SCIENTIFIC COMMITTEE
TENTH REGULAR SESSION

Majuro, Republic of the Marshall Islands
6-14 August 2014

ANALYSIS OF WCPO LONGLINE OBSERVER DATA TO DETERMINE FACTORS IMPACTING CATCHABILITY AND CONDITION ON RETRIEVAL OF OCEANIC WHITE-TIP, SILKY, BLUE, AND THRESHER SHARKS (SUPPLEMENTARY INFORMATION)

WCPFC-SC10-2014/EB-WP-01a
(SUPPLEMENTARY INFORMATION)

B Caneco[1], C Donovan[1], and S Harley[2]

[1] DMP Stats, The Coach House, Mount Melville House, St Andrews, Scotland KY 168 NT, UK

[2] Secretariat of the Pacific Community, Oceanic Fisheries Programme.

DMP STATS

DMP STATS The Coach House, Mount Melville House, St. Andrews, Scotland KY 168 NT, UK +44(0) 1334 477 544 mail@dmpstats.com

CONTENTS

1 EXECUTIVE SUMMARY

We present here an overview of the data-acquisition & treatment, exploratory data analysis and preliminary modelling conducted for phase 1 of the "*Analysis of WCPO longline observer data to determine factors impacting catchability and condition on retrieval of blue, mako, oceanic white-tip, silky, hammerhead and thresher sharks*" [*contract CC14/169*].

Key points from this exercise follow.

Exploratory Data Analysis (EDA)

- There is substantial confounding between some potential covariates and the key predictors of interest for the catch-rates and condition of sharks, even considered at the fishery level.
- In light of this, separate analyses were in some cases conducted at the sub-fishery level (i.e. fleet level) for the individual species groups.

Preliminary Modelling

- Three response types were considered for the key questions:
  - i. A binary condition response at the individual catch level to investigate influences on shark mortality at time of retrieval.
  - ii. A CPUE response (catch-per-1000 hooks) at the set level to consider gross influences on species catch rates.
  - iii. A CPUE response (catch-per-100 hooks) at the hook level to investigate catch rates at the hook position level.

- The models applied to these response respectively were:
  - i. Logit-link, binomial errors, Generalized Linear Models (GLMs) & Generalised Additive Models (GAMs) were fitted with to investigate structural relationships between condition and the available covariates.
  - ii. Log-link, Tweedie errors, Generalized Linear Models (GLMs) & Generalised Additive Models (GAMs) fitted with to investigate structural relationships between set level catch-rates and the available covariates.
  - iii. Log-link, Tweedie errors, Generalized Linear Models (GLMs) & Generalised Additive Models (GAMs) fitted with to estimate hook-position catch rates conditioning on fishery or flag.

Gear effects on M1 FM Fishery catch rate

A general overview of the significant variables is given in the following table:

	Set-level catch-rate models				
	Fishery	M1			
	Species	OCS	FAL	THR	BSH
Temporal	Year				
	Month				
"Key" variables	Wire Trace				
	Shark Lines				
	Hook Types				
	Shark Bait				
Further set characteristics	Trip ID				
	Set ID				
	Flag				
	Set Start Time				
	Soak time				
	Hooks between floats				
	Hook position				
Oceanographic	SST				
	Height				
	Current (u)				
	Current (v)				
	Isodepth				
	Wind Stress				
	Latitute				
	Longitude				

Colour key
Not signficant
Significant and positive
Significant and negative
Significant and complex

In terms of the key gear variables, wire-trace, shark-lines, shark-bait and hook types:

- Hook-type could not be considered due to a lack of contrast.
- Catch rates of OCS were significantly higher in the presence of shark-lines.
- Catch rates of FAL were significantly higher in the presence of shark-lines & sharkbait.
- There were no significant associations between THR catch-rates and the key gear variables.
- Catch rates of BSH were significantly higher in the presence of shark-bait.

Gear effects on M1 FM Fishery mortality on retrieval

A general overview of the significant variables is given in the following table:

	Condition at retrieval models				
	Fishery	M1			
	Species	OCS	FAL	THR	BSH
Temporal	Year	Not significant	Significant and complex	Significant and positive	Not significant
	Month	Not significant	Not significant	Not significant	Not significant
"Key" variables	Wire Trace	Not significant	Significant and positive	Not significant	Significant and positive
	Shark Lines	Not significant	Significant and positive	Significant and positive	Significant and positive
	Hook Types	Not significant	Not significant	Not significant	Not significant
	Shark Bait	Not significant	Not significant	Not significant	Not significant
Further set characteristics	Trip ID	Not significant	Not significant	Not significant	Not significant
	Set ID	Not significant	Not significant	Not significant	Not significant
	Flag	Not significant	Not significant	Not significant	Not significant
	Set Start Time	Not significant	Not significant	Significant and negative	Significant and complex
	Soak time	Not significant	Not significant	Significant and positive	Significant and complex
	Hooks between floats	Not significant	Significant and positive	Not significant	Not significant
	Hook position	Not significant	Not significant	Not significant	Not significant
Oceanographic	SST	Not significant	Not significant	Significant and complex	Significant and complex
	Height	Not significant	Not significant	Not significant	Significant and complex
	Current (u)	Not significant	Significant and complex	Not significant	Significant and negative
	Current (v)	Not significant	Not significant	Significant and complex	Not significant
	Isodepth	Significant and complex	Not significant	Not significant	Not significant
	Wind Stress	Not significant	Significant and complex	Not significant	Not significant
	Latitute	Significant and complex	Significant and complex	Significant and complex	Significant and complex
	Longitude	Significant and complex	Significant and negative	Significant and complex	Significant and negative

Colour key

Not signficant (grey)
Significant and positive (red)
Significant and negative (blue)
Significant and complex (yellow)

In terms of the key gear variables, wire-trace, shark-lines, shark-bait and hook types:

- Hook-type could not be considered due to a lack of contrast.
- There were no significant effects on OCS mortality rates related to the key gear variables.
- The probability of being dead on retrieval was significantly higher for FAL under shark-lines and wire-traces.
- The probability of being dead on retrieval was significantly higher for THR under shark-lines.
- The probability of being dead on retrieval was significantly higher for BSH under shark-lines and wire-traces.

Hook position M1 FM Fishery

- The catch-rates of OCS are significantly and markedly higher on the position 0 hooks. The catch rates are estimated to be 0.128 (95% CI: [0.11, 0.15]) per 100 hooks at position 0, dropping to 0.03 (95% CI [0.23, 0.39]) at position 1. This is approximately 47% (95% CI: [40%-54%]) of the total catch over the first 12 hook positions.
- The catch-rates of FAL differed significantly with hook-position, with a clear decreasing trend with increasing hook-position. The catch rates are estimated to be 0.214 (95% CI: [0.19, 0.24]) per 100 hooks at position 0, dropping to 0.19 (95% CI [0.17, 0.22]) at position 1. This is respectively 18% (95% CI: [15.7%-20%]) and 15.7% (95% CI: [14%-18%]) of the total catch over the first 12 hook positions.
- There was a generally increasing catch rate of THR with increasing hook-position, with position 0 being the lowest. The catch rates are estimated to be 0.044 (95% CI: [0.03, 0.06]) per 100 hooks at position 0. This is approximately 3.5% (95% CI: [2.7%-4.5%]) of the total catch over the first 12 hook positions.
- There was no monotonic trend of BSH catch rates with respect to hook position, with positions 2-4 being highest.
- The catch rates of BSH are estimated to be 0.099 (95% CI: [0.08, 0.11]) per 100 hooks at position 0, rising to 0.12 (95% CI [0.11, 0.14]) at position 1. This is respectively 7.3% (95% CI: [6.3%-8.5%]) and 9% (95% CI: [7.8%-10.3%]) of the total catch over the first 12 hook positions.

DMP STATS The Coach House, Mount Melville House, St. Andrews, Scotland KY 168 NT, UK +44(0) 1334 477 544 mail@dmpstats.com

Gear effects on M2 FV Fiji Fishery catch rate

A general overview of the significant variables is given in the following table:

	Set-level catch-rate models				
	Fishery		M1		
	Species	OCS	FAL	THR	BSH
Temporal	Year				
	Month				
"Key" variables	Wire Trace				
	Shark Lines				
	Hook Types				
	Shark Bait				
Further set characteristics	Trip ID				
	Set ID				
	Flag				
	Set Start Time				
	Soak time				
	Hooks between floats				
	Hook position				
Oceanographic	SST				
	Height				
	Current (u)				
	Current (v)				
	Isodepth				
	Wind Stress				
	Latitute				
	Longitude				

Colour key

Not signficant

Significant and positive

Significant and negative

Significant and complex

In terms of the key gear variables, wire-trace, shark-lines, shark-bait and hook types:

- Catch rates of OCS were significantly higher in the presence of wire-trace and shark-lines.
- Catch rates of OCS were significantly higher on type-C hooks compared to type-J hooks.
- Catch rates of THR were significantly higher on type-J hooks compared to type-C hooks.
- Catch rates of BSH were significantly higher in the presence of shark-lines.

Gear effects on M2 FV Fiji Fishery mortality on retrieval

A general overview of the significant variables is given in the following table:

	Condition at retrieval models				
	Fishery	M1			
	Species	OCS	FAL	THR	BSH
Temporal	Year	Not significant	Not significant		Not significant
	Month	Not significant	Not significant		Not significant
"Key" variables	Wire Trace	Not significant	Not significant		Significant and positive
	Shark Lines	Not significant	Not significant		Not significant
	Hook Types	Significant and complex	Not significant		Significant and complex
	Shark Bait	Not significant	Not significant		Not significant
Further set characteristics	Trip ID				
	Set ID				
	Flag				
	Set Start Time	Not significant	Not significant		Significant and negative
	Soak time	Not significant	Not significant		Not significant
	Hooks between floats		Not significant		Not significant
	Hook position	Not significant	Significant and positive		Significant and positive
Oceanographic	SST	Not significant	Not significant		Significant and negative
	Height	Not significant	Not significant		Not significant
	Current (u)	Not significant	Not significant		Not significant
	Current (v)	Not significant	Not significant		Not significant
	Isodepth	Not significant	Not significant		Not significant
	Wind Stress	Not significant	Not significant		Not significant
	Latitute	Not significant	Not significant		Not significant
	Longitude	Not significant	Not significant		Not significant

Colour key

Not signficant (grey)
Significant and positive (red)
Significant and negative (blue)
Significant and complex (yellow)

In terms of the key gear variables, wire-trace, shark-lines, shark-bait and hook types:

- The probability of being dead on retrieval was significantly higher for OCS and BSH with type-J hooks, compared to type-C hooks.
- The probability of being dead on retrieval was significantly higher for BSH under shark-lines.

Hook position M2 FV Fiji Fishery

- There is a markedly higher catch rate for OCS at hook-position 0 compared to other positions. The catch rate at position 0 is 0.091 per 100 hooks [0.066 to 0.127, 95% CI]. This translates to approximately 63.2% [45.6% to 87.8%, 95% CI] of the catch of this species observed in this fisheries data. Beyond position 0, there is little clear pattern other than catches being generally near zero for positions 7+.
- The catch-rates of FAL differed significantly with hook-position, with a clear decreasing trend with increasing hook-position. The catch rates are estimated to be 0.067 (95% CI: [0.05, 0.09]) per 100 hooks at position 0, dropping to 0.016 (95% CI [0.009, 0.03]) at position 1. This is respectively 47.6% (95% CI: [35.6%-63.2%]) and 11.5% (95% CI: [6.5%-20.3%]) of the total catch over the first 19 hook positions.
- There was no particular trend catch rate of THR with increasing hook-position, with position 0 being relatively low. The catch rates are estimated to be 0.004 (95% CI: [0.001, 0.012]) per 100 hooks at position 0. This is approximately 5.3% (95% CI: [1.5%-18.8%]) of the total catch over the first 19 hook positions.
- There was a decreasing trend of BSH catch rates with respect to hook position. The catch rates of BSH are estimated to be 0.13 (95% CI: [0.1, 0.17]) per 100 hooks at position 0, falling to 0.08 (95% CI [0.06, 0.11]) at position 1. This is respectively 22.4% (95% CI: [17.3%-28.8%]) and 14.4% (95% CI: [10.5%-19.8%]) of the total catch over the first 19 hook positions.

Gear effects on M2 FV US Fishery catch rate

A general overview of the significant variables is given in the following table:

	Set-level catch-rate models				
	Fishery	M1			
	Species	OCS	FAL	THR	BSH
Temporal	Year				
	Month				
"Key" variables	Wire Trace				
	Shark Lines				
	Hook Types				
	Shark Bait				
Further set characteristics	Trip ID				
	Set ID				
	Flag				
	Set Start Time				
	Soak time				
	Hooks between floats				
	Hook position				
Oceanographic	SST				
	Height				
	Current (u)				
	Current (v)				
	Isodepth				
	Wind Stress				
	Latitute				
	Longitude				

Colour key

Not signficant

Significant and positive

Significant and negative

Significant and complex

In terms of the key gear variables, wire-trace, shark-lines, shark-bait and hook types:

- Only hook-types could be contrasted.
- Catch rates of OCS and THR were significantly higher on hook type-C compared to type-T.

DMP STATS The Coach House, Mount Melville House, St. Andrews, Scotland KY 168 NT, UK +44(0) 1334 477 544 mail@dmpstats.com

Gear effects on M2 FV US Fishery mortality on retrieval

A general overview of the significant variables is given in the following table:

	Condition at retrieval models				
	Fishery	M1			
	Species	OCS	FAL	THR	BSH
Temporal	Year	Not significant	Significant and negative	Not significant	Significant and negative
	Month	Not significant	Not significant	Not significant	Not significant
"Key" variables	Wire Trace				
	Shark Lines				
	Hook Types	Not significant	Not significant	Not significant	Not significant
	Shark Bait				
Further set characteristics	Trip ID				
	Set ID				
	Flag				
	Set Start Time	Not significant	Significant and complex	Not significant	Significant and complex
	Soak time	Significant and complex	Significant and positive	Not significant	Significant and complex
	Hooks between floats	Significant and complex	Not significant	Not significant	Significant and complex
	Hook position	Significant and complex	Significant and complex		Significant and complex
Oceanographic	SST	Not significant	Significant and complex	Not significant	Significant and complex
	Height	Not significant	Significant and complex	Not significant	Significant and complex
	Current (u)	Significant and positive	Significant and complex	Not significant	Significant and complex
	Current (v)	Significant and negative	Significant and complex	Not significant	Significant and complex
	Isodepth	Significant and complex	Significant and complex	Not significant	Significant and complex
	Wind Stress	Significant and negative	Significant and complex	Not significant	Significant and complex
	Latitute	Not significant	Significant and complex	Not significant	Significant and complex
	Longitude	Not significant	Significant and complex	Not significant	Significant and complex

Colour key
Not signficant
Significant and positive
Significant and negative
Significant and complex

In terms of the key gear variables, wire-trace, shark-lines, shark-bait and hook types:

- Only hook-type could be contrasted and there were no significant differences in terms of condition at retrieval for any species considered.

Hook position M2 FV US Fishery

- There is no particular relationship between the hook-positions and the catch rates for OCS. The shallowest position (1) is estimated at 0.03 per 100 hooks [0.018 to 0.049, 95% CI]. This translates to approximately 6.9% [4.21% to 11.4%, 95% CI] of the catch of this species observed in this fisheries data.
- The catch-rates of FAL tended to be flat with respect to increasing hook position, with some marked fluctuations at the deeper positions (15 to 17). The shallowest position (1) is estimated at 0.06 per 100 hooks [0.044 to 0.085, 95% CI]. This translates to approximately 4.1% [2.95% to 5.7%, 95% CI] of the catch of this species observed in this fisheries data.
- The catch-rates of THR tended towards mild increases respect to increasing hook position, with some marked fluctuations at the deeper positions (13 to 17). The shallowest position (1) is estimated at 0.003 per 100 hooks [0.001 to 0.012, 95% CI]. This translates to approximately 1.4% [0.4% to 4.9%, 95% CI] of the catch of this species observed in this fisheries data.
- The catch-rates of BSH tended towards mild increases respect to increasing hook position, with some marked fluctuations at the deeper positions (13 to 17). The shallowest position (1) is estimated at 0.07 per 100 hooks [0.054 to 0.1, 95% CI]. This translates to approximately 2.3% [1.7% to 3.1%, 95% CI] of the catch of this species observed in this fisheries data.

Gear effects on M7 HD Fishery catch rate

A general overview of the significant variables is given in the following table:

	Set-level catch-rate models				
	Fishery	M1			
	Species	OCS	FAL	THR	BSH
Temporal	Year				
	Month				
"Key" variables	Wire Trace				
	Shark Lines				
	Hook Types				
	Shark Bait				
Further set characteristics	Trip ID				
	Set ID				
	Flag				
	Set Start Time				
	Soak time				
	Hooks between floats				
	Hook position				
Oceanographic	SST				
	Height				
	Current (u)				
	Current (v)				
	Isodepth				
	Wind Stress				
	Latitute				
	Longitude				

Colour key

Not signficant

Significant and positive

Significant and negative

Significant and complex

In terms of the key gear variables, wire-trace, shark-lines, shark-bait and hook types:

- Catch rates of OCS were significantly higher on type-C hooks compared to type-T hooks (with type-J being intermediate and not distinct from T in particular).
- Catch rates of BSH were significantly higher on type-J hooks compared to type-T & type-C hooks.
- Catch rates of BSH were significantly higher in the presence of wire-trace.

Gear effects on M7 HD Fishery mortality on retrieval

A general overview of the significant variables is given in the following table:

	Condition at retrieval models				
	Fishery	M1			
	Species	OCS	FAL	THR	BSH
Temporal	Year	Significant and complex	Significant and complex	Significant and complex	Significant and complex
	Month	Significant and complex	Significant and complex	Significant and complex	Significant and complex
"Key" variables	Wire Trace	Not signficant	Not signficant	Not signficant	Significant and negative
	Shark Lines				
	Hook Types	Not signficant	Not signficant	Not signficant	Not signficant
	Shark Bait				
Further set characteristics	Trip ID				
	Set ID				
	Flag				
	Set Start Time	Not signficant	Not signficant	Not signficant	Not signficant
	Soak time	Significant and complex	Significant and positive	Significant and complex	Significant and positive
	Hooks between floats	Significant and complex	Not signficant	Significant and complex	Significant and complex
	Hook position	Significant and positive	Significant and positive	Not signficant	Not signficant
Oceanographic	SST	Not signficant	Not signficant	Significant and positive	Significant and complex
	Height	Not signficant	Significant and negative	Significant and complex	Not signficant
	Current (u)	Significant and positive	Not signficant	Not signficant	Significant and negative
	Current (v)	Not signficant	Not signficant	Significant and complex	Significant and complex
	Isodepth	Not signficant	Not signficant	Significant and positive	Significant and complex
	Wind Stress	Not signficant	Significant and complex	Significant and complex	Not signficant
	Latitute	Not signficant	Significant and complex	Significant and positive	Significant and complex
	Longitude	Not signficant	Significant and negative	Significant and complex	Significant and complex

Colour key
Not signficant
Significant and positive
Significant and negative
Significant and complex

In terms of the key gear variables, wire-trace, shark-lines, shark-bait and hook types:

- The probability of being dead on retrieval was significantly lower for BSH with wire-traces.

Hook position M7 HD Fishery

- There is no particular relationship between the hook-positions and the catch rates of OCS. The shallowest position (1) is estimated at 0.03 per 100 hooks [0.018 to 0.049, 95% CI]. This translates to approximately 6.9% [4.21% to 11.4%, 95% CI] of the catch of this species observed in this fisheries data.
- The catch-rates of FAL differed significantly with hook-position, tending to be high at both low and high hook-positions. The catch rates are estimated to be 0.009 (95% CI: [0.008, 0.01]) per 100 hooks at position 1, dropping to 0.008 (95% CI [0.007, 0.009]) at position 2. This is respectively 10.7% (95% CI: [9.7%-12%]) and 9.9% (95% CI: [8.8%-11%]) of the total catch over the first 15 hook positions. There was evidence of correlations in the errors unaccounted for in these models, likely giving confidence intervals that are spuriously narrow.
- There was a clear increasing catch rate of THR with increasing hook-position, with position 1 being the lowest. The catch rates are estimated to be 0.005 (95% CI: [0.004, 0.006]) per 100 hooks at position 1. This is approximately 1.2% (95% CI: [1%-1.3%]) of the total catch over the first 15 hook positions. There was evidence of correlations in the errors unaccounted for in these models, likely giving confidence intervals that are spuriously narrow.
- There was a clear increasing catch rate of BSH with increasing hook-position, with position 1 being the lowest. The catch rates are estimated to be 0.16 (95% CI: [0.156, 0.164]) per 100 hooks at position 1. This is approximately 4.3% (95% CI: [4.2%-4.5%]) of the total catch over the first 15 hook positions. There was evidence of correlations in the errors unaccounted for in these models, likely giving confidence intervals that are spuriously narrow.

2 PROJECT OVERVIEW

Section 2.1 provides a description of the overarching project that the current report is a component of. There are two main phases to the project, with this study addressing the first phase only.

2.1 OVERARCHING SHARK CATCH, CONDITION AND MITIGATION PROJECT

The following paraphrases the project description in the associated Terms of Reference (ToR).

In response to the stock assessments of oceanic white-tip and silky sharks, the WCPFC-SC tasked SPC with examining potential mitigation options for these two shark species which are taken in tropical, and sub-tropical longline fisheries. These analyses are documented in Bromhead et al. (2013) and OFP (2012). These analyses were hindered by the paucity and unbalanced nature of the observer data. The analyses did identify that there was some deliberate shark targeting using shallow lines attached to the floats, appropriately called 'shark lines', but conceded that the data made it difficult to make many other conclusions regarding wire leaders or hook types.

The purpose of this consultancy is to further develop the analyses undertaken by Bromhead et al. (2013) in five main ways:

1. Inclusion of observer data from longline fishery in American Samoa;
2. Expansion of the species considered to include blues, makos, thresher, and hammerhead sharks;
3. Consideration of other potential influencing factors on catch rates (catchability), e.g., hook depth and soak time;
4. Analysis of these factors and other factors on the reported condition of sharks at the time of retrieval;
5. Monte Carlo simulation of some 'what-if' scenarios on predicted impacts of potential mitigation measures that integrate estimates of catchability and condition with post-release survival estimates from the literature.

2.2 SCOPE OF PRESENT STUDY

The study reported up here is a subset of the project described in Section 2.1. Specifically this study addresses the following goals:

i. Estimate the effect of various fishing gear specifications on catch rates for the shark species groups based on the observer data sets described in Table 1. Gear characteristics of particular interest include leader material, hook type, hooks between floats, soak time, and use of shark lines and bait. The analysis should also consider other variable that can impact catch rates such as location, season, oceanographic variables, and flag/vessel effects.

Table 1: Species and fisheries under consideration for this study.

Species/Species groups	Regional Fisheries
Oceanic whitetip	Fiji and American Samoa
Silky shark	Marshall Islands and Federated States of Micronesia
Blue shark	Hawaii (deepset fishery only)
Hammerheads	
Thresher sharks	
Mako sharks	

ii. Estimate the effect of various fishing gear specifications on condition at time of retrieval for the same species and same observer data sets described above. Data will grouped into alive and dead only (i.e., the full suite of life-status categories used in the observer data base will not be considered). Variables similar to those considered under (i) above should be considered.

iii. Estimate the potential effect of eliminating shallow hooks from branchlines on for the same species and same observer data sets described above. This analysis could be integrated into that undertaken under (i) above or undertaken separately.

iv. **A future phase of the project will address the following goal** - Monte Carlo simulations of some 'what-if' scenarios of potential mitigation measures by integrating estimates of catchability and condition from (i – iii) above, and post-release survival estimates from the literature.

3 SUMMARY OF DATA AND PRE-TREATMENT

Observer and oceanographic data were provided by SPC in of three separated datasets:

- Effort and fishing gear data at a set level, comprising key factors such as start time, soak time, hooks between floats, hook type, shark bait use, wire trace use and shark line use.
- Shark catch data at a hook level (i.e. hook position between floats in a set), describing the condition (dead or alive) of caught sharks.
- Oceanographic estimates data, comprising relevant environmental variables such as sea surface temperature and height, wind stress, current speed and isotherm depths (mainly sourced from GODAS[3] and ASCAT[4]).

Data pre-treatment involved primarily checking the provided data for inconsistencies and errors before combining the datasets together for the exploratory data analysis. Data was subsequently processed and reshaped accordingly for each type of response under analysis. For example, data for the analysis of catch rates of a given shark species by a fishery/fleet was obtained by summing over the number caught in each of the sets observed in that fishery/fleet, allocating zeros to sets with no catches of the species under consideration.

[3] NCEP Global Ocean Data Assimilation System

[4] Advanced Scatterometer on board the METOP-A European Earth observation satellite

4 MODELLING OVERVIEW

Model selection in terms of the systematic components, was conducted in a number of ways:

1. In the first instance, cases of highly related covariates were identified using naïve Generalized Linear Models (naïve meaning without interaction terms and assuming independence of errors). These were used as the basis for calculating Generalized Variance Inflation Factors (GVIFs). Particularly large GVIFs indicated variables that might be collinear, leading to inflated variances around parameter estimates. Model selection methods may resolve this in an automated manner, but from an interpretative point of view, a qualitative selection between such variables is preferred.
2. Subject to this *a priori* reduction in potential covariates, a speculative "saturated" model was specified, including interaction terms (up to level four in some instances). The interactions were restricted to variables thought to logically have potential interactions, rather than consider all possible interaction terms. These saturated models are presented in their respective sections, but typically consisted of greater than 20 terms, consisting of quantitative covariates, factor covariates and various interactions therein.
3. Examination of preliminary models indicated non-linearities ought to be considered in all models. Hence, Generalized Additive Models (Hastie & Tibshirani, 1986, 1990) were the favoured modelling framework - specifically the implementations offered in mgcv (Wood, 2000; 2003; 2004; 2006 & 2011) which offer tools for optimising complexity and model selection. Model selection for GAMs was via the *Null space penalization*, whereby all terms are retained in their model, but their relevance in the model is reduced by penalties accordingly (Marra & Wood, 2011).

The models fitted were of two main types:

1. For catch rate responses (set and hook-level), Generalized Additive Models (GAMs) were used. To account for the prevalence of zeros in the data, Tweedie distributions were used, which are a generalisation of several exponential family distributions. In particular they permit a compound Poisson-Gamma distribution. The continuous, skewed, catch rate data observed here would be suited by a Gamma distribution, but for a mass at zero, which is not permitted under a Gamma. Tweedie parameters were estimated in each case to parameterise this compound Poisson-Gamma error distribution. The default log-link function was adopted. Smooth terms were specified where appropriate (both univariate and bivariate) and were penalised regression splines – the smoothing parameters being optimised via multiple Generalized Cross-Validation.
2. For condition models, the response was alive/dead leading to logit-link binomial errors GAMs as befits binary response data. Smooth terms were specified where appropriate (both univariate and bivariate) and were penalised regression splines – the smoothing parameters being optimised via multiple Generalized Cross-Validation.

The following primary assumptions were considered in the modelling process:

1. Non-linearities: where considered *a priori* likely, these were accommodated via smoother terms specified in the GAMs. The combination of optimised smoothing parameters and term selection via Null space penalisation is to account for complexities in the systematic component.
2. Zero-rich data and general adequacy of assumed error distributions: a compound Poisson-Gamma distribution was utilised in the GAMs via Tweedie distributions, which can accommodate levels of mass at zero. The Tweedie parameter was estimated as part of the process and the resulting error distributions checked for adequacy via diagnostics on quantile residuals. Quantile residuals were also used to test the adequacy of the assumed error distribution in the binary response GAMs.
3. Independence of errors: *a priori* the likelihood of correlated errors seems high given the nature of sets. This was checked via quantile residuals with runs tests and autocorrelation functions, where the sets were ordered in a logical fashion. In the vast majority of cases the correlation in the errors was ignorable once proper account had been taken of the preponderance of zero catch observations.

Table 2: Explanatory variables and corresponding abbreviations used in model structures throughout this preliminary report.

Variable	Abbreviation in modelling
Year (2008-2103)	yy
Month within year (1-12)	mm
Hook type (types C, J, T)	hook_type
Wire trace (yes/no)	wire_trace
Shark lines (yes/no)	nbshark_lines
Shark bait (yes/no)	sharkbait
Latitude (decimal)	lat
Longitude (decimal)	lon
Start time for set	set_start_time
Sea Surface Temperature (SST °C)	sst
Sea Surface Height (m)	height
Current speed (v and u, m/s)	vcur, ucur
Depth of 20^{o}C degree isotherm	isodepth
Wind stress (Newton/m^{2})	windStress
Number of hooks between floats	hk_bt_flt
Number of hooks used on the set	hook_set
Position of the hook to nearest float	hook_pos
Soak time (hours)	soak

5 MARSHALL ISLANDS AND FEDERATED STATES OF MICRONESIA FISHERY (M1_FM)

5.1 FISHERY DATA DESCRIPTION

Observer data from the Marshal Islands and Federated States of Micronesia longline fishery comprises trips carried out by the Chinese and Micronesian fleets while operating in the EEZs of Marshal Islands (MH) and Micronesia (FM) (Figure 1), with a larger number of observed sets occurring in the MH's EEZ. The data time-series for this fishery extends from 1993 to 2012, with the majority of available observed trips of both fleets occurring in the period 2004-2008. There are no considerable changes in the number of observed trips over months and, with the exception of the Chinese fleet in 2004, the number of sets per trip appears to have increased from 2000 onwards (Figure 1). While the bulk of the observed trips were taken in the Chinese fleet, there are no substantial differences in terms of temporal and spatial distributions of trips and sets between the two fleets (Figure 1).

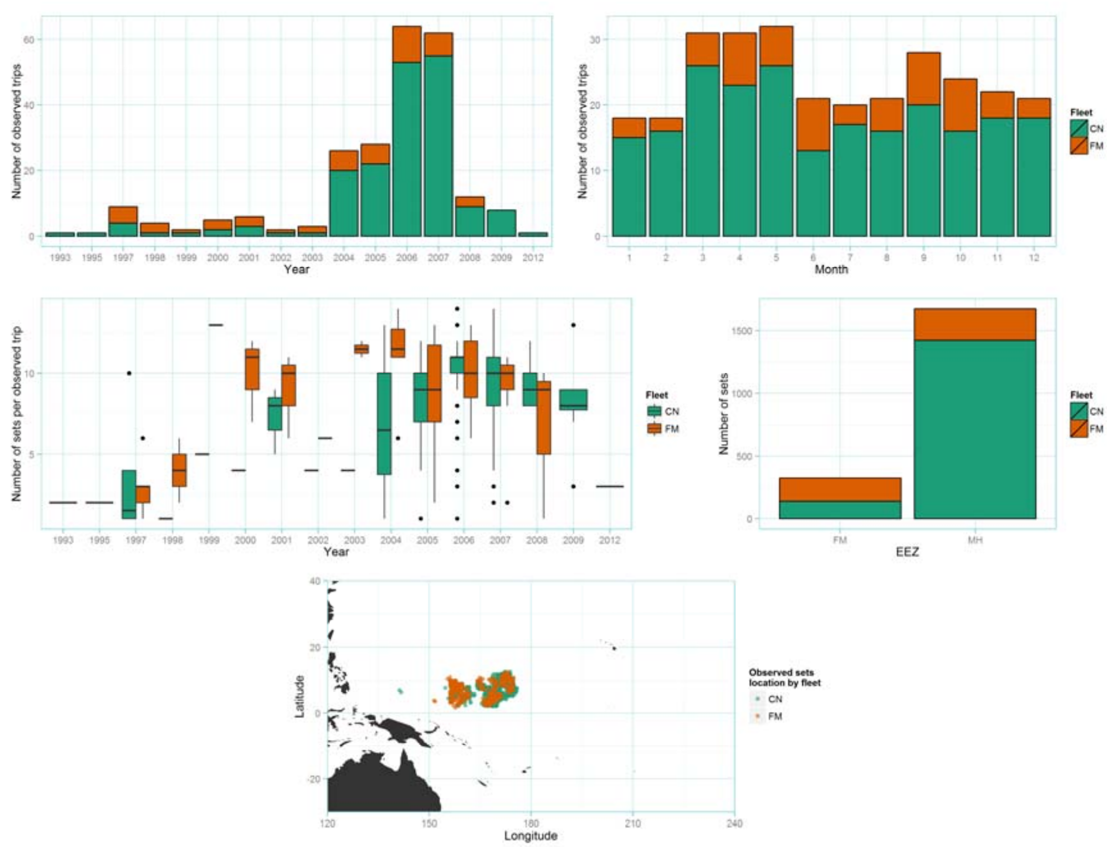


Figure 1: Temporal and spatial distribution of fishing effort by fleet (CN = Chinese, FM = Micronesian) in observed trips of the Marshall Islands and Federated States of Micronesia fishery (file "M1_FM_tripsAndSetsDesc.png").

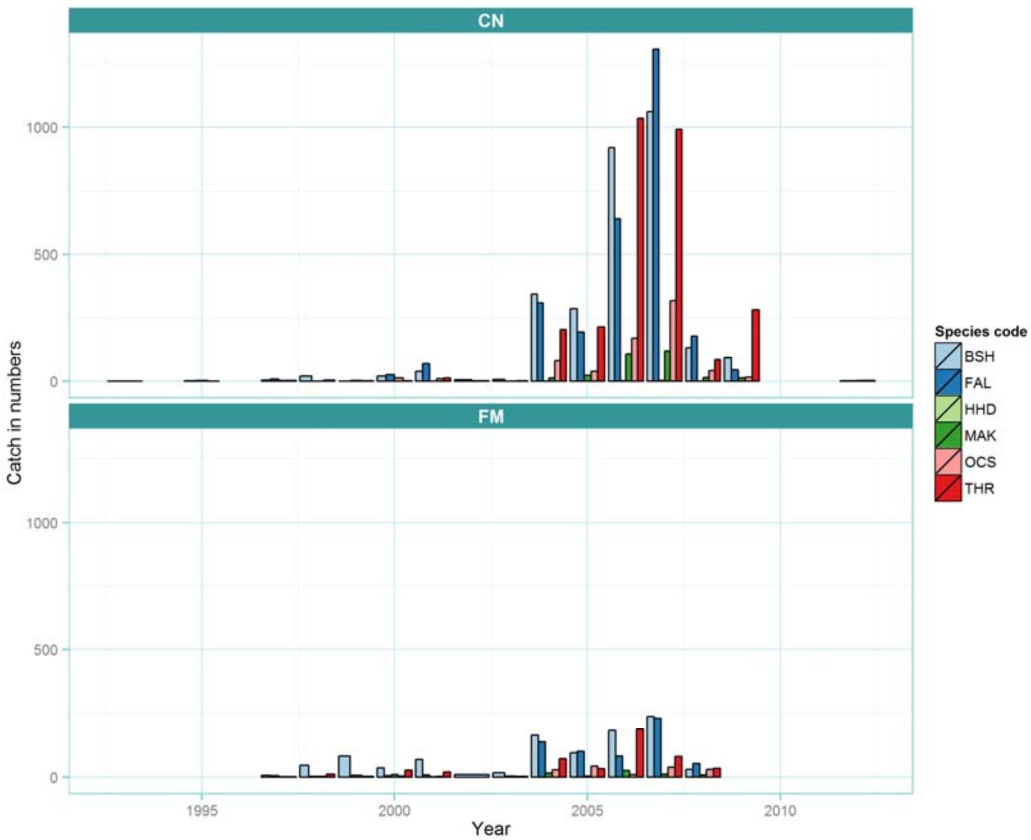


Figure 2: Bycatch of shark species over time in observed trips for fleets operating in the Marshall Islands and Federated States of Micronesia longline fishery (file "M1_FM_catchNumSpYrFlg.png").

Furthermore, blue sharks, silky sharks and thresher sharks are the species with higher bycatch numbers across years in trips from both fleets (Figure 2). For example, the Chinese fleet caught over 1250 silky sharks and over 1000 blue sharks in trips observed during 2007, while catching around 300 oceanic whitetip sharks in trips from the same year. On the other hand, catches of Hammerhead sharks are virtually absent in this fishery for the whole time-series duration.

Figure 3 shows that the type of hook used in the M1_FM fishery is unknown until 2009, and from that year on there are very few sets with positively identified hook types. The absence of information about hook type lead to the exclusion of this key factor from the analysis. To avoid confounding effects due to the lack of contrast between categories of wire leader, shark bait and shark line use in earlier and later periods of the time-series, the analysis of the M1_FM fishery was constrained to the period 2004-2007. Further to the temporal and spatial similarities in observed trips from the two fleets, the identical contrasts in key fishing factors (Figure 3) meant that the two fleets could be used in the same statistical models.

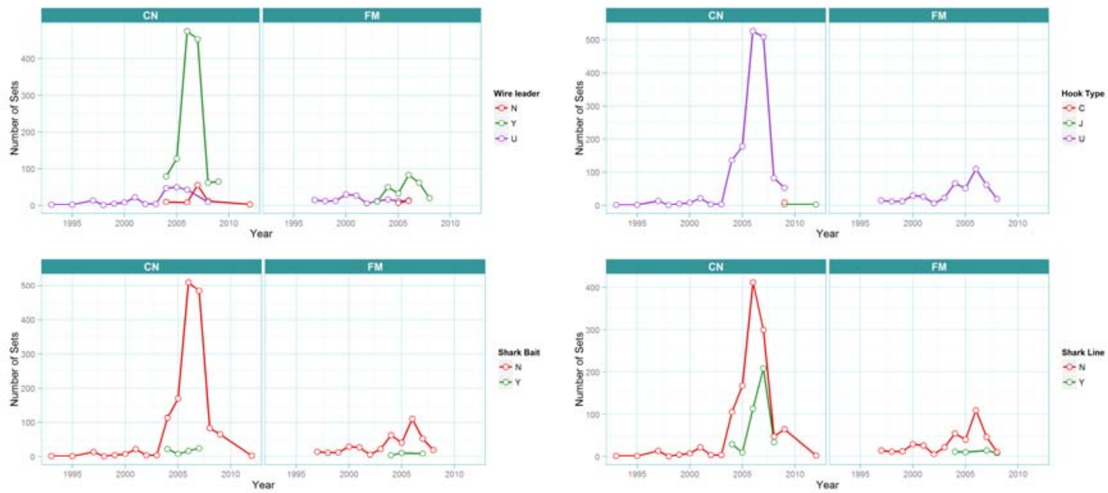


Figure 3: Contrast of key gear types over time in observed trips by fleets operating in the Marshall Islands and Federated States of Micronesia fishery (File "M1_FM_SampSizeAndKeyFactors.png").

5.2 EXPLANATORY VARIABLES EDA

There were a few observations with low values of hook between floats, but they were considered to be within the possible range of values of that covariate (Figure 4). There were no other extreme values in the remaining covariates.

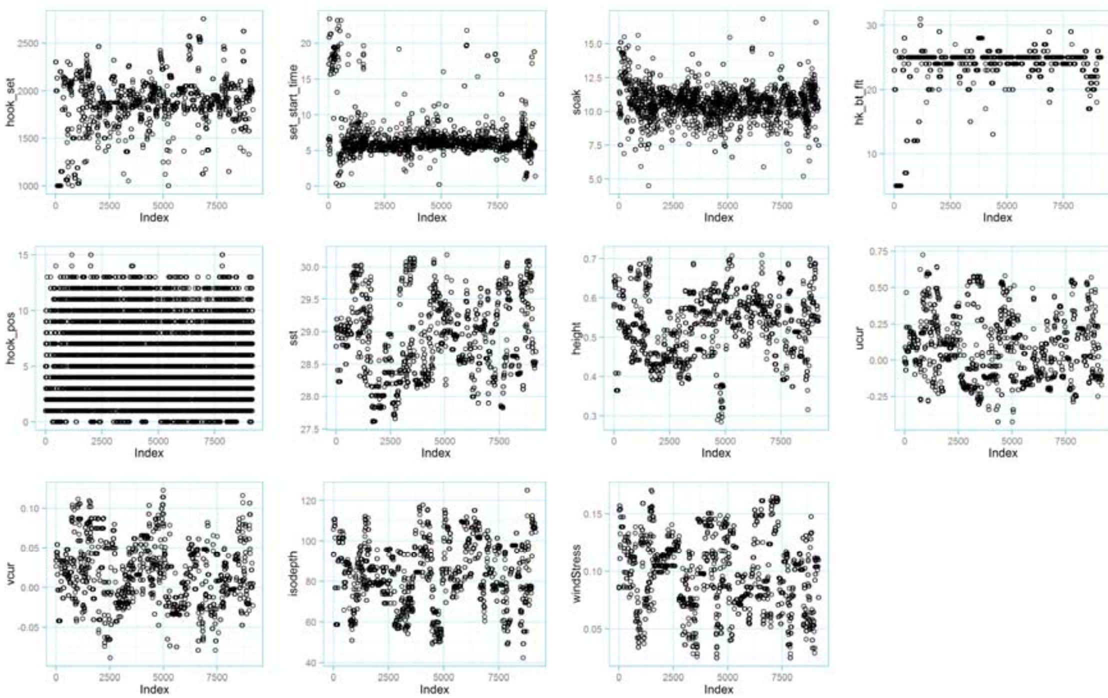


Figure 4: Cleveland dotplots of some of the explanatory variables considered for the analysis of shark bycatch in the Marshall Islands and Federated States of Micronesia longline fishery. Plots constructed from data sorted by set starting date (File “M1_FM_covsClevPlots.png”)

Figure 5 shows the inexistence of considerably strong correlations between the explanatory variables from observations used for the M1_FM analysis. Furthermore, several of the oceanographic variables (e.g. sea surface temperature, sea surface height and wind stress) appear to have non-linear relationships with month and latitude (Figure 5). Changes in wind stress levels also seem to be accompanied by changes in horizontal current speed and sea surface temperature and height.

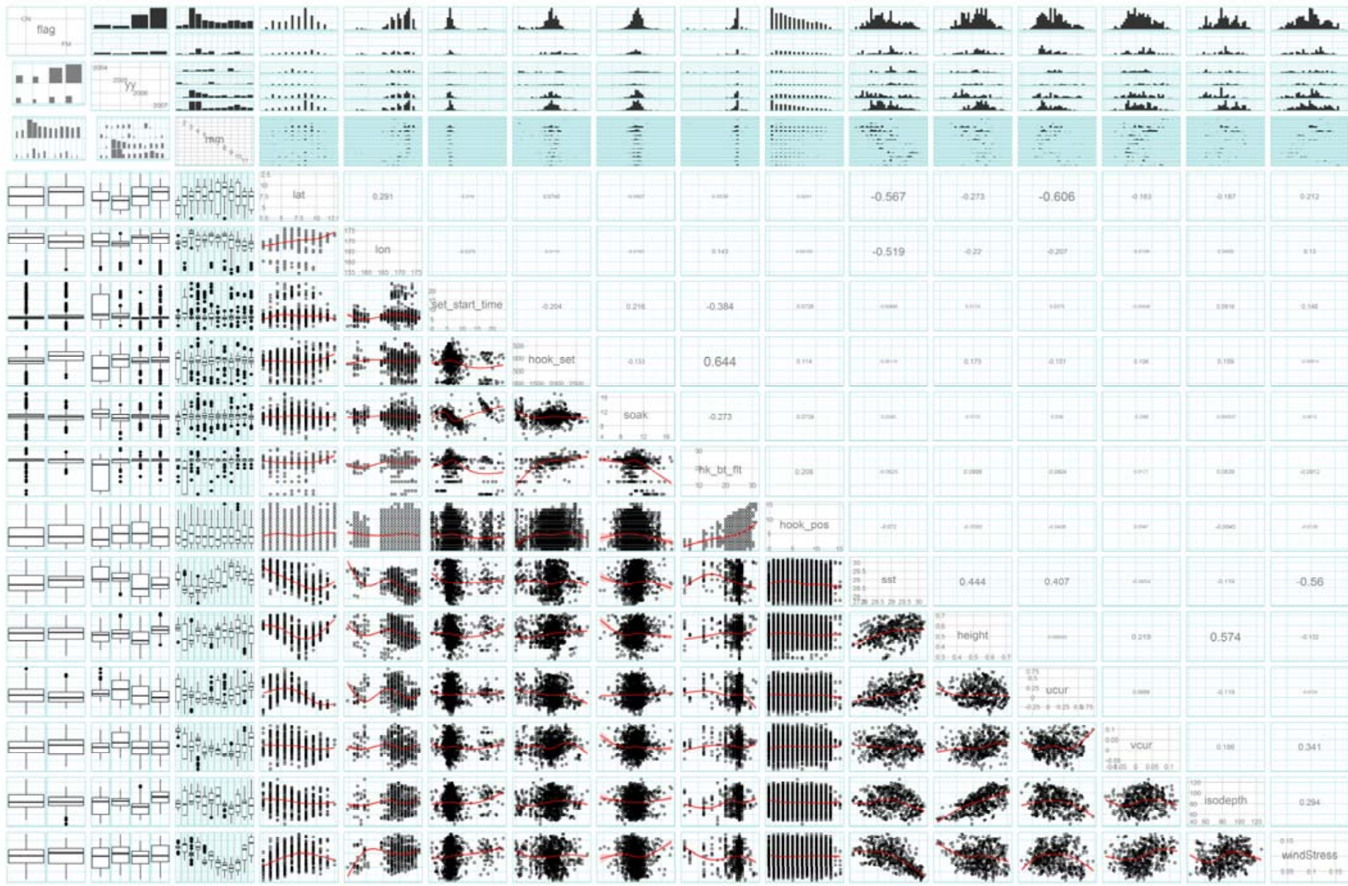


Figure 5: Matrixplot of relevant covariates considered for the analysis of shark bycatch in the Marshall Islands and Federated States of Micronesia longline fishery. Plots presented in panels depend on the data types under comparison. Pairs of continuous variables are displayed as scatterplots in the lower panels (with an added LOESS smoother to help visualisation) and Pearson correlation coefficients in the upper panels (with font size proportional to correlation coefficient). Continuous Vs Discrete covariates are represented by boxplots and barplots, while pairs of discrete covariates are displayed as fluctuation plots (lower panels) and barplots (upper panels) (file “M1_FM_CovsPairsPlot.png”)

5.3 CN & FM COMBINED FLEET ANALYSIS - MODELLING

Presented here are results for three models covering set-level catch-rates, catch-rates with respect to hook position and the condition of sharks (alive/dead) at retrieval. Four shark species are considered: Oceanic whitetips (OCS), silky (FAL), thresher (THR) and blue (BSH).

Details regarding modelling methods are presented previously, but all models fall within the Generalized Additive Modelling framework. The set of parametric and smoothing terms selected between are given in the following table – the specific model fitted for each species and model is presented in the relevant results section.

Table 3: Model terms considered. "s" indicates a smoothing term. ":" indicates an interaction between associated terms.

Parametric terms	Smooth terms
yy	s(set_start_time)
mm	s(sst)
hook_type	s(height)
wire_trace	s(ucur)
nbshark_lines	s(vcur)
hook_type:wire_trace	s(isodepth)
hook_type:nbshark_lines	s(windStress)
wire_trace:nbshark_lines	s(lat,lon)
hook_type:wire_trace:nbshark_lines	s(soak,hk_bt_flt)
	s(soak):nbshark_lines
	s(soak):sharkbait

Reference model – In the interests of brevity, one species' models are presented in more detail than the others. This will be referred to as the *reference case* and will be the Oceanic whitetip shark (OCS). In particular model diagnostics and non-significant results will be presented for this model, but not others. The same modelling approach has been applied for similar models, but only significant results (including assumption violations) will be reported upon.

5.4 CN & FM COMBINED FLEET ANALYSIS - OCEANIC WHITE SHARKS (OCS) [REFERENCE CASE]

5.4.1 Catch rate of Oceanic Whitetip Sharks (OCS) [reference case]

5.4.1.1 Data Exploration

While predominantly absent in sets from observed trips in the M1_FM fishery, within-set bycatch levels of oceanic whitetip sharks ranged largely between 1 shark/set (in about 200 sets) and 3 sharks/set (in about 20 sets) (Figure 6). However, in a few instances, up to 9 and 11 oceanic whitetips were observed in a single set.

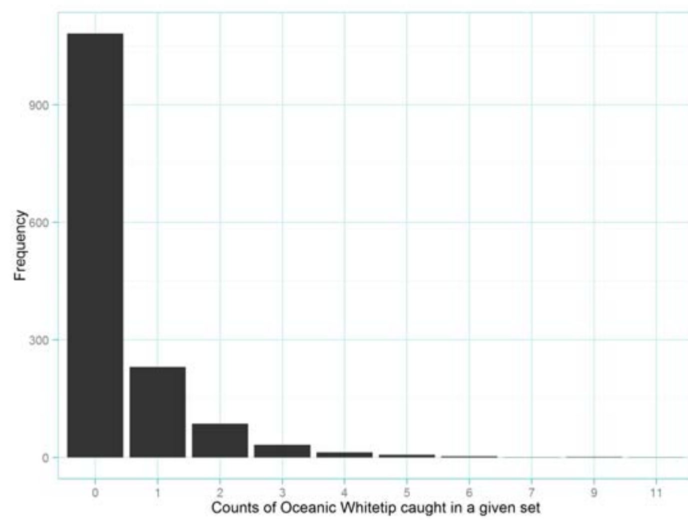


Figure 6: Distribution of counts per set of oceanic whitetip shark caught in observed trips of the M1_FM longline fishery (M1_FM-OCS_CatchNumFreqs.png)

The predominance of zero catches in observed sets hampers the investigation of bivariate patterns between catch rates of oceanic whitetip sharks and each of the covariates considered in the present analysis (Figure 7). Yet, these plots indicate that:

- The presence of wire leader, shark bait and shark lines may lead to higher catch rates of this species;
- Catch rates are higher in the months of May, June and July;

- Higher catch rates appear occur within 10-11 hours of soak time and when sets started to be around 5-6am.

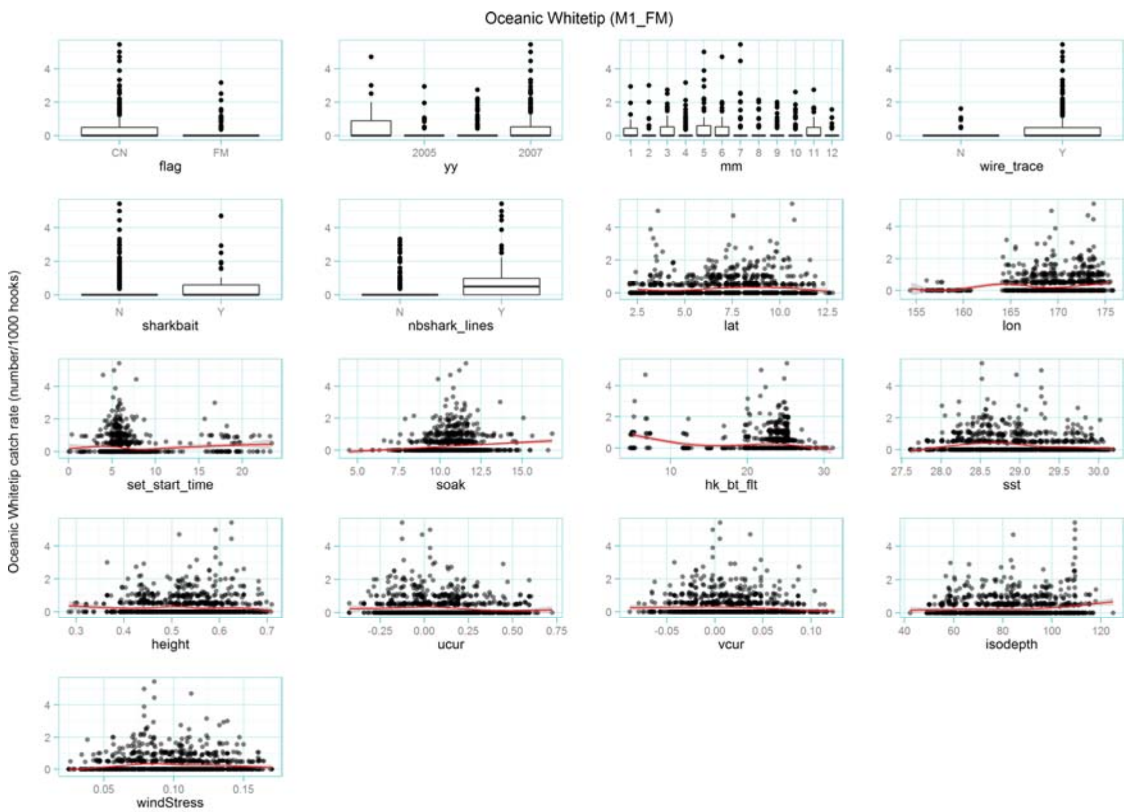


Figure 7: Distributions of set-level catch-rate responses against the explanatory variables considered for the analysis of oceanic whitetip sharks caught in the M1_FM fishery (file “M1_FM-OCS_CatchRatesVsCovs.png”).

5.4.1.2 Modelling results

Catch rate (per 1000 hooks) was modelled as a function of set-level characteristics and associated oceanographic variables. The initial set of potential covariates was thinned on the basis of gross collinearity, as determined by Generalized Variance Inflation Factors (GVIFs). A range of potential interactions and non-linear terms were specified *a priori* as per Table 3. These were further thinned during the modelling process in some cases where estimation was not possible e.g. interactions for factor variables where certain combinations of factors were not observed in the dataset.

The model fitted was a GAM with log-link and Tweedie-distributed errors. An initial model run was conducted for the purposes of estimating the governing Tweedie parameter – subsequently refitted with the estimated Tweedie parameter specified. Model "selection" was conducted within the GAM using shrinkage, so model terms of little relevance have their effective degrees of freedom shrunk towards zero. In terms of parametric components, backwards selection (p-value<0.05) was used, but not on the key factors of wire-trace, shark-bait, shark-lines and hook type (where applicable). These were generally retained in the model for interpretative interest. The results of this process are summarised in Table 4.

The Tweedie parameter was estimated to be 1.016, with the model returning an adjusted-R^2 of 0.342 and a deviance explained of 39.8%.

Table 4: Model terms from the generalized additive model. DF = Degrees of Freedom. EDF=Effective Degrees of Freedom. Shrinkage has been applied as a means of model selection. Terms of little relevance have their EDF shrunken.

Model terms	EDF	DF	Chi or F-stat	*p*-value
yy		3	3.81	0.010
mm		11	3.23	0.000
wire_trace		1	0.08	0.777
nbshark_lines		1	84.83	0.000
sharkbait		1	0.06	0.799
flag		1	0.00	0.956
s(set_start_time)	7.36	9	1.65	0.026
s(sst)	5.22	9	2.20	0.000
s(height)	5.66	9	1.09	0.065
s(ucur)	0.69	9	0.12	0.183
s(vcur)	0.43	9	0.05	0.279
s(isodepth)	4.03	9	3.54	0.000
s(windStress)	3.88	9	1.98	0.000
s(lon,lat)	17.18	29	2.66	0.000
s(soak,hk_bt_flt)	9.26	29	1.18	0.000

The model assumptions were assessed. Non-linearities are assumed to be captured via the GAM which optimises complexity. The adequacy of the assumed Tweedie distribution was assessed via Quantile Residuals (Dunn & Smyth, 1996), as indicated in Figure 8 – these ought to be approximately Normally distributed, producing an approximately straight-line. Potential correlation in the errors was similarly assessed on quantile residuals, both by examination of acf plots and a runs test. The data were ordered by set date within trip IDs, so the correlations are sought via sets that are temporally, and likely spatially, close.

In this case the Tweedie distribution is acceptable and there is no significant residual correlation once this zero rich distribution has been accounted for. The runs-test returns a *p*-value of 0.85.

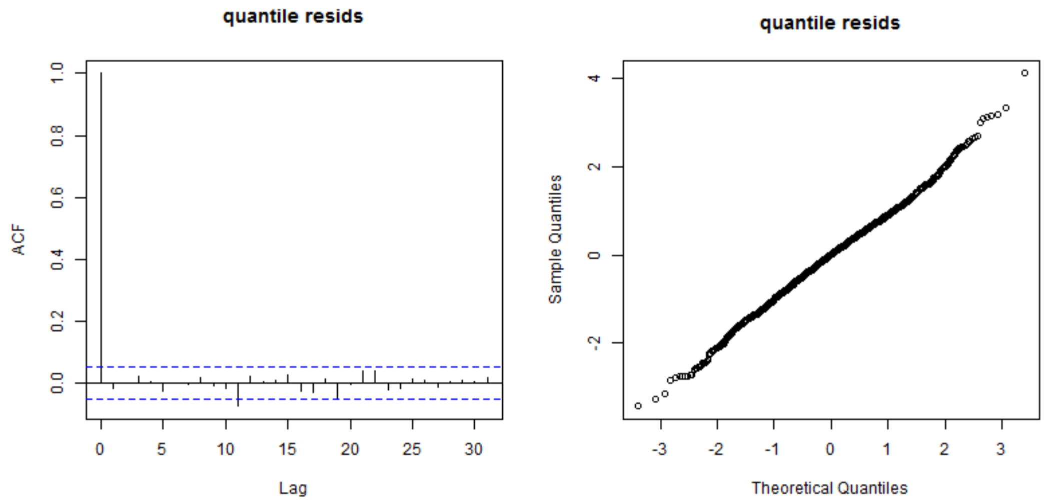


Figure 8: diagnostic plots for the quantile residuals from the GAM. The error distribution is assumed to be a Tweedie distribution, with key parameter estimated.

The following are link-scale estimates of the components of the GAM. All terms in the model are represented here regardless of practical/statistical significance. Generally only top-line results will be presented outside of the reference case. Where applicable, the parametric terms are presented first, followed by smooth terms and then interactions. The following observations can be made:

- There is a significantly higher catch rate when shark-lines are present.
- Shark-bait and wire-trace have negligible partial contributions to catch.
- There is temporal variability in terms of months and years, with high uncertainty about estimates.
- There is a general and significant increase in catch rate with respect to increasing isodepth.
- There is a significant and complex relationship between catch-rate and time that the sets were started.
- SST and wind-stress have significant unimodal relationships with catch rate.
- There is a significant spatial component to catch-rates which is not explained by the other covariates as evidenced by the significant latitude-longitude interaction surface. Generally catch rates increase to the East, but highest levels are observed at high latitudes in in the mid-longitude area.

- There is some catch relationship with both soak-time and the hooks between floats, the majority of the pattern being explicable by numbers of hooks between floats. In particular, baskets with few hooks have higher catch rates.

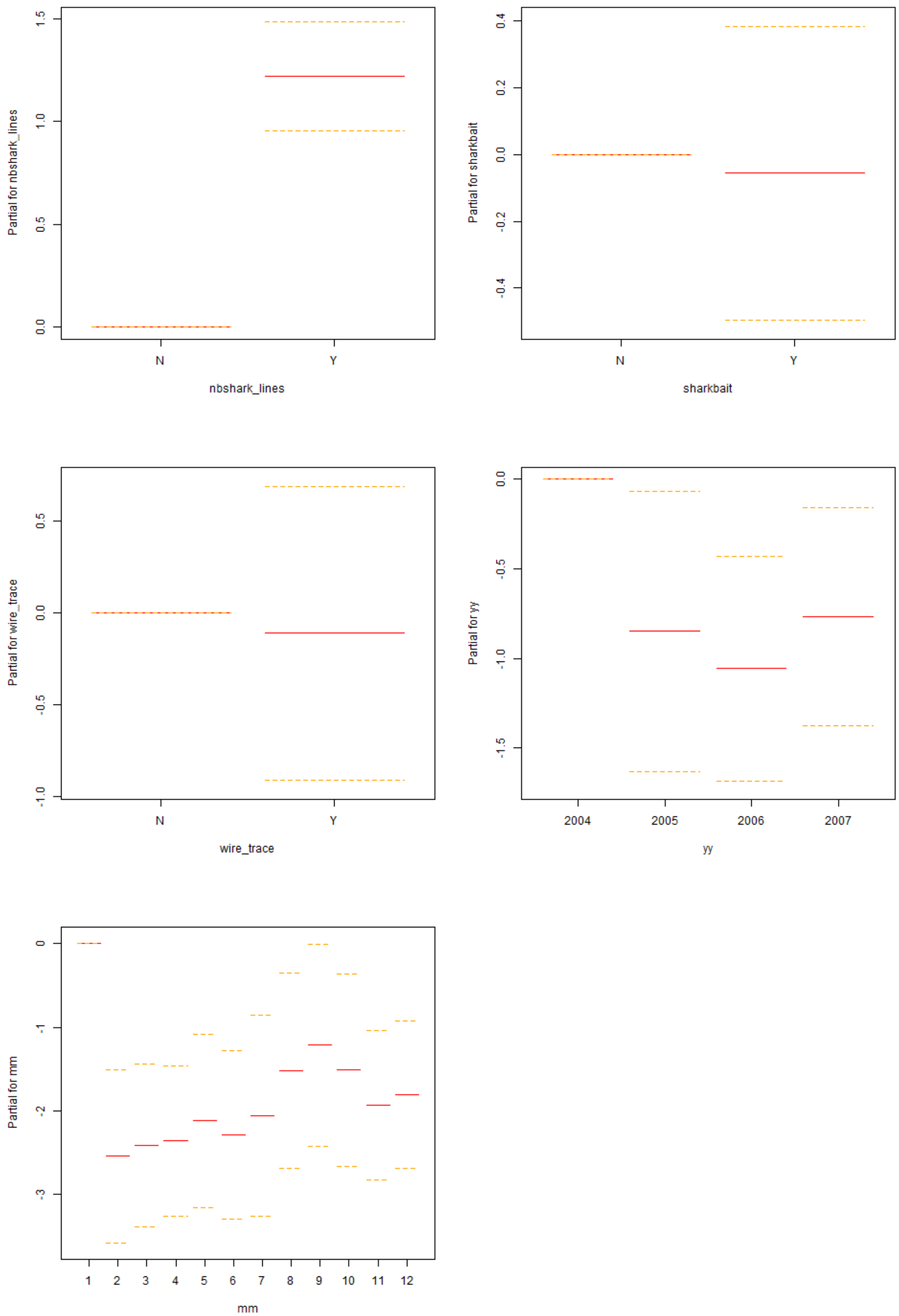


Figure 9: estimates and approximate 95% confidence intervals for parametric terms within the GAM.

DMP STATS The Coach House, Mount Melville House, St. Andrews, Scotland KY 168 NT, UK +44(0) 1334 477 544 mail@dmpstats.com

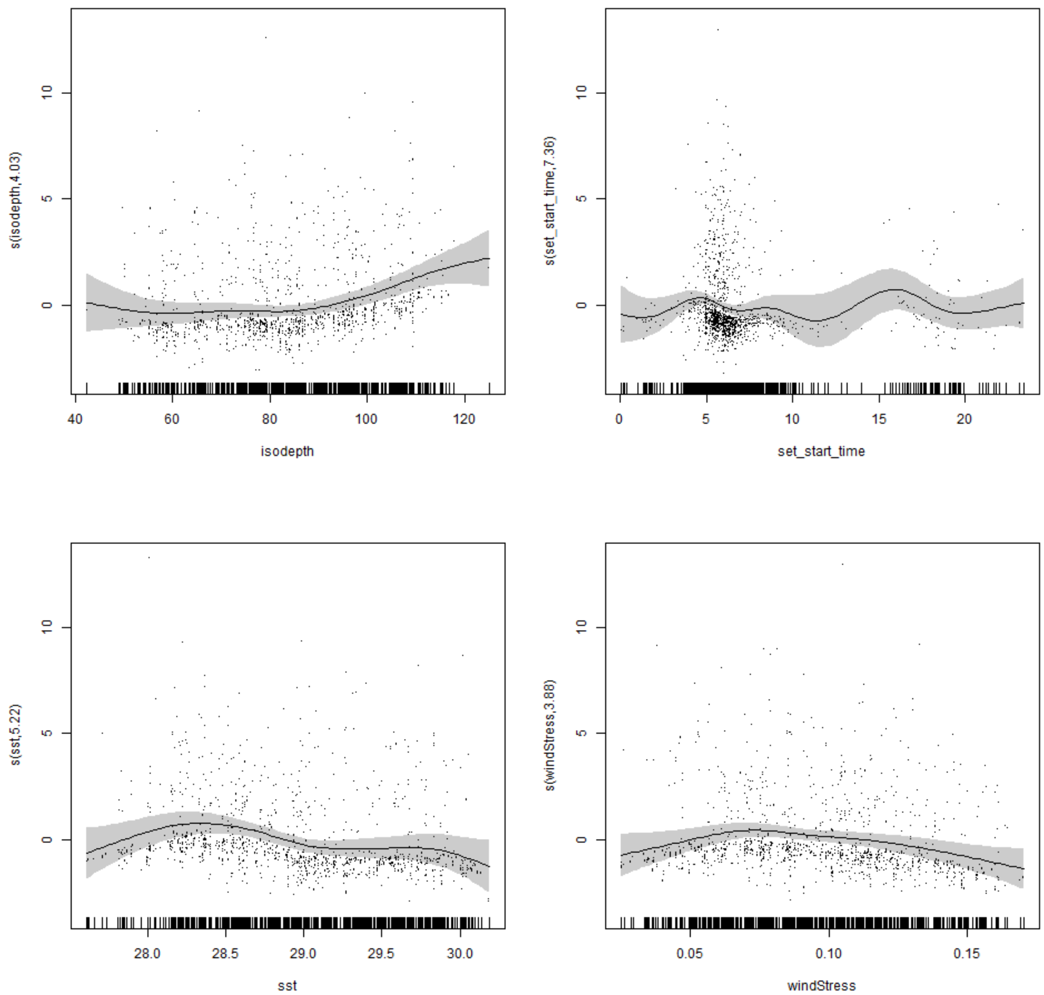


Figure 10: estimated smooths and approximate 95% confidence envelopes for smooth terms within the GAM.

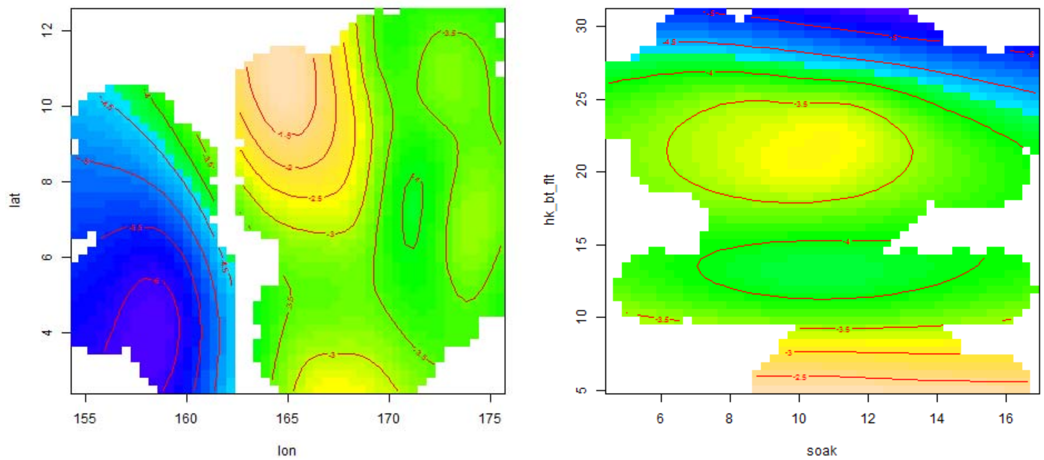


Figure 11: estimated smooth surfaces for the bivariate smooth components within the GAM.

5.4.2 Hook-level catch rate of oceanic whitetip sharks [reference case]

5.4.2.1 Data Exploration

Figure 12 shows that catch rates of oceanic whitetip sharks tend to be larger at lower hook position numbers (i.e. at shallower hooks). In general, there appears to be no clear patterns of change in catch rates at different hook positions that could be explained by changes in covariates considered in the present analysis. Still, and despite the considerable degree of overlapping between observations, there is sparse indication for larger catch rates of this species at shallower hooks when (Figure 12): (i) sets were performed in 2007; (ii) shark lines are in use; and (iii) sets took between 9 and 12 hours of soak time.

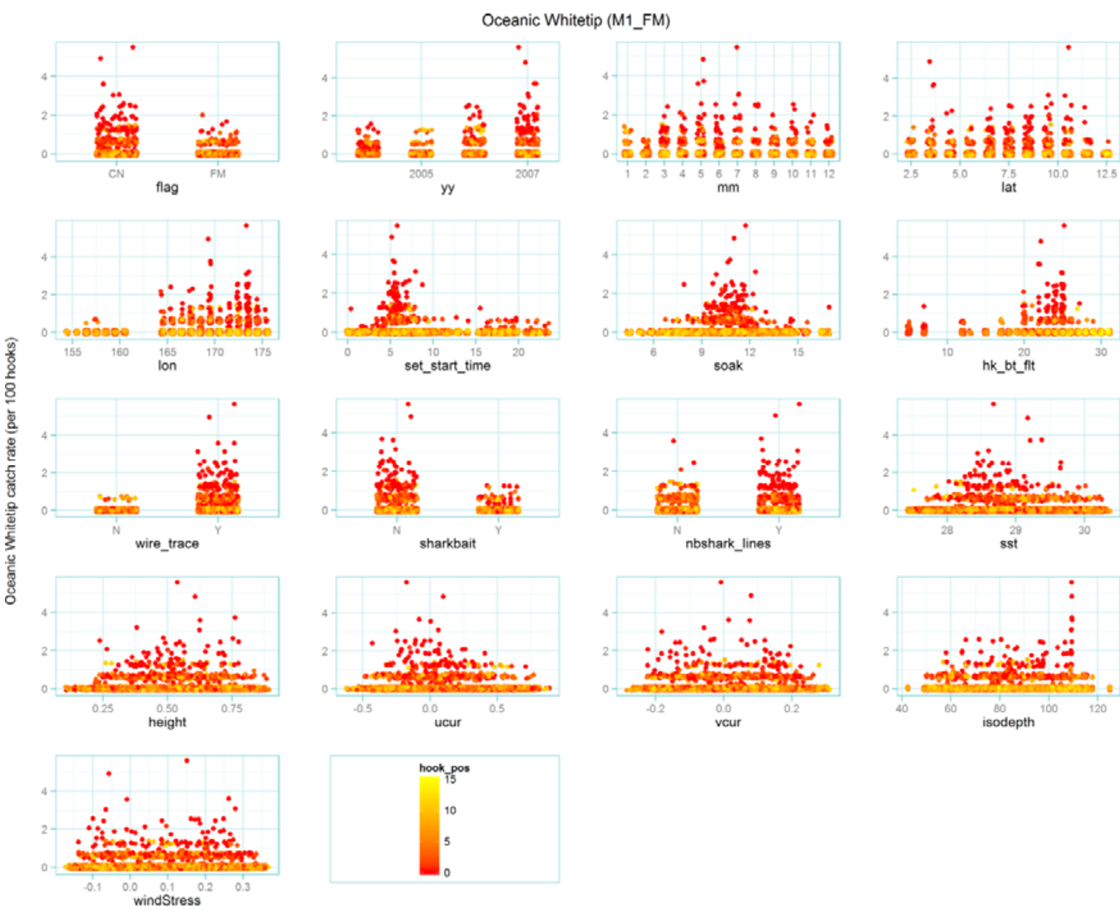


Figure 12: Distributions of per-hook catch-rate responses against the explanatory variables considered in the analysis, for oceanic whitetip sharks caught in the M1_FM fishery. Dots were slightly jittered to minimise overplotting and help visualisation and their colour convey the basket hook position of the respective observation (file "M1_FM-OCS_perHookRatesVsCovs.png").

5.4.2.2 Modelling results

The effect of hook position on catch rates was conditioned solely on species and fishery. A GAM was fitted with log-link and Tweedie error distribution. An initial model run was conducted for the purposes of estimating the governing Tweedie parameter – subsequently refitted with the estimated Tweedie parameter specified

The Tweedie parameter was estimated to be 1.109, with the model returning an adjusted-R^2 of 0.04 and a deviance explained of 17.3%.

The model assumptions were assessed. Non-linearities are not relevant as the hook position is fitted as a factor variable. The adequacy of the assumed Tweedie distribution was assessed via Quantile Residuals as indicated in Figure 13– these ought to be approximately Normally distributed, producing an approximately straight-line. Potential correlation in the errors was similarly assessed on quantile residuals, both by examination of acf plots and a runs test. The data were ordered by set date within trip IDs, so the correlations are sought via sets that are temporally, and likely spatially, close.

In this case the Tweedie distribution is acceptable and there is no significant residual correlation once this zero rich distribution has been accounted for. The runs-test returns a *p*-value of 0.084.

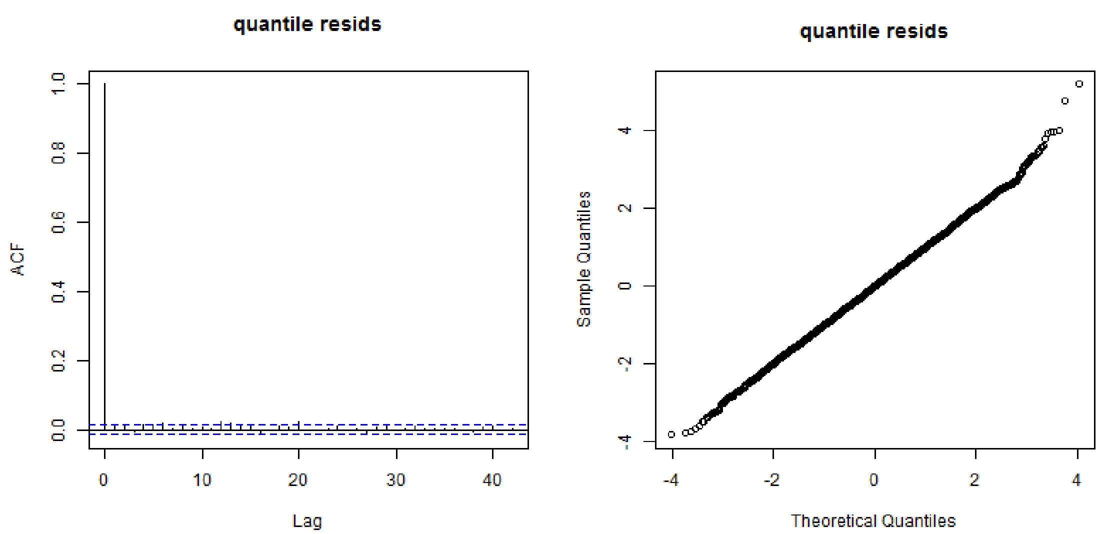


Figure 13: diagnostic plots for the quantile residuals from the GAM. The error distribution is assumed to be a Tweedie distribution, with key parameter estimated.

Estimated catch rate for each hook position and approximate 95% confidence intervals are presented in Figure 14 and Table 5.

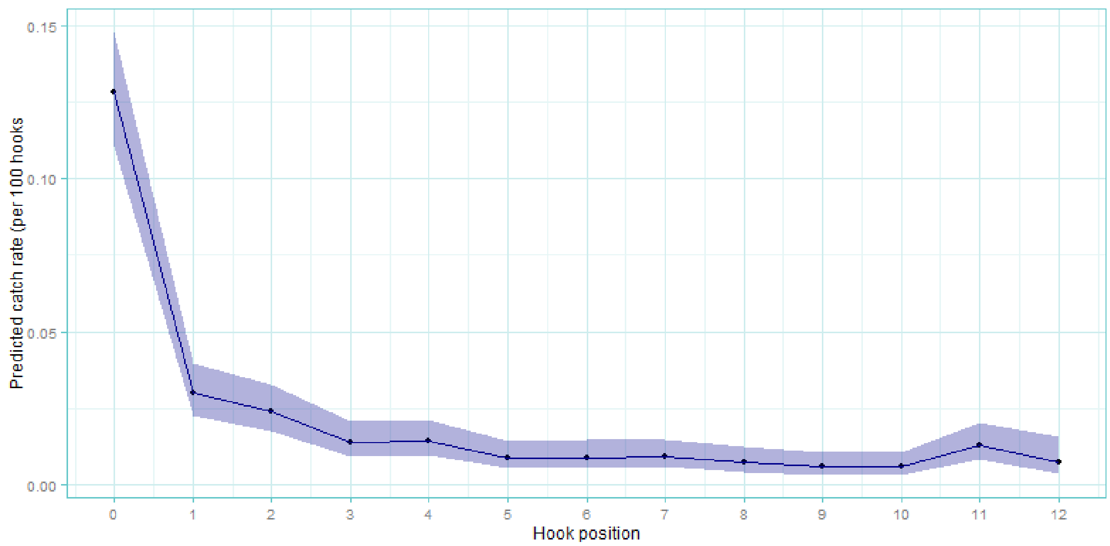


Figure 14: estimated mean catch rate (per 100 hooks) at hook-position with approximate 95% confidence envelope.

