Vol. 15, pp. 235–250 (2018)
doi: 10.2142/biophysico.15.0_235

Biophysics and Physicobiology

https://www.jstage.jst.go.jp/browse/biophysico/

Regular Article

Data analytic study of the homothermal maintenance mechanism of *Skunk Cabbage*: Capturing pre-equilibrium characteristics using extended poisson model

Shuji Kawasaki[1] and Kikukatsu Ito[2,3]

[1]*Faculty of Science and Engineering, Iwate University, Morioka, Iwate 020-8550, Japan*
[2]*Faculty of Agriculture, Iwate University, Morioka, Iwate 020-8550, Japan*
[3]*Agri-Innovation Research Center, Iwate University, Morioka, Iwate 020-8550, Japan*

Received May 18, 2018; accepted September 4, 2018

A wild plant called *Skunk Cabbage* is known to heat itself and keep its body warm before spring. We study its homothermal maintenance mechanism from a mesoscopic point of view. We take the increment process of the temperature time series and consider it as 'elastic' force that always tries to backlash its temperature to an intrinsic target temperature. We then propose a kind of extended Poisson distribution for the model of the 'elastic' force. The hypothesis testing result by Kolmogorov-Smirnov test suggests that the proposed distribution is a plausible candidate of the model for the 'elastic' force, on the temperature range in which the system is in equillibrium. In addition, it turns out that the parameters in the model captures well the linear behaviour of the expectations of the 'elastic' force at each of the present temperatures and similarly, the constancy of the variances of the force. Especially, the linearity of the expected increments over displacements of tempertures indicates that the backlash might be considered to be like the elastic force of a spring as described by Fuch's law.

Key words: homothermal maintenance, exothermic and endothermic reactions, extended Poisson distribution, least-squares fitting, Kolmogorov-Smirnov test

Correspondinng author: Shuji Kawasaki, Faculty of Science and Engineering, Iwate University, 3-18-8, Ueda, 020-8550 Morioka, Japan. e-mail: shuji@iwate-u.ac.jp

Background and Problem

Skunk Cabbage is a plant of the family Araceae that grows wildly in cold marshes. It has the special feature of being equipped with a heating organ called a *spadix*. The organ generates heats continuously for several days before spring and retains its temperature at about 22~24°C, in order to nurture its seeds effectively [1]. We have been interested in the temperature control mechanism and have obtained partial results on the mechanism from various points of view. Microscopically, we have shown from the molecular or cellular biological points of view that particular molecules governing mitochondorial respiration in exothermic cells perform an important function in heat generation [2–5].

Regarding the temperature control, or more specifically, the homothermal maintenance, we have found that the spadix temperature and the amount of respiratory metabolism have an inverse correlation. In the temperature range in which the homothermal maintenance works well, the lower (resp., higher) the spadix temperature is, the higher (resp., lower) the respiratory activity becomes. We thus have come to believe that the maintenance is based on so called *Le Chatelier*'s principle [6, p. 150]: when the equilibrium state of a system is subjected to a disturbance, the composition of the system adjusts to tend to minimise the effect of the disturbance. We will consider this principle with the distur-

◀ **Significance** ▶

As a temperature control mechanism of the *Skunk Cabbage*, we have found that, around the Central Temperature (CT), the operation to rebound the temperatures to the CT is done linearly with respect to temperature displacements from the CT in average; also, variance takes a constant value there. Both sample mean and sample variance show good agreement with their model parameters of the proposed mixed Poisson model.

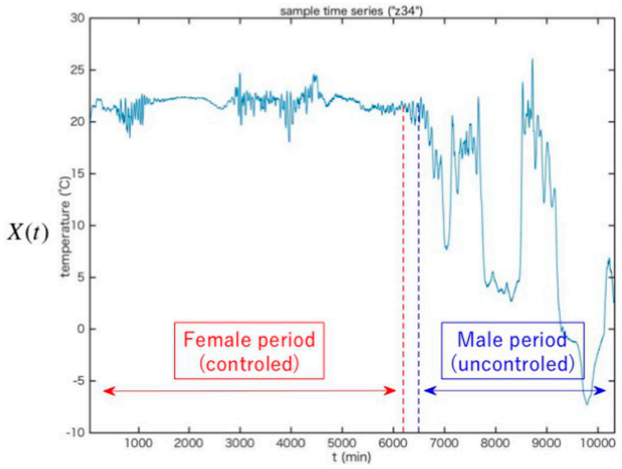

Figure 1 A sample of time series data $X(t)$ of spadix temperatures.

bance being temperature change.

In addition, macroscopically, we have revealed from the applied mathematics point of view that the time series data of the spadix temperature presents a chaotic behaviour [7,8]. In particular, the temperature control mechanism has an intrinsic dynamics characterised by a 'Zazen Attracter' [7]. It is the first chaotic phenomenon discovered in plants [9–11].

Figure 1 shows sample data of the spadix temperature time series. Its first part presents the homothermal maintenance stage (the female stage) in which the temperature control is performed effectively. In the latter part, the control no longer works well. The phase of the control system thus changes.

From these results, the elements of the microscopic and macroscopic mechanisms have become gradually become understood. However, knowing the microscopic genes or enzymes committed to the mechanism does not lead immediately to an understanding of the macroscopic chaos, and vice versa; at present, only fragmentary elements have been obtained, but these have not been linked organically. Consequently, a complete understanding of the mechanism remains to be acquired. We do not know the relationship of the chaos and the homothermal maintenance.

On the other hand, a plant does not have nervous system in general, and thus can neither sense ambient temperature nor provide temperature information as feedback to the nervous system to determine the control at each instant. That is, the mechanism must be a unidirectional one that achieves the appropriate control of the spadix temperature in a simple manner. However, with the present situation, as described above, we do not yet have a higher standpoint to be able to infer the biological strategy of the Skunk Cabbage for the homothermal maintenance.

Given this situation, we investigate the mechanism from another point of view in this paper. We analyse the time series data of the spadix temperature to determine the kinds of behaviour the plant tends to present when the ambient temperature falls or rises. Thus, we might be confronted with the problem of how an 'elastic' force is working when the spadix temperature falls to lower or rises to higher than its target temperature. Does it resemble a simple harmonic oscillation?

Of course, it cannot be a simple harmonic oscillation, because the 'elastic' force in the spadix is not of constant periodicity. What kind of principle is behind the background? Such arguments might capture what the microscopic phenomena leads to, and also, what causes the macroscopic phenomena; thus, we might need some mesoscopic study that bridges the micro- and macroscopic phenomena.

In this paper, by analysing the above-mentioned 'elastic' force, we find that the control of homothermal maintenance is always achieved in accordance with the balance of the exothermic and endothemic potentials. When the temperature has fallen (or, risen), even though the 'elastic' force to raise (resp., lower) the temperature is stronger, the opposite force to lower (resp., raise) the temperature exists the same time. This conflict can be considered to be necessary in order to enable the unidirectional temperature control without feedback.

As a result, to describe the 'elastic' force in terms of the greater and lesser strength of exo- and endothermic potentials, we propose a mixed Poisson model. On this basis, we obtain another understanding of the mechanism through which the 'elastic' force is generated. Essentially, it resembles seasaw. Having a higher potential at one of the two states automatically implies the state changes to another (opposite) one of lower potential. In addition, the higher or lower the spadix temperature becomes beyond the target temperature, the stronger the 'elastic' force to return the temperature to the target becomes; see Materials and Methods section. Indeed, the average behaviour of the 'elastic' force, as a function of the present temperature, fits well with its theoretical values derived from the Poisson model. This might also imply the plausibility of the model.

At the end of this introductory section, we remark on the basic scope of this study. Although we analyse the time series data like in Figure 1 above, we do not consider fitting a time series model to the data. Such model fitting might be considered in the future, but we do not intend to do so, at least not in a classical way. This is because, as mentioned in the scientific articles above, the homothermal maintenance mechanism is considered to be driven according to chaos, which is a nonlinear system. In fact, in our preliminary investigation, the data have presented a sharp self-similarity (see [12,13] for self-similarity). Thus, we do not expect that conventional linear models like autoregressive (AR) and autoregressive moving-average (ARMA) to work for the data here. However, nonlinear models like fractionally integrated ARMA [12,13] might be candidates of the model. But, rather, our purpose is not just the time series modeling. Our final goal, though, is not just the time series modeling,

which is an identification of the time series class, but to reveal the homothermal maintenance mechanism, including the relationship with chaos dynamics, of the mysterious yet charming plant.

This goal requires approaching from several points of views as described above, and a key concept is the phase transition model with a linear system of equilibrium state (we suppose that a neighbourhood of the central temperature (CT) corresponds to this state) and a nonlinear system of nonequilibrium state (outside of the CT neibourhood might corresponds to this state). Indeed, it will be revealed in this paper that the time series representing an equilibrium state shows a linear behaviour on average in a neighbourhood of the CT, indicating that the behaviour of the equilibrium state is linear.

The remainder of the paper is organised as follows. We begin with explaining the basic facts and overviewing the 'elastic' force in Materials and Methods section. The extended Poisson model is proposed, along with the distribution of the 'elastic' force, in Results and Discussion section, and then the model is validated through hypothesis testing. We then show that two characteristics of the average behaviour are captured well by the Poisson parameters. Conclusion section presents concluding remarks.

Materials and Methods

Time series data of spadix temperatures

We have several time series data of spadix temperatures, collected in the field, like Shizukuishi, Iwate prefecture, Japan (see Supplementary Tables S1). The dataset and MATLAB program codes used in the study are uploaded at ResearchGate [14,15] and are open for free access.

Let us denote the spadix temperature by $X(t)$, $t \geq 0$ and let $\boldsymbol{X}$ be its value range. Figure 1 presents a sample of $X(t)$. As seen in the figure, the spadix temperature has roughly remained constant throughout the female period. It is reported [16] that the respiratory metabolic system of the plant is activated when the spadix temperature falls, and inactivated when the temperature rises; in addition, the lower (resp., higher) the temperature falls (rises), the stronger (weaker) the activity becomes. The activity is thus inversely correlated with the temperature increases. $\boldsymbol{X}$ is 15~30°C in broadest case, but typically a little bit less range may be supposed in this paper.

The activation is considered to be the operation that raises the temperature, and the inactivation the operation that lowers it. We have thus speculated that during the female period, there may be a critical point of temperature $x = x_0 \in \boldsymbol{X}$ such that if $x < x_0$ (resp., $x > x_0$) then an action works so as to raise (lower) the temperature. Let us call such an x_0 the CT.

