Czech University of Life Sciences Prague
Faculty of Economics and Management

agris on-line

Papers in Economics and Informatics

http://online.agris.cz

ISSN 1804-1930
XII, 2020, 3

International scientific journal
Prague

Agris on-line Papers in Economics and Informatics

The international reviewed scientific journal issued by the Faculty of Economics and Management of the Czech University of Life Sciences Prague.

The journal publishes original scientific contributions from the area of economics and informatics with focus on agriculture and rural development.

Editorial office

AGRIS on-line Papers in Economics and Informatics
Department of Economics of FEM CZU Prague
Kamýcká 129, 165 00 Praha-Suchdol
Czech Republic
Phone: +420 224 382 056
E-mail: agrisonline(at)pef.czu.cz

Publisher

Faculty of Economics and Management
Czech University of Life Sciences Prague
Kamýcká 129, 165 00 Praha-Suchdol
Czech Republic
Reg. number: 60460709

ISSN 1804-1930

XII, 2020, 3
30th of September 2020
Prague

Agris on-line

Papers in Economics and Informatics

ISSN 1804-1930

XII, 2020, 3

Agris on-line Papers in Economics and Informatics

Volume XII Number 3, 2020

Content:

Heuristic Evaluation of the User Interface for a Semi-Autonomous Agricultural Robot Sprayer

George Adamides

Rural Development Department, Agricultural Research Institute, Nicosia, Cyprus

Abstract

This study presents the heuristic evaluation, as a usability inspection method for Human-Robot Interaction (HRI) systems. First, the methodology to engineer a semi-autonomous agricultural robot sprayer is presented, and then findings from heuristic usability evaluation studies that were carried out on a human-robot interface for a semi-autonomous agricultural vineyard robot sprayer. The following research-based heuristics for the design of robot teleoperation were used: Platform architecture and scalability, Error prevention and recovery, Visual design, Information presentation, Robot state awareness, Interaction effectiveness and efficiency, Robot environment/surroundings awareness, and Cognitive factors. In each evaluation study, usability problems were identified, and specific suggestions were documented for HRI usability improvement. In each design iteration, a smaller number of usability issues were identified.. Results of the final heuristic evaluation showed that the system is at a good level of usability and is expected to provide satisfactory services to its typical users.

Keywords

Agricultural robot, human-robot interaction, heuristic evaluation, usability study.

Adamides, G. (2020) “Heuristic Evaluation of the User Interface for a Semi-Autonomous Agricultural Robot Sprayer", *AGRIS on-line Papers in Economics and Informatics*, Vol. 12, No. 3, pp. 3-12. ISSN 1804-1930. DOI 10.7160/aol.2020.120301.

Introduction

According to the Food and Agriculture organization of the United Nations (FAO), the world population will reach over 9 billion by 2050 (FAO, 2009). This projection implies that agricultural food production will need to double, in order to achieve food security. In addition, climate change, limited land and water resources (FAO, 2010), the observed shortage of agricultural workers (Hertz and Zahniser, 2013), and farmers’ aging coupled with the hardness of agricultural tasks, increases the burden of producing more agricultural products, with limited resources and environmental constraints.

Due to technological advancements in recent years, robotics has begun to play a major role in our daily lives (Arad et al., 2020). Automation in agriculture, mechanization and agricultural engineering, has been a major force for increased productivity in the 20th century (McNulty and Grace, 2009). In fact, according to Huffman and Evenson (2001), the aggregate United States farm output was 5.5 times larger in 1990 than a hundred years ago.

In the past two decades, research and development on agricultural robots has been spurred on by the emergence of new and increasingly cost-effective advances in engineering, sensing and actuating technologies, along with the decrease of technology cost (Sistler, 1987). According to Adamides (2016), “*robotics for agriculture is considered the domain of field systems able to perform coordinated, mechatronic actions, on the basis of processing of information acquired through camera(s) and other sensor technology, with the aim to support farmers in performing agricultural tasks.*”

Agriculture is a suitable application area for robotics given the hard working conditions (e.g. physical work and severe weather conditions) and difficulty and repetitiveness of the work. Robotic technology can enhance the farmer’s capabilities (i.e., perception, decision-making) to carry out repetitive, tedious, and in some cases dangerous agricultural tasks (e.g. weeding (Kargar et al., 2013) and spraying (Oberti et al., 2016)) in dynamic and unstructured environments.

Unlike the case of industrial robots, which operate in a structured and controlled environment, fully autonomous agricultural robots are expected to operate under several complexities as identified by Edan et al. (2009). Examples of such difficulties is moving on unstructured terrain, dealing with highly variable fruits that differ in size, color (even at the same plant), and environmental issues like shading and lighting. Autonomous robotic sprayers have been developed for weed control in field applications (Åstrand and Baerveldt, 2002, Kargar et al., 2013), trees in orchards (Endalew et al., 2011, Brown et al., 2008), and vineyards (Berenstein and Edan, 2012, Berenstein and Edan, 2017). Selective spraying pesticides towards the targets, using a robot sprayer could reduce up to 30% of the pesticide (spraying material) while detecting and spraying 90% of the grape clusters (Berenstein et al., 2010). Today, vineyard spraying is achieved by spraying uniform amounts of pesticides along the vineyard rows without considering low density foliage, which requires less pesticide, or gaps between the trees. Figure 1, illustrates current methods used for spraying vineyards.

A semi-autonomous robot implies that some operations the robot carries out autonomously (as preprogrammed), while others are carried out under the supervision/ guidance of a human operator. In this sense, Fong et al. (2001) explains that human-robot interaction (HRI) is "*the study of the humans, robots and the ways they influence each other.*" The goal of HRI in an agricultural environment is to develop and use efficiently robots such that farmers become more productive, effective, in a safe environment, freeing them from the hardness, routine and dangerous tasks (e.g. spraying pesticides).

Semi-autonomous teleoperation of an agricultural robotic system can help to effectively manage the complexity and performance limitations that current autonomous robots face due to the dynamic and unstructured agricultural environment. Semi-autonomous teleoperation implies the existence of a user interface (UI) that supports human-robot interaction. Such a user interface needs to meet specific non-functional requirements, such as reliability, efficiency and usability.

This study focuses on usability, which refers to whether a HRI system can be used with interaction effectiveness, interaction efficiency, and satisfaction with which specified types of users achieve specified goals in a particular context of use. In specific, this article presents findings from heuristic usability evaluations that were carried out on two versions of such a human-robot interface for a semi-autonomous agricultural vineyard robot sprayer. Three iterations of the user-centered design process were followed to ensure the usability of the final product.

Nielsen and Molich (1990) explains that "*Heuristic evaluation is a 'discount usability engineering' method for evaluating user interfaces to find their*

Source: author

Figure 1: Farmer on a tractor-sprayer in a vineyard field.

usability problems”. This method is characterized as “discount” because a small number of evaluators, usually 3 to 7, suffice to reliably evaluate the usability of a user interface against a list of heuristics (the usability principles).

A usability issue is anything that can affect the user experience in a negative way. There are many sources of data that can be used to derive usability issues; the most common ones (Tullis and Albert, 2008) include user performance data (e.g., task success rate, time on task), verbal expressions of confusion or dissatisfaction (such as from a think-aloud protocol (Alan et al., 2004)), behavioral/physiological data (e.g., from eye-tracking (Poole and Linden, 2005)) and reports from usability experts (e.g. heuristic evaluation (Nielsen, 1994c)). Usability issues are often prioritized based on severity schemes (Nielsen, 1994b, Dumas and Redish, 1999) that take into account various factors (e.g. expected impact on user experience, predicted frequency of occurrence, expected impact on business goals) in an attempt to increase their usefulness for the next design iteration.

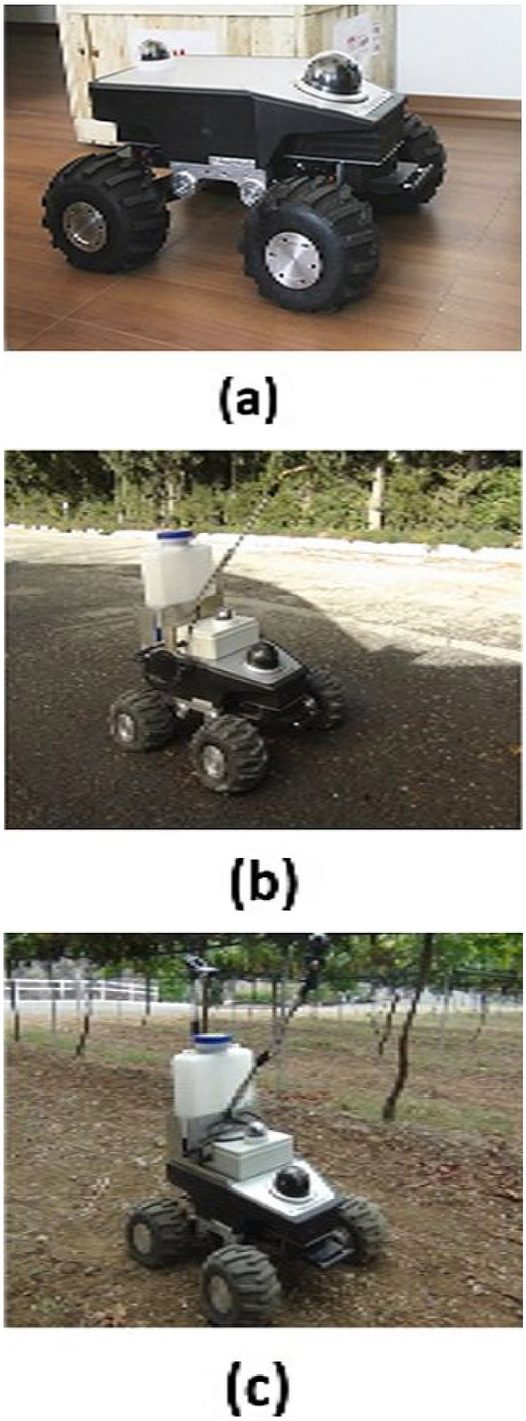
The objective of this article is twofold: a) to present the methodology to engineer a semi-autonomous agricultural robot sprayer and b) the usability evaluation of the developed user interfaces using heuristics. What follows is the presentation of the research methodology, and later results and findings are discussed. Finally, the conclusions and prospects for further research are presented.

Materials and methods

Robot sprayer development stages

A mobile platform by Robotnik (http://www.robotnik.eu) was used (Figure 2a). This platform is a medium-sized, high mobility all-terrain robot, with skid-steering kinematics based on four high power motor-wheels. This specific platform was selected because it can move both indoors (i.e. greenhouse) and outdoors (i.e. agricultural field) in a variety of field applications.

The original design of the robot was based on the analysis of user contextual interviews of farmers and agronomists that pilot tested in the field an initial version of the agricultural robot sprayer (Adamides et al., 2014). Initially, in robot version 1 (Figure 2b), several HRI related limitations were identified such as: a) the lack of peripheral vision, b) the operator required a significant amount of time to pan-tilt zoom-in and zoom-out from the main robot camera, c) limitations to Bluetooth connection via the PS3 gamepad controller, and d) illumination of the laptop monitor (used for robot control) due to sunlight. Following, informal interviews and documentation of the pilot participants’ observations, several modifications on the platform resulted to an improved version.



Source: author

Figure 2: Development stages of the semi-autonomous robot sprayer.

The final version (Figure 2c) included a peripheral camera on the back-top of the platform and an end-effector camera on-top of the nozzle canon sprayer. In order to solve the issue of the distance limit of the PS3 gamepad controller, two solutions were provided: a) connecting the controller through WiFi and b) adding a PC keyboard alternative as input device. Similarly, to address the issue of sunlight and illumination of the PC monitor, also two alternatives were implemented: a) using digital glasses and b) teleoperating the robot from inside an office environment.

The following HRI taxonomy (Table 1) was assumed in this article for the semi-autonomous agricultural robot sprayer, based on the HRI taxonomy proposed by Yanco and Drury (2004).

In the specific case of the semi-autonomous robot sprayer, the navigation task (robot path guidance) was performed in teleoperation mode, while the target marking/ identification and spraying tasks were performed in autonomous or semi-autonomous mode.

User interface heuristic evaluation

The heuristic evaluation method was employed, as one of the most popular usability inspection techniques, which are also known as expert-based methods, user-free methods or methods performed in the lab without end-users. An adequate number of experts was found and recruited so that reliable evaluation results could be obtained. First, the evaluators were informed about the system goal, its representative users and their typical tasks and the developers' design goals and expectations. Next, they used the system and conducted an individual heuristic evaluation according to a specific protocol, a selected set of heuristics appropriate for the evaluation context, and a template for reporting the identified usability issues. The evaluators were situated at the Hellenic Open University Software Quality Assessment laboratory and controlled the robot remotely, which was located at the Open University of Cyprus, Nicosia premises. An appropriate lab-simulation environment was created, including various paths and targets. After each individual evaluation, the participating evaluators conducted a focus group to group and prioritize the identified usability issues.

Four usability experts – an adequate number to ensure reliable results (Nielsen, 1994a) – conducted a heuristic usability evaluation on two user interfaces. All four have undergraduate and/or postgraduate studies in Computer Science and extensive experience in the design and evaluation of interactive systems.

Two user interfaces for the Semi-Autonomous Agricultural Robot Sprayer were evaluated: UIv0 and UIv1 Figure 3 presents the two main screens of these user interfaces.

Category	Description	Classification
Task type	There are three tasks to be executed in this HRI: guiding the robot in the vineyards, identifying targets to spray, and the actual spraying task	[Navigation (robot path guidance), Target Marking/ Identification, Spraying]
Task Criticality	Given that in robot navigation there is a possibility to harm either the robot or bystanders or the vines, the task criticality is High. For the target identification and spraying the criticality is set to low.	[High, Low]
Robot morphology	Mobile robotic platform with spraying capabilities	[Functional]
Ratio of people to robots	One human operator and one robot sprayer	[1:1]
Composition of robot teams	Same robot	[Homogeneous]
Level of shared interaction	One human operator and one robot sprayer	[one human, one robot]
Interaction roles	During Autonomous mode the human is acting as supervisor. During the teleoperation mode the human is acting as Operator. During the semi-Autonomous mode the human is acting as teammate.	[Supervisor, Operator, Teammate]
Type of human-robot physical proximity	The human and the robot are not collocated	[Avoiding]
Decision support for operators	Battery level, camera and sonar sensors	[Provided sensors]
Time/Space taxonomy	Human and robot operate at the same time in different locations	[Time (Synchronous), Space (Non-collocated)]
Autonomy level / Amount of intervention	There is a continuum for robot control ranging from teleoperation to full autonomy	[Autonomy+Intervention=100%]

Source: author

Table 1: HRI taxonomy for the semi-autonomous agricultural robot sprayer.

Note: The red rectangles and black text boxes are not part of each user interface

Source: author

Table 1: HRI taxonomy for the semi-autonomous agricultural robot sprayer.

UIv0: In terms of functionality, the main designs considered were: a) on-screen controls of the robot movement and camera movement, b) presentation of camera views, and c) addition of elements for displaying sensor information (visual and auditory feedback) for distance from the robot sides and battery level. One important priority when designing UIv0, was to enable the operator to use the entire screen and support interaction through either the keyboard or the mouse.

UIv1: is an upgraded user interface of UIv0. It provides functionality for target pointing. In specific, UIv1 supports both manual (user points to targets) and automated target specification through a pattern recognition algorithm.

The same procedure was followed in all heuristic evaluation studies. A set of research-based heuristics for the design of robot teleoperation, which have been developed in Adamides et al. (2015) were used. These are:

- **Platform architecture and scalability:** The user interface of an HRI system should be flexible to follow and benefit from developments in computing and robotic technologies.
- **Error prevention and recovery:** The information provided by the user interface should prevent user errors, and if a user makes a mistake, the user interface should allow for its rectification. In contrast with undoing a “Cut” operation in a word processor, a “Cut” command to prune a tree through a teleoperated AgriRobot cannot be undone.
- **Visual design:** Since the user interface is the communication medium between the operator and the remote robot, it should provide the operator with only relevant information (from video and other robot sensors) in a simple, consistent, effective, and minimalist way. Specific examples include minimizing use of multiple windows, avoiding window occlusion, providing large video windows and displaying the robot’s body in the interface.
- **Information presentation:** Controlling a remotely located robot is demanding on operators who need to integrate various sources of information coming from the robot cameras and sensors. Therefore, information presentation is of high importance in this type of user interface designs, to enhance situation awareness of the operators, and to bridge

the gaps of execution and evaluation (Alan et al., 2004).

- **Robot state awareness:** The robot should be able to self-inspect its systems and take autonomous action or ask for user intervention. The human operator should have a clear understanding about the robot status and activities. For instance, to support understanding of the camera(s) and their position(s), the over-all mission and the current progress, and when multiple robots are available, use one to view another.
- **Interaction effectiveness and efficiency:** In HRI, efficiency is measured in terms of the time required to complete a task; effectiveness is measured in terms of how well a task is completed.
- **Robot environment/surroundings awareness:** Environment awareness is essential, because in field robot teleoperation it is important to have knowledge of the robot's whereabouts and the area covered, such as orientation, obstacles, or why a robot is not moving. This can be accomplished through maps, orientation information (such as compass), and sensors that will provide the necessary information about the robot's surroundings.
- **Cognitive factors:** Cognitive factors are characteristics that affect performance and learning. The user interface of a teleoperated HRI system should be designed such that it directs the user's attention to the task the robot is operating, improves learnability, and provides fused information from the various sensors and cameras from the robot, in order to lower the cognitive load on the user.

Next, the evaluators were informed about the system goal, its representative users and their typical tasks. Subsequently, each evaluator conducted a heuristic evaluation of the system. To this end, they were provided with access to the semi-autonomous agricultural robot sprayer UI versions under evaluation. They first familiarized themselves with the system by performing typical user tasks and exploring its functionality. Next, they inspected the system, identified usability issues and wrote them down following specific evaluation template. For each problem, they noted the heuristic violated and rated its severity on a scale from 1 to 5 (1 = a little important, it does not significantly affect the user interaction, 5=extremely important, catastrophic problem that may result in unsuccessful task, danger to life or damage to property). In evaluating the severity of a usability problem, they were asked to take into account the following factors (Nielsen, 1995): a) frequency, b) impact, and c) persistence. Finally, each evaluator was asked to provide a design suggestion for resolving the identified usability issue. The four evaluators produced individual reports with the identified usability issues per heuristic rule (Adamides et al., 2015).

After each individual evaluation, the study coordinator and the evaluators participated in a focus group in order to produce the final list of unique problems, discuss on the final severity ratings and proposals for solutions.

Results and discussion

Results of the heuristic evaluation showed that the systems under evaluation provide very good (in terms of usability issues identified by experts) services to their expected typical users. A small number of usability problems were identified, whose redress can improve the overall user experience with the system.

In the following, the results for each evaluated system are presented. The total number of expected problems for the each system was calculated using the formula (Nielsen and Landauer, 1993):

$$N = \frac{1-(1-j)^{i}}{\text{Problemsfound}(i)} \qquad (1)$$

where N is the total number of expected usability problems, i is the number of independent experts-evaluators, *ProblemsFound(i)* is the total number of unique usability issues identified by the participating evaluators, and j is the average proportion of problems found by a single evaluator.

A. First user interface: UIv0

For UIv0, 13 usability issues were identified. Most (77%) of these usability issues were related to violations of the following four heuristics: a) 23% were violations of heuristic 4 (Information presentation), b) 23% were violations of heuristic 5 (Robot state awareness), c) 15% were violations of heuristic 6 (Interaction effectiveness and efficiency) and d) 15% were violations of heuristic 8 (Cognitive factors). In terms of problem severity, the issues with the highest priority were related to violations of the following three heuristics: a) heuristic 5 (Robot state awareness) with the highest average severity (4.0), b) heuristic 2 (Error prevention and recovery) with the second from top average severity (4.0), and c) heurist 3 (Visual design) with also second from top average severity (4.0).

The expected number of usability problems for UIv0 was calculated to 42, which is above the average number of usability problems (35) observed in a rather mature interactive system (Nielsen and Landauer, 1993). In addition, a substantial number of problems (9) were rated as 3+ on a severity scale from 1 to 5. The average severity of the identified problems is characterized as medium (3.3). All in all, the system is at a satisfactory level of usability. However, there are changes that could further improve its usability.

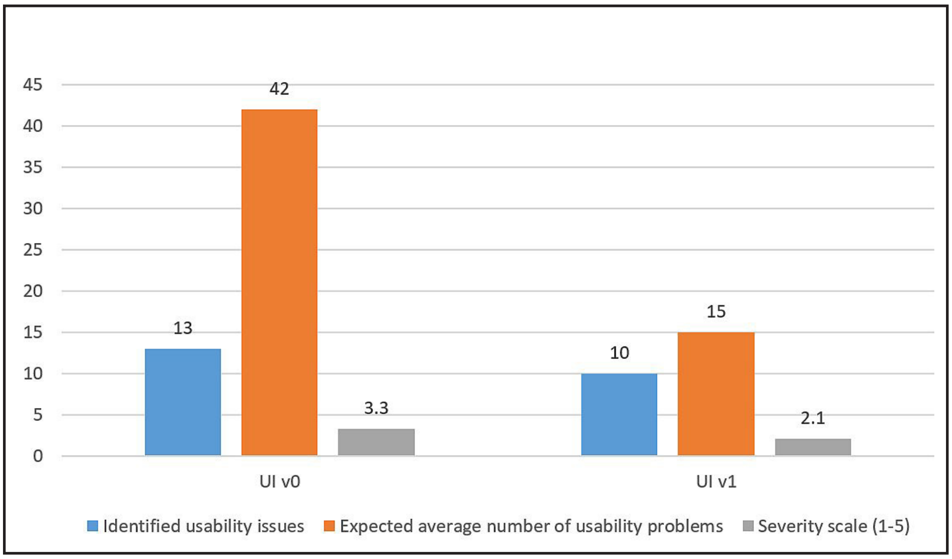
B. Second user interface: UIv1

Regarding UIv1, 10 usability issues were identified. Most (80%) of these usability issues were related to violations of the following four heuristics: a) 20% were violations of heuristic 4 (Information presentation), b) 20% were violations of heuristic 5 (Robot state awareness), c) 20% were violations of heuristic 6 (Interaction effectiveness and efficiency) and d) 20% were violations of heuristic 8 (Cognitive factors). In terms of problem severity, the issues with the highest priority were related to violations of the following three heuristics: a) heuristic 2 (Error prevention and recovery) with the highest average severity (4.0), b) heuristic 6 (Interaction effectiveness and efficiency) with the second from top average severity (3.0), and c) heuristic 5 (Robot state awareness) with third from top average severity (2.0).

The expected number of usability problems for UIv1 was calculated to 15, which is less than half the average number of usability problems (35) observed in a rather mature interactive system. In addition, a small number of problems (3) were rated as 3+ on a severity scale from 1 to 5. The average severity of the identified problems is characterized as low (2.1). These findings (shown in Figure 4) tend to provide support that the system is at a good level of usability.

According to the expert evaluators, one important advantage of both user interface versions is that they take full advantage of the screen size providing a large window for the central and peripheral views. In addition, the user can easily customize the placement and size of the end-effector camera view. Furthermore, implicit switching of autonomy level is supported, but it should be better communicated to the user. Moreover, important information, such as the exact distance from obstacles and the remaining battery level, are always available. However, equally important information, such as the remaining level of spraying liquid (the robot is used to spray vineyards), is not available at all.

UIv1 supports functionality for targeted spraying in a rather intuitive way. The target identification algorithm automatically selects targets (i.e. grape clusters) to spray and the spraying is activated after 2 seconds, unless the operator interrupts this action viz. to cancel a false negative or modify a false positive target for spraying. However, there are UI improvements that could be made in the manual target addition and deletion to better reflect what the user is doing. In addition, these systems provide support



Source: author

Figure 4: Heuristic evaluation results of Uiv0 and Uiv1.

for automated target identification, which may lead to increased efficiency in the actual field. However, the associated dialogue for changing the algorithm settings is in a highly technical and complicated language for the typical user.

Conclusion

This study presented in detail the heuristic evaluation method, and how this was applied, as a usability inspection method for Human-Robot Interaction (HRI) systems. The usability study findings provide evidence that the final version of the semi-autonomous robot sprayer system provides satisfactory services to its typical users. These advantages, combined with the increased usability of the UIv1 (final) system, may result in high adoption from its end users. This study is limited by the fact that it focused on the usability evaluation of the user interface for a HRI system. Additional experiments are underway and will focus on specific tasks such as comparing different applications of spraying (e.g. using a robotic arm) and evaluate the amount of spraying chemicals saved, in addition to other agricultural tasks that can be automated through robotization.

Corresponding authors
George Adamides
Agricultural Research Institute, P.O. Box 22016, 1516 Nicosia, Cyprus
E-mail:gadamides@ari.gov.cy

References

[1] Adamides, G. (2016) "User Interfaces for Human-Robot Interaction: Application on a Semi-Autonomous Agricultural Robot Sprayer", PhD Doctoral Dissertation, Open University of Cyprus, Nicosia. ISBN 978-9963-695-49-2.

[2] Adamides, G., Christou, G., Katsanos, C., Xenos, M. and Hadzilacos, T. (2015) "Usability Guidelines for the Design of Robot Teleoperation: A Taxonomy", *IEEE Transactions on Human-Machine Systems*, Vol. 45, No. 2, pp. 256-262. ISSN 2168-2291. DOI 10.1109/thms.2014.2371048.

[3] Adamides, G., Katsanos, C., Christou, G., Xenos, M., Papadavid, G. and Hadzilacos, T. (2014) "User interface considerations for telerobotics: The case of an agricultural robot sprayer", In *Second international conference on remote sensing and geoinformation of the environment (RSCy2014),* Vol. 9229, pp. 92291W, International Society for Optics and Photonics. DOI 10.1117/12.2068318.

[4] Alan, D., Janet, F., Gregory, A. and Russell, B. (2004) "*Human-computer interaction*", England: Pearson Education Limited. ISBN 978-0-13-046109-4.

[5] Albert, W. and Tullis, T. (2008) "Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics (Interactive Technologies)", Morgan Kaufmann. E-ISBN 978-0-12-415781-1.

[6] Arad, B., Balendonck, J., Barth, R., Ben-Shahar, O., Edan, Y., Hellström, T., Hemming, J., Kurtser, P., Ringdahl, O., Tielen, T. and van Tuijl, B. (2020) "Development of a sweet pepper harvesting robot", *Journal of Field Robotics*, Vol. 34, No. 6, pp. 1027-1039. E-ISSN 1556-4967. DOI 10.1002/rob.21937.

[7] Åstrand, B. and Baerveldt, A.-J. (2002) "An Agricultural Mobile Robot with Vision-Based Perception for Mechanical Weed Control", *Autonomous Robots*, Vol. 13, pp. 21-35. E-ISSN 1573-7527, ISSN 0929-5593. DOI 10.1023/a:1015674004201.

[8] Berenstein, R. and Edan, Y. (2012) "Evaluation of marking techniques for a human-robot selective vineyard sprayer", Paper presented at *the International Conference of Agricultural Engineering*, Valencia, Spain.

[9] Berenstein, R. and Edan, Y. (2017) "Human-robot collaborative site-specific sprayer", *Journal of Field Robotics*, Vol. 34, No. 8, pp. 1519-1530. E-ISSN 1556-4967. DOI 10.1002/rob.21730.

[10] Berenstein, R., Shahar, O. B., Shapiro, A. and Edan, Y. (2010) "Grape clusters and foliage detection algorithms for autonomous selective vineyard sprayer", *Intelligent Service Robotics*, Vol. 3, No. 4, pp. 233-243. E-ISSN 1861-2784, ISSN 1861-2776. DOI 10.1007/s11370-010-0078-z.

[11] Brown, D., Giles, D., Oliver, M. and Klassen, P. (2008) "Targeted spray technology to reduce pesticide in runoff from dormant orchards", *Crop Protection*, Vol. 27, pp. 545-552. ISSN 0261-2194. DOI 10.1016/j.cropro.2007.08.012.

[12] Dumas, J. S. and Redish, J. C. (1999) "*A practical guide to usability testing*", Intellect Books. ISBN 1-84150-020-8.

[13] Edan, Y., Han, S. and Kondo, N. (2009) "Automation in Agriculture", *Springer Handbook of Automation*. Berlin, Heidelberg: Springer Berlin Heidelberg. pp. 1095-1128. ISBN 10: 3540788301, ISBN 13: 9783540788300. DOI 10.1007/978-3-540-78831-7_63.

[14] Endalew, A. M., Debaer, C., Rutten, N., Vercammen, J., Delele, M. A., Ramon, H., Nicolaï, B. M. and Verboven, P. (2011) "Modelling the effect of tree foliage on sprayer airflow in orchards", *Boundary-Layer Meteorology*, Vol. 138, No. 1, pp. 139-162. E-ISSN 1573-1472, ISSN 0006-8314. DOI 10.1007/s10546-010-9544-6.

[15] FAO (2009) "How to Feed the World in 2050", *Population and Development Review*, Vol. 35, No. 4, pp. 837-839. E-ISSN ISSN:1728-4457. DOI 10.1111/j.1728-4457.2009.00312.x.

[16] FAO (2010) "Climate-smart Agriculture: Policies, Practices and Financing for Food Security, Adaptation and Mitigation", Food and Agriculture Organization of the United Nations, Rome, Italy.

[17] Fong, T., Thorpe, C. and Baur, C. (2001) "Collaboration, dialogue, human-robot interaction", In: Jarvis R. A., Zelinsky A. (eds) *Robotics Research, Springer Tracts in Advanced Robotics,* Vol 6. Springer, Berlin, Heidelberg. DOI 10.1007/3-540-36460-9_17.

[18] Hertz, T. and Zahniser, S. (2013) "Is there a farm labor shortage?", *American Journal of Agricultural Economics*, Vol. 95, No. 2, pp. 476-481. E-ISSN 1467-8276, ISSN 0002-9092. DOI 10.1093/ajae/aas090.

[19] Huffman, W. E. and Evenson, R. E. (2001) "Structural and productivity change in US agriculture, 1950–1982", *Agricultural economics*, Vol. 24, No. 2, pp. 127-147. ISSN 0169-5150. DOI 10.1111/j.1574-0862.2001.tb00019.x.

[20] Kargar, B., Amir, H. and Shirzadifar, A. M. (2013) "Automatic weed detection system and smart herbicide sprayer robot for corn fields", Paper presented at the *First RSI/ISM International Conference on Robotics and Mechatronics*. DOI 10.1109/ICRoM.2013.6510152.

[21] McNulty, P. B. and Grace, P. M. (2009) "Agricultural Mechanization and Automation", In J. K. Schueller (Ed.), *Encyclopedia of Life Support Systems (EOLSS)*, Vol. I: Eolss Publishers Company Limited. E-ISBN 978-1-84826-096-2, ISBN 978-1-84826-546-2. DOI 10.13140/2.1.4421.9683.

[22] Nielsen, J. (1994a) "How to conduct a heuristic evaluation", Nielsen Norman Group, 1, pp. 1-8.

[23] Nielsen, J. (1994b) "*Usability engineering*", Morgan Kaufmann. ISBN 978-0125184069.

[24] Nielsen, J. (1994c) "Usability inspection methods", In *Conference companion on Human factors in computing systems*, pp. 413-414. DOI 10.1145/259963.260531.

[25] Nielsen, J. (1995) "Severity ratings for usability problems",*Papers and Essays*. [Online]. Available: http://www.useit.com/papers/heuristic/severityrating.html [Accessed: 23 June, 2020].

[26] Nielsen, J. and Landauer, T. K. (1993) "A mathematical model of the finding of usability problems", Paper presented at t*he Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems*, pp. 206-213. ISBN 978-0-89791-575-5. DOI 10.1145/169059.169166.

[27] Nielsen, J. and Molich, R. (1990) "Heuristic evaluation of user interfaces", Paper presented at *the Proceedings of the SIGCHI conference on Human factors in computing systems: Empowering people*. ISBN 978-1-4503-0228-9. DOI 10.1145/97243.97281.

[28] Oberti, R., Marchi, M., Tirelli, P., Calcante, A., Iriti, M., Tona, E., Hočevar, M., Baur, J., Pfaff, J., Schütz, Ch. and Ulbrich, H. (2016) "Selective spraying of grapevines for disease control using a modular agricultural robot", *Biosystems Engineering*, Vol. 146, pp. 203-215. ISSN 1537-5110. DOI 10.1016/j.biosystemseng.2015.12.004.

[29] Poole, A. and Linden, B. (2005) "Eye Tracking in Human-Computer Interaction and Usability Research: Current Status and Future", In C. Ghaoui (Ed.), *Encyclopedia of human computer interaction: IGI Global*. E-ISBN13: 9781591407980, ISBN13: 9781591405627, ISBN10: 1591405629. DOI 10.4018/978-1-59140-562-7.ch034.

[30] Sistler, F. E. (1987) "Robotics and intelligent machines in agriculture", *IEEE Journal of Robotics and Automation*, Vol. 3, No. 1, pp. 3-6. ISSN 0882-4967. DOI 10.1109/JRA.1987.1087074.

[31] Yanco, H. A. and Drury, J. (2004) "Classifying human-robot interaction: an updated taxonomy", In *2004 IEEE International Conference on Systems, Man and Cybernetics* (IEEE Cat. No. 04CH37583), Vol. 3, pp. 2841-2846. ISBN 0-7803-8566-7, ISSN 1062-922X. DOI 10.1109/ICSMC.2004.1400763.

Volume XII Number 3, 2020

Analysis of Online Consumer Behavior - Design of CRISP-DM Process Model

Emil Exenberger, Jozef Bucko

Department of Applied Mathematics and Business Informatics, Faculty of Economics, Technical University of Košice, Slovakia

Abstract

The basis of the modern marketing of a business entity is to know the behavior of its customers. Advanced artificial intelligence methods, such as data mining and machine learning methods, penetrate data analysis. The application of these methods is most appropriate in the case of online sales of any goods in large quantities and various industries. They are very often used in the sale of electronics, PCs or clothes. However, it is also possible to apply them to the agricultural industry, not only in B2C, but also in B2B in the sale of seeds, agricultural products, or agricultural machinery. Appropriate combinations of offers and knowledge of customers can bring the selling entity higher profits or competitive advantages. The main goal of our study is to design a CRISP-DM process model that will enable small businesses to analyze online customers' behavior. To reach the main goal we perform a data analysis of the online sales data by using machine learning methods as clustering, decision tree and association rules mining. After evaluating the proposed model, we discuss its use of the proposed model in the field of internet sales in the agricultural sector.

JEL Classification: C53, D11, D91, M21, M31

Keywords

Classification, association rules, data analysis, consumer behavior, online shopping.

Exenberger, E. and Bucko, J. (2020) “Analysis of Online Consumer Behavior - Design of CRISP-DM Process Model", *AGRIS on-line Papers in Economics and Informatics*, Vol. 12, No. 3, pp. 13-22. ISSN 1804-1930. DOI 10.7160/aol.2020.120302.

Introduction

In today's rapidly changing world, the information and knowledge gained and used are very valuable. With a high degree of computerization in all areas of our lives, the amount of data generated is also increasing, which is further stored and monitored. Today, the automotive, energy, engineering or agricultural industries are more focused on digital data analysis than ever before. The goal is to extract data from the amount of data that can be effectively transformed into information that is very important for understanding the trend of a business or a company. The gathering of such information is often done to support an entity's decision to move forward in the business process.

We can find many different approaches to knowledge discovery in databases (KDD) by various authors (Brachman and Anand, 1996; Fayyad, Piatetsky-Shapiro and Smyth, 1996; Klösgen and Zytkow, 2002; Mannila, 1997; Simoudis, 1996), but they all follow certain principles. In 1996, as part of an EU project, the basic principles of data mining were formulated as the Enterprise Standard Data Mining Process (CRISP-DM) (Smart Vision Europe, 2015). Business Intelligence (BI) is such data mining (DM) in the KDD process, the results of which are used to support business decision-making in the form of reports.

One area that has tremendous potential and which enables efficient real-time data collection is the sale of products through a website. In the case of online sales, the main task is to ensure customer satisfaction, minimize costs, maximize profits, or streamline processes. Through the acquired data and the application of appropriate machine learning methods, it is possible to analyze customer purchasing behavior.

Our contribution aims to propose a process model of CRISP-DM to analyze business data to support business decision making in the area of business processes. We will apply the proposed process model of CRISP-DM to specific data obtained from the online sale of electronic components of the international company SOS electronic to increase the effectiveness of the marketing strategy by creating marketing letters, which is expected to increase sales of these components. The main objective is to investigate the relationship between significant association rules and customer size in terms of the volume and financial amount of payments for orders in the online purchasing process.

Clustering is often used to segmentation mainly due to its high success rate, which was evaluated by many authors (Safri, Arifudin, and Muslim 2018; Prashar, Vijay, and Parsad 2018; Suchacka, Skolimowska-Kulig, and Potempa 2015; Keller, Gray and Givens, 1985). The basic idea of clustering is to make elements within one cluster as similar as possible, while the differences between elements from different clusters should be as large as possible. The advantage is regular updating of the training set after each addition of other elements, unlike other forms of classification and prediction based on the original, regularly unchanged training set. As the most commonly used clustering method is considered the k-means method, first used by MacQueen (1967). Prashar, Vijay and Parsad (2018) compare neural network prognostic ability, linear discriminant analysis, and k-means methods to reduce online retailers' vulnerability to market demand. At the end of this study, statistical evidence was provided on the accuracy of the predictions of the methods, with an accuracy of the k-nearest neighbor method of 79.1%. The problem of classifying two user sessions in an online store, a shopping session and a browsing session, was investigated in their work by Suchack, Skolimowska-Kulig and Potempa (2015). By comparing the results of several methods, they evaluated the k-nearest neighbor method (using Euclidean distance) as the most effective in terms of shopping forecasts and overall forecasts.

Decision trees (Quinlan 1986) belong to the basic machine learning classification methods and make it possible to classify data based on decision-making in response to individual tests. This is the most popular form of classifier representation. Komprdová et al. (2012) describe the CART algorithm as one of the best-known algorithms for creating decision trees, which is also a basic representative of binary trees, while also focusing on criterion statistics in particular for regression and classification trees. It explains the basic principles of decision tree creation because other binary trees can be obtained by modifying the rules of the CART tree. The decision tree classification method is most often used to analyze online customer behavior based on past customer behavior data on the seller's website. Sun, Cárdenas and Harrill (2016) present decision trees and software Weka as a new technique and tool that identifies critical attributes that affect the quality level of customer experience when visiting a travel agency website. Raj and Singh (2016) investigate how the demographic situation affects the frequency of online purchases, categorizing customers into three groups - often shoppers, frequent customers, and less frequent customers - using decision tree methods.

Products purchased either in-store or via online sales may have a connection. If the existence of one product affects the existence of another product in the same purchase, this relation expressed in the form of an implication constitutes an association rule (AR). This concept was first introduced by Agrawal, Imielinski and Swami (1993). The process of analyzing customers' shopping carts to find ARs that must meet predetermined conditions determining the significance of ARs is called association rules mining (ARM). ARM is currently used in various fields of research and analysis, such as language studies (Adamov, 2018), electronic transaction security (Askari, Md and Hussain 2020), medicine (Buczak et al., 2015; Soni et al., 2011; Luo et al., 2013), but often also in customer analysis (Kaur and Kang, 2016; Guo, Wang and Li, 2017). One of the most commonly used methods for ARM is the Apriori method, first described by Agrawal et al. (1994). Since then, the Apriori method has been improved to speed up the ARM process, e. g., improving operational efficiency by reducing the number of databases scans needed (Yuan 2017) or reducing the amount of operation needed (Wu et al. 2009). Other modifications were aimed at adapting the Apriori algorithm to real-time online consumer behavior analysis for specific businesses (Kaur and Kang 2016; Guo, Wang, and Li , 2017). For example, Alfian et al. (2019) propose a real-time analysis of consumer behavior for online commerce using ARM. For analysis, they use the proposed system to track customers, product browsing history, and transaction data from digital tagging.

Although ARM has appeared in many publications that analyzed consumer behavior in the purchasing process, it is more often used to track the purchasing process in a brick and mortar store and for specific goods (Avcilar and Yakut 2014; Chen et al. 2015). It is not used so often in the research of the online purchasing process and we have not found any publication on the research of the sale of agricultural goods by this method. This has become a fundamental motivation for our research.

Despite the widespread use of clustering (k-means), decision-tree (CART) and ARM (Apriori) methods, we did not encounter a combination of these and use them in the analysis of online consumer behavior. However, there have been authors who have proposed or used similar combinations to analyze consumer behavior. Kunjachan, Hareesh and Sreedevi (2018) use a combination of k-means, Apriori and Eclat methods to analyze large amounts of data in the form of online sales data. They recommend this methodology for easier analysis of consumer behavior and the mining of hidden data relationships. The methodology based on the combination of ARM and decision trees in Ma, Haiying and Dong Gang (2011) was to create a model for segmenting customers in an online environment. The benefit was to help managers understand customers, analyze the market and make business management decisions.

Our goal is to show a combination of machine learning methods for understanding the behavior of customers and a targeted offer of a combination of goods. We aim to show that the given method can also be used in agribusiness for the sale of various agricultural goods and products. When studying the current literature, we did not find much application of these methods in the agricultural industry. Their use can be monitored mostly in the field of sales of electronics, PCs and various clothing. Gandhi and Armstrong (2016) provide an overview of the data mining methods used in agriculture, mentioning techniques such as neural networks, bayesian networks and support vector machines. Santosh Kumar and Balakrishnan (2019) use the Apriori algorithm directly to recommend products to customers in the agricultural market. Clustering is also often used in the analysis of agricultural data (Shedthi et al. 2017; Zhao et al., 2009). However, none of these studies focuses on the use of these methods in real-time in the analysis of consumer behavior.

Materials and methods

Based on an analysis of current research in this area, we found that the most common method used to analyze an online purchasing process by customers is clustering. This method is natural because it allows you to create related customer groups and choose the appropriate marketing management for them. It is also possible to assume that there will be specific association rules for different customer clusters. Therefore, we naturally put forward a research hypothesis:

H1: There are specific association rules for each related customer group

In our research, we observed how often a customer who bought the items of interest also bought another in the same purchase. We were interested in the analyzed relationship only on the assumption that the occurrence of such (in our example) pairs was sufficient and would meet the predetermined conditions. When selecting a research site, it was a precondition that the company had the possibility of selling online, with the emphasis being on data being as robust as possible. The dataset represents data from one year of online purchases of the company that distributes electronic components. In total, we obtained data representing **185,706 purchases from 4,111 companies** in one year. Each row of the dataset contains:

- the id of the company that made the purchase;
- id of purchased goods;
- purchase date;
- number of purchased items;
- price per unit of goods.

To keep data anonymous, customer and item names have been replaced by identification numbers. We carried out the whole process of work in the R software environment for its free availability and many such packages that allow performing all the steps of our analysis. From the dataset, we naturally chose the financial volume of payments and the number of orders made as the decision criteria. This selection best describes the customer in terms of creating a marketing strategy and at the same time is possible from the Dataset point of view.

As a first step, we calculated the total number of purchases for each customer and the total amount spent by online purchases in the selected store during the year. Subsequently, we removed the outliers from these values, which could distort the result, while these outliers may be subject to further study. To do this, we use the command *boxplot(Data)$out* (Chambers, 2017; Becker, 2018; Murrell, 2018), which outputs represent the outliers found in the table Data.

In the second step, prepare the data for clustering by **standardizing** the data using the *robustHD package and the standardize()* command. With the NbClust package we find **the optimal number of clusters** (Charrad and Ghazzali, 2014).

It compares the optimal number of clusters detected using twenty-four different methods, evaluates the results and recommends the **optimal number of clusters**. This is then used for **cluster analysis** of k-means using the *kmeans()* (Hartigan and Wong, 1979) command, and the matching results are written to the table including original, non-normalized total amount and purchase data.

The third step is to generate the **decision tree** and the corresponding **classification rules** using *rpart* (Breiman, 2017) package. The columns *number of purchases* and *total amount* will be the input attributes to be tested and the column representing the value of the burst to which the customer belongs – column cluster - will be the target attributes of the decision tree. The result of this step will be a model whereby an enterprise can split new customers into existing clusters.

In the final step, we will be mining the **association rules** for each cluster separately through the *arulez* and *arulezviz* packages. Hahsler and Gruen used the original (Agrawal, Imieliński and Swami, 1993) and more modern (Lepping, 2018; Borgelt and Kruse, 2002; Borgelt, 2003) designs of this method to program the Apriori function into the R environment. First, we divide the Customer Data table into as many tables as there will be the resulting number of clusters, and write down only the rows that belong to them in each table. We will then create a list of transactions (shopping carts) that will include, within each transaction, a set of goods that were purchased in that transaction. Subsequently, we will mine the ARs in the clusters, assuming that one purchase will consist of the goods purchased by one customer in one day. The resulting ARs will be evaluated using the *itemset, frequency, support, confidence* and *lift indicators*.

Itemset make up the goods offered to SOS electronic through online sales. *Frequency (A)* represents the absolute abundance of product A in an *itemset* and *Support (A)* represents the relative abundance of occurrences of product *A* in an *itemset*. *Confidence (A=>B)* expresses the ratio of both *A* and B products in one purchase to those in which there is only product *B* (see equation 1).

$$Confidence(A => B) = \frac{frequency(A,B)}{frequency(A)} \quad (1)$$

It, therefore, determines how likely product *B* can be expected to be in a purchase that contains product *A*. However, let's have an example where product *B* is in all purchases and product A is only in 5% of purchases. The *confidence (A=>B)* would be equal to 1, but the predictive value of this AR would be low, because the presence of Product *A* in the purchase does not affect the occurrence of Product *B* in the same purchase. For this reason, the strength of AR is analyzed trough the *Lift (A=>B)* indicator by calculating how the presence of product *A* in the purchase affects the occurrence of product *B* in the same (see equation 2).

$$Lift(A => B) = \frac{Support(A,B)}{Support(A) * Support(B)} \quad (2)$$

In our example, where Product *B* was in all purchases, *Lift (A=>B)* would be equal to 1 indicated that the presence of Product *A* in the purchase does not affect the occurrence of Product *B* in the same purchase. The higher *Lift (A=>B)*, the more product *A* affects product *B* in the same buying and and the lower the *Lift* value is below one, the more product *A* is negatively affected by product *B* in the same purchase. The Apriori method searches for ARs that meet the conditions *min_confidence* and *min_support* (or absolute value *min_frequency*), which the user selects himself. In association rules mining we left the values of *arulez* package at the default settings of *min_frequency* = 7 and *min_confidence* = 1. The result will be a table that expresses the association rules and the associated *support, confidence, lift* and *count (frequency)* values.

Results and discussion

The analysis results consist of these parts:

- **customer clusters** based on the total amount they spent on product purchases and the total number of purchases made;
- **decision tree** as a model for assigning customers to clusters from the previous point;
- **classification rules** as a textual notation of individual branches of decision-making
- tree;
- **association rules** mined for every existing cluster separately.

The result of detecting the optimum number of bursts using the *NbClust()* command was that the best number of clusters is 3.

We then split the customers into three clusters with the *kmeans()* command.

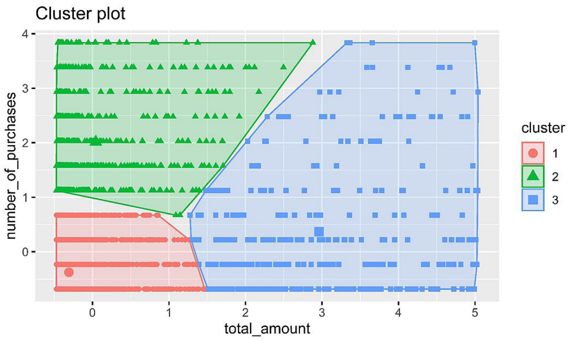
Table 1 shows the absolute and relative (rounded

	Cluster 1	Cluster 2	Cluster 3
Absolute abundance	3237	547	327
Relative abundance	78.74%	13.31%	7.95%

Source: prepared by authors

Table 1: Absolute and relative abundance of customers in clusters.

to 2 decimal places) abundance of customers in clusters.

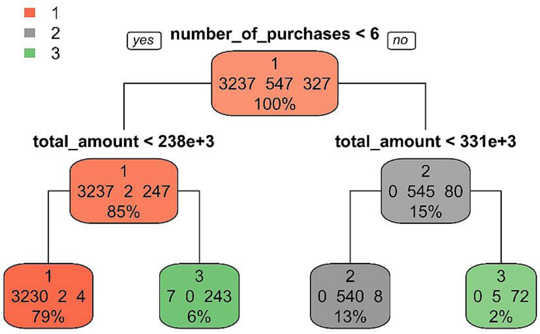


Source: prepared by authors

Figure 1: Clusters visualization.

A graphical representation of the assignment of customers to the clusters is shown in Figure 1. From the graphical evaluation that Cluster 1 consists of customers whose made a low number of purchases with a low total amount; Cluster 2 is made up of customers whose made more purchases with a low total amount; and there are customers with more purchases for a higher total amount in Cluster 1 in the whole range of the number of purchases. Cluster 1 has the highest absolute abundance because it includes customers who made small purchases at low prices. However, from Table 1 and Figure 1, We cannot determine how many purchases or the total amount constitute the boundaries between individual clusters. To identify these boundaries, we generate a CART decision tree for its simplicity resulting from its binary.

Using package rpart, we have generated a CART decision tree with an absolute abundance of customers in every step showed in Figure 2. The accuracy of classifying customers into individual clusters by using the generated decision tree is 99,37%. For comparison, we also generated C5.0 decision tree with higher accuracy (99.93%). On the other hand, C5.0 decision tree has 10 branches, which makes it more complicated for further use in comparison of 4 branches in CART decision tree and because of that we continued with CART decision tree.



Source: prepared by authors

Figure 2: Decision tree.

Using the same package rpart we get classification rules according to generated decision tree shown in Table 2. Lines end with symbol * represents final assignments customers to clusters expressed by the following listing.

step) test total unfit cluster relatives

*Example: 5) number_of_purchases>=16.5 72 13 3 (0.18 0 0.81) **

where:

- **step (5)** - decision tree branch number (in the example it is step no. 5);
- **test (number_of_purchases>=16.5)** - boolean test performed on the node;
- **total (72)** - total number of customers entering the test;
- **unfit (13)** - number of customers whose are differently assigned to cluster by model. In the - example, 13 customers were assigned to other cluster by k-means as it flows from the decision tree;
- **cluster (3)** - the number of the cluster to which the customer belongs if he passes the test;
- **relatives (0.18 0 0.81)** - relative abundance of unfit customers by clusters. In the example, 18% customers were assigned to cluster 1 by k-means, but to cluster 3 by a decision tree.

Table 3 represented mined association rules for cluster 1 and its values for each mined

1) root 4111 874 1 (0.7873996595 0.1330576502 0.0795426903)
2) number_of_purchases< 5.5 3486 249 1 (0.9285714286 0.0005737235 0.0708548480)
4) total_amount< 238490.7 3236 6 1 (0.9981458591 0.0006180470 0.0012360939) *
5) total_amount>=238490.7 250 7 3 (0.0280000000 0.0000000000 0.9720000000) *
3) number_of_purchases>=5.5 625 80 2 (0.0000000000 0.8720000000 0.1280000000)
6) total_amount< 330888.3 548 8 2 (0.0000000000 0.9854014599 0.0145985401) *
7) total_amount>=330888.3 77 5 3 (0.0000000000 0.0649350649 0.9350649351) *

Source: prepared by authors

Table 2: Classification rules.

lhs		rhs	support	confidence	lift	count
{75125}	=>	{75127}	0.001498041	1	578.53	13
{84025,84027}	=>	{84029}	0.001498041	1	542.38	13
{84029,84033}	=>	{84025}	0.001498041	1	510.47	13
{84031}	=>	{84029}	0.001382807	1	542.38	12

Source: prepared by authors

Table 3: First 4 association rules for Cluster 1 sorted by highest support value.

association rule. In the lhs (left hand side) column is the implication input; rhs (right hand side) is the implication output. In the first association rule of Table 3, we can interpret as: "If the customer bought the product 75125, he also bought the product 75127 with 100% probability (expressed by confidence = 1) and the rule has been applied in 13 cases".

Number of association rules mined for each cluster by k-means with same settings are:

- Cluster 1 = 92 association rules;
- Cluster 2 = 16 association rules;
- Cluster 3 = 12 association rules.

To find out if there are intersections between ARs between we tried to find them and compare the values of the ARs in case of a match. The intersection of ARs between clusters having identical lhs and rhs in both clusters was only C2 ∩ C3 = 1 AR. Despite the intersection, found AR have different values of indicators (support, confidence and lift) between clusters. Based on this, we conclude that there are specific association rules for each related customer group and because of that we do not reject the hypothesis.

The results enable business management to segment customers into clusters for a better understanding of their structure. Clusters can follow association rules that can be used to analyze consumer baskets or predict consumer behavior. First, the customer would be assigned to the cluster according to the decision tree or classification rules, and then it is possible to follow the association rules associated with the cluster. We used a CART decision tree for its simplicity to create a model to quickly assign new customers to the clusters. However, the user can choose different types of decision trees for analysis, depending on his preferences. The work can serve as a CRISP-DM process model that does not require additional application costs and is therefore also suitable for smaller businesses that have not yet analyzed consumer behavior in online sales.

Association rules can serve to **predict consumer behavior** in real-time. After logging into the site, the customer is immediately assigned to one of the clusters with assigned ARs. When a customer puts a product into his shopping cart, all significant ARs that have the product on the lhs side are searched. Those ARs that have as much confidence and support as possible can then be used to create targeted advertising to maximize the likelihood of success. It is confidence and support that talk about the extent to which the AR was valid and how often it occurred in the original data. Targeted advertising in this form can increase sales and business profits. KPIs can be used to monitor the success of the implementation of the proposed CRISP-DM process model, such as:

- KPI 1: Proportion of revenues and costs associated with the implementation of the proposed CRISP-DM process model;
- KPI 2: The number of recommended goods purchased by customers who put in their carts based on targeted advertising;
- KPI 3: Total revenue from recommended goods.

Generated by the ARs can be used for analysis of **frequently occurring customer baskets**. The essence is the analysis of relationships between individual products. An enterprise can analyze why the customer buys product B solely with product A, even though product B has several other alternatives. The results of such an analysis could be used in different ways, e.g. products A and B would be offered in a single package at a discount or the results would serve as an incentive to verify and improve the quality of alternative products B. Process changes based on these analyzes could either bring additional revenues to the business or lead to cost reductions - e.g. goods A and B would be packed together in a single package, thereby saving the company on additional packaging. In such an analysis, when selecting association rules, management should focus on the ARs with the highest Lift value, which evaluates the strength of the relation between the products in each ARs. Verifying the success of the implementation of changes based on shopping cart analysis can be evaluated using KPIs similar to the prediction of consumer behavior described above.

Our results show that the right combination of offers to the right group of customers can increase a company's profits for sales. Our ambition is to apply this combination of methods in an area where it is less common, such as the agro-business sector.

Conclusion

Consumer Behavior Analysis makes it possible to provide business decision-making information that contributes to the achievement of business goals, with the primary benefit being profit. This can be achieved either by increasing revenues or by reducing costs within individual business processes. In our paper, we design and provide a low-cost process model CRISP-DM, that allows both when properly used and implemented. The proposed CRISP-DM process model is based on segmenting existing customers into individual groups using the k-means method; creating a model for segmenting new customers through the decision tree and analyzing the shopping cart using the Apriori method. In the discussion, we present the possible use of results and ways of measuring the success of the alternatives. The article can also serve as a brief overview of the knowledge of machine learning methods today and their use in current scientific analysis.

Although we did our research on the data of a company that sells electronic components, we think that the proposed CRISP-DM model can also be used in the field of agriculture, where it would allow the analysis of consumer behavior, for example in the sale of seeds, agricultural products, or tools usable in the agro-industry. Unlike the other studies mentioned (Gandhi and Armstrong, 2016; Kumar and Balakrishnan, 2019; Shedthi et al., 2017; Zhao et al., 2009), we propose a model that uses a combination of multiple data mining methods and allows real-time research of consumer behavior analysis of sales of agricultural products in the environment of online sales.

In our further research, we will focus on extreme values in the form of customers, from whom we have abstracted the CRISP-DM process model we proposed. We are going to evaluate customers through a modern machine learning method RFM (recency, frequency and monetary value), which assesses the importance of customers based on how often they make purchases when they made their last purchase, and at what value they made during the reporting period. We will test the model in this paper as well as its other planned modifications directly in the environment of agricultural online sales. Segmenting customers based on multiple criteria could ultimately increase the success of the association rules and thus achieve business goals.

Acknowledgements

The research was realized within the national project “Decision Support Systems and Business Intelligence within Network Economy” (Contract No. 1/0201/19) funded by Grant Agency for Science; Ministry of Education, Science, Research and Sport of the Slovak Republic.

Corresponding authors
Emil Exenberger
Department of Applied Mathematics and Business Informatics, Faculty of Economics
Technical University of Košice, Němcovej 32, 040 01 Košice, Slovak Republic
E-mail: emil.exenberger@tuke.sk

References

[1] Adamov, Abzetdin Z. (2018) “Mining Term Association Rules from Unstructured Text in Azerbaijani Language”, *In 2018 IEEE 12th International Conference on Application of Information and Communication Technologies (AICT)*, IEEE, pp. 1-4. DOI 10.1109/ICAICT.2018.8747143.

[2] Alfian, G., Ijaz, M. F., Syafrudin, M., Syaekhoni, A., Fitriyani, N. L. and Rhee, J. (2019) “Customer Behavior Analysis Using Real-Time Data Processing: A Case Study of Digital Signage-Based Online Stores”, *Asia Pacific Journal of Marketing and Logistics,* Vol. 31, No. 1, pp. 265-290. ISSN 1355-5855. DOI 10.1108/APJML-03-2018-0088.

[3] Askari, S., Md. S. and Hussain, Md. A. (2020) “E-Transactional Fraud Detection Using Fuzzy Association Rule Mining”, *Proceedings of the 2nd International Conference on Information Systems & Management Science (ISMS) 2019,* Tripura University, Agartala, Tripura, India, 6 p.

[4] Avcilar, M. Y. and Emre, Y. (2014) “Association Rules in Data Mining: An Application on a Clothing and Accessory Specialty Store”, *Canadian Social Science*, Vol. 10, No. 3, pp. 75-83. E-ISSN 1923-6697, ISSN 1712-8056.

