PROCESSING OF MULTI-MODAL ENVIRONMENTAL SIGNALS RECORDED FROM A 'SMART' BEEHIVE

Gordon Hunter	Faculty of Science, Engineering and Computing, Kingston University, UK
Donald Howard	Farnborough College of Technology, UK
Sylvain Gauvreau	École Nationale Supérieure d'Ingénieurs de Caen (ENSI-Caen), France
Olga Duran	Faculty of Science, Engineering and Computing, Kingston University, UK
Rosa Busquets	Faculty of Science, Engineering and Computing, Kingston University, UK

1 INTRODUCTION

Environmental factors, including air pollution, noise, and decline in biodiversity, have become issues of major concern over recent decades. Air pollution and other environmental contaminants (such as pesticides) have led to concerns relating to the health and well-being of human, animal and plant populations, whilst changes in temperature and rainfall patterns raise issues of possible rises in sea levels, coastal erosion and changes to sustainable plant and animal populations. For example, the population of bees has experienced a marked decline in many countries, which is likely to have very serious consequences for agriculture and other plant life. Bees could also be sensitive to other environmental factors such as pollution, and our recent [1, 2] and present work is a first step towards monitoring bees for obtaining information from the wider environment.

In this paper, we discuss the analysis and processing of multi-modal (sound, temperature, humidity, natural light level and air quality) signals recorded over several months from a sensor system of our own design. This sensor system was originally planned and constructed to monitor the health and well-being of honeybees in a beehive. However, we noted that same sensor system could additionally provide useful information concerning the local natural environment – for example, variations in air quality over time. We apply various signal processing methodologies both to individual signals and to the relationships between them, and discern some interesting patterns within the signals, including some relating to interactions between the environment and the activities of people living and working in the area. This work shows how a relatively simple and low-cost sensor system can be used to perform monitoring of the local environment, with a view to improving or preserving its quality, or at least limiting damage to it due to human interventions. Our sensor network (with Raspberry Pi microcomputer) cost approximately GBP £ 100 per system, or approximately GBP £ 200 per unit if audio and video recording, plus additional local data storage, were required [2]. This should make the system reasonably affordable to farmers or environmental NGOs (e.g) in developing countries, for whom commercially-produced environmental monitoring systems may be too expensive.

The remainder of this paper is structured as follows. In section 2, we describe the design, implementation and deployment of our sensor system, then outline the signal processing methodologies we have employed to analyse the output from the sensors. In section 3, we present and discuss our results, and in section 4 we present our conclusions and suggest future developments to this work.

2 MULTI-MODAL SENSOR NETWORK

2.1 Design and Implementation

As noted above, our multi-modal sensor system was originally intended to monitor the health and well-being of honeybees in a beehive [1, 2]. The system included sensors to monitor the temperature and light levels both inside and outside the hive, relative humidity and levels of a variety of gases and vapours (including hydrocarbons, CO, CO_2, NO_x) of environmental importance and smoke [3] inside the hive, the mass (weight) of the hive, plus microphones both inside the hive and a video camera monitoring the hive entrance. Four beehives, in close proximity to each other (a few metres apart), all sited at a semi-rural suburban location approximately 10 miles (16 km) from the centre of London, were each fitted with such a sensor system. Each hive had a Raspberry Pi (RPi) micro-computer with local memory storage, which a further RPi was used as a "Master" to oversee the systems and to collate the data. Data was then downloaded to a nearby PC connected to the Kingston University network. Although this was not in a densely-populated position, it was relatively close to a major trunk road (the A3) connecting London to the South coast of England, and to four sizeable London suburbs (Kingston, Putney, Richmond and Wimbledon). Details of the sensors used, and the frequencies at which they were sampled plus the volumes of data they generated, are given in Table 1 below.

Sensors used in/on each hive, with RPi	**Sampling frequency and size of daily generated data file.**
RPi camera, hive entrance video	12 seconds of 640x480 pixel mp4 video every minute, while exterior light level is above threshold level (850 MBytes/day)
USB microphone, interior sound	12 minutes of 44100Hz recording every 20 minutes, recorded 24 hours per day (6 GBytes/day)
HX711 with 4x50kg gauges, hive mass (weight)	Every minute (19 kBytes/day)
DHT22, external temperature and humidity	Every minute (45 kBytes/day)
MCP3008 ADC	-
Photoresistor, exterior light level	Every minute (23 kBytes/day)
Photoresistor, interior light level	Every minute (23 kBytes/day)
MQ-135, interior gas sensor	Every minute (23 kBytes/day)
TMP36, interior temperature	every minute (35 kBytes/day)

Table 1 : The set of sensors, etc. deployed in/on each beehive, together with their sampling rates and the corresponding volumes of data generated each day.

