



Crop raw yield monitor data.

# Yield Mapping and Use of Yield Map Data

**Dave Franzen**, NDSU Extension Soil Specialist

**Francis Casey**, Associate Professor

**Nathan Derby**, Research Associate

## Yield maps...may be predictive of yield potential and soil nutrient variability for future crop management decisions

Combine yield monitors are growing in popularity. Data are utilized to build field yield maps in certain years or yield frequency maps during multiple years. Maps serve as location-year record of management and may be predictive of yield potential and soil nutrient variability for future crop management decisions.

### Yield Monitors

Yield monitors are available for purchase from the combine manufacturer or an independent yield monitor manufacturing company. Most U.S. grain yield monitors measure grain flowing through the clean grain auger into the hopper and grain moisture continuously in the grain flow. Data generated with a yield monitor are only as good as the correct installation, calibration and maintenance of the unit and its components.

Even though yield monitor manufacturers normally sell computer software to help analyze yield data, growers may want to conduct their own analyses using commonly available data management software and geographic information software (GIS) programs. Being able to export your data into useful files is important in analyzing yield data in data management software. Choose a yield monitor capable of exporting data in .txt format because this format is imported easily into most database spreadsheets and GIS software programs.

A number of yield monitor data errors are associated with each data set generated within a field. These include partial combine header passes, the lag between the beginning of grain cut and when the monitor measures the grain yield, and the time the header takes to move from the down position during the completion of a pass to the up position that terminates the data flow.

The grain flow within the combine may bunch up and spurt on occasion instead of flowing continuously.

Considerable small-scale variability occurs within any yield monitor data set because of these and other sources of error. One of the easiest data-cleaning operations is to eliminate outliers in the data set using a spreadsheet, such as Microsoft Excel. Outliers are data that, because of gaps or spurts in the

For a current list of suppliers and their links, see <http://mpac.missouri.edu/links/yieldmonitor.htm>

**For tips on correct installation and operation of yield monitors, see the manufacturer's recommendations and these suggested publications:**

**"Elements of precision agriculture: Basics of yield monitor installation and operation,"** by Shearer et al., publication PA-1, University of Kentucky. <[www.bae.uky.edu/precag/PrecisionAg/Exten\\_pubs/pa1.pdf](http://www.bae.uky.edu/precag/PrecisionAg/Exten_pubs/pa1.pdf)>

**"Precision Farming Tools: Yield Monitor"** by Grisso et al., publication 442-502, Virginia Cooperative Extension. <[www.ext.vt.edu/pubs/bse/442-502/442-502.html](http://www.ext.vt.edu/pubs/bse/442-502/442-502.html)>

**NDSU**  
**Extension Service**  
North Dakota State University

Fargo, North Dakota 58108

DECEMBER 2008

grain stream, are unreasonably low or high yields. For example, in a field of spring wheat that varies from 20 to 50 bushels per acre (bu/acre), some yield points may be as high as 150 bu/a. These actual yields most likely do not exist, so choosing a more reasonable maximum yield is desirable.

In fields with soluble salt problems, zero yield is possible. Cleaning the lower end of the yield sort is more difficult. In this case, the area of actual zero yield would need to be identified within the data set and only those data not within the problem area should be considered for cleaning.

In fields where actual grain yields are not zero, choosing some lower limit of yield also would be recommended. A data sort from high to low within a spreadsheet then can be conducted, and yields above and below the limits can be deleted from the data set. Some precision-ag consultants have their own proprietary cleaning software. A public arena download of yield data-cleaning software is available from the USDA-ARS in Missouri (Suddoth, 2007) at [www.ars.usda.gov/services/software/download.htm?softwareid=20](http://www.ars.usda.gov/services/software/download.htm?softwareid=20).

