Transgenic Res (2014) 23:795–807
DOI 10.1007/s11248-014-9810-3

ORIGINAL PAPER

Establishment and optimization of a regionally applicable maize gene-flow model

Ning Hu · Jichao Hu · Xiaodong Jiang · Zongzhi Lu · Yufa Peng · Wanlong Chen · Kemin Yao · Ming Zhang · Shirong Jia · Xinwu Pei · Weihong Luo

Received: 9 November 2013 / Accepted: 3 June 2014 / Published online: 25 June 2014

Abstract Because of the rapid development of transgenic maize, the potential effect of transgene flow on seed purity has become a major concern in public and scientific communities. Setting a proper isolation distance in field experiments and seed production is a possible solution to meet seed-quality standards and ensure adventitious contamination of products is below a specific threshold. By using a Gaussian plume model as basis and data recorded by meteorological stations as input, we have established a simple regionally applicable maize gene-flow model for prediction of the maximum threshold distances (MTD) at which gene-flow frequency is equal to or lower than a threshold value of 1 or 0.1 % ($\mathrm{MTD}_{1\%}$, $\mathrm{MTD}_{0.1\%}$). After optimization of the model variables, simulated outcrossing rate was a good fit to data obtained from field experiments ($y = 1.156x$, $R^2 = 0.8913$, $n = 30$, $P < P_{0.01}$). In the process of model calibration, it was found that only 15.82 % of the total amount of the pollen released by each plant participated in the dispersal process. The variable "*a*" for genetic pollen competitiveness between donor and recipient was introduced into our model, for the "Zinuo18" and "Su608" used, "*a*" was 17.47. Finally, the model was successfully used in the spring maize-growing region of Northeast China. The range of $\mathrm{MTD}_{1\%}$ and $\mathrm{MTD}_{0.1\%}$ in this region varied from 10 m to 49 m and from 17 m to 125 m, respectively.

Keywords Maize · Gene-flow model · Threshold value · Maximum threshold distance

---

Ning Hu, Jichao Hu, and Xiaodong Jiang contributed equally to this publication.

Electronic supplementary material The online version of this article (doi:10.1007/s11248-014-9810-3) contains supplementary material, which is available to authorized users.

N. Hu · W. Luo (✉)
College of Agriculture, Nanjing Agricultural University, 1 Weigang, Nanjing 210095, Jiangsu, People's Republic of China
e-mail: lwh@njau.edu.cn

N. Hu · J. Hu · X. Jiang · W. Chen · K. Yao
Jiangsu Key Laboratory of Agricultural Meteorology,
College of Applied Meteorology,
Nanjing University of Information Science &
Technology, Nanjing 210044, Jiangsu,
People's Republic of China

Z. Lu
Institute of Plant Protection, Jilin Academy of
Agricultural Sciences, Gongzhuling 136100, Jilin,
People's Republic of China

Y. Peng
Institute of Plant Protection, Chinese Academy of
Agricultural Sciences, Beijing 100193, People's Republic
of China

Introduction

Throughout the world, maize (*Zea mays* L.) is one of the most important food and feed crops. Genetically modified (GM) crops have been developed rapidly in the past two decades, and maize has become the second largest GM crop worldwide. The cultivation area of GM maize in 2011 reached 51 million ha, accounting for 32 % of the total acreage of GM crops (James 2011).

The development of GM maize has also resulted in significant progress in the study of the gene flow and outcrossing of maize. In general, the frequency of gene flow decreases progressively with increasing distance, and can be expressed as an exponential curve (Sanvido et al. 2008; Ma et al. 2004). We surveyed the data on maize gene flow published internationally in recent decades and found that the maximum threshold distance (MTD) at which the gene-flow frequency is equal to or lower than a threshold value of 1 % ($\mathrm{MTD}_{1\%}$) ranges from 3 to 60 m (Ma et al. 2004; Bannert et al. 2008; Goggi et al. 2007; Della Porta et al. 2008; Di and Liu 2008; Lu et al. 2005) and the maximum distance at which the gene-flow frequency is equal to or lower than a threshold value of 0.1 % ($\mathrm{MTD}_{0.1\%}$) ranges from 110 to 119 m (Goggi et al. 2007; Di and Liu 2008; Lu et al. 2005). To study the co-existence of GM and non-GM maize after large-scale commercialization of GM maize, scientists in the UK, Spain, and France have collected samples from non-GM maize surrounding GM maize-planting areas. The results showed that $\mathrm{MTD}_{1\%}$ and $\mathrm{MTD}_{0.1\%}$ were less than 30 and 80 m, respectively (Pla et al. 2006; Weekes et al. 2007).

Obviously, these MTD figures have important reference value for setting appropriate isolation distances and taking appropriate control measures both in maize seed production and in environmental release of GM maize. However, results derived from one place or one year would not be consistent, because of wide environmental variation, which would significantly affect gene flow. Therefore, it is valuable to establish a model, particularly a regionally applicable model with predictive function. Models describing maize gene flow have been reported in the literature; these can be classified as two types—statistical models, for example exponential, logarithmic, and power functions (Ivanovska et al. 2009; Goggi et al. 2006; Gustafson et al. 2006; Allnutt et al. 2008) and mechanistic models based on the theory of atmospheric diffusion, including the Gaussian plume model (GPM), the Eulerian advection–diffusion conservation equation, the Lagrangian random flight model, the localised near field (LNF) model, the large eddy simulation (LES) model, and the normal inverse Gaussian model (Dupont et al. 2006; Dietiker et al. 2011; Jarosz et al. 2004; Arritt et al. 2007; Chamecki et al. 2009; Angevin et al. 2008; Klein et al. 2003; Coléno et al. 2009; Loos et al. 2003). We compared these models and chose the GPM as the basis of our model for two reasons:

1. the GPM has been commonly used to describe spore and pollen dispersal; and
2. fewer variables are involved and the calculation is much simpler.

In our previous work, a rice gene-flow model was established on the basis of the GPM, and used to calculate the maximum distance of gene flow in the southern rice growing region of China (Yao et al. 2008).

The purpose of this study was to establish a maize gene-flow model based on the GPM by using historical meteorological data from a variety of locations as input. As a result, the model can be used over a larger scale to predict MTDs at different places for a historical period. Here we report the detailed results obtained in this study.

Materials and methods

Plant materials

Morphological markers are commonly used to study outcrossing between different varieties or species. In

---

M. Zhang
Biotechnology Research Center, Jilin Academy of Agricultural Sciences, Changchun 130124, Jilin, People's Republic of China

S. Jia (✉) · X. Pei (✉)
Laboratory for Crop Genetic Engineering, Biotechnology Research Institute, Chinese Academy of Agricultural Sciences, 12 Zhongguancun South Street, Beijing 100081, People's Republic of China
e-mail: srjia@126.com

X. Pei
e-mail: peixw@mail.caas.net.cn

this study we used seed colour as a marker. The male parent used was Zinuo18 (“Pz”), with purple seed colour (complete dominant, homozygous); the female parent used was Jidan35 (“Yj”, used in the Gongzhuling experiment) or Su608 (“Ys”, used in the Lishui experiment), with a yellow seed colour (recessive, homozygous).

