

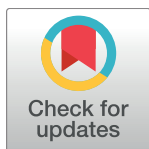
RESEARCH ARTICLE

Predicting forest insect flight activity: A Bayesian network approach

Stephen M. Pawson^{1*}, Bruce G. Marcot², Owen G. Woodberry³

1 Scion, Riccarton, Christchurch, New Zealand, **2** US Forest Service, Pacific Northwest Research Station, Portland, Oregon, United States, **3** Bayesian Intelligence, Monash University LPO, Clayton, Melbourne, Victoria, Australia

* steve.pawson@scionresearch.com



Abstract

Daily flight activity patterns of forest insects are influenced by temporal and meteorological conditions. Temperature and time of day are frequently cited as key drivers of activity; however, complex interactions between multiple contributing factors have also been proposed. Here, we report individual Bayesian network models to assess the probability of flight activity of three exotic insects, *Hylurgus ligniperda*, *Hylastes ater*, and *Arhopalus ferus* in a managed plantation forest context. Models were built from 7,144 individual hours of insect sampling, temperature, wind speed, relative humidity, photon flux density, and temporal data. Discretized meteorological and temporal variables were used to build naïve Bayes tree augmented networks. Calibration results suggested that the *H. ater* and *A. ferus* Bayesian network models had the best fit for low Type I and overall errors, and *H. ligniperda* had the best fit for low Type II errors. Maximum hourly temperature and time since sunrise had the largest influence on *H. ligniperda* flight activity predictions, whereas time of day and year had the greatest influence on *H. ater* and *A. ferus* activity. Type II model errors for the prediction of no flight activity is improved by increasing the model's predictive threshold. Improvements in model performance can be made by further sampling, increasing the sensitivity of the flight intercept traps, and replicating sampling in other regions. Predicting insect flight informs an assessment of the potential phytosanitary risks of wood exports. Quantifying this risk allows mitigation treatments to be targeted to prevent the spread of invasive species via international trade pathways.

OPEN ACCESS

Citation: Pawson SM, Marcot BG, Woodberry OG (2017) Predicting forest insect flight activity: A Bayesian network approach. PLoS ONE 12(9): e0183464. <https://doi.org/10.1371/journal.pone.0183464>

Editor: Daniel Doucet, Natural Resources Canada, CANADA

Received: November 28, 2016

Accepted: August 4, 2017

Published: September 27, 2017

Copyright: This is an open access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the [Creative Commons CC0](https://creativecommons.org/licenses/by/4.0/) public domain dedication.

Data Availability Statement: All Bayesian networks are supplied via the ABNMS BN repository (<http://www.abnms.org/bnrepo/>). These are the relevant links: Hylurgus Flight Activity: <http://www.abnms.org/bnrepo/bn?bnId=191>; Hylastes Flight Activity: <http://www.abnms.org/bnrepo/bn?bnId=192>; Arhopalus Flight Activity: <http://www.abnms.org/bnrepo/bn?bnId=193>. The minimal data set for this study can also be accessed from Dryad (doi:[10.5061/dryad.2j1f3](https://doi.org/10.5061/dryad.2j1f3)).

Funding: This work was funded by the New Zealand Ministry for Business, Innovation, and

Introduction

Understanding the abiotic factors that control flight behavior may allow the prediction of flight activity as a function of forecast weather conditions. Such modelling could provide an assessment of the potential for colonization by dispersing bark beetles or wood borers of recently harvested *Pinus radiata* D. Don logs. Understanding the potential for colonization is an important step toward a risk-based approach to managing the phytosanitary requirements for wood exports [1].

New Zealand has 1.7 million ha of managed plantation forests, predominantly (90%) single species, even-age stands of the exotic *P. radiata* [2]. Exotic wood boring and bark beetles are

Employment (C04X1204) and the Stakeholders in Methyl Bromide Reduction Inc. An Agmardt grant enabled the participation of Bruce Marcot with in-kind contribution through U.S. Forest Service, Pacific Northwest Research Station. The manuscript was reviewed by STIMBR prior to publication. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

sometimes intercepted at New Zealand's borders, and some have eluded detection, subsequently established, and are now spread widely throughout the country [3]. Three colonizers that are now abundant in plantation forests are two species of bark beetles (Scolytinae), *Hylurgus ligniperda* Fabricius and *Hylastes ater* (Paykull), and one wood borer (Cerambycidae), *Arhopalus fesus* (Mulsant). All three species are saprophytes that colonize recently dead or dying trees [4–6].

Seasonal flight activity has been extensively studied, particularly for *H. ligniperda* and *H. ater*. Results vary, likely due to the influence of local resources and differences in habitat and abiotic conditions between continents. Reay *et al.* [7] reported spring and autumn flight activity peaks of *H. ligniperda* in New Zealand's central North Island, and Tribe [8] observed an autumnal (April/May) peak in flight activity (April/May) in South Africa, whereas Mausel *et al.* [9] report a single peak of flight activity in spring (September/October) in two regions of Chile and a bi-modal activity pattern in a third region. Comprehensive daily sampling of *H. ligniperda* across eight regions in New Zealand indicate the existence of both unimodal and bimodal phenology patterns (Scion, unpublished data). Reay *et al.* [7] report a single autumn peak of *H. ater* flight activity in New Zealand, whereas Mausel *et al.* [9] observed both spring and autumn peaks in Chile. Reay *et al.* [7] acknowledged the potential of missing spring flight activity as their sampling began in October. Similarly Sopow *et al.* [10] report a strong autumn peak of activity, but like Reay *et al.* [7] their sampling did not encompass early spring conditions. Recent sampling confirms a strong late summer/autumn emergence of *H. ater*, with smaller spring flights in some regions in late August/September (Scion, unpublished data). *Arhopalus fesus* has a single period of flight activity in New Zealand that begins in late October and concludes in late April (Scion, unpublished data).

Despite the attention devoted to the seasonal phenology of these species, little is known regarding the abiotic factors that determine intraday flight activity patterns. Both *H. ligniperda* and *H. ater* reportedly have a generally crepuscular (active during the hours of dawn and dusk); In New Zealand Kerr *et al.* [11] sampled every three hours over two independent 72-hour periods and found activity peaks for both species from 07:30 to 10:30 and again at 19:30 to 22:30. Similar crepuscular flight patterns have been observed in other Scolytidae, e.g., *Dendroctonus brevicornis* LeConte in California, USA [12], *Pityophthorus juglandis* Blackman in California, USA [13] and *Hylastes nigrinus* (Mannerheim) in Oregon, USA [14]. However, patterns can be seasonal; e.g., in Israel *Orthotomicus erosus* (Wollaston) and *Pityogenes calcaratus* (Eichhoff) were crepuscular in spring and summer but had a unimodal midday peak of flight activity in winter [15].

Arhopalus fesus is a nocturnal species; small outdoor cage studies conducted in New Zealand using an infrared movement detector indicated peak activity from 21:00 to 24:00, with some evidence of an additional peak in male activity at dawn in New Zealand [16]. Field trapping in New Zealand at intervals of three hours corroborates these captive studies with maximum trap catch of *A. fesus* occurring in 19:30 to 22:30 and 22:30 to 1:30 sampling periods [11]. To date, the abiotic factors that govern the initiation of flight behavior of *H. ligniperda*, *H. ater*, and *A. fesus* have not been quantified. However, morning flight activity of *H. ligniperda* and *H. ater* was reported by Kerr *et al.* [11] to be coincident with increasing temperature.

Few studies have quantitatively assessed the influence of abiotic factors on the flight of bark beetles and wood borers. Temperature is frequently cited as a strong predictor of flight activity, however, light intensity, wind speed, rainfall, and relative humidity also are reported to influence forest insect flight activity [12, 14, 17, 18]. Although Chen *et al.* [13] observed strong individual relationships between *P. juglandis* flight activity and temperature, barometric pressure, light intensity and wind speed, the probability of flight was influenced both collectively and interactively between these four factors [13]; hence, complex interactions make the prediction

of flight activity challenging. Such factors may also interact with insects' chemical host finding cues and subsequently influence trap sensitivity, thus biasing interpretation of flight behavior. For example, low wind speeds could prevent the formation of chemical plumes, precluding effective orientation towards traps by flying insects [18].

Empirical models that quantify the potential risk of infestation by forest insects are needed to develop a risk management framework to guide future application of phytosanitary treatments for export wood products. To support such a framework we developed empirically-based Bayesian network (BN) models to predict the flight activity of three exotic forest insects *H. ligniperda*, *H. ater*, and *A. ferox* in New Zealand as a function of meteorological and temporal factors in recently clear-cut *P. radiata* plantations. BNs are graphical networks of variables connected by logical, correlational, or causal relationships quantified by conditional probability tables [19]. They are used widely in ecological and environmental sciences for diagnosis, forecasting, prediction, and other applications [20], including modelling of invasive invertebrates [21]. We chose to use BNs because they express conditions and results as probabilities, particularly for depicting errors of false presence and false absence. This lends BNs to risk analysis and risk management applications [22]. Further, BNs are robust to missing data and to conditions of data multicollinearity and nonlinearity that otherwise violate assumptions in more traditional multivariate modelling approaches. BNs also can be structured from a combination of empirical data and expert knowledge, and can be updated with new information as it comes to hand [23].

Methods

Study sites and insect trapping

Four sites were established in *P. radiata* plantation forests in Canterbury, New Zealand (S1 Fig, S1 Table). One meteorological station and three flight activity traps were installed at each site. Traps were located equidistant in a circle 40 m from the meteorological station. Each trap consisted of an ethanol and alpha-pinene baited black colored flight intercept panel trap with an electronically controlled stepper motor that rotated a circular carousel to separate trap catch into individual plastic containers (6286PTCL—SQ PET JAR 58MM 233ML, Stowers, New Zealand) on an hourly basis (S2 Fig). The panel trap and plastic containers were coated with an insecticide (cypermethrin), applied at the start of each trapping period, to kill insects and to minimize the potential for any movement between plastic containers by captured individuals. All traps at each site were established two weeks before sampling was initiated to ensure any trap establishment effects, e.g., soil disturbance, did not influence trap efficacy. The trial was run over three non-consecutive periods, with the first in spring (23 September and 14 October 2014) and then twice in summer (19 November to 18 December 2014, and 13 January to 11 February 2015). Each of the focal species has a slightly different phenology, hence sampling periods were spread to encompass periods of known flight activity for each species. Traps were maintained on a daily basis, with insects identified and counted in the field or brought to the laboratory when large numbers were present in a given sample hour. If a trap fault occurred during a 24-hour period, all affected sampling periods were discarded.

Variables affecting flight activity

A summary of variables used to model flight activity is provided in Table 1. Meteorological data were collected using a 2.5-m metal tower (Scottech, Hamilton, New Zealand). Data from sensors were recorded on a CR1000 (Campbell Scientific, Logan, USA) data logger with measurements taken every minute. Sensors included RM Young wind monitor, model 05103 (RM Young Company, Michigan, USA); Apogee quantum sun calibration sensor, model sq-110

Table 1. Temporal and abiotic variables used to model the probability of insect flight.

Variable name	Description	Units of measure
<i>Temporal</i>		
Day of year	Day of year as integer	Integer
Time since sunrise	Elapsed time between the start of the trapping hour and sunrise. Corrected for geospatial differences in sunrise	Integer (minutes)
Time since sunset	The elapsed time between the start of the trapping hour and sunset. Corrected for geospatial differences in sunset	Integer (minutes)
<i>Abiotic</i>		
Maximum hourly temperature	Maximum temperature within the sample hour	°C
Photon flux density	Average radiant flux of solar radiation within the wavelengths of 400 to 700 nm	μmol photons m ⁻² s ⁻¹
Rainfall	Accumulated rainfall within the sample hour	mm
Relative humidity	Average humidity within the sample hour	%
Temperature range	Difference between maximum and minimum temperature within the sampling hour	°C
Wind speed	Average instantaneous wind speed per hour	m ⁻¹ s ⁻¹

<https://doi.org/10.1371/journal.pone.0183464.t001>

photosynthetic flux density sensor (Apogee Instruments, Logan, USA); CSI temperature and relative humidity probe, model hc2s3 (Campbell Scientific, Logan, USA); and CSI rain gauge, model tb4 (Campbell Scientific, Logan, USA). All meteorological variables were summarized as hourly average, maximum, and minimum values using the R-package xts [24], except rainfall that was calculated as an hourly sum. Time since sunrise and sunset was calculated using the sunriset function in R-maptools [25]. This calculates the time of sunrise (accuracy ± 1 minute) corrected for location and atmospheric refraction using the formula of Meeus [26].

To define the relationship between meteorological variables and forest insect flight activity, each was regressed across a series of discrete intervals defined using an unsupervised discretization of temperature data, as per the method of Blackburn *et al.* [27]. Variables were discretized into 20 intervals using the unsupervised Minimum Description Length (MDL) method in Weka (vers. 3.6) [28]. *Arhopalus fesus* flight activity was concentrated at low wind speeds, and to improve sensitivity to this we left-shifted the discretization to increase bin numbers at low wind speeds. The maximum observed insect flight activity in each interval was used to fit the relationship with individual meteorological variables. Outliers in models were identified as observations with Cook's distances > 1, using R-nlreg [29]. A Gaussian equation ($Y = k \times \exp(-1/2 * (\text{temperature} - \mu)^2 / \sigma^2)$) was used to fit the relationship between maximum interval flight activity and temperature for all species using the R-nls function. The adjusted R² for the non-linear models was calculated using Wherry's formula [30]. A negative exponential equation ($Y = y_0 * \exp(\text{wind speed}/b)$) was used to fit the relationship between wind speed and *A. fesus* using a generalized non-linear least square, R-nlme [31]. Generalized Additive Models (GAM) with a gamma distribution (inverse link) were used to fit the relationship between flight activity and wind speed, photon flux density (PAR), and relative humidity for *H. ligniperda*, using R-mgcv [32]. Graphical model validation tools were used to check the model assumptions of variance homogeneity and normality. The significance of individual GAM models was assessed by a likelihood ratio test of the model against the null, intercept only, model. The adjusted R² was extracted from the model summary. All non-linear modelling was done in R-Version 3.2.2 [33].