Table 5: Estimates of catch rates for each of the hook positions. Confidence intervals are approximate 95%. Percentage of total catch is calculated from overall estimated catch over hook positions presented. Confidence intervals for percentages are based on this naïve translation of cate rate.

Hook position	Catch rate (per 100 hooks)	SE	Lower CI	Upper CI	Catch (as % of total)	Lower CI (as % of total)	Upper CI (as % of total)
0	0.1282	1.0771	0.1109	0.1483	46.53	40.22	53.82
1	0.0298	1.153	0.0225	0.0394	10.8	8.17	14.28
2	0.0239	1.1699	0.0176	0.0325	8.68	6.38	11.8
3	0.0138	1.2236	0.0093	0.0206	5.02	3.38	7.46
4	0.0142	1.2218	0.0096	0.021	5.16	3.48	7.64
5	0.0086	1.2844	0.0053	0.0141	3.13	1.92	5.11
6	0.0089	1.2838	0.0054	0.0145	3.22	1.97	5.25
7	0.0092	1.2782	0.0057	0.0149	3.35	2.07	5.42
8	0.0072	1.3171	0.0042	0.0124	2.61	1.52	4.48
9	0.0058	1.3564	0.0032	0.0106	2.12	1.17	3.85

10	0.0057	1.3748	0.003	0.0106	2.06	1.11	3.85
11	0.0127	1.2625	0.008	0.02	4.59	2.91	7.25
12	0.0075	1.4417	0.0037	0.0154	2.73	1.33	5.58

The following can be observed from these:

- There is a markedly higher catch rate for this species at hook-position 0 compared to other positions.
- The catch rate at position 0 is 0.128 per 100 hooks [0.111 to 0.148, 95% CI]. This translates to approximately 46.5% [40.22% to 53.8%, 95% CI] of the catch of this species observed in this fisheries data.
- Beyond position 0, the catch rates are similar with some tendency towards higher catches at positions 1 & 2.

5.4.3 Condition of oceanic whitetip sharks at time of retrieval [reference case]

5.4.3.1 Data Exploration

While there seems to be no difference between the Chinese and Micronesian fleets in the condition of caught oceanic whitetip sharks at the time of retrieval, the proportion of dead sharks appears to change with the year at which the sets were observed (Figure 15). In addition, off the key fishing gear factors considered in the present analysis, the proportion of dead sharks appear to increase when wire leader is in use. Apart from a possible reduction in the death rate of oceanic whitetip sharks at hauling when wind stress levels remain at around 0.1 Newton/m^2, the graphical exploration suggests a lack of clear relationships between the condition of caught sharks and oceanographic covariates (Figure 15).

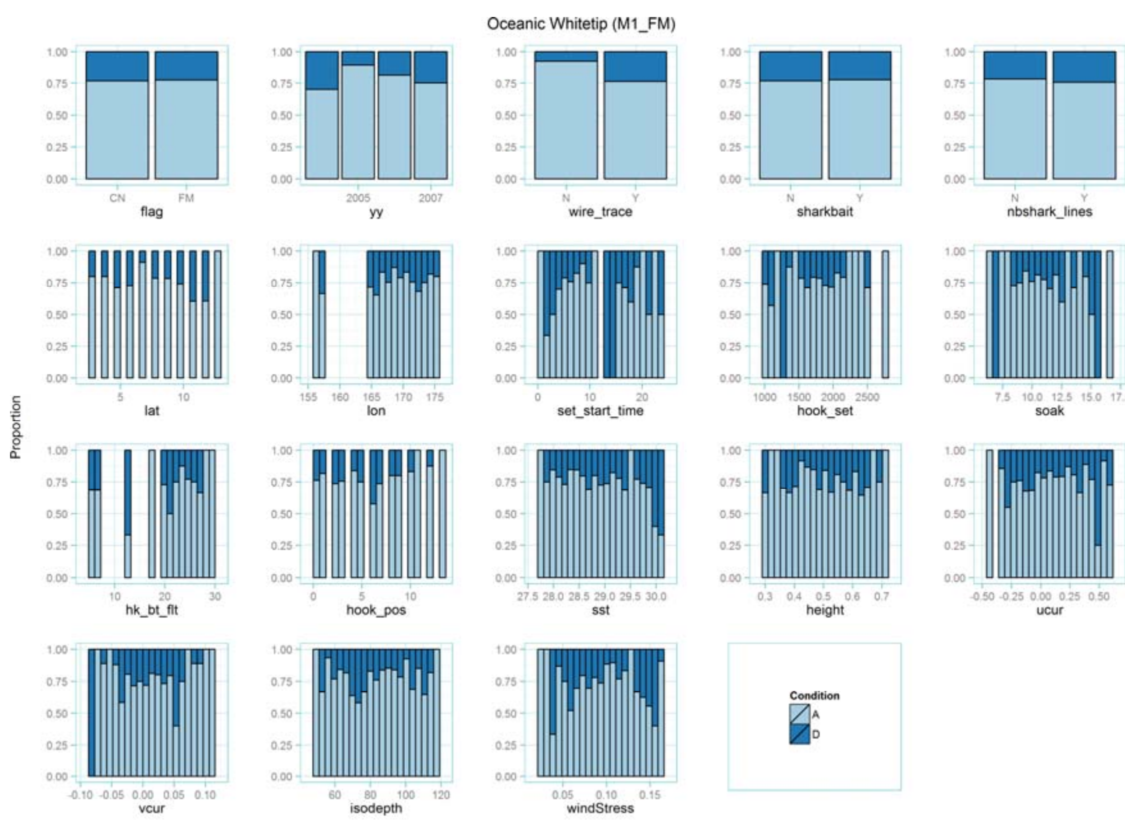


Figure 15: Distributions of condition responses (D= dead, A = alive), expressed in terms of proportion, against the explanatory variables considered for the analysis of oceanic whitetip sharks caught in the M1_FM fishery (file “M1_FM-OCS_CondtnVsCovs.png”).

5.4.3.2 Modelling results

Condition (dead/alive) was modelled as a function of set-level characteristics and associated oceanographic variables. The initial set of potential covariates was thinned on the basis of gross collinearity, as determined by Generalized Variance Inflation Factors (GVIFs). A range of potential interactions and non-linear terms were specified *a priori* as per Table 3. These were further thinned during the modelling process in some cases where estimation was not possible e.g. interactions for factor variables where certain combinations of factors were not observed in the dataset.

The model fitted was a GAM with logit-link and binomially-distributed errors. Model "selection" was conducted within the GAM using shrinkage, so model terms of little relevance have their effective degrees of freedom shrunk towards zero. In terms of parametric components, backwards selection (p-value<0.05) was used, but not on the key factors of wire-trace, shark-bait, shark-lines and hook type (where applicable). These were generally retained in the model for interpretative interest. The results of this process are summarised in Table 4.

In terms of fit to the data, the software generated adjusted R^2 and deviance explained were 0.062 and 9.31% respectively, indicating poor fit. However these measures are of dubious utility for binomial error models. Raw predictive power to the current dataset was also assessed via confusion matrices, providing Sensitivity (true positive rate) = 0.72, Specificity (true negative rate) = 0.59 and a misclassification rate of 0.38 when using the response mean as a decision boundary. These are against the training data and will be over-estimates –cross-validated estimates would be preferred with future analysis.

Table 6: Model terms from the generalized additive model. DF = Degrees of Freedom. EDF=Effective Degrees of Freedom. Shrinkage has been applied as a means of model selection. Terms of little relevance have their EDF shrunken.

Model terms	EDF	DF	Chi or F-stat	*p*-value
yy		3	2.39	0.495
flag		1	0.02	0.890
sharkbait		1	0.98	0.323
nbshark_lines		1	0.06	0.803
wire_trace		1	1.71	0.191
s(set_start_time)	0.00	9	0.00	0.633
s(sst)	0.96	9	2.54	0.043
s(height)	0.00	9	0.00	0.303
s(ucur)	1.12	9	1.90	0.139
s(vcur)	0.00	9	0.00	0.571
s(isodepth)	4.68	9	14.10	0.004
s(windStress)	0.00	9	0.00	0.461
s(soak)	0.00	9	0.00	0.765
s(hook_pos)	2.96	9	5.70	0.100
s(hk_bt_flt)	1.01	9	2.15	0.109
s(lon,lat)	4.39	29	11.96	0.004
s(soak,hk_bt_flt)	0.00	27	0.00	0.477

The model assumptions were assessed. Non-linearities are assumed to be captured via the GAM which optimises complexity. The adequacy of the assumed binomial distribution was assessed via Quantile Residuals (Dunn & Smyth, 1996), as indicated in Figure 16 – these ought to be approximately Normally distributed, producing an approximately straight-line. Potential correlation in the errors was similarly assessed on quantile residuals, both by examination of acf plots and a runs test. The data were ordered by set date within trip IDs, so the correlations are sought via sets that are temporally, and likely spatially, close. The residuals are distributionally acceptable and little autocorrelation is in evidence, the runs-test providing a *p*-value of 0.042.

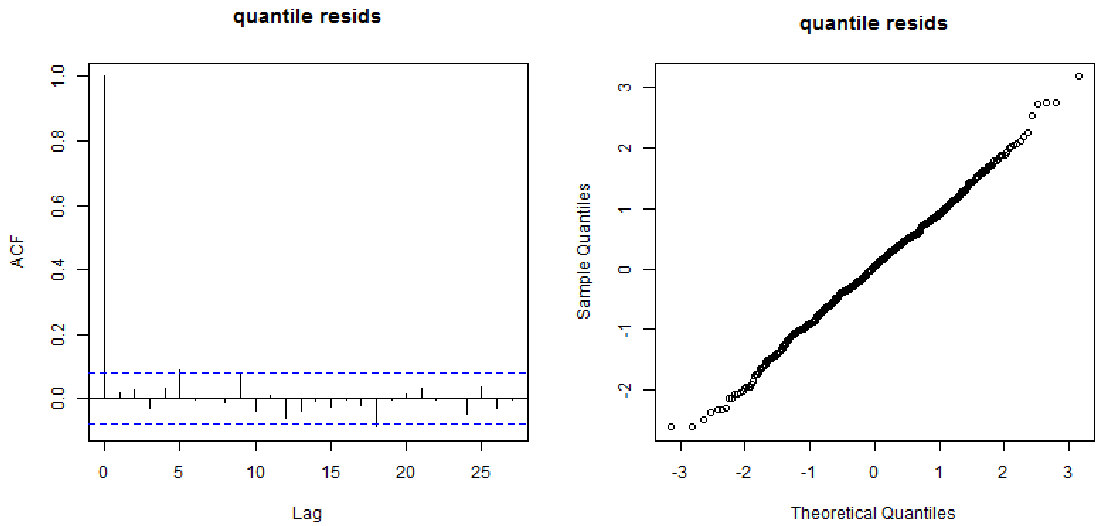
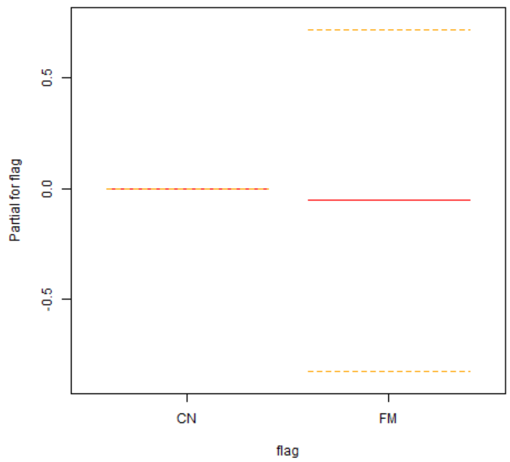


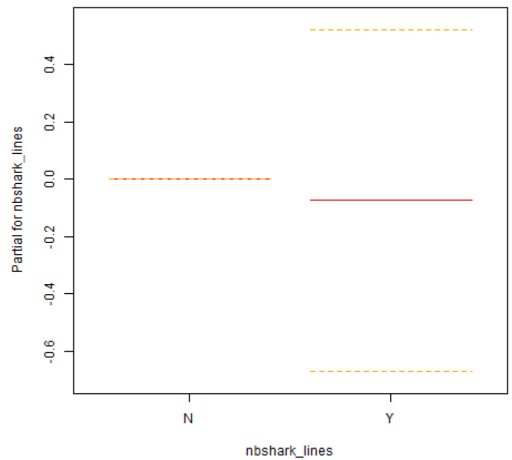
Figure 16: diagnostic plots for the quantile residuals from the GAM. The error distribution is assumed to be a Tweedie distribution, with key parameter estimated.

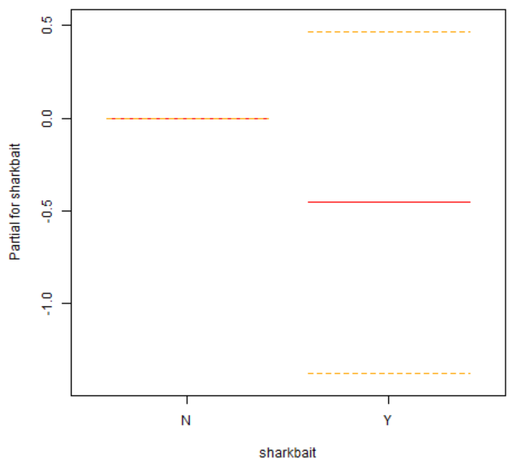
The following are link-scale estimates of the components of the GAM. All terms in the model are represented here regardless of practical/statistical significance although near-zero EDF relationships are excluded. Generally only top-line results will be presented outside of the reference case. Where applicable, the parametric terms are presented first, followed by smooth terms and then interactions.

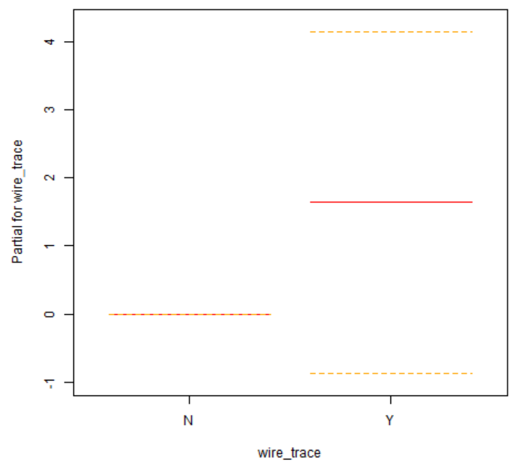
The following observations can be made:

- Little significant pattern with regards the key factors of wire-trace, shark-bait and shark-lines.
- No significant differences between flags.
- Some significant relationship with isodepth, with a tendency towards higher mortality with higher isodepth.
- Higher mortality with lower numbers of hooks in baskets and a tendency towards higher mortality in higher latititudes.


Partial for flag
0.5
0.0
-0.5
CN
FM
flag


Partial for nbshark_lines
0.4
0.2
0.0
-0.2
-0.4
-0.6
N
Y
nbshark_lines


Partial for sharkbait
0.5
0.0
-0.5
-1.0
N
Y
sharkbait


Partial for wire_trace
4
3
2
1
0
-1
N
Y
wire_trace

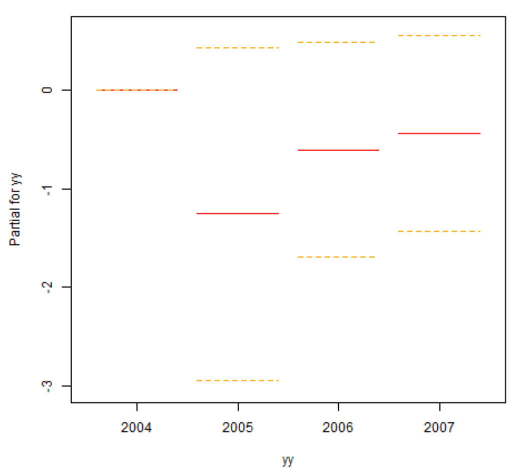

Partial for yy
0
-1
-2
-3
2004
2005
2006
2007
yy

Figure 17: estimates and approximate 95% confidence intervals for parametric terms within the GAM.

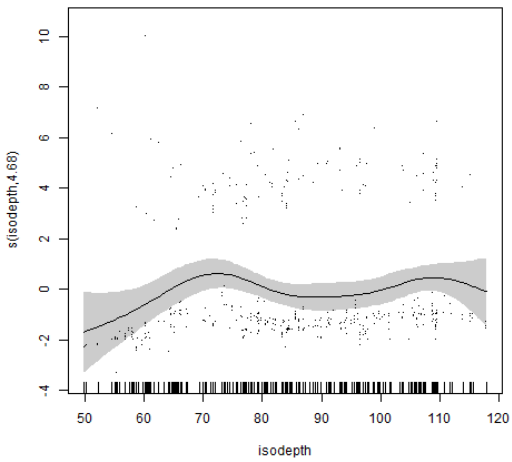


Figure 18: estimated smooths and approximate 95% confidence envelopes for smooth terms within the GAM.

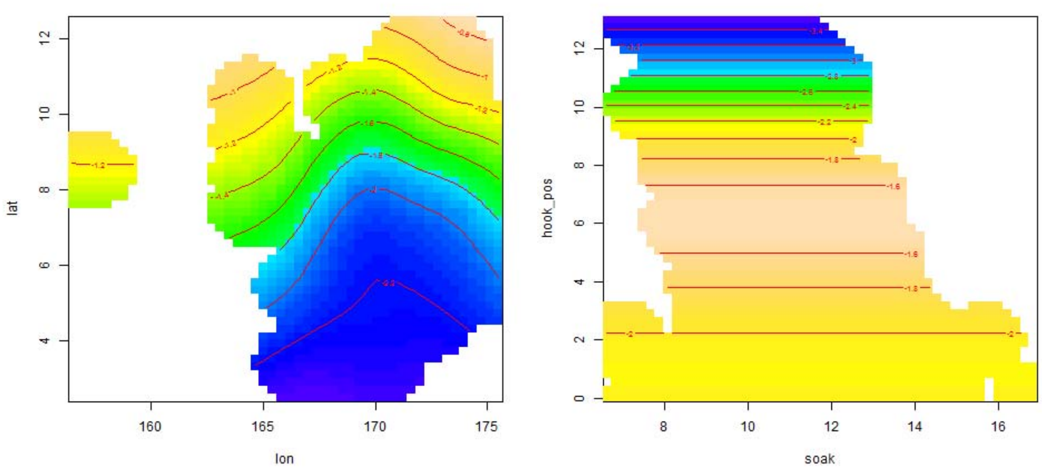


Figure 19: estimated smooth surfaces for the bivariate smooth components within the GAM.

5.5 CN & FM COMBINED FLEET ANALYSIS – SILKY SHARKS (FAL)

5.5.1 Catch rate of Silky Sharks (FAL)

5.5.1.1 Data Exploration

While predominantly absent in sets from observed trips in the M1_FM fishery, within-set bycatch levels of silky sharks decrease gradually from 1 shark/set (in about 300 sets) to 14 sharks/set (in about 10 sets), with one of the sets catching the maximum of 35 silky sharks (Figure 20).

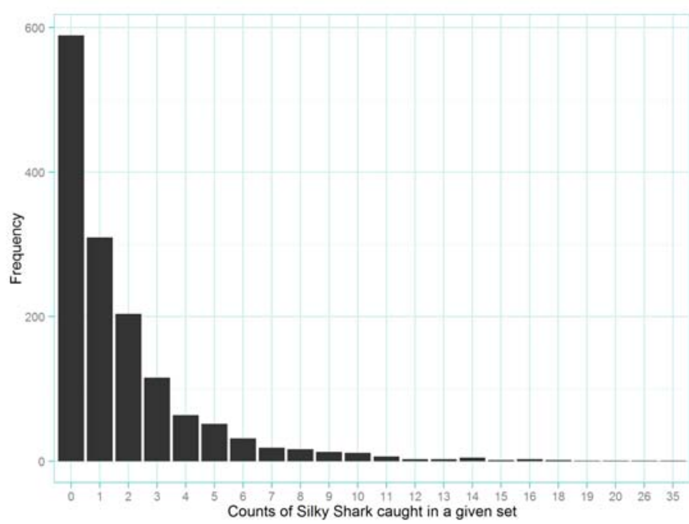


Figure 20: Distribution of counts per set of silky shark caught in observed trips of the M1_FM longline fishery (M1_FM-FAL_CatchNumFreqs.png)

Figure 21 shows no substantial patterns in bivariate scatterplots between catch rates of silky sharks and the covariates considered in the present analysis. Yet, there are some indications for higher catch rates when:

- observed trips took place in 2004 and 2007;
- wire leader, shark bait and shark lines were in use;
- sets were performed in April and between September-November;
- the average sea surface temperature remained around 28.5^{o}C;
- and average horizontal current speed (ucur) reached values of approximately 0.12 m/s.

DMP STATS The Coach House, Mount Melville House, St. Andrews, Scotland KY 168 NT, UK +44(0) 1334 477 544 mail@dmpstats.com

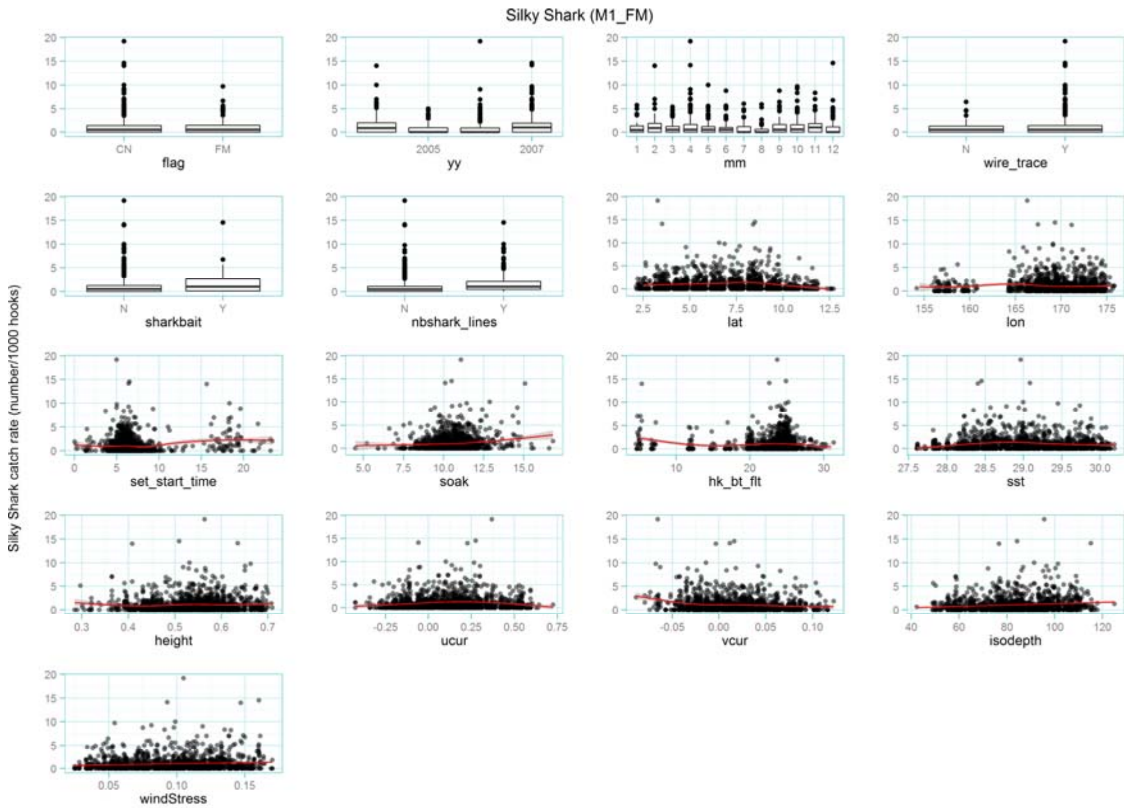


Figure 21: Distributions of set-level catch-rate responses against the explanatory variables considered for the analysis of silky sharks caught in the M1_FM fishery (file "M1_FM-FAL_CatchRatesVsCovs.png")

5.5.1.2 Modelling results

Modelling details are as per the reference case. The model returned an adjusted-R^2 of 0.248 and a deviance explained of 30.6%.

Table 7: Model terms from the generalized additive model. DF = Degrees of Freedom. EDF=Effective Degrees of Freedom. Shrinkage has been applied as a means of model selection. Terms of little relevance have their EDF shrunken.

Model terms	EDF	DF	Chi or F-stat	*p*-value
yy		3	9.93	0.000
mm		11	4.92	0.000
wire_trace		1	0.12	0.732
nbshark_lines		1	12.61	0.000
sharkbait		1	4.46	0.035
flag		1	2.79	0.095
s(set_start_time)	1.14	9	0.69	0.004
s(sst)	7.75	9	3.36	0.000
s(height)	2.08	9	0.38	0.135
s(ucur)	4.37	9	1.11	0.023
s(vcur)	0.97	9	3.27	0.000
s(isodepth)	4.10	9	0.90	0.047
s(windStress)	0.00	9	0.00	0.885
s(lon,lat)	17.34	29	2.10	0.000
s(soak,hk_bt_flt)	15.91	29	1.22	0.001

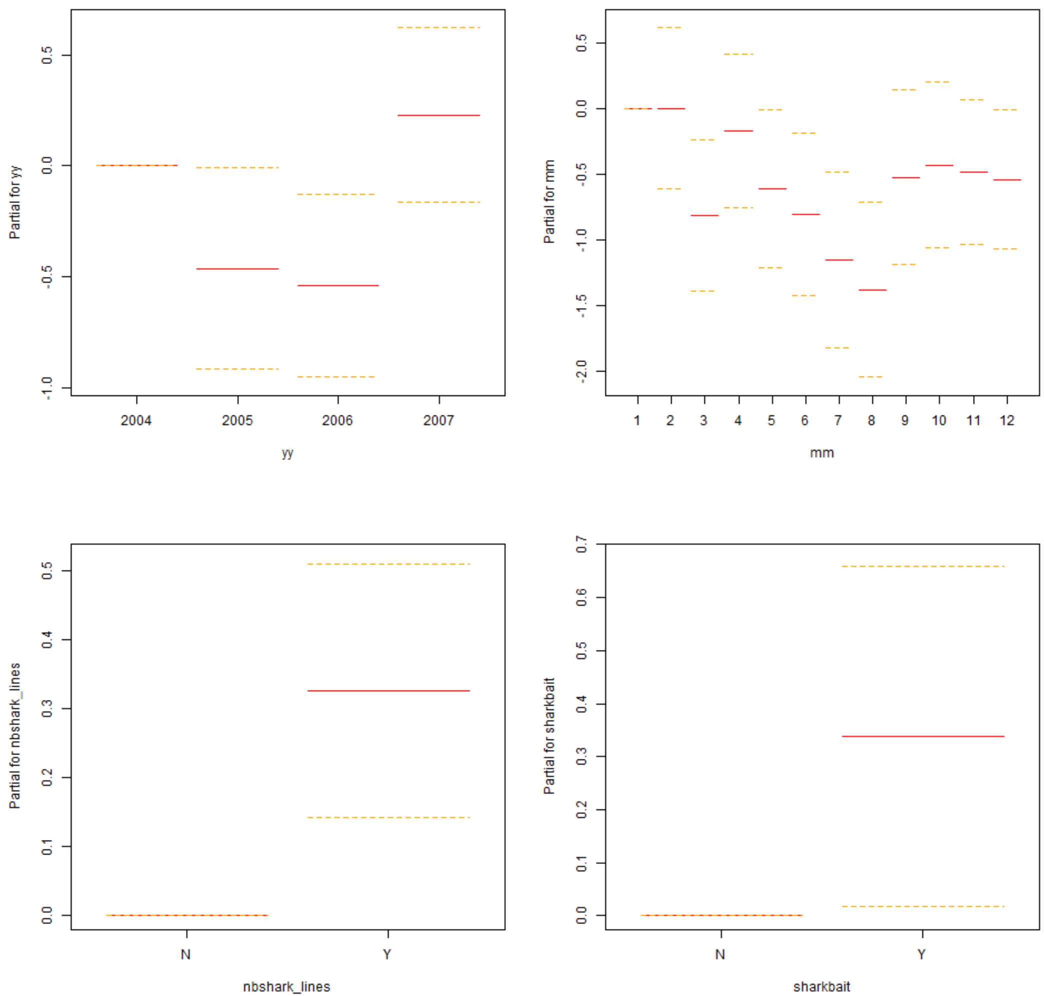


Figure 22: estimates and approximate 95% confidence intervals for parametric terms within the GAM.

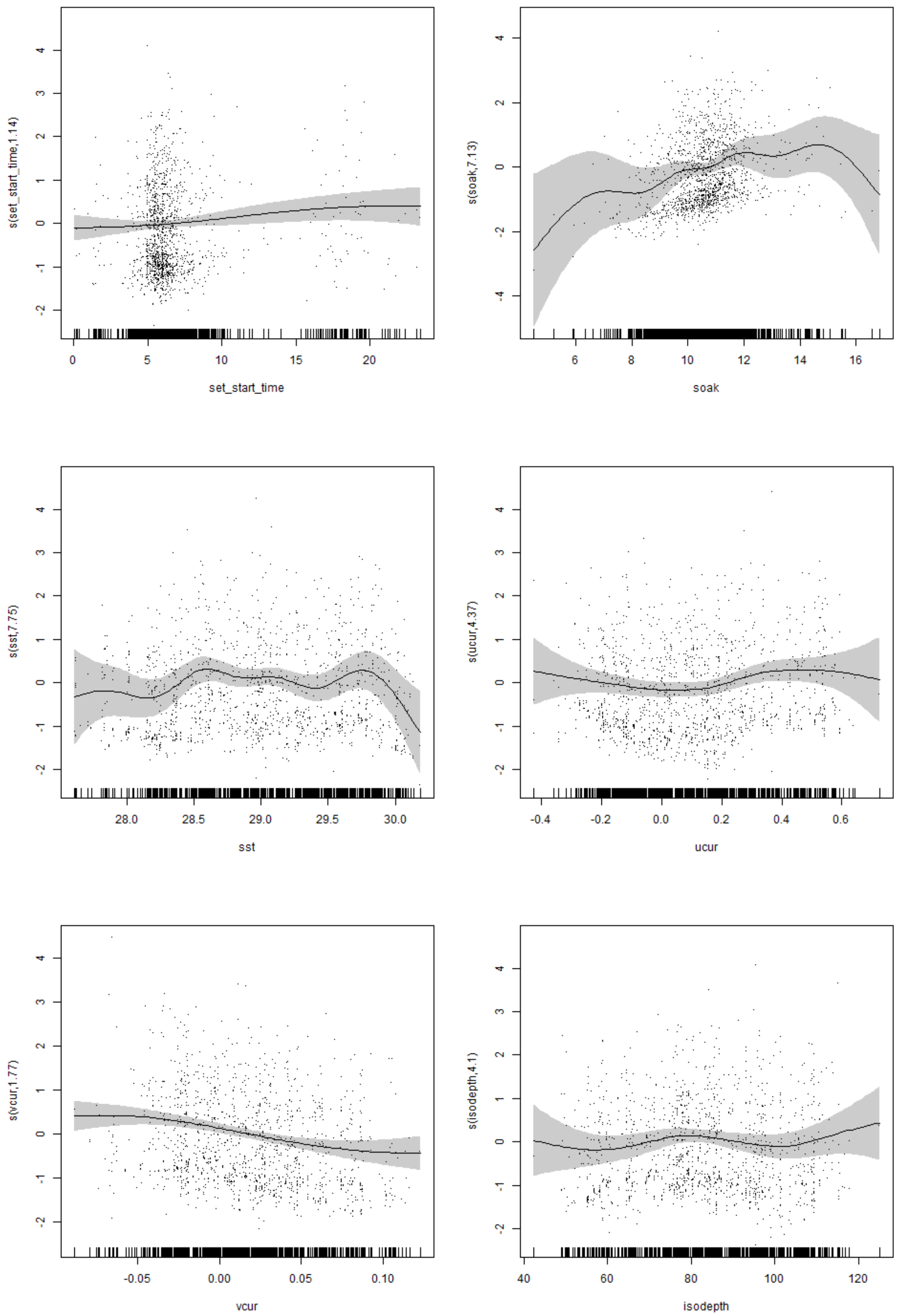


Figure 23: estimated smooths and approximate 95% confidence envelopes for smooth terms within the GAM.

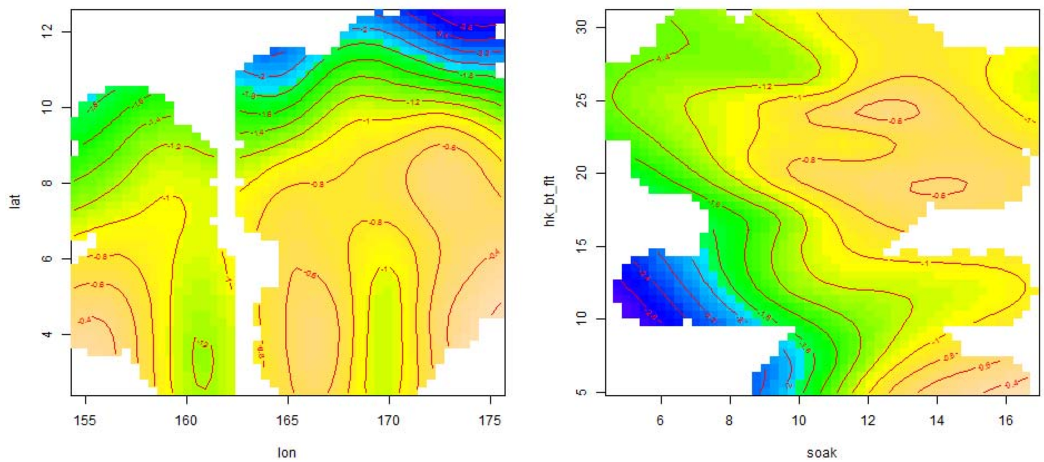


Figure 24: estimated smooth surfaces for the bivariate smooth components within the GAM.

5.5.2 Hook-level catch rate of silky sharks

5.5.2.1 Data Exploration

Figure 25 shows that catch rates of silky sharks tend to be larger at lower hook position numbers (i.e. at shallower hooks), but there are also a few deeper hooks (i.e. higher hook position values) with large catch rates. In general, there appears to be no clear patterns of change in catch rates at different hook positions across the values of the covariates considered in the present analysis. Still, and despite the considerable degree of overlapping between observations, there is loose indication for larger catch rates of this species at shallower hooks when (Figure 25): (i) sets were performed in 2007; (ii) shark lines are in use; and (iii) average sea surface temperatures were around 28.5^{o}C.

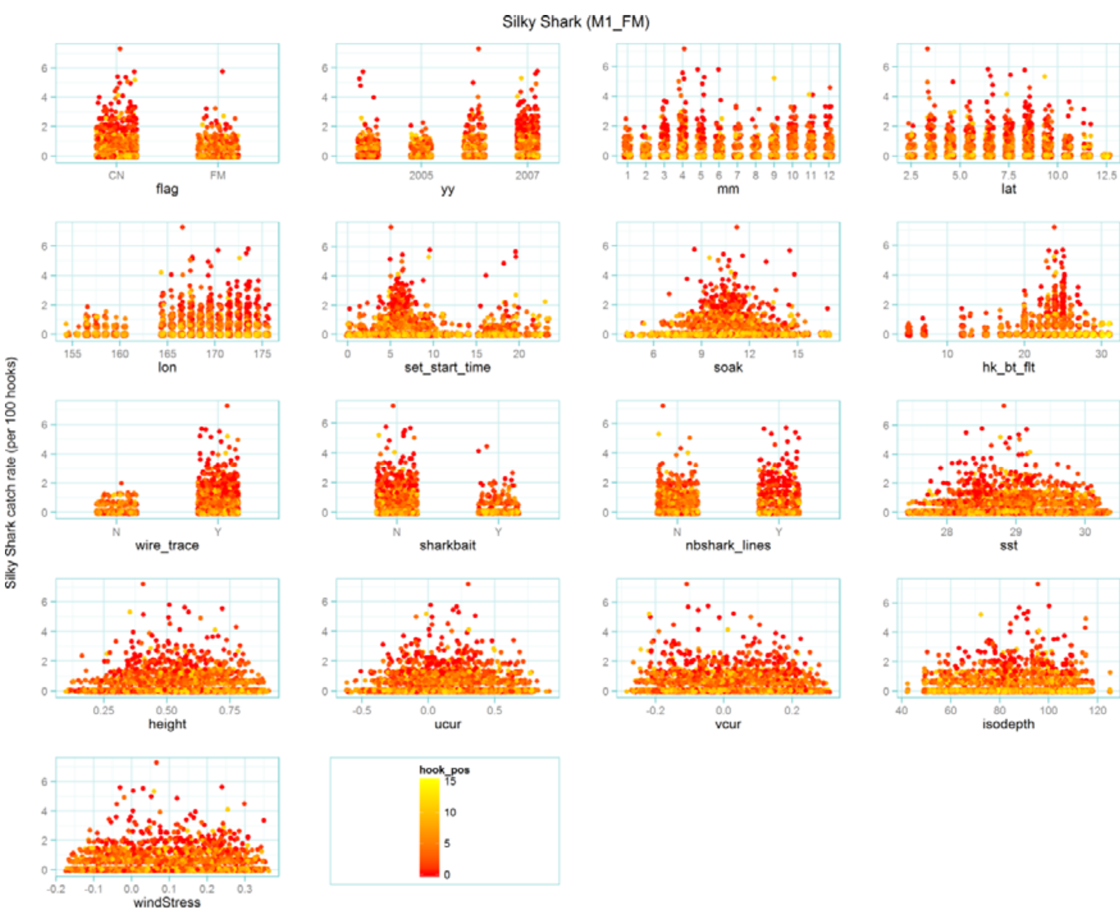


Figure 25: Distributions of per-hook catch-rate responses against the explanatory variables considered in the analysis, for silky sharks caught in the M1_FM fishery. Dots were slightly jittered to minimise overplotting and help visualisation and their colour convey the basket hook position of the respective observation (file "M1_FM-FAL_perHookRatesVsCovs.png").

5.5.2.2 Modelling results

Modelling details are as per the reference case. The model returned an adjusted-R^2 of 0.026 and a deviance explained of 6.91%.

Table 8: Estimates of catch rates for each of the hook positions. Confidence intervals are approximate 95%. Percentage of total catch is calculated from overall estimated catch over hook positions presented. Confidence intervals for percentages are based on this naïve translation of cate rate.

Hook position	Catch rate (per 100 hooks)	SE	Lower CI	Upper CI	Catch (as % of total)	Lower CI (as % of total)	Upper CI (as % of total)
0	0.2135	1.0624	0.1896	0.2404	17.76	15.77	20
1	0.1910	1.0658	0.1686	0.2164	15.89	14.02	18
2	0.1452	1.0748	0.1260	0.1672	12.08	10.48	13.91
3	0.1225	1.0817	0.1051	0.1429	10.19	8.74	11.89
4	0.0906	1.0948	0.0758	0.1081	7.53	6.31	8.99
5	0.0736	1.1045	0.0606	0.0895	6.13	5.04	7.44
6	0.0571	1.1194	0.0458	0.0712	4.75	3.81	5.93
7	0.0627	1.1143	0.0507	0.0775	5.21	4.22	6.45
8	0.0397	1.1431	0.0305	0.0516	3.3	2.54	4.29
9	0.0435	1.1380	0.0338	0.0561	3.62	2.81	4.66
10	0.0386	1.1512	0.0293	0.0509	3.21	2.44	4.24
11	0.0781	1.1129	0.0634	0.0964	6.5	5.27	8.02
12	0.0460	1.1845	0.0330	0.0641	3.82	2.74	5.33

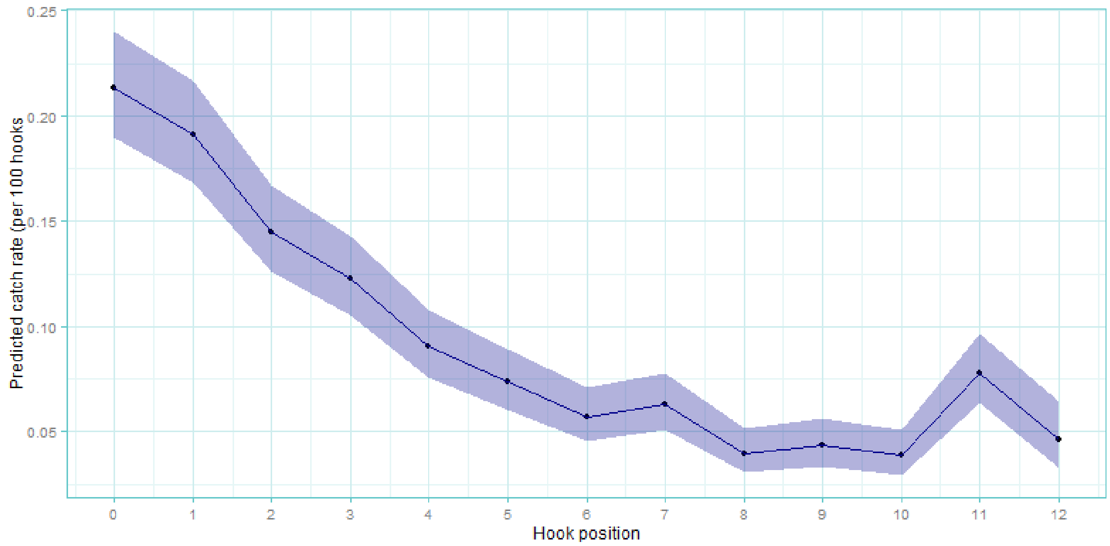


Figure 26: estimated mean catch rate (per 100 hooks) at hook-position with approximate 95% confidence envelope.

5.5.3 Condition of silky sharks at time of retrieval

5.5.3.1 Data Exploration

While there seems to be no difference between the Chinese and Micronesian fleets in the condition of caught silky sharks at the time of retrieval, the proportion of dead sharks appears to have decreased in 2005 (Figure 27). In addition, off the key fishing gear factors considered in the present analysis, the proportion of dead sharks appear to decrease marginally when shark lines and shark bait were in use. The graphical investigation of the data suggests a lack of clear relationships between the condition of caught sharks and fishing gear features and oceanographic covariates (Figure 27).

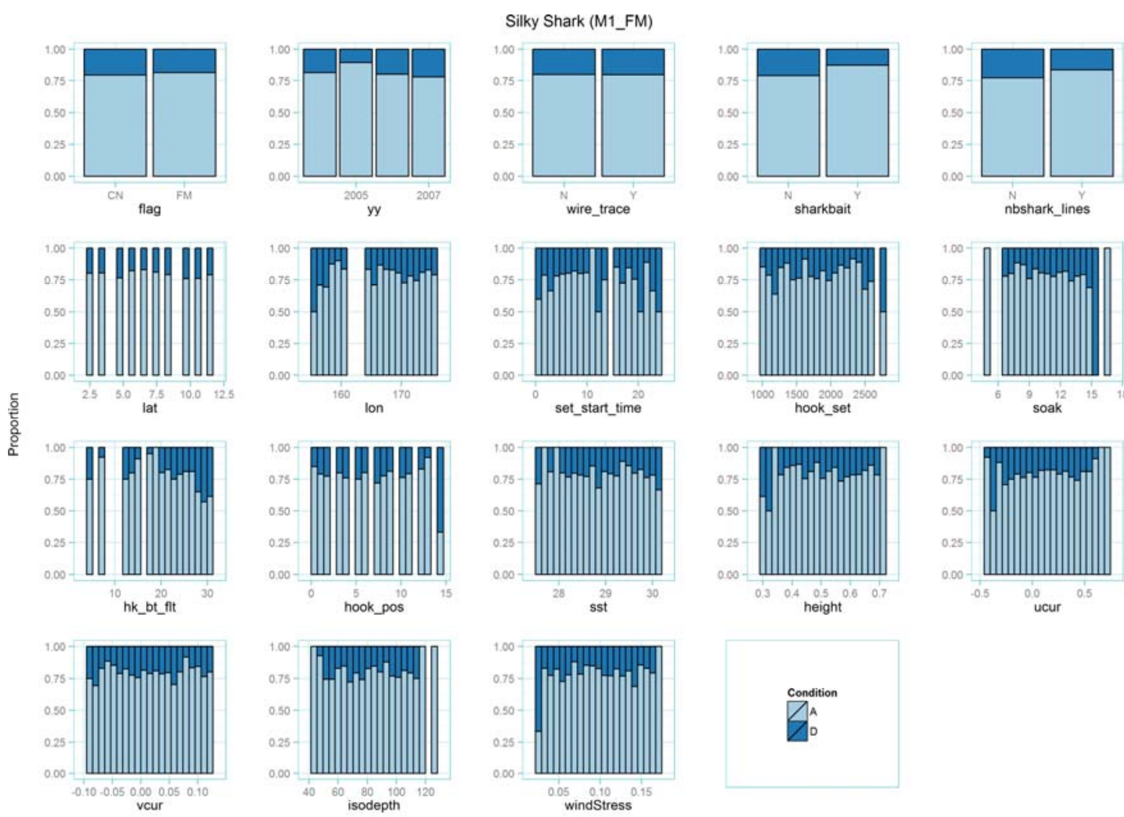


Figure 27: Distributions of condition responses (D= dead, A = alive), expressed in terms of proportion, against the explanatory variables considered for the analysis of silky sharks caught in the M1_FM fishery (file "M1_FM-FAL_CondtnVsCovs.png")

5.5.3.2 Modelling results

Modelling details are as per the reference case. The model returned an adjusted-R^2 of 0.053 and a deviance explained of 7.28%. Raw predictive power to the current dataset was also assessed via confusion matrices on the training data, providing Sensitivity (true positive rate) = 0.64, Specificity (true negative rate) = 0.60 and a misclassification rate of 0.39 when using the response mean as a decision boundary. These are against the training data and will be over-estimates.

Table 9: Model terms from the generalized additive model. DF = Degrees of Freedom. EDF=Effective Degrees of Freedom. Shrinkage has been applied as a means of model selection. Terms of little relevance have their EDF shrunken.

Model terms	EDF	DF	Chi or F-stat	*p*-value
yy		3	26.44	0.000
flag		1	0.93	0.334
sharkbait		1	0.80	0.371
nbshark_lines		1	7.83	0.0052
wire_trace		1	8.58	0.003
s(set_start_time)	4.68	9	8.34	0.066
s(sst)	0.00	9	0.00	0.612
s(height)	0.00	9	0.00	0.601
s(ucur)	7.60	9	19.93	0.003
s(vcur)	0.00	9	0.00	0.482
s(isodepth)	0.00	9	0.00	0.730
s(windStress)	3.08	9	9.19	0.012
s(hook_pos)	1.84	9	4.48	0.065
s(lon,lat)	8.19	29	18.40	0.003
s(soak,hk_bt_flt)	20.85	29	59.75	0.000

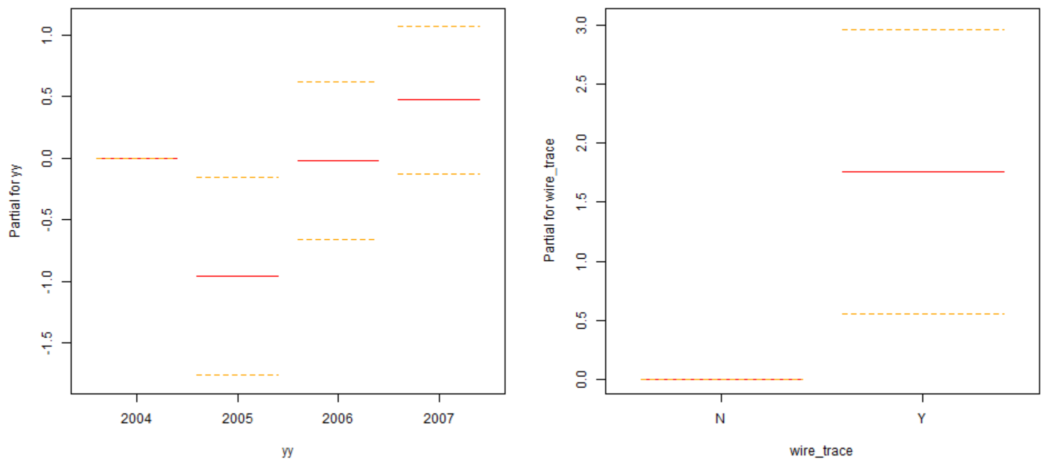


Figure 28: estimates and approximate 95% confidence intervals for parametric terms within the GAM.

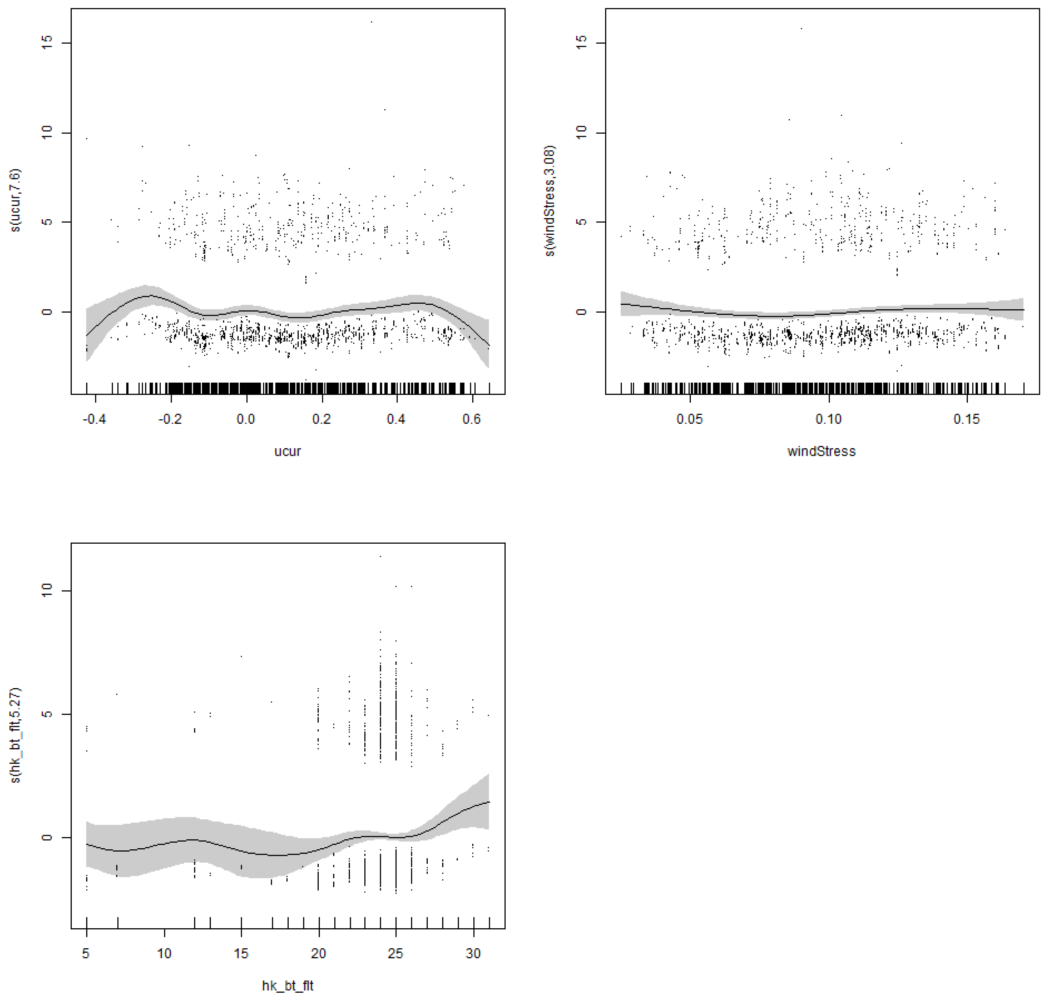


Figure 29: estimated smooths and approximate 95% confidence envelopes for smooth terms within the GAM.

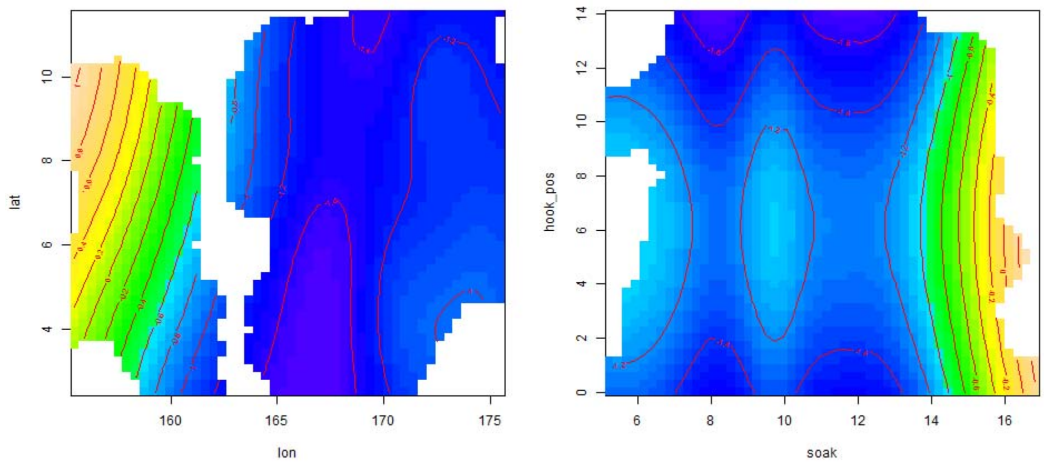


Figure 30: estimated smooth surfaces for the bivariate smooth components within the GAM.

DMP STATS The Coach House, Mount Melville House, St. Andrews, Scotland KY 168 NT, UK +44(0) 1334 477 544 mail@dmpstats.com

5.6 CN & FM COMBINED FLEET ANALYSIS – THRESHER SHARKS (THR)

5.6.1 Catch rate of Thresher Sharks (THR)

5.6.1.1 Data Exploration

While predominantly absent in sets from observed trips in the M1_FM fishery, within-set bycatch levels of thresher sharks decay gradually from 1 shark/set (in about 300 set) to 13 sharks/set (in about 10 sets), with one of the sets catching the maximum of 34 thresher sharks (Figure 31).

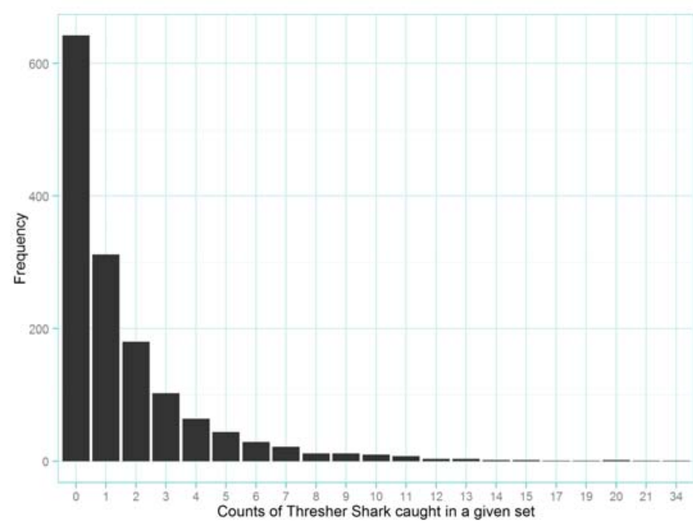


Figure 31: Distribution of counts per set of thresher shark caught in observed trips of the M1_FM longline fishery (M1_FM-THR_CatchNumFreqs.png)

Figure 32 shows no demarked patterns in bivariate scatterplots between catch rates of thresher sharks and the covariates considered in the present analysis. Yet, there are some indications for higher catch rates when:

- observed trips took place in 2004 and 2007 and between the period of March-June;
- wire leader and shark lines were in use;
- sets were performed between 7.5-10 degrees of latitude and within 170-175 degrees of longitude

- the average sea surface temperature varied between 28^{o}C - 29^{o}C, and sea surface height remained within 0.4-0.5 meters
- average horizontal current speed (ucur) reached values of approximately -0.1 m/s;
- the average depth of the 20^{o}C isotherm was situated at around 80 meters, and the average wind stress levels remain slightly over 0.10 Newton/m^2

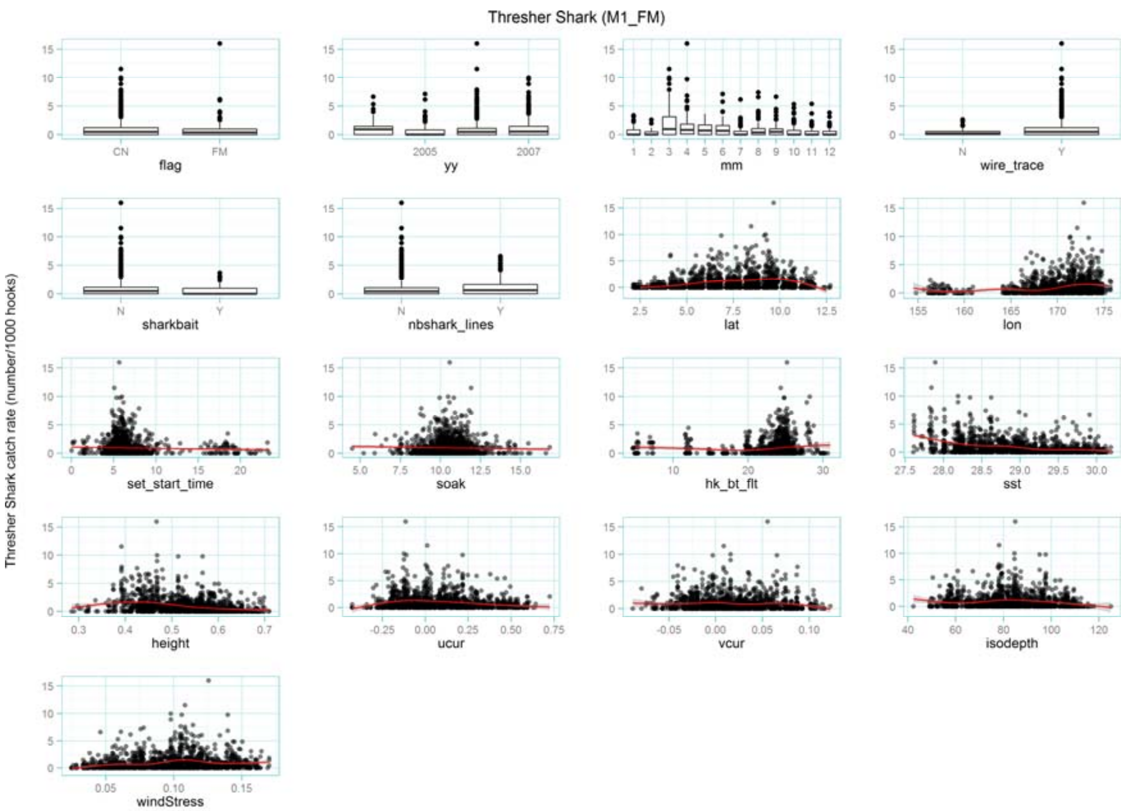


Figure 32: Distributions of set-level catch-rate responses against the explanatory variables considered for the analysis of thresher sharks caught in the M1_FM fishery (file “M1_FM-THR_CatchRatesVsCovs.png”)

5.6.1.2 Modelling results

Modelling details are as per the reference case. The model returned an adjusted-R^2 of 0.393 and a deviance explained of 45%.

Table 10: Model terms from the generalized additive model. DF = Degrees of Freedom. EDF=Effective Degrees of Freedom. Shrinkage has been applied as a means of model selection. Terms of little relevance have their EDF shrunken.

Model terms	EDF	DF	Chi or F-stat	*p*-value
yy		3	4.76	0.003
mm		11	3.34	0.000
wire_trace		1	2.65	0.104
nbshark_lines		1	0.49	0.486
sharkbait		1	0.01	0.933
flag		1	2.33	0.127
s(set_start_time)	4.93	9	2.47	0.000
s(sst)	8.05	9	4.09	0.000
s(height)	5.10	9	3.12	0.000
s(ucur)	7.88	9	3.43	0.000
s(vcur)	4.23	9	1.09	0.029
s(isodepth)	2.80	9	1.83	0.000
s(windStress)	1.70	9	0.44	0.073
s(lon,lat)	21.71	29	5.52	0.000
s(soak,hk_bt_flt)	1.79	29	0.23	0.020

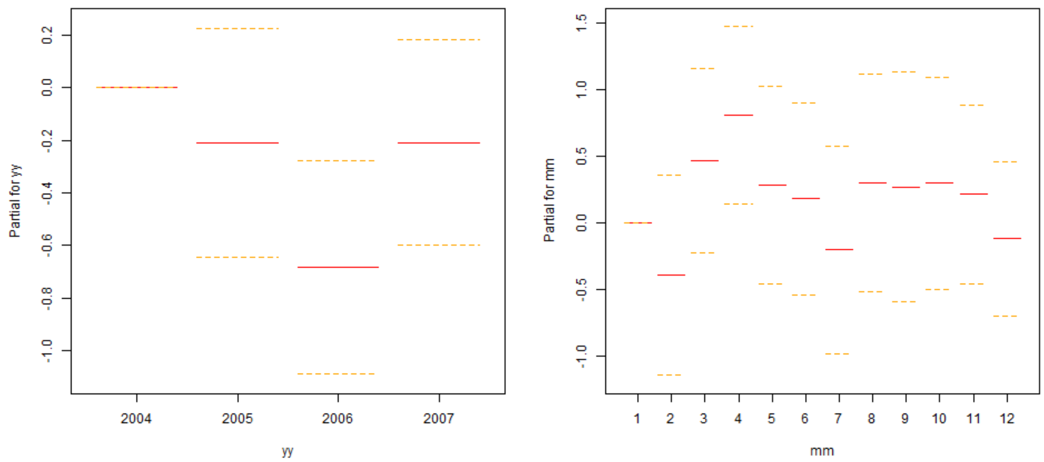


Figure 33: estimates and approximate 95% confidence intervals for parametric terms within the GAM.

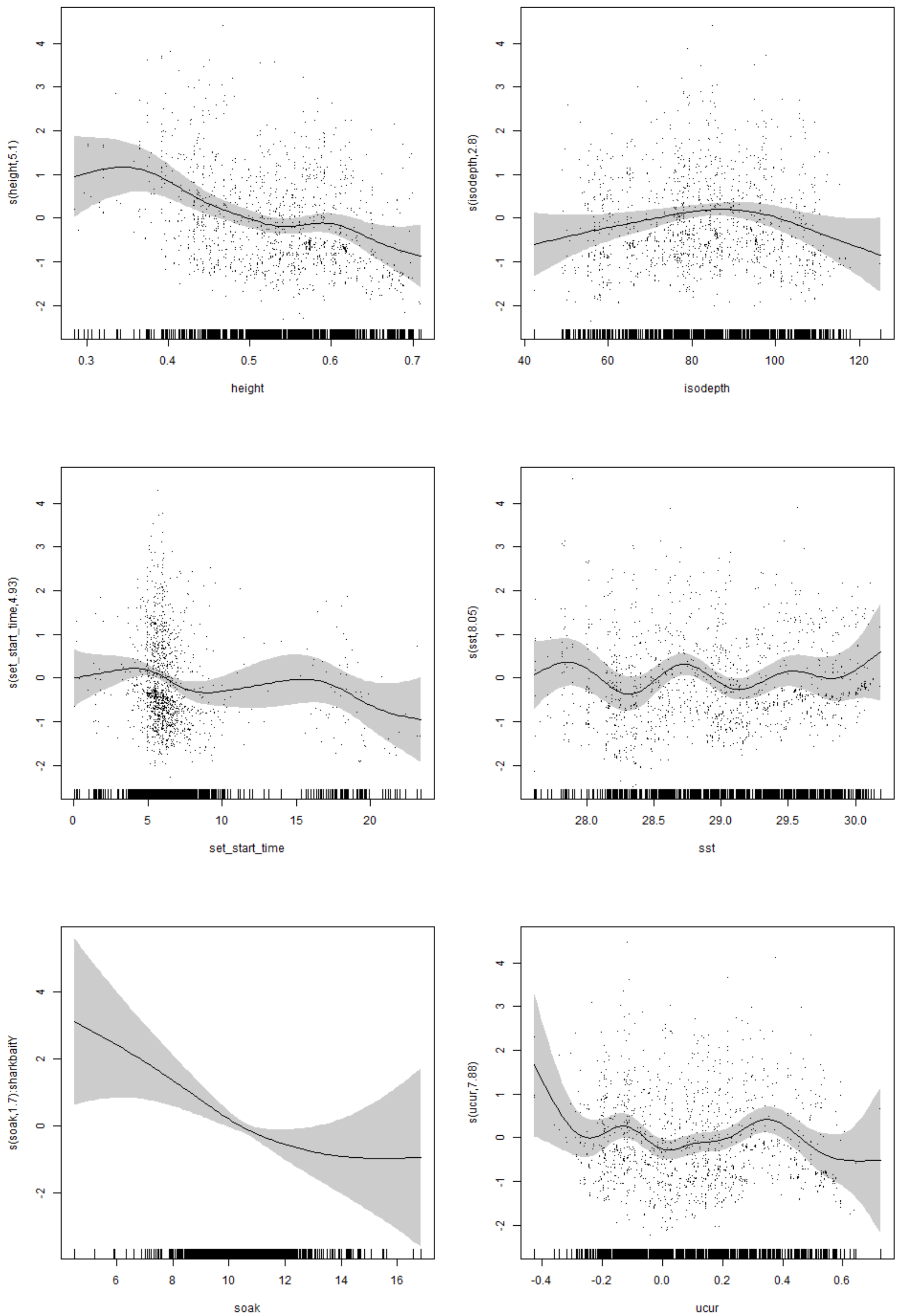


Figure 34: estimated smooths and approximate 95% confidence envelopes for smooth terms within the GAM.

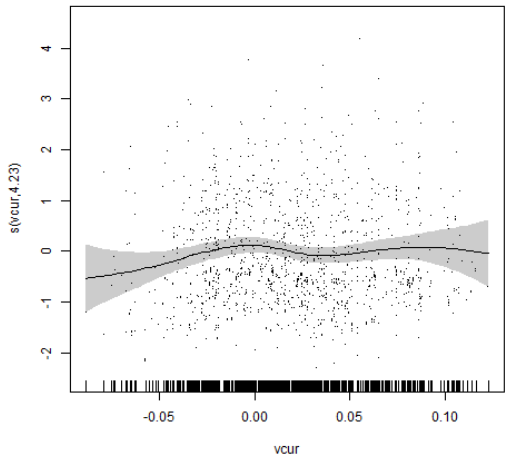


Figure 35: estimated smooths and approximate 95% confidence envelopes for smooth terms within the GAM.

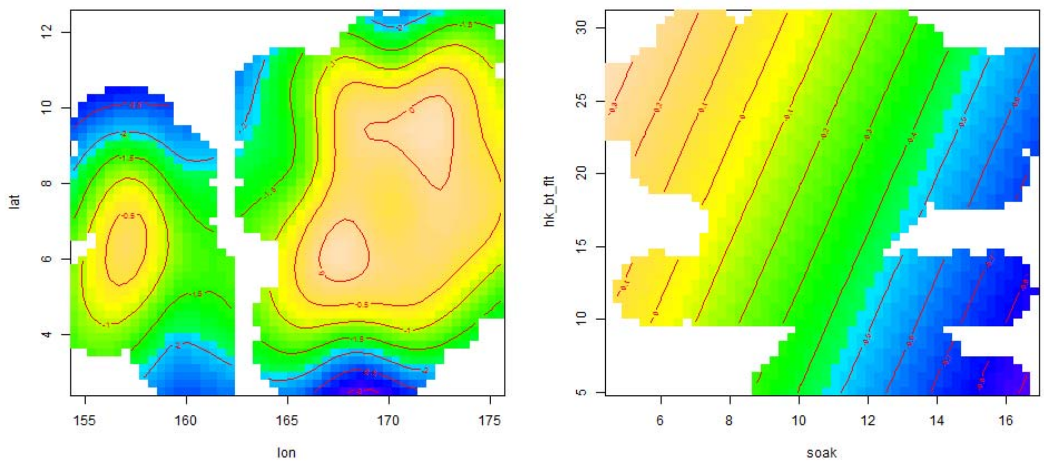


Figure 36: estimated smooth surfaces for the bivariate smooth components within the GAM.

5.6.2 Hook-level catch rate of thresher sharks

5.6.2.1 Data Exploration

There are no clear patterns between catch rates of thresher sharks and hook positions at which they were caught (Figure 37). Furthermore, there are no demarked trends in how the relationship between catch rates and hook positions changes with varying levels of the covariates considered in the present analysis.

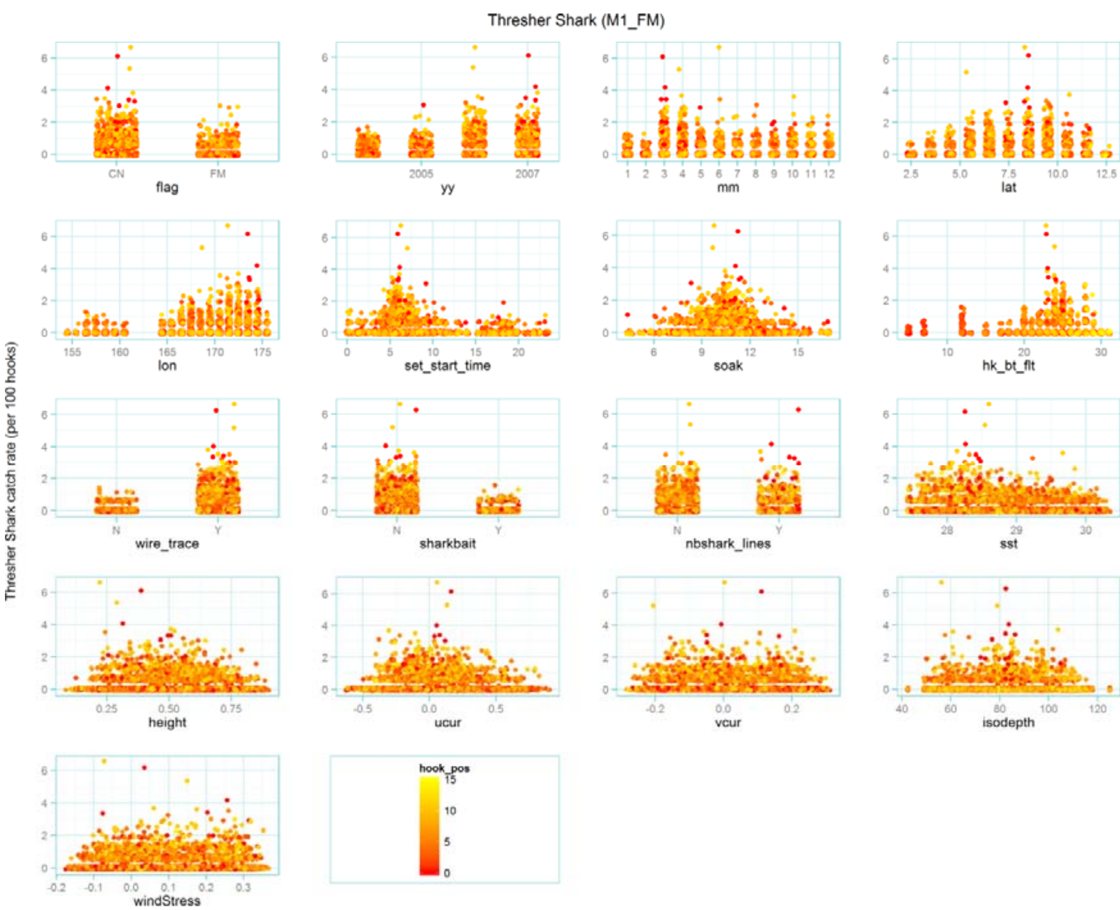


Figure 37: Distributions of per-hook catch-rate responses against the explanatory variables considered in the analysis, for thresher sharks caught in the M1_FM fishery. Dots were slightly jittered to minimise overplotting and help visualisation and their colour convey the basket hook position of the respective observation (file "M1_FM-THR_perHookRatesVsCovs.png").

5.6.2.2 Modelling results

Modelling details are as per the reference case. The model returned an adjusted-R^2 of 0.007 and a deviance explained of 1.98%.

Table 11: Estimates of catch rates for each of the hook positions. Confidence intervals are approximate 95%. Percentage of total catch is calculated from overall estimated catch over hook positions presented. Confidence intervals for percentages are based on this naïve translation of cate rate.

Hook position	Catch rate (per 100 hooks)	SE	Lower CI	Upper CI	Catch (as % of total)	Lower CI (as % of total)	Upper CI (as % of total)
0	0.0436	1.1342	0.0340	0.0557	3.48	2.72	4.45
1	0.0537	1.1202	0.0430	0.0671	4.29	3.44	5.36
2	0.0858	1.0945	0.0719	0.1024	6.85	5.74	8.18
3	0.0942	1.0908	0.0795	0.1117	7.53	6.35	8.92
4	0.0949	1.0909	0.0800	0.1125	7.58	6.39	8.99
5	0.1037	1.0868	0.0881	0.1221	8.28	7.04	9.75
6	0.0979	1.0904	0.0826	0.1160	7.82	6.6	9.27
7	0.1044	1.0876	0.0886	0.1231	8.34	7.07	9.83
8	0.0873	1.0964	0.0729	0.1046	6.97	5.82	8.35
9	0.0861	1.0979	0.0717	0.1034	6.88	5.73	8.26
10	0.1149	1.0873	0.0975	0.1354	9.18	7.79	10.81
11	0.1665	1.0757	0.1443	0.1921	13.3	11.53	15.34
12	0.1190	1.1129	0.0965	0.1468	9.51	7.71	11.73

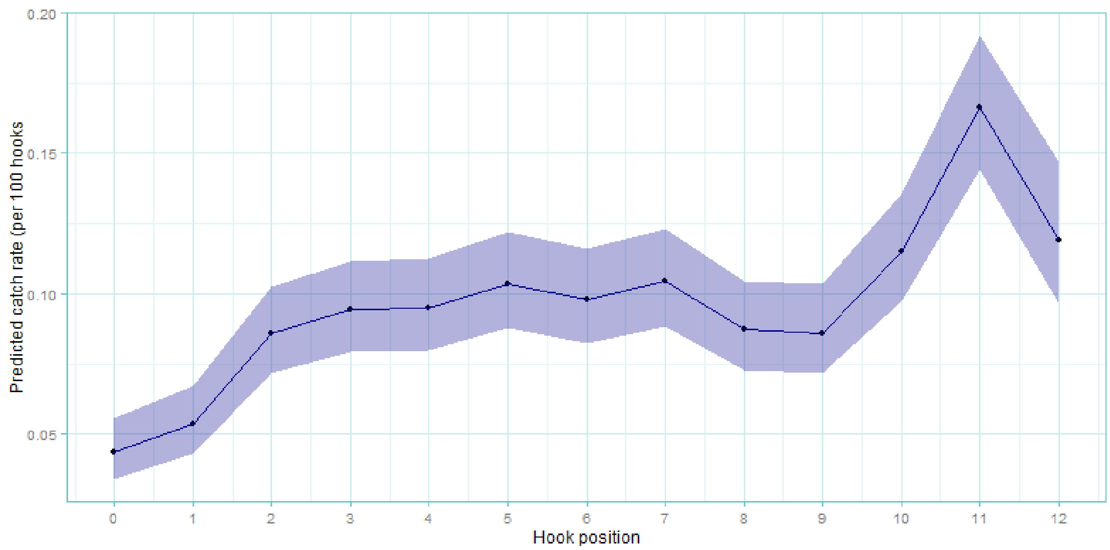


Figure 38: estimated mean catch rate (per 100 hooks) at hook-position with approximate 95% confidence envelope.

