doi: 10.17221/83/2016-PPS

Using the LOESS Method to Describe the Effect of Temperature on Development Rate

Junhe LIU[1], Yan YAN[2], Mingfu YU[1], Megha N. PARAJULEE[3†], Peijian SHI[4], Jiayang LIU[1*] and Zihua ZHAO[5*]*

[1]*Department of Biological Engineering, Huanghuai University, Zhumadian, Henan, P.R. China;* [2]*Landscape Research Institutes of Zhumadian, Zhumadian, Henan, China;* [3]*Texas A & M AgriLife Research and Extension Center, Lubbock, Texas, USA;* [4]*Collaborative Innovation Center of Sustainable Forestry in Southern China of Jiangsu Province, Nanjing Forestry University, Xuanwu District Nanjing, P.R. China;* [5]*Department of Entomology, College of Plant Protection, China Agricultural University, Beijing, P.R China*

**Corresponding authors: zhzhao@cau.edu.cn; jyliu@uga.edu; shi_peijian@163.com*

Abstract

Liu J., Yan Y., Yu M., Parajulee[†] M.N., Shi P., Liu J., Zhao Z. (2017): Using the LOESS method to describe the effect of temperature on development rate. Plant Protect. Sci., 53: 226–231.

Temperature has a significant influence on development rates of insects and mites. Many parametric models were built to describe the temperature-dependent development rates. However, these models provided different shapes of the curves of development rate versus temperature. For different datasets, investigators have to spend much time on considering which the parametric model is the best for describing the temperature-dependent development rates. In the present study, we encourage investigators to use an important non-parametric model, the LOESS method, which belongs to local regression methods. The LOESS method is used to fit some published data on the development rate of aphids to check the goodness-of-fit. We find that the LOESS method is very flexible for fitting the given datasets. Thus, we consider that the LOESS method can be used to describe the effect of temperature on the development rate of insects or mites.

Keywords: non-parametric model; local regression; weighted least squares; Aphididae

In general, one insect or mite species needs to experience different developmental stages in its lifecycle. Completing each developmental stage requires a period of time, which is referred to as development time (or development duration). The reciprocal characteristic of development time is referred to as development rate. It has been demonstrated that temperature has a strong influence on development rate (Campbell *et al.* 1974; Taylor 1981). Development rate is regarded as a linear function of temperature over the range of moderate temperatures (Campbell *et al.* 1974; Zhao *et al.* 2012). However, this linear relationship between development rate and temperature does not fit if temperatures were not in this range. Several non-linear models have been developed for describing the effect of temperature on development rate (e.g., Logan *et al.* 1976; Sharpe & DeMichele 1977; Schoolfield *et al.* 1981; Taylor 1981; Wang *et al.* 1982; Lactin *et al.* 1995; Brière *et al.* 1999; Ikemoto 2005, 2008; Shi *et al.* 2011a, b). These non-linear models are like a two-sided sword: (1) they represent abundant tools for describing the temperature-dependent development rates; (2) it takes to spend much time on choosing one among so many models. Although some comparative studies were done to recommend the best of these models (e.g., Smits *et al.* 2003; Kontodimas *et al.* 2004; Kim *et al.* 2009; Shi & Ge 2010), these studies did not draw a corresponding conclusion yet.

It is seemingly difficult to use a single model to describe the effect of temperature on development