Bayesian network model development and calibration

We developed BN structures for each species using supervised machine learning techniques. The objective of supervised learning is to develop a probabilistic classifier model of a target

variable (here, flight activity) based on predictor observations (temperature, time of day, etc.). Classifier models provide a predicted probability distribution over a set of classes (in our case, true or false flight activity), given a set of inputs. The process of developing a BN from data involves three stages: discretization of continuous variables into range states; identification of relationships between variables (that is, the network structure); and parameterization of the probability tables.

For each model training run, the continuous meteorological variables were discretized using a supervised approach relative to the target variable (flight activity) to allocate discrete intervals for subsequent Bayesian network modelling. Using a supervised discretization (the Fayyad *et al.* [34] Minimum Description Length (MDL) method in Weka version 3.6 [28]). The supervised discretization process differs from that used above for the discretization of meteorological variables by providing an information theoretic metric that minimizes the joint entropy between the continuous variable and target variable [34]. It selects a discretization that balances the trade-off between fit to data and model complexity (i.e., more data will justify a greater number of bins). For some variables, the discretization process did not identify useful splits because there was no relationship or the correlation was not strong enough in the data to justify the additional model complexity; in such cases the average value is returned and we excluded such variables from our BN models as they are not informative of flight activity.

We explored three procedures to generate the BN structure: naïve Bayes [35]; CaMML [36], which uses an information theoretic metric to infer causal structure; and Tree Augmented Naïve Bayes (TAN) [37], discussed in detail below. The final network selection was based on a comparison of classification accuracy. In all cases, the TAN structure performed best, and here we report only the results of these models, however models from each of the three approaches are available online at the Australasian Bayesian Network Modelling Society's BN repository. TAN structure learning involves augmenting the naïve Bayes structure (where all predictive variables are children of the target variable), with dependences between the predictive variables (using a tree structure) to represent correlations between them. TAN and naïve Bayes networks are designed specifically for model prediction and not for denoting causal linkages. We chose this approach because our main objective was to predict insect flight activity. Building TAN structures and conducting calibration and validation assessments of individual models were done using the GeNIe (www.bayesfusion.com/#!/genie-modeler/lf73d) Application Programming Interface (API).

Conditional probability tables of TAN models were learnt using the expectation maximization (EM) algorithm [38–40], available in Netica (norsys.com) and GeNIe (and most BN modeling tools). EM iterates to maximize the log likelihood of the BN given the training data, calculating conditional probability values from case data by integrating over missing values [41]. We then conducted sensitivity analysis on the models, using variance reduction [23], to determine the degree to which predictor variables account for flight probability. As more data are collected, it is possible to revise and improve the BN by updating the conditional probabilities or by relearning the BN discretizations and structure. We tested the degree to which each BN model was calibrated to the case files for each species by calculating 5 model performance indices (logarithmic loss and quadratic (Brier) loss, lower values of which denote better model performance; and spherical payoff, Gini coefficient, and area under the ROC curve, higher values of which denote better model performance) and 3 classification error rates (Type I or false positives, Type II or false negatives, and overall confusion error) [42]. Type I errors represented prediction of insect flight when there was no flight (i.e., false positive), Type II errors represented prediction of no flight when flight occurred (i.e., false negative), and overall confusion error is the total error rate over all outcomes. Because Type II error is more egregious in the context of biosecurity policy (wanting to ensure absence of pest insects), we developed a

novel ROC curve to depict Type II classification error rates at four probability thresholds of predicting no flight occurrence: 50, 90, 95, and 99%, whereas traditional ROC curves address false positives. We identified the best models on the basis of the lowest type II error rates and the five performance metrics.

Model validation

One possible concern with calibrating BN structures and parameters to case files by using machine-learning algorithms such as EM is with overfitting the model to the cases observed [43]. Several approaches have been suggested for testing and avoiding overfitting in BNs [44, 45] but none specifically address Type II errors. We devised a novel approach using cross-validation [46] that directly addresses our central concern of minimizing Type II errors as follows. We ran 100 iterations of 4-fold cross-validation analysis, each iteration randomizing the training and testing samples, and relearning the model discretizations, structure, and parameters (see above). We initially ran 3-, 4-, and 5-fold cross validations on each species model and found that error results were nearly identical for each fold, so we settled on 4-fold as the best balance in sample size between training and testing subsets. We recorded the Type I, Type II, and overall confusion error rates for each iteration across a range of probability thresholds from 0 to 1 in 0.001 increments, and then calculated the average, standard deviation, minimum, and maximum error values across all iterations. In addition, we also calculated the number of standard deviations between the calibration and the average of the 100 runs of 4-fold cross validation and plotted the outcome as a means of visualizing any bias between calibration and validation results.

Model influence runs

We conducted influence runs [47] to determine the maximum influence that each covariate can have on the calculated posterior probability of flight in our model. This was performed by sequentially selecting each state of the input variable, updating the BN, and recording the resulting target variable's posterior probability. Influence runs quantify the range of effects that each predictor variable set can have on the target outcome, whereas sensitivity analysis tests variance reduction effects from incremental changes of each predictor variable from their normative prior probability value settings.

Results

Trap results

A total of 10,135 *H. ligniperda*, 224 *H. ater*, and 553 *A. ferus* were trapped during the course of 7,144 individual hours of sampling spread across 78 non-continuous days at the four sites. In total there were 759 (11%), 141 (2%), and 197 (3%) positive trap catch hours of *H. ligniperda*, *H. ater*, and *A. ferus* respectively. Average catch rates per trap varied between the three sample periods. *Hylurgus ligniperda* flight activity was more prevalent in spring and early summer, *H. ater* more in late summer, and *A. ferus* most pronounced in the second summer sampling period (Fig 1).

Although insecticide was used in trapping containers to reduce the potential for internal movement between containers, 9 individuals were clearly inconsistent with the other 7,144 hourly trap samples and the known behavior of these species. Movement between containers is the most parsimonious explanation. Specific anomalous catches of note were *H. ligniperda* activity at -0.7 and 3.6°C that are substantially less than all other cases of positive *H. ligniperda* flight activity (S3 Fig) and 7 instances of *A. ferus* activity recorded during the day when photon

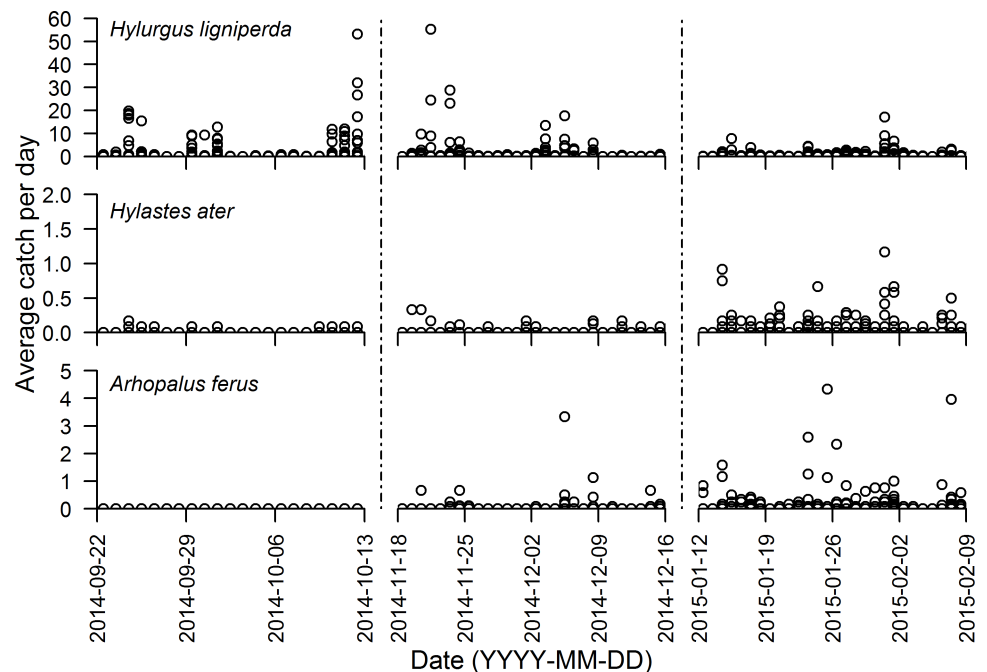


Fig 1. Hourly catch per trap averaged across all traps at the four study sites. Separate panels indicate the three discontinuous time periods when sampling was undertaken.

<https://doi.org/10.1371/journal.pone.0183464.g001>

flux density exceeded $500 \mu\text{mol m}^{-2} \text{s}^{-1}$ (S4 Fig). These discrepancies were removed from non-linear regression and BN modeling; however, for completeness they are shown in the overall capture outcomes (Figs 1 and 2, S3–S5 Figs). Trap catch of *H. ligniperda* and *H. ater* indicated a generally crepuscular flight activity pattern, with *H. ater* exhibiting a more pronounced decline in midday flight activity than *H. ligniperda* (Fig 2). *Arhopalus fesus* was clearly nocturnal with peak activity occurring within the first hour post sunset (Fig 2).

Environmental correlates

Average hourly values of each abiotic variable across all sites and time periods are summarized in S6 Fig. Minimum average hourly temperature across all days and sites occurred just before sunrise and peaked at midday. Wind speed was correlated with temperature ($r = 0.27$, $n = 14,138$, $P < 0.001$) and peaked in early afternoon. Relative humidity was inversely correlated with temperature ($r = -0.66$, $n = 14,138$, $p < 0.001$). Photon flux density, as a proxy for light intensity, was unimodal with respect to time of day with peak intensity at midday. Hourly accumulated rainfall was the most variable abiotic factor between sites (greatest standard error of the mean, S6 Fig) and showed no consistent daily pattern.

A Gaussian relationship fitted the observed flight activity of *H. ligniperda* ($k = 83.18$, $\mu = 18.57$, $\sigma = 3.40$), *H. ater* ($k = 3.16$, $\mu = 18.29$, $\sigma = 3.46$), and *A. fesus* ($k = 10.16$, $\mu = 17.28$, $\sigma = 3.07$) as a function of maximum hourly temperature (S3 Fig). Peak flight activity occurred between 17 and 18°C for all species, with flight activity observed at a minimum hourly maximum temperature of 6.3 , 6.3 , and 6.4°C for *H. ligniperda*, *H. ater*, and *A. fesus* respectively (excluding outliers discussed above). No flight activity was observed when maximum temperature within an hour exceeded 31.6 , 26.3 , and 24.7°C for *H. ligniperda*, *H. ater*, and *A. fesus* respectively.

A generalized additive model with a Gamma distribution (Table 2), and a negative exponential model ($y_0 = 11.93$, $b = -3.65$), best described the relationships between flight activity

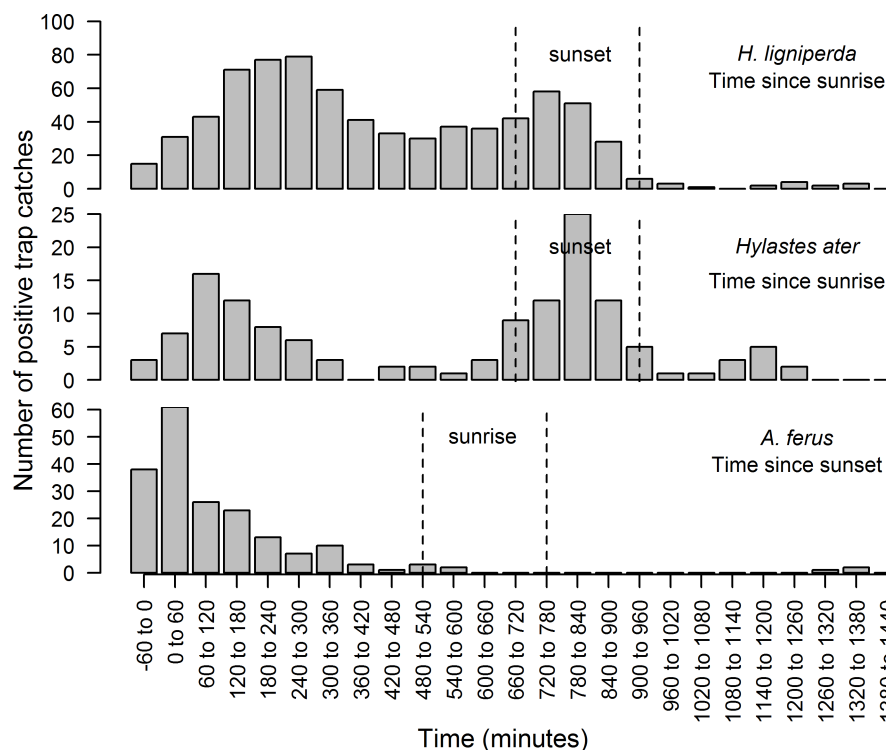


Fig 2. Number of positive trap catch hours as a function of time since sunrise in hourly bins for *H. ligniperda* and *H. ater* and for time since sunset for *A. ferus*. Because day length varied as a function of the day of the year during the study a range is provided that encompasses the period when sunrise or sunset occurred. Dashed lines indicate the period where sunset occurred as a function of time since sunrise for *H. ligniperda* and *H. ater*. Similarly for the nocturnal *A. ferus* these dashed lines indicate the period when sunrise occurred as a function of time since sunset.

<https://doi.org/10.1371/journal.pone.0183464.g002>

and wind speed, for *H. ligniperda* and *A. ferus*, respectively (S5 Fig). Observed flight activity of *H. ligniperda* was low at very low wind speeds, but increased quickly with rising wind speed and peaked at 2 m s^{-1} . *Arhopalus ferus* was capable of active flight at 15 m s^{-1} , however most flight activity of all species occurred at wind speeds $< 10 \text{ m s}^{-1}$. No clear relationship was observed between the upper bound of *H. ater* flight activity and wind speed.

