



Heriot-Watt University
Research Gateway

Multimodal Data Fusion of Spatial Fields in Sensor Networks

Citation for published version:

Zhang, P, Peters, GW, Nevat, I, Teo, KB & Wang, Y 2020, Multimodal Data Fusion of Spatial Fields in Sensor Networks. in *2019 IEEE SENSORS.*, 8956540, IEEE.
<https://doi.org/10.1109/SENSORS43011.2019.8956540>

Digital Object Identifier (DOI):

[10.1109/SENSORS43011.2019.8956540](https://doi.org/10.1109/SENSORS43011.2019.8956540)

Link:

[Link to publication record in Heriot-Watt Research Portal](#)

Document Version:

Peer reviewed version

Published In:

2019 IEEE SENSORS

Publisher Rights Statement:

© 2020 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

General rights

Copyright for the publications made accessible via Heriot-Watt Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

Heriot-Watt University has made every reasonable effort to ensure that the content in Heriot-Watt Research Portal complies with UK legislation. If you believe that the public display of this file breaches copyright please contact open.access@hw.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Multimodal Data Fusion of Non-Gaussian Spatial Fields in Sensor Networks

Pengfei Zhang¹, Gareth W. Peters² and Ido Nevat³

1. Institute for Infocomm Research (I2R), Singapore

2. Department of Actuarial Mathematics and Statistics, Heriot-Watt University, Edinburgh, UK.

3. TUMCREATE, Singapore

Abstract—We develop a robust data fusion algorithm for field reconstruction of multiple physical phenomena. The contribution of this paper is twofold: First, we demonstrate how multi-spatial fields which can have any marginal distributions and exhibit complex dependence structures can be constructed. To this end we develop a model where a latent process of these physical phenomena is modelled as Multiple Gaussian Process (MGP), and the dependence structure between these phenomena is captured through a Copula process. This model has the advantage of allowing one to choose any marginal distributions for the physical phenomenon. Second, we develop an efficient and robust linear estimation algorithm to predict the mean behaviour of the physical phenomena using *rank correlation* instead of the conventional *linear Pearson correlation*. Our approach has the advantage of avoiding the need to derive intractable predictive posterior distribution and also has a tractable solution for the rank correlation values. We show that our model outperforms the model which uses the conventional linear Pearson correlation metric in terms of the prediction mean-squared-errors (MSE). This provides the motivation for using our models for multimodal data fusion.

Keywords: Sensor Networks, Copula, Multiple Output Gaussian Process, Rank Correlation

I. INTRODUCTION

The term “Internet-of-Things” (IoT) describes several technologies and research disciplines in which the Internet extends into the physical world [1], [2]. IoT networks consist of sensors that can collect different types of data modalities from the environment. For example, sensors can measure temperature, humidity or pollution particles from environment at same time. Therefore, it has been increasingly important problem to study multimodal sensor networks where different modalities exhibit different statistical distributions. In addition, the correlation between different data can also be taken into account in order to make more accurate inference. However, both of these two tasks are difficult and challenging.

Many works have been developed to understand the dependence of multimodal data in sensor networks. Classical methods include the linear dependence structure between different fields, resulting

in linear correlated output, namely multiple output Gaussian Process (GP) [3], [4]. However, these classical methods suffer from two main drawbacks that make it infeasible to solve real world challenging problems:

- 1) The marginal distribution of GP is Gaussian. However, in many practical cases the Normality assumption is violated. For example, wind field is typically modelled as Weibull distribution [5], and a Poisson distribution is widely used to model discrete counts of data, e.g., the number of pollution particles in the field [6].
- 2) GP models only capture the linear Pearson correlation dependence, and do not allow for more complex dependence structures. However, more complicated nonlinear dependence structures might exist in real physical contexts. For example, the extreme pressure in spatial regions and extreme rainfalls cannot be captured using linear dependence structures [7].

It is therefore necessary to develop new models to incorporate both the non-Gaussian marginals as well as non-linear dependence structure of multimodal fields. Developing such a model is the main focus of this paper. A general framework of modeling dependencies is to use Copula functions [8]. Copula models have become popular because of their ability to separate the marginal distribution from the dependence structure of multivariate distributions. This allows nonlinear dependence structures to be captured and modelled. We develop a hierarchical model where the MGP is used as a latent process while the marginal distribution can be any process, and the dependence between these processes is captured via Copula. We study bivariate processes in this paper, however, it can be easily extended to multiple processes.

II. BACKGROUND FEATURES OF THE MODEL FORMULATION

In this section we present important definitions of some key components used in the model construction, namely related to non-parametric Gaussian

Processes, linear dependent Gaussian Processes and parametric Copula models. We also discuss some properties of rank correlations related to Copula processes. These definitions are essential for our wireless sensor network models and our problems and solutions as well.

Definition 1 (Gaussian Process [4]): A Gaussian process is a collection of random variables, any finite number of which have a joint Gaussian Distribution.

