Early prediction of bumblebee flight task using machine learning

Morton Williams, Samuel; Aldabashi, Nawaf; Palego, Cristiano; Woodgate, Joe; Makinson, James; Cross, Paul

Computers and Electronics in Agriculture

DOI:
https://doi.org/10.1016/j.compag.2021.106065

Published: 01/05/2021

Peer reviewed version

Early Prediction of Bumblebee Flight Task Using Machine Learning

S. M. Williams[a,*], N. Aldabashi[a], C. Palego[a], J. L. Woodgate[b], J. C. Makinson[c], P. Cross[a]

[a] *College of Environmental Sciences and Engineering, Bangor University, Wales, LL572DG*
[b] *School of Biological and Chemical Sciences, Queen Mary University of London, London, England, E1 4NS*
[c] *Hawkesbury Institute for the Environment, Western Sydney University, Penrith, New South Wales, Australia*

Abstract

This work demonstrates the development of a neural network algorithm able to determine the function of a bee's flight within six measurements (≈18 seconds with current radar technology) of its relative position on leaving a nest. Engineering advancements have created technology to track individual insects, unlocking research possibilities to investigate how bumblebees react to their environment in more detail. This includes how they discover and make use of resources. The development of an intelligent algorithm would allow for the automated monitoring of resource use and nest health. An imbalance of bee flight tasks may indicate a shortage of resources or over-reliance on a plant that may soon stop flowering. Recent developments using drones to track insects can benefit from an intelligent target acquisition system given limited drone battery life. Such knowledge will also benefit the tracking itself by allowing for customised flight parameters to match target flight patterns. Data captured by these tracking techniques are taxing to parse manually using human expertise. Artificial intelligence can produce meaningful knowledge faster with equal precision. In this work, a comparison between a neural network (NN), random forest (RF), and support vector machine (SVM) is provided to distinguish the best model for the task by comparing cross entropy loss and accuracy across the dataset, showing improved results as time goes on. In situations where the radar lost sight of the target, a purpose-built filter was created to mitigate signal losses. The generated model provides results with a peak accuracy of 92%. This model,

*Corresponding author
Email address: eeu816@bangor.ac.uk (S. M. Williams)

combined with the filter, create an opportunity to monitor the number of bees leaving the nest for each flight task with smaller, cheaper, and stationary receiver solutions with shorter ranges by removing the need to track a bee for its entire flight to ascertain its errand.

Keywords: Bombus, Pollinator, Decline, Machine Learning, Insect Tracking, Harmonic Radar

---

1. Highlights

- An accurate model to predict bumblebee errand within 6 measurements.
- Best model chosen from Neural Network, Random Forest, and Support Vector Machine.
- Final hit rate of over 90%.
- Includes a filtering process to mitigate losses from radar technology.
- Useful to support insect tracking systems, pollination services, and nest monitoring.

2. Introduction

Despite a long history of using nest-building insects such as bumblebees (*Bombus terrestris*) in human agriculture, our understanding of them has remained limited. It has become clear that pesticides, parasites, and climate change have all had a significant impact on bee behaviour and population decline (Thompson, 2003; Williams, Paul H. and Osborne, Juliet L., 2009; Potts et al., 2010). In particular, it has been shown that pathogen spillover from commercially produced bumblebees to native wild bees can happen when the two groups share common flower food sources (Colla et al., 2006).

Historically, only 6 of 19 British Bumblebee species maintained their ranges between 1960 and 1980 (Williams, 1982). Bumblebee species are projected to decline significantly in North America by 2070 (Sirois-Delisle and Kerr, 2018). Bumblebee species richness reduction is expected to impair pollination services, with a consequential effect on food yields and human welfare.

Maximising the efficiency of colonies in the face of new pressures is now an opportunity to learn more about bees while increasing investment return (Potts et al., 2016). Using radar technologies allows for recommendations to

be made based on the real-time evaluation of colony activity. Bumblebees, in particular, are important pollinators for soft fruit such as tomato (*Lycopersicon esculentum*) with up to 50 colonies used per hectare during the growing season. The value of these crops is estimated to be €12 billion (Hayo H.W. Velthuis and Adriaan van Doorn, 2006).

Recent technological advances are bringing the possibility of intensive, life-long tracking of animal movements (Kays et al., 2015). This may lead to an explosion in the acquisition of movement data comparable to the effect of DNA sequencing (Nathan et al., 2008). However, for this increase in data availability to lead to an equivalent leap forward in our understanding of behaviour, we need appropriate analysis techniques that are scalable to very large datasets and which provide genuine insight into the behaviours being measured.

Classification of movement data into distinct activity types is one such area in which machine learning can speed up analysis and improve accuracy. Using bumblebees as an example in a commercial setting, if many individuals across multiple nests were tracked across lifetimes to determine pollination efficiency, the volume of data would necessitate advanced tools able to contextualise the information.

However, making sure that any classification matches are similar to that done by experts is paramount. This leads to the importance of machine learning algorithms using human insights, such as the classification of exploration and exploitation flights, to ensure that the outputs make biological sense.

The use of machine learning in aiding the monitoring of insect nest health has become a key area of research in recent years. Studies involving the prediction of insect species via sound, using low-cost camera equipment for inter-nest tracking, and behavioural analysis are just some examples of how machine learning is aiding in understanding the complex behaviour of insect colonies (Kawakita and Ichikawa, 2019; Boenisch et al., 2018; Blut et al., 2017).

These studies, and most others, are limited to observing colonies in contained spaces, without any positional information, or through metrics such as nest weight, temperature, and humidity (Rafael Braga et al., 2020).

The use of Radio-Frequency (RF) technologies bypasses these limitations and represents a novel way to monitor, predict, and understand these creatures. Principally, RF offers a potential understanding of how to encourage bees to utilise local flora efficiently. Similarly, noting how the local balance of flora affects nest longevity and reproductive success. Commercially, it will contribute to best practices for planting crops, wild-flowers, field margins, and gardens to

maximise pollination potential and benefit bee populations.

This work is concerned with the rapid classification of bumblebee flight task using data relating to detailed bee positional information during flight. In particular, the data acquisition methods as described by Woodgate (Woodgate et al., 2016) provide an opportunity for additional investigation focused on the automatic early classification based on flight patterns.

Given recent developments in tracking technologies for individual insects, such as by Shearwood (Shearwood et al., 2018), classification of flight could serve as an auxiliary aid to deploy tracking techniques that match the characteristics of the flight at hand.