5.6.3 Condition of thresher sharks at time of retrieval

5.6.3.1 Data Exploration

There are no strong patterns between the condition of caught silky sharks at the time of retrieval and most of the covariates considered in the present analysis (Figure 39). There are however some indications that the proportion of dead sharks might have increased when observed sets were performed under certain conditions, like e.g.:

- When sets were carried out in 2006 and 2007;
- When longlines were set within 5-11 degrees of latitude;
- If average sea surface height was slightly above 0.3 meters.

In addition, it is worth noting that the overall proportion of dead thresher sharks at hauling time is consistently larger (around 50%) than other species under identical conditions (around 25% in oceanic whitetip, Section 0; around 20% in silky sharks, Section 5.5.3; and around 15% in blue sharks, Section 5.7.3).

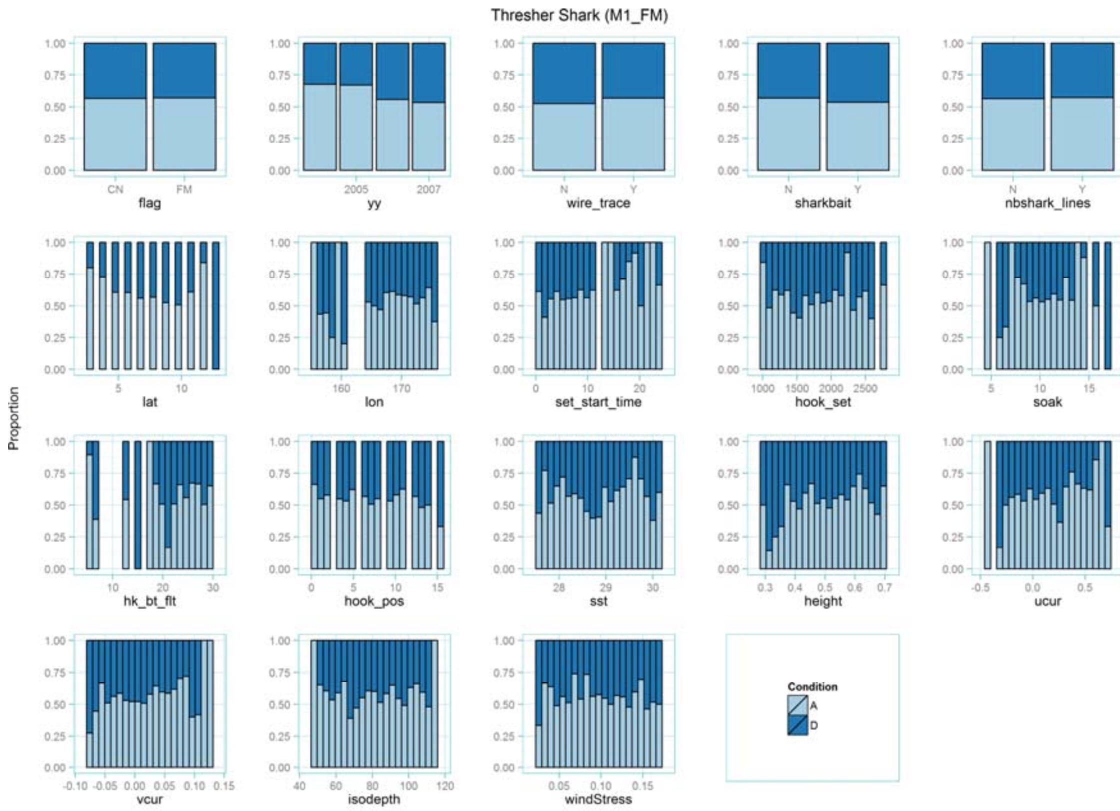


Figure 39: Distributions of condition responses (D= dead, A = alive), expressed in terms of proportion, against the explanatory variables considered for the analysis of thresher sharks caught in the M1_FM fishery (file "M1_FM-THR_CondtnVsCovs.png")

5.6.3.2 Modelling results

Modelling details are as per the reference case. The model returned an adjusted-R^2 of 0.095 and a deviance explained of 9.64%. Raw predictive power to the current dataset was also assessed via confusion matrices on the training data, providing Sensitivity (true positive rate) = 0.69, Specificity (true negative rate) = 0.59 and a misclassification rate of 0.37 when using the response mean as a decision boundary. These are against the training data and will be over-estimates.

Table 12: Model terms from the generalized additive model. DF = Degrees of Freedom. EDF=Effective Degrees of Freedom. Shrinkage has been applied as a means of model selection. Terms of little relevance have their EDF shrunken.

Model terms	EDF	DF	Chi or F-stat	*p*-value
yy		3	15.52	0.0014
flag		1	0.92	0.3380
sharkbait		1	0.23	0.6340
nbshark_lines		1	7.83	0.0052
wire_trace		1	1.25	0.2638
s(set_start_time)	1.71	9	8.11	0.0038
s(sst)	7.09	9	45.63	0.0000
s(height)	0.00	9	0.00	0.5512
s(ucur)	3.98	9	5.39	0.1830
s(vcur)	3.88	9	9.02	0.0270
s(isodepth)	0.65	9	1.36	0.1296
s(windStress)	3.47	9	6.37	0.0817
s(lon,lat)	19.47	29	69.99	0.0000
s(soak,hk_bt_flt)	19.93	29	47.93	0.0000

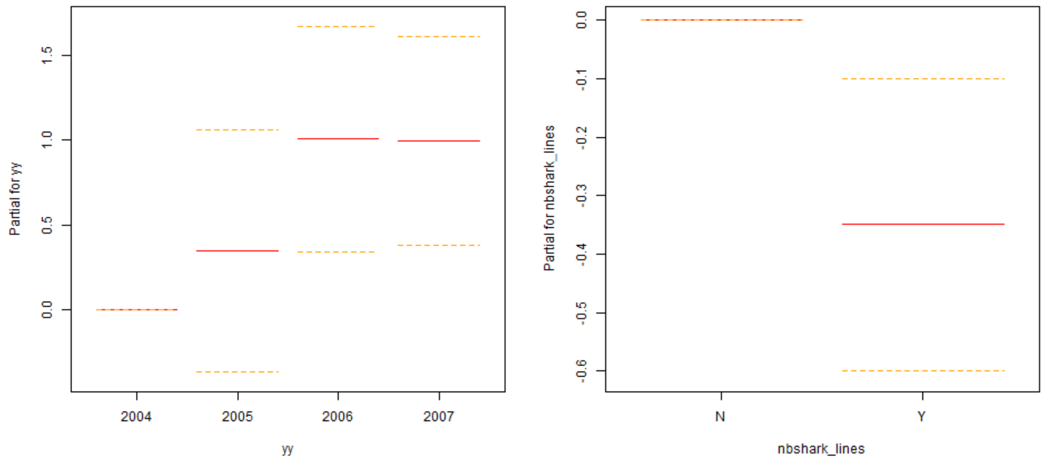


Figure 40: estimates and approximate 95% confidence intervals for parametric terms within the GAM.

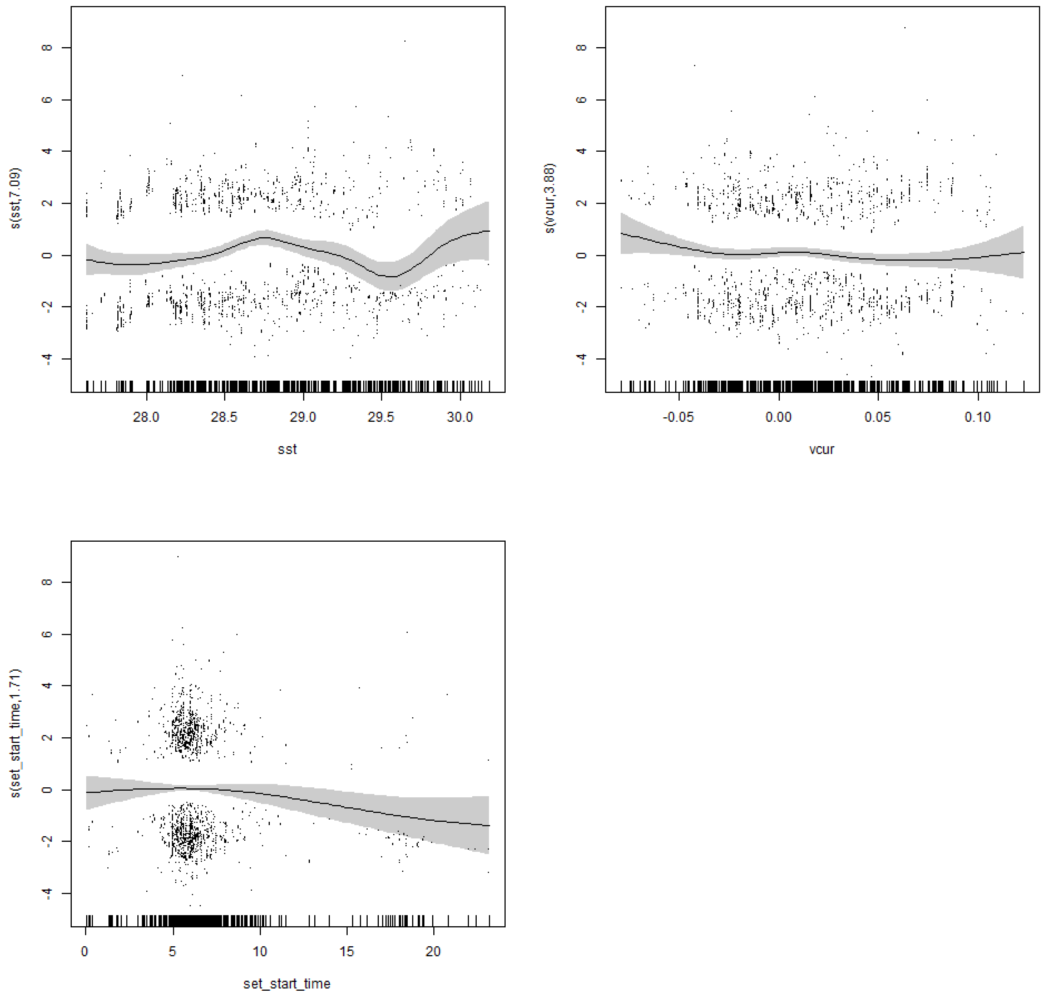


Figure 41: estimated smooths and approximate 95% confidence envelopes for smooth terms within the GAM.

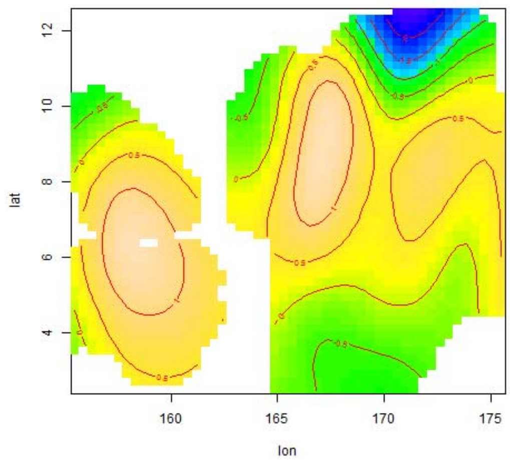


Figure 42: estimated smooth surfaces for the bivariate smooth components within the GAM.

5.7 CN & FM COMBINED FLEET ANALYSIS – BLUE SHARKS (BSH)

5.7.1 Catch rate of Blue Sharks (BSH)

5.7.1.1 Data Exploration

While mostly absent in sets from observed trips in the M1_FM fishery, within-set bycatch levels of blue sharks decay gradually from 1 shark/set (in about 340 set) to 12 sharks/set (in about 5 sets), with one of the sets catching the maximum of 19 blue sharks (Figure 43).

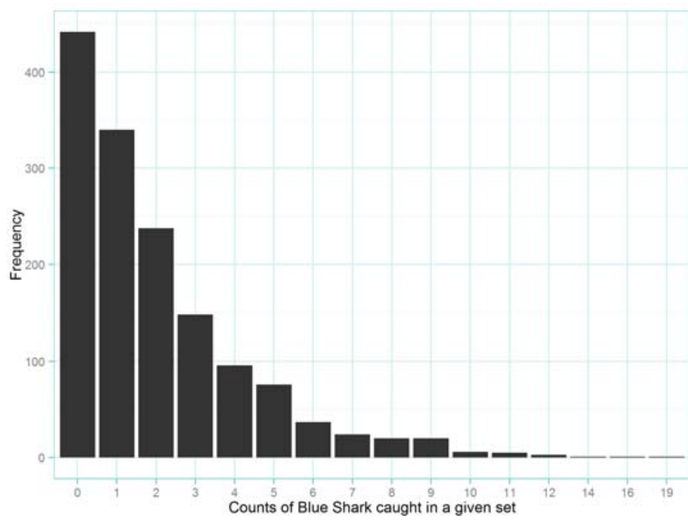


Figure 43: Distribution of counts per set of blue shark caught in observed trips of the M1_FM longline fishery (M1_FM-BSH_CatchNumFreqs.png)

Figure 44 suggests some extent of increasing catch rates of blue sharks when longlines were set under certain conditions, e.g.:

- When observed trips took place in 2004 and 2007 and in April or June;
- If wire leader and shark lines were in use, and shark bait was not used
- as sets were performed at northerly locations (i.e. with increasing degrees of latitude)
- as values of average sea surface temperature and average sea surface height reached their lower end values;
- average horizontal current speed (ucur) ranged between -0.25 and -0.1 m/s;
- when average wind stress levels exceeded 0.08 Newton/m^2.

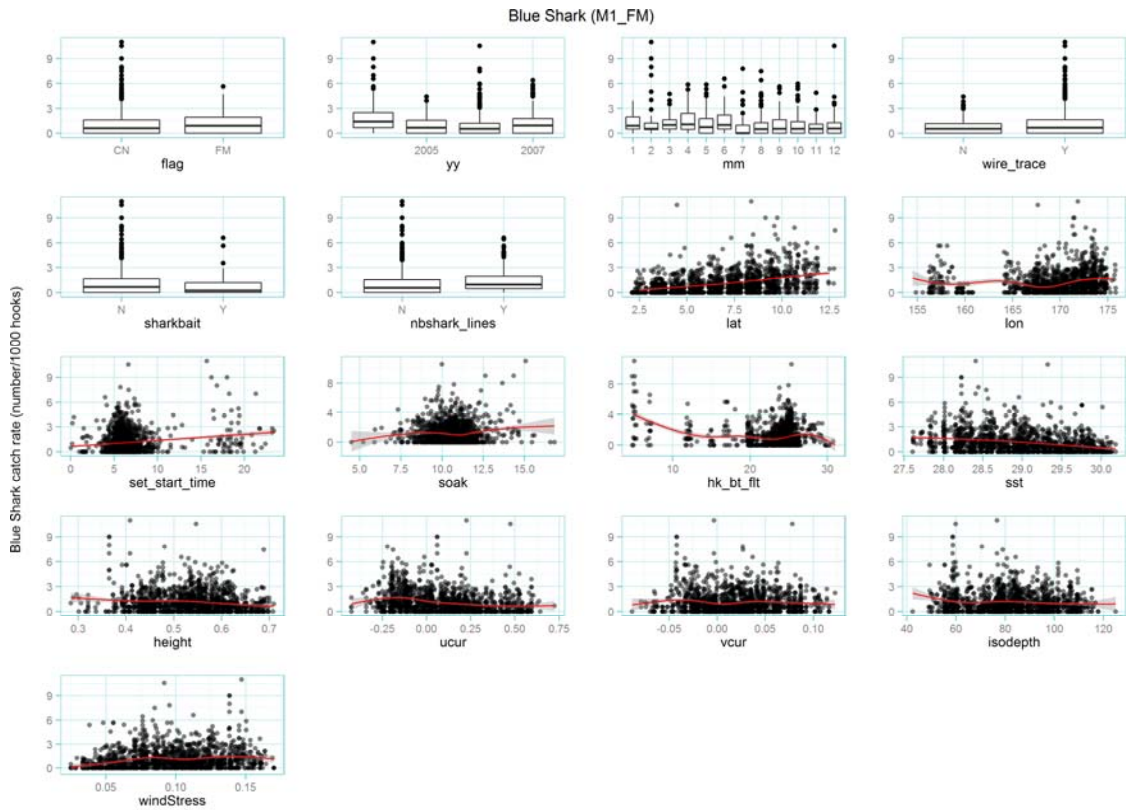


Figure 44: Distributions of set-level catch-rate responses against the explanatory variables considered for the analysis of blue sharks caught in the M1_FM fishery (file "M1_FM-BSH_CatchRatesVsCovs.png")

5.7.1.2 Modelling results

Modelling details are as per the reference case. The model returned an adjusted-R^2 of 0.365 and a deviance explained of 38.9%.

Table 13: Model terms from the generalized additive model. DF = Degrees of Freedom. EDF=Effective Degrees of Freedom. Shrinkage has been applied as a means of model selection. Terms of little relevance have their EDF shrunken.

Model terms	EDF	DF	Chi or F-stat	*p*-value
yy		3	13.82	0.000
mm		11	2.47	0.005
wire_trace		1	0.49	0.483
nbshark_lines		1	1.87	0.172
sharkbait		1	11.78	0.001
flag		1	0.26	0.610
s(set_start_time)	0.00	9	0.00	0.757
s(sst)	6.50	9	2.69	0.000
s(height)	8.29	9	3.37	0.000
s(ucur)	8.69	9	3.53	0.000
s(vcur)	0.00	9	0.00	1.000
s(isodepth)	1.97	9	1.31	0.000
s(windStress)	5.50	9	2.40	0.000
s(lon,lat)	17.54	29	2.51	0.000
s(soak,hk_bt_flt)	12.32	29	1.31	0.000

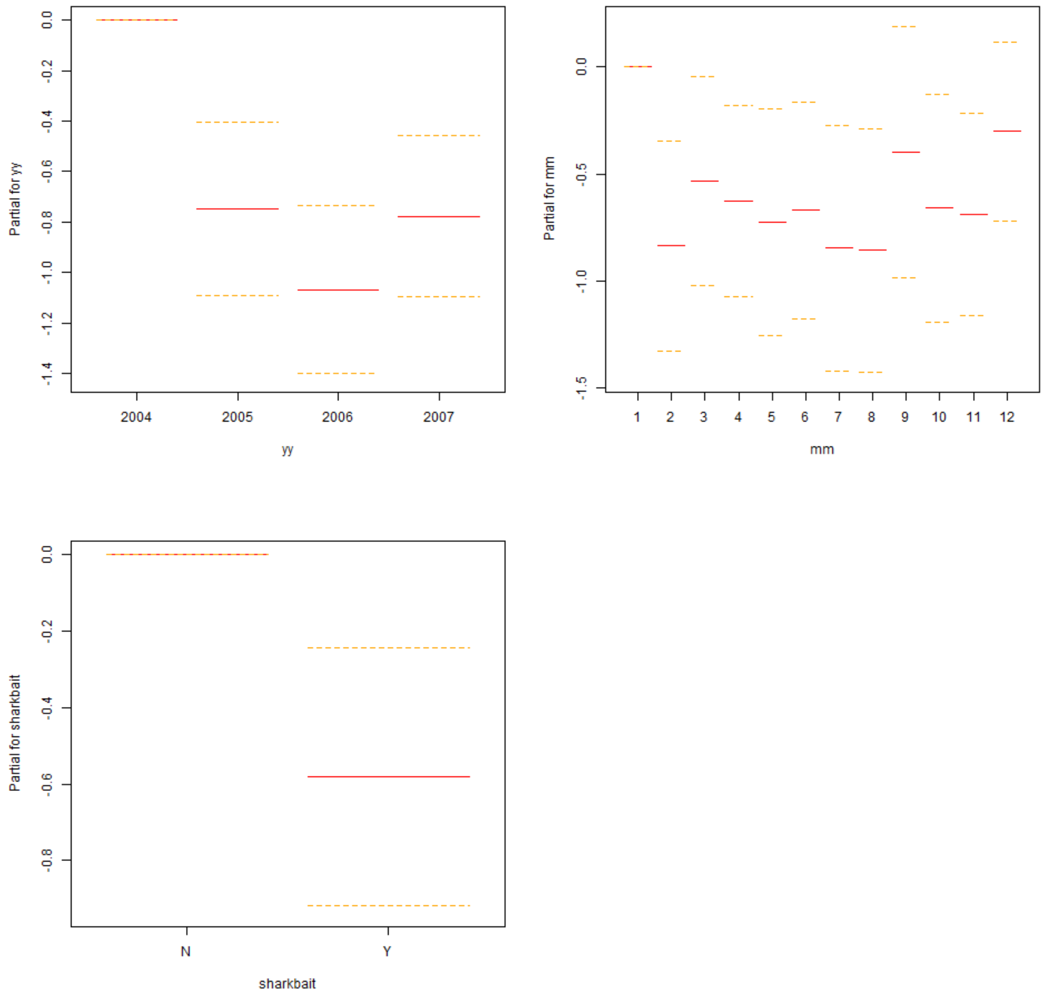


Figure 45: estimates and approximate 95% confidence intervals for parametric terms within the GAM.

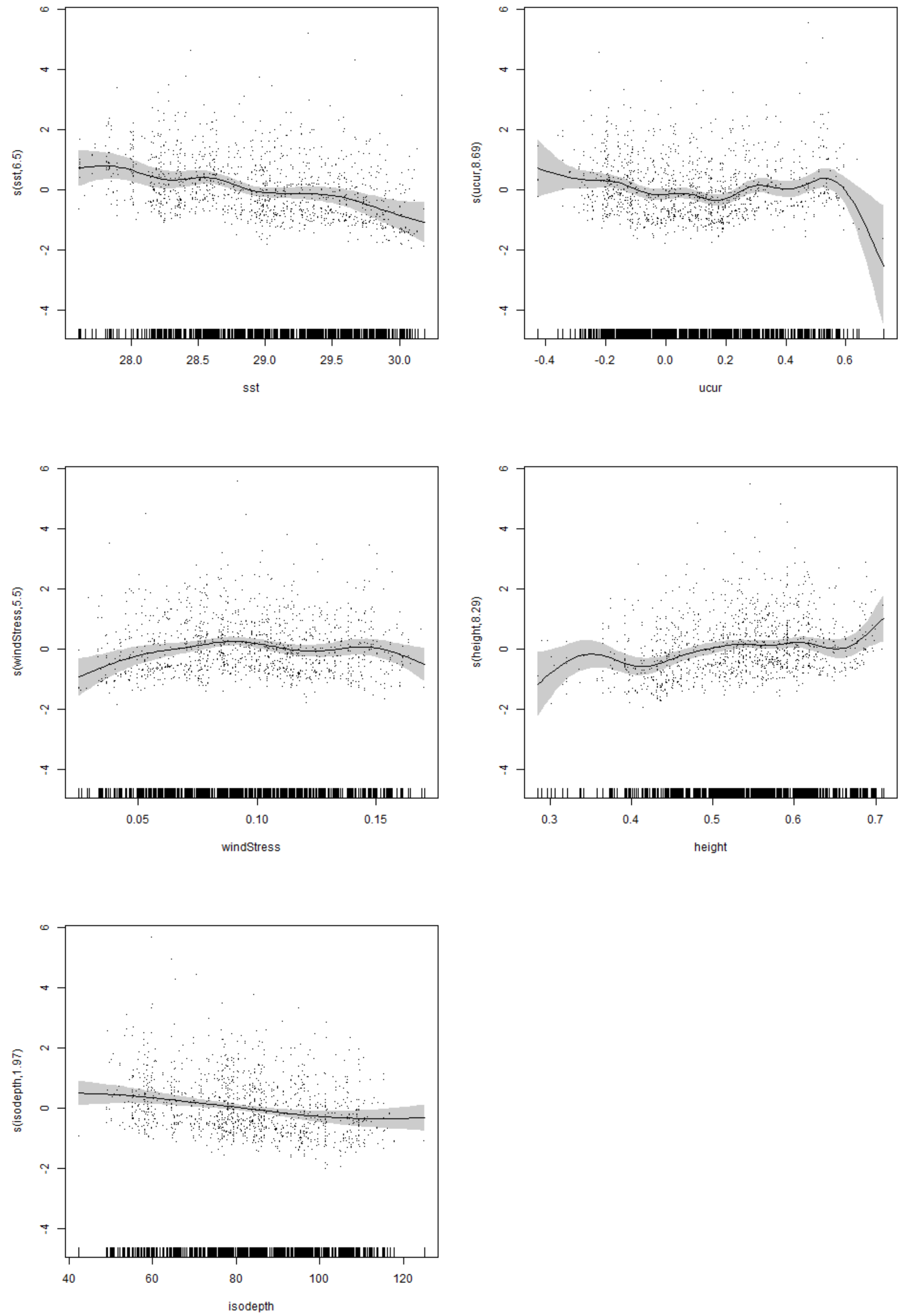


Figure 46: estimated smooths and approximate 95% confidence envelopes for smooth terms within the GAM.

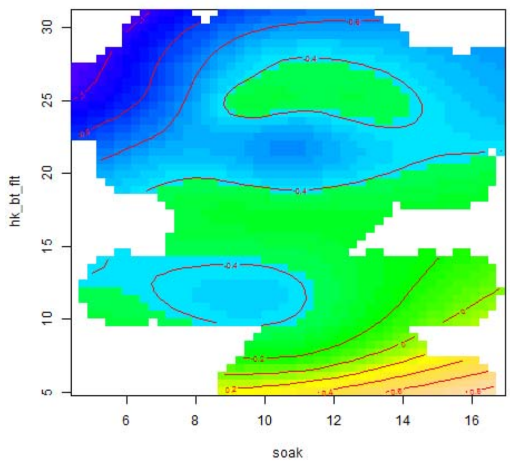


Figure 47: estimated smooth surfaces for the bivariate smooth components within the GAM.

5.7.2 Hook-level catch rate of blue sharks

5.7.2.1 Data Exploration

There are no clear patterns between catch rates of blue sharks and hook positions at which they were caught (Figure 48). Furthermore, there are no demarked trends in how the relationship between catch rates and hook positions changes with varying levels of the covariates considered in the present analysis.

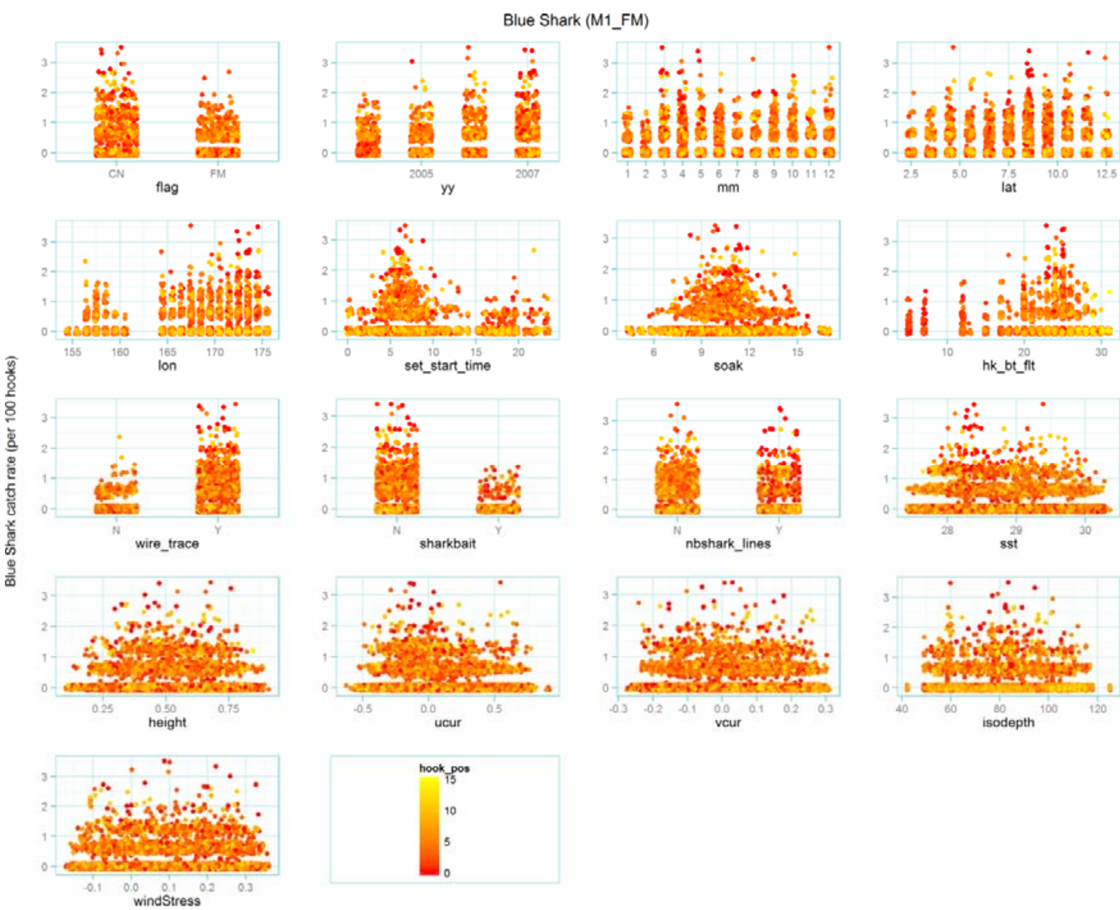


Figure 48: Distributions of per-hook catch-rate responses against the explanatory variables considered in the analysis, for blue sharks caught in the M1_FM fishery. Dots were slightly jittered to minimise overplotting and help visualisation and their colour convey the basket hook position of the respective observation (file "M1_FM-OCS_perHookRatesVsCovs.png").

5.7.2.2 Modelling results

Modelling details are as per the reference case. The model returned an adjusted-R^2 of 0.006 and a deviance explained of 1.2%.

Table 14: Estimates of catch rates for each of the hook positions. Confidence intervals are approximate 95%. Percentage of total catch is calculated from overall estimated catch over hook positions presented. Confidence intervals for percentages are based on this naïve translation of cate rate.

Hook position	Catch rate (per 100 hooks)	SE	Lower CI	Upper CI	Catch (as % of total)	Lower CI (as % of total)	Upper CI (as % of total)
0	0.0985	1.0799	0.0848	0.1146	7.27	6.25	8.45
1	0.1218	1.0725	0.1062	0.1397	8.98	7.83	10.3
2	0.1354	1.0690	0.1188	0.1544	9.99	8.76	11.38
3	0.1394	1.0686	0.1224	0.1588	10.28	9.03	11.71
4	0.1263	1.0722	0.1102	0.1448	9.31	8.13	10.68
5	0.1108	1.0767	0.0958	0.1280	8.17	7.07	9.44
6	0.1040	1.0798	0.0895	0.1209	7.67	6.6	8.92
7	0.1024	1.0805	0.0880	0.1191	7.55	6.49	8.79
8	0.0726	1.0948	0.0608	0.0867	5.35	4.48	6.39
9	0.0678	1.0987	0.0564	0.0815	5	4.16	6.01
10	0.0693	1.1010	0.0573	0.0836	5.11	4.23	6.17
11	0.1029	1.0880	0.0872	0.1214	7.59	6.43	8.95
12	0.1049	1.1097	0.0855	0.1286	7.73	6.31	9.48

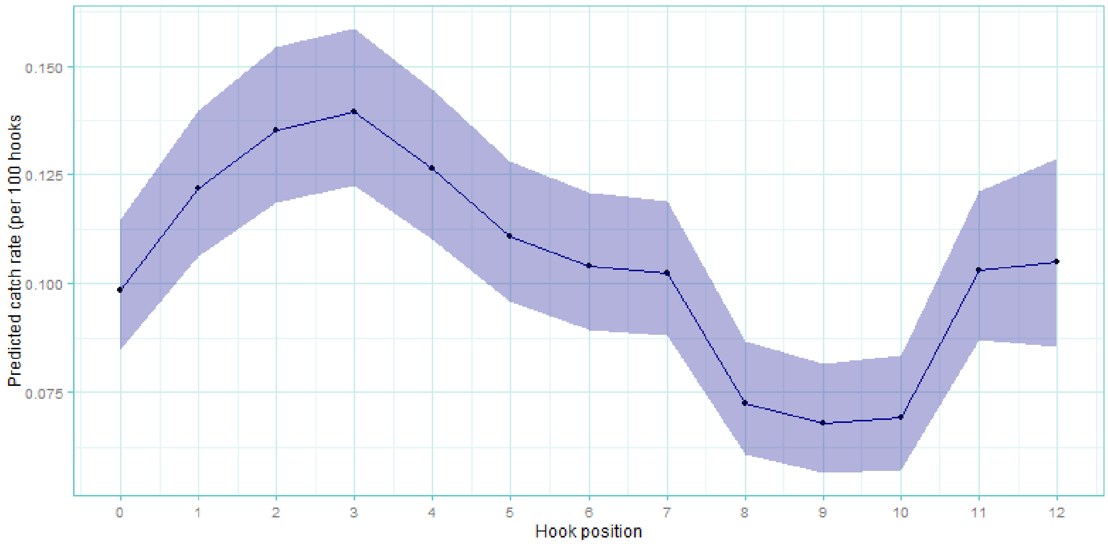


Figure 49: estimated mean catch rate (per 100 hooks) at hook-position with approximate 95% confidence envelope.

5.7.3 Condition of blue sharks at time of retrieval

5.7.3.1 Data Exploration

There are no strong patterns between the condition of caught blue sharks at the time of retrieval and most of the covariates considered in the present analysis (Figure 50). There are however some indications that the proportion of dead blue sharks sharks might have increased when observed sets were performed under certain conditions, like e.g. (Figure 50):

- When sets were performed by Chinese vessels and carried out in 2006;
- When wire leader and shark lines were in use, and when no shark bait was used;
- If longlines were set within 9-11 degrees of latitude;
- When sets were released between 12:00-14:00 hours;
- If average sea temperature was below 28.5^{o}C and surface height was below 0.4 meters.

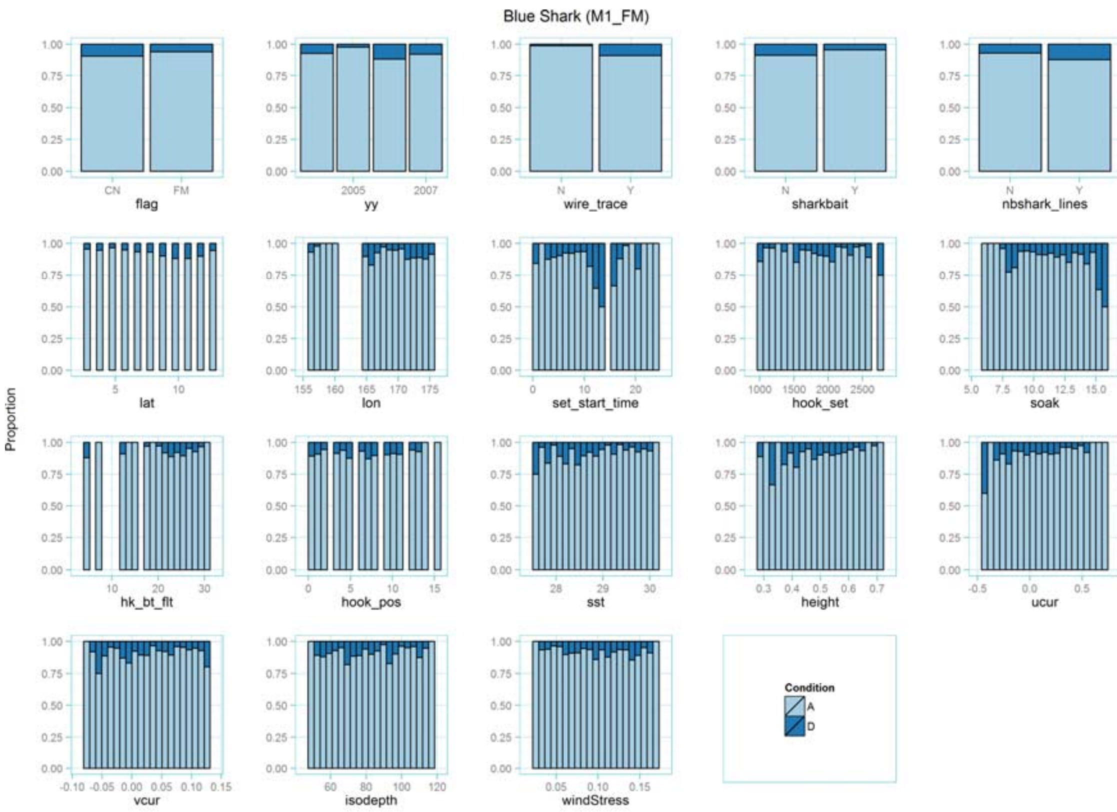


Figure 50: Distributions of condition responses (D= dead, A = alive), expressed in terms of proportion, against the explanatory variables considered for the analysis of blue sharks caught in the M1_FM fishery (file "M1_FM-BSH_CondtnVsCovs.png")

5.7.3.2 Modelling results

Modelling details are as per the reference case. The model returned an adjusted-R^2 of 0.095 and a deviance explained of 14.8%. Raw predictive power to the current dataset was also assessed via confusion matrices on the training data, providing Sensitivity (true positive rate) = 0.75, Specificity (true negative rate) = 0.66 and a misclassification rate of 0.34 when using the response mean as a decision boundary. These are against the training data and will be over-estimates.

Table 15: Model terms from the generalized additive model. DF = Degrees of Freedom. EDF=Effective Degrees of Freedom. Shrinkage has been applied as a means of model selection. Terms of little relevance have their EDF shrunken.

Model terms	EDF	DF	Chi or F-stat	*p*-value
yy		3	5.01	0.171

flag		1	0.44	0.505
sharkbait		1	0.53	0.466
nbshark_lines		1	4.49	0.034
wire_trace		1	7.40	0.007
s(set_start_time)	7.20	9	18.86	0.004
s(sst)	4.48	9	22.14	0.000
s(height)	7.00	9	26.43	0.000
s(ucur)	4.00	9	19.82	0.000
s(vcur)	3.85	9	6.01	0.134
s(isodepth)	0.00	9	0.00	0.460
s(windStress)	0.00	9	0.00	0.605
s(lon,lat)	3.06	29	7.08	0.028
s(soak,hk_bt_flt)	18.70	29	39.56	0.001

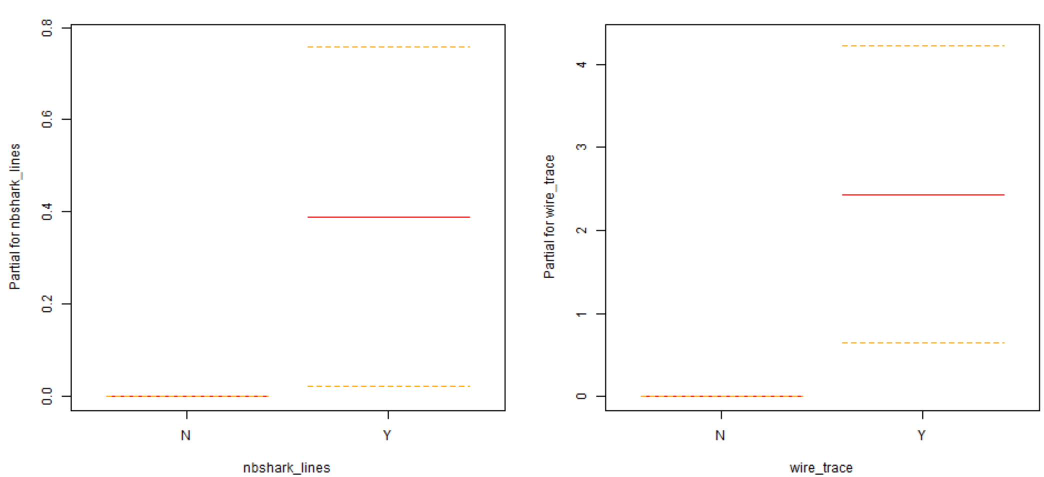


Figure 51: estimates and approximate 95% confidence intervals for parametric terms within the GAM

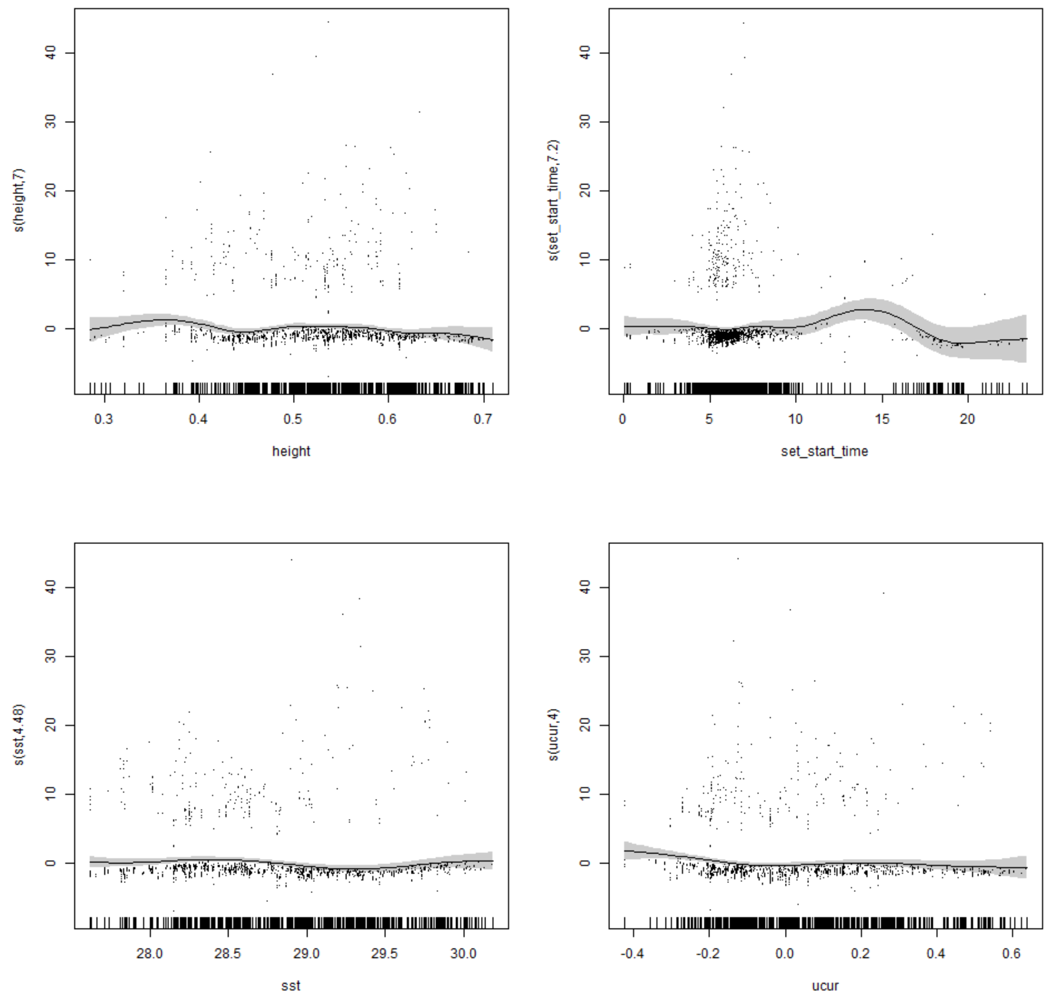


Figure 52: estimated smooths and approximate 95% confidence envelopes for smooth terms within the GAM.

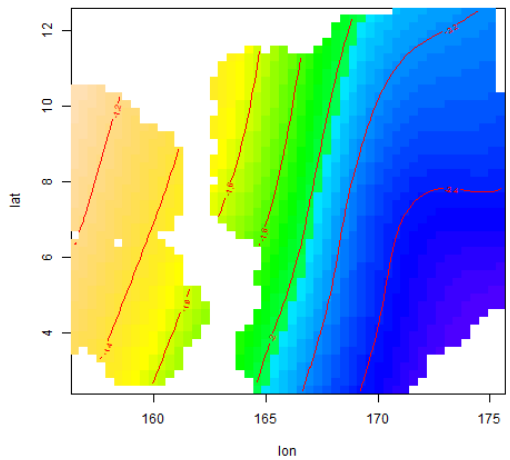


Figure 53: estimated smooth surfaces for the bivariate smooth components within the GAM.

6 FIJI AND AMERICAN SAMOA FISHERY (M2_FV)

6.1 FISHERY DATA DESCRIPTION

Observer data from the Fiji and American Samoa (M2_FV) fishery comprises trips carried out by the Fijian fleet (while operating mainly in the EEZs of Fiji, Solomon Islands and Vanuatu) and the United States fleet (operating exclusively within the EEZ of American Samoa). Observed trips from the US fleets are only available for the period 2006-2011, sensibly during which the total number of observed trips in the fishery increased to the highest levels in the time-series (Figure 54). The number of observed trips dropped substantially in last two years. Observer data in this fishery seem to cover well the whole duration of the year, with no demarked changes in number of sets between months.

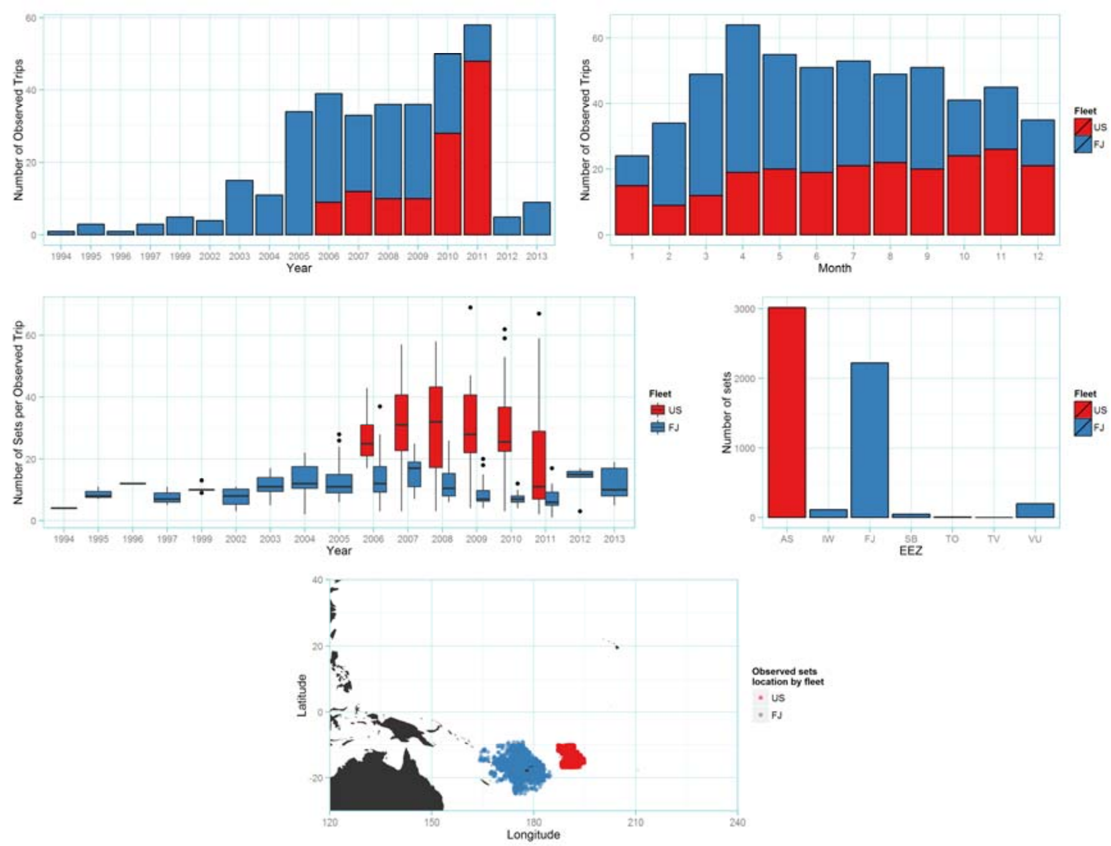


Figure 54: Temporal and spatial distribution of fishing effort by fleet in observed trips of the Fiji and American Samoa fishery (file "M2_FV_tripsAndSetsDesc.png").

Figure 54 also shows relevant differences between the two fleets operating in the fishery in terms of fishing strategies. In particular, the US fleet performs a larger number of sets per trip than the Fijian and, in geographical terms, the observed trips from each fleet operated in completely separated areas of the fishery. Furthermore, observed trips of Fijian fleets have used wire leader, shark lines and shark bait in their sets for extended periods of the time-series (generally during 2003-2013), while none of those key gear features were applied in observed US vessels (Figure 55). There are also differences in the types of hooks used by each fleet – while the Fijian fleet used hook types “C” and “J”, the hook type “C” was extensively used by the US fleet together with hook type “T”, with the former hook type being only employed in 2010 and 2011 (Figure 55).

To avoid the potential occurrence of artefacts in the results due to the underlining discrepancies in fishing strategies employed by the two fleets, the subsequent analysis of shark byctach in the M2_FV fishery was carried out separately for each fleet.

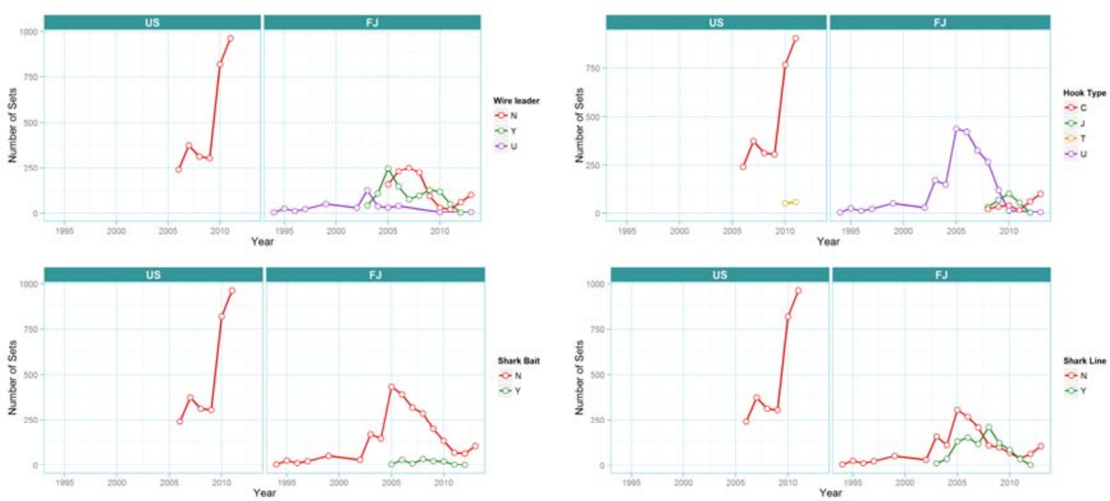


Figure 55: Contrast of key gear types over time in observed trips by fleets operating in the Fiji and American Samoa fishery (File “M2_FV_SampSizeAndKeyFactors.png”).

In general, blue sharks and silky sharks were the species with higher within-year bycath in the time-series of observed trips from both fleets, followed by oceanic whitetip sharks, mako sharks and thresher sharks (Figure 56). For example, observed US flagged trips caught around 5000 blue sharks and over 2000 silky sharks in 2011, which was double the bycatch of mako, oceanic whitetip and thresher sharks in the same trips. With fewer observed trips, the Fijian trips caught around 400 blue sharks and a similar number of silky sharks in 2009, followed by about 200 specimens of each of mako and oceanic whitetip species (Figure 56).

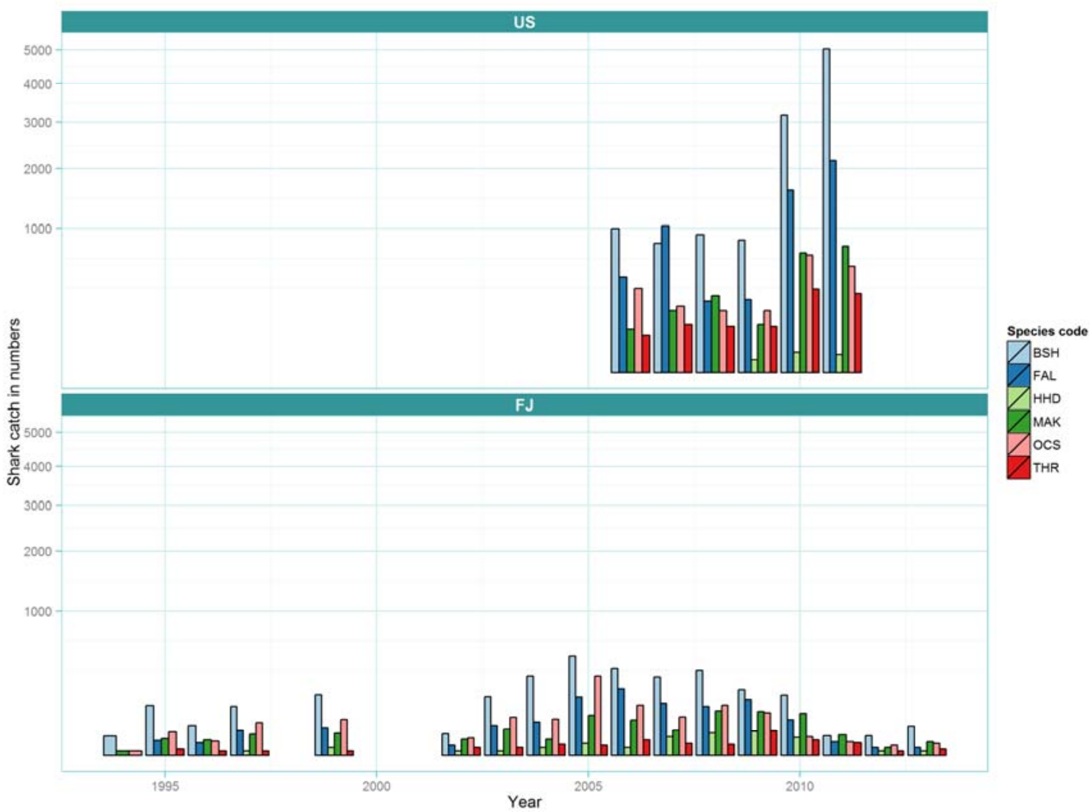


Figure 56: Bycatch of shark species over time in observed trips for fleets operating in the Fiji and American Samoa longline fishery (file "M2_FV_catchNumSpYrFlg.png").

6.2 FIJIAN FLEET ANALYSIS – EXPLANATORY VARIABLES EDA

Despite the time-series of the original dataset spanning from years 1994-2013, the absence of information about the type of hook and leader in use by the Fijian fleet until year 2007 (Figure 55) meant the analysis had to be restricted to data starting from 2008. There was however good overall contrast between the usage of the key fishing features (i.e. wire leader, hook type, shark bait and shark lines) in sets observed during the period 2008-2012.

There were a few observations with large values of soak time (over 20 hours) that were considered too extreme in relation to the normal range of soak times in the dataset – these observations were excluded from the analysis (Figure 57). There were no other extreme values in the remaining covariates.

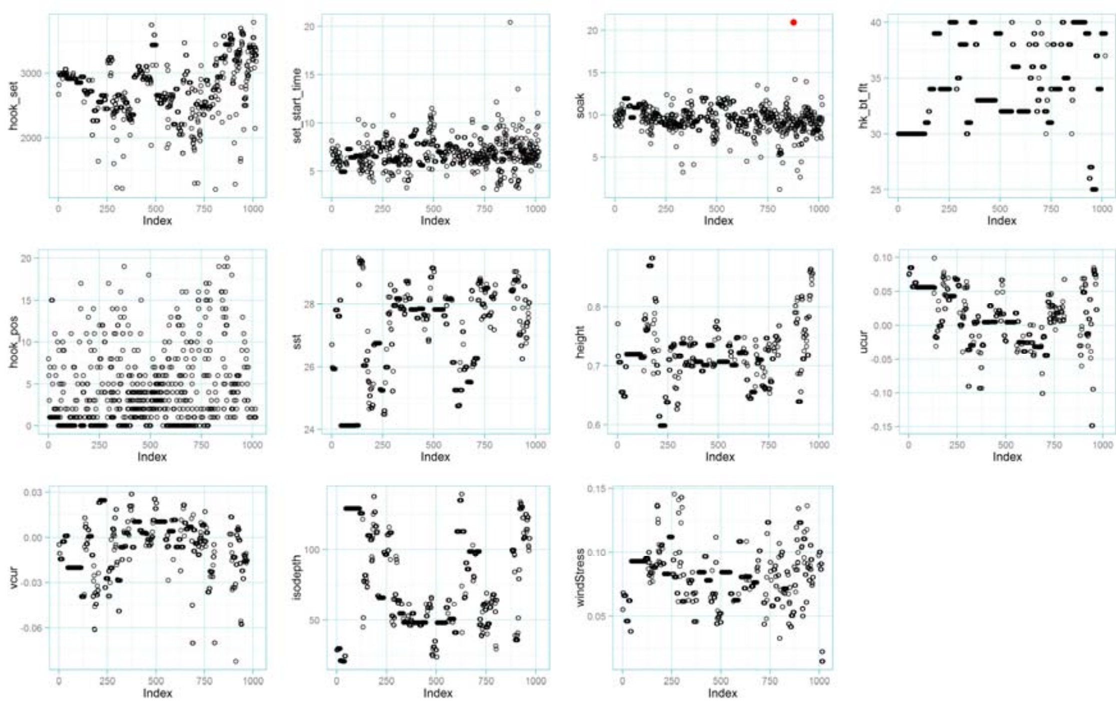


Figure 57: Cleveland dotplots of some of the explanatory variables considered under the analysis of shark bycatch by the Fijian fleet in the Fiji and American Samoa fishery. Plots constructed from data sorted by set starting date, and values deemed to be too extreme for the normal range of the covariate, and posteriorly excluded from the analysis, are coloured in red (File "M2_FV_FJ_Catch_covsClevPlots.png").

Figure 58 shows the existence of mild negative correlation (-0.74) between average sea surface temperature and average depth of the 20^oC isotherm, which can translate into a strong collinearity between these two covariates – collinearity was further investigated through the estimation of VIFs. The remaining covariates from observations used for the analysis of the Fijian fleet in the M2_FV fishery show weak correlation between them.

Furthermore, several of the oceanographic variables (e.g. sea surface temperature, sea surface height and wind stress) appear to have non-linear relationships with month and latitude (Figure 58). Changes in the average depth of the 20^oC isotherm also appears to be followed by changes in average sea surface height and in horizontal (ucur) and vertical (vcur) current speeds.

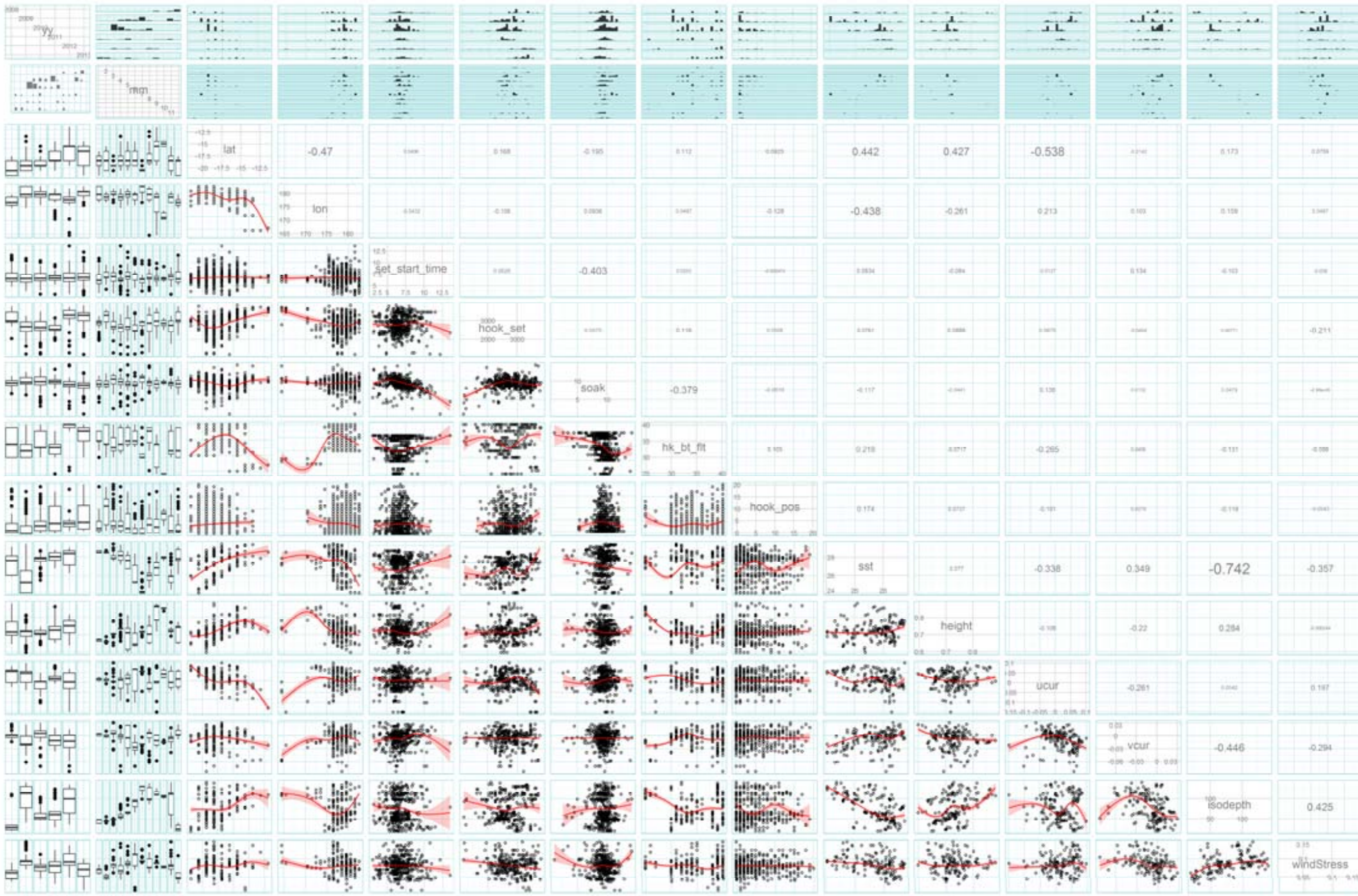


Figure 58: Matrixplot of relevant covariates considered for the analysis of shark bycatch by the Fijian fleet in the Fiji and American Samoa fishery. Plots presented in panels depend on the data types under comparison. For example, pairs of continuous variables are displayed as scatterplots in the lower panels (with an added GAM-based smoother to help visualisation) and Pearson correlation

coefficients in the upper panels (with font size proportional to correlation coefficient). Continuous Vs Discrete covariates are represented by boxplots and barplots, while pairs of discrete covariates are displayed as fluctuation plots (lower panels) and barplots (upper panels) (file "M2_FV_FJ_Catch_CovsPairsPlot.png")

6.3 FIJIAN FLEET ANALYSIS – MODELLING

Presented here are results for three models covering set-level catch-rates, catch-rates with respect to hook position and the condition of sharks (alive/dead) at retrieval. Four shark species are considered: Oceanic whitetips (OCS), silky (FAL), thresher (THR) and blue (BSH).

Details regarding modelling methods are presented previously, but all models fall within the Generalized Additive Modelling framework. The set of parametric and smoothing terms selected between are given in the following table – the specific model fitted for each species and model is presented in the relevant results section.

Oceanic variables were not used in models for the Fijian flagged data. A substantial proportion of catch information did not have associated oceanic data values. Two options arise: a) inclusion of oceanic data at the loss of 20-30% of catch data, or b) exclusion of oceanic data in order to use 100% of the catch information. Option b) was favoured as it permitted more data for the estimation of parameters relating to the key gear variables. Spatial smoothers were still included in the modelling mix, which may act as a proxy for much of the excluded oceanic information.

Table 16: Model terms considered. “s” indicates a smoothing term. “:” indicates an interaction between associated terms.

Parametric terms	Smooth terms
yy	s(set_start_time)
mm	s(sst)
hook_type	s(height)
wire_trace	s(ucur)
nbshark_lines	s(vcur)
hook_type:wire_trace	s(isodepth)
hook_type:nbshark_lines	s(windStress)
wire_trace:nbshark_lines	s(lat,lon)
hook_type:wire_trace:nbshark_lines	s(soak,hk_bt_flt)
	s(soak):nbshark_lines
	s(soak):sharkbait

Reference model – In the interests of brevity, one species' models are presented in more detail than the others. This will be referred to as the *reference case* and will be the Oceanic whitetip shark (OCS). In particular model diagnostics and non-significant results will be presented for this model, but not others. The same modelling approach has been applied for similar models, but only significant results (including assumption violations) will be reported upon.

6.4 FIJIAN FLEET ANALYSIS – OCEANIC WHITETIP SHARKS (OCS)

6.4.1 Catch rate of oceanic whitetip sharks [reference case]

6.4.1.1 Data Exploration

While predominantly absent in sets from observed trips of the Fijian fleet in the M2_FV fishery, within-set bycatch levels of oceanic whitetip sharks ranged mainly between 1 shark/set (in about 20 sets) and 2 sharks/set (in about 10 sets) (Figure 59). However, in a few instances, up to 8 and 15 oceanic whitetips were observed in a single set.

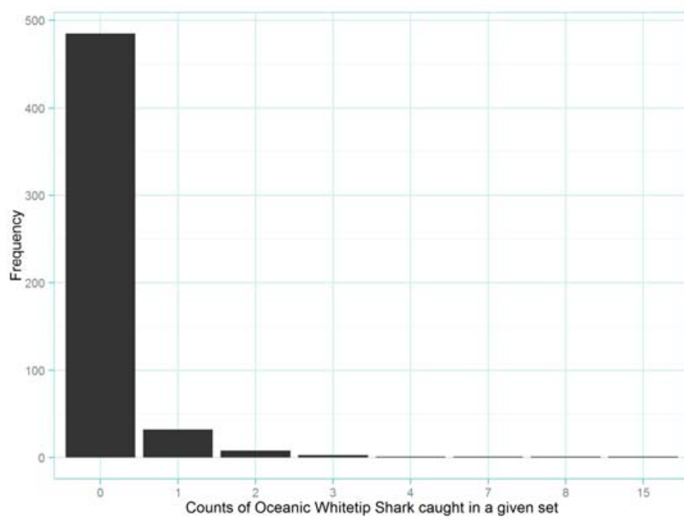


Figure 59: Distribution of counts per set of oceanic whitetip sharks caught in observed trips of the Fijian fleet operating in M2_FV longline fishery (file M2_FV_FJ_OCS_CatchNumFreqs)

The predominance of zero catches in observed sets hampers the investigation of bivariate patterns between catch rates of oceanic whitetip sharks and each of the covariates considered in the present analysis (Figure 60). Yet, these plots indicate that, on average:

- The presence of hook type “J”, wire leader, shark bait and shark lines may have led to higher catch rates of this species;
- Catch rates were higher in 2009;

- Higher catch rates appear within -20 and -17.5 degrees latitude, and around 180 degrees longitude.

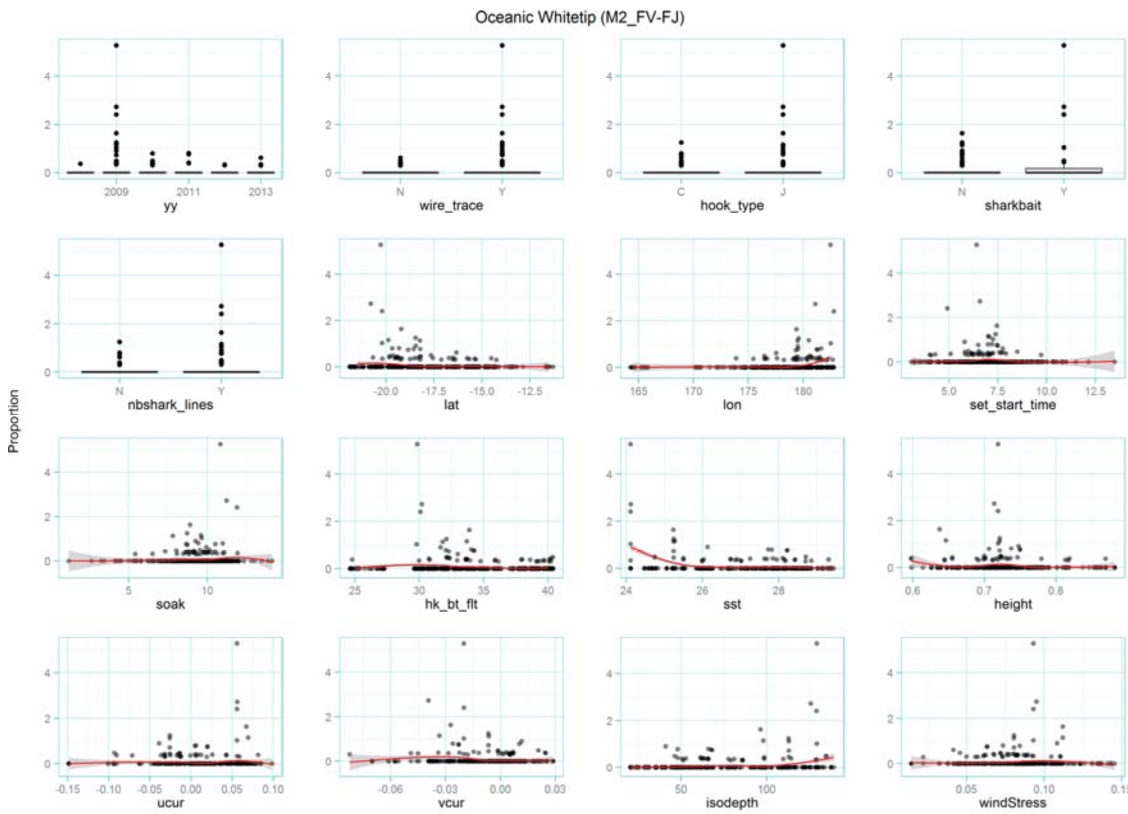


Figure 60: Distributions of set-level catch-rates against the explanatory variables considered for the analysis of oceanic whitetip sharks caught by the Fijian fleet in the M2_FV fishery (file “M2_FV_FJ_OCS_CatchRatesVsCovs.png”)

6.4.1.2 Modelling results

Catch rate (per 1000 hooks) was modelled as a function of set-level characteristics and associated oceanographic variables. The initial set of potential covariates was thinned on the basis of gross collinearity, as determined by Generalized Variance Inflation Factors (GVIFs). A range of potential interactions and non-linear terms were specified *a priori* as per Table 3. These were further thinned during the modelling process in some cases where estimation was not possible e.g. interactions for factor variables where certain combinations of factors were not observed in the dataset.

The model fitted was a GAM with log-link and Tweedie-distributed errors. An initial model run was conducted for the purposes of estimating the governing Tweedie parameter – subsequently refitted with the estimated Tweedie parameter specified. Model "selection" was conducted within the GAM using shrinkage, so model terms of little relevance have their effective degrees of freedom shrunk towards zero. In terms of parametric components, backwards selection (*p*-value<0.05) was used, but not on the key factors of wire-trace, shark-bait, shark-lines and hook type (where applicable). These were generally retained in the model for interpretative interest. The results of this process are summarised in Table 4.

The Tweedie parameter was estimated to be 1.019, with the model returning an adjusted-R^2 of 0.637 and a deviance explained of 73%.

Table 17: Model terms from the generalized additive model. DF = Degrees of Freedom. EDF=Effective Degrees of Freedom. Shrinkage has been applied as a means of model selection. Terms of little relevance have their EDF shrunken.

Model terms	EDF	DF	Chi or F-stat	*p*-value
yy		5	1.14	0.3361
wire_trace		1	5.27	0.0221
nbshark_lines		1	6.11	0.0138
hook_type		1	5.43	0.0202
sharkbait		1	1.09	0.2960
s(set_start_time)	7.91	9	4.98	0.0000
s(lon,lat)	26.82	27	1.71	0.0121
s(soak,hk_bt_flt)	26.19	29	1.93	0.0009

DMP STATS The Coach House, Mount Melville House, St. Andrews, Scotland KY 168 NT, UK +44(0) 1334 477 544 mail@dmpstats.com

The model assumptions were assessed. Non-linearities are assumed to be captured via the GAM which optimises complexity. The adequacy of the assumed Tweedie distribution was assessed via Quantile Residuals (Dunn & Smythe, 1996), as indicated in Figure 8 – these ought to be approximately Normally distributed, producing an approximately straight-line. Potential correlation in the errors was similarly assessed on quantile residuals, both by examination of acf plots and a runs test. The data were ordered by set date within trip IDs, so the correlations are sought via sets that

In this case the Tweedie distribution is acceptable and there is no significant residual correlation once this zero rich distribution has been accounted for. The runs-test returns a *p*-value of 0.28.

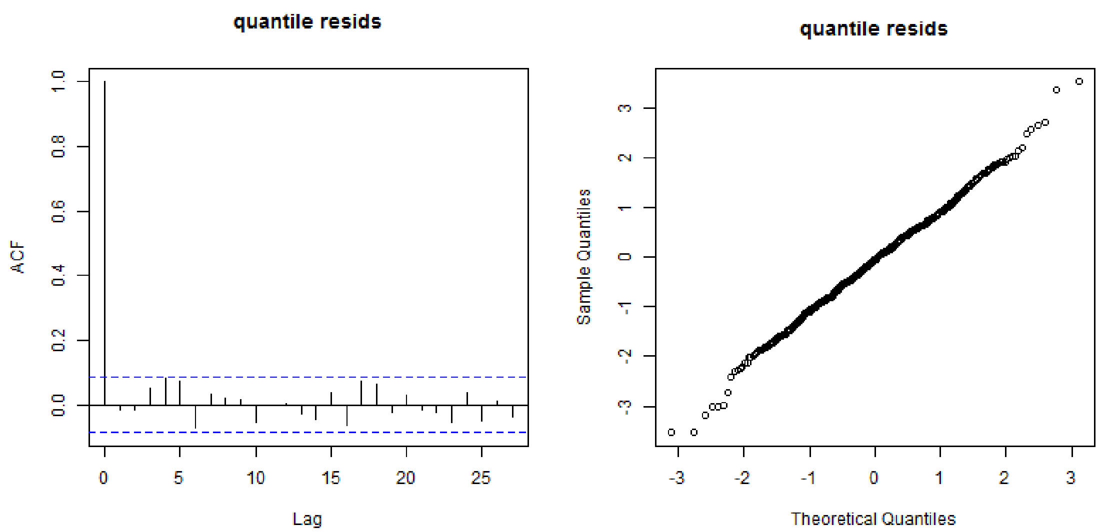


Figure 61: diagnostic plots for the quantile residuals from the GAM. The error distribution is assumed to be a Tweedie distribution, with key parameter estimated.

The following are link-scale estimates of the components of the GAM. All parametric terms in the model are represented here regardless of practical/statistical significance. For smoothers, those with EDFs shrunken to zero are excluded as the plots are flat by definition. Generally only top-line results will be presented outside of the reference case. Where applicable, the parametric terms are presented first, followed by smooth terms and then interactions.

The following observations can be made:

- There is a significantly higher catch rate when shark-lines or wire-trace are present.
- Shark-bait has a negligible partial contribution to catch.
- There were significantly higher levels of catch on hook type C versus hook type J.
- There is temporal variability in terms of years, with high uncertainty about estimates.
- There is a significant and complex relationship between catch-rate and time that the sets were started.
- SST and wind-stress have significant unimodal relationships with catch rate.
- There is a significant spatial component to catch-rates which is not explained by the other covariates as evidenced by the significant latitude-longitude interaction surface. Generally catch rates increase to the East, but highest levels are observed at high latitudes in in the mid-longitude area.
- There is some catch relationship with both soak-time and the hooks between floats. In particular, baskets with few hooks have higher catch rates, particularly coupled with soak times around 10 hours.