A Gaussian process is completely specified by its mean function, $\mu(x)$, and covariance function $k(x, x')$ and denoted by

$$f(x) \sim \mathcal{GP}(\mu(x), k(x, x')).$$

Definition 2 (Linearly Dependent Gaussian Processes [9]): Given two Gaussian processes f_1 and f_2 , if the correlation between $f_1(x), \forall x$ in the domain of f_1 and $f_2(x'), \forall x'$ in the domain of f_2 is :

$$k(f_1(x), f_2(x')) = \mathbb{E}[(f_1(x) - \mu_1(x))(f_2(x') - \mu_2(x'))],$$

then the two random processes are said to be linearly dependent. The dependency structure of the two dependent Gaussian processes is captured via a kernel matrix K :

$$K = \begin{bmatrix} K_1 & K_{12} \\ K_{21} & K_2 \end{bmatrix},$$

where K_1 and K_2 are the correlation matrices within process 1 and process 2 respectively; K_{12} and K_{21} are the correlation matrices which capture the cross dependency between process 1 and process 2.

Furthermore, it will be useful to define a Copula distribution for a multivariate random vector as it provides a means to study dependence structures which are scale-free measures of dependence or concordance, see discussions in [10]. In general the term Copula is a Latin noun that means ‘‘a link, tie, bond’’ which in the context in which we consider it in this work, is used to link marginal distributions to form a joint dependent distribution model.

Definition 3 (Copula Distribution): A function $C : [0, 1] \times \dots \times [0, 1] \mapsto [0, 1]$ is a Copula if it satisfies:

- C is grounded;
- for ever $i \in \{1, \dots, n\}$ and any $u_i \in [0, 1]$ one has

$$C(1, \dots, 1, u_i, 1, \dots, 1) = u_i$$

i.e. the marginals are uniform.

- C is n -increasing, such that for all $(x_1, \dots, x_n), (y_1, \dots, y_n) \in [0, 1]^n$ with $x_i \leq y_i$ one has

$$\sum_{i_1=1}^2 \dots \sum_{i_n=1}^2 (-1)^{i_1+\dots+i_n} C(u_{1i_1}, \dots, u_{ni_n}) \geq 0$$

where $u_{j1} = x_j$ and $u_{j2} = y_j$ for all $j \in \{1, \dots, n\}$.

We note that in a bivariate context for instance the notion of groundedness is defined as follows.

Definition 4 (Grounded Function): Consider S_1 and S_2 as non-empty subsets of $[-\infty, \infty]$. Suppose that S_i has at least element a_i , for $i \in \{1, 2\}$. Then a function $G : S_1 \times S_2 \mapsto \mathbb{R}$ is grounded if

$$G(x, a_2) = 0 = G(a_1, y), \quad \forall (x, y) \in S_1 \times S_2$$

Furthermore, one can also state well known related results as follows for combinations of strictly increasing and decreasing functions, see [8, 2nd Edition, Theorem 2.4.4].

Proposition 1 (Influence of Increasing and Decreasing Transformations of the Marginals): Consider two continuous random variables X_1 and X_2 with joint Copula given by C_{X_1, X_2} . If $T_1(\cdot)$ and $T_2(\cdot)$ are two strictly monotone functions defined on $\text{Ran}X_1$ and $\text{Ran}X_2$, respectively. Then $C_{T_1(X_1), T_2(X_2)}$ is characterized by one of the following combinations:

- If T_1 is strictly increasing and T_2 is strictly decreasing, then

$$C_{T_1(X_1), T_2(X_2)}(u_1, u_2) = u_1 - C_{X_1, X_2}(u_1, 1 - u_2)$$

- If T_1 is strictly decreasing and T_2 is strictly increasing, then

$$C_{T_1(X_1), T_2(X_2)}(u_1, u_2) = u_2 - C_{X_1, X_2}(1 - u_1, u_2)$$

- If T_1 and T_2 are strictly decreasing, then

$$C_{T_1(X_1), T_2(X_2)}(u_1, u_2) = u_1 + u_2 - 1 + C_{X_1, X_2}(1 - u_1, 1 - u_2)$$

Remark 1: It was shown in [10] that all the axioms that a concordance measure (measure of dependence) should satisfy, as outlined by [11], are also uniquely characterized by a Copula formulation. This means that all known measures of dependence such as familiar correlations, associations, tail dependence and beyond can be captured uniquely by the Copula function.

Under Copula, rank correlations have the following properties, such as Spearman correlation.

Proposition 2 (Spearman’s Rho Rank Correlation Under Monotonic Marginal Transforms): Consider two continuous random variables X_1 and X_2 with joint copula given by C_{X_1, X_2} with copula density $c(u_1, u_2)$ when it exists. If $T_1(\cdot)$ and $T_2(\cdot)$ are two strictly monotone functions defined on $\text{Ran}X_1$ and $\text{Ran}X_2$, respectively. Then the Spearman’s rho rank correlation between X_1 and X_2 , denoted by ρ_{X_1, X_2}^S , is given after transformation by:

- If T_1 and T_2 are strictly increasing, then

$$\rho_{T_1(X_1), T_2(X_2)}^S = \rho_{X_1, X_2}^S.$$

- If T_1 is strictly increasing and T_2 is strictly decreasing, then

$$\rho_{T_1(X_1), T_2(X_2)}^S = 3 - 12 \int_0^1 \int_0^1 C(u_1, 1-u_2) du_1 du_2$$

