Modelling and inference of maize pollen emission rate with a Lagrangian dispersal model using Monte Carlo method

Otmane Souhar[1] , Alexis Marceau[2] and Benjamin Loubet[2]

[1]LIMA Laboratory, University of Chouaib Doukkali, BP. 20, 24000 El Jadida, Morocco and [2]National Institute of Agronomic Research, UMR EcoSys INRAE-AgroParisTech, 78850, Thiverval-Grignon, France

Abstract

This work explores the uncertainty of the inferred maize pollen emission rate using measurements and simulations of pollen dispersion at Grignon in France. Measurements were obtained via deposition of pollen on the ground in a canopy gap; simulations were conducted using the two-dimensional Lagrangian Stochastic Mechanistic mOdel for Pollen dispersion and deposition (SMOP). First, a quantitative evaluation of the model's performance was conducted using a global sensitivity analysis to analyse the convergence behaviour of the results and scatter diagrams. Then, a qualitative study was conducted to infer the pollen emission rate and calibrate the methodology against experimental data for several sets of variable values. The analysis showed that predicted and observed values were in good agreement and the calculated statistical indices were mostly within the range of acceptable model performance. Furthermore, it was revealed that the mean settling velocity and vertical leaf area index are the main variables affecting pollen deposition in the canopy gap. Finally, an estimated pollen emission rate was obtained according to a restricted setting, where the model studied includes no deposition on leaves, no resuspension and with horizontal pollen fluctuations either taken into account or not. The estimated pollen emission rate obtained was nearly identical to the measured quantity. In conclusion, the findings of the current study show that the described methodology could be an interesting approach for accurate prediction of maize pollen deposition and emission rates and may be appropriate for other pollen types.

Introduction

Nowadays, coverage of genetically modified crops is increasing rapidly which induces contamination risk in conventional or organic crops. Furthermore, there is great variability in cross-pollination rates due to many involved factors which describe pollen and canopy characteristics as well as to their interactions with its environment. It is therefore difficult to understand and account for the risks of cross-pollination only by experimental approaches (Treu and Emberlin, 2000). For these reasons and in order to minimize and monitor the risk of this potential and complex problem, various numerical methods have been developed and discussed in the literature (e.g., Dietiker *et al.*, 2011; Marceau *et al.*, 2011; Torimaru *et al.*, 2012). Maize pollen grains are roughly spherical with diameters around 90 µm and therefore have a high settling velocity compared with other pollen types (Raynor *et al.*, 1972; Di-Giovany *et al.*, 1995). As a consequence, dispersed maize pollen quantities are small: 95% of the emitted pollen is deposited within 10 m of the source and 99% at 30 m (Jarosz *et al.*, 2003). In order to simulate the dispersal of maize pollen as well as to estimate quantities of released pollen, many physically-based models have been developed:

The Lagrangian stochastic approach was introduced by Jarosz *et al.* (2004), Dupont *et al.* (2006) proposed an Eulerian approach, while Klein *et al.* (2003) adopted a statistical approach. Furthermore, several atmospheric dispersion models have been developed recently (see, e.g., Dietiker *et al.*, 2011; Zhang *et al.*, 2014; Baklanov *et al.*, 2017; Zink *et al.*, 2017). However, the numerical dispersion models need to be combined with a flow model, whether empirical, Lagrangian or Eulerian, describing the wind profiles and turbulence. In this setting, Arritt *et al.* (2007) have developed a three-dimensional Lagrangian model to simulate the dispersion of pollen and have validated the approach on daily measurements of deposition on the ground. Dupont *et al.* (2006) have compared Lagrangian and Eulerian models with experimental two-dimensional dispersion measurements in the wind direction. The Eulerian model is based on a flow model, Aquilon, in which an advection-diffusion conservation equation is integrated, while the Lagrangian stochastic model, SMOP (Stochastic Mechanistic mOdel for Pollen dispersion and deposition), uses an empirical model of flow. Astini *et al.* (2009) predict effective rates of cross-pollination at the end of the pollination season using the Lagrangian model of Arritt *et al.* (2007) combined with the pollination model of Lizaso *et al.* (2003). However, the model input

CAMBRIDGE UNIVERSITY PRESS

variables are subject to many sources of uncertainties including errors of measurement, absence of information and weak or partial understanding of the driving forces and mechanisms. These impose limits on our confidence in the response (i.e., outputs) of the model. Following an Eulerian model, Dupont *et al.* (2006) developed an empirical sensitivity analysis for some variables: settling velocity of pollen, Schmidt number, turbulence parameters and percentage of pollen re-entrained in the air after deposition on leaves. Contrary to our setting this methodology does not take into account possible interaction between variables. Good modelling practice requires that the specialists provide an evaluation of the confidence in their models by applying a Sensitivity Analysis (SA). Such an approach is used to study how a variation in the output of a model can be apportioned, quantitatively or qualitatively, in different sources of input variation. For performing a SA of a model, three approach classes are available: *Screening Designs* (SD), *Local Sensitivity Analysis* (LSA) and *Global Sensitivity Analysis* (GSA) described by Saltelli *et al.* (2000a). The Monte Carlo method, a GSA, has been applied herein because of its robustness and performance as stated by Saltelli *et al.* (2000a, 2000b). However, its application involves multiple simulations that make it expensive in terms of time taken to perform the analysis. In order to reduce the computation time, the Latin hypercube sampling technique has been used to create a matrix of values used as inputs to the SMOP model. Such a design is able to sample parameter values with a good homogeneous distribution in the parameters' domain as proved by McKay *et al.* (1979) and Iman (1992).

Emission of maize pollen could be estimated using plastic bags placed on the tassels (Hall *et al.*, 1982; Jarosz *et al.*, 2003). As a shortcoming of this method, tassels in closed environment conditions will induce a greenhouse effect, and as a consequence, this method provides a measure of potential emitted pollen rather than the actual emitted pollen. Due to this fact, Raupach (1989) and later Jarosz *et al.* (2005), Aylor *et al.* (2006) and Dupont *et al.* (2006) provided an alternative approach, which we use in this work, leading to a real emission rate that could be inferred from both measurements and simulations of pollen dispersion.

The objective of the current paper is to present a methodology to infer pollen emission rate taking into account the uncertainty of input variables of the SMOP model on deposition in a small canopy gap. Uncertainty of variables is related to retention of pollen by the plant canopy as well as to pollen aerodynamic characteristics.

Material and methods

Model description

The Lagrangian stochastic SMOP model has been under development in the National Institute of Agronomic Research, France since 2003 (Jarosz *et al.*, 2004). The model is designed to predict pollen dispersion and deposition downwind and within a flowering field. Displacement of individual pollen grains is given by the following two-dimensional stochastic differential equations.

$$du = a_u dt + b_u d\xi_u \quad dx = u \, dt \tag{1}$$

$$dw = a_w d + b_w d\xi_w \quad dz = (w - W_S) \, dt \tag{2}$$

where t is time; u and w are the horizontal and vertical air particle velocity components, respectively; $d\xi_u$ and $d\xi_w$ are random

numbers drawn from Gaussian distributions with mean zero and variance dt; W_s is the gravitational settling velocity of pollen grains (terminal velocity of pollen grains in still air); and a_u, b_u, a_w and b_w are the Langevin coefficients (depending on horizontal and vertical averages of the air velocity components, U and W, respectively; the horizontal and vertical Eulerian velocity variances, σ_u^2 and σ_w^2, respectively; the shear stress, $\overline{u'w'}$; and the Lagrangian velocity time scale, T_L). All of these parameters are described in Loubet (2000) and Loubet *et al.* (2006).

The SMOP model computes pollen trajectories taking into account both canopy and pollen characteristics, and micrometeorological parameters (Loubet, 2000; Jarosz *et al.*, 2004; Loubet *et al.*, 2006). The mean wind speed profile which is implemented in the model above and in the canopy is defined, respectively, as (Raupach *et al.*, 1996) and (Garratt, 1992).

$$u(z) = \begin{cases} u(h_c) \times \exp\left(\gamma \times \frac{z}{h_c} - 1\right) & z < h_r \\ \frac{u_*}{k}\left[\ln\left(\frac{z-d}{z_0}\right) - \Psi_m\left(\frac{z-d}{L}\right) + \Psi_m\left(\frac{z_0}{L}\right)\right] & z \geq h_r \end{cases} \tag{3}$$

where h_c is the canopy height; u_* is the friction velocity; k is the von Karman constant; L is the Monin-Obukhov length; h_r is the height where the exponential and logarithmic profiles of the model join which is the height of the inflexion point of the wind speed profile; Ψ_m is the stability correction function given by Paulson (1970) and Dyer (1974); $z_0 = 0.1 \times h_c$ is the roughness length and $d = 0.7 \times h_c$ is the displacement height (Kaimal and Finnigan, 1994); and $\gamma = 2$ is the attenuation factor (Raupach *et al.*, 1996; Loubet *et al.*, 2006). Required meteorological records were provided by an Agroclim station located in the experimental domain at Grignon, France. Data should be retrieved from the web site http://www6.paca.inra.fr/agroclim where global missions and objectives of Agroclim stations are given by the web site https://www6.paca.inrae.fr/agroclim/content/download/3214/32033/version/1/file/Plaquette+AGROCLIM+2020+FR_V1.pdf. For more information about the model, please visit the following website: https://www6.versailles-grignon.inra.fr/ecosys_eng/Staff/Photos-and-CV/L/Loubet-B/MODDAAS-SMOP.

