# Picking the Low Fruit

Tony Grift

As a scientist and engineer, I wonder about many things, but recently the driving forces behind research in agricultural engineering have been on my mind. Most of us are under the impression that we do research to improve the world, elevate or at least maintain our standard of living, increase productivity and efficiency, secure the future for our successors, or some other lofty goal. However, over the last two decades, I've noticed that our research output does not always address such worthy problems. Instead, in too many research papers, the authors are trying to solve the lesser problems of finishing a degree, keeping a job, getting a promotion, or pleasing a department head.

To explore the current state of ag research, I thought it might be more interesting to look at what is wrong rather than looking at what works. To obtain some objective data, I searched the ASABE Technical Library using ag-related terms in titles and keywords (the complete list can be found in the sidebar). Here is a sample: "water" resulted in 3834 hits, "soil" in 2375, "manure" in 1077, "corn" in 933, "energy" in 854, "fertilizer" in 249, "ethanol" in 244, "GPS" in 218, "health" in 209, and "robot" in 167. In addition, "model" resulted in 3179 hits and "simulat" (to cover both "simulate" and "simulation") resulted in 1171, for a total 4356. Now guess how many hits "validation," the necessary counterpart of modeling, produced. Look it up in the sidebar – that is not a typo! This discovery made me wonder why there is so much emphasis on modeling. Could there be ulterior motives at work? I am afraid so. Modeling is inexpensive and controllable, whereas measurements can be tedious, expensive, erroneous, and yield unexpected results. This reasoning drove me to categorize papers that in my view are below par. Here are a few of these categories:

## Modeling madness

In *modeling madness* papers, the researchers conveniently focus on modeling and leave validation for "future research." Many journals no longer review papers that are solely about modeling, and with good reason: they have literally seen enough of them. NASA has gone through a similar process in funding modeling work. I am not arguing that modeling is useless (I wouldn't dare; my department head is a modeling "enthusiast"!), and I know and agree with some of the arguments in favor of it. Nevertheless, the numbers show that validation is too often downplayed or not even attempted, as if it were an afterthought. To me, unvalidated models are like modern art: colorful, even beautiful to some, but useless to the rest of us.

## Fit and forget

The counterpart of the modeling madness paper is the *fit and forget* paper, in which data are collected and a curve is fitted without attempting to understand the underlying process. The researcher measures something, records some data, and does an ordinary least squares fit, preferably with a high $R^2$ value. The result is often termed a "statistical model," which I find insulting to those who create models that have a basis in physics, but let's leave that discussion for another time. The question here is why the authors of *fit and forget* papers do not take it further and try to find a justification for that parsimonious model with the high $R^2$. If we do not try to discover the underlying mechanism, then what have we really learned? The obsession with high $R^2$ values can also result in completely illogical results. I remember reviewing a paper in which the researchers fitted an $\exp(\exp(x))$ curve through their data. It certainly yielded a high $R^2$ value, but nature does not create these kinds of relationships. What is the point of plotting such lines without contemplating their real meaning?

## Proof of concept

In a typical *proof of concept* paper, the researcher uses PVC tubes to represent real corn stalks, simulates sunlight with halogen lamps, and represents weeds with pieces of cardboard. In a particularly egregious example of this approach, two years of university research were wasted because the PI on a project was adamant about "harvesting the crop indoors" to avoid having to do the real thing. The reason was simple: harvesting Miscanthus takes place in the winter—not a fun time to be outside in the Midwest. I have yet to see a paper in which someone brought a "proof of concept" to fruition. Making a concept work outdoors is a huge step, and yet, as with the validation of models, it is hugely downplayed by researchers.

## Testing

In *testing* papers, a product such as a GPS receiver is driven around on a four-wheeler to assess its accuracy. Am I wrong in assuming that this is the manufacturer's job? Sure, I understand that agriculture places some constraints on GPS accuracy, but not only have we produced 218 papers on GPS accuracy, we have devoted entire sessions at our national meetings to this perceived "problem"! Could it be that the low cost of GPS receivers has something to do with their popularity as a researchable device? Are we solving a problem here or just creating one?

### Tool time

Over the years, cameras have become so ubiquitous and inexpensive that we find them all over the ag engineering landscape. I agree that there are areas where cameras are useful and that machine vision itself is a respectable field. In fact, high-speed video of agricultural processes can be literally eye-opening for those who are under the delusion that the mind can accurately predict how stuff works in agricultural machines. For example, I never realized the chaotic behavior of fertilizer particles being accelerated along a vane until I saw high-speed video of the process. However, in spite of my enthusiasm for visualization as a way to get our minds around a complex problem, cameras are overused. I am sure this is a skewed number, but the number of graduate student applications that get past my inbox filter and state that the applicants have machine vision expertise runs high. This is clearly a popular field, especially in Asia, and I wonder why. The answer might be related to why the field of mathematics is very advanced in Russia but not so much in the Caribbean. What would you work on when there is little money, nothing else to do for eight months of the year, and no palm trees in sight?