Table 2. Results from the final General Additive Models (GAMs) for the flight activity of *H. ligniperda*. GAMs have a parametric component and a smoothing part, hence the distinction between parametric coefficients and the smoothing terms. $s()$ = smooth term for a continuous variable, SE = standard error of the estimate, t = t -statistic, P = P -value, edf = estimated degrees of freedom, F = F -statistic. Wdspd = Wind speed, PAR = Photon flux density, and RH = Relative humidity. Significant values are denoted with $P < 0.05$ = *, $P < 0.01$ = **, $P < 0.001$ = ***.

Parametric coefficients	Estimate	SE	t	P	
Intercept					
Wdspd	0.14	0.020	7.01	< 0.001	***
PAR	0.04	7.5×10^{-3}	5.63	< 0.001	***
RH	0.09	0.015	5.51	< 0.001	***
Approximate significance of smoother terms	edf	F		P	
$s(\text{Wdspd})$	4.10	8.80		0.016	*
$s(\text{PAR})$	3.76	3.83		0.022	*
$s(\text{RH})$	4.05	4.59		0.010	*

<https://doi.org/10.1371/journal.pone.0183464.t002>

Table 3. Description of the best model for each species.

Model	Number of nodes	Number of linkages	Total number of probabilities
<i>H. ligniperda</i>	10	17	190
<i>H. ater</i>	7	11	106
<i>A. ferus</i>	7	11	104

<https://doi.org/10.1371/journal.pone.0183464.t003>

A generalized additive model with a Gamma distribution best fitted the relationship between *H. ligniperda* and photon flux density and relative humidity (Table 2, S4 Fig). Peak flight activity of *H. ligniperda* occurred at very low and intermediate photon flux densities, and at 42% relative humidity. No clear relationship was observed between the upper bound of *H. ater* and *A. ferus* flight activity and either photon flux density or relative humidity (S4 Fig).

Bayesian network models

Bayesian network models are summarized (Table 3) and the best model for each species (Fig 3) was selected by lowest total classification error rate. The probability of *H. ligniperda* flight was most sensitive (comprising at least half of all model sensitivity) to maximum hourly temperature and time since sunrise (Table 4). The probability of *H. ater* flight was most sensitive to time since sunrise and the time of year (Day of year) (Table 4). The probability of *A. ferus* flight was most sensitive to time since sunset (Table 4).

Model calibration

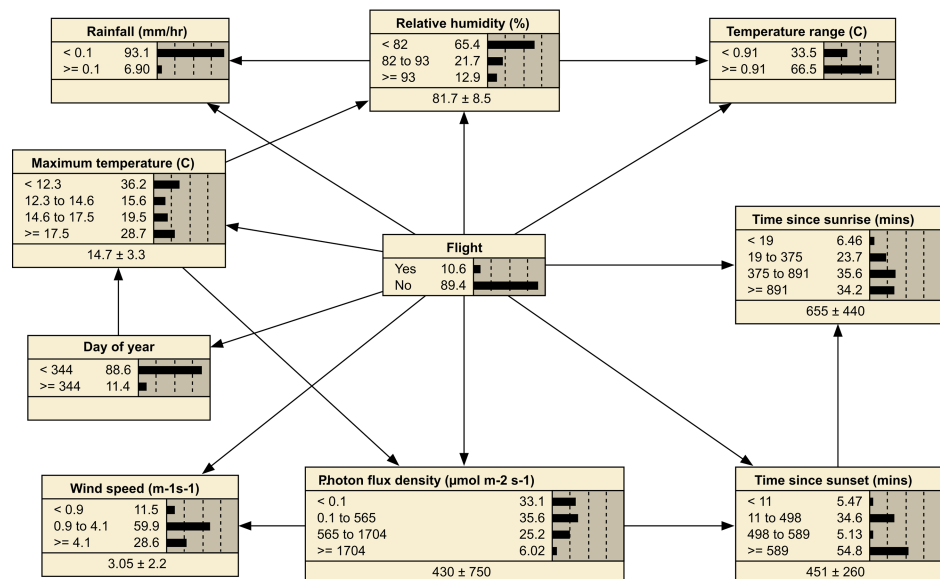
BN model calibration results (Table 5) suggest better model fit for *H. ater* and *A. ferus* than for *H. ligniperda*, in particular with the latter showing greater values of logarithmic loss and quadratic loss and slightly lower values of spherical payoff, all denoting poorer model fit. *Arhopalus ferus* had better model fit denoted by higher values of the Gini coefficient and area under the ROC curve (Fig 4). However, consistently across all model thresholds denoting no flight, the *H. ligniperda* model had highest Type I and the lowest Type II errors, and the *H. ater* model had highest Type II errors. Overall, calibration results suggested that the *H. ater* and *A. ferus* models had the best fit for low Type I and overall errors, and *H. ligniperda* had the best fit for low Type II errors.

As expected, across the four model thresholds denoting no flight, Type I errors increased and Type II errors decreased, and overall errors increased. At the most stringent criterion of no flight—that is, when the "no flight" state in the output node achieved $\geq 99\%$ probability—the *H. ligniperda* model performed best with minimal Type II errors, but at the cost of a high Type I error rate. The *A. ferus* model performed second best with a low Type II error rate but a higher Type I error, and the *H. ater* model performed worst with high Type II and Type I errors. No species model achieved $<15\%$ simultaneously for Type I and II error rates across the four model thresholds.

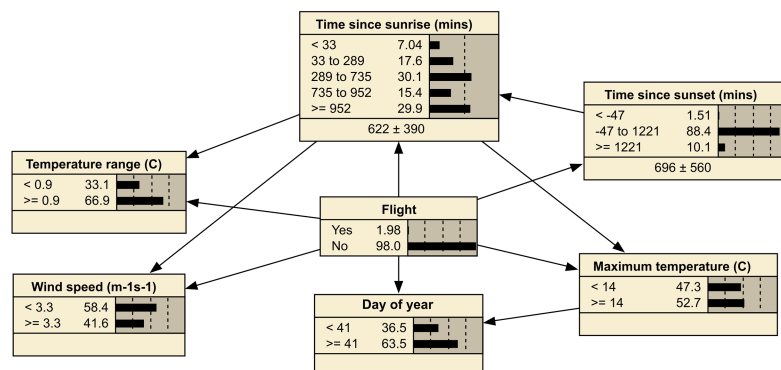
Model validation

We compared validation to calibration in a novel way. Cross-validation resulted in patterns similar to the calibration results (Table 5, Fig 5) with *H. ligniperda* again showing lowest Type II error rates and *H. ater* showing highest, and no species model achieving $<17\%$ for both Type I and II error rates across the four model thresholds. In general, Type I and II error rates were complementary; reducing one increases the other. This pattern is seen also in comparing calibration and cross-validation results (the green line in Fig 5), where high concordance (low standard deviation difference between the two results) occurred generally with one error type

Hylurgus ligniperda



Hylastes ater



Arhopalus ferus

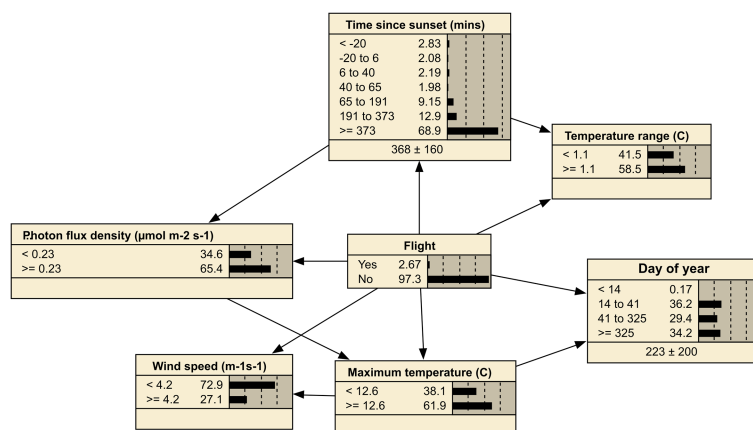


Fig 3. Bayesian networks models of flight activity of three forest insect species.

<https://doi.org/10.1371/journal.pone.0183464.g003>

but not the other. False positive rates calculated from the all-data calibration (the red line in Fig 5, left column) generally coincided well with the false positive rates calculated from cross-validation, for all three species. False negative rates from calibration (the red line in Fig 5, right column) coincided well with cross-validation for *H. ligniperda* but less so for the other two species, that is, where calibration error rates mostly were less than cross-validation error rates, suggesting some degree of model overfitting.

Influence runs

Influence runs (Fig 6) suggest that a number of variables are individually associated with decreased flight probability of all three species. However, only a few variables have major influence on increased flight probability: increasing maximum hourly temperature ($\geq 17.5^\circ\text{C}$) and moderate time since sunrise (11–498 min) for *H. ligniperda*; shorter time since sunset (sampling hours that began <47 minutes before actual sunset) for *H. ater*; and lower days of the year (<14) and shorter time since sunset (sampling hours that began 20 min before to 6 min after actual sunset) for *A. ferus*. Other variables also are associated with increased flight probability for all three species, although to lesser degrees.

Table 4. Sensitivity analysis of the Bayesian network models (Fig 3), showing degree of sensitivity of insect flight probability to each predictor variable.

Variable	Mutual Information	Percent	Variance of Beliefs
<i>Hylurgus ligniperda</i>			
Maximum hourly temperature	0.07298	14.9	0.008755
Time since sunrise	0.06282	12.9	0.006895
Time since sunset	0.05696	11.7	0.005719
Photon flux density	0.04507	9.23	0.005316
Relative humidity	0.03162	6.48	0.003277
Temperature range	0.01956	4.01	0.002218
Average wind speed	0.01171	2.4	0.00139
Day of year	0.00813	1.66	0.000802
Rainfall	0.00374	0.766	0.00038
<i>Hylastes ater</i>			
Time since sunrise	0.0104	7.42	0.0003
Day of year	0.00939	6.7	0.000262
Time since sunset	0.00466	3.33	0.000205
Maximum hourly temperature	0.00753	5.37	0.000179
Wind speed	0.00331	2.36	8.08E-05
Temperature range	0.00209	1.49	4.97E-05
<i>Arhopalus ferus</i>			
Time since sunset	0.05908	33.3	0.004525
Photon flux density	0.03249	18.3	0.001138
Day of year	0.02303	13	0.001088
Maximum hourly temperature	0.00833	4.69	0.000246
Temperature range	0.0029	1.64	0.000106
Wind speed	0.00263	1.48	8.09E-05
Time since sunrise	0.00209	1.49	4.97E-05

<https://doi.org/10.1371/journal.pone.0183464.t004>

Table 5. Summary of BN model performance. Model performance is assessed at different predictive thresholds with both calibration (entire dataset) and validation (4-fold cross validation) results presented.

Metric	<i>Hylurgus ligniperda</i>				<i>Hylastes ater</i>				<i>Arhopalus ferus</i>			
Logarithmic loss	0.244				0.074				0.068			
Quadratic loss	0.148				0.035				0.038			
Spherical payoff	0.918				0.982				0.979			
Gini coefficient	0.737				0.738				0.915			
Area under ROC	0.870				0.868				0.958			
Model threshold predicting no flight	50	90	95	99	50	90	95	99	50	90	95	99
Calibration												
Type I error rate (%)	3	31	39	56	0	5	8	30	1	5	9	15
Type II error rate (%)	73	12	6	2	100	48	38	12	59	23	14	6
Total error rate (%)	10	29	35	50	2	6	8	29	3	6	9	14
Validation												
Average type I error rate (%)	2	30	37	55	<1	3	10	37	<1	6	9	17
Average type II error rate (%)	80	17	10	4	100	81	51	22	71	30	20	10
Average total error rate (%)	10	28	35	49	2	5	11	36	3	6	10	17

<https://doi.org/10.1371/journal.pone.0183464.t005>

Discussion

Chen *et al.* [13] highlight the importance of first and second order interactions among multiple environmental factors as regulators of flight activity in *P. juglandis*, the invasive walnut twig borer. Our BN models provide an alternative approach to model the effects of multiple interacting abiotic factors (both meteorological and temporal) on insect flight activity, and thereby

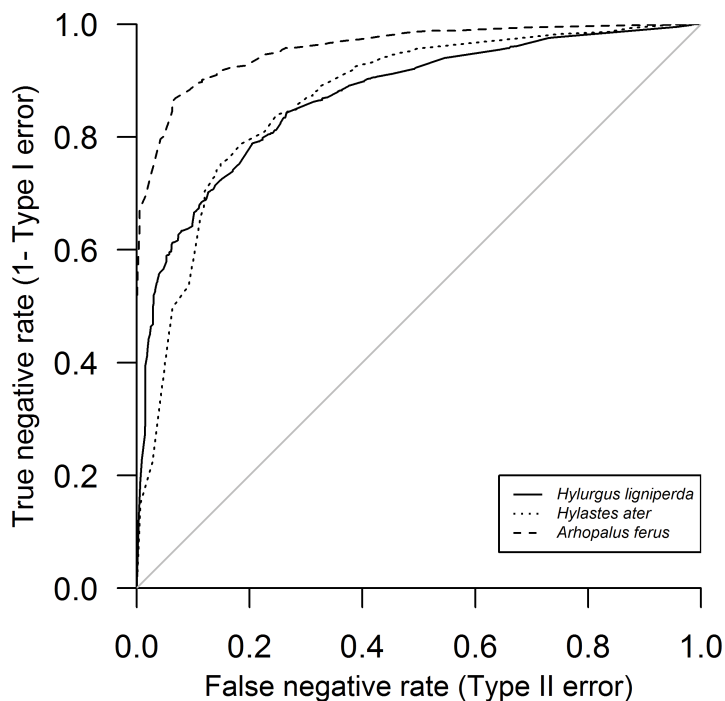


Fig 4. Modified receiver operating curve (ROC) showing model predictions of the true negative state as a function of the type II (false negative) error rate.

