Review on the Report: "Warm Season Turfgrass Nitrogen Rate and Irrigation BMP Verification" by Drs. Laurie E. Trenholm, John L. Cisar, J. Bryan Unruh, and Jerry B. Sartain

Xu-Feng Niu
Professor of Statistics
Florida State University
Tallahassee, FL 32306-4330
November-December 2012

1. Background

The research project "Warm Season Turfgrass Nitrogen Rate and Irrigation BMP Verification" started in 2004 and was supported by FDEP Contract WM869. The four Principle Investigators (PIs) are Professors from the Institute of Food and Agricultural Sciences, University of Florida: Dr. Laurie E. Trenholm; Dr. John L. Cisar, Ft. Lauderdale Research and Education Center (FLREC); Dr. J. Bryan Unruh, West Florida Research and Education Center (WREC); and Dr. Jerry B. Sartain.

The main purposes of the project under review were to assess the effects of lawn and landscape management practices on Florida natural water resources, mitigate any potential nonpoint source pollution of water resources from fertilization and other lawn-grass maintenance practices, and study leaching of nitrates and phosphorus from turf under various management routines for verifying, or modifying, statewide Best Management Practices (BMPs) to protect water quality. Based on the report, the research was conducted in 3 locations statewide: the Ft. Lauderdale Research and Education Center, the Plant Science Research and Education Unit in Citra, and the West Florida Research and Education Center in Jay. The whole research consisted of the following six parts:

- Project 1, Nitrate Leaching from Newly Sodded Grasses;
- Project 2, Nitrate Leaching from Established Grasses;
- Project 3, Nitrate Leaching Due to Nitrogen Source and Timing;

• Project 4, Nitrate Leaching During Fall/Winter Months;
• Project 5, P test kits – Phosphorous Requirements and Leaching
• Project 6, Slow Release Nitrogen Study

The reviewer of this report had a full-day meeting with the four PIs (one on telephone) and the FDEP management team supervising this project on November 8, 2012. The four PIs presented their research in the meeting. My review and comments on the report and the project will focus on the following two aspects: experimental designs of the project and statistical data analysis.

2. Experimental Designs.

Mainly three types of experiment have been used in the research project: randomized complete block designs, split-plot designs, and split-split-plot designs. All of the three types of designs are natural choices for this kind of research.

Randomized complete block designs are probably the most popular and useful designs in many fields, in which the experimental units are grouped into blocks by some nuisance factors that may affect the measured result, but are not of primary interest. Therefore, within each block, the conditions are relatively homogeneous, but between blocks, large differences may exist.

Split-plot designs were originally developed by Fisher (1925) for use in agricultural experiments, which nowadays are also widely used in social studies or educational research. A split-plot experiment is a blocked randomization experiment, where the blocks themselves serve as experimental units for a subset of the factors. In general, the blocks in a split-plot design are referred to as whole plots, while the experimental units within blocks are called split plots, split units, or subplots.

In practice, two levels of randomization are conducted in a split-plot experiment, corresponding to the two levels of experimental units. One randomization is conducted to determine the assignment of block-level treatments (Factor A) to whole plots, and another randomization is used to decide the assignments of Factor-B treatments to split-plot experimental units. The restriction on randomization mentioned in split-plot designs may be extended to more than one block factor. For the case where the restriction is on two block factors the resulting design is called a split-split-plot design.

a). An example of randomized complete block experimental designs conducted by the PIs:

JAY, FL – Dr. J. Bryan Unruh, Project 1, Report Page 54

"The study was conducted at the University of Florida's West Florida Research and Education Center, (30°46'3N, 87°08'W) on native soil identified by the USGA survey as Fuquay loamy sand. The experiment consisted of 24-4m x 3m) turfgrass plots in a **randomized complete block with 4 replications**. One of three N rates; 12.232, 24.464, and 48.928 N kg ha-1yr-1 in soluble granular form custom blended using 10-10-10, supplemented with 0-46-0, and 0-0-62 for available P and K rates of 48.928 kg ha-1yr-1 was applied to half of the 24 plots at day of seeding 16 May 2005 (Trial 1), 03 May 2005 (Trial 2) with the other half receiving one of the three N rates 30 days after seeding 16 June 2005 (Trial 1), 02 June 2006 (Trial 2)."

b). Examples of split-plot experimental designs conducted by the PIs:

1). FT. LAUDERDALE – Dr. John Cisar, Project 2, Report Page 96.

"Field plot design was a randomized complete block, split-plot design with four replicates. Main blocks (8 x 4 m) consisted of one of two irrigation regimes: 2.5 mm daily (Low) except when daily precipitation > 6.4 mm (irrigation turned off), and 13.0 mm three times weekly (High). Sub plots (2 x 4 m) consisted of four nitrogen (N) rates (98, 196, 294, and 588 kg N ha-1 yr-1) and (49, 98, 196, and 294 kg N ha-1 yr-1) on the St. Augustinegrass and bahiagrass respectively."

