An investigation of the effects of environmental and observational variables on haul-out counts of Harbour seals (*Phoca vitulina vitulina*) in Ireland

Maria Rakka¹ and Cóilín Minto^{1*}

March 23rd 2015

¹Marine and Freshwater Research Centre, Galway-Mayo Institute of Technology, Dublin Road, Galway, Ireland.

*Contact email: <u>coilin.minto@gmit.ie</u>







Executive summary

Purpose

Animal counts may be affected by a suite of environmental, behavioural and populationlevel influences. The success of a monitoring design will depend, among other things, on the ability to account for such influences and mitigate their effects where necessary. Here we investigated the influence of observational and environmental effects on standardized harbour seal haul-out counts from the National Parks and Wildlife Service's annual monitoring programme conducted among 16 locations over five years from 2009-2013. The aims were to: (1) understand the effects of observational and environmental variables on the counts and (2) derive standardised indices having accounted for these effects.

Procedure

At the location level, we modelled the response of haul-out counts per tidal state (landbased data) or per boat-based survey as a function of 10 principle observational and environmental variables including: day of the year, time of the day, tidal state, weather and disturbance. Poisson generalized linear/additive/mixed models were used to: (1) infer which factors had significant effects on the observed counts; (2) estimate the direction and magnitude of the effect (e.g., whether the counts increase or decrease significantly with a given variable) and (3) use this information to provide standardised counts per location.

Results

A diversity of location-specific variables was found to influence the haul-out counts. Broadly, however, day of the year, tidal state and hour of the day were the most consistently significant *observational variables* retained in the best fitting models. The influence of tidal state was somewhat consistent across sites with counts increasing up to low water and decreasing after low water; the influences of day of the year and hour of the day were more variable across sites with some sites having higher counts earlier in the season, others later and similarly so for the time of the day.

Tidal range, weather and wind direction were the three most consistently significant *environmental variables* retained in the best fitting models. Highest counts were typically observed on Spring tides; the effects of wind direction and weather being more diverse and location-dependent. Disturbance effects were retained in the models for three locations (i.e., Donegal Bay, Kinvara Bay, and Moy Estuary) where higher counts were associated with the absence of disturbance.

Effects were broadly similar across modelling techniques though the generalized linear mixed effects model was preferred owing to its accounting for visit-level correlation among the counts. Differences in the performance of the mixed effects models among survey locations indicate differing between-visit variability across locations, reflecting the uniqueness of the characteristics of each haul-out location.

1

Six of 12 land-based locations had similar trends in counts between standardised and raw mean and maxima; other cases where the trends differed reflected different environmental and observational effects accounted for within the standardized counts (e.g., Emlagh Point, Moy Estuary, and Kinvara Bay).

Recommendations

Based on the findings of this analysis, we recommend the following:

- (1) The influence of environmental and observational variables on the haul-out counts should be monitored at a frequency appropriate to the resources available. Our analyses, albeit based on a large and exceptionally valuable dataset, covers a time period of only 5 years. Inferences on the variables influencing the counts must therefore be appropriately judicious.
- (2) Out of a potential ten observational and environmental variables, the chosen model often retained considerably fewer significant variables (3 locations: 1 variable; 4 locations: 2 variables; 2 locations: 3 variables; 2 locations: 6 variables; 1 location: 7 variables). This could reflect the sample size to date but also an inherent adequacy of the standardised programme design that is also reflected in the less common effect of environmental variables.
- (3) Especially concerning boat-based surveys, the addition of new survey years will further enhance the ability of the modelling procedure to detect patterns and trends from the data with statistical confidence.
- (4) The predominance of day of the year effects reflects a design feature to cover the period of optimal haul-out behaviour. We recommend this continues. Our findings may assist in further refinement of the monitoring design at a location-level, though sites with significant visit-level variability (e.g., Ballysadare Bay, Mannin Bay, Oranmore Bay) should have the day of the year effects continuously monitored since this variability reflects considerable variance around the estimated day of the year effects.
- (5) Given the typical dome shaped effect of tidal state centred on low water it is important that the programme maintains this approach to conducting haul-out counts.
- (6) Should future modifications of the survey design occur, the interpretations of data from the modified programme should be cognisant of the design changes.

Abstract

Prevailing environmental and observational conditions during field surveys can influence the resulting counts of wild animals, including through the effects of variability in natural behaviour. Here we investigate the influence of environmental and observational covariates on standardised daytime haul-out counts of Harbour seals (Phoca vitulina vitulina) obtained during the annual moult season in Ireland between 2009 and 2013. Commencing with a dataset of 3,688 count records, the effects of ten recorded variables (including among others: day of the year, tidal state, tidal range, hour of the day, disturbance) on harbour seal counts were investigated for 12 terrestrial monitoring locations using generalized linear (GLM), additive (GAM) and mixed effects (GLMM) models. Model selection from all combinations of main effects resulted in a set of best fitting and, importantly given the sample size, most parsimonious models per survey region. Overall the GLM and GAM results were similar in terms of the explanatory variables retained which for many locations included day of the year, hour of the day, tidal state and to a lesser extent tidal range, disturbance and weather, though these latter variables may also be important for certain locations. The best fitting GLMM models were typically simpler, containing fewer explanatory variables but with significant visit-level variation estimated. The magnitude of the effect of each covariate was investigated via the proportional change in the deviance by omitting a given variable, which again highlighted the importance of day of the year, hour of the day and tidal state as highly influential variables that had a large effect on the recorded counts of harbour seals. Annual predictions from the best fitting models provided standardised alternatives to raw count summary statistics. These predictions showed similar year-driven effects to the raw statistics in six monitoring locations and differing year effects at the remaining six sites. A reduced analysis of boat-based survey data from south-western Ireland reflected the lower number of count replicates available and only included the statistical investigation of day of the year and tidal range effects on the counts. This work, which is based on extensive field data gathered under standardised conditions, provides key statistical guidance concerning ongoing monitoring methodology for harbour seals at a diverse range of sites in Ireland.

Table of Contents

Introduction	5
Materials and Methods	6
Data	6
Modelling	
By-location analysis of land-based monitoring data	10
Model selection and checking	12
Explanatory variable effects	12
By-location analysis of boat-based monitoring data	13
Standardised counts	14
Results	14
Model checking	14
Model selection for land-based count data	14
Explanatory variable effects	15
Standardised counts predicted by the modelling approach	19
By-location boat-based analyses	22
Discussion	23
Model checking	23
Model selection for land-based count data	24
Variable retention and influence	24
Standardised counts	26
By-location boat-based analyses	26
General discussion	26
Acknowledgements	27
References	27
Appendices	

Introduction

Counts of animal species may vary in response to a variety of environmental, demographic, behavioural and anthropogenic influences (Clark, 2007). Observational surveys designed to monitor the abundance of a species or a representative part thereof must therefore consider the inclusion of information on the environmental conditions under which counts are conducted. It is essential that the potential effects of such conditions are examined and better understood in order to contribute to a more robust knowledge of species biology, to discern potential trends from otherwise raw observational data and to optimise the design of monitoring programmes. In designing species monitoring programmes some conditions such as the general time of the year or the time of the day may be controlled for, whereas factors such as the weather conditions and human disturbance may be beyond observer control. When inferring appropriate population descriptors or trends from such observational data it is essential that the influence of both environmental and observational variables are accounted for (Maunder and Punt, 2004; Cosgrove *et al.*, 2014).

The Harbour seal (*Phoca vitulina vitulina*) is one of the two pinniped species commonly occurring and breeding in Ireland. Significant regional declines in the species' population trends worldwide (e.g., Northern Ireland: Wilson *et al.*, 2002; United Kingdom: Lonergan *et al.* 2007, Canada: Bowen *et al.*, 2003) have raised concerns about an emerging large-scale population decrease. In Ireland harbour seals are legally protected under the Wildlife Acts 1976 to 2012 and under the EU Habitats Directive (92/43/EEC) as a species of Community interest. Compliance with the Habitats Directive requires among other things (a) the establishment of designated Special Areas of Conservation (SAC) for the protection of the species and (b) the effective investigation and monitoring of harbour seal populations including demographic trends, habitat and distribution, and the identification of potential threats to the species.

A systematic nationwide study of harbour seal population size and distribution in Ireland reported a minimum of 2,905 animals ashore during the 2003 moult season (Cronin *et al.*, 2007) with highest occurrences along the south-west, west and north-west coasts where most of the important haul-out areas are located. Harbour seal haul-out behaviour provides opportunities for such population assessment, especially during the moulting season when the highest proportion of the population can be found ashore at regular haul-out locations (Watts, 1996; Cronin & Ó Cadhla, 2007). However the number of harbour seals available for counting within a prescribed season is influenced by a wide range of factors which may directly or indirectly affect haul-out behaviour, such as individual preferences, tidal state and weather (Watts, 1996; Huber *et al.*, 2001). Moreover, due to the variety of habitats that are used as haul-out locations by individual harbour seals the relevance of each factor may also depend on the specifics of the study area (Thompson *et al.*, 1997), a key component which cannot be overlooked when seeking to interpret acquired population count data.

Since 2009 the National Parks & Wildlife Service (NPWS) of Ireland's Department of Arts, Heritage and the Gaeltacht has been carrying out a national programme of annual sitebased monitoring for harbour seals, building on initial regionalised survey efforts (e.g., Heardman *et al.*, 2006) and contributing to national species surveillance objectives. This survey effort originally began across 14 coastal locations during the moult season that were expected to collectively represent approximately 40-50% of the national population (NPWS, 2012) and it subsequently expanded to cover 17-19 such locations where possible. The monitoring programme prioritises coverage of the most important accessible haul-out locations, many of which are situated within designated SAC sites for the species. The survey effort carried out so far has been conducted in a systematically designed framework by professional observers (see NPWS, 2012 and further details in the *Data* section of the Materials and Methods). The resulting 5-year dataset compiled to date provided a rich opportunity to investigate the influence of environmental and observational variables on harbour seal counts at a location-by-location level, further facilitating cross-location or regional comparisons.

Due to the need for an efficient and scientifically robust plan for the monitoring of harbour seal populations in Ireland and considering the seasonal and annual variability associated with local factors influencing haul-out behaviour, the aims of the current study were to:

- Determine the key environmental and observational factors influencing harbour seal counts at 16 haul-out locations monitored annually by NPWS personnel along the south-west, west and north-west coasts of Ireland;
- 2) Produce standardised by-location estimates of relative abundance having accounted for the influence of measured environmental and observational covariates on the count data.

In addition to their immediate utility for NPWS scientific and regional management, the results of these investigations will feed into simulations on the power to detect trends under a range of monitoring programme designs (Phase 2 of the current project).

Materials and Methods

At first we present an overview of the dataset, followed by an exposition of the suite of analyses conducted per monitoring location.

Data

Response (count)

The dataset used in the statistical analysis consisted of replicated standardised counts of harbour seals gathered from 16 regional locations (12 surveyed using land-based methods, 4 using boat-based methods) over the period 2009-2013 with the exception of Gweebarra Bay, which was added to the programme in 2013. Given that most survey locations or bays

contain numerous spatially-referenced haul-out sites, the entire dataset put forward for the analysis numbered 3,688 count records.

For land-based monitoring locations, a count record consisted of the total number of animals (adult and juveniles recorded in and out of the water) observed at the monitoring location during a given state of the tide. As indicated above the data were recorded at a fine spatial resolution, i.e., per geo-located haul-out site which is typically a rocky skerry, intertidal sandbank, reef or shoal. Count records were then summed across individual haul-out sites to provide an overall count of harbour seals per tidal state for the monitoring location. All visible animals were thus counted at hourly intervals starting at two hours before the local time of low tide (Low Water) and finishing two hours after Low Water; this ideally resulted in 5 replicate counts per location visit (example shown in Figure 1). A given land-based monitoring location was typically visited three times per annual moult season, providing 15 tidal-state counts per year (Table 1; the independence of these observations is considered in the *Modelling* section below).

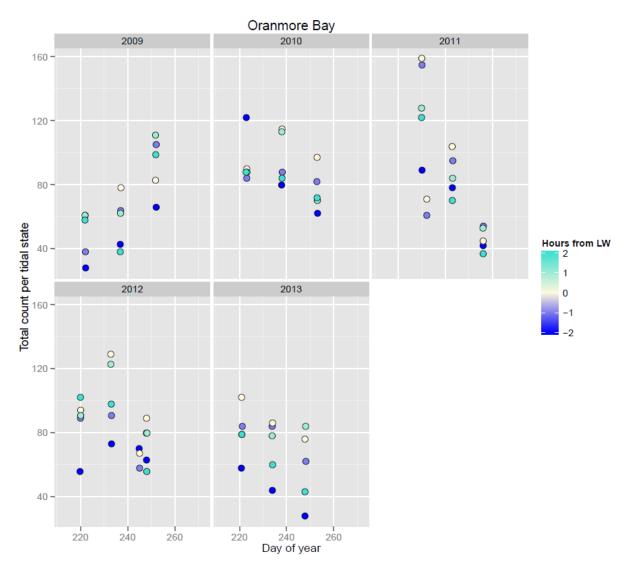


Figure 1. Example of haul-out counts per tidal state (hours from low water) for a given location (Oranmore Bay) over the survey time period (similar plots for all locations are provided in Appendix 1).

Table 1. The number of harbour seal counts conducted per monitoring location per year. Land-
based counts comprise total summed records per tidal state (i.e., a value of 15 denotes three
location visits with five counts per visit); boat-based counts comprise total summed records per
survey (i.e., across tidal states).

Location	2009	2010	2011	2012	2013	
						Total
Land-based monitoring locations						
Adrigole Harbour	6	1	3	3	2	15
An Baile Láir, Inverin, Loughaunbeg	10	8	15	15	15	63
Ballysadare Bay	13	15	15	15	15	73
Cashla Bay	10	10	10	15	15	60
Donegal Bay	15	14	18	15	16	78
Emlagh Point, Roonagh, Louisburgh	15	5	8	7	13	48
Gweebarra Bay	-	-	-	-	15	15
Kinvara Bay	15	15	19	15	15	79
Mannin Bay	10	9	15	15	15	64
Moy Estuary	15	14	16	15	15	75
Oranmore Bay	15	15	17	18	15	80
Westport Bay, Clew Bay	15	15	14	15	15	74
Boat-based monitoring locations						
Bantry Bay (Inner)	3	3	3	3	4	16
Dunmanus Bay	1	1	1	2	1	6
Kenmare River	2	3	3	1	3	12
Roaringwater Bay	3	2	3	2	2	12

In addition to the land-based monitoring, boat-based surveys were performed at a selection of key locations for the species in the south-west of Ireland (Table 1), two of which have been the subject of regular NPWS survey effort since 1985 (inner Bantry Bay and Kenmare River; Heardman *et al.*, 2006). This alternative approach was required because of the inaccessibility and poor visibility of many local haul-out sites from land-based vantage points. Given the larger bay areas to be covered in such instances the boat-based surveys took place across a range of tidal states but centred around the local time of Low Water. Consequently a single summed count across all haul-out sites within the location formed the response per survey. Note that for the Kenmare River location, land-based counts from the outer north shore (Illaunsillagh, West Cove) and south shore (Illaunameanla, Ballycrovane Harbour) were excluded from the boat-based analyses and, given their sporadic recording over tidal states and years, were not analysed further in the current study.

Explanatory variables

Ten principal explanatory variables (both environmental and observational) accompanied each land-based and boat-based count of harbour seals (Table 2). While tidal-state-level recorded variables were included in the boat-based dataset, since the individual haul-out counts were summed to the survey visit level (i.e., a single location total per survey day) tidal-state-level variables were not included in the corresponding data analysis.

Variable	Raw values/levels	Model treatment (if different)
Date	Date format	Year (categorical) Day-of-year centred on day 238 (August 26 th) (continuous)
Time of count	Decimal hour of day (continuous)	
Tidal state	LW±{0, 1, 2, 3}	Hours from Low Water (continuous)
Tidal range	Spring tides, neap tides (categorical)	
Weather	Cloudy, partial cloud, rain, showers, sunny (categorical)	
Wind direction	N, NW, W, SW, S, SE, E, NE (categorical)	
Wind force	Beaufort scale (discrete)	Mean wind speed (km.hr ⁻¹) (continuous)
Precipitation type	None, fog/mist, rain (categorical)	
Precipitation intensity	Intermittent light, intermittent heavy, continuous light, continuous heavy (categorical)	
Disturbance	None, aircraft, dinghy, fishing boat, walker, other (categorical)	Present or Absent (categorical binary)

Table 2. Explanatory environmental and observational variable treatment in modelling the effectson harbour seal counts.

Most explanatory variables were included directly in the analysis as they appeared in the database but a small selection, typically consisting of multilevel categorical variables, were simplified. For example, wind force (Beaufort scale) was converted to mid-scale wind speed (km.hr⁻¹) in order to reduce the number of parameters to be estimated during the modelling process. Similarly, the Disturbance variable had been recorded according to six classes observed in the field by NPWS personnel (Table 2) and then further qualified in text/comment form. But for a given haul-out site observed disturbances may have only been recorded once in the entire five-year time series. Including the variable in such a mode in the model would have resulted in the accompanying parameter being estimated from a single observation. To mitigate such eventualities, the Disturbance variable was collapsed into a Disturbance 'presence/absence' variable (Table 2). In addition, since most of the categories in the 'Precipitation type' and 'Precipitation intensity' variables were already captured by the 'Weather' variable, the former descriptors were excluded from the analysis.

Appendix 1 contains by-location plots of all variables used in the analyses.

Modelling

Rationale

The goal of the modelling process was to investigate the effects of key observational and environmental variables (Table 2) on the haul-out counts of harbour seals such that they may be accounted for (a) in standardising the raw counts and (b) in ongoing survey design where the variables can be controlled for. The underlying idea was to derive models that account for the observed variability in the raw counts recorded by field personnel. As a simplified example, consider a five-year set of single counts:

$$y_{2009} = 100, y_{2010} = 100, y_{2011} = 50, y_{2012} = 100, y_{2013} = 100$$
,

where the counts were conducted on spring tides except in 2011 where the count was obtained on a neap tide. If we knew that the counts are halved on neap tides, all else being equal, a standardised count (expected for a spring tide) for 2011 would be 100 and we would conclude that the counts appear constant across time. Of course the NPWS harbour seal monitoring dataset consists of far more observations per annum (Table 1) with a range of accompanying environmental covariates (Table 2) but the rationale behind the analysis is essentially the same: to standardise the observational counts to those expected under fixed conditions.

By-location analysis of land-based monitoring data

Generalized Linear Models

As the response data are raw counts, a natural first choice for the distribution of the data is the Poisson distribution (Zuur *et al.*, 2009). The Poisson distribution is governed by a rate parameter - here the average number of animals per tidal state - which is allowed to vary as a function of the environmental and observational variables in a Poisson generalized linear model (GLM) (McCullagh and Nelder, 1989). The response variable is count per tidal state y_i and the Poisson GLM is given by:

Poisson distribution:	$y_i \sim \text{Pois}(\lambda_i)$,	
Linear predictor:	$\eta_i = X_i \beta,$	(1)
Log-link:	$\ln \lambda_i = \eta_i,$	

where: λ_i is the rate parameter (mean count per tidal state) of the Poisson distribution, which varies as a function of the covariates X_i through parameter vector β (estimated during fitting). The linear predictor η_i is related to the rate via the log-link. Both categorical and continuous variables (Table 2) are included in X_i . In addition to the variables included in Table 2, quadratic effects of day-of-the-year, hours from Low Water and hour-of-the-day were included as potential explanatory variables. As we were also interested in standardising the counts to a yearly value, categorical year effects were included in all models (Maunder & Punt, 2004).

Given the large number of potential covariates (Table 2) and comparatively fewer observations per location per year (Table 1) a key component to the success of the modelling was to balance the number of potential explanatory variables included in the analysis with the number of observations available for statistical testing. It is possible that interactions could exist between the variables (e.g., the influence of certain weather conditions differing by tidal state or by day-of-the-year). Despite that, given the relatively small number of observations per survey year and the standardised methodological approach to controllable variables inherent in the monitoring programme design, interactions among the variables were not considered within the current analysis. The models therefore consisted of an examination of the main effects of the explanatory variables in X_i . Changes in variable effects by year (year interactions), particularly those pertaining to day-of-the-year (DOY), are further discussed in the model selection section of the Discussion.

Generalized additive models

To allow for non-linear influence of the continuous explanatory variables over that of the quadratic variables in the GLM, Poisson-distributed generalized additive models (GAMs) (Hastie and Tibshirani, 1990; Wood, 2006) including smoother terms on all continuous variables were also estimated. The maximum degrees of freedom per smoother (which controls the model flexibility) was set as the maximum number of observations (e.g., 5 for tidal state).

Generalized linear mixed effects models

The five land-based counts of harbour seals recorded per survey visit (one per tidal state) are non-independent. Therefore, factors other than those observed may also influence the counts obtained on a given visit. For example, variability in seal haul-out behaviour or periods of severe weather might heavily influence the use of more exposed and/or more sheltered haul-out sites within comparatively discrete survey locations. To account for this correlation structure we also fitted generalized linear mixed effects models (GLMMs) (Breslow and Clayton, 1993) to the dataset, where the linear predictor (1) is extended to:

$$\eta_i = X_i \beta + b_{\nu_i} \tag{2}$$

where v_i is the visit ID of the *i*th observation and the visit-level random effects are assumed to be normally distributed with mean zero and variance σ_b^2 to be estimated within the

model $b_{v_i} \sim N(0, \sigma_b^2)$. Random effects here reflect additional visit-level variability not captured by the directly observed covariates included in the model (Table 2).

Model selection and checking

Including quadratic terms, the number of potential explanatory variables for many of the land-based monitoring locations was 12. Including or leaving out each of these variables resulted in 2¹²= 4,096 potential models per location; with due consideration of each of the three modelling methods (GLM, GAM, GLMM) this resulted in over 12,000 potential model fits per location. Model selection was conducted by first identifying a 'global model' (Table 3), i.e., the model containing the maximum number of variables for each location. All potential valid sub-models were then fitted in order to create a set of candidate models. The *best model* for each location was then selected by using the corrected Akaike Information Criterion (AICc) rather than its more general (AIC) form (Burnham and Anderson, 2002); the AICc provides a measure of the quality of the model fit to the provided dataset but also accounts for the relatively small number of observations per location. Simpler sub-models of the *best model* and the *best model*) were identified and chosen as the *overall best fitting model* - which we defined as the most parsimonious, best-performing model on which we could reliably base our inferences.

The basic assumptions of generalized linear models, including the appropriateness of the distributional assumption and a linear relationship between the response and the linear predictor, were tested by investigating a set of diagnostic plots for each model (i.e., residuals versus fitted, quantile-quantile plot, scale versus location, predicted versus observed plots). Outliers and overly influential points were identified through Cook's distance plots (Cook and Weisberg, 1982).

Explanatory variable effects

The importance of the covariate effects for each location was assessed in two ways: first by investigating the estimated terms included in the parameter vector β (Equation 1) of the overall best fitting model. The parameters were plotted across locations, including their 95% confidence intervals, in order to compare the relevant environmental factors influencing the counts at each monitoring location. Retention of a given covariate in the best fitting model demonstrated the statistical significance of its inclusion. It is equally important to investigate the comparative influence of the retained variables on the model's predicted counts. For instance, a small tidal range effect and large day-of-the-year effect might have similar significance in the model output but a change in the latter would be expected to have a greater effect on the count predictions. In order to take such effects into account, the proportional increase in the deviance per degree of freedom (Abeare, 2009) was used as a measure of relative influence of the variables in each model. We conducted this analysis

for the GLM and GLMM models only since the estimated degrees of freedom of the smooth terms of the GAMs often increase on omission of other explanatory variables.

	Adrigole Harbour	An Baile Láir	Ballysadare Bay	Cashla Bay	Donegal Bay	Emlagh Point	Gweebarra Bay	Kinvara Bay
Year	~ ~ ~	~ ~ ~	v v v	v v v	• • •	~ ~ ~		v v v
Hours from Low Water	~ ~ ~	• • •		• • •	• • •	• • •	~ ~ ~	~ ~ ~
Hours from Low Water ²		v v	v v	v v	v v	~ ~	~ ~	v v
Day of the year	~ ~ ~	• • •	• • •	• • •	• • •	• • •	~ ~	~ ~ ~
Day of the year ²		~ ~	v v	v v	v v	~ ~		~ ~
Hour of the day	~ ~ ~	• • •	~ ~ ~	• • •	• • •	• • •	~ ~ ~	• • •
Hour of the day ²		~ ~	v v	v v	~ ~	~	~ ~	v v
Tide		• • •	~ ~ ~	• • •		• • •	~ ~ ~	v v v
Disturbance Presence	~ ~ ~	• • •	~ ~ ~	• • •	• • •	• • •		v v v
Weather	~ ~ ~	• • •	~ ~ ~	• • •	• • •	• • •	~ ~ ~	v v v
Wind Speed	~ ~ ~	• • •	~ ~ ~	• • •	• • •	• • •	~ ~ ~	v v v
Wind Direction		• • •	• • •	• • •	• • •	• • •		
Precipitation Type								
Precipitation Intensity								
Observer Number								
Visit	~	~	 	~	~	~	 	~
	Mannin	Моу	Oranmore	Westport	Bantry	Kenmare	Dunmanus	Roaringwate
	Bay	Estuary	Вау	Вау	Вау	River	Вау	Вау
Year	~ ~ ~	• • •	 <td>• • •</td><td> </td><td> </td><td> </td><td> </td>	• • •	 	 	 	
Hours from Low Water	• • •	• • •	 <td>• • •</td><td></td><td></td><td></td><td></td>	• • •				
Hours from Low Water ²	~ ~	~ ~	 ✓ 	v v				
Day of the year	~ ~ ~	• • •	v v v	• • •	~ ~	~ ~		 ✓
Day of the year ²	~ ~	v v	v v	v v	v v	~ ~		 ✓
Hour of the day	~ ~ ~	• • •	• • •	• • •				
Hour of the day ²	~ ~	~ ~	v v	v v				
Tide		• • •	~ ~ ~	• • •		~ ~		~ ~
Disturbance Presence		• • •	~ ~ ~	• • •				
Weather	~ ~ ~	• • •	~ ~ ~	• • •				
Wind Speed	~ ~ ~	• • •	• • •	• • •				
Wind Direction	~ ~ ~	• • •	• • •	• • •				
Precipitation Type								
Precipitation Intensity								
Observer Number								
Visit	~	~	~	~				

Table 3. Global model variable inclusion by harbour seal count location and model type (GLM: blacktick, GAM: red tick, GLMM: green tick).

By-location analysis of boat-based monitoring data

The boat-based dataset was restricted to a maximum of 16 counts per location over the five-year sampling period and generally no more than 2-3 survey visits per year (Table 2). Based on the land-based data analyses we chose to include only year, day-of-the-year and tidal range as potential explanatory variables in the analysis of boat-based survey data. No GLMMs were fitted for the boat-based areas since, unlike the land-based monitoring data, individual location counts were not replicated within a survey visit.

Standardised counts

Following verification of the modelling assumptions and estimation of relevant parameters, each overall best fitting model was used to produce a predicted standardised count per location per year for a representative set of conditions, based on the mean of continuous variables and the most represented categorical variable values for a given location.

Results

Model checking

Some non-conformity with the Poisson distributional assumption, as shown by departures in the quantile-quantile plots and estimated dispersion parameter, was observed (Appendix 2, e.g., Donegal Bay and Oranmore Bay). Three locations (excluding Dunmanus Bay, where there is only one boat-based survey per year – the second count in 2012 (Table 1) was a land-based count and was thus excluded) displayed significantly underdispersed count data (i.e., variance of counts was less than the mean of counts in the Poisson), while counts for eight locations were overdispersed (i.e., variance of counts was greater than the mean of counts in the Poisson). This is demonstrated in Appendix 2 titles for GLMs and GAMs with the significance of dispersion parameter placed in parentheses.

The diagnostics were typically improved upon in GLMM fits, particularly for Gweebarra Bay where only a single year's count data were available. Persistent (across GLM, GAM and GLMM models) residual outliers (i.e., residual values less than -2 or greater than +2) did occur in some locations (Ballysadare Bay - low residuals, Kinvara - high and low, Mannin Bay - low, Moy Estuary - low, Westport Bay - high, Donegal Bay - low and high). Robust versions of the models could be fitted to this subset of locations in order to investigate the influence of these residuals on the estimated parameters.

Model selection for land-based count data

A diversity of environmental and observational variables were retained in the modelling process for the overall best fitting models (Table 4). As judged by the minimum AICc the GLM, GAM and GLMM models were best fitted across a total of 4, 4 and 4 locations, respectively (Table 5). GLMM models that fitted the best were those that had a greater amount of random effect variability (i.e., visit-level variation).

Table 4. Environmental and observational variables retained in the overall best fitting models percount location and model type (GLM: black tick, GAM: red tick, GLMM: green tick).

		_	ole our	A	n Ba Láir		Bal	lysac Bay		Cas	shla	Bay	D	oneg Bay	-	1	mlaį Poin		Gv	veeb Bay		Kin	vara l	Bay
Year	~	~	<i>v</i>	-	~	~	-	<i>v</i>	~	-	~	~	~	~	~	-	~	~		541		~	~	~
Hours from Low Water							-	~				~	-		~	~		~				~	~	~
Hours from Low Water ²													-		~	-		~				-		~
Day of the year			~												2			2					2	~
Day of the year ²				Ĵ		5							5											5
Hour of the day							Ĵ			Ĵ	5		Ĵ	5					5		~			
Hour of the day ²							Ĵ		Ĵ					-										
Hour of the day							1		Ť															
Tide				~	~	~	~	~		~	~								~	~		~	~	~
Disturbance Presence									~			~	-	~								-	~	~
Weather							-	~	~				-		~							~	~	~
Wind Speed							-	~												~		~	~	~
Wind Direction							-	~	~	~	~					-	~	~				~	~	~
Precipitation Type																								
Precipitation Intensity																								
Observer Number																								
Visit			~			~			~			~			~			~			~			~
	м	lanı	nin		Mo	/	Or	anm	ore	w	estp	ort	Bar	ntrv	Bay	Du	nma	nus	К	enm	are	Roa	ringw	ater
		Bay	/	E	stua	ry		Вау			Вау		Dai	iu y	Duy		Вау			Rive	r		Вау	
Year	~	~	~	~	~	~	~	~	~	~	~	~	~	~	~	~	~		~	~		~	×	
Hours from Low Water	~	~	~	~	~	~	~	~	~	~	~	~												
Hours from Low Water ²	~		~	~		~	~		~	~		~												
Day of the year	~	~		~	~	~				~	~	~	~	~										
Day of the year ²	~									~		~	~											
Hour of the day	~	~		~	~	~	-	~	~															
Hour of the day ²				~		~																		
Tide	~									~	~	~												
Disturbance Presence	~			~	~	~	-	~																
Weather				~	~	~	-	~		~	~	~												
Wind Speed	~									~	~	~												
Wind Direction	~	~		~	~	~	~	~	~	~	~	~												
Precipitation Type																								
Precipitation Intensity																								
Observer Number																								
Visit			~			~			~			~												

Explanatory variable effects

On examination of the deviance of the best fit models upon inclusion of given variables, the relative influence of the environmental and observational variables was shown to differ quantitatively and qualitatively by survey location (Figures 2 and 3).

For many GLM and GLMM model fits, at least two covariates were seen to be of comparatively greater influence on the raw count data obtained in the field. A **day-of-the-year** effect (linear or linear plus quadratic) was the most influential factor in 5 GLMs (An

Báile Láir, Kinvara Bay, Mannin Bay, Westport Bay, Adrigole Harbour) and 4 GLMMs (An Báile Láir, Emlagh Point, Westport Bay, Adrigole Harbour). **Hours from Low Water** was the

Location	GLM	GAM	GLMM
An Baile Láir	369.9	369.5	372.6
Ballysadare Bay	821.8	819.9	791.0
Cashla Bay	433.4	432.6	436.5
Emlagh Point	297.0	298.3	301.2
Kinvara Bay	729.8	726.8	733.4
Mannin Bay	484.3	481.4	471.9
Moy Estuary	688.5	693.5	692.3
Oranmore Bay	824.7	834.2	688.8
Westport Bay	595.1	595.7	598.9
Adrigole Harbour	99.7	99.7	106.2
Gweebarra Bay	131.4	132.6	98.9
Donegal Bay	1078.1	974.6	980.3

Table 5. AICc values for the best fitting models per model category per land-based count location.The best fitting model in each case (i.e., lowest AICc value) is highlighted via grey shading.

most influential covariate in 2 GLMs (Oranmore Bay, Donegal Bay) and 4 GLMMs (Cashla Bay, Kinvara Bay, Mannin Bay, Donegal Bay). **Hour-of-the-day** (linear or quadratic) was the most influential variable in 2 GLMs (Moy Estuary, Gweebarra Bay) and 3 GLMMs (Ballysadare Bay, Oranmore Bay, Gweebarra Bay). **Survey year** was the most influential covariate in GLMs for Cashla Bay and Emlagh Point while **tidal range** was the most influential covariate in the Ballysadare Bay GLM. **Disturbance** was the most influential covariate in the Moy Estuary GLMM.

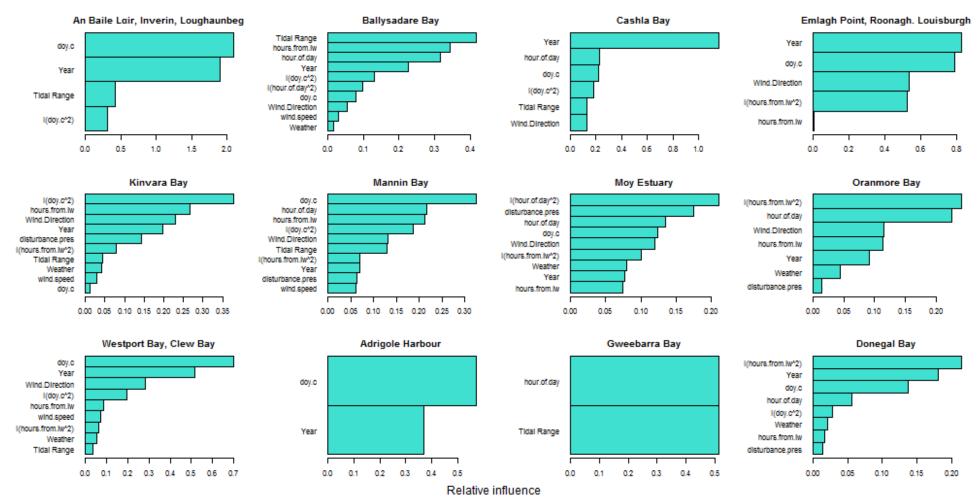


Figure 2: Relative influence of the selected environmental and observational variables on the deviance of the best **GLM** model for each land based count location. Care should be taken in interpretation and comparing between count locations since the scales of computed relative influence (x axes) shown are not consistent. *doy.c* denotes day of the year centred on day 238.7

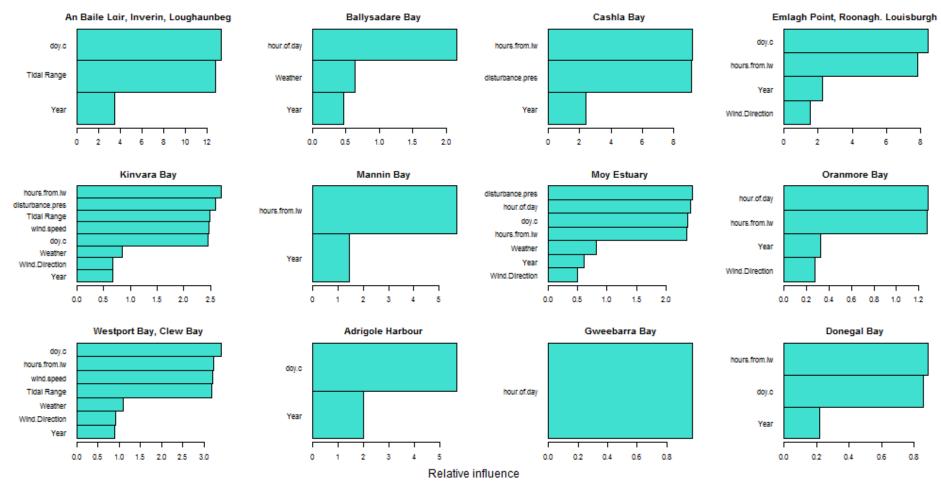


Figure 3: Relative influence of the selected environmental and observational variables on the deviance of the best **GLMM** model for each land based count location. Care should be taken in interpretation and comparing between count locations since the scales of computed relative influence (x axes) shown are not consistent. *doy.c* denotes day of the year centred on day 238.

The effects of the tested variables on harbour seal counts were similar across the three model methods (Appendix 3). In most of the land-based count locations, harbour seal counts appeared lower in surveys performed later during the moult season (beginning-mid September) with the exception of Kinvara Bay and Oranmore Bay where counts peaked in the end of August and Adrigole Harbour where counts were higher towards the end of the survey period (Figure 4). In addition, counts in most of the survey locations displayed similar patterns with respect to tidal state (hours from low water) with maximum counts being observed closer to low water (Appendix 3). Counts were also maximised during spring tides in most of the locations, except for An Baile Láir where maximum counts were recorded during neap tides (Figure 4).

	Best model	Hours from low water	Day of the year	Hour of the day	Wind speed	Tidal range	Disturbance	Weather	Wind direction		
Adrigole Harbour	GLM/GAM		~							1	
An Baile Láir	GAM		× .			Neap					
Ballysadare Bay	GLMM			1				Cloudy			
Cashla Bay	GAM			1		Spring			SW	~	-
Donegal Bay	GAM	~	N	1			None				
Emlagh Point	GLM	~	- N						SE	5	
Gweebarra Bay	GLMM			× .							
Kinvara Bay	GAM	~	~		× .	Spring	None	Showers	NW	~	-
Mannin	GLMM	~									L
Moy Estuary	GLM	1	× .	×			None	Sunny	S	~	1~
Oranmore Bay	GLMM	~~		1					NW	-*-	L
Westport Bay	GLM		- MA		1	Spring		Cloudy	W		
Bantry Bay	GLM		~								
Dunmanus Bay	GLM										
Kenmare River	GLM										
Roaringwater Bay	GLM/GAM										

Figure 4: Summarized effects for the environmental and observational variables retained in the best overall models by harbour seal location. Symbols represent the form of harbour seal count response for increases in continuous variables and the covariate levels at which harbour seal counts are maximized for discrete variables.

Standardised counts predicted by the modelling approach

Inspection of the constructed plots for the standardised count estimates across the models, count locations and survey years revealed similar trends from GLM, GAM and GLMM models in eight out of 11 locations with the exception of Ballysadare Bay, Cashla Bay and Mannin Bay (Figure 5). 95% confidence intervals around the counts estimated from the three model types were comparatively wide and overlapping in most cases. For locations with significant visit-level variation, confidence intervals surrounding the predicted mean from the GLMM model were wider (Figure 5). Raw mean and maximum counts per year estimated directly from the dataset appeared slightly closer to the GLMM estimates (6 out of 11 locations), compared to the GLM and GAM model estimates (5 out of 11 locations).

Cases of deviance among modelled and raw trends from the count data, e.g., at Emlagh Point (Figure 5), demonstrated the effect of the standardisation on the predicted counts. The raw mean and maxima for this location in 2010 were obtained from replicate counts within a single survey visit early in the moult season that year (Table 1). The GLM, GAM and GLMM models accounted for a general negative relationship between the harbour seal counts and day-of-the-year for this location (Appendix 3). They thus standardised the comparatively higher 2010 data to a lower value. The 2011 counts differed less, owing to a broader range of day-of-the-year data contributing to the raw mean. These anomalous yet natural features in the harbour seal monitoring dataset are addressed further in the Discussion below.

Although overall trends of the standardised counts differ considerably across survey locations in some cases similar patterns can be observed. For example, standardised counts in 2011 appear lower compared to 2012 for An Baile Láir and Adrigole Harbour (Figure 5). Moreover, a general increase in the standardised counts for 2012 can be observed in most of the locations, excluding Cashla Bay where counts remained at the same levels as in 2011 and Emlagh Point and the Moy Estuary where the standardised counts exhibited a decrease. However, in some of the locations overlapping confidence intervals among years suggest that this apparent increase might not be statistically significant, e.g., for Ballysadare Bay (Figure 5).

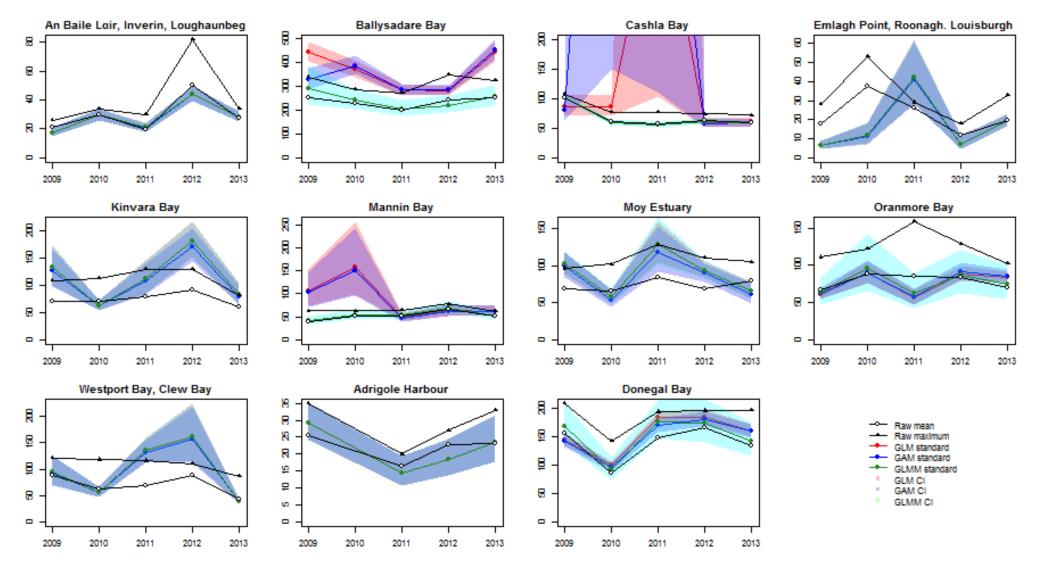


Figure 5: Standardised harbour seal counts generated by the three modelling methods (GLM, GAM, GLMM), associated 95% Confidence Intervals (CI) and raw mean and maximum counts for each land based location. In order to better illustrate trends in predicted and raw counts, the count scale (y axis) in each graph was allowed to vary according to the count magnitude at each location. Care should therefore be taken in interpretation and comparing between count locations/years since the scales shown are not consistent.

By-location boat-based analyses

The models fitted to the four boat-based locations were considerably simpler (Table 3). Dayof-the-year linear and quadratic effects were retained in the best model for Bantry Bay only (Table 4, Figure 6).

