

## NON-PARAMETRIC ANALYSIS OF YIELD RISK IN LITHUANIAN CROP FARMING

Tomas BALEŽENTIS<sup>1</sup>, Dalia ŠTREIMIKIENĖ<sup>2</sup>,  
Rasa MELNIKIENĖ<sup>3</sup>, Zhiqian YU<sup>4</sup>

<sup>1,2,3</sup>Lithuanian Institute of Agrarian Economics, Kudirkos g. 18-2, LT-03105 Vilnius, Lithuania

<sup>4</sup>School of Economics and Statistics, Guangzhou University, No. 230, Wai Huan Xi Road,  
Panyu District, Guangzhou, 510006, Guangdong, P. R. China

E-mails: <sup>1</sup>tomas@laei.lt; <sup>2</sup>dalia@mail.lai.lt (corresponding author);

<sup>3</sup>rasa.melnikiene@laei.lt; <sup>4</sup>yzq\_8866@hotmail.com

Received 08 January 2017; accepted 20 April 2017

**Abstract.** Socioeconomic development requires meeting the goals of food security. Yield risk constitutes an important factor of farming business viability. As the Central and Eastern European countries have been affected by both economic and environmental transformations, there is a need to develop a robust methodology for assessment of yield risks in order to propose convincing guidelines for both farmers and government institutions in regards to risk management and viability of agricultural business in general. This paper attempts to devise non-parametric measures of yield risk for Lithuanian crop farming. The research covers the period of 2000–2015. County-level data from Statistics Lithuania are employed for the analysis. The non-parametric analysis of yield risk relies on information diffusion theory and linear moving average. The results indicate that there exist differences in yield trends, yield loss rates and yield risk among crops and regions. Maize, buckwheat and winter rape exhibited the highest yield risk. These results shed light on the extent of yield risks underlying crop farming in Lithuania and, to a certain extent, can be contrasted to situation in Central and Eastern European countries. Indeed, the obtained results can be applied in decision making at different levels of management.

**Keywords:** risk management, yield risk, Lithuania, crop farming, diffusion theory, non-parametric analysis.

**JEL classification:** C14, Q15.

### Introduction

Agricultural business depends on multiple external factors affecting both crop yields and prices. As agricultural sector is related to such important issues as food security and viability of rural areas, public support aims to manage or reduce the agricultural risks and thus ensure implementation of the goals of sustainable rural development (Breustedt *et al.* 2008; Bokusheva 2011; Shi, Jiang 2016; Bokusheva *et al.* 2016). By managing agricultural risks, farmers are able to streamline their activities and ensure stability of

their income. In the European Union (EU), the Common Agricultural Policy (CAP) also includes measures aimed at managing agricultural risks (e.g., income risk).

Among different types of agricultural risks, yield risk is important as it captures the fluctuations in agricultural output (Shi *et al.* 2013; Chavas *et al.* 2014; Chhatre *et al.* 2016; Ker *et al.* 2016). Such shocks constitute an immediate factor for fluctuations in farm income. Therefore, farmers need to fathom the level of risk they are to face, whereas government can use the estimates of risk in order to adjust public support schemes. Direct payments under the CAP are not differentiated across the regions of Lithuania. The payments for less favoured areas, though, can be used to increase attractiveness of farming in certain areas. All in all, appraisal of yield risk is likely to contribute to decision making in farming business (Peleckis *et al.* 2015).

Central and East European countries face certain additional circumstances that stress the need for risk management, among other implications. Specifically, these countries are the new Member States of the EU. Following the accession to the EU, these countries faced serious structural adjustments in terms of farm size and specialisation. In addition, climate change has altered the farming conditions (IPCC 2014; Povilaitis *et al.* 2013). Therefore, there is a need to assess the impacts of such shifts in regards to the agricultural risks.

The estimation of yield risk can be based on different approaches. In principle, one can follow the classification by Yuan *et al.* (2015), who made a distinction among probability-based and indicator-based measures of risk. Indicator-based measures are basically the aggregates of various indicators describing the extent and/or likelihood of the hazard. For instance, Girdziute *et al.* (2014) applied factor analysis to derive a composite indicator of agricultural risk. As regards the probability-based indicators, these are based on the estimation of statistical distributions (Kuziak 2016; Piontek 2016). The latter type of measures allows to draw conclusions about the extreme events which would be hard to estimate otherwise. Probability-based measures can further be divided into parametric and non-parametric ones. Parametric measures rely on parameterized statistical distributions, while non-parametric measures follow distribution-free approach and require no assumptions about the family of the underlying distributions. These approaches might yield different results. Accordingly, there is a need to apply different approaches to obtain more robust results.

Among the non-parametric approaches, information diffusion theory (Huang 1997, 2005) can be employed to estimate the yield risk as well as other types of risk. The latter approach has been applied in different areas. Hao *et al.* (2012) and Xie *et al.* (2016) applied the diffusion theory for analysis of drought risk in China. Liu *et al.* (2013) employed the diffusion theory approach to estimate and combine risks for multiple hazards. Hao *et al.* (2014) utilised the same approach in order to estimate the risk of biological disasters in grasslands. Chen *et al.* (2015) followed the same vein when identifying risks associated with rice production in China.

This paper seeks to estimate the yield risk for different crops in Lithuania. Indeed, Lithuania is endowed with the largest area of agricultural land if opposed to the other Baltic States. Specifically, the areas of agricultural land amounted to 2.95 million ha,

1.87 million ha, and 0.97 million ha in Lithuania, Latvia, and Estonia, respectively, as of 2014 (FAO 2017). Due to geoclimatic conditions, Lithuania is rather similar to the other CEE countries (which does not fully apply for Estonia and Latvia) in terms of crop yields. However, Lithuania still lags behind the Western and some Northern European countries. For instance, the data from Farm Accountancy Data Network (European Commission 2017) show that the average wheat yield for 2004–2013 amounted to 7 t in Germany, 5 t in Poland, 4.3 t in Lithuania, 3.7 t in Latvia, and 3.2 t in Estonia. Baležentis and Kriščiukaitienė (2016) looked into the issue of yield risk in Lithuanian crop farming by means of parametric analysis, namely by estimation of probability-based measures. The current paper applies the non-parametric approach to gain further insights into the issue of yield risk in Lithuanian crop farming under the changing environmental and economic situation. The research covers the period of 2000–2015. County-level data from Statistics Lithuania (2016) are employed for the analysis.