Sensitivity analysis methods

Modelling a real-life process with a computer program, one is often faced with the problem of what values should be used for the input variables. Therefore, the output results have uncertainty associated with the input values. Stochastic approaches are usually used to obtain probabilistic output based on 'randomly' distributed parameters. Analysts desiring to conduct a sensitivity analysis using such an approach are obliged to select a set of values for their model inputs. This selection may be done in different ways: the random sampling is usually used in cases where the only general behaviour of flows are accounted for; the stratified sampling and the Latin hypercube sampling techniques are preferably chosen when studies need more accuracy: (i) what is the uncertainty in the output given the uncertainty in the input? and (ii) how important are the individual elements of the input with respect to the uncertainty in the output? (McKay *et al.*, 1979; Iman, 1992). Latin hypercube sampling is a particular stratified sampling technique and in the absence of correlations between input variables, it generates a set of values by dividing the sample space of each variable into m intervals of equal probability (where m is the number of required model runs) and one value is randomly selected from each interval.

Regression analysis sensitivity

Using the least squared method, the model input variables X_j ($1 \leq j \leq p$) and the model output Y are used to construct a linear model which can be written as:

$$Y_k = b_0 + \sum_{j=1}^{p} b_j . X_{jk} + \varepsilon_k = \hat{Y}_k + \varepsilon_k \quad k = 1, \ldots, m, \quad (4)$$

where p is the number of the input variables; b_j ($1 \leq j \leq p$) is the ordinary regression coefficients (ORC) of the model output; ε_k is the regression residual; and $\hat{Y}_k$ is the regression-estimate of the model output for k^{th} model run which represents the best linear prediction. The signs of the coefficients b_1, b_2, …, b_p indicate whether Y_k increases (i.e., a positive coefficient) or decreases (i.e., a negative coefficient) as the corresponding X_{jk} values increases. The main drawback of the ORC method is the dependence of coefficients on physical units of variables and therefore these quantities cannot be compared with each other (Helton *et al.*, 2006). Because of this, the ORC in Eqn (4) is expressed in the standardized regression coefficients (SRC) or normalized regression coefficients (NRC) forms (Saltelli *et al.*, 2000a) as:

$$\frac{\hat{Y} - \bar{Y}}{\hat{S}} = \sum_{j=1}^{p} b_j^{(s)} \cdot \frac{X_j - \bar{X}_j}{\hat{S}_j},$$

and

$$\frac{Y}{\bar{Y}} = \sum_{j=1}^{p} b_j^{(n)} \cdot \frac{X_j}{\bar{X}_j},$$

where $[\bar{Y}, \hat{S}]$ and $[\bar{X}_j, \hat{S}_j]$ are the means and standard deviations of $(Y_1, \ldots, Y_m)$ and $(X_{j1}, \ldots, X_{jm})$; and $b_j^{(s)}$ and $b_j^{(n)}$ are the SRC and NRC coefficients, respectively. These new coefficients may be obtained easily from the ORC:

$$b_j^{(s)} = b_j . \frac{\hat{S}_j}{\hat{S}} \text{ and } b_j^{(n)} = b_j . \frac{\bar{X}_j}{\bar{Y}}.$$

$b_j^{(s)}$ and $b_j^{(n)}$ measure respectively the effect of moving each input variable away from its mean by a fixed fraction of its standard deviation and the relative change $\Delta Y / \hat{Y}$ of Y due to relative change of $\Delta X_j / \hat{X}_j$, while the other variables remain constant. The regression study is usually associated with the coefficient of determination R^2 which is computed using the variance of the predicted values in relation to the variance of the observed ones. More specifically, if R^2 is close to 1 then the linear relationship is strong. The usefulness of the coefficient R^2 in sensitivity analysis is limited by the fact that when additional variables are added to the regression equation, the value of R^2 increases even when these variables do not significantly improve the regression equation (Janssen, 1994). Because of this, the adjusted R^2_{adj} is usually used instead:

$$R^2_{adj} = 1 - (1 - R^2) \times \frac{m - 1}{m - p - 1}, \quad (5)$$

where m is the model runs and p is the number of the input variables. The adjusted R^2_{adj} provides a measure of the extent to which the regression model can match the observed data and it is more appropriate for a multiple linear regression than the coefficient of determination R^2 (Janssen *et al.*, 1992; Janssen, 1994).

Linear correlation analysis sensitivity

The linear correlation coefficient (LCC) measures the strength of the linear relationship between the input variable X_j and the output Y; it can be written as:

$$\rho(X_j, Y) = \frac{\sum_{k=1}^{m} (X_{jk} - \bar{X}_j)(Y_k - \bar{Y})}{\sqrt{\sum_{k=1}^{m} (X_{jk} - \bar{X}_j)^2 \sum_{k=1}^{m} (Y_k - \bar{Y})^2}}, \quad (6)$$

where $\bar{X}_j = \sum_{k=1}^{m} X_{jk}/m$ and $\bar{Y} = \sum_{k=1}^{m} Y_k/m$. LCC is always ranged between −1 and +1, values close to either these extremes indicate a strong correlation, while small absolute values indicate little or no correlation.

Statistical performance indices

The quantitative evaluation of models is usually made through statistical performance analysis that is presented and widely used in many works. The statistical indices which we will use here are defined as (Hanna *et al.*, 1993; Chang & Hanna, 2004):

the fractional bias (FB) where the general expression is given by

$$FB = 2 \times \frac{\overline{C}_p - \overline{C}_o}{\overline{C}_p + \overline{C}_o}, \quad (7)$$

where C_p and C_o are model predictions and observations of pollen concentration; $\overline{C}$ is the average over the dataset. FB produces values ranged along a linear scale and the systematic bias refers to the arithmetic difference between C_p and C_o, advantages of FB is that it is a dimensionless number, symmetrical and bounded; the fractional bias values are ranged between −2.0 (extreme under-prediction) to +2.0 (extreme over-prediction).

Like FB, the geometric mean bias (MG) is a measure of mean bias and indicate only systematic errors which lead to always over-estimation or underestimation of the measured values. MG is defined as

$$MG = \exp{(\overline{\ln{(C_p)}} - \overline{\ln{(C_o)}})}, \quad (8)$$

It is clear to see that MG value greater than 1 implies that the model underestimates and an MG value less than 1 that the model overestimates.

The normalized mean square error (NMSE) and the geometric mean variance (VG) are measures of scatter and reflect both systematic and unsystematic (random) errors. NMSE is an estimator of the overall deviations between predicted and measured values.

$$NMSE = \frac{\overline{(C_o - C_p)^2}}{\overline{C}_o \times \overline{C}_p}, \quad (9)$$

The normalization by the product $\overline{C}_o \times \overline{C}_p$ ensures that the NMSE will not be biased towards models that over-predict or under-predict. Smaller values of NMSE denote better model performance, however high NMSE values do not necessarily mean

that a model is completely wrong. The VG is given by

$$VG = \exp \overline{((\ln(C_o) - \ln(C_p))^2)} \qquad (10)$$

According to Ahuja and Kumar (1996) and Kumar *et al.* (1993), the performance of a model can be deemed acceptable if: |FB| < 0.30, 0.75 < MG < 1.25, 1.00 ≤ VG < 4.00 and NMSE < 1.50.

A perfect model would have MG and VG = 1; and FB, and NMSE = 0. However, MG and VG are known to be strongly influenced by extremely high and low values, while FB and NMSE are more influenced by high observed and predicted concentrations (Chang and Hanna, 2004).

Methodology: sensitivity analysis and inferred pollen emission rates

The main objective of this paper is to infer the pollen emission rate from measurements and simulations of pollen deposition rates in a canopy gap for hybrid maize of 2.02 m height. The experiment was carried out unfortunately only over 1 day on 8 August 2008 on a 0.5 ha (120.4 × 41 m²) maize crop in Grignon, France (48°50′N, 1°56′E, 101 m a.s.l.) with a row spacing of 0.8 m and plant density of 9.5 plants/m², however, the developed methodology should be applied for multiple days. Measurements were done when released pollens by plants were high and the leaf area density is expressed in horizontal and vertical projections where ranges of variability are defined below. The ground deposition rate was recorded in a canopy gap of 2.4 m diameter at seven distances (0.0, 0.4, 0.8, 1.2, 1.6, 2.0 and 2.4 m) on a line parallel to the row and to the wind direction, starting where the airflow entered the gap and ending at the downwind edge of the plants encircling the gap. A container was placed at the height of the tassel base in the first inter row of maize upwind of the canopy gap to measure pollen emission rate. These measurements will be considered to provide good approximations of the real pollen emission rates. Measurements were recorded during 30 min intervals with six intervals recorded between 10:15 and 13:15 UTC where the maximum pollen emission rate typically occurs (Scott, 1970; Gregory, 1973).