doi: 10.17221/83/2016-PPS

rates of different insects and mites. That is, a single universal non-linear model used to describe the temperature-dependent development rates of all species of insects and mites might be impossible. The majority of the investigators usually choose models based on the data in practice. It is necessary to point out that all the non-linear models built for describing the temperature-dependent development rates are parametric models. The shape of the curve of a parametric model is known when its formulation is given. However, sometimes these models do not fit the experimental development rate data very well. It seems that those data are authentic although there are probably some small experimental errors. For example, it is found that the aforementioned non-linear parametric models are not applicable to the development rate data of the cabbage beetle, *Colaphellus bowringi* Baly (Coleoptera: Chrysomelidae), at eight temperatures ranging from 16°C to 30°C (Kuang *et al.* 2011). The mortality of this insect species at every temperature excluding 30°C is low (lower than 50% for the egg, pupal and larval stage). However, temperature dependent development models are not intended to predict or account for mortality, only the rate of development with respect to temperature. Since all non-linear models are chosen mainly according to whether they can fit the data well, some non-parametric fitting methods, such as local regression models (Cleveland *et al.* 1991; Loader 1999; Zhao *et al.* 2013) and generalised additive models (Hastie & Tibshirani 1990), can reflect the relationship between development rate and temperature. We need not know the strict formulation before using these models to fit the data. And these models will provide the fitted results only according to the data themselves. Thus, non-parametric models are in fact considered to let the data speak. It indicates that we need not to know the complex mechanism of the effect of temperature on development rate but we can use a non-parametric model to reflect the relationship between development data and temperature very well. Frankly speaking, although the Sharpe-Schoolfield-Ikemoto model (Sharpe & DeMichele 1977; Schoolfield *et al.* 1981; Ikemoto 2005, 2008; Shi *et al.* 2011b) and another model proposed recently by Ratkowsky *et al.* (2005) based on thermodynamics have perfectly combined the temperature-dependent development rates to the probability of an enzyme being in the active state, the mechanisms of temperature on the enzymes that control development and the types of these controlling enzymes are still unknown. Therefore, it is necessary to use a non-parametric fitting method. In the present study, we employ an important local regression model, LOESS (Cleveland 1979; Cleveland *et al.* 1991; Jacoby 2000), to depict the effect of temperature on development rate.

MATERIAL AND METHODS[(1)]

The LOESS method was pioneered by Cleveland (1979), Cleveland *et al.* (1988), Cleveland and Grosse (1991). Suppose r_i for $i = 1$ to n are observed development rates. Suppose T_i for $I = 1$ to n are observed temperatures. The local regression model related to the temperature-dependent development rates can be described as:

$$r_i = g(T_i) + \varepsilon_i \tag{1}$$

where: $g(\bullet)$ – smooth function of T; ε_i – random variables with mean 0 and variance σ^2

Let $\hat{r}_i$ represent an esti≤mate of r_i, T_i is the daily average temperature, i is daily temperature.

Let $\Delta_k(T_i)$ denote the Euclidean distance of T_i to T_k. Let $\Delta_{(k)}(T_i)$ denote the values of these distances ordered from smallest to largest, and let

$$T(u) = \begin{cases} (1-u^3)^3 & \text{for } 0 \le u < 1 \\ 0 & \text{for } u \ge 1 \end{cases} \tag{2}$$

denote the weight function.

Local regression requires a smoothing parameter, α ($\alpha \le 1$). And the integer obtained from truncating $\alpha \bullet n$ represents the number of points used to perform each local regression. Obviously, the parameter, α, limits the proportion of observations that is to be used in each local regression. Let q be that integer. We define a weight for (T_k, r_k) by

$$w_k(T_i) = \mathrm{T}\left(\frac{\Delta_k(T_i)}{\Delta_{(q)}(T_i)}\right) \tag{3}$$

It decreases or remains constant as T_k increases in distance from T_i.

(1) For each i compute the estimates, $\hat{\beta}_{ij}$, $j = 0$ to d, of the parameters in a polynomial regression of degree d of r_i on T_i, which is fitted by weighted least squares with weight $w_k(T_i)$ for (T_k, r_k), i.e., minimising:

$$\sum_{k=1}^{n} w_k(T_i)\left(r_k - \sum_{j=0}^{d}\left(\hat{\beta}_{ij} T_k^j\right)\right) \tag{4}$$

[(1)]We mainly referred to the study of Cleveland (1979) in this section.

doi: 10.17221/83/2016-PPS

Then we compute $\hat{r}_i$ by

$$\hat{r}_i = \sum_{j=0}^{d} (\hat{\beta}_{ij} T_i^j) = \sum_{k=1}^{n} (I_{ik} r_k) \quad (5)$$

(2) Let B denote the bisquare weight function that is defined by

$$B(v) = \begin{cases} (1 - v^2)^2 & \text{for } |v| < 1 \\ 0 & \text{for } |v| \geq 1 \end{cases} \quad (6)$$

Let

$$e_i = r_i - \hat{r}_i \quad (7)$$

be the residuals from the current fitted values. Let s be the median of the $|e_i|$. Define robustness weights by

$$\zeta_k = B(e_k/6s) \quad (8)$$

(3) Compute new $\hat{r}_i$ for each i by fitting a d^{th} degree polynomial using weighted least squares with weight $\zeta_k \cdot w_k(T_i)$ at (T_k, r_k).