Since small spatial errors occur among yield measurements, yield monitor data usually are displayed as maps in larger zones of varying yield ranges. These yield maps are useful to producers because they visually show that yields vary throughout the field and where the yields are high, low and in between without the clutter of individual dots that vary widely in value. For visual display, choosing yield increments that make sense is helpful. For example, in a corn field, choosing display ranges of 30 bu/acre in a field that averages 200 bu/acre is probably appropriate. Displaying ranges of 10 bu/acre increments would create a map that would be too “busy” and hard to interpret. In contrast, a spring wheat field that averaged 40 bu/acre would need increments of about 10 bu/acre or less to be meaningful. Thirty bu/acre increments may not show important variation from the average.

Although the yield map itself is a useful tool that illustrates the management and climate results for the current growing season, yield maps also may be predictive of future yield potential and might be related to the availability of certain soil nutrients, such as nitrogen. Single-year yield maps for predictive purposes have not been nearly as useful as multiyear yield maps. A multiyear yield map also is called a “yield frequency map.”

Whether a field has had a history of a single crop or a diverse crop rotation, the same general procedure should be followed to create a yield frequency map. A field that has been in wheat continuously for 10 years might average 80 bu/acre one year and 20 bu/acre another year. The actual bushels for the field therefore cannot be used when the data sets are combined. If the field was corn one year, soybeans the next, wheat the next and then sunflowers the year after, these yields obviously cannot be added to each other spatially with any meaning. The range of yields in any year therefore must be “normalized.”

Normalization is a simple mathematical exercise that converts bu/acre into relative yield. In the example year of high wheat yield with highest cleaned yield of 80 bu/acre, divide each yield by 80. The range of yields is therefore from 0 to 1. If the next year is canola and the highest canola yield is 3,500 pounds/acre, divide each yield by 3,500. The range of yields is from 0 to 1.

Another way to normalize yield data is to normalize it at the end of analysis and not at the beginning. Using the following figures as examples, first impose a grid on the yield data that makes some sense. A field should have at least 40 grids to produce a meaningful map at the end of the exercise. For example, a user might choose to superimpose 40 1-acre grids over a 40-acre field. A quarter-section might have 2-acre grids. A 20-acre field might have ½-acre grids.

To create the grids and the average yield within a grid, use a software program such as Surfer (Golden Software Co., Golden, Colo.) or ArcGIS (ESRI GIS Software Co., Redlands, Calif.) that can import spatial data and then convert them to estimated values. This estimation feature usually is used for taking less dense data and estimating values at small distances. However, it also can be used to take densely sampled data, such as the thousands of points of yield data, and average them within a less dense grid of your choosing. In Surfer, the resulting grid file can be saved in an ASCII text file and then uploaded into a spreadsheet.

Within the spreadsheet, the grid is given a +1, -1 or 0 value, depending on whether the average of the grid is greater than the field average, less than the field average or within ½ bu/acre of the field average. Giving the +1, -1 or 0 is a normalization procedure. Then these normalized grids can be exported into a spreadsheet and summed by grid with other years of data that have been treated



20	30	25	22	23	25	28	29	30	31	32	29	28	30	40	44	45	48	47	35	30	28	26	25
25	28	23	20	24	26	30	32	33	35	40	36	29	35	38	46	47	50	52	53	43	35	28	28
28	28	26	25	28	30	34	35	38	40	40	42	38	36	40	45	48	53	45	43	38	38	36	36
30	30	32	34	35	38	38	37	40	40	43	45	46	44	45	46	48	46	45	48	49	49	45	46
32	32	33	35	35	36	38	40	42	42	44	46	44	44	38	37	40	42	41	45	47	50	48	47
34	34	35	36	36	35	33	40	32	38	35	36	40	42	43	44	45	46	38	37	40	41	42	
38	38	39	39	40	38	42	45	56	50	52	53	38	35	37	38	40	42	44	46	37	38	39	40
38	38	38	38	36	36	37	32	46	47	50	42	39	38	36	40	42	44	46	38	36	35	39	38
36	38	40	42	44	45	48	50	52	46	38	36	35	43	44	46	48	50	52	53	54	52	48	48
35	35	36	38	39	42	44	44	46	42	40	38	37	36	38	39	40	42	44	46	48	48	52	48
34	34	36	38	40	42	43	41	44	38	36	38	39	38	39	39	41	42	41	43	44	45	44	43

An example of raw yield monitor data from a field of spring wheat in bu/acre.

