Sensors & Transducers, Vol. 250, Issue 3, March 2021, pp. 39-43

Sensors & Transducers

Published by IFSA Publishing, S. L., 2021
http://www.sensorsportal.com

First Results of the BeE-Nose on Mid-Term Duration Hive Air Monitoring for Varroa Infestation Level Estimation

Andreas KÖNIG

Institute of Integrated Sensor Systems, TU Kaiserslautern,
Erwin-Schrödinger-Str. 12, Kaiserslautern, 67663, Germany
Tel.: +496312053696, fax: +496312053889
E-mail: koenig@eit.uni-kl.de

Received: 16 February 2021 /Accepted: 22 March 2021 /Published: 31 March 2021

Abstract: Bees are recognized as an indispensable link in the human food chain and general ecological system. Numerous threats, from pesticides to parasites, endanger bees and frequently lead to hive collapse. The varroa destructor mite is a key threat to bee keeping and the monitoring of hive infestation level is of major concern for effective treatment. Sensors and automation, e.g., as in condition-monitoring and Industry 4.0, with machine learning offer help. In numerous activities a rich variety of sensors have been applied to apiary/hive instrumentation and bee monitoring. Quite recent activities try to extract estimates of varroa infestation level by hive air analysis based on gas sensing and gas sensor systems. In our work in the IndusBee4.0 project [8, 11], an hive-integrated, compact autonomous gas sensing system for varroa infestation level estimation based on low-cost highly integrated gas sensors was conceived and applied. This paper adds to [11] with the first results of a mid-term duration investigation from July to September 2020 until formic acid treatment. For the regarded hive more than 79 % of detection probability based on the SGP30 gas sensor readings have been achieved.

Keywords: Multi-modal bee health monitoring, Varroa infestation level estimation, Gas sensing, Machine learning, Apiary intelligence.

1. Introduction

Major issues from environmental pollution to invasive species are threatening our ecological system and the human food supply. Insects, and honey bees in particular, play a decisive role, e.g., for pollination. The varroa mite parasite is a major threat to bee keeping and the cause of many bee colony losses. The monitoring of the varroa infestation level is one important task of conventionally operating bee keepers. Though there is a community practicing treatment free bee keeping [1], the majority of bee keepers follows standard treatment practice, e.g., by formic acid, which needs to know the right time to start treatment based on the hive infestation level.

Sensors and automation approaches, e.g., condition-monitoring, and Industry 4.0, can both alleviate hive keeping and also make it much more effective (see, e.g., [8, 11]). In addition to common monitoring modalities, like temperature, moisture, visual/IR images, or sound patterns, the hive air attracted attention, both due to therapeutic interest for patients with respiratory ailments [3], and for extended hive condition monitoring based on simple and cost effective gas sensors, in particular MOX sensors [2], e.g., for CO_2 or VOC concentration measurement. The aspect of estimating the varroa infestation level by means of such gas sensors combined with pattern recognition and machine learning techniques has attracted numerous researchers. The approach shows potential to be generalized to other parasites or illness, e.g., the emerging pest of SHB (Small Hive Beetle), or foulbrood.

In [4-7] interesting activity based on a multi-gas sensor system of Figaro gas sensors can be found. A principal and useful relation between the gas sensor readings and varroa infestation level was reported. The ground truth in [4] was obtained by employing the flotation method to the same hive. The external measurement setup, however, reportedly implied challenges, e.g., with regard to dew point issues and context dependence due to day time of measurement [6], and monitoring times were rather limited, i.e., continuous monitoring for a whole bee season or development cycle of a hive was not reported yet.

In our IndusBee4.0 project, started end 2017, in-hive cost-effective integrated sensor systems and machine learning based data analysis, continuous state or condition monitoring, and automated decision making is pursued [8, 11] for longer period, preferably covering the decisive part of the bee season from the onset of breeding activity to the first formic acid treatment [11] and finding effective cues in the monitoring data to estimate the desired infestation level. In Section 2, the ground truth obtaining will be outlined, Section 3 describes the details of the data acquisition, and Section 4 gives the first results obtained from July to September.