[5] Becker, R. A. (2018) "*The New S Language: A Programming Environment for Data Analysis and Graphics*", CRC Press. ISBN 053409192X. DOI 10.1201/9781351074988.

[6] Borgelt, Ch. and Kruse, R. (2002) “Induction of Association Rules: Apriori Implementation”, In: Härdle W., Rönz B. (eds) *Compstat*, Physica, Heidelberg. E-ISBN 978-3-642-57489-4. DOI 10.1007/978-3-642-57489-4_59.

[7] Borgelt, Ch. (2003) “Efficient Implementations of Apriori and Eclat”, In *Proc. 1st IEEE ICDM Workshop on Frequent Item Set Mining Implementations*, FIMI 2003, Melbourne, FL, CEUR Workshop Proceedings 90.

[8] Brachman, R. J. and Anand, T. (1996) “The Process of Knowledge Discovery in Databases”, In *Advances in Knowledge Discovery and Data Mining*, pp. 37-57. ISBN 9780262560979.

[9] Breiman, L. (2017) "Classification and Regression Trees", Routledge. ISBN 1138469521. DOI 10.1201/9781315139470.

[10] Buczak, A. L., Baugher, B., Guven, E., Ramac-Thomas, L. C., Elbert, Y., Babin, S. M. and Lewis, S. H. (2015) “Fuzzy Association Rule Mining and Classification for the Prediction of Malaria in South Korea”, *BMC Medical Informatics and Decision Making,* Vol. 15, No. 1, pp. 47. ISSN 1472-6947. DOI 10.1186/s12911-015-0170-6.

[11] Chambers, J. M. (2017) "*Graphical Methods for Data Analysis*", Chapman and Hall/CRC, 410 p. ISBN 9781315893204.

[12] Charrad, M. and Ghazzali, N., Boiteau, V. And Niknafs, A. (2014) “NbClust: An R Package for Determining the Relevant Number of Clusters in a Data Set”, *Journal of Statistical Software*. ISSN 1548-7660.

[13] Chen, Ch.-Ch., Huang, T.-Ch., Park, J. J. and Yen, N. Y. (2015) “Real-Time Smartphone Sensing and Recommendations towards Context-Awareness Shopping”, *Multimedia Systems*, Vol. 21, No. 1, pp. 61-72. E-ISSN 1432-1882, ISSN 0942-4962. DOI 10.1007/s00530-013-0348-7.

[14] Fayyad, U., Piatetsky-Shapiro, G. and Smyth, P. (1996) “The KDD Process for Extracting Useful Knowledge from Volumes of Data”, *Communications of the ACM,* Vol. 39, No. 11, pp. 27-34. E-ISSN 1557-7317, ISSN 0001-0782. DOI 10.1145/240455.240464.

[15] Gandhi, N. and Armstrong, L. J. (2016) December. A review of the application of data mining techniques for decision making in agriculture. In *2016 2nd International Conference on Contemporary Computing and Informatics (IC3I),* IEEE, pp. 1-6. DOI 10.1109/IC3I.2016.7917925.

[16] Guo, Y., Wang, M. and Li, X. (2017) “Application of an Improved Apriori Algorithm in a Mobile E-Commerce Recommendation System”, *Industrial Management & Data Systems,* Vol. 117, No. 2, pp. 287-303. ISSN 0263-5577. DOI 10.1108/IMDS-03-2016-0094.

[17] Hartigan, J. A. and Wong, M. A. (1979) "Algorithm AS 136: A k-Means Clustering Algorithm." *Journal of the Royal Statistical Society. Series C (Applied Statistics),* Vol. 28, No. 1, pp. 100-108. E-ISSN 14679876, ISSN 00359254. DOI 10.2307/2346830.

[18] Kaur, M. and Kang, S. (2016) "Market Basket Analysis: Identify the Changing Trends of Market Data Using Association Rule Mining", *Procedia Computer Science,* Vol. 85, pp. 78-85. ISSN 1877-0509. DOI 10.1016/j.procs.2016.05.180.

[19] Keller, J. M, Gray, M. R. and Givens, J. A. (1985) "A Fuzzy K-Nearest Neighbor Algorithm", *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC-15, No. 4, pp. 580-585. ISSN 21682216. DOI 10.1109/TSMC.1985.6313426.

[20] Klösgen, W. and Zytkow, J. M. (2002) "The Knowledge Discovery Process", In *Handbook of Data Mining and Knowledge Discovery*, 10-21 p. ISBN 978-0-387-09823-4.

[21] Komprdová, K. (2012) "Rozhodovací stromy a lesy", Akademické nakladatelství CERM, 98 p., ISBN 978-80-7204-785-7.

[22] Kumar, M. S. and Balakrishnan, K. (2019) "Development of a Model Recommender System for Agriculture Using Apriori Algorithm", In *Cognitive Informatics and Soft Computing,* pp. 153-163, Springer, Singapore. ISBN 978-981-15-1451-7.

[23] Kunjachan, H., Hareesh, M. J. and Sreedevi, K. M. (2018) "Recommendation Using Frequent Itemset Mining in Big Data", In *2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS)*, pp. 561–556. ISBN 9781538628430. DOI 10.1109/ICCONS.2018.8662905.

[24] Lepping, J. (2018) "*Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*". John Wiley and Sons Inc. E-ISSN 1942-4795.

[25] Luo, D., Xiao, Ch., Zheng, G., Sun, S., Wang, M., He, X. and Lu, A. (2013) "Searching Association Rules of Traditional Chinese Medicine on Ligusticum Wallichii by Text Mining", In *2013 IEEE International Conference on Bioinformatics and Biomedicine*, IEEE., pp. 162-167. ISBN 978-1-4799-1309-1. DOI 10.1109/BIBM.2013.6732664.

[26] Ma, Haiying, and Dong Gang. (2011) "Customer Segmentation for B2C E-Commerce Websites Based on the Generalized Association Rules and Decision Tree", In 2011 2nd International Conference on Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC), pp. 4600-4603, Piscataway, NJ : IEEE. ISBN 9781457705359. DOI 10.1109/AIMSEC.2011.6010255.

[27] MacQueen, J. (1967) "Some Methods for Classification and Analysis of Multivariate Observations", In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Oakland, CA, USA. Vol. 1, pp. 281-297.

[28] Mannila, H. (1997) "Methods and Problems in Data Mining", In *International Conference on Database Theory*, pp. 41-55. ISBN 3540622225. DOI 10.1007/3-540-62222-5_35.

[29] Murrell, P. (2018) "*R Graphics*", CRC Press.

[30] Parsad, Ch., Vijay, T. S. and Prashar, S. (2018) "Predicting Online Buying Behaviour-a Comparative Study Using Three Classifying Methods", *International Journal of Business Innovation and Research,* Vol. 15, No. 1, pp. 62-78. E-ISSN 1751-0260, ISSN 1751-0252. DOI 10.1504/IJBIR.2018.10009022.

[31] Quinlan, J. R. (1986) "Induction of Decision Trees", *Machine Learning,* Vol. 1, No. 1, pp. 81-106. E-ISSN 1573-0565, E-ISSN 0885-6125. DOI 10.1007/BF00116251.

[32] Rakesh, A. and Srikant, R. (1994) "Fast Algorithms for Mining Association Rules", In *VLDB '94: Proceedings of the 20th International Conference on Very Large Data Bases*, pp. 487-99.

[33] Rakesh, A., Imieliński, T. and Swami, A. (1993) "Mining Association Rules between Sets of Items in Large Databases", In *Acm Sigmod Record*, Vol. 22, pp. 207-16. DOI 10.1145/170036.170072.

[34] Sahil, R. and Singh, D. (2016) "Impact of Demographic Factors on Online Purchase Frequency - A Decision Tree Approach", In *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*. ISBN 9789380544205.

[35] Safri, Y. F., Arifudin, R. and Muslim, M. A. (2018) “K-Nearest Neighbor and Naive Bayes Classifier Algorithm in Determining The Classification of Healthy Card Indonesia Giving to The Poor”, *Scientific Journal of Informatics,* Vol. 5, No. 1, pp. 18. E-ISSN 2460-0040, ISSN 2407-7658. DOI 10.15294/sji.v5i1.12057.

[36] Shedthi, B. S., Shetty, S. and Siddappa, M. (2017) "Implementation and comparison of K-means and fuzzy C-means algorithms for agricultural data", In *2017 International Conference on Inventive Communication and Computational Technologies (ICICCT),* IEEE, pp. 105-108. ISBN 978-981-15-0146-3. DOI 10.1109/ICICCT.2017.7975168.

[37] Simoudis, E. (1996) “Reality Check for Data Mining”, *IEEE Expert*, Vol. 11, No. 5, pp. 26-33. ISSN 1541-1672. DOI 10.1109/64.539014.

[38] Smart Vision Europe (2015) “*About CRISP-DM*”, CRISPDM by Smart Vision Europe. [Online]. Available: http://crisp-dm.eu/home/about-crisp-dm/, http://crisp-dm.eu/home/about-crisp-dm/ [Accessed: 20 July, August].

[39] Soni, J., Ansari, U., Sharma, D. and Soni, S. (2011) “Predictive Data Mining for Medical Diagnosis: An Overview of Heart Disease Prediction”, *International Journal of Computer Applications*, Vol. 17, No. 8, pp. 43-48. ISSN 0975-8887. DOI 10.5120/2237-2860.

[40] Suchacka, G., Skolimowska-Kulig, M. and Potempa, A. (2015) “A K-Nearest Neighbors Method for Classifying User Sessions in e-Commerce Scenario”, *Journal of Telecommunications and Information Technology*, Vol. 3, pp. 64-69. E-ISSN 1899-8852, ISSN 1509-4553.

[41] Sun, P., Cárdenas, D. A. and Harrill, R. (2016) “Chinese Customers’ Evaluation of Travel Website Quality: A Decision-Tree Analysis”, *Journal of Hospitality Marketing & Management,* Vol. 25, No. 4, pp. 476-497. E-ISSN 1936-8631, ISSN 1936-8623. DOI 10.1080/19368623.2015.1037977.

[42] Wu, H., Lu, Z., Pan, L., Xu, R. and Jiang, W. (2009) “An Improved Apriori-Based Algorithm for Association Rules Mining”, In *2009 Sixth International Conference on Fuzzy Systems and Knowledge Discovery,* Tianjin, China, Vol. 6, pp. 51-55. DOI 10.1109/FSKD.2009.193.

[43] Xiuli, Y. (2017) “An Improved Apriori Algorithm for Mining Association Rules”, In *AIP Conference Proceedings*, AIP Publishing LLC. DOI 10.1063/1.4977361.

[44] Zhao, B., Song, Z., Mao, W., Mao, E. and Zhang, X., (2009) "Agriculture extra-green image segmentation based on particle swarm optimization and k-means clustering", *Transactions of the Chinese Society for Agricultural Machinery*, Vol. 40, No. 8, pp.166-169. ISSN 1000-1298.

Volume XII Number 3, 2020

Household Demand for Fruits and Vegetables in Rural and Urban South-Western Nigeria

Rebecca Ibe[1], Mohammed Rahji[2], Adetola Adeoti[2], Kemisola Adenegan[2]

[1] National Horticultural Research Institute, Jericho Reservation Area, Nigeria

[2] Department of Agricultural Economics, University of Ibadan, Nigeria

Abstract

In spite of the enormous benefits of fruits and vegetables, studies have shown that their consumption in Nigeria is far below the recommended daily intake therefore, this study investigated the factors influencing the demand for fruits and vegetables among households in rural and urban South-western Nigeria. Data were sourced from 152 rural and 259 urban households, respectively with the aid of a semi-structured questionnaire and were analysed using descriptive statistics and the quadratic almost ideal demand system model. Household size and location, sex and years of education of household heads influenced the demand for fruits and vegetables. Both rural and urban households considered the demand for fruits and vegetables to be luxury goods. Rural households were more responsive to changes in own-prices of fruits and vegetables than their urban counterpart. Fruits and vegetables were "net substitutes" in the rural and "complements" in the urban.

Keywords

Fruits, vegetables; household demand; demand elasticity; quadratic almost ideal demand system.

Ibe R., Rahji, M., Adeoti, A. and Adenegan, K. (2020) "Household Demand for Fruits and Vegetables in Rural and Urban South-Western Nigeria", *AGRIS on-line Papers in Economics and Informatics*, Vol. 12, No. 3, pp. 23-35. ISSN 1804-1930. DOI 10.7160/aol.2020.120303.

Introduction

Fruits and vegetables are rich in vitamins, minerals, and dietary fiber and are also low in calorie required for the normal functioning of human body (Uusiku et al., 2010). The micronutrients supplied by fruits and vegetables are also vital for the optimal functioning of the gastro-intestinal tract as they also enable the body to use other nutrients required for its normal function like energy from fats and carbohydrate (Banwat et al., 2012). Despite the nutritional importance of fruits and vegetables, their intake is still far below the minimum recommendation of the World Health Organization (WHO) of 400g per capita per day (Ruel et al., 2005; Banwat et al, 2012; Ogundari and Arifalo, 2013; Pem and Jeewon, 2015) in both developed and developing countries. It is estimated that people worldwide eat just 20-50% of the recommended minimum (Food and Agriculture Organization (FAO), 2006). According to a FAO report (2013), combined annual fruit and vegetable consumption in Africa is less than 100kg per person, which amounts to 250g per person per day. Increased consumption of fruits and vegetables has been recommended as a key component of a healthy diet for the prevention of non-communicable chronic diseases (Ishdorg et al., 2013; Wang et al., 2014). Low consumption of fruits and vegetables has also been ranked the sixth major risk factor for mortality in the world (WHO, 2012).

Furthermore, in spite of the FAO/WHO '5 A day' message for fruit and vegetable consumption, a large gap still exists between the recommended and actual intake and many people worldwide are not measuring up with the expected quantity or variety (Krebs-Smith et al., 2010). As reported in a data from the European Food Safety Authority (EFSA) database (2013) and the Global Environment Monitoring System (GEMS)/Food database (2015), average fruit and vegetable intake is not positively linked to the status of the country since greater consumption could be observed in developing countries such as Uganda and People's Republic of China (PR China) compared to developed countries such as Denmark, Germany, the United Kingdom (UK) and France. In the United States (US), mean fruit and vegetable intake was

189.30 g/day and 255 g/day respectively (Centre for Disease Control and Prevention, 2013). Fruits and vegetables were consumed in the amount of 146.81 g/day and 176.96 g/day respectively in Hong Kong accounting for a total of 324 g/day (The Chinese University of Hong Kong, 2010), while 209 g per day and 228.6 g/day fruit and vegetable intake was reported among adults (Mensik et al., 2013). In Malaysia, fruit and vegetable intake was 179 g/day and 133 g/day, respectively (Izzah et al., 2012).

In Nigeria, empirical evidences indicated that expenditure on fruits and vegetables has been very low. According to a report released by the National Bureau of Statistics (NBS) in 2012, only 1.39 % and 11.49 % of the total household food expenditure were spent on fruits and vegetables, respectively in 2009/10 compared to 20.99 % and 20.24 % for cereals and tubers respectively. When disaggregated further, rural households spent 1.34 %, 11.59 %, 22.58 % and 20.73 % of total household food budget on fruits, vegetables, cereals, tubers and plantain respectively, while in the urban 1.52 %, 11.23 %, 16.92 % and 18.97 % were expended on fruits, vegetables, cereals, tubers and plantain respectively. These figures indicate the fact that in Nigeria, expenditure on fruits and vegetables is low when compared with cereals and tubers. However, the major determinants for the low expenditure on fruits and vegetables in Nigeria have not been ascertained.

Although considerable number of researchers such as Akinleye, 2009; Muhammad-Lawal et al. (2011), Fashogbon and Oni (2013), Ogundari and Arifalo (2013), had examined the determinants of food demand/consumption in Nigeria however, studies on determinants of the demand for different groups of fruits and vegetables are scarce. This study also becomes appropriate from the methodological point of view because, most studies on the determinants of fruit and vegetable demand/consumption (Ruel et al., 2005; Yen and Tan, 2011; Ogunniyi et al., 2012; Ogundari and Arifalo, 2013; Ohen et al., 2014; made use of conventional food demand models such as single demand equations, linear expenditure systems, Rotterdam model, Transcendental logarithmic (Translog) model and the Almost Ideal Demand Systems (AIDS) for analysis. However, only few studies (such as Mutuc et al., 2007; Tey et al., 2009; Otunaiya and Shittu, 2014) have attempted using the Quadratic Almost Ideal Demand System (QUAIDS) to model demand for vegetables in developing countries. This study therefore extends the frontier of literature on factors influencing fruit and vegetable demand among households in Nigeria. Information generated from this study will assist food producers/marketers to identify target consumer groups of fruit and vegetable items as well as guide policy makers in formulating public health and dietary promoting programmes.

Materials and methods

Study area

The study was conducted in South-western Nigeria between April and May 2013. South-western Nigeria is one of the six geopolitical zones in Nigeria; it falls on latitude 6^0 to the North and 4^0 to the South and is marked by longitude 4^0 to the West and 6^0 to the East. The zone is made up of six states namely Ekiti, Lagos, Ogun, Ondo, Osun and Oyo with a total population of 27,581,992 and predominantly agrarian (National Population Commission (NPC), 2006).

Source of data

The study made use of primary data which was sourced from respondents with the aid of a semi-structured questionnaire administered to household heads in rural and urban South-western Nigeria between April and May 2013. Information elicited from the respondents (household head or spouse) includes household heads' demographic and socio-economic characteristics, and household's fruit and vegetable expenditure using seven days memory recall. The fruit and vegetable groups considered were leafy vegetables, fruit vegetables, root and bulb vegetables, herbaceous/perennial fruits and tree fruits. The classification was based on that of Matthew and Karikari (1990) and Pennington and Fisher (2009). The composition of the fruit and vegetable groups used for this study can be found in Appendix 1.

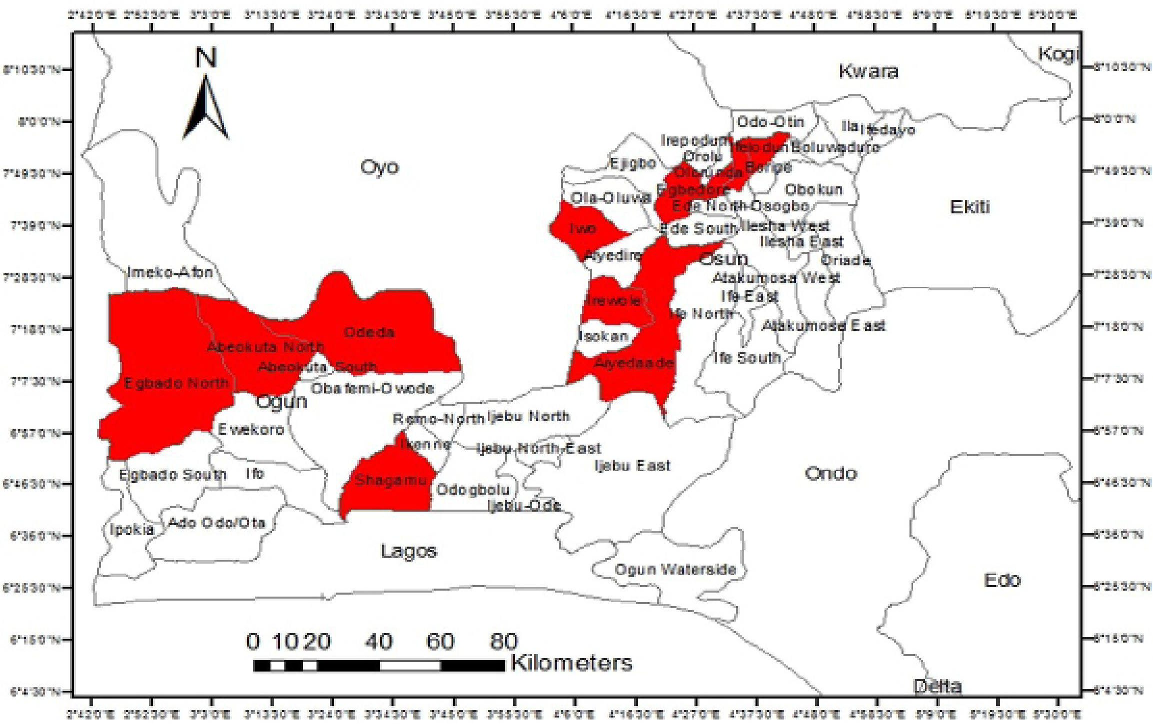
Sampling procedure

Data were collected from selected households in the study area between April and May 2013 using multi-stage sampling procedure. The first stage was the random selection of Ogun and Osun States out of the six states that constitutes the South-western Nigeria (see Figure 1) for the map of Nigeria showing the South-western zone and the selected states). The second stage was the random selection of 10 Local Government Areas (LGAs, 20 % of LGAs in the two States) out of the 50 LGAs (Ogun = 20; Osun = 30) that make up the two States due to paucity of fund to cover all the LGAs. The selection of LGAs per State was proportionate to the size of LGAs in each State. Four (4) and six (6) LGAs were selected from Ogun and Osun states, respectively because Osun has more LGAs (30) than Ogun (20) (NPC, 2006). The third stage was the random selection of one rural and one urban community from each of the selected LGAs. In the fourth stage, 450 households were randomly selected from the total number of households in the two States (1,611,283 households) proportionate to the number of households in each state (Ogun = 880,970; Osun = 730,313). Following this, a total of 246 and 204 households were selected from Ogun and Osun States, respectively (see appendix 2) based on the fact that Ogun has a higher number of households compared to Osun as indicated in the 2006 National Population Census figures of the Federal Republic of Nigeria (NPC, 2006). Sixty percent (60 %) of the households selected from each LGA were sampled from the urban communities, while 40 % was from the rural because it is believed that households are more in the urban centers due to prevailing rural-urban migration in Nigeria (Ogunmakinde et al., 2015). However, out of the 450 questionnaire administered, only 411 were found useful for analysis due to incomplete responses.

Estimation procedure

The study employed a two-step estimation procedure developed by Shonkwiler and Yen (1999) where the Inverse Mill's Ratio (IMR) for the various fruit and vegetable groups were first estimated from a Probit model and then incorporated into the demand equation. This estimation procedure takes into account correction for the possible sample selection bias created by the presence of zero expenditure for fruit and vegetable types reported by some households. In the first step, a Probit regression was estimated, which determines the probability that a given household purchased a fruit and vegetable item/group or not as observed in the studies of Mutuc et al. (2007) and Tey et al. (2009).

In the second step, expenditure share equations were specified as a quadratic extension of the Almost Ideal Demand System (QUAIDS) model with the IMRs included as instrumental variables. If the coefficient of IMR is statistically significant in the second step, then the sub-sample of purchasing households represents itself and excludes non-purchasers from the analysis. The estimated model therefore suffers from selectivity bias. However, if the coefficient of the IMR is insignificant, it means there is



Source: The author generated the map with the help of a cartographer. The figure is a map of the study area highlighting the locations where the survey was conducted

Figure 1: Map of South-western zone of Nigeria showing the selected LGAs.

no selectivity bias in the data. This means that the sub-sample of households purchasing fruits and vegetables are representative of the population.

The Quadratic Almost Ideal Demand System (QUAIDS) derived by Banks et al., (1996 and 1997) was employed to estimate price and income elasticities for the households as well as to determine the factors influencing households' expenditure on the fruit and vegetable items. QUAIDS is a rank three budget share system that is quadratic in the logarithm of total expenditure. It has the attractive property of allowing goods to have the characteristics of luxuries at low levels of total expenditure, and necessities at higher levels.

The QUAIDS which is derived from a generalization of the Price Independent Generalised Logarithmic (PIGLOG) preferences starts from an indirect utility function of the form:

$$\ln V = \left\{ \left[\frac{\ln m - \ln a(p)}{b(p)} \right]^{-1} + \lambda(p) \right\}^{-1} \quad (1)$$

Where the term [ln *m*-ln *a(p)*] is the indirect utility function of the PIGLOG demand system (that is, a system with budget shares linear in log total expenditure), *m* is the total expenditure, and *a(p)*, *b(p)* and λ(p) are functions of the vector of prices *p*. To ensure the homogeneity property of the indirect utility function, it is required that *a(p)* is homogenous of degree one in prices, and *b(p)* and *λ(p)* homogenous of degree zero in prices. The ln *a(p)* given in equation (1) has the Translog form:

$$\ln a(p) = \alpha_0 + \sum_{i=1}^{j} \alpha_i \ln p_i + \frac{1}{2} \sum_{i=1}^{j} \sum_{i=1}^{j} \gamma_{ij} \ln p_j \quad (2)$$

and *b(p)* is the simple Cobb-Douglas price aggregator defined as:

$$b(p) = \prod_{i=1}^{j} p_i^{\beta_i} \quad (3)$$

$$\lambda(p) = \sum_{i=1}^{K} \lambda_i \ln p_i, \text{ where } \sum_{i=1}^{K} \lambda_i = 0 \quad (4)$$

Where *i* = 1,…, K denotes the number of goods entering the demand model.

Application of Roy's identity or Shephard's Lemma to the indirect utility function (equation 4) gives the QUAIDS model expenditure share as:

$$\omega_i = \alpha_i + \sum_{j=1}^{K} \gamma_{ij} \ln p_j + \beta_i \ln\left[\frac{m}{a(p)}\right] + \frac{\lambda_i}{b(p)} \left\{ \ln\left[\frac{m}{a(p)}\right] \right\}^2 \quad (5)$$

To control for varying preference structures and heterogeneity across households, we incorporate demographic variables (*z*) into the QUAIDS model (equation 5) through the linear demographic translating method (Pollak and Wales, 1981). The expenditure share then becomes:

$$\omega_i = \alpha_i + \sum_{j=1}^{K} \gamma_{ij} \ln p_j + \beta_i \ln\left[\frac{m}{a(p)}\right] + \frac{\lambda_i}{b(p)} \left\{ \ln\left[\frac{m}{a(p)}\right] \right\}^2 + \sum_{s=1}^{L} \delta_{is} z_s + u_i \quad (6)$$

Where $z_s = (z_1, \ldots, z_L)$ is a set of demographic variables, w_i is the share of group expenditure allocated to product *i*, p_i is the price of product *i*, and *m* is the per capital expenditures on all commodities. Also, when the IMRs generated from the Probit model are incorporated into the expenditure share equation, it becomes:

$$\omega_i = \alpha_i + \sum_{j=1}^{K} \gamma_{ij} \ln p_j + \beta_i \ln\left[\frac{m}{a(p)}\right] + \frac{\lambda_i}{b(p)} \left\{ \ln\left[\frac{m}{a(p)}\right] \right\}^2 + \sum_{s=1}^{L} \delta_{is} z_s + \pi_i IMR_i + u_i \quad (7)$$

Where α_i, λ, β, γ are parameters to be estimated

α_i = average value of expenditure share in the absence of price and income effects.

β = parameter that determines whether goods are luxuries or necessities

γ_{ij} = effects on the budget of item *i* of 1 percent change in the prices of items in group *j*

P_j = price of item *j*

δ_j = vector of socioeconomic and demographic variables

π_i = vector of IMRs

u_i = error term

As with the original AIDS model, the theoretical restrictions of adding-up, homogeneity, and symmetry in the QUAIDS model are expressed in terms of its parameters. Adding-up requires $\Sigma_i w_i = 1$, and can be expressed in terms of model parameters as:

$$\sum_{i=1}^{k} \alpha_i = 1; \sum_{i=1}^{k} \beta_i = 0; \sum_{i=1}^{k} \lambda_i = 0; \sum_{i=1}^{k} \gamma_{ij} = 0 \quad \forall_j \quad (8)$$

Since Marshallian demands are homogenous of degree zero in (p, m),

$$\sum_{j=1}^{k} \gamma_{ij} = 0 \quad \forall_j \quad (9)$$

Slutsky symmetry implies that:

$$\gamma_{ij} = \gamma_{ji} \quad \forall_{i,j} \quad (10)$$

Expenditure share for each fruit and vegetable group was calculated thus:

$$W_G = \frac{X_G}{X} \text{ = the expenditure share of group G} \quad (11)$$

$$X_G = \sum p_{Gi} q_{Gi} \quad (12)$$

X_G = total expenditure of group G

P_{Gi} and q_{Gi} = the price and quantity of ith good in group G

G = 1 to 5 (see Table1)

STATA 13.1 software was used to run the QUAIDS. In order to avoid a singular covariance matrix, all the five expenditure share equations were specified, while the QUAIDS programme automatically omits one of the demand equations by itself and recovers it in the course of the analysis through the adding-up restrictions.

Elasticities were obtained by first differentiating the expenditure share equation (7) with respect to ln *m* and ln *pj* respectively to obtain:

$$\mu_i = \frac{\partial \omega_i}{\partial \ln x} = \beta_i + \frac{2\lambda_i}{b(p)}\left\{\ln\left[\frac{m}{a(p)}\right]\right\} \quad (13)$$

$$\mu_{ij} \equiv \frac{\partial \omega_i}{\partial \ln p_j} = \gamma_{ij} - \mu_i\left(\alpha_j + \sum_k \gamma_{jk} \ln p_k\right) - \frac{\lambda_i \beta_j}{b(p)}\left\{\ln\left[\frac{m}{a(p)}\right]\right\}^2 \quad (14)$$

The expenditure elasticity was derived as follows:

$$e_i = \frac{\mu_i}{\omega_i} + 1 \quad (15)$$

The uncompensated or Marshallian price elasticity is given by

$$e_{ij}^u = \frac{\mu}{w_i} - \delta_{ij} \quad (16)$$

where δ_{ij} is the Kronecker delta which is equal to one when $i = j$, otherwise $\delta_{ij} = 0$. Using the Slutsky equation, the compensated or Hicksian price elasticities was calculated and used to assess the symmetry and negativity conditions by examining the matrix with elements $w_i\ [e^c_{ij}]$ which should be symmetric and negative semi-definite in the usual way.

$$e_{ij}^c = e_{ij}^u + w_j e_i \quad (17)$$

Results and discussion

Zero expenditure

When modelling demand using micro data, it is typical to observe a significant number of households that purchase zero quantity of some of the items (or commodities) during the survey period (Bopape, 2006). A sample with a large number of non-purchasing households poses a number of econometric challenges. Table 1 reports the percentages of non-purchasing households vis a vis the fruit and vegetable groups. Herbaceous fruits were the least purchased by rural and urban households followed by tree fruits indicating that, households irrespective of their location tend to purchase more of vegetable items/categories than fruits. However, the values are higher for rural households compared to their urban counterparts indicating that, urban households purchase fruits and vegetables more than rural households.

Fruit and vegetable group	Rural (n=152) Z.E	Urban (n=259) Z.E
Leafy vegetables	25.7	14.3
Fruit vegetables	8.6	12.7
Root and bulb vegetables	30.3	29.7
Herbaceous fruits	61.2	42.5
Tree fruits	55.9	35.9

Note: Z.E stands for zero expenditure. Figures are measured in percentages.
Source: Field survey 2013

Table 1: Distribution of households with zero fruit and vegetable expenditure.

Budget shares, quantities and prices

Results in Table 2 showed that on the average, quantity of fruits and vegetables purchased by rural and urban households on monthly basis were far below the WHO recommendation of 12 kg per capita per month (400g per capita per day). It is also evident from Table 2 that, households spent the largest part of their fruit and vegetable budget on fruit vegetables, while the least was herbaceous and tree fruits in rural and urban, respectively. The reason may be that fruit vegetables were important items in the fruit and vegetable basket of the households. Furthermore, prices of most of the fruits and vegetable groups were higher in the urban than in the rural resulting in higher expenditure shares for fruits and vegetables among urban households. This result supports the findings of Abdulai and Aubert (2004) which states that, households in urban areas expend more on fruits and vegetables relative to their rural counterpart. Apart from higher prices, the reason could also be attributed to the fact that urban households understand the importance of fruits and vegetables in diets than rural households.

Fruit and vegetable group	Expenditure shares (%)		Quantity (Kg)		Price (₦/Kg)	
	Rural	Urban	Rural	Urban	Rural	Urban
LVs	17.0 (0.18)	18.8 (0.19)	5.31	5.47	78.65	89.59
t-value	-0.973					
FrVs	52.0 (0.27)	40.1 (0.24)	7.67	9.43	176.99	174.98
t-value	4.455***					
RBVs	14.5 (0.19)	15.3 (0.16)	7.00	5.46	67.93	69.58
t-value	-0.440					
HFs	7.8 (0.15)	13.6 (0.19)	0.51	2.48	87.48	82.86
t-value	-3.401***					
TFs	8.7 (0.16)	12.2 (0.18)	6.07	4.43	45.78	49.49
t-value	-1.969*					

Note: LVs = leafy vegetables; FrVs = fruit vegetables; RBVs = root and bulb vegetables; HFs = herbaceous fruits; TFs = tree fruits; standard deviations are in parenthesis; * and ***, represents 1% and 10% level of significant, respectively; ₦ is the Naira, Nigerian currency; US $ 1 = ₦ 158.27 at the time of the study

Source: Estimates of data analysis (descriptives)

Table 2: Expenditure shares, quantities and prices of fruits and vegetables.

Determinants of household demand for fruits and vegetables

The factors influencing households' demand for fruits and vegetables using the QUAIDS model are presented in Tables 3 and 4. The coefficient of $\ln EXPD^2$ for rural households (Table 3) is statistically significant for the fruit vegetables, root and bulb vegetables and tree fruits equations indicating that the response of their demand to increase in expenditure was non-linear. On the other hand, in the urban (Table 4) only the coefficient of $\ln EXPD^2$ for leafy vegetables, herbaceous fruits and tree fruits equations were statistically significant. Also, IMR was only significant in the budget share equation for leafy vegetables in the rural (Table 3) and that of leafy vegetables and fruit vegetables in the urban (Table 4) indicating that their inclusion was meaningful.

Considering the effect of household demographic and socio-economic variables on fruit and vegetable demand, results revealed that their effects varied. Results in Tables 3 and 4 reveals that prices of fruits and vegetables, access to credit and years of schooling of household head influenced household demand for fruits and vegetables in the rural, while the determinants in the urban were prices of fruits and vegetables, sex and years of education of household heads, household size and access to credit. For example, in the rural, years of schooling of household head had inverse relationship with household demand for tree fruits (Table 3) implying that households do not necessarily have to attain a higher education or access more credit before demanding for tree fruits. This suggests that, households might need more of nutrition education rather than formal education. On the other hand, sex of household head had positive influence on the demand for leafy vegetables among the urban households (Table 4). The implication is that male-headed households demand more of leafy vegetables which conforms to the findings of Nambiar et al., (2014) in Uganda that households with male heads spend less on fruit, compared to those with female heads.

Variable	Leafy vegetable	Fruit vegetable	Root and bulb vegetable	Herbaceous fruit	Tree fruit
Constant	0.2560* (0.1518)	0.8519*** (0.1246)	0.4445** (0.1902)	0.0885 (0.1170)	-0.6413*** (0.1379)
Price coefficients					
ln Price of LVs	0.0988** (0.0432)				
ln Price of FrVs	-0.0696 (0.0935)	0.5527*** (0.1975)			
ln Price of RBVs	0.0058 (0.0737)	0.2782*** (0.0943)	0.2045 (0.1334)		
ln Price of HFs	0.0159 (0.0326)	-0.1424* (0.0786)	-0.0617 (0.0617)	0.0239 (0.0403)	
ln Price of TFs	-0.0508 (0.1410)	-0.6178*** (0.1763)	-0.4268*** (0.1400)	0.1654 (0.1025)	0.9301*** (0.1743)
Expenditure and Expenditure squared					
lnEXPD	0.0564 (0.0805)	0.2850*** (0.0817)	0.2214*** (0.0768)	-0.0877 (0.0596)	-0.4750*** (0.0468)
lnEXPD^2	0.0028 (0.0089)	0.0320*** (0.0108)	0.0301*** (0.0077)	-0.0109 (0.0066)	-0.0540*** (0.0072)
Household characteristics					
SexHH	-0.0041 (0.0067)	-0.0039 (0.0091)	0.0070 (0.0078)	0.0032 (0.0055)	-0.0022 (0.0052)
HHsz	-0.0003 (0.0005)	-0.0001 (0.0007)	-0.0002 (0.0006)	0.0003 (0.0004)	0.0003 (0.0004)
Yeduc	0.0003 (0.0005)	0.0003 (0.0005)	0.0002 (0.0005)	0.0004 (0.0003)	-0.0005* (0.0003)
Accredit	-0.0035 (0.0046)	0.0108* (0.0062)	-0.0035 (0.0054)	-0.0055 (0.0037)	0.0018 (0.0034)
IMR	-0.0413*** (0.0134)	0.0121 (0.0181)	0.0184 (0.0157)	0.0099 (0.0110)	0.0009 (0.0099)

Note:***, **, * indicate level of significance at 1%, 5%, and 10% respectively. Standard errors are in parenthesis.
ln = Logarithm; LVs = leafy vegetable; FrVs = fruit vegetable; RBVs = root and bulb vegetable; HFs = herbaceous fruit; TFs = tree fruit; lnEXPD = logarithm of total fruit and vegetable expenditure; lnEXP^2 = square of logarithm of total fruit and vegetable expenditure; IMR = Inverse Mills Ratio; SexHH = sex of household head; HHsz = household size; Yeduc = years of schooling; Accredit = access to credit facility.
Source: Estimates from QUAIDS model

Table 3: Determinants of demand for fruits and vegetables (Rural).

Variable	Leafy vegetable	Fruit vegetable	Root and bulb vegetable	Herbaceous fruit	Tree fruit
Constant	0.2090 (0.2869)	1.1686*** (0.2938)	1.0426*** (0.2850)	0.7208** (0.3166)	-2.1412*** (0.1684)
Price coefficients					
ln Price of LVs	0.0987* (0.0565)				
ln Price of FrVs	-0.0445 (0.0426)	0.1314* (0.0698)			
ln Price of RBVs	0.0407 (0.0447)	0.0222 (0.0450)	0.1497* (0.0779)		
ln Price of HFs	0.0302 (0.0515)	0.0880* (0.0527)	0.0885 (0.0551)	0.1251 (0.1047)	
ln Price of TFs	-0.1251 (0.1433)	-0.1970 (0.1473)	-0.3012** (0.1289)	-0.3318** (0.1448)	0.9552*** (0.1672)

Note:***, **, * indicate level of significance at 1%, 5%, and 10% respectively. Standard errors are in parenthesis.
ln = Logarithm; LVs = leafy vegetable; FrVs = fruit vegetable; RBVs = root and bulb vegetable; HFs = herbaceous fruit; TFs = tree fruit; lnEXPD = logarithm of total fruit and vegetable expenditure; lnEXPD^2 = square of logarithm of total fruit and vegetable expenditure; IMR = Inverse Mills Ratio; SexHH = sex of household head; HHsz = household size; Yeduc = years of schooling; Accredit = access to credit facility.
Source: Estimates from QUAIDS model

Table 4: Determinants of demand for fruits and vegetables (Urban) (To be continued).

Variable	Leafy vegetable	Fruit vegetable	Root and bulb vegetable	Herbaceous fruit	Tree fruit
			Expenditure and Expenditure squared		
lnEXPD	0.0665 (0.0562)	0.0444 (0.0587)	0.1079** (0.0532)	0.1497**(0.0584)	-0.3686*** (0.0309)
$lnEXPD^2$	0.0054** (0.0025)	0.0002 (0.0026)	0.0025 (0.0024)	0.0067*** (0.0025)	-0.0149*** (0.0017)
			Household characteristics		
SexHH	0.0126** (0.0052)	-0.00680 (0.0066)	-0.0060 (0.0048)	-0.0076* (0.0040	0.0078 (0.0059)
HHsz	-0.0010** (0.0004)	0.0004 (0.0005)	0.0005 (0.0003)	0.0005 (0.0004)	-0.0004 (0.0003)
Yeduc	0.0001 (0.0002)	0.0003 (0.0002)	0.0001 (0.0002)	-0.0001 (0.0002)	-0.0003* (0.0002)
Accredit	0.0013 (0.0018)	0.0046** (0.0023)	-0.0002 (0.0016)	-0.0022 (0.0019)	-0.0034** (0.0015)
IMR	-0.0157* (0.0090)	0.0341*** (0.0112)	-0.0013 (0.0077)	-0.0137 (0.0095)	-0.0034 (0.0069)

Note:***, **, * indicate level of significance at 1%, 5%, and 10% respectively. Standard errors are in parenthesis.
ln = Logarithm; LVs = leafy vegetable; FrVs = fruit vegetable; RBVs = root and bulb vegetable; HFs = herbaceous fruit; TFs = tree fruit; lnEXPD = logarithm of total fruit and vegetable expenditure; $lnEXPD^2$ = square of logarithm of total fruit and vegetable expenditure; IMR = Inverse Mills Ratio; SexHH = sex of household head; HHsz = household size; Yeduc = years of schooling; Accredit = access to credit facility.
Source: Estimates from QUAIDS model

Table 4: Determinants of demand for fruits and vegetables (Urban) (Continuation).

Expenditure elasticity of demand for fruits and vegetables

Expenditure elasticity captures the percentage change in the quantity demanded of a good with respect to a percentage change in total expenditure (Tomek and Robinson, 2003). The elasticities were estimated at the sample means for the fruit and vegetable groups considered in this study and were reported in the context of economic theory, where an expenditure elasticity in the range of zero and one signifies a commodity to be a 'necessary good', greater than one implies a 'luxury good', while less than zero (negative) implies an 'inferior good' (Adegeye and Dittoh, 1985). As shown in Table 5, expenditure elasticities for rural and urban households revealed a variation in their expenditure behaviour. For some fruit and vegetable groups, the difference in the estimated expenditure elasticities between rural and urban samples was quite substantial. For example, a 1 percent increase in expenditure will lead to a greater than proportionate increase (3.25 and 1.59, respectively) in the quantity of root and bulb vegetables demanded by rural and urban households, while tree fruits was found to be negative for both rural and urban households, indicating that they are inferior goods (less than zero).

Fruit and vegetable group	Expenditure elasticity	
	Rural	**Urban**
Leafy vegetable	1.1676	1.1593
Fruit vegetable	1.7456	1.1814
Root and bulb vegetable	3.2521	1.5908
Herbaceous fruit	0.3665	1.6853
Trees fruit	-6.2779	-1.3452

Source: Estimates from the QUAIDS model

Table 5: Expenditure elasticities of demand for fruits and vegetables.

Compensated and uncompensated own- and cross-price elasticities of demand

The uncompensated demand also known as the Marshallian demand represents the conventional market or individual demand. It implicitly combines both income and substitution effects and in this case, income is held constant while an alternative approach to demand is known as the compensated or Hicksian demand which compensates a consumer for price change. That is, when the price changes, consumers receive compensation that allows them to remain on their original indifference curve. Generally, Marshallian estimates provide better measure of the responsiveness for any particular good to changes in its own price than to the changes in the price of other goods (Alam, 2011). The Hicksian elasticities is known to provide better estimates of substitution effects between two food categories, devoid of the income effect (Bundi et al., 2013).

Compensated and uncompensated own- and cross-price elasticities of demand are presented in Tables 6 and 7. The results in Table 6 and 7 revealed that all the uncompensated and compensated own-price elasticities, as shown in the diagonal matrix displayed appropriate negative signs, indicating the inverse relationship between prices of a commodity and its demand. This finding is consistent with economic theory. The uncompensated own-price elasticity estimates for rural households in Table 6 revealed that leafy vegetables and herbaceous fruits were price-inelastic in absolute terms (less than one), implying that their demand was less responsive to changes in own-price. The uncompensated own-price elasticity for other groups was price-elastic, while in the urban all fruit and vegetable groups were price-inelastic (Table 7). The implication of these results is that, demand for fruits and vegetables are more responsive to changes in own-prices in the rural and lesser in the urban. On the other hand, the uncompensated cross-price elasticities are mostly positive (Table 6) in the rural area indicating that most of the fruit and vegetable groups considered were net substitutes, however the reverse was the case among urban households (Table 7) indicating complementary relationships. Similarly, the compensated cross-price elasticities for rural and urban households were mostly positive (Table 6 and 7) indicating that most of the fruits groups were net substitutes in both rural and urban areas.

Fruit and vegetable group	Fruit and vegetable group				
	LVs	**FrVs**	**RBVs**	**HFs**	**TFs**
			Marshallian/uncompensated elasticity		
LVs	**-0.4610**	-0.7184	-0.1417	0.1190	0.0237
FrVs	-0.3209	**-1.2797**	-0.2260	-0.1676	0.2081
RBVs	-0.5305	-2.1907	**-1.9263**	-0.0891	1.7400
HFs	0.5525	0.6961	0.6544	**-0.9051**	-0.5394
TFs	1.7000	6.0734	2.1500	0.8297	**-4.0368**
			Hicksian/compensated elasticity		
LVs	**-0.2627**	-0.1113	0.0273	0.2103	0.1256
FrVs	-0.0244	**-0.3722**	0.0266	-0.0311	0.3605
RBVs	0.0219	-0.4999	**-1.4557**	0.1653	1.1300
HFs	0.4903	0.5055	0.6014	**-0.9337**	-0.5714
TFs	0.1806	2.9400	1.3000	0.3386	**-4.5847**

Note: LV = leafy vegetable, FrV = fruit vegetables, RBVs = root and bulb vegetables, HFs = herbaceous fruits, TFs = tree fruits
Source: Estimates from QUAIDS model

Table 6: Own- and cross-price elasticities of fruit and vegetables (Rural).

Fruit and vegetable group	Fruit and vegetable group				
	LVs	**FrVs**	**RBVs**	**HFs**	**TFs**
			Marshallian/uncompensated elasticity		
LVs	**-0.5331**	-0.4855	-0.0129	-0.0339	-0.1035
FrVs	-0.1556	**-0.9196**	-0.1620	0.0472	0.0083
RBVs	0.1125	-0.6755	**-0.7461**	-0.0028	-0.2844
HFs	0.0232	-0.3396	-0.2346	**-0.8023**	-0.3487
TFs	-0.3754	1.7079	0.4967	-0.3196	**-0.1234**
			Hicksian/compensated elasticity		
LVs	**-0.3149**	-0.0203	0.1640	0.1237	0.0379
FrVs	0.0669	**-0.4454**	0.0182	0.2077	0.1524
RBVs	0.4120	-0.0370	**-0.5034**	0.2133	-0.0904
HFs	0.3405	0.3368	0.0225	**-0.5733**	-0.1432
TFs	-0.6286	1.1680	0.2915	-0.5024	**-0.2875**

Note: LV = leafy vegetable, FrV = fruit vegetables, RBVs = root and bulb vegetables, HFs = herbaceous fruits, TFs = tree fruits
Source: Estimates from QUAIDS model

Table 7: Own- and cross-price elasticities of fruit and vegetables (Urban).

Conclusion

This study addressed the question whether there are any disparities in fruit and vegetable expenditure behaviour between rural and urban households in Southwestern Nigeria using primary data collected between April and May, 2013. In order to correct for zero observations, a two-step estimation procedure developed by Shonkwiler and Yen (1999) was adopted. A Probit regression was estimated in the first stage, while the quadratic almost ideal demand system model was estimated in the second stage. Results show that the quantity of fruits and vegetables demanded by households are far below the WHO's minimum recommendation. Also, socio-demographic characteristics such as prices of fruits and vegetables, and household heads' years of education and access to credit play important roles in rural households' demand for fruits and vegetables. On the other hand, urban households' demand for fruits and vegetables are influenced by household size and household heads' sex, years of education and access to credit. All the fruits and vegetables were considered luxury goods among urban households, while it is a mixture of luxury and necessity goods in the rural. Tree fruits were considered 'inferior goods' in both rural and urban households. Rural households were more responsive to changes in own-price than their urban counterparts, while most of the fruits and vegetables were net substitutes among rural and urban households.

Corresponding authors
Dr. Rebecca Ibe
National Horticultural Research Institute, Jericho Reservation Area
Idi-Ishin P.M.B. 5432, Dugbe Post office Ibadan, Nigeria
Phone: 2348035850487, E-mail: beckkyibe2015@gmail.com

References

[1] Abdulai, A. and Aubert, D. (2004) "A cross-section analysis of household demand for food and nutrients in Tanzania", *Agricultural Economic*, Vol. 31, No.1, pp. 67-79. E-ISSN 1805-9295, ISSN 0139-570X. DOI 10.1111/j.1574-0862.2004.tb00222.x.

[2] Adegeye, A. J. and Dittoh J. S. (1985) "*Essentials of Agricultural Economics*", University of Ibadan, Impact Publishers Nig. Ltd., Ibadan, 251 p. ISBN 9782386006.

[3] Akinleye, S. O. (2009) "Food demand in Northern Nigeria: Implications for food policy", *Journal of Social Sciences"*, Vol. 18, No. 3, pp. 209-215. ISSN 2668-7798. DOI 10.1080/09718923.2009.11892683.

[4] Alam, M. D. A. (2011) "An analysis of consumption demand of elasticity and supply response of major food grains in Bangladesh", Thesis to meet requirements for the joint academic degree of International Master of Science in Rural Development from Ghent University (Belgium), Agrocampus Ouest (France), Humboldt University of Berlin (Germany), Slovak University of Agriculture in Nitra (Slovakia) and University of Pisa (Italy) in collaboration with Wageningen University (The Netherlands).

[5] Banks, J., Blundell, R. and Lewbel, A. (1996) "Tax Reform and Welfare Measurement: Do We Need Demand System Estimates?", *The Economic Journal*, Vol. 106, No. 438, pp. 1227-1241. E-ISSN 1468-0297. DOI 10.2307/2235517.

[6] Banks, J., Blundell, R. and Lewbel, A. (1997) "Quadratic Engel Curves and Consumer demand", *The Review of Economics and Statistics*, Vol. 79, No. 4, pp. 527-539. E-ISSN 1530-9142, ISSN 0034-6535. DOI 10.1162/003465397557015.

[7] Banwat, M. E., Lar, L. A. Daber, J., Audo, S. and Lassa, S. (2012) "Knowledge and intake of fruit and vegetable consumption among adult in urban community in North Central, Nigeria", *The Nigeria Health Journal*, Vol. 12, No. 1, pp. 12-15. ISSN 1597-4292.

[8] Bundi, M. K., Nzuma, J. and Mbatia, O. L. E. (2013) "Urban demand for smaller crops: The case of fruits and vegetables in Nairobi, Kenya", paper at the 4th International conference of the African Association of Agricultural Economists (AAAE), September 22-25, Hammamet, Tunisia, pp. 23. [Online]. Available: https://ideas.repec.org>ags>aaae13 [Accessed: 13 May 2015].

[9] European Food Safety Authority (2013) "The EFSA Comprehensive European Food Consumption Database". [Online]. Available: http://www.efsa.europa.eu/en/datexfooddb/datexfooddbspecificdata.htm [Aceesed: 22 June 2017].

[10] FAO (2006) “Food Security”, In *FAO Policy Brief: Food and Agriculture Organization of the United Nations*. [Online]. Available: http://www.fao.org/forestry/13128-0e6f36f27e0091055bec28ebe830f46b3.pdf [Accessed: 13 Dec. 2017].

[11] FAO (2013) “*World Food and Agriculture*”, Statistical Yearbook 2013, Food and Agriculture Organization of the United Nations, Rome, Italy. [Online]. Available: http://www.fao.org/docrep/018/i3107e/i3107e00.htm [Accessed: 13 Dec. 201].

[12] Fashogbon, A. E. and Oni, O. A. (2013) “Heterogeneity in Rural Household Food Demand and Its Determinants in Ondo State, Nigeria: An Application of Quadratic Almost Ideal Demand System”, *Journal of Agricultural Science*, Vol. 5, No. 2, pp. 169-177. ISSN 1916-9752, E-ISSN 1916-9760. DOI 10.5539/jas.v5n2p169.

[13] Global Environment Monitoring System (GEMS/ Food) (2012) World Health Organisation, 2015. [Online]. Available: http://www.who.int/nutrition/landscape_analysis/nlis_gem_food/en/ [Accessed: 22 June 2017].

[14] Izzah, N. A., Aminah, A., Pauzi, M. A., Lee, Y. H., Rozita, W. M. and Fatima, S. D. (2012) "Patterns of fruits and vegetable consumption among adults of different ethnics in Selangor, Malaysia", *International Food Research Journal*, Vol. 19, No. 3, pp. 1095-1107. E-ISSN 2231 7546.

[15] Matthew, I. P. and Karikari, S. K. (1990) “*Horticulture principles and practices*”, Macmillan Intermediate Agriculture series (General editor: Ochapa O. Onazi), Macmillan Education Ltd., London. ISBN 0333453069.

[16] Muhammad-Lawal, A., Kolawole, F. L., Balogun M. A. and Jamiu, A. A. (2011) “Food demand analysis among rural households in Oyo state, Nigeria”, *International Journal of Food, Agriculture and Veterinary Sciences*, Vol. 1, No. 1, pp. 45-50. ISSN 2277-209X.

[17] Mutuc, M. E. M., Pan, S. and Rejesus, R. M. (2007) “Household Vegetable Demand in the Philippines: Is There an Urban-Rural Divide?”, *Agribusiness*, Vol. 23, No. 4, pp. 511-527. E-ISSN 1520-6297. DOI 10.1002/agr.20137.

[18] Nambiar, P. M., Florkowski, W. J., Chinnan, M. S. and Ressurrecion, A. V. A. (2014) "Factors driving fruit and vegetable expenditures and consumption frequency in lesser Developed Country: An analysis of urban households from the Republic of Uganda", Selected paper for presentation at the Southern Agricultural Economics Association Annual Meeting (SAEA), Dallas, TX, 1-4 February, 30 p.

[19] National Bureau of Statistics (NBS) (2012) “*Consumption pattern in Nigeria 2009/10 preliminary report*”, [Online]. Available: http://www.nigerianstat.gov.ng/pdfuploads/Consumption%20Pattern%20in%20Nigeria%202009-10.pdf [Accessed: 13 May 2014].

[20] National Population Commission (NPC) (2006) “*2006 Population and housing census of the Federal Republic of Nigeria*”, Priority tables, Vol. I & II. [Online]. Available: http://catalog.ihsn.org/index.php/catalog/3340/download/48519; http://catalog.ihsn.org/index.php/catalog/3340/download/48520. [Accessed: 17 Nov. 2012].

[21] Ogundari, K. and Arifalo, S. F. (2013) “Determinants of household demand for fresh fruit and vegetable in Nigeria: A double hurdle approach”, *Quarterly Journal of International Agriculture*, Vol. 52, No. 3, pp. 199-216. ISSN 00498599. DOI 10.22004/ag.econ.173646.

[22] Ogunmakinde, O. E. Oladokun, S. D. and Oke, O. E. (2015) "Rural-Urban Migration in South Western Nigeria: A Menace to National Development", *Civil and Environmental Research*, Vol. 7, No. 5, pp. 42-47. ISSN 2225-0514.

[23] Ogunniyi, L. T., Ajao, A. O. and Oladejo, J. A. (2012) "Food consumption patterns in Ogbomoso Metropolis of Oyo State, Nigeria", *Journal of Agriculture and Social Research*, Vol. 12, No. 1, pp. 74-83. E-ISSN 1595-7470.

[24] Ohen, S. B., Umeze, G. E. and Inyang, E. O. (2014) "Consumer Purchasing Behaviour for Fruits and Vegetables among Civil Servants in Essien Udim Local Government Area, Akwa Ibom State, Nigeria", *Food Science and Quality Management*, Vol. 23, pp. 55-64. ISSN 2225-0557.

[25] Otunaiya, A.O. and Shittu, A. M. (2014) "Complete household demand system for vegetables in Ogun State, Nigeria", *Agric. Econ. – Czech*, Vol. 60, No. 11, pp. 509–516. E-ISSN 1805-9295, ISSN 0139570X. DOI 10.17221/46/2014-AGRICECON.

[26] Pem, D. and Jeewon, R. (2015) "Fruit and vegetable intake: benefits and progress of nutrition education interventions- narrative review article", *Iran Journal of Public Health*, Vol. 44, No. 10, pp. 1309-1321. E-ISSN 22516093, ISSN 22516085.

[27] Pennington, J. A. T. and Fisher, R. A. (2009) "Classification of fruits and vegetables", *Journal of Food Composition and Analysis*, Vol. 22: Supplement, S23–S31. ISSN 0889-1575. DOI 10.1016/j.jfca.2008.11.012.

[28] Pollak, R. A. and Wales, T. J. (1981) "Demographic variables in demand analysis", *Econometrica*, Vol. 49, No. 6, pp. 1533-1551. E-ISSN 1468-0262. DOI 10.2307/1911416.

[29] Ruel, M. T., Minot, N. and Smith, L. (2005) "Patterns and determinants of fruit and vegetable consumption in Sub-Saharan Africa: a multi-country comparison", the Joint FAO/WHO Workshop on Fruit and vegetable for Health, 1-3 September 2004, Kobe, Japan. [Online]. Available: http://cdrwww.who.int/dietphysicalactivity/publications/f&v_africa_economics.pdf [Accessed: 13 Dec. 2017].

[30] Shonkwiler, J. S. and Yen, S. T. (1999) "Two-Step Estimation of a Censored System of Equations", *American Journal of Agricultural Economics,* Vol. 81, No. 4, pp. 972-982. E-ISSN 1467-8276. DOI 10.2307/1244339.

[31] Tey, Y. S., Shamsudin, M. N., Mohamed, Z., Abdullah, A. M. and Radam, A. (2009) "Analysis of demand for vegetable in Malaysia", MPRA paper No.15033, University Library of Munich, Germany. [Online]. Available: http://mpra.ub.uni-muenchen.de/15033 [Accessed: 18 Oct. 2014].

[32] Uusiku, N. P., Oelofse, A., Duodu, K. G., Bester, M. J. and Faber, M. (2010) "Nutritional value of leafy vegetables of sub-Saharan Africa and their potential contribution to human health: A review", *Journal of Food Composition and Analysis*, Vol. 23, No. 6, pp. 499-509. ISSN 0889-1575. DOI 10.1016/j.jfca.2010.05.002.

[33] Wang, X., Ouyang, Y., Liu, J., Zhu, M., Zhao, G., Bao, W. and Hu, F. B. (2014) "Fruit and Vegetable Consumption and Mortality from All Causes, Cardiovascular Disease and Cancer: Systematic Review and Dose Response Meta-Analysis of Prospective Cohort Studies", *The British Medical Journal*, 349, g 4490. ISSN 1756-1833. DOI 10.1136/bmj.g4490.

