Yap *et al. BMC Plant Biology* 2011, **11**:23
http://www.biomedcentral.com/1471-2229/11/23

BMC Plant Biology

METHODOLOGY ARTICLE Open Access

Functional mapping of reaction norms to multiple environmental signals through nonparametric covariance estimation

John S Yap[1], Yao Li[2], Kiranmoy Das[3], Jiahan Li[3], Rongling Wu[4,3]*

Abstract

Background: The identification of genes or quantitative trait loci that are expressed in response to different environmental factors such as temperature and light, through functional mapping, critically relies on precise modeling of the covariance structure. Previous work used separable parametric covariance structures, such as a Kronecker product of autoregressive one [AR(1)] matrices, that do not account for interaction effects of different environmental factors.

Results: We implement a more robust nonparametric covariance estimator to model these interactions within the framework of functional mapping of reaction norms to two signals. Our results from Monte Carlo simulations show that this estimator can be useful in modeling interactions that exist between two environmental signals. The interactions are simulated using nonseparable covariance models with spatio-temporal structural forms that mimic interaction effects.

Conclusions: The nonparametric covariance estimator has an advantage over separable parametric covariance estimators in the detection of QTL location, thus extending the breadth of use of functional mapping in practical settings.

Background

The phenotype of a quantitative trait exhibits *plasticity* if the trait differs in phenotypes with changing environment [1-7]. Such environment-dependent changes, also called *reaction norms*, are ubiquitous in biology. For example, thermal reaction norms show how performance, such as caterpillar growth rate [8] or growth rate and body size in ectotherms [9], varies continuously with temperature [10]. Another example is the flowering time of *Arabidopsis thaliana* with respect to changing light intensity [11]. However, QTL mapping of reaction norms is difficult to model because of the inherent complexity in the interplay of a multitude of factors involved. An added difficulty is in their being "infinite-dimensional" as they require an infinite number of measurements to be completely described [12]. Wu et al. [13] proposed a functional mapping-based model which addresses the latter difficulty by using a biologically relevant mathematical function to model reaction norms. The authors considered a parametric model of photosynthetic rate as a function of light irradiance and temperature and studied the genetic mechanism of such process. They showed through simulations that in a backcross population with one or two-QTLs, their method accurately and precisely estimated the QTL location(s) and the parameters of the mean model for photosynthesis rate. For a backcross population with one QTL, the mean model consists of two surfaces that describe the photosynthetic rate of two genotypes. However, in their model, they assumed the covariance matrix to be a Kronecker product of two AR(1) structures, each modeling a reaction norm due to one environmental factor. This type of covariance model is said to be *separable*. Although computationally efficient because of the minimal number of parameters to be estimated, this model only captures separate reaction norm effects but fails to incorporate interactions. A more general approach is therefore needed.

* Correspondence: rwu@hes.hmc.psu.edu
[4]Center for Computational Biology, Beijing Forestry University, Beijing 100083, PR China
Full list of author information is available at the end of the article

In the context of longitudinal data, Yap et al. [14] proposed a nonparametric covariance estimator in functional mapping. It was nonparametric in the sense that the covariance matrix has an unconstrained set of parameters to be estimated and not the usual distribution-free sense in nonparametric statistics. This estimator can be obtained by employing a modified Cholesky decomposition of the covariance matrix which yields component matrices whose elements can be interpreted and modeled as terms in a regression [15]. A penalized likelihood procedure is used to solve the regression with either an L_1 or L_2 penalty [16]. Penalized likelihood in regression is a technique used to obtain minimum mean squared error (MSE) of estimated regression coefficients by balancing bias and variance. L_1 or L_2 penalties, which are functions of the regression covariates, are included in a regression model in order to shrink coefficients towards estimates with minimum MSE. In the case of the L_1 penalty, some of the coefficients are actually shrunk to zero. Thus, with the L_1 penalty, a more parsimonious regression model is obtained. The use of penalized likelihood with L_1 or L_2 penalties is particularly useful when there is multi-collinearity among the covariates in the regression i.e. when there are near linear dependencies or high correlations among the regressors or predictor variables. An iterative procedure is implemented by using the ECM algorithm [17] to obtain the final estimator. Through Monte Carlo simulations, this nonparametric estimator is found to provide more accurate and precise mean parameters and QTL location estimates than the parametric AR(1) form for the covariance model, especially when the underlying covariance structure of the data is significantly different from the assumed model.

The question of how to incorporate interaction effects in a model with multiple factors has not, to our knowledge, been thoroughly explored in the biology literature, especially in the context of genetic mapping that incorporates interactions of function-valued traits. The spatio-temporal literature, however, has a wealth of publications that developed more general models such as *nonseparable* covariance structures which are used to model the underlying interactions of random processes in the space and time domains (see [18,19]). A nonseparable covariance cannot be expressed as a Kronecker product of two matrices like separable structures can. The random processes being modeled may be the concentration of pollutants in the atmosphere, groundwater contaminants, wind speed, or even disposable household incomes. The main significance of the covariance in this context is in providing a better characterization of the random process to obtain optimal *kriging* or prediction of unobserved portions of it. It therefore seems natural to consider the utilization of nonseparable structures in the simulation and modeling of reaction norms that react to two environmental factors. More concretely, we consider the photosynthetic rate as a random process, and the irradiance and temperature as the spatial (one dimension) and temporal domains, respectively.

The remaining part of this paper is organized as follows: We first describe the functional mapping model proposed by Wu et al. [13] for reaction norms. Then, we formulate separable and nonseparable models used in spatio-temporal analyses and present a simulation study using some nonseparable structures. Lastly, the new model and its implications for genetic mapping are discussed. From hereon, the terms covariance matrix, covariance structure or covariance function are used interchangeably.

Functional Mapping of Reaction Norms

Reaction Norms: An Example

Wolf [20] described a reaction norm as a surface landscape determined by genetic and environmental factors. The surface is characterized by a phenotypic trait as a function of different environmental factors such as temperature, light intensity, humidity, etc., and corresponds to a specific genetic effect such as additive, dominant or epistatic [21]. At least in three dimensions, the features of the surface such as “slope”, “curvature”, “peak valley”, and “ridge”, can be described graphically to help visualize and elucidate how the underlying factors affect the phenotype.

An example of reaction norms that illustrate a surface landscape is *photosynthesis* [13], the process by which light energy is converted to chemical energy by plants and other living organisms. It is an important yet complex process because it involves several factors such as the age of a leaf (where photosynthesis takes place in most plants), the concentration of carbon dioxide in the environment, temperature, light irradiance, available nutrients and water in the soil. A mathematical expression for the rate of single-leaf photosynthesis, P, without photorespiration [22] is

$$P = \frac{\alpha I + P_m}{2\theta} - \frac{\sqrt{b^2 - 4\theta\alpha I P_m}}{2\theta} \tag{1}$$

where $b = (\alpha I + P_m$, $\theta \in (0,1)$ is a dimensionless parameter, α is the photochemical efficiency, I is the irradiance, and P_m is the asymptotic photosynthetic rate at a saturating irradiance. P_m is a linear function of the temperature, T

$$P_m = \begin{cases} P_m(20)P(T) & T \geq T* \\ 0 & T < T^*, \end{cases} \tag{2}$$

where $P(T) = \frac{T - T^*}{20 - T^*}$, $P_m(20)$ is the value of P_m at the reference temperature of 20°C and T^* is the temperature at which photosynthesis stops. T^* is chosen over a range of temperatures, such as 5°C-25°C, to provide a good fit to observed data.

Wu et al. [13] studied the reaction norm of photosynthetic rate, defined by Eqs. (1) and (2), as a function of irradiance (I) and temperature (T). That is, the authors considered $P = P(I, T)$. We assume that $T^* = 5$ so that the reaction norm model parameters are (α, $P_m(20)$, θ). The surface landscape that describes the reaction norm of P (I,T), with parameters (α, $P_m(20)$, θ) = (0.02, 1, 0.9), is shown in Figure 1. As stated earlier, each reaction norm surface corresponds to a specific genetic effect. Thus, if a QTL is at work, the genetic effects produce different surfaces defined by distinct sets of model parameters corresponding to different genotypes.