Deposition rates were measured using containers (diameter = 50 mm, height = 70 mm) filled with 30 ml of an electrolyte solution (Coulter Isoton, Beckman USA) and the amount of pollen in the containers were counted by using an automatic counter (Coulter Multisizer III, Beckman, USA). The wind speed and global radiation were recorded at 5 m above the canopy while relative humidity, rainfall, water vapour pressure, water vapour pressure deficit, presence of wetness and temperature were recorded at plant height. Each meteorological variable was averaged over successive 30 min intervals, as for the deposition measurements.

To start, sensitive analysis of the pollen deposition in the canopy gap according to seven input variables of SMOP was performed. These variables describe pollen and canopy characteristics as well as pollen interactions with its environment. There are four continuous variables: two variables of the Gaussian distribution of pollen settling velocity (mean, *meanW$_s$*, ranged from 0.15 m/s, which corresponds to dry pollen to 0.35 m/s, which corresponds to fresh pollen and standard deviation, *StdW$_S$*, ranged from 0.01 to 0.06 m/s) (Aylor, 2002; Loubet *et al.*, 2007; Marceau *et al.*, 2012); horizontal and vertical projections of the leaf area density LAD_x and LAD_z, respectively (LAD_x ranged from 0 to 0.55 m²/m³ and LAD_z ranged from 0 to 0.65 m²/m³, Drouet (2003)). More specifically, LAD_x influences deposition due to the gravitational settling, and for

example, $LAD_x = 0$ m²/m³ means that no deposition occurs within the canopy due to the gravitational settling. While LAD_z influences deposition due to the inertial impaction, $LAD_z = 0$ m²/m³ means that no deposition occurs within the canopy due to the inertial impaction. In addition, there are three 2-level variables (Yes or No): deposition on leaves (D_f); resuspension of pollen on vegetation and ground (R_p); and horizontal pollen fluctuations due to the wind (1D). In case when no deposition on leaves will be considered ($D_f = 1$) we impose the probability of deposition trajectories equal to zero. A detailed description of the three 2-level variables D_f, R_p and 1D as well as a description of their eight possible combinations are given in Table 1.

Due to the difficulty to describe correctly the turbulence flow in the canopy gap, as we detail in the Discussion section, the following SMOP parameters were estimated using meteorological data; wind speed (4.06 m/s), canopy height (2.02 m), σ_u/u_* and σ_w/u_* above the canopy were set to 2.6 and 1.25, respectively, neutral conditions were assumed. The crop length upwind and downwind of the gap (105 and 13 m) and the assumed released pollen rate to the atmosphere was set at 100 grains/m²/s. Figure 1 describes the geometric configuration of the experimental study.

The output of the model considered here is the pollen deposition rate in the centre of the canopy gap which is a typical practice configuration for measurements of pollen deposition on the ground required to infer the pollen emission rate. To do this and to examine how the uncertainty in the variable estimates affects the deposition rate, the Latin hypercube sampling is used. This technique, as noted above, allows exploration of the full variable space with a greatly reduced number of model runs (Iman, 1992). Therefore, only 105 sets of variables are created to run the model.

In the first approach, the sensitivity methods described in the previous subsections were applied independently to the seven variables. As a second approach, the three 2-level variables were replaced by a unique discrete variable that has been named hereafter as *interaction* representing the eight combinations of the three 2-level variables. To finish, for a set of SMOP parameters, inferences of pollen emission rate were obtained corresponding to the time intervals where pollen deposition rates were measured. Pollen emission rate was inferred by assuming that the ratio between the simulated pollen deposition rate ($D_{simulated}$) and the measured pollen deposition rate (D_{real}) is equal to the ratio between the simulated pollen release rate ($S_{simulated}$ = 100 grains/m²/s) and the inferred pollen emission rate (S_{real}).

$$\frac{S_{simulated}}{S_{real}} = \frac{D_{simulated}}{D_{real}} \Leftrightarrow S_{real} = S_{simulated} \times \frac{D_{real}}{D_{simulated}} \qquad (11)$$

For a given time interval, the pollen release rate is inferred from the measured deposition rate at seven distances in the canopy gap following Eqn (11). These inferred estimations are used to compute an average and a standard error of pollen emission rate. To assess the influence of SMOP input variables on the inferred pollen release rate, SA is done using Eqn (11) for different interactions.

Results

In the current work, the sensitivity analysis was done according to the following seven variables described previously in the methodology section (*meanW$_S$*, *StdW$_S$*, LAD_x, LAD_z, D_f, R_p, 1D). Results of general regression and correlation analysis are shown in

Table 1. Description of the 2-level parameters which constitute the eight interactions used to conduct a sensitivity analysis of the model with a description of each of the eight possible combinations of the three variables D_f, R_p and $1D$

Variable	No (0)	Yes (1)
D_f	Pollen deposition on leaves and pollen impacted on stems	No pollen deposition on leaves and no pollen impacted on stems
R_p	Pollen rebound: 75% of impacted pollen return in the air + pollen rebound on the ground in the function of pollen velocity	No pollen rebound on leaves and ground among the 75% of the impacted pollen
$1D$	Pollen trajectories are sensitive to both horizontal and vertical fluctuations due to the turbulence	Pollen trajectories are sensitive only to vertical fluctuations due to the turbulence
Combination of D_f R_p $1D$	Description of the included model interaction	
000	Deposition on leaves, 75% of the impacted pollen on leaves and ground rebound, and taking into account horizontal and vertical pollen fluctuations due to the turbulence	
001	Deposition on leaves, 75% of the impacted pollen on leaves and ground rebound, and taking into account only vertical pollen fluctuations due to the turbulence	
010	Deposition on leaves, 0% of the impacted pollen on leaves and ground rebound, and taking into account horizontal and vertical pollen fluctuations due to the turbulence	
011	Deposition on leaves, 0% of the impacted pollen on leaves and ground rebound, and taking into account only vertical pollen fluctuations due to the turbulence	
100	No deposition on leaves, 75% of the impacted pollen on the ground rebound, and taking into account horizontal and vertical pollen fluctuations due to the turbulence	
101	No deposition on leaves, 75% of the impacted pollen on the ground rebound, and taking into account only vertical pollen fluctuations due to the turbulence	
110	No deposition on leaves, 0% of the impacted pollen on the ground rebound, and taking into account horizontal and vertical pollen fluctuations due to the turbulence	
111	No deposition on leaves, 0% of the impacted pollen on the ground rebound, and taking into account only vertical pollen fluctuations due to the turbulence	
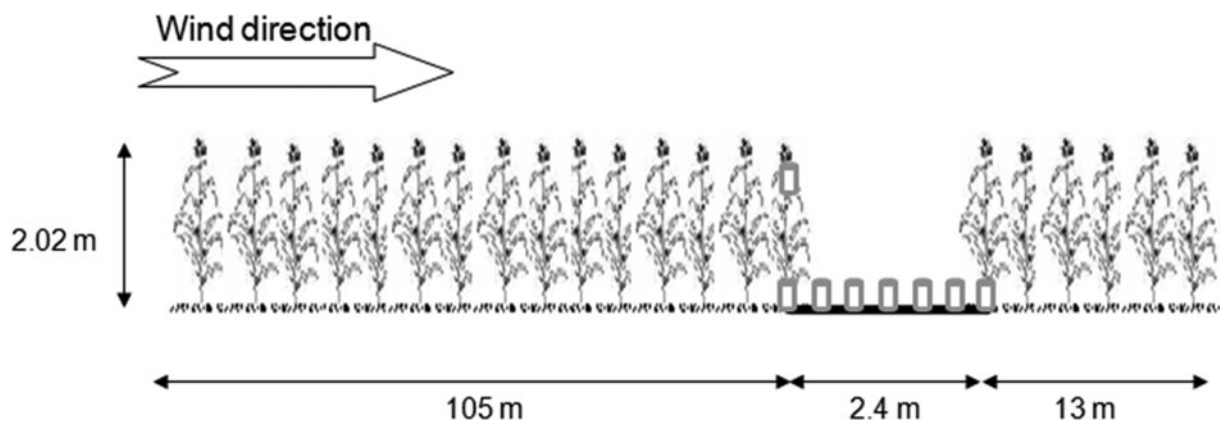
Fig. 1. Configuration of the computational domain with the mean wind direction and canopy dimensions. Seven containers were placed in the canopy gap at ground level, on a line parallel to the wind direction. Another container was placed in the first inter-row of maize, upwind of the canopy gap and at the height of the tassel base.

Table 2. For this general analysis, SRC method shows that the pollen deposition is mainly affected by $1D$, D_f and $meanW_S$ variables, while NRC shows that the more important variables are D_f, LAD_z and $meanW_S$. Results given by LCC method show the linear relationship between the input variables and the output pollen deposition. LCC method reveals that the output is mainly affected by $1D$, D_f, R_p and $meanW_S$, and is weakly dependent on LAD_z. Because of the difficulties to assess accurately the impact of each 2-level variable and to take into account interactions between the three 2-level variables a unique discrete variable was created corresponding to a combination of the three variables.

Interaction values are interpreted as follows. The first character corresponds to the variable D_f; the second character corresponds to the variable R_p; and the third character corresponds to the variable $1D$ (e.g., interaction 010 means that the model includes deposition on leaves, 0% of the impacted pollen on leaves and ground rebound, and taking into account horizontal and vertical pollen fluctuations due to the turbulence).