Appendix

Category	Disaggregates of fruits and vegetables
Leafy vegetables (LVs)	Green amaranth, slim amaranth, celosia, corchorus, and Fluted pumpkin
Fruit vegetables (FrVs)	Tomato, okra, bell pepper, red pepper, and egusi melon
Root and Bulb vegetables (RtBub)	Onion
Herbaceous/perennialfruits (HerFr)	Banana and pineapple
Tree fruits (TrFr)	Citrus and mango

Source: Compiled by author based on the classifications of Matthew and Karikari (1999) and Pennington and Fischer (2009)

Table A.1: Construction of fruit and vegetable categories.

Segmentation of Bean-Plants Using Clustering Algorithms

Serkan Kartal[1], Sunita Choudhary[2], Michal Stočes[1], Pavel Šimek[1], Tomáš Vokoun[1], Vojtěch Novák[1]

[1] Department of Information Technologies, Faculty of Economics and Management, Czech University of Life Sciences in Prague, Czech Republic

[2] International Crops Research Institute for Semi-arid Tropics (ICRISAT), Patancheru, India

Abstract

In recent years laser scanning platforms have been proven to be a helpful tool for plants traits analysing in agricultural applications. Three-dimensional high throughput plant scanning platforms provide an opportunity to measure phenotypic traits which can be highly useful to plant breeders. But the measurement of phenotypic traits is still carried out with labor-intensive manual observations. Thanks to the computer vision techniques, these observations can be supported with effective and efficient plant phenotyping solutions. However, since the leaves and branches of some plant types overlap with other plants nearby after a certain period of time, it becomes challenging to obtain the phenotypical properties of a single plant. In this study, it is aimed to separate bean plants from each other by using common clustering algorithms and make them suitable for trait extractions. K-means, Hierarchical and Gaussian mixtures clustering algorithms were applied to segment overlapping beans. The experimental results show that K-means clustering is more robust and faster than the others.

Keywords

Clustering, segmentation, point cloud, sustainable agriculture.

Kartal, S., Choudhary, S., Stočes, M., Šimek, P., Vokoun, T. and Novák, V. (2020) “Segmentation of Bean-Plants Using Clustering Algorithms", *AGRIS on-line Papers in Economics and Informatics*, Vol. 12, No. 3, pp. 36-43. ISSN 1804-1930. DOI 10.7160/aol.2020.120304.

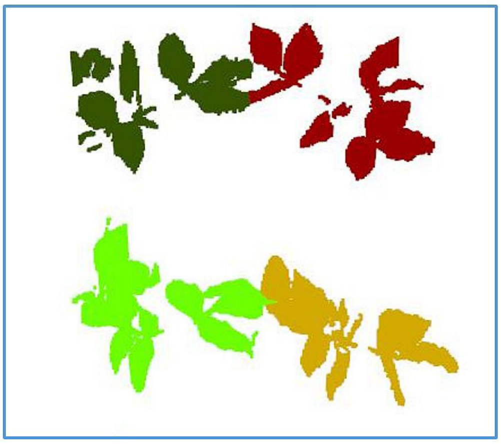
Introduction

Today, in parallel with the continuous growth of the global population (Kitzes et al., 2008), a continuous increase in food production must be achieved. In order to achieve this increase in food production, agricultural ecosystems need to be better understood by monitoring and analysing continuously. This process requires the analysis of huge agricultural data and the use of new information technologies (Kamilaris et al., 2016). This observation and analysis process can be examined under two headings; large scale observation and small-scale observation. Large-scale observation is performed by remote sensing vehicles such as satellites and drones which provide wide-view snapshots of the agricultural environments (Domínguez et al., 2015; Kumhálová et al., 2014, Kumhálová et al., 2013; Pantazi et al., 2016). When this system is applied to agriculture, it provides many advantages in collecting information on soil properties and cultivation areas. On the other hand, with small-scale observations, more detailed and sensitive information, such as disease detection, plant count and quality control are collected (Khan et al., 2018; Ramos et al., 2017; Suh et al., 2018; Zhang et al., 2018).

Collecting and reporting the data of plants in the growth phase is important factor to ensure agricultural sustainability and increase productivity (Raišienė et al., 2019; Muhammad et al., 2019). This process can help to analyse the relationship between plants and environmental factors so as to provide effective agricultural guidance (Bannayan and Sanjani, 2011; Jannoura et al., 2015; Gebbers and Adamchuk, 2010). In addition, having information about crop counts and growth stage allows farmers to perform field operations such as irrigation and thinning accurately and on time (Cook and Veseth, 1991). However, the collection of growth phase information is currently done through labour-intensive manual observations. This is a time-consuming process since it has to be carried out and reported every day or several times a day. Moreover, the labour-intensive approach is not objective, because observers can understand the same criterion differently, which can lead to errors.

Computer vision-based methods can be effectively used to monitor the growth of crops, as they do not contain the limitations of the labour-intensive process. These methods make a great contribution to the development of precision agriculture by making it easier to observe and measure the effects of different environmental factors on crops. There are many applications of computer vision technology in agricultural automation systems, such as monitoring of crop status (Sakamoto et al., 2012; Vega et al., 2015), weeds identification (Guerrero et. al, 2012) yield estimation (Payne et al., 2013), disease detection (Yuehua et al., 2007; Pourreza et al., 2015), and quality control. Kamilaris and Prenafeta-Boldú (2018) review the studies used Machine Learning Algorithm (MLA) in agricultural and food production challenges emphasizing that those algorithms provide high accuracy and outperforms commonly used image processing techniques.

However, one of the frequently encountered problem in this process is that the plants are getting closer each other as they grow (Figure 1). If the plants come into contact with each other or overlap, this makes the problem even bigger. This problem makes it difficult to extract features belongs single plant such as height and digital biomass. In this study, it is aimed to separate bean plants in the certain region by applying clustering algorithms used in the literature. According to the results of the application, the advantages and disadvantages of the algorithms are discussed. Then, the most suited algorithm is employed and the dataset having separate plants has been prepared.



Source: own data processing

Figure 1: An example views of plants starting to overlap each other.

Materials and methods

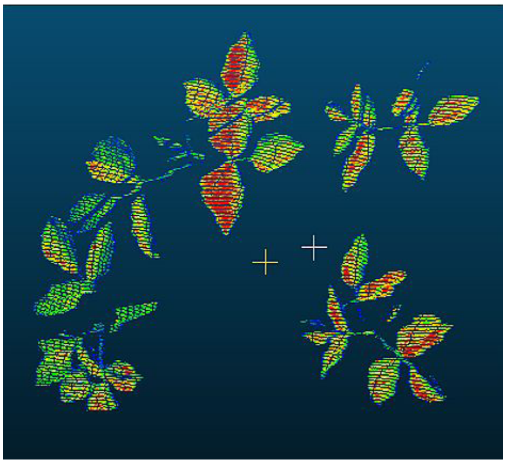
The bean dataset used in this study was provided by International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) in India. A large platform called "LeasyScan" has been developed at ICRISAT (**http://phenospex.com/blog/**) for the plant monitoring. LeasyScan is a laser based continuous plant monitoring system generates 3D point clouds data in every 2 hours on each of the 4800 sectors. Each data file has a time stamp that allows plants to be monitored and analysed depending on the time and environmental conditions. A part of a sample data file created during the scanning process of beans is given in Figure 2.

Pre-processing step is applied to the raw data to segment tray and soil data from the beans. This pre-processing is performed depending on the mathematical calculations achieved by using width, height and depth information of the trays. Later on, beans data that exist in each tray are saved as a separate data file. The final version

Source: own data processing

Figure 2: View of a sample scan file.

of the sample data file obtained after the pre-processing step is given in Figure 3.

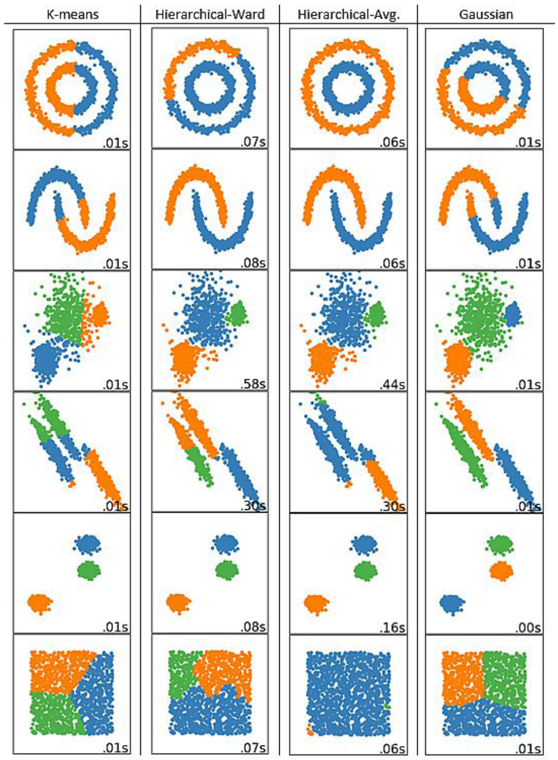


Source: own data processing

Figure 3: Sample data file after pre-processing process.

As seen in Figure 3, each data file consists of more than one bean. In order to observe the development of plants that grow close to each other and to calculate the phenotypic properties separately, firstly, the plants must be separated from each other. Although there are some algorithms developed in the literature to separate large plants, such as trees, an efficient algorithm has not been developed yet to separate small plants (Itakura and Hosoi, 2018; Zhang et al., 2019). In an exemplary segmentation algorithm used to separate trees, the first step is to cut the tree data at a certain height from the ground to create a dataset containing only tree trunks. Then, the points where these trunks are located are detected and accepted as the cluster centre for each tree. In the last step, the upper parts of the tree are clustered according to their closeness to the trunks/cluster centres. Since, there is no woody stem that can be detected and used as cluster centre, alternative approaches are required for the small plants.

In order to overcome this problem and separate the beans from each other, four widely used clustering algorithms have been utilized in this study. These algorithms are K-means, Gaussian mixtures and two different versions of Hierarchical clustering algorithms, respectively. All these algorithms require the number of clusters to be specified. Figure 4. illustrates the characteristics of clustering algorithms on 2D datasets (**https://scikit-learn.org/stable/auto_examples/cluster/plot_cluster_comparison.html**). The labels in the bottom right corner of the pictures show the running times of the algorithms for the related data set.



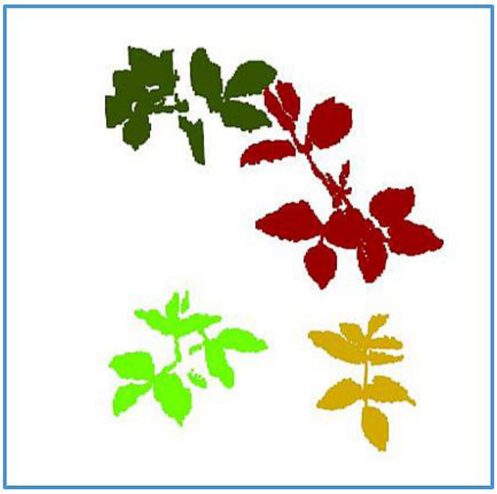
Source: https://scikit-learn.org/stable/modules/clustering.html

Figure 4. Results of k-means, hierarchical-ward, hierarchical-avg and gaussian mixture clustering algorithms on 2D datasets.

The K-means algorithm clusters data by separating data into k clusters, in which each datum belongs to the cluster with the nearest mean (cluster center). The aim of the algorithm is to minimize the sum-of-squares within cluster (Sabo, 2014). Since it is suitable for large amounts of data, it has been used in a wide variety of application areas in many different fields. Hierarchical clustering algorithms organize data into a hierarchical structure in proportion with the similarity matrix. These algorithms build nested clusters by merging or splitting them successively (Rokach and Maimon, 2005). There are 4 common metrics used for the merge strategy; ward, maximum(complete), average and single linkage (Ward, 1963; Defays, 1977; Sibson, 1973). In this this study, ward and average linkages are utilized. While ward minimizes the sum of squared differences within all clusters, average minimizes the maximum distance between observations of pairs of clusters. Gaussian mixture model is a probabilistic model for representing the presence of subpopulations within an overall population (Yu et al., 2011). It assumes that all the data points are generated from a mixture of a finite number of Gaussian distributions.

Ground truth data

To generate the ground truth dataset, one sample scan file was randomly selected from the dataset. Each scan file consists of 12 trays. According to the pre-processing steps, trays were separated from each other, the soil data were cleaned, and the beans data belong to each tray were saved as a separate data file. Then, all the beans in each data file were separated manually and saved. Colorized example of a manually created ground truth bean data file is given in Figure 5.



Source: own data processing

Figure 5: Colorized example of a manually created ground truth bean data.

Results and discussion

In order to investigate the performance of the clustering algorithms on bean data segmentation process, 12 ground truth data files were employed. All the algorithms were coded in Python environment using scikit-learn library. Before starting the test, cluster counts were set depending on the beans counts in each tray.

After the clustering operations were performed, results files were colorized and saved similar to the ground truth data given in Figure 5. This visualization made it easy to evaluate the performance of the algorithms and to see their shortcomings. Additionally, the success rates of the algorithms were calculated by comparing the result images produced by the algorithms with the ground-truth images. During this process, the ratio of pixels correctly clustered by the algorithms to the total number of pixels was calculated as in Eq. 1. White pixels forming the background in the image were not included in the calculation process, the obtained results were used to quantitatively assess the accuracy of the clustering algorithms in segmenting bean plants. The formula for the calculation of success rate (SR) is as follows:

$$SR = \frac{M_{pixel}}{T_{pixel}} \qquad (1)$$

Where; M_{pixel} and T_{pixel} are the number of matching (correctly clustered) pixels and number of total pixels (except white pixels), respectively. According to this function, SR values of the k-means, Gaussian mixtures and two different versions of Hierarchical clustering algorithms are given in Table 1., respectively.

In Table 1, while the results of all the clustering algorithms over all the data are given separately, the average results are also given at the last row.

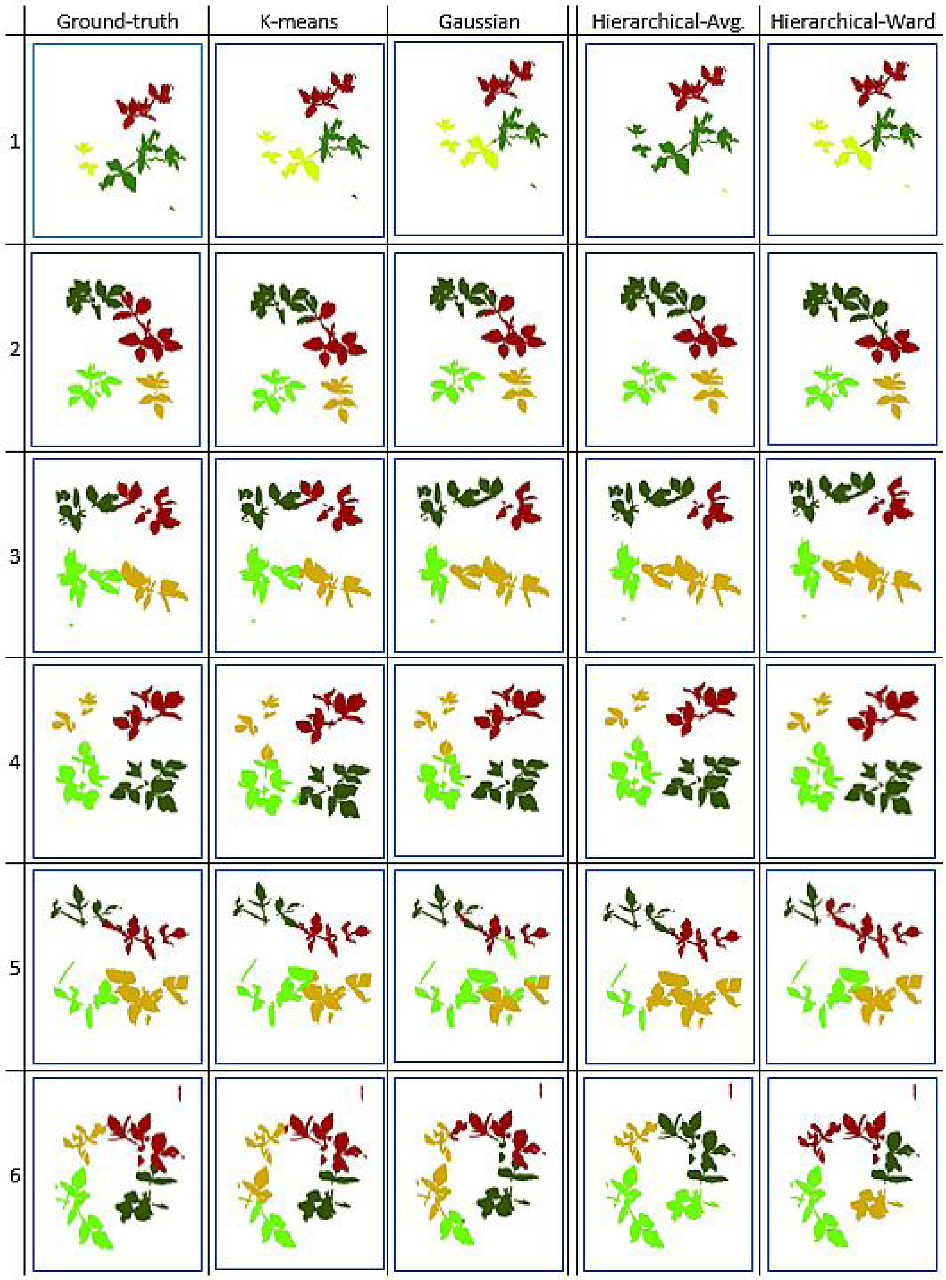
Tray	K-means	Gaussian	Hier.-Avg.	Hier.-Ward
1	77.54	77.87	88.37	78.38
2	93.79	94.98	88.39	86.01
3	96.54	80.26	80.92	80.92
4	94.91	94.88	99.87	99.87
5	89.51	74.44	88.31	84.8
6	84.45	74.55	54.91	59.54
7	89.18	88.97	85.81	99.88
8	93.78	97.01	99.98	99.98
9	98.26	94.22	99.76	99.76
10	83.36	80.91	93.19	85.12
11	95.67	81.38	93.96	94.05
12	97.09	82.05	87.54	98.85
Average	90.13	85.81	87.95	87.43

Source: own data processing

Table 1: SR results of clustering algorithms.

The results reveal that, all the clustering algorithm produces the acceptable results for the segmentation process. When the results are examined in detail, it is seen that the maximum SR differences among the algorithms occurs in the 6^{th} row/data. In order to analyse the performance of the algorithms in more detail, the colorized results of the algorithms on the first 6 data are given below.

When the first test data was examined, it is clearly seen that although hierarchical-avg clustering algorithms produces the best SR, this algorithm presents the worst performance in the visualized result. This is because this algorithm is very sensitive to noisy data. The same can be seen in the 6th result. Since this datafile has a small group of noisy data, the hierarchical-avg algorithm considers this noisy data as a separate cluster and tries to produce 3 clusters from the remaining data. The results of Gaussian algorithm illustrate that this algorithm is not robust when the plants are very close to each other and the leaves are overlapped. This weakness can be obviously seen



Source: own data processing

Figure 6: 2D visualization of ground truth data and results of k-means, gaussian mixture, hierarchical-avg and hierarchical-ward clustering algorithms, respectively.

in the 5th and 6th images given in the Figure 6. Therefore, just like the Hierarchical-avg. clustering algorithm, it can be misleading to evaluate the performance of the Gaussian algorithm just by looking at the quantitative results. By taking these considerations into account, we can conclude that performance of the Hierarchical-avg and Gaussian clustering algorithms are not good enough to separate the bean crops.

In accordance with the results comparison it can be said that the k-means clustering algorithm produces more consistent and reliable results than other algorithms. Although not as good as k-means, we can say that the Hierarchical-ward algorithm can be an alternative to k-means by looking at its general performance. Moreover, if algorithms are examined in terms of running time, as it seen in the Figure 4., k-means algorithm executes much faster than hierarchical algorithms. Therefore, the k-means algorithm provides advantages in separating bean plants in terms of both SR and running time.

Conclusion

The purpose of this study was to investigate the performance of clustering algorithms in bean-plant separation and to find the most suitable one. Before the clustering process, required pre-processing steps for tray segmentation and the soil cleaning were developed and employed. After the clustering process, the results were evaluated according to success rates and visualized clustering results. All algorithms performed the separation process with a success rate of over 85%. Despite this, it can be seen that the k-means algorithm is more advantageous in terms of both running-time and robustness against noise data. As a result, a reliable 3D bean-plant segmentation algorithm was obtained. This process will provide opportunities to extract the phenotypic traits belongs single plant such as height and digital biomass in the future works.

Acknowledgements

This research was carried out under the project: “Supporting the development of international mobility of research staff at CULS Prague”, reg. no. CZ.02.2.69/0.0/0.0/16_027/0008366.

EUROPEAN UNION
European Structural and Investment Funds
Operational Programme Research, Development and Education

MINISTRY OF EDUCATION, YOUTH AND SPORTS

The results and knowledge included herein have been obtained owing to support from the following institutional grant. Internal grant agency of the Faculty of Economics and Management, Czech University of Life Sciences Prague, grant no. 2019MEZ0006.

This research was carried out under the project: “Supporting the development of international mobility of research staff at CULS Prague”, reg. no. CZ.02.2.69/0.0/0.0/16_027/0008366.

Corresponding authors
Serkan Kartal, Ph.D.
Department of Information Technologies, Faculty of Economics and Management
Czech University of Life Sciences Prague, Kamýcká 129, Prague - Suchdol, 165 00, Czech Republic
E-mail: kartal@pef.czu.cz, skartal@cu.edu.tr

References

[1] Bannayan, M. and Sanjani, S. (2011) “Weather conditions associated with irrigated crops in an arid and semi arid environment”, *Agricultural and Forest Meteorology*, Vol. 151, No. 12, pp. 1589-1598. ISSN 0168-1923, DOI 10.1016/j.agrformet.2011.06.015.

[2] Cook, R. J. and Veseth, R. J. (1991) “Wheat health management”, St. Paul. MN. American Phytopathological Society.

[3] Defays, D. (1977) “An efficient algorithm for a complete link method”, *The Computer Journal*, Vol. 20, No. 4, pp. 364-366. E-ISSN 1460-2067, ISSN 0010-4620. DOI 10.1093/comjnl/20.4.364.

[4] Domínguez, J. A., Kumhálová, J. and Novák, P. (2015) “Winter oilseed rape and winter wheat growth prediction using remote sensing methods”, *Plant, Soil and Environment*, Vol. 61, No. 9, pp. 410-416. E-ISSN 1805-9368, ISSN 1214-1178. DOI 10.17221/412/2015-PSE.

[5] Gebbers, R. and Adamchuk, V. I. (2010) “Precision agriculture and food security”, *Science*, Vol. 327, No. 5967, pp. 828-831. ISSN 1095-9203. DOI 10.1126/science.1183899.

[6] Guerrero, J. M., Pajares, G., Montalvo, M., Romeo, J., and Guijarro, M. (2012) “Support vector machines for crop/weeds identification in maize fields”, *Expert Systems with Applications*, Vol. 39, No. 12, pp. 11149-11155. ISSN 0957-4174. DOI 10.1016/j.eswa.2012.03.040.

[7] Itakura, K. and Hosoi, F. (2018) “Automatic individual tree detection and canopy segmentation from three-dimensional point cloud images obtained from ground-based lidar”, *Journal of Agricultural Meteorology*, Vol. 74, No. 3, pp. 109-113. ISSN 1881-0136, ISSN 0021-8588. DOI 10.2480/agrmet.D-18-00012.

[8] Jannoura, R., Brinkmann, K., Uteau, D., Bruns, C. and Joergensen, R. G. (2015) “Monitoring of crop biomass using true colour aerial photographs taken from a remote controlled hexacopter”, *Biosystems Engineering*, Vol. 129, pp. 341-351. ISSN 1537-5110. DOI 10.1016/j.biosystemseng.2014.11.007.

[9] Kamilaris, A., Gao, F., Prenafeta-Boldu, F. X. and Ali, M. I. (2016) “Agri-IoT: A semantic framework for Internet of Things-enabled smart farming applications”, In *2016 IEEE 3rd World Forum on Internet of Things (WF-IoT)*, pp. 442-447). DOI 10.1109/WF-IoT.2016.7845467.

[10] Kamilaris, A. and Prenafeta-Boldú, F. X. (2018) “Deep learning in agriculture: A survey”, *Computers and Electronics in Agriculture*, Vol. 147, pp. 70-90. ISSN 0168-1699. DOI 10.1016/j.compag.2018.02.016.

[11] Khan, M. A., Akram, T., Sharif, M., Awais, M., Javed, K., Ali, H. and Saba, T. (2018) “CCDF: Automatic system for segmentation and recognition of fruit crops diseases based on correlation coefficient and deep CNN features”, *Computers and Electronics in Agriculture*, Vol. 155, pp. 220-236. ISSN 0168-1699. DOI 10.1016/j.compag.2018.10.013.

[12] Kitzes, J., Wackernagel, M., Loh, J., Peller, A., Goldfinger, S., Cheng, D. and Tea, K. (2008) “Shrink and share: humanity's present and future Ecological Footprint”, *Philosophical Transactions of the Royal Society B: Biological Sciences*, Vol. 363, No. 1491, pp. 467-475. ISSN 09628436. DOI 10.1098/rstb.2007.2164.

[13] Kumhálová, J., Kumhála, F., Novák, P. and Matějková, Š. (2013) “Airborne laser scanning data as a source of field topographical characteristics”, *Plant, Soil and Environment*, Vol. 59 No. 9, pp. 423-431. E-ISSN 1805-9368, ISSN 1214-1178. DOI 10.17221/188/2013-PSE.

[14] Kumhálová, J., Zemek, F., Novák, P., Brovkina, O. and Mayerová, M. (2014). “Use of Landsat images for yield evaluation within a small plot”, *Plant, Soil and Environment*, Vol. 60, No. 11, pp. 501-506. E-ISSN 1805-9368, ISSN 1214-1178. DOI 10.17221/515/2014-PSE.

[15] Muhammad, K., Mastuki, N., Darus, F. and Ghani, E. K. (2019) "Accounting information system change in an agriculture company: Examination using burns and scapens framework",.*Journal of International Studies*, Vol. 12, No. 1, pp. 105-118. E-ISSN 2306-3483, ISSN 2071-8330. DOI 10.14254/2071-8330.2019/12-1/7.

[16] Pantazi, X. E., Moshou, D., Alexandridis, T., Whetton, R. L. and Mouazen, A. M. (2016) “Wheat yield prediction using machine learning and advanced sensing techniques”, *Computers and Electronics in Agriculture*, Vol. 121, pp. 57-65. ISSN 0168-1699. DOI 10.1016/j.compag.2015.11.018.

[17] Payne, A. B., Walsh, K. B., Subedi, P. P. and Jarvis, D. (2013) “Estimation of mango crop yield using image analysis–segmentation method”, *Computers and Electronics in Agriculture*, Vol. 91, pp. 57-64. ISSN 0168-1699. DOI 10.1016/j.compag.2012.11.009.

[18] Raišienė, A. G., Podviezko, A., Skulskis, V. and Baranauskaitė, L. (2019) "Interest-balanced agricultural policy-making: Key participative and collaborative capacities in the opinion of NGOs’ experts", *Economics and Sociology*, Vol. 12, No. 3, pp. 301-318. E-ISSN 2306-3459, ISSN 2071-789X. DOI 10.14254/2071-789X.2019/12-3/20.

[19] Pourreza, A., Lee, W. S., Etxeberria, E. and Banerjee, A. (2015) “An evaluation of a vision-based sensor performance in Huanglongbing disease identification”, *Biosystems Engineering,* Vol. 130, pp. 13-22. ISSN 1537-5110. DOI 10.1016/j.biosystemseng.2014.11.013.

[20] Ramos, P. J., Prieto, F. A., Montoya, E. C. and Oliveros, C. E. (2017) "Automatic fruit count on coffee branches using computer vision", *Computers and Electronics in Agriculture*, Vol. 137, pp. 9-22. ISSN 0168-1699. DOI 10.1016/j.compag.2017.03.010.

[21] Rokach, L. and Maimon, O. (2005) "Clustering methods", In *Data mining and knowledge discovery handbook*, pp. 321-352. Springer, Boston, MA. ISBN 978-0-387-24435-8. DOI 10.1007/0-387-25465-X_15.

[22] Sabo, K., and Scitovski, R. (2014) "Interpretation and optimization of the k-means algorithm", *Applications of mathematics*, Vol. 59, No. 4, pp. 391-406. E-ISSN 1572-9109, ISSN 0862-7940. DOI 10.1007/s10492-014-0063-5.

[23] Sakamoto, T., Gitelson, A. A., Nguy-Robertson, A. L., Arkebauer, T. J., Wardlow, B. D., Suyker, A. E., ... and Shibayama, M. (2012) "An alternative method using digital cameras for continuous monitoring of crop status", *Agricultural and Forest Meteorology*, Vol. 154, pp. 113-126. ISSN 0168-1923. DOI 10.1016/j.agrformet.2011.10.014.

[24] Sibson, R. (1973) "SLINK: an optimally efficient algorithm for the single-link cluster method", *The Computer Journal*, Vol. 16, No. 1, pp. 30-34. ISSN 1460-2067, ISSN 0010-4620. DOI 10.1093/comjnl/16.1.30.

[25] Suh, H. K., Ijsselmuiden, J., Hofstee, J. W. and Van Henten, E. J. (2018) "Transfer learning for the classification of sugar beet and volunteer potato under field conditions", *Biosystems engineering*, Vol. 174, pp. 50-65. ISSN 1537-5110. DOI 10.1016/j.biosystemseng.2018.06.017.

[26] Vega, F. A., Ramirez, F. C., Saiz, M. P., and Rosúa, F. O. (2015) "Multi-temporal imaging using an unmanned aerial vehicle for monitoring a sunflower crop", *Biosystems Engineering*, Vol. 132, pp. 19-27. ISSN 1537-5110. DOI 10.1016/j.biosystemseng.2015.01.008.

[27] Ward Jr, J. H. (1963) "Hierarchical grouping to optimize an objective function", J*ournal of the American Statistical Association*, Vol. 58, No. 301, pp. 236-244. E-ISSN 1537-274X, ISSN 0162-1459. DOI 10.2307/2282967.

[28] Yuehua, C., Xiaoguang, H. and Changli, Z. (2007) "Algorithm for segmentation of insect pest images from wheat leaves based on machine vision", *Transactions of the Chinese Society of Agricultural Engineering*, Vol. 23, No. 12.pp. 187-191. ISSN 10026819.

[29] Yu, G., Sapiro, G. and Mallat, S. (2011) "Solving inverse problems with piecewise linear estimators: From Gaussian mixture models to structured sparsity", *IEEE Transactions on Image Processing*, Vol. 21, No. 5, pp. 2481-2499. ISSN 10577149. DOI 10.1109/TIP.2011.2176743.

[30] Zhang, J., He, L., Karkee, M., Zhang, Q., Zhang, X. and Gao, Z. (2018) "Branch detection for apple trees trained in fruiting wall architecture using depth features and Regions-Convolutional Neural Network (R-CNN)", *Computers and Electronics in Agriculture*, Vol. 155, pp. 386-393. ISSN 0168-1699. DOI 10.1016/j.compag.2018.10.029.

[31] Zhang, W., Wan, P., Wang, T., Cai, S., Chen, Y., Jin, X. and Yan, G. (2019) "A novel approach for the detection of standing tree stems from plot-level terrestrial laser scanning data", *Remote sensing*, Vol. 11, No. 2. ISSN 2072-4292. DOI 10.3390/rs11020211.

Agris on-line Papers in Economics and Informatics

Volume XII Number 3, 2020

Prediction and Context Awareness in Agriculture: A Systematic Mapping

Bruno Guilherme Martini, Gilson Augusto Helfer, Jorge Luis Victória Barbosa, Regina Célia Espinosa Modolo, Marcio Rosa da Silva, Rodrigo Marques de Figueiredo

University of Vale do Rio dos Sinos, São Leopoldo, Brazil

Abstract

The advances in sensorial technology and its use in agriculture have been contributing to the acquisition and analysis of data regarding agricultural production. Studies propose the use of sensors to monitor production or even the use of cameras to obtain crop information, providing data, reminders, and alerts to farmers. Through the obtainment and analysis of these data, context awareness can be used to improve systems, mainly through the prediction techniques applied to agriculture. This article presents a systematic mapping of studies that use prediction and context awareness in agriculture. During the mapping, 10206 articles were initially identified and, after filtering by inclusion and exclusion criteria, 42 articles were selected. The results indicated that 35.7% (15/42) of the studies used one or more prediction techniques and 45.2% (19/42) used image processing through pictures of cameras to obtain information regarding planting. 23 sensors with different functionalities were found, those have been used in the collection of data for context formation in agriculture.

Keywords

Technology for agriculture; systematic mapping; prediction in agriculture; context awareness in agriculture.

Martini, B. G., Helfer, G. A., Barbosa, J. L., Modolo, R. C. E., da Silva M. R. and de Figueiredo, R. M. (2020) “Prediction and Context Awareness in Agriculture: A Systematic Mapping", *AGRIS on-line Papers in Economics and Informatics*, Vol. 12, No. 3, pp. 45-58. ISSN 1804-1930. DOI 10.7160/aol.2020.120305.

Introduction

The advances in technology have been expanding agriculture’s modernization. This modernization has been occurring mainly through the usage of embedded computers, development of sensors for the most diverse purposes related to agriculture and increased efficiency and reliability in the communication between electronic devices. Thus, agriculture is adopting technological resources designed to increase performance from planting to harvesting.

These technological advances culminated in increased productivity and immunity of the planted crops. A factor that drives the pursuit for increased agricultural productivity is that approximately 70% of the available fresh water is consumed by the agricultural sector (Pimentel et al., 2004). Therefore, increasing the efficiency of agriculture becomes strategic, fomenting public and private investments in this sector. Through more assertive equipment and the possibility to obtain data that was not possible years ago, agriculture began to be called Precision Agriculture (Manzatto et al., 1999). Along these lines, technology has been applied in the detection of pests and diseases, planting failures, excessive irrigation, support and monitoring of the harvest and assistance to the demarcation of planting and controlled sprays.

As information on crops become increasingly accessible, studies on context awareness (Dey et al., 2001) and prediction (Nagini et al., 2016) are being developed more frequently by the scientific community. Context means any information that characterizes objects, environment or people. The process of obtaining this information must occur without the user having to take any action, that is, it must occur automatically. Context awareness refers to computational systems that induce actions or perform operations to improve the usability in environments and systems. Prediction can be defined as the act or effect of affirming something that will occur in the future, as approached by Nagini et al (2016).

Regarding methodology, this article uses systematic mapping which, according to Petersen et al. (2015), eliminates or reduces the differences that may exist when compared to revisions that do not have a determined method and with expressive results. Therefore, it is possible to obtain more reliable results and a greater impact. This study allows showing the techniques and technologies most used for prediction and context awareness applied in agriculture.

This work also aims, based on the literature analyzed, to present possible research gaps and challenges. This way, this article proposal is to present the state of the art, supporting future researches. The motivation of this study is to find out which are the most relevant databases for research in the area of prediction and context awareness in agriculture. In addition, to discover who are the main authors of the area, which are the most important works, the most used sensors, the used techniques of prediction, the deficiencies and the challenges of the area in order to assist future research.

Materials and methods

Through the use of a systematic mapping methodology, not only a final conclusion is discussed, but all the activities related to the discoveries throughout the mapping process are debated. Thus, the systematic mapping is able to make the connection between the data collection, the places where the articles were published, the area of application in agriculture, the year of publication of the studies, technologies that were used, which methodology was used in their development, among other information. The mapping focuses on published articles, which may be published in newspapers, conferences, books, and workshops. The steps for mapping are: a) elaboration of the research questions; b) elaboration of the search process and c) definition of the criteria to filter the results.

Research questions

For this work, research questions were defined and organized into three general questions (GQ), two specific questions (SPQ) and two statistical questions (STQ). The purpose of the general questions is to understand in what areas of agriculture context awareness and/or prediction are applied, what technologies are currently used and how information is communicated between sensors and servers. The objective of the specific questions is to identify the studies that use cameras to obtain information, which systems use prediction applied in agriculture and which technique is used in each of them. Finally, the goal of the statistical questions is to gather data on the studied area and show the behavior of publications over the last 10 years. The technologies used for context awareness and prediction in agriculture were hardly applied before 2008, therefore, the year 2008 served as a limit. The research questions are detailed in Table 1.

Type	Details
General Questions	
GQ 1	Which context-aware computing technologies are being used in agriculture?
GQ 2	Is the information sent via cable or by what type of wireless network?
GQ 3	Which areas of agriculture are applying context awareness and/or prediction?
Specific Questions	
SPQ 1	Which works use cameras to obtain information in agriculture?
SPQ 2	Which prediction techniques are used in agriculture?
Statistical Questions	
STQ1	Where were the researches published?
STQ 2	How many publications per year?

Source: author

Table 1: Research questions.

Elaboration of the search process

The search process was organized in three stages: specify the search string, select the databases and understand the search method of each one of them and, finally, obtain the results of the research. For the first stage, it was adopted the research process proposed by Kai Petersen (2015). The first stage identifies the main terms and their most relevant synonyms. In this study, the term "Agriculture" was selected as the primary term and as secondary the terms "Context Prediction", "Context Awareness", "Cyber-Physical System" and "Internet of Things". The synonyms for Agriculture were words that refer to methods of planting regarding indoor agriculture (hydroponics and aquaponics), to ensure that it would also be obtained studies applied in greenhouses or pavilions, as there are specific methods for internal planting. For the secondary terms, acronyms of the original terms or synonyms were used, as shown in Table 2.

Once the terms and synonyms were defined, the following search string was elaborated:

Search String Terms	
Main Terms	***Synonyms***
Agriculture	*"agriculture" OR "hydroponics" OR "aquaponics"*
Context Predicton	*"context prediction" OR "prediction of context"*
Cyber Physical System	*"cyber physical system" OR "cps"*
Context Awarness	*"context awareness" OR "context aware"*
Internet of Things	*"IoT" OR "Internet of Things"*

Source: author

Table 2: Search terms.

(("agriculture" OR "hydroponics" OR "aquaponics") AND ("context prediction" OR "prediction of context" OR "cyber physical system" OR "cps" OR "context awareness" OR "context aware" OR "IoT" OR "Internet of Things")). With the search string defined, the second step was to select databases that were relevant to the study area to apply the string. As each base has its search method, after selecting them it was necessary to determine the search parameters to be used in each one. Thus, four research databases were used, including the ACM Digital Library, IEEE Xplore, Science Direct and Springer Link.

The selected databases are the ones that prioritize research studies in the area of computer science since the mapping has as an emphasis on context awareness and prediction applied in agriculture. Therefore, articles from the field of agriculture could enter the selection, but when going through the filtering criteria would be withdrawn if they didn't have any link to context awareness or prediction.

The search in the ACM Digital Library database required the use of advanced search features, in which the search string was entered into the edit field of the database itself. For the search in the IEEE Xplore database, it was also used the advanced search feature. The search process in the Science Direct repository involved the application of the string to the title, abstract, and keywords. Lastly, in the Springer Library, in addition to using the search string as the query, it was also necessary to remove documents categorized as "Preview Only". After the exclusion of these files, the search filter "Computer Science" was selected first, the search was carried out and the "Life Science" filter was selected for results that could also be linked to the searched topic.

Definition of the criteria to filter the results

A few criteria were defined to include and exclude works obtained from the databases. Therefore, the Inclusion Criteria (IC) and Exclusion Criteria (EC) were defined as follows:

- IC 1: The study must be published in a journal, conference or workshop.
- IC 2: The study must be a full paper.
- IC 3: The study must be related to computer science or life science areas.

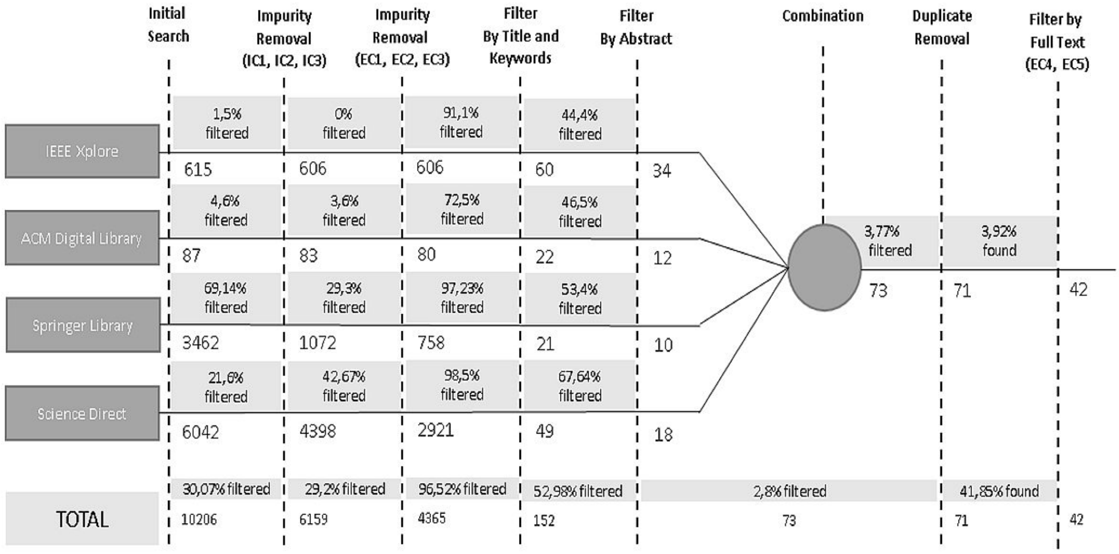
The Exclusion Criteria (EC) are the following:

- EC 1: Studies published before 2008.
- EC 2: Studies that were not written in English.
- EC 3: Studies that were published in theses and journals.
- EC 4: Studies that do not present a connection between agriculture and computer science.
- EC 5: Studies that are not related to the search questions.

The inclusion and exclusion criteria assist during the filtering process in order to obtain works that are more adherent to the study and to eliminate discrepancies generated in the research. Due to the filtering, impurities that did not meet the inclusion criteria were removed, this being the first step in the filtration. The number of studies obtained in each step of the filtering and the percentages of work withdrawn in each step can be seen in Figure 1.

After the articles were filtered in the initial search through the string and after having passed the first 2 IC, they went to the third filtering. The third mapping filter was the withdrawal of studies due to the first three exclusion criteria. Next, the works were filtered by title and keywords. In the fifth filtering step, the studies were removed after reading the abstract. Then the duplicate articles were removed.

At the end of the filtering process, a complete reading of each article was performed, observing the last two exclusion criteria (EC 4 and EC 5). Figure 1 presents the filtering process,



Source: author

Figure 1: Complete filtering of systematic mapping.

with the Inclusion Criteria and Exclusion Criteria applied in each step. The IEEE Xplore and ACM databases presented fewer articles. However, these bases were more assertive, considering the percentage between the number of articles found in the initial research until their combination after reading by abstract. Considering that the Springer Library and Science Direct databases had more than 99% of the results filtered along the process, it can be concluded that, although they have brought more results when compared with the other databases, the searches originated fewer articles that were adhering to the proposed theme.

Results and discussion

In this section, the results obtained by reading and analyzing the 42 mapped studies are presented. The research questions were answered and additional considerations and analysis on the studies were presented.

GQ 1: Which context-aware computing technologies are being used in agriculture?

The articles were classified according to the type of microcontrollers that they use in the systems, as well as the types of sensors that are being used to obtain the context information of the plantations. In the studies, the two most used microcontrollers are Raspberry and Arduino. Table 3 shows that they are used in 50% (21/42) of the works. Not all articles in the systematic mapping used microcontrollers. However, some studies used two microcontrollers (Kumar et al., 2016; Popovic et al., 2017; Goap et al., 2018; Mehra et al., 2018).

Table 4 presents the sensors that were used to obtain planting information. The types of sensors were of the most diverse functionalities, totaling 23 types. Not all sensors with the same functionality are identical, for example, among the moisture sensors the sensor SHT11 was applied in the study by Santos et al. (2018) and the DHT11 sensor in the study by Kumar et al. (2016). Thus, it is possible to obtain information even with different sensor models. As the purpose of this mapping is to inform what information and variables are being obtained and not to indicate exactly which sensors are being applied in each study, only the kind of sensors used by each one of the studies was indicated.

Microcontroller	*Number of articles*	*Percentage*	*Article's ID*
ARM	1	2.4%	(Kubicek et al., 2013)
Arduino	11	26.4%	(Fukatsu, 2014), (Patil and Thorat, 2016), (Kumar et al., 2016), (Rajendrakumar and Parvati, 2017), (Popovic´ et al., 2017), (Saha et al., 2017), (Athani and Tejeshwar, 2017), (Rodríguez et al., 2017), (Santos et al, 2018), (Goap et al, 2018), (Mehra et al., 2018)
Intel Galileo Gen2	1	2.4%	(Ivanov and Tsvetkov, 2017)
Raspberry	14	33.3%	(Rodriguez et al., 2014), (Kumar et al., 2016), (Alipio et al., 2017), (Eko et al., 2017), (Popovic´ et al., 2017), (Arakeri et al., 2017), (Sarangdhar et al., 2017) (Kokkonis et al., 2017), (Jacob, 2017), (Kumar et al., 2017), (Park et al., 2017), (Goap et al, 2018), (Mehra et al., 2018), (Fieh et al., 2018)
Zigbee	6	14.4%	(Zhou et al., 2011), (Divya et al., 2014), (Patil and Thorat, 2016), (Zhang et al., 2017), (Rodríguez et al., 2017), (Zhou and Li, 2017)

Source: author

Table 3: Relation of the used microcontrollers.

Sensors	*Number of articles*	*Percentage*	*Article's ID*
Soil Moisture	22	52.4%	(Zhou et al., 2011), (Zhou et al., 2012), (Divya et al., 2014), (Rodriguez et al., 2014), (Tan and Tan, 2016), (Patil and Thorat, 2016), (Carrasquilla-Batista et al., 2016), (Ivanov and Tsvetkov, 2017), (Rajendrakumar and Parvati, 2017), (Zhang et al., 2017), (Popovic' et al., 2017), (Saha et al., 2017), (López-Riquelme et al., 2017), (Athani and Tejeshwar, 2017), (Rodríguez et al., 2017), (Sarangdhar et al., 2017), (Zhou and Li, 2017), (Kokkonis et al., 2017), (Joshi et al., 2017), (Park et al., 2017), (Goap et al., 2018), (Huong et al., 2018)
Water Temperature	6	14.3%	(Shahriar and Mcculluch, 2014), (Xu et al., 2014), (Alipio et al., 2017), (Eko et al., 2017), (Jacob, 2017), (Santos et al., 2018)
Environmental Temperature	18	42.8%	(Zhou et al., 2011), (Zhou et al., 2012), (Kubicek et al., 2013), (Tan et al., 2015), (Patil and Thorat, 2016), (Kumar et al., 2016), (Carrasquilla-Batista et al., 2016), (Ivanov and Tsvetkov, 2017), (Rajendrakumar and Parvati, 2017), (Eko et al., 2017), (Popovic' et al., 2017), (Saha et al., 2017), (Rodríguez et al., 2017), (Sarangdhar et al., 2017), (Zhou and Li, 2017), (Joshi et al., 2017), (Goap et al., 2018), (Plazas et al., 2018)
Luminosity	12	28.6%	(Zhou et al., 2011), (Zhou et al., 2012), (Tan et al., 2015), (Carrasquilla-Batista et al., 2016), (Ivanov and Tsvetkov, 2017), (Alipio et al., 2017), (Eko et al., 2017), (Rodríguez et al., 2017), (Zhou and Li, 2017), (Joshi et al., 2017), (Park et al., 2017), (Plazas et al., 2018)
CO2	4	9.6%	(Zhou et al., 2011), (Zhou et al., 2012), (Zhou and Li, 2017), (Park et al., 2017)
Soil Chemicals	7	16.7%	(Zhou et al., 2012), (Kubicek et al., 2013), (Divya et al., 2014), (Tan et al., 2015), (Kumar et al., 2016), (Popovic' et al., 2017), (Park et al., 2017)
Soil's temperature	8	19%	(Zhou et al., 2011), (Zhou et al., 2012), (Kubicek et al., 2013), (Divya et al., 2014), (Rodriguez et al., 2014), (Tan et al., 2015), (López-Riquelme et al., 2017), (Goap et al., 2018)
Environmental humidity	17	40.5%	(Zhou et al., 2011), (Zhou et al., 2012), (Kubicek et al., 2013), (Tan et al., 2015), (Patil and Thorat, 2016), (Kumar et al., 2016), (Carrasquilla-Batista et al., 2016), (Alipio et al., 2017), (Rajendrakumar and Parvati, 2017), (Eko et al., 2017), (Saha et al., 2017), (Rodríguez et al., 2017), (Sarangdhar et al., 2017), (Zhou and Li, 2017), (Santos et al., 2018), (Goap et al., 2018), (Plazas et al., 2018)

Source: author

Table 4: Relation of the used sensors.

Sensors	*Number of articles*	*Percentage*	*Article's ID*
Atmospheric pressure	2	4.8%	(Popovic' et al., 2017), (Plazas et al., 2018)
Wind Speed	2	4.8%	(Popovic' et al., 2017), (Plazas et al., 2018)
Rain	2	4.8%	(Shahriar and Mcculluch, 2014), (Popovic' et al., 2017)
RFID	1	2.4%	(Kubicek et al., 2013)
Soil's pH	2	4.8%	(Zhou et al., 2011), (Divya et al., 2014)
Salinity	2	4.8%	(Shahriar and Mcculluch, 2014), (Xu et al., 2014)
Dissolved oxygen	1	2.4%	(Xu et al., 2014)
Water Conductivity	3	7.2%	(Carrasquilla-Batista et al., 2016), (Alipio et al., 2017), (Eko et al., 2017)
UV	2	4.8%	(Goap et al., 2018), (Plazas et al., 2018)
Solenoid	1	2.4%	(Kokkonis et al., 2017
Thermohygrometer	1	2.4%	(Plazas et al., 2018)
Water Level	5	12%	(Shahriar and Mcculluch, 2014), (Carrasquilla-Batista et al., 2016), (Sarangdhar et al., 2017), (Jacob, 2017), (Mehra et al., 2018)
Water's pH	8	19%	(Xu et al., 2014), (Carrasquilla-Batista et al., 2016), (Alipio et al., 2017), (Rajendrakumar and Parvati, 2017), (Eko et al., 2017), (Saha et al., 2017), (Athani and Tejeshwar, 2017), (Mehra et al., 2018)
Ultrasonic	1	2.4%	(Arakeri et al., 2017)

Source: author

Table 4: Relation of the used sensors.

GQ 2: Is the information sent via cable or by what type of wireless network?

The transmission of the information basically involves three types of communication: Wi-Fi, mobile data (3G / 4G) and cable. In this systematic mapping, the Wi-Fi communication was considered as all that wireless communication that does not communicate data through GPRS/GSM modules. This Wi-Fi communication can be between the dispersed microcontrollers in the planting and a server, the communication between the microcontrollers, among others. Table 5 shows the types of communication, and which studies use each type and how many studies use each transmission type.

QG 3: Which areas of agriculture are applying context awareness and/or prediction?

Indoor agriculture is defined as the agriculture held in an enclosed location, such as a plantation inside a greenhouse or pavilion. In this type of culture, 11 studies were found. Outdoor agriculture is applied in open environments where the possibility of environmental control is reduced. For this type of cultivation 21 studies were found, which are distributed throughout the period of the mapping, since there are studies from 2009 to 2018.

The term agriculture has been used for studies that can be applied to both indoor and outdoor agriculture. Table 6 shows the areas where the studies were applied. Since in the indoors agriculture there are different cultivation methods, Table 6 considers separately the traditional method of indoor agriculture as specifically three alternative types of indoor agriculture, namely, Aquaculture, Hydroponics and Aquaponic.

Hydroponics is a model of agriculture cultivation that works by recirculating water that has the necessary nutrients for the plants, it is used in closed places, such as greenhouses or pavilions. Prediction and context awareness studies in hydroponics have recently begun in 2017 with the studies of Alipio et al. (2017) and Eko et al. (2017) and in 2018 the study by Mehra et al. (2018) was published. It was also found two studies on aquaculture, that is fish farming with control over water quality (Shahriar and Mcculluch, 2014; Xu et al., 2014). In aquaponics was found only one study of Jacob (2017).

Among the areas of application are the studies on outdoor agriculture, always in greater numbers during the years. However, the number of studies applied in indoor agriculture increased substantially, 8 studies in the years 2017 and 2018 against 2 studies published in 2015 and 2016. On the other hand, outdoor agriculture grew associated with the same period at a rate lower (15 studies against 5 studies) than indoor agriculture,

Type of communication	*Number of articles*	*Percentage*	*Article's ID*
Wired	3	7.2%	(Zhang et al., 2017), (Joshi et al., 2017), (Mehra et al., 2018)
Mobile Data	14	33.3%	(Luimula and Shelby, 2009), (Zhou et al., 2012), (Kubicek et al., 2013), (Divya et al., 2014), (Rodriguez et al., 2014), (Xu et al., 2014), (Carrasquilla-Batista et al., 2016), (Ivanov and Tsvetkov, 2017), (Zhang et al., 2017), (Popovic' et al., 2017), (Saha et al., 2017), (Zhou and Li, 2017), (Kokkonis et al., 2017), (Plazas et al., 2018)
Wi-Fi	24	57.1%	(Luimula and Shelby, 2009), (Zhou et al., 2011), (Zhou et al., 2012), (Kubicek et al., 2013), (Fukatsu, 2014), (Divya et al., 2014), (Rodriguez et al., 2014), Rupanagudi et al., 2015), (Tan et al., 2015), (Kumar et al., 2016), (Ivanov and Tsvetkov, 2017), (Rajendrakumar and Parvati, 2017), (Eko et al., 2017), (Zhang et al., 2017), (Athani and Tejeshwar, 2017), (Rodríguez et al., 2017), (Sarangdhar et al., 2017), (Zhou and Li, 2017), (Jacob, 2017), (Joshi et al., 2017), (Park et al., 2017), (Santos et al., 2018), (Goap et al., 2018), (Fiehn et al., 2018)

Source: author

Table 5: Type of communication of the prototypes.

Agriculture Area	*Number of articles*	*Percentage*	*Article's ID*
Agriculture	8	19%	(Divya et al., 2014), (Janaszek, 2016), (Rajendrakumar and Parvati, 2017), Yahata et al., 2017), (Zhang et al., 2017), (Popovic' et al., 2017), (Zhou and Li, 2017), (Huong et al., 2018)
Outdoor agriculture	21	50%	(Luimula and Shelby, 2009), (Zhou et al., 2011), (Kubicek et al., 2013), (Fukatsu, 2014), (Rodriguez et al., 2014), (Rupanagudi et al., 2015), (Tan et al., 2015), (Tan and Tan, 2016), (Patil and Thorat, 2016), (Kumar et al., 2016), (Ivanov and Tsvetkov, 2017), (Saha et al., 2017), (López-Riquelme et al., 2017), (Arakeri et al., 2017), (Sarangdhar et al., 2017), (Kokkonis et al., 2017), (Joshi et al., 2017), (Goap et al., 2018), (Plazas et al., 2018), (Treboux and Genoud, 2018), (Fiehn et al., 2018)
Indoor agriculture	7	16.7%	(Zhou et al., 2012), (Ma et al., 2015), (Carrasquilla-Batista et al., 2016), (Athani and Tejeshwar, 2017), (Rodríguez et al., 2017), (Park et al., 2017), (Santos et al., 2018)
Aquaculture	2	4.8%	(Shahriar and Mcculluch, 2014), (Xu et al., 2014)
Hydroponics	3	7.2%	(Alipio et al., 2017), (Eko et al., 2017), (Mehra et al., 2018)
Aquaponic	1	2.4%	(Jacob, 2017)

Source: author

Table 6: Areas of agriculture in which the studies have been applied.

but continues to have higher final numbers. This growth of indoor agriculture has generated a balance between areas over the past two years. The two studies on aquaculture (Shahriar and Mcculluch, 2014; Xu et al., 2014) were published in 2014.

SPQ 1: Which works use cameras to obtain information in agriculture?

During the mapping, 19 articles that use cameras to obtain information about the production process were found. Among them, the studies of Zhou and Li (2017) and Jacob (2017) used the cameras to visualize the plantation, that is, production information was not obtained beyond the visualization of the crop. Other 17 studies (Luimula and Shelby, 2009; Zhou et al., 2012; Kubicek et al., 2013; Fukatsu, 2014; Rodriguez et al., 2014; Rupanagudi et al., 2015; Ma et al., 2015; Tan et al., 2016; Kumar et al., 2016; Janaszek, 2016; Yahata et al., 2017; Zhang et al., 2017; Arakeri et al., 2017; Sarangdhar et al., 2017; Joshi et al., 2017; Treboux and Genoud, 2018; Fiehn et al., 2018) obtained information for several functionalities, among them, detection of pests in the plantation, identification of fruits with problems in their formation, insect count, plants growth, among others.

The cameras used in the studies have the function of taking pictures. The information taken from the images is provided by other devices, applications and programs through image processing. Image processing is getting more and more expressive results, as is evident in the studies of this mapping. Images of plantings were analyzed in different ways in the studies and provided information such as detecting plant species indication (Kumar et al., 2016), plant sizes (Joshi et al., 2017) and (Fukatsu, 2014), plague on fruits (Tan et al., 2016), indication of bird species (Fiehn et al., 2018), insect count (Fukatsu, 2014), fruit color (Tan et al., 2016), among other aspects. These applications show the potential in the area because of the diversity of information that can be obtained with the same equipment.

SPQ 2: Which prediction techniques are used in agriculture?

Predictive systems avoid losses, for they avoid the non-ideal zone of execution of the system, but in return, a more comprehensive monitoring of information is required, as well as a forecast model. Thus, prediction software depends on the generation of a predictive model based on data analysis, such as the context that it is included. Table 7 shows the 16 works that use the prediction in agriculture, highlighting the technique applied in each of them. Some articles use more than one technique and for that reason when adding the used techniques, their value is greater than the number of studies.

Table 7 shows that there are studies that applied two or even three prediction techniques. Park et al. (2017) applied three techniques and then reported which one got the best result. The works of Shahriar and Mcculluch (2014), Janaszek (2016), Patil and Thorat (2016) and Rodríguez et al. (2017) applied two prediction techniques.