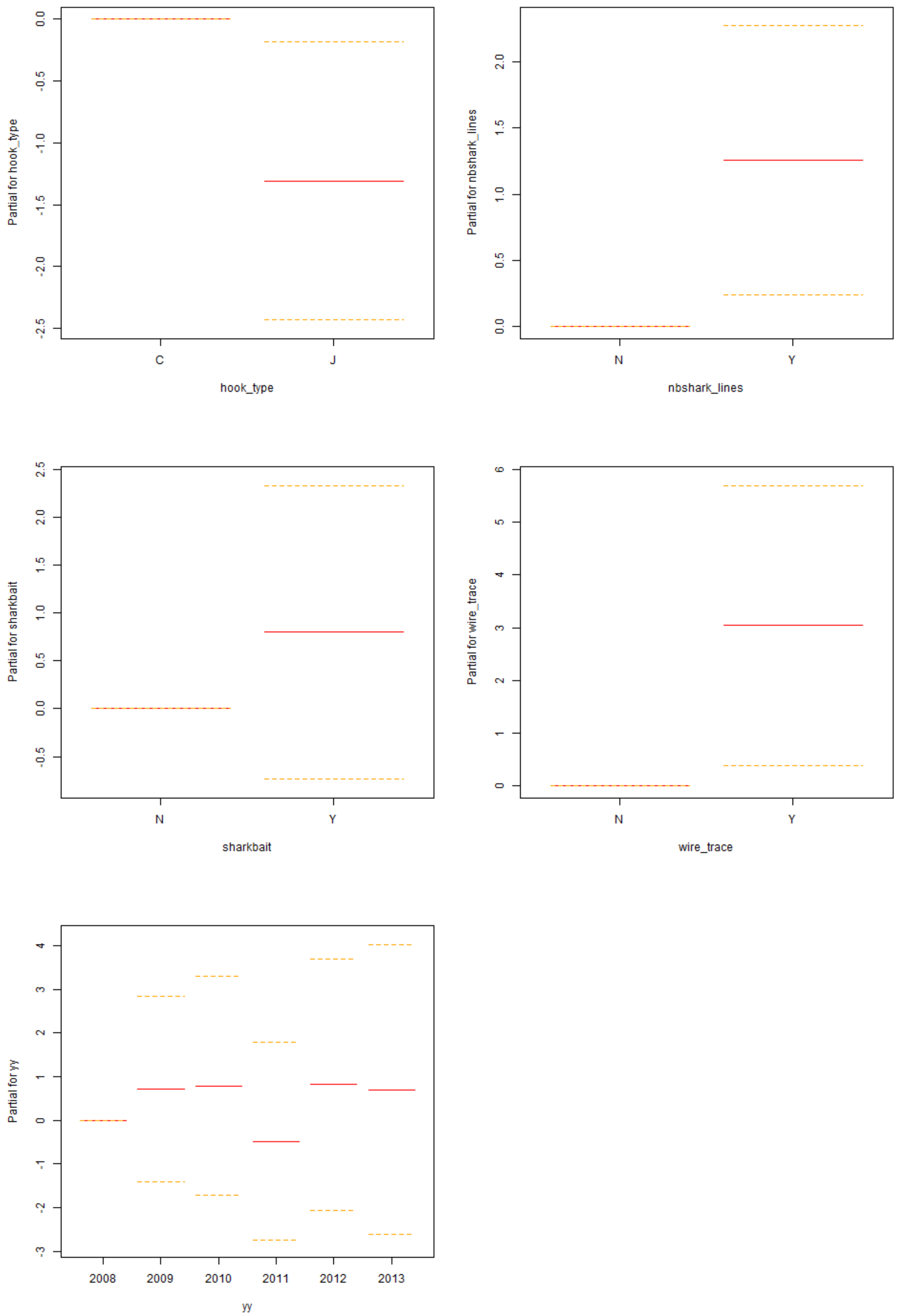


Figure 62: estimates and approximate 95% confidence intervals for parametric terms within the GAM.

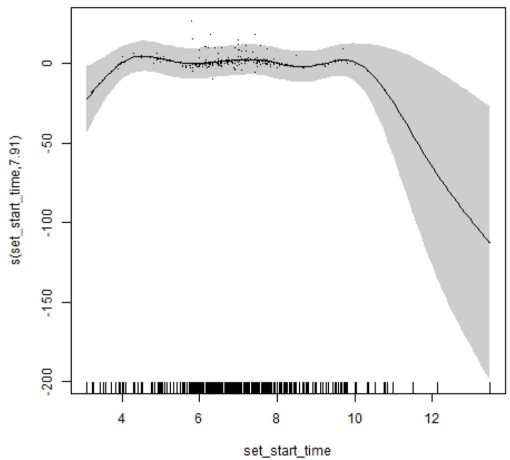


Figure 63: estimated smooths and approximate 95% confidence envelopes for smooth terms within the GAM.

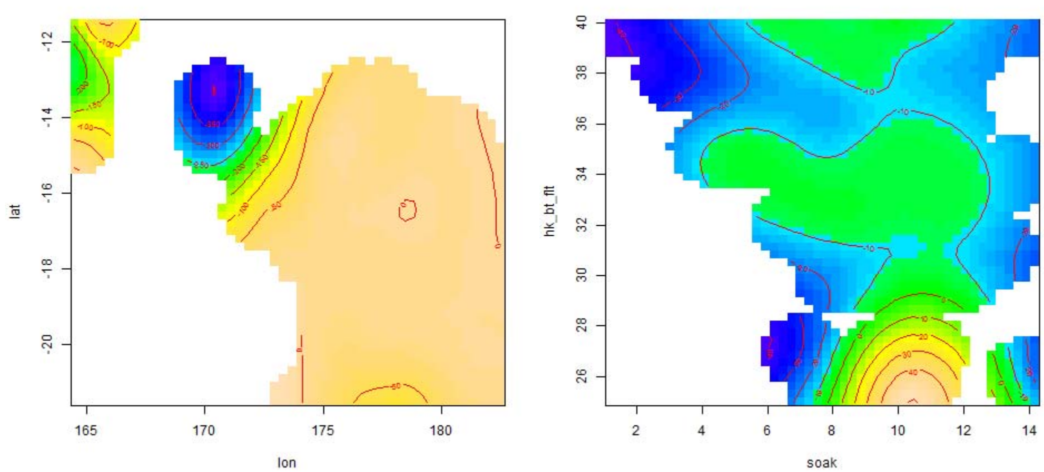


Figure 64: estimated smooth surfaces for the bivariate smooth components within the GAM.

6.4.2 Hook-level catch rate of oceanic whitetip sharks

6.4.2.1 Data Exploration

Figure 65 shows that catch rates of oceanic whitetip sharks in the Fijian fleet tend to be larger at lower hook position numbers (i.e. at shallower hooks). In general, there appears to be no clear patterns of change in catch rates at different hook positions across the values of the covariates considered in the present analysis.

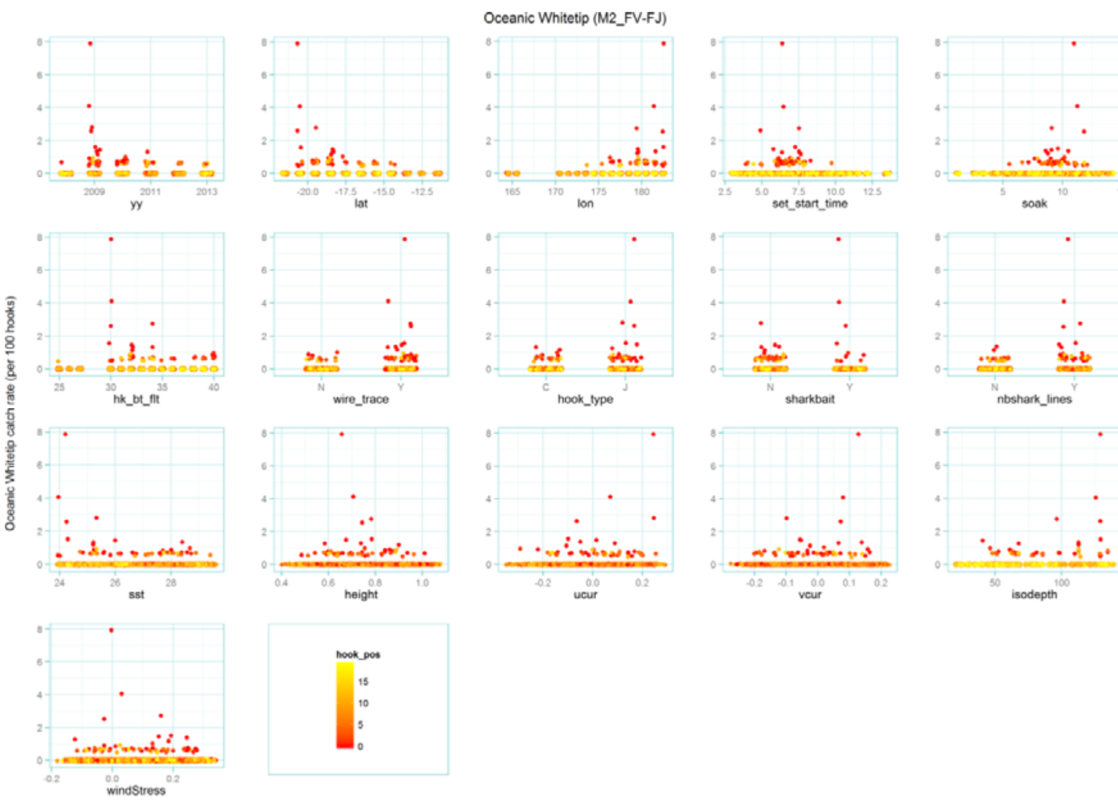


Figure 65: Distributions of per-hook catch-rate responses against the explanatory variables considered in the analysis, for oceanic whitetip sharks caught by the Fijian fleet operating in the M7_FV fishery. Dots were slightly jittered to minimise overplotting and help visualisation and their colour convey the basket hook position of the respective observation (file “M2_FV-FJ-OCS_perHookRatesVsCovs.png”).

6.4.2.2 Modelling results

The effect of hook position on catch rates was conditioned solely on species and fishery. A GAM was fitted with log-link and Tweedie error distribution. An initial model run was conducted for the purposes of estimating the governing Tweedie parameter – subsequently refitted with the estimated Tweedie parameter specified

The Tweedie parameter was estimated to be 1.215, with the model returning an adjusted-R^2 of 0.02 and a deviance explained of 35.3%.

The model assumptions were assessed. Non-linearities are not relevant as the hook position is fitted as a factor variable. The adequacy of the assumed Tweedie distribution was assessed via Quantile Residuals as indicated in Figure 13– these ought to be approximately Normally distributed, producing an approximately straight-line. Potential correlation in the errors was similarly assessed on quantile residuals, both by examination of acf plots and a runs test. The data were ordered by set date within trip IDs, so the correlations are sought via sets that are temporally, and likely spatially, close.

In this case the Tweedie distribution is acceptable and there is no significant residual correlation once this zero rich distribution has been accounted for. The runs-test returns a *p*-value of 0.79.

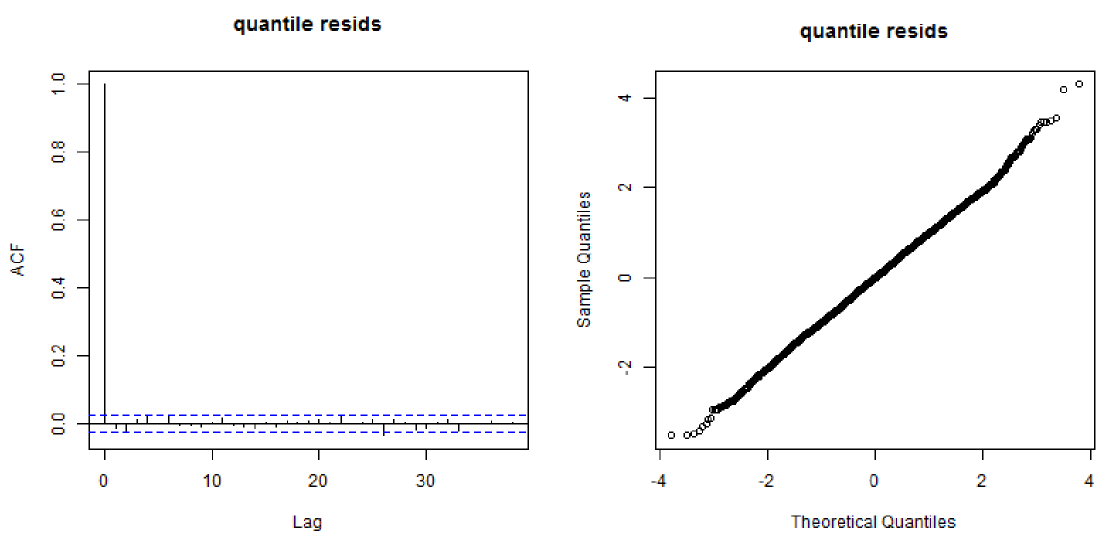


Figure 66: diagnostic plots for the quantile residuals from the GAM. The error distribution is assumed to be a Tweedie distribution, with key parameter estimated.

Estimated catch rate for each hook position and approximate 95% confidence intervals are presented in Figure 14 and Table 18. There are a number of (deeper) hook positions that had no recorded catches against them. No variance estimates are possible for these, so no confidence intervals are calculated.

Table 18: Estimates of catch rates for each of the hook positions. Confidence intervals are approximate 95%. Percentage of total catch is calculated from overall estimated catch over hook positions presented. Confidence intervals for percentages are based on this naïve translation of cate rate.

Hook position	Catch rate (per 100 hooks)	SE	Lower CI	Upper CI	Catch (as % of total)	Lower CI (as % of total)	Upper CI (as % of total)
0	0.0914	1.1822	0.0658	0.1269	63.24	45.55	87.79
1	0.0113	1.4626	0.0054	0.0238	7.82	3.71	16.48
2	0.0084	1.5333	0.0036	0.0194	5.8	2.51	13.42
3	0.0017	2.2312	0.0003	0.0081	1.17	0.24	5.62
4	0.0015	2.292	0.0003	0.0079	1.07	0.21	5.45
5	0.009	1.5148	0.004	0.0204	6.25	2.77	14.1
6	0.0068	1.591	0.0027	0.0169	4.7	1.89	11.68
7	0.0018	2.1958	0.0004	0.0083	1.23	0.26	5.73
8	0	NA	0	NA	0	0	NA
9	0	NA	0	NA	0	0	NA
10	0	NA	0	NA	0	0	NA
11	0.0013	2.4342	0.0002	0.0074	0.9	0.16	5.13
12	0	NA	0	NA	0	0	NA
13	0.0047	1.7397	0.0016	0.0138	3.22	1.09	9.54
14	0.0038	1.8392	0.0011	0.0125	2.62	0.79	8.65
15	0	NA	0	NA	0	0	NA
16	0.0029	2.2526	6.00E-04	0.0141	1.99	0.41	9.78
17	0	NA	0	NA	0	0	NA
18	0	NA	0	NA	0	0	NA
19	0	NA	0	NA	0	0	NA

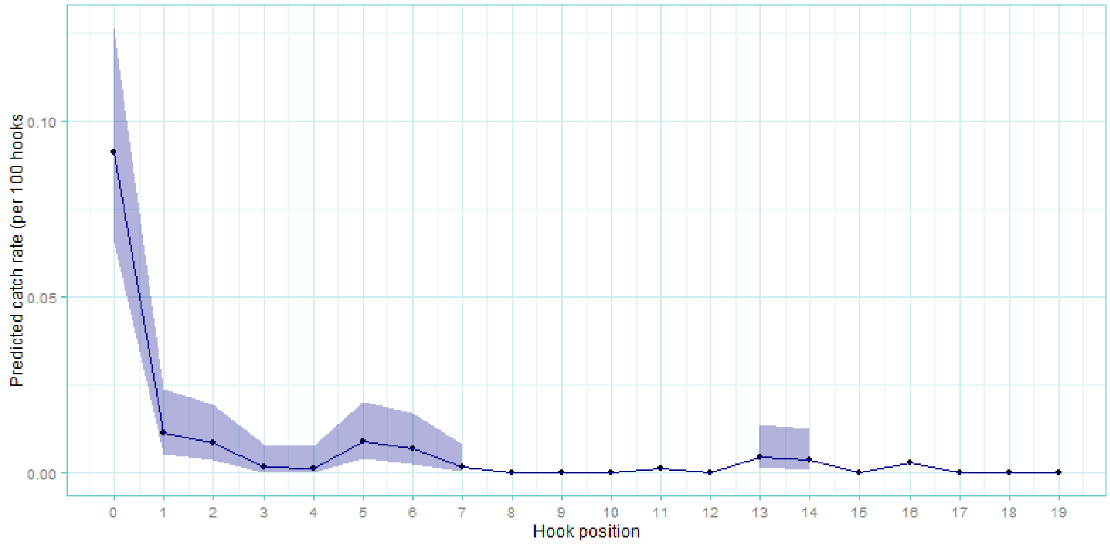


Figure 67: estimated mean catch rate (per 100 hooks) at hook-position with approximate 95% confidence envelope.

The following can be observed from these:

- There is a markedly higher catch rate for this species at hook-position 0 compared to other positions.
- The catch rate at position 0 is 0.091 per 100 hooks [0.066 to 0.127, 95% CI]. This translates to approximately 63.2% [45.6% to 87.8%, 95% CI] of the catch of this species observed in this fisheries data.
- Beyond position 0, there is little clear pattern other than catches being generally near zero for positions 7+.

6.4.3 Condition of oceanic whitetip sharks at time of retrieval

6.4.3.1 Data Exploration

There are no strong patterns between the condition of caught oceanic whitetip sharks at the time of retrieval and most of the covariates considered in the present analysis (Figure 68). There are however some indications that the proportion of dead sharks might have increased when observed sets were performed under certain conditions, like e.g. when hook type "J" was in use.

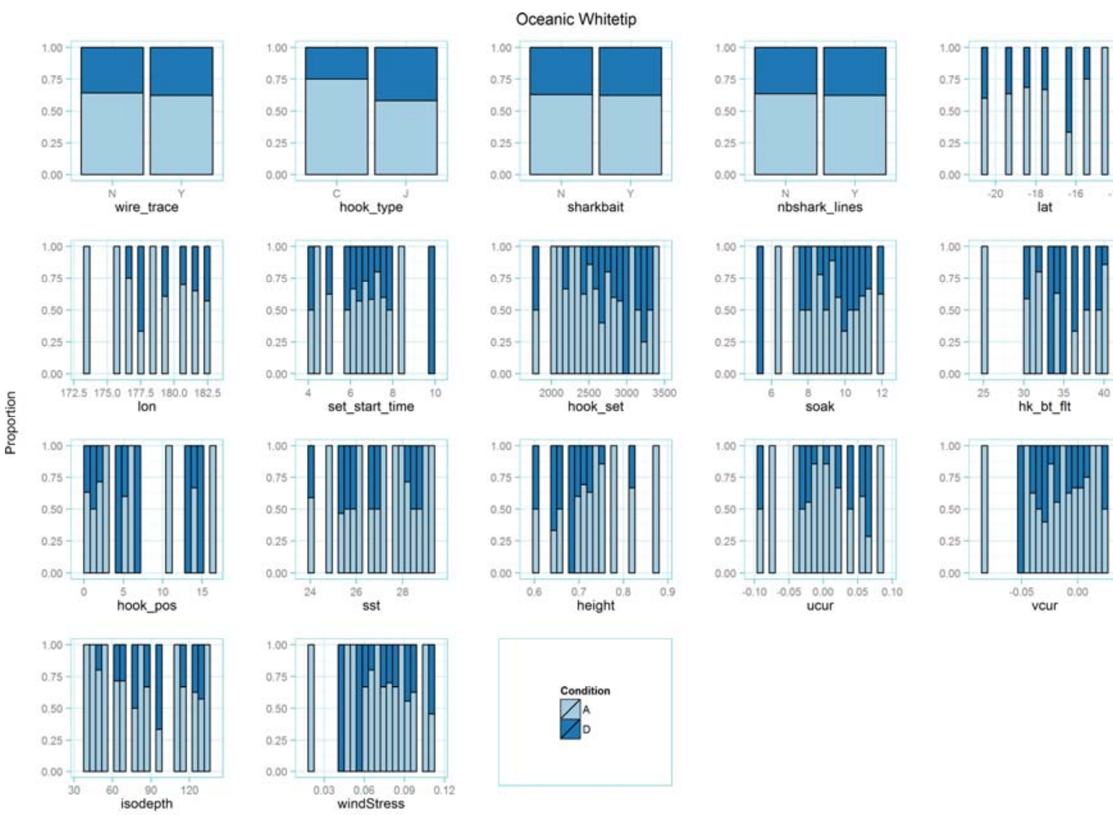


Figure 68: Distributions of condition responses (D= dead, A = alive), expressed in terms of proportion, against the explanatory variables considered for the analysis of oceanic whitetip sharks caught by the Fijian fleet in the M2_FV fishery (file "M2_FV_FJ_OCS_CondtnVsCovs.png")

6.4.3.2 Modelling results

Condition (dead/alive) was modelled as a function of set-level characteristics and associated oceanographic variables. The initial set of potential covariates was thinned on the basis of gross collinearity, as determined by Generalized Variance Inflation Factors (GVIFs). A range of potential interactions and non-linear terms were specified *a priori* as per Table 3. These were further thinned during the modelling process in some cases where estimation was not possible e.g. interactions for factor variables where certain combinations of factors were not observed in the dataset.

The model fitted was a GAM with logit-link and binomially-distributed errors. Model "selection" was conducted within the GAM using shrinkage, so model terms of little relevance have their effective degrees of freedom shrunk towards zero. In terms of parametric components, backwards selection (p-value<0.05) was used, but not on the key factors of wire-trace, shark-bait, shark-lines and hook type (where applicable). These were generally retained in the model for interpretative interest. The results of this process are summarised in Table 4.

In terms of fit to the data, the software generated adjusted R^2 and deviance explained were 0.027 and 6.39% respectively, indicating poor fit. However these measures are of dubious utility for binomial error models. Raw predictive power to the current dataset was also assessed via confusion matrices, providing Sensitivity (true positive rate) = 0.97, Specificity (true negative rate) = 0.21 and a misclassification rate of 0.51 when using the response mean as a decision boundary. These are against the training data and will be over-estimates –cross-validated estimates would be preferred with future analysis.

Table 19: Model terms from the generalized additive model. DF = Degrees of Freedom. EDF=Effective Degrees of Freedom. Shrinkage has been applied as a means of model selection. Terms of little relevance have their EDF shrunken.

Model terms	EDF	DF	Chi or F-stat	*p*-value
sharkbait		1	0.00	0.9787
nbshark_lines		1	2.18	0.1394
wire_trace		1	1.89	0.1694
hook_type		1	4.81	0.0284
s(hook_pos)	0.00	9	0.00	0.9067
s(soak,hk_bt_flt)	0.00	29	0.00	0.6397

The model assumptions were assessed. Non-linearities are assumed to be captured via the GAM which optimises complexity. The adequacy of the assumed binomial distribution was assessed via Quantile Residuals (Dunn & Smythe, 1996), as indicated in Figure 16 – these ought to be approximately Normally distributed, producing an approximately straight-line. Potential correlation in the errors was similarly assessed on quantile residuals, both by examination of acf plots and a runs test. The data were ordered by set date within trip IDs, so the correlations are sought via sets that are temporally, and likely spatially, close. The residuals are distributionally acceptable and little autocorrelation is in evidence, the runs-test providing a *p*-value of 0.46.

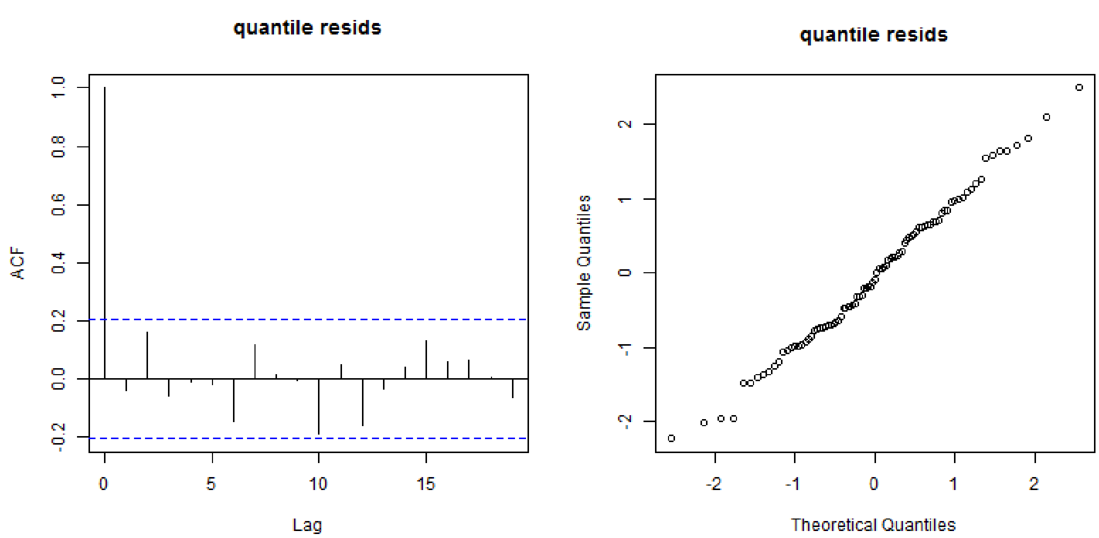


Figure 69: diagnostic plots for the quantile residuals from the GAM. The error distribution is assumed to be a Tweedie distribution, with key parameter estimated.

The following are link-scale estimates of the components of the GAM. All terms in the model are represented here regardless of practical/statistical significance although near-zero EDF relationships are excluded. Generally only top-line results will be presented outside of the reference case. Where applicable, the parametric terms are presented first, followed by smooth terms and then interactions. The following observations can be made:

- Only hook type shows any significant relationship with mortaility, with J-type being higher than C-type.
- Oceanic covariates were not considered due to missing data and the spatial smooth, which may act as a proxy for these, is shrunk to EDF zero i.e. not relevant.

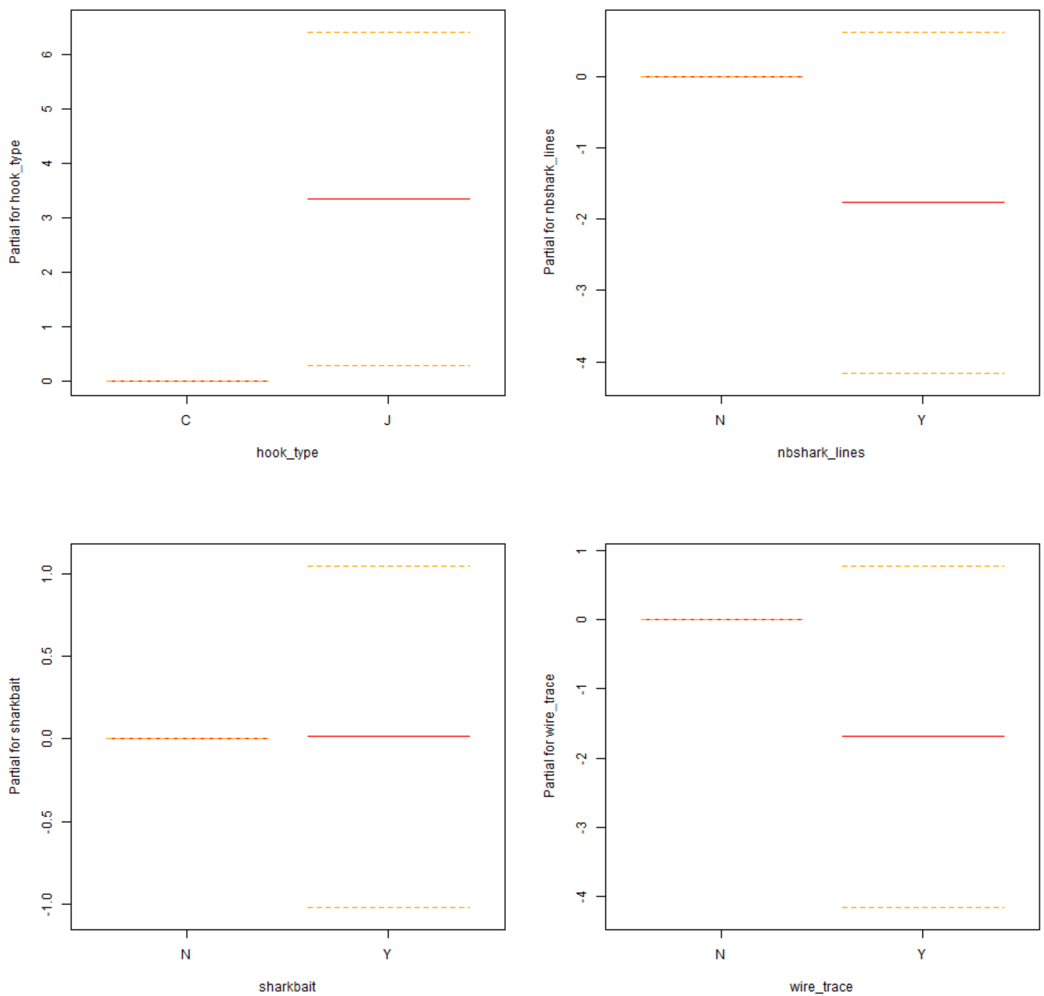


Figure 70: estimates and approximate 95% confidence intervals for parametric terms within the GAM.

6.5 FIJIAN FLEET ANALYSIS – SILKY SHARKS (FAL)

6.5.1 Catch rate of Silky Sharks (FAL)

6.5.1.1 Data Exploration

While predominantly absent in sets from observed trips of the Fijian fleet in the M2_FV fishery, within-set bycatch levels of silky sharks ranged mainly between 1 shark/set (in about 30 sets) and 2 sharks/set (in about 10 sets) (Figure 71). In a few instances there were up to 5 silky sharks observed in a single set.

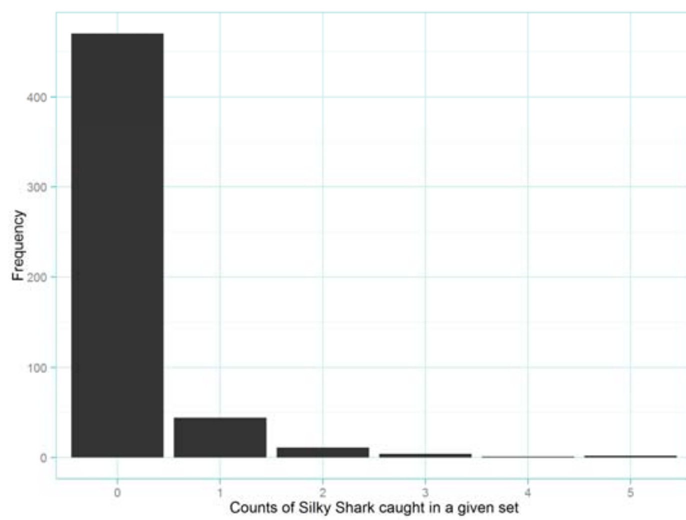


Figure 71: Distribution of counts per set of silky sharks caught in observed trips of the Fijian fleet operating in the M2_FV longline fishery (file M2_FV_FJ_FAL_CatchNumFreqs)

The predominance of zero catches in observed sets hampers the investigation of bivariate patterns between catch rates of silky sharks and each of the covariates considered in the present analysis (Figure 72). Yet, these plots indicate that, on average, higher catch rates were likely to occur when:

- Hook type “J”, wire leader and shark bait were in use;
- Sets were performed in 2009 or 2010;
- Longlines were laid within -20 and -17.5 degrees latitude, and around 180 degrees longitude.

- Soak time was around 10 hours and number of hooks between floats ranged between 35-40 hooks;
- Sets were performed with an average sea surface temperature around 28ºC.

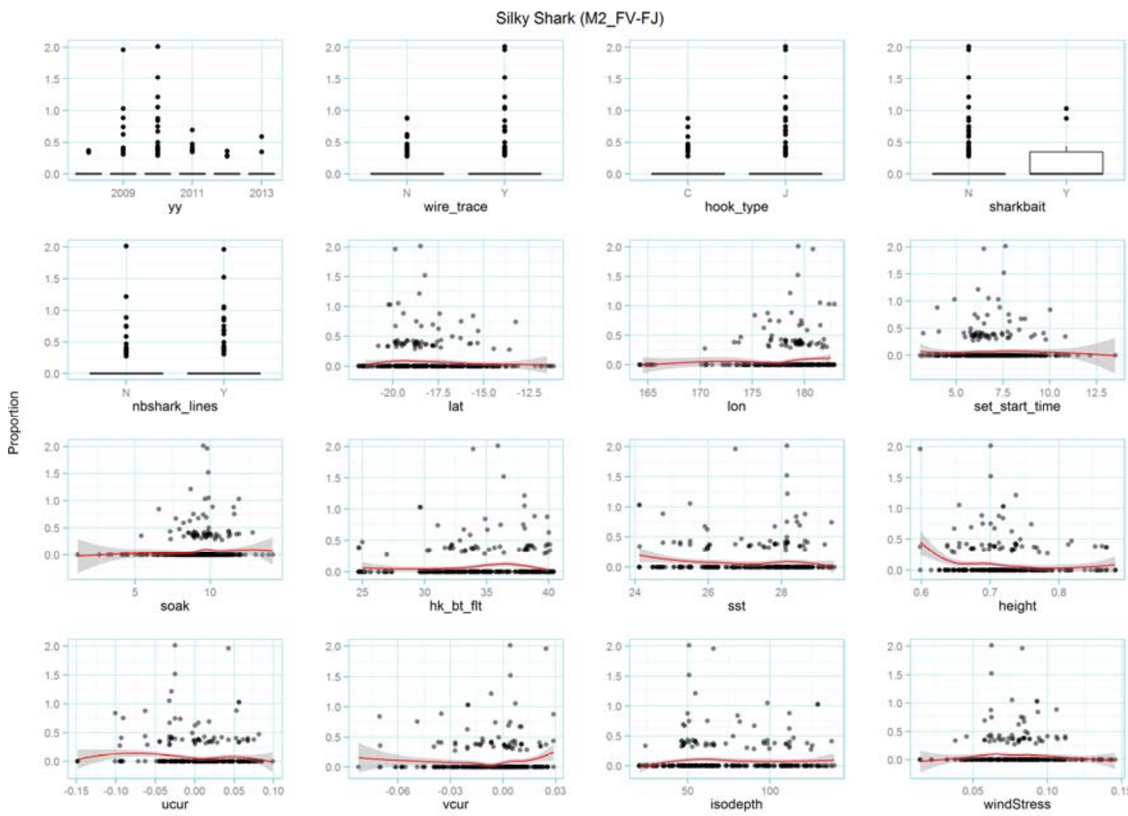


Figure 72: Distributions of set-level catch-rates against the explanatory variables considered for the analysis of oceanic silky caught by the Fijian fleet in the M2_FV fishery (file “M2_FV_FJ_FAL_CatchRatesVsCovs.png”)

6.5.1.2 Modelling results

Modelling details are as per the reference case. The model returned an adjusted-R^2 of 0.313 and a deviance explained of 50%.

Table 20: Model terms from the generalized additive model. DF = Degrees of Freedom. EDF=Effective Degrees of Freedom. Shrinkage has been applied as a means of model selection. Terms of little relevance have their EDF shrunken.

Model terms	EDF	DF	Chi or F-stat	*p*-value
yy		5	2.06	0.0694
wire_trace		1	0.12	0.7300
nbshark_lines		1	1.44	0.2309
hook_type		1	0.12	0.7314
sharkbait		1	0.44	0.5066
s(set_start_time)	7.31	8	1.05	0.3183
s(lon,lat)	26.09	29	0.66	0.8081
s(soak,hk_bt_flt)	27.10	29	1.03	0.2928

None of the covariates considered indicate statistical significance and are not considered further in detail.

6.5.2 Hook-level catch rate of silky sharks

6.5.2.1 Data Exploration

Figure 73 shows that catch rates of silky sharks in the Fijian fleet tend to be larger at lower hook position numbers (i.e. at shallower hooks). In general, there appears to be no clear patterns of change in catch rates at different hook positions across the values of the covariates considered in the present analysis. Still, and despite the considerable degree of overlapping between observations, there is loose indication for larger catch rates of this species at shallower hooks when wire leader, hook type "J" and shark lines are in use (Figure 73).

Figure 73: Distributions of per-hook catch-rate responses against the explanatory variables considered in the analysis, for silky sharks caught by the Fijian fleet operating in the M7_FV fishery. Dots were slightly jittered to minimise overplotting and help visualisation and their colour convey the basket hook position of the respective observation (file "M2_FV-FJ-FAL_perHookRatesVsCovs.png").

6.5.2.2 Modelling results

Modelling details are as per the reference case. The model returned an adjusted-R^2 of 0.027 and a deviance explained of 22.7%.

Table 21: Estimates of catch rates for each of the hook positions. Confidence intervals are approximate 95%. Percentage of total catch is calculated from overall estimated catch over hook positions presented. Confidence intervals for percentages are based on this naïve translation of cate rate.

Hook position	Catch rate (per 100 hooks)	SE	Lower CI	Upper CI	Catch (as % of total)	Lower CI (as % of total)	Upper CI (as % of total)
0	0.0673	1.1578	0.0505	0.0896	47.45	35.6	63.23
1	0.0163	1.3361	0.0093	0.0288	11.53	6.53	20.34
2	0.012	1.4001	0.0062	0.0232	8.45	4.37	16.35
3	0.0114	1.4125	0.0058	0.0223	8.01	4.07	15.76
4	0.0052	1.6561	0.0019	0.0139	3.65	1.36	9.81
5	0.0031	1.9124	0.0009	0.011	2.17	0.61	7.73
6	0.005	1.67	0.0018	0.0137	3.53	1.29	9.64
7	0.0013	2.673	0.0002	0.0089	0.91	0.13	6.28
8	0.0036	1.8174	0.0011	0.0117	2.57	0.8	8.29
9	0.0043	1.7308	0.0015	0.0127	3.07	1.05	8.99
10	0.009	1.4727	0.0042	0.0191	6.32	2.96	13.5
11	0	NA	0	NA	0	0	NA
12	0.0016	2.4403	0.0003	0.0091	1.12	0.19	6.43
13	0	NA	0	NA	0	0	NA
14	0	NA	0	NA	0	0	NA
15	0.0017	2.5592	0.0003	0.011	1.23	0.19	7.73
16	0	NA	0	NA	0	0	NA
17	0	NA	0	NA	0	0	NA
18	0	NA	0	NA	0	0	NA
19	0	NA	0	NA	0	0	NA

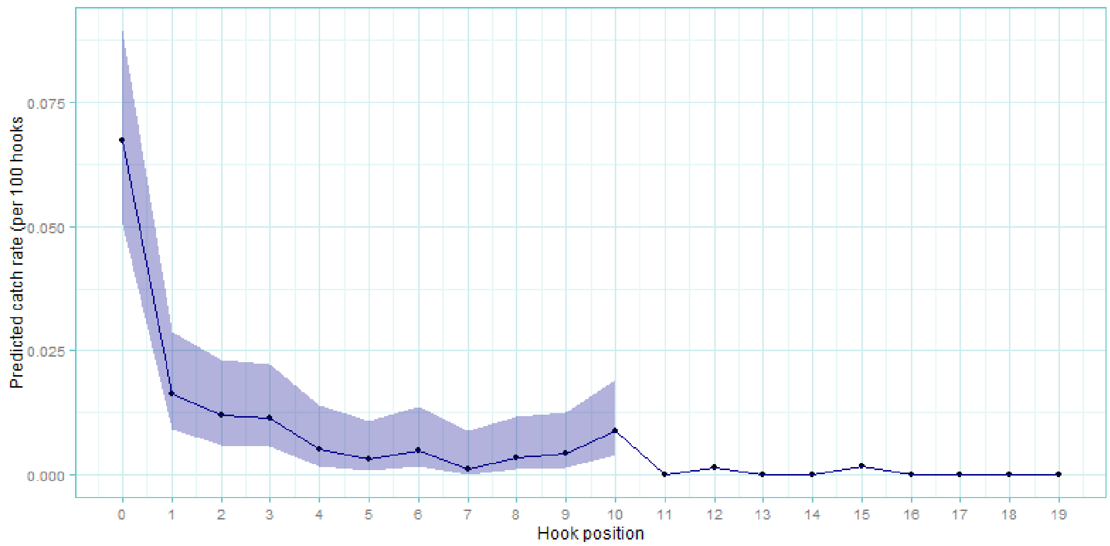


Figure 74: estimated mean catch rate (per 100 hooks) at hook-position with approximate 95% confidence envelope.

6.5.3 Condition of silky sharks at time of retrieval

6.5.3.1 Data Exploration

There are no strong patterns between the condition of caught silky sharks at the time of retrieval and most of the covariates considered in the present analysis (Figure 75). There are however some indications that the proportion of dead silky sharks might have increased when observed sets were performed under certain conditions, like e.g.:

- If hook type "C" is in use;
- When sets are performed;
- If longlines were set around 175 degrees of longitude and -17.5 degrees of latitude.

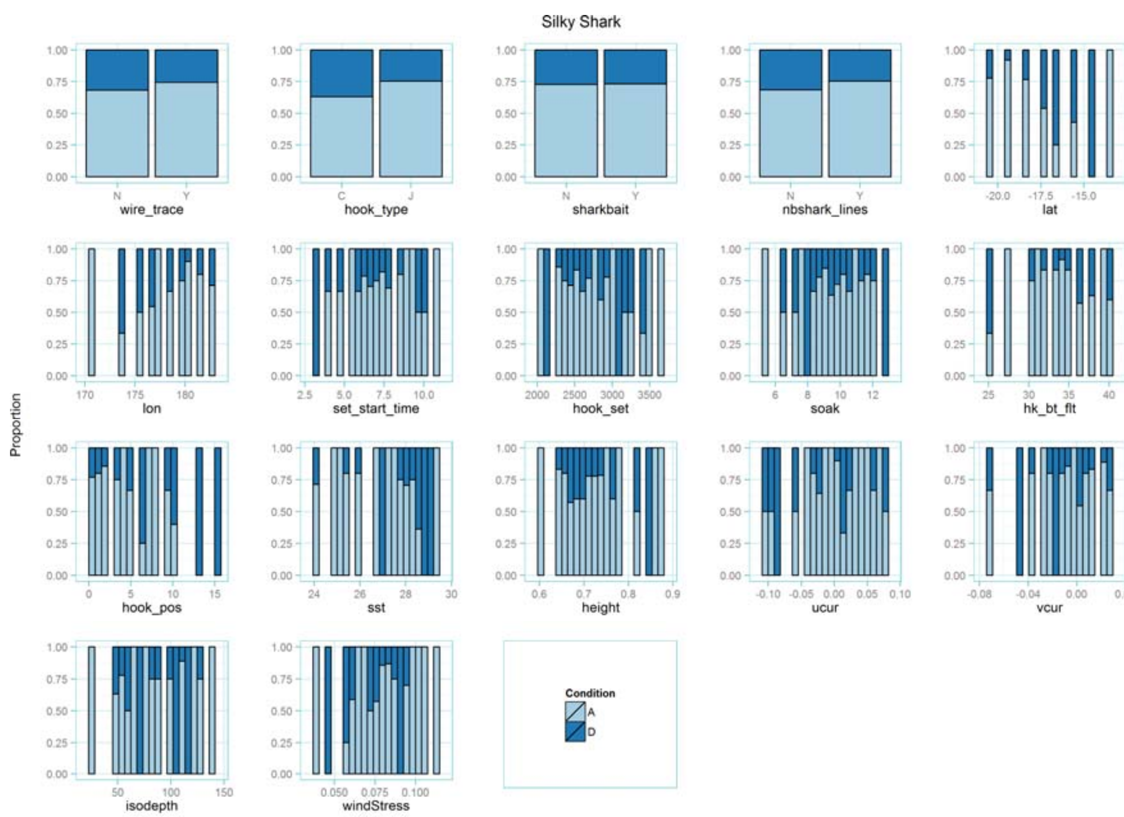


Figure 75: Distributions of condition responses (D= dead, A = alive), expressed in terms of proportion, against the explanatory variables considered for the analysis of silky sharks caught by the Fijian fleet in the M2_FV fishery (file "M2_FV_FJ_FAL_CondtnVsCovs.png")

6.5.3.2 Modelling results

Modelling details are as per the reference case. The model returned an adjusted-R^2 of 0.071 and a deviance explained of 12.1%. Raw predictive power to the current dataset was also assessed via confusion matrices on the training data, providing Sensitivity (true positive rate) = 0.56, Specificity (true negative rate) = 0.72 and a misclassification rate of 0.33 when using the response mean as a decision boundary. These are against the training data and will be over-estimates.

Table 22: Model terms from the generalized additive model. DF = Degrees of Freedom. EDF=Effective Degrees of Freedom. Shrinkage has been applied as a means of model selection. Terms of little relevance have their EDF shrunken.

Model terms	EDF	DF	Chi or F-stat	*p*-value
sharkbait		1	0.00	0.9769
nbshark_lines		1	0.31	0.5762
wire_trace		1	0.45	0.5009
hook_type		1	1.36	0.2437
s(hook_pos)	1.56	9	6.05	0.0194
s(soak,hk_bt_flt)	1.56	27	3.30	0.0621

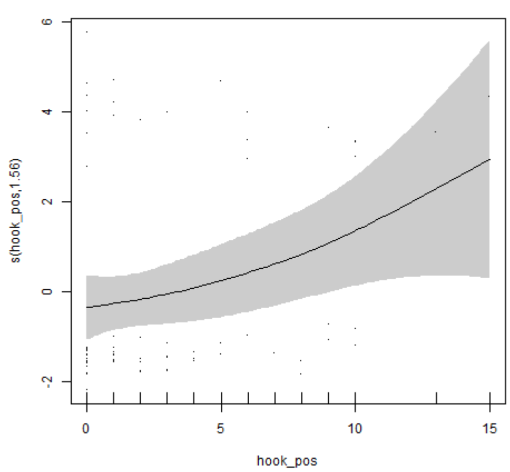


Figure 76: estimated smooths and approximate 95% confidence envelopes for smooth terms within the GAM.

DMP STATS The Coach House, Mount Melville House, St. Andrews, Scotland KY 168 NT, UK +44(0) 1334 477 544 mail@dmpstats.com

6.6 FIJIAN FLEET ANALYSIS – THRESHER SHARKS (THR)

6.6.1 Catch rate of thresher sharks (THR)

6.6.1.1 Data Exploration

While predominantly absent in sets from observed trips of the Fijian fleet in the M2_FV fishery, within-set bycatch levels of thresher sharks ranged mainly between 1 shark/set (in about 20 sets) and 2 sharks/set (in about 5 sets) (Figure 77). In a few instances there were up to 4 thresher sharks observed in a single set.

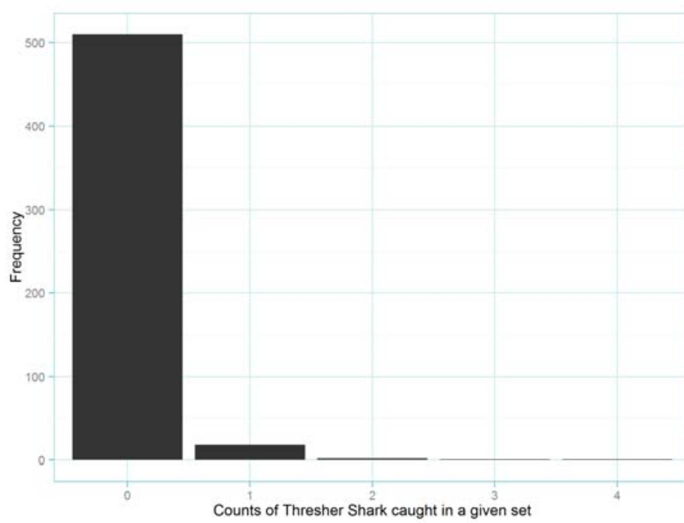


Figure 77: Distribution of counts per set of thresher shark caught in observed trips of the Fijian fleet operating in the M2_FV longline fishery (file M2_FV_FJ_THR_CatchNumFreqs)

The predominance of zero catches and the small number of specimens in observed sets hampers the investigation of bivariate patterns between catch rates of thresher sharks and each of the covariates considered in the present analysis (Figure 78).

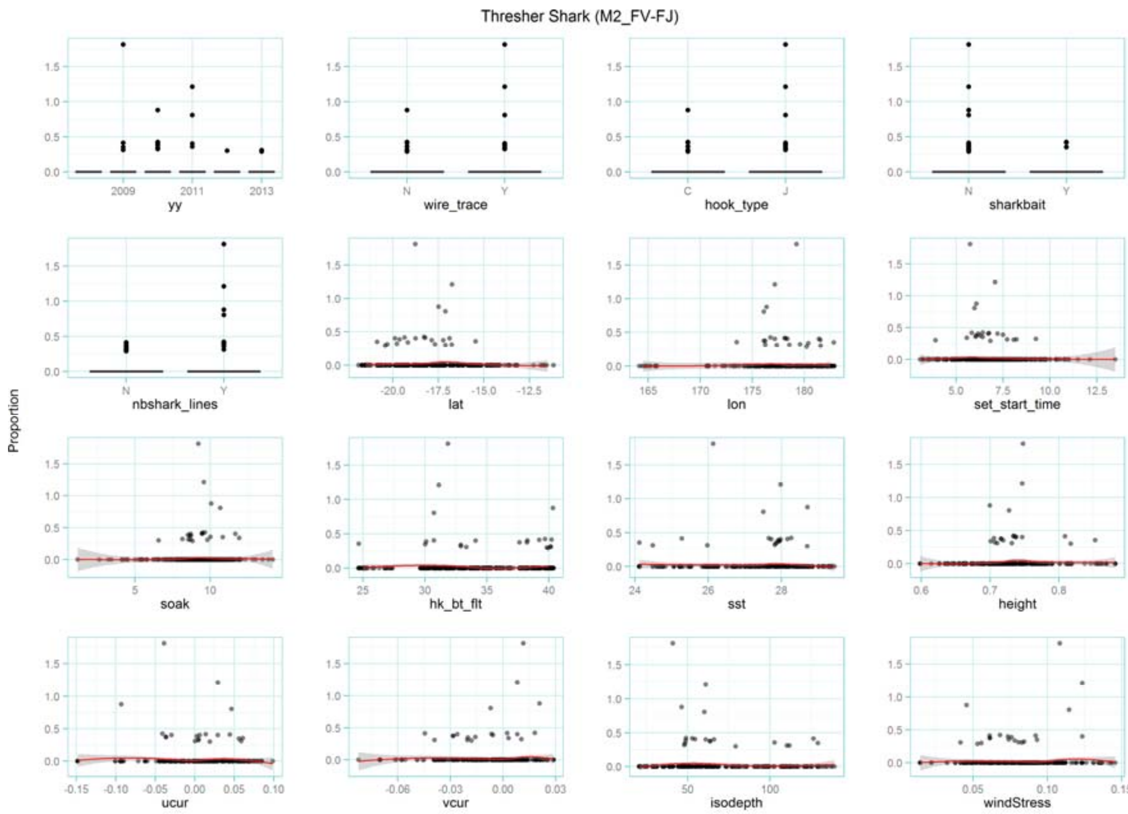


Figure 78: Distributions of set-level catch-rates against the explanatory variables considered for the analysis of thresher sharks caught by the Fijian fleet in the M2_FV fishery (file "M2_FV_FJ_THR_CatchRatesVsCovs.png")

6.6.1.2 Modelling results

Modelling details are as per the reference case. The model returned an adjusted-R^2 of 0.296 and a deviance explained of 64.2%.

Table 23: Model terms from the generalized additive model. DF = Degrees of Freedom. EDF=Effective Degrees of Freedom. Shrinkage has been applied as a means of model selection. Terms of little relevance have their EDF shrunken.

Model terms	EDF	DF	Chi or F-stat	*p*-value
wire_trace		1	2.16	0.1425
nbshark_lines		1	0.49	0.4859
hook_type		1	4.60	0.0324
sharkbait		1	2.38	0.1235
s(lon,lat)	22.88	29	1.63	0.0024
s(soak,hk_bt_flt)	26.94	27	2.67	0.0000

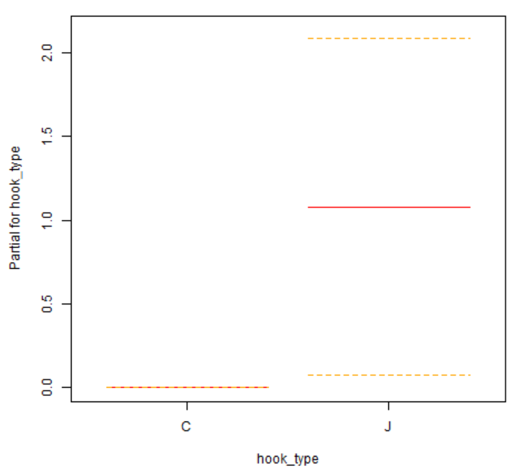


Figure 79: estimates and approximate 95% confidence intervals for parametric terms within the GAM.

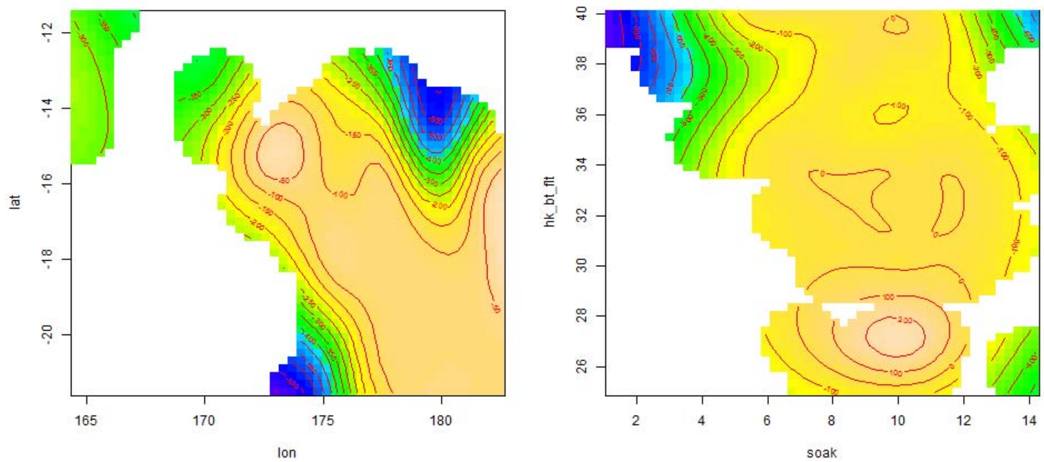


Figure 80: estimated smooth surfaces for the bivariate smooth components within the GAM.

6.6.2 Hook-level catch rate of thresher sharks

6.6.2.1 Data Exploration

Figure 81 shows that, as opposed to the trend in other species, catch rates of thresher sharks in the Fijian fleet tend to be larger at higher hook position numbers (i.e. at deeper hooks). However, there are no demarked trends in how the relationship between catch rates and hook positions changes with varying levels of the covariates considered in the present analysis.

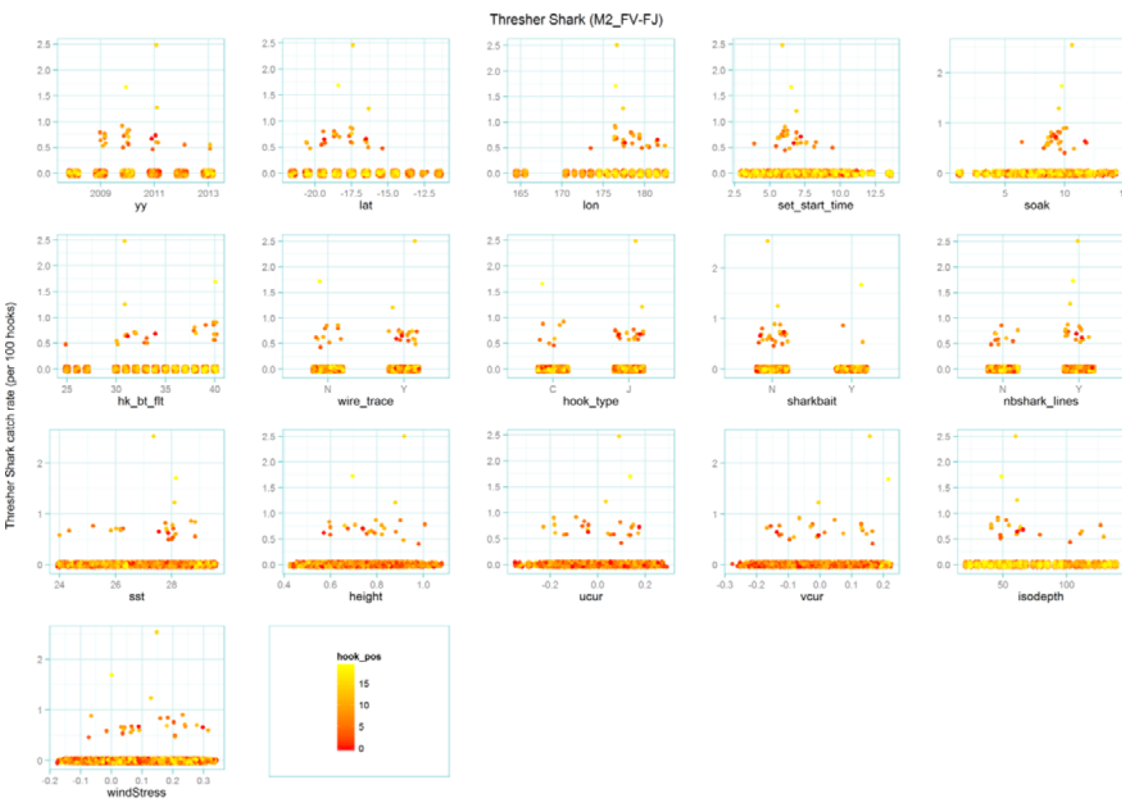


Figure 81: Distributions of per-hook catch-rate responses against the explanatory variables considered in the analysis, for thresher sharks caught by the Fijian fleet operating in the M7_FV fishery. Dots were slightly jittered to minimise overplotting and help visualisation and their colour convey the basket hook position of the respective observation (file "M2_FV-FJ-THR_perHookRatesVsCovs.png").

6.6.2.2 Modelling results

Modelling details are as per the reference case. The model returned an adjusted-R^2 of <0.001 and a deviance explained of 14.6%.

Table 24: Estimates of catch rates for each of the hook positions. Confidence intervals are approximate 95%. Percentage of total catch is calculated from overall estimated catch over hook positions presented. Confidence intervals for percentages are based on this naïve translation of cate rate.

Hook position	Catch rate (per 100 hooks)	SE	Lower CI	Upper CI	Catch (as % of total)	Lower CI (as % of total)	Upper CI (as % of total)
0	0.0035	1.8982	0.001	0.0124	5.35	1.52	18.81
1	0	NA	0	NA	0	0	NA
2	0	NA	0	NA	0	0	NA
3	0.0047	1.7752	0.0015	0.0144	7.08	2.3	21.82
4	0.0034	1.9225	9.00E-04	0.0121	5.09	1.41	18.34
5	0.0038	1.864	0.0011	0.0129	5.76	1.7	19.52
6	0.0022	2.1513	5.00E-04	0.0101	3.41	0.76	15.29
7	0.0036	1.884	0.0011	0.0126	5.52	1.59	19.09
8	0	NA	0	NA	0	0	NA
9	0.004	1.8352	0.0012	0.0133	6.14	1.87	20.19
10	0.002	2.2468	4.00E-04	0.0095	2.96	0.61	14.48
11	0.0072	1.6243	0.0028	0.0185	10.85	4.19	28.07
12	0.0035	1.8996	0.001	0.0124	5.34	1.52	18.78
13	0	NA	0	NA	0	0	NA
14	0.0021	2.2736	4.00E-04	0.0105	3.18	0.64	15.89
15	0.0123	1.5402	0.0053	0.0286	18.59	7.97	43.34
16	0	NA	0	NA	0	0	NA
17	0	NA	0	NA	0	0	NA
18	0	NA	0	NA	0	0	NA
19	0.0137	1.9153	0.0038	0.0489	20.73	5.8	74.1

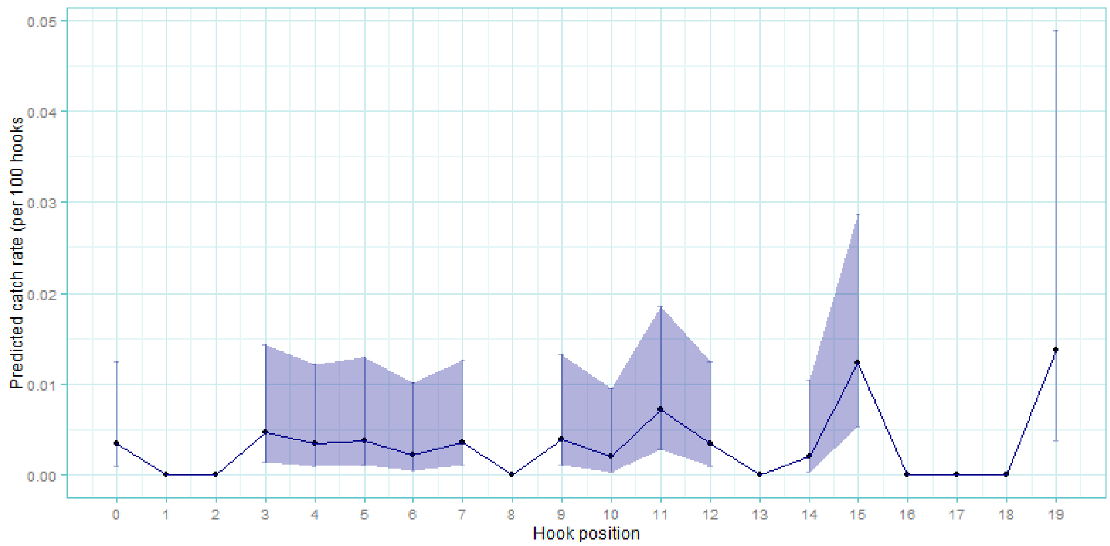


Figure 82: estimated mean catch rate (per 100 hooks) at hook-position with approximate 95% confidence envelope.

6.6.3 Condition of thresher sharks at time of retrieval

6.6.3.1 Data Exploration

There are no strong patterns between the condition of caught thresher sharks at the time of retrieval and most of the covariates considered in the present analysis (Figure 83). There are however some indications that the proportion of dead sharks might have increased when observed sets were performed under certain conditions, like e.g. when hook type "J" and shark lines were in use and shark bait was not employed.

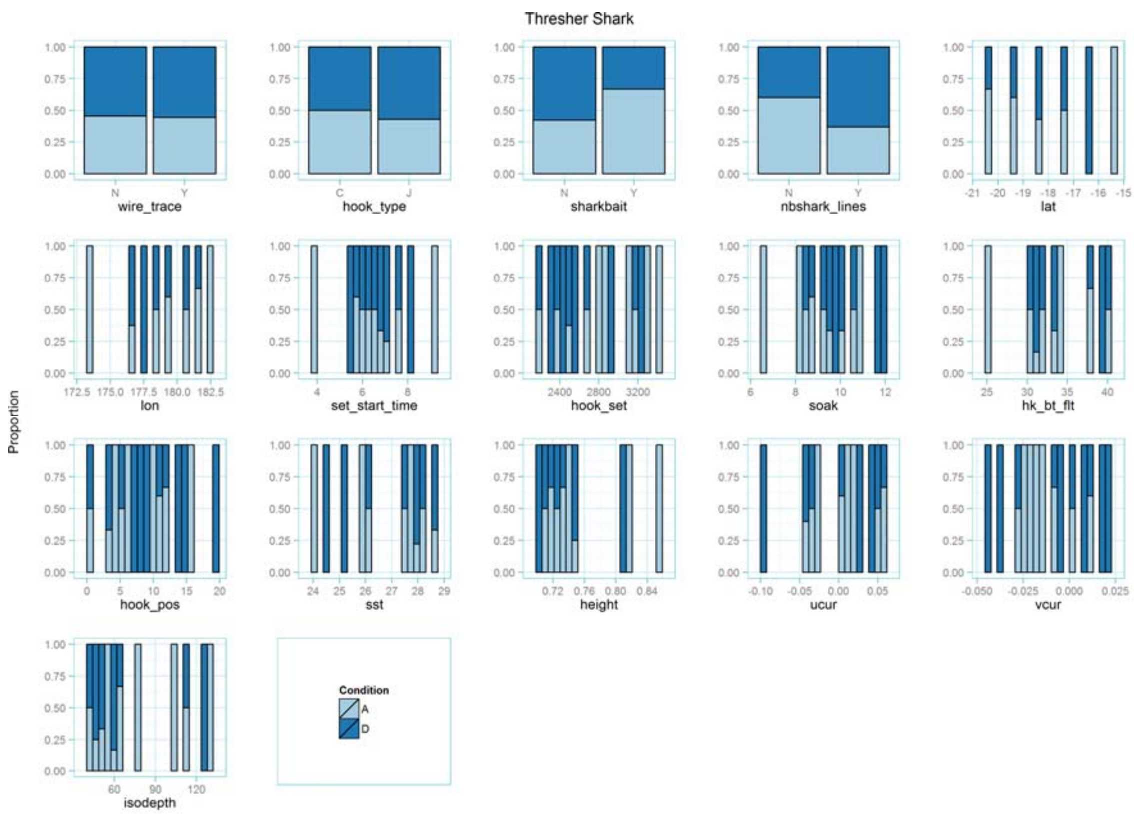


Figure 83: Distributions of condition responses (D= dead, A = alive), expressed in terms of proportion, against the explanatory variables considered for the analysis of thresher sharks caught by the Fijian fleet in the M2_FV fishery (file "M2_FV_FJ_THR_CondtnVsCovs.png")

6.6.3.2 Modelling results

There were only 29 observations of condition at retrieval for this species and fishery, hence no formal analysis was conducted.

6.7 FIJIAN FLEET ANALYSIS – BLUE SHARKS (BSH)

6.7.1 Catch rate of blue sharks

6.7.1.1 Data Exploration

While predominantly absent in sets from observed trips of the Fijian fleet in the M2_FV fishery, within-set bycatch levels of blue sharks ranged mainly between 1 shark/set (in about 110 sets) and 5 sharks/set (in about 10 sets) (Figure 84). In a few instances there were up to 9 and 11 blue sharks observed in one single set.

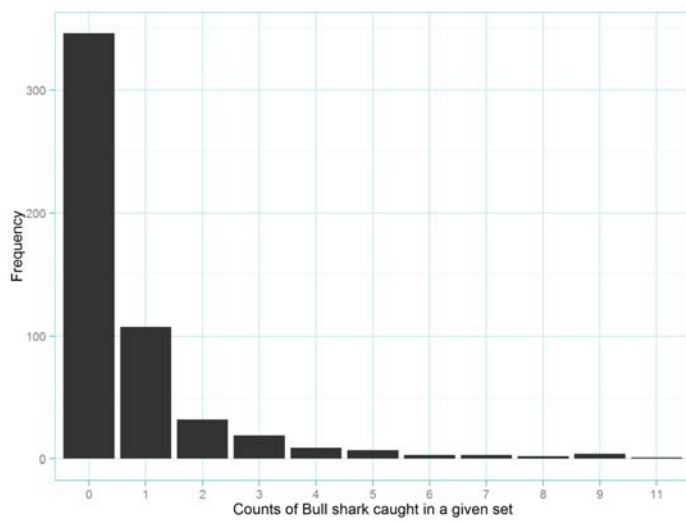


Figure 84: Distribution of counts per set of blue shark caught in observed trips of the Fijian fleet operating in the M2_FV longline fishery (file M2_FV_FJ_BSH_CatchNumFreqs).

The predominance of zero catches in observed sets hampers the investigation of bivariate patterns between catch rates of blue sharks and each of the covariates considered in the present analysis (Figure 85Figure 85). Yet, these plots indicate that, on average, higher catch rates were likely to occur when:

- Hook type “J” and wire leader were in use;
- Sets were performed in 2009 or 2010;

- Longlines were laid within -20 and -17.5 degrees latitude, and around 180 degrees longitude.
- The number of hooks between floats ranged between 30-35 hooks;
- Sets were performed with an average sea surface temperature around 28^{o}C, average sea surface high remained around 0.7 meters and the average wind stress varied between 0.05-0.1 Newton/m^2.

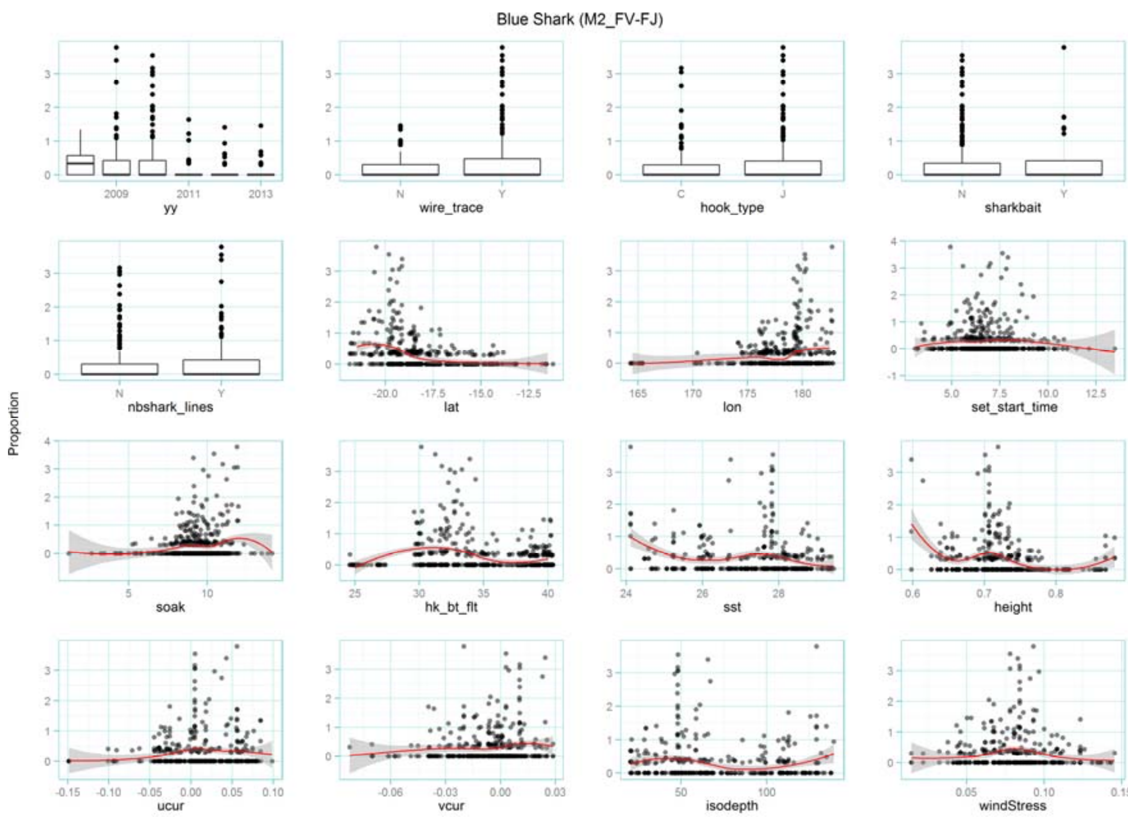


Figure 85 Distributions of set-level catch-rates against the explanatory variables considered for the analysis of blue sharks caught by the Fijian fleet in the M2_FV fishery (file "M2_FV_FJ_BSH_CatchRatesVsCovs.png")

6.7.1.2 Modelling results

Modelling details are as per the reference case. The model returned an adjusted-R^2 of 0.466 and a deviance explained of 49.6%.

Table 25: Model terms from the generalized additive model. DF = Degrees of Freedom. EDF=Effective Degrees of Freedom. Shrinkage has been applied as a means of model selection. Terms of little relevance have their EDF shrunken.

Model terms	EDF	DF	Chi or F-stat	*p*-value
yy		5	3.46	0.0044
wire_trace		1	11.59	0.0007
nbshark_lines		1	1.59	0.2082
hook_type		1	0.09	0.7608
sharkbait		1	0.02	0.8974
s(set_start_time)	3.94	9	0.62	0.1826
s(lon,lat)	9.70	29	2.76	0.0000
s(soak,hk_bt_flt)	11.49	29	1.17	0.0000

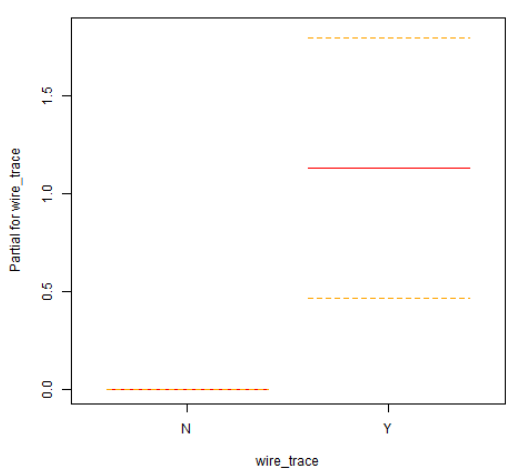


Figure 86: estimates and approximate 95% confidence intervals for parametric terms within the GAM.

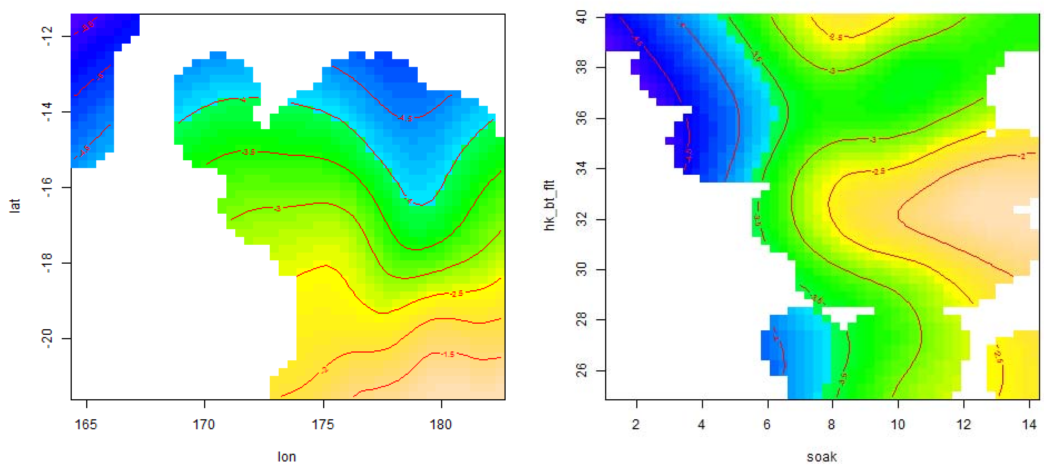


Figure 87: estimated smooth surfaces for the bivariate smooth components within the GAM.

6.7.2 Hook-level catch rate of blue sharks

6.7.2.1 Data Exploration

Figure 88 shows that catch rates of blue sharks in the Fijian fleet tend to be larger at lower hook position numbers (i.e. at shallower hooks). However, there are also a few deeper hooks (i.e. higher hook position values) with large catch rates. There are no clear trends in how the relationship between catch rates and hook positions changes with varying levels of the covariates considered in the present analysis.

Figure 88: Distributions of per-hook catch-rate responses against the explanatory variables considered in the analysis, for blue sharks caught by the Fijian fleet operating in the M7_FV fishery. Dots were slightly jittered to minimise overplotting and help visualisation and their colour convey the basket hook position of the respective observation (file "M2_FV-FJ-BSH_perHookRatesVsCovs.png").

6.7.2.2 Modelling results

Modelling details are as per the reference case. The model returned an adjusted-R^2 of 0.04 and a deviance explained of 14.2%.

Table 26: Estimates of catch rates for each of the hook positions. Confidence intervals are approximate 95%. Percentage of total catch is calculated from overall estimated catch over hook positions presented. Confidence intervals for percentages are based on this naïve translation of cate rate.

