TEM Journal. Volume 10, Issue 2, Pages 627-638, ISSN 2217-8309, DOI: 10.18421/TEM102-18, May 2021.

# Multi-Hazard Risk Assessment Policies in the Agrarian Sector using Business Continuity Data

Athanasios Podaras [1], Dana Nejedlová [1], Pavla Vrabcová [2]

[1] Department of Informatics, Faculty of Economics, Technical University of Liberec, Voroněžská 13, 460 01 Liberec 1, Czech Republic
[2] Department of Forestry Technologies and Construction, Faculty of Forestry and Wood Sciences, Czech University od Life Sciences Prague, Kamýcká 129, 165 00 Prague – Suchdol, Czech Republic

***Abstract* – Inspired by the negative consequences of the COVID-19 pandemic on agriculture, the primary objective of the current paper is the utilization of real business continuity template data to, firstly, construct a business intelligence multidimensional and multi-hazard risk assessment data model and, secondly, conduct an aggregate pivot descriptive analysis regarding the influence of the pandemics and other hazards on selected agrarian industries. The second part includes an additional predictive regression analysis model regarding the influence of multiple hazards on the recovery time of interrupted due to these hazards key agricultural operations. The pilot multidimensional model can inspire agribusiness practitioners to assess efficiently the risks posed by multiple hazards to key agricultural activities. Data from specific agrarian sectors have been collected for the present study.**

***Keywords* – agricultural operations, COVID-19, hazards, multidimensional data, pivot analysis, regression, risk assessment.**

DOI: 10.18421/TEM102-18
https://doi.org/10.18421/TEM102-18

**Corresponding author:** Athanasios Podaras,
*Department of Informatics, Faculty of Economics, Technical University of Liberec, Voroněžská 13, 460 01 Liberec 1, Czech Republic.*
**Email:** athanasios.podaras@tul.cz

*Received: 11 April 2021.*
*Revised: 02 May 2021.*
*Accepted: 12 May 2021.*
*Published: 27 May 2021.*



## 1. Introduction

In today's turbulent business environment, key business operations are highly exposed to a diversity of hazards. The agribusiness industries are highly influenced by a multitude of hazard types, which are distinguished as natural, biological man-made [1] and technical [2]. As reported by Lorencová et al. [3], the interactions between agriculture, the environment, and society are very complicated and pluralistic. Agricultural and forestry enterprises not only in the Czech Republic face a number of problems, namely, adverse weather (extreme fluctuations in drought, heat, etc.), climate change [3], [4], the reduction of subsidies and their differences [5], declining acreage of agricultural land [6], aging workforce [7], rising seed costs, fertilizers, plant protection products, fuels, labor or rental of agricultural land.

The present COVID-19 global outbreak has resulted in significant financial losses for the global economy. Crucial business and industrial sectors have been highly affected. The agricultural sector and its sub-sectors or industries have also experienced the negative consequences of the pandemic. According to a recent statement of the Food and Agriculture Organization of the United Nations [8], "*the pandemic is impacting global food systems, disrupting regional agricultural value chains, and posing risks to household food security*". Additionally, the pandemic has significantly influenced the conducting of "*basic and applied research in agricultural sciences*" [9]. Bioeconomy, which is strongly bound to agrarian resilience and sustainability stemming from the effective use of information and communication technologies [10], has also experienced the negative effect of the COVID-19.

After the COVID-19 outbreak, the reconsideration of the primary sustainable development goals (SDGs) as they had been initially defined by the United

Nations has been proved to be inevitable, since "*the COVID-19 has brought to the fore the fact that the SDGs as currently designed are not resilient to shocks imposed by pandemics*" [11]. Jámbor et al. [12] summarize the impacts of COVID-19 in the following areas: limited food supplies, panic shopping on the demand side and stockpiling, labor shortages, food security and food security, protectionist movements, supply problems, increased e-shopping of agricultural products and more.

It is thus crucial to rethink current and address novel BCM policies [13], which can be considered as proactive tools to minimize the negative impact of diverse hazard types and, due to the vast global negative effect of epidemics on agriculture, a closer analysis regarding the impact of the biological hazards category on the agri-business sector and its industries is considered to be a task of paramount importance.

The adoption of standard proactive business continuity [14] management policies can serve as a highly effective tool towards these existing hazard-related challenges. A recent review study regarding the negative consequences of diverse hazard types on the agricultural sector, highlight the strong positive influence of the immediate action to the continuity of agribusiness, stating that "*emergency response that enables continuity of business for agriculture sectors effectively and rapidly ensures that disasters are handled in both an economically and environmentally sustainable way*" [15]. Continuity of business, which results in a resilient to the multitude of hazards and sustainable agribusiness, relies among other things, on the adoption of modern information and communication technologies as well as big data and multidimensional database solutions such as business intelligence systems. Information processing in the agrarian sector facilitates effective decision-making and highly boosts agricultural productivity [16].

The primary goal of the current agri-business continuity study is two-fold, its outputs stem from the incorporation of available dairy [34], and cattle feed yard [35] business continuity template data. The first fold is focused on the creation of a hazard risk assessment multidimensional data model through the observation and conceptualization of the involved dimensions, facts, hierarchies, and data granularity regarding the hazards and the affected key business functions in both agricultural industries. Moreover, the second fold deals with critical online analytical processing and regression analysis tasks regarding the agribusiness hazard risk mitigation and their influence on the recovery time required for a set of affected key business functions. The proposed data pattern and the examples explained via the utilization of spreadsheet data analysis in order to demonstrate its practical value are expected to serve as pilot knowledge-based agri-business continuity management tools for more effective recovery policies starting from the currently investigated industries.

The rest of the paper is structured as follows: theoretical background focuses on the analysis of the literature regarding the importance of big data collection and processing to conduct crucial agri-business decisions and, in particular, decisions regarding the charting of multi-hazard risk assessment policies in agriculture, including a reference of the current state of business continuity practices in the agrarian sector. Chapter Material and methods is focused on the analysis of the necessary tools, methods and research framework, which has been considered to derive the present results. Chapter Results is devoted to the analysis and presentation of the research output. Section 4 is used as an explanatory discussion part with respect to the limitations as well as the possibilities to further expand the current research achievements. The paper is finalized with a summarized conclusion regarding the present accomplishments as well as the implications for future research.