Figure 1 : Left image: the location of the beehives in the Kingston University apiary in South West London, on which the sensor networks are hosted. Viewed from the west, the hive entrances face south. The distance from the nearest hive to the furthest is of the order of 30 metres, with the nearest three hives being within a total distance about 15 metres of each other. The lower level box (second from front or from right) hosts the power supply required for our sensor networks. The right image shows the entrance to one hive (with bees), with the external light level sensor suspended a few centimetres away.

Data was collected 24 hours per day from the hives whilst they contained honeybee colonies over the period March 2017 to August 2018, although during the Winter of 2017-18, relatively little bee activity occurred. The data analysed in this paper were collected between May and August 2017.

2.2 Analysis Methodology

We have employed a variety of time series and signal processing approaches to analyse the signals produced by the sensors, both individually, and relationships between signals from different sensors. These are outlined below. We created a suite of tools written in Python to implement these methods.

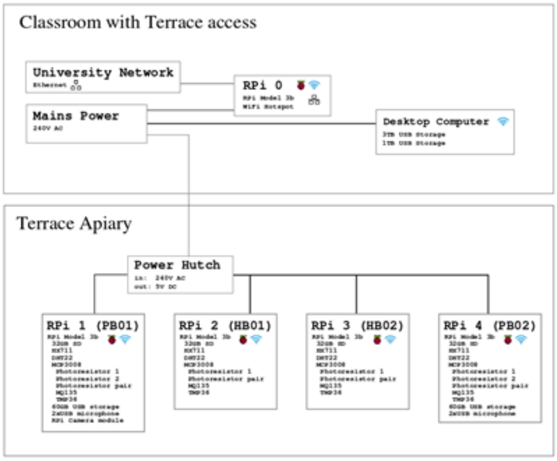

Figure 2 : Schematic representation of our sensor network, power supply, etc. The code in parentheses after the number indexing each Raspberry Pi indicates which hive that Pi is deployed in, and the sensors in that hive are then listed below (see also Table 2). The various “slave” Pi computers are connected to the “master” Pi (RPi 0) via the Pi’s onboard WiFi, and thence to the University network via the master Pi’s Ethernet port. The power is provided from a 240V 50Hz AC mains, converted to 5V DC by a transformer/rectifier in the “Power Hutch”. (Taken from [2].)

2.2.1 Autocorrelation

For a signal sampled at regular, discrete time intervals, $x[n] = x(t = n / f_s)$, where f_s is the sampling frequency, and $n = 0, 1, 2, 3, \ldots$, the autocorrelation function $\rho[x\,;\tau]$ at time lag τ samples, using a time window of N samples, is given by :

$$\boldsymbol{\rho}[\,\boldsymbol{x}\,;\,\boldsymbol{\tau}\,] = \frac{\mathbf{1}}{\boldsymbol{N}} \sum_{\boldsymbol{n}=\mathbf{0}}^{\boldsymbol{N}-\mathbf{1}} \boldsymbol{x}[\boldsymbol{n}]\,\boldsymbol{x}[\boldsymbol{n}-\boldsymbol{\tau}]$$

which can, if desired, be normalized to ensure that $-1 \leq \rho[x\,;\tau] \leq 1$. If normalized, values close to 1 indicate that, at lag τ, the signal is strongly “in phase” with its lagged form, indicating that τ samples correspond to a periodic component in the signal, whilst values close to -1 indicate that the signal is strongly “out of phase” with its lagged form at that value of τ. Autocorrelation can hence be used as a means of detecting periodic components within signals.

2.2.2 Cross-correlation

Cross-correlation is similar to autocorrelation, but the signal $x[n]$ is compared with a lagged version of another signal, $y[n]$, instead of with itself. The definition of the cross-correlation between $x[n]$ and $y[n]$ is given by :

$$\boldsymbol{\rho}_{\boldsymbol{xy}}[\,\boldsymbol{x}\,;\boldsymbol{y}\,;\,\boldsymbol{\tau}\,] = \frac{\mathbf{1}}{\boldsymbol{N}} \sum_{\boldsymbol{n}=\mathbf{0}}^{\boldsymbol{N}-\mathbf{1}} \boldsymbol{x}[\boldsymbol{n}]\,\boldsymbol{y}[\boldsymbol{n}-\boldsymbol{\tau}]$$

which, again, can be normalised if desired. In a similar manner to the autocorrelation, if normalised, values close to +1 indicate that the two signals are strongly “in phase” with a lag of τ samples, whilst a value close to -1 indicate the two signals are strongly “out of phase” at that value of the lag. Cross-correlation can hence be used to investigate whether one signal seems to be a delayed version of another, but such relationships may or may not be causal.

