Biometrics, 2025, 81(1), ujae156
https://doi.org/10.1093/biomtc/ujae156
Biometric Methodology

Gaussian processes for time series with lead–lag effects with applications to biology data

Wancen Mu[1], Jiawen Chen [1], Eric S. Davis [2], Kathleen Reed[3], Douglas Phanstiel [2,3,4], Michael I. Love[1,5], Didong Li[1,*]

[1]Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, United States, [2]Curriculum in Bioinformatics and Computational Biology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, United States, [3]Curriculum in Genetics and Molecular Biology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, United States, [4]Department of Cell Biology and Physiology, University of North Carolina at Chapel Hil, Chapel Hill, NC 27599, United States, [5]Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27514, United States

[*]Corresponding author: Didong Li, Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA (didongli@unc.edu).

ABSTRACT

Investigating the relationship, particularly the lead–lag effect, between time series is a common question across various disciplines, especially when uncovering biological processes. However, analyzing time series presents several challenges. Firstly, due to technical reasons, the time points at which observations are made are not at uniform intervals. Secondly, some lead–lag effects are transient, necessitating time-lag estimation based on a limited number of time points. Thirdly, external factors also impact these time series, requiring a similarity metric to assess the lead–lag relationship. To counter these issues, we introduce a model grounded in the Gaussian process, affording the flexibility to estimate lead–lag effects for irregular time series. In addition, our method outputs dissimilarity scores, thereby broadening its applications to include tasks such as ranking or clustering multiple pairwise time series when considering their strength of lead–lag effects with external factors. Crucially, we offer a series of theoretical proofs to substantiate the validity of our proposed kernels and the identifiability of kernel parameters. Our model demonstrates advances in various simulations and real-world applications, particularly in the study of dynamic chromatin interactions, compared to other leading methods.

KEYWORDS: identifiability; interpretability; irregular time series; parameter inference.

1 INTRODUCTION

The lead–lag effect is a widespread phenomenon where changes in one time series (the leading series) influence another time series (the influenced follower series) after a certain delay. This phenomenon is prevalent in a variety of fields, such as climate science (De Luca and Pizzolante, 2021), healthcare (Feng et al., 2021; He and Huang, 2024; Huang and He, 2023; Runge et al., 2019; Zhu and Gallego, 2021), economics (Skoura, 2019; Wang et al., 2017), and finance (Bacry et al., 2013; Da Fonseca and Zaatour, 2017; Hayashi and Koike, 2018; Hoffmann et al., 2013; Ito and Sakemoto, 2020). For example, Harzallah and Sadourny (1997) explore the lead–lag relationship between the Indian summer monsoon and various climate variables, such as sea surface temperature, snow cover, and geopotential height. However, this dynamic, rooted in causal regulatory relationships, is notably observed in biological processes, which serves as the primary inspiration for the proposed method. The lead–lag effect manifests itself in various processes, including development (Ding and Bar-Joseph, 2020; Gerstein et al., 2010; Strober et al., 2019), disease-associated genetic variation (Krijger and De Laat, 2016), and responses to various interventions and treatments (Faryabi et al., 2008; Lu et al., 2021).

To provide a clearer context for this concept, we consider temporal-omics datasets, particularly those used in the identification of potential enhancer–promoter pairs in various biological processes (Fulco et al., 2019; Moore et al., 2020; Schoenfelder and Fraser, 2019; Whalen et al., 2016). Enhancers are short *cis*-regulatory DNA sequences that are distal to the genes they regulate, as opposed to the promoter sequence that is adjacent to a particular gene. Since enhancers and promoters coordinate to generate time- and context-specific gene expression programs, that is, the amount of RNA produced at what time, we hypothesize that the dynamic patterns of enhancer regulation are typically followed by corresponding gene program patterns, albeit with a certain time lag. Approximately a million candidate human enhancers have been identified (Moore et al., 2020), though the intricate dynamics of how and which enhancers interact with promoters in specific processes are still an area of active and ongoing research.

Addressing this challenge, our study focuses on the multi-omics time series data presented by Reed et al. (2022). This dataset includes measurements of enhancer activity, promoter transcription, and 3D chromatin structure at 7 time points, from 0 to 6 hours (a later time point is excluded from this

Received: April 10, 2024; Revised: November 8, 2024; Accepted: December 27, 2024

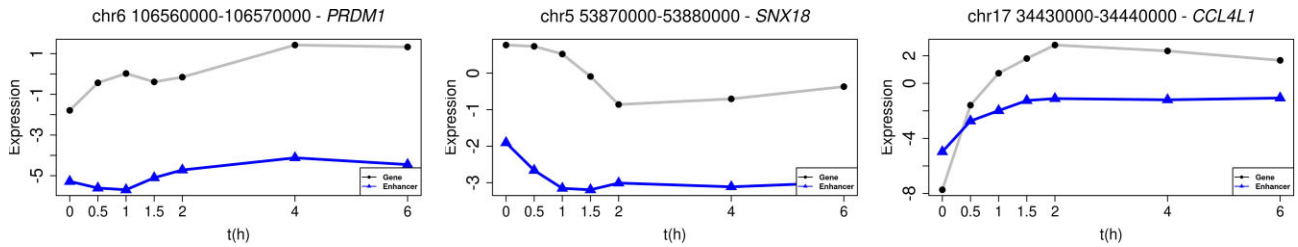

FIGURE 1 Examples of enhancer–promoter pairs with different dynamic patterns. The title for each subfigure represents the enhancer's genomic location and the gene name.

analysis). These data reveal complex dynamic patterns, as depicted in Figure 1. Our goal is to identify functional human enhancer–promoter pairs from over 20 000 candidate pairs, estimate the time lag, and quantify the extent of the lead–lag effect between enhancer activity and gene expression over time. The biological motivation for this goal is to uncover a relationship in genomics time series for a subset of enhancer–gene pairs, in particular those in which the enhancer state changes after cell activation and causes a subsequent change in expression of a target gene. Measuring the lag between enhancer activity and gene expression over time for these candidate enhancer–promoter pairs helps in understanding the sequence of regulatory events set off by activation, and how the time scale of gene regulation may vary across the genome.

In the context of our motivating application involving multi-omics time-series data, and similar scenarios across various fields, there are key challenges that need careful consideration. Firstly, a common issue is the scarcity of time-point observations. This limitation is compounded by the irregular intervals between successive measurements, often due to the high cost of experimental design or sensor failures. Additionally, in cases where there is an abundance of time-point data, the focus may shift to pinpointing local time delays within smaller segments of the time series. These factors underscore the importance of accurately quantifying the uncertainty associated with the extent of the lead–lag relationship, particularly when dealing with limited temporal datasets.

Existing methods for aligning time-lagged time series to explore the magnitude of lead–lag effects are numerous and varied. This exploration can extend to tasks such as ranking leaders or followers among multiple time series (Wu et al., 2010) and performing network clustering of time series based on the pairwise lead–lag relationships (Bennett et al., 2022). For example, Wu et al. (2010) rank the leaders of the streams to the sea surface temperature. Widely used models such as time-lagged cross-correlation (TLCC, Shen, 2015) by comparing the cross-correlation coefficients, dynamic time warping (DTW, Berndt and Clifford 1994), and its differentiable approximation, soft-DTW (Cuturi and Blondel, 2017), and soft-DTW divergence (Blondel et al., 2021) have proven useful by comparing distance loss among multiple pairwise time series after alignment (De Luca and Pizzolante, 2021). TLCC, for instance, operates by sliding one signal over another and identifying the shift position with the highest correlation. However, these methods often require the same time intervals for each time series, which makes them less effective for irregular time series, potentially leading to distortions of the true underlying relationships.