## 2. Theoretical Background

Many researchers in the domain [17], [18], have underlined the big data processing potential in agricultural research and practice. Regarding the practical use of big data in agribusiness, Lytos et al. [16] have conducted a recent review regarding the importance of big data collection and processing within the agricultural domain. The research team highlights the need to rely on computerized systems and various data collection and processing tools to foster agricultural productivity. Other researchers consider the incorporation of modern big data technologies as the key factor to reduce the risk of various hazard types that threaten the agricultural sector [19]. The researchers delineate the architecture of an excellent technical big data collection and processing solution, which combines satellite and sensor data technologies to mitigate disaster risk from human and natural hazards in the agricultural sector, and support the possibility to transfer the technique in the crisis management domain.

Another important research regarding the importance of big data processing in fostering critical decisions in agriculture has been conducted by Kamilaris et al. [17]. The researchers indicate the need to "*employ the recent practice of big data analysis, in order to solve various relevant problems*" in agriculture. In the specific study, the need to employ big data tools to foster decision making related to critical agricultural issues such as, food protection and security, ensuring of animal food quality and safety, protection and insurance of small farmers and animal disease recognition has been thoroughly explained. Additionally, researchers also require the existence of big data volumes to generate new knowledge in agricultural sciences. For these reasons, several data repositories devoted to data availability and easy data access for researchers are listed in the study conducted by [9].

Kamilaris et al. [17] state that open and free to access agricultural big data volumes could "*create tremendous opportunities for research and development towards smarter and more sustainable farming*." Finally, the need to associate knowledge and information management with risk assessment has been underlined in the business field [20] in general and in the agricultural area [21] in particular when agricultural risk management policies are combined with business continuity data. Risk mitigation agricultural research results related to the COVID-19 have been also published. Sharma et al. [21] analyzed the impact of COVID-19 on the global agricultural supply chain using risk mitigation strategies that can be useful for the post-COVID era in a specific field.

The primary concern which has triggered the current investigation stems from a research conducted with respect to the application of business continuity management in the Czech agricultural enterprises, where it has been concluded that "*agrarian organizations are not interested in applying BCM according to standards*" [22].

Additionally, the recent developments regarding the COVID-19 outbreak has triggered the necessity to reconsider the importance of business continuity management in charting risk assessment policies within the agricultural sector. Due to the fact that a special BCM concern is to tackle pandemic issues, agricultural BCM template data can be a valuable tool for ameliorating business continuity and hazard risk assessment policies, since "*data collection is an important activity throughout the BCM development process*" [23].

A limitation regarding the research around the agricultural business continuity strategies is the lack of available agri-business continuity template data. Despite this limitation, research outputs regarding the application of business continuity policies in the agricultural domain can rely even on small data volumes due to the fact that for understanding the concept of big data in agriculture one should initially explore small data [24].

## 3. Materials and Methods

### *3.1. Business Intelligence and Multidimensional Data Processing Concepts*

The concept of multidimensional data models is used to design business intelligence data warehouse [25] schemas. Considering the ground theory related to the multidimensional data schema, the key terminology includes facts, dimensions, hierarchies and granularity. Based on the study of Caniupán et al. [26] it is that "*facts correspond to events which are usually associated with numeric values known as measures and are referenced using the dimension elements*", while "*dimensions are modelled as hierarchies of elements, where each element belongs to a category. The categories are also organized into a hierarchy called hierarchy schema*."

Additionally, Vaisman and Zimanyi [25] refer to various types of hierarchies depending on the relationship between parent and child attributes in each dimension. For example, strict hierarchies are described as dimensional hierarchy types in which one-to-many relationship is observed between a parent and child hierarchy levels. Non-strict hierarchy is a term used to delineate a many-to-many relationship between parent and child hierarchy levels. The aforementioned concepts are utilized to design the multidimensional schema. Moreover, online analytical processing tasks such as pivoting, drilling-down, rollup, slicing, and dicing are also considered as crucial information processing and multidimensional data analysis tools for the completion of the current research.

### *3.2. Research Steps and Expected Results*

The present contribution relies on a sequential dual-path research framework based on the collected data from two agri-business continuity templates regarding the dairy [34] and the cattle [35] sectors (Figure 1). In the first path, the data are used to design a multidimensional business continuity hazard risk analysis schema.

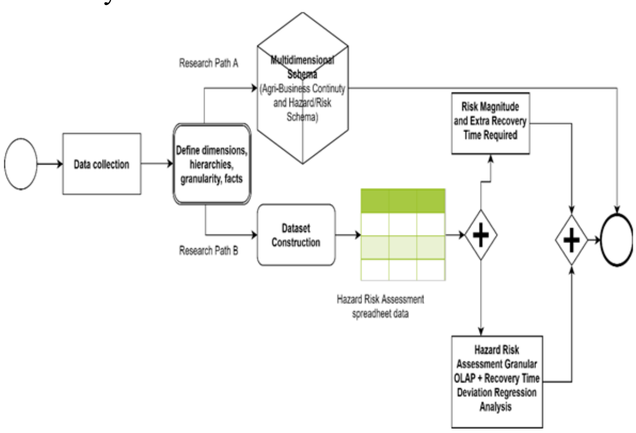


*Figure 1. The Delineated Summarized Research Framework*

The second path delineates the computation of the hazard risk assessment and the additional recovery time required for a set of business functions which are affected by specific hazard types. Moreover, regression analysis models are used to predict the % increase the aforementioned computed additional recovery time.

### *3.3. Dataset Preparation*

The constructed dataset is composed of the following variables:

Hazard Type: As stated in the introductory part of the current paper, the hazard types are classified as human, technological or technical, biological and natural. Thus, the defined hazard types are mapped to the specific hazards and each hazard can be triggered by several causes.

Risk Exposure Level: The specific column expresses the degree of vulnerability of one or more business functions to a specific hazard, hazard type or cause. For the needs of the current research, the following risk exposure quantification mapping has been determined:

- High Risk Exposure: Since a 0–1 scale has been selected for the risk exposure level, high risk exposure is defined as 0.8;
- Medium Risk Exposure: the medium value risk exposure is defined as 0.5;
- Low Risk Exposure: the low risk exposure level equals to 0.2.

Key Business Functions (*KBFs*) Impacted and Impact of a given hazard: the impact of a given hazard type on specific key business functions have been estimated According to the following ratio (Eq. 1):

$$I_H = n_{i_{KBFs}} / n_{KBFs}, \quad (1)$$

where:

$I_H$ = impact of the hazard to one or more *KBFs*,
$n_{i_{KBFs}}$= number of impacted *KBFs* while $n_{KBFs}$ is the total number of the *KBFs* in a given agribusiness industry.

