sciendo

DOI: 10.2478/bile-2018-0013

Biometrical Letters
Vol. 55 (2018), No. 2, 197-214

Use of classification and regression trees (CART) for analyzing determinants of winter wheat yield variation among fields in Poland

M. Iwańska[1], A. Oleksy[3], M. Dacko[4], B. Skowera[5], T. Oleksiak[2], E. Wójcik-Gront[1]

[1]Department of Experimental Design and Bioinformatics, Warsaw University of Life Science – SGGW, Poland; e-mail: m.iwanska123@gmail.com
[2]Plant Breeding and Acclimatization Institute – National Research Institute (IHAR-PIB), Radzików, Poland
[3]Institute of Plant Production, University of Agriculture in Krakow, Poland
[4]Institute of Economic and Social Sciences, University of Agriculture in Krakow, Poland
[5]Department of Ecology, Climatology, and Air Protection, Faculty of Environmental Engineering and Land Surveying, University of Agriculture in Krakow, Poland

SUMMARY

Wheat is one of the modern world's staple food sources. Its production requires good environmental conditions, which are not always available. However, agricultural practices may mitigate the effects of unfavorable weather or poor-quality soils. The influence of environmental and crop management variables on yield can be evaluated only based on representative long-term data collected on farms through well-prepared surveys.The authors of this work analyzed variation in winter wheat yield among 3868 fields in western and eastern Poland for 12 years, as dependent on both soil/weather and crop management factors, using the classification and regression tree (CART) method. The most important crop management deficiencies which may cause low wheat yields are insufficient use of fungicides, phosphorus deficiency, non-optimal date of sowing, poor quality of seeds, failure to apply herbicides, lack of crop rotation, and use of cultivars of unknown origin not suitable for the region. Environmental variables of great importance for the obtaining of high yields include large farm size (10 ha or larger) and good-quality soils with stable pH. This study makes it possible to propose strategies supporting more effective winter wheat production based on the identification of characteristics that are crucial for wheat cultivation in a given region.

Key words: classification and regression trees (CART), determinants of yield, winter wheat, yield modeling, farm data

1. Introduction

Winter wheat is one of the main crops used for food production worldwide. Poland produces 11 Mt of winter wheat grains; it is the fourth largest wheat producer in the European Union (after France, Germany and the UK) and 15th in the world (CSO 2017). The worldwide yield of winter wheat has been growing steadily for a number of years. In Poland the total yield rose from 3.6 t ha^{-1} in 1995 to almost 5.0 t ha^{-1} in 2014 (FAOSTAT 2014). The upward trend results from the genetic improvement of cultivars and more intensive agricultural practices (Loyce et al., 2008; Montesino-San Martína et al., 2014; Rozbicki et al., 2015). Due to climate change, environmental conditions have become more unfavorable for wheat production (Reidsma et al., 2010; Gornott and Wechsung, 2016; Montesino-San Martína et al., 2014). Thus, researchers have been making efforts to identify the influence of growing conditions and weather on wheat yields and to find ways to counterbalance the environmental stress by means of agricultural practices. However, most research done so far has concerned only the few most important agricultural practices, taking into account weather conditions over a short period of time, usually three years (Rozbicki et al., 2015; Fallahi et al., 2008; Girma et al., 2007; Lazaro et al., 2010). The yield of crops depends on many interrelated variables characterizing soil quality, weather conditions and the intensity of agricultural practices. Their influence on yield can be evaluated only based on representative data collected on farms through well-prepared surveys (Ferraro et al., 2009; Zheng et al., 2009; Delmotte et al., 2011).When the data cover a wide time range—several years, for instance—with different environmental conditions and a considerable number of diversified farms, classic statistical methods might not be adequate to find cause-and-effect relationships between all variables and the outcome (yield). In this case data mining methods such as CART (Classification And Regression Trees) might be very useful (Krupnik et al., 2015; Sileshi et al., 2010; Roel et al., 2007; Breiman et al., 1984, Dacko et al., 2016). There is a great difference between the potential and real yield of winter wheat in Poland. Moreover, the yield is subject to high variation across the country. It is extremely important to understand the source of limitations on effective grain production. To date, however, there have been few studies of the influence of many different environmental conditions and varied agricultural practices on yield on a national scale using CART (Ferraro et al., 2009; Zheng et al., 2009; Delmotte et al., 2011). Thus, the aims of this work were: i) to illustrate the use and usefulness of the CART method in analyzing winter wheat yield variability among fields across Poland as caused by both soil/weather and crop management factors; ii) to compare the contributions of environmental and management factors on winter wheat yield variation in Poland (based on the importance of predictors evaluated in CART); and iii) to propose strategies supporting more effective winter wheat production.