Beyond quantifying the lead–lag effect between pairwise time series, there is often a desire to estimate time lag considering the irregularities in time-series data; one approach is to first impute values at regular time points and then apply methods such as TLCC to calculate the time lag. Common techniques for this separate modeling of each time series include using splines (Rehfeld et al., 2011) or separate Gaussian processes (SGP, Li et al., 2021). However, such methods do not leverage shared information across time series, potentially leading to suboptimal results. Hierarchical GP models such as Magma (Leroy et al., 2022, 2023) integrate shared information across multiple time series but are primarily designed for prediction and clustering individual time series. This is in contrast to our primary objective, which is to cluster pairwise time series by a measure of similarity. More direct approaches for irregular time series include the use of the negative and positive lead–lag estimator (NAPLES, Ito and Nakagawa 2020) and the lead–lag method (Hoffmann et al., 2013), which utilize the Hayashi–Yoshida covariance estimator to estimate a fixed time lag. Alternatively, multitask GPs through latent variable models introduce additional variables to account for temporal discrepancies such as GPLVM (Dürichen et al., 2014), which may limit flexibility due to the presumed independence between tasks and temporal sequences (also known as separable kernels). Moreover, more complex misalignment techniques (Kazlauskaite et al., 2019; Mikheeva et al., 2022) might not perform well with limited time observations for biology data where simpler linear mismatch assumptions suffice, and they can face operational challenges such as Cholesky decomposition failures. Table 1 compares the aforementioned methods, highlighting their suitability for limited or irregular time points, their ability to quantify the magnitude of lead–lag effects, and estimate time lags:

In response to these challenges and limitations of existing methods, we propose the Gaussian process for lead–lag time-series (GPlag) model. GPlag introduces a novel class of GP kernels for time series with lead–lag effects, featuring interpretable parameters that estimate both the time lag and the degree of lead–lag effects. These parameters are crucial for tasks such as identifying associated factors, and ranking or clustering time series based on lead–lag relationships. Our model also provides theoretical justification for the identifiability and interpretability of these key parameters, and integrates a Bayesian framework for

TABLE 1 An overview of related existing methods.

Model	Quantify the magnitude of lead–lag effect		Estimate time lag	
	Allows for limited time points	Allows for irregular time points	Allows for limited time points	Allows for irregular time points
TLCC	X		X	
DTW	X		X	
Lead–lag			X	X
SGP	X*	X*	X*	X*
Splines	X*	X*	X*	X*
GPLVM		X*		X*
MAGMA		X*		X*
GPlag	X	X	X	X

X* means yes if followed by TLCC or DTW.

greater flexibility in handling time series with constraints such as limited observations or irregular time gaps.

2 METHODS

2.1 Gaussian process

Gaussian processes (GPs) are widely used for modeling dependency structures, such as spatial statistics and time-series modeling. In generic terms, a GP models a random function where any finite collection of realizations (ie, n observations) follows a multivariate normal (MVN) distribution.

Definition 1 *$f : \Omega \to \mathbb{R}$ follows a GP with mean function μ and covariance function K, denoted by $f \sim \mathrm{GP}(\mu, K)$, if for any $x_1, \cdots, x_n \in \Omega$, $[f(x_1), \cdots, f(x_n)] \sim N(v, \Sigma)$, where $v_i = \mu(x_i)$ and $\Sigma_{ij} = K(x_i, x_j)$.*

Covariance functions, also known as kernels, are widely studied when $\Omega \subset \mathbb{R}^p$, such as the radial basis function (RBF, also known as the squared exponential kernel), the exponential kernel, and the Matérn kernel (Stein, 1999).

2.2 Pairwise time-series GP model

For the sake of clarity, we initiate pairwise time-series models, and the extension to multiple time series is delayed to Section 2.5. Given a pair of time series $Y(t) := [Y_1(t), Y_2(t)]^\top$, where $t \in \mathbb{R}$, we model y as a multioutput GPs: $Y = F(t) + \varepsilon$. Here, $F = [F_1, F_2]^\top \sim \mathrm{GP}(\mu, \widetilde{K})$ and $\varepsilon \sim N(0, \tau^2 \mathrm{I}_2)$ represent the measurement error. The mean function is defined as μ, and $\widetilde{K} : \mathbb{R} \times \mathbb{R} \to \mathrm{PD}(2)$ is a cross-covariance kernel, with $\mathrm{PD}(2)$ representing the space of all 2 by 2 positive definite matrices. μ is often assumed to be zero, which can be achieved by a vertical shift (Gelfand et al., 2010), so we assume $\mu = 0$ in the remaining sections.

The cross-covariance kernel $\widetilde{K}$ is a crucial component of the model, expected to incorporate both the relationship between Y_1 and Y_2, and the time lag. Constructing valid cross-covariance kernels is a recognized challenge (Apanasovich and Genton, 2010; Genton and Kleiber, 2015; Gneiting et al., 2010), even more so when kernels must have certain identifiable and interpretable parameters in biological applications.

We address this by transforming the problem to its "dual" state, converting the multioutput GP into a single-output GP via $y = f(t, l) + \epsilon, l = 1, 2$, where $f \sim N(0, K)$ is a single-output GP, $\epsilon \sim N(0, \tau^2)$. We denote the 2 time series as $F_1(t) = f(t, 1)$ and $F_2(t) = f(t, 2)$. Here, K is a covariance kernel of a single-output GP, but the domain expands to $\mathbb{R} \times \{1, 2\}$ instead of $\mathbb{R}$. Essentially, we transform a multioutput GP over $\mathbb{R}$ into a single-output GP over $\mathbb{R} \times \mathscr{C}$, where $\mathscr{C} = \{1, 2\}$ is an index set of time series.

In this newly formed domain $\mathbb{R} \times \mathscr{C}$, a family of inseparable, semi-stationary kernel has been developed by Li et al. (2021), that is, $K((t, l), (t', l')) = K((t + h, l), (t' + h, l'))$ for any $h \in \mathbb{R}$. Consequently, K can be reformulated as a function on $\mathbb{R} \times \mathscr{C} \times \mathscr{C}$ as $K'(t, l, l') := K((t, l), (0, l'))$. However, for simplicity, we will denote both as K in the subsequent sections. A rich family of semi-stationary kernels can be found in Li et al. (2021). This family of kernels, although rich, was not specifically designed for time series with a time lag. To account for the time lag, the following theorem ensures that any existing semi-stationary kernel can induce a kernel with a parameter that measures the time lag, denoted by s, which ensures that our proposed kernel, with the inclusion of this time lag parameter, remains a valid kernel within the GP framework.

Theorem 1 *Let K' be a semi-stationary kernel on $\mathbb{R} \times \mathscr{C}$, then the induced kernel defined as*

$$K((t, l), (t', l')) := K'(t - t' - \mathbf{1}_{\{l \neq l'\}} s, l, l') \qquad (1)$$

is positive definite, and thus serves as a covariance kernel for a GP, where $s \in \mathbb{R}$ measures the time lag between 2 time series.

The following corollary provides some concrete candidates of kernels for GPlag.

Corollary 1 *If $\phi : \mathbb{R}_{\geq 0} \to \mathbb{R}$ is a completely monotone function and $\psi : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is a nonnegative function with a completely monotone derivative, then*

$$K((t, l), (t', l')) = \frac{\sigma^2}{\psi(\mathbf{1}_{\{l \neq l'\}})^{1/2}} \phi\left(-\frac{|t - t' - \mathbf{1}_{\{l \neq l'\}} s|^2}{\psi(\mathbf{1}_{\{l \neq l'\}})}\right)$$

is a valid kernel, where $\sigma^2 > 0$ measures the spatial variance and $s \in \mathbb{R}$ measures the time lag.

The proof is given in Section 1.2. Since the key idea in the above theorem and corollary is motivated by Gneiting (2002), we call the kernel family constructed from Corollary 1 the

Gneiting family. In particular, Gneiting enumerates 4 candidates for ϕ and 3 for ψ in Gneiting (2002). Exploring each amalgamation of ϕ and ψ families allows for further tailoring by adjusting parameters intrinsic to the family, culminating in the desired kernel. Furthermore, more kernels can be constructed by combining kernels from Gneiting family and other simple kernels. Analogues of the RBF, exponential kernel, and Matérn kernel in Euclidean space are provided in the following definitions.