According to the data collected, each hazard type, hazard, and cause influence 1 or more *KBFs*. For example, in the dairy farm sector, the pandemic cause which can result in loss of hourly labor has an impact on three out of 8 total key business functions. As a result, the impact of pandemic is estimated as 3/8 or $I_H$ = 0.375. Considering the high risk exposure level of the impacted functions with this hazard type/cause the Risk Magnitude (*RM*) value of the specific hazard can be computed as follows (Eq. 2) [13]:

$$RM = I_H \cdot P_H, \quad (2)$$

where:

$RM$ = risk magnitude of the given hazard type/cause,
$I_H$ = impact of the hazard on one or more $KBFs$,
$P_H$ = probability that the hazard will affect the specific $KBFs$.

In the above stated example, the risk magnitude is computed as follows:

$$RM = 0.85 \cdot 0.375 = 0.31875.$$

It can be realized that the highest $RM = 0.85 \cdot 1 = 0.85$, in the case that the total of the *KBFs* are influenced by and are highly exposed to a specific hazard type.

Another variable of paramount importance for the current research is the Recovery Time estimation for a set of *KBFs* that are impacted by the specific hazard type or cause. For every individual *KBF*, the recovery time is mapped to the level of importance or criticality of the function. For example, in the case of the Cattle Feed Yards industry, the *KBF* Cattle Movement and Holding is considered as Most Critical function, and as a consequence it must be rapidly resumed of a specific hazard triggering its interruption. In both the currently studied industries, three levels of criticality are assigned, namely, most critical, moderately critical, and least critical. Using the approach suggested by Supriadi and Sui-Pheng [13], the following recovery time values can be assigned to each level:

- Most Critical: Recovery Time < 24 hours;
- Moderately Critical: 24 hours ≤ Recovery Time < 72 hours;
- Least Critical: 72 hours ≤ Recovery Time ≤ 168 hours.

For a set of business functions affected by a hazard type or cause, the recovery time computation is defined as the following ratio (Eq. 3):

$$RT = \sum n_{KBFs_L} \cdot t_{max_L} / n_{KBFs} \quad (3)$$

where:

$n_{KBFs_L}$= number of *KBFs* that belong to a specific level $L$,
$t_{max_L}$= the maximum recovery time at the given *KBF* level,
$n_{KBFs}$ = total number of the *KBFs* impacted by a specific hazard type but in ideal conditions where the hazard has not yet affected these functions. Using the formula proposed in Podaras and Nejedlova [37], by combining *RT* and *RM* values, the additional recovery time ($RT_{AD}$) required due to hazards' emergence and the negative impact on specific functions, will be computed as follows (Eq. 4):

$$RT_{AD} = RT + (RT \cdot RM) \quad (4)$$

The utilized dataset includes 56 BCM observations from both agricultural industries. The main dataset variables are the Hazard, Cause, Type, *KBFs* Impacted, and Risk Exposure for a set of *KBFs* affected by a combination of {Hazard, Cause, and Type}. The variables *KBFs* Impacted and Risk Exposure are then used for estimating the Risk Magnitude, the Extended Recovery Time and the Time Deviation between the suggested recovery time in ideal conditions and the recovery time required when the crisis occur.

## 4. Results

### 4.1. Path A: Conceptualization of the Multidimensional Data Model

The utilized template data from two different agricultural sectors, namely, the dairy farms and the cattle feed yards have revealed the existence of three different dimensions that is, industry, hazard and risk exposure level. Each dimension is composed of different levels of hierarchy. The following dimensions have been distinguished for the construction of the risk/hazards multidimensional schema (Figure 2). To the best of our knowledge, a multi-hazard risk assessment multi-dimensional model has not been suggested in the currently presented form in the past.

Dimension 1-Industry: The industry dimension includes strict hierarchies [25] between the levels of industry, key business function and sub-functions. Each agribusiness industry is composed of multiple key business functions and each key business function is composed of several sub-functions. Thus, an one to many (1:N) relationship between the parent and the child level can be defined.

Dimension 2-Risk Exposure Level: The specific dimension includes three different risk exposure levels of each key business function with specific hazard types. However, the risk exposure levels are dimensional attributes which do not follow a hierarchical or granular structure. The levels of risk exposure are attributes which share a common level of hierarchy.

Dimension 3-Hazard: The hazard dimension follows a similar granular structure as the dimension industry. However, non-strict hierarchies [25] can be observed in this dimension between hazard causes and hazards due to the fact that the relationship between a parent hierarchy level and a child hierarchy level is many to many (N:M). Each hazard cause is related to one or more hazards, but each hazard is also related to many hazard causes. Consequently, non-strict hierarchies should be considered for the conceptualization of the final multidimensional model.

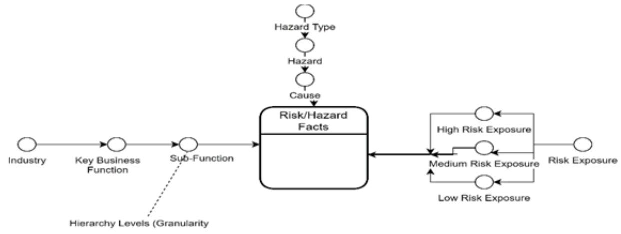


*Figure 2. The conceptualized multidimensional schema for agri-business continuity based on hazard risk assessment*

### 4.2. Path B: Pivot Analysis

#### 4.2.1. Aggregate Computations Regarding the Number of Key Business Functions Impacted

The first important agri-business continuity online analytical processing report is based on a spreadsheet pivot analysis regarding the average aggregate computation of the key business functions that are impacted by the four different hazard types in both industries. The OLAP report computations (Table 1) and the corresponding conducted graph (Figure 3) are below illustrated.