Let $y(t) = X(t+1) - X(t)$ be increments of $X(t)$. The thermal sensor is a needle-like device and is connected to a digital thermometer. The sensor is inserted into the spadix of each of the individuals, and the thermometer displays the temperatures in 0.1°C steps. Thus $X(t)$ and $y(t)$ are in 0.1°C steps in this paper. The sensors continue to measure the temperature data every minute and record it in micromemory for several days. Figure 1 presents a sample time series of $X(t)$.

We are interested in $y(t)$ rather than $X(t)$ itself. As mentioned above, the plant tries to raise (lower) its temperature when the present temperature is lower (higher) than the CT x_0. The strength of the backlash may depend on the displacements $(X(t) - x_0)$ and the amount of backlash may appear as the jumps $y(t)$; What embodies the action, backlash from the present temperature x toward x_0, is just the $y(t)$. Thus we will investigate the temperature control mechanism through $y(t)$, rather than through $X(t)$ itself. In addition, the formation of the histogram of $y(t)$ is in fact better than that of $X(t)$.

Figures 2 (a) and (b) are the histograms of $X(t)$ and $y(t)$ of the same sample, respectively. For this sample, the value range of $X(t)$ is approximately 18~25°C; the range depends on the individuals, with some being 21~27°C and others 20~25°C, for example. All of the samples used in this study are taken from the same place and same period (data lengths are different). As for $y(t)$, every sample has zero median and mode, and the concentricity of the histogram is high.

The histogram of $y(t)$ in Figure 2 (b) is nearly symmetric. The jumps $y(t)$ to higher temperatures from lower and to lower temperatures from higher have similar value frequencies. Thus, the 'elastic' force for both directions are basically the same, on the whole.

Overview of the 'Elastic' force

We would like to characterise the temperature dependence of $y(t)$ on the present temperatures x or $|x - x_0|$. This leads to

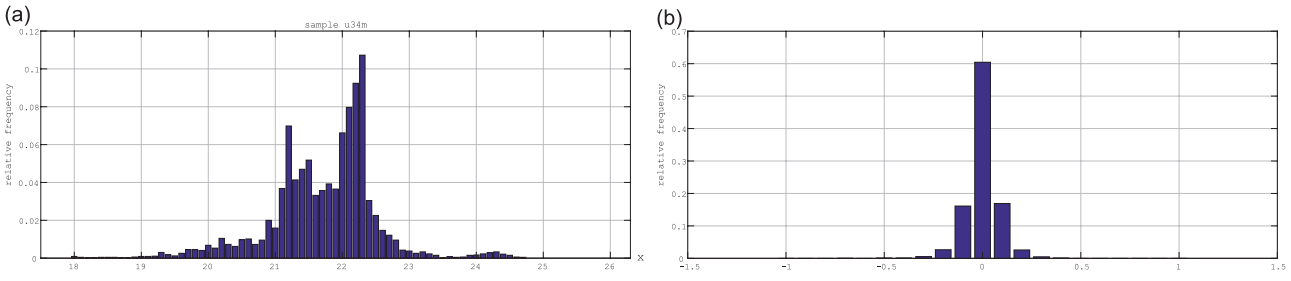

Figure 2 Histogram of $X(t)$ (top) and $y(t)$ (bottom).

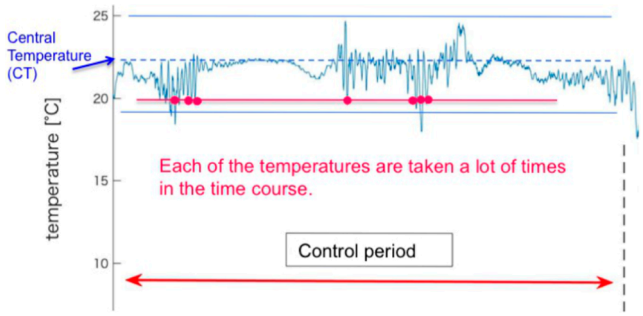

Figure 3 Investigate tendency of increments $Y(x)$.

considering the tendency on the set of values

$$Y(x) \triangleq \{y(t)|X(t)=x,\ t\in\boldsymbol{T}\},\quad x\in\boldsymbol{X} \tag{1}$$

where $\boldsymbol{T}$ is the female period and $\boldsymbol{X}$ is the value set of $X(t)$ for $t\in\boldsymbol{T}$. An image of the $Y(x)$ is shown in Figure 3. In the time series data, we determine $\boldsymbol{T}$ by sight. We often call the elements of $Y(x)$ 'elastic force' at x, since they correspond to the action on the present temperatures $X(t)=x$ to backlash to CT.

We draw the histogram of $Y(x)$ for each of the x values to recognise the tendency visually. However, they are field data of natural phenomena and are often noisy, with the result that it is difficult to recognise the tendency in a straightforward way. Here we have only one sample time series for each of the individual plants. We then consider taking a kind of averaged data by considering the following $Y_\Delta(x)$ instead of $Y(x)$: for $\Delta=0$, 0.1, 0.2, ...

$$Y_\Delta(x) \triangleq \{y(t)|\,|X(t)-x|\le\Delta,\ t\in\boldsymbol{T}\}. \tag{2}$$

Taking $\Delta=0$ reduces $Y_\Delta(x)$ to $Y(x)$. The data $X(t)$ we are using in this paper have lengths of thousands. For every x, $Y_\Delta(x)$ consists of tens to hundreds of elements (In data analysis, we consider $Y_\Delta(x)$ only for those values of x for which $Y_\Delta(x)$ contains more than 20 elements). As in the case of $Y(x)$, we call $Y_\Delta(x)$ the 'elastic force' in the neighbourhood of x.

In Figures 4 (a)~(c), the histograms of $Y_\Delta(x)$ with $\Delta=0.4$ of the sample same as in Figures 1 and 2 are depicted. Figure 4 (a) is for those values of x that are less than the CT, $x=19.8$, 20.4, 20.8 and 21.0°C. Similarly, (b) is for values close to the CT, $x=21.6$, 22.0, 22.2, 22.4°C, and (c) is greater values, $x=22.6$, 22.8, 23.0, 23.2°C. As a result of the 'averaging' over $x\pm0.4$, noise relating to the original $Y(x)$ is suppressed somewhat, revealing the above tendency more clearly.

With $Y(x)$ or $Y_\Delta(x)$ so defined, we now define the CT. We set it to be a particular value $x=x_0\in\boldsymbol{X}$ such that $|\mathbf{E}[Y(x)]|$ or $|\mathbf{E}[Y_\Delta(x)]|$ are minimal. Typically, it is $\simeq0$. Actually, a very precise definition of the CT may not be necessary, for the time being. Later a method of determining the CT in a more objective way will be suggested. See Results and Discussion section.

Of the histograms in Figure 4, those in (b) are mostly symmetric and have high concentricity, in contrast to those in (a) and (b). The histograms in (a) have the positive part superior to the negative part, and conversely, those in (c) have the negative part superior to the positive part; in both cases, the concentricity is not as high as in (b).

In (a), the plant might try to raise the temperature to the CT from below as much as possible, resulting in a distribution that is biased to the right with high dispersion. In (b), the plants tend to remain in their current states close to the CT and do not jump substantially,with the result that the concentricity around the origin is high. In (c), the plant might attempt to lower the temperature to the CT from above as much as possible, resulting in a distribution that is biased to the left with high dispersion.

Using a boxplot can be helpful, in general, in understanding the rough outline of a distribution, but in the present case, this does not work. While the expected values $\mathbf{E}[Y_\Delta(x)]$ are very close to zero for every x, standard deviations are relatively large, resulting in boxes that appear to be similar. It is for that reason that we examine the histograms here, as in Figure 4.

Results and Discussion

Model of the 'elastic' force: extended poisson distribution

In the previous section, we recognised the temperature control visually from the 'elastic' force towards the CT in the female period. This result might suggest that

- The larger the displacement of the present temperature from the CT is, the stronger the 'elastic' force is.

This might imply that the control system has a potential energy, which is thought to be a function of displacements $|x-x_0|$, of the 'elastic' force, and the greater the displacements are, the greater the energy become *automatically*. In this section, we consider modelling of the 'elastic' force for each value of x.

First, let us take note of an important observation regarding the 'elastic' force. In the histogram in Figure 5, we see that when the present temperature is in less than the CT, not only does the 'elastic' force raise the temperature towards the CT, but also, the force to fall further is in effect exists at the same time (although the latter is weaker than the former). Similarly, when the present temperature is greater than the CT, both the 'elastic' force to raise the temperature and the force to lower it are in effect. As for the coexistence of the 'elastic' force for both directions, we recall the chemical reactions taking place in the spadix [16]. This is an exothermic reaction generating heats through a dehydrogenase reaction changing starch absorbed from the roots into protons and water. This is, in fact, a reversible reaction, i.e. the endothermic reaction also occurs at the same time. These reactions are represented by the following expression:

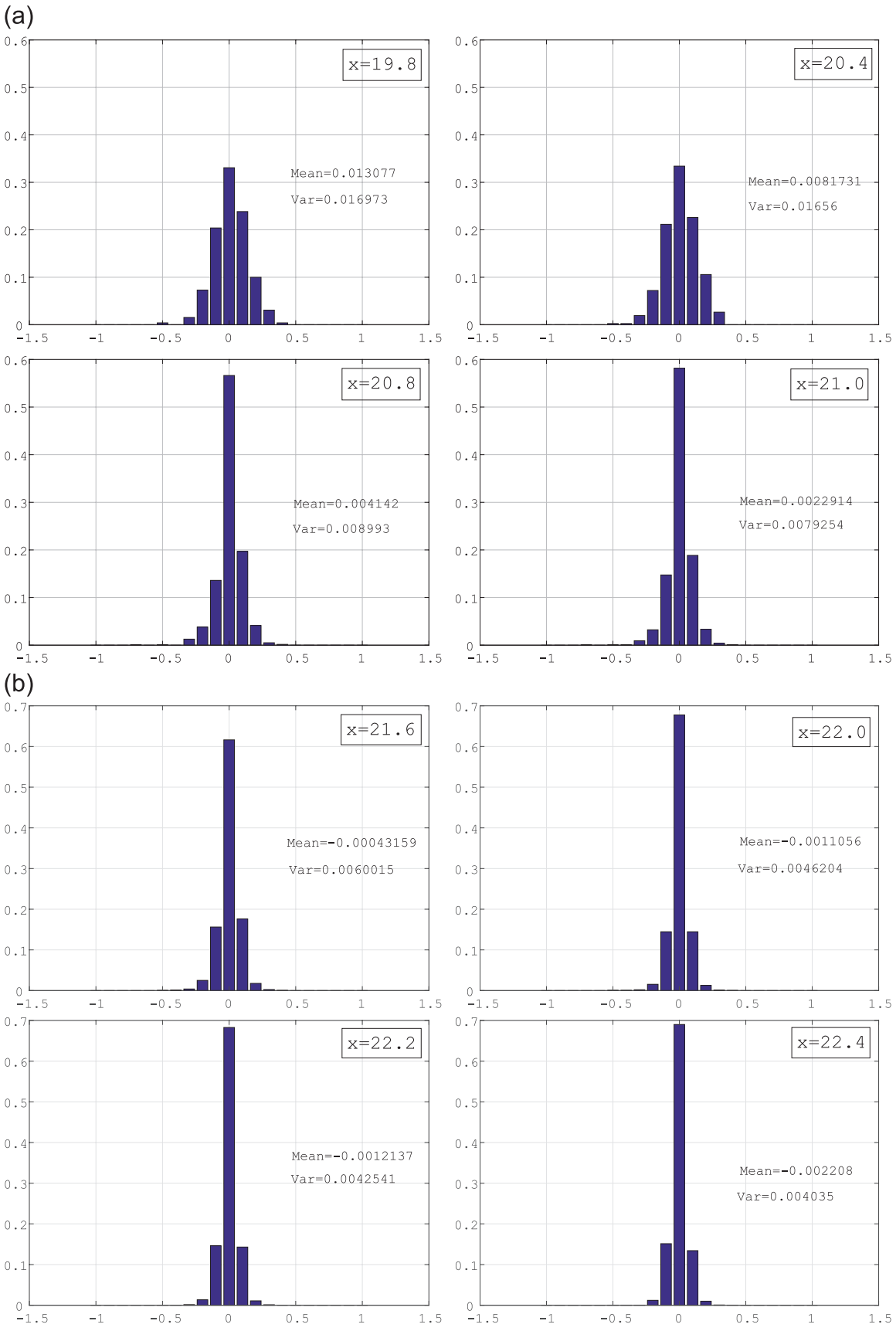

Figure 4 Histogram of $Y(x)$ for several x, $\Delta = 0.4$. (a) lower side than CT. (b) neighbourhood of CT.

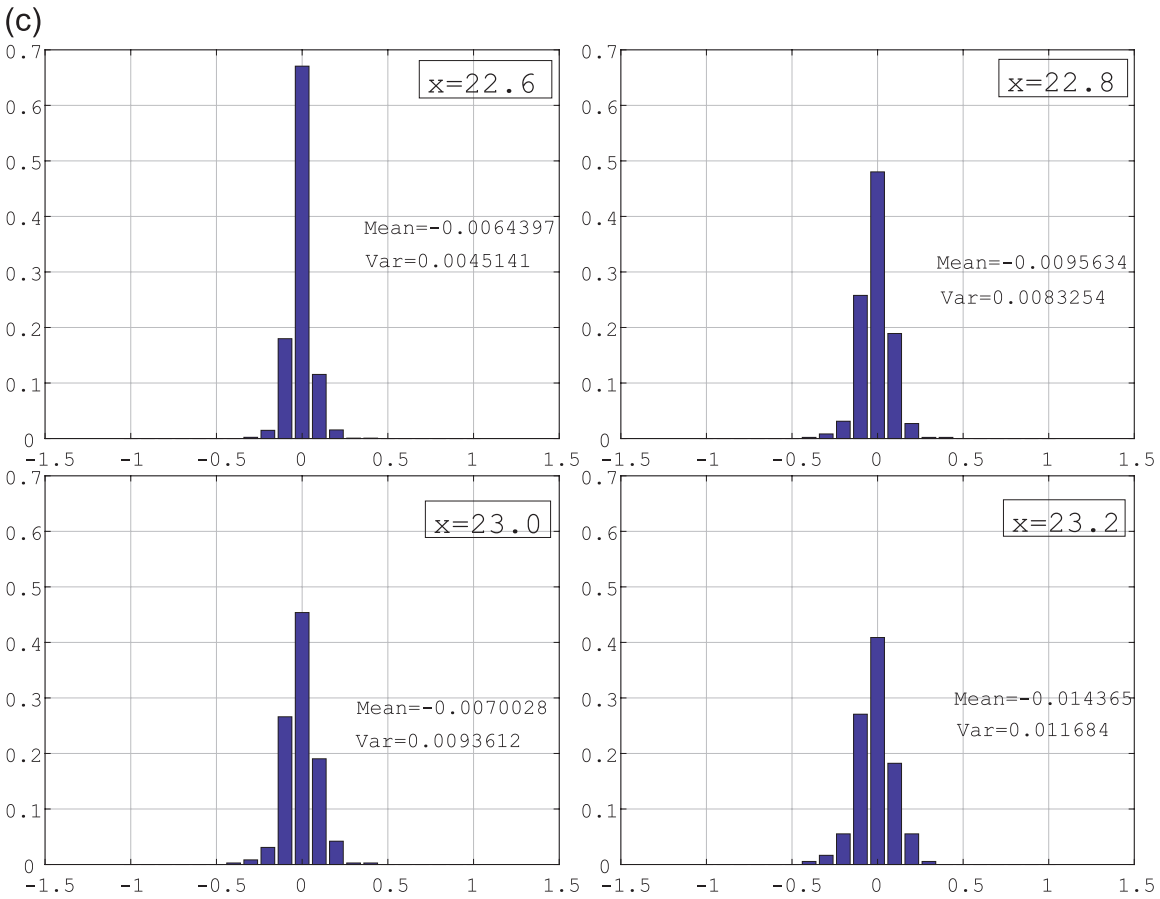

Figure 4 (c) higher side than CT.

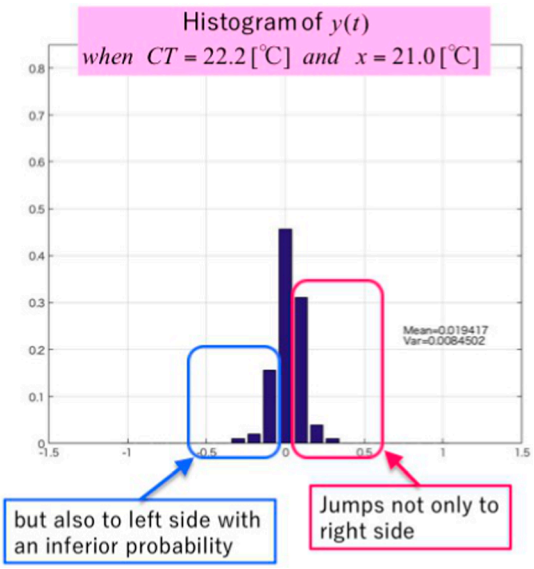

Figure 5 Histogram of $Y(x)$, $x = 21.0$°C and CT = 22.0°C.

$$RH_2 \underset{k_1'}{\overset{k_1}{\rightleftharpoons}} R + 2H^+ + 2e^- \xrightarrow[+\frac{1}{2}O_2]{k_2} H_2O, \tag{3}$$

where RH_2 represents a carbohydrate (starch) for a respiratory substrate R, H^+ is a proton, and e^- is an electron. The first half of the expression is the reversible reaction, with the reaction constants k_1 for exothermicity and k_1' for endothermicity. The plant thus generates heat by decomposing starch (energy source) absorbed from the roots. The latter half reaction, finally generating water, is an irreversible reaction. In this paper, we are thinking of the modeling only for the first half, the state referred to as pre-equilibrium.

Our assumptions for pre-equilibrium modeling are thus as follows:

- The exo- and endothermic reactions are taking place at the same time. The total increase or decrease of the spadix temperature is a result of balancing of the two reaction intensities. Depending on the magnitudes of the displacements from the CT, the 'elastic' force in opposite directions towards the CT change.

Further:

- When x is less (resp., greater) than the CT, the 'elastic' force works at a constant rate so as to restore the equilibrium state, i.e. to return to the CT, but owing to randomness, the force fluctuates around the rate.

Based on the above assumptions, we considered the following model of spadix temperature distribution caused by exo- and endothermicity coexistence and their balancing. Note that the temperature data are in 0.1°C steps.

Let $k=0, 1, 2, \ldots$. Letting the temperature be raised by 0.1°C degrees results in particular virtual exothermic factors being generated in appropriate units, and letting the temperature be lowered by 0.1°C degrees results in particular virtual endothermic factor being generated in appropriate units.

Then, if the virtual exothermic factors are generated in k units, the temperature increases by $0.1\times k$°C degrees, and if the virtual endothermic factors are generated in k units, the temperature decreases by $0.1\times k$°C degrees. We denote the generation rates of the factors by $\lambda_1=\lambda_1(x)$ and $\lambda_2=\lambda_2(x)>0$ [mol/s] respectively. Our model then assumes that the numbers of units of generation per unit time obey a Poisson distribution of intensity λ_1 and λ_2, respectively:

$$p_1(k,x)=P\left(\begin{array}{l}k \text{ units of amount of the}\\ \text{exothermic factor generated}\end{array}\ \middle|\ X(t)=x\right)$$
$$=\frac{[\lambda_1(x)]^k}{k!}e^{-\lambda_1(x)}, \tag{4}$$

$$p_2(-k,x)=P\left(\begin{array}{l}k \text{ units of amount of the}\\ \text{endothermic factor generated}\end{array}\ \middle|\ X(t)=x\right)$$
$$=\frac{[\lambda_2(x)]^k}{k!}e^{-\lambda_2(x)} \tag{5}$$

for $k=0, 1, 2, \ldots$. In general, a Poisson distribution represents the probabilities of the number k of events occuring in a unit time, and the Poisson intensity $\lambda>0$ is the mean of the number of events generated [17,18]. In the present case, we use two Poisson distributions for each of the exo- and endothermic factors. This raises the question of how to represent the number of units of the endothermic factors? If we were to think of only the exo- or endothermic factors independently, we might use only the usual Poisson distribution. Then, however, it would be difficult to represent the exo- and endothermicity coexistence and their balancing. It would thus be useful to define a Poisson distribution such as taking nonpositive values are formally considered. That is, endothermic factor generation is thought of as if exothermic factors are generated in negative amounts.