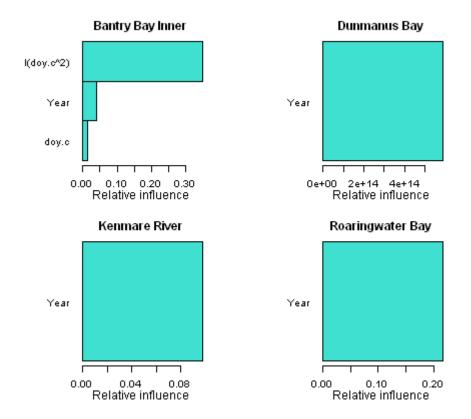


Figure 6: Relative influence of the selected observational variables on the deviance of the best GLM model for each boat based count location. Very large values of relative influence reflect the single observation and estimate per year for that location resulting in a fitted deviance close to zero. Care should be taken in interpretation and comparing between count locations since the scales of computed relative influence (x axes) shown are not consistent.

Concerning standardised counts of the boat-based survey locations, very similar trends were observed between the two modelling methods (GLM and GAM; Figure 7).

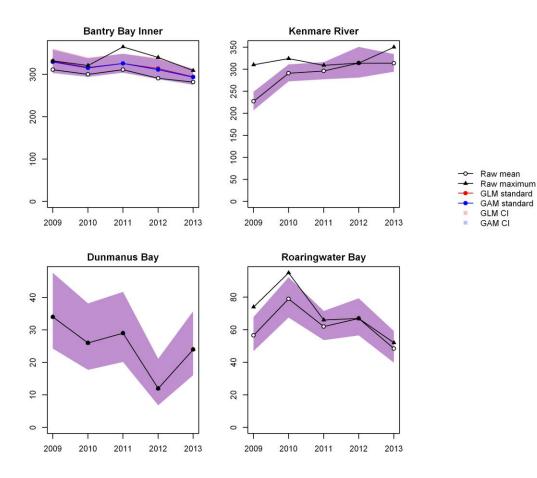


Figure 7: Standardised harbour seal counts generated by the two modelling methods (GLM, GAM), associated 95% Confidence Intervals (CI) and raw mean and maximum counts for each boat based location. In order to better illustrate trends in predicted and raw counts, the count scale (y axis) in each graph was allowed to vary according to the count magnitude at each location. Care should therefore be taken in interpretation and comparing between count locations/years since the scales shown are not consistent.

Discussion

Model checking

Count data often display greater variability than that predicted by simple parametric models such as models using the Poisson distribution (Richards, 2008). Such overdispersed data might either indicate the existence of important environmental, observational or natural covariates not accounted for in the model, or a lack of independence among study subjects (Hilbe, 2011). As demonstrated by the improvement of the diagnostic plots upon fitting GLMM models (Appendix 2), overdispersion observed in some of the harbour seal survey locations might be due to variability at the 'visit' level, which is taken into consideration within the random effects of the GLMM models. A simple causative factor in this regard could be differences in haul-out behaviour of the study animals (e.g., in the timing of hauling out or in the proportion of the local population doing so). The use of observation level random effects has been shown to produce more accurate parameter estimates when overdispersion is caused by random noise surrounding the Poisson data or clustering in the count data, i.e., the existence of non-independent observations (e.g., Harrison, 2014).

Model selection for land-based count data

The three modelling methods used in this project (GLM, GAM, GLMM), although based on the same distributional assumption, display varying sensitivity to the characteristics of different datasets (e.g., GAM: non-linear effects of explanatory variables; GLMM: nonindependent observations; and random effects in the GLMM). As judged via the AICc the GLMM did not consistently out-perform the GLM and the GAM in some survey locations (Table 5). In such cases where the GLMM did not out-perform other modelling approaches the estimated random effects variance (i.e., visit-level variability) was very small (i.e., often zero) and the model effectively reverted to a GLM. GLMMs, however, encapsulated the nonindependence of the counts at the visit level, provided more stable standardised counts particularly in cases where there were significant outlier data (Figure 5, Appendix 1) and they were generally simpler in terms of the number of coefficients than either the GLM or GAM models. For these reasons the Poisson GLMM is considered the most suitable modelling framework, of those considered, for fitting and interpreting these data.

For those land-based count locations where three survey visits were conducted annually, including both visit-level random effects and year specific day-of-the-year effects in the GLMM models was not feasible because both use the same information, albeit tacitly. Rather than allowing day-of-the-year effects to vary by year we opted, from both a practical and fundamental perspective, (a) to assume that day-of-the-year effects were constant across years and (b) to model the visit-level variability with random effects. It would be important for future analysis that the values of the random effects in relation to day-of-the-year are monitored for systematic departures from the random effects assumption, which would indicate a true change in the day of year effect. With additional years of consistent standardised data, such effects could be modelled by blocking day-of-the-year effects across sets of years.

Variable retention and influence

Cases where different variables were retained by the three model methods further illustrate the comparative performance of the GLMM models wherein a visit-level effect was accounted for. For example, the GLM model for Cashla Bay indicated the survey year as the most influential explanatory variable on the count data. This reflected unusually high counts recorded at the site across two survey visits in 2009, whereas the GLMM for this location described a relatively high variance at the visit-level, which is captured in the random effects of the model. Among the tested variables the: year, day-of-the-year and hours from Low Water (tidal state) had the greatest influence on harbour seal counts in most of the locations (Table 4, Figures 2&3). While both latter variables can be controlled in designing monitoring approaches, their influence on harbour seal counts varied in different locations (Figures 2&3) which makes them important factors to consider in survey planning per location. Environmental variables such as weather, wind direction, wind speed and tidal range were only influential covariates for count data at some locations (e.g., Kinvara Bay, Westport Bay; Figure 3). This may reflect a comparative efficacy in methodology which currently seeks to restrict harbour seal surveys to certain weather limits for example (NPWS, 2010, 2012). In cases where environmental conditions remained relatively influential via model fitting this might be due to the particular geographic characteristics of each location (e.g., exposure, intertidal topography) that could alter the influence of such covariates on seal haul-out behaviour, which would be inherently reflected in the harbour seal counts. To further investigate such differences, the direction of the estimated effects (e.g., per weather condition) per location and model type are provided in Appendix 3.

The presence of disturbance was retained in GLMM fits for three of the 12 land-based survey locations investigated (Cashla Bay, Kinvara Bay, Moy Estuary; Table 4) where it had an important influence on the recorded harbour seal counts (Figure 3). Whereas the three locations did not exhibit a higher incidence of disturbance compared to the rest of the land-based survey locations, in all three cases where disturbance was retained in the GLMM models the majority of disturbance records were described as "dinghy passing" (Appendix 1). According to observations provided by the NPWS observers these occurrences very often resulted in shore-based harbour seals entering the water and moving away, thus directly affecting the subsequent counts. Disturbance could be further investigated in these data by modelling the proportion of animals in and out of the water with and without disturbance. Additional analyses could also include finer spatial-scale analyses of haul-out site preference (Appendix 1) in the presence of certain categories of disturbance.

Whereas the uniqueness of each count location may explain potential differences in the influence of specific variables on the harbour seal counts, the general trends in the main environmental and observational variables (i.e., day-of-the-year and hours from Low Water) resembled the expected patterns. A peak in the use of haul-out sites has been reported to occur during the moulting season (Cronin & Ó Cadhla, 2007) with subsequent decreases in the number of seals on shore (Cordes *et al.*, 2011). Moreover, seal counts were generally expected to increase in the hours close to the time of Low Water as more space is gradually made available on the haul-out sites. Whether symmetry exists around Low Water at individual haul-out locations can be investigated by comparing the estimated GLM and GAM effects of tidal state (Appendix 3).

Standardised counts

In locations with a well-balanced distribution of important environmental covariates affecting counts across years, the standardised and raw summary statistics differed minimally (e.g., Donegal Bay). In contrast, in cases with a few counts in more extreme environmental or observational conditions (e.g., Emlagh Point, Oranmore Bay), the latter may heavily influence the raw values but were nevertheless standardised for in the means predicted by the models. Furthermore, in survey locations where the predicted counts differed between the GLM, GAM and GLMM models (e.g., Ballysadare Bay, Cashla Bay, Mannin Bay) the means predicted by the GLM and GAM methods were usually accompanied by notably wide confidence intervals, while GLMM models generated trends that were considerably closer to the raw counts (Figure 5). Such survey locations usually display high variability at the visit-level which is accounted for better via the GLMM models, providing improved standardised counts. For Ballysadare Bay and Mannin Bay, this was reflected in the AICc value which was lower for the GLMM model (Table 5). However this was not the case for Cashla Bay, where the GAM model performed better according to the AICc criterion (Table 5).

By-location boat-based analyses

The relatively smaller amount of data obtained by boat surveys led to simpler models, as highlighted by the case of Dunmanus Bay where a total of six counts (Table 1) did not allow for the inclusion of more variables other than year (Tables 3 & 4). Moreover, in locations with more data, counts were not balanced across different levels of the included variables e.g., counts in Bantry Bay were only performed during neap tides which explains the exclusion of tidal range in the respective models (Tables 3 & 4). This is an indication of the difficulty in balancing the conditions at which counts are made while covering these locations. However, considering the poor visibility which limits the ability to perform land based surveys in these locations, it is important that these areas are covered by using the existing design, especially as the addition of future surveys will gradually allow for more indepth statistical analysis.

General discussion

The set of land-based count locations that were subject to generalised modelling consisted of 12 very different locations in terms of their: spatial extent, habitat, topography, and geography. Through the current analysis different environmental and observational covariates were shown to influence harbour seal counts at each location. This phenomenon has previously been reported in the literature and has been attributed to the uniqueness of the characteristics of each haul-out site (Thompson *et al.*, 1997; Cronin, 2007). As a result of the intrinsic variability demonstrated in the current dataset trends in the recorded count data were also different across survey locations, highlighting the importance of performing such detailed analysis at a location level. Statistical analysis at a higher level (i.e., by collating and merging the data across all locations) would enable an overall examination of the harbour seal population trends but this approach would create complications in the analysis

per-se. Namely the differences in covariate influences per survey location would require the addition of numerous interactions within the models which would be demanding in terms of parameter numbers given the number of years in the present dataset.

Although peaks in harbour seal counts vary not only regionally but also temporally (i.e., from year to year) due to a variety of environmental and biological factors (Huber *et al.*, 2001), identifying the observational covariates which influence such peaks is important for survey planning, as explored in Phase 2 of the current project.

Acknowledgements

This work was initiated and funded in 2014 by the National Parks & Wildlife Service of the Department of Arts, Heritage and the Gaeltacht under the project "Statistical modelling and power analysis of NPWS Harbour seal monitoring data". We wish at the outset to thank all National Parks & Wildlife Service observers and external personnel who carried out the very substantial survey effort and data entry that contributed to this project, namely Dermot Breen, Carl Byrne, Helen Carty, Cameron Clotworthy, Pat Dawson, Pascal Dower, Leonard Floyd, Emma Glanville, Patrick Graham, Clare Heardman, Gerry Higgins, Tara Keena, James Kilroy, Emer Magee, Lee McDaid, Eoin McGreal, Frank McMahon, Jacinta Murphy, Louise O'Boyle, Irene O'Brien, Oliver Ó Cadhla, Aonghus O'Donaill, Declan O'Donnell, Ger O'Donnell, Barry O'Donoghue, Danny O'Keeffe, Michael O'Sullivan, Tim Roderick, Andrew Speer, Raymond Stephens, Rebecca Teesdale and Fiona Wheeldon. Also thanks to all NPWS District Conservation Officers and senior line management who oversaw regional implementation of the harbour seal monitoring programme in 2009-2013. We are also grateful to Dr Oliver Ó Cadhla, Dr Ferdia Marnell and Dr Eamonn Kelly of the Department of Arts, Heritage and the Gaeltacht and Dr Ian O'Connor and Dr Rick Officer of GMIT for their support towards all stages of the project.

References

- Abeare, S.M. (2009). Comparisons of boosted regression tree, glm and gam performance in the standardization of yellowfin tuna catch-rate data from the Gulf of Mexico longline fishery. Master's thesis. Louisiana State University and Agricultural and Mechanical College.
- Bowen, W.D., Ellis, S.L., Iverson, S.J. & Boness, D.J. (2003). Maternal and newborn lifehistory traits during periods of contrasting population trends: implications for explaining the decline of harbour seals (Phoca vitulina), on Sable Island. *Journal of Zoology*, 261: 155–163.

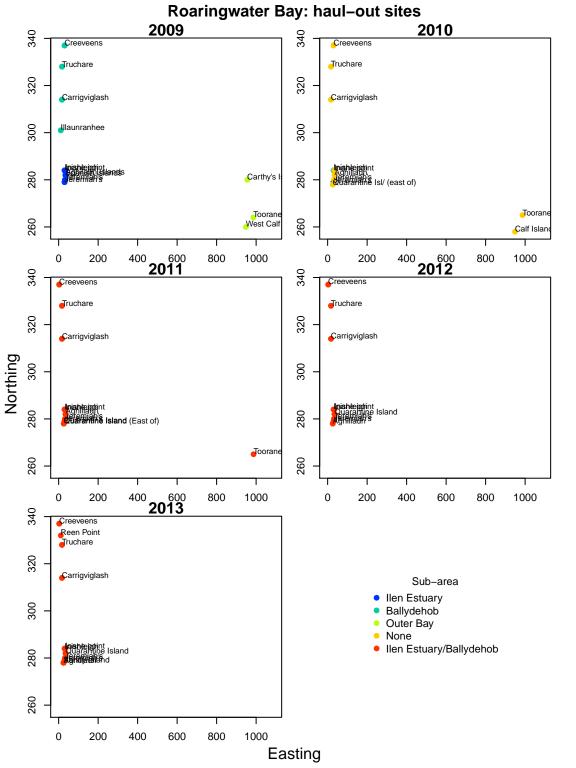
- Breslow, N.E. & Clayton, D.G. (1993). Approximate Inference in Generalized Linear Mixed Models. *Journal of the American Statistical Association*, 88: 9-25.
- Burnham, K.P. & Anderson, D.R. (2002). Model *selection and multimodel inference: a practical information-theoretic approach*. Springer.
- Clark, J.S. (2007). Models for Ecological Data. Princeton University Press.
- Cordes, L.S., Duck, C.D., Mackey, B.L., Hall, A.J. & Thompson, P.M. (2011). Long-term patterns in harbour seal site-use and the consequences for managing protected areas. *Animal Conservation*, 14: 430-438.
- Cook, R.D. & Weisberg, S. (1982). Residuals and influence in regression. Chapman & Hall.
- Cosgrove, R., Sheridan, M., Minto, C. & Officer, R. (2014). Application of finite mixture models to catch rate standardization better represents data distribution and fleet behavior. *Fisheries Research*, 153: 83-88.
- Cronin, M. A. (2007). Abundance, habitat use and haul out behaviour of harbour seals (*Phoca vitulina vitulina* L.) in southwest Ireland 2007. PhD thesis. University College Cork, Ireland. 263 pp.
- Cronin, M., Duck, C., Ó Cadhla, O., Nairn, R., Strong, D. & O'Keeffe, C. (2007). An assessment of population size and distribution of harbour seals in the Republic of Ireland during the moult season in August 2003. *Journal of Zoology*, 273(2): 131-139.
- Cronin, M. A. & Ó Cadhla. O. (2007). NPWS Phocid monitoring methods and interval assessment. Recommendations for the monitoring of the harbour seal and grey seal populations in the Republic of Ireland. 46 pp.
- Harrison, X.A. (2014). Using observation-level random rffects to model overdispersion in count data in ecology and evolution. PeerJ2:e616; DOI: 10.7717/peerj.616
- Hastie, T. J. & Tibshirani, R. J. (1990). Generalized Additive Models. Chapman & Hall/CRC.
- Heardman, C., O'Donnell, D. & McMahon, D. (2006). The status of the harbour seal *Phoca* vitulina L. in inner Bantry Bay, Co Cork and inner Kenmare River, Co. Kerry: 1964-2004. *Irish Naturalists' Journal* 28(5), 181-191.
- Hilbe, J.M. (2011).Negative binomial regression. 2nd edition. Cambridge: Cambridge University Press.
- Huber, H.R., Jeffries, S.J., Brown, R.F., DeLong, R.L., Van Blaricom, G. (2001). Correcting aerial survey counts of harbour seals (*Phoca vitulina richardsi*) in Washington and Oregon. *Marine Mammal Science*, 17: 276–293.

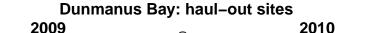
- Lonergan, M., Duck, C.D., Thompson, D., Mackey, B.L., Cunningham, L. & Boyd, I.L. (2007). Using sparse survey data to investigate the declining abundance of British harbour seals. *Journal of Zoology (Lond.)*, 271: 261–269.
- Maunder, M.N., Punt, A.E., (2004). Standardising catch and effort data: a review of recent approaches. *Fisheries Research*, 70: 141-159.
- McCullagh, P., & Nelder, J. A. (1989). Generalized Linear Models, Second Edition. Chapman and Hall, London.
- NPWS. (2010). Harbour seal population monitoring 2009-2012: Report No. 1. Report on a pilot monitoring study carried out in southern & western Ireland, 2009. National Parks & Wildlife Service, Department of the Environment, Heritage and Local Government, Dublin. 11pp.
- NPWS. (2012). Harbour seal pilot monitoring project, 2011. National Parks & Wildlife Service, Department of the Environment, Heritage and Local Government, Dublin. 15pp.
- Richards, S. A. (2008). Dealing with overdispersed count data in applied ecology. *Journal of Applied Ecology*, 45: 218-227.
- Thompson, P.M., Tollit, D.J., Wood, D., Corpe, H.M., Hammond, P.S., & Mackay, A. (1997). Estimating harbour seal abundance and status in an estuarine habitat in north-east Scotland. *Journal of Applied Ecology*, 34: 43-52
- Watts, P. (1996). The diel hauling-out cycle of harbour seals in an open marine environment: correlates and constraints. *Journal of Zoology*, 240(1), 175–200.
- Wilson, S.C., Pierce, G.J., Higgins, C. & Armstrong, M.J. (2002). Diet of harbour seals *Phoca vitulina* of Dundrum Bay, north-east Ireland. *Journal of Marine Biology Association*, 82: 1009–1018.
- Wood, S.N. (2006). *Generalized Additive Models: An Introduction with R*. Chapman & Hall/CRC.
- Zuur, A.F., Ieno, E.N., Walker, N., Saveliev, A.A., Smith, G.M. (2009). *Mixed Effects Models* and Extensions in Ecology with R. Springer.

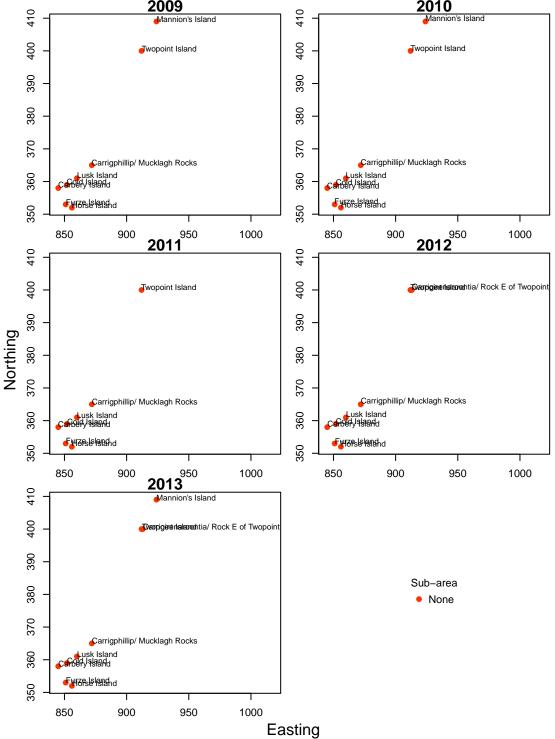
Appendices

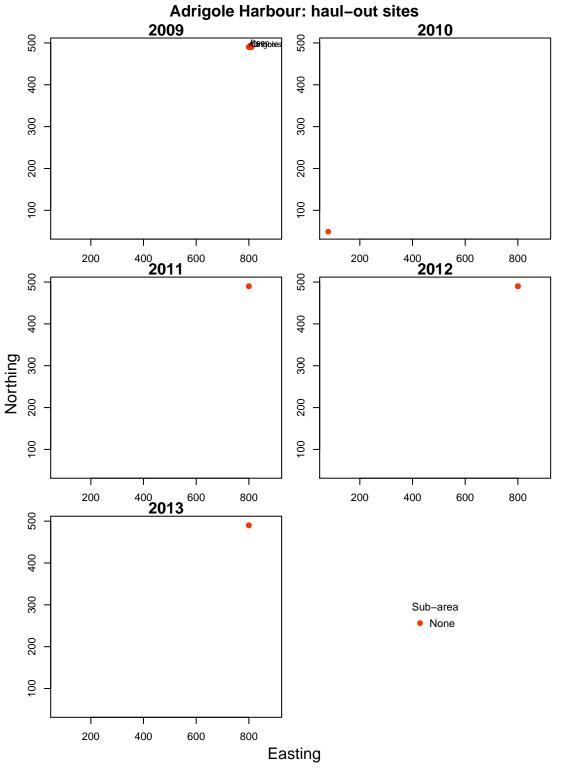
Appendix 1

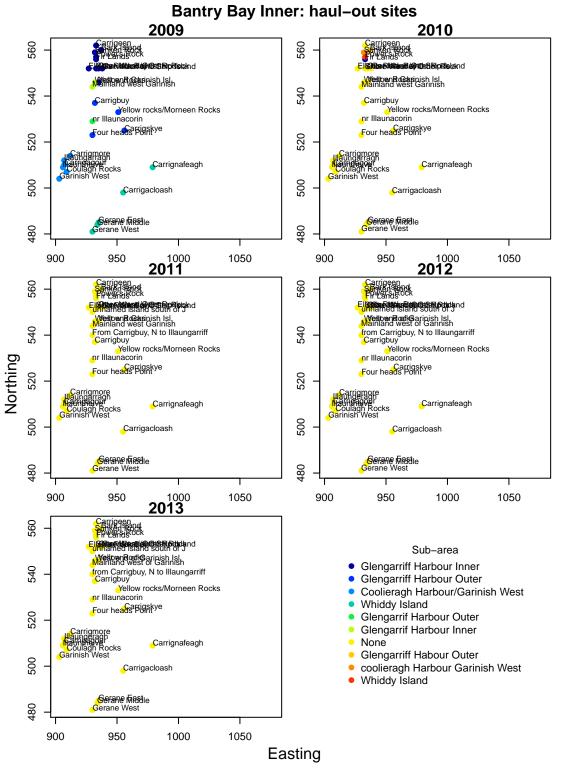
Exploratory visualisation of the NPWS Harbour seal monitoring programme database