- If T_1 is strictly decreasing and T_2 is strictly increasing, then

$$\rho_{T_1(X_1), T_2(X_2)}^S = 3 - 12 \int_0^1 \int_0^1 C(1-u_1, u_2) du_1 du_2$$

- If T_1 and T_2 are strictly decreasing, then

$$\rho_{T_1(X_1), T_2(X_2)}^S = 12 \int_0^1 \int_0^1 C(1-u_1, 1-u_2) du_1 du_2 - 3$$

Proof: The proof of each result follows directly from the application of the identity for Spearman's rho linear correlation written in terms of a copula as denoted in [12],

$$\begin{aligned} \rho_{X_1, X_2}^S &= 12 \int_0^1 \int_0^1 C(u_1, u_2) du_1 du_2 - 3 \\ &= 3 - 6 \int_0^1 \int_0^1 \left[u_1 \frac{\partial C}{\partial u_1}(u_1, u_2) + u_2 \frac{\partial C}{\partial u_2}(u_1, u_2) \right] \\ &\quad \times du_1 du_2 \\ &= 12 \mathbb{E}[U_1 U_2] - 3 \\ &= \frac{\mathbb{E}[U_1 U_2] - \mathbb{E}[U_1] \mathbb{E}[U_2]}{\sqrt{\text{Var}(U_1) \text{Var}(U_2)}} \end{aligned} \quad (1)$$

and then application of Proposition 1 to obtain for each case enumerated. ■

As with Kendall's tau rank correlation, for many Copula families the explicit solution for the Copula based expression for the Spearman rank correlation is known explicitly in terms of the Copula parameters.

Furthermore, it will be often useful to link the rank correlation such as Spearman's rho to the notion of linear correlation that we will denote generically as ρ . In general, where the joint dependence structure of the multivariate distribution is specified in terms of a correlation matrix, such as elliptical families where ρ is a model parameter. Then one obtains ρ^S and ρ as given by the identity:

$$\rho^S(X_1, X_2) = \rho(F_1(X_1), F_2(X_2)).$$

In certain cases there is also a direct relationship known between rank and linear correlations such as in the multivariate Gaussian Copula case in which the Spearman correlation ρ_S is obtained in terms of ρ linear correlation according to the expression

$$\rho = 2 \sin\left(\frac{\pi}{6} \rho^S\right). \quad (2)$$

Based on these definitions we can now present our hierarchical model for multimodal spatial fields.

III. HIERARCHICAL BAYESIAN MODEL FOR MULTIPLE MODALITY SPATIAL RANDOM FIELDS

The sensor network is deployed in \mathbb{R}^2 to monitor various physical phenomena. Based on the observations collected by the sensors, we wish to make predictions about the physical quantities at any location in space, denoted $\mathbf{x}_* \in \mathbb{R}^2$. To make the exposition simple we only consider two physical phenomena, but our model can be generalised to any number of modalities.

- 1) The two physical phenomena of interest, denoted $Z^{(1)}(\mathbf{x}_i)$ and $Z^{(2)}(\mathbf{x}_j)$, are correlated via two latent dependent GPs, $f^{(1)}(\mathbf{x}_i)$ and $f^{(2)}(\mathbf{x}_j)$, at any point $(\mathbf{x}_i, \mathbf{x}_j) \in \mathbb{R}^2$ through a Copula process which we will specify later. The two latent GPs $f^{(1)}(\mathbf{x}_i), f^{(2)}(\mathbf{x}_j)$ are coupled as per Definition 2:

$$\begin{aligned} &\left(f^{(1)}(\mathbf{x}_i), f^{(2)}(\mathbf{x}_j) \right) : \mathbb{R}^2 \times \mathbb{R}^2 \mapsto \mathbb{R} \times \mathbb{R} \text{ s.t.} \\ &\left(f^{(1)}(\mathbf{x}_i), f^{(2)}(\mathbf{x}_j) \right) \\ &\sim \mathcal{GP} \left(\begin{bmatrix} \mu^{(1)}(\mathbf{x}_i) \\ \mu^{(2)}(\mathbf{x}_j) \end{bmatrix}, K(\mathbf{x}_i, \mathbf{x}_j; \Psi) \right), \end{aligned}$$

where $\mu^{(1)}(\mathbf{x}_i), \mu^{(2)}(\mathbf{x}_j) \in \mathbb{R}$ are the mean functions of each of the two GPs. The spatial dependence between any two points is given by the covariance function $K(\mathbf{x}_i, \mathbf{x}_j; \Psi) : \mathbb{R}^2 \times \mathbb{R}^2 \mapsto \mathbb{R}$, parameterised by Ψ [4] and,

$$K = \begin{bmatrix} K^{(1)} & K^{(1,2)} \\ K^{(2,1)} & K^{(2)} \end{bmatrix}.$$

- 2) The two physical phenomena $Z^{(1)}(\mathbf{x}_i)$ and $Z^{(2)}(\mathbf{x}_j)$ are associated with $f^{(1)}(\mathbf{x}_i)$ and $f^{(2)}(\mathbf{x}_j)$ through the following Gaussian Copula processes:

$f^{(1)}(\mathbf{x}_i)$ and $f^{(2)}(\mathbf{x}_j)$ are mapped to $[0, 1]$ through univariate normal CDFs. Denote the resulting data as $U^{(1)}(\mathbf{x}_i)$ and $U^{(2)}(\mathbf{x}_j)$. Then we take inverse CDFs at $U^{(1)}(\mathbf{x}_i)$ and $U^{(2)}(\mathbf{x}_j)$ and denote the resulting data as $Z^{(1)}(\mathbf{x}_i)$ and $Z^{(2)}(\mathbf{x}_j)$.

