Technical Note

Note: Technical notes are aimed at promoting discussion. The views expressed are not necessarily those of the editors or the Australian Acoustical Society.

# REMOTE BEEHIVE MONITORING USING ACOUSTIC SIGNALS

**Amro Qandour[1], Iftekhar Ahmad[1], Daryoush Habibi[1], and Mark Leppard[2]**

**[1]Centre for Communications and Electronics Research (CCER), Edith Cowan University, Joondalup,WA 6027, Australia**

**[2]Communications Engineer, Rio Tinto, Perth, WA 6000, Australia**

**a.qandour@ecu.edu.au**

Recent developments in Wireless Sensor Networks (WSNs) have led to their use in remote data acquisition and automatic data analysis applications, which have proven to be an invaluable tool in a diverse range of fields including biosecurity. Further indications have been found that honeybee health can be monitored and determined through the use of acoustic analysis. In this paper, we present a system that has the ability to remotely detect the presence of pest infestation on a colony of honeybees by comparing the acoustic fingerprint of a hive to a fingerprint of known status. This will aid the goals of increasing surveillance programs by reducing the labour time and costs that are associated with managing and maintaining monitoring programs. Other benefits of the system proposed in this article include the ability to make available a collection of deterministic, standardised and nondiscriminatory statistical data for the purpose of research into determining the causes of colony collapse disorder.

## INTRODUCTION

The honeybee (Apis-mellifera) [1] is unquestionably considered as the most important and significant contributor among the animal pollinators, playing an essential role in the prosperity of the world's ecosystems and indeed to life itself. It is estimated that the honeybee is responsible for the pollination of over 90% of global commercial pollination services, and approximately 35% of the world's food crops [2]. The honeybee is probably best known for its production and storage of honey; however the economic value of the pollinator is not attributed solely to the hive produce, but largely to the products derived as a direct result from honeybee pollination. This constitutes an estimated $2 billion in revenue per year for Australia and $198 billion worldwide. In recent times, there have been rapid increases in agricultural development and human population, both of which are heavily dependent on the success of the honeybee industry. This has led to greater than ever demands for honeybee pollination, placing mounting strain on managed honeybee colony populations worldwide [3].

Conversely to this trend of increasing global demand, bee colonies around the world are under an increasing number of threats from a range of sources. The rapid spread of exotic pests such as the Varroa-destructor, better known as the Varroa-mite, is undoubtedly the biggest mortal threat to honeybees [4]. The Varroa-mite has already proven to be extremely damaging to the international honeybee industry as it has advanced throughout the world, and alarmingly, since the first reports of the arrival of Varroa-destructor in New Zealand in early 2000, Australia is now the only country free of the pest.

The work reported in this Technical Note has been undertaken with the support received from the AAS education award in 2013.

| | Freq. (Hz) | Signal Pattern | Sender | Possible Sig. |
|---|---|---|---|---|
| **Tooting** | 300 ~ 500 | Pulse sequence | Queen | Prevent hatching of further queens and trigger quacking |
| **Quacking** | 300 ~ 350 | Pulse sequence | Queen | Presence detection, viabilityof confined queens |
| **Hissing** | 300 ~ 3600 | Single pulse | Colony | Warning signal |
| **Piping** | 100 ~ 2000 | Single pulse | Scout | Triggers colony hissing, prepare for swarming |
| **Recruit** | 200 ~ 350 | Pulse sequence | Forager | Existence and quality of valuable food source |

Figure 1. Acoustic signatures of honeybee colonies

### Invasive pest surveillance

Currently there are surveillance programs operating on a state-by-state basis aimed at the early detection of the Varroa-mite and other foreign pests and threats arriving in Australia. This current beehive surveillance and monitoring is achieved through the use of bait hives, which are located around Australia's major harbours and ports. These hives are situated such that any foreign bee infested with unwanted diseases or pests arriving at a port will inhabit the bait hives before spreading further. It has been shown [5] that the early detection of pests is imperative if an infestation is to be contained and eliminated. Though there are treatments available, it is commonly agreed upon within the honeybee communities that the prevention of an outbreak is better than a cure. Through regular manual examinations of the bait hives and the bee colonies that settle in them, inspectors are able to ascertain if there is any potential threat and subsequently intervene with the necessary steps to eliminate the threat. To satisfy this requirement of increasing surveillance intensity with the current method of monitoring would necessitate large numbers of qualified inspectors to travel to every site individually. This would require large amounts of manpower, technical expertise, time and consequently money in order for their continued success. The lack of any better option is largely a result of the honeybee industry residing outside of the focus of modern technological developments.