<https://doi.org/10.1371/journal.pone.0183464.g004>

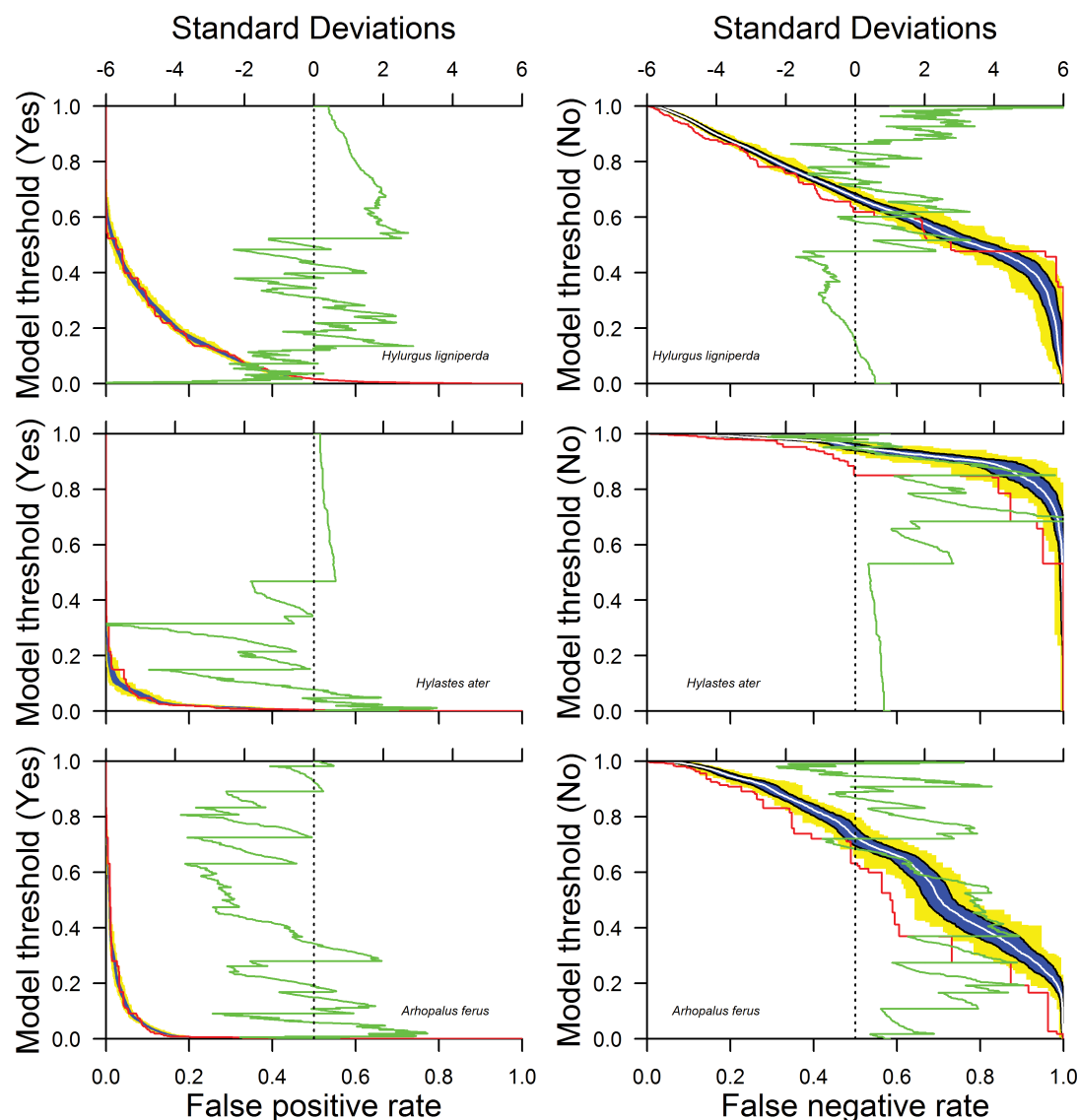


Fig 5. False positive (Type I error) rates as a function of the threshold required for the model to predict 'Yes', and false negative (Type II error) rates as a function of the threshold required for the model to predict 'No' for *H. ligniperda*, *H. ater*, and *A. ferus*. Red lines indicate the relationship for the calibration dataset (i.e., full casefile), blue shading indicates the range of the first standard deviation for 100 runs of 4-fold cross validation, with the inner white line denoting mean outcomes. Yellow indicates the maximum and minimum values observed during those 100 runs. The green curve represents the number of standard deviations between the calibration and the average of the 100 runs of 4-fold cross validation.

<https://doi.org/10.1371/journal.pone.0183464.g005>

determining the overall and individual explanatory power of predictor variables, although results varied amongst species.

In our best performing BN models, *Hylurgus ligniperda* flight activity was most sensitive to maximum hourly temperature, photon flux density, and time since sunrise and sunset (Table 4). Kerr *et al.* [11] also observed temporal patterns with distinct morning and evening peaks of *H. ligniperda* and *H. ater* activity during short-term (two periods of three-hourly sampling over three days) summer monitoring periods. Similarly, meteorological factors and time of day were key determinants of the *P. juglandis* flight activity, as reported by Chen *et al.* [13]. Such concordance highlights the need to consider multivariate models to predict flight activity patterns.

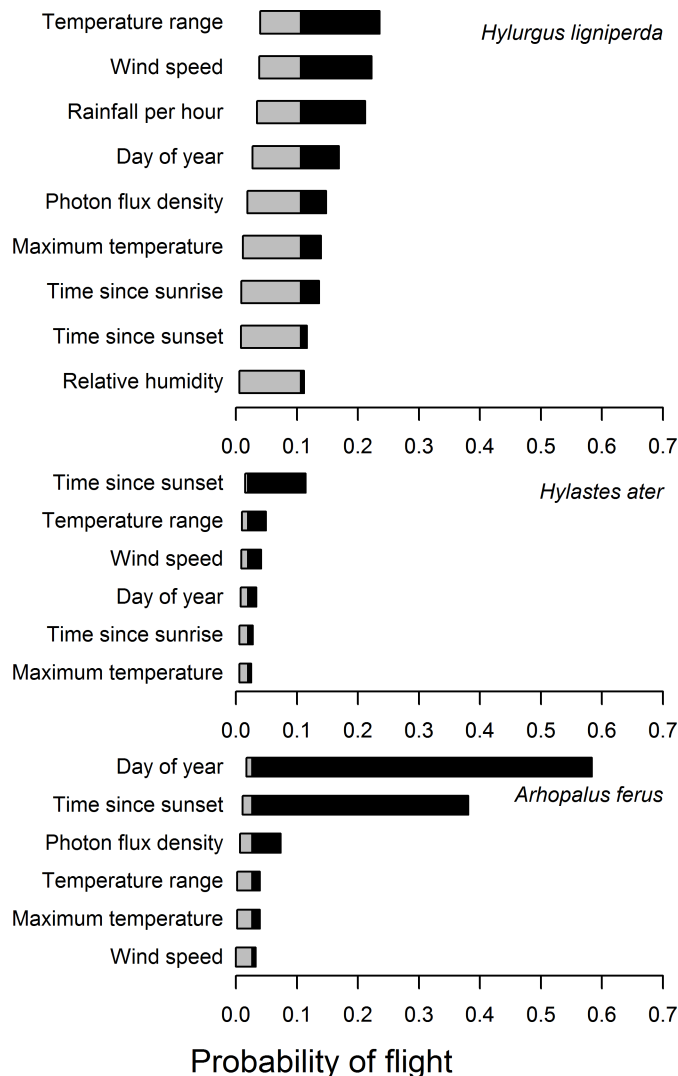


Fig 6. Model influence runs. Presented as tornado diagrams [48] that have a grey and a black portion, to the left and to the right, of the expected ("normative") outcome where all other covariates are set to their prior probability (expected, normative) distributions. Gray portions of the bars show the variables potential range influence on reducing flight probability; black portions show its potential range of influence on increasing flight probability. Variables are sorted in order of decreasing overall influence for each model.

<https://doi.org/10.1371/journal.pone.0183464.g006>

Sensitivity of *H. ater* and *A. ferus* flight activity to meteorological variables was low compared to the effect of time since sunrise and sunset and the day of the year, although *H. ater* flight activity was moderately sensitive to maximum hourly temperature (Table 4). Short-term activity studies of individual caged cerambycids have demonstrated a range of temporal patterns including diurnal activity, e.g., *Callidiellum rufipenne* (Motchulsky) [49] and *Gaurotes virginea* (L.) [50], nocturnal activity, e.g., *A. ferus* [16], and cathemeral activity, e.g., *Semanotus bifasciatus* (Motchulsky) [49] and *Nadezhdiella cantori* (Hope) [51]; however more precise prediction of flight as a function of meteorological factors is currently lacking. The overall low sensitivity of flight activity of *H. ater* and *A. ferus* to meteorological variables in our study may reflect (1) the small fraction of positive flight activity hours within the calibration dataset, and (2) lower attractiveness of *H. ater* and *A. ferus* to alpha-pinene and ethanol baited panel traps

relative to *H. ligniperda*. Improving trap sensitivity is one of the most important options to reduce Type II modelling errors of flight activity for these species (see uncertainties below). To test the effect of trap sensitivity on the performance of Bayesian network models requires additional datasets from multiple trap types. Model performance could then be evaluated and potentially enhanced by adding an additional node for trap type.

Uncertainty as useful information

Our models are built from field data collected from one region (Canterbury) of New Zealand. The robustness of applying the models to other regions likely depends on the similarity of environmental conditions (especially weather patterns and climate regimes) to the Canterbury field site. New Zealand has complex weather patterns with strong gradients in precipitation from west to east. The most similar meteorological regions to Canterbury are Marlborough, Wairarapa, Hawke's Bay and East Cape [52]; however, the robustness of models between regions would need to be verified by field sampling. The wetter regions down the west coasts of both islands are likely to be most dissimilar to Canterbury [52].

Prediction of flight activity is inherently complex with many interacting factors, e.g., temperature, light, and wind speed, influencing flight [53]. In such circumstances a BN modeling approach is advantageous in that it explicitly shows the degree of uncertainty in the prediction (in this case the flight activity outcome states), denoted as the distribution of posterior probabilities [42]. Because we can assess the implications of uncertainty and error as they propagate through the network, we can identify predictor variables least well understood that could have major influences on prediction outcomes. We identified such variables in our use of sensitivity tests and influence runs, which can be part of a broader assessment of the value of further information [54, 55] given monitoring and study costs and the degree to which they could reduce model error, particularly Type II error. Also, the BN models can be incrementally updated and improved by incorporating new data using the EM algorithm and, where warranted, by restructuring the networks.

Potential effect of sampling errors on model performance

Parameter uncertainty and prediction accuracy of our flight activity models depend on sampling bias, specifically the sensitivity and specificity of the flight intercept panel traps we used to sample flying beetles. The species we modelled are not known to possess long-range aggregation or sexual pheromone communication systems and our traps relied on kairomones (host volatiles) of ethanol and alpha-pinene as attractants. Host volatiles are generally less powerful than pheromones, and that reduces the sensitivity of traps [56]. Lower trap sensitivity increases the likelihood that our model would predict that conditions are not suitable for flight even though flight occurred, leading to a Type II error. Low trap sensitivity is one potential explanation for the high Type II error rates when the model threshold to predict 'No' is low. Increasing the models' predictive threshold partially compensates for this error, but in the absence of data on true detectability, prediction uncertainty remains. Increasing trap sensitivity may reduce type II errors, however more sensitive traps have the potential to sample flying insects from a larger area [57]. Increased sample areas may then be less representative of abiotic prediction variables measured by the meteorological station. The impact of this trade-off between trap sensitivity and model performance is currently unknown.

Implications for risk management

Prediction of forest insect flight activity over time provides an opportunity to estimate the potential rate of post-harvest log colonization by insects during storage in the forest, at

processing sites, or at ports prior to export. Understanding such risks allows phytosanitary treatments to be targeted effectively, thus reducing the potential for on-going spread of such species via international trade. To do this requires an understanding of the form of the relationship between individual log colonization rates and flight activity. Probability of flight must be combined with the infestation rate to make an assessment of overall phytosanitary risk. The form of this relationship is currently unknown, but it is likely to depend on the volume of logs in a given area. Larger concentrations of logs increase the odor plume that will attract beetles to a site, however they also dilute the instantaneous population of flying insects amongst a greater volume of logs, hence reducing infestation rate. Furthermore, knowledge of the ecology of individual species can be used as a management tool to modify the form of the relationship and reduce the likelihood of log colonization. For example, stacking logs above the ground will reduce the rate of colonization irrespective of flight activity, as *H. ligniperda* and *H. ater* are known to prefer logs in ground contact [9]. On a wider scale, understanding landscape configurations of forest age classes as dispersal corridors or barriers can inform on occurrence of forest insect species [58].

Evidence that conditions were not suitable for insect flight from the time of harvest to export could be used to support the case for temporal periods of low pest pressure when phytosanitary treatments are not necessary, as the risk of log colonization is below the maximum pest limit of importing countries. Setting an appropriate threshold for log exports requires an understanding of the potential risk of entry, establishment, and spread of a specific pest in an importing country [59]. Risk is generally gauged by the probability of an adverse event times its consequence [60]. Our results provide the first part of a fuller pest risk assessment which would entail applying the flight probabilities from our models to further analyses of infestation rates of trees and logs, and then applying those rates to utility estimates of the costs of infestation that pertain to phytosanitary treatments and monitoring as incurred by the wood products export industry. Such scientific assessments might then be used to inform policy and management who would set acceptable risk levels, particularly for false negatives [61]—predicted absence of insect flight activity when in fact insects activity occurred—that have consequences for maintaining export market access.

Selection of acceptable false negative rates has cost implications for export standards and phytosanitary treatments, and also policy implications for releasing potentially infested wood products to trading partners. As an example, in dealing with such decisions, some export sectors use specific sampling protocols to identify acceptable risk levels, e.g., no infested items within a sample of 600 provides an assurance (with 95% confidence) that the infestation rate is less than 0.5% [62]. Such a monitoring framework could be based on a specific risk attitude, market implications of false negatives, an acceptable level of detection, and information on overall infestation rates. In this way, key uncertainties in risk analysis, particularly with false negatives, can play important roles in informing and guiding policy [63].