(4) Repeatedly execute steps 2 and 3 a total of t times. The final $\hat{r}_i$ are robust locally weighted regression fitted values.

We use the R package (http://www.r-project.org/) to carry out the LOESS fitting. There are two outputs needed to explain in detail: equivalent number of parameters and residual standard error.

Let

$$\hat{r} = Lr \quad (9)$$

and let

$$\delta_1 = \text{Trace } (I - L)^T (I - L) \quad (10)$$

The definitions of these two outputs are:

$$\text{Equivalent number of parameters} = \text{Trace } (L^T L) \quad (11)$$

$$\text{Residual standard error} = (\text{RSS}/\delta_1)^{1/2} \quad (12)$$

RESULTS AND DISCUSSION

In the present study, we used the LOESS method to fit some published development rate data of Aphididae (Table 1 and Figure 1). The LOESS model can fit the data well, and it does not need a strict formulation before fitting the data. The coefficients of determination (i.e., R^2) from sample A to sample F were 0.9993, 0.9971, 0.9967, 0.9980, 0.9851, and 0.9960, respectively. The curve shapes produced by the LOESS method can be different. The curve shapes cannot be produced by a single parametric model, such as the logistic model or the performance model. We append the program for investigators to execute the LOESS fitting. Just to use this program, we can use the LOESS method to accurately describe the temperature-dependent development rates.

The advantages have been mentioned above for using the LOESS method to describe the effect of temperature on development rate. The disadvantages of using it are two aspects: (1) it is easy to lead to the overfitting, i.e., it is very flexible that it can also fit some data with experimental errors; (2) it cannot be used to predict the lower developmental threshold (hereafter read as LDT) and the sum of effective temperatures (hereafter read as SET) which are often estimated by a classical linear model. The classical linear model for describing the temperature-dependent development rates is:

$$r = a + bT + \varepsilon \quad (13)$$

where: r – development rate; T – constant temperature; a and b – constants; ε – random error item

The estimates of the LDT and SET are defined by:

$$\begin{cases} \widehat{\text{LDT}} = -\hat{a}/\hat{b} \\ \widehat{\text{SET}} = -1/\hat{b} \end{cases} \quad (14)$$

where: the letter 'a' with a 'hat' represents the estimate of respective lettering

The LDT denotes the temperature below which development rate equals zero. There is another temperature concept known as upper developmental threshold (UDT) beyond which development rates also equal zero, and that is similar to the LDT. Insects are supposed to develop between LDT and UDT. However, the LOESS method based on the experi-

Table 1. Summary of the cited data

Sample	Species	Temperature range (°C)	Literature
A	*Aphis gossypii* Glover	15–32	Satar *et al.* (2005)
B	*Lipaphis erysimi* Kaltenbach	8.3–35.1	Liu and Meng (1990)
C	*Myzus persicae* Sulzer	6.2–30.0	Liu and Meng (1990)
D	*Myzus persicae* Sulzer	8–28	Chen *et al.* (2002)
E	*Myzus persicae* Sulzer	8–32	Qin and Li (2006)
F	*Rhopalosiphum pseudobrassicae* Davis	10–30	Kawada (1964)

