

The application of various software programs for mapping yields in precision agriculture

*Miroslaw Zagórda, Maria Walczykova**

University of Agriculture in Krakow,
Faculty of Production and Power Engineering

Abstract. Yields mapping can be done in GIS programs as well as in specialized software for agriculture. These can be programs designed for preparing informative and application maps and also software designed for integrated management of agricultural production that have a built-in GIS module. In the presented study, representatives of those three groups were taken into consideration – QGIS, Agro-Map, and FarmTMWorks Office (now Trimble Ag Software). The aim of the study was to evaluate the compatibility of the obtained yield maps in terms of basic statistical and geostatistical parameters, such as the average yield, the coefficient of variability, placement on the field zones with different yield ranges, and their percentage share in the area of the field surface. Calculations and visualization of the results were done for winter wheat, winter rape, and corn, cultivated in three-year rotation on one field. All three presented programs allow preparation of raw yield monitoring data so that yield maps produced with identical assumptions of interpolation parameters differ slightly. The difference in the use of these applications is mainly due to the further use of the resulting maps, the time required to learn the program, the task execution time, and the cost of purchase.

1 Introduction

In decision making and managing a farm, it is important to have knowledge about spatial variability of agricultural production and relationships between different phenomena. The basis of spatial dependence is the assumption that “everything is related to everything else, but near things are more related than distant things” [1], and this is fundamental for precision agriculture. One of the most widely used precision farming technologies is yield mapping [2]. Yield maps are a key component due to their usefulness in both the development and evaluation of precision management strategies [3]. These maps contain a wealth of information and can be an important tool for making informed decisions on paddock management [4]. Yield sensors in combined harvesters provide a cheap and readily available source of yield data for researchers, providing more data, and improving the effectiveness of yields as a measure for description of management zones [5]. Yield maps provide feedback information on how crops respond to certain soil and crop management practices and are used to determine recommendation rates for many inputs. Crop yield is used as an important integrator of varying crop and soil variables, such as

nutrients, moisture, and pest problems. This establishes the grounds for managing farmlands based on crop and attributes [6]. Crop yield maps reflect stable yield patterns and annual random yield variation. Therefore, a single-year yield map is useful for the interpretation of possible causes of yield variation but may be of limited value for more strategic site-specific management decisions over medium- to long-term periods [7].

Yield monitor data contain systematic and random sources of measured yield variation, including more stable yield variability related to climate and soil-landscape features, variable management-induced yield variability, and measurement errors associated with the yield mapping process itself. Management induced variation includes random events that typically occur in small patches, such as planter skips, poor crop establishment, non-uniform fertilizer application, herbicide damage, lodging, or pest damage. The most significant measurement errors associated with raw yield datasets include grain flow and moisture sensor errors, inappropriate GPS recordings and combine movement, yield surges, operator errors, unknown crop width, other outlying values, and also data processing errors [2,4]. Based on the review of the literature, Arslan and Colvin [6] provided a classification of sources of errors in yield

*Corresponding author: rtwalczy@cyf-kr.edu.pl

data into four categories: sensor errors, errors due to operating conditions, operator errors, as well as data reduction and yield mapping errors. Researchers have reported that 10 to 30% [8] or even to 50% [3] of the observations in a given field contain significant errors and should be removed.

Methods for removing of these outliers from the raw yield data have not been standardized, although many different filtering techniques have been suggested to address specific error types. Sudduth and Drummond [3] describe several separate filtering techniques implemented in Yield Editor to remove measurements that are not consistent with the actual crop. Using these techniques, the following amendments can be done: correction for the transport time between the location where the crop is harvested and the location where flow rate is sensed; removing observations when entering and exiting each transect; removing data points collected at greater velocities and less than the specified limit; eliminating data where rapid velocity changes have occurred from one point to the next point along a transect; and deleting points with above and below the yield thresholds, removing data points with swath width readings lower than the minimum.

Dobermann et al. [7] processed raw data through a cleaning algorithm that deleted the following erroneous values: header status up, start and end pass delays (8 s) for both headlands and stop-and-go segments within the field, short segments (<12 data points), frequency distribution outliers (values outside the mean \pm 3 SD), co-located yield records caused by global positioning system (GPS) drift, and local neighbourhood outliers. According to Gozdowski et al. [9], data cleaning of outliers is usually a method allowing the identification of erroneous observations for the whole set of data at once. Most often, the method based on the distribution presented as histogram is used. Nevertheless, the above cited authors found that, in order to get more reliable yield maps, the classical statistical method should be aided by the spatial statistics using Moran's I autocorrelation coefficient.

