Finding suitable weather indices for novel fisheries index insurance using machine learning

Working Paper not for Distribution

Nathaniel Grimes

2024-10-13

Index insurance is a financial tool gaining traction for application in fisheries. It will cover fishers losses under extreme weather events that impact fishery productivity. This is the first assessment to determine the feasibility of such programs and whether suitable indices exist. Catch and revenue data from 74 California fisheries are matched to 20 environmental variables using three prediction models: linear regression, LASSO regression, and random forests. The models are used to calculate the utility improvements of index insurance for fishers. Random forests provide the most significant improvements in utility, increasing fisher welfare by 14%. The results suggest that index insurance can improve fisher welfare, but the extent of success is not consistent across all fisheries. Specifically tailored insurance contracts must be created for each fishery using the best available data and models

Table of contents

1 Introduction **2**

2 Insurance Model **4**

3 Data **5**
3.1 Fishery Data 5
3.2 Environmental Data 6

4 Methods **7**
4.1 Linear Models 7
4.2 LASSO Regression 8
4.3 Random Forests 8

4.4 Weather variables of importance 9

5 Results **9**

6 Discussion **10**

7 Appendix **11**

References **13**

1 Introduction

Predicting fishery output from weather variables is notoriously difficult. It is widely established that climate and weather affect fishing populations (Lehodey *et al.* 2006), but most stock assessment models use little to no year to year environmental data (Privitera-Johnson and Punt 2020). Variations in environmental conditions are now the leading cause of fishery closures and disaster relief payouts in the United States (Bellquist *et al.* 2021). Disaster declarations are becoming more frequent straining a slow, inequitable system (Holland and Leonard 2020; Jardine *et al.* 2020). Calls for new financial tools to alleviate fisher income shocks have grown (Mumford *et al.* 2009; Sethi 2010).

Index insurance has risen as a prime candidate to protect fishing communities during disasters (Watson *et al.* 2023). Index insurance is a financial product that pays out when an independently verified index, such as rainfall or temperature, falls below a predetermined threshold. The index is chosen to be highly correlated with the asset being insured. However, it is difficult to establish clear, concise weather impacts on fishery productivity. The biological dynamics of the system can lead to lower stock health persisting for years after a negative shock (**Hilborn2003?**). Individual fisheries can cover enormous areas in ocean basins. The expansive spatial coverage of fisheries makes it unclear where and how much specific weather variables impact biological abundance. In addition, if weather impacts have been observed, they are most likely highly non-linear adding further complexity. The greatest impediment to the development of fishery insurance policies is reconciling these challenges to find suitable indices that can predict fishery productivity (Watson *et al.* 2023).

Recent expansions in oceanic remote sensing has led to a plethora of new environmental indices that could be used to predict fishery productivity. Fishery data collection continues to improve with better reporting systems with longer and more detailed catch histories. This study aims to leverage these improved data sources to create suitable indices for fisheries index insurance using machine learning.

The difficulty in modelling fishery productivity with environmental indices leads to basis risk. Formally, basis risk is the probability that policyholders experience a harmful shock to their income, but the index does not trigger. Basis risk lowers demand for index insurance and remains a significant roadblock in setting up new programs (Binswanger-Mkhize 2012; Clarke 2016; Clement *et al.* 2018).

It is impossible to completely eliminate basis risk, and there exists a wide range in exisiting agriculture products. Well designed policies can capture up to 90% of the income variation as shown in Kenyan pasture grazing indices (Jensen *et al.* 2019). Abysmal correlations are prevalent in the Rainfall Index Insurance for Pasture, Rangeland, and Forage (RI-PRF) program where correlations as low as 0.071 exist in California, which leads to 46% additional points of basis risk (Keller and Saitone 2022). The program has a 26% probability of not paying out when damages are suffered in Nebraska and Kansas (Yu *et al.* 2019). Subsidies covering up to 60% of ranchers paid premiums are needed to stimulate demand in the RI-PRF program (Goodrich *et al.* 2019).