### Technological intervention

Information from relevant sources [2, 6] have indicated that honeybees change their acoustic behaviour as a result of being exposed to certain stressors. This being so, it is reasonable that a potential solution to the pest infestation problem could be to use wireless sensor devices that can automatically detect the presence of infected colonies based on the colony's measured acoustic signature. Such a device would operate as a remote surveillance system and would aid current and future surveillance programs by providing the inspectors with tools such as advanced warnings and detailed analysis of hive health. In close to real-time, the system could provide alerts as to the arrival of a new colony to a bait hive, as well as the current health status of the new colony. This information would ideally include as much detail as possible regarding the what, where and when such warning has occurred. Consequently, this would allow inspectors to utilise their time in the most efficient and informed way, could potentially save large amounts of money, and revolutionise national biosecurity monitoring programs.

## BACKGROUND AND RELEVANT WORKS

Honeybees have been observed to produce a variety of different sounds [7, 8] as forms of communication within the colony. Most of the sounds produced have been characterised by a low fundamental frequency between 300 and 600Hz and their corresponding harmonics [8]. The sounds produced by honeybees are one of the primary forms of communication within the colony; however communication is also achieved through the use of chemical means [9]. As shown in Figure-1, there is a range of sounds of different acoustic frequencies used by bees for a variety of reasons. Interestingly to note, it is not only the range of frequencies that are produced that determine the meaning of the noise, but also the acoustic structure in terms of signal pattern. The accurate quantification of the characteristics of these signal patterns and frequency ranges will be the key in developing a system that can identify possible threats to the hive and colony.

Various studies [10, 11] have shown that the health, status and activity of a honeybee colony can be determined through the analysis of the acoustic characteristics of the hive. Through the analysis of these studies it becomes unquestionable that honeybee hive acoustics change to reflect the current status and circumstance of the hive. The idea of this project was to design and develop a system which could understand and recognise the different acoustic characteristics produced by a healthy colony and a colony infested with Varro-mites.

## SYSTEM OVERVIEW

This section deals with the main components of the beehive monitoring system (system architecture) and acoustical analysis techniques executed using dedicated acoustical software.

### System architecture

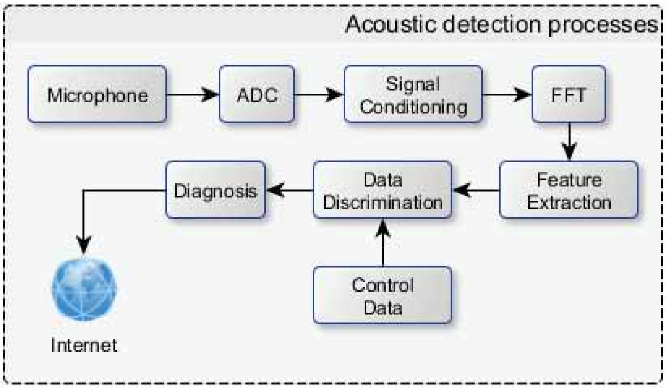


Figure 2. Acoustical detection process for honeybee colonies

The proposed honeybee monitoring system is designed to automatically acquire, process, and analyse audio data from remote honeybee hives to help alleviate the time and energy required for manual monitoring. To achieve this, the system must be as reliable and self-functional as possible. The key components of the honeybee monitoring system involve:

#### *Sensor node*

The Beagleboard [12] is an ideal platform for the honeybee system due to its miniaturised size (remote deployment) and intensive computational power (required for acoustical analysis tasks).