2.2.3 Fourier Analysis

Fourier analysis is a well-established method of investigating and processing periodic components in signals. For signals which are sampled at regular, discrete time intervals, $x[n] = x(\, t = n\,/\,f_s)$, where f_s is the sampling frequency, and n = 0, 1, 2, 3, … , the Discrete Fourier Transform (DFT) X based on N consecutive samples is defined as :

$$X[k] = \sum_{n=0}^{N-1} x[n]\, e^{\frac{-2\pi ikn}{N}}$$

for k = 0, 1, 2, 3, … , $(N - 1)$ with $i^2 = -1$, and the inverse transform is :

$$x[n] = \frac{1}{N} \sum_{k=0}^{(N-1)} X[k] e^{\frac{2\pi ikn}{N}}$$

for n = 0, 1, 2, 3, … , $(N - 1)$ [4]. For $N = 2^m$, with m a positive integer a highly efficient implementation of the DFT can be performed using Cooley & Tukey's Fast Fourier Transform (FFT) algorithm [5], which is what we have used in this study. In the above formulae, k can be regarded as a frequency in units of "per sample", such that there are f_s samples per second.

3 RESULTS AND DISCUSSION

We applied the time series and signal processing methods to the signals from each sensor. Some preliminary results from the mass, temperature and humidity monitoring, including the identification of a swarm from one hive, were reported in [2]. In this current paper, we report a more detailed analysis, including discussion of signals relating to the local environment. Many of the signals sampled minute by minute proved somewhat noisy, so these were first smoothed using an appropriate moving average (low pass) filter.

3.1 SOUND SIGNALS

The sound signals recorded in the hives during June and July 2017 were analysed using the FFT methodology. The spectral envelope obtained is shown in Figure 3a below. This can be compared with the corresponding spectral envelope (Fig. 3b) obtained from a beehive located elsewhere, for which the sound data is publically available on the internet (https://www.youtube.com/watch?v=vKEUJaQcPrg).

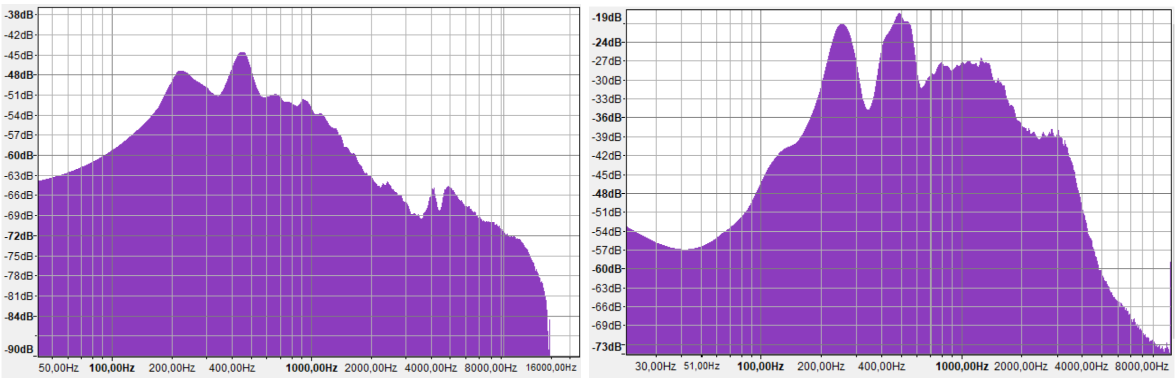

Figure 3a (Left-hand image) the frequency spectral envelope of the sounds recorded from one of our beehives during June and July 2018; 3b (Right-hand image) the corresponding frequency spectral envelope of the sounds recorded from a beehive in the sound file publicly available on the internet at https://www.youtube.com/watch?v=vKEUJaQcPrg. Note the scales are not the same on the two graphs.

It can been seen that, whilst the two spectral envelopes are not identical, there are considerable similarities between them. Both show major peaks around 220-260 Hz, a dip in amplitude around 350 Hz, followed by another peak between 450 and 500 Hz. A further dip in amplitude occurs in both envelopes around 550 to 650 Hz, followed by a rise to a rough "plateau" between about 700 and 1500 Hz. As the frequency rises further, there is a steady decline in amplitude in both cases up to about 3000 Hz, a second "plateau" between 3000 and 4000 Hz, followed by another decline in amplitude up to around 10 000 Hz. The first peak around 220-260Hz would appear to be consistent with both the "warble" (225 – 285 Hz) and the "worker piping" bands noted in Howard et al [6] and previously described by Boys [7] and Seeley & Tautz [8]. There seems little evidence of the "Queen piping" described by Kirchner [9] as occurring at a fundamental frequency of around 300Hz with harmonics at multiples thereof. The second peak around 450 to 500 Hz is consistent with the fundamental of the hive background sound frequency of *Apis mellifera carnica* described by Favre [10]. The sound data recorded will be further investigated to study if distinctive acoustic profiles could be detected in the run-up to the swarming events observed (see section 3.5 below).