Yield maps can be processed in commercial or free GIS programs (e.g., QGIS, GRASS, or ArcView GIS), in commercial applications provided by the yield monitor producers for preparing not only yield maps but also informative and application maps for performing variable rate application of some agricultural supplies (for instance Agro-Map) or in programs designed to support the management of farm production processes like currently available 365FarmNet, FarmTMWorks Office (now Trimble Ag Software), Agrar Office, SGIS Agronomy, SMSTMSoftware, etc. that have an in-built GIS module. As far as the GIS programs are concerned, a great deal of expertise is required, and these applications are typically used by companies providing services to agricultural producers in the area of precision farming. The other can be used directly by agricultural producers. Programs built as an integrated system of information on agricultural production space, having a built-in GIS module, usually contain several interconnected modules for field data, field operations, used machinery, supplies, machinery maintenance,

employment, storage, business economics, soil mapping, yields, and proposals of field operations with possibility of processing application for different technological processes like fertilization, liming, planting etc. [10-11].

The aim of the study was to evaluate the compatibility of maps processed from the same raw yield data in three programs - QGIS, Agro-Map, and FarmTMWorks, in terms of basic statistical and geostatistical parameters and the fragmentation, i.e. the percentage share of zones with different yield ranges in the area of the field surface. The data were processed in the way as an ordinary user would do it, according to the method of processing indicated in each program. Methods of possible data cleaning and the choice of interpolation parameters in those three applications were presented as well as the use of created yield maps as a basis for preparing application maps for mineral fertilization. Calculations and visualization of the results were done for winter wheat, winter rape, and corn, cultivated in three-year rotation on one field.

2 Material and methods

The data for yield mapping were recorded for 3 years in a field of 19 hectares. The harvest was carried out with CLAAS Lexion 430 combine harvester equipped with an optical crop monitoring system. Data were recorded by the system at a frequency of 0.2 Hz at the average speed of 5.21 km·h⁻¹, which resulted in the average distance between measurement points of 7.24 m. The spacing of points between successive runs resulted from the working width of the harvesting unit, and in the case of wheat and oilseed rape, it was, on average, 5.15 m (with a header width of B=5.2 m) and 4.46 m for corn (with a header width of B=4.5 m). The procedure for removing erroneous measurements was based on the assumption that the zero yield and the maximum yields, which exceed the potential yielding of the examined plant, would be discarded, taking into account the variety, the year of cultivation, and the site. The results of post-registration variety testing of COBORU (Research Center for Cultivar Testing) were applied and the following maximum limit values were assumed: for winter wheat, 12 t·ha⁻¹ (for 14% m.c.) [14], winter oilseed rape 6 t·ha⁻¹ (for 8 % m.c.) [15], and corn 16 t·ha⁻¹ (for 14% m.c.) [16]. Filtration of other parameters such as working speed, mass flow, and grain moisture content depends on the capabilities of the used computer application. On the basis of the data prepared that way, basic statistics such as the average yield, standard deviation, coefficient of variation, and percentage share of zones with different yield ranges in the area of the field surface were determined.

3 Results and discussion

3.1. Processing the yield data in QGIS program

The yield data is first read from the monitoring device mounted on the combine by means of a simple view program that the manufacturer provides to the user free

of charge. In case of the presented research, it was Agro-Map program, however, in a full version. In the QGIS, there is virtually no format on the list of spatial data formats that can be imported into the program in which the files are created in combine harvesters of different manufacturers, unless it is a file with a *.shp* extension. Therefore, it was necessary to convert the file and this was made in the MS Excel spreadsheet by adapting the data format to the requirements of the QGIS program. To remove the outliers, i.e. zero and unrealistic values of yields, histograms were made and the yields to the value of one standard deviation as well as exceeding the potential yield of the given crop were discarded (Fig. 1). For the data prepared that way, basic statistics was determined (Table 1). On the basis of this data, it can be stated that the most erroneous measurements took place during winter rape harvesting. Removal of the incorrect measurements resulted in a decrease in the yield variation coefficient, calculated from individual measurements, by approximately 46%, 44%, and 32%, respectively, for wheat, rape and maize. The Inverse Distance Weighted (IDW) interpolation method (Fig. 2) was used to prepare yield maps, which, besides kriging, is the most commonly used.

Five equal yield ranges that are considered optimum were chosen [12], and the values of interpolation parameters have been set. The interpolation parameters for modification are the power factor that determines the effect of the values and distances of the neighbourhood points on the interpolation results and the number of nearest points to the centre of the raster model cell [13]. Decisive in determination of these parameters is the combine working width, speed, and the frequency of yield data registration. The power factor was assumed to be 2, as a standard value [13]. On the basis of the prior information about the spacing of the measurements points, the pixel size of 10x10 m was used for interpolation, which corresponded to about 2 working widths of the harvester. At the surface of one pixel, there were about 4 measurement points. Based on the maps created (Fig. 2), the percentages of areas with different yield ranges (Table 2) were determined. As can be seen on the maps (Fig. 2), two yield areas with similar percentages were prevailing in winter wheat and oilseed rape, while in corn one yield range dominated with a high, almost 78% share, indicating very well-balanced yields.

