Scientific Papers Series Management, Economic Engineering in Agriculture and Rural Development
Vol. 15, Issue 1, 2015
PRINT ISSN 2284-7995, E-ISSN 2285-3952

CLUSTER ANALYSIS OF NATURAL DISASTER LOSSES IN POLISH AGRICULTURE

Grzegorz STRUPCZEWSKI

The Cracow University of Economics, Department of Risk Management and Insurance, 27 Rakowicka Street, Cracow 31-510, Poland, Phone: (+48) 608377802,
E-mail: strupczg@uek.krakow.pl

Corresponding author: strupczg@uek.krakow.pl

Abstract

Agricultural production risk is of special nature due to a great number of hazards, relative weakness of production entities on the market and high ambiguity which is greater than in industrial production. Natural disasters occurring very frequently, at simultaneous low percentage of insured farmers, cause damage of such sizes that force the state to organise current financial aid (for instance in the form of preferential natural disaster loans). This aid is usually not sufficient. On the other hand, regional diversity of the risk level does not positively affect the development of insurance. From the perspective of insurance companies and policymakers it becomes highly important to investigate the spatial structure of losses in agriculture caused by natural disasters. The purpose of the research is to classify the 16 Polish voivodeships into clusters in order to show differences between them according to the criterion of level of damage in agricultural farms caused by natural disasters. On the basis of the cluster analysis it was demonstrated that 11 voivodeships form quite a homogeneous group in terms of size of damage in agriculture (the value of damage in cultivations and the acreage of destroyed cultivations are two most important factors determining affiliation to the cluster), however, the profile of loss occurring in other five voivodeships has a very individual course and requires separate handling in the actuarial sense. It was also proved that high value of losses in agriculture in the absolute sense in given voivodeships do not have to mean high vulnerability of agricultural farms from these voivodeships to natural risks.

Key words: *agriculture, cluster analysis, natural disasters, insurance, losses, risk management*

INTRODUCTION

Economic stability of the agricultural economy sector may be disturbed by various random events, such as natural disasters or diseases in cultivations and animals. Weather is a significant factor in agriculture, which does not submit to traditional methods of risk control. Weather conditions are a serious source of uncertainty in agribusiness. Drought or excessive precipitation threaten harvest practically worldwide. As a result of global climate change variability of temperature in the world increases and apart from this more and more often weather phenomena take the form of extreme events. These risks affect not only cultivation, but also the efficiency of breeding farms, the use of artificial fertilisers or demand for different types of agricultural products. This means that natural risk touches upon different areas of economy directly or indirectly related to agriculture. Governments of countries remain neutral to hazards, by organising and financing diverse forms of aid for farmers aggrieved by natural disasters.

As an instrument of stabilisation of agricultural farms (micro perspective) and agricultural economy as a whole (macro perspective), agricultural insurance can perform the role of development stimulus, improvement in quality and improvement in the degree of agribusiness modernisation, and, as a consequence – growth in its competitive capacity on the EU market (Strupczewski, 2014:69) [12].

Geographic location of Poland at the contact of continental climate and Atlantic climate impact, and large surface of the country, cause considerable regional differences in potential consequences of natural risks. Often aggregated data analysis to the level of a country does not reflect variability of local losses. It creates a hazard of inadequate determination of the potential effect of natural hazards on agriculture, and hence the adaptation needs towards natural disasters.

Uneven distribution of natural risks posing a threat to cultivation and agricultural real estate is one of development barriers of agricultural insurance.

The paper can act as a guide to policymakers who are interested in understanding the structure of losses in agriculture, as it can influence public financial aid for farmers affected by natural disasters. This article highlights certain issues that both policymakers and insurance companies can utilize further for their own purposes to design better risk management tools of mitigating natural disaster losses in agriculture. Finally, by providing relevant data on nature of financial losses in agriculture, the paper delivers broader perspective on factors determining agriculture development.

The purpose of the research is an attempt to classify voivodeships into clusters showing differences between them according to the criterion of sizes of damage in agricultural farms caused by natural disasters.

The subject of the research are losses in agriculture (cultivation and fixed assets) caused by various kinds of natural risk, which occurred in the period 2010-2013 in Poland.

Natural threats in agriculture – outline of the issue

Classification of kinds of risks in agriculture

Understanding the sources and the nature of risk is the main condition of building an appropriate policy of risk management. Necessary factors are: the analysis of distribution, frequency of incidence and the financial effects of hazards most important from the point of view of the continuity of farms operation. However, the starting point for advanced analyses should be preparing an appropriate typology of risks that could occur in agricultural operations.

Agriculture, to a greater extent than other branches of economy, is exposed to natural risk, especially because of the fact that taking this kind of risk may take catastrophic sizes. The specific nature of agricultural production risk results from many reasons rooted in specific nature of agriculture as a branch of economy (Pope, 2003:128; Stroinski, 2006:22-23). [9,10]

The demand for agricultural raw materials is characterised by higher concentration than consumption demand for ready-made food products. It means that farmers and final recipients are the takers of prices, which are affected by price shocks on the international market. Many agricultural producers are guided by the guess that markets are not fair for them because of asymmetry of the market force.

(i)Agricultural produce are goods characterised by low price flexibility of demand AND supply. Profit flexibility of many agricultural raw materials is also at a low level as compared to production goods or services. For this reason, various kinds of market shocks have greater impact on the agricultural sector.

(ii)The dependence of a production cycle on biological factors, which are characterised by long period between decision-making and obtaining final effects.

(iii)Scarce possibility of alternative use of arable land (except for land located in the vicinity of cities).

(iv)Direct exposure of agricultural cultivations to continuous weather conditions (rain, sunlight, wind, frost, hailstorm, diseases, pests) and limited possibilities of protection against them.

(v)Limited impact of an agricultural producer on the location of cultivations, and consequently some group of farms may be exposed to recurring losses arising from repeated events in a given area.

(vi)Small possibilities of farm's property protection arising from "open" nature of conducted activities (theft, loss, vandalism).

(vii)Agricultural and animal production does not give the possibility of exact planning of sizes of production and potential revenue (fluctuations of agricultural market prices, fluctuations in crops).

(viii)Substantial impact of policy of the state on the earned income from agricultural operations (e.g. Common Agricultural Policy of the EU, minimum prices system, subsidising agricultural production, protection actions).