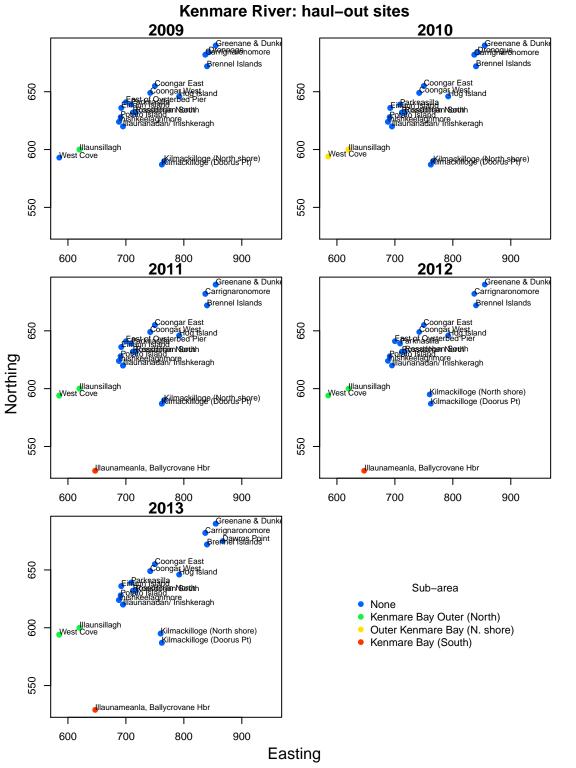


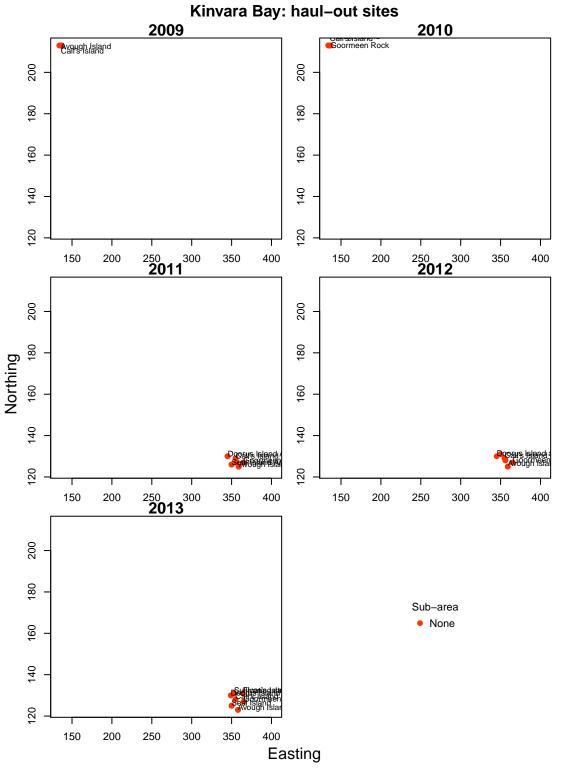


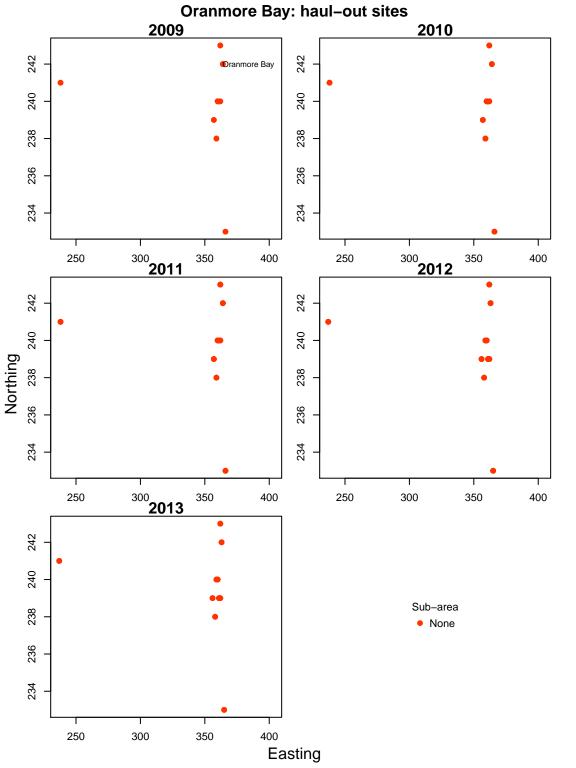


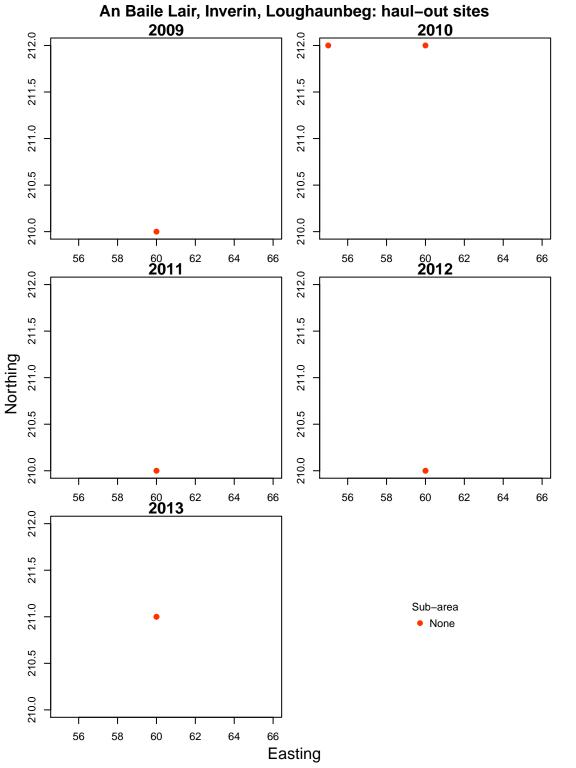


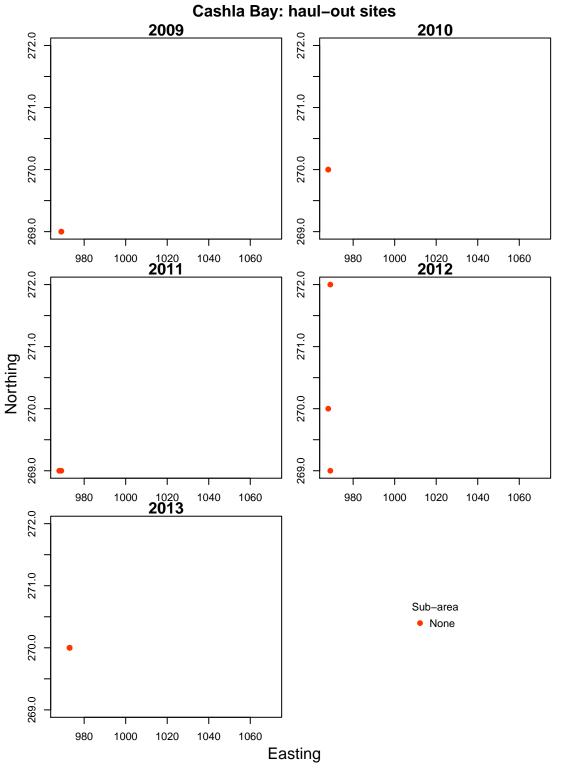


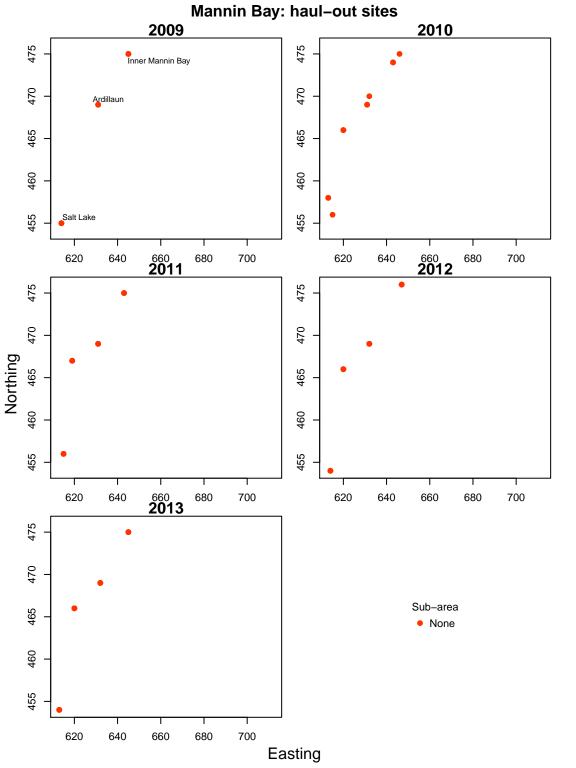


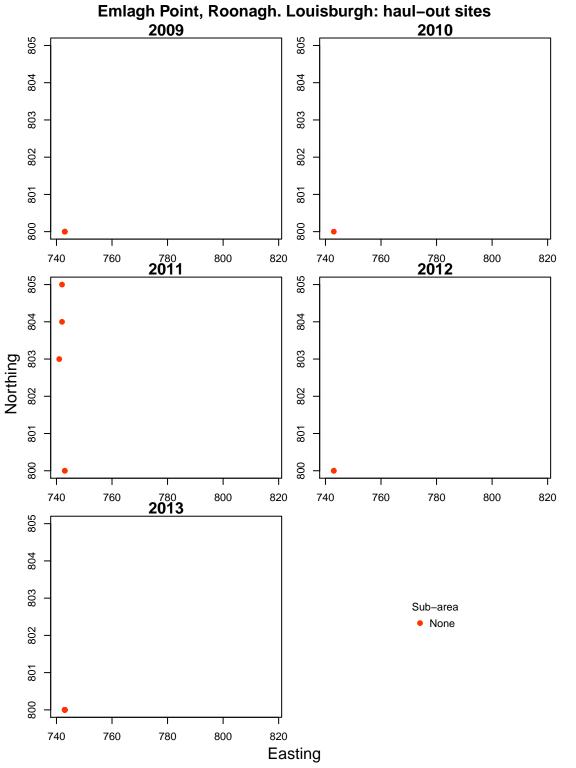


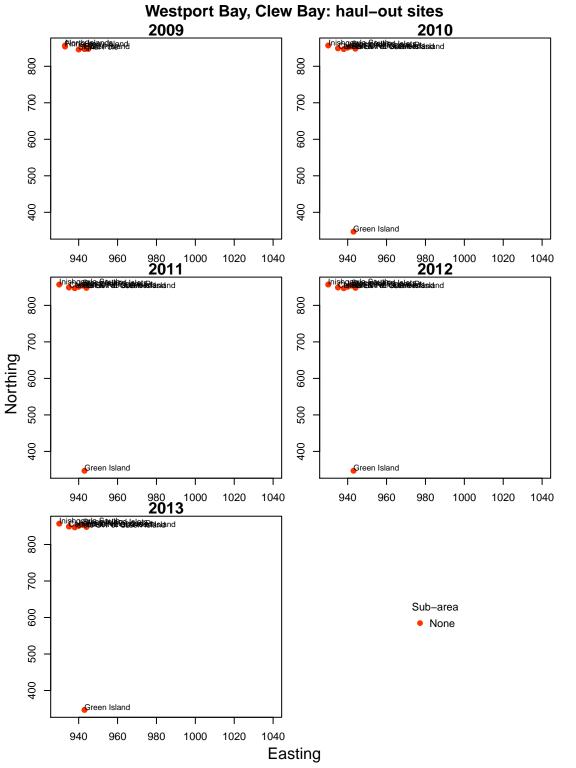


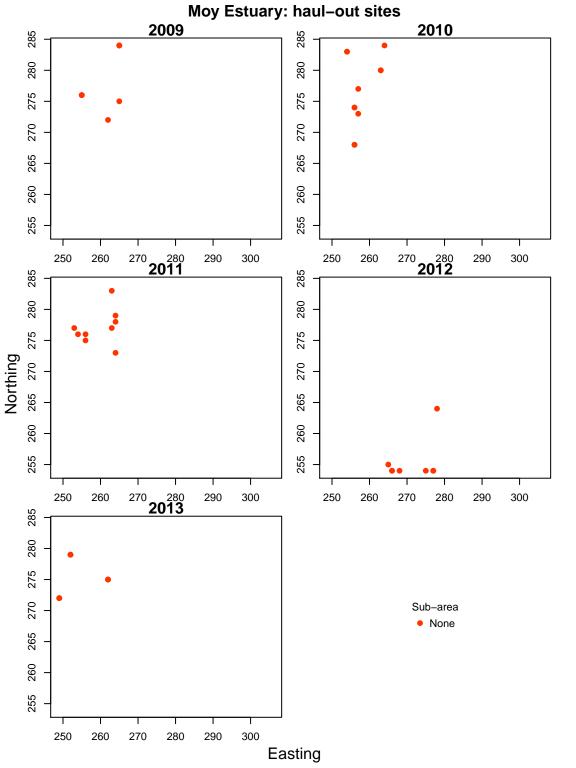


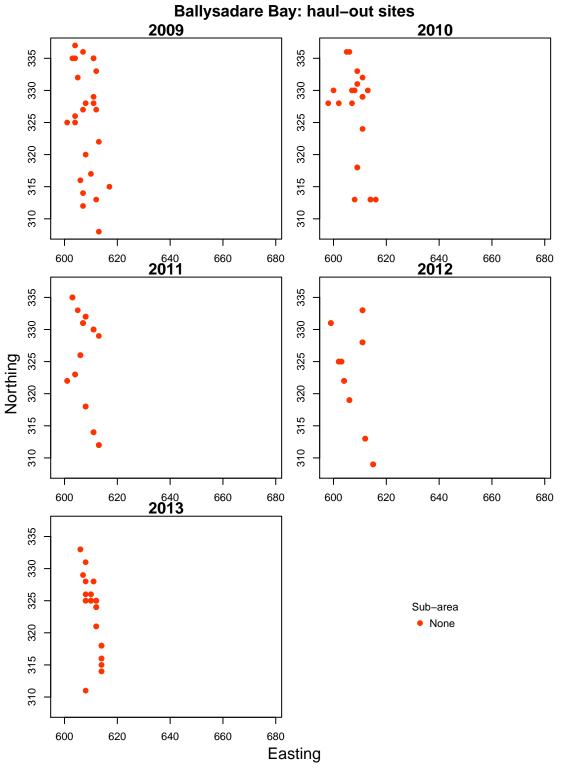


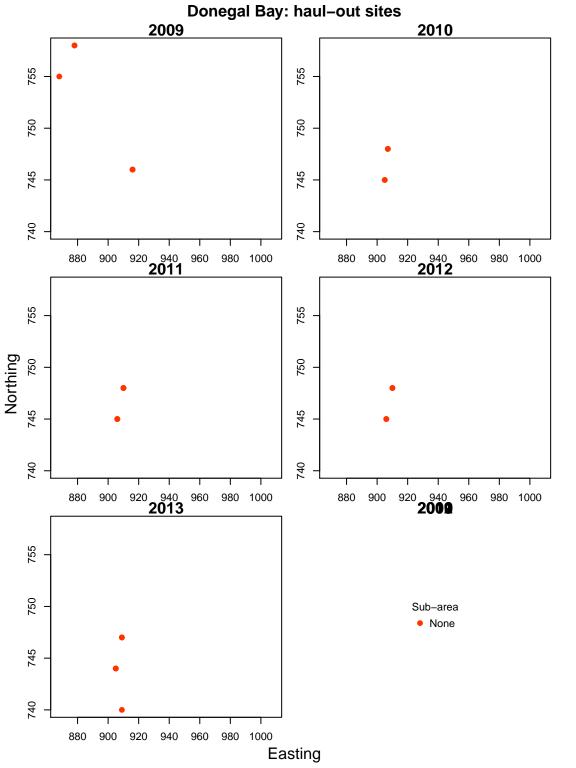


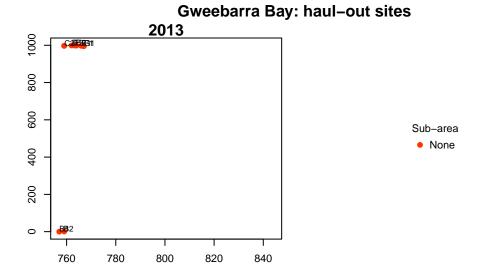




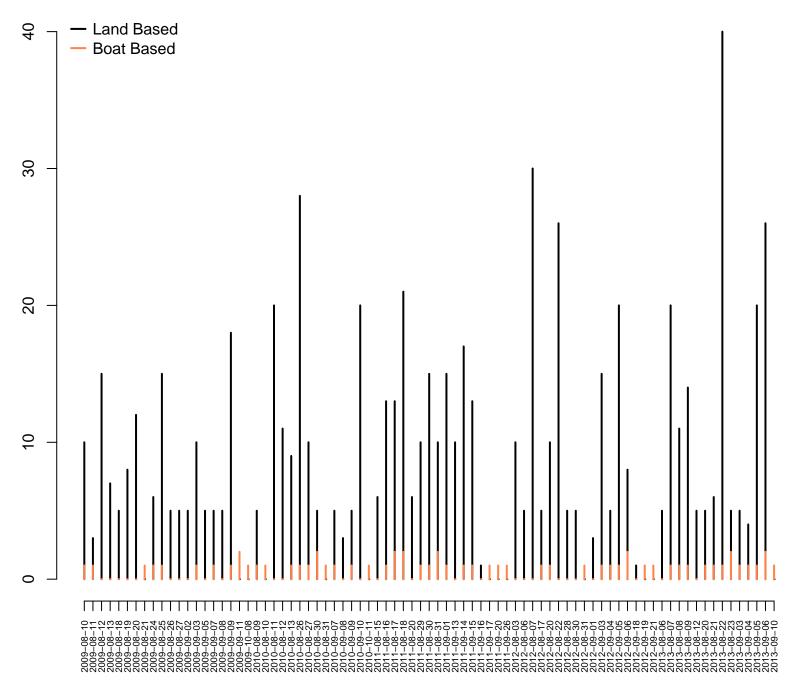








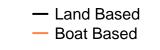
Easting

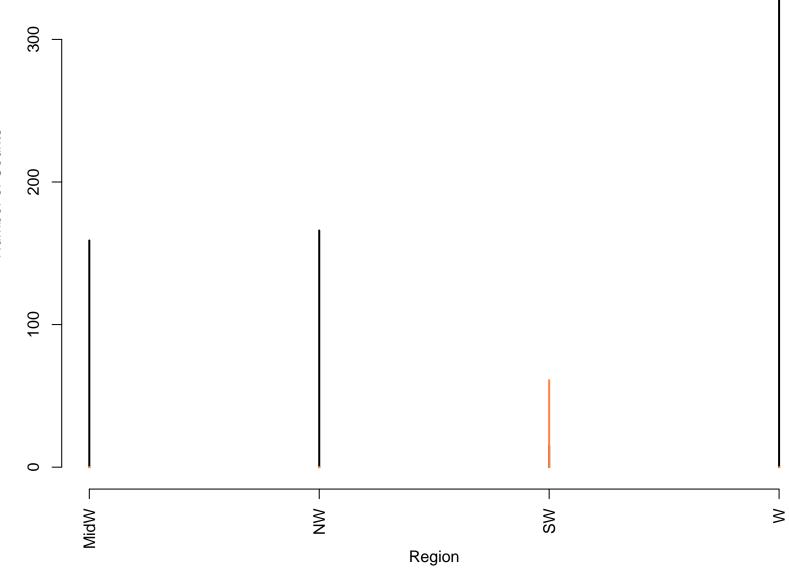


Date

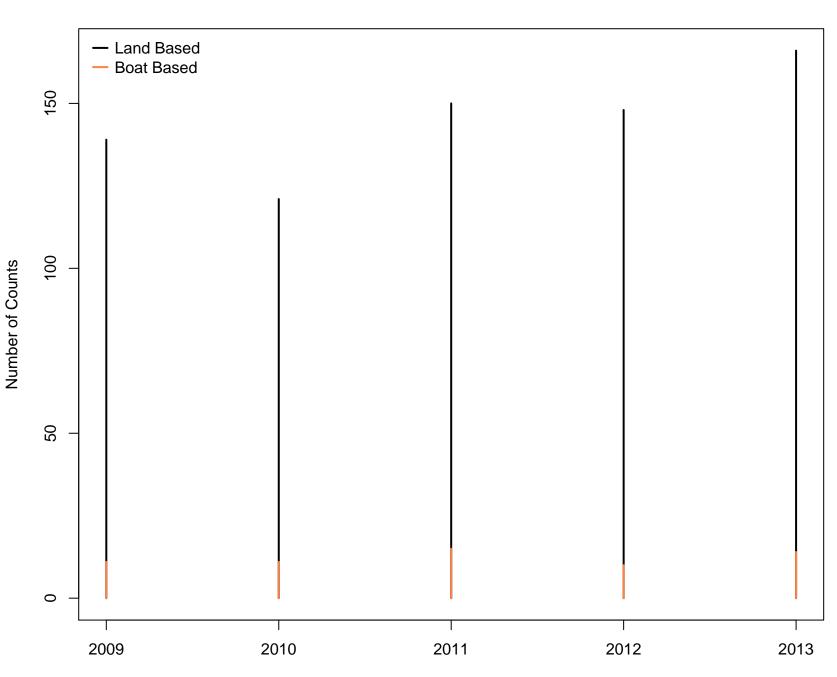
Number of Counts



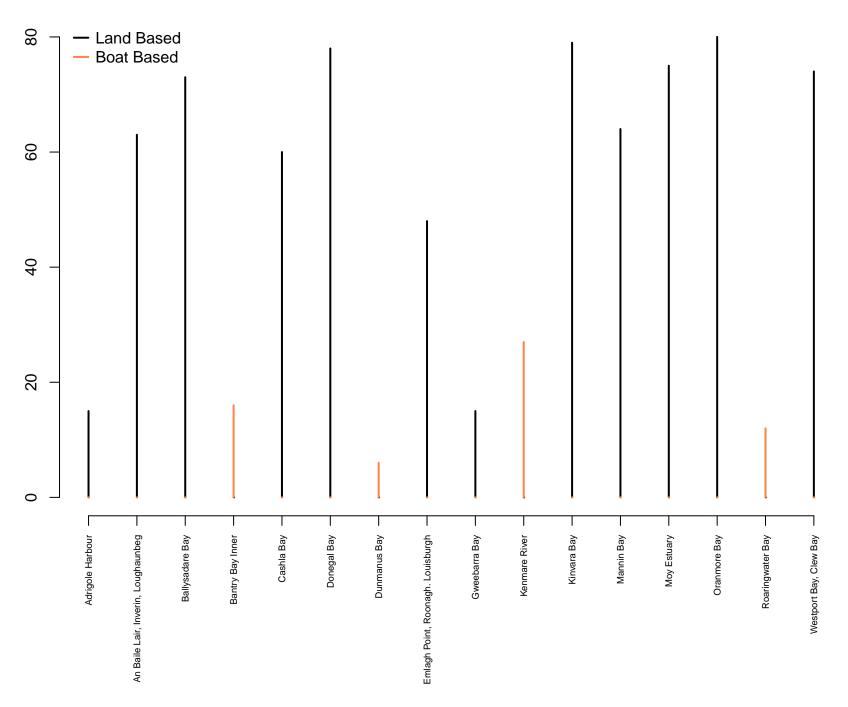


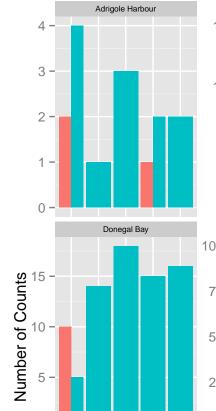


Number of Counts



Location





Mannin Bay

2013 -

2012 -

5 -

0

15 -

10 -

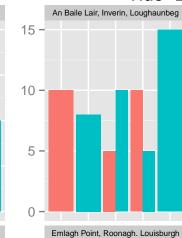
5 -

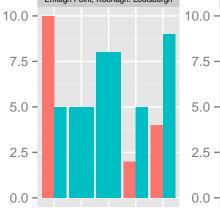
0

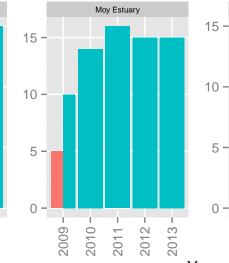
2009 -

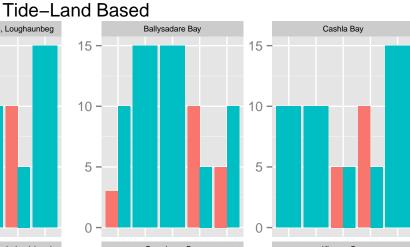
2010 -

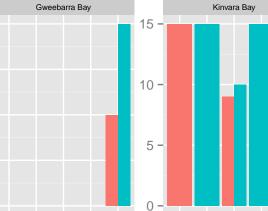
2011 -











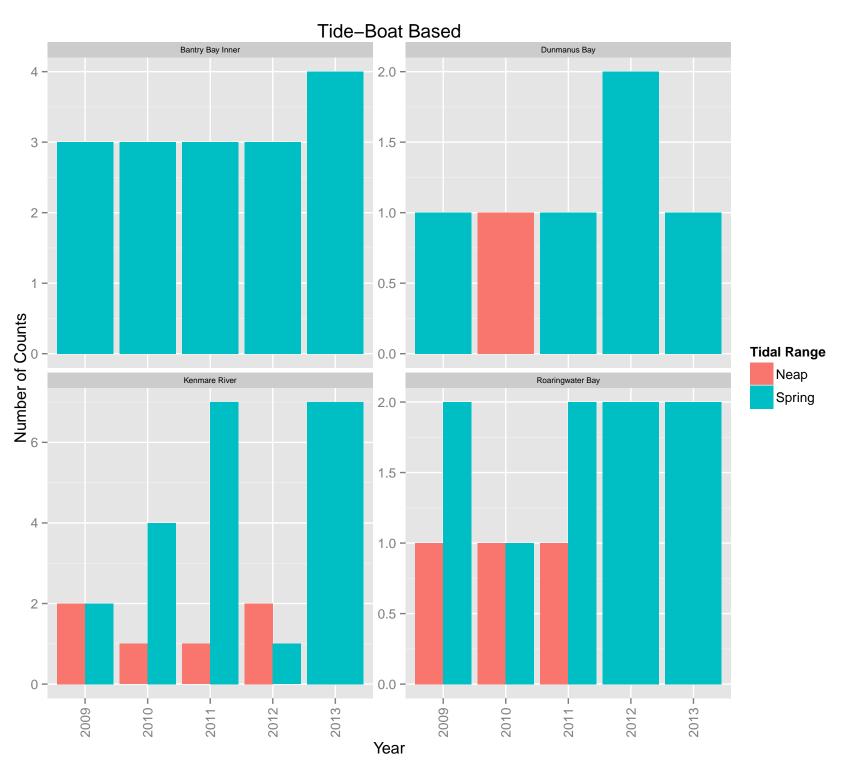


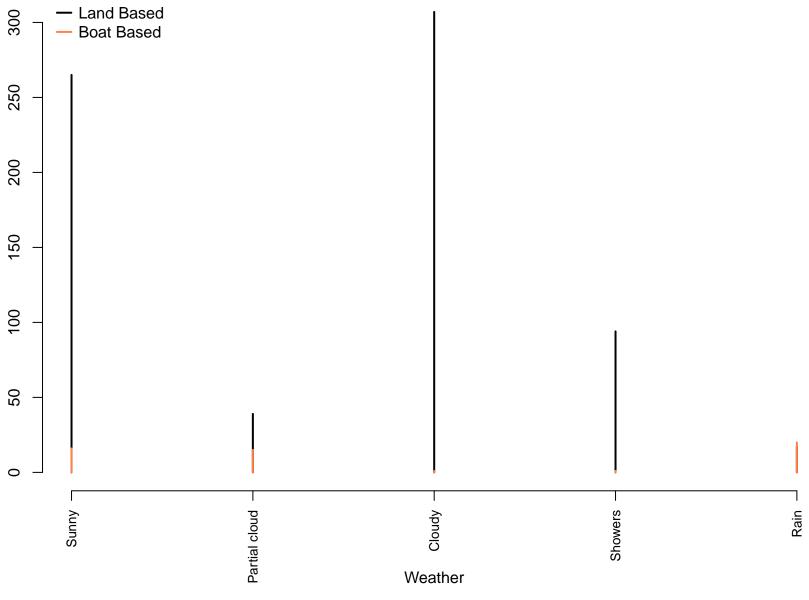


Year

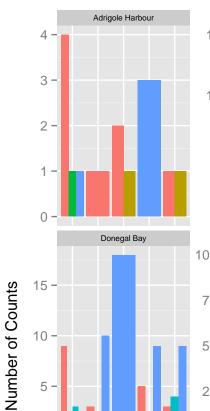
5 -

0



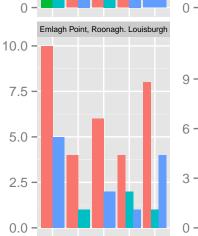


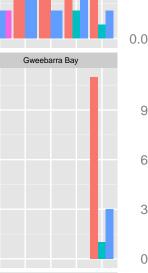
Weather

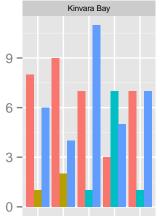


Weather-Land Based









Westport Bay, Clew Bay

2011 -

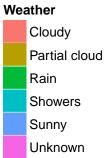
2012 -

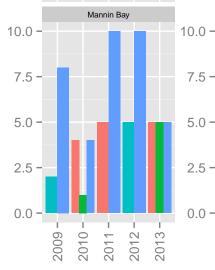
2013 -

2010 -

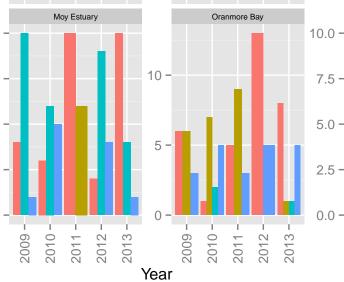
2009 -

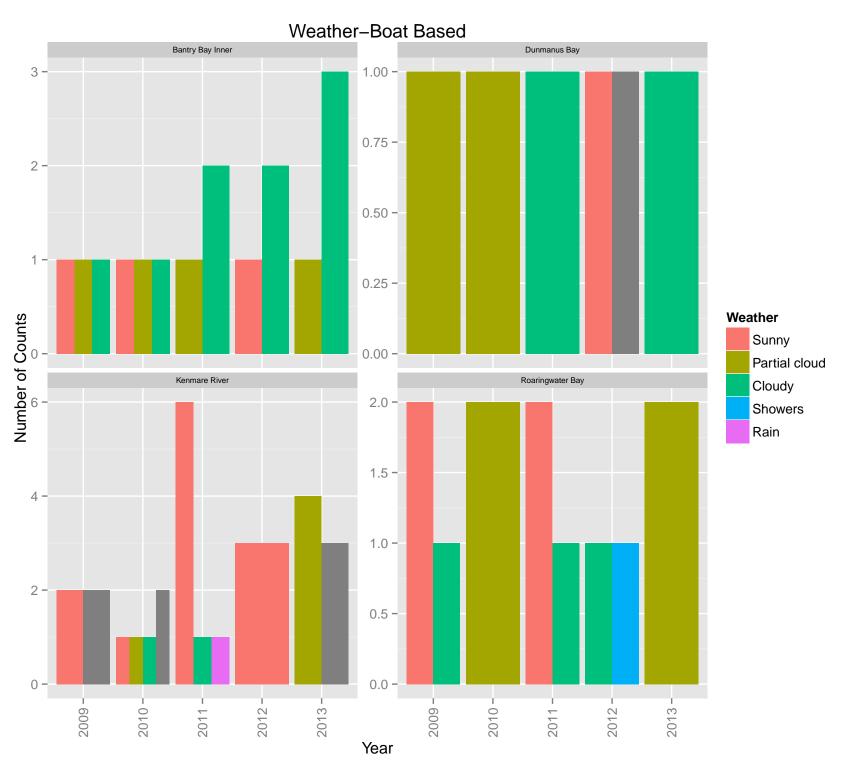
Cashla Bay



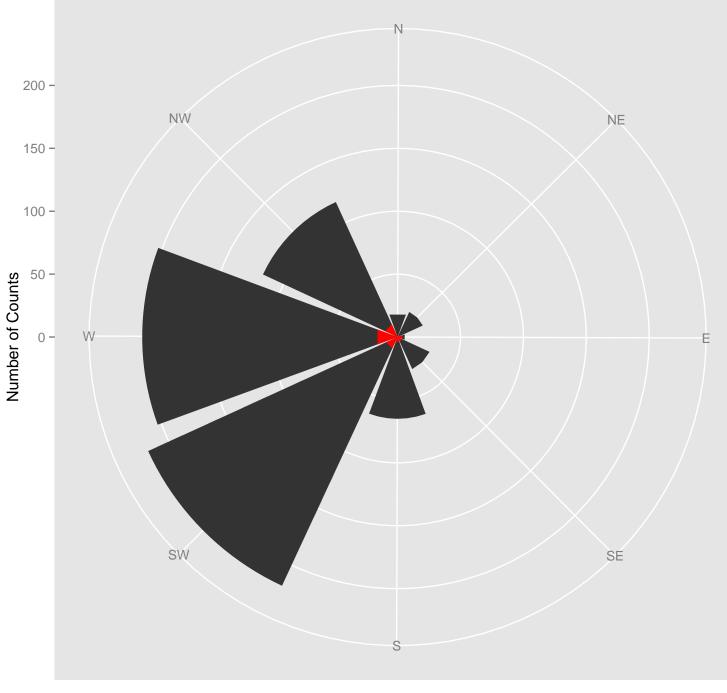


0

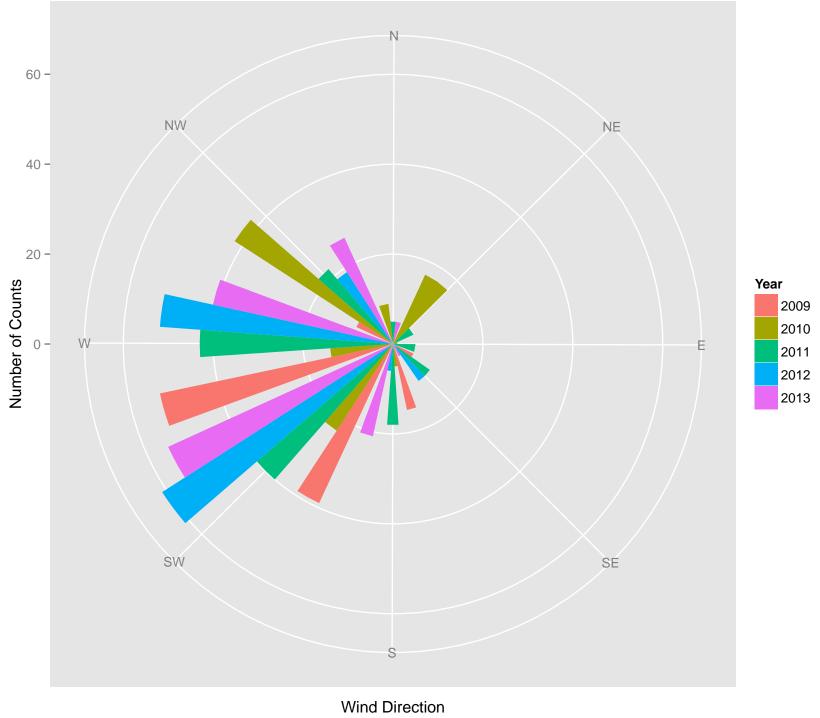




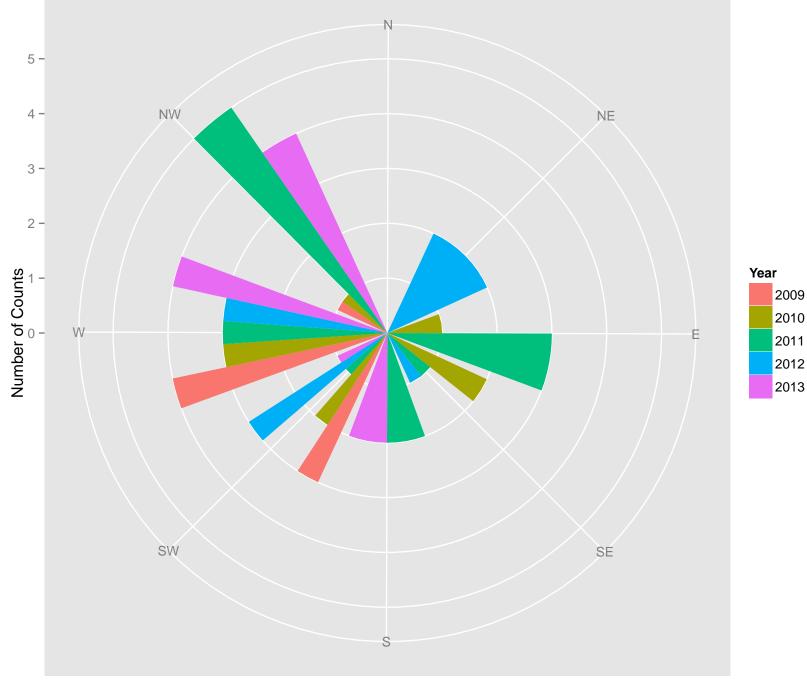
Wind Direction

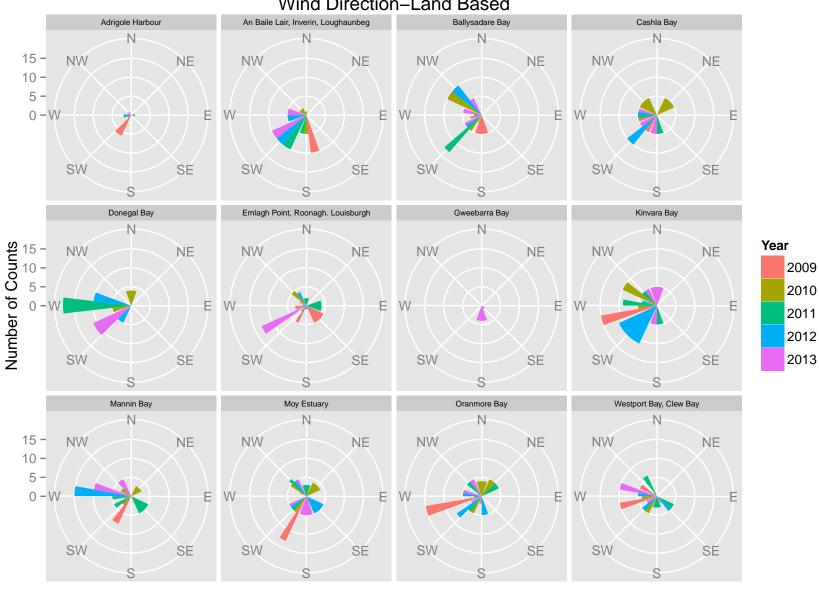


Wind Direction-Land Based



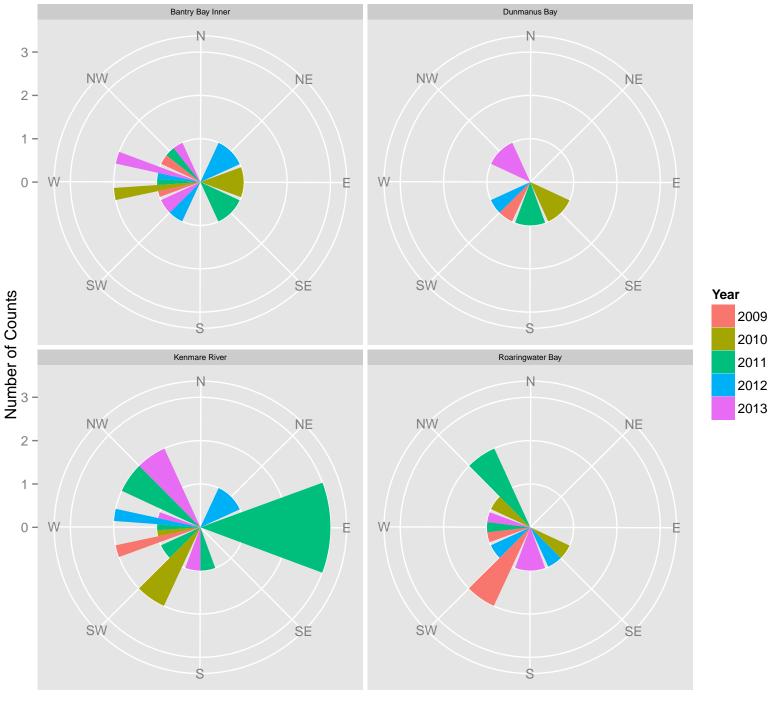
Wind Direction-Boat Based



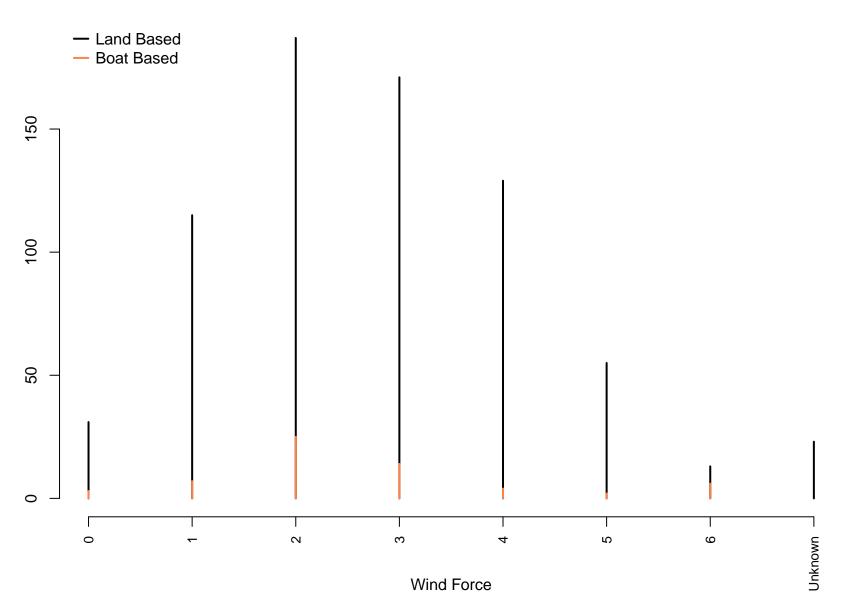


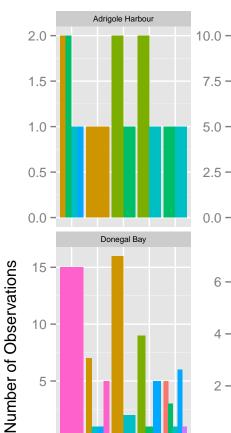
Wind Direction-Land Based

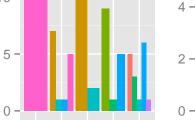
Wind Direction–Boat Based



Wind Force



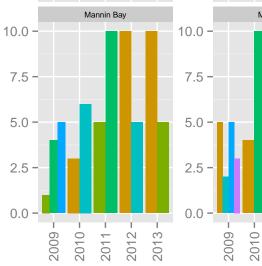




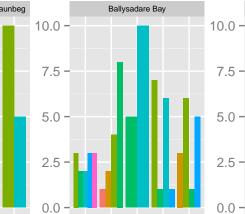
6

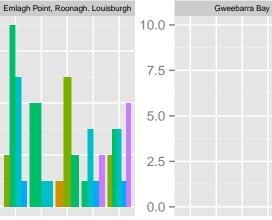
Moy Estuary

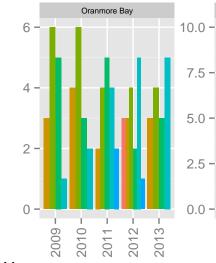
2011 -2012 -

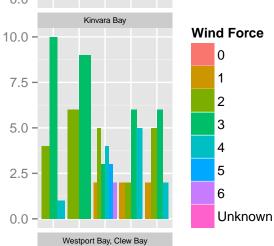




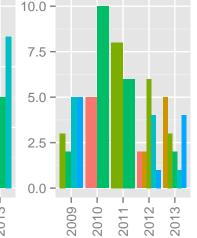






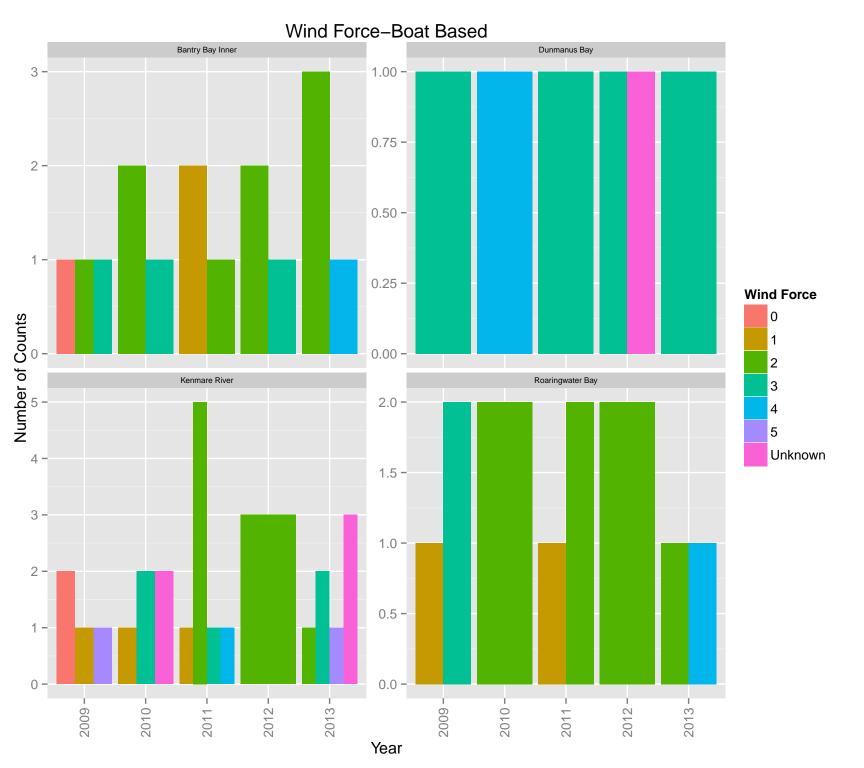


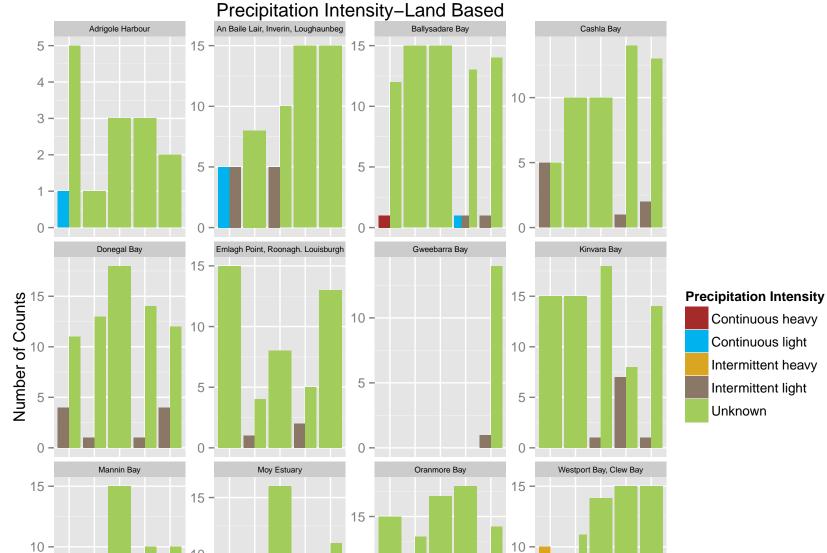
Cashla Bay



Year

2013 -





10 -

5 -

0

Year

2012 -

2013 -

2009 -

2010 -

ļ ł.

2011

5 -

0

2009 -2010 -

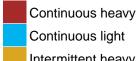
2013 -

2012 -

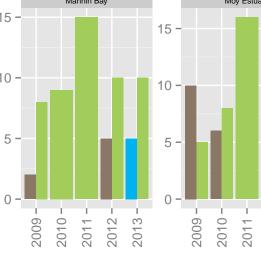
2012 -2013 -

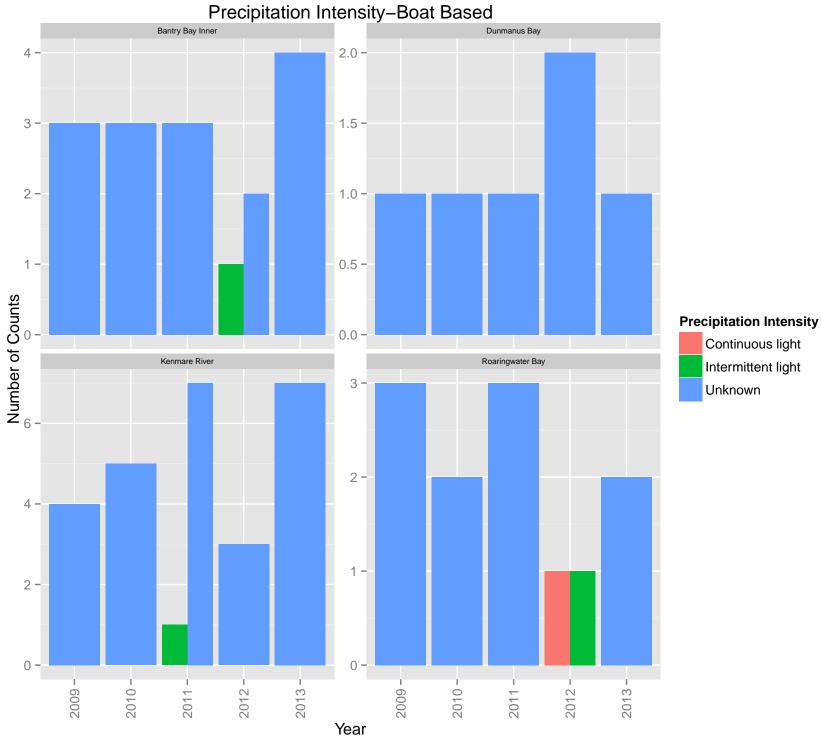
Ľ,

2011

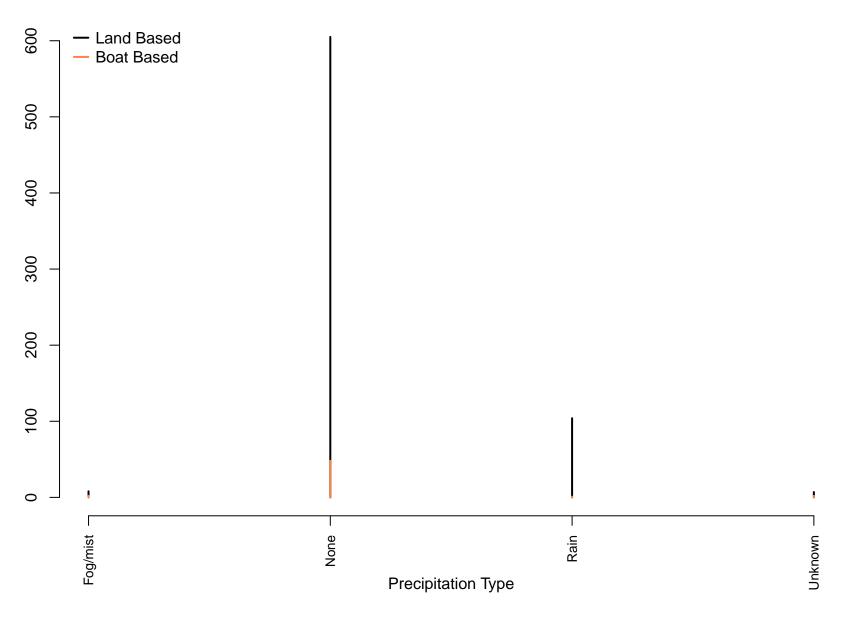


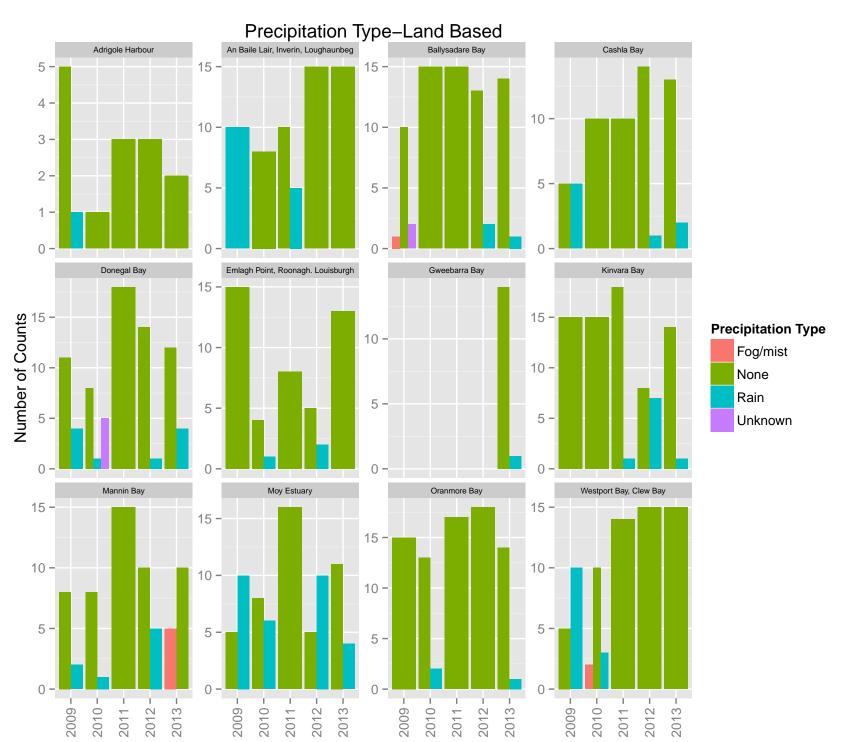
Continuous light Intermittent heavy Intermittent light