Then, through a combination of the two Poisson distributions, we represent the total amount of temperature increase or decrease, as in Figure 6. For example, the total temperature increasing by 0.1°C degrees consists of all of the combinatorial patterns (number of exothermicity factors, number of endothermicity factors) $=\{(k+1,k)|k=0,1,2,\ldots\}$,

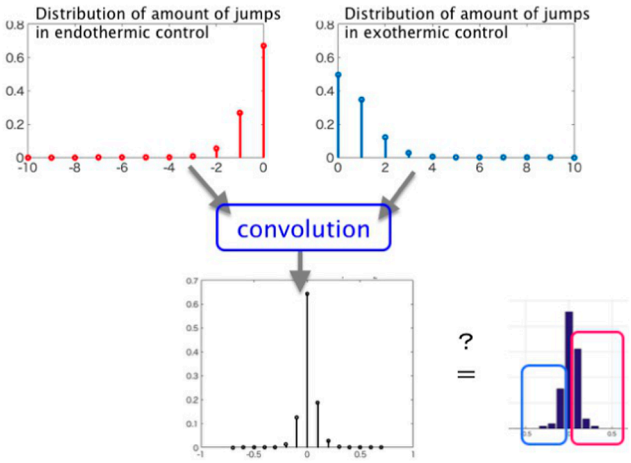

Figure 6 An extended Poisson distribution representing the simultaneous generation of exothermal and endothermal factors.

while the total temperature decreasing by 0.2°C degrees consists of all of the combinatorial patterns (number of exothermicity factors, number of endothermicity factors) $=\{(k,k+2)|k=0,1,2,\ldots\}$. Hence, we are led to consider

$$p_x(k)\triangleq P(y(t)|_{X(t)=x}=0.1\times k)$$
$$=\sum_{l\geq 0}p_1(k+l,x)\,p_2(-l,x),\quad k\in\mathbf{Z}. \tag{6}$$

An example of this extended Poisson distribution with $\lambda_1=0.3$ and $\lambda_2=0.2$ is shown in Figure 6. It can be seen that $\lambda_2<\lambda_1$ corresponds to the greater or lesser probability weights. In addition, the expectation and variance of the extended Poisson distribution is given as follows (see Mean and Variance in Supplementary Text S1):

$$\mathbf{E}_x[y]=\sum_{k\in\mathbf{Z}}(0.1k)\,p_x(k)=0.1[\lambda_1-\lambda_2] \tag{7}$$

$$\mathrm{Var}_x[y]=\sum_{k\in\mathbf{Z}}(0.1k-\mathbf{E}_x[y])^2\,p_x(k)=(0.1)^2[\lambda_1+\lambda_2]. \tag{8}$$

The expectation can take either nonnegative or nonpositive real number values.

The left-hand column of Figure 7 (a) contains the histograms $y_\Delta(x)$, $\Delta=0$ and the fitting results of the extended Poisson distribution for the lower temperatures $x=19.8$, 20.4 and 20.8°C for the same sample as in Figure 1. We derived the fitting for the histogram using the least squares method. The right-hand column of Figure 7 (a) is $y_\Delta(x)$, $\Delta>0$. Similarly, Figure 7 (b) shows the neighbourhood of the CT, $x=21.0$, 21.2 and 21.4°C, and Figure 7 (c) shows for greater values of x, $x=22.4$, 22.6 and 23.0°C. In our search for the best values for λ_1 and λ_2, we take the pitch of their values as 0.005. The data we used in this paper showed no difference for results with smaller pitches.

For the histograms in (b), the fit is not as bad as it might be, i.e. not bad with $\Delta=0$. However, the fit for the histo-

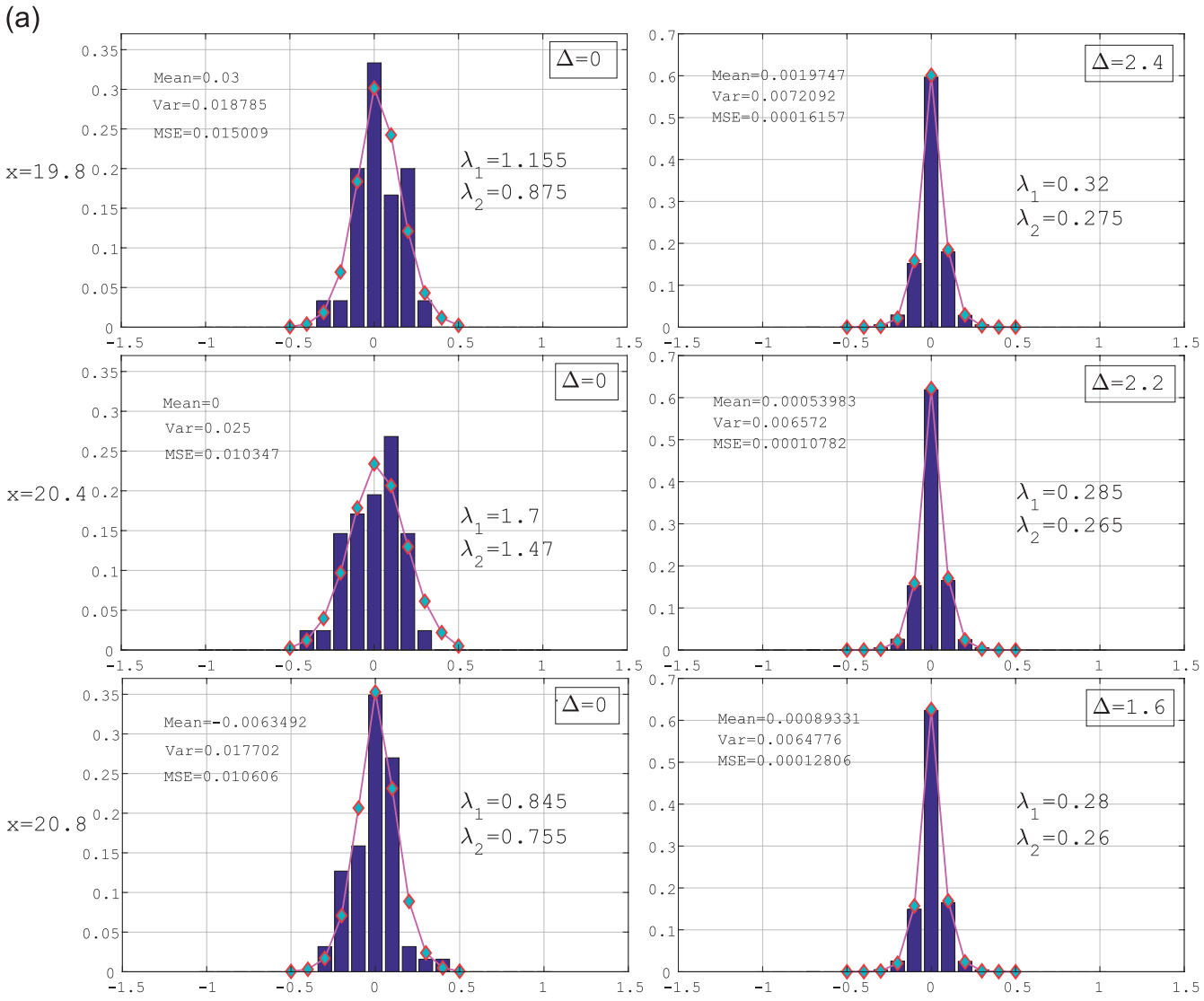

Figure 7 Fitting the extended Poisson dist. to $Y_\Delta(x)$. (a) $x = 19.8, 20.4, 20.8$°C, from top to bottom.

grams in (a) and (c) is not good. Because the data length of $Y_\Delta(x)$, $\Delta = 0$ is not sufficient, the histogram itself presents no meaningful shape. However, for $\Delta > 0$, the histograms show meaningful shapes, and the fits are good. For $\Delta = 0$, we should not say that 'the model fails', but rather, that the histogram shaping itself is not yet perfected.

For $\Delta > 0$, indeed the histogram shaping is better than for $\Delta = 0$, suggesting that we might be able to argue for a particular tendency. However, how are we to consider results drawn from the $y_\Delta(x)$, $\Delta > 0$? Does it have the same tendency as that of $\Delta = 0$?

Unfortunately, we cannot say that the tendency is precisely the same in general. However,we can say that the distribution of $Y_\Delta(x)$, $\Delta > 0$ presents an average behaviour in the neighbourhood of x and an approximate distribution of $Y(x)$. In any event, we can consider the tendency at those values of x, as they are only for values near the CT, as shown in Figure 7 (b). Therefore, we consider $Y_\Delta(x)$, $\Delta > 0$, to model the average behaviour in the neighbourhood of each x. Below we will show that the average behaviour is linear with respect to x, suggesting that the tendency of $\Delta = 0$ and $\Delta > 0$ are close to each other. In addition, the model is at least not rejected when testing the hypothesis as in the next section.

Before going to the Hypothesis testing, let us remark plausibility of the extended Poisson distribution. The temperature of an object may be considered to vary continuously in general, apart from the restriction that measuring devices can record the temperatures only in discrete values. On the other hand, the Poisson distribution is a probabilistic model for counting data taking nonnegative integer values. Nevertheless, we think the modeling the temperature displacements with the Poisson-like distribution is meaningful, due to the following reasons:

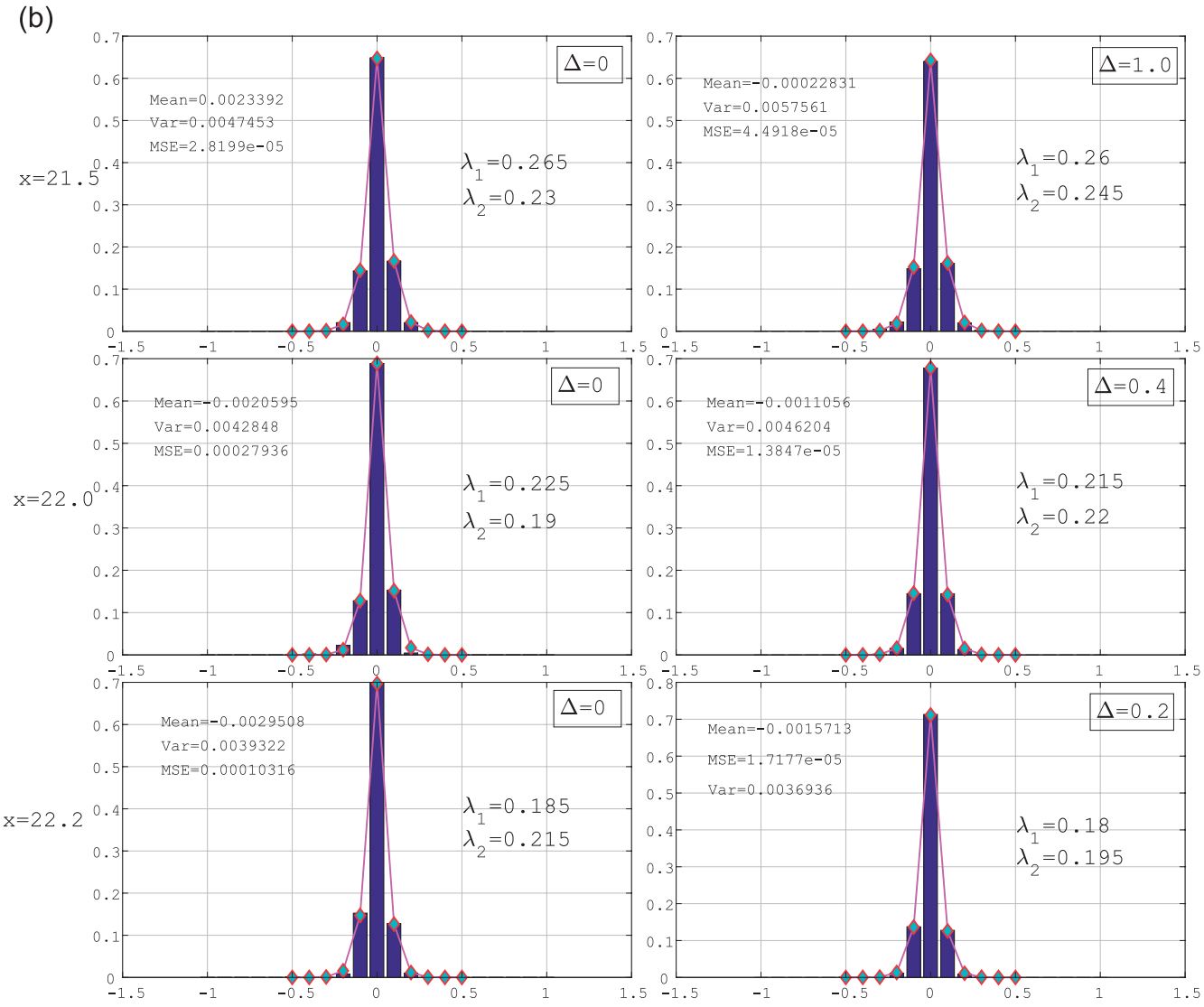

Figure 7 (b) $x = 21.5, 22.0, 22.2$°C, from top to bottom.

1. We use the extended Poisson distribution as an approximation of a continuous distribution model. The smaller the breadth of the *classes* in the histogram becomes, the closer the histogram may approximate a continuous distribution. In fact, it is well known that some discrete-valued distributions like Poisson, binomial, etc. approach to a continuous normal distribution.
2. As in the Supplementary Material, the histogram of the increment set $Y(x)$ or $Y_\Delta(x)$ are of high kurtosis, which may be expected to approach to a non-Gaussian distribution. This corresponds to the fact that the temperature control is based on a nonlinear system and not on a classical diffusion. This further indicates that the innovation sequences $Y(x)$ or $Y_\Delta(x)$ are distributed according to a heavy-tailed distribution like stable distributions [19, see Chapter 6 and Fig. 6.3]. The stable distribution and other heavy-tailed distributions are known to be generated by certain normalized sum of Poisson distributions [20]. Therefore, the Poisson distributions can be considered to be 'seeds' of the heavy-tailed distributions, and it is natural to use Poisson distributions as a model for the innovations.
3. In this paper, the available data are only of the 0.1°C step, anyway. At least for this discrete data, fitting the Poisson-like distribution is not strange; the model fits very well, indeed. Rather, we might accept the fitting result and consider why the model fits in such a good manner.

Hypothesis testing for the extended poisson model

In this section, we perform the hypothesis testing for the plausibility of Extended Poisson model $p_x(k)$ fitted to the data $Y_\Delta(x)$. Since the hypothesis testing is not the mainstream of the paper although it is necessary, readers who are not interested in it may skip this section.

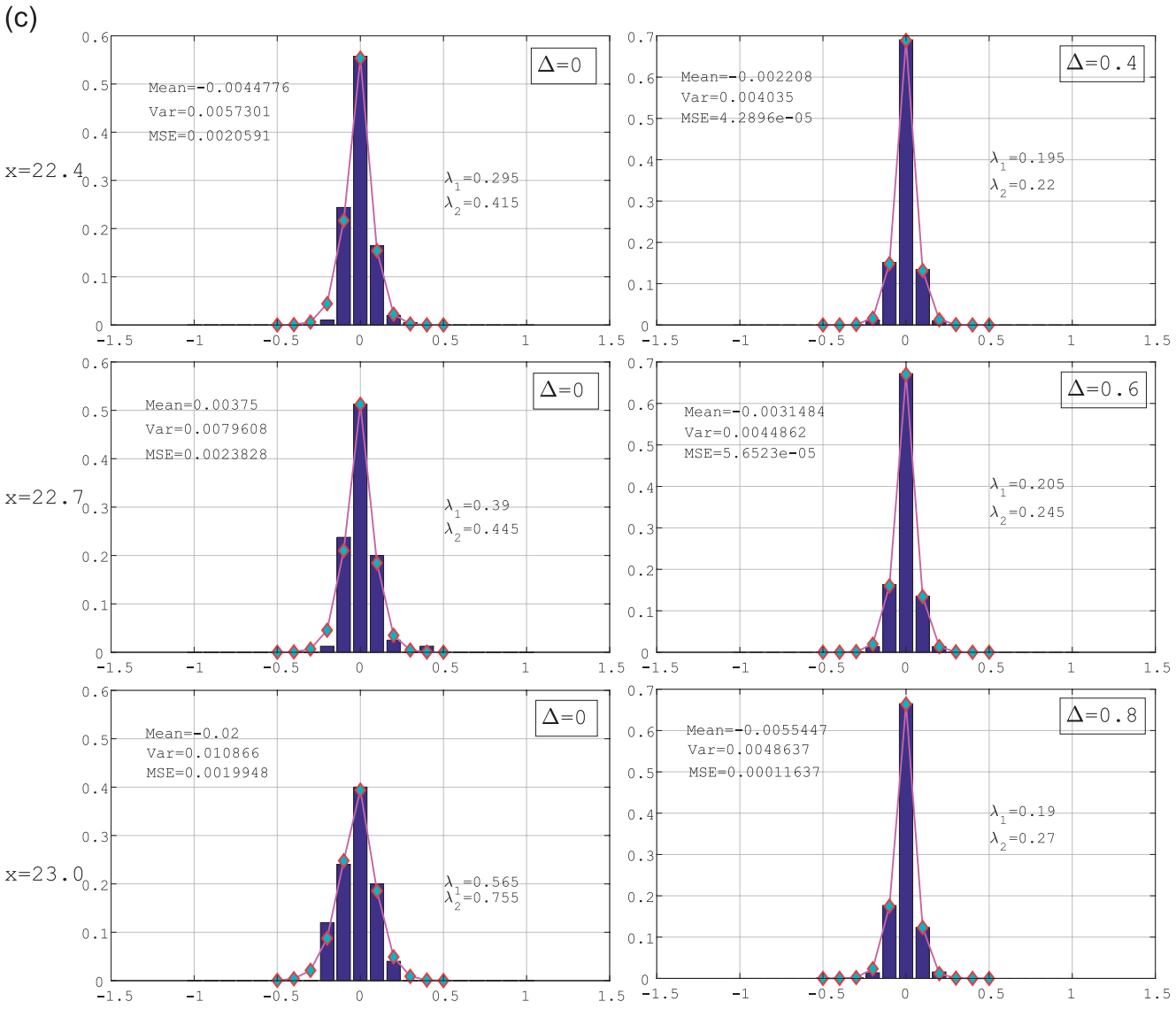

Figure 7 (c) $x = 22.4$, 22.7, 23.0°C, from top to bottom.

We use the Kolmogorov-Smirnov (KS) test [21–23], which is well-suited for fitting a histogram of real data to a model distribution. The KS test measures a discrepancies between two empirical cumulative distribution function (CDF). Here we take, as one of the CDF, the prospective model CDF i.e. the extended Poisson distribution.

We denote the empirical CDF of $Y_\Delta(x)$ as $\hat{F}_n = \hat{F}_n(z)$, $z \in \mathbf{R}$, for data length n, i.e.

$$\hat{F}_n(z) = \frac{1}{n}\sum_{j=1}^{n} I_j(z), \quad I_j(z) = \begin{cases} 1 & y_j \le z, \\ 0 & y_j > z \end{cases} \quad \text{for} \quad y_j \in Y_\Delta(x) \tag{9}$$

and the CDF of the estimated extended Poisson distribution as $F(z)$, respectively.

We take the null hypothesis H_0: $\hat{F}_n \sim F$ and the alternative hypothesis H_1: otherwise. For the frequency of the data $Y_\Delta(x)$ and its estimated frequency as shown above, the KS test considers the statistic

$$D_n \triangleq \sup_z |\hat{F}_n(z) - F(z)|. \tag{10}$$

Then, it is shown that $\sqrt{n}\,D_n$ asymptotically obeys the KS distribution given by

$$L(z) = 1 - 2\sum_{j=1}^{\infty}(-1)^{j-1} e^{-2j^2 z^2}, \quad z \in \mathbf{R}. \tag{11}$$

The KS test then rejects H_0 when $\sqrt{n}\,D_n > c$ for percentiles c of $L(z)$. For example, the 95% and 99% percentiles for large values of n are 1.358 and 1.628, respectively [24, p. 70].

Supplementary Table S2 shows the results of the KS test for samples 1 to 12. For each sample, we performed the test for several values of Δ and for significance levels 5% and

1%. For all samples, we found that as Δ becomes larger the range of x for which H_0 is not rejected becomes broader. Some samples are 'good', meaning that H_0 is not rejected for most values of x as they stand, i.e. with $\Delta = 0$, whereas other samples gain more 'not rejected' gradually as Δ increases. Anyway, 'not rejected' increases with Δ for every sample. We increased the Δ until all or most of x become 'not rejected'. For all samples, we take the range of x for which the size $|Y_\Delta(x)|$ is more than 30. Sample 5 is a special one that has all 'not rejected' for $\Delta = 0$. From the Table, we consider that our extended Poisson model to be validated as a plausible candidate of the model for these sample data, for the temperatures around CT.

Finally we mention the inapplicability of the χ^2 test. Although we wanted to use the χ^2 test [22,23] together with the KS test, we found it difficult to use. This is because the histograms of $Y_\Delta(x)$ basically have high concentricity, i.e. they have very small tail probabilities, whereas the number of classes $k = -5$, ..., 5 for the histogram of $Y_\Delta(x)$ itself is small, due to the fact that our data have a 0.1°C step precision, which might not be minute enough. Classes with small tail probabilities (around $k = -5$ or 5), as is suggested in general and small expected frequencies (e.g., less than 5) should be summed up into a new single class in order to preclude inaccuracy in the χ^2 test [25]. But this reduces the effective number of classes.

Hence, the degree of freedom of the χ^2 distribution-which is given by (number of effective classes)-1-(number of estimated parameters in model fitting) and must be one or more-ends up being zero in many cases. This defect might be overcome if the data precision, which is presently 0.1°C, is improved by using a smaller step. We hope for future availability of such precise data.