## 1. Methods

This section presents the key concepts for the analysis of yield risk. Specifically, the non-parametric analysis of yield risk relies on the three main elements: 1) information diffusion theory; 2) yield loss rate; and 3) trend estimation via linear moving average (LMA).

### 1.1. Information diffusion approach and non-parametric modelling of risk

Parametric modelling of yield risk requires assumptions regarding the shape of the underlying statistical distribution (e.g., Gaussian, log-normal, Burr etc.). Such a choice involves a kind of uncertainty as the chosen distribution might not be flexible enough to capture certain peculiarities of the phenomenon analysed. Therefore, another type of models, viz. non-parametric ones, can be applied in order to describe the probabilities of extreme events, which, indeed, one of the most important measures of risk. The non-parametric models allow for a less restrictive analysis in the sense of the shape of statistical distributions. Indeed, the analysis can be made more robust by employing and comparing multiple techniques.

Huang (2005) proposed a non-parametric information diffusion approach, which is closely related to kernel density estimation. The said approach has been applied in modelling agricultural risks (Hao *et al.* 2012, 2014; Chen *et al.* 2015). Let  $X$  be a sample, which is diffused to set  $U$ , which, indeed, can be regarded as the universe of discourse of  $X$ . Specifically, let  $X = (x_1, x_2, \dots, x_m)$  and  $U = (u_1, u_2, \dots, u_n)$ . Furthermore, let indexes  $i = 1, 2, \dots, m$  and  $j = 1, 2, \dots, n$  be used to index the elements of  $X$  and  $U$ , respectively. The likelihood to observe the value of  $u_j$  at a certain sample point,  $x_i$ , is defined as follows (Huang 2005; Chen *et al.* 2015):

$$f_i(u_j) = \frac{1}{h\sqrt{2\pi}} \exp\left(-\frac{(x_i - u_j)^2}{2h^2}\right), \quad (1)$$

where  $h$  is diffusion coefficient (bandwidth parameter). The diffusion set can vary with respect to the choice of the step size. In our case, we follow Chen *et al.* (2015) and

define  $U = \{0, 0.01, 0.02, \dots, 1\}$ . Huang (2005) proposed the following procedure to determine the value of the diffusion coefficient:

$$h = \begin{cases} 0.8146(b - a), & 1 < m \leq 5, \\ 0.5690(b - a), & m = 6, \\ 0.4560(b - a), & m = 7, \\ 0.3860(b - a), & m = 8, \\ 0.3362(b - a), & m = 9, \\ 0.2986(b - a), & m = 10, \\ 2.6851(b - a)/(m - 1), & 11 \leq m. \end{cases} \quad (2)$$

where  $b = \max_i x_i$  and  $a = \min_i x_i$  are upper and lower bounds of the sample  $X$ , respectively.

Chen et al. (2015) argued that information entropy theory can provide theoretical basis for calculation of the diffusion coefficient  $h$ . They defined  $h$  as follows:

$$h = \begin{cases} (e^H / \sqrt{2\pi e})(b - a)/(m - 1), & m > 11 \\ (e^H / \sqrt{2\pi e})(b - a), & m \leq 11 \end{cases} \quad (3)$$

where  $H = \ln m$ ,  $b = \max_i x_i$  and  $a = \min_i x_i$  are upper and lower bounds of the sample  $X$ , respectively. Having determined the value of  $h$ , one can proceed with estimation of the points of the density function. First, Eq. 1 is employed to estimate the likelihoods of observing sample values for a given diffusion vector  $f_i(u_j)$ , for  $\forall i, j$ . Second, for each element of the sample, normalised likelihood is computed by considering all the elements of the diffusion vector:

$$\mu_{x_i}(u_j) = \frac{f_i(u_j)}{\sum_{j=1}^n f_i(u_j)}. \quad (4)$$

Value  $\mu_{x_i}(u_j)$ , thus, indicates the relative likelihood to observe value of  $x_i$  given the diffusion vector. Then, the likelihood of observing a certain value of the diffusion vector can be obtained by considering the normalised likelihoods associated with all the elements of the sample:

$$q(u_j) = \frac{\mu_{x_i}(u_j)}{\sum_{i=1}^m \mu_{x_i}(u_j)}. \quad (5)$$

Next, let us define the sum of the likelihoods associated with the elements of the diffusion vector as:

$$Q = \sum_{j=1}^n q(u_j). \quad (6)$$

The latter value allows for computation of the relative frequencies (probabilities) of observing particular values within the diffusion vector given the observed sample:

$$p(u_j) = \frac{q(u_j)}{Q}. \quad (7)$$

By looking at the probabilities rendered by Eq. 7, one can define an instance of the survival function, which indicates the probability to observe values exceeding  $u_j$ :

$$p(u > u_j) = \sum_{k=1}^j p(u_k). \tag{8}$$

The probabilities resulting from application of Eq. 8 can be used to measure the risk of different degrees of hazard. Specifically, by adjusting value  $j$  and thus moving along the elements of the diffusion vector, one can look at probabilities of observing different values (intervals within the universe of discourse) of the variable under consideration. In order to provide an overall measure of risk, one can consider the average hazard:

$$\bar{u} = \sum_{j=1}^n p(u_j)u_j. \tag{9}$$

### 1.2. Yield loss rate

In this research, we focus on yield loss risk. In order to obtain comparable measures of risk, we consider a dimensionless measure of yield loss, i.e. yield loss rate (Chen *et al.* 2015). The yield loss rate,  $r_t$ , is defined in terms of the observed yield,  $x_t$ , and expected yield,  $E(x_t)$ :

$$r_t = \frac{\max(E(x_t) - x_t, 0)}{x_t}. \tag{10}$$

### 1.3. Linear moving average and trend analysis

In order to facilitate the risk analysis, one needs to obtain the expected value of a certain variable (e.g., yield). Such techniques as ordinary least squares, autoregressive-moving average models, splines or kernel smoothing (Goodwin, Mahul 2004; Ye *et al.* 2015) can be applied in order to estimate the expected values. In this research, we rely on the linear moving average (LMA) approach, as defined by Zhang and Wang (2010). Indeed, LMA is employed in order to estimate of expected yields. In its essence, LMA relies on the two well-known techniques, viz. linear regression and moving average. Combination of the said technique allows one to obtain a non-linear trend.

LMA is similar to the moving average approach as it is carried out for sub-samples of a certain length (length of the sub-sample time series is referred to as step). To formally present the approach, let  $k$  stand for the step size. As a result, the original time series is partitioned into  $n - k + 1$  sub-samples, with  $n$  being the number of observations within the original sample.

Index  $i = 1, 2, \dots, n - k + 1$  is then introduced to keep track of the sub-samples. Further on,  $t$  stands for the time period within a certain sub-sample. The time index then serves in construction of the linear trend within each sub-sample:

$$y_i(t) = a_i + b_i t + e_i^t, \tag{11}$$

where  $y_i(t)$  represents the variable modelled (e.g. yield, price) and  $t$  indicates time periods which, indeed, differ across sub-samples:

$$t = \begin{cases} 1, 2, \dots, k, & i = 1, \\ 2, 3, \dots, k + 1, & i = 2, \\ \vdots & \vdots \\ n - k + 1, n - k + 2, \dots, n & i = n - k + 1. \end{cases} \quad (12)$$

As one can note from Eq. 12, different time periods enter different numbers of sub-samples. More specifically, observations at the two endpoints of a time series are represented in fewer sub-samples if opposed to those located in the middle of a time series. Indeed, the quantity of fitted values associated with each observation relies on both sample size and step size. Let us define  $q$  so that it represents the quantities of fitted values for each observation. Depending on the sample size, the values of  $q$  are as follows:

$$q = \begin{cases} 1, 2, \dots, \underbrace{k, \dots, k}_{n-2(k-1)}, \dots, 2, 1, & k \leq \frac{n}{2}, \\ 1, 2, \dots, \underbrace{n - k + 1, \dots, n - k + 1}_{2k-n}, \dots, 2, 1, & k > \frac{n}{2}. \end{cases} \quad (13)$$

The LMA renders the (overall) fitted values given the sub-sample fitted values. Recalling that  $q$  captures the number of fitted values for each time point, the trend is estimated as the average of  $q$  values:

$$\hat{y}(t) = \frac{1}{q} \sum_{j=1}^q \hat{y}_j(t), \quad (14)$$

where  $\hat{y}(t)$  is the trend for the  $t$ -th time period and  $\hat{y}_j(t)$  is the sub-sample fitted value from Eq. 11. In our case,  $\hat{y}(t)$  serves as the expected yield in Eq. 10.

## 2. Results

In general, there have been increasing trends in crop yields observed in Lithuania during 2000–2015. This can be attributed to several intertwined factors. First, increasing rates of application of fertilizers and other agrochemicals contributed to a persistent increase in crop yields. Second, climate change has resulted in higher mean annual temperatures (Povilaitis *et al.* 2013), which has also contributed to increasing yields. Third, accession to the EU meant implementation of the CAP in Lithuania along with deeper integration into the global markets. These have resulted in improved farming practices (including increased rates of application of agrochemicals) and expansion of areas sown in general. While the improved farming practices are likely to reduce the yield risk, the effect of expansion in areas sown is rather arbitrary. Therefore, we will further quantify the yield risk by means of non-parametric modelling in order to identify its variation across crops and regions.

Table 1 presents the dynamics in yields of different crops in Lithuania during 2000–2015. As one can note, all the crops followed upwards trends in yields. However, the rates of growth varied across the crops significantly. The lowest rate of growth is observed for potatoes (3.3%). The two crops cultivated in less fertile regions, namely

winter rye and buckwheat, exhibited relatively low rates of growth of 19.2% and 11.6%, respectively. Another group of crops comprised winter triticale, spring triticale, oats, mixed cereals, and spring rape. The rates of growth ranged in between 32% and 46.6% for the latter group of crops. Finally, the third group of crops encompasses winter wheat, winter barley, spring wheat, spring barley, maize, legumes, and winter rape. The rates of growth in yields fluctuated around 60% for most of these crops, whereas that for winter barley stood at 106.4%. Coefficient of variation (CV) enables to compare the crops in terms of temporal fluctuations in yields. Most of the crops showed the CVs ranging in between 0.18 and 0.23. The exceptions include winter rye (0.15), buckwheat (0.25), legumes (0.27), and maize (0.4). Therefore, in spite of the generally positive change in the yields, certain crops featured relatively higher volatility in the yields during the period of 2000–2015. This calls for further analysis of crop-specific trends in yield variation.

As it was described in Section 1.3, the analysis relies on application of the LMA. The step size is a crucial parameter for the analysis. Different step sizes were tested in order to ensure both the smoothness and flexibility of the trend. As a result, the value of 6 years was chosen for the LMA. Figure 1 presents the case of mixed cereal yield in Vilnius county. Clearly, the LMA trend allows for flexibility in the direction of change, as represented by a kink at year 2006. Suchlike modelling is particularly relevant to Lithuania, where less fertile areas experienced declining yields prior to accession to the EU as agricultural production there had not followed the modern practices due to insufficient investments. The LMA was applied for each crop and county in order to obtain the expected yields.