-1	-1	-1	+1	-1
-1	-1	+1	+1	+1
-1	-1	+1	+1	+1
-1	+1	+1	+1	-1
-1	+1	-1	+1	+1
-1	+1	-1	+1	+1

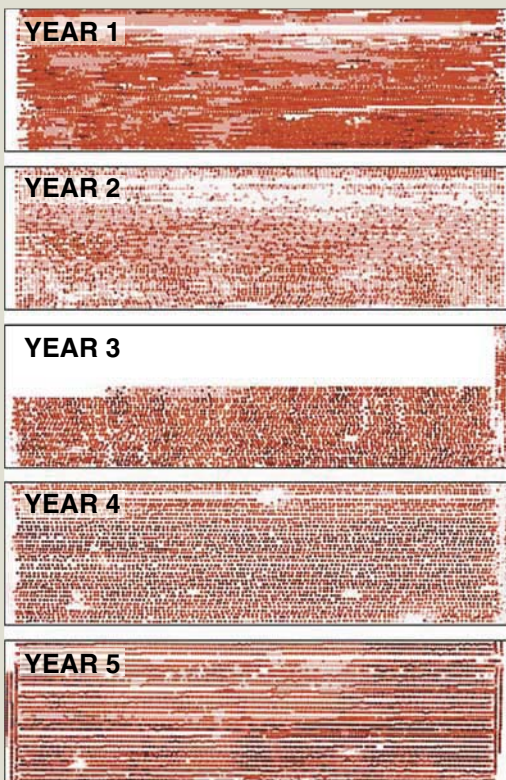
20	30	25	22	23	25	28	29	30	31	32	29	28	30	40	44	45	48	47	35	30	28	26	25
25	28	23	20	24	26	30	32	33	35	40	36	29	35	38	46	47	50	52	53	43	35	28	28
28	28	26	25	28	30	34	35	38	40	40	42	38	38	36	40	45	48	53	45	43	38	38	36
30	30	32	34	35	38	38	37	40	40	43	45	46	44	45	46	48	46	45	48	49	49	45	46
32	32	33	35	35	36	38	40	42	42	44	46	44	44	38	37	40	42	41	45	47	50	48	47
34	34	35	36	36	35	35	33	40	32	38	35	36	40	42	43	44	45	46	38	37	40	41	42
38	38	39	39	40	38	42	45	56	50	52	53	38	35	37	38	40	42	44	46	37	38	39	40
38	38	38	38	36	36	37	32	46	47	50	42	39	38	36	40	42	44	46	38	36	35	39	38
36	38	40	42	44	45	48	50	52	46	38	36	35	43	44	46	48	50	52	53	54	52	48	48
35	35	36	38	39	42	44	44	46	42	40	38	37	36	38	39	40	42	44	46	48	48	52	48
34	34	36	38	40	42	43	41	44	38	36	38	39	38	39	39	41	42	41	43	44	45	44	43

During the normalization process, a grid is superimposed over the data. The software recognizes the set boundaries for averaging the data inside each grid box.

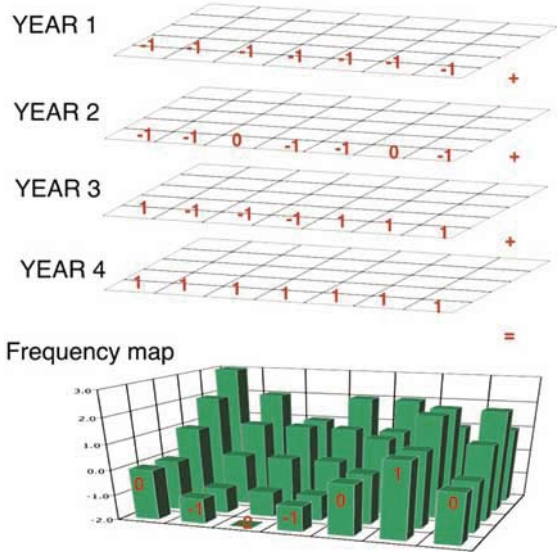
30.1	29.9	33.7	46.7	30.4
29.6	37	41.7	45.4	43
34.2	37.3	40.7	42.1	44
38.2	42.9	42	42	37.8
38.3	45.9	38.5	46	49.8
36.4	41.6	38	41.2	44

The data are averaged within each grid box independently of the other data.