Likelihood

We consider a backcross design with one QTL. Extensions to more complicated designs and the two-QTL case, as in [13], are straightforward. Assume a backcross plant population of size n with a single QTL affecting the phenotypic trait of photosynthetic rate. The photosynthetic rate for each progeny i (i = 1, ..., n) is measured at different irradiance (s = 1, ..., S) and temperature (t = 1, ..., T) levels. This choice of variables is adopted for consistency in later discussions as we will be working with spatio-*t*emporal covariance models. The set of phenotype measurements or observations can be written in vector form as

$$\mathbf{y}_i = \underbrace{[y_i(1,1), \ldots, y_i(1,T),}_{\text{irradiance 1}} \ldots, \underbrace{[y_i(S,1), \ldots, y_i(S,T)',}_{\text{irradiance S}} \quad (3)$$

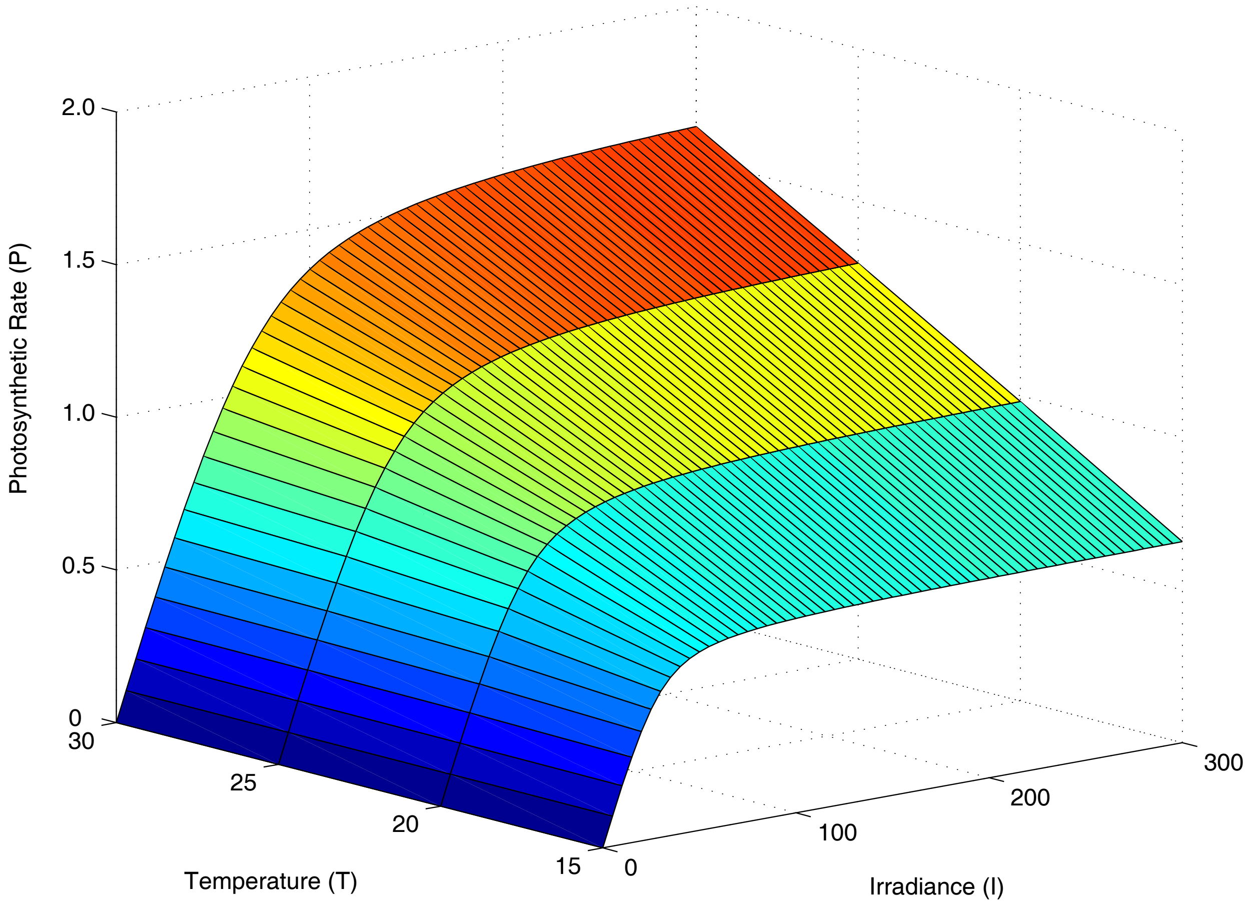

Figure 1 Reaction norm surface of photosynthetic rate as a function of irradiance and temperature. Model is based on equations (1) and (2) with parameters (α, $P_m(20)$, θ) = (0.02, 1, 0.9). Adapted from [13].

The progeny are genotyped for molecular markers to construct a genetic linkage map for the segregating QTL in the population. This means that the genotypes of the markers are observed and will be used, along with the phenotype measurements, to predict the QTL. With a backcross design, the QTL has two possible genotypes (as do the markers) which shall be indexed by k = 1, 2. The likelihood function based on the phenotype and marker data can be formulated as

$$L(\Omega) = \prod_{i=1}^{n} \left[\sum_{k=1}^{2} p_{k|i} f_k(\mathbf{y}_i \mid \Omega) \right] \tag{4}$$

where $p_{k|i}$ is the conditional probability of a QTL genotype given the genotype of a marker interval for progeny i. We assume a multivariate normal density for the phenotype vector y_i with genotype-specific means

$$\begin{aligned} \mu_k = [&\underbrace{\mu_k(1,1), \ldots, \mu_k(1,T),}_{\text{irradiance 1}} \\ &\ldots, [\underbrace{\mu_k(S,1), \ldots, \mu_k(S,T)',}_{\text{irradiance S}} \end{aligned} \tag{5}$$

and covariance matrix $\Sigma = \mathrm{cov}(\mathrm{y}_i)$.

Mean and Covariance Models

The mean vector for photosynthetic rate in (5) can be modeled using equations (1) and (2) as

$$\begin{aligned} \mu_k(s, t) = & \frac{\alpha_k s + P_{mk}}{2\theta_k} \\ & - \frac{\sqrt{b_k^2 - 4\theta_k \alpha_k s P_{mk}}}{2\theta_k} \end{aligned} \tag{6}$$

Where $b_k = \alpha_k s + P_{mk}$,

$$P_{mk}(t) = \begin{cases} P_{mk}(20)P(t) & t \geq T^* \\ 0 & t < T^* \end{cases} \tag{7}$$

$P(t) = \dfrac{t - T^*}{20 - T^*}$ and k = 1, 2.

Wu et al. [13] used a separable structure (Mitchell et al., 2005) for the $ST \times ST$ covariance matrix Σ as

$$\Sigma_{AR(1)} = \Sigma_1 \otimes \Sigma_2 \tag{8}$$

where Σ_1 and Σ_2 are the (S×S) and (T×T) covariance matrices among different irradiance and temperature levels, respectively, and ⊗ is the Kronecker product operator. Note that Σ_1 and Σ_2 are unique only up to multiples of a constant because for some |c| > 0, $c\Sigma_1 \otimes (1/c)\Sigma_2 = \Sigma_1 \otimes \Sigma_2$. Each of Σ_1 and Σ_2 is modeled using an AR(1) structure with a common error variance, σ^2, and correlation parameters ρ_k (k = 1, 2):

$$\Sigma_k = \sigma^2 \begin{bmatrix} 1 & \rho_k & \cdots & \rho_k^{S-1} \\ \rho_k & 1 & \cdots & \rho_k^{S-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_k^{S-1} & \rho_k^{S-2} & \cdots & 1 \end{bmatrix} \tag{9}$$

Separable covariance structures, however, cannot model interaction effects of each reaction norm to temperature and irradiance. Thus, there is a need for a more general model for this purpose.

Yap et al. [14] proposed to use a data-driven nonparametric covariance estimator in functional mapping. The authors showed that using such estimator provides better estimates for QTL location and mean model parameters when compared to AR(1). Huang et al. [16] showed that the nonparametric estimator works well for large matrices. Functional mapping of reaction norms when there are two environmental signals necessitates the use of large covariance matrices that result from Kronecker products of smaller matrices. Here, we are interested in determining whether the nonparametric covariance estimator of Yap et al. [14] will still work well in this reaction norm setting.

It should be noted that unlike parametric models, e.g. AR(1), there are no parameters being estimated in the nonparametric covariance estimator. The entries of the matrix are determined based on the data. This is different from a model-dependent covariance matrix model with one parameter for each of its elements. Due to over-parametrization, such a model may not lead to convergence to yield reliable results.

Note that with (6)-(9), $\mathbf{\Omega} = \mathbf{\Omega_1} \cup \mathbf{\Omega_2}$ in (4), where $\mathbf{\Omega_1} = \{\alpha_1, P_{m1}(20), \theta_1, \sigma^2, \rho_1\}$ and $\mathbf{\Omega_1} = \{\alpha_2, P_{m2}(20), \theta_2, \sigma^2, \rho_2\}$. These model parameters may be estimated using the ECM algorithm [17], but closed form solutions at the CM-step are be very complicated. A more efficient method is to use the Nelder-Mead simplex algorithm [23] which can be easily implemented using softwares such as Matlab.

