Finding suitable weather indices for novel fisheries index insurance

Manuscript Draft for Fish and Fisheries Submission

Nathaniel Grimes

2025-11-30

Index insurance is a financial tool gaining traction for application in fisheries. However, identifying weather indicies to serve as triggers remains a significant barrier to widespread implementation. Weak correlations between weather variables and catch will lead to substaintial basis risk. This paper is the first to empirically evaluate the feasibility of index insurance for fisheries. I use the Market Squid Fishery in California as a case study to test insurance contracts built on five different prediction models with twelve weather variables. Only three weather variables in single index contracts improved fisher welfare at an actuarially fair premium. Only the index with krill abundance was profitable for insurance companies. All four multivariatve models improved utility from a range of 5%-35% at actuarially fair premiums. When charging a competitive market premium, contracts built with random forest, regularized random forest, and support vector machine algorithms were profitable for insurance. Thus, index insurance is feasible in a private market, and provides fishers with protection against environmental variability. However, bolstering fisher trust in products built from complex machine learning algorithms remains a significant step.

1 Introduction

Fishers face enormous environmental variability that impacts their livelihoods. Extreme weather events that degrade stock health such as stagnant upwelling or marine heatwaves can lead to significant losses in fishery productivity (Smith *et al.* 2023; Szuwalski *et al.* 2023; Villaseñor-Derbez *et al.* 2024). These shocks can devastate fishing communities, leading to income loss and food insecurity (Holland and Leonard 2020; Jardine *et al.* 2020). The need for financial tools to protect fishers from these shocks has never been greater (Mumford *et al.* 2009; Sethi 2010; Kasperski and Holland 2013; Sumaila *et al.* 2021). Index insurance has risen as a leading candidate to protect fishing communities during such disasters, but accurately

1

identifying suitable weather indices for fisheries has impeded widespread deployment (Watson *et al.* 2023; Hobday *et al.* 2025). This paper is the first to empirically construct index insurance contracts to protect against fluctuations in fishery biological productivity.

Index insurance is a financial product that pays out when an independently verified index, such as sea surface temperature, falls below a predetermined threshold called the trigger. The index is ideally chosen to be highly correlated with the asset being insured. In fisheries, catch or yield will be the primary underwritten asset. The trigger represents a critical value where the underlying asset is expected to suffer a loss. Fishers pay a premium that transfers income from good periods to bad periods. Index insurance is preferable to indemnity because it is faster to administer, cheaper to monitor, and avoids moral hazards (Barnett and Mahul 2007).

The COAST program is the only extant index insurance product for fisheries (Sainsbury *et al.* 2019). It is a pilot program that insures small island nations in the Caribbean from hurricanes. It uses indices of wave height, wind speed, and storm surge to initiate payouts once they fall below the trigger. There is growing interest to expand index insurance to cover other environmental shocks, but significant market barriers have limited widespread implementation (Watson *et al.* 2023; Hobday *et al.* 2025).

The primary market barrier is empirically constructing contracts that sufficiently reduce basis risk. Basis risk is the potential for a policyholder's actual loss to not align with the contract payouts. For example, if a fisher suffers poor catch, but the weather index does not exceed the trigger, then the fisher will not receive payouts. The opposite may occur where insurance companies payout despite fishers receiving high catch. Basis risk lowers demand for index insurance and leads to exorbitant premium rates (Binswanger-Mkhize 2012; Clarke 2016; Clement *et al.* 2018).

Identifying triggers and indices that are highly correlated to loss is the best way to eliminate basis risk (Jensen *et al.* 2019). Agricultural index insurance has refined techniques to build and price contracts over the last 20 years (Jensen *et al.* 2016). The most common design strategy is to match individual farms to the nearest meteorological stations, then use linear regression to model yield-weather distributions (Miranda and Farrin 2012; Spicka and Hnilica 2013; Conradt *et al.* 2015; Dalhaus *et al.* 2018; Benso *et al.* 2023). Premiums are then calculated based on the payout distribution of prior years for a given contract.

However, fisheries possess more difficult data and empirical challenges than agriculture. Predicting fishery output from weather variables is notoriously difficult. It is widely established that climate and weather affect fish populations (Lehodey *et al.* 2006), but even the most advanced fishery stock assessment models use little to no environmental data for projections (Privitera-Johnson and Punt 2020). Spatial distributions of fisheries can cover large areas making it unclear where and how much specific weather variables impact biological abundance. Additionally, if weather to catch connections can be established, they are often nonlinear and interact with other variables. The biological dynamics of fish reproduction can lead to lower stock health persisting for years after a negative shock (Hilborn *et al.* 2003).

Fishery management is also unique challenge to navigate empirically that is absent from agriculture insurance. Good fisheries management provides the necessary data to model weather to catch dependency, but can also sever the connection. For example, many fisheries in the United States transitioned from open access to individual transferable quota systems in the early 2000s (Costello and Polasky 2008). Distortions in catch would be attributable to shifts in management regimes rather than weather. In ensuing years after the implementation of quota systems, catch would then become contingent on management choices instead of weather.

Despite these challenges, innovations in fishery and oceanographic data collection, and increased exposure of fishers to environmental variability, suggest a new opportunity has arisen to reconsider the use of index insurance in fisheries. New methods including machine learning algorithms could better capture the complex, nonlinear relationships between weather and fishery productivity.

This paper is the first to empirically evaluate the feasibility of index insurance for fisheries using established univariate, linear methods and novel machine learning techniques. I use the Market Squid Fishery in California as a case study to test insurance contracts built on five different prediction models with twelve weather variables. I evaluate whether the contracts improve fisher welfare by determining their marginal willingness to pay. Through the marginal willingness to pay, I assess whether insurance companies could feasibly offer contracts within a private market to remain solvent.

Only one index, krill abundance, in univariate linear models improved fisher welfare at the market premium. All four multivariate models improved utility from a range of 5%-35% improvement at actuarially fair rates. When charging a competitive market premium, the contracts built with random forest, regularized random forest, and support vector machine algorithms were profitable for insurance companies. The leading variable of influence in the nonlinear models included upwelling measures, regional El Nino metrics, and krill abundance. Communicating to fishers the relative influence of each variable is crucial to build trust and transparency when using complex models. Overall, there is potential for index insurance to be feasible in a private market, and provide fishers with protection against environmental variability.

These results contribute to three literatures. Fisheries risk management includes formal and informal means of mitigating risk. However, developing formal financial tools to manage fisheries risk has been limited. Prior studies have explored the theoretical application of insurance in fisheries (Herrmann *et al.* 2004; Mumford *et al.* 2009; Sethi 2010; Watson *et al.* 2023; Hobday *et al.* 2025), but few have empirically tested the feasibility of such programs. Hobday *et al.* (2025) provides a compelling case for index insurance in fisheries as well as a comprehensive review of the potential barriers. My study provides an evaluation of a new formal risk mitigation tools that could be used by fishers to manage environmental risk.

Within the index insurance literature, there has been a growing application of machine learning to improve index design (Feng *et al.* 2019; Cesarini *et al.* 2021). Payout schedules limited to linear payouts introduce unnecessary basis risk that can be mitigated with neural networks

3

directly predicting the triggers (Chen *et al.* 2024). Machine learning models can also improve estimation of crops yields (Klompenburg *et al.* 2020). Improved estimation of crop yields leads to better insurance performance (Schmidt *et al.* 2022). My study further demonstrates the value of machine learning for index design, but reinforces the challenge of communicating complex models to policyholders and additional considerations that influence pricing.

The application of machine learning is growing in fisheries as researchers explore data questions beyond formal stock assessments. Neural networks are able to predict fishing activity in the global fleet using vessel GPS data (Souza *et al.* 2016). Ensemble models built through combinations of random forests, boosted trees, and dynamic linear models improve Bristol Bay sockeye salmon forecasts by 15% compared to a standard lagged regression model (Ovando *et al.* 2022). Environmental variables of importance to groundfish populations in Alaska were uncovered using single index varying coefficient models regularized with LASSO (Correia 2021). Random Forests models better predict fish catch in Indonesia than traditional linear models (Rahman *et al.* 2022). Random Forests models have also been used to predict fishery catch in the California Market Squid Fishery to isolate important biological and ecological variables (Akselrud 2024). My study builds on this literature by demonstrating the value of machine learning for fisheries even in small data sets.

The rest of the paper is structured as follows. Section 2 describes the design of index insurance contracts used in this study as well as demonstrating how imperfect prediction models lead to basis risk. Section 3 describes the algorithms used to predict fishery productivity, and how variables of importance were identified. Section 4 describes the characteristics of the market squid fishery that make it suitable for index insurance. Section 5 includes the data collection, transformations, and sources. Fisheries data comes from newly open-access sources provided by the California Department of Fish and Wildlife. Weather variables are collected from a variety of sources to explore a wide range of potential indices. Section 6 presents the primary findings of the paper including the performance of all models. Future steps and considerations for broader implementation of fishery index insurance are outlined in Section 7.

2 Insurance Framework

Index insurance is designed to protect policyholders from suffering losses below critical yield thresholds. In fisheries index insurance, actuaries must first estimate how fishery harvest is affected by weather variables. For example, fisheries harvest, y can be modeled with a simple technology function where fisher inputs, x interact with a stock of biomass, B, which is a function of random weather shocks, $\tilde{w}$ as shown in Equation 1:

$$y = f(x, B(\tilde{w})) + \epsilon \tag{1}$$

The harvest technology function, $f(x, B(\tilde{w}))$, is increasing and concave in inputs and biomass. All other idiosyncratic shocks to harvest that are not biological are captured in ϵ. Assuming

inputs are selected optimally conditional on the expected biomass, it is clearer to show the effects of weather on harvest strictly in terms of the random variable $\tilde{w}$:

$$y = f(B(\tilde{w})) + \epsilon \tag{2}$$

The weather and yield relationship, $f(B(\tilde{w}))$ can be empirically estimated by $g(\tilde{w})$, with a model error term η[1]. Subbing the estimator into Equation 2 gives:

$$y = g(\tilde{w}) + \eta + \epsilon \tag{3}$$

Formally, basis risk is the combined stochastic effects of η and ϵ. The model error, η, arises from the inability to perfectly estimate the weather to yield relationship. The idiosyncratic error, ϵ, arises from all other shocks to yield that are not related to weather. Well designed contracts minimize the negative effects of design risk stemming from η. For example, a perfectly estimated weather to harvest relationship would minimize η to zero.

The estimation of $g(\tilde{w})$ is extremely difficult in fisheries. However, the estimation does not need to be perfect or even statistically significant to provide value to fishers. The model only needs to capture enough of the weather to harvest relationship to provide sufficient protection with the insurance contract.

Structural forms of all models are presented in detail in Section 3. For now, the general model output will be represented as $\hat{y}(w)$ stemming from estimator $\hat{g}(w)$, where w is the observed weather variables and $\hat{y}$ is the predicted harvest.

Index insurance contracts are designed as put options in Equation 4 that initiate payouts once a predicted models' estimates, $\hat{y}(w)$, fall below a critical yield threshold, $c * \mathbb{E}[y]$.

$$I(w) = \max(c * \mathbb{E}[y] - \hat{y}(w), 0) \tag{4}$$

Equation 4 indicates fishers will receive the difference between the long run average of catch and the model predicted outcome from observed weather variables. The coverage level, c, represents the level of protection offered by the contract. The main results of the paper set coverage to 100%, where any predicted catch below long run average triggers a payout. In the appendix, multiple coverage levels are tested at higher levels to mimic disaster level thresholds..

Premiums are calculated by determining the expected payout of the contract times a premium loading factor, m as shown in equation Equation 5.

