Multivariate landing model for the CPS fishery

Felipe J. Quezada

August 2, 2021

1 Introduction

- Research question:
  - **Big question:** What is the effect of climate change on fish landings?
  - **Narrow question:** How does changes in the presence of sardine/anchovy/market squid/herring affect landings at ports located in the US west coast?
- Contribution:
  - Interaction between species.
  - Better understanding of fishers species portfolio

2 Coastal Pelagic Species fishery

Mean revenues by species are shown in Figure **??**, while mean landings by species are shown in Figure **??**.

Figure 2 shows landing over time by species, while Figure **??** shows landing and prices over time for selected CPS species.

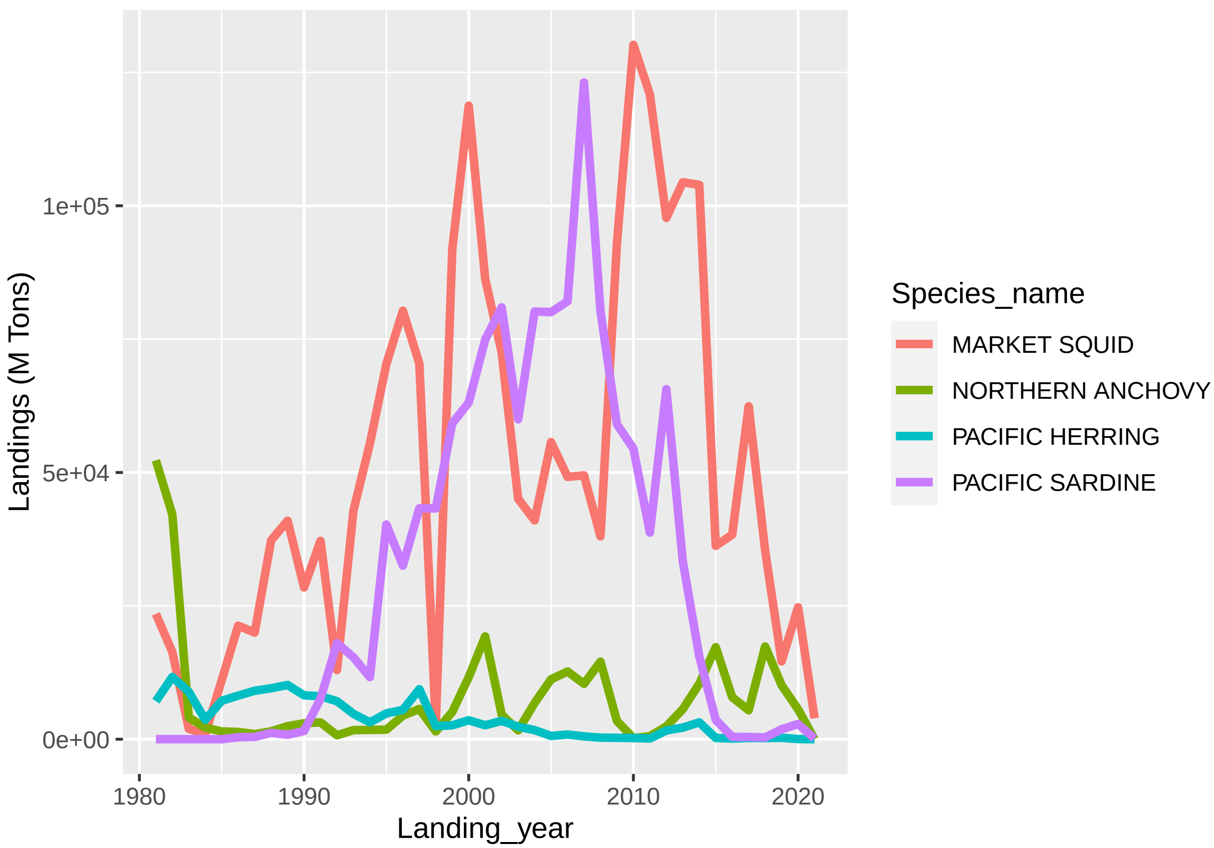


Figure 1: Total annual landing by CPS species.

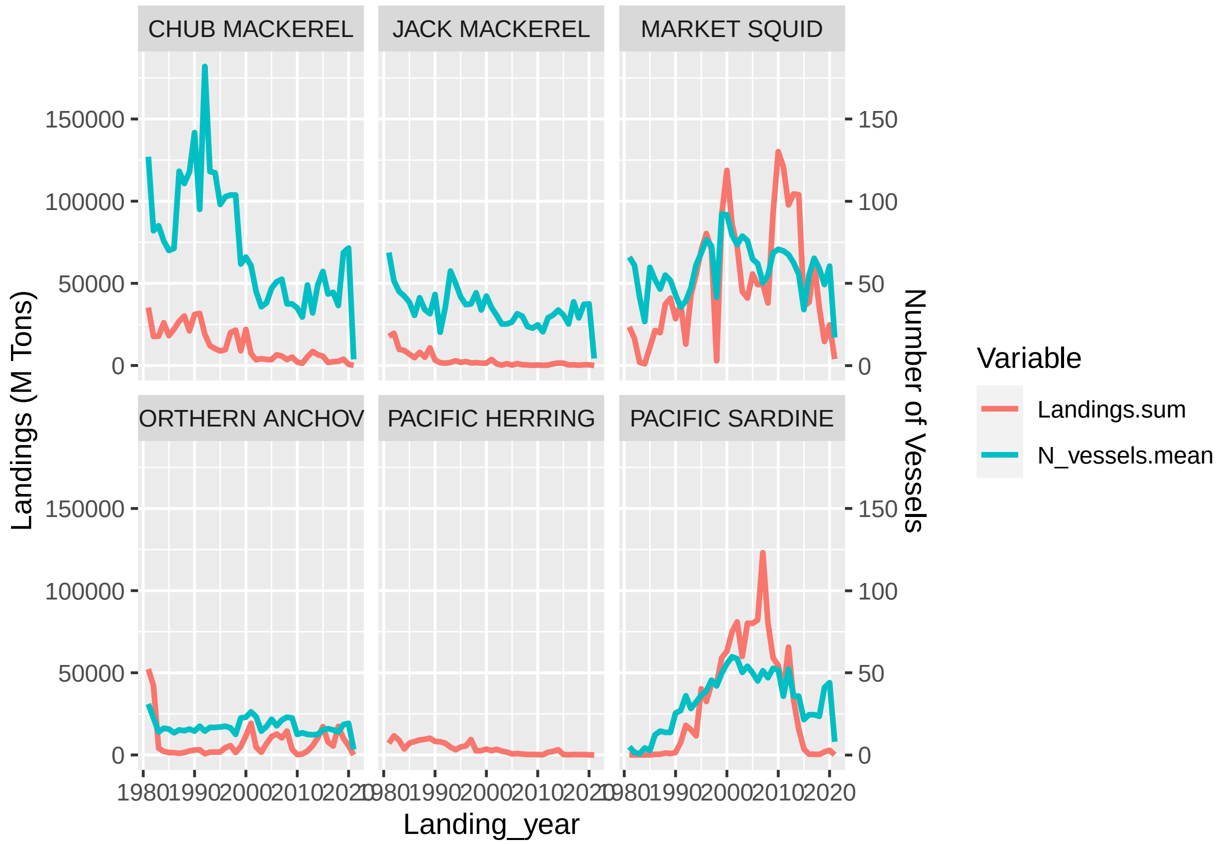


Figure 2: Total annual landing by CPS species.

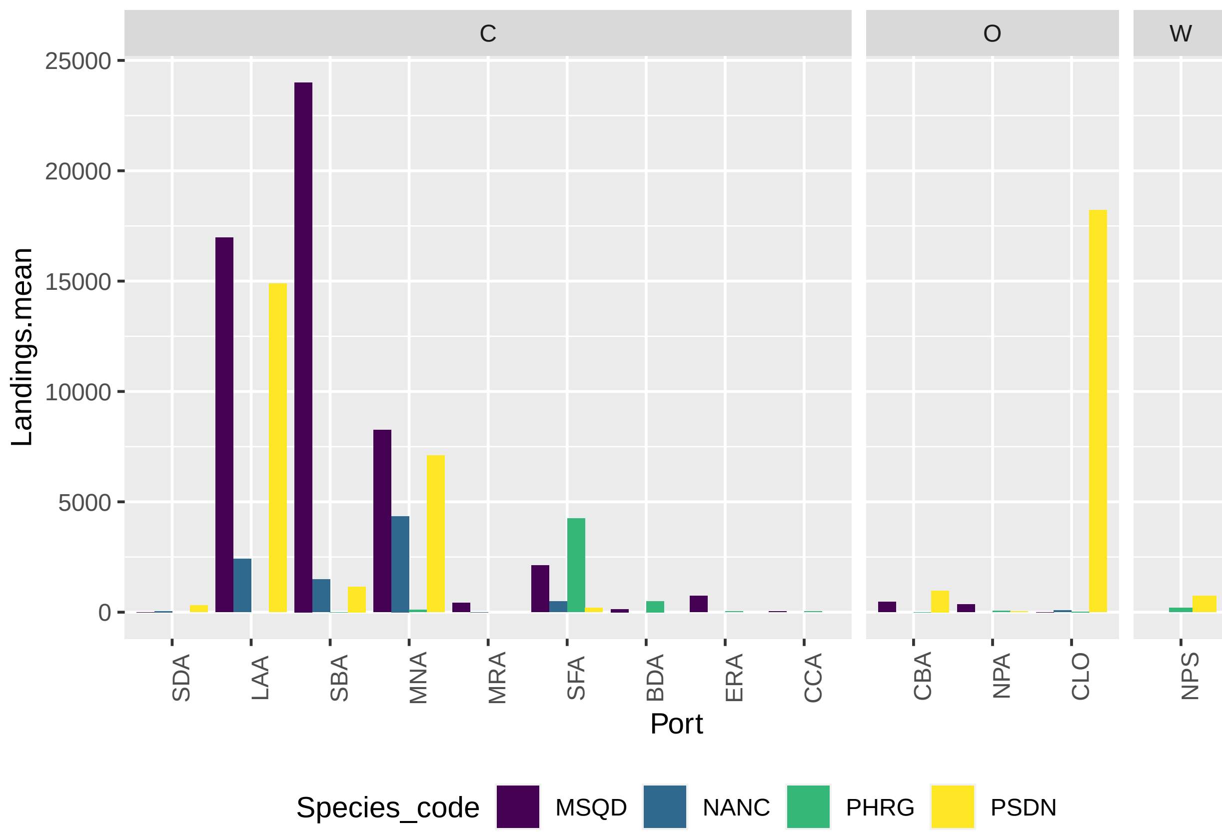


Figure 3: Landing by Port.

