A Statistical Learning Approach to Pasture, Rangeland, Forage (PRF) Insurance Coverage Selection

Samuel D. Zapata
Associate Professor, Department of Agricultural Economics, Texas A&M University, McAllen, TX.

Xavier Villavicencio
Extension Associate, Department of Agricultural Economics, Texas A&M AgriLife Extension Service, Texas A&M University, Weslaco, TX

Anderson Xicay
Extension Assistant, Department of Agricultural Economics, Texas A&M AgriLife Extension Service, Texas A&M University, Weslaco, TX

Selected Paper prepared for presentation at the 2023 Agricultural & Applied Economics Association Annual Meeting, Washington DC: July 23-25, 2023

A Statistical Learning Approach to Pasture, Rangeland, Forage (PRF) Insurance Coverage Selection

Abstract

This study contributes to the existing climate risk management literature by developing and testing different statistical learning algorithms to assist producers to effectively participate in the Pasture, Rangeland, Forage (PRF) insurance program. The proposed methods are expected to outperform existing coverage selection strategies by improving the predictive power of the chosen coverage policies. Furthermore, the proposed coverage selection methods can be viewed as comprehensive and universal solutions to the PRF problem given that they are applicable to every participating state. Additionally, the optimization methods proposed in this study can be extended to analyze the decision process of other Rainfall Index-based insurance programs and used in a wide range of resource allocation decisions.

***Keywords*:** Shrinkage selection models, Ensemble learning methods, Mixed integer nonlinear programming, Risk management, Texas.

JEL Classifications: Q18, C61, Q15

Introduction

The Pasture, Rangeland, Forage (PRF) is a pilot insurance program launched in 2007 as a tool for livestock and forage producers to mitigate the risk of forage loss associated with the lack of precipitation. Currently, the program is available in 48 states and over 250 million acres were enrolled in 2022, making it the largest crop insurance program in the U.S. in terms of the number of participating acres. However, compared to traditional crop insurance programs, the PRF is an index-based insurance that utilizes aggregated weather data to estimate the relative temporal precipitation within a specific area. Actual farm level precipitation or forage production is not measured. Thus, given exogenous precipitation index realizations, the indemnity received depends merely on farmers' *ex-ante* selection of the program coverage parameters.

Farmers can tailor their policies by selecting appropriate grid intervals, coverage level and productivity factor. These three parameters determine the amount of protection, premium rates, subsidy level, and expected indemnity payments. Regardless of their economic implications, only a handful of studies have been conducted to untangle the relationships between decision parameters and the impacts they have on expected returns and risk. Previous studies have focused on providing selection recommendations for specific geographical areas, evaluated a limited number of decision parameters, and omitted parameter selection restrictions (Jimenez Maldonado, 2011; Diersen, Gurung and Fausti, 2015; Steward, 2018; Westerhold et al., 2018). Zapata and Garcia (2022) made the first approximation to incorporate all the PRF decision variables and program restrictions in the selection process, through a mixed integer nonlinear programming model based on risk-efficient portfolio selection strategies.

However, studies such as Brodie et al. (2009) suggest that traditional portfolio optimization methods suffer from a lack of empirical stability, since the results are sensitive to variations in expected returns, volatilities, or correlations in the model variables. In this sense, it has been proposed to use alternative methods that seek to reduce this volatility and improve the predictive power of the selected portfolio. Novel data-driven methods such as statistical learning algorithms have been identified to overcome the actual performance of traditional portfolio optimization methods. The main objective of this study is to explore and test several statistical learning algorithms to develop robust coverage selection strategies for the PRF program.

Parameters Selection Conundrum

Insurable acres are assigned to a coordinate system with grids equal to 0.25 degrees in latitude by 0.25 degrees in longitude (or approximately 17 by 17 miles at the equator). Precipitation is interpolated for each grid, and it is estimated for eleven overlaying 2-month index intervals (i.e., *January-February, February-March, ..., November-December*). Estimated interval precipitation levels for each individual grid are transformed into a precipitation index (I_j), where a value of 100 is assigned to the index interval historical average precipitation. Distribution of the percent of value (w_j) is used to select a disjoint set of the 11 available 2-month index intervals. The sum of the percent of values must add up to 100 percent, and no more than a given state-based percent (typically between 50 to 70 percent) can be assigned to a particular index interval. In addition, no less than 10 percent can be allocated to the selected index intervals. Different premium rates (α_j) are associated with each index interval, and the premium amount depends on the percent of value assigned to each index interval.

An indemnity payment is triggered when the index interval precipitation index falls below the chosen coverage level (λ). Coverage levels between 70 to 90 in 5-unit increments are available. Higher premium rates and lower subsidy levels (s) are associated with higher coverage levels, and vice versa. Given the county base value (v), farmers have the option to modify the amount of protection by adjusting the coverage level and productivity factor (γ). Namely, policyholders can choose a productivity factor between 60 to 150 percent in 1-unit increments.