Table 1. Yield change for selected crops in Lithuania, 2000–2015

Crop	Yield, t/ha					Growth, %	Std. Dev.	Mean	CV
	2000	2004	2008	2012	2015				
Winter wheat	3.56	4.12	4.76	5.17	5.71	60.5	0.83	4.0	0.21
Winter triticale	2.77	3.18	3.27	3.80	4.07	46.6	0.56	2.9	0.19
Winter rye	2.34	2.54	2.76	2.80	2.79	19.2	0.35	2.3	0.15
Winter barley	2.13	3.78	3.94	4.42	4.39	106.4	0.75	3.3	0.23
Spring wheat	2.62	3.45	3.01	3.87	4.20	60.6	0.59	3.3	0.18
Spring barley	2.43	2.93	2.88	3.36	4.00	64.3	0.56	2.9	0.19
Spring triticale	2.32	2.72	2.33	2.90	3.06	32.0	0.46	2.4	0.19
Oats	1.87	2.23	2.07	2.27	2.54	35.8	0.37	2.0	0.19
Buckwheat	0.89	0.59	0.76	0.83	0.99	11.6	0.19	0.8	0.25
Mixed cereals	1.79	2.08	1.91	2.16	2.43	35.6	0.37	1.9	0.19
Maize	2.85	2.20	4.25	5.65	4.71	65.2	1.75	4.4	0.40
Legumes	1.84	1.88	1.70	1.83	2.89	57.2	0.48	1.8	0.27
Winter rape	2.22	2.70	2.72	3.38	3.51	58.0	0.56	2.4	0.23
Spring rape	1.38	1.82	1.58	1.99	1.91	38.3	0.29	1.6	0.18
Potatoes	16.39	12.88	14.79	16.96	16.93	3.3	2.67	14.0	0.19

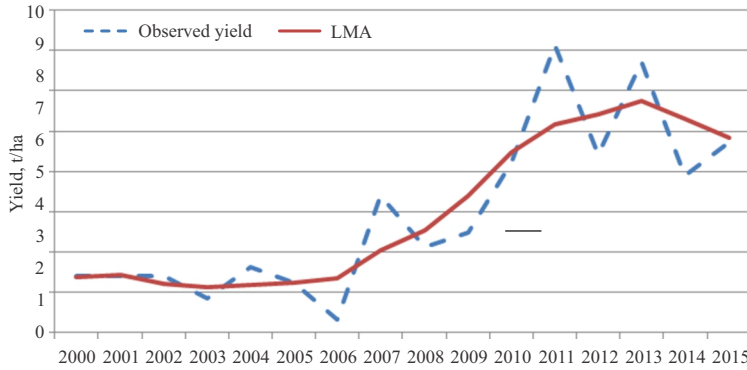


Fig. 1. Dynamics in mixed cereal yield in Vilnius county and the corresponding LMA trend, 2000–2015

The expected values of yields serve as means to compute the yield loss ratios (Eq. 10). These variables can describe the degree of loss experienced during a certain time period. CVs for the average loss rates resemble the spatial differences in this variable. Table 2 summarizes the yield loss ratios for each crop and county.

Table 2. Mean yield loss rates for different counties, 2000–2015

Crop											Average	SD	CV
	Alytus	Kaunas	Klaipėda	Marijampolė	Panevėžys	Šiauliai	Tauragė	Telšiai	Utena	Vilnius			
Winter wheat	0.06	0.07	0.11	0.06	0.10	0.07	0.06	0.08	0.09	0.08	0.08	0.02	0.24
Winter triticale	0.07	0.06	0.11	0.07	0.11	0.08	0.08	0.07	0.12	0.09	0.09	0.02	0.25
Winter rye	0.08	0.06	0.09	0.05	0.07	0.08	0.07	0.03	0.07	0.06	0.07	0.02	0.25
Winter barley	0.18	0.10	0.14	0.10	0.15	0.06	0.10	0.11	0.22	0.11	0.13	0.05	0.36
Spring wheat	0.07	0.05	0.07	0.04	0.06	0.06	0.04	0.06	0.06	0.05	0.06	0.01	0.18
Spring barley	0.09	0.05	0.05	0.04	0.06	0.05	0.05	0.03	0.08	0.07	0.06	0.02	0.33
Spring triticale	0.11	0.06	0.08	0.07	0.07	0.09	0.07	0.06	0.11	0.08	0.08	0.02	0.22
Oats	0.08	0.07	0.07	0.09	0.09	0.08	0.08	0.08	0.09	0.10	0.08	0.01	0.10
Buckwheat	0.13	0.13	0.19	0.13	0.13	0.30	0.13	0.25	0.14	0.11	0.16	0.06	0.38
Mixed cereals	0.11	0.10	0.12	0.06	0.11	0.09	0.08	0.06	0.09	0.12	0.09	0.02	0.23
Maize	0.09	0.09	0.43	0.06	0.20	0.17	0.13	0.20	1.28	0.31	0.30	0.36	1.23
Legumes	0.13	0.07	0.10	0.08	0.11	0.08	0.04	0.09	0.11	0.12	0.09	0.03	0.28
Winter rape	0.15	0.10	0.18	0.12	0.15	0.07	0.14	0.10	0.23	0.12	0.14	0.04	0.33
Spring rape	0.08	0.06	0.10	0.04	0.11	0.07	0.05	0.08	0.14	0.09	0.08	0.03	0.35
Potatoes	0.10	0.07	0.11	0.08	0.06	0.12	0.10	0.08	0.08	0.06	0.09	0.02	0.24



The data in Table 2 suggest that the mean yield loss rates varied in between 0.06 for spring wheat and spring barley and 0.3 for maize. This indicates that the shocks in crop yields (if compared to the expected values) vary across the crops. Looking at the CVs, the highest regional variation is observed for maize (CV of 1.23). Oats exhibited the lowest CV of 0.1, which indicates relatively low differences in yield loss rates across the counties. Maize showed the highest value of the CV. Table 1 has shown that the trend in yields varied across the counties and crops. The analysis of Table 2 once again confirmed that the differences in yield loss rates also prevail along regional and temporal dimensions. Thus, it is important to quantify yield risks in order to develop proper policies and business decisions.