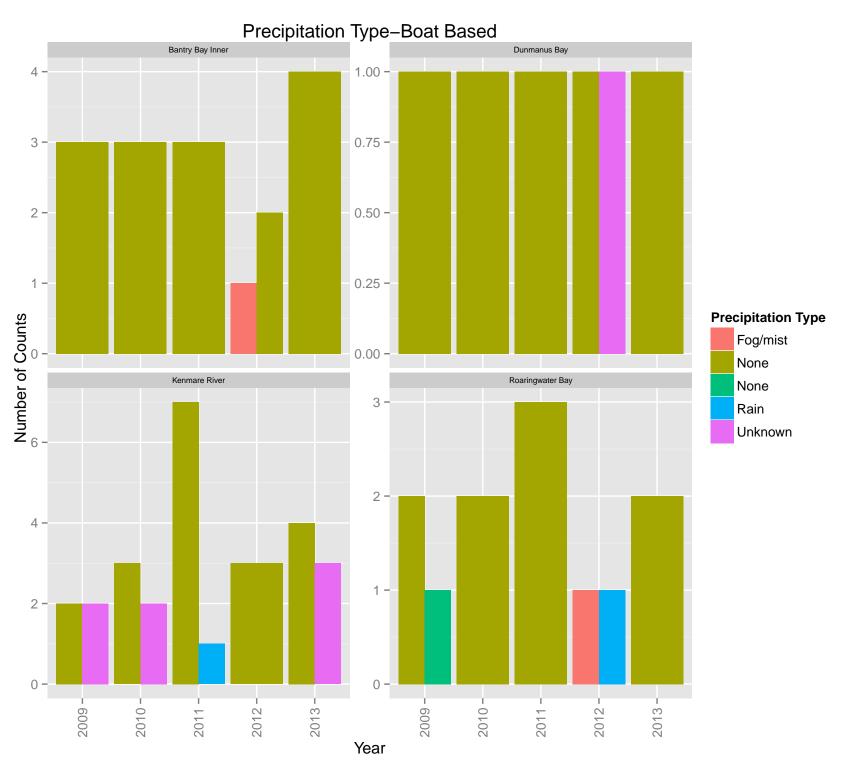


Precipitation Type

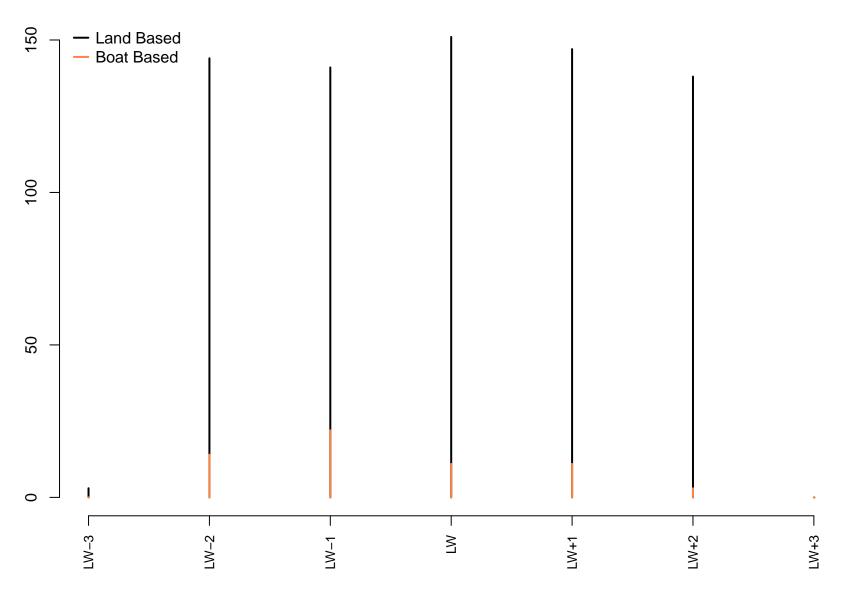


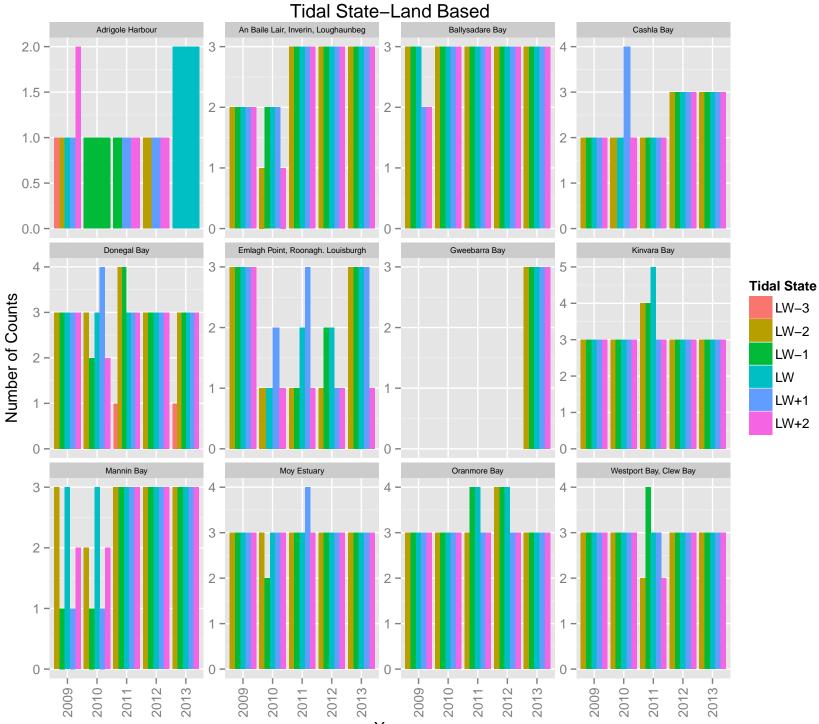


Year



Tidal State



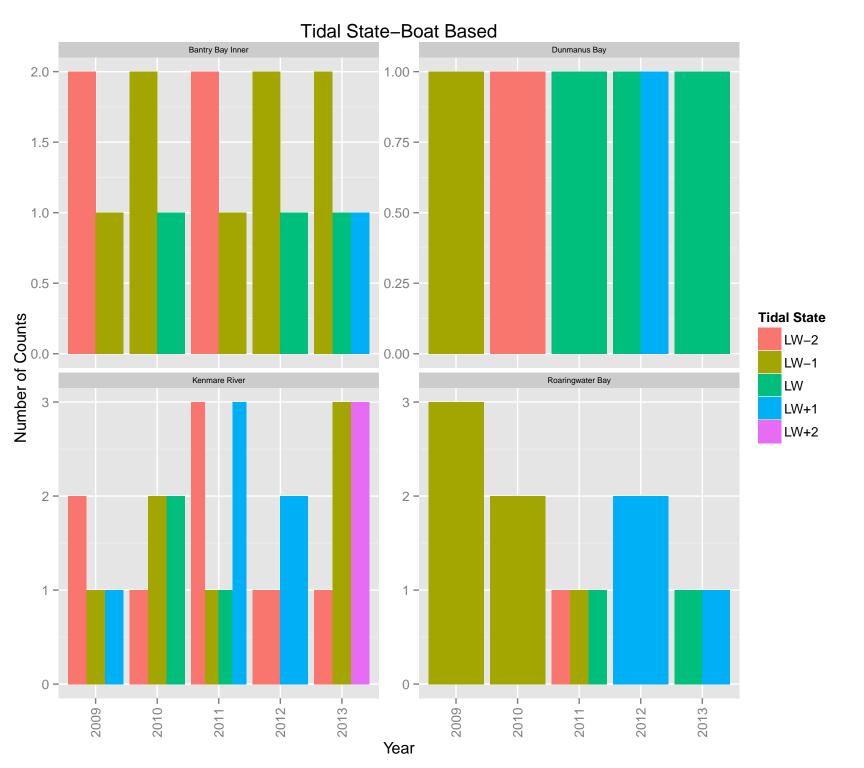


Year

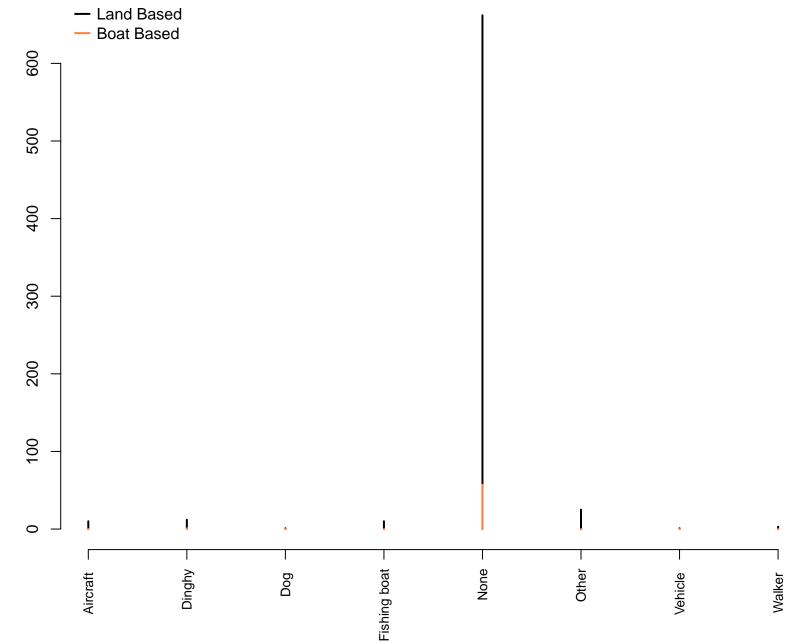
LW-1 LW LW+1 LW+2

LW-3

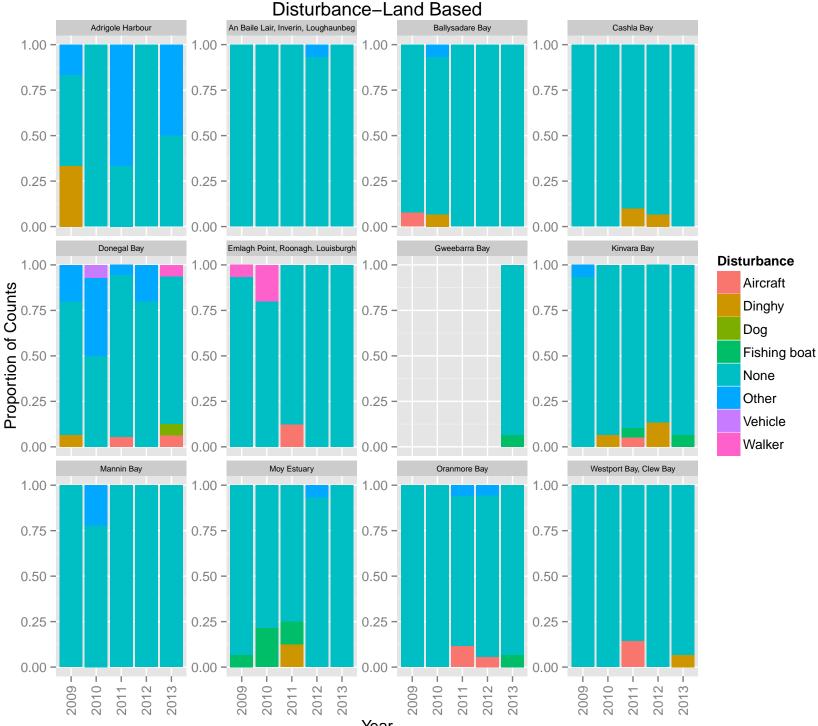
LW–2



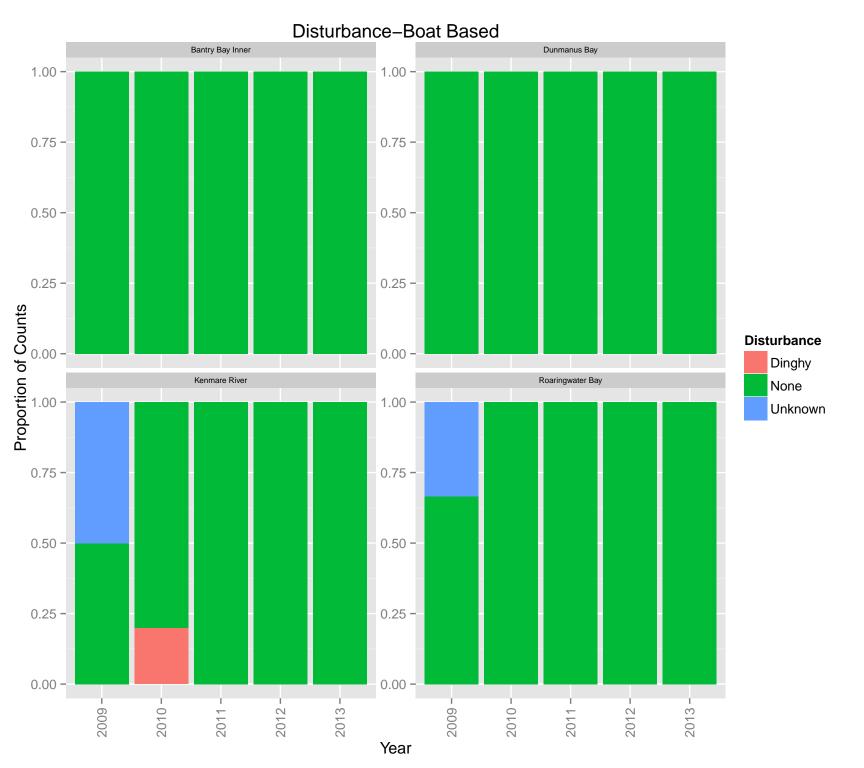
Disturbance



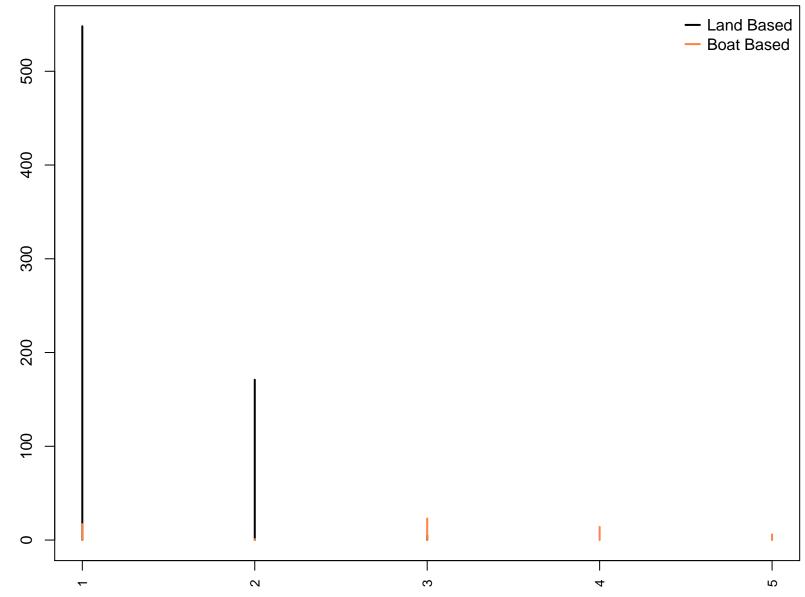
Number of Counts



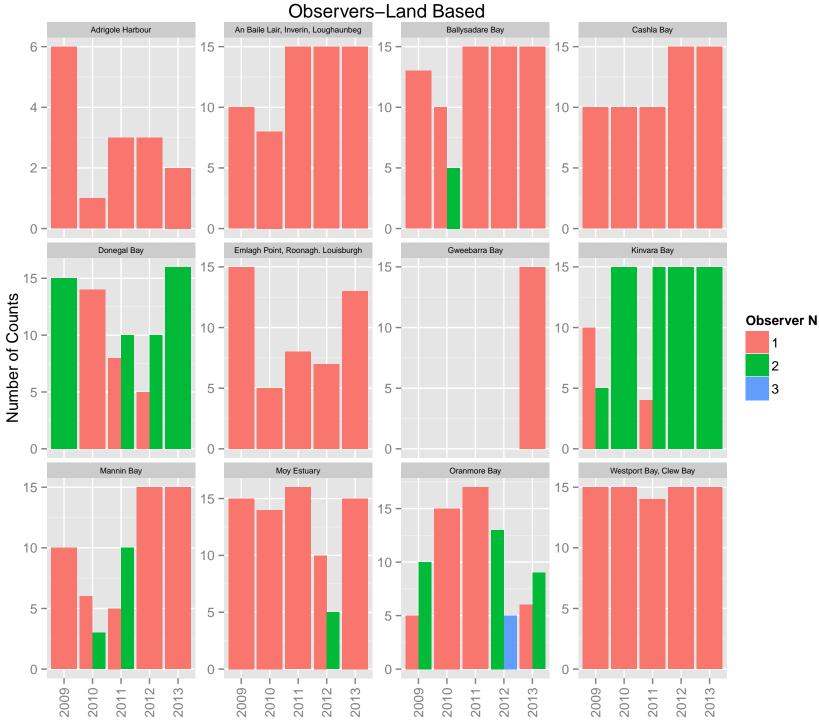
Year



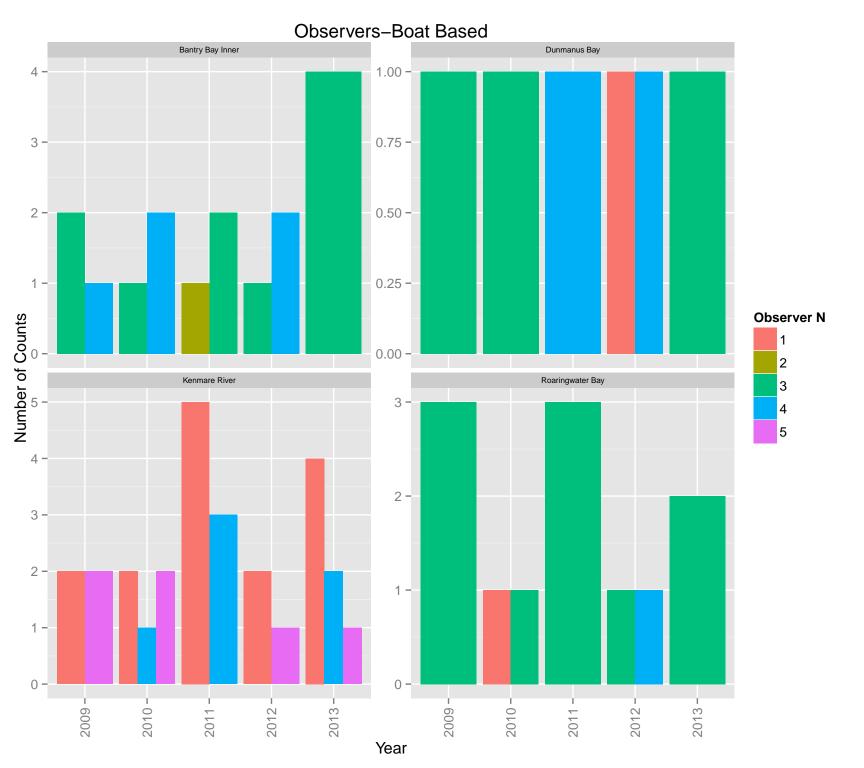
Number of Observers

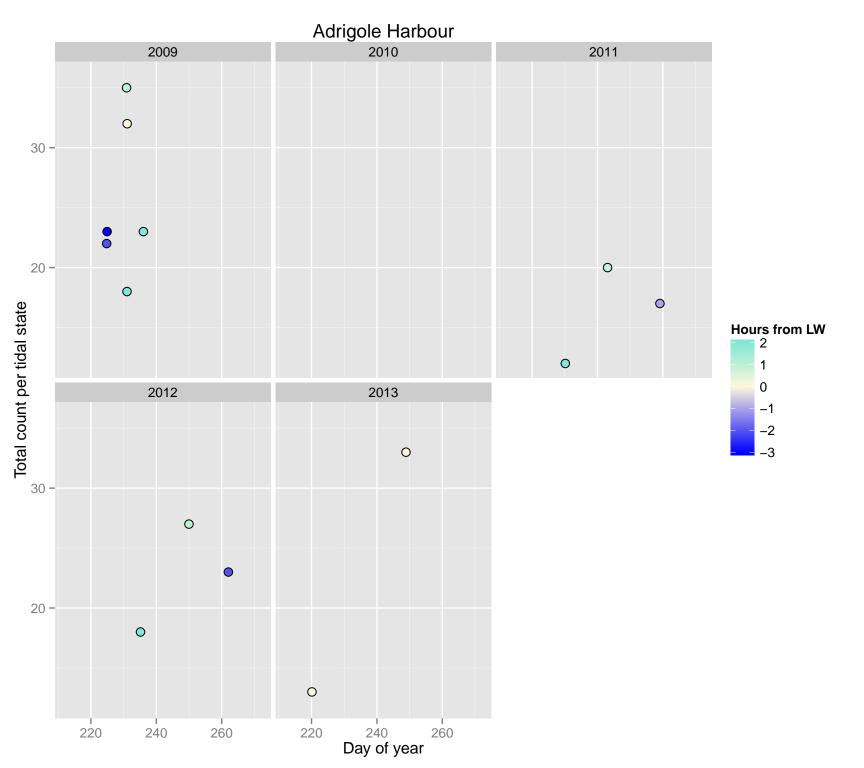


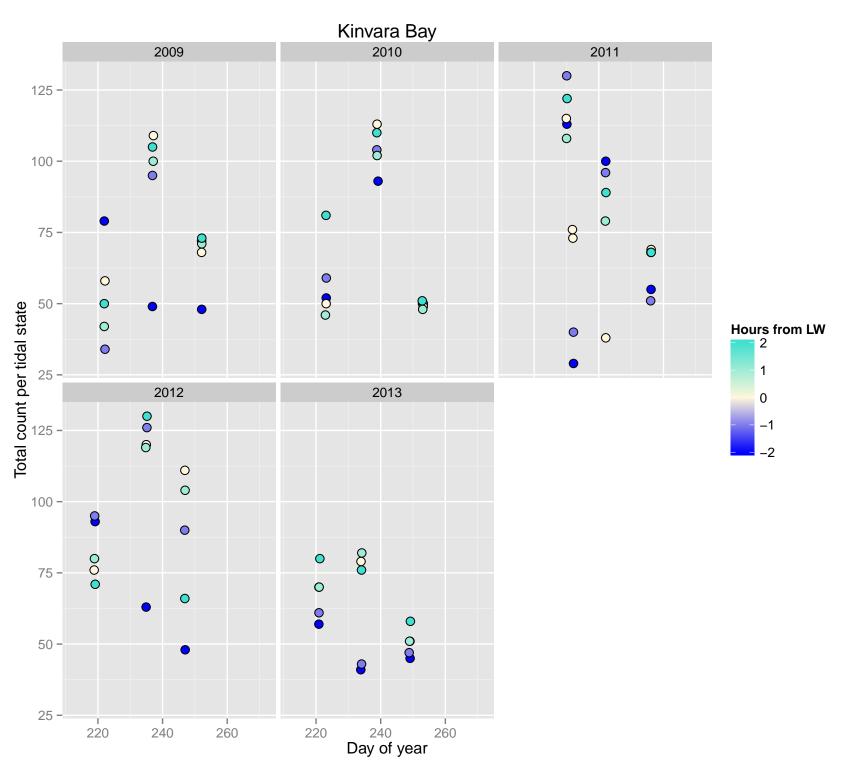
Number of Counts

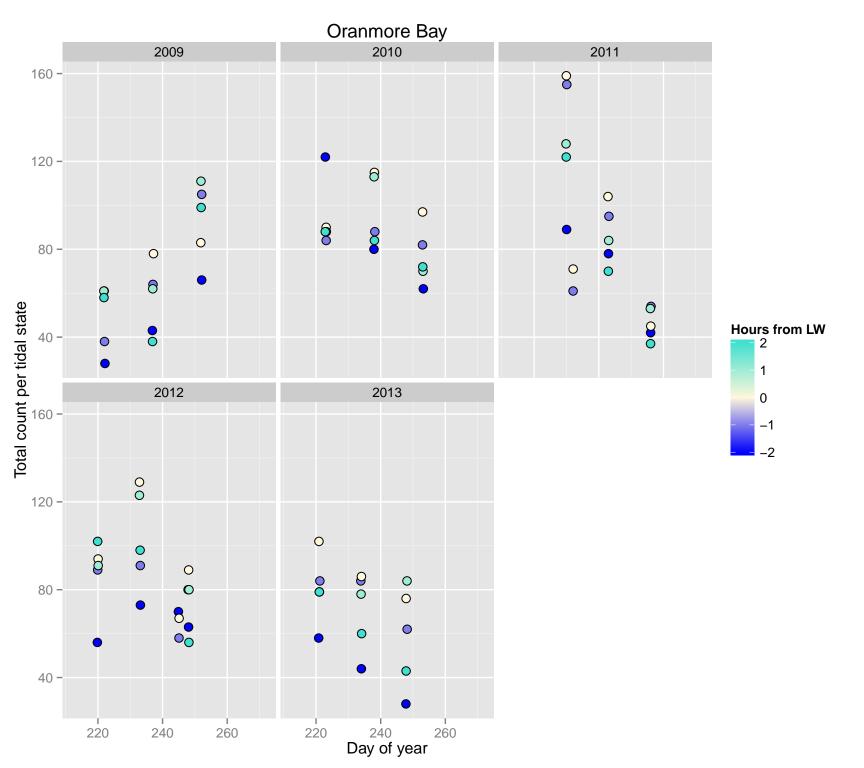


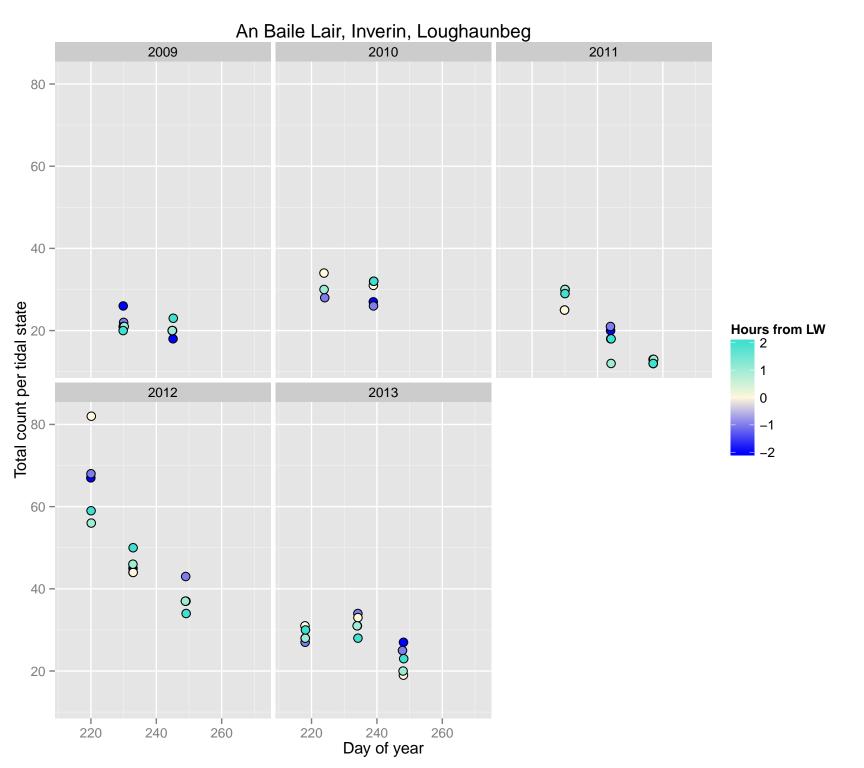
Year

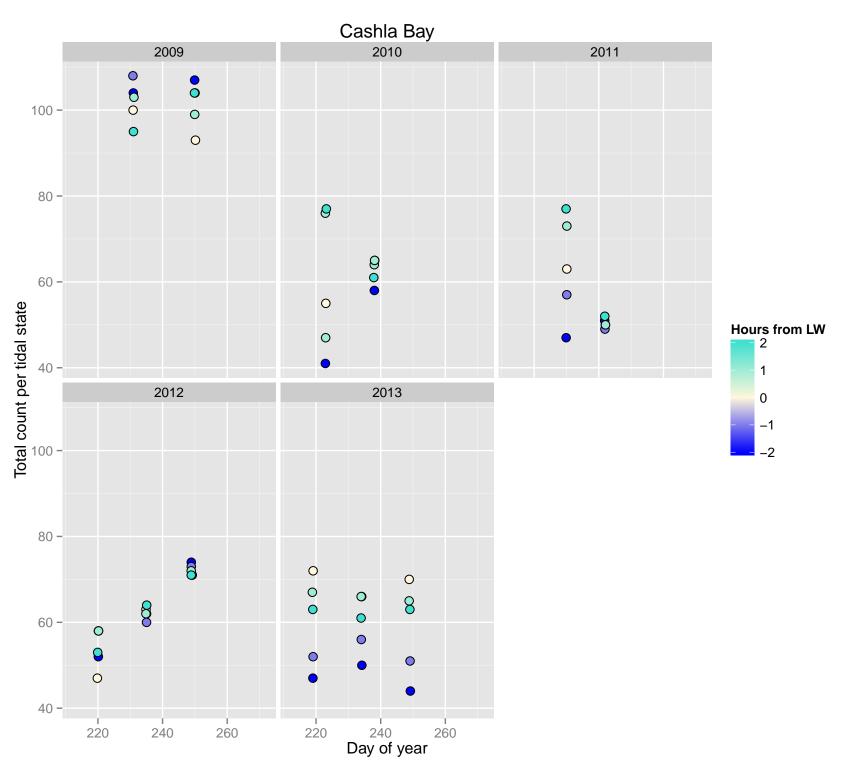


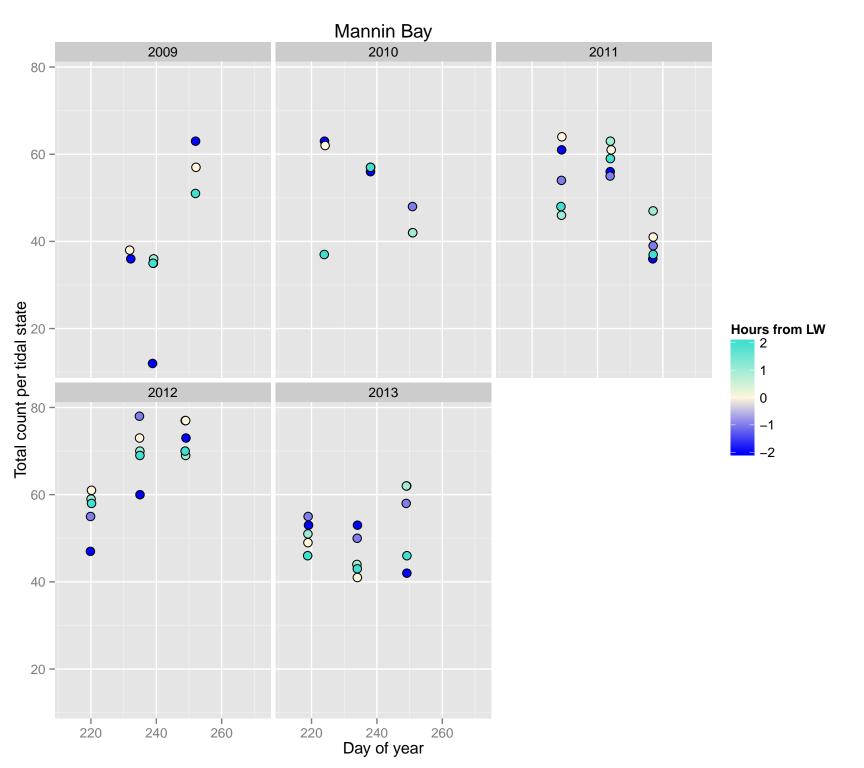


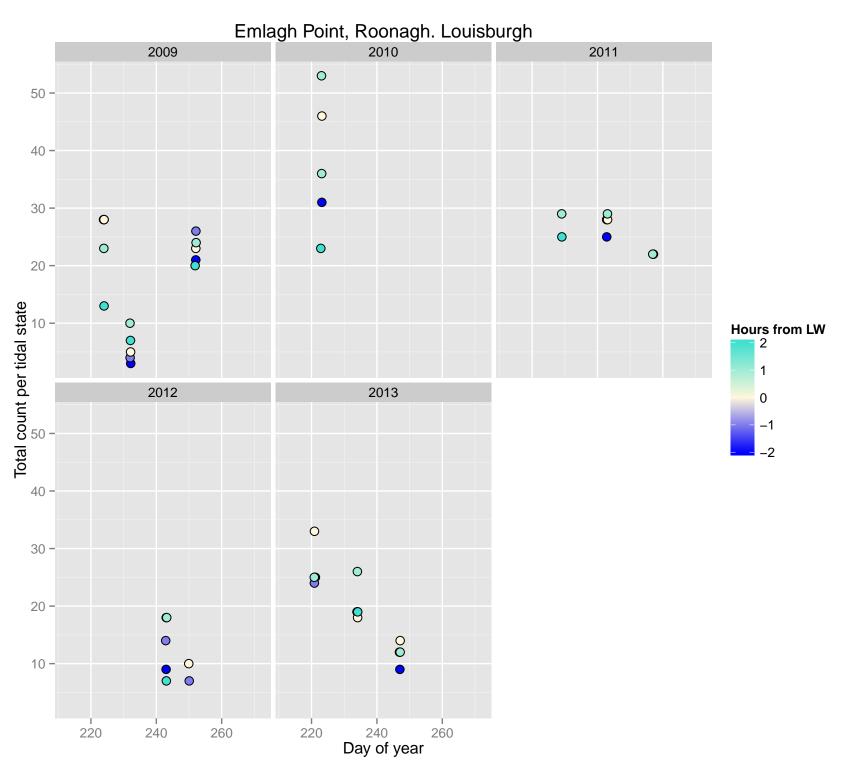


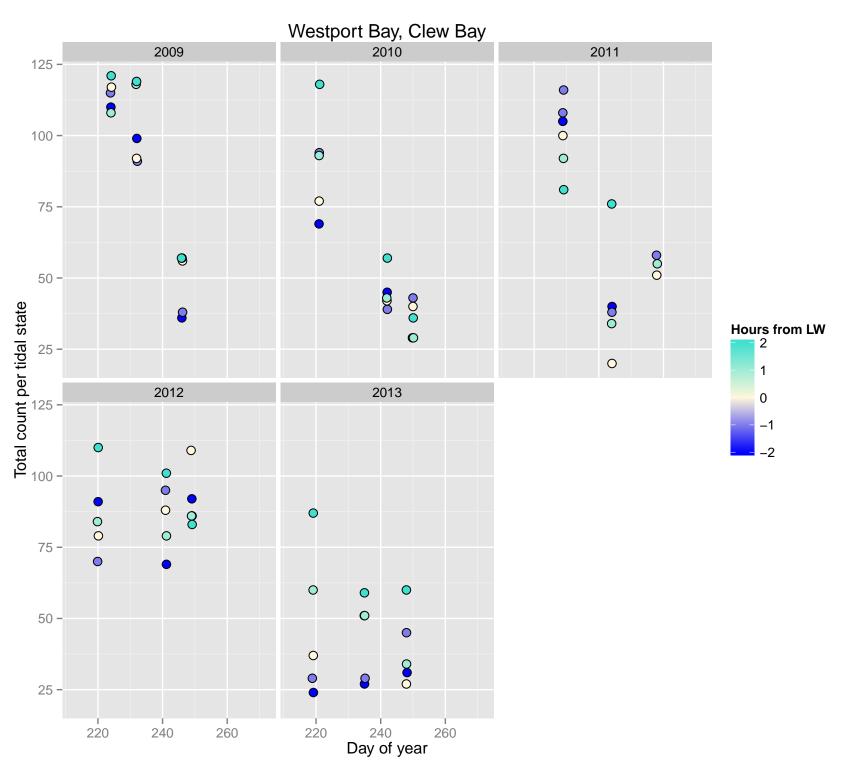


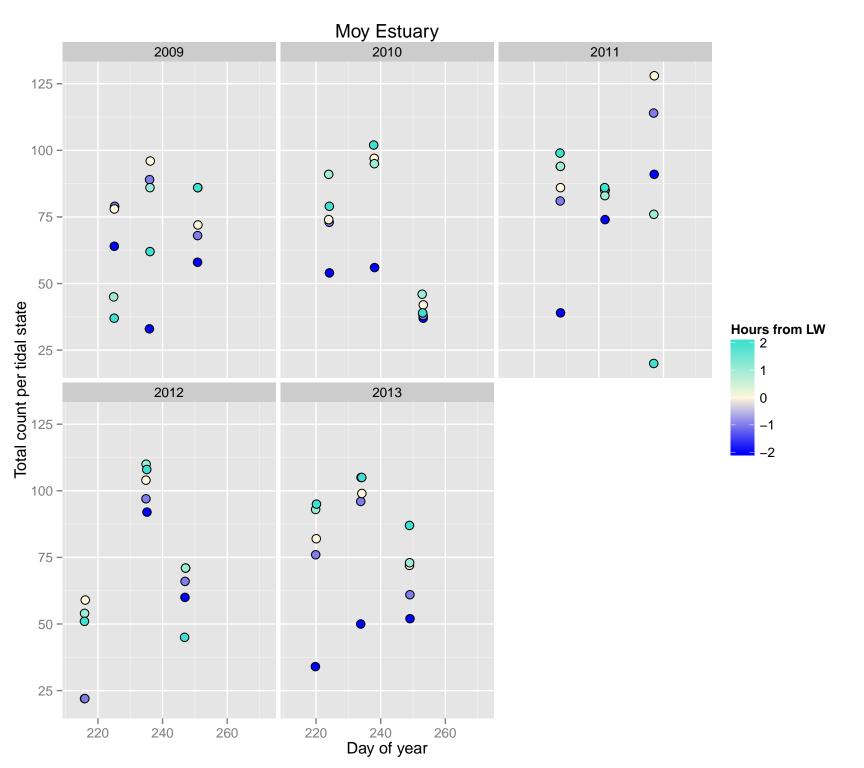


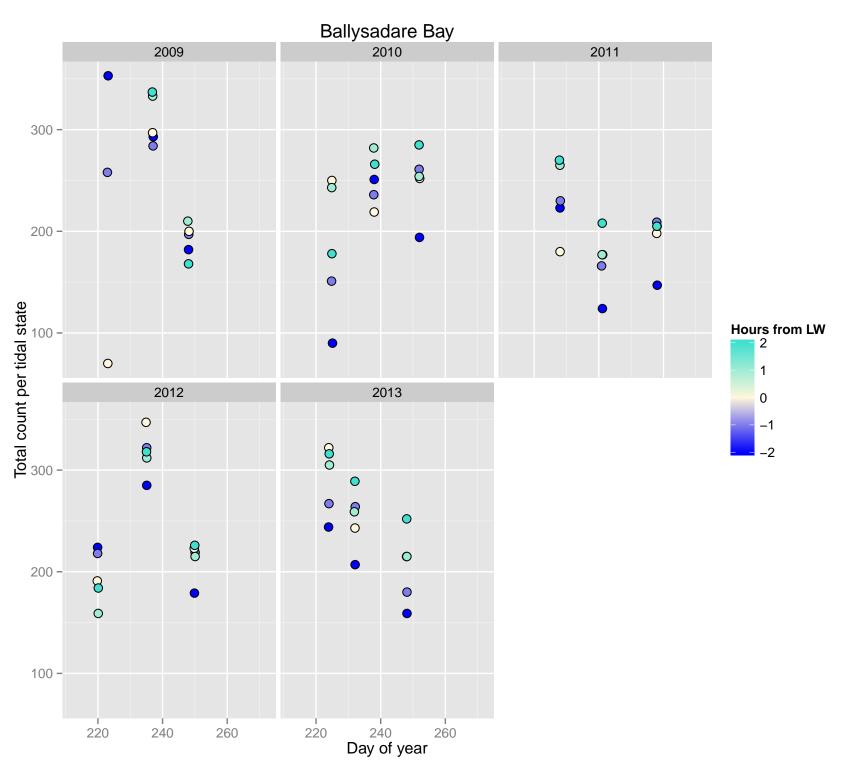


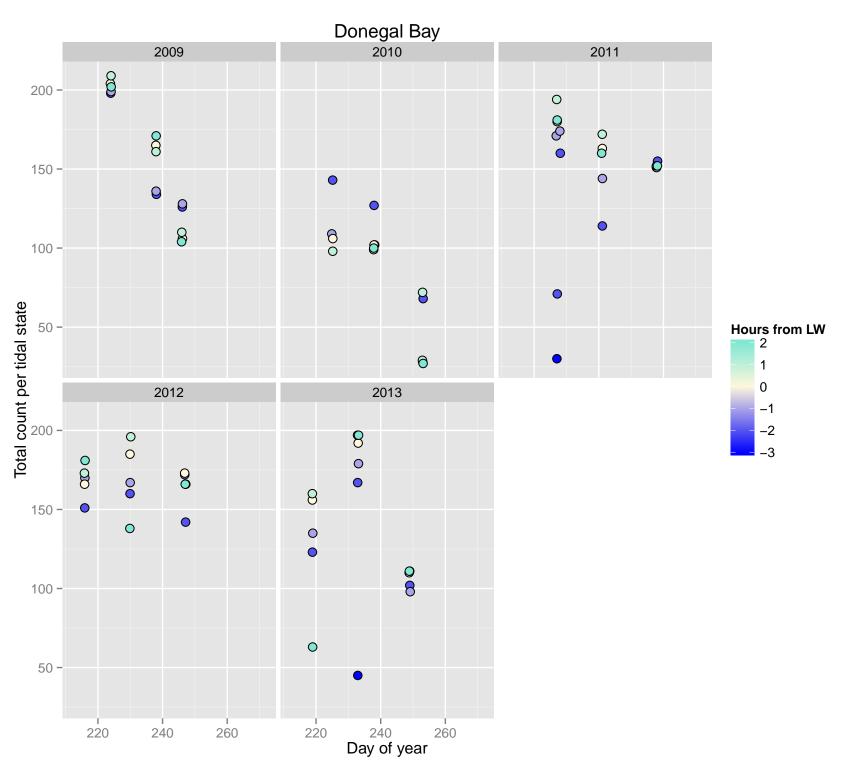


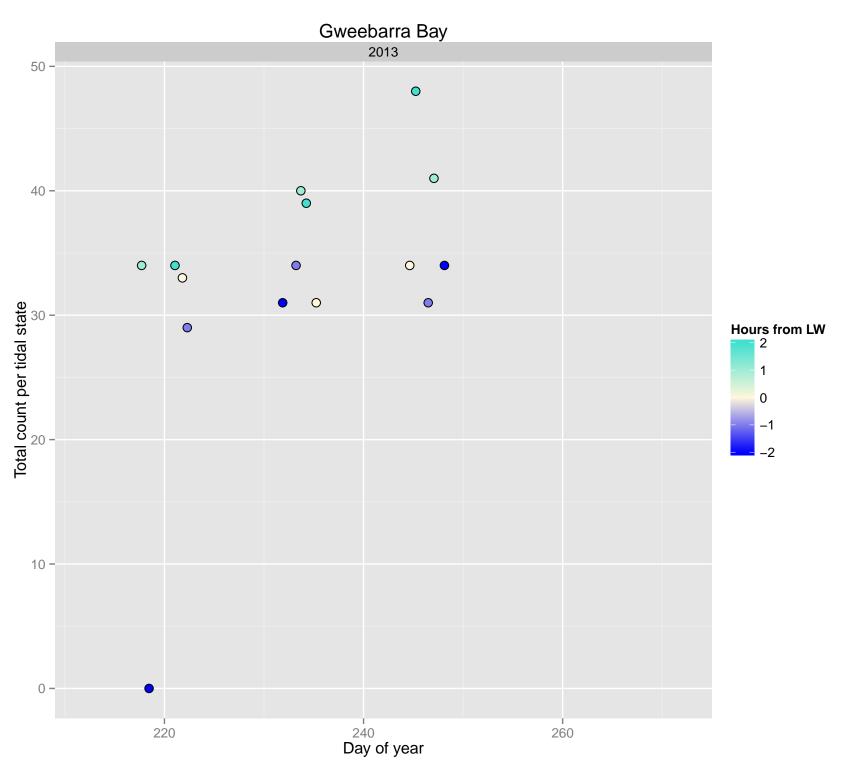


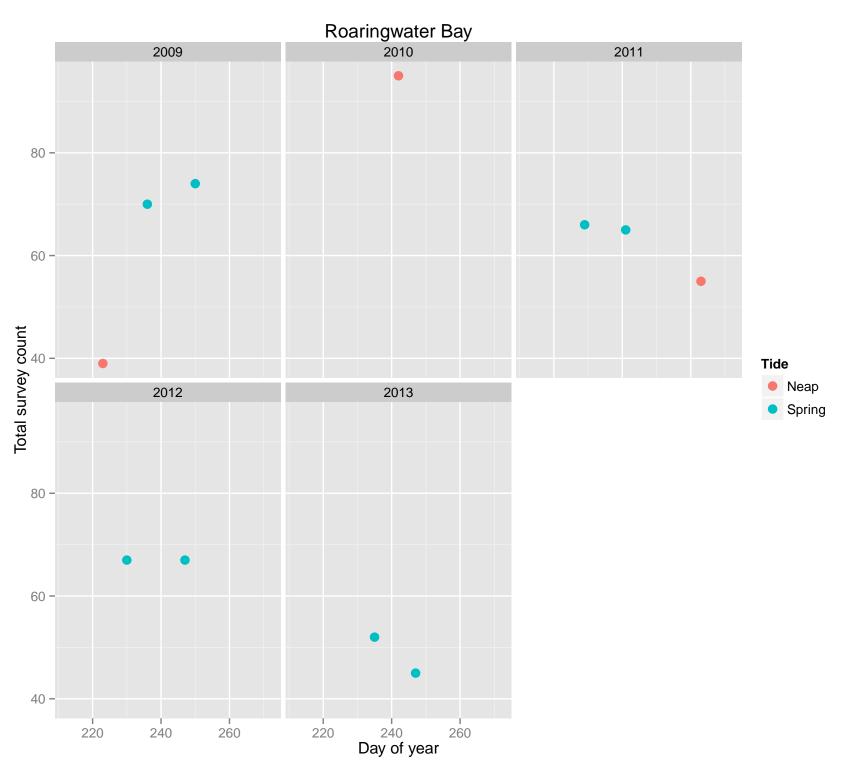


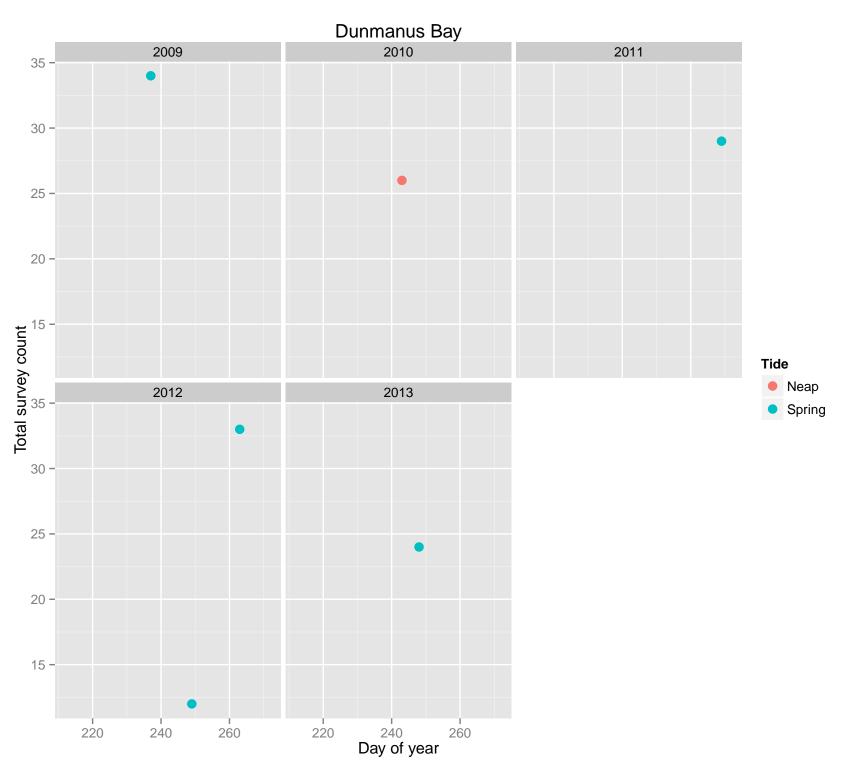


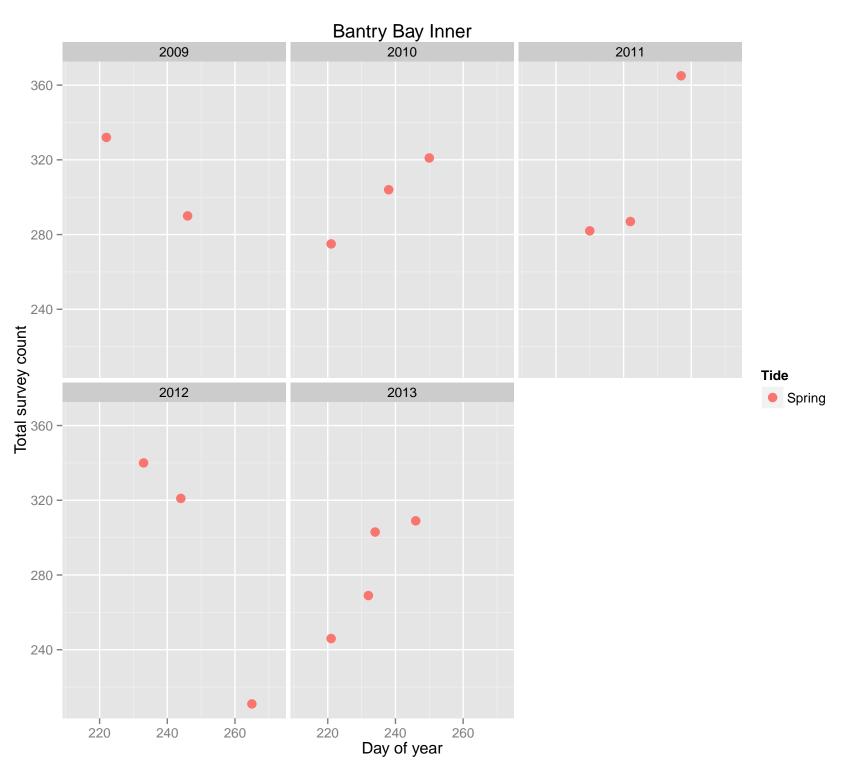


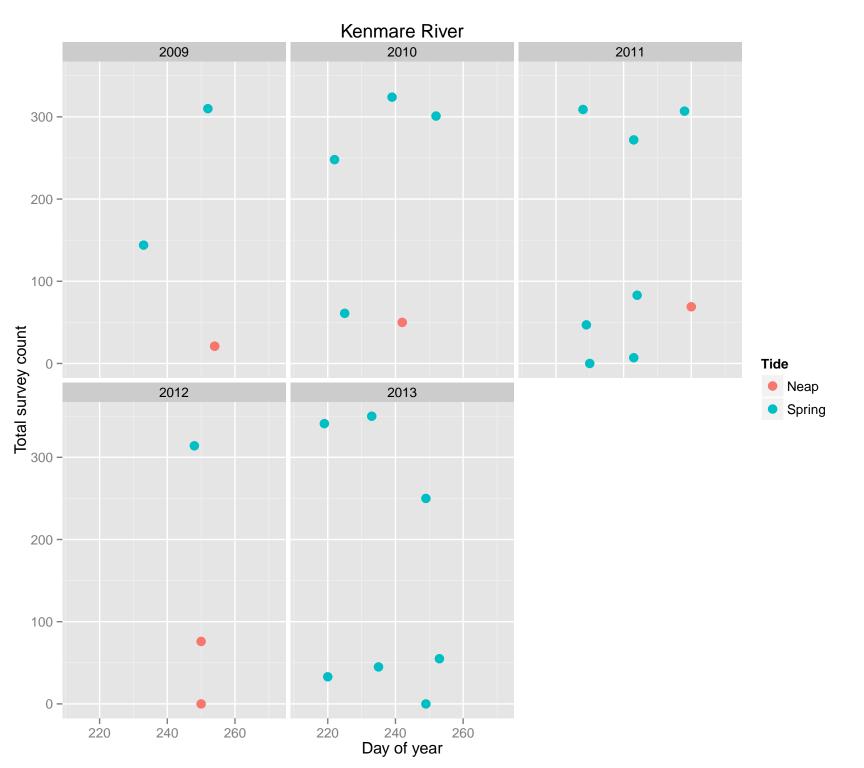


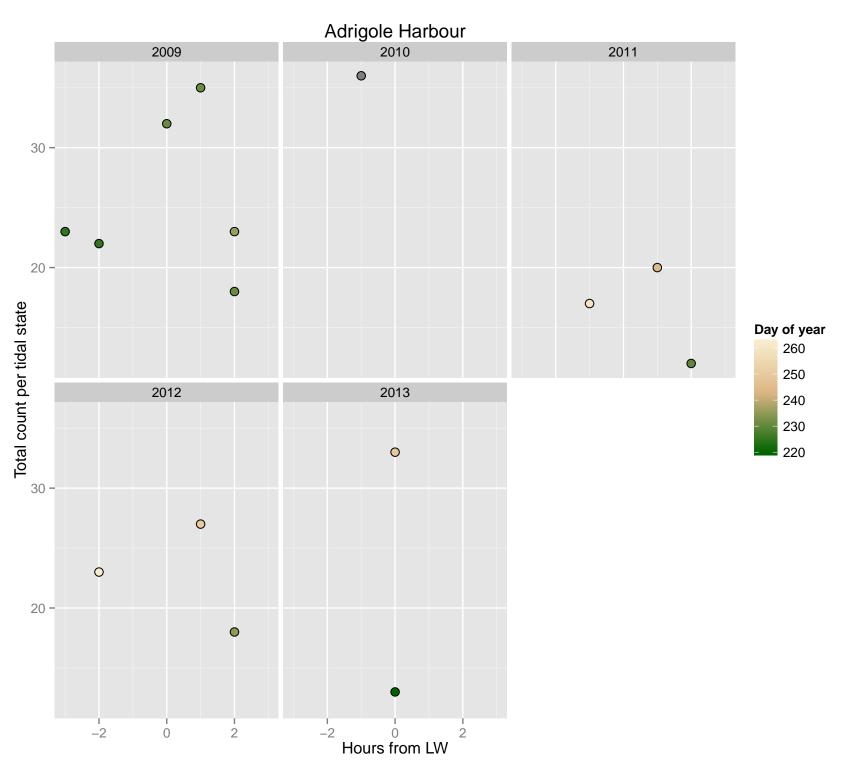


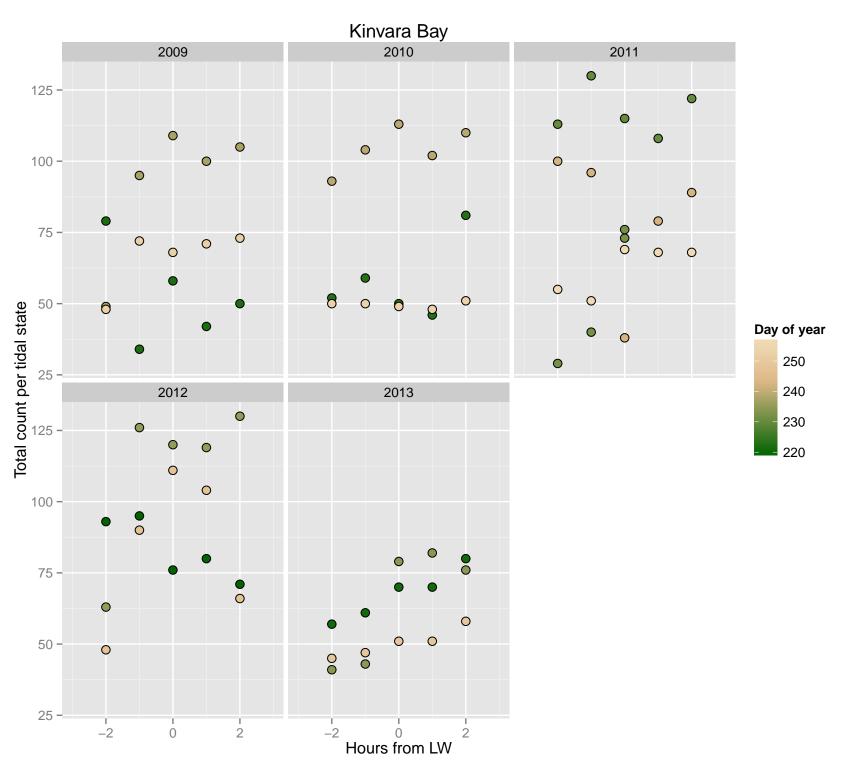


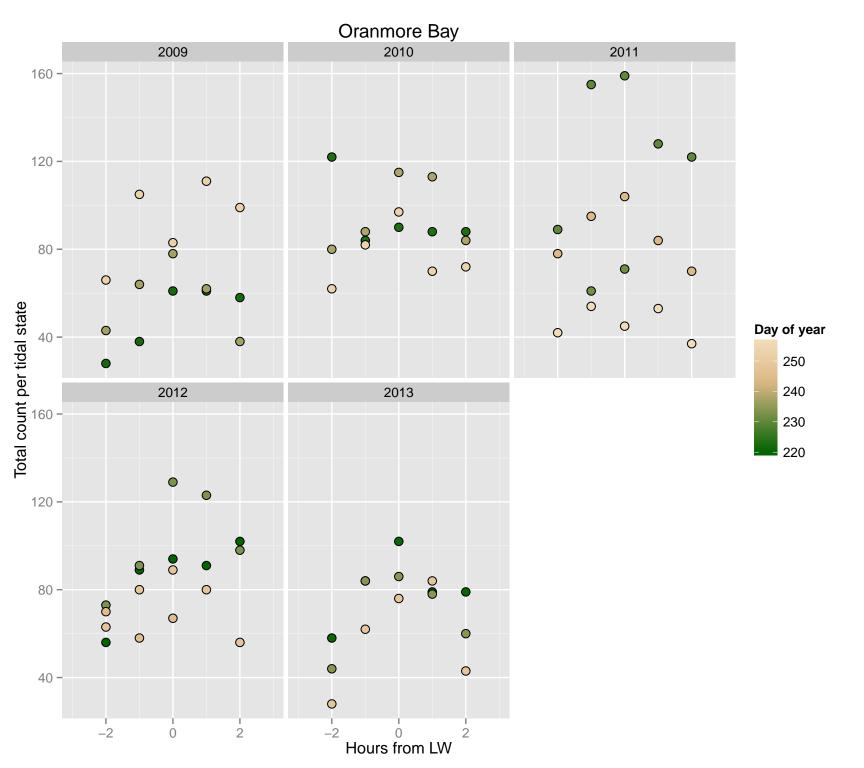


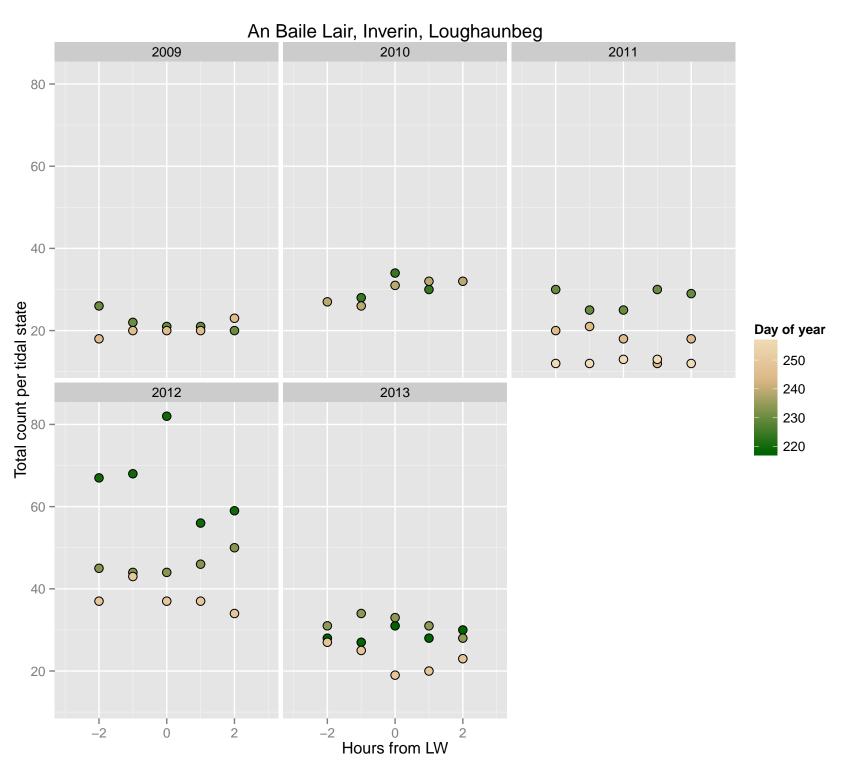


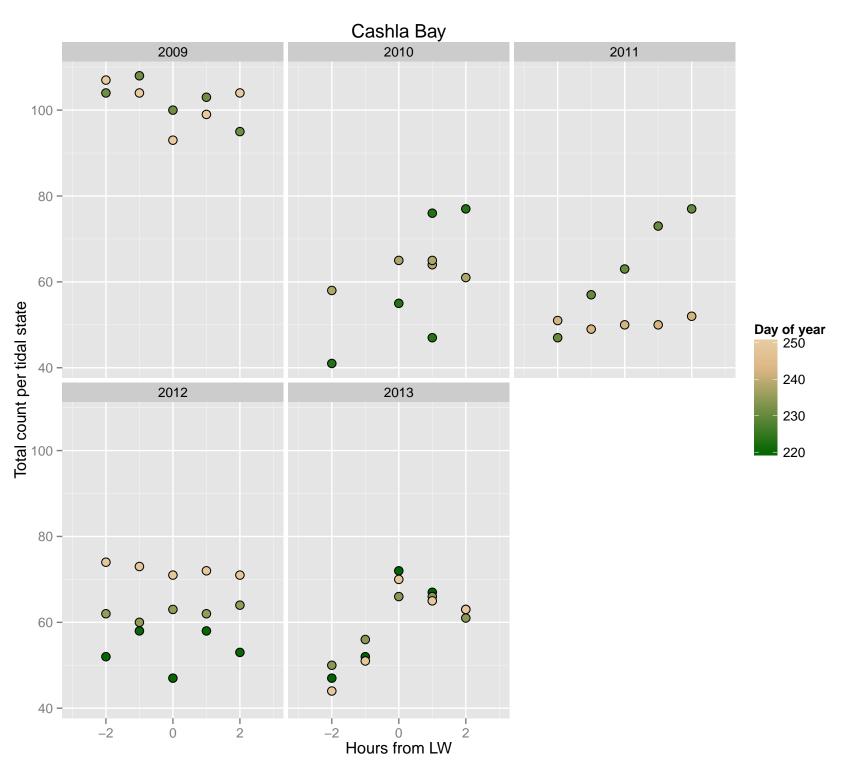


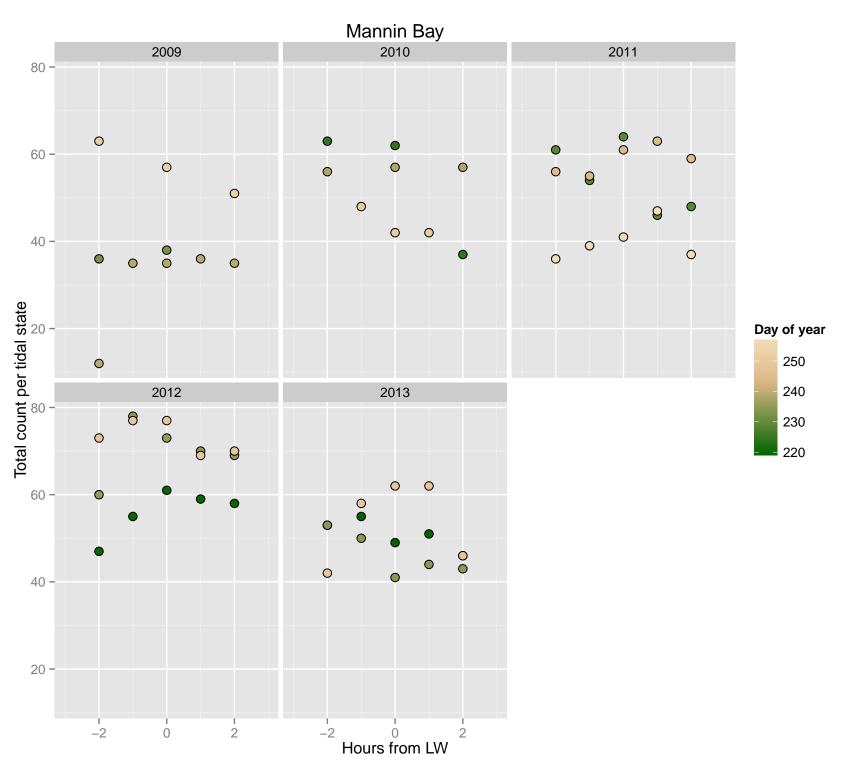


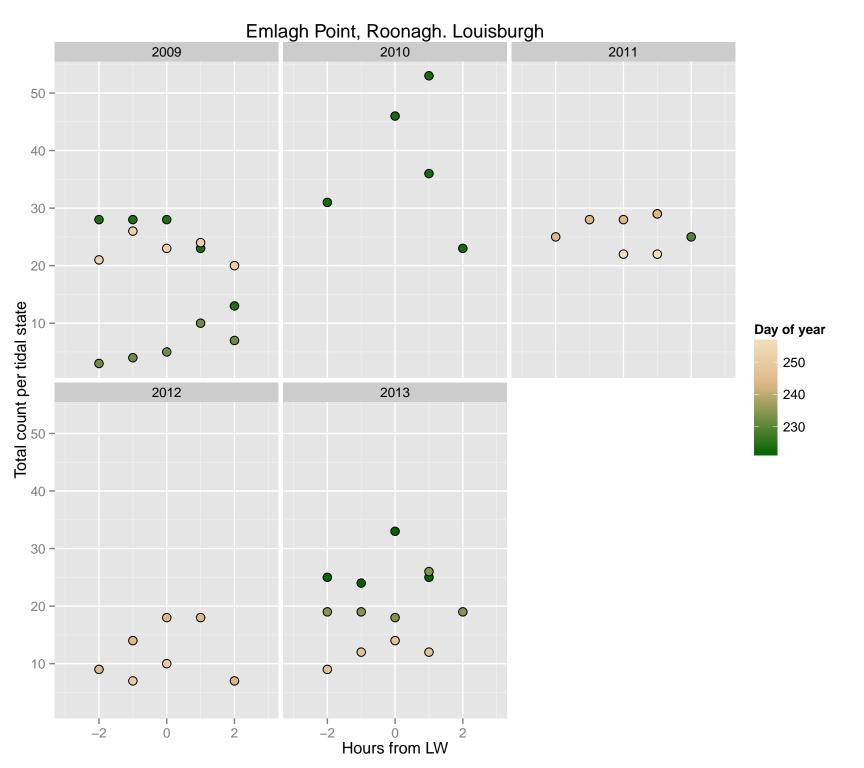


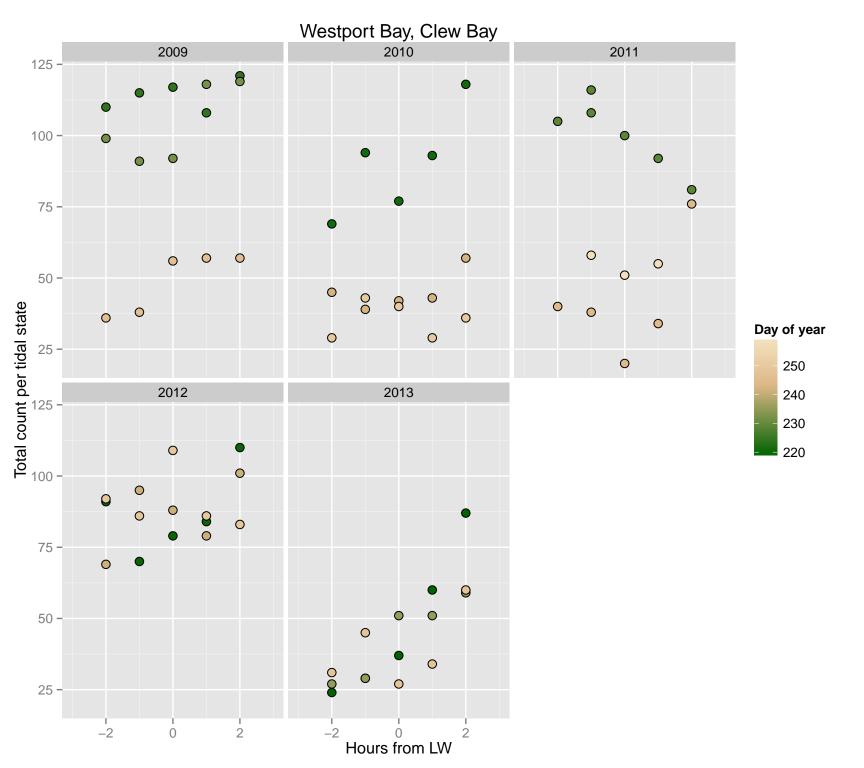


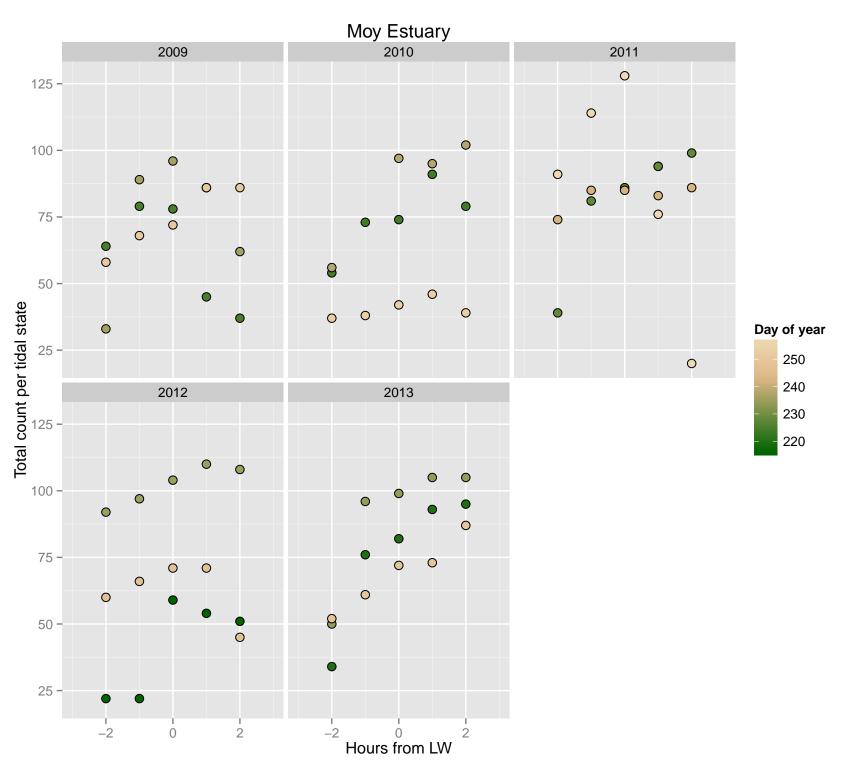


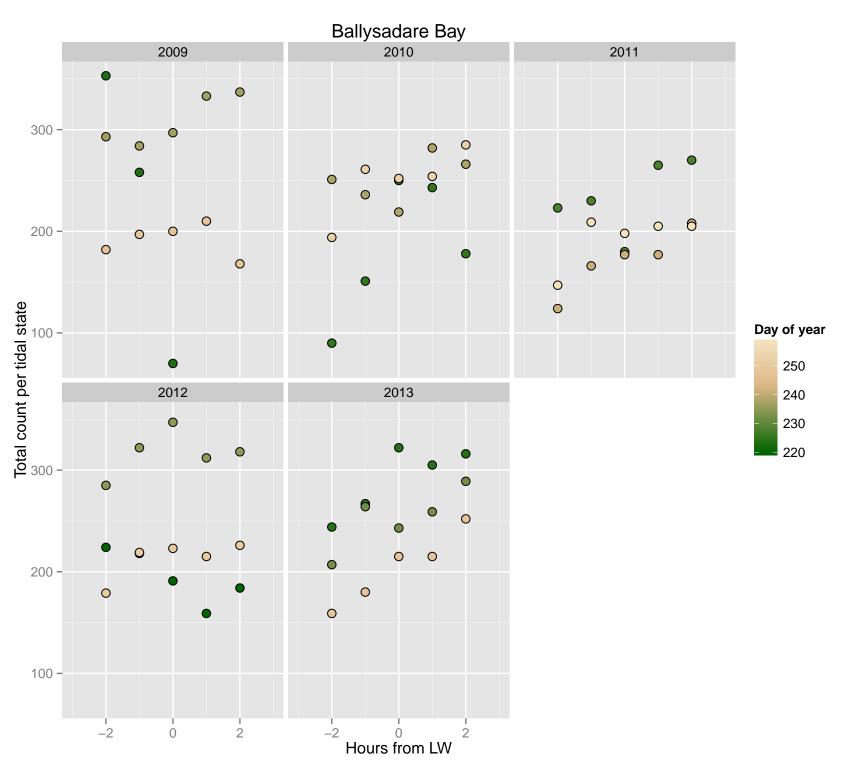


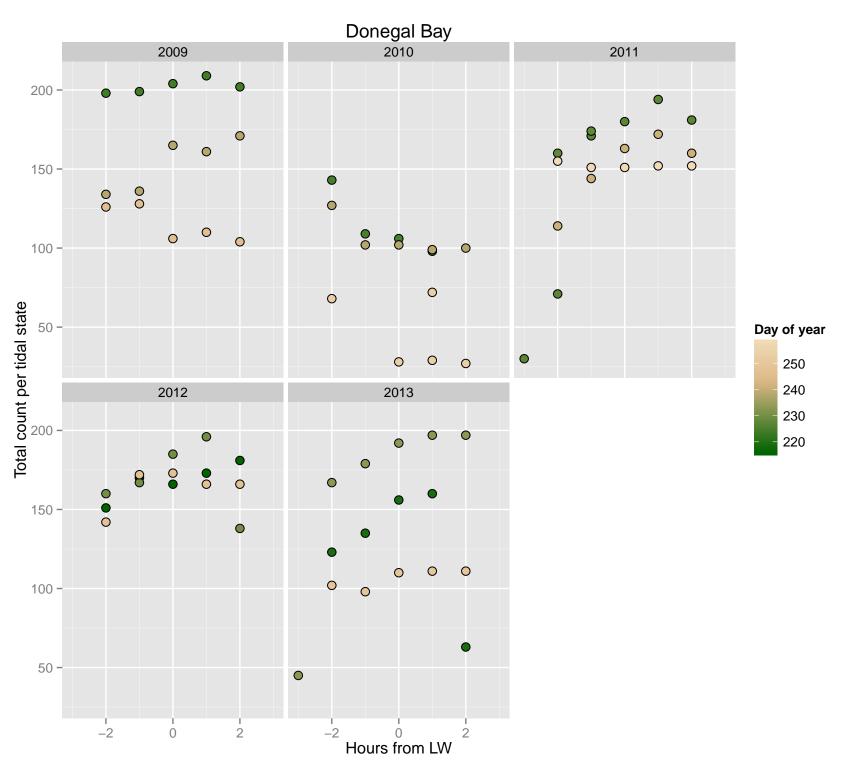


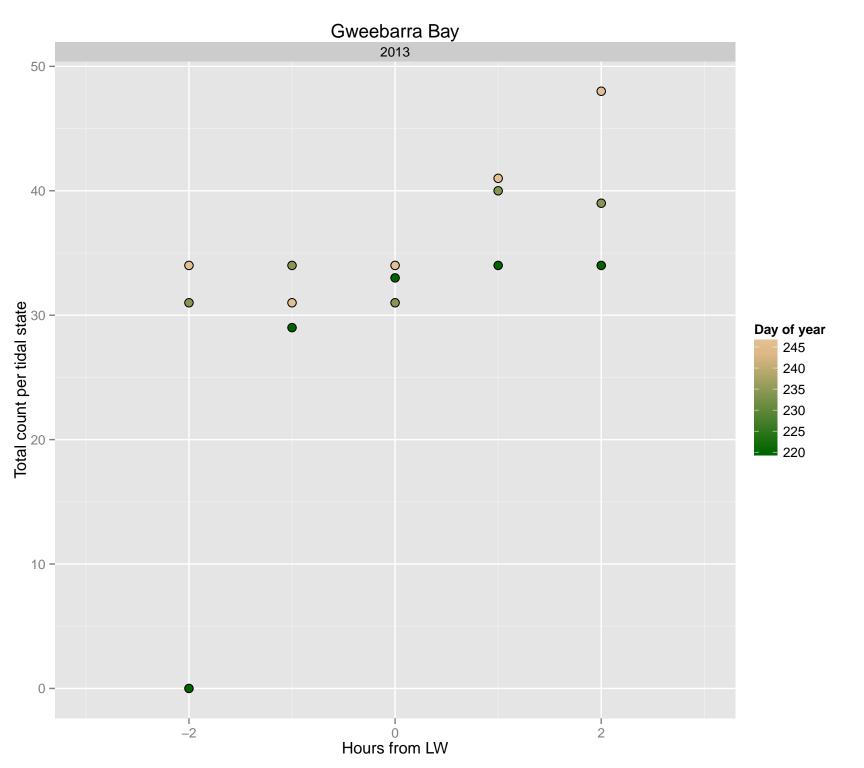


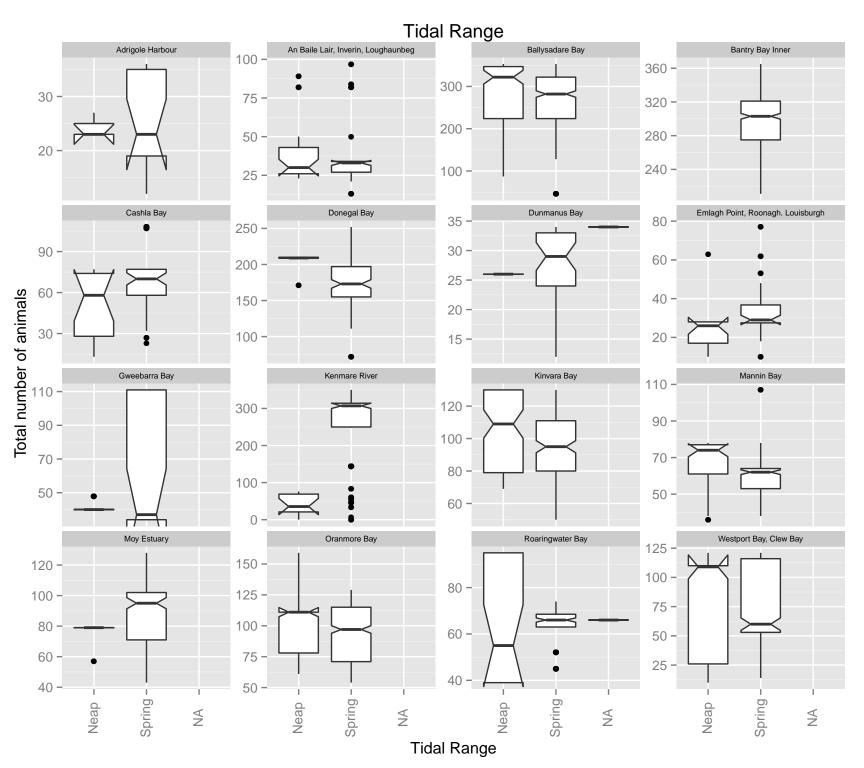


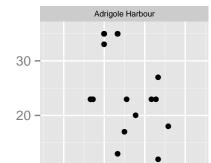


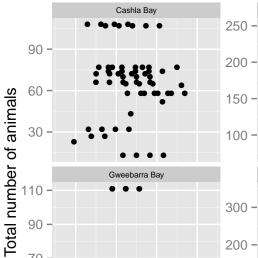


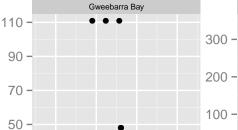


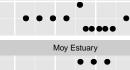


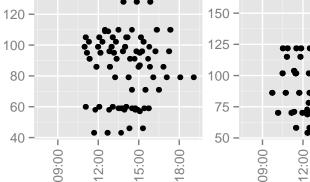




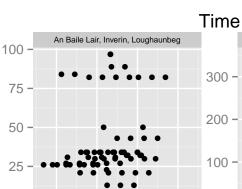


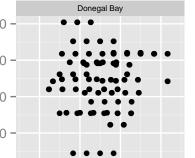






0 -



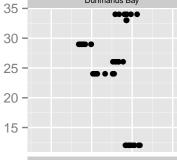


Kenmare River

Oranmore Bay

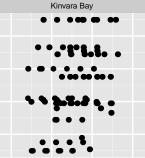
5:00

18:00

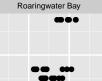


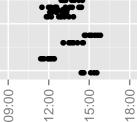
Ballysadare Bay

Dunmanus Bay

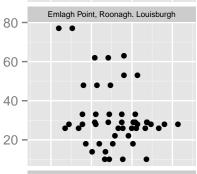


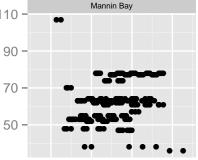
110

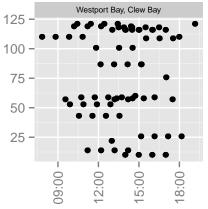












Time of count

120 -

100 -

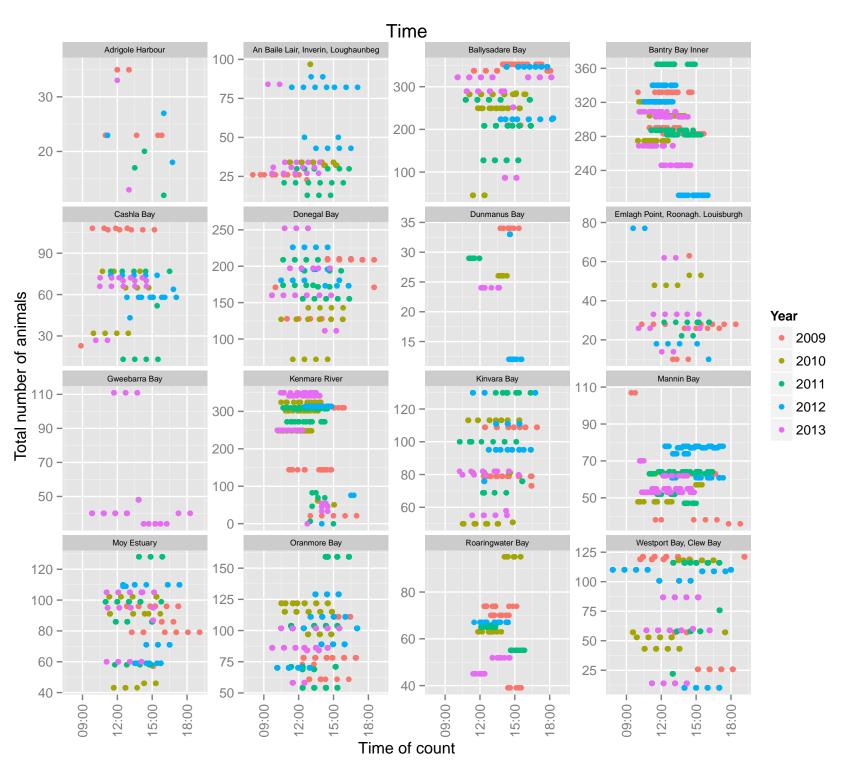
80 -

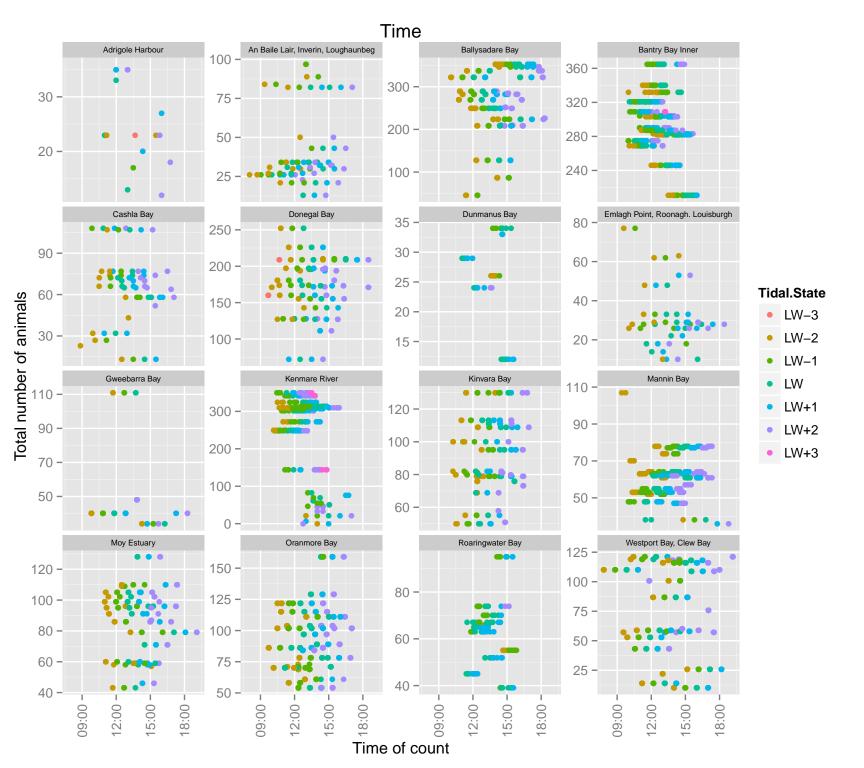
60 -