2. Varroa Monitoring Options

There are several standard methods available for conventional varroa infestation level assessment for the required ground truth. They all have in common, that they imply substantial effort for the bee keeper and deliver results only at larger time steps. The analysis of hive debris including mites, dropping from the hive bottom and collected on a slider or tray, is most common. Usually, three days are expended until a manual, or more recently (semi) automated vision-based analysis, of the debris for the number of varroa mites can be conducted. The hive infestation level can be estimated from this count [8]. Another common approach, also denoted as flotation method, used in [4], extracts a bee sample from the hive and effectively drowns them to separate bees and varroa to count the mites. The powder sugar and the CO_2-based sedation method are two alternative, more bee-friendly, variants. Sample adequateness will probably depend on the location of extraction in the hive. A more recent principle approach tries to scrutinize in and out going bees at the flight hole for varroa mites clinging to them, e.g., [8, 9].

Here, the conventional established tray analysis was employed and more frequently repeated for the instrumented hives.

3. Data Acquisition by BeE-Nose

For the continued and unobtrusive measurements in individual bee colonies and a complete apiary, a modular monitoring system has been conceived, that networks the instrumented hives and combines them with hive keeper assistance systems, as outlined in [8, 11]. The block diagram of the IndusBee4.0 Apiary Monitoring system is shown in [11, Fig. 1]. The basic building block is the SmartComb in-hive autonomous measurement system [11, Fig. 4]. In the monitored hive, the SmartComb is located in the middle of the second super of a three super hive, i.e., in the very center. The monitoring can be controlled and read-out via WLAN. The regarded hive has been instrumented also with hive and honey room scales in addition to temperature, moisture, sound patterns, and gas sensors readings. The focus of this report here is on the SGP30 gas sensor, with fortunate properties of baseline determination [12], and its monitoring results. At a later point, context from the other sensors will be included in the analysis.

In the following experiments, the SGP30 gas sensor readings of time intervals from July to September 2020, due to issues in employed libraries and several required restarts, have been merged.

4. Experiments and First Results

For the first hive monitoring investigations by BeE-Nose, the data was acquired in the interval from 9^{th} July to 11^{th} September with about 6 readings per minute.

The development of the varroa population was monitored more frequently than usual by established tray analysis as outlined in Section 2. The obtained varroa count development confirmed the validity of the chosen observation time. The regarded instrumented hive had released a strong swarm end of April.

Fig. 1 shows the SGP30 gas sensor recordings over the July to September campaign. The two raw data outputs for ethanol and hydrogen as well as the computed e-CO_2 and TVOC outputs are displayed. This is complemented by the scaled varroa ground truth, which was obtained concurrently to gas measurements by following the standard procedure of counting the daily average drop of varroa on the tray, here 0, 2.5, 8.6, 14, and scaling this to an estimate of the actual population in the hive by the standard factor 150. For the sake of visibility, in the Fig. 1 the floored resulting values were additionally scaled by a factor of 10. As can be seen, in this hive the varroa population ramps up during the campaign until the critical level for required treatment, commonly more than 10 per day, is reached in September.

Fig. 2 shows a scatter plot from the immediate SGP30 sensor data, i.e., ethanol and hydrogen outputs, picking in a hold-out approach every thousandth sample from the available 1003919 of the baseline corrected readings as training set of a size of 1003 samples, in this first step without further preprocessing or feature extraction. A corresponding test set of the same size has been picked from the same sensor readings with a temporal displacement of

500 samples to the training data. The four varroa level estimations obtained from counting and related varroa population size estimate as displayed in Fig. 1 have been translated to four classes from '1, No Varroa', '2, Low Varroa', '3, Mid Varroa', to '4, Treatment !'. However, the boundaries between classes 1, 2 and 3 are rather arbitrary, but make classification definitely much harder and error prone. Thus, with regard to the actual binary issue *ok* or *treatment needed* alert, this was simplified to just two classes '1, Ok', here summarizing former classes 1 to 3 and former class 4 as '2, Treatment !' in the remaining conducted classification experiments.

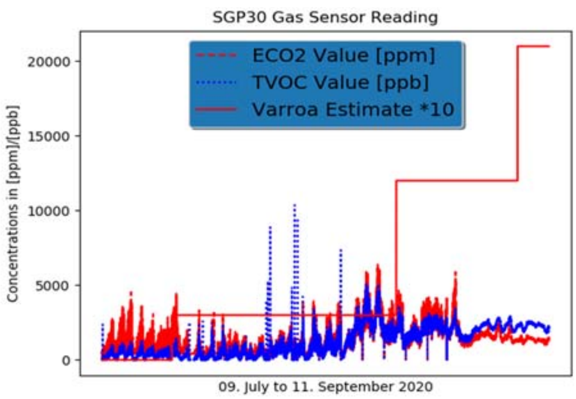

Fig. 1. SGP30 gas sensor recordings over the July to September campaign for the computed e-CO_2 and TVOC values complemented by the scaled varroa ground truth obtained from tray analysis.