Recent insect tracking technology developed at Bangor University can be mounted on a drone and is significantly more portable than harmonic radar (Shearwood et al., 2020a,b). Classification of initial bee flights could allow this type of drone tracking system to prioritise targets, and consequently maximise tracking focused flight by making best use of a limited battery life (approx. 40 minutes). As a conceptual design (Fig. 1), an intelligent target acquisition system could provide the drone with detailed flight behaviours and predicted bee tasks to augment the drone's own flight parameters to match the nimbler target.

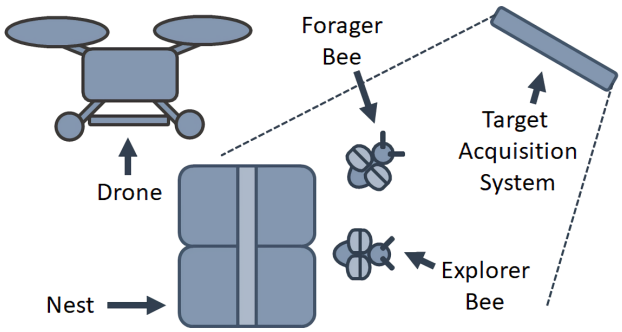

Figure 1: Conceptual design of a proposed system with target acquisition to prioritise targets for a battery-limited drone.

A robust algorithm for the early classification of flight purpose could make it possible to monitor how colonies of bees divide their labour resources between exploration for new floral resources and exploitation of those already known, in near-real-time. This would allow researchers or commercial users of bumblebee

pollinators to monitor the efficiency of pollination, colony health and predict future needs in time to respond flexibly to them: an increase in exploration flights might suggest that currently known resources are insufficient to support the colony; while a drop in exploitation flights could predict upcoming starvation. Furthermore, these methods would scale well, opening up the potential to monitor pollination services over large areas or allow researchers to investigate interactions between colonies in resource exploitation.

Woodgate et al. (2016) gathered data using harmonic radar at Rothamsted Research Station. Locations where bees stopped to forage were recorded by manually visiting the GPS coordinates to confirm exploitation behaviour (Woodgate et al., 2016). The authors described an algorithm for classifying flights into those which explore for new resources versus those focused on foraging from known floral sources. This algorithm was inherently simple and effective. An *exploitation* flight was a flight which consisted of a single loop where the bee stopped for a length of time at a location it had previously investigated. All other flights, including those with multiple loops to and from the nest, were considered *exploration* flights. In this work, we are concerned with improving the data acquisition of similar studies and investigate the automated early classification of these flights.

Individual bees seemed to follow a loose pattern of some initial exploration flights followed by periods of exclusive exploitation flights. This, in turn, could be included in the dataset as information to assist with prediction. This is true only if each bee can be uniquely always identified. RF techniques do not easily afford this. Therefore, it was decided to exclude this so that the algorithm generated can function with a wide range of technologies, some of which may not uniquely identify the bee.

For their assessment to work, Woodgate et al. required knowledge of the bees flight from start to end. The goal of this work was to generate an algorithm that could predict the bumblebee's task much earlier in its flight.

3. Materials and Methods

The dataset gathered by harmonic radar contains 244 flights, each flight consisting of polar coordinates of the bee's position and a timestamp taken when the radar scanned (360 degrees every 3s). Two classes of data exist within the dataset, namely flights labelled as those exploiting existing resources and those exploring for new resources.

Some fixed conditions could influence the final outcome such as experiments being undertaken during daylight hours in good weather. The primary concern with the dataset was that there were too many potential variables to fully describe all drivers that might affect bee behaviour. This range includes things as simple as local temperature up to small undetectable air currents.

Variable such as this were excluded to form a baseline. The algorithm was designed to work in ignorance of such factors so that a bumblebee flying in windy weather can be predicted with the same accuracy as a bee flying in calm weather.

The acquired data were taken as coordinates representing the distance in meters from the source nest. Using this data in this study allowed for the extraction of meta-data such as current speed, average speed, distance from the nest, and perpendicular distance from the average bearing.

An additional metric was used as described in the original work, named digressiveness. This measure was a numerical representation of flight efficiency, with a value of 1 representing a perfectly efficient flight. Given that the work here focuses on machine learning, a small adjustment was made to this metric to enable better normalisation. This original equation is as shown in Equation (1) with the changes made as in Equation (2). Flight distance is the sum of all vectors between detected positions, and the optimal distance represents a straight line between start and endpoints.

$$D = \frac{Flight\,Distance}{2 * Optimal\,Distance} \tag{1}$$

$$D = \frac{Flight\,Distance}{Optimal\,Distance} - 1 \tag{2}$$

These changes were made as the emphasis is on early detection of a task, where the insect may not as yet have returned to the nest, therefore only the outward journey is important necessitating the removal of the factor of two in the original equation. Similarly, having a perfect flight represented by zero allows for better use of the metric in learning models. Neural networks, in particular, expect the data to be normalised between zero and one. By having the most efficient flight defined as a value of one, the normalisation would lose the full range of potential values.

With this new equation, a perfect flight would be D = 0 rather than D = 1 as per the original specification, allowing better normalisation of digressiveness

over the series. A simple demonstration is present in Fig. 2, showing the digressiveness value for a selection of lines.

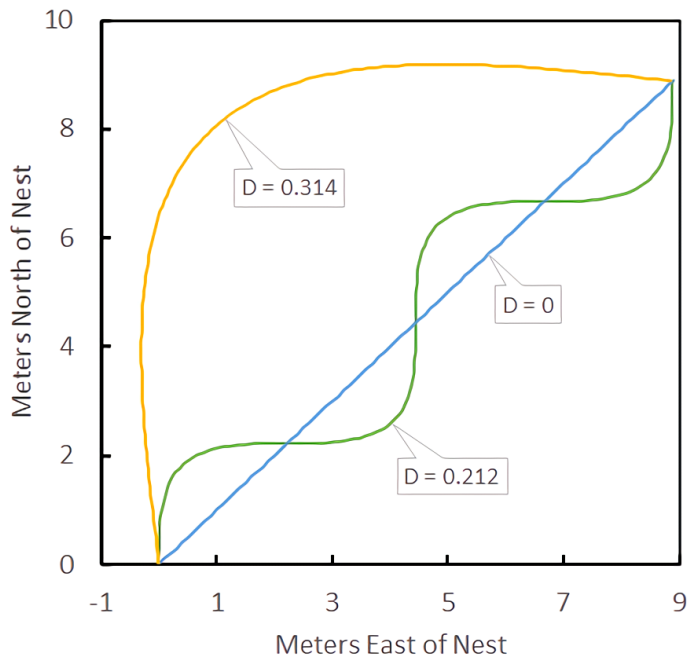

Figure 2: Example digressiveness metrics for hypothetical bee flight patterns.