Figure 4: Annual average landing by area port. *Notes:* BDA = Bodega Bay; BRA = Brookings; CBA = Coos Bay; CCA = Crescent City; CLO = Columbia River (OR); ERA = Eureka; LAA = Los Angeles; MNA = Monterey; MRA = Morro Bay; NPA = Newport; NPS = North Puget Sound; SBA = Santa Barbara; SDA = San Diego; SFA = San Francisco.

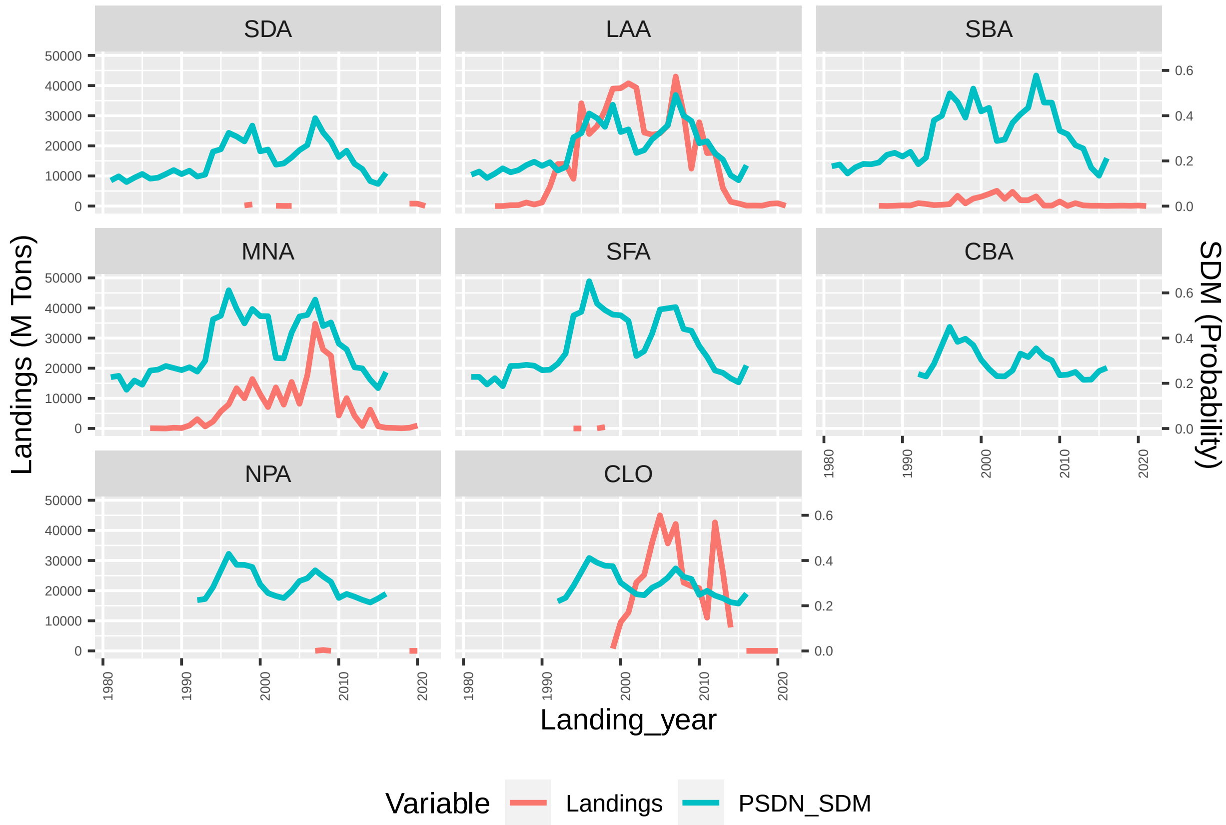
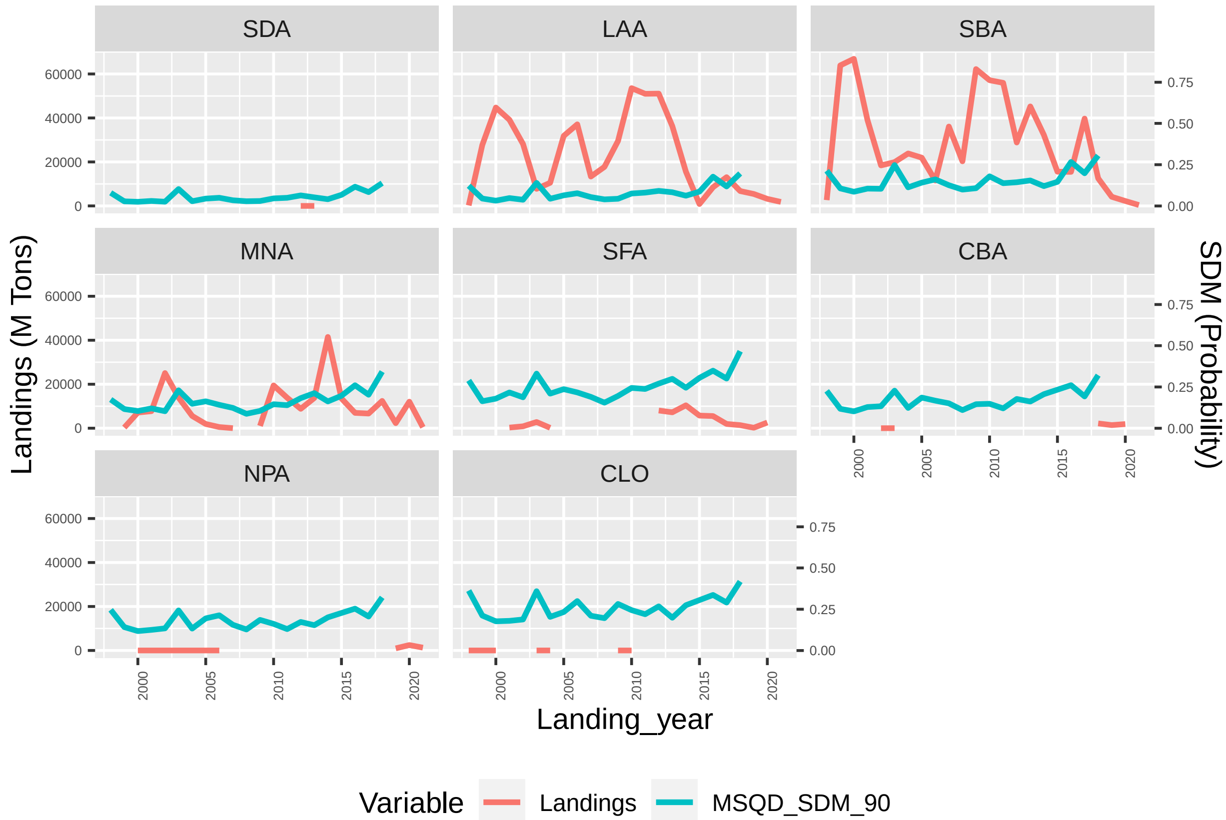
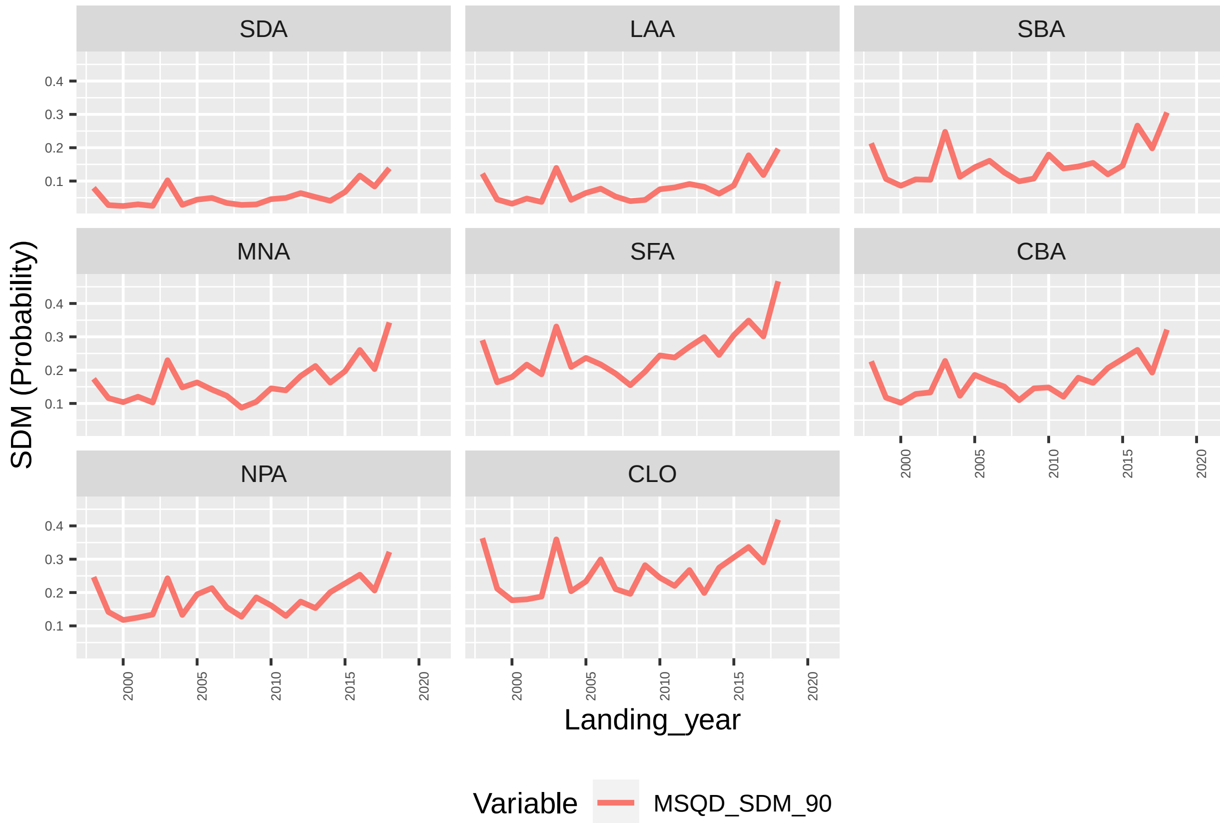