Definition 2 *The following covariance kernels are called time lag radial basis function (LRBF), time lag exponential kernel (LExp), and time lag Matérn kernel (LMat), respectively:*

$$K((t,l),(t',l')) = \frac{\sigma^2}{(\mathbf{1}_{\{l\neq l'\}}a^2+1)^{1/2}} e^{-\frac{b\left(t-t'-\mathbf{1}_{\{l\neq l'\}}s\right)^2}{\mathbf{1}_{\{l\neq l'\}}a^2+1}}, \quad (2)$$

$$K((t,l),(t',l')) = \frac{\sigma^2}{\mathbf{1}_{\{l\neq l'\}}a^2+1} e^{-b|t-t'-\mathbf{1}_{\{l\neq l'\}}s|}, \quad (3)$$

$$K((t,l),(t',l')) = \frac{\sigma^2 2\left\{\frac{b}{2}|t-t'-\mathbf{1}_{\{l\neq l'\}}s|\right\}^{\nu}}{(\mathbf{1}_{\{l\neq l'\}}a^2+1)^{\nu+1/2}\Gamma(\nu)} \times K_\nu\left(b|t-t'-\mathbf{1}_{\{l\neq l'\}}s|\right). \quad (4)$$

As an illustration, LRBF can be derived by setting $\phi(t) = e^{-bt}$ and $\psi(t) = a^2t + 1$ in Corollary 1; LExp is the product of a kernel in Gneiting family with $\phi(t) = e^{-bt^{1/2}}$, $\psi(t) \equiv 1$, and a simple kernel $\frac{1}{\mathbf{1}_{\{l\neq l'\}}a^2+1}$; LMat is the product of a kernel in Gneiting family with $\phi(t) = \frac{(bt^{1/2})^\nu}{2^{\nu-1}\Gamma(\nu)}K_\nu(bt^{1/2})$, $\psi \equiv 1$, and a simple kernel $\frac{1}{(\mathbf{1}_{\{l\neq l'\}}a^2+1)^{\nu+1/2}}$. Common choices for ν are 1/2, 3/2, 5/2, where the Bessel function K_ν admits a closed form, similar to the standard Matérn kernel, with the specific choice depending on the data. Higher values of ν are suited for smoother processes. Additional candidate kernels are presented in the Web Appendix A Section 1.1, along with a detailed explanation on how they are constructed.

2.3 Interpretation of parameters

The interpretability of kernel parameters is one of the major advantages of GP models. In the kernels defined above, σ^2, the spatial variance, controls the point-wise variance; b, the length scale parameter, governs the temporal dependency; the smoothness parameter ν determines the smoothness of the sample path. In this paper, we assume ν to be given, which is a common assumption in GP literature (Rasmussen, 2004). More importantly, a is called the dissimilarity parameter, where a smaller a signifies a lesser degree of dissimilarity, which corresponds to a stronger lead–lag effect. For 2 distinct time series $l \neq l'$, when $a = \infty$, the covariance $K((t,l),(t',l')) = 0$, indicating independence between time series or no lead–lag effect at all. On the other hand, when $a = 0$, $K((t,1),(t',2)) = \sigma^2 e^{-b(t-t'-s)^2}$, implying that the 2 time series are the same up to a time lag, the strongest lead–lag effect. The interpretation of a in GPlag parallels measures lead–lag effect in other classical methods such as DTW and TLCC: $a = 0$ in GPlag aligns with a loss of 0 in DTW and a correlation of 1 in TLCC to signify the strongest lead–lag effect. Similarly, $a = \infty$ in GPlag corresponds to a loss of infinity in DTW and a correlation of 0 in TLCC for the weakest effect. The time-lag parameter s measures the time lag between 2 time series. In the next subsection, we discuss the implementation of GPlag and the inference of these parameters, especially a and s. Further visualization of sample paths of above kernels with different combinations of parameters is provided in Figures S5-S7. Theoretical support for the identifiability of a and s is provided in Section 3.

2.4 Implementation

In this section, we outline the algorithm to estimate the parameters $\{a, b, s, \sigma^2, \tau^2\}$ proposed in the previous model, given the observations $\{(t_i, c_i, y_i)\}_{i=1}^n$, $c_i \in \{1, 2\}$, $n = n_1 + n_2$. The most direct approach is to find the maximum likelihood estimate (MLE) given by:

$$\begin{aligned} l(a,b,s,\sigma^2,\tau^2) &:= \log p(y_1,\cdots,y_n|c_1,\cdots,c_n,t_1,\cdots,t_n,a,b,s,\sigma^2,\tau^2) \\ &= -\frac{n}{2}\log 2\pi - \frac{1}{2}\log|\Sigma+\tau^2\mathrm{I}_n| \\ &\quad - \frac{Y^\top(\Sigma+\tau^2\mathrm{I}_n)^{-1}Y}{2}, \end{aligned} \quad (5)$$

where $Y = [y_1, \cdots, y_n]^\top$ denotes the vector of observations, and Σ is the $(n_1 + n_2) \times (n_1 + n_2)$ covariance matrix specified in Equations (2)-(4). Optimization methods such as L-BFGS-B (Byrd et al., 1995) and Adam (Kingma and Ba, 2014) are widely used. The L-BFGS-B algorithm allows for setting lower and upper bounds on the parameters, enhancing the stability of the numeric optimization. The parameters a, b, σ^2, τ^2 are constrained to be positive and s can be set to a reasonable range depending on specific cases. For the initial values of b, σ^2, τ^2, we suggest running a standard GP regression first and using the estimates as our initial values. For s, we recommend taking the estimate from TLCC, and we recommend $a = 1$ as the initialization. The illustration of the MLE process is provided in Web Appendix A Algorithm 1.

The extension to Bayesian inference is straightforward: we can simply sample from the posterior $\pi(a,b,s,\sigma^2,\tau^2|\{t_i,c_i,y_i\}_{i=1}^n) \propto \pi(a,b,s,\sigma^2,\tau^2)\, p(Y|\{y_i,c_i,t_i\}_{i=1}^n, a,b,s,\sigma^2,\tau^2)$, where $\pi(a,b,s,\sigma^2,\tau^2)$ is the joint prior distribution. We suggest using independent inverse Gamma priors for a, b, σ^2, τ^2 and a Gaussian prior for s. The algorithm is depicted in Algorithm 3 in Supplement 1.4 and has been implemented in Python using GPyTorch (Gardner et al., 2018) and the Pyro (Bingham et al., 2019) package. This implementation enables rapid, fully Bayesian inference and takes advantage of GPU acceleration.

2.5 Extension to multi-time series

The kernels introduced in Section 2.2 can be extended to accommodate any number of time series, denoted by $L \geq 2$. Following the same rationale, it suffices to define a kernel on $\mathbb{R} \times \mathscr{C} \times \mathscr{C}$, where $\mathscr{C} = \{1, 2, \cdots, L\}$. The following theorem generalizes LRBF, LExp, and LMat for any L by aligning $A = (a_{ll'})$ with the criteria for a distance metric, parameterized by

$a_{ll'} \geq 0$, b, s_l, σ^2, τ^2 under the constraints $a_{ll'} = 0 \Longleftrightarrow l = l'$, $a_{ll'} = a_{l'l}$, $a_{ll'} + a_{l'k} \geq a_{lk}$ for any $l, l', k = 1, \cdots, L$ and $s_1 = 0$:

Theorem 2 *The following covariance functions are extensions of LRBF, LExp, and LMat:*

$$K((t,l),(t',l')) = \frac{\sigma^2}{(a_{ll'}^2+1)^{1/2}} e^{-\frac{b\left(t-t'+s_l-s_{l'}\right)^2}{a_{ll'}^2+1}},$$

$$K((t,l),(t',l')) = \frac{\sigma^2}{(a_{ll'}^2+1)} e^{-b\left|t-t'+s_l-s_{l'}\right|},$$

$$K((t,l),(t',l')) = \frac{\sigma^2 2\left\{\frac{b}{2}\left|t-t'+s_l-s_{l'}\right|\right\}^{\nu}}{(a_{ll'}^2+1)^{\nu+1/2}\Gamma(\nu)} \times K_{\nu}\left(b\left|t-t'+s_l-s_{l'}\right|\right).$$

The positive definite, symmetric, and triangle inequality constraints on $a_{ll'}$ ensure the positive definiteness of K. And the constraint on s_1 is due to our treating the first time series as the baseline, which means that the time lag is relative to the first time series. The inference is similar to the previous case, with the only difference being that the optimization method involved in MLE needs to consider these additional constraints. In other words, it is a constrained optimization problem. Nevertheless, it is not intractable, as all the constraints on $a_{ll'}$ are linear and can be accommodated by existing optimization methods, including L-BFGS-B and Adam, as illustrated in Web Appendix A Algorithm 2 with 3 time series.