Temperature dependence of equilibrium characteristics captured by poisson parameters

Here, we consider the temperature dependence of the (conditional) expectation $\eta(x) \triangleq [\mathbf{E}[y(t)|X(t) = \mathrm{x}]$. We are interested in whether the dependence relationship can be described well by the Poisson parameters at least partially.

We first plotted the sample mean of $Y(x)$, $x \in \boldsymbol{T}$ as in Figure 8 (top). This does not appear to provide us with helpful insights into the relationship. Therefore, as in previous sections, we consider taking the expectation under the 'average' with Δ. We thus consider

$$\eta_\Delta(x) \triangleq \mathbf{E}_x[Y_\Delta(x)] = \mathbf{E}[y(t)| \, |X(t) - x| \le \Delta] \tag{12}$$

for $\Delta = 0$, 0.1, 0.2, ... Especially, when $\Delta = 0$, it reduces to $\eta(x)$. Figure 8 plots the $\eta_\Delta(x)$: (middle) is $\Delta = 0.4$ and (bottom) is $\Delta = 0.8$.

As seen in the figure, the larger the Δ, the clearer the linearity tendency of $\eta_\Delta(x)$ over x becomes. In every case, the left half of the straight line contains positive values, and the right half negative. This corresponds to the fact that below CT, the jumps take positive values towards CT, and above CT, the jumps take negative values towards CT. The red line in the neighbourhood of CT in the Figure is the model fitted using least squares. Also, the figure may suggest that we can now determine the CT as the intersection point of this red line and $\eta_\Delta = 0$ as well.

While it might be desirable to test the significance of the linearity itself, one cannot apply the basic form of the statistical inference for linear regression, because $Y(x)$ or $Y_\Delta(x)$ follow a extended Poisson distribution, rather than a normal distribution. Even though one might think of applying the generalised linear model [5] in which regression errors might not necessarily be normal, our extended Poisson model is in fact not even the generalized linear model. Hence, it is difficult to apply a hypothesis testing for linearity, and we do not consider such testing in this paper. Instead, we look at the correlation coefficient $\mathrm{Cov}[\boldsymbol{X}, \eta_\Delta] / [\mathrm{Var}[\boldsymbol{X}]\mathrm{Var}[\eta_\Delta]]^{1/2}$ as a measure of goodness of fit to a straight line. We thus report our results in this paper from the point of view of data analysis.

The correlation coefficient in the above-mentioned linearity takes real values in [−1, 1]. A value closer to −1 or +1 indicates a better fit. In particular, a value just equal to −1 or +1 implies that a straight line fits the data perfectly. In the present case, if we take Δ to be somewhat large then the correlation coefficient appears to approach −1. In fact, the values are $0.95 - 0.99$. This suggests, from data analysis, the plausibility of the linearity. The linearity,

$$\eta_\Delta(x) = -k(x - x_0) \tag{13}$$

for a constant $k > 0$ and CT x_0, is the same form as the elastic force of a spring to return to its original length when it is expanded or contracted. The temperature control is done based on the principle of a spring (*Fuch's law*) *in average*.

This linearity in the average of the jumps $Y_\Delta(x)$ of the temperature control is captured by the parameters of extended Poisson model. That is, the fitted linear model and the parameter $0.1[\lambda_1(x) - \lambda_2(x)]$ in Eq. (7) take mostly very close values. See Figures 8 and 9. The estimability itself of $\eta_\Delta(x)$ by $0.1[\lambda_1(x) - \lambda_2(x)]$ was expected *for each x*. What is found here is the linearity of $\eta_\Delta(x)$, and the additional plausibility of the extended Poisson model: there could be cases such that the model happens to appear to be good for *for each x*, but turns out not to be, when it is thought of as a model for all $x \in \boldsymbol{X}$. However our model, proposed for each x, turns out to be good for all $x \in \boldsymbol{X}$, simultaneously as well. Thus, the dependence of $\eta_\Delta(x)$ on x is captured by $\lambda_1(x) - \lambda_2(x)$ as the linearity. In addition, this might be a reason for investigating $Y_\Delta(x)$ instead of $Y(x)$.

Let us further consider the temperature control mechanism at equilibrium. The control is based on the balancing of exothermic reaction and endothermic reaction in pre-equilibrium in the left half of (3). Let $[RH_2]$ be concentration of molecules RH_2, and let $[R]$ and $[H^+]$ be the concentration of

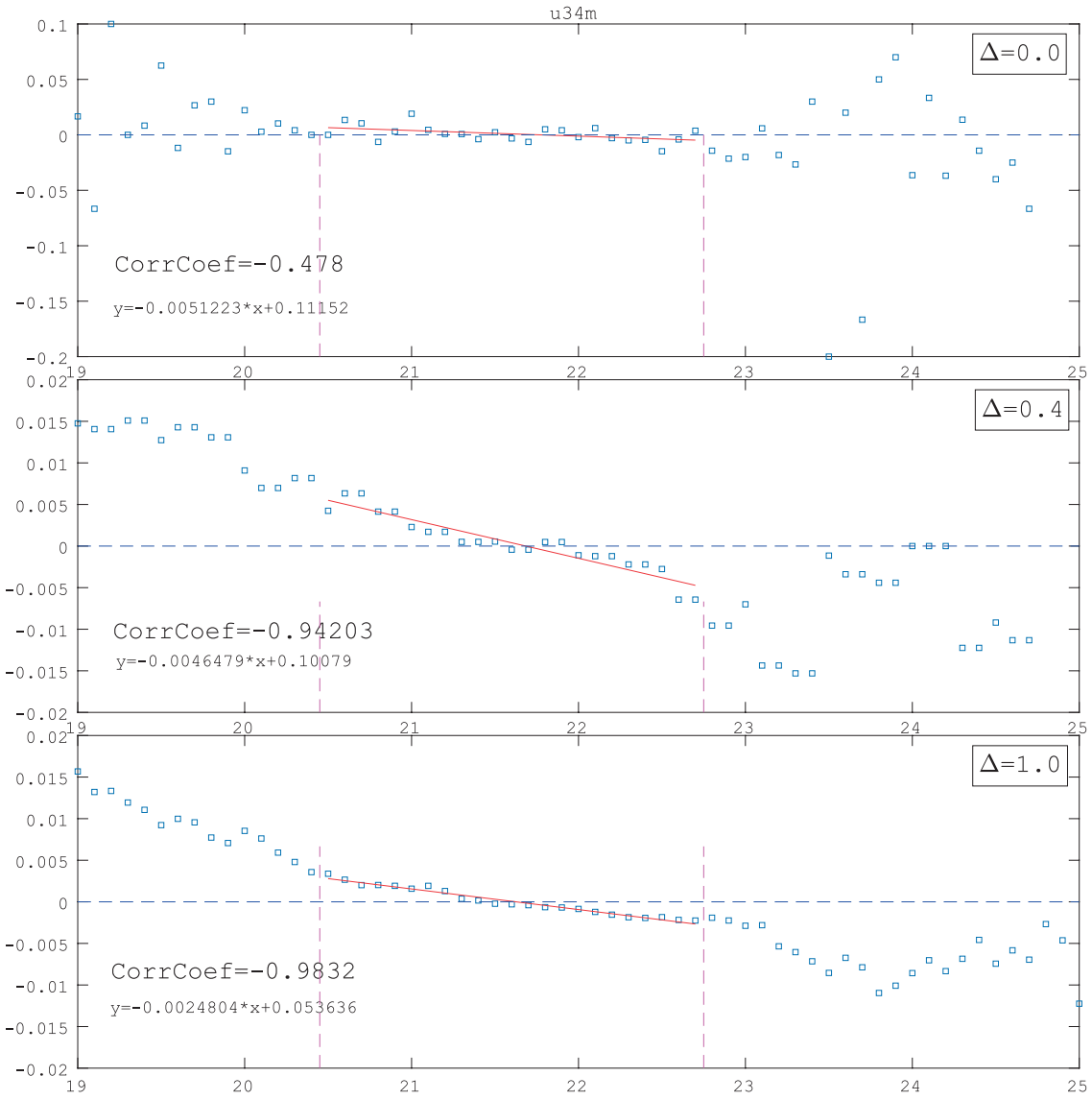

Figure 8 Linear behaviour of $\eta_\Delta(x)$ in pre-equilibrium. $\Delta = 0$, 0.4 and 1.0, from top to bottom.

R and H^+, respectively. Then, from chemical reaction kinetics [6], we have

$$\begin{cases} \dfrac{d[RH_2]}{dt} = -k_1[RH_2] + k_1'[R][H^+]^2[e^-]^2, \\ \dfrac{d[R]}{dt} = k_1[RH_2] - k_1'[R][H^+]^2[e^-]^2. \end{cases} \tag{14}$$

Adding the two equations leads to $[RH_2] + [R] \equiv \text{const.}$ This indicates that the kinetics is such that if the exothermic reaction is exacerbated, then the reverse reaction, i.e. the endothermic reaction, is automatically switched to promotion, as on a seesaw. $[R] = const. - [RH_2]$ implies that when the exothermicity is exacerbated and hence $[RH_2]$ is much consumed, or, $[RH_2]$ remains small, then $[R]$ becomes large, and hence endothermicity promotion begins. This might be considered to have merits as follows:

- In order for plants that do not have sensors to perceive environmental information to perform the appropriate temperature control, the automatic switching of exothermic reaction ⇌ endothermic reaction occurs.
- If a particular amount of energy was consumed for the switching itself, a loss of the energy occurs. In an intensely cold environment, the plant would prefer to use energy mostly for exothermicity itself as far as possible. The above mechanism, which does not have additional function modules for the switching, causes no loss of energy.

