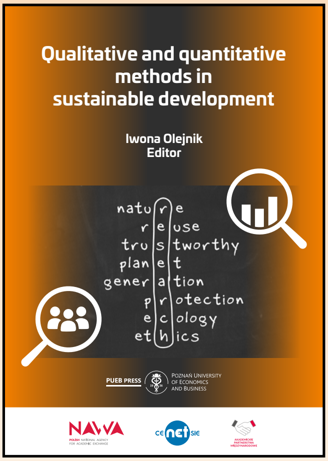


Qualitative and quantitative methods in sustainable development

Iwona Olejnik
Editor

POLISH NATIONAL AGENCY FOR ACADEMIC EXCHANGE

AKADEMICKIE PARTNERSTWA MIĘDZYNARODOWE

eISBN 978-83-8211-072-2
https://doi.org/10.18559/978-83-8211-072-2

PUEB PRESS
Poznań University of Economics and Business

Qualitative and quantitative methods in sustainable development, pp. 179-205
https://doi.org/10.18559/978-83-8211-072-2/05

5.

DATA ENVELOPMENT ANALYSIS METHODS IN SUSTAINABLE AGRICULTURAL DEVELOPMENT RESEARCH

Katarzyna Smędzik-Ambroży
Poznań University of Economics and Business

Agnieszka Sapa
Poznań University of Economics and Business

Abstract: Sustainable development of business entities can be analysed in terms of three dimensions, i.e., economic, social and environmental ones. The economic dimension of sustainable development can be assessed, *inter alia*, by entities' technical efficiency defined as the relation of outputs to inputs. One of the methods that is used to assess the technical efficiency of business entities compared to other entities is the Data Envelopment Analysis (DEA) method.

The aim of the chapter is to determine the relative technical efficiency of representative agricultural farms from the individual European Union countries in 2018. Moreover, the scale efficiency indexes and the area of scale effects (increasing or decreasing) of the analysed farms were also determined. In the study the data from the Farm Accountancy Data Network (FADN) for 2018 were applied.

In order to achieve the assumed research goals, the input-oriented DEA model was used, and the technical efficiency indexes of farms were estimated with the assumption of constant return to scale (CRS) and variable return to scale (VRS). This allowed, among others, for indicating the countries with farms achieving the highest technical efficiency (Belgium, Spain, Italy, Malta and Netherlands assuming CRS, and Belgium, Spain, Italy, Malta and Netherlands, Greece, Ireland, Romania and Slovenia assuming VRS), the lowest technical efficiency (the Czech Republic and Slovakia) within surveyed group of farms. All relatively inefficient farms (except Slovakia) functioned in the area of increasing economies of scale.

Keywords: DEA method, economics sustainability, effect of scale, farms, technical efficiency.

Katarzyna Smędzik-Ambroży, Agnieszka Sapa

5.1. DEA—theoretical background

In production process inputs are converted into effects. Efficiency describes how effectively the company transforms inputs into effects. The measurement of efficiency is very important for the company because it informs, e.g., if the assumed goals were achieved or not. Moreover, it allows to compare the achieved efficiency level of a particular company with the results of other similar units. One of the popular methods used to determine the relative efficiency of the examined units is the Data Envelopment Analysis (DEA) method. The DEA method was proposed by Charnes, Cooper and Rhodes in 1978, the so-called CCR model (abbreviation of the first letters of the authors' surnames) (Charnes, Cooper, & Rhodes, 1978). It is an extension of Farrell's (1957) work on technical efficiency estimation.

The DEA method is used to measure the relative efficiency of the selected objects (units) in a situation where the units use many inputs simultaneously and achieve many effects. Before proceeding with further analysis, the notion of relative efficiency and the difference between efficiency (i.e., the relation of effects to inputs) and relative efficiency should be introduced. Table 5.1 presents four objects and each of them is described by one effect (y) and one input (x). The highest efficiency is attributed to unit A, as it transforms a given input into an effect in the best way. To determine the relative efficiency index, all examined units should be compared to the best units, in this case to object A.

Table 5.1. Efficiency and related efficiency of the selected objects A–D

Unit	Input (x)	Output (y)	Efficiency (y/x)	Related efficiency indicator (%)
A	18	125	6,94	100
B	16	44	2,75	40
C	17	80	4,71	68
D	11	23	2,09	30

Source: (Domagała, 2007, p. 24–25).

The efficient unit, i.e., the one whose relative efficiency index is 100% (unit A), determines the efficiency frontier (also the production possibilities curve of the examined objects, Figure 5.1). All units on this curve achieve a relative efficiency equal to 100%, i.e., they are efficient units. The remaining units are inefficient units, e.g., units C, D, E (see Figure 5.1). When the efficiency frontier is indicated, it can be determined how the inefficient units can approach the efficiency frontier. The improvement of the relative efficiency index of inefficient units can be reached by:

- increasing (maximizing) the effects with unchanged inputs (the so-called outputs orientation; unit C'),
- reducing (minimizing) inputs for given effects (the so-called inputs orientation; unit C"),
- the lack of orientation on effects or inputs (the so-called mixed approach; units lying between units C" and C').

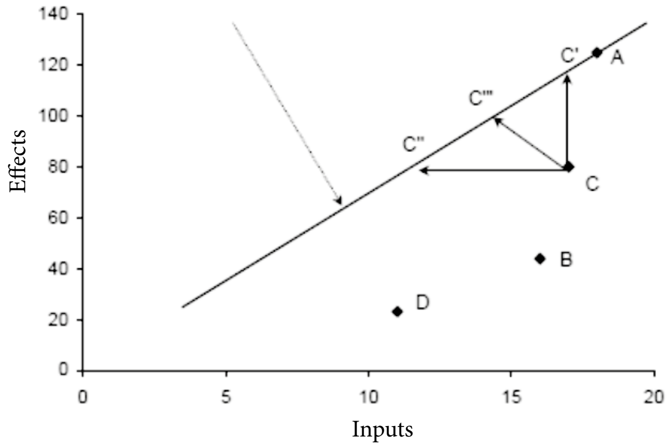


Figure 5.1. The production possibilities curve

Source: (Domagała, 2007, pp. 24–25).

In practice, units are rarely described by one input and one effect. This means that it becomes impossible to assess the efficiency of the examined units, as in the example above (Table 5.1, Figure 5.1). Then the DEA method is useful, as in this method it is assumed that the measure of relative efficiency is the relation of the weighted sum of the outputs and the weighted sum of the inputs. This is a relative efficiency determined for a specific set of objects called DMUs (Decision Making Units).

$$relative\, efficiency = \frac{\text{weighted sum of the outputs}}{\text{weighted sum of the inputs}}$$

The DEA method is a frontier non-parametric method of estimating efficiency. This method allows for identification of the efficiency frontier (the production possibilities curve). The efficiency frontier is constructed in a non-parametric manner using linear programming techniques. This means that the efficiency frontier is spread over the best units (efficient DMUs, called frontiers) in a given studied group of units. Thus, the efficiency frontier is not spread on units determined by a specific production function or by the adopted specific values of inputs and outputs. It is difficult to indicate both an efficiency benchmark level and an objective (functional) relationship between inputs and outputs. So, identification of the units' efficiency only within the group of units and setting out the efficiency benchmarks based on the best units in the examined group can be a useful solution. Since the DEA method does not require the adoption of many predetermined limitations and assumptions, it has become popular and is widely used in social science research, both at the microeconomic and macroeconomic levels.

The DEA method enables to rank the surveyed units according to their efficiency indexes and to determine the relative differences between them in this respect. The object with the highest efficiency index (the so-called frontier) is selected within an examined group of units. The remaining units from the analysed group are considered inefficient units. The efficient unit is the reference (the benchmark) for the other objects, the efficiency indicators of which were evaluated in relation to this efficient unit.

The basic DEA-CCR model (the input-oriented model) can be presented as:
where:

$$\max h_0 = \frac{\sum_{r=1}^{s} u_r y_{ro}}{\sum_{i=1}^{m} v_i x_{io}}$$

assumed that:

$$x = \frac{\sum_{r=1}^{s} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} \le 1; \; j = 1, \dots, n,$$

$$v_r v_i \ge 0; \quad r = 1, \dots, s; \quad i = 1, \dots, m.$$

n—number of decision-making units DMUs ($j = 1, \dots, n$),
s—number of outputs ($r = 1, \dots, s$),
m—number of inputs ($i = 1, \dots, m$),
x_{ij}—value of i input of decision-making unit j,
y_{rj}—value of r output of decision-making unit j,
v_i—decision variable; weight related to i input,
u_r—decision variable; weight related to r input,
o—index of examined decision making unit, $1 \le o \le n$,
h_o—efficiency index of object o.