It is noted that additional metrics could be used that describes the bumblebees' trajectory from the nest, which would likely aid in classification such as the exact heading from North which would highlight the bearing of flowers the bee is targetting. However, this work aimed to strip positional information from the data to allow any generated models to work with other colonies with different geographical features. In particular, by not including bearing as a metric the characteristics of the landscape for this nest are not dominant in determining predictions. Because the data is expressed in more general terms such as distance from the nest, current speed, and average speed the data from other colonies can be substituted effectively.

There is a caveat when using the data in that localisation error can occur (Woodgate et al., 2016). For instance, occlusions such as those caused by the bee flying behind a tree would mask the detection leading to dropouts in the dataset. Other perturbations exist, requiring the filtering of data. An example

of this is shown in Fig. 3, detailing the missing elements of an otherwise strong track. In the case of flight 83, there is a large distance gap which is unaccounted for that must not be allowed to affect the data. Similarly, in flight 130, the bee was likely busy collecting pollen from shrubbery for some time which led to it being occluded. In the current work, the concern is with when the bee is both moving and able to be seen therefore both tracks must be filtered down to the relevant parts.

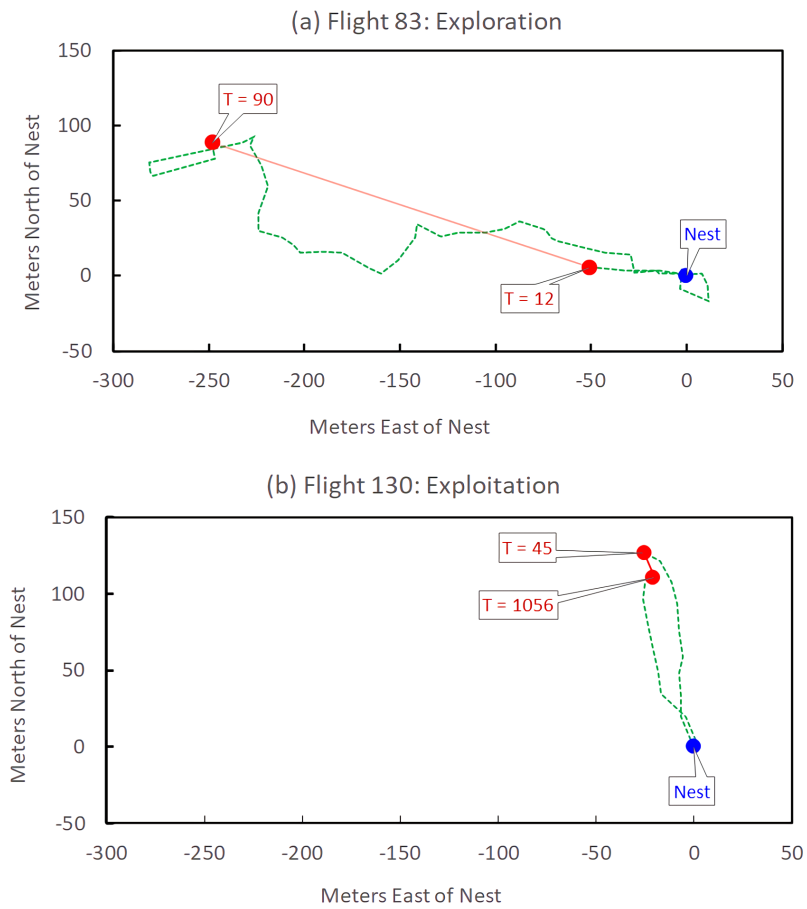

Figure 3: Details of (a) flight 83 and (b) flight 130 showing lost segments of flight (as solid red lines), with T being time since track beginning.

As the work was focused on the early prediction of bumblebee flight task, a sampling window of 50 positional readings was used. This is the first 50

positions read by the radar if they satisfy the data filtering process.

The chosen approach for filtering the data was to focus on three variables; current speed, perpendicular deviation from current bearing, and digressiveness. Perpendicular deviation takes the bearing of the flight, including the current point, and evaluates the absolute distance between the average bearing and the bee's true position, serving as a method of understanding how quickly the bee changes direction without using specific positional information.

The distribution of unfiltered data was explored to find the point at which growth becomes exponential within the set. Final values for the filter were determined to be a max speed of the bumblebee of 8 meters per second, a perpendicular deviation of no more than 8 meters, and a digressiveness of 2 or less.

The filtration of digressiveness is key as some rare flights had multiple, short distance occlusion losses. Coupled with the radar's accuracy ($\pm 2m$) this could give the appearance of the bee looping multiple times in a short distance, skewing results.

It is also noted that the limitations on speed are less than with similar research where the maximum speed of a bumblebee is around 15 meters a second (Osborne et al., 1999), however, this could be due to previous research noting a range of 3 to 15 meters a second depending on environmental conditions such as wind resistance.

Presented in Fig. 4 are the normalised distributions of the dataset. Some machine learning algorithms, such as neural networks, struggle to separate two classes differentiated by exceptionally small values, especially given intrinsic rounding errors. The filter allows better distribution of the data for machine learning with much less obscured by having compacted values.

The filtered data was divided into two similar-sized sets for both classes (n≈1130 each) before being inputted into the following models;

- A random forest classifier (RF (Ho, 1995)) of 1000 nodes initialised to a random state.
- A support vector machine classifier (SVM (Pedregosa Fabian et al., 2011)) with a radial basis function kernel, regularization of 1, and a kernel coefficient of 0.5.
- A neural network (NN) with two hidden layers of scaled exponential linear units (SELU (Klambauer et al., 2017)) with a final sigmoid activation layer

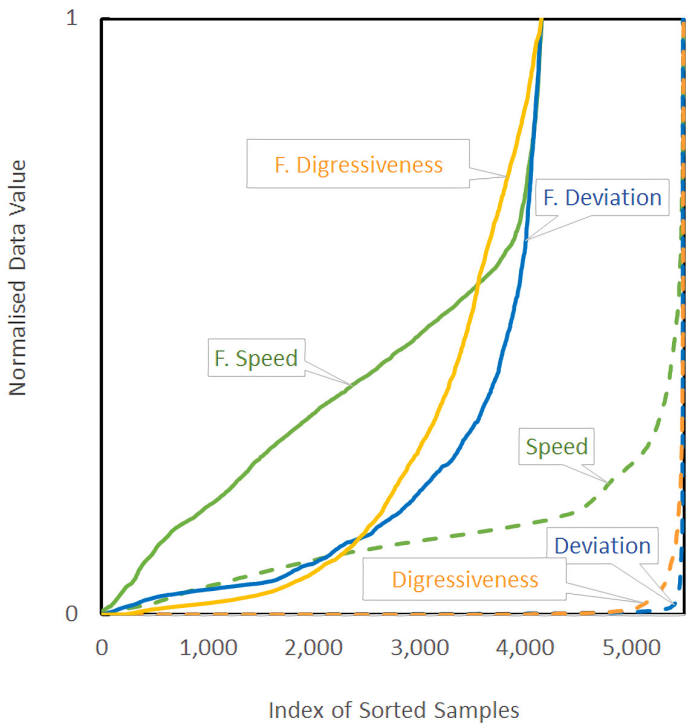

Figure 4: Filtered (Solid line) and unfiltered (Dashed line) normalised data distributions of the data set.

and a dropout rate of 0.5.