3.2 Temperature and Relative Humidity

The previous study [2] showed a largely periodic daily variation of temperature and relative humidity over time – confirming the well-known relationship that, for approximately constant absolute humidity, high temperature coincides with low relative humidity, and vice versa [11]. This is shown for one of our hives in Figure 4 below.

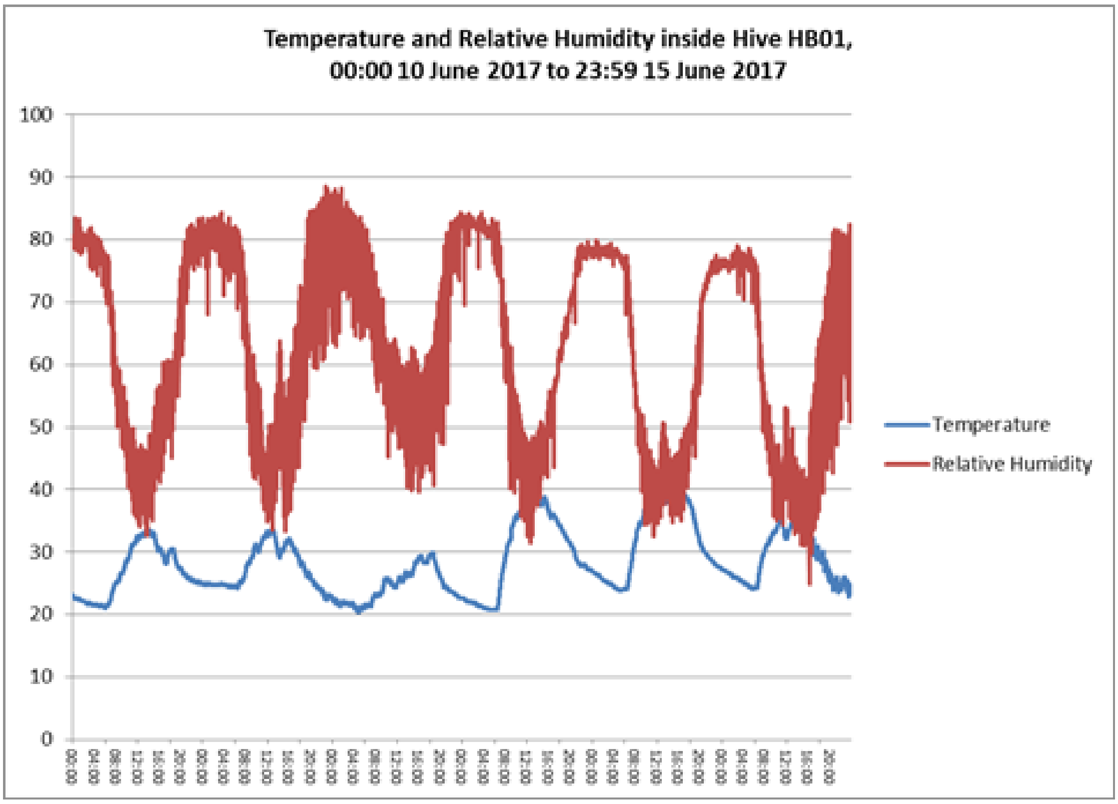

Figure 4. The variation of temperature (in ºC) and relative humidity (% of saturation) inside hive HB01 over the period 10 June to 15 June 2017, measured by our bespoke sensor system. A daily periodic component to both signals is clear.

Previous authors [12, 13, 14] have noted that the temperature in a hive tends to rise in the period immediately before the occurrence of a swarm. This can be observed in our data, and is discussed further in section 3.5 below.

3.3 External Light Level and Temperature

Investigation of the variation of external light level and external temperature over time not surprisingly showed a daily periodic component to both signals (see Figure 5), with higher temperatures on the whole corresponding to higher light levels. Since the data was collected in summer (year 2017), and during a period of mostly warm, fine weather, the daily duration of strong external light level was considerably longer than that of near darkness. Scrutiny of the cross-correlation between these signals confirmed this relationship between temperature and light level, with periodicity 1440 minutes (24 hours), but with the temperature signal lagging behind the light level by 145 minutes (see Figure 6). This confirms the traditional view that, even in Summer, bright mornings can be cool, whilst daily temperatures tend to peak in mid afternoon rather than at midday.