Contract design can mitigate basis risk through providing more options so that individuals can better select policies that protect them. Policyholders choose lower trigger levels when correlations between between index and asset are low (Lichtenberg and Iglesias 2022). Lower trigger levels correspond to protection against more catastrophic shocks. Increased contract flexibility reduces basis risk by only small amounts. Yu *et al.* (2019) found that more flexible contracts could account for only 5-9% of basis risk. Farms in Kansas closer to weather stations had better predictive impacts of rain on yield (Yu *et al.* 2019).

Designing indices with stronger correlations to fishery losses is the most effective way to reduce basis risk (Jensen *et al.* 2019). Agricultural researchers continually seek new methods and data sources to improve the correlation between loss and weather variables. Quantile regressions improve Kazak wheat farmers utility between 0.1-22% over linear models depending on the underlying measure of utility (Conradt *et al.* 2015). Remote sensing variables leveraging the latest satellite data on vegetative cover and rainfall provide better coverage than county wide averages (Dalhaus and Finger 2016; Valverde-Arias *et al.* 2020).

Machine learning has exploded as a new method to define better indices in agriculture index insurance (Feng *et al.* 2019; Cesarini *et al.* 2021; Schmidt *et al.* 2022; Chen *et al.* 2024). Machine learning models excel in index insurance because indemnity contracts only need predictive relationships. Whereas, fishery stock assessments build complex models with biological foundations to accurately inform management of future fish stocks, index insurance can look retroactively at data to uncover relationships and test out of sample predictive quality. Machine learning may be necessary in fisheries index insurance to uncover any valuable relationships between weather and productivity.

The application of machine learning is growing in fisheries as researchers explore data questions beyond formal stock assessments. Ensemble models built through combinations of random forests, boosted trees, and dynamic linear models improved 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). The expected non-linear interactions of weather and fishery productivity merit the use of machine learning in fisheries.

This study will provide the first comprehensive examination of the necessary features of weather indices for fisheries index insurance. We contribute to the growing ocean adaption and blue finance literature. While our intent is to improve fisher welfare with a new financial program, we also expand the assessment of which weather variables have predictive performance in highly productive upwelling ecosystems.

The rest of the paper is structured as follows. Section 2 describes the insurance model tested in this study. Section 3 describes the data collection, transformations, and sources. Fisheries data comes from newly open-access sources provided by the California Department of Fish and Wildlife. Section 4 describes the algorithms used to predict fishery productivity and evaluate the utility of index insurance. Section 5 presents the welfare results as well as extracting which variables contribute to the prediction models to infer more interpretable results for fishers. Section 6 discusses the results and implications for the future of fisheries index insurance.

2 Insurance Model

Insurance contracts are specified by calculating payout functions ($I(\omega)$) based on independently measured weather variables. Neural networks have been used to provide non-linear payoff schedules that better reduce basis risk (Chen *et al.* 2024). While it has been shown that linear payoff functions inherently lead to basis risk and therefore lower demand (Clarke 2016; Jensen *et al.* 2016), we maintain their use to preserve measures of interpretability that are clearer for a first analysis of fishery index insurance.

Payouts will be issued when prediction models predict negative deviations from the long run average. The three prediction models ($k \in \{\text{LR,LA,RF}\}$ are a linear regression (LR), a LASSO regression (LA), and a random forest (RF).

$$I(\omega) = \max(0, (\bar{\pi} - \hat{\pi}_t^k(\omega)) \cdot l) \tag{1}$$

Where k is the prediction model, l is the level of coverage, $\hat{\pi}_t^k(w)$ is the predicted fishing variable from ω weather variables, and $\bar{\pi}$ the long run average of the fishing variable. The premium ρ is calculated as the expected value of the payout function times the premium loading factor m (Equation 2).