doi: 10.17221/83/2016-PPS

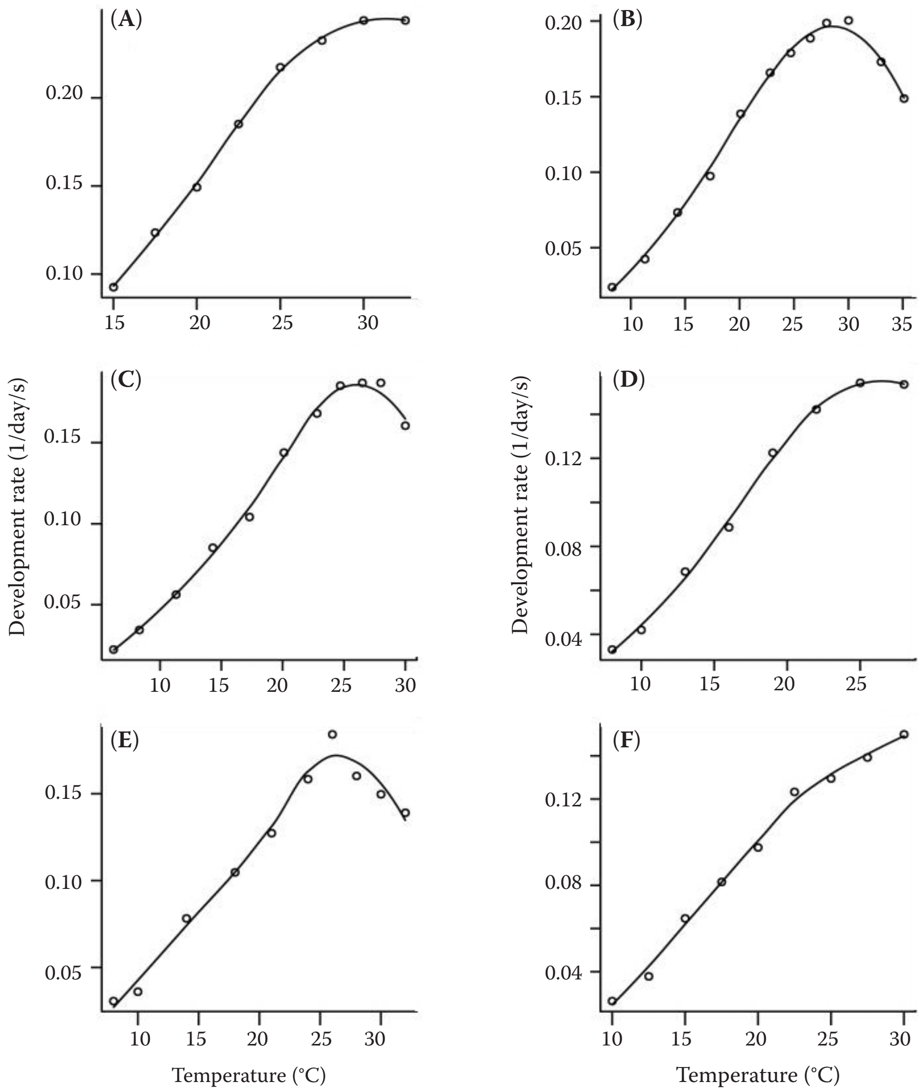

Figure 1. Comparison between the observations and predicted values by the loess method. Data source is shown in Table 1. The open circles represent the observations; and the solid lines represent the predicted values by the loess method. The smoothing parameter is defined as 0.75 for each sample

mental data cannot obtain these three parameters which are often used in entomology. It is necessary to point out that not all investigators believe in the existence of LDT, SET, and UDT. And some believe only in one or two. We can find the evidence from the non-linear models. The Logan model assumes only the existence of UDT. If the investigators want to calculate the LDT, they have to use the linear model for its estimation (e.g., Bonato *et al.* 2007; Eliopoulos *et al.* 2010). The Sharpe-Schoolfield-Ikemoto model assumes only the existence of LDT (Ikemoto 2005, 2008; Shi *et al.* 2011b). The Wang-Lan-Ding model (Wang *et al.* 1982), Lactin (Lactin *et al.* 1995), and performance models (Shi & Ge 2010; Shi *et al.* 2011a) are models whereas the existence of both LDT and UDT is assumed. In addition, the data presented in Figure 1 could conform well to other models, particularly those evaluated by Shi and Ge (2010).

In summary, there are two debates about the existence of LDT, SET, and UDT. Then we cannot reject the use of the LOESS method though it cannot be used to predict these three parameters (or theoretical concepts in a sense). Investigators had better deliberately design the temperature range to make these temperatures meaningful for experiments. In this case, the LOESS method can thoroughly show its advantages. For example, Liu and Meng (1990) carried out an experiment to explore the effect of temperature on the development of two species of

doi: 10.17221/83/2016-PPS

aphids (*Lipaphis erysimi* Kaltenbach, *Myzus persicae* Sulzer). They set the temperature range from 6.2°C to 37.0°C. That should be enough, because the suitable temperatures for these two species of aphids to develop in nature should be in this temperature range. As an important non-parametric fitting method, it should be reasonable and feasible for us to use the LOESS method to describe the effect of temperature on the development rate of insects or mites.