Hypothesis Tests

The features of the surface landscape are important because they can be used as a basis in formulating hypothesis tests. Let H_0 and H_1 denote the null and alternative hypotheses, respectively. Then the existence of a QTL that determines the reaction norm curves can be formulated as

$$H_0 : \alpha_1 = \alpha_2, P_{m1}(20) = P_m(20), \theta_1 = \theta_2,$$

versus

H_1 : at least one of the equalities

above does not hold

This means that if the reaction norm curves are distinct (in terms of their respective estimated parameters), then a QTL possibly exists. The estimated location of the QTL is at the point at which the log-likelihood ratio obtained using the null and alternative hypotheses is maximal. Of course a slight difference in parameter estimates does not automatically mean a QTL exists. The significance of the results can be determined by permutation tests [24] which involves a repeated application of the functional mapping model on the data where the phenotype and marker associations are broken to simulate the null hypothesis of no QTL. A significance level is then obtained based on the maximal log-likelihood ratio at each application to infer the presence or absence of a QTL (see ref. [25] for more details). A procedure described in ref. [26] can be used to test the additive effects of a QTL. Other hypotheses can be formulated and tested such as the genetic control of the reaction norm to each environmental factor, interaction effects between environmental factors on the phenotype, and the marginal slope of the reaction norm with respect to each environmental factor or the gradient of the reaction norm itself. The reader is referred to Wu et al. [13] for more details.

Spatio-Temporal Covariances

We investigate the use of parametric and nonseparable spatio-temporal covariance structures in functional mapping of photosynthetic rate as a reaction norm to the environmental factors irradiance and temperature. As stated earlier, the main idea is to model irradiance as a one-dimensional spatial variable and temperature as a temporal variable. The choice of which environmental signal is modeled as temporal or spatial is arbitrary. For more about spatio-temporal modeling, we refer the reader to [27,19].

Basic Ideas, Notation, and Assumptions

We consider a real-valued spatio-temporal random process given by

$$Y(s,t), (s,t) \in \mathbb{R}^d \times \mathbb{R}, d \in \mathbb{Z}^+ \tag{10}$$

where observations are collected at coordinates

$$(s_1,t_1),(s_2,t_2),\dots,(s_N,t_N)$$

to characterize unobserved portions of the process. This collection of coordinates are not necessarily ordered fixed levels of each trait. We will only be concerned with the case $d = 1$. Aside from those mentioned earlier, Y may also represent ozone levels, disease incidence, ocean current patterns or water temperatures. In our setting, Y represents photosynthetic rate.

If var $(Y(s, t)) < \infty$ for all $(s, t) \in \mathscr{R} \times \mathscr{R}$, then the covariance, cov $(Y(s, t), Y(s + u, t + v))$, where u and v are spatial and temporal lags, respectively, exists. We assume that the covariance is *stationary* in space and time so that for some function C,

$$\text{cov}\,(Y(s,t), Y(s+u,t+v)) = C(u,v). \tag{11}$$

This means that the covariance function C depends only on the lags and not on the values of the coordinates themselves. Stationarity is often assumed to allow estimation of the covariance function from the data [18]. We also assume that the covariance function is *isotropic* which means that it depends only on the absolute lags and not in the direction or orientation of the coordinates to each other. The covariances considered in this paper are *positive (semi-) definite* as they satisfy the following condition: for any $(s_1, t_1), \dots, (s_k , t_k) \in \mathscr{R} \times \mathscr{R}$, any real coefficients $a_1, \dots, a_k$, and any positive integer k,

$$\sum_{i=1}^{k}\sum_{j=1}^{k} a_i a_j C(s_i - s_j, t_i - t_j) \geq 0 \tag{12}$$

Note that $C(u, 0)$ and $C(0, v)$ correspond to purely spatial and purely temporal covariance functions, respectively.

In spatio-temporal analysis, the ultimate goal is optimal prediction (or kriging) of an un-observed part of the random process $Y(s, t)$ using an appropriate covariance function model. We utilize a covariance model to calculate the mixture likelihood associated with functional mapping.

Separable and Nonseparable Covariance Structures

Separable Covariance Structures

A covariance function $C(u, v|\theta)$ of a spatio-temporal process is *separable* if it can be expressed as

$$C(u,v \mid \theta) = C_1(u \mid \theta_1) C_2(v \mid \theta_2) \tag{13}$$

where $C_1(u|\theta_1)$ and $C_2(v|\theta_2)$ are purely spatial and purely temporal covariance functions, respectively, and $\theta = (\theta_1, \theta_2)'$. This representation implies that the observed joint process can be seen as a product of two independent spatial and temporal processes.

A more general definition for separability is as a Kronecker product (equation (8)). From equation (8), it can be shown that $\Sigma_{AR(1)}^{-1} = \Sigma_1^{-1} \otimes \Sigma_2^{-1}$ and $|\Sigma_{AR(1)}| = |\Sigma_1|^{d_2} |\Sigma_2|^{d_1}$, where $|\cdot|$ denotes the determinant of a matrix; d_1 and d_2 are the dimensions of Σ_1 and Σ_2, respectively. This illustrates the computational advantage of using separable models in likelihood estimation where the inverse and determinant of the covariance matrix are calculated. For a large covariance matrix of dimension *UV*, its inverse can be calculated from the inverses of its Kronecker component matrices, Σ_1 and Σ_2, with dimensions *U* and *V*, respectively. Thus, the inversion of a 100 × 100 matrix, for example, may only require the inversion of two 10 × 10 matrices. A similar argument can be used for the determinant. $\Sigma_{AR(1)}$ can be put in the form (13) as

$$\begin{aligned} C(u, v \mid \sigma^2, \rho_1, \rho_2) &= \sigma^2 \rho_1^u \ . \ \sigma^2 \rho_2^v \\ &= \sigma^4 \rho_1^u \rho_2^v, \end{aligned} \qquad (14)$$

where $u = 1, ..., U$, $v = 1, ..., V$. Note that this model assumes equidistant or regularly spaced coordinates. Thus, two consecutive or closest neighbor coordinates will have the same correlation structure as another even if their respective distances are different. A more appropriate model might be

$$C(u, v \mid \sigma^2, \rho_1, \rho_2, a, b) = \sigma^4 \rho_1^{u/a} \rho_2^{v/b} \qquad (15)$$

where *a* and *b* are scale parameters. In this model, the scale parameters correct for the uneven distances between coordinates.

Nonseparable Covariance Structures

Here, we present some nonseparable covariance models that were derived in two different ways. The details of the derivation are omitted as they are rather complicated and lengthy.

The following nonseparable covariance models were derived by Cressie and Huang [18] using the Fourier transform of the spectral density and by utilizing Bochner's Theorem [28]:

$$\begin{aligned} C(u, v) = &\frac{\sigma^2}{\sqrt{(a^2 v^2 + 1)}} \\ &\times \exp\left(-\frac{b^2 u^2}{a^2 v^2 + 1} \right), \end{aligned} \qquad (16)$$

$$C(u, v) = \frac{\sigma^2 (a \mid v \mid + 1)}{(a \mid v \mid + 1)^2 + b^2 \mid u \mid^2} \qquad (17)$$

$$\begin{aligned} C(u, v) = &\sigma^2 \exp(-a \mid v \mid - b^2 \mid u \mid^2) \\ &\times \exp(-c \mid v \mid \mid u \mid^2), \end{aligned} \qquad (18)$$

where *a*, $b \geq 0$ are scaling parameters of time and space, respectively; $c \geq 0$ is an interaction parameter of time and space, and $\sigma^2 = C(0, 0) \geq 0$. Note that when $c = 0$, (18) reduces to a separable model.

Gneiting [27] developed an approach that can produce nonseparable covariance models without relying on Fourier transform pairs. One such model is

$$\begin{aligned} C(u, v) = &\frac{\sigma^2}{(a \mid v \mid^{2\alpha} + 1)^\tau} \\ &\times \exp\left(-\frac{b \mid u \mid^{2\beta}}{(a \mid v \mid^{2\alpha} + 1)^{\beta\gamma}} \right), \end{aligned} \qquad (19)$$

with $(u, v) \in \mathscr{R} \times \mathscr{R}$ and where *a*, $b > 0$ are scaling parameters of space and time, respectively; α, $\beta \in (0, 1]$ are smoothness parameters of space and time, respectively; γ 0[1]; $\tau \geq 1/2$; and $\sigma^2 \geq 0$. γ is a space-time interaction parameter which implies a separable structure when 0 and a nonseparable structure otherwise. Increasing values of γ indicates strengthening spatio-temporal interaction.