To summarize, the model for the data generated has the following two-step process:

Step 1 :

$$\left[U^{(1)}(\mathbf{x}_i), U^{(2)}(\mathbf{x}_j) \right] := \left[F_1\left(f^{(1)}(\mathbf{x}_i)\right), F_2\left(f^{(2)}(\mathbf{x}_j)\right) \right].$$

Step 2 :

$$\begin{aligned} &\left[Z^{(1)}(\mathbf{x}_i), Z^{(2)}(\mathbf{x}_j) \right] \\ &:= \left[H_1^{-1}\left(U^{(1)}(\mathbf{x}_i)\right), H_2^{-1}\left(U^{(2)}(\mathbf{x}_j)\right) \right]. \end{aligned}$$

F_1 and F_2 are marginal CDFs of f_i and f_j , and $H_1^{-1}(u_i)$ and $H_2^{-1}(u_j)$ represent some inverse CDFs which may be different, $u_i, u_j \in [0, 1]$.

Denote the joint CDF of $Z^{(1)}(\mathbf{x}_i), Z^{(2)}(\mathbf{x}_j)$ as H_{12} and the joint CDF of $f^{(1)}(\mathbf{x}_i), f^{(2)}(\mathbf{x}_j)$ as F_{12} .

The above is a Copula process and by Sklar's Theorem,

$$C^{GA}(u_i, u_j) = F_{12}(F_1^{-1}(u_i), F_2^{-1}(u_j)).$$

$$C^{GA}(u_i, u_j) = H_{12}(H_1^{-1}(u_i), H_2^{-1}(u_j)).$$

$C^{GA}(u_i, u_j)$ refers to Gaussian Copula.

- 3) **Sensors observations:** there are n_1 sensors measuring the first physical phenomenon and n_2 sensors measuring the second physical phenomenon over a 2-D region $\mathcal{X} \subseteq \mathbb{R}^2$, at locations $\mathbf{x}_i \in \mathcal{X}, i = \{1, \dots, n_1\}$ and $\mathbf{x}_j \in \mathcal{X}, j = \{1, \dots, n_2\}$, assumed known. Each sensor collects a noisy observation of the respective physical process:

$$Y^{(1)}(\mathbf{x}_i) = Z^{(1)}(\mathbf{x}_i) + W,$$

$$Y^{(2)}(\mathbf{x}_j) = Z^{(2)}(\mathbf{x}_j) + V,$$

where W and V are i.i.d Gaussian noises: $W \sim N(0, \sigma_w^2)$, $V \sim N(0, \sigma_v^2)$.

- 4) We denote by \mathbf{Y} the observation vector of the two physical phenomena, as follows:

$$\mathbf{Y} = \left[\underbrace{Y_1^{(1)}, Y_2^{(1)}, \dots, Y_{n_1}^{(1)}}_{\text{phenomenon 1}}, \underbrace{Y_1^{(2)}, Y_2^{(2)}, \dots, Y_{n_2}^{(2)}}_{\text{phenomenon 2}} \right]^\top.$$

IV. ESTIMATION OBJECTIVES

The goal is to derive a low complexity algorithm to perform multimodal spatial field reconstitution, given noisy observations of the two physical phenomena \mathbf{Y} . the objective is to make predictions for the intensities $f_*^{(1)} := f^{(1)}(\mathbf{x}_*)$ and $f_*^{(2)} := f^{(2)}(\mathbf{x}_*)$ of the phenomena at any location \mathbf{x}_* in the field. To obtain this, we define the following estimation objective: The Minimum Mean Squared Error (MMSE) estimator of the joint predicted values of intensities at any location \mathbf{x}_* :

$$\hat{\mathbf{f}}_* = \mathbb{E}[\mathbf{f}_* | \mathbf{Y}, \mathbf{x}, \mathbf{x}_*, \Theta] = \int_{-\infty}^{\infty} \mathbf{f}_* p(\mathbf{f}_* | \mathbf{Y}, \mathbf{x}, \mathbf{x}_*, \Theta) d\mathbf{f}_*,$$

We define the following shorthand notations:

$\mathbf{x}_* := (\mathbf{x}_*^{(1)}, \mathbf{x}_*^{(2)})$ - test locations.

$\mathbf{f}_* := (f_*^{(1)}, f_*^{(2)})$ - predictions of the intensities at \mathbf{x}_* .

$\mathbf{Z}_* := (Z_*^{(1)}, Z_*^{(2)})$, predictions of the two phenomena at \mathbf{x}_* .