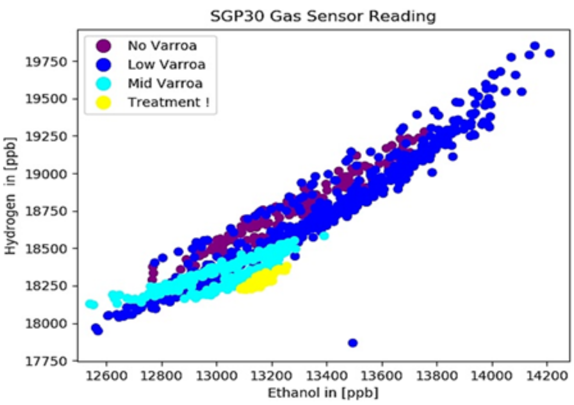

Fig. 2. Scatter plot of the training subset of SGP30 ethanol and hydrogen gas sensor readings over the July to September campaign.

The scatter plot in Fig. 2 shows for four assumed classes a mediocre support of the hypo-thesis, that the SGP30 readings give a cue on varroa infestation level. Unfortunately, even reducing to a two class problem or alert detector for treatment, inter class distance is small and a substantial overlap in the currently acquired sensor data can be observed. This advocates both the use of additional sensors as well as efficient feature computation [13-15].

Fig. 3 correspondingly shows a scatter plot for the e-CO_2 and TVOC sensor readings, training set, with a comparable outcome as observed in Fig. 2.

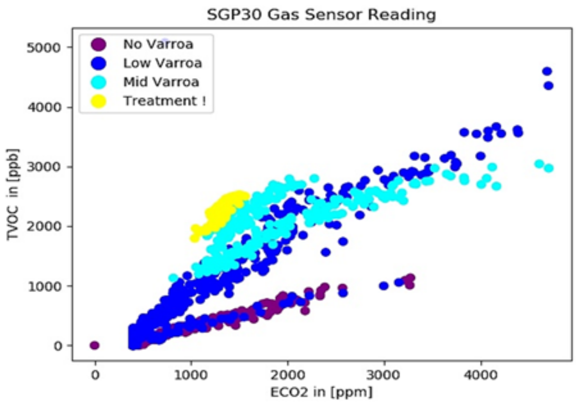

Fig. 3. Scatter plot of the training subset of SGP30 e-CO_2 and TVOC gas sensor readings over the July to September campaign.

Classification experiments were conducted for both SGP30 ethanol/hydrogen and e-CO_2/TVOC data sets for several common and relevant classifiers and the results are summarized in Table 1.

Table 1. First classification results for the SGP30 data and a voting kNN-classifier (k = 3), RNN, PNN, and SVMs.

Classifier	No. of Classes	Resubstitution	Generali-zation
SGP30 **ethanol and hydrogen** output			
kNN	4	89.0 %	79.8 %
kNN	2	98.5 %	97.4 %
RNN	2	100 %	96.91 %
PNN	2	100 %	97.41 %
SVM	2	93.15 %	93.15 %
NLSVM	2	98.00 %	97.91 %
SGP30 **e-CO_2 and TVOC** output			
kNN	4	88.33 %	82.85 %
kNN	2	99.1 %	98.6 %
RNN	2	100 %	97.81 %
PNN	2	100 %	97.61 %
SVM	2	98.4 %	98.0 %
NLSVM	2	98.9 %	98.6 %
SGP30 **ethanol, hydrogen, e-CO_2 & TVOC output**			
kNN	2	99.1 %	98.7 %

The classifications were predominantly carried out based on the sklearn package of Python, RNN and PNN were employed in the proprietary QuickCog tool. RNN does not require parameters, PNN kernel width was set to 2 in resubstitution and 4 and 7 in generalization for the reported results. For the kNN, k=3 was determined and employed. The linear SVM was employed with C = 1.0, the nonlinear SVM with rbf-kernel, C = 1.01, and gamma = $2.6 \cdot 10^{-5}$.

The kNN runs for the four class case give the worst results of 79.8 % and 82.85 % in generaliza-tion, which is not unexpected and due to the granularity of currently determined varroa counts and resulting class thresholds, and the overlap at class boundaries.

