The Data-Intensive Farm Management Project:
Using Precision Technology to Get the Information to Use Precision Technology Profitably

David S. Bullock
University of Illinois
Department of Agricultural and Consumer Economics

Invited Presentation at the InfoAg Conference
St. Louis, July 23, 2019
This research was funded in part by a USDA National Institute of Food and Agriculture Food Security Program grant, award number 2019-68004-24769.
There are dozens of these input application “decision tool” software packages on the market:
Lots of claims about how much money farmers can make by using their “decision tools” to farm data-intensively and site-specifically:
Pioneer’s Encirca Nitrogen service claims a **74% ROI increase** over the standard nitrogen management practice.
“Adapt-N has been tested and calibrated in replicated strip trials to ensure accuracy across weather, soil, and field management permutations. Independent research validates a $30-per-acre grower profit using the tools. ‘We pride ourselves on superior accuracy and validation,’ says Steve Sibulkin, CEO of ATC.”

Source: https://www.precisionag.com/systems-management/taking-nitrogen-to-the-max/
Climate Corporation stated, “Today customers experience an average savings of $12.31 per acre using the FieldView nitrogen management tool” (Eathington, 2018).

Source: https://climate.com/blog/next-evolution-of-digital-farming/
Building Seed Advisor
Pre-Commercial Seed Selection and Placement Technology

+6 bu/ac advantage*

Nearly 80% win rate*

2018 testing on 100,000 acres across 3 states

*Based on 2017 farmer field trials on 10,000 acres.
NEW!

SEE NITROGEN IN A NEW WAY!

Introducing best-in-class weather and nitrogen management tools

Maximize in-season N-decisions to adapt to field-level weather

Analyze 5 seasons of Nitrogen loss by soil texture in each field

Tap into DTN’s weather network powered by 22,500 stations

» Factoring in multiple weather conditions from your local weather station
» Demonstrated $30/acre profit increase & reduced N losses
» Powered by Adapt-N
Three basic analytical methods behind the decision-tool software: yield-based algorithms, crop growth models, empirical, regression-based models.

<table>
<thead>
<tr>
<th>Corporate Name</th>
<th>Product Name</th>
<th>Analytical Engine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monsanto/Climate Corp DuPont/Pioneer</td>
<td>FieldView™ Encirca®</td>
<td>Suite of crop and soil models</td>
</tr>
<tr>
<td>YARA ProAgrica DTN EFC Systems</td>
<td>Adapt-N SST DTN Nitrogen FieldAlytics</td>
<td>Adapt-N (crop and soil models with yield-based algorithm)</td>
</tr>
<tr>
<td>Trimble Agriculture AgLeader</td>
<td>GreenSeeker OptRx</td>
<td>Algorithms with sensor-based vegetative indices</td>
</tr>
<tr>
<td>AgLeader</td>
<td>SMS™</td>
<td>Partially yield-based: user can create own formula-based prescriptions</td>
</tr>
<tr>
<td>Nutrien Ag Solutions™ SMART Fertilizer Mgt FieldApex Midwest Reg Climate Ctr U Nebraska</td>
<td>Echelon SMART! Fert Mgt Software FieldApex U2U Split-N Maize-N</td>
<td>Yield-based, integrating crop characteristics, soil tests, and other field data</td>
</tr>
<tr>
<td>US Midwest Land-grant Universities</td>
<td>Maximum Return to Nitrogen</td>
<td>Regression-based: Regional averages from on-farm strip trials</td>
</tr>
</tbody>
</table>
No doubt many, maybe all of you CCAs use one or more of these products to provide management advice to your clients.
Today I want to talk about

• the **Data-Intensive Farm Management** research project (DIFM),

• DIFM’s relationship to current farm management decision tools,

• how conducting on-farm field trials under DIFM protocols can help your farmer-clients make better management decisions, and

• developing a cyber-infrastructure for globally-coordinated on-farm precision experimentation.
First, just a little bit of “theory”: 
To know optimal input application rates, you need to know 1) prices, and 2) the yield response function.
Of course, in real life things are a little more complicated:
Yield depends on

- Managed inputs \((x)\)
- Spatially dependent characteristics \((c)\)
- Unmanaged time-dependent variables \((z)\)

\[ y = f(x, c, z) \]
Managed Inputs

$x = (x_1, x_2, x_3, \ldots)$

- Nitrogen fertilizer (pre-plant)
- Herbicide (post-emergence)
- Equipment choices
Spatially-dependent characteristics

\[ c = (c_1, c_2, c_3, \ldots) \]

- Soil depth
- Soil texture
- Slope of ground
Unmanaged time-dependent variables

\[ z = (z_1, z_2, z_3, \ldots) \]

- July rainfall
- June temperature
- European corn borer population (possibly also spatially dependent)
Therefore, to use the AgEcon 101 picture,

- Every site in every year will have its own economically optimal input application rates:

\[
\begin{align*}
\text{Area A, 2019} & : f(x_1, c^A, z^{2019}) \\
\text{Area B, 2019} & : f(x_1, c^B, z^{2019}) \\
\text{Area C, 2019} & : f(x_1, c^C, z^{2019}) \\
\text{Area D, 2019} & : f(x_1, c^D, z^{2019})
\end{align*}
\]
The Problem:

WE DON’T KNOW THE PICTURE!
DIFM scientists have come to the conclusion that nobody knows very much about the God function (maybe God, of course).
Why not?