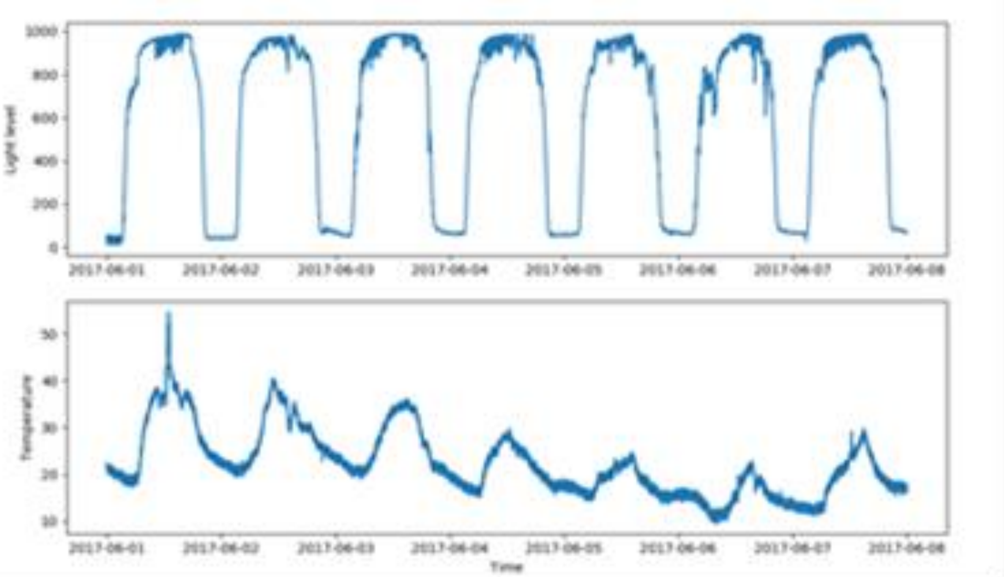

Figure 5 : Variation of external light level (upper graph) and external temperature (lower graph) over time for 7 days in June 2017. A clear daily periodic component can be observed in both signals.

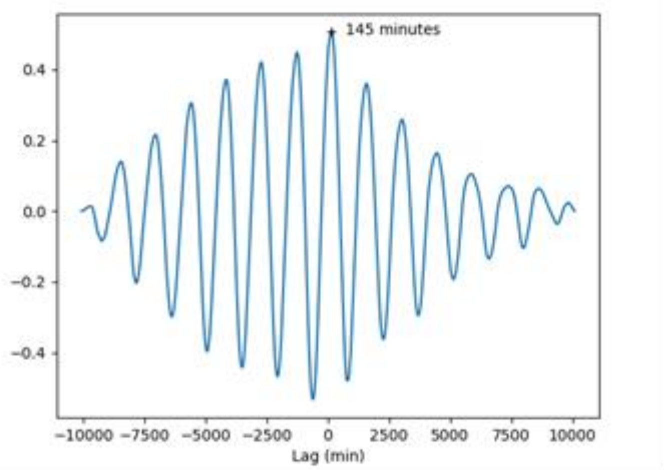

Figure 6 : Cross-correlation between external light level and temperature against time lag in minutes. The temperature lags behind the light level by approximately 145 minutes, corresponding to a little less than two and a half hours – indicating that the peak in daily temperature tends to occur in the early-mid afternoon and the daily low tends to occur some hours after midnight. The time gap between peaks is exactly 1440 minutes, or precisely 24 hours.

3.4 Air Quality over Time

The signal from the gas sensor (MQ-135) showed a clear daily periodic component (see Figure 7), and its Fourier analysis showed clear spikes at frequencies of 0.0007 min^{-1} and 0.0001 min^{-1}, corresponding to a daily (24 hour) and weekly (7 day) component respectively. Whilst part of the daily variation in gas level could be attributed to temperature variation, the resistance of the MCQ-135 will only vary slightly for temperatures between 10°C and 30°C (the approximate temperature range over the period during which this data was collected) [3]. Thus, fluctuations in air pollution may be detected here, and would explain the weekly (7 day) periodic component. Although the hives (and hence the sensors) are located in a semi-rural suburb of South-West London, they are relatively close to several major suburban population centres (Kingston, Putney, Richmond and Wimbledon) and a major trunk road (the A3) connecting London to the South coast of England. Therefore, traffic levels will tend to be heavy – and hence air quality poor – around the morning and early evening “rush hours” each working day, but this effect will be less marked at weekends (hence the weekly component).