Technical efficiency indicators, estimated using the DEA method, range from 0 to 1. Depending on whether the input-oriented or the output-oriented DEA model is used, the difference between the value of the estimated indicator for a given object and unity means:

- how much a given object should proportionally reduce its inputs without changing its outputs in order to achieve full efficiency (input-oriented model),

– how much a given object should proportionally increase its outputs without changing the level of used inputs in order to achieve full efficiency (output-oriented model).

For example, if in input-oriented DEA model the estimated value of the technical efficiency index for a given object is 0.7, it means that the object should reduce its inputs by 30%. In other words, the object should reduce their inputs to 70% of the current inputs level to manage its inputs in a fully efficient way and to achieve the same relation of outputs to inputs as the best units (frontiers).[1] In the case of output-oriented DEA model, the estimated technical efficiency index at the level of 0.7 for a given object means that it should increase its effects by 30% in order to achieve full efficiency with the given inputs.

It is worth underlining that regardless of the model orientation (input or output), the higher the efficiency index of a given unit, the higher its efficiency. It should also be emphasized that the achieved efficiency indicators in DEA methods are always relative, so adding a new unit to the analysis may change their values.

Analysis with DEA can assume constant or variable returns to scale. The basic version of DEA model assumes constant returns to scale (DEA–CRS). CRS implies linearity between inputs and outputs, meaning that doubling the inputs used, will double the outputs, which is rare in practice. Banker, Charnes and Cooper modified the basic DEA model by introducing the assumption of variable returns to scale (VRS). This is BCC model (DEA–BCC). Thus, the adoption of this approach allows to estimate the technical efficiency without the CRS assumption.

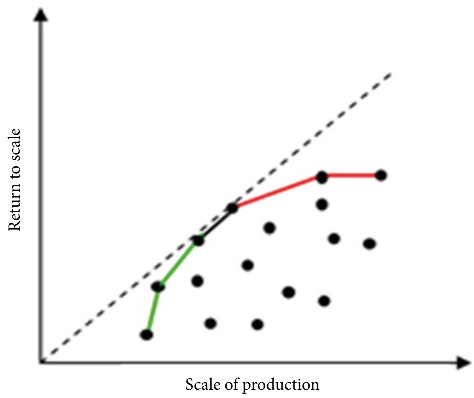


Figure 5.2. Relation between the scale of production and return to scale

Source: Own elaboration based on (Guzik, 2009; Czyżewski, Smędzik-Ambroży & Mrówczyńska-Kamińska, 2020).

[1] It should be emphasized that in the case of the input-oriented model, the main assumption is full complementarity of inputs and zero degree of their substitution (Guzik, 2009).

Considering VRS assumption, it should be stated that a company may experience increasing return to scale (IRS) or decreasing return to scale (DRS). IRS means a situation where increasing the inputs by t-times affects the outputs increase by more than t-times (green line, Figure 5.2). In some companies, especially with a very huge scale of production, the decreasing return to scale (DRS) can be observed. It results from the law of diminishing marginal productivity and is an unfavourable situation (red line, Figure 5.2). It means that an increase in inputs by t-times is accompanied by outputs increase by less than t-times. The company's goal is to achieve the optimal production scale and after reaching that the company can operate within decreasing return to scale conditions. The type of return to scale (CRS or DRS) in which a given unit operates can also be determined using the DEA method. In this case, the model with non-increasing returns to scale (NIRS) can be applied.

The next index the DEA enables to estimate is the production scale efficiency index. It is the relation between the unit efficiency index assuming constant return to scale and the unit efficiency index assuming variable return to scale. The production scale efficiency index takes values from 0 to 1. The index informs how much less inputs could be used if the outputs volume were optimal. The index equal to 1 means the optimal production (output) scale in a given group of units.

Summarizing the above, it can be concluded that the essence of the DEA model, is, among others:

- searching for the best units in the examined group of objects (the best ones create the efficiency frontier) in a situation where there are many outputs and inputs, which can be expressed in non-monetary units,
- identifying inefficient units (in Farrel's efficiency sense) and creating their ranking,
- indicating the inefficient objects the distance from the efficiency frontier (i.e., the size of their inefficiency; output surplus or input deficit),
- indicating units with an optimal production scale in the examined group of units (the production scale efficiency),
- determining the area of economies of scale within a given unit.

The basic DEA model (DEA–CCR) became the basis for modifying and developing this method. It allowed for the creation of many varieties of DEA, e.g., the aforementioned BCC model (Bankers, Charnes, & Cooper, 1984), but also the SBM model (slack-based model) (Charnes, Cooper, Golany, Seiford, & Stutz, 1985), CEM model (cross-efficiency model) (Sexton, Silkman, & Hogan, 1986) and many others.

The DEA procedure is presented in the second section. It allows for imagining the main steps you have to take to achieve the assumed goals of the research. The third part of this chapter includes the case study where the DEA method is employed. The application of the DEA input-oriented model in its classic basic variant including CCR model (assuming constant return to scale—CRS) and the BBC model (assuming the variable return to scale—VRS) is presented. In the last part of this chapter, you can find some tasks and questions related to the DEA method.

5.2. DEA procedure: main steps

5.2.1. Aims of research and data (inputs and outputs) selection

Step 1
At the beginning, you have to define the aim of you research. If your goal is to assess efficiency of units in comparison to other units, the DEA method is appropriate. You can evaluate efficiency of units at micro or macro level, so you can compare efficiency of particular firms, farms, institutions, sectors or countries.

Step 2
When you have the aim of your research, in the next step you have to define and prepare the input and output variables. Because of DEA features, you are allowed to take multiple inputs and outputs. You can consider, e.g., three inputs and two outputs in the efficiency estimation instead of the single efficiency calculation (only one output divided by only one input).

The number of inputs and outputs are not limitless and depends on the number of analysed entities. One of the rough rules of numbs in the DEA method is to take the number of units equal or greater than three times the sum of the total number of inputs and outputs ($3 \cdot (\text{inputs} + \text{outputs})$).

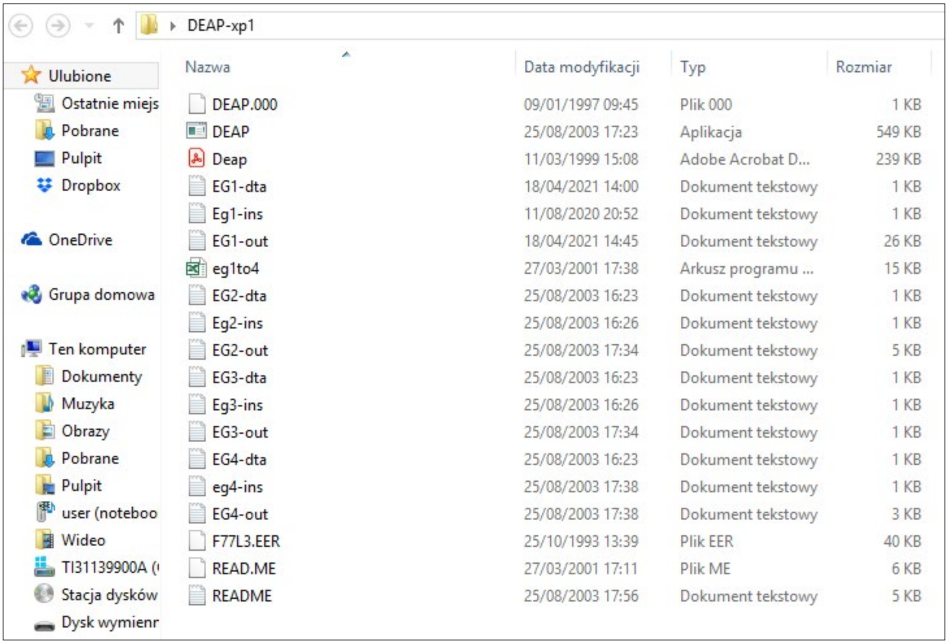
It is important to know, that DEA does not accept negative or zero values for inputs and outputs. In this case, you can employ technics to avoid such a situation, e.g., zero value can be replaced by very low values such as 0.01.

When preparing the data set, you have to remember that each analysed unit must have the same number of inputs and outputs in order to be compared.

You can use primary or secondary data. The first group of data can be gathered during your own primary survey. The second group of data can be retrieved, e.g., from public databases that embrace different range of data at local, country, international or sectoral level, etc. (World Bank, UNCTAD, FADN).

Next, your selected outputs and inputs should be prepared in a table form in Excel file. And considering the requirements of the DEA software, the selected outputs should be put in columns before columns with inputs.