(ix)Seasonal nature of production and capital intensity and inconvenience of warehousing of ready-made products.
(x)Agrarian culture and agricultural production methods are diverse worldwide.
(xi)Growth in mechanisation in agriculture denoting risk of accidents at work.
The above described specific characteristics of agricultural production, conditioning agricultural production on climatic and biological factors, the dynamics of natural factors, conditioning product prices on market conditions – are key determinants of typology of risk present in the agricultural production sector.
Risks found in agriculture can be divided into two basic groups (Pope, 2003:127) [9]: (a)price risk – resulting from agricultural market liberalisation, (b)production risk – related to the occurrence of unfavourable unforeseeable events (e.g. natural disasters, embargoes for export of agricultural products).
Considering the range of impact, OECD introduces the following risk typology (OECD, 2008) [7]:
-*specific risk* (idiosyncratic), referring in micro-scale to single business entities. The risk level depends on individual decisions and is partially conditioned by the level of knowledge and management skills;
-*common risk* (interdependent) in the mezoeconomic scale and owing to risk factors affecting groups of entities with common characteristics (e.g. farms specialising in specified direction of production) or entities in the areas limited territorially (local communities);
-*system risk* is present in the macroeconomic scale and covering events potentially influencing the overall business entities or their substantial part in a regional scale. System risk, also called basic, is determined by external forces and cannot be controlled by single persons or entities. Within this risk category there can be distinguished, among others, market, political, currency, inflation risk and a number of factors related to the forces of nature.
Hardaker *et al*.(1997) [2]mentioned six risks typical of agribusiness:
-*personal risk* – accident at work or death of a farm owner,
-*material risk* – destruction or loss of tangible assets of a farm,
-*production risk* – variability of income arising from the impact on the financial result of diverse internal and external factors,
-*price risk* – variability of purchase prices of agricultural produce,
-*institutional risk* – the possibility of unexpected change in the agricultural market as a result of public administration intervention,
-financial risk – loss of liquidity, changes in interest rates, depletion of own capital.
Jerzak (2006) [4] when proposing his own typology lists only four, though quite extensive groups of agricultural risks:
-*natural risk* (the presence of natural disasters),
-*technological risk* (technical progress),
-*organisational risk* (planning, controlling and organising agricultural production),
-*economic risk* (the impact of macroenvironment on prices, means of production, agricultural policy, demand, form of ownership, structure of income).

Table 1. Risk categories in agriculture

Type of risk	Micro (specific)	Mezo (trade)	Macro (general)
Market /price	---	Change in land prices, new requirements of the food industry	Changes in product prices and means of production prices caused by shocks, commercial policy, endogenic changeability
Production	Hail, frosts, non-infectious diseases, personal risk, assets risk	Rain, land sliding, environment contamination	Flood, drought, plagues, infectious diseases, technologies
Financial	Changes in income beyond farms	---	Changes in interest rates and value of financial assets
Institutional - legal	Legal liability	Changes in local policy (regulations)	Changes in regional and national policy, regulations of environmental protection, payments for agriculture

Source: Majewski et al., 2008:167[5]

Scientific Papers Series Management, Economic Engineering in Agriculture and Rural Development
Vol. 15, Issue 1, 2015
PRINT ISSN 2284-7995, E-ISSN 2285-3952

A synthesis of the so far presented risks classifications may be a holistic and multi-dimensional matrix of kinds of risk in agriculture presented in Table 1.

Natural risks in agriculture

In plant production the basic risk factors are unfavourable weather conditions, such as: drought, hurricane, hailstorm, spring frosts or heavy rains.

They cause not only direct losses in yield of cultivated plants, but may also worsen the quality of products, sometimes substantially (e.g. hail damage fruit, deterioration in biochemical parameters under drought conditions, intensifying outbreaks of plants diseases at excessive precipitation).

The natural and random nature, partially connected with the course of weather, is also observed in the case of occasional intensification of presence of pathogens of cultivated plants (weeds, insects) which may lead to any unpredicted falls in harvest or generate increased costs of plant protection (Majewski et al., 2008:168). [5]

Table 2.Natural risks in agriculture

Content	Fire	Hurricane	Flood	Hailstorm
General risk characteristics	-fires in agriculture constitute 25% of all fires, -small fires prevail (90%), medium fires are approx. 10%	-sudden occurrence -mass losses occur often -variable intensity of risk during the year and in long-term periods	-may reach mass and catastrophic sizes -mostly floods caused by precipitation occur -for the crops the most severe flood is in June-July	-mainly damage to cultivations -small fluctuations of loss ratio in the long-term perspective -occurrence from May to August
Frequency of damages	-10 fires per 1000 agricultural buildings, - 1 fire per 206 farms	-11.7 damage per 1000 agricultural buildings	-1.22% of area of cultivations in Poland annually on average	-1.32% of area of cultivations in Poland annually on average
Risk intensity	Average degree of damage of a brick building 38%	Damage relate mainly to roofing and the roof structure	The average degree of cultivation damage from 36 to 75%	The average degree of cultivation damage from 25 to 35%
Average structure of material losses	-buildings 55% -cultivation 16% -dead inventory 12% -home movables 10% -livestock 6% -building materials 1%	-buildings 85% -movable property 9% -cultivation 6%	-cultivation 97% -movable property 2% -buildings 1%	Types of cultivations most vulnerable to hailstorm: orchards, tobacco, hemp, vegetables
Regional division of losses	-lack	-the largest hazard: Central Pomerania, the Suwałki Region, the Beskidy region, Bieszczady Mountains, the Mazovia region	-the largest risk along the course of the Oder and the Vistula rivers	-southern Poland

Source: prepared by the author on the basis of (Stroinski, 2006) [10]

Risk factors in animal production are first of all related to the hazard of the presence of epidemic diseases of systemic nature. Directly, they can cause significant losses in animal herds (in extreme cases a total elimination of animals in the herd) in the areas limited to the regional or local scale. Indirectly, they may cause decrease in demand and prices of particular products, thus adversely affect the situation of all producers on a national or supranational scale.

Stroinski (2006) [10] published interesting study of hazards present in agriculture from the point of view of insurance companies.

The most important information regarding fire, flood, hurricane and hailstorm risk are gathered in table 2.

For buildings in an agricultural farm the largest hazard are hurricanes and fires, as they are characterised by the frequency of occurrence and intensity of impact. While fires create normally individual risk, hurricanes may cause mass damage, though concentrated on a limited area. Agricultural cultivation can be harmed as a result of flood or hailstorm, namely risks towards which it is difficult to use effective prevention methods within a broad area. No wonder that in the light of the data, flood may destroy even 75% of cultivation (regardless of its type). The destructive impact of hailstorm mostly affects orchards, vegetables, tobacco and hemp. The presence of flood and hailstorms is subject to clear regional division, as opposed to fires.