---

[1]Fisheries scientists might argue that you should first estimate $b(\tilde{w})$ and then estimate the relationship between yield and biomass. However, that would introduce an additional model error term and is not feasible for the case study. Market squid have short lifespans. Estimating their abundance is not available. Therefore the data does not exist in our context. Also insurance companies do not have access to the confidential data sources needed to build catch per unit effort and thus biomass estimates.

$$\rho(w) = \mathbb{E}[I(w)]m \tag{5}$$

I use burn rate analysis to calculate the expected payout in Equation 5. Burn rate analysis uses past realizations of weather variables to create a distribution of payouts that would have been observed if the policy was active over that time. The mean of the payout distribution is used as the statistic for $\mathbb{E}[I(w)]$. Burn rate analysis is a common method used by actuaries to price index insurance contracts, and is suitable in this setting because of the short time period of analysis (Jewson *et al.* 2005).

2.1 Utility evaluation

Evaluating the performance of index insurance contracts through a utility framework provides more accurate measures for the impact of basis risk on fishers willingness to pay for insurance (Conradt *et al.* 2015; Kenduiywo *et al.* 2021). Extremely risk averse policyholders will avoid insurance all together with any level of basis risk as the downside risk of not receiving payouts when necessary outweighs the benefits of having insurance (Clarke 2016). Therefore, assessing contracts based on utility will provide more information on fisher welfare especially in extreme events than simply comparing the incidence of false negatives.

I calculate expected utility with, u_i and without insurance, u_{ni}, in the test set as:

$$
\begin{aligned}
u_i &= \frac{1}{n_{test}} \sum_{t}^{n_{test}} u(y_t + I_t(w_t) - \rho_t) \\
u_{ni} &= \frac{1}{n_{test}} \sum_{t}^{n_{test}} u(y_t)
\end{aligned}
\tag{6}
$$

Where y_t is the observed harvest per fisher in year t, $I_t(w_t)$ is the payout from the insurance contract based on weather variables w_t, and ρ_t is the premium paid for insurance. The number of years in the test set determines n_{test}. The utility function, $u()$, is an increasing, concave function with risk aversion. I use a negative exponential utility function with constant absolute risk aversion in the main results of the paper. Constant absolute risk aversion is preferred in this case as it remains defined in the event of negative net income arising from large premiums and low harvest. Other utility specifications with constant relative risk aversion are included in the appendix. The expected utility with insurance, u_i, includes the payouts from the insurance contract minus the premiums paid. The expected utility without insurance, u_{ni}, is the utility from the catch.

I determine whether fishers are better off with insurance by calculating the percent change in utility from having insurance in Equation 7:

$$u_{rr} = \frac{u_i - u_{ni}}{u_{ni}} \tag{7}$$

Positive values of u_{rr} indicate fishers are better off with insurance, whereas negative values indicate fishers are worse off with insurance.

I use two loading factors, m to determine the viability of insurance contracts. As a base, I use actuarially fair premiums where the premium equals the expected payout of the contract i.e. $\rho = \mathbb{E}[I(w)]$.

Second, I calculate the premium loading factor that makes fishers indifferent between having insurance and not having insurance where $u_{rr} = 0$ as m^*. The premium loading factor is the multiple of the expected payout that insurance companies charge to cover administrative costs and profit. Varying m only affects the premiums paid. The payout schedule is solely determined empirically by the models. Values of m^* greater than 1 indicate fishers are willing to pay more than the expected payout for insurance, whereas values below 1 indicate subsidies will be needed to stimulate demand.

m^* will be called the "market" premium as it represents the upper bound an insurance company could charge for insurance. I calculate the loss ratio, $lr = \sum I(\omega) / \sum \rho$, at both actuarially fair and market premiums to analyze the market viability of a contract. Loss ratios above 1 indicate insurance companies lost money on the contract, whereas loss ratios below 1 indicate insurance companies are profitable. If a contract at m^* leads to $lr < 1$, then the contract is viable in a free market. Insurers can use any m less than the market m^* and still encourage fishers to purchase insurance.

3 Methods

I use five models to predict yearly catch $\hat{y}(w)$ that serves as the index within Equation 4. Simple linear models are the base model given its ubiquitous use in index insurance policies. Simple linear models create single index contracts that use only one weather variable. I refer to all simple linear regression models as "univariate" to differentiate between the multivariate models. The four multivariate models will be LASSO regression, random forest, regularized random forest, and support vector machine as predictor models.

I use a rolling window cross validation method to assess the out of sample predictive quality of each model, evaluate utility with insurance in the test set, and tune hyperparameters of the multivariate models (Ovando *et al.* 2022; Akselrud 2024). The data is split into an initial training window from 1990-2013 to provide enough data to accurately calibrate each model. From the trained model, a single prediction is made with environmental data from the next year. Whether the contract pays out according to Equation 4 is recorded. The premium for that year is calculated using burn rate pricing on the training set. In the next period, the training window rolls forward one year to include 1990-2014. The model is retrained and a

prediction is made for 2015. This procedure continues until the end of the data set in 2023. Expected utility is the calculated for the test set based on the payout schedule and premiums calculated as in Equation 6.

The rolling window validation strategy is well suited for time series data where temporal autocorrelation may be present. I do not provide a validation window for parameter tuning as the time series is short. Withholding observations may lead to more overfitting whereas I can use the test set to tune hyperparameters. Additionally, the rolling window mimics policywriters' strategies where they update their actuarial models each year to provide the most accurate pricing information.

Root mean squared error will serve as the metric to evaluate out of sample predictive performance of each model in the hyperparameter tuning.

3.1 Univariate Models

Perfect regression coefficients mimic the optimal choice of scale in index insurance contracts (Mahul 1999). Combined with the ease of implementation, simple linear models on single weather indices are the most common design choice for index insurance policies (Benso *et al.* 2023). Regression models offer a basic starting place to consider the viability of fishery index insurance.

Annual catch is regressed on each environmental variable individually using Ordinary Least Squares (OLS) as shown in Equation 8.

$$y_t = \beta_0 + \beta w_t + \epsilon_t \tag{8}$$

Linear models are preferred in index insurance because they provide a simple, transparent relationship between the weather variable and catch. For example, it is possible to show the output of a contract built on linear regression in terms of the weather variable directly in Equation 9.

$$I(w) = \max(\beta_1 c * (\mathbb{E}[y] - \frac{\beta_0}{\beta_1}) - w, 0) \tag{9}$$

The trigger, $(\mathbb{E}[y] - \frac{\beta_0}{\beta_1})$ is now expressed directly in units of the weather variable while still interpreted as a yield threshold. This structure allows for clear communication to fishers about payouts. For example, if sea surface temperature is the index, then the trigger can be expressed as 12 degrees Celsius. Any realized temperature, w, below 12 degrees Celsius would initiate a payout.

3.2 LASSO Regression

Least Absolute Shrinkage and Selection Operator (LASSO) regression is a popular regularization technique to assist model selection. It attempts to minimize the residual sum of squared errors through OLS, but adds a penalty constraint on the absolute sum of selected coefficient values (Equation 10).

$$\hat{\beta}^{lasso} = \arg\min_{\beta} \left\{ \sum_{i=1}^{n}(y_i - \beta_0 - \sum_{j=1}^{p} w_{ij}\beta_j)^2 + \lambda \sum_{j=1}^{p} |\beta_j| \right\} \tag{10}$$

Where, y_i is catch, β the regression coefficients, n, the number of observations, p the number of weather variables, and w the total collection of weather variables. The λ is the penalty term that controls the amount of shrinkage. Models are trained using the `glmnet` package in R. The λ that minimizes RMSE is selected through a grid search method. This choice is to ensure the most parsimonious model that still captures the most important weather variables. Reducing the number of variables used to define a trigger payout will simplify the contract substantially.

3.3 Random Forests

Random Forests are tree-based ensemble models that capture non-linear interactions through recursive partitioning. They are less sensitive to over fitting through the aggregation of many trees. Each tree is built on a bootstrapped sample of the training data. At each split, a random subset of predictor variables is chosen to determine the best split. The final prediction is the average of all the trees in the forest. Random forests have been shown to outperform traditional regression models in fisheries applications (Ovando *et al.* 2022; Rahman *et al.* 2022; Akselrud 2024).

The high dimmensionality of the data and few observations will make random forests susceptible to overfitting. I address this high p low n problem through feature selection and regularized random forests. Certain weather variables are likely to be highly correlated with each other. I select the indices with the highest correlation to catch for five categories of variables: Temperature, regional indices, upwelling, prey, and paralarvae abundance.

Regularized random forest provides an agnostic, data driven methodology for feature selection. Regularized random forest computes permutation importance scores for each variable during out of bag testing. If the increase in overall performance for a given variable is lower than a specific threshold, it is removed from the model. Similar to LASSO, it allows for an unbiased assessment of important variables. I can compare whether my feature selection on expertise is accurate by examining the variables selected by the regularized random forest.

I tune three parameters to optimize the random forest model. The number of variables randomly sampled at each split, *mtry*, the minimum number of observations in a terminal node,

nodesize, and the maximum depth of each tree. I manually construct a parameter grid to search for the combination of parameters that minimize RMSE in the test set. Max depth is set to small values, while *nodesize* has higher than default thresholds to reduce overfitting bias. The small number of observations in the data set ($n = 34$ years) necessitates more conservative parameter choices. Models are built using the `ranger` package in R.

Regularized random forest requires one additional tuning parameter along with the three used in standard random forests. The regularization parameter, γ_{rrf}, controls the threshold a variable will be added contingent on the permutation score. It is normalized from 0-1 with 1 acting as the most aggressive regularization. Models are built with `grrf` package in R.

3.4 Support Vector Macine

Support Vector Machine (SVM) constructs vectors that define an error insenstive margin when used for regression. The algorithm then minimizes a loss function on the slack points outside the support vectors. I use a radial basis function kernel to map the predictors into a higher-dimensional feature space, enabling the algorithm to capture nonlinear patterns without explicitly defining transformations of the input variables. Support Vector Machines work well in time series data with low observations and high dimensionality (Sapankevych and Sankar 2009).

Model complexity is controlled by the regularization parameter, C, which controls the width of the support vectors. A smaller C allows for a wider margin, placing more data points within the support vectors and enhances generalization. A larger C aims to minimize the error all training points correctly, which may lead to overfitting. The kernel coefficient, γ_{svm}, defines the influence of individual training examples. A small γ_{svm} implies a broader influence, leading to less sensitivity to outliers, while a large γ_{svm} focuses on nearby points, capturing more intricate patterns but risking overfitting. I use a grid search method to tune C and γ_{svm} to minimize RMSE in the test set. Models are built using the `kernlab` package in R.

3.5 Weather variables of importance

Machine learning algorithms are inherently "black boxes" that sacrifice interpretability for predictive accuracy. Fishers will be less likely to purchase complicated products that do not correspond to their experiences. Extracting the relative contribution of weather variables will assist translating products to fishers. Additionally, it can help ground-truth the chosen variables with previous biological modelling.

I extract importance measures using permutation for all the multivariate models. Permutation randomly modifies a variable's value then calculates the change in the model performance. As models are updated in each period, I will have n_{test} models each with their own permutation scores. I examine the importance of each variable in a model over the rolling training window.

4 Market Squid Fishery

The California Market Squid Fishery is an ideal candidate to assess the viability of index insurance due to its high economic value, simple management system, excellent data, and "corn-like" biological attributes.