The information diffusion model was then applied on the yield loss rates in order to non-parametrically estimate the underlying probability densities. Indeed, Eq. 7 can be employed to compute the probabilities associated with each element of the diffusion vector. In our case, the elements of the diffusion vector correspond to the values of the yield loss rate. Eqs. 2 and 3 present the two options for determination of the bandwidth parameter  $h$ . We follow the two approaches for calculation of the bandwidth parameter in order to ensure robustness of the results. Therefore, the two models are considered: Model 1 employs Eq. 2, whereas Model 2 employs Eq. 3. Thereafter, the measures of yield risk are based on the two aforementioned approaches.

The differences between the two approaches for estimation of the bandwidth parameter can be illustrated by considering the resulting densities. Figure 2, thus, depicts the two densities for yield loss ratio for potatoes in Vilnius county. The densities are based on Eq. 7. As one can note, the shape of the probability densities are rather similar, yet Model 1 shows somewhat increased probabilities to observe values of 0.45–0.55. These differences can also induce changes in the measures of risk.

The probabilities associated with different degrees of yield loss can be used to compute the expected yield loss rate (Eq. 9). Figure 3 presents the mean values of the expected yield loss for each crop. Again, the two models have been applied to estimate these values. As one can note, there exist differences between the levels of the expected yield loss rates rendered by the two models. Anyway, these differences do not render decisive changes in ranking of the average values.

Irrespective of the model applied, the lowest expected yield loss rate is observed for spring barley. Indeed, the expected yield loss rate amounted to 9.2% and 11.2%, depending on the model applied. However, the latter crop does not play an important role in the farming business in Lithuania. The second lowest expected yield loss rate is observed for spring wheat. Depending on the model used, the expected yield loss rate was 9.5% or 11.6%. The latter crop has gained more popularity in Lithuania as it can be seen from an increasing share of areas sown under this crop (Baležentis, Kriščiukaitienė 2016). Winter rye also showed low value of the expected yield loss rate under both approaches. The expected loss amounted to 10% and 11.8%, depending on the model used. However, the share of area sown under the latter crop has dropped significantly during 2000–2015. Spring triticale showed the fourth lowest level of yield risk, as represented by the mean expected yield loss rate (12.1% and 14.3%).

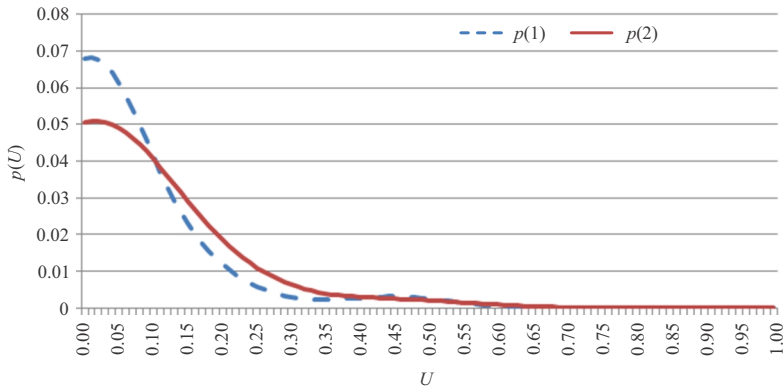


Fig. 2. Probability densities for yield loss rates for potatoes in Vilnius county

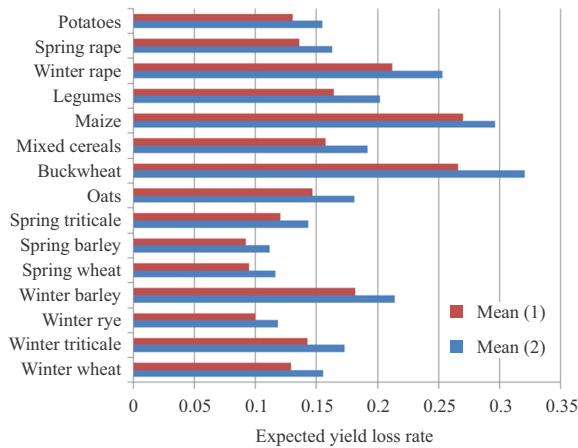


Fig. 3. Expected yield loss rates for different crops in Lithuania

The ranking of winter wheat and potatoes varied across the two approaches, yet the latter two crops still showed the 10<sup>th</sup> or 11<sup>th</sup> largest values of the expected yield loss rates. Specifically, the expected yield loss rate amounted to 12.9% and 15.6% for winter wheat, and 13.1% and 15.5% for potatoes. Indeed, winter wheat remains as the main crop cultivated in Lithuania and even increased its share in the total area sown during 2000–2015. As regards potatoes, area sown under this crop has shrunken significantly during the said period. Spring rape exhibited somewhat higher yield risk. The mean expected yield loss rate was 13.6% or 16.3% for Models 1 and 2, respectively.

Winter triticale featured the 8<sup>th</sup> highest level of yield risk among the crops analysed. Depending on the model used, the mean expected yield loss rate was 14.3% or 17.3%. Oats followed with mean expected yield loss rates of 14.7% and 18.1%. Mixed cereals showed higher yield risk under both models. Specifically, the mean expected yield loss rates were 15.8% and 19.2% depending on the model applied. Legumes followed with mean loss rates of 16.4% and 20.2%. Winter barley was specific with even higher mean

expected yield loss rates of 18.2% and 21.4%. Winter rape has seen an increasing share of the area sown during 2000–2015, yet this crop also exhibited rather high yield risk of 18.2% or 21.4%, depending on the model used.

