**Introduction**

The walleye pollock fishery in Alaska is the largest in the United States by volume and generated a wholesale gross value of $1.4 billion in 2008 (2). Chinook salmon are classified as a prohibited species catch in this fishery, meaning their incidental capture is strictly regulated (1). In response to record high bycatch in 2007, Amendment 91 established a hard cap of 60,000 Chinook salmon for the entire fishery, divided by sector and season (A and B), with the consequence that exceeding this limit results in a complete fishery closure. Additionally, a performance limit of 47,591 salmon is allocated across seasons and sectors, with a rule that if any sector exceeds its share in three out of seven consecutive years, it is permanently restricted to that limit (2). This approach balanced the need to incentivize bycatch reduction while accounting for natural variability in salmon encounters. However, while the bycatch in the fisheries was reduced, in response to low Chinook abundance Amendment 110 was introduced to add an adaptive mechanism where the performance limit is further reduced during periods of low salmon abundance based on an established index (1). To enforce these rules, 100% observer coverage is mandated on all vessels within the pollock fishery, ensuring compliance and accurate monitoring of bycatch levels (1).

Efforts to reduce Chinook salmon bycatch in the pollock fishery employ a variety of strategies, including fixed closure areas, short-term closures in high-bycatch zones, and salmon bycatch excluders in trawl nets (1). While these measures have contributed to bycatch reduction, they can also have unintended commercial and ecological consequences. For example, time-area closures designed to protect one species may concentrate fishing effort elsewhere, potentially impacting other species (4), and they often restrict fishing in areas where bycatch risk is low, as they are typically based on historical rather than real-time data (3). In contrast, dynamic ocean management leverages eco-informatics and near real-time data streams to support adaptive fishing practices, allowing for a more responsive and precise approach to bycatch mitigation (3). This strategy aligns with industry experience, which has shown that cooperative data-sharing is a highly effective method for reducing salmon bycatch (1). To this end, providing the industry with models that incorporate environmental covariates to predict species distribution across longitude, latitude, and depth would offer a valuable resource for further refining bycatch avoidance strategies.

Depth is of particular interest as Chinook salmon and walleye pollock occupy overlapping ranges. Pollock are found from the seafloor to midwater and near-surface depths, with most catches occurring between 50 and 300 meters using pelagic trawls (5). This aligns with the 0–500 meter range observed for Chinook salmon in tagging studies (6), highlighting a key factor driving bycatch. However, models of salmon occupancy patterns that can take into account real-time environmental conditions could help fishers refine their operations, allowing them to target specific depths where bycatch risk is lower.

While several studies have examined depth occupancy in Chinook salmon, they have primarily focused on understanding the factors influencing depth use rather than developing inferential tools for prediction (7). Machine learning has also been applied, but mainly to analyze how environmental covariates influence depth occupancy, rather than generating practical predictive models (7). Given that fish behavior in response to environmental factors is inherently stochastic, an effective model would not aim to pinpoint exact depths, but rather estimate the likelihood of salmon occupying different depth ranges within the water column. Framed in this way, the problem becomes an ideal application for a probabilistic deep learning classifier, capable of modeling uncertainty and providing flexible, data-driven depth distribution predictions.

The goal of this study is to develop a probabilistic deep learning classifier capable of predicting Chinook salmon depth occupancy in near-real time. To achieve this, we will first identify key environmental covariates from existing literature that can be measured or predicted and then use pop-up satellite archival tagging data from Chinook salmon in the Gulf of Alaska (7) to build and evaluate the model. This approach aims to provide a practical, data-driven tool for improving bycatch mitigation strategies in the pollock fishery.

**Methods**

**Data**

The data used is a series of tracks from 111 Chinook salmon (*Oncorhynchus tshawytscha*) caught and monitored between 2013 and 2022 (6)(8). These tracks were obtained from pop-up satellite archival tags which collect temperature, light level, and depth information at specified (sub day) intervals. Depth information was extracted from these tracks.