$$\rho(\omega) = \mathbb{E}[I(\omega)]m \tag{2}$$

Utility measures offer the most insightful evaluation of index insurance policies (Kenduiywo *et al.* 2021). It captures value added for policyholders, not just measures of payout frequency as other measures of basis risk. Constant absolute risk aversion allows more consistent comparison for different levels of wealth. Fishing variables may vary extensively from fishery to fishery. We normalize utility by dividing all measured payouts and fishing variables by the maximum observed value. Expected utility for a given fishery is the average utility over all years in the sample for any variable of interest π_t. Fishers are allowed to choose insurance coverage levels l to ensure feasible contracts.

$$
\begin{aligned}
\mathbb{E}[U_b] &= \frac{1}{n}\sum_{t}^{T}\frac{1-e^{-a\pi_t}}{a} && \text{No Insurance} \\
\mathbb{E}[U_i] &= \max_{l}\frac{1}{n}\sum_{t}^{T}\frac{1-e^{(-a(\pi_t+I(\omega,l)-\rho(w))}}{a} && \text{Insurance} \\
U_r &= \frac{\mathbb{E}[U_i]-\mathbb{E}[U_b]}{\mathbb{E}[U_b]}\cdot 100 && \text{Percent Change in Utility}
\end{aligned}
\tag{3}
$$

We will compare the percent change in fisher utility with insurance (U_i) versus without insurance (U_b) for each prediction method. The variable of interest π_t, will be fishing revenue, landings, and revenue per fisher to test what measures of fishery productivity are most suitable for index insurance.

No data on fishery insurance suppliers exists to create a market equilibrium. We iteratively vary the premium loading $m \in [1,2]$ to create a range of coverage values fishers will be willing to pay for a given m. Then, the amount of coverage purchased times the premium loading factor approximately equals the expected revenue an insurance company would receive. Insurance companies could then examine their own administrative and legal costs to determine whether the feasible contract is profitable.

3 Data

This study attempts to cover breadth, not depth in possible indices. Each fishery has unique ecological characteristics that interact with environmental variables in different and non-linear ways. By studying a wide collection of fisheries and environmental variables we can uncover the potential feasibility of index insurance for fisheries holistically, and then further refine measures with ecologically sound models in future applications.

3.1 Fishery Data

Landings revenue, and participation data comes from the West Coast Fish data package (Free *et al.* 2022). It is a reconstruction of California Fish and Wildlife Department catch data combined with PacFin receipts for Washington and Oregon. The last three years of data are updated from the CDFW Marine Fisheries Data Explorer (MFDE). Names are matched to each species within the West Coast Fish data package.

We select California fisheries with a minimum of 30 years of consecutive, non-confidential catch records at both the state and port-complex level. Unclassified catch records are dropped i.e. "Other Sharks" and similar categories. Fisheries with an average revenue from 2010-2019 greater than \$100,000 at the state and \$75,000 at port-complex level are analyzed. Twenty four fisheries at the state level and 50 fisheries at the port complex level meet these criteria. These fisheries contain the most economically important fisheries in California and their mean values are shown in Table 3 and at the port complex in Table 4.

Fisheries have complex spatial dynamics. Agriculture has clear, quantifiable impacts of weather in grids that are well suited for index insurance. Drought on a single farm directly leads to crop loss for that farm. Whether there is sufficient spatial coverage to identify impacts down to an individual farm remains a challenge in agriculture (Dalhaus and Finger 2016; Leppert *et al.* 2021; Stigler and Lobell 2024). Fish and fishers can move thousand of miles in a given year, thus more consideration must be given to the location of weather impacts in fisheries. We spatially refine catch histories using the California CDFW fishing blocks records from the MFDE Data Explorer. Summarized catch histories of all landed fish within each block provide an average representation of effort for a given fishery. Spatial catch history is measured at both the state and port-complex level. The spatial location refines the location of environmental variables. Local weather is more likely to affect fishery productivity and catch than observations thousands of miles away.