Step 3



Picture 5.1. Folder DEAP-xp1

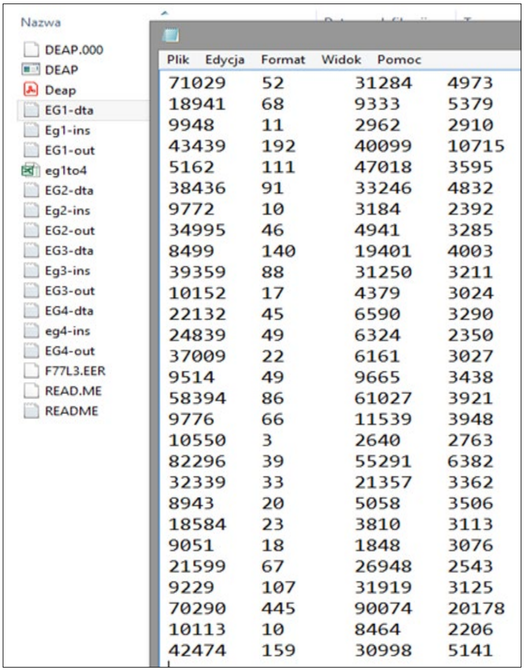
To engage DEA model and make all calculations you have to install the software. The core calculations are made using the DEAP computer program, used by the Centre for Efficiency and Productivity Analyses at the University of Queensland. DEAP is a free software constructed by Tim Coelli and can be downloaded from: https://economics.uq.edu.au/cepa/software (Centre for Efficiency). After the installing process, the DEAP-xp1 folder (Picture 5.1) is available.

Of course, you can download DEAP software at the very beginning of the research process. Here is the last time you have to do it, if you want to employ DEA model.

Step 4

In the next stage, the prepared data file should be imported to DEA program. It will be saved in the EG1-dta file (Picture 5.2).

Remember that you cannot put text data, but numbers in this file. Therefore, the names of variables (inputs, outputs) and names of DMUs (analysed units) cannot be included in the table. So, in the final file, the analysed units will be numbered, i.e., 1, 2, 3, etc. A properly prepared data file looks like the one in Picture 5.2.


Nazwa
DEAP.000
DEAP
Deap
EG1-dta
Eg1-ins
EG1-out
eg1to4
EG2-dta
Eg2-ins
EG2-out
EG3-dta
Eg3-ins
EG3-out
EG4-dta
eg4-ins
EG4-out
F77L3.EER
READ.ME
README

Plik Edycja Format Widok Pomoc

71029	52	31284	4973
18941	68	9333	5379
9948	11	2962	2910
43439	192	40099	10715
5162	111	47018	3595
38436	91	33246	4832
9772	10	3184	2392
34995	46	4941	3285
8499	140	19401	4003
39359	88	31250	3211
10152	17	4379	3024
22132	45	6590	3290
24839	49	6324	2350
37009	22	6161	3027
9514	49	9665	3438
58394	86	61027	3921
9776	66	11539	3948
10550	3	2640	2763
82296	39	55291	6382
32339	33	21357	3362
8943	20	5058	3506
18584	23	3810	3113
9051	18	1848	3076
21599	67	26948	2543
9229	107	31919	3125
70290	445	90074	20178
10113	10	8464	2206
42474	159	30998	5141

Picture 5.2. A ready-to-use spreadsheet in EG1-dta.txt file

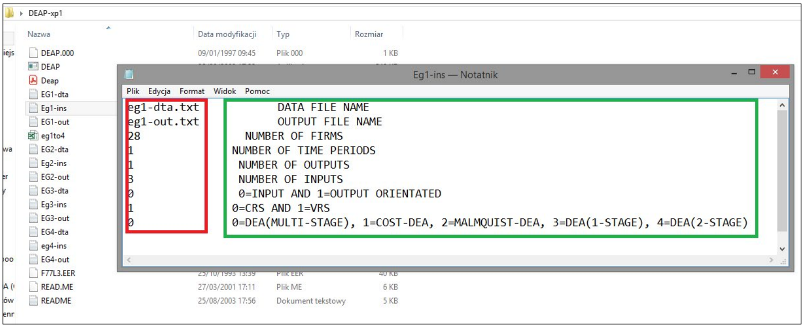
5.2.2. Model calibration and calculation

Step 5
To run a DEA model, the user has to calibrate the model. There are some parameters the researcher should consider. These parameters can be found in the Eg1-ins file (Picture 5.3). In the right column there are the names of parameters (green frame in Picture 5.3) and in the left column there is a place [space] to be filled with the values of parameters according to the particular data set (red frame in the Picture 5.3):

– number of firms (number of analysed objects),
– number of time periods (numbers of analysed years),
– number of outputs (numbers of engaged outputs),
– number of inputs (numbers of engaged inputs),
– 0 = input and 1 = output oriented (you can choose between input-oriented model or output-oriented. 0 means the assumption of input-oriented model and 1 means the assumption of the output-oriented model),
– 0 = CRS and 1 = VRS (you can choose the model with constant returns to scale (CRS) or the model with variables returns to scale (VRS); 0 means the assumption of CRS and 1 means the assumption VRS. If the user puts 1, the program

gives the efficiency indicators for both the CRS and VRS models, which is necessary to get the production scale efficiency indicators. If the user puts 0, the program calculates the efficiency indicators only for CRS model),

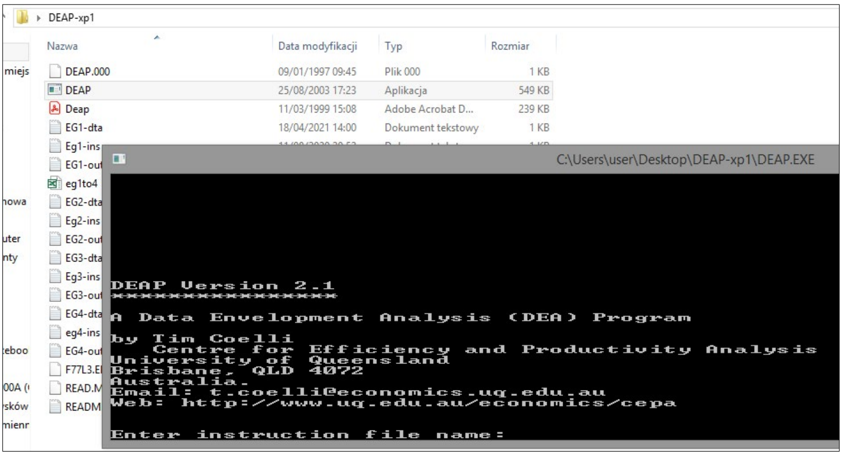
– the description of the econometric procedure for determining effectiveness. A more detailed description of that part of the procedure can be found in the DEAP manual included with the software folder (Coelli, 1996).



Picture 5.3. Model calibration in Eg1-ins file

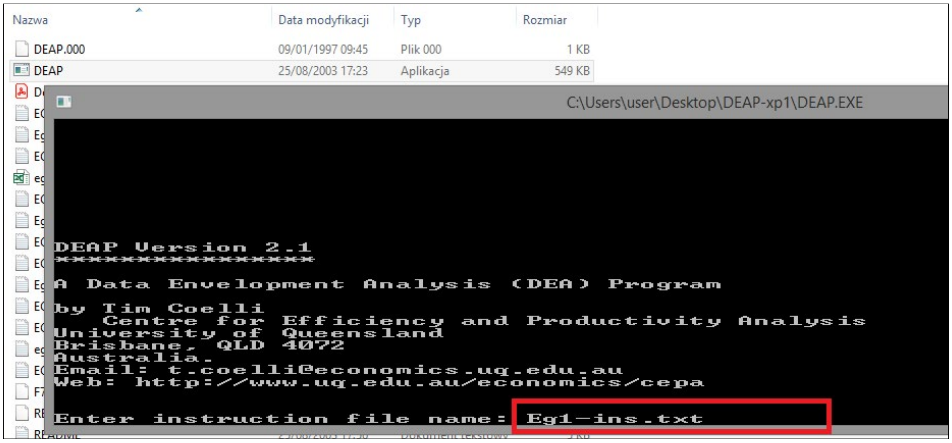
Step 6

The next step is to let the program make calculations. To obtain the final results, the DEAP file should be opened, and the black frame will appear (Picture 5.4).



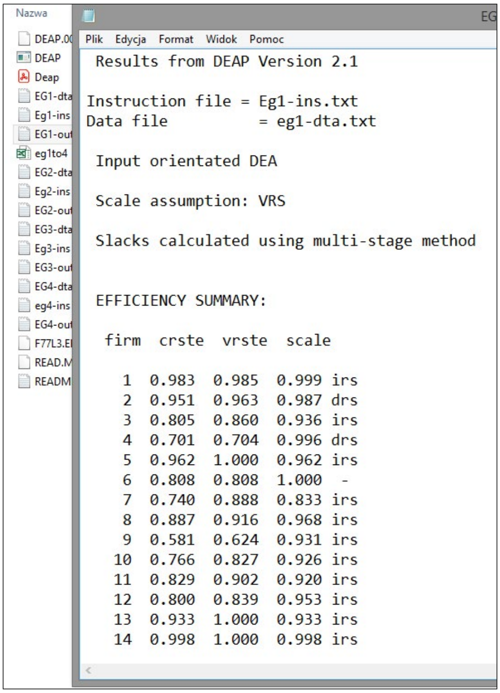
Picture 5.4. The DEAP file opened (i)

Enter the name of the configuration file, i.e., Eg1-ins.txt in the place of the blinking cursor (red frame in the Picture 5.5). Next, press enter, and the window will close automatically.