2. Material and methods

2.1. Sampling of farms and empirical data set

The material taken for statistical analysis consisted of long-term (1992–2003) production data from individual farms. The data were obtained in research conducted by the Laboratory of Economics of Seed and Plant Breeding (IHAR-PIB) in Radzików, Poland. The survey was carried out among farms holding agricultural accounting data for the needs of the Institute of Agricultural Economics and Food Economy (IAFE-IERiGŻ). The selection of farms for the study was carried out by IAFE-IERiGŻ employees, to include representative farms for a given region. The surveys were completed by IAFE-IERiGŻ agricultural accountants. The interviewers filled out forms prepared as MS Excel worksheets, including data on a farm’s production region, the area of agricultural land, and the age and education of the farmer. Each survey covered one farm, and concerned one field, the cultivation of one cultivar of one crop species (wheat or other cereal species) and information on seed certification grades, field characteristics, production factors and crop management factors. The winter wheat yield (treated as the dependent variable in the CART analysis) was expressed as grain yield in dt (0.1 ton) per hectare for a field. The determinants of winter wheat yield variation used for the construction of classification and regression trees (CART) were factors related to soil/weather and crop management (19 predictors; Table 1). All calculations and diagrams in the CART model in this paper were built using STATISTICA ver. 12 software (StatSoft, Inc., 2014).

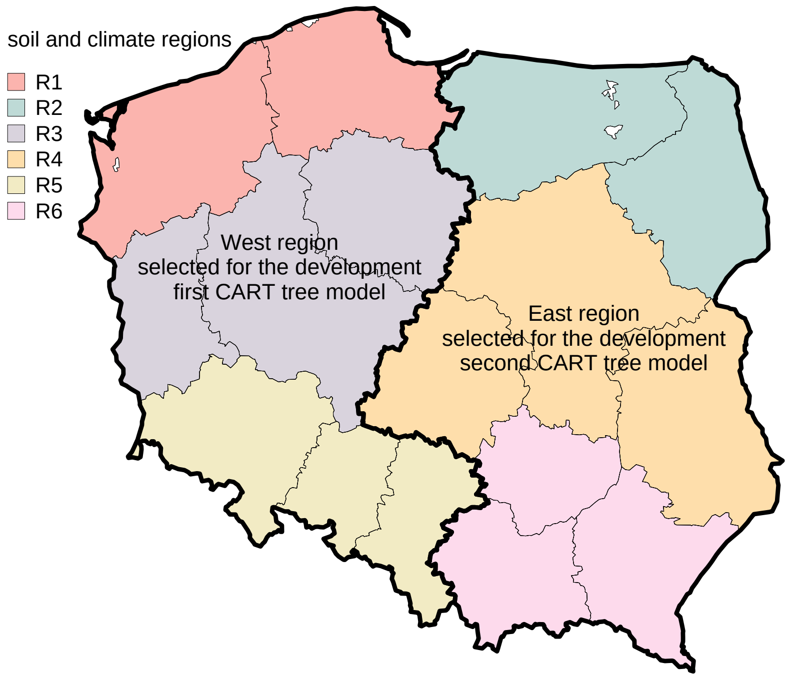

Figure 1. COBORU soil and climate regions in Poland (authors' graphic)

2.2. CART model description

In the analysis an interactive CART method was used. This is a method commonly used in data mining. A model is created that predicts the value of a target variable (yield) based on several input variables (environmental, genetic and management variables). The process of CART regression tree construction is based on seeking opportunities for the dichotomous division of observations into

Table 1. Factors (hypothetically) affecting the grain yield of winter wheat as used in the CART model

Factor (predictor variable) name	Variable type	Description of variables
Fertilizers applied 1. Nitrogen (N); 2. Phosphorus (P); 3. Potassium (K)	categorical variable on an ordinal scale	not used (0 kg ha^{-1}); low (≤ 50 kg ha^{-1}); moderate (50–100 kg ha^{-1}); high (> 100 kg ha^{-1})
4. Manure fertilizations	categorical variable on an ordinal scale	just before sowing; under forecrop; before pre-forecrop; 3 years ago; 4 years ago; 5 years ago; more than 5 and less than 10 years ago; more than 10 years ago
Pesticides applied 5. Fungicide; 6. Insecticide; 7. Herbicide	categorical variable on an ordinal scale	used (1, 2, 3 treatments); not used
8. Sowing date*	categorical variable on an ordinal scale	early; optimum; late
9. Seeding rate	categorical variable on an ordinal scale	low (140–200 kg ha^{-1}); optimum (201–250 kg ha^{-1}); high (251–300 kg ha^{-1}); very high (>301 kg ha^{-1})
10. Area of agricultural land (size of farm)	categorical variable on an ordinal scale	up to 3 ha; 3–5 ha; 5–8 ha; 8–10 ha; 10–15 ha; 15–20 ha; 20–25 ha; 25–35 ha; 35–50 ha; 50–100 ha; more than 100 ha
11. Forecrop	categorical variable	Cereal; Roots and tuber crops; Pulses and perennial legumes; Winter oilseed rape
12. Soil valuation class	categorical variable on an ordinal scale	I (best quality soils); II (very good quality); IIIa–IIIb (good quality); IVa–IVb (medium quality); V (poor quality); VI (poorest quality)
13. Soil pH	categorical variable on an ordinal scale	Strong acidic; Acidic; Slightly acidic; Neutral; Alkaline
14. Seed quality	qualitative variable	certified seeds; non-certified seeds
15. Cultivar	qualitative variable	National Registered**; Unknown national; Unknown foreign
Hydrothermal coefficient for; 16. April, 17. May, 18. June, 19. July***	categorical variable on an ordinal scale	$k \leq 0.4$ (extremely dry); $0.4 < k \leq 0.7$ (very dry); $0.7 < k \leq 1.0$ (dry); $1.0 < k \leq 1.3$ (fairly dry); $1.3 < k \leq 1.6$ (optimal); $1.6 < k \leq 2.0$ (fairly moist); $2.0 < k \leq 2.5$ (moist); $2.5 < k \leq 3.0$ (very moist); $k > 3.0$ (extremely wet)