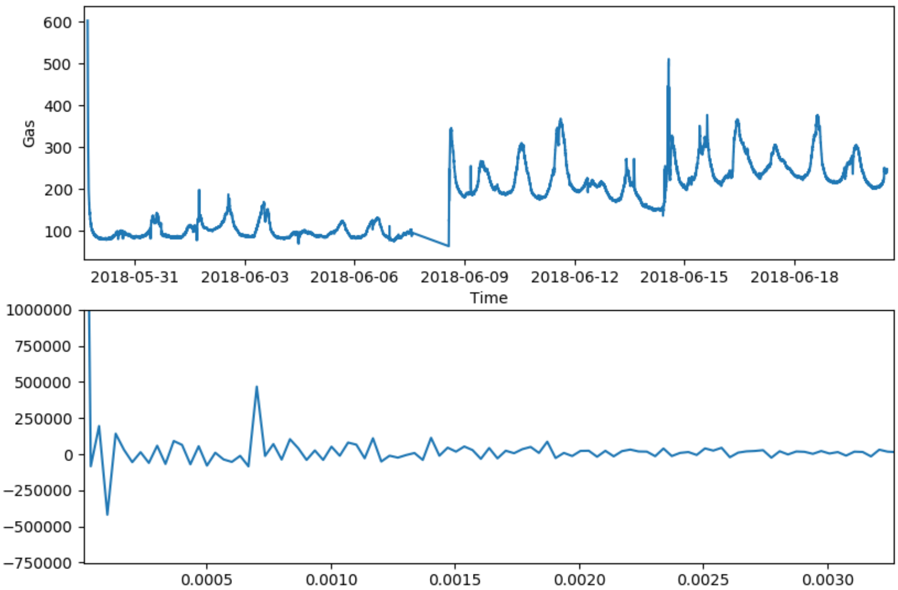

Figure 7 : Upper graph : Variation of gas sensor signal over time (over several days), with clear evidence of a daily periodic component. The FFT of this signal (lower graph) confirms this periodic element (peak at frequency of approximately 0.0007 min^{-1}) and also the presence of a weekly periodic element (peak at frequency of approximately 0.0001 min^{-1}).

3.5 Hive Mass and Temperature over Time

Detailed investigation of the mass of one particular hive over time showed clear indications of first one quite large swarm (13 June 2017) followed by a smaller swarm a few days later (16 June 2017). Such secondary swarms, sometimes known as “casts” from the same hive are not uncommon, according to experienced beekeepers. Study of the (smoothed) hive temperature over time indicated that the first swarm had occurred from one hive immediately following a rise in hive daily averaged temperature (to a value above about 29^{o}C), in agreement with previous findings of [12, 13, 14] – see Figure 8. Those two swarms from the same hive were the only ones detected from any of the four hives over the course of our period of data collection. Photographs showing the “beard” of bees from the first swarm (which had temporarily settled in a nearby apple tree) are displayed in Figure 9. Studying the variation in hive mass over time over the course of a “regular” day – one where no inspection of the hive occurred, and no swarm was detected – showed a drop in hive mass by about

300 to 500g in the morning, as worker bees left the hive to forage. These bees return later in the day, laden with pollen and nectar, so that the hive mass gradually increases as the day progresses, to the extent that the mass of the hive will typically be around 300 to 400g greater by late evening than it was just before dawn. The hive mass tends to remain relatively stable overnight, possibly showing a small decrease due to the bees feeding and metabolising food reserves.