Market squid is the second most valuable in California with an average annual ex-vessel value of $48 million since 2013. By landing weight, it is the largest fishery in California with total catch exceeding half a million tons since 2013. Fishing predominately occurs in the Santa Barbara Channel during October and November. It is vulnerable to large fluctuations in catch, which threatens livelihoods during disastrous years. Index insurance would offer much needed stabilization to fishers in this volatile fishery.

The fishery has been consistently managed since 2005 with a permit entry system and a maximum annual catch limit of 118,000 metric tons. Prior to 2005, the fishery was open access with fishers only paying for a licence and a landing tax. The number of licensed fishers dramatically declined in the implementation of the permit system in 2005. We adjust for the change in permits by using the catch per fisher as the outcome variable to allow comparison between pre-2005 and post-2005 management regimes.

Fishery catch data is publicly available from the California Department of Fish and Wildlife (CDFW) Marine Fisheries Data Explorer (MFDE). The data contains catch records from 1980 to 2024 at the port-complex and state level. Additionally, the data contains spatial records of catch with specified fishing blocks that can be used to refine the location of catch. I use this information to spatially match environmental variables as shown in Figure 1 for sea surface temperature. This procedure ensures that only the most relevant environmental data is used to predict fishery harvest in the models.

Market squid is a short lived species with a one year life cycle (Butler and Fuller 1999). Sea surface temperature and upwelling primarily affect recruitment leading to large fluctuations in abundance by the time harvest occurs in the fall (Chasco *et al.* 2022; Suca *et al.* 2022). Year to year variation does not appear to persist removing dynamic considerations. This leads market squid to possess "corn-like" attributes akin to crops protected in agricultural index insurance. Planting cycles each year are independent from prior weather outcomes. A similar logic holds for market squid and allows the analysis to focus solely on contemporaneous weather effects.

Other independent variables demonstrate correlations to squid abundance. The Rockfish Recruitment and Ecosystem Assessment Survey (RREAS), conducted by the Southwest Fishery Science Center, provides independent surveys that collect tow samples that contain squid paralarvae and krill. Each of these metrics could be used by insurance companies to index fishery catch. Krill abundances are significantly correlated with fall landings (Ralston *et al.* 2018). Paralarvae measures were the most important variable in random forest regressions to predict squid fishery catch (Akselrud 2024). These variables are also indicative of the potential to include non-weather variables that could serve as triggers.
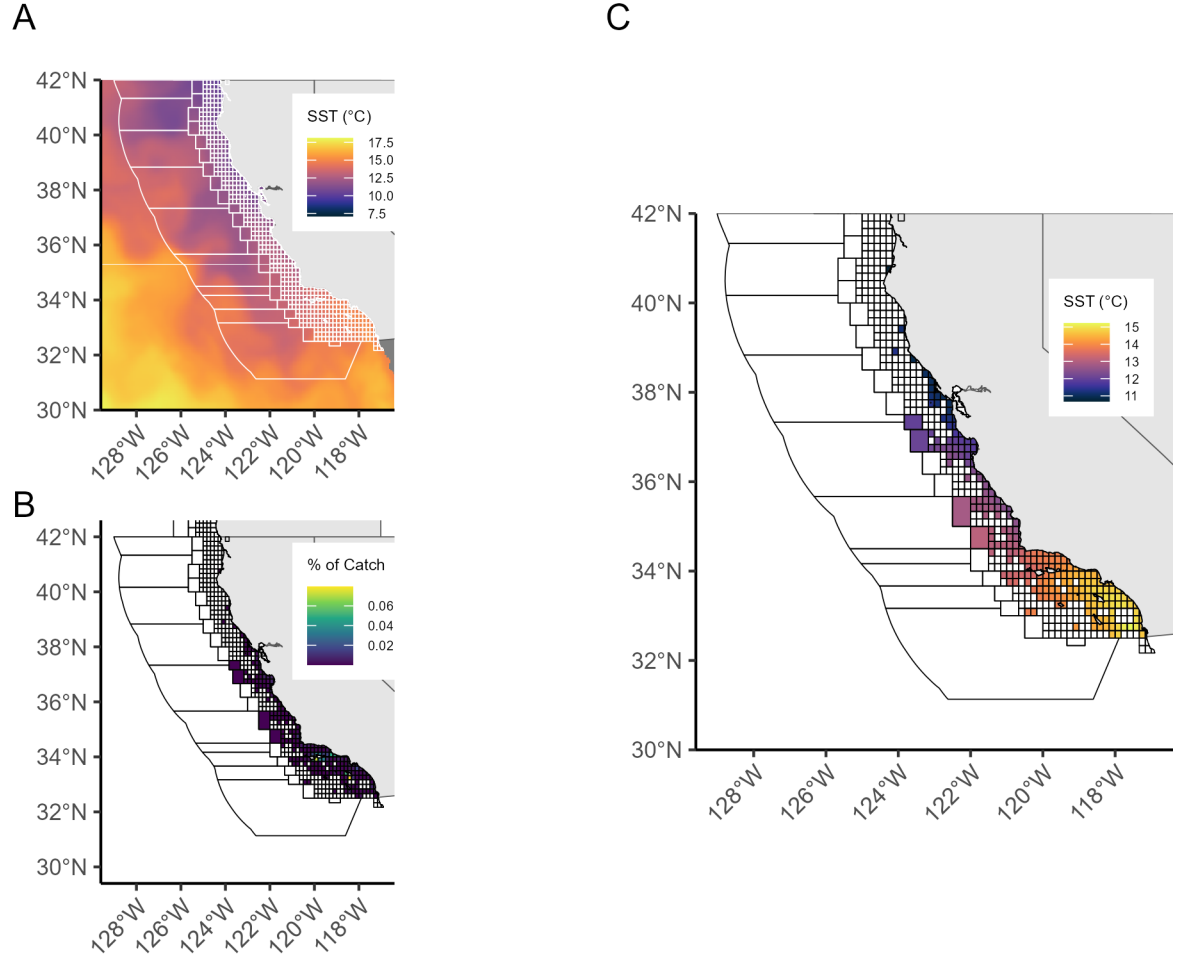
Figure 1: Demonstration of spatially matching catch data with weather variables. Panel A shows sea surface temperature at a 5km resolution. Panel B is the proportion of historical market squid catch by fishing block. Panel C is the intersection of fishing grids and weather variables. Sea surface temperature values are then averaged within each block, and the weighted sum by percent of historical catch is aggregated for each year.

5 Data

The dataset contains 34 annual observations (1990-2023) of squid catch per fisher and 12 weather variables. Catch histories extend to 1980, but to ensure a balanced dataset I limit observations to 1990, which is the earliest data available for squid and krill abundance in RREAS.

Fishery catch is publicly available through the California Department of Fish and Wildlife Marine Data Explorer. State wide catch is divided by the number of fishers to create metric tons of catch per fisher. Catch per fisher is detrened using a robust method of moments estimator. Detrending removes long term trends such as improved fishing technology, and any potential effects of the management shift in 2005 so that deviations from long run average are more directly attributable to weather fluctuations. Observations are scaled relative to the last observed value to avoid negative values.

Environmental data to define indices are summarized in Table A1. Descriptions of the sources of data and calculations are provided below.

Sea surface temperature data comes from the NOAA DHW that provides 5-km resolution of monthly temperature from 1985 to 2023. The 5-km grids are aggregated within the nearest California fishing block to provide an annual time series.

The Habitat Compression Index measures the area extent of water below average temperatures thresholds along the US West Coast (Schroeder *et al.* 2022). Habitat compression is a measure of the spatial extent of cold water habitats that are important for fish species. The index is broken down into four distinct oceanographic regions ranging from 3.5 degrees to 5.5 degrees latitude in size with coverage out to 150 km offshore. I use the cumulative habitat compression index that sums the index value in each month to provide a yearly time series of habitat compression for each fishery. The cumulative index showed stronger correlations with biological productivity measures than monthly measures (Schroeder *et al.* 2022)

I use five different sources to define the strength of upwelling. Monthly observations of Coastal Upwelling Transport Index (CUTI) and Biological Effective Upwelling Transport Index (BEUTI) are averaged over the spring months of March, April, and May. Those months align most closely with squid egg mixing and dispersion. Both indices create measures of vertical movement in the mixed layer at 1 degree latitude intervals extending 75 km along the entire US West Coast (Jacox *et al.* 2018). The closest layer to the surface was used in this analysis as the correlation between surface index values and deeper index values are high. CUTI examines the physical measures of wind, Ekman transport, and cross-shore geostrophic transport to indicate the strength of upwelling in a given month. BEUTI adds nitrate concentration in its calculation to capture more biological effects of upwelling. Fishing blocks are matched to the nearest 1 degree latitude interval to provide a monthly time series of upwelling for each fishery. Seasonal strengths of upwelling are captured by averaging CUTI and BEUTI within each quarter of the year.

In addition to direct measurements of upwelling, I characterize three upwelling events: Spring Transition Index, Relaxation, and Frequency following methods from Suca *et al.* (2022). Each variable was a significant predictor in Suca *et al.* (2022) recruitment models. Spring transition is the date when the California Current transitions to a more stable upwelling system after the turbulent winter months. Mathematically, it is the first day of the year when the cumulative sum of BEUTI reaches its minimum value. Relaxation is the count of the number of days when CUTI exceeds 1.0 $m^2 s^{-1}$ followed by three or more consecutive days of CUTI values <0.5 $m^2 s^{-1}$. Frequency is the proportion of days when CUTI is measured between 2.0 and 0.5 $m^2 s^{-1}$. Frequency is indicative of moderate upwelling events that stimulate primary production without pushing production away from the coast.

Broader climate conditions have also been observed to impact squid abundance. The Pacific Decadal Oscillation (PDO) and El Nino Southern Oscillation (ENSO) have been linked to squid abundance (Perretti and Sedarat 2016). ENSO has been used effectively in crop index insurance in Peru. Larger climatic patterns may also be influential in squid recruitment. PDO data is taken from the PDO ERSST V5 dataset. One measure of ENSO data is taken from the multivariate ENSO Index Version 2 (MEI.v2). The last measure is the Oceanic Nino Index (ONI) that is the 3 month running mean of ERSST.v5 SST anomalies in the Nino 3.4 region.

Squid paralarvae and krill abundance are calculated from the Rockfish Recruitment and Ecosystem Assessment Survey (RREAS). Midwater trawls sample yearly at various stations along the California coast. I use delta-generalized linear models to estimate log catch per unit effort of paralarvae and krill for the whole state in each year[2]. Catch per unit effort serves as a proxy for prey and squid paralarvae abundance.

6 Results

The multivariate models better predict weather-harvest dependency than univariate models (Figure 2). However, overfitting is a concern especially for the multivariate models. All multivariate models had lower RMSE than every univariate linear model in the training set (Panel A). In the test set (Panel B), multivariate models remained more accurate as a group, but the discrepancy between the types of models was much smaller. Support vector machine models had the lowest out of sample RMSE for the multivariate models followed by random forests, and then LASSO.

The best out of sample predictive model was frequency of upwelling events with the lowest RMSE out of all models at 265 mt per fisher. Two other groups of single weather indices performed well. Climate regional measures, (PDO, ENSO, ONI) had lower error in both the

---

[2]RREAS does not have consistent temporal or spatial coverage. From 1983-2003 all surveys were done in a narrow window near Monterrey Bay. In 2004, more surveys locations were standardized into the Southern California Blight. Surveys were not conducted in 2014-2016 for the southern stations due to budget constraints. Using catch per unit effort for all stations provides a useful statewide proxy. Future use to make estimates more port specific would be valuable, but need to adjust for the data gaps

test and training sets than most other single variables. Krill abundance has the lowest RMSE in the training set, but had the highest RMSE in the test set out of all models.