3 IDENTIFIABILITY OF GPLAG KERNEL PARAMETER

To support our interpretation to $A := (a_{ll'}) \in \mathbb{R}^{L\times L}$ and $S := (s_l) \in \mathbb{R}^L$ in Section 2, we discuss the identifiability, that is, different parameters lead to different models. Note that not all parameters of commonly used GP models are identifiable, which means that different parameter values can lead to the same model. For example, in the classical Matérn covariance function, the spatial variance σ^2 and length scale parameter b are not identifiable, meaning that it is not possible to distinguish between different values of these parameters based on the data alone (Li et al., 2023; Stein, 1999; Zhang, 2004). This can be a limitation of GP models and it is important to consider when interpreting the results. In the case of GPlag, if the key parameters A and S are not identifiable, that is, the GPlag determined by (A, S) is equivalent to the one determined by $(\widetilde{A}, \widetilde{S})$, then it is unconvincing to interpret A and S as time series dissimilarity and time lag. Without identifiability, A and S do not admit any clear interpretation and the results of the model would be hard to interpret.

In this section, we prove that the parameters A and S in GPlag model are identifiable, allowing them to be used to as measures of the lead–lag effect between time series (measured by A) and the time lag (measured by S). Here, we assume all parameters are positive, finite, real numbers, and the domain is fixed (also known as the infill domain). The fixed domain assumption is reasonable for multiomics time series, as the dynamics of gene regulation in such experiments occur during a fixed window of time. Investigators will decide, based on available budget, the time points needed to capture the dynamics during, say, activation or differentiation. If allocated more budget, they will then fill in the gaps between existing observations, to have better resolution over the same time period. We begin by introducing the necessary definitions.

Definition 3 *Two GPs K and $\widetilde{K}$ are said to be equivalent, denoted by $K \equiv \widetilde{K}$, if the corresponding Gaussian random measures are equivalent to each other. That is, 2 Gaussian random measures are absolutely continuous with respect to each other.*

If $K \equiv \widetilde{K}$, then K can never be correctly distinguished from $\widetilde{K}$ regardless of how dense the observed time points are (Zhang, 2004). Focusing on a parametric family of covariance functions K_θ, if there exists $\theta_1 \neq \theta_2$ such that $K_{\theta_1} \equiv K_{\theta_2}$, then any estimator of θ based on n observations $\{t_i, y_i\}_{i=1}^n$, denoted by $\widehat{\theta}_n$, cannot be weakly consistent (Dudley, 1989).

Definition 4 *θ is said to be identifiable if $K_\theta \equiv K_{\widetilde{\theta}} \Longleftrightarrow \theta = \widetilde{\theta}$.*

As a result, it suffices to show the following necessary and sufficient conditions for 2 GPlags being equivalent, which suggests identifiable parameters for further interpretation:

Theorem 3 *Let K and $\widetilde{K}$ be 2 LMat kernels with parameters (σ^2, b, A, S) and $(\widetilde{\sigma}^2, \widetilde{b}, \widetilde{A}, \widetilde{S})$ and ν is assumed to be known, then*

$$K \equiv \widetilde{K} \Longleftrightarrow (\sigma^2 b^{2\nu}, A, S) = (\widetilde{\sigma}^2 \widetilde{b}^{2\nu}, \widetilde{A}, \widetilde{S}).$$

That is, the identifiable parameters, also known as the microergodic parameters, are $\{\sigma^2 b^{2\nu}, A, S\}$. In particular, let $\nu = \frac{1}{2}$, the microergodic parameters for LExp kernel are $\{\sigma^2 b, A, S\}$.

The proof is given in Appendix 1.3. Although σ^2 and b in GPlag are not identifiable, the proposed interpretation of the model does not rely on these 2 parameters. Instead, the parameter A is interpreted as a measure of dissimilarity between time series under a time lag represented by the parameter S. Note that GPs are commonly used for regression tasks. For this purpose, even though some kernel parameters such as σ^2 and b are not identifiable in these models, the prediction performance remains asymptotically optimal when the kernel parameters are misspecified (Kaufman and Shaby, 2013), suggesting that regression is distinct from parameter inference in GPs.

Moreover, the identifiability of kernel parameters in GPs depends on the dimension of the domain. For time series, the domain is 1-dimensional, but our theorem also holds for 2-dimensional and 3-dimensional domains. In dimensions greater than 3, it becomes a challenging problem and we treat it as a future work since our focus of this work is 1-dimensional time series.

4 SIMULATION STUDIES

The main focus of our experiments was to illustrate that GPlag could accurately estimate both the time lag and the extent of lead–lag effects. This capability aided in the ranking or

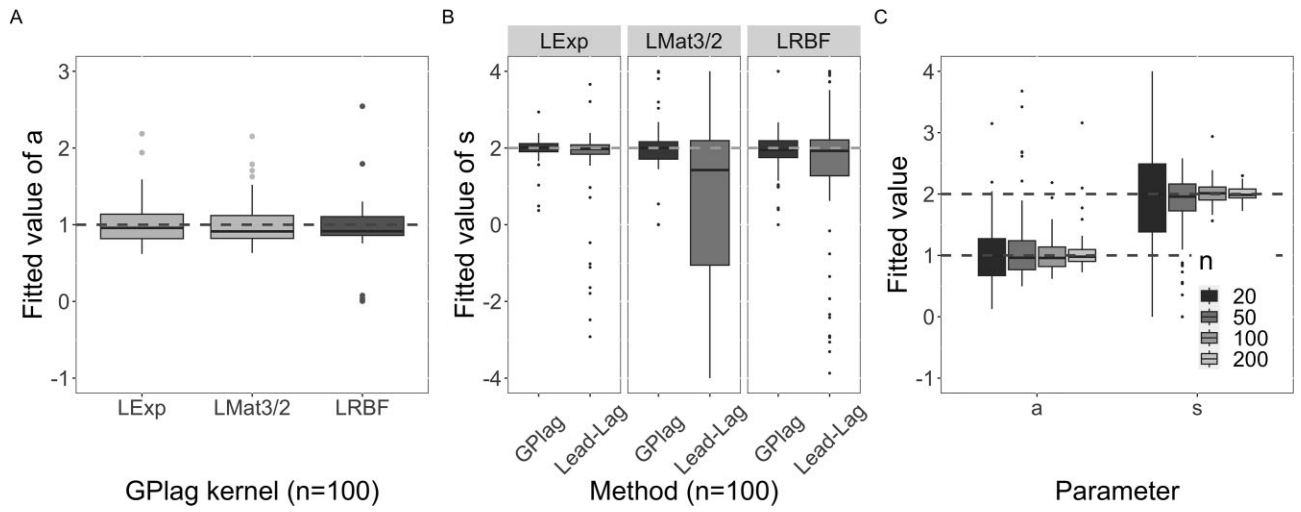

FIGURE 2 Parameter estimation. (A) MLE of a in GPlag with data ($n = 100$) generated from 3 different kernels. (B) s estimated by GPlag and lead–lag with data ($n = 100$) generated from 3 different kernels. (C) MLEs of GPlag parameters a and s in the LExp kernel with different n. The true parameter values were represented by the dashed horizontal lines. All boxplots were from 100 replicates.

clustering of multiple time series that existed in a lead–lag relationship with a target time series. We demonstrated the theoretical properties of GPlag using synthetic data, as these properties were best illustrated under a fully controlled environment. All code and additional implementation details were available in Supplement Section 2.1.

Baseline models (1) For estimating the time lag, we compared with lead–lag. (2) By utilizing the GP framework, we were able to predict values for 2 time series at unobserved time points. Then, we evaluated the prediction error against 3 baselines: natural cubic splines with 5 breakpoints, SGP, and Magma. (3) To evaluate the effectiveness of the estimate of lead–lag effects, we employed 4 baseline methods: TLCC, DTW, soft-DTW, and soft-DTW divergence. TLCC and DTW could output aligned time series, allowing us to calculate correlation coefficients from the aligned data. On the other hand, soft-DTW and soft-DTW divergence measured the amount of warping needed to align 2 time series and provided a distance loss metric. We used both correlation coefficients and distance loss as measures of dissimilarity between the time series.