Figure 5: SDM mean for Pacific Sardine by area port. *Notes:* BDA = Bodega Bay; BRA = Brookings; CBA = Coos Bay; CCA = Crescent City; CLO = Columbia River (OR); ERA = Eureka; LAA = Los Angeles; MNA = Monterey; MRA = Morro Bay; NPA = Newport; NPS = North Puget Sound; SBA = Santa Barbara; SDA = San Diego; SFA = San Francisco.

```
## Warning: Removed 13 row(s) containing missing values (geom_path).
```




```
## Warning: Removed 3 row(s) containing missing values (geom_path).
```



3 Empirical model

3.1 Data

We use as outcome variable landing data public available from PacFIN. It is a yearly panel data for the period 1980 - 2020 that contains landings of commercial species by ports located in the west coast of the United States (i.e. Washington, Oregon and California).

Our main treatment variables are species presence. This variables were obtained from Species Distributiuon Models (SDM). Future forecast on this variable allows us to simulate a fishery in the future.

Additional explanatory variables are the annual catch limit (ACL) for the species in consideration and

Summary statistics for the whole data set is shown in Table 1.

3.2 Model

In general, landings are conditional to biological stocks (affected by climate change), harvest cost, prices and regulations.

- Outcome variable:
  - Pacific sardine landings by vessel, port and year. In this case we would predict landing by vessel i that land in port j
- Treatment variable:

Table 1: Descriptive statistics.

	Mean	Std.Dev	Min	Max
ACL_MSQD.mean	108213.70	2459.39	107047.00	113398.00
ACL_PSDN.sum	42014.25	55826.16	0.00	186791.00
ACT_PSDN.sum	42503.50	55335.76	0.00	186791.00
Landing_year	2001.89	11.65	1981.00	2021.00
Landings	293.31	2172.46	0.00	66890.30
MSQD_SDM_120	0.14	0.06	0.02	0.37
MSQD_SDM_60	0.22	0.11	0.03	0.56
MSQD_SDM_90	0.18	0.08	0.02	0.47
N_dealers	20.38	13.01	0.75	68.00
N_vessels	42.10	31.42	0.75	182.00
Price	1.37	1.79	0.00	28.53
PSDN_SDM	0.30	0.11	0.10	0.65
Revenue	347993.43	1835052.50	0.00	49987499.00

- Change spatial distribution of species due to climate.
  * Smith et al. [2021] relate port-level landings to the probaility of presence from a sardine species distribution model (SDM), and then makes projection to quantify future changes in landings.
  * Specifically, Smith et al. [2021] use mean monthly probability of presence of sardine within 60 km of port.

- Explanatory variables
  - Harvest costs (e.g., distances and fuel cost)
  - Own price and substitutes
  - Effort
    * Fishing effort from matching PacFIN data to Global Fishing Watch
    * "VESSEL_NUM" in PacFIN data: "It can be a USCG VID (ex: 1234567 or AK1234nn) or MISSING or UNKNOWN if vessel ID not provided or invalid. It is also "Null" if no vessel was used."
  - Regulations.
    * Incorporate ACL in the model (Smith et al. [2021] obtained this information from CPS Fisheries Management Plan, and Federal Register), maybe with a function that have a ceiling limit, or as **censored data** (Stan code).
    * Port capacity.
    * Closures.
- Random-coefficients?
  - By port: Different intercept / coefficient by vessel port.
- Some considerations:
  - Vessels likely to have contract with ports. Less flexibility where they land.
  - Harvest happen near-shore, less probability of longer trips if species move further.
  - Multivariate framework allows us to consider interrelation between species and CPS fleet.
    * Sardine harvest affect squid harvest?
    * If Sardine opens... Sardine more preferred? Sequentially or simultaneous harvest?

3.2.1 Empirical strategy

- Statistical model:

- Bayesian hierarchical model
  - Hierarchical effects by port.
  - Uncertainty from modeling the process (as well as from the imperfect observation of the process).
  - Model the zeros in data
    - Included in the estimation.
    - N/A landings were converted to zero.
    - Closures.
    - Port restrictions (i.e. no infrastructure)
- Smith et al. [2021] estimate with GAM framework (allows for non-linear relationships)
- Some details:
  - Spatial autocorrelation between ports through species abundance? Between areas where vessel harvest?
    - Morris et al. [2019] include spatial errors in a bayesian framework.
    - How to deal with non-linearities
    - How far vessel travel for each species? (See animation here)

3.2.2 Bayesian model

We estimate a Hierarchial Bayesian Hurdel model for each species. The model have the following structure:

$$[\alpha_i, \beta_i, \sigma^2_\alpha, \sigma^2_\beta, \Sigma | q_{i,t}] \propto \text{multivariate normal}(q_{i,t} | \alpha_i + \beta_i SDM_{i,t}, \Sigma) \times \text{normal}(\alpha_i | \mu_\alpha, \sigma^2_\alpha)$$
$$\times \text{normal}(\beta_i | \mu_\beta, \sigma^2_\beta) \times [\mu_\alpha][\mu_\beta][\sigma^2_\alpha][\sigma^2_\beta][\Sigma],$$

where $q_{i,t}$ is the observed landings in port $i \in (1, \ldots, L)$ at year t, L is the total number of ports, $SDM_{i,t}$ is the SDM output for sardine observed at port i during the landing year t, α_i is the coefficient estimate by port, β_i is the effect of SDM on landing by port, and Σ is a matrix covariances.

- Changes on estimations:
  - SDM outputs of Market Squid included (using first 60km within coast)
  - Separate equation depending on species
  - Time fixed effects included.
  - Separate correlation matrix (0.37 positive correlation between sardine and squid)