#### *Sensors*

The bee acoustics are acquired by using an electret microphone situated within the hive of the honeybee colony. The sound (honeybees) is picked up by the microphone and processed by the beagleboard using acoustical algorithms in order to discriminate between a healthy or infected hive.

#### *Radio transmission*

A low-power radio transceiver [13, 14] (Zigbee Link) is used to transmit the acoustical data from remote beehive sites to gateways. The data contains alarm messages in case an infection is detected, and diagnostics status to verify system operation (e.g. remaining power percentage).

#### *Algorithm for acoustical analysis*

The process depicted in Figure 2 is used to analyse the acoustic signatures of a honeybee colony. Training data (control) consist of sound samples of healthy and infected bee colonies. Acoustical comparison between training and live data (collected in real-time) will allow us to determine if the acoustic fingerprint of the hive correlate to the acoustic fingerprints of a hive infested with pests.

### Acoustical analysis techniques

This section describes the techniques used to determine the acoustical properties of a beehive colony.

#### *Acoustic features*

The most commonly used acoustical features in acoustical analysis applications (e.g. honeybee monitoring) are:

- Peak Frequency (PF): PF can be defined as the frequency contains the highest (most) power for a given window of audio.
- Spectral Centroid (SC): SC is also known as the mean frequency – or gravity center – of the power spectrum of a frame.
- Bandwidth (B): The bandwidth of a signal is the range of frequencies present in a signal.
- Root Variance Frequency (RVF): The RVF feature component describes the convergence of the power spectrum for a given sample.

#### *Data discrimination*

One of the most fundamental goals of the honeybee monitoring system is to be able to accurately determine the status of hive health purely through the analysis of its acoustic fingerprint. The honeybee system is designed with binary categorical discriminant analysis functionality (rather than regression analysis). The classification tools used to perform this type of analysis are:

- Principle Component Analysis (PCA): PCA is an exploratory data analysis tool used for making predictive models [15]. Commonly implemented as a form of dimensionality reduction, it involves finding the eigenvalues and eigenvectors of the covariance matrix of the meansubtracted data set [8, 11, 16].
- Support Vector Machines (SVM): SVM are machine learning models built around algorithms designed to analyse data and recognise patterns. The application of SVMs to binary classification problems have been shown to perform exceptionally even for large dimensional vectors [15].
- Linear Discriminant Analysis (LDA): LDA is used to find the linear combination of a set of features which maximises the separability between the classes [17].

## ACOUSTIC MODEL

This section describes the acoustical model used to analyse the audio signatures of beehive colonies. The aim is to be able to differentiate between a healthy or an infected beehive colony. The main processes involved are: a. control training, b. audio classification, and c. feature discrimination.

### Training

The classifier training relies on input-signals from known data-sets (various infected/healthy beehive sound samples) in order to build accurate training data-sets [18, 19]. Four approaches were taken to perform the training of the classifiers. Firstly, the feature sets were passed through a PCA algorithm which narrowed the four features down to two. This reduced feature set is then put through an LDA algorithm and as well as an SVM algorithm. In these two scenarios, the LDA and SVM classifiers use the features chosen by PCA to establish their discriminant functions. They will be referred to herein as PCA_LDA and PCA_SVM.

Table 1. Selection from the feature list of the control data-set. Label "1" represents infected samples. Label "−1" represents healthy samples

| $F_{ID}$ | PF (Hz) | SC (Hz) | B (Hz) | RVF (Hz) | Label |
|---|---|---|---|---|---|
| 1 | 1036 | 960 | 146 | 7211 | 1 |
| 2 | 998 | 971 | 168 | 7098 | 1 |
| 3 | 942 | 954 | 198 | 7239 | 1 |
| 4 | 1052 | 1025 | 125 | 4210 | -1 |
| 5 | 1114 | 992 | 157 | 4767 | -1 |
| 6 | 1028 | 1011 | 158 | 4709 | -1 |

In the other two methods, the original feature sets are passed into an LDA algorithm and an SVM algorithm without first going through PCA. This means that the results of the last two methods are purely dependant on the features chosen by LDA and PCA. The idea of this analysis was to compare which of the four overall methods had the highest accuracy of classification, and thereby establish the best method of generating classifier functions for distinguishing between infected (labelled as 1) and healthy (labelled as -1) honeybee colonies as depicted by Table 1.