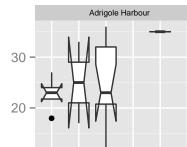
80 -

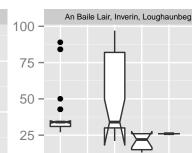
60 -

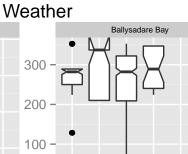
40 -

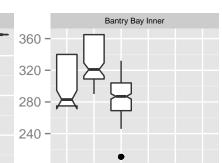


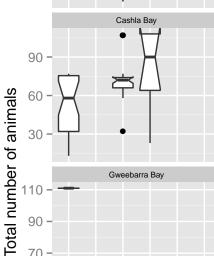












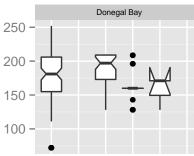
10

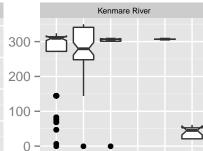
90

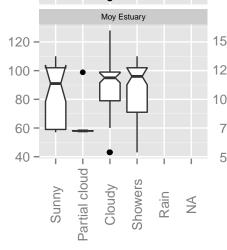
70

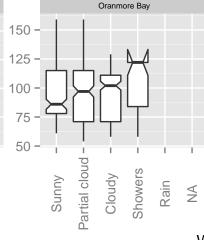
50 -

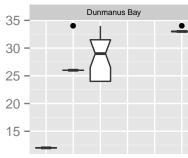
Gweebarra Bay

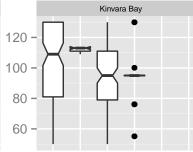












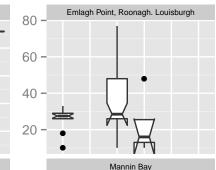
Roaringwater Bay

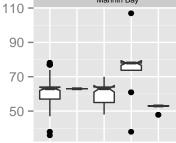
Showers

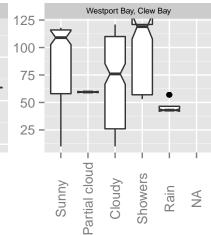
Rain

ΨN

Cloudy







Weather

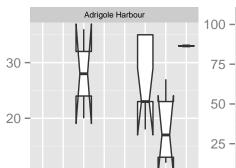
80 -

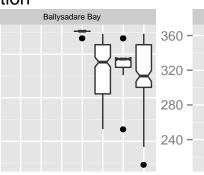
60 -

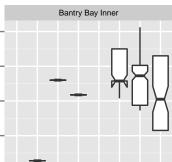
40 -

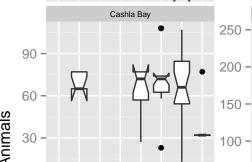
Partial cloud

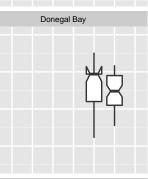
Sunny

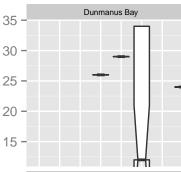


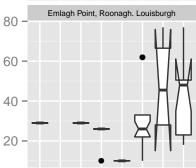












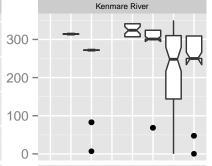
120 -

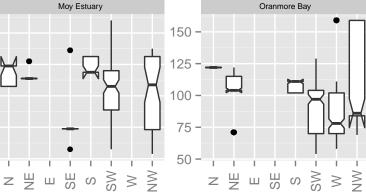
100 -

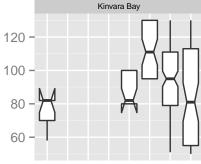
80

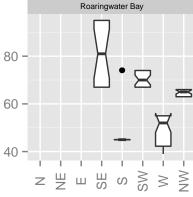
60 -

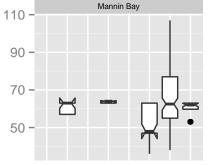
40 -

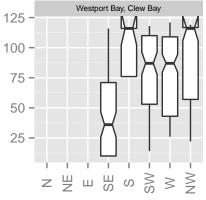






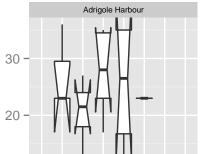


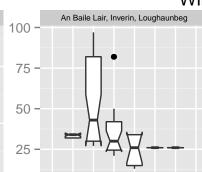


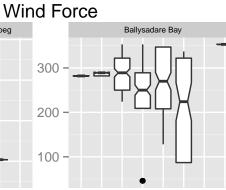


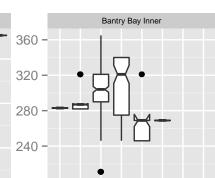
Wind Direction

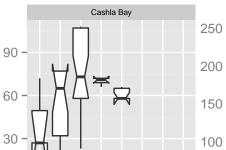
Wind Direction











Gweebarra Bay

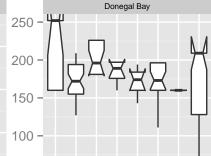
Number of Animals

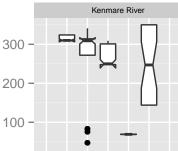
110 -

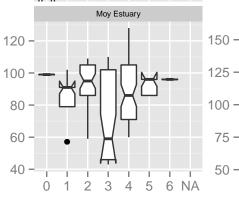
90 -

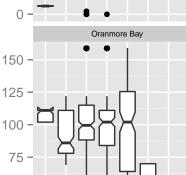
70 -

50 -









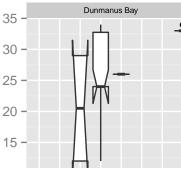
1 1

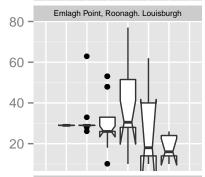
3 4 5

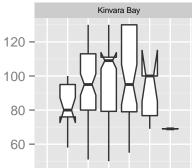
2

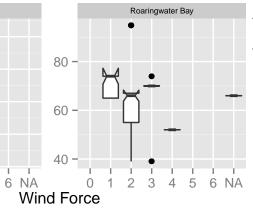
1

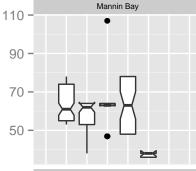
0

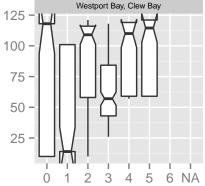


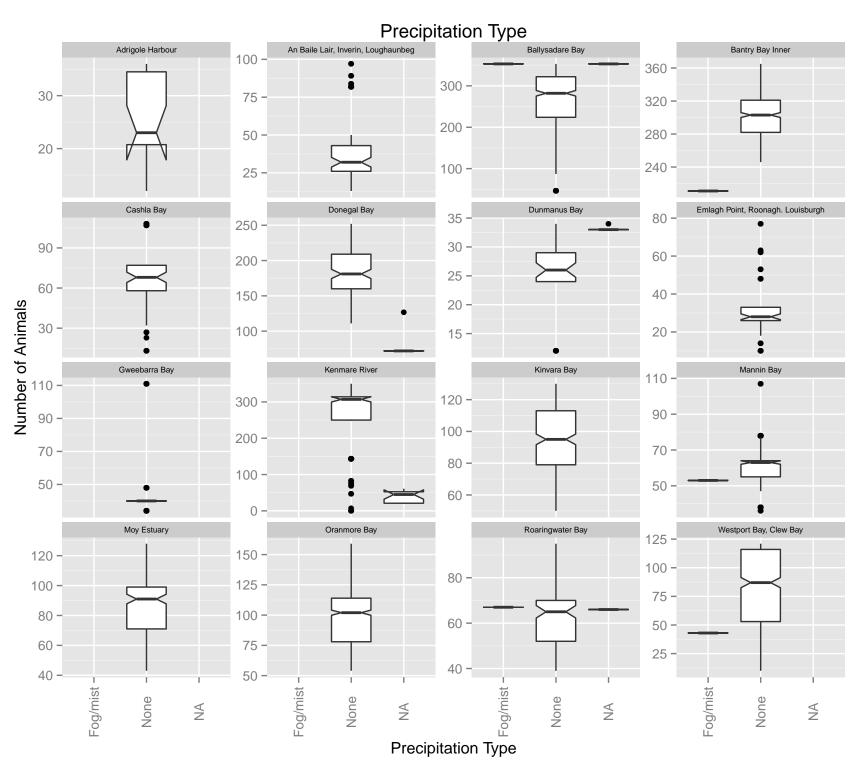


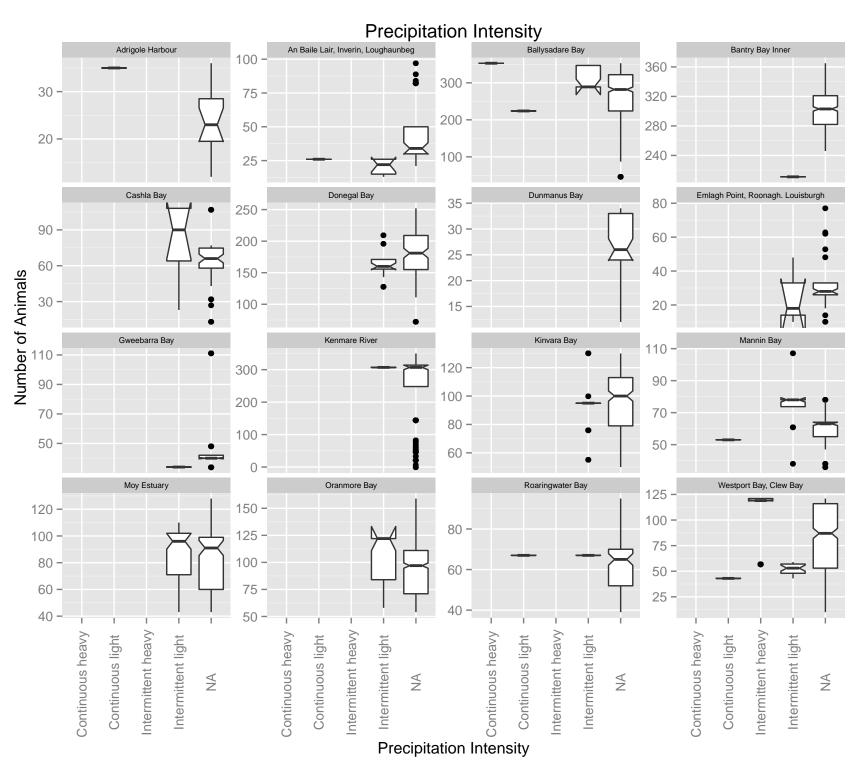


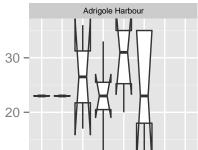


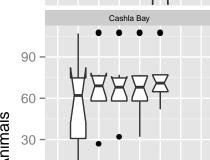


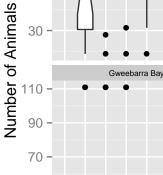


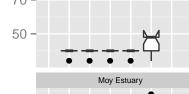


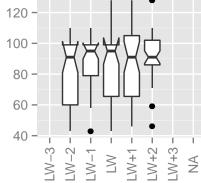


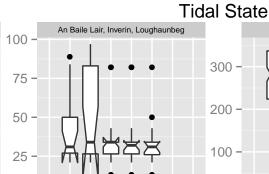


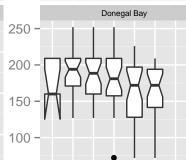


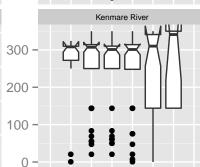


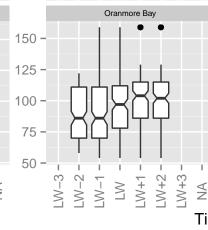


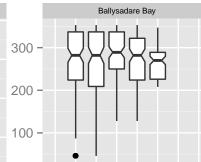


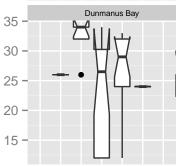


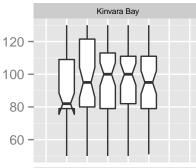