Our models pertain to environmental correlates of insect flight activity, namely temperature, wind speed, humidity, and other variables that are mediated by forest stand structure, age, proximity to existing stands, and topographic location. Results of the influence analyses (Fig 6) suggest that flight activity could be variously lowered for the three insect species by storing harvested logs in areas with lower maximum temperature, relative humidity, wind speed, and direct solar radiation. Such conditions occur within even-age *Pinus radiata* forest stands >15 years old, with canopy closure >20% that can intercept wind and solar insolation. Hence, temporary in-forest storage of logs is best confined to cut blocks sufficiently small in size that are embedded in mature stands. This conclusion is supported by Mausel *et al.* [9] observation that logs stored in mature stands had lower levels of *H. ligniperda* and *H. ater* colonization.

Additional steps could refine our models. This includes additional flight activity data to improve model performance and generalize its application to wider geographic regions. An economic "value of information" analysis would assist this by focusing effort on predictor variables that are least well understood but have the most influence on prediction outcomes.

Finally the models should be field tested in an operational framework whereby flight activity is predicted and infestation of recently harvested logs are monitored. Such analysis can be used within a risk management framework to help inform future phytosanitary treatment needs for the global movement of wood commodities, specifically to determine areas of low pest prevalence and to calculate maximum post-fumigation exposure risk of insect occurrence and infestation.

Our approach has relevance, in that it can be applied to fundamental ecological studies of specific organisms and to specific applied ecological problems, including aspects of the invasion process. We recommend to readers that Bayesian networks should be applied more broadly in the management of complex ecological problems. This recommendation is based on their ability to utilize both quantitative and qualitative data simultaneously, their robustness to missing data and conditions of data multicollinearity and nonlinearity that can otherwise violate assumptions in more traditional multivariate modelling, and their ease of adoption by decision makers due to their interactive, graphical interface.

Supporting information

S1 Fig. Location of study sites in Canterbury, New Zealand. Green shading indicates areas of exotic plantation forests.

(PDF)

S2 Fig. Separator trap used in study.

(PDF)

S3 Fig. Average catch per trap per hour as a function of maximum hourly temperature.

Red symbols indicate the maximum trap catch per interval, solid blue symbols represent model outliers with a Cook's distance > 1 , and open blue symbols indicate suspect trap catch removed from analysis (see [results](#)). The adjusted R^2 provides an estimate of the variance explained by the non-linear curve fitting and are only provided when a non-linear response is identified.

(PDF)

S4 Fig. Average catch per trap per hour as a function of photon flux density and relative humidity. Red symbols indicate the maximum trap catch per interval, solid blue symbols represent model outliers with a Cook's distance < 1 , and open blue symbols indicate suspect trap catch removed from analysis (see [results](#)). Model predictions are only shown when a significant relationship was present between the predictor and response variable. The adjusted R^2 provides an estimate of the variance explained by the non-linear curve fitting and are only provided when a non-linear response is identified.

(PDF)

S5 Fig. Average catch per trap per hour as a function of wind speed. Red symbols indicate the maximum trap catch per interval, solid blue symbols represent model outliers with a Cook's distance < 1 , and open blue symbols indicate suspect trap catch removed from analysis (see [results](#)). Model predictions are only shown when a significant relationship was present between the predictor and response variable. The adjusted R^2 provides an estimate of the variance explained by the non-linear curve fitting and are only provided when a non-linear

response is identified.
(PDF)

S6 Fig. Mean hourly meteorological conditions as a function of the time of day averaged across all sites. The grey polygon defines the standard error of the mean for each hourly measurement.
(PDF)

S1 Table. Location of individual traps in the four study sites. Two sites were established in Ashley Forest, the first on McGibbons Rd in a recently clear-felled stand with a predominant southeast aspect at 300 m elevation. The second site was on Mt Grey Rd also in a recently clear-felled stand, but with a predominantly eastern aspect at 381m elevation. The third site was on a flat, recently clear-felled site in the West Melton Forest Rd at 104 m elevation. The fourth site was on a flat, recently clear-felled site in McLeans Forest at 59 m elevation. As recent clearfells all sites had no limited to no forest structure, i.e., newly planted sites with 0.3 m seedlings or young trees up to 0.7 m.
(PDF)

S2 Table. Title and description of nodes used in Bayesian network models of *Hylurgus ligniperda*, *H. ater* and *Arhopalus fesus* flight activity. N/A = the node does not pertain to the species model.
(PDF)

S3 Table. Conditional probability tables for each node in the Bayesian network model of *Hylurgus ligniperda* flight activity as discretised from case data using the expectation maximization algorithm. S3A Table. Conditional probability table for node day of year. S3B Table. Conditional probability table for node relative humidity (%). S3C Table. Conditional probability table for node temperature range (°C). S3D Table. Conditional probability table for node rainfall (mm/hr). S3E Table. Conditional probability table for node maximum temperature (°C). S3F Table. Conditional probability table for node time since sunrise (mins). S3G Table. Conditional probability table for node Photon flux density ($\mu\text{mol photons m}^{-2}\text{s}^{-1}$). S3H Table. Conditional probability table for node Wind speed ($\text{m}^{-1}\text{s}^{-1}$). S3I Table. Conditional probability table for node time since sunset (mins).
(PDF)

S4 Table. Conditional probability tables for each node in the Bayesian network model of *Hylastes ater* flight activity as discretised from case data using the expectation maximization algorithm. S4A Table. Conditional probability table for node flight. S4B Table. Conditional probability table for node temperature range (°C). S4C Table. Conditional probability table for node wind speed ($\text{m}^{-1}\text{s}^{-1}$). S4D Table. Conditional probability table for node day of year. S4E Table. Conditional probability table for node maximum temperature (°C). S4F Table. Conditional probability table for node maximum temperature (°C). S4G Table. Conditional probability table for node time since sunrise (mins).
(PDF)

S5 Table. Conditional probability tables for each node in the Bayesian network model of *Arhopalus fesus* flight activity as discretised from case data using the expectation maximization algorithm. S5A Table. Conditional probability table for node. S5B Table. Conditional probability table for node time since sunrise (mins). S5C Table. Conditional probability table for node Photon flux density ($\mu\text{mol photons m}^{-2}\text{s}^{-1}$). S5D Table. Conditional probability table for node wind speed ($\text{m}^{-1}\text{s}^{-1}$). S5E Table. Conditional probability table for node maximum temperature (°C). S5F Table. Conditional probability table for node day of year. S5G Table.

Conditional probability table for node temperature range (°C).
(PDF)

Acknowledgments

The authors thank Environment Canterbury and Rayonier[Matariki for access to forest stands to conduct this experiment, and Brooke O'Connor, Jessica Kerr, and Jeremy Burnett for the establishment of the trial, daily maintenance of the separator traps and identification of samples. Martin Bader for assistance with GAM modelling. This work was funded by the New Zealand Ministry for Business, Innovation, and Employment (C04X1204) and the Stakeholders in Methyl Bromide Reduction Inc. An Agmardt grant enabled the participation of Bruce Marcot with in-kind contribution through U.S. Forest Service, Pacific Northwest Research Station.

Author Contributions

Conceptualization: Stephen M. Pawson.

Data curation: Stephen M. Pawson, Owen G. Woodberry.

Formal analysis: Stephen M. Pawson, Bruce G. Marcot, Owen G. Woodberry.

Funding acquisition: Stephen M. Pawson.

Methodology: Stephen M. Pawson, Bruce G. Marcot, Owen G. Woodberry.

Project administration: Stephen M. Pawson.

Writing – original draft: Stephen M. Pawson.

Writing – review & editing: Bruce G. Marcot, Owen G. Woodberry.

References

1. IPPC. ISPM22: Requirements for the establishment of areas of low pest prevalence. Rome: Secretariat of the International Plant Protection Convention, Food and Agriculture Organization of the United Nations, 2005.
2. FOA, MPI. 2015/16 Facts and Figures New Zealand Plantation Forest Industry. Wellington: New Zealand Forest Owners Association and Ministry for Primary Industries, 2016.
3. Brouckhoff EG, Bain J, Kimberley M, Knizek M. Interception frequency of exotic bark and ambrosia beetles (Coleoptera: Scolytinae) and relationship with establishment in New Zealand and worldwide. Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere. 2006; 36(2):289–98.
4. Hosking GP, Bain J. *Arhopalus ferus* (Coleoptera: Cerambycidae); Its biology in New Zealand. New Zealand Journal of Forestry Science. 1977; 7(1):3–15.
5. Brouckhoff EG, Hosking GP. *Arhopalus tristis* (F.) (Coleoptera: Cerambycidae) (= *Arhopalus ferus* (Mulsant)): Burnt pine longhorn beetle. Forest and Timber Insects in New Zealand. 2001; 27 (Revised):1–8.
6. Milligan RH. *Hylastes ater* (Paykull) (Coleoptera: Scolytidae), Black Pine Beetle. Forest and Timber Insects in New Zealand. 1978; 29:1–8.
7. Reay SD, Walsh PJ. Observations of the flight activity of *Hylastes ater* and *Hylurgus ligniperda* (Curculionidae: Scolytinae) in *Pinus radiata* forests in the central North Island, New Zealand. New Zealand Entomologist. 2001; 24:79–85.
8. Tribe GD. Phenology of *Pinus radiata* log colonization by the red-haired pine bark beetle *Hylurgus ligniperda* (Fabricius) (Coleoptera, Scolytidae) in the south-western cape province. Journal of the Entomological Society of Southern Africa. 1991; 54(1):1–7.
9. Mausel DL, Gara RI, Lanfranco D, Ruiz C, Ide S, Azat R. The introduced bark beetles *Hylurgus ligniperda* and *Hylastes ater* (Coleoptera: Scolytidae) in Chile: seasonal flight and effect of *Pinus radiata* log placement on colonization. Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere. 2007; 37(1):156–69. <https://doi.org/10.1139/x06-215>

10. Sopow SL, Bader MKF, Brockerhoff EG. Bark beetles attacking conifer seedlings: Picking on the weakest or feasting upon the fittest? *Journal of Applied Ecology*. 2015; 52(1):220–7. <https://doi.org/10.1111/1365-2664.12368>
11. Kerr JL, Kelly D, Bader MK-F, Brockerhoff EG. Olfactory cues, visual cues, and semiochemical diversity interact during host location by invasive forest beetles. *Journal of Chemical Ecology*. 2016. <https://doi.org/10.1007/s10886-016-0792-x> PMID: 27832345
12. Fettig CJ, Shea PJ, Borys RR. Diurnal flight pattern of *Dendroctonus brevicornis* LeConte (Coleoptera: Scolytidae) in northern California. *Journal of Entomological Science*. 2004; 39(3):453–6.
13. Chen Y, Seybold SJ. Crepuscular flight activity of an invasive insect governed by interacting abiotic factors. *PLoS ONE*. 2014; 9(8). <https://doi.org/10.1371/journal.pone.0105945> PMID: 25157977
14. Daterman GE, Rudinsky JA, Nagel WP. Flight patterns of bark and timber beetle. Oregon State University, Agricultural Experiment Station Technical Bulletin. 1965;(87).
15. Mendel Z, Boneh O, Shenhar Y, Rivon J. Diurnal flight patterns of *Orthotomicus erosus* and *Pityogenes calcaratus* in Israel. *Phytoparasitica*. 1991; 19(1):23–31. <https://doi.org/10.1007/BF02981008>
16. Suckling DM, Gibb AR, Daly JM, Chen X, Brockerhoff EG. Behavioural and electrophysiological responses of *Arhopalus tristicus* to burnt pine and other stimuli. *Journal of Chemical Ecology*. 2001; 27(6):1091–104. PMID: 11504016
17. Haack RA. Voltinism and diurnal emergence. Flight patterns of *Ips calligraphus* (Coleoptera: Scolytidae) in Florida. *Florida Entomologist*. 1985; 68(4):658–67.
18. Östrand F, Anderbrant O. From where are insects recruited? A new model to interpret catches of attractive traps. *Agricultural and Forest Entomology*. 2003; 5(2):163–71. <https://doi.org/10.1046/j.1461-9563.2003.00174.x>
19. Darwiche A. Modeling and reasoning with Bayesian networks. New York: Cambridge University Press; 2009. 560 p.
20. Landuyt D, Broekx S, D'hondt R, Engelen G, Aertsens J, Goethals PLM. A review of Bayesian belief networks in ecosystem service modelling. *Environmental Modeling and Software*. 2013; 46:1–11.
21. Boets P, Landuyt D, Everaert G, Broekx S, Goethals PLM. Evaluation and comparison of data-driven and knowledge-supported Bayesian belief networks to assess the habitat suitability for alien macroinvertebrates. *Environmental Modelling & Software*. 2015; 74:92–103.
22. Fenton N, Neil M. Risk assessment and decision analysis with Bayesian networks. Boca Raton, FL: CRC Press; 2012. 524 p.
23. Marcot BG, Steventon JD, Sutherland GD, McCann RK. Guidelines for developing and updating Bayesian belief networks applied to ecological modeling and conservation. *Canadian Journal of Forest Research*. 2006; 36:3063–74.
24. Ryan JA, Ulrich JM. xts: Extensible time series. Version 0.9–7. 2008.
25. Bivand R, Lewin-Koh N. Maptools: Tools for reading and handling spatial objects. R package version 0.8–36. 2015.
26. Meeus J. Astronomical algorithms. Virginia: Willmann-Bell Inc; 1991.
27. Blackburn TM, Lawton JH, Perry JN. A method of estimating the slope of upper bounds of plots of body size and abundance in natural animal assemblages. *Oikos*. 1992; 65(1):107–12. <https://doi.org/10.2307/3544892>
28. Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH. The WEKA Data Mining Software: An Update. The Association for Computing Machinery's Special Interest Group on Knowledge Discovery and Data Mining (SIGKDD). 2009; 11(1).
29. Brazzale AR. The R Core Team. Package nlreg: Higher order inference for nonlinear heteroscedastic models. R package version 1.2–2. 2014.
30. Myers JL, Well AD, Lorch RF. Research design and statistical analysis. Third ed. Great Britain: Routledge; 2010. 1–809 p.
31. Pinheiro J, Bates D, Debroy S, Sarkar D. The R Core Team. Package nlme: Linear and nonlinear mixed effect models. R package version 3.1–121. 2015.
32. Wood S. Mixed GAM computation vehicle with GCV/AIC/REML smoothness estimation, Package: mgcv, Version: 1.8–7. 2015.
33. R Development Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2013.
34. Fayyad UM, Irani KB, editors. Multi-interval discretisation of continuous-valued attributes for classification learning. Thirteenth International Joint Conference on Artificial Intelligence; 1993: International Joint Conferences on Artificial Intelligence.