2). JAY, FL – Dr. J. Bryan Unruh, Project 2, Report page 129.

"The study was conducted at the University of Florida's West Florida Research and Education Center, (30°46'3N, 87°08'W) on stands of Common Centipedegrass and Floratam St. Augustinegrass grown on native soil identified by the USGA survey as Fuquay loamy sand. The experiment consisted of 64-6m x 3m turfgrass split plots with 4 replications. Whole plots (Irrigation Regime and Turfgrass variety) measure 12m x 12m; and split plots measure 3m x 6m. Sod was established the week of 01 May 2005. Treatments were applied (2005 Yr 1) 08 Jun, 09 Aug, (2006 Yr 2) 14 April, 14 Jun, 11 Aug, (2007 Yr 3) 16 Apr, 14 Jun, and 14 Aug. Grass splits received 4 N rates. Common Centipedegrass 6.103, 12.206, 24.413, and 48.826 N kg ha-1yr-1, and Floratam St. Augustinegrass 16.259, 32.713, 65.085, and 97.652 N kg ha-1yr-1 from soluble urea, applied as a spray with a volume of 1222 L per hectare."

c). An example of split-split-plot experimental designs conducted by the PIs:

Project 2, Report page 168.

"The research was conducted at the Ft. Lauderdale Research and Education Center, the Plant Science Research and Education Unit in Citra, and the West Florida Research and Education Center in Jay. St. Augustinegrass cv. 'Floratam' was planted in each location and allowed to establish prior to initiation of treatments. Plots were 4 m x 4 m with a lysimeter in the center of each plot as previously described in this report."

"**Experimental design in each location was a split plot with 4 replications. For the analysis that included location, the design was a split-split plot, with locations as whole plots, irrigation as sub plots and NR as sub-sub plots.** Following correction for outliers, all data were analyzed with PROC GLM or PROC REG (SAS Institute, 1999) Significance was determined at the 5% probability level."

Main comments on the experiment designs:

In the reviewer's opinion, all of the three types of designs are natural choices for the research conducted in this project. In some PIs' reports, experiment design types were not clearly stated. Moreover, the reviewer would like to see some brief discussions on how randomizations were done in the experiments, which were not provided in the current report.

3. Statistical Data Analysis

Statistical data analyses were performed by the four PIs and results of the analyses were presented in the report. For data obtained from the three types of experimental designs, analysis of variance (ANOVA) and regression techniques were mainly used in the research.

As mentioned in the background section, one main purpose of the research is to study leaching of nitrates and phosphorus from turf under various management routines that result in different nitrates and phosphorus application rates. Based on the report, results showed that the effects of treatments were quite often not significant at the significant level $\alpha = 0.05$. On concern of FDEP is that the results such as "because of a large amount of variability in the samples over time, there is less than 95% confidence that 2 means are different" may lead to some misunderstanding and suggest that fertilization rates be increased from current BMPs.

In the reviewer's opinion, the analyses performed by the four PIs could be improved by including covariates and environmental factors in the current ANOVA and regression analysis. For example, the following covariates may be considered in the modeling and analysis:

1) **Rainfall**. It is well known that rainfall is a source for nitrates and phosphorus. For instance, on page 6 of the main report Dr. John Cisar mentioned:
"During Trial 1, seven separate rainfall events occurred, with a cumulative total of 108 mm. Nitrate-N and orthophosphate concentrations within the rainfall for five of the rain events ranged from 0.05-0.15 and 0.05-0.80 mg l-1 respectively, with means of 0.11 and 0.35 mg l-1 for nitrate-N and orthophosphate respectively. Rainfall contributed a total of 11.9 and 38.3 mg m-2 of nitrate-N and orthophosphate over the Trial."

   On page 21 of the report, Dr. John Cisar stated:
"Nitrate-N concentrations were influenced by fertilizer treatments during the second half of Trial 2 (Tables 9 and 10). For the most part, nitrate-N concentrations were similar except for shortly after a large rain event occurring on 29 September resulting in 46 mm of rainfall (Figure 3). For the remainder of Trial 2, nitrate-N concentrations were highest in percolate collected from St. Augustinegrass fertilized at 30DAI (mean = 14.7 mg l-1, Figure 3). These concentrations were similar to concentrations from St. Augustinegrass fertilized 0DAI (mean = 2.68 mg l-1), but higher than concentrations from St. Augustinegrass not fertilized (mean = 0.60 mg l-1, P=0.03). Again, the difference in means was driven by rainfall events (Figures 3 and 4)."