Examining fisher utility provides clearer measures of whether a model sufficiently protects fishers from environmental risk. First, I present the results of the contracts with actuarilly fair premiums, then select only models that improved utility to calculate the market premium.

The majority of univariate models made fishers worse off as the fishers experienced high degrees of downside risk where they did not receive payouts when necessary (Figure 3 Panel A). Only three weather indices improved fisher utility: Upwelling frequency (35%), relaxation of upwelling (12%), and krill abundance (8%). However, frequency and relaxation were highly unprofitable for insurance companies as each had a loss ratio greater than 250% (Figure 3 Panel B), which implies that insurance companies would have to pay out $2.50 for every $1 of premium collected. The upwelling indices paid large sums at times when fishers did not experience loss. The large payouts act more as wealth transfers than risk protection. Krill remained profitable for insurance companies with a loss ratio of 55% at actuarilly fair rates.

Every multivariate model improved fisher utility while reducing the insurance company loss ratio (Figure 3) compared to the univariate models. Contracts built with SVM and LASSO both improved fisher utility by 37% and 28% respectively at actuarilly fair rates, but each led to large losses for the insurance companies. Insurance companies remained solvent with both random forest models, and fishers experienced 8-14% gains in utility.

Figure 4 shows the market premium an insurance company could charge at the upper limit of demand for each model in Panel A. The ensuing loss ratio at that market premium is shown in Panel B. Even at the upper limit of demand, insurance companies would not be profitable with the upwelling models and LASSO. The loss ratios remain above 1 implying insurance companies paid more out than they received at the market premium. Fishers are most willing to pay for a contract with the support vector machine as it has the highest premium. Insurance companies would be willing to provide contracts at the market premiums because the loss ratio is 0.75. Random forest contracts are also profitable for insurance companies, though fishers are slightly less willing to pay for them. The accuracy of random forest to avoid false positives helps insurance companies avoid excessive payouts, allowing insurers to remain profitable even at lower market premiums.

Next, I examine the payout and premium schedules over the testing period to understand how contracts improve utility. Univariate models payout less frequently (Figure 5 Right Panel). The krill index paid out once during the lowest catch year in 2019. The large one, time payout provides a significant gain in utility. Due to fisher risk aversion, fishers are more willing to pay higher premiums to protect against that single devastating loss (Figure 6). The upwelling indices pay more in later years. The frequency index pays out every year from 2018-2023 with the largest payout in 2020 (197 mt per fisher). These payout schedules by chance correspond to COVID-19 disruptions. However, the losses experienced by fishers are not driven by weather risk, but rather a coincidence that the upwelling indices predicted low catch during those years.
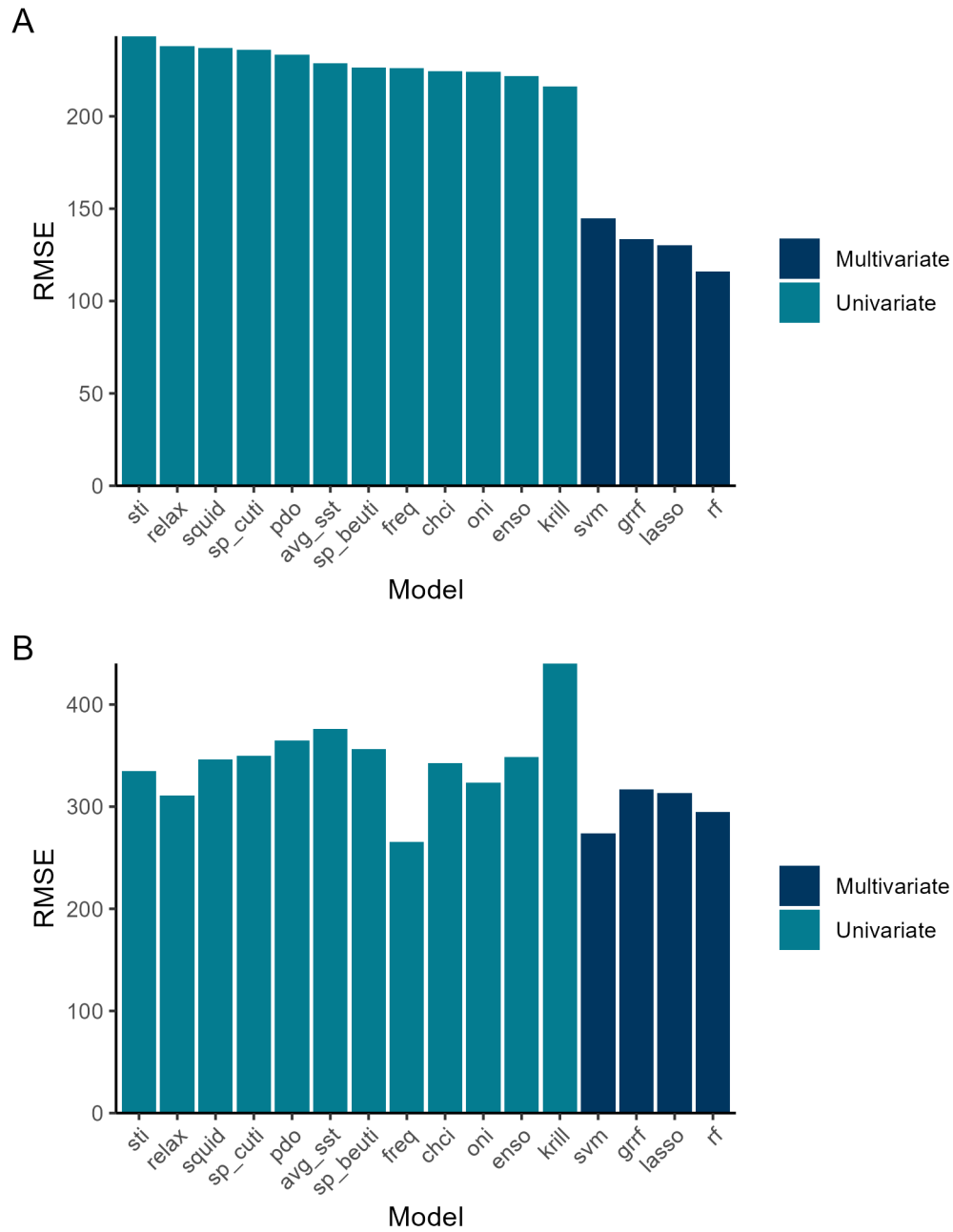
Figure 2: The Root Mean Square Error of weather models used to calculate insurance contracts. Panel A depicts the average RMSE across all training set windows. Panel B is the total RMSE in the test set using individual predictions from the rolling window. Model types are separated into single index univariate models (blue) and multivariate models (green)
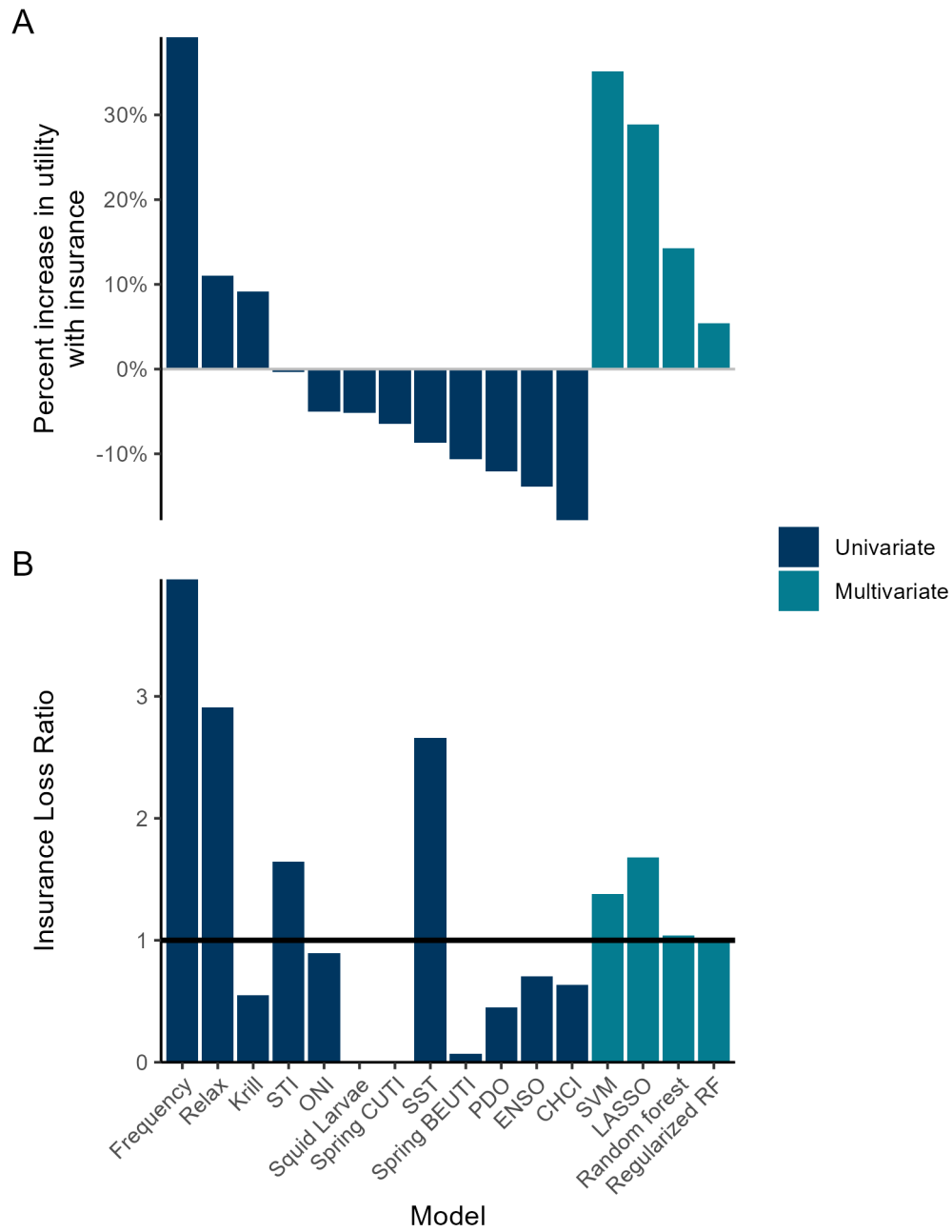
Figure 3: Percent increase in utility with insurance (Panel A) and the insurance loss ratio (Panel B) for each model using actuarially fair premiums. Models that fall above the horizontal line in Panel B are not profitable for insurance companies.
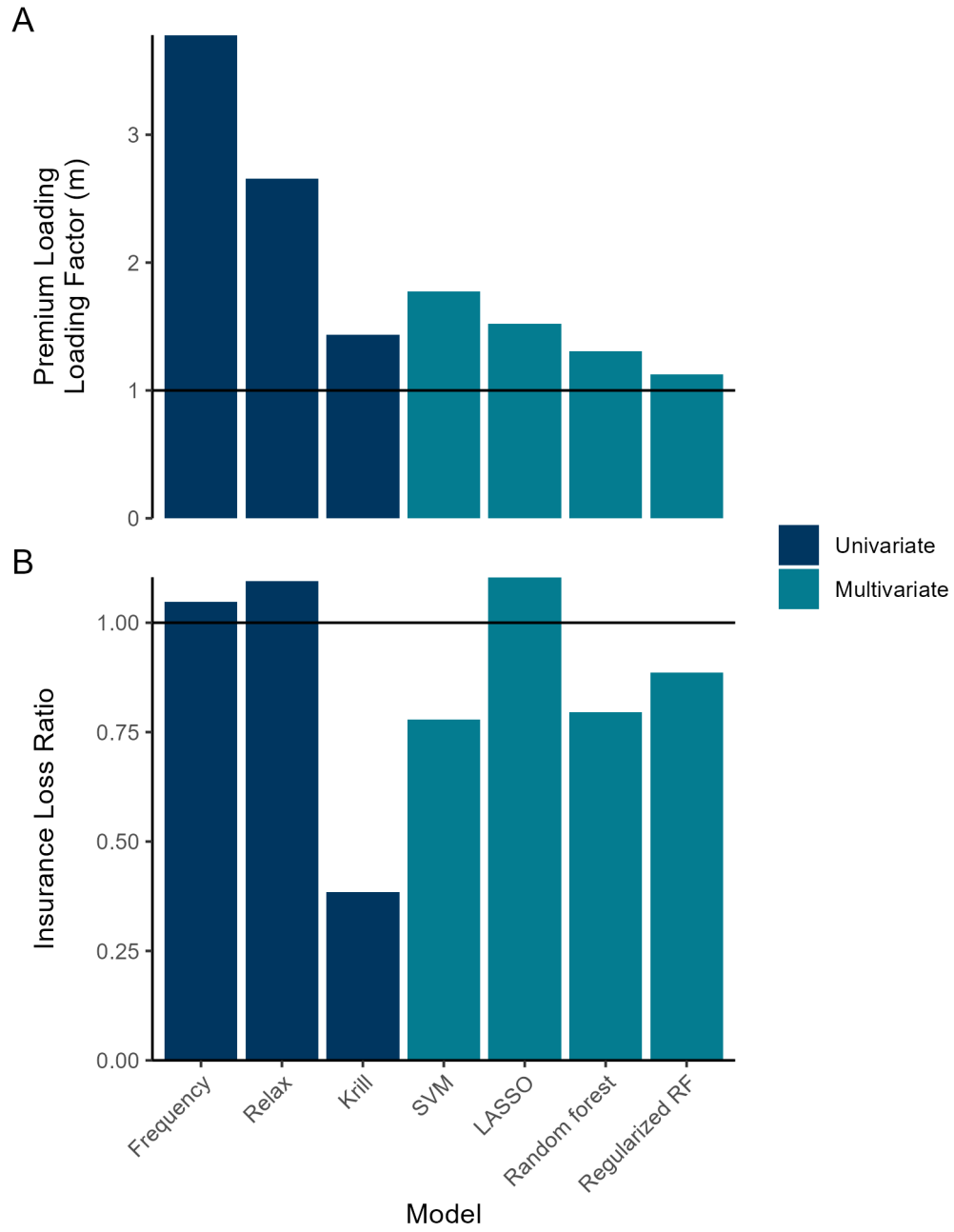
Figure 4: Market premiums rate (m^*) that make fishers indifferent between having insurance and not having insurance (Panel A). The corresponding loss ratio for contracts at the market rate is in Panel B.

