



# Inferring ethylene temporal and spatial distribution in an apple orchard (*Malus domestica* Borkh): a pilot study for optimal sampling with a gas sensor

João Valente<sup>1</sup> · Rodrigo Almeida<sup>2</sup> · Lammert Kooistra<sup>2</sup>

Received: 8 April 2019 / Revised: 13 October 2020 / Accepted: 29 October 2020 / Published online: 2 December 2020  
© The Author(s) 2020

## Abstract

Apples emit a volatile organic compounds during the ripening process named ethylene, which can be used to infer the optimal harvest time. Currently, the fruit ethylene emission is assessed in controlled facilities, thus laborious and expensive. This article pioneers the study of assessing ethylene emissions in uncontrolled environments. However, understanding how the ethylene spatial temporal dynamics in an open field, its still elusive. Therefore, this paper provides a model from an (*Malus domestica* Borkh) apple orchard for simulation and analysis of ethylene behaviour. We demonstrate that the model is able to explain the ethylene emissions behaviour in an orchard field when subject to different wind speeds, directions and ripeness stages. Based on that we have investigated different sampling schemes—regular and random—for capturing the variability of ethylene in an orchard using an electrochemical gas sensor. These results show that a random sampling scheme performs 25% better than an equivalent regular-defined grid. Moreover, the measurements acquired locally in the rows tend to be 10% more reliable than in other locations from the orchard. Finally, the ethylene variability can be assessed with a confidence of 75% using 4 and 16 sampling points.

**Keywords** Harvest time date · Ethylene gas simulation · Open-source software · Electrochemical gas sensor · Apple orchard

## 1 Introduction

For farmers, determining the optimal harvesting time (OHT) is key. Harvesting immature fruits results in poor quality when ripe and higher susceptibility to mechanical damage. Harvesting overripe fruits results in soft and flavourless products, with a very short shelf-life (Kader 1999).

Maturity indices exist to support farmers decision making across the growing season, such as for finding the OHT. Maturity indices can be obtained via destructive and non-destructive methods and take into account chemical composition (e.g., total acidity), physical properties [e.g., firmness (Zhang et al. 2008)], fruit physiological changes [e.g., ethylene emission rate (Kathirvelan and Vijayaraghavan 2017)], and chronological features [e.g., days after blooming (Knee 2002)]. Destructive methods are methods which the fruits or vegetable are subject to a test that degrades, e.g., penetration, compression, the product leaving it not consumable, while non-destructive methods use remote sensing techniques, e.g., acoustic impulse response, near-infrared (Arefi et al. 2015).

The ethylene emission rate can be detected using gas chromatography techniques, electrochemical gas sensors and optical sensors (Cristescu et al. (2012)). Research so far using electrochemical gas sensors has been aiming at developing methods to assess volatile organic compounds (VOC) emissions using concentration chambers, apples (Brezmes et al. 2000; Łysiak 2014; Pathange et al. 2006;

---

Communicated by Youn Young Hur, Ph.D.

---

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s13580-020-00316-9>) contains supplementary material, which is available to authorized users.

---

✉ João Valente  
joao.valente@wur.nl

<sup>1</sup> Information Technology Group, Wageningen University and Research, Wageningen, The Netherlands

<sup>2</sup> Laboratory of Geo-Information Science and Remote Sensing, Wageningen University and Research, Wageningen, The Netherlands

Saevels et al. 2003), pears (Ma et al. 2016; Brezmes et al. 2000; Zhang et al. 2008), tomatoes (Wang and Zhou 2007; Mandarin Gómez et al. (2006), banana (Kathirvelan and Vijayaraghavan 2017), mango (Lihuan et al. 2017), and grapes (Athamneh et al. 2008) among other fruits have been studied with this methodology. Moreover, the measurement of ethylene emission was done using enclosures of strawberries in the plant in order to track the emissions of the same fruit over time. The results of such research show promise, especially in postharvest processing (Iannetta et al. 2006).

Current methods for assessing fruit maturity require the sampling of individual fruit in the field and further assessment in the lab. That process is both labor intensive, since it requires an operator to physically go to the field and sample fruits, and dependent on the individual fruits that are sampled (Sun et al. 2019).

This work aims to investigate how ethylene could be assessed in an uncontrolled environment like an orchard field, using an electronic nose to improve current assessment techniques. The advantage of measuring ethylene directly in an orchard field in comparison with the current methods are enhanced in Table 1.

The main problem is that the sensors measurements are highly affected by the conditions in the field such as temperature, sunlight, among others (Kader (2002)). Moreover, the sensor measurements sensibility is also subject to the interference and the environment conditions, such as relative humidity, wind speed and wind direction (Popoola et al. 2016). Finally, electrochemical sensors are prone to cross-sensitivity with other agricultural derived anthropogenic greenhouse gases (GHG) emissions—Carbon dioxide (CO<sub>2</sub>) (Vermeulen et al. 2012), nitrous oxide (N<sub>2</sub>O) (Smith et al. 2007), and Methane (CH<sub>4</sub>). Through CH<sub>4</sub>, the agricultural production also contributes to secondary GHGs such as tropospheric ozone (O<sub>3</sub>) (Dentener et al. 2005).

The study presented in this paper is a first attempt at understanding the potential and the limitations of electrochemical sensors measurements in real-time in an open field,

creating with it a theoretical framework from which further work can be developed.

For the best of our knowledge, no studies have been made so far regarding open field sampling schemes and its application when subject to the wind speed and direction. Therefore, we present a study where it can be shown:

1. How the spatial arrangement from the orchard trees influence the ethylene distribution behavior.
2. How ethylene distributes within the apple orchard when subject to different wind speeds.
3. Where (within the orchard) the ethylene concentrates.
4. Which sampling scheme proposed performs better by checking if a sample is significantly different from the mean of the population (z-score).

## 2 Materials and methods

In Monroy and Gonzalez-Jimenez (2017) it is stated that the high complexity of the dispersion of gases and the high sensibility to very small changes in the measurement environment make it difficult to create reproducible, and therefore, meaningful experiments. For that purpose a 3D gas dispersion simulator denoted as GADEN was developed (Monroy et al. 2017). This simulator not only emulates the gas on a three-dimensional space but takes into account obstacles and air flow dynamics.

Therefore, a model from an orchard field was built from empirical knowledge of in situ measurements taken in the real environment: Area size, number of trees, tree dimensions, space between trees, space between rows, fruit load per tree, and wind speed. The complete GADEN setup and configuration parameters used in the simulations and respective files are available in a public repository.

## 3 Study area

The study area and apple orchard is located in the Wageningen Plant Research for Flowerbulbs, Nurserystock and Fruits in Randwijk, The Netherlands (Fig. 1). A test plot of 0.17 ha of apple trees was selected (study area enhanced in red). The plot is used to conduct trials on novel apple cultural practices, using substrate that is placed in the ground where the trees are planted. It has a length of 5 m in between tree rows and 1.1 m between trees in the row which results in 14 lines and about 300 trees in the plot. The apple (*Malus domestica* Borkh) cultivar selected was Junami because is a very popular product in Netherlands and one of the most consumed apples due to its refreshing fruity flavor. Furthermore, it has a very good reputation for what shelf time regards.

