhu (2010). This is because the 'black-box' models ignore the internal structure of the network system and hence cannot provide enough information on identify-

✉ Jian Li
   lijian@ouc.edu.cn

   Bao Jiang
   jiangbao@ouc.edu.cn

1   College of Economics, Ocean University of China, Qingdao 266100, China

2   Marine Development Studies Institute, Ocean University of China, Qingdao 266100, China

3   Department of Mathematical Sciences, Tsinghua University, Beijing 100084, China

ing the root causes of the inefficiency of the network system Esmaeilzadeh and Matin (2019). For decision makers, it is not enough to evaluate whether the network system is efficient or not from a holistic perspective. In order to improve the performance of the network system comprehensively, decision makers need to investigate the running states of sub-stages to make the optimal decisions Chen (2009). Therefore, many researchers have put forward that the traditional 'black-box' models should be converted into 'glass-box' models in which the internal structure of the network system should be taken into account in the efficiency evaluation Kao (2014).

In order to transform 'black-box' models into 'glass-box' models, many scholars have devoted great efforts to the study of network DEA models. For example, Fǎre and Grosskopf (2000) proved the usefulness of network DEA by analyzing several network models with different structures. Seiford and Zhu (1999) established the two-stage network DEA models to assess sub-stages of 55 top banks in US. Cook et al. (2014) summarized the latest researches of network DEA models.

Among all the network structures, the two-stage network structure has been considered as the most fundamental and straightforward structure because it can be expanded to some more intricate structures and has been widely used in numer-

ous practical studies Kao and Hwang (2008). For example, Seiford and Zhu (1999) decomposed the entire operation process of a commercial bank into two sub-stages: profitability and marketability, and the efficiencies of two sub-stages and the whole stage were calculated by CCR models respectively. Some other examples contain the performance of bank branch affected by information technology (Chen and Zhu 2004) and the performance of two-stage Major League Baseball (Sexton and Lewis 2003).

Some of the above studies are based on the two-stage network DEA models in which two sub-stages are regarded as independent individuals with their own inputs and outputs respectively. However, this research method does not consider the influence of the relationship between the two sub-stages within the whole stage on efficiency evaluation of the two-stage network system. From the perspective of supply chain, due to the close cooperation between upstream and downstream enterprises, each individual does not exist independently, whose inputs and outputs are affected by neighboring enterprises. For instance, for the purpose of reaching an efficient status, the downstream enterprise may choose to cut down its inputs (intermediates). But this action means that the outputs of the upstream enterprise will be reduced, which may lead to the decrease of the efficiency of the upstream enterprise. Thus, there will be potential conflict among the two sub-stages, which is caused by their relationship.

In order to solve this potential conflict caused by the relationship between two sub-stages, a framework that the efficiency of the whole stage is equal to the product of the efficiencies of two sub-stages was proposed by Kao and Hwang (2008). This reasonable relationship reflects the close connection between the two sub-stages within the whole stage, which can effectively cope with this potential conflict. Moreover, the two-stage network DEA models considering their relationship are more reliable in identifying the efficiencies of the network system and consequently are capable of more accurately finding out the root causes of the inefficiency of the network system (Kao and Hwang 2008). Therefore, it is essential to establish the two-stage network DEA models that conform to their relationship Kao (2016).

Recently, some burgeoning issues such as green supply chain and reverse supply chain have received increasing attention in many researches Dyckhoff and Allen (2001). Due to the novelty of evaluation indexes, these emerging research objects with network structure, such as green supply chain, involve not only traditional evaluation indexes such as cost and time, but also some new evaluation indexes such as customer satisfaction, social benefits and flexibility. However, these new evaluation indexes are difficult to quantify accurately and hence considered as inaccurate data, which will lead to the failure of the traditional two-stage network DEA models (Sadjadi and Omrani 2008). Therefore, in order

to accurately evaluate these emerging issues, it is urgent to handle the problem of inaccurate data.

In order to solve the imprecise data, many stochastic DEA models have been built up based on the probability theory by Sengupta (1982), Banker (1993), Sueyoshi (2000) and Olesen and Petersen (2016). Furthermore, fuzzy methods have been used to deal with the inaccurate data in DEA. Kao (2011) developed the fuzzy forms of the two-stage network DEA models. Some other researchers also have been devoted to the application of fuzzy method in two-stage network DEA, including Tavana (2018), Hatami-Marbini and Saati (2018) and Wanke et al. (2018).

Nowadays, there is a tendency to use uncertainty theory established by Liu (2007) to cope with the problem of inaccurate data. The core advantage of uncertainty theory lies in the reasonable construction of belief degree to solve the problem of inaccurate data on the basis of overcoming the limitations of human beings Liu (2012). Moreover, uncertainty theory has constantly been applied in many fields of science and engineering. Liu (2009) put forward uncertain programming in 2009. Many other scholars have also been devoted to the application in the field of programming from uncertain machine scheduling (Liu 2010) and uncertain multi-objective programming (Liu and Chen 2015) to uncertain multilevel programming model (Liu and Yao 2015).

In recent years, uncertainty theory has been also studied and proved by many experts in the field of DEA. Wen et al. (2014) proposed the first uncertain DEA model. Then, Lio and Liu (2018) introduced an uncertain DEA model containing imprecise inputs and outputs. Jiang et al. (2019) further established an uncertain DEA model to recognize the scale efficiency in uncertain situation. In addition, Jiang et al. (2020) developed two new uncertain DEA models capable of further identifying specific scale efficiency. However, up to now, there is no research on the uncertainty theory in the two-stage network DEA models.

Thus, the purpose of this paper is to apply and extend the uncertainty theory to the two-stage network DEA models. In general, the contributions of this paper are as follows. Firstly, this article expands the application of uncertainty theory in the field of DEA. Secondly, this paper expands the evaluation method of unobservable variables and the scope of application of DEA models by establishing the uncertain two-stage network DEA models. At last, for the contribution of practice, the new models can provide more constructive opinions for decision makers to identify the root causes of the inefficiency of the network system in complex and uncertain environment and to further enhance the performance of the network system.

The rest of this article is organized as follows. Section 2 will provide some basic theoretical knowledge of uncertainty theory. The uncertain two-stage network DEA models and their equivalent forms will be presented and proved in Sect. 3.