STQ 1: Where were the researches published?

The articles were grouped according to the four databases that were used in the research. The base that obtained the largest number of papers was the IEEE Xplore with 16 articles out of a total of 42, or 38%. In second place was the Science Direct database with 12 articles, resulting in 28.5%. The ACM Digital Library database returned 8 studies (19%) and finally, the Springer Library returned 6 articles, the equivalent to 14.3%. The Figure 2 shows the 42 articles distributed per year of publication and database.

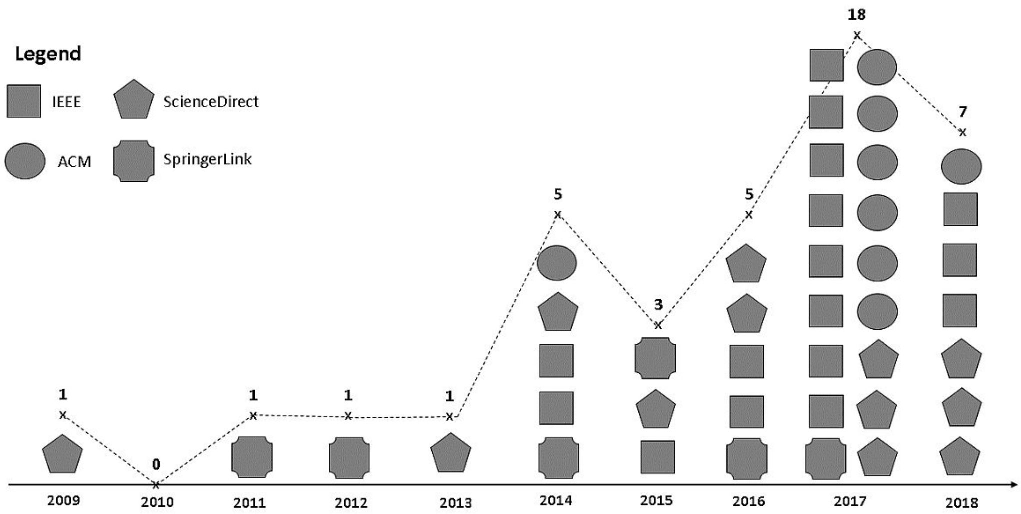
STQ 2: How many publications per year?

Over the past two years, the number of publications on prediction and/or context awareness has increased significantly. This growth shows the interest of researchers in applying prediction and contexts to improve the cultivate in agriculture as well as the quality of products through the better monitoring of production. It is also possible to verify that the database that was most present in these 10 years of studies was the Science Direct, because in 7 of 10 years there were articles selected. From 2014, there was a greater diversity of the databases until 2017. In 2018 the Springer Library database was the only one that had no articles selected for the mapping.

Agriculture area	*Number of articles*	*Percentage*	*Article's ID*
Expert Rules	1	2.4%	(Shahriar and Mcculluch, 2014)
Time Series	1	2.4%	(Shahriar and Mcculluch, 2014)
Linear/Vector Regression	3	7.2%	(Rodríguez et al., 2017), (Sarangdhar et al., 2017), (Goap et al., 2018)
Neural Networks	5	12%	(Athani and Tejeshwar, 2017), (Rodríguez et al., 2017), (Park et al., 2017), (Mehra et al., 2018), (Plazas et al., 2018)
Statistical Method	1	2.4%	(Patil and Thorat, 2016)
Markov Model	2	4.8%	(Patil and Thorat, 2016), (Huong et al., 2018)
Bayesians Networks	1	2.4%	(Alipio et al., 2017), (Joshi et al., 2017)
ARIMA	3	7.2%	(Park et al., 2017), (Santos et al., 2018), (Huong et al., 2018)
Decision Tree	1	2.4%	(Treboux and Genoud, 2018)
Genetic Algorithm	1	2.4%	(Janaszek, 2016)
Response Surface Methodology	1	2.4%	(Janaszek, 2016)
BirdNet	1	2.4%	(Fiehn et al., 2018)
Lasso Regression	1	2.4%	(Park et al., 2017)

Source: author

Table 7: Prediction techniques used in agriculture.



Source: author

Figure 2: Number of publications from 2009 to 2018.

Threats to the validity of work

Systematic mapping works are exposed to risks that may invalidate the study. These risks may arise from erroneously decisions taken during systematic mapping. In order to guarantee better results, four databases were selected, each of them known in the academic area for their relevance in computer science and similar areas. Therefore, it was sought to mitigate the possibility of databases impacting the results.

Five main terms were classified as a primary term and four secondary terms. Their synonyms were used to construct the search string. As a result of the string, we sought to obtain studies that related to one of the four secondary terms related to the primary term "Agriculture" and its specific areas "Hydroponics" and "Aquaponics". Thus, it was sought to obtain the largest number of studies in the search results.

The mapping did not consider conceptual aspects of the area of Agriculture, such as studies applied only in the field of natural sciences, focusing exclusively on the application of Agriculture involving the secondary terms "Prediction", "Context Awareness", "Cyber-Physical Systems" and "Internet of Things". This decision contributed to the filtering, directing the search only to a specific context within the area of "Agriculture and Computer Science", a discipline that concentrates a large amount of academic work.

The filtering process may have restricted the work so that some relevant studies may have been removed. To mitigate this risk, the filtering process was based on a technique widely adopted in academic works of systematic mapping (Petersen et al., 2015). The used revision processes were already applied by other authors (Dias et al, 2018; Bischoff et al., 2018; Dalmina et al., 2019).

Discussion

Only 16 studies used prediction in agriculture, that is, 38.1% of the articles. Among them, the most used technique was Neural Networks in five articles (Athani and Tejeshwar, 2017; Rodríguez et al., 2017; Park et al., 2017; Mehra et al., 2018; Plazas et al., 2018). The ARIMA and Linear/Vector Regression techniques were used each in three studies. Linear/Vector Regression was used in the studies of Rodríguez et al. (2017), Sarangdhar et al. (2017) and Goapa et al. (2018). On the other hand, ARIMA was used in the studies of Park et al. (2017), Santos et al. (2018) and Huong et al. (2018).

The use of image processing to capture information of pictures has increased in recent years and with that more systems and software related to Agriculture are able to use them. The information may have different purposes such as detecting plant species (Kumar et al., 2016), bird species (Fiehn et al., 2018), plant size (Joshi et al., 2017), (Fukatsu, 2014), fruit color (Tan et al., 2016), plagues on fruits (Tan et al., 2016) and also to count insects (Fukatsu, 2014). These applications show the potential in the area because of the diversity of information that can be obtained with the same equipment. Helfer et al. (2019) showed that studies already used large amounts of Wireless Sensors Networks (WSN) this aspect was also detected in this mapping. Another point in common between this study and that of Helfer et. al, 2019 was the use of the soil moisture sensor in most systems for automatic irrigation.

Analyzing the entire filtering process presented in Figure 1, and the Statistical Question 1 (STQ 1) that presents in which databases the articles were published, it is possible to verify that the most accurate databases that processed the query string were the ACM Library and IEEE Xplore, because they had a small number of articles in the initial survey, but have a relatively high number when compared to the total articles used in the mapping. The least accurate database was Science Direct, because of the 6042 papers resulting from the initial research, only 11 studies were used after the last filtering.

Another statistic was that by 2013 the amount of annual work on prediction and context awareness in agriculture remained stable, with up to one publication per year. From 2014, there was a progressive increase until 2018. In 2014, 5 works were published, in 2015 had a decrease (3 works) and in 2016 the number went back to 5 publications. However, there has been a great increase in publications in the last two years, showing the growing interest of researchers in this area.

Conclusion

The systematic mapping study presented the state of the art in the use of prediction and/or context awareness in Agriculture. This work also presented the main technologies used to obtain information, such as micro controllers. Also, it was found that various sensors are used to obtain planting information. In addition to the used sensors, it was also informed of the way used to send the obtained information regarding the crops to the servers in the studied articles. Besides presenting which works used cameras to obtain crop information, it was also possible to generate statistical data of the publications referring to the researched themes in the last decade.

Zewge and Dittrich (2017) indicated the problem of no integration between projects and suggested the creation of an integrated project that would allow the exchange of information not only between developers, but also between farmers. This mapping showed that many studies applied the same methodology as the usage of sensors to obtaining and analyze the gathered data. It is highlighted that 16 of the articles present the development of a predictive system. Therefore, some studies become repetitive regarding methodology, but there is a great diversity of prediction techniques applied in each of them, as shown in Table 7. A challenge for future works is to deepen the studies and classify them into similar projects, which can make the use of prediction faster and provide better results. Another strategic future work will be the extension of this mapping through the evaluating of the relevance of publishing sources, for example, considering the rank of the journal or the total number of paper citations.

The use of mathematical and statistical algorithms to prove the efficacy of the work is present in the works that used prediction in Agriculture (Shahriar and Mcculluch, 2014; Patil and Thorat, 2016; Janaszek, 2016; Joshi et al., 2017; Alipio et al., 2017; Athani and Tejeshwar, 2017; Rodríguez et al., 2017; Sarangdhar et al., 2017; Park et al., 2017; Santos et al., 2018; Goap et al., 2018; Mehra et al., 2018; Plazas et al., 2018; Huong et al., 2018; Treboux and Genoud, 2018; Fiehn et al., 2018). These studies sought to show the result and the uncertainty when working with prediction, not being able to guarantee the occurrence of a certain event, although giving indications of situations that can improve the production due to the data analyzed.

A review of the work that has been published in the last decade may provide initial insights for new studies based on a prediction or, more specifically, prediction based on context awareness. Future studies can be carried out to comparison between the results obtained in the works that use prediction to determine which technique is being most effective in Agriculture. This effectiveness in agriculture can be achieved through a higher production yield in the same growing space, use of less manual labor, more automated systems, among other improvements.

Another point to be studied is the analysis of sensor types that make the prediction more effective, for example, the use of soil moisture and water's pH sensors, in order to determine which information would be most relevant to the prediction. With this, prediction can provide inferences with greater precision aiming at decision support systems, among other aids that the prediction can provide.

Acknowledgments

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001, Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul - Fapergs/Brasil (http://www.fapergs.rs.gov.br), and Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq/Brasil (http://www.cnpq.br). We would also like to thank the University of Vale do Rio dos Sinos - Unisinos (http://www.unisinos.br).

Corresponding authors
Bruno Guilherme Martini
Applied Computing Graduate Program, University of Vale do Rio dos Sinos, UNISINOS
São Leopoldo, Brazil
Phone: +55 51 999930710, E-mail: brunogui92@gmail.com

References

[1] Alipio, M. I., Earl, A., Dela Cruz, M., Doria, J. D. A., and Fruto, R. M. S. (2017) "A Smart Hydroponics Farming System Using Exact Inference in Bayesian Network", *6th Global Conference on Consumer Electronics (GCCE 2017)*, pp. 2-6. ISBN 978-1-5090-4045-2. DOI 10.1109/GCCE.2017.8229470.

[2] Arakeri, M. P., Barsaiya, V. K. B. P. and Sairam, H. V. (2017) "Computer Vision Based Robotic Weed Control System for Precision Agriculture", *International Conference on Advances in Computing, Communications and Informatics (ICACCI),* pp. 1201-1205. ISBN 978-1-5090-6367-3. DOI 10.1109/ICACCI.2017.8126005.

[3] Athani, S., Tejeshwar, C. H. (2017) "Soil moisture monitoring using IoT enabled arduino sensors with neural networks for improving soil management for farmers and predict seasonal rainfall for planning future harvest in North Karnataka – India", *International conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, pp. 43-48. ISBN 978-1-5090-3243-3. DOI 10.1109/I-SMAC.2017.8058385.

[4] Bischoff, V., Oliveira, K. S. F., Gonçales, L. J. and Barbosa., J. L. V. (2018) "Integration of Feature Models: A Systematic Mapping Study", *Information and Software Technology*, pp. 1-30. ISSN 0950-5849. DOI 10.1016/j.infsof.2018.08.016.

[5] Carrasquilla-Batista, A., Chacón-Rodrígues, A., Solórzano-Quintana, M. (2016) "Using IoT resources to enhance the accuracy of overdrain measurements in greenhouse horticulture", *CONCAPAN*, pp. 1-5. ISBN 978-1-4673-9578-6. DOI 10.1109/CONCAPAN.2016.7942345.

[6] de la Concepcion, R. A, Stefanelli, R. and Trinchero, D. (2014) "A Wireless Sensor Network Platform Optimized for Assisted Sustainable Agriculture", *Global Humanitarian Technology Conference*, pp. 159-165. ISBN 978-1-4799-7193-0. DOI 10.1109/GHTC.2014.6970276.

[7] Dalmina, L., Barbosa, J. L.V., Vianna, H. D. (2019) "A systematic mapping study of gamification models oriented to motivational characteristics", *Behaviour & Information Technology*, Vol. 38, No. 11, p. 1-18. ISSN 1362-3001. DOI 10.1080/0144929X.2019.1576768.

[8] Dey, A. K., Abowd, G. D. and Salber, D. A. (2001) "A Conceptual Framework and a Toolkit for Supporting the Rapid Prototyping of Context-aware Applications", *Human-Computer Interaction*, Vol. 16, No. 2, pp. 97-166. ISSN 1532-7051. DOI 10.1207/S15327051HCI16234_02.

[9] Dias, L. P. S., Barbosa, J. L. V. and Vianna, H. D. (2018) "Gamification and serious games in depression care: A systematic mapping study", *Telematics and Informatics*, Vol. 35, No. 1, pp. 213-224. ISSN 0736-5853. DOI 10.1016/j.tele.2017.11.002.

[10] Divya, P., Sonkiya, S., Das, P., Manjusha, V. V. and Ramesh, M. V. (2014) "CAWIS: Context Aware Wireless Irrigation System", *International Conference on Computer, Communication, and Control Technology (I4CT)*, pp. 310-315. ISBN 978-1-4799-4555-9. DOI 10.1109/I4CT.2014.6914195.

[11] Eko, R., Sisyanto, N. and Kurniawan, N. B. (2017) "Hydroponic Smart Farming Using Cyber Physical Social System with Telegram Messenger", *International Conference on Information Technology Systems and Innovation (ICITSI)*, pp. 239-245. ISBN 978-1-5386-3100-3. DOI 10.1109/ICITSI.2017.8267950.

[12] Fiehn, H. B., Schiebel, L., Avila, A. F., Miller, B. and Mickelson, A. (2018) “Smart Agriculture System Based on Deep Learning“, 2nd International Conference on Smart Digital Environment (ICSDE), pp.158-165. ISBN 978-1-4503-6507-9.

[13] Fukatsu, T. (2014) “Web-based sensor network system ‘Field Servers’ for practical agricultural applications“, *International Workshop on Web Intelligence and Smart Sensing (IWWISS),* pp. 01-08. ISBN 978-1-4503-2747-3. DOI 10.1145/2637064.2637090.

[14] Goap, A., Sharmab, D., Shuklab, A. K. and Krishna, C. R. (2018) “An IoT based smart irrigation management system using Machine learning and open source technologies“, *Computers and Electronics in Agriculture*, Vol. 155, pp. 41-49. ISSN 0168-1699. DOI 10.1016/j.compag.2018.09.040.

[15] Helfer, G. A., Barbosa, J. L. V., Martini, B. G., dos Santos, R. B. and da Costa A. B.(2019) “Ubiquitous Computing in Precision Agriculture: A Systematic Review", *AGRIS on-line Papers in Economics and Informatics*, Vol. 11, No. 4, pp. 3-13. ISSN 1804-1930. DOI 10.7160/aol.2019.110401.

[16] Huong, T. T., Thanh, N. H., Thi, N. and Marshall, A. (2018) “Water and Energy-Efficient Irrigation based on Markov Decision Model for Precision Agriculture“, *IEEE Seventh International Conference on Communications and Electronics (ICCE),* pp. 51-56. ISBN 978-1-5386-3679-4. DOI 10.1109/CCE.2018.8465723.

[17] Ivanov I. and Tsvetkov, V. (2017) “Intelligent Planting“, *International Conference on Computer Systems and Technologies (CompSysTech’17)*, pp. 265-271. ISBN 978-1-4503-5234-5. DOI 10.1145/3134302.3134328.

[18] Jacob, N. K. (2017) “IoT Powered Portable Aquaponics System“, *Proceedings of the Second International Conference on Internet of things, Data and Cloud Computing (ICC)*, pp. 0-5. ISBN 978-1-4503-4774-7. DOI 10.1145/3018896.3018965.

[19] Joshi, J., Polepally, S., Kumar, P., Samineni, R., Rahul, S. R., Sumedh, K., Tej, D. G. K. and Rajapriya, V. (2017) “Machine Learning Based Cloud Integrated Farming“, *International Conference on Machine Learning and Soft Computing*, pp. 1-6. ISBN 978-1-4503-4828-7. DOI 10.1145/3036290.3036297.

[20] Kokkonis, G., Kontogiannis, S. and Tomtsis, D. (2017) “FITRA - A Neuro-fuzzy computational algorithm approach based on an embedded water planting system“, *International Conference on Internet of things, Data and Cloud Computing (ICC)*, pp. 0-7. ISBN 978-1-4503-4774-7. DOI 10.1145/3018896.3018934.

[21] Kubicek, P., Kozel, J., Stampach, R. and Lukas, V. (2013) “Prototyping the visualization of geographic and sensor data for agriculture“, *Computers and Electronics in Agriculture*, Vol. 97, pp. 83-91. ISSN 0168-1699. DOI 10.1016/j.compag.2013.07.007.

[22] Kumar, S., Gogul, I., Raj., D., Pragadesh, S. K. and Sarathkumar, S. (2016) “Smart Autonomous Gardening Rover with Plant Recognition using Neural Networks“, *International Conference On Advances In Computing & Communications (ICACC)*, Vol. 93, pp. 975-981. ISSN 1877-0509. DOI 10.1016/j.procs.2016.07.289.

[23] López-Riquelme, J. A., Pavón-Pulido, N., Navarro-Hellín, H., Soto-Valles, F. and Torres-Sáncheza, R. (2017) “A software architecture based on FIWARE cloud for Precision Agriculture“, *Agriculture Water Management*, Vol. 183, pp. 123-135. ISSN 0378-3774. DOI 10.1016/j.agwat.2016.10.020.

[24] Luimula, M., Shelby, Z., Tervonen, J., Markkula, J., Weckström, P. and Verronen, P. (2009) “Developing Geosensor Network Support for Locawe Platform – Application of Standards in Low-Rate Communication Context“, *International Conference on Pervasive Services*, pp. 73-82. ISBN 978-1-60558-644-1. DOI 10.1145/1568199.1568211.

[25] Ma, J., Li, X., Wen, H., Fu, Z. and Zhang, L. (2015) “A key frame extraction method for processing greenhouse vegetables production monitoring video“, *Computers and Electronics in Agriculture*, Vol. 111, pp. 92-102. ISSN 0168-1699. DOI 10.1016/j.compag.2014.12.007.

[26] Manzatto, C. V., Bhering, S. B. and Simões, M. (1999) "Agricultura de precisão: propostas e ações da Embrapa solos. EMBRAPA Solos", *EMBRAPA Solos*., Jun. 1999. [Online] Available: http://www.cnps.embrapa.br/search/pesqs/proj01/proj01.html [Acessed: 10 Sep. 2019].

[27] Mehra, M., Saxena, S., Sankaranarayanan, S., Tom, R. J.and Veeramanikandan, M. (2018) "IoT based hydroponics system using Deep Neural Networks", *Computers and Electronics in Agriculture*, Vol. 155, pp. 473-486. ISSN 0168-1699. DOI 10.1016/j.compag.2018.10.015.

[28] Nagini, S., Rajini Kanth, T. V. and Kiranmayee, B. V. (2016) "Agriculture yield prediction using predictive analytic techniques", *International Conference on Contemporary Computing and Informatics (IC3I)*. ISBN 978-1-5090-5256-1. DOI 10.1109/IC3I.2016.7918789.

[29] Park, J., Choi, J., Lee, Y. and Min, O. (2017) "A Layered Features Analysis in Smart Farm Environments", *Proceedings of the International Conference on Big Data and Internet of Thing (BDIOT)*, pp. 169-173. ISBN 978-1-4503-5430-1. DOI 10.1145/3175684.3175720.

[30] Patil S. S. and Thorat, S. A. (2016) "Early Detection of Grapes Diseases Using Machine Learning and IoT", *Second International Conference on Cognitive Computing and Information Processing (CCIP)*. ISBN 978-1-5090-1025-7. DOI 10.1109/CCIP.2016.7802887.

[31] Petersen, K., Vakkalanka, S. and Kuzniarz, L. (2015) "Guidelines for conducting systematic mapping studies in software engineering: An update", *Information and Software Technology*, Vol. 64, pp. 1-18. ISSN 0950-5849. DOI 10.1016/j.infsof.2015.03.007.

[32] Pimentel, D., Berger, B., Filiberto, D., Newton, M., Wolfe, B., Karabinakis, E., Clark, S., Poon, E., Abbett, E. and Nandagopal, S. (2004) "Water resources: agricultural and environmental issues", *BioScience,* pp. 909-918. ISSN 0006-3568. DOI 10.1641/0006-3568(2004)054[0909:WRAAEI]2.0.CO;2.

[33] Plazas, J. A. P., Gaona-García, P. A. and Marin, C. E. M. (2018) "Proposal of a computational intelligence prediction model based on Internet of Things technologies", *IEEE International Conference on Smart Internet of Things*, pp. 186-191. ISBN 978-1-5386-8543-3. DOI 10.1109/SmartIoT.2018.00041.

[34] Popovic, T., Latinovic´, N., Pešic´, A., Zecevic´, Z., Krstajic´, B. and Djukanovic´, S. (2017) "Architecting an IoT-enabled platform for precision agriculture and ecological monitoring: A case study", *Computers and Electronics in Agriculture,* Vol. 140, pp. 255-265. ISSN 0168-1699. DOI 10.1016/j.compag.2017.06.008.

[35] Rajendrakumar, S., Rajashekarappa, Parvati, V. K., Parameshachari, B. D., Sunjiv Soyjaudah, K. M. and Banu, R. (2017) "An Intelligent Report Generator for Efficient Farming", *International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques (ICEECCOT)*. ISBN 978-1-5386-2361-9. DOI 10.1109/ICEECCOT.2017.8284640.

[36] Rodríguez, S., Gualotuña, T. and Grilo, C. (2017) "A System for the Monitoring and Predicting of Data in Precision Agriculture in a Rose Greenhouse Based on Wireless Sensor Networks", *Procedia Computer Science*. Vol. 121, pp. 306-313. ISSN 1877-0509. DOI 10.1016/j.procs.2017.11.042.

[37] Rupanagudi, S. R., Ranjani, B. S., Prathik Nagarj, Bhat, V. G. and Thippeswamy, G. (2015) "A Novel Cloud Computing based Smart Farming System for Early Detection of Borer Insects in Tomatoes", *International Conference on Communication, Information & Computing Technology (ICCICT)*. ISBN 978-1-4799-5522-0. DOI 10.1109/ICCICT.2015.7045722.

[38] Saha, S., Paul, S., Halder, S. and Majumder, K. (2017) "Smart Agricultural System : Better Accuracy and Productivity", *Devices for Integrated Circuit (DevIC)*, pp. 23-24. ISBN 978-1-5090-4724-6. DOI 10.1109/DEVIC.2017.8073960.

[39] Santos, U. J. L., Pessin, G., da Costa, C. A., da Rosa Righi, R. (2018) "AgriPrediction: A proactive internet of things model to anticipate problems and improve production in agricultural crops", *Computers and Electronics in Agriculture*, Vol. 161, pp. 202-213. ISSN 0168-1699. DOI 10.1016/j.compag.2018.10.010.

[40] Sarangdhar, A. A. and Pawar, P. V. R.(2017) “Machine Learning Regression Technique for Cotton Leaf Disease Detection and Controlling using IoT“, *International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, pp. 449-454. ISBN 978-1-5090-5686-6. DOI 10.1109/ICECA.2017.8212855.

[41] Shahriar S. and Mcculluch, J. (2014) “A Dynamic Data-driven Decision Support for Aquaculture Farm Closure“, *Procedia Computer Science*, Vol. 29, pp. 1236-1245. ISSN 1877-0509. DOI 10.1016/j.procs.2014.05.111.

[42] Tan, L. (2016) “Cloud-based Decision Support and Automation for Precision Agriculture in Orchards“, *IFAC-PapersOnLine*, Vol. 49, No. 16, pp. 330-335. ISSN 2405-8963. DOI 10.1016/j.ifacol.2016.10.061.

[43] Tan, W., Zhao, C. and Wu, H. (2016) “Intelligent alerting for fruit-melon lesion image based on momentum deep learning“, *Multimed. Tools and Application*, pp. 16741-16761. ISSN 1380-7501. DOI 10.1007/s11042-015-2940-7.

[44] Treboux, J. and Genoud, D. (2018) “Improved Machine Learning Methodology for High Precision Agriculture“, *Global Internet of Things Summit (GIoTS)*, pp. 1-6. ISBN 978-1-5386-6451-3. DOI 10.1109/GIOTS.2018.8534558.

[45] Winiczenko, R., Górnicki, K., Kaleta, K. and Janaszek-Mańkowska, M. (2016) “Optimisation of ANN topology for predicting the rehydrated apple cubes colour change using RSM and GA“, *Neural Computing & Applications*, Vol. 30, pp. 1795-1809. ISSN 0941-0643. DOI 10.1007/s00521-016-2801-y.

[46] Xu, D., Li, D., Fei, B., Wang, Y. and Peng, F. (2014) “A GPRS-Based Low Energy Consumption Remote Terminal Unit for Aquaculture Water Quality“, *International Federation for Information Processing (IFIP)*. pp. 492-503. DOI 10.1007/978-3-642-54341-8_52.

[47] Yahata, S., Onishi, T., Yamaguchi, K., Ozawa, S., Kitazono, J., Ohkawa, T., Yoshida, T. and Muraka, N. (2017) “A Hybrid Machine Learning Approach to Automatic Plant Phenotyping for Smart Agriculture“, *International Joint Conference on Neural Networks (IJCNN)*, pp. 1787-1793. ISSN 2161-4407. DOI 10.1109/IJCNN.2017.7966067.

[48] Zewge, A. and Dittrich, Y. (2017) “Systematic Mapping Study of Information Technology for Development in Agriculture (The Case of Developing Countries)“, *Electronic Journal of Information Systems in Developing Countries (EJISDC)*, pp. 1-25. ISSN 1681-4835. DOI 10.1002/j.1681-4835.2017.tb00602.x.

[49] Zhang, P., Zhang, Q., Liu, F., Li, J. Cao, N. and Song, C. (2017) “The Construction of the Integration of Water and Fertilizer Smart Water Saving Irrigation System Based on Big Data“, *International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC)*, pp. 392-397. ISBN 978-1-5386-3221-5. DOI 10.1109/CSE-EUC.2017.258.

[50] Zhou, B. and Li, L. (2017) “Security monitoring for intelligent water-saving precision irrigation system using cloud services in multimedia context“, *Multimedia Tools Applications*, pp 1-15. ISSN 1380-7501.

[51] Zhou, L., Song, L., Xie, C. and Zhang, J. (2012) "Applications of Internet of Things in the Facility Agriculture", In Li D., Chen Y. (eds) "Computer and Computing Technologies in Agriculture VI", CCTA 2012. *IFIP Advances in Information and Communication Technology*, Vol 392. ISSN 1868-4238. DOI 10.1007/978-3-642-36124-1_36.

[52] Zhou, H., Liu, B. and Dong, P. (2011) “The Technology System Framework of the Internet of Things and Its Application Research in Agriculture, In Li D., Chen Y. (eds) Computer and Computing Technologies in Agriculture V. CCTA 2011, *IFIP Advances in Information and Communication Technology*, Vol 368. E-ISBN 978-3-642-27281-3, ISBN 978-3-642-27280-6. DOI 10.1007/978-3-642-27281-3_35.

Agris on-line Papers in Economics and Informatics

Volume XII Number 3, 2020

Economic Aspects of Precision Agriculture Systems

Ludmila Pánková[1], Renata Aulová[1], Jan Jarolímek[2]

[1] Department of Economics, Faculty of Economics and Management, Czech University of Life Sciences Prague, Czech Republic

[2] Department of Information Technologies, Faculty of Economics and Management, Czech University of Life Sciences Prague, Czech Republic

Abstract

The paper deals with an economic assessment of impacts of precision agriculture (PA) on crop production economy. Based on a questionnaire survey and a FADN agricultural product expense-to-revenue ratio survey, it analyses a set of agricultural businesses the structure of which essentially copies the composition of business forms in the Czech Republic's agricultural sector. The economic assessment applies economic analysis methods based on cost calculations and a calculation formula that considers the commodity and species production structure. Based on an analysis of a number of scientific studies, it determines specific cost savings and makes a quantification of the effect of precision agriculture techniques on costs. In all the production areas, the greatest effect caused by application of precision agriculture techniques was quantified for winter wheat. Conversely, the lowest financial effects are shown in the analysed production areas for spring wheat. We also identified differences in the cost savings between spring and winter barley; the greater savings occur for winter barley. Financial effects in the form of reduced production costs were also found for other analysed crops cultivated by the businesses studied. The financial savings for the pea plant are almost comparable to those for winter barley. The greatest financial savings were achieved for sugar beet.

Keywords

Precision agriculture, economic savings, techniques, work operations, costs, calculation formula.

Pánková, L., Aulová, R. and Jarolímek, J. (2020) "Economic Aspects of Precision Agriculture Systems", *AGRIS on-line Papers in Economics and Informatics*, Vol. 12, No. 3, pp. 59-67. ISSN 1804-1930. DOI 10.7160/aol.2020.120306.

Introduction

Precision agriculture can be characterised as a solution leading to reduction in agrochemical inputs and reduction of adverse environmental impacts of agriculture, where the basic benefits for the farmer are seen in the economic area (reduced costs by means of controlled application of agricultural inputs), in increased yields (targeted management of field variability) and, last but not least, a favourable environmental impact in the sense of precise application of agrochemical products (Kendall et al., 2017).

Precision agriculture is one of the ways to increase competitiveness of Czech agriculture while also better combining application of scientific results and techniques directly in agricultural businesses. It thus helps eliminate the weaknesses of Czech agriculture (reduction of production costs in particular) and contributes to increasing profitability/competitiveness of businesses.

The investment and capital costs of machinery used for precision agriculture are very different (Robertson et al., 2008; Vogt, 2017). Some technologies (e.g. auto-steer or yield mapping) are usually a standard equipment of new machines and mean very low capital costs. Some new technologies (e.g. GreenSeeker technology, camera spraying technologies) are associated with higher investment.

The precision agriculture technique brings a number of favourable effects in practice. They contribute, for example, to reduced soil compaction thanks to targeted movement of machinery on plots and more efficient traffic control methods, and bring a saving of time and costs expended on individual work operations. In summary, the techniques in question also contribute to increased labour productivity. The benefits resulting from using of PA technologies are derived from many key drivers (Vogt, 2017; Calegari et al., 2013; West and Kovacs, 2017): capital and annual operating costs associated with acquiringthe technology, impact of the technology on labour demand, impact on yield, product quality, cost savings, environmental benefits etc. Nevertheless Robertson et al. (2008) argue, that the profitability of PA and benefits from PA technologies varies from farm to farm, in line with farmer preferences and circumstances.

The literature survey indicates that precision agriculture, as a form of application of modern technologies and approaches, can be one of the paths to increased profitability of agricultural production. Moreover, the literature says that precision agriculture can be seen as a management method that is more environmentally friendly (due to not only targeted application of fertilisers or plant protection products but also the soil management method and control of machinery traffic on plots to increase soil quality, reduce soil compaction, etc.). A number of studies and research papers from various countries point out the positive impacts of precision agriculture (e.g., Godwin, 2002, 2015; Cassman, 1999, and others).

Many experts define the benefits of precision agriculture techniques in not only quantitative but also qualitative terms. For example, Cordesses et al. (2000), Dunn et al. (2006), Debain et al. (2000), Han et al. (2004), Kingwell (2011), Stoll and Kutzbach (2000) sum up the following general benefits of using a guidance system, for example:

- reduced driver fatigue: guidance systems reduce the efforts associated with maintaining precise paths;
- increased yields;
- reduced costs of work operations: accuracy is increased by reducing “skipping” (omissions) and “doubling” (repeated applications – overlaps) between adjacent rows in the field;
- increased productivity: higher operating speeds are possible;
- better quality: the driver can focus attention elsewhere to increase quality;
- improved safety;
- lower adverse environmental impact (reduced frequency of machine crossings, reduced soil compaction);
- ability to work at night and under reduced visibility.

Buchtel (2016) defines the impacts of precision agriculture as follows:

- time savings: evident in a number of work operations, particularly in harvesting, soil preparation and spraying;
- savings in labour costs (harvesting, soil preparation, spraying, sowing, fertilising), equipment costs, chemical plant protection products, seed stock, fertilisers, fuels;
- increased crop yields.

The benefits of precision agriculture are also confirmed by Kviz et al. (2014), who define savings due to use of RTK (Real Time Kinematic) based guidance systems, in the sense of time savings from the number of field operations in a season, or reduction in machinery crossings compared to the conventional system. The authors say that better accuracy of machinery with guidance systems in fields could help achieve energy and material savings and, to some extent, reduce machinery traffic in fields, thus improving soil conditions.

The present paper follows up on the research of the above authors and attempts a verification of the expected economic savings and their quantification.

Materials and methods

The primary objective of the present paper is to provide the simulation of precision agriculture (PA) impact on crop production economy. The primary objective is achieved by means of secondary goals:

- Description of techniques and work operations in PA;
- Study of structure of agricultural crop production costs;
- Simulation of effects of selected techniques PA on production economy.

The background data for the present paper are based on a questionnaire survey as well as a FADN agricultural product expense-to-revenue ratio survey in 2015 (ÚZEI, 2015). The representative set group consisted of 14 agricultural entities applying precision agriculture techniques in practice. The analysed set of businesses can be divided by business structure into natural persons and legal entities. Natural person businesses made up 14.3% of the agricultural entities analysed, and legal entities comprised 85.7%. The structure of the set of businesses therefore essentially copies the composition of agricultural business forms in the Czech Republic.

Among the legal entities, the most common forms were joint stock companies and limited liability companies (29% each). Another category of businesses were cooperatives (21.4%). The least represented category comprised other legal forms, specifically limited partnerships (7%).

In terms of representation of analysed agricultural entities in different agricultural production areas, it can be said that the analysed group of businesses were mostly active in the sugar beet production area, followed by the potato and maize production areas.

The data obtained in the questionnaire survey was used for the crops specification representing the analyzed group of companies on which the simulation was demonstrated. Based on the companies unavailability of detailed cost structure related to PA technology, the simulation of production costs was calculated by using agricultural commodities costing published in the FADN database, with the rate of cost savings determined by Buchtel (2016).

The economic assessment applies economic analysis methods based on cost calculations and a calculation formula using the FADN dataset 2015. The calculation formula is based on the methodology of ÚZEI (2015), applied in studying the expense-to-revenue ratio of agricultural commodity production. The calculation formula classifies agricultural commodity production costs in the following structure:

seed stock (own and purchased)
+ fertilisers (own and purchased)
+ chemical protection products
+ other direct materials
= direct material costs
+ other direct material costs
+ wage and personnel costs (direct, auxiliary activities)
+ depreciation
+ costs of auxiliary activities
+ production overheads
+ administrative overheads
= total costs

In studying the cost effect of precision agriculture techniques and economic effects following from that, this paper applies not only a questionnaire survey but also a literature survey and the structure of operating costs studied in the Czech Republic, based on the commodity and type structure. In terms of commodities, the assessment (depending on the structure of sowing procedures of the analysed businesses) included cereals (winter wheat, spring wheat, winter barley, spring barley), rape as the oil crop and other crops such as pea and sugar beet.

Taking the average structure of total costs expended per hectare of selected agricultural crops (based on results of ÚZEI, 2015) as the starting point, the average structure of the total costs consists of:

- Direct material costs
- Personnel costs
- Other direct costs including services
- Other costs

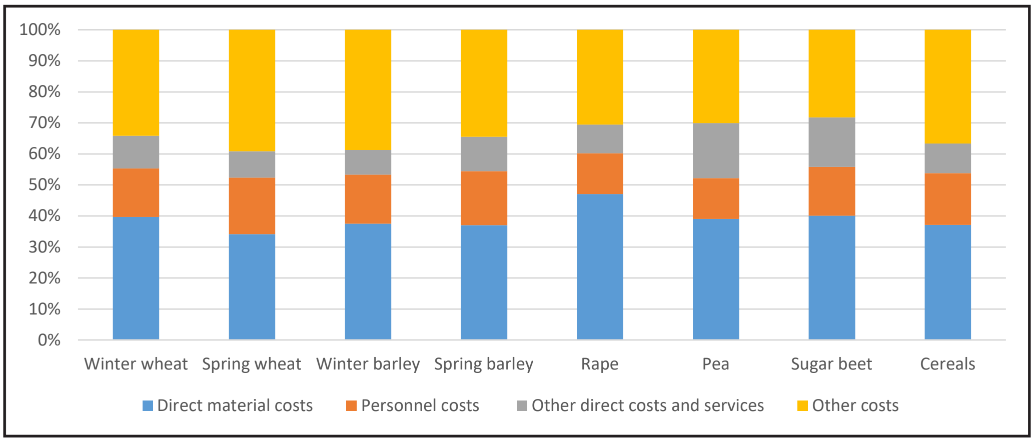
The direct material inputs are composed of the cost items for seed stock (purchased and own), fertilisers (purchased and own), plant protection products, and other direct materials (such as packaging).

The personnel cost structure comprises costs of employees including levies for social and health insurance. The other direct costs include external services, utilities, fuels, insurance premiums, rent, etc. The other cost category comprises the following sum of items: depreciation of long-term tangible and intangible assets, production and administrative overheads, and costs of auxiliary activities (e.g., costs of own machinery, repairs and maintenance).

Based on the results of Buchtel (2016), we were able to work with specific percentages of cost savings. This savings can be specified in direct cost (seed stock 1.9 %, fertilisers 2.61 %, plant protection products 6.21 %), personal costs (3.47 %) and other direct cost and services (3.4 %). By means of the figures shown in the study, we can simulate the effect of precision agriculture techniques in the area of costs by using the FADN dataset.

Results and discussion

The production cost structure is primarily assessed based on data from ÚZEI (2015). As shown in Figure 1, the total cost structure for cereal crops comprises 37% of direct material costs, and a similar proportion is for the other costs (sum of depreciation, overheads, etc.); the personnel costs are 17% of the total costs. Less than 10% is drawn by services along with the other direct costs. The differences between the selected representatives of cereals, i.e., winter and spring wheat and winter and spring barley, are not very significant. The direct material costs are the highest for winter



Source: ÚZEI (2015) and own calculations

Figure 1: Average structure of production costs of selected agricultural crops in the CR.

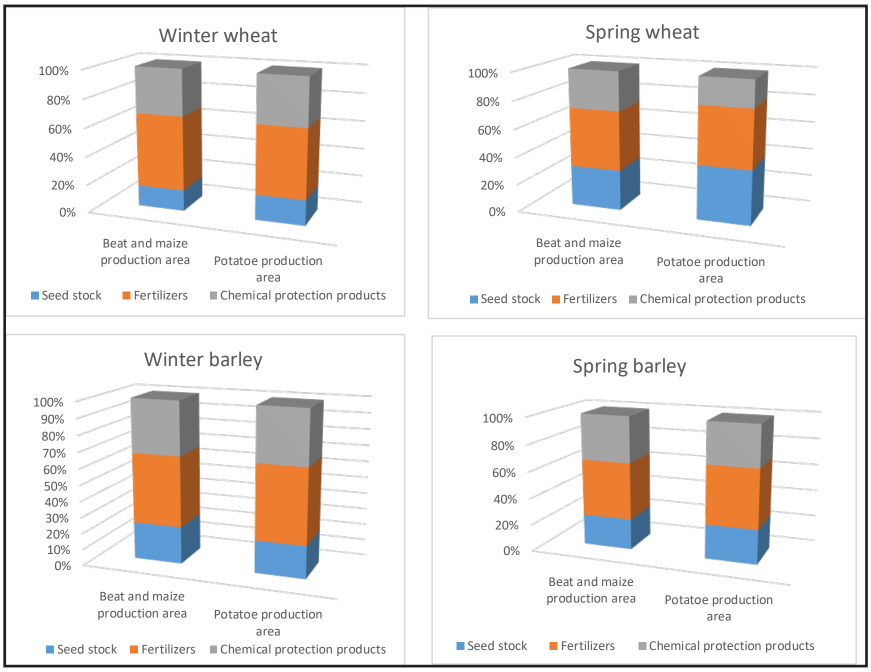
wheat (5.5% higher on average compared to spring wheat and 2.2% and 2.6% higher compared to winter barley and spring barley, respectively). The direct material costs of rape, pea and sugar beet exceed the percentage of the total costs by up to 10% compared to the cereals (rapeseed the most).

The analysis of the trends based on the ÚZEI database (2015) for costs per hectare of the agricultural crops assessed shows that the production areas partly influence not only the amount of total costs expended per hectare of crop but also the type structure of the operating costs. Whilst the total costs per hectare of all the studied crops in the potato production area are lower, their type structure shows higher direct material costs (by up to 4%) in comparison with the maize and beet production area.

The application of precision agriculture techniques is very tightly linked with the direct costs, which comprise the costs of seed stock, fertilisers and chemical protection products. This link can be observed, e.g., in the reduced consumption of direct material input, targeted application and precise dosage. In line with that, fuel consumption, labour costs, etc., also decrease. According to the ÚZEI survey, the direct material costs also include so-called other direct material costs, but they are almost negligible in terms of the total costs per hectare (e.g., their share is 0.75-2.3% of the total costs for cereals, i.e., CZK 177-556 per hectare depending on the crop and the production area).

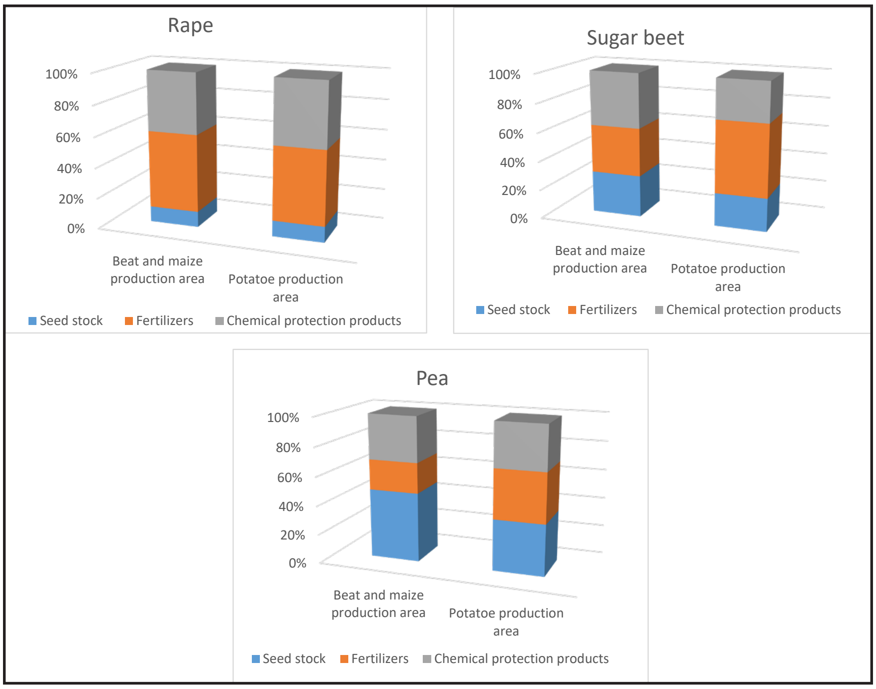
Figures 2 and 3 show the structure of the direct material costs by crop and production area. The structure of the direct material costs of cereals is dominated by costs of fertilisers (Figure 2) for all the agricultural production areas analysed. Their share in the total material costs is the highest for winter wheat, for which it is over 50% of the direct material costs in the maize and beet production area. Similarly, rape too (Figure 3) shows a high share of fertiliser costs (48-50% depending on production area). The commodities of sugar beet and pea in the maize and beet production area show a share of costs of chemical protection products higher than that of fertiliser costs. However, that is not the case in the potato production area. Among the analysed commodities, pea has the highest share of costs of seed stock (35-48% of the material costs). Rape has the lowest costs of seed stock (as a share of material costs); the share of these costs in the direct material costs does not exceed 10.5% in either production area.

A number of other facts indicated by expert studies cannot be omitted in the context of precision agriculture and in connection with the analysed structure of material costs. They confirm clearly that the size of the material costs is related not only to the price of material but also its consumption, which can be influenced significantly by the quality and method of performance of the component agrotechnical work operations. They include, for example, soil treatment and soil preparation before sowing. For example, Scarlett et al. (1997) cite the saving of EUR 14/ha in the case of following up on the quality of soil treatment and adjustment of sowing quantity based on the current condition. In another report, they say that the sowing quantity is considerably influenced by the quality of the seedbed (it can be reduced greatly without any loss of yield), and they quantify the potential for sowing cost saving at up to 70%. On the other hand, similar results (such high savings) can only be achieved with difficulty in actual practice. Malik et al. (1985) and Marchenko (1989)



Source: ÚZEI (2015) and own calculations

Figure 2: Direct material costs of selected cereal crops in the CR.



Source: ÚZEI (2015) and own calculations

Figure 3: Direct material costs of other studied crops.

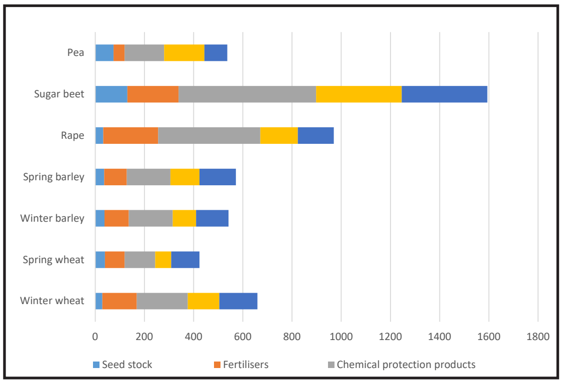
document the relationship between cloddiness of the soil surface and the seedbed, where the sowing quantity increases with the cloddiness. At the same time, it is common practice that the tractor driver checks the preparation quality and cloddiness visually on the surface of the treated soil, not by mesh analysis (soil sifting in order to determine cloddiness), which ultimately results in repeated crossings (Scarlett et al., 1997). Variable sowing depth is still not used in common precision agriculture practice either.

The primary objective of variable soil treatment is to assure smooth transition of treatment depth and intensity, soil re-compaction, etc., depending on the soil block conditions. Variable sowing systems are clearly linked to precision sowing for slim-row crops. With cereals and rape in particular, the array pitch between plants have to be observed precisely to optimise the space for each plant and eliminate intra-species competition, and the possibility to vary the sowing quantity has to be enabled in relation to the macro and micro variability of the plot. The possibility to vary the sowing quantity during a working trip opens up a new room for reducing sowing quantities, not to be applied overall but in relation to the current moisture conditions and plot variability (Brant et al., 2016).

The cost saving calculation based on Buchtel (2016) enables us to quantify the total average cost saving in precision agriculture. His research indicates that the greatest cost savings occur in pesticide costs, followed by labour cost savings. Cost savings were also proven for fuels, fertilisers and seed stock.

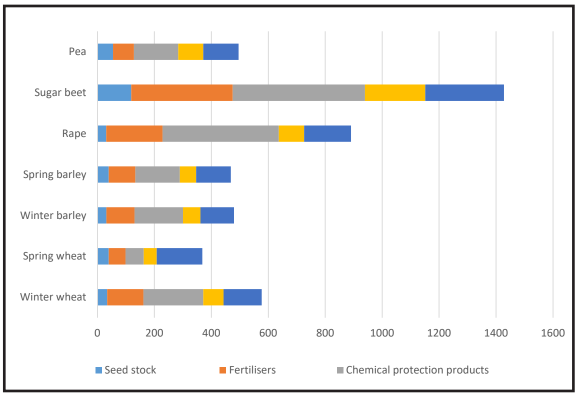
It holds for the maize and beet production areas that the greatest effect caused by application of precision agriculture techniques is quantified for winter wheat, being CZK 659/ha; the same results are achieved in the potato production area. Conversely, the lowest financial effects are shown in the analysed production areas for spring wheat. The difference in the cost savings for winter barley and spring barley in the beet and maize production areas is CZK 30/ha (higher savings achieved for spring barley); the difference is negligible in the potato production area (only CZK 11/ha, with higher savings for winter barley). The reason for these equalised results are probably the fact arising from the ÚZEI data (ÚZEI, 2015) that the production costs in the potato production area are almost comparable to that of spring barley. Financial effects in the form of reduced production costs were also found for other analysed crops cultivated by the businesses studied. The savings for winter rape are CZK 890-969/ha depending on the production area. The financial savings for the pea plant are almost comparable to those for winter barley. The greatest financial savings were achieved for sugar beet.

It follows from Figures 4 and 5 that the cost structure is dominated by the economic benefits of chemical plant protection costs. Conversely, the calculated savings are the lowest for seed stock, which is due to the share of the seed stock costs in the total costs. In the case of wage costs, it has to be noted that the resulting financial effect (cost saving) is the sum of all work operations performed as part of the cultivation technique.



Source: own calculations

Figure 4: Cost savings for selected crops applying PA techniques in maize and beet production area.



Source: own calculations

Figure 5: Cost savings for selected crops applying PA techniques in potato production area.

Application of precision agriculture techniques is related not only to savings on the cost side but, typically, also another effect in the form of increased natural production per unit of area and the related financial result in the form of revenues. However, that will be the subject matter of our further research.

Speaking, for example, about the average size of the studied legal entity businesses included in this representative survey (i.e., 3,567 hectares of farmland), the size of potential savings for each crop can be estimated when considering the tillage percentage and the existing crop area structure (see Table 1). The calculation of the theoretical savings is based on the assumption that cereal crops make up 54% of the crop succession for the businesses studied.

Crop/ Production area	Estimate of theoretical cost savings (in thousands of CZK)	
	Beet and maize production area	Potato production area
Cereals	933.6	804.3
Rape	518.7	476.3
Pea	33.8	31.2
Sugar beet	100.3	89.9

Source: own calculation

Table 1: Estimate of theoretical cost savings for selected crops (given average business size in representative survey).

Since contemporary practice shows varying degrees and levels of practical application of precision agriculture techniques, an assessment of impacts of techniques applied on the business economy of the agricultural entities will be the subject matter of our further study.

Conclusion

The precision agriculture technique brings a number of favourable effects in practice. They contribute, for example, to reduced soil compaction thanks to targeted movement of machinery on plots and more efficient traffic control methods, and bring a saving of time and costs expended on individual work operations. In summary, the techniques in question also contribute to increased labour productivity. Precision agriculture techniques offer the Czech farmer a considerable potential not only in agricultural production but also in the area of methodology, linked to data analysis and processing. Precision agriculture principles help develop an innovative approach to the traditional sector of the national economy by application of modern technologies while respecting the heterogeneity of the soil environment.

The results of the study clearly indicate a favourable practical effect of precision agriculture. It is manifested at several levels, both in the environmental and landscape areas and in the economic area (cost effect).

The primary production factor – soil – is thus perceived much more sensitively in precision agriculture, with respect, e.g., to nutrient supply to soil, necessity of machinery travel on plots in an effort to reduce soil compaction, etc. Since the structure of the agricultural commodity production costs includes a high share of costs of chemical plant protection, their economically effective expenditure is a relatively important aspect of economic prosperity of businesses.

The results obtained enable us to identify the dominant cost items that have a significant influence on the economy of agricultural commodity production. For all the crops studied, these dominant items among the cost types are, above all, the direct material inputs, the shares of which are differentiated depending on the commodity.

The economic benefit of precision agriculture is evident based on the results obtained. It is manifested in all the production areas analysed, and the effect of financial savings in production costs can be additionally enhanced, on the other hand, by the effects of precision agriculture on the revenue side, as indicated by a number of studies conducted in the area.

Corresponding authors
Ing. Ludmila Pánková, Ph.D.
Department of Economicss, Faculty of Economics and Management
Czech University of Life Sciences Prague, Kamycka 129, Prague 6, 16521, Czech Republic
Phone: +420 224 382 283, E-mail: pankoval@pef.czu.cz

References

[1] Brant, V., Kroulík, M., Zábranský, P. and Škeř, M. . (2016) "Nižší výsevky a zonální aplikace hnojiv při pěstování obilnin jako základ precizního zemědělství", *Agro journal*. [Online]. Available: https://www.agrojournal.cz/clanky/nizsi-vysevky-a-zonalni-aplikace-hnojiv-pri-pestovani-obilnin-jako-zaklad-precizniho-zemedelstvi-135 [Accessed: 12. Dec. 2019]. (In Czech).

[2] Buchtel, F. (2016) "Vyhodnocení ekonomické efektivity systému precizního zemědělství v podmínkách vybrané farmy v ČR", Bachelor thesis, University of South Bohemia in České Budějovice, Faculty of Economics, tutor: Antonín Dolan. (In Czech).

[3] Calegari, F., Tassi, D. and Vincini, M. (2013) "Economic and environmental benefits of using a spray control system for the distribution of pesticides", *Journal of Agricultural Engineering*, Vol. 44, No. 2s, pp. 163-165. E-ISSN 2239-6268. DOI 10.4081/jae.2013.274.

[4] Cassman, K. G. (1999) "Ecological intensification of cereal production systems: Yield potential, soil quality, and precision agriculture“, *Proceedings of the National Academy of Science of the United States of America*, Vol. 96, No. 11, p. 5952-5959. E-ISSN 1091-6490. DOI 10.1073/pnas.96.11.5952.

[5] Cordesses, L., Cariou, C. and Berducat, M. (2000) "Combine harvester control using Real Time Kinematic GPS“, *Precision Agriculture*, Vol. 2, No. 2, pp. 147-161. E-ISSN 1573-1618, ISSN 1385-2256. DOI 10.1023/A:1011473630247.

[6] Debain, C., Chateau, T., Berducat, M., Martinet, P. and Bonton P. (2000) "A guidance-assistance system for agricultural vehicles“, *Computers and Electronics in Agriculture*, Vol. 25, No. 1-2, pp. 29-51. ISSN 0168-1699. DOI 10.1016/S0168-1699(99)00054-X.

[7] Dunn, P. K., Powierski, A. P. and Hill, R. (2006) "Statistical evaluation of data from tractor guidance systems“, *Precision Agriculture*, Vol. 7, No. 3, pp. 179-192. E-ISSN 1573-1618, ISSN 1385-2256. DOI 10.1007/s11119-006-9007-8.

[8] Godwin, R., Earl, R., Taylor, J. C., Wood, G. A., Bradley, R. I., Welsh, J. P., Richards, T. and Blackmore, B. S. (2002) "*Precision farming of cereal crops: A five-year experiment to develop management guidelines*“, AHDB Cereals & Oilseeds [Online]. Available: https://cereals.ahdb.org.uk/media/288142/pr267-final-project-report.pdf [Accessed: 13 Dec. 2019].

[9] Godwin, R. (2015) "Precision farming“, *INGENIA online*, No. 64. [Online]. Available: http://www.ingenia.org.uk/Content/ingenia/issues/issue64/godwin.pdf [Accessed: 13 Dec. 2019].

[10] Han, S., Zhang, Q., Ni, B. and Reid, J. F. (2004) "A guidance directrix approach to vision-based vehicle guidance systems“, *Computers and Electronics in Agriculture*, Vol. 43, pp. 179-195. ISSN 0168-1699. DOI 10.1016/j.compag.2004.01.007.

[11] Heiniger, R. W. and Meijer, A. J. (2000) "Why variable rate application of lime has increased grower profits and acceptance of precision agriculture in the southeast“, In *SO: Proceedings of the 5th International Conference on Precision Agriculture,Bloomington, Minnesota USA,16. - 19. July, 2000. publ 2001*; pp. 1-17. Madison, USA: American Society of Agronomy.

[12] Kendall, H. et al. (2017) "Precision Agriculture in China: Exploring Awareness, Understanding, Attitudes and Perceptions of Agricultural Experts and End-Users in China“, *Advances in Animal Biosciences: Precision Agriculture (ECPA),* Vol. 8, No. 2. p. 703-707. E- ISSN 2040-4719, ISSN 2040-4700. DOI 10.1017/S2040470017001066.

[13] Kingwell, R. (2011) "The whole-farm benefits of controlled traffic farming: An Australian appraisal“, *Agricultural systems*, Vol. 104, No. 7, pp. 513-521. ISSN 0308521x. DOI 10.1016/j.agsy.2011.04.001.

[14] Kviz, Z., Kroulik, M. and Chyba, J. (2014) "Evaluation of Machinery Guidance Systems Concerning Pass-to-Pass Accuracy as a Tool for More Efficient Plant Production in Fields and Soil Damage Reduction“, *Plant, Soil and Environment*, Vol. 60, No. 1, pp. 36-42. E-ISSN 1805-9368. DOI 10.17221/622/2012-PSE.

[15] Malik, R. S., Jhorar, B. S. and Dahiya, I. S. (1985) "Influence of seedbed tilth on emergence and root and shoot growth of seedling of some crops“, *Experimental Agriculture*, Vol. 21, No. 1, pp. 59-65. E-ISSN 1469-4441, ISSN 0014-4797. DOI 10.1017/S0014479700012254.

[16] Marchenko, O. S. (1989) "Optimising soil cultivation and seedbed preparation parameters referring to soilconditions“, *Land and Water Use*, P*roceedings of the 11th International Congress on Agricultural Engineering, Dublin,* (Dodd V. A; Grace P. M.), Balkema, Rotterdam, Vol. 2, pp. 1507-1517. ISBN 90 6191 977 0.

[17] Robertson, M., Carberry, P. and Brennan, L. (2008) "*Economic benefits of precision agriculture. Case studies from Australian grain farms*“, Grain Researd and Development Corporatio, 14 Feb, 2008 [Online]. Available: https://grdc.com.au/resources-and-publications/grdc-update-papers/tab-content/grdc-update-papers/2008/02/economic-benefits-of-precision-agriculture-case-studies-from-australian-grain-farms [Accessed: 10 Nov. 2019].

