A Geostatistical framework to Interpolate Sustainable Aviation Data

Remi Perrichon, Xavier Gendre, Thierry Klein

A Geostatistical framework to Interpolate Sustainable Aviation Data

Rémi Perrichon (corresponding author)

École Nationale de l'Aviation Civile (ENAC), Université de Toulouse, France.

E-mail: remi.perrichon@enac.fr

Xavier Gendre

Institut de Mathématiques de Toulouse (IMT), Université de Toulouse, France.

Pathway.com, Paris, France.

Thierry Klein

École Nationale de l'Aviation Civile (ENAC), Université de Toulouse, France.

Institut de Mathématiques de Toulouse (IMT), Université de Toulouse, France.

Summary. In the sustainable aviation literature, it is common to rely on multiple data sets, including traffic, noise, and meteorological data. Spatial interpolation of these data is a frequent practice. This article presents a comprehensive comparison of interpolation techniques for aviation data, focusing on stochastic methods. Two primary case studies are investigated: the interpolation of noise in the vicinity of airports (two-dimensional interpolation) and the interpolation of weather values across multiple pressure layers (three-dimensional interpolation). In the first case, standard geostatistical methods are particularly well-suited. The interpolation of meteorological values is more complex. The underlying physics gives rise to a trend and complex anisotropies that are challenging to estimate with the basic geostatistical framework. In the absence of outliers and if meteorological data are sampled on a sufficiently dense regular grid, trilinear interpolation performs comparably to advanced statistical methods. Otherwise, statistical methods appear to be more flexible.

Keywords: Aircraft Noise, Condensation Trails, Interpolation, Kriging

1. Introduction

The swift expansion of worldwide aviation activities has rendered their adverse ecological effects a global apprehension. Notably, the first report from the Intergovernmental Panel on Climate Change (IPCC) on a specific industrial subsector is the one on aviation and its consequences on the atmosphere that was written by Penner et al. (1999). In addition to CO_2 emissions, Lee et al. (2009) have shown that non-CO_2 effects are substantial yet generally challenging to estimate. More specifically, Lee et al. (2021) have highlighted that the largest positive (warming) climate forcings adding to that of CO_2 are those from contrail cirrus and from NO_x-driven changes in the chemical composition of the atmosphere. However, the environmental impact of aviation is not limited to the climate. Namely, the noise produced by aircraft during their operation represents an ecological, economic, and social problem which is increasingly documented in the literature as shown by Franssen (2004), Cohen and Coughlin (2008), Salvi (2008), Zheng et al. (2020).

Contributions on aviation sustainability commonly involve the use of multiple data sources, such as weather, noise, and traffic data. Often, there is a requirement for *spatial interpolation*, primarily due to the disparate levels of granularity in the observation of these data. Many examples of spatial interpolation can be found in the scientific literature on contrails and noise pollution, two topics for which it is necessary to provide some contextual background:

- **Contrails**. As put by Kärcher (2018), *condensation trails* (contrails) are line-shaped ice clouds generated by jet aircraft cruising in the upper troposphere at 8-13 km altitude. Depending on surrounding atmospheric conditions, contrails can be short- or long-lived. The theory of *contrail formation* is now well documented. According to Paoli and Shariff (2016), the formation stage of contrails lasts for about 10 minutes. The thermodynamic mixing model of Schumann (1996) has shown that temperatures typically below 233 K (≈ -40 °C) provides a threshold below which either short-lived or long-lived contrails appear behind jet aircraft. *Contrail occurrence* is predicted with confidence if ambient pressure, relative humidity, water vapour, heat emissions, and propulsive characteristics of aircraft engines are known. Aircraft typically form *persistent* contrails when flying through pockets of

air that are cold and have relative humidity greater than 100% with respect to ice, so-called ice supersaturated regions. A presentation of such regions is to be found in the work of Gierens et al. (2012).

To assess the actual environmental impact of contrail, *contrail detection* and *contrail tracking* are key. A method proposed by Vazquez-Navarro et al. (2010) follows the evolution of contrails from their linear stage until they are indistinguishable from natural cirrus clouds. In a recent work, Chevallier et al. (2023) have introduced a procedure to detect contrails and identify the aircraft that generated them. To account for contrail lifetimes that are underestimated due to the limitation of satellite data, Gierens and Vazquez-Navarro (2018) have complemented some previous works with a Weibull distribution model that describes the survival rate of contrails.

In many analyses, it is often needed to perform an interpolation of weather values. For example, Duda et al. (2004) have used linear interpolation to obtain weather data over the Great Lakes at a finer spatial and temporal resolution whereas Schumann (2012) has relied on linear interpolation to input ambient meteorological conditions in its Contrail Cirrus Prediction Tool (CoCiP). More recently, linear interpolation has been used by Gierens et al. (2020) to compare ERA-5 and MOZAIC data. An analogous approach is employed by Wilhelm et al. (2021, 2022). Aviation contrail climate effects in the North Atlantic have been assessed by Teoh et al. (2022). Linear interpolation is utilized to associate each waypoint of a flight to meteorological data. Discussion on the relevance of linear interpolation is often limited to the vertical dimension. Notably, Gierens et al. (2020) have argued that transforming pressure to perform the interpolation is not of key importance for the final value of the Equitable Threat Score (ETS), a measure introduced to quantify the degree of agreement between in situ and reanalysis data.

- **Noise**. As put by Sabatini and Gardi (2023), the development and improvement of airport facilities, including their design and redesign, depend on the calculation and measurement of aircraft noise. There is actually a multitude of aircraft noise prediction models, each designed for specific purposes. Some authors such as Fil-

ippone (2014) have suggested distinguishing between *theoretical methods* that rely on a physical model of noise production and propagation and *best practice methods* that rely almost exclusively on measurement databases (fly-over or other measurements), which are augmented with other sub-models. Recently, the use of ADS-B data for the computation of noise around airports has been a research topic of great interest as shown by Pretto et al. (2022). As airports recognize the importance of sharing noise-related data, Gasco et al. (2017) have argued that aircraft noise predictions are becoming more accessible. The main media for this communication are noise maps, periodic reports, and systems for visualizing data from Noise Measurement Terminals (NMTs). As explained by Genescà et al. (2013), the placement of these NMTs is chosen so that the measured noise levels are representative of the acoustic influence of the airport on the population. Measured and predicted aircraft noise are regularly compared, for instance by Simons et al. (2022), Bendarkar et al. (2022), Huynh et al. (2022), Jäger et al. (2021), Arnone et al. (2023). For each NMT, the difference between modeled and measured values are usually computed (*local agreement*). Yet, a spatial (*global*) agreement is more complicated to get since, obviously, noise measurements are taken at a limited number of specific locations. In this regard, having noise maps based solely on data collected by the NMTs would be valuable in order to visualize the empirical spatial distribution of noise measurements. These interpolated values can then be compared with the output of a more comprehensive acoustic model. Moreover, acoustic research frequently employs geostatistical methods for noise mapping in urban areas. Typical examples are the works of Aumond et al. (2018) focusing on the XIIIth district of Paris and Tsai et al. (2009) focusing on Taiwan. Several interpolation methods have been compared by Harman et al. (2016) and Can et al. (2014).

While the contrail literature emphasizes the simplicity and good performance of linear interpolation, the scientific literature on noise pollution shows a preference for statistical methods. Linear interpolation is indeed simple to implement and easily generalizes to multiple dimensions. The quality of predictions is particularly good if the grid being interpolated is already of high resolution, which is often the case with reanalysis data

used in the literature on contrails. Yet, as early pinpointed by Myers (1994), *deterministic interpolation* and more specifically linear interpolation, is only one option among many others. In environmental sciences, a great number of other methods have recently been reviewed by Li and Heap (2014). Most of them are *stochastic*, as explained by Webster and Oliver (2007). One reason is the ability of the statistical framework to provide accurate predictions, to quantify uncertainties and most of all to enable the use of covariates. Hence, several questions arise.

What should be expected from the geostatistical approach for interpolating meteorological data in contrail studies? Why is the geostatistical approach particularly interesting to interpolate noise values in the vicinity of airports? What similarities exist between these two applications and when should the geostatistical framework be chosen?

The objectives of this work are of several kinds:

(a) Presenting a general statistical framework for the interpolation of aviation data, assumptions on which this framework is based, its implementation and limitations.
(b) Confirming that linear interpolation is a highly competitive method to interpolate fine resolution weather data. We compare linear interpolation to other methods for several weather variables: temperature, relative humidity, and the two components of horizontal wind.
(c) Confirming findings from the contrail literature, especially regarding the poor estimation of relative humidity values in regions for which air is supersaturated with respect to ice. The power-law correction to humidity fields introduced by Teoh et al. (2022) is discussed.

The paper is organized as follow. First, the data sources and considered samples for the contrail and noise case studies are presented in Section 2. The distinctions between deterministic interpolation methods and stochastic methods are elucidated in Section 3. The geostatistical framework for interpolation over a two-dimensional (2D) domain is introduced in Section 4. The 2D case is especially important for the noise case study. The three-dimensional (3D) case is developed in Section 5. The 3D case is particularly relevant for the contrail case study. Finally, Section 6 is devoted to the correction of humidity fields. This question, although seemingly peripheral, is directly related to the

reliability of the interpolated values. Without correcting relative humidity values beyond 100%, it is likely that even a good interpolation method would yield disappointing results.