The Boxplot (Fig. 2) shows clearly the effect of D_f variable on the simulated pollen deposition rate for the eight interactions. Furthermore, interactions when deposition on leaves is taken into account lead to much lower levels of the simulated pollen
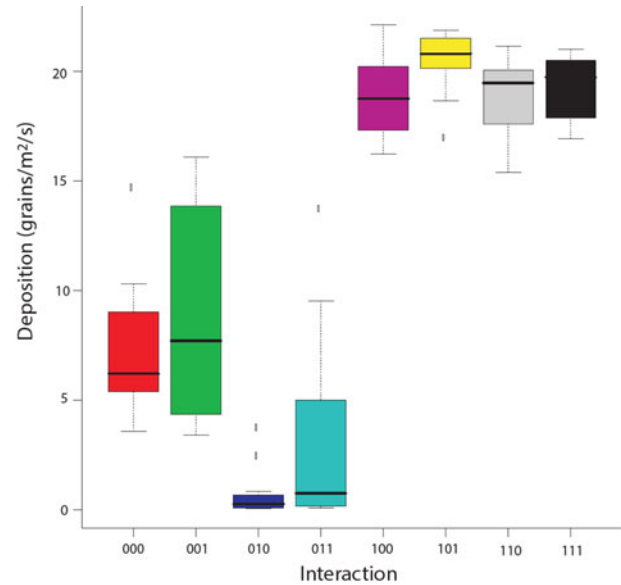
Table 2. General regression and correlation analysis: linear correlation coefficient (LCC), standardized regression coefficients (SRC) and normalized regression coefficients (NRC) are computed using ordinary regression coefficients (ORC) and equations given in the section on sensitivity analysis methods

| Variable	ORC	SRC	NRC	LCC
$meanW_S$	9.89	0.27	0.20	0.55
$StdW_S$	14.50	0.07	0.04	−0.16
LAD_x	−2.95	−0.06	−0.07	−0.14
LAD_z	−8.54	−0.10	−0.23	−0.22
D_f	14.35	0.40	0.59	0.69
R_p	−3.09	−0.10	−0.12	0.59
$1D$	1.19	0.66	0.05	−0.85

Fig. 2. Boxplots of 105 model run representing the simulated pollen deposition rate (grains/m²/s) in the canopy gap and classed per interaction. The 105 sets of variables are divided per interaction (defined in Table 1) as: 000: 14 model runs; 001: 12 model runs; 010: 11 model runs; 011: 15 model runs; 100: 13 model runs; 101: 14 model runs; 110: 14 model runs; 111: 12 model runs. The endpoints of the box are formed by the lower and upper quartiles of the data, that is, X0.25 and X0.75. The line within the box represents the median, X0.50. The bar above the box extends to the minimum of X0.75 + 1.5(X0.75- X0.25) and the maximum value. The bar at the bottom of the box extends to the maximum of X0.25–1.5(X0.75- X0.25) and the minimum value. The observations falling outside of these bars are shown in marks.

deposition rates. Also, the variance of these deposition rates is more important when horizontal fluctuation is not taken into account. Deposition corresponding to interactions 010 and 011 are lower because there is deposition on leaves and no resuspension inducing strong retention of pollen by plants which shows the influence of the R_p variable when deposition on leaves is taken into account.

To explain the variation of deposition observed for each interaction, regression and correlation analysis were conducted on the four remaining continuous variables. Corresponding results are shown in Figs 3 and 4, and Table 3. The scatterplots (Fig. 3) and the adjusted R_{adj}^2 coefficient of determination (Table 3) show that the SMOP-modelled pollen deposition rate at the seven positions in the canopy gap may be estimated with a linear regression model except for interaction 100 ($R_{adj}^2 < 0.6$) for which the model is then non-linear. Table 3 shows that variation in $meanW_S$ has a significant effect on outputs for models in six out of eight interaction cases (with the exception of 010 and 011). NRC and ORC are positive which means an increase of $meanW_S$ leads to an increasing of deposition in the gap. The variation in LAD_z has a significant effect on outputs for models in three interactions (000, 001 and 011), while variation in LAD_x has a significant effect on outputs for models in two interactions (000 and 001). Variation in $StdW_S$ is only significant for interaction 010. The correlation analysis results shown in Fig. 4 are similar to regression analysis, taking into account the effects of interactions. When no deposition on leaves is taken into account, the $meanW_S$ variable is the main correlated variable with deposition rate; while LAD_z variable is more correlated when deposition on leaves is taken into account.

The quality and performance of models are usually presented by drawing a scatter diagram using predicted and observed values (when observations are available). Observed deposition rates were calculated from estimations of the number of pollen grains collected in each container and time of exposure. Figure 5 shows the measured deposition rates as a function of downwind distance in the canopy gap at different times. The figure reveals that deposition rates are approximately unchangeable with distance in the canopy gap moreover the standard deviation of pollen deposition rates are ranged from 60 grains/m²/s at 10:15 UTC to 15 grains/m²/s at 12:45 UTC, while the mean of deposition rates are ranged from 524 grains/m²/s at 10:15 UTC to 84 grains/m²/s at 12:45 UTC. Figure 6 represents a scatter diagram between measured and inferred pollen emission rates for interactions 110 and 111

which seem to be good parameterizations of the model as shown in Table 4. The figure provides an immediate visualization of the overall model performance. This direct visual inspection shows that the inferred pollen emission rate tended to be slightly larger than those measured. This is due to several factors as explained in the discussion section. In fact, most of the points are predicted within a factor of two, the model appears then to have a good performance qualitatively for these two interactions. This is re-affirmed from the values of statistical performance measures given in Table 4, which provides a quantitative evaluation of the model including the parameters FB, MG, NMSE and VG to specify which of the eight 2-level interactions gives good results.

Figure 7 shows the measured and inferred emission rates using the SMOP model for interactions 001, 110 and 111 as a function of time. Measured and inferred emission rates decreased with time and the inferred ones are higher than the measured ones. These profiles show the quality of the model parameterization for interaction 111. Furthermore, measured rates decrease with time and are ranged at 10:15 UTC from 590 grains/m²/s at 0.8 m to 438 grains/m²/s at 2.0 m, and at 12:45 UTC from 107 grains/m²/s at 1.6 m to 60 grains/m²/s at 2.4 m. However, the inferred emission rates, for interaction 110 (resp. 111), are ranged from 1022 to 615 grains/m²/s (resp. from 792 to 536 grains/m²/s) at 10:15 UTC, and from 224 to 138 grains/m²/s (resp. from 169 to 104 grains/m²/s) at 12:45 UTC. These profiles show also the quality of the model parameterization for interaction 111. Although the inferred pollen rates are higher than the measured ones, these results are similar to those found by Jarosz *et al.* (2005) and Marceau *et al.* (2012).

Table 4 summarizes the SMOP model performance: all statistical measures presented in this table are based on measured *v*.
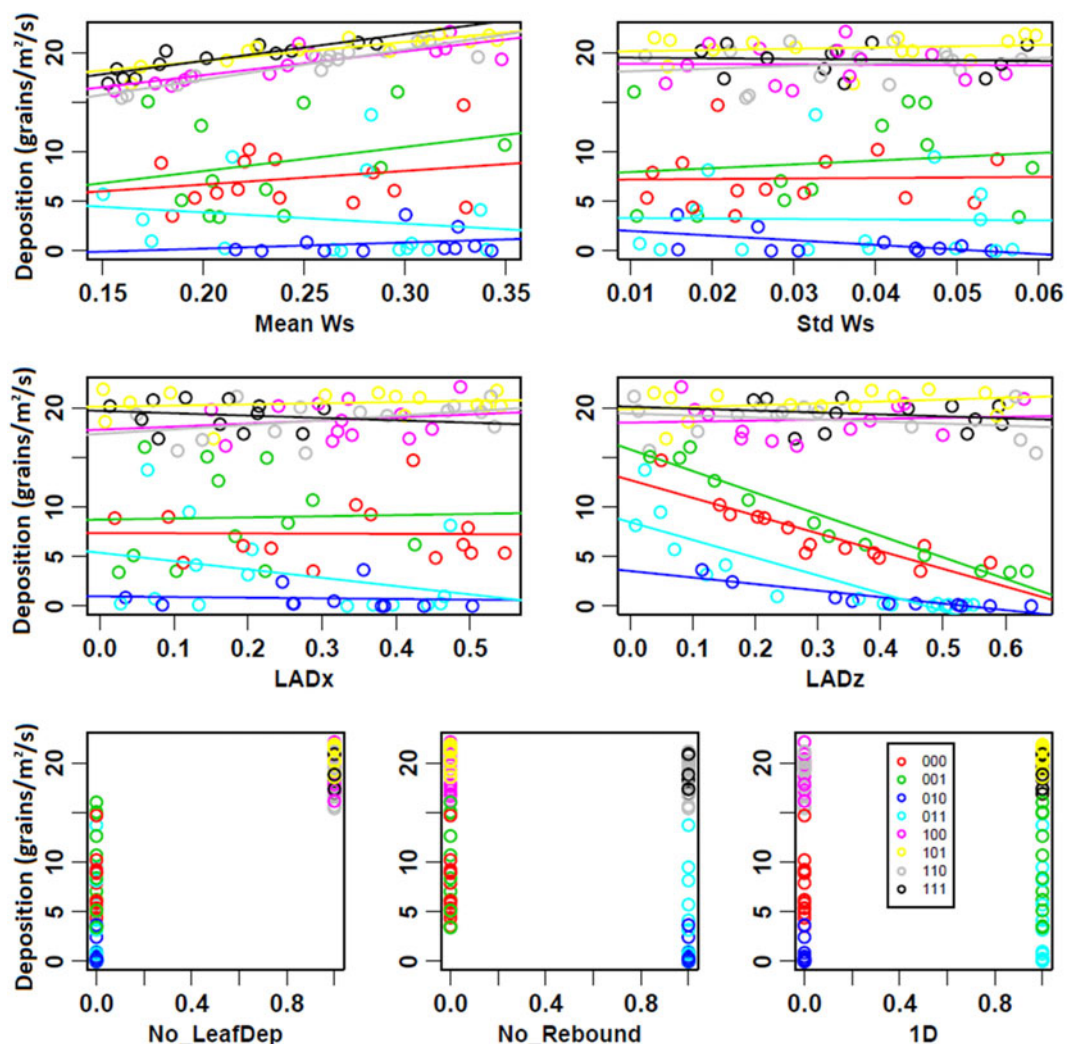
Fig. 3. Simulated pollen deposition rate (grains/m²/s) in the canopy gap. The assumed pollen release rate to the atmosphere is 100 grains/m²/s. This scatterplot represents 105 model runs classed per interaction using a colour for each one. Lines represent the linear regression.