4 Results

4.1 Pacific Sardine model

4.1.1 Compare models

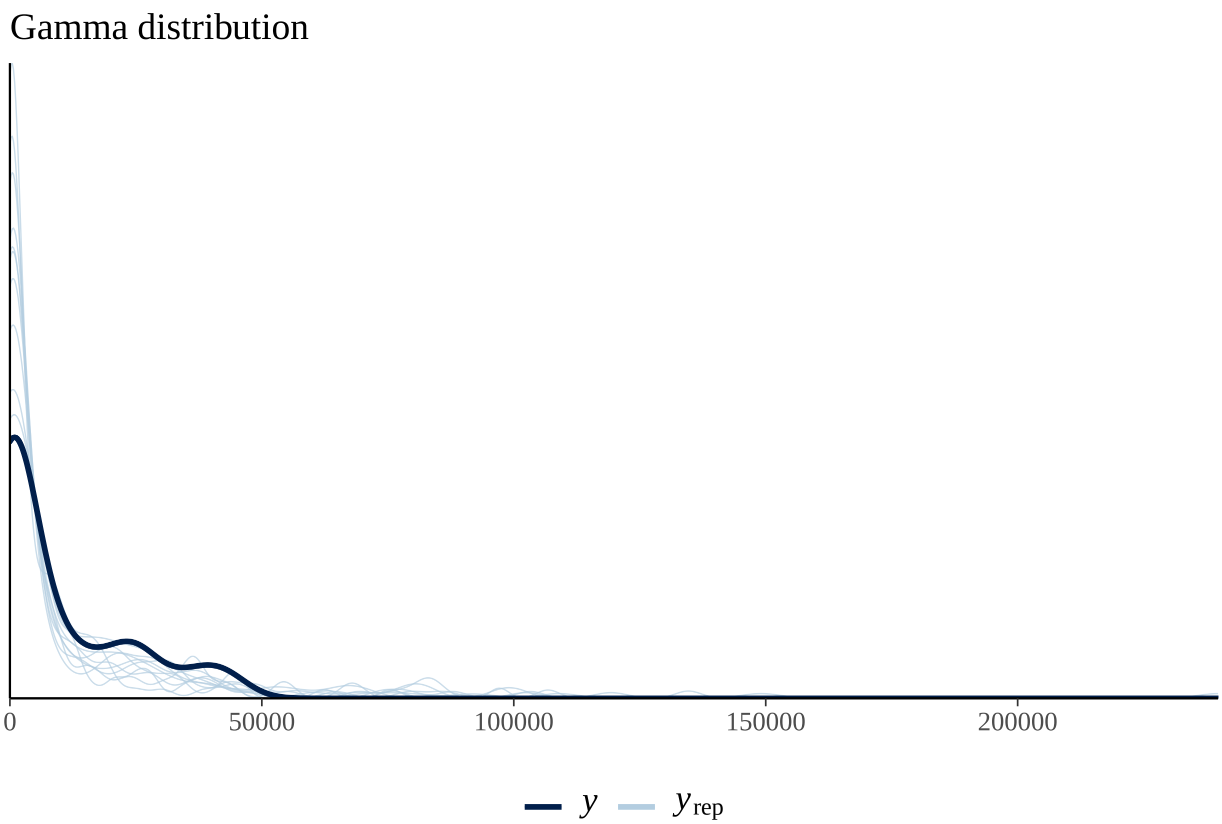
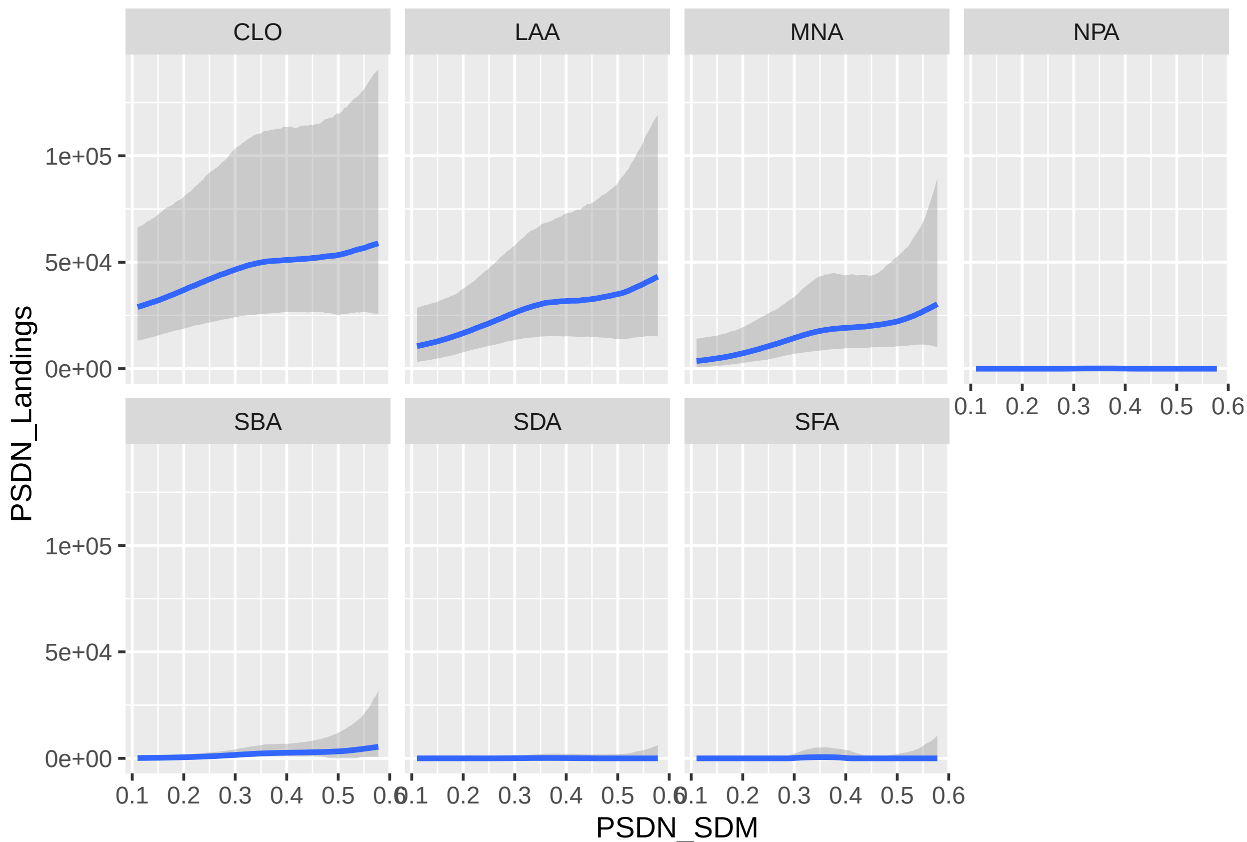
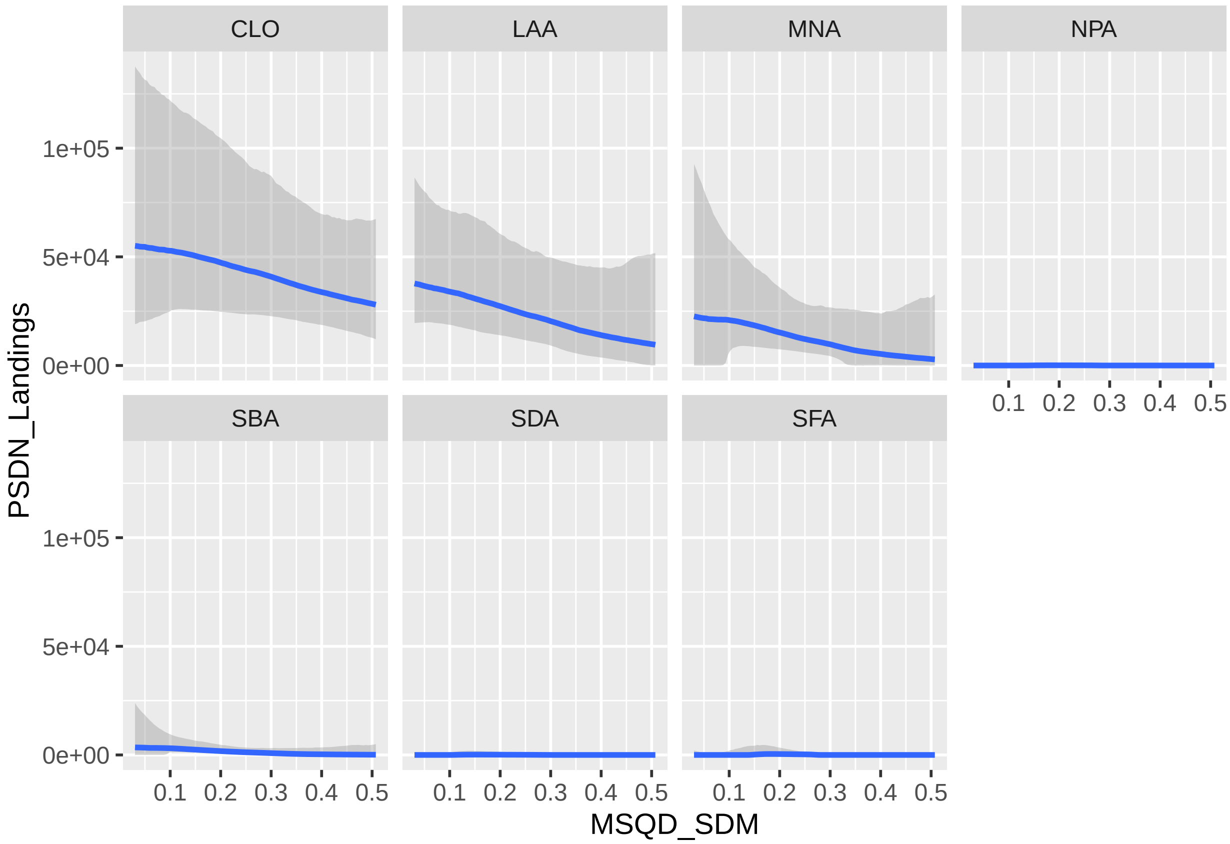


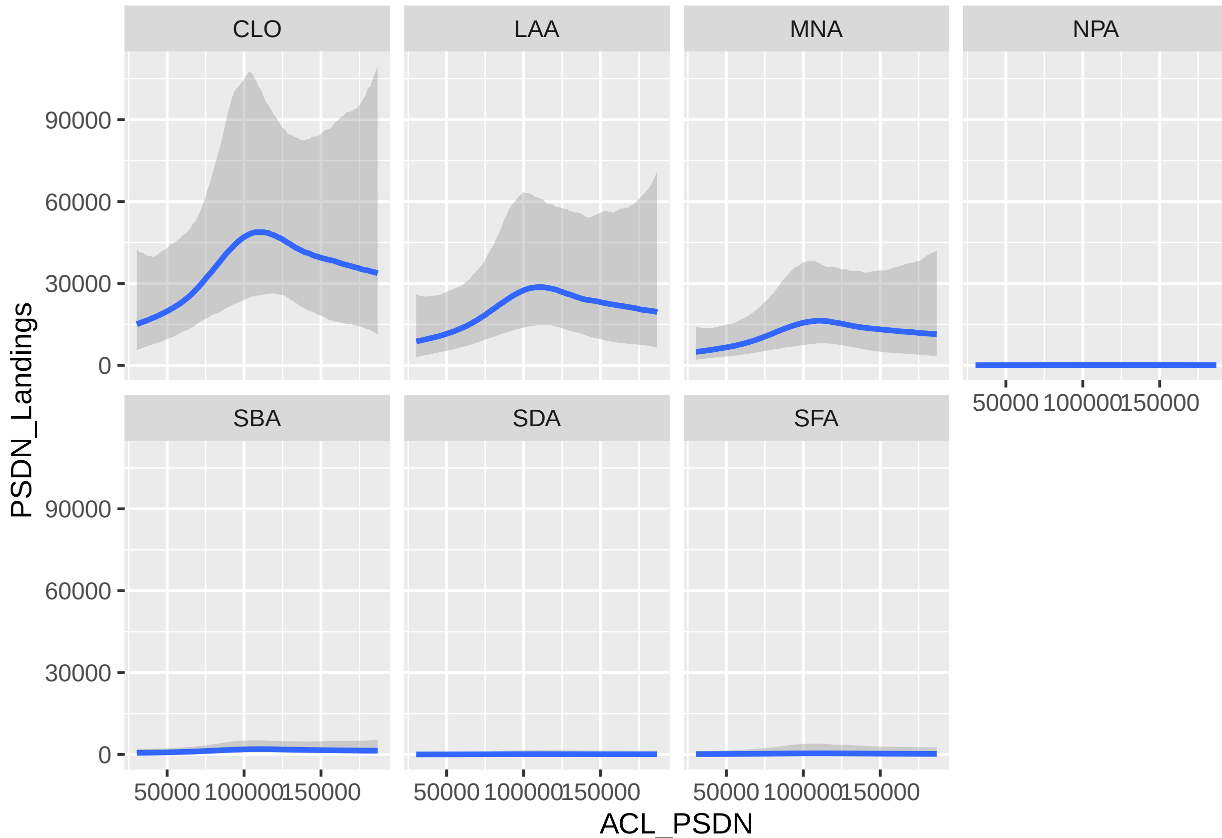
Figure 6: Graphical posterior predictive checks. (Gamma distribution)