Parameter estimation and inference. To validate the identifiability of our model, we simulated a pair of time series from GPlag with 3 kernels in Equations (2)-(4), with 100 replicates. We set the mean to zero for both series and specified parameters as $b = 0.3$, $a = 1$, $s = 2$, $\sigma^2 = 4$ for each kernel. For LMat, we set the smoothness parameter as $\nu = 3/2$. The number of measurements within each time series varied across {20, 50, 100, 200} and the time points were generated followings $t_i = \frac{50}{n-1}(i-1) + \epsilon$ and $t_i' = t_i + s$, where $i \in \{1, 2, \ldots, n\}$ and $\epsilon \sim \text{Unif}(-\frac{1}{4}, \frac{1}{4})$ represents noise.

Our results showed that the MLE of a and s were very close to their truth among all 3 kernels when $n = 100$. This was consistent with the theoretical analysis in Theorem 3 (Figure 2A and B) that GPlag could accurately recover the true lead–lag effects. Compared to lead–lag, GPlag exhibited similar or smaller variance, specifically when the kernel was Matérn with $\nu = 3/2$. Additionally, we performed simulations with a varying number of time points and found that the variance of the MLEs of a and s decreased as the sample size increased, which empirically confirmed the consistency of the MLEs (Figure 2C). Plots for other parameters of secondary interest, including σ^2, b, and $\sigma^2 b^{2\nu}$, are provided in Figure S1. We also performed simulations with different settings of $b = 0.3, 5, 10$, and found that the consistency of the MLEs for a and s was maintained across these different values of b (see Figure S2 for more details).

Prediction. We assessed the accuracy of GPlag's predictions by generating data from 2 different processes. In the first, we used the kernel LExp in Equation (3) with parameters $b = 1$, $a = 0.3$, $s = 2$, and $\sigma^2 = 4$. In the second, we generated pairwise time series from linear functions with non-Gaussian noise (t-distributed): $y = 2t + 3 + 5\epsilon_1$ and $y = 2(t - 20) + 3 + 5\epsilon_2$ where $t = 0, 1, 2, \ldots, 100$. The noise term ϵ_1 and ϵ_2 was drawn from a t-distribution with 5 degrees of freedom. We then randomly selected 50% of the data as training data for model fitting and used the remaining 50% as testing data for prediction in both settings. We applied the best linear unbiased prediction (BLUP; see the Supplement Section 2.1 for more details) and used the mean squared error (MSE) on the testing set as the metric to evaluate the prediction accuracy.

Among all methods, GPlag achieved the highest log-likelihood and smallest MSE (Figure 3A and B, Figure S3). The advantage of GPlag over SGP and splines was that it modeled 2 time series together, which allowed for sharing of information between time series through a. When $a \to \infty$, it was equivalent to SGP, while a small estimated a indicated that 2 time series were similar, and the covariance function then allowed for more information to be shared between them. Although Magma adopted a hierarchical GP to specifically model the mean function for irregular data, it did not consider time-lag information across time series.

Clustering or ranking time series based on dissimilarity to target time series.

We evaluated the relationship between the magnitude of parameter a and the degree of lead–lag effects (the level of synchrony with a designated target signal) to assess the effectiveness

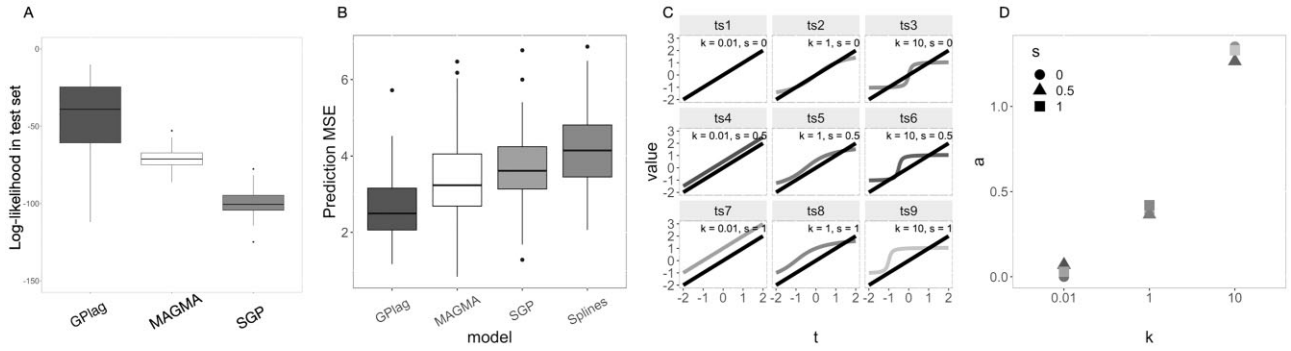

FIGURE 3 (A-B) Evaluation of prediction performance for time series derived from LExp Kernel, as assessed by log-likelihood (A) and MSE (B). (C) Line plots of 9 features time series versus target time series with various orders of distortion and time lag. The black line represents the target time series. (D) Estimation of a in 9 pairwise comparisons.

of GPlag. We simulated 9 time series from the function $f_k(t) = \frac{\arctan(k(t+s))}{\arctan(k)}$ for every pair of k and s, where $k \in \{0.01, 1, 10\}$ and $s \in \{0, 0.5, 1\}$. Among them, the time series simulated from the function of $k = 0.01$, $s = 0$ was treated as the target time series, which was close to a linear function (Figure 3C, ts1). Each time series had 50 time points ranging from $[-2, 2]$. Here, k determined the degree of distortion, where a smaller k (eg, 0.001) corresponded to a curve close to the identity transformation, and a larger k represented the cases with more upward/downward distortion. By comparing different columns in Figure 3C, we could see that the degree of dissimilarity between the target time series and other time series was controlled by k, whereas s controlled the time lag along the t-axis.

Nine pairs of time series were fitted by GPlag with the LExp kernel. We observed that 3 pairs of time series in each column with the same k were grouped together (Figure 3D), and a increased as k increased. This implied that GPlag could accurately rank the degree of lead–lag effects among multiple pairs of time series based on the parameter a, through precise estimation of the time-lag parameter s. A larger estimate of a indicated a stronger dissimilarity between pairs of time series. To further quantify the clustering performance of lead–lag effects among the 9 pairs of time series, we used the Adjusted Rand Index (ARI, Hubert and Arabie 1985) and Normalized Mutual Information (NMI, Strehl and Ghosh 2002) as evaluation metrics. To define a cluster from the pairwise dissimilarity parameter (correlation coefficients for TLCC and distance loss for DTWs), we used K-means (Shannon, 1948) with 3 clusters. As the focus of this work was on determining pairwise dissimilarity rather than the clustering algorithm, we opted for a simple clustering method. GPlag achieved the highest ARI and NMI in clustering the lead–lag effects among 9 pairwise time series (Table 1, columns 2 and 3). We also tested GPlag using the LRBF and LMat kernels, and found that ARI and NMI scores remained at 1, consistent with the LExp results. The inferred a values are shown in Figure S4, further indicating that the results are not sensitive to the choice of kernel. The results for the competitors are based on their default parameter settings provided in the package.

Extension to 3 time series. To demonstrate the applicability of GPlag to more than 2 time series, we conducted simulation studies using 3 time series generated from the LExp kernel. In this scenario, all constraints were linear, allowing us to utilize optimizers that support linear constraints. Similar to the pairwise simulation setting, we set the mean to zero for all series, and the time points were generated according to the equations $t_i = \frac{50}{49}(i-1) + \epsilon$ and $t_i' = t_i + s_i$, where $i \in \{1, 2, \ldots, 50\}$ and $\epsilon \sim \text{Unif}(-\frac{1}{4}, \frac{1}{4})$ represented noise. The specified parameter values were $b = 0.3$, $a_{12} = a_{13} = a_{23} = 1$, $s_2 = 2$, $s_3 = 4$, $\sigma^2 = 4$. The maximum likelihood estimate (MLE) of the parameters closely approximated the true values (represented by the dashed line), providing further validation of the theoretical properties. Additionally, the proposed method achieved the lowest MSE in this particular setup (see Figure 4).