inferred pollen deposition rates, the calculated values are mostly within the range of acceptable model performance for interactions 110 and 111 corresponding to the fact that the model includes no deposition on leaves, no resuspension, and taking (or not) into account horizontal pollen fluctuations. And as presented in Fig. 7, these two interactions are used to infer maize pollen emission rates. For the remaining interactions, the results are out of the desired ranges.

Discussion

The SMOP model has been evaluated in many experimental studies with various hybrids, for example, studies conducted in France at Montargis and Grignon in 2000–2002, 2008 and 2009 (Jarosz et al., 2004, 2005; Marceau et al., 2011, 2012). During these experiments, airborne concentration and deposition rates of maize pollen were measured at several locations within and downwind from various maize fields. All of these validations were done without a detailed sensitivity analysis of the model, such as that conducted in the current work, to pinpoint uncertainties in maize pollen deposition rates as well as in inference of pollen emission rates. Note that Jarosz et al. (2004) and Marceau et al. (2012) carried out an SA of the model but only to the settling velocity, and showed that measured and modelled deposition rates performed at downwind distance $x = 3$ m and $x = 10$ m were very sensitive to this uncertain variable.

Using the surface wetness index measurements Jarosz et al. (2003) and Marceau et al. (2012) have shown that the start of the pollen release in the morning appeared to coincide with the drying of the crop, and the largest concentration and deposition usually occurs at around 10:00 UTC (Scott, 1970; Gregory, 1973; Marceau et al., 2011) which is the time when our experiment was done. Furthermore, Marceau et al. (2011) and Sofiev et al. (2013) have explored and quantified the effects of both meteorological variables and variety on pollen emission. Their data analysis revealed the effects of temperature and relative humidity on both diurnal and seasonal patterns of emission. The rate of pollen emission was higher when temperatures were high and lower when the humidity was high. In addition, they have shown that the pollen emission rate is higher when the wind speed is high. However, direct measurement of pollen release rate is not easy to quantify as this flux remains difficult to estimate. Jarosz et al. (2003) and Marceau et al. (2011) inferred
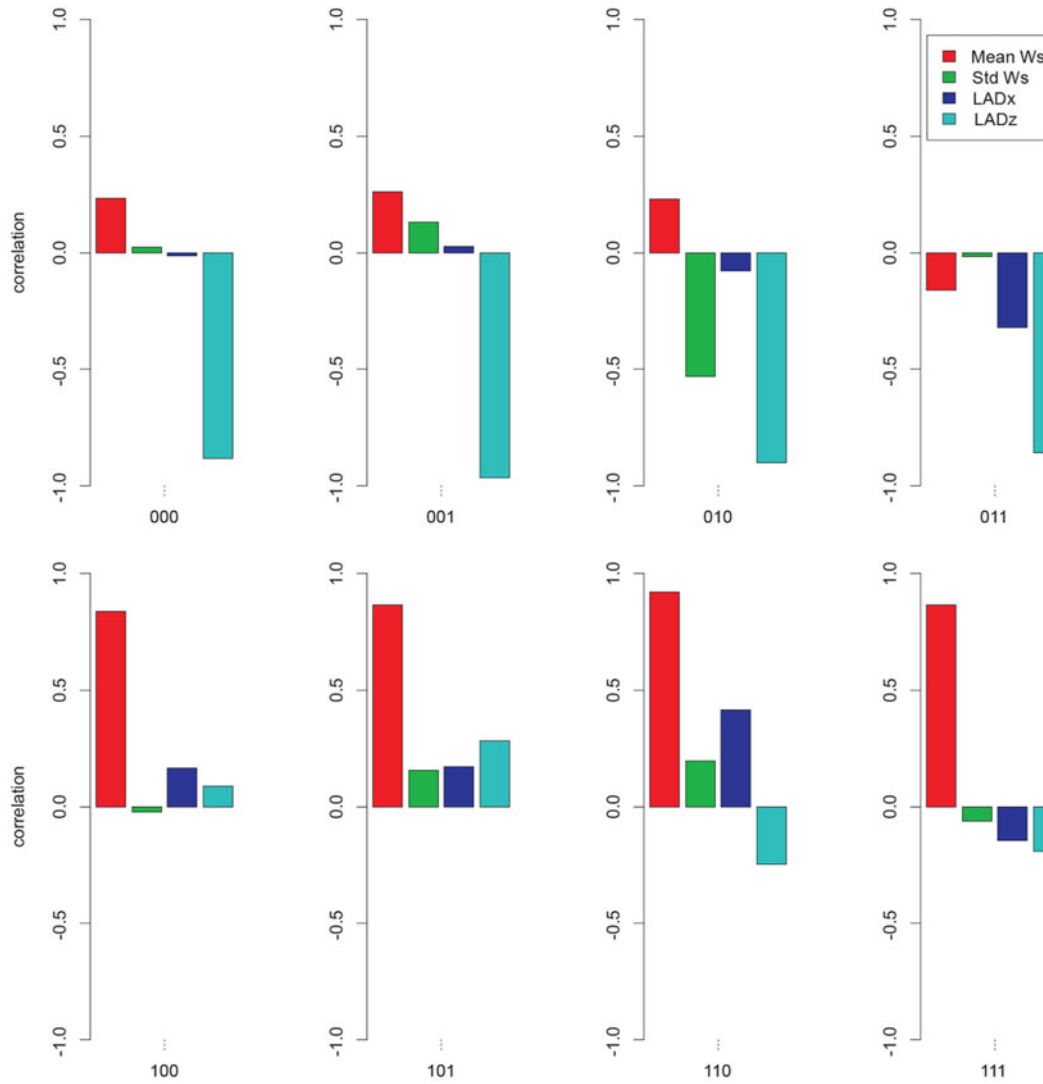
Fig. 4. Linear correlation coefficients for each interaction, between simulated deposition rate in the canopy gap and the four continuous variables: mean of pollen settling velocity, *meanW$_S$* (m/s); standard deviation, *StdW$_S$* (m/s); horizontal projection of the leaf area density, *LAD$_x$* (m²/m³); and vertical projection, *LAD$_z$* (m²/m³).

Table 3. Summary of regression analysis classed per interaction: ordinary regression coefficients (ORC) and normalized regression coefficients (NRC)

Variable	meanWs		StdW$_S$		LADx		LADz		
Interaction	ORC	NRC	ORC	NRC	ORC	NRC	ORC	NRC	R^2_{adj}
000	22.90*	0.76	5.58	0.02	−3.87*	−0.17	−19.95*	−0.82	0.95
001	7.58*	0.20	21.46	0.09	7.39*	0.15	−22.39*	−0.79	0.94
010	4.64	1.70	−32.75*	−1.50	−6.01	−0.59	−12.12	−3.01	0.87
011	−31.20	−2.54	−63.42	−0.72	−2.70	−0.22	−21.30*	−1.96	0.77
100	24.55*	0.32	−13.25	−0.02	0.19	0.01	1.26	0.02	0.58
101	20.52*	0.27	23.84	0.04	0.68	0.01	2.30	0.03	0.76
110	33.26*	0.44	4.95	0.01	1.86	0.03	−0.81	−0.01	0.82
111	30.92*	0.33	−9.80	−0.18	−4.48	−0.04	−1.99	−0.04	0.78

*indicates the most significant parameters at the 1% level on the basis of the *T*-test statistic. The last column shows the adjusted R^2_{adj} computed per interaction.

the release rate by combining pollen measured directly on the tassel and deposition measured above the source using the slope of the linear regression.