4.1.2 Effect of SDM's on Sardine landings

4.1.3 Results by port




CLO
LAA
MNA
NPA
SBA
SDA
SFA
PSDN_Landings
1e+05
5e+04
0e+00
0.1 0.2 0.3 0.4 0.5
MSQD_SDM



CLO
LAA
MNA
NPA
SBA
SDA
SFA
PSDN_Landings
90000
60000
30000
0
50000 100000 150000
ACL_PSDN

4.1.4 Interactive effects between Pacific Sardine and Market Squid SDMs

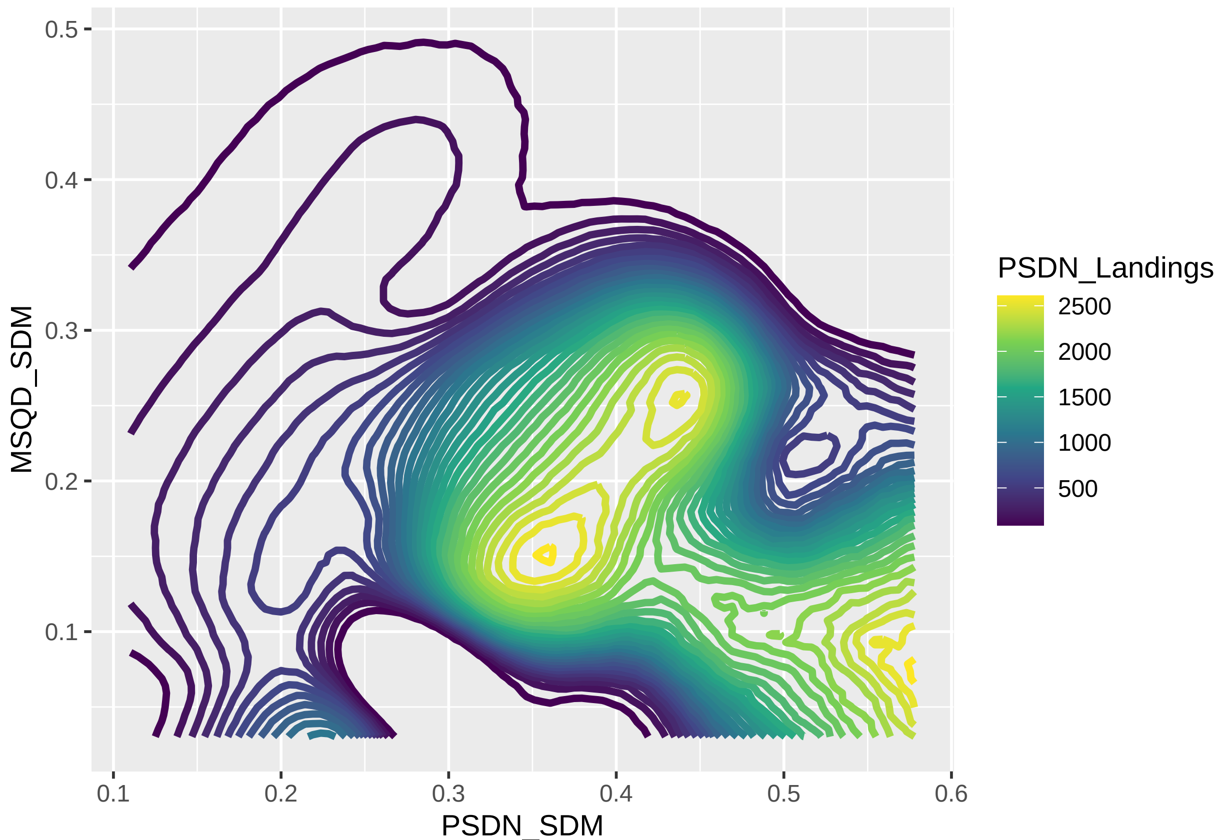
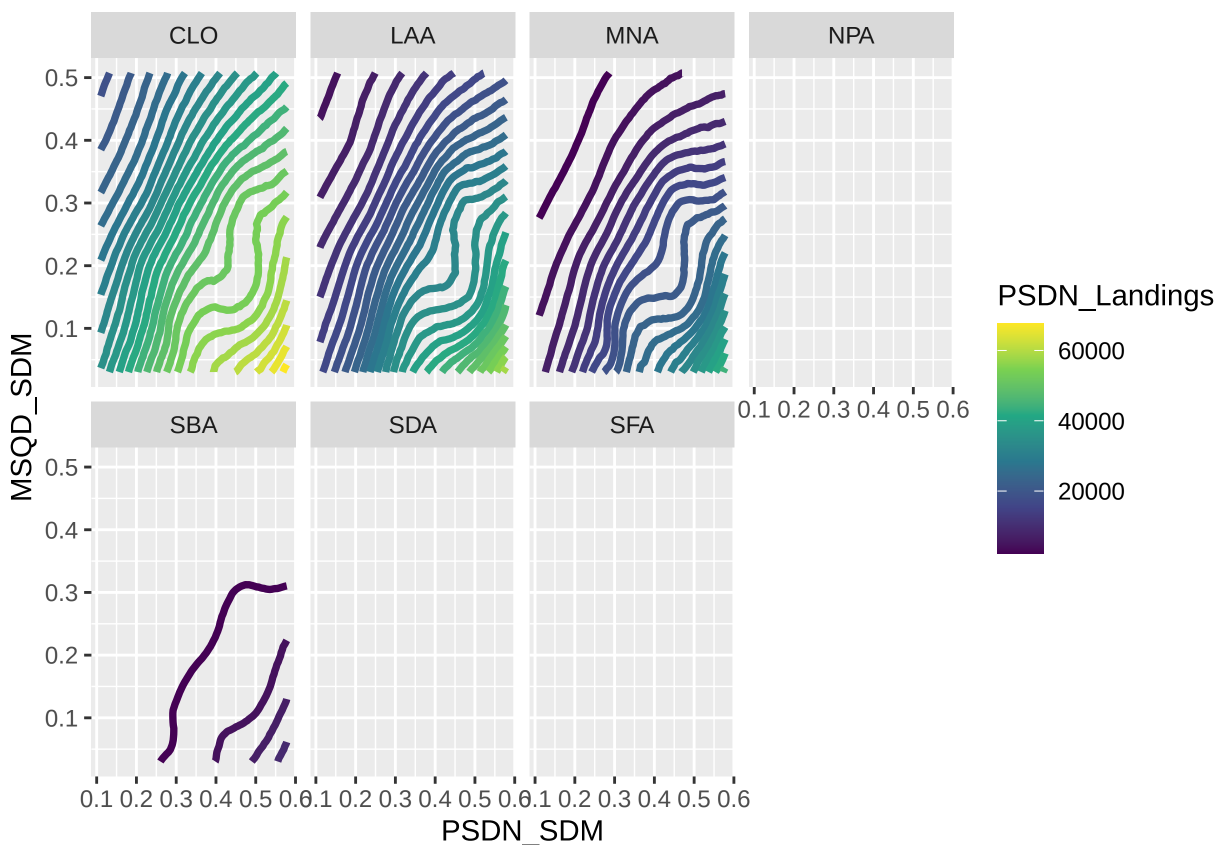


Figure 7: Conditional effect of SDMs on Pacific Sardine landings.

Table 2: Comparison of models fit1 to fit3 based on approximate leave-one-out cross-validation. Market squid landigns model.

	ELPD-LOO	ELPD-Diff	SE-Diff	P-LOO	LOOIC	Akaike-Weight
fit_qMSQD_gamma	-890.66	0.00	0.00	36.45	1,781.32	1.00
fit_qMSQD_lognormal	-903.40	-12.74	8.03	37.49	1,806.79	0.00

Note. ELPD-LOO = expected log posterior predictive density (higher is better); ELPD-DIFF = difference in ELPD values compared to the best model. SE-DIFF = standard error of the ELPD difference. P-LOO = effective number of model parameters (lower is better); LOOIC: leave-one-out information criterion (lower is better); Akaike-Weight = Model weight based on the LOOIC values (higher is better).