Picture 5.5. The DEAP file opened (ii)

The researcher can find the results of the calculation in the EG1-out file (Picture 5.6).



Picture 5.6. Results in the EG1-out file (i)

In Table 5.2. there are some abbreviations that help you to calibrate the model and read the results.

Table 5.2. Abbreviations used in DEA calculation

CRS	Constant Returns to Scale
VRS	Variable Returns to Scale
CRSTE	Constant Returns to Scale Technical Efficiency
VRSTE	Variable Returns to Scale Technical Efficiency
Scale	Scale Efficiency
IRS	Increasing Returns to Scale
DRS	Decreasing Returns to Scale

Own elaboration.

5.2.3. Results interpretation

Step 7
The next step is the interpretation of the results that are obtained in the calculation process. The table with the results is presented in Picture 5.7.

The first column (firms) contains the numbers of the analysed units. The second column (crste) includes technical efficiency indicators for selected units assuming the constant returns to scale (CRS). The third column (vrste) presents the technical efficiency indexes of the analysed units assuming variable returns to scale (VRS). In the next column (scale) there are the scale efficiency indicators, while in the last column there are the scale effects areas (IRS or DRS).

If the production scale of the analysed objects is optimal, then the scale effects area is not marked in the last column (e.g., no abbreviation in the fifth column for firm no. 1, Picture 5.7). The scale effects area (IRS or DRS) can only be determined in units with a non-optimal production scale, i.e., those where the scale efficiency index is lower than 1 (present abbreviation in the fifth column for firms no. 2–7, Picture 5.7).

```
EFFICIENCY SUMMARY:

 firm  crste  vrste  scale

    1  1.000  1.000  1.000  -
    2  0.311  0.473  0.656 irs
    3  0.557  0.917  0.607 irs
    4  0.311  0.315  0.985 irs
    5  0.097  0.614  0.159 irs
    6  0.555  0.627  0.886 irs
    7  0.543  1.000  0.543 irs
    8  1.000  1.000  1.000  -
    9  0.156  0.551  0.284 irs
   10  0.845  0.944  0.895 irs
   11  0.382  0.788  0.485 irs
   12  0.555  0.771  0.719 irs
   13  0.843  1.000  0.843 irs
   14  1.000  1.000  1.000  -
   15  0.220  0.663  0.331 irs
   16  1.000  1.000  1.000  -
   17  0.196  0.575  0.340 irs
   18  1.000  1.000  1.000  -
   19  1.000  1.000  1.000  -
   20  0.686  0.836  0.820 irs
   21  0.291  0.680  0.428 irs
   22  0.754  0.921  0.819 irs
   23  0.692  1.000  0.692 irs
   24  0.583  0.912  0.640 irs
   25  0.203  0.706  0.288 irs
   26  0.260  0.341  0.762 drs
   27  0.495  1.000  0.495 irs
   28  0.584  0.634  0.920 irs

 mean  0.576  0.795  0.700

Note: crste = technical efficiency from CRS DEA
      vrste = technical efficiency from VRS DEA
      scale = scale efficiency = crste/vrste
```

Picture 5.7. Results in the EG1-out file (ii)

When interpreting the results, it is worth remembering that all the obtained indicators are relative, i.e., their values are estimated in comparison to other entities. Adding or subtracting another unit could change the obtained results of all analysed units. This applies to both the efficiency indicators (assuming CRS or VRS) and the scale efficiency indicators. It should also be emphasised that results equal to 1 are always present for units achieving the highest technical efficiency (the highest relation of outputs to inputs in the analysed sample) and / or the scale efficiency. Adding or subtracting one unit can cause that fully efficient units (efficiency score equals to 1) could not be such units in a new research sample.

There are some main elements / aspects that should be indicated and interpreted:

- efficiency scores assuming constant returns to scale technology; units which are granted 100% efficiency scores (efficiency equals to 1) and units which are inefficient (column 2, Picture 5.7);
- efficiency scores assuming variable returns to scale technology; units which are granted 100% efficiency scores and units which are inefficient (column 3, Picture 5.7);
- scale of production in relation to efficient and inefficient units assuming constant or variable returns to scale (column 4, Picture 5.7);
- the level of input reduction (or augment of outputs) in order to become efficient maintaining output (or input) level;
- mean values for the analysed group of units in relation to efficiency score.

You can find a more detailed interpretation in two examples in part 3.3. The first example is devoted to the analysis of the relative economic efficiency of farms in the European Union. Then the analysis is in-depth at more disaggregated level as the second example is concentrated only on relative economic efficiency of crops farms in European Union.

5.3. Comparison of farms' efficiency in the European Union: case study no. 1

5.3.1. Aims of research and data selection from FADN

Step 1

In our case study: the main aim of the research is to determine the relative economic efficiency of representative agriculture farms from the European Union countries in 2018. In other words, we want to compare agricultural farms in the EU considering their efficiency. The production scale efficiency and areas of scale effects (IRS or DRS) will also be estimated.

Step 2
In the next step, the input and output variables should be defined. In our case, to estimate efficiency index for agricultural farm and compare them, the FADN database is used. FADN (Farm Accountancy Data Network) is a system for collecting accountancy data from agricultural farms in each country of the European Union.

The data collected under FADN structure are used for the annual description of the income of farms operating in the individual EU countries, the analysis of the activity of farms and the assessment of the effects of implemented and planned changes in the EU agricultural policy. The FADN was developed as a harmonized system of sample surveys, using precisely defined terms with a precisely developed method of selecting a sample of farms and transparent control procedures. As a result, the data obtained by the FADN are reliable and representative, what determines the possibility of reflecting the actual results of farms operating in the EU countries (Goraj & Olewnik, 2011). The data collected under the FADN is publicly available and published on the website: https://ec.europa.eu/agriculture/rica/database/database_en.cfm.

To achieve comparability between the variables used in the FADN from individual EU countries, each variable that occurs in the FADN database is precisely defined. These variables are described by symbols SE with a specific number, e.g., SE011. To explain the individual abbreviations, take a look at the diagrams presenting the creation of individual variables of FADN available on the website: https://ec.europa.eu/agriculture/rica/annex003_en.cfm.

Technical efficiency can express the economic dimension of the farm sustainability. For the measurement of economics sustainability of farms, the researchers use e.g.: the value of income per person or farm, the number of the holding's expenses, less often the farm wage level. Among other measures of economic dimension of farm sustainability, the employment and professional activity indicators, workforce productivity, fixed asset capital intensity, and energy intensity indicators, investment level, outlays on research and development activity are the most commonly used in the literature.

In our case study, on the input side, we use annual values of three variables: 1) depreciation in EUR (SE 360), 2) labour input in hours (SE 011), and 3) land inputs expressed in Utilized Agricultural Area in hectares (SE 025).

On the output side the annual farm net income in EUR (SE 420) is applied.

As mentioned before, the selected input and output variables for individual EU countries from FADN database should be downloaded and saved in Excel file. Considering the requirements of the software DEA, the selected outputs should be put before inputs in data table (Table 5.3).

Table 5.3. Output and input values of farms in selected EU countries in 2018

No.	Country	Farm Net Income	Utilised Agricultural Area	Depreciation	Labour input
1	Belgium	71029	52	31284	4973
2	Bulgaria	18941	68	9333	5379
3	Cyprus	9948	11	2962	2910
4	Czech Republic	43439	192	40099	10715
5	Denmark	5162	111	47018	3595
6	Germany	38436	91	33246	4832
7	Greece	9772	10	3184	2392
8	Spain	34995	46	4941	3285
9	Estonia	8499	140	19401	4003
10	France	39359	88	31250	3211
11	Croatia	10152	17	4379	3024
12	Hungary	22132	45	6590	3290
13	Ireland	24839	49	6324	2350
14	Italy	37009	22	6161	3027
15	Lithuania	9514	49	9665	3438
16	Luxembourg	58394	86	61027	3921
17	Latvia	9776	66	11539	3948
18	Malta	10550	3	2640	2763
19	Netherlands	82296	39	55291	6382
20	Austria	32339	33	21357	3362
21	Poland	8943	20	5058	3506
22	Portugal	18584	23	3810	3113
23	Romania	9051	18	1848	3076
24	Finland	21599	67	26948	2543
25	Sweden	9229	107	31919	3125
26	Slovakia	70290	445	90074	20178
27	Slovenia	10113	10	8464	2206
28	United Kingdom	42474	159	30998	5141

Source: Own survey based on FADN database.

