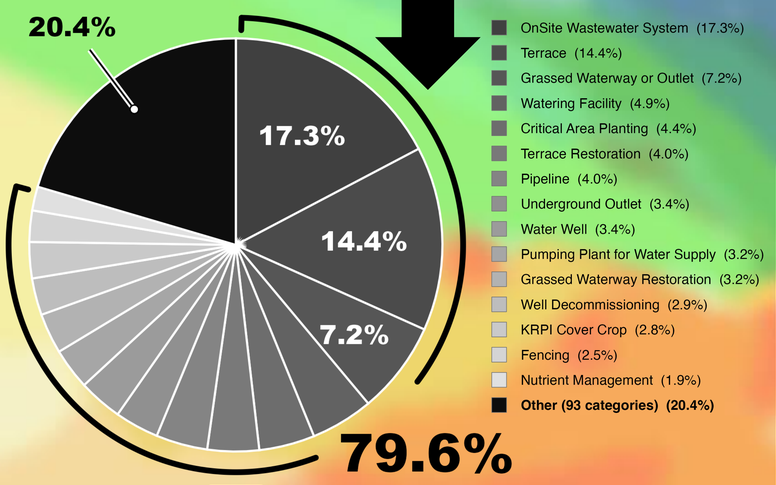
Mining the Mind: Expert Elicitation as a Form of Dimension Reduction

Robert M. Sholl[1], Trevor J. Hefley[2]
Department of Statistics, Kansas State University

The Problem

Kansas Department of Agriculture (KDA) has been tracking cost-shared best management practices (BMPs) implemented state-wide for over two decades with **no strict naming or grouping scheme**.

The result is **108 unique practices**, many of which are duplicates under different names. **Until this is reduced researchers cannot determine clear effects of BMPs on water quality from the KDA database.**



The Solution

Categories can be **condensed into higher order groups** using dimension reduction techniques. Typical approaches include Principal Component Analysis methods, tree-based grouping, and regularization techniques such as Elastic Net[2].

Our approach involves **converting expert opinion on practice exclusions and groupings into data to act as a shrinkage method.**

Eliciting Data

Expert opinions are gathered through interviews, workshops, or digital surveys, then formed into categorical data. Our method for expert data leverages likelihood-based measurement error models; expertise can then be quantified via Fisher Information. As such we can **define an expert to be useful** if their **opinions reduce the variance in parameter estimates[1].**

Acknowledgements

KANSAS STATE UNIVERSITY | Kansas Water Institute

References

[1]Lele, S.R. and Allen, K.L. (2006), On using expert opinion in ecological analyses: a frequentist approach. Environmetrics, 17: 683-704. https://doi.org/10.1002/env.786
[2]Nanga, S. , Bawah, A. , Acquaye, B. , Billa, M. , Baeta, F. , Odai, N. , Obeng, S. and Nsiah, A. (2021) Review of Dimension Reduction Methods. *Journal of Data Analysis and Information Processing*, **9**, 189-231. doi: 10.4236/jdaip.2021.93013.

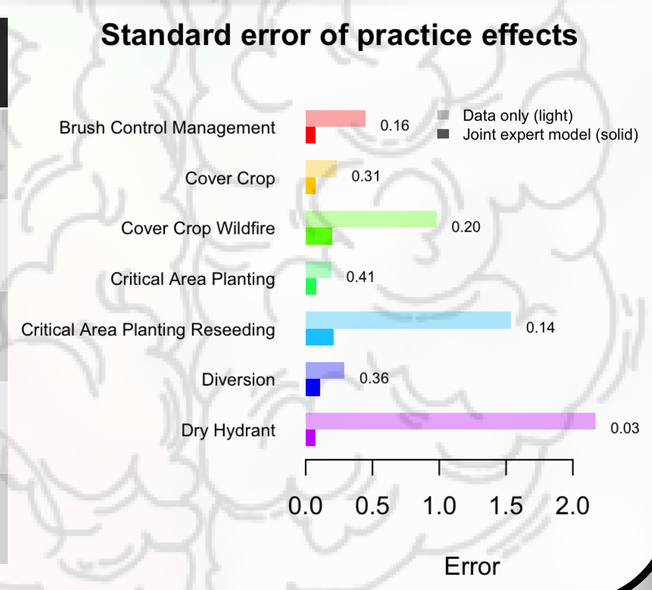
Let's Collaborate!

rmsholl@ksu.edu
rmshksu.github.io
linkedin.com/in/robertmsholl



Results

Expert	Band Ratio
1	0.857
2	0.015
3	0.922
4	0.421
5	0.667



Conclusions & Next Steps

The academic expert **(Expert 2) was the most informative** by the confidence band ratio metric. **This could be due to the design of the interview.** Future studies would want to carefully consider their elicitation techniques.
Our study shows that expert opinion can be used as a form of dimension reduction, in a fully frequentist framework. We have also shown that an individual expert can **outperform** traditional variable selection methods.

Error Model

We use a non-Bayesian implementation of a typical measurement error model, considering the possibility of expert data contributing bias[1];

$$L(\beta, \eta; O, E, X) = \prod_{i=1}^{n} g(E_i \mid O_i; \eta) \prod_{i=1}^{n} f(O_i; X_i, \beta).$$

Gamma Response

The response variable is an estimate of phosphorous loads in every HUC12 in Kansas, generated from a spatial model that assumes a Gamma distribution (strictly positive values);

$$f(x; \alpha, \lambda) = \frac{\lambda\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}.$$

Quantifying Expertise

An expert is deemed 'useful' if the addition of **their opinion contributes to an increase in the Fisher Information** contained in the parameter estimates[1].

This is the same as when the confidence band for an estimate produced using expert opinion is less than the confidence band from only the pure observations. That is to say:

$$\frac{\mathrm{CI}_{\mathrm{Expert}}}{\mathrm{CI}_{\mathrm{Observations}}} < 1 \implies \text{Informative Expert.}$$

Meet our Experts

This approach allows for **anyone or anything to be considered an 'expert'.**

Our study uses five expert datasets:

1. A joint elicitation of two field experts.
2. An interview with an academic expert.
3. A non-expert applying a heuristic.
4. Principle Component Analysis.
5. Least Absolute Shrinkage Selection Operator.