Acknowledgements. We thank Dr Takaya Ikemoto for providing us with one useful reference.

Appendix

Before using this program, we need to download the R package from http://www.r-project.org/. Then we type the following code in the R Console:

```
T <- c(15, 17.5, 20, 22.5, 25, 27.5, 30, 32.5)
D <- c(10.8, 8.1, 6.7, 5.4, 4.6, 4.3, 4.1, 4.1)
T <- T[!is.na(T)]
D <- D[!is.na(D)]
r <- 1/D
plot(T,r,xlab=expression(paste("Temperature (",degree,"C)")),ylab=c("Development rate (1/days)"),pch=1)
r.loess <- loess(r~T,span=0.75,data.frame(T=T,r=r))
r.hat <- predict(r.loess,data.frame(T=T))
Rsquare <- 1-sum((r-r.hat)^2)/sum((r-mean(r))^2)
T.predict <- seq(min(T),max(T),by=0.001)
r.predict <- predict(r.loess,data.frame(T=T.predict))
lines(T.predict,r.predict,col=1)
```

Here, T is a vector that is used to save the data of experimental temperatures; D is a vector that is used to save the data of development times. We take the data of Satar *et al*. (2005) as an example. Users can replace T and D with other data needed to be analysed. r is a vector that is used to save the observed development rates; r.hat is a vector that is used to save the predicted development rates by the LOESS method at T; Rsquare is the coefficient of determination, as an indicator of goodness-of-fit; T.predict is a vector that is used to save the temperatures given by the users if they want to predict the development rates at these temperatures. Default of T.predict is the range from the lowest experimental temperature to the highest experimental temperature by an increment of 0.001°C. r.predict is a vector that is used to save the predicted development rates at T.predict.

References

Brière J.F., Pracros P., Le Roux A.Y., Pierre J.S. (1999): A novel rate model of temperature-dependent development for arthropods. Environmental Entomology, 28: 22–29.

Bonato O., Lurette A., Vidal C., Fargues J. (2007): Modelling temperature-dependent bionomics of *Bemisia tabaci* (Q-biotype). Physiology Entomology, 32: 50–55.

Campbell A., Frazer B.D., Gilbert N., Gutierrez A.P., Mackauer M. (1974): Temperature requirements of some aphids and their parasites. Journal of Applied Ecology, 11: 431–438.

Chen W., Cui Z., Li B. (2002): Effect of temperature on population parameters of *Myzus persicae* (Sulzer) (Homoptera: Aphididae). Hubei Agricultural Science, 2: 68–69. (in Chinese)

Cleveland W.S. (1979): Robust locally weighted regression and smoothing scatterplots. Journal of the American Satistical Association, 74: 829–836.

Cleveland W.S., Grosse E. (1991): Computational methods for local regression. Statistics and Computing, 1: 47–62.

Cleveland W.S., Devlin S.J., Grosse E. (1988): Regression by local fitting: methods, properties, and computational algorithms. Journal of Econometrics, 37: 87–114.

Cleveland W.S., Grosse E., Shyu W.M. (1991): Local regression models. In: Chambers J.M., Hastie T.J. (eds): Statistical Models. London, Chapman & Hall: 309–376.

Eliopoulos P.A., Kontodimas D.C., Stathas G.J. (2010): Temperature-dependent development of *Chilocorus bipustulatus* (Coleoptera: Coccinellidae). Environmental Entomology, 39: 1352–1358.

Hastie T.J., Tibshirani R.J. (1990): Generalized Additive Models. London, Chapman & Hall.

Ikemoto T. (2005): Intrinsic optimum temperature for development of insects and mites. Environmental Entomology, 34: 1377–1387.

Ikemoto T. (2008): Tropical malaria does not mean hot environments. Journal of Medical Entomology, 45: 963–969.

Jacoby W.G. (2000): Loess: a nonparametric, graphical tool for depicting relationships between variables. Electoral Studies, 19: 577–613.

Kawada K. (1964): The influence of constant temperature on the development, longevity and reproduction of the turnip aphid, *Rhopalosiphum pseudobrassicae* Davis. Berichte des Ohara Institutes für landwirtschaftliche Biologie, 12: 243–250.