5 GENOMICS APPLICATION: DYNAMIC CHROMATIN INTERACTIONS

In this section, we applied GPlag to a time-resolved multiomics dataset involving enhancers and genes in human cells activated by interferon gamma from a resting state (Reed et al., 2022). Originally, the dataset comprised over 20 000 pairs. Each pair consisted of 1 time series for each enhancer and likewise 1 time series for each gene, at 7 irregularly spaced time points (0, 0.5, 1, 1.5, 2, 4, and 6 hours after a perturbation). The time series captured gene expression measurements for genes and histone tail acetylation measurements (H3K27ac), a measure of activity for the enhancers, both approximately variance stabilized. It is widely believed that the alterations in acetylation occur prior to the changes in expression of the regulated genes (Reed et al., 2022). We focused only on pairs that became functional after stimulation, by applying a filter on the total range of stabilized values across the timescale, yielding 4776 dynamic pairs of enhancer–gene time series. Subsequently, the objective of our analysis, after estimating the model parameters of GPlag for these 4776 pairs, was to identify the subset of all possible enhancer–gene pairs that represent *functional* enhancer–gene pairs, that is, the enhancer responds to cell activation and regulates the expression of the gene, after some time lag. Many types of time profiles can be captured by such a time-series experiment, for example, increasing across the range, decreasing, or increasing then decreasing, etc. Particular functional

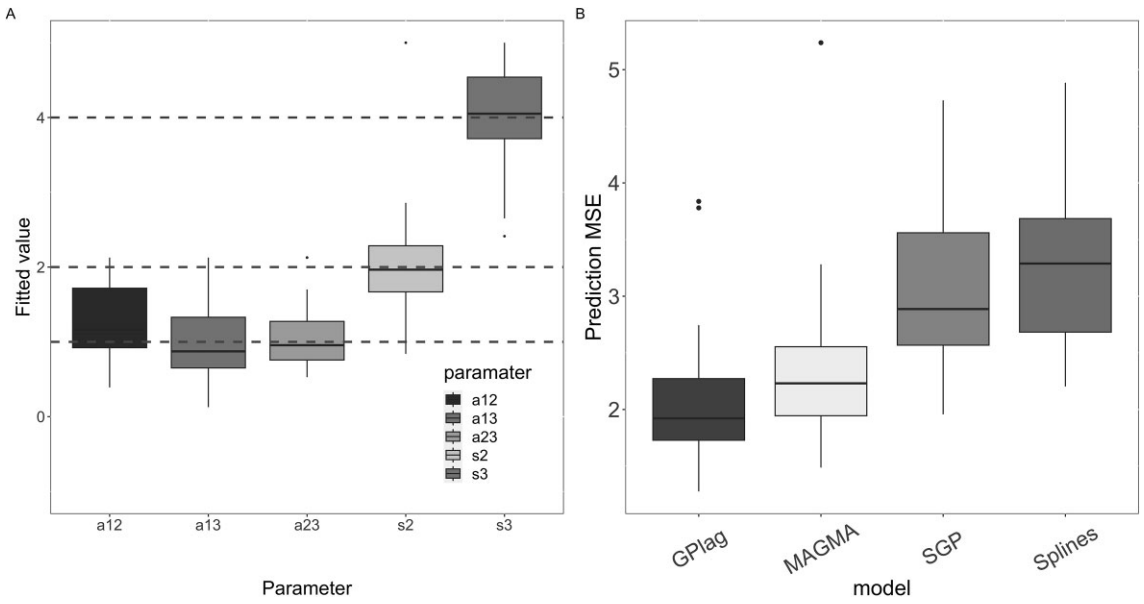

FIGURE 4 Fitting performance on 3 time series generated from LExp with $b = 0.3$, $a_{12} = a_{13} = a_{23} = 1$, $s_2 = 2$, $s_3 = 4$. (A) MLE of a_{12}, a_{13}, a_{23}, s_1, s_2. The true parameter values are represented by the dashed horizontal lines. (B) Prediction MSE on the held-out dataset, from 100 replicates.

TABLE 2 Evaluation of GPlag versus TLCC, DTW, soft-DTW, and soft-DTW divergence on various datasets.

	Synthetic data clustering		Dynamic Chromatin Interactions	
Model	**ARI**	**NMI**	***P*-value: association test**	***P*-value: enrichment test**
TLCC	0.35	0.58	0.051	0.13
DTW	0.4	0.55	0.292	1
Soft-DTW	0.27	0.49	0.957	0.57
Soft-DTW divergence	0.07	0.39	0.907	0.90
GPlag	**1.00**	**1.00**	**0.019**	**0.05**

Since the above algorithms are deterministic, the standard deviation is not reported.

enhancer–gene pairs will have subtle differences in the dynamics of the coordinated change in activity and expression, compared to other functional pairs, which we hope to capture with our GPlag model.

To assess the performance of GPlag in ranking enhancer–gene connections, we split the pairs into 2 groups according to the parameter *a*. The top 20% pairs were designated as the candidate enhancer–gene group, with the remaining pairs forming the negative group. To compare against baselines, we split all pairs into 2 groups using their respective metric of similarity and a 20% cutoff. Subsequently, we assessed the performance from 3 perspectives.

Firstly, we compared the Hi-C (High-throughput Chromosome Conformation Capture, Lieberman-Aiden et al. 2009) counts between the candidate enhancer–gene group and the negative group. Since Hi-C counts measured the frequency of physical interactions between enhancers and genes in 3-dimensional chromatin structure, we expected the candidate enhancer–gene group to have higher Hi-C counts, while the Hi-C information was not used in our ranking of pairs. Our results showed candidate enhancer–gene pairs exhibited significantly higher Hi-C counts compared to the other pairs (Table 2, columns 4 and 5), with GPlag producing the smallest *p*-value (Wilcoxon rank-sum test, $p = 0.019$).

Secondly, we analyzed the enrichment of *looped* enhancer–gene pairs between the 2 groups. A *looped* enhancer–gene pair is 1, where the enhancer element is physically brought into close proximity with the regulated gene, enabling the enhancer's activity to influence the expression of the associated gene. Looping in this dataset was determined using Hi-C counts, but the determination of loops uses spatial pattern detection and statistical models to determine if the signal rises to the level of a "loop." We observed that candidate enhancer–gene pairs identified by GPlag had a significantly higher likelihood of being "looped" enhancer–gene pairs compared to the matched pairs from the other group, selected based on genomic distance (Davis et al., 2023) (χ^2 test, $p = 0.05$). However, no significant difference in enrichment was found between the 2 groups of pairs identified using baseline methods.

The lack of statistical significance observed in DTW-based methods may be due to the limited information available for alignment because of the small number of observations. Additionally, with irregularly spaced time points, TLCC might introduce bias when estimating time lag and coefficients. Given

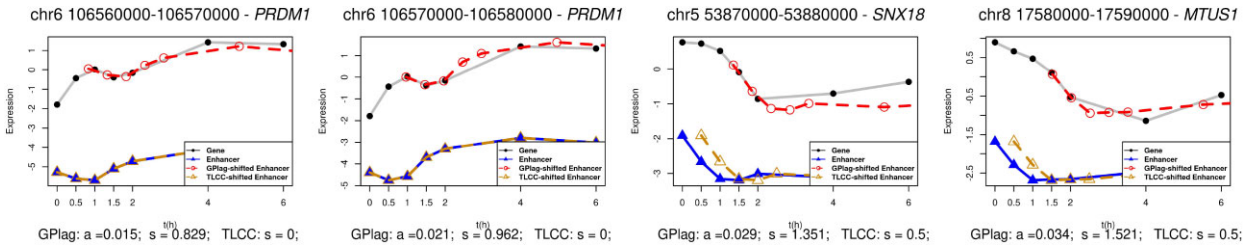

FIGURE 5 Visualization of GPlag modeling output on selected pairs. The title for each subfigure represents the enhancer's genomic location and the gene name. The solid circle line represents gene expression; the solid triangle line represents enhancer activity; the dashed circle line represents the enhancer shifted by GPlag; and the dashed triangle line represents the enhancers shifted by TLCC.

the constraint of TLCC, which can only generate time lag estimates from the observed time points, its performance was compromised when applied to this dataset with irregular time gaps (Figure 5). However, GPlag, by estimating a continuous time lag, was able to identify lags that restored strong similarity between the pairwise time series.