As regards the results based on Model 1, buckwheat was the second most risky crop (mean expected yield loss rate of 26.6%), whereas maize appeared as the most risky crop (mean expected yield loss rate of 27%). The order is reversed if Model 2 is applied: maize and buckwheat show the mean expected yield loss rates of 29.6% and 32.1%, respectively.

As the previous analysis has suggested the presence of spatial variation in yields and yield loss rates, we compute CVs for the expected yield loss rates. This variable, thus, identifies the extent of spatial variation in yield risk. The results are presented in Figure 4. Note that the values are presented for the two models, as it was explained before.

As Figure 4 suggests, maize showed the highest regional variation in yields risk under both models applied. More specifically, the CV under Models 1 and 2 were 0.56 and 0.48, respectively. Spring rape and winter barley came next with CVs fluctuating around the values of 0.4. The crops with the highest regional differences in the yield risk deserve additional attention in terms of their re-allocation and identification of varieties adapted to the conditions of certain regions (e.g. different soil types). The lowest regional variation in yield risk is observed for oats. This finding is supported by both models used for the information diffusion approach. In order to further demonstrate the applicability of the information diffusion theory for analysis of yield risk and discuss the regional differences in yield risk, we will further discuss the county-level results for crops featuring the highest CVs.

Following Zhang and Wang (2010), we define the four classes of hazard associated with different degrees of yield loss. Low-level hazard is defined for cases where yield loss rate exceeds 5%, yet remains 15% at most. Medium-level hazard is defined for yield loss rates above 15% and less or equal to 25%. High-level hazard is defined for

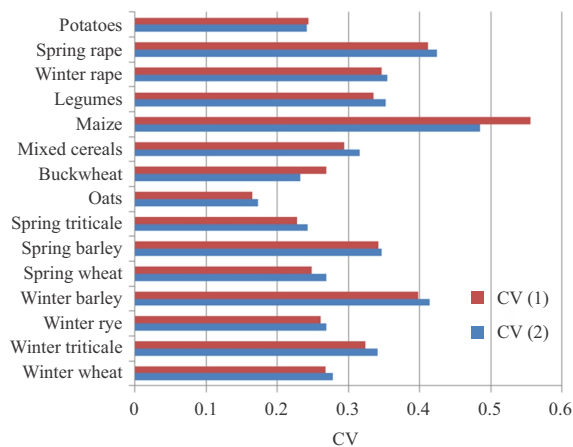


Fig. 4. Coefficients of variation (CVs) for expected yield loss across different crops

yield loss rates above 25% and less or equal to 35%. Finally, the catastrophic hazard is associated with yield loss rates exceeding 35%. The mean hazard can be calculated by assigning mean hazard levels to each level of hazard and summing up the products of probabilities and mean levels. We define the mean hazard in lines with Zhang and Wang (2010) by assigning mean hazard levels of 10%, 20%, 30%, and 40% to probabilities of low, medium, high, and catastrophic hazards, respectively. Basically, the measure of the mean hazards indicates the degree of yield risk without considering low levels of hazard (below 5%).

As maize showed the highest regional variation in yield risk, Table 3 presents the county-level data on different levels of hazards. Vilnius, Utena, and Klaipėda counties showed the highest level of yield risk. Specifically, the expected yield loss rates ranged in between 0.45 and 0.52 according to Model 1 and in between 0.48 and 0.51 according to Model 2. The mean hazard indicator ignores extremely low hazard and thus shows lower values if opposed to the expected yield loss rate. For instance, the highest expected yield loss rate value under Model 1 is 0.46 (the one observed for Klaipėda county), whereas the corresponding mean hazard indicator is 0.3. Looking at Model 2, the corresponding values for Utena county are 0.51 and 0.32.

As one can note, the non-parametric diffusion approach is quite sensitive to extreme values of yield loss. Counties specific with the highest yield risk show rather high prob-

Table 3. County-level measures of yield risk for maize

Hazard	Alytus	Kaunas	Klaipėda	Marijampolė	Panevėžys	Šiauliai	Tauragė	Telšiai	Utena	Vilnius
Model 1										
Low	0.37	0.38	0.12	0.41	0.25	0.30	0.37	0.30	0.09	0.12
Medium	0.12	0.16	0.12	0.13	0.20	0.17	0.23	0.15	0.09	0.12
High	0.04	0.06	0.11	0.03	0.13	0.09	0.08	0.08	0.10	0.12
Catastrophic	0.12	0.10	0.58	0.06	0.26	0.22	0.08	0.23	0.66	0.57
Mean	0.12	0.13	0.30	0.10	0.21	0.18	0.14	0.18	0.32	0.30
Expected loss	0.14	0.15	0.46	0.11	0.28	0.22	0.16	0.22	0.52	0.45
Model 2										
Low	0.35	0.33	0.10	0.38	0.20	0.27	0.31	0.27	0.10	0.11
Medium	0.18	0.21	0.11	0.20	0.17	0.19	0.24	0.19	0.10	0.11
High	0.08	0.10	0.10	0.07	0.14	0.12	0.13	0.12	0.10	0.11
Catastrophic	0.13	0.13	0.62	0.07	0.37	0.25	0.13	0.25	0.65	0.61
Mean	0.15	0.16	0.31	0.13	0.24	0.20	0.17	0.20	0.32	0.31
Expected loss	0.17	0.18	0.49	0.14	0.33	0.24	0.19	0.25	0.51	0.48

abilities of the catastrophic risk. This is caused by a flexible shape of the estimated probability distribution. Indeed, this phenomenon is also related to low quantities of maize output and areas sown under maize there. The carried out analysis can be reiterated for different crops thereby identifying the most risky regions.