* Sowing dates were adjusted according to the COBORU soil and climate regions in Poland
**Grana, Begra, Salwa, Gama, Emika, Jawa, Alba, Parada, Almarii, Kamila, Juma, Kobra, Roma, Mikon, Elena, Sakwa, Korwetta, Zyta
***Sielianinov hydrothermal coefficient /Source: own study/

subsets (nodes) having different mean values of the dependent variable (Dacko et al., 2016) so that the reduction in impurity (variance) of the dependent variable is the largest. The method requires a considerably large dataset (Dacko et al., 2016; Ferraro et al., 2009; Topal et al., 2010). In addition, in the CART model it is possible to estimate and rank the importances of the predictors used. The interactive method of CART trees incorporates the researcher's interactions into the final choice of the predictor and its specific value. This is justified, for example, where several predictors give the same reduction in the impurity of dependent variable. Then, based on expert knowledge, the most reasonable predictor is selected. This mechanism was used in this work to find the variables most responsible for variation in winter wheat yield in Poland. A collection of yield data was evaluated for 3868 fields across the years 1992–2003. For each wheat yield entry a set of 19 qualitative predictors was used, describing seed characteristics, hydrothermal and soil conditions, and crop management treatments applied by the farmers.

3. Results

The variation in winter wheat grain yield in the analyzed period (1992–2003) was relatively high. The distributions of yields over the twelve years, however, were different between the western and eastern regions of Poland (Fig. 2). From the preliminary analysis, one predictor was seen to have played a key role in differentiating yields of winter wheat. This was location, expressed in terms of regions of Poland with different soil and climate conditions, reflecting in this context the division into western and eastern Poland (see Fig. 1: COBORU soil-climate regions). For this reason, it was decided to construct two separate models: one for the western regions (R1, R3, R5) and one for the eastern regions (R2, R4, R6). Based on the estimated coefficients of variation (CV%), in the western regions (R1, R3 and R5; Fig. 1) the standard deviation of yield was 27% of the mean, while in the eastern regions (R2, R4 and R6) it was close to 25% of the mean. The analyzed production data for the western regions included 1833 fields on which winter wheat was grown. Grain yields in these regions ranged from 14.7 to 75.0 dt ha^{-1}, with an average yield of 43.6 dt ha^{-1} and a standard deviation of 11.6 dt ha^{-1}. In the eastern parts of the country production data were analyzed from 2035 fields used for the cultivation winter wheat, and grain yields ranged from 10.0 to 66.0 dt ha^{-1} with an average yield of 36.7 dt ha^{-1} and a standard deviation of 9.3 dt ha^{-1}. The distribution of winter wheat yields in both parts of the country was normal (Fig. 2). In the western regions more than 87% of the fields had yields ranging from 30.0 to 60.0 dt ha^{-1} (Fig. 2, part A), while in the eastern regions just over 83% of the fields had yields ranging from 30.0 to 50.0 dt ha^{-1} (Fig. 2, part B). For both of these macro-regions of Poland, variation in winter wheat grain yield occurred due to the wide range of variability in soil and hydrothermal conditions and in the intensity of application of crop management inputs.

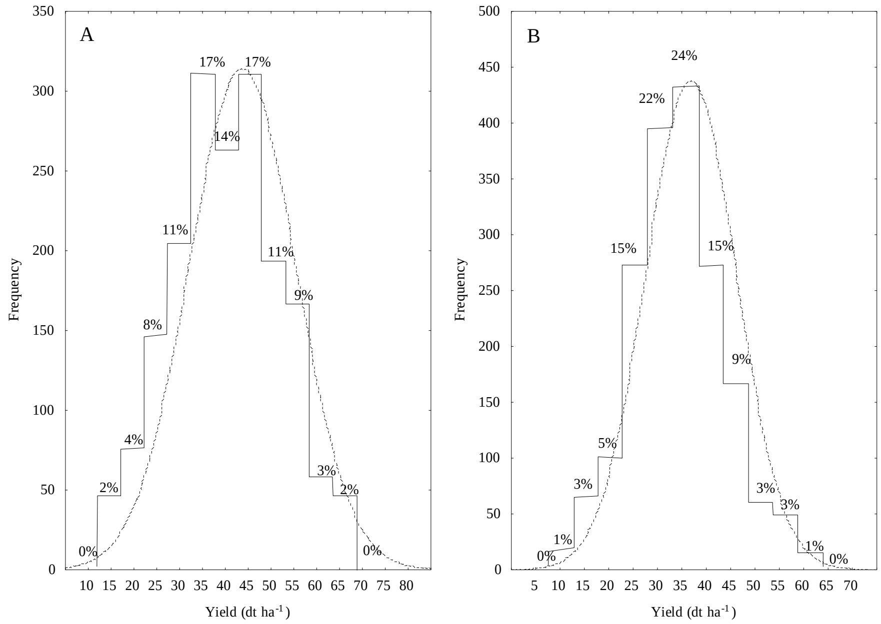

Figure 2. Distribution of winter wheat yield**: A -** in western **and B - in** eastern regions of Poland

Interactive CART regression trees were employed twice, separately for each of the two selected macro-regions of Poland. In the analysis of production data from individual farms, crop management factors as well as soil and hydrothermal coefficient factors (19 predictors; Table 1) were taken into account to select the most important predictors contributing to variation in winter wheat yield over the period of twelve years (1992–2003) for western and eastern regions separately.