Thus, we may say that the mechanism of switching of the exothermic reaction ⇌ endothermic reaction is remarkably

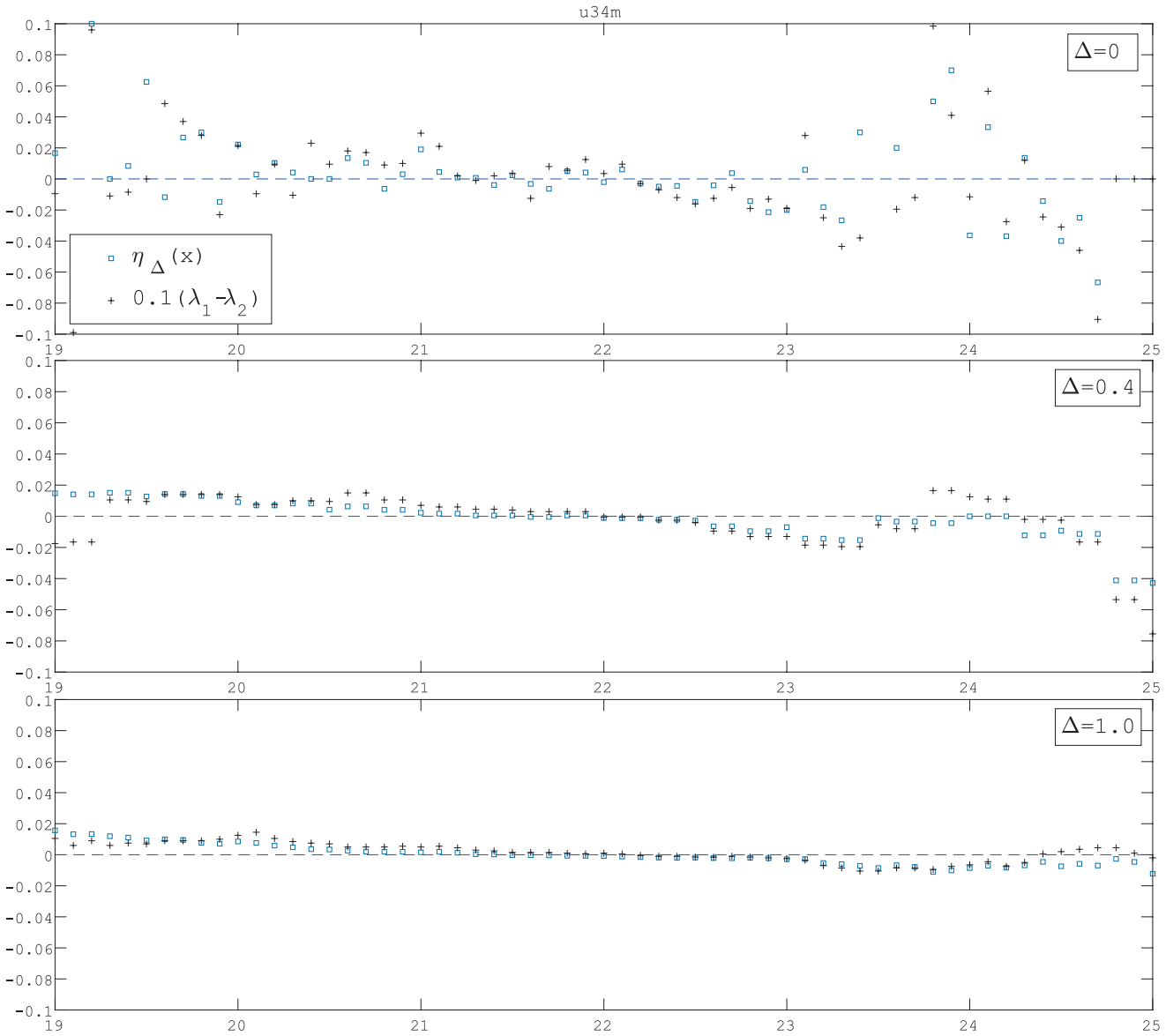

Figure 9 The larger the Δ becomes, the better the fit of $0.1(\lambda_1 - \lambda_2)$ to $\eta_\Delta(x)$. $\Delta = 0$, 0.4 and 0.8, from top to bottom (The same sample as Fig. 10).

simple and well constructed.

We have pointed out above that in the pre-equilibrium, the 'elastic' force is a linear function of x on an interval around the CT. It is the behaviour of the expectation $\eta_\Delta(x)$. How about the variance? Because $[\lambda_1(x) + \lambda_2(x)]$ essentially represents the variance as in Eq. (8), we plot the variance of $Y_\Delta(x)$, $x \in \boldsymbol{X}$ and its prospective model $(0.1)^2[\lambda_1(x) + \lambda_2(x)]$ in Figure 10.

In the Figure, we can see that on the linearity region of $0.1(\lambda_1 - \lambda_2)$, the variances are mostly constant. In addition, the model of variance, $(0.1)^2[\lambda_1(x) + \lambda_2(x)]$ show good agreement with the variance. Here as well, the larger the Δ, the clearer the tendency of $(0.1)^2[\lambda_1(x) + \lambda_2(x)] = const.$ becomes. What is observed commonly from these is that for particular large values of Δ or greater, the constant region is much broader than we expected. We anticipated that such a constant region would be just a small neighbourhood of the CT. The plant can be considered to be coping with the homothermal maintenance by equipping this broad constant region. Because the constant region is considered to be exactly the homothermal maintenance region, the broad region might be a merit for the plant.

A more detailed analysis, such as a quantitative evaluation of the constant region or a characterisation of the region width of each individual, might be desirable as future work.

Conclusion

In this paper, as a partial study to elucidate the temperature control mechanism of Skunk Cabbage, we considered

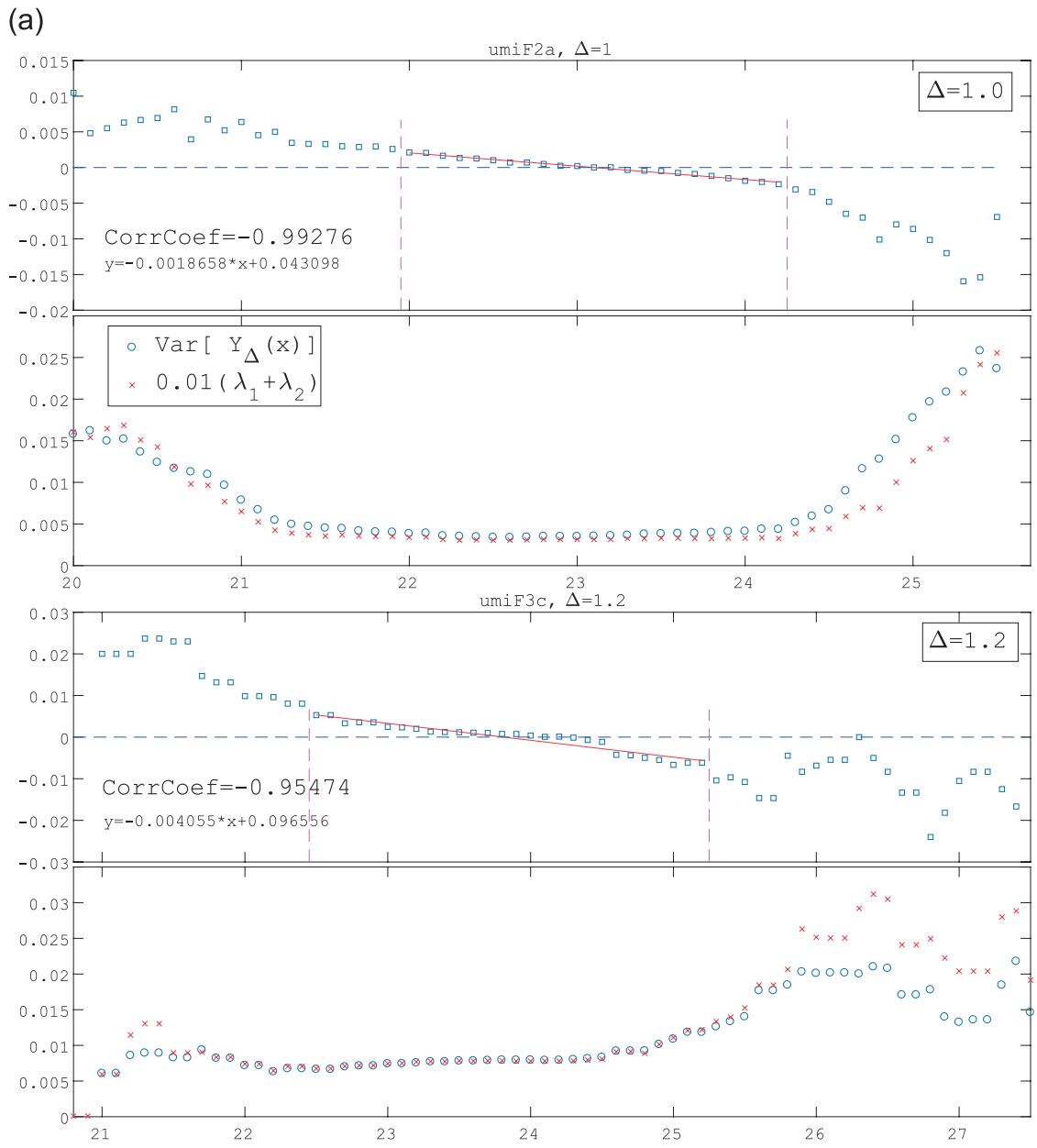

Figure 10 $\eta_\Delta(x)$ and $(0.1)^2(\lambda_1+\lambda_2)$. (a) (top) sample 12 with $\Delta = 1.0$, (bottom) sample 6 with $\Delta = 1.2$.

the modelling of the homothermal maintenance in pre-equilibrium, through a data analytic study. For a kind of averaged increments $Y_\Delta(x)$ for each x, we proposed an extended Poisson model. Through hypothesis testing, we showed the plausibility of the model. It turned out that the temperature dependence of the expectation of $Y_\Delta(x)$ can be considered to be linear, and the linearity agrees well with $0.1[\lambda_1-\lambda_2]$. In addition, the temperature dependence of the variance of $Y_\Delta(x)$ can be considered to be a constant with unexpectedly broad range. This constant also agrees well with the prospective model $(0.1)^2[\lambda_1+\lambda_2]$. The proposed extended Poisson model might thus be not only a good candidate for the distribution of each x, but also a good means of investigating the mechanism. We strongly hope that we will be able to develop a method to show the theoretical plausibility, i.e. hypothesis testing framework, for the linearity of the mean or constancy of variance for this extended Poisson distribution.