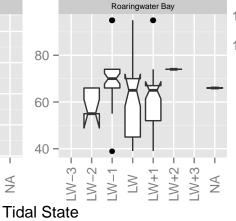


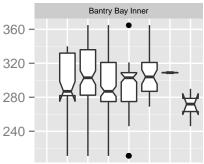


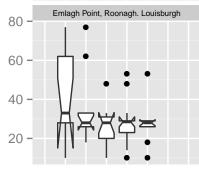


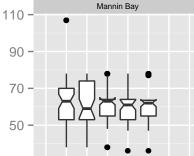


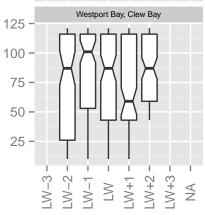


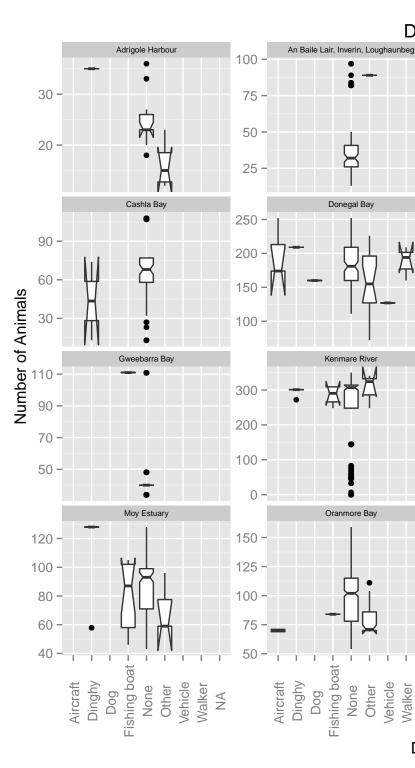


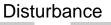


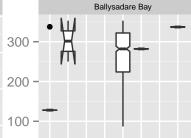


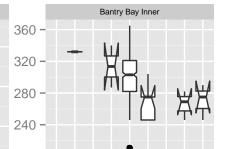


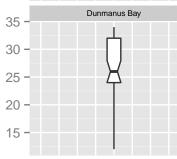


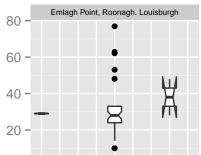


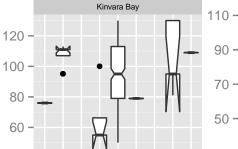


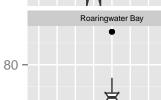


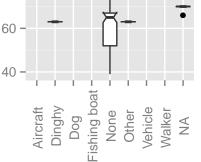


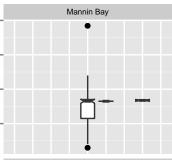


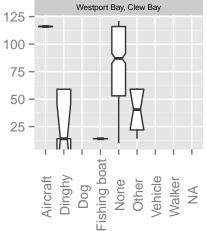












 \succeq

凶

Vehicle

None Other Walker

ΑN

Disturbance

Observer Number 100 An Baile Lair, Inverin, Loughaunbeg Ballysadare Bay 100 • 360

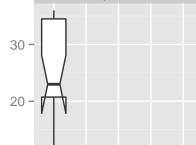
35

30 -

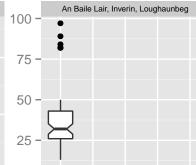
25 -

20 -

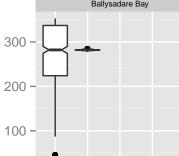
15 -



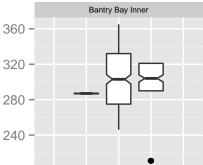
Adrigole Harbour

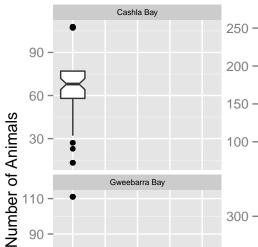


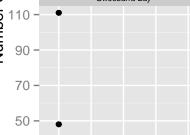
Donegal Bay

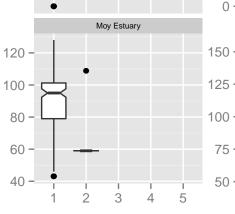


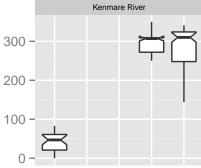
Dunmanus Bay

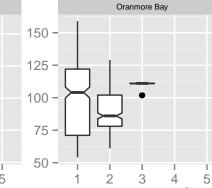


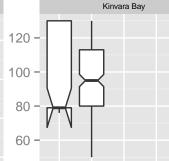


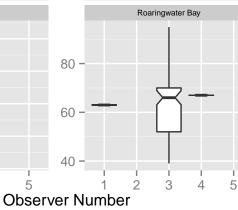


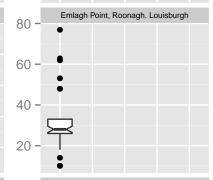


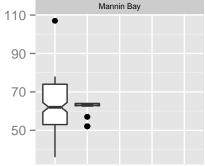


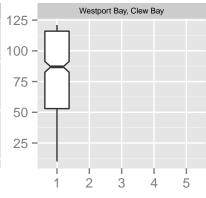








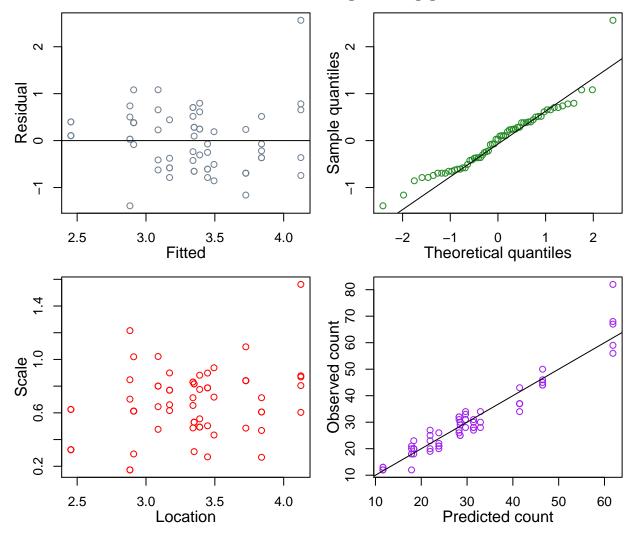




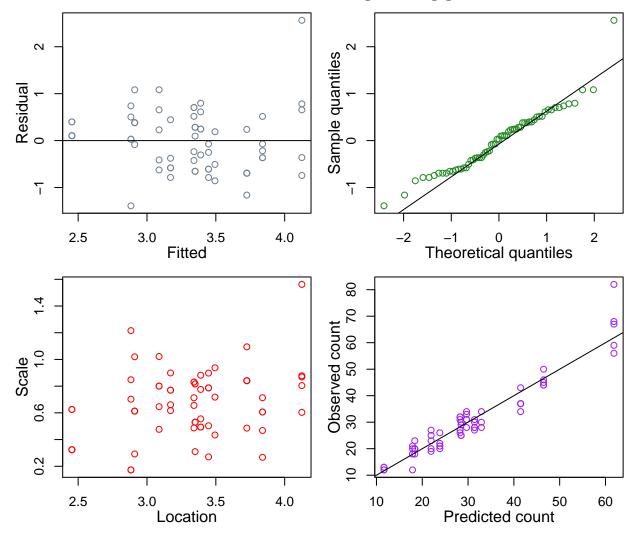
Appendix 2

By-location GLM, GAM and GLMM fit diagnostic plots. The estimated dispersion parameter for GLM and GAM fits is displayed in the title.

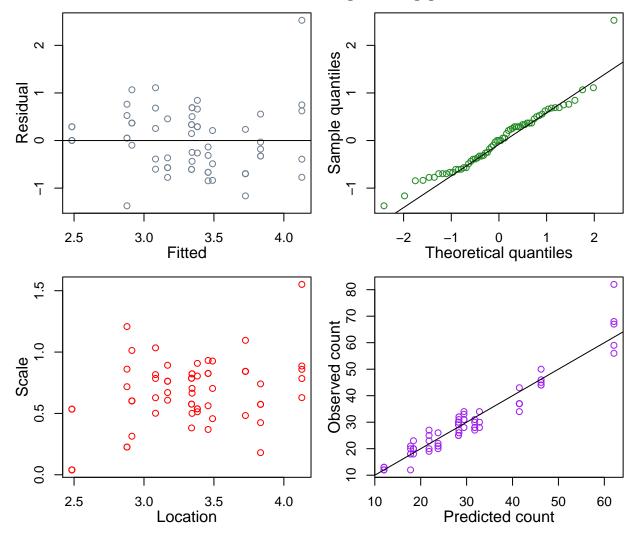
An Baile Lair, Inverin, Loughaunbeg glm 0.416



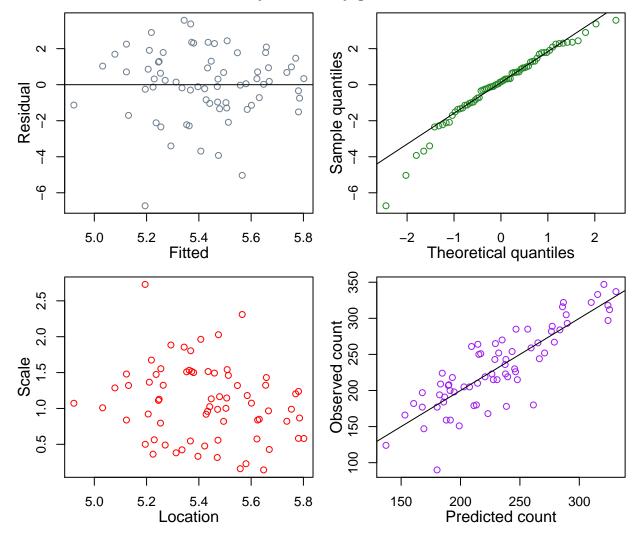
An Baile Lair, Inverin, Loughaunbeg glmm



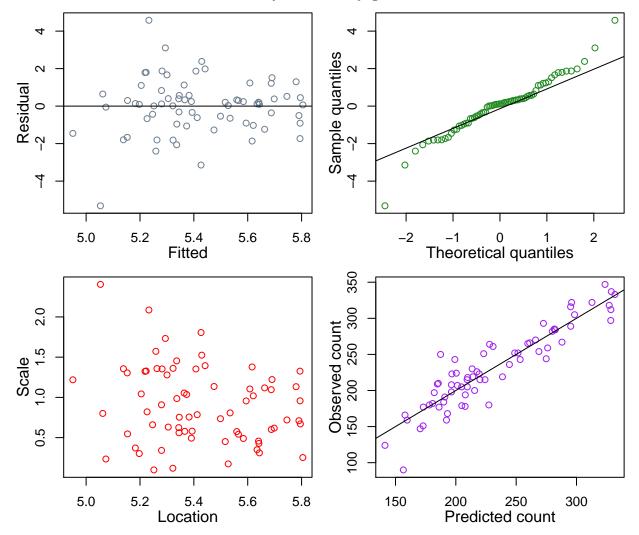
An Baile Lair, Inverin, Loughaunbeg gam 0.416



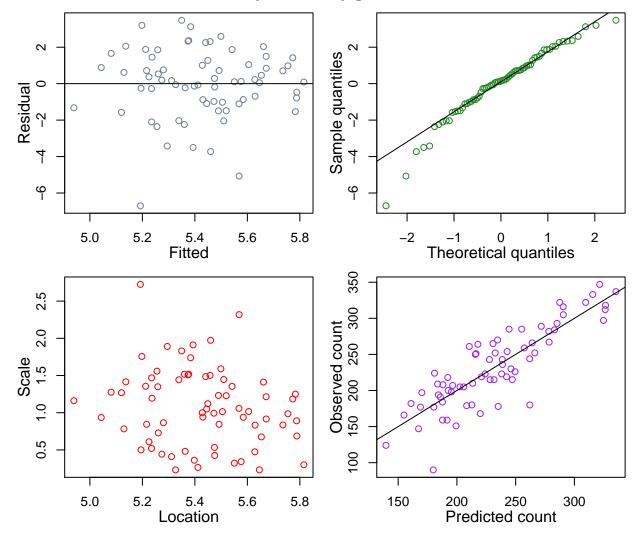
Ballysadare Bay glm 3.627



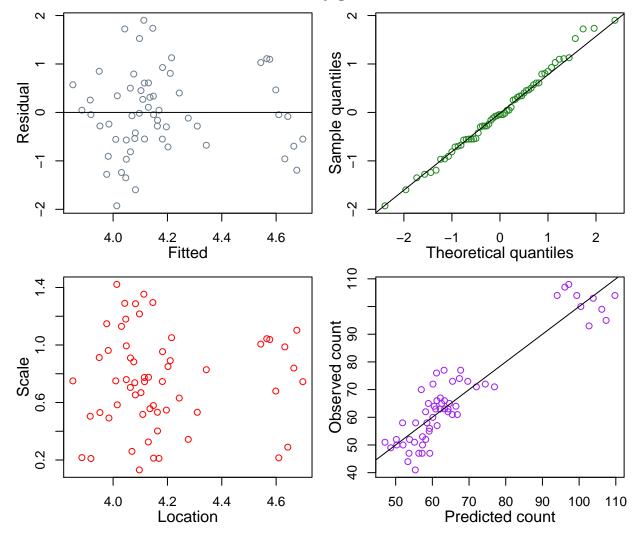
Ballysadare Bay glmm



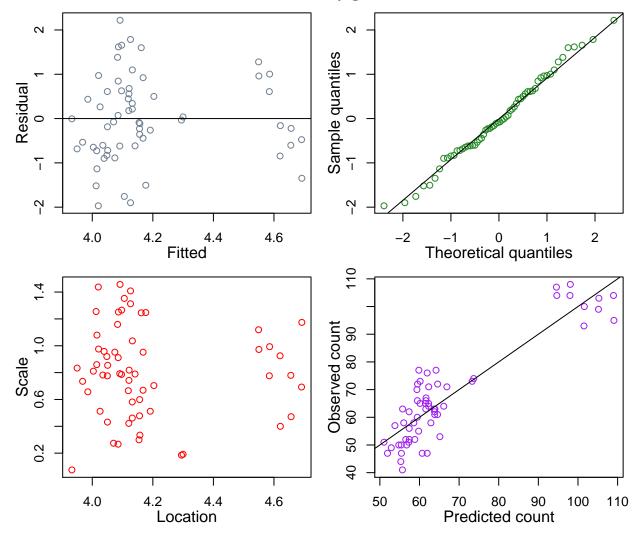
Ballysadare Bay gam 3.545



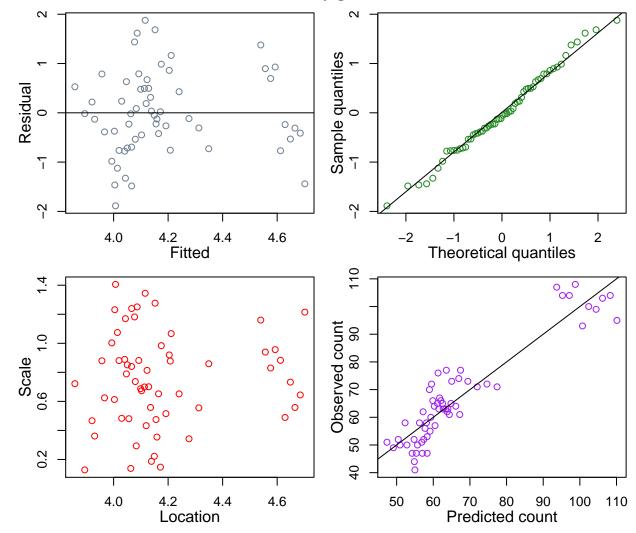
Cashla Bay glm 0.698



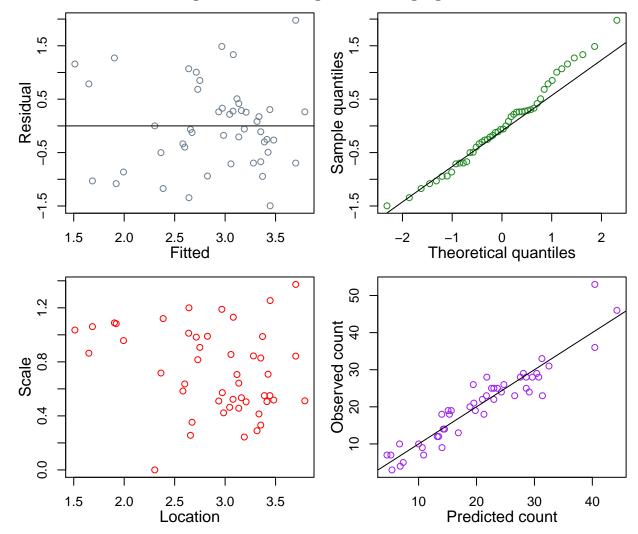
Cashla Bay glmm



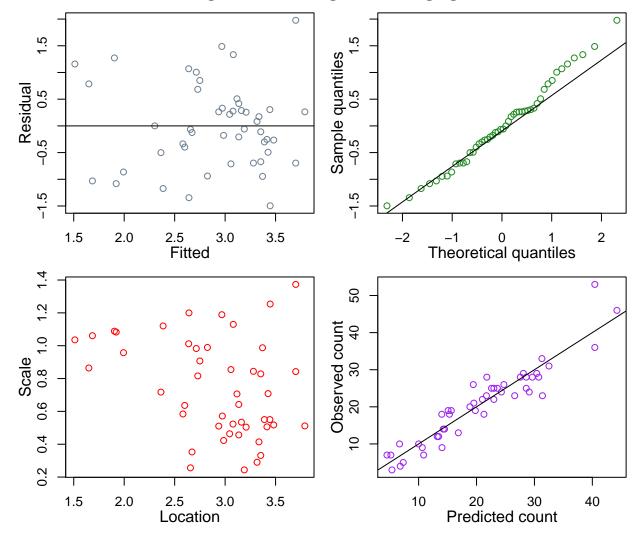
Cashla Bay gam 0.691



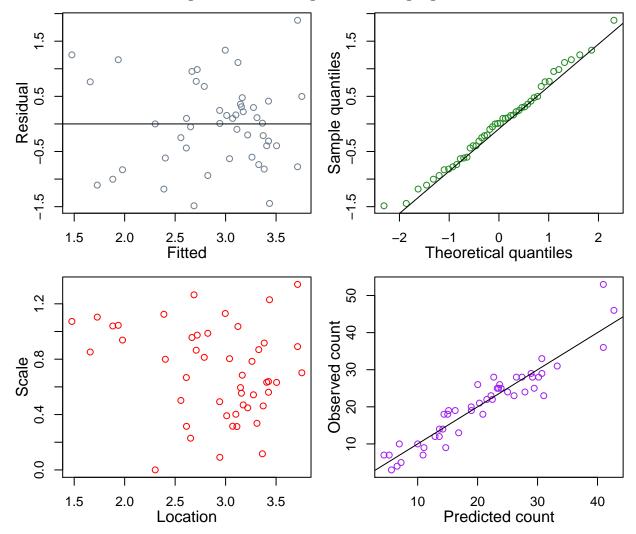
Emlagh Point, Roonagh. Louisburgh glm 0.605



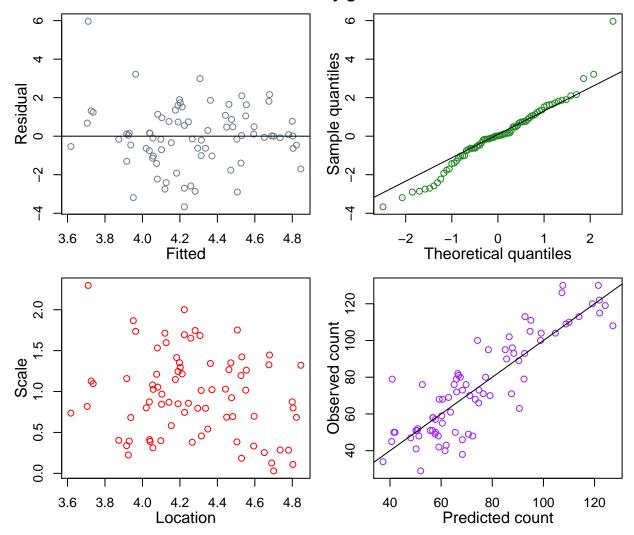
Emlagh Point, Roonagh. Louisburgh glmm



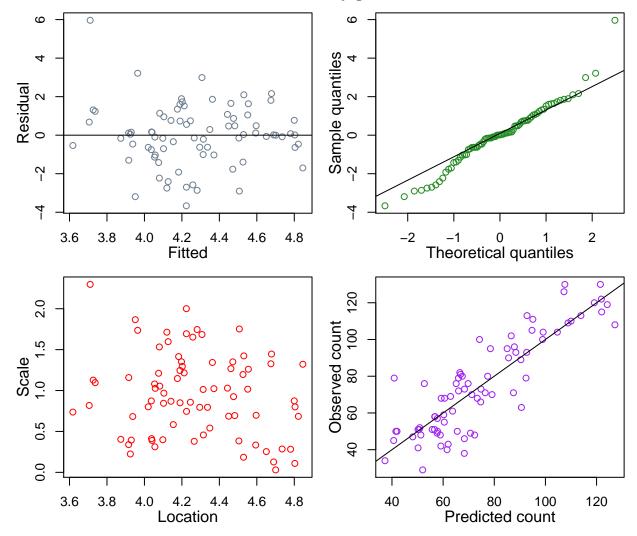
Emlagh Point, Roonagh. Louisburgh gam 0.569