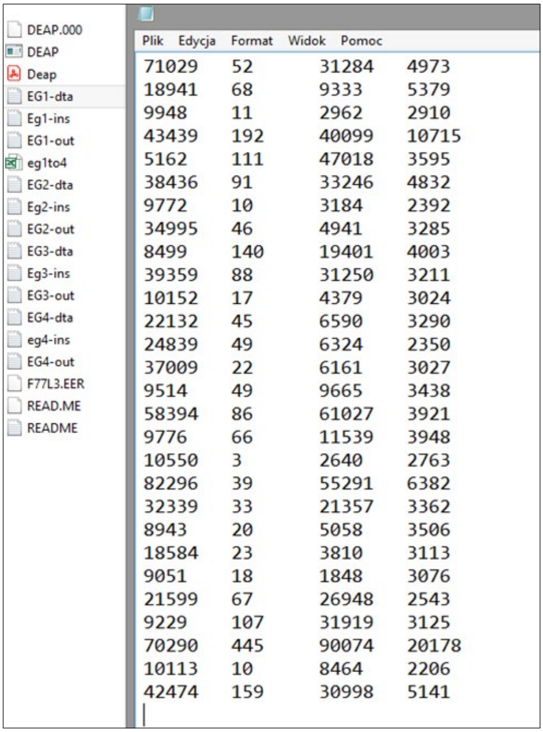
Step 3

The next step: install DEAP software following steps presented in point 5.2.1. *Step 3*. You can skip that step if you already have DEAP software.

Step 4

In the next step, the downloaded data from FADN prepared in a suitable table (see Table 5.3) should be imported to DEA program and saved in the EG1-dta file (Picture 5.8).

As it was said, you have to put only numbers not text data in this file. Therefore, the names of inputs (Utilised Agricultural Area; Depreciation, Labour Input) and output (Farm Net Income), and names of DMUs (Belgium, Bulgaria, Cyprus, etc.) cannot be included in the table. In our case the prepared data file looks like in Picture 5.8.



71029	52	31284	4973
18941	68	9333	5379
9948	11	2962	2910
43439	192	40099	10715
5162	111	47018	3595
38436	91	33246	4832
9772	10	3184	2392
34995	46	4941	3285
8499	140	19401	4003
39359	88	31250	3211
10152	17	4379	3024
22132	45	6590	3290
24839	49	6324	2350
37009	22	6161	3027
9514	49	9665	3438
58394	86	61027	3921
9776	66	11539	3948
10550	3	2640	2763
82296	39	55291	6382
32339	33	21357	3362
8943	20	5058	3506
18584	23	3810	3113
9051	18	1848	3076
21599	67	26948	2543
9229	107	31919	3125
70290	445	90074	20178
10113	10	8464	2206
42474	159	30998	5141

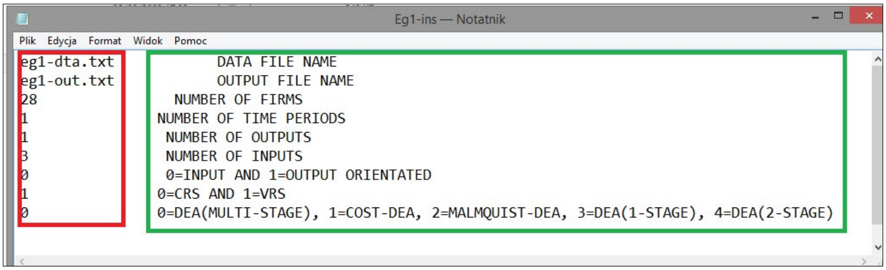
Picture 5.8. A ready-to-use spreadsheet in EG1-dta.txt file

5.3.2. Model calibration and calculation

Step 5

In the next step the model is calibrated according to our assumption. The calibration is made in the Eg1-ins file. Considering our research goal and data set, the left column should be fulfilled (red frame in Picture 5.9):

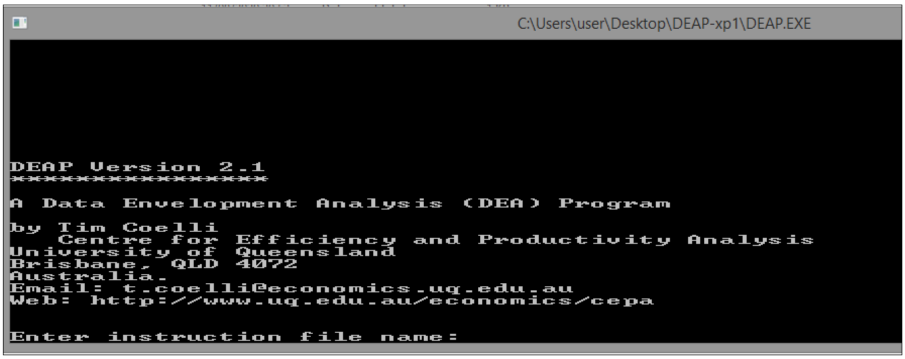
- 28 (number of firms; number of analysed objects, in the presented case study 28 countries)
- 1 (number of times; 1 because the case study concerns only 2018),
- 1 (number of outputs; 1 because there is only one output Farm Net Income in the case study),
- 3 (number of inputs; there are three inputs in our case study: Utilized Agricultural Area, Depreciation and Labour Input),
- 0 (model orientation; in our case study, the input-oriented model was assumed),
- 1 (model with variables return to scale is adopted),
- 0 (the multistage econometric procedure was applied).



Picture 5.9. Model calibration in Eg1-ins file

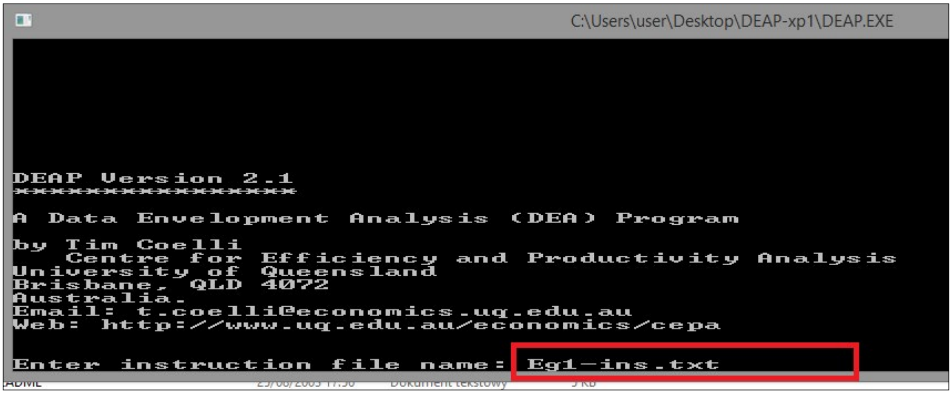
Step 6

To calculate and to obtain final result, the DEAP file should be opened, and the black frame will appear (Picture 5.10).



Picture 5.10. The DEAP file opened (i)

In the place of the blinking cursor (red frame in Picture 5.11), enter the name of the configuration file, i.e., Eg1-ins.txt. Next, press enter, and the window will close automatically.

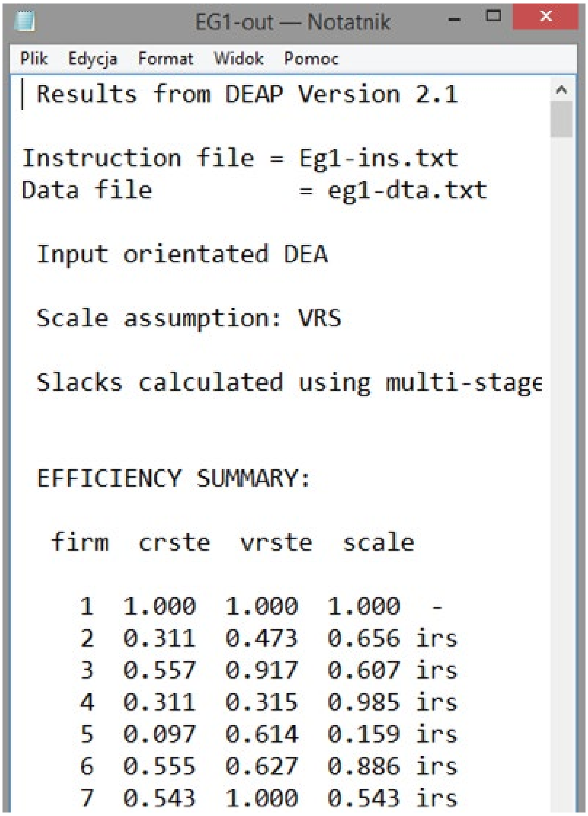


C:\Users\user\Desktop\DEAP-xp1\DEAP.EXE

```
DEAP Version 2.1
******************
A Data Envelopment Analysis (DEA) Program
by Tim Coelli
   Centre for Efficiency and Productivity Analysis
University of Queensland
Brisbane, QLD 4072
Australia.
Email: t.coelli@economics.uq.edu.au
Web: http://www.uq.edu.au/economics/cepa
Enter instruction file name: Eg1-ins.txt
```

Picture 5.11. The DEAP file opened (ii)

The final results of calculation can be found in the EG1-out file (Picture 5.12). Having it the researcher should interpret the obtained values.