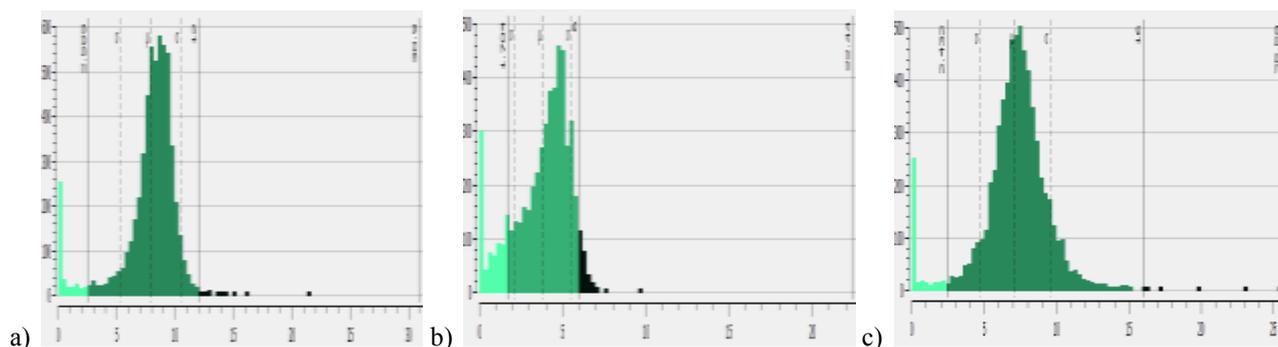


Fig. 1. Histograms for: a) winter wheat, b) winter rape, c) corn.

Table 1. Descriptive statistics calculated from individual observations – results from QGIS program.

Specifications	Total number of observations	Removed observations %	Average yield $t \cdot ha^{-1}$ based on		Standard deviation $SD t \cdot ha^{-1}$ based on		CV %	
			total observ.	observ. cleaned from errors	total observ.	observ. cleaned from errors	total observ.	observ. cleaned from errors
Winter wheat	5563	7.55	7.852	8.300	2.589	1.455	32.97	17.53
Winter rape	5086	18.93	3.797	4.197	1.704	1.061	44.88	25.32
Corn	6437	5.75	7.126	7.453	2.432	1.743	34.13	23.39

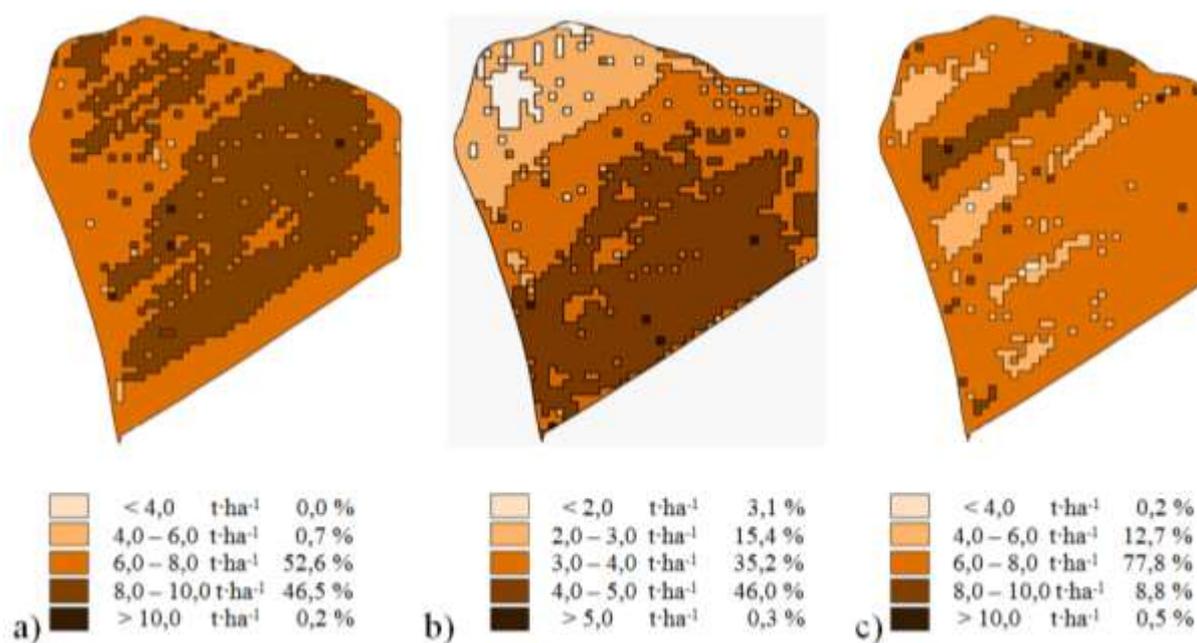


Fig. 2. Maps of the spatial distribution of yield prepared in QGIS: a) winter wheat, b) winter rape, c) corn.