In this study, the inferred pollen emission rate is calculated using Eqn (11) for only two validated interactions 110 and 111. However, before conducting this inference, we have to inspect
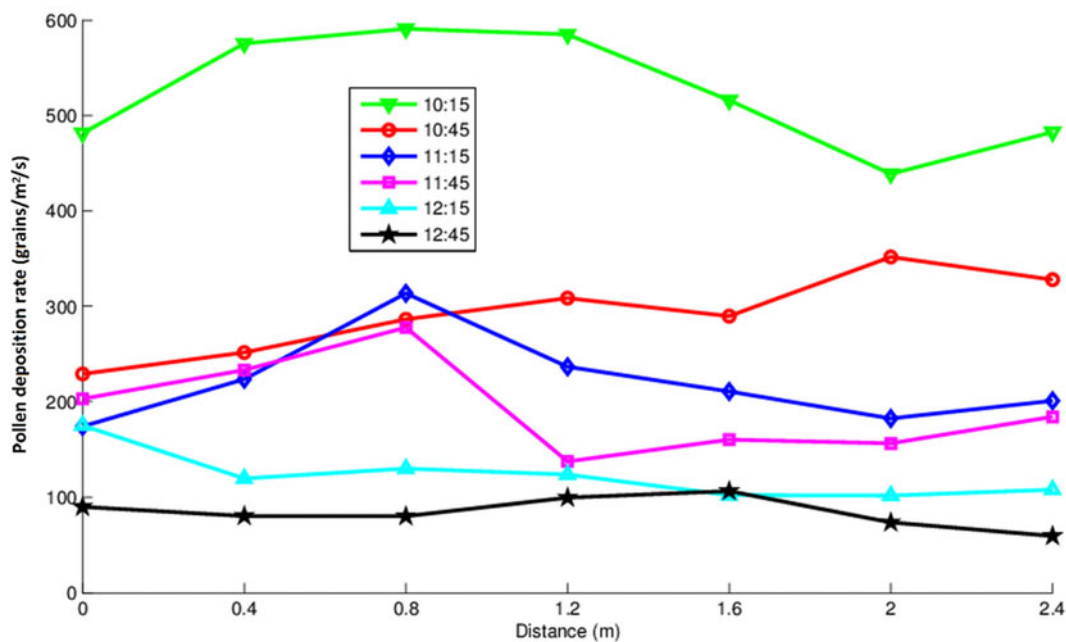
Fig. 5. Measured deposition rate (grains/m²/s) as a function of downwind distance in the canopy gap at different time. Measurements were recorded during 30-min intervals with six intervals recorded between 10:15 and 13:15 UTC and were calculated from estimations of the number of pollen grains collected in each container and time of exposure.
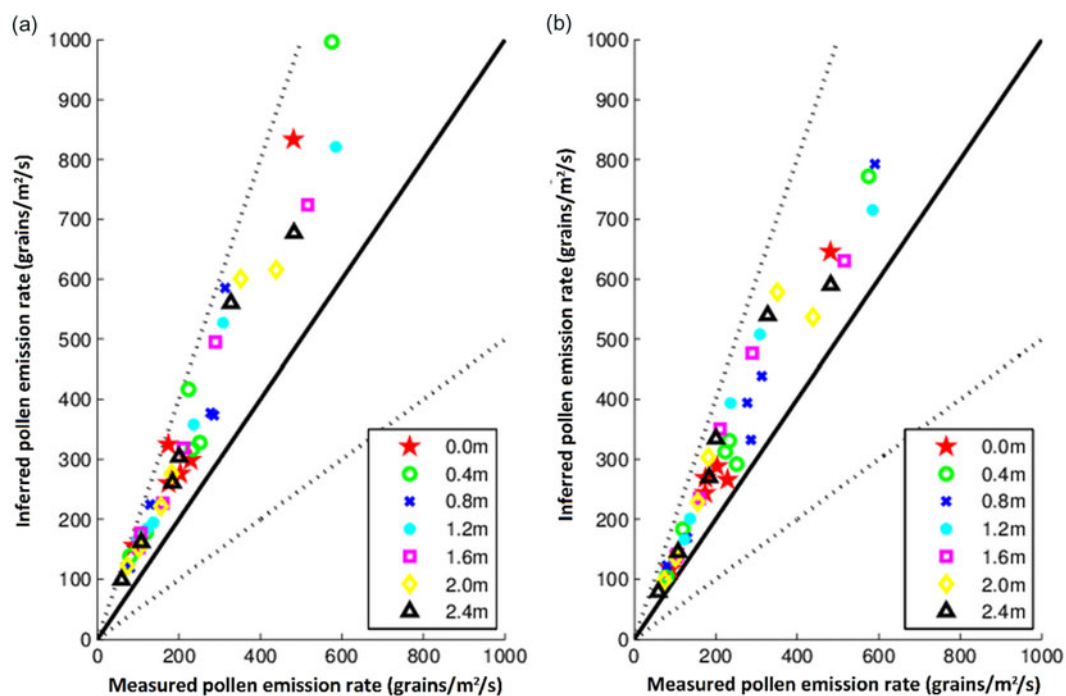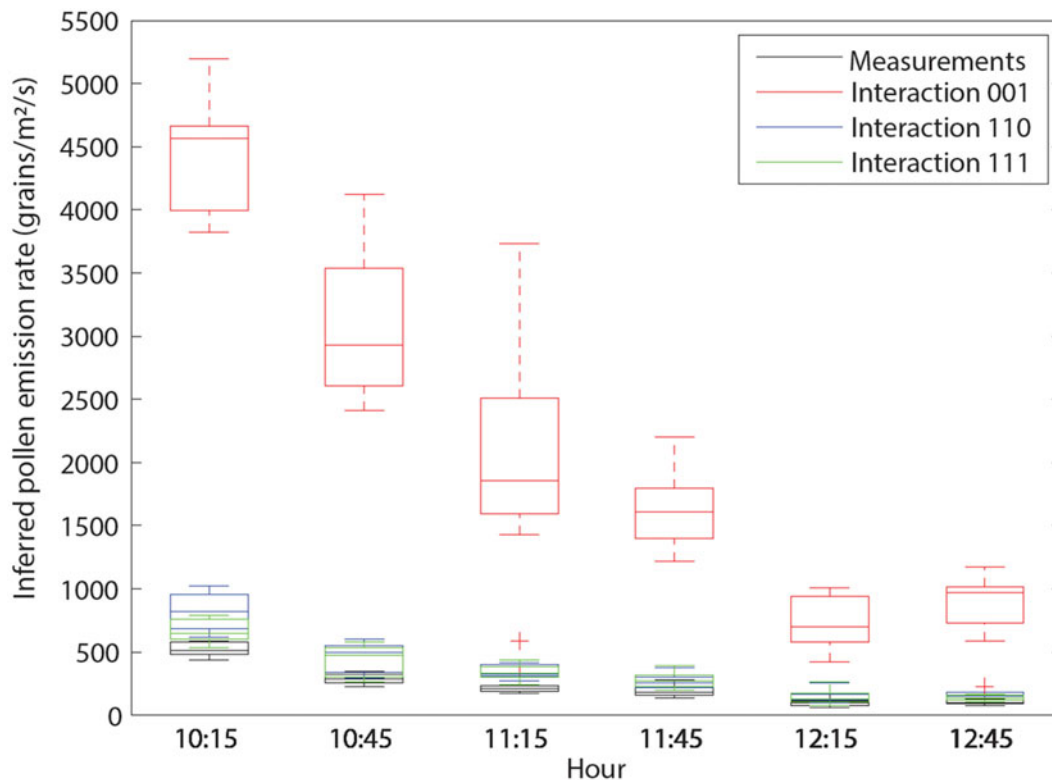

Fig. 6. Measured pollen emission rate versus inferred pollen emission rate (grains/m²/s) using the model at seven distances in the canopy gap showed only for (a) interaction 110 and (b) interaction 111 as defined in Table 1. The solid line is the one-to-one between observed and inferred pollen emission rates whereas the dotted lines correspond to factor of two (i.e., $y = 0.5 \times x$ and $y = 2 \times x$).

the influence of environmental variables on the simulated pollen deposition rates by using a global sensitivity analysis. Table 2 shows that increasing the $meanW_S$ seems to increase the pollen deposition rate in the canopy gap; this may be explained by the fact that the horizontal wind speed in the canopy gap is low, and inertia and crossing trajectories effect of heavy particles is

small (Snyder and Lumley, 1971; Reynolds, 2000). The same result is also shown for detailed sensitivity analysis by including interactions (Fig. 3 and Table 3). Nevertheless, turbulence intensity increases immediately downwind of a roughness change (Gash, 1986) and, in particular, much more in a canopy gap. One can see that deposition rates decrease when the leaf area

Table 4. Summary of the four statistical performance parameters: the fractional bias (FB), the geometric mean bias (MG), the normalized mean square error (NMSE) and the geometric mean variance (VG) classed per interaction as defined in Table 1