### Simulation

In *simulation* papers, the authors try to find a reason to use their favorite simulation software, such as MATLAB, ANSYS, or FLUENT. I have seen quite a few papers in which the researchers are clearly trying to find a hole for their favorite peg. It's the old "if you only have a hammer, everything looks like a nail" conundrum. EDEM is the latest tool that allows discrete element modeling with sometimes spectacular results. Note the word "spectacular." The outputs of these dynamic models sure are fun to watch, but how does one validate them? If the model's output looks similar to the process that it mimics, then is that a sufficient measure of the model's appropriateness?

### Safety in numbers

In *safety in numbers* papers, an idea is milked until the cow dies. With similar research reported in a variety of journals, there is bound to be some overlap, which is understandable. However, relentless reporting of small incremental steps on the same topic results in a huge number of papers, only a few of which have rigor and value. The term "salami science" is also appropriate here.

### Nothing lost, nothing gained

I have always regarded with suspicion any paper in which an index is defined. I call such efforts *nothing lost, nothing gained* papers. What do we really learn from an index? Sure, I understand that we do not want to wait until an idea is set in stone before using it, especially if it has some utility. For instance, the normalized difference vegetative index (NDVI) has been used widely, and successfully. However, imagine how much more useful it could be if we understood the underlying reason why it works: we could not only segment vegetation from backgrounds but possibly distinguish skin cancer from healthy skin.

## Technical Library Hits

| | | | | | |
|---|---|---|---|---|---|
| Water | 3,834 | Environment | 709 | Fertilizer | 249 |
| Model | 3,179 | Spray | 646 | Ethanol | 244 |
| Soil | 2,375 | Test | 641 | GPS | 218 |
| Air | 1,929 | Animal | 628 | Milk | 213 |
| Quality | 1,861 | Precision | 606 | Implement | 211 |
| Irrigation | 1,432 | Drainage | 592 | Health | 209 |
| Analysis | 1,303 | Image | 577 | Calibration | 207 |
| Crop | 1,204 | Tillage | 508 | Bacteria | 205 |
| Simulat | 1,171 | Tractor | 474 | Biological | 184 |
| Application | 1,167 | Safety | 440 | Electronic | 172 |
| Process | 1,088 | Vision | 429 | Robot | 167 |
| Manure | 1,077 | Fruit | 424 | Validation | 145 |
| Design | 1,067 | Food | 416 | Statistic | 136 |
| Measure | 989 | Housing | 379 | Optic | 134 |
| Waste | 955 | Soybean | 348 | Regression | 116 |
| Corn | 933 | Forest | 333 | Orchard | 104 |
| Energy | 854 | Infrared | 322 | Guidance | 85 |
| Erosion | 847 | Economic | 310 | Fish | 81 |
| Plant | 803 | Material | 306 | Ergonomic | 41 |
| Agriculture | 757 | Structure | 290 | | |
| Sensor | 714 | Remote | 278 | | |

The scientific method is "a body of techniques for investigating phenomena, acquiring new knowledge, or correcting and integrating previous knowledge." Where is the word "model" in that statement? How often do we find papers in which researchers dare to correct themselves or others? And few papers bother to repeat research. Repeating older research makes a lot of sense, but it's not glamorous. Researchers prefer to develop new and innovative devices, which is laudable, but unfortunately more and more often we merely integrate off-the-shelf components. As a result, more and more research involves integrating or testing off-the-shelf components. The new research paradigm seems to consist of obtaining a large grant, buying an expensive instrument, mounting it on a four-wheeler, collecting data (preferably on a warm sunny day), analyzing the data with some fun new tool, and then publishing five papers about it.

We are becoming less and less "applied" in our research, probably because the word "applied" does not have good connotations. Mathematics is the queen of sciences, and maybe that's why we assume that abstract ideas are somehow superior. I must admit that I am in awe of mathematics. However, mathematics is only useful when it's applied to the betterment of human (and animal) kind, and that's where engineers come in. We should take pride in being applied, in our ability to build useful instruments and machines, and to come up with tangible results that really do some good in the world. If we keep pushing ourselves away from the workshop, away from the field and farm, away from manure and steel, then we will become mere theorists who talk brilliantly about the challenges facing the world but who don't really do anything about them. I don't want to be part of that club.

ASABE member Tony Grift is an associate professor of agricultural and biological engineering, University of Illinois at Urbana-Champaign, Urbana, USA; grift@illinois.edu.