35. Duda RO, Hart PE. Pattern classification and scene analysis. J Wiley and Sons. 1973.
36. Wallace CS, Korb KB. Learning linear causal models by MML sampling. In: Gammerman A, editor. Causal Models and Intelligent Data Management: Springer-Verlag; 1999. p. 89–111.
37. Friedman N, Geiger D, Goldszmidt M. Bayesian network classifiers. Machine Learning. 1997; 29:131–63.
38. Korb KB, Nicholson AE. Bayesian artificial intelligence: CRC press; 2004.
39. Russell SJ, Norvig P. Artificial intelligence: a modern approach: Prentice-Hall, Inc.; 1995. 932 p.
40. Neapolitan RE. Learning Bayesian Networks: Prentice-Hall, Inc.; 2003.
41. Marcot BG. Characterizing species at risk I: Modeling rare species under the Northwest Forest Plan. Ecology and Society. 2006; 11(2).
42. Marcot BG. Metrics for evaluating performance and uncertainty of Bayesian network models. Ecological Modelling. 2012; 230:50–62. <https://doi.org/10.1016/j.ecolmodel.2012.01.013>
43. Han M, Wang Z, Liu Y. Selecting one dependency estimators in Bayesian network using different MDL scores and overfitting criterion. Journal of Information Science and Engineering. 2014; 30:371–85.
44. Cawley GC, Talbot NLC. Preventing over-fitting during model selection via Bayesian regularisation of the hyper-parameters. Journal of Machine Learning Research. 2007; 8:841–61.
45. Jakulin A, Rish I. Bayesian learning of Markov network structure. Lecture Notes in Computer Science. 2006; 4212:198–209.
46. Arlot S, Celisse A. A survey of cross-validation procedures for model selection. Statistics Surveys. 2010; 4:40–79.
47. Amstrup SC, Marcot BG, Douglas DC. A Bayesian network modeling approach to forecasting the 21st century worldwide status of polar bears. In: DeWeaver ET, Bitz CM, Tremblay L-B, editors. Arctic sea ice decline: observations, projections, mechanisms, and implications Geophysical Monograph 180. Washington, D.C.: American Geophysical Union; 2008.
48. Goodwin P, Wright G. Decision analysis for management judgment. Third edition. New York: John Wiley & Sons; 2004. 477 p.
49. Iwata R, Maro T, Yonezawa Y, Yahagi T, Fujikawa Y. Period of adult activity and response to wood moisture content as major segregating factors in the coexistence of two conifer longhorn beetles, *Callidiellum rufipenne* and *Semanotus bifasciatus* (Coleoptera: Cerambycidae). European Journal of Entomology. 2007; 104(2):341–5.
50. Starzyk JR. The daily activity rhythm of *Gaurotus virginea* (L.) (Coleoptera, Cerambycidae). Folia Biologica. 1968; 16(3):267–82. PMID: [5702338](https://pubmed.ncbi.nlm.nih.gov/5702338/)
51. Wang Q, Zeng W, Chen L, Li J, Yin X. Circadian reproductive rhythms, pair-bonding, and evidence for sex-specific pheromones in *Nadezhdiella cantori* (Coleoptera: Cerambycidae). Journal of Insect Behavior. 2002; 15(4):527–39. <https://doi.org/10.1023/A:1016333317564>
52. Leathwick JR, Wilson G, Stephens RTT. Climate surfaces for New Zealand. Hamilton: Landcare Research Contract Report: LC9798/126, 2002.
53. Chen Y, Why A, Batista G, Mafra-Neto A, Keogh E. Flying insect classification with inexpensive sensors. Journal of Insect Behavior. 2014; 27(5):657–77. <https://doi.org/10.1007/s10905-014-9454-4>
54. Maxwell SL, Rhodes JR, Runge MC, Possingham HP, Ng CF, McDonald-Madden E. How much is new information worth? Evaluating the financial benefit of resolving management uncertainty. Journal of Applied Ecology. 2015; 52(1):12–20. <https://doi.org/10.1111/1365-2664.12373>
55. Williams BK, Eaton MJ, Breininger DR. Adaptive resource management and the value of information. Ecological Modelling. 2011; 222(18):3429–36. <https://doi.org/10.1016/j.ecolmodel.2011.07.003>
56. Nadel RL, Wingfield MJ, Scholes MC, Lawson SA, Slippers B. The potential for monitoring and control of insect pests in Southern Hemisphere forestry plantations using semiochemicals. Annals of Forest Science. 2012; 69(7):757–67. <https://doi.org/10.1007/s13595-012-0200-9>
57. Byers JA. Active space of pheromone plume and its relationship to effective attraction radius in applied models. Journal of Chemical Ecology. 2008; 34:1134–145. <https://doi.org/10.1007/s10886-008-9509-0> PMID: [18584255](https://pubmed.ncbi.nlm.nih.gov/18584255/)
58. Holland JD. Isolating spatial effects on beta diversity to inform forest landscape planning. Landscape Ecology. 2010; 25(9):1349–62. <https://doi.org/10.1007/s10980-010-9499-5>
59. Rassati D, Faccoli M, Petrucco Toffolo E, Battisti A, Marini L. Improving the early detection of alien wood-boring beetles in ports and surrounding forests. Journal of Applied Ecology. 2015; 52:50–8.
60. Aven T. Foundations of risk analysis. West Sussex, England: John Wiley & Sons; 2003. 190 p.

61. Hanson N. Using biological data from field studies with multiple reference sites as a basis for environmental management: The risks for false positives and false negatives. *Journal of Environmental Management*. 2011; 92(3):610–9. <https://doi.org/10.1016/j.jenvman.2010.09.026> PMID: 21035245
62. MPI. Biosecurity risk management of entry pathways for pasture pests. Wellington: Ministry for Primary Industries, 2013.
63. Rouse HL, Norton N. Managing scientific uncertainty for resource management planning in New Zealand. *Australasian Journal of Environmental Management*. 2010; 17(2):66–76.

Table S1: Location of individual traps in the four study sites

Two sites were established in Ashley Forest, the first on McGibbons Rd in a recently clear-felled stand with a predominant southeast aspect at 300 m elevation. The second site was on Mt Grey Rd also in a recently clear-felled stand, but with a predominantly eastern aspect at 381m elevation. The third site was on a flat, recently clear-felled site in the West Melton Forest Rd at 104 m elevation. The fourth site was on a flat, recently clear-felled site in McLeans Forest at 59 m elevation. As recent clearfells all sites had no limited to no forest structure, i.e., newly planted sites with 0.3 m seedlings or young trees up to 0.7 m.

Forest	Location	Forest Compartment	Harvest Year	NZTM East	NZTM North
Ashley Forest	McGibbons Rd	29	2012	1562257	5219568
Ashley Forest	McGibbons Rd	29	2012	1562274	5219605
Ashley Forest	McGibbons Rd	29	2012	1562216	5219563
Ashley Forest	McGibbons Rd	29	2012	1562232	5219536
Ashley Forest	Mt Grey Rd	62	2012	1562429	5222459
Ashley Forest	Mt Grey Rd	62	2012	1562427	5222418
Ashley Forest	Mt Grey Rd	62	2012	1562415	5222496
Ashley Forest	Mt Grey Rd	62	2012	1562390	5222467
West Melton	Thompsons Rd	12	2012	1547946	5187228
West Melton	Thompsons Rd	12	2012	1547905	5187228
West Melton	Thompsons Rd	12	2012	1547948	5187188
West Melton	Thompsons Rd	12	2012	1547985	5187227
McLean's Island	Norton Ave	28	2011	1555927	5188432
McLean's Island	Norton Ave	28	2011	1555887	5188427
McLean's Island	Norton Ave	28	2011	1555930	5188393
McLean's Island	Norton Ave	28	2011	1555966	5188427

Table S2. Title and description of nodes used in Bayesian network models of *Hylurgus ligniperda*, *H. ater* and *Arhopalus ferus* flight activity. N/A = the node does not pertain to the species model.

Node title	Node description	States		
		<i>Hylurgus ligniperda</i>	<i>Hylastes ater</i>	<i>Arhopalus ferus</i>
Flight	The occurrence of flight during that hour.	Yes No	Yes No	Yes No
Day of year	The day of the year was expressed as a decimal (001-366). Because the sampling occurred across three discontinuous periods we incorporate the day of the year as a measure of the time of year to indicate if there are seasonal differences to the flight activity.	<344 ≥344	<41 ≥41	<14 >14 to 41 >41 to 325 ≥325
Maximum temperature (C)	The maximum temperature recorded during the hour. Measurements during the hour were taken every minute.	<12.3 >12.3 to 14.6 >14.6 to 17.5 ≥17.5	<14.0 ≥14.0	<12.6 ≥12.6
Photon flux density (μmol photons m ⁻² s ⁻¹)	Measure of the average photons in the 400-700nm range per square metre per second during the sampling hour. The average is of observations taken every minute.	<0.1 >0.1 to 565 >565 to 1704 ≥1704	N/A	<0.23 ≥0.23
Rainfall (mm/hr)	Cumulative rainfall across the sampling hour.	<0.1 ≥0.1	N/A	N/A
Relative humidity (%)	The average relative humidity calculated from minute observations during the hour	<82 >82 to 93 ≥93	N/A	N/A
Temperature range (C)	The range (in °C) between the maximum and minimum temperatures recorded during the sampling hour. Measurements during the hour were taken every minute.	<0.91 ≥0.91	<0.90 ≥0.90	<1.1 ≥1.1
Time since sunrise (mins)	The time in minutes since sunrise as measured from the start of the sampling hour. For example, if sunrise was at 5:50am, the trap hour beginning at 6am would have a value of -10 minutes, however the 5am trap hour would have a value of 50 minutes.	<19 >19 to 375 >375 to 891 ≥891	<33 >33 to 289 >289 to 735 >735 to 952 ≥952	N/A
Time since sunset (mins)	The time in minutes since sunset as measured from the start of the sampling hour. For example, if sunset was at 8:30pm, the trap hour beginning at 8pm would have a value of -30 minutes, however the 9pm trap hour would have a value of 30 minutes.	<11 >11 to 498 >498 to 589 ≥589	<-47 >-47 to 1221 ≥1221	<-20 >-20 to 6 >6 to 40 >40 to 65 >65 to 191 >191 to 373 ≥373
Wind speed (m-1s-1)	Average wind speed during the hour, calculated from minute observations.	<0.9 >0.9 to 4.1 ≥4.1	<3.3 ≥3.3	<4.2 ≥4.2

Table S3. Conditional probability tables for each node in the Bayesian network model of *Hylurgus ligniperda* flight activity as discretized from case data using the expectation maximization algorithm

Table S3A. Conditional probability table for node Day of year

Flight	Outcome	
	< 344	>= 344
Yes	0.970938	0.029062
No	0.876216	0.123784

Table S3B. Conditional probability table for node relative humidity (%)

Flight	Maximum Temperature	Outcome		
		< 82	>82 to 93	>= 93
Yes	< 12.3	0.466667	0.333333	0.200000
Yes	>12.3 to 14.6	0.792453	0.188679	0.018868
Yes	>14.6 to 17.5	0.848958	0.140625	0.010417
Yes	>= 17.5	0.970954	0.029046	2.07E-08
No	< 12.3	0.374411	0.378336	0.247253
No	>12.3 to 14.6	0.679849	0.169491	0.150659
No	>14.6 to 17.5	0.649958	0.252297	0.097744
No	>= 17.5	0.970644	0.026165	0.003191

Table S3C. Conditional probability table for node temperature range (°C)

Flight	Relative humidity	Outcome	
		< 0.91	>= 0.91
Yes	< 82	0.091703	0.908297
Yes	>82 to 93	0.393443	0.606557
Yes	>= 93	0.888888	0.111112
No	< 82	0.192201	0.807799
No	>82 to 93	0.587761	0.412239
No	>= 93	0.718202	0.281798

Table S3D. Conditional probability table for node rainfall (mm/hr)

Flight	Relative humidity	Outcome	
		< 0.1	>= 0.1
Yes	< 82	0.983988	0.016012
Yes	>82 to 93	0.934426	0.065574
Yes	>= 93	0.777777	0.222223
No	< 82	0.978365	0.021635
No	>82 to 93	0.886348	0.113652
No	>= 93	0.758772	0.241228

Table S3E. Conditional probability table for node maximum temperature (°C)

Flight	Day of year	Outcome			
		< 12.3	>12.3 to 14.6	>14.6 to 17.5	>= 17.5
Yes	< 344	0.039456	0.066667	0.244898	0.648979
Yes	>= 344	0.045455	0.181818	0.545454	0.227273
No	< 344	0.381021	0.157028	0.198568	0.263384
No	>= 344	0.532319	0.234474	0.111534	0.121673

Table S3F. Conditional probability table for node time since sunrise (mins)