Computer Simulation

We investigated the performances of the following nonseparable covariances structures that were presented in the preceding section

$$\begin{aligned} C_1(u, v) = &\frac{\sigma^2}{\sqrt{(a^2 v^2 + 1)}} \\ &\times \exp\left(-\frac{b^2 u^2}{a^2 v^2 + 1} \right), \end{aligned} \qquad (20)$$

$$C_2(u, v) = \frac{\sigma^2 (a \mid v \mid + 1)}{(a \mid v \mid + 1)^2 + b^2 \mid u \mid^2}, \qquad (21)$$

$$\begin{aligned} C_3(u, v) = &\frac{\sigma^2}{(a \mid v \mid + 1)} \\ &\times \exp\left(-\frac{b \mid u \mid}{(a \mid v \mid + 1)^{\gamma/2}} \right), \end{aligned} \qquad (22)$$

where a, b ≥ 0; $\gamma \in$ 0[1] and $\sigma^2 > 0$. C_1 and C_2 correspond to (16) and (17), respectively, and C_3 is a special case of (19) with $\alpha = 1/2$, $\beta = 1/2$ and $\tau = 1$.

We generated photosynthetic rate data using these nonseparable covariances to simulate interaction effects between the two environmental signals in functional mapping of a reaction norm. The generated data was analyzed using the nonparametric estimator Σ_{NP} proposed by Yap et al. [14] using an L_2 penalty, and $\Sigma_{AR(1)}$ (equation (8)). Note that the underlying covariance structures were very different from the assumed model, $\Sigma_{AR(1)}$, and we therefore expected to get biased estimates. The issue we wanted to address was the extent to which the bias cannot be ignored and an alternative estimator such as Σ_{NP} may be more appropriate.

Covariance fit was assessed using entropy (L_E) and quadratic (L_Q) losses:

$$L_E(\Sigma,\Sigma) = \mathrm{tr}(\Sigma^{-1}\Sigma) - \log\left|\Sigma^{-1}\Sigma\right| - m$$

and

$$L_Q(\Sigma,\Sigma) = \mathrm{tr}(\Sigma^{-1}\Sigma - I)^2$$

where $\hat{\Sigma}$ is the estimate of the true underlying covariance Σ [14,16,29-31]. Each loss function is 0 when $\hat{\Sigma} = \Sigma$ and large values suggest significant bias.

Using a backcross design for the QTL mapping population, we randomly generated 6 markers equally spaced on a chromosome 100 cM long. One QTL was simulated between the fourth and fifth markers, 12 cM from the fourth marker (or 72 cM from the leftmost marker of the chromosome). The QTL had two possible genotypes which determined two distinct mean photosynthetic rate reaction norm surfaces defined by equations (1) and (2) (see also Figure 1). The surface parameters for each genotype were (α_1, $P_{m1}(20)$, θ_1) = (0.02, 2, 0.9) and (α_2, $P_{m2}(20)$, θ_2) = (0.01, 1.5, 0.9). Phenotype observations were obtained by sampling from a multivariate normal distribution with mean surface based on irradiance and temperature levels of {0, 50, 100, 200, 300} and {15, 20, 25, 30}, respectively, and covariance matrix $C_l(u, v)$, l = 1, 2, 3 with a = 0.50, b = 0.01 for C_1, a = 1.00, b = 0.01 for C_2, a = 1.00, b = 0.01, c = 0.60 for C_3 and σ^2 = 1.00 for all three covariances.

Figure 2 shows the reaction norm surfaces of photosynthetic rate as functions of irradiance and temperature that were used in the simulation. Within the considered domain of values for irradiance and temperature, one surface lies above the other. These surfaces differ only in terms of the α_2 and $P_{m1}(20)$ parameters.

The functional mapping model was applied to the marker and phenotype data with n = 200, 400 samples. The surface defined by equations (1) and (2) was used as mean model with Σ_{NP} and $\Sigma_{AR(1)}$ as covariance models to analyze the data generated using C_l(u, v). 100 simulation runs were carried out and the averages on all runs of the estimated QTL location, mean parameter estimates, entropy and quadratic losses, including the respective Monte carlo standard errors (SE), were recorded. Tables 1 and 2 present the results of these simulations. The results show that using Σ_{NP} yields reasonably accurate and precise parameter estimates. The results for $\Sigma_{AR(1)}$ are similar to Σ_{NP} except that the average losses, given by L_E and L_Q, are inflated for C_1 and C_2. Figure 3 shows box plots of the log-likelihood values under the alternative model. These plots reveal biased estimates of C_1 and C_2 by $\Sigma_{AR(1)}$ and the degrees of bias are consistent with the average losses. The results for the log-likelihood values under the null model are very similar but are not shown. We also provided the covariance and corresponding contour plots of $C_l(u, v)$, l = 1, 2, 3 and the $\Sigma_{AR(1)}$ estimates of these in Figure 4 and 5. We only provided plots for $C_l(u, v)$, l = 1, 2, 3 and $\Sigma_{AR(1)}$ to illustrate the behavior of these parametric models. We did not include plots for the estimated Σ_{NP} because there are no parametric estimates for this model and we did not record all elements of the estimated Σ_{NP} in the simulation runs.

We conducted further simulations using C_1 as the underlying covariance structure of the data with n = 400. This was the case where $\Sigma_{AR(1)}$ performed the worst. We considered two scenarios: increased variance parameter, σ^2, or increased irradiance and temperature levels (finer grid). That is,

1. σ^2 = 2, 4 with irradiance and temperature levels of {0, 50, 100, 200, 300} and {15, 20, 25, 30}, respectively.
2. σ^2 = 1, 2 with irradiance and temperature levels of {0, 50, 100, 150, 200, 250, 300} and {15, 18, 21, 24, 27, 30}, respectively.

We included an analysis of the simulated data using C_1 as the covariance model to ensure the results are not false-positives. The results of the simulation are shown in Tables 3 and 4. The tables include columns for the log-likelihood values under the null (H_0) and alternative (H_1) hypotheses as well as the maximum of the log-likelihood ratio (max*LR*). Max*LR* is used in permutation tests to assess significance of QTL existence (see Section 2.3). Under scenarios (1) or (2), i.e. increased variance parameter σ^2 or increased irradiance and temperature levels, using Σ_{NP} yields significantly more accurate and precise estimates of the QTL location compared to $\Sigma_{AR(1)}$: In Table 3, when σ^2 = 4, the estimates of the true QTL location of 72 were 71.64 and 74.20 for NP and $\Sigma_{AR(1)}$, respectively; In Table 4, when σ^2 = 2, the estimates were 72.13 and 78.44. Although for $\Sigma_{AR(1)}$, max*LR* appears to be more accurate, the log-likelihood ratios are still significantly different from the estimates given by C_1. Again, this is reflected in the inflated average losses. Note that the max*LR* estimates are larger for $\Sigma_{AR(1)}$ when compared to those for Σ_{NP}. We do not expect this to be always the case. In other instances, the max*LR* estimates for $\Sigma_{AR(1)}$ may be smaller than those for Σ_{NP} . However, in those instances, we expect the max*LR* estimates for Σ_{NP} to still be more accurate and precise than