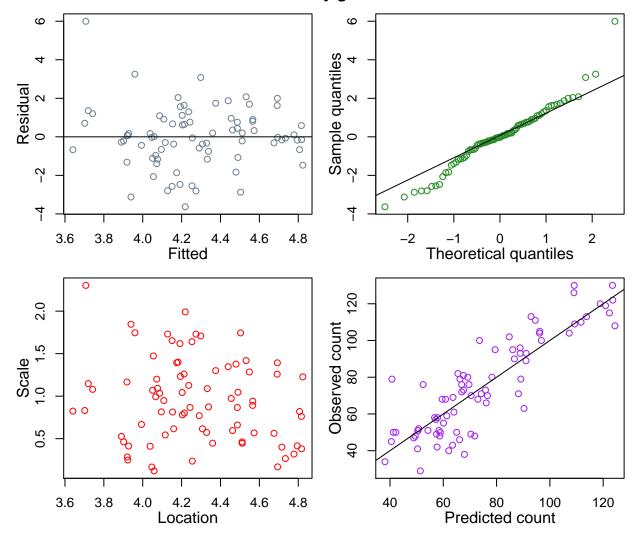
Kinvara Bay glm 2.5



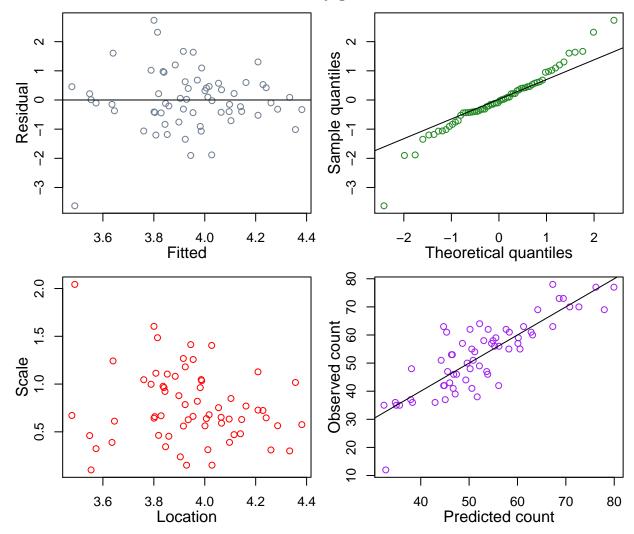
Kinvara Bay glmm



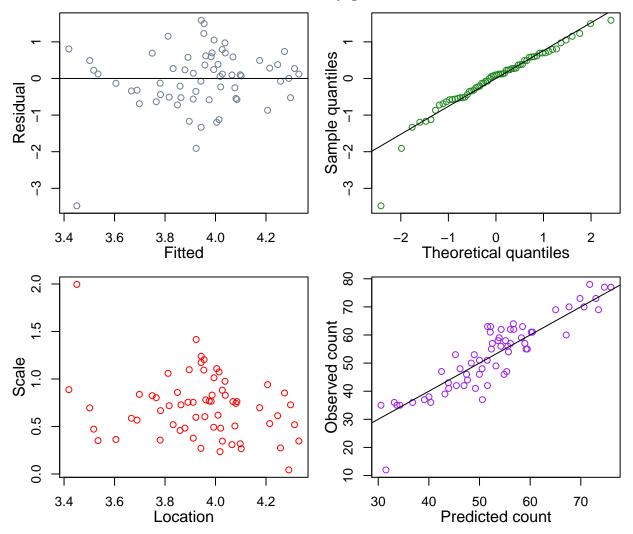
Kinvara Bay gam 2.436



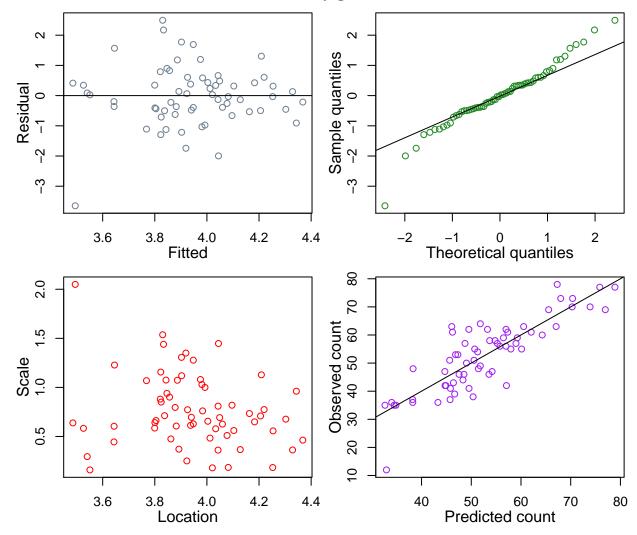
Mannin Bay glm 1.005



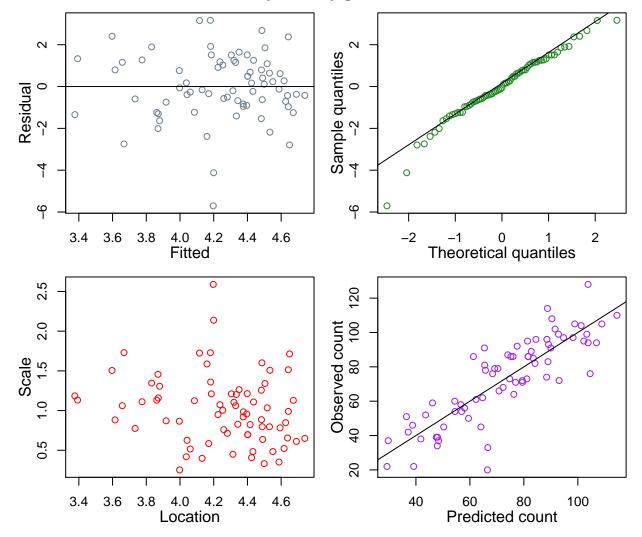
Mannin Bay glmm



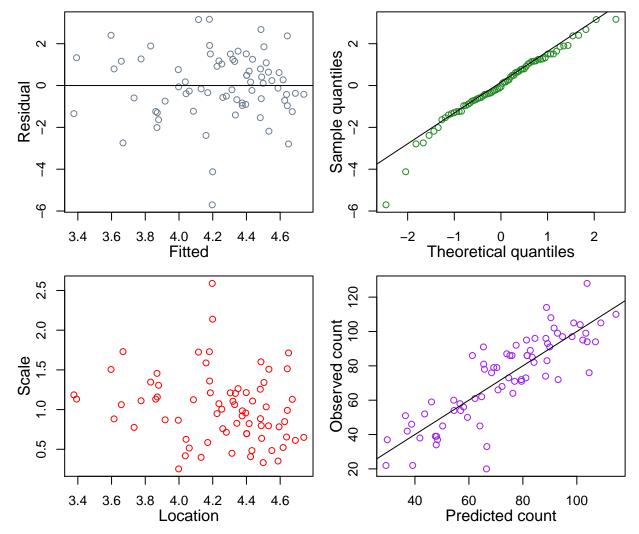
Mannin Bay gam 0.965



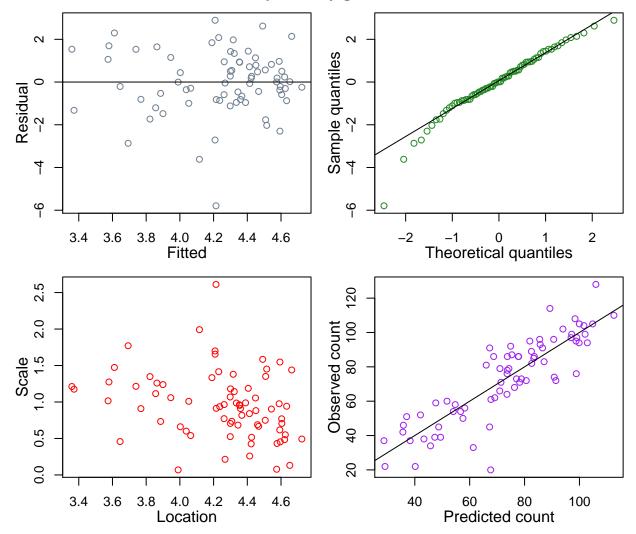
Moy Estuary glm 2.439



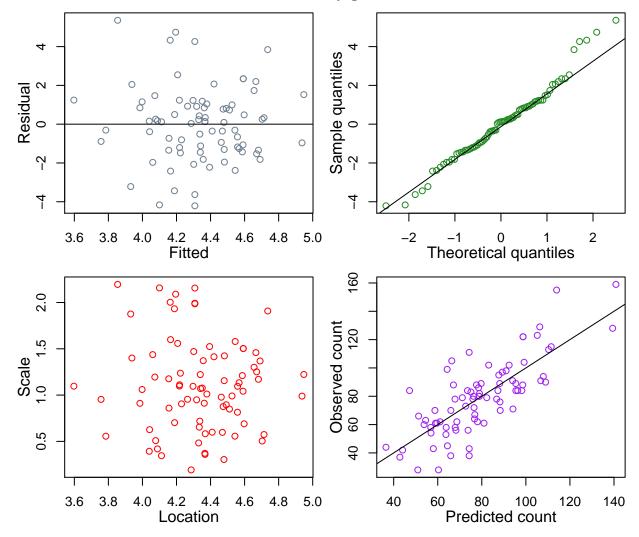
Moy Estuary glmm



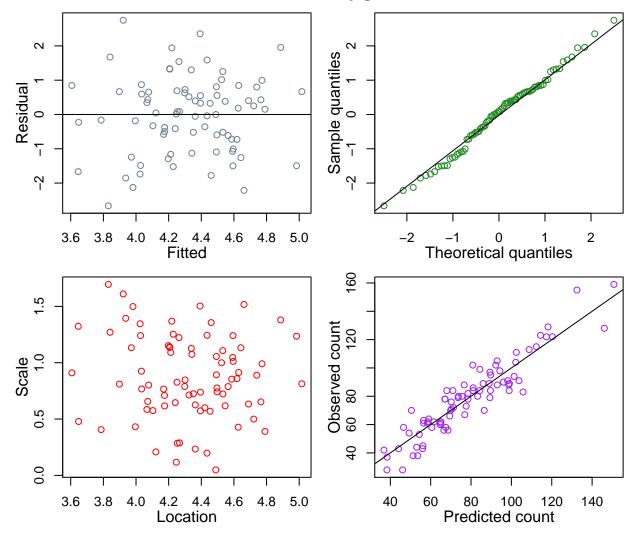
Moy Estuary gam 2.15



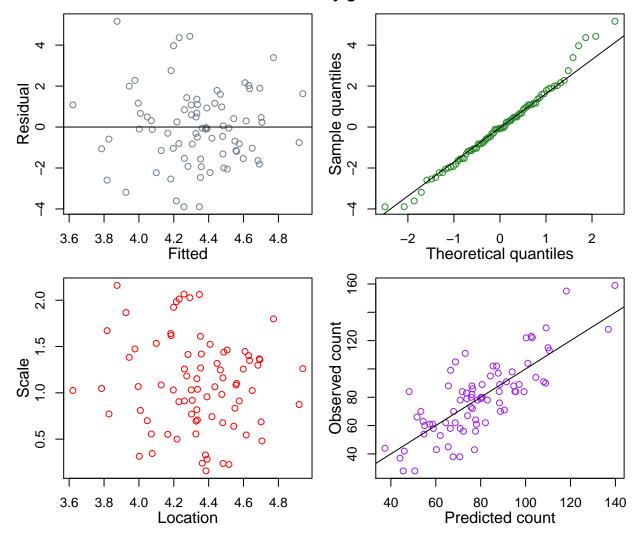
Oranmore Bay glm 3.598



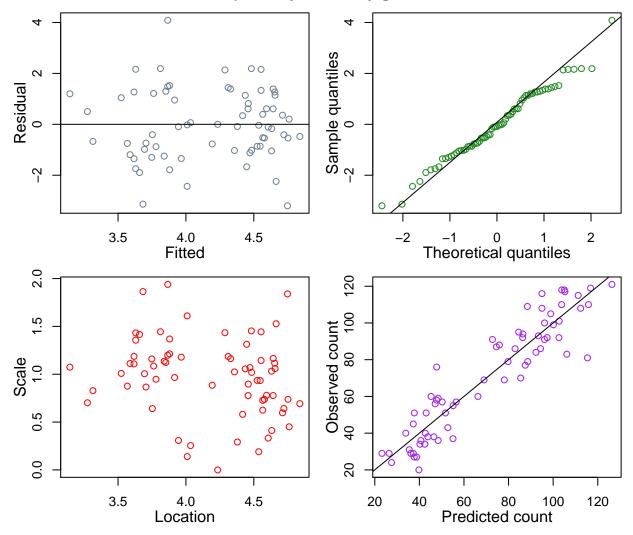
Oranmore Bay glmm



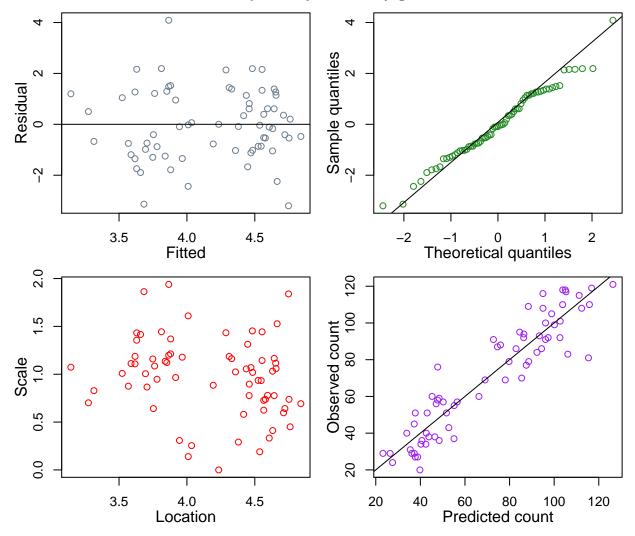
Oranmore Bay gam 3.421



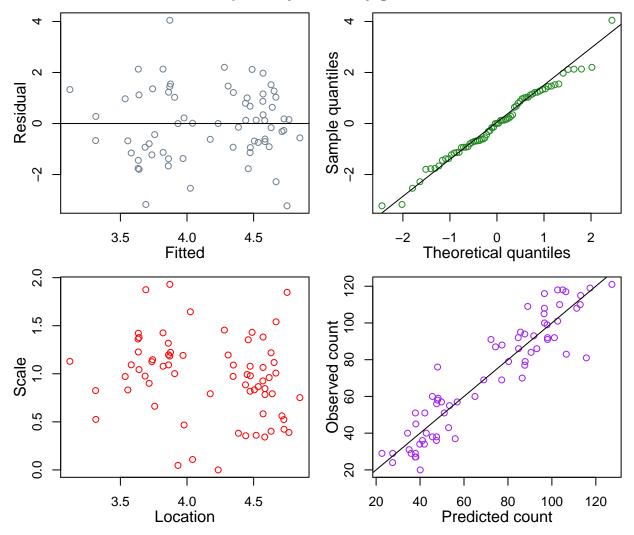
Westport Bay, Clew Bay glm 1.894



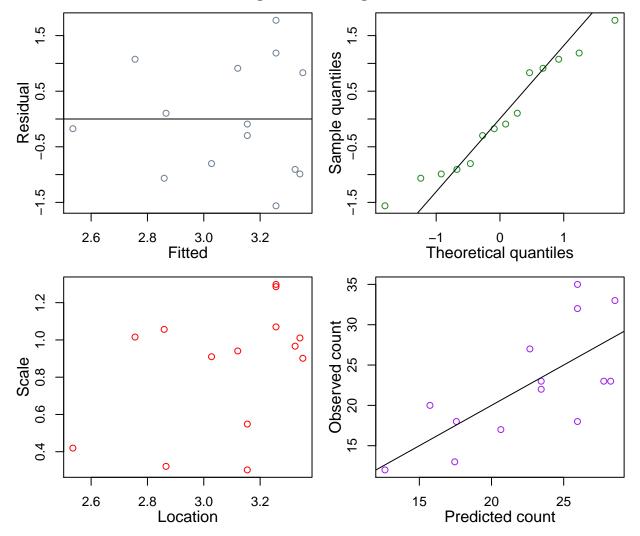
Westport Bay, Clew Bay glmm



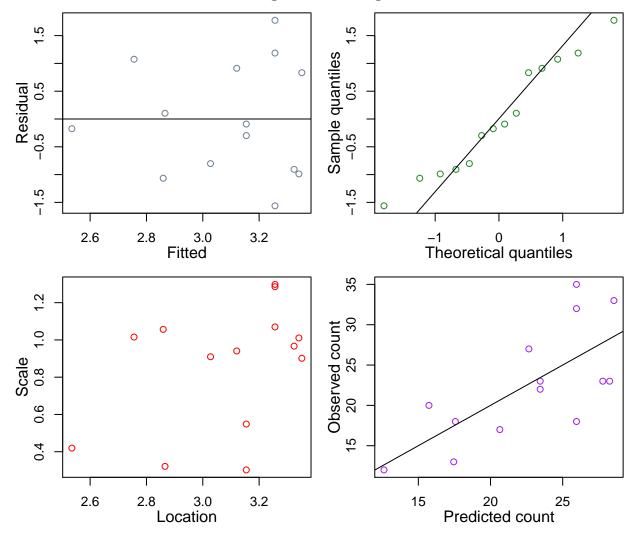
Westport Bay, Clew Bay gam 1.867



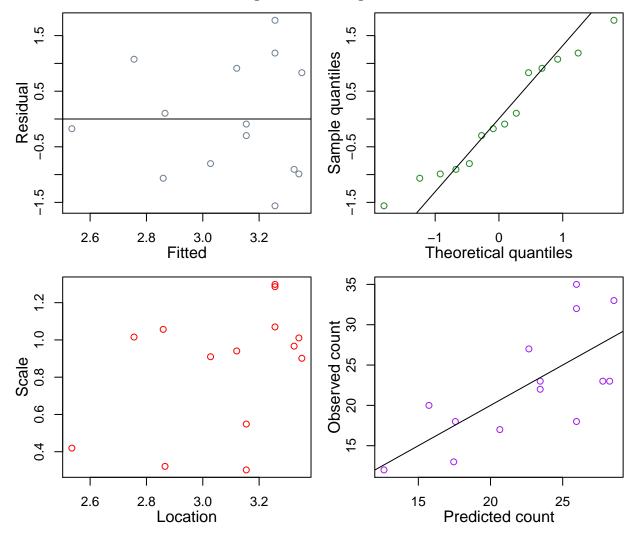
Adrigole Harbour glm 0.957



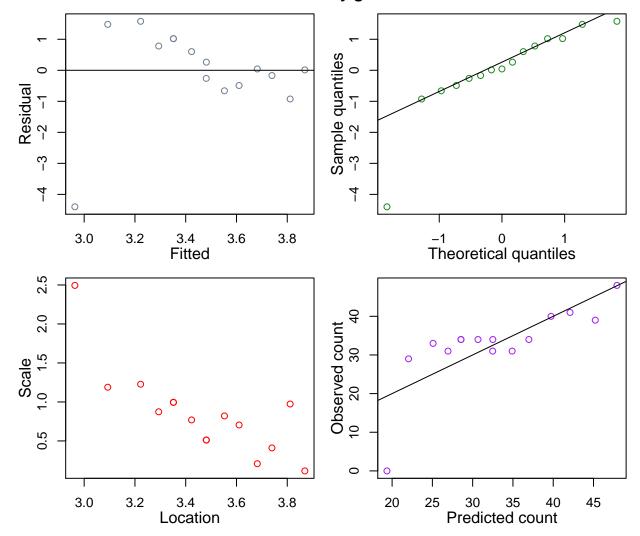
Adrigole Harbour glmm



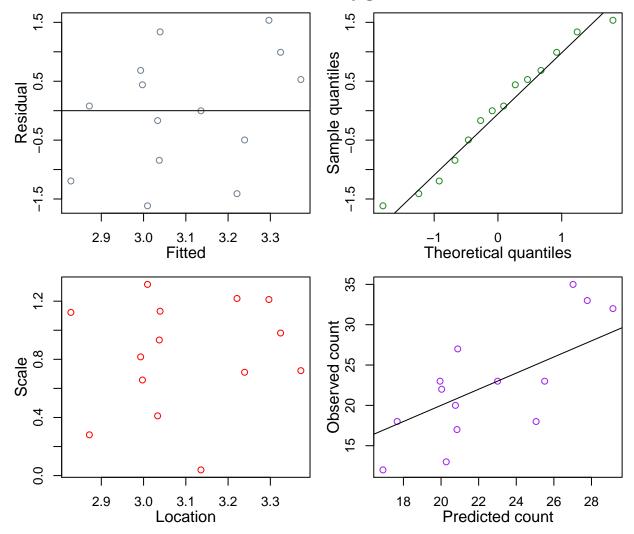
Adrigole Harbour gam 0.957



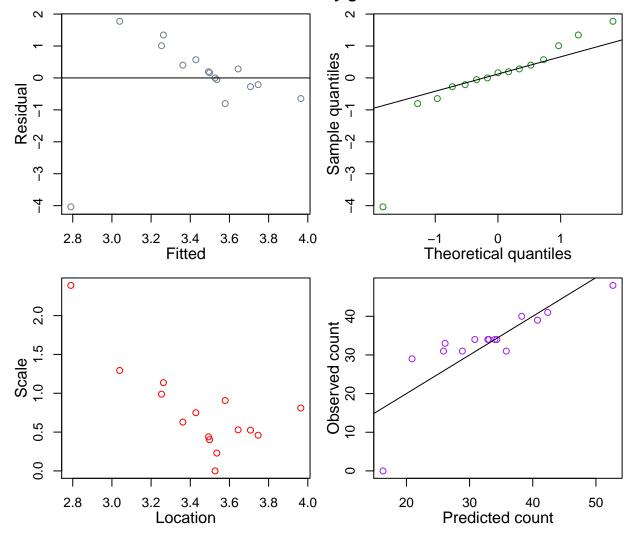
Gweebarra Bay glm 1.926



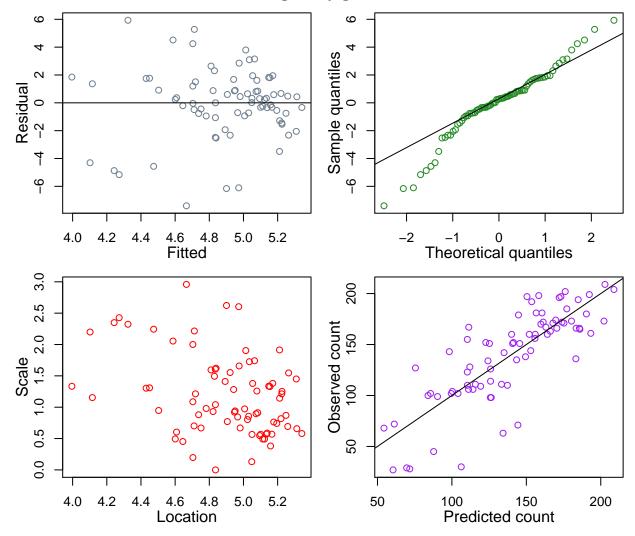
Gweebarra Bay glmm



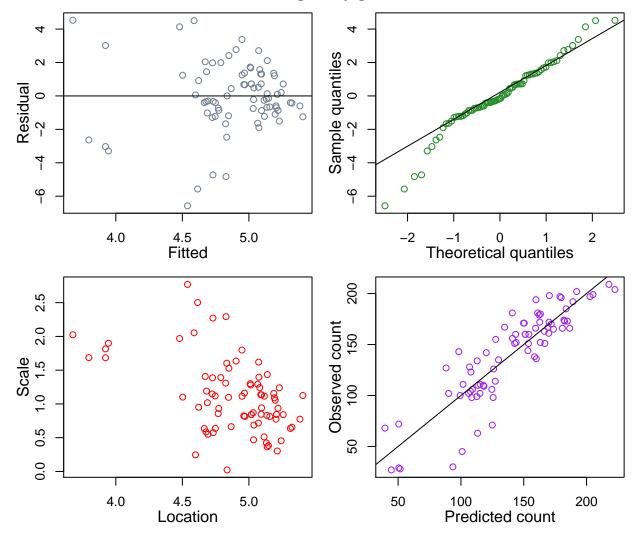
Gweebarra Bay gam 1.617



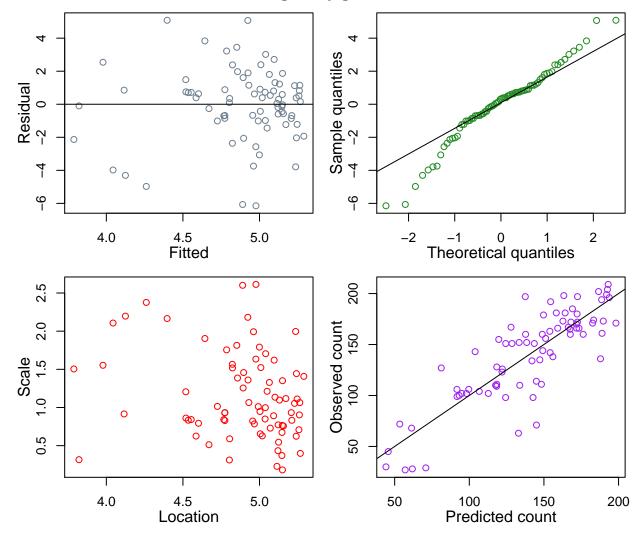
Donegal Bay glm 6.071



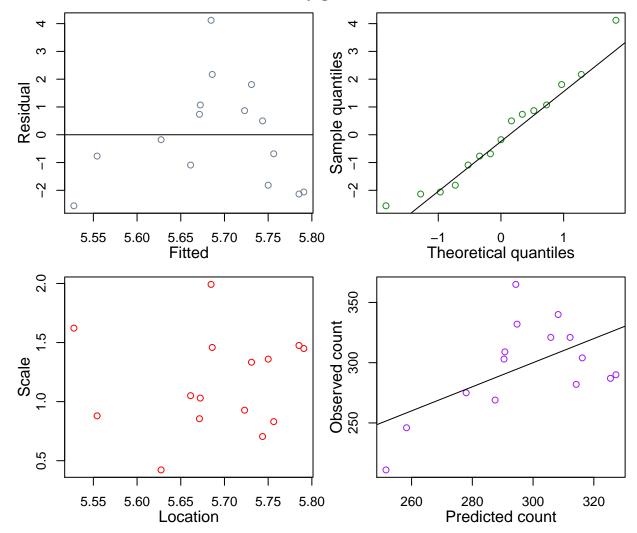
Donegal Bay glmm



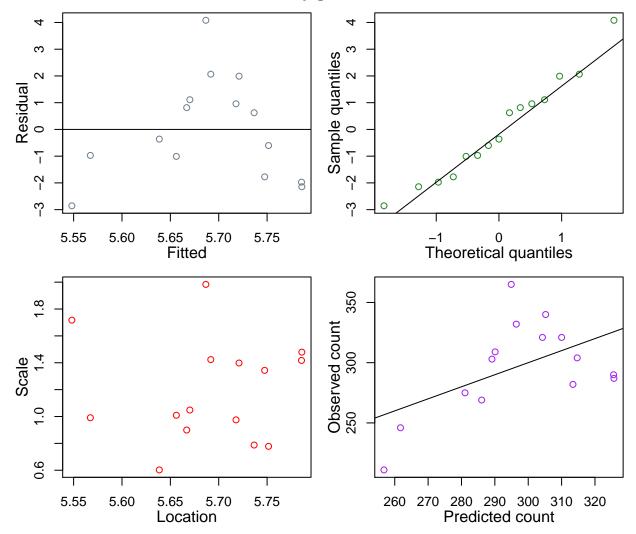
Donegal Bay gam 4.748



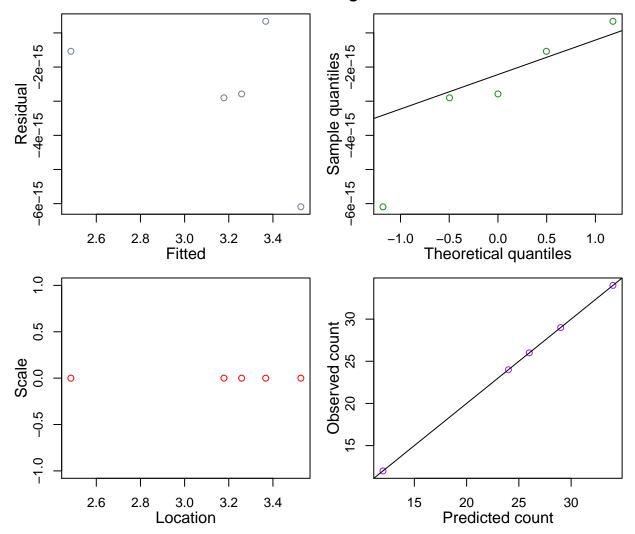
Bantry glm 3.238

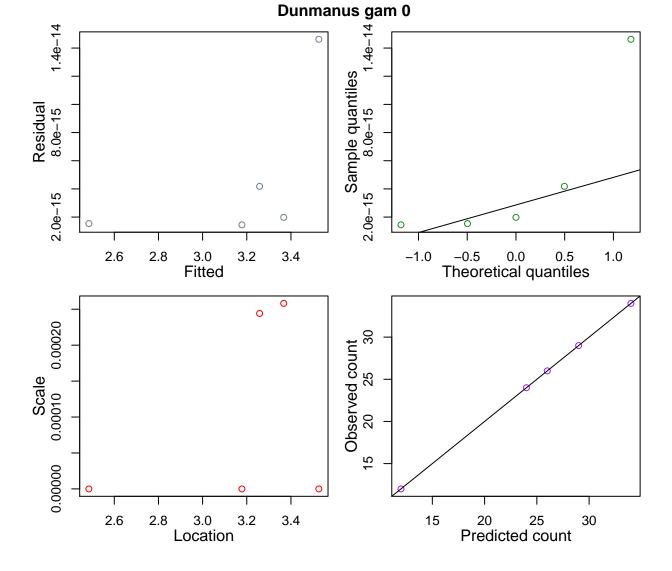


Bantry gam 3.357

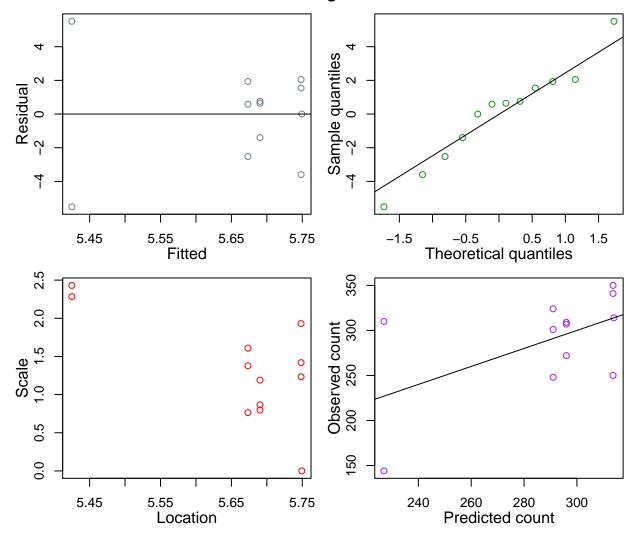


Dunmanus glm 0

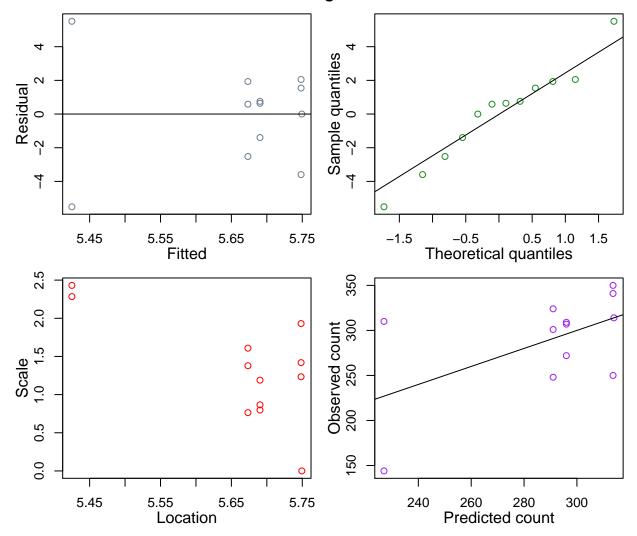




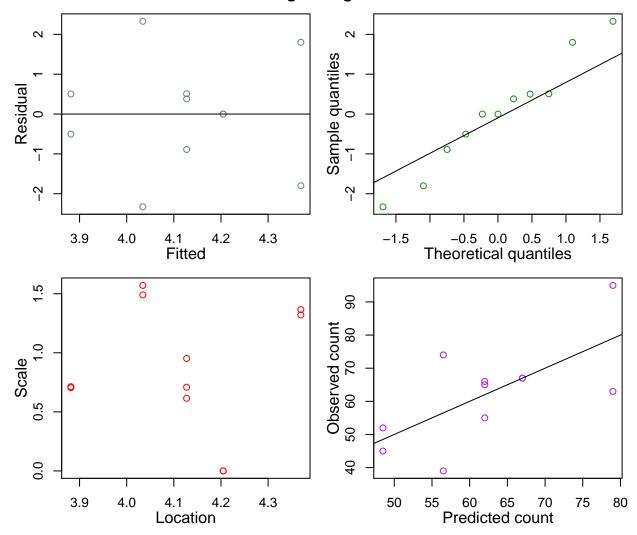
Kenmare glm 7.798



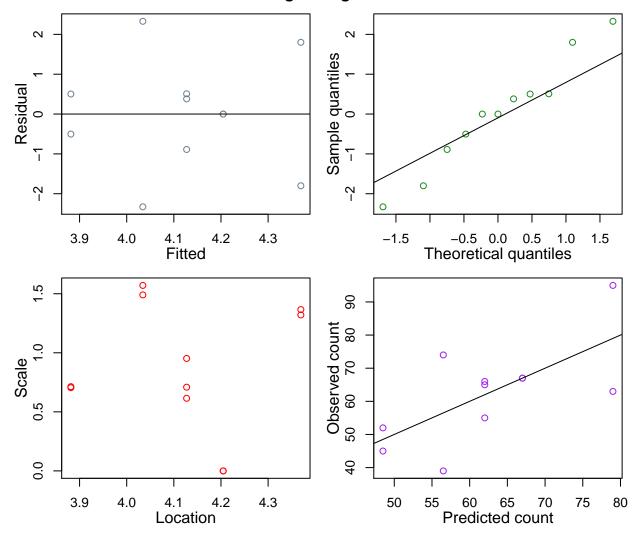
Kenmare gam 7.798



Roaringwater glm 1.729



Roaringwater gam 1.729



Appendix 3

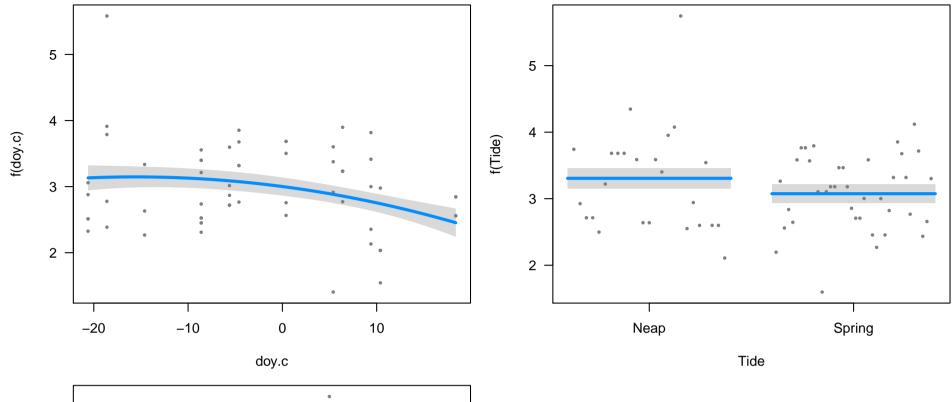
By land-based location GLM, GAM and GLMM effects plots. Partial residuals are included to visualize the influence of given data points on the estimates. The corrected Akaike's Information Criterion for a given fit is displayed in the title.

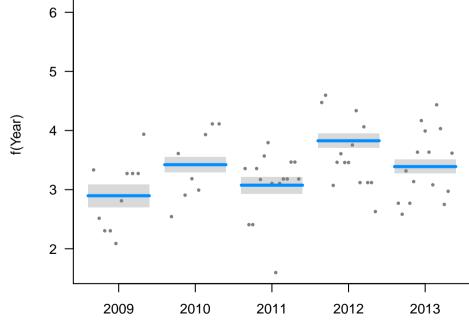
Variable abbreviations are given by:

- *hours.from.lw*: hours from low water
- *doy.c*: day-of-the-year centred on day 238
- *Tide*: tidal state (spring or neap)
- *disturbance.pres*: disturbance present or not
- *npts*: number per tidal state, which is the response

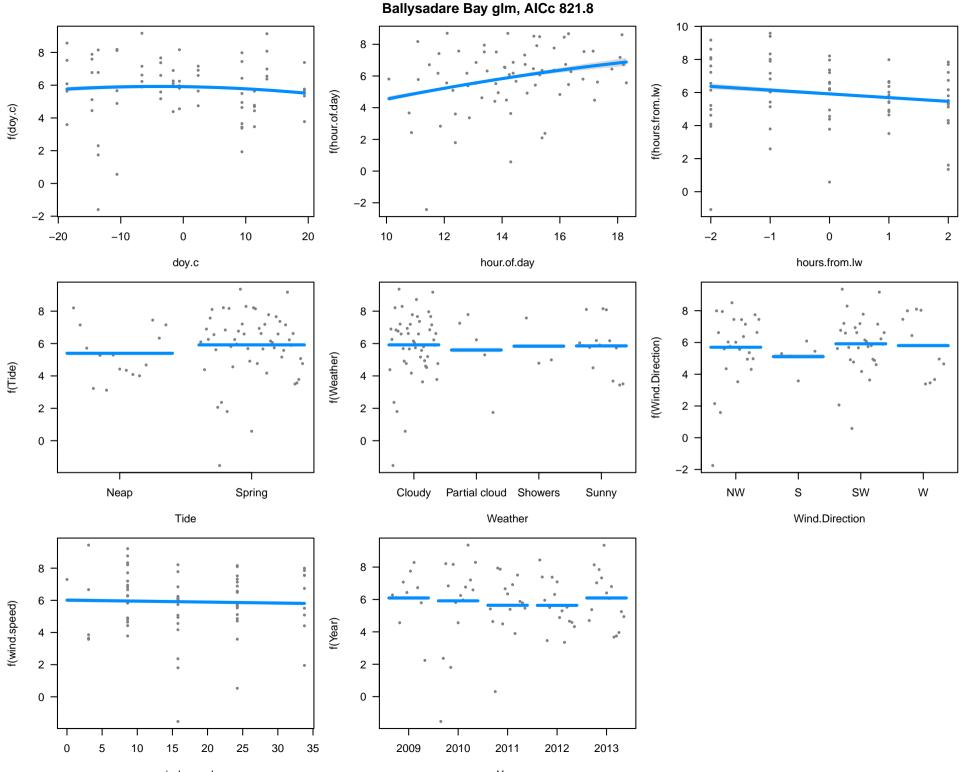
f() denotes functional form over a given variable.

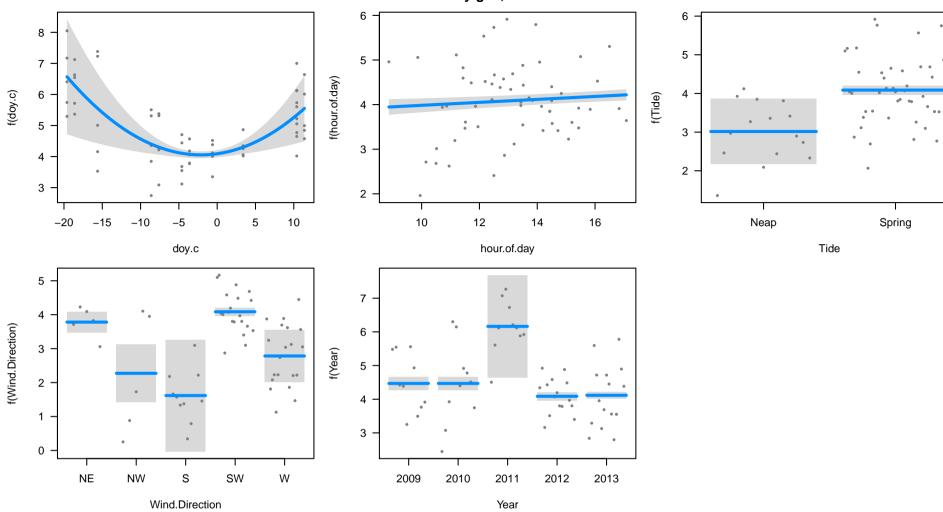
An Baile Lair, Inverin, Loughaunbeg glm, AICc 369.9





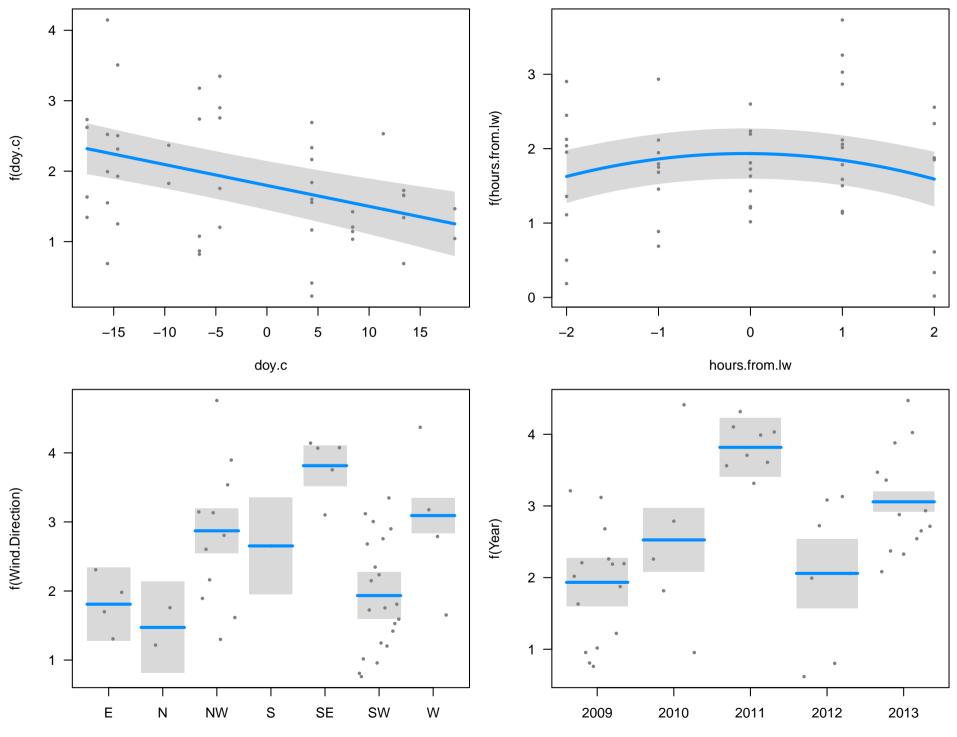
Year





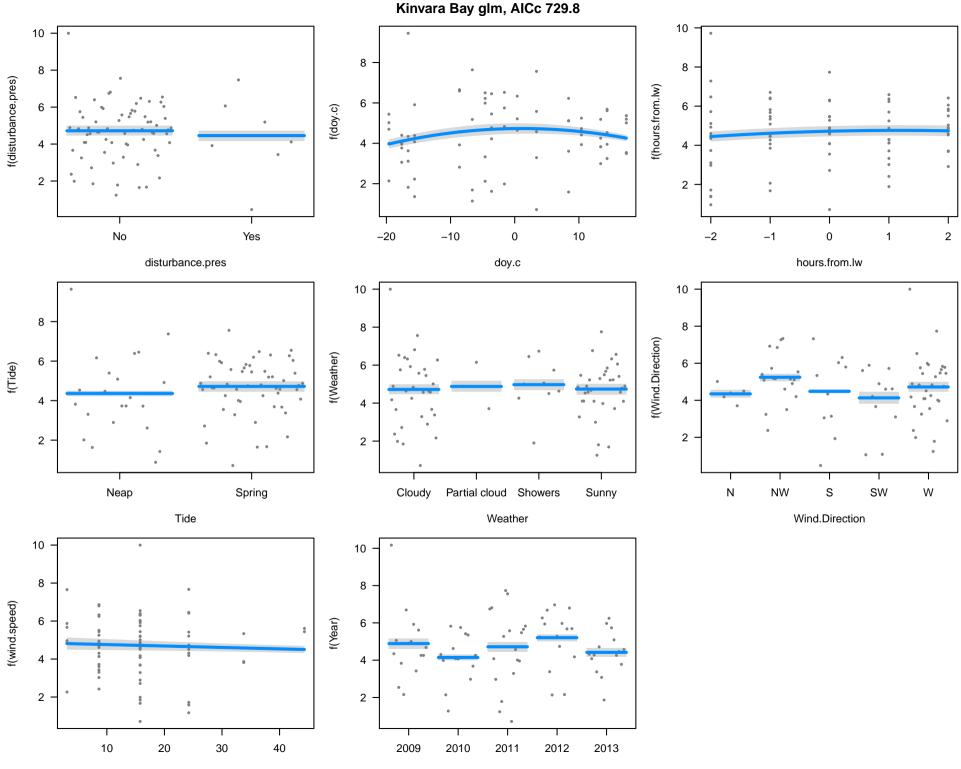
Cashla Bay glm, AICc 433.4

Emlagh Point, Roonagh. Louisburgh glm, AICc 297



Wind.Direction

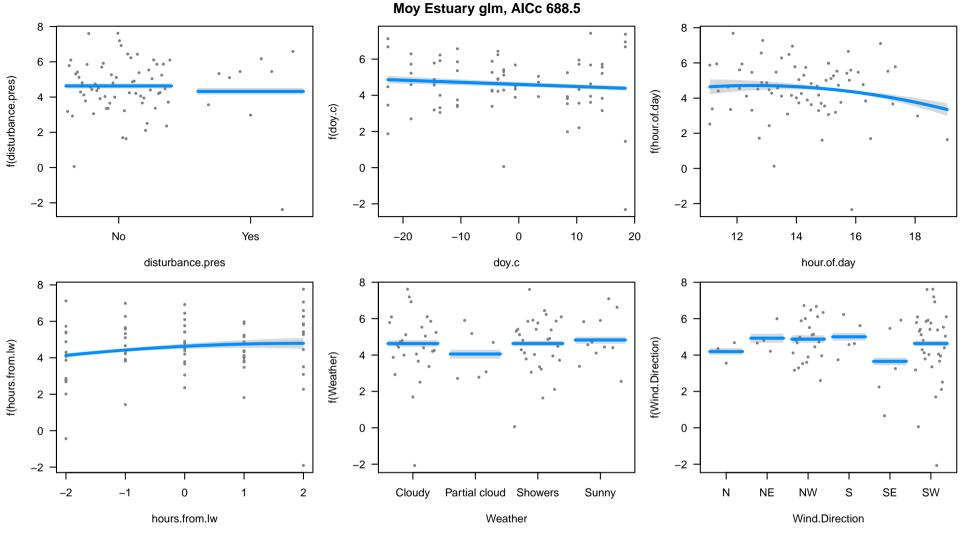
Year

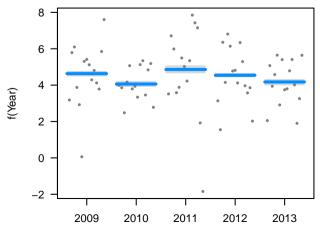


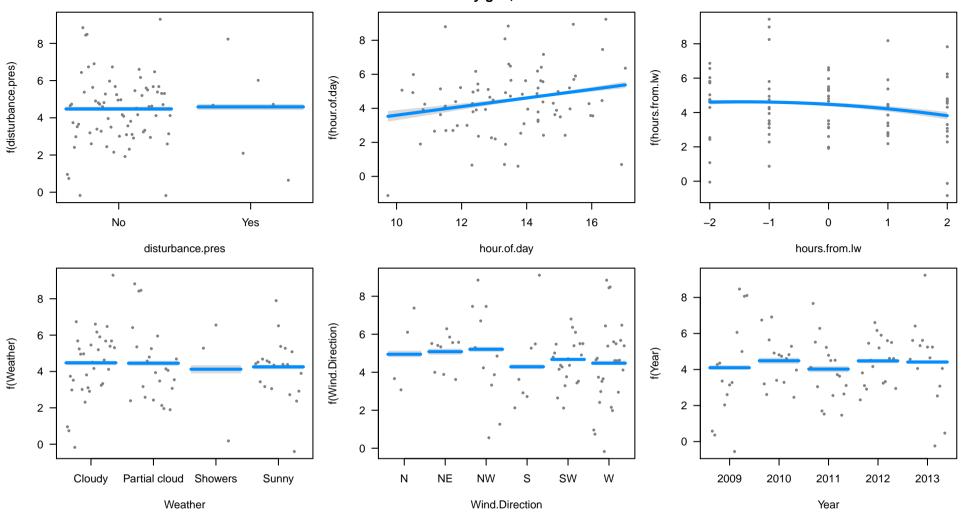


f(disturbance.pres) f(hour.of.day) f(doy.c) No Yes -20 -10 disturbance.pres doy.c hour.of.day f(Wind.Direction) f(hours.from.lw) f(Tide) 2 --1 Spring -2 -1 Neap NE NW SE SW W hours.from.lw Wind.Direction Tide . f(wind.speed) f(Year)

Mannin Bay glm, AICc 484.3





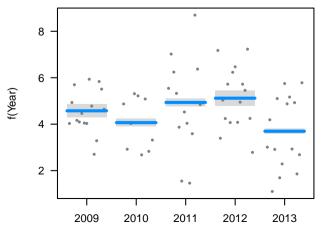


Oranmore Bay glm, AICc 824.7

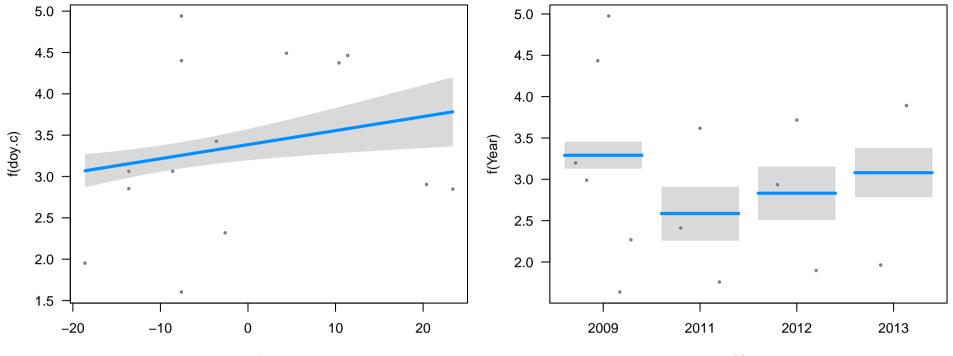
. . 8 8 8 6 f(hours.from.lw) 6 6 f(doy.c) f(Tide) 4 4 4 2 2 -2 --20 -10 0 10 20 -2 -1 0 1 2 Neap Spring doy.c hours.from.lw Tide . . 8 8 . . . 6 f(Wind.Direction) 6 6 f(wind.speed) f(Weather) 4 4 4 ... 2 2 · 2 · 0 Rain Sunny NW SE SW W Cloudy Showers s 10 15 20 25 30 35 0 5 Wind.Direction wind.speed

Westport Bay, Clew Bay glm, AICc 595.1

Weather



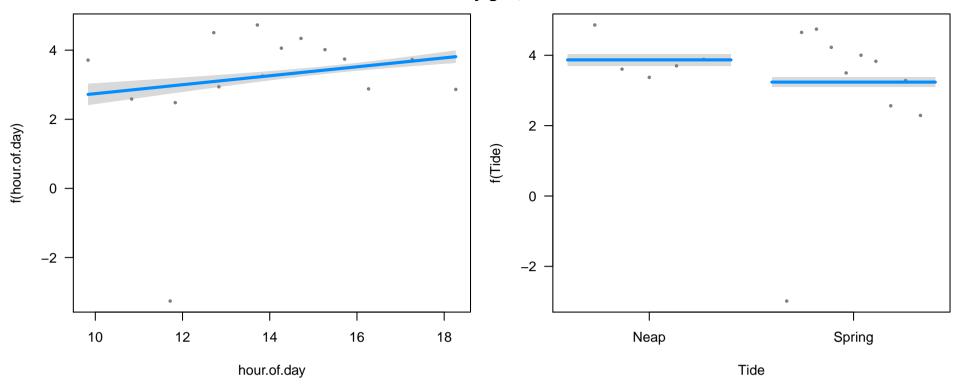
Adrigole Harbour glm, AICc 99.7

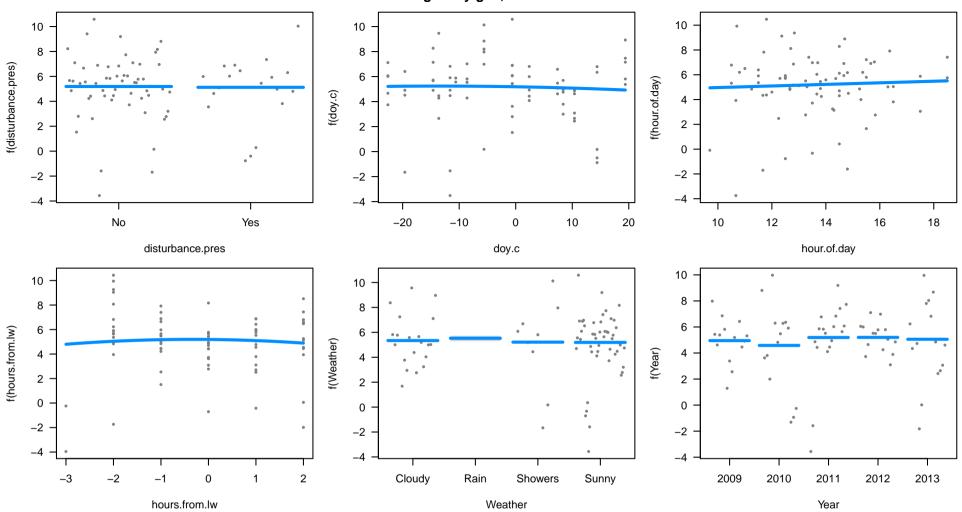




Year

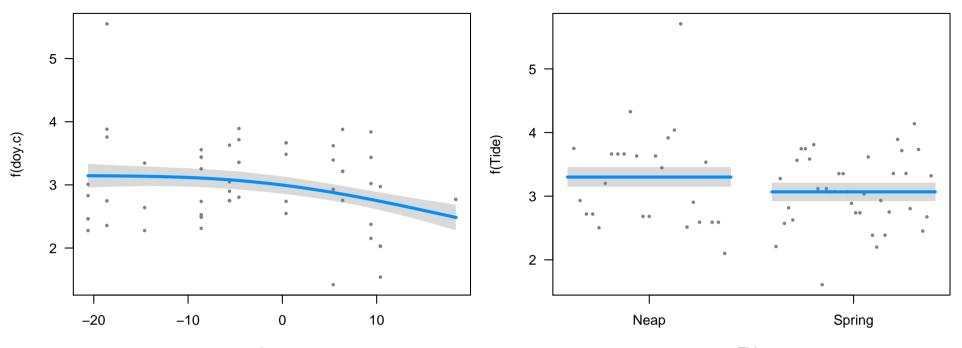
Gweebarra Bay glm, AICc 131.4





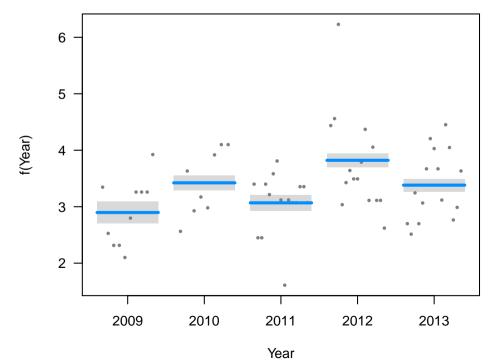
Donegal Bay glm, AICc 1078.1

An Baile Lair, Inverin, Loughaunbeg gam, AICc 369.5

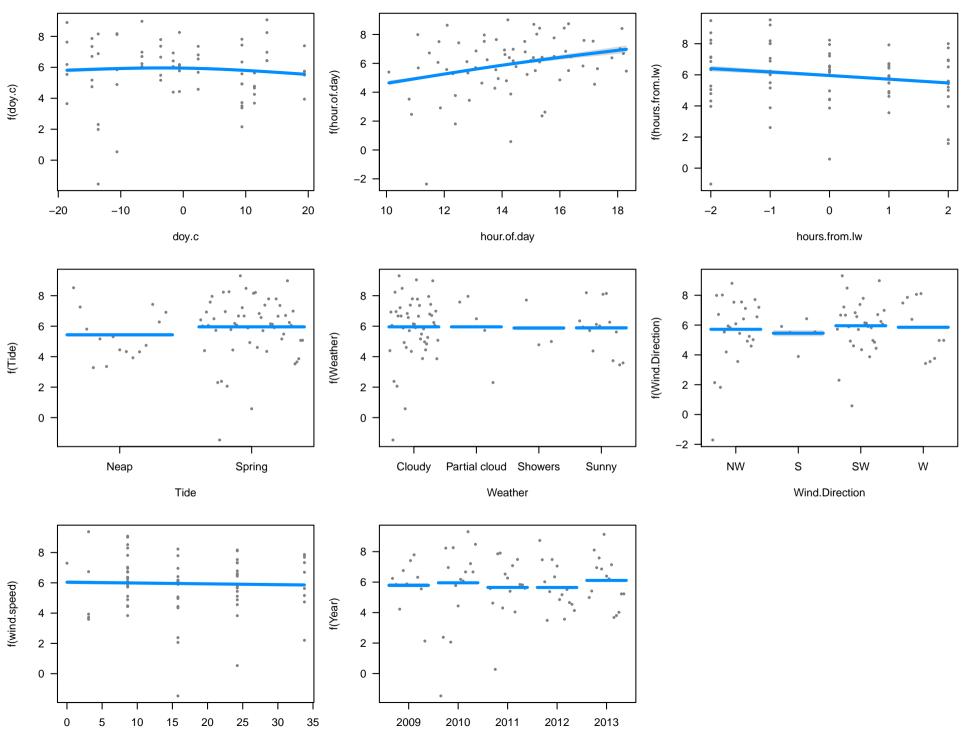




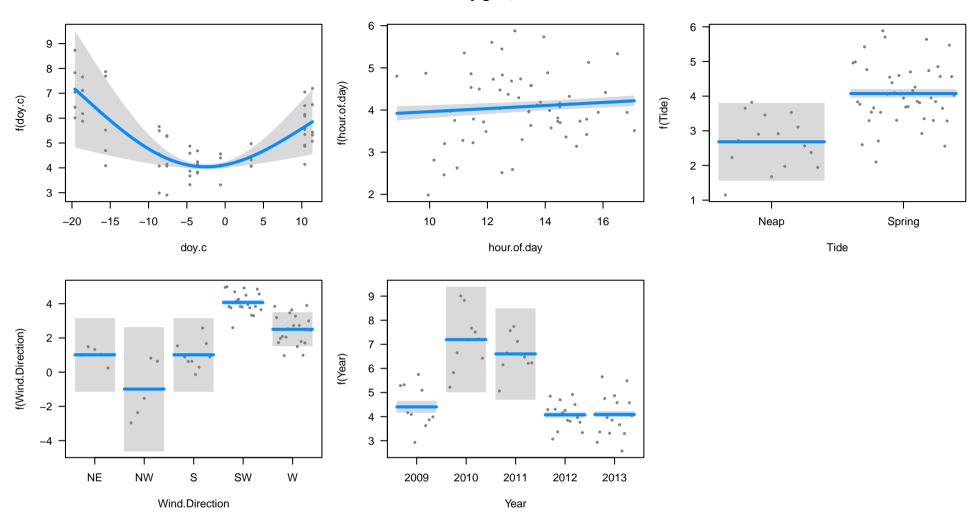




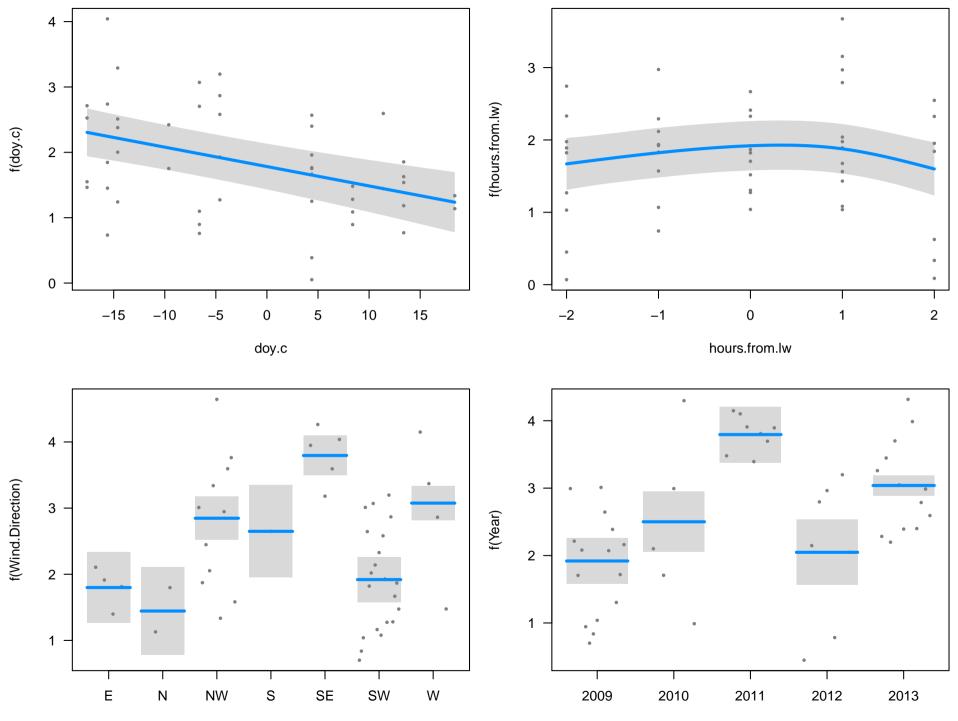
Ballysadare Bay gam, AICc 819.9



Cashla Bay gam, AICc 432.6



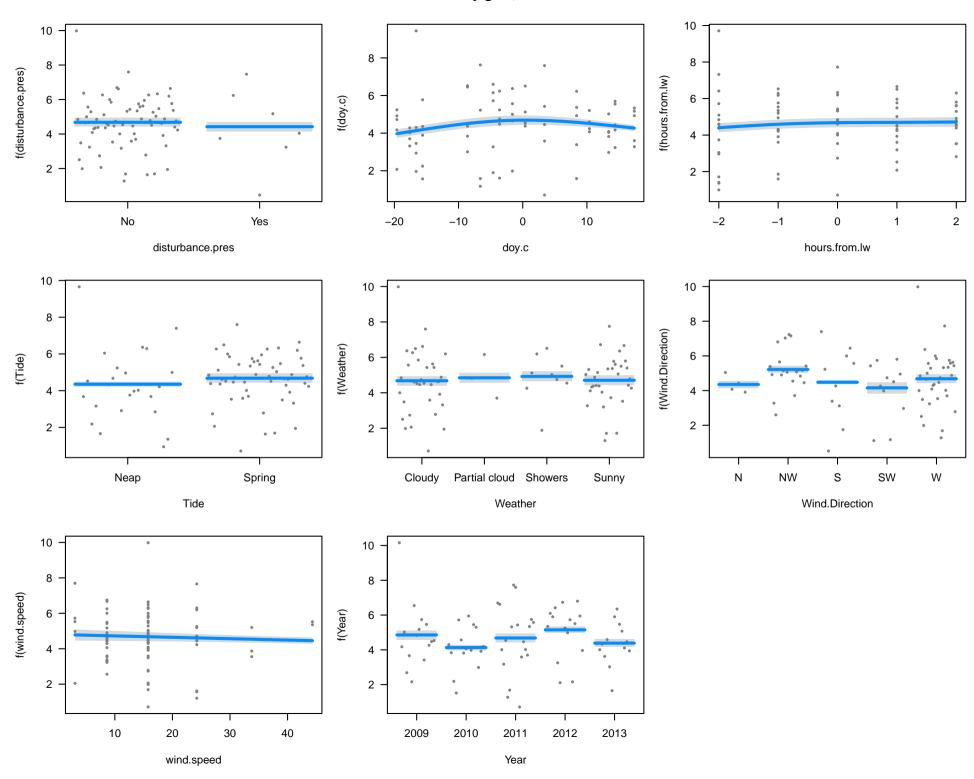
Emlagh Point, Roonagh. Louisburgh gam, AICc 298.3