### Classification

Both the LDA and SVM classification algorithms return prediction values for the given test data after applying their respective methods [20, 21]. The test data consist of a 10 second recording of either healthy or infected honeybee hive samples. The result of the prediction for both methods is an array of the same length as the number of frames observed. This means the classifying functions output either a '1' or a '1' for each frame depending on whether that frame matches the fingerprint of an infected hive or a healthy hive as seen in Table 1.

### Implementation

A PCA algorithm was developed so that the feature set extracted from the system could be tested for suitability for use in a classifier system. To do this, a number of control experiments were conducted in order to establish what types of results could be expected from different inputs. The first experiment was conducted primarily to confirm that the set of acoustic features that had been extracted for use in the system were suitable enough to allow reasonably unique and independent fingerprinting functions to be generated for sets of predominantly similar audio data.

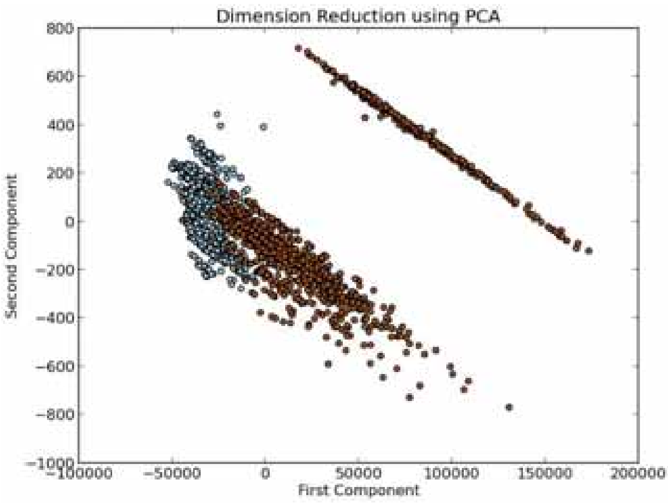


Figure 3. Separability of acoustic features using Principle Component Analysis (PCA) - healthy (blue) / infected (red) / first-component (RVF) / second-component (B)

As depicted by Figure 3, there is a clear separation (i.e. minimal overlapping between the coloured regions/points) between healthy and infected features in the acoustical domain of a beehive colony (mainly between RVF and B). This analysis confirms the suitability of PCA in determining the highest abnormalities in the acoustical spectrum components (i.e. healthy and infected feature separation). The next stage involves determining the linear relationship (classifier function) that best describes the highest degree of separability between healthy and unhealthy colonies using the principle components selected by the PCA. Once established (i.e. relationship is found), a set of test data is fed into the classifier function in order to make future predictions on the status of the bee colony involved.

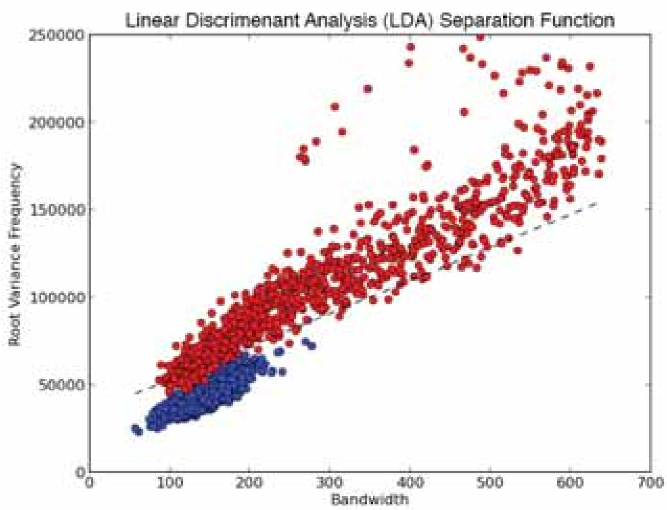


Figure 4. Classification analysis using Linear Discriminant Analysis (LDA) - healthy (blue) / infected (red)

Figure 4 shows the plot of the discriminant function generated by the LDA algorithm. The results indicate the best separation between healthy (blue) and infected (red) samples are derived from the bandwidth and root variance frequency components. The dashed line shows the threshold of classification based on the features chosen (i.e. B & RVF) to generate the corresponding plot.