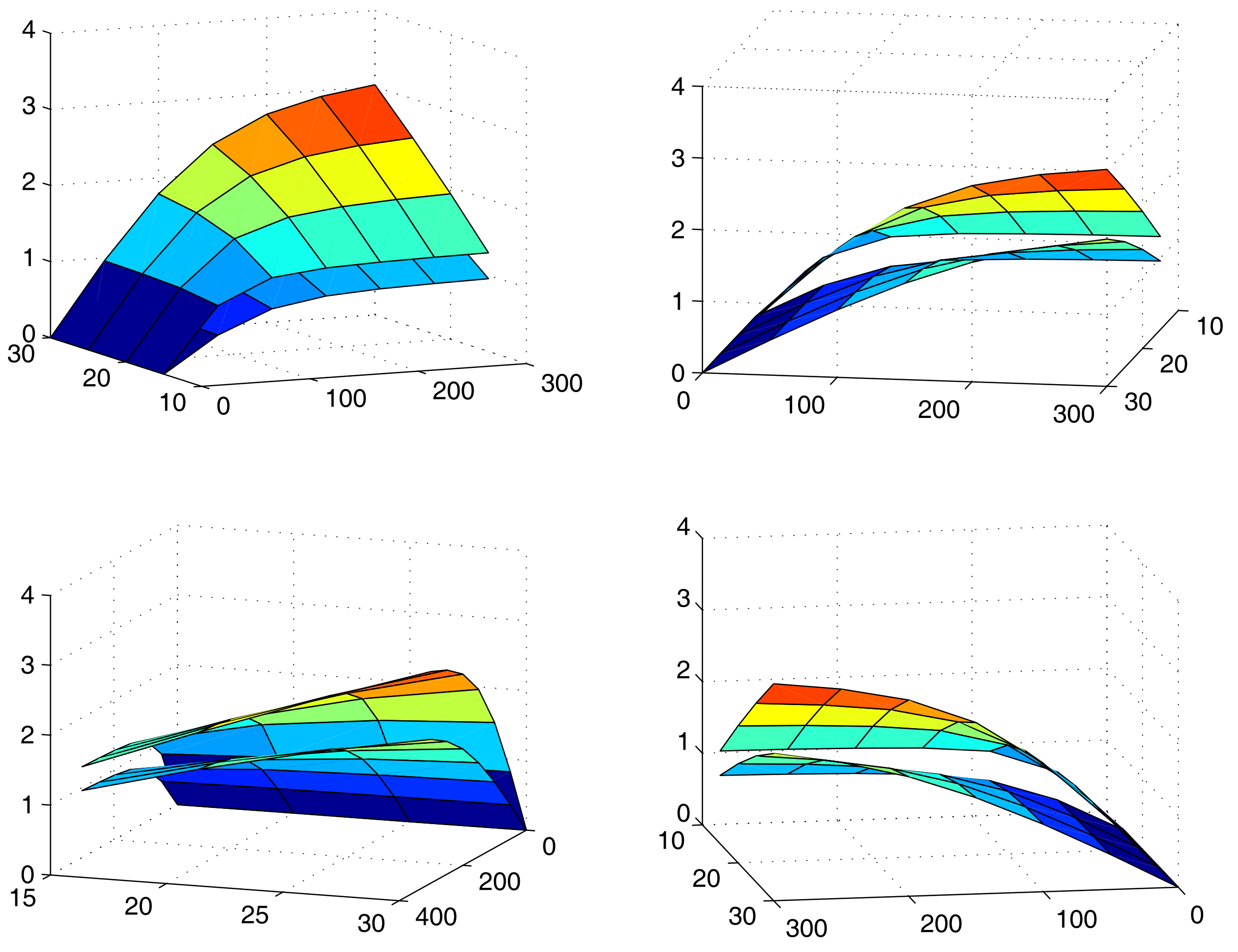

Figure 2 Reaction norm surfaces of photosynthetic rate as functions of irradiance and temperature. Models are based on equations (1) and (2) with parameters (α_1, $P_{m1}(20)$, θ_1) = (0.02, 2, 0.9) and (α_2, $P_{m2}(20)$, θ_2) = (0.01, 1.5, 0.9) as used in the simulation.

Table 1 Averaged QTL position, mean curve parameters, entropy and quadratic losses and their standard errors (given in parentheses) for two QTL genotypes in a backcross population under different sample sizes (*n*) based on 100 simulation replicates (Σ_{NP})

		QTL	QTL genotype 1			QTL genotype 2				
Covariance	*n*	Location	$\hat{\alpha}_1$	$\hat{P}_{m1}(20)$	$\hat{\theta}_1$	$\hat{\alpha}_2$	$\hat{P}_{m2}(20)$	$\hat{\theta}_2$	L_E	L_Q
C_1	200	71.68	0.02	2.02	0.90	0.01	1.52	0.88	1.04	2.03
		(0.28)	(0.00)	(0.01)	(0.00)	(0.00)	(0.02)	(0.01)	(0.01)	(0.02)
	400	72.16	0.02	2.00	0.90	0.01	1.52	0.88	0.53	1.06
		(0.23)	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)
C_2	200	71.88	0.02	2.00	0.90	0.01	1.53	0.88	1.00	1.96
		(0.29)	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.01)	(0.02)
	400	71.92	0.02	2.00	0.90	0.01	1.52	0.89	0.52	1.02
		(0.17)	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)
C_3	200	72.12	0.02	2.01	0.89	0.01	1.54	0.87	0.88	1.70
		(0.37)	(0.00)	(0.01)	(0.01)	(0.00)	(0.02)	(0.01)	(0.01)	(0.02)
	400	72.08	0.02	2.01	0.90	0.01	1.52	0.89	0.48	0.94
		(0.20)	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)
	True:	72.00	0.02	2.00	0.90	0.01	1.50	0.90		

Table 2 Averaged QTL position, mean curve parameters, entropy and quadratic losses and their standard errors (given in parentheses) for two QTL genotypes in a backcross population under different sample sizes (*n*) based on 100 simulation replicates ($\Sigma_{AR(1)}$)

		QTL	QTL genotype 1			QTL genotype 2				
Covariance	*n*	Location	$\hat{\alpha}_1$	$\hat{P}_{m1}(20)$	$\hat{\theta}_1$	$\hat{\alpha}_2$	$\hat{P}_{m2}(20)$	$\hat{\theta}_2$	L_E	L_Q
C_1	200	72.32	0.02	2.03	0.90	0.01	1.53	0.87	19.43	681.78
		(0.45)	(0.00)	(0.01)	(0.01)	(0.00)	(0.02)	(0.01)	(0.07)	(6.16)
	400	71.72	0.02	2.03	0.90	0.01	1.51	0.89	19.45	684.11
		(0.27)	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.05)	(4.40)
C_2	200	71.96	0.02	2.01	0.90	0.01	1.55	0.87	4.83	58.60
		(0.34)	(0.00)	(0.01)	(0.00)	(0.00)	(0.02)	(0.01)	(0.02)	(1.01)
	400	71.84	0.02	2.01	0.90	0.01	1.52	0.89	4.83	58.61
		(0.20)	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.02)	(0.77)
C_3	200	72.00	0.02	2.01	0.89	0.01	1.54	0.87	0.60	1.51
		(0.35)	(0.00)	(0.01)	(0.01)	(0.00)	(0.02)	(0.01)	(0.00)	(0.10)
	400	71.96	0.02	2.01	0.89	0.01	1.52	0.89	0.60	1.43
		(0.22)	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)	(0.08)
	True:	72.00	0.02	2.00	0.90	0.01	1.50	0.90		