3.2 Environmental Data

Fisheries are highly sensitive to marine heatwaves and water temperature. Sea surface temperature is a natural variable to first consider in fisheries index insurance. Sea surface temperature data comes from the NOAA DHW data set that provides 5-km resolution of monthly temperature from 1985 to 2023. The 5-km grids are averaged within the nearest California fishing block to provide an annual time series of temperature for each fishery. Temperature is lagged from 1 to 3 years prior to account for residual impacts that carry over due to fishery biological dynamics.

Upwelling provides vital nutrients to stimulate primary productivity. The coast of California is a highly productive ecosystem due to its patterns of upwelling (Chelton *et al.* 1982; Huyer 1983). We capture upwelling through monthly observations of Coastal Upwelling Transport Index (CUTI) and Biological Effective Upwelling Transport Idnex (BEUTI). 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. Spring upwelling in early March and April are especially important to a wide array of fish species. Yearly average and amplitude values (the difference between minimum observed upwelling and maximum) are also calculated. These indices are the most temporally limited datasets in this analysis, only extending from 1988 to 2023.

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 oceangraphic regions ranging from 3.5 degrees to 5.5 degrees lattitude in size with coverage out to 150 km offshore. We 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)

The final environmental variables are the Pacific Decadal Oscillation (PDO) and the El Nino Southern Oscillation (ENSO). Both indices are well known to affect marine ecosystems and fisheries. Both indices are averaged over a given year. PDO data is taken from the PDO ERSST V5, and ENSO data is taken from the multivariate ENSO Index Version 2 (MEI.v2).

Summary statistics for the environmental data are presented in Table 5. In total, 74 fisheries with 35 years of catch data and 20 weather variables are spatially matched with annual coverage from 1988 to 2023.

4 Methods

We use three models to predict yearly fishing revenue, landings, and catch per fisher at state and port-complex levels. Linear models are used as the base model given its ubiquitous use in index insurance policies. We compare utility improvements with the adoption of more robust LASSO regression and random forest models.

In all class of models, the final utility maximization choice of coverage leverage is found through a box constrained quasi-Newton Method using the optim function in R. Choices are constrained to be non-negative. Premium schedules are found by the model output below the trigger values in Equation 1 and then averaged over the total fishery data.

4.1 Linear Models

Perfect regression coefficients mimic the optimal choice of scale in index insurance contracts (**Mahul1999?**). Combined with the ease of implementation, linear models on single weather indices are the most common design choice for index insurance policies. They offer a basic starting place to consider the viability of fishery index insurance.

Yearly aggregated fishing variables are regressed on each environmental variable individually. Weather variables are spatially matched to the location of catch. We perform a 10 fold cross validation method to determine the best individual weather variable based on root mean square error (RMSE). To preserve the time series element of the data, we used a rolling split to partition the training and testing data. For example, the first fold contains the first 70% of data as training (1988-2011), and the last 30% as testing (2013-2023). The final date of the training set is extended in each fold until the year 2020 to create 10 folds. Models with the lowest average RMSE are selected and trained on the full set before being passed to the utility optimization procedure in Equation 3.

4.2 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 Ordinary Least Squares (OLS), but adds a penalty constraint on the absolute sum of selected coefficient values (Equation 4).

$$\hat{\beta}^{lasso} = \arg\min_{\beta} \left\{ \sum_{i=1}^{n} (y_i - \beta_0 - \sum_{j=1}^{p} \omega_{ij}\beta_j)^2 + \lambda \sum_{j=1}^{p} |\beta_j| \right\} \tag{4}$$

Where, y_i is our fishing variable, β the regression coefficients, n, the number of observations, p the number of predictors, and ω 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 LASSO regression model is trained on 200 bootstrapped samples of the training data. The optimal λ is selected through a grid search method that selects the minimum RMSE. This choice is to ensure the most parsimonious model that still captures the most important weather variables. LASSO is particularly well suited for this research design as the absolute value of the penalty term shrinks coefficients to zero. Overfitting is a concern with so few observations in the initial training set; the shrinkage towards zero will help minimize this bias.

4.3 Random Forests

While LASSO offers us the ability to simultaneously explore a wide collection of weather variables including lagged effects, it remains linear in its predictions. 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.