Statistical parameters	000	001	010	011	100	101	110	111	Performance measures Chang and Hanna (2004)
FB	0.97	1.42	0.91	1.30	0.43	0.38	0.02	0.09	|FB| < 0.30
MG	2.92	6.19	2.87	4.97	1.57	1.44	1.02	1.15	0.75 < MG < 1.25
NMSE	1.23	4.07	1.26	2.86	0.36	0.40	0.13	0.08	NMSE < 1.50
VG	3.26	7.34	5.04	8.88	2.20	1.65	1.20	1.27	$1.00 \leq VG < 4.00$

Fig. 7. Measured and inferred pollen emission rates (grains/m²/s) estimated using the restricted setting where our model includes no deposition on leaves, no resuspension and either taking both horizontal and vertical pollen fluctuations into account or only the vertical, which correspond to 110 and 111 interactions versus the interaction 001 as defined in Table 1.

density, which is expressed in horizontal and vertical projections, increases. Indeed, when modelling pollen trajectories, not all of the pollen deposited on leaves and ground rebound. The global regression shows that SRC, NRC and LCC methods provide a different order of the more influenced variables on the simulated deposition rates (Table 2). This may be due to the discrete variables as well as to the variances of the inputs and the output pollen deposition rate, for example, the variance of LAD_z is equal to approximately 169 times the variance of $StdW_S$; and the variance of LAD_x is equal to approximately 8 times the variance of $meanW_S$. More accurately, SRC and NRC coefficients are related as follows.

$$b_j^{(s)} = b_j^{(n)} \cdot \frac{V_{Xj}}{V_Y},$$

where V_{Xj} and V_Y are the coefficients of variation for X_j and Y, respectively. This global sensitivity analysis showed that the real

influence of each of the seven input variables on deposition rates cannot be concluded.

Jarosz *et al.* (2004) have shown that for larger settling velocity the model simulates the deposition rates better near the source but does not simulate the concentration profiles correctly, especially at $x = 10$ m, where measured concentrations are greatly underestimated for large $meanW_S$. These results suggest that discrepancies in concentration and deposition reported by the model near the source are unlikely to be solely due to pollen settling velocity. To better understand this discrepancy and since no definite results could be concluded from the global sensitivity analysis, eight combinations of the three 2-level variables were introduced. By considering these combinations, the SMOP model may be represented as a linear model, which is established by the coefficient of determination R^2_{adj} (Table 3). Simulated pollen deposition rates are given in Figs 2 and 3, and the calculated values tended to be in the same range as those reported by Aylor *et al.* (2006). Figure 2 shows great variability of pollen deposition; furthermore,

when the canopy and ground are perfect sinks (i.e., no pollen rebound on leaves and ground among 75% of the impacted pollen) and horizontal fluctuations are accounted for, the amount of pollen reaching the deposition cup will decrease almost to zero (deposition near zero). One could interpret that an effect of an increase of probability of impaction is due to increased particle horizontal velocity. This also explains the larger correlation coefficient of LAD_z in Fig. 4 for 001 when compared to 011. Meanwhile, the interactions 100 and 101 show positive correlations between deposition rate and LAD_z while all other combinations show negative correlations (Fig. 4). This is probably because the 100 and 101 interactions correspond to maximum rebound on plants and ground while 110 and 111 have rebound only on the ground. This is the only situation for which the increase of LAD_z show an effect via lowering the wind speed profile in the canopy (which depends on LAD_z), while in any other case, the deposition by impaction dominates. In the literature, little information is known about pollen resuspension from either leaves or the ground. Aylor *et al.* (2003) have shown experimentally that pollen could be quite easily dislodged from maize leaves by either leaf shaking, roll-off or small wind speed (0.2–0.5 m/s). However, in the current experiment, the wind speed was fixed at 4.06 m/s which explain the neglected effect of $1D$ variable on the deposition rate.

Daily measured pollen emission rates and therefore deposition rates depend on the local weather conditions (e.g., Scott, 1970; Gregory, 1973; Marceau *et al.*, 2011). This explains the results given in Fig. 5: measured deposition rates are approximately unchangeable with distance in the canopy gap and do not decrease as is the case when we consider deposition rates downwind the source (Jarosz *et al.*, 2003; Marceau *et al.*, 2011; Torimaru *et al.*, 2012). This result is certainly due to the turbulence intensity in the canopy gap and may explain the discrepancy between measured and simulated pollen deposition rates (Figs 6 and 7). Indeed, the empirical parameterization of turbulence field in the transition zone used is not the correct one and future versions of the model should take this conclusion into account.

Variations in input variables (Fig. 7) have a large effect on the inferred pollen emission rates. Dupont *et al.* (2006) showed that a simplified parameterization of turbulence in the area between two different canopies leads to underestimates of pollen deposition close to the field. However, the current work found the opposite, i.e. the simulated pollen deposition rates overestimate the measured ones (Figs 6 and 7). The meteorological conditions during these experiments were not reported in sufficient detail to draw definite conclusions about quantities of maize pollen released and distance of pollen dispersal in the canopy gap. Various assumptions are cited below to explain this overestimation:

(1) the surface is dynamically heterogeneous that could perhaps lead to overestimation of the turbulence intensity in the canopy gap (Flesch *et al.*, 1995): larger turbulence intensity would induce larger vertical diffusion, which favours larger deposition rates near the source. Indeed, it is difficult to describe correctly the turbulence flow in the transition zone at the downwind edge of the source and in particular in a canopy gap (Gash, 1986; Heisler and DeWalle, 1988; McCartney and Lacey, 1991). They showed that, immediately downwind of a roughness change, the turbulence intensity increases three times over normal conditions. By taking into account these results, the SMOP model includes an empirical parameterization of the turbulence field for heterogeneous

landscapes. However, this parameterization, which was also used in the current simulation, is possibly not valid for the canopy gap that is established by an approximately constant deposition rate with distance (Fig. 5),

(2) by considering interactions 110 and 111, i.e. where there is no deposition on leaves and no rebound of pollen, the variable $meanW_S$ is the only significant variable (Fig. 3); therefore, lower $meanW_S$ leads to less deposition. The settling velocity, as well as many of the biophysical characteristics of pollen, is mainly determined by pollen water content (Di-Giovanni *et al.*, 1995; Aylor, 2003; Loubet *et al.*, 2007; Marceau *et al.*, 2012). Indeed, pollen water content affects pollen mass, diameter and density which determine pollen settling velocity as formulated by Di-Giovanni *et al.* (1995). In the current work, the simple model proposed by Marceau *et al.* (2012) was used to predict accurately pollen water content at emission as a function of air vapour pressure deficit. In our opinion, this simple model computes with uncertainty pollen water content, which affects the pollen settling velocity. Hence, the overestimation of the deposition rate would be linked to an erroneous parameterization of intervals of the two variables describing the Gaussian distribution of pollen settling velocity,

(3) biased emission and deposition rates measurements; the cups used to collect the pollen grains influence their deposition on the ground. McCartney and Lacey (1991) showed experimentally that numbers of spores collected on horizontal microscope slides openly exposed were almost double to those on slides placed on a table or at the bottom of large cups. Hence, collecting pollen in containers could have led to an underestimation of the deposition rate in the current work (McCartney *et al.*, 1991), and

(4) the quantity of pollen deposited in the source itself was estimated to range from 17 to 50% (Jarosz *et al.*, 2004). Then, by taking into consideration interactions 110 and 111 where there is no pollen deposition on leaves, the simulated deposition rates automatically overestimate the measured deposition ones in the canopy gap.

The real pollen emission rate at a given time is always difficult to determine. To this end, one should use the inference methodology set out in the current paper for calibration of the parameters by combining the Lagrangian stochastic model with a high-resolution version of the meteorological flow model, if they have to be used in complex experiments. Without any calibration of the model as detailed in this work, the inference of pollen emission rate appears to be uncertain.

Conclusion

The current study has shown that the inferred emission rate overestimates the measured quantities. Sensitivity Analysis has shown that simulated pollen deposition rates are mainly influenced by and correlated with the mean gravitational settling velocity and to the leaf area density. The overestimated emission rate could possibly be as a result of (i) biased emission and deposition rates measurements, (ii) gravitational settling velocity of pollens, (iii) pollen deposition on leaves or pollen resuspension from the leaves or the ground. However, the model overestimation of the emission rate could also be due to a parameterization of the turbulent field in the canopy gap.

References

Ahuja S and Kumar A (1996) Evaluation of MESOPUFF-II SOx transport and deposition in the Great Lakes Region. *AWMA Speciality Conference on Atmospheric Deposition to the Great Lakes, V IP-72, October 28–30,* pp. 283–299.

Arritt RW, Clark CA, Goggi AS, Sanchez HL, Westgate ME and Riese JM (2007) Lagrangian numerical simulations of canopy air flow effects on maize pollen dispersal. *Field Crops Research* **102**, 151–162.

Astini JP, Fonseca JP, Clark C, Lizaso J, Grass L, Westgate M and Arritt R (2009) Predicting out crossing in maize hybrid seed production. *Agronomy Journal* **101**, 373–380.

Aylor DE (2002) Settling speed of corn (*Zea mays*) pollen. *Journal of Aerosol Science* **33**, 1601–1607.

Aylor DE (2003) Rate of dehydration of corn (Zea mays L.) pollen in the air. *Journal of Experimental Botany* **54**, 2307–2312.