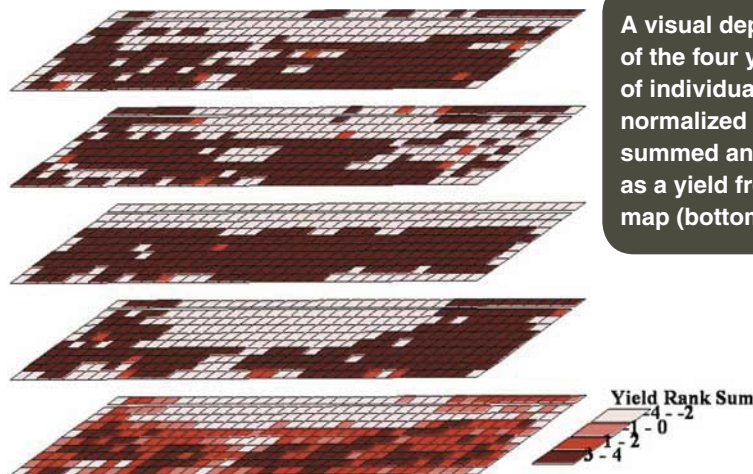
If the data average within the grid is higher than average, the grid is given a +1. If average, it is given a value of 0. If the average within the grid is lower than average, the grid is given a -1. Afterward, this data can be exported into a spreadsheet for summing multiple years of yield data that have been handled similarly. The resulting summation can be imported into mapping software to construct the yield frequency map.



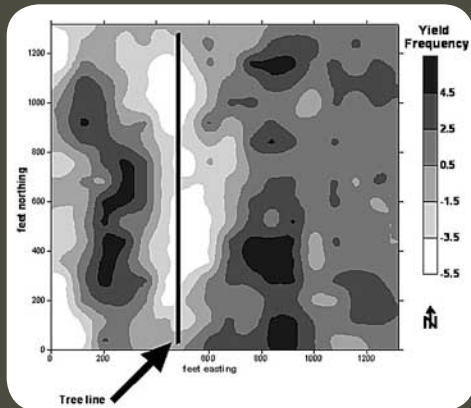
A series of five years of crop raw yield monitor data of the same field. Missing data can be normalized by weighting yields by years of data so the large area of missing data in the third yield map is divided only by 4 instead of 5, or normalized by the summation multiplied by 1.25, to compare favorably with data from five years of maps instead of the four it has available.



Each year's data are superimposed with the same grid. Yields are averaged within each grid. +1 is given to grids greater than average for the field, -1 for grids lower than the average and 0 for the rare grid that is within a half-bushel of average. These grids then are summed individually within a spreadsheet and remapped into the mapping software.



A visual depiction of the four years of individually normalized yields summed and mapped as a yield frequency map (bottom map).



A yield frequency map depicting a corn and soybean rotation in Illinois in a five-year period (from Franzen, 2008). Each individual year did not show all of these patterns. Meaningful long-term trends in yield were revealed with accompanying relationships to soil phosphorus, potassium and pH (acidity or alkalinity) levels only following the combination of years.

in the exact same manner. In this way, two years, three years or 20-plus years of data can be combined to create a more meaningful yield frequency map.