**Table 1** Comparison between conventional ethylene assessment practices in controlled facilities, and improved ones in uncontrolled outdoor field

Highlights	Current	Ideal
Sampling frequency	Daily <sup>x</sup>	Real-time
Human support	Field & laboratory <sup>y</sup>	Field <sup>z</sup>
Data availability	Discrete	Continuous
Assessment type	Invasive	Remote
Costs	High	Low

<sup>x</sup>Sample are taken in a daily basis

<sup>y</sup>Individual fruits picking and chemical composition analysis

<sup>z</sup>Sensors placing and removing from the field



**Fig. 1** Picture of the selected study area in Randwijk, The Netherlands

Therefore, the interest in carrying out further studies within this cultivar.

### 3.1 Building the environment

A total of 18 trees were simulated as cylinders with a diameter of 0.3 m. This shape was chosen since it allows for wind flow in between rows and lines. It is important to note that the effect of small leaves and branches in between stems is not taken into account in these simulations. From the internal volume of this environment a grid mesh was generated with a 0.1 m resolution. The edges of the environment were set as outlets—where airflow and gas are allowed to leave the environment—while  $x = 0$  and  $y = 0$  were set as environment inlets.

### 3.2 Ethylene emission simulations

For ethylene emission, three phases were defined according to literature: pre-climacteric, entering climacteric and climacteric (Lougheed and Franklin 1971; Reid et al. 1973). Fruit position was determined using empirical knowledge of the field experiments, especially when it comes to the height of the fruit. Fruit load was determined with average yield values of commercial apple plots. The parameters

used to model the trees ( $n = 18$ ) from the apple orchard are described in Table 2.

For a matter of simplification an artificial center was defined in each tree for the total emissions of ethylene. This center ( $P$ ), can be described as the average position of emission sources of the tree, and is defined by a height ( $h$ ) and direction in relation to the main stem ( $dir$ ). This  $dir$  parameter in relation to the stem is defined in order to make the distribution of this parameter uniform, and therefore, the number of directions must be divisible by the number of

**Table 2** Summary of ethylene emission parameters based in the work of Reid et al. (1973, Lougheed and Franklin (1971) and expert knowledge

Parameter	Symbol	Unit	Distribution
Ethylene emission	$e_i$	$\mu\text{L}/\text{h kg}$	
Pre-climacteric	$e_1$		$e_1 \sim \mathcal{N}(2.7, 1.6^2)$
Entering climacteric	$e_2$		$e_2 \sim \mathcal{N}(7.3, 4.5^2)$
Climacteric	$e_3$		$e_3 \sim \mathcal{N}(20, 15^2)$
Fruit load per tree	$l$	$\text{kg}$	$l \sim \mathcal{N}(21.5, 12.8^2)$
Emission per tree	$E_i$	$\mu\text{g}/\text{s kg}$	$E_i = e_i * l$
Position	$P$	$(h, dir)$	
Height	$h$	$\text{m}$	$h \sim \mathcal{N}(1.4, 0.5^2)$
Direction	$dir$		$dir \sim \mathcal{U}(1, 6)$

trees. In this case six directions were defined, each one is the 6th part of a circle, the equivalent to 60°. This number of directions was defined to reduce the computational workload during the simulations. The six wind speed directions are sufficient to cover all potential cases and avoid occlusions.

The simulated model is illustrated in Fig. 2.

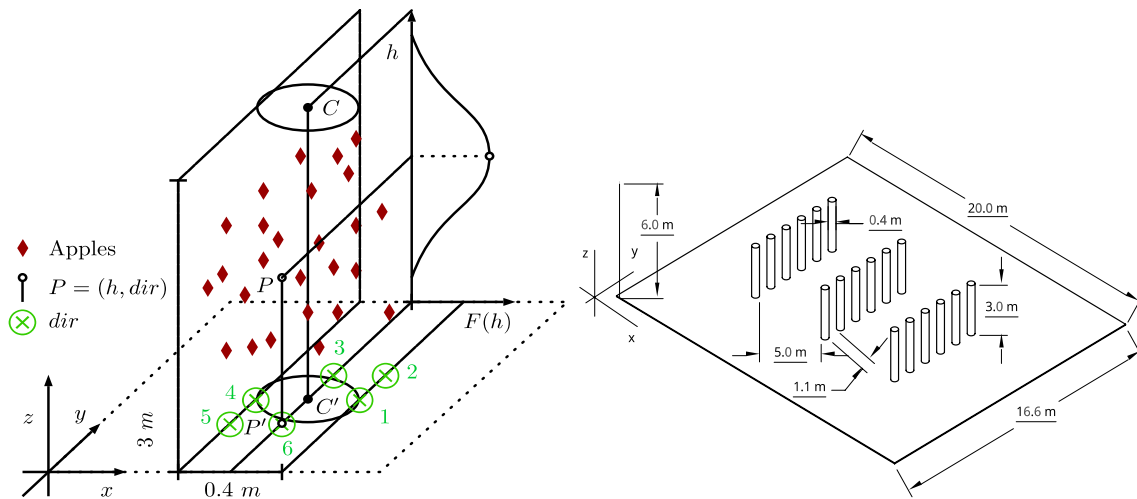
### 3.3 Wind simulations

For the wind speed parameters, the observations made in the field and also the records of a nearby weather station were used to determine the three wind speed scenarios (low, mean and high). Additionally, one of the constraints of the simulator used is that if stochastic turbulence occurs in the environment, the simulator result will not necessarily match a real world scenario. This happens especially when an adjective

airflow is not present (mass air movement) e.g. with very low or null wind speeds (Monroy et al. 2017). Therefore to summarize three scenarios for wind speed were constructed (0, 2 and 5 m/s) and at two different wind directions ( $\vec{x}$ ,  $\vec{y}$ ) which amounts to 5 complete wind simulations.

### 3.4 Optimal sampling

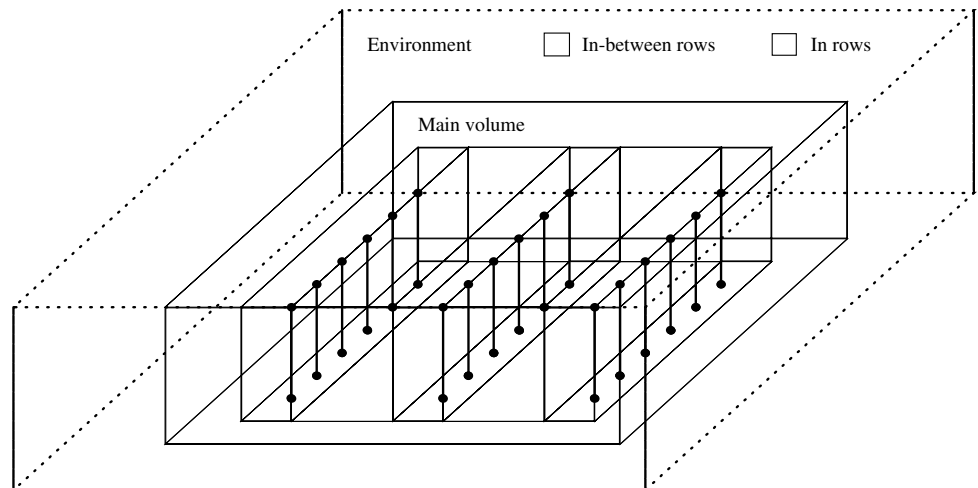
In order to determinate the best sampling scheme, four distinct zones (or volumes) - in rows, in-between rows, and in the environment- were defined in the simulated environment (Fig. 3). These are the potential orchard field elements where ethylene can be measured with sensor. In each zone a statistical analysis focusing on the mean, standard deviation, variance, maximum value, and number of occupied cells, across simulation time steps was carried out.