The size of each dataset is the number of points present in the first 50 readings in each flight, filtered based on the discussed parameters. There was an approximate 3:1 ratio of available points for exploitation and exploration respectively pre and post-filtering. Neural networks are sensitive to imbalanced classes of data (Buda et al., 2018). In this instance, we used undersampling of the larger class to even the dataset. To strengthen the conceptual results of the work, we include learning outcomes for both filtered and unfiltered data. This allows the demonstration of patterns discovered from filtered data matching patterns in unfiltered data.

To confirm prediction validity, both accuracy and loss were evaluated. Binary cross-entropy loss, also known as log loss, was used as a loss function for all the models generated (Ma et al., 2004). This loss can loosely be interpreted as the proportion of correct predictions produced by the model in a set, in addition to its confidence in those predictions. A perfect loss would have a value of zero, as demonstrated in Equation 3. In this case, i is the index of a given prediction, y_i represents the target value output and $\hat{y}_i$ is the predicted value.

$$loss_i = y_i \cdot log\, \hat{y}_i + (1 - y_i) \cdot log(1 - \hat{y}_i) \tag{3}$$

To calculate the loss for both the RF and SVM, these models were created in scikit-learn with SVM probabilities determined by Platt's method (Platt and Others, 1999).

For the original learning outcomes, 8:2 ratio was used to split data into training and testing sets for the models, with the neural networking taking a small sample (20%) of its training set to act as a validation set. In this method, the flights were disassembled into constituent context points and reformed into randomised composited sets.

In the case of the neural network, the check-pointed model with the highest accuracy against the validation set was used as the final outcome to limit any residual overfitting.

For predictions over time, 180 flights were used as training with 64 flights held for testing. As the filtering process could shorten a flight to less than 50 measurements, three separate instances of each model were trained. One triplet of models would contain one NN, one SVM, and one RF. Each triplet of models shared the same set of flights. The average accuracy across sets of triplets was used as the final measurement.

In essence, each of the three datasets was a random permutation of the original set, split further as discussed into a training and testing set. This served the purpose of reducing possible bias introduced by the training data having a disproportionate number of flights with the full 50 possible values versus the testing set.

An important observation of this data is that measurements were taken every 3 seconds by the radar. The initial measurement was not guaranteed to be the first possible point of the bumblebee leaving the nest as the bee could be obscured by the nest itself or masked by the rotation of the radar. To counteract this, predictions are made from the second point onwards so as not to make faulty assumptions as to a target's current speed.

It is also prudent to mention that Woodgate et al noted that their classification of flight task used an algorithm created to match human observation (Woodgate et al., 2016). However, they note that such classification was not necessarily suited to the nuances of bee behaviour. Notably, they discuss the trade-off between capturing what was happening precisely while also creating a method that was as simple and universal as possible. This creates the possibility that where the machine learning structures disagree with the original method, some of these instances may be due to the more complex nature of machine learning and in fact represent the ground truth. This also formed part of our study.

To investigate further, the unfiltered data were explored using Ward's method of clustering to build a set of data labels by calculating the incremental sum of squares. Following this, the method works by creating a simple nearest centroid classifier to estimate a label based on proximity to the nearest Ward cluster centroid (Jr., 1963). As shown in a dendrogram present in Fig. 5, there are indications that there exist multiple sub-flight types, with ten clusters providing the best distinctions in our study. The goal was to determine if more than the described two categories of flight exist by predicting them over a small set of flights. This also allows exploration of whether multiple flight categories can be present in a single flight, indicating the possibility of either adaptive tasks or strict tasks.

4. Results and Discussion

A final accuracy, on the initial 50 readings of each flight, was achieved as 91%, 81%, and 85% for the neural network, support vector classifier, and random

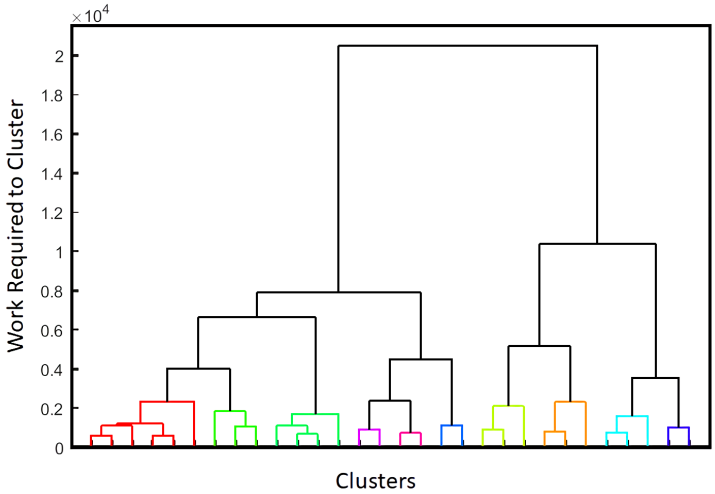

Figure 5: Dendrogram of unfiltered data, with the cluster threshold set to ten to match cluster analysis.

forest classifier respectively. This is the result when flights were disassembled and individual points re-composited to form sets, stripping them of intra-flight relationships. These are detailed in Table 1, showing that the random forest performs very similarly regardless of filtering the data, even doing somewhat better with unfiltered data. Both the neural network and the support vector machine benefit from filtering the data, however, the SVM does have lower loss with unfiltered data. These results do indicate that the patterns within the data exist both with and without filtering.