Table 2. Percentage share of areas with different yield ranges – based on yield maps in Fig. 2, processed in QGIS program.

Specifications	Yield ranges (numbered according to the maps - from the smallest to the largest values)				
	1	2	3	4	5
Winter wheat	0.0	0.7	52.6	46.5	0.2
Winter rape	3.1	15.4	35.2	46.0	0.3
Corn	0.2	12.7	77.8	8.8	0.5

3.2. Processing yield data in Agro-Map

The raw yield data for this application may come from practically all currently used yield monitoring devices. Agro-Map program offers a kriging model for the creation of a contoured map, which is best for the cartographic representation of analyses. It provides a function that examines raw data after selection of the interpolation procedure. The yield values are organized into classes and the frequency limit, below which it is assumed that values are erroneous, is set at 3%, which proceeds invisibly [17]. When starting interpolation, the user assumes the resolution of the map (cell size), creates a legend, and introduces interpolation parameters (search radius, minimum, and maximum neighbours). Taking into account the recording parameters presented in the description of the work in the QGIS program, there were about 4 measurement points on the surface of 10 x 10 m. Consequently, it was reasonable to assume 20 m of the search radius and the number of adjacent points at the level of 16, i.e. interpolation was made for 3 runs on the distance of about 20 m. During map creation, the program proposes a range of yield values, without taking into account zero and unreasonable high yields. These suggested value limits can be altered at will if they do

not appear sensible. In the dialog window, the moisture content of the harvested crop should also be corrected during the interpolation. Removing erroneous values caused by other factors (too small or too high working speed, zero moisture content of grain) before the interpolation is possible only manually in the source file, where the recorded data are in the form of a table. Table 3 presents the basic statistics computed for the two data cleaning variants - according to program hints and after manual removal of the measurements mentioned above. A method of removing erroneous measurements other than those used in QGIS application resulted in minor differences in mean yields as well as in the coefficient of variation. The maps in Figure 3 refer to the data cleaned from all measurements, which were considered to be incorrect in connection with the above-mentioned rules. Table 4 presents the percentages of areas with different yield ranges for both variants. As a consequence of the results included in Table 3, the differences between yields are insignificant in the variants, but this time, there is a greater fragmentation of the field than in the QGIS program (Table 4).

Table 3. Descriptive statistics calculated from individual observations – results from Agro-Map application.

Specifications	Removed observations %		Average yield t·ha ⁻¹		Standard deviation SD t·ha ⁻¹		Coefficient of variability CV %	
	(A) suggested by the program	(B) without all zero values	(A) suggested by the program	(B) without all zero values	(A) suggested by the program	(B) without all zero values	(A) suggested by the program	(B) without all zero values
Winter wheat	4.82	5.27	8.136	8.143	1.812	1.785	22.27	21.92
Winter rape	9.24	9.79	3.889	3.896	1.393	1.390	35.82	35.68
Corn	6.09	6.38	7.399	7.397	1.659	1.657	22.42	22.40