In the next section, we will perform real-life experiments to test the validity of our acoustic model in detecting the presence of the Varro-mite pest in beehive colonies.

## EXPERIMENTAL ANALYSIS

One of the most important factors to keep in consideration during the analysis of these results is that the training and test data was limited to a small number of low quality samples. This essentially means that the sounds had gone through a number of re-sampling processes by the time it was recorded onto our system for analysis, and as such, they had obviously undergone a significant degradation in quality. Additionally, with limited samples available, the choice was made to use the 5 healthy honeybee hive recordings obtained from [22] as both the training and test data for the healthy hive classification. Due to the fact that only one infested hive audio sample could be obtained, the training data was generated using 5 recordings of the same sample. In this section, we reveal the acoustical patterns associated with beehive colonies infected with the Varro-mite pest. We illustrate these patterns using classification techniques widely used in the acoustical analysis domain:

**Support vector machine (SVM)**

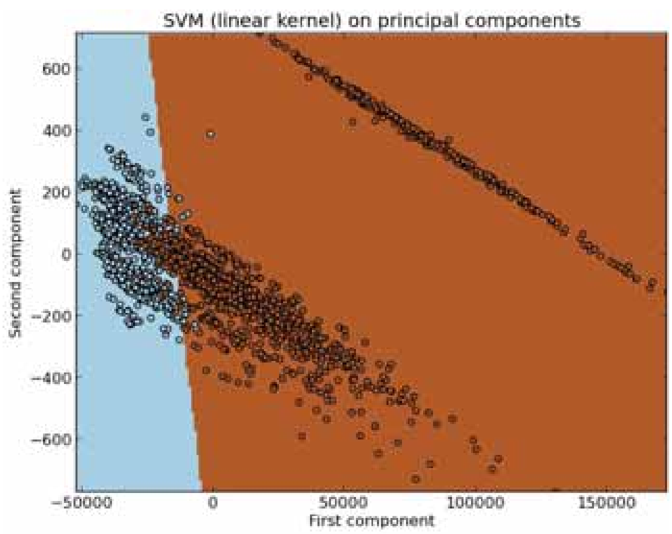


Figure 5. A depiction of healthy region (blue) vs infected region (red) using SVM - first-component (RVF) / second-component (B)

Figure-5 illustrates the regions where healthy (blue region) or infected (red region) components (PCA output data) reside within. This method of separation is generated using automatic scripts implemented on the target system (beehive node) and requires little human intervention in the final deployment.

**Linear discriminant analysis (LDA)**

Another similar approach which can be used to classify acoustic features is Linear Discriminant Analysis (LDA). This method is used to determine the highest separability function which can be generated from a given data set. The results in Figures 6,7 illustrate the use of separability functions in order to make future predictions on unknown data sets, and to classify the data as either healthy or infected.

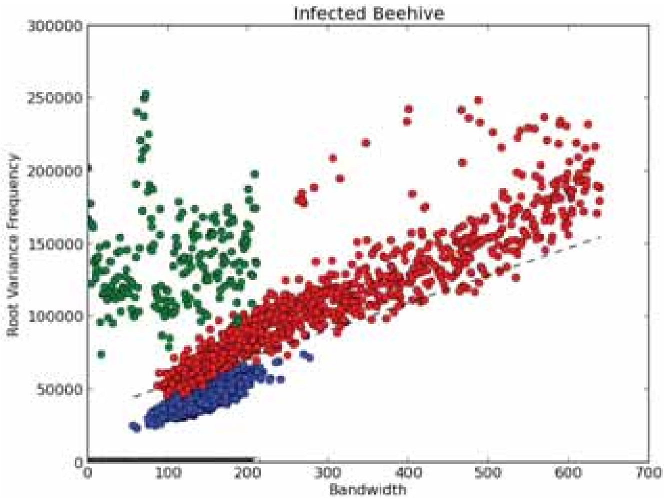


Figure 6. Varro-mite infected beehive acoustical response computed using LDA - healthy (blue) / infected (red) / test-data (green)