Aylor DE, Schultes NP and Shields EJ (2003) An aerobiological framework for assessing cross-pollination in maize. *Agricultural and Forest Meteorology* **119**, 111–129.

Aylor DE, Boehm MT and Shields EJ (2006) Quantifying aerial concentrations of maize pollen in the atmospheric surface layer using remote-piloted airplanes and Lagrangian stochastic modelling. *Journal of Applied Meteorology and Climatology* **45**, 1003–1015.

Baklanov A, Smith Korsholm U, Nuterman R, Mahura A, Nielsen KP, Sass BH, Rasmussen A, Zakey A, Kaas E, Kurganskiy A, Sørensen B and González-Aparicio I (2017) Enviro-HIRLAM online integrated meteorology-chemistry modelling system: strategy, methodology, developments and applications. *Geoscientific Model Development* **10**, 2971–2999.

Chang JC and Hanna SR (2004) Air quality model performance evaluation. *Journal of Meteorology and Atmospheric Physics* **87**, 167–196.

Di-Giovanni F, Kevan PG and Nasr ME (1995) The variability in settling velocities of some pollen and spores. *Grana* **34**, 39–44.

Dietiker D, Stamp P and Eugster W (2011) Predicting seed admixture in maize combining flowering characteristics and a Lagrangian stochastic dispersion model. *Field Crops Research* **121**, 256–267.

Drouet JL (2003) MODICA And MODANCA: modelling the three-dimensional shoot structure of graminaceous crops from two methods of plant description. *Field Crops Research* **83**, 215–222.

Dupont S, Brunet Y and Jarosz N (2006) Eulerian modelling of pollen dispersal over heterogeneous vegetation canopies. *Agricultural and Forest Meteorology* **141**, 82–104.

Dyer AJ (1974) A review of flux-profile relationships. *Boundary-Layer Meteorology* **7**, 363–371.

Flesch TK, Wilson JD and Yee E (1995) Backward-time Lagrangian stochastic dispersion models and their application to estimate gaseous emissions. *Journal of Applied Meteorology* **34**, 1320–1332.

Garratt JR (1992) *The Atmospheric Boundary Layer.* Cambridge, UK: Cambridge University Press.

Gash JHC (1986) Observation of turbulence downwind of a forest-heath interface. *Boundary-Layer Meteorology* **36**, 227–237.

Gregory PH (1973) *Microbiology of the Atmosphere.* New York, Toronto: John Wiley & Sons.

Hall AJ, Vilella F, Trapani N and Chimenti C (1982) The effect of water stress and genotype on the dynamics of pollen shedding and silking in maize. *Field Crops Research* **5**, 349–363.

Hanna SR, Chang JC and Strimaitis DG (1993) Hazardous gas model evaluation with field observations. *Atmospheric Environment* **27A**, 2265–2285.

Heisler GM and DeWalle DR (1988) Effects of windbreak structure on wind flow. *Agriculture, Ecosystem and Environment* **18**, 41–69.

Helton JC, Johnson JD, Sallaberry CJ and Storlie CB (2006) Survey of sampling-based methods for uncertainty and sensitivity analysis. *Reliability Engineering and System Safety* **91**, 1175–1209.

Iman RL (1992) Uncertainty and sensitivity analysis for computer modeling applications. *Reliability technology* **28**, 153–168.

Janssen PHM (1994) Assessing sensitivities and uncertainties in models: a critical evaluation. *Proceedings of the 75th Anniversary Conference of WAU, April 5–7*, Wageningen, The Netherlands, pp. 344–361.

Janssen PHM, Heuberger PSC and Sanders R (1992) UNCSAM 1.1: A Software Package for Sensitivity and Uncertainty Analysis. National Institute of Public Health and Environmental Protection, Bilthoven, The Netherlands. Report No. 959101004.

Jarosz N, Loubet B, Durand B, McCartney A, Foueillassar X and Huber L (2003) Field measurements of airborne concentration and deposition rate of maize pollen. *Agricultural and Forest Meteorology* **119**, 37–51.

Jarosz N, Loubet B and Huber L (2004) Modelling airborne concentration and deposition rate of maize pollen. *Atmospheric Environment* **38**, 5555–5566.

Jarosz N, Loubet B, Durand B, Foueillassar X and Huber L (2005) Variations in maize pollen emission and deposition in relation to microclimate. *Environmental Science and Technology* **39**, 4377–4384.

Kaimal JC and Finnigan JJ (1994) *Atmospheric Boundary Layer Flows Their Structure and Measurement.* Oxford: Oxford University Press.

Klein EK, Lavigne C, Foueillassar X, Gouyon PH and Laredo C (2003) Corn pollen dispersal: quasi-mechanistic models and field experiments. *Ecological Monographs* **73**, 131–150.

Kumar A, Luo J and Bennett G (1993) Statistical evaluation of lower flammability distance using four hazardous release models. *Process Safety Progress* **12**, 1–11.

Lizaso JI, Westgate ME, Batchelor WD and Fonseca A (2003) Predicting potential kernel set in maize from simple flowering characteristics. *Crop Science* **43**, 892–903.

Loubet B (2000) *Modélisation du dépôt sec d'ammoniac atmosphérique à proximité des sources* (PhD thesis). Paul Sabatier University, Toulouse, France.

Loubet B, Cellier P, Milford C and Sutton MA (2006) A coupled dispersion and exchange model for short-range dry deposition of atmospheric ammonia. *Quarterly Journal of the Royal Meteorological Society* **132**, 1733–1763.

Loubet B, Jarosz N, Saint-Jean S and Huber L (2007) A method for measuring the settling velocity distribution of large biotic particles. *Aerobiologia* **23**, 159–169.

Marceau A, Loubet B, Andrieu B, Durand D, Foueillassar X and Huber L (2011) Modelling diurnal and seasonal patterns of maize pollen emission in relation to meteorological factors. *Agricultural and Forest Meteorology* **151**, 11–21.

Marceau A, Saint-Jean S, Loubet B and Huber L (2012) Biophysical characteristics of maize pollen: variability during emission and consequences on cross-polination risks. *Field Crops Research* **127**, 51–63.

McCartney, HA and Lacey ME (1991) Wind dispersal of pollen from crops of oilseed rape (Brassica napus L.). *Journal of Aerosol Science* **22**, 467–477.

McKay MD, Beckmann RJ and Conover WJ (1979) A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* **21**, 239–245.

Paulson CA (1970) The mathematical representation of wind speed and temperature profiles in the unstable atmospheric surface layer. *Journal of Applied Meteorology* **9**, 857–861.

Raupach MR (1989) Applying Lagrangian fluid-mechanics to infer scalar source distributions from concentration profiles in plant canopies. *Agricultural and Forest Meteorology* **47**, 85–108.

Raupach MR, Finnigan JJ and Brunet Y (1996) Coherent eddies and turbulence in vegetation canopies: the mixing-layer analogy. *Boundary-Layer Meteorology* **78**, 351–382.

Raynor GS, Eugene CO and Janet VH (1972) Dispersion and deposition of corn pollen from experimental sources. *Agronomy Journal* **64**, 420–427.

Reynolds AM (2000) Prediction of particle deposition on to rough surfaces. *Agricultural and Forest Meteorology* **104**, 107–118.

Saltelli A, Chan K and Scott EM. (2000a) Probability and Statistics series. In *Sensitivity Analysis.* John Wiley and Sons.

Saltelli A, Tarantola S and Compolongo F (2000b) Sensitivity analysis as an ingredient of modelling. *Statistical Science* **15**, 377–395.

Scott RK (1970) The effect of weather on the concentration of pollen within sugar-beet seed crops. *Annals of Applied Biology* **66**, 119–127.

Snyder WH and Lumley JL (1971) Some measurements of particle velocity autocorrelation functions in a turbulent flow. *Journal of Fluid Mechanics* **48**, 41–71.

Sofiev M, Siljamo P, Ranta H, Linkosalo T, Jaeger S, Rasmussen A, Rantio-Lehtimaki A, Severova E and Kukkonen J (2013) A numerical model of birch pollen emission and dispersion in the atmosphere. *International Journal of Biometeorology* **57**, 45–58.

Torimaru T, Wennstrom U, Lindgren D and Wang XR (2012) Effects of male fecundity, interindividual distance and anisotropic pollen dispersal on mating success in a Scots pine (*Pinus sylvestris*) seed orchard. *Heredity* **108**, 312–321.

Treu R and Emberlin J (2000) Pollen dispersal in the crops maize, oilseed rape, potatoes, sugar beet and wheat. In a report for the Soil Association from the National Pollen Research Unit (Worcester, University College Worcester).

Zhang R, Duhl T, Salam MT, House JM, Flagan RC, Avol EL, Gilliland FD, Guenther A, Chung SH, Lamb BK and VanReken TM (2014) Development of a regional-scale pollen emission and transport modeling framework for investigating the impact of climate change on allergic airway disease. *Biogeosciences* **11**, 1461–1478.

Zink K, Kaufmann P, Petitpierre B, Broennimann O, Guisan A, Gentilini E and Rotach MW (2017) Numerical ragweed pollen forecasts using different source maps: a comparison for France. *International Journal of Biometeorology* **61**, 23–33.