We tune two hyperparameters to create the best performing random forest for each fishery: The number of variables to consider at each split and the minimum number of observations in a leaf node. We use a grid search method to find the best hyperparameters based on RMSE through the year based cross validation method presented in the linear models. The final model is trained on the full dataset and passed to the utility optimization procedure in Equation 3.

4.4 Weather variables of importance

Machine learning algorithms are inherently "black boxes" that sacrifice interpreability 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 bioloigcal modelling.

The cross-validation in the linear models provides a simple weather variable comparison. We calculate the frequency a given weather variable is chosen as the best performing linear model.

We use `vip` package in R to extract importance measures for both the LASSO and random forests [1]. Feature extraction will occur for each fishery product, but the importance of each will be normalized then aggregated in order to compare all features.

5 Results

Index insurance can improve fisher utility even with univariate linear models, but the extent of success is not consistent across all fisheries. At actuarially fair premiums, $m = 1$, the average utility improvement for fishers with linear models is approximately 2% at the state level of all fishing dependent variables (Table 1). LASSO increases the utility gains of index insurance through better performing indices. The improvements slightly depend on the fishery dependent variable. Targeting fishing revenues provides more than 1% additional gains in utility over landings. Revenues may be more difficult to predict as they incorporate external market factors such as price and demand, but revenues more closely match the financial interests of fishers.

Random forests provide the most significant improvements in utility. The ability to capture the non-linear interactions of weather variables with fishery productivity provides much more accurate trigger states that more closely align with lost value. All three fishery dependent trigger variables with random forest models improved fisher utility by 14%.

Results are consistent when looking at the port-complex level, though LASSO and random forest have 2% less utility improvements than at the state level (Table 2). Large discrepncies can exist between port-complexes for the same insurance policies. For example, an insurance contract for Chinook Salmon based on metric tons at the California state level improved utility

[1]Nathan note: I need to read more exactly how this package will extract between permutations or variance measures

Table 1: Average relative percent improvement in utility for 24 fisheries at the state level. Standard deviations in utilty are included in parathesis. Fishers

	Landings	Revenue	Revenue per Fisher
Linear	2.4% (2.7)	2.1% (2.8)	2.4% (2.7)
LASSO	6.3% (6.4)	7.3% (8.3)	8% (6.2)
Random Forest	13.9% (10.8)	14.7% (10)	14% (8.6)

Table 2: Average relative percent improvement in utility for 50 fisheries at the port complex level.

	Landings	Revenue	Revenue per Fisher
Linear	1.9% (2.6)	1.9% (3.4)	2.7% (3.9)
LASSO	4.9% (4.8)	6% (5)	6.7% (5)
Random Forest	11.8% (6.9)	12.8% (6.2)	12.5% (6.9)

by 21%. However, different port complexes choose different levels of coverage (**?@tbl-chnk**). Eureka was the most extreme difference by opting for no insurance. The weather variability was not captured sufficiently by the LASSO model in Eureka and could not provide enough smoothing to incentivize fishers to purchase insurance. This observation emphasizes the benefit of creating policies at the port-complex level as the local variations capture basis risk exposure to each unique community [2].

6 Discussion

Remaining things do to

- Variable importance
- Feasibility of interprebable models
- Alter m off actuarilly fair
- Change utility models
- Make a payout frequency table
- Show the income smoothing effect as a distribution

[2]Nathan note: Should compare the eureka port-complex data with the cali lasso model to see if they would want insurance with that. Kind of a robustness check

Table 3: Summary statistics of catch from 1988-2023 for California fisheries.