**Fig. 2** The CAD model of the orchard used for GADEN (right). Theoretical model of an individual tree (left). The average position of emission sources (apple fruit) is defined as P. The center coordinates

of the stem is given by C. The direction (dir) is the projection of point P in the xy-plane (P'). Taking  $C' = (x_c, y_c)$  as the center stem location and dir,  $P' = (x_c + 0.3 * \cos \frac{dir * 2\pi}{6}, y_c + 0.3 * \sin \frac{dir * 2\pi}{6})$

**Fig. 3** Delimitation of zones: environment, main volume, in rows and in-between rows





### 3.5 Statistical analysis

For statistical analysis of the simulated data, a normal distribution is a key pre-requisite. The ethylene concentration distribution is as expected of gas concentration data in an empty environment, positively skewed. In order to obtain a normal distribution of the data regarding ethylene concentration, a Box–Cox transformation was applied with the  $\lambda$  parameter being optimized for each simulation run.

The Box-Cox transformation (Box and Cox 1964), with the  $\lambda$  parameter being optimized using the profile likelihood function, is defined below,

$$y_i^{(\lambda)} = \begin{cases} \frac{y_i^\lambda - 1}{\lambda} & \text{if } \lambda \neq 0, \\ \log(y_i) & \text{if } \lambda = 0. \end{cases}$$

When doing statistical tests on this data, it is important to note that two different random variables are at play,  $O$  and  $C$ :

$O$  = Occupation of a cell,  $O \in \mathbb{B}$ .

$C$  = Ethylene concentration of a cell,  $C \in \mathbb{R}_{>0}$

The probability of a certain cell to be occupied with ethylene is given:

$$P(O = 1) = \frac{\text{Number of occupied cells}}{\text{Total number of cells}}$$

Finally, the probability of an occupied cell to have a  $c$  and the probability of a cell to be occupied with a  $c$  is given by,

$$P(C = c | O = 1) = \frac{P(O = 1 \cap C = c)}{P(O = 1)} \Leftrightarrow$$

$$P(O = 1 \cap C = c) = P(C = c | O) * P(O = 1)$$

The statistic used for the purpose of analyzing the results of different sampling schemes was the  $z$ -test. The population used, that is the ethylene concentration of the occupied cells in each simulation, is assumed as approximately normally distributed after applying the Box-Cox transformation. Additionally, all parameters of this population are known: mean ( $\mu$ ) and standard deviation ( $\sigma$ ). For a certain sample of  $n$  cells with a sample mean  $\bar{s}$ , the  $z$ -test computes the  $z$ -score:

$$z - \text{score} = \frac{\bar{s} - \mu}{\frac{\sigma}{\sqrt{n}}}$$

Moreover, two hypothesis have been defined: The mean of the sample isn't significantly different from the mean of the population ( $H_0$ ), and the mean of the sample is significantly different from the mean of the population ( $H_1$ ).

Given a significance level of 95%  $H_0$  is rejected in favor of  $H_1$  if  $|z - \text{score}| > 1.96$ . This is a two-tail  $z$ -test. The  $z$ -score is a measurement of deviation of the sample mean from the population mean taking into account both the

sample size ( $n$ ) and the population standard deviation ( $\sigma$ ). With this statistic, it is straightforward to make conclusions on the confidence level of a certain sample to estimate the population mean. If  $H_0$  is rejected then the sample mean deviates significantly from the population mean, which in the case of sampling scheme design, informs us that the sampling positions chosen are not appropriate to estimate the population mean.

In order to create an appropriate and optimal sampling scheme for measuring ethylene concentration in an orchard, two attempts were made: using a random sampling scheme, and using a regular grid. In all cases, the number of samples ( $n$ ) was varied from 1 to 4 to 16, which are sample sizes that are feasible in real conditions. This sampling scheme analysis has the ultimate goal of trying to maximize the information value of a set of samples, minimizing the cost of that sampling.

### 3.6 Random sampling

In order to test the performance of a random sampling scheme, an iterative approach was taken. For each simulation 100 sample sets were selected per time step. Each sample set consisted of 1, 4 or 16 randomly selected points. The simulations were performed over 20 time steps of 15 s each. Only the presence of ethylene was analysed in these samples. If there was ethylene present in at least one of the elements of the sample set, that sample was taken into account and recorded.

Secondly, other 100 sample sets were randomly selected out of the occupied cells. This sample sets were then used to perform the  $z$ -test and the result recorded. This iterative process provides the probability of finding an occupied cell and also the probability of a sample of occupied cells to belong to the population where it was taken from, using  $\mu$  as defined before and a significance level of 95%. This process was also repeated for each of the zones defined in Fig. 3. Finally, a composite probability was computed.

The process described above is useful for determining the confidence level of a sample related to the population where it was sampled from, but as mentioned before, the goal is to differentiate between different stages. For that, it is relevant to run the same process but sampling from populations that are known as being different. For instance, to determine if a random sample from the main volume of simulation  $n$  would, when applying the  $z$ -test considering the population to be the simulation  $m$ , return a non-significant difference between means. That will provide a confidence level in differentiating different populations using only a sample, and additionally, differentiation between fruit maturity stages.

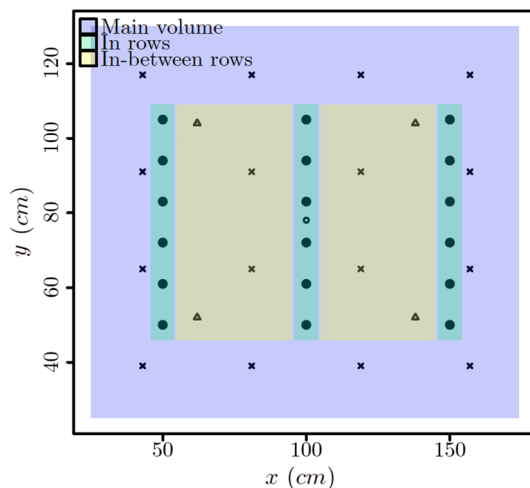
### 3.7 Regular grid sampling

In order to test the performance of a regular grid sampling scheme, a grid composed of 1, 4 and 16 points was constructed, and its center points (Fig. 4). The same process as in the random sampling evaluation using the  $z$ -test was done with the regular grid sampling points.