As the dataset was limited to 244 flights, it was not possible to disaggregate the dataset by either ambient temperature or time of day. For example, the random forest improved to 90% accuracy and 0.27 loss when the time of day was included. However, this may be an overstatement of the algorithm's capabilities. After approximately 6pm, all recorded flights (6 total flights) were exploitation flights. This could mean that bees leave the nest that late in the day only for food or that there was not enough data to capture the truth of the matter, so the algorithm would likely (incorrectly) assign exploitation to all flights in this slot. More flights would provide the correct ratio of classes for the algorithm to attribute a label. Given this limitation, time was not used though it may be reincorporated as the dataset is further expanded. This also supports previous arguments in favour of keeping the variables as general as possible.

Table 1: Learning Results: Strongest results in bold.

Method	Filtration	Accuracy	Loss*
Neural Network	Filtered	**91%**	**0.42**
	Unfiltered	75%	0.60
Random Forest	Filtered	85%	0.35
	Unfiltered	**87**%	**0.32**
Support Vector Machine	Filtered	**81**%	0.59
	Unfiltered	71%	**0.55**

*Binary Cross Entropy Loss (Ma et al., 2004).

More interesting for this study was the accuracy versus flight-time. As each subsequent point contains more context about the flight, such as a more refined value for average speed, it is expected to see an increase in accuracy over time.

It proved prudent to evaluate the results as an average of models across three sets. To take the example of the neural network, results were 50%, 69%, and 47% for each triplet respectively on the initial prediction (2 measurements taken.) While some sway is to be expected, there is a 22% accuracy shift between the weakest and strongest result. For comparison, the RF managed 74%, 71%, and 73%. This is a much more typical sway with a gap of 3% accuracy.

However, as time progresses this shift tapers out such that at the peak accuracy of the neural network, the results are 94%, 93%, and 90%. Comparatively, at this point, the RF returned an accuracy of 79%, 81%, and 72%, now producing a 9% shift in accuracy.

The likely reason behind these shifts is due to imperfect filtering of the data, a smaller number of test flights than ideal, a residual error left from the radar itself, and the discussed random flight order in the dataset. As previously mentioned, the first reading of the bee was not always adjacent to the nest which created the possibility of feed-forward error as points taken for learning purposes contain continuously more context about the flight. Flights were validated on whether their initial distance from the nest (x) was less than or equal to the distance between the first and second reading (y). Assuming acceleration upon leaving the nest, the bee must have been airborne for longer than the three second rotation of the radar if x is greater than y. 53% of recorded flights began with the bee already beyond y distance from the nest, making speed at these early points unknowable.

With a larger number of flights recorded an adaptive filter could be developed

such that these issues may be addressed. For example, having a lower speed limit at the start of a flight. This might be expected as the bee accelerates away from the nest and would serve to curtail the use of flights where the bee had already covered some distance. With 244 flights, such an adaptive filter would lack the proper context to correctly determine a meaningful value for this speed.

The binary cross-entropy loss across model triplets is much more stable, though the trend for SVM results to spike remains. This supports the idea that accuracy issues stem from data points which are either incorrectly filtered or form part of a set that was labelled incorrectly by the original algorithm.

As shown in Fig. 6, the loss values are strongest with the random forest. The SVM results lack stability, and while both the neural network and random forest start competitive, the random forest in time outpaces the competition.

Conversely, with accuracy the results favour the neural network. It is worth noting that both the random forest and support vector classifiers manage the initial classification with higher accuracy than the neural network. However, the neural network reaches its highest accuracy faster, with 81% at 4th measurement (12 seconds) and a peak of 92% at the 6th measurement (18 seconds.) The neural network is also the strongest model over the entirety of the dataset.

Observing both loss and accuracy together, the neural network is strongest. Before 18 seconds, the neural network and random forest are almost equal in loss value but the neural network quickly climbs to its peak accuracy whereas the random forest takes much longer to reach similar results.

Another way of interpreting these results lies in looking at model predictions on specific flights. Fig. 7 shows two flights, one exploration (flight 233) and one exploitation (flight 89.) Flight 233 had a total flight prediction accuracy of 90%, though even at its third prediction it made a mistake. Flight 83 had a final accuracy of 75% and made a substantial error on its ninth prediction by producing the incorrect results with almost certainty.

Both flights produce a majority vote in favour of the correct result within the first four measurements, supporting the initial findings presented here.

Interestingly, these early results provide context surrounding some of the incorrect predictions created by this approach. Flight 64 is an exploration flight, yet the neural network predicted only 10% of the points correctly. Context is provided in Fig 8 and shows that for the first 12 measurements, this flight bears a striking resemblance to flight 89. However, the full flight plot shows the characteristic looping and backtracking associated with exploration flight.

As mentioned, cluster analysis was also performed. No filter was used, how-

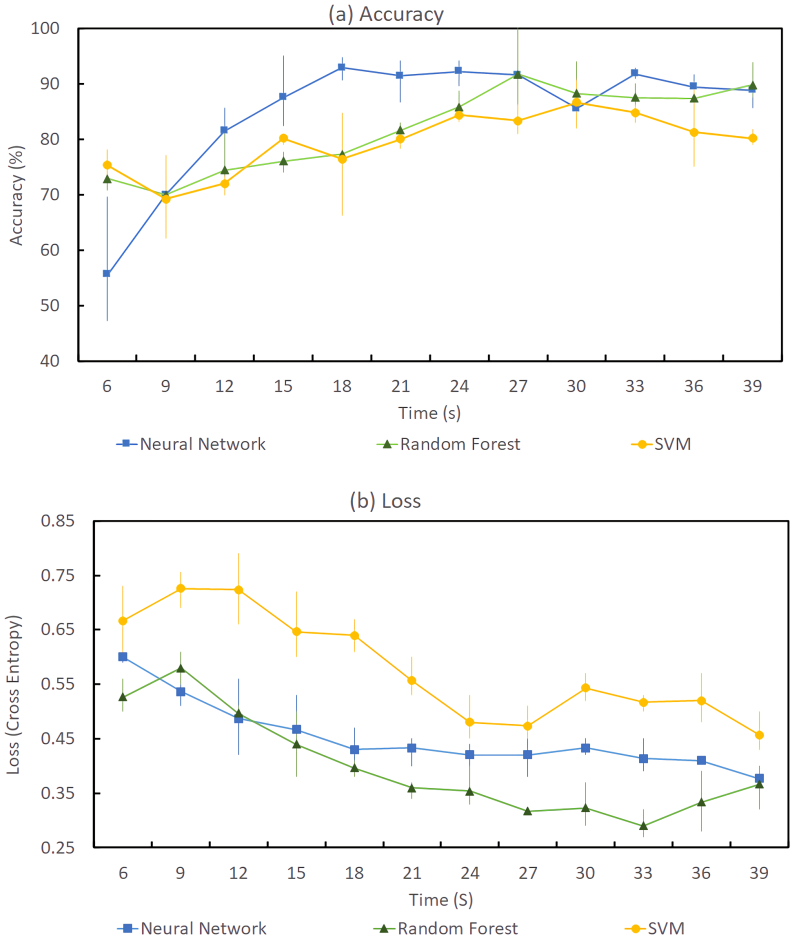

Figure 6: (a) Average accuracy and (b) loss over time of the models trained on the dataset. Error bars are for the min and max values across model triplets.