2. Data sources and considered samples

2.1. The contrail case study

2.1.1. Trajectory data

The Measurement of Ozone and Water Vapour on Airbus In-service Aircraft (MOZAIC) programme started in 1994 (Marenco et al. (1998)). It was transferred into the European Research Infrastructure IAGOS (In-service Aircraft for a Global Observing System) in 2008 (Petzold et al. (2015)). IAGOS operates a global-scale monitoring system for atmospheric trace gases, aerosols and clouds utilising the existing global civil aircraft. As of 2020, the MOZAIC/IAGOS fleet has visited 330 airports. Due to the high frequency of the flights covered by IAGOS, data are highly representative of the altitude band and flight corridors frequented by passenger aircraft.

The focus is made on the IAGOS L2 time series data product, that is to say data that have been submitted to final quality control (level 2) (Boulanger et al. (2018)). One single file is provided for each flight. Data include both the position of the aircraft every 4 seconds (longitude, latitude, altitude) as well as the air temperature, relative humidity, and the two components of the wind.

IAGOS data are the ground truth we consider to compare interpolation methods. To put it differently, an interpolation method is said to be good if the interpolated values along a flight path are close (in a sense to be defined) to the measured values. IAGOS data are reliable enough to serve as ground truth in this context as assessed by Neis et al. (2015).

2.1.2. Weather data

ERA5 is the fifth generation European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis for the global climate and weather for the past four to seven decades (Hersbach et al. (2020)). Reanalysis combines model data with observations from across the world into a globally complete and consistent data set using the laws of

physics.

The focus is made on a data set named *ERA5 hourly data on pressure levels from 1940 to present*. In this data set, several weather variables are available on an hourly basis for 37 pressure levels on a $0.25° \times 0.25°$ longitude-latitude grid. Despite the fine spatial and temporal granularity of ERA5 data, they exhibit a number of well-documented limitations.

More specifically, in situ measurements of weather and modelled data may differ at the tropopause level. Differences between in situ measurements provided by IAGOS and reanalysis data of ERA-Interim (an old version of the ERA5 data set we consider) have been quantified by Reutter et al. (2020). Temperature values are found to be very similar as well as water vapour volume mixing ratio values. However, IAGOS water vapour volume mixing ratio values show a larger variability and stronger extreme values, which has a consequence on the values of relative humidity with respect to ice. Crucially, ERA-Interim and IAGOS behave differently when relative humidity with respect to ice exceeds 100% (ice supersaturated regions). This assessment is also made by Gierens et al. (2020). A review of existing studies that have identified the limitations of humidity fields provided by the ECMWF ERA5 product is given in the supplementary material of Teoh et al. (2022).

Limitations of humidity fields have a very important consequence on the comparison of interpolation methods: even a perfect interpolation based on ERA5 data would not be able to retrieve measured weather values (IAGOS). It is expected that interpolation errors will be more significant for relative humidity values than for temperature ones, and this, regardless of the quality of the interpolation. Section 6 details the right procedure to correct humidity fields. Without further clarification, the relative humidity referred to in this work is always the relative humidity with respect to ice.

2.1.3. Considered sample

The sample of interest for the contrail case study is made of 1,212 flights. The steps of data acquisition are briefly recalled in the following.

Flights are first downloaded from the IAGOS data portal (observational data access).

The query period is the year 2019. Flights from the IAGOS-CORE project are selected. Two variables are kept : relative humidity with respect to ice and relative humidity with respect to liquid water. The query area goes from $-180°$ to $180°$ in longitude and from $-90°$ to $90°$ in latitude. 2,290 flights are retrieved.

Out of these flights, we perform some additional subsetting. Using quality labels provided in IAGOS, we only use data with validity flag "0" which means that the measurement is reliable as stated by Gierens et al. (2020). For a given flight, if more than 90% of the points have validity flag "0", invalid points are dropped and the flight is retained. Otherwise, the entire flight is discarded. To avoid creating a data set for each weather variable, the criterion is considered for all four variables jointly. At the end of this procedure, there are 1,365 flights remaining.

For each remaining flight, corresponding ERA5 data are downloaded thanks to the Climate Data Store (CDS) Application Programming Interface (API) with a margin of 2 degrees of longitude and latitude. The flights for which this margin is not respected without coordinate transformations are excluded from the analysis for simplicity. At the end of this step, the final sample consists of 1,212 flights. To speed up the calculations, each flight is linearly interpolated to retain only 100 points. Records associated to water supersaturation are discarded to compare interpolation results.

2.2. The noise case study

The Chicago Department of Aviation's Airport Noise Management System (ANMS) is a comprehensive system to provide actual measurement of the aircraft noise levels in Chicago neighborhoods and suburban communities around O'Hare and Midway. This integrated system includes a network of about 40 permanent noise monitors that measure the noise environment and a system directly connected to the Federal Aviation Administration's (FAA) air traffic control radar that collects aircraft flight tracks.

Noise monitors record noise events based on threshold exceedance. Each noise event starts at the time the noise level exceeds a decibel threshold, typically slightly above the background or ambient noise level, and ends at the time the noise level returns to the threshold. Once the noise events are collected, they are correlated to actual aircraft

operations.

The focus is made on monthly reports made available by Chicago O'Hare International Airport. Our particular focus lies on the month of December 2022.

3. Mathematical framework

Deterministic and statistical methods share similarities when it comes to defining a spatial interpolation problem. This shared intuition is presented in Subsection 3.1. The geostatistical framework and its expected benefits are detailed in Subsection 3.2. Typical assumptions related to the geostatistical framework are stated in Subsection 3.3. The two case studies we consider are respectively presented in Subsection 3.4 and Subsection 3.5.

3.1. A shared intuition based on weighted averages

As stated by Webster and Oliver (2007), nearly all interpolation methods can be seen as weighted averages of data. Raw data come as a collection of n values denoted $\{z(\boldsymbol{s}_i), i = 1, ..., n\}$ over a *region of interest* hereinafter referred to as $D \subset \mathbb{R}^d$ (in this work, $d = 2$ or $d = 3$). Note that s_i is a location on D and $z(\boldsymbol{s}_i)$ is its associated value. To get the predicted value $z^*(\boldsymbol{s}_0)$ of an unknown location $\boldsymbol{s}_0$, the following formula is commonly used:

$$z^*(\boldsymbol{s}_0) = \sum_{i=1}^{n} \lambda_i z(\boldsymbol{s}_i). \tag{1}$$

Choosing an interpolation method boils down to choosing a procedure to compute the weights $\lambda_1, ..., \lambda_n$. Weights depend on $\boldsymbol{s}_0$ and the locations of known points. One basically predicts unobserved values from observed data for which the only exogenous variables are their spatial locations.

Bilinear interpolation is a good place to start when $D \subset \mathbb{R}^2$. Known points are typically sampled on a rectilinear grid. 4 known points are considered to describe the rectangle within which the interpolation takes place: $\boldsymbol{s}_{1,1}$ (bottom left corner), $\boldsymbol{s}_{1,2}$ (upper left corner), $\boldsymbol{s}_{2,1}$ (bottom right corner), $\boldsymbol{s}_{2,2}$ (upper right corner). Known points have coordinates $(x_1, y_1), (x_1, y_2), (x_2, y_1), (x_2, y_2)$. In between, the unknown location s_0

writes (x_0, y_0). The weights $\lambda_{1,1}, \lambda_{1,2}, \lambda_{2,1}, \lambda_{2,2}$ must satisfy the transposed linear system:

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ x_1 & x_1 & x_2 & x_2 \\ y_1 & y_1 & y_2 & y_2 \\ x_1y_1 & x_1y_2 & x_2y_1 & x_2y_2 \end{bmatrix} \begin{bmatrix} \lambda_{1,1} \\ \lambda_{1,2} \\ \lambda_{2,1} \\ \lambda_{2,2} \end{bmatrix} = \begin{bmatrix} 1 \\ x_0 \\ y_0 \\ x_0y_0 \end{bmatrix}. \tag{2}$$

At the cost of increased computational complexity, *bicubic interpolation* gives an interpolated surface that is smoother than corresponding surfaces obtained by bilinear interpolation. More generally, known points may be the vertices of a convex quadrilateral. Linear interpolation can be performed in higher dimensions. When points are scattered, one may triangulate the set of known points. Once the triangle containing the point to be interpolated is identified, it is possible to perform linear interpolation. It involves a calculation of barycentric coordinates.

When known points are not located on a rectilinear grid, methods based on inverse functions of distance are also very popular (abbreviated IDW for Inverse Distance Weighting). In this setting, weights are defined by:

$$\lambda_i = \frac{1}{\text{dist}(\boldsymbol{s}_i, \boldsymbol{s}_0)^\beta}, \ \beta > 0, \tag{3}$$

and are scaled so that they sum to 1. Data points near to $\boldsymbol{s}_0$ carry larger weight than those further away. When β is large, the relative weights diminish rapidly as the distance increases, and so the interpolation is sensibly local. Interpolation is exact. Because the weights never become zero, there are no discontinuities.

Deterministic methods are rightly appreciated for their simplicity and their good practical results. Yet, for each method, spatial dependence is considered in a rigid manner. It is also impossible to associate a degree of certainty with each interpolated value. No covariate is taken into account. These reasons sometimes lead to preferring statistical methods to deterministic ones.