*Table 1. The estimated average number and percentage of the KBFs impacted by all the possible hazard types in both agricultural industries*

| Hazard Type | *CF* | *DF* |
|---|---|---|
| biological | 3.75 | 4.25 |
| man-made | 2.64 | 1.70 |
| natural | 3.85 | 4.00 |
| technological | 4.50 | 1.00 |
| AVERAGE *KBFs* per Industry | **3.26** | **2.25** |
| % Average Affected *KBFs* per Industry | **40.00%** | **22.50%** |

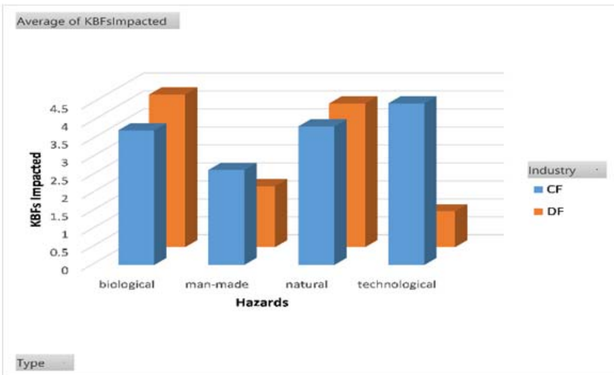


*Figure 3. The aggregate computations (average number) regarding the number of KBFs affected by different hazard types in the CF/DF Industries*

The observation of the Table and the corresponding chart reveals that the most influential hazard type in the Dairy Farms industry is the biological in which the pandemic cause is included. In the Cattle Feed Yards industry, the technological hazard affects mostly the key business functions. In both industries, the natural hazards type is considered the second most influential category.

TEM Journal. Volume 10, Issue 2, Pages 627-638, ISSN 2217-8309, DOI: 10.18421/TEM102-18, May 2021.

Drilling down the hazard dimension, more detailed reports can be generated according to which the granularity depth reaches the hazard level. In this way, similar results can be reported in more detail (Figure 4, Figure 5). Taking into consideration the drilling own potential of the OLAP operations, more detailed results have been conducted regarding the biological hazards type and the *pandemic* cause which is of paramount importance for agri-business and the negative consequences of COVID-19 crisis. The following report has been generated, (Table 2).

*Table 2. The estimated average number and percentage of the KBFs impacted by the pandemic cause in both agricultural industries*

| Type | Cause | Hazard | CF | DF |
|---|---|---|---|---|
| Biological | Pandemic | Loss of hourly employees | 3.00 | |
| | | Loss of management | 8.00 | |
| | | Loss of hourly labor | | 3.00 |
| | | Loss of management | | 3.00 |
| | | Loss of Vet services | | 1.00 |
| **Total** | | | **5.50** | **2.33** |
| **% average *KBFs* affected by the pandemic cause** | | | **68.75%** | **23.33%** |

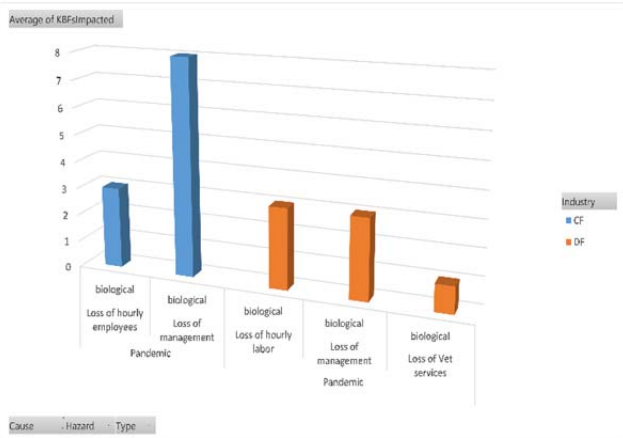


*Figure 4. The average number of KBFs impacted by the combination {Hazard Type, Cause, Hazard Effect}*

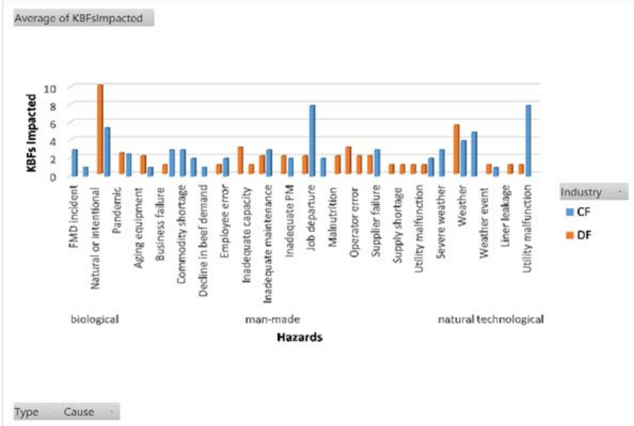


*Figure 5. The average number of KBFs impacted by the combination {Hazard Type, Cause}*

From the above stated OLAP reports, it has been concluded that the most influential hazard is based on the combination of {biological type, pandemic cause and loss of management hazard effect} affecting all (8) of the key business functions in the cattle feed yards industry. The same combination has also a strong impact on the dairy industry. Moreover, the average number of impacted key business functions in the CF industry is 5.5, while in the DF industry is 2.33 when referring to the impact of the combination {biological type, pandemic cause} on the key business functions.

In the dairy sector, the loss of hourly labor hazard due to the pandemic cause (hazard type = biological) triggers a 30% increase in the recovery time for three (3) out of ten affected key business functions and it's the most. Respectively, the hazard that triggers the highest recovery time increase ratio in the cattle feed yard industry is the loss of management with 20% increase. The specific hazard affects all the key business functions (8/8). The results can be visualized in a pivot analysis (Table 3) and the corresponding chart (Figure 6).

*Table 3. Estimated average increase in the recovery time due to the effect of {Type, Cause, Hazard Effcet} on a set of KBFs in both industries*

| Type | Cause | Hazard | CF | DF |
|---|---|---|---|---|
| Biological | Pandemic | Loss of hourly employees | 7.50 | |
| | | Loss of management | 20.00 | |
| | | Loss of hourly labor | | 30.00 |
| | | Loss of management | | 7.50 |
| | | Loss of Vet services | | 2.50 |
| **Total % average increase in the recovery time due to pandemic** | | | **13.75** | **13.33** |

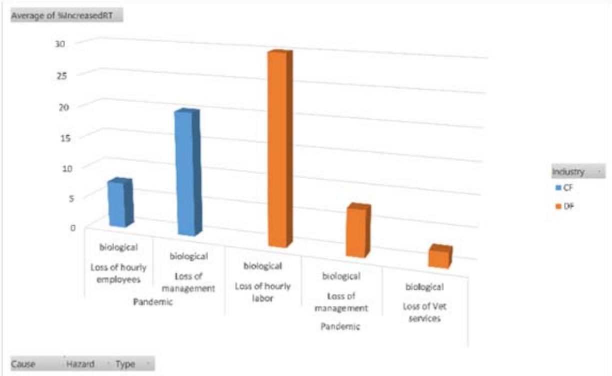


*Figure 6. Estimated average increase in the recovery time due to the effect of {Type, Cause, Hazard Effcet} on a set of KBFs for each agricultural industry*

In the dairy sector, the *loss of hourly labor* hazard due to *the pandemic* cause (Hazard Type= Biological) triggers a 30% increase in the recovery time for three (3) out of ten affected key business functions. Respectively, the hazard that triggers the highest recovery time increase ratio in the cattle feed yard industry is the *loss of management* with 20% increase. The specific hazard affects all key business functions (8/8).