The consistent large payouts lead insurance companies to significantly increase premiums as shown in Figure 6.

Multivariate predictive models initiate more frequent payouts, but payout larger quantities when fishers experience low catch (Figure 5). The flexibility of the underlying predictive models drives the payout schedule. However, with larger payouts, premiums must increase to compensate insurance suppliers. Premiums with multivariate models are high and remain consistent throughout the ten years of the test set even with updates each year (Figure 6). Contracts built with LASSO models differ the most from the other multivariate models. LASSO contracts pay out large sums in 2015 and 2023 when fishers experienced high catch. Insurance companies suffer losses in those years limiting profits even at the upper bound of market premiums.

The SVM and Random Forest contracts provide the largest utility gains while offering competitive market premiums. Overall, both models minimize downside risk for both fishers and insurance suppliers. There are two key differences in the performance of both models that are more easily observed when the net insurance payouts, payout minus the premium, is compared to the realization of fisher catch as shown in Figure 7. SVM contracts payout more during down years compared to random forests. Larger payouts in periods of higher loss provide large gains in utility. The higher payouts lead to the corresponding higher premiums.

Random forests may be more accurate in identifying when to payout as the amount of false positives and negatives favors random forests compared to SVM. The notable exception is 2020 and 2021. Fishers experienced low catch primarily due to COVID-19 closures and market disruptions. COVID-19 was entirely independent of weather variables. Random forests paid a small amount in 2020 (7 mt per fisher) as weather indicators predicted a mild year in 2020. However, SVM paid out a larger amount (58 mt per fisher) in 2020. The large payout in 2020 was not necessary to protect fishers from weather risk, but it did provide significant utility gains as fishers were inadvertently protected from the COVID-19 disruptions.

6.1 Variable Importance

Multivariate models are more accurate, improve fisher utility, and allow insurance companies to earn a profit. However, the payout schedule is determined by complicated, non-linear interactions. Frequency, ENSO, and krill were consistently the most important variables in the non-linear multivariate models Figure 8. Frequency became more influential across all models in later years. Sea surface temperature was important only within the random forests models.

I can evaluate ex-post whether my feature selection for random forests [3] was accurate by comparing the variables of importance in the regularized random forest to those in the standard random forest. All measures selected based on expertise were included within the regularized
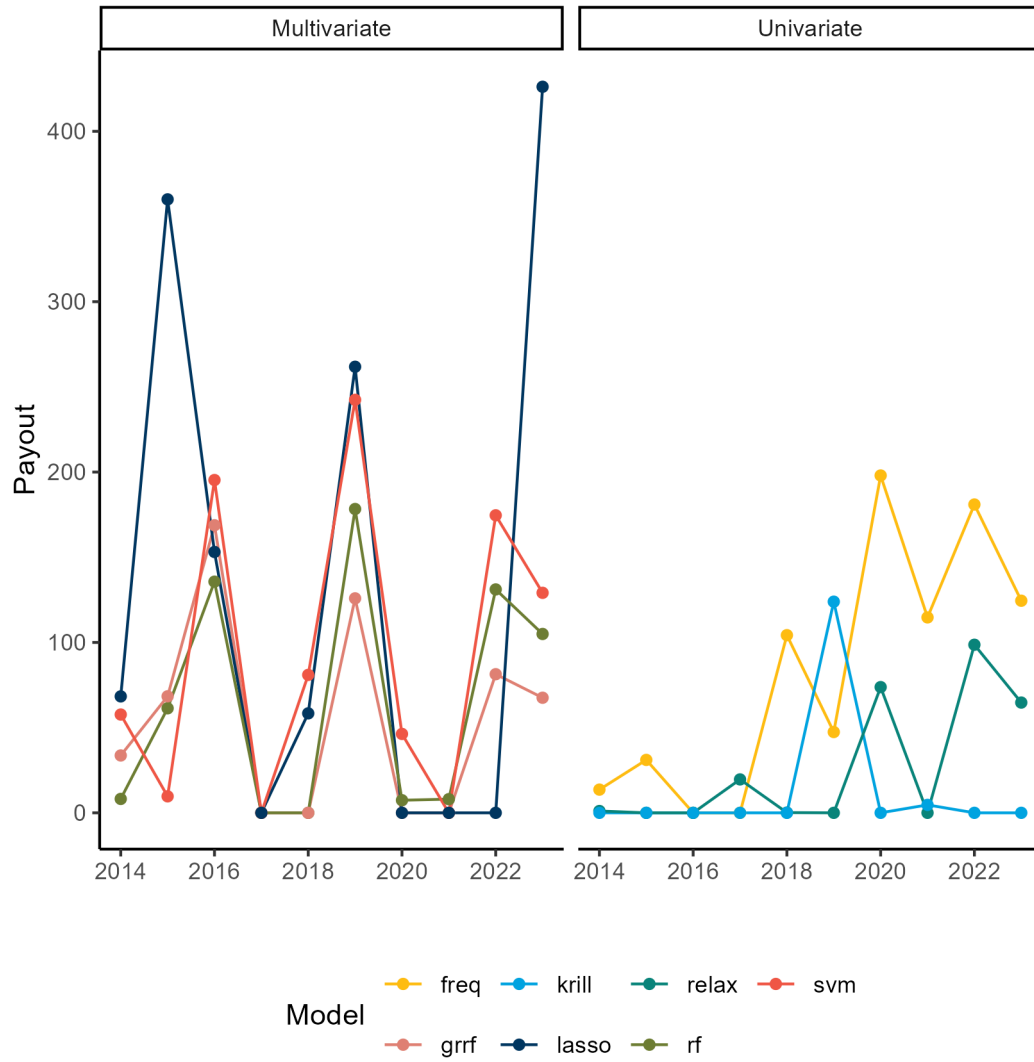
---