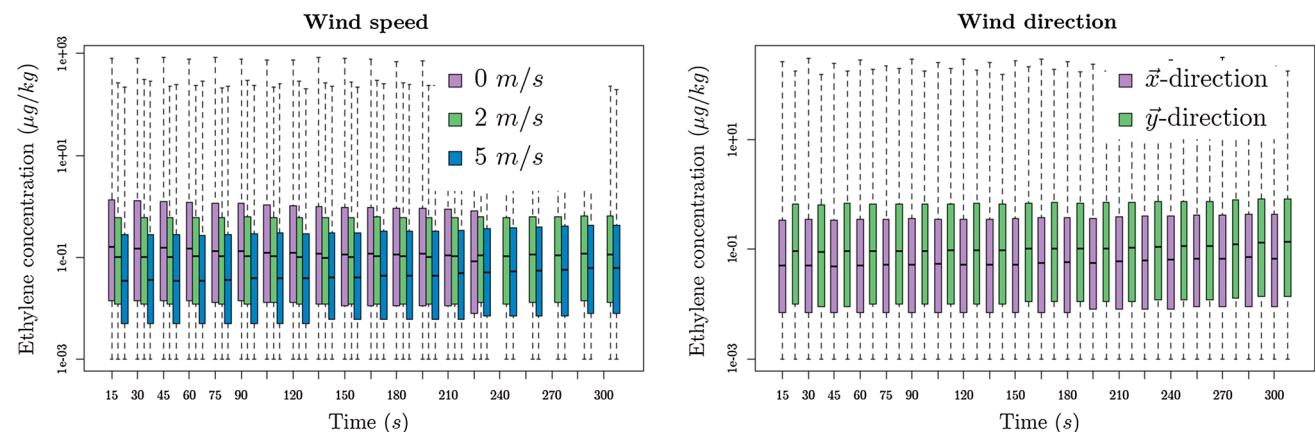
## 4 Results

### 4.1 Ethylene distribution behaviour

After running the simulations, plots of the maximum ethylene concentration measured in the  $xy$ ,  $xz$  and  $yz$ -plane were constructed. The simulation results for pre-climacteric (Fig.



**Fig. 4** Sampling points with a regular grid using  $n=1$  ( $\square$ ),  $n=4$  ( $\Delta$ ) and  $n=16$  ( $\times$ ) in the main volume ( $xy$ -plane). All the points have  $z=1.6$  The ( $\bullet$ ) represents the tree positions



**Fig. 5** The average, standard deviation and ethylene concentration limits. For the wind speed: 0 m/s ( $0.13 \pm 0.52$   $\mu\text{g}/\text{kg}$ ), 2 m/s ( $0.10 \pm 0.30$   $\mu\text{g}/\text{kg}$ ), and 5 m/s ( $0.04 \pm 0.16$   $\mu\text{g}/\text{kg}$ ). For wind direction:  $x$  ( $0.05 \pm 0.19$   $\mu\text{g}/\text{kg}$ ) and  $y$  ( $0.11 \pm 0.35$   $\mu\text{g}/\text{kg}$ )

S1), entering climacteric stage (Fig. S2), and for the climacteric stage (Fig. S3) are available online as supplementary material. The heat maps depict the maximum ethylene concentration in the orchard through different wind speeds and provide the gas dispersion visualization in three perspectives. This is aimed at uncovering the spatial distribution of the gas in the simulated apple orchard field.

In general, higher maximum concentration values can be found very close or on the ethylene emission sources positions with higher emission rate. The behaviour of gas distribution in a low wind speed condition is relevant: the spread of gas in relation to the ethylene emission source is quite small in all directions. Nevertheless, the apparent range of maximum ethylene concentration decreases with an increase in wind speed, which is expected. When wind speed is high, a shift occurs on this position and the emission source appears to be further away from the trees than it actually is. In the  $x$  direction this is more discernible in all stages.

Moreover, a shadow effect of the simulated trees is also clearly visible, especially when the wind flows in the  $y$  direction, causing a concentration of gas after the last tree in the row. This effect appears to be less important in the  $x$  direction due to the fact that the distance between rows is 5 times the distance between trees. This can be stated visually in Fig. S1-3 for the maximum wind speed defined, but also quantified in Fig. 5.

It should be noticed that the gas distribution is absent along the  $z$ -plane above the tree level (3m) in all simulations. This might be surprising but it is also expected since ethylene has about the same molecular weight as a standard air mixture. However, in these simulations, no  $\vec{z}$  direction wind flow was considered, which might occur in a real setting with a more complex wind flow, that can cause gas dispersion above the tree canopy.

Furthermore, the maximum concentration of ethylene can be found roughly on the same positions in every stage that has the same wind conditions. This is expected but still an important point when considering sampling strategies since with this, one can conclude that the same sampling strategy can be applied to different ethylene emission stages with confidence.

In Fig. 6 it's quickly concluded that the average concentration of ethylene in the occupied cells of the environment is more or less stable in all simulations across time. This suggests that a steady state is reached regarding the ethylene concentration distribution. Pre-climacteric simulations show the lowest average ethylene concentration at the same time that lower wind speeds increase the average ethylene concentration (see Table 3).

The standard deviation is also constant across time in most simulations. The exception are the simulations with very low wind speeds (0 m/s), where there is an increase in the standard deviation in the environment and a decrease in all the other zones, more sharply represented in-between rows (see Fig. 5). Additionally, in the environment, a higher ethylene emission and higher wind speeds return a higher standard deviation of the ethylene emission concentration.

The maximum recorded ethylene concentration is more variable across time than the other displayed statistics but is again relatively stable. In this instance, higher emission rates and lower wind speeds return the highest maximum ethylene concentration. The maximum ethylene concentration recorded in all simulations is observed in the climacteric stage, and at low wind speeds. The maximum values are on or close to the emission position which is in the rows. However, this maximum value decreases with time in the low wind speed simulations which again suggests a dilution effect with time when wind is not a factor.

The number of ethylene occupied cells in the environment is another important metric since it provides information of the probability of finding an ethylene filled cell. The difference in number of occupied cells between simulations occurs in-between the rows. This happens since the wind direction plays a fundamental role determining if the volume in between the rows gets occupied with ethylene or not as is clearly visible in the Fig. S1-3: when wind flows in the  $x$  direction, gas is distributed in between rows while wind in the  $y$  direction keeps the gas in the rows.

When looking for higher concentrations of ethylene the zones that can give the most reliable estimation of the ethylene concentration are those measurements obtained in-rows. Moreover, when measuring in an outdoor environment—such as, a fruit orchard—wind is an important factor to have into consideration before employing any sampling scheme. Figure 7 shows that when deciding about the most reliable zone to take samples from the in-rows still prevail over

in-between rows, because the ethylene concentration is in general higher in rows.

## 4.2 Comparing sampling strategies

The aim of this comparison is to find the best strategy to estimate the average concentration of ethylene in the environment. For this purpose, and for each simulation and time step, the statistic analysis explained in Section Statistical analysis is applied.