#### 4.2.2. Recovery Priority Level Modification Due to Pandemic

An additional final observation regarding the effect of the combination {pandemic cause, biological hazard type} on specific *KBFs* groups is that in some cases the recovery priority level is modified. This effect can be visualized using additional pivot tables (Table 4).

*Table 4. Average increase in the recovery time for the KBF sets for which the recovery priority level is expected to be modified only in the dairy sector due to the influence of the combined factors {Type, Cause, Hazard Effect}*

| Type | Cause | Hazard | KBFs Impacted | DF |
|---|---|---|---|---|
| Biological | Pandemic | Loss of hourly labor | 3.00 | 30.00 |
| | | Loss of Vet services | 1.00 | 2.50 |
| **Total % Increase in the Recovery Time** | | | | **16.25** |

The pandemic effect on the dairy sector regarding the groups of *KBFs* for which the recovery priority level has been modified, resulted to loss of hourly labor causing 30% increase in the recovery time for a group of three *KBFs* and to loss of Vet services causing 2.5% increase in recovery time for one (1) *KBF*.

In the first case, the affected *KBFs* have been the Parlor Operations (PO – Most Critical), the Herd Management (HM-Moderately Critical) and the Calf Rearing (CF – Moderately Critical). Based on the equation (3) and the recommended by BCM experts [13] mapping between the recovery time and the recovery priority levels, the average recovery time for the given group is 56 hours. Thus, the proposed priority level is moderately critical *KBF* group. However, the influence of the pandemic resulted in the modification of the recovery priority level and 30% increase of the recovery time. The new recovery time based on (4) has been estimated as 72.8 hours (least critical group). The loss of Vet services caused by the pandemic resulted to a 2.5% increase in the recovery time for the function Animal Health/Hospital Management (AH – Moderately Critical). The initially proposed recovery time has been < 72 hours while the recovery time after the influence of the pandemic is estimated as 73.8 hours (Least Critical *KBF*). In the cattle feed yard industry no recovery priority modifications have been observed in case of either pandemic or any other cause of any hazard type. The only cases regarding recovery prioritization levels have been observed within the dairy sector and can be caused by biological and/or man-made hazard types. The average values regarding the % increased recovery time comparing pre-hazard and post-hazard expected recovery time, as well as the number of *KBFs* affected in the dairy sector, can be plotted using pivot analysis and the constructed spreadsheet dataset (Figure 7, Figure 8).

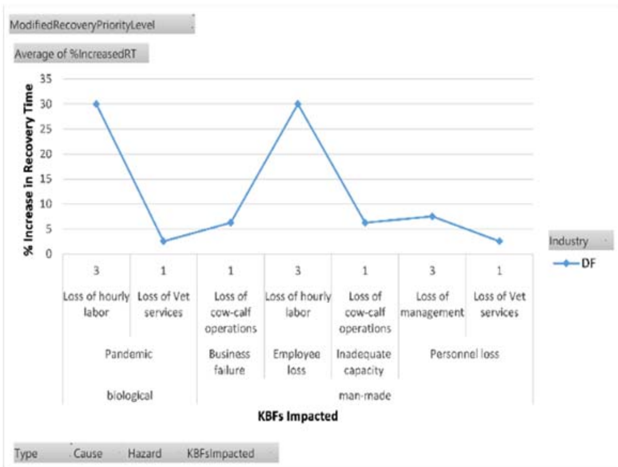


*Figure 7. Expected recovery priority level modification in the dairy sector*

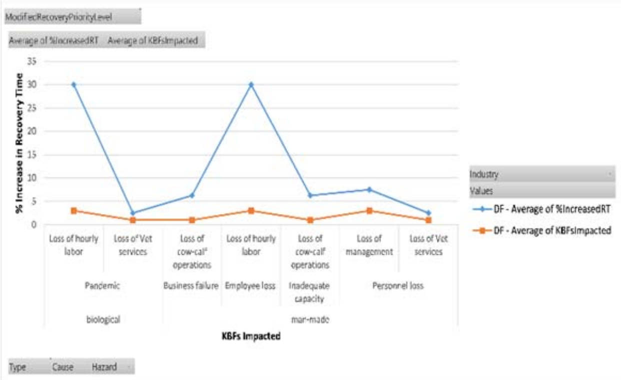


*Figure 8. Trend of the estimated average KBFs impacted and the estimated average % increase in the recovery time for KBFS with modified recovery priority level after the influence of the combined factors {hazard type, cause, hazard effect}*

## 5. Discussion and Limitations

The utilization of agri-business continuity template data, based on the above analyzed results, facilitates the construction of a conceptual multidimensional business intelligence data pattern and the OLAP pivot analysis regarding the agribusiness hazard risk assessment process. Therefore, descriptive hazard risk assessment data processing has been proved highly beneficial for the selected industries. However, the rule-based knowledge discovery process requires also the generation of rules which can activate the predictive potential of data mining [17]. Despite the fact that part of the constructed dataset can be utilized as input for machine learning regression analysis, which has been considered as highly robust for critical situations within the agribusiness sector [27], it is believed that data from more agricultural industries including more observations, data can significantly increase the predictive potential of the data that are exported from the agri-BCM templates or the proposed multidimensional data model.

### *5.1. Regression Analysis to Predict the %Increased Recovery Time*

Selected input variables have been utilized to predict the % increased recovery time, namely, the KBFs Impacted and the Recovery Priority Level. The main limitation is the low number of observations which has not permitted the inclusion of more than two input variables (explanatory variables) to predict the % increased recovery time (explained variable) for a set of *KBFs*. Multiple Linear Regression analysis has been conducted [27] using the R-Package [28].