Hook position	Catch rate (per 100 hooks)	SE	Lower CI	Upper CI	Catch (as % of total)	Lower CI (as % of total)	Upper CI (as % of total)
0	0.1295	1.1386	0.1005	0.1671	22.36	17.34	28.84
1	0.0835	1.1744	0.0609	0.1145	14.42	10.52	19.76
2	0.0663	1.1973	0.0466	0.0943	11.44	8.04	16.28
3	0.0558	1.2163	0.038	0.0819	9.63	6.56	14.14
4	0.0622	1.204	0.0432	0.0895	10.74	7.46	15.45
5	0.0322	1.2922	0.0195	0.0531	5.55	3.36	9.17
6	0.0224	1.3574	0.0123	0.0408	3.87	2.13	7.05
7	0.0238	1.3458	0.0133	0.0426	4.11	2.29	7.35
8	0.0115	1.5271	0.005	0.0264	1.99	0.87	4.56
9	0.0109	1.5455	0.0046	0.0255	1.88	0.8	4.41
10	0.0089	1.6175	0.0035	0.0228	1.53	0.6	3.93
11	0.0076	1.679	0.0028	0.021	1.31	0.48	3.63
12	0.012	1.514	0.0053	0.0271	2.07	0.92	4.67
13	0.0062	1.8006	0.002	0.0197	1.07	0.34	3.4
14	0.013	1.5151	0.0058	0.0294	2.25	1	5.08
15	0.0088	1.7011	0.0031	0.025	1.52	0.54	4.31
16	0.0085	1.8484	0.0025	0.0282	1.46	0.44	4.87
17	0.0081	2.0623	0.002	0.0335	1.4	0.34	5.77
18	0.0081	2.1817	0.0018	0.0374	1.4	0.3	6.45
19	0	NA	0	NA	0	0	NA

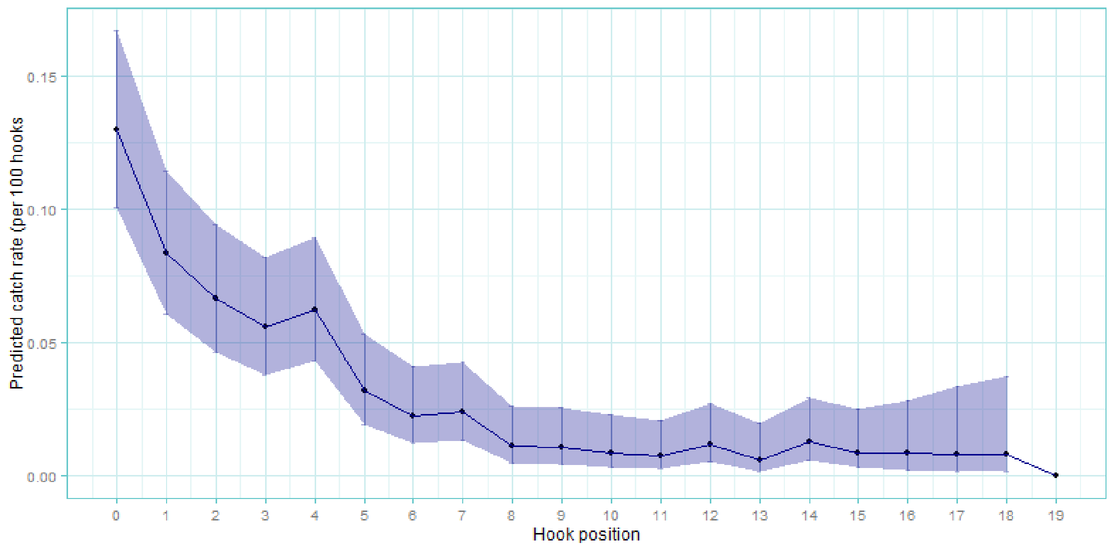


Figure 89: estimated mean catch rate (per 100 hooks) at hook-position with approximate 95% confidence envelope.

6.7.3 Condition of blue sharks at time of retrieval

6.7.3.1 Data Exploration

There are no strong patterns between the condition of caught blue sharks at the time of retrieval and most of the covariates considered in the present analysis (Figure 90). There are however some indications that the proportion of dead sharks might have increased when observed sets were performed under certain conditions, like e.g. (i) when hook type “J” were in use; and (ii) if wind stress levels reached levels around 0.1 Newton/m^2. Interestingly, the proportion of dead blue sharks also appears to increase as catching hooks were positioned deeper in the water column.

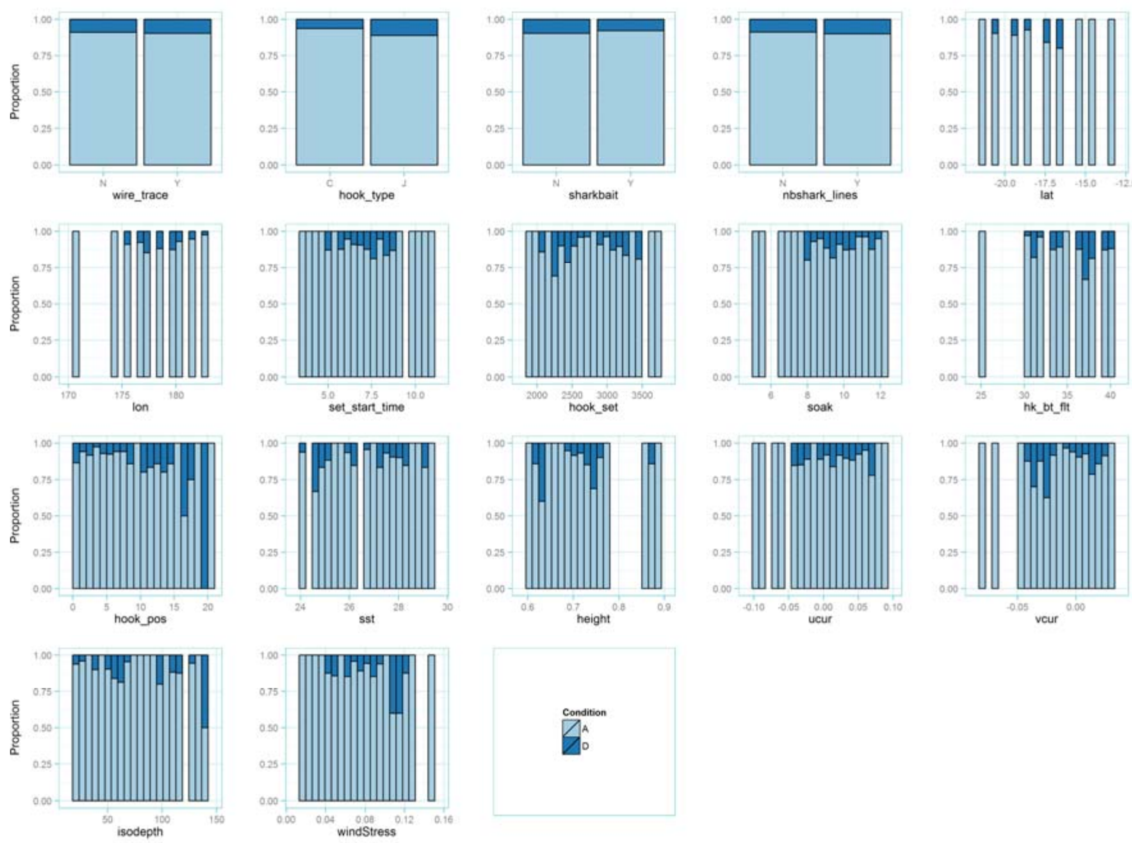


Figure 90: Distributions of condition responses (D= dead, A = alive), expressed in terms of proportion, against the explanatory variables considered for the analysis of blue sharks caught by the Fijian fleet in the M2_FV fishery (file “M2_FV_FJ_BSH_CondtnVsCovs.png”)

6.7.3.2 Modelling results

Modelling details are as per the reference case. The model returned an adjusted-R^2 of 0.31 and a deviance explained of 46.2%. Raw predictive power to the current dataset was also assessed via confusion matrices on the training data, providing Sensitivity (true positive rate) = 0.97, Specificity (true negative rate) = 0.75 and a misclassification rate of 0.23 when using the response mean as a decision boundary. These are against the training data and will be over-estimates.

Table 27: Model terms from the generalized additive model. DF = Degrees of Freedom. EDF=Effective Degrees of Freedom. Shrinkage has been applied as a means of model selection. Terms of little relevance have their EDF shrunken.

Model terms	EDF	DF	Chi or F-stat	*p*-value
sharkbait		1	1.13	0.2872
nbshark_lines		1	0.50	0.4814
wire_trace		1	6.18	0.0129
hook_type		1	7.34	0.0067
s(set_start_time)	3.20	9	3.33	0.0313
s(sst)	1.81	9	5.60	0.0000
s(height)	8.00	8	4.13	0.0721
s(ucur)	0.00	9	0.00	0.7606
s(vcur)	0.90	5	0.87	0.0670
s(isodepth)	0.00	9	0.00	0.0125
s(windStress)	0.00	9	0.00	0.6480
s(soak)	0.00	9	0.00	0.2680
s(hook_pos)	0.35	9	0.66	0.0299
s(hk_bt_flt)	0.00	9	0.00	0.2282
s(lon,lat)	13.28	27	9.21	0.0167

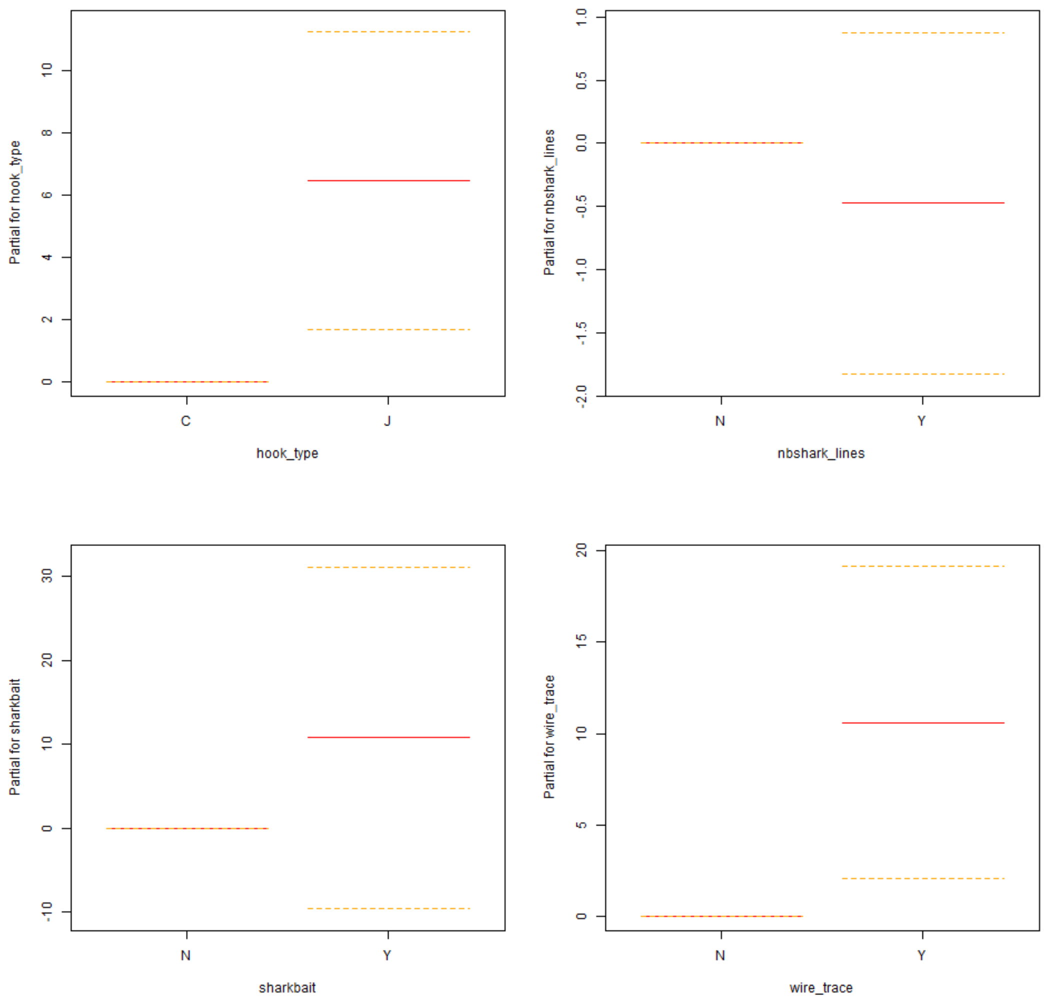


Figure 91: estimates and approximate 95% confidence intervals for parametric terms within the GAM

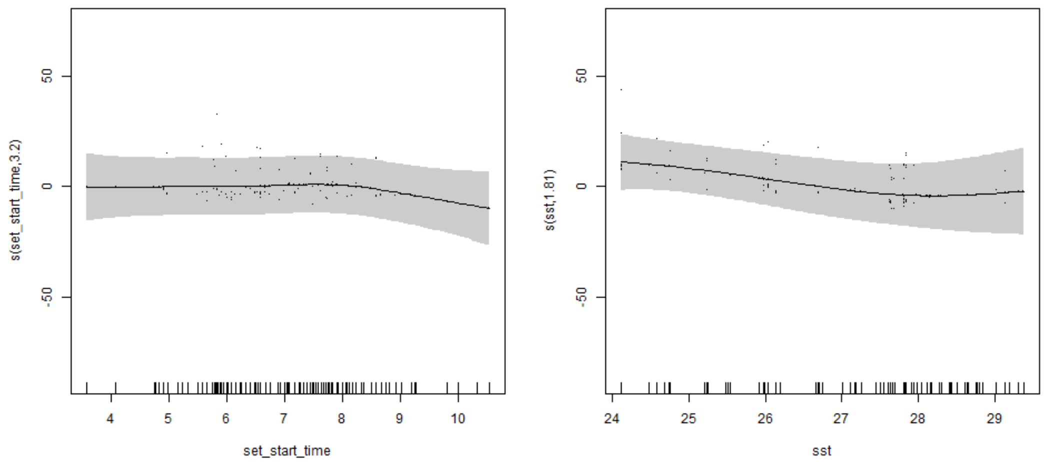


Figure 92: estimated smooths and approximate 95% confidence envelopes for smooth terms within the GAM.

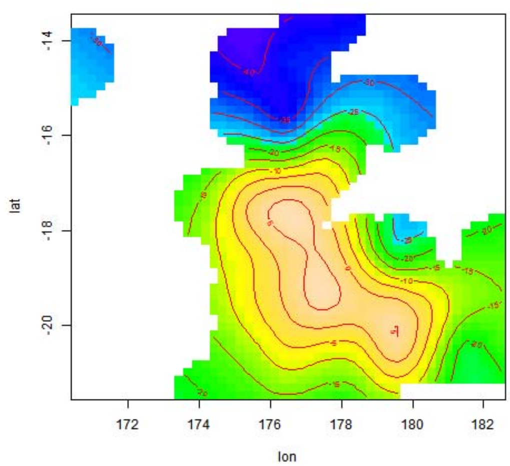


Figure 93: estimated smooth surfaces for the bivariate smooth components within the GAM.

6.8 US FLEET ANALYSIS – EXPLANATORY VARIABLES EDA

Off the key fishing features potentially driving the bycatch of sharks in the US fleet of the M2_FV fishery, hook type (with two categories, hook “C” and hook “T”) was the only factor that could be included in the present analysis (Figure 55). Moreover, to avoid confounding effects between years and the use of hook type “C”, the analysis was restricted to years 2010 and 2011.

There were one observation with a large of soak time (around 16 hours) that was considered to be too far from normal range of soak times (Figure 94) – this observation was excluded from the analysis. On the other hand, there were a few observations (111) with what could be judged as extreme low levels of average sea surface height (below 0.5 meters) (Figure 94). However, the cost of losing information by excluding a substantial amount of data points outweighed the potential influential nature of those observations (which can be assessed during the modelling phase), and therefore these observations were kept in the analysis. There were no other extreme values in the remaining covariates.

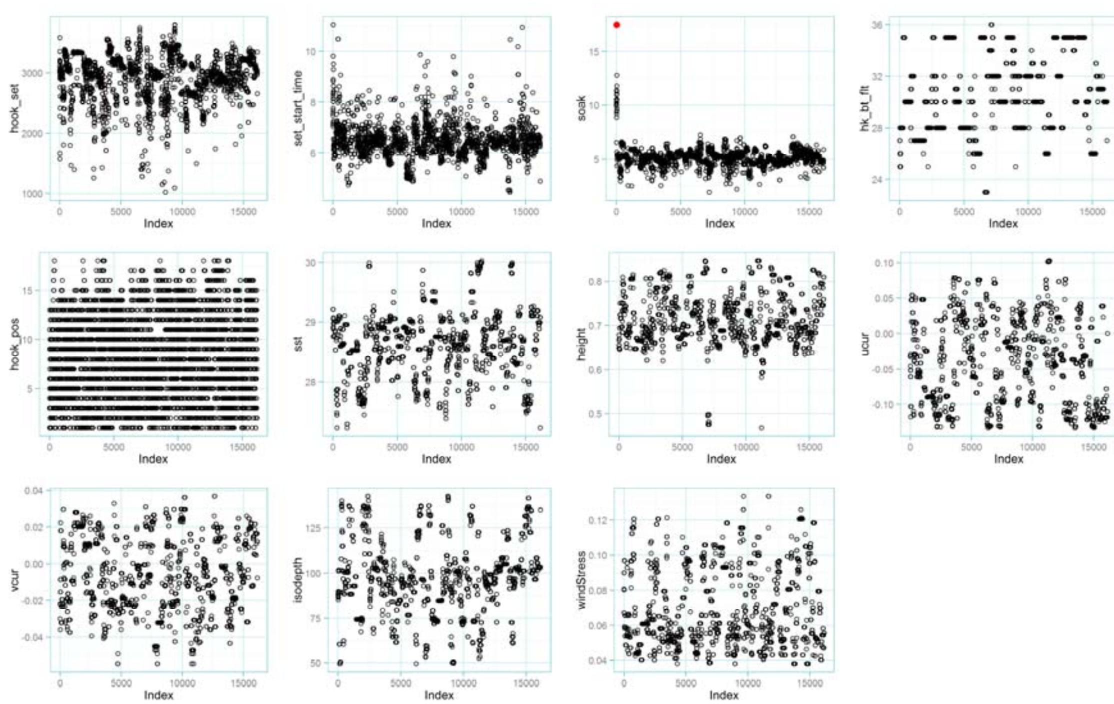


Figure 94: Cleveland dotplots of some of the explanatory variables considered for the analysis of shark bycach by the US fleet in the Fiji and American Samoa fishery. Plots constructed from data sorted by set starting date, and values deemed to be too extreme for the normal range of the covariate, and posteriorly excluded from the analysis, are coloured in red (File “M2_FV_US_Catch_covsClevPlots.png”)

Figure 95 shows the existence of mild negative correlation (-0.72) between latitude and average sea surface height, which can translate into a strong collinearity between these two covariates – collinearity was further investigated through the estimation of VIFs. The remaining covariates from observations used for the analysis of the Fijian fleet in the M2_FV fishery show generally weak to non-existent correlation between them.

Furthermore, several of the oceanographic variables (e.g. sea surface temperature, horizontal current speed and wind stress) appear to have non-linear relationships with month and latitude (Figure 95). Changes in the average sea surface temperature also appears to be followed by changes in average current speeds, average sea surface height and average depth of the 20°C isotherm.

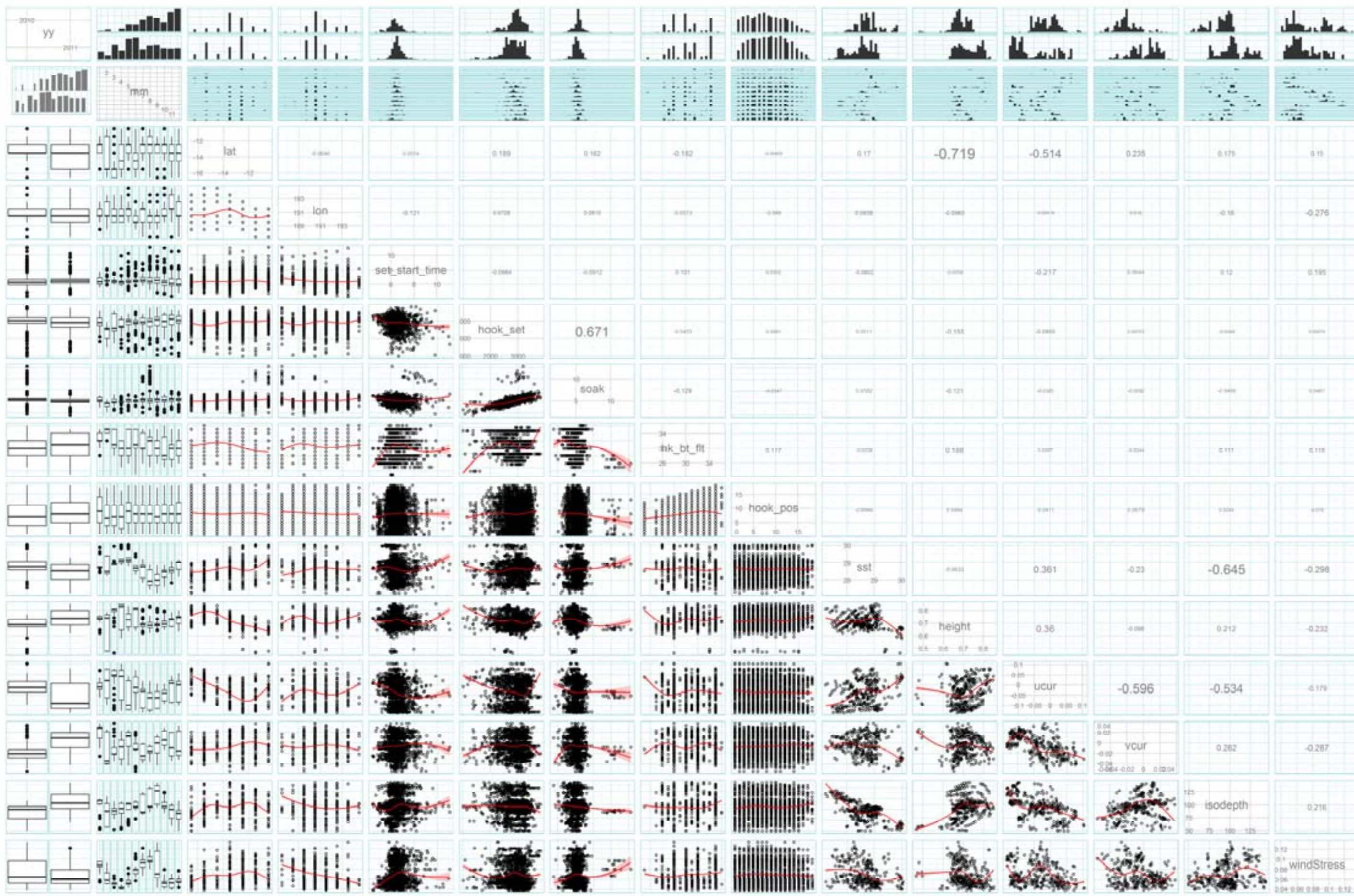


Figure 95: Matrixplot of relevant covariates considered for the analysis of shark bycatch by the US fleet in the Fiji and American Samoa fishery. Plots presented in panels depend on the data types under comparison. For example, pairs of continuous variables are displayed as scatterplots in the lower panels (with an added GAM-based smoother to help visualisation) and Pearson correlation coefficients in

the upper panels (with font size proportional to correlation coefficient). Continuous Vs Discrete covariates are represented by boxplots and barplots, while pairs of discrete covariates are displayed as fluctuation plots (lower panels) and barplots (upper panels) (file “M2_FV_US_CovsPairsPlot.png”

6.9 US FLEET ANALYSIS – MODELLING

Presented here are results for three models covering set-level catch-rates, catch-rates with respect to hook position and the condition of sharks (alive/dead) at retrieval. Four shark species are considered: Oceanic whitetips (OCS), silky (FAL), thresher (THR) and blue (BSH).

Details regarding modelling methods are presented previously, but all models fall within the Generalized Additive Modelling framework. The set of parametric and smoothing terms selected between are given in the following table – the specific model fitted for each species and model is presented in the relevant results section.

Table 28: Model terms considered. "s" indicates a smoothing term. ":" indicates an interaction between associated terms.

Parametric terms	Smooth terms
yy	s(set_start_time)
mm	s(sst)
hook_type	s(height)
wire_trace	s(ucur)
nbshark_lines	s(vcur)
hook_type:wire_trace	s(isodepth)
hook_type:nbshark_lines	s(windStress)
wire_trace:nbshark_lines	s(lat,lon)
hook_type:wire_trace:nbshark_lines	s(soak,hk_bt_flt)
	s(soak):nbshark_lines
	s(soak):sharkbait

Reference model – In the interests of brevity, one species' models are presented in more detail than the others. This will be referred to as the *reference case* and will be the Oceanic whitetip shark (OCS). In particular model diagnostics and non-significant results will be presented for this model, but not others. The same modelling approach has been applied for similar models, but only significant results (including assumption violations) will be reported upon.

6.10 US FLEET ANALYSIS – OCEANIC WHITETIP SHARKS (OCS) [REFERENCE CASE]

6.10.1 Catch rate of oceanic whitetip sharks [reference case]

6.10.1.1 Data Exploration

While predominantly absent in sets from observed trips of the US fleet in the M2_FV fishery (Figure 96), within-set bycatch levels of oceanic whitetip sharks ranged between 1 shark/set (in about 10 sets) and 16 sharks/set (in about 10 sets). However, in a few instances, up to 48 and 56 oceanic whitetips were observed in a single set.

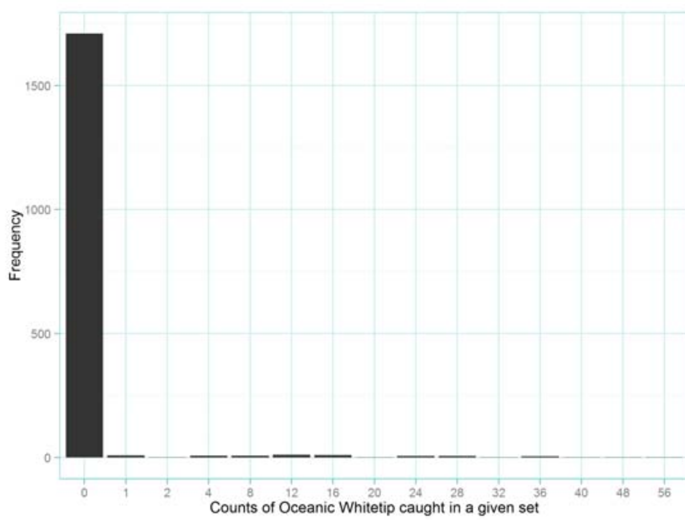


Figure 96: Distribution of counts per set of oceanic whitetip sharks caught in observed trips of the US fleet operating in M2_FV longline fishery (file M2_FV_US_OCS_CatchNumFreqs).

The predominance of zero catches and the relatively small number of specimens present the in observed sets hampers the investigation of bivariate patterns between catch rates of oceanic whitetip sharks and each of the covariates considered in the present analysis (Figure 97). However, it is still possible to verify that, at least at a graphical level, there are no clear relationships between catch rates of this species yielded by the US fleet operating in the M2_FV fishery and the available covariates.

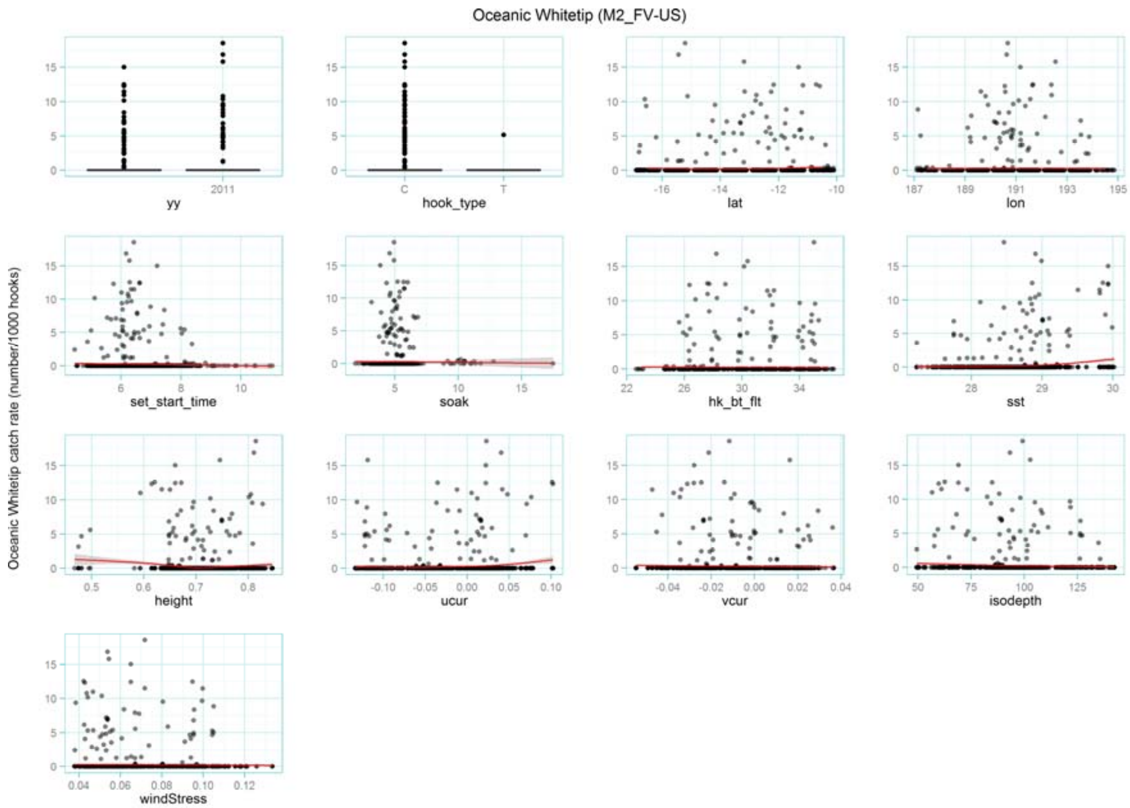


Figure 97: Distributions of set-level catch-rates against the explanatory variables considered for the analysis of oceanic whitetip sharks caught by the US fleet in the M2_FV fishery (file “M2_FV-US-OCS_CatchRatesVsCovs.png”)

6.10.1.2 Modelling results

Catch rate (per 1000 hooks) was modelled as a function of set-level characteristics and associated oceanographic variables. The initial set of potential covariates was thinned on the basis of gross collinearity, as determined by Generalized Variance Inflation Factors (GVIFs). A range of potential interactions and non-linear terms were specified *a priori* as per Table 3. These were further thinned during the modelling process in some cases where estimation was not possible e.g. interactions for factor variables where certain combinations of factors were not observed in the dataset.

The model fitted was a GAM with log-link and Tweedie-distributed errors. An initial model run was conducted for the purposes of estimating the governing Tweedie parameter – subsequently refitted with the estimated Tweedie parameter specified. Model "selection" was conducted within the GAM using shrinkage, so model terms of little relevance have their effective degrees of freedom shrunk towards zero. In terms of parametric components, backwards selection (p-value<0.05) was used, but not on the key factors of wire-trace, shark-bait, shark-lines and hook type (where applicable). These were generally retained in the model for interpretative interest. The results of this process are summarised in Table 4.

The Tweedie parameter was estimated to be 1.44, with the model returning an adjusted-R^2 of ~0 and a deviance explained of 22.5%.

Table 29: Model terms from the generalized additive model. DF = Degrees of Freedom. EDF=Effective Degrees of Freedom. Shrinkage has been applied as a means of model selection. Terms of little relevance have their EDF shrunken.

Model terms	EDF	DF	Chi or F-stat	*p*-value
yy		1	2.37	0.1240
hook_type		1	5.17	0.0231
s(lon,lat)	28.16	29	1.41	0.0589
s(soak,hk_bt_flt)	26.04	29	1.37	0.0411

The model assumptions were assessed. Non-linearities are assumed to be captured via the GAM which optimises complexity. The adequacy of the assumed Tweedie distribution was assessed via Quantile Residuals (Dunn & Smythe, 1996), as indicated in Figure 8 – these ought to be approximately Normally distributed, producing an approximately straight-line. Potential correlation in the errors was similarly assessed on quantile residuals, both by examination of acf plots and a runs test. The data were ordered by set date within trip IDs, so the correlations are sought via sets that are temporally, and likely spatially, close.

In this case the Tweedie distribution is acceptable and there is no significant residual correlation once this zero rich distribution has been accounted for. The runs-test returns a *p*-value of 0.77.

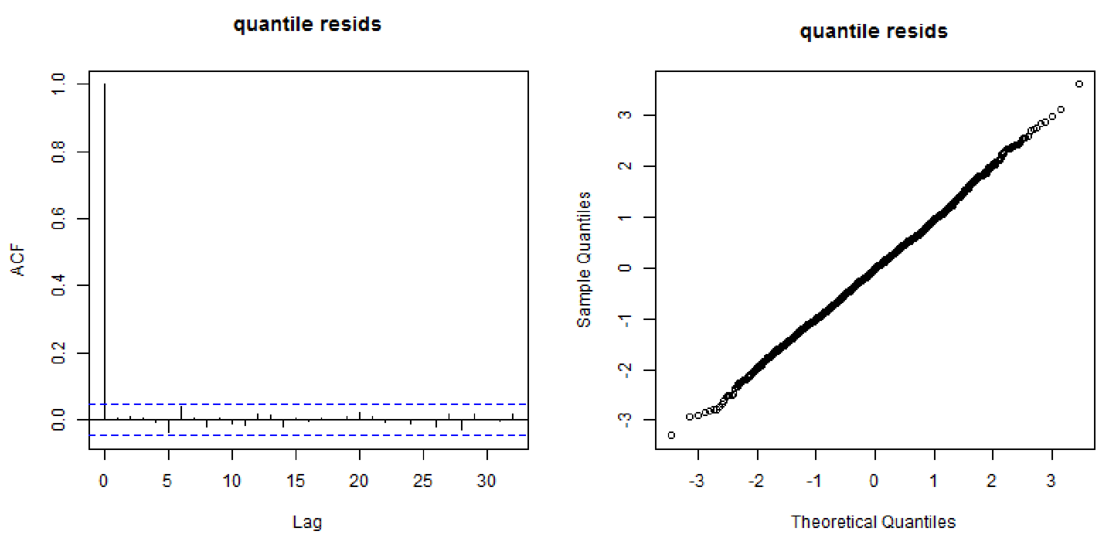


Figure 98: diagnostic plots for the quantile residuals from the GAM. The error distribution is assumed to be a Tweedie distribution, with key parameter estimated.

DMP STATS The Coach House, Mount Melville House, St. Andrews, Scotland KY 168 NT, UK +44(0) 1334 477 544 mail@dmpstats.com

The following are link-scale estimates of the components of the GAM. All parametric terms in the model are represented here regardless of practical/statistical significance. For smoothers, those with EDFs shrunken to zero are excluded as the plots are flat by definition. Generally only top-line results will be presented outside of the reference case. Where applicable, the parametric terms are presented first, followed by smooth terms and then interactions.

The following observations can be made:

- There is a significantly lower catch rate under hook type T compared to hook type C.
- Catches were on average lower in 2011 than 2010, but high uncertainty about the estimates.
- There is a significant spatial component to catch-rates which is not explained by the other covariates as evidenced by the significant latitude-longitude interaction surface. The pattern observed is complex.
- There is some catch relationship with both soak-time and the hooks between floats, with peak catch rates being with intermediate basket-sizes and soak times.

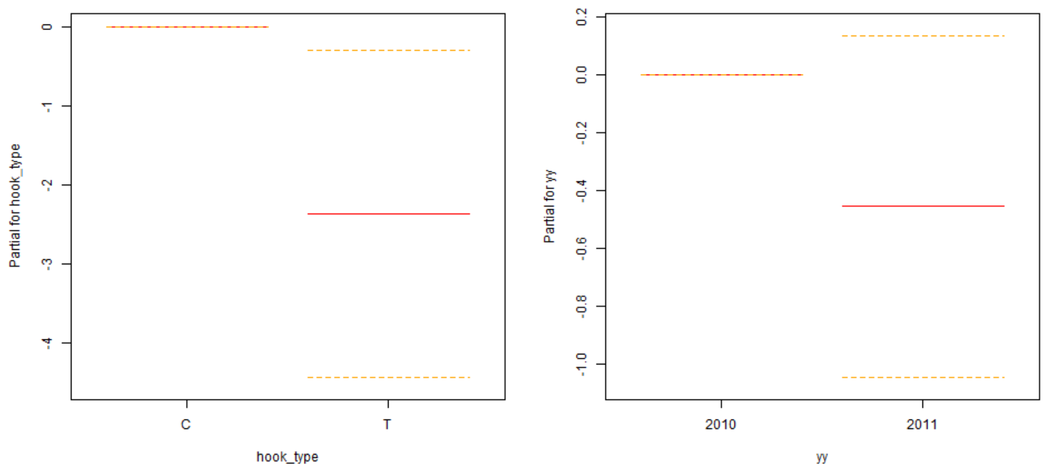


Figure 99: estimates and approximate 95% confidence intervals for parametric terms within the GAM.

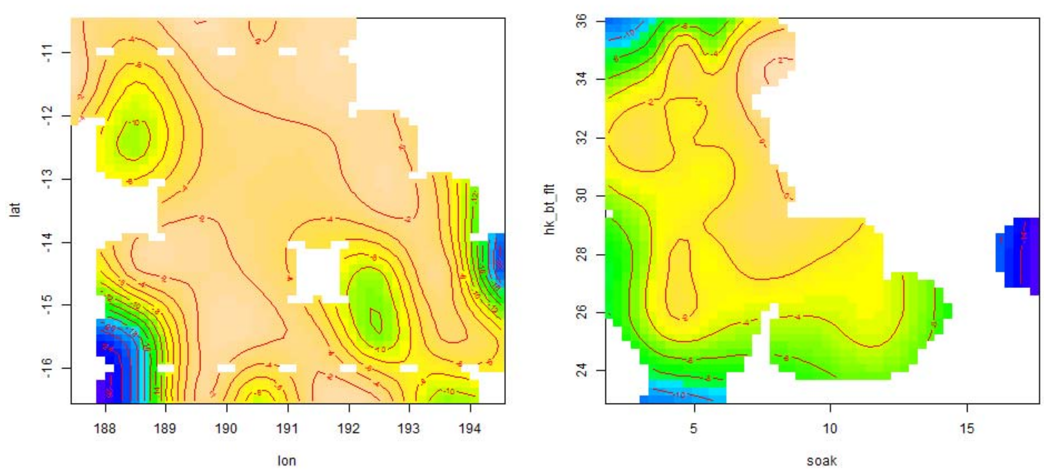


Figure 100: estimated smooth surfaces for the bivariate smooth components within the GAM.

6.10.2 Hook-level catch rate of oceanic whitetip sharks

6.10.2.1 Data Exploration

There are no clear patterns between catch rates of oceanic whitetip sharks and hook positions at which they were caught (Figure 101). Furthermore, there are no demarked trends in how the relationship between catch rates and hook positions changes with varying levels of the covariates considered in the analysis.

Figure 101: Distributions of per-hook catch-rate responses against the explanatory variables considered in the analysis, for oceanic whitetip sharks caught by the US fleet operating in the M7_FV fishery. Dots were slightly jittered to minimise overplotting and help visualisation and their colour convey the basket hook position of the respective observation (file "M2_FV-US-OCS_perHookRatesVsCovs.png").

6.10.2.2 Modelling results

The effect of hook position on catch rates was conditioned solely on species and fishery. A GAM was fitted with log-link and Tweedie error distribution. An initial model run was conducted for the purposes of estimating the governing Tweedie parameter – subsequently refitted with the estimated Tweedie parameter specified

The Tweedie parameter was estimated to be 1.109, with the model returning an adjusted-R^2 of ~0 and a deviance explained of 0.5%.

The model assumptions were assessed. Non-linearities are not relevant as the hook position is fitted as a factor variable. The adequacy of the assumed Tweedie distribution was assessed via Quantile Residuals as indicated in Figure 13– these ought to be approximately Normally distributed, producing an approximately straight-line. Potential correlation in the errors was similarly assessed on quantile residuals, both by examination of acf plots and a runs test. The data were ordered by set date within trip IDs, so the correlations are sought via sets that are temporally, and likely spatially, close.

In this case the Tweedie distribution is acceptable and there is no significant residual correlation once this zero rich distribution has been accounted for. The runs-test returns a *p*-value of 0.088.

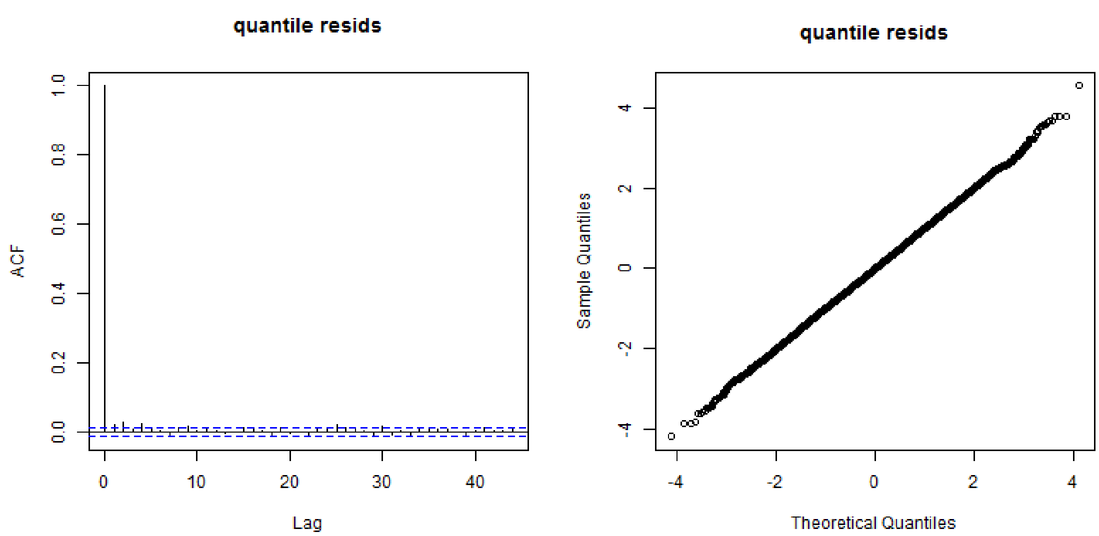


Figure 102: diagnostic plots for the quantile residuals from the GAM. The error distribution is assumed to be a Tweedie distribution, with key parameter estimated.

Estimated catch rate for each hook position and approximate 95% confidence intervals are presented in Figure 14 and Table 5.

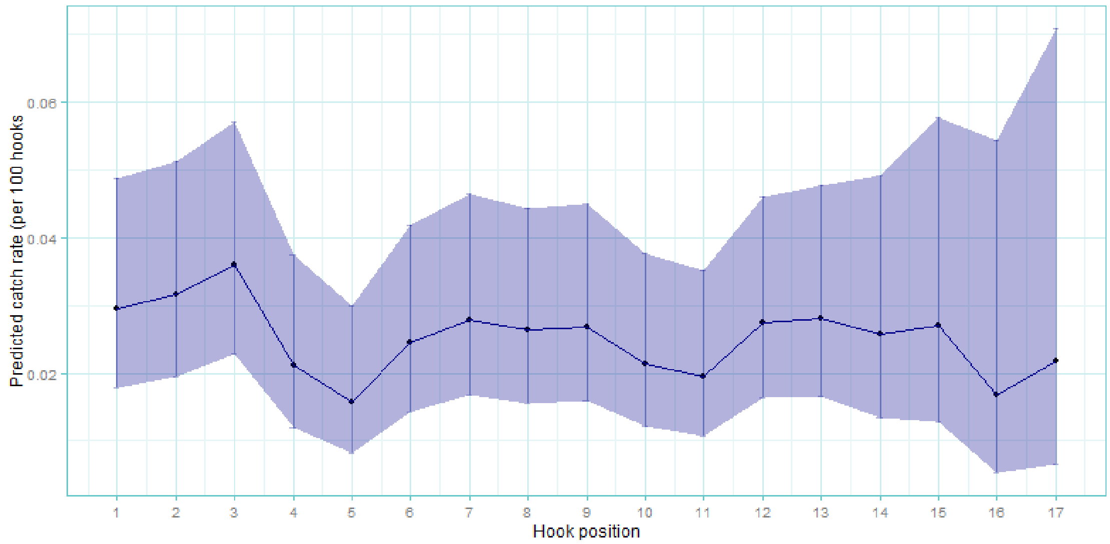


Figure 103: estimated mean catch rate (per 100 hooks) at hook-position with approximate 95% confidence envelope.

Table 30: Estimates of catch rates for each of the hook positions. Confidence intervals are approximate 95%. Percentage of total catch is calculated from overall estimated catch over hook positions presented. Confidence intervals for percentages are based on this naïve translation of cate rate.

Hook position	Catch rate (per 100 hooks)	SE	Lower CI	Upper CI	Catch (as % of total)	Lower CI (as % of total)	Upper CI (as % of total)
1	0.0296	1.2898	0.0179	0.0487	6.93	4.21	11.4
2	0.0316	1.2811	0.0194	0.0513	7.4	4.55	12.03
3	0.0361	1.2643	0.0228	0.0572	8.47	5.35	13.41
4	0.0211	1.3392	0.0119	0.0374	4.94	2.78	8.75
5	0.0157	1.3905	0.0082	0.0299	3.67	1.92	7
6	0.0244	1.3169	0.0142	0.0418	5.71	3.33	9.79
7	0.0278	1.2981	0.0167	0.0464	6.52	3.91	10.87
8	0.0263	1.306	0.0156	0.0444	6.16	3.65	10.39
9	0.0267	1.3039	0.0159	0.0449	6.25	3.72	10.52
10	0.0213	1.3373	0.0121	0.0377	5	2.83	8.83
11	0.0194	1.3524	0.0108	0.0351	4.55	2.52	8.23

12	0.0275	1.3002	0.0164	0.046	6.44	3.85	10.77
13	0.0281	1.3125	0.0165	0.0478	6.58	3.86	11.21
14	0.0258	1.3904	0.0135	0.0492	6.05	3.17	11.53
15	0.0271	1.4725	0.0127	0.0578	6.34	2.97	13.54
16	0.0168	1.8247	0.0052	0.0545	3.93	1.21	12.78
17	0.0217	1.8327	0.0066	0.071	5.08	1.55	16.64

The following can be observed from these:

- There is no particular relationship between the hook-positions and the catch rates.
- The shallowest position (1) is estimated at 0.03 per 100 hooks [0.018 to 0.049, 95% CI]. This translates to approximately 6.9% [4.21% to 11.4%, 95% CI] of the catch of this species observed in this fisheries data.

6.10.3 Condition of oceanic whitetip sharks at time of retrieval [reference case]

6.10.3.1 Data Exploration

There are no strong patterns between the condition of caught oceanic whitetip sharks at the time of retrieval in observed trips of the US fleet operating in the M2_FV fishery and the covariates considered in the present analysis (Figure 104).

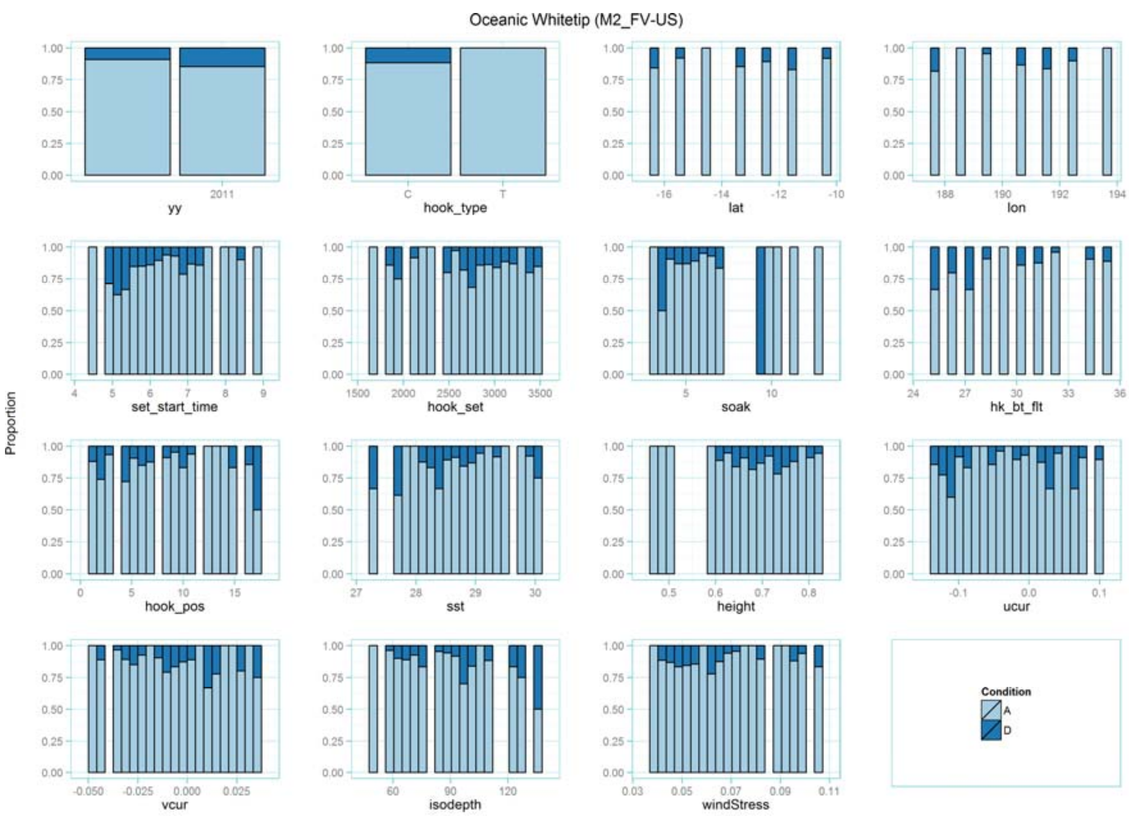


Figure 104: Distributions of condition responses (D= dead, A = alive), expressed in terms of proportion, against the explanatory variables considered for the analysis of oceanic whitetip sharks caught by the US fleet in the M2_FV fishery (file "M2_FV-US-OCS_CondtnVsCovs.png")

6.10.3.2 Modelling results

Condition (dead/alive) was modelled as a function of set-level characteristics and associated oceanographic variables. The initial set of potential covariates was thinned on the basis of gross collinearity, as determined by Generalized Variance Inflation Factors (GVIFs). A range of potential interactions and non-linear terms were specified *a priori* as per Table 3. These were further thinned during the modelling process in some cases where estimation was not possible e.g. interactions for factor variables where certain combinations of factors were not observed in the dataset.

The model fitted was a GAM with logit-link and binomially-distributed errors. Model "selection" was conducted within the GAM using shrinkage, so model terms of little relevance have their effective degrees of freedom shrunk towards zero. In terms of parametric components, backwards selection (p-value<0.05) was used, but not on the key factors of wire-trace, shark-bait, shark-lines and hook type (where applicable). These were generally retained in the model for interpretative interest. The results of this process are summarised in Table 4.

In terms of fit to the data, the software generated adjusted R^2 and deviance explained were 0.306 and 40.2% respectively, indicating poor fit. However these measures are of dubious utility for binomial error models. Raw predictive power to the current dataset was also assessed via confusion matrices, providing Sensitivity (true positive rate) = 0.91, Specificity (true negative rate) = 0.72 and a misclassification rate of 0.26 when using the response mean as a decision boundary. These are against the training data and will be over-estimates –cross-validated estimates would be preferred with future analysis.

Table 31: Model terms from the generalized additive model. DF = Degrees of Freedom. EDF=Effective Degrees of Freedom. Shrinkage has been applied as a means of model selection. Terms of little relevance have their EDF shrunken.

Model terms	EDF	DF	Chi or F-stat	*p*-value
yy		1	0.01	0.9329
hook_type		1	0.00	1.0000
s(set_start_time)	1.43	8	0.60	0.0942
s(sst)	0.00	9	0.00	0.0029
s(height)	0.00	9	0.00	0.1083
s(ucur)	1.97	9	2.28	0.0002
s(vcur)	3.37	8	3.54	0.0010
s(isodepth)	8.00	9	7.04	0.0001
s(windStress)	1.71	8	1.44	0.0154
s(hook_pos)	8.00	9	22.31	0.0000
s(soak,hk_bt_flt)	20.04	27	10.97	0.0122

The model assumptions were assessed. Non-linearities are assumed to be captured via the GAM which optimises complexity. The adequacy of the assumed binomial distribution was assessed via Quantile Residuals (Dunn & Smythe, 1996), as indicated in Figure 16 – these ought to be approximately Normally distributed, producing an approximately straight-line. Potential correlation in the errors was similarly assessed on quantile residuals, both by examination of acf plots and a runs test. The data were ordered by set date within trip IDs, so the correlations are sought via sets that are temporally, and likely spatially, close. The residuals are distributionally acceptable and little autocorrelation is in evidence, the runs-test providing a *p*-value of 0.926.

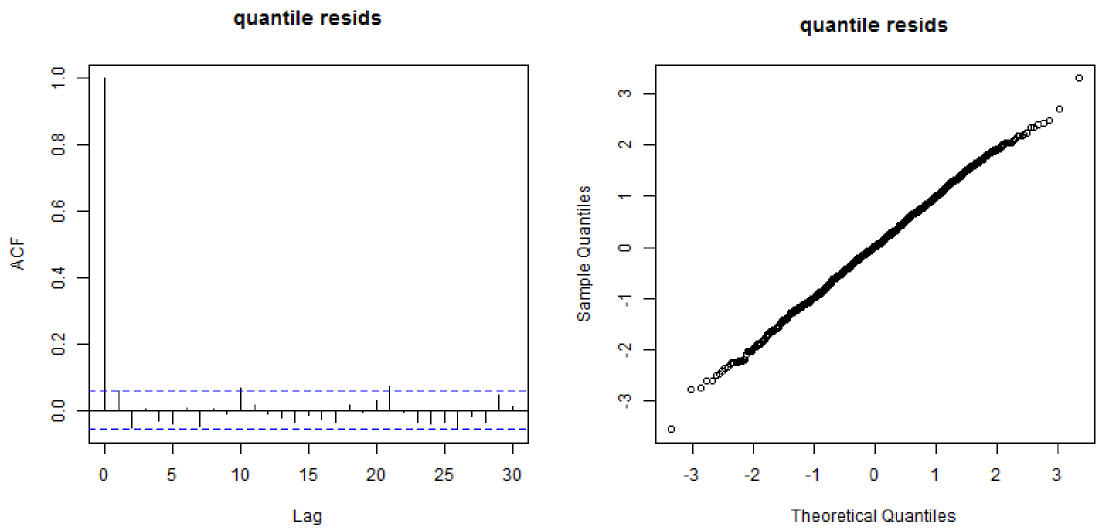


Figure 105: diagnostic plots for the quantile residuals from the GAM. The error distribution is assumed to be a Tweedie distribution, with key parameter estimated.

The following are link-scale estimates of the components of the GAM. All terms in the model are represented here regardless of practical/statistical significance although near-zero EDF relationships are excluded. Generally only top-line results will be presented outside of the reference case. Where applicable, the parametric terms are presented first, followed by smooth terms and then interactions.

The following observations can be made:

- Little significant pattern with regards the key factors of wire-trace, shark-bait and shark-lines.
- A complex non-monotone relationship with isodepth.
- Generally increasing morality rates with current (u).
- Generally increasing morality rates with current (v) and wind-stress.
- A marked dip in mortality with hook positions around 12-14.
- Higher mortality with lower numbers of hooks in baskets.

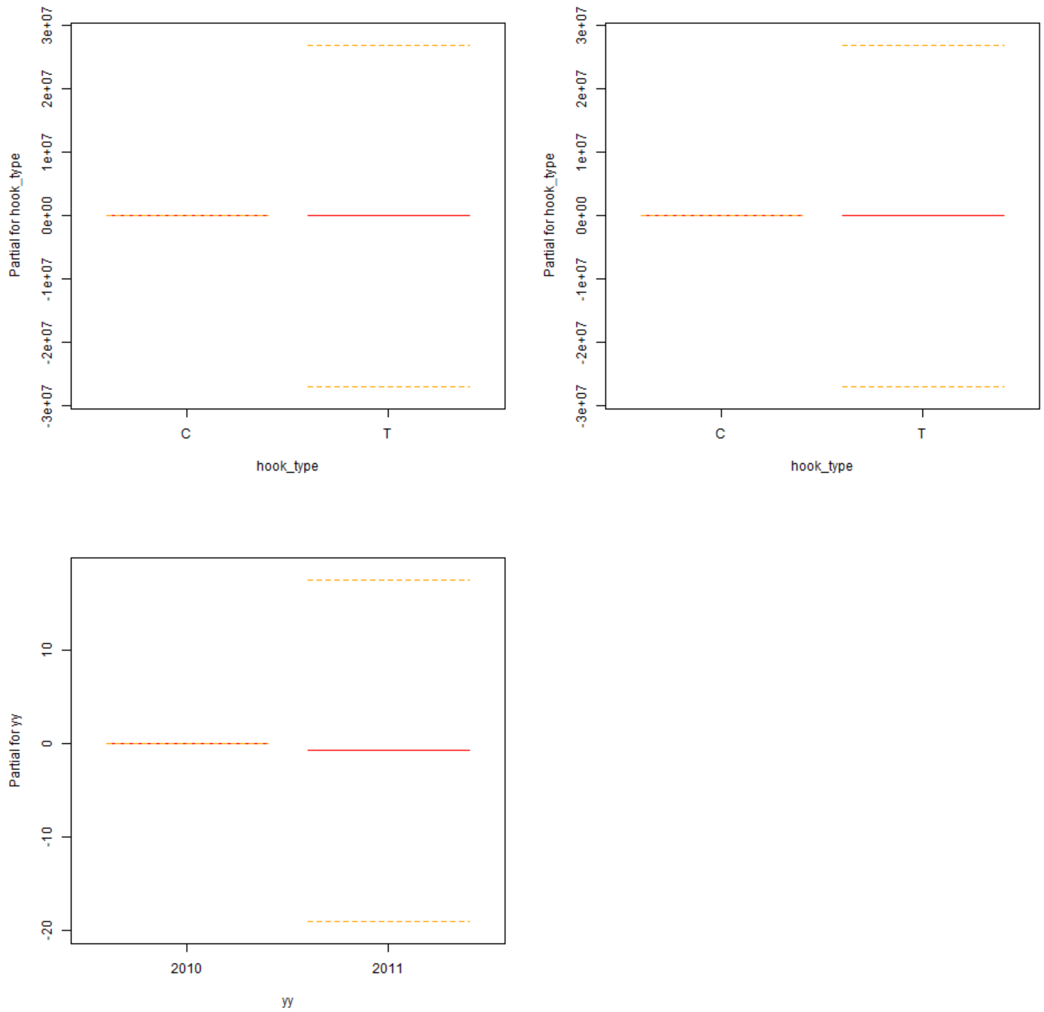


Figure 106: estimates and approximate 95% confidence intervals for parametric terms within the GAM.

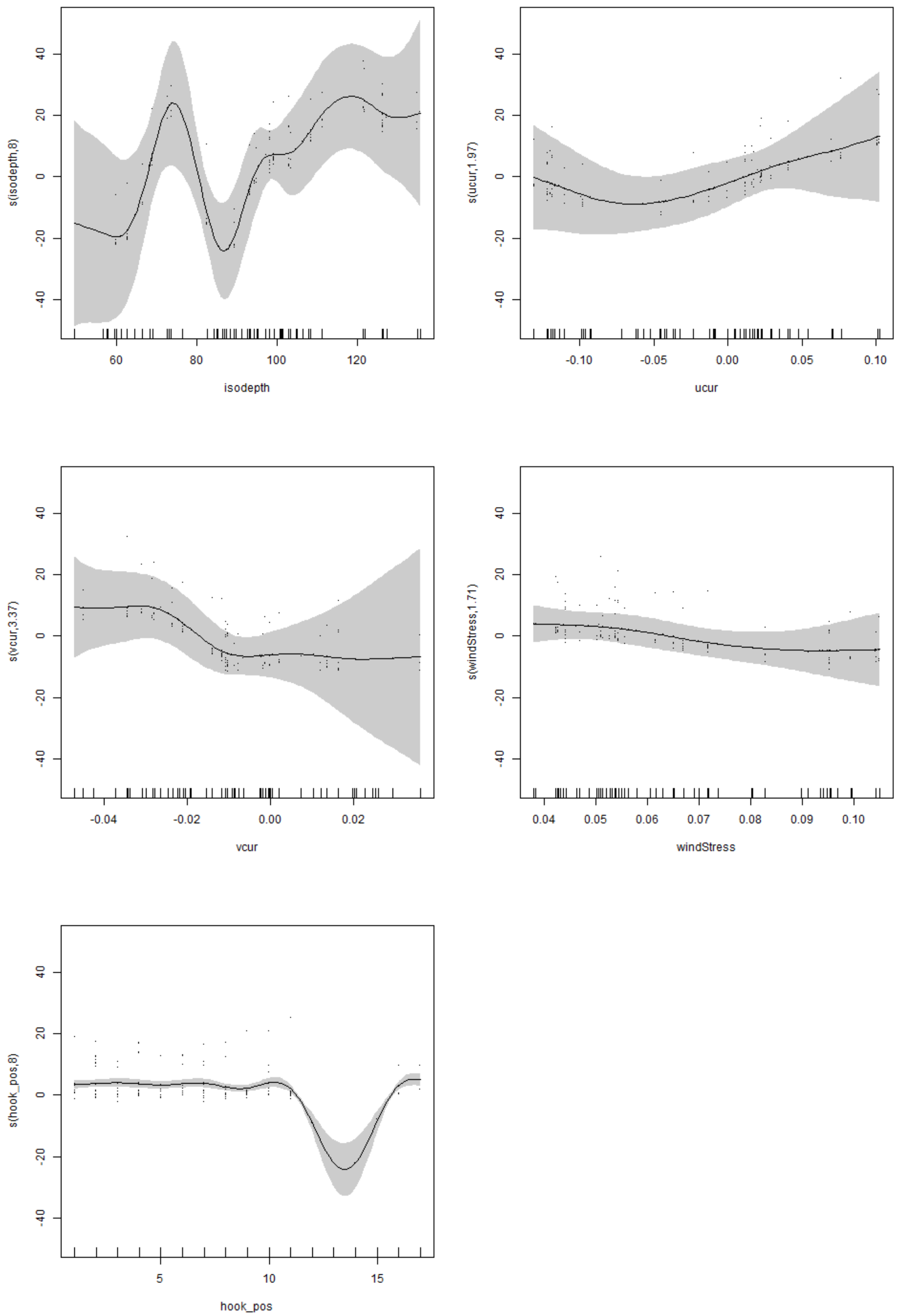


Figure 107: estimated smooths and approximate 95% confidence envelopes for smooth terms within the GAM.

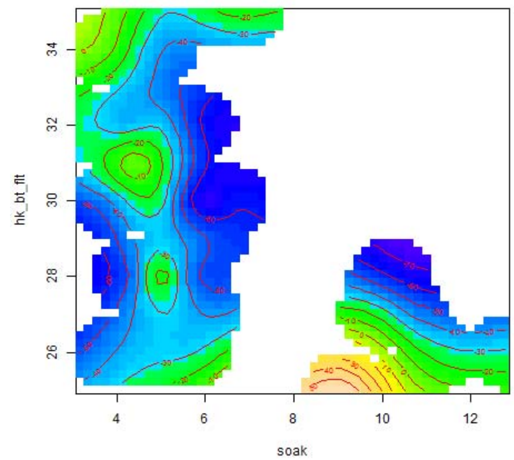


Figure 108: estimated smooth surfaces for the bivariate smooth components within the GAM.

6.11 US FLEET ANALYSIS – SILKY SHARKS (FAL)

6.11.1 Catch rate of Silky Sharks (FAL)

6.11.1.1 Data Exploration

While predominantly absent in sets from observed trips of the US fleet in the M2_FV fishery (Figure 109), within-set bycatch levels of silky sharks ranged between 4 shark/set (in about 30 sets) and 32 sharks/set (in about 10 sets). However there were up to a maximum of 84 oceanic whitetips in one single set.

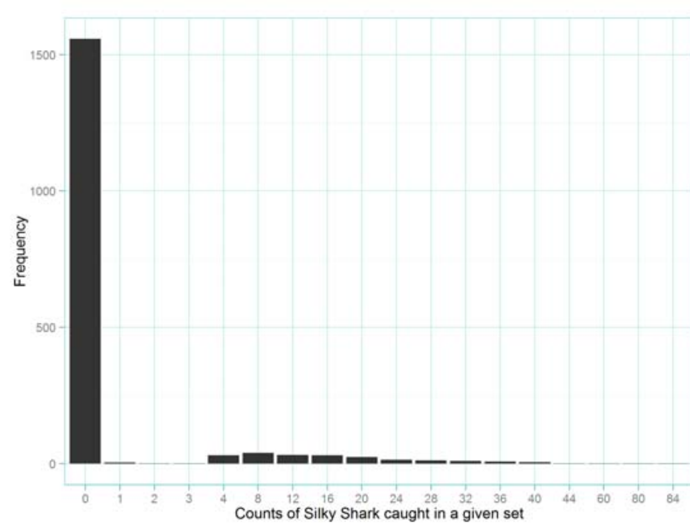


Figure 109: Distribution of counts per set of silky sharks caught in observed trips of the US fleet operating in M2_FV longline fishery (file M2_FV_US_FAL_CatchNumFreqs).

The predominance of zero catches in observed sets hampers the investigation of patterns between catch rates of silky sharks and each of the covariates considered in the present analysis (Figure 110). However, it is still possible to verify that, at least at a graphical and bivariate level, there are no clear relationships between catch rates of this species yielded by the US fleet operating in the M2_FV fishery and the values of the available covariates.

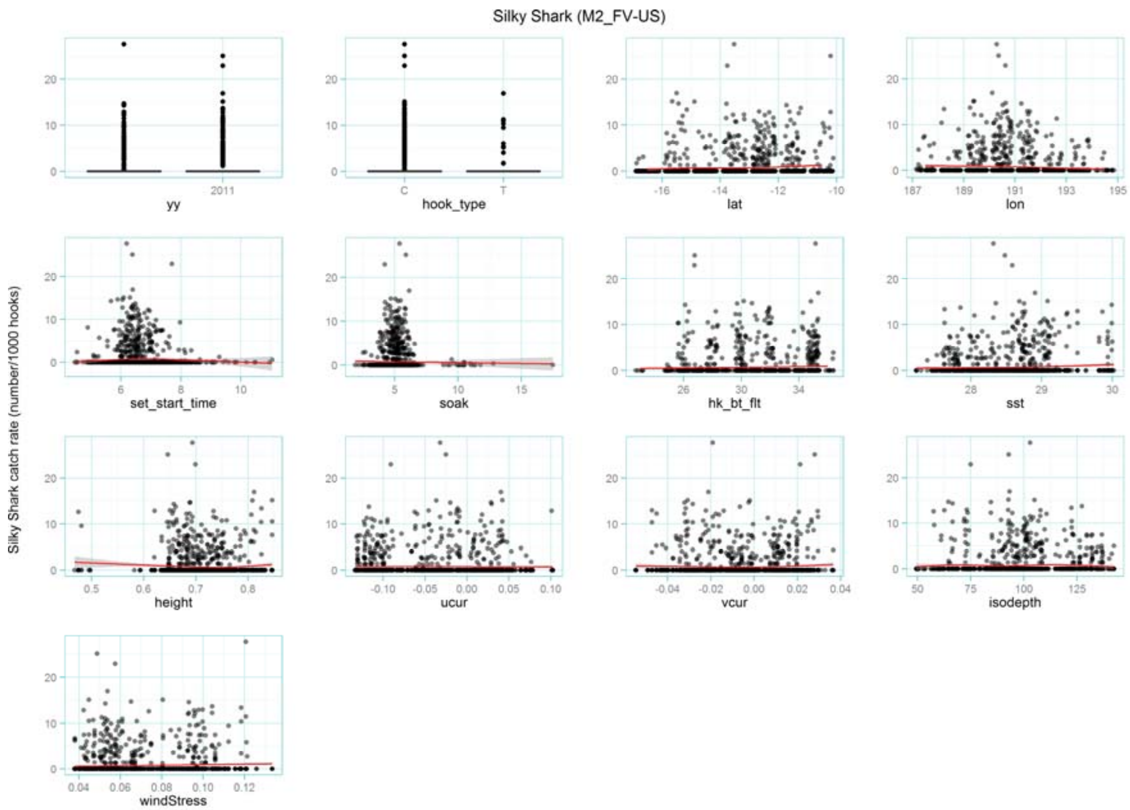


Figure 110: Distributions of set-level catch-rates against the explanatory variables considered for the analysis of oceanic silky caught by the US fleet in the M2_FV fishery (file “M2_FV-US-FAL_CatchRatesVsCovs.png”)

6.11.1.2 Modelling results

Modelling details are as per the reference case. The model returned an adjusted-R^2 of 0.03 and a deviance explained of 15.7%.

Table 32: Model terms from the generalized additive model. DF = Degrees of Freedom. EDF=Effective Degrees of Freedom. Shrinkage has been applied as a means of model selection. Terms of little relevance have their EDF shrunken.

Model terms	EDF	DF	Chi or F-stat	*p*-value
yy		1	2.32	0.1279
hook_type		1	3.22	0.0728
s(set_start_time)	6.19	9	1.59	0.0216
s(sst)	0.00	9	0.00	1.0000
s(height)	7.55	9	2.34	0.0029
s(ucur)	0.59	9	0.10	0.2024
s(vcur)	6.93	9	1.44	0.0502
s(isodepth)	0.00	9	0.00	1.0000
s(windStress)	8.59	9	2.12	0.0135
s(lon,lat)	13.45	29	1.14	0.0003
s(soak,hk_bt_flt)	23.87	29	1.40	0.0119

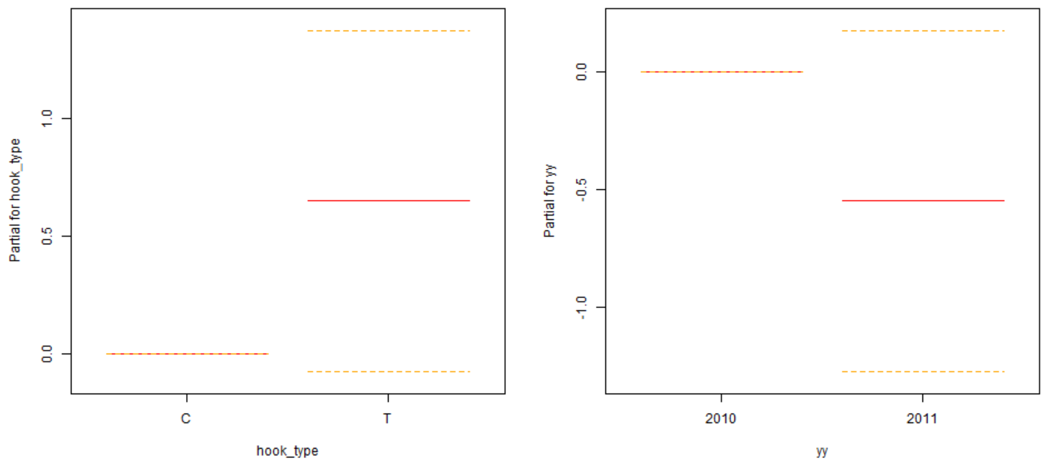


Figure 111: estimates and approximate 95% confidence intervals for parametric terms within the GAM.

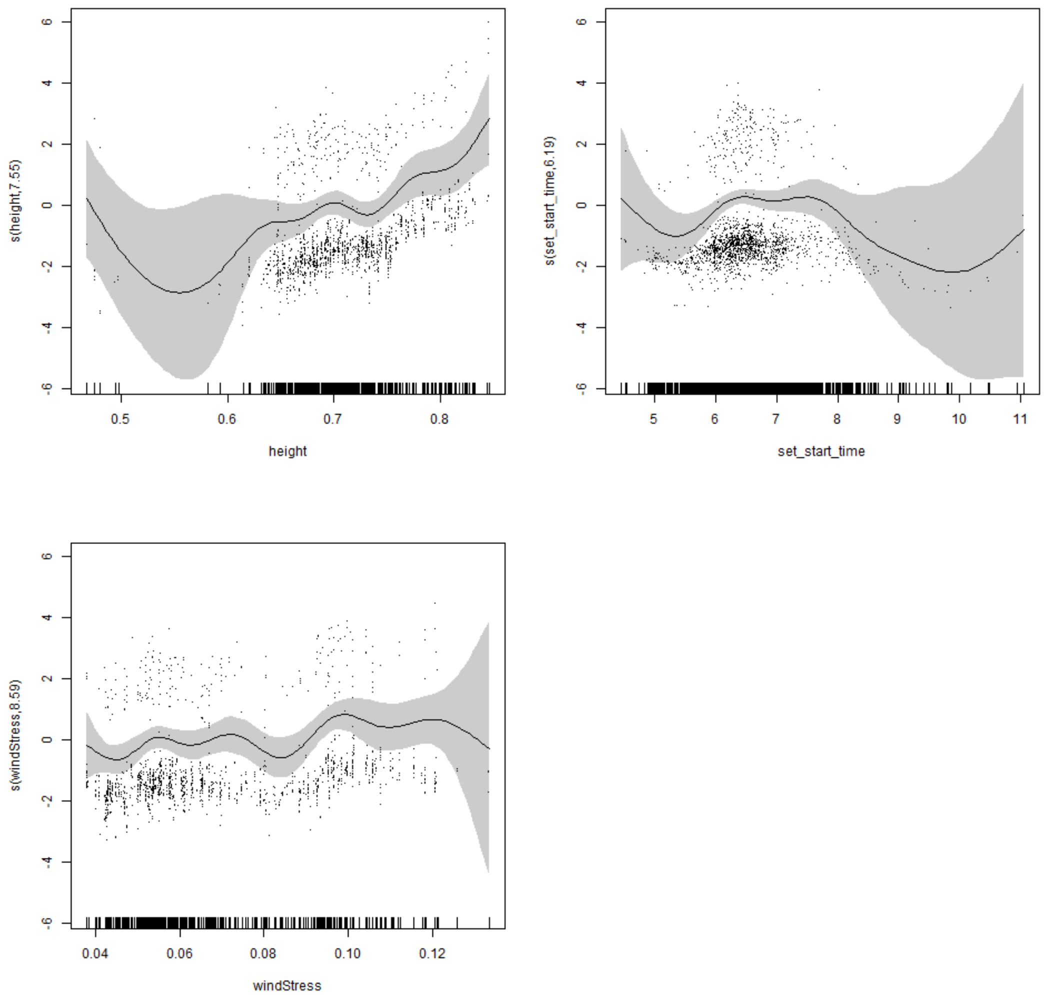


Figure 112: estimated smooths and approximate 95% confidence envelopes for smooth terms within the GAM.

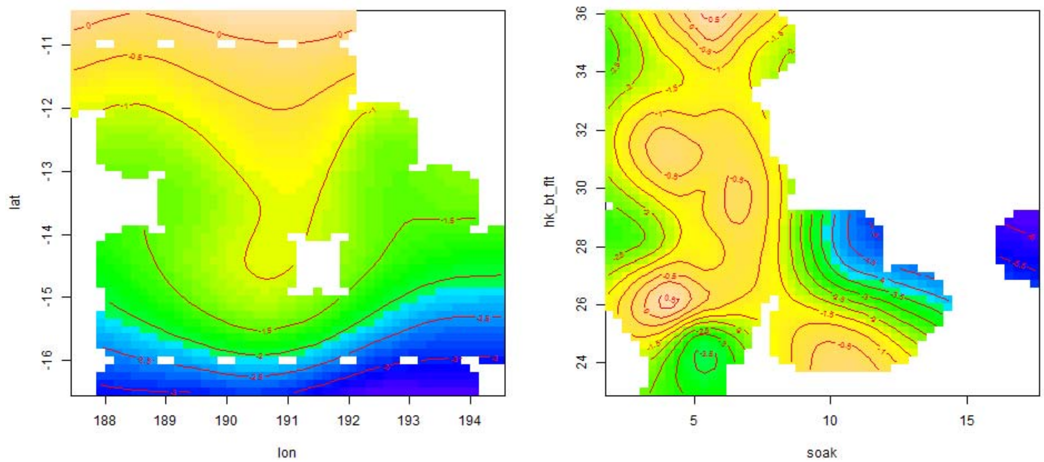


Figure 113: estimated smooth surfaces for the bivariate smooth components within the GAM.