Reducing to two classes, i.e., an alert detector for *treatment needed*, however, returns more than 97 % in all cases, which is encouraging. Table 2 and Table 3 show the confusion matrices of the kNN classifications of SGP30 ethanol and hydrogen output for four and two classes of Table 1, respectively. The confusion between classes 3 and 4, and 1 and 2, respectively, gives only 82.6 % recognition (True Positives, TP) of the class 2 *treatment needed*.

Table 2. Confusion matrix for the SGP30 data and a voting kNN-classifier with k = 3 and four classes.

SGP30 **ethanol and hydrogen** output				
	1	**2**	**3**	**4**
1	118	38	1	0
2	45	403	56	0
3	0	36	223	14
4	0	0	12	57

Table 3. Confusion matrix for the SGP30 data and a voting kNN-classifier with k = 3 and two classes.

SGP30 **ethanol and hydrogen** output		
	1	**2**
1	920	14
2	12	57

This is repeated in Table 1 for e-CO_2 and TVOC output and the corresponding confusion matrices are given in Table 4 and Table 5.

Table 4. Confusion matrix for the SGP30 data and a voting kNN-classifier with k = 3 and four classes.

SGP30 **e-CO_2 and TVOC** output				
	1	**2**	**3**	**4**
1	141	16	0	0
2	62	394	47	1
3	0	33	232	8
4	0	0	5	64

Table 5. Confusion matrix for the SGP30 data and a voting kNN-classifier with k = 3 and two classes.

SGP30 **e-CO_2 and TVOC** output		
	1	**2**
1	925	9
2	5	64

The confusion between class 3 and 4, and 1 and 2, respectively, now gives 92.75 % recognition (TP) of the class 2 *treatment needed* for e-CO_2 and TVOC output. Finally, in the last row of Table 1, one more experiment just for kNN with k=3 can be observed, where the four sensor outputs have been merged, which left resubstitution unchanged and gave a minor increase in generalization from 98.6 % to 98.7 %.

5. Discussion

The employed SGP30 gas sensor and the related measurements indicate a moderate correlation with the manually determined onset of varroa infestation in the bee hive. However, the SGP30 sensor readings cannot realistically be expected to give a direct indication of the present varroa population, the observed correlation or dependency can be accredited to indirect indication of changes in bee hive climate due to the imposed influence and stress of the growing mite population on the bee colony. This clearly advocates to add additional sensors, e.g., as inspired by [4-7], but in hive-integrated embodiment. The popular Bosch SENSORTEC BME680, which even allows temperature modulation, has already been included [11] and will be incorporated in the future analysis. This will in general also demand for effective domain-specific feature computation as, e.g., presented in [13-15], which will also give credit to the temporal nature of the data. For instance, analyzing the temporal sequence of occurrence of current misclassifications, it became obvious, that these were predominantly singular or temporarily isolated events. As the treatment decision allows to accumulate decisions over a longer intervall, e.g., several hours, singularities could be masked by a voting approach for a more robust alert.

Further, the number of manual inspections and related varroa counts are still sparse or coarse and would ask for a more fine grained, i.e., more frequent, determination in future measurement campaigns. Also, in the current investigation, samples have been adopted from any time of the 24 h observation cycle in the regarded campaign of three months. In the light of [6], the context of time of day and colony activity should be definitely included in the choice of classification data for improved detection and alert generation capability.

6. Conclusions

IoT, Machine Learning, and Artificial Intelligence massively move in the agro-tech domain, striving for fully automated farming solutions [10]. The importance of bees in agriculture and the overall ecosystem as well as the stringent need for technological support and alleviation of bee keeping is undisputed [4-7, 8, 9, 11]. In this work, a step towards continuous varroa infestation level determination by a low-cost, small, and unobtrusive in-hive monitoring system with gas sensor extension [11] has been achieved. First results obtained by this BeE-Nose from a mid-term observation time in a hold-out approach harmonize with those reported in [4-7], which were based on comparably macroscopic hive-external equipment, shorter observation time, but richer sensor palette.

Though *ad hoc* use of this first SGP30 data already delivered a moderate leverage for the infestation level estimation and treatment alert generation, the role of context, robustness, and stability of the results, varroa threshold settings, as well as generalization issues to different hives have to be thoroughly investigated in the next steps.