The blue and red features are used to differentiate between the healthy (blue) or infected (red) beehive acoustic characteristics. The green features represent a data set from an unknown beehive. By visual inspection, there is a clear indication that the data set (green) from Figure-6 resides mostly within the infected zone. Similarly in Figure 7, the green features statistically seem to indicate the beehive is healthy.

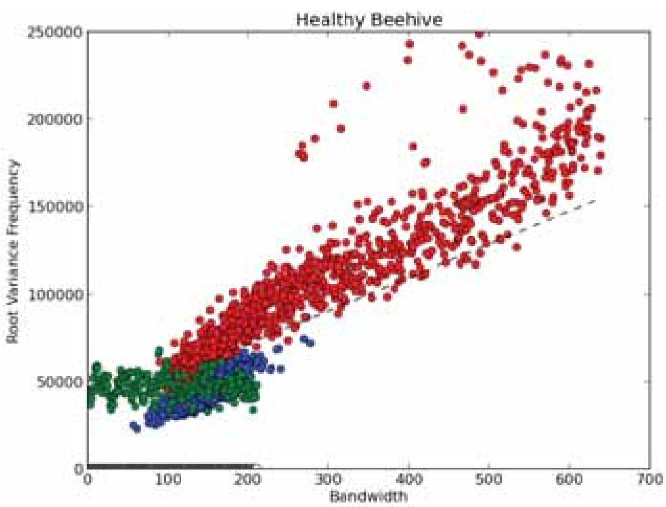


Figure 7. Healthy beehive acoustical response computed using LDA - healthy (blue) / infected (red) / test-data (green)

### Discussion

Our beehive detection system demonstrated considerable accuracy with all four features used as discriminant functions. This however came at the cost of a higher computational time. With an improved quantity and quality of training data it would most likely be found that the discriminating capabilities of the system could be significantly improved. Also, it has been shown that both LDA and SVMs are both capable of generating predictions accuracy percentages which are better than "chance" based percentages. Depending on the outcome of these future tests, it may be decided that PCA is only needed to be incorporated as far as the development stages, to help identify which features should be extracted and used in the final design.

## CONCLUSION

In this paper, we presented a system prototype for remote monitoring of beehives. This is rather a non-intrusive approach of dealing with pest infestation problem in the honeybee industry. The developed prototype is capable of capturing and analysing of acoustic samples collected from beehives. We showed how to extract features and train a classifier that can predict the infestation status of a beehive. However, further research is still required to complete the system. Several ways in which this research can be improved include:

1. Acquisition of more bee samples: One of the most important aspects for the next phase of the projects development will be to attain a more comprehensive set of control data to train the system with.

2. Expanded feature set and classifiers: There are a much greater range of features that could be used and tested for use in the honeybee system. It would therefore be a desirable process in the next stages of development that an expanded number of features be extracted from the audio samples, in order to create larger feature sets from which to compute the acoustic fingerprints.

3. Memory/Data management: Since the memory resources available to the system are limited, further methods of ensuring efficient memory and data management, as well as minimising any redundant processes should be implemented.


## ACKNOWLEDGEMENT

The authors wish to acknowledge and thank the AAS committee for providing the necessary funds to realise the work presented in this technical note. Furthermore, we would like to extend our sincere gratitude to all peer-reviewers involved in the AAS for all their efforts and valuable feedback.