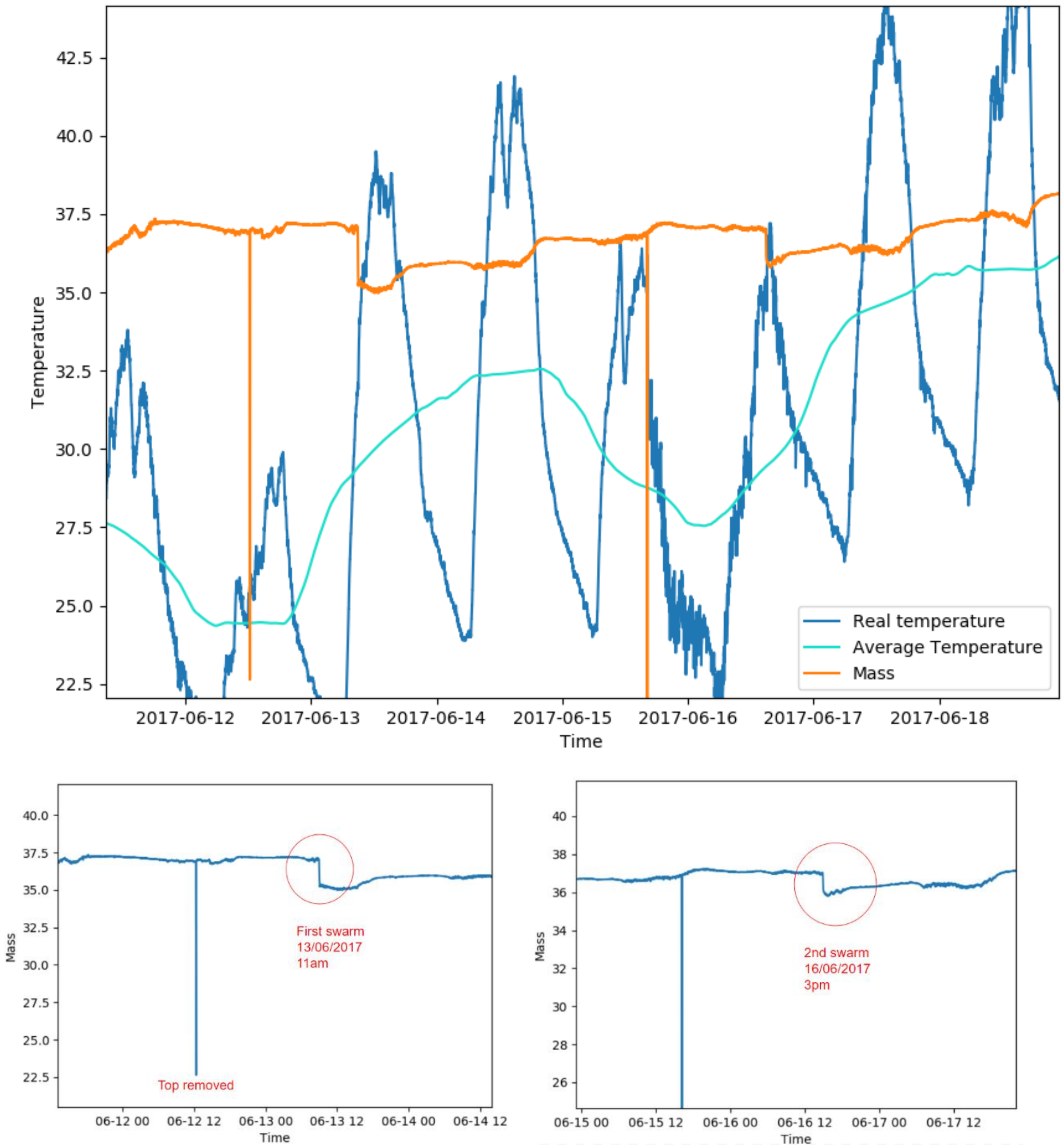

Figure 8. Upper graph : Hive mass, unsmoothed hive temperature and daily average (smoothed) temperature over time for one particular hive over the period 12 June 2017 to 18 June 2017. It can be observed that the first swarm (with a quite sudden mass loss of about 1.9 kg) occurred just after the daily averaged hive temperature first rose above 29°C on 13 June 2017, with a smaller second “cast” swarm (a loss of about 1.2 kg) occurring on 16 June 2017, after the average temperature again rose above 29°C following a brief drop below 28°C. Lower graphs : close-ups of the changes in hive mass around the time of each swarm. The very large but very short-term drops in hive mass (of around 15 kg) correspond to the “roof” of the hive (together with a 10 kg weight used to keep it firmly in place) being removed for inspections of the hive.

(a) (b) (c)

Figure 9. (a) a view of the “beard” of bees from the swarm around 2pm on 13 June 2017, after it had settled in an apple tree by a wall, around 50 metres from the hive it had come from, (b) and (c) close-ups of the “beard”, showing individual bees.

4 CONCLUSIONS AND FUTURE WORK

We have shown that, despite being originally designed for a different purpose (namely monitoring the health and well-being of honeybees), our relatively low-cost sensor, data collection, storage and analysis system [1, 2] could also be used to monitor signals important to the local environment – notably in relation to air quality - and investigate relations between such signals, and patterns they display over time. We intend to continue and extend this work, both in the environmental and bee-related contexts (e.g. to study the effects of pollution and pesticides on bee well-being [15, 16]) , and investigate a wider range of signal modalities (e.g. including factors such as rainfall and wind, for which we already have some limited data) and widen the analysis methodology to include intelligent pattern recognition techniques. Evidence suggests that, to date, African honeybees have been less susceptible to the problems affecting their European and North American cousins, and we are currently in discussions with a University and an NGO in West Africa with a view to investigating this in more detail.