EG1-out — Notatnik

Plik Edycja Format Widok Pomoc

```
 Results from DEAP Version 2.1

Instruction file = Eg1-ins.txt
Data file         = eg1-dta.txt

 Input orientated DEA

 Scale assumption: VRS

 Slacks calculated using multi-stage


 EFFICIENCY SUMMARY:

  firm  crste  vrste  scale

    1  1.000  1.000  1.000  -
    2  0.311  0.473  0.656 irs
    3  0.557  0.917  0.607 irs
    4  0.311  0.315  0.985 irs
    5  0.097  0.614  0.159 irs
    6  0.555  0.627  0.886 irs
    7  0.543  1.000  0.543 irs
```

Picture 5.12. Results in the EG1-out file (i)

5.3.3. Results interpretation

Step 7

The results of our case study are presented in Picture 5.13. The first column (firms) contains the numbers of the EU countries. And as mentioned before, the objects are numbered, in our case from 1 to 28 (Member countries of the European Union in 2018, Picture 5.13 and Table 5.3). The second one (crste) includes technical efficiency indicators for representative farms in these selected countries assuming the constant returns to scale (CRS). The third column (vrste) presents the technical efficiency indexes of farms in the individual EU countries assuming variable returns to scale (VRS). In the next column (scale) there are scale efficiency indicators, while in the last one there are the scale effect areas (IRS or DRS).

In our sample the most efficient farms were in Belgium (no. 1), Spain (no. 8), Italy (no. 14), Malta (no. 16) and Netherlands (no. 19) in 2018 (Picture 5.13 and Table 5.3). The representative farms from these countries, regardless of the assumption about the returns to scale (CRS or VRS), achieved the efficiency indicators equal to 1. It means that they were located on the so-called frontier curve and constituted benchmarks for farms from other EU countries; they are granted 100% efficiency score. In other words, farms from these countries (Belgium, Spain, Italy, Malta and Netherlands) fully efficiently use land, labour and capital inputs to achieve assumed output (expressed in Farm Net Income in the case study). The farms from these countries have also an optimal scale of production, as evidenced by the scale efficiency index equal to 1 (column 4 in Picture 5.13).

```
EFFICIENCY SUMMARY:

 firm  crste  vrste  scale

   1   1.000  1.000  1.000   -
   2   0.311  0.473  0.656  irs
   3   0.557  0.917  0.607  irs
   4   0.311  0.315  0.985  irs
   5   0.097  0.614  0.159  irs
   6   0.555  0.627  0.886  irs
   7   0.543  1.000  0.543  irs
   8   1.000  1.000  1.000   -
   9   0.156  0.551  0.284  irs
  10   0.845  0.944  0.895  irs
  11   0.382  0.788  0.485  irs
  12   0.555  0.771  0.719  irs
  13   0.843  1.000  0.843  irs
  14   1.000  1.000  1.000   -
  15   0.220  0.663  0.331  irs
  16   1.000  1.000  1.000   -
  17   0.196  0.575  0.340  irs
  18   1.000  1.000  1.000   -
  19   1.000  1.000  1.000   -
  20   0.686  0.836  0.820  irs
  21   0.291  0.680  0.428  irs
  22   0.754  0.921  0.819  irs
  23   0.692  1.000  0.692  irs
  24   0.583  0.912  0.640  irs
  25   0.203  0.706  0.288  irs
  26   0.260  0.341  0.762  drs
  27   0.495  1.000  0.495  irs
  28   0.584  0.634  0.920  irs

 mean  0.576  0.795  0.700

Note: crste = technical efficiency from CRS DEA
      vrste = technical efficiency from VRS DEA
      scale = scale efficiency = crste/vrste
```

Picture 5.13. Results in the EG1-out file (ii)

If the analysis is based on variable returns to scale effects (VRS), the farms from Greece (no. 7), Ireland (no. 13), Romania (no. 23) and Slovenia (no. 27) are also among farms with the highest technical efficiency (Picture 5.13 and Table 5.3). However, farms from these countries do not achieve the optimal production scale, because the adequate scale efficiency indexes differ from one (column 4, Picture 5.13). All these farms operate in the area of increasing economies to scale (column 5 in Picture 5.13). That allows to conclude that increasing the inputs (and further production volume as a consequence) would result in an over-proportional increase in the farm output (Farm Net Income in this case study).

In our case study, only the farms from Slovakia (no. 26) operate in decreasing returns to scale (column 5, Picture 5.13). It means, that the increase in inputs (and agricultural production as a consequence) causes less than proportional increase in output (Farm Net Income in our case study). Therefore, it is not economically justified to further increase (expand) production by these farms.

In our research sample, assuming VRS, the farms from the Czech Republic (no. 4) and Slovakia (no. 26) had the lowest technical efficiency in 2018 (Picture 5.13). For the Czech Republic, the technical efficiency index of farms was equal to 0.315. Assuming input-oriented model, it means that the same output (Farm Net Income in the case study) can be achieved by input reduction by 68.5% (1 – 0.315 = 0.685). The input reduction by 68.5% would cause these farms to achieve technical efficiency equal to 1. In Slovakia, a 65.9% (1 – 0.341 = 65.9) reduction in farm inputs would make farms achieve efficiency index at the level of 1 while maintaining output level. Thus, the higher the technical efficiency index, the lesser the need to reduce inputs to achieve full technical efficiency at a given output. The other efficiency indicators for the surveyed objects (in the case study for representative agricultural holdings from the individual EU countries) should be interpreted in a similar way.

The lowest scale efficiency of representative farms, amounting to only 0.159, was recorded in Denmark (no. 5) (Picture 5.13 and Table 5.3). This means that adjusting the production volume to the optimal level would allow them to save as much as 84.10% of inputs (1–0.159).

The average values for the studied group of units can also be assessed (red frame in Picture 13). Depending on the adopted assumption, the efficiency ratio was 0.576 and 0.795 for CRS and VRS, respectively. Assuming the VRS, in the EU countries' farms, it was necessary to reduce inputs on farms by 20.5% (1 – –0.795 = 0.205) to achieve full technical efficiency (equal to 1) at a given output level (Farm Net Income in the case study) in 2018. Adjusting the production volume to the optimal scale in surveyed EU farms, would allow them to save as much as 30% of the current inputs (1 – 0.70 = 0.30).

5.4. Comparison of crops farm efficiency in the European Union: case study no. 2

5.4.1. Aims of research and data selection from FADN

Step 1
The main aim of the research is to determine the relative technical efficiency of representative farms specialized in field crops from the European Union in 2018. The production scale efficiency and areas of scale effects (IRS or DRS) will be also estimated.

Step 2
To calculate efficiency index, the input and output variables should be defined first. The data are retrieved from FADN database. On the input side, we use annual values of three variables: 1) intermediate consumption value in EUR (SE 275), 2) labour input in hours (SE 011), and 3) land inputs expressed in Utilized Agricultural Area in hectares (SE 025). On the output side the total production value of farm in EUR (SE 131) is applied. The data should be downloaded and saved in Excel in the form as presented in Table 5.4.