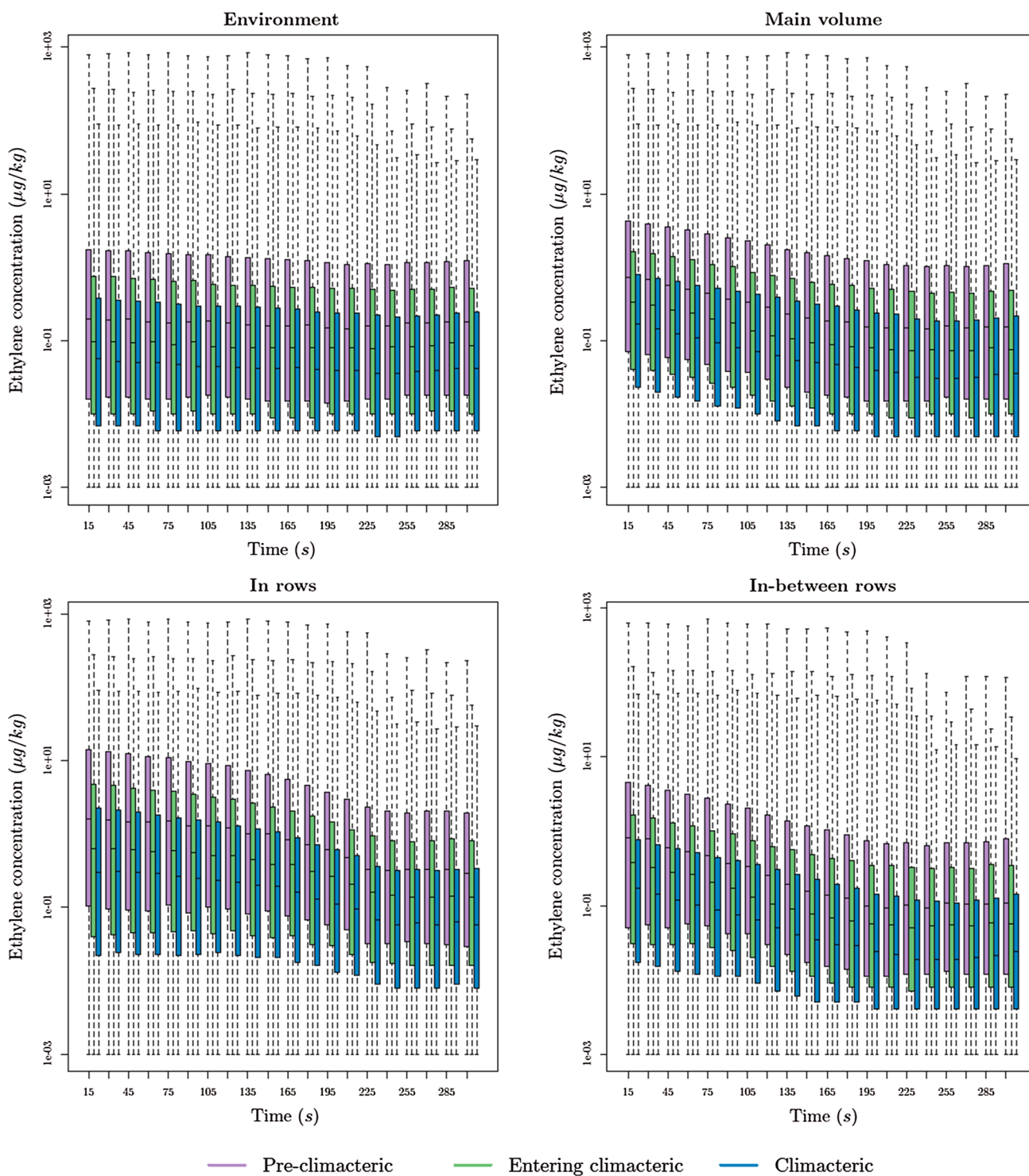
## 4.3 Random sampling

When testing a one point random sample in the simulated environment to determine the average ethylene concentration we can easily conclude that the sample point is significant if located in an ethylene filled cell. That is, the confidence level of 1 random sample is very close to 100% in all zones and simulations and also across time as shown in Fig. 8. The issue is that the probability of a randomly sampled cell to be ethylene occupied is low in most zones. In the rows the average probability of finding ethylene is higher than in the other zones (see Fig. 8). With 4 samples the result is more or less the same as with 1 samples. The confidence level of the sample decreases very slightly. Nevertheless, when looking at ethylene existence, the probability is in general much higher in all zones and simulations, when compared with only 1 samples. This also results in a composite level that is higher in rows as can be appreciated in Fig. 8. Finally, the confidence level remains mainly unchanged from 16 to 4 random samples. The probability of existence of ethylene is however much higher than with only 4 samples, especially in the rows. This results in a composite confidence level that is also very high in the rows (see Fig. 8).

## 4.4 Regular grid

The regular grid  $z$ -score is plotted in Fig. 9 for the different amount of points, and it's relevant to note that it increases in general with the number of points. Using only the center point, almost all the simulations fall within the confidence interval for the  $z$ -score while this is not true for the other number of points. There is also a clear tendency to underestimate the population mean since overall the  $z$ -score is more negative than positive. Overestimation only occurs when wind speeds are high and ethylene emission is low.

When making the comparison between different simulations with the regular grid points the same effect that was determined with the random points is concluded: it's more likely to confound a higher ethylene emission stage with a sample from a lower emission stage than the reverse. Additionally, it appears that between climacteric simulations the confidence level is high in all number of regular grid points.