In future work, more sophisticated feature computation [13-15] and multiple gas sensor employment, starting with adding BME680 information, will be regarded. Exploitation of further cues from acoustical data [8, 11] and merging those with the gas sensor information for increase of robust varroa infestation level estimation and treatment alert generation will be investigated (see [11], Fig. 10).

Potential generalization to other pests and illnesses as the SHB or variations of foulbrood open further opportunities.

References

[1]. C. Hudson, S. Hudson, Notes on Treatment Free Beekeeping, January 2018 (https://beemonitor.files.wordpress.com/2018/04/notes-on-treatment-free-beekeeping-jan-2018.pdf)

[2]. C. Jaeschke, O. Gonzalez, M. Padilla, K. Richardson, J. Glöckler, J. Mitrovics, B. Mizaikoff, A Novel Modular System for Breath Analysis Using Temperature Modulated MOX Sensors, *MDPI Proceedings*, Vol. 49, Issue 14(1), 2019, pp. 1-3.

[3]. K. Speer., Stockluftanalyse, *Technische Universität Dresden*, accessed March 2020, https://www.chm.tu-dresden.de/lc3/dateien/stockluft.pdf

[4]. A. Szczurek, M. Maciejewska, B. Bąk, J. Wilk, J. Wilde, M, Siuda, Detecting varroosis using a gas sensor system as a way to face the environmental threat, *Science of the Total Environment*, Vol. 722, 2020, pp. 137866.

[5]. B. Bąk, J. Wilk, P. Artiemjew, J. Wilde, M. Siuda, Diagnosis of Varroosis Based on Bee Brood Samples Testing with Use of Semiconductor Gas Sensors, *Sensors*, Vol. 20, Issue 14, 2020.

[6]. Szczurek A., Maciejewska M., Zajiczek Ż., Bąk, B., Wilk J., Wilde J.**,** Siuda M.**,** The Effectiveness of Varroa destructor Infestation Classification Using an E-Nose Depending on the Time of Day, *Sensors*, Vol. 20, Issue 9, 2020.

[7]. A. Szczurek, M. Maciejewska, B. Bąk, J. Wilk, J. Wilde, M. Siuda, Detection Level of Honeybee Desease: Varroosis Using a Gas Sensor Array, in *Proceedings of the 5th International Conference on Sensors Engineering and Electronics Instrumentation Advances (SEIA'2019)*, Tenerife (Canary Islands), Spain, 25-27 September 2019, pp. 255-256.

[8]. A. König, IndusBee4.0 – Integrated Intelligent Sensory Systems for Advanced Bee Hive Instrumentation and Hive Keeper's Assistance Systems, *Sensors & Transducers Journal*, Vol. 237, Issue 9-10, 2019, pp. 109-121.

[9]. L. Chazette, M. Becker, H. Szczerbicka, Basic algorithms for bee hive monitoring and laser-based mite control, in *Proceedings of IEEE Symposium Series on Computational Intelligence (SSCI)*, Athens, 2016, pp. 1-8.

[10]. L. Rembert, How AI and the IoT are improving farming sustainability (https://www.embedded.com/how-ai-and-the-iot-are-improving-farming-sustainability/), accessed Oct. 20, 2020.

[11]. A. König, BeE-Nose – An In-Hive Multi-Gas-Sensor Extension to the IndusBee4.0 System for Hive Air Quality Monitoring and Varroa Infestation Level Estimation, in 'Advances in Signal Processing: Reviews', Book Series (S. Y. Yurish, Ed.), *IFSA Publishing*, Vol. 2, 2021, pp. 443-463.

[12]. Sensirion, Datasheet SGP30 Indoor Air Quality Sensor (http://www.sensirion.com).

[13]. A. Vanarse, J. I. Espinosa-Ramos, A. Osseiran, A. Rassau, N. Kasabov, Application of a Brain-Inspired Spiking Neural Network Architecture to Odor Data Classification, *Sensors*, Vol. 20, Issue 10, 2020, 2756.

[14]. T. Nowotny, A. Z. Berna, R. Binions, S. Trowell, Optimal feature selection for classifying a large set of chemicals using metal oxide sensors, *Sensors and Actuators B: Chemical*, Vol. 187, 2013, pp. 471-480.

[15]. E. Martinelli, C. Falconi, A. D'Amico, C. Di Natale, Feature extraction of chemical sensors in phase space, *Sensors and Actuators B: Chemical*, Vol. 95, Issue 1-3, 2003, pp. 132–139.