[3]And to an extent SVM

Figure 5: Payouts in the test set by each contract. Multivariate models (left panel) have more flexible payout schedules than utility improving univariate models (right panel). Individual models are colored by model type.
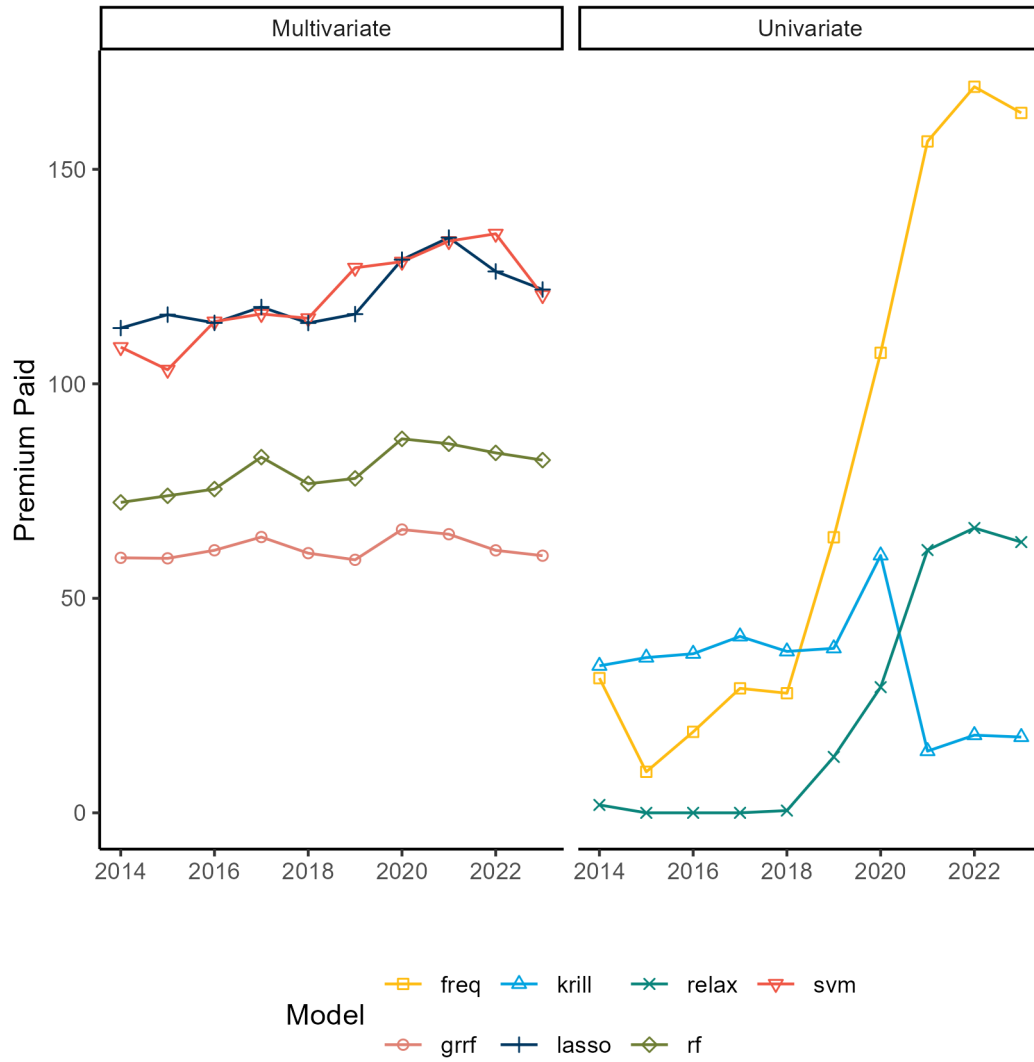
Figure 6: Premiums in the test set by each contract. Multivariate models (left panel) charge higher premiums to compensate for the more frequent payouts compared to univariate models (right panel).
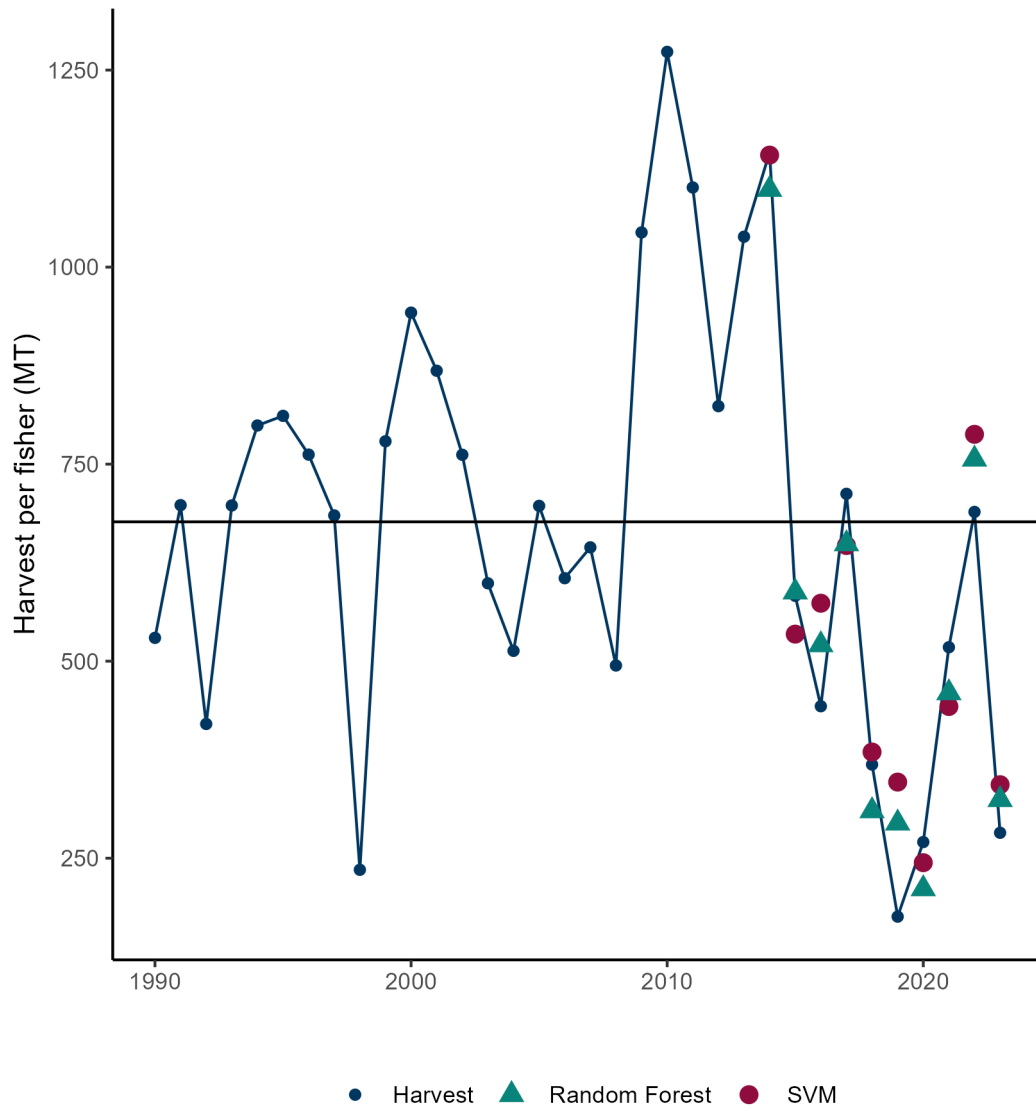
Figure 7: Net insurance payouts in the testing set (red points) from the SVM model and random forest (green points) compared to the detrended catch per fisher data from the California Market Squid Fishery (blue line and points). Insurance variables are in units of harvest per fisher.

random forest. The only missing information that was valuable in the regularized random forest were PDO and ONI. Perhaps more information could be captured with the inclusion of the other regional climate variables, but each are highly correlated and capture similar patterns. Overall, the feature selection based on correlation and expert knowledge was effective for both the random forest and SVM, and did not exclude any of the primary weather variables.
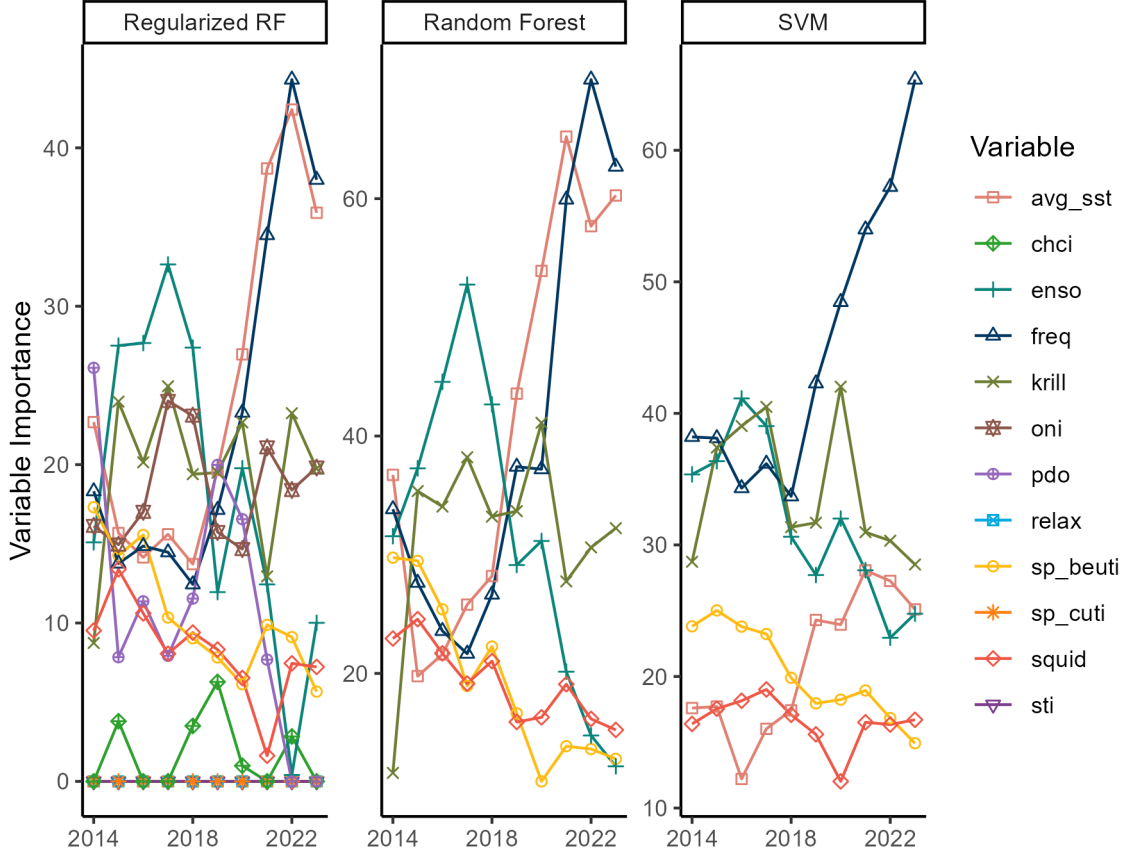

Figure 8: Variable importance from permutation for regularized random forest (left panel), random forest (middle panel), and SVM (right panel) in each iteration of the training window. Variables are colored and shaped by type.

LASSO coefficients can be aggregated and divided by the absolute total to get a measure of relative importance. Frequency is the most important variable in LASSO in all years comprising about 75% of the explanatory power of the model Figure 9. Spring CUTI was the second most important variable in all years followed by MEIv2 El Nino. Together, those three variables explained more the 90% of the models predictive power. A multiple linear regression contract with those three variables could increase predictive power beyond a single index while

remaining clearly interpretable. However, given the lack of profitability for the overall LASSO it is unlikely that such a model would be market viable. The remaining environmental variables do not provide much additional value to the model. Spring Transition Index, PDO, CHCI, and SST are dropped from the model during some training windows (Zoomed panel of Figure 9).
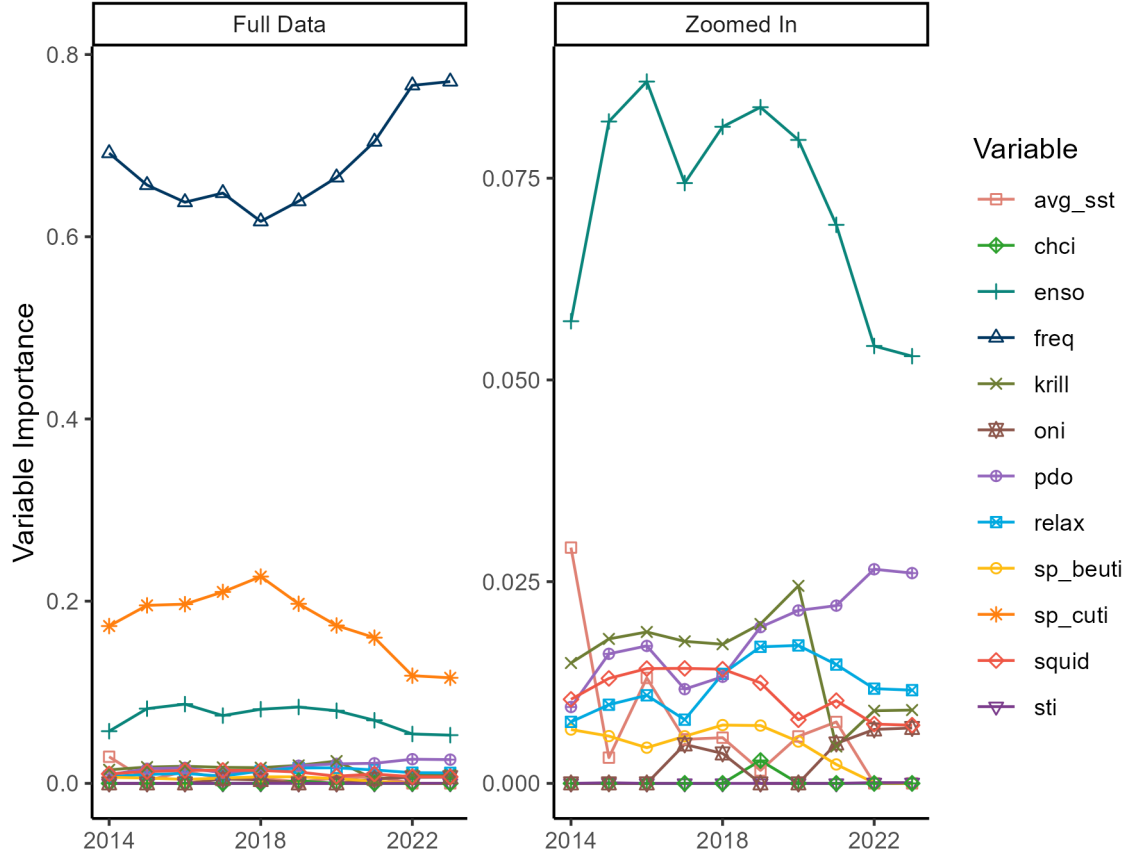
Figure 9: Relative importance of LASSO regression coefficients for each model trained in the training set. The right panel focuses on variables with less than 10% contribution.