Flight	Time since sunset (mins)	Outcome			
		< 19	>19 to 375	>375 to 891	>= 891
Yes	< 11	1.75E-07	1.75E-07	0.947368	0.0526317
Yes	>11 to 498	4.76E-07	4.76E-07	0.142857	0.857142
Yes	>498 to 589	0.944443	5.56E-07	5.56E-07	0.055556
Yes	>= 589	0.003026	0.564296	0.432678	1.51E-08
No	< 11	3.00E-08	3.00E-08	0.729730	0.27027
No	>11 to 498	0.040098	4.09E-09	0.078969	0.880933
No	>498 to 589	0.600575	0.014368	2.87E-08	0.385057
No	>= 589	0.041551	0.404432	0.542013	0.0120037

Table S3G. Conditional probability table for node Photon flux density ($\mu\text{mol photons m}^{-2}\text{s}^{-1}$)

Flight	Maximum Temperature ($^{\circ}\text{C}$)	Outcome			
		< 0.1	>0.1 to 565	>565 to 1704	>= 1704
Yes	< 12.3	0.366666	0.433333	0.200000	3.33E-07
Yes	>12.3 to 14.6	0.094340	0.584905	0.320755	1.89E-07
Yes	>14.6 to 17.5	0.062500	0.421875	0.484375	0.031250
Yes	>= 17.5	0.035270	0.331950	0.547718	0.0850623
No	< 12.3	0.490188	0.434851	0.074961	3.92E-09
No	>12.3 to 14.6	0.390772	0.369115	0.230697	0.0094162
No	>14.6 to 17.5	0.342523	0.333333	0.279866	0.0442774
No	>= 17.5	0.155073	0.228462	0.412891	0.203574

Table S3H. Conditional probability table for node Wind speed ($\text{m}^{-1}\text{s}^{-1}$)

Flight	Wind speed ($\text{m}^{-1}\text{s}^{-1}$)	Outcome		
		< 0.9	>0.9 to 4.1	>= 4.1
Yes	< 0.1	0.244444	0.600000	0.155556
Yes	>0.1 to 565	0.066667	0.789474	0.143860
Yes	>565 to 1704	0.005263	0.794737	0.200000
Yes	>= 1704	2.13E-07	0.553191	0.446808
No	< 0.1	0.227018	0.598187	0.174795
No	>0.1 to 565	0.108551	0.651307	0.240142
No	>565 to 1704	0.010578	0.504937	0.484485
No	>= 1704	2.62E-08	0.319372	0.680628

Table S3I. Conditional probability table for node time since sunset (mins)

Flight	Photon flux density ($\mu\text{mol photons m}^{-2}\text{s}^{-1}$)	Outcome			
		< 11	>11 to 498	>498 to 589	>= 589
Yes	< 0.1	0.622222	0.377778	2.22E-07	2.22E-07
Yes	>0.1 to 565	0.101754	0.0140351	0.0631579	0.821053
Yes	>565 to 1704	2.63E-08	2.63E-08	2.63E-08	1
Yes	>= 1704	2.13E-07	2.13E-07	2.13E-07	0.999999
No	< 0.1	0.0984031	0.866638	0.0276219	0.00733708
No	>0.1 to 565	0.0465219	0.193177	0.125831	0.634471
No	>565 to 1704	7.05E-09	7.05E-09	7.05E-09	1
No	>= 1704	2.62E-08	2.62E-08	2.62E-08	1

Table S4. Conditional probability tables for each node in the Bayesian network model of *Hylastes ater* flight activity as discretized from case data using the expectation maximization algorithm.

Table S4A. Conditional probability table for node flight

Flight	
Yes	0.019767
No	0.980233

Table S4B. Conditional probability table for node temperature range (°C)

Flight	Time since sunrise (mins)	Outcome	
		< 0.9	>= 0.9
Yes	< 33	0.666666	0.333334
Yes	33 to 289	0.057693	0.942307
Yes	289 to 735	0.190477	0.809524
Yes	735 to 952	0.192308	0.807692
Yes	>= 952	0.307693	0.692307
No	< 33	0.262525	0.737475
No	33 to 289	0.133112	0.866889
No	289 to 735	0.183827	0.816173
No	735 to 952	0.530144	0.469857
No	>= 952	0.520057	0.479943

Table S4C. Conditional probability table for node wind speed (m⁻¹s⁻¹)

Flight	Time since sunrise (mins)	Outcome	
		< 3.3	>= 3.3
Yes	< 33	0.666666	0.333334
Yes	33 to 289	0.884615	0.115385
Yes	289 to 735	0.619048	0.380953
Yes	735 to 952	0.769231	0.230769
Yes	>= 952	0.999999	7.69E-07
No	< 33	0.733467	0.266533
No	33 to 289	0.535774	0.464226
No	289 to 735	0.349788	0.650212
No	735 to 952	0.668899	0.331100
No	>= 952	0.756017	0.243983

Table S4D. Conditional probability table for node Day of year

Flight	Maximum Temperature (°C)	Outcome	
		< 41	>= 41
Yes	< 14	0.473684	0.526316
Yes	>= 14	0.803279	0.196721
No	< 14	0.166418	0.833582
No	>= 14	0.532014	0.467986

Table S4E. Conditional probability table for node maximum temperature (°C)

Flight	Time since sunrise (mins)	Outcome	
		< 14	>= 14
Yes	< 33	3.33E-06	0.999997
Yes	33 to 289	0.019231	0.980769
Yes	289 to 735	0.238096	0.761904
Yes	735 to 952	0.134615	0.865385
Yes	>= 952	0.461538	0.538462
No	< 33	0.625251	0.374750
No	33 to 289	0.318636	0.681364
No	289 to 735	0.283028	0.716972
No	735 to 952	0.588517	0.411483
No	>= 952	0.680038	0.319962

Table S4F. Conditional probability table for node maximum temperature (°C)

Flight	Time since sunset (mins)	Outcome		
		< -47	-47 to 1221	>= 1221
Yes		0.085106	0.687943	0.22695
No		0.01373	0.887729	0.098541

Table S4G. Conditional probability table for node time since sunrise (mins)

Flight	Time since sunset (mins)	Outcome				
		< 33	33 to 289	289 to 735	735 to 952	>= 952
Yes	< -47	8.33E-07	8.33E-07	8.33E-07	0.999997	8.33E-07
Yes	-47 to 1221	0.030928	0.536082	0.134021	0.164948	0.134021
Yes	>= 1221	3.12E-07	3.12E-07	0.250000	0.749999	3.12E-07
No	< -47	1.04E-07	1.04E-07	0.072917	0.927083	1.04E-07
No	-47 to 1221	0.080393	0.193652	0.286773	0.097793	0.341389
No	>= 1221	1.45E-08	1.45E-08	0.493469	0.506531	1.45E-08

Table S5. Conditional probability tables for each node in the Bayesian network model of *Arhopalus fesus* flight activity as discretized from case data using the expectation maximization algorithm.

Table S5A. Conditional probability table for node

Flight	
Yes	0.0266629
No	0.973337

Table S5B. Conditional probability table for node time since sunrise (mins)

Flight	Time since sunset (mins)	Outcome						
		< -20	-20 to 6	6 to 40	40 to 65	65 to 191	191 to 373	>= 373
Yes		0.01579	0.294737	0.052632	0.210526	0.247368	0.131579	0.047368
No		0.028691	0.013264	0.02105	0.014562	0.087226	0.128604	0.706603

Table S5C. Conditional probability table for node Photon flux density ($\mu\text{mol photons m}^{-2}\text{s}^{-1}$)

Flight	Time since sunset (mins)	Outcome	
		< 0.23	>= 0.23
Yes	< -20	0.999997	3.33E-06
Yes	-20 to 6	~1	1.79E-07
Yes	6 to 40	0.899999	0.100001
Yes	40 to 65	~1	2.50E-07
Yes	65 to 191	~1	2.13E-07
Yes	191 to 373	~1	4.00E-07
Yes	>= 373	1.11E-06	0.999999
No	< -20	0.492462	0.507538
No	-20 to 6	0.978261	0.021739
No	6 to 40	0.979452	0.020548
No	40 to 65	0.990099	0.009901
No	65 to 191	0.971901	0.028099
No	191 to 373	0.980942	0.019058
No	>= 373	0.079372	0.920628

Table S5D. Conditional probability table for node wind speed ($\text{m}^{-1}\text{s}^{-1}$)

Flight	Maximum temperature (C)	Outcome	
		< 4.2	>= 4.2
Yes	< 12.6	0.888888	0.111111
Yes	>= 12.6	0.877907	0.122093
No	< 12.6	0.831106	0.168894
No	>= 12.6	0.657473	0.342527

Table S5E. Conditional probability table for node maximum temperature (°C)

Flight	Photon flux density ($\mu\text{mol photons m}^{-2}\text{s}^{-1}$)	Outcome	
		< 12.6	>= 12.6
Yes	< 0.23	0.1	0.9
Yes	>= 0.23	1.00E-06	0.999999
No	< 0.23	0.593955	0.406045
No	>= 0.23	0.287556	0.712444

Table S5F. Conditional probability table for node Day of year

Flight	Maximum temperature (°C)	Outcome			
		< 14	14 to 41	41 to 325	>= 325
Yes	< 12.6	5.56E-07	0.444444	5.56E-07	0.555555
Yes	>= 12.6	0.040698	0.796511	5.81E-08	0.162791
No	< 12.6	3.71E-09	0.133259	0.44098	0.425761
No	>= 12.6	0.001179	0.489863	0.214286	0.294672

Table S5G. Conditional probability table for node temperature range (°C)

Flight	Time since sunset (mins)	Outcome	
		< 1.1	>= 1.1
Yes	< -20	0.333334	0.666666
Yes	-20 to 6	0.678571	0.321429
Yes	6 to 40	0.200001	0.799999
Yes	40 to 65	0.750000	0.250000
Yes	65 to 191	0.638298	0.361702
Yes	191 to 373	0.440000	0.560000
Yes	>= 373	0.333334	0.666666
No	< -20	0.59799	0.40201
No	-20 to 6	0.804348	0.195652
No	6 to 40	0.582192	0.417808
No	40 to 65	0.861386	0.138614
No	65 to 191	0.657851	0.342149
No	191 to 373	0.600897	0.399103
No	>= 373	0.314426	0.685574

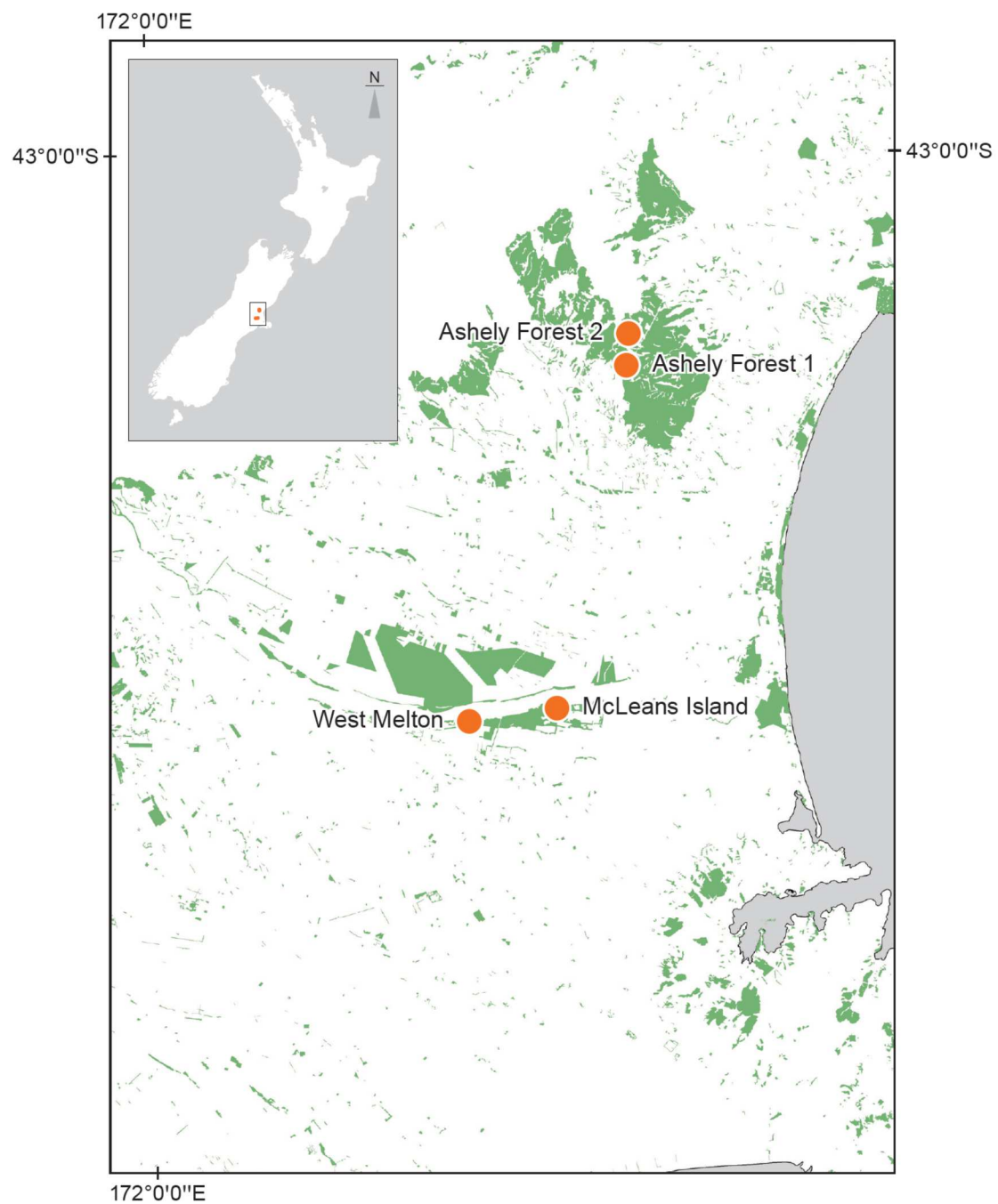


Figure S1. Location of study sites in Canterbury, New Zealand. Green shading indicates areas of exotic plantation forests.