Table 5.4. Output and input values of crop farms in selected EU countries in 2018 prepared for DEA analysis

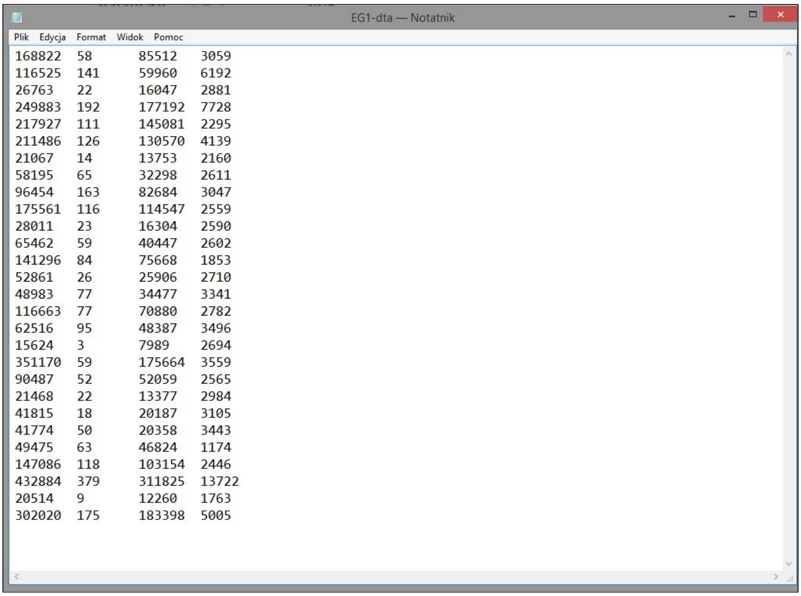
No.	Country	Production	Utilised Agricultural Area	Intermediate consumption	Labour input
1	Belgium	168822	58	85512	3059
2	Bulgaria	116525	141	59960	6192
3	Cyprus	26763	22	16047	2881
4	Czech Republic	249883	192	177192	7728
5	Denmark	217927	111	145081	2295
6	Germany	211486	126	130570	4139
7	Greece	21067	14	13753	2160
8	Spain	58195	65	32298	2611
9	Estonia	96454	163	82684	3047
10	France	175561	116	114547	2559
11	Croatia	28011	23	16304	2590
12	Hungary	65462	59	40447	2602
13	Ireland	141296	84	75668	1853
14	Italy	52861	26	25906	2710
15	Lithuania	48983	77	34477	3341
16	Luxembourg	116663	77	70880	2782
17	Latvia	62516	95	48387	3496
18	Malta	15624	3	7989	2694

No.	Country	Production	Utilised Agricultural Area	Intermediate consumption	Labour input
19	Netherlands	351170	59	175664	3559
20	Austria	90487	52	52059	2565
21	Poland	21468	22	13377	2984
22	Portugal	41815	18	20187	3105
23	Romania	41774	50	20358	3443
24	Finland	49475	63	46824	1174
25	Sweden	147086	118	103154	2446
26	Slovakia	432884	379	311825	13722
27	Slovenia	20514	9	12260	1763
28	United Kingdom	302020	175	183398	5005

Source: Own survey based on FADN database.

Step 3
Install DEAP software following point 5.2.1 *Step 3*.

Step 4
Import data prepared to for DEA and saved in the EG1-dta file (Picture 5.14). The names of variables (inputs, outputs) and names of countries (analysed units) were numbered, i.e., 1, 2, 3, etc. A properly prepared data file looks like the one in Picture 5.14.


EG1-dta — Notatnik

Plik Edycja Format Widok Pomoc

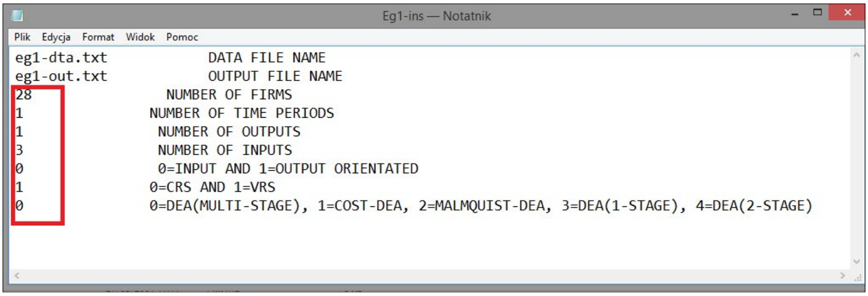
168822	58	85512	3059
116525	141	59960	6192
26763	22	16047	2881
249883	192	177192	7728
217927	111	145081	2295
211486	126	130570	4139
21067	14	13753	2160
58195	65	32298	2611
96454	163	82684	3047
175561	116	114547	2559
28011	23	16304	2590
65462	59	40447	2602
141296	84	75668	1853
52861	26	25906	2710
48983	77	34477	3341
116663	77	70880	2782
62516	95	48387	3496
15624	3	7989	2694
351170	59	175664	3559
90487	52	52059	2565
21468	22	13377	2984
41815	18	20187	3105
41774	50	20358	3443
49475	63	46824	1174
147086	118	103154	2446
432884	379	311825	13722
20514	9	12260	1763
302020	175	183398	5005

Picture 5.14. A ready-to-use spreadsheet in EG1-dta.txt file

5.4.2. Model calibration and calculation

Step 5

Calibrate model according to the assumption. The calibration is made in the Eg1-ins file.



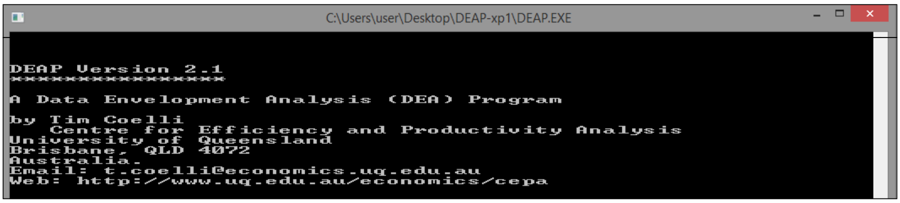
Picture 5.15. Model calibration in Eg1-ins file

Considering the research goal and data set, the left column should be fulfilled (red frame in Picture 5.15):

- 28 (number of firms; number of analysed objects, in the presented case study 28 countries),
- 1 (number of times; 1 because the case study concerns only 2018),
- 1 (number of outputs; 1 because there is only one output production value in this example),
- 3 (number of inputs; there are three inputs in the example: Utilized Agricultural Area, Intermediate Consumption, Labour Input),
- 0 (model orientation; in our case study, the input-oriented model was assumed),
- 1 (model with variables return to scale is adopted),
- 0 (the multistage econometric procedure was applied).

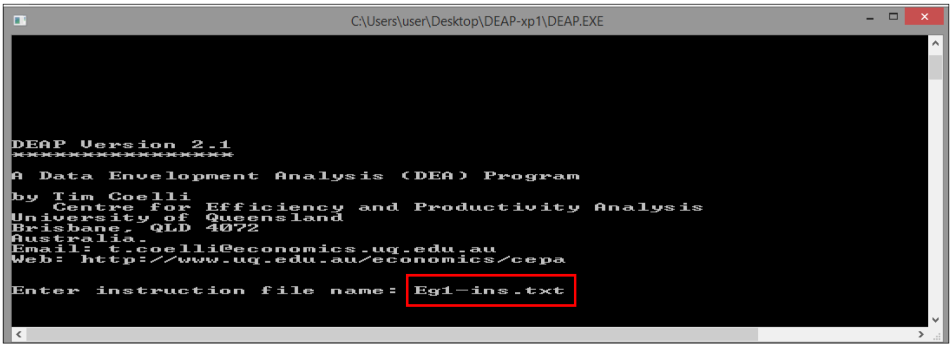
Step 6

To calculate and to obtain final result the DEAP file should be opened, and the black frame will appear (Picture 5.16).



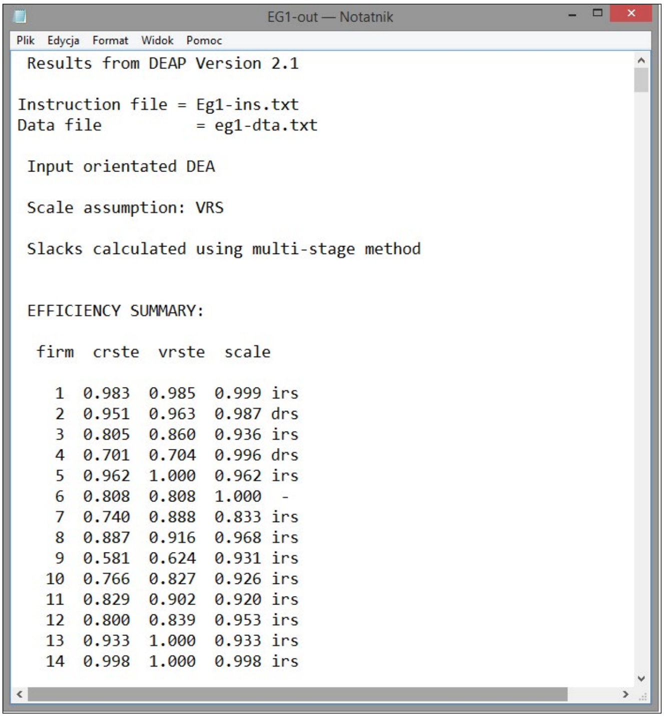
Picture 5.16. The DEAP file opened (i)

In the place of the blinking cursor (red frame in the Picture 5.17), enter the name of the configuration file, i.e., Eg1-ins.txt. Next, press enter, and the window will close automatically.