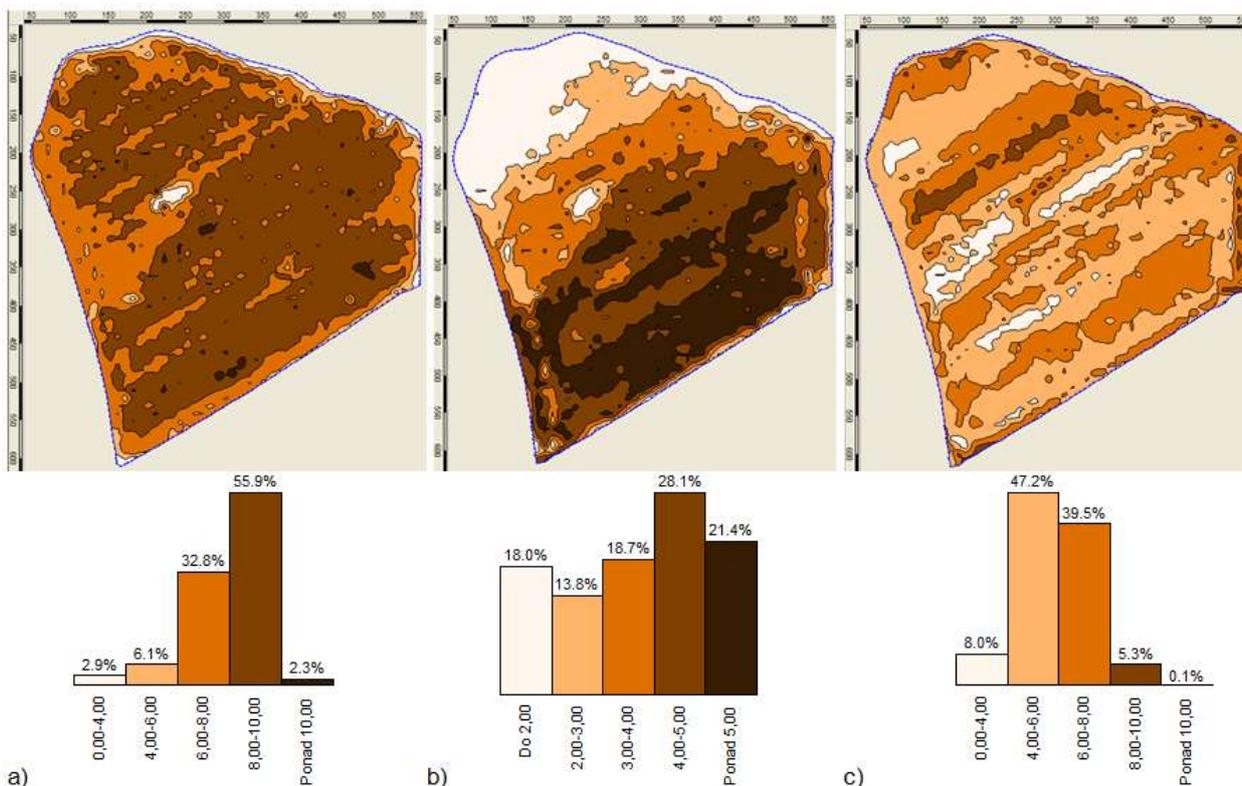


Fig. 3. Maps of the spatial distribution of yield with legend, prepared in Agro-Map: a) winter wheat, b) winter rape, c) corn.

Table 4. Percentage share of areas with different yield ranges – based on yield maps in Fig. 3, processed in Agro-Map application.

Specifications	Yield ranges (numbered according to the maps - from the smallest to the largest values)									
	1		2		3		4		5	
	(A)	(B)	(A)	(B)	(A)	(B)	(A)	(B)	(A)	(B)
Winter wheat	2.9	2.8	6.1	6.2	32.8	32.8	55.9	55.9	2.3	2.3
Winter rape	18.0	16.1	13.8	13.7	18.7	17.2	28.1	25.1	21.4	27.9
Corn	8.0	8.0	47.2	47.2	39.5	39.4	5.3	5.3	0.1	0.1

3.3. Processing data in FarmTMWorks Office

As in the Agro-Map application, raw yield data for mapping can come from practically all currently used crop monitoring devices (AGCO, AgLeader, Case IH, Claas Lexion, New Holland, Trimble, etc.). The first step in the mapping procedure is to discard the zero values of all parameters given in the filter by entering the minimum value of 0.1 and entering the maximum

possible yield value for the given plant (Fig. 4a). After critical evaluation of the remaining parameters in the displayed window (Fig. 4b), further adjustments are made to the filter (Fig. 4c). As a minimum yield, the standard deviation is recommended. Table 5 includes values of the basic statistics obtained in this program, and they are similar to the results obtained in the QGIS program (Table 1). Both methods of interpolation proposed by the program, i.e. AVERAGE (AVG) and

IDW, were used for the mapping. The maps shown in Figure 5 were prepared by the IDW method, using interpolation parameters identical to those of QGIS. In order to show the impact of the interpolation technique on the fragmentation of the resulting yield map, i.e. the percentage share of areas with different yielding ranges, Table 6 also includes the results obtained by the AVERAGE method. It has been shown that the

differences are minimal. In order to obtain more continuous boundary lines of homogeneous areas in the field, the edge smoothing function (possible in the range from 0 - 100%) and the function of removing small area polygons (size can be provided) were used. Figure 6 shows the results for yield maps at 50% smoothing, and removed areas less than 0.1 ha, and Table 7, the figures for both methods of interpolation.