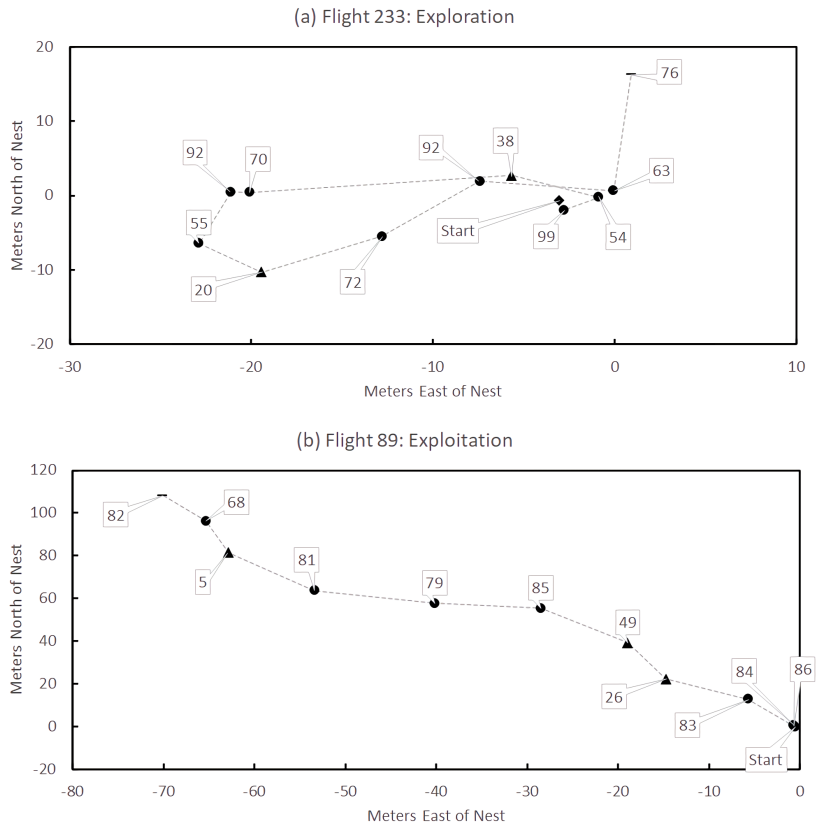

Figure 7: First 12 positions of two flights, (a) flight 233 and (b) flight 89, with confidence scores. Scores below 50 indicate an incorrect prediction represented by ▲. A score of 0 would indicate total confidence in an incorrect prediction, a score of 100 indicating perfect prediction. Correct predictions are marked by ●.

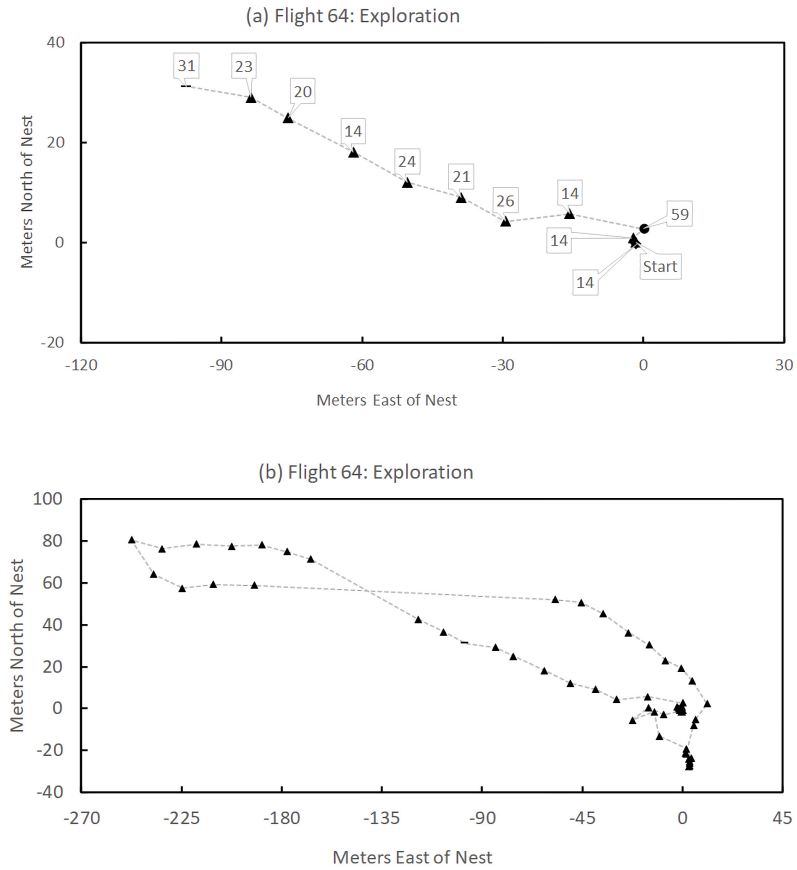

Figure 8: (a) First 12 samples of flight 64 with predictions plus (b) total plot of flight 64 showing characteristics of exploration.

ever as with the models themselves, only the first 50 measurements were evaluated. The goal was to determine if there is justification for there being more than two types of flight present in the dataset and whether flight 64, in particular, could be classed as a hybrid flight.

Ten clusters were formed from the entire dataset minus flights 64, 89, and 233. With these cluster centroids, predictions were undertaken using the one nearest neighbour classifier approach. Results of predictions for flight 64 in Fig. 9 show that segments of the flight remain intact, rather than there being randomly clustered data. This indicates that even a relatively simple algorithm can segment a flight into similar sections. Furthermore, there appear to be repetitive clusters based on flight parameters. Using the example of cluster 5 (C5), it is present before the major arc of the flight and also immediately following, indicating the same behaviour both before and after this arc.

It is important to note the functions of the clusters would require further analysis. This would also likely need a larger sample of flights to provide better definitions. However, it is important to evaluate the possibility of hybrid flights, or perhaps flights that do not fall under the umbrella of either exploitation or exploration.