Thirdly, to evaluate whether our candidate enhancer–promoter pairs were detected as having a functional relationship in an orthogonal experimental data type, we downloaded the genomic locations of response expression quantitative trait loci (response eQTL) found in human macrophages exposed to Interferon - gamma (IFN $_\gamma$, Alasoo et al., 2018) [ie, using one of the 2 stimuli used in the Reed et al. (2022) experiment, and in the same cell type]. Because the number of eQTLs was small compared to the number of enhancers, we used a relatively stringent threshold to label a putative functional enhancer–promoter pair: we split the candidate enhancer–promoter pairs into groups according to posterior mean correlation (> 0.98). There were 74 H3K27ac peaks overlapping with macrophage response eQTL, among which 12 pairs were in our defined candidate enhancer–promoter pair sets. Therefore, there was a significant enrichment of our candidate enhancer–promoter pairs to eQTLs (χ^2 test $p = 0.18 \times 10^{-3}$), further suggesting the relevance of the enhancer–promoter pairs highly ranked by GPlag.

6 DISCUSSION

GPlag is a flexible and interpretable model specifically designed for irregular time series with limited time points that exhibit lead–lag effects, a scenario often encountered in biology. We believe that our method's adaptability renders it valuable for time-series applications in various areas such as longitudinal gene expression, epigenetics, proteomics, cytometry, and microbiome studies. It can be applied to a variety of problems, including but not limited to ranking or clustering time series based on the degree of lead–lag effects and interpolation or forecasting. The core feature of our methodology is the transform of the multioutput GP into a single-output GP, which is grounded in theoretical support, affirming its validity as a covariance kernel for GP. This transform leverages the shared information among time series and enables simultaneous modeling of the time-lag parameter s and dissimilarity parameters a, thereby achieving desirable interpretability and expanding the application of the model. Additionally, our model can be extended to accommodate multiple time series, beyond just pairwise time series, enhancing its applicability and versatility.

The work here suggests some exciting future directions. First, the extension to a nonstationary kernel allows for more flexibility when modeling nonstationary time series with more heterogeneity. The 2 main challenges are to find a valid nonstationary kernel with interpretable parameters and perform efficient and effective statistical inference. Second, for time-series data with large n, that is, dense time points, the vanilla GP is computationally expensive with $O(n^3)$ complexity (Rasmussen, 2004). In this situation, we can apply scalable GP methods such as nearest-neighbor GP (NNGP, Datta et al. 2016) or variational inference GP (Tran et al., 2015) to reduce the complexity to $O(n \log n)$ for large datasets. In addition, incorporating additional parameters, such as varying the variance σ^2 and length scale b across different time series, could potentially enhance model performance. Finally, the question of whether the parameters in Theorem 3 can be consistently estimated remains open. Even for simpler kernels such as the Matérn, consistent estimators for identifiable parameters (eg, $\sigma^2 b^{2\nu}$ and ν) were only recently developed (Loh and Sun, 2023; Loh et al., 2021). For more complex kernels, like the ones proposed in this paper, constructing consistent estimators is an open and challenging problem.

ACKNOWLEDGMENTS

W.M. and J.C. contributed equally to this paper.

SUPPLEMENTARY MATERIALS

Supplementary material is available at *Biometrics* online.

Web Appendices A, referenced in Section 2, Web Appendices B, referenced in Section 4, Section 5, and the code are available with this paper at the Biometrics website on Oxford Academic and at https://github.com/Wancen/GPlag.

FUNDING

W.M. and M.I.L. were supported by National Institutes of Health R01HG009937, D.H.P. was supported by National Institutes of Health R35-GM128645, and D.L. was supported by National Institutes of Health R01 AG079291, R56 LM013784,

R01 HL149683, UM1 TR004406, R01 LM014407, and P30 ES010126.

CONFLICT OF INTEREST

None declared.

DATA AVAILABILITY

The data that support the findings in this paper are available in Gene Expression Omnibus (GEO), at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE201376

REFERENCES

Alasoo, K., Rodrigues, J., Mukhopadhyay, S., Knights, A. J., Mann, A. L., Kundu, K. et al. (2018). Shared genetic effects on chromatin and gene expression indicate a role for enhancer priming in immune response. *Nature Genetics*, 50, 424–431.

Apanasovich, T. V. and Genton, M. G. (2010). Cross-covariance functions for multivariate random fields based on latent dimensions. *Biometrika*, 97, 15–30.

Bacry, E., Delattre, S., Hoffmann, M. and Muzy, J.-F. (2013). Some limit theorems for Hawkes processes and application to financial statistics. *Stochastic Processes and their Applications*, 123, 2475–2499.

Bennett, S., Cucuringu, M. and Reinert, G. (2022). Lead–lag detection and network clustering for multivariate time series with an application to the US equity market. *Machine Learning*, 111,1–42.

Berndt, D. J. and Clifford, J. (1994). Using dynamic time warping to find patterns in time seriesIn: *Proceedings of the 3rd International Conference on Knowledge Discovery and Data Mining*, 10, 359–370. Washington DC, USA: AAAI Press.

Bingham, E., Chen, J. P., Jankowiak, M., Obermeyer, F., Pradhan, N., Karaletsos, T. et al. (2019). Pyro: deep universal probabilistic programming. *The Journal of Machine Learning Research*, 20, 28:1–28:6.

Blondel, M., Mensch, A. and Vert, J.-P. (2021). Differentiable divergences between time series. In: *Proceedings of the 24th International Conference on Artificial Intelligence and Statistics*. 3853–3861. PMLR.

Byrd, R. H., Lu, P., Nocedal, J. and Zhu, C. (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, 16, 1190–1208.

Cuturi, M. and Blondel, M. (2017). Soft-DTW: a differentiable loss function for time-series. In: *Proceedings of the 34th International Conference on Machine Learning*, 894–903. PMLR.

Da Fonseca, J. and Zaatour, R. (2017). Correlation and lead–lag relationships in a Hawkes microstructure model. *Journal of Futures Markets*, 37, 260–285.

Datta, A., Banerjee, S., Finley, A. O. and Gelfand, A. E. (2016). Hierarchical nearest-neighbor Gaussian process models for large geostatistical datasets. *Journal of the American Statistical Association*, 111, 800–812.

Davis, E. S., Mu, W., Lee, S., Dozmorov, M. G., Love, M. I. and Phanstiel, D. H. (2023). matchRanges: generating null hypothesis genomic ranges via covariate-matched sampling. *Bioinformatics*, 39, btad197.

De Luca, G. and Pizzolante, F. (2021). Detecting leaders country from road transport emission time-series. *Environments*, 8,18.

Ding, J. and Bar-Joseph, Z. (2020). Analysis of time-series regulatory networks. *Current Opinion in Systems Biology*, 21, 16–24.

Dudley, R. M. (1989). *Real Analysis and Probability*, Boca Raton, FL, USA: Chapman and Hall/CRC.

Dürichen, R., Pimentel, M. A., Clifton, L., Schweikard, A. and Clifton, D. A. (2014). Multitask Gaussian processes for multivariate physiological time-series analysis. *IEEE Transactions on Biomedical Engineering*, 62, 314–322.

Faryabi, B., Vahedi, G., Chamberland, J.-F., Datta, A. and Dougherty, E. R. (2008). Optimal constrained stationary intervention in gene regulatory networks. *EURASIP Journal on Bioinformatics and Systems Biology*, 2008, 1–10.

Feng, Z., Sivak, J. A. and Krishnamurthy, A. K. (2021). Two-stream attention spatio-temporal network for classification of echocardiography videos. In: *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, 1461–1465. Piscataway, NJ, USA: IEEE.

Fulco, C. P., Nasser, J., Jones, T. R., Munson, G., Bergman, D. T., Subramanian, V. et al. (2019). Activity-by-contact model of enhancer–promoter regulation from thousands of CRISPR perturbations. *Nature genetics*, 51, 1664–1669.

Gardner, J., Pleiss, G., Weinberger, K. Q., Bindel, D. and Wilson, A. G. (2018). GPyTorch: blackbox matrix–matrix Gaussian process inference with GPU acceleration. In: *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, 31, PMLR.