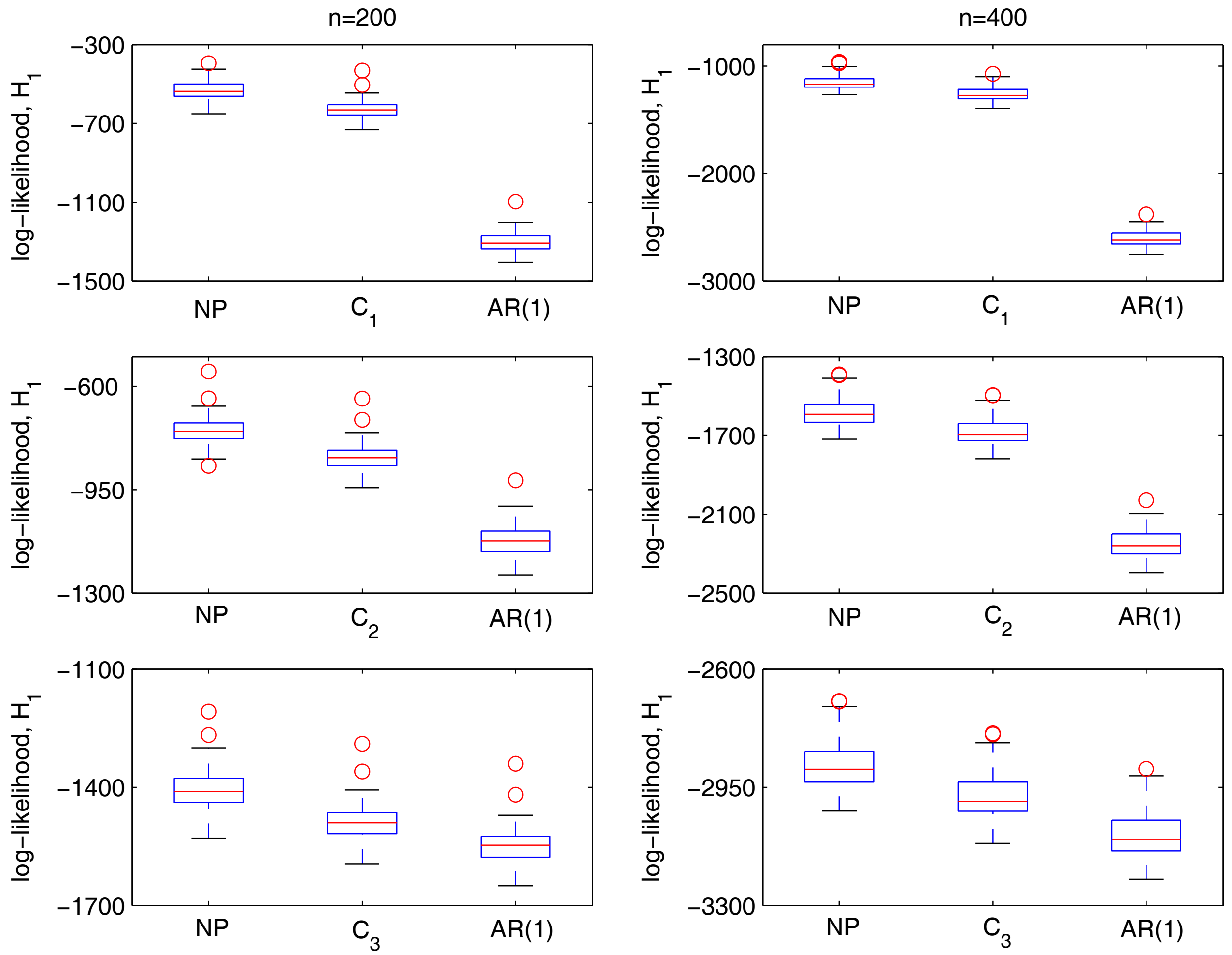

Figure 3 Boxplots of the values of the log-likelihood under the alternative model, *H*$_1$. Significantly biased estimates by $\Sigma_{AR(1)}$ are apparent for C_1.

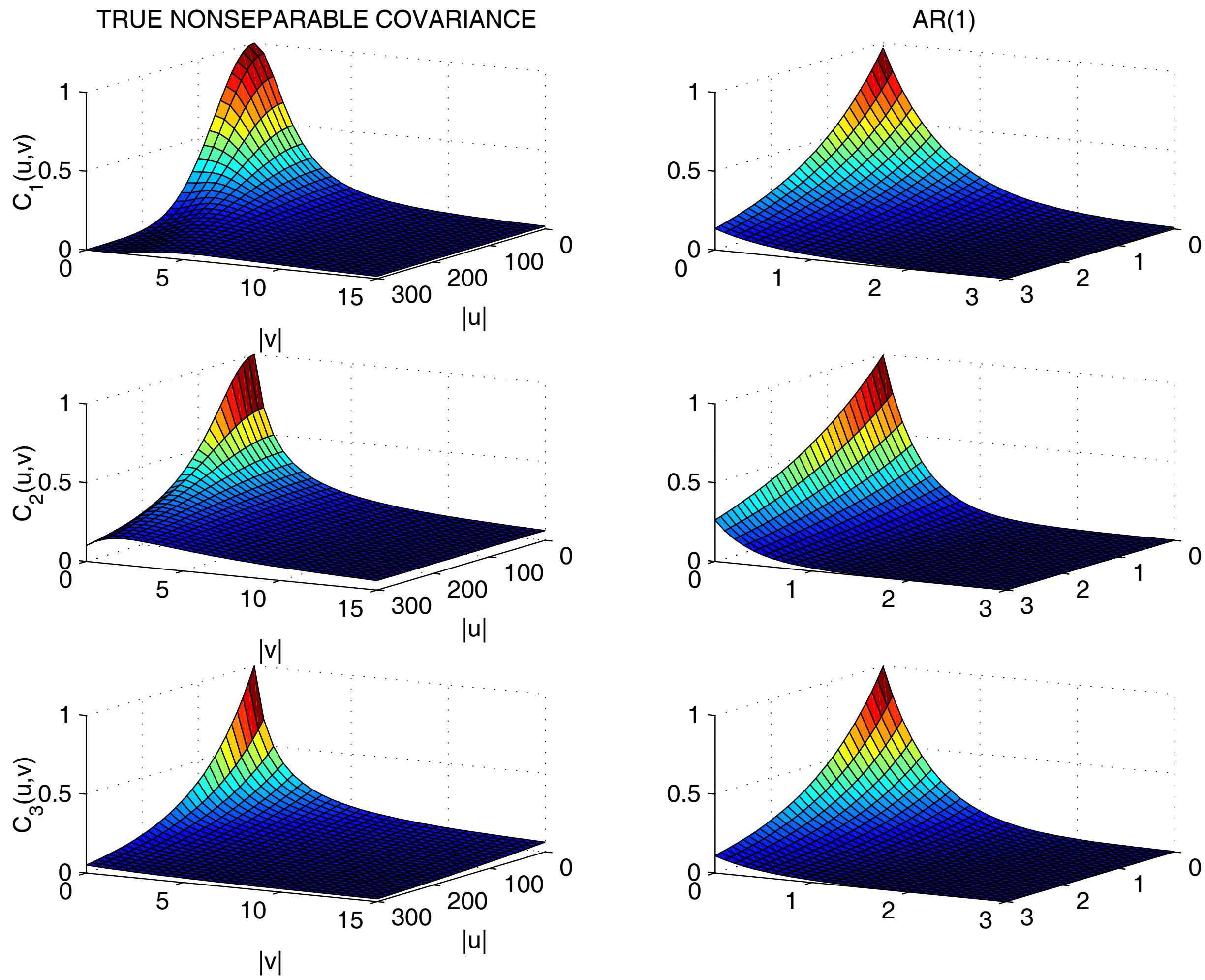

Figure 4 Covariance plots. Plots of C_l, l = 1, 2, 3 versus irradiance (|u|) and temperature (|v|) lags are on the left column. On the right column are the estimates of C_l by $\Sigma_{AR(1)}$.

those for $\Sigma_{AR(1)}$, unless the true underlying covariance structure is $\Sigma_{AR(1)}$, which is not likely.

Discussion

In this paper, we studied the covariance model in functional mapping of photosynthetic rate as a reaction norm to irradiance and temperature as environmental signals. In the presence of interaction between the two signals simulated by nonseparable covariance structures, our analysis showed that Σ_{NP} is a more reliable estimator than $\Sigma_{AR(1)}$ particularly in QTL location estimation. The advantage of Σ_{NP} over $\Sigma_{AR(1)}$ is greater when the variance of the reaction norm process and the number of signal levels increase.

Σ_{NP} was developed in the context of a one dimensional (longitudinal) vector which has an ordering of variables. The phenotype vector we considered here consists of observations based on two levels of irradiance and temperature measurements, i.e.,

$$\mathbf{y}_i = \underbrace{[y_i(1,1),\ldots,y_i(1,T)}_{\text{irradiance 1}}, \ldots, \underbrace{[y_i(S,1),\ldots,y_i(S,T)'}_{\text{irradiance S}}, \tag{23}$$

This vector has no natural ordering like in longitudinal data. However, our simulation results still suggest that Σ_{NP} can be directly applied to observations that have no variable ordering such as (23). The process by which Σ_{NP} was obtained in Yap et al. [14] was based on non-mixture type of longitudinal covariance estimators. This process is flexible and can potentially accommodate other estimators that can handle unordered data or are invariant to variable permutations. See for example