In order to check the robustness of the results, we contrast the non-parametric approach to the parametric one. Specifically, Baležentis and Kriščiukaitienė (2016) applied the normal and logistic distributions to estimate the measures of risk for different crops. However, the two approaches cannot be compared directly as different yield loss ratios were defined in the aforementioned study and this one. We therefore use ranks rather than exact values of indicators reflecting the degree of expected loss. Specifically, the average relative risk premia are used for the parametric approach, whereas the expected yield loss rates are considered for the non-parametric approach. The averages for the country are used for the analysis (Table 4).

Table 4. Rankings of the crops based on the average measures of risk according to parametric and non-parametric approaches

Crop	Normal distribution	Logistic distribution	Non-parametric Model 1	Non-parametric Model 2
Winter wheat	12	13	11	10
Winter triticale	8	9	8	8
Winter rye	13	10	13	13
Winter barley	4	3	4	4
Spring wheat	14	14	14	14
Spring barley	15	14	15	15
Spring triticale	10	5	12	12
Oats	10	12	7	7
Buckwheat	2	2	2	1
Mixed cereals	6	8	6	6
Maize	1	1	1	2
Legumes	5	5	5	5
Winter rape	3	4	3	3
Spring rape	8	10	9	9
Potatoes	7	7	10	11
Rank Correlation				
Normal distribution	1			
Logistic distribution	0.91	1		
Non-parametric Model 1	0.96	0.81	1	
Non-parametric Model 2	0.94	0.78	0.99	1

**Note:** the results for the normal and logistic distributions (average relative risk premia) are based on Baležentis and Kriščiukaitienė (2016); expected yield loss rates are used for the non-parametric approach.

The six most risky crops coincide across the non-parametric and parametric approaches. Further on, the rank correlation indicates that the results are highly consistent across the two approaches. The lowest values of rank correlation are observed among the results for the logistic distribution and the two non-parametric models (0.78 and 0.81). Therefore, the results obtained can be considered as rather robust ones.

## **Conclusions**

The carried out analysis indicates that crops show different levels of yield risk in Lithuania. What is more, there exist differences in spatial variation of the yield risk for different crops. Specifically, winter rape shows the highest yield risk irrespectively of the region considered. Therefore, these crops require improved varieties and farming practices to be applied in order to ensure stability of income flows.

As regards regional variation of the yield risk, maize and spring rape and winter barley show the highest discrepancies. Therefore, the cropping patterns of these crops can be further improved by considering movement towards different regions and introducing varieties that are more suitable for certain regions.

The indicators of risk can be applied to support decision making facilitated by both farmers and public bodies. Farmers can be informed of the yield risk in their region in order to adapt their crop-mix. Indeed, suchlike analysis can also be carried out at lower level of aggregation if data are available. As regards government institutions, the scope of the support schemes can be adjusted in regards to the spatial differences in yield risk.

There have been certain limitations pertinent to the present study. First, the analysis was focused on yield risk. Indeed, yield risk can be offset by price fluctuations. However, this depends on crops and situation in international markets. Second, a single country was considered in this research. Therefore, further research could seek for a number of extensions. The analysis can be supplemented by applying different techniques for estimation of yield risks. Furthermore, yield and price risks can be considered simultaneously in order to reflect the income risk. Finally, more regions could be included in the analysis.

The carried out analysis contributes to the literature by unveiling the patterns of yield risk in Lithuania, which has faced serious transformations in terms of economic transition, institutional shifts, and climate change. These results can be compared to other regions in Central and Eastern Europe in order to devise the tailored policy measures aimed at ensuring viability of the farming business. The evidence-based policy measures can ensure effective use of public funds and increase the viability of rural areas in general. These results are important in regards to adjusting national and international food security policies with regards to risk mitigation. The management of risk is even more important in the presence of the climate change.

## **Acknowledgement**

Zhiqian Yu gratefully acknowledges financial support from the Youth Talents Project of Guangdong Province (2015WQNCX122) and the Youth Academic Team Project at Guangzhou University (201602XSTD).