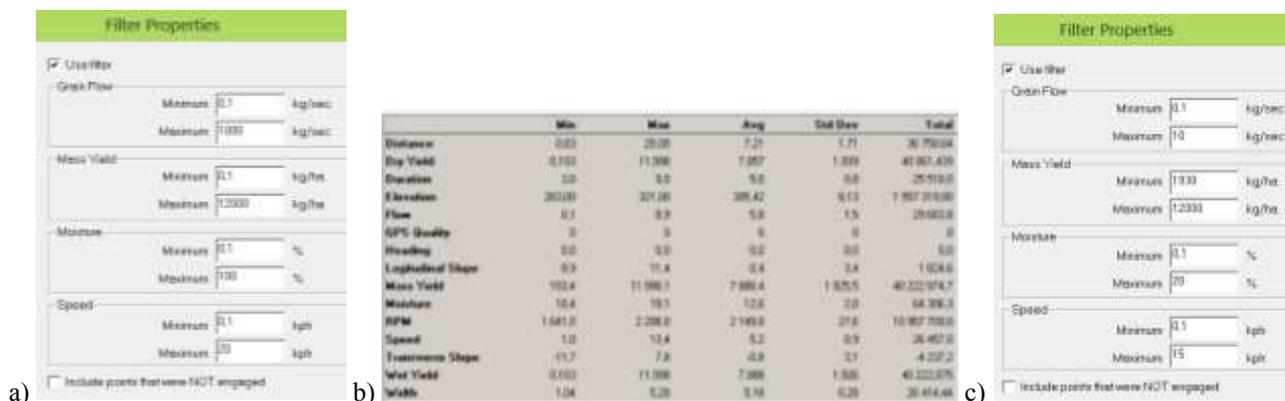


Fig. 4. Filtering data in tFarm™Works Office: a) filter pre-setting, b) statistics of the filtered data, c) correction of the filter settings.

Table 5. Descriptive statistics calculated from individual observations – results from Farm™Works Office.

Specifications	Removed observations %	Average yield t·ha ⁻¹	Standard deviation SD t·ha ⁻¹	Coefficient of variability CV %
Winter wheat	10.9	8.04	1.57	19.5
Winter rape	17.4	4.15	1.10	26.5
Corn	6.01	7.32	1.79	24.5

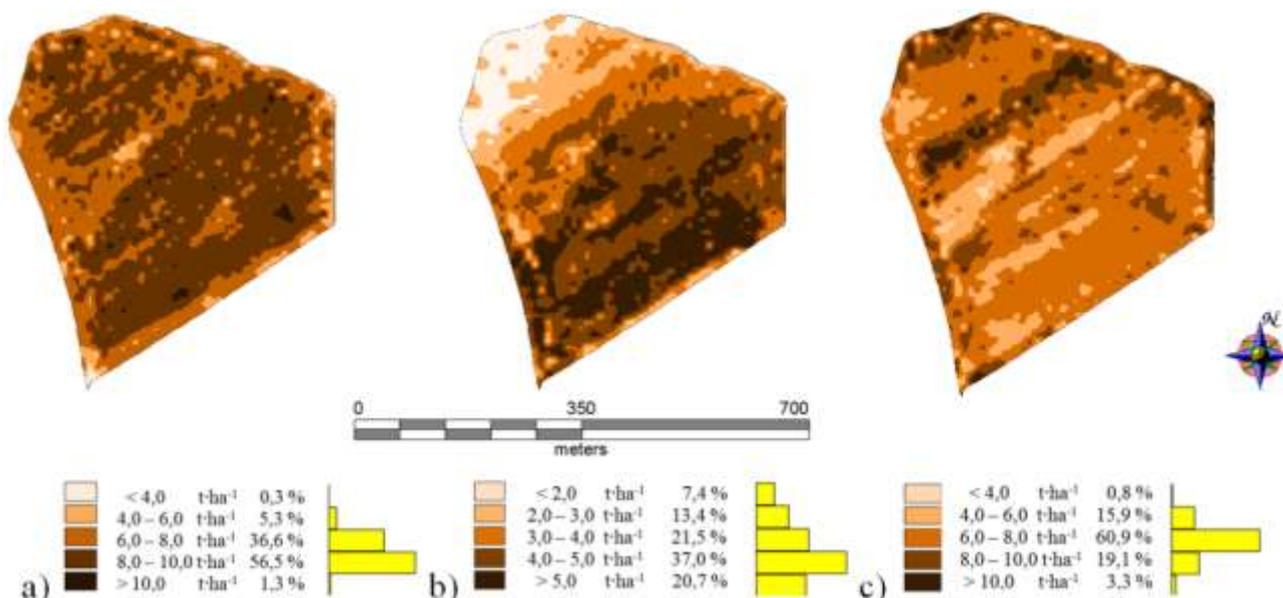


Fig. 5. Maps of the spatial distribution of yield with legend, prepared in Farm™Works Office: a) winter wheat, b) winter rape, c) corn.