In Sect. 4, a numerical example will be given for illustration. The final section is the conclusion remarks.

2 Preliminaries

In order to better expand the application of uncertainty theory in two-stage network DEA models, we will introduce the basic knowledge and concepts of uncertainty theory. The fundamental and essential definition of uncertainty theory is uncertain measure. The uncertain measure $\mathcal{M}$ is defined as a set function on a σ-algebra $\mathcal{L}$ over a nonempty set Γ by the following three axioms Liu (2007):

Axiom 1. (Normality Axiom) $\mathcal{M}\{\Gamma\} = 1$ for the universal set Γ.

Axiom 2. (Duality Axiom) $\mathcal{M}\{\Lambda\} + \mathcal{M}\{\Lambda^c\} = 1$ for any event Λ.

Axiom 3. (Subadditivity Axiom) For every countable sequence of events $\Lambda_1, \Lambda_2, \ldots$, we have

$$\mathcal{M}\left\{\bigcup_{i=1}^{\infty} \Lambda_i\right\} \leq \sum_{i=1}^{\infty} \mathcal{M}\{\Lambda_i\}.$$

Besides, Liu (2009) proposed the product axiom for the product uncertain measure in 2009 as follows:

Axiom 4. (Product Axiom) Let $(\Gamma_k, \mathcal{L}_k, \mathcal{M}_k)$ be uncertainty spaces for $k = 1, 2, \ldots$. The product uncertain measure $\mathcal{M}$ is an uncertain measure satisfying

$$\mathcal{M}\left\{\prod_{k=1}^{\infty} \Lambda_k\right\} = \bigwedge_{k=1}^{\infty} \mathcal{M}_k\{\Lambda_k\}$$

where Λ_k are arbitrarily chosen events from $\mathcal{L}_k$ for $k = 1, 2, \ldots$, respectively.

Definition 1 (Liu 2007) Uncertain variable is a measurable function τ from an uncertainty space $(\Gamma, \mathcal{L}, \mathcal{M})$ to the set of real numbers such that

$$\{\tau \in B\} = \{\gamma \in \Gamma \mid \tau(\gamma) \in B\}$$

is an event for any Borel set B of real numbers.

Definition 2 (Liu 2007) The uncertainty distribution Ψ of an uncertain variable τ is defined by

$$\Psi(x) = \mathcal{M}\{\tau \leq x\}$$

for any real number x.

We will introduce some common uncertainty distributions based on the definitions above. Firstly, we will introduce the most common one called linear uncertainty distribution, which is

$$\Psi(x) = \begin{cases} 0, & \text{if } x \leq a \\ (x-a)/(b-a), & \text{if } a < x \leq b \\ 1, & \text{if } x > b; \end{cases}$$

and the second common uncertainty distribution is called zigzag uncertainty distribution, which is

$$\Psi(x) = \begin{cases} 0, & \text{if } x \leq a \\ (x-a)/[2(b-a)], & \text{if } a < x \leq b \\ (x+c-2b)/[2(c-b)], & \text{if } b < x \leq c \\ 1, & \text{if } x > c. \end{cases}$$

An uncertainty distribution $\Psi(x)$ is deemed to be regular (Liu 2010) if it is a continuous function and strictly increasing with respect to x at which $0 < \Psi(x) < 1$, and

$$\lim_{x \to -\infty} \Psi(x) = 0, \qquad \lim_{x \to \infty} \Psi(x) = 1.$$

If τ is an uncertain variable with regular uncertainty distribution $\Phi(x)$, the inverse function $\Psi^{-1}(\alpha)$ is called the inverse uncertainty distribution of τ (Liu 2010).

Definition 3 (Liu 2009) Uncertain variables $\tau_1, \tau_2, \ldots, \tau_k$ are said to be independent provided that

$$\mathcal{M}\left\{\bigcap_{i=1}^{k}(\tau_i \in B_i)\right\} = \bigwedge_{i=1}^{k} \mathcal{M}\{\tau_i \in B_i\}$$

for any Borel sets $B_1, B_2, \ldots, B_k$ of real numbers.

The following theorem proposed by Liu (2010) can calculate the inverse uncertainty distribution of a strictly monotonous function of independent uncertain variables with regular uncertainty distributions as below:

Theorem 1 (Liu 2010) *Let $\tau_1, \tau_2, \ldots, \tau_k$ be independent uncertain variables with regular uncertainty distributions $\Psi_1, \Psi_2, \ldots, \Psi_k$, respectively. If f is strictly increasing with respect to $\tau_1, \tau_2, \ldots, \tau_q$ ($q \leq k$) and strictly decreasing with respect to $\tau_{q+1}, \tau_{q+2}, \ldots, \tau_k$, then $\tau = f(\tau_1, \tau_2, \ldots, \tau_k)$ is an uncertain variable with an inverse uncertainty distribution*

$$\Phi^{-1}(\alpha) = f(\Psi_1^{-1}(\alpha), \ldots, \Psi_q^{-1}(\alpha), \Psi_{q+1}^{-1}(1-\alpha), \ldots, \Psi_k^{-1}(1-\alpha)).$$

Expected value can be used to represent the size of τ, which is the average value of uncertain variable in the sense of uncertain measure. The formal definition is shown below.

Definition 4 (Liu 2007) The expected value of uncertain variable τ is defined by

$$E[\tau] = \int_0^{+\infty} \mathcal{M}\{\tau \geq x\}dx - \int_{-\infty}^0 \mathcal{M}\{\tau \leq x\}dx$$

provided that at least one of the two integrals is finite.