4.2 Market Squid model

4.2.1 Compare models

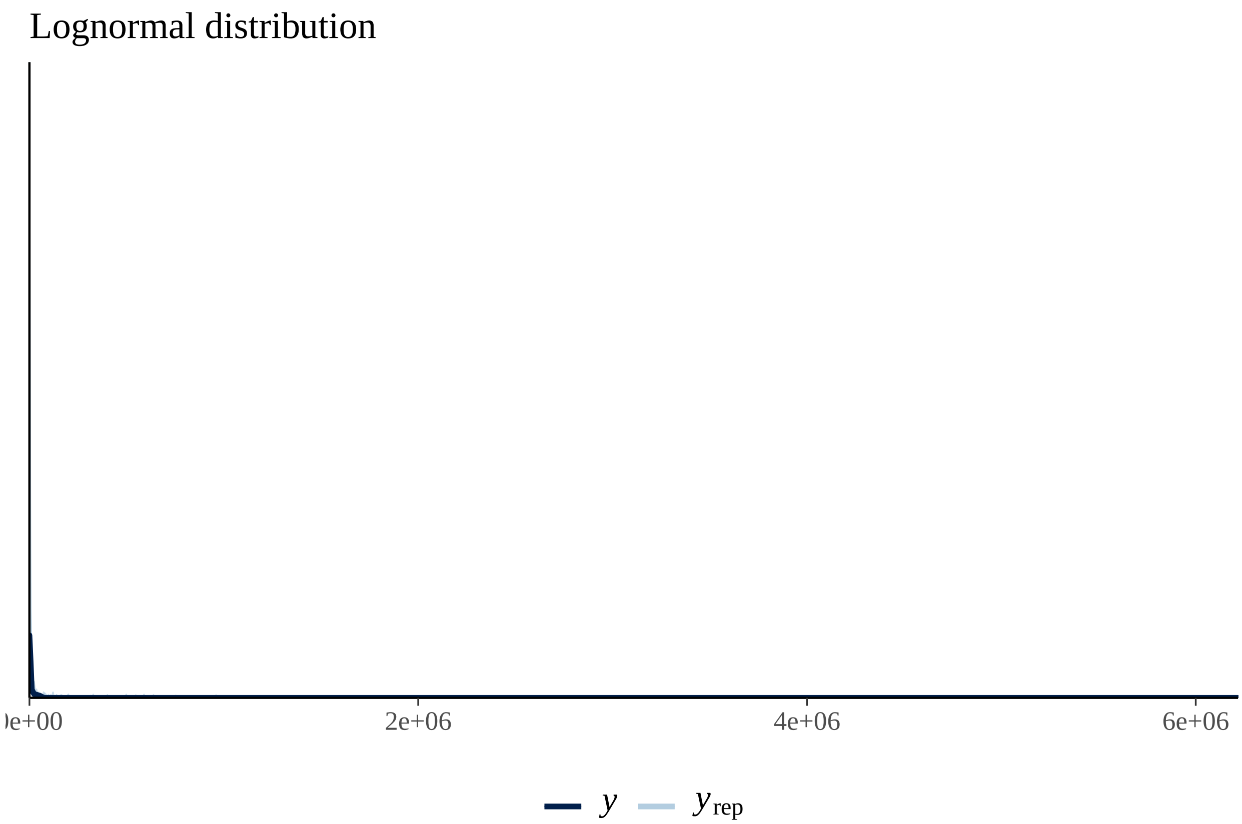


Figure 8: Graphical posterior predictive checks for Market Squid landings model. (Gamma distribution)

4.2.2 Compare models

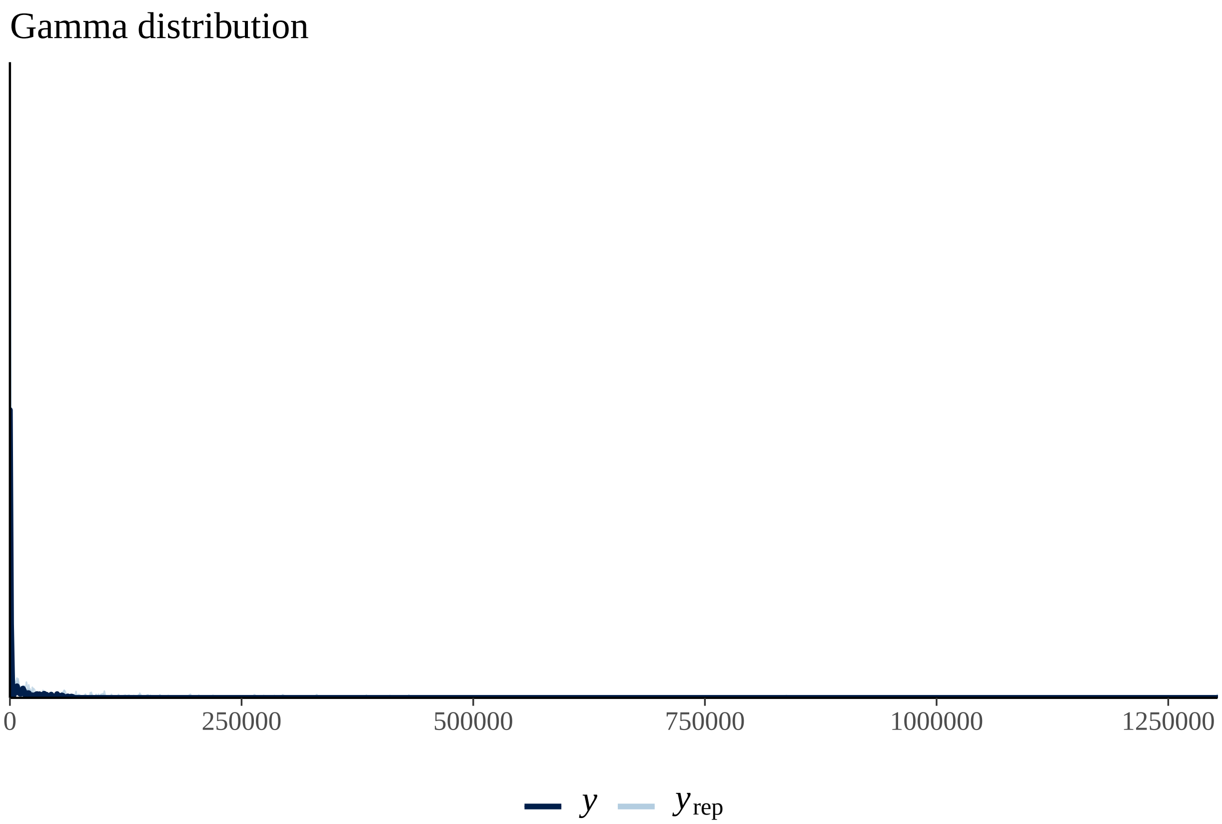


Figure 9: Graphical posterior predictive checks for Market Squid landings model. (Gamma distribution)