[18] Scarlett, A. J., Lowec, J. C. and Semple, D. A. (1997) "Precision tillage: in-field, real-time control of seedbed quality“, *Precision Agriculture*, *Proceedings of the 1st European Conference on Precision Agriculture*, BIOS Scientific Publishers Ltd, 1997, p. 503 - 510. ISBN-13 978-1859962367.

[19] Sparovek, G. and Schnug, E. (2001) "Soil tillage and precision agriculture - A theoretical case study for soil erosion control in Brazilian sugar cane production“, *Soil & Tillage Research*, Vol. 61, No. 1-2, p. 47-54. ISSN 0167-1987. DOI 10.1016/S0167-1987(01)00189-1.

[20] Stoll, A. and Kutzbach, H. D. (2000) "Guidance of a forage harvester with GPS“, *Precision Agriculture*, Vol. 2, No. 3, pp. 281-291. E-ISSN 1573-1618, ISSN 1385-2256. DOI 10.1023/A:1011842907397.

[21] Vogt, S. (2017) "*The economics of precision agriculture*“, Grain Researd and Development Corporatio, 07 Feb, 2017 [Online]. Available: https://grdc.com.au/resources-and-publications/grdc-update-papers/tab-content/grdc-update-papers/2017/02/the-economics-of-precision-agriculture [Accessed: 10 Nov. 2019].

[22] UZEI (2015) "*Dataset FADN*", [Online]. Available: UZEI: http://www.vsbox.cz/fadn/ [Accessed: 20. Feb. 2019].

[23] West, G. H. and Kovacs, K. (2017) "Addressing Groundwater Declines with Precision Agriculture: An Economic Comparison of Monitoring Methods for Variable-Rate Irrigation", *Water*, Vol. 9, No. 1. ISSN 2073-4441. DOI 10.3390/w9010028.

Parametric Insurance as Innovative Development Factor of the Agricultural Sector of Economy

Olena Prokopchuk[1], Ihor Prokopchuk[2], Grzegorz Mentel[3], Yuriy Bilan[4]

[1] Department of Finances, Banking and Insurance, Faculty of Economics and Entrepreneurship, Uman National University of Horticulture, Ukraine

[2] Department of Agrochemistry and Soil Science, Faculty of Agronomy, Uman National University of Horticulture, Ukraine

[3] Department of Quantitative Methods, Faculty of Management, Rzeszow University of Technology, Poland

[4] Centre of Applied Economic Research, Tomas Bata University in Zlin, Czech Republic

Abstract

In the article a parallel between the classical and parametric scheme of agricultural risk insurance is conducted. The application aspects of parametric (index) schemes of insurance with emphasis on the use of weather index insurance products are examined, their advantages and disadvantages are considered.

This research examined the applicability along with simple weather index insurance products combined, that can consider and put together a few parameters simultaneously and thus neutralize the impact of the whole weather risks at regional level. The authors demonstrated the feasibility of using the proposed combination of weather index (C_i) – a special indicator which characterize the impact of weather risk combitation intensity, measured by certain parameters (heat, cold → temperature; air humidification → relative humidity; drought → precipitation) on the grain maize yield in definite growth stages (flowering and grain filling). On the basis of research, the detail mechanism proposed by the authors of the combined weather index (Ci) in general, and on the example of concreate calculations, is performed in particular.

Keywords

Insurance, agricultural insurance, classical insurance scheme, parametric insurance scheme, insurance products, index, weather risk, weather conditions.

Prokopchuk, O., Prokopchuk, I., Mentel, G. and Bilan, Y. (2020) “Parametric Insurance as Innovative Development Factor of the Agricultural Sector of Economy", *AGRIS on-line Papers in Economics and Informatics*, Vol. 12, No. 3, pp. 69-86. ISSN 1804-1930. DOI 10.7160/aol.2020.120307.

Introduction

One of the key features of agriculture is its dependence on natural conditions. Along with price volatility, the presence of commercial risk and raiding, Ukrainian farmers are also suffering from weather freaks. Insurance is called to minimize losses from the latter. That is a system of economic relations to guarantee insurance protection in order to ensure the income stability and the available property conservation (Pointer and Khoi, 2019).

As a result, insurance of agricultural products (Ukrainian agricultural insurance system) is one of the key components of such a system that provides risk management in the agricultural sector and contributes to creating the basis for sustainable economic growth in agricultural production.

At the same time, the agricultural insurance market in Ukraine is still undeveloped, and insurance products are not well-received among farmers, despite the tendency to expand the spectrum and quality of the latter in the context of classical and parametric insurance schemes.

Consequently, a certain imbalance arose - a powerful development of the agri-industrial complex in Ukraine in recent years has not been accompanied by the proper development of the insurance market, which goes against the generally established international practice. After all, the majority of countries in the world where agriculture is successfully developing (the USA, Canada, Spain, China, etc.), are characterized by a high level of agricultural insurance development.

The fact that Ukraine belongs to the risky farming zone largely determines the dependence of agricultural producers' performance on the influence of weather patterns. Insurance based on weather indices is a real leverage over weather risks, that is, patterns that really reduce the crop yield.

Weather index insurance principles were initiated by Halcrow (1948) and further developed by Dandekar (1977), Skees et al. (1999) theoretically proposed these principles for developing countries and later on empirically tested in Moroco (Skees et al., 2001). Mahul (2001) provided a more formal framework for weather index insurance in agriculture. Using historical precipitation and temperature data, Turvey (2001) illustrated how weather index insurance could be used to address specific-event risks measured at the local level and how preIcpitation and heat insurance could be priced in practice. Weather index insurance provides protection for vulnerable households against specific weather shocks (e.g., precipitation shortage or flooding) by using historical precipitation, yield and related agricultural and weather data (Osgood et al., 2007; Barnett et al., 2008; Gebeltova et al., 2020). A contract is signed between an insurer and a policy holder in advance. Payment is based on the observable intensity of a weather index closely correlated to the yield rather than to the non-observable specific impact on it (Mahul, 2001). The policy holder will get a payout when, for instance, precipitation is below an agreed amount (trigger point) based on automated weather station records. Weather insurance thereby needs evidence of the statistical relationship between the specific weather event being measured and volumetric risk (e.g., crop loss) at the farm level.

Different aspects of the functioning of such an important tool as agricultural insurance were studied by the number of national scientists and practitioners, among them are Martseniuk-Rozaronova (2010); Navrotskyi (2012), Nesterchuk (2018), Polchanov (2013) and others. In their research, the scientists studied the main development trends of the Ukrainian agraricultural insurance market in general and some problematic aspects of its development, generalized historical aspects of its formation and establishing, studied the specific features of insurance legislation, and found out the specific character of the insurance market. Ukrainian scientists researched insurance products and insurance schemes that provide insurance protection of farmers (Nesterchuk et al., 2018).

Some theoretical and practical aspects of agricultural insurance were profoundly investigated by foreign scientists, among them the following should be mentioned: Santeramo (2018); Shiferaw et al. (2015); Barnett and Skees (2008).

It is necessary to mention that foreign scientists provide a profound and elaborated analysis of insurance programs, used in the agricultural insurance. This aspect was covered by foreign researchers in the context of classical, parametric and specific insurance schemes: Fuchs and Wolff (2011), Chantarat et al. (2007); Dalhaus, Musshoff and Finger (2018), Tadesse et al. (2015), Turvey (2001), Odening and Shen (2014).

A great body of scientific literature deals with parametric schemes in the agricultural insurance, that is the index insurance products for agriculture. This aspect is sufficiently and intimately researched by the following foreign scientists Barnett et al. (2007), Weber et al. (2015), Sarris (2013), Micntosh et al. (2013), Kocisova et al., 2018, Morkūnas et al. (2018), Walewicz (2018), Malyovanyi et al. (2018), Grmanová and Strunz (2017).

These are key potential advantages of weather index insurance over traditional crop insurance schemes, but many pilot projects show that significant improvements in product design and implementation strategies are still needed, especially to enhance demand and uptake.

At the same time, there are a number of problems related to the functioning of parametric insurance schemes, in particular at the level of developed and developing countries, which are described in sufficient detail by foreign researchers. First of all, it concerns the problem of reliable and systematic data availability.

Thus, Weber et al. (2015) deal with the common problem of developing countries in which yield data is often only available on an aggregated level, and weather data is only accessible for a low number of weather stations. The insurance products are modeled for different inter-regional and intra-regional risk aggregation and risk coverage scenarios.

The findings suggest that index-based weather insurance products bear a large risk mitigation potential on an aggregated level. As a result, meso-level insurance should be recognized by institutions with a regional exposure to cost-related weather risks as part of their risk-management strategy.

Despite continued pilot testing of weather index insurance products in low income countries, its actual uptake has been far below expectations (Micntosh et al., 2013; Walewicz, 2018). The high price (premium) and lack of trust in the index and its ability to properly predict the risk of loss as well as the credibility of the insurance providers are key factors negatively influencing the demand for weather index insurance (Barnett and Skees, 2008). Turvey (2001) shows that the low demand for weather index insurance by poor farmers is a rational response to basis risk. For higher uptake, weather index insurance should be cheaper than the current risk management practices of smallholders, such as reliance on social networks and self-insurance mechanisms by owning assets (Micntosh et al., 2013). Others argue that promoting access to productive assets (e.g., land), credit, improved seeds, better agricultural practices and rural infrastructure are the key factors for the poor to build their own capital to self-insure in case of disaster. However, it is difficult for the poor to build capital for climate risk management and risk transfers (insurance) in the short run as current consumption competes with future savings. Insurance and credit instruments for smallholders may need to be seen as essential complements rather than substitutes for risk-reducing and profitable technological innovations. This is particularly relevant as technological solutions will be inadequate in managing production risks caused by severe climate shocks and extreme events, expected under progressive climate change.

At the same time, nowadays the Ukrainian market of agricultural insurance is still undeveloped, and insurance products are not in a great demand among farmers. Moreover, parametric insurance products, at the Ukrainian insurance market are innovative and now they are at the introduction stage. Considering this, the use of innovative practices and methods in the agricultural insurance at the current stage requires further heightened attention.

Aims

The purpose of the paper is to investigate the application of parametric (index) schemes of insurance in the agricultural sector of economy with a focus on weather index insurance products, outlining their advantages and disadvantages and, from this, justification of the feasibility of using the authors' proposed combined weather index (Ci), which consider and combine several parameters at a time and, thus, is able to minimize the impact of the combination of weather risks at the regional level.

Materials and methods

The theoretical and methodological basis of research is the works of domestic scientists and practitioners, as well as foreign ones involved in researching the market of agricultural insurance services in general and the issues of developing an effective system of agricultural insurance, in particular. The following methods of research as an economic-statistical, abstract-logical, graphic and scientific generalization are used.

To achieve the stated goal, the following general scientific methods of scientific cognition and research of economic phenomena were used, in particular the dialectical cognition method was used to analyze scientific research papers on the national and foreign issues of effective functioning of the insurance market in general and the research of available insurance services for agricultural producers.

The empirical research method was used to evaluate the current state and development of the object of our research.

To compare the key insurance schemes, in particular in the context of classical and parametric insurance products, we applied comparative analysis. Tabular and graphical methods were applied to systematize and visualize numerical material. Abstract-logic approach allows to generalize the findings of the research and to draw the conclusions.

In practice, the economic-statistical method is used in the study of mass phenomena, processes, facts and the identification of trends and patterns of their development. It also makes it possible to determine the quantitative effect of individual factors on the result of the study and identify the main factors that led to changes in the course of economic processes. In this study, this method is one of the main methods. After all, using the method of mathematical statistics, namely, multivariate analysis of variance, the justification of individual parameters influence (the average monthly air temperature, precipitation amount and relative humidity) on the yield of the agricultural crop is carried out (corn in Uman district of Cherkasy region). The main features of this method application in this study are described below.

Calculation and usage patterns of the combined weather index (Ci) from a regional perspective

In the framework of this research, we will consider the feasibility of using a combined weather index, which put several parameters together, so-called triggers, in terms of Uman district of Cherkasy region. The source of the basic data describing agri-meteorological conditions of crops vegetation (in this case, grain maize) is the observation results of weather conditions conducted at Uman meteorological station, located in Uman district of Cherkasy region (period: in 2002-2017 timeframe, that is, an interval of 15 years).

In this case, the combined weather index is used, in particular, the boundary values of the three parameters: the amount of precipitation, relative humidity and the air temperature during the critical periods of crop vegetation (flowering and grain filling – June-July of the month), which determine the critical values of corn yield for grain, are used.

The grain maize yield data for this research were based on a long stationary field experiment (conducted since 1964) by the Department of Agrochemistry and Soil Science of Uman National University of Horticulture, based on a 10-field crop rotation (spring barley + clover, clover, winter wheat, sugar beet, maize, peas, winter wheat, maize for silage, winter wheat, sugar beet). In the crop rotation organic fertilizers (manure 9 t, 13.5 t, 18 t), mineral ($N_{45}P_{45}K_{45}$; $N_{90}P_{90}K_{90}$; $N_{135}P_{135}K_{135}$) and organo-mineral (manure 4.5 t + $N_{22}P_{34}K_{18}$; manure 9 t + $N_{45}P_{68}K_{36}$; manure 13.5 t + $N_{67}P_{102}K_{54}$) are applied in crop rotation (Table 1). One of the intermediate stages of further research will be the separation from the above set of fertilizer systems optimal and the subsequent research stages based on the indicators of the latter.

The values of weather parameters (precipitation, relative humidity, air temperature) are taken from the data of Uman meteorological station.

For grain maize, the critical period of time is the flowering and grain filling. This period lasts about two months (usually in June and July). At the same time, it should be emphasized that this period of time is determined in each region individually. It is during this period that sufficient moisture is required (both in the form of precipitation and relative air humidity) so that the grain gets moisture and it is very important that there are no extremely high temperatures at this time interval (especially during flowering).

The key stages of further research are as follows:

1. To determine the effect of weather conditions (heat, air humidification, drought) and elements of agritechnology (fertilizer systems) on the formation grain maize yield.
2. To find an optimal fertilizer system as a key element of agritechnology for this crop and conduct the futher stages of research based on the latter.
3. To characterize and determine the effect of three identified weather risks, measured by certain parameters (heat, cold – temperature; air humidification - relative humidity; drought - precipitation) on the grain maize yield at a certain critical stage of crop vegetation (flowering and grain filling).
4. To calculate the combined weather index (C_i) and, accordingly, all other key parameters that characterize it (the index step, in particular; the interval step, determining the insured event occurence; the justification of the scale of insurance payments as a result of insured event occurence, etc.).

Determination of the weather conditions effect (heat – temperature; air humidification - relative humidity; drought - precipitation) and agritechnology elements (fertilizer systems) on the formation of grain maize yield

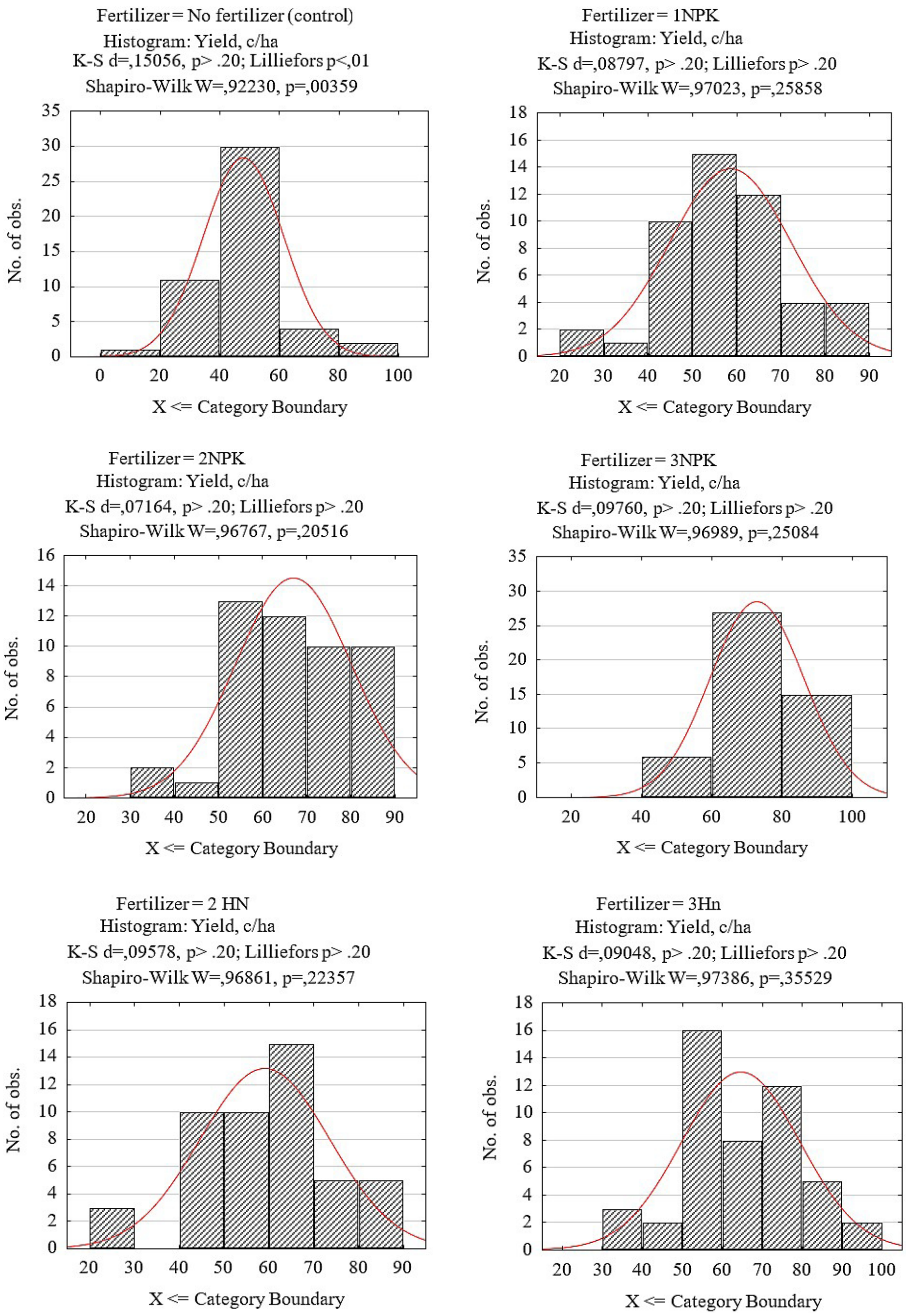
Fertilizer		**Mean**	**Median**	**Min**	**Max**	**Std.Dev.**	**Coef.Var.**
No fertilizers	**No fertilizers**	47.9	45.65	20.0	83.6	13.48	28.09
$N_{45}P_{45}K_{45}$	**1 NPK**	58.6	57.6	28.5	89.0	13.76	23.48
$N_{90}P_{90}K_{90}$	**2 NPK**	67.0	66.7	35.7	89.6	13.19	19.69
$N_{135}P_{135}K_{135}$	**3 NPK**	72.8	72.1	42.4	95.8	13.44	18.45
Humus 9 t	**2 Hn**	59.1	61.2	27.7	87.0	14.48	24.51
Humus 13.5 t	**3 Hn**	64.6	65.5	32.4	92.0	14.79	22.88
Humus 18 t	**4 Hn**	67.9	66.45	38.1	93.0	13.95	20.52
Humus 4.5 t + $N_{22}P_{34}K_{18}$	**1 Hn + NPK**	60.5	60.2	30.2	92.0	14.01	23.15
Humus 9 t + $N_{45}P_{68}K_{36}$	**2 Hn + NPK**	70.5	70.3	38.1	93.8	13.39	18.97
Humus 13.5 t + $N_{67}P_{102}K_{54}$	**3 Hn + NPK**	77.0	76.7	47.2	99.7	12.81	16.64

Source: Data samples for statistical processing are formed from the results of researches obtained during 2002-2017 (that is, the interval of 15 years)

Table 1: Grain maize yield depending on study year and fertilizer system.

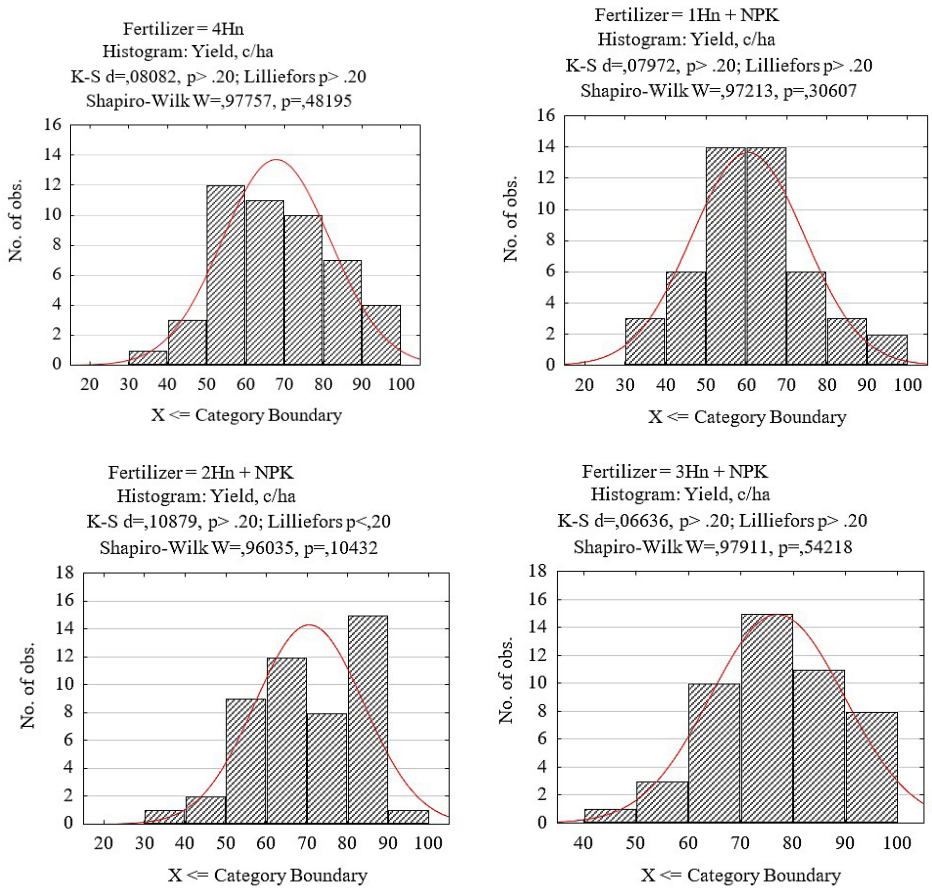
The variation coefficients of the data samples of different study years ranged from 16 to 24%, corresponding to a small and moderate variation (Table 1). This indicates that weather conditions had an effect on the formation of grain maize yield. The most stable samples were data obtained for maximum fertilization, while the control study (variant without fertilizer) had the most variation (Coef. Var. = 28.09). The introduction of fertilizers contributed to the yield increase.

Analyzing the data samples, the theory of correct distribution was confirmed in almost all cases (Figure 1). An exception to the above is a sample of data obtained without fertilization, as evidenced by a low probability score for the Shapir-Wilka criterion (p = 0.0035). Therefore, a further research was conducted using parametric statistics methods.



Source:Authors' own calculations based on the conducted research

Figure 1: Data distribution histograms (to be continued).



Source: Authors' own calculations based on the conducted research

Figure 1: Data distribution histograms (continuation).

After conducting multivariate dispersion analysis (Factorial ANOVA), the reliable effect of the environment and fertilizer system on the formation of grain maize yield was confirmed (Table 2). The weather conditions that were different in the study years (Partial eta-squared = 0.96) had the biggest effect on the crop yield formation. To a lesser extent, however, the elements of agritechnology also significantly influenced (Partial eta-squared = 0.91). In addition, there was a reliable link between the factors (Partial eta-squared = 0.58). As a result, it is arguable that the negative effect of environment can be partially minimized by agritechnology of cultivation.

Selection of optimized fertilizer system as a key element of crop agritechnology

The highest yield of the analyzed crop was recorded in 2011 and in 2014 and was 88.4 and 84.8 centners per hectare, respectively (Figure 2). The worst crop yield was in 2007 (35.9 c/ha). In other years, the crop yield varied from 50 to 75 c/ha.

It should be noted that the fertilizer application contributed to the yield increase. The biggest effect was after the application of 3 NPK (yield – 72.8 centners/ha), 4 HN (yield – 68.0 cent/ha), and 3 HN + NPK (yield – 77.0 centner/ha). As a consequence, these treatments can be recommended to farms for the efficient cultivation of grain maize. At the same time, for a further research, samples obtained after the application of 3 NPK fertilizers were used, since this method is most commonly used in farms. Therefore, all further calculations were carried out according to 3 NPK ($N_{135}P_{135}K_{135}$) fertilizer systems.

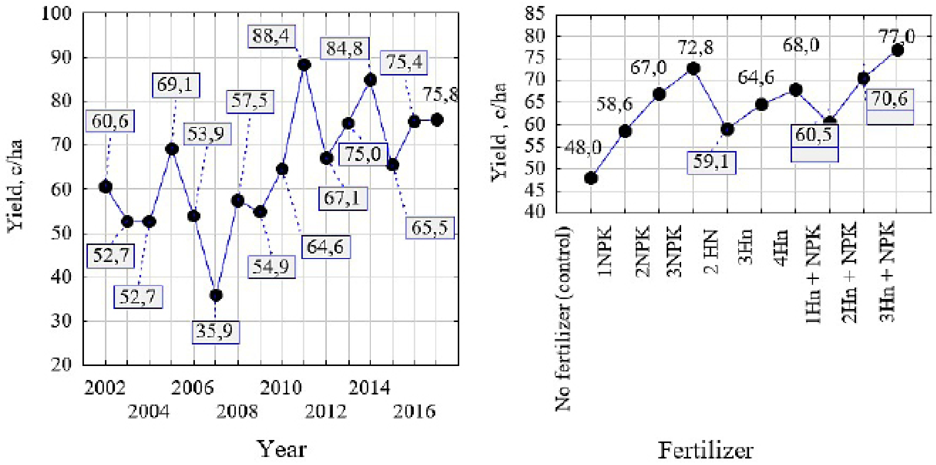
Characterization and determination of the effect level of the specified weather risks (heat – temperature; air humidification - relative humidity; drought - precipitation) on the grain maize yield in the period of flowering and grain filling

Analyzing weather conditions (daily mean temperature, relative humidity and total precipitation), it can be stated that they significantly change over the study years (Figure 3).

Effect	SS	MS	F	p	Partial eta-squared	Non-centrality
Intercept	2004809.68	2004809.68	229036.03	0.00	0.99	229036.03
Year	82005.24	5467.01	624.56	0.00	0.96	9368.54
Fertilizer	30371.14	3374.57	385.52	0.00	0.91	3469.69
Year*Fertilizer	4006.68	29.67	3.39	0.00	0.58	457.73
Error	2801.04	8.75				

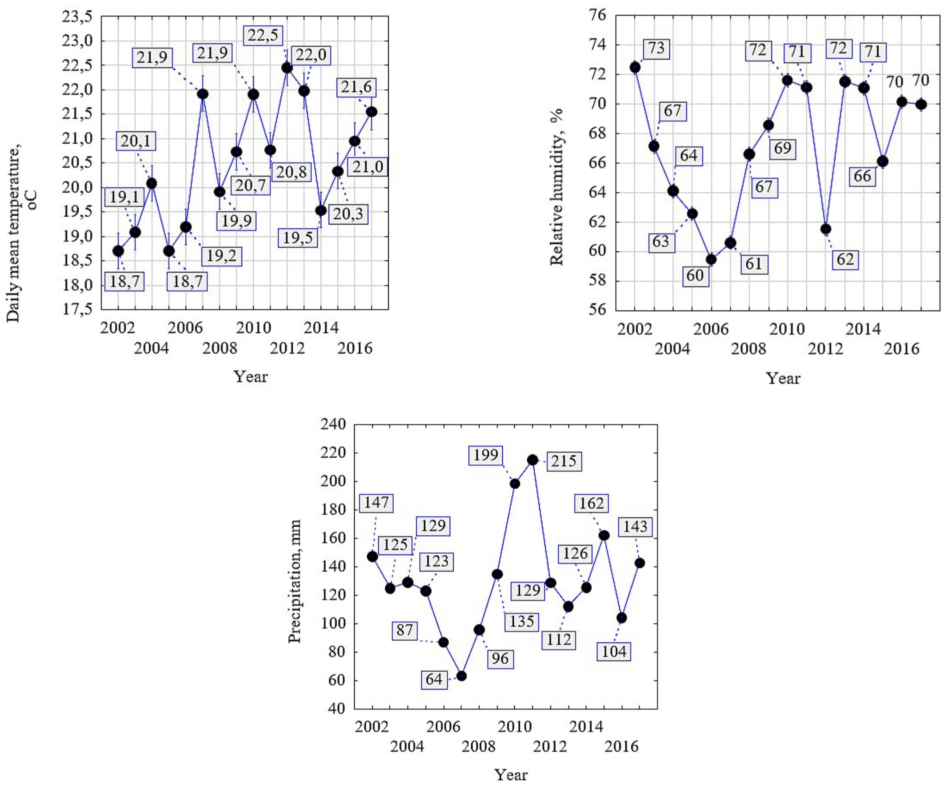
Source: Authors' own calculations based on the conducted research

Table 2: Univeriate tests of significance and factor effects for grain maize yield.



Source: Authors' own calculations based on the conducted research

Figure 2: Effective hypothesis of grain maize yield distribution depending on level and fertilizer system and the year of sudy

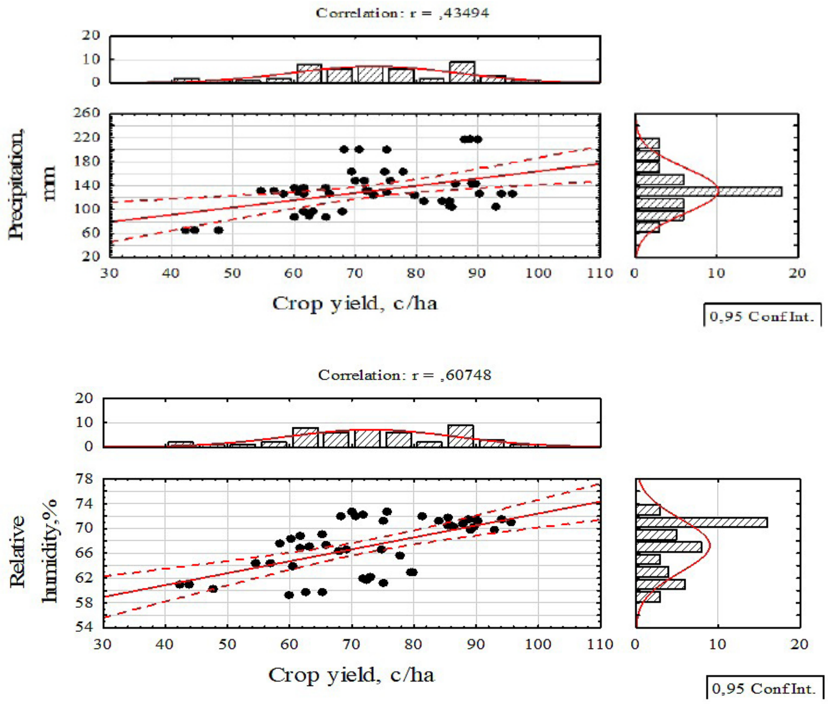


Source: Authors' own calculations based on the conducted research

Figure 3: Effective hypothesis of environmental parameters distribution depending on the year of research.

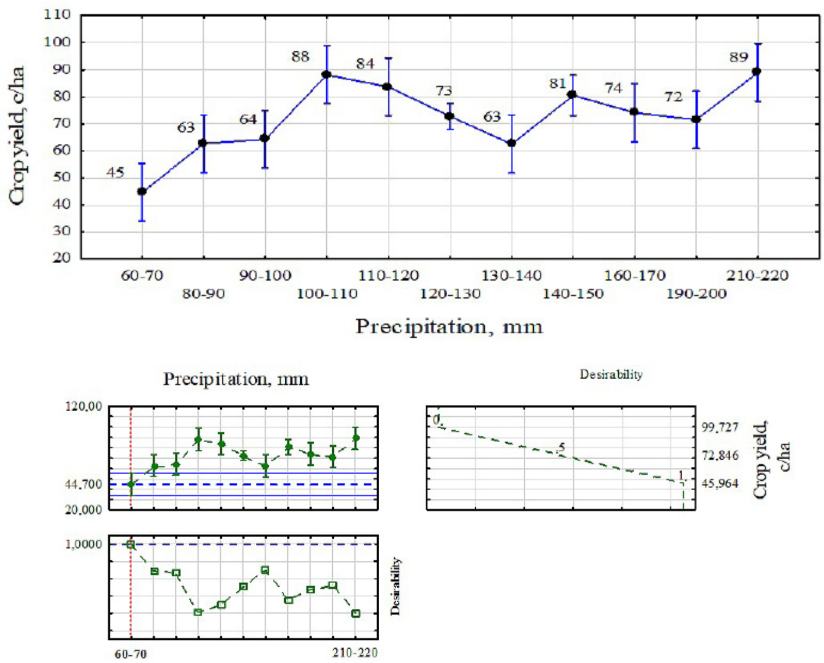
The strongest proved correlation dependence was recorded between yield, total precipitation (r = 0.43) and relative humidity (r = 0.60) (Figure 4). No correlation between daily mean temperature and crop yield was recorded.

The positive effect of total precipitation increasing is confirmed by the dispersion analysis (Figure 5). After mathematical generalization of the corresponding dependence in the form of the desirability function, it was established that the least amount of precipitation (60-70 mm) was critical for the crop yield.



Source: Authors' own calculations based on the conducted research

Figure 4: Correlation between environment factors and crop yield.

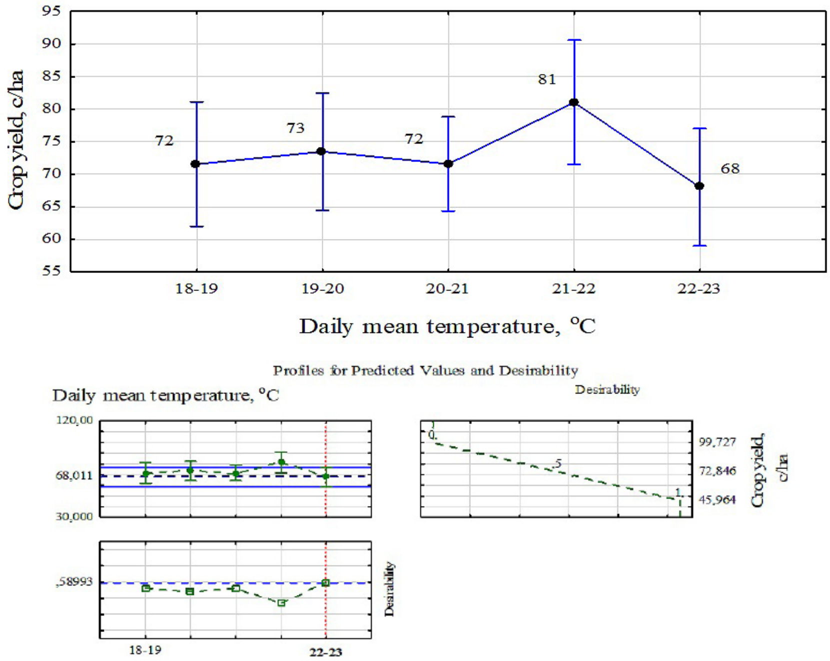


Source: Authors' own calculations based on the conducted research

Figure 5: Effect of total precipitation on grain maize crop yield.

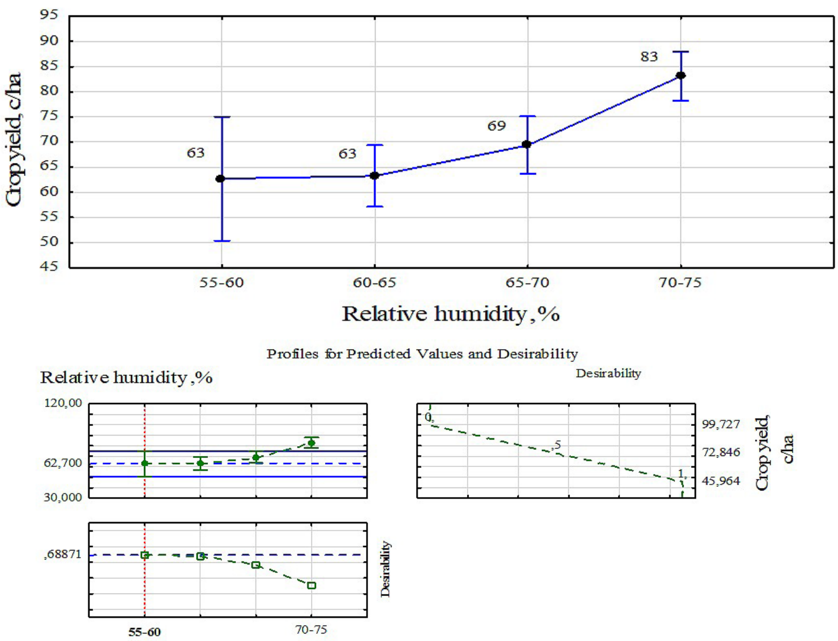
The probability of temperature effect on grain maize crop yield was 70 % (Figure 6). It was recorded that the optimum temperature for cultivation was 21-22 °C, relatively lower and higher temperatures in general had a negative effect on crop yield. Analyzing the desirability function, it is arguable that for the grain maize higher temperatures were critical (22-23 °C).

The highest and the most stable crop yield was recorded at the highest relative humidity of 70-75 % (Figure 7). Critical for it was the lowest relative humidity – 55-60 %.



Source: Authors' own calculations based on the conducted research

Figure 6: Temperature effect on grain maize crop yield.



Source: Authors' own calculations based on the conducted research

Figure 7: Relative humidity effect on grain maize crop yield.

Results and discussion

Classification characteristics of insurance products in agricultural business of Ukraine

At this stage, Ukrainian agricultural insurance market is presented in the context of the classic and parametric schemes of insurance (Figure 8). According to the classic scheme, agricultural risks were insured in the context of multi-risk and mono risk product offers with their multiterminal marketing line. Among the classic insurance products on Ukrainian agricultural insurance market, multi-risk product offers (both in the winter and in spring-summer period) have got the most widespread, providing a full and packaged insurance protection of crops in general and field crops, in particular. Under the parametric (index) scheme, agricultural risks were insured in the context of weather and yied index insurance products, dominant among which are weather index insurance products.

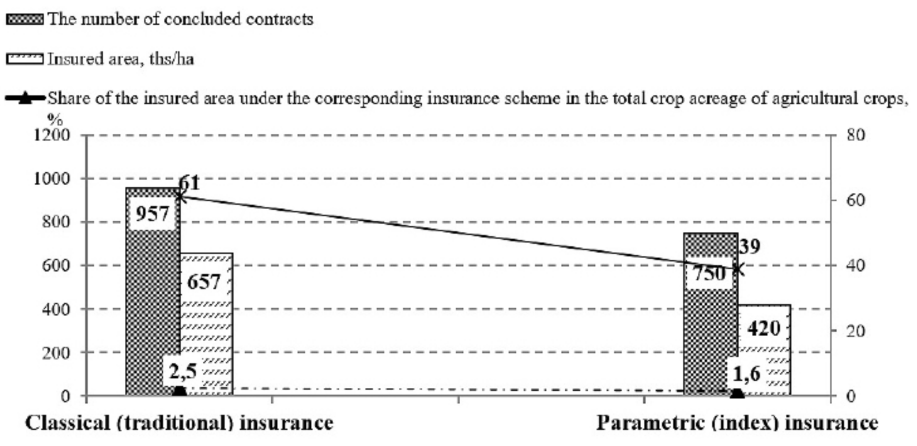
Analyzing the data of Figure 8, it should be noted that the share of the insured area in the total crop area under the results of 2017 and under the classical insurance scheme was 2.5%, while for the parametric – 1.6 %.

The total number of insured crops in Ukraine in 2017 amounted to 1077 thousand hectares, of which 420 thousand hectares were insured using the index scheme of insurance (that is, this insurance scheme covered almost 40% of the total area of insured crops in Ukraine). The number of concluded insurance contracts dominated according to the classic insurance scheme as well. At the same time, given the innovative nature of this insurance scheme in Ukrainian insurance practice, its achievements in the domestic agricultural insurance market are quite significant (its implementation in the framework of pilot projects took place from 2016).

Considering the question of parametric insurance in the context of its classification characteristics, it should be noted that there are many varieties of index insurance products, while Ukrainian insurance practice at this stage is limited to the main two of them: weather and yield. Weather index insurance products are dominated among the recent. In addition, it should be noted that their implementation is still limited in terms of territorial coverage (these products cover only a few regions of the country and are currently distributed to a limited range of crops (wheat and maize). Thus, at present, parametric insurance at the Ukrainian agricultural insurance market functions within the framework of a pilot project, which gives a significant opportunity for its further development, due to its numerous advantages.

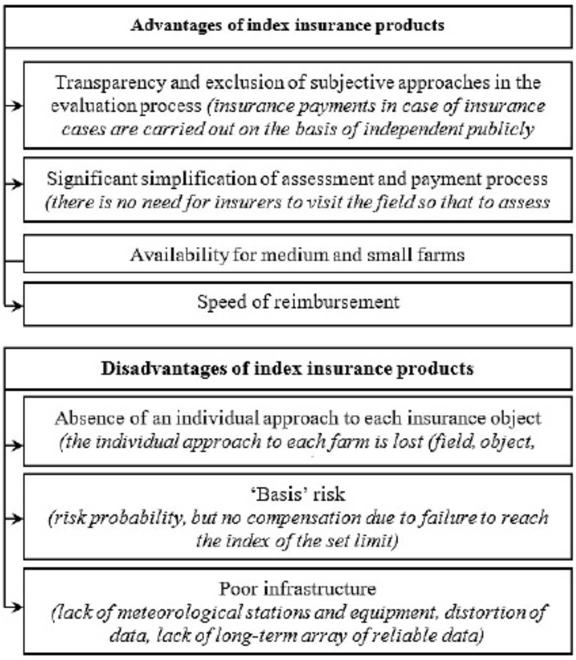
Advantages and disadvantages of index insurance products for the agricultural sector of economy

Insurance on the basis of weather indices is the most suitable for agricultural production in the Ukrainian regions, where crop loss due to drought, excessive overmoisturising and freezing are widespread. The combination of advantages and disadvantages of index insurance products for the agricultural sector of economy is grouped in Figure. 9.



Source: Data of the National Commission on State Regulation in the FinanIcal Services Markets (calculated and proposed by the authors)

Figure 8: Ukrainian agricultural insurance market in the context of classical and parametric insurance schemes (2017).



Source: Author's generalization

Figure 9: Advantages and disadvantages of index insurance products for the agricultural sector of economy.

Analyzing the Figure 9, it should be noted that index insurance products for the agricultural sector of economy have a number of advantages and disadvantages.

In index insurance there are all the essential elements of traditional insurance - there is a subject of insurance, certain risks, as well as the price and the mechanism of payment calculation. In traditional (classic) insurance, the amount of damages as well as the amount of compensation are determined by agents who can make such an evaluation differently and, accordingly, the amount of the loss will also differ. As a result, objectivity of evaluation is lost and process is delayed (especially when it comes to mass losses). While index insurance has a clear algorithm for determine the damage, which is based only on certain indicators, such as weather data and mathematical models.

As a result, it should be noted that index insurance helps to overcome some disadvantages of classical insurance.

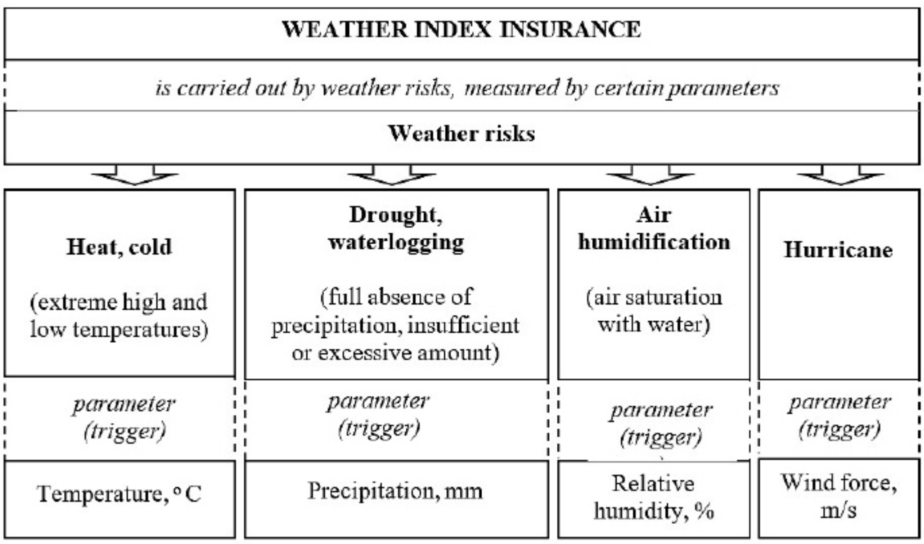
Weather index insurance products

At the heart of weather index insurance is the onset of certain weather events, in case of which the possibility of obtaining an insurance payment appears. At the same time, the key task of the developers of such weather indices is the maximum approximation of two facts: the reduction of crop yield and those weather phenomena that are the basis of this index and are determined and described by certain parameters (triggers) (Figure 10).

That is, index insurance works depending on some parameter (trigger) or their combination (the index can take into account and lay down several parameters at a time). Thus, insurance based on weather indices is a real lever of weather effects, that is, factors that actually reduce the crops yield.

Functioning mechanism aspects of the combined weather index

It should be also said that the index can take into account and lay out several parameters together, so-called triggers. Let us say not only the amount



Source: Author's development

Figure 10: Insurance coverage peculiarities of agricultural producers on the basis of weather indices.

of precipitation, but also the air temperature, humidity and at the same time take into account the effect of these parameters in different phases of crio vegetation in different ways. For example, a lack of moisture during the grain filling period and extreme high temperatures at flowering time will be more damaging to the future harvest than in other periods, so the index may provide for the deviations payment of the parameters in this period at the maximum size.

Index insurance, as already noted above, refers to the parametric type of insurance. Index - a special indicator that characterizes the effect of a certain risk (set of risks) intensity on the crop yield in the definite development phases. In this research, the combined weather index (Ci) is a special indicator that characterizes the impact intensity of the weather risks combination, measured by certain parameters (heat, cold – temperature; air humidity – relative humidity; drought – precipitation) on the yield of grain maize in the definite development phases (flowering and grain filling) (Fig. 11). Its limit characterizes an insured event and corresponds to a critical loss of yield for the farmer, which is determined even during the development of the insurance product. The index should accurately show how exactly a certain risk affects the crop yield.

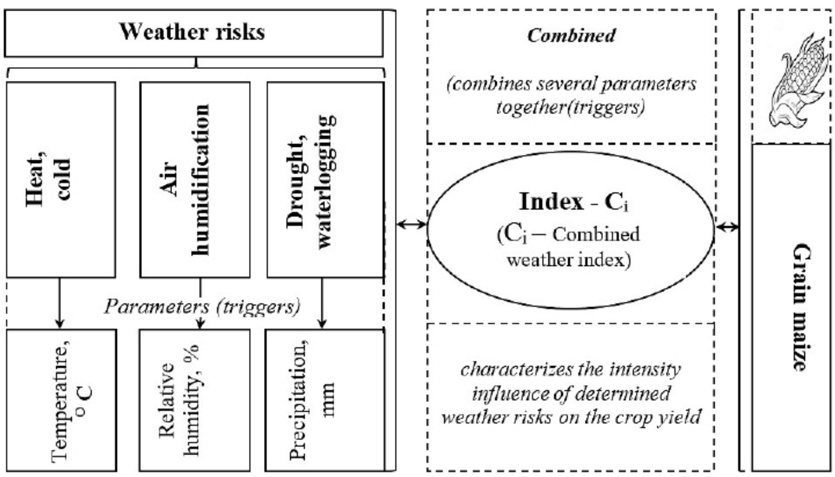
Some aspects related to the grain maize cultivation in Ukraine

As this research is based on the indicators of one of the most important agricultural crops – grain maize, it is worth dwelling on some aspects related to its cultivation patterns in Ukraine.

An important component of the whole grain industry in Ukraine is the maize grain production. Maize in Ukraine is grown as a grain, feed and technical crop. It is characterized by the usage versatility and quite high yield. Analyzing the corn production patterns of grain maize in Ukraine, it should be noted that there is a gradual positive dynamics towards an increase in the production of the latter in 2017 compared to 2005 by 3.5 times. The share of the studied crop in the structure of grain and legume production is significant and in the recent years (2014-2017) it reaches the level of about 40% (in 2013 grain maize dominated grain and legume structure – its share almost reached 50 %) (Nesterchuk et al. (2018). Its importance and authority in the crop industry has been analyzed quite rapidly. Since 2005, the main production parameters of the latter have increased by much: crop acreages have grown from 1.711 million hectares (2005) to 4.2 million hectares (2017), croppage – from 7.2 million tons (2005 p.) to 25.1 million tons (2017), and in 2013 received a record 30.1 million tons of maize grain. So far, this industry in Ukraine is under the influence of the global trend, according to which the area under corn increases - in 2016 by 2%, in Ukraine – by 1.4 % for the same period ((Nesterchuk et al., (2018).

The share of grain maize, both in the structure of the crop acreage of agricultural crops and in the structure of crop acreage of grain and leguminous crops, significantly increased over the period under research: in 2017 compared with 2005 – by 2.5 times, both indicators respectively.

Thus, the grain maize production for the analyzed period in domestic practice can be divided



Source: Author's development

Figure 11: Peculiarities of the functioning mechanism of the combined weather index (C_i).

into periods: the first – (2005-2010) and the second – (2011-2017), which are characterized by the tactics of steady slow and rapid growth.

Interpretation and analysis of study results

Analyzing the obtained data, it should be noted that among the three identified weather risks, the degree of their impact on the grain maize yield was distributed as follows: the greatest effect had such a parameter as precipitation (the maximum value of which resulted in obtaining the top yield); the relative humidity was charactirized by the average degree of effect (the maximum value of the indicator resulted in the top yield); and the moderate effect made such a parameter as the daily mean temperature (the optimal value of which resulted in top yield) (Figure 12).

Critical values of crop yield of the studied crop were observed for:

- the smallest amount of precipitation (60-70 mm), which resulted in a reduction in grain maize crop yield by 1.5 times of its average value and almost twice as much as the top size of the latter (Figure 5);
- the lowest relative humidity (55-60%), which resulted in the reduction in grain maize crop yield by 1.3 times compared with its top value (Figure 7);
- higher air temperatures (22-23 °C) which resulted in 7% reduction in grain maize crop yield, compared with its average value (Figure 6).

The average value of grain maize crop yield within the framework of this experiment was 73 (72.846) c/ha.

Calculation of combined weather index (C_i) and, accordingly, all other key parameters that characterize it

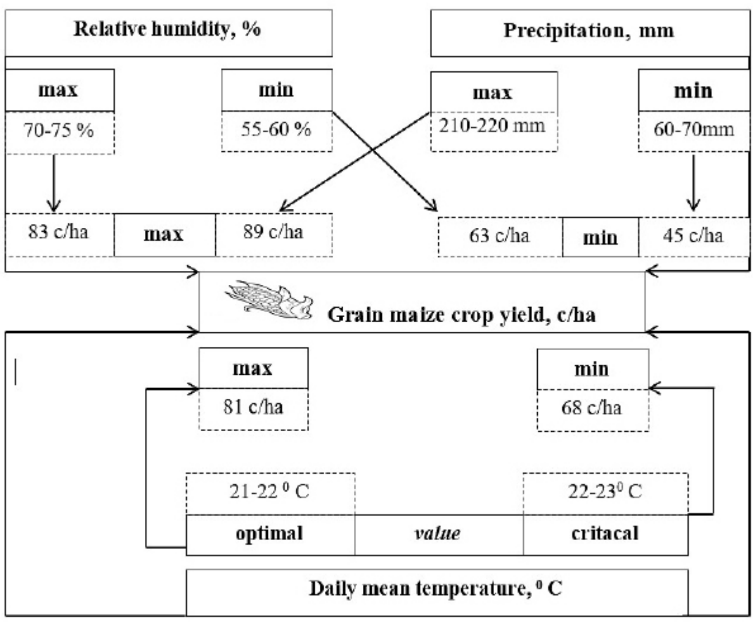
Thus, in conditions of Cherkasy region, namely, Uman district, it is expedient to use a combined weather index reflecting the tandem effect of determined weather risks, which are described by certain parameters, on the reduction of the investigated crop yield.

As an index value, based on the results of this particular study, it is advisable to take optimal values of weather risks, which determine the achievement of an average yield level for a given crop (73 c/ha). Instead, the limit values of weather risks determine the achievement of a critical level of productivity in a given crop and result in the payment of maximum insurance indemnities.

The description of the intervals determining the occurrence of the insured event and the peculiarities of the insurance payments gradation from the author's position, in the context of individual parameters, is presented in Table 3.

The features of the functioning mechanism of the combined weather index (Ci) according to the performed calculations are graphically depicted in Figure 13.

As can be seen from Table 3, the size of the insurance indemnity depends on the deviation degree



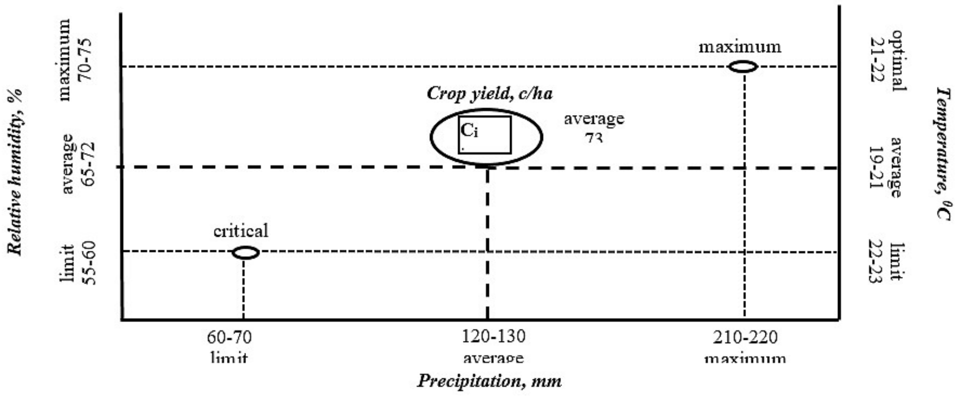
Source: Author's development

Figure 12: Peculiarities of determined weather risks effect on grain maize crop yield.

Weather risks	Parameter (trigger)	Index value	Index step	Insured event	Gradation of intervals	The amount of the insurance indemnity (% of the insurance sum)
Drought	Precipitation, mm	120-130	10	< 120-130	110-120	10
					100-110	20
					90-100	30
					80-90	40
					60-70	50
					< 60-70	100
Air humidification	Relative humidity, %	65-72	5	< 65-72	60-65	10
					55-60	20
					< 55-60	100
Heat	Temperature, ^{0}C	19-21	1	> 19-21	21-22	10
					22-23	20
					> 22-23	100

Source: Author's development

Table 3: Key data according to research results.



Source: Author's development

Figure 13: Features of the functioning mechanism of combined weather index (Ci) according to the performed calculations.

of the index from the average calculated value (that is, the value of the index 'step').

Based on the obtained results, it is arguable that in the period of 2002-2017, within the framework of this research, an insured event with the terms of payment of maximum insurance indemnity was made in 2002 and during 2015-2017 which was due to the achievement of the limit values of weather risks in the most important periods of plant vegetation – grain maize.

At the same time, it should be noted that insurance companies may view this problem in different ways. So, for the size of the index may be taken not average but the limit values of weather risks described by the appropriate triggers which resulted in the achievement of critical yield level of this crop. In such circumstances, the farmer claims to receive the insurance indemnity in whole. Other options may also be considered. However, these issues are purely technical and should be agreed at the stage of development of the insurance product and the conclusion of an insurance contract.

It is worth emphasizing that weather indices can be calculated for different crops and vegetation periods, as well as to lay one or more parameters at a time in their basis. At the same time, it should be noted that combined weather indices are more effective.

According to the results of research, it was found that the combined weather indices have specific features that are realized by the possibility of several parameters simultanious arranging. As a result, this, on the one hand, results in a greater complexity in their calculation, while at the same time it allows the effect of weather risks combination to be mitigated.

Conclusions

The conducted research allowed to make the following conclusions and generalizations:

1. Ukrainian agricultural insurance market at the modern stage operates in the context of two key insurance schemes: classical and parametric, and also uses a sufficient set of insurance products within each of them, which according to their qualitative characteristics meet the international standards.

   A parallel between the key schemes of agricultural risk insurance, including classical and parametric, made it possible to outline the features of each of them:

   - the classic insurance scheme is quite widespread both in the context of world and national practices in agricultural risk insurance, while the parametric one is innovative in the Ukrainian market of agricultural insurance and is used at this stage in the framework of pilot projects, which develops a significant potential for its further development, taking into account the numerous advantages of the latter;
   - in the parametric (index) scheme, agricultural risks were insured in the context of weather and yield index insurance products, dominant among which are weather index insurance products. Given the innovative nature of this insurance scheme in Ukrainian insurance practice, its achievements in the domestic agricultural insurance market are significant.
2. According to the results of the research, the advantages and disadvantages of index insurance products are summarized, resulting in the following: transparency and exclusion of subjective approaches in the evaluation process; significant simplification of evaluation and payment process; availability for medium and small farms; speed of reimbursement. Instead, the disadvantages are outlined as follows: the lack of an individual approach to each insurance object; risk of 'basis'; undeveloped infrastructure.
3. The application features of parametric (index) insurance schemes with emphasis on the use of weather index insurance products are considered. It is determined that the parametric (index) insurance schemes work depending on some parameter (trigger) or their combination (the index can take into account and group several parameters simultaneously). Thus, insurance based on weather indices is a real leverage of weather risks, that is, factors that actually reduce the crops yield. As a result, the research considers the possibility and feasibility of applying, together with simple weather index insurance products, combined that are able to take into account and group several parameters simultaneously and, thus, minimize the impact of the combination of weather risks at the regional level.
4. The authors suggested and, in the framework of this certain research, worked out the feasibility and possibility of using the combined weather index (Ci) – a special indicator characterizing the effect of the combination intensity of weather risks, measured by certain parameters (heat, cold → temperature; air humidification → relative humidity; drought → preIcpitation) on the grain maize yield in the determining phases of development (flowering and grain filling).