Table 6. Percentage share of areas with different yield ranges – based on yield maps in Fig. 5, processed in FarmTMWorks Office.

Specifications	Yield ranges (numbered according to the maps - from the smallest to the largest values)									
	1		2		3		4		5	
	AVG	IDW	AVG	IDW	AVG	IDW	AVG	IDW	AVG	IDW
Winter wheat	0.3	0.3	4.5	5.3	36.7	36.6	57.4	56.5	1.1	1.3
Winter rape	7.8	7.4	13.2	13.4	20.8	21.5	37.3	37.0	20.9	20.7
Corn	0.5	0.8	14.5	15.9	63.2	60.9	19.2	19.1	2.6	3.3

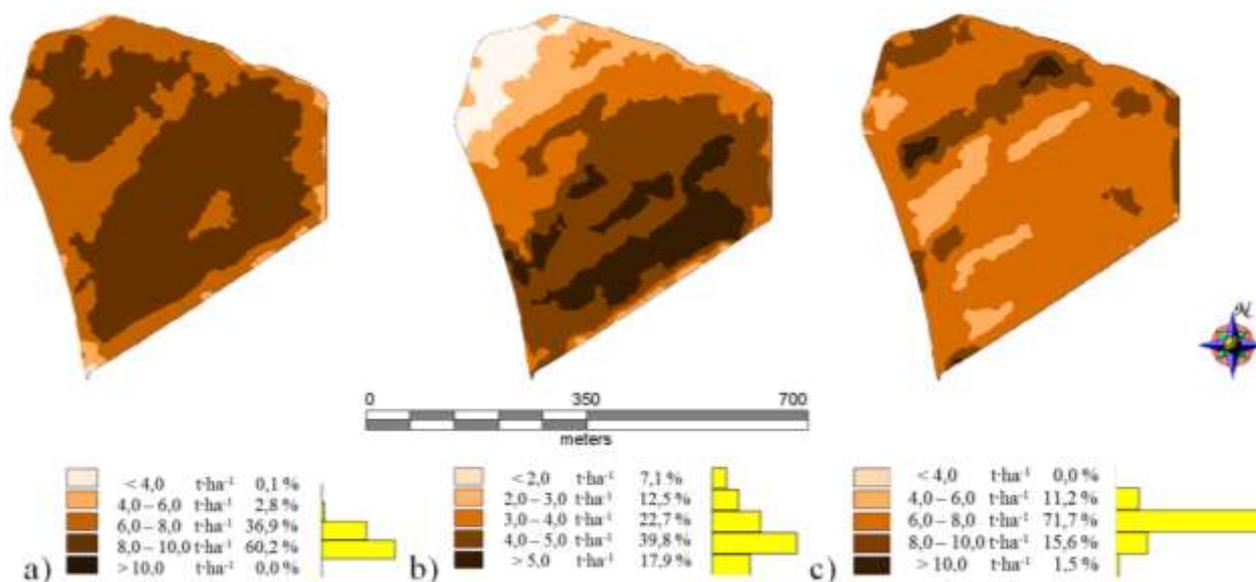


Fig. 6. Maps of spatial distribution of yield with legend, prepared in FarmTMWorks Office (smoothed 50%, with removed fragments smaller than 0.1 ha): a) winter wheat, b) winter rape, c) corn.

Table 7. Percentage share of areas with different yield ranges – based on yield maps in Fig. 6, processed in FarmTMWorks Office (smoothed 50%, with removed fragments smaller than 0.1 ha).

Specifications	Yield ranges (numbered according to the maps - from the smallest to the largest values)									
	1		2		3		4		5	
	AVG	IDW	AVG	IDW	AVG	IDW	AVG	IDW	AVG	IDW
Winter wheat	0.1	0.1	2.4	2.8	38.0	36.9	59.5	60.2	0.0	0.0
Winter rape	7.9	7.1	11.9	12.5	21.9	22.7	39.6	39.8	18.7	17.9
Corn	0.0	0.0	11.5	11.2	71.0	71.7	16.3	15.6	1.2	1.5