7 Discussion

Index insurance is feasible for the California Market Squid Fishery. Contracts built with a single index may be the preferred option as they are easy to implement and have clear, interpretable triggers, but the inconsistency of payouts is undesirable for both fishers and insurers. Out of twelve single index models, only three would have improved fisher utility over

the last decade. Fishery productivity is rarely driven by individual weather patterns. It is unsurprising that using a single measure would be provide sufficient risk protection.

Despite the preference for single index models, the performance of single index models is lackluster even in agriculture. For example, in the United States as part of the Rangeland, Pasture, Forage - Rainfall Index (RPF-RI) program rainfall to production correlations may be as low as 0.07 (Keller and Saitone 2022). Significant subsidies are needed to support the RPF-RI due to the high levels of basis risk (Goodrich *et al.* 2019). Other agricultural IBI products do demonstrate improved utility for farmers (Conradt *et al.* 2015; Dalhaus *et al.* 2018; Kenduiywo *et al.* 2021), but few studies compare out of sample performance nor consider the supply side incentives of insurance companies (Chen *et al.* 2024). Without assessing supply side viability may contribute to the low uptake of index insurance in agriculture and high premiums (Binswanger-Mkhize 2012; Miranda and Farrin 2012). For example, Kenyan farmers gain utility at actuarilly fair rates for maize rainfall contracts, but lose utility at the commercial premiums (Jensen *et al.* 2016). Single index fishery contracts demonstrate similar patterns that could prohibit effective deployment.

Machine learning models have been promoted as a response to the failures of single index agricultural index insurance (Cesarini *et al.* 2021). Few studies have used machine learning to define weather-yield relationships, and as far as I am aware, no policy currently exists that uses multivariate machine learning models. Fishery machine learning models are market viable and provide large improvements to fisher utility. Fishers are willing to consider complex models so long as the basis risk remains low [Emlab Study]. The model variable importance scores can help elucidate the "black-box" of machine learning models. The variables that were most predictive as single indices remained important in the multivariate models. Upwelling frequency, relaxation, and krill abundance were key explanatory variables in all models. Even though it is impossible to define a clear trigger for each variables in these models, explaining that these intuitive variables contribute the most to the determination of payouts will entrust fishers to the contracts.

Models with high basis risk often lead to high premiums. However, in this paper, the premiums fishers paid for multivariate models was higher than single index models. This phenomena originates from a unique form of data leakage that does not interfere with model performance, but does influence premium calculation. Nonlinear machine learning models accurately identify extreme losses in the training sets. Payouts in the training set will be much higher leading to more expensive premiums. Fishers are willing to pay these higher premiums as the multivariate models accurately pay out large sums in bad times. However, this does emphasize the need to avoid overfitting as inaccurate out of sample predictions will lead fishers to pay high premiums without appropriate compensation when needed. All index insurance contracts that plan to use machine learning to identify weather to yield relationships will be affected by this phenomena.

Hyperparameter tuning and feature selection are essential to reduce overfitting. The rolling window strategy used in this paper demonstrates an effective way to limit overfitting and data leakage. Feature selection requires expertise in identifying potential indices that affect catch.

The choice of fishery is an essential component to expand index insurance. Each fishery is unique. Different weather variables will be appropriate for fisheries outside of California Market Squid. Diversifying the weather variable pool, spatially and ecologically, will allow insurance companies to diversify their risk more seamlessly than fishers are able. For example, insurance companies could create contracts for sea surface temperature in different ocean basins. The risk of a marine heatwave to strike both the California Current Ecosystem and the Gulf of Mexico is low. Additionally, species of fish react differently to weather variables (Free *et al.* 2019, 2023). Offering contracts that cover different species will spread out risk for insurance companies to avoid fund insolvency.

The selection criteria used in this paper to justify choosing California Market Squid for analysis is not restrictive, but does demonstrate some considerations that must be assessed when selecting fisheries for index insurance.

Fast growing species with short life cycles are more likely to be influenced by contemporaneous weather variables, and thus easier to empirically identify weather to catch relationships. However, longer lived species may also be insurable. The most important consideration is whether an environmental variable that influences catch can be identified. For example, Dungeness Crab have a four year life cycle. Environmental variables during the megalopae stage affect adult crab abundance and therefore harvest four years later (Shanks *et al.* 2010; Magel *et al.* 2020). If the variables are sufficiently correlated with catch, then an insurance contract could designed based on weather conditions four years prior.

An interesting consideration for the contract would be the timing of the payout. Should the contract pay out when the weather event occurs, or when fishers experience low catch? Paying out when the weather event occurs would provide fishers with more time to adjust their fishing strategies. However, the longer the lag between the weather event and the payout, the more likely it is that other confounding factors could influence catch.

Management provides the most interesting considerations. As demonstrated in Section 4, well managed fisheries often have the most robust data available. However, the change in regulations influence harvest more directly than weather stochasticity. Empirically disentangling weather to catch relationships with changes in regulations is difficult.

Well managed fisheries attempt to smooth biological risk by avoiding boom and bust cycles. Fishers remain vulnerable to fluctuations in management decisions such as quotas, season closures, or lost fishing days. These actions more directly impact fisher livelihoods than weather. Instead of designing contracts solely on weather variables, a new category of indices directly related to fishery management could be used. For example, quotas are set based on the underlying health of the fishery. An index insurance contract could use the quota as the index. This will provide a direct measure of lost catch for fishers while remaining an independent measure. Currently, insurance companies have demonstrated reticence on using management measures as they perceive management actions to introduce human influence. However, quotas are set based on rigorous scientific assessment outside the influence of fisher actions. The index would

remain independent of policyholder influence, and is calculated in a transparent, objective manner.

Other triggers that originate from management could also be used. Dungeness crab fishers stated interest in insuring against the number of days the fishing season is delayed by management. Preseason Salmon run strength predictions could be another. The exact nature and suitability require rigorous assessment and cooperation with management to ensure independence.

This paper demonstrates fisheries index insurance is feasible in a private market, but one-to-one application of techniques used in agriculture is unlikely to be successful. Insurance companies currently lack the expertise to navigate the complexity of fisheries. Partnerships between fisheries scientists, economists, and insurance companies will be essential to design effective contracts. As fishing communities face increasing environmental and economic uncertainty, index insurance provides a promising tool to help fishers remain financially resilient.

A Appendix

Table A1: Summary of environmental variables used to predict fishery catch.

	Variable Name	Temporal Resolution	Spatial Resolution	Source
Temperature	Sea Surface Temperature	Monthly	5-km	NOAA Coral Bleaching Degree Heating Week
	Cummulative Habitat Compression Index	Annual	1-degree latitude	Integrated Ecosystem Assessment
Upwelling	Spring BEUTI	Daily	1-degree latitude	Jacox et al., 2018
	Spring CUTI	Daily	1-degree latitude	Jacox et al., 2018
	Frequency	Daily	1-degree latitude	Suca et al., 2022
	Relaxation	Daily	1-degree latitude	Suca et al., 2022
	Spring Transition Index	Daily	1-degree latidue	Suca set al., 2022
Climate	Pacific Decadal Oscillation	Monthly	Regional	PDO ERSST V5
	Oceanic Nino Index	Monthly	Regional	MEI.v2
	El Nino Southern Oscillation Index	Monthly	Regional	ONI
Biological	Squid Paralarvae	Annual	Statewide	RREAS
	Krill Abundance	Annual	Statewide	RREAS

Table A2: Model performance and utility results when coverage level of contract is set to 80% of average catch. Instances where no payouts are issued in the test set are indicated by "No payouts". Loss ratios of market premiums that would need to be subsized are indicated by "-".

Model	% Utility	Loss Ratio	Train RMSE (mt per fisher)	m^*	Loss Ratio at m^*
SST	29.1	1.17	228.7	1.1	1.079
Spring BEUTI	-4.2	0.00	226.3	No payouts	No payouts
STI	-0.2	1.44	243.4	0.3	-
Frequency	63.5	1.53	226.1	1.5	1.013
Relaxation	-0.4	0.00	238.0	No payouts	No payouts
Spring CUTI	0.0	0.00	236.1	No payouts	No payouts
CHCI	-6.5	0.48	224.6	0	-
ENSO	-11.1	0.82	221.8	0.5	-
PDO	-12.1	0.65	233.5	0.1	-
ONI	-1.3	0.94	224.0	0.9	-
Krill	0.0	NaN	216.0	NA	NA
Squid	-3.4	0.00	237.1	No payouts	No payouts
Random Forest	-3.1	0.00	115.8	No payouts	No payouts
Regularized RF	-2.3	0.00	133.6	No payouts	No payouts
Support Vector Machine	3.4	0.45	144.7	1.8	0.253
LASSO	19.9	2.84	130.2	2.9	0.963

References

Akselrud, C.I.A. (2024) Random forest regression models in ecology: Accounting for messy biological data and producing predictions with uncertainty. *Fisheries Research* **280**, 107161.

Barnett, B.J. and Mahul, O. (2007) Weather index insurance for agriculture and rural areas in lower-income countries. *American Journal of Agricultural Economics* **89**, 1241–1247.

Benso, M.R., Gesualdo, G.C., Silva, R.F., et al. (2023) Review article: Design and evaluation of weather index insurance for multi-hazard resilience and food insecurity. *Natural Hazards and Earth System Sciences* **23**, 1335–1354.

Binswanger-Mkhize, H.P. (2012) Is there too much hype about index-based agricultural insurance? *Journal of Development Studies* **48**, 187–200.

Butler, J. and Fuller, D. (1999) BUTLER ET AL.: AGE AND GROWTH OF MARKET SQUID OFF CALIFORNIA CalCOFl rep AGE AND GROWTH OF MARKET SQUID (LOLIGO OPALESCENS) OFF CALIFORNIA DURING 1998.

Cesarini, L., Figueiredo, R., Monteleone, B. and Martina, M.L.V. (2021) The potential of machine learning for weather index insurance. *Natural Hazards and Earth System Sciences* **21**, 2379–2405.

Chasco, B.E., Hunsicker, M.E., Jacobson, K.C., Welch, O.T., Morgan, C.A., Muhling, B.A. and Harding, J.A. (2022) Evidence of temperature-driven shifts in market squid doryteuthis opalescens densities and distribution in the california current ecosystem. *Marine and Coastal Fisheries* **14**, 1–13.

Table A3: Model performance and utility results when coverage level of contract is set to 90% of average catch. Instances where no payouts are issued in the test set are indicated by "No payouts". Loss ratios of market premiums that would need to be subsized are indicated by "-".