The interactive CART model explaining the winter wheat yield variation for the western regions of Poland led to a regression tree consisting of 11 terminal nodes and 10 splitting nodes (Fig. 3).

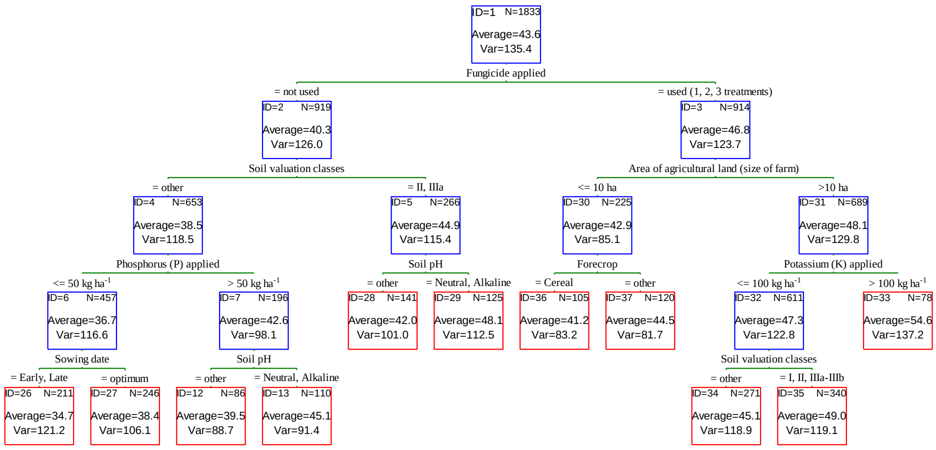

Figure 3. Interactive CART regression tree explaining winter wheat grain yield variation in western regions of Poland

In these western regions of the country, applied fungicide protection treatments were the most important factor determining yield variation (Fig. 3). The numbers of unprotected and fungicide-protected fields were similar, and average yields over the period of twelve years were 43.6 dt ha^{-1}. However, the failure to use fungicide treatments to protect crops led to a decrease in the average yield by 6.5 dt ha^{-1}. On farms that used such extensive cultivation (no fungicides), the soil valuation class, as recorded in the agricultural land register, was an important determinant of the level of yield. Based on this criterion (soil quality class) a further division of the CART regression tree occurred. In better soil conditions, recognized in Poland as typical for the growing of wheat and in accordance with the soil requirements of that plant (soil valuation classes: II – very good quality and IIIa – good quality), grain yields were higher on average by 6.4 dt ha^{-1} than on the poorer soils assigned to the soil valuation classes IIIb – good quality, IVa–IVb – medium quality, V – poor quality and VI – poorest quality. As shown in the tree diagram (Fig. 3), the yield of wheat grown on better soils was significantly higher. In turn, on weaker soils the amount of phosphorus applied had a significant influence on the yield. Plantations not fertilized with phosphorus (0 kg ha^{-1}) or fertilized with a low dose (up to 50 kg ha^{-1}), compared with those fertilized with phosphorus in a moderate dose (50–100 kg ha^{-1}) or high dose (above 100 kg ha^{-1}) gave lower yields on average by 14% (5.9 dt ha^{-1}). In fields not fertilized or fertilized with only low doses of phosphorus, sowing date was an important crop management factor for wheat yield. Early or late sowing, in comparison with the optimal sowing date in western Poland, gave lower yields on average by 3.8 dt ha^{-1}. When a higher dose of phosphorus was applied the yield depended on soil pH, being higher by 5.6 dt ha^{-1} on neutral and alkaline soils than on more acidic soils. Let us return to the first division of the CART tree, which suggested that higher yields of winter wheat are favored by the use of protection treatment against fungal diseases. Where such treatment was undertaken, the most important factor for further differentiation was the area of agricultural land (size of farm). Smaller farms (up to 10 ha) achieved lower yields on average by 5.2 dt ha^{-1} than farms above 10 ha. On smaller farms, which are inherently characterized by lower production intensity, the yield of winter wheat depended on the forecrop (position in the crop rotation cycle). Lower yields, by 3.3 dt ha^{-1}, were recorded where wheat was cultivated after cereal crops, which in theory are commonly regarded as worse forecrops for winter wheat. An increase in winter wheat yield was favored by forecrops of winter oilseed rape, root and tuber crops, and pulses and perennial legumes. On larger farms, the level of potassium (K) fertilization proved to be an important crop management factor for the winter wheat yield.