Field experiment design

Field experiments were conducted in experimental fields in Gongzhuling (124°49′22″N, 43°30′17″E), Jilin province (Northeast China), and Lishui (119°10′36″ N, 31°34′49″E), Jiangsu province (East China), in 2010. Data generated in the Gongzhuling and Lishui experiments were used to establish and validate the model.

Seeds of the donor and recipient were sown on the same day in the Gongzhuling experiment, with a 35-day delay after sowing local varieties to ensure no external pollen interference from surrounding fields within 5 km. The dates of tassel flowering and ears silking of the donor were 2 and 1 days later, respectively, than for the recipient. In the Lishui experiment, there was no other maize planting within 3 km. Seeds of the donor were sown five days later than those of the recipient, to meet the requirement for synchronisation of the flowering season. The flowering date of the donor was one day earlier than that of the recipient, but silking occurred on the same day. In other words, flowering and silking of donor and recipient were well synchronised in both experiments.

As shown in Fig. 1, the field experiment was designed in the shape of concentric circle with the pollen donor planted in the inner circle (radius: $r = 15$ m in Gongzhuling; $r = 10$ m in Lishui) and the recipient planted in the outer circle (radius: $r = 80$ m in Gongzhuling; $r = 30$ m in Lishui). Outside the outer circle was an additional fan-shaped area specially designed to enable observation of pollen deposition (radius: $r = 230$ m, undersurface was maize in Gongzhuling; $r = 270$ m, undersurface was blank in Lishui). The total areas of the experiments in Gongzhuling and Lishui were 5.7 and 5.9 ha, respectively. The planting density was 60,000 plants ha^{-1} in both experiments. Meteorological data during the whole flowering season are summarized in the supplementary file (Appendix A).

Artificial crossing experiment

To quantify the relationship between outcrossing rate and the ratio of pollen grains released by donor and recipient, an artificial cross experiment was conducted in an independent field experiment in Lishui. Before silking, the ears of “Ys” were covered with paper bags. During flowering, the tassels of “Pz” and “Ys” were covered with paper bags and pollen was collected from 8:00–8:10 am. Pollen grains of “Pz” and “Ys” were weighed to form 15 combinations of pollen mixtures, including “Pz” pollen-to-“Ys” pollen 0:100, 1:99, 5:95, 10:90, 20:80, 30:70, 40:60, 50:50, 60:40, 70:30, 80:20, 90:10, 95:5, 99:1, and 100:0. The pollen mixture was individually pollinated on three ears of “Ys”. After maturation, the seeds were collected, and outcrossing rate was determined by checking seed colour. The experiment was repeated three times.

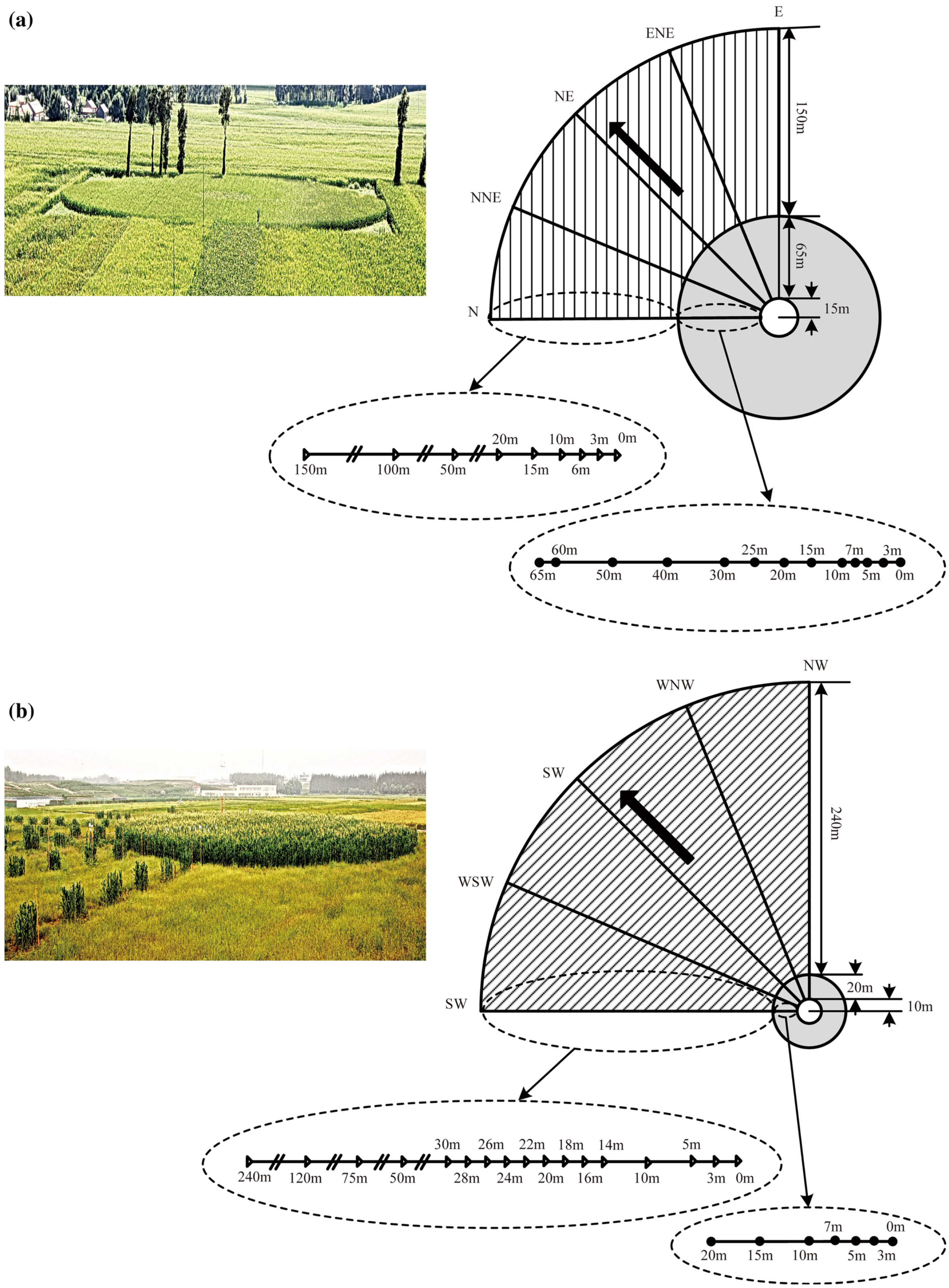
Measurement and observation

Pollen viability

To determine whether pollen longevity is an important characteristic determining the probability of fertilization when pollen travels over long distances, we used the 2,3,5-triphenyltetrazolium chloride (TTC) method to measure the time course of pollen viability, in accordance with Duncan and Widholm (2004). At the flowering stage, tassels collected from the field were brought to the laboratory. Pollen grains were deposited on glass slides which were then maintained in the plant-growth chamber at different temperatures (20–35 °C) and relative humidity (60–90 %) for 0–480 min. At different times, pollen grains on slides were stained with 0.5 % TTC solution, checked under a microscope, and pollen viability and half life were determined. Each treatment was repeated three times. Results indicated that the half life of maize pollen ranged from 52 to 155 min (unpublished, Jiang), depending on temperature and relative humidity.