2) **Grass Quality**. It is also well known that leaching rates is affected by grass quality. As the PIs mentioned on Page 3 of the report: "Bowman et al. (2002) observed a 92% reduction in NO3–N leaching from St. Augustinegrass grown in lysimeters from a second treatment application as compared to the first treatment. Nitrogen was applied as ammonium nitrate at a rate of 5.0 g·m-2 on both treatment dates. The authors attributed the sharp reduction in NO3–N leaching following the second treatment application to development of a more extensive root system, a larger microbial population to increase N immobilization, and less water percolation in the columns."

3) **Season.** On Page 38, Dr. John Cisar mentioned: "Seasonality may have influenced results:

- Trials 1 and 3 bracketed an active growth period. During this time less leaching occurred in general. Although not significant, a trend for less nitrate-N leaching coupled with significantly better quality was determined for turfgrass fertilized 30 DAI.

- Trial 2 corresponds to a less active growth period. Perhaps at this time, the turfgrass is less resource efficient. This would explain the higher nitrate-N concentrations and greater leaching due to fertilizer treatments documented during the 31-60 DAI."

Other covariates and environmental factors may be evaluated and included in the investigation of relationship between leaching rates and application rates.

3.1 Model Selection.

When many covariates are considered and included in the analysis, model selection criteria and procedures should be used to choose the best model. Inference and conclusions should be based on the chosen best model.

In ANOVA and regression analysis, many models may be considered to describe the relationship between the response Y and predictors $X = (X_1, \cdots, X_p)$. The general linear regression model (with ANOVA model as a special case) has the form:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \cdots + \beta_p X_{pi} + \varepsilon_i, \qquad i = 1, 2, \cdots, n, \tag{1}$$

where the random errors ε_i's are assumed to be independent and identically distributed with mean zero and variance σ^2. It is usually assumed that the random errors have a normal distribution.

When different models are fitted to the observations $\{ (X_{1i}, \cdots, X_{pi}; Y_i), \ i = 1, 2, \cdots, n \}$, model selection techniques need to be used to decide which model fit the data better. Statistical inferences such as estimation and prediction will then be based on the best model selected. The Bayesian Information Criterion (SBC) suggested by Schwartz (1978) is one of popular criteria for model comparison. For a fitted model (linear or nonlinear) with p parameters, the SBC is defined as

$$\text{SBC(p)} = -2 \ \log(\text{maximum likelihood function}) + p \times \log(\text{n}),$$

where the likelihood function is based on the distribution assumption of the model such as normal or log-normal or other distribution families, and *n* is the sample size. When the random errors ε_i's have a normal distribution, the SBC(p) has the simplified form:

$$\text{SBC(p)} = n \times \log\left(\sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 / (n-p-1)\right) + p \times \log(\text{n}), \tag{2}$$

where $\hat{Y}_i$ is the fitted value based on one of the candidate models and $\sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2$ is the **Residual Sum of Squares (RSS)** based on the fitted candidate model.

Intuitively, there are two parts in (2), the first part is

$$n \times \log\left(\sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 / (n-p-1)\right) = n \times \log \hat{\sigma}^2,$$

which is a measure of the goodness-of-fit of the candidate model. In general, increasing the number of parameters in a model will improve the goodness-of-fit of the model to the data regardless how many parameters in the **true model** that generated the data. When a model with too many predicators (significant or not significant ones) is fitted to a data set, we may get a perfect fit but the model will be useless for inference such as prediction. In statistics, fitting a model with too many unnecessary parameters is called *over-fitting*. The second part in SBC, $p \times \log(\text{n})$, puts a penalty term on the complexity of a candidate model, which will increase when the number of parameters in a candidate model increases. Thus the criterion SBC requires a candidate model fitting the data well and penalizing the complexity of the model. **For a group of candidate models, the SBC value can be calculated for each of the models and the preferred model is the one with the lowest SBC value.**

3.2 Outlier detection and treatments.

One important topic in ANOVA and regression analysis is how to detect and handle outliers. Some outliers were observed in the figures presented in the report, such as in Figure 3 (Nitrate-N concentrations during Trial 2 were influenced by fertilizer level) and in Figure 35 (N Timing Blackout Test: Nitrate Leaching from 8 June – October 6, 2009.)

One comment provided by FDEP to the four PIs is "Outliers are defined as > 2 SD by the researchers, however, such a deviation caused by heavy rains, or very high concentrations due to a low volume first flush effect may well be legitimate N loads that must not be removed."
The reviewer agrees with this comment. After rainfall and other covariates are included in the analysis, some outlier may be explained by the effects of the covariates. Outliers should be identified and handled based on the residuals from the selected best model, not based on the original data.

References

Fisher, R. A. (1925). Statistical Methods for Research Workers. Edinburgh: Oliver and Boyd.

Schwartz, G. (1978), Estimating the Dimension of a Model, *The Annals of Statistics*, 6, 461-464.