Using association rule induction [29], the *KBFs impacted* has been proved the most important variable compared to the hazard type and the risk exposure variables when the target variable (predicted) is the risk exposure level. The association rule induction computations have been based on the decision tree induction [30] concepts and the criteria of entropy and information gain. The entropy of information is computed according to following formula (Eq. 5):

$$Entropy = \sum_{i=1}^{n} -P(v_i) \cdot log_2 P(v_i) \qquad (5)$$

When the aforementioned formula is applied to a dataset, then it quantifies the entropy of the value of a particular attribute of this dataset. $P(v_i)$ is the probability that the value of the attribute is $i$ and there are $n$ possible values of this attribute. Remainder of bits is another important measure, which quantifies the amount of information contained in the correct prediction about the value of a target attribute when the value of the other particular attribute $A$ is known. $Remainder(A)$ (Eq. 6) is equal to the entropy of an attribute $A$ weighted by the frequencies $n_i$ of values of an attribute $A$ ($k$ is the number of instances, i.e. the number of rows in the dataset table):

$$RemainderOfBits = \sum_{i=1}^{n} (n_i/k) \cdot Entropy(A)_i \qquad (6)$$

Information gain (Eq. 7) is then defined as the difference between the original information requirement, which is the entropy of the target attribute $A$, and the new information requirement Remainder regarding the value of attribute $A$.

$$Gain(A) = Entropy(A) - Remainder(A) \qquad (7)$$

The computations have been based on excel spreadsheet data and the conducted results are illustrated in the form of chart (Figure 9).

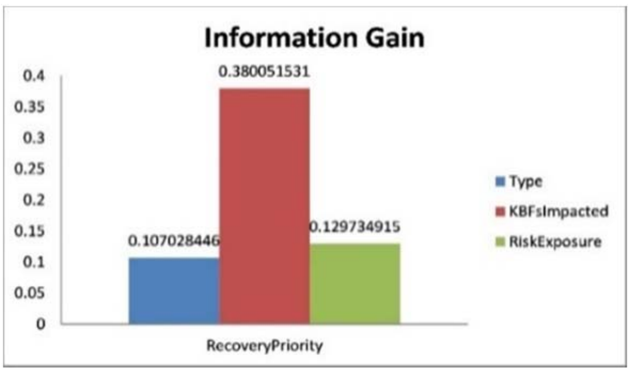


*Figure 9. The association rule induction using information gain regarding the variables KBFS impacted, risk exposure and hazard type when predicting recovery priority level for the affected KBF group*

In order to mitigate the risk of overfitting and conduct unbiased regression analysis, the dataset has been split into training (80% of the observations) and testing (20% of the observations) and used 10folds cross validation [31] in order to measure the accuracy of our predictions. Evaluation metrics for the conducted regression analysis has been the $R^2$ and the Root Mean Square Error (*RMSE*) [31] and the Mean Absolute Error (*MAE*) [32]. The conditions tested to prevent possible overfitting in the results was the number of records compared to the number of variables (min. 15 records per variable) [33] and the condition $RMSE_{\mathrm{TESTING}} < RMSE_{\mathrm{TRAINING}}$ [32]. Both conditions have been covered. The regression analysis results for the simple multiple regression and the 10folds cross validation indicated robust and promising predictive potential ($R^2_{\mathrm{10FOLDS\ REP.CV}}$ = 0.6774) (Table 5). However, the precondition for improving the predictive accuracy of the specific rule is to include more variables, which requires more BCM Template observations from more agri-business industries.

*Table 5. Accuracy measurement based on evaluation metrics for the multiple linear regression model (% increased recovery time predictor)*

| **Multiple Linear Regression** | $R^2$ | ***RMSE*** | ***MAE*** |
|---|---|---|---|
| Simple Model (testing_set over training_set) | 0.6080 | 8.1852 | 6.1545 |
| 10Folds CV | 0.6789 | 7.9733 | 6.1200 |
| Repeated 10 Folds CV (Repeats = 3) | 0.6774 | 8.4672 | 6.2541 |

The regression – based rule induction is thus expressed via the following formula (8) and the regression plot can be also visualized (Figure 10).

$$\%IncreasedRT = 10.9817 + 5.6983\ KFS\ Impacted - 5.9017\ Recovery\ Priority \tag{8}$$

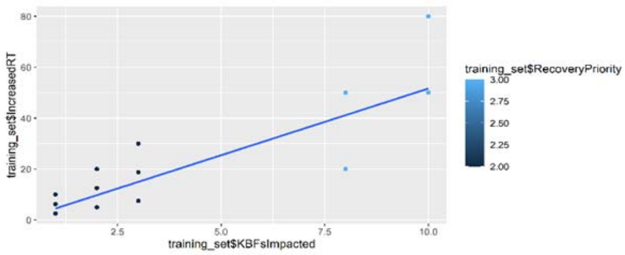


*Figure 10 The % increased recovery time for affected KBFS multiple linear regression plot*

## 5.2. Further Validation – Comparison of Recovery Priority Levels before and after applying Regression Analysis

The above conducted regression analysis prediction model has been further validated for its accuracy in predicting whether the *KBFs* Impacted by biological hazards, and the pandemic in particular, different results than the previously conducted as a result of BCM agribusiness continuity template data processing will be inferred. Using the regression analysis formula (8) new recovery time has been computed for the dataset records that include the impact of biological hazards to specific *KBF* groups or individual *KBFs*. A comparative Table (Table 6) has been utilized to show and that the predictive