3.2. The geostatistical framework and its expected benefits

Geostatistical methods assume that $\{z(\boldsymbol{s}_i), i = 1, ..., n\}$ is a collection of *regionalized values*. A regionalized value is the outcome of a random mechanism. Each location $\boldsymbol{s}$

on D is associated to the realisation $z(\boldsymbol{s})$ of a random variable $Z(\boldsymbol{s})$. Values are said to be regionalized because they exhibit some spatial correlation. A family of real-valued random variables $\{Z(\boldsymbol{s}), \boldsymbol{s} \in D\}$ is traditionally called a *spatial random field*.

In the sequel, we assume that the first moment as well as the usual second-order moments of the random fields of interest are well-defined. In other words, the *expectation* of the random field is defined as a non-random function of s that coincides at each point with the expectation of the random variable at that point:

$$\mu(\boldsymbol{s}) = \mathbb{E}(Z(\boldsymbol{s})). \tag{4}$$

The *covariance function* of the random field is defined as a non-random function of $\boldsymbol{s}_i$ and $\boldsymbol{s}_j$. For any pair of values $(\boldsymbol{s}_i, \boldsymbol{s}_j)$, it coincides with the covariance between the random variables at those two points:

$$C(\boldsymbol{s}_i, \boldsymbol{s}_j) = Cov(Z(\boldsymbol{s}_i), Z(\boldsymbol{s}_j)) = \mathbb{E}\left[(Z(\boldsymbol{s}_i) - \mu(\boldsymbol{s}_i))(Z(\boldsymbol{s}_j) - \mu(\boldsymbol{s}_j))\right], \ \forall \boldsymbol{s}_i, \boldsymbol{s}_j \in D. \tag{5}$$

Kriging refers to predicting the value of a non-observed point $\boldsymbol{s}_0$ of a random field Z with a linear predictor. The general idea is (unsurprisingly) to produce a weighted average:

$$Z^*(\boldsymbol{s}_0) = \sum_{i=1}^{n} \lambda_i Z(\boldsymbol{s}_i). \tag{6}$$

The weighting is found ensuring that the expected prediction error is zero (unbiasedness of the kriging predictor) and its variance minimum.

The theoretical benefits of kriging are well known. Kriging takes into account the geometric characteristics, the number and organization of locations. It considers the structure of the spatial correlation that is deduced from known points. The weights kriging uses are not calculated on the basis of an arbitrary rule that can be used in some cases but not others, but rather on the behavior of the function that represents the structure of spatial correlation. Kriging makes it possible to quantify how accurate are the predictions using the prediction error variance. Interpolation is exact and covariates may be used.

The framework for inference in geostatistics is not the same as the one of classical multivariate statistics. Instead of having the realization of a random sample of inde-

pendent and identically distributed random variables, there are no multiple independent realizations of the same field. Assumptions are necessary.

3.3. Typical assumptions

The most famous assumption in geostatistics is the *second-order stationary* hypothesis. The hypothesis is as follows:

(a) The expectation exists and is constant, and therefore does not depend on the location s: $\mu(\boldsymbol{s}) = \mu$.
(b) The covariance exists for every pair of random variables $Z(\boldsymbol{s})$ and $Z(\boldsymbol{s}+\boldsymbol{h})$, and only depends on the vector $\boldsymbol{h}$ that joins the locations $\boldsymbol{s}$ and $(\boldsymbol{s}+\boldsymbol{h})$, but not specifically on them: $C(Z(\boldsymbol{s}), Z(\boldsymbol{s}+\boldsymbol{h})) = C(\boldsymbol{h}),\ \forall \boldsymbol{s} \in D,\ \forall\ \boldsymbol{h} \in \mathbb{R}^d$ such that $\boldsymbol{s}+\boldsymbol{h} \in D$.

When the covariance function only depends on the distance between the locations $\boldsymbol{s}$ and $\boldsymbol{s}+\boldsymbol{h}$, it is called *isotropic*. When it depends on both the distance and direction of vector $\boldsymbol{h}$, it is described as *anisotropic*.

In geostatistics, it is common to define the *variogram*, that is, the variance of the first differences of the random field:

$$2\gamma(\boldsymbol{s}_i - \boldsymbol{s}_j) = \mathbb{V}(Z(\boldsymbol{s}_i) - Z(\boldsymbol{s}_j)),\ \forall \boldsymbol{s}_i, \boldsymbol{s}_j \in D. \tag{7}$$

The function γ is called the *semivariogram*. As this function is all that is needed for modeling and making spatial predictions. Some authors have called γ a variogram which may be confusing.

In the case of second-order stationarity, the covariance function and the semivariogram are equivalent when it comes to defining the structure of spatial dependence displayed by the phenomenon. If the hypothesis holds, $\gamma(\boldsymbol{h}) = C(0) - C(\boldsymbol{h})$. It is convenient from a mathematical standpoint but unrealistic in many practical applications.

A very important case in which the second-order stationarity assumption doesn't hold is the so-called *drift* situation in which the mean of the random field depends on location. Difficulties associated with the drift case are addressed in the sequel and are of key important for both case studies.

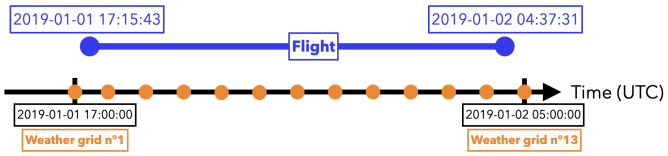

Fig. 1. Adding weather data to this flight involves several weather grids.

A relaxed version of the second-order stationarity hypothesis is also commonly used: *intrinsic stationarity*. A random field is said to be intrinsically stationary (or simply intrinsic) if, for any given translation vector $\boldsymbol{h}$, the first-order increments $Z(\boldsymbol{s}+\boldsymbol{h})-Z(\boldsymbol{s})$ are second-order stationary.

Often, it is enough to estimate spatial dependence in the neighborhood of the points for which we need a prediction. It is possible to relax the second-order stationarity hypothesis (respectively the intrinsic hypothesis). In this view, *quasi-stationarity* is defined by Journel and Huijbregts (2004) to be a second-order (resp. intrinsic) stationarity hypothesis that holds within a neighbourhood.

3.4. Full problem statement for weather interpolation

For a given flight, the position of the aircraft is recorded for a finite set of observation points (IAGOS data are presented in Subsection 2.1.1). The goal is to match each of these points with a temperature value, a relative humidity value and wind values (ERA5 data are presented in Subsection 2.1.2). We need to interpolate a three-dimensional weather grid that is available every hour. As the duration of each flight spans over several hours, the interpolation problem involves multiple weather grids as schematized in Figure 1.

The complete problem is a spatio-temporal interpolation task. Since time is not the focus of this work, temporal interpolation is simplified to the fullest extent. For each point, the closest ERA5 grid in time is retained, and various spatial interpolation methods are compared.

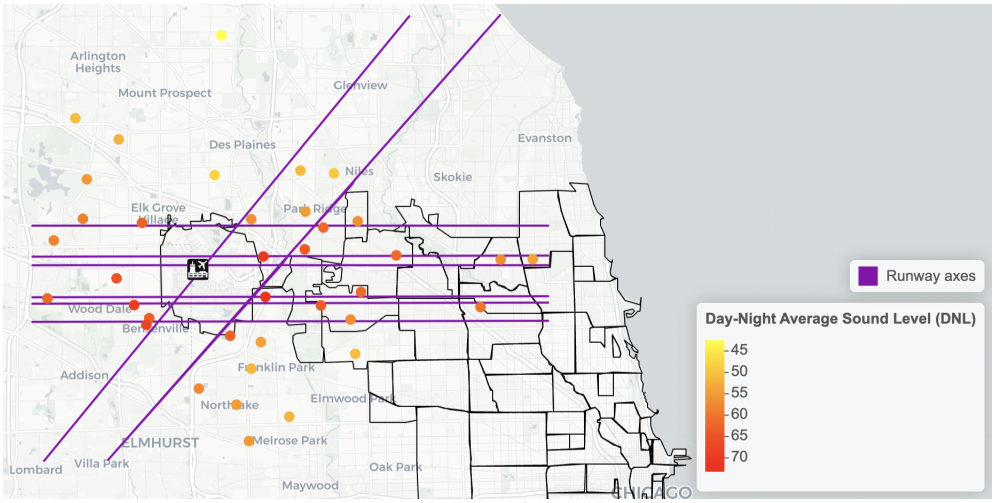

Fig. 2. Location of noise monitors in the vicinity of Chicago O'Hare International Airport. The data presented summarizes the Day-Night Average Sound Levels (DNLs) in December 2022. Current community area boundaries in Chicago are reported in black. Runway axes are in violet.

3.5. Full problem statement for noise interpolation

Noise monitors are located on the ground, in the vicinity of Chicago O'Hare International Airport (noise data are presented in Subsection 2.2). Here, the objective is to interpolate noise values around the airport, involving both interpolation and extrapolation. Both are two-dimensional.

4. Interpolation over a two-dimensional (2D) domain

Interpolation over a two-dimensional (2D) domain is mostly relevant for the noise case study. In this section, the *region of interest* D is two-dimensional. The locations of noise monitoring stations are shown in Figure 2.