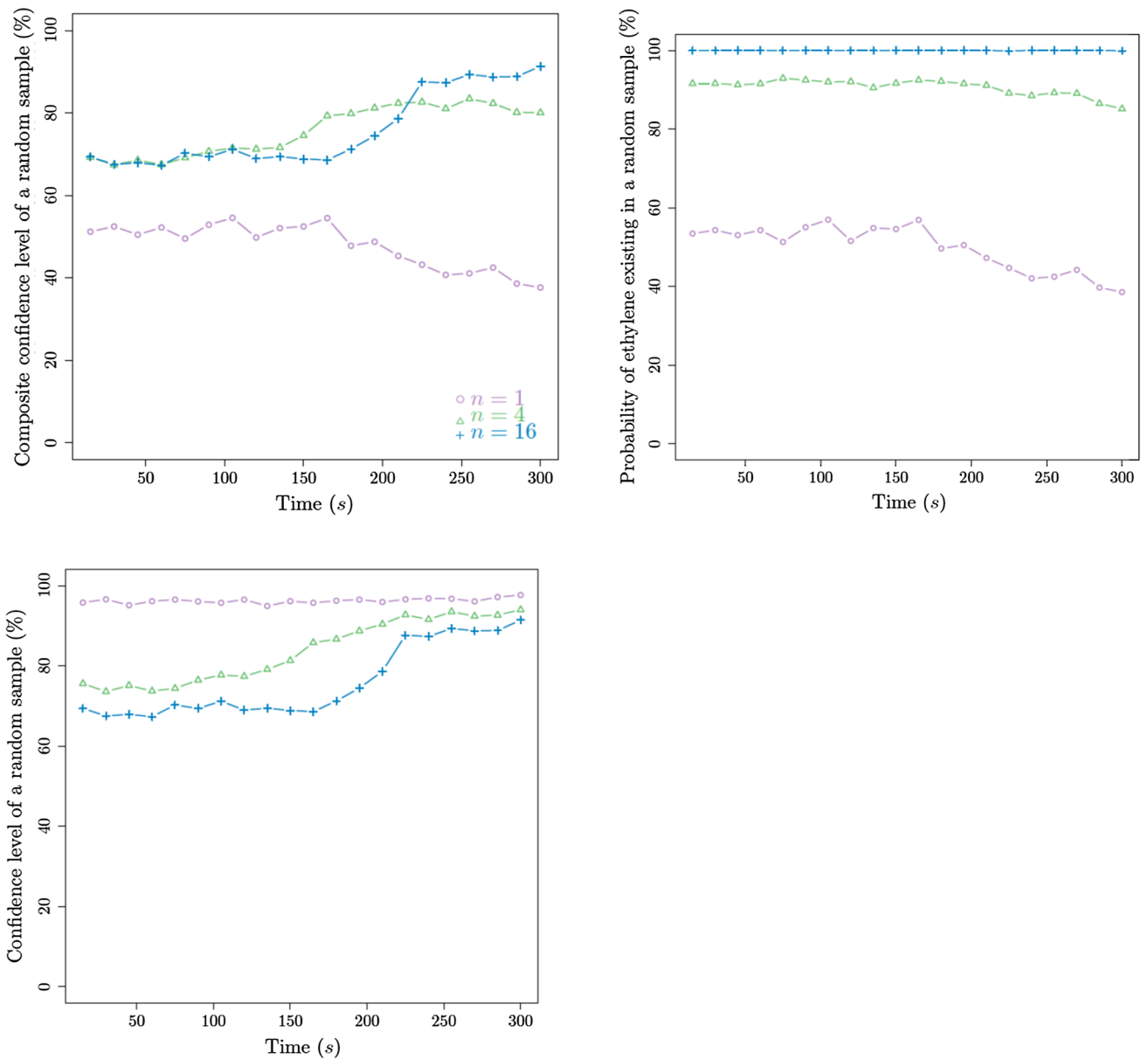
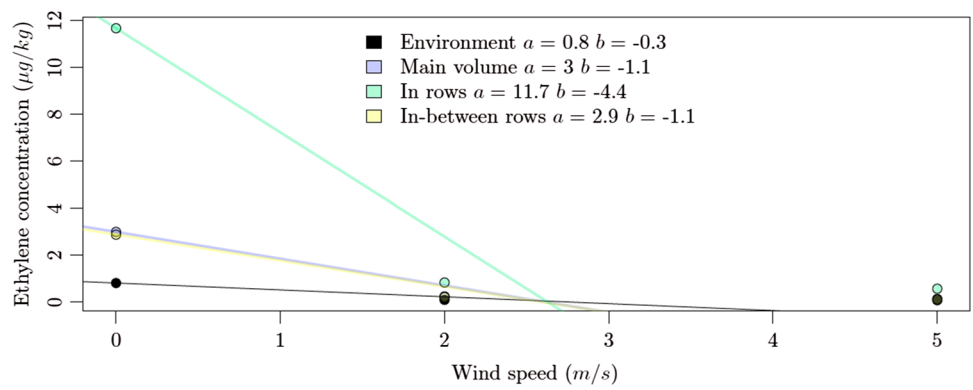
**Fig. 6** Average, standard deviation, maximums and minimums of concentration of ethylene ( $\mu\text{g}/\text{kg}$ ) in the different simulations and zones

**Table 3** Average, standard deviation, absolute values of concentration of ethylene ( $\mu\text{g}/\text{kg}$ ) in the different simulations and zones

	Environment	Main volume	In rows	In-between rows
Pre-climacteric	$0.04 \pm 0.14$	$0.07 \pm 0.18$	$0.17 \pm 0.52$	$0.06 \pm 0.14$
Entering climacteric	$0.09 \pm 0.29$	$0.14 \pm 0.39$	$0.38 \pm 1.17$	$0.13 \pm 0.33$
Climacteric	$0.17 \pm 0.67$	$0.30 \pm 0.98$	$0.90 \pm 3.28$	$0.28 \pm 0.85$



**Fig. 7** Relation between wind speed and average ethylene concentration in the 4 different zones. The coloured lines represent the trend line for each zone, as given by the equation  $y = a + bx$  where  $b$  is the decrease in average ethylene concentration ( $\mu\text{g}/\text{kg}$ ) per additional unit of wind speed ( $\text{m}/\text{s}$ )



**Fig. 8** Mean confidence level, probability of existence, composite level of the ethylene concentration random samples across time

**Fig. 9** Comparison of z-score between regular grid of 1, 4 and 16 points as displayed in Fig. 4. The lines show discontinuities when no ethylene concentration is present at any of the measurement points. The dotted lines represent the confidence interval defined in the z-test. Very positive z-scores ( $z$ -score  $> 1.96$ ) suggest an overestimation of the population mean while very negative ( $z$ -score  $< -1.96$ ) suggest an underestimation of the population mean

Finally, it should be noticed that the 4 and 16 points have similar confidence profiles in-rows, and in-between-rows with regard to random sampling. Actually, both fall in the fourth quarter of confidence, with an average difference of about 10%, where we can infer that 16 samples is not a better sample set versus a 4 samples. This approaches confidence matching is noticed in Fig. 10.