$\mathbf{x} := (\mathbf{x}^{(1)}, \mathbf{x}^{(2)})$ - sensor locations.

$\mathbf{f} := (\mathbf{f}_{1:n_1}^{(1)}, \mathbf{f}_{1:n_2}^{(2)})$

- realizations of the Gaussian Processes at \mathbf{x} .

To derive the above estimation objectives, the joint predictive density $p(\mathbf{f}_* | \mathbf{Y}, \mathbf{x}, \mathbf{x}_*, \Theta)$ needs to be evaluated first.

A. Predictive posterior density of the spatial intensities

The predictive posterior density is given by

$$p(\mathbf{f}_* | \mathbf{Y}, \mathbf{x}, \mathbf{x}_*, \Theta) = \int p(\mathbf{f}_* | \mathbf{f}, \mathbf{x}, \mathbf{x}_*, \Theta) p(\mathbf{f} | \mathbf{Y}, \mathbf{x}, \mathbf{x}_*, \Theta) d\mathbf{f}$$

$$= \int p(\mathbf{f}_* | \mathbf{f}, \mathbf{x}, \mathbf{x}_*, \Theta) \frac{p(\mathbf{Y} | \mathbf{f}, \mathbf{x}, \mathbf{x}_*, \Theta) p(\mathbf{f} | \mathbf{x}, \mathbf{x}_*, \Theta)}{\int p(\mathbf{Y} | \mathbf{f}, \mathbf{x}, \mathbf{x}_*, \Theta) p(\mathbf{f} | \mathbf{x}, \mathbf{x}_*, \Theta) d\mathbf{f}} d\mathbf{f}.$$

Unfortunately, the predictive posterior density cannot be evaluated analytically as this involves a $(n_1 + n_2)$ -dimensional integral that is intractable. Instead, in the following we develop the Spatial Best Linear Unbiased Estimator (S-BLUE), .

B. Spatial Best Linear Unbiased Estimator (S-BLUE) Field Reconstruction Algorithm

We develop the spatial field reconstruction via S-BLUE, which enjoys a low computational complexity and is the optimal estimator (in terms of minimizing the MSE) out of all linear estimators. The big advantage of the S-BLUE is that it does not require calculating the predictive posterior density, but only the first two cross moments of the model. The S-BLUE is the optimal (in terms of minimizing Mean Squared Error (MSE)) of all linear estimators and is given by the solution to the following optimization problem:

$$\hat{\mathbf{f}}_* := \hat{a} + \hat{\mathbf{B}}\mathbf{Y}_{1:N} = \arg \min_{a, \mathbf{B}} \mathbb{E} \left[(f_* - (a + \mathbf{B}\mathbf{Y}_{1:N}))^2 \right], \quad (3)$$

where $\hat{a} \in \mathbb{R}$ and $\hat{\mathbf{B}} \in \mathbb{R}^{1 \times N}$.

The optimal linear estimator that solves (3) is given by

$$\hat{\mathbf{f}}_* = \mathbb{E}_{f_* \mathbf{Y}_{1:N}} [f_* \mathbf{Y}_{1:N}] \mathbb{E}_{\mathbf{Y}_{1:N}} [\mathbf{Y}_{1:N} \mathbf{Y}_{1:N}^T]^{-1} (\mathbf{Y}_{1:N} - \mathbb{E}[\mathbf{Y}_{1:N}]), \quad (4)$$

and the Mean Squared Error (MSE) is given by

$$\sigma_*^2 = k(\mathbf{x}_*, \mathbf{x}_*) - \mathbb{E}_{f_* \mathbf{Y}_{1:N}} [f_* \mathbf{Y}_{1:N}] \mathbb{E}_{\mathbf{Y}_{1:N}} [\mathbf{Y}_{1:N} \mathbf{Y}_{1:N}^T]^{-1} \times \mathbb{E}_{\mathbf{Y}_{1:N}} [f_* \mathbf{Y}_{1:N}]. \quad (5)$$

To evaluate (4-5) we need to calculate the cross-correlation $\mathbb{E}_{f_* \mathbf{Y}_{1:N}} [f_* \mathbf{Y}_{1:N}]$, auto-correlation $\mathbb{E}_{\mathbf{Y}_{1:N}} [\mathbf{Y}_{1:N} \mathbf{Y}_{1:N}^T]$ and $\mathbb{E}[\mathbf{Y}_{1:N}]$.

Note, here without loss of generality we calculate the correlation for zero-mean Gaussian process ($\mu_f(x_*) = 0$). For the case where $\mu_f(x_*) \neq 0$, it is easy to shift the estimation by $\mu_f(x_*)$.

C. Copula Fitting

We adopt the approach in [13] to fit Gaussian Copula on \mathbf{Y} . The approach is below:

- 1) Estimate the rank correlation, either $\rho_S(Y_i, Y_j)$ or $\rho_K(Y_i, Y_j)$, for each marginal pair of variables. Then transform to the linear correlation measure;

- 2) Construct the estimated sample pseudo correlation matrix \hat{R}^* with (i, j) -th element given by Eq. (2).
- 3) The pseudo correlation matrix \hat{R}^* must be made positive definite with unit diagonal entries and off-diagonal entries in the range $[-1, 1]$.