Suppose that τ is defined as an uncertain variable with uncertainty distribution Ψ. Afterwards, Liu (2007) Liu (2010) represented the formulas about the expected values of τ as below:

$$E[\tau] = \int_0^{+\infty} (1 - \Psi(x))dx - \int_{-\infty}^0 \Psi(x)dx,$$

$$E[\tau] = \int_0^1 \Psi^{-1}(\alpha)d\alpha.$$

Theorem 2 (Liu and Ha 2010) *Assume $\tau_1, \tau_2, \ldots, \tau_k$ are independent uncertain variables with regular uncertainty distributions $\Psi_1, \Psi_2, \ldots, \Psi_k$, respectively. If $f(\tau_1, \tau_2, \ldots, \tau_k)$ is strictly increasing with respect to $\tau_1, \tau_2, \ldots, \tau_q$ ($q \leq k$) and strictly decreasing with respect to $\tau_{q+1}, \tau_{q+2}, \ldots, \tau_k$, then the expected value of $\tau = f(\tau_1, \tau_2, \ldots, \tau_k)$ is*

$$E[\tau] = \int_0^1 f(\Psi_1^{-1}(\alpha), \ldots, \Psi_q^{-1}(\alpha), \Psi_{q+1}^{-1}(1-\alpha),$$
$$\ldots, \Psi_k^{-1}(1-\alpha))d\alpha.$$

3 The uncertain two-stage network DEA models

In this part, we will introduce the uncertain two-stage network DEA models considering the relationship between the two sub-stages within the whole stage. Similar to the traditional two-stage network DEA models, the structure of the uncertain two-stage network DEA models is also composed of two sub-stages, which is shown in Fig. 1. The inputs are transformed into intermediates through the first stage, and then the intermediates are totally put into the second stage. In the new models, inputs, intermediates and outputs are considered to be uncertain variables.
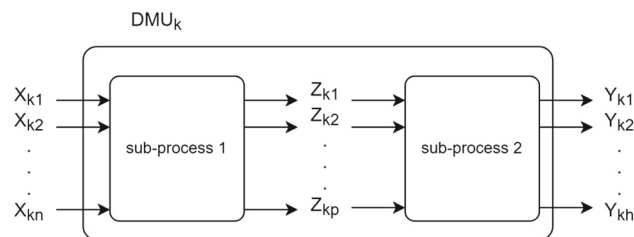

Fig. 1 Two-stage network system

In order to build the whole stage model, the first stage model and the second stage model of the uncertain two-stage network DEA models, some basic symbols need to be defined as below:

DMU_k: the kth DMU, $k = 1, 2, \ldots, j$

DMU_o: the target DMU

$\tilde{x}_k = (\tilde{x}_{k1}, \tilde{x}_{k2}, \ldots, \tilde{x}_{kn})$: the uncertain inputs vector of DMU_k, $k = 1, 2, \ldots, j$

$\tilde{x}_o = (\tilde{x}_{o1}, \tilde{x}_{o2}, \ldots, \tilde{x}_{on})$: the uncertain inputs vector of DMU_o

$\tilde{y}_k = (\tilde{y}_{k1}, \tilde{y}_{k2}, \ldots, \tilde{y}_{kh})$: the uncertain outputs vector of DMU_k, $k = 1, 2, \ldots, j$

$\tilde{y}_o = (\tilde{y}_{o1}, \tilde{y}_{o2}, \ldots, \tilde{y}_{oh})$: the uncertain outputs vector of DMU_o

$\tilde{z}_k = (\tilde{z}_{k1}, \tilde{z}_{k2}, \ldots, \tilde{z}_{kp})$: the uncertain intermediates vector of DMU_k, $k = 1, 2, \ldots, j$

$\tilde{z}_o = (\tilde{z}_{o1}, \tilde{z}_{o2}, \ldots, \tilde{z}_{op})$: the uncertain intermediates vector of DMU_o

$u \in \Re^n$: the vector of input weights

$v \in \Re^h$: the vector of output weights

$w \in \Re^p$: the vector of intermediate weights

3.1 The whole stage model

In order to establish the whole stage model more reasonably, we will take into account the proportional constraint of each sub-stage based on the relationship between the two sub-stages within the whole stage. Moreover, the multiplier of the intermediate $\tilde{z}$ we take in this model is always the same regardless of whether it plays the role of input or output throughout the process. Based on the above statement and uncertainty theory, this model is formulated as below:

$$\begin{cases} \max_{u,v,w} c_w = E\left[\frac{u^T \tilde{y}_o}{v^T \tilde{x}_o}\right] \\ s.t. \quad E\left[\frac{u^T \tilde{y}_k}{v^T \tilde{x}_k}\right] \leq 1, \quad k = 1, 2, \ldots, j \\ \quad\quad E\left[\frac{w^T \tilde{z}_k}{v^T \tilde{x}_k}\right] \leq 1, \quad k = 1, 2, \ldots, j \\ \quad\quad E\left[\frac{u^T \tilde{y}_k}{w^T \tilde{z}_k}\right] \leq 1, \quad k = 1, 2, \ldots, j \\ \quad u \geq 0 \\ \quad v \geq 0 \\ \quad w \geq 0. \end{cases} \quad (1)$$

Definition 5 *(Uncertain DEA Efficiency)* If and only if the optimal solution $\overline{c_w}$ of the model (1) can reach 1, DMU_o can be considered as efficient.

The efficiency of DMU_0 can be calculated by evaluating the optimal value of the whole stage model. The equivalent form of model (1) is proved as below:

Theorem 3 *For each k, let inputs $\tilde{x}_{k1}, \tilde{x}_{k2}, \ldots, \tilde{x}_{kn}$, outputs $\tilde{y}_{k1}, \tilde{y}_{k2}, \ldots, \tilde{y}_{kh}$ and intermediates $\tilde{z}_{k1}, \tilde{z}_{k2}, \ldots, \tilde{z}_{kp}$ be independent uncertain variables with regular uncertainty distributions $\Psi_{k1}, \Psi_{k2}, \ldots, \Psi_{kn}, \Phi_{k1}, \Phi_{k2}, \ldots, \Phi_{kh}$*

and $\Omega_{k1}, \Omega_{k2}, \ldots, \Omega_{kp}$, respectively. Then the whole stage model (1) is equal to the following form:

$$
\begin{cases}
\max_{u,v,w} \; c_w = \int_0^1 \dfrac{\sum_{t=1}^h u_t \Phi_{ot}^{-1}(\alpha)}{\sum_{i=1}^n v_i \Psi_{oi}^{-1}(1-\alpha)} \mathrm{d}\alpha \\[3mm]
s.t. \quad \int_0^1 \dfrac{\sum_{t=1}^h u_t \Phi_{kt}^{-1}(\alpha)}{\sum_{i=1}^n v_i \Psi_{ki}^{-1}(1-\alpha)} \mathrm{d}\alpha \le 1, \quad k = 1, 2, \ldots, j \\[3mm]
\qquad \int_0^1 \dfrac{\sum_{q=1}^p w_q \Omega_{kq}^{-1}(\alpha)}{\sum_{i=1}^n v_i \Psi_{ki}^{-1}(1-\alpha)} \mathrm{d}\alpha \le 1, \quad k = 1, 2, \ldots, j \\[3mm]
\qquad \int_0^1 \dfrac{\sum_{t=1}^h u_t \Phi_{kt}^{-1}(\alpha)}{\sum_{q=1}^p w_q \Omega_{kq}^{-1}(1-\alpha)} \mathrm{d}\alpha \le 1, \quad k = 1, 2, \ldots, j \\[3mm]
\qquad u = (u_1, u_2, \ldots, u_h) \ge 0 \\[2mm]
\qquad v = (v_1, v_2, \ldots, v_n) \ge 0 \\[2mm]
\qquad w = (w_1, w_2, \ldots, w_p) \ge 0
\end{cases}
\tag{2}
$$

where $\Psi_{o1}, \Psi_{o2}, \ldots, \Psi_{on}, \Phi_{o1}, \Phi_{o2}, \ldots, \Phi_{oh}$ and $\Omega_{o1}, \Omega_{o2}, \ldots, \Omega_{op}$ are the regular uncertainty distributions of $\tilde{x}_{o1}, \tilde{x}_{o2}, \ldots, \tilde{x}_{on}, \tilde{y}_{o1}, \tilde{y}_{o2}, \ldots, \tilde{y}_{oh}$ and $\tilde{z}_{o1}, \tilde{z}_{o2}, \ldots, \tilde{z}_{op}$, respectively.

Proof Since the function $u^T \tilde{y}_k / v^T \tilde{x}_k$ is strictly increasing with respect to $\tilde{y}_k$ and strictly decreasing with respect to $\tilde{x}_k$ for each k. According to Theorem 1, we can conclude that the inverse uncertainty distribution of $u^T \tilde{y}_k / v^T \tilde{x}_k$ is

$$
G_k^{-1}(\alpha) = \frac{\sum_{t=1}^h u_t \Phi_{kt}^{-1}(\alpha)}{\sum_{i=1}^n v_i \Psi_{ki}^{-1}(1-\alpha)}.
$$

From Theorem 2, we obtain

$$
E\left[\frac{u^T \tilde{y}_k}{v^T \tilde{x}_k} \right] = \int_0^1 \frac{\sum_{t=1}^h u_t \Phi_{kt}^{-1}(\alpha)}{\sum_{i=1}^n v_i \Psi_{ki}^{-1}(1-\alpha)} \mathrm{d}\alpha, \quad k = 1, 2, \ldots, j.
$$

Similarly, for each k, according to the above proof, it can be drawn at once that the function

$$
E\left[\frac{w^T \tilde{z}_k}{v^T \tilde{x}_k} \right] = \int_0^1 \frac{\sum_{q=1}^p w_q \Omega_{kq}^{-1}(\alpha)}{\sum_{i=1}^n v_i \Psi_{ki}^{-1}(1-\alpha)} \mathrm{d}\alpha
$$

is strictly increasing with respect to $\tilde{z}_k$ and strictly decreasing with respect to $\tilde{x}_k$. Identically, for each k, it can be inferred very quickly that the function

$$
E\left[\frac{u^T \tilde{y}_k}{w^T \tilde{z}_k} \right] = \int_0^1 \frac{\sum_{t=1}^h u_t \Phi_{kt}^{-1}(\alpha)}{\sum_{q=1}^p w_q \Omega_{kq}^{-1}(1-\alpha)} \mathrm{d}\alpha
$$

is strictly increasing with respect to $\tilde{y}_k$ and strictly decreasing with respect to $\tilde{z}_k$.

The theorem has been proved. $\qquad\square$

3.2 The first stage model

The first stage model is established based on the whole stage model. Therefore, the first stage model is trying to find the maximum efficiency of the first stage while maintaining the efficiency of the whole stage. Then the first stage model is formulated as follows:

$$
\begin{cases}
\max_{u,v,w} \; c_1 = E\left[\dfrac{w^T \tilde{z}_o}{v^T \tilde{x}_o} \right] \\[3mm]
s.t. \quad E\left[\dfrac{u^T \tilde{y}_0}{v^T \tilde{x}_0} \right] = c_w \\[3mm]
\qquad E\left[\dfrac{u^T \tilde{y}_k}{v^T \tilde{x}_k} \right] \le 1, \quad k = 1, 2, \ldots, j \\[3mm]
\qquad E\left[\dfrac{w^T \tilde{z}_k}{v^T \tilde{x}_k} \right] \le 1, \quad k = 1, 2, \ldots, j \\[3mm]
\qquad E\left[\dfrac{u^T \tilde{y}_k}{w^T \tilde{z}_k} \right] \le 1, \quad k = 1, 2, \ldots, j \\[2mm]
\qquad u \ge 0 \\[1mm]
\qquad v \ge 0 \\[1mm]
\qquad w \ge 0.
\end{cases}
\tag{3}
$$

Definition 6 *(Uncertain DEA Efficiency)* If and only if the optimal solution $\overline{c_1}$ of the model (3) can reach 1, DMU$_o$ can be considered as efficient.