MATERIALS AND METHODS

The classification of voivodeships into clusters showing differences between them according to the criterion of sizes of damage in agricultural farms caused by natural disasters has been the goal of this research work.

In this purpose, the losses registered in agriculture (cultivation and fixed assets) caused by various kinds of natural risk, and occurred in the period 2010-2013 in Poland have been taken into consideration.

The analysis has been based on the data of the Agency for Restructuring and Modernisation of Agriculture (ARMA) collected in connection with the aggrieved farmers filing applications for payment of the so-called "natural disaster loans", i.e. preferential loans for resuming production in agricultural farms and special departments of agricultural production where damage were caused by drought, hail, heavy rain, negative effects of wintering, spring frosts, flood, hurricane, lightning, soil slide or avalanche. The collected data relate to: the value of damage in cultivations, the value of damage in fixed assets, the number of aggrieved agricultural farms, area of damaged cultivations.

Voivodeships were adopted as research objects. The statistical analysis included total values including the whole temporary range of the research. Arithmetic mean could distort the results in the event of extreme events.

Research procedure includes the following stages of procedure:(1)identification of the objective of the research, (2) definition of research hypotheses, (3) preliminary data analysis, (4) identification and selection of independent variables, (5) ordering data set with agglomeration method, (6)validation of agglomeration results with the use of k-means, (7)verification of hypotheses and formulation of conclusions.

Implementation of the scheduled research will make it possible to empirically verify the following research hypotheses:

H1.High value of loss in agriculture in the absolute perspective in a given voivodeship does not have to mean high vulnerability of agricultural farms from this voivodship to natural risks.

H2.There is spatial diversity of distribution of damage in agriculture caused by natural risks.

The article consists of three basic parts and introduction and summary of conducted research. After formulating the research problem, objective and research hypotheses in the introduction, the special character of agricultural activity was discussed, along with the typology of kinds of risk connected with agriculture, and a review was made of the most important natural hazards affecting agribusiness in Poland. The next part of the article presents a number of analytical statements showing forming losses in agriculture across voivodeships. Then, cluster analysis was conducted by voivodeships with the use of agglomeration method and k-means, which was presented in detail in the third part of the study. At the end a summary was made of the concerned issues, with particular focus on the issues of agricultural insurance.

The details about the mathematical models used in this research are presented within the paragraph Results and Discussions.

RESULTS AND DISCUSSIONS

Preliminary data analysis

Record losses in agriculture in the period of 2010-2013 were observed in the Mazowieckie Voivodeship – both in cultivations (PLN 1.3 billion) and in fixed assets (PLN 372 million). They constituted 27% of all damage in agricultural farms (Fig. 1 and Fig. 2).

The region is regularly affected by natural disasters, which is proven by an exceptional level of losses in subsequent years. In critical 2013 the share of value of destroyed cultivations in the Mazovia region in relation to the whole country was 56%, and concerning fixed assets – 81%. At the same time it is difficult to indicate one main cause of such large damage. The problem is rather high intensity and the frequency of the presence of such hazards as: flood, hailstorm, hurricane, intensive rainfall, frosts.

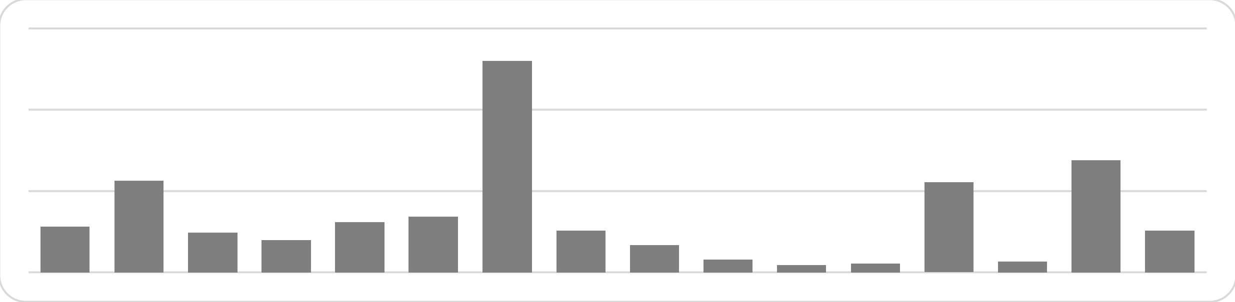

Fig. 1. The size of losses in cultivations in the period of 2010-2013 (million PLN)
Source: prepared by the author on the basis of ARMA data.

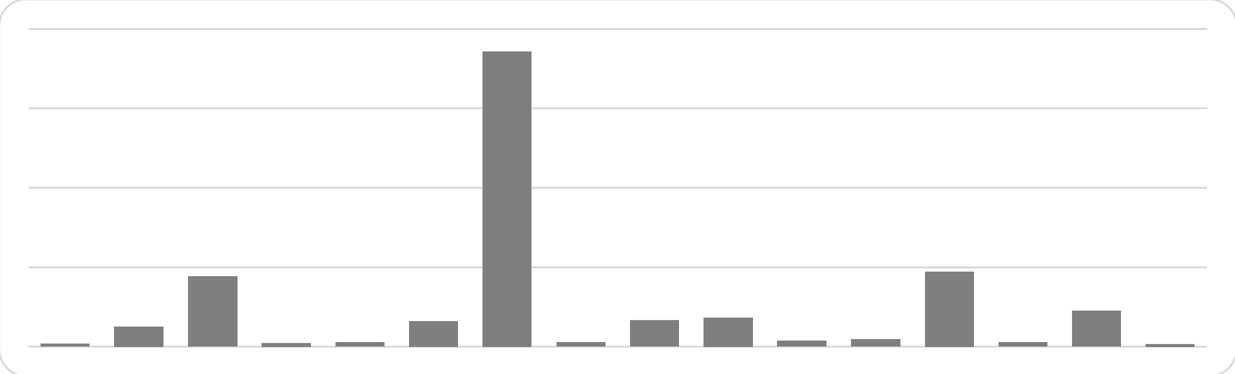

Fig. 2. The size of losses in fixed assets in the period 2010-2013 (million PLN)
Source: prepared by the author on the basis of ARMA data.