Looking at the composition of the three flights in question, again shown in Fig. 9, shows that it is likely that 64 does fall under a hybrid category. Like the exploitation flight 89 that it was mischaracterised as, it contains four component clusters rather than the two present in the exploration flight 233. On the two overlapping clusters C5 and C6, flight 64 falls directly between flight 89 and 233 in terms of proportion. The exploration flight is strictly composed of these two clusters, with an overwhelming majority of C6 (88%). Flight 64 is much more similar to 89 with 52% and 38% respectively.

This is evidence that hybrid flights exist. In this particular case, it could be that the bee went in search of additional food after a first visit. Further work could investigate correlating clusters with other behaviours.

It is important to note that the final validation for the algorithm will be to test in the field. Real-time execution of the algorithm with third-party assessment would refine measurement precision. This could be in the form of observing bee pollen load on return to differentiate between foraging and other flight types. Additional harmonic radar data gathering will enable further classification of the detected clusters (for example predation and disease). Early classification of behaviour types can improve unmanned drone air-time both in prioritising bees to save battery life and in enabling adaptive flight to match

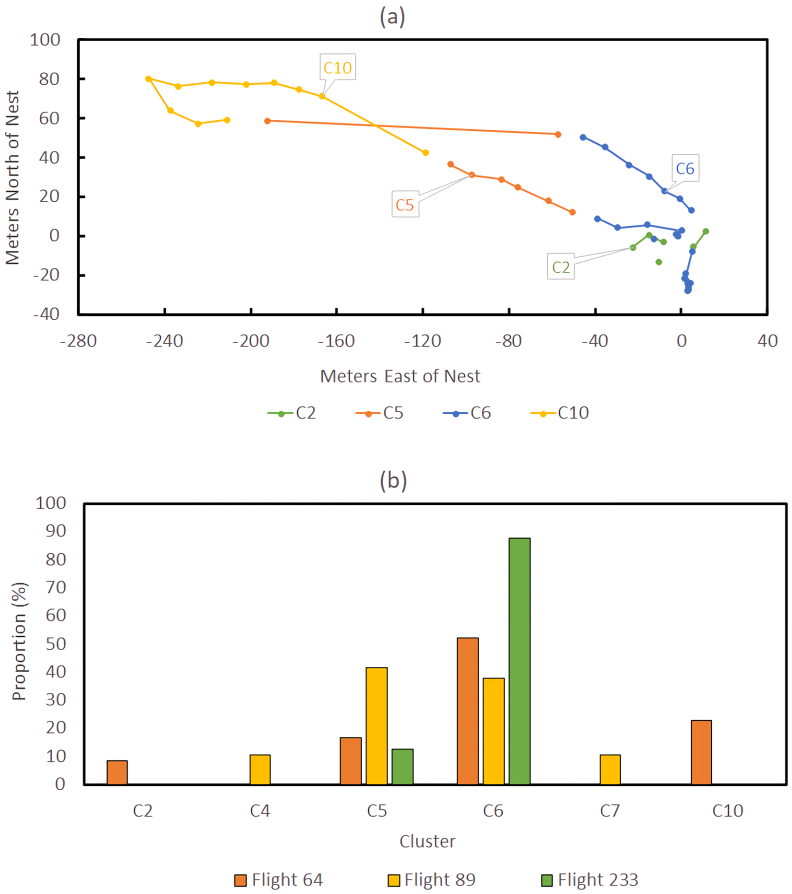

Figure 9: (a) Flight 64 labelled with nearest centroid classifier and (b) proportional constitutions of flights 64, 89, and 233 of their component clusters.

bee flight patterns. This in turn will serve as more proof that the algorithm is correctly predicting to match field observations.

5. Conclusion

This work shows that a rapid automated prediction of bumblebee flight task can be achieved with strong accuracy. Results become more than 80% certain after just 4 measurements with a peak accuracy of 92% achieved by the 6th measurement. This is a faster assessment than previously possible and does not require knowledge of the bee's full flight. Predictions distinguish between two classes, previously labelled exploration and exploitation. Potential exists for these to be conglomerates of multiple flight types. This algorithm will aid future systems to gather more data for the distinction beyond the two included classes. Prediction is generalised to other bumblebee nests because the raw data was abstracted as current and average speed, perpendicular deviation from average bearing, distance from the nest, and digressiveness.

These results open pathways to expand on the radar tracking of insects by allowing fast determination of flight errand. This could allow for automatic prioritisation of exploring bees over foraging bees for longer range tracking to build maps of nest resource acquisition. Similarly, it would allow for a shorter range system to sit near a nest to monitor the number of bees leaving for each task over time, without needing knowledge of a bee's destination.

With more development, this would lead to being able to monitor resource use and pollination efficiency in near-real-time so that interventions can be made to improve them, such as moving nests to help them make better use of crops or providing supplementary food when colonies are in need.

Further work could augment the process to predict the future needs of the colony and nest health. Too many exploration flights might suggest they are not getting enough food; too few exploration flights might suggest the colony is over-reliant on a small number of food sources and will be in trouble if those plants stop flowering. Other deviations from normal behaviour may in turn indicate other issues such as disease or pesticide ingestion.

This could scale to monitoring multiple colonies over a large area, such as an entire farm, and allow for moving colonies to areas that are not getting enough pollination. In addition, it would be possible to monitor wild bees to work out where conservation resources should be concentrated and researchers can start to look at how the foraging decision of one colony affect another. This could provide insights into how the pattern of nest foraging vs exploring affect the nest's overall health.

Given the 3-second time delay of the tracking system, a faster system might

be able to produce better results. The additional resolution offered by more measurements in a shorter time frame offers a potential boost in performance. This may result in classification being possible in a shorter period than the demonstrated 12 to 18 seconds which would further reduce the range requirement of a classification device, thereby potentially making them easier and cheaper to manufacture.

These models represent proof of concept that real-time evaluation of bumblebee task can be carried out successfully and could aid in automated tracking solutions for bumblebees and other colony insects.

6. Acknowledgements

This work was supported by KESS 2. Knowledge Economy Skills Scholarships (KESS 2) is a pan-Wales higher level skills initiative led by Bangor University on behalf of the HE sector in Wales. It is part funded by the Welsh Government's European Social Fund (ESF) convergence programme for West Wales and the Valleys.

Part funding by WEFO Solar Photovoltaic Academic Research Consortium II (SPARC II) is gratefully acknowledged.

The original radar work was funded by European Research Council Advanced Grant no. 339347:RadarSpacePollinator and special thanks are extended to Lars Chittka. J. L. Woodgate is currently supported by EPSRC program grant Brains-on-Board (EP/P006094/1).