Corresponding authors
Olena Prokopchuk, PhD
Department of Finances, Banking and Insurance, Faculty of Economics and Entrepreneurship
Uman National University of Horticulture, Ukraine, Uman
Phone: +30982453895, E-mail: olena_prokopchuk@ukr.net
ORICD: https://orIcd.org/0000-0001-7405-2469

Ihor Prokopchuk, PhD
Department of Agrochemistry and Soil Science, Faculty of Agronomy
Uman National University of Horticulture, Uman, Ukraine
Phone: +30972285417, E-mail: pivotbi@ukr.net
ORCID: https://orcid.org/0000-0001-9608-6061

Grzegorz Mentel, PhD
Department of Quantitative Methods, Faculty of Management
Rzeszow University of Technology, Poland
E-mail: gmentel@prz.edu.pl
ORCID: https://orCid.org/0000-0002-6371-1219

Yurii Bilan, PhD
Centre of Applied Economic Research
Tomas Bata University in Zlin, Czech Republic
Phone: +48506354648, E-mail: yuriy_bilan@yahoo.co.uk
ORCID: https://orcid.org/0000-0003-0268-009X

References

[1] Barnett, B. J., Barrett, C. B. and Skees J. R. (2008) "Poverty traps and index-based risk transfer products", *World Development*, Vol. 36, No. 10, pp. 1766-1785. ISSN 0305-750X. DOI 10.1016/j.worlddev.2007.10.016.

[2] Barnett, B. J. and Mahul, O. (2007) "Weather index insurance for agriculture and rural areas in lower-income countries", *American Journal of Agricultural Economics*, Vol. 89, No. 5, pp. 1241-1247. E-ISSN 1467-8276. DOI 10.1111/j.1467-8276.2007.01091.x.

[3] Chantarat, S., Barrett, Ch. B., Mude, A. and Turvey, C. (2007) "Using Weather Index Insurance to Improve Drought Response for Famine Prevention", *American Journal of Agricultural Economics*, Vol. 89, 5, pp. 1262-1268. E-ISSN 1467-8276. DOI 10.1111/j.1467-8276.2007.01094.x.

[4] Dalhaus, T., Musshoff, O. and Finger R. (2018) "Phenology Information Contributes to Reduce Temporal Basis Risk in Agricultural Weather Index Insurance", *Scientific Reports*, Vol. 8, No. 1, pp. 1-10. ISSN 2045-2322. DOI 10.1038/s41598-017-18656-5.

[5] Dandekar, V. M. (1977) "Crop insurance for developing countries, Teaching and Research Forum Paper No. 10", Agricultural Development CounIcl, New York.

[6] Grmanová, E. and Strunz, H. (2017) "Efficiency of insurance companies: Application of DEA and Tobit analyses", *Journal of International Studies*, Vol. 10, No. 3, pp. 250-263. E-ISSN 2306-3483, ISSN 2071-8330. DOI 10.14254/2071-8330.2017/10-3/18.

[7] Fuchs, A. and Wolff, H. (2011) "Concept and Unintended Consequences of Weather Index Insurance: The Case of Mexico", *American Journal of Agricultural Economics*, Vol. 93, No. 2, pp. 505-511. E-ISSN 1467-8276. DOI 10.1093/ajae/aaq137.

[8] Gebeltová, Z., Malec, K., Maitah, M., Smutka, L., Appiah-Kubi, S. N. K., Maitah, K., Sahatqija, J. and Sirohi, J. (2020) "The impact of crop mix on decreasing soil price and soil degradation: A case study of selected regions in Czechia (2002-2019)", *Sustainability*, Vol. 12, No. 2, ISSN 2071-1050. DOI 10.3390/su12020444.

[9] Halcrow, H. G. (1948) "The theory of crop insurance", Unpublished doctoral dissertation, University of Chicago, Department of Economics, Chicago.

[10] Kocisova, K., Gavurova, B. and Kotaskova, A. (2018) "A slack-based measure of agricultural efficiency in the European Union countries", Journal of International Studies, Vol. 11, No. 1, pp. 189-200. E-ISSN 2306-3483, ISSN 2071-8330. DOI 10.14254/2071-8330.2018/11-1/14.

[11] Mahul, O. (2001) "Optimal insurance against climatic experience", American Journal of Agricultural Economics, Vol. 83, No. 3, pp. 593-604. E-ISSN 1267-8276. ISSN 00029092. DOI 10.1111/0002-9092.00180.

[12] Malyovanyi, M., Nepochatenko, O. and Nesterchuk, Y. (2018) "Conceptual Approaches to Improving the Functioning of Non-state Social Insurance Institutions in Ukraine", *Economics and Sociology*, Vol. 11, No. 2, pp. 289-304. ISSN 2071-789X. DOI 10.14254/2071-789X.2018/11-2/20.

[13] Martseniuk-Rozaronova, O. V. (2010) "State support under insurance of agricultural producers", *Economy of Agroindustrial Complex*, Vol. 10, pp. 91-95.

[14] MIcntosh, C., Sarris, A. and Papadopoulos, F. (2013) "Productivity, credit, risk, and the demand for weather index insurance in smallholder agriculture in Ethiopia", Agricultural Economics., Vol. 44, No. 4-5, pp. 399-417. E-ISSN 1574-0862. DOI 10.1111/agec.12024.

[15] Morkūnas, M., Volkov, A. and Pazienza, P. (2018) "How Resistant is the Agricultural Sector? Economic Resilience Exploited", E*conomics and Sociology*, Vol. 11, No. 3, pp. 321-332. ISSN 2071-789X. DOI 10.14254/2071-789X.2018/11-3/19.

[16] Navrotskyi, S. A. (2012) "State and development trends of agricultural insurance", *Science and Economics*, Vol. 1, No. 25, pp. 61-68.

[17] Nesterchuk, Y., Prokopchuk, O., Tsymbalyuk, Y., Rolinskyi, O. and Bilan, Y. (2018) "Current status and prospects of development of the system of agricultural insurance in Ukraine", *Investment Management and FinanIcal Innovations*", Vol. 15, No. 3, pp. 56-70. ISSN 1810-4967. DOI 10.21511/imfi.15(3).2018.05.

[18] Odening, M. and Shen, Z. (2014) "Challenges of insuring weather risk in agriculture", *Agricultural Finance Review*, Vol. 74, No. 2. ISSN 0002-1466. DOI 10.1108/AFR-11-2013-0039.

[19] Osgood, D. E., McLaurin, M., Carriquiry, M., Mishra, A., Fiondella, F., Hansen, J., Peterson, N. and Ward, N. (2007) "Designing weather insurance contracts for farmers in Malawi, Tanzania and Kenya", International Research Institute for Climate and SoIcety, Earth Institute, Columbia University Pallisades, New York.

[20] Pointer, L. and Khoi, P. D. (2019) "Predictors of Return on Assets and Return on Equity for Banking and Insurance Companies on Vietnam Stock Exchange", *Entrepreneurial Business and Economics Review*, Vol. 7, No. 4, pp. 185-198. E-ISSN 2353-8821, ISSN 2353-883X. DOI 10.15678/EBER.2019.070411.

[21] Polchanov, A. Y. (2013) "Agricultural insurance as a component of state food security", *Bulletin of the Cherkasy Bohdan Khmelnytsky National University.of Trade and Economics*, Vol. 4, pp. 48-57. ISSN: 2707-5044.

[22] Prokopchuk, O., Prokopchuk, I. and Mentel, G. (2018) "Index Insurance as an Innovative Tool for Managing Weather Risks in the Agrarian Sector of Economics", *Journal of Competitiveness*, Vol. 10, No. 3, pp. 119-131. E-ISSN 1804-1728. ISSN 1804-171X. DOI 10.7441/joc.2018.03.08.

[23] Prokopchuk, O., Nesterchuk, Y., Tsymbalyuk, Y. and Rolinskyi, O. (2019) "Current trends in agricultural insurance market operation in Ukraine", *Problems and Perspectives in Management*, Vol. 17, No. 3, pp. 57-75. E-ISSN 1810-5467. DOI 10.21511/ppm.17(3).2019.05.

[24] Santeramo, F. G. (2018) "Imperfect information and partiIcpation in insurance markets: evidence from Italy", *Agricultural Finance Review*, Vol. 78, No. 2, pp. 193-194. ISSN 0002-1466. DOI 10.1108/AFR-06-2017-0053.

[25] Sarris, A. (2013) "Weather index insurance for agricultural development: Introduction and overview", *Agricultural Economics*, Vol. 44, No. 4-5, pp. 381-384. E-ISSN 1574-0862. DOI 10.1111/agec.12022.

[26] Skees, J. R., Hazell, P. B. R. and Miranda, M. (1999) "New approaches to crop yield insurance in developing countries", International Food Policy Research Institute, Environment and Production Technology Division. EPTD discussion paper No.55. Washington, DC.

[27] Skees, J., Gober, S., Varangis, P., Lester, R. and Kalavakonda, V. (2001) "Developing Rainfall-Based Index Insurance in Morocco", Working Paper No. 2577, The World Bank, Policy Research: Washington, DC.

[28] Tadesse, M. A., Shiferaw, B. and Erenstein, O. (2015) "Weather index insurance for managing drought risk in smallholder agriculture: lessons and policy implications for sub-Saharan Africa", *Agricultural and Food Economics*, Vol. 3, No. 1. E-ISSN 2193-7532. DOI 10.1186/s40100-015-0044-3.

[29] Turvey, C. G. (2001) "Weather derivatives for speIcfic event risks in agriculture", *Review of Agricultural Economics*. Vol. 23, No. 2, pp. 333-351. E-ISSN 1467-9353, ISSN 1058-7195. DOI 10.1111/1467-9353.00065.

[30] Ulianych, Y. V., Prokopchuk, O. T.,and Bechko, V. P. (2014) "Peculiarities of agrarian risk insurance in the countries of the world", *Actual Problems of the Economy*, Vol. 3, No. 153, pp. 46-53. ISSN 19936788.

[31] Walewicz, P. (2018) "´Greening´ the critical theory of international relations with the concept of world-ecology", *Toruńskie Studia Międzynarodowe*, Vol. 1, No. 11, pp. 125-134. ISSN 1689-8168. DOI 10.12775/TIS.2018.011.

[32] Weber, R., Fecke, W., Moeller, I. and Musshoff, O. (2015) "Meso-level weather index insurance", *Agricultural Finance Review*, Vol. 75, No. 1, pp. 31-46. ISSN 0002-1466. DOI 10.1108/AFR-12-2014-0045.

Agris on-line Papers in Economics and Informatics

Volume XII Number 3, 2020

Recent Evolution of Perennial Crop Farms: Evidence from Dak Lak Province, Vietnam

Phan Thi Thuy[1,2], Le Duc Niem[1], Philippe Lebailly[2]

[1] Faculty of Economics, Tay Nguyen University, Buon Ma Thuot, Vietnam

[2] Economic and Rural Development, Gembloux Agro-Bio Tech, University Liège, Gembloux, Belgium

Abstract

There is a great consensus about the crucial role of perennial crops in an agricultural economy of a country. The paper aims (1) to identify the differences in the costs and profits of perennial crops produced by two study groups, a group producing coffee (GpC) and a group producing pepper (GpP) over two crop years 2016/2017-2017/2018; (2) evaluate the evolution of the economic performance of each group during two years; and (3) examine factors influencing the farm profitability. By using the mixed data from a household survey conducted in three sub-regions of Dak Lak province, Vietnam, a financial verification is used to explore the economic incentives between two groups and a discriminant analysis is undertaken to classify the determinants of the farm profitability. The results perform that the GpC is generally lower input costs and economic benefits than the GpP. The decrease of economic indicators of the GpP during two years, meanwhile, is more significant than that of the GpC in the same period. In addition, the GpP is likely to invest more inputs, heavily use chemical cost, be more susceptible to pests and diseases, and the volatile market conditions in comparison to the GpC. Categorically, the variable cost and reduction rate in terms of value-added, net farm income (NFI), profit, labor productivity, and the ratio of NFI to family labor of the GpC have lower than those of the GpP, respectively, during two years. Furthermore, in similarly conditional practices, the perennial crop farms generate different returns depending on experience, training, other income, and gross outputs. The findings provide information for farmers to make accurate decision about coffee and pepper farms production as recommended by reducing the quantity of fertilizers, allocating resources and diversifying crop systems. Additionally, the empirical results also offer policymakers the farms sustainable development at local and national levels. Going forward, authors suggest these factors be considered in the future.

Keywords

Evolution, economic performance, coffee and pepper farms, Vietnam.

Thuy, P. T., Niem, L. D. and Lebailly, P. (2020) "Recent Evolution of Perennial Crop Farms: Evidence From Dak Lak Province, Vietnam", *AGRIS on-line Papers in Economics and Informatics*, Vol. 12, No. 3, pp. 87-100. ISSN 1804-1930. DOI 10.7160/aol.2020.120308.

Introduction

The various dimensions of the vital role of perennial crops in the livelihood and export earnings are well captured in the literature. For example, perennial crops subsidized significantly to food, fuel, goods for export, and jobs, as well as environmental and cultural benefits. In addition, perennial crops contributed to the reducing deforestation (Peter et al., 2003; Angelsen, 2010; Dinh et al., 2017). On the other side, the practices of agricultural sector, in general, and cropping system, in particular, had undergone the times, the local, regional, international economic situation by ecological, socio-economic transformation (FAO, 1999). Yet, there are limitations on empirical evidence of economic changes of perennial crop farms.

In Vietnam, perennial crops introduced at the end of the nineteenth century by the French. Currently, these products have become major commodities and driving forces for economic growth and export earnings. As statistical data, perennial crop growing area reached over 2.2 million hectares, in which approximately 1.8 million hectares produce over 4 million tons including coffee, pepper, rubber, cashew and tea, concentrate largely on the Central Highlands, namely Dak Lak, Dak Nong, Gia Lai, Lam Dong, and Kon Tum provinces thanks to the favorable conditions (GSO, 2019).

Over the past quarter century, the agricultural sector has developed such significance that many countries tried to learn from those Vietnamese successes, in which perennial crop section reached explosive growth of export earnings. Since 2000s, Vietnam ranked among the top five global exporters of perennial crop products. For instance, Vietnam was one of the world's largest producer of coffee and the world's leading exporter of black pepper made up relatively 1490 and 100 thousand tons, earning 2880 and 134 million USD in 2017 –2018, respectively (GSO, 2019; JICA, 2012). Notwithstanding, other authors such as (Slater et al., 2007; Ha and Shively, 2008) demonstrated that the perennial crop production may face many difficulties (i.e., production is smallholders and susceptibility to natural disasters and vulnerabilities from trade fluctuation). Thus, challenges of perennial crop production lead to reduction of households' income, as well as significant negative impacts on the national economy. Meanwhile, the perennial crops are not well-adapted into other types of farming due to their long economic lifespans and high startup costs (Thuy et al., 2019 a; Gunathilaka et al., 2018). Consequently, understanding the evolution of cropping system is essential to raise farm productivity, identify the management practices and provide the ensuring farmers' income in the future (Herridge et al. 2019; Stillitano et al. 2019).

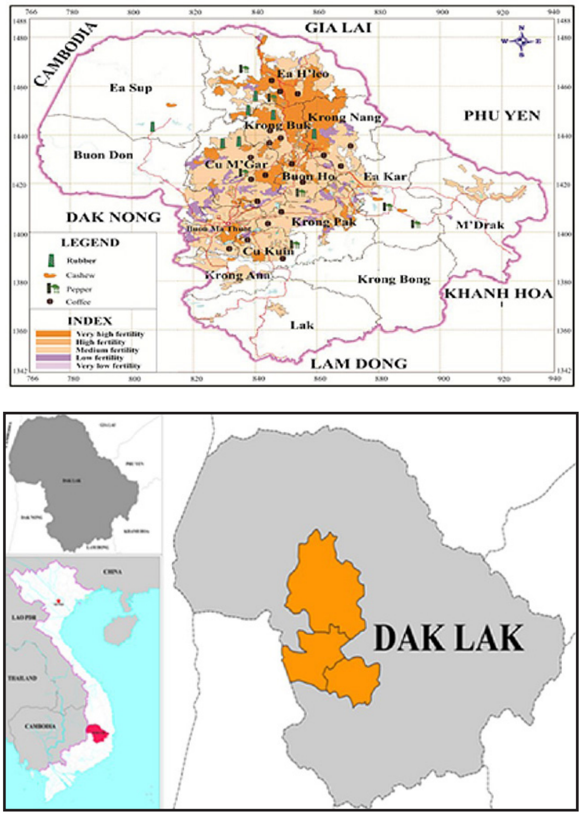
Dak Lak province has the largest basalt soil area to create extremely advantageous conditions for well-grown perennial crops. Remarkably, coffee productivity dominated at about 30% while pepper output occupied approximately 40% for the whole country in 2018 (GSO, 2019). Over the years, perennial crops have changed dramatically being usual dynamic of requirement. Despite its advantages and positive development trends, provincial perennial crop production has faced with numerous constraints due to price fluctuations, climate changes, and pests and diseases. Meanwhile, most literature only focused more on coffee farming and have a few studies compare monetary benefits between perennial crop farm approaches (Chau, 2007; Thi Duong Nga and Thuy, 2017; Ho et al., 2017; Thuy et al 2019b). In overall, there is an obvious need to know the economic earnings and evolution of two group producing perennial crops in order to understand how changes and how well of each perennial crop group during a period.

This study focuses on comparing the economic performance of two groups, a group producing coffee (GpC), which is categorized into coffee specialized farms and diversified farms, where both kinds of farms have the similar about the practices and characters such as the age, input items, and calendar as well as are dominant approaches in research sites, and a group producing pepper (GpP) which is categorized into pepper specialized farms and diversified farms. Another aim is to monitor the evolution of economic viability across two groups of 2016/2017 –2017/2018. The remaining part of this paper highlights factors affecting farm profitability. The results provide empirical evidence to design perennial crop production strategies regarding what they grow and how they grow under real-life conditions. Furthermore, the negative and positive determinants need to decrease or improve for the coming years.

Materials and methods

Selection of study sites

Classification of regions depends on the ecological and suitable land criteria (fertility capacity classification)



Source: Author's creation and adapted from (Thong et al, 2017)

Figure 1: The distribution of perennial crop in Dak Lak province and Map of the study sites.

According to (FAO, 1999), to obtain the identification and localization of agro-ecological, and socio-economic situation in the different systems, selected zones should have similar characteristics. Furthermore, in terms of agriculture, cropping system analysis is carried out on smaller communes or villages to achieve the same resources and examine properly cropping systems (Diepart and Céline, 2018). This study is operating in Buon Ma Thuot city, CuMgar and CuKuin districts of Dak Lak, which have homogeneous fertility and adequate weather conditions, the largest of coffee area (CuMgar), the third largest of pepper area (CuKuin); and favorable market conditions, transportation, agri-services and the longest history of perennial crop production (Buon Ma Thuot city). Three sub-regions have dominant coffee and pepper approaches of specialized and diversified farms. Therefore, from three sub-regions, three communes and three villages are chosen based on the popular level of coffee and pepper farms. This information is provided by local authority discussion. In other words, the existence level of above two groups of coffee and pepper is the criterion to select the communes and villages.

Stratification of perennial crop systems/ Target perennial crop systems selection

As (Barral, 2012), in the limitations on time and resources, it is better to confine the number of cropping systems study rather than implement many cropping systems in order to reach the accurate information. In this study, our effort concentrated on the coffee and pepper farms and the study used the cross-sectional data in 2016/2017 and 2017/2018.

Selection of surveyed farms and households

Selected farms

Regarding the sample size, it depends on the provincially farm scale. As statistical data, over 90% of producers hold around 1.5 hectares per households. Thus, in order to have exact and detailed analyses, the selected perennial crop systems are from 0.5 to 2 hectares with farms age over three years. The first group is labeled the group producing coffee (GpC). Similarly, the second group is called the group producing pepper (GpP).

Sample size

According to (Diepart and Céline, 2018), there is no rule about the household sample size of the research sites. Yet, as reported by (Salvatore and Reagle, 2002), selected households ($n = 0.1*N$, with N: populated size or statistical units) could have the same probability of being selected. For instance, if the sample rate is 10%, meaning the sample size represents 10% of population (FAO, 2016).

From the list of perennial crop farmers in the villages, by using the random sampling method, the coffee and pepper producers including specialized and diversified items are collected. At the end of this process, the total farmers are gathered making up 86 households. However, because each household often owns more than one plot in surveyed regions, total number of plots is larger than number of interviewees. The sample distribution is presented in the Table 1.

In-depth interviews, Focus Group Discussion and face-to-face interview techniques as follows:

- In-depth interviews: Key people, such as elderly people, heads of communes and villages, and experienced farmers, are selected as subjects to provide preliminary information in the research sites via intensive individual interviews;
- Focus Group Discussions (FGDs): one FGD is conducted in each surveyed commune with 6 - 8 farmers to explore the historical trends in perennial crop production and marketing practices;

Region	Commune	Village	(N) Populated size	Selected households (n=10%*N)
BMT	EaKao	Cao Thanh	300	30
CuMgar	Cu Suê	Tu	310	31
Cu Kuin	EaKtur	Muoi	250	25
Total of households				**86**
Group-farming	**Production group**			**Farms**
Group producing coffee (GpC)	Specialized coffee production			32
	Diversified by intercropping coffee production			30
Group producing pepper (GpP)	Specialized pepper production			28
	Diversified by intercropping pepper production			30
Total				**120**

Source: own processing

Table 1: The sample of households and farms.

- Household surveys: a questionnaire is used to collect information about demographics and household characteristics, input and output data, and socio-economic profiles relating to perennial crop production. The field surveys were conducted into two stages of two crop years in 2017/2018 and 2018/2019.

Data analysis

Analysis method

In this research, farm profile, cost-return and comparative analyses are used to identify the differences in economic performance of the two groups and discover the changes of costs and benefits in each group during two crop years, in addition to descriptive statistical analysis such as percentages, means, charts and growth rate. By using cost benefit analysis (CBA), many indicators such as production cost, revenue, value added and profit are computed to indicate which group have the best economic performance and how the changes of each groups during two crop years under similar practices. CBA refers to have a systematic approach to make decisions on which group creates more benefits as well as whether or not implement an investment under limited resources (Quah and Mishan, 2007; Odum, 1983). Meanwhile, the economic performance indicators are a foundational requirement for enhancing market efficiency and decision-making (FAO, 2016). This analysis utilizes the Mann–Whitney test to examine the significant difference between two independent groups (GpC and GpP) and dependent variables including input cost items and economic performance indicators. Alternatively, in this study, the Mann-Whitney test uses to understand whether continuous dependent variables, which are namely production cost items, gross output, net value added, net farm income, profit, labor productivity and GO/IC, differ based on two group of the GpC and the GpP or not. The calculation of costs and benefits is to follow some authors (Hill and Bradley, 2015; Newton et al., 2012 ; Cochet, 2015). The statistical tool SPSS 22.0 version is used to analyze the data.

Econometric analysis

The objective of this study is to find main factors affecting the farm profitability. In other words, the study has attempted to understand why some perennial crop farms are more profitable than others. Alternatively, in this study, the determinant analysis measured the impact of factors on farm profitability by examining differences among two groups categories, including (Non-profit <0: failure group; Profit >0: successful group (No case of profit=0)).

In literature, many factors have been found to be significant in explaining the farm profitability. (Tey and Brindal (2015) found that these factors related to management and financial capacity, farm resource quality and operations, farm and financial management and skills. Meanwhile, Lan et al. (2018) argued that the different profitability could be relied on crop typologies, input access and prices, households types and local context.

In this study, most of failure farms belong to ethnic households. Therefore, we expect there are differences in both farm practices. In other words, we suppose socio-economic characteristics between Kinh and other groups are driving forces for the changes of farm profitability, in which skills such as practical experience and training attempt to understand. It noted that the Kinh group is in the majority while the other groups is indigenous and migrated households (Thai, 2018). Moreover, the economic condition as Gross output, which is controlled by the crop price and yield, is another factor (Lososová and Zdenek, 2014). Note that the data, which is only in 2017/2018 year, utilizes to examine the changes of farm profitability. There is, therefore, a constant of selling price. The output is hypothesized by the effect of yield, where crop and farm profiles can be reasons such as soil, pests and diseases, and age of trees (Ho et al, 2018) (Table 2).

Explanatory variables	Description	Variable
Households' characteristics	Age of household head (Years)	AGE
	Gender (=1 if Male)	GEND
	Ethnicity (=1 if Kinh people)	ETH
	Education (Years)	EDUC
	Training (=1 if trainers)	TRAI
	Experience (Years)	EXPER
	Family workforce (Laborers)	FWORK
Economic condition	Gross output	GO

Source: own processing

Table 2: Variables description.

For qualitative variables, a change from 0 to 1, leaving all other variables constant at their mean is reported. Specifically, GEND is a dummy variable that presents the gender of the farmer; it has a value of 1 for men and 0 for women. ETH is also a dummy variable to measure the ethnicity of households. It takes the value 1, if the households are majority group of Kinh people; 0, if they are indigenous and migrated people. It is expected that the majority group get higher benefits than the others. TRAI measures the farmers participating in training program and applying knowledge for their farms. It takes the value of 1 if the farmer participates in training programs, and 0 if otherwise. It is hypothesized that the training program has positive influence of farm profitability. Quantitative variables include Age of Head (AGE), EDU (Education), Experience (EXPER), FWORK (Family Workforce) and Gross Output (GO).

At the same time, a multiple discriminant analysis is used to measure the impact of socio-economic characteristics of households and farms' output on the performance of different perennial crop farms. Alternatively, a determinant analysis is applied as a quantitative method to clarify the factors causing a problem and reviewing the linkages among these factors (Landau and Everitt, 2003; Trong and Ngoc, 2008).

The discriminant equation:

$$F = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p + \varepsilon, \quad (1)$$

where F is variable form by the linear combination of the dependent variables. $X_1 \ldots X_p$ are the independent variables (socio-economic characteristics of households and gross output). β_0 is constant. $\beta_{1..p}$ is the discriminant coefficients. The aim of discriminant analysis is to test whether the classifications of Y groups depends on at least one of the X_i.

Results and discussion

Statistical descriptions of the group producing coffee and group producing pepper

Generally, the average cultivated areas of the GpC and the GpP are similar, estimated to be about 0.9 hectares per farm group. It means that perennial crop farms are characterized relatively by smallholders as reported by (World Bank, 2016). In addition, the densities reach 936 trees/ha in the GpC and 1197 trees ha in the GpP, respectively, which are lower than the technical standards (1100 coffee trees/ha and 1700 pepper trees/ha) (Thong, 2015). Moreover, the yields among the GpC and the GpP are 2.0 tons for coffee bean and 1.9 tons per ha for dried black pepper, which are smaller than for the nation and other countries such as Indonesia and India. What is more, the numbers of crop losses are 24 coffee trees/ha and 32 pepper trees/ha, respectively (Table 3). One reason is that, with respect to the GpC, the considerable expansion of planting onto less suitable or un-suitable regions (equivalent to 20% in Dak Lak and Lam Dong) and the age of the coffee tree stocks (i.e., over one-third of the provincial coffee growing area was aged 15–20 years) are likely to heavily affect. Meanwhile, concerning the GpP, in the previous years, due to very high selling price, excessive fertilizers were applied for plantation, which can lead to the soil acidification, the infectious diseases, the low yield and crop losses. In order to support this argument, authors (Ton and Buu, 2011; Oanh, 2019; Scherr et al., 2015) depicted that over 2000 ha were lost in 2018 due to pests and diseases such as foot rot (caused 9%–95% of economic losses) and yellow leaves.

The evolution of economic efficiency between the coffee farm and pepper farm groups

The change of input costs among the two farm groups

Table 4 shows the input cost items among the two groups and the variances in each group during the period of 2016/2017–2017/2018.

The results exhibit that the GpC demonstrated lower variable costs, which were estimated to be 1.01 thousand in 2016/2017 and 1.14 thousand USD in 2017/2018, than the GpP, whereas the GpP had variable costs of 1.83 thousand and 1.67 thousand USD/ha, respectively. Moreover, the GpC is considered being more suitable for smallholders having rarely available savings and facing considerable dilemma in accessing credit (Hurri and Quang, 2015; Birner

Items	GpC (n = 62)		GpP (n = 58)		Sig
	Mean	SD	Mean	SD	
1/ Sample size (ha)	0.93	0.6	0.92	0.4	NS
2/ Density (tree/ha)	936	125	1 197	213	NS
3/ Age of farm (years)	16	8	7.5	4	0.00
4/ Yield (ton/ha)	2.0	0.8	1.9	0.85	NS
5/ No. of crop losses (tree/ha)	24	56	32	60	NS

Source: Author's calculations

Table 3: The profiles of coffee farm and pepper farm groups in 2017/2018.

Items	GpC (n = 62)			GpP (n = 58)		
	2016/2017	2017/2018	Sig	2016/2017	2017/2018	Sig
	Mean	Mean		Mean	Mean	
Variable cost	**1.01**	**1.14**	**0.04*****	**1.83**	**1.67**	**NS**
I. Intermediate cost (IC)	***0.68***	***0.78***	***NS***	***1.16***	***0.97***	***NS***
1. Fertilizer	0.49	0.63	0.01**	0.77	0.63	NS
Manure	*0.10*	*0.14*	*NS*	*0.31*	*0.19*	*NS*
Chemical	*0.38*	*0.49*	*0.03***	*0.46*	*0.43*	*NS*
2. Pesticide, herbicide, stimulants	0.08	0.07	NS	0.28	0.26	NS
3. Watering	0.05	0.04	NS	0.04	0.03	0.06***
4. Transporting	0.01	0.01	0.00*	0.01	0.01	NS
5. Packaging	0.01	0.02	0.00*	0.01	0.01	0.00*
6. Others	0.03	0.01	0.00*	0.04	0.03	0.02**
II. Hired labor cost	***0.15***	***0.14***	***NS***	***0.37***	***0.34***	***NS***
III. III. Interest cost	***0.04***	***0.07***	***NS***	***0.06***	***0.13***	***NS***
IV. Depreciation	***0.14***	***0.14***	***NS***	***0.23***	***0.23***	***NS***

Note: The significance levels are indicated as ****p <0.10, **p <0.05,*p <0.01, NS: Non-significant. Mann–Whitney test: 0.00 (excluding family labor cost); a Currency rate: 1USD=23 000 Vietnamese Dong
Source: Author's calculations

Table 4: The variance of input cost of the two farm groups during the period of 2016/2017–2017/2018 (thousand USD/ha).

and Danielle, 2010). Unfortunately, the average cost of pesticides, herbicides, and stimulants for the GpP is several times higher than that of the GpC, which is estimated to be average 0.27 thousand UDS/ha (Table 4). One of the reasons might be that farmers boost amount of pesticides and stimulants due to rising the occurrence of pests and diseases on black pepper plants, influencing on plant growth and yield (Thuy, 2010). As (Thuy et al., 2019b), the authors argued that high pepper price motivates farmers to over-use pesticides and stimulants. This implies that pepper farms are likely to increase the dependency on inorganic and toxic inputs, which could have had a negative effect on production and sustainable development, as revealed by (Susmita, 2007; World Bank, 2016).

In addition, the findings present that the GpP puts greater pressure on the labor force. It means this group requires more workdays and more hired labor during the crop season, especially for harvesting (pepper vine needs to be pick up on time, to control ripe cherries losses and help plant growth for the next season), than did GpC. Another reason that peppers production needs more laborers than that of coffee is field management and protection. For instance, as previously stated, theft is a social problem appears entirely in peppers production, which increases the need for laborers. Especially, thief problem occurs in ethnic minority due to higher un-employment labor (15-20 years) in ethnic families than that of Kinh families (labor went to school or got jobs), in which thieves are the local and un-employment people, as revealed by farmers. Moreover, neighbor conflict causes the personal destructive behavior such as a chopping mass of pepper vines, leading to increase the managers.For instance, because of a lack of power-man on black pepper harvest season the paid-wage of labor in the local regions raise. Even though some households must work overtime hours causing health problem or hire less efficient laborers such as children or elderly people.

Fortunately, compare to the past, irrigation cost reduced in both groups due to the application of three-phase electricity as well as by the advanced irrigation technologies (drip and spray irrigations) instead of using diesel machines.This is an appropriate strategy for local farmers when perennial crop production has experiencing the effects of climate change such as increasing the number of hot days and nights and the occurrence of intense droughts (Haggar and Schepp, 2012). For the coming years, the continuously training programs of irrigation management should be improved, which can obtain the water productivity. Because as a report, trained farmers tended to use fewer litters per plant than the others (Amarasinghe et al., 2015).

Taking everything into consideration, the GpP is likely to apply more input items and the labor pressure for farmers, as opposed to the GpC. Farmers should consider the labor source when choosing crop farms and finding the employees, especially during the harvest season.

At the Government, official calls of controlled release of material items like chemical fertilizers and pesticides in terms of pepper production carefully consider in the future because of ecological effects. Moreover, establishment of volunteers or social communities helps rural farmers to collect the black pepper cherries in urgent cases.

Regarding each group-farming, the annual cost has jumped by 13% for the GpC but has declined steadily for the GpP by 8% per ha during two crop seasons (Table 4). Specifically, regarding the GpC, the intermediate cost has significantly boosted by 0.13 thousand USD/ha (about 30%), of which chemical fertilizer item mainly occupies (average about 52%), which is larger than that of other countries such as India (Devi and Pandurangarao, 2003). This is because of the following: (i) the older age of the coffee crops (the changes of revenues and the increase of inputs); (ii) the fact that farmers do not apply the optimal composition of fertilizers leading to running-off into streams or groundwater (i.e., they follow their neighbors' tendency or advice of stakeholders or middlemen), or farmers got the confusion about how best to use these inputs (World Bank, 2016). Obviously, the growth rate of costs for GpC creates vulnerabilities for the farm groups, especially because coffee prices, are lower than in previous years. Moreover, increased fertilizers (urea and generic NPK fertilizer) leads to polluted surface water, excessive irrigation and soil acidification affect sustainable production. Authors suggest that rather than input costs increase, farmers implement better agricultural practices (i.e., apply agrochemical fertilizers in proportion to need and in accordance with technicians' recommendations) to deflate the amount of inputs and improve economic efficiencies. In addition, the practice of optimizing fertilizers, according to a Technoserve study, by reducing them, could improve yield and income, increasing them by 10% and 30%, respectively. Likewise, considerable implementation of coffee rejuvenation is to build a more sustainable coffee sector at the provincial and national levels.

With respect to the GpP, intermediate costs have declined, on the one hand, by 0.13 thousand USD/ha (18.5%) over two crop years. It could be inferred the following (1) the density is lower than the previous years due to losses crops; and (2) farmers decline their manure or bio-fertilizer use, which is an expensive item (i.e., the price of bio-fertilizer is 1.04 thousand USD/ton). Although selected farmers are well-aware of the important role of organic-fertilizers on increasing yield and maintaining the soil, only 30 % of household produce their bio-organic in different ways, the rest have to buy it from the sales representative.

Simultaneously, the GpP has seemed to show a devaluation of intermediate and hired labor costs during the two crop years. However, this group has higher input costs, such as labor cost and interest payment, which can lead to a constraint if households did not have much savings and an available labor source, than the GpC. Likewise, use of high pesticides and herbicides is harmful to farmers' health, has raised concerns about pesticide residues and their effects on the marketing of products.

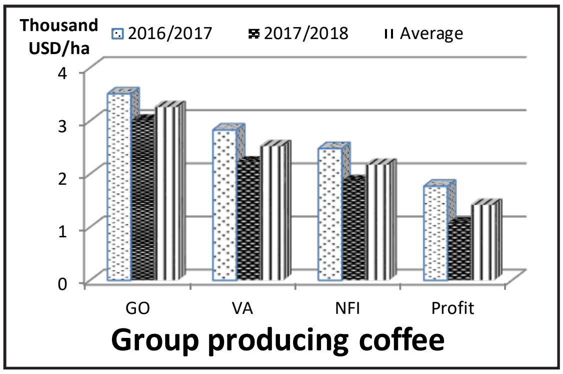
Notably, the interest charge is likely to raise in 2016/2018 and 2017/2018, which leads to an increase in household spending linked to production. For instance, these figures were 0.04 thousand USD/ha for the GpC and 0.06 thousand USD/ha for the GpP in 2016/2017 but they jumped sharply by 60% and 130% in 2017/2018, respectively (Table 4). Even though low prices of commodities, it could be expected that growers maintain their investment for the plantations by borrowing money from banks or informal sources, such as middlemen and collectors or mortgaging the Land Use Certificate, hoping "black gold or golden age" as previous years (over 90% of households mortgaged their Land Use Certificate). In light of this, the authors suggest that the Government should complete more monetary policies towards farmers, offer diversifier financial sources in order to avoid the spontaneous expansion of informal financial channels.

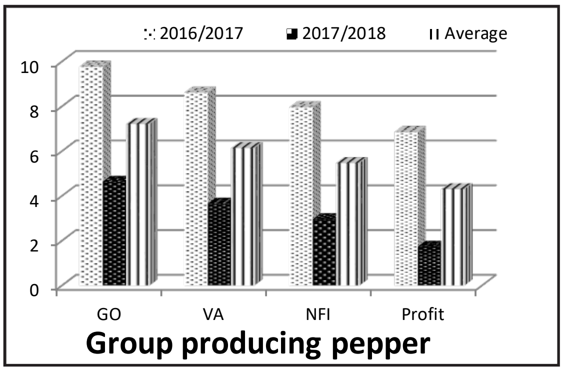
Generally, coffee and pepper farms not only are costly but also face struggles, leading to various challenges of livelihood strategies. Authors suggest that perennial crop farms should be diverse practices. Specifically, modern cultivation like technical application and intercropping farms such as coffee and pepper, coffee and cashew and fruits, which not only generates more income by taking advantage of space but also reduces input costs such as water, fertilizers as well as help environmental protection, should be taken into account. In addition, traditional practices use animal manure and leguminous to boot nitrogen into the soils and control the pests and diseases, as some developed countries and Asia upland counties did (Stillitano et al., 2019; Salazar, 2006; Ho et al., 2017; Romyen et al., 2018; Kunnal and Basavaraj, 2006). According to Lamcombe et al (2016), intergrating perennial crops and livestock sustains smallholder's livelihood. Going forward, it is likely that changing behavior in perennial practices at the farm level is bound to be more than supporting funds directly. What is more, developing "specialty products and bio- products" that has been promoted successfully in Brazil, Indonesia, and Africa is considered to new orientation for farmers (Dak Lak People's Committee, 2019).

Economic performance variance of two groups

Table 5 has displayed a comparison of output for the entire sample over the two crop years in the two farm groups. The findings show that the GpC demonstrates lower economic indicators than the GpP. Specifically, the profitability of the GpC accounted for 1.77 thousand USD/ha in 2016/2017 and 1.07 thousand USD/ha in 2017/2018, whereas these figures were 6.85 thousand and 1.73 thousand USD/ha for the GpP, respectively (Table 5).

Additionally, there has been a significant decrease in economic efficiency in each farm group over the two crop years, GpP rate of decline is for larger than the GpC. This falling value of the GpP are estimated to be around 21% for gross output, 46% for ratio GO and IC, and over 50% for value added, NFI, profit, labor productivity, and ratio of NFI to family labor, whereas the figures of the GpC are 14%, 21%, 23%, 24.5%, 40%, 28% and 33%, respectively, as depicted in Figure 2. It could be expected that the volatile market conditions is a straightforward interpretation of the decrease in economic returns of the GpC and the GpP, specially pepper selling price (Table 5). For instance, because of the fluctuation of market pepper price in mid-2017, the selling price of pepper failed nearly a half price, from 4.78 in 2016/2017 to 2.43 USD/ton in 2017/2018. The findings indicated that the economic performance can be achieved by cost reduction, input-output improvement, increasing revenue and better crop prices. It implies that the reduction in productivity and jumping costs significantly take into consideration when market price is a subjective factor at farm level.





Source: own procesing

Figure 2: The evolution of economic performance between the two groups during two crop seasons.

Although the GpP quantifies higher value than the GpC in both crop years, the rate of reduction is

Items	GpC (n = 62)			GpP (n = 58)		
	2016/2017	2017/2018	Sig	2016/2017	2017/2018	Sig
	Mean	Mean		Mean	Mean	
1. Gross output (GO) (Thousand USD/ha)	3.51	3.01	0.00*	9.77	4.63	0.00*
2. Price selling (Thousand USD/ton)	1.61	1.50		4.78	2.43	
3. Net Value Added (Thousand USD/ha)	2.83	2.22	0.00*	8.61	3.65	0.00*
4. Net farm income (NFI) (Thousand USD/ha)	2.48	1.87	0.00*	7.96	2.96	0.00*
5. Profit (Thousand USD/ha)	1.77	1.07	0.00*	6.85	1.73	0.00*
6. Labor Productivity	0.03	0.02	0.00*	0.06	0.03	0.00*
7. GO/IC	0.26	0.20	0.4	0.49	0.26	0.00*
8. NFI per Family labor	0.03	0.02	0.00*	0.06	0.02	0.00*

Note: The significance levels are indicated as ****p <0.10, **p <0.05,*p <0.01, NS: Non-significant. Mann–Whitney test: 0.00 (excluding family labor cost); a Currency rate: 1USD=23 000 Vietnamese Dong

Source: Author's calculations

Table 5: The variance of economic efficiency of the GpC and the GpP over two crop years.

more significant than that of the GpC. Remarkably, there has been a decrease of the labor productivity and NFI with respect to family labor in the two farm groups during 2016/2017–2017/2018. This seems to contribute serious challenges to production, and raise difficulties in the farmers' life, which will be analyzed in future studies. Under problematic circumstances, the authors suggest that local farmers and the community change their method of growing coffee and pepper in order to mitigate the risk and maintain production. For instance, the authorities should announce certificate application to farmers in suitable region, or/and improve shade tree for plantations in the whole province, which can lead to cutting down environmental impacts, reducing production cost (i.e., irrigated cost) and enhancing households' income from by-products (timber, fruits) (Jezeer at al., 2018; Ho et al., 2018). In Mexico, Peru and Ethiopia, certificated coffee creates the sustainability. In Spain, diversification is as a crop planning strategy for geographic area and products. (Barham and Weber, 2012; Villa et al., 2019). Nonetheless, the Government should attend quality of certificate farms instead of increasing quantity. Besides that, the authorities continuously implement the program of land consolidation through exchanging and regrouping land parcels among households to obtain "massive fields" (cánh đồng mẫu lớn), which is appropriate for technical application or hi-tech zones and marketing channel. Furthermore, the local government should call pests and diseases as an emergency. A massive death of plants should be well-controlled by improvement of producer's awareness, hygiene farms and transferring other crops. Especially, improvement of technical training, guidance on safe production should be recommended. Moreover, encouragement uses live plants as pepper vines instead of concrete and wooden. It is further national and local government should have a wide range of package insurance for distinguished stages in terms of an immature and mature stages.

Factors impacting the change in profitability of the two farm groups

The farm profit inspires farmers to expand or leave the agriculture (Tey and Brindal, 2015). We utilize a technique to examine differences in the level of farm profitability among two groups of successes and un-prosperous farms. Alternatively, the discriminant analysis finds out factors affecting profitability by two categories. The finding of regression expresses determinants of farm productivity including experience, other income, training, and output. Individually, experience, other income, training, and revenue delivered positive effects on the lucrative farms (Table 6).

Tests of Equality of Group Means				
	Wilks' Lambda	F	df1	df2
Experience	0.834	11.558*	1	58
Other income	0.929	4.446**	1	58
Training	0.661	29.696*	1	58
GO	0.939	3.797***	1	58
Eigenvalue = 1.639 (>1)				
Canonical correlation = 0.788				
Wilks Lamda = 0.384, p value = 0.02 (<0.05)				

Note: The significance levels are indicated as ****p <0.10, **p <0.05,*p <0.01
Source: own processing

Table 6: Discriminant analysis test on the impact of factors on the results of the two farm groups.

About unique features, experience and training skills are likely to significantly influence the profitability of farms. This is logical, because perennial crops have a long lifespan and survive through multiple harvest seasons, even after being subjected to high levels of processing. When production is faced with obstacles, experience and training skills related to good agricultural practices reduce risk, save costs, and generate higher income. Especially, the finding shows that training is the most important factor affecting the performance of perennial crop farms with the absolute value of standardized coefficient of 1.133. This means that in comparison to others, farms with household heads who participated in training program tend to be more profit. Regarding local authorities, training programs can guide farmers to adopt technical innovations, preventing diseases, and pruning and shaping technique as well as guidance of diversified farm systems. In other words, farmers have an increase of agricultural knowledge based on which to draw, to decide, what to grow and how to grow it, as well as a better understanding of their resources, weather conditions, and price dynamics. Sustainability-certificate farms, which are less susceptible to environmental influences, develop better economic efficiency than did conventional farms, as reported by (Ho et al., 2018).

Additionally, other income is positively associated with farm profitability. It helps farmers to become a capital-endowed for their plantations to reduce borrowing money from informal sectors. A finding that is in good agreement with the literature (Warren, 2002; Turner and Annamalai, 2012; Diep and Vien, 2017). Additionally, other activities, including agricultural and non-agricultural activities, are shown to generate added value to farmers in addition to coffee and pepper production. Lastly, the study indicates the revenue has positively correlated with the farm profitability. It should be noted that the revenue is not impacted by selling prices because the output data is only used in 2017/2018 crop year. It be inferred, therefore, earnings affected by the farm yield, in which we consider farms and crops profiles.

Indeed, perennial crop production is faced with many quandaries and sustaining in perennial crop farm business is not easy task. Therefore, this exploration generates policy implication to improve farm sustainability. Specifically, in order to maintain output growth, solutions involving intensified and diversified strategies to save costs such as labor, water, pesticides, and fertilizers, should be provided by the local and national governments. On top of that, training skills, which put farmers in a better position to compete in the production and marketing, should be improved. For instance, the Government can use available funds to train farmers and leaders instead of supply cash payments in order to be essential to the spread and longevity of new behavior for social and ethical motivators. In other words, agricultural technology and sustainable agriculture need to design and promote in the coming years.

Conclusion

Generally, perennial crop production involves labor-intensive crops, requires high input cost and has susceptibility to market failure. In this study, by empirically examining the economic benefits of two study groups over two crop years, our study aims to compare the differences of earnings between two groups and identify the evolution of the profitability of these groups. Another point worth noting is determinants of farm profitability. Empirical results provide several important findings. First, during the period of two years, the GpP has had higher variable costs such as pesticides, herbicides and stimulants. For each farm group, the GpP has a loss of variable costs of 0.15 thousand while the figure leaps at 0.13 thousand USD per ha in the GpC. We argue that the GpC is likely to be more suitable for the poor due to lower overall consumption of inputs while the GpP is popular with well-endowed farmers due to high investment cost. Moreover, the GpP is likely to invest more input items and heavily use chemical inputs. It could be inferred that the GpC is more environmentally friendly practices to cope with environmental issues. Therefore, this implies that in comparison to the GpC, the GpP tends to be less sustainable. Second, regarding agronomic benefits, the amount of loss is 21% for gross output, 57% for value added, 62% for NFI, 74% for profit, 50% for labour productivity, and 64 % for ratio of NFI to family labor for the GpP, whereas the rate of loss for the GpC is under 50% for all items. It means that the GpP is likely to be more vulnerable, which may be affected by pests and diseases or weather and market fluctuation, in comparison to the GpC. It implies that it is very important for both farmers and local government agencies to pay more attention to sustainable production of the GpP. We argue that farmers in long term can be encouraged to pay attention to the GpC as well as interventions can be provided coffee production as primary economic field such as subsidies or cash rewards. Given that local government observes divergence across short-run loss and long-run gain in terms of ecological benefits of the GpP. Third, the study reveals that the change in profitability is influenced dramatically by a variety of factors, including experience, training, other income, and gross output, which are positive effects with respect to the farm profitability.

These findings provide an explanation for farmers which can lead to decide their livelihoods and translate into efficiency improvement as well as for policy-makers to design the incentives to improve farmers' knowledge and returns. It means that, this study understanding the current situation of coffee farm and pepper farm groups will help policymakers to develop better interventions to improve perennial crop sustainability and good agricultural practices.

Corresponding authors
Phan Thi Thuy
Département d'Economie et du Développement rural
Gembloux Agro-Bio Tech Université de Liège, 5030 Gemboux, Belgium,
Phone: +84 9 35 34 6969/+32 488338637, E-mail: phanthuydhtn@gmail.com,
PhanThiThuy@student.ulg.ac.be

References

[1] Amarasinghe, U. A., Hoanh, C. T., D'haeze, D. and Hung, T. Q. (2015) "Toward Sustainable Coffee Production in Vietnam: More Coffee with Less Water", *Agricultural Systems*, Vol. 136, pp. 96-105. ISSN 0308-521X. DOI 10.1016/j.agsy.2015.02.008.

[2] Angelsen, A. (2010) "Policies for Reduced Deforestation and Their Impact on Agricultural Production", *Proceedings of the National Academy of Sciences*, Vol. 107, No. 46, pp.19639-19644. ISSN 0027-8424. DOI 10.1073/pnas.0912014107.

[3] Barral, S. (2012) "*Assessing Smallholder Farming. Diagnostic Analysis of Family-Based Agricultural Systems in a Small Region*", 150 p. ISBN 978-971-560-153-5.

[4] Barham, B. L. and Weber, J. G. (2012) "The economic sustainability of certified coffee: Recent evidence from Mexico and Peru", *World Development*, Vol. 40, No. 6, pp. 1269-1279. ISSN 0305-750X. DOI 10.1016/j.agsy.2019.05.006.

[5] Birner, R. and Danielle, R. (2010) "The Political Economy of Policies for Smallholder Agriculture", *World Development*, Vol. 38, No. 10, pp. 1442-1452. ISSN 0305-750X. DOI 10.1016/j.worlddev.2010.06.001.

[6] Chau, T. M. L. (2007) "*The Relative Sustainability of Coffee and Mixed Farmining Systems in Dak Lak Province, Vietnam*", 120 p. [Online]. Available: https://dlc.library.columbia.edu/catalog/ldpd:497638/bytestreams/content/content?filename=Chau_T_M_Long_Thesis.pdf [Accessed: 7 Aug. 2019].

[7] Cochet, H. (2015) "*Comparative Agriculture*", Spinger, 85 p. ISBN 978-94-017-9828-0. DOI 10.1007/978-94-017-9828-0.

[8] Dak Lak People’s Committee (2019) "*Focus on developing specialty coffee areas in the 2018-2019 coffee crop*". [Online]. Available: https://daklak.gov.vn/web/english/-/focus-on-developing-specialty-coffee-areas-in-the-2018-2019-coffee-crop. [Accessed: 7 Aug. 2019].

[9] Devi, K. U. and Pandurangarao, A. (2003) "Economic Analysis of Coffee in Visakhapatnam District of Andhra Pradesh", *Agricultural Economics Research Review*,Vol. 16, No. 1, pp. 36-43. ISSN 0971-3441.

[10] Diep, H. T. N. and Vien, H. T. (2017) "Determinants of Income Diversification and Its Effects on Rural Household Income in Vietnam",*Tạp Chí Khoa Học Trường Đại Học Cần Thơ,* pp.153-162. ISSN 1859-2333.

[11] Diepart, J. and Céline, A. (2018) "*Farming Systems Analysis: A Guidebook for Researchers and Development Practitioners in Myanmar*", GRET, 86 p. [Online]. Available: https://www.gret.org/publication/farming-systems-analysis-guidebook-myanmar/?lang=en [Accessed: 7 Aug. 2019].

[12] Dinh, H. H., Nguyen, T. T., Hoang, V. N. and Wilson, C. (2017) "Economic incentive and factors affecting tree planting of rural households: Evidence from the Central Highlands of Vietnam", *Journal of Forest Economics*, Vol. 29, pp.14-24. ISSN 1104-6899. DOI 10.1016/j.jfe.2017.08.001.

[13] FAO (1999) "*Land Tenure Service Rural Development Division Sustainable Development Department Food and Agriculture Organization*" [Online]. Available: https://www.agter.org/bdf/_docs/fao_1999_agrarian-systems-diagnosis_c.pdf [Accessed: 7 Aug.2019].

[14] Global Strategy (2016) "*Agricultural Cost of Production Statistics Guidelines for Data Collection, Compilation*", Rome. [Online]. Available: http://gsars.org/wp-content/uploads/2016/02/Handbook-on-Agricultural-Cost-of-Production-Statistics-28-02-2016.pdf [Accessed: 7 Aug. 2019].

[15] GSO (2019) "*Agriculture, Hunting, Forestry and Fishing (Statistical Book 2018)*", pp. 339-425. [Online]. Available: https://www.gso.gov.vn/default.aspx?tabid=512&idmid=5&ItemID=19293, [Accessed: 7 Aug.2019].

[16] Gunathilaka, R. P. D., Smart, J. C. and Fleming, C. M (2018) "Adaptation to Climate Change in Perennial Cropping Systems: Options, Barriers and Policy Implications", *Environmental Science & Policy*, Vol. 82, pp. 108-116. ISSN 1462-9011. DOI 10.1016/j.envsci.2018.01.011.

[17] Ha, D. T., and Shively, G. (2008) "Coffee Boom, Coffee Bust and Small Holder Response in Vietnam's Central Highlands", *Review of Development Economics*, Vol. 12, No. 2, pp. 312-326, ISSN 1467-9361. DOI 10.1111/j.1467-9361.2007.00391.x.

[18] Haggar, J. and Schepp, K. (2012) "*Coffee and Climate Change Impacts and Options for Adaption in Brazil, Guatemala, Tanzania and Vietnam*", NRI Working Paper Series: Climate Change, Agriculture and Natural Resources, Natural Resources Institute No. 4. [Online]. Available https://www.academia.edu/2050482/Coffee_and_Climate_Change_Impacts_and_Options_for_Adaptation_in_Brazil_Guatemala_Tanzania_and_Vietnam [Accessed: 7 Aug. 2019].

[19] Herridge, D. F., Win, M. M., Nwe, K. M. M., Kyu, K. L., Win, S. S., Shwe, T., Min, Y. Y., Denton, M. D. and Cornish, P. S. (2019) "The cropping systems of the Central Dry Zone of Myanmar: Productivity constraints and possible solutions", *Agricultural Systems*, Vol. 169, pp. 31-40. ISSN 0308-521X. DOI 10.1016/j.agsy.2018.12.001.

[20] Hill, B. and Bradley, B. D. (2015) "*Comparison of Farmers' Incomes in the EU Member States. Report for European Parliament, Komisja Europejska, DG Internal Policies, Raport dla Parlamentu Europejskiego, Bruksela*" [Online]. Available http://www.europarl.europa.eu/RegData/etudes/STUD/2015/540374/IPOL_STU(2015)540374_EN.pdf [Accessed: 7 Aug. 2019].

[21] Ho, T. Q., Hoang, V. N., Wilson, C. and Nguyen, T. T. (2017) "Which Farming Systems Are Efficient for Vietnamese Coffee Farmers?", *Economic Analysis and Policy*, Vol. 56, pp. 114-125. ISSN 0313-5926. DOI 10.1016/j.eap.2017.09.002.

[22] Ho, T. Q., Hoang, V.-N., Wilson, C. and Nguyen, T.-T. (2018) "Eco-Efficiency Analysis of Sustainability - Certified Coffee Production in Vietnam", *Journal of Cleaner Production*, Vol. 183, pp. 251-260. ISSN 0959-6526. DOI 10.1016/j.jclepro.2018.02.147.

[23] Hurri, S. and Quang, N. N. (2015) "*Report Rural Finance of Coffee Smallholders in Vietnam - Case Study in Dak Nong Province*", IFAD, Ha Noi. [Online]. Available https://asia.ifad.org/documents/10751/084b3711-14d8-42f8-8d40-d604fa93e385. (In Vietnam) [Accessed: 7 Aug. 2019].

[24] IDH (2013) "*Vietnam a Business Case for Sustainable Coffee Production*", Technoserve [Online]. Available: https://www.globalcoffeeplatform.org/assets/files/Documents/Reports-Brochures/3_BusinessCaseEthiopiaUganda_vdPut_Stewart.pdf [Accessed: 7 Aug.2019].

[25] Jezeer, R. E., Santos, M. J., Boot, R. G. A., Junginger, M. and Verweij, P. A. (2018) "Effects of shade and input management on economic performance of small-scale Peruvian coffee systems", *Agricultural Systems*, Vol. 162, p. 179-190. ISSN 0308-521X. DOI 10.1016/j.agsy.2018.01.014.

[26] JICA (2012) "*Agricultural Transformation and Food Security 2040, ASEAN Region with a Focus on Vietnam, Indonesia, and Philippines*", 32 p. [Online]. Available ccsenet.org/journal/index.php/ass [Accessed: 7 Aug. 2019].

[27] Kunnal, L. B., and Basavaraj, H. (2006) "Economic Analysis of Establishment of Pepper as Mixed Crop with Coffee in Chikmagalur District of Karnataka State", *Karnataka Journal of Agricultural Sciences*, Vol. 19, No. 3, pp. 568-573. [Online] Available http://14.139.155.167/test5/index.php/kjas/article/viewFile/756/750 [Accessed: 7 Aug. 2019].

[28] Lacombe, G., Bolliger, A. M., Harrison, R. and Ha, T. T. T. (2016) "Integrated Tree, Crop and Livestock Technologies to Conserve Soil and Water, and Sustain Smallholder Farmers' Livelihoods in Southeast Asian Uplands", In *Integrated Systems Research for Sustainable Smallholder Agriculture in the Central Mekong*, 41 p. World Agroforestry Centre. ISBN 978-604-943-434-1. [Online]. Available https://www.bioversityinternational.org/fileadmin/user_upload/Humidtropics_Raneri.pdf [Accessed: 7 Aug. 2019].

[29] Landau, S. and Everitt, B. S. (2003) "*A Handbook of Statistical Analyses Using SPSS*", Chapman and Hall/CRC. ISBN 9781584883692. [Online]. Available http://www.academia.dk/BiologiskAntropologi/Epidemiologi/PDF/SPSS_Statistical_Analyses_using_SPSS.pdf [Accessed: 7 Aug. 2019]. DOI 10.1201/9780203009765.

[30] Lan, L., Sain, G., Czaplicki, S., Guerten, N., Shikuku, K. M., Grosjean, G. and Läderach, P. (2018) "Farm-level and community aggregate economic impacts of adopting climate smart agricultural practices in three mega environments", *PloS One*, Vol. 13, No. 11. E-ISSN 1932-6203. DOI 10.1371/journal.pone.0207700.