In comparison with plantations that used no fertilizers or applied them in low doses (≤ 50 kg ha^{-1}) or moderate doses (50–100 kg ha^{-1}), the use of potassium fertilizers in high doses (above 100 kg ha^{-1}) led to the highest yield of grain (54.6 dt ha^{-1}). However, let us recall that this result was achieved by larger farms (over 10 ha of agricultural land) applying protective treatment against fungal pathogens. When the fertilizer dose was lower (< 100 kg ha^{-1}) the level of wheat yield was determined by the soil quality class. A relatively high average yield (49.0 dt ha^{-1}), when potassium fertilizer use was less intensive, was possible only on the best soil valuation classes (I – best quality soils, II – very good quality, IIIa–IIIb – good quality) subject to the conditions relating to fungicide protection and area of agricultural land.

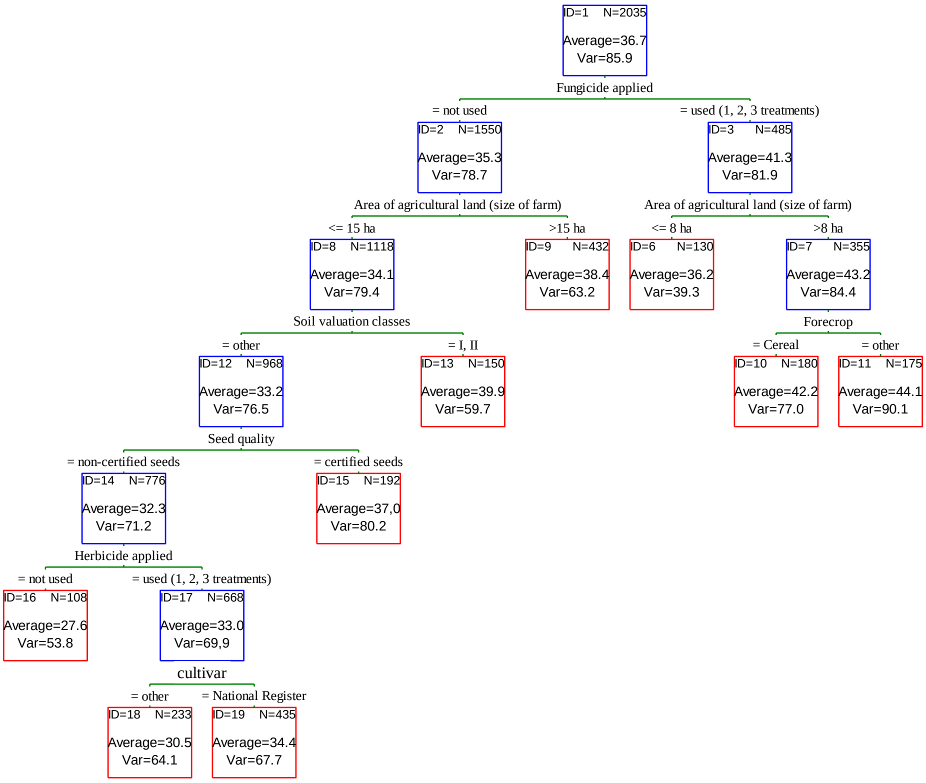

Figure 4. Interactive CART regression tree explaining winter wheat grain yield variation in eastern regions of Poland

In summary, in the discussed CART interactive tree model (Fig. 3), there were two divisions based on environmental factors: soil valuation class and soil pH. Among the crop management factors, the divisions that followed were based on the levels of phosphorus and potassium fertilizers applied, the sowing date and the forecrop.

Significant predictors explaining winter wheat yield variation in the western part of Poland also included the size of farms (expressed by the area of agricultural land). This criterion provides a more detailed division of the plantations subject to protective treatment against fungal pathogens, i.e. with more intensive cultivation technologies.

The interactive CART regression tree model explaining the variation of winter wheat yields in the eastern part of Poland was less extensive: it consisted of 8 splitting nodes and 9 terminal nodes (Fig. 4). The average grain yield over 12 years was 36.7 dt ha^{-1}. Also here, against the background of other crop management treatments, the application of fungicides proved to be a key determinant of yield variability (Fig. 4). However, such treatments were used much less often than in the western regions. Fungicides were applied on only 24% of the 2035 fields studied, although the effect was a statistical improvement in crop yield by 6.0 dt ha^{-1}. An important predictor of the yield of winter wheat grain in the eastern part of the country was the size of farms, expressed by the area of agricultural land. On the basis of this variable, subsequent tree divisions occurred both within fields protected with fungicides and within non-protected fields. On larger farms (with more than 8 ha of agricultural land) where fungicides were used, the yield of wheat averaged 43.2 dt ha^{-1}. The interactive CART tree diagram also shows that the yield was improved by 10.0 dt ha^{-1} by the use of appropriate forecrops: winter oilseed rape, root and tuber crops, and pulses and perennial legume plants. The model showed that the highest average yield (44.1 dt ha^{-1}) was obtained by those farmers in eastern Poland who managed larger areas of agricultural land, used active fungicide protection and selected a forecrop favorable to winter wheat.