After we fit the Copula, we could estimate the length scale l for square exponential kernel $k(\cdot, \cdot)$ which minimize $\sum(R_{i,j}^* - \exp(-(X_i, X_j)/l^2))$.

D. Cross-correlation and auto-correlation derivations

In this section, we derive the cross correlation and auto correlations of the terms required in (4) and (5).

1) *Cross-correlation between a test point and sensors observations* $\mathbb{E}_{\mathbf{f}_*, \mathbf{Y}_{1:N}} [\mathbf{f}_* \mathbf{Y}_{1:N}]$: it has been shown that Kendall (ρ_K) or Spearman (ρ_S) correlation are robust approximation of population correlation ρ . For the bivariate normal distribution, there is analytic relationship between these variables as shown in Section II. In [14], it was shown that ρ_K and ρ_S are invariant to impulse noise.

Proposition 3 (Cross Correlation 1): *The cross correlation between a test point and sensors observations* $\mathbb{E}_{\mathbf{f}_*, \mathbf{Y}_{1:N}} [\mathbf{f}_* \mathbf{Y}_{1:N}] = \rho_{\mathbf{f}_*, \mathbf{f}}^S * \sigma_{\mathbf{f}_*} * \sigma_{\mathbf{Y}_{1:N}}$

Proof: According to the expectation definition, $\mathbb{E}_{\mathbf{f}_*, \mathbf{Y}_{1:N}} [\mathbf{f}_* \mathbf{Y}_{1:N}] = \mathbb{E}_{\mathbf{f}_*, \mathbf{Y}_{1:N}} [\mathbf{f}_* H(F^{-1}(\mathbf{f}))]$, this quantity is intractable. Another way of expressing the cross correlation is $\mathbb{E}_{\mathbf{f}_*, \mathbf{Y}_{1:N}} [\mathbf{f}_* \mathbf{Y}_{1:N}] = \rho_{\mathbf{f}_*, \mathbf{Y}_{1:N}} * \sigma_{\mathbf{f}_*} * \sigma_{\mathbf{Y}_{1:N}}$ where $\rho_{\mathbf{f}_*, \mathbf{Y}_{1:N}}$ is the population correlation between a test point \mathbf{f}_* and observation $\mathbf{Y}_{1:N}$. It is also difficult to get the population correlation. We use the Spearman rank correlation $\rho_{\mathbf{f}_*, \mathbf{Y}_{1:N}}^S$ to approximate this quantity. Also according to [14], ρ_K and ρ_S and robust approximation to ρ that are invariant to impulse noise. Through this way, the properties of Spearman correlation can be used.

According to Proposition 2, If T_1 and T_2 are strictly increasing, then

$$\rho_{T_1(X_1), T_2(X_2)}^S = \rho_{X_1, X_2}^S. \quad (6)$$

By definition, $H(F^{-1}(\mathbf{f}))$ is strictly increasing function on \mathbf{f} , therefore $\rho_{\mathbf{f}_*, \mathbf{Y}_{1:N}}^S = \rho_{\mathbf{f}_*, H(F^{-1}(\mathbf{f}))}^S = \rho_{\mathbf{f}_*, \mathbf{f}}^S$. Therefore, $\mathbb{E}_{\mathbf{f}_*, \mathbf{Y}_{1:N}} [\mathbf{f}_* \mathbf{Y}_{1:N}] = \rho_{\mathbf{f}_*, \mathbf{f}}^S * \sigma_{\mathbf{f}_*} * \sigma_{\mathbf{Y}_{1:N}}$. ■

2) *Correlation of sensors observations* $\mathbb{E}_{\mathbf{Y}_{1:N}} [\mathbf{Y}_{1:N} \mathbf{Y}_{1:N}^T]$:

Proposition 4 (Cross Correlation 2): *The cross correlation between sensors observations* $\mathbb{E}_{\mathbf{Y}_{1:N}} [\mathbf{Y}_{1:N} \mathbf{Y}_{1:N}^T] = \rho_{\mathbf{f}, \mathbf{f}}^S * \sigma_{\mathbf{Y}_{1:N}} * \sigma_{\mathbf{Y}_{1:N}}$

Proof: Similarly as Proposition 3,

$$\begin{aligned} \mathbb{E}_{\mathbf{Y}_{1:N}} [\mathbf{Y}_{1:N} \mathbf{Y}_{1:N}^T] &= \rho_{\mathbf{Y}_{1:N}}^S * \sigma_{\mathbf{Y}_{1:N}} * \sigma_{\mathbf{Y}_{1:N}} \\ &= \rho_{F^{-1}(H(\mathbf{f})), F^{-1}(H(\mathbf{f}))}^S * \sigma_{\mathbf{Y}_{1:N}} * \sigma_{\mathbf{Y}_{1:N}} \\ &= \rho_{\mathbf{f}, \mathbf{f}}^S * \sigma_{\mathbf{Y}_{1:N}} * \sigma_{\mathbf{Y}_{1:N}}. \end{aligned}$$

■

3) *Expected value of the observations* $\mathbb{E}_{\mathbf{Y}_{1:N}} [\mathbf{Y}_{1:N}]$: $\mathbb{E}_{\mathbf{Y}_{1:N}} [\mathbf{Y}_{1:N}]$ is based on the distribution of marginals. For example, if $\mathbf{Y}_{1:N}$ has exponential marginal, then $\mathbb{E}_{\mathbf{Y}_{1:N}} [\mathbf{Y}_{1:N}] = 1/\lambda$. If $\mathbf{Y}_{1:N}$ has gamma marginal, then $\mathbb{E}_{\mathbf{Y}_{1:N}} [\mathbf{Y}_{1:N}] = \alpha/\beta$, where α and β are the shape and rate parameters of Gamma distribution.