While depth is measured on the tag at intervals more frequent that 15 mins, upon upload to the satellite the data is aggregated to a per 15 minute granularity in order to reduce the amount of information going over the satellite. We had a few tags that had been recovered and had the full data streams and upon comparison with the aggregated data determined that the measurements uploaded give a sense of the central tendency of the fish within a range of values. Therefore in order to not over-represent the precision of the uploaded data we sampled drew depth measurements from a normal distribution centered at the uploaded depth and with a standard deviation of 10% of that uploaded depth. From there we assigned each measurement to a depth bin in increments of 25, 50, 75, 100, 150, 200, 250, 300, 400, and 500 meters where each bin indicates its lower bound and does not include depths from the bins at shallower depths. In summary this gave us for each 15 minutes of tag deployment a sampled depth bin that the fish was likely in during that 15 minute interval. Before introduction to the model these depth bins were each given a float “id” between 0 and 1 with bin 25 corresponding to 0.1, 50 to 0.2, and so on.

Environmental data was derived from the Global Ocean Biogeochemistry Hindcast dataset (10.48670/moi-00019) and the Global Ocean Physics Reanalysis (10.48670/moi-00021) from the E.U. Copernicus Marine Service Information. Statistics were aggregated per Uber h3 resolution 4 cell and depth bin (as with the depth measurements) in the Northern Pacific. Features pulled were chlorophyll, net primary production, nitrate, oxygen, phosphate, silicate, elevation, mixed layer thickness, salinity, temperature and north and easterly current.

Before being included in the models all environmental features were rescaled to run between 0 and 1 with chlorophyll, net primary production, and mixed layer thickness being log scaled before doing so. This was intended to normalize the data for introduction into the neural network.

Temporal features were derived using the “suntimes” and “ephem” packages in Python with the former being used to compute day/night features and the latter to compute lunar cycle features. Time in all cases was represented by taking the cycle in question, decomposing it to radians, and then providing the sine and cosine of that feature to the models. As such we built a seasonal feature indicating the number of days through the year, a day night feature indicating how far through the diurnal cycle (cosines at 0 for sunrise and sunset) a time point was, and a final set of sines and cosines indicating progress through the lunar cycle. Given the range of sines and cosines are -1 to 1 no further normalization was done.

**Building the Models**

We approached the model building by building a log-odds model (?).

The first step was to split into training and validation sets. Given the low number of individuals in the sample and the intention to demonstrate the technique's value as an EDA tool it was decided to not maintain a hold out test set.

72 individuals were randomly selected for training and 39 for validation.

The next step in building a log-odds model is to decide on the formulation of our choices. In our case we took each of the individual depth bins as a choice where depth bins below the elevation at the site in question were not considered. As such our model ends up predicting the probability, given the data, of occurrence in any one of those depth bins. Training data was derived by identifying the actual depth bin occupied.

Next, we needed to determine the specifics of the contrast sampling. For this example, after inspecting the distribution of number of choices per salmon and number of choices per decision, we decided on random sampling (with replacement) 5,000 decisions per individual and 10 choices per decision.

Over a validation/training split of 39, 72 this resulted in 5,550,000 contrasts of which 3,600,000 were used in training and the rest in validation.

For each of the three models trained, the hyperparameters for the internal log-odds component of the model were parametrized in the following way:

|  |  |
| --- | --- |
| Component | Options |
| Layers | 2, 3 |
| Units per Layer | 24, 32 |
| Batch Size | 200,000 |
| Learning Rate | 0.001 |

We proceeded by grid search and used 6 separate seeds for each combination. Models were trained in Keras using an Adam optimizer for 75 epochs. Given this is a log-odds model we used categorical cross entropy as the loss function. Training was done on AWS Batch using Fargate instances of 2 vcpu's and 4 GB of memory.

Lowest loss (categorical cross entropy) at the end of the 75 epochs over the validation dataset was used to select the best set of parameters for each of the models trained.

**Visualizations**

Visualizations were either built using the training and validation data or by using a dataset built by rebuilding features but over every single H3 resolution 4 cell in a specified area (???) and over the course of a full year (2022). This second dataset allows us to see how the model behaves over the full course of the year in a full space as opposed to just over the training and testing data.

**Results**

**Modeling**

**Table 1: Model Selection**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Model | Features | NLP-C Train | NLP-C Val | NLP-D Train | NLP-D Val |
| Null |  |  |  | 1.835 | 1.740 |
| A | depth\_bin | 0.470 | 0.526 | 1.412 | 1.457 |
| B | A + season | 0.438 | 0.487 | 1.330 | 1.368 |
| C | B + diel | 0.429 | 0.480 | 1.313 | 1.352 |
| D | C + nitrate, salinity, mlt | 0.426 | 0.474 | 1.308 | 1.339 |

18 separate models were trained to explore the feature space with four (table 1) being the most salient. Judging off NLP-D Val we see a significant jump in going from a null model (random guessing) to a model aware of the depth bin (1.740 to 1.457). This is expected as the distribution of fish across depth bins was highly skewed toward the shallower depths (table 2) and the model was able to capture this skew.