On smaller farms (up to 15 ha of agricultural land) where the wheat crop was not protected by fungicides, grain yields averaged 34.1 dt ha^{-1} and depended on the soil valuation class. For soil in class IIIa (good quality) and poorer, crop yields were differentiated by the seed quality. Where non-certified seed was used (the vast majority of cases) the yield was lower by 4.7 dt ha^{-1} compared with farms using certified seed. Where non-certified seed was used, the next crop management factor in the division of the CART tree was herbicide protection. In fields where weed control was limited, the winter wheat yield was higher (33.0 dt ha^{-1}) and depended on the cultivar. Cultivation based on domestically bred cultivars, entered in the national COBORU register, led to relatively high grain yields compared with unknown foreign cultivars. In the light of the CART model, the worst results were obtained by smaller farms with poor-quality soils that did not apply protection against diseases and weeds, and used non-certified seed (27.6 dt ha^{-1}).

The CART models provide a synthetic assessment of the importance of yield predictors in the creation of the sequence of trees for eastern and western Poland. For each predictor, the model calculates the corresponding reduction in impurity for yield variability. The predictors can then be compared with each other using a scale from 0 to 1, creating a ranking of predictors. According to this hierarchy, key variables that explain the level of winter wheat yield include the area of agricultural land and the soil valuation class, and highly important ones include cultivar, use of potassium and phosphorus fertilizers, seed quality, soil pH and seeding rate (Table 2). Fungicides produced the greatest reduction in impurity in CART for eastern Poland. However, in comparison with other predictors in the sequence of trees created for eastern Poland (for alternative splits) their importance was not the highest.

Table 2. Ranking of predictors by importance

Predictor	Validity (western Poland)	Validity (eastern Poland)	Mean validity	Validity category
Area of agricultural land (size of farm)	0.96	0.71	0.84	Key
Soil valuation class	1.00	0.61	0.80	
Cultivar	0.51	1.00	0.76	Very important
Potassium (K) fertilizer applied	0.98	0.37	0.68	
Phosphorus (P) fertilizer applied	0.84	0.30	0.57	
Seed quality	0.65	0.47	0.56	
Soil pH	0.86	0.26	0.56	
Seeding rate	0.50	0.57	0.54	
Herbicide applied	0.26	0.72	0.49	Important
Hydrothermal coefficient for April	0.62	0.35	0.49	
Nitrogen (N) fertilizer applied	0.59	0.37	0.48	
Fungicide applied	0.60	0.35	0.48	
Hydrothermal coefficient for July	0.55	0.37	0.46	
Forecrop	0.53	0.32	0.43	
Hydrothermal coefficient for May	0.48	0.35	0.42	Less important
Hydrothermal coefficient for June	0.38	0.26	0.32	
Manure application	0.29	0.17	0.23	
Sowing date	0.13	0.19	0.16	Not important
Insecticide applied	0.10	0.13	0.12	

Source: own study

4. Discussion

The CART analysis revealed that environment and crop management variables were very important in explaining winter wheat yield variability in Poland; however, the hierarchy of these variables differed between the western and eastern parts of the country. Eight predictive variables were significant in explaining winter wheat yield variation for the west of Poland. These included variables characterizing soil quality, pH and crop management (the use of fungicides, phosphorus and potassium fertilizer rates, forecrop, date of sowing, and farm size). In the case of eastern Poland the seven most important variables were the use of fungicides and herbicides, farm size, forecrop, seed quality, cultivar and soil quality.

CART analysis aimed at explaining soy yield variability performed by Zheng et al. (2010) indicated that agronomic practices were equally important as soil quality. Lobell et al. (2002, 2005), Tittonell et al. (2008) and Zheng et al. (2009) showed that management practices were the most important factors explaining the variation in yield of various agricultural crops in different locations (51% to 93% explained variability).

In this study, the CART models explained 20% (r=0.44) of winter wheat yield variation in the case of eastern Poland and 22% (r=0.47) in western Poland. These low percentages can be explained by the large and very diversified area covered (the whole of Poland) and the long period of time (12 years) with varied weather conditions. Zheng et al. (2010) concluded that the contribution of soil and crop management variables to explaining agricultural crop yield variation depends on the areal extent of the experiments. An earlier study of Zheng et al. (2009) conducted in a smaller area with several fields in one village showed that soil and crop management variables explained 81% of soy yield variation, while crop management variables alone explained 76% of the variation. However, later Zheng et al. (2010) found in experiments conducted at regional level that crop management variables explained only 22.7% of soy yield variation. This indicates that the contribution of crop management variables to explaining yield variation decreases when the area covered by the experiments becomes larger. This was later confirmed by Zhang et al. (2012) in one-year experiments conducted in Fengqiu county in Henan province (China). Using CART, the authors explained 25% of wheat yield variation with variables related to crop management and soil. When only crop management variables were taken into account they explained 19.4% of the variation. Thus, low values of explained yield variation with the use of CART is not unusual when large areas are considered. For our 12-year database the values were 20% and 22% respectively for the eastern and western parts of Poland. CART cannot be used for precise prediction. However, it allows the user to determine the most important variables and their relative performance in explaining variation in yields of a crop. Krupnik et al. (2015) studied the influence of crop management and environment on wheat yield variation in