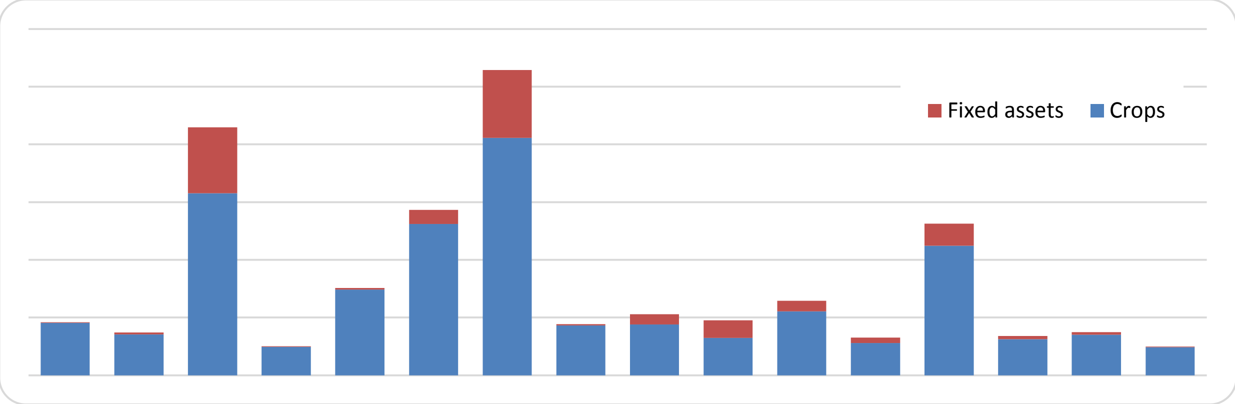

Fig. 3. The average value of losses per 1 ha of damaged area in the period 2010-2013 (thousand PLN)
Source: prepared by the author on the basis of ARMA data.

The areas endangered to a significant extent with natural disasters also include voivodeships: Wielkopolskie, Świętokrzyskie and Kujawsko-Pomorskie (losses in cultivations in the range of PLN 550-700 million, and in fixed assets – below PLN 100 million).

The analysis of average level of damage per 1 hectare of destroyed cultivations reveals substantial deviation in the result of four voivodeships (Mazowieckie, Lubelskie, Małopolskie, Świętokrzyskie) from others (Fig. 3). This may prove concentration of agricultural production within cultivations representing the highest value (e.g. vegetables, orchards) or multiple occurrences of accidents on this area.

Considering the average volume of losses recorded in the aggrieved agricultural farms (Fig. 4), it is possible to note a clear division into three groups of voivodeships.

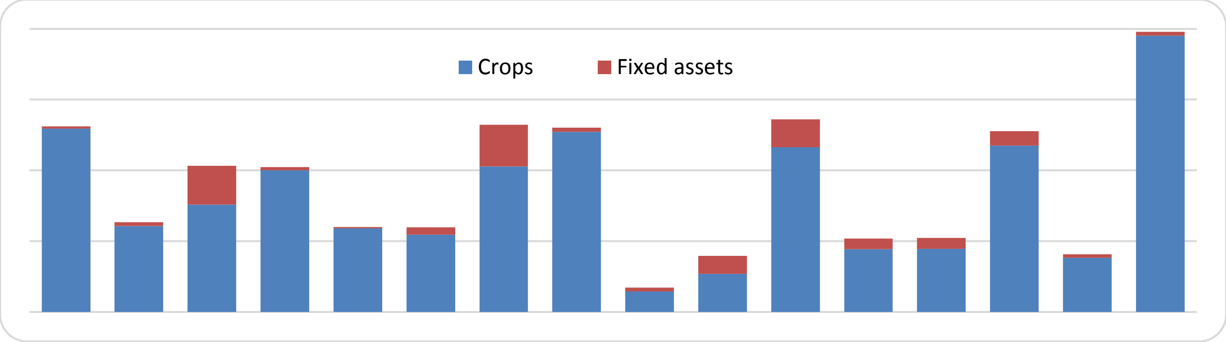

Fig. 4. The average value of loss per 1 aggrieved farm in the period 2010-2013 (thousand PLN)
Source: prepared by the author on the basis of ARMA data.

The largest sensitivity to destruction occurred in the Zachodniopomorskie Voivodeship (nearly PLN 80 thousand losses per farm in the period of 4 analysed years). The second group is formed by farms with average susceptibility to damage (from PLN 40 to 55 thousand) – Voivodeships: Dolnośląskie, Lubelskie, Lubuskie, Mazowieckie, Opolskie, Pomorskie and Warmińsko-Mazurskie. In the remaining regions the value of this ratio does not exceed PLN 26 thousand – group of low-susceptibility to damage.

The size of losses in particular voivodeships should be also analysed in the context of potential of a given region measured by the total number of agricultural farms as well as whole arable lands. The relation of damage size and resources related to agriculture in the region allows determining the average annual susceptibility to natural risk ratio (Fig. 5).

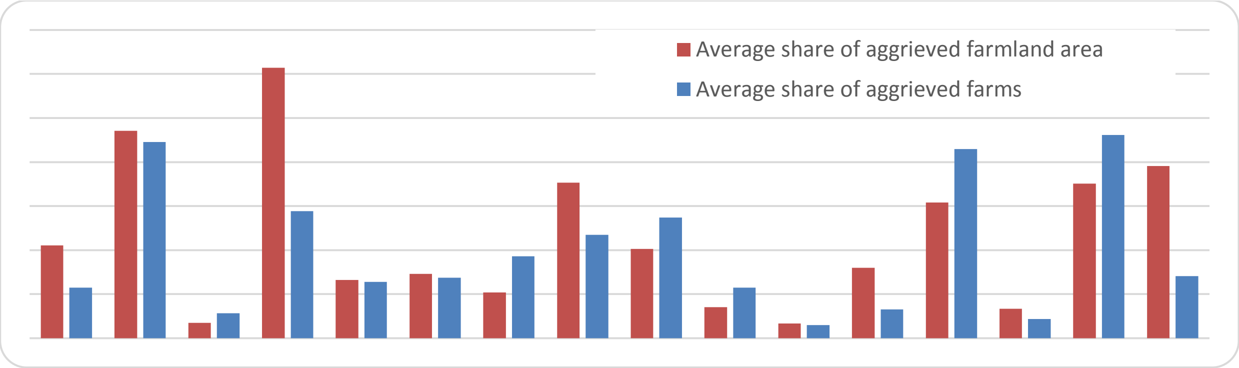

Fig. 5. Susceptibility to natural risk by voivodeships
Source: prepared by the author on the basis of ARMA data.