The efficiency of DMU$_0$ can be calculated by evaluating the optimal value of the first stage model. The equivalent form of model (3) is proved as below:

Theorem 4 *For each k, let inputs $\tilde{x}_{k1}, \tilde{x}_{k2}, \ldots, \tilde{x}_{kn}$, outputs $\tilde{y}_{k1}, \tilde{y}_{k2}, \ldots, \tilde{y}_{kh}$ and intermediates $\tilde{z}_{k1}, \tilde{z}_{k2}, \ldots, \tilde{z}_{kp}$ be independent uncertain variables with regular uncertainty distributions $\Psi_{k1}, \Psi_{k2}, \ldots, \Psi_{kn}, \Phi_{k1}, \Phi_{k2}, \ldots, \Phi_{kh}$ and $\Omega_{k1}, \Omega_{k2}, \ldots, \Omega_{kp}$, respectively. Then the first stage model (3) is equal to the following form:*

$$
\begin{cases}
\max_{u,v,w} c_1 = \int_0^1 \dfrac{\sum_{q=1}^p w_q \Omega_{oq}^{-1}(\alpha)}{\sum_{i=1}^n v_i \Psi_{oi}^{-1}(1-\alpha)} \, d\alpha \\[2ex]
s.t. \quad \int_0^1 \dfrac{\sum_{t=1}^h u_t \Phi_{ot}^{-1}(\alpha)}{\sum_{i=1}^n v_i \Psi_{oi}^{-1}(1-\alpha)} \, d\alpha = c_w \\[2ex]
\int_0^1 \dfrac{\sum_{t=1}^h u_t \Phi_{kt}^{-1}(\alpha)}{\sum_{i=1}^n v_i \Psi_{ki}^{-1}(1-\alpha)} \, d\alpha \le 1, \quad k = 1, 2, \ldots, j \\[2ex]
\int_0^1 \dfrac{\sum_{q=1}^p w_q \Omega_{kq}^{-1}(\alpha)}{\sum_{i=1}^n v_i \Psi_{ki}^{-1}(1-\alpha)} \, d\alpha \le 1, \quad k = 1, 2, \ldots, j \\[2ex]
\int_0^1 \dfrac{\sum_{t=1}^h u_t \Phi_{kt}^{-1}(\alpha)}{\sum_{q=1}^p w_q \Omega_{kq}^{-1}(1-\alpha)} \, d\alpha \le 1, \quad k = 1, 2, \ldots, j \\[2ex]
u = (u_1, u_2, \ldots, u_h) \ge 0 \\[1ex]
v = (v_1, v_2, \ldots, v_n) \ge 0 \\[1ex]
w = (w_1, w_2, \ldots, w_p) \ge 0
\end{cases} \tag{4}
$$

where $\Psi_{o1}, \Psi_{o2}, \ldots, \Psi_{on}, \Phi_{o1}, \Phi_{o2}, \ldots, \Phi_{oh}$ and $\Omega_{o1}, \Omega_{o2}, \ldots, \Omega_{op}$ are the regular uncertainty distributions of $\tilde{x}_{o1}, \tilde{x}_{o2}, \ldots, \tilde{x}_{on}, \tilde{y}_{o1}, \tilde{y}_{o2}, \ldots, \tilde{y}_{oh}$ and $\tilde{z}_{o1}, \tilde{z}_{o2}, \ldots, \tilde{z}_{op}$, respectively.

Proof Since the function $u^T \tilde{y}_k / v^T \tilde{x}_k$ is strictly increasing with respect to $\tilde{y}_k$ and strictly decreasing with respect to $\tilde{x}_k$ for each k. According to Theorem 1, we can conclude that the inverse uncertainty distribution of $u^T \tilde{y}_k / v^T \tilde{x}_k$ is

$$
J_k^{-1}(\alpha) = \frac{\sum_{t=1}^h u_t \Phi_{kt}^{-1}(\alpha)}{\sum_{i=1}^n v_i \Psi_{ki}^{-1}(1-\alpha)}.
$$

From Theorem 2, we obtain

$$
E\left[\frac{u^T \tilde{y}_k}{v^T \tilde{x}_k}\right] = \int_0^1 \frac{\sum_{t=1}^h u_t \Phi_{kt}^{-1}(\alpha)}{\sum_{i=1}^n v_i \Psi_{ki}^{-1}(1-\alpha)} \, d\alpha, \quad k = 1, 2, \ldots, j.
$$

Similarly, for each k, according to the above proof, it can be drawn at once that the function

$$
E\left[\frac{w^T \tilde{z}_k}{v^T \tilde{x}_k}\right] = \int_0^1 \frac{\sum_{q=1}^p w_q \Omega_{kq}^{-1}(\alpha)}{\sum_{i=1}^n v_i \Psi_{ki}^{-1}(1-\alpha)} \, d\alpha
$$

is strictly increasing with respect to $\tilde{z}_k$ and strictly decreasing with respect to $\tilde{x}_k$. Identically, for each k, it can be inferred very quickly that the function

$$
E\left[\frac{u^T \tilde{y}_k}{w^T \tilde{z}_k}\right] = \int_0^1 \frac{\sum_{t=1}^h u_t \Phi_{kt}^{-1}(\alpha)}{\sum_{q=1}^p w_q \Omega_{kq}^{-1}(1-\alpha)} \, d\alpha
$$

is strictly increasing with respect to $\tilde{y}_k$ and strictly decreasing with respect to $\tilde{z}_k$.

The theorem has been proved. $\square$

3.3 The second stage model

Similarly, the efficiency of the second stage model is based on the efficiencies of the first stage and the whole stage. While maintaining the efficiencies of the whole stage and the first stage, the efficiency of the second stage reaches the maximum. The second stage model is presented as follows:

$$
\begin{cases}
\max_{u,v,w} c_2 = E\left[\dfrac{u^T \tilde{y}_o}{w^T \tilde{z}_o}\right] \\[2ex]
s.t. \quad E\left[\dfrac{u^T \tilde{y}_0}{v^T \tilde{x}_0}\right] = c_w \\[2ex]
E\left[\dfrac{w^T \tilde{z}_0}{v^T \tilde{x}_0}\right] = c_1 \\[2ex]
E\left[\dfrac{u^T \tilde{y}_k}{v^T \tilde{x}_k}\right] \le 1, \quad k = 1, 2, \ldots, j \\[2ex]
E\left[\dfrac{w^T \tilde{z}_k}{v^T \tilde{x}_k}\right] \le 1, \quad k = 1, 2, \ldots, j \\[2ex]
E\left[\dfrac{u^T \tilde{y}_k}{w^T \tilde{z}_k}\right] \le 1, \quad k = 1, 2, \ldots, j \\[1ex]
u \ge 0 \\[1ex]
v \ge 0 \\[1ex]
w \ge 0.
\end{cases} \tag{5}
$$

Definition 7 (*Uncertain DEA Efficiency*) If and only if the optimal solution $\overline{c_2}$ of the model (5) can reach 1, DMU$_o$ is considered as efficient.