Pollen release

Two variables, total amount of pollen released from anthers per plant per day (Q_{day}, grains plant^{-1} day^{-1}) and the percentage of florets flowering per plant per


(a)
E
ENE
NE
NNE
N
150m
65m
15m
20m 10m 3m 0m
150m 100m 50m 15m 6m
60m 25m 15m 7m 3m
65m 50m 40m 30m 20m 10m 5m 0m
(b)
NW
WNW
SW
WSW
SW
240m
20m
10m
30m 26m 22m 18m 14m 5m
240m 120m 75m 50m 28m 24m 20m 16m 10m 3m 0m
7m 0m
20m 15m 10m 5m 3m

◀**Fig. 1** Photographs and schematic diagrams of field experiments on maize gene flow at Gongzhuling (**a**) and Lishui (**b**). *E*, *ENE*, *NE*, etc., represent compass directions. *Large arrows* represent prevailing wind direction during the flowering season. *White* and *gray* denote areas in which pollen donors and recipients, respectively, were planted. *Cross hatching* denotes areas used for observation of pollen deposition, with *diagonal cross hatching* denoting an area without crops (blank) and *vertical cross hatching* denoting an area with maize growing. *Filled circles* and *arrow heads* denote sites for observation of outcrossing and pollen deposition, respectively

hour (R_{hour}, %), were measured to estimate pollen release.

During the flowering period, 30 donor and recipient plants were randomly chosen for observation of Q_{day}. Tassels were covered with paper bags and pollen grains were collected and weighed each day at 18:00. Finally, the weight of pollen grains per plant per day was converted into the total number of pollen grains per plant per day on the basis of the weight of a single pollen grain (0.362 μg grain^{-1}, unpublished data, Jiang).

Ten additional donor and recipient plants were chosen to measure R_{hour}. The number of flowering florets out of the total number of florets was investigated every hour.

The diurnal pattern of pollen release during the whole flowering season could be obtained from Q_{day} and R_{hour} as follows:

$$Q_{\mathrm{release}} = Q_{\mathrm{day}} \times R_{\mathrm{hour}} \tag{1}$$

Pollen release was assessed on an hourly basis over the periods when pollen was released. Most of the plants and florets started releasing pollen at 0500 hours and stopped at approximately 1800 hours. To approximate the dynamics of pollen release, a normal function was used (Marceau et al. 2011). Variables were estimated for each of the three varieties measured, by use of the least-squares method:

$$Q_{\mathrm{release}}(t,d) = \frac{Q_s}{2\pi\sigma_t\sigma_d}\exp\left[-\frac{(t-\mu_t)^2}{2\sigma_t^2} - \frac{(d-\mu_d)^2}{2\sigma_d^2}\right] \tag{2}$$

where Q_{release} is the hourly flux of pollen release (grains plant^{-1} h^{-1}), and d and t are the day after the first-flowering dates and the time after sunrise, respectively. Q_s, σ, and μ are the optimized variables representing total pollen production (grains plant^{-1} season^{-1}), standard deviation, and peak location of the normal curve. In this study, $Q_s = 1.45 \times 10^7$ grains plant^{-1} season^{-1}, $\sigma_t = 2.04$, $\sigma_d = 1.53$, $\mu_t = 4.49$, and $\mu_d = 6.10$ for "Pz"; $Q_s = 1.36 \times 10^7$ grains plant^{-1} season^{-1}, $\sigma_t = 2.20$, $\sigma_d = 0.98$, $\mu_t = 5.37$, and $\mu_d = 7.20$ for "Yj"; and $Q_s = 1.01 \times 10^7$ grains plant^{-1} season^{-1}, $\sigma_t = 2.29$, $\sigma_d = 0.98$, $\mu_t = 5.41$, and $\mu_d = 7.20$ for "Ys". Measured and simulated pollen release are compared in the supplementary file (Appendix B).

Pollen deposition rate

Pollen deposition rate was monitored daily within periods of 7 and 9 days in Gongzhuling and Lishui, respectively. The sampling sites were located at different distances (0–240 m, as shown in Fig. 1) along five lines downstream of the prevailing wind direction. Each line was 22.5° apart (Fig. 1). Three glass slides coated with Vaseline were placed horizontally at the position of the average height of the ears. The number of pollen grains taken from 30 fields of view was counted under a microscope, and the pollen deposition rate (grains m^{-2} h^{-1}) was calculated.

Measurement of outcrossing rate

The sampling sites were individually located at different distances starting from the edge of inner circle, as shown in Fig. 1. Outcrossing rate was measured on the basis of seed colour as morphological marker. Because the frequency of gene flow at longer distances was relatively low, more than 10,000 maize grains were taken beyond a distance of 15 m to ensure a rare outcrossing event could be detected.

Outcrossing rate can be calculated by use of Eq. (3):

$$G = N_{\mathrm{donor}}/(N_{\mathrm{donor}} + N_{\mathrm{recipient}}) \times 100\ \% \tag{3}$$

where G is outcrossing rate (%), and N_{donor} or $N_{\mathrm{recipient}}$ are the numbers of purple or yellow seeds observed, respectively.

Micrometeorological data measurement

Wind speed and direction at a height of 10 m were measured by use of a mechanical anemometer (010C; Met One Instruments, OR, US) and a wind vane (WP4; Delta-T Devices, Cambridge, UK), respectively. Global radiation was measured by use of a net radiation sensor (CNR4; Kipp & Zonen, Delft, The

Netherlands). All the aforementioned data were automatically collected at intervals of 1 s and stored per half hour by use of a data logger (CR3000; Campbell Scientific, UT, USA).

Maize gene-flow model

Model establishment

The ratio of pollen grains deposited from the donor and recipient at a given location (denoted "*Rqq*" hereafter) is a major factor determining the outcrossing rate or gene-flow frequency. Therefore, outcrossing rate, G, can be described by use of Eq. (4):

$$G = f(Rqq) \tag{4}$$

where Rqq is expressed as $Rqq = q_{\mathrm{donor}}/(q_{\mathrm{recipient}} + q_{\mathrm{donor}})$, in which q_{donor} and $q_{\mathrm{recipient}}$ (grains m^{-2} season^{-1}) represent the quantity of pollen grains from the donor and recipient, respectively, deposited at a given location throughout the flowering season.