Mazovia, where the largest losses occurred in the absolute perspective, paradoxically belongs to voivodeships with the lowest vulnerability ratios, both in the area perspective (average participation of the aggrieved arable lands) and concerned (average participation of aggrieved farms). The most aggrieved voivodeship turned out to be the Lubuskie Voivodeship, where on average damage occurred on every eighth hectare of cultivations. A little lower range of damage in cultivations was recorded in the Kujawsko-Pomorskie Voivodeship (9.4%) and the Zachodniopomorskie Voivodeship (7.8%). As regards the percentage of aggrieved farms, the highest ratios were observed in three voivodeships: the Wielkopolskie Voivodeship (9.2%), the Kujawsko-Pomorskie Voivodeship (8.9%) and the Świetokrzyskie Voivodeship (8.6%).

The above conclusions confirm at the same time H1 hypothesis made in the introduction that distributions of losses in the absolute and relative perspective do not correspond. A proper assessment of sizes of damage in voivodeships requires reference to the potential of the region and the degree of agricultural development within its area.

Determination of comparison objects and selection of diagnostic variables

In accordance with the previously outlined objective of the research and formulated hypotheses, objects of comparisons were voivodeships as the highest level of the administrative division of Poland. The selection of diagnostic variables was made on the basis of the statistical criterion that includes the information value of variables. In the statistical criterion two factors are taken into account: discriminatory ability of variables and their capacity (Panek et al., 2013:21-23) [8]

Discriminatory ability of variables (variability in respect of the examined objects) is measured by means of a classic variability coefficient. From the data set we eliminate variables, whose variability coefficient does not exceed the stated threshold value adopted at the level of 0.1. All the variables were accepted (table 3).

Table 3. Definitions and descriptive statistics of variables

Variable name	Description	Arithmetic mean	Standard deviation	Variability coeff. (%)
LOSS_CROP	Total value of damage in cultivations in the period of 2010-2013 (PLN)	337,755,177	319,263,338	94.53
LOSS_ASSETS	Total value of damage in fixed assets in the period of 2010-2013 (PLN)	48,322,944	90,982,657	188.28
AREA	Total area of destroyed cultivations (ha) in the period of 2010-2013	151,257	132,105	87.34
FARMS	The total number of aggrieved agricultural farms in the period of 2010-2013	14,288	13 606	95.23
SHARE_FARMS	Average percentage of aggrieved farms in the overall number of farms in the voivodeship	0.039	0.029	73.64
SHARE_AREA	Average percentage of area of damaged cultivations in the overall area of arable lands in the voivodeship	0.046	0.034	74.59

Source: prepared by the author.

Capacity (information potential) of variables means the level of correlation with other variables. Information capacity of variable is the greater, the weaker it is correlated with other variables and at the same time the stronger it is correlated with variables that are not taken into account in the finally adopted set of diagnostic variables. Correlation is interpreted as transfer of the same information in compared objects. The basic verification method of information capacity of quantitative variables is a matrix of Pearson's linear correlation coefficients (see table 4). On the contrary, a complex tool of information capacity analysis of variables-the so-called parametric method – was prepared by Z. Heellwig in 1968[3].

At the beginning the critical value of the correlation coefficient r* should be determined, above which two variables will be assessed as excessively mutually correlated. This can be done by means of a formal method using the procedure of verification of significance of correlation of diagnostic variables. At the beginning r* value is determined using the following formula (Panek et al., 2013:23) [8]:

$$r^* = \sqrt{\frac{t_{\alpha,s}^2}{t_{\alpha,s}^2 + n - 2}}$$

where:

$t_{\alpha,\, p}$ – value from distribution table t - Student for $s=n-2$ degrees of freedom and the adopted level of significance α (α = 0.05).

From the distribution board of t - Student the value was read of statistics t = 2.7764 for the level of significance α = 0.05 and s=4 degrees of freedom.

Then the critical value r* was calculated:

$$r^* = \sqrt{\frac{2{,}7764^2}{2{,}7764^2 + 6 - 2}} = 0.8114$$

Border value of the correlation coefficient was calculated by means of a formal method and is thus 0.8114.

Further stages of selection of variables are determined by the so-called parametric method.

Table 4. Pearson's linear correlation coefficients matrix between the variables

No.	Variables	1	2	3	4	5	6
1	LOSS_CROP	1.0000	0.8456	0.4707	0.7240	0.4792	0.1277
2	LOSS_ASSETS	0.8456	1.0000	0.0075	0.4813	0.0906	-0.2317
3	AREA	0.4707	0.0075	1.0000	0.5666	0.7642	0.6873
4	FARMS	0.7240	0.4813	0.5666	1.0000	0.7617	0.1762
5	SHARE_FARMS	0.4792	0.0906	0.7642	0.7617	1.0000	0.7056
6	SHARE_AREA	0.1277	-0.2317	0.6873	0.1762	0.7056	1.0000

Source: prepared by the author.

The procedure of a parametric method proceeds according to the following scheme (Panek et al., 2013:24):

1.Determination of the median of each R correlation matrix column:

$$R_{j'} = M_j\left(r_{jj'}\right) \quad ,j,j' = 1,2,\ldots,m$$

Application of position measure, such as the median, allows to increase the resistance of the obtained results to values of diverging diagnostic variables (Mlodak, 2006:31) [6].

2.Finding a column, for which R_j is the highest.

3.In the indicated column, the selection of elements with absolute values greater than values the threshold value r * (0.8114) and the identification of lines corresponding to these elements.

The variable corresponding to the distinguished column is called a central variable, while the variables corresponding to the distinguished lines – satellite variables (of a given central variable). Satellite variables duplicate information included in the central variable and therefore they should be removed from further analysis.

4. Reduction in correlation matrix R by crossing out columns and lines corresponding to central and satellite variables.

5. Repetition of steps 1-4 until exhausting the set of acceptable diagnostic variables.

The final set of diagnostic variables will include all the identified central variables and isolated variables (i.e. variables, which were not substantially correlated with any other variable).

As a result of performing the above described procedure, variable LOSS_ASSETS was rejected. The remaining five variables were qualified for further research taking into account the criteria of discriminatory ability and information capacity.

Stimulation of variables

The application of a multi-dimensional comparative analysis requires that the diagnostic variables have a uniform nature – a stimulating factor (Panek et al., 2013:33) [8]. On the basis of substantive premises it may be concluded that all the variables found in the study are destimulants, as their high values in the examined objects are undesirable from the point of view of a given phenomenon (the higher values of measures describing the number and the value of damage in agriculture, the worse for the voivodeship).