## REFERENCES


[1] F. Ruttner et al., *Biogeography and taxonomy of honeybees*. Springer-Verlag, 1988. (document)

[2] Honeybee varroa vibration and ccd. [Online]. Available: http://modernsurvivalblog.com/natural-disasters/the-honeybee-varroa-vibration-and-ccd/ (document)

[3] A.-M. Klein, B. E. Vaissiere, J. H. Cane, I. Steffan-Dewenter, S. A. Cunningham, C. Kremen, and T. Tscharntke, "Importance of pollinators in changing landscapes for world crops," *Proceedings of the Royal Society B: Biological Sciences*, vol. 274, no. 1608, pp. 303–313, 2007. (document)

[4] D. Sammataro, U. Gerson, and G. Needham, "Parasitic mites of honey bees: life history, implications, and impact," *Annual review of entomology*, vol. 45, no. 1, pp. 519–548, 2000. (document)

[5] P. Boland, "A review of the national sentinel hive program," *Biosecurity Australia*, 2005. (document)

[6] T. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," *Magnetics in Japan, IEEE Translation Journal on*, vol. 2, no. 8, pp. 740–741, Aug 1987. (document)

[7] L. Fahrenholz, I. Lamprecht, and B. Schricker, "Calorimetric investigations of the different castes of honey bees, apis mellifera carnica," *Journal of Comparative Physiology B*, vol. 162, no. 2, pp. 119–130, 1992. (document)

[8] M. Hrncir, F. G. Barth, and J. Tautz, "32 vibratory and airborne-sound signals in bee communication (hymenoptera)," *Insect Sounds and Communication: Physiology, Behaviour, Ecology, and Evolution*, p. 421, 2005. (document)

[9] M. L. Winston, *The biology of the honey bee*. Harvard University Press, 1991. (document)

[10] D. Atauri Mezquida and J. Llorente Martínez, "Short communication. platform for bee-hives monitoring based on sound analysis. a perpetual warehouse for swarm apos; s daily activity," *Spanish Journal of Agricultural Research*, vol. 7, no. 4, pp. 824–828, 2009. (document)

[11] Y. Le Conte, M. Ellis, and W. Ritter, "Varroa mites and honey bee health: can varroa explain part of the colony losses?" *Apidologie*, vol. 41, no. 3, pp. 353–363, 2010. (document)

[12] G. Coley, "Beaglebone black system reference manual," 2013. [Online]. Available: http://elinux.org/Beagleboard:BeagleBoneBlack (document)

[13] Z. Alliance, "Ieee 802.15. 4, zigbee standard," *On http://www.zigbee.org*, 2009. (document)

[14] J. A. Gutierrez, E. H. Callaway, and R. L. Barrett, *Low-rate wireless personal area networks: enabling wireless sensors with IEEE* 802.15. 4. IEEE Standards Association, 2004. (document)

[15] V. N. Vapnik and V. Vapnik, *Statistical learning theory*. Wiley New York, 1998, vol. 2. (document)

[16] M. Shen, L. Cui, N. Ostiguy, and D. Cox-Foster, "Intricate transmission routes and interactions between picorna-like viruses (kashmir bee virus and sacbrood virus) with the honeybee host and the parasitic varroa mite," *Journal of General Virology*, vol. 86, no. 8, pp. 2281–2289, 2005. (document)

[17] I. Jolliffe, *Principal component analysis*. Wiley Online Library, 2005. (document)

[18] D. Albanese, R. Visintainer, S. Merler, S. Riccadonna, G. Jurman, and C. Furlanello, "mlpy: Machine learning python," *arXivpreprint arXiv:1202.6548*, 2012. (document)

[19] T. E. Oliphant, *A Guide to NumPy*. Trelgol Publishing USA, 2006, vol. 1. (document)

[20] B. Scholkopft and K.-R. Mullert, "Fisher discriminant analysis with kernels," *Neural networks for signal processing IX*, 1999. (document)

[21] I. Steinwart and A. Christmann, *Support vector machines*. Springer, 2008. (document)

[22] M. Ward, "Bee hive hums recorded to monitor insects' health," Technology correspondent, BBC News, http://www.bbc.com/news/technology-16114890. (document)