Because maize pollen is large and heavy, most of the grains are deposited close to their source without further dispersal (Sears and Stanley-Horn, 2000; Pleasants et al. 1999). Therefore, the total number of pollen grains released by each plant is divided into two parts. We define the portion of pollen grains deposited near the pollen-releasing plant as $Q_{\mathrm{deposition}}$ (grains plant^{-1} h^{-1}). The remainder of the pollen, that participating the dispersal process, which is carried above the canopy by wind and dispersed over longer distances, is defined as $Q_{\mathrm{dispersed}}$ (grains plant^{-1} h^{-1}). The ratio of $Q_{\mathrm{dispersed}}$ to Q_{released} is defined as pollen dispersal rate (R_{d}, %):

$$Q_{\mathrm{dispersed}} = Q_{\mathrm{released}} - Q_{\mathrm{deposited}} = Q_{\mathrm{released}} \cdot R_{\mathrm{d}} \tag{5}$$

A single plant is defined as a point source continuously releasing pollen grains. On the basis of the GPM, the pollen concentration at a location downwind released by a point source can be expressed as (Arya 1999):

$$C(x, y, z, i, j, z_{\mathrm{H}}, t) = \frac{Q_{\mathrm{dispersed}}(i, j, z_{\mathrm{H}}, t)}{2\pi \bar{u}_H \sigma_y \sigma_z} \exp\left[-\frac{(y-j)^2}{2\sigma_y^2}\right] \exp\left[-\frac{[z - z_H + v_d(x-i)/\bar{u}_H]^2}{2\sigma_z^2}\right] \tag{6}$$

where t is time, C (grains m^{-3}) is the pollen concentration, $Q_{\mathrm{dispersed}}$ (grains plant^{-1} s^{-1}) is the total amount of pollen released by a single source in unit time, z_H (m) is the mean height of the tassels, $\bar{u}_{\mathrm{H}}$(m s^{-1}) is the mean wind speed at the height of tassels, v_{d} (m s^{-1}) is the deposition velocity, σ_y (m) and σ_z (m) are atmospheric diffusion parameters, i.e., the standard deviations of the pollen concentration in the crosswind and vertical directions, respectively.

In Eq. (6), the dominant wind direction per hour and the location of the pollen source determines the coordinate system. The downwind direction is defined as the positive x-axis whereas the crosswind direction is defined as the y-axis, and the height above the ground is defined as the z-axis. The pollen source is located at the coordinates (i, j, H), and the coordinates (x, y, z) define a given location to which the pollen is dispersed. If the prevailing wind direction during the whole flowering season is not consistent with the dominant wind direction per hour, and the angle between these two directions is "θ", the pollen concentration should be modified as follows:

$$\begin{aligned} C(x, y, z, i, j, z_{\mathrm{H}}, t) = C(&x' \cos\Delta\theta \\ &+ y' \sin\Delta\theta, y' \cos\Delta\theta \\ &- x' \sin\Delta\theta, z, i, j, z_{\mathrm{H}}, t) \end{aligned} \tag{7}$$

The pollen dispersed downwind to different distances from each point source (a single plant) is calculated individually. The figures derived from numerous individual point sources are then summed as the total amount of pollen dispersed to a given location:

$$C_{\mathrm{donor}}(x, y, z, t) = \int_0^{2\pi} \int_0^{r_{\mathrm{donor}}} C(x, y, z, r\cos\theta, r\sin\theta, z_{\mathrm{H}}, t) \cdot \mathrm{d}r \cdot \mathrm{d}\theta \tag{8}$$

$$C_{\mathrm{recipient}}(x, y, z, t) = \int_0^{2\pi} \int_{r_{\mathrm{donor}}}^{r_{\mathrm{recipient}}} C(x, y, z, r\cos\theta, r\sin\theta, z_H, t) \cdot \mathrm{d}r \cdot \mathrm{d}\theta \tag{9}$$

where r_{donor} and $r_{\mathrm{recipient}}$ are the radii of the inner and outer circles, respectively, and the polar coordinate (r, θ) defines the location of each point source.

Settling speed and deposition velocity are important variables in pollen deposition. In this study, we calculated settling speed and deposition velocity by use of Stokes Law (Aylor 2002) and the resistance

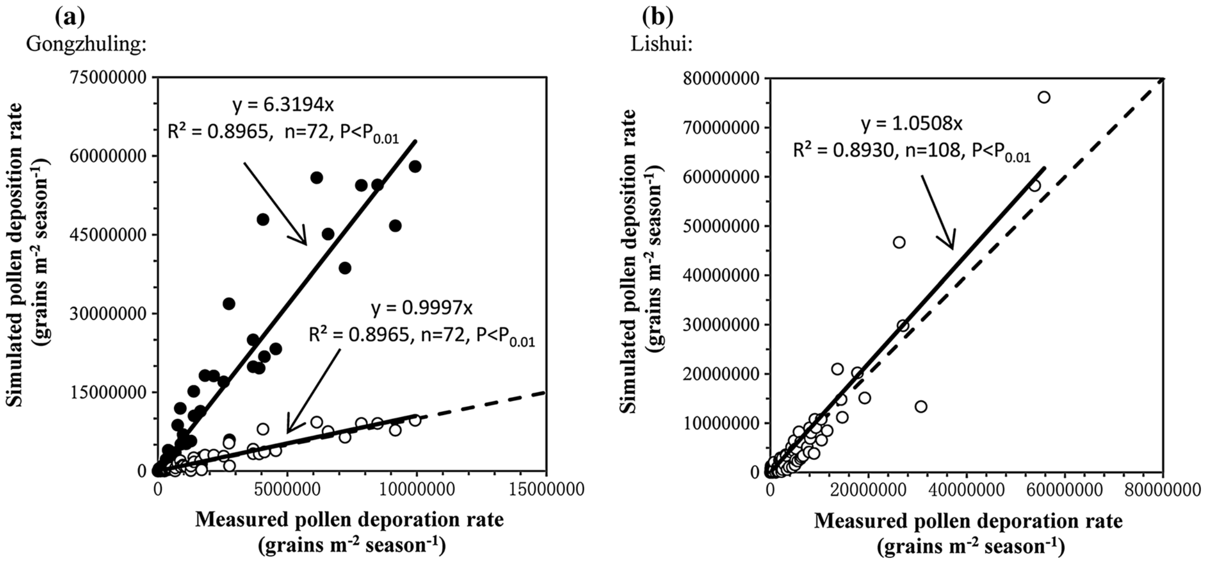


Fig. 2 Calibration (**a**) and validation (**b**) of pollen dispersal, R_d. *Solid circles* are used to indicate that all pollen grains released by each point source participate in the dispersal process without deposition ($R_d = 1$); *open circles* indicate only 15.82 % of the pollen grains released by each point source participate in the dispersal process ($R_d = 15.82$ %). The *solid line* is the linear regression between simulated and measured deposition; the *short dashed line* is a 1:1 reference line