6.11.2 Hook-level catch rate of silky sharks

6.11.2.1 Data Exploration

Figure 114 shows that catch rates of silky sharks in the US fleet tend to be larger at higher hook position numbers (i.e. at deeper hooks). However, there are no demarked trends in how the relationship between catch rates and hook positions changes with varying levels of the covariates considered in the present analysis.

Figure 114: Distributions of per-hook catch-rate responses against the explanatory variables considered in the analysis, for silky sharks caught by the US fleet operating in the M7_FV fishery. Dots were slightly jittered to minimise overplotting and help visualisation and their colour convey the basket hook position of the respective observation (file "M2_FV-US-FAL_perHookRatesVsCovs.png").

6.11.2.2 Modelling results

Modelling details are as per the reference case. The model returned an adjusted-R^2 of ~0 and a deviance explained of 0.91%.

Table 33: Estimates of catch rates for each of the hook positions. Confidence intervals are approximate 95%. Percentage of total catch is calculated from overall estimated catch over hook positions presented. Confidence intervals for percentages are based on this naïve translation of cate rate.

Hook position	Catch rate (per 100 hooks)	SE	Lower CI	Upper CI	Catch (as % of total)	Lower CI (as % of total)	Upper CI (as % of total)
1	0.0612	1.1838	0.0439	0.0851	4.1	2.95	5.71
2	0.0921	1.1516	0.0699	0.1215	6.18	4.69	8.15
3	0.0757	1.1662	0.056	0.1023	5.08	3.76	6.86
4	0.08	1.162	0.0596	0.1073	5.37	4	7.2
5	0.0773	1.1646	0.0573	0.1042	5.18	3.85	6.99
6	0.074	1.168	0.0546	0.1003	4.97	3.66	6.73
7	0.0818	1.1602	0.0612	0.1095	5.49	4.1	7.35
8	0.075	1.1669	0.0554	0.1015	5.03	3.72	6.81
9	0.0771	1.1647	0.0572	0.104	5.18	3.84	6.98
10	0.0686	1.1741	0.0501	0.094	4.6	3.36	6.3
11	0.0719	1.1703	0.0528	0.0978	4.82	3.54	6.56
12	0.0831	1.1593	0.0622	0.111	5.57	4.17	7.45
13	0.0874	1.1631	0.065	0.1175	5.86	4.36	7.88
14	0.0977	1.1837	0.0702	0.1359	6.55	4.71	9.12
15	0.1657	1.1739	0.121	0.2269	11.12	8.12	15.22
16	0.0411	1.4569	0.0197	0.086	2.76	1.32	5.77
17	0.1808	1.247	0.1173	0.2787	12.13	7.87	18.7

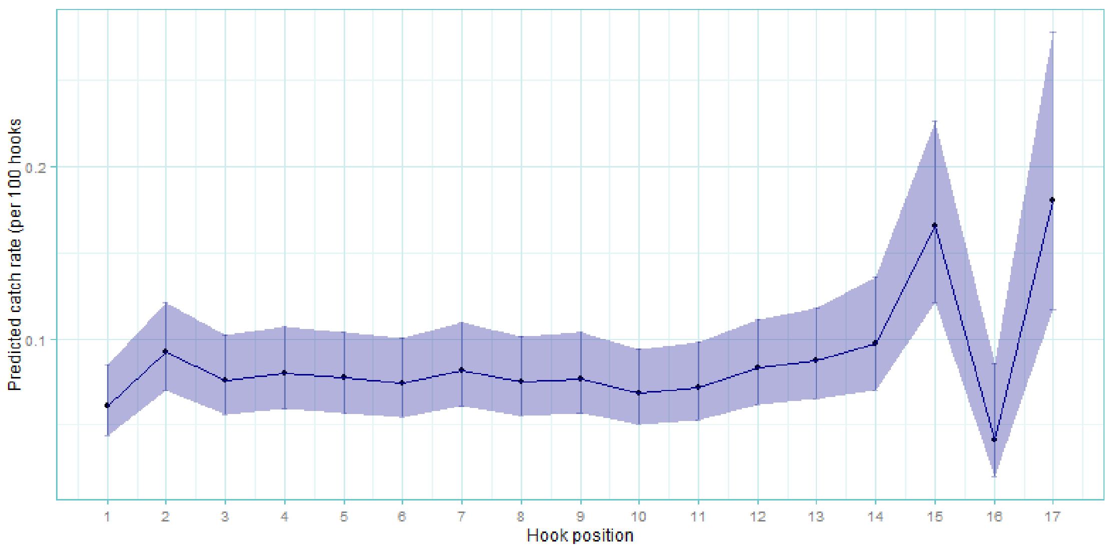


Figure 115: estimated mean catch rate (per 100 hooks) at hook-position with approximate 95% confidence envelope.

6.11.3 Condition of silky sharks at time of retrieval

6.11.3.1 Data Exploration

There are no strong patterns between the condition of caught silky sharks at the time of retrieval and most of the covariates considered in the present analysis (Figure 116). There are however sparse indications that the proportion of dead sharks might have increased when observed sets were performed under certain conditions, like e.g. (i) when hook type "C" was employed instead of hook type "T"; and (ii) if stets were performed when average sea surface temperatures were outside the region of 28.5^{o}C.

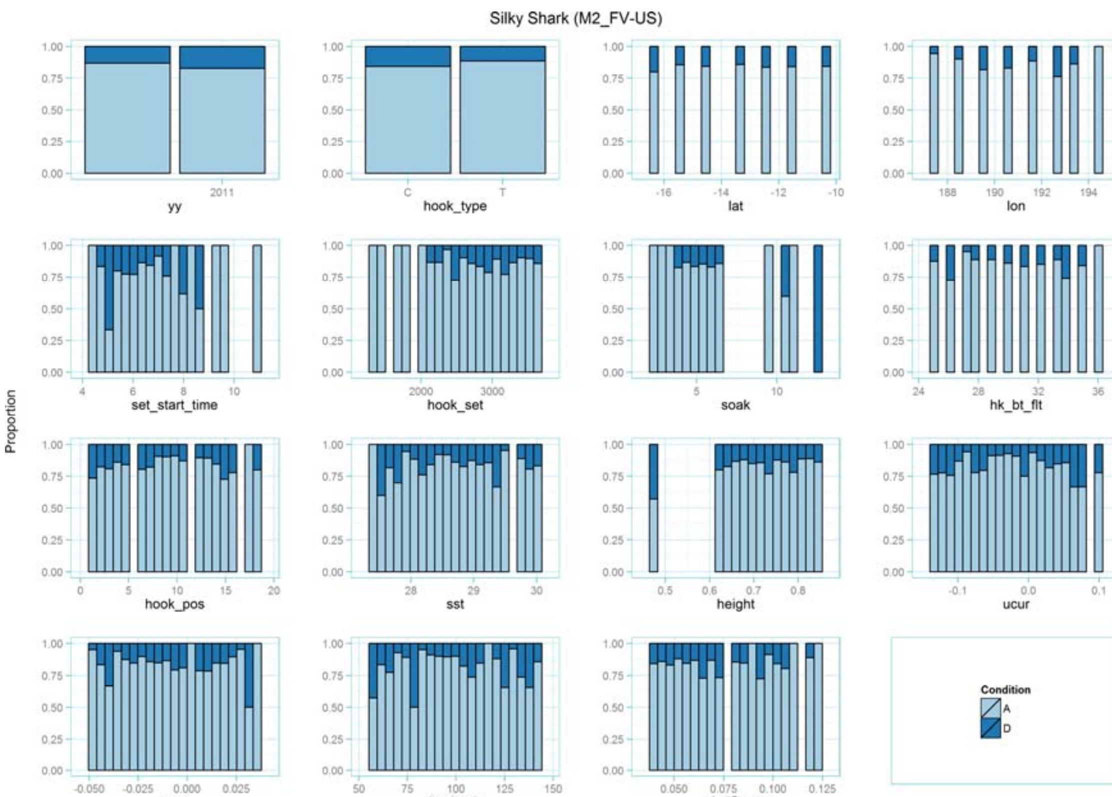


Figure 116: Distributions of condition responses (D= dead, A = alive), expressed in terms of proportion, against the explanatory variables considered for the analysis of silky sharks caught by the US fleet in the M2_FV fishery (file "M2_FV-US-FAL_CondtnVsCovs.png")

6.11.3.2 Modelling results

Modelling details are as per the reference case. The model returned an adjusted-R^2 of 0.23 and a deviance explained of 27.4%. Raw predictive power to the current dataset was also assessed via confusion matrices on the training data, providing Sensitivity (true positive rate) = 0.78, Specificity (true negative rate) = 0.72 and a misclassification rate of 0.26 when using the response mean as a decision boundary. These are against the training data and will be over-estimates.

Table 34: Model terms from the generalized additive model. DF = Degrees of Freedom. EDF=Effective Degrees of Freedom. Shrinkage has been applied as a means of model selection. Terms of little relevance have their EDF shrunken.

Model terms	EDF	DF	Chi or F-stat	*p*-value
yy		1	8.88	0.0029
hook_type		1	2.15	0.1427
s(set_start_time)	5.68	9	21.72	0.0001
s(sst)		8.00	9	53.32
s(height)	8.56	9	45.46	0.0000
s(ucur)	8.00	9	42.37	0.0000
s(vcur)	8.00	9	53.03	0.0000
s(isodepth)	8.00	8	51.84	0.0000
s(windStress)	5.09	9	28.23	0.0000
s(soak,hook_pos)	28.43	29	116.20	0.0000
s(lat,lon)	25.87	27	94.08	0.0000
s(soak,hk_bt_flt)	26.58	28	96.94	0.0000

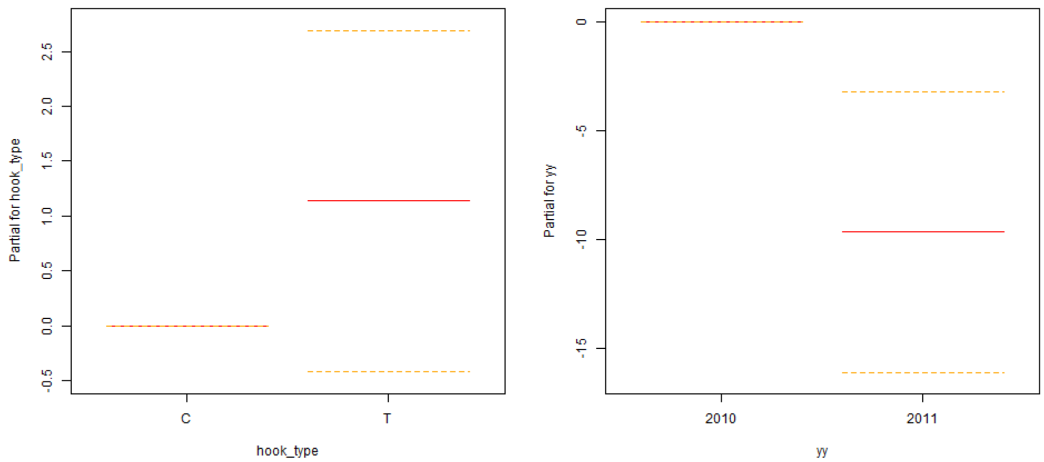


Figure 117: estimates and approximate 95% confidence intervals for parametric terms within the GAM.

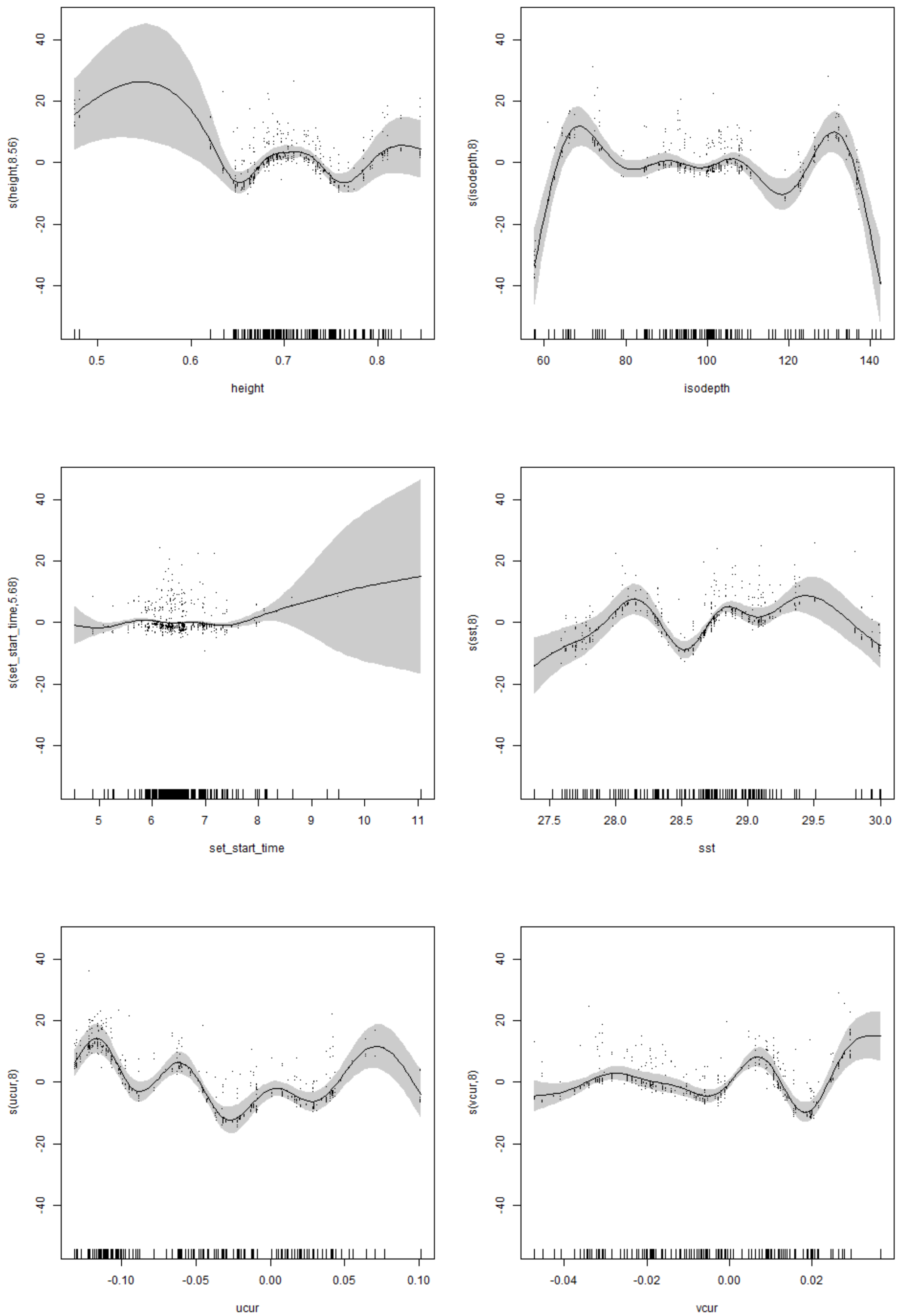


Figure 118: estimated smooths and approximate 95% confidence envelopes for smooth terms within the GAM.

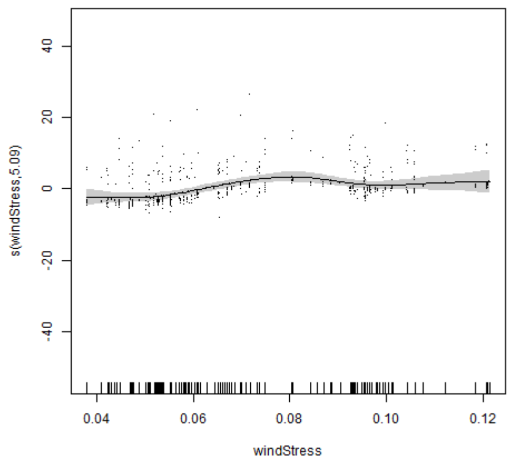


Figure 119: estimated smooths and approximate 95% confidence envelopes for smooth terms within the GAM.

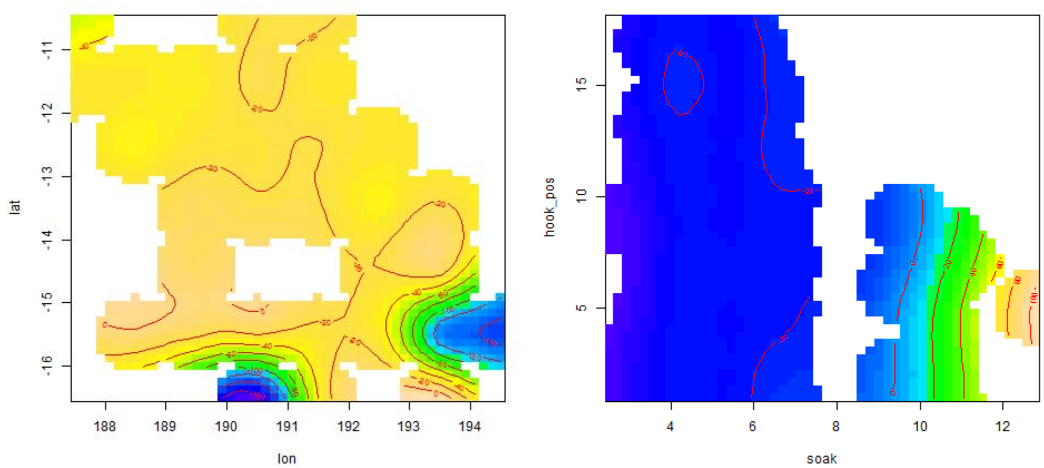


Figure 120: estimated smooth surfaces for the bivariate smooth components within the GAM.

6.12US FLEET ANALYSIS – THRESHER SHARKS (THR)

6.12.1 Catch rate of thresher sharks (THR)

6.12.1.1 Data Exploration

While predominantly absent in sets from observed trips of the US fleet in the M2_FV fishery, within-set bycatch levels of thresher sharks ranged mainly between 4 shark/set (in about 10 sets) and 12 sharks/set (in about 5 sets) (Figure 121). However, there were up to a maximum of 64 thresher sharks in one single set.

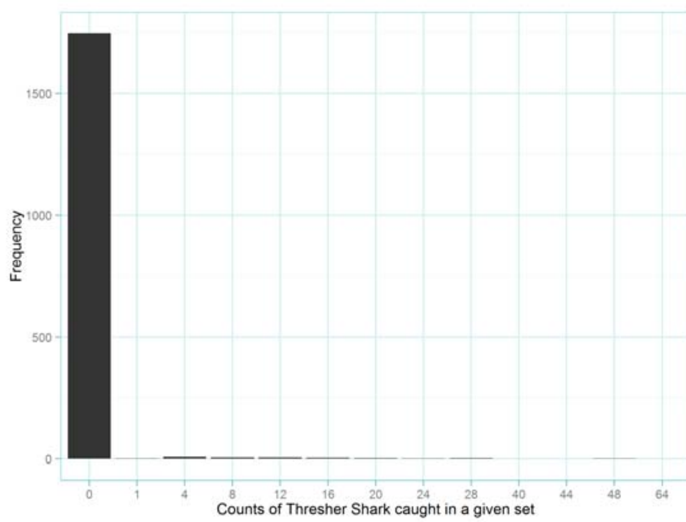


Figure 121: Distribution of counts per set of thresher sharks caught in observed trips of the US fleet operating in M2_FV longline fishery (file M2_FV_US_THR_CatchNumFreqs).

The predominance of zero catches and the relatively small number of specimens in observed sets hinders the investigation of bivariate patterns between catch rates of thresher sharks and each of the covariates considered in the present analysis (Figure 122).

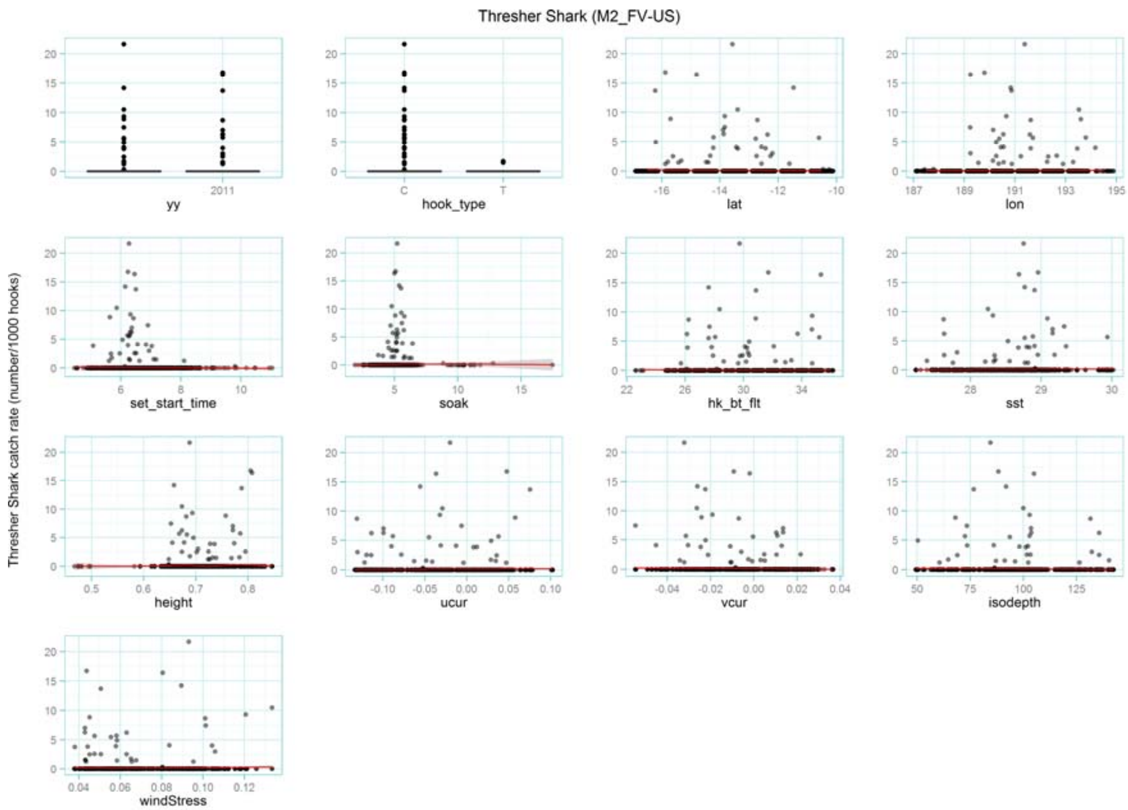


Figure 122: Distributions of set-level catch-rates against the explanatory variables considered for the analysis of thresher sharks caught by the US fleet in the M2_FV fishery (file "M2_FV-US-THR_CatchRatesVsCovs.png")

6.12.1.2 Modelling results

Modelling details are as per the reference case. The model returned an adjusted-R^2 of ~0 and a deviance explained of 34%.

Table 35: Model terms from the generalized additive model. DF = Degrees of Freedom. EDF=Effective Degrees of Freedom. Shrinkage has been applied as a means of model selection. Terms of little relevance have their EDF shrunken.

Model terms	EDF	DF	Chi or F-stat	*p*-value
yy		1	2.03	0.1545
hook_type		1	3.92	0.0478
s(lon,lat)	26.94	28	1.19	0.1856
s(soak,hk_bt_flt)	26.12	29	1.40	0.0344

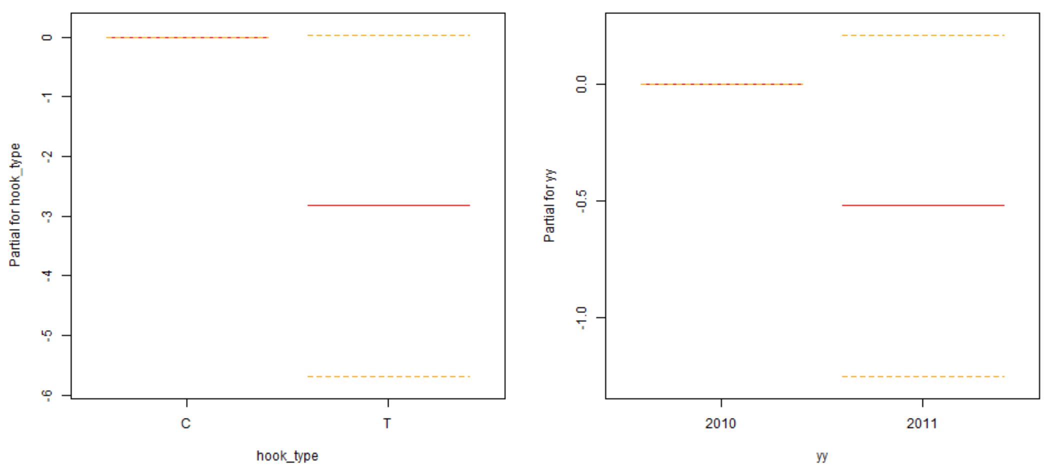


Figure 123: estimates and approximate 95% confidence intervals for parametric terms within the GAM.

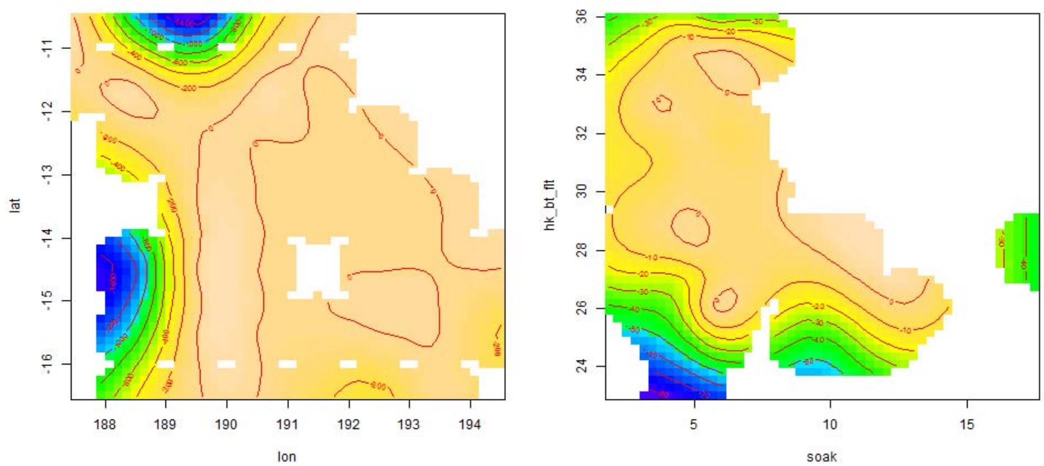


Figure 124: estimated smooth surfaces for the bivariate smooth components within the GAM.

6.12.2 Hook-level catch rate of thresher sharks

6.12.2.1 Data Exploration

There are no clear patterns between catch rates of thresher sharks and hook positions at which they were caught (Figure 125). Furthermore, there are no demarked trends in how the relationship between catch rates and hook positions changes with varying levels of the covariates considered in the analysis.

Figure 125: Distributions of per-hook catch-rate responses against the explanatory variables considered in the analysis, for thresher sharks caught by the US fleet operating in the M7_FV fishery. Dots were slightly jittered to minimise overplotting and help visualisation and their colour convey the basket hook position of the respective observation (file "M2_FV-US-THR_perHookRatesVsCovs.png").

6.12.2.2 Modelling results

deviance explained of 3.5%.

Table 36: Estimates of catch rates for each of the hook positions. Confidence intervals are approximate 95%. Percentage of total catch is calculated from overall estimated catch over hook positions presented. Confidence intervals for percentages are based on this naïve translation of cate rate.

Hook position	Catch rate (per 100 hooks)	SE	Lower CI	Upper CI	Catch (as % of total)	Lower CI (as % of total)	Upper CI (as % of total)
1	0.0034	1.8893	0.001	0.012	1.41	0.4	4.9
2	0.0049	1.7241	0.0017	0.0144	2.02	0.7	5.88
3	0.016	1.3896	0.0084	0.0306	6.55	3.44	12.49
4	0.007	1.5981	0.0028	0.0176	2.87	1.14	7.19
5	0.0179	1.3688	0.0097	0.0331	7.31	3.95	13.53
6	0.0204	1.3453	0.0114	0.0366	8.35	4.67	14.94
7	0.0098	1.5022	0.0044	0.0217	3.99	1.8	8.86
8	0.0139	1.4188	0.007	0.0276	5.68	2.86	11.28
9	0.0113	1.4649	0.0054	0.024	4.63	2.19	9.79
10	0.0138	1.42	0.007	0.0275	5.65	2.84	11.23
11	0.0103	1.4887	0.0047	0.0225	4.21	1.93	9.17
12	0.0068	1.61	0.0027	0.0173	2.77	1.09	7.06
13	0.027	1.3174	0.0157	0.0463	11.01	6.42	18.9
14	0.0163	1.4927	0.0074	0.0357	6.65	3.03	14.59
15	0.0043	2.3416	8.00E-04	0.0227	1.75	0.33	9.25
16	0.0135	1.9442	0.0037	0.0497	5.52	1.5	20.3
17	0.048	1.5395	0.0206	0.1119	19.62	8.42	45.7

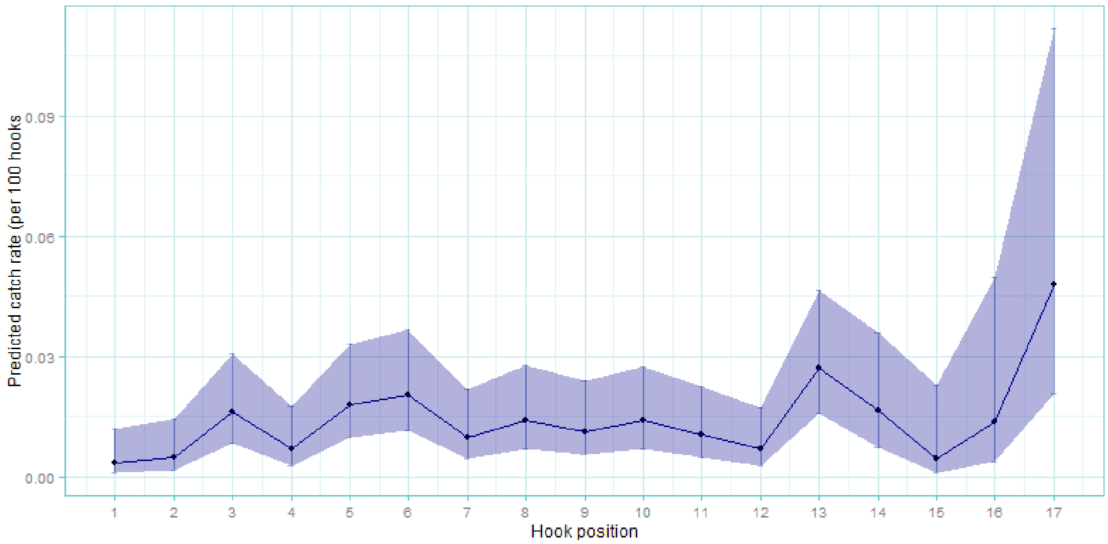


Figure 126: estimated mean catch rate (per 100 hooks) at hook-position with approximate 95% confidence envelope.

6.12.3 Condition of thresher sharks at time of retrieval

6.12.3.1 Data Exploration

There are no strong patterns between the condition of caught thresher sharks at the time of retrieval in observed trips of the US fleet operating in the M2_FV fishery and the covariates considered in the present analysis (Figure 127).

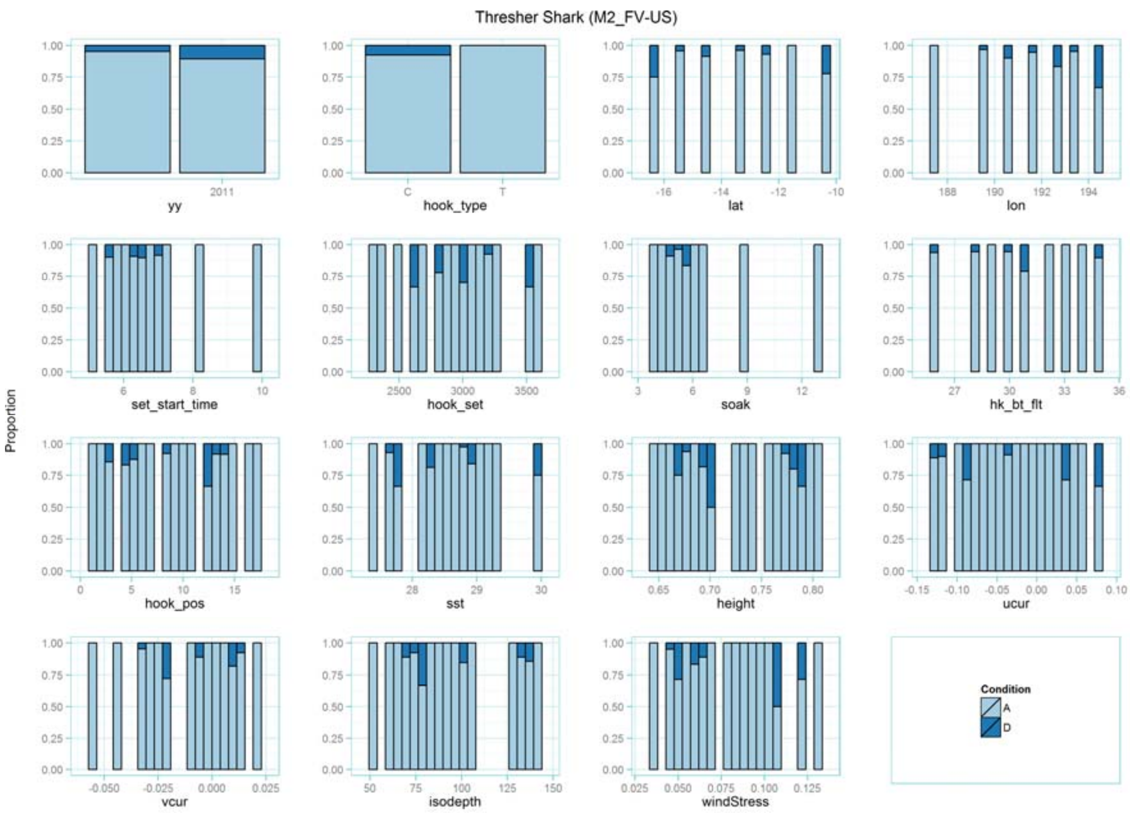


Figure 127: Distributions of condition responses (D= dead, A = alive), expressed in terms of proportion, against the explanatory variables considered for the analysis of thresher sharks caught by the US fleet in the M2_FV fishery (file "M2_FV-US-THR_CondtnVsCovs.png")

6.12.3.2 Modelling results

Modelling details are as per the reference case. The model returned an adjusted-R^2 of 0.89 and a deviance explained of 92.5%. Raw predictive power to the current dataset was also assessed via confusion matrices on the training data, providing Sensitivity (true positive rate) = 1, Specificity (true negative rate) = 0.99 and a misclassification rate of 0.013 when using the response mean as a decision boundary. These are against the training data and will be over-estimates.

Table 37: Model terms from the generalized additive model. DF = Degrees of Freedom. EDF=Effective Degrees of Freedom. Shrinkage has been applied as a means of model selection. Terms of little relevance have their EDF shrunken.

Model terms	EDF	DF	Chi or F-stat	*p*-value
yy		1	0.26	0.6089
hook_type		1	0.00	1.0000
s(set_start_time)	0.00	8	0.00	0.0488
s(sst)	0.00	8	0.00	0.1665
s(height)	0.00	9	0.00	0.0694
s(ucur)	0.00	9	0.00	0.0423
s(vcur)	5.12	8	0.08	0.1614
s(isodepth)	0.00	9	0.00	0.0298
s(windStress)	0.50	8	0.01	0.0366
s(soak)	0.00	9	0.00	0.0765
s(hook_pos)	0.00	9	0.00	0.4496
s(hk_bt_flt)	0.00	5	0.00	0.4874
s(soak,hook_pos)	11.11	26	0.40	0.1314

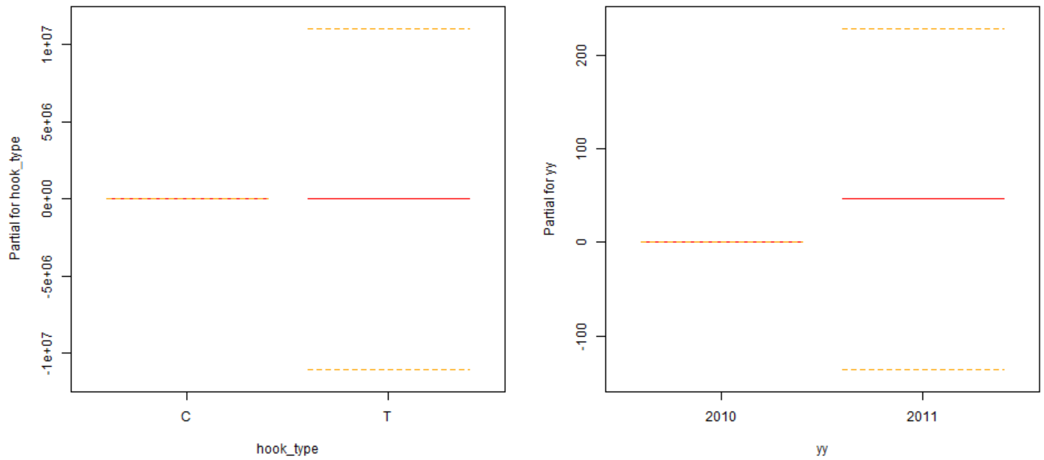


Figure 128: estimates and approximate 95% confidence intervals for parametric terms within the GAM.

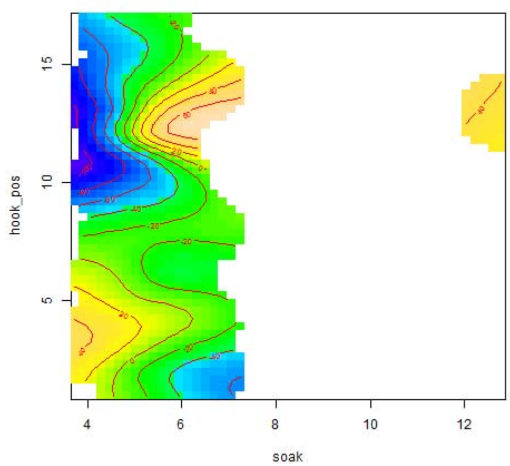


Figure 129: estimated smooth surfaces for the bivariate smooth components within the GAM.

6.13US FLEET ANALYSIS – BLUE SHARKS (BSH)

6.13.1 Catch rate of blue sharks

6.13.1.1 Data Exploration

While predominantly absent in sets from observed trips of the US fleet in the M2_FV fishery (Figure 130), within-set bycatch levels of blue sharks ranged between 4 shark/set (in about 80 sets) and 44 sharks/set (in about 10 sets). However, there were up to a maximum of 80 blue sharks in one single set.

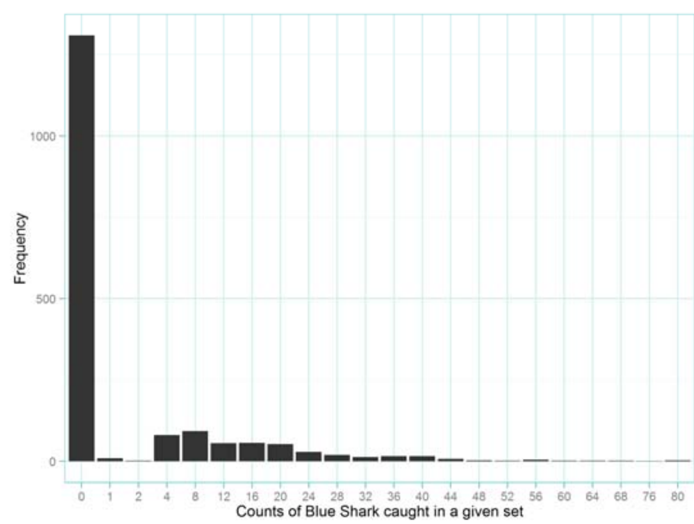


Figure 130: Distribution of counts per set of blue sharks caught in observed trips of the US fleet operating in M2_FV longline fishery (file M2_FV_US_BSH_CatchNumFreqs).

The predominance of zero catches in observed sets hinders the investigation of patterns between catch rates of blue sharks in the US fleet and each of the covariates considered in the analysis (Figure 131). However, there are sparse indications for higher catch rates when e.g.:

- Observed trips took place in 2011;
- Hook type “T” was in use;
- Sets were performed below -14 degrees of latitude.
- The average sea surface temperature remained around 28.5^{o}C;
- The average wind stress levels were below 0.06 Newton/m^2

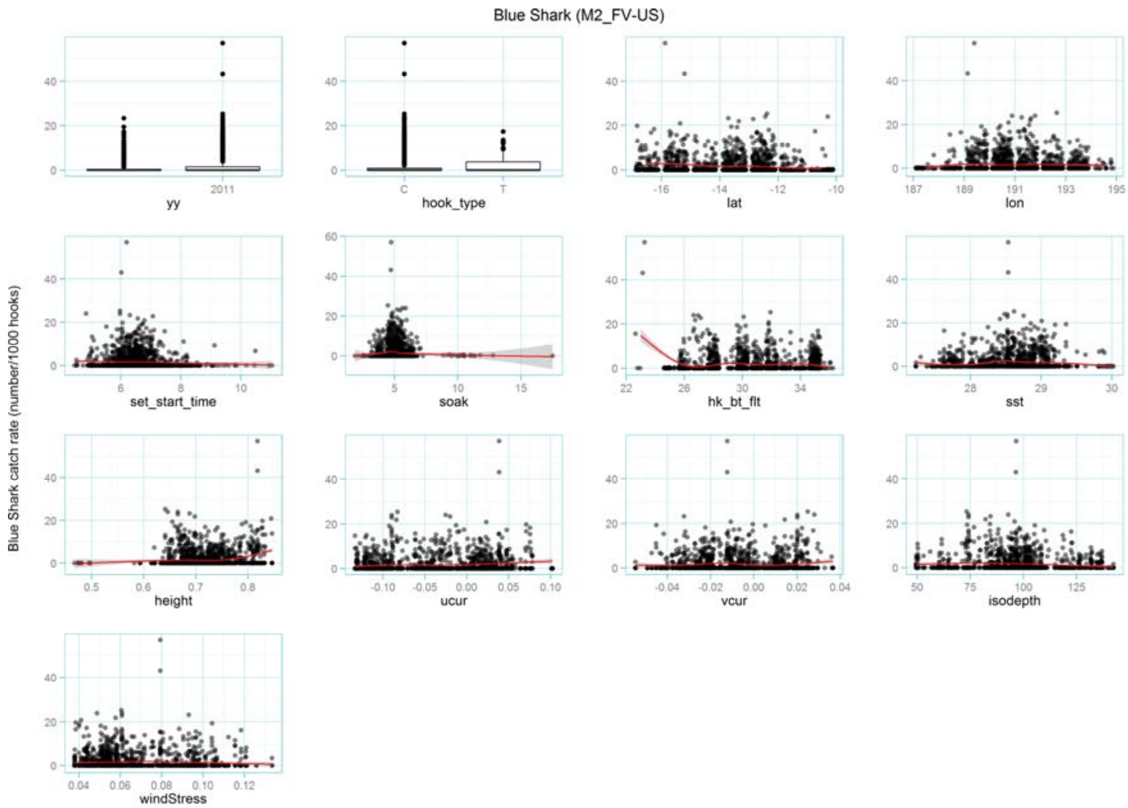


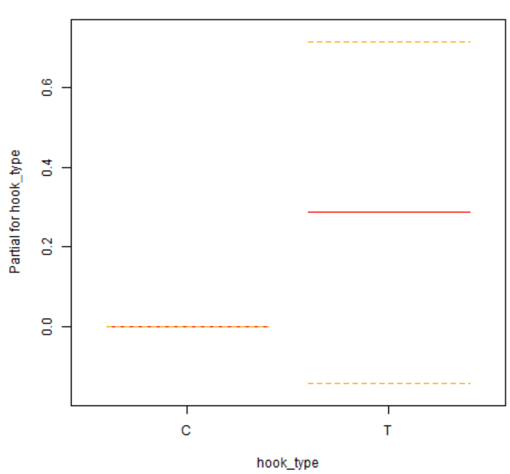
Figure 131: Distributions of set-level catch-rates against the explanatory variables considered for the analysis of blue sharks caught by the US fleet in the M2_FV fishery (file "M2_FV_FJ_BSH_CatchRatesVsCovs.png")

6.13.1.2 Modelling results

Modelling details are as per the reference case. The model returned an adjusted-R^2 of 0.16 and a deviance explained of 19.7%.

Table 38: Model terms from the generalized additive model. DF = Degrees of Freedom. EDF=Effective Degrees of Freedom. Shrinkage has been applied as a means of model selection. Terms of little relevance have their EDF shrunken.

Model terms	EDF	DF	Chi or F-stat	*p*-value
yy		1	0.30	0.5862
hook_type		1	1.80	0.1799
s(set_start_time)	5.17	9	1.32	0.0260
s(sst)	6.67	9	2.72	0.0002
s(height)	7.54	9	2.77	0.0004
s(ucur)	0.00	9	0.00	0.8011
s(vcur)	8.02	9	1.74	0.0293
s(isodepth)	7.42	9	3.24	0.0000
s(windStress)	0.77	9	0.29	0.0567
s(lon,lat)	16.24	29	1.40	0.0001
s(soak,hk_bt_flt)	8.51	29	0.53	0.0320



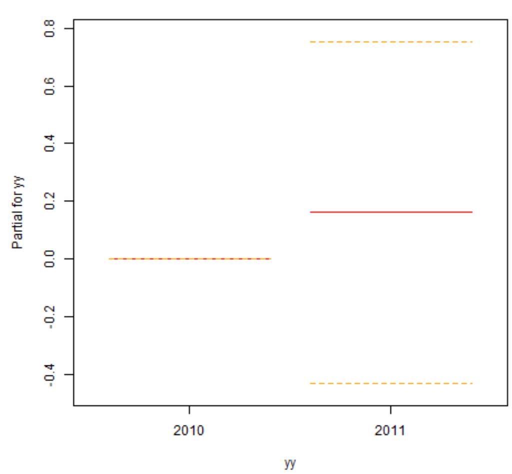


Figure 132: estimates and approximate 95% confidence intervals for parametric terms within the GAM.

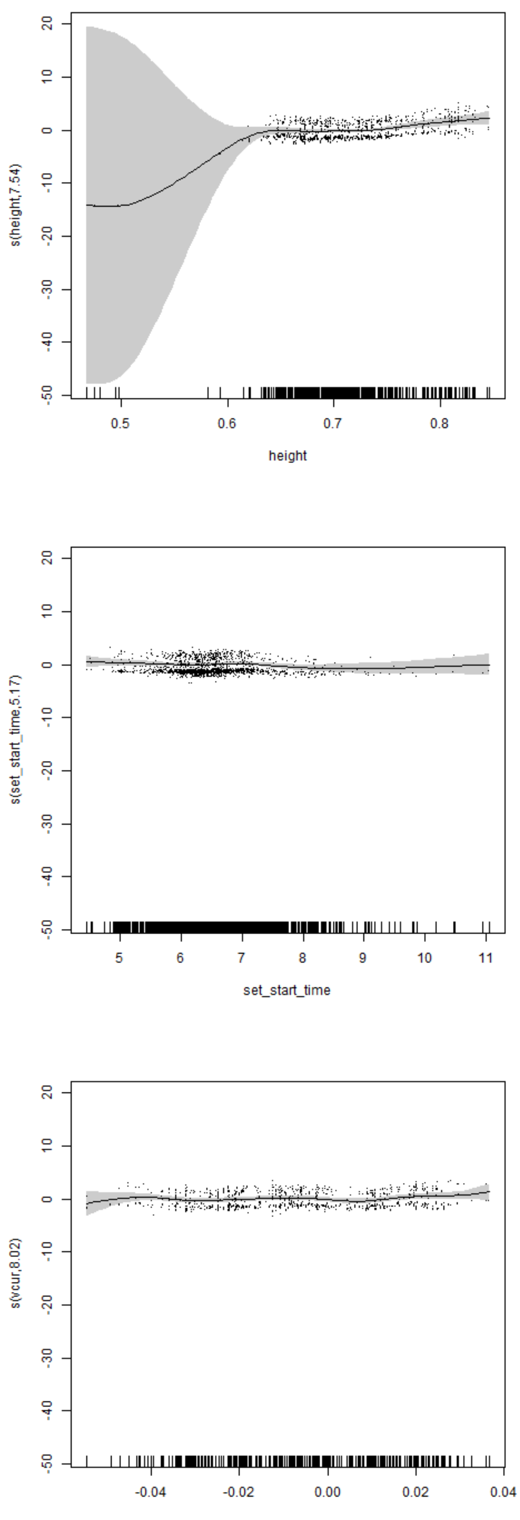


Figure 133: estimated smooths and approximate 95% confidence envelopes for smooth terms within the GAM.

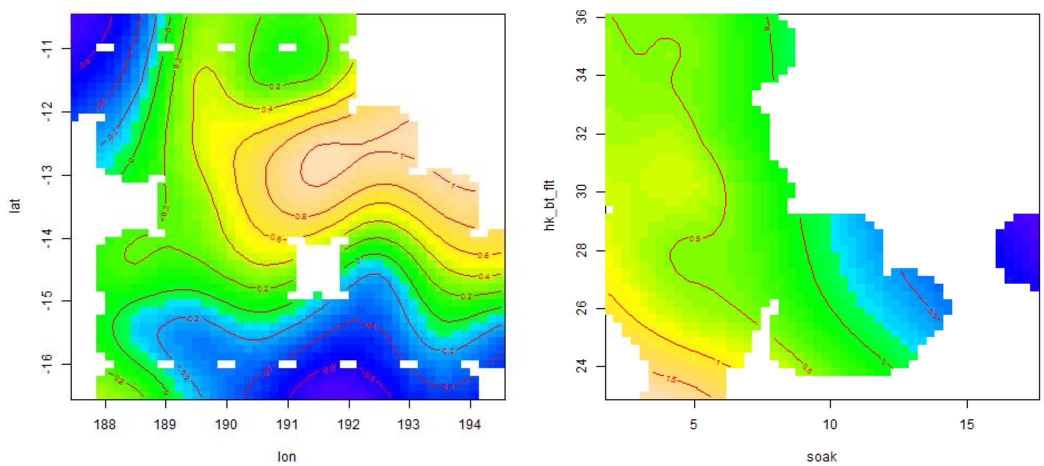


Figure 134: estimated smooth surfaces for the bivariate smooth components within the GAM.

6.13.2 Hook-level catch rate of blue sharks

6.13.2.1 Data Exploration

Figure 135 shows that catch rates of blue sharks in the US fleet tend to be larger at higher hook position numbers (i.e. at deeper hooks). However, there are no demarked trends in how the relationship between catch rates and hook positions changes with varying levels of the covariates considered in the present analysis.

Figure 135: Distributions of per-hook catch-rate responses against the explanatory variables considered in the analysis, for thresher sharks caught by the US fleet operating in the M7_FV fishery. Dots were slightly jittered to minimise overplotting and help visualisation and their colour convey the basket hook position of the respective observation (file "M2_FV-US-BSH_perHookRatesVsCovs.png").

6.13.2.2 Modelling results

Modelling details are as per the reference case. The model returned an adjusted-R^2 of 0.003 and a deviance explained of 1.4%.

Table 39: Estimates of catch rates for each of the hook positions. Confidence intervals are approximate 95%. Percentage of total catch is calculated from overall estimated catch over hook positions presented. Confidence intervals for percentages are based on this naïve translation of cate rate.

Hook position	Catch rate (per 100 hooks)	SE	Lower CI	Upper CI	Catch (as % of total)	Lower CI (as % of total)	Upper CI (as % of total)
1	0.0735	1.169	0.0541	0.0998	2.3	1.69	3.12
2	0.139	1.1251	0.1103	0.1751	4.35	3.45	5.48
3	0.1695	1.114	0.1371	0.2094	5.3	4.29	6.55
4	0.1552	1.1188	0.1246	0.1935	4.86	3.9	6.05
5	0.1677	1.1146	0.1356	0.2075	5.25	4.24	6.49
6	0.1869	1.109	0.1526	0.2289	5.85	4.78	7.16
7	0.2086	1.1036	0.172	0.253	6.53	5.38	7.92
8	0.1805	1.1107	0.147	0.2218	5.65	4.6	6.94
9	0.1791	1.1112	0.1456	0.2201	5.6	4.56	6.89
10	0.1775	1.1116	0.1443	0.2184	5.56	4.52	6.84
11	0.2027	1.1049	0.1667	0.2465	6.34	5.22	7.71
12	0.1557	1.1188	0.1249	0.194	4.87	3.91	6.07
13	0.2492	1.1	0.2067	0.3004	7.8	6.47	9.4
14	0.2074	1.1287	0.1636	0.263	6.49	5.12	8.23
15	0.3386	1.1238	0.2693	0.4256	10.6	8.43	13.32
16	0.1146	1.2724	0.0715	0.1837	3.59	2.24	5.75
17	0.2895	1.1958	0.2039	0.4111	9.06	6.38	12.87

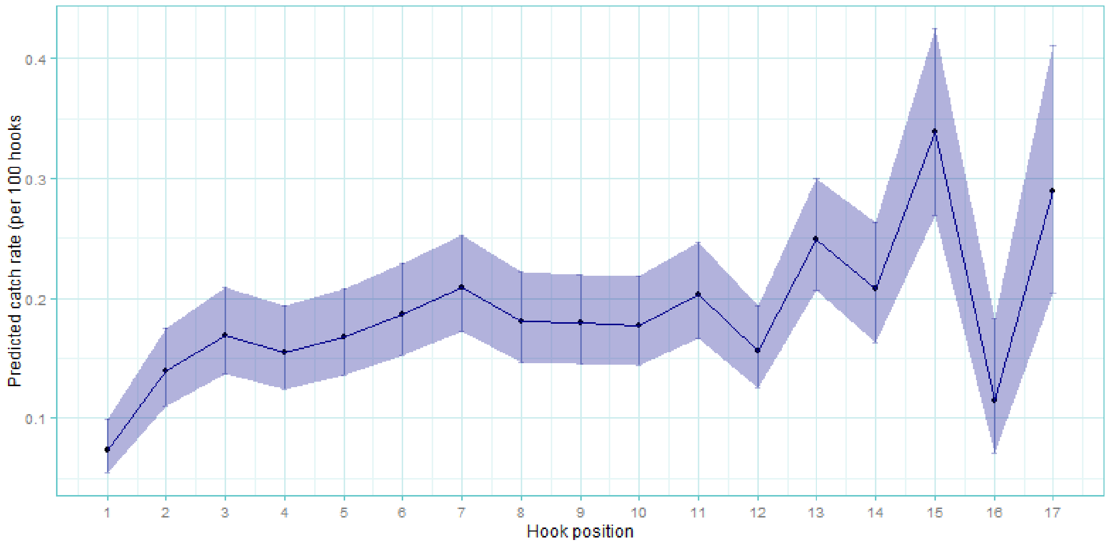


Figure 136: estimated mean catch rate (per 100 hooks) at hook-position with approximate 95% confidence envelope.

6.13.3 Condition of blue sharks at time of retrieval

6.13.3.1 Data Exploration

There are no strong patterns between the condition of caught blue sharks at the time of retrieval and most of the covariates considered in the present analysis (Figure 137). There are however slight indications that the proportion of dead sharks might have increased when observed sets were performed under certain conditions, like e.g. (i) as the number of hooks in the set (i.e. the length of the whole set) surpasses the 2000 hooks; and (ii) if sets were performed at a local average sea surface temperatures outside the region of 28.5^{o}C.

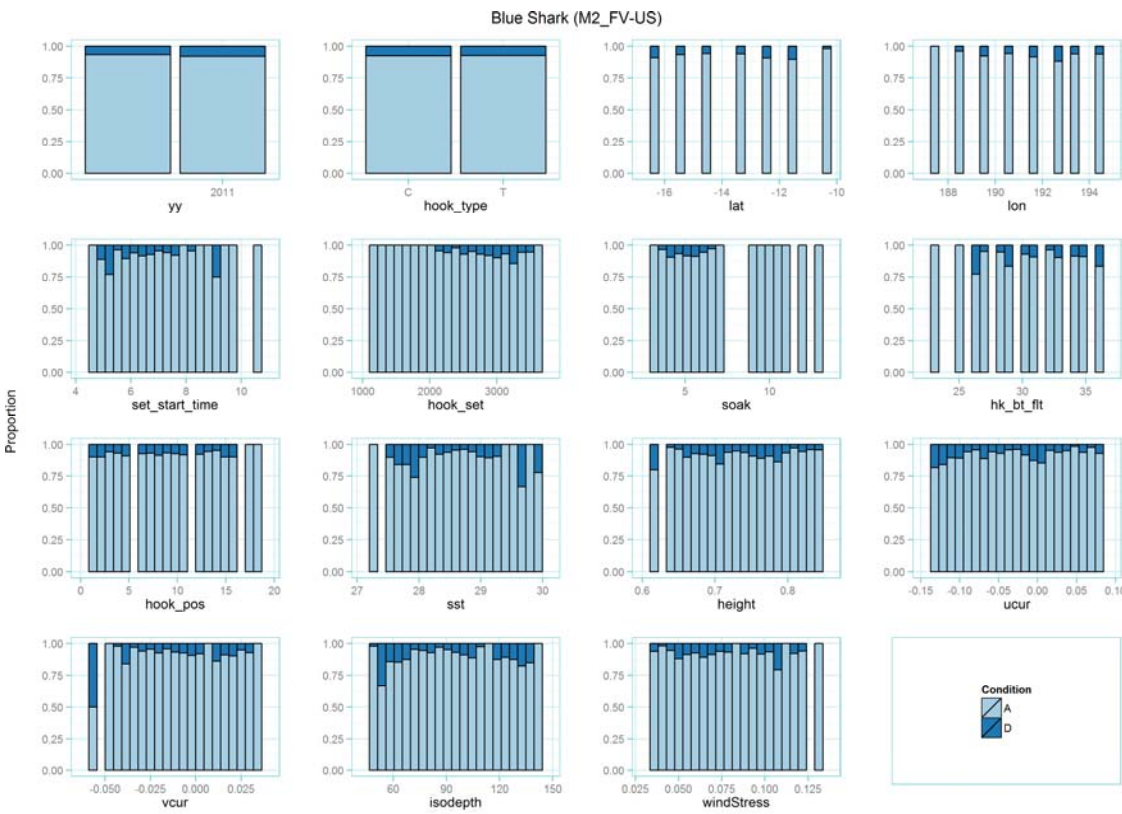


Figure 137: Distributions of condition responses (D= dead, A = alive), expressed in terms of proportion, against the explanatory variables considered for the analysis of blue sharks caught by the US fleet in the M2_FV fishery (file “M2_FV-US-BSH_CondtnVsCovs.png”)

6.13.3.2 Modelling results

Modelling details are as per the reference case. The model returned an adjusted-R^2 of 0.184 and a deviance explained of 26.2%. Raw predictive power to the current dataset was also assessed via confusion matrices on the training data, providing Sensitivity (true positive rate) = 0.78, Specificity (true negative rate) = 0.74 and a misclassification rate of 0.25 when using the response mean as a decision boundary. These are against the training data and will be over-estimates.

Table 40: Model terms from the generalized additive model. DF = Degrees of Freedom. EDF=Effective Degrees of Freedom. Shrinkage has been applied as a means of model selection. Terms of little relevance have their EDF shrunken.

Model terms	EDF	DF	Chi or F-stat	*p*-value
yy		1	7.73	0.0054
hook_type		1	0.64	0.4233
s(set_start_time)	7.72	9	46.82	0.0000
s(sst)	8.56	9	46.38	0.0000
s(height)	7.98	9	58.80	0.0000
s(ucur)	7.98	9	55.30	0.0000
s(vcur)	6.02	9	33.63	0.0000
s(isodepth)	7.17	9	51.69	0.0000
s(windStress)	5.80	9	15.64	0.0065
s(soak,hook_pos)	26.96	29	127.31	0.0000
s(lat,lon)	26.14	29	158.53	0.0000
s(soak,hk_bt_flt)	27.00	28	167.56	0.0000

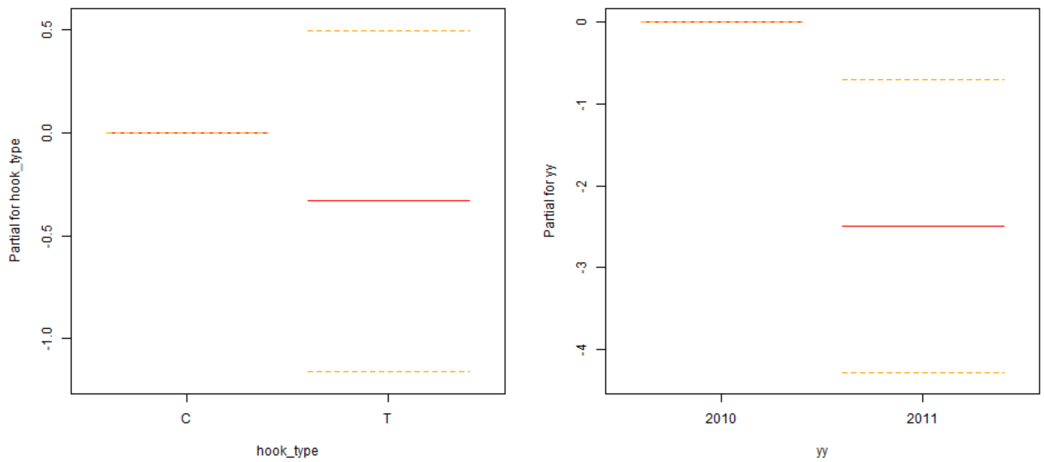


Figure 138: estimates and approximate 95% confidence intervals for parametric terms within the GAM

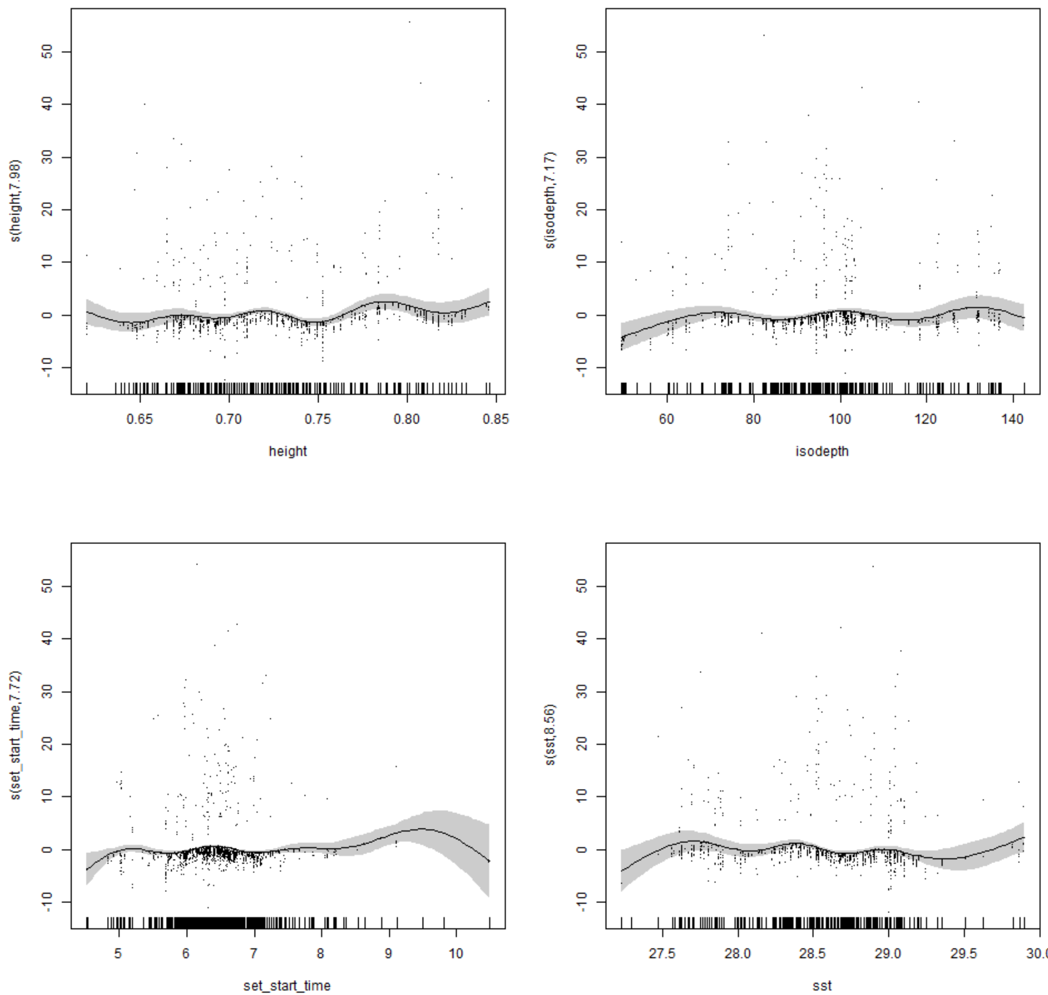


Figure 139: estimated smooths and approximate 95% confidence envelopes for smooth terms within the GAM.

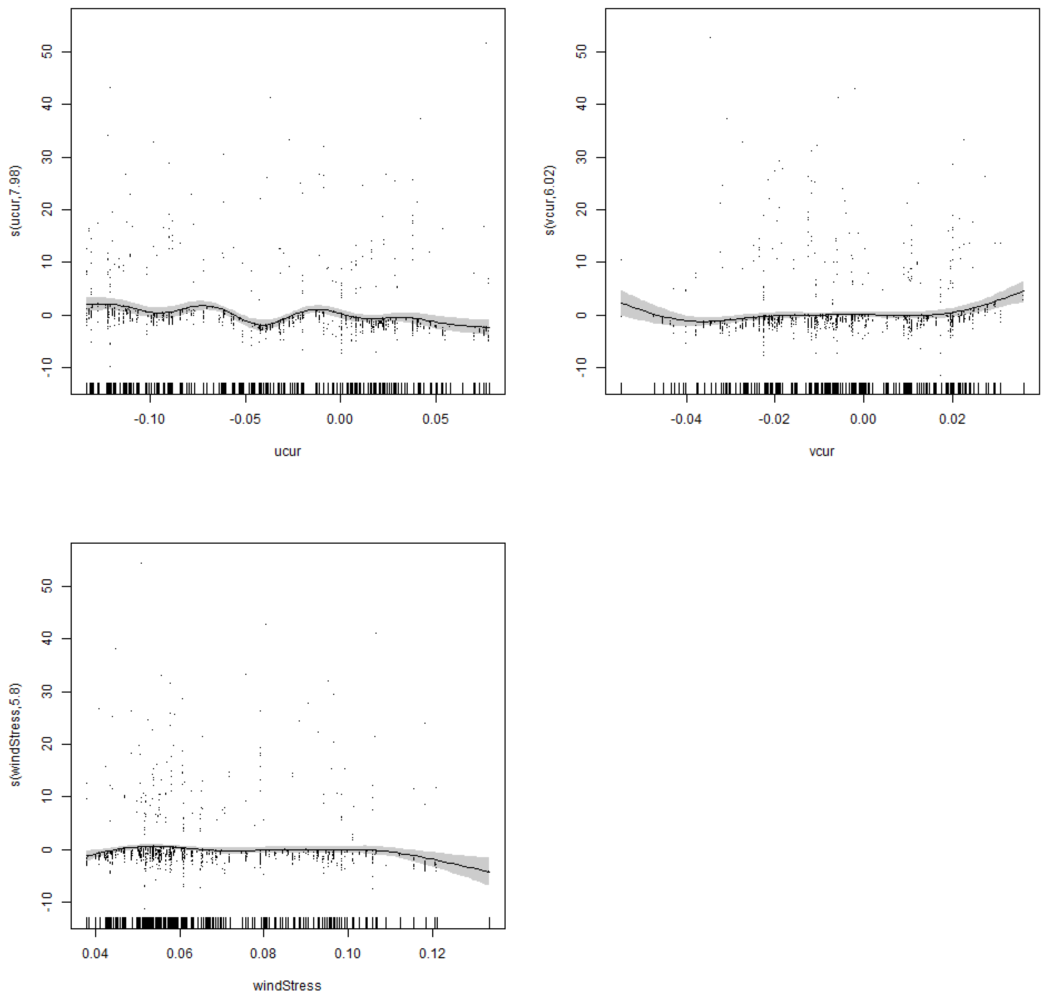


Figure 140: estimated smooths and approximate 95% confidence envelopes for smooth terms within the GAM.

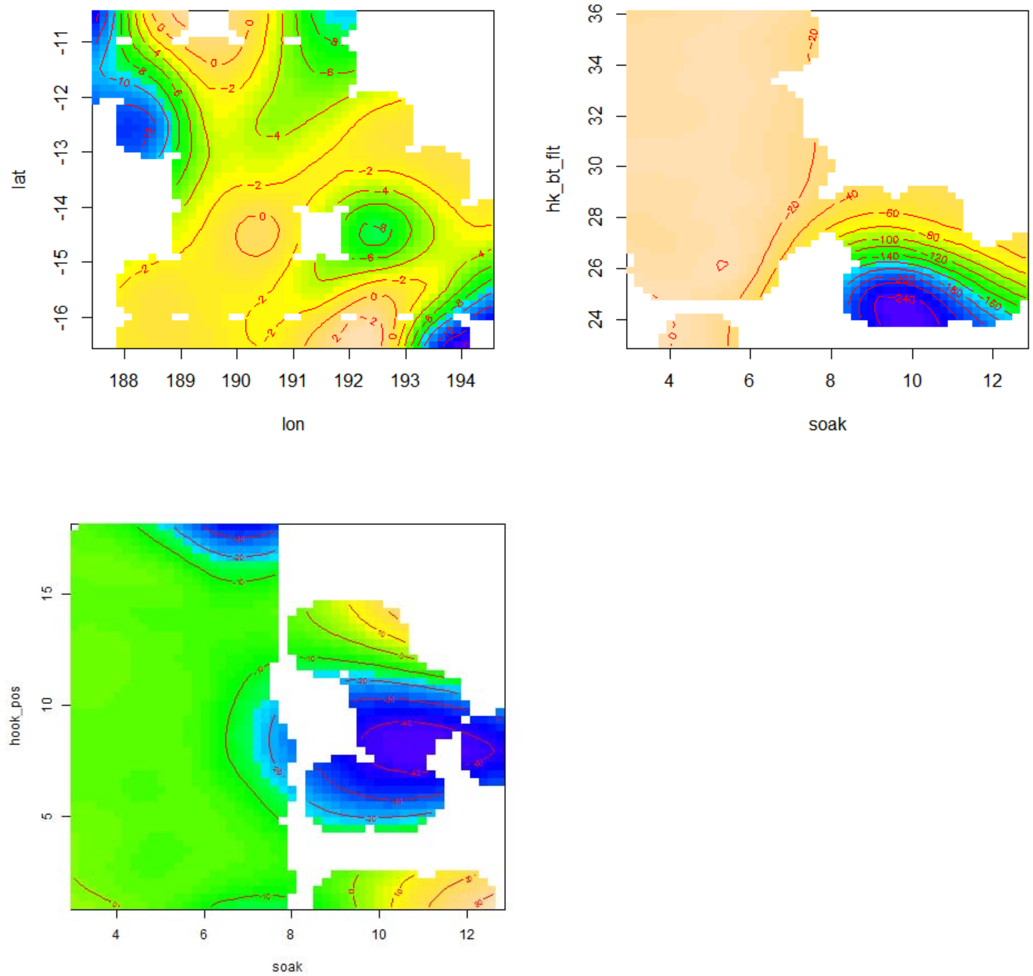


Figure 141: estimated smooth surfaces for the bivariate smooth components within the GAM.

7 HAWAIIAN DEEP WATER FISHERY (M7_HD)

7.1 FISHERY DATA DESCRIPTION

Observer data from the Hawaiian deep water fishery was originated from trips carried out by US fleets while mainly operating in the EEZ of Hawaii (HW) and within International Waters (IW) (Figure 142). The data time-series for this fishery extends from 1994 to 2011, with the majority of available observed trips occurring in the period 2001-2011. There are no considerable changes in the number of observed trips over months, and the number of sets per trip has increased slowly through the years, with vessels performing mainly between 10 and 15 sets per trip (Figure 142).

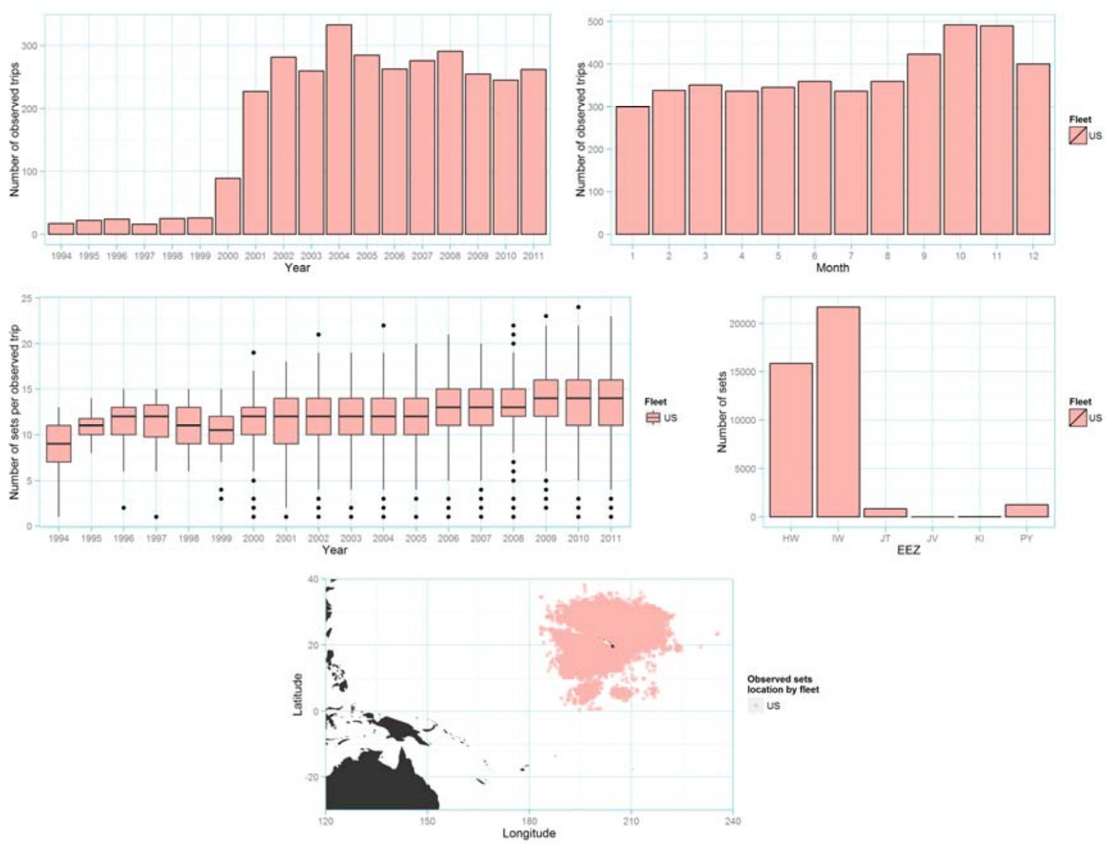


Figure 142: Temporal and spatial distribution of fishing effort in observed trips by the US fleet operating in the Hawaiian deepwater fishery (file “M7_HD_tripsAndSetsDesc.png”).

The blue shark was by far the species with higher bycatch numbers in observed US trips across the entire time-series (Figure 143), with thresher sharks being the second most abundant species in terms of shark bycatch. For example, this fishery caught over 21'000 blue sharks in trips observed during 2007, while catching around 2000 thresher sharks and about 200 oceanic whitetip sharks in trips from the same year. On the other hand, catches of Hammerhead sharks are virtually absent in the M7_HD fishery (Figure 143).