The efficiency of DMU$_0$ can be calculated by evaluating the optimal value of the second stage model. The equivalent form of model (5) is proved as below:

Theorem 5 *For each k, let inputs $\tilde{x}_{k1}, \tilde{x}_{k2}, \ldots, \tilde{x}_{kn}$, outputs $\tilde{y}_{k1}, \tilde{y}_{k2}, \ldots, \tilde{y}_{kh}$ and intermediates $\tilde{z}_{k1}, \tilde{z}_{k2}, \ldots, \tilde{z}_{kp}$ be independent uncertain variables with regular uncertainty distributions $\Psi_{k1}, \Psi_{k2}, \ldots, \Psi_{kn}, \Phi_{k1}, \Phi_{k2}, \ldots, \Phi_{kh}$ and $\Omega_{k1}, \Omega_{k2}, \ldots, \Omega_{kp}$, respectively. Then the second stage model (5) is equal to the following form:*

$$\begin{cases}
\max\limits_{u,v,w} c_2 = \int_0^1 \dfrac{\sum_{t=1}^h u_t \Phi_{ot}^{-1}(\alpha)}{\sum_{q=1}^p w_q \Omega_{oq}^{-1}(1-\alpha)}\,\mathrm{d}\alpha \\[2mm]
s.t. \quad \int_0^1 \dfrac{\sum_{t=1}^h u_t \Phi_{ot}^{-1}(\alpha)}{\sum_{i=1}^n v_i \Psi_{oi}^{-1}(1-\alpha)}\,\mathrm{d}\alpha = c_w \\[2mm]
\quad \int_0^1 \dfrac{\sum_{q=1}^p w_q \Omega_{oq}^{-1}(\alpha)}{\sum_{i=1}^n v_i \Psi_{oi}^{-1}(1-\alpha)}\,\mathrm{d}\alpha = c_1 \\[2mm]
\quad \int_0^1 \dfrac{\sum_{t=1}^h u_t \Phi_{kt}^{-1}(\alpha)}{\sum_{i=1}^n v_i \Psi_{ki}^{-1}(1-\alpha)}\,\mathrm{d}\alpha \le 1, \quad k = 1, 2, \dots, j \\[2mm]
\quad \int_0^1 \dfrac{\sum_{q=1}^p w_q \Omega_{kq}^{-1}(\alpha)}{\sum_{i=1}^n v_i \Psi_{ki}^{-1}(1-\alpha)}\,\mathrm{d}\alpha \le 1, \quad k = 1, 2, \dots, j \\[2mm]
\quad \int_0^1 \dfrac{\sum_{t=1}^h u_t \Phi_{kt}^{-1}(\alpha)}{\sum_{q=1}^p w_q \Omega_{kq}^{-1}(1-\alpha)}\,\mathrm{d}\alpha \le 1, \quad k = 1, 2, \dots, j \\[2mm]
\quad u = (u_1, u_2, \dots, u_h) \ge 0 \\[1mm]
\quad v = (v_1, v_2, \dots, v_n) \ge 0 \\[1mm]
\quad w = (w_1, w_2, \dots, w_p) \ge 0
\end{cases} \tag{6}$$

where $\Psi_{o1}, \Psi_{o2}, \dots, \Psi_{on}, \Phi_{o1}, \Phi_{o2}, \dots, \Phi_{oh}$ and $\Omega_{o1}, \Omega_{o2}, \dots, \Omega_{op}$ are the regular uncertainty distributions of $\tilde{x}_{o1}, \tilde{x}_{o2}, \dots, \tilde{x}_{on}, \tilde{y}_{o1}, \tilde{y}_{o2}, \dots, \tilde{y}_{oh}$ and $\tilde{z}_{o1}, \tilde{z}_{o2}, \dots, \tilde{z}_{op}$, respectively.

Proof Since the function $u^T \tilde{y}_k / v^T \tilde{x}_k$ is strictly increasing with respect to $\tilde{y}_k$ and strictly decreasing with respect to $\tilde{x}_k$ for each k. According to Theorem 1, we can conclude that the inverse uncertainty distribution of $u^T \tilde{y}_k / v^T \tilde{x}_k$ is

$$Q_k^{-1}(\alpha) = \frac{\sum_{t=1}^h u_t \Phi_{kt}^{-1}(\alpha)}{\sum_{i=1}^n v_i \Psi_{ki}^{-1}(1-\alpha)}.$$

From Theorem 2, we obtain

$$E\left[\frac{u^T \tilde{y}_k}{v^T \tilde{x}_k}\right] = \int_0^1 \frac{\sum_{t=1}^h u_t \Phi_{kt}^{-1}(\alpha)}{\sum_{i=1}^n v_i \Psi_{ki}^{-1}(1-\alpha)}\,\mathrm{d}\alpha, \quad k = 1, 2, \dots, j.$$

Similarly, for each k, according to the above proof, it can be drawn at once that the function

$$E\left[\frac{w^T \tilde{z}_k}{v^T \tilde{x}_k}\right] = \int_0^1 \frac{\sum_{q=1}^p w_q \Omega_{kq}^{-1}(\alpha)}{\sum_{i=1}^n v_i \Psi_{ki}^{-1}(1-\alpha)}\,\mathrm{d}\alpha$$

is strictly increasing with respect to $\tilde{z}_k$ and strictly decreasing with respect to $\tilde{x}_k$. Identically, for each k, it can be inferred very quickly that the function

$$E\left[\frac{u^T \tilde{y}_k}{w^T \tilde{z}_k}\right] = \int_0^1 \frac{\sum_{t=1}^h u_t \Phi_{kt}^{-1}(\alpha)}{\sum_{q=1}^p w_q \Omega_{kq}^{-1}(1-\alpha)}\,\mathrm{d}\alpha$$

is strictly increasing with respect to $\tilde{y}_k$ and strictly decreasing with respect to $\tilde{z}_k$.

The theorem has been verified. $\qquad\square$

4 A numerical example

In this section, the expected results of the uncertain two-stage network DEA models we have established are shown as follows. We will use a set of data to test the uncertain two-stage network DEA models. The original data which is shown in Table 1 is composed of three uncertain inputs, three uncertain intermediates and three uncertain outputs of five DMUs. The results of the efficiency evaluation are shown in Table 2, including the efficiency values for the whole stage, the first stage and the second stage.