4.1. Characterizing spatial dependence

Stochastic interpolation methods require estimating the spatial dependence between observations. In practice, the empirical (or experimental) semivariogram is the instrument used to estimate the structure of spatial variability existing in the phenomenon of interest. The most famous estimator of the semivariogram is the Matheron's method of

moments estimator:

$$\hat{\gamma}(\boldsymbol{h}) = \frac{1}{2\#N(\boldsymbol{h})} \sum_{N(\boldsymbol{h})} [Z(\boldsymbol{s}_i + \boldsymbol{h}) - Z(\boldsymbol{s}_i)]^2 \tag{8}$$

where $Z(\boldsymbol{s}_i)$ are the values of the characteristic of interest at the points $\boldsymbol{s}_i$ when we also know the value of this characteristic at $\boldsymbol{s}_i + \boldsymbol{h}$, and $\#N(\boldsymbol{h})$ is the number of pairs of locations separated by a vector $\boldsymbol{h}$ (the *lag*). In practice, nothing can be said about the variogram at lag distances smaller than the minimum distance.

4.2. Structural analysis

A given empirical semivariogram may not meet the require theoretical properties of a valid semivariogram. An important one is *conditional negative-definiteness*. The absence of this property can result in negative mean-squared errors of prediction. The so-called *structural analysis* step then concerns the fitting of a valid model to the empirical semivariogram. This step is necessary to make accurate spatial predictions.

Fitting a semivariogram model can be done by eye or using statistical procedures. A common practice is first to manually choose the models that best capture the main features of the empirical semivariogram and then to perform statistical fitting. Statistical fitting is usually based on least squares methods or maximum likelihood. Note that a Gaussian assumption is crucial for the maximum likelihood estimation procedure.

4.3. Kriging

In the kriging procedure, the weighting (see Equation 6) is found ensuring that the expected prediction error must be zero (unbiasedness of the kriging predictor) and its variance minimum. As a result, the mean-squared prediction error is minimized (spatial prediction is said to be *optimal* in this sense). Unbiasedness translates into the linear constraint $\sum_{i=1}^{n} \lambda_i = 1$.

Simple Kriging (SK) is the special case of kriging a second-order stationary random field with a known mean. Kriging a second-order stationary random field with an unknown constant mean is called *Ordinary Kriging* (OK). *Universal Kriging* (UK) deals with the non-stationary (in mean) case and is developped in Subsection 4.4. The OK

problem is solved by the method of Lagrange multipliers. The optimization goes:

$$\min_{\lambda_1,\ldots,\lambda_n,\alpha} \mathbb{V}\left[Z^*(\boldsymbol{s}_0) - Z(\boldsymbol{s}_0)\right] - \alpha\left(\sum_{i=1}^{n}\lambda_i - 1\right), \tag{9}$$

α being the Lagrange multiplier. This results in a system of $n+1$ equations. The system of equations can be presented in matrix form:

$$\begin{pmatrix} \gamma(\boldsymbol{s}_1-\boldsymbol{s}_1) & \gamma(\boldsymbol{s}_1-\boldsymbol{s}_2) & \cdots & \gamma(\boldsymbol{s}_1-\boldsymbol{s}_n) & 1 \\ \gamma(\boldsymbol{s}_2-\boldsymbol{s}_1) & \gamma(\boldsymbol{s}_2-\boldsymbol{s}_2) & & \gamma(\boldsymbol{s}_2-\boldsymbol{s}_n) & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \gamma(\boldsymbol{s}_n-\boldsymbol{s}_1) & \gamma(\boldsymbol{s}_n-\boldsymbol{s}_2) & & \gamma(\boldsymbol{s}_n-\boldsymbol{s}_n) & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ \alpha \end{pmatrix} = \begin{pmatrix} \gamma(\boldsymbol{s}_0-\boldsymbol{s}_1) \\ \gamma(\boldsymbol{s}_0-\boldsymbol{s}_2) \\ \vdots \\ \gamma(\boldsymbol{s}_0-\boldsymbol{s}_n) \\ 1 \end{pmatrix}.$$

The minimized mean-squared prediction error is the *kriging variance*. Prediction intervals can be constructed from the kriging variance with an additional Gaussian assumption. The effect of estimating the semivariogram parameters on the mean-squared prediction error is often unaccounted for.

4.4. Drift

In geostatistics, a drift or trend may be long-range or short-range but is viewed as a smooth systematic non-random variation. A common practice is to break down the random field into the sum of two components:

$$Z(\boldsymbol{s}) = \mu(\boldsymbol{s}) + \varepsilon(\boldsymbol{s}) \tag{10}$$

where $\mu(\boldsymbol{s})$ denotes the deterministic part of the random field (the drift, that is unknown) and $\varepsilon(\boldsymbol{s})$ the stochastic part that can be treated as second-order stationary. This *trend-fluctuation-noise decomposition* is not unique.

As stated by Hristopulos (2020), trend estimation methods can be *empirical* or *process-based*. In the former approach, the trend function is determined from general knowledge of the process or the exploratory analysis of the data. In the latter, the functional form of the trend and possibly the coefficients of the trend function are inferred from a model.

In geostatistics, one would generally assumes that both macro and micro variations are important. As a consequence, *UK* was proposed by Matheron in 1969. For UK, the deterministic component writes:

$$\mu(\boldsymbol{s}) = \sum_{j=1}^{p} a_j f_j(\boldsymbol{s}) \tag{11}$$

where $f_1, ..., f_p$ are p known functions and $a_1, ..., a_p$ unknown coefficients (or parameters). Functions are generally polynomials in the spatial coordinates. As an example, a basic trend for $D \subset \mathbb{R}^2$ would be written:

$$\mu(\boldsymbol{s}) = a_1 + a_2 x + a_3 y + a_4 x^2 + a_5 y^2 \tag{12}$$

for $\boldsymbol{s} = \begin{pmatrix} x & y \end{pmatrix}^\top \in D$. In the UK approach, the stochastic part is assumed to be a zero-mean intrinsically stationary random process with known variogram 2γ. The UK system is then no more than an augmentation of the OK system. In practice, the variogram must be estimated because it is not known. When external variables are considered instead of or additional to functions of spatial coordinates, UK is called *Kriging with External Drift* (KED). Formally, any external variable is treated as a deterministic function assumed known everywhere in the domain of interest.

When the trend is particularly complex (which is commonly the case with weather data), other methods may be used. A recent one called Spatial Regression with Partial Differential Equation regularization (SR-PDE) is introduced in Subsection 4.6.

4.5. Anisotropy

The hypothesis of isotropy remarkably simplifies the modeling of spatial dependence. Yet, this hypothesis is violated when the empirical semivariogram depends on the direction of $\boldsymbol{h}$. In this case, there is an underlying physical process evolving differentially in space to take into account. Geometric anisotropy is the only case for which isotropy can be restored with a simple coordinate transformation. Conceptually, geometric anisotropy is obtained by some stretching of an isotropic model. Speaking in terms of semivariogram, geometric anisotropy is characterized by:

$$\gamma(\boldsymbol{h}) = \gamma_{\text{iso}}(\|\boldsymbol{A}\boldsymbol{h}\|_2)$$

where the matrix $\boldsymbol{A}$ defines the transformation from the initial space to the isotropic space. A linear transformation of the coordinates in enough to solve the anisotropy problem. As put by Chilès and Delfiner (2012), in $\mathbb{R}^2$, denoting by θ_1 and $\theta_1 + \frac{\pi}{2}$ the main directions of anisotropy, a rotation θ_1 defines a new coordinate system with axes parallel to the anisotropy directions. The matrix A is then:

$$\boldsymbol{A} = \underbrace{\begin{pmatrix} b_1 & 0 \\ 0 & b_2 \end{pmatrix}}_{\boldsymbol{T}} \underbrace{\begin{pmatrix} \cos(\theta_1) & \sin(\theta_1) \\ -\sin(\theta_1) & \cos(\theta_1) \end{pmatrix}}_{\boldsymbol{R}},$$

where $\boldsymbol{T}$ is matrix of scaling factors and $\boldsymbol{R}$ a rotation matrix. Empirically, the *variogram map* can be used to estimate the anisotropy parameters. To estimate a variogram map, variogram values are computed for a grid of separation vectors creating square cells.

4.6. A more advanced framework

In many applications related to the interpolation of meteorological or industrial data, the spatial variations of the phenomenon of interest are highly complex. As stated by Sangalli (2021), it may be typically due to:

- The complex physics of the phenomenon under study (for instance, the velocity field of blood flow in human arteries)
- An external source that generates strong anisotropies and non-stationarities in the observed quantity of interest (for instance, prevailing winds play a huge role in environmental and climate data)
- The complicated conformation of the planar domain where the data are observed (for instance, a domain with holes or with a strong concavity).
- Non-planar domains (for instance, the cerebral cortex)

Relative humidity and noise may typically involve a complex physics. Yet, the domain is planar and relatively simple. The approach introduced by Sangalli (2021) is called Spatial Regression with Partial Differential Equation (SR-PDE) regularisation.

Unlike the classic geostatistical approach, whose common assumptions are given in Subsection 3.3, the SR-PDE approach assumes that spatial field is deterministic. The spatial structure of the phenomenon is modelled via a PDE in a regularising term.