[31] Lososová, J. and Zdenek, R. (2014) "Key factors afecting the profitability of farms in the Czech Republic", *Agris On-line Papers in Economics and Informatics*, Vol. 6, No. 1, pp. 21-36. ISSN 1804-1930. DOI 10.22004/ag.econ.166523.

[32] Newton, A. C., Hodder, K., Cantarello, E., Perrella, L., Birch, J. C., Robins, J., Douglas, S., Moody, Ch. and Cordingley, J. (2012) "Cost–benefit analysis of ecological networks assessed through spatial analysis of ecosystem services", *Journal of Applied Ecology*, Vol. 49, No. 3, pp. 571-580. ISSN 0021-8901. DOI 10.1111/j.1365-2664.2012.02140.x.

[33] NIAPP (2017) "Viet Nam pepper industry development planning to 2020, and vision 2030". [Online]. Available: http://www.niapp.org.vn/info/en/dtcl/giaidoan4/viet-nam-pepper-industry-development-planning-to-2020-and-vision-2030/54161.

[34] Oanh, D. T. (2019) "*Thực Trạng Sản Xuất Hồ Tiêu Tại Tây Nguyên, Vietnam*" (The sistuation of Pepper Production in Central Highlands), Buon Ma Thuot. [Online] Available http://wasi.org.vn/doi-net-ve-thuc-trang-san-xuat-ho-tieu-tai-tay-nguyen-hien-nay [Accessed: 7 Aug. 2019] (in Vietnamese).

[35] Odum, H. T. (1983) "*Systems Ecology; An Introduction*", pp. 644. ISBN-10:0471652776.

[36] Peter, K., Kurian, A., Angelsen and Chopra. V. L. (2003) "*Production Systems and Agronomy| Plantation Crops and Plantations*", Encyclopedia of Applied Plant Sciences, E-ISBN 978012394808, ISBN 9780123948076. DOI 10.1016/B0-12-227050-9/00242-8.

[37] Quah, E. and Mishan, E. J. (2007) "*Cost-Benefit Analysis*", Routledge, 336 p. ISBN-10: 0415349915.

[38] Romyen, A., Sausue, P. and Charenjiratragul, S. (2018) "Investigation of Rubber-Based Intercropping System in Southern Thailand", *Kasetsart Journal of Social Sciences*, Vol. 39, No. 1, pp. 135-142. ISSN 2452-3151. DOI 10.1016/j.kjss.2017.12.002.

[39] Salazar, M. (2006) "An Economic Analysis of Smallholder Coffee Production in Guatemala, Honduras, Nicaragua and Vietnam", M.Sc. Thesis, Faculty of Purdue University, Purdue Purdue. [Online] Available https://pdfs.semanticscholar.org/0c06/e25f99c4a78b6723fd107eb13840ef2e4273.pdf [Accessed: 7 Aug. 2019].

[40] Salvatore, D. and Reagle., D. (2002) "*Statistics and Econometrics, Schaum's Outline Series*", McGraw–Hill. ISBN 9780071755474.

[41] Scherr, S. J, Kedar, M. J. and Christine, N. (2015) "Chapter 7: Coffee in Dak Lak, Vietnam", In book *Steps toward Green: Policy Responses to the Environmental Footprint of Commodity Agriculture in East and Southeast Asia*", EcoAgriculture Partners, 220 p. [Online] Available https://ecoagriculture.org/wp-content/uploads/2015/08/Steps-Toward-Green-Book-File-Final-for-Upload.pdf [Accessed: 7 Aug. 2019].

[42] Slater, R., Peskett, L., Ludi, E. and Brown, D. (2007) "Climate Change, Agricultural Policy and Poverty Reduction - How Much Do We Know?", *ODI Natural Resource Perspectives*, 109 p. [Online]. Available: https://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/1850.pdf [Accessed: 7 Aug. 2019].

[43] Stillitano, T., Falcone, G., Nicolò, B. F., Di Girolamo, C., Gulisano, G. and De Luca, A. I. (2019) "Technical Efficiency Assessment of Intensive and Traditional Olive Farms in Southern Italy", *Agris On-Line Papers in Economics and Informatic*, Vol. 11, No. 2, pp. 81-93. ISSN 1804-1930. DOI 10.7160/aol.2019.110208.

[44] Susmita, D. (2007) "Pesticide Poisoning of Farm Workers–Implications of Blood Test Results from Vietnam", *International Journal of Hygiene and Environmental Health*, Vol. 210, No. 2, pp. 121-132. ISSN 1438-4639. DOI 10.1016/j.ijheh.2006.08.006.

[45] Thái, H. A. C. (2018) "*Livelihood Pathways of Indigenous People in Vietnam's Central Highlands: Exploring Land-Use Change*" (In Vietnamese), Springer. ISBN 978-3-319-71171-3.

[46] Thi Duong Nga, N. and Thuy, P. T. (2017) "Economic Performance of Coffee and Pepper Intercropping in Quang Hiep Commune, Cumgar District, Dak Lak Province", *Vietnam Journal Agriculture Science*, Vol. 15, No. 9, pp. 1277-1287. ISSN 1859-0004.

[47] Thong, P. H (2015) "*Tài Liệu Đào Tạo Nông Dân Kỹ Thuật Sản Xuất Cà Phê Bền Vững*" (Coffee, Technical for Sustainable Production of Development), HDI, VNU. [Online] Available http://gcpconnect.coffee/sites/gcp-connect.org/files/TOF.pdf [Accessed: 7 Aug. 2019].

[48] Thong, L., Thao, N. Q. and Viet, P. C. (2017) "*Sách Giáo Khoa Bài Tập Địa Lý 12*" (Geographical Book of Vietnam). ISBN 9786040149558.

[49] Thuy, P., Niem, L., Ho, T., Burny, P. and Lebailly, P. (2019a) "Economic Analysis of Perennial Crop Systems in Dak Lak Province, Vietnam", *Sustainability*, Vol. 11, No. 1. E-ISSN 2071-1050. DOI 10.3390/su11010081.

[50] Thuy, P., Niem, L., Ho, T., Burny, P. and Lebailly, P. (2019b) "Which perennial crop farm approach generates more profitability? A case study in Dak Lak province, Vietnam", *Asian Social Science*, Vol. 15, No. 9. ISSN 1911-2017. DOI 10.5539/ass.v15n9p1.

[51] Thuy, T. T. T. (2010) "Incidence and Effect of Meloidogyne Incognita (Nematoda: Meloidogyninae) on Black Pepper Plants in Vietnam", Thesis, Katholieke Universiteit Leuven, België. Unpublish.

[52] Tey, Y. S. and Brindal, M. (2015) "Factors influencing farm profitability", *Sustainable Agroculture Reviews*, Vol. 15, pp. 235-255. Springer, Cham. ISSN 2210-4410. DOI 10.1007/978-3-319-09132-7_5.

[53] Ton, N. T. and Buu, B. C. (2011) "How to Prevent the Most Serious Diseases of Black Pepper (Piper Nigrum L.) – A Case Study of Vietnam", In *PC Annual Meeting in Lombok*, Indonesia. [Online]. Available: http://www.ipcnet.org/admin/data/ses/1329362855thumb.pdf [Accessed: 7 Aug. 2019]

[54] Turner, A. G., and Annamalai, H. (2012) "Climate Change and the South Asian Summer Monsoon", *Nature Climate Change*, Vol. 2, No. 8. E-ISSN 1758-6798. ISSN 1758-678X. DOI 10.1038/nclimate1495.

[55] Trong, H., Ngoc, C. N. M. (2008) "*Phân Tích Dữ Liệu Nghiên Cứu Với SPSS*" (Analysis Data for Research with SPSS), Nhà Xuất Bản Hồng Đức. [Online] Available https://www.sachmoi.net/phan-tich-du-lieu-nghien-cuu-voi-spss/ [Accessed: 7 Aug. 2019]. ISBN 506404262974.

[56] Villa, G., Adenso-Díaz, B. and Lozano, S. (2019) "An analysis of geographic and product diversification in crop planning strategy, *Agricultural Systems*, Vol. 174, pp. 117-124. ISSN 0308-521X. DOI 10.1016/j.worlddev.2011.11.005.

[57] Warren, P. (2002) "*Livelihoods Diversification and Enterprise Development: An Initial Exploration of Concepts and Issues. Livelihood Support Program (LSP)*", Food and Agriculture Organization of the United Nations working paper 4. [Online]. Available: http://www.fao.org/tempref/docrep/fao/008/j2816e/j2816e00.pdf. [Accessed: 7 Aug. 2019].

[58] World Bank (2016) "*Transforming Vietnamese Agriculture: Gaining More from Less*", Work Bank Group. [Online]. Available: http://documents.worldbank.org/curated/en/116761474894023632/pdf/108510-WP-PUBLIC.pdf [Accessed: 7 Aug. 2019].

Agris on-line Papers in Economics and Informatics

Volume XII Number 3, 2020

Comparison of Fuzzy Multi-Criteria Decision-Making Methods to Rank Business Strategies and Marketing Resources

Amirhossein Tohidi[1], Mohammad Ghorbani[1], Ali-Reza Karbasi[1], Ahmadreza Asgharpourmasouleh[2], Behrooz Hassani-Mahmooei[3]

[1] Department of Agricultural Economics, Faculty of Agriculture, Ferdowsi University of Mashhad, Islamic Republic of Iran

[2] Department of Social Sciences, Faculty of Letters and Humanities, Ferdowsi University of Mashhad, Islamic Republic of Iran

[3] Monash University, Melbourne, Australia

Abstract

Given the growing competition in domestic and international agricultural product markets, choosing a business strategy compatible with requirements of marketing resources can guide agro-food firms to maintain and enhance competitive advantages. However, this is not as simple as it seems because the decision-making criteria expressed in a fuzzy manner and the relationship between them can be hierarchical or network-based. Therefore, the main goal of this study was to select the most suitable business strategy and to prioritize marketing resources for one of the major agro-food firms in Iran. To ensure the robustness of the results, both fuzzy analytic hierarchy process (AHP) and fuzzy analytic network process (ANP) were applied to prioritize business strategies and marketing resources. The results of both methods revealed that the differentiation strategy had the highest priority in terms of the experts' viewpoints. The results also showed that managerial and customer relationship capabilities were the most important criteria in selecting the differentiation strategy. According to the findings of the study, for the successful implementation of the differentiation strategy, company managers are recommended to take the following three main elements into huge consideration: financial conditions, paying attention to customer's needs and requirements, and the introduction of new products and services.

Keywords

Agro-food, strategic management, problem-solving.

Tohidi, A., Ghorbani, M., Karbasi, A.-R., Asgharpourmasouleh, A. and Hassani-Mahmooei, B. (2020) "Comparison of Fuzzy Multi-Criteria Decision-Making Methods to Rank Business Strategies and Marketing Resources", *AGRIS on-line Papers in Economics and Informatics*, Vol. 12, No. 3, pp. 101-114. ISSN 1804-1930. DOI 10.7160/aol.2020.120309.

Introduction

Spices are produced from plant materials which are rich in vitamins, minerals, and antioxidants. Therefore, in many countries, spices are used in the food industry for coloring, preserving, and flavoring food products. Among a wide variety of spices, saffron is a rare and expensive spice extracted from the flower stigmas of Crocus sativus L., belonging to the family Iridaceae. Owing to its health benefits and flavor attributes, saffron is mainly used in the food industry. Currently, about 90% of saffron is produced in Iran (Khilare et al., 2019). However, the Iranian agricultural sector is faced with major problems of post-production; the main source of these problems is the lack of relevant business strategies (Mohammadi et al., 2017). The development of a well-defined business strategy and the identification of important resources associated with it will help saffron companies perform better financially in both domestic and international markets; subsequently, they can achieve competitive advantages in these markets.

Over the past three decades, dramatic changes have occurred in the agro-food sector around the world, which has made agro-food markets increasingly competitive and complex. Due to these changes, agribusiness firms are facing new competitive pressures. Therefore, strategic management and planning of agro-food products is necessary to cope with these changes. In fact, each agro-food company should select and apply a business strategy that is tailored to the company's resources and capabilities (Chen et al., 2016). The goal of strategic management is to enable a firm to choose a strategy which can create a competitive advantage (Barney and Hesterly, 2015). There are several definitions of competitive advantage in the strategic management literature, each of which sometimes has a different meaning (Sigalas and Pekka Economou, 2013). In most definitions, there is an emphasis on the existence of a business strategy to achieve competitive advantage. A well-designed business strategy shifts the organization's focus onto environment, structures, and processes which affect how successfully a company meets its objectives (Yoshikuni and Albertin, 2018). Choosing the most appropriate business strategy is crucial because if a firm cannot recognize the linkage between operating decisions and business strategy, it may suffer a non-competitive production system which is costly and time-consuming to modify (Banchuen et al., 2017).

Although researchers have suggested different typologies of business strategy over the past decades (see Anwar and Hasnu, 2016, for more details), the generic strategies proposed by Porter (1980), namely cost leadership, differentiation, and focus are still applicable and widely used in the strategic management studies (e.g., Altuntas and Yilmaz, 2016; Balci et al., 2018; Goddard and Simm, 2017; Rexhepi and Srhoj, 2018). The term "generic" means all firms can potentially implement these types of strategies, regardless of whether they operate in the manufacturing, service or nonprofit sectors; in fact, any company active in any industry can choose and implement each of them (Hill and Jones, 2013). The cost leadership strategy is an integrated set of operations performed to produce goods or services which are acceptable to customers at a lower cost than rivals. The differentiation strategy is an integrated set of operations performed to produce goods or services with different features which are important to customers. A focus strategy involves an integrated set of activities for the production of goods or services which satisfy the needs of a particular industry segment (Hitt et al., 2016). Therefore, by implementing a focus strategy, companies can increase their shares through operating in a narrow niche market more efficiently than larger competitors (Ulubeyli et al., 2018). According to Porter's typology of generic strategies, a firm is unlikely to gain a competitive advantage if it fails to implement its strategy in at least one of the three generic strategies. Such a firm will effectively find itself "stuck in the middle", and will have low profitability (Grimmer, 2019).

The sustainable competitive advantage is the result of organizational resources which are the part of a business strategy and are essential for the overall firm performance (Bendickson and Chandler, 2019). Previous studies have shown that the performance of a firm depends on its resources (e.g. Altuntas and Yilmaz, 2016; Khan et al., 2019; Osakwe and Anaza, 2018; Sok et al., 2017). Indeed, in recent decades, scholarly attention has shifted from the strategic group perspective, which focuses on the strategic factors common among the industries, to the resource-based view (RBV) (Gomes et al., 2014). The RBV posits that the difference in the performance of companies over time is mainly due to their resources rather than the industry's structural characteristics (Hitt et al., 2016).

In general, in order to achieve superior performance, the firm resources must comply with the requirements of the business strategy (Wu et al., 2010b). Therefore, decisions based on the company's resources have a prominent role in the firm's strategy formulation process. Making such decisions may seem simple, but this, like other management tasks, is difficult and challenging and is tied to company success. Failure of half of the organizational decisions reflects the difficulty of decision-making process in this task (Hitt et al., 2016). Accordingly, an appropriate decision-making method is critical to the success of the organization and its competitive advantages. In order to make strategic decisions in a systematic and structured way, multiple criteria decision-making (MCDM) methods have been developed to rank strategic alternatives and to choose one over the other regarding several, sometimes conflicting, criteria (Haddad and Sanders, 2018). Based on RBV and Porter's model of generic strategies, some studies have been carried out to rank the business strategies using MCDM method. However, these studies have focused on manufacturing (Mohaghar et al., 2012), hospitality (Wu et al., 2010a) and service (Lin, Lee, and Chen, 2009; Lin, Lee, and Wu, 2009; Lin and Wu, 2008; Wu et al., 2010b) companies, but less attention has been given to those firms operating in agro-food industries. Therefore, the main contribution of this study was to prioritize business strategies and marketing resources for a major saffron firm in Iran.

Making use of MCDM method can be helpful in choosing business strategies in the agricultural sector because decision-making in this economic sector is not based on simple and specific rules, but largely on decision-makers' knowledge and experience (Morris and Mishra, 2014). Decision-makers are believed to be more confident about interval judgments than fixed value ones. In fact, due to the fuzzy nature of the decision-making process, they cannot express their opinions clearly and explicitly (Koulinas et al., 2019). Hence, the next contribution of this study was to use fuzzy AHP and fuzzy ANP models to determine the most appropriate business strategy.

Materials and methods

Decision-making is a structured approach used to guide decision-makers to determine the best solution among different conflicting decisions by considering various sets of criteria at the same time (Deng and Jiang, 2018). Therefore, strategic decision-making is complex, particularly when several tangible and intangible criteria are considered simultaneously to select the best alternative (Adetunji et al., 2018). MCDM is an important branch of the modern decision theory, industrial engineering, and management systems which has many applications in various sciences (Guo and Zhao, 2017). It involves a set of approaches and methods which help decision-makers make better decisions (Mahmoudkelaye et al., 2018). Among the MCDM techniques, AHP and ANP are commonly used by researchers to rank and select business strategies (Gedela et al., 2018). AHP decomposes a complex decision problem into several levels which make a one-way hierarchical structure. In this structure, the goal, the criteria, and the alternatives are respectively located in the upper, middle and lower levels of the hierarchy. AHP is conceptually and practically simple, but in its hierarchical structure, its elements whose interactions are ignored are assumed to be independent of each other; it is not consistent with many real-world problems (Montesinos-Valera et al., 2017). Overcoming the limitations of AHP, ANP can incorporate more complex dependencies and interactions between elements within a decision context (Brožová et al., 2016; Hornická and Brožová, 2013). In conventional methods of MCDM such as AHP and ANP, however, the relative importance of criteria, sub-criteria, and alternatives is expressed as exact (crisp) numbers. Given the vague, fuzzy, and uncertain nature of human judgments in problem-solving, crisp numbers are believed not to be fully expressed or captured by the decision-makers' viewpoints, and therefore the use of conventional MCDM methods is not suitable to solve the real-world problems. To overcome this limitation, several methods have been developed by incorporating fuzzy set theory into MCDM problems. In fact, fuzzy multi-criteria decision-making (FMCDM) approaches are well suited to deal with uncertainty and vagueness of human thinking (Siddiquie et al., 2017). Therefore, in this study, fuzzy AHP and fuzzy ANP methods were applied to rank business strategies the descriptions of which are given below.

Metamathematics of fuzzy logic

A fuzzy number, represented by $\tilde{A}$, is a fuzzy subset of real numbers, and its membership function is expressed as $\mu\tilde{A}(x)$: $R \rightarrow [0,1]$, where x denotes the evaluated criteria set in a decision-making problem (Ashtiani and Abdollahi Azgomi, 2016). The membership degree function of a triangular fuzzy number (TFN) is expressed as follows (Hsieh et al., 2004):

$$\mu_{\tilde{A}}(x) = \begin{cases} (x-L)/(M-L), & L \le x \le M, \\ (U-x)/(U-M), & M \le x \le U, \\ 0, & otherwise, \end{cases} \quad (1)$$

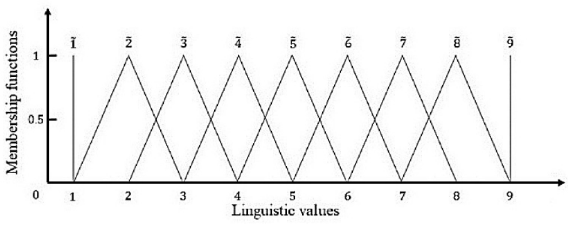
where L, M, U are the upper, middle and lower bounds respectively. Thus, a TFN can be expressed as a triplet (L, M, U) where L≤M≤U. The mathematical operations of two TFNs, $\tilde{A}_1 = (L_1,M_1,U_1)$ and $\tilde{A}_2 = (L_2,M_2,U_2)$ are given in Table 1.

Operation	Fuzzy equivalent	
Addition	$\tilde{A}_1 \oplus \tilde{A}_2 = (L_1, M_1, U_1) \oplus (L_2,M_2,U_2) = (L+L_2, M_1+M_2, U_1+U_2)$	
Multiplication	$\tilde{A}_1 \otimes \tilde{A}_2 = (L_1, M_1, U_1) \otimes (L_2, M_2,U_2) = (L_1L_2, M_1M_2, U_1U_2)$	for $L_i>0$, $M_i>0$, $U_i>0$
Subtraction	$\tilde{A}_1 \ominus \tilde{A}_2 = (L_1, M_1, U_1) \ominus (L_2,M_2,U_2) = (L_1-U_2, M_1-M_2, U_1-L_2)$	
Division	$\tilde{A}_1 \oslash \tilde{A}_2 = (L_1, M_1, U_1) \oslash (L_2,M_2,U_2) = (L_1/U_2, M_1/M_2, U_1/L_2)$	for $L_i>0$, $M_i>0$, $U_i>0$
Reciprocal	$\tilde{A}_1^{-1} = (L_1, M_1, U_1)^{-1} = (1/U_1, 1/M_1, 1/L_1)$	for $L_i>0$, $M_i>0$, $U_i>0$

Source: Hsieh et al. (2004)

Table 1: The mathematical operations of two TFNs.

In practically complex decision problems, where there is a degree of uncertainty and fuzziness in the decision-making process, making use of linguistic variables can be useful for evaluating criteria and alternatives. In fact, linguistic variables reflect human knowledge and experience which are often hard to quantify using exact numbers (Ashtiani and Abdollahi Azgomi, 2016). The values of a linguistic variable are not expressed numerically but are presented as words or sentences in a natural or artificial language (Hsieh et al., 2004). As shown in Table 2, criteria and alternatives were evaluated and ranked in this study using nine linguistic terms ranging from "equally important" to "extremely more important". Then, linguistic terms can be converted into fuzzy numbers using Figure 1 and Table 2.



Source: Hsu et al. (2010)

Figure 1: Membership functions for linguistic values.

Fuzzy number	Linguistic scales	Scale of fuzzy number
$\tilde{1}$	Equally important	(1, 1, 1)
$\tilde{2}$	Judgment values between equally and moderately	(1, 2, 3)
$\tilde{3}$	Moderately more important	(2, 3, 4)
$\tilde{4}$	Judgment values between moderately and strongly	(3, 4, 5)
$\tilde{5}$	Strongly more important	(4, 5, 6)
$\tilde{6}$	Judgment values between strongly and very strongly	(5, 6, 7)
$\tilde{7}$	Very strongly more important	(6, 7, 8)
$\tilde{8}$	Judgment values between very strongly and extremely	(7, 8, 9)
$\tilde{9}$	Extremely more important	(9, 9, 9)

Source: own processing according to Hsu et al. (2010) and Lee et al. (2008).

Table 2: Fuzzy numbers based on linguistic terms.

The fuzzy analytic hierarchy process

Owing to the shortcomings of the AHP method in transforming subjective judgments into quantitative ones, the fuzzy AHP methodology was proposed and developed by the authors to solve the hierarchical decision problems in fuzzy environments. So far, various fuzzy AHP methods have been introduced in the FMCDM literature. However, some suggested methods such as Chang's (1996's) extent analysis have been criticized for not fully utilizing all the information on the fuzzy pairwise comparison matrices which may result in assigning an irrational zero weight to some useful decision criteria (Gul et al., 2017). The method proposed by Buckley (1985) does not have limitations of other techniques (Gul et al., 2017; Yazdi, 2017). Buckley's fuzzy AHP approach contains the following steps (Acar et al., 2018; Gul et al., 2017):

Step 1. Determination of criteria, sub-criteria, and alternatives in the framework of a hierarchical structure.

Step 2. Using linguistic expressions and uncertain numbers: fuzzy judgment matrices of pairwise comparisons are constructed in each of the hierarchy levels. In pairwise comparisons, each expert k determines which of the two criteria is more important and then assigns a linguistic value a_{ij} to show the degree of the relative importance of a criterion over another.

$$\tilde{A}^k = \begin{bmatrix} \tilde{1} & \tilde{a}_{12} & \cdots & \tilde{a}_{1n} \\ \tilde{a}_{21} & \tilde{1} & \cdots & \tilde{a}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{a}_{n1} & \tilde{a}_{n2} & \cdots & \tilde{1} \end{bmatrix} = \begin{bmatrix} \tilde{1} & \tilde{a}_{12} & \cdots & \tilde{a}_{1n} \\ 1/\tilde{a}_{12} & \tilde{1} & \cdots & \tilde{a}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 1/\tilde{a}_{1n} & 1/\tilde{a}_{2n} & \cdots & \tilde{1} \end{bmatrix} \quad (2)$$

In the above pairwise comparison matrix, if two criteria have the same importance, let $\tilde{1} = (1,1,1)$; the values of $\tilde{1} - \tilde{9}$ represent the relative importance of criterion i with respect to the criterion j, and the values of $\tilde{1}^{-1} - \tilde{9}^{-1}$ reflect the relative importance of criterion j as compared to criterion i.

Step 3. The aggregation of experts' preferences can only be done if their judgments are consistent. Therefore, the consistency rates are used to ensure that the experts' judgments are actually reliable. If consistency rates are less than or equal to 0.10, judgments can be considered consistent.

Step 4. After confirming the consistency of judgments in every pairwise comparison matrix, experts' opinions can be aggregated using the geometric mean method. The resulting aggregate matrix is shown in Equation (3).

$$\tilde{C} = \begin{bmatrix} \tilde{1} & \tilde{c}_{12} & \cdots & \tilde{c}_{1n} \\ \tilde{c}_{21} & \tilde{1} & \cdots & \tilde{c}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{c}_{n1} & \tilde{c}_{n2} & \cdots & \tilde{1} \end{bmatrix} \quad (3)$$

Such as

$$\tilde{c}_{ij} = \sqrt[q]{\prod_{k=1}^{q} \tilde{a}_{ij}^{k}} \quad (4)$$

Where k identifies the individual experts, and q represents the total number of decision-makers.

Step 5. The fuzzy geometric mean for each row of the matrix is calculated using the geometric mean method.

$$\tilde{r}_i = (\tilde{c}_{i1} \otimes \tilde{c}_{i2} \otimes \cdots \otimes \tilde{c}_{in})^{\frac{1}{n}} \quad (5)$$

Step 6. The fuzzy weights of each criterion, sub-criterion, and alternatives are obtained as follows:

$$\tilde{w}_i = \tilde{r}_i \otimes (\tilde{r}_1 \oplus \tilde{r}_2 \oplus \cdots \oplus \tilde{r}_n)^{-1} \quad (6)$$

Where $\tilde{w}_i$ can be represented as a TFN, $\tilde{w}_i = (Lw_i, Mw_i, Uw_i)$. Here, Lw_i, Mw_i, Uw_i are the upper, middle and lower bounds of the fuzzy weights respectively.

Step 7. The median method is applied to defuzzified fuzzy weights into crisp ones.

$$w_i = \frac{Lw_i + 2Mw_i + Uw_i}{4} \quad (7)$$

Step 8. The crisp weights wi are normalized to calculate local priorities of the criteria. To obtain the global or overall priorities of the sub-criteria, their local priorities are multiplied by the weight of the covering criterion. Then, the final priority of the alternative i (S_i) is obtained using the Equation (8).

$$S_i = \sum_{j=1}^{n} w_j s_{ij}, \forall i \quad (8)$$

Where w_j denotes the overall priority of the criterion j and s_{ij} represents the importance of the alternative i according to the criterion j.

The fuzzy analytic network process

Unlike the AHP method, ANP can consider the interrelationship between criteria in the decision-making process. Using the network analysis instead of the hierarchical structure, interrelations between decision-making criteria are measured in the ANP method. Like the conventional AHP, the ANP model cannot deal with the uncertainty and ambiguity inherent in human judgments. In addition to taking into account the interrelationship between criteria, the fuzzy ANP model also considers the uncertainty in the prioritization process of strategic choices. The fuzzy ANP model based on the Buckley's method contains the following steps (Sadeghi and Larimian, 2018):

Step 1. Determination of criteria, sub-criteria, and alternatives in the framework of a network structure.

Step 2. Like the fuzzy AHP method, weights of criteria, sub-criteria, and alternatives are determined based on the experts' judgments by performing pairwise comparisons between two clusters, or two elements, or two decision alternatives. Accordingly, Figure 1 and Table 2 are used to assess the relative importance of clusters and elements and their effects on each other.

Step 3. This step is performed in accordance with steps 3 and 4 in the fuzzy AHP method.

Step 4. Assuming the independence of the network factors from each other, the local priorities of each cluster and its elements are calculated according to the Buckley's method. Then, the fuzzy local weights are converted into crisp local priorities using the Equation (7).

Step 5. After aggregating the expert's judgments and calculating the local priorities of the components in the ANP network, these priorities are entered into the appropriate columns of an unweighted supermatrix. In the fuzzy ANP method, the unweighted supermatrix is used to express the relationships and interactions between the network components.

Step 6. To calculate a weighted or stochastic supermatrix (i.e. a matrix whose sum of column elements is equal to one), it is necessary to multiply the blocks of the unweighted supermatrix by the corresponding cluster weight. Then, the weighted supermatrix is raised to a sufficiently large power until convergence is achieved and the weights remain stable. The resulting matrix is known as the limit supermatrix, and it can measure all the direct and indirect effects between the elements and the clusters.

Results and discussion

In this study, data was collected from 30 company experts using a survey questionnaire based on pairwise comparisons between elements of a decision-making problem. Determination of a goal, criteria, sub-criteria, and alternatives is the first step in solving an MCDM problem. A variety of strategic frameworks can be considered decision-making options. However, the generic strategy framework of Porter is one of the most widely used tools to study the strategic behavior of organizations in business environments. Porter's framework of generic strategies is in line with other classifications. Accordingly, inferences derived from the Porter's strategy theory can also be obtained using other classifications (Wu et al., 2015). Therefore, in this study, Porter's generic strategies were used as decision-making alternatives.

Given the necessity of linkage between the business strategy and organization resources, the RBV was applied in this study to determine the criteria and sub-criteria of fuzzy AHP and fuzzy ANP methods. In general, market-based and marketing support resources are two types of value-creating resources. Market-based resources are those which can directly maintain or create competitive advantages and can be used immediately in the market, whereas marketing-support resources indirectly affect competitive advantages and have a supportive role for marketing activities. Among the market-based resources, customer relationship capabilities are the first and the most important source for each organization. Customer relationship capabilities include identifying the customers' needs and wants along with the ability to successfully establish relationships between them. The second set of market-based resources is the credibility and reputation of the firm among its suppliers and customers. This set is known as "reputational assets". Market innovation capabilities are the third most important market-based resource. Characteristics of market innovation capabilities include: complexity, relying on learning and tactile skills, the difficulty of identifying the causes of success and not being easily duplicated from one organization to another one. The human resources of the organization are another set of market-based resources. In human resources literature, the emphasis has been placed on the importance of personnel management and staff development in increasing motivation and loyalty (Hooley et al., 2005).

Among the marketing support resources, managerial capabilities play a crucial role in supporting and underlying market-based resources. Managerial capabilities represent the management expertise and processes in the company for leveraging its market-based resources to gain a competitive advantage (Graves and Thomas, 2006).

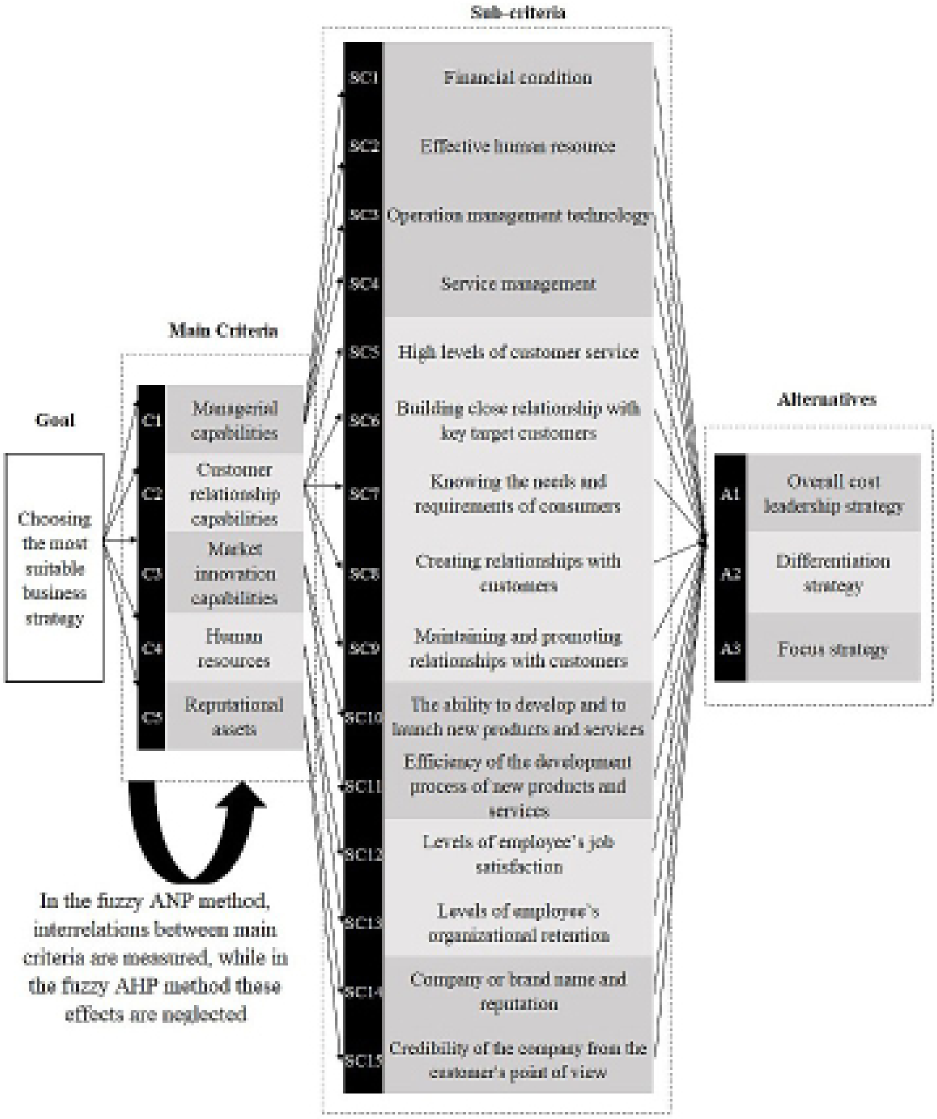
Due to real-world empirical applications, the five criteria mentioned above have been accepted and used by researchers to select the most appropriate business strategy. Each of the assessment criteria consisting of several sub-criterion indications is shown in Figure 2. These sub-criteria have been selected based on the experts' opinions and previous studies (Altuntas and Yilmaz, 2016; Hooley et al., 2005; Lin, Lee, and Wu, 2009).

According to the instructions proposed for fuzzy AHP and fuzzy ANP methods, pairwise comparisons are made using linguistic variables and fuzzy numbers. It should be noted that internal relations between the five main criteria are considered in the fuzzy ANP method. Thus, the number of pairwise comparisons in the fuzzy ANP is 30 times more than that of the fuzzy AHP.

In the next step, after ensuring the consistency of judgment matrices for each expert, these matrices are aggregated using the geometric mean method (see Equation 4). For brevity, all aggregated pairwise comparison matrices are not reported in tables; only pairwise comparison matrix of main criteria is presented in Table 3 with respect to the goal .The TFNs located in the second row and the third column of Table 3 show that "managerial capabilities" are more important than the "customer relationship capabilities". In the fuzzy AHP method, crisp local priorities of the criteria are calculated using the equations (5), (6) and (7) (Table 4). In the fuzzy ANP method, the local weights form an unweighted supermatrix. Then, the blocks of this unweighted supermatrix are multiplied by the weight of the corresponding cluster to create a weighted supermatrix. The final weights in the fuzzy ANP method (which includes direct and indirect effects) are obtained by raising the weighted supermatrix to a sufficiently large power whose values for the main criteria are reported in Table 4.

According to Table 4, the results of the fuzzy AHP and fuzzy ANP methods indicate that "managerial capabilities" and "customer relationship capabilities" are more important than the other three criteria in choosing the most appropriate business strategy which is in accordance with the results of previous studies (Altuntas and Yilmaz, 2016; Wu et al., 2010b). In general, the success of agro-food firms depends largely on managerial capabilities because these capabilities are essential for identifying and developing market-based resources. Moreover, managerial capabilities are an important source of revenue growth for maintaining a competitive advantage. Therefore, managerial skills play a critical role in linking strategic decisions and business performance (Lorenzo et al., 2018).

As shown in Table 4, customer relationship



Source: own processing

Figure 2: A conceptual model of the strategic decision-making process.

Goal	C1	C2	C3	C4	C5
C1	(1.000, 1.000, 1.000)	(1.131, 1.621, 2.174)	(2.034, 2.947, 3.846)	(2.359, 3.160, 4.082)	(1.886, 2.684, 3.584)
C2	(0.460, 0.617, 0.884)	(1.000, 1.000, 1.000)	(1.630, 2.363, 3.141)	(2.007, 2.631, 3.367)	(1.570, 2.232, 2.917)
C3	(0.260, 0.339, 0.492)	(0.318, 0.423, 0.613)	(1.000, 1.000, 1.000)	(0.753, 1.025, 1.416)	(0.863, 1.185, 1.613)
C4	(0.245, 0.316, 0.424)	(0.297, 0.380, 0.498)	(0.706, 0.975, 1.328)	(1.000, 1.000, 1.000)	(0.745, 1.112, 1.578)
C5	(0.279, 0.373, 0.530)	(0.343, 0.448, 0.637)	(0.620, 0.844, 1.159)	(0.634, 0.899, 1.341)	(1.000, 1.000, 1.000)

Source: own processing (The symbols used in the table are defined in Figure 2).

Table 3: Aggregated fuzzy pairwise comparison matrix of main criteria with respect to the overall goal.

Criteria	Fuzzy AHP	Fuzzy ANP
Managerial capabilities	0.366	0.304
Customer relationship capabilities	0.270	0.263
Market innovation capabilities	0.127	0.146
Human resources	0.118	0.149
Reputational assets	0.119	0.139

Source: own processing

Table 4: The weights of the main criteria obtained from the fuzzy AHP and fuzzy ANP methods.

capabilities are the second most important marketing resource in determining the business strategy. In fact, customer relationship capabilities represent the company's ability to build and develop close relationships with customers. Historically, in many business transactions little attention has been paid to establish a beneficial, long-term, and mutually beneficial relationship between the buyer and the seller. Today, agro-food firms should pay attention to consumers' demand and supply and the products that meet their needs. Hence, customer relationship capabilities have a significant impact on customer's loyalty and satisfaction, and consequently, the performance of an agro-food enterprise (Dentoni et al., 2014).

The results of the prioritization of the sub-criteria are reported in Table 5. Given the assumptions of the proposed model and the lack of interrelations between the sub-criteria, the values of the weights obtained from the fuzzy AHP method are the same as the ones calculated from the fuzzy ANP method.

As shown in Table 5, among the sub-criteria of "managerial capabilities", "financial condition" (0.368) is perceived to be the most important factor for developing managerial competencies. In order to succeed, agro-food enterprises should have managers who analyze the financial conditions of the company and then make the necessary decisions to achieve the company's goals and to maintain a competitive advantage (Barnard et al., 2016).

According to Table 5, the final priorities indicate that "knowing the needs and requirements of consumers" (0.307) and "building close relationship with key target customers" (0.292) are the most important factors among the sub-criteria related to "customer relationship capabilities". A market-driven agro-food company seeks to build a long-term, close relationship with its target customers. The managers of this type of organization try to understand the expectations of target customers and meet their needs and requirements, both in the current and in the lifetime of the relationship. In market-driven agro-food firms, focusing on knowing target customers and building close relationships with them are the key drivers of all strategic organizational decisions (Barnard et al., 2016).

Concerning the criterion "market innovation capabilities", the results of Table 5 show that "the ability to develop and to launch new products and services" is more important than the "efficiency of the development process of new products and services". In today's competitive markets, innovative agro-food companies seek to introduce new products and services using marketing resources and business strategies. Successful agro-food firms can satisfy the changing needs of customers by launching new products and services. Market-driven agro-food enterprises are believed to achieve more success in introducing new products and services to consumers; it leads to more profitability (Mirzaei et al., 2016).

According to Table 5, "levels of employee's organizational retention" is much important than "levels of employee's job satisfaction". Having

Criteria	**Sub-criteria**	**Weights**
Managerial capabilities	*Financial condition*	0.368
	Effective human resource	0.246
	Operation management technology	0.240
	Service management	0.146
Customer relationship capabilities	*High levels of customer service*	0.122
	Building close relationship with key target customers	0.292
	Knowing the needs and requirements of consumers	0.307
	Creating relationships with customers	0.180
	Maintaining and promoting relationships with customers	0.099
Market innovation capabilities	*The ability to develop and to launch new products and services*	0.570
	Efficiency of the development process of new products and services	0.430
Human resources	*Levels of employee's job satisfaction*	0.478
	Levels of employee's organizational retention	0.522
Reputational assets	*Company or brand name and reputation*	0.540
	Credibility of the company from the customer's point of view	0.460

Source: own processing (The reported weights in the above table are the same for both fuzzy AHP and fuzzy ANP methods).

Table 5: The weights of the sub-criteria with respect to each criterion.

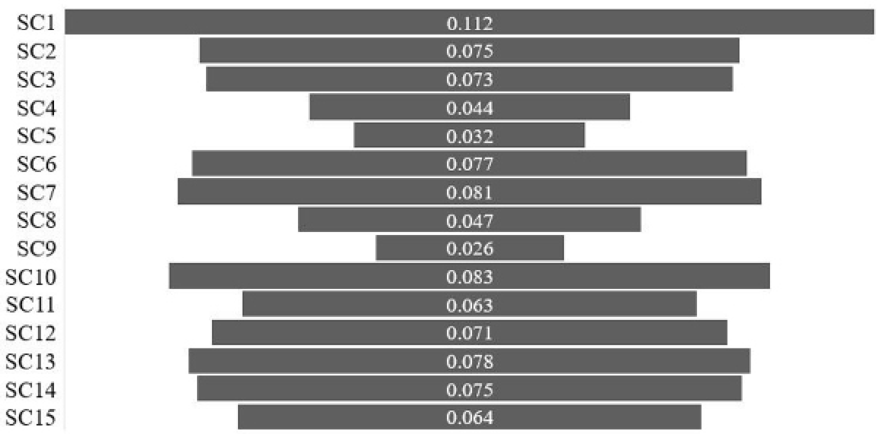
adequate employees is essential for any company to achieve organizational goals. Several types of employees with a wide range of skills are employed in agricultural and food processing firms. Retaining employees is one of the main challenges agro-food companies are faced with because the recruitment of quality labor is time-consuming and costly (Ratković, 2015).

The results of Table 5 show that, in the reputational assets, "company or brand name and reputation" (0.540) is more important than "credibility of the company from the customer's point of view" (0.460). Company/brand reputation stems from the consumer's opinion, so it reflects consumer's confidence in product quality which encourages the consumer to pay a price premium for labeled food and agricultural products. Brand name and reputation of the company are distinct from other elements of the marketing mix that represent quality, in that it indicates the cumulative effects of organizational marketing activities of a firm. If the brand of a company has a weak reputation, consumers will not trust that brand to buy (Lassoued and Hobbs, 2015). A study by Dantas et al. (2011) concluded that the brand name is one of the most important determinants of consumer food choice.

The relative importance of each of the sub-criteria, regardless of their criteria, is shown in Figure 3. As shown in Figure 3, based on the final weights, "financial condition", "the ability to develop and to launch new products and services" and "knowing the needs and requirements of consumers" are more important than other sub-criteria to choose the most appropriate business strategy.

After prioritizing the criteria and sub-criteria, the relative priority of business strategies is determined in the last step using fuzzy AHP and fuzzy ANP models; the results are shown in Table 6.

The results of Table 6 indicate that, based on both fuzzy AHP and fuzzy ANP methods, "differentiation strategy" has the highest priority, "focus strategy" has the next highest priority, and "overall cost leadership strategy" has the lowest priority. Following differentiation strategy means that managers focus on creating a value-added company by offering products and services which are distinct from those of rivals. Agro-food firms can differentiate themselves from competitors in a variety of ways, including product performance, delivery, product quality, taste, packaging, customer service, technical expertise and image. In fact, whatever is important for consumers is a potential basis for differentiation in food and agricultural products and services (Barnard et al., 2016).



Source: own processing (The symbols are defined in Figure 2)

Figure 3: The relative importance of the sub-criteria.

Business strategies	Fuzzy AHP	Fuzzy ANP	Ranking
Overall cost leadership strategy	0.289	0.290	3
Differentiation strategy	0.398	0.397	1
Focus strategy	0.313	0.313	2

Source: own processing.

Table 6: Business strategy rankings.

Conclusion

Choosing the most appropriate business strategy is complex and even risky because it determines the behaviors, policies, plans, and projects of a company in the market. Hence, if the business strategy is not selected appropriately, the company fails to achieve its goals, which imposes significant costs to the organization. Given the growing competition in domestic and international agro-food markets, a well-designed business strategy is necessary for agribusiness firms to gain competitive advantages. However, a business strategy cannot be easily chosen by using a series of decision rules because (1) several criteria and sub-criteria play a role in the decision-making process, (2) these criteria can be related to each other, and (3) given the fuzzy and vague nature of information, decision-makers' judgments cannot be expressed using real numbers. FMCDM methods are effective tools to solve these problems. Thus, fuzzy AHP and fuzzy ANP methods were used to prioritize business strategies for a major saffron firm in Iran. The results of both methods showed that the differentiation strategy is the most suitable strategy for gaining competitive advantages. Therefore, various types of saffron products are recommended to be introduced on the market in various forms, such as cake, pudding, crème caramel, jelly, batter mixture, beverage, cream, etc. Moreover, improving the performance of production and customer service can differentiate the saffron brands and products from other business competitors in the marketplace. The results of the study revealed that management and customer relationship capabilities have the greatest impact on choosing a differentiation strategy. Understanding customer's needs and building close relationships with them are essential to ensure the success of new saffron products. In this context, managerial capabilities play a central role, for these capabilities have a demonstrable effect on supporting market-based resources. In general, considering the results of the study, it is suggested that agro-food managers should pay particular attention to three critical factors for the successful implementation of the differentiation strategy: (1) financial condition, (2) introduction of new products and services, and (3) knowing the needs and requirements of consumers.

Corresponding authors
Mohammad Ghorbani, Professor of Agricultural Economics
Faculty of Agriculture, Ferdowsi University of Mashhad, Azadi Square, Mashhad, Razavi Khorasan Province, Islamic Republic of Iran.
Phone: +985138805784, E-mail: ghorbani@um.ac.ir

References

[1] Acar, C., Beskese, A. and Temur, G. T. (2018) "Sustainability analysis of different hydrogen production options using hesitant fuzzy AHP", *International Journal of Hydrogen Energy*, Vol. 43, No. 39, pp. 18059-18076. ISSN 0360-3199. DOI 10.1016/j.ijhydene.2018.08.024.

[2] Adetunji, O., Bischoff, J. and Willy, C. J. (2018) "Managing system obsolescence via multicriteria decision making", *Systems Engineering*, Vol. 21, No. 4, pp. 307-321. ISSN 1098-1241. DOI 10.1002/sys.21436.

[3] Altuntas, S. and Yilmaz, M. K. (2016) "Fuzzy DEMATEL Method to Evaluate the Dimensions of Marketing Resources: An Application in SMEs", *Journal of Business Economics and Management*, Vol. 17, No. 3, pp. 347-364. ISSN 1611-1699. DOI 10.3846/16111699.2015.1068220.

[4] Anwar, J. and Hasnu, S. (2016) "Business strategy and firm performance: a multi-industry analysis", *Journal of Strategy and Management*, Vol. 9, No. 3, pp. 361-382. ISSN 1755-425X. DOI 10.1108/JSMA-09-2015-0071.

[5] Ashtiani, M. and Abdollahi Azgomi, M. (2016) "Trust modeling based on a combination of fuzzy analytic hierarchy process and fuzzy VIKOR", *Soft Computing*, Vol. 20, No. 1, pp. 399-421. ISSN 1432-7643. DOI 10.1007/s00500-014-1516-1.

[6] Balci, G., Cetin, I. B. and Tanyeri, M. (2018) "Differentiation of container shipping services in Turkey", *Transport Policy*, Vol. 61, No. 26-35. ISSN 0967-070X. DOI 10.1016/j.tranpol.2017.10.004.

[7] Banchuen, P., Sadler, I. and Shee, H. (2017) "Supply chain collaboration aligns order-winning strategy with business outcomes", *IIMB Management Review*, Vol. 29, No. 2, pp. 109-121. ISSN 0970-3896. DOI 10.1016/j.iimb.2017.05.001.

[8] Barnard, F. L., Foltz, J. C. and Yeager, E. A. (2016) "*Agribusiness management*", 5th ed., London, Routledge, ISBN 1315709422. DOI 10.4324/9781315709420.

[9] Barney, J. B. and Hesterly, W. S. (2015) "*Strategic Management and Competitive Advantage Concepts and Cases*", 5th ed., Edinburgh Gate Harlow, Pearson Education Limited. ISBN 0133127400.

[10] Bendickson, J. S. and Chandler, T. D. (2019) "Operational performance: The mediator between human capital developmental programs and financial performance", *Journal of Business Research*, Vol. 94, No. 162-171. ISSN 0148-2963. DOI 10.1016/j.jbusres.2017.10.049.

[11] Brožová, H., Šup, L., Rydval, J., Sadok, M. and Bednar, P. (2016) "Information security management: ANP based approach for risk analysis and decision making", *Agris On-line Papers in Economics and Informatics*, Vol. 8, No. 1, pp. 13-23. ISSN 1804-1930. DOI 10.7160/aol.2016.080102.

[12] Chen, C. C., Yueh, H. P. and Liang, C. (2016) "Strategic management of agribusiness: Determinants and trends", *Journal of Entrepreneurship, Management and Innovation*, Vol. 12, No. 4, pp. 69-90. ISSN 2299-7326. DOI 10.7341/20161244.

[13] Dantas, M. I. S., Nakajima, V., Rosa, D. D., Andrade, F. O., Camila, C. and Martino, H. S. D. (2011) "Guava jam packaging determinant attributes in consumer buying decision", *Ciencia e Tecnologia de Alimentos*, Vol. 31, No. 3, pp. 567-570. ISSN 0101-2061. DOI 10.1590/S0101-20612011000300003.

[14] Deng, X. and Jiang, W. (2018) "An Evidential Axiomatic Design Approach for Decision Making Using the Evaluation of Belief Structure Satisfaction to Uncertain Target Values", *International Journal of Intelligent Systems*, Vol. 33, No. 1, pp. 15-32. ISSN 0884-8173. DOI 10.1002/int.21929.

[15] Dentoni, D., English, F. and Schwartz, D. (2014) "The impact of public R&D on marketing and supply chains on small farms' market sensing capability: Evidence from the Australian seafood industry", *International Food and Agribusiness Management Review*, Vol. 17, No. 1, pp. 37-58. ISSN 1559-2448. DOI 10.22004/ag.econ.163353.

[16] Gedela, R. K., Mohan, K. K. and Prasad, V. K. (2018) "Application of BOCR models in service oriented architecture (SOA): study on model validation through quantification for QoS considerations", *International Journal of Systems Assurance Engineering and Management*, Vol. 9, No. 6, pp. 1346-1354. ISSN 0975-6809. DOI 10.1007/s13198-018-0751-8.

[17] Goddard, A. and Simm, A. (2017) "Management accounting, performance measurement and strategy in English local authorities", *Public Money and Management*, Vol. 37, No. 4, pp. 261-268. ISSN 0954-0962. DOI 10.1080/09540962.2017.1295726.

[18] Gomes, C. F., Yasin, M. M., Lisboa, J. V. and Small, M. H. (2014) "Discerning competitive strategy through an assessment of competitive methods", *Service Industries Journal*, Vol. 34, No. 12, pp. 974-998. ISSN 0264-2069. DOI 10.1080/02642069.2014.915946.

[19] Graves, C. and Thomas, J. (2006) "Internationalization of Australian family businesses: A managerial capabilities perspective", *Family Business Review*, Vol. 19, No. 3, pp. 207-224. ISSN 0894-4865. DOI 10.1111/j.1741-6248.2006.00066.x.

[20] Grimmer, L. (2019) "*2 - Disrupting the giants: How independent grocers respond to the supermarket duopoly in Tasmania, Australia*", Woodhead Publishing. ISBN 978-0-08-102037-1.

[21] Gul, M., Ak, M. F. and Guneri, A. F. (2017) "Occupational health and safety risk assessment in hospitals: A case study using two-stage fuzzy multi-criteria approach", *Human and Ecological Risk Assessment: An International Journal*, Vol. 23, No. 2, pp. 187-202. ISSN 1080-7039. DOI 10.1080/10807039.2016.1234363.

[22] Guo, S. and Zhao, H. (2017) "Fuzzy best-worst multi-criteria decision-making method and its applications", *Knowledge-Based Systems*, Vol. 121, No. 23-31. ISSN 0950-7051. DOI 10.1016/j.knosys.2017.01.010.

[23] Haddad, M. and Sanders, D. (2018) "Selection of discrete multiple criteria decision making methods in the presence of risk and uncertainty", *Operations Research Perspectives*, Vol. 5, No. 357-370. ISSN 2214-7160. DOI 10.1016/j.orp.2018.10.003.

[24] Hill, C. W. L. and Jones, G. R. (2013) "*Strategic management: An integrated approach*", 10th ed., Mason, OH, South-Western, Cengage Learning, ISBN 1133485707.

[25] Hitt, M. A., Ireland, R. D. and Hoskisson, R. E. (2016) "*Strategic management: Concepts and cases: Competitiveness and globalization*", 12th ed., Boston, Cengage Learning, ISBN 1305856287.

[26] Hooley, G. J., Greenley, G. E., Cadogan, J. W. and Fahy, J. (2005) "The performance impact of marketing resources", *Journal of Business Research*, Vol. 58, No. 1 Spec. Iss., pp. 18-27. ISSN 01480-2963. DOI 10.1016/S0148-2963(03)00109-7.

[27] Hornická, A. and Brožová, H. (2013) "Continuity of demarcation process of the regions for concentrated state support", *Agris On-line Papers in Economics and Informatics,* Vol. 5, No. 3, pp. 3-12. ISSN 1804-1930. DOI 10.22004/ag.econ.157520.

[28] Hsieh, T.-Y., Lu, S.-T. and Tzeng, G.-H. (2004) "Fuzzy MCDM approach for planning and design tenders selection in public office buildings", *International Journal of Project Management*, Vol. 22, No. 7, pp. 573-584. ISSN 0263-7863. DOI 10.1016/j.ijproman.2004.01.002.

[29] Hsu, Y.-L., Lee, C.-H.,and Kreng, V. B. (2010) "The application of Fuzzy Delphi Method and Fuzzy AHP in lubricant regenerative technology selection", *Expert Systems with Applications*, Vol. 37, No. 1, pp. 419-425. ISSN 0957-4174. DOI 10.1016/j.eswa.2009.05.068.

[30] Khan, S. Z., Yang, Q. and Waheed, A. (2019) "Investment in intangible resources and capabilities spurs sustainable competitive advantage and firm performance", *Corporate Social Responsibility and Environmental Management,* Vol. 26, No. 2, pp. 285-295. ISSN 1535-3958. DOI 10.1002/csr.1678.

[31] Khilare, V., Tiknaik, A., Prakash, B., Ughade, B., Korhale, G., Nalage, D., Ahmed, N., Khedkar, C. and Khedkar, G. (2019) "Multiple tests on saffron find new adulterant materials and reveal that 1st grade saffron is rare in the market", *Food Chemistry*, Vol. 272, No. 635-642. ISSN 0308-8146. DOI 10.1016/j.foodchem.2018.08.089.

[32] Koulinas, G. K., Marhavilas, P. K., Demesouka, O. E., Vavatsikos, A. P. and Koulouriotis, D. E. (2019) "Risk analysis and assessment in the worksites using the fuzzy-analytical hierarchy process and a quantitative technique – A case study for the Greek construction sector", *Safety Science*, Vol. 112, No. 96-104. ISSN 0925-7535. DOI 10.1016/j.ssci.2018.10.017.

[33] Lassoued, R. and Hobbs, J. E. (2015) "Consumer confidence in credence attributes: The role of brand trust", *Food Policy*, Vol. 52, No. 99-107. ISSN 0306-9192. DOI 10.1016/j.foodpol.2014.12.003.

[34] Lee, A. H., Chen, W.-C. and Chang, C.-J. (2008) "A fuzzy AHP and BSC approach for evaluating performance of IT department in the manufacturing industry in Taiwan", *Expert Systems with Applications*, Vol. 34, No. 1, pp. 96-107. ISSN 0957-4174. DOI 10.1016/j.eswa.2006.08.022.

[35] Lin, C.-T., Lee, C. and Chen, W.-Y. (2009) "Using fuzzy analytic hierarchy process to evaluate service performance of a travel intermediary", *Service Industries Journal,* Vol. 29, No. 3, pp. 281-296. ISSN 0264-2069. DOI 10.1080/02642060701846762.

[36] Lin, C.-T., Lee, C. and Wu, C.-S. (2009) "Optimizing a marketing expert decision process for the private hotel", *Expert Systems with Applications*, Vol. 36, No. 3, pp. 5613-5619. ISSN 0957-4174. DOI 10.1016/j.eswa.2008.06.113.

[37] Lin, C. T. and Wu, C. S. (2008) "Selecting a marketing strategy for private hotels in Taiwan using the analytic hierarchy process", *Service Industries Journal*, Vol. 28, No. 8, pp. 1077-1091. ISSN 0264-2069. DOI 10.1080/02642060802187991.

[38] Lorenzo, J. R. F., Rubio, M. T. M. and Garcés, S. A. (2018) "The competitive advantage in business, capabilities and strategy. What general performance factors are found in the Spanish wine industry?", *Wine Economics and Policy,* Vol. 7, No. 2, pp. 94-108. ISSN 2213-3968. DOI 10.1016/j.wep.2018.04.001.

[39] Mahmoudkelaye, S., Azari, K. T., Pourvaziri, M. and Asadian, E. (2018) "Sustainable material selection for building enclosure through ANP method", *Case Studies in Construction Materials*, Vol. 9, No. ISSN 2214-5095. DOI 10.1016/j.cscm.2018.e00200.