Lack of data
How could this be? Haven’t agronomists forever been doing experiments to figure out response functions?
Earl Heady, throughout the 50s, 60s, 70s:
Fact: Despite all this research, until fairly recently, it’s a stretch to say that university-led N fertilizer management recommendations were based on data or statistical analysis.
In the past, university fertilizer management recommendations could only be rules of thumb. Until about 2007: the “1.2 Rule”:

Profit-Maximizing N Rate?

My goal is 200 bushels per acre.

What yield do you want?

Then multiply that by 1.2!

$200 \times 1.2 = 240$ lbs N fert per acre.
Empirical work behind yield-based fertilizer recommendations often wasn’t based on good data, or good data analysis:

Reprinted from Stanford (1966)
But over the past couple of decades, empirical studies have consistently shown that yield-base recommendation algorithms are unreliable:

Figure 4. Example of the relationship between corn grain yield and EONR found in states across the Corn Belt. The graph is for SC sites in Iowa (0.10 price ratio). Points on the left axis represent sites where there was no response to N fertilizer rate.

Source:
Not surprising theoretically: Here lower yielding area needs more N:

\[ y_1^* \text{ and } y_2^* \]

\[ f^1(N) \quad \text{Slope} = \frac{w}{p} \]

\[ f^2(N) \quad \text{Slope} = \frac{w}{p} \]
Historically, it was too labor-intensive, too expensive to generate the needed data:
Laying out small plots using tape measures and orange flags:
Applying inputs by hand or with special equipment:
grad student labor...
As a result,

• Most studies short-term, in few locations, with small plots
• How do the results of a field trial in Boone County, Iowa in 1996 imply for N management in Montgomery County, Ohio in 2019?
Other ways scientists have been trying to learn about $f(x,c,z)$:
Crop-modeling approach
What’s a “crop growth model?”

-A bunch of equations in a computer
-Basically, the model is trying to be $f(x, c, z)$
Climate Corp FieldView, Pioneer Encirca, Yara Adapt-N are all largely based on crop growth models.
A challenge they face: crop growth models don’t “move” well.
Adapt-N was built at Cornell, mostly with data from small-plot trials in the northeast U.S.

But the field characteristics ($c$) in New York can be quite different from those in Kansas. So the modeler has to have Kansas data to “calibrate” the model for Kansas. Not easy, and the data have always been expensive to generate.

\[ C_{NY} \neq C_{KS} \]
And that's a trick.

So FieldView, Enirca, Adapt-N are trying to run experiments here.

And that's a trick. to make recommendations there.
The “empirical” approach, starting with MRTN...
Since 2006, universities using MRTN approach, based on hundreds of Midwest strip trials:

Source:
Problem: MRTN rates are recommendations for whole regions—but in individual fields, they are sometimes well off the mark:

Figure 1. Nitrogen rate responses of corn following soybean in 51 on-farm trials in Illinois in 2017. Yellow triangles show the calculated optimum N rate (EONR) and yield at that rate for that trial, and the green circles show the MRTN rate and the yield at that N rate.

Source: http://bulletin.ipm.illinois.edu/?p=4095
DIFM is trying to build on what the MRTN project is doing by applying some important additional strategies:

- “Checkerboard” trials, not strip trials. More data, can put more confidence in statistical analysis
- Generating data to provide site-specific, not region-specific recommendations
- Generating data in the same fields for which we are providing management advice
- We are generating lots of characteristics data
How is DIFM different from all the other “Ag Big Data” talk out there?
After all, folks are gathering field characteristics data all over the place.
It’s good to gather field characteristics data. But,
Key Point: If you don’t vary managed input application rates, you don’t know how to look at $c$ and then determine optimal rates:

Can’t identify what these curves look like if you don’t vary $N$ much.
DIFM varies managed input rates, so we can see how yield responds to different rates:
Key: We got a USDA grant, and have been running very big “checkerboard” trials very inexpensively,
2018 Central Illinois soybean population trial, on 150+ acres:
2018 Ohio corn trial, varying UAN28:
Design of 2019 wheat trial, varying N on 160+ acre Washington wheat field:
100-acre cotton seed trial in south Texas:
A 100-hectare corn fertilizer and seed rate trial in Argentina:
How does DIFM run these trials, and how can CCAs work with us?
Farmer provides shapefile of field’s boundary
Farmer provides some other information, if available:

AB-line

Past yield file
DIFM designs trial, creates a shapefile, which is just an Rx map:

The farmer uploads the design’s shapefile:
Then the farmer just drives.
Farmer enjoyed taking video of seed rate monitor while he put the experiment in the ground:
Seed monitor and as-planted map. Yes, this works!
Use yield monitor at harvest:
We are starting to generate data.