References

Blut, C., Crespi, A., Mersch, D., Keller, L., Zhao, L., Kollmann, M., Schellscheidt, B., Fülber, C., Beye, M., 2017. Automated computer-based detection of encounter behaviours in groups of honeybees. Scientific Reports doi:10.1038/s41598-017-17863-4.

Boenisch, F., Rosemann, B., Wild, B., Dormagen, D., Wario, F., Landgraf, T., 2018. Tracking all members of a honey bee colony over their lifetime using learned models of correspondence. Frontiers Robotics AI doi:10.3389/frobt.2018.00035.

Buda, M., Maki, A., Mazurowski, M.A., 2018. A systematic study of the class imbalance problem in convolutional neural networks. Neural Networks 106, 249 – 259. doi:https://doi.org/10.1016/j.neunet.2018.07.011.

Colla, S.R., Otterstatter, M.C., Gegear, R.J., Thomson, J.D., 2006. Plight of the bumble bee: Pathogen spillover from commercial to wild populations. Biological Conservation 129, 461 – 467. doi:https://doi.org/10.1016/j.biocon.2005.11.013.

Hayo H.W. Velthuis, Adriaan van Doorn, 2006. A century of advances in bumblebee domestication and the economic and environmental aspects of its commercialization for pollination. Apidologie 37, 421–451. doi:10.1051/apido:2006019.

Ho, T.K., 1995. Random decision forests, in: Proceedings of the International Conference on Document Analysis and Recognition, ICDAR. doi:10.1109/ICDAR.1995.598994.

Jr., J.H.W., 1963. Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association 58, 236–244. doi:10.1080/01621459.1963.10500845.

Kawakita, S., Ichikawa, K., 2019. Automated classification of bees and hornet using acoustic analysis of their flight sounds. Apidologie doi:10.1007/s13592-018-0619-6.

Kays, R., Crofoot, M.C., Jetz, W., Wikelski, M., 2015. Terrestrial animal tracking as an eye on life and planet. Science 348. doi:10.1126/science.aaa2478.

Klambauer, G., Unterthiner, T., Mayr, A., Hochreiter, S., 2017. Self-normalizing neural networks, in: Advances in Neural Information Processing Systems. `arXiv:1706.02515`.

Ma, Y.D., Liu, Q., Qian, Z.B., 2004. Automated image segmentation using improved PCNN model based on cross-entropy, in: 2004 International Symposium on Intelligent Multimedia, Video and Speech Processing, ISIMP 2004. doi:10.1109/isimp.2004.1434171.

Nathan, R., Getz, W.M., Revilla, E., Holyoak, M., Kadmon, R., Saltz, D., Smouse, P.E., 2008. A movement ecology paradigm for unifying organismal movement research. Proceedings of the National Academy of Sciences 105, 19052–19059. doi:10.1073/pnas.0800375105, `arXiv:https://www.pnas.org/content/105/49/19052.full.pdf`.

Osborne, J., Clark, S., Morris, R., Williams, I., Riley, J., Smith, A., Reynolds, D., Edwards, A., 1999. A landscape-scale study of bumble bee foraging range and constancy, using harmonic radar. Journal of Applied Ecology 36, 519–533. doi:10.1046/j.1365-2664.1999.00428.x.

Pedregosa Fabian, F., Michel, V., Grisel Olivier, O., Blondel, M., Prettenhofer, P., Weiss, R., Vanderplas, J., Cournapeau, D., Pedregosa, F., Varoquaux, G., Gramfort, A., Thirion, B., Grisel, O., Dubourg, V., Passos, A., Brucher, M., Perrot, Édouardand, M., Duchesnay, Édouard, Duchesnay Edouard, F., 2011. Scikit-learn: Machine Learning in Python. Technical Report.

Platt, J., Others, 1999. Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. Advances in large margin classifiers .

Potts, S.G., Biesmeijer, J.C., Kremen, C., Neumann, P., Schweiger, O., Kunin, W.E., 2010. Global pollinator declines: trends, impacts and drivers. Trends in Ecology Evolution 25, 345 – 353. doi:https://doi.org/10.1016/j.tree.2010.01.007.

Potts, S.G., Ngo, H.T., Biesmeijer, J.C., Breeze, T.D., Dicks, L.V., Garibaldi, L.A., Hill, R., Settele, J., Vanbergen, A., 2016. The assessment report of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services on pollinators, pollination and food production. Technical Report. Bonn, Germany.

Rafael Braga, A., G. Gomes, D., Rogers, R., E. Hassler, E., M. Freitas, B., A. Cazier, J., 2020. A method for mining combined data from in-hive sensors, weather and apiary inspections to forecast the health status of honey bee colonies. Computers and Electronics in Agriculture doi:10.1016/j.compag.2019.105161.

Shearwood, J., Aldabashi, N., Eltokhy, A., Franklin, E.L., Raine, N.E., Zhang, C., Palmer, E., Cross, P., Palego, C., 2020a. C-Band Telemetry of Insect Pollinators Using a Miniature Transmitter and a Self-Piloted Drone. IEEE Transactions on Microwave Theory and Techniques doi:10.1109/tmtt.2020.3034323.

Shearwood, J., Hung, D.M.Y., Cross, P., Preston, S., Palego, C., 2018. Honey-Bee Localization Using an Energy Harvesting Device and Power Based Angle of Arrival Estimation, in: IEEE MTT-S International Microwave Symposium Digest. doi:10.1109/MWSYM.2018.8439173.

Shearwood, J., Williams, S., Aldabashi, N., Cross, P., Freitas, B.M., Zhang, C., Palego, C., 2020b. Localization and tracking bees using a battery-less transmitter and an autonomous unmanned aerial vehicle, in: IEEE MTT-S International Microwave Symposium Digest. doi:10.1109/IMS30576.2020.9223950.

Sirois-Delisle, C., Kerr, J.T., 2018. Climate change-driven range losses among bumblebee species are poised to accelerate. Scientific Reports doi:10.1038/s41598-018-32665-y.

Thompson, H.M., 2003. Behavioural effects of pesticides in bees - Their potential for use in risk assessment. doi:10.1023/A:1022575315413.

Williams, P.H., 1982. The distribution and decline of british bumble bees (bombus latr.). Journal of Apicultural Research 21, 236–245. doi:10.1080/00218839.1982.11100549.

Williams, Paul H., Osborne, Juliet L., 2009. Bumblebee vulnerability and conservation world-wide. Apidologie 40, 367–387. doi:10.1051/apido/2009025.

Woodgate, J.L., Makinson, J.C., Lim, K.S., Reynolds, A.M., Chittka, L., 2016. Life-long radar tracking of bumblebees. PLoS ONE doi:10.1371/journal.pone.0160333.