	Landings (mt)		Revenue (USD)		MT per Fisher		Number of Fishers	
Species	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Albacore	2055.4	2646.5	$3,574,984	4051538.3	5.3	3.1	305.5	307.7
Bluefin tuna	760.5	1131.1	$1,118,896	1137192.5	9.3	12.1	76.3	46.1
California spiny lobster	314.6	69.6	$7,752,729	5402314.5	1.7	0.6	202.0	39.8
Chinook salmon	1576.1	1291.7	$10,383,616	7643140.9	1.5	1.1	1146.6	1022.2
Chub mackerel	11304.3	10208.0	$1,944,342	1709450.2	80.2	59.5	133.9	78.6
Dungeness crab	5712.8	3334.7	$29,115,960	23070113.4	11.6	7.9	534.5	121.7
Market squid	51363.6	34634.6	$27,517,541	24057582.4	462.1	289.5	109.4	23.4
Nom. Cabezon	35.3	38.7	$306,669	321874.4	0.2	0.1	199.0	115.1
Nom. Calif halibut	374.6	140.9	$2,605,362	873979.0	0.9	0.3	427.1	87.8
Nom. California sheephead	46.7	38.8	$342,986	231782.3	0.4	0.2	134.1	88.5
Nom. Lingcod	359.0	421.0	$388,778	276453.1	0.4	0.3	675.4	374.0
Nom. Shortspine thornyhead	56.4	70.0	$432,850	465662.0	0.5	0.6	116.2	35.6
Northern anchovy	7696.1	10157.9	$860,871	643585.4	230.8	246.9	32.7	9.8
Opah	102.4	100.0	$198,182	257248.5	2.5	3.6	71.0	46.8
Pacific bonito	1090.7	1694.6	$544,835	886379.0	8.7	10.3	108.2	168.6
Pacific sardine	20665.4	22757.5	$2,574,205	2614838.8	309.1	325.9	53.3	21.0
Red sea urchin	7947.2	6067.5	$11,443,321	7684762.0	32.7	12.7	231.1	139.5
Ridgeback prawn	172.8	146.5	$634,305	473336.8	7.7	5.7	26.1	12.6
Rock crab	620.2	146.3	$1,750,541	597289.3	3.6	1.2	180.7	47.4
Sablefish	2753.2	1823.2	$5,759,976	2768043.7	10.4	6.9	269.6	70.3
Spotted prawn	153.1	76.1	$3,206,726	1989792.4	3.7	2.3	52.6	31.1
Swordfish	808.3	554.7	$5,671,198	3401866.3	6.7	3.2	129.3	84.3
White seabass	109.2	76.4	$635,876	429179.6	0.8	0.7	150.3	45.1
Yellowfin tuna	8853.7	17403.5	$10,230,193	20959459.4	67.4	89.0	74.0	73.9

7 Appendix

Table 4: Summary statistics of catch from 1988-2023 for California fisheries split betwe species and port complex

		Landings (mt)		Revenue (USD)		MT per Fisher	
Species	Port	Mean	SD	Mean	SD	Mean	SD
Albacore	Eureka	355.0	377.4	$672,866	634894.5	5.4	3.4
	San Francisco	175.8	368.5	$322,734	639778.0	2.6	2.2
Bluefin tuna	San Diego	25.2	38.5	$163,167	334643.0	0.8	1.7
	Los Angeles	47.2	36.2	$303,840	172626.8	0.7	0.3
	San Diego	17.9	12.3	$117,380	51099.8	0.6	0.3
	San Francisco	140.4	64.0	$1,053,595	721430.4	1.3	0.6