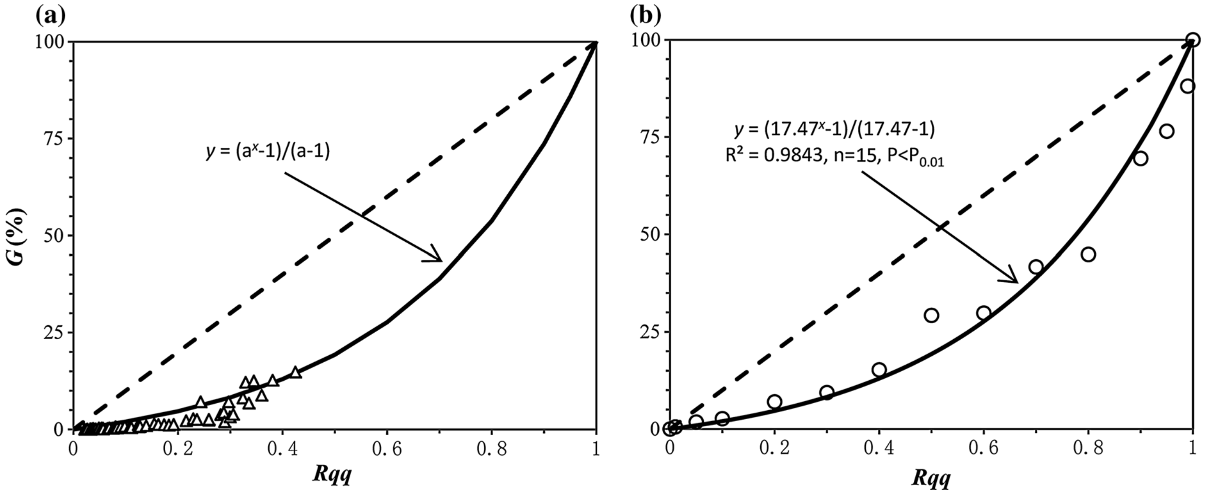


Fig. 3 Relationship between the ratio of pollen grains from “Pz” and “Ys” (*Rqq*) and outcrossing (*G*) obtained in the Gongzhuling experiment (**a**) and in the artificial crossing experiment (**b**). The *solid line* indicates the non-linear regression between *Rqq* and *G*; the *dashed line* is a reference line: $y = x \times 100$ %

model (Slinn 1982), respectively. Here, the average v_d was determined to be 0.25 m s^{-1}. The pollen deposition rate (D, grains plant^{-1} h^{-1}) is given by Eq. (10) and the total amount of pollen deposited from donor and recipient could be easily summed for the whole flowering season.

$$D = v_d \cdot C \tag{10}$$

It should be noted that pollen grains from recipients deposited in the recipient planting area (the outer circle) include two components—one dispersed from other recipient plants located upwind and the other deposited from its own tassel. Therefore, $Q_{\text{deposition}}$ from its own tassel should be added to $D_{\text{recipient}}$.

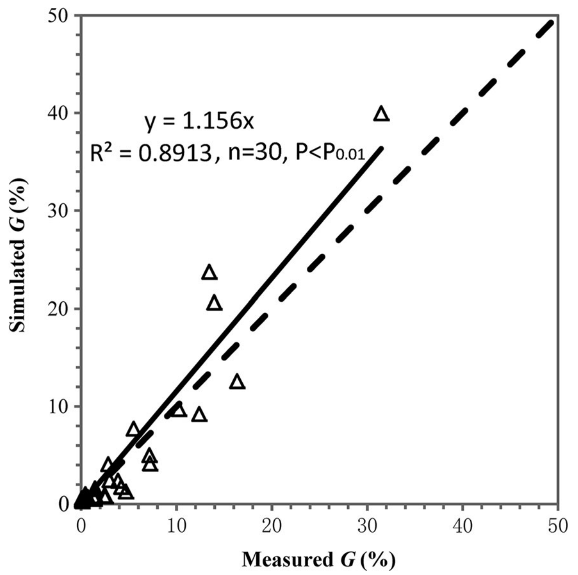


Fig. 4 Simulated versus measured outcrossing (G) in the Lishui experiment. The *solid line* is the regression line; the *dashed line* is a 1:1 reference line

Meteorological data

To determine wind speed at the tassel height, the wind speed at the 10 m level recorded by meteorological stations was converted in accordance with Menteith and Unsworth (1990).

The atmospheric dispersal parameters σ_y and σ_z were calculated as follows (Jiang et al. 2004):

$$\sigma_y = \gamma_1 x^{\alpha_1} \quad \text{and} \quad \sigma_z = \gamma_2 x^{\alpha_2} \tag{11}$$

where α_1, γ_1, α_2, and γ_2 are variables depending on atmospheric stability (Appendix C). In this study, stability levels were classified by the range of wind velocity and global radiation (Appendix D).

Statistical analysis

Results from experiments and data simulated by the model were used for regression analysis. A routine t test was used to determine the significance of differences at $P_{0.05}$ and $P_{0.01}$.

Application of the maize gene-flow model

In this study, the provinces Liaoning, Jilin, Heilongjiang, and part of Inner Mongolia (approximately 1.26 million square kilometres) in the spring maize-growing region of Northeast China were targeted for application of the model, because this is a major maize-production region, producing 1/3 of the maize grown annually in China.

A total of 101 local stations in this region were chosen and historical meteorological data from these stations from 1997–2010, including global radiation and wind speed and direction were collected and used as input to the model. The location of the study areas and the 101 meteorological stations are shown in Appendix E.

During application of the model a threshold value of 1 or 0.1 % was used to calculate the threshold distances (TDs) at which gene-flow frequency is equal to or lower than the threshold value in different parts of Northeast China.

Because of climate fluctuation, TDs vary substantially from year to year. For a given location, the TDs were calculated by the model from 14 years of meteorological data. The longest TD was described as the maximum threshold distance (MTD).

Results

Calibration and validation of $Q_{\text{dispersed}}$

During calibration of $Q_{\text{dispersed}}$ by use of the results obtained from the Gongzhuling experiment, we found that if it was assumed that all the pollen grains released from each point source participated in the dispersal process without deposition, i.e., $R_{\text{d}} = 1$, then the model-simulated value was significantly higher than the measured value, $y = 6.3194x$ ($R^2 = 0.8965$, $n = 72$, $P < P_{0.01}$) (Fig. 2a). Apparently, some of the pollen released did not participate in the dispersal process. When pollen dispersal rate, R_{d}, was adjusted to 15.82 %, the simulated and measured values were well matched, $y = 0.9997x, R^2 = 0.8965, n = 72, P < P_{0.01}$ (Fig. 2a).

The data generated in the Lishui experiment were further used for validation of R_{d}. The results demonstrated that the simulated and measured values were in agreement ($y = 1.0508\ x$, $R^2 = 0.8930$, $n = 108$, $P < P_{0.01}$, Fig. 2b).