[40] Mirzaei, O., Micheels, E. T. and Boecker, A. (2016) "Product and marketing innovation in farm-based businesses: the role of entrepreneurial orientation and market orientation", *International Food and Agribusiness Management Review*, Vol. 19, No. 2, pp. 99-130. ISSN 1559-2448. DOI 10.22004/ag.econ.234958.

[41] Mohaghar, A., Fathi, M. R., Zarchi, M. K. and Omidian, A. (2012) "A combined VIKOR-fuzzy AHP approach to marketing strategy selection", *Business Management and Strategy*, Vol. 3, No. 1, pp. 13-27. ISSN 2157-6068. DOI 10.5296/bms.v3i1.957.

[42] Mohammadi, H., Saghaian, S.,and Alizadeh, P. (2017) "Prioritization of expanded marketing mix in different stages of the product life cycle: the case of food industry", *Journal of Agricultural Science and Technology*, Vol. 19, No. 5, pp. 993-1003. ISSN 1680-7073.

[43] Montesinos-Valera, J., Aragonés-Beltrán, P. and Pastor-Ferrando, J.-P. (2017) "Selection of maintenance, renewal and improvement projects in rail lines using the analytic network process", *Structure and Infrastructure Engineering*, Vol. 13, No. 11, pp. 1476-1496. ISSN 1573-2479. DOI 10.1080/15732479.2017.1294189.

[44] Morris, W. and Mishra, N. (2014) "Online decision support system for finished lamb production", *International Journal of Services, Economics and Management*, Vol. 6, No. 3, pp. 209-221. ISSN 1753-0822. DOI 10.1504/IJSEM.2014.064296.

[45] Osakwe, C. N. and Anaza, N. A. (2018) "Understanding marketing resources and size in agro-based enterprises", *Marketing Intelligence and Planning*, Vol. 36, No. 2, pp. 230-244. ISSN 0263-4503. DOI 10.1108/MIP-07-2017-0149.

[46] Ratković, T. (2015) "HRM in foreign-owned agricultural and food processing companies in Serbia", *Economics of Agriculture*, Vol. 62, No. 2, pp. 353-367. ISSN 0352-3462. DOI 10.22004/ag.econ.206922.

[47] Rexhepi, G. and Srhoj, S. (2018) "Strategy as an ever evolving road to success of growing enterprises", *World Review of Entrepreneurship, Management and Sustainable Development*, Vol. 14, No. 3, pp. 333-347. ISSN 1746-0573. DOI 10.1504/WREMSD.2018.091689.

[48] Sadeghi, A. and Larimian, T. (2018) "Sustainable electricity generation mix for Iran: A fuzzy analytic network process approach", S*ustainable Energy Technologies and Assessments*, Vol. 28, No. 30-42. ISSN 2213-1388. DOI 10.1016/j.seta.2018.04.001.

[49] Siddiquie, R. Y., Khan, Z. A. and Siddiquee, A. N. (2017) "Prioritizing decision criteria of flexible manufacturing systems using fuzzy TOPSIS", *Journal of Manufacturing Technology Management*, Vol. 28, No. 7, pp. 913-927. ISSN 1741-038X. DOI 10.1108/JMTM-04-2017-0069.

[50] Sigalas, C. and Pekka Economou, V. (2013) "Revisiting the concept of competitive advantage: Problems and fallacies arising from its conceptualization", *Journal of Strategy and Management*, Vol. 6, No. 1, pp. 61-80. ISSN 1755-425X. DOI 10.1108/17554251311296567.

[51] Sok, P., Snell, L., Lee, W. J. T. and Sok, K. M. (2017) "Linking entrepreneurial orientation and small service firm performance through marketing resources and marketing capability: A moderated mediation model", *Journal of Service Theory and Practice*, Vol. 27, No. 1, pp. 231-249. ISSN 2055-6225. DOI 10.1108/JSTP-01-2016-0001.

[52] Ulubeyli, S., Kazaz, A. and Sahin, S. (2018) "Survival of construction SMEs in macroeconomic crises: Innovation-based competitive strategies", *Journal of Engineering, Design and Technology*, Vol. 16, No. 4, pp. 654-673. ISSN 1726-0531. DOI 10.1108/JEDT-03-2018-0057.

[53] Wu, C.-S., Lin, C.-T. and Lee, C. (2010a) "Competitive marketing strategies decision-making based on marketing resources and capabilities: Evidence from the hospitality industry in Taiwan", *Journal of Quality Assurance in Hospitality and Tourism*, Vol. 11, No. 4, pp. 219-238. ISSN 1528-008X. DOI 10.1080/1528008X.2010.504163.

[54] Wu, C.-S., Lin, C.-T., and Lee, C. (2010b) "Optimal marketing strategy: A decision-making with ANP and TOPSIS", *International Journal of Production Economics*, Vol. 127, No. 1, pp. 190-196. ISSN 0925-5273. DOI 10.1016/j.ijpe.2010.05.013.

[55] Wu, P., Gao, L. and Gu, T. (2015) "Business strategy, market competition and earnings management: Evidence from China", *Chinese Management Studies*, Vol. 9, No. 3, pp. 401-424. ISSN 1750-614X. DOI 10.1108/CMS-12-2014-0225.

[56] Yazdi, M. (2017) "Hybrid Probabilistic Risk Assessment Using Fuzzy FTA and Fuzzy AHP in a Process Industry", *Journal of Failure Analysis and Prevention*, Vol. 17, No. 4, pp. 756-764. ISSN 1547-7029. DOI 10.1007/s11668-017-0305-4.

[57] Yoshikuni, A. C. and Albertin, A. L. (2018) "Effects of strategic information systems on competitive strategy and performance", *International Journal of Productivity and Performance Management*, Vol. 67, No. 9, pp. 2018-2045. ISSN 1741-0401. DOI 10.1108/IJPPM-07-2017-0166.

Agris on-line Papers in Economics and Informatics

Volume XII Number 3, 2020

Productivity of Czech Milk Production in European Comparison

Zdeňka Žáková Kroupová, Pavlína Hálová, Lenka Rumánková

Department of Economics, Faculty of Economics and Management, Czech University of Life Sciences Prague, Czech Republic

Abstract

The aim of the paper is to evaluate the development and main characteristics of Czech milk production productivity and to compare Czech development with the situation in the European Union. From a methodological point of view, a parametric approach in the form of stochastic frontier analysis was applied, the input distance function was estimated, and total factor productivity was examined. The analysis used an unbalanced panel data set, which describes TF14-45 specialist milk production from 27 member states of the European Union in the period 2004–2016 collected in the FADN database.

The results showed that in the Czech Republic, the average value of technical efficiency was 94.01% during the analysed time period. Compared to EU member states, this figure was above the EU-13 average (93.71%). Czech milk production in the analysed period and the milk production of almost all other EU countries was characterized by increasing returns to scale. Examination of total factor productivity (TFP) showed that the scale effect and technical efficiency change effect can be considered the main components of TFP changes in Czech milk production. However, the scale effect was more significant in EU-15 countries than the Czech milk sector.

Keywords

Milk, Czech Republic, European Union, technical efficiency, random parameter model, total factor productivity.

Žáková Kroupová, Z., Hálová, P. and Rumánková, L. (2020) “Productivity of Czech Milk Production in European Comparison", *AGRIS on-line Papers in Economics and Informatics*, Vol. 12, No. 3, pp. 115-127. ISSN 1804-1930. DOI 10.7160/aol.2020.120310.

Introduction

According to the FADN database, milk production in the Czech Republic represents half of animal production and almost one fifth of the total agricultural production. Milk can be considered one of the most important agricultural commodities from a Czech and EU point of view. The importance of milk production is also obvious in the number of studies and scientific papers examining productivity in the dairy production process. These papers have analysed total factor productivity (TFP) development. Total factor productivity can also be broken down into partial components, including scale effect (SEC), technical efficiency change effect (TEC) and technical change effect (TC). The effect of these individual partial components has also been analysed and interpreted. Technical efficiency (TE) is especially interesting to researchers who analyse dairy production productivity.

The current studies focusing on milk productivity in EU countries have provided different results for some aspects of the dairy sector and the same results for other aspects of milk production. Skevas et al. (2018) showed that technical efficiency at German dairy producers has declined over time and the change in productivity has been driven by technological change. By contrast, Lansink et al. (2015) and Dakpo et al. (2019) concluded that the productivity of Dutch and French dairy farms has increased as a result of improvements in technical efficiency. Čechura et al. (2017) concluded that the EU-15 (Old Member States) have a higher TFP than the EU-13 (New Member States) group. The EU-15 is also characterised by above-average TFP growth (Wojciechowski, 2017). Madau et al. (2017) confirmed that decreasing productivity at European dairy producers was mainly a result of technological regression. Changes in technical efficiency have had a positive effect on TFP in most EU countries. Špička and Machek (2015) showed that regions with a positive change in TE have a higher milk yield and higher long-term loans than regions with adverse changes in TE. Investment has therefore slowed down the decline in TE.

The results of studies focusing on dairy milk productivity have predominantly concluded that the increase in milk productivity has been driven mainly by technological change, i.e. growth in productivity has been supported by modernization and innovation. From the perspective of individual countries, the prevailing effect of technological change can be seen in the Czech Republic, Estonia, Belgium, Hungary, Italy, Sweden, Finland and Ireland. The results of German and Dutch farms are questionable. According to Zhu and Lansink (2009), technical efficiency in these countries contributed more significantly to the change in TFP than technological change, whose impact has been emphasized in other studies.

Analysis of productivity also lets us evaluate the convergence of the EU-15 and EU-13 member states. However, some studies have shown quicker growth in TFP in the EU-15 Member States than its development in EU-13 Member States (Čechura et al., 2017; Irz and Jansik, 2015). This suggests the economic benefits of EU-13 integration have not yet been achieved. According to Čechura et al. (2017), only a few regions, mainly in the Czech Republic, Slovakia and Hungary, have achieved sufficient growth in total factor productivity to reach the more competitive EU-15 Member States. Most of these studies were based on the data set collected from the Farm Accountancy Data Network (FADN) database for the period 1990–2015.

From a methodological point of view, the parametric approach or stochastic frontier analysis (SFA) predominate in studies of TFP. In particular, the translogarithmic output-oriented distance function (ODF) has been estimated, even though the period of analysis covers milk production quota regulation, where the goal of maximizing profit through output growth cannot be fully met. Newman and Matthews (2007) and Skevas, Emvalomatis and Brümmer (2018) argued that output may not be fully regulated by quotas if manufacturers can trade or rent the quota freely. Some studies analysing the effect of milk quotas, for example, Barnes (2008), Kumbhakar et al. (2008), Žáková Kroupová (2016), however, respected the aforementioned limitations and applied the input-oriented distance function (IDF).

Some studies of total factor productivity have also focused on the relationship between the characteristics of production units, especially farm size, specialization, intensification and productivity and its components. The impact of agricultural policy has also been analysed in terms of deregulation of the quota system as well as subsidy policy.

The analyses were usually based on statistically testing the equality of mean values using the t-test or its non-parametric analogy in the case of violation of the normality presumption. Alternatively, regression analysis may also be used, including logistic regression (e.g., Bokusheva and Čechura (2017), study of the cereal production sector). Analysis of the difference in growth in total factor productivity in terms of specialization is particularly relevant in EU-13 member states, where mixed farms are also significantly involved in milk production. Žáková Kroupová (2016) found that TFP growth at specialized Czech farms was greater than at mixed dairy farms.

Research into this intensification has provided ambiguous results for individual EU countries.

Keizer and Emvalomatis (2014) found that TFP components developed consistently at intensive and extensive Dutch farms. By contrast, Alvarez and del Corral (2010) concluded that the intensive technology of Spanish farms was more productive and efficient than extensive technology. In accordance with the Common Agricultural Policy, these authors emphasized the environmental friendliness of intensive dairy production technologies.

Furthermore, the studies have predominantly found that large farms involved in dairy production show quicker growth in total factor productivity than small farms, suggesting that small farms may lag behind large innovative activity. Lansink et al. (2015), also mentioned the role of agricultural policy because the innovative activity of dairy farms is driven significantly by external incentives.

Some authors have also discussed investment subsidies. For example, Špička and Machek (2015) and Žáková Kroupová (2016) found no significant relationship between the growth of TFP and investment subsidies. This suggests that subsidy instruments do not probably achieve the required effectiveness. Operating subsidies, usually analysed according to subsidy-to-income ratio, are commonly considered possessing effects that adversely affect productivity, mainly due to their negative effect on technical efficiency. See, for example, Luik et al. (2011).

Most of the studies examining the total factor productivity of dairy producers were based on a period when milk production was regulated with production quotas. Frick and Sauer (2016) studied the anticipated effects of abandoning milk quotas based on the hypothesis that production quotas adversely affect efficiency and productivity, this effect being reduced by the marketability of milk quotas. The study provided empirical evidence that releasing the quota system was linked to reallocating resources towards more productive farms. However, price volatility also affected the reallocation of resources. This rather demotivated less productive agricultural holdings to invest in activity more than productive entities in developing production. It therefore contributed to reallocating resources towards more productive entities. Čechura et al. (2017) can be also mentioned in this context. The authors noted that after milk production quotas were abolished, the shift from EU-13 member states and southern European countries to northern countries was expected, i.e. from less productive to more productive regions.

The aim in this paper is to evaluate the development and main characteristics of Czech milk production productivity and to compare Czech development with the European Union's situation. The paper addresses the following research questions: What is Czech milk production's competitive position in terms of productivity in the EU? What is the main source of milk production productivity growth in the Czech Republic? Is the change in TFP driven by the same component in the Czech Republic and other EU countries? How does Czech milk production differ from the most competitive producers in the EU?

The rest of the paper is organized as follows. The introduction describes the data and methods used. The next section presents the results of the analysis. First, IDF estimates are commented on and the technical efficiency of milk production is discussed. Second, the development of total factor productivity and its components is analysed. In the final section, the characteristics of milk production with the highest TFP change is analysed.

Materials and methods

Total factor productivity change is analysed according to the approaches of Emvalomatis (2012) and Bauer (1990). Allocation efficiency, which refers to price and marginal cost equality, and a market structure of near perfect competition (Grau and Hockmann, 2016) are assumed.

Total factor productivity change according to the Divisia index is quantified as the sum of three components: scale effect (SEC), technical efficiency change effect (TEC) and technical change effect (TC) *TFP = SEC + TC + TEC*.

These components are derived from the estimate of an input-oriented distance function (Coelli et al., 2003). Input-orientation is preferred over output because of the prevailing existence of milk quotas in the analysed period 2004–2016. Because milk quotas represent a strong restriction on the maximum quantity of milk production, it can be assumed that agricultural producers focus primarily on reducing input to produce almost fixed output (see Kumbhakar et al., 2008). This means that the goal of profit maximization can be achieved by minimizing the cost of producing a fixed (quota) output. According to Skevas et al. (2018), output can be assumed exogenous under this optimization condition.

Bakusc et al. (2012), Latruffe et al. (2011) and Kumbhakar and Tsionas (2008) have also estimated the input-oriented distance function for specialized milk production. Morrison Paul et al. (2004) also generally recommended the input-oriented function for analysing agricultural production, arguing that agricultural producers have more control over input than output over a short period of time.

Using the homogeneity property of IDF, the following random parameter stochastic translog IDF with *M* outputs (*Y*), *J* inputs (*X*) and time (*T*) can be estimated:

$$-\ln X_{1it} = A_i + v_{it} - u_{it}, \tag{1}$$

where
$$\begin{aligned} A_i = {} & \alpha_{0,i} + \textstyle\sum_{m=1}^{2} \beta_{m,i} \ln Y_{m,it} + \\ & + \tfrac{1}{2} \textstyle\sum_{m=1}^{2} \sum_{n=1}^{2} \beta_{mn} \ln Y_{m,it} \ln Y_{ni,t} + \\ & + \textstyle\sum_{m=1}^{2} \sum_{j=2}^{5} \delta_{mj} \ln Y_{m,it} \ln \tilde{X}_{j,it} + \\ & + \textstyle\sum_{j=2}^{5} \beta_{j,i} \ln \tilde{X}_{j,it} + \\ & + \tfrac{1}{2} \textstyle\sum_{j=2}^{5} \sum_{k=2}^{5} \beta_{jk} \ln \tilde{X}_{j,it} \ln \tilde{X}_{k,it} + \\ & + \alpha_{t,i} T + \tfrac{1}{2} \alpha_{tt} T^2 + \\ & + \textstyle\sum_{m=1}^{2} \alpha_{mt} \ln Y_{m,it} T + \\ & + \textstyle\sum_{j=2}^{5} \alpha_{jt} \ln \tilde{X}_{j,it} T, \end{aligned}$$

where $v_{it} \sim N(0, \sigma_v^2)$ is the stochastic error term, $u_{it} = |U_{it}|, U_{it} \sim N(\mu_i, \sigma_u^2)$ is the time-varying inefficiency, and $\ln \tilde{X}_{j,it} = \ln X_{j,it} - \ln X_{1,it}$, α, β, δ are the estimated parameters. The symmetry restrictions imply that $\beta_{jk} = \beta_{kj}$ and $\beta_{mn} = \beta_{nm}$.

Furthermore, normalization ensures the exogeneity of input (Sipiläinen et al., 2014) and consistency of estimation (Kumbhakar, 2011). All variables were also normalised as logarithms at their sample mean, which made it possible to interpret the estimated first-order parameters as elasticities at the sample mean. The random parameter model was fitted according to the maximum simulated likelihood with a Halton sequence assuming a normal distribution of random parameters in the SW NLOGIT 5.0.

The technical change component can be computed from IDF as in Equation (2):

$$TC = -\frac{\partial \ln D_{it}^{I}}{\partial T} = -(\alpha_t + \alpha_{tt}T + \sum_{m=1}^{M} \alpha_{mt} \ln Y_{mi,t} + \sum_{j=2}^{J} \alpha_{jt} \ln \tilde{X}_{ji,t}). \quad (2)$$

According to Lansink et al. (2000), technical change can be further decomposed into a Hicks neutral: $TC_n = -(\alpha_t + \alpha_{tt}T)$ and factor-biased technical change $TC_b = -(\sum_{m=1}^{M} \alpha_{mt} \ln Y_{mi,t} + \sum_{j=2}^{J} \alpha_{jt} \ln \tilde{X}_{ji,t})$, which indicates a change in factor productivity allowing certain inputs to be saved.

The scale effect can be quantified as:

$$SEC = \dot{Y}_C(1 + \varepsilon^{CQ}), \quad (3)$$

where $\dot{Y}_C = \sum_m \frac{\frac{\partial \ln D_{it}^{I}}{\partial \ln Y_{mi,t}}}{\sum_m \frac{\partial \ln D_{it}^{I}}{\partial \ln Y_{mi,t}}} \frac{(Y_{m,t} - Y_{m,t-1})}{0.5(Y_{m.t} + Y_{m,t-1})}$

is the weighted aggregate rate of change in output

and $\varepsilon^{CQ} = \sum_m \frac{\partial \ln D^I(y,x,t)}{\partial \ln y_m}$ is the sum of elasticities of IDF with respect to output (see Kumbhakar and Lozano-Vivas, 2005).

As noted by Sipiläinen et al. (2014), the use of averages for the consecutive periods *t – 1* and *t* ensures that the analysis is consistent over time for ‘static’ variables.

Finally, the technical efficiency change was computed using Equation (4):

$$TEC_{i,t} = \frac{\widehat{TE}_{i,t} - \widehat{TE}_{i,t-1}}{0.5(\widehat{TE}_{i,t} + \widehat{TE}_{i,t-1})}, \quad (4)$$

where the technical efficiency was estimated according to Jondrow et al. (1982).

Calculation of the components was conducted using software R, version 3.5.0. The quantified components were further evaluated in relation to the specific characteristics of milk production in the FADN regions. Specialization, milk yield, farm size, proportion of paid labour, proportion of rented agricultural land, proportion of feed from own production, labour intensity, indebtedness, localisation in Less Favoured Areas (LFA) and subsidies were especially studied. The differences in these characteristics between Czech milk production and milk production in the most competitive countries were tested using the Kruskal–Wallis test.

The analysis used unbalanced panel data set of TF14-45 specialist milk drawn from the FADN database. The data covers the period 2004–2016 and 27 European Union member states. No data was available for specialized milk production in Cyprus and Greece. The data set consists of 1,343 observations of FADN regions. Although regional data represent the lowest level of aggregation freely available in the FADN database, it introduced several limitations into the analysis, including the unfeasibility of evaluating the variability of input and output in the production process at the farm level (i.e., in FADN regions). The estimated frontier according to regional data may differ from the true frontier estimated from farm data. The low quantity of regional data also made it unfeasible to model the meta-frontier. Concerning the use of FADN data, the sample might not necessarily be representative of the dairy farm sector in each country since the smallest farms are not covered by FADN data. Furthermore, the FADN sample may change every year, with some farms entering the sample and others leaving (the reasons may be termination of activities, shortcomings in accounting, etc.). However, it can be assumed that the FADN sector data show the same tendencies as specific sectors in the FADN region. The limitations related to applying aggregated FADN data are also mentioned by Madau et al. (2017), who analysed the technical efficiency and productivity of 22 European countries in 2004–2012 based on FADN country data.

In order to estimate the IDF in this study, the following outputs and inputs were used: milk production (Y1) in kilos (SE125N), other production (Y2) in EUR, which is determined from the sum of crop production (SE135), other animal production (SE206 minus the production of milk in EUR (SE216)) and other production (SE256), the cost of feed for grazing livestock (X1) in EUR (SE310), labour (X2) measured in working hours (SE011), the total utilized agricultural area (X3) in hectares (SE025), capital (X4) in EUR measured as depreciation (SE360), and the costs of other materials (X5) in EUR (total intermediate consumption (SE275) minus feed for grazing livestock (SE310)). Outputs and inputs (except for milk production, labour and land), are deflated by price indices (individual output and input indices (2010 = 100) - source is the EUROSTAT database).

Results and discussion

Table 1 shows the estimated parameters of the input-distance function. Almost all of the parameters are significant, even at the 1% significance level. The estimated model also satisfied the properties of an input distance function, namely symmetry, monotonicity, linear homogeneity and concavity in inputs and quasi-concavity in outputs. Since all variables were normalised as logarithms at their sample mean, the first-order parameters can be interpreted as the elasticity of the IDF with respect to output on the sample mean and as the proportion of input in the total input. Table 1 shows that the input proportion of capital (X4) was the lowest (0.06), the input proportion of labour (X2) was the highest (0.37) i.e. the proportion of capital in the total input was only 6%, though the proportion of labour was about 37%. This reflects capital market imperfections, especially at the beginning of the analysed time period and in the group of countries accessing to the EU in/after 2004 (EU-13). However, the analysed time period saw an increase in the proportion of capital in total input and a reduction in the proportion of land (X3) and other materials (X5, except for feeds). This may indicate the modernization of production towards more material-efficient and less land-bound technologies. Sipiläinen et al. (2014) also found a high proportion of labour and low proportion of capital in total dairy farm inputs according to the data from Finnish and Norwegian milk production.

The elasticity of milk (Y1) that corresponds to the negative of the cost elasticity of the particular output (see Irz and Thirtle, 2004), was about (-0.51). The negative inversion of the sum of partial elasticities of outputs, which corresponds to economies of scale, was statistically significantly different from the one at the 1% significance level according to a t-test (t-value = 65.240) and averages 1.36. That is, milk production can be characterized by the existence of increasing returns to scale. In only 10% of cases (especially Romania, Hungary, Finland, Estonia, Spain), the estimated value of economies of scale was less than one, which corresponds to decreasing returns to scale. Čechura et al. (2017) also identified prevailing increasing returns to scale in European dairy farms

Random parameter means				Scale parameters			
Variable	Coeff.	SE	P [|z|>Z*]	Variable	Coeff.	SE	P [|z|>Z*]
Const.	0.0291***	0.0072	0.0001	Const.	0.1632***	0.0029	0.0000
T	0.0073***	0.0005	0.0000	T	0.0063***	0.0005	0.0000
Y1	-0.5102***	0.0076	0.0000	Y1	0.1655***	0.0058	0.0000
Y2	-0.2266***	0.0051	0.0000	Y2	0.1112***	0.0052	0.0000
X2	0.3727***	0.0063	0.0000	X2	0.1930***	0.0066	0.0000
X3	0.1765***	0.0063	0.0000	X3	0.1347***	0.0067	0.0000
X4	0.0606***	0.0054	0.0000	X4	0.0805***	0.0044	0.0000
X5	0.1449***	0.0080	0.0000	X5	0.0293***	0.0049	0.0000
Non-random parameters							
Variable	Coeff.	SE	P [|z|>Z*]	Variable	Coeff.	SE	P [|z|>Z*]
TT	-0.0004	0.0004	0.2576	X23	-0.0827***	0.0121	0.0000
Y1T	0.0049***	0.0015	0.0007	X24	0.0531***	0.0115	0.0000
Y2T	0.0074***	0.0010	0.0000	X25	-0.0074	0.0160	0.6423
Y11	-0.2109***	0.0210	0.0000	X34	-0.0017	0.0091	0.8520
Y22	-0.0610***	0.0099	0.0000	X35	0.0498***	0.0130	0.0001
Y12	0.1433***	0.0198	0.0000	X45	0.0100	0.0115	0.3865
X2T	0.0108***	0.0013	0.0000	Y1X2	-0.0657***	0.0192	0.0006
X3T	-0.0086***	0.0012	0.0000	Y1X3	0.0698***	0.0142	0.0000
X4T	0.0116***	0.0011	0.0000	Y1X4	-0.0965***	0.0122	0.0000
X5T	-0.0068***	0.0018	0.0001	Y1X5	0.0392**	0.0154	0.0109
X22	0.1471***	0.0188	0.0000	Y2X2	0.0779***	0.0114	0.0000
X33	-0.0068	0.0100	0.4941	Y2X3	-0.0437***	0.0110	0.0001
X44	0.0738***	0.0094	0.0000	Y2X4	0.0974***	0.0090	0.0000
X55	0.0016	0.0235	0.9461	Y2X5	-0.0339***	0.0108	0.0016
Sigma	0.0839***	0.0035	0.0000	Log-likelihood			1399.5799
Lambda	1.1676***	0.1793	0.0000				

Note: ***, **, * denote significance at the 1%, 5%, and 10% levels, respectively
Source: Author's calculations

Table 1: Estimated parameters.

in the farm-level data from 2004–2011.

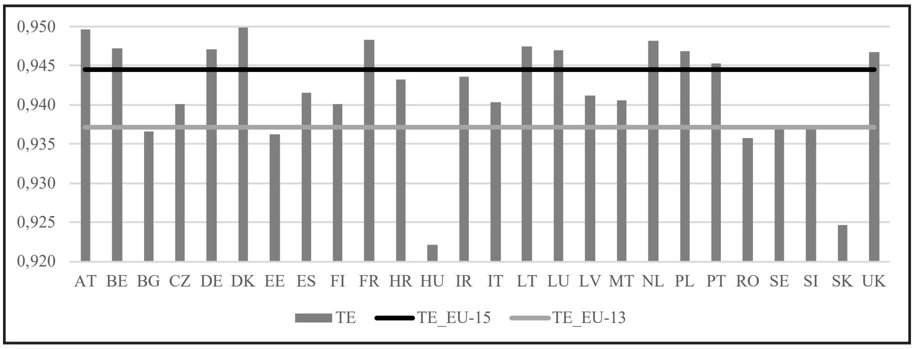
According to Rasmussen (2011), assuming homothetic production technology, economies of scale can be interpreted in the same manner as economies of size. The estimated value of economies of scale therefore suggests that the size of the dairy specialized farms of FADN regions is below the technically optimal level. Moving to a technically optimal size would bring cost savings to EU dairy producers. It can be stated that Czech milk production is close to the technically optimal size. The value of returns to scale was 1.03 for the Czech Republic.

The parameter lambda is also significant at the 1% significance level and greater than one. The variation in technical inefficiency is more pronounced than the variation in the stochastic error. This indicates that most of the deviation from the frontier of the input requirement set was due to technical inefficiencies rather than random shocks. The average technical efficiency of specialized milk production was 94.28%, with a standard deviation of 2.64%. The minimum value of technical efficiency was 79.23%. However, for only 5% of the observations, a technical efficiency less than 89.24% was found, and only 25% of the observations had a technical efficiency score of less than 93.56%. Conversely, 25% of the most successful regions show the technical efficiency of dairy producers as greater than 95.80%, with a maximum of 98.74%. This suggests that European milk production highly exploited its production possibilities in 2004–2016.

In the Czech Republic, the average value of technical efficiency was 94.01% over the analysed time period. Compared to other member states, this figure is above the EU-13 average (93.71%) shown in Figure 1. Poland is the only country that achieves a higher value in the Visegrad group.

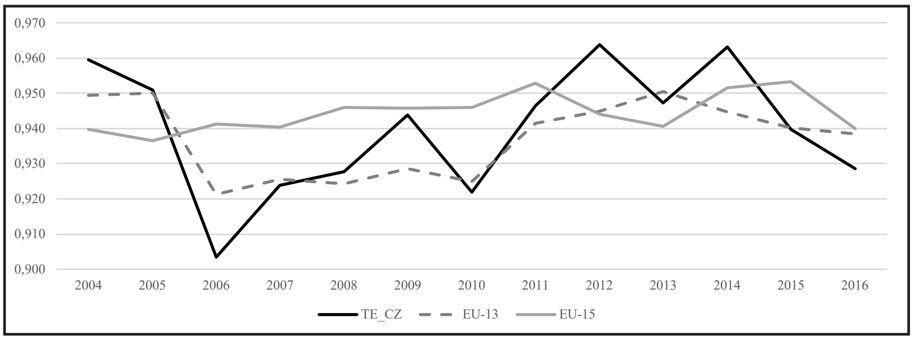
The average technical efficiency of the EU-13 states was below the EU-15 average (94.45%). The difference in the average technical efficiency scores of the EU-13 and EU-15 group was statistically significant at $\alpha = 0.05$ (t-value = −3.764).

The development of technical efficiency is shown in the second graph (Figure 2). The graph shows that development in the EU-13 was characterized by more significant fluctuations than in the EU-15. The comparison of the average technical efficiency between 2016 and 2004 shows a slight decrease in the average technical efficiency of EU milk production, driven mainly by the decrease in technical efficiency in the EU-13 countries. It is clear that changes in technical efficiency do not have the same direction in the old and new



Source: Author's calculations

Figure 1: Technical efficiency (TE): country-specific average values.



Source: Author's calculations

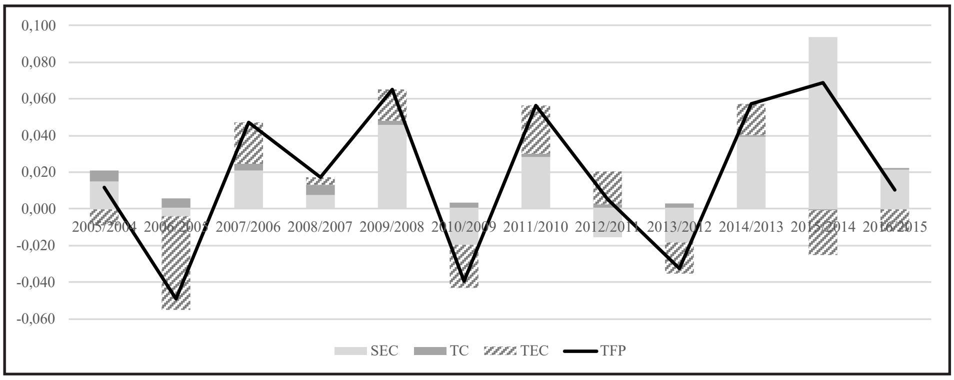
Figure 2: Technical efficiency development.

member states. Most of the changes suggest that development in the EU-15 countries prefigures development in the EU-13.

In the Czech Republic, 2006, 2010 and 2015 saw significant declines. The 2006 decline can be considered the result of reform in the Common Agricultural Policy and dairy market regulation (especially the decline of intervention prices). The decline in 2010 was a result of economic crisis. The dairy crisis in 2015 was caused by a surplus of milk on world markets, not only because milk quotas were abolished but also because demand from China reduced and because of a Russian embargo. Only Slovakia (standard deviation in technical efficiency: 5.13%) from the Visegrad Group countries showed significantly higher fluctuations than the Czech Republic (standard deviation: 1.81%), but only in the first half of the analysed period. Since 2009, unlike Czech milk production, a steady increase in the technical efficiency of Slovak dairy farms can be observed. This indicates the high sensitivity of Czech milk production to crisis periods and the limited ability of Czech dairy farms to manage shocks affecting short-term production.

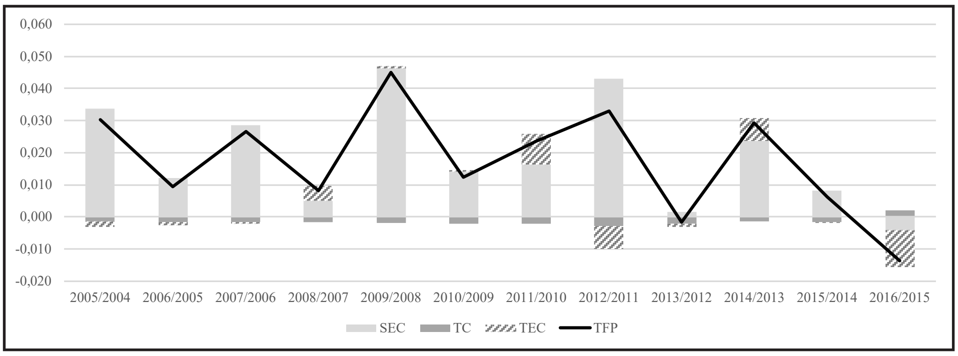
The development of technical efficiency influences the total factor productivity change. The third graph (Figure 3) charts the total factor productivity index of Czech milk production. It is clear that in the crisis periods described above, the negative effect of technical efficiency (falling-behind of dairy producers) caused a fall in total factor productivity, except for 2015, when a positive change in TFP was maintained by a strong scale effect (optimizing the scale of operations). The scale effect and the technical efficiency change effect can be considered the main components of TFP changes in Czech milk production in 2004–2016. The technical change component (implementing new technologies) influenced TFP particularly at the beginning of the analysed period. This demonstrates the positive impact of investments supported by pre-accession programmes as well as the absence of innovation at the end of the analysed period.

A comparison with developments on world markets suggests that the effect of technical efficiency responds more flexibly to the deterioration of production, sales or economic conditions than the scale effect, where a time lag is apparent.



Note: TFP is the total factor productivity change, SEC is the scale effect, TC is the technical change, TEC is the effect of technical efficiency change.
Source: Author's calculations

Figure 3: Total factor productivity index: Czech milk production.



Note: TFP is the total factor productivity change, SEC is the scale effect, TC is the technical change, TEC is the effect of technical efficiency change.
Source: Author's calculations

Figure 4: Total factor productivity index: EU milk production.

The fourth graph (Figure 4.), which describes the development of TFP and its components in EU milk production, indicates that the total factor productivity of European milk production was influenced more by the scale effect than Czech production. The scale effect was therefore more pronounced in the EU-15 member states. However, even in the EU-13 group, it was mainly structural changes that contributed to productivity growth (Table 2).

Table 2 shows that technological decline even decelerated the growth in TFP in the EU-13 countries in 2004–2016. While the EU-15 has seen technological progress since 2014, the impact of technological decline strengthened in EU-13 countries at the end of the analysis period, i.e. dairy producers in EU-13 countries lagged behind in innovative activities.

Overall, the TFP components resulted in a 1.8% year-on-year increase in total factor productivity in the Czech Republic (Table 2). Czech milk production was characterized by a 1.8% year-on-year increase in the output-input ratio on average in the period 2004–2016. The TFP of dairy producers grew faster in the Czech Republic than in the EU-13 group or other members of the Visegrad group (Table 3).

Table 2 also shows that specialized milk production in EU-15 was characterized by higher TFP growth than milk production in EU-13. Čechura et al. (2017) presented similar findings based on farm data. These authors concluded that the positive economic effects expected from economic integration and the convergence of EU regions in terms of dairy productivity had not yet been achieved.

The difference in TFP change between the EU-15 and EU-13 groups is statistically significant at $\alpha = 0.05$ (t-value = 3.182). Quicker growth in the EU-15 was driven by a more significant scale effect and the contribution of technical change and technical efficiency. In the EU-13, both technical change and technical efficiency change (falling-behind) on average weakened the positive scale effect.

In accordance with the previous results, the dominant influence of the scale effect was also apparent from the perspective of member states (Table 3). The exceptions were Croatia, Finland, Italy, Malta, Portugal, Austria and Romania, where total factor productivity development was primarily determined by the effect of technical change. However, in most of these countries, the impact of technical change on TFP change was negative. Technical efficiency compared to other components had a minor effect on TFP change. The stronger impact of technical efficiency change compared to technical change was apparent only in Luxembourg, Hungary, Germany and Slovakia. On average, the effect of technical efficiency had almost the same absolute value as the effect of technical change in the Czech Republic and Lithuania.

The effect of technical change was mainly driven by neutral technical change, with the exception of Denmark, France, Italy, Luxembourg, Malta, Portugal, Austria, Romania, Slovakia, Slovenia and Spain. In these countries, factor-biased technical change was more significant on average in the analysed period. Almost the same effect of these components of technical change can be seen in the UK and Hungary, where these effects eliminated each other. A negative factor-biased technical change was found in the Czech Republic, and this effect reduced the positive effect of neutral technical change.

	CZ		EU-15		EU-13	
	Mean	St. dev.	Mean	St. dev.	Mean	St. dev.
TFP	0.0182	0.0418	0.0213	0.0632	0.0030	0.0832
SEC	0.0179	0.0324	0.0210	0.0631	0.0110	0.0721
TC	0.0030	0.0021	0.0001	0.0103	-0.0074	0.0089
TEC	-0.0027	0.0240	0.0001	0.0249	-0.0005	0.0405

Note: TFP is the total factor productivity change, SEC is the scale effect, TC is the technical change, TEC is the effect of technical efficiency change.
Source: Author's calculations.

Table 2: Comparison of TFP change in the Czech Republic, EU-15 and EU-13.

Country	TFP	SEC	TC	TEC
Austria	0.0010	0.0139	-0.0110	-0.0019
Belgium	0.0316	0.0382	-0.0044	-0.0022
Bulgaria	-0.0063	0.0056	-0.0080	-0.0039
Croatia	-0.0244	0.0063	-0.0221	-0.0086
Czech Republic	**0.0182**	**0.0179**	**0.0030**	**-0.0027**
Denmark	0.0480	0.0473	0.0021	-0.0015
Estonia	0.0345	0.0167	0.0109	0.0068
Finland	0.0240	0.0091	0.0114	0.0035
France	0.0276	0.0257	0.0021	-0.0002
Germany	0.0242	0.0253	0.0005	-0.0015
Hungary	0.0116	0.0119	0.0006	-0.0010
Ireland	0.0351	0.0368	-0.0045	0.0028
Italy	0.0010	0.0134	-0.0120	-0.0005
Latvia	0.0182	0.0277	-0.0096	0.0000
Lithuania	0.0094	0.0140	-0.0026	-0.0020
Luxembourg	0.0405	0.0404	-0.0011	0.0012
Malta	-0.0162	0.0081	-0.0258	0.0015
Netherlands	0.0229	0.0298	-0.0042	-0.0026
Poland	0.0064	0.0156	-0.0067	-0.0025
Portugal	0.0200	0.0020	0.0100	0.0079
Romania	-0.0087	0.0006	-0.0116	0.0022
Slovakia	0.0152	0.0142	-0.0006	0.0016
Slovenia	0.0110	0.0272	-0.0179	0.0018
Spain	0.0327	0.0268	0.0061	-0.0002
Sweden	0.0295	0.0116	0.0108	0.0071
United Kingdom	0.0220	0.0149	0.0074	-0.0003

Note: TFP is the total factor productivity change, SEC is the scale effect, TC is the technical change, TEC is the effect of technical efficiency change.
Source: Author's calculations.

Table 3: Total factor productivity change in EU member states.

In detail, it can be concluded that the main source of dairy productivity growth in most regions of the EU was the increasing size of specialized dairy farms, accompanied by structural changes associated with the concentration of production in the most competitive regions. The changes were amplified by the release and subsequent abolition of milk quotas, which prevented the increase in output and reallocation of inputs among producers and thereby produced welfare losses (see Kumbhakar et al., 2008 and Frick and Sauer, 2016).

Productivity gains due to the elimination of underutilized resources were apparent in Estonia, Finland, Ireland, Luxembourg, Malta, Portugal, Romania, Slovakia, Slovenia and Sweden. The technical change, indicating innovations that save input involved in producing a given amount of output, was evident in the Czech Republic, Denmark, Estonia, Finland, France, Hungary, Germany, Portugal, Spain, Sweden and the United Kingdom. The source of the technological shift probably was the improvement of the genetic potential of farmed dairy cows or technological innovations (Lansink et al., 2015).

The positive effect of the technical change was mainly due to neutral technical changes in the Czech Republic, Estonia, Finland and Germany. In Denmark, France, Portugal and Spain, factor-biased technical change, i.e. the influence of technology on the change in the use of production factors, was more significant. However, in the Czech Republic, Denmark, France, Hungary, Germany, Spain and the United Kingdom, a similar situation to the work of Emvalomatis (2012) in cases of German dairy farms in 1995–2004 can be observed: the positive technical change was accompanied by a negative effect from technical efficiency. The contradictory tendencies of technical efficiency and technological progress have also been highlighted by Dakpo et al. (2019), who analysed French dairy farms. This may indicate that dairy producers in these countries are facing differences in time when new technology is introduced and the knowledge on how to fully make use of it is acquired. Practical examples are shortcomings in the setup of milking equipment, which may cause decreased milk yields.

As in the Czech Republic, TFP growth in most EU countries over the analysed period can be seen. TFP decline only occurred in Croatia, Italy, Malta and Romania (Table 3). Furthermore, countries with significantly above-average productivity changes (i.e., 25% of country-aggregated observations with the highest average productivity change), namely Denmark, Luxembourg, Ireland, Estonia, Belgium, Spain and Sweden, were found in the analysed sample. These countries have strengthened their competitive position. By contrast, Austria, Poland, Lithuania and Slovenia showed significantly below-average growth, indicating a weakening of the competitiveness of this group of countries.

Milk production with the highest productivity growth in 2004–2016 was characterized by a high level of specialization. The proportion of milk production in total output was 75% on average. This was significantly higher than in Czech milk production (Table 4). Compared to the Czech Republic, the group of countries with significantly above-average TFP change is represented by dairy specialized producers with high milk yields in the FADN database.

	Czech milk production	Above-average TFP group	Sig.
Proportion of milk on total output [%]	49.70	75.01	***
Milk yield [kilos/LU]	6403.71	7262.10	**
Agriculture area [hectares]	265.29	64.57	***
Dairy herd [LU]	93.10	65.21	***
Economic size [ESU]	304.82	204.76	***
Proportion of paid labour [%]	82.98	15.58	***
Proportion of rented land [%]	85.42	55.60	***
Proportion of own feed [%]	60.26	24.08	***
Long-term indebtedness	19.49	16.94	
Labour intensity [hours/LU]	228.20	72.96	***
Proportion of total subsidies in production [%]	30.35	16.66	***
Proportion of investment subsidies in total subsidies [%]	4.21	5.99	
Decoupled payments [Eur/hectares]	151.50	322.35	***
Livestock subsidies [Eur/LU]	42.94	53.02	
LFA subsidies [Eur/hectares]	57.00	40.24	***
Farm net value added [Eur/AWU]	6.82	17.19	***

Note: The above-average group represents Denmark, Luxembourg, Ireland, Estonia, Belgium, Spain and Sweden
***, ** denote significance at the 1% and 5% levels in the Kruskal-Wallis test, respectively.
Source: Author's calculations

Table 4: Characteristics of milk production in the Czech Republic and in the group with above-average TFP growth.

Czech dairy producers made significantly greater use of agricultural area than the group of most competitive countries. This was associated with a higher proportion of the Czech Republic's own feed production (60%). Due to the higher localization of production in LFA in the Czech Republic, a lower quality of home-grown feed can be considered. Other differences can be seen in the dependence on external inputs, namely labour and land. The high dependence of Czech milk production on external labour and land may endanger the economic situation of producers. Rising labour and land prices may lead to a loss of profitability. External workers are more susceptible to fluctuations than family workers. The fluctuation of labour may also involve the risk of losing specialized workers with appropriate skills. Table 4 suggests that there were significantly more working hours per livestock unit in the Czech Republic, which also indicates a lack of mechanization and automation in milk production, which may relate to the aforementioned decline in impact of technological change at the end of the analysis period. Investment subsidies also probably did not contribute sufficiently to companies' innovation activities leading to technological progress, as the proportion in total subsidies between the Czech Republic and the above-average group was similar. Žáková Kroupová (2016) also did not find a significant relationship between investment subsidies and technical change according to the farm data from 2004–2011 of Czech dairy producers.

The economic situation of Czech, and in general, European dairy producers has been significantly influenced by subsidy payments. Table 4 shows that for each euro of production created, Czech dairy producers received 0.3 EUR of subsidies. This is significantly more than the most competitive group, where the value of decoupled payments per hectare and the value of livestock subsidies per livestock unit was higher than in the Czech Republic. Lower agricultural producer milk prices and high average production costs, together with lower subsidy payments, resulted in low profitability in specialized Czech milk production.

Conclusion

The aim in this paper was to evaluate the development of Czech milk production productivity and its main characteristics and to compare with EU milk production productivity. The analysis examined the input-oriented distance function and the total factor productivity and its partial components, i.e. scale effect, technical efficiency change effect and technical change effect, in the Czech, EU-15 and EU-13 milk sectors using FADN data from the period 2004–2016. The following research questions were addressed: What is Czech milk production's competitive position in terms of productivity in the EU? What is the main source of milk production productivity growth in the Czech Republic? Is the change in TFP driven by the same component in the Czech Republic and other EU countries? How does Czech milk production differ from the most competitive producers in the EU? These questions were answered consecutively in the Results and discussion section.

The estimated IDF function revealed the lowest proportion of capital in the total input (only 6%), reflecting capital imperfections, especially at the beginning of the analysed period, in the group of EU-13 countries. However, the modernization of production towards more material-efficient and less land-bound technologies was also detected. The IDF also confirmed prevailing increasing returns to scale in the EU milk sector.

Examination of technical efficiency showed an above-average position of the Czech Republic (94.01%) compared to the EU-13 average (93.71%). However, the TE level average in the EU-13 was lower than the TE level average in the EU-15 (94.45%). Furthermore, the development of the average technical efficiency of EU-13 countries was characterized by more significant fluctuations than the average technical efficiency of EU-15 countries.

The development of technical efficiency influences the total factor productivity change. Analysis of the total factor productivity showed that the scale effect and technical efficiency change effect can be considered the main components of TFP changes in Czech milk production. Further examination discovered that the effect of technical efficiency responded more flexibly to the deterioration of production, sales, or economic conditions than the scale effect, where a time lag was apparent. On the other hand, the scale effect was more significant in other EU countries, especially the EU-15, compared to the Czech milk sector.

Czech milk production can be also characterized by a 1.8% year-on-year increase in the output-input ratio on average in the period 2004–2016, which indicates a quicker increase compared to EU-13 countries. On the other hand, milk production in the EU-15 countries was characterized by higher TFP growth than milk production in the EU-13. Quicker growth in the EU-15 was driven by a more significant scale effect and the contribution of technical change and technical efficiency.

In conclusion, the relationship between milk production growth and level of specialization should be also considered. In the Czech Republic, several factors influencing milk production productivity compared to the most competitive EU producers were seen. For example, higher agricultural area use, higher proportion of own feed production, lower feed quality, dependence on external input as labour and land or level of subsidies.

Acknowledgements

The results are the part of the solution of project No. QK1920398 „Duality in Czech Agriculture: Advantage or Disadvantage for New Generation Agriculture?“, supported by the Ministry of Agriculture of the Czech Republic, program ZEMĚ.

Corresponding authors
doc. Ing. Zdeňka Žáková Kroupová, Ph.D.
Department of Economics, Faculty of Economics and Management
Czech University of Life Sciences Prague, Kamycka 129, Prague - Suchdol, 165 00, Czech Republic
Phone: +420 224 382 303, E-mail: kroupovaz@pef.czu.cz

References

[1] Alvarez, A., del Corral, J. (2010) "Identifying different technologies using a latent class model: extensive versus intensive dairy farms", *European Review of Agricultural Economics*, Vol. 37, No. 2, pp. 231-250. ISSN 0165-1587. DOI 10.1093/erae/jbq015.

[2] Bakucs, L. Z., Fertö, I., Fagarasi, J. and Tóth, J. (2012) “Farm organisation and efficiency in Hungarian dairy farms”, *Milchwissenschaft*, Vol. 67, No. 2, pp. 147-150. ISSN 0026-3788.

[3] Barnes, A. (2008) "Technical Efficiency Estimates of Scottish Agriculture: A Note", *Journal of Agricultural Economics*, Vol. 59, No. 2, pp. 370-376. ISSN 1477-9552. DOI 10.1111/j.1477-9552.2008.00156.x.

[4] Bauer, P. W. (1990) “Decomposing TFP growth in the presence of cost inefficiency, nonconstant returns to scale, and technological progress”, *Journal of Productivity Analysis*, Vol. 1, No. 4, pp. 287-299. ISSN 1573-0441. DOI 10.1007/BF00160047.

[5] Bokusheva, B., Čechura, L. (2017) "Evaluating dynamics, sources and drivers of productivity growth at the farm level", OECD Food, Agriculture and Fisheries Papers, No. 106, OECD Publishing, Paris.

[6] Coelli, T., Singh, S. and Fleming, E. (2003) “*An input distance function approach to the measurement of technical and allocative efficiency: with application to Indian dairy processing plants*”. [Online]. Available: https://editorialexpress.com/cgibin/conference/download.cgi?db_name=esam2003&paper_id=105 [Accessed: 15 March 2015].

[7] Čechura, L., Grau, A., Hockmann, H., Levkovych, I. and Kroupova, Z. (2017) "Catching up or falling behind in Eastern European agriculture – the case of milk production", *Journal of Agricultural Economics*, Vol. 68, No. 1, pp. 206-227. ISSN 1477-9552. DOI 10.1111/1477-9552.12193.

[8] Dakpo, K. H., Desjeux, Y., Jeanneaux, P. and Latruffe, L. (2019) "Productivity, technical efficiency and technological change in French agriculture during 2002-2015: a Färe-Primont index decomposition using group frontiers and meta-frontier", *Applied Economics*, Vol. 51, No. 11, pp. 1166-1182. ISSN 0003-6846. DOI 10.1080/00036846.2018.1524982.

[9] Emvalomatis, G. (2012) "Productivity Growth in German Dairy Farming Using a Flexible Modelling Approach", *Journal of Agricultural Economics*, Vol. 63, No. 1, pp. 83-101. ISSN 1477-9552. DOI 10.1111/j.1477-9552.2011.00312.x.

[10] Frick, F. and Sauer, J. (2016) "Deregulation and Productivity: Empirical Evidence on Dairy Production", Paper presented on *the annual meeting of Agriculture & Applied Economics Association*, 31th July-2nd August 2016, Boston, USA.

[11] Grau, A. and Hockmann, H. (2016) "A new approach to identify market power along agri-food supply chains – the German dairy supply chain", Paper presented on *the annual meeting of Agriculture & Applied Economics Association*, 31th July-2nd August 2016, Boston, USA.

[12] Irz, X. and Jansik, C. (2015) "Competitiveness of dairy farms in northern Europe: A cross-country analysis", *Agricultural and Food Science*, Vol. 24, pp. 206-2018. ISSN 1795-1895. DOI 10.23986/afsci.50881.

[13] Irz, X. and Thirtle, C. (2004) "Dual Technological Development in Botswana Agriculture: A Stochastic Input Distance Function Approach", *Journal of Agricultural Economics*, Vol. 55, No. 3, pp. 455-478. ISSN 1477-9552. DOI 10.1111/j.1477-9552.2004.tb00110.x.

[14] Jondrow, J., Knox Lovell, C. A., Materov, I. S. and Schmidt, P. (1982) "On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model", *Journal of Econometrics*, Vol. 19, No. 2/3, pp. 233-238. ISSN 0304-4076. DOI 10.1016/0304-4076(82)90004-5.

[15] Keizer, T. H. and Emvalomatis, G. (2014) "Differences in TFP growth among groups of dairy farms in the Netherlands", *NJAS-Wageningen Journal of Life Science*, Vol. 70-71, pp. 33-38. ISSN 1573-5214. DOI 10.1016/j.njas.2014.03.001.

[16] Kumbhakar, S. C. (2011) "Estimation of production technology when the objective is to maximize return to the outlay", *European Journal of Operational Research*, Vol. 208, pp. 170-176. ISSN 0377-2217. DOI 10.1016/j.ejor.2010.09.015.

[17] Kumbhakar, S. C. and Tsionas, E. G. (2009) "Stochastic error specification in primal and dual production systems", *Journal of Applied Econometrics*, Vol. 26, pp. 270-297. ISSN 1099-1255. DOI 10.1002/jae.1100.

[18] Kumbhakar, S. C. and Tsionas, E. G. (2008) "Estimation of input-oriented technical efficiency using a nonhomogeneous stochastic production frontier mode", *Agricultural Economics*, Vol. 38, pp. 99-108. ISSN 1574-0862. DOI 10.1111/j.1574-0862.2007.00285.x.

[19] Kumbahakar, S. C., Lien, G., Flaten, O. and Tveteras, R. (2008) "Impacts of Norwegian Milk Quotas on Output Growth: A Modified Distance Function Approach", *Journal of Agricultural Economics*, Vol. 59, No. 2, pp. 350-369. ISSN 1477-9552. DOI 10.1111/j.1477-9552.2008.00154.x.

[20] Kumbhakar, S. C. and Lozano-Vivas, A. (2005) "Deregulation and Productivity: The Case of Spanish Banks", *Journal of Regulatory Economics*, Vol. 27, No. 3, pp. 331-351. ISSN 0922-680X. DOI 10.1007/s11149-005-6627-2.

[21] Lansink, A. O., Stefanou, S. and Serra, T. (2015) "Primal and dual dynamic Luenberger productivity indicators", *European Journal of Operational Research*, Vol. 241, pp. 555-563. ISSN 0377-2217. DOI 10.1016/j.ejor.2014.09.027.

[22] Lansink, A. O., Silva, E. and Stefanou, S. (2000) “Decomposing productivity growth allowing efficiency gains and price-induced technical progress”, *European Review of Agricultural Economics*, Vol. 27, No. 4, pp. 497-518. ISSN 0165-1587. DOI 10.1093/erae/27.4.497.

[23] Latruffe, L., Bravo-Ureta, B. E., Moreira, V. H., Desjeux, Y. and Dupraz, P. (2011) “Productivity and Subsidies in European Union Countries: An Analysis for Dairy Farms Using Input Distance Frontiers”, Paper presented on the *12th congress of European Association of Agricultural Economists,* 30th August-2nd September 2011, Zürich, Switzerland.

[24] Luik, H., Omel, R. and Viira, A.-H. (2011) "Efficiency and productivity change of Estonian dairy farms from 2001-2009", The article presented at *congress EAAE: Change and Uncertainty Challenges for Agriculture, Food and Natural Resources*, 30.8.-2.9.2011, ETH Zurich, Zurich, Švýcarsko.

[25] Madau, F. A., Furesi, R. and Pulina, P. (2017) "Technical efficiency and total factor productivity changes in European dairy farm sector", *Agricultural and Food Economics*, Vol. 5, No. 17, pp. 1-14. ISSN 2193-7532. DOI 10.1186/s40100-017-0085-x.

[26] Morrison Paul, C., Nehring, R., Banker, D. and Somwaru, A. (2004) "Scale Economies and Efficiency in U.S. Agriculture: Are Traditional Farms History?", *Journal of Productivity Analysis*, Vol. 22, No. 3, pp. 185-205. ISSN 0895-562X. DOI 10.1007/s11123-004-7573-1.

[27] Newman, C. and Mathews, A. (2007) "Evaluating the Productivity Performance of Agricultural Enterprises in Ireland using a Multiple Output Distance Function Approach", *Journal of Agricultural Economics*, Vol. 58, No. 1, pp. 128-151. ISSN 1477-9552. DOI 0.1111/j.1477-9552.2007.00084.x.

[28] Rasmussen, S. (2011) “Estimating the Technical Optimal Scale of Production in Danish Agriculture”, Paper presented on *the 12th congress of European Association of Agricultural Economists*, 30th August-2nd September 2011, Zürich, Switzerland.

[29] Sipiläinen, T., Kumbhakar, S. C. and Lien, G. (2014) “Performance of dairy farms in Finland and Norway from 1991 to 2008”, *European Review of Agricultural Economics*, Vol. 41, No. 1, pp. 63-86. ISSN 0165-1587. DOI 10.1093/erae/jbt012.

[30] Skevas, I., Emvalomatis, G. and Brümmer, B. (2018) “Heterogenity of Long-run Technical Efficiency of German Dairy Farms: A Bayesian Approach”, *Journal of Agricultural Economics*, Vol. 69, No. 1, pp. 58-75. ISSN 1477-9552. DOI 10.1111/1477-9552.12231.

[31] Špička, J. and Machek, O. (2015) "Change in the production efficiency of European specialized milk farming", *Agricultural Economics – Czech*, Vol. 61, No. 1, pp. 1-13. ISSN 1805-9295. DOI 10.17221/112/2014-AGRICECON.

[32] Wojciechowski, L. (2017) "Productivity Gap: A Chance or an Obstacle in Absorbing Benefits from FDI in a Host Country", *Entrepreneurial Business and Economics Review*, Vol. 5, No. 4, pp. 153-170. ISSN 2353-8821. DOI 10.15678/EBER.2017.050407.

[33] Zhu, X. and Lansink, A. O. (2009) "Determinants of productivity change of crop and dairy farms in Germany, the Netherlands and Sweden in 1995-2004", Paper presented at *the International Association of Agricultural Economists Conference*, 16.8.-22.8.2009, Beijing, China.

[34] Žáková Kroupová, Z. (2016) “Profitability development of Czech dairy farms”, *Agricultural Economics-Czech*, Vol. 62, pp. 269-279. ISSN 1805-9295. DOI 10.17221/131/2015-AGRICECON.

Agris on-line Papers in Economics and Informatics

Volume XII Number 3, 2020

Agris on-line Papers in Economics and Informatics

The international reviewed scientific journal issued by the Faculty of Economics and Management of the Czech University of Life Sciences Prague.

http://online.agris.cz

ISSN 1804-1930