Overall, there are up to 100,555 possible selection combinations of the coverage level, productivity factor, and index intervals, without considering the percent of value assigned to each index interval. It is assumed that the goal of risk-averse farmers is to maximize their expected net returns from participating in the PRF program:

(1) $$E(R) = E\{\sum_{j=1}^{11} w_j v \lambda \gamma [\max\left(\frac{\lambda - \mathrm{I}_j}{\lambda}, 0\right) - (1-s)\alpha_j]\},$$

subject to a set of individual risk preference and parameter selection constraints, $\boldsymbol{f}(w_j, \lambda, \gamma) \leq \boldsymbol{X}$.

Proposed Selection Methods

The Mean-variance approach presented in Zapata and Garcia (2022) serves as a benchmark for our study. In this coverage selection strategy, the producer's objective function is to minimize the payout variance $\mathrm{V} = \sum_{j=1}^{11} \sum_{k=1}^{11} w_j w_k \sigma_{jk}$, subject to a given expected return, $E(R)$ in Eq. 1.

In this study, we proposed different shrinkage and meta-learning algorithms to the context of the PRF program. The proposed methods are tested on an array of representative grids across Texas precipitation zones. Compared to existing portfolio selection strategies, the proposed algorithms could significantly improve risk management decisions, both in terms of expected returns and risk.

The proposed methods include three shrinkage and two ensemble learning coverage selection strategies. Namely, Least Squares Shrinkage (LSS) penalizes the magnitude of the percent of value allocated to each index interval. Minimum Selection Shrinkage (MSS) limits the number of index intervals chosen. The third strategy, Least Squared and Selection Shrinkage

(L3S), combines the penalties of the two aforementioned shrinkage methods. These methods tend to shrink the size and number of coefficients and reduce the sensitivity of the optimal solution to changes in the parameters of the problem.

In terms of ensemble learning methods, they resample the feasible search space of the optimization problem to obtain robust solutions and better predictions in terms of smaller out-of-sample mean square errors. Bagging Selection (BS) used bootstrapping techniques to identify and aggregate alternative coverage solutions. Similar to BS, Random Set Selection (RSS) uses bootstrap sampling, but at each iteration a random subset of feasible decision parameter levels is considered.

Empirical Application and Preliminary Results

We implemented the proposed methodologies in different grids located uniformly throughout the state of Texas. Grids where the observed returns of the PRF program were high, medium and low were selected and the received payouts of a common strategy used by farmers, were compared with the proposed coverage selection strategies.

In particular, the average indemnity received by farmers in the last 5 years (2018-2022) was used as the expected value to be obtained in the optimization procedure. For the shrinkage methods, the value of the penalty parameters were calculated using a recursive cross-validation approach in the training set,

For illustration purposes, we considered a grid in Anderson County. We focused on conventional, non-irrigated forage for hay use as these are the most common practices in the region. Current actuarial data and past precipitation indexes (i.e., 1948-2017) were used to identify the optimal set of parameters for each year.

Compared to a recommended naïve coverage selection strategy (i.e., assigning the same percent of value to 6 index intervals), the highest payout gains are observed with the shrinkage models: MSS with a Sharpe index (mean/standard deviation) higher by 320%, L3S (307%) and LSS (141%), while the ensemble learning methods reported less gains, for which the results were 1.7% better with BS, -36% with RSS.

Conclusions

This is one of the first studies that combines shrinkage and ensemble learning techniques in the selection of an optimal portfolio for crop insurance. Preliminary results suggest that the proposed methods outperform common strategies used by farmers, in terms of higher observed returns or less variance, measured by the Sharpe ratio obtained with each proposed technique, except the RSS.

This study provides improved decision-making tools for farmers to effectively participate in the PRF program. Furthermore, the proposed coverage selection methods can be viewed as universal solutions to the PRF problem given that they are applicable to any participating state. Additionally, the proposed coverage selection strategies can be extended to analyze the decision process of other Rainfall Index-based insurance programs such as the Apiculture and Annual Forage insurance programs.

References

Brodie, J., Daubechies, I., De Mol, C., Giannone, D., & Loris, I. (2009). Sparse and stable Markowitz portfolios. Proceedings of the National Academy of Sciences, 106(30), 12267-12272.

Diersen, M., P. Gurung, and S. Fausti Optimal Allocation of Index Insurance Intervals for Commodities. Selected Paper. Atlanta, GA: Southern Agricultural Economics Association, January 31–February 3, 2015.

Jimenez Maldonado, A.J. “Developing Decision Rules for the Rainfall Index Insurance Program: An Application to Pennsylvania Producers.” Ph.D. Dissertation, The Pennsylvania State University, May 2011.

Stewart, S.A. “Financial Outcomes from Selection of Insurance Intervals.” Master Thesis. Utah State University, Logan, 2018.

Westerhold, A., C. Walters, K. Brooks, M. Vanderveer, J. Volesky, and W. Schacht “Risk Implications from the Selection of Rainfall Index Insurance Intervals.” Agricultural Finance Review 78,5(2018):514–31.

Zapata, S.D. and J. M. García (2022). “Risk-Efficient Coverage Selection Strategies for the Pasture, Rangeland, Forage (PRF) Insurance Program.” Journal of Agricultural and Applied Economics 54, 286–305.