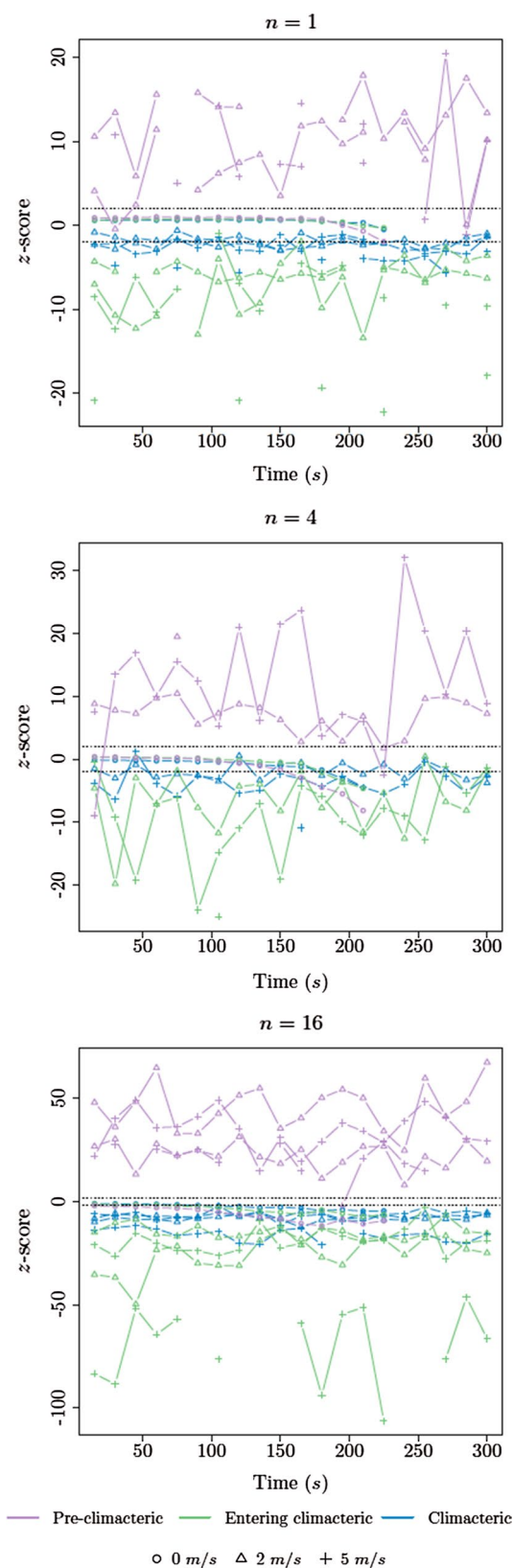
## 5 Discussion

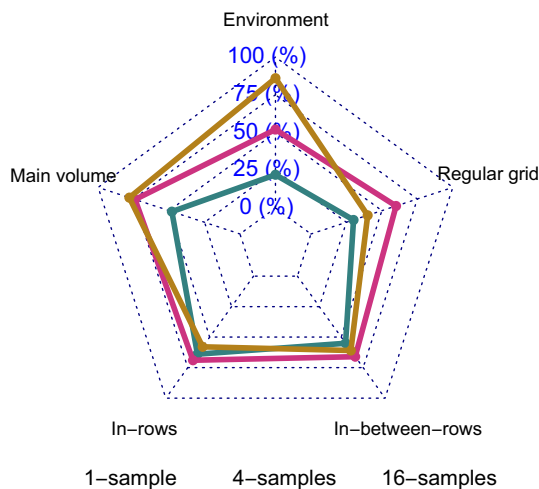
The spatial–temporal variation of ethylene in an apple orchard is not a parameter easy to assess because of its fast dynamics and unpredictable weather conditions on the field. This process is quite complex since a number of uncertainties are present. This was previously noticed in related research (Monroy and Gonzalez-Jimenez 2017). The research approach that was used was also to adopt a simulated model from the environment and emission sources (Monroy et al. 2017).

An effort was made to obtain an approximate computational model from the apple orchard to infer the spatial–temporal ethylene distribution when subject to different wind speed and directions. This is a key parameter in this study as already pointed out by Popoola et al. (2016). The field was modelled with respect to the field data acquired by plant research experts using conventional measuring instruments and invasive techniques (Cristescu et al. 2012).

Through the simulation we could verify some expected behaviours such as that for low wind speeds the gas remains close to the ethylene emission source; concentration decrease with the wind speed; the ethylene plume shifts in the wind direction as previously indicated in (Villa et al. 2016); pre-climacteric shows lower average ethylene concentration as enhanced in Paul et al. (2011); and gas dilution effect can be appreciated over the time for low-wind speeds (Génard and Gouble 2005).

It was interesting to observe in simulations that ethylene emission increases exponentially over the fruit maturity stages and that higher concentration values might be found within the orchard rows. Nevertheless, it should be enhanced that in-between-rows the variations are smaller. Which indicates that quantifying the magnitude in that zone will be less efficient—because the absolute ethylene concentration is smaller—but that the measurements will be more constant. Another important behaviour observed





**Fig. 10** Confidence level of the random (Environment, Main volume, In-Rows, In-between-rows) and Regular sampling strategies, versus sample numbers

it is that the gas distribution is absent in the z-plane, which indicates that the measurement should not be done above the trees.

A process to take in consideration when carrying out the experiments outdoor is the ethylene dispersion when subject to wind, and the dilution with time. It was verified that for wind speeds greater than 3 m/s the detection is less probable to occur. Moreover, it was verified that the average concentration achieves the steady-state in an open environment under ideal conditions.

The sampling scheme to be adopted plays an important role in the fields measurements because it defines the amount of resources and time that must be employed. More samples mean more execution time and consequently increase the cost and complexity of the operations. It's clear that a regular grid performs worse than a random sampling.

This study provide analytical evidences that a random sampling with 16 points will give a better estimate. Nevertheless, it was also proved that the sampling points optimization from 16 to 4 has an affordable confidence cost as shown in Fig. 10. The sampling strategy to adopt it will be dictated by the amount of resources and the time window available to carry out the field practices.

Having in consideration the results of this study, for future work further environment variables will be added to the current orchard model, e.g., cross sensitivity from other gases (Popoola et al. 2016). This step will give more insights how adapt the sampling scheme to sensors uncertainties and later to make a model uncertainty analysis (Uusitalo et al. 2015).

## 6 Conclusions

In conclusion a novel apple orchard simulator for ethylene emission analysis was developed based in open-source software and build using biophysical parameters from a (*Malus domestica* Borkh) apple orchard. The simulator was successful tested empirically. Different sampling schemes have been simulated to derive the best way to assess the ethylene concentration in an open field. This computational approach can be extended to other crops where there is the need to study further ethylene emissions in an uncontrolled environments and develop strategies to assess it. The authors are confident that this simulator could be useful for further research and for education purpose.

**Acknowledgements** This work was supported by the SPECTORS project (143081) which is funded by the European cooperation program INTERREG Deutschland-Nederland.

**Author contribution statement** JV perceived and planned the study. Development and statistical analysis was done by RA. The manuscript was written by RA, JV. LK applied for the funding acquisition. All authors reviewed and finalized the manuscript.

**Data availability** The open-source code with all the files necessary to run the simulations can be found in a public repository in <https://github.com/rodrigoalmeida94/gaden>

## Compliance with ethical standard

**Conflict of interest** The authors declare that there is no conflict of interest regarding the publication of this paper.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Arefi A, Moghaddam PA, Mollazade K, Hassanpour A, Valero C, Gowen A (2015) Meakness detection in agricultural crops: destructive and nondestructive tests: a review. *Compr Rev Food Sci Food Safety* 14:657–680. <https://doi.org/10.1111/1541-4337.12152>
- Athamneh AI, Zoecklein BW, Mallikarjunan K (2008) Electronic nose evaluation of cabernet sauvignon fruit maturity. *J Wine Res* 19:69–80. <https://doi.org/10.1080/09571260802164061>