Bangladesh. They concluded that the most important predictor was the nitrogen fertilization rate and environment, regardless of the sowing date. In that work, environments with high and low efficiency of wheat production were identified. However, even fields with salty soils in a hot climate sometimes gave reasonable yields, which means that unfavorable environmental conditions can be counterbalanced by optimal crop management. From many variables—landform, use of manure, forecrop, straw tillage, sowing date, number of sown seeds, N, P, K fertilization, herbicides, insecticides, irrigation, lodging—Zhang et al. (2012) identified N fertilization as the main predictor of wheat yield variation. In their work N fertilization rates were very high (257.6 kg N ha^{-1}), in contrast to our study, where very low or no N fertilization was used, and it turned out to be moderately important in explaining winter wheat yield variation. The same was observed in survey-based studies carried out in Germany by Macholdt and Honermeier (2017). According to their findings, more important variables were the choice of cultivar, forecrop and chemical plant protection. Similarly, in our study the most important proved to be the use of fungicides. The same was found by Bertelsen et al. (2001), Mercer and Ruddock (2005) and Loyce et al. (2008). In another study by Lobell et al. (2005) based on CART analysis it was N fertilization that was the main predictor in explaining wheat yield variation, followed by the date of sowing, and the regression model explained 44% of the variation. However, those authors also state that the predictors' importance depended on weather conditions, and in the following year the time between planting and first irrigation was the most important. Thus, it is always desirable to include long-term data to avoid potentially untypical climatic conditions encountered in a shorter time frame. The present work includes data collected across 12 years. The year variable was replaced by the Sielianinov coefficient, which combines the sum of precipitation and the air temperature during spring-summer plant vegetation. The hierarchy of predictors in our study was similar to the results obtained by Macholdt and Honermeier (2017).

5. Conclusions

In the analysis the authors evaluated the influence of environmental and crop management variables on winter wheat yield variation in Poland. The most important variables which led to low yields of wheat in 1992–2003 were non-use of fungicides, low-quality soils with phosphorus deficiency, and incorrect sowing date. Additionally, in the eastern part of Poland, important variables included a high level of farm fragmentation, poor quality of seeds and absence of herbicide protection. The results show that there is room for improvement in terms of better crop management, especially on small farms in the east of Poland, or the use of better-quality soils with stable pH for crop cultivation.

The main conclusion from this work is that the key factor in achieving higher yields is the use of fungicides in two or three treatments; this applies across the whole of Poland. This strategy is the most efficient on farms with cultivated areas above 10 ha, combined with high rates of potassium fertilization (above 100 kg ha^{-1}). In farms with cultivated areas up to 10 ha which use fungicides, higher yields can be achieved with the use of forecrops other than cereals. When no fungicides are used, especially in the western part of Poland, higher yields can be obtained on higher-quality soils (class II and IIIa) with stable pH. On lower-quality soils it is necessary to use phosphorus fertilization with rates above 50 kg ha^{-1} on soils with pH neutral or alkaline. In the eastern part of Poland, where environmental conditions for winter wheat cultivation are less favorable and farm fragmentation is greater, when no fungicides are used the crop should be grown on the best available soils and using certified seeds. Otherwise, farmers should take into account the use of herbicides and the selection of cultivars recommended for their region.

In this work, CART was used to explain the influence of environmental and crop management variables on winter wheat yield based on a large and complex database. This makes it possible to propose strategies leading to more effective winter wheat production in different regions of Poland.

REFERENCES

Bertelsen J.R., de Neergaard E., Smedegaard-Petersen V. (2001): Fungicidal effects of azoxystrobin and epoxiconazole on phylloshere fungi, senescence and yield of winter wheat. Plant Pathol. 50: 190-205.

Breiman L., Friedman J.H., Olshen R.A., Stone C.J. (1984): Classification and Regression Trees. Chapman and Hall (Wadsworth, Inc.), US New York.

CSO. Central Statistical Office. Concise Statistical Yearbook of Poland. (2017): Available online: https://danepubliczne.gov.pl/dataset/5c9f136c-025d-4b82-b03c-d8d7148dfe09/resource/cd90dfe3-1665-4a4d-b034-25e0ead1b389/download/malyrocznikstatystycznypolski2017.pdf (accessed on 4 January 2018).

Dacko M., Zając T., Synowiec A., Oleksy A., Klimek-Kopyra A., Kulig B. (2016): New approach to determine biological and environmental factors influencing mass of a single pea (Pisum sativum L.) seed in Silesia region in Poland using a CART model. Eur. J. Agron. 74: 29-37.

Delmotte S., Tittonell P., Mouret J.C., Hammond R., Lopez-Ridaura S. (2011): On farm assessment of rice yield variability and productivity gaps between organic and conventional cropping systems under Mediterranean climate. European Journal of Agronomy, 35(4): 223-236.

Fallahi H.A., Nasseri A., Siadat A. (2008): Wheat Yield Components are Positively Influenced by Nitrogen Application under Moisture Deficit Environments. International Journal of Agriculture & Biology, 10: 673-676.