Model	% Utility	Loss Ratio	Train RMSE (mt per fisher)	m^*	Loss Ratio at m^*
SST	29.1	1.59	228.7	1.3	1.25
Spring BEUTI	-10.4	0.00	226.3	No payouts	No payouts
STI	-0.3	1.64	243.4	0.3	-
Frequency	58.7	2.68	226.1	2.6	1.046
Relaxation	6.5	3.07	238.0	2.8	1.112
Spring CUTI	-2.6	0.00	236.1	No payouts	No payouts
CHCI	-14.4	0.61	224.6	0.1	-
ENSO	-12.9	0.75	221.8	0.4	-
PDO	-12.1	0.52	233.5	0.1	-
ONI	-3.7	0.92	224.0	0.8	-
Krill	-4.6	0.00	216.0	No payouts	No payouts
Squid	-4.5	0.00	237.1	No payouts	No payouts
Random Forest	10.3	0.84	115.8	1.4	0.596
Regularized RF	-1.6	0.61	133.6	0.9	-
Support Vector Machine	21.4	1.27	144.7	1.8	0.716
LASSO	30.2	1.78	130.2	1.9	0.956

Table A4: Model performance and utility results with log utility. Instances where no payouts are issued in the test set are indicated by "No payouts". Loss ratios of market premiums that would need to be subsized are indicated by "-".

Model	% Utility	Loss Ratio	Train RMSE (mt per fisher)	m^*	Loss Ratio at m^*
SST	0.2	2.66	228.7	1.3	2.128
Spring BEUTI	-0.7	0.07	226.3	0.2	-
STI	0.0	1.64	243.4	1.1	1.478
Frequency	2.4	3.96	226.1	4	1.003
Relaxation	0.7	2.91	238.0	2.8	1.052
Spring CUTI	-0.4	0.00	236.1	No payouts	No payouts
CHCI	-0.6	0.64	224.6	0.4	-
ENSO	-0.5	0.71	221.8	0.5	-
PDO	-0.5	0.45	233.5	0.2	-
ONI	-0.1	0.89	224.0	0.9	-
Krill	-0.1	0.55	216.0	0.9	-
Squid	-0.3	0.00	237.1	No payouts	No payouts
Random Forest	0.4	1.04	115.8	1.2	0.895
Regularized RF	0.1	1.00	133.6	1.1	0.947
Support Vector Machine	1.4	1.38	144.7	1.5	0.895
LASSO	2.0	1.68	130.2	1.6	1.028

Chen, Z., Lu, Y., Zhang, J. and Zhu, W. (2024) Managing weather risk with a neural network-based index insurance. *Management Science* **70**, 4306–4327.

Clarke, D.J. (2016) A theory of rational demand for index insurance. *Journal: Microeconomics* **8**, 283–306.

Clement, K.Y., Botzen, W.J.W., Brouwer, R. and Aerts, J.C.J.H. (2018) A global review of the impact of basis risk on the functioning of and demand for index insurance. *International Journal of Disaster Risk Reduction* **28**, 845–853.

Conradt, S., Finger, R. and Spörri, M. (2015) Flexible weather index-based insurance design. *Climate Risk Management* **10**, 106–117.

Correia, H.E. (2021) Semiparametric model selection for identification of environmental covariates related to adult groundfish catches and weights. *Scientific Reports* **11**, 1–14.

Costello, C. and Polasky, S. (2008) Optimal harvesting of stochastic spatial resources. *Journal of Environmental Economics and Management* **56**, 1–18.

Dalhaus, T., Musshoff, O. and Finger, R. (2018) Phenology information contributes to reduce temporal basis risk in agricultural weather index insurance. *Scientific Reports* **8**, 1–10.

Feng, P., Wang, B., Liu, D.L., Waters, C. and Yu, Q. (2019) Incorporating machine learning with biophysical model can improve the evaluation of climate extremes impacts on wheat yield in south-eastern australia. *Agricultural and Forest Meteorology* **275**, 100–113.

Free, C.M., Anderson, S.C., Hellmers, E.A., et al. (2023) Impact of the 2014–2016 marine heatwave on US and canada west coast fisheries: Surprises and lessons from key case studies. *Fish and Fisheries* **24**, 652–674.

Free, C.M., Thorson, J.T., Plinsky, M.L., Oken, K.L., Wiedenmann, J. and Jensen, O.P. (2019) Impacts of historical warming on marine fisheries production. *Science* **365**, 979–983.

Goodrich, B., Yu, J. and Vandeveer, M. (2019) Participation patterns of the rainfall index insurance for pasture, rangeland and forage programme. *The Geneva Papers on Risk and Insurance - Issues and Practice* **45**, 29–51.

Herrmann, M., Greenberg, J., Hamel, C. and Geier, H. (2004) Extending federal crop insurance programs to commercial fisheries: The case of bristol bay, alaska, sockeye salmon. 352–366 pp.

Hilborn, R., Quinn, T.P., Schindler, D.E. and Rogers, D.E. (2003) Biocomplexity and fisheries sustainability. *PNAS* **100**, 6564–6568.

Hobday, A.J., Little, L.R., Watson, J.R. and Spillman, C.M. (2025) Parametric insurance for climate adaptation in fisheries and aquaculture. *Reviews in Fish Biology and Fisheries* **35**, 175–185.

Holland, D.S. and Leonard, J. (2020) Is a delay a disaster? Economic impacts of the delay of the california dungeness crab fishery due to a harmful algal bloom. *Harmful Algae* **98**.

Jacox, M.G., Edwards, C.A., Hazen, E.L. and Bograd, S.J. (2018) Coastal upwelling revisited: Ekman, bakun, and improved upwelling indices for the u.s. West coast. *Journal of Geophysical Research: Oceans* **123**, 7332–7350.

Jardine, S.L., Fisher, M.C., Moore, S.K. and Samhouri, J.F. (2020) Inequality in the economic impacts from climate shocks in fisheries: The case of harmful algal blooms. *Ecological Economics* **176**.

Jensen, N.D., Barrett, C.B. and Mude, A.G. (2016) Index insurance quality and basis risk:

Evidence from northern kenya. *American Journal of Agricultural Economics* **98**, 1450–1469.

Jensen, N., Stoeffler, Q., Fava, F., et al. (2019) Does the design matter? Comparing satellite-based indices for insuring pastoralists against drought. *Ecological Economics* **162**, 59–73.

Jewson, S., Brix, A. and Ziehmann, C. (2005) *Weather derivative valuation: The meteorological, statistical, financial, and mathematical foundations.* Cambridge University Press.

Kasperski, S. and Holland, D.S. (2013) Income diversification and risk for fishermen. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 2076–2081.

Keller, J.B. and Saitone, T.L. (2022) Basis risk in the pasture, rangeland, and forage insurance program: Evidence from california. *Amer. J. Agr. Econ* **104**, 1203–1223.

Kenduiywo, B.K., Carter, M.R., Ghosh, A. and Hijmans, R.J. (2021) Evaluating the quality of remote sensing products for agricultural index insurance. *PLoS ONE* **16**.

Klompenburg, T. van, Kassahun, A. and Catal, C. (2020) Crop yield prediction using machine learning: A systematic literature review. *Computers and Electronics in Agriculture* **177**.

Lehodey, P., Alheit, J., Barange, M., et al. (2006) Climate variability, fish, and fisheries.

Magel, C.L., Lee, E.M.J., Strawn, A.M., Swieca, K. and Jensen, A.D. (2020) Connecting crabs, currents, and coastal communities: Examining the impacts of changing ocean conditions on the distribution of u.s. West coast dungeness crab commercial catch. *Frontiers in Marine Science* **7**, 1–16.

Mahul, O. (1999) Optimum area yield crop insurance. *American Journal of Agricultural Economics* **81**, 75–82.

Miranda, M.J. and Farrin, K. (2012) Index insurance for developing countries. *Applied Economic Perspectives and Policy* **34**, 391–427.

Mumford, J.D., Leach, A.W., Levontin, P. and Kell, L.T. (2009) Insurance mechanisms to mediate economic risks in marine fisheries. *ICES Journal of Marine Science* **66**, 950–959.

Ovando, D., Cunningham, C., Kuriyama, P., Boatright, C. and Hilborn, R. (2022) Improving forecasts of sockeye salmon (oncorhynchus nerka) with parametric and nonparametric models. *Canadian Journal of Fisheries and Aquatic Sciences* **79**, 1198–1210.

Perretti, C.T. and Sedarat, M. (2016) The influence of the el niño southern oscillation on paralarval market squid (doryteuthis opalescens). *Fisheries Oceanography* **25**, 491–499.

Privitera-Johnson, K.M. and Punt, A.E. (2020) A review of approaches to quantifying uncertainty in fisheries stock assessments. *Fisheries Research* **226**, 105503.

Rahman, L.F., Marufuzzaman, M., Alam, L., Bari, M.A., Sumaila, U.R. and Sidek, L.M. (2022) Application of machine learning to investigate the impact of climatic variables on marine fish landings. *National Academy Science Letters* **45**, 245–248.

Ralston, S., Dorval, E., Ryley, L., Sakuma, K.M. and Field, J.C. (2018) Predicting market squid (doryteuthis opalescens) landings from pre-recruit abundance. *Fisheries Research* **199**, 12–18.

Sainsbury, N.C., Turner, R.A., Townhill, B.L., Mangi, S.C. and Pinnegar, J.K. (2019) The challenges of extending climate risk insurance to fisheries. *Nature Climate Change* **9**, 896–897.

Sapankevych, N. and Sankar, R. (2009) Time series prediction using support vector machines: A survey. *IEEE Computational Intelligence Magazine* **4**, 24–38.

Schmidt, L., Odening, M., Schlanstein, J. and Ritter, M. (2022) Exploring the weather-yield nexus with artificial neural networks. *Agricultural Systems* **196**.

Schroeder, I.D., Santora, J.A., Mantua, N., et al. (2022) Habitat compression indices for monitoring ocean conditions and ecosystem impacts within coastal upwelling systems. *Ecological Indicators* **144**, 109520.

Sethi, S.A. (2010) Risk management for fisheries. *Fish and Fisheries* **11**, 341–365.

Shanks, A., Roegner, G.C. and Miller, J. (2010) Using megalopae abundance to predict future commercial catches of dungeness crabs (cancer magister) in oregon. *California Cooperative Oceanic Fisheries Investigations Reports* **51**, 106–118.

Smith, K.E., Burrows, M.T., Hobday, A.J., et al. (2023) Biological impacts of marine heatwaves. *Annual Review of Marine Science* **15**, 1–27.

Souza, E.N.D., Boerder, K., Matwin, S. and Worm, B. (2016) Improving fishing pattern detection from satellite AIS using data mining and machine learning. *PLoS ONE* **11**.

Spicka, J. and Hnilica, J. (2013) A methodical approach to design and valuation of weather derivatives in agriculture. *Advances in Meteorology* **2013**.

Suca, J.J., Santora, J.A., Field, J.C., et al. (2022) Temperature and upwelling dynamics drive market squid (doryteuthis opalescens) distribution and abundance in the california current. *ICES Journal of Marine Science* **79**, 2489–2509.

Sumaila, U.R., Walsh, M., Hoareau, K., et al. (2021) Financing a sustainable ocean economy. *Nature Communications* **12**.

Szuwalski, C.S., Aydin, K., Fedewa, E.J., Garber-Yonts, B. and Litzow, M.A. (2023) The collapse of eastern bering sea snow crab. *Science* **382**, 306–310.

Villaseñor-Derbez, J.C., Arafeh-Dalmau, N. and Micheli, F. (2024) Past and future impacts of marine heatwaves on small-scale fisheries in baja california, mexico. *Communications Earth and Environment* **5**.

Watson, J.R., Spillman, C.M., Little, L.R., Hobday, A.J. and Levin, P.S. (2023) Enhancing the resilience of blue foods to climate shocks using insurance. *ICES Journal of Marine Science* **80**, 2457–2469.