Nom. Calif halibut

	Santa Barbara	97.9	44.4	$705,370	187398.0	1.1	0.3
Chinook salmon	Bodega Bay	329.6	309.3	$2,258,504	1989960.2	1.0	0.6
	Eureka	103.6	178.4	$649,797	1000836.3	0.4	0.4
	Fort Bragg	343.9	404.8	$2,320,248	2367832.3	1.1	1.4
	Monterey	313.7	272.7	$1,896,626	1161527.4	0.9	0.6
	Morro Bay	72.9	81.4	$493,105	515982.2	0.6	0.5
	San Francisco	493.7	348.0	$3,339,479	2075261.9	1.2	0.8
Chub mackerel	Los Angeles	10320.4	9344.1	$1,782,814	1561212.8	150.3	110.6
Dungeness crab	Bodega Bay	572.9	554.1	$3,390,721	3510592.1	6.7	6.2
	Eureka	3571.6	2232.7	$16,072,727	12672636.9	13.9	10.7
	Fort Bragg	247.3	191.4	$1,376,466	1437305.5	5.6	3.9
	Monterey	91.7	104.4	$692,846	868788.1	3.0	2.8
	San Francisco	1174.3	1276.2	$7,090,167	8035652.0	7.2	6.3
California spiny lobster	Los Angeles	98.3	21.3	$2,326,853	1471851.3	1.3	0.5
	San Diego	99.0	22.5	$2,179,454	1244460.4	1.5	0.6
	Santa Barbara	116.8	49.4	$3,239,108	2893952.5	2.0	0.9
Market squid	Los Angeles	16986.5	14774.8	$8,750,561	8927983.2	278.8	175.1
	Santa Barbara	24534.0	20112.5	$12,790,364	13303266.4	412.4	269.4
Opah	San Diego	58.6	66.2	$121,629	182250.9	2.5	3.6
Rock crab	Los Angeles	101.5	105.4	$259,159	196512.5	2.3	1.7
	Morro Bay	75.9	65.7	$183,343	124726.4	3.5	2.4
	San Diego	67.1	35.0	$163,988	79313.9	2.3	0.9
	Santa Barbara	344.2	173.1	$1,036,433	641375.7	5.3	3.0
Ridgeback prawn		155.5	139.5	$562,758	453595.2	7.9	5.6
Red sea urchin	Los Angeles	1229.8	1020.2	$1,890,768	1362341.7	18.0	4.8
	Santa Barbara	3971.6	2451.7	$6,081,764	3987202.8	32.0	14.5
	Bodega Bay	90.7	90.1	$186,193	166543.1	4.1	2.7
	Eureka	886.0	607.5	$1,634,560	691917.3	12.9	5.4
	Fort Bragg	574.5	368.0	$1,221,487	633518.2	13.5	13.6
	Los Angeles	227.5	552.3	$321,915	461035.2	13.6	50.7
	Monterey	316.0	204.2	$642,984	415849.4	7.2	6.1

Sablefish

	Morro Bay	229.8	185.6	$679,713	939824.4	8.1	5.4
	San Diego	25.9	22.7	$137,336	144319.5	2.1	2.3
	San Francisco	336.2	320.6	$559,985	268977.2	6.8	4.6
	Santa Barbara	63.3	74.2	$352,529	476423.6	3.0	2.9
Nom. California sheephead	San Diego	13.9	11.3	$119,436	84159.3	0.7	0.5
	Santa Barbara	20.9	22.7	$134,335	125858.2	0.4	0.4
Spotted prawn		53.4	29.2	$1,116,708	873801.1	3.5	2.4
Swordfish	Los Angeles	315.9	360.1	$2,072,740	1972406.2	5.8	5.2
	San Diego	178.4	131.0	$1,416,146	905587.3	3.0	1.3
	Santa Barbara	66.1	91.5	$466,637	585529.3	2.3	1.9
White seabass	Los Angeles	31.5	30.9	$146,996	111988.1	1.2	1.0
	San Diego	13.1	26.4	$73,176	87486.8	0.6	0.7
	Santa Barbara	54.3	43.1	$331,544	269526.5	1.1	1.0

References

Bellquist, L., Saccomanno, V., Semmens, B.X., Gleason, M. and Wilson, J. (2021) The rise in climate change-induced federal fishery disasters in the united states. *PeerJ* **9**.

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

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.

Chelton, D.B., Bernal, P.A. and McGowan, J.A. (1982) Large-scale interannual physical and biological interaction in the california current. *Journal of Marine Research* **40**, 1095–1125.

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.

Table 5: Summary statistics of environmental variables from 1988-2023 for California fisheries.