4) *MMSE estimate of predicted intensity values:* The MMSE estimate of the predictions at any location \mathbf{x}_* is given by

$$\begin{aligned} \sigma_*^2 &= k(\mathbf{x}_*, \mathbf{x}_*) - \mathbb{E}_{\mathbf{f}_*, \mathbf{Y}_{1:N}} [f_* \mathbf{Y}_{1:N}] \mathbb{E}_{\mathbf{Y}_{1:N}} [\mathbf{Y}_{1:N} \mathbf{Y}_{1:N}]^{-1} \\ &\quad \times \mathbb{E}_{\mathbf{Y}_{1:N}} f_* [\mathbf{Y}_{1:N} f_*]. \end{aligned} \quad (7)$$

where all the quantities have been derived in the above subsections.

V. SIMULATION RESULTS

In this section, we present simulation results to compare the performance between robust BLUE (R-BLUE) performance and linear BLUE (L-BLUE) performance. L-BLUE is developed by approximating $\rho_{\mathbf{f}_*, \mathbf{Y}_{1:N}}$ in Proposition 3 and $\rho_{\mathbf{Y}_{1:N}, \mathbf{Y}_{1:N}}$ in Proposition 3 with $\rho_{\mathbf{f}_*, \mathbf{f}_{1:N}}^P$ and $\rho_{\mathbf{f}_{1:N}, \mathbf{f}_{1:N}}^P$ respectively. ρ^P denotes the linear Pearson-Norman correlation. We using MSE as the performance metrics. The comparison between R-BLUE and L-BLUE for single GP is studied first, followed by comparison between L-BLUE and R-BLUE for bivariate GP fields. Lastly, we also summarize the MSE comparison for many different realisations.

A. Linear-BLUE and Robust-BLUE Comparison

In this section, we compare the MSE performance of the Linear BLUE and the Robust BLUE in terms of single GP field reconstruction accuracy. We run 1000 realisations and the MSE for robust BLUE is 1.8617 and linear BLUE is 2.0323.

We also run the comparison between Linear BLUE and Robust BLUE for bivariate Gamma process setting. We generate bivariate GP as shown in Fig. 1, then we transform them into bivariate Gamma processes as shown in Fig. 2. We then reconstruct the GP from the gamma process realisations for two processes as shown in Figs. 3 and 4. After 1000 iterations, the MSE for GP1 and GP2 using robust BLUE are 1.1711 and 1.1963 respectively. The MSE for GP1 and GP2 using linear BLUE are 1.2029 and 1.2035 respectively.

We also test the robustness when one of the points is corrupted by impulsive noise. In this case, we purposely distorted a single observation by adding 30 to its real value. We then ran 1000 iterations, and the MSE for GP1 and GP2 using robust BLUE are 1.1946 and 6.2631 respectively. The MSE for GP1 and GP2 using linear BLUE are 1.2256 and 6.8120 respectively.