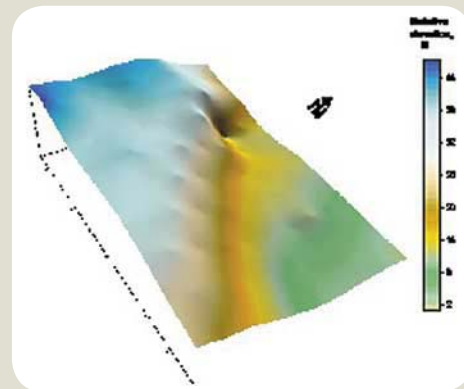
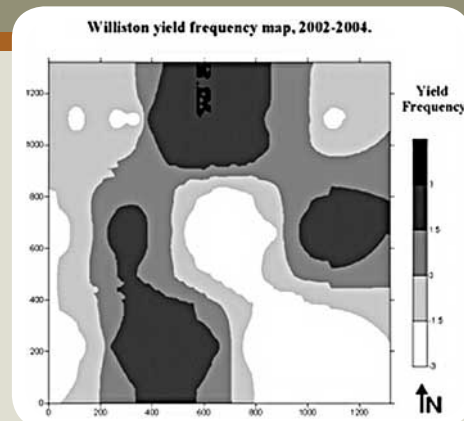
The yield frequency map has been used as a zone layer for revealing residual soil nitrate and other nutrients in North Dakota. The yield frequency map also can help reveal areas that require additional management, such as a change in nitrogen application timing or a change in drainage if permitted. It can reveal the yield drag due to hedgerows, saline spots, compacted areas or locations containing harmful levels of sodium.

The yield frequency map also can be used conceivably to predict the relative performance of farms and evaluate future production potential for renting or purchasing decisions. If the soils on a farm that you manage have a certain level of productivity, this potential productivity probably also is possible on farms within several miles of that farm. This information can help determine whether a farm is worth renting or buying.

The yield maps are a valuable archive of field performance and the changes that management might have had on the fields. These archived maps and data can be shared with bankers to help secure loans and with future renters or purchasers of the land. They can be passed down to the next generation so that the things growers learn about their fields will not be lost to the heirs.

Many growers also use a yield monitor to help develop nutrient strategies on specific farms. Different rates of nitrogen, for example, can be applied in strips across soils within a field, and the yields can help the grower determine what the best rate might be for a soil or the field. In addition, new products can be tested on a small strip within the field and the efficacy of their use

**Yield frequency map at Williston, N.D. (top), with lower yields on the ridge slope (southeast to northwest direction) and higher yield in the more level areas above and below the ridge slope (southwest and northeast) (elevation image on the bottom).**



can be much more easily seen this way than by using a weigh wagon. The many data points of each strip also lend themselves to the use of statistics offered within spreadsheets to provide confidence to the grower that the differences seen were statistically real or not. Variability within a field can mask differences or create differences unless statistics are used to help eliminate the clutter.

## References

- Suddoth, 2007. Yield editor software. USDA-ARS. [www.ars.usda.gov/services/software/download.htm?softwareid=20](http://www.ars.usda.gov/services/software/download.htm?softwareid=20)
- Franzen, 2008. Summary of forty years of grid sampling in two Illinois fields. NDSU Technical Report.

The NDSU Extension Service does not endorse commercial products or companies even though reference may be made to tradenames, trademarks or service names.

This publication may be copied for noncommercial, educational purposes in its entirety with no changes. Requests to use any portion of the document (including text, graphics or photos) should be sent to NDSU. [permission@ndsu.edu](mailto:permission@ndsu.edu). Include exactly what is requested for use and how it will be used.

**For more information on this and other topics, see: [www.ag.ndsu.edu](http://www.ag.ndsu.edu)**

County commissions, North Dakota State University and U.S. Department of Agriculture cooperating. North Dakota State University does not discriminate on the basis of race, color, national origin, religion, sex, disability, age, Vietnam Era Veterans status, sexual orientation, marital status, or public assistance status. Direct inquiries to the Vice President for the Division of Equity, Diversity and Global Outreach, 205 Old Main, (701) 231-7708. This publication will be made available in alternative formats for people with disabilities upon request, (701) 231-7881.

1M-12-08