The data available this time was of 0.1°C steps, enabling us to treat the data as if it obeyed a discrete distribution, and we considered a extended Poisson distribution on $0.1\times k$ ($k\in\mathbf{Z}$). However, we think that this treatment is only provisional, because if we think of the exothermic or endothermic reactions at the molecular level, a huge number of molecules (in fact, of the order of Avogadro number) are involved in the reactions. That is, a single generation of product of the

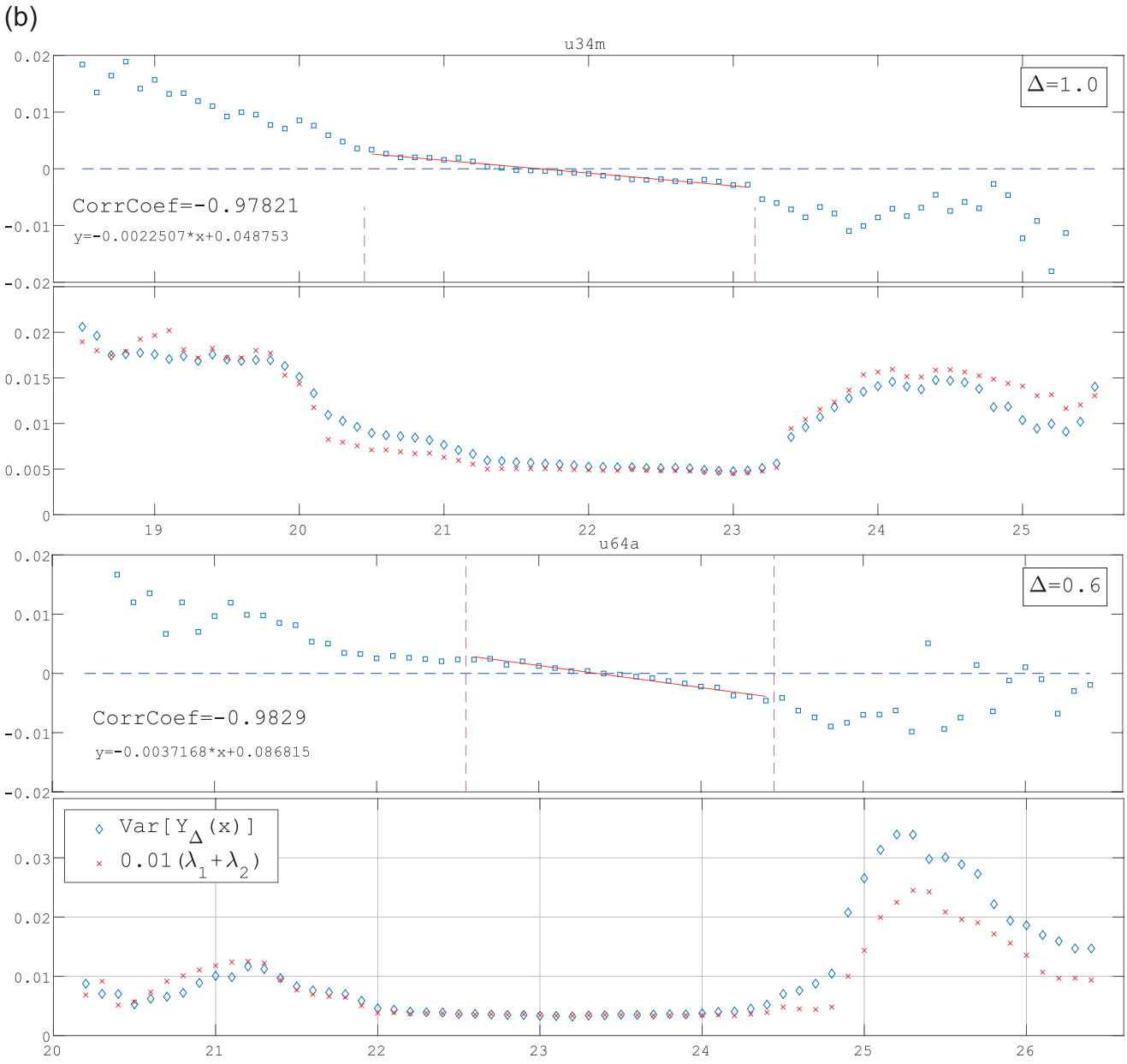

Figure 10 (b) (top) sample 10 with $\Delta = 1.0$, (bottom) sample 7 with $\Delta = 0.6$.

exothermal or endothermal reactions at one moment causes a minute increase or decrease in temperature; the limit of the sum of such minute increases or decreases will appear as a macroscopic increase or decrease of temperature. In the present case, the relevant distribution is not Gaussian. Hence, the limit theorem will be a noncentral limit theorem. The extended Poisson distribution has a cusp-like peak at the origin for small values of the intensities λ_1 and λ_2. Thus, in the female period, there corresponds a singular distribution due to the noncentral limit theorem and the critical phenomena. A continuous model might actually be desirable for understanding these elements of the mechanism.

Nevertheless, the discrete model treatment in this paper using the extended Poisson model is still an important step, because a discrete model can be expected to serve as a particular case reducible from a continuous model to be developed in the future.

In addition, we introduced a kind of averaging by Δ in the data in order to extract clearer tendencies. We would like to make the meaning, plausibility or appropriate magnitude of the Δ clearer in the future.

Finally, we would like to proceed to the analysis of the nonequiilibrium state, i.e. the latter half of the reactions in Eq. (3). To do that, we think we expect to have to elucidate the relationship of the extended Poisson model and the 'elastic' force, as well as the related dynamical system that

can describe the seesaw from various points of view, such as molecular biology, data analysis, phase transition in non-equilibrium physics and mathematical modelling. Among them, the approach from physical point of view as in [26,27] may be of a useful clue.

There also are thermogenic plants like *Lotus*, *Dracunculus vulgaris* other than the Skunk Cabbage, that perform the control for homothermal maintenance, as reported in [28]. However,to the best of our knowledge, mathematical or data analyses as in this paper have not been done so far. It is desired in the future to clarify if such other plants have a similar mechanism characterized by the pre-equilibrium.

Acknowledgements

This research was supported by JSPS KAKENHI, Grant Number JP15594730.

Conflicts of Interest

S. K. and K. I. declare hereby that they have no conflict of interests.

Author Contribution

S. K. conducted the mathematical and computational analyses and drafted the manuscript. K. I. indicated the problem, collected the data and interpreted the results and revised the manuscript.

References

[1] Knutson, R. M. Heat production and temperature regulation in eastern skunk cabbage. *Science* **186**, 764–747 (1974).
[2] Ito, K., Ogata, T., Kakizaki, Y., Ellliot, C., Albury, M. S., Moore, A. L. Identification of a gene for pyruvate-insensitive mitochondrial alternative oxidase expressed in the thermogenic appendices in *Arum maculatum*. *Plant Physiol.* **157**, 1721–1732 (2011).
[3] Ito, K., Takahashi, H., Umekawa, Y., Imamura, T., Kawasaki, S., Ogata, T., *et al.* Metabolite profiling reveals tissue- and temperature-specific metabolomic responses in thermoregulatory male florets of *Dracunculus vulgaris* (Araceae). *Metabolomics* **9**, 919–930 (2013).
[4] Kakizaki, Y., Moore, A. L. & Ito, K. Different molecular bases underlie the mitochondrial respiratory activity in the homoeothermic spadices of *Symplocarpus renifolius* the transiently thermogenic appendices of *Arum maculatum*. *Biochem. J.* **445**, 237–246 (2012).
[5] Onda, Y., Kato, Y., Abe, Y., Ito, T., Ito-Inaba, Y., Morohashi, M., *et al.* Pyruvate-sensitive AOX exists as a non-covalently associated dimer in the homeothermic spadix of the skunk cabbage, *Symplocarpus renifolius*. *FEBS Lett.* **581**, 5852–5858 (2007).
[6] Atkins, A. & De Paula, J. *Physical Chemistry for the life Sciences, 2nd ed.* (Oxford University Press, New York, 2009).
[7] Ito, T. & Ito, K. Nonlinear dynamics of homeothermic temperature control in skunk cabbage, Symplocarpus foetidus. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **72**, 051909 (2005).
[8] Takahashi, K., Ito, T., Onda, Y., Endo, T., Chiba, S., Ito, K., *et al.* Modeling of the thermoregulation system in the skunk cabbage: *Symplocarpus foetidus*. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* **76**, 031918 (2007).
[9] Skunk Cabbage generates plenty of heat, CERN Courier, Feb. 8, 2006. (See also http://cerncourier.com/cws/article/cern/29508).
[10] Zen and the art of central heating maintenance, New Scientist, **16** Nov. 2005. (See also https://www.newscientist.com/article/mg18825265-000-zen-and-the-art-of-central-heating-maintenance/).
[11] The Zen of Skunk Cabbage, Science AAAS, **310**: 1239, 2005. (See also http://www.sciencemag.org/news/2005/11/zen-skunk-cabbage).
[12] Beran, B. *Statistics for Long-Memory Processes* (Chapman & Hall/CRC Press, Florida, 1994).
[13] Pipiras, V. & Taqqu, M. *Long-Range Dependence and Self-Similarity* (Cambridge University Press, Cambridge, 2017).
[14] The dataset (MATLAB format) used in the Figures in the paper are available at: https://www.researchgate.net/publication/315090574_Data_set_for_Data_analytic_study_of_Skunk_Cabbage, with DOI: 10.13140/RG.2.2.23852.05768.
[15] The MATLAB codes used to generate Figures in the paper are available at: https://www.researchgate.net/publication/315090590_MATLAB_code_files_for_Skunk_Cabbage_analysis, with DOI: 10.13140/RG.2.2.23852.05768.
[16] Umekawa, Y., Seymour, R. S. & Ito, K. The biochemical basis for thermoregulation in heat-producing flowers. *Sci. Rep.* **6**, 24830 (2016).
[17] Billingsley, P. *Probability and Measure, 3rd ed.* (John Wiley & Sons, New York, 1995).
[18] Feller, W. *An Introduction to Probability Theory and Its Applications 2nd ed.* (John Wiley & Sons, New York, 1971).
[19] Klafte, J. & Sokolov, I. M. *First Steps in Random Walks* (Oxford University Press, New York, 2011).
[20] Samorodnitsuky, G. & Taqqu, M. *Stable Non-Gaussian Random Processes* (Chapman & Hall, Florida, 1994).
[21] Durbin, J. *Distribution Theory for Tests Based on the Sample Distribution Function* (SIAM, Philadelphia, 1973).
[22] Lehmann, E. L. *Elements of Large Sample Theory* (Springer, New York, 1999).
[23] Lehmann, E. L. & Romano, J. P. *Testing Statistical Hypotheses, 3rd ed.* (Springer, New York, 2005).
[24] Lindley, D. V. & Scott, W. F. *New Cambridge Statistical Tables, 2nd ed.* (Cambridge University Press, New York, 1995).
[25] Tate, M. W. & Hyer, L. A. Inaccuracy of the X^2 Test of Goodness of Fit When Expected Frequencies Are Small. *J. Am. Stat. Assoc.* **68**, 836–841 (1973).
[26] Kruse, J. & Adams, M. A. Three parameters comprehensively describe the temperature response of respiratory oxygen reduction. *Plant Cell Environ.* **31**, 954–967 (2008).
[27] Kruse, J., Rennenberg, H. & Adams, M. A. Steps towards a mechanistic understanding of respiratory temperature responses. *New Phytol.* **189**, 659–677 (2011).
[28] Seymour, R. S., Lindshau, G. & Ito, K. Thermal clamping of temperature-regulating flowers reveals the precision and limits of the biochemical regulatory mechanism. *Planta* **231**, 1291–1300 (2010).