**Table 2: Depth Skew**

|  |  |
| --- | --- |
| Depth Bin | Proportion of Samples |
| [0, 25] | 46.4% |
| (25, 50] | 16.6% |
| (50, 75] | 13.4% |
| (75, 100] | 10.7% |
| (100, 150] | 13.0% |
| (150, 200] | 4.9% |
| (200, 250] | 1.3% |
| (250, 300] | 0.6% |
| (300, 400] | 0.2% |
| (400, 500] | <0.1% |

Next, we see another substantial improvement in NLP-D Val by adding our season features (1.457 to 1.368). This is also expected as there are strong changes in depth occupancy of fish through the seasons (figure 1). We did observe a difference between the observed and predicted proportions of fish per depth bin in the validation data, however this is due to differences in those proportions between the validation and training data sets. However, in general, across both sets the pattern is the same – fish tend to move deeper in the winter months with a peak in the depth <=25 bin in the months of May and June.

A graph of a graph of a number of different colored squares

AI-generated content may be incorrect.

A graph of different colored squares

AI-generated content may be incorrect.

**Figure 1:** Expected and actual proportion of fish in each depth bin by month over the validation dataset using model B.

Next, model C sees a small improvement over B (1.368 to 1.352) with the inclusion of our diel features. While we clearly see a meaningful trend in the depth occupancy over the course of a day (figure 2) this feature has a smaller effect because there is a large degree of variation in this pattern across fish with some in the data exhibiting no diel pattern or even the opposite of the pattern described here.

A graph showing the value of a number of radians

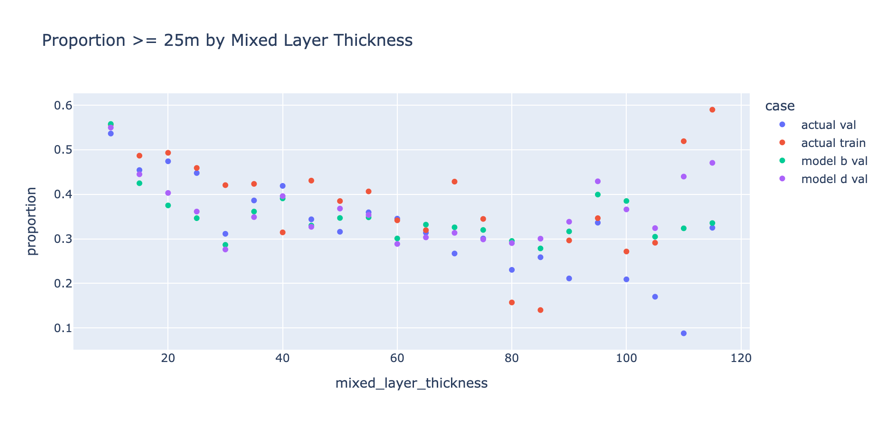
AI-generated content may be incorrect.

A graph of a graph

AI-generated content may be incorrect.

**Figure 2: Diel Variation in August** x-axis is the passage of the day in radians (starting at night moving to day at 0 and then passing through the day). Note the variation between the training set and the validation set.

Our final model (1.352 to 1.339 NLP-D Val) includes three environmental covariates, nitrate, salinity, and mixed layer thickness. These features were chosen from the slew of options as they were the only ones with an appreciable pattern in the data that also individually added value to the model in addition to the season and diel features. One interesting note is that the “value” of these features is somewhat hidden by the fact that they are all captured to some extent by the seasonality feature itself. For example for average proportion by mixed layer thickness in depth bin >= 25m (figure 3) we see no real difference between model b and model d indicating that seasonality is accounting for the broad patterns that mixed layer thickness can identify.

A graph with numbers and points

AI-generated content may be incorrect.A graph with colored dots

AI-generated content may be incorrect.

**Figure 3: Environmental Features** proportion by each feature for depth >= 25m for actuals in validation and training and two models. Note that models b and d are largely the same even though b does not actually possess a mixed layer thickness feature.

Finally, it should be noted that while elevation did not, in fact, end up as a feature this was because when elevation was added as a feature to the model it had no effect on the NLP-C Val. It should be noted however that there is an implicit addition of elevation in the model at low depths because while the log-odds will remain the same the number of options will shrink thereby boosting all the probabilities in lower depths. Due to this and a great deal of variation in the actual proportions by depth while elevation could drop the NLP-C over the training data it could not do so over the validation data.