Simulation of outcrossing rate

To determine the extent of outcrossing it is necessary to elucidate the quantitative relationship between the

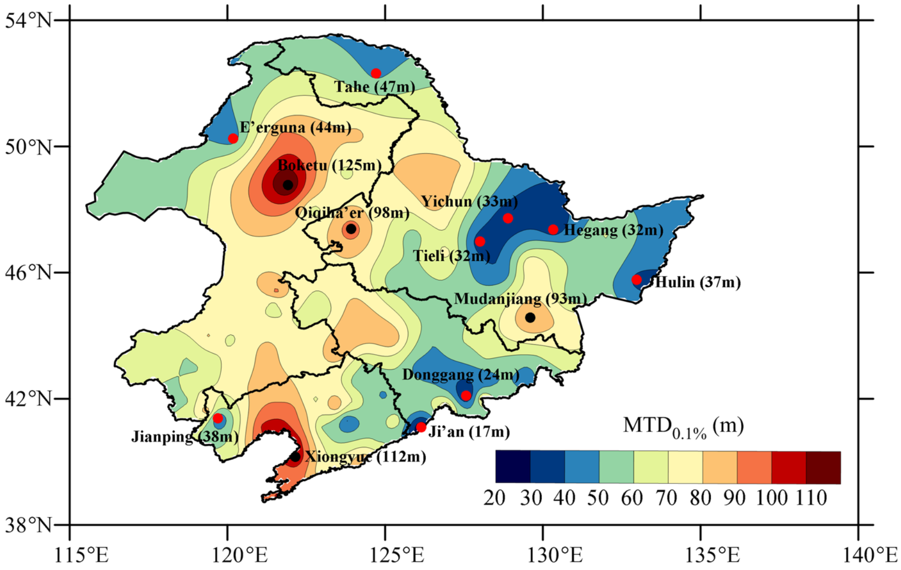


Fig. 5 The spatial distribution of maximum threshold distances ($MTD_{0.1\%}$) of maize gene flow from 1997 to 2010 in the spring maize-growing region of Northeast China. $MTD_{0.1\%}$ is the maximum distance at which the gene-flow frequency is equal to or lower than a threshold value of 0.1 %. *Black circles* and *red circles* represent spots with the highest and lowest MTD, respectively. (Color figure online)

ratio of pollen grains deposited from donor and recipient (*Rqq*) and outcrossing rate (*G*), as described by Eq. (4).

In the Gongzhuling experiment, we calculated *Rqq* at given distances, as described in Eqs. (5)–(10), and the results were compared with *G* observed from the field experiment. It was found there was a non-linear relationship between these variables (Fig. 3a). This non-linear equation could be expressed as an exponential curve, $G = (a^{Rqq} - b)/c \times 100\ \%$. Because it also met the requirements $G = 0\ \%$ when $Rqq = 0$ and $G = 100\ \%$ when $Rqq = 1$, the equation above could be simplified as follows:

$$G = (a^{Rqq} - 1)/(a - 1) \times 100\ \% \tag{12}$$

where "*a*" is a coefficient expressing pollen genetic competitiveness between donor and recipient.

To quantify the value of "*a*" in Eq. (12), a series of pollen mixture combinations was artificially crossed with the female parent "Ys". Results are shown in Fig. 3b, in which the value of "*a*" is calculated to be 17.47, and $R^2 = 0.9843$, $n = 15$, $P < P_{0.01}$.

Validation of outcrossing rate measured in the Lishui experiment

For validation of the relationship between *G* and *Rqq*, the value $a = 17.47$ obtained above was incorporated into the process for calculation of *G* at a given location in the Lishui experiment. The calculated value of *G* was then compared with the field observed data at the same location. Results indicated they were a good fit, $y = 1.156x$, $R^2 = 0.8913$, $n = 30$, $P < P_{0.01}$ (Fig. 4).

Application of the maize gene-flow model

In this study, the model was applied to the spring maize-growing region in Northeast China. Data indicated there was a wide range of spatial and temporal variations of MTD because the climate conditions vary substantially place by place and year by year.

$MTD_{1\%}$ and $MTD_{0.1\%}$ in this region are shown in Appendix F. Results indicated that the $MTD_{1\%}$ during 1997–2010 among the 101 sites was within the range 10–49 m, and $MTD_{0.1\%}$ ranged from 17 to 125 m.

The spatial distribution pattern of MTDs (Fig. 5) showed that the whole region could be divided into three parts. The central part was a region with higher MTDs whereas the MTDs in the southeast part and in the northwest part were lower. This was probably because the study area was surrounded by three mountain ranges: the Xiao Xing'an Mountains to the Northeast, the Greater Xing'an Mountains to the northwest, and the Changbai Mountains to the southeast. In China, the southeast monsoon usually occurs in the summer season, and affects the climate in this region. Later, it affected maize pollen dispersal and gene flow. As a result, the MTDs in the open plain region and the windward side of mountains were higher, whereas the MTDs in the region to the leeward side of the mountains were relatively low. In conclusion, the southeast monsoon and topography were the key factors affecting patterns of MTDs distribution.

Discussion

Characteristics of the model

In this work, we established a regionally applicable maize gene-flow model to predict maximum threshold distances at different locations in 1997–2010. Threshold distance (TD) is calculated on a yearly basis and results in 14 TDs for each location. The highest among these 14 TDs was used to represent the MTD for a given site. In brief, what we considered in terms of risk assessment was a “worst case” scenario. To meet the requirement of this concept, two methods were used in this study. One was to use homozygous donor and recipient as material, because a genetically hemizygous or heterozygous seed colour locus will result in underestimation of outcrossing rate. The other was to adjust sowing date to synchronize the flowering of donor and recipient.

The model is based on the GPM, which is suitable for describing small-scale movement of particles in the atmosphere on an even undersurface with stationary turbulence (Loos et al. 2003). Compared with other mechanistic models, this model has fewer variables and uses global radiation, wind speed, and wind direction as input; these can easily be obtained from the national meteorological network. As a result, the function of the model is extended, which enables the model to be extrapolated over a wide range of conditions. In contrast, the statistical model based on observed data generally lacks wide applicability, and other mechanistic models have more variables, which results in complex solutions (Dupont et al. 2006).

In the process of model application, local meteorological data recorded by 101 stations during the period 1997 to 2010 were collected and used as input to calculate the MTDs at different locations in the spring maize-growing region of Northeast China. Results indicated that the range of $\mathrm{MTD}_{1\%}$ and $\mathrm{MTD}_{0.1\%}$ varied from 10 to 49 m and 17 to 125 m, respectively. We also surveyed internationally published data on maize gene flow and discovered that $\mathrm{MTD}_{1\%}$ is 3, 6, 28, 35, or 60 m (Ma et al. 2004; Bannert et al. 2008; Goggi et al. 2007; Della Porta et al. 2008; Di and Liu 2008; Lu et al. 2005) and $\mathrm{MTD}_{0.1\%}$ is 100, 112, or 119 m (Goggi et al. 2007; Di and Liu 2008; Lu et al. 2005). Data obtained from non-GM maize surrounding a large-scale commercial GM maize-planting area show that $\mathrm{MTD}_{1\%}$ and $\mathrm{MTD}_{0.1\%}$ are less than 30 and 80 m, respectively (Pla et al. 2006; Weekes et al. 2007). The data predicted by our model generally coincide with these internationally published data.