The efficiency of the whole stage is the chief observation object for decision makers, which enables them to understand the network system from a holistic perspective. As is shown in Table 2, we can observe the efficiencies (c_w^*) of the whole stage of five DMUs by the uncertain two-stage network DEA models. It can be clearly seen that the optimal value (c_w^*) of DMU_1 reaches 1.0000 which means that DMU_1 is efficient at the whole stage. By comparison, the efficiency (c_w^*) of the whole stage of the DMU_2 is 0.2549 which is lower than 1.0000 and considered to be inefficient at the whole stage. The analyses of the DMU_3, DMU_4 and DMU_5 can be shown in the same way.

After identifying the efficiency of the whole stage, the efficiencies of the internal sub-stages also need to be analyzed in order to find out the root causes of the inefficiency of the network system. As shown in Table 2, the efficiencies (c_1^*, c_2^*) of two sub-stages of DMU_1 are both 1.0000, which is the real reason why DMU_1 is efficient at the whole stage. That is because only when the efficiency of each sub-stage reaches 1.0000, the whole stage can be efficient as a whole. In other words, the efficiencies of the two sub-stages of DMU_1 can correspond to the relationship between the two sub-stages that the efficiency of the whole stage is equal to the product of the efficiencies of the two sub-stages. For decision makers, efficient DMU_1 can be viewed as a benchmark to further improve other inefficient DMUs.

Different from DMU_1, the efficiency (c_1^*) of the first stage of DMU_2 is 1.0000 while the efficiency (c_2^*) of the second stage is 0.2545. It is easy to draw the conclusion that the inefficiency of the whole stage of DMU_2 is due to the inefficiency of the second stage. Similarly, the efficiency (c_2^*) of the second stage of DMU_3 is 1.0000, whereas the efficiency (c_1^*) of the first stage is 0.1779, which can be regarded as the root cause of the inefficiency of DMU_3. Through the comparison of the above two DMUs, we can easily find out that DMU_2 and DMU_3 are similarly inefficient at the whole stage. But the reasons for the inefficiency are quite different, which are due to the different internal inefficient stages. Since the reasons for each DMU being inefficient are different, it is very important to be able to identify the inefficient sub-stages, which can prevent decision makers from allocating finite resources to efficient stages and neglecting inefficient stages.

Table 1 Five DMUs with three inputs, three intermediates and three outputs, in which $\mathcal{L}(a, b)$ represents linear uncertain variable and $\mathcal{Z}(a, b, c)$ represents zigzag uncertain variable

DMU_k	1	2	3	4	5
Input1	$\mathcal{L}(4, 6)$	$\mathcal{L}(10, 11)$	$\mathcal{L}(28, 30)$	$\mathcal{L}(30, 31)$	$\mathcal{L}(14, 16)$
Input2	$\mathcal{L}(5, 6)$	$\mathcal{L}(9, 11)$	$\mathcal{L}(29, 30)$	$\mathcal{L}(30, 32)$	$\mathcal{L}(15, 16)$
Input3	$\mathcal{L}(4, 5)$	$\mathcal{L}(10, 12)$	$\mathcal{L}(30, 31)$	$\mathcal{L}(28, 30)$	$\mathcal{L}(14, 15)$
Intermediates1	$\mathcal{Z}(28, 29, 30)$	$\mathcal{Z}(59, 60, 62)$	$\mathcal{Z}(27, 28, 30)$	$\mathcal{Z}(30, 31, 32)$	$\mathcal{Z}(38, 39, 40)$
Intermediates2	$\mathcal{Z}(28, 29, 30)$	$\mathcal{Z}(59, 60, 61)$	$\mathcal{Z}(28, 29, 31)$	$\mathcal{Z}(28, 30, 31)$	$\mathcal{Z}(38, 40, 42)$
Intermediates3	$\mathcal{Z}(30, 31, 32)$	$\mathcal{Z}(60, 61, 62)$	$\mathcal{Z}(30, 31, 32)$	$\mathcal{Z}(27, 28, 30)$	$\mathcal{Z}(40, 41, 42)$
Output1	$\mathcal{L}(400, 402)$	$\mathcal{L}(199, 200)$	$\mathcal{L}(398, 400)$	$\mathcal{L}(298, 300)$	$\mathcal{L}(199, 200)$
Output2	$\mathcal{L}(399, 400)$	$\mathcal{L}(200, 202)$	$\mathcal{L}(400, 401)$	$\mathcal{L}(300, 302)$	$\mathcal{L}(200, 201)$
Output3	$\mathcal{L}(403, 404)$	$\mathcal{L}(199, 200)$	$\mathcal{L}(396, 398)$	$\mathcal{L}(300, 301)$	$\mathcal{L}(198, 200)$

Table 2 The results of evaluating the efficiencies by the uncertain two-stage network DEA models

DMU_k	1	2	3	4	5
Optimal value c_w^* for the whole stage	1.0000	0.2549	0.1778	0.0509	0.0410
Optimal value c_1^* for the first stage	1.0000	1.0000	0.1779	0.1331	0.4439
Optimal value c_2^* for the second stage	1.0000	0.2545	1.0000	0.3828	0.0924

By further observation, we can find out more things worthy of our attention from DMU_4 and DMU_5. The efficiency (c_1^*) of the first stage of DMU_4 is 0.1331 while the efficiency (c_2^*) of the second stage is 0.2545. The result for DMU_5 can be interpreted in a similar way. Different from DMU_2 and DMU_3, the DMU_4 and DMU_5 are not only inefficient at the whole stage, but also at all the sub-stages. It is a common situation faced by most policy makers. However, due to the limited resources, it is impossible for policy makers to invest resources in all the inefficient sub-stages to improve performance. Therefore, decision makers can firstly choose to improve the inefficient sub-stages with lower efficiency. This is because the inefficient sub-stage with lower efficiency is more likely to be improved when the same resources are invested. Moreover, a slight improvement in the relatively low efficiency of the sub-stage will result in a higher improvement of the overall performance. This fully demonstrates the importance of identifying efficiencies of sub-stages for decision makers.