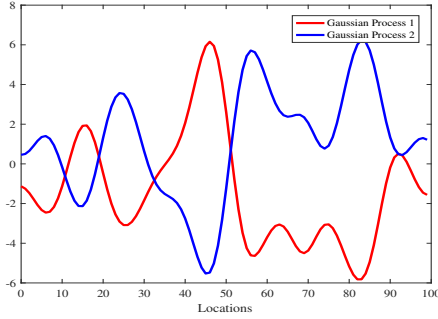


Fig. 1: Gaussian process realisations for both modality

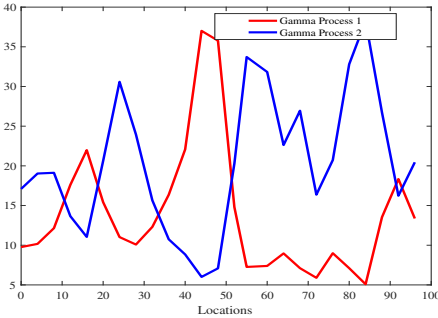


Fig. 2: Gamma process realisations for both modality

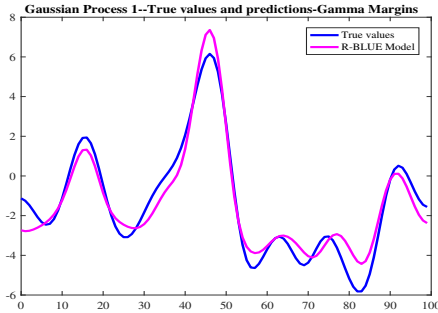


Fig. 3: Gaussian process 1 predictions

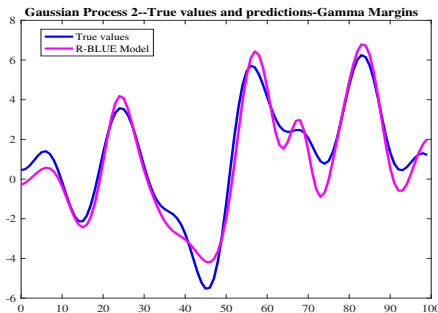


Fig. 4: Gaussian process 2 predictions

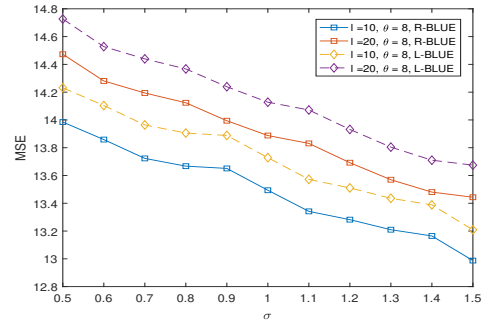


Fig. 5: Comparison of MSE with different L and σ

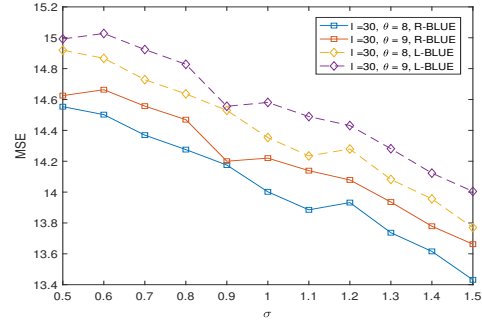


Fig. 6: Comparison of MSE with different θ and σ

B. MSE performance for various parameters under impulsive noise

We also compared the average MSE performance for different sets of parameters, including the length scale (l) and scaling factor (θ) as well as the noise σ . We added impulsive noise equal to amplitude 20 at location 11 of signal and showed the robustness of R-BLUE compared with L-BLUE. Both Figs. 5 and 6 show that the R-BLUE provides smaller MSE compared with L-BLUE despite in the presence of impulsive noise.

VI. CONCLUSIONS

We developed efficient data fusion algorithm for field reconstruction in multimodal sensor networks where complex dependence exists between multimodal fields. We developed low complexity Robust-BLUE method for field reconstruction where dependence is captured through rank correlation. Through extensive simulations, we showed the accuracy of using R-BLUE method and better performance over L-BLUE method which uses traditional Pearson correlation metric in terms of the prediction of mean-squared-errors (MSE).

REFERENCES

- [1] M. Chui, M. Löffler, and R. Roberts, "The internet of things," *McKinsey Quarterly*, vol. 2, no. 2010, pp. 1–9, 2010.

- [2] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of things (iot): A vision, architectural elements, and future directions," *Future generation computer systems*, vol. 29, no. 7, pp. 1645–1660, 2013.
- [3] M. A. Osborne, S. J. Roberts, A. Rogers, and N. R. Jennings, "Real-time information processing of environmental sensor network data using bayesian gaussian processes," *ACM Transactions on Sensor Networks (TOSN)*, vol. 9, no. 1, p. 1, 2012.
- [4] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press, 2005.
- [5] K. Conradsen, L. B. Nielsen, and L. P. Prahm, "Review of Weibull statistics for estimation of wind speed distributions," *Journal of Climate and Applied Meteorology*, vol. 23, no. 8, pp. 1173–1183, 1984.
- [6] W. Y. Yi, K. M. Lo, T. Mak, K. S. Leung, Y. Leung, and M. L. Meng, "A survey of wireless sensor network based air pollution monitoring systems," *Sensors*, vol. 15, no. 12, pp. 31392–31427, 2015.
- [7] S.-C. Kao and R. S. Govindaraju, "Trivariate statistical analysis of extreme rainfall events via the plackett family of copulas," *Water Resources Research*, vol. 44, no. 2, 2008.
- [8] R. B. Nelsen, *An introduction to copulas*. Springer Science & Business Media, 2007.
- [9] L. Bo and C. Sminchisescu, "Twin gaussian processes for structured prediction," *International Journal of Computer Vision*, vol. 87, no. 1-2, pp. 28–52, 2010.
- [10] M. Taylor, "Multivariate measures of concordance," *Annals of the Institute of Statistical Mathematics*, vol. 59, no. 4, pp. 789–806, 2007.
- [11] M. Scarsini, "On measures of concordance." *Stochastica*, vol. 8, no. 3, pp. 201–218, 1984.
- [12] B. Schweizer and E. F. Wolff, "On nonparametric measures of dependence for random variables," *The annals of statistics*, pp. 879–885, 1981.
- [13] M. G. Cruz, G. W. Peters, and P. V. Shevchenko, *Fundamental aspects of operational risk and insurance analytics: A handbook of operational risk*. John Wiley & Sons, 2014.
- [14] W. Xu, Y. Hou, Y. Hung, and Y. Zou, "Comparison of spearman's rho and kendall's tau in normal and contaminated normal models," *arXiv preprint arXiv:1011.2009*, 2010.