Picture 5.17. The DEAP file opened (ii)

The final results of calculation can be found in the EG1-out file (Picture 5.18). Having it the researcher should interpret the obtained values.


EG1-out — Notatnik

Plik Edycja Format Widok Pomoc

Results from DEAP Version 2.1

Instruction file = Eg1-ins.txt
Data file = eg1-dta.txt

Input orientated DEA

Scale assumption: VRS

Slacks calculated using multi-stage method

EFFICIENCY SUMMARY:

firm	crste	vrste	scale	
1	0.983	0.985	0.999	irs
2	0.951	0.963	0.987	drs
3	0.805	0.860	0.936	irs
4	0.701	0.704	0.996	drs
5	0.962	1.000	0.962	irs
6	0.808	0.808	1.000	-
7	0.740	0.888	0.833	irs
8	0.887	0.916	0.968	irs
9	0.581	0.624	0.931	irs
10	0.766	0.827	0.926	irs
11	0.829	0.902	0.920	irs
12	0.800	0.839	0.953	irs
13	0.933	1.000	0.933	irs
14	0.998	1.000	0.998	irs

Picture 5.18. Results in the EG1-out file (i)

5.4.3. Results interpretation

Step 7

The most efficient farms specialized in field crops in 218 were in Netherlands (no. 19), and Portugal (no. 22) (Picture 5.14). The representative farms from these countries, regardless of the assumption about the returns to scale (CRS or VRS), achieved the efficiency indicators equal to 1. It means that they were located on the so-called frontier curve and constituted benchmarks for farms from the other EU countries; they are granted 100% efficiency score. In other words, farms from these countries (Netherlands, Portugal) fully efficiently use land, labour and capital inputs to achieve assumed output (expressed in total production value). The farms from these countries have also an optimal scale of production, as evidenced by the scale efficiency index equal to 1 (column 4 in Picture 5.19).

If the analysis is based on variable returns to scale effects (VRS), the farms from Denmark (no. 5), Ireland (no. 13), Italy (no. 14), Malta (no. 18), Finland (no. 24), Slovakia (no. 26), Slovenia (no. 27) are also among farms with the highest technical efficiency (Picture 5.19). However, farms from these countries do not achieve the optimal production scale, because the adequate scale efficiency indexes differ from one (column 4, Picture 5.19). These farms without farms from Slovakia (no. 26) operate in the area of increasing economies to scale (column 5 in Picture 5.19). That allows to conclude that increasing the inputs (and further production volume as a consequence) would result in an over-proportional increase in the farms output (total production value). In this case study, only the farms from Slovakia (no. 26), Bulgaria (no. 2), Czech Republic (no. 4) and United Kingdom (no. 28) operate in decreasing returns to scale (column 5, Picture 5.19). It means, that the increase in input (and agricultural production as a consequence) causes less than proportional increase in output (total production value). Therefore, it is not economically justified to further increase (expand) production by these farms. In our research sample, assuming VRS, the farms from Estonia (no. 9) and Latvia (no. 17) had the lowest technical efficiency in 2018 (Picture 5.19). For Estonia, the technical efficiency index of farms was equal to 0.624. Assuming input-oriented model, it means that the same output (total value of production) can be achieved by input reduction by 37.6% (1 – 0.624 = 0.376). The input reduction by 37.6% would cause these farms to achieve technical efficiency equals to 1. In Latvia, a 33.8% (1 – 0.662 = 0.338) reduction in farm inputs would make farms achieve efficiency index at the level of 1 while maintaining output level. The lowest scale efficiency of representative farms, amounting to 0.528, was recorded in Finland (no. 24) (Picture 5.14). This means that adjusting the production volume to the optimal level would allow them to save as much as 47.2% of inputs (1 – 0.528).

```
EFFICIENCY SUMMARY:

 firm  crste  vrste  scale

   1  0.983  0.985  0.999 irs
   2  0.951  0.963  0.987 drs
   3  0.805  0.860  0.936 irs
   4  0.701  0.704  0.996 drs
   5  0.962  1.000  0.962 irs
   6  0.808  0.808  1.000  -
   7  0.740  0.888  0.833 irs
   8  0.887  0.916  0.968 irs
   9  0.581  0.624  0.931 irs
  10  0.766  0.827  0.926 irs
  11  0.829  0.902  0.920 irs
  12  0.800  0.839  0.953 irs
  13  0.933  1.000  0.933 irs
  14  0.998  1.000  0.998 irs
  15  0.696  0.721  0.966 irs
  16  0.819  0.841  0.974 irs
  17  0.637  0.662  0.963 irs
  18  0.976  1.000  0.976 irs
  19  1.000  1.000  1.000  -
  20  0.863  0.891  0.968 irs
  21  0.775  0.838  0.925 irs
  22  1.000  1.000  1.000  -
  23  0.991  0.991  1.000  -
  24  0.528  1.000  0.528 irs
  25  0.713  0.774  0.921 irs
  26  0.690  1.000  0.690 drs
  27  0.816  1.000  0.816 irs
  28  0.822  0.823  0.999 drs

mean  0.824  0.888  0.931
```

Picture 5.19. Results in the EG1-out file (ii)

The average values for the studied group of units can also be assessed. Depending on the adopted assumption, the efficiency ratio was 0.824 and 0.888 for CRS and VRS, respectively. Assuming the VRS, in the EU countries' farms, it was necessary to reduce inputs on farms by 11.2% (1 – 0.888 = 0.112) to achieve full technical efficiency (equal to 1) at a given output level (total value of production) in 2018. Adjusting the production volume to the optimal scale in surveyed EU farms, would allow them to save 6.9% of the current inputs (1 – 0.931 = 0.069).

Questions / tasks

The DEA method can also be successfully used to assess the performance of non-agricultural units such as banks, hospitals, commercial enterprises, etc. However, it should be remembered that DMUs should operate in similar conditions and be comparable in terms of technology used, scale of operation and specialization. Using publicly available statistical databases and the DEA method, try to do the following exercises:

1. Compare efficiency of the agricultural farms in two EU selected countries in 2004–2007 (use average data for 2004–2017 from the EUFADN database).
2. Evaluate and compare efficiency of farms specialized in field crops and in horticulture within the EU-28 countries in 2017 (use averaged data from the EU-FADN database).
3. Which of the EU countries has the highest efficiency of the food industry measured by the relation of income and the number of employees and the value of fixed capital involved in production process (use Eurostat database)?

References

Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, *30*(9), 1078-1092.

Centre for Efficiency and Productivity Analysis (CEPA). Page which describes the computer program DEAP Version 2.1. Retrieved July 15, 2018 from https://economics.uq.edu.au/cepa/software

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, *2*(6), 429-444.

Charnes, A., Cooper, W. W., Golany, B., Seiford, L., & Stutz, J. (1985). *Foundations of Data Envelopment Analysis for Pareto-Koopmans efficient empirical productions functions.* (Texas University at Austin Center for Cybernetic Studies No. CCS-504).

Coelli T. J. (1996). *A guide to DEAP version 2.1.: A Data Envelopment Analysis (computer) program*. (Centre for Efficiency and Productivity Analysis (CEPA) Working Papers No. 8/96).

Czyżewski, B., Smędzik-Ambroży, K., & Mrówczyńska-Kamińska, A. (2020). Impact of environmental policy on eco-efficiency in country districts in Poland: How does the decreasing return to scale change perspectives?. *Environmental Impact Assessment Review*, *84*, 106431.

Domagała, A. (2007). Metoda Data Envelopment Analysis jako narzędzie badania względnej efektywności technicznej. *Badania Operacyjne i Decyzje*, (3-4), 21-34.

Farm Accountancy Data Network (FADN). *FADN Public Database*. Retrieved July 5, 2018 from https://ec.europa.eu/agriculture/rica/database/database_en.cfm

Farm Accountancy Data Network (FADN). *Annex: Standard results indicators*. Retrieved July 5, 2018 from https://ec.europa.eu/agriculture/rica/annex003_en.cfm

Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society: Series A (General)*, *120*(3), 253-281.

Goraj, L., & Olewnik, E. (2011). *FADN i Polski FADN*. Warszawa: Wydawnictwo Instytutu Ekonomiki Rolnictwa i Gospodarki Żywnościowej – Państwowy Instytut Badawczy. Retrieved July 20, 2018 from http://fadn.pl/wp-content/uploads/2011/06/FADN_polski_FADN.pdf

Guzik, B. (2009). *Podstawowe modele DEA w badaniu efektywności gospodarczej i społecznej*. Poznań: Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu.

Sexton, T. R., Silkman, R. H., & Hogan, A. J. (1986). Data envelopment analysis: Critique and extensions. *New Directions for Program Evaluation*, (32), 73-105.