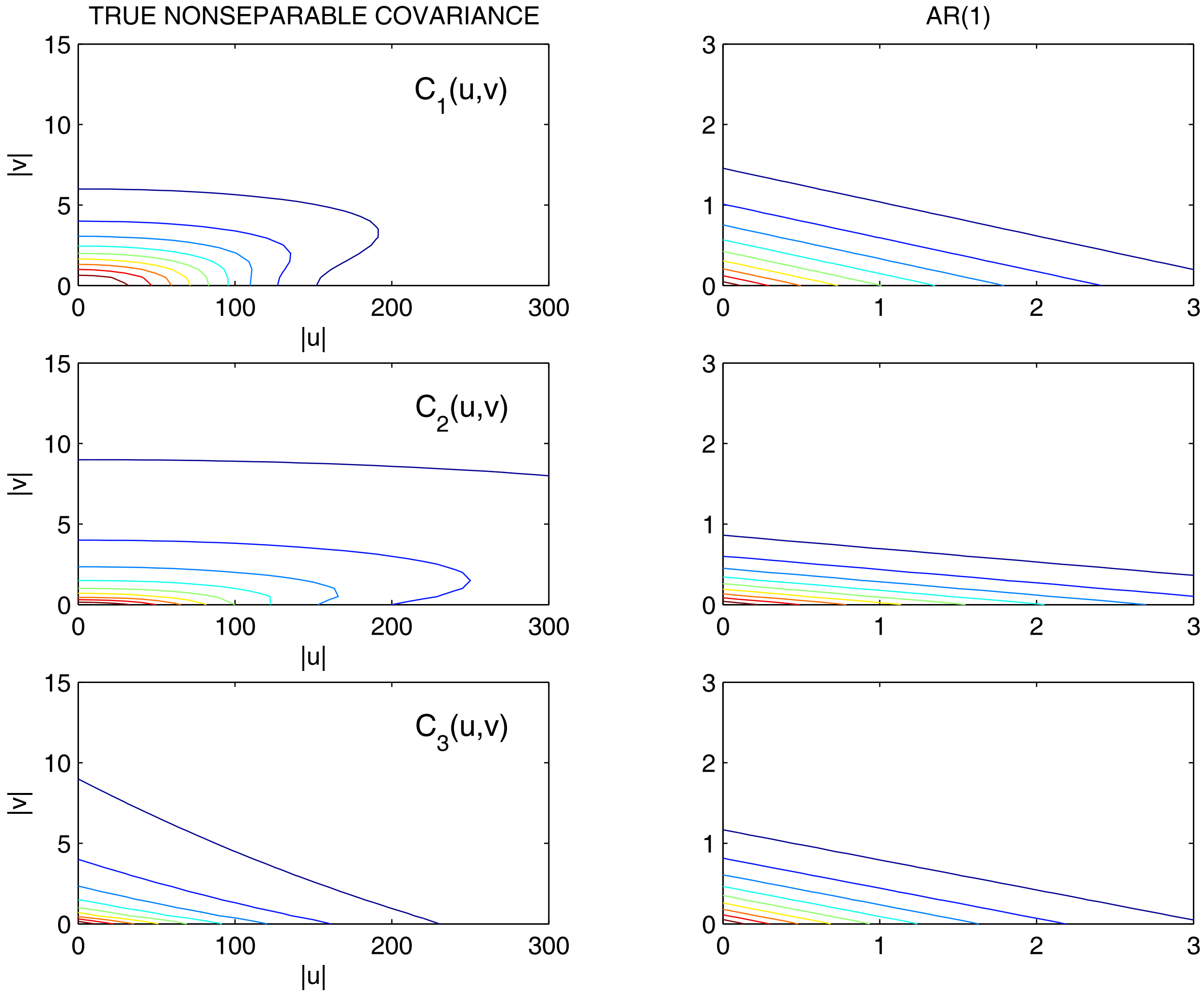

Figure 5 Contour plots. Contour plots of C_l, l = 1, 2, 3 on the left column. On the right column are the contour plots of the estimates of C_l by $\Sigma_{AR(1)}$.

the sparse permutation invariant covariance estimator (SPICE) proposed by Rothman et al. [32].

In the presence of interactions, nonseparable covariances can possibly be used in place of Σ_{NP}, but they should closely reflect the structure of the data. Unfortunately, as with any parametric model, this is not often the case. In fact, it is not even known whether the data exhibits interactions or not. Before deciding on what model to use, one might utilize tests for separability [33,34]. If separable models are appropriate, then there are many options. Otherwise, it is difficult to choose from a number of complex nonseparable covariances because there are no available general guidelines as yet that can help one decide which model to use. The covariance C_3 that was used in the simulations had an easy to interpret interaction parameter $\gamma \in$ 0[1]. However, despite an interaction "strength" of $\gamma = 0.6$, the separable model, $\Sigma_{AR(1)}$, estimated the data generated by C_3 quite well. Thus, the trade-o between using a nonseparable model instead of a separable one may not be worth it. Another option is to use separable approximations to nonseparable covariances [35]. The nonseparable covariances that we considered were assumed to be stationary and isotropic. These two assumptions may not always hold for real data. Although not specifically addressed here, using Σ_{NP} may work for data that do not satisfy these assumptions.

Finally, we only considered two environmental signals with interactions: irradiance and temperature. However, the reaction norm of photosynthetic rate is a very complex process because there are really more environmental signals at play other than these two. Theoretically, the spatial domain of spatio-temporal nonseparable covariance models can be extended to more than one

Table 3 Averaged QTL position, mean curve parameters, log-likelihood values, maximum log-likelihood ratios (max*LR*), entropy and quadratic losses and their standard errors (given in parentheses) for two QTL genotypes in a backcross population based on 100 simulation replicates (C_1 with n = 400 and σ^2 = 2, 4)

		QTL	QTL genotype 1			QTL genotype 2			log-likelihood				
Covariance	σ^2	**Location**	$\hat{\alpha}_1$	$\hat{P}_{m1}(20)$	$\hat{\theta}_1$	$\hat{\alpha}_2$	$\hat{P}_{m2}(20)$	$\hat{\theta}_2$	H_0	H_1	**max*LR***	L_E	L_Q
$\Sigma_{AR(1)}$	2	72.40	0.02	2.05	0.89	0.01	1.52	0.87	-5437	-5373	128.51	19.45	684.37
		(0.44)	(0.00)	(0.01)	(0.01)	(0.00)	(0.02)	(0.01)	(7.36)	(7.31)	(2.45)	(0.05)	(4.44)
	4	74.20	0.02	2.11	0.88	0.01	1.52	0.84	-8175	-8141	65.55	19.44	683.82
		(0.69)	(0.00)	(0.02)	(0.01)	(0.00)	(0.03)	(0.02)	(7.32)	(7.31)	(1.80)	(0.05)	(4.46)
C_1	2	71.96	0.02	2.01	0.90	0.01	1.54	0.88	-4088	-4021	133.41	0.01	0.13
		(0.29)	(0.00)	(0.01)	(0.00)	(0.00)	(0.02)	(0.01)	(7.17)	(7.16)	(2.15)	(0.00)	(0.02)
	4	71.96	0.02	2.03	0.89	0.01	1.57	0.86	-6822	-6788	69.07	0.01	0.13
		(0.44)	(0.00)	(0.01)	(0.01)	(0.00)	(0.03)	(0.02)	(7.16)	(7.16)	(1.57)	(0.00)	(0.02)
NP	2	72.16	0.02	2.01	0.89	0.01	1.54	0.87	-3967	-3912	109.79	0.53	1.05
		(0.29)	(0.00)	(0.01)	(0.00)	(0.00)	(0.02)	(0.01)	(6.87)	(6.89)	(1.66)	(0.00)	(0.01)
	4	71.64	0.02	2.01	0.89	0.01	1.57	0.84	-6713	-6684	59.92	0.53	1.04
		(0.49)	(0.00)	(0.01)	(0.01)	(0.00)	(0.03)	(0.02)	(6.89)	(6.93)	(1.27)	(0.00)	(0.01)
	True:	72.00	0.02	2.00	0.90	0.01	1.50	0.90					

dimensions i.e., d > 1 in (10). For example, a two dimensional spatial domain models an area on a flat surface while a three dimensional domain models space. There are spatio-temporal models for these. However, this extension cannot be used to increase the number of signals in a reaction norm unless the signals have the same unit of measurement or one assumes separability or no interaction among the signals. For example, carbon dioxide concentration cannot be added as a signal, in addition to irradiance and temperature, when modeling photosynthetic rate as a reaction norm in the functional mapping setting because it does not have the same unit as irradiance or temperature. Thus, it is difficult to simulate data from more than two signals with interactions. However, Σ_{NP} can theoretically handle covariances associated with more than two signals that may involve interactions. The computer code for the model will be available from http://statgen.psu.edu.