**Outcomes**

This modeling effort demonstrates the value of a depth model in reducing Chinook salmon bycatch, particularly in the walleye pollock fishery. At its core, the model provides a risk assessment for fishing in specific areas, times, and depths, quantified as the summed probability over depths. In theory, this could guide “optimized” fishing patterns with minimal bycatch risk. However, such an approach would be overly simplistic, as fishers and managers must consider additional factors like abundance levels, fuel costs, vessel speeds, localized closures, and weather. Our model assumes fish are present; an abundance model could indicate that a high-risk area poses little actual threat due to low expected numbers. Thus, this tool should complement, not replace, other decision-making inputs. We aim to explore how this model informs choices at multiple levels—from broad seasonal planning and spatial trends to real-time decisions on trawl depths and timing—demonstrating its value as an informational resource.

**Seasonal Choices**

A map with red and blue squares

AI-generated content may be incorrect.

**Figure 4: Seasons with Minimal Risk** locations colored by the season that sees the minimal probability (at some point in the season) for depth range (25, 50].

In determining the months with minimal risk in the (25, 50] depth range and then mapping those months to their walleye pollock seasons (figure 4) we can see that season B is best for most near shore fishing in the Gulf of Alaska whereas in the EBS or far offshore season A captures the points in time with minimal risk.

A screenshot of a map

AI-generated content may be incorrect.

**Figure 5: Risk Sensitivity** locations are here colored by the difference in minimal risk between the best and worst months at that location. Risk differences are binned in increments of 5% absolute difference in probability.

We can also investigate with the model how sensitive timings are in each area over the months (figure 5). What we find is that in near shore regions in the GOA the difference in minimum probability in depth bin (25, 50] across months is around 5% whereas off coast it can range from 10-15%.

**Spatial Choices**

Rather than looking at what times are best at each location we can instead look across locations at a specific point in time (figure 6). Comparing two months from each of the seasons (August in season B and February from season A) we can see substantial differences in the patterning of risk across space for depth bin (25, 50]. Specifically in the August we see the risk increase as you move away from shore whereas in February we see the opposite and we also see a much steeper gradient in risk.

A map with a number of dots

AI-generated content may be incorrect.A screenshot of a map

AI-generated content may be incorrect.

**Figure 6: Fixed Month Spatial Variability in Minimum Risk** here we have per location in the months of August and February the minimum risk observed in that month for depth bin (25, 50]. Note that in the summer the risk falls off as you approach coastline in the GOA whereas in the winter the opposite is true.

**Local Temporal**

**A graph with different colored lines

AI-generated content may be incorrect.**

**Figure 7: Depth Occupancy Over Time** the predicted probability of each depth bin at lat=52.3 and lon=-128.4 over the course of 2022. Note the large diel variance between July and September in the [0, 25] depth bin.

Plotting the timeline of probability per depth bin across a full year for one h3 cell (84129c1ffffffff) shows us a great deal of interesting detail (figure 7). We see a gradual rise in the depth occupancy from September through July, and then a quite significant diel pattern beginning in mid-July that tapers off in September.

A graph of a graph

AI-generated content may be incorrect.A graph of different colored lines

AI-generated content may be incorrect.

**Figure 8: Depth Occupancy Over a Day** same as figure 7 but over a single day (times are in UTC).

We can zoom in on a specific set of days in July and August to look at the diel pattern more closely (figure 8). We see that the diel pattern peaks at 9a.m. UTC (midnight AKST) in both cases but that in July fish are rising to the full [0, 50] set of bins as opposed to August where they rise only to the [0, 25] bin and drop out of the (25, 50] like the rest of the deeper bins.

**Discussion**

This modeling exercise demonstrates that it is possible to construct an informative depth occupancy model that can aid fishermen in making decisions to mitigate bycatch risk. The key term here is inform—decision-making in fisheries involves multiple auxiliary objectives and additional constraints within which fishermen operate. Thus, an effective model must provide actionable insights tailored to these constraints.

The ability to tailor information to different levels of summarization or detail was demonstrated in the examples presented in our results.