Lots and lots of data.
U.S. Trials:

University of Illinois
University of Nebraska
Louisiana State University
Montana State University
Texas A&M University
Washington State University
Advanced Ag Alliance (New York)
2019 U.S. Trials:
Gather or create characteristics maps:

- Electroconductivity
- Elevation
- Organic Matter
- pH
- Satellite imagery
- Soil samples (when available)
- SSURGO soil types...
Taking a look at the data from one field, to illustrate the kinds of questions we are trying to answer:
Results from a recent central Illinois corn trial
Yield map

Some characteristics data (electroconductivity, soil type)
If we are talking uniform-rate management, his farmer’s usual N rate of 160 was too low on this field (given the year’s weather).

Applying 180 instead of 160 gets us (on average) a 15 bu increase in yield. With $4 corn and 0.35/pound N, that’s $4x15 - $0.35x20 = $53 per acre left on the table.

These results are what statisticians call “statistically significant” -- we are very confident that what we are seeing here is not just due to chance.
This field trial shows our most dramatic result of what a field trial can do for a farmer. We aren’t saying we can save every farmer $53 per acre. On many of our trials, the farmer’s usual rates were fine.

But we could tell them that the field trial was verifying that their existing management choices were looking good.
On the other hand, as a whole, this field was showing no response to seed rates beyond the lowest trial rate of 27000. The farmer’s usual rate was 35000. Clearly, at least given the year’s weather, he was wasting a lot of money on seed.
Not much evidence that electro-conductivity tells us much about optimal N rates.
The site-specific Rx we would give the farmer. Actually, the farmer pretty much will do best with a uniform strategy of (27000, 180) on the field.
Where to from here?

We are beginning to develop a cyber-infrastructure, to facilitate the running and analysis of thousands of trials per year. We want this to be global, with close collaboration among universities, farmers, and the private sector.
Data Analysis and Simulation System

Crop Growth Model to simulate \( y = f(x,c,z) \)

Statistical Algorithms to estimate reduced forms of \( y = f(x,c,z) \)

Artificial Intelligence, Neural Networks

Data Generation System

Characteristics Data Production System
- Remote and proximal sensors,
- Robotic soil and plant phenotype sensors

Optimization Algorithms: basis of decision-aid software

Central Database contains \( x, c, z, y \)

Basic field and management information

GPS-equipped variable-rate applicators implement trial as-applied data,

Yield monitored and mapped

Crop grows

Decision Support System/Information Portal
- Consultants, farmers use together
- Provides intuitive explanations for agronomics behind the results
- Probability-based for decisions under uncertainty
- Simulations of varying prices and weather
- Site-specific “prescription” maps

Get a rather complicated diagram:
Broad Schema of Cyber-Infrastructure:

- Central Database contains $x, c, z, y$

Data Analysis and Simulation System

Data Generation System
- $x$ and $y$: On-farm Precision Experiment System
- $c$: Characteristics Data Generation
- $z$: Weather Data System

Decision Support System/Information Portal
Key: To scale this up to thousands of trials, we are working to (semi-) automatize much of this system.
How do we get farm- and site-specific advice to farmers?
Every farmer can’t have someone with a PhD analyze the farm’s data to make recommendations.
We are creating software that CCAs can use to help their clients run trials, and come to sensible, statistically-reliable conclusions about what the data mean.
Idea: with a little training, you learn to...

Upload farmer info to the cloud

- Planter width: _____
- Combine width: _____
- N applicator width: ______
- Farmer’s usual N rate: _____
- Farmer’s usual seed rate: _____

Boundary file
Let our computer program design the trial, and send that shapefile to the farmer’s variable-rate equipment:
Farmer puts the trial in the ground

Data generated

DIFM

“As-planted” or “as-applied” data back to us
DIFM analyzes the data,
sends you a report, that you can use, along with your agronomic expertise, to give your client advice based on data from his own field.
Allows you to simulate different price and weather situations
Works with you to think about the agronomic intuition behind the results. (Not a black box.)
DIFM later meets with and your farmers to discuss results.
You help us learn, we try to help you and your clients learn.
We think that you’ll be able to show your clients that you are making them money.

Data, not testimonials:

“Last year, I got 205 bu/acre. But this year, I used Brand X, and I got 220 bu/acre!!!”
If you and your clients would like to partner with the DIFM project in some on-farm research in 2020
Talk to me here, or send me an email:

dsbullocc@Illinois.edu