Gelfand, A. E., Diggle, P., Guttorp, P. and Fuentes, M. (2010). *Handbook of Spatial Statistics*. Boca Raton, FL, USA: CRC Press.

Genton, M. G. and Kleiber, W. (2015). Cross-covariance functions for multivariate geostatistics. *Statistical Science*, 30, 147–163.

Gerstein, M. B., Lu, Z. J., Van Nostrand, E. L., Cheng, C., Arshinoff, B. I., Liu, T. et al. (2010). Integrative analysis of the *Caenorhabditis elegans* genome by the modENCODE project. *Science*, 330, 1775–1787.

Gneiting, T. (2002). Nonseparable, stationary covariance functions for space–time data. *Journal of the American Statistical Association*, 97, 590–600.

Gneiting, T., Kleiber, W. and Schlather, M. (2010). Matérn cross-covariance functions for multivariate random fields. *Journal of the American Statistical Association*, 105, 1167–1177.

Harzallah, A. and Sadourny, R. (1997). Observed lead–lag relationships between Indian summer monsoon and some meteorological variables. *Climate Dynamics*, 13, 635–648.

Hayashi, T. and Koike, Y. (2018). Wavelet-based methods for high-frequency lead–lag analysis. *SIAM Journal on Financial Mathematics*, 9, 1208–1248.

He, Q. and Huang, H.-H. (2024). A framework of zero-inflated Bayesian negative binomial regression models for spatiotemporal data. *Journal of Statistical Planning and Inference*, 229, 106098.

Hoffmann, M., Rosenbaum, M. and Yoshida, N. (2013). Estimation of the lead–lag parameter from non-synchronous data. *Bernoulli*, 19, 426–461.

Huang, H.-H. and He, Q. (2023). Statistical modeling of *Peromyscus maniculatus* (deer mouse) amounts per trap with spatiotemporal data. *Japanese Journal of Statistics and Data Science*, 6, 847–860.

Hubert, L. and Arabie, P. (1985). Comparing partitions. *Journal of Classification*, 2, 193–218.

Ito, K. and Nakagawa, K. (2020). NAPLES; Mining the lead-lag Relationship from Non-synchronous and High-frequency Data. arXiv, preprint arXiv:2002.00724, preprint: not peer reviewed.

Ito, K. and Sakemoto, R. (2020). Direct estimation of lead–lag relationships using multinomial dynamic time warping. *Asia-Pacific Financial Markets*, 27, 325–342.

Kaufman, C. and Shaby, B. A. (2013). The role of the range parameter for estimation and prediction in geostatistics. *Biometrika*, 100, 473–484.

Kazlauskaite, I., Ek, C. H. and Campbell, N. (2019). Gaussian process latent variable alignment learning. In: *Proceedings of the 22nd International Conference on Artificial Intelligence and Statistics*. 748–757. PMLR.

Kingma, D. P. and Ba, J. (2014). Adam: a method for stochastic optimization. In *Proceedings of the International Conference on Learning Representations*, arXiv, preprint arXiv:1412.6980.

Krijger, P. H. L. and De Laat, W. (2016). Regulation of disease-associated gene expression in the 3D genome. *Nature Reviews Molecular Cell Biology*, 17, 771–782.

Leroy, A., Latouche, P., Guedj, B. and Gey, S. (2022). MAGMA: inference and prediction using multi-task Gaussian processes with common mean. *Machine Learning*, 111, 1821–1849.

Leroy, A., Latouche, P., Guedj, B. and Gey, S. (2023). Cluster-Specific Predictions with Multi-Task Gaussian Processes. *Journal of Machine Learning Research*, 24, 1–49.

Li, D., Jones, A., Banerjee, S. and Engelhardt, B. E. (2021). Multi-group Gaussian processes. arXiv, preprint arXiv:2110.08411, preprint: not peer reviewed.

Li, D., Tang, W. and Banerjee, S. (2023). Inference for Gaussian Processes with Matérn Covariogram on Compact Riemannian Manifolds. *Journal of Machine Learning Research*, 24, 1–26.

Lieberman-Aiden, E., van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A. et al. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*, 326, 289–293.

Loh, W.-L. and Sun, S. (2023). Estimating the parameters of some common Gaussian random fields with nugget under fixed-domain asymptotics. *Bernoulli*, 29, 2519–2543.

Loh, W.-L., Sun, S. and Wen, J. (2021). On fixed-domain asymptotics, parameter estimation and isotropic Gaussian random fields with Matérn covariance functions. *The Annals of Statistics*, 49, 3127–3152.

Lu, J., Dumitrascu, B., McDowell, I. C., Jo, B., Barrera, A., Hong, L. K. et al. (2021). Causal network inference from gene transcriptional time-series response to glucocorticoids. *PLoS Computational Biology*, 17, e1008223.

Mikheeva, O., Kazlauskaite, I., Hartshorne, A., Kjellström, H., Ek, C. H. and Campbell, N. (2022). Aligned multi-task Gaussian process. In: *Proceedings of the 25th International Conference on Artificial Intelligence and Statistics*, 2970–2988. PMLR.

Moore, J. E., Purcaro, M. J., Pratt, H. E., Epstein, C. B., Shoresh, N., Adrian, J. et al. (2020). Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature*, 583, 699–710.

Rasmussen, C. E. (2004). *Gaussian Processes in Machine Learning*. Berlin, Germany: Springer.

Reed, K. S., Davis, E. S., Bond, M. L., Cabrera, A., Thulson, E., Quiroga, I. Y. et al. (2022). Temporal analysis suggests a reciprocal relationship between 3D chromatin structure and transcription. *Cell Reports*, 41, 111567.

Rehfeld, K., Marwan, N., Heitzig, J. and Kurths, J. (2011). Comparison of correlation analysis techniques for irregularly sampled time series. *Nonlinear Processes in Geophysics*, 18, 389–404.

Runge, J., Nowack, P., Kretschmer, M., Flaxman, S. and Sejdinovic, D. (2019). Detecting and quantifying causal associations in large nonlinear time series datasets. *Science Advances*, 5, eaau4996.

Schoenfelder, S. and Fraser, P. (2019). Long-range enhancer–promoter contacts in gene expression control. *Nature Reviews Genetics*, 20, 437–455.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27, 379–423.

Shen, C. (2015). Analysis of detrended time-lagged cross-correlation between two nonstationary time series. *Physics Letters A*, 379, 680–687.

Skoura, A. (2019). Detection of lead–lag relationships using both time domain and time-frequency domain; an application to wealth-to-income ratio. *Economies*, 7, 28.

Stein, M. L. (1999). *Interpolation of Spatial Data: Some Theory for Kriging*. Berlin, Germany: Springer Science and Business Media.

Strehl, A. and Ghosh, J. (2002). Normalized mutual information clustering. *Journal of Machine Learning Research*, 3, 583–617.

Strober, B., Elorbany, R., Rhodes, K., Krishnan, N., Tayeb, K., Battle, A. et al. (2019). Dynamic genetic regulation of gene expression during cellular differentiation. *Science*, 364, 1287–1290.

Tran, D., Ranganath, R. and Blei, D. M. (2015). The variational Gaussian process. In *Proceedings of the Internation Conference on Learning Representations*, arXiv:1511.06499.

Wang, D., Tu, J., Chang, X. and Li, S. (2017). The lead–lag relationship between the spot and futures markets in China. *Quantitative Finance*, 17, 1447–1456.

Whalen, S., Truty, R. M. and Pollard, K. S. (2016). Enhancer–promoter interactions are encoded by complex genomic signatures on looping chromatin. *Nature Genetics*, 48, 488–496.

Wu, D., Ke, Y., Yu, J. X., Yu, P. S. and Chen, L. (2010). Detecting leaders from correlated time series. In: *Database Systems for Advanced Applications*, (eds. H., Kitagawa, Y., Ishikawa, Q., LI and Watanabe, C.), 352–367. Berlin, Germany: Springer.

Zhang, H. (2004). Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. *Journal of the American Statistical Association*, 99, 250–261.

Zhu, J. and Gallego, B. (2021). Evolution of disease transmission during the COVID-19 pandemic: patterns and determinants. *Scientific Reports*, 11, 1–9.

Received: April 10, 2024; Revised: November 8, 2024; Accepted: December 27, 2024