* Broad-scale seasonal planning – The model identifies areas preferable for fishing in season A versus season B, providing a high-level decision support tool for large-scale planning.
* Localized spatial risk assessment – If a specific point in time is considered, the model helps assess spatial risk gradients within that period. For example, in August and February, results indicate that in summer, for depths between 25–50 meters, risk variation across space is minimal, making timing the more significant factor. However, in winter, risk varies substantially across space, meaning offshore fishing can provide a clear advantage.
* Temporal risk assessment – If location flexibility is limited, the model can be used to investigate risk differences across time, both seasonal and diurnal. In figures 7 and 8, reductions in bycatch risk were observed through careful selection of fishing windows, particularly in summer.
* Depth selection strategies – If neither location nor timing is flexible, selecting the appropriate depth can still reduce risk. Figure 8 illustrates unexpected depth occupancy patterns. In this case, lower depths (below 25–50 meters) showed increased risk during the day, suggesting that in the highlighted region, fishing between 25–50 meters may be the optimal strategy.

This ability to provide context-specific information on depth occupancy is a primary advantage of this approach. Prior studies have largely examined general shifts in mean depth occupancy over time or analyzed diurnal depth-use patterns at a population-wide level. While informative, these models lack the granularity needed for real-time, context-specific decision-making. In contrast, the present model accounts for fine-scale environmental and temporal dependencies, allowing for more precise risk assessments.

Other studies have developed models that are highly context-specific, but they have primarily focused on predicting the most likely depths at which Chinook will be found, rather than assessing risk across the entire water column. By incorporating probabilistic representations of depth occupancy across multiple environmental and temporal dimensions, this model provides a more comprehensive framework for evaluating bycatch risk.

**Challenges and Future Directions**

Despite its advantages, the model presents significant usability and digestibility challenges due to the large data volume it generates. Predictions cover every depth bin, at every hour, across all days within a specific year, for the entire Gulf of Alaska and Eastern Bering Sea—resulting in an overwhelming dataset. Processing such extensive data requires sophisticated computational resources and data analysis tooling. Without a structured means of distilling insights, the model’s practical utility remains limited.

To address this, an application layer is necessary. This layer must:

* Provide hierarchical data navigation, starting from high-level summaries and progressively refining insights to fine-scale details.
* Offer multiple views and filters, allowing decision-makers to incorporate specific constraints and easily interpret how different variables affect risk.
* Enable interactive visualization tools, so fishermen can dynamically adjust filters and quickly extract relevant risk assessments based on operational constraints.

Additionally, enhancing the model with expanded data sources could improve predictive capabilities. Environmental covariates, such as mixed-layer thickness, showed minimal effect in validation, likely due to their correlation with seasonality. More targeted data collection across environmental gradients—ensuring coverage across different seasons and times of day—could allow for better differentiation of these effects.

Another critical area for improvement is stock-specific risk assessment. Chinook salmon exist in multiple substocks, some more vulnerable than others. If fishermen had access to fine-scale depth occupancy assessments that differentiate between high-risk and low-risk substocks, they could refine their avoidance strategies further. Expanding data collection efforts to include genetic or tag-based identification of these substocks would be a valuable step forward.

Importantly, these improvements do not require changes to the model’s methodology. Instead, they involve the inclusion of additional data and features. Because this model leverages machine learning, incorporating new information would enhance predictive accuracy automatically without requiring fundamental structural modifications.

Finally, risk assessment is most effective when paired with abundance forecasts. The current model estimates the likelihood of fish occurring at different depths, given their presence, but it does not predict overall abundance. Avoiding Chinook hotspots is a primary strategy for bycatch reduction, and integrating abundance data would allow fishermen to identify low-risk alternatives with greater confidence. Even in high-risk depth zones, if overall Chinook abundance is predicted to be low, the practical risk of bycatch remains manageable.

**References**

1. Bering Sea Salmon Bycatch Update

2. Fisheries of the Exclusive Economic Zone Off Alaska; Chinook Salmon Bycatch Management in the Bering Sea Pollock Fishery

3. A dynamic ocean management tool to reduce bycatch and support sustainable fisheries

4. What are we protecting? Fisher behavior and the unintended consequences of spatial closures as a fishery management tool

5. https://www.adfg.alaska.gov/index.cfm?adfg=walleyepollock.printerfriendly

6. Behavior and thermal environment of Chinook salmon Oncorhynchus tshawytscha in the North Pacific Ocean, elucidated from pop-up satellite archival tags

7. Chinook salmon depth distributions on the continental shelf are shaped by interactions between location, season, and individual condition

8. Understanding the behavior and ecology… Michael