*Table 6. The comparison of the recovery priority levels for sets of KBFS before and after the application of the regression analysis prediction function*

| **Industry** | **Hazard** | **Cause** | ***KBFs* Impacted** | **Average Recovery Time After Hazard (Hours)** | **Average Recovery Time Regression Prediction (Hours)** | **Recovery Priority Level Before Regression** | **Recovery Priority Level After Regression** | **Modified Recovery Priority Level for Key Business Functions Affected (0=No, 1=Yes)** |
|---|---|---|---|---|---|---|---|---|
| CF | Significant beef market decline | National off-site BSE/FEAD/FMD incident | 1 | 26.40 | 25 | 2 | 2 | 0 |
| CF | Loss of hourly employees | Pandemic | 3 | 25.80 | 28 | 2 | 2 | 0 |
| CF | Loss of hourly employees | FMD incident | 3 | 25.80 | 28 | 2 | 2 | 0 |
| CF | Loss of management | Pandemic | 8 | 108 | 125 | 3 | 3 | 0 |
| DF | Onsite FEAD/FMD outbreak | Natural or intentional | 10 | 146.88 | 123 | 3 | 3 | 0 |
| DF | Loss of hourly labor | Pandemic | 3 | 72.80 | 65 | 3 | 2 | 1 |
| DF | Loss of management | Pandemic | 3 | 180.60 | 185 | 3 | 3 | 0 |
| DF | Loss of Vet services | Pandemic | 1 | 73.80 | 71 | 3 | 2 | 1 |

accuracy in terms of recovery priority level ranking is more than average 60% (3 out of 5 correct predictions) when referring to the pandemic. The model is highly precise regarding the rest of the hazard causes, types and consequences. Both of the inaccurate predictions are related to the Dairy Sector. For the rest of the data records regarding the effect of biological hazards, the predictions of the regression formula have been proved accurate (75% accuracy, 6 out of 8 records). Regarding the multi-hazard risk assessment value of the formula, a 95% accuracy rate has been estimated, since only 3 out 56 indicated a modified risk exposure level.

### 5.3. Tackling Current Limitations Based on Future Research Activities

As it has been already mentioned, more agri-business continuity template data can be proved more precise in predicting the impact of pandemic and other hazard types on the recovery of key business functions. The multidimensional pattern can remain stable no matter the data size. However, future activities will eliminate the current limitations:

- Use of multi-criteria decision making [10] for multi-hazard risk assessment. The aspect of more stakeholders is demanded towards this direction. More precise results can be inferred but currently MCDM has not been considered to be essential since we relied exclusively on the available template data. MCDM quantification is based on thorough interviews and situation analysis;
- Multi-hazard risk assessment regarding individual *KBFs* and not only groups of *KBFs*. The specific activity is also considered as important in conducting future research outputs. The available data has not permitted the completion of this activity. However, the specific task is supported by the proposed multidimensional schema;
- Consider more detailed multi-hazard risk assessment and recovery time estimations for charting a more cost-effective and costly balanced [36] recovery policy.

## 6. Conclusion

The current contribution illustrated the potential of agri-business continuity template data to be utilized as input for a) designing standard business intelligence multidimensional risk assessment data models and b) creating datasets the analysis of which can facilitate the multi-hazard risk assessment policies in agribusiness. The case of two selected agricultural industries, namely, the dairy farms and the cattle feed yards, has been used for the design of the multidimensional data solution schema including dimensions, facts, hierarchy and granularity, and the investigation of the impact of multiple hazards-with a special focus on the pandemics on the recovery of key business functions in the event of their unforeseen interruption. The analysis of the spreadsheet data using OLAP pivot aggregate computations indicated that the presence of a pandemic during the recovery process of a set of key business functions can result in an approximate (13%) increase in the recovery time in both industries, when compared to the recovery time proposed when the hazard's effect is not considered (ideal conditions). In the dairy farms industry, an additional modification of the recovery priority level has been observed for two sets of key business functions, where the pandemic can result in loss of labor and a 30% increase in the recovery time and to loss of vet services and a 2.5% increase in the recovery time respectively. Using machine learning multiple regression analysis, the spreadsheet dataset potential to predict the % increase in recovery time has been demonstrated, considering the number of key business functions impacted and the risk exposure level as explanatory variables. The evaluation metrics results indicated an approximate accuracy level of 67% using the simple as well as the repeated 10 folds cross- validation. However, the incorporation of more data in our future research from multiple industries can further boost the predictive potential of the specific approach. Finally, the fact that the generated regression formula has been 65%, 75%, and 95% accurate in predicting whether the recovery priority level for a given set of business functions will be modified or not (binary decision) when the pandemic, any biological hazard and any hazard category will affect the recovery process of key agribusiness functions respectively, it is also an implication of a highly promising and robust predictive decision making technique in terms of multi-hazard risk assessment policy formulation based on BCM template data. The future research will focus on dealing with the current issue of limited data records via electronic questionnaires, and on developing a physical data warehouse repository for multi-hazard risk assessment for agribusiness sectors to ensure resilient and sustainable agri-business continuity policies in the post-COVID period.

**Acknowledgements**

*This work was supported by the Czech Ministry of Agriculture [grant number QK1920391]; the work was also supported by the Faculty of Economics, Technical University of Liberec, Czech Republic.*

**References**

[1]. Komarek, A. M., De Pinto, A., & Smith, V. H. (2020). A review of types of risks in agriculture: What we know and what we need to know. *Agricultural Systems*, *178*, 102738.

[2]. El Bilali, H., & Allahyari, M. S. (2018). Transition towards sustainability in agriculture and food systems: Role of information and communication technologies. *Information Processing in Agriculture*, *5*(4), 456-464.

[3]. Lorencová, E., Frélichová, J., Nelson, E., & Vačkář, D. (2013). Past and future impacts of land use and climate change on agricultural ecosystem services in the Czech Republic. *Land Use Policy*, *33*, 183-194.

[4]. Davidson, D. (2016). Gaps in agricultural climate adaptation research. *Nature Climate Change*, *6*(5), 433-435.

[5]. Prášilová, M., & Procházková, R. (2020). Structural changes of Czech agriculture and the impact of these on inner foodstuffs self-sufficiency of Czech Republic. *Vision*, 2905-2919.

[6]. Janků, J., Sekáč, P., Barakova, J., & Kozak, J. (2016). Land use analysis in terms of farmland protection in the Czech Republic. *Soil and Water Research*, *11*(1), 20-28.

[7]. Urbancova, H., & Čermáková, H. (2015). The costs of Age Management in agricultural companies. *Agricultural Economics*, *61*(1), 14-22.

[8]. FAO. (2021). *Novel Coronavirus (COVID-19)*. The Food and Agriculture Organization of the United Nations. Retrieved from: http://www.fao.org/2019-ncov/q-and-a/en/ [accessed: 07 March 2021].

[9]. Hoogenboom, G. (2020). Agronomic data and the coronavirus epidemic. *The Journal of Agricultural Science*, *158*(3), 161-162.

[10]. Therond, O., Duru, M., Roger-Estrade, J., & Richard, G. (2017). A new analytical framework of farming system and agriculture model diversities. A review. *Agronomy for Sustainable Development*, *37*(3), 1-24.