Owing to the fact that variables are destimulants and are measured on a quotient scale, a quotient transformation was used that transformed them into stimulating factors – also measured on the quotient scale. The form of this transformation is as follows (Panek et al., 2013:33) [8].:

$$x_{ij}^S = b\left[x_{ij}^D\right]^{-1}$$

where:

x_{ij}^D – the value of j variable destimulants in i object,

x_{ij}^S – the value of j variable after transformation into a stimulating factor in i object,

b – constant used in an arbitrary manner, here $b = 1$.

Standardisation of variables

Standardisation transformation is intended to obtain the comparability of variables (at least in the aspect of units of measurement) and standardisation of their scope of variability. It is required in the case of taxonomic methods (Panek et al., 2013:35) [8]. Considering the nature of variables, standardisation was selected by way of classic standardisation, as a result of which the arithmetic mean assumes the value of 0, and standard deviation the value of 1.

Elimination of negative variables values

Obtaining the required in taxonomic research characteristics of a positive value of variables took place as a result of applying the following transformation Grabinski et al., 1989:28) [1].:

$$z'_{ij} = \begin{cases} z_{ij} \text{ gdy } min_{i,j}\{z_{ij}\} > 0 \\ z_{ij} + \varepsilon \text{ gdy } min_{i,j}\{z_{ij}\} \leq 0 \end{cases}$$

provided that:

$$\varepsilon = -min_{i,j}\{z_{ij}\} + \frac{1}{5}S(z)$$

where:

S (z) – standard deviation calculated from all elements of the matrix of standardised input data.

Parameter value ε amounted to 1.132327.

The finally prepared set of variables after conducting their stimulation, standardisation and elimination of negative values is presented in Table 5.

Table 5. Independent variables prepared for agglomeration analysis

Voivodeship	LOSS_CROP	AREA	FARMS	SHARE_FARMS	SHARE_AREA
Dln	0.645586	0.573141	0.995709	1.035040	0.645908
Kpm	0.371626	0.241502	0.479977	0.275939	0.340303
Lbl	0.732207	2.226029	0.778582	2.084942	3.413902
Lbs	0.885006	0.451108	1.069958	0.418430	0.282433
Łdz	0.597954	0.844580	0.602148	0.929440	0.974390
Młp	0.548763	1.327969	0.555624	0.864228	0.890044
Maz	0.215078	0.564508	0.437886	0.640983	1.214919
Opo	0.702475	0.602459	1.053565	0.511949	0.423000
Pdk	1.023227	0.923260	0.447793	0.439883	0.667284
Pdl	2.072078	1.433456	0.826258	1.029156	1.751128
Pom	3.635462	4.314020	4.226787	3.984677	3.567498
Ślk	2.927351	1.761199	1.512076	1.800490	0.820876
Świ	0.376334	0.717144	0.440202	0.285800	0.470894
Wma	2.361958	1.585224	2.845524	2.706754	1.839381
Wlk	0.320473	0.200000	0.402772	0.266990	0.424546
Zpm	0.701659	0.351635	1.442374	0.842533	0.390730

Source: prepared by the author.

Ordering the data set by means of agglomeration method

Agglomeration, as one of the methods of hierarchy clusters analysis, allows to group similar objects. The measure of similarity are distances between the objects, and most often the so-called Euclidean distance is used. It is a particular case of Minkowski metric, applicable to variables measured on the ordinal quotient scale. It measures section length $d_{ii'}$ connecting objects in multi-dimensional space, which can be expressed by the formula (Panek et al., 2013:44) [8].:

$$d_{ii'} = \sqrt{\sum_{j=1}^{m} \left(z_{ij} - z_{i'j}\right)^2}$$

The results of analysis have the form of a tree diagram, which graphically illustrates clusters of similar objects owing to defined diagnostic variables. The system of connections in the tree diagram makes it possible to specify mutual location of objects with respect to each other and groups of objects created in subsequent steps of the procedure.

From among the existing agglomeration methods, it was decided to use two: **the farthest neighbourhood method (full binding)** and **the Ward's method**. This selection results from the intention of the Author, so that the effect of grouping are "clumps" of non-one-element objects[35].

The tree diagram contains larger average distances between bindings, thanks to which the results of agglomeration are more legible (Panek et al., 2013:108) [8]..

On the other hand, the advantage of the Ward's method is its high efficiency. It results from the use of approach based on variance analysis. In pursuit of minimising the sum of

[35]A reverse result than the intended one would be grouping results resembling "chains" of objects, created, for example, as a consequence of using the nearest neighbourhood method. "Snowball" effect is created, where a big group "attracts" single observations.

squares of deviations inside clusters ESS[36], the pairs of clusters are chosen, which as a result of connection will give a cluster with minimum diversity.

As a result of the agglomeration procedure by the **method of full binding** (using the Euclidean distance) a tree diagram was obtained shown in Fig. 6.

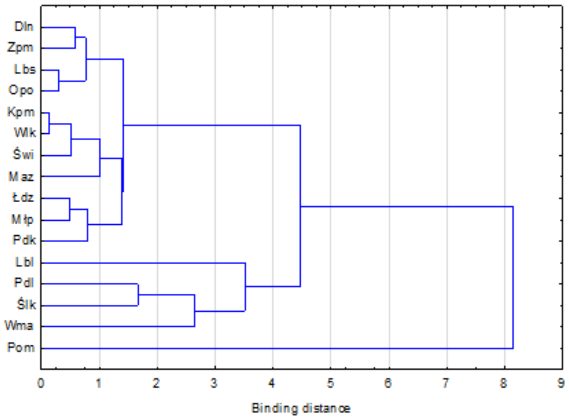

Fig. 6. A tree diagram was made by means of a full binding method
Source: own calculations made in the Statistica program.

In order to determine the place of 'cut-off" the diagram was analysed of binding distance in respect of binding stages (Fig. 7).

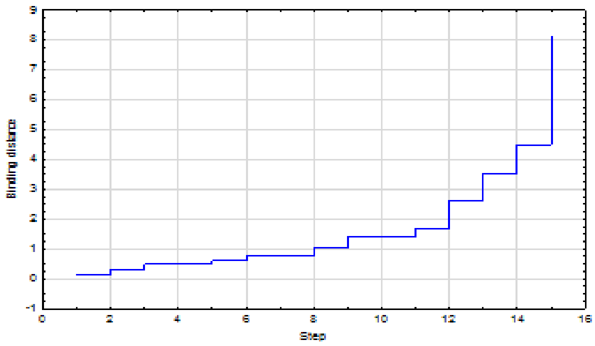

Fig. 7. The diagram of binding distance in respect of binding stages – full binding method
Source: own calculations made in the *Statistica* program.