Let $\boldsymbol{w}_i = (w_{i1}, ..., w_{iq})^\top \in \mathbb{R}^q$ be q covariates observed at $\boldsymbol{s}_i$. We consider the model:

$$z(\boldsymbol{s}_i) = \boldsymbol{w}_i^\top \boldsymbol{\beta} + f(\boldsymbol{s}_i) + \varepsilon(\boldsymbol{s}_i), \ i = 1, ..., n, \tag{13}$$

where $\boldsymbol{\beta} \in \mathbb{R}^q$ is a vector of unknown regression coefficients, $f : \mathcal{D} \to \mathbb{R}$ is an unknown deterministic field that captures the spatial structure of the phenomenon under study and $\varepsilon(\boldsymbol{s}_1), ..., \varepsilon(\boldsymbol{s}_n)$ are uncorrelated errors with zero mean and finite variance. Sangalli et al. (2013) proposed to estimate the vector $\boldsymbol{\beta}$ and f by minimising the following regularised sum-of-square-error functional:

$$\sum_{i=1}^{n} \left[z(\boldsymbol{s}_i) - \boldsymbol{w}_i^\top \boldsymbol{\beta} - f(\boldsymbol{s}_i) \right]^2 + \lambda \int_{\mathcal{D}} (\Delta f)^2 \, d\boldsymbol{s}, \tag{14}$$

where λ is a positive smoothing parameter and Δ the Laplace operator. The Laplace operator provides a simple measure of the local curvature of the field f. The functional is well defined for $\boldsymbol{\beta} \in \mathbb{R}^q$ and $f \in H^2(\mathcal{D})$, where $H^2(\mathcal{D})$ is the Sobolev space of functions $g : \mathcal{D} \to \mathbb{R}$ such that g and its first and second derivatives are in $L^2(\mathcal{D})$. It is assumed that $\mathcal{D}$ has boundary $\partial\mathcal{D} \in \mathcal{C}^2$.

When the domain is bounded, the use of appropriate boundary conditions guarantees the uniqueness of the solution (Sangalli et al. (2013), Azzimonti et al. (2014)). We denote by $V(\mathcal{D})$ the subspace of $H^2(\mathcal{D})$ with the chosen boundary conditions.

Numerical discretisation procedures are used because there is no analytical solution. The numerical discretisation reduces the estimation problem to the solution of a linear system. The spatial domain $\mathcal{D}$ is represented by an appropriate mesh, and the functions over $\mathcal{D}$ are approximated by a finite system of bases defined on this mesh. Convenient meshes of the spatial domain are typically obtained by constrained Delaunay triangulation when the planar domain is complex.

4.7. Results for the noise case study

We compare four approaches to interpolate noise measurements:

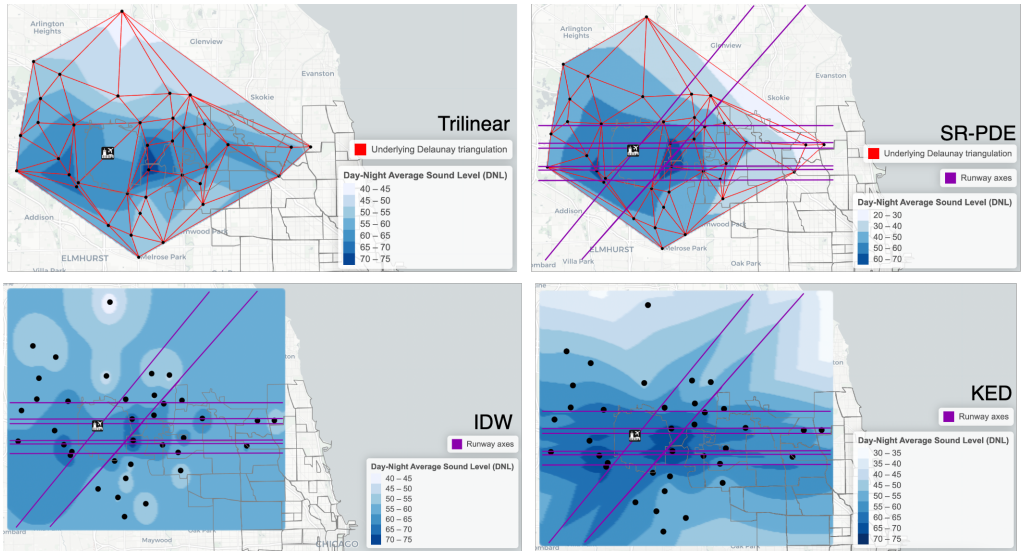

Fig. 3. Interpolation of noise measurements in the vicinity of Chicago O'Hare International Airport (December 2022) using Trilinear interpolation [upper left], IDW [lower left], KED [lower right], and SR-PDE [upper right].

- Linear interpolation based on Delaunay triangulation
- IDW interpolation with the Euclidean distance and $\beta = 2$ (see Equation 3)
- KED (see Subsection 4.4) with a drift given by:

$$\mu(\boldsymbol{s}) = a_1 + a_2 x + a_3 y + a_4 xy + a_5 \left\|\boldsymbol{s} - \boldsymbol{s}^{\text{air}}\right\|_2 + a_6 \min_{\boldsymbol{s}^{\text{run}} \in \mathcal{R}} \left\|\boldsymbol{s} - \boldsymbol{s}^{\text{run}}\right\|_2 , \qquad (15)$$

  where x is the longitude, y the latitude, $\boldsymbol{s}^{\text{air}}$ the airport location, and $\mathcal{R}$ the union runway axes. There is no apparent issue of anisotropy.

- SR-PDE (Subsection 4.6) with two covariates: the distance to the airport and the distance to the closest runway axis. The smoothing parameter λ is selected using generalised cross-validation introduced by Craven and Wahba (1978).

Resulting noise maps are depicted in Figure 3. When possible, noise extrapolation is performed.

In light of the obtained noise map, it is evident that trilinear interpolation, as well as the interpolation obtained through SR-PDE, are highly dependent on the underlying triangulation. The noise levels obtained consist of broken lines that do not correspond to

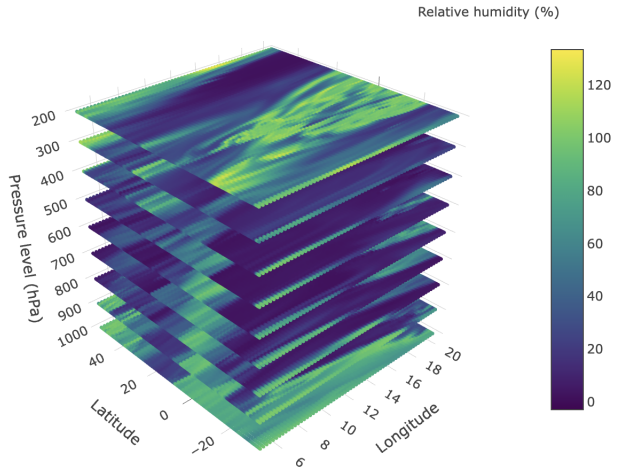

Fig. 4. Weather grid with relative humidity values on 2019-01-01 00:00:00.

a credible diffusion of noise from the acoustical perspective. The inclusion of covariates in the SR-PDE approach is not sufficient to achieve satisfactory interpolation, even though the smoothing parameter value selected by GCV is very low. The interpolation obtained through IDW is not credible either. By design, it does not consider the distance to the airport, which would allow for concentric noise levels centered around the airport. Crucially, kriging with external drift provides a very satisfactory noise map as it takes into account the distance to the nearest runway axis and the distance to the airport.

5. Interpolation over a three-dimensional (3D) domain

Interpolation over a three-dimensional (3D) domain is especially suitable for the weather case study. Every hour, raw weather data are given on a three-dimensional regular grid. A weather grid example is shown in Figure 4. For weather data, two spatial coordinates and a third dimension are available. The third dimension is the pressure level in hectopascal (hPa).

In order to simplify distance calculations, longitude and latitude values are transformed using the widespread Web Mercator projection. To use the Euclidean distance, it is necessary to transform the pressure level. To go from a pressure level p to an altitude alt in meters (m), the following formula is provided by the National Oceanic and

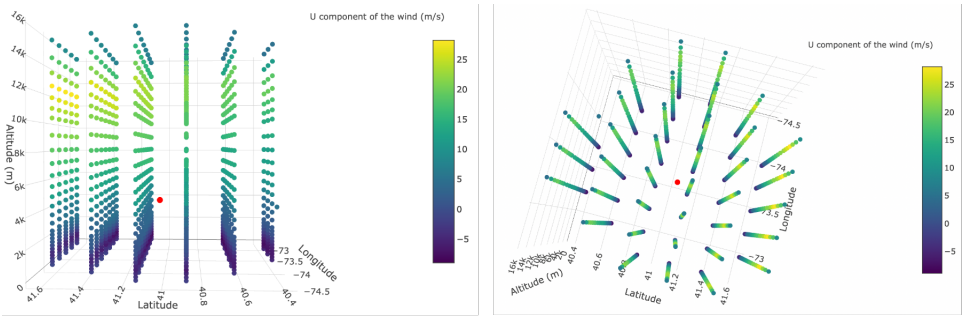

Fig. 5. Representation of the neighborhood of a trajectory point, shown in red, for which we want to predict the values of temperature, relative humidity, and wind. Two views are provided: from the side [left] and from the top [right].

Atmospheric Administration (NOAA):

$$\text{alt} = \frac{145366.45\left[1 - \left(\frac{p}{1013.25}\right)^{0.190284}\right]}{3.281}.$$

This formula is based on the International Standard Atmosphere (ISA).

In the case of meteorological data, the area of interest is particularly vast and covers both oceanic and continental zones. It would be at best ambitious, at worst erroneous, to assume second-order stationarity, given that spatial dependence is likely to change significantly across the entire domain. Therefore, for each grid, a neighborhood of 1,000 points is considered around the point of interest in the spirit of quasi-stationarity defined in Subsection 3.3. Figure 5 illustrates the neighborhood of a trajectory point.