- Box GEP, Cox DR (1964) An analysis of transformations. *J R Stat Soc Ser B Methodol* 26:211–252
- Brezmes J, Llobet E, Vilanova X, Saiz G, Correig X (2000) Fruit ripeness monitoring using an electronic nose. *Sensors Actuat B Chem* 69:223–229. [https://doi.org/10.1016/s0925-4005\(00\)00494-9](https://doi.org/10.1016/s0925-4005(00)00494-9)
- Cristescu SM, Mandon J, Arslanov D, De Pessemier J, Hermans C, Harren Frans J M (2012) Current methods for detecting ethylene in plants. *Ann Botany* 111:347–360. <https://doi.org/10.1093/aob/mcs259>
- Dentener F, Stevenson D, Cofala J, Mechler R, Amann M, Bergamaschi P, Raes F, Derwent R (2005) The impact of air pollutant and methane emission controls on tropospheric ozone and radiative forcing: CTM calculations for the period 1990–2030. *Atmos Chem Phys* 5:1731–1755. <https://doi.org/10.5194/acp-5-1731-2005>
- Génard M, Gouble B (2005) ETHY. a Theory of Fruit Climacteric Ethylene Emission. *Plant Physiol* 139:531–545. <https://doi.org/10.1104/pp.105.063339>
- Gómez AH, Wang J, Guixian H, Pereira AG (2006) Electronic nose technique potential monitoring mandarin maturity. *Sensors Actuat B Chem* 113:347–353. <https://doi.org/10.1016/j.snb.2005.03.090>
- Iannetta Pietro P M, Laarhoven L-J, Medina-Escobar N, James EK, McManus MT, Davies HV, Harren Frans J M (2006) Ethylene and carbon dioxide production by developing strawberries show a correlative pattern that is indicative of ripening climacteric fruit. *Physiol Plant* 127:247–259. <https://doi.org/10.1111/j.1399-3054.2006.00656.x>
- Kader AA (1999) Fruit maturity, ripening, and quality relationships. *Acta Horticult* 485:203–208
- Kader AA (2002) Postharvest technology of horticultural crops, vol 3311. University of California Agriculture, Natural Resources
- Kathirvelan J, Vijayaraghavan R (2017) An infrared based sensor system for the detection of ethylene for the discrimination of fruit ripening. *Infr Phys Technol* 85:403–409. <https://doi.org/10.1016/j.infrared.2017.07.022>
- Knee M (2002) Fruit quality and its biological basis, vol 9. Crc Press, Cambridge
- Lihuan S, Liu W, Xiaohong Z, Guohua H, Zhidong Z (2017) Fabrication of electronic nose system and exploration on its applications in mango fruit (M. Indica Cv. Datainong) quality rapid determination. *J Food Meas Char* 11:1969–1977. <https://doi.org/10.1007/s11694-017-9579-1>
- Lougheed EC, Franklin EW (1971) Ethylene Production of Apples at 20 c After Storage at 0 c. *Can J Plant Sci*, 51(2). NRC Research Press: 170–72. <https://www.nrcresearchpress.com/doi/pdfplus/10.4141/cjps71-031>
- Łysiak G (2014) Measurement of ethylene production as a method for determining the optimum harvest date of ‘Jonagored’ apples. *Folia Horticult*. <https://doi.org/10.1515/fhort-2015-0002>
- Ma L, Wang L, Chen R, Chang K, Wang S, Xinran H, Sun X et al (2016) a low cost compact measurement system constructed using a smart electrochemical sensor for the real-time discrimination of fruit ripening. *Sensors* 16:501. <https://doi.org/10.3390/s16040501>
- Monroy J, Gonzalez-Jimenez J (2017) Gas classification in motion: an experimental analysis. *Sensors Actuat B. Chem* 240:1205–1215. <https://doi.org/10.1016/j.snb.2016.09.013>
- Monroy J, Hernandez-Bennets V, Fan H, Lilienthal A, Gonzalez-Jimenez J (2017) GADEN: A 3d gas dispersion simulator for mobile robot olfaction in realistic environments. *Sensors* 17:1479. <https://doi.org/10.3390/s17071479>
- Pathange PL, Parameswarakumar M, Marini RP, O’Keefe S, Vaughan D (2006) Non-destructive evaluation of apple maturity using an electronic nose system. *J Food Eng* 77:1018–1023. <https://doi.org/10.1016/j.jfoodeng.2005.08.034>
- Paul V, Rakesh P, Srivastava GC (2011) The fading distinctions between classical patterns of ripening in climacteric and non-climacteric fruit and the ubiquity of ethylene—an overview. *J Food Sci Technol* 49:1–21. <https://doi.org/10.1007/s13197-011-0293-4>
- Popoola Olalekan AM, Stewart GB, Mead MI, Jones RL (2016) Development of a baseline-temperature correction methodology for electrochemical sensors and its implications for long-term stability. *Atmos Environ* 147:330–343. <https://doi.org/10.1016/j.atmosenv.2016.10.024>
- Reid MS, Rhodes MJC, Hulme AC (1973) Changes in Ethylene and Co2 during the ripening of apples. *J Sci Food Agric* 24:971–979. <https://doi.org/10.1002/jsfa.2740240815>
- Saevens S, Jeroen L, Berna AZ, Veraverbeke EA, Di Natale C, Nicolai BM (2003) Electronic nose as a non-destructive tool to evaluate the optimal harvest date of apples. *Postharvest Biol Technol* 30:3–14. [https://doi.org/10.1016/s0925-5214\(03\)00059-0](https://doi.org/10.1016/s0925-5214(03)00059-0)
- Smith P, Martino D, Cai Z, Gwary D, Janzen H, Kumar P, McCarl B et al (2007) Policy and technological constraints to implementation of greenhouse gas mitigation options in agriculture. *Agric Ecosyst Environ* 118:6–28. <https://doi.org/10.1016/j.agee.2006.06.006>
- Sun M, Yang X, Zhang Y, Wang S, Wong MW, Ni R, Huang D (2019) Rapid and visual detection and quantitation of ethylene released from ripening fruits: the new use of grubbs catalyst. *J AgricFood Chem* 67:507–513. <https://doi.org/10.1021/acs.jafc.8b05874>
- Uusitalo L, Lehtikoinen A, Helle I, Myrberg K (2015) An overview of methods to evaluate uncertainty of deterministic models in decision support. *Environ Modell Softw* 63:24–31. <https://doi.org/10.1016/j.envsoft.2014.09.017>
- Vermeulen SJ, Campbell BM, Ingram John SI (2012) Climate change and food systems. *Ann Rev Environ Resour* 37:195–222. <https://doi.org/10.1146/annurev-environ-020411-130608>
- Villa T, Farhad S, Mn K, Ma L, Felipe G (2016) Development and validation of a Uav based system for air pollution measurements. *Sensors* 16:2202. <https://doi.org/10.3390/s16122202>
- Wang J, Zhou Y (2007) Electronic-Nose technique: potential for monitoring maturity and shelf life of tomatoes. *New Zealand J Agric Res* 50:1219–1228. <https://doi.org/10.1080/00288230709510405>
- Zhang H, Jun W, Sheng Y (2008) Prediction of soluble solids content, firmness and pH of pear by signals of electronic nose sensors. *Anal Chim Acta* 606:112–118. <https://doi.org/10.1016/j.aca.2007.11.003>

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.