Because seed purity higher than a specific threshold value is commonly used as a standard in conventional agricultural practice, we used 1 and 0.1 % to calculate MTDs by use of the model. This means that gene-flow frequency should be equal to or below 1 or 0.1 % at distances beyond $\mathrm{MTD}_{1.0\%}$ or $\mathrm{MTD}_{0.1\%}$, respectively. In other words, the 99 or 99.9 % seed purity required for seed production or for basic seed will be met satisfactorily (MOA 2008). It should be noted that some rare gene-flow incidents may occur at longer distances, which will result in fluctuation of outcrossing rate; this is commonly known as a “fat tail” phenomenon. Bannert and Stamp (2007) reported that the $\mathrm{MTD}_{0.01\%}$ in a two-year experiment varied by more than a factor of two (125 m in 2003; 287 m in 2004). Therefore, for GM crops improved with common agronomic traits, we believe it is unnecessary to further increase the threshold value standard from 0.1 to 0.01 %. However, for GM crops with specific safety concerns, for example plants used for pharmaceutical and industrial purposes, more stringent isolation and control measures should certainly be taken to avoid pollen-mediated gene flow, to prevent its derived seed products entering the food chain.

Optimization of model variables

The variables are the key factors determining the accuracy of the model. To improve the accuracy of the model we conducted a series of biological and meteorological observations, during which more than 95,000 data were accumulated. Here, we consider four variables worthy of discussion:

“Rqq” and “a”

The ratio of pollen grains deposited by donor and recipient (*Rqq*), representing pollen quantitative competitiveness, was commonly regarded as a key factor determining gene-flow frequency and/or outcrossing rate in previous studies (Loos et al. 2003; Della Porta et al. 2008). However, we have found that another important biological factor, pollen genetic competitiveness, should also be considered. We therefore introduced a value "a" into our model. For the donor and recipient combination we used ("Pz" and "Ys"), the value of "a" is 17.47, as a result of which the shape of the plot of "Rqq" against "a" is an upward parabola (Fig. 3). If pollen genetic competitiveness is equal for donor and recipient, then theoretically the curve shape will be linear ("a" = 1). In contrast, the curve shape will be a downward parabola if "a" is less than 1. Clearly, when a new combination of donor and recipient is used, the value of "a" should be determined correspondingly. This could be achieved simply by conducting a pollination experiment with a mixture of donor and recipient pollen.

$Q_{dispersed}$ versus $Q_{released}$

Q_{released} is highly variable day by day depending on weather conditions, therefore, hourly data for Q_{released} per plant during the whole flowering period is used as input in the model calculation. In the "Materials and methods" section we defined Q_{released} as two parts, i.e., $Q_{\mathrm{dispersed}}$ and $Q_{\mathrm{deposited}}$. Results show that only 15.82 % of the pollen grains participate in the dispersal process. However, it should be noted that part of the 15.82 % of pollen grains participating in the dispersal process is deposited within the donor area whereas the other part will be dispersed further and deposited outside the donor area. Although a fixed value "R_{d}" is used to express this in our study, we believe the ratio of pollen grains deposited within and outside the donor area may vary, depending on donor size and climatic factors, particularly wind speed (Marceau et al. 2011; Jarosz et al. 2005; Marceau et al. 2012; Viner et al. 2010; van Hout et al. 2008). How climatic factors affect the ratio "R_{d}" for maize should be studied further. With increased donor size, the total amount of pollen deposited inside the donor area will increase. Simultaneously, the pollen grain dispersed outside the donor area will also increase. It is assumed that when equilibrium is reached (we refer to this as "critical size of pollen donor" or "critical strength of pollen source") the amount of pollen dispersed outside the donor area will remain unchanged. It may explain why, in some experiments, the gene-flow frequency downwind is no longer increased when the size of the donor is increased to a specific level (Di and Liu 2008; Ma et al. 2004). In our study, the radius of pollen donor size is 15 m in the Gongzhuling experiment and 10 m in the Lishui experiment. The critical size of the pollen donor for maize must be further studied in the future.

Pollen longevity

Pollen longevity is one of the determinants affecting fertilisation ability and seed setting at longer distances (Luna et al. 2001; Aylor et al. 2003). According to our measurement, the half life of maize pollen ranges from 52 to 155 min at 20–35 °C and 60–90 % relative humidity (unpublished, Jiang). It is much longer than that of rice pollen, 3–5 min (Koga et al. 1971). For rice, therefore, pollen longevity is an important model variable whereas for maize it is not a crucial factor limiting gene flow. For this reason, the longevity of maize pollen is ignored in our model.

Atmospheric stability

The Gaussian model is perhaps the most commonly used type of model for predicting the dispersion of particulate plumes. The width and height of Gaussian plumes depend on the atmospheric dispersal variables, as affected by weather conditions. Commonly, these variables are determined by means of site-specific measurement. However, this is not readily applicable on a large-scale because the measurements cannot be provided by local meteorological stations. An alternative that is often adequate is to use wind velocity and global radiation to predict atmospheric dispersal variables. However, the precision of this method should be further improved by future study. Moreover, the Gaussian model assumes that particulate dispersion has a Gaussian distribution, meaning that particulate distribution has a normal probability distribution. However, this assumption is not valid for uneven surfaces, for example windbreaks. Coupled with its dependence on steady-state meteorological conditions, use of this model is highly limited.

As mentioned above, several variables, for example critical donor size and the of value "a", must be optimized further. The question of whether the model can be used practically with an extrapolation function may be raised. To answer this question, first, it should be emphasized that the model was established by using field experiment data from Gongzhuling (124°49′22″ N, 43°30′17″ E) and validated by use of data generated from Lishui (119°10′36″ N, 31°34′49″ E), which is 1,400 km from the former site. This demonstrates the model itself has the function of extrapolation. Second, on the basis of the data of this study, there are large differences among $\mathrm{MTD}_{0.1\%}$ for 101 sites during 14 years, i.e. 125–17 m = 108 m, indicating that the spatial distribution of MTDs differs location by location. Although MTDs may change to some extent after optimization of variables, the features of the spatial distribution of MTDs are unlikely to change. In conclusion, we believe that the model and the methodology we have reported in this paper are useful and suitable for prediction of gene flow.

Acknowledgments This work is part of the project "GMO biosafety evaluation" supported by the "National Key Programs of China, Breeding of GMOs" (2009ZX08011-019B, 2014ZX08011-001) and the "Foundation of Jiangsu Key Laboratory for Agricultural Meteorology" (KYQ1203). Partial financial support was also obtained from Jiangsu Higher Education Institutions (PAPD) and the Ministry of Education of China (grant PCSIRT). We are grateful to Professor Fen-fen Feng for her technical assistance in observation of biological characteristics, to Ms Bo Pang, Ms Yexian Zhou, and Mr Peijian Shi for helping with observations, and to Professor Chuangen Lü for supporting the field experiment at Lishui. Kind help and suggestions by Professor Qianhua Yuan during preparation of the manuscript are also gratefully acknowledged.