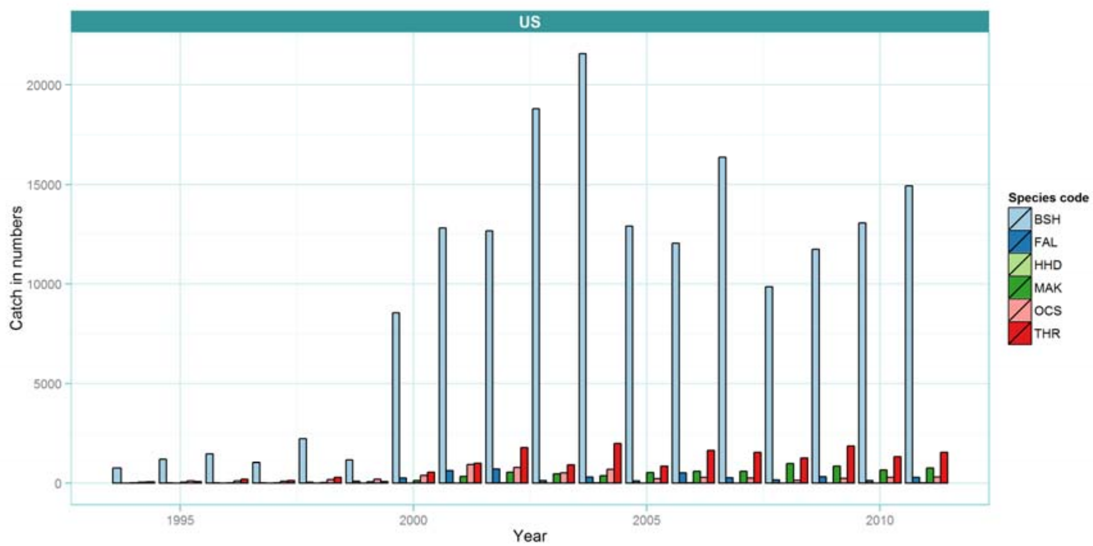


Figure 143: Bycatch of shark species over time in observed trips by the US fleet operating in the Hawaiian deepwater fishery (file "M7_HD_catchNumSpYrFlg.png").

The observer data collected from the M7_HD fishery offers an 18 years-long time-series (comprising nearly 40'000 unique sets) with vast amounts of contrasting data on the type of branchline leader employed and the different types of hooks ("C", "J" and "T") used in this fishery (Figure 144). On the other hand, the complete absence of use of shark lines and shark bait by the US vessels operating in the fishery meant that these two key fishing factors could not be considered in the analysis.

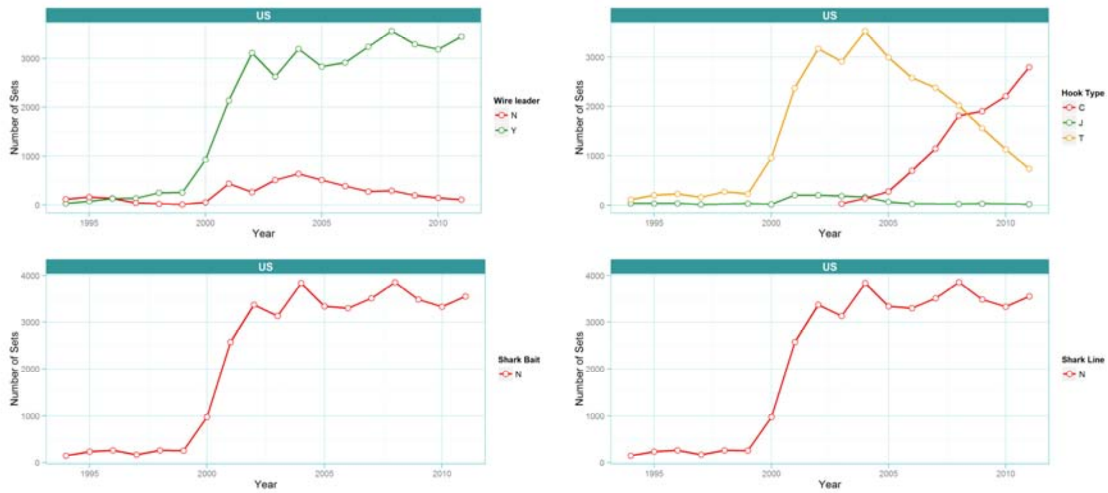


Figure 144: Contrast of key gear types over time in observed trips by the US fleet operating in the Hawaiian deepwater fishery (File "M7_HD_SampSizeAndKeyFactors.png").

7.2 EXPLANATORY VARIABLES EDA

Figure 145 presents the covariate data for the observations used for the analysis of shark bycatch in the Hawaiian deep fishery. There were a few observations with large values of wind stress (over 0.3 Newton/m^2) that were considered too extreme in relation to the normal range of wind stress levels in the dataset – these observations were excluded from the analysis. Other observations with slightly wide covariates values (e.g. in current speeds, ucur and vcur) were considered to be within the expected range of values and therefore kept in the analysis.

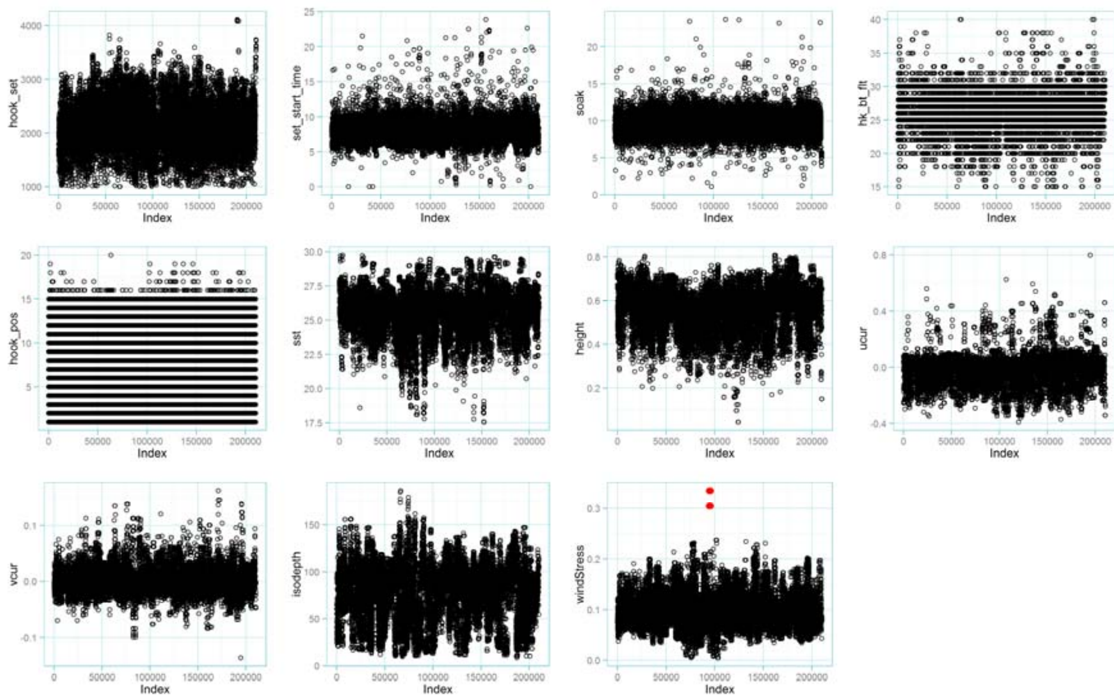


Figure 145: Cleveland dotplots of relevant covariates considered for the analysis of shark bycatch in the Hawaiian deepwater fishery. Plots constructed from data sorted by set starting date (File "M7_HD_covsClevPlots.png")

Figure 146 shows the presence of very weak to non-existent correlations between the explanatory variables from observations used for the analysis of the M7_HD fishery.

Furthermore, with respect to the geographical region underpinning the M7_HD fishery, several of the oceanographic variables (e.g. sea surface temperature, sea surface height and wind stress) appear to have non-linear relationships with month, latitude and longitude (**Figure 58**). Changes in the average depth of the 20^{o}C isotherm also appears to be followed by changes in average sea surface height and in horizontal (ucur) and vertical (vcur) current speeds. In addition, there appears to be a (non-linear) relationship between wind stress and current speed.

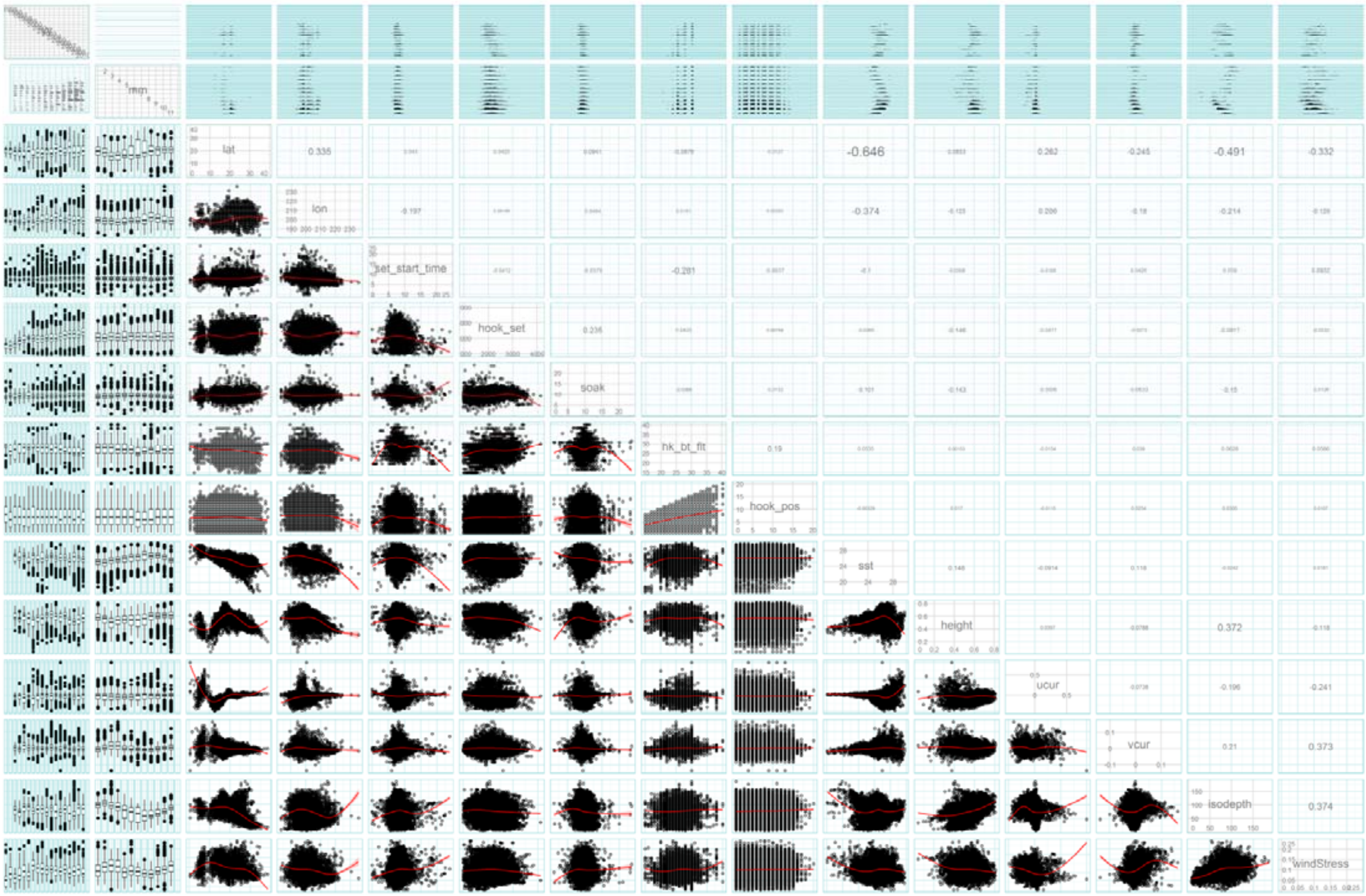


Figure 146: Matrixplot of relevant covariates considered for the analysis of shark bycatch in the Hawaiian deepwater fishery. Plots presented in panels depend on the data types under comparison. For example, pairs of continuous variables are displayed as scatterplots in the lower panels (with an added LOESS smoother to help visualisation) and Pearson correlation coefficients in the upper panels (with font size proportional to

correlation coefficient). Continuous Vs Discrete covariates are represented by boxplots and barplots, while pairs of discrete covariates are displayed as fluctuation plots (lower panels) and barplots (upper panels) (file "M7_HD_CovsPairsPlot.png")

8 HAWAIIAN DEEP WATER FISHERY – MODELLING

Presented here are results for three models covering set-level catch-rates, catch-rates with respect to hook position and the condition of sharks (alive/dead) at retrieval. Four shark species are considered: Oceanic whitetips (OCS), silky (FAL), thresher (THR) and blue (BSH).

Details regarding modelling methods are presented previously, but all models fall within the Generalized Additive Modelling framework. The set of parametric and smoothing terms selected between are given in the following table – the specific model fitted for each species and model is presented in the relevant results section.

In contrast to the previous fisheries, M7 HD has no shark-lines or shark-bait. The pool of terms considered for models is correspondingly smaller.

Table 41: Model terms considered. "s" indicates a smoothing term. ":" indicates an interaction between associated terms.

Parametric terms	Smooth terms
yy	s(set_start_time)
mm	s(sst)
hook_type	s(height)
wire_trace	s(ucur)
hook_type:wire_trace	s(vcur)
hook_type:wire_trace	s(isodepth)
	s(windStress)
	s(lat,lon)
	s(soak,hk_bt_flt)

Reference model – In the interests of brevity, one species' models are presented in more detail than the others. This will be referred to as the *reference case* and will be the Oceanic whitetip shark (OCS). In particular model diagnostics and non-significant results will be presented for this model, but not others. The same modelling approach has been applied for similar models, but only significant results (including assumption violations) will be reported upon.

DMP STATS The Coach House, Mount Melville House, St. Andrews, Scotland KY 168 NT, UK +44(0) 1334 477 544 mail@dmpstats.com

8.1 OCEANIC WHITE SHARKS (OCS)

8.1.1 Catch rate of oceanic white sharks (OCS)

8.1.1.1 Data Exploration

While predominantly absent in sets from observed trips in the M7_HD fishery, within-set bycatch levels of oceanic whitetip sharks ranged mainly between 1 shark/set (in about 3000 sets) and 3 sharks/set (in about 200 sets) (Figure 147). However there were up to a maximum of 92 oceanic whitetips in one single set.

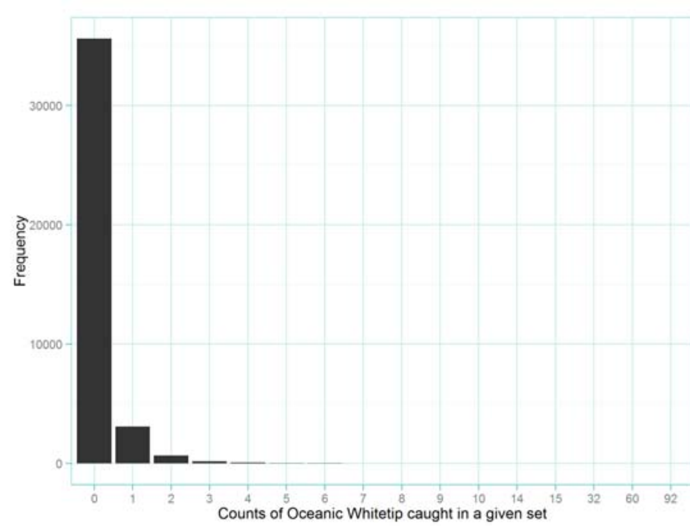


Figure 147: Distribution of counts per set of oceanic whitetip shark caught in observed trips of the M7_HD longline fishery (M7_HD-OCS_CatchNumFreqs.png)

The predominance of zero catches in observed sets hinders the clear visualisation of patterns between catch rates of oceanic whitetip sharks and the covariates considered in the present analysis (Figure 148). However, there are some sparse indications that, on average, higher catch rates were likely to occur when, e.g.:

- Hook types “C” or “T” were employed, or wire leader were in use;
- Sets were performed between 1998 and 2002;
- Longlines were laid within 0-10 degrees latitude, or within 190-200 degrees longitude.

- Soak time varied between 8-12 hours;
- Average sea surface temperature was over 25^{o}C, or the average depth of the 20^{o}C isotherm was roughly between 50-100 meters.

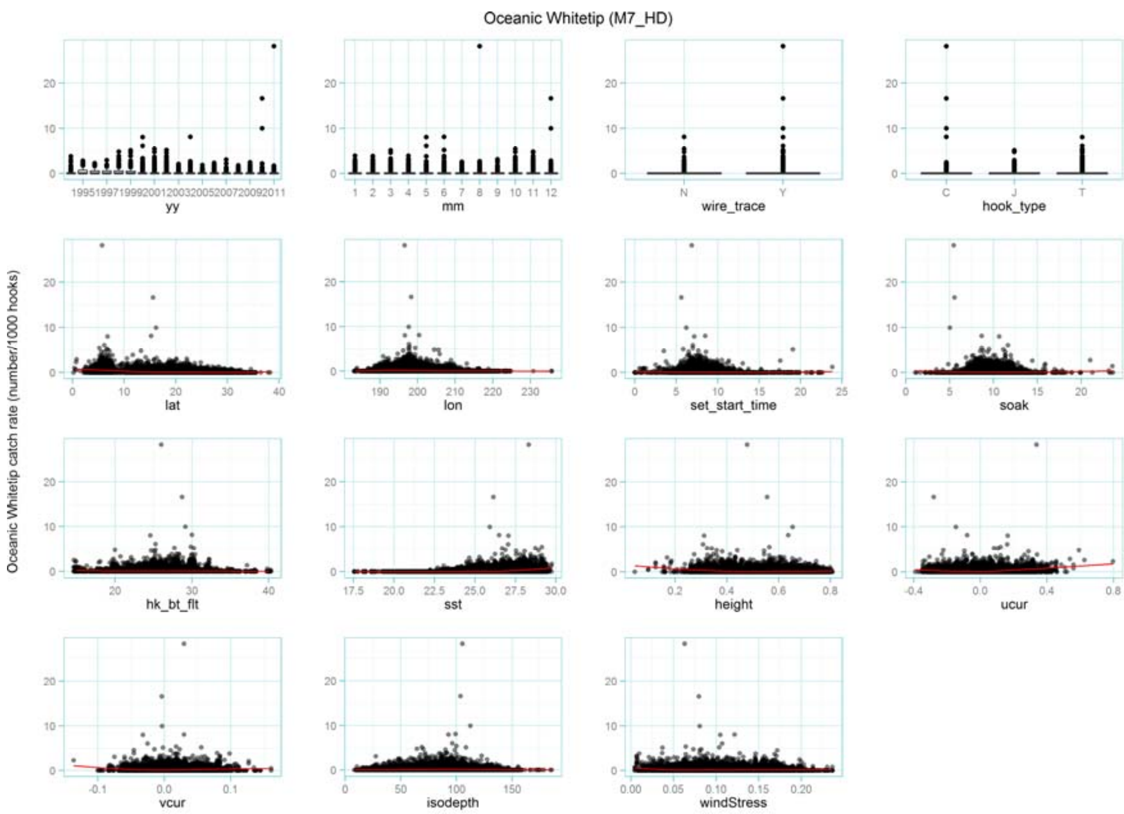


Figure 148: Distributions of set-level catch-rate responses against the explanatory variables considered for the analysis of oceanic whitetip sharks caught in the M7_HD fishery (file “M7_HD-OCS_CatchRatesVsCovs.png”).

8.1.1.2 Modelling results

Catch rate (per 1000 hooks) was modelled as a function of set-level characteristics and associated oceanographic variables. The initial set of potential covariates was thinned on the basis of gross collinearity, as determined by Generalized Variance Inflation Factors (GVIFs). A range of potential interactions and non-linear terms were specified *a priori* as per Table 3. These were further thinned during the modelling process in some cases where estimation was not possible e.g. interactions for factor variables where certain combinations of factors were not observed in the dataset.

The model fitted was a GAM with log-link and Tweedie-distributed errors. An initial model run was conducted for the purposes of estimating the governing Tweedie parameter – subsequently refitted with the estimated Tweedie parameter specified. Model "selection" was conducted within the GAM using shrinkage, so model terms of little relevance have their effective degrees of freedom shrunk towards zero. In terms of parametric components, backwards selection (p-value<0.05) was used, but not on the key factors hook-type & wire-trace. This was generally retained in the model for interpretative interest. The results of this process are summarised in Table 4.

The Tweedie parameter was estimated to be 1.034, with the model returning an adjusted-R^2 of 0.164 and a deviance explained of 32.7%.

Table 42: Model terms from the generalized additive model. DF = Degrees of Freedom. EDF=Effective Degrees of Freedom. Shrinkage has been applied as a means of model selection. Terms of little relevance have their EDF shrunken.

Model terms	EDF	DF	Chi or F-stat	*p*-value
yy		15	60.49	0.0000
mm		11	14.70	0.0000
hook_type		2	6.44	0.0016
wire_trace		1	3.76	0.0525
s(lon,lat)	24.92	29	104.52	0.0000
s(soak,hk_bt_flt)	27.64	29	8.58	0.0000

The model assumptions were assessed. Non-linearities are assumed to be captured via the GAM which optimises complexity. The adequacy of the assumed Tweedie distribution was assessed via Quantile Residuals (Dunn & Smythe, 1996), as indicated in Figure 8 – these ought to be approximately Normally distributed, producing an approximately straight-line. Potential correlation in the errors was similarly assessed on quantile residuals, both by examination of acf plots and a runs test. The data were ordered by set date within trip IDs, so the correlations are sought via sets that are temporally, and likely spatially, close.

In this case the Tweedie distribution is acceptable and there is no significant residual correlation once this zero rich distribution has been accounted for. The runs-test returns a *p*-value of 0.143.

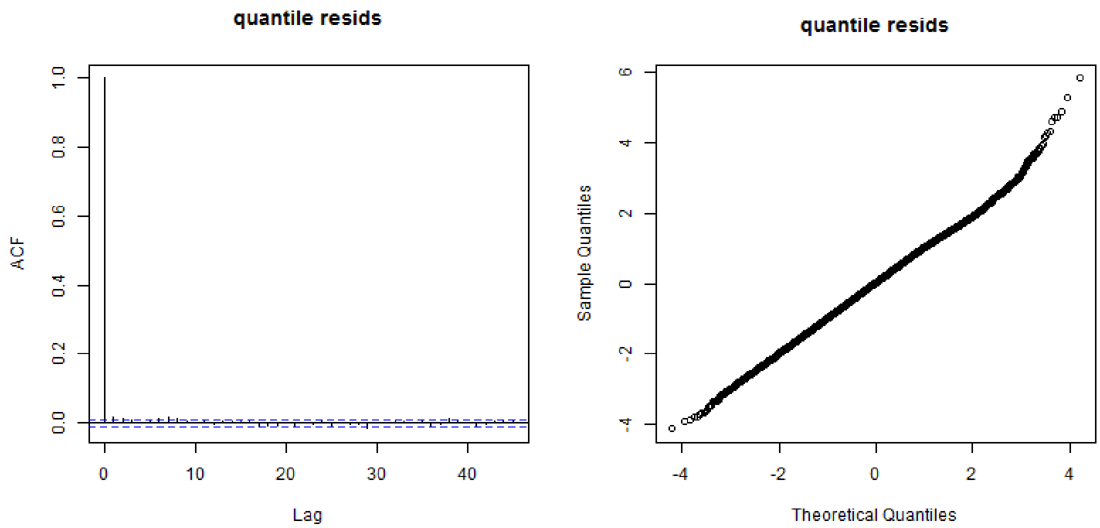


Figure 149: diagnostic plots for the quantile residuals from the GAM. The error distribution is assumed to be a Tweedie distribution, with key parameter estimated.

The following are link-scale estimates of the components of the GAM. All parametric terms in the model are represented here regardless of practical/statistical significance. For smoothers, those with EDFs shrunken to zero are excluded as the plots are flat by definition. Generally only top-line results will be presented outside of the reference case. Where applicable, the parametric terms are presented first, followed by smooth terms and then interactions.

The following observations can be made:

- There is a significantly lower catch rate under hook type T compared to hook type C.
- Catches were on average lower in 2011 than 2010, but high uncertainty about the estimates.
- There is a significant spatial component to catch-rates which is not explained by the other covariates as evidenced by the significant latitude-longitude interaction surface. The pattern observed is complex.
- There is some catch relationship with both soak-time and the hooks between floats, with peak catch rates being with intermediate basket-sizes and soak times.

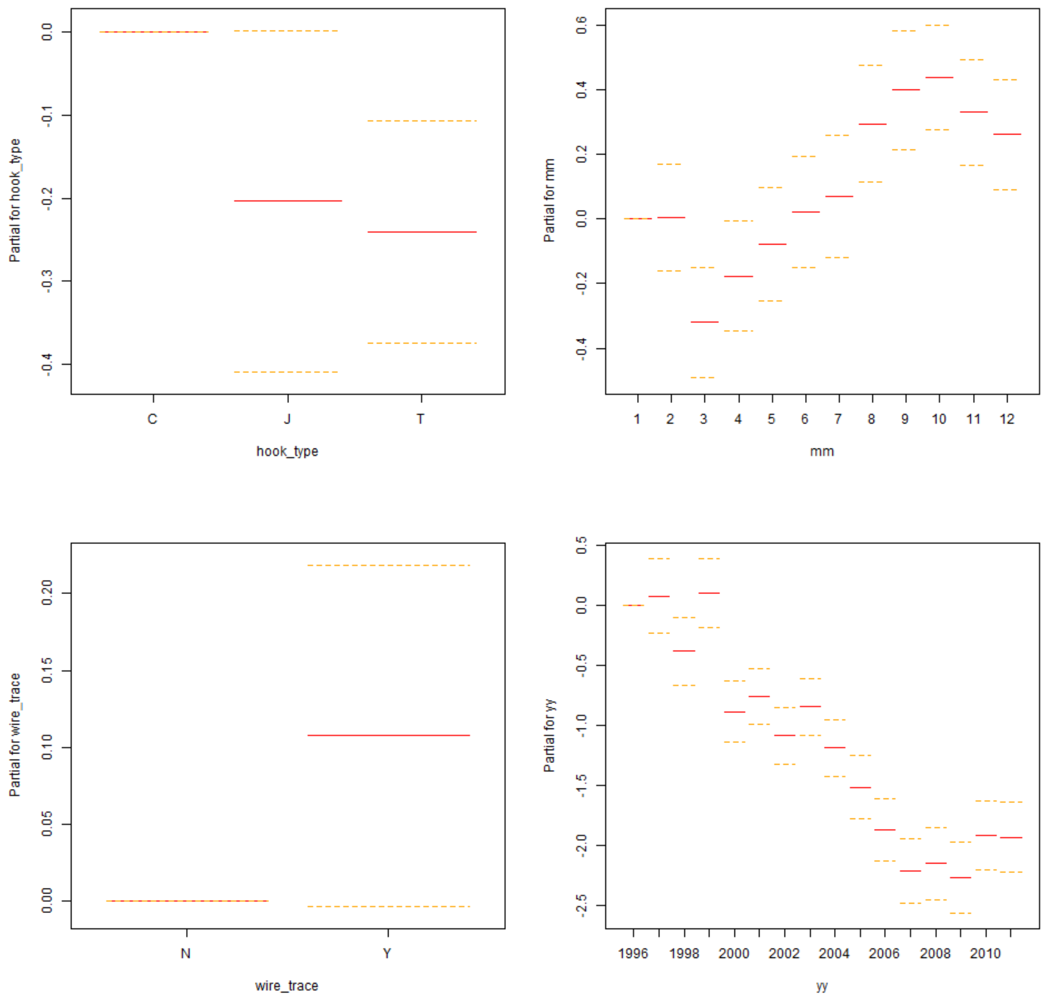


Figure 150: estimates and approximate 95% confidence intervals for parametric terms within the GAM.

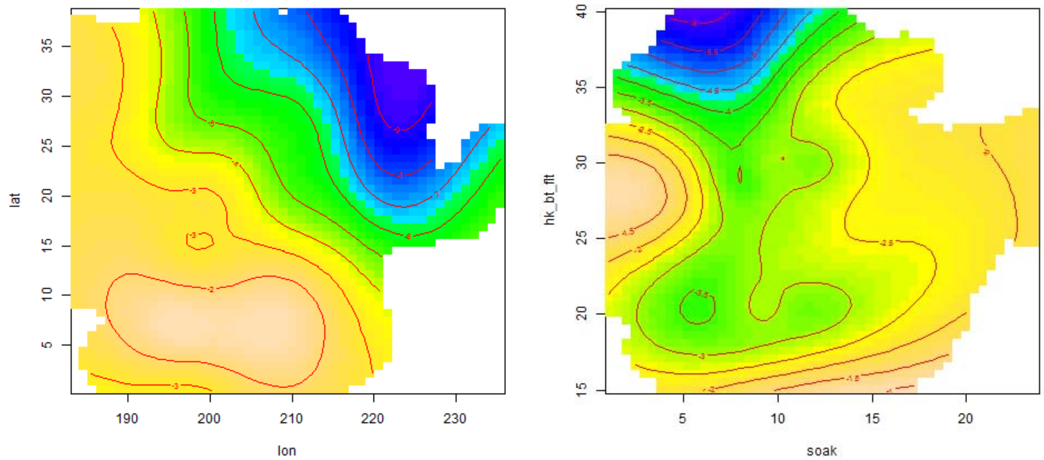


Figure 151: estimated smooth surfaces for the bivariate smooth components within the GAM.

8.1.2 Hook-level catch rate of oceanic whitetip sharks

8.1.2.1 Data Exploration

There are no clear patterns between catch rates of oceanic whitetip sharks and hook positions at which they were caught (Figure 152). Furthermore, there are no demarked trends in how the relationship between catch rates and hook positions changes with varying levels of the covariates considered in the analysis.

Figure 152 : Distributions of per-hook catch-rate responses against the explanatory variables considered in the analysis, for oceanic whitetip sharks caught in the M7_HD fishery. Dots were slightly jittered to minimise overplotting and help visualisation and their colour convey the basket hook position of the respective observation (file "M7_HD-OCS_perHookRatesVsCovs.png").

8.1.2.2 Modelling results

The effect of hook position on catch rates was conditioned solely on species and fishery. A GAM was fitted with log-link and Tweedie error distribution. An initial model run was conducted for the purposes of estimating the governing Tweedie parameter – subsequently refitted with the estimated Tweedie parameter specified

The Tweedie parameter was estimated to be 1.046, with the model returning an adjusted-R^2 of 0.002 and a deviance explained of 2.1%.

The model assumptions were assessed. Non-linearities are not relevant as the hook position is fitted as a factor variable. The adequacy of the assumed Tweedie distribution was assessed via Quantile Residuals as indicated in Figure 13– these ought to be approximately Normally distributed, producing an approximately straight-line. Potential correlation in the errors was similarly assessed on quantile residuals, both by examination of acf plots and a runs test. The data were ordered by set date within trip IDs, so the correlations are sought via sets that are temporally, and likely spatially, close.

In this case the Tweedie distribution is acceptable and there is no significant residual correlation once this zero rich distribution has been accounted for. The runs-test returns a *p*-value of 0.98.

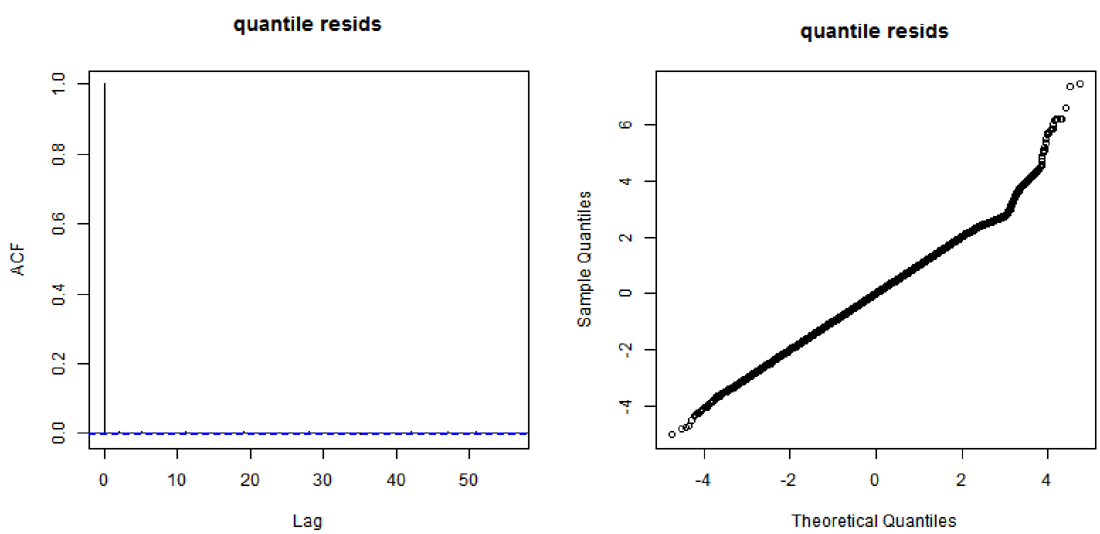


Figure 153: diagnostic plots for the quantile residuals from the GAM. The error distribution is assumed to be a Tweedie distribution, with key parameter estimated.

Estimated catch rate for each hook position and approximate 95% confidence intervals are presented in Figure 14 and Table 5.

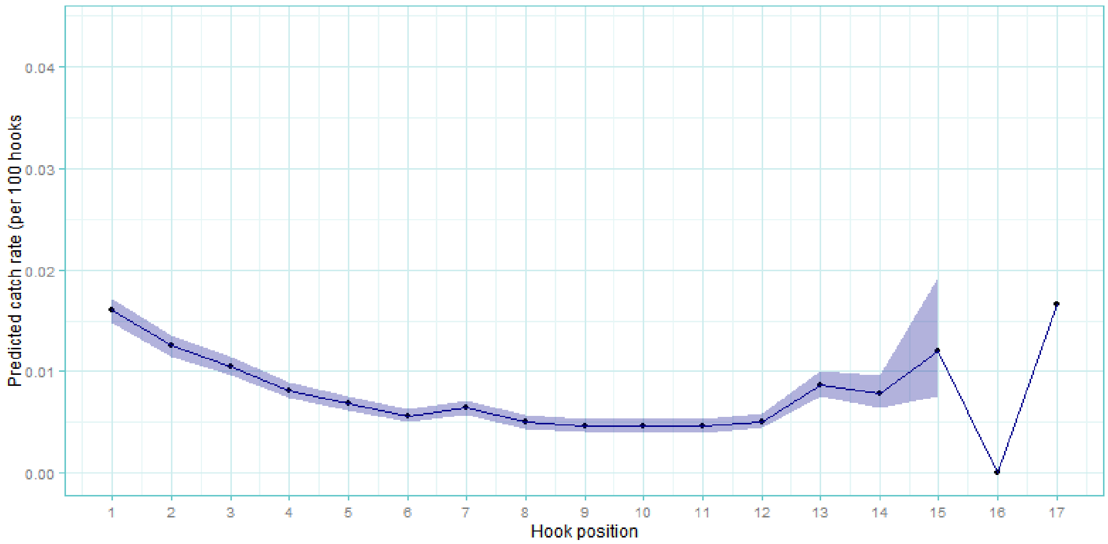


Figure 154: estimated mean catch rate (per 100 hooks) at hook-position with approximate 95% confidence envelope.

Table 43: Estimates of catch rates for each of the hook positions. Confidence intervals are approximate 95%. Percentage of total catch is calculated from overall estimated catch over hook positions presented. Confidence intervals for percentages are based on this naïve translation of cate rate.

Hook position	Catch rate (per 100 hooks)	SE	Lower CI	Upper CI	Catch (as % of total)	Lower CI (as % of total)	Upper CI (as % of total)
1	0.016	1.0386	0.0148	0.0172	11.81	10.97	12.72
2	0.0125	1.0435	0.0115	0.0136	9.23	8.49	10.03
3	0.0105	1.0473	0.0096	0.0115	7.75	7.08	8.49
4	0.0081	1.0536	0.0074	0.009	6.02	5.44	6.67
5	0.0068	1.0584	0.0061	0.0076	5.05	4.52	5.65
6	0.0056	1.0644	0.005	0.0063	4.14	3.67	4.68
7	0.0064	1.0604	0.0057	0.0072	4.72	4.2	5.29
8	0.0051	1.0678	0.0044	0.0057	3.74	3.28	4.25
9	0.0047	1.0708	0.0041	0.0053	3.44	3.01	3.94
10	0.0046	1.0719	0.004	0.0053	3.42	2.99	3.92
11	0.0046	1.0742	0.004	0.0053	3.43	2.98	3.95
12	0.0051	1.0756	0.0045	0.0059	3.8	3.3	4.39

13	0.0087	1.072	0.0076	0.01	6.44	5.62	7.38
14	0.0079	1.1068	0.0065	0.0096	5.82	4.77	7.1
15	0.012	1.2684	0.0075	0.0192	8.89	5.58	14.16
16	0	NA	0	NA	0	0	NA
17	0.0166	1.6392	0.0063	0.0438	12.29	4.67	32.38

The following can be observed from these:

- There is no particular relationship between the hook-positions and the catch rates.
- The shallowest position (1) is estimated at 0.03 per 100 hooks [0.018 to 0.049, 95% CI]. This translates to approximately 6.9% [4.21% to 11.4%, 95% CI] of the catch of this species observed in this fisheries data.

8.1.3 Condition of oceanic whitetip sharks at time of retrieval

8.1.3.1 Data Exploration

There are no substantial patterns between the condition of caught oceanic whitetip sharks at the time of retrieval and most of the covariates considered in the present analysis (Figure 155). There are however some slight indications that, on average, the proportion of dead oceanic whitetip sharks might have increased: (i) if sets took place in June; (ii) when observed sets used brachlines with wire leaders; or (iii) as the number of hooks between float used in the set were larger.

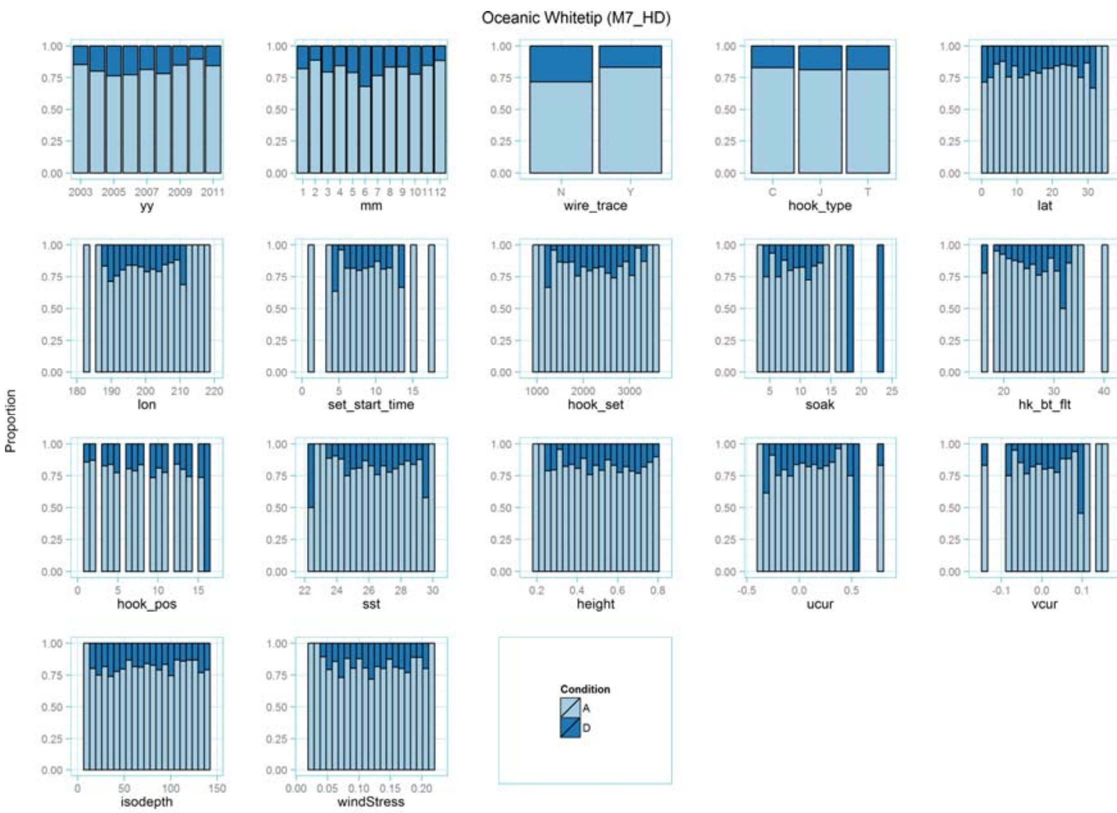


Figure 155: Distributions of condition responses (D= dead, A = alive), expressed in terms of proportion, against the explanatory variables considered for the analysis of oceanic whitetip sharks caught in the M7_HD fishery (file "M7_HD-OCS_CondtnVsCovs.png").

8.1.3.2 Modelling results

Condition (dead/alive) was modelled as a function of set-level characteristics and associated oceanographic variables. The initial set of potential covariates was thinned on the basis of gross collinearity, as determined by Generalized Variance Inflation Factors (GVIFs). A range of potential interactions and non-linear terms were specified *a priori* as per Table 3. These were further thinned during the modelling process in some cases where estimation was not possible e.g. interactions for factor variables where certain combinations of factors were not observed in the dataset.

The model fitted was a GAM with logit-link and binomially-distributed errors. Model "selection" was conducted within the GAM using shrinkage, so model terms of little relevance have their effective degrees of freedom shrunk towards zero. In terms of parametric components, backwards selection (p-value<0.05) was used, but not on the key factors of wire-trace, shark-bait, shark-lines and hook type (where applicable). These were generally retained in the model for interpretative interest. The results of this process are summarised in **Table 4**.

In terms of fit to the data, the software generated adjusted R^2 and deviance explained were 0.042 and 5.89% respectively, indicating poor fit. However these measures are of dubious utility for binomial error models. Raw predictive power to the current dataset was also assessed via confusion matrices, providing Sensitivity (true positive rate) = 0.64, Specificity (true negative rate) = 0.60 and a misclassification rate of 0.39 when using the response mean as a decision boundary. These are against the training data and will be over-estimates –cross-validated estimates would be preferred with future analysis.

Table 44: Model terms from the generalized additive model. DF = Degrees of Freedom. EDF=Effective Degrees of Freedom. Shrinkage has been applied as a means of model selection. Terms of little relevance have their EDF shrunken.

Model terms	EDF	DF	Chi or F-stat	*p*-value
yy		8	16.98	0.0303
mm		11	28.62	0.0026
hook_type		2	0.26	0.8792
wire_trace		1	3.46	0.0628
s(set_start_time)	0.00	9	0.00	0.5901
s(sst)	0.12	9	0.13	0.2959
s(height)	0.00	9	0.00	0.4587
s(ucur)	2.39	9	8.86	0.0081
s(vcur)	2.29	9	5.57	0.0512
s(isodepth)	0.00	9	0.00	0.6042
s(windStress)	0.00	9	0.00	0.5197
s(soak)	6.48	9	23.05	0.0005
s(hook_pos)	1.44	9	15.45	0.0001
s(hk_bt_flt)	2.13	9	11.53	0.0014

The model assumptions were assessed. Non-linearities are assumed to be captured via the GAM which optimises complexity. The adequacy of the assumed binomial distribution was assessed via Quantile Residuals (Dunn & Smythe, 1996), as indicated in Figure 16 – these ought to be approximately Normally distributed, producing an approximately straight-line. Potential correlation in the errors was similarly assessed on quantile residuals, both by examination of acf plots and a runs test. The data were ordered by set date within trip IDs, so the correlations are sought via sets that are temporally, and likely spatially, close. The residuals are distributionally acceptable and little autocorrelation is in evidence, the runs-test providing a *p*-value of 0.031.

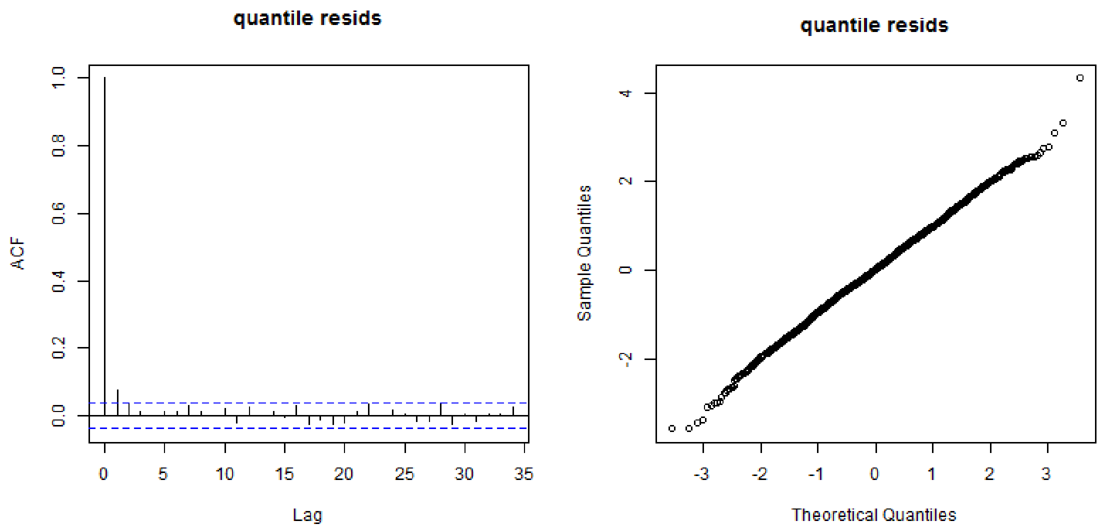


Figure 156: diagnostic plots for the quantile residuals from the GAM. The error distribution is assumed to be a Tweedie distribution, with key parameter estimated.

The following are link-scale estimates of the components of the GAM. All terms in the model are represented here regardless of practical/statistical significance although near-zero EDF relationships are excluded. Generally only top-line results will be presented outside of the reference case. Where applicable, the parametric terms are presented first, followed by smooth terms and then interactions.

The following observations can be made:

- Little significant pattern with regards the key factors of wire-trace, shark-bait and shark-lines.
- A complex non-monotone relationship with isodepth.
- Generally increasing morality rates with current (u).
- Generally increasing morality rates with current (v) and wind-stress.
- A marked dip in mortality with hook positions around 12-14.
- Higher mortality with lower numbers of hooks in baskets.

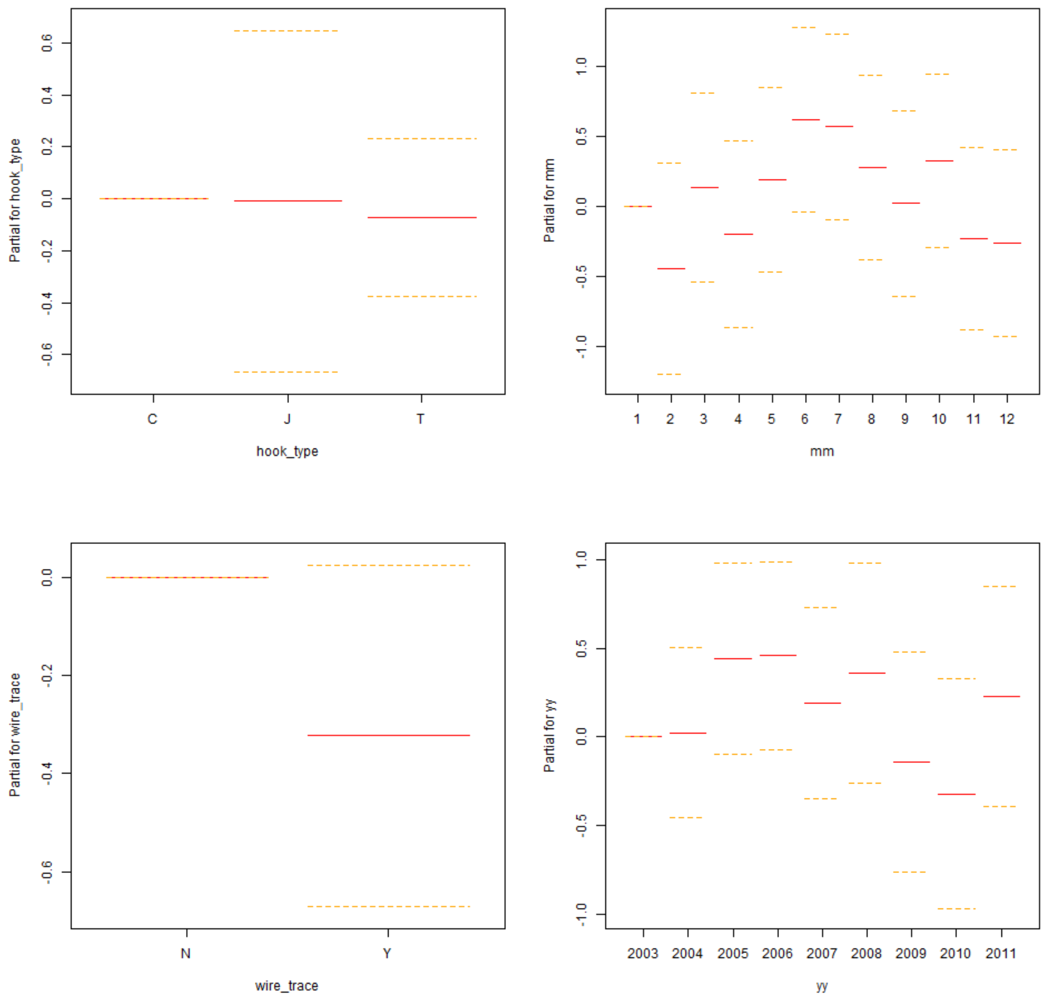


Figure 157: estimates and approximate 95% confidence intervals for parametric terms within the GAM.

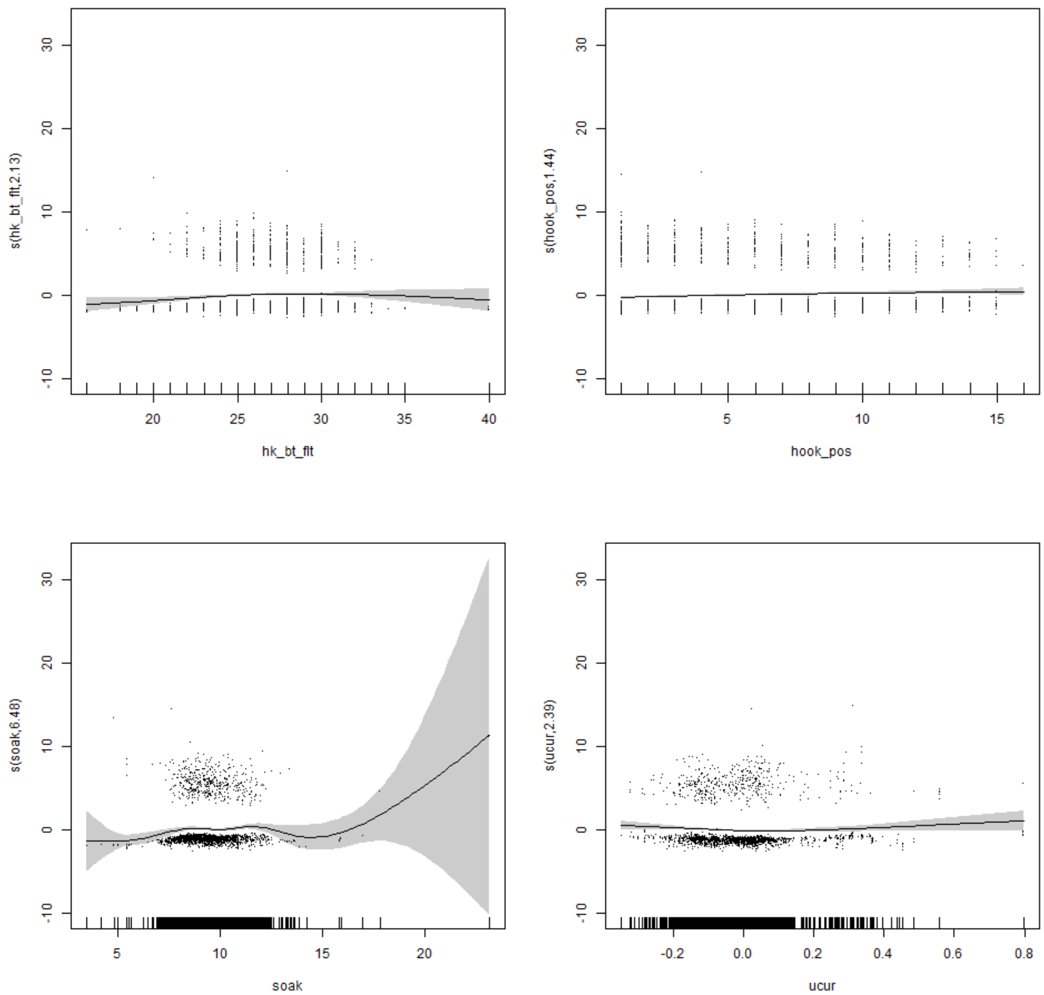


Figure 158: estimated smooths and approximate 95% confidence envelopes for smooth terms within the GAM.

8.2 SILKY SHARKS (FAL)

8.2.1 Catch rate of silky Sharks (FAL)

8.2.1.1 Data Exploration

While predominantly absent in sets from observed trips in the M7_HD fishery, within-set bycatch levels of silky sharks ranged mainly between 1 shark/set (in about 1200 sets) and 3 sharks/set (in about 150 sets) (Figure 159). However there were up to a maximum of 68 silky sharks in one single set.

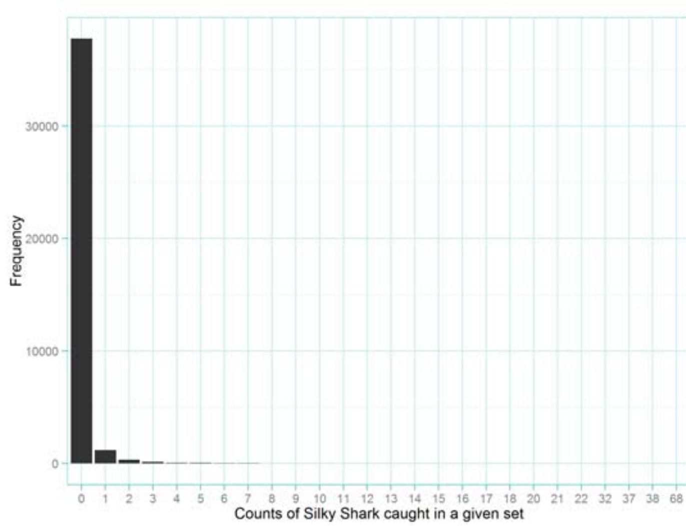


Figure 159: Distribution of counts per set of silky shark caught in observed trips of the M7_HD longline fishery (M7_HD-FAL_CatchNumFreqs.png)

The predominance of zero catches in observed sets hinders the clear visualisation of patterns between catch rates of silky sharks and the covariates considered in the present analysis (Figure 160). However, there are some sparse indications that, on average, higher catch rates were likely to occur when, e.g.:

- Hook types “C” or “T” were employed, or wire leader were in use;
- Sets were performed between 2000 and 2002;
- Longlines were laid within 0-10 degrees latitude, and within 190-200 degrees longitude.
- Soak time varied between 5-14 hours;

- The number of hooks between floats ranged between 25-30
- Average sea surface temperature was over 25^{o}C, and the average depth of the 20^{o}C isotherm was roughly between 50-100 meters.

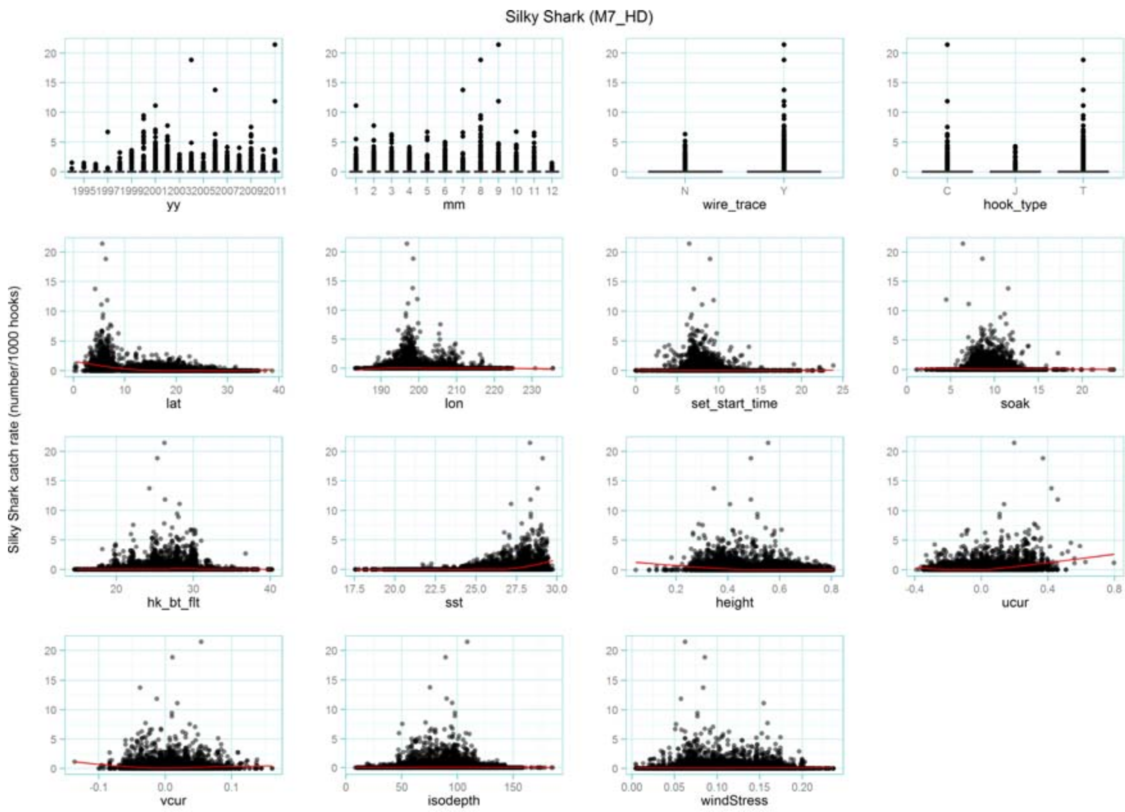


Figure 160: Distributions of set-level catch-rate responses against the explanatory variables considered for the analysis of silky sharks caught in the M7_HD fishery (file "M7_HD-FAL_CatchRatesVsCovs.png")

8.2.1.2 Modelling results

Modelling details are as per the reference case. The model returned an adjusted-R^2 of 0.27 and a deviance explained of 52.9%.

Table 45: Model terms from the generalized additive model. DF = Degrees of Freedom. EDF=Effective Degrees of Freedom. Shrinkage has been applied as a means of model selection. Terms of little relevance have their EDF shrunken.

Model terms	EDF	DF	Chi or F-stat	*p*-value
yy		15	6.78	0.0000
mm		11	12.65	0.0000
hook_type		2	2.28	0.1018
wire_trace		1	0.14	0.7127
s(lon,lat)	24.98	29	129.97	0.0000
s(soak,hk_bt_flt)	27.65	29	1.80	0.0028

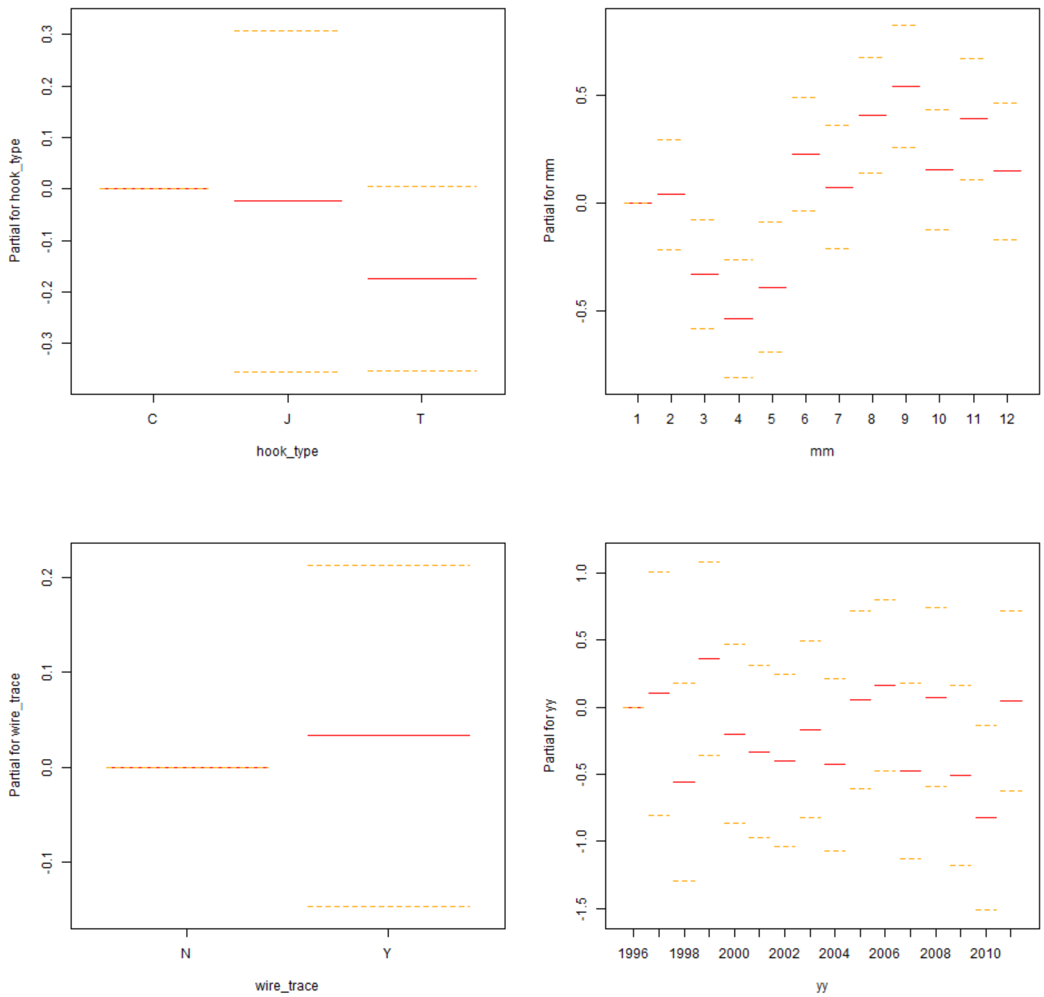


Figure 161: estimates and approximate 95% confidence intervals for parametric terms within the GAM.

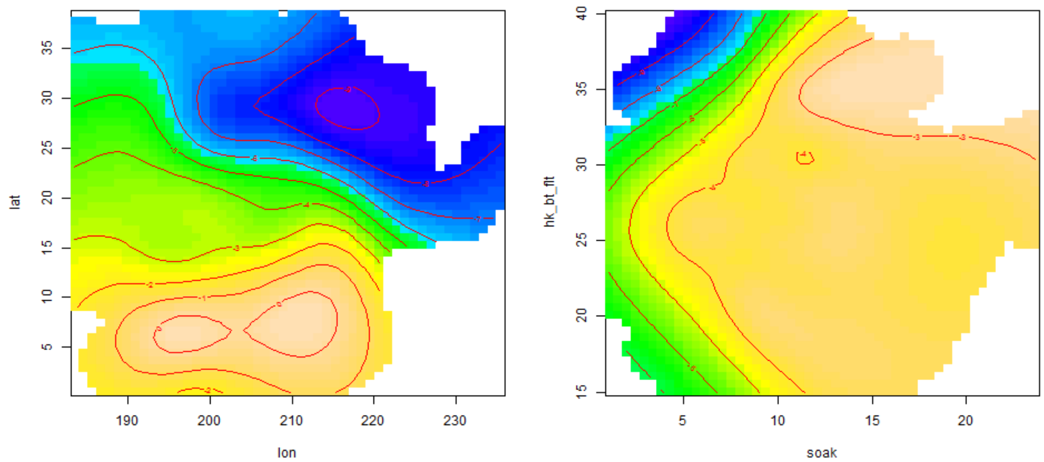


Figure 162: estimated smooth surfaces for the bivariate smooth components within the GAM.

8.2.2 Hook-level catch rate of silky sharks

8.2.2.1 Data Exploration

There are no clear patterns between catch rates of silky shark and hook positions at which they were caught (Figure 163). Furthermore, there are no demarked trends in how the relationship between catch rates and hook positions changes with varying levels of the covariates considered in the analysis.

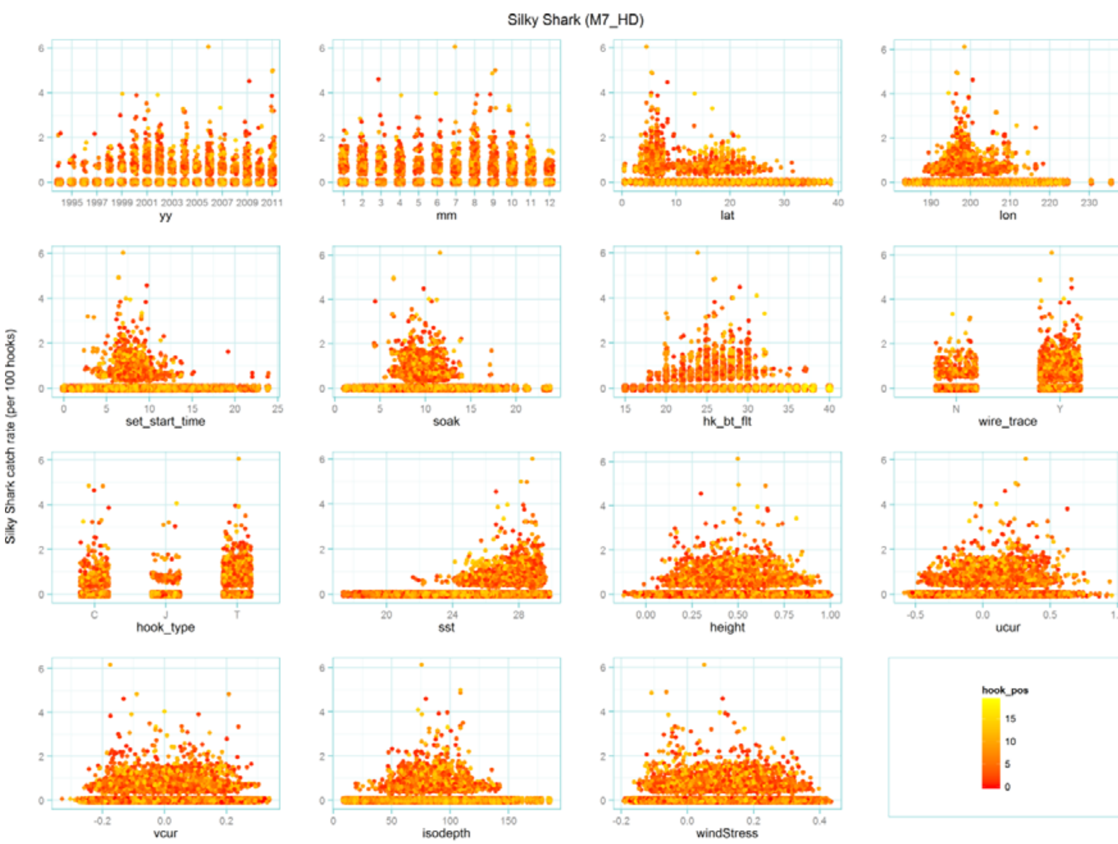


Figure 163: Distributions of per-hook catch-rate responses against the explanatory variables considered in the analysis, for silky sharks caught in the M7_HD fishery. Dots were slightly jittered to minimise overplotting and help visualisation and their colour convey the basket hook position of the respective observation (file “M7_HD-FAL_perHookRatesVsCovs.png”).

8.2.2.2 Modelling results

Modelling details are as per the reference case. The model returned an adjusted-R^2 of <0.0001 and a deviance explained of 1.14%.

Table 46: Estimates of catch rates for each of the hook positions. Confidence intervals are approximate 95%. Percentage of total catch is calculated from overall estimated catch over hook positions presented. Confidence intervals for percentages are based on this naïve translation of cate rate.

Hook position	Catch rate (per 100 hooks)	SE	Lower CI	Upper CI	Catch (as % of total)	Lower CI (as % of total)	Upper CI (as % of total)
1	0.009	1.0556	0.0081	0.01	10.77	9.68	11.97
2	0.0083	1.0579	0.0074	0.0092	9.85	8.82	11
3	0.0063	1.0658	0.0055	0.0071	7.47	6.6	8.47
4	0.0059	1.0677	0.0052	0.0067	7.01	6.17	7.97
5	0.005	1.0733	0.0043	0.0057	5.91	5.14	6.79
6	0.0043	1.0782	0.0037	0.005	5.15	4.44	5.97
7	0.0041	1.0801	0.0035	0.0048	4.89	4.2	5.68
8	0.0037	1.0845	0.0031	0.0043	4.36	3.72	5.12
9	0.004	1.0817	0.0034	0.0046	4.72	4.05	5.51
10	0.0038	1.0844	0.0032	0.0045	4.53	3.87	5.31
11	0.0041	1.0837	0.0035	0.0049	4.94	4.22	5.78
12	0.0036	1.0956	0.003	0.0043	4.33	3.62	5.17
13	0.0061	1.0932	0.0051	0.0072	7.21	6.06	8.59
14	0.0065	1.1271	0.0052	0.0083	7.79	6.16	9.85
15	0.0093	1.3407	0.0052	0.0165	11.07	6.23	19.67

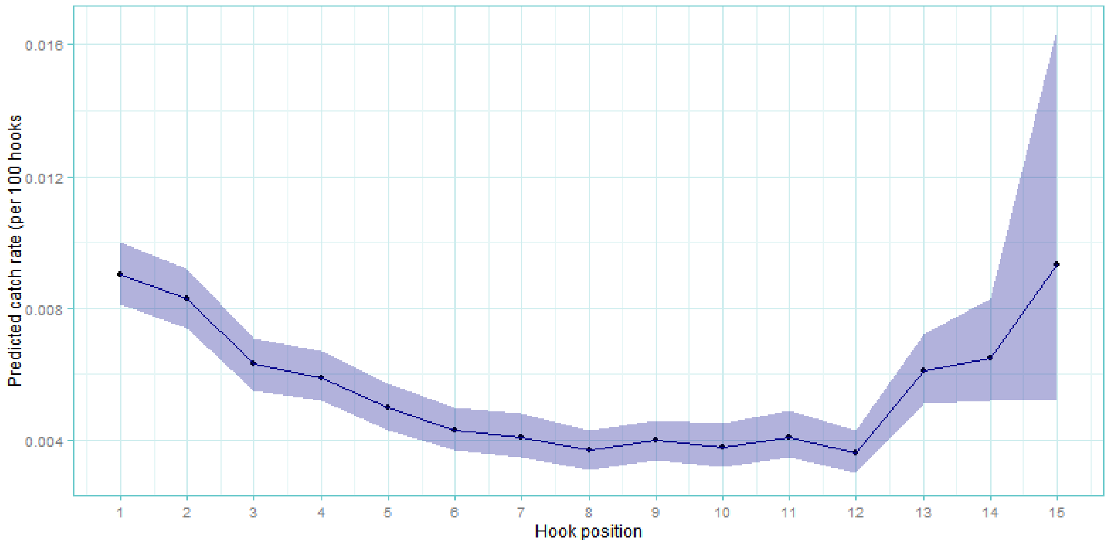


Figure 164: estimated mean catch rate (per 100 hooks) at hook-position with approximate 95% confidence envelope.

8.2.3 Condition of silky at time of retrieval

8.2.3.1 Data Exploration

There are no substantial patterns between the condition of caught silky sharks at the time of retrieval and most of the covariates considered in the present analysis (Figure 165). There are however some slight indications that, on average, the proportion of dead silky sharks might have increased: (i) as sets were performed at locations nearer the equator and towards westerly regions of the fishery; or (ii) if longlines were set when wind stress levels were higher or lower than 0.1 Newton/m^2.

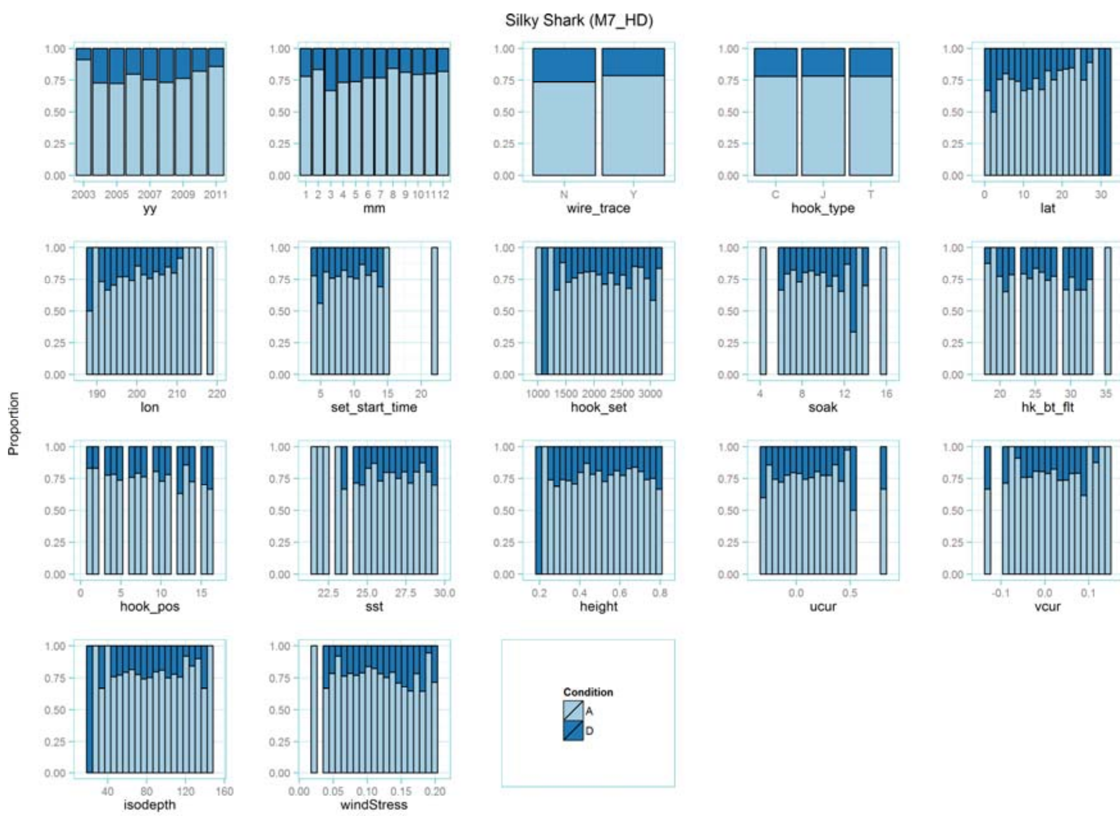


Figure 165: Distributions of condition responses (D= dead, A = alive), expressed in terms of proportion, against the explanatory variables considered for the analysis of silky sharks caught in the M7_HD fishery (file "M7_HD-FAL_CondtnVsCovs.png").

8.2.3.2 Modelling results

Modelling details are as per the reference case. The model returned an adjusted-R^2 of 0.06 and a deviance explained of 8.09%. Raw predictive power to the current dataset was also assessed via confusion matrices on the training data, providing Sensitivity (true positive rate) = 0.67, Specificity (true negative rate) = 0.59 and a misclassification rate of 0.39 when using the response mean as a decision boundary. These are against the training data and will be over-estimates.

Table 47: Model terms from the generalized additive model. DF = Degrees of Freedom. EDF=Effective Degrees of Freedom. Shrinkage has been applied as a means of model selection. Terms of little relevance have their EDF shrunken.