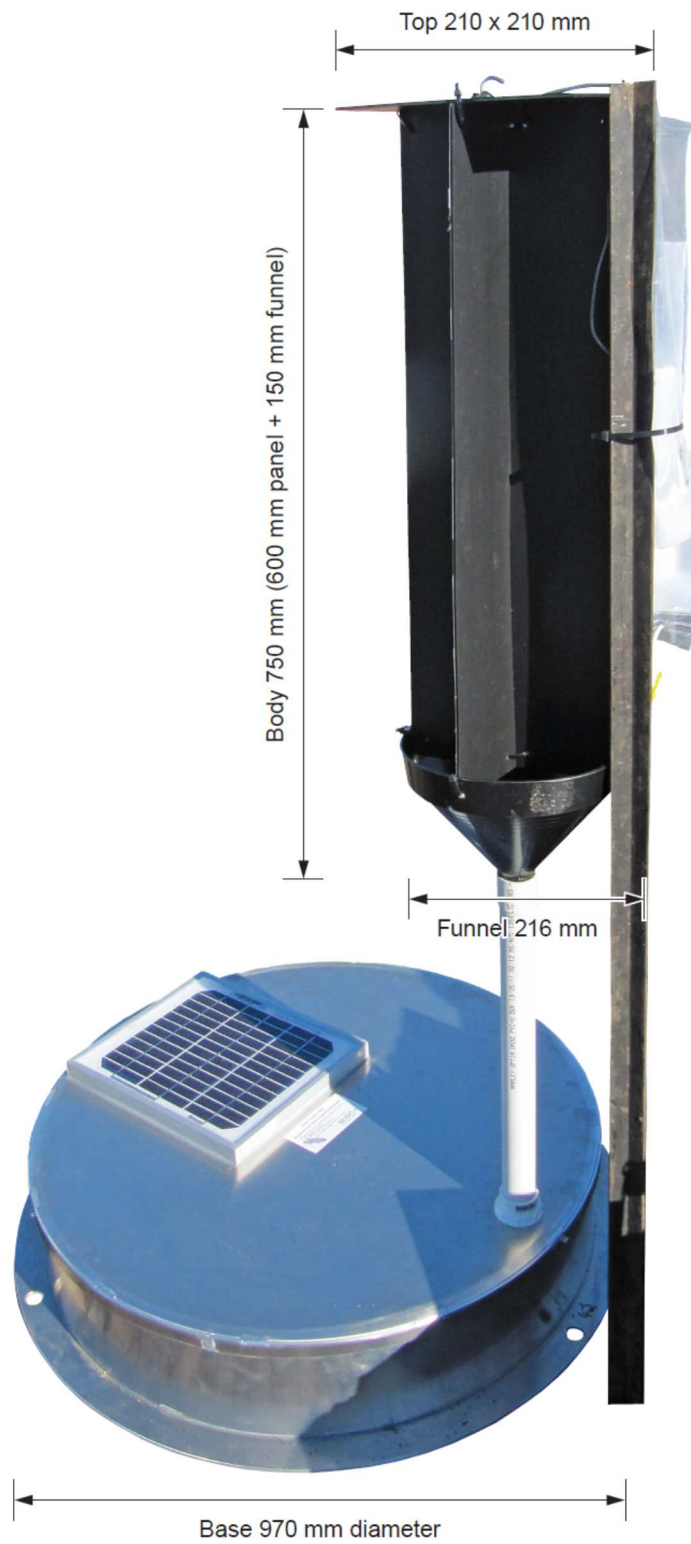


Figure S2. Separator trap used in study

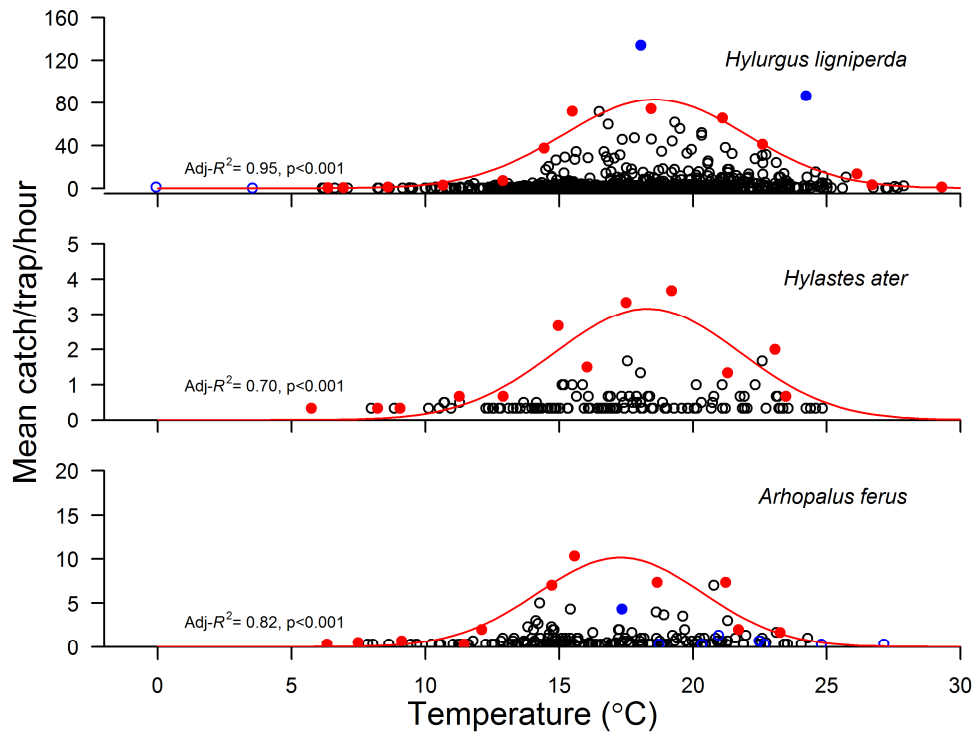


Fig S3. Average catch per trap per hour as a function of maximum hourly temperature.

Red symbols indicate the maximum trap catch per interval, solid blue symbols represent model outliers with a Cook's distance > 1 , and open blue symbols indicate suspect trap catch removed from analysis (see results). The adjusted R^2 provides an estimate of the variance explained by the non-linear curve fitting and are only provided when a non-linear response is identified.

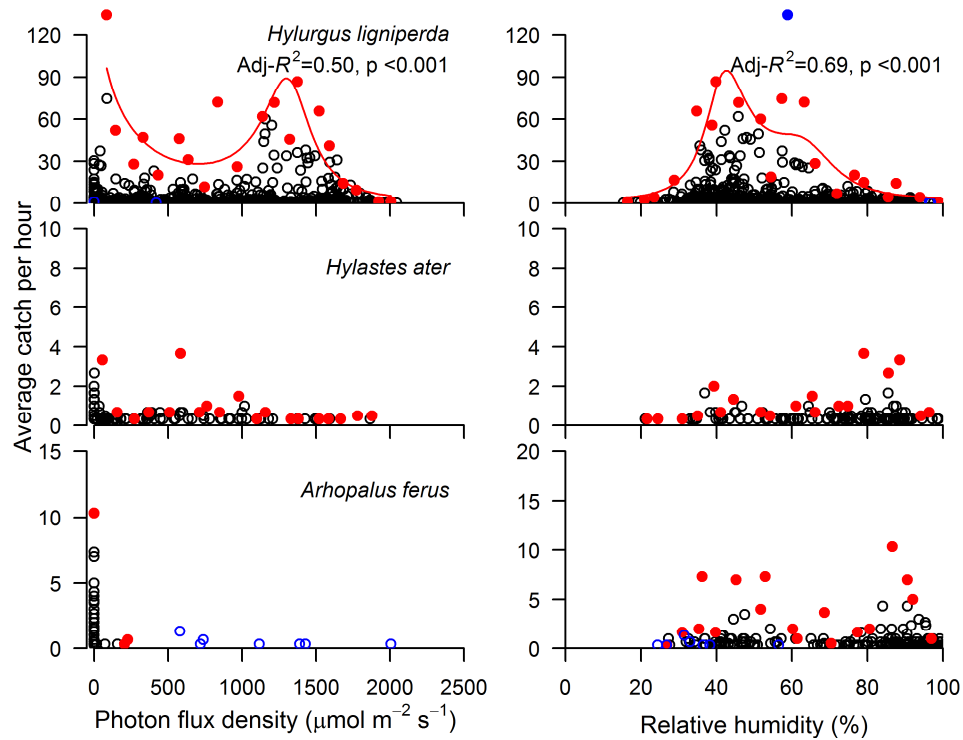


Fig S4. Average catch per trap per hour as a function of photon flux density and relative humidity.

Red symbols indicate the maximum trap catch per interval, solid blue symbols represent model outliers with a Cook's distance < 1 , and open blue symbols indicate suspect trap catch removed from analysis (see results). Model predictions are only shown when a significant relationship was present between the predictor and response variable. The adjusted R^2 provides an estimate of the variance explained by the non-linear curve fitting and are only provided when a non-linear response is identified.

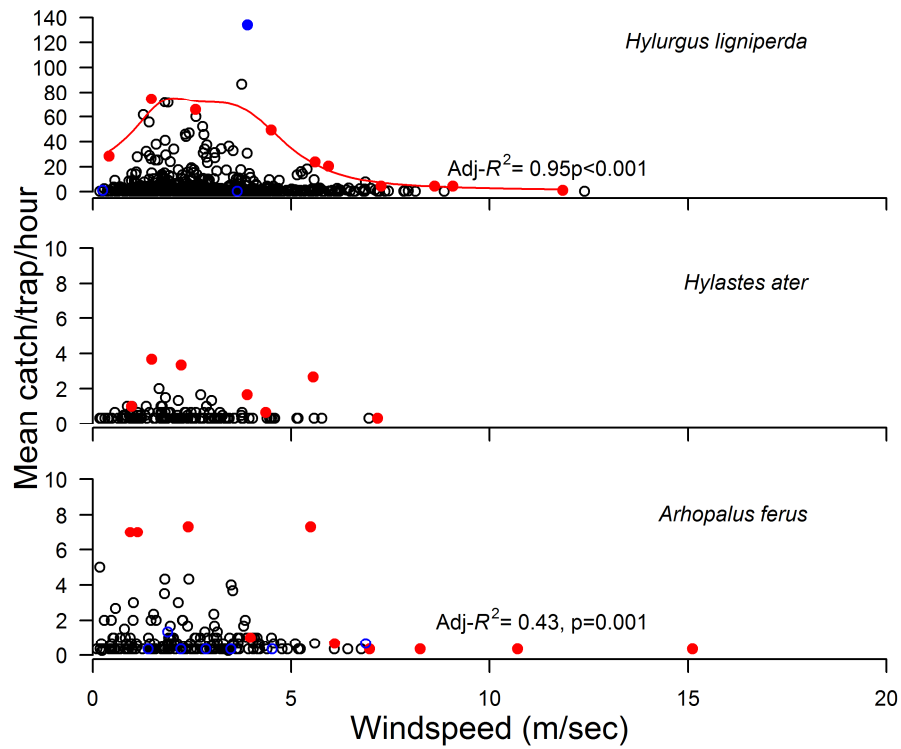


Fig S5. Average catch per trap per hour as a function of wind speed.

Red symbols indicate the maximum trap catch per interval, solid blue symbols represent model outliers with a Cook's distance < 1 , and open blue symbols indicate suspect trap catch removed from analysis (see results). Model predictions are only shown when a significant relationship was present between the predictor and response variable. The adjusted R^2 provides an estimate of the variance explained by the non-linear curve fitting and are only provided when a non-linear response is identified.

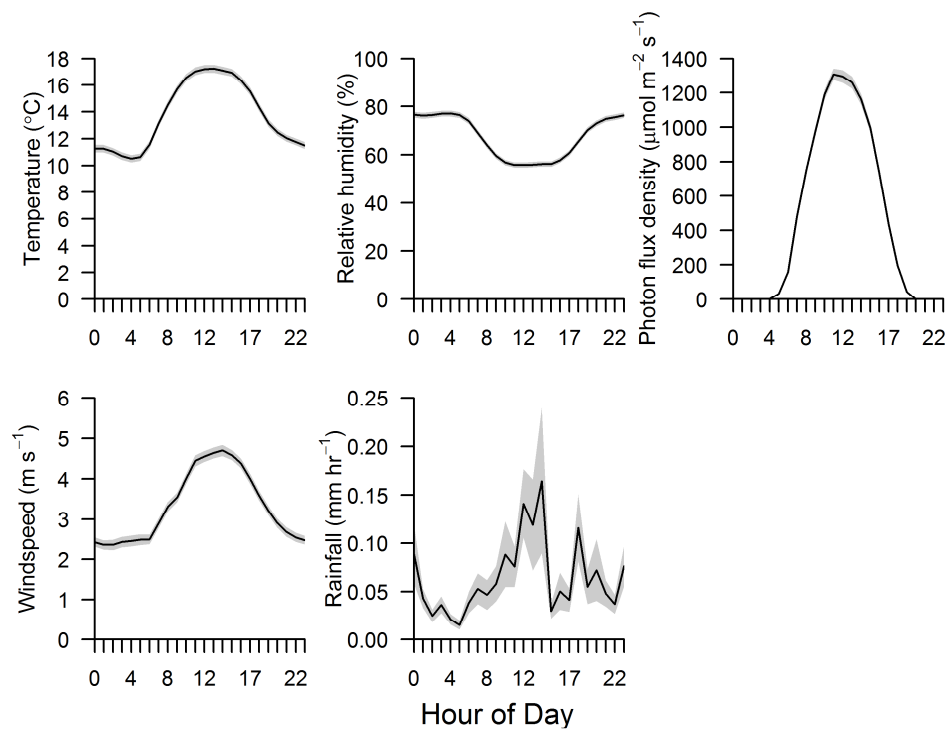


Figure S6. Mean hourly meteorological conditions as a function of the time of day averaged across all sites. The grey polygon defines the standard error of the mean for each hourly measurement.