References

Allnutt TR, Dwyer M, McMillan J, Henry C, Langrell S (2008) Sampling and modeling for the quantification of adventitious genetically modified presence in maize. J Agric Food Chem 56:3232–3237

Angevin F, Klein EK, Choimet C, Gauffreteau A, Lavigne C et al (2008) Modelling impacts of cropping systems and climate on maize cross-pollination in agricultural landscapes: the MAPOD model. Eur J Agron 28:471–484

Arritt RW, Clark CA, Goggi AS, Lopez Sanchez H, Westgate ME et al (2007) Lagrangian numerical simulations of canopy air flow effects on maize pollen dispersal. Field Crop Res 102:151–162

Arya SP (1999) Air pollution meteorology and dispersion. Oxford University Press, Oxford

Aylor DE (2002) Settling speed of corn (*Zea mays*) pollen. J Aerosol Sci 33:1601–1607

Aylor DE, Schultes NP, Shields EJ (2003) An aerobiological framework for assessing cross-pollination in maize. Agric For Meteorol 119:111–129

Bannert M, Stamp P (2007) Cross-pollination of maize at long distance. Eur J Agron 27:44–51

Bannert M, Vogler A, Stamp P (2008) Short-distance cross-pollination of maize in a small-field landscape as monitored by grain color markers. Eur J Agron 29:29–32

Chamecki M, Meneveau C, Parlange MB (2009) Large eddy simulation of pollen transport in the atmospheric boundary layer. J Aerosol Sci 40:241–255

Coléno FC, Angevin F, Lécroart B (2009) A model to evaluate the consequences of GM and non-GM segregation scenarios on GM crop placement in the landscape and cross-pollination risk management. Agric Syst 101:49–56

Della Porta G, Ederle D, Bucchini L, Prandi M, Verderio A et al (2008) Maize pollen mediated gene flow in the Po valley (Italy): source–recipient distance and effect of flowering time. Eur J Agron 28:255–265

Di H, Liu ZJ (2008) Gene flow of bar transgenic maize (*Zea mays* L.). Chin Agric Sci Bull 24:111–113

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

Duncan DR, Widholm JM (2004) Osmotic induced stimulation of the reduction of the viability dye 2,3,5-triphenyltetrazolium chloride by maize roots and callus cultures. J Plant Physiol 161:397–403

Dupont S, Brunet Y, Jarosz N (2006) Eulerian modelling of pollen dispersal over heterogeneous vegetation canopies. Agric For Meteorol 141:82–104

Goggi AS, Caragea P, Lopez-Sanchez H, Westgate M, Arritt R et al (2006) Statistical analysis of outcrossing between adjacent maize grain production fields. Field Crop Res 99:147–157

Goggi AS, Lopez-Sanchez H, Caragea P, Westgate M, Arritt R et al (2007) Gene flow in maize fields with different local pollen densities. Int J Biometeorol 51:493–503

Gustafson DI, Brants IO, Horak MJ, Remund KM, Rosenbaum EW et al (2006) Empirical modeling of genetically modified maize grain production practices to achieve European Union labeling thresholds. Crop Sci 46:2133–2140

Ivanovska A, Todorovski L, Debeljak M, Dzeroski S (2009) Modelling the outcrossing between genetically modified and conventional maize with equation discovery. Ecol Model 220:1063–1072

James C (2011) Global status of commercialized Biotech/GM crops: 2011. ISAAA, New York

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

Jarosz N, Loubet B, Durand B, Foueillassar X, Hubert L (2005) Variations in maize pollen release and deposition in relation to microclimate. Environ Sci Technol 39:4377–4384

Jiang WM, Sun JN, Cao WJ, Jiang RB (2004) Introduction to air pollution meteorology. China Meteorological Press, Beijing

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

Koga Y, Akihama T, Fujimaki H, Yokoo M (1971) Studyies on the longevity of pollen grain of rice, *Oriza sativa* L. I. Morphological change of pollen grains after shedding. Cytologia 36:104–110

Loos C, Seppelt R, Meier-Bethke S, Schiemann J, Richter O (2003) Spatially explicit modelling of transgenic maize pollen dispersal and cross-pollination. J Theor Biol 225:241–255

Lu XB, Sun HW, Yang CL, Liu KQ, Shang YF et al (2005) Gene flow of transgenic corn to cultivated relatives in China. Acta Ecol Sin 25:2450–2453

Luna VS, Figueroa MJ, Baltazar MB, Gomez LR, Townsend R et al (2001) Maize pollen longevity and distance isolation requirements for effective pollen control. Crop Sci 41:1551–1557

Ma BL, Subedi KD, Reid LM (2004) Extent of cross fertilization in maize by pollen from neighbouring transgenic hybrids. Crop Sci 44:1273–1282

Marceau A, Loubet B, Andrieu B, Durand B, Foueillassar X et al (2011) Modelling diurnal and seasonal patterns of maize pollen emission in relation to meteorological factors. Agric For Meteorol 151:11–21

Marceau A, Saint-Jean S, Loubet B, Foueillassar X, Huber L (2012) Biophysical characteristics of maize pollen: variability during emission and consequences on cross-pollination risks. Field Crop Res 127:51–63

Menteith JL, Unsworth MH (1990) Principles of environmental physics. Edward Arnold, London

Pla M, La Paz JL, Penas G, Garcia N, Palaudelmas M et al (2006) Assessment of real-time PCR based methods for quantification of pollen-mediated gene flow from GM to conventional maize in a field study. Transgenic Res 15:219–228

Pleasants JM, Hellmich RL, Lewis LC (1999) Pollen deposition on milkweed leaves under natural conditions. In: Monarch butterfly research symposium, Chicago

Sanvido O, Widmer F, Winzeler M, Streit B, Szerencsits E et al (2008) Definition and feasibility of isolation distances for transgenic maize cultivation. Transgenic Res 17:317–335

Sears MK, Stanley-Horn D (2000) Impact of Bt corn pollen on monarch butterfly populations. In: Fairbairn C, Scoles G, McHughen A (eds) Proceedings of the 6th International Symposium on the Biosafety of Genetically Modified Organisms. University Extension Press, Canada, pp 120–130

Slinn WGN (1982) Predictions for particle deposition to vegetative surfaces. Atmos Environ 16:1785–1794

van Hout R, Chamecki M, Brush G, Katz J, Parlange MB (2008) The influence of local meteorological conditions on the circadian rhythm of corn (*Zea mays* L.) pollen emission. Agric For Meteorol 148:1078–1092

Viner BJ, Westgate ME, Arritt RW (2010) A model to predict diurnal pollen shed in maize. Crop Sci 50:235–245

Weekes R, Allnutt T, Boffey C, Morgan S, Bilton M et al (2007) A study of crop-to-crop gene flow using farm scale sites of fodder maize (*Zea mays* L.) in the UK. Transgenic Res 16:203–211

Yao K, Hu N, Chen W, Li R, Yuan Q et al (2008) Establishment of a rice transgene flow model for predicting maximum distances of gene flow in southern China. New Phytol 180:217–228