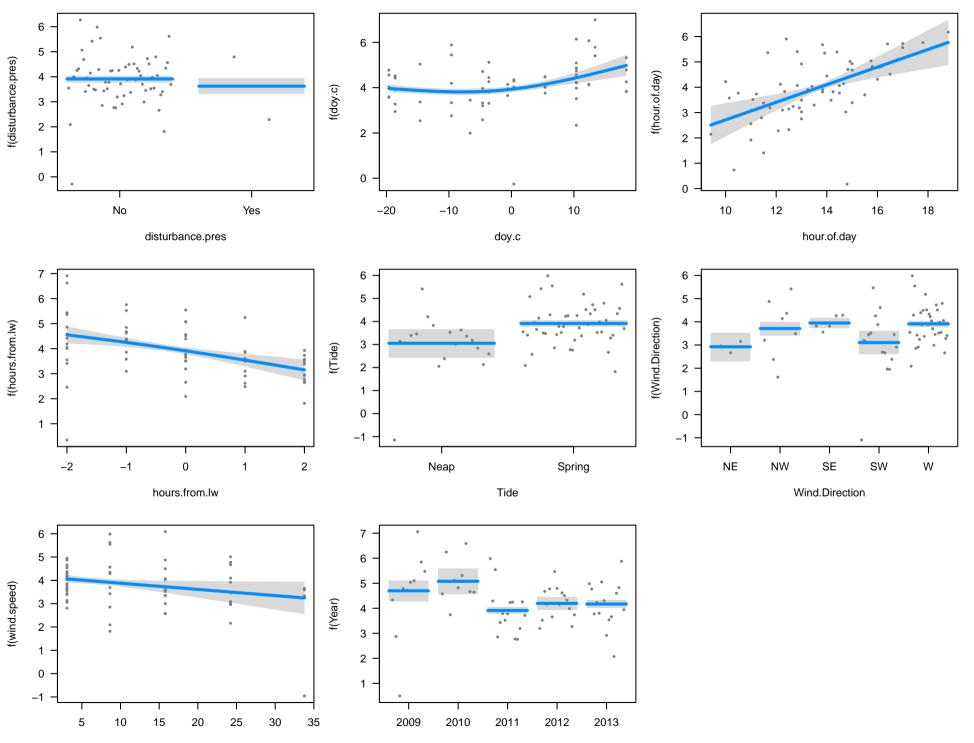
Wind.Direction

Year

Kinvara Bay gam, AICc 726.8

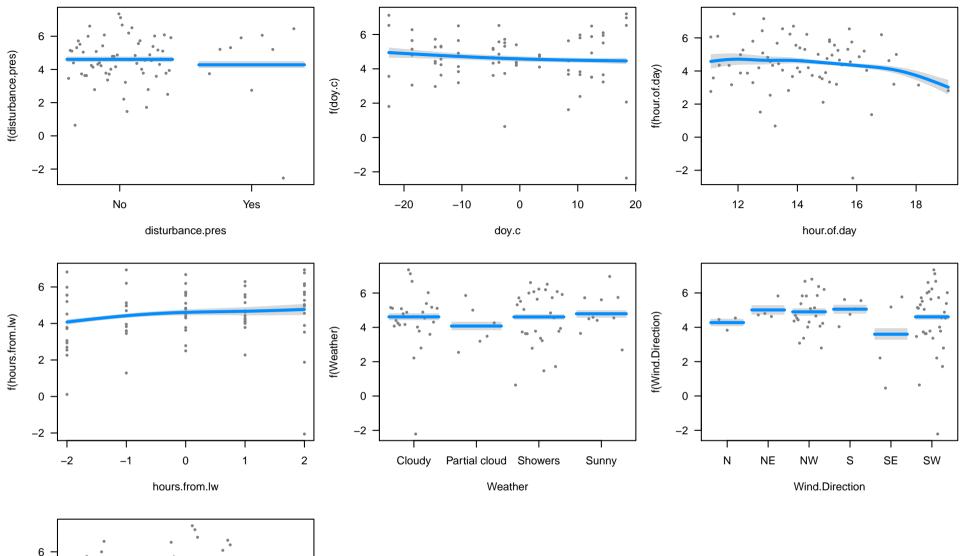


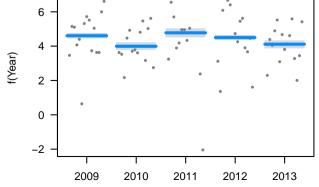
Mannin Bay gam, AICc 481.4



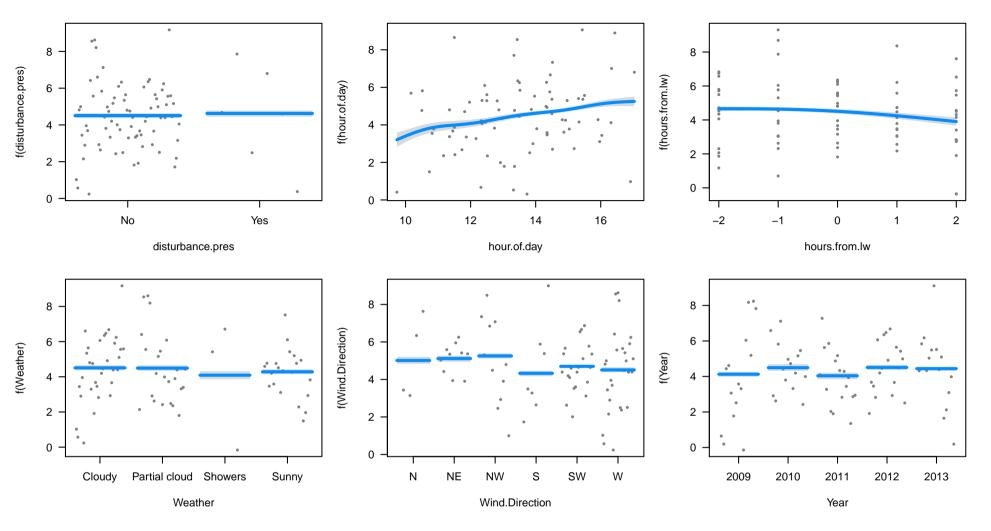
wind.speed

Moy Estuary gam, AICc 693.5

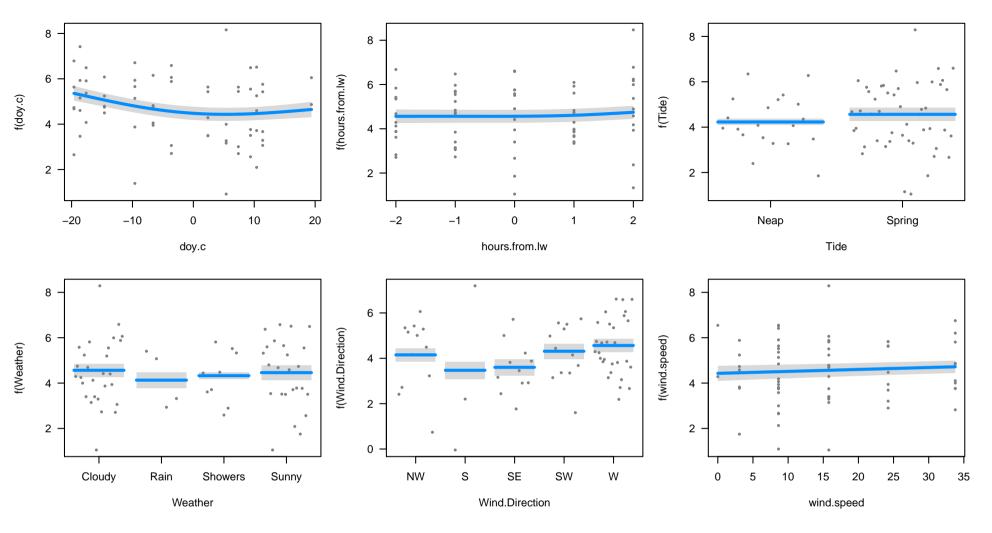


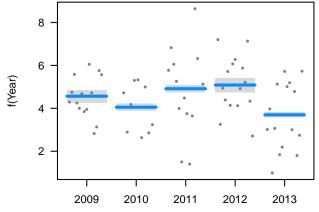


Oranmore Bay gam, AICc 834.2

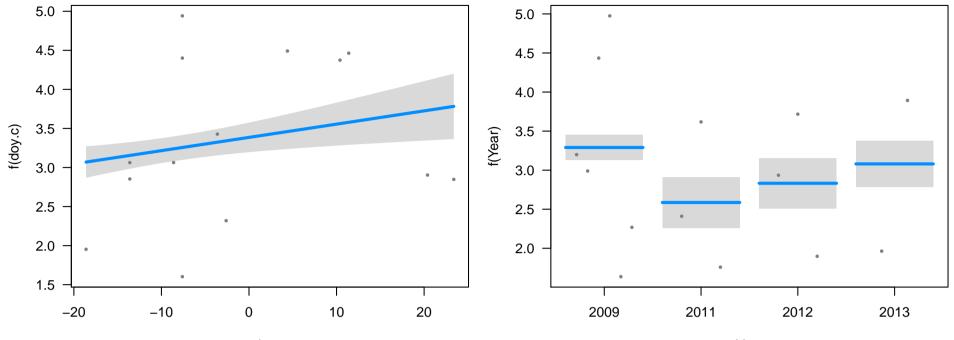


Westport Bay, Clew Bay gam, AICc 595.7





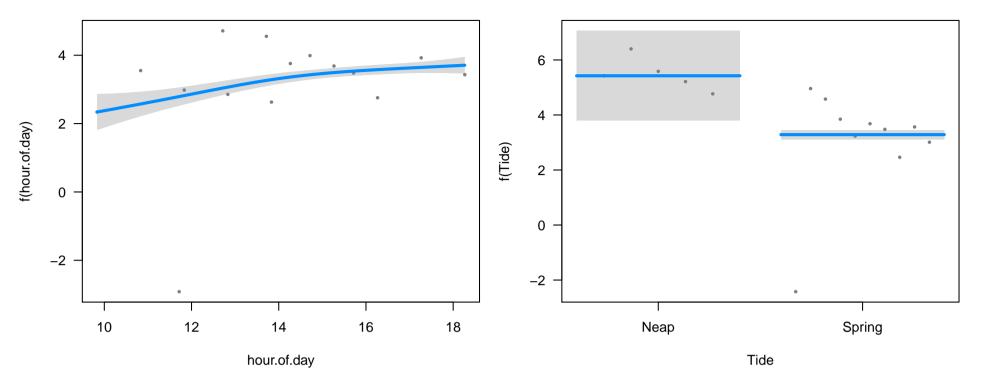
Adrigole Harbour gam, AICc 99.7

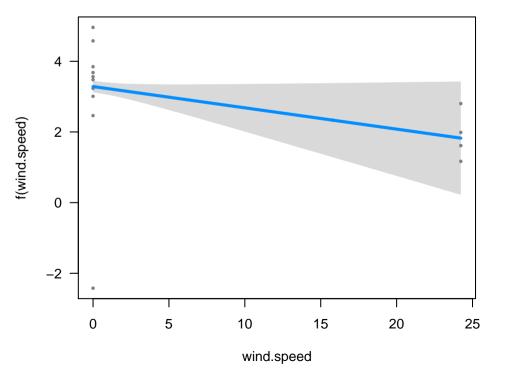




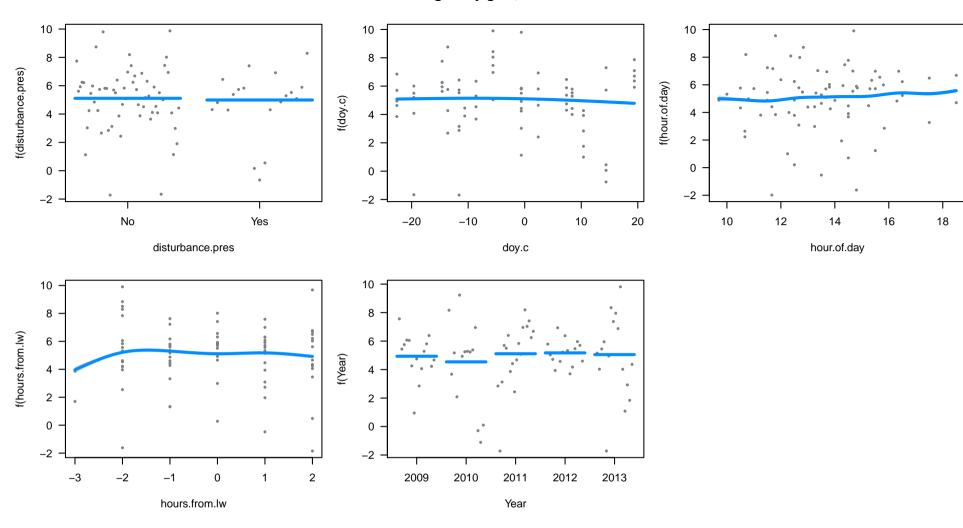
Year

Gweebarra Bay gam, AICc 132.6

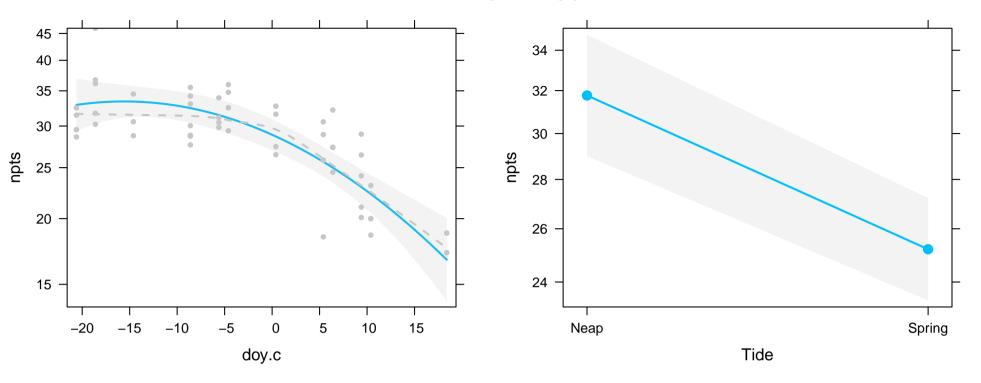


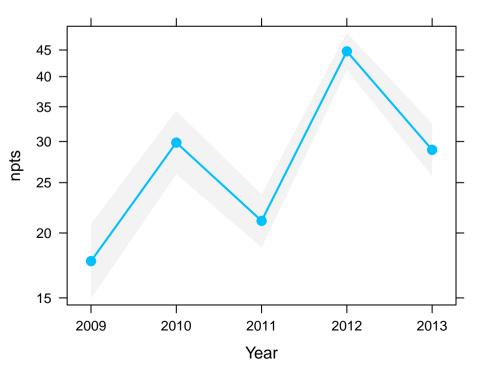


Donegal Bay gam, AICc 974.6

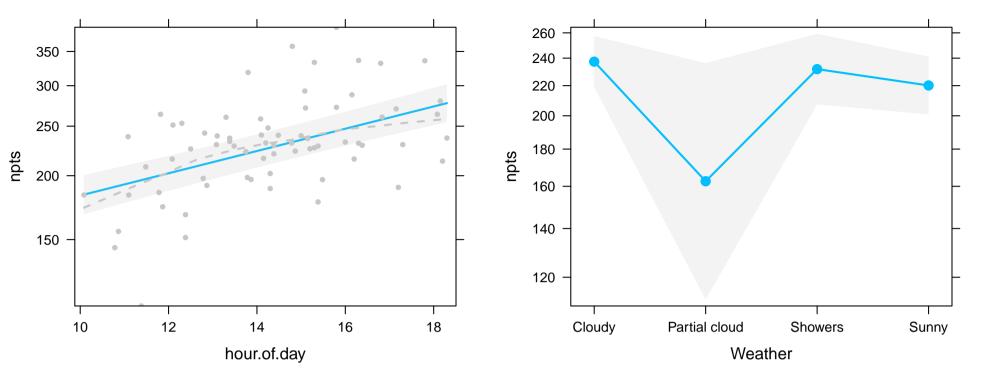


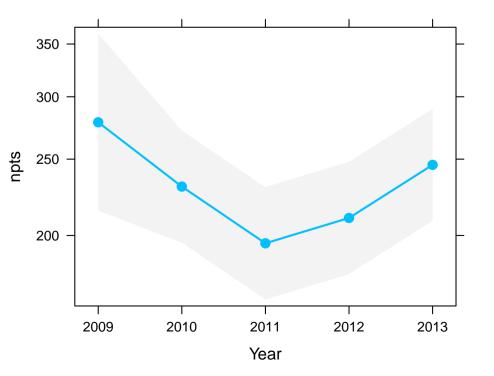
An Baile Lair, Inverin, Loughaunbeg glmm, AICc 372.6



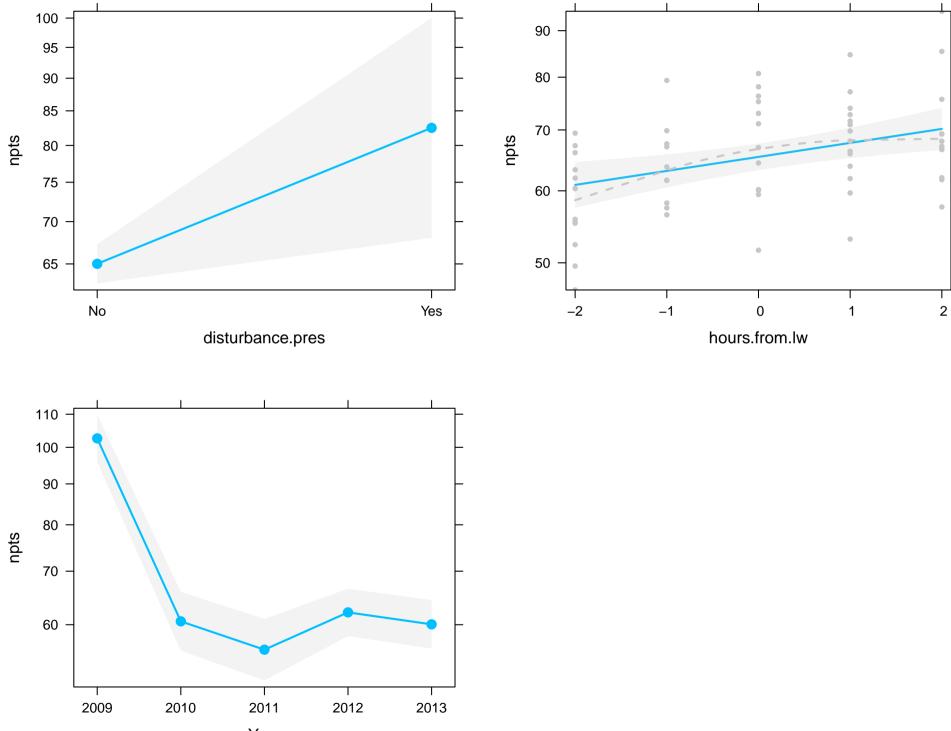


Ballysadare Bay glmm, AICc 791



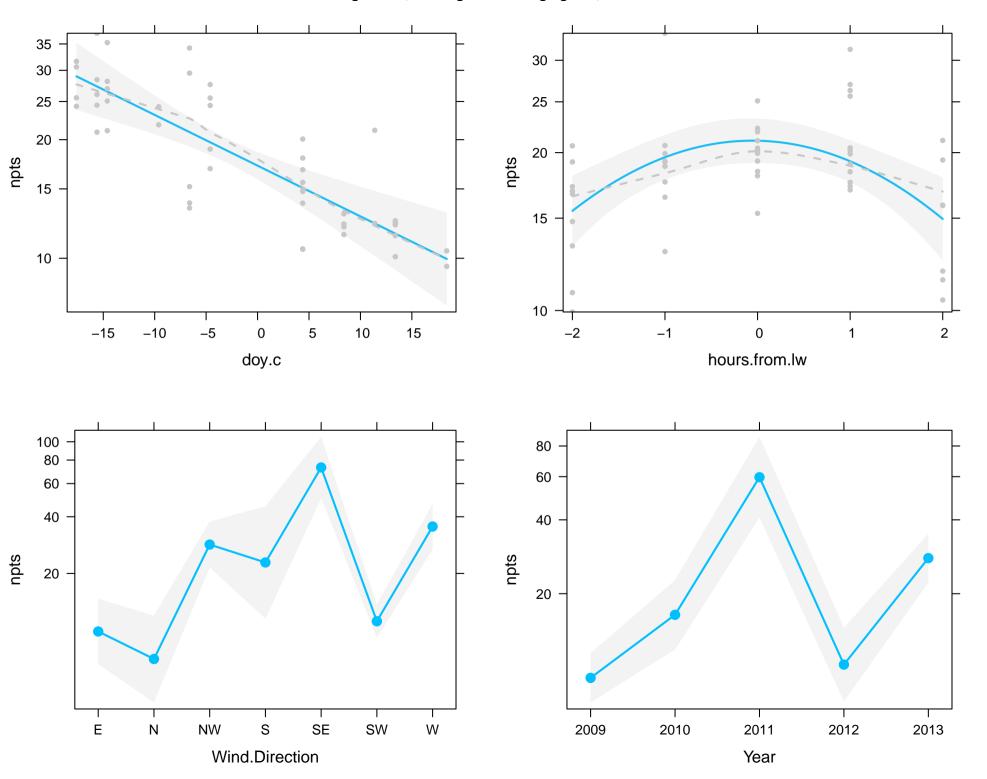


Cashla Bay glmm, AICc 436.5

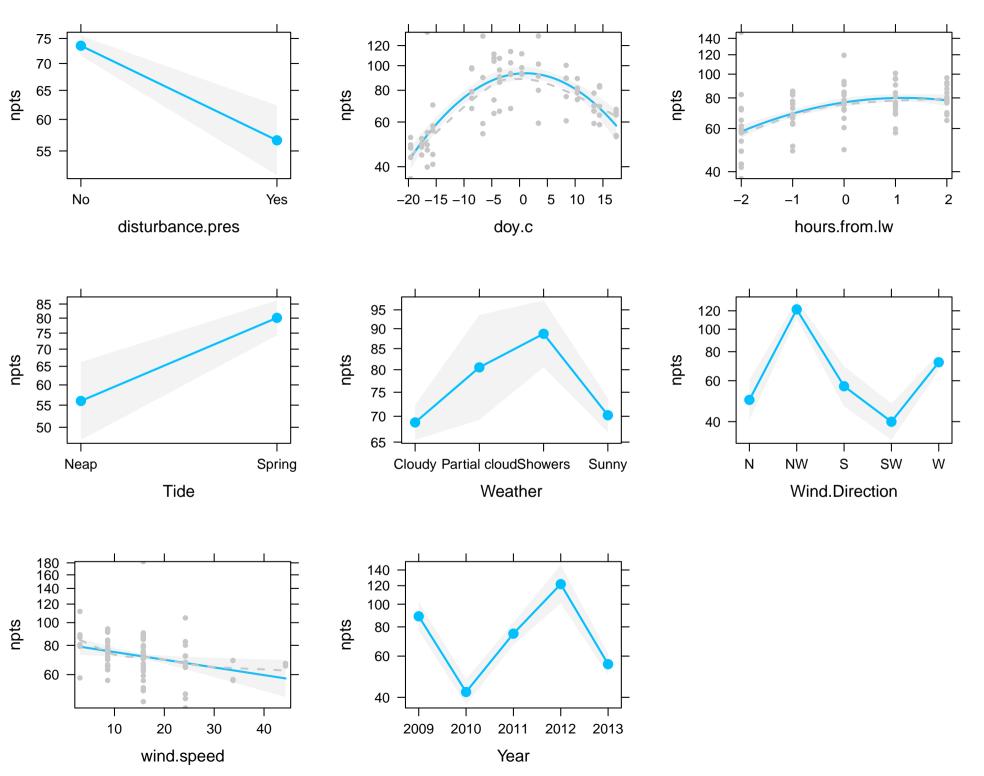


Year

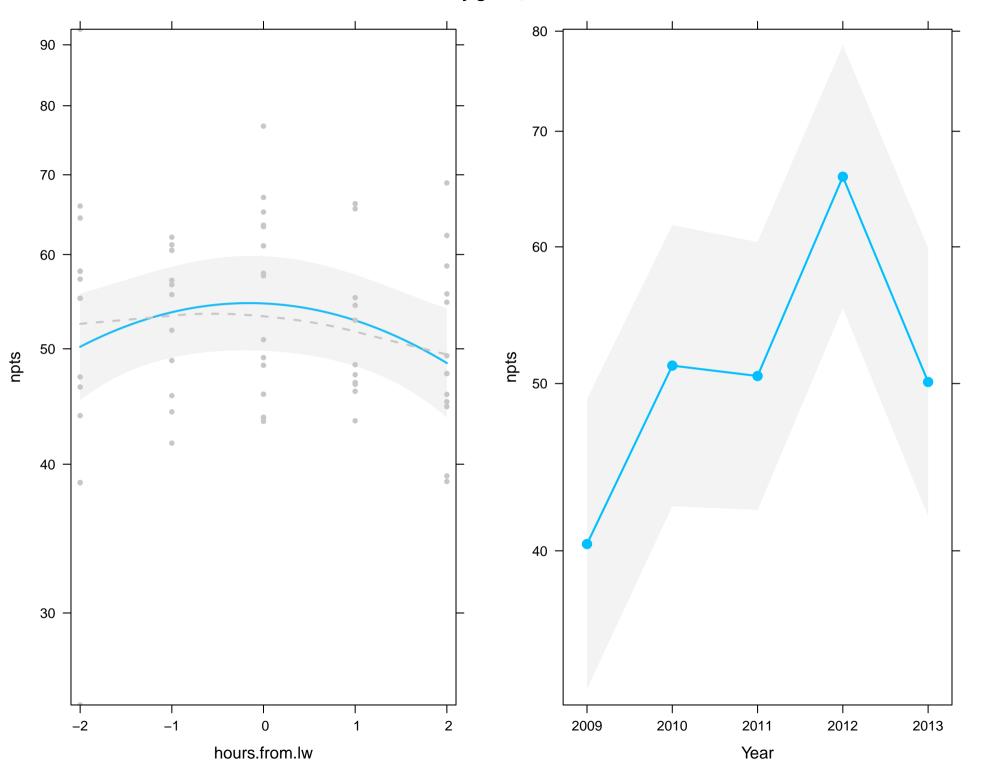
Emlagh Point, Roonagh. Louisburgh glmm, AICc 301.2



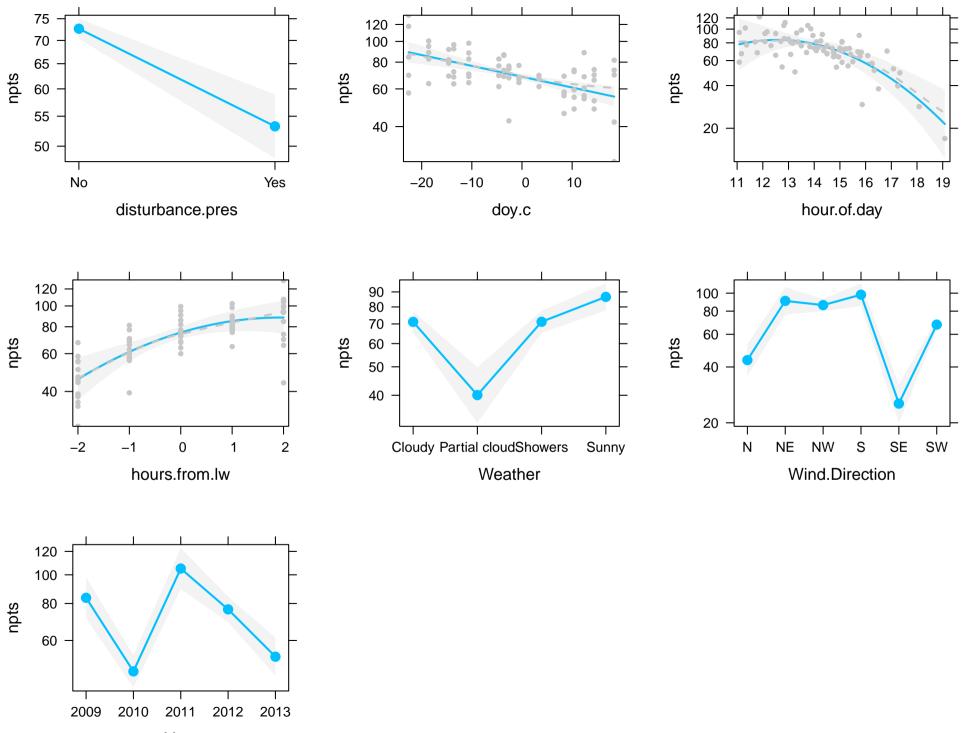
Kinvara Bay glmm, AICc 733.4



Mannin Bay glmm, AICc 471.9

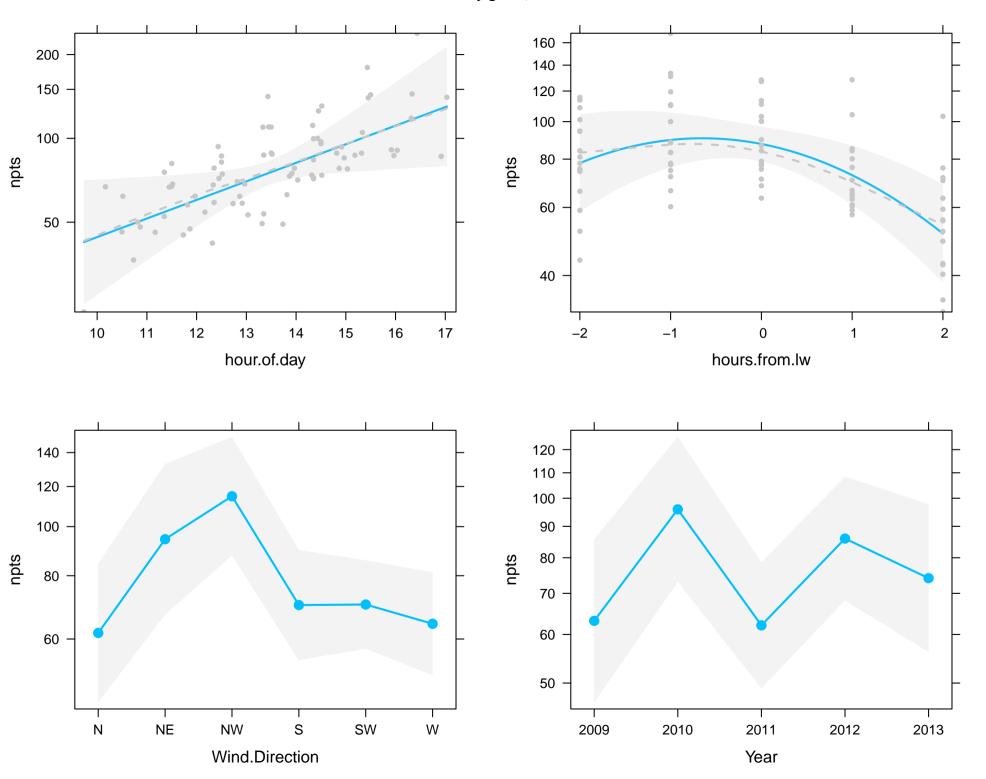


Moy Estuary glmm, AICc 692.3

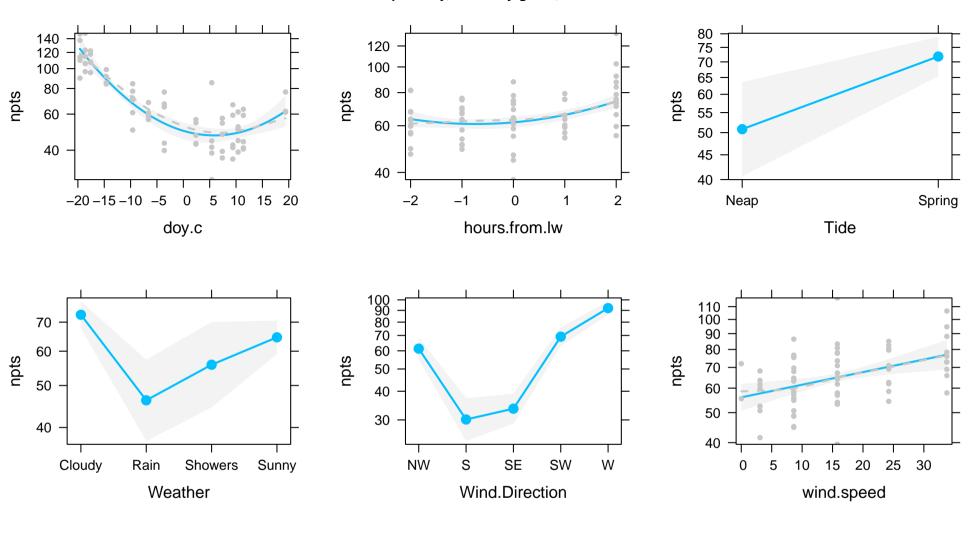


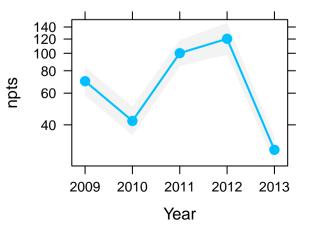


Oranmore Bay glmm, AICc 688.8

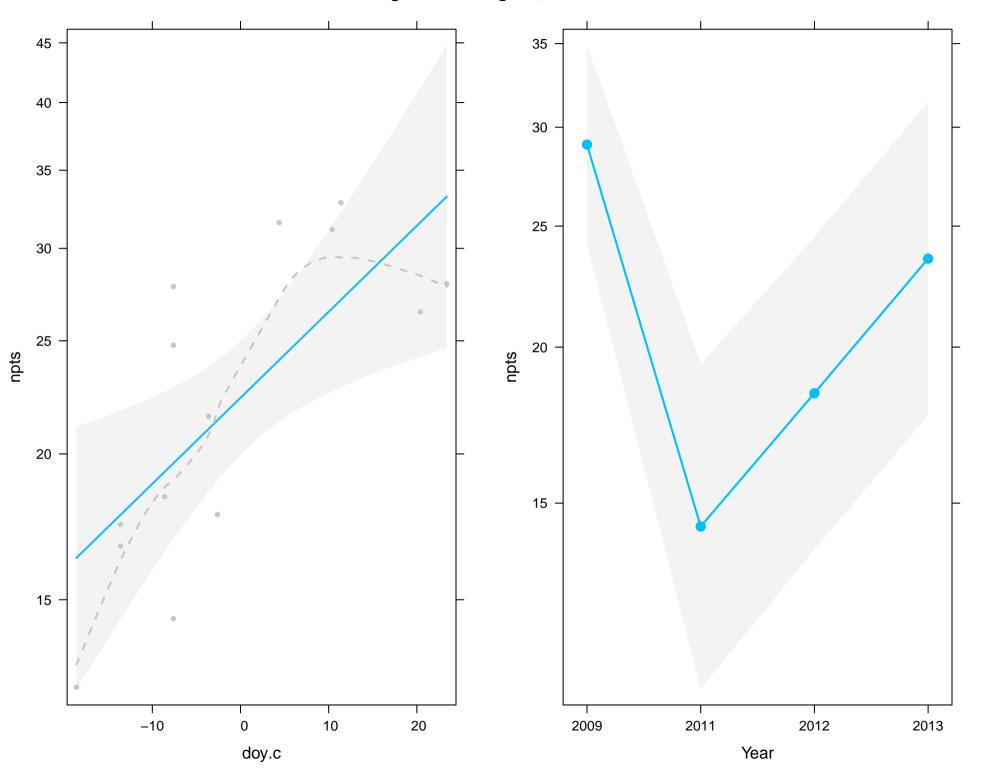


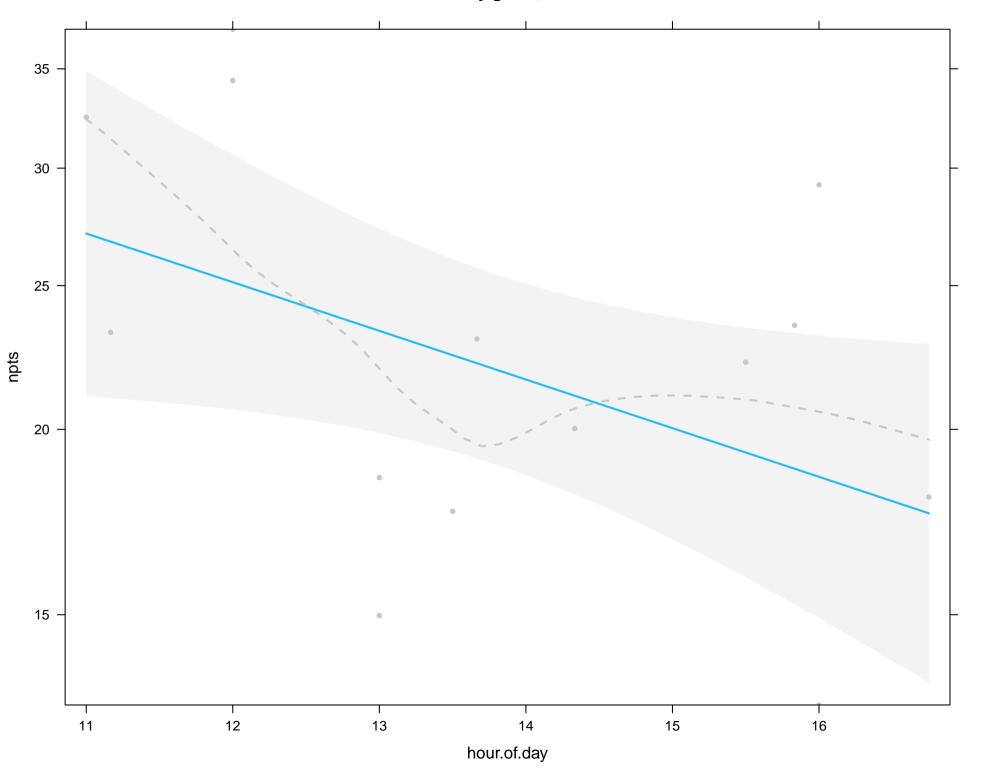
Westport Bay, Clew Bay glmm, AICc 598.9





Adrigole Harbour glmm, AICc 106.2





Donegal Bay glmm, AICc 980.3