3.4. Use of yield maps for creating application maps for fertilization

The simplest method of preparing application maps is to prepare them on the basis of the previously made map of yield or soil nutrients content. On the basis of the prepared yield maps in all the discussed programs, it is possible to make application maps for mineral fertilization. One way to plan for mineral fertilization is to use the method of returning to the soil nutrients taken by crop that is based on information concerning yield (grain and straw) and the amount of mineral content contained therein [18] [19]. In the case of yield maps prepared in the presented programs, it is possible to assign the amount of taken nutrients calculated with the methods well-known in agronomy to the respective yield ranges. The yield map is converted into a fertilizer dosing (application) map that is exported to

the spreader control panel. For technical reasons related to the application of fertilizers, application maps should be adapted to the equipment in terms of resolution (usually it is the spreader's working width), the format (file format supported by the dose control system), and the minimum size of polygons required for the variable rate application (e.g., the minimum size of the polygon in the field at the spreader with a working width of 20 m will be 20 m x 20 m, i.e. 0.04 ha). In FarmTMWorks Office and the QGIS, it is also possible to process data in such a way (smoothing out contours, removing small areas) so that the resulting yield map can be converted into the application map of fertilization by adding the appropriate fertilizer dose attribute to the particular yield ranges. The yield map prepared in Agro-Map with the use of the kriging method contains many small areas and requires

additional processing by the user, which involves manually removing small areas on the application map. After that, the calculated fertilizer doses are assigned to the respective yielding ranges on the map.

5 Closing remarks

All three presented programs enable preparation of raw yield monitoring data so that yield maps produced with identical assumptions of interpolation parameters differ slightly. The difference in the use of these applications is mainly due to the further use of the resulting maps, the time required to learn the program, task execution time, and cost of acquisition. Of the three presented applications, it is only possible to add more layers of information on the field in order to determine the management zones as accurately as possible in the QGIS program. The Agro-Map application enables smart soil sampling on the soil nutrient content based on the created yield maps and preparing application maps for variable rate mineral fertilization as well as other supplies (seeds, herbicides, etc.). FarmTMWorks Office software also enables assessment of the economic efficiency of the areas with different yields. In each of these cases, a significant personal effort is required to learn how to handle these programs, but the most demanding is the QGIS application. The time of preparation of the yield map, including the upload of the original data, when carried out the way presented in the study, lasts from 15 to 70 minutes for the expert user in the program - the least in the Agro-Map program, the most in the QGIS program. As far as purchase costs are concerned, the QGIS is free and the highest acquisition cost is related to FarmTMWorks Office software, nevertheless, as mentioned previously, mapping is just one of the many farm management support modules that this program includes.

References

1. H. J. Miller, *Annals of the Association of Am. Geogr.* 94 (2), 284-89 (2004)
2. J.L. Ping, A. Doberman, *Prec. Agric.* 6, 193-212 (2005)
3. Sudduth K.A., Drummond S.T. *Agron. J.* 99:1471-1482 (2007)
4. T.P. Robinson, G. Metternicht, *Agric. Syst.* 85, 19-41 (2005)
5. S. Nawar, R. Corstanje, G. Halcro, D. Mulla, A.M. Mouazen, *Adv. in Agron.* 175-244 (2017)
6. S. Arslan, T. Colvin, *Prec. Agric.* 3, 135-154 (2002).
7. A. Dobermann, J.L. Ping, Y.I. Adamchuk, G.C. Simbahan, R.B. Ferguson, *Agron. J.* 95, 1105-1120 (2003)
8. Y. Rong-Cai, H. Zhinqui, *ASA-CSAA-SSSA Annual Meeting*, Long Beach, CA (2014)
9. D. Gozdowski, S. Samborski, E.S. Dobers, *Colloq. Biometr.* 40, 41-51 (2010)
10. V. Rataj, J. Galambošová, M. Macák, L. Nozdrovický, *Profi Press SK* (2014)
11. M. Walczykova, M. Zagórda, *Polish Acad. of Arts and Sci. Commiss. on Agric., Forest. and Vet. Sc.* 18, Monograph Vol. 22, 306-314 (2016)
12. R. Grisso, M. Alley, S. Phillips, Ph. McClellan, *Virginia Coop. Exten., Precision Farming*, Publication 442-500 (2002)
13. R. Pudełko, *Pamiętnik Puławski (Polish Journal of Agronomy)*, z. 146: 99-111 (2007)
14. Zboża ozime (Winter grain) 2016. <http://www.coboru.pl/DR/pdopublikacjecentra lne.aspx>
15. Rzepak ozimy (Winter rape) 2016. <http://www.coboru.pl/DR/pdopublikacjecentra lne.aspx>
16. Kukurydza (Corn) 2016. <http://www.coboru.pl/DR/pdopublikacjecentra lne.aspx>
17. *Agro-Map Manual* (2004)
18. S. Mercik, *Wydawnictwo SGGW, Warszawa (Publishing House Warsaw University of Life Sciences)*, pp. 288. (2004)
19. Kopeć M., Gondek K., *Materials for training in fertilizer advisory*, Kraków (2002)