5 ACKNOWLEDGEMENTS

This work has been supported by the Industry Strategy Fund (ISCF P1178-180) of Kingston University (KU) London. Both Don Howard and Sylvain Gauvreau are grateful to the KU ISCF fund for financial support over the period of this work, and Sylvain Gauvreau is also grateful to the ERASMUS educational exchange scheme of the EU for giving his the opportunity to participate in this project. Gordon Hunter would like to thank Simon Sylvester of the Twickenham & District Beekeepers’ Association for some helpful comments regarding the life cycle of bees and bee colonies. We would all like to thank the Biodiversity team at Kingston University to allow us to access and use the University beehives, and the School of Education. Kingston University, for allowing us to use some indoor space close to the hives to host some of our equipment.

6 REFERENCES

1. G. Hunter, D. Howard and O. Duran (2016) “A User-Centred Design Approach Towards Developing an Intelligent Living Environment for Honeybees”, Proceedings of British Computer Society Human Computer Interaction Conference, July 2016, Bournemouth, U.K., British Computer Society
2. D. Howard, O. Duran and G. Hunter (2018) “A low-cost multi-modal sensor network for the monitoring of honeybee colonies/hives”, Proceedings of 2nd International Workshop on Intelligent Systems for Agriculture Production and Environment Protection (ISAPEP'18), Rome, Italy, June 2016

3. Olimex Ltd., Technical Data sheet for MQ-135 Gas Sensor https://www.olimex.com/Products/Components/Sensors/SNS-MQ135/resources/SNS-MQ135.pdf [Last accessed 31 March 2019]
4. P.A. Lynn and W. Fuerst (1998) "Introductory Digital Signal Processing with Computer Applications" (2nd Edition) John Wiley and Sons, Chichester, U.K.
5. J.W. Cooley & J.W. Tukey (1965) "An Algorithm for the Machine Calculation of Complex Fourier Series", *Mathematics of Computation,* Vol. 19, pp 297-301
6. D. Howard, O. Duran, G. Hunter and S. Stebel (2013) "Signal Processing the Acoustics of Honeybees (*Apis Mellifera*) to Identify the 'Queenless' State in Hives", Proceedings of the Institute of Acoustics, Vol. 35, Part 1, pp 290 – 297
7. R. Boys (1999) "Listen to the Bees", available on-line at https://beedata.com.mirror.hiveeyes.org/data2/listen/listenbees.htm [Last accessed 30 March 2019]
8. T.D. Seeley and J. Tautz (2001) "Worker Piping in Honeybee Swarms and its Role in Preparing for Liftoff", Journal of Comparative Physiology A, Vol. 187, pp 667 - 676
9. W.H. Kirchner (1993) "Acoustical Communication in Honeybees", Apidologie, Vol. 24, pp 297-307
10. D. Favre (2011) "Mobile Phone-Induced Honeybee Worker Piping", Apidologie, Vol. 42, pp 270 – 279
11. A.L. Buck (1981) "New Equations for Computing Vapor Pressure and Enhancement Factor", Journal of Applied Meteorology, 20, pp 1527 – 1532
12. S. Ferrari, M. Silva, M. Guarino & D. Berckmans (2008) "Monitoring of swarming sounds in bee hives for early detection of the swarming period", Computers and Electronics in Agriculture, Vol. 64, pp 72-77
13. D.S. Kridi et al (2016) "Applications of Wireless Sensor Networks for Beehive Monitoring and In-Hive Thermal Patterns Detection", Computers and Electronics in Agriculture, Vol. 127, pp 221 – 235
14. A. Zacepins et al (2016) "Remote Detection of the Swarming of Honey Bee Colonies by Single-Point Temperature Monitoring", Biosystems Engineering, Vol. 148, pp 76 – 80
15. P.R. Whitehorn, S. O'Connor, F.L. Wackers, D. Goulson (2012) "Neonicotinoid Pesticide Reduces Bumble Bee Colony Growth and Queen Production", Science, Vol. 336 (6079), pp 351-352
16. A.N. Arce, A.R. Rodriges, J.Yu, T.J. Colgan, Y. Wurm, R.J. Gill (2018) "Foraging Bumblebees Acquire Preference for Neonicotinoid-Treated Food with Prolonged Exposure", Proceedings of the Royal Society B: Biological Sciences, Vol. 285 (1885) 20180655 http://dx.doi.org/10.1098/rspb.2018.0655 [Last accessed 31 March 2019]