## References

- Baležentis, T.; Kriščiukaitienė, I. 2016. *Production and price risk in Lithuanian crop farming*. Lithuanian Institute of Agrarian Economics, Vilnius.
- Bokusheva, R. 2011. Measuring dependence in joint distributions of yield and weather variables, *Agricultural Finance Review* 71(1): 120–141. <https://doi.org/10.1108/00021461111128192>
- Bokusheva, R.; Kogan, F.; Vitkovskaya, I.; Conratt, S.; Batyrbayeva, M. 2016. Satellite-based vegetation health indices as a criteria for insuring against drought-related yield losses, *Agricultural and Forest Meteorology* 220: 200–206. <https://doi.org/10.1016/j.agrformet.2015.12.066>
- Breustedt, G.; Bokusheva, R.; Heidelbach, O. 2008. Evaluating the potential of index insurance schemes to reduce crop yield risk in an arid region, *Journal of Agricultural Economics* 59(2): 312–328. <https://doi.org/10.1111/j.1477-9552.2007.00152.x>
- Chavas, J. P.; Shi, G.; Lauer, J. 2014. The effects of GM technology on maize yield, *Crop Science* 54(4): 1331–1335. <https://doi.org/10.2135/cropsci2013.10.0709>
- Chen, Q.; Zhang, J.; Zhang, L. 2015. Risk assessment, partition and economic loss estimation of rice production in China, *Sustainability* 7(1): 563–583. <https://doi.org/10.3390/su7010563>
- Chhatre, A.; Devalkar, S.; Seshadri, S. 2016. Crop diversification and risk management in Indian agriculture, *Decision* 43(2): 167–179. <https://doi.org/10.1007/s40622-016-0129-1>
- European Commission. 2017. *FADN public database* [online], [cited 5 December 2016]. Available from Internet: [http://ec.europa.eu/agriculture/rica/database/database\\_en.cfm](http://ec.europa.eu/agriculture/rica/database/database_en.cfm)
- FAO. 2017. *FAOSTAT, land* [online], [cited 5 December 2016]. Available from Internet: <http://www.fao.org/faostat/en/#data/RL>
- Girdziute, L.; Slavickiene, A.; Vaitkevicius, S. 2014. Integrated risk assessment at Lithuanian farms, *Transformations in Business & Economics* 13(1): 143–154.
- Goodwin, B. K.; Mahul, O. 2004. *Risk modeling concepts relating to the design and rating of agricultural insurance contracts*. World Bank Policy Research Working Paper, 3392.
- Hao, L.; Yang, L. Z.; Gao, J. M. 2014. The application of information diffusion technique in probabilistic analysis to grassland biological disasters risk, *Ecological Modelling* 272: 264–270. <https://doi.org/10.1016/j.ecolmodel.2013.10.014>
- Hao, L.; Zhang, X.; Liu, S. 2012. Risk assessment to China's agricultural drought disaster in county unit, *Natural Hazards* 61(2): 785–801. <https://doi.org/10.1007/s11069-011-0066-4>
- Huang, C. 1997. Principle of information diffusion, *Fuzzy Sets and Systems* 91(1): 69–90.
- Huang, C. F. 2005. *Risk assessment of natural disaster – theory & practice*. Beijing: Science Press.
- IPCC. 2014. *Climate Change 2014*. A synthesis report. Contribution of working groups I, II and III to the fifth assessment report of the intergovernmental panel on climate change. Geneva: IPCC.
- Ker, A. P.; Tolhurst, T. N.; Liu, Y. 2016. Bayesian estimation of possibly similar yield densities: implications for rating crop insurance contracts, *American Journal of Agricultural Economics* 98(2): 360–382. <https://doi.org/10.1093/ajae/aav065>
- Kuziak, K. 2016. Assessing systemic risk with Beta approach, *Transformations in Business & Economics* 15(2A): 305–315.
- Liu, B.; Siu, Y. L.; Mitchell, G.; Xu, W. 2013. Exceedance probability of multiple natural hazards: risk assessment in China's Yangtze River Delta, *Natural Hazards* 69: 2039–2055. <https://doi.org/10.1007/s11069-013-0794-8>
- Peleckis, K.; Peleckienė, V.; Peleckis, K. 2015. Crop yield insurance—negotiating between government, farmers and insurance companies, *Economic Science for Rural Development* 37: 15–25.

Piontek, K. 2016. Measuring and testing skewness of financial return distributions, *Transformations in Business & Economics* 15(2A): 316–328.

Povilaitis, V.; Lazauskas, S.; Feizienė, D.; Kukujevas, A.; Feiza, V. 2013. Maize productivity as influenced by different nitrogen levels and climate change, *Journal of Food, Agriculture & Environment* 11(2): 803–807.

Shi, G.; Chavas, J. P.; Lauer, J. 2013. Commercialized transgenic traits, maize productivity and yield risk, *Nature Biotechnology* 31(2): 111–114. <https://doi.org/10.1038/nbt.2496>

Shi, H.; Jiang, Z. 2016. The efficiency of composite weather index insurance in hedging rice yield risk: evidence from China, *Agricultural Economics* 47(3): 319–328. <https://doi.org/10.1111/agec.12232>

Statistics Lithuania. 2016. *Official statistics portal* [online], [cited 5 December 2016]. Available from Internet: <http://osp.stat.gov.lt/en/home>

Xie, Z.; Xu, J.; Deng, Y. 2016. Risk analysis and evaluation of agricultural drought disaster in the major grain-producing areas, China, *Geomatics, Natural Hazards and Risk* 7(5): 1691–1706. <https://doi.org/10.1080/19475705.2015.1091387>

Ye, T.; Nie, J.; Wang, J.; Shi, P.; Wang, Z. 2015. Performance of detrending models of crop yield risk assessment: evaluation on real and hypothetical yield data, *Stochastic Environmental Research and Risk Assessment* 29(1): 109–117. <https://doi.org/10.1007/s00477-014-0871-x>

Yuan, X. C.; Tang, B. J.; Wei, Y. M.; Liang, X. J.; Yu, H.; Jin, J. L. 2015. China's regional drought risk under climate change: a two-stage process assessment approach, *Natural Hazards* 76(1): 667–684. <https://doi.org/10.1007/s11069-014-1514-8>

Zhang, Q.; Wang, K. 2010. Evaluating production risks for wheat producers in Beijing, *China Agricultural Economic Review* 2(2): 200–211.

**Tomas BALEŽENTIS** is a Principal Researcher at Lithuanian Institute of Agrarian Economics and holds PhD from Vilnius University, Lithuania, and University of Copenhagen, Denmark. Dr Baležentis has published over 100 peer-reviewed papers on multi-criteria decision making, efficiency analysis, and agricultural economics in such outlets as *Agricultural Economics*, *European Journal of Operational Research*, *Applied Energy*, and *Journal of Productivity Analysis*.

**Dalia ŠTREIMIKIENĖ** is a Principal Researcher at Lithuanian Institute of Agrarian Economics. She has been included in Rooster of experts in United Nations Framework Convention on Climate Change (UNFCCC) as a world class expert on greenhouse gas accounting, forecasting and mitigation measures on greenhouse gas emissions and assessment. D. Streimikiene has an extensive experience in international and Lithuanian research projects.

**Rasa MELNIKIENĖ** is the Director of Lithuanian Institute of Agrarian Economics. She has graduated as a PhD from Vilnius Gediminas Technical University. She had served as a member of the Seimas of the Republic of Lithuania. Her research areas include rural development and agricultural economics.

**Zhiqian YU** is an Assistant Professor of Economics at School of Economics and Statistics, Guangzhou University. She holds a PhD degree in Economics from Jinan University (2014), and had studied as a visiting PhD student at the Copenhagen Business School. Her research fields include efficiency and productivity analysis, banking management, and public education.