[11]. Ibn-Mohammed, T., Mustapha, K. B., Godsell, J., Adamu, Z., Babatunde, K. A., Akintade, D. D., ... & Koh, S. C. L. (2021). A critical analysis of the impacts of COVID-19 on the global economy and ecosystems and opportunities for circular economy strategies. *Resources, conservation, and recycling*, *164*, 105169.

[12]. Jámbor, A., Czine, P., & Balogh, P. (2020). The Impact of the Coronavirus on Agriculture: First Evidence Based on Global Newspapers. *Sustainability*, *12*(11), 1-10.

[13]. Supriadi, L. S. R., & Pheng, L. S. (2018). Business continuity management (BCM). In *Business Continuity Management in Construction* (pp. 41-73). Springer, Singapore.

[14]. Farr, M., & Bailey, D. (2019). Uniting business continuity management and operational risk management. *Journal of business continuity & emergency planning*, *12*(4), 294-300.

[15]. Culhane, M., Cardona, C., Goldsmith, T. J., Charles, K. S., Suskovic, G., Thompson, B., & Starkey, M. (2018). Building an all-hazards agricultural emergency response system to maintain business continuity and promote the sustainable supply of food and agricultural products. *Cogent Food & Agriculture*, *4*(1), 1550907.

[16]. Lytos, A., Lagkas, T., Sarigiannidis, P., Zervakis, M., & Livanos, G. (2020). Towards smart farming: Systems, frameworks and exploitation of multiple sources. *Computer Networks*, *172*, 107147.

[17]. Kamilaris, A., Kartakoullis, A., & Prenafeta-Boldú, F. X. (2017). A review on the practice of big data analysis in agriculture. *Computers and Electronics in Agriculture*, *143*, 23-37.

[18]. Wolfert, S., Ge, L., Verdouw, C., & Bogaardt, M. J. (2017). Big data in smart farming–a review. *Agricultural systems*, *153*, 69-80.

[19]. Řezník, T., Lukas, V., Charvát, K., Charvát, K., Křivánek, Z., Kepka, M., ... & Řezníková, H. (2017). Disaster Risk Reduction in Agriculture through Geospatial (Big) Data Processing. *ISPRS International Journal of Geo-Information*, *6*(8).

[20]. Yarovenko, H., Bilan, Y., Lyeonov, S., & Mentel, G. (2021). Methodology for assessing the risk associated with information and knowledge loss management. *Journal of Business Economics and Management*, *22*(2), 369-387.

[21]. Sharma, R., Shishodia, A., Kamble, S., Gunasekaran, A., & Belhadi, A. (2020). Agriculture supply chain risks and COVID-19: mitigation strategies and implications for the practitioners. *International Journal of Logistics Research and Applications*, 1-27.

[22]. Hájek, P., & Urbancová, H. (2013). Using of Business Continuity Standards in Agriculture, Industry and ICT. *AGRIS on-line Papers in Economics and Informatics*, *5*(665-2016-44979), 55-67.

[23]. Engemann, K. J., & Henderson, D. M. (2014). *Business continuity and risk management: essentials of organizational resilience*. Rothstein Publishing..

[24]. Coble, K. H., Mishra, A. K., Ferrell, S., & Griffin, T. (2018). Big data in agriculture: A challenge for the future. *Applied Economic Perspectives and Policy*, *40*(1), 79-96.

[25]. Vaisman, A., & Zimányi, E. (2014). Data warehouse systems. *Data-Centric Systems and Applications*.

[26]. Caniupán, M., Bravo, L., & Hurtado, C. A. (2012). Repairing inconsistent dimensions in data warehouses. *Data & Knowledge Engineering*, *79*, 17-39.

[27]. Muangprathub, J., Boonnam, N., Kajornkasirat, S., Lekbangpong, N., Wanichsombat, A., & Nillaor, P. (2019). IoT and agriculture data analysis for smart farm. *Computers and electronics in agriculture*, *156*, 467-474.

[28]. Peng, R. D. (2020). *R Programming for Data Science*. Leanpub, Canada. Retrieved from: https://leanpub.com/rprogramming [accessed: 07 March 2021].

[29]. Ordonez, C., & Zhao, K. (2011). Evaluating association rules and decision trees to predict multiple target attributes. *Intelligent Data Analysis*, *15*(2), 173-192.

[30]. Russell, S., & Norvig, P. (2020). Artificial intelligence: a modern approach 4th ed. Pearson Education UK.

[31]. Andryeyev, O., Onus, U., Casas, V., & Mitschele-Thiel, A. (2019, May). Experimental validation of air-to-ground propagation models for low-altitude platforms. In *2019 15th International Conference on Distributed Computing in Sensor Systems (DCOSS)* (pp. 591-595). IEEE.

[32]. Pošík, P. (2015). *Bias-variance trade-off. Crossvalidation. Regularization.* Retrieved from: https://cw.fel.cvut.cz/old/_media/courses/ae3m33ui/lectures/biasvariance-slides.pdf [accessed: 10 February 2021].

[33]. Babyak, M. A. (2004). What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. *Psychosomatic medicine*, *66*(3), 411-421.

[34]. PRPC. (2013). *Texas Dairy Farm Business Continuity Guide*. Panhandle Regional Planning Commission. Retrieved from: http://www.theprpc.org/Programs/RegionalServices/Dairy-industry/prpc_dairy/bcdoc/Texas%20Dairy%20Farm%20BCG%20Template.docx. [accessed: 05 March 2021].

[35]. PRPC. (2013). *Texas Cattle Feedyard Business Continuity Guide Supplement*. Panhandle Regional Planning Commission. Retrieved from: http://theprpc.org/Programs/RegionalServices/Cattle-Feedyard-Industry/PRPC_CattleFeed/bcdoc/Texas%20Cattle%20Feedyard%20BCG%20Supplement.docx. [accessed: 05 March 2021].

[36]. Rittinghouse, J., Ransome, J. F., & CISM, C. (2011). Business continuity and disaster recovery for infosec managers. Elsevier.

[37]. Podaras, A., & Nejedlová, D. (2017). A Risk Management Framework for Business Continuity in Agriculture. In *Proceedings of the 8th International Conference on Information and Communication Technologies in Agriculture, Food & Environment (HAICTA 2017)*. September 21-24, Greece, Chania, Crete.