4.2.3 Effect of SDM's on Market SquidSardine landings

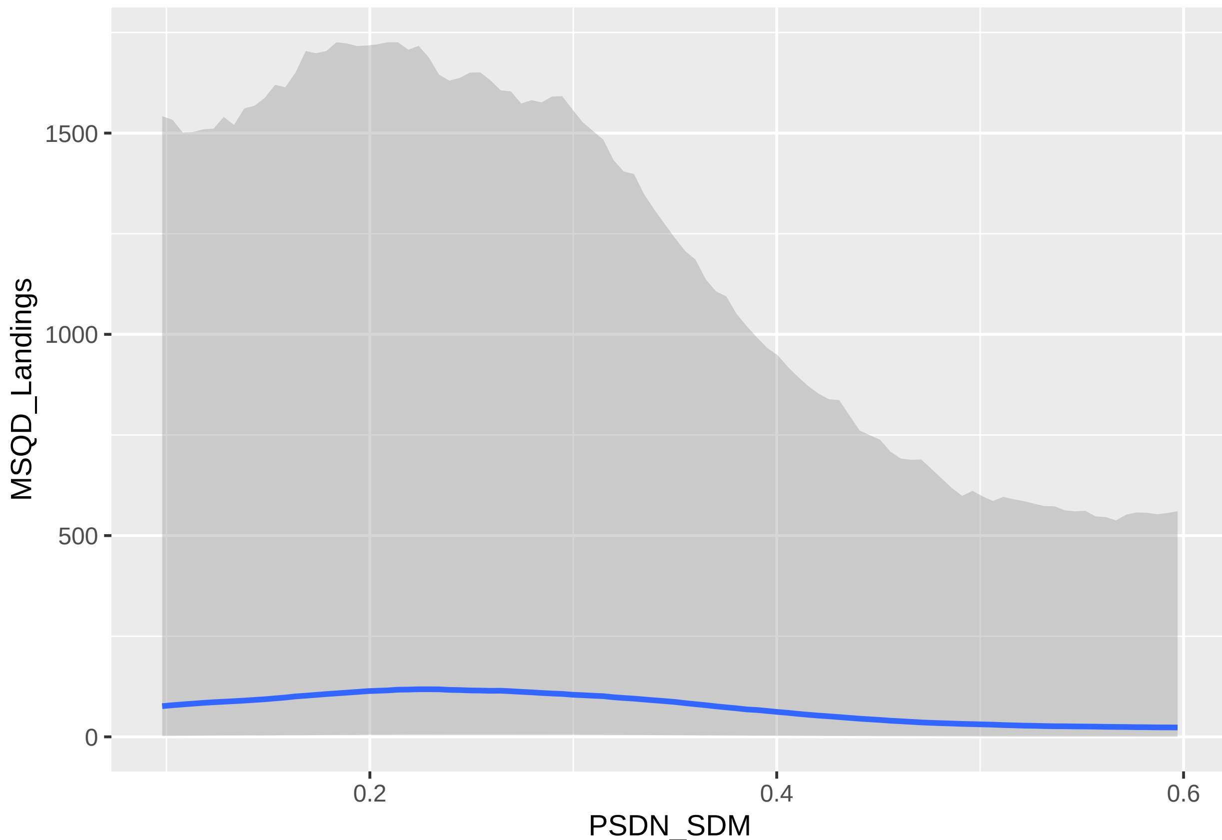
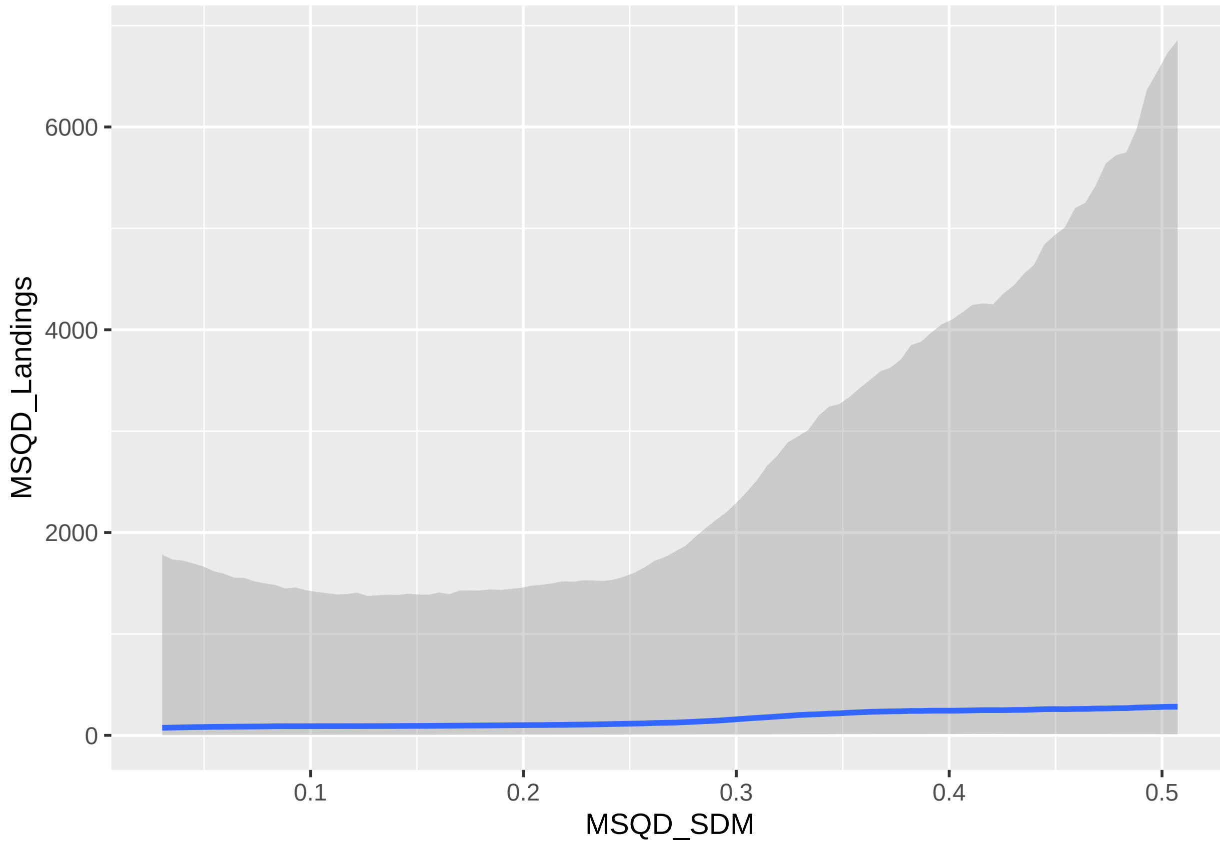
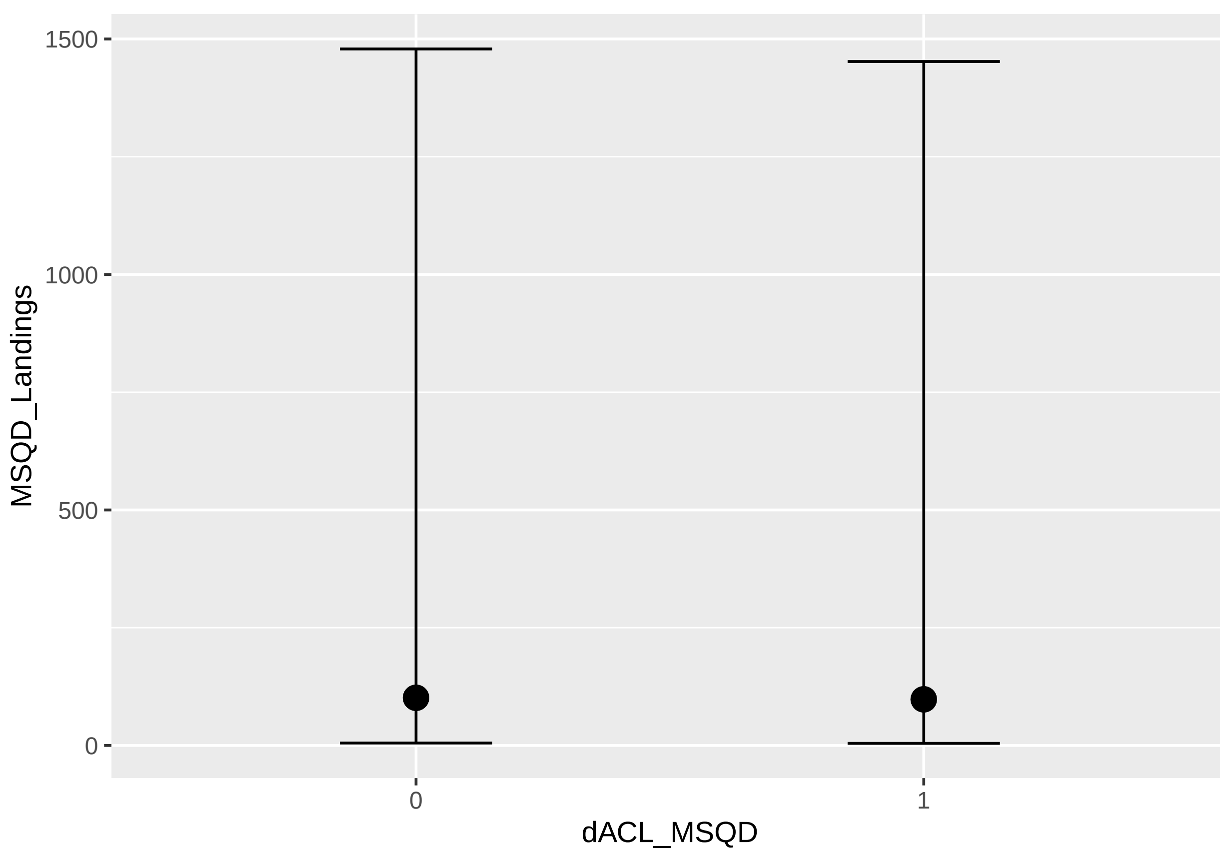
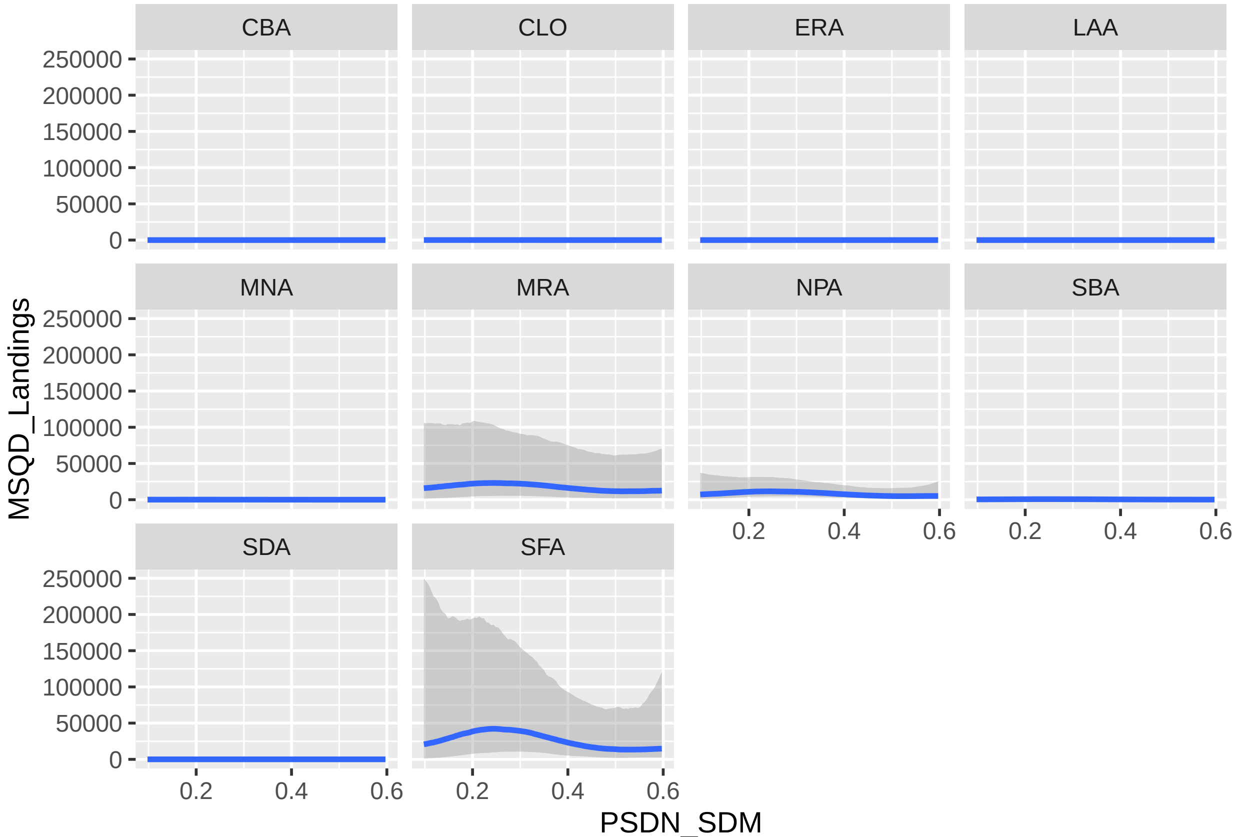