Table 4 Averaged QTL position, mean curve parameters, log-likelihood values, maximum log-likelihood ratios (max*LR*), entropy and quadratic losses and their standard errors (given in parentheses) for two QTL genotypes in a backcross population based on 100 simulation replicates (C_1 with n = 400, increased irradiance and temperature levels, and σ^2 = 1, 2)

		QTL	QTL genotype 1			QTL genotype 2			log-likelihood				
Covariance	σ^2	**Location**	$\hat{\alpha}_1$	$\hat{P}_{m1}(20)$	$\hat{\theta}_1$	$\hat{\alpha}_2$	$\hat{P}_{m2}(20)$	$\hat{\theta}_2$	H_0	H_1	**max*LR***	L_E	L_Q
$\Sigma_{AR(1)}$	1	72.16	0.02	2.04	0.90	0.01	1.48	0.88	-1278	-1063	430.01	223	64090
		(0.36)	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(14.01)	(14.15)	(4.78)	(0.45)	(261.88)
	2	78.44	0.02	2.15	0.91	0.01	1.48	0.86	-6992	-6876	231.86	222	63923
		(0.84)	(0.00)	(0.02)	(0.00)	(0.00)	(0.02)	(0.01)	(14.08)	(14.16)	(3.62)	(0.44)	(257.89)
C_1	1	71.76	0.02	2.01	0.90	0.01	1.51	0.89	4913	5068	309.86	0.01	0.31
		(0.18)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)	(11.04)	(11.10)	(3.17)	(0.00)	(0.04)
	2	71.76	0.02	2.01	0.90	0.01	1.52	0.88	-821.08	-743.76	154.64	0.01	0.31
		(0.24)	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(11.10)	(11.12)	(2.22)	(0.00)	(0.04)
NP	1	71.73	0.02	2.01	0.90	0.01	1.51	0.89	5431	5537	212.64	2.34	4.55
		(0.18)	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)	(0.00)	(11.22)	(11.11)	(2.20)	(0.01)	(0.03)
	2	72.13	0.02	2.01	0.90	0.01	1.49	0.89	-336	-273	127.37	2.37	4.53
		(0.34)	(0.00)	(0.01)	(0.00)	(0.00)	(0.01)	(0.01)	(10.44)	(10.42)	(1.72)	(0.01)	(0.03)
	True:	72.00	0.02	2.00	0.90	0.01	1.50	0.90					

Acknowledgements

This work is partially supported by NSF grant IOS-0923975, the Changjiang Scholarship Award and "One-thousand Person Plan" Award at Beijing Forestry University.

Author details

[1]Department of Statistics, University of Florida, Gainesville, FL 32611 USA. [2]Department of Statistics, West Virginia University, Morgantown, WV 26506, USA. [3]Center for Statistical Genetics, Pennsylvania State University, Hershey, PA 17033, USA. [4]Center for Computational Biology, Beijing Forestry University, Beijing 100083, PR China.

Authors' contributions

JY participated in the design of the study, performed the statistical analysis, and wrote the manuscript. YL, KD and JL participated in the statistical analysis. RW conceived of the study, participated in its design and coordination, and wrote the manuscript. All authors read and approved the final manuscript.

Received: 21 November 2009 Accepted: 26 January 2011
Published: 26 January 2011

References

1. Via S, Gomulkievicz R, de Jong G, Scheiner SM, *et al*: **Adaptive phenotypic plasticity: Consensus and controversy.** *Trends in Ecology and Evolution* 1995, **10**:212-217.
2. Scheiner SM: **Genetics and evolution of phenotypic plasticity.** *Annual Reviews of Ecology and Systematics* 1993, **24**:35-68.
3. Schlichting CD, Smith H: **Phenotypic plasticity: Linking molecular mechanisms with evolutionary outcomes.** *Evolutionary Ecology* 2002, **16**:189-201.
4. West-Eberhard MJ: *Developmental Plasticity: An Evolution* Oxford University Press, New York; 2003.
5. Wu RL: **The detection of plasticity genes in heterogeneous environments.** *Evolution* 1998, **52**:967-977.
6. Wu RL, Grissom JE, McKeand SE, O'Malley DM: **Phenotypic plasticity of fine root growth increases plant productivity in pine seedlings.** *BMC Ecology* 2004, **4**:14.
7. de Jong G: **Evolution of phenotypic plasticity: Patterns of plasticity and the emergence of ecotypes.** *New Phytologist* 2005, **166**:101-117.
8. Kingsolver JG, Izem R, Ragland GJ: **Plasticity of size and growth in fluctuating thermal environments: comparing reaction norms and performance curves.** *Integrative and Comparative Biology* 2004, **44**:450-460.
9. Angilletta MJ Jr, Sears MW: **Evolution of thermal reaction norms for growth rate and body size in ectotherms: an introduction to the symposium.** *Integrative and Comparative Biology* 2004, **44**:401-402.
10. Yap JS, Wang CG, Wu RL: **A simulation approach for functional mapping of quantitative trait loci that regulate thermal performance curves.** *PLoS ONE* 2007, **2(6)**:e554.
11. Stratton D: **Reaction norm functions and QTL-environment interactions for flowering time in Arabidopsis thaliana.** *Heredity* 1998, **81**:144-155.
12. Kirkpatrick M, Heckman N: **A quantitative genetic model for growth, shape, reaction norms, and other infinite-dimensional characters.** *Journal of Mathematical Biology* 1989, **27**:429-450.
13. Wu J, Zeng Y, Huang J, Hou W, Zhu J, Wu RL: **Functional mapping of reaction norms to multiple environmental signals.** *Genetical Research* 2007, **89**:27-38.
14. Yap JS, Fan J, Wu RL: **Nonparametric covariance estimation in functional map-ping of quantitative trait loci.** *Biometrics* 2009, **65**:1068-1077.
15. Pourahmadi M: **Joint mean-covariance models with applications to longitudinal data: Unconstrained parameterisation.** *Biometrika* 1999, **86(3)**:677-690.
16. Huang J, Liu N, Pourahmadi M, Liu L: **Covariance selection and estimation via penalised normal likelihood.** *Biometrika* 2006, **93**:85-98.
17. Meng X-L, Rubin D: **Maximum likelihood estimation via the ECM algorithm: A general framework.** *Biometrika* 1993, **80**:267-278.
18. Cressie N, Huang H-C: **Classes of nonseparable, spatio-temporal stationary covariance functions.** *Journal of the American Statistical Association* 1999, **94**:1330-1340.
19. Gneiting T, Genton M, Guttorp P: **Geostatistical space-time models, stationary, separability and full symmetry.** In *Statistical Methods for Spatio-temporal Systems (Monographs on Statistics and Applied Probability).* Edited by: Finkenstadt B, Held L, Isham V. Chapman 2006:.
20. Wolf JB: **The geometry of phenotypic evolution in developmental hyperspace.** *Proceedings of the National Academy of Sciences of the USA* 2002, **99**:15849-15851.
21. Wu RL, Ma C-X, Casella G: *Statistical Genetics of Quantitative Traits: Linkage, Maps, and QTL* Springer-Verlag, New York; 2007.
22. Thornley JHM, Johnson IR: *Plant and Crop Modelling: A Mathematical Approach to Plant and Crop Physiology* Clarendon Press, Oxford; 1990.
23. Nelder J, Mead R: **A simplex method for function minimization.** *Computer Journal* 1965, **7**:308-313.
24. Doerge RW, Churchill GA: **Permutation tests for multiple loci affecting a quantitative character.** *Genetics* 1996, **142**:285-294.
25. Ma C, Casella G, Wu RL: **Functional mapping of quantitative trait loci underlying the character process: A theoretical framework.** *Genetics* 2002, **161**:1751-1762.
26. Wu RL, Ma C-X, Lin M, Casella G: **A general framework for analyzing the genetic architecture of developmental characteristics.** *Genetics* 2004, **166**:1541-1551.
27. Gneiting T: **Nonseparable, stationary covarience functions for space-time data.** *Journal of the American Statistical Association* 2002, **97**:590-600.
28. Bochner S: *Harmonic Analysis and the Theory of Probability* University of California Press, Berkley and Los Angeles; 1955.
29. Wu WB, Pourahmadi M: **Nonparametric estimation of large covariance matrices of longitudinal data.** *Biometrika* 2003, **90**:831-844.
30. Huang J, Liu L, Liu N: **Estimation of large covariance matrices of longitudinal data with basis function approximations.** *Journal of Computational and Graphical Statistics* 2007, **16**:189-209.
31. Levina E, Rothman A, Zhu J: **Sparse estimation of large covariance matrices via a nested lasso penalty.** *Annals of Applied Statistics* 2008, **2**:245-263.
32. Rothman A, Bickel P, Levina E, Zhu J: **Sparse permutation invariant covariance estimation.** *Electronic Journal of Statistics* 2008, **2**:494-515.
33. Mitchell MW, Genton MG, Gumpertz ML: **Testing for separability of space-time covariences.** *Envirometrics* 2005, **16**:819-831.
34. Fuentes M: **Testing separability of spatial-temporal covariance functions.** *Journal of Statistical Planning and Inference* 2005, **136**:447-466.
35. Genton M: **Separable approximations of space-time covariance matrices.** *Envirometrics* 2007, **18**:681-695.

doi:10.1186/1471-2229-11-23
Cite this article as: Yap *et al.*: Functional mapping of reaction norms to multiple environmental signals through nonparametric covariance estimation. *BMC Plant Biology* 2011 **11**:23.