The place where there is clear flattening (longer vertical line), determines the optimum cut-off point. Step 13 was decided to be taken with the binding distance equal to 3. This means that four clusters were formed.

Then agglomeration was made with the **Ward's method**, whose course is illustrated in the tree diagram in Fig. 8. Like before, after analysis of the binding distance diagram in respect of binding stages (Fig. 9) a cut-off point was selected at the distance equal to 3 and the thirteenth step of agglomeration procedure, which resulted in the formation of four clusters.

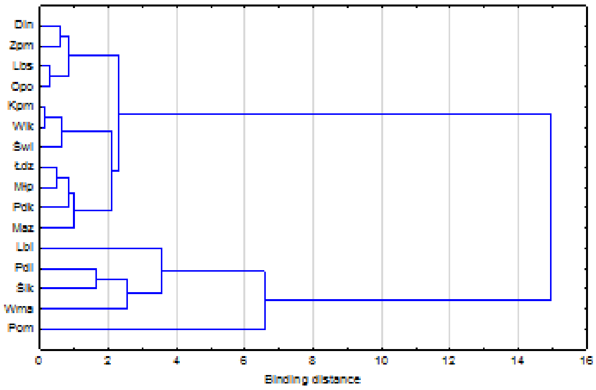

Fig. 8. The tree diagram made by means of the Ward's method
Source: own calculations made in the *Statistica* program.

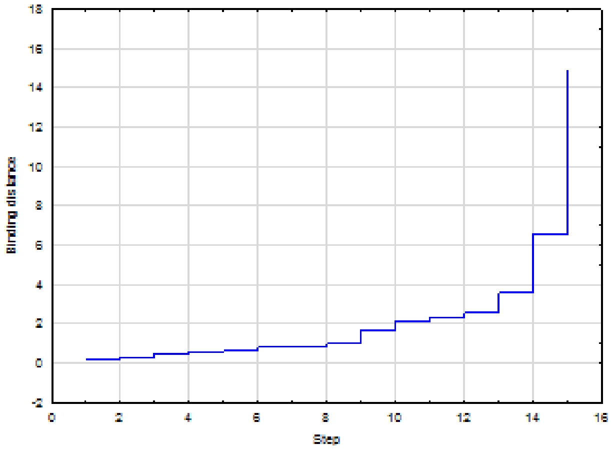

Fig. 9. The diagram of binding distance in respect of binding stages – the Ward's method
Source: own calculations made in the *Statistica* program.

Table 6 presents a comparison of agglomeration results with the Ward's method and full binding method.

Both agglomeration methods rendered identical results of voivodeship division into clusters.

The created group significantly vary among themselves. Apart from two one-element clusters, a cluster was created with 11 voivodeships. What is interesting, such a great cluster is created already at the initial stage of the agglomeration process at a relatively limited binding distance (full binding method: 1.5; Ward's method: 2.3). This proves very

[36] ESS (Error Sum of Squares).

similar properties of regions included in it and the lack of possibility to separate a greater number of smaller clusters.

Table 6. Summary results of agglomeration procedure

Full binding method		Ward's method	
Cluster number	Elements of cluster	Cluster number	Elements of cluster
1	POM	1	POM
2	LBL	2	LBL
3	PDL, ŚLK, WMA	3	PDL, ŚLK, WMA
4	DLN, KPM, LBS, ŁDZ, MAZ, MŁP, OPO, PDK, ŚWI, WLK, ZPM	4	DLN, KPM, LBS, ŁDZ, MAZ, MŁP, OPO, PDK, ŚWI, WLK, ZPM

Source: prepared by the author

Grouping with the k-means method

The k-means method is the most often used non-hierarchical taxonomic method of grouping. Its result is division, in which no cluster is a sub-cluster of another set. The starting point of the analysis is setting the *a priori* specified number of clusers (k) that will be formed in a way to minimise the intra-group variability and maximise the inter-group variability. Research procedure is of iterative nature, where the researcher has the opportunity to impose the upper limit of the number of iterations, after which stopping the process of grouping takes place.

K-means analysis is supposed to verify the correctness of the grouping results with the agglomeration method.

This research contains grouping with cases up to 4 clusters, specifying the maximum number of iterations for 15. The selection of initial centres of clusters took place by way of maximising initial distances between clusters.

As a result of analysis with the k-means method, the following division of voivodeships into clusters was obtained:

Cluster 1: PDL, ŚLK, WMA

Cluster 2: DLN, KPM, LBS, ŁDZ, MŁP, MAZ, OPO, PDK, ŚWI, WLK, ZPM

Cluster 3: LBL

Cluster 4: POM

The essence of each cluster can be recognised as a result of means analysis within each of them (Fig. 10). When interpreting the diagram it should be remembered that the data for analysis were subject to stimulation (with destimulants into stimulating factors), therefore, the highest means values indeed mean their lowest levels in raw data (and thus a more preferred situation).

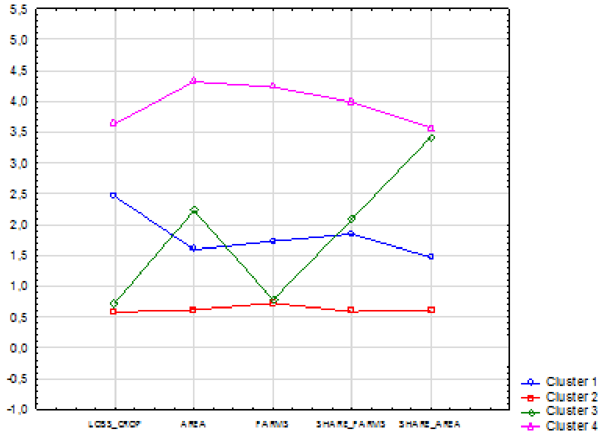

Fig. 10. The diagram of means of each cluster
Source: prepared by the author

The Pomorskie Voivodeship (Cluster 4) is a region, in which the lowest values of losses in agriculture were recorded in each of five examined variables. Cluster 3 represents the Lubelskie Voivodeship, where the percentage of area of arable lands affected by losses in relation to the total area of arable lands within the voivodeship belonged to the lowest. On the contrary, losses in cultivations expressed by value were shaped there on a relatively high level. Voivodeships characterised by relatively low values of losses in cultivations with slightly increased values of average other variables created cluster no. 1 (the Podlaskie, Śląskie and Warmińsko-Mazurskie Voivodeships). The remaining 11 voivodeships create cluster 2, characterised by the highest average values of every diagnostic variable. These are thus regions, where agricultural farms suffered due to natural risks to the greatest degree.