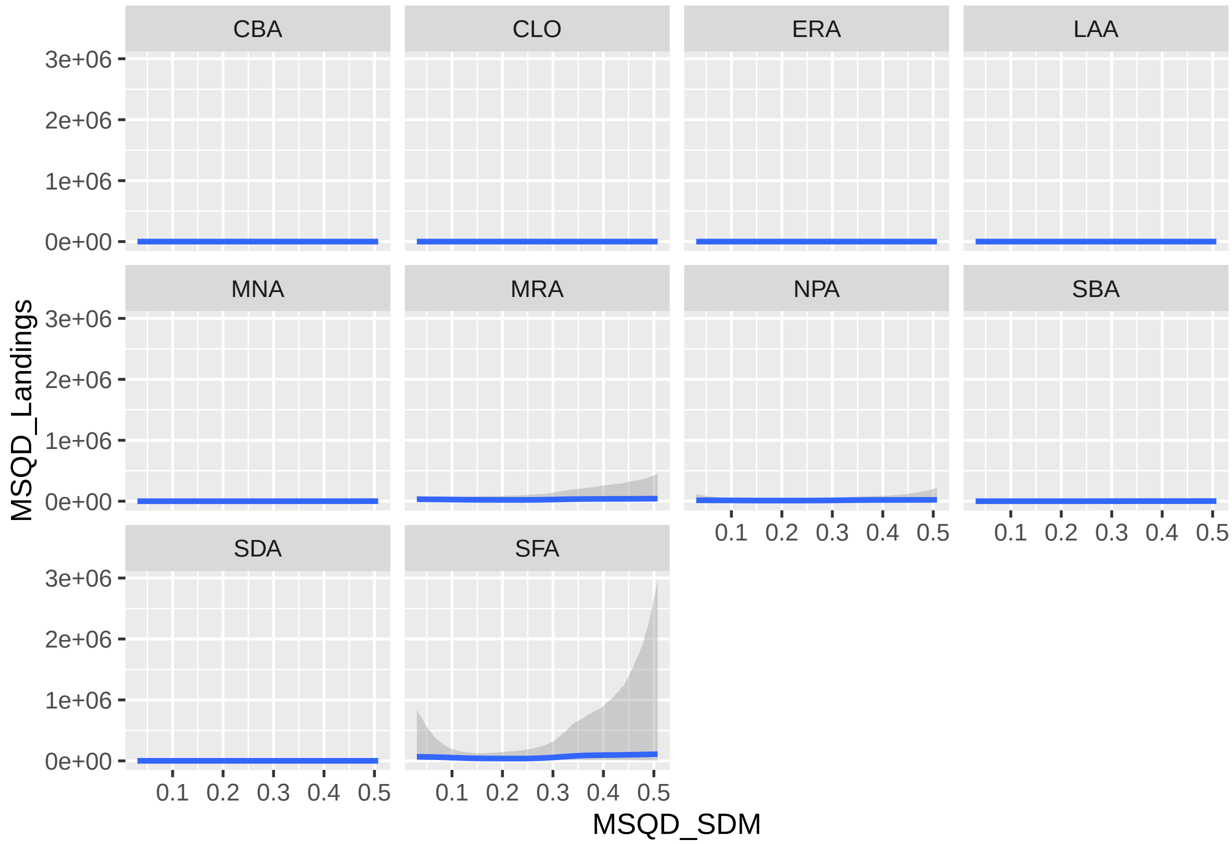
Figure 10: Conditional effect of SDMs on Pacific Sardine landings.


6000
4000
2000
0
MSQD_Landings
0.1
0.2
0.3
0.4
0.5
MSQD_SDM


1500
1000
500
0
MSQD_Landings
0
1
dACL_MSQD

4.2.4 Results by port




CBA
CLO
ERA
LAA
MNA
MRA
NPA
SBA
SDA
SFA
3e+06
2e+06
1e+06
0e+00
0.1
0.2
0.3
0.4
0.5
MSQD_Landings
MSQD_SDM

4.2.5 Interactive effects between Pacific Sardine and Market Squid SDMs

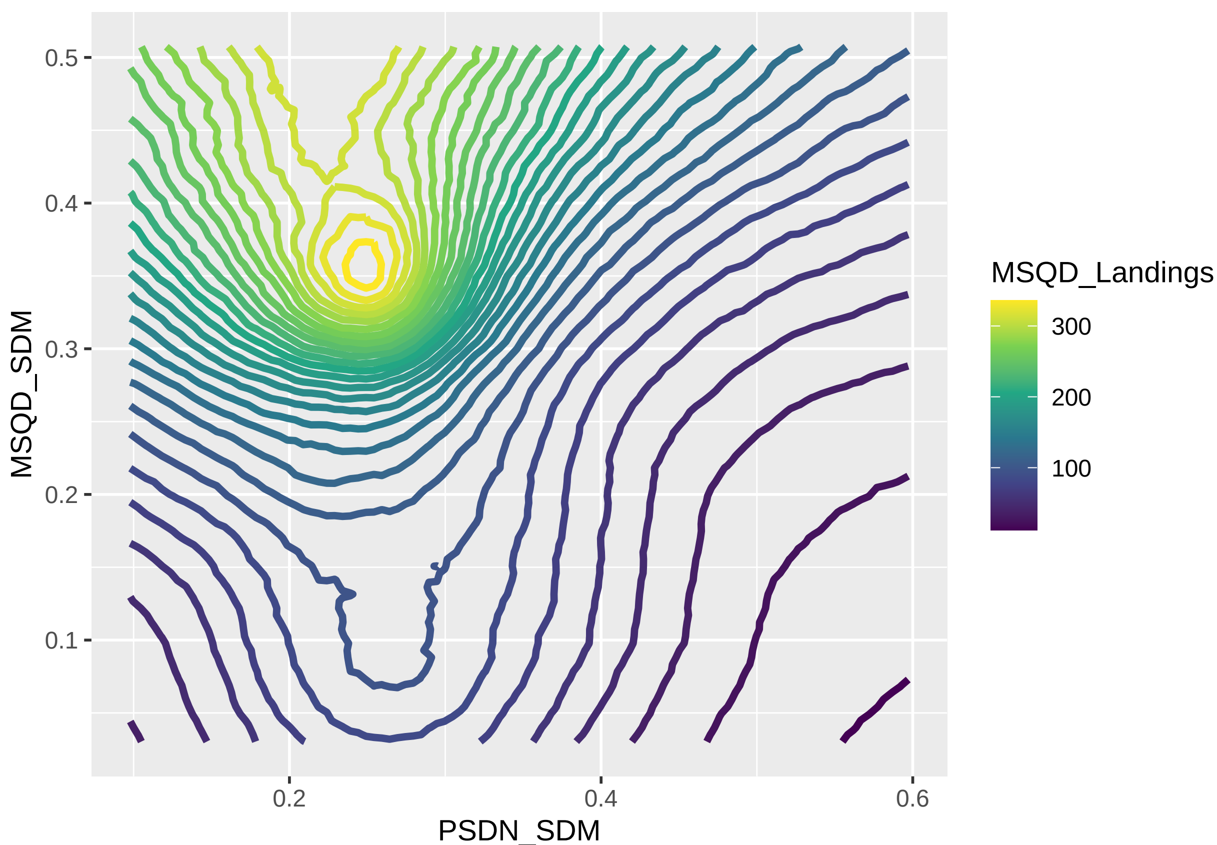
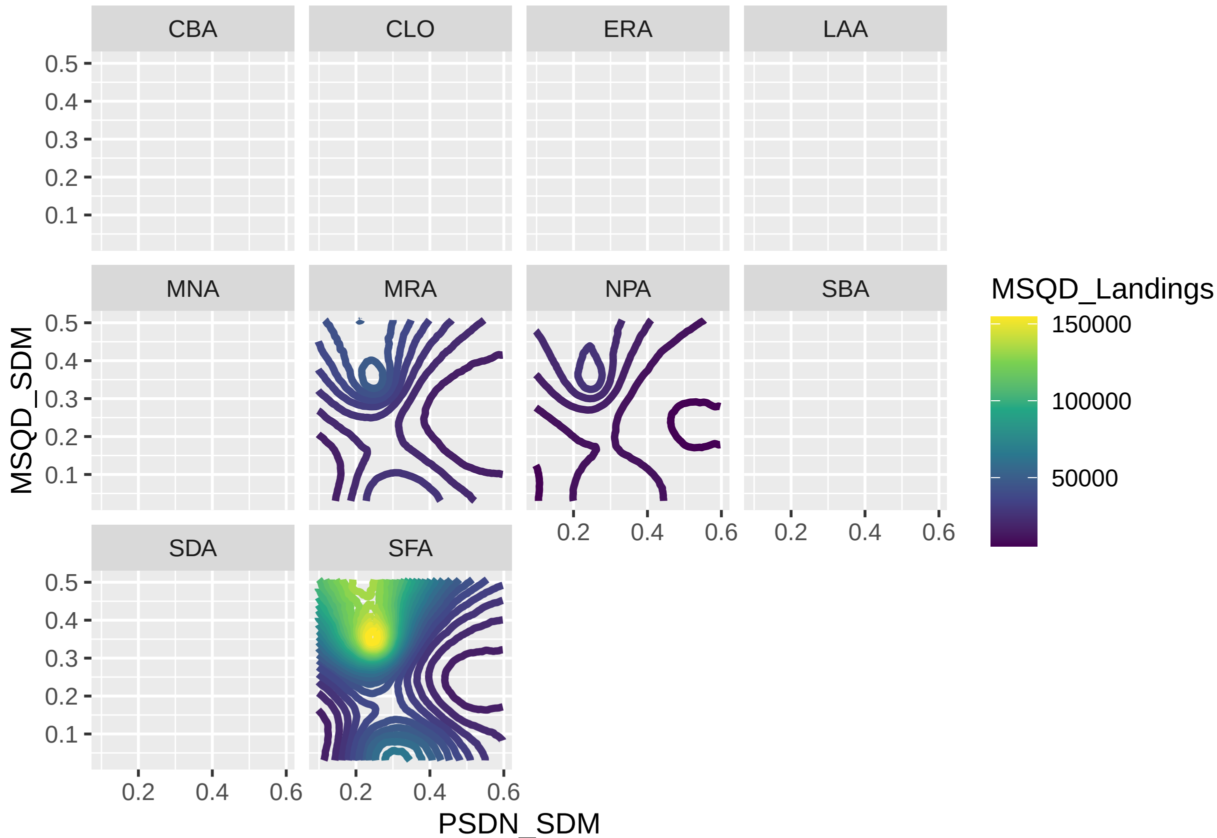


Figure 11: Conditional effect of SDMs on Pacific Sardine landings.



5 Conclusions

Appendix

Mitzi Morris, Katherine Wheeler-Martin, Dan Simpson, Stephen J Mooney, Andrew Gelman, and Charles DiMaggio. Bayesian hierarchical spatial models: Implementing the besag york mollié model in stan. *Spatial and spatio-temporal epidemiology*, 31:100301, 2019.

James A Smith, Barbara Muhling, Jonathan Sweeney, Desiree Tommasi, Mercedes Pozo Buil, Jerome Fiechter, and Michael G Jacox. The potential impact of a shifting pacific sardine distribution on us west coast landings. *Fisheries Oceanography*, 2021.