Model terms	EDF	DF	Chi or F-stat	*p*-value
yy		8	20.71	0.0079
mm		11	29.91	0.0016
hook_type		2	1.27	0.5299
wire_trace			1	0.33
s(set_start_time)	0.00	9	0.00	1.0000
s(sst)	0.86	9	4.66	0.0154
s(height)	4.86	9	15.61	0.0024
s(ucur)	0.00	9	0.00	0.6093
s(vcur)	0.46	9	0.60	0.2365
s(isodepth)	0.00	9	0.00	0.5662
s(windStress)	1.78	9	5.02	0.0378
s(soak)	4.71	9	14.20	0.0053
s(hook_pos)	2.08	9	13.95	0.0004
s(hk_bt_flt)	0.00	9	0.00	0.7107
s(lat,lon)	22.40	29	39.93	0.0060

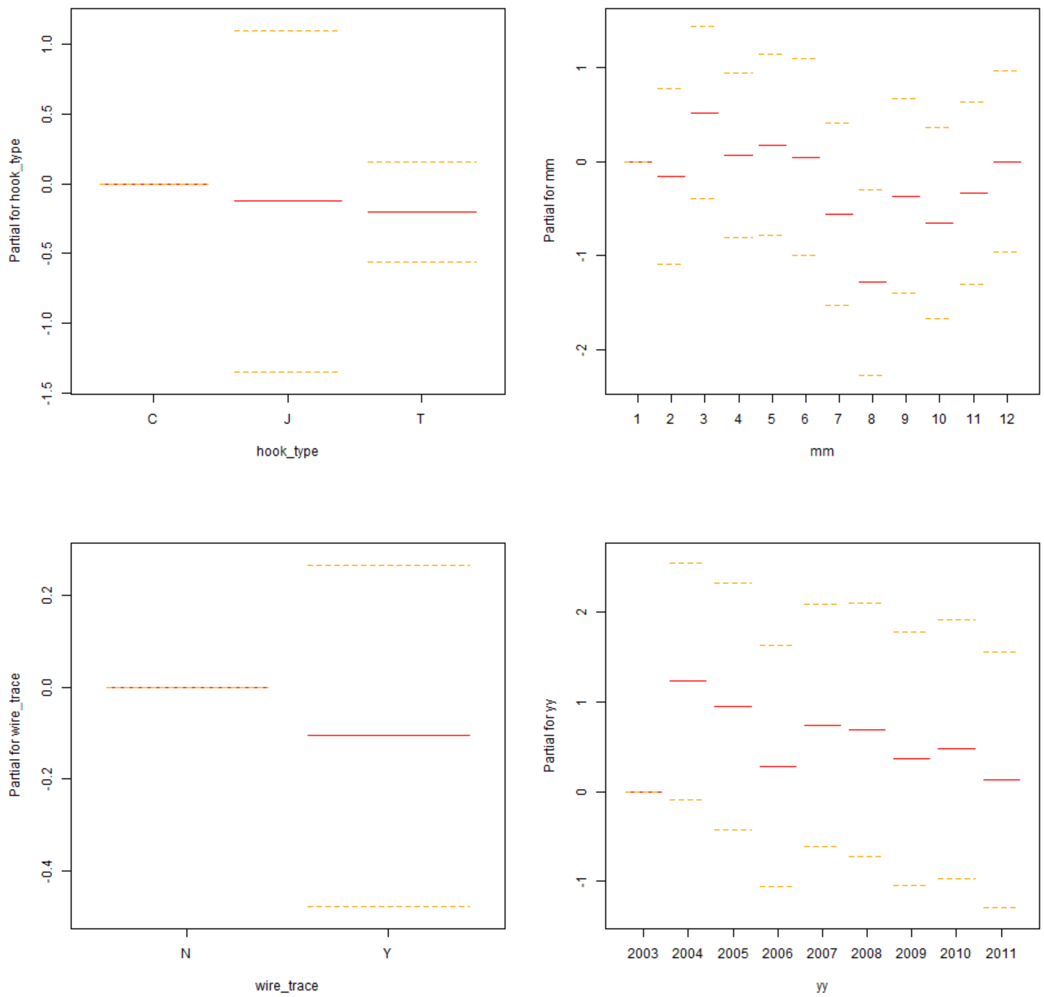


Figure 166: estimates and approximate 95% confidence intervals for parametric terms within the GAM.

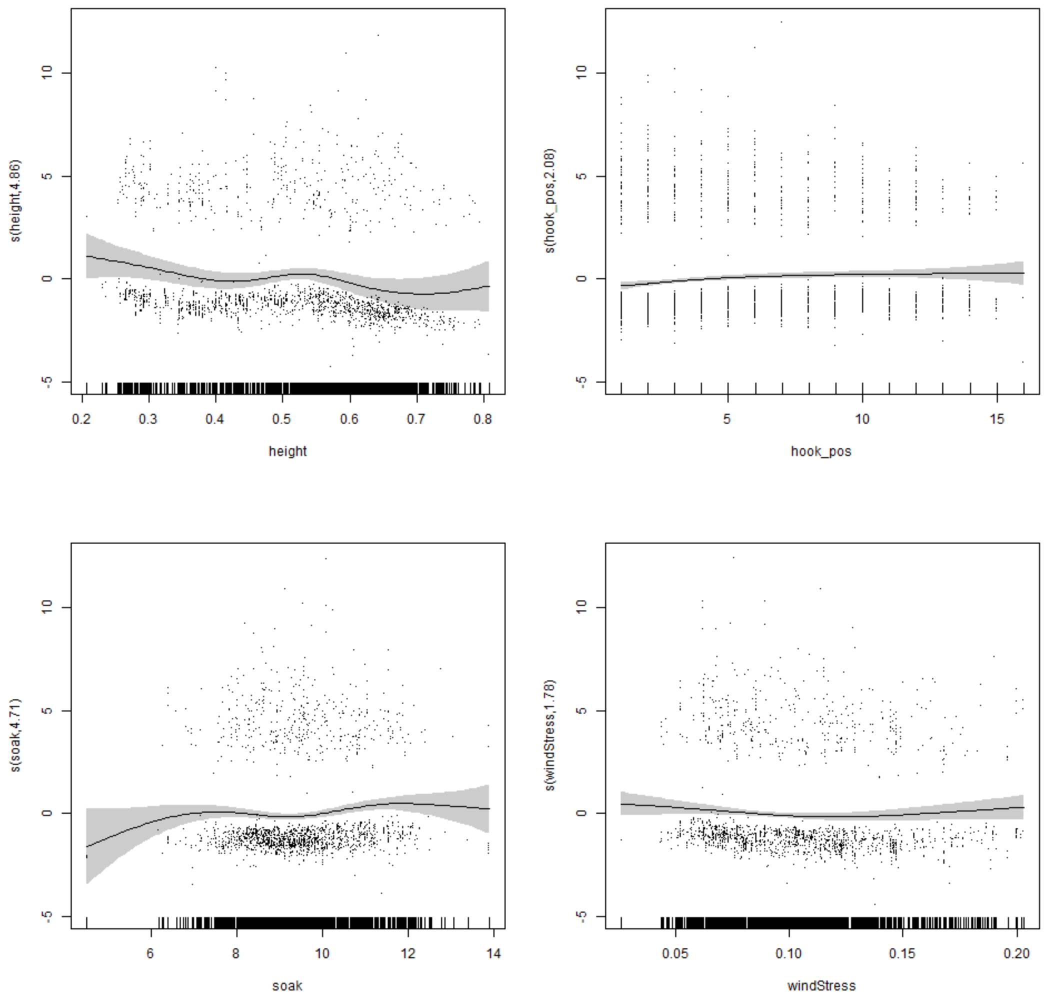


Figure 167: estimated smooths and approximate 95% confidence envelopes for smooth terms within the GAM.

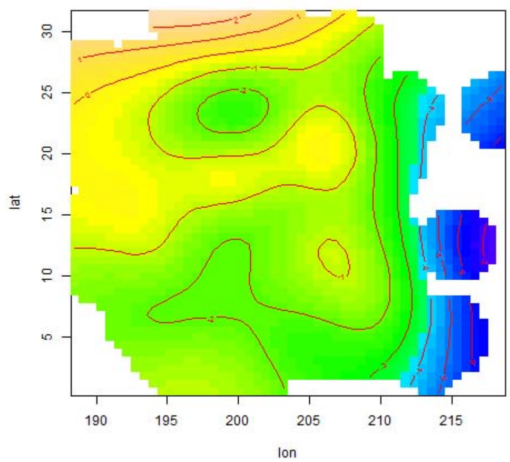


Figure 168: estimated smooth surfaces for the bivariate smooth components within the GAM.

8.3 THRESHER SHARKS (THR)

8.3.1 Catch rate of thresher sharks (THR)

8.3.1.1 Data Exploration

While predominantly absent in sets from observed trips in the M7_HD fishery, within-set bycatch levels of tresher sharks ranged mainly between 1 shark/set (in about 6200 sets) and 5 sharks/set (in about 130 sets) (Figure 169). However there were up to a maximum of 94 silky sharks in one single set.

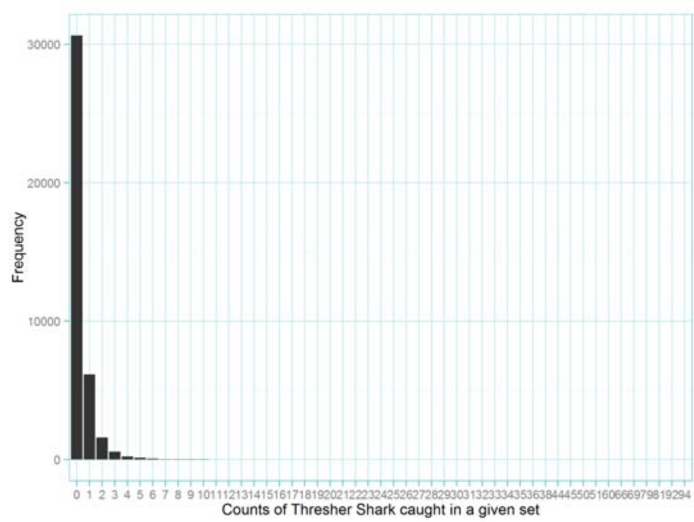


Figure 169: Distribution of counts per set of thresher shark caught in observed trips of the M7_HD longline fishery (M7_HD-THR_CatchNumFreqs.png)

The predominance of zero catches in observed sets hinders the clear visualisation of patterns between catch rates of thresher sharks and the covariates considered in the present analysis (Figure 170). However, there are some sparse indications that, on average, higher catch rates were likely to occur when, e.g.:

- Wire leader and hook types "C" or "T" were in use;
- Longlines were set within 5-10 degrees latitude, and within 190-205 degrees longitude;
- Sets started between 5:00-10:00 hours.

- Soak time varied between 7-12 hours;
- Average sea surface temperature was over 25^{o}C, and the average depth of the 20^{o}C isotherm was roughly between 50-100 meters.

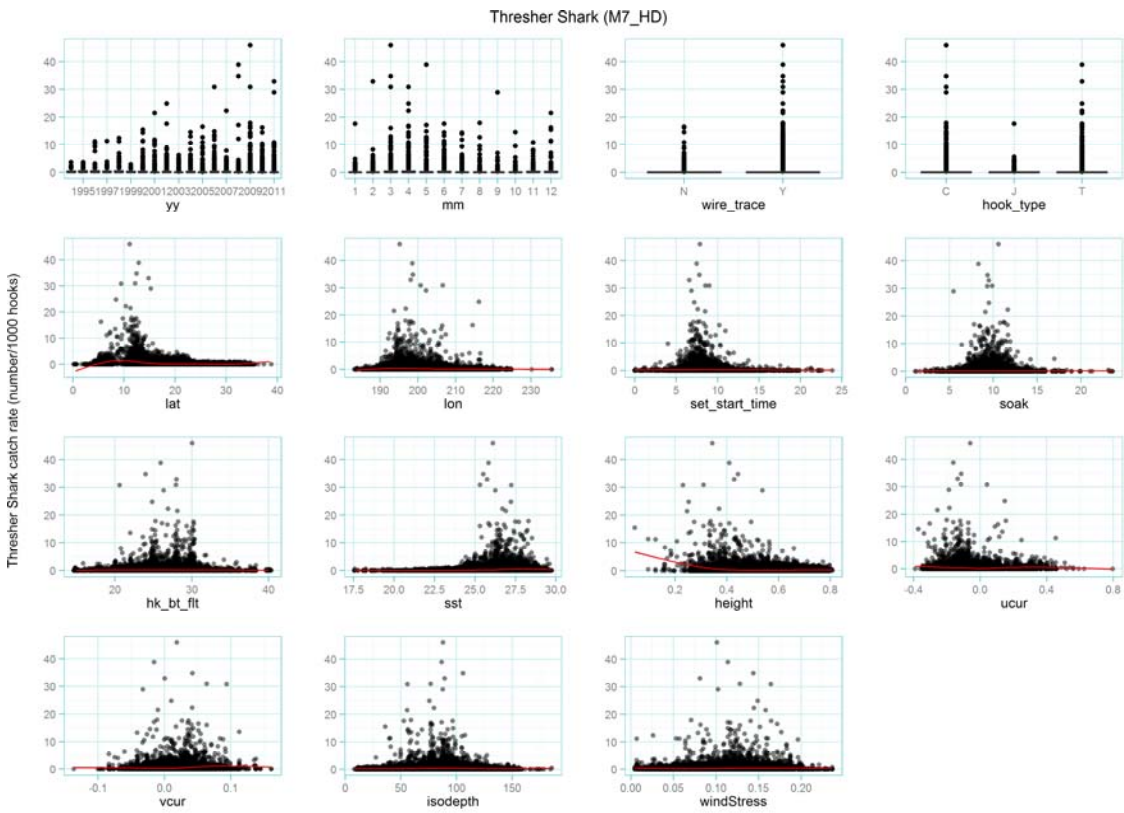


Figure 170: Distributions of set-level catch-rate responses against the explanatory variables considered for the analysis of thresher sharks caught in the M7_HD fishery (file M7_HD-THR_CatchRatesVsCovs.png")

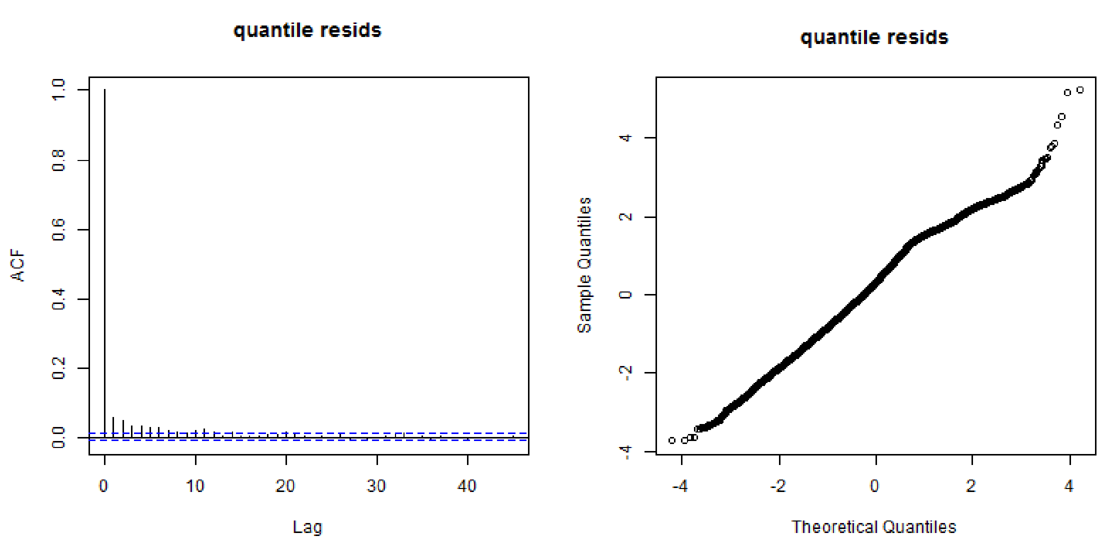
8.3.1.2 Modelling results

Modelling details are as per the reference case. The model returned an adjusted-R^2 of 0.18 and a deviance explained of 33.3%.

Table 48: Model terms from the generalized additive model. DF = Degrees of Freedom. EDF=Effective Degrees of Freedom. Shrinkage has been applied as a means of model selection. Terms of little relevance have their EDF shrunken.

Model terms	EDF	DF	Chi or F-stat	*p*-value
yy		15	2.52	0.0010
mm		11	6.49	0.0000
hook_type		2	1.11	0.3292
wire_trace		1	0.49	0.4851
s(lon,lat)	26.84	29	51.96	0.0000
s(soak,hk_bt_flt)	23.22	29	0.99	0.1885

Residual autocorrelation under this model is highly significant (runs-test *p*-value <<0.001) and persists at low levels over a few lags. Convergence times for mixed models equivalents were prohibitive for this current study. Significant results under this model should be treated with caution, as positive autocorrelation is likely to produce unduly small variance estimates.



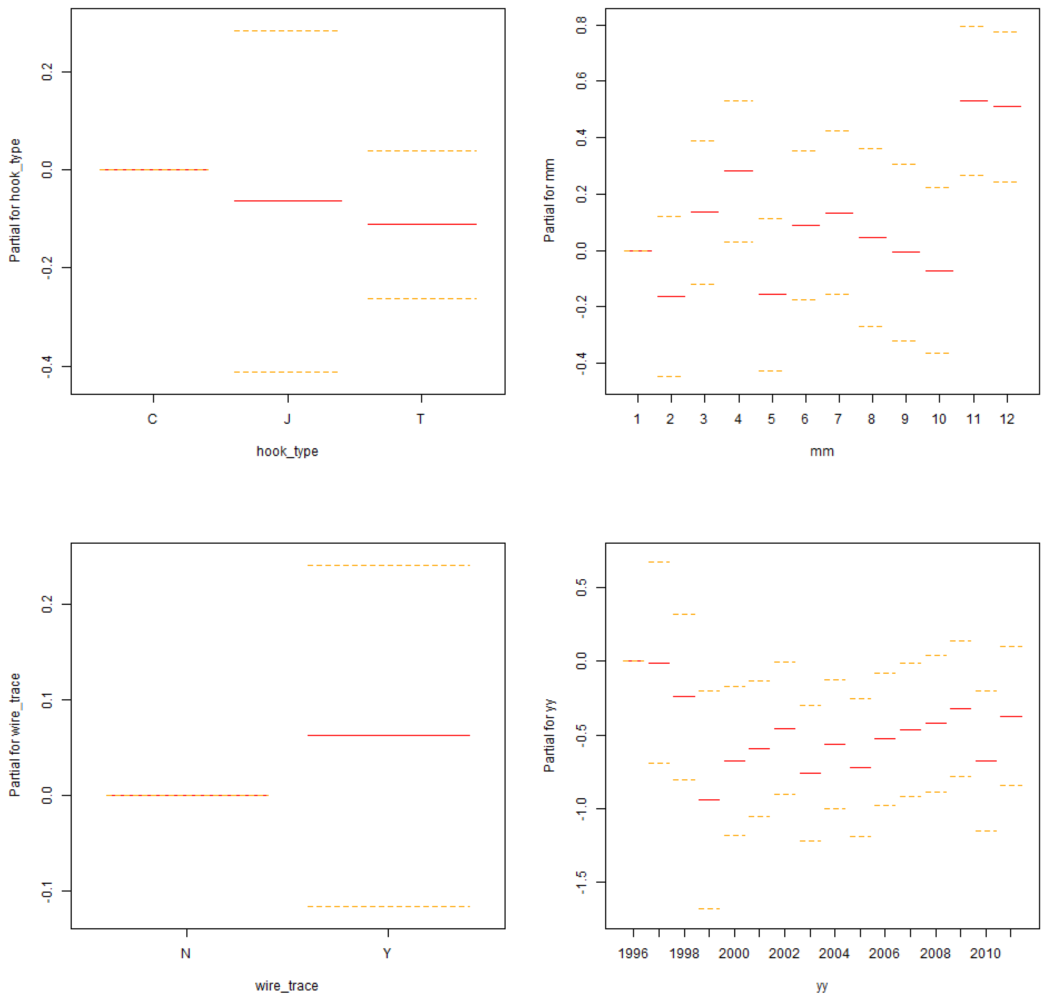


Figure 171: estimates and approximate 95% confidence intervals for parametric terms within the GAM.

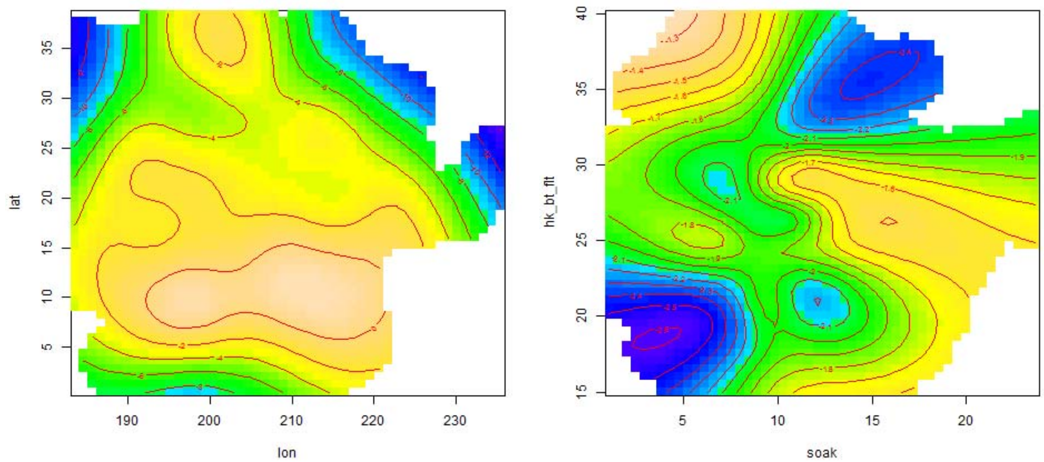


Figure 172: estimated smooth surfaces for the bivariate smooth components within the GAM.

8.3.2 Hook-level catch rate of thresher sharks

8.3.2.1 Data Exploration

Figure 173 suggests that catch rates of thresher sharks in the M7_HD fishery tend to be larger at higher hook position numbers (i.e. at deeper hooks). However, there are no demarked trends in how the relationship between catch rates and hook positions changes with varying levels of the covariates considered in the present analysis.

Figure 173: Distributions of per-hook catch-rate responses against the explanatory variables considered in the analysis, for thresher sharks caught in the M7_HD fishery. Dots were slightly jittered to minimise overplotting and help visualisation and their colour convey the basket hook position of the respective observation (file “M7_HD-THR_perHookRatesVsCovs.png”).

8.3.2.2 Modelling results

Modelling details are as per the reference case. The model returned an adjusted-R^2 of ~0 and a deviance explained of 3.5%.

Table 49: Estimates of catch rates for each of the hook positions. Confidence intervals are approximate 95%. Percentage of total catch is calculated from overall estimated catch over hook positions presented. Confidence intervals for percentages are based on this naïve translation of cate rate.

Hook position	Catch rate (per 100 hooks)	SE	Lower CI	Upper CI	Catch (as % of total)	Lower CI (as % of total)	Upper CI (as % of total)
1	0.0049	1.0733	0.0042	0.0056	1.16	1.01	1.33
2	0.0091	1.0551	0.0082	0.0101	2.16	1.94	2.39
3	0.0123	1.0479	0.0112	0.0135	2.93	2.67	3.21
4	0.0153	1.0434	0.0141	0.0166	3.63	3.34	3.95
5	0.0191	1.0392	0.0177	0.0206	4.53	4.2	4.89
6	0.0205	1.038	0.0191	0.0221	4.87	4.53	5.24
7	0.0226	1.0363	0.0211	0.0243	5.38	5.02	5.77
8	0.0233	1.0359	0.0217	0.0249	5.53	5.16	5.92
9	0.0253	1.0347	0.0236	0.027	6	5.61	6.42
10	0.0267	1.0343	0.025	0.0286	6.35	5.95	6.79
11	0.0344	1.0315	0.0323	0.0365	8.17	7.69	8.68
12	0.0322	1.0348	0.0302	0.0345	7.66	7.17	8.19
13	0.0643	1.0313	0.0605	0.0683	15.27	14.38	16.23
14	0.0527	1.0479	0.0481	0.0578	12.53	11.43	13.73
15	0.0582	1.1371	0.0452	0.0749	13.83	10.75	17.79

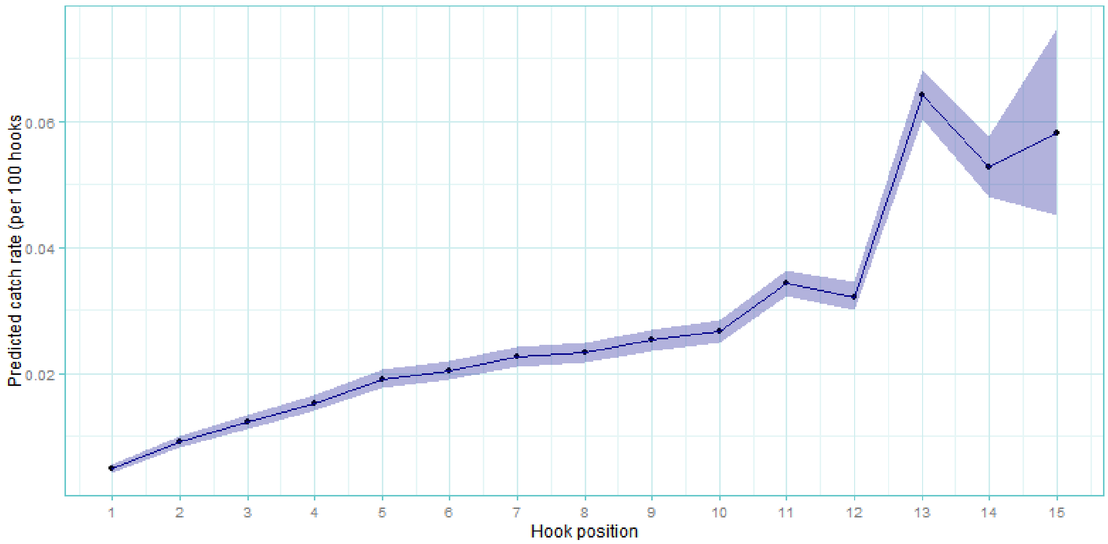


Figure 174: estimated mean catch rate (per 100 hooks) at hook-position with approximate 95% confidence envelope.

8.3.3 Condition of thresher sharks at time of retrieval

8.3.3.1 Data Exploration

There are no substantial patterns between the condition of caught thresher sharks at the time of retrieval and most of the covariates considered in the present analysis (Figure 175). There are however some slight indication that, on average, the proportion of dead silky sharks might have increased when observed sets were performed under certain conditions, e.g.:

- If sets took place in year 2009;
- If hook types "C" or "J" were in use;
- When sets were located below 12 degrees of latitude;
- If the average sea surface height remained below 0.3 meters.

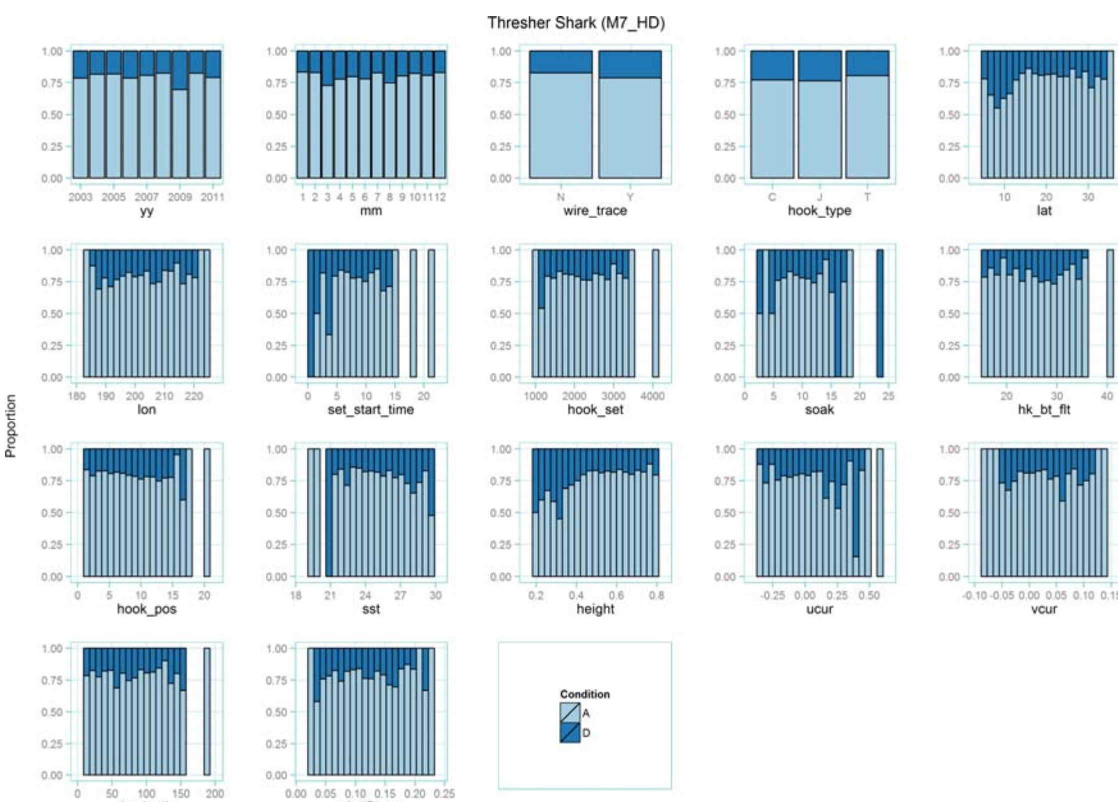


Figure 175: Distributions of condition responses (D= dead, A = alive), expressed in terms of proportion, against the explanatory variables considered for the analysis of thresher sharks caught in the M7_HD fishery (file "M7_HD-THR_CondtnVsCovs.png").

8.3.3.2 Modelling results

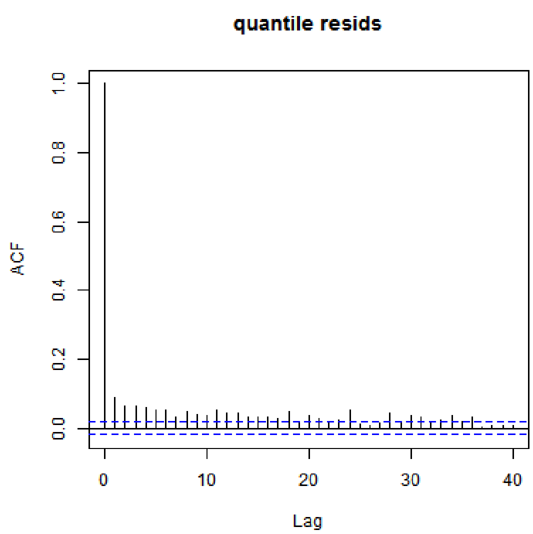
Modelling details are as per the reference case. The model returned an adjusted-R^2 of 0.08 and a deviance explained of 7.5%. Raw predictive power to the current dataset was also assessed via confusion matrices on the training data, providing Sensitivity (true positive rate) = 0.56, Specificity (true negative rate) = 0.67 and a misclassification rate of 0.35 when using the response mean as a decision boundary. These are against the training data and will be over-estimates.

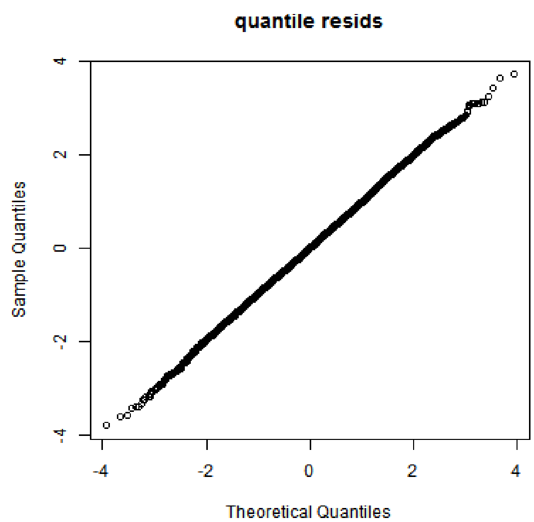
Table 50: Model terms from the generalized additive model. DF = Degrees of Freedom. EDF=Effective Degrees of Freedom. Shrinkage has been applied as a means of model selection. Terms of little relevance have their EDF shrunken.

Model terms	EDF	DF	Chi or F-stat	*p*-value
yy		8	43.04	0.0000
mm		11	26.70	0.0051
hook_type		2	0.96	0.6203
wire_trace		1	1.07	0.3007
s(set_start_time)	8.88	9	34.80	0.0000
s(sst)	3.91	9	17.32	0.0003
s(height)	5.62	9	25.57	0.0000
s(ucur)	0.00	9	0.00	1.0000
s(vcur)	7.18	9	57.83	0.0000
s(isodepth)	7.12	9	17.30	0.0108
s(windStress)	7.82	9	64.60	0.0000
s(soak,hook_pos)	2.69	29	17.78	0.0000
s(lat,lon)	24.52	29	136.23	0.0000
s(soak,hk_bt_flt)	22.12	28	69.27	0.0000

DMP STATS The Coach House, Mount Melville House, St. Andrews, Scotland KY 168 NT, UK +44(0) 1334 477 544 mail@dmpstats.com

Residual autocorrelation under this model is highly significant (runs-test *p*-value <<0.001) and persists at low levels over long lags. Convergence times for mixed models equivalents were prohibitive for this current study. Significant results under this model should be treated with caution, as positive autocorrelation is likely to produce unduly small variance estimates.





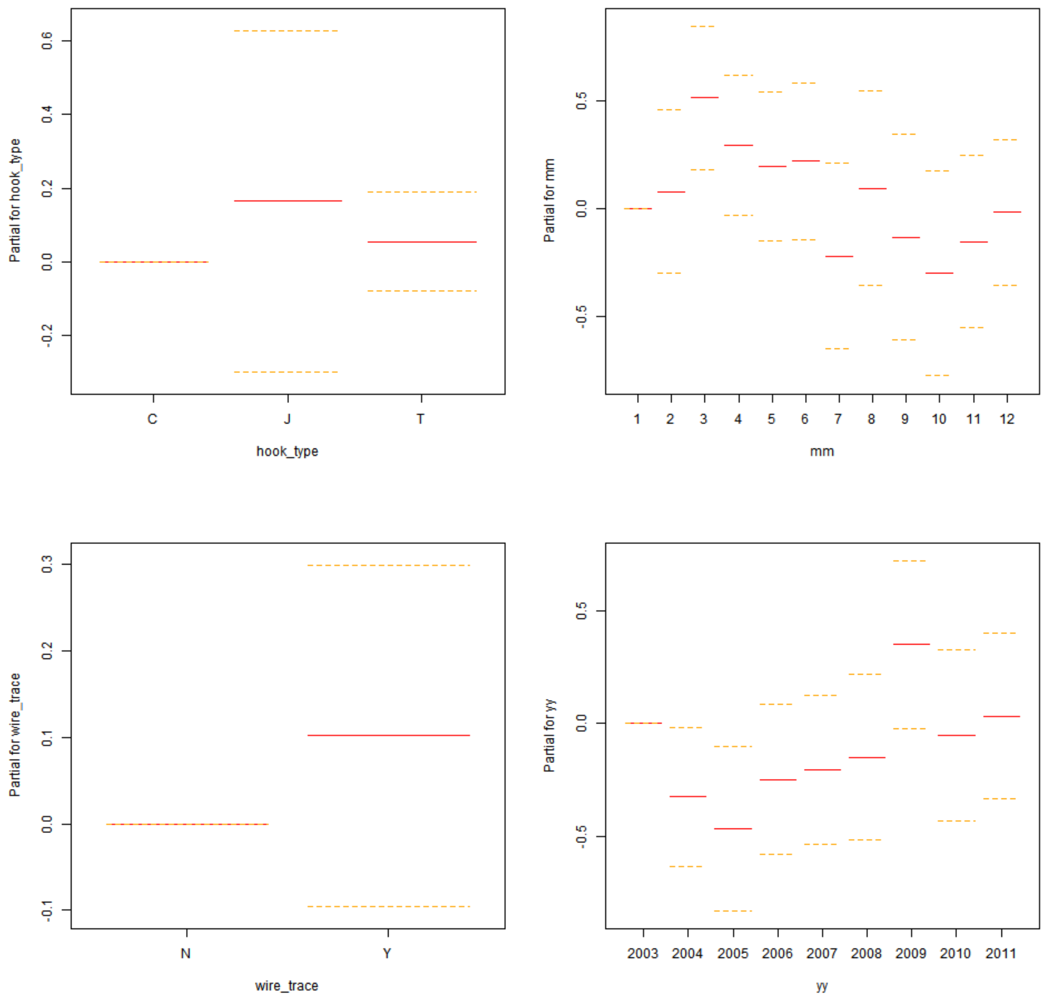
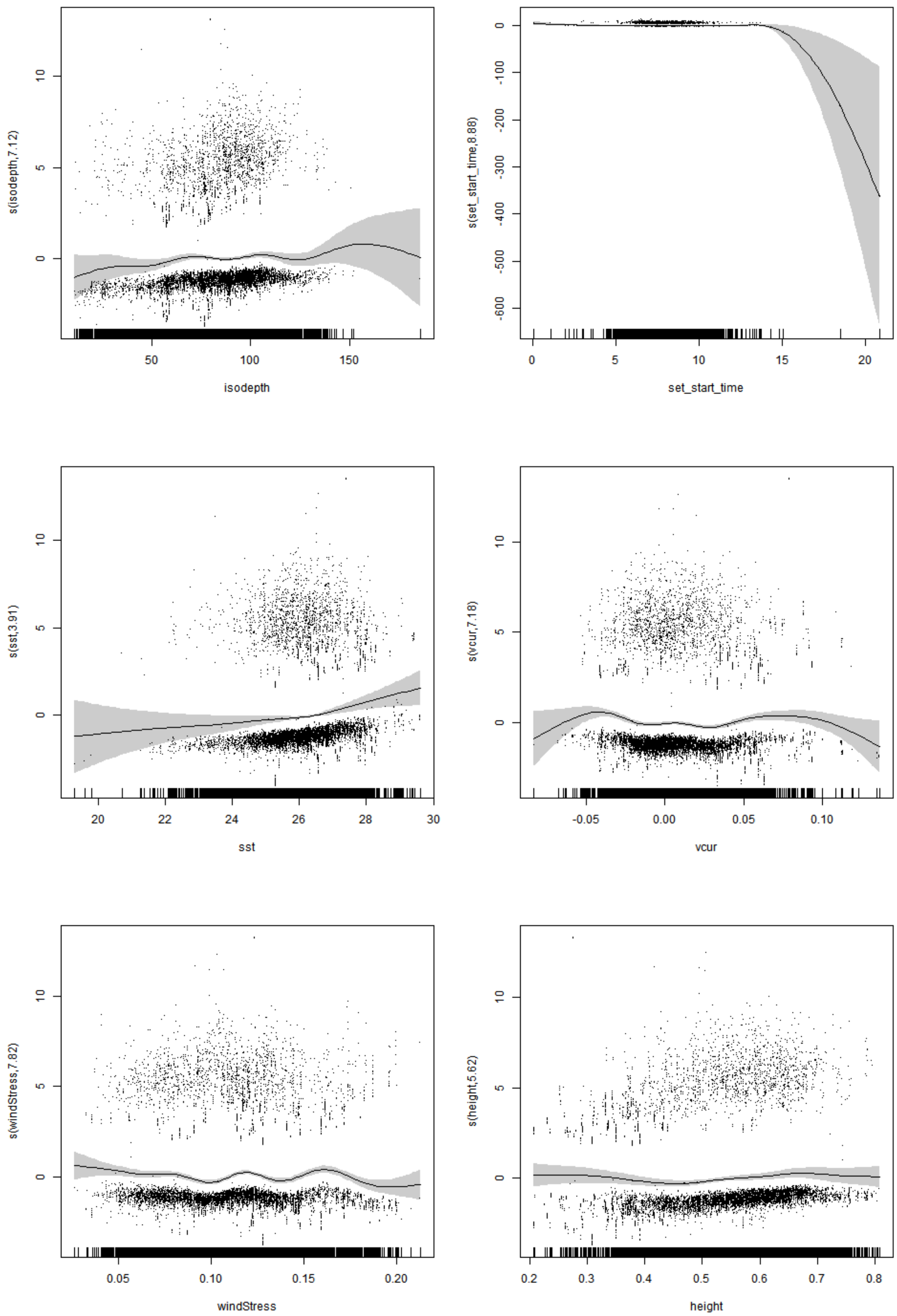


Figure 176: estimates and approximate 95% confidence intervals for parametric terms within the GAM.


s(isodepth,7.12)
isodepth
s(set_start_time,8.88)
set_start_time
s(sst,3.91)
sst
s(vcur,7.18)
vcur
s(windStress,7.82)
windStress
s(height,5.62)
height

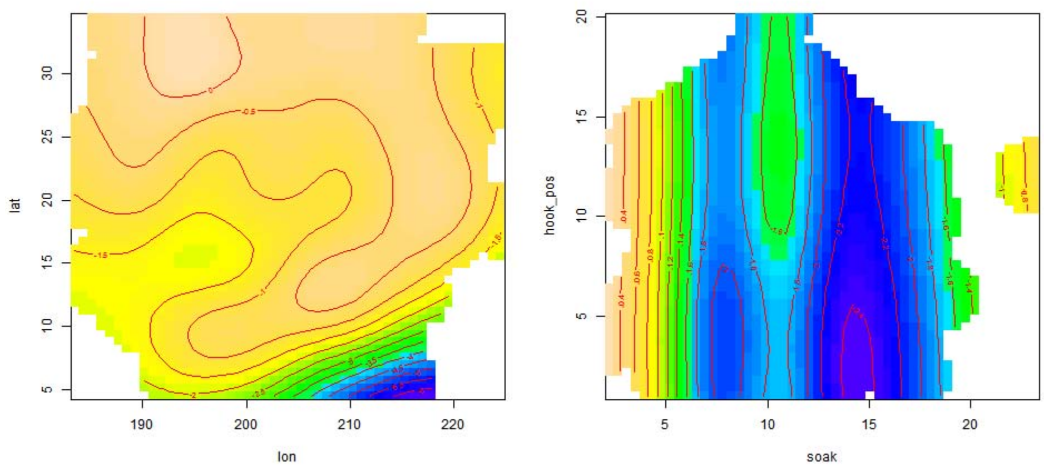


Figure 177: estimated smooth surfaces for the bivariate smooth components within the GAM.

8.4 BLUE SHARKS (BSH)

8.4.1 Catch rate of blue sharks (BSH)

8.4.1.1 Data Exploration

Catch data of blue shark in the M7_HD fishery is the only analysed study case where whithin-set absence of bycatch (in about 5000 sets) is not the predominant catch/set ratio in observed trips (Figure 178). Instead, the presence of 1 shark/set was the most common occurrence (in about 6500 sets), followed by the 2 sharks/set in just under 6000 sets. In fact, there is an approximately exponential decay in the frequency of sets with increasing numbers of sharks/set, which reaches a maximum of 180 blue sharks caught in a single set (Figure 178).

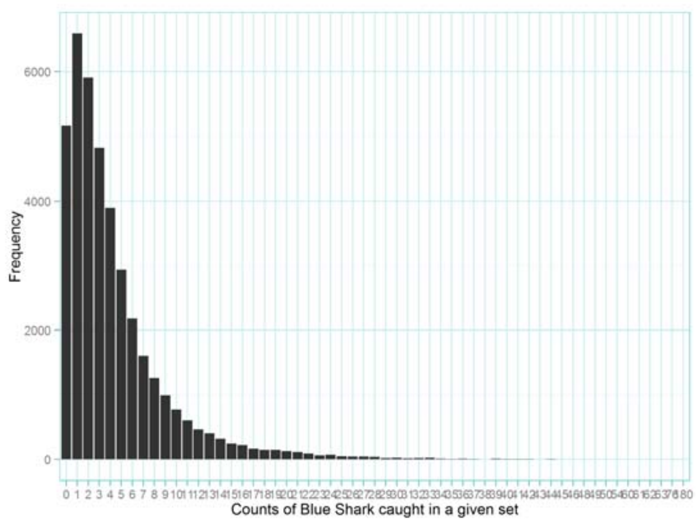


Figure 178: Distribution of counts per set of blue shark caught in observed trips of the M7_HD longline fishery (M7_HD-BSH_CatchNumFreqs.png)

Bivariate plots presented in Figure 179 suggest that, on average, higher catch rates of blue shark in the M7_HD fishery were likely to occur when, e.g.:

- Sets were performed between the months of September and December
- Wire leader and hook type "J" were in use;
- Longlines were set within 10-25 degrees latitude, or within 190-200 degrees longitude;

- Average sea surface temperature was around 27^{o}C, the average sea surface height was over 0.5 meters, or the average depth of the 20^{o}C isotherm was roughly between 60-90 meters.

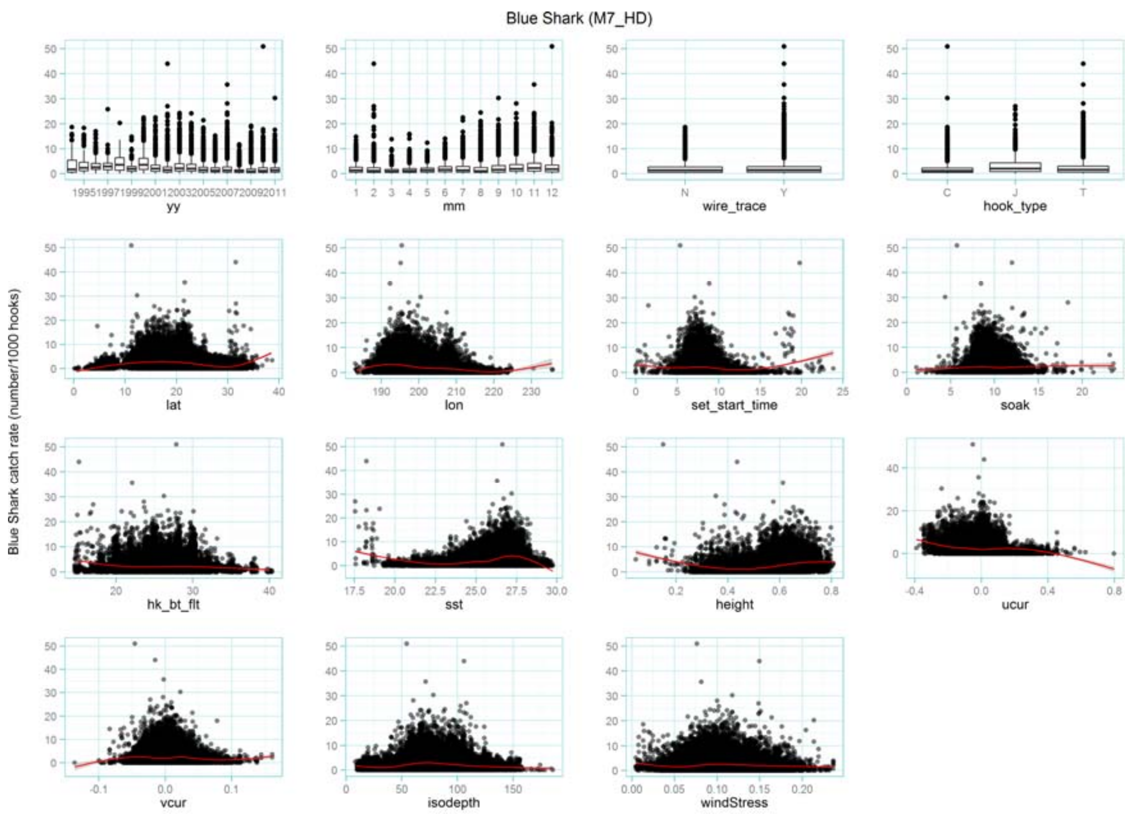


Figure 179: Distributions of set-level catch-rate responses against the explanatory variables considered for the analysis of blue sharks caught in the M7_HD fishery (file “M7_HD-BSH_CondtnVsCovs.png”)

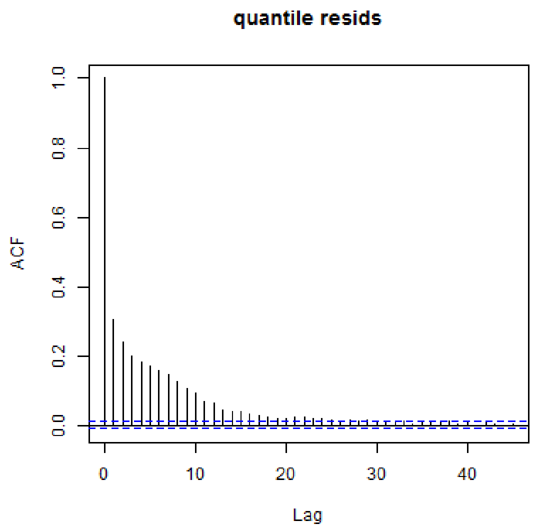
8.4.1.2 Modelling results

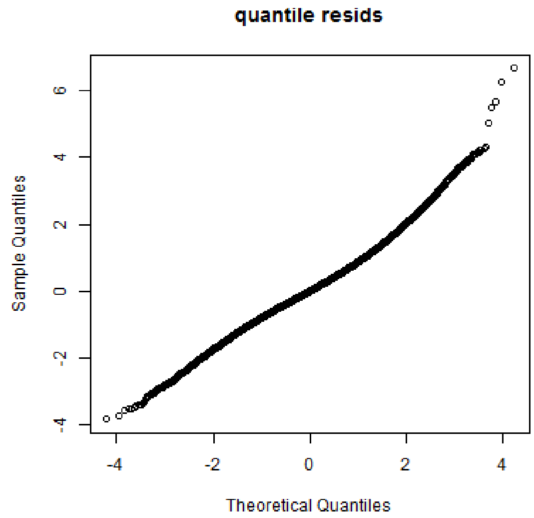
Modelling details are as per the reference case. The model returned an adjusted-R^2 of 0.40 and a deviance explained of 42.1%.

Table 51: Model terms from the generalized additive model. DF = Degrees of Freedom. EDF=Effective Degrees of Freedom. Shrinkage has been applied as a means of model selection. Terms of little relevance have their EDF shrunken.

Model terms	EDF	DF	Chi or F-stat	*p*-value
yy		17	116.66	0.0000
mm		11	656.02	0.0000
hook_type		2	23.89	0.0000
wire_trace		1	255.75	0.0000
s(lon,lat)	28.40	29	518.63	0.0000
s(soak,hk_bt_flt)	28.00	29	37.87	0.0000

Residual autocorrelation under this model is highly significant (runs-test *p*-value <<0.001) and persists at strong levels over long lags. Convergence times for mixed models equivalents were prohibitive for this current study. Significant results under this model should be treated with caution, as positive autocorrelation is likely to produce unduly small variance estimates.


quantile resids
ACF
1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
40
Lag


quantile resids
Sample Quantiles
6
4
2
0
-2
-4
-4
-2
0
2
4
Theoretical Quantiles

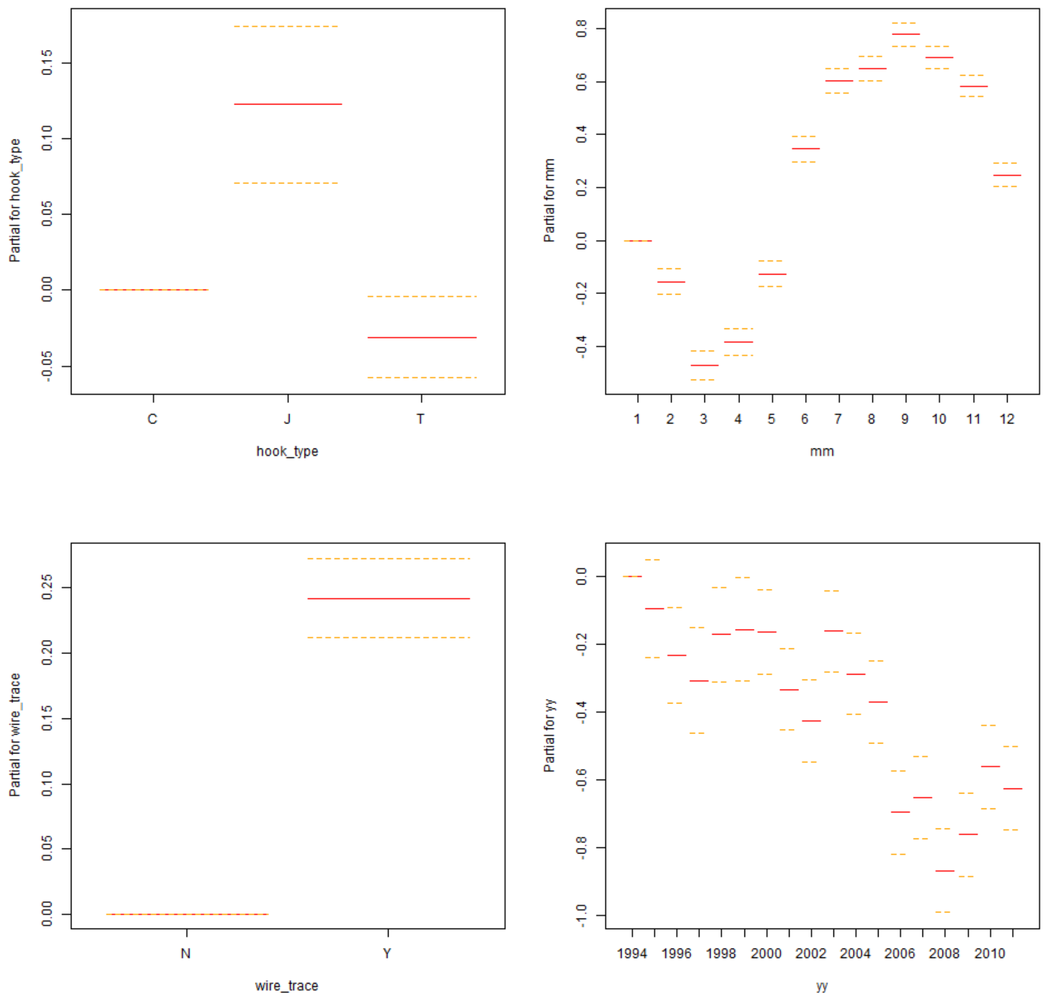


Figure 180: estimates and approximate 95% confidence intervals for parametric terms within the GAM.

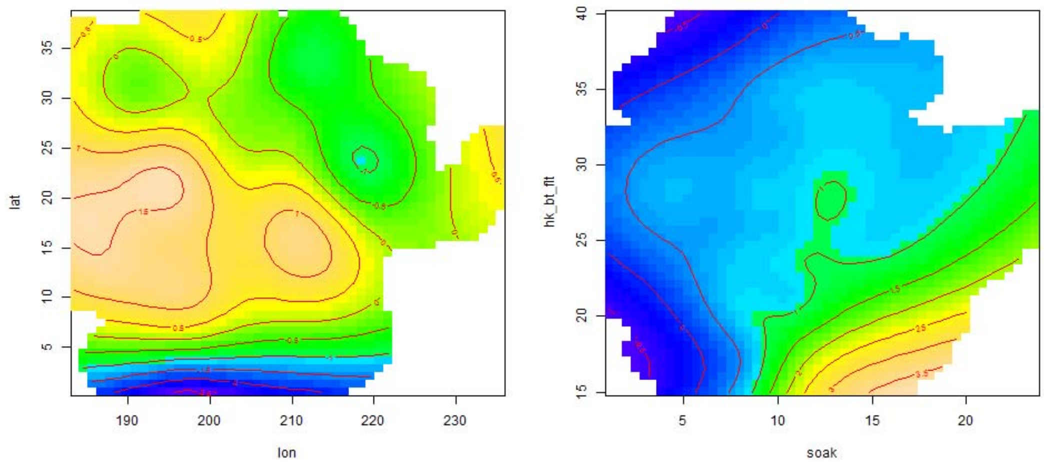


Figure 181: estimated smooth surfaces for the bivariate smooth components within the GAM.

8.4.2 Hook-level catch rate of blue sharks

8.4.2.1 Data Exploration

Figure 182 suggests that catch rates of blue sharks in the M7_HD fishery tend to be larger at higher hook position numbers (i.e. at deeper hooks). Furthermore, and despite the considerable degree of overlapping between data points, there is sparse indication for larger catch rates of this species at deeper hooks if more than 25 hooks between floats were used, or when sets were performed with average sea surface temperatures above 22°C (Figure 182).

Figure 182: Distributions of per-hook catch-rate responses against the explanatory variables considered in the analysis, for blue sharks caught in the M7_HD fishery. Dots were slightly jittered to minimise overplotting and help visualisation and their colour convey the basket hook position of the respective observation (file "M7_HD-BSH_perHookRatesVsCovs.png").

8.4.2.2 Modelling results

Modelling details are as per the reference case. The model returned an adjusted-R^2 of 0.006 and a deviance explained of 0.8%.

Residual autocorrelation under this model is highly significant (runs-test *p*-value <<0.001) and persists at strong levels over long lags. Convergence times for mixed models equivalents were prohibitive for this current study. Significant results under this model should be treated with caution, as positive autocorrelation is likely to produce unduly small variance estimates.

Table 52: Estimates of catch rates for each of the hook positions. Confidence intervals are approximate 95%. Percentage of total catch is calculated from overall estimated catch over hook positions presented. Confidence intervals for percentages are based on this naïve translation of cate rate.

Hook position	Catch rate (per 100 hooks)	SE	Lower CI	Upper CI	Catch (as % of total)	Lower CI (as % of total)	Upper CI (as % of total)
1	0.16	1.0129	0.1561	0.1641	4.35	4.24	4.46
2	0.202	1.0115	0.1975	0.2066	5.49	5.37	5.61
3	0.2335	1.0108	0.2286	0.2384	6.34	6.21	6.48
4	0.2411	1.0106	0.2362	0.2461	6.55	6.42	6.69
5	0.2422	1.0106	0.2372	0.2472	6.58	6.45	6.72
6	0.2365	1.0107	0.2316	0.2415	6.43	6.29	6.56
7	0.2319	1.0108	0.2271	0.2368	6.3	6.17	6.44
8	0.2142	1.0112	0.2096	0.219	5.82	5.69	5.95
9	0.2187	1.0111	0.214	0.2235	5.94	5.82	6.07
10	0.2104	1.0115	0.2057	0.2151	5.72	5.59	5.85
11	0.232	1.0113	0.227	0.2372	6.31	6.17	6.45
12	0.22	1.0124	0.2148	0.2254	5.98	5.84	6.12
13	0.3656	1.012	0.3572	0.3742	9.93	9.71	10.17
14	0.2895	1.0187	0.2792	0.3001	7.87	7.59	8.16
15	0.3826	1.0476	0.3493	0.419	10.4	9.49	11.39

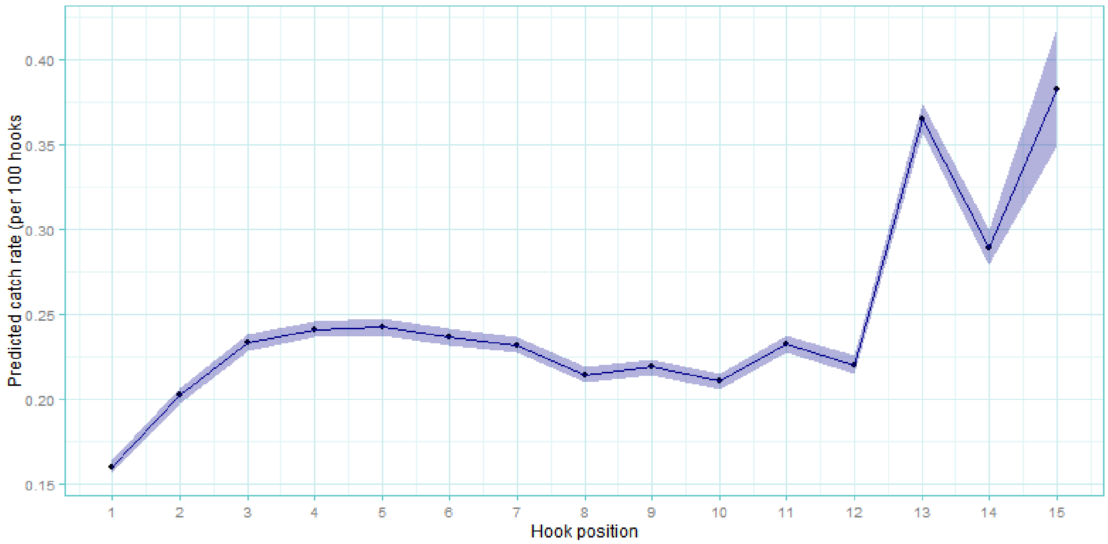


Figure 183: estimated mean catch rate (per 100 hooks) at hook-position with approximate 95% confidence envelope.

8.4.3 Condition of blue sharks at time of retrieval

8.4.3.1 Data Exploration

There are no substantial patterns between the condition of caught blue sharks at the time of retrieval and most of the covariates considered in the present analysis (Figure 184). There are however some slight indications that, on average, the proportion of dead blue sharks might have increased when observed sets were performed under certain conditions, e.g.:

- When sets were located below 12 degrees of latitude;
- If the average sea surface height remained below 0.3 meters.

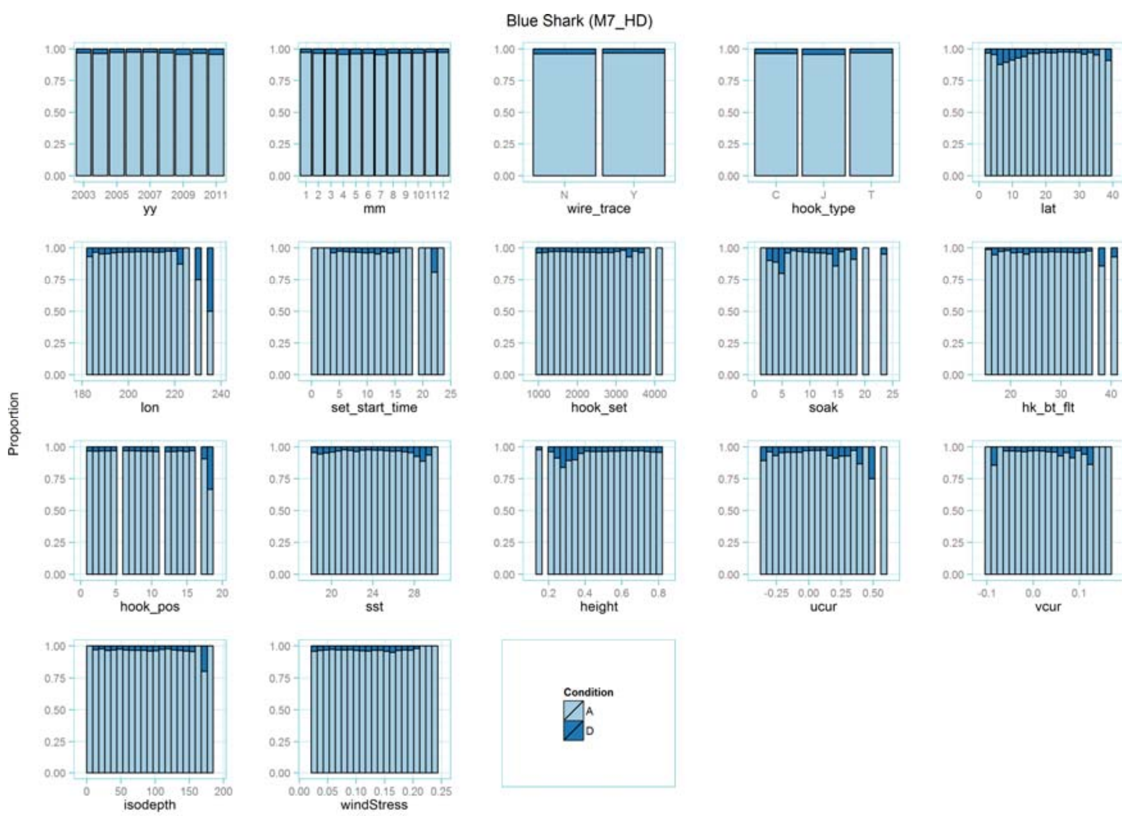


Figure 184: Distributions of condition responses (D= dead, A = alive), expressed in terms of proportion, against the explanatory variables considered for the analysis of blue sharks caught in the M7_HD fishery (file “M7_HD-BSH_CondtnVsCovs.png”)

8.4.3.2 Modelling results

Raw predictive power to the current dataset was also assessed via confusion matrices on the training data, providing Sensitivity (true positive rate) = 0.56, Specificity (true negative rate) = 0.62 and a misclassification rate of 0.38 when using the response mean as a decision boundary. These are against the training data and will be over-estimates.

Table 53: Model terms from the generalized additive model. DF = Degrees of Freedom. EDF=Effective Degrees of Freedom. Shrinkage has been applied as a means of model selection. Terms of little relevance have their EDF shrunken.

Model terms	EDF	DF	Chi or F-stat	*p*-value
yy		8	87.79	0.0000
mm		11	48.04	0.0000
hook_type		2	4.48	0.1062
wire_trace		1	13.92	0.0002
s(set_start_time)	4.48	9	30.49	0.0000
s(sst)	5.57	9	50.13	0.0000
s(height)	0.98	9	1.66	0.1474
s(ucur)	5.91	9	28.22	0.0000
s(vcur)	4.63	9	15.81	0.0022
s(isodepth)	6.07	9	36.37	0.0000
s(windStress)	0.82	9	1.92	0.0909
s(soak,hook_pos)	23.64	29	90.22	0.0000
s(lat,lon)	25.20	29	170.91	0.0000
s(soak,hk_bt_flt)	20.91	28	57.06	0.0000

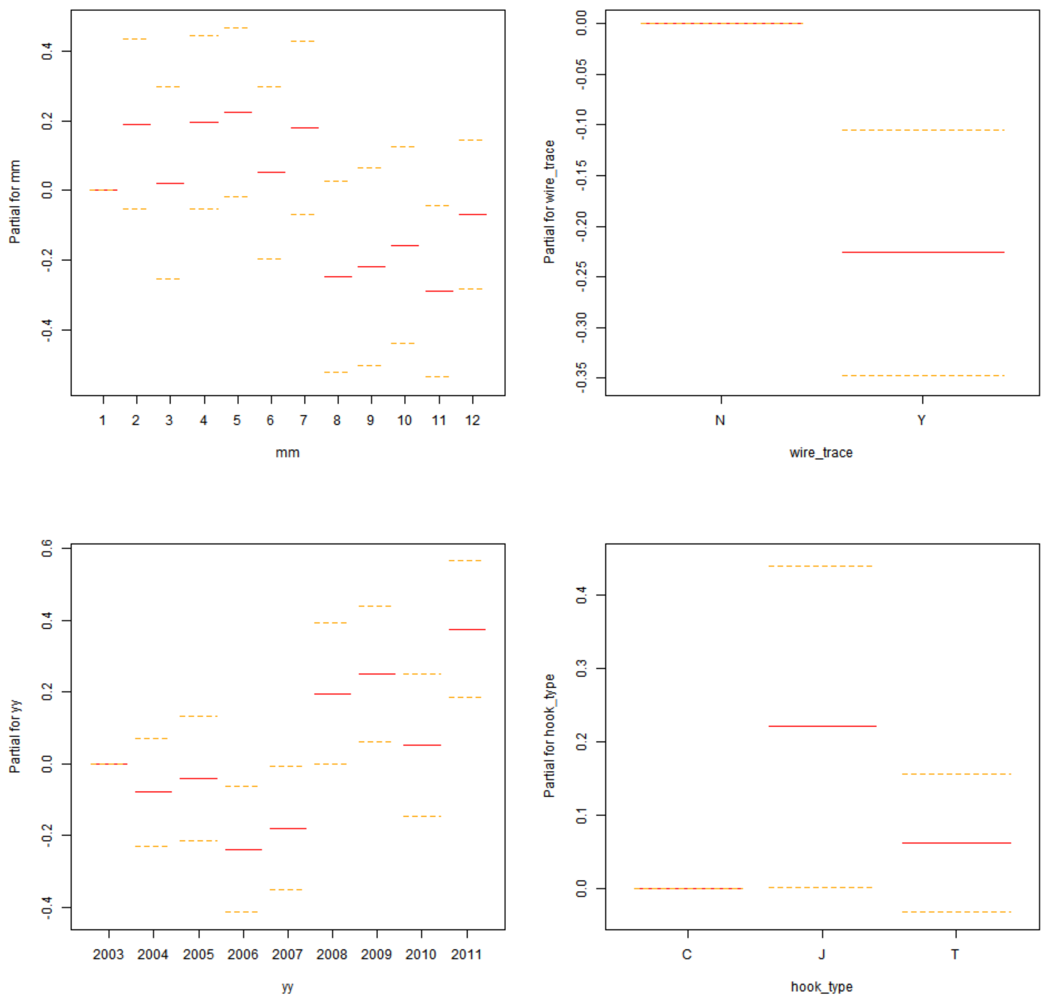


Figure 185: estimates and approximate 95% confidence intervals for parametric terms within the GAM.

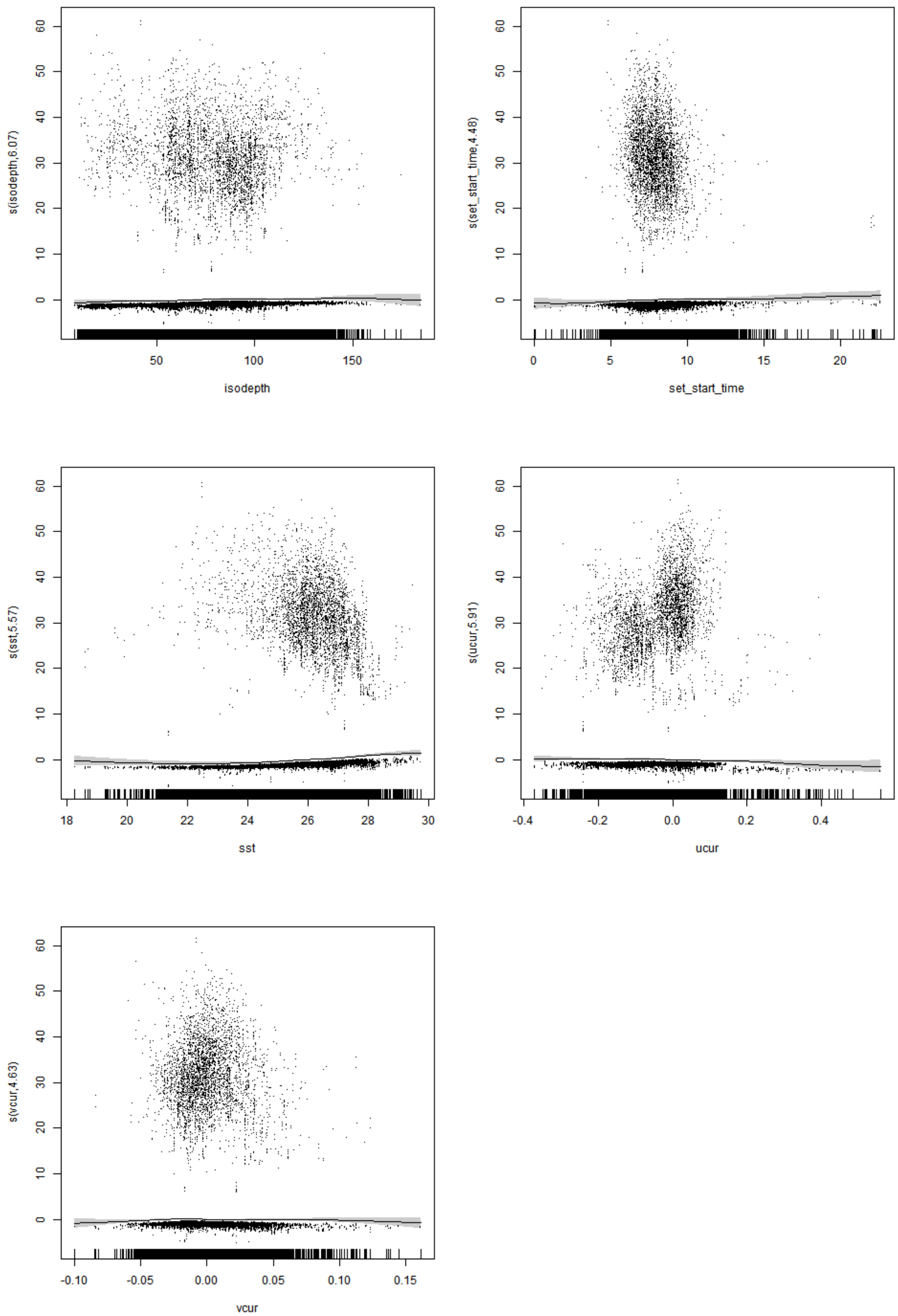


Figure 186: estimated smooths and approximate 95% confidence envelopes for smooth terms within the GAM.

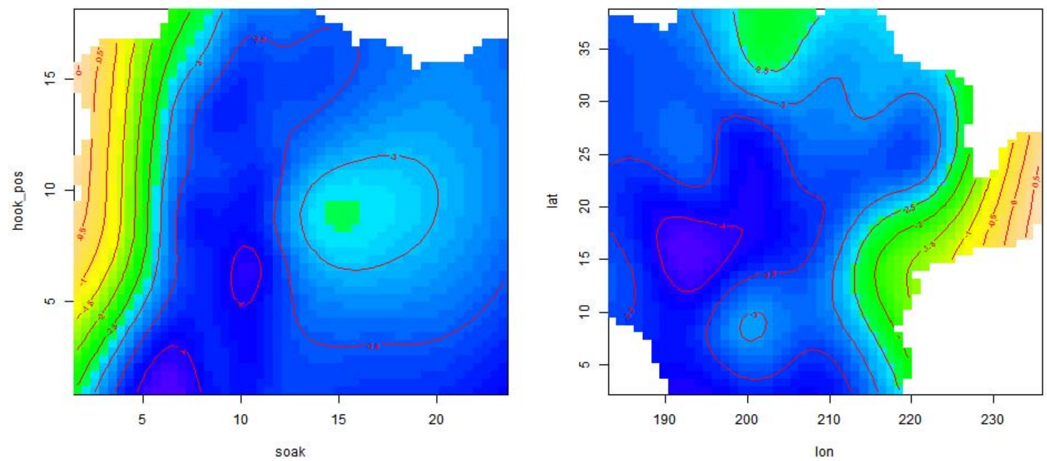


Figure 187: estimated smooth surfaces for the bivariate smooth components within the GAM.

9 REFERENCES

Bromhead, D., Rice, J. and Harley, S. (2013) *Analyses of the potential influence of four gear factors (leader type, hook type, "shark" lines and bait type) on shark catch rates in WCPO tuna longline fisheries.* SCIENTIFIC COMMITTEE NINTH REGULAR SESSION Pohnpei, Federated States of Micronesia, 6-14 August 2013WCPFC-SC9-2013/EB-WP-02 rev 1 (22 July 2013) Oceanic Fisheries Programme, Secretariat for the Pacific Community.

Dunn, K. P., and Smyth, G. K. (1996). Randomized quantile residuals. Journal of Computational and Graphical Statistics 5, 1-10.

Marra, G. and S.N. Wood (2011) Practical variable selection for generalized additive models Computational Statistics and Data Analysis 55,2372-2387

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Smyth, G. Hu, Y., Dunn, P., Phipson, B. and Chen, Y. (2013). statmod: Statistical Modeling. R package version 1.4.18. http://CRAN.R-project.org/package=statmod

Wood, S.N. (2011) Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. Journal of the Royal Statistical Society (B) 73(1):3-36

Wood, S.N. (2004) Stable and efficient multiple smoothing parameter estimation for generalized additive models.Journal of the American Statistical Association. 99:673-686.

Wood, S.N. (2006) Generalized Additive Models: An Introduction with R. Chapman and Hall/CRC.

Wood, S.N. (2003) Thin-plate regression splines. Journal of the Royal Statistical Society (B) 65(1):95-114.

Wood, S.N. (2000) Modelling and smoothing parameter estimation with multiple quadratic penalties. Journal of the Royal Statistical Society (B) 62(2):413-428.