These conclusions confirm the authenticity of

hypothesis H2, in accordance with which there is clear spatial diversity of distribution of damage in agriculture caused by natural risks.
The results of variance analysis, supplementing the interpretation of the k-means method show (F and q values) that LOSS_CROP variables and AREA constitute the main criterion determining affiliation to clusters (table 7).

Table 7. The results of variance analysis

Variable name	Between SS	df	Internal SS	df	Statistics F	p-value
LOSS_CROP	15.01131	3	0.988692	12	60.73197	0.000000
AREA	14.86739	3	1,132606	12	52.50687	0.000000
FARMS	12.62868	3	3.371317	12	14.98368	0.000232
SHARE_FARMS	13.78141	3	2.218585	12	24.84721	0.000020
SHARE_AREA	14.46465	3	1.535347	12	37.68439	0.000002

Source: prepared by the author

CONCLUSIONS

Agricultural production risk is of special nature due to a great number of hazards, relative weakness of production entities on the market, greater than in industrial production unpredictability of variability of phenomena (Strupczewski, 2014:596) [11]. In addition, it is intensified by the difficulty of adjusting once undertaken actions resulting from a long cycle of agricultural production.
Natural disasters occurring very frequently, at simultaneous low percentage of insured farmers, cause damage of such sizes that force the state to organise current financial aid (for instance in the form of preferential natural disaster loans). This aid is not sufficient.
Regional diversity of the risk level, substantially does not positively affect the development of insurance. In voluntary insurance, the premium should reflect the real level of exposition to risk. This means that in areas with over-average risk there will be small demand for insurance caused by high cost of protection. On the other hand, relatively low premiums in the areas with low hazard may prove an insufficient incentive in the absence of the sense of need to buy insurance. The subsidising mechanism of premiums by the state, which is present in insurance of agricultural cultivations and farm animals, should aim at equallizing the level of premium within the whole country, contributing to the increase in commonness of insurance.
Apart from the problem of diversity of territorial intensity of risk level, there are two issues difficult to solve, limiting the possibility of insurance of agricultural producers: asymmetry of information and moral hazard.
The first factor is related to information asymmetry between the producer desiring to obtain insurance and the insurance company. It involves subjective and endogenic (namely dependent only on the agricultural producer) conditions affecting the management result, including income being the object of possible insurance. The problem is that the insuring party, i.e. the agricultural producer, knows much more about the potential risk and its factors in production than the insurer. Also much depends on their diligence and other volumes hard to observe and assess. Such asymmetry concerning information and real risk assessment may involve the problem of temptations of abuses. This is a risk due to which the tendency of insurance companies to enter into such insurance contracts decreases.
The temptation of abuses (the so-called moral hazard) is present when the insured party, after buying an insurance policy, as a result of this changes the way of production and management, neglects diligence, resigns knowingly from welfare or otherwise tries to increase the potential dimension or probability of losses, and hence damages. It is about intentional actions leading to risk and losses. On an agricultural farm these can include defined negligence in the use of procedures, in untimely e.g. use of chemicals, such as fertilisers, plant pesticides, in feeding, in

counteracting diseases, etc. As a result, it is assumed that insurance in agriculture should only cover the events and accidents leading to unintended losses, where it is possible to exclude possible impact of subjective factor, dependent on the farmer (Majewski et al., 2008:51) [5].

On the basis of the cluster analysis it was demonstrated that 11 voivodeships form quite a homogeneous group in terms of size of damage in agriculture (the value of damage in cultivations and the acreage of destroyed cultivations are two most important factors determining affiliation to the cluster), however, the profile of loss occurring in other five voivodeships has a very individual course and requires separate handling in the actuarial sense.

REFERENCES

[1]Grabinski, T., Wydymus, S, Zelias, A., 1989, Methods of numerical taxonomy in modeling socio-economic phenomena. PWN, Warsaw.

[2]Hardner, J., Huirne, B., Ruud, B.M., Anderson, J. R.,1997, Coping with risk in agriculture. CAB International, Wallingford.

[3] Helwig, Z., 1968, Application of the taxonomy method to the typological division of countries due to their level of development as well as resources and structure of qualified staff. In: Statistical Review 15 (4): 307-327.

[4] Jerzak, M.A., 2006, Basic issues of risk in agriculture. In: Jerzak, M.A., Czyzewski, A., (EDS.): Economic conditions of the use of market prices stabilization tools in risk management in agriculture. Poznan University of Life Sciences Publ., Poznan: 101-118.

[5]Majewski, E., Was, A., Cyganski, L., 2008, Risk factors in a farm. In: Hamulczuk, M. (Ed.): Price risk management and the possibility of stabilizing the income of agricultural producers. IERiGZ Publ., Warsaw: No. 113.

[6] Mlodak, A., 2006, Taxonomic analysis in regional statistics. Difin, Warsaw.

[7]OECD, 2008, An Assessment on Risk Exposure in Agriculture – a literature review. Working Party on Agricultural Policies and Markets, Trade and Agriculture Directorate, Committee for Agriculture.

[8] Panek, T., Zwierzchowski, J., 2013, Statistical methods of multidimensional comparative analysis. Theory and Applications. Warsaw School of Economics Publ., Warsaw.

[9] Pope, R.D., 2003, Risk and agriculture. Some issues and evidence. In: MEYER D.J. (ed.): The economics of risk. Upjohn Institute for Employment Research, Kalamazoo.

[10]Stroinski, E., 2006, Life and non-life insurance in agriculture. Academy of Finance Publ., Warsaw.

[11]Strupczewski G., 2014, Corporate Non-life Insurance Claims: Empirical Evidence from the Polish Market. In: European Financial Systems 2014. Proceedings of the 11th International Scientific Conference. Masaryk University, Brno.

[12]Strupczewski G., 2014, Why enterprises buy insurance? Theoretical aspects of corporate demand for property insurance. In: Baaken T., Teczke J. (eds.): Managing disruption and destabilization. International Management Foundation, Cracow-Muenster.

[13]Zmudzka, E., 2004, Climate background of agricultural production in Poland in the second half of the twentieth century. In: Acta Agrophysica 3(2): 399-408.