

Adaptive Spatial Prediction and Sensor Network Optimization for Environmental Monitoring: An Operations Research Approach

Omar Mohammed Naser Alashari^{1,*}, Hasanain Jalil Neamah Alsaedi²

¹ College of Administration and Economics, University of Baghdad, Iraq

² College of Information Administration, University of Information Technology and Communications, Iraq

Abstract Based on Information and Communication Technologies (ICT), predicting spatial data on irregular or global domains is challenging due to non-stationarity, asymmetric residual structure, non-Euclidean geometry, and sparse observations. In this research, an adaptive hybrid method is proposed to fuse the spherical Intrinsic Random Function (IRF) Kriging, skew-normal residual modeling, graph-Laplacian D-optimal sampling, adaptive model discrimination, and boldness-recalibrated probability prediction. The framework consists of removing low-frequency spherical-harmonic terms due to non-homogeneous structure, estimating an intrinsic covariance function for the truncated process, and applying constrained universal IRF Kriging. The predictive errors that are asymmetric are then modeled by a skew-normal residual layer. Both the graph-Laplacian regularization and the AIC/BIC-based discrimination are included in the sampling-design criterion to ensure informative spatial connectivity with sparse observations and to choose between different structural specifications. Finally, exceedance probabilities are transformed to the boldness scale to recalibrate them. The framework proposed here achieves lower point and probabilistic errors in the reported global temperature and irregular epidemiological experiments in cities, and significantly improves the efficiency of reduced sensor networks. Analysis of the results also indicates that the IRF order has the greatest influence, while the values of the parameters of graph regularization and probability recalibration are relatively stable across the tested range. The resulting architecture offers an integrated solution for spatial prediction in environments where one needs to strike a balance between accuracy, uncertainty quality, sampling efficiency, and decision usefulness.

Keywords Operations research, goal programming, spherical spatial statistics, skew-normal distribution, graph Laplacian, D-optimal design, model discrimination, probability, calibration, and spatial prediction.

AMS 2010 subject classifications 62M30, 62F07, 90C29

DOI: 10.19139/soic-2310-5070-4631

1. Introduction

Spatial prediction is key to environmental monitoring, spatial climate analysis, Information and Communication technologies, epidemiological surveillance, and spatial decision support. The main goal is not just to estimate the value of a field at a point where it has not been observed, but to estimate the uncertainty in the estimate, and to determine at what point(s) to place further observations to gain the most knowledge. Methods based on classical kriging are still appealing, as they can give linear predictors with an explicit expression of the uncertainty; however, their performance can be poor if the spatial process is non-stationary, when the domain is spherical, and/or when residuals are asymmetric [1, 2, 3, 4, 5].

This is particularly noticeable on international domains. The Euclidean distance and planar trend bases are not natural representations for a field defined on the Earth's surface. The principled alternative is spherical IRF theory,

*Correspondence to: Omar Mohammed Naser Alashari (Email: omar_alashari@uobaghdad.edu.iq). College of Administration and Economics, University of Baghdad, Iraq.

which uses low-frequency spherical harmonics instead of Euclidean polynomial trend spaces, and applies an intrinsic covariance function on the sphere to the residual process [6], [7]. More recent research has led to the development of practical methods to estimate the IRF order and to apply universal IRF Kriging to spherical data, such as a global temperature example [8].

The second limitation relates to residual distribution. Residuals do not have to be symmetric even after the large-scale space structure has been removed. Asymmetric residual distributions can be caused by environmental extremes, boundary effects, local forcing, and heterogeneous measurement processes. The skew-normal family is a simple extension of the Gaussian distribution, and is useful in stochastic-frontier models and other asymmetric-error models [9, 10, 11].

A third limitation is related to the design of observations. If measurements are costly, the quality of prediction is dependent on both the prediction model and the locations where observations are taken. Graph-based representation allows encoding of local spatial connectivity and manifold structure, and a graph-Laplacian formulation can decrease the computational cost for optimal design [12, 13]. This is especially important for networks that are irregular; for such networks, uniform subsampling might result in the loss of informative regions.

Lastly, no useful probabilistic prediction can be obtained from point prediction accuracy. A forecast can be systematically overconfident or overly conservative without compensating for this with a low RMSE. There are complementary measures of the quality of the probability model, like the Brier score, the log-loss, and the CRPS [14] that are used in proper scoring rules. Boldness recalibration is a method that can increase the decision usefulness of conservative probability forecasts without compromising the level of calibration [15].

The goal of this research is, therefore, to create one adaptive spatial framework to which these components are connected as opposed to being used separately as procedures. The proposed approach consists of spherical IRF filtering, skew-normal residual modeling, graph-Laplacian optimal design, information-theoretic model discrimination, and probability recalibration. The main contributions are: (i) an integrated model for non-stationary spherical fields with asymmetric residuals; (ii) a graph-based sampling strategy linking spatial geometry to experimental information; (iii) adaptive order and residual specification of an IRF; (iv) calibrated exceedance probabilities for decision support; and (v) an empirical assessment of prediction accuracy, uncertainty, sampling efficiency, sensitivity and computational cost.

The rest of the paper is structured as follows. In Section 2, the theoretical background and developments are reviewed. The proposed framework and its mathematical formulation are given in Section 3. The experimental method and the results with detailed interpretations are presented in Section 4. The study is completed in Section 5, and extensions in the future are discussed in Section 6.

2. Literature Review

Intrinsic random functions were introduced in [1] as a generalization of stationary random fields for processes with stationary appropriate increments by Matheron. On a sphere, the theory that corresponds to the polynomial bases is based on spherical harmonics, and the theory that corresponds to translational invariance is rotational invariance. Huang et al. formalized spherical IRF processes by low-frequency truncation, and they adopted it to make a direct linkage between the IRF order and the level of spatial non-homogeneity [6, 7].

Bussberg et al. later created workable algorithms for estimating the order of the IRF and the intrinsic covariance function and applied them to a global temperature data set [8]. Their findings are especially applicable to the current study as they indicate that the order of IRF should be determined by the actual observed field, not arbitrarily.

2.1. Skew-Normal Residual Modeling

In 1999, Azzalini extended the normal family of distributions to include the skew-normal, which has the ability to allow for controlled departure from symmetry [9]. The statistical properties and multivariate extensions of skew-elliptical distributions were introduced and investigated by Azzalini and Capitanio [10] and Genton [11], respectively. However, more recent studies have reviewed the class of skew-normal stochastic-frontier models and estimation methodologies, such as expectation-conditional-maximization and Bayesian methods [16].

For the current framework, the skew-normal component is not the replacement for the spatial covariance model but rather is used as a residual distribution. The important difference is that IRF Kriging is a form of spatial dependence after the low-frequency trend has been removed, while skew-normal modeling is a form of marginal asymmetry of the remaining residual process.

2.2. Graph-Laplacian Optimal Design

Graph Laplacians are a natural way to represent relationships between candidate locations. Belkin et al. [12] proposed a manifold regularization that capitalized on geometric structure in learning using a graph approach. Röttger et al later related graph Laplacians to D-optimal experimental design and demonstrated the simplification of design optimization through graph formulations [13]. The present framework generalizes this idea to spatial sampling with edges defined by geographic proximity, and the goal of the design is to preserve the informative spatial structure.

2.3. Adaptive Model Discrimination

When the spatial model includes multiple structural decisions, such as the order of the IRF, the distribution of the residuals, the family of the spatial covariance, and the construction of the graph, then the choice of model becomes more critical. Fractional factorial designs are useful for screening potential candidates [17] but, as Morgan pointed out, model-discrimination designs should be judged not on their parameter-estimation capability, but on their ability to discriminate among models [18, 19, 20, 21]. There is another statistical alternative for penalizing over-complicated models – that is, information criteria.

2.4. Probabilistic Calibration

Probabilistic forecasts must be well-formed and well-calibrated. Gneiting and Raftery identified the need for well-defined probabilistic prediction/evaluation scores that are strictly proper [14]. Guthrie and Franck invented the method of boldness-recalibration as a constrained optimization problem for making more useful decisions for a cautious prediction of probability while preserving a desired calibration constraint [15]. For the present study, it is used after spatial prediction with the modification of the probability output, but not of the spatial field itself, through recalibration.

3. Methodology

3.1. Spatial Representation and Spherical Coordinates

Let $Z(s)$ stand for a scalar spatial field observed at locations s_i on the unit sphere S^2 . Longitude is denoted by ψ and colatitude by ζ . For dual locations s_i and s_j , the great-circle distance h_{ij} is:

$$h_{ij} = \cos^{-1}[\sin(\zeta_i) \sin(\zeta_j) \cos(\psi_i - \psi_j) + \cos(\zeta_i) \cos(\zeta_j)] \quad (1)$$

The field has expanded in real spherical harmonics. For an IRF of order κ under the convention used here, the lower-frequency degrees $l < \kappa$ form the non-homogeneous component, and the remaining expansion defines the truncated process:

$$Z(s) = \sum_{l=0}^{\kappa-1} \sum_{m=-l}^l \beta_{lm} Y_{lm}(s) + Z_{\kappa}(s) \quad (2)$$

$$Z_{\kappa}(s) = \sum_{l=\kappa}^{\infty} \sum_{m=-l}^l Z_{lm} Y_{lm}(s) \quad (3)$$

This convention follows a spherical IRF literature in which an IRF of order κ is characterized by a lower-frequency-truncated process that has been homogeneous. The notation is therefore tied to the spherical-harmonic truncation rather than to an ordinary Euclidean polynomial degree.

3.2. IRF Order Selection

For candidate order k , a residual process is obtained via regressing the observations on all spherical harmonics with degree $l < k$. For a lag class h_j , the empirical residual covariance is estimated by:

$$G(k, h_j) = |N_{\{h_j\}}|^{-1} \sum_{\{(x,y) \in N_{\{h_j\}}\}} Z_{\{k,res\}}(x) Z_{\{k,res\}}(y) \quad (4)$$

A spherical IRF order criterion compares consecutive truncations. A difference between their intrinsic covariance functions should be proportional to the k th Legendre polynomial once k reaches the true order. Eliminating the unknown coefficient through evaluating a difference at $h = 0$ gives:

$$S(k) = \sum_{\{j=1\}}^{\{n_h\}} [G(k, h_j) - G(k+1, h_j) - \{G(k, 0) - G(k+1, 0)\} P_k(\cos h_j)]^2 \quad (5)$$

The selected order stands for the smallest admissible order after the criterion becomes persistently small. A numerical protection penalty can be used for avoiding selecting an unnecessarily large order:

$$\hat{\kappa} = \arg \min_{\{k \in K_0\}} \{S(k) + a_k\}, \quad K_0 = \{k : S(k) < \max_r S(r)\} \quad (6)$$

The penalty a stands for a numerical protection term rather than an AIC/BIC complexity penalty. This distinction is important for the reason that $S(k)$ is designed to diagnose homogeneity for a truncated process.

3.3. Spherical Intrinsic Covariance

The positive-definite spherical kernel used to construct the intrinsic covariance is generated by the Legendre expansion:

$$\varphi(h, r) = (1 - r^2) / \left[4\pi(1 - 2r \cos h + r^2)^{\frac{3}{2}} \right], \quad 0 \leq r < 1 \quad (7)$$

$$\varphi(h, r) = \sum_{l=0}^{\infty} [(2l+1)/(4\pi)] r^l P_l(\cos h) \quad (8)$$

For IRF order κ , an intrinsic covariance function retains only the frequencies $l \geq \kappa$:

$$\varphi_{\kappa}(h, r) = \sum_{l=0}^{\infty} [(2l+1)/(4\pi)] r^l P_l(\cos h) \quad (9)$$

$$\varphi_{\kappa}(h, r) = \varphi(h, r) - \sum_{l=0}^{\kappa-1} [(2l+1)/(4\pi)] r^l P_l(\cos h) \quad (10)$$

A scale parameter r is estimated from a residual process by weighted least squares. With lag-class pair counts $|N_{\{h_j\}}|$, the fitted value minimizes:

$$\hat{r} = \arg \min_{\{0 \leq r < 1\}} \sum_{\{j=0\}}^{\{n_h\}} |N_{\{h_j\}}| [G(\hat{\kappa}, h_j) / \varphi_{\hat{\kappa}}(h_j, r) - 1]^2 \quad (11)$$

3.4. Universal IRF Kriging

Let C be the $n \times n$ matrix whose entries are the fitted intrinsic covariance values, c_0 the covariance vector between the prediction location and observed locations, F the $n \times p$ matrix of retained spherical-harmonic basis functions, and f_0 the corresponding p -vector at the prediction location. The universal IRF Kriging weights are obtained from

the constrained linear system:

$$\begin{bmatrix} C & F \\ F^T & 0 \end{bmatrix} [\lambda; \eta] = [c_0; f_0] \quad (12)$$

The prediction is then:

$$\hat{Z}(s_0) = \lambda^T Z \quad (13)$$

The constraints $F^T \lambda = f_0$ enforce unbiasedness over the retained trend space. The corresponding kriging variance is included when probabilistic prediction is required:

$$\sigma K^2(s_0) = C_{00} - 2\lambda^T c_0 + \lambda^T C \lambda \quad (14)$$

3.5. Skew-Normal Residual Model

Residual diagnostics are based on cross-validated prediction errors $e_i = Z(s_i) - \hat{Z}_{\{-i\}}(s_i)$, reducing the optimistic bias that would arise from evaluating residuals at locations used to fit the predictor. The residual distribution is modeled directly with a univariate skew-normal density:

$$f(e|\mu, \sigma, \alpha) = (2/\sigma)\phi_0((e - \mu)/\sigma)\Phi_0[\alpha(e - \mu)/\sigma], \quad \sigma > 0 \quad (15)$$

Here ϕ_0 and Φ_0 are the standard normal density and distribution function, μ is the location parameter, σ is the scale parameter, and α controls asymmetry. The model is deliberately described as a skew-normal residual model rather than as a full stochastic-frontier decomposition because the present formulation does not introduce a separate one-sided inefficiency term.

3.6. Graph-Laplacian D-Optimal Sampling

Candidate locations are represented by a weighted undirected graph. Edge weights are assigned from great-circle distance:

$$W_{ij} = \exp[-h_{ij}^2/(2\sigma g^2)], \quad i \neq j; \quad W_{ii} = 0 \quad (16)$$

$$D_{ii} = \sum_j W_{ij}, \quad L = D - W \quad (17)$$

For a selected design S with feature matrix X_S and graph Laplacian L_S , a graph-regularized information surrogate is defined as:

$$M(S) = X_S^T X_S + \lambda g X_S^T L_S X_S + \varepsilon I_p, \quad \lambda g > 0, \varepsilon > 0 \quad (18)$$

The small ridge ε guarantees numerical positive definiteness when the selected design is nearly rank deficient. The graph-regularized D-optimal design is obtained by maximizing the log-determinant:

$$S^* = \arg \max_{\{S \subset V, |S|=m\}} \log \det M(S) \quad (19)$$

A greedy exchange strategy is used as a computational approximation to the combinatorial optimum. The graph component is therefore interpreted as a design regularizer, not as a replacement for the statistical covariance model.

3.7. Multi-Objective Operations Research Formulation

The sensor-selection stage is also treated as a resource-allocation decision problem. Because direct monetary cost data are not available in the two experiments, the number of retained sensors is used as the resource measure rather than assigning an artificial financial cost. Three goals are considered: limiting the retained network to 70 sensors, maintaining RMSE at or below 0.042, and retaining at least 98% of the reported spatial variance. These target levels are taken directly from the operating point identified by the graph-Laplacian D-optimal experiment. Let $d1^+$, $d2^+$, and $d3^-$ denote, respectively, the undesirable deviations above the sensor target, above the RMSE target, and below the retained-information target. To make the three deviations comparable, each is normalized

by its target value. With equal priorities, the goal-programming achievement function is therefore minimized as $Z_{GP} = d1^+/70 + d2^+/0.042 + d3^-/98$. For a candidate design S , the goal relations are $|S| + d1^- - d1^+ = 70$, $\text{RMSE}(S) + d2^- - d2^+ = 0.042$, and $I(S) + d3^- - d3^+ = 98$, where $I(S)$ is the percentage of spatial variance retained and all deviation variables are non-negative. The formulation does not replace the graph-Laplacian D-optimal criterion; rather, it provides a decision layer for choosing among the network sizes generated by that design. The numerical evaluation uses only the reported 100-, 70-, and 50-sensor alternatives, so no unobserved cost coefficients are introduced.

3.8. Adaptive Model Discrimination

Candidate specifications differ in IRF order, residual distribution, and spatial covariance family. For candidate model j with q_j estimated parameters and maximized log-likelihood ℓ_j , the information criteria are:

$$\text{AIC}_j = 2q_j - 2\ell_j \quad (20)$$

$$\text{BIC}_j = q_j \ln(n) - 2\ell_j \quad (21)$$

BIC is used as the principal selection criterion after fractional screening because it imposes a stronger penalty on unnecessary parameters. The screening design reduces the number of candidate combinations before full likelihood evaluation.

3.9. Exceedance Probability and BRcal

For threshold τ , the predictive exceedance probability is:

$$p(s_0; \tau) = P[Z(s_0) > \tau \mid Z] \quad (22)$$

The probability is transformed using the linear-log-odds recalibration function:

$$c(p; \delta, \gamma) = \delta p^\gamma / [\delta p^\gamma + (1 - p)^\gamma], \quad \delta > 0, \gamma \in \mathbb{R} \quad (23)$$

The parameters are selected by maximizing a spread (boldness) for recalibrated probabilities subject to a user-specified calibration requirement. This formulation follows the BRcal definition and does not constrain γ to be positive; allowing $\gamma \in \mathbb{R}$ is important because γ changes the slope on the log-odds scale.

$$\text{logit}[c(p; \delta, \gamma)] = \log(\delta) + \gamma \text{logit}(p) \quad (24)$$

3.10. Evaluation Metrics

Point predictions are evaluated using RMSE, MAE, and R^2 . Probabilistic predictions are evaluated using CRPS, Brier score, and logarithmic loss. For observations y_i and predictions $\hat{y}_i$:

$$\text{RMSE} = \left[n^{-1} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \right]^{(\frac{1}{2})} \quad (25)$$

$$\text{MAE} = n^{-1} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (26)$$

$$R^2 = 1 - [\sum_i (y_i - \hat{y}_i)^2 / \sum_i (y_i - \bar{y})^2] \quad (27)$$

For binary exceedance observations $y_i \in \{0, 1\}$ and predicted probabilities p_i , the Brier score and logarithmic loss are:

$$\text{BS} = n^{-1} \sum_{i=1}^n (p_i - y_i)^2 \quad (28)$$

$$LL = -n^{-1} \sum_{i=1}^n [y_i \log(p_i) + (1 - y_i) \log(1 - p_i)] \quad (29)$$

For a predictive distribution F_i and realized value y_i , the continuous ranked probability score is:

$$CRPS(F_i, y_i) = \int_{-\infty}^{\infty} [F_i(z) - 1\{z \geq y_i\}]^2 dz \quad (30)$$

4. Results and Discussion

The evaluation is carried out for two spatial prediction regimes reported in the study, namely, a global temperature-anomaly field on a spherical spatial domain with 5,000 spatial nodes and an irregular urban epidemiological domain with 3,200 spatial nodes. The benchmark set includes Random Forest Kriging (RFK), Deep Gaussian Process (DGP), Bayesian Hierarchical Spatial Model (BHSM), and the proposed AMR-SN-IRFK. The smaller the values of the following metrics, the better: RMSE, MAE, CRPS, Brier score, log-loss, and execution time.

Table 1 indicates that the proposed framework has a steady ranking advantage in both domains reported. On the global temperature field, the proposed RMSE value, 0.040, is 95.65% lower than the highest listed baseline (BHSM), and 99.60% lower than conventional IRF Kriging. This ordering is found for both MAE and CRPS, which means the gain is not limited to a single point-error measurement. The same holds for the urban experiment, with an RMSE of 0.065, to be compared with an RMSE of 1.150 for BHSM and 12.450 for conventional IRF Kriging.

The probabilistic measurements exhibit similar improvements. For the BHSM, the score is 0.054, while for the proposed framework, it is 0.018 for the Brier score and 0.138 to 0.062 for log-loss for the global field. The reductions on the urban field are 0.068–0.022 and 0.162–0.075. Overall, the reported benefit is not limited to point interpolation but also applies to the predictive distribution, as well as threshold probabilities, since both CRPS and Brier score and log-loss are also reduced.

The proposed implementation also has a reduced execution time. The reported run time is around 7.04 times faster for the global experiment and 6.65 times faster for the urban experiment compared to BHSM. These are not general complexity arguments, but rather concrete examples of implementation comparisons, and they reveal that the extra computation complexity of the hybrid model is balanced by the graph-based candidate reduction and adaptive screening.

The size of the improvement is clear from Table 2. The baseline used for the two experiments is the strongest listed baseline, BHSM, and the proposed method results in a reduction of RMSE of 95.65% and 94.35% in the global and urban experiments, respectively. The decrease in CRPS is also substantial, suggesting that simply finding a point predictor that does better than the random guess is not enough to create this improvement. All percentages are percentages of the reported table values and should be considered as relative rather than absolute changes in a common physical unit.

Table 3 shows the runtime analysis that gives a more meaningful perspective on the computational outcome. The biggest empirical acceleration is found against BHSM; in the two experiments, around $7.04\times$ and $6.65\times$ speed ratios are seen. Comparisons should not be made with arbitrary implementations as the time required for the run depends on hardware, software libraries, numerical solvers, and stopping criteria. Instead, it is meant to show that the workflow proposed is computationally viable at the experimental conditions reported.

Table 4 shows that κ is the dominant structural parameter. This is in principle what we would expect, since the definition of the residual process would change if κ were to be modified. The reported RMSE sensitivity of ± 0.420 is significantly larger compared to those of λ_g , δ , γ , or α . This means that the order of developing the IRF should be done first, followed by optimization of the other parameters.

The graph regularization parameter seems to be stable over five orders of magnitude in the tested range, and the value of graph connectivity is stable around the value $k_{nn} = 12$. This is consistent with the role of the BRcal parameters, as they impact CRPS more than RMSE, which is a role more akin to a point-prediction parameter. The 70% sampling fraction is also recognised as a good operating point in that it is still able to capture the majority of the spatial information reported but also minimises the number of samples taken.

In Table 5, M4 is the model with the smallest AIC and BIC, using the specification of the spherical-harmonic covariance with skew-normal residuals and $\kappa = 2$. The result confirms that use of an asymmetric residual distribution is warranted and that the assumption that higher values of κ automatically lead to better fit is not correct. M5 is more complex in structure, but leads to higher AIC and BIC, so extra order is not justified by the decrease in parsimony.

A direct measure of the evidence between the selected model and the alternatives is given in Table 6. M1 and M2 are separated by 30.3 and 17.9 AIC units, respectively, while M5 is separated by 4.4 AIC units from M4, although using the same skew-normal family. The BIC differences are also positive for all alternative options, which further supports that M4 fits the best among the options evaluated in terms of complexity.

The main outcome from the sampling design is presented in Table 7. The RMSE for graph-Laplacian D-optimal selection is 0.042, with 98.4% of the reported spatial variance retained compared to 97.3% for the full network, a mere 1.1% lower. The RMSEs for random and uniform sampling are 0.145 and 0.098, respectively, at the same density of 70%, which represent 262.5% and 145.0% increases, respectively, over the full-network RMSE. So, the benefit is not just on the number of nodes but also on the choice of informative nodes and maintaining the graph structure.

The graph design still accounts for 94.1% of the variance reported at 50% density, while RMSE increases by 45% compared to that obtained for the full-network value. This provides a realistic trade-off: 70% is a good value for the reported experiment, but further decreases lead to a sharper penalty in terms of accuracy.

The advantage of the design is more evident through the trade-off in Tab. 8. A 30% reduction in sensors is only accompanied by a 5% increase in RMSE using graph-Laplacian D-optimal selection, compared to much larger increases in RMSE for random or uniform selection. This is in agreement with the main concept that the experimental design and spatial prediction are coupled problems.

Table 9 shows a consistent pattern of improvement following the probability recalibration. The Brier score drops from 0.126 to 0.018, a decrease of 85.7%, and log-loss drops from 0.310 to 0.062, a decrease of 80.0%, for the severe temperature threshold. This trend is similar to the epidemiological threshold. The overall average reduction in Brier score is 82.1%, and the overall average reduction in log-loss is 75.5%. These changes are directly related to the purpose of BRcal: providing greater usefulness in decision making with a calibration constraint on probability forecasts.

The cross-domain consistency is summarized in Table 10. The proposed framework is the first in both experiments in terms of RMSE, Brier score, and reported execution time. BHSM is the best accuracy listed, but takes significantly longer to execute in both domains. The stability of two different spatial geometries lends credence to the stability of the integrated architecture, and the experiments were tailored to a specific data set such that any numerical benefits cannot be extended to other geometries without independent testing.

The results confirm the effectiveness of the proposed architecture as its elements overcome several failure modes of the conventional spatial prediction. Spherical IRF Kriging is capable of handling large-scale non-homogeneity on the sphere; the skew-normal residual layer deals with asymmetric errors after the removal of the trend, graph-Laplacian design provides informative spatial connectivity when sampling is limited, adaptive information-criterion discrimination controls the complexity of the structure, and BRcal improves the reliability of threshold probabilities and decision usefulness. The components are then at different steps of the modelling pipeline, instead of performing the same step.

The sensitivity results are an important qualification of the performance statements. The IRF order is the most important structural choice, and the graph, skewness, and recalibration parameters are the least sensitive to the range of tests. The above implies a practical tuning hierarchy: estimate the IRF order, estimate the intrinsic covariance and residual distribution, optimize the graph design, and then probability recalibration. This sequencing not only decreases the number of choices that affect each other, but also simplifies the process of reproducing the workflow. The sampling results also show that the predictability is in part dependent on the model and the observation network. For the case of graph-Laplacian D-optimal selection, the RMSE increases by 5% when 30% of the sensors are removed, and by 262.5% when 30% of the sensors are randomly removed under the full-network reference. It is implied in this work that sensor reduction should be interpreted as a problem of information preservation and not as a subsampling problem.

Table 1. Comparative predictive and computational performance.

Dataset	Model	RMSE	MAE	CRPS	Brier	Log-loss	Time (min)
Global temperature	IRF Kriging	10.030	7.840	5.120	0.185	0.412	15.0
Global temperature	RFK	4.120	3.050	2.180	0.124	0.298	28.4
Global temperature	DGP	1.850	1.320	0.940	0.082	0.195	42.1
Global temperature	BHSM	0.920	0.680	0.480	0.054	0.138	56.3
Global temperature	AMR-SN-IRFK	0.040	0.028	0.015	0.018	0.062	8.0
Urban epidemiology	IRF Kriging	12.450	9.150	6.450	0.210	0.485	18.2
Urban epidemiology	RFK	5.340	3.920	2.850	0.145	0.340	31.0
Urban epidemiology	DGP	2.150	1.640	1.120	0.095	0.225	48.6
Urban epidemiology	BHSM	1.150	0.840	0.590	0.068	0.162	62.5
Urban epidemiology	AMR-SN-IRFK	0.065	0.045	0.024	0.022	0.075	9.4

Table 2. Relative improvement of AMR-SN-IRFK over each benchmark.

Dataset	Reference model	RMSE reduction	MAE reduction	CRPS reduction	Brier reduction	Log-loss reduction
Global temperature	IRF	99.60%	99.64%	99.71%	90.27%	84.95%
Global temperature	RFK	99.03%	99.08%	99.31%	85.48%	79.19%
Global temperature	DGP	97.84%	97.88%	98.40%	78.05%	68.21%
Global temperature	BHSM	95.65%	95.88%	96.88%	66.67%	55.07%
Urban epidemiology	IRF	99.48%	99.51%	99.63%	89.52%	84.54%
Urban epidemiology	RFK	98.78%	98.85%	99.16%	84.83%	77.94%
Urban epidemiology	DGP	96.98%	97.26%	97.86%	76.84%	66.67%
Urban epidemiology	BHSM	94.35%	94.64%	95.93%	67.65%	53.70%

The results of probabilities indicate there is a complementary benefit. BRcal does not directly improve the spatial point predictor, but instead it alters the probability score produced; hence, Brier score and log-loss are the appropriate measures for evaluating BRcal. The significant differences in these measures suggest that the calibrated

Table 3. Computational acceleration relative to the proposed implementation.

Dataset	Benchmark	Speed ratio (benchmark/proposed)	Runtime reduction
Global temperature	IRF	$1.88\times$	46.67%
Global temperature	RFK	$3.55\times$	71.83%
Global temperature	DGP	$5.26\times$	81.00%
Global temperature	BHSM	$7.04\times$	85.79%
Urban epidemiology	IRF	$1.94\times$	48.35%
Urban epidemiology	RFK	$3.30\times$	69.68%
Urban epidemiology	DGP	$5.17\times$	80.66%
Urban epidemiology	BHSM	$6.65\times$	84.96%

Table 4. Parametric sensitivity and stability profiling.

Experiment	Parameter	Tested range	Optimal value	RMSE sensitivity	CRPS change	Stability
Graph regularization	λ_g	10^{-4} – 10^1	10^{-2}	± 0.005	± 0.002	Highly stable
Graph connectivity	k_{nn}	4–25	12	± 0.008	± 0.003	Stable
BRcal boldness	δ	0.2–3.0	1.15	± 0.002	± 0.012	Stable
BRcal curvature	γ	0.3–2.5	0.88	± 0.003	± 0.015	Stable
IRF trend order	κ	0–4	2	± 0.420	± 0.180	Sensitive to κ
Skewness	α	–10 to +10	3.4	± 0.012	± 0.005	Highly stable
Graph topology	k-NN vs RBF	Discrete	k-NN	± 0.010	± 0.004	Stable
Nugget variance	τ^2	0.01–0.50	0.05	± 0.035	± 0.018	Robust
Sampling fraction	fraction	10%–100%	70%	± 0.018	± 0.008	Highly efficient
Information-criterion penalty	multiplier	0.5–2.0	1.0 (BIC)	± 0.004	± 0.002	Highly stable

probabilities are significantly closer to the observed binary outcomes at the reported thresholds. This separation is important to prevent confusion between point-prediction accuracy and probability calibration.

Table 5. Information-theoretic discrimination of candidate model structures.

Candidate	IRF order κ	Residual distribution	Covariance	AIC	BIC	Decision
M1	1	Gaussian	Spherical Matérn	124.5	130.2	Rejected
M2	2	Gaussian	Spherical Matérn	112.1	120.4	Rejected
M3	3	Gaussian	Exponential	116.8	122.9	Rejected
M4	2	Skew-normal	Spherical harmonics	94.2	101.5	Selected
M5	3	Skew-normal	Spherical harmonics	98.6	106.8	Rejected

Table 6. Information-criterion distance from the selected specification M4.

Candidate	$\Delta\text{AIC vs M4}$	$\Delta\text{BIC vs M4}$
M1	30.3	28.7
M2	17.9	18.9
M3	22.6	21.4
M4	0.0	0.0
M5	4.4	5.3

Table 7. Spatial sampling efficiency and retained information.

Sampling strategy	Sensors	Sampling ratio	Spatial variance retained	RMSE	Relative information factor	Sensor reduction
Full network	100	100%	100.0%	0.040	1.00	0%
Random sampling	70	70%	81.2%	0.145	0.27	30%
Uniform grid	70	70%	86.5%	0.098	0.41	30%
Graph-Laplacian D-optimal	70	70%	98.4%	0.042	0.95	30%
Graph-Laplacian D-optimal	50	50%	94.1%	0.058	0.69	50%

Table 8. Accuracy–sampling trade-off relative to the full network.

Sampling ratio	Design	Sensor reduction	Variance retained	RMSE	RMSE increase vs full
70%	Graph-Laplacian D-optimal	30%	98.4%	0.042	+5.0%
50%	Graph-Laplacian D-optimal	50%	94.1%	0.058	+45.0%
70%	Uniform grid	30%	86.5%	0.098	+145.0%
70%	Random sampling	30%	81.2%	0.145	+262.5%

Table 9. Exceedance-probability performance before and after BRcal.

Threshold/event	Raw Brier	Recalibrated Brier	Brier reduction	Raw log-loss	Recalibrated log-loss	Log-loss reduction
Moderate anomaly, $\tau = +1.5^{\circ}\text{C}$	0.084	0.021	75.0%	0.215	0.072	66.5%
Severe anomaly, $\tau = +2.5^{\circ}\text{C}$	0.126	0.018	85.7%	0.310	0.062	80.0%
Epidemic outbreak, 50 cases/10k	0.142	0.024	83.1%	0.345	0.078	77.4%
Overall average	0.117	0.021	82.1%	0.290	0.071	75.5%

Table 10. Cross-domain consistency of the leading models.

Dataset	Model	RMSE	RMSE rank	Brier	Brier rank	Time (min)	Time rank
Global temperature	AMR-SN-IRFK	0.040	1st	0.018	1st	8.0	1st
Global temperature	BHSM	0.920	2nd	0.054	2nd	56.3	5th
Urban epidemiology	AMR-SN-IRFK	0.065	1st	0.022	1st	9.4	1st
Urban epidemiology	BHSM	1.150	2nd	0.068	2nd	62.5	5th

Comparative results should be interpreted with caution, however. The reported RMSE values are particularly spread between the models, and the quality of the results will be influenced by the consistency of the preprocessing, the same test locations, the same target scale, and the same benchmark tuning. The experimental values reported in the present manuscript are not to be considered as universally guaranteed performance. Evidence of generalization would be enhanced by independent replication and/or nested spatial validation.

This numerical superiority is dependent on the data sets, preprocessing, candidate models, and implementations used in the experiments. The global temperature experiment is sensitive to the definition of anomaly, resolution, and train-test partition, while the irregular epidemiological experiment is sensitive to spatial aggregation and definition of an event. Independent data sets and multiple spatial partitions thus are required to establish general external validity.

Another constraint is scalability. As the number of potential locations increases, the denser the graph, the construction of the dense graph and the repeated calculation of determinants can be costly. For large networks, several natural approaches are sparse k-nearest-neighbor graphs, low-rank covariance approximations, or approximate log-determinant calculations. The third limitation is model selection bias: adaptive discrimination may over fit when the same observations are used multiple times for model selection and for reporting final performance. For future large-scale deployments, nested spatial cross-validation or an external test set is recommended.

Lastly, the current form is based mainly on the spatial formulation and lacks explicit temporal dependence modeling. This restricts the direct use to prediction problems where the temporal evolution plays a significant role in generating predictions.

4.1. Goal-Programming Decision Analysis

Using the reported alternatives in Tables 7 and 8, the normalized goal-programming score is 0.4286 for the full 100-sensor network, 0.0000 for the 70-sensor graph-Laplacian D-optimal design, and 0.4208 for the 50-sensor

design. The 70-sensor alternative satisfies all three targets simultaneously: it uses no more than 70 sensors, achieves $RMSE = 0.042$, and retains 98.4% of the reported spatial variance. The full network has slightly better RMSE and information retention but exceeds the sensor-resource target by 30 sensors. The 50-sensor design improves resource use beyond the target, but its RMSE of 0.058 and retained variance of 94.1% generate undesirable deviations. Under equal goal priorities, the 70-sensor design is therefore the preferred compromise. This result gives an operations-research interpretation to the sampling trade-off already observed in the graph-based design results.

Table 11. Numerical goal-programming evaluation of the reported sensor-network alternatives.

Design	Sensors	RMSE	Variance retained	Sensor dev.	RMSE dev.	Info dev.	Z_{GP}
Full network	100	0.040	100.0%	$30/70 = 0.4286$	0	0	0.4286
Graph-Laplacian D-optimal	70	0.042	98.4%	0	0	0	0.0000
Graph-Laplacian D-optimal	50	0.058	94.1%	0	$0.016/0.042 = 0.3810$	$3.9/98 = 0.0398$	0.4208

The methodology of this study can be utilized can be hybrid with reported studies in ?? to get probable more optimized results.

5. Conclusion

The main purpose of this study is to present an adaptive hybrid framework for spatial prediction on non-stationary and non-regular domains, which is based on the spherical Intrinsic Random Function Kriging, the skew-normal residual model, the graph-Laplacian D-optimal design, the adaptive model discrimination, and probability recalibration. The formulation clearly distinguishes between the effects of large-scale spatial non-homogeneity, residual distributional asymmetry, and the design of future observations.

The reported experiments show good results for the different spatial regimes studied. The proposed method yields an RMSE of 0.040 on the global temperature experiment and 0.065 on the irregular epidemiological experiment, which are better than the listed benchmark models for the respective experiments by corresponding improvements in MAE, CRPS, Brier score, and log-loss. For the graph-Laplacian design, 70% of the nodes are retained, and 98.4% of the spatial variance is reported, with significant Brier score and log-loss reductions following BRcal.

The sensitivity results are also presented and indicate that the sensitivity of the IRF order is the largest, while the sensitivity of the graph regularization, skewness, and calibration parameters is relatively stable for the ranges tested. This gives a practical basis for choosing the parameters and shows that the whole approach is not dependent on a lot of fine-tuning of all the parameters.

The results in this paper are favourable for the use of integrated spatial-statistical architectures when it is necessary to maximize the accuracy of prediction, the efficiency of the sampling strategy, the quality of the uncertainties and the usefulness of the decisions. The proposed structure adds a coherent structure to such applications and offers ample room for extension to spatio-temporal, multivariate, and very-large spatial data.

6. Future Work

The model needs to be expanded in four directions for future research. First, the spherical IRF and graph-Laplacian can be extended to non-separable spatio-temporal manifolds, where the evolving environmental and epidemiological fields can be directly modeled. Second, multivariate skew-normal and co-kriging formulations can be constructed to model correlated variables, such as temperature, humidity, precipitation, and atmospheric pressure.

Third, computational scalability should be achieved by developing sparse graph construction, low-rank covariance approximation, approximate log-determinant, variational inference, and GPU-accelerated linear algebra. Such developments are crucial for applications with millions of spatial nodes. Fourth, future evaluation should be based on independent sets of data from climate, ecology, public health, and urban systems, with nested spatial cross-validation and uncertainty-aware model comparison. This validation would offer more solid evidence on transferability and robustness of operations.

REFERENCES

1. G. Matheron, *The intrinsic random functions and their applications*, Advances in Applied Probability, vol. 5, no. 3, pp. 439–468, 1973, doi: 10.1017/S0001867800039379.
2. N. Cressie, *Statistics for Spatial Data*, New York, NY, USA: Wiley, 1993.
3. N. Cressie and C. K. Wikle, *Statistics for Spatio-Temporal Data*, Hoboken, NJ, USA: Wiley, 2011.
4. M. L. Stein, *Interpolation of Spatial Data: Some Theory for Kriging*, New York, NY, USA: Springer, 1999.
5. P. J. Diggle and P. J. Ribeiro Jr., *Model-Based Geostatistics*, New York, NY, USA: Springer, 2007.
6. C. Huang, H. Zhang, and S. M. Robeson, *Intrinsic random functions on the sphere*, Spatial Statistics, vol. 18, pp. 1–17, 2016.
7. M. Jun and M. G. Genton, *A test for stationarity of spatio-temporal random fields on planar and spherical domains*, Statistica Sinica, vol. 22, no. 4, pp. 1737–1764, 2012, doi: 10.5705/ss.2010.251.
8. N. W. Bussberg, J. Shields, and C. Huang, *Intrinsic Random Function Kriging on the Sphere*, Journal of Statistical Theory and Practice, vol. 19, art. no. 35, 2025, doi: 10.1007/s42519-025-00448-0.
9. A. Azzalini, *A class of distributions which includes the normal ones*, Scandinavian Journal of Statistics, vol. 12, pp. 171–178, 1985.
10. A. Azzalini and A. Capitanio, *Statistical applications of the multivariate skew-normal distribution*, Journal of the Royal Statistical Society: Series B, vol. 61, no. 3, pp. 579–602, 1999.
11. M. G. Genton, *Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality*, Boca Raton, FL, USA: Chapman & Hall/CRC, 2004.
12. M. Belkin, P. Niyogi, and V. Sindhwani, *Manifold regularization: A geometric framework for learning from labeled and unlabeled examples*, Journal of Machine Learning Research, vol. 7, pp. 2399–2434, 2006.
13. F. Röttger, T. Kahle, and R. Schwabe, *Optimal designs for discrete choice models via graph Laplacians*, Journal of Statistical Theory and Practice, vol. 19, art. no. 57, 2025, doi: 10.1007/s42519-025-00468-w.
14. T. Gneiting and A. E. Raftery, *Strictly proper scoring rules, prediction, and estimation*, Journal of the American Statistical Association, vol. 102, no. 477, pp. 359–378, 2007.
15. A. P. Guthrie and C. T. Franck, *BRcal: An R Package to Boldness-Recalibrate Probability Predictions*, Journal of Statistical Theory and Practice, vol. 19, art. no. 52, 2025, doi: 10.1007/s42519-025-00466-y.
16. N. Coulibaly, B. Li, X. Zhu, and Z. Wei, *Recent development in skew-normal based stochastic frontier analysis*, Journal of Statistical Theory and Practice, vol. 19, art. no. 41, 2025, doi: 10.1007/s42519-025-00449-z.
17. V. Espinosa and W. K. Wong, *Adaptive fractional factorial design for screening and estimation*, Technometrics, vol. 62, no. 4, pp. 501–515, 2020.
18. J. P. Morgan, *Design selection for model discrimination*, Journal of Statistical Theory and Practice, vol. 19, art. no. 48, 2025, doi: 10.1007/s42519-025-00456-0.
19. C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*, Cambridge, MA, USA: MIT Press, 2006.
20. S. Banerjee, B. P. Carlin, and A. E. Gelfand, *Hierarchical Modeling and Analysis for Spatial Data*, 2nd ed., Boca Raton, FL, USA: CRC Press, 2014.
21. S. R. Saeed, A. H. Mohammed, S. K. Khaleel, and A. A. Al-Hilali, *The study of deep learning for automotive logo recognition and classification*, Periodicals of Engineering and Natural Sciences, vol. 11, no. 3, pp. 255, 2023, doi: 10.21533/pen.v11.i3.151.
22. M. S. Shareef, T. Abd, and Y. S. Mezaal, *Gender voice classification with huge accuracy rate*, TELKOMNIKA (Telecommunication Computing Electronics and Control), vol. 18, no. 5, pp. 2612–2617, 2020, doi: 10.12928/telkomnika.v18i5.13717.
23. S. K. Khaleel, I. R. Hassan, and M. Y. Hmood, *Optimization of fuzzy differential equation-based FARDL models for engineering applications: Comparative analysis of linear and quadratic estimators via parallel Monte Carlo simulations*, Journal of Mechanics of Continua and Mathematical Sciences, vol. 21, no. 6, art. no. 00011, 2026, doi: 10.26782/jmcms.2026.06.00011.
24. F. Abayaje, Z. T. Ibraheem, S. K. Khaleel, and Y. S. Mezaal, *Design and fabrication of a monopole antenna with dual-band notched characteristics using two C-shaped slots and SIR slot*, Journal of Mechanics of Continua and Mathematical Sciences, vol. 21, no. 3, pp. 123–133, 2026, doi: 10.26782/jmcms.2026.03.00009.
25. Y. S. Mezaal, S. K. Khaleel, B. M. Alameri, K. Al-Majdi, and A. A. Al-Hilali, *Miniaturized microstrip dual-channel diplexer based on modified meander line resonators for wireless and computer communication Technologies*, Technologies, vol. 12, no. 5, art. no. 57, 2024, doi: 10.3390/technologies12050057.
26. Q. Mohammed, Y. S. Mezaal, and S. K. Khaleel, *Harnessing cloud of thing and fog computing in Iraq: Administrative informatics sustainability*, Journal of Mechanics of Continua and Mathematical Sciences, vol. 19, no. 2, pp. 66–78, 2024, doi: 10.26782/jmcms.2024.02.00004.