5 Conclusions

In the research of the economics, the key element to improve the performance of network system is to identify the internal root causes of the inefficiency. Moreover, due to the novelty of evaluation indexes, the traditional DEA models cannot deal with the emerging indexes such as social welfare, which cannot be quantified accurately. Thus, based on uncertainty theory, this paper establishes the uncertain two-stage network DEA models, which have the most basic structure that can be extended to other more complex network models. The new models can accurately identify the root causes of the inefficiency of the network system and deal with the problem of inaccurate data in uncertain environments.

In addition to theoretical expansion, the models can also be used in empirical application. In reality, there are many systems with complex network structure, such as maritime supply chain. Generally speaking, the maritime supply chain consists of two sub-stages: ports and shipping companies. For decision makers, in addition to identifying the overall efficiency of the maritime supply chain, it is also important to explore its internal structure and hence identify the internal root causes of the inefficiency. By the uncertain two-stage network DEA models, decision makers can identify the inefficient supply chain members and make the right resource allocation decisions in uncertain environment.

In general, the contributions of this paper are as follows. Firstly, this article expands the application of uncertainty theory in the field of DEA. Secondly, this paper expands the evaluation method of unobservable variables and the scope of application of DEA models by establishing the uncertain two-stage network DEA models. Despite some theoretical and practical contributions, this research concentrates on only a simple two-stage network structure. In practice, the parallel structure is also common, especially for enterprises with multiple homogenous sub factories. Therefore, in the future, our research will focus on the efficiency evaluation of parallel structure systems with inaccurate data.

Compliance with ethical standards

References

Banker RD (1993) Maximum likelihood, consistency and DEA: statistical foundations. Manag Sci 39(10):1265–1273

Charnes A, Cooper WW, Rhodes E (1978) Measuring the efficiency of decision making units. Eur J Oper Res 2(6):429–444

Chen C (2009) A network-DEA model with new efficiency measures to incorporate the dynamic effect in production networks. Eur J Oper Res 194(3):687–699

Chen Y, Zhu J (2004) Measuring information technology's indirect impact on firm performance. Inf Technol Manag J 5(1–2):9–22

Cook WD, Zhu J (2010) Network DEA: additive efficiency decomposition. Eur J Oper Res 207(2):1122–1129

Cook WD, Tone K, Zhu J (2014) Data envelopment analysis: prior to choosing a model. Omega 44:1–4

Dyckhoff H, Allen K (2001) Measuring ecological efficiency with data envelopment analysis (DEA). Eur J Oper Res 132(2):312–325

Esmaeilzadeh A, Matin RK (2019) Multi-period efficiency measurement of network production systems. Measurement 134:835–844

Fåre R, Grosskopf S (2000) Network DEA. Soc Econ Plan Sci 34(1):35–49

Hatami-Marbini A, Saati S (2018) Efficiency evaluation in two-stage data envelopment analysis under a fuzzy environment: a common-weights approach. Appl Soft Comput 72:156–165

Jiang B, Lio W, Li X (2019) An uncertain DEA model for scale efficiency evaluation. IEEE Trans Fuzzy Syst 27(8):1616–1624

Jiang B, Zou Z, Lio W, Li J (2020) The uncertain DEA models for specific scale efficiency identification. J Intell Fuzzy Syst 38:3403–3417

Kao C (2011) Efficiencies of two-stage systems with fuzzy data. Fuzzy Sets Syst 176(1):20–35

Kao C (2014) Network data envelopment analysis: a review. Eur J Oper Res 239(1):1–16

Kao C (2016) Efficiency decomposition and aggregation in network data envelopment analysis. Eur J Oper Res 255(3):778–786

Kao C, Hwang S (2008) Efficiency decomposition in two-stage data envelopment analysis: an application to non-life insurance companies in Taiwan. Eur J Oper Res 185(1):418–429

Lio W, Liu B (2018) Uncertain data envelopment analysis with imprecisely observed inputs and outputs. Fuzzy Optim Decis Mak 17:357–373

Liu B (2007) Uncertainty theory, 2nd edn. Springer, Berlin

Liu B (2009) Theory and practice of uncertain programming, 2nd edn. Springer, Berlin

Liu B (2009) Some research problems in uncertainty theory. J Uncertain Syst 3(1):3–10

Liu B (2010) Uncertainty theory: a branch of mathematics for modeling human uncertainty. Springer, Berlin

Liu B (2012) Why is there a need for uncertainty theory. J Uncertain Syst 6:3–10

Liu B, Chen X (2015) Uncertain multiobjective programming and uncertain goal programming. J Uncertain Anal Appl 3:10

Liu Y, Ha M (2010) Expected value of function of uncertain variables. J Uncertain Syst 4(3):181–186

Liu B, Yao K (2015) Uncertain multilevel programming: algorithm and application. Comput Ind Eng 89:235–240

Olesen OB, Petersen NC (2016) Stochastic data envelopment analysis-a review. Eur J Oper Res 251(1):2–21

Sadjadi SJ, Omrani H (2008) Data envelopment analysis with uncertain data: an application for iranian electricity distribution companies. Energy Policy 36(11):4247–4254

Seiford LM, Zhu J (1999) Profitability and marketability of the top 55 US commercial banks. Manag Sci 45(9):1270–1288

Sengupta JK (1982) Efficiency measurement in stochastic input-output systems. Int J Syst Sci 13(3):273–287

Sexton TR, Lewis HF (2003) Two-stage DEA: an application to major league baseball. J Prod Anal 19(2–3):227–249

Sueyoshi T (2000) Stochastic DEA for restructure strategy: an application to a Japanese petroleum company. Omega 28(4):385–398

Tavana M (2018) Efficiency decomposition and measurement in two-stage fuzzy DEA models using a bargaining game approach. Comput Ind Eng 118:394–408

Wanke P, Kalam Azad A, Emrouznejad A (2018) Efficiency in BRICS banking under data vagueness: a two-stage fuzzy approach. Glob Financ J 35:58–71

Wen M, Guo L, Kang R, Yang Y (2014) Data envelopment analysis with uncertain inputs and outputs. J Appl Math 2:1–7