Ferraro D.O., Rivero D.E., Ghersa C.M. (2009): An analysis of the factors that influence sugarcane yield in Northern Argentina using classification and regression trees. Field Crops Research, 112(2-3): 149-157.

Girma K., Holtz S.L., Arnall D.B., Fultz L.M., Hanks T.L., Lawles K.D., Mack C.J., Owen K.W., Reed S.D., Santillano J., Walsh O., White M.J., Raun W.R. (2007): Weather, Fertilizer, Previous Year Yield, and Fertilizer Levels Affect Ensuing Year Fertilizer Response of Wheat. Agronomy Journal, 99:1607-1614.

Gornott C., Wechsung F. (2016): Statistical regression models for assessing climate impacts on crop yields: A validation study for winter wheat and silage maize in Germany. Agric. For. Meteorol. 217: 89-100.

Krupnik T.J., Ahmed Z.U., Timsina J., Yasmin S., Hossain F., Al Mamun A., Mridha A.I., McDonald A.J. (2015): Untangling crop management and environmental influences on wheat yield variability in Bangladesh: an application of non-parametric approaches. Agricultural Systems, 139: 166-179.

Lazaro L., Abbate P.E., Cogliatti D.H., Andrade F.H. (2010): Relationship between yield, growth and spike weight under phosphorus deficiency and shading. Journal of Agricultural Science, 148: 83-93.

Lobell D.B., Ortiz-Monasterio J.I., Asner G.P., Naylor R.L., Falcon W.P. (2005): Combining field surveys, remote sensing, and regression trees to understand yield variations in an irrigated wheat landscape. Agronomy Journal, 97(1): 241-249.

Lobell D., Ortiz-Monasterio J., Addams C., Asner G. (2002): Soil, climate, and management impacts on regional wheat productivity in Mexico from remote sensing. Agric. For. Meteorol. 114: 31–43.

Loyce C., Meynard J.M., Bouchard C., Rolland B., Lonnet P., Bataillon P., Bernicot M.H., Bonnefoy M., Charrier X., Debote B., Demarquet, T., Duperrier B., Félix I., Heddadj D., Leblanc O., Leleu M., Mangin P., Méausoone M., Doussinault G. (2008): Interaction between cultivar and crop management effects on winter wheat diseases, lodging, and yield. Crop protection, 27(7): 1131-1142.

Macholdt J., Honermeier B. (2017): Yield Stability in Winter Wheat Production: A Survey on German Farmers' and Advisors' Views. Agronomy, *7*(3): 45.

Mercer P.C., Ruddock A. (2005): Disease management of winter wheat with reduced doses of fungicides in Northern Ireland. Crop Prot. 24: 221-228.

Montesino-San Martína M., Olesen J.E., Porter J.R. (2014): A genotype, environment and management (GxExM) analysis of adaptation in winter wheat to climate change in Denmark. Agric. For. Meterol. 187: 1-13.

Reidsma P., Ewert F., Lansink A.O., Leemans R. (2010): Adaptation to climate change and climate variability in European agriculture: The importance of farm level responses. Eur. J. Agron. 32: 91-102.

Roel A., Firpo H., Plant R.E. (2007): Why do some farmers get higher yields? Multivariate analysis of a group of Uruguayan rice farmers. Computers and Electronics in Agriculture 58(1): 78-92.

Rozbicki J., Ceglinska A., Gozdowski D., Jakubczak M., Cacak-Pietrzak G., Mądry W., Golba J., Sobczynski G., Studnicki M., Drzaga T. (2015): Influence of the cultivar, environment and management on the grain yield and bread-making quality in winter wheat. J. Cereal Sci. 61: 126-132.

Sileshi G., Akinnifesi F.K., Debusho L.K., Beedy T., Ajayi O.C., Mong'omba S. (2010): Variation in maize yield gaps with plant nutrient inputs, soil type and climate across sub-Saharan Africa. Field Crops Research 116(1-2): 1-13.

Tittonell P., Shepherd K.D., Vanlauwe B., Giller K.E. (2008): Unravelling the effects of soil and crop management on maize productivity in smallholder agricultural systems of western Kenya—an application of classification and regression tree analysis. Agric. Ecosyst. Environ. 123: 137-150.

Topal M., Aksakal V., Bayram B., Yağanoğlu A.M. (2010): An analysis of the factors affecting birth weight and actual milk yield in Swedish red cattle using regression tree analysis. J. Anim. Plant Sci. 20: 63-69.

Zhang J., Liu Q., Xu M., Zhao B. (2012): Effects of soil properties and agronomic practices on wheat yield variability in Fengqiu County of North China Plain. African Journal of Agricultural Research 7(11): 1650-1658.

Zheng H., Chen L., Han X., Ma Y., Zhao X. (2010): Effectiveness of phosphorus application in improving regional soybean yields under drought stress: A multivariate regression tree analysis. African Journal of Agricultural Research *5*(23): 3251-3258.

Zheng H., Chen L., Han X., Zhao X., Ma Y. (2009): Classification and regression tree (CART) for analysis of soybean yield variability among fields in Northeast China: the importance of phosphorus application rates under drought conditions. Agric. Ecosyst. Environ 132: 98-105.