Kim T., Ahn J., Lee J.H. (2009): Temperature-dependent developmental model of *Neoseiulus californicus* (McGregor) (Acari, P

hytoseiidae). Journal of Applied Entomology, 133: 284–291.

Kontodimas D.C., Eliopoulos P.A., Stathas G.J, Economou L.P. (2004): Comparative temperature-dependent devel-

doi: 10.17221/83/2016-PPS

opment of *Nephus includens* (Kirsch) and *Nephus bisignatus* (Boheman) (Coleoptera: Coccinellidae) preying on *Planococcus citri* (Risso) (Homoptera: Pseudococcidae): evaluation of a linear and various nonlinear models using specific criteria. Environmental Entomology, 33: 1–11.

Kuang X., Parajulee M.N., Shi P., Ge F., Xue F. (2011): Testing the rate isomorphy hypothesis using five statistical methods. Insect Science, 19: 121–128.

Lactin D.J., Holliday N.J., Johnson D.L., Craigen R. (1995): Improved rate model of temperature-dependent development by arthropods. Environmental Entomology, 24: 68–75.

Liu S., Meng X. (1990): The relationships between temperature and development rates of *Myzus persicae* and *Lipaphis erysimi*. Acta Phytophylacica Sinica, 17: 169–175. (in Chinese)

Loader C. (1999): Local Regression and Likelihood. New York, Springer.

Logan J.A., Wollkind D.J., Hoyt S.C., Tanigoshi L.K. (1976): An analytic model for description of temperature dependent rate phenomena in arthropods. Environmental Entomology, 5: 1133–1140.

Qin X., Li Z. (2006): Effect of temperature on growth, development, reproduction and survival of *Myzus persicae*. Chinese Agricultural Science Bulletin, 22: 365–370. (in Chinese)

Ratkowsky D.A., Olley J., Ross T. (2005): Unifying temperature effects on the growth rate of bacteria and the stability of globular proteins. Journal of Thermal Biology, 233: 351–362.

Satar S., Kersting U., Uygun N. (2005): Effect of temperature on development and fecundity of *Aphis gossypii* Glover (Homoptera: Aphididae) on cucumber. Journal of Pest Science, 78: 133–137.

Schoolfield R.M., Sharpe P.J.H., Magnuson C.E. (1981): Non-linear regression of biological temperature-dependent rate models based on absolute reaction-rate theory. Journal of Thermal Biology, 88: 719–731.

Sharpe P.J.H., DeMichele D.W. (1977): Reaction kinetics of poikilotherm development. Journal of Thermal Biology, 64: 649–670.

Shi P., Ge F. (2010): A comparison of different thermal performance functions describing temperature-dependent development rates. Journal of Thermal Biology, 35: 225–231.

Shi P., Ge F., Sun Y., Chen C. (2011a): A simple model for describing the effect of temperature on insect development rate. Journal of Asia-Pacific Entomology, 14: 15–20.

Shi P., Ikemoto T., Egami C., Ge F., Sun Y., Cao H. (2011b): A modified program for estimating the parameters of the SSI model. Environmental Entomology, 40: 462–469.

Smits N., Brière J.F., Fargues J. (2003): Comparison of nonlinear temperature-dependent development rate models applied to *in vitro* growth of entomopathogenic fungi. Mycological Research, 107: 1476–1484.

Taylor F. (1981): Ecology and evolution of physiological time in sects. American Naturalist, 117: 1–23.

Wang R.S., Lan Z.X., Ding Y.Q. (1982): Studies on mathematical models of the relationship between insect development and temperature. Acta Ecologica Sinica, 2: 47–57. (in Chinese)

Zhao Z.H., He D.H., Hui C. (2012): From the inverse density-area relationship to the minimum patch size of a host-parasitoid system. Ecological Research, 27: 303–308.

Zhao Z.H., Guan X.Q., He D.H., Ge F. (2013): Species composition and diversity of parasitoids and hyper-parasitoids in different wheat agro-farming systems. Journal of Insect Science, 13 (162): 1–8. doi: 10.1673/031.013.16201

Received: 2016–05–27
Accepted after corrections: 2017–02–08
Published online: 2017–05–11