Weather Index	Mean	SD	Temporal Resolution	Spatial Resolution	Source
CUTI Amp	1.3	0.6	Monthly	1 degree latitude	Jacox et al., 2018
CUTI Avg	0.5	0.3	Monthly	1 degree latitude	Jacox et al., 2018
CUTI Fall	0.2	0.3	Monthly	1 degree latitude	Jacox et al., 2018
CUTI Summer	0.6	0.3	Monthly	1 degree latitude	Jacox et al., 2018
CUTI Spring	0.7	0.4	Monthly	1 degree latitude	Jacox et al., 2018
CUTI Winter	0.2	0.3	Monthly	1 degree latitude	Jacox et al., 2018
BEUTI Amp	15.8	10.6	Monthly	1 degree latitude	Jacox et al., 2018
BEUTI Avg	4.1	3.9	Monthly	1 degree latitude	Jacox et al., 2018
BEUTI Fall	1.0	2.0	Monthly	1 degree latitude	Jacox et al., 2018
BEUTI Summer	4.2	4.7	Monthly	1 degree latitude	Jacox et al., 2018
BEUTI Spring	9.1	7.9	Monthly	1 degree latitude	Jacox et al., 2018
BEUTI Winter	1.9	4.2	Monthly	1 degree latitude	Jacox et al., 2018
Cummulative Habitat Compression Index	4.8	2.3	Yearly	1 degree latitude	Integrated Ecosytem Assessment
Average Sea Surface Temperature	14.2	2.0	Monthly	5x5 km	NOAA Coral Bleaching Degree Heating Week
Sea Surface Temperature Lag 1 Year	14.2	2.0	Monthly	5x5 km	NOAA Coral Bleaching Degree Heating Week
Sea Surface Temperature Lag 2 Years	14.2	2.0	Monthly	5x5 km	NOAA Coral Bleaching Degree Heating Week
Sea Surface Temperature Lag 3 Years	14.2	2.0	Monthly	5x5 km	NOAA Coral Bleaching Degree Heating Week
Sea Surface Temperature Lag 4 Years	14.2	2.0	Monthly	5x5 km	NOAA Coral Bleaching Degree Heating Week
ENSO	-0.1	0.7	Monthly	Regional	MEI.v2
Pacific Decadal Oscillation	-0.3	1.0	Monthly	Regional	PDO ERSST V5

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

Dalhaus, T. and Finger, R. (2016) Can gridded precipitation data and phenological observations reduce basis risk of weather index-based insurance? *Weather, Climate, and Society* **8**, 409–419.

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., Poulsen, C.V., Bellquist, L.F., Wassermann, S.N. and Oken, K.L. (2022) The CALFISH database: A century of california's non-confidential fisheries landings and participation data. *Ecological Informatics* **69**, 101599.

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.

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**.

Huyer, A. (1983) Coastal upwelling in the california current system. *Progress in Oceanography* **12**, 259–284.

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.

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**.

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

Leppert, D., Dalhaus, T. and Lagerkvist, C.J. (2021) Accounting for geographic basis risk in heat index insurance: How spatial interpolation can reduce the cost of risk. *Weather, Climate, and Society* **13**, 273–286.

Lichtenberg, E. and Iglesias, E. (2022) Index insurance and basis risk: A reconsideration. *Journal of Development Economics* **158**.

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.

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.

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.

Stigler, M. and Lobell, D. (2024) Optimal index insurance and basis risk decomposition: An application to kenya. *American Journal of Agricultural Economics* **106**, 306–329.

Valverde-Arias, O.R., Esteve, P., Tarquis, A.M. and Garrido, A. (2020) Remote sensing in an index-based insurance design for hedging economic impacts on rice cultivation. *Hazards Earth Syst. Sci* **20**, 345–362.

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.

Yu, J., Vandeveer, M., Volesky, J.D. and Harmoney, K. (2019) Estimating the basis risk of rainfall index insurance for pasture, rangeland, and forage. *Journal of Agricultural and Resource Economics* **44**, 179–193.