Handling three dimensions follows essentially the same framework as in the two-dimensional case. Defining an equation for the three-dimensional trend does not present any specific challenges. Yet, incorporating anisotropy introduces a higher level of complexity because defining a rotation in $\mathbb{R}^3$ requires more parameters than in $\mathbb{R}^2$.

5.1. Anisotropy when $D \subset \mathbb{R}^3$

In the three-dimensional case, the matrix $\boldsymbol{A}$ can be written:

$$\boldsymbol{A} = \begin{pmatrix} b_1 & 0 & 0 \\ 0 & b_2 & 0 \\ 0 & 0 & b_3 \end{pmatrix} \begin{pmatrix} \cos(\theta_3) & \sin(\theta_3) & 0 \\ -\sin(\theta_3) & \cos(\theta_3) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta_2) & \sin(\theta_2) \\ 0 & -\sin(\theta_2) & \cos(\theta_2) \end{pmatrix} \begin{pmatrix} \cos(\theta_1) & \sin(\theta_1) & 0 \\ -\sin(\theta_1) & \cos(\theta_1) & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

In geostatistics, the primary focus is on the specific case where the principal axis of anisotropy is the vertical axis. In this case, $\theta_2 = \theta_3 = 0$. From our case study perspective, vertical anisotropy accounts for the fact that meteorological variability evolves differently from one pressure level to another.

5.2. SR-PDE when $D \subset \mathbb{R}^3$

The SR-PDE approach has been extended by Arnone et al. (2023) to nonconvex 3D domains. As a consequence, tetrahedral meshes are used to approximate the 3D domain of interest.

5.3. Results for the contrail case study

Multiple interpolation methods are being compared with the ultimate goal of assigning temperature, relative humidity, and wind values to every point along the trajectories presented in Subsection 2.1.3. As these flights are chosen in a manner that provides onboard measured values, it becomes feasible to quantify the difference between the measured value and the predicted value. Figure 6 depicts the discrepancies between the measured and predicted values for each meteorological variable across the sample of 1,212 flights. As expected, interpolation errors are larger for relative humidity (see Subsection 2.1.2).

Regardless of the variable of interest, the several versions of IDW interpolation do not yield good results. Basic geostatistical methods fail to accurately capture the trend which explains significant discrepancies from the reference values. As a result, predictions made by Ordinary Kriging (OK) and Universal Kriging (UK) are not satisfactory even if vertical anisotropy is considered. The best results are provided by trilinear interpolation and SR-PDE interpolation. The performance of these two methods can be explained by the fine spatial granularity of ERA5 reanalysis data and the absence of outliers. Temperature values are predicted much more accurately than relative humidity values, as observed in the existing literature on contrails (Subsection 2.1.2). One can illustrate the good results of these two methods with a specific flight (Figure 7). The SR-PDE method is conceptually more complex than trilinear interpolation. In the specific case of ERA5

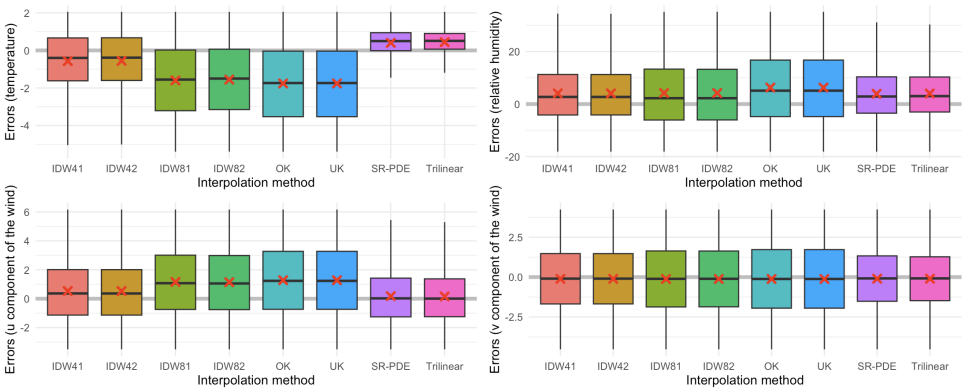

Fig. 6. For each variable of interest, boxplots depict the discrepancies from the measured values for each interpolation method. A positive difference indicates that the reference value is greater than the predicted value. The mean is indicated by the red cross. IDW41 refers to an IDW interpolation restricted to four nearest neighbors with β=1 (see Equation 3). OK refers to Ordinary Kriging, UK to Universal Kriging, and SR-PDE to the Spatial Regression with Partial Differential Equation regularization interpolation.

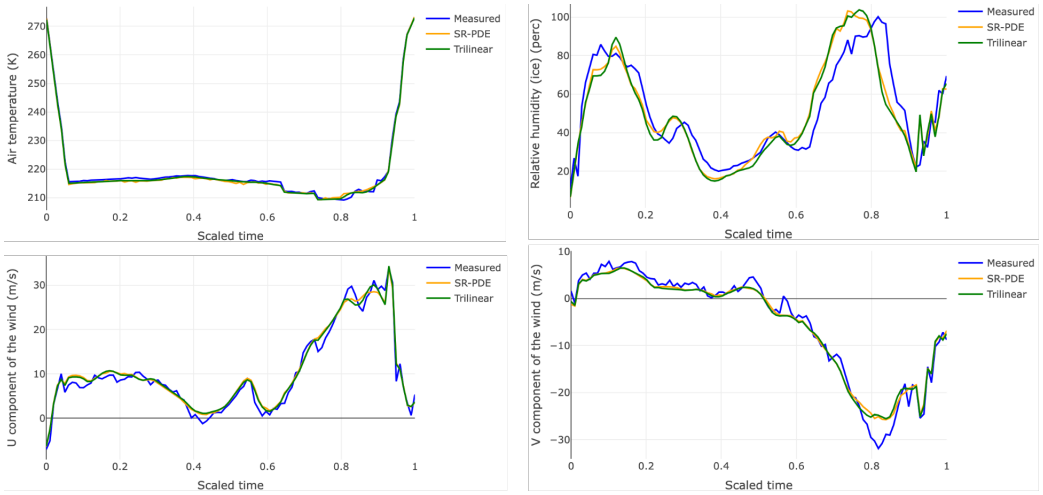

Fig. 7. For each variable of interest and for a specific flight, measured and predicted values.

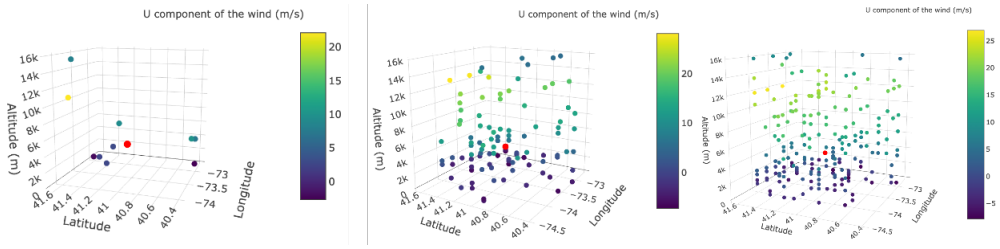

Fig. 8. Representation of the neighborhood of a trajectory point (in red) with different missing rates. Only 10 points [left], 100 points [middle], or 200 points [right] are remaining.

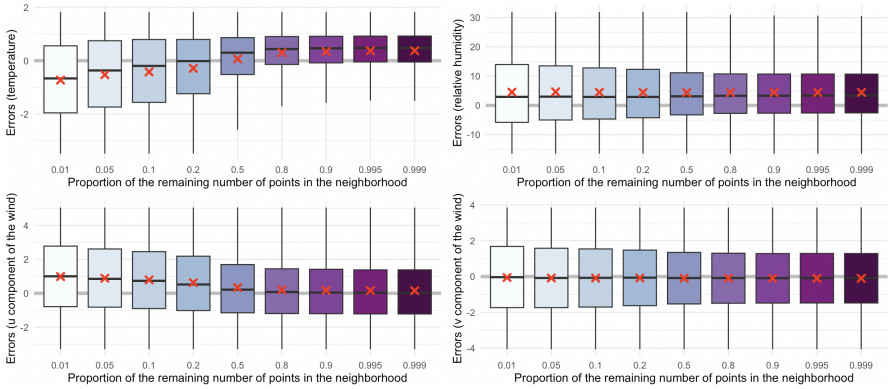

Fig. 9. For each variable of interest, boxplots depict the discrepancies from the measured values for each missing proportion.

reanalysis data, its use is of limited interest. However, as soon as meteorological data are scattered, it is clear that interpolation using SR-PDE becomes interesting because it offers more flexibility than linear interpolation on a mesh. To illustrate this point, one can calculate the prediction error when meteorological data become less abundant. Starting from the complete neighborhood (1,000 points), a certain proportion of points is randomly removed. This proportion ranges from 0.01 (almost all points remaining) to 0.99 (only 10 points remaining). Visualizations are provided in Figure 8. As seen in Figure 9, the number of points remaining in the neighborhood does not have the same effect on the prediction error depending on the variable of interest. For each variable, a higher density of points in the neighborhood helps reducing the interquartile range. For temperature values, it seems that the median error shifts from negative to positive, indicating that the interpolation of ERA5 data tends to very slightly underestimate the

measured temperature when the point density is high. Regarding relative humidity, the effect of a large number of points is primarily to reduce the interquartile range. For the u component of the wind, the median error decreases and stabilizes around zero as the number of points increases in the neighborhood. For the v component of the wind, it seems that the number of points has no significant effect on the error.

6. Correction to ERA5 humidity fields

The humidity fields of ERA5 products have limitations, the main one being that relative humidity values are underestimated in ice supersaturated regions (see Subsection 2.1.2). As Teoh et al. (2022) pointed out, there are several ways to address this issue in the literature on contrails (see Supplement, Section 3). Previous studies that used contrail cirrus prediction models uniformly increased humidity fields by dividing them with an enhancement factor that is strictly less than 1. This approach is not ideal as it may artificially increase the coverage of ice supersaturated regions. Additionally, the adjusted distribution of relative humidity values may not align with one derived from in-situ measurements, if available. To overcome these limitations, Teoh et al. (2022) have developed a new correction. The objective of this new approach is twofold: to better predict the location and extent of ice supersaturated regions and to better predict relative humidity values within these regions. To determine if the objective is achieved, ice supersaturated regions occurrence is measured from the 2019 IAGOS campaign and compared to the one derived from ERA5. The Equitable Threat Score (ETS) is also provided based on Appendix A of Gierens et al. (2020). Given a location $\boldsymbol{s}_i$, the corrected relative humidity value, denoted $\widetilde{z}(\boldsymbol{s}_i)$, is given by:

$$\widetilde{z}(\boldsymbol{s}_i) = \begin{cases} \frac{z(\boldsymbol{s}_i)}{a} & \text{if } \frac{z(\boldsymbol{s}_i)}{a} \leq 1 \\ \min\left\{\left(\frac{z(\boldsymbol{s}_i)}{a}\right)^b, \text{UB}\right\} & \text{if } \frac{z(\boldsymbol{s}_i)}{a} > 1, \end{cases} \tag{16}$$

where $z(\boldsymbol{s}_i)$ is the original relative humidity value (from ERA5), a and b two coefficients and UB a chosen upper bound (based on IAGOS data). Coefficients a and b are chosen to minimize a Cramér–von Mises criterion. To be more precise, the criterion formula proposed by Teoh et al. (2022) is the famous Cramér-von Mises test statistic used to

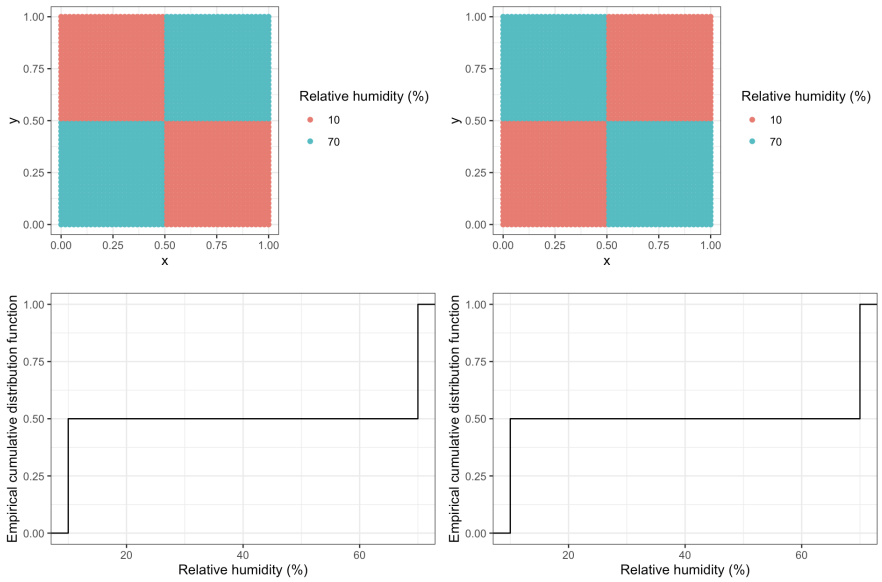

Fig. 10. A simple example where two humidity fields would yield the same empirical distribution function.

perform a goodness-of-fit test.

It is misleading, if not ill-suited to the objectives, to use this formula, and this is for three main reasons.

- It is straightforward to produce a scenario in which two distinct humidity fields yield two empirical distributions that are identical. An illustration of such a scenario is given in Figure 10. Ensuring that the two empirical distribution functions align is not sufficient for the spatial agreement of the two fields. In practice, the limited number of IAGOS flights does not allow for the adoption of more ambitious methods. Indeed, to compare spatial distributions of relative humidity for a given area at a given time, it would be necessary to have a sufficient number of IAGOS flights in that area, which is not the case. Since relative humidity is quite variable, it is not reasonable to group IAGOS flights temporally to obtain more measurement points in a given area. As a consequence, the adequacy of empirical cumulative distribution functions meets practical imperatives, although it would be preferable to adopt another method theoretically.

- More worryingly, the test statistic of Teoh et al. (2022) does not correspond to the

correct statistical framework. Let $X_1, ..., X_n$ be a sample of random variables that are independent and identically distributed according to some unknown cumulative distribution function F. A goodness-of-fit test is used if the problem is to test that F is actually a reference distribution (say, a normal or an exponential distribution). However, regarding relative humidity values, the reference distribution is estimated from IAGOS data. In other words, it is not known a priori. The framework of two-sample problems must be used in this case (Gibbons and Chakraborti (2010)).

Let X and Y be two populations with cumulative distribution functions F_X and F_Y, respectively. We have a random sample of size N drawn from the X population and another random sample of size M drawn independently from the Y population. The hypothesis of interest in the two-sample problem is that the two samples are drawn from identical populations:

$$H_0 : \forall x,\ F_X(x) = F_Y(x). \tag{17}$$

The most general two-sided alternative is:

$$H_A : \exists x,\ F_X(x) \neq F_Y(x). \tag{18}$$

In this framework, the Cramér–von Mises criterion is quite different, as early demonstrated by Anderson (1962). Let r_i and s_j be the ranks in the pooled sample of ordered observations of the first and second sample. The Cramér–von Mises criterion is:

$$T = \frac{1}{(N+M)^2} \left\{ \frac{N}{M} \sum_{i=1}^{N} \left(r_i - \frac{N+M}{N} i \right)^2 + \frac{M}{N} \sum_{j=1}^{M} \left(s_j - \frac{N+M}{M} j \right)^2 \right\}. \tag{19}$$

- The use of the Cramér–von Mises criterion is interesting from the perspective of statistical inference because it is possible determine its distribution under the null hypothesis with some assumptions. A key assumption is independence. Assuming that downsampling IAGOS data is sufficient to limit the temporal correlation, it remains that relative humidity values are highly spatially correlated (that is why it is even possible to perform spatial interpolation). Indeed, at a given point in time, two relative humidity values are all the more similar the closer they are

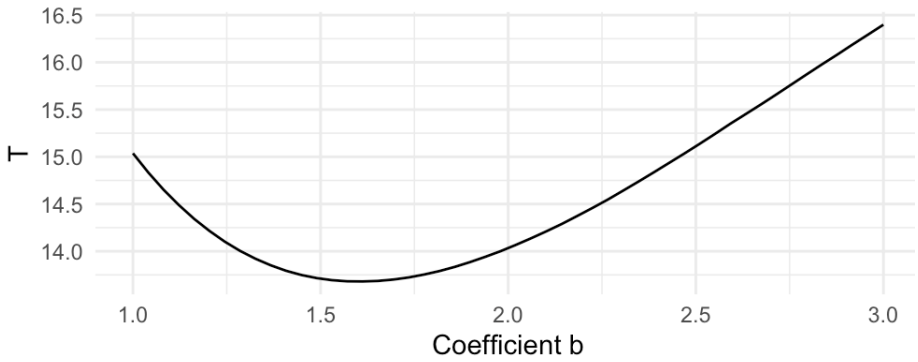

Fig. 11. The Cramér–von Mises criterion value versus a range of b coefficients for a set to its optimal value

in space. Using the Cramér–von Mises criterion is still possible to quantify the discrepancy between IAGOS and ERA5 data, but nothing can be said from the inference perspective if the right assumptions are not used.

Since ERA5 data have been interpolated (Section 5.3), it is possible to search for the values of the coefficients a and b that minimize the correct Cramér–von Mises criterion (Equation 19). We focus on the geographical area studied by Teoh et al. (2022) ($-50°$W, $40°$N, $-10°$W, $75°$N) and estimate the cumulative distribution function of relative humidity on IAGOS measurements in this area for 2019 as well as the one based on interpolated ERA5 data (trilinear interpolation). The optimal a is found to be 0.9367 (Teoh et al. (2022) found $a = 0.9779$). When a is set to the optimal value, the Cramér–von Mises criterion evolves in a comparable manner to what was found by Teoh et al. (2022), as one may notice in Figure 11. The minimum is found when $b = 1.6122$ (Teoh et al. (2022) found $b = 1.776$). As already pinpointed by Teoh et al. (2022), optimal coefficients are only valid for a specific area at a given time. Notably, it is possible to compute these coefficients for each season.

6.1. A seasonal approach to the correction

Still considering the geographical area studied by Teoh et al. (2022), one may be interested in checking the seasonal variations of the correction coefficients. The

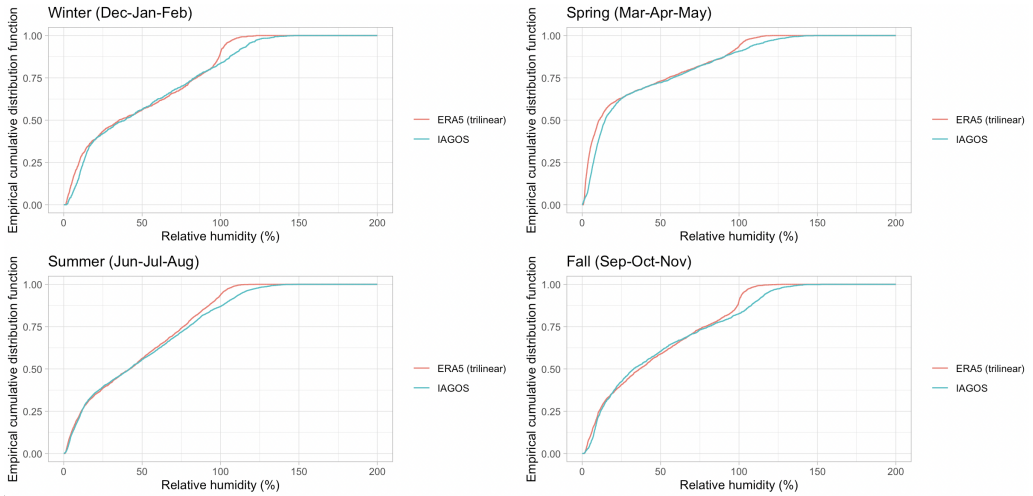

Fig. 12. For each season, empirical cumulative distribution function of relative humidity for each data set (IAGOS data and interpolated ERA5 data).

empirical cumulative distribution functions of relative humidity for each data set (IAGOS data and interpolated ERA5 data) clearly show a kind of ceiling value for relative humidity around 100% (Figure 12). Interestingly, it is not possible to find the optimal coefficients a and b for each season, as clearly observed in Figure 13. While the values of coefficients a and b are of a similar magnitude for summer, winter, and fall, it is evident that the optimal coefficients for spring ($a = 0.6837$, $b = 0.2$) are not satisfactory.

It becomes much clearer when observing the difference in the cumulative distribution functions of relative humidity for spring compared to the other seasons (Figure 14). While the other three seasons exhibit a similar pattern where the difference between the two distributions is mainly noticeable for relative humidity values beyond 100%, it is noteworthy that in spring, this difference is also very pronounced for low relative humidity values. As a result, to minimize the Cramér–von Mises criterion, it is crucial for the coefficient a to be very small in order to correct the disparity between the empirical distribution functions at low relative humidity values. However, this low value of a affects the threshold at which the coefficient b comes into play (see Equation 20). The value of b for which the Cramér–von Mises criterion is minimized is smaller than 1. This is indeed necessary to counterbalance

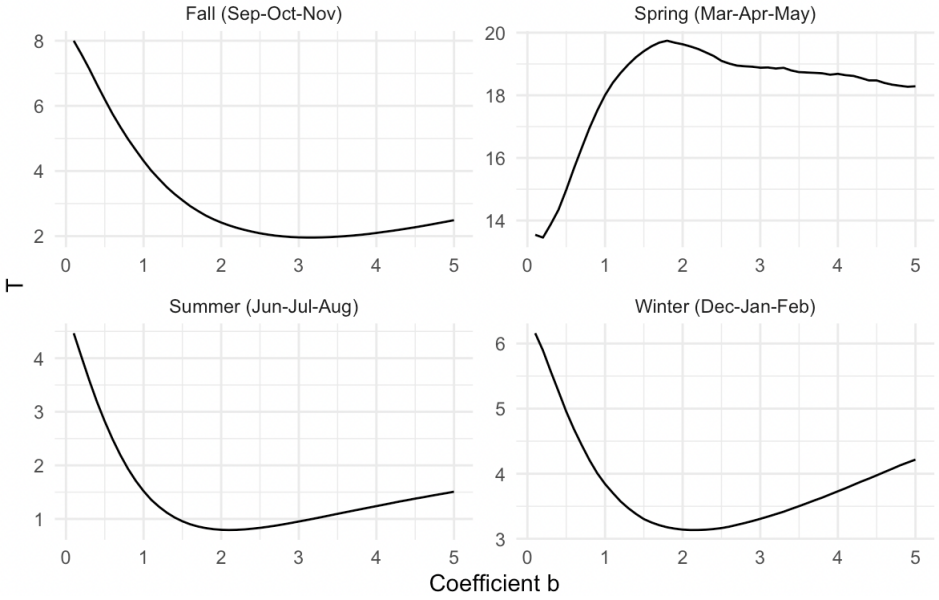

Fig. 13. For each season, the Cramér–von Mises criterion value versus a range of b coefficients for a set to its optimal value.

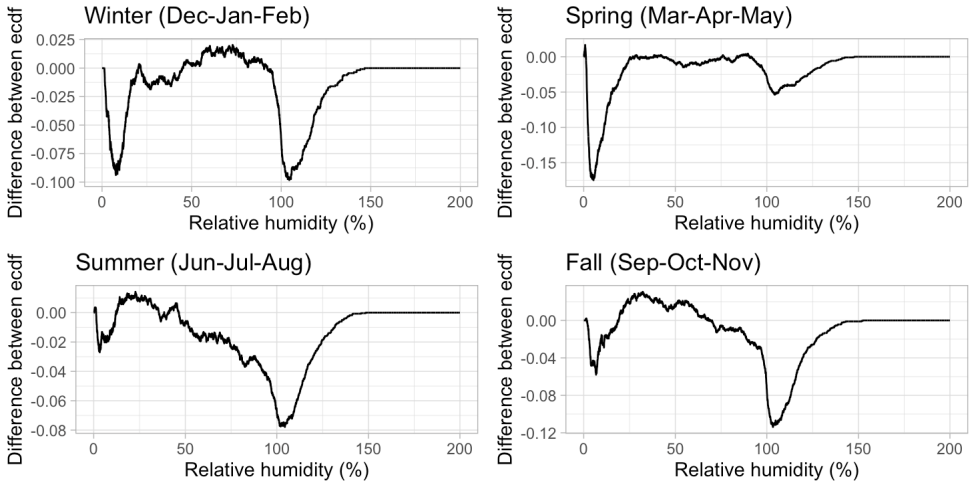

Fig. 14. For each season, the difference between empirical cumulative distribution functions (ecdf).

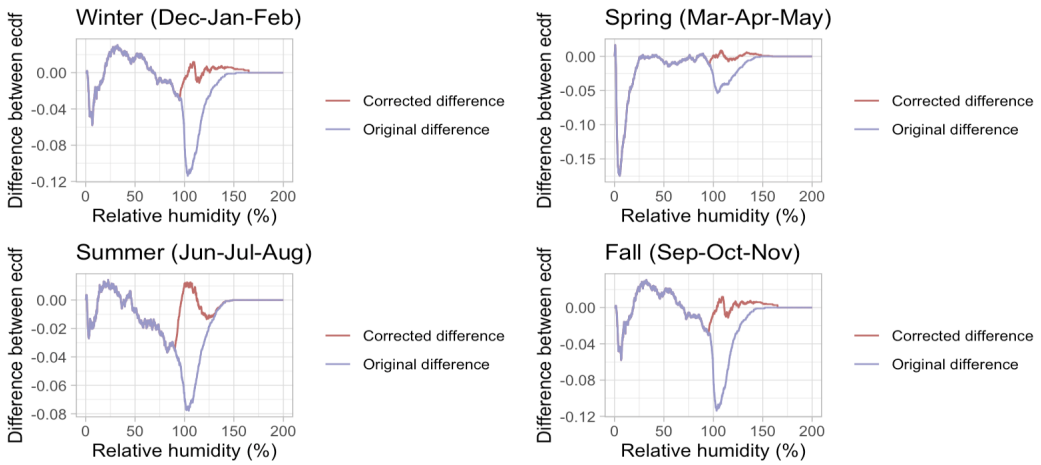

Fig. 15. For each season, the difference between empirical cumulative distribution functions (ecdf) and the corrected difference.

the very small value of a. It is possible to address this issue with a slight modification of Equation 20. Given a location $\boldsymbol{s}_i$, the new corrected relative humidity value, denoted $\check{z}(\boldsymbol{s}_i)$, is given by:

$$\check{z}(\boldsymbol{s}_i) = \begin{cases} z(\boldsymbol{s}_i) & \text{if } \frac{z(\boldsymbol{s}_i)}{a} \leq 1 \\ \min\left\{\left(\frac{z(\boldsymbol{s}_i)}{a}\right)^b, \text{UB}\right\} & \text{if } \frac{z(\boldsymbol{s}_i)}{a} > 1, \end{cases} \tag{20}$$

where $z(\boldsymbol{s}_i)$ is the original relative humidity value (from ERA5), a and b two coefficients and UB a chosen upper bound (based on IAGOS data). With this new formula, it is possible to obtain good correction coefficients, as evidenced by the corrected difference in empirical distribution functions shown in Figure 15.

7. Conclusions

The interpolation of noise or meteorological data can be done using deterministic or stochastic methods. In both cases, choosing an interpolation method generally involves deciding how known values in the domain of interest should be weighted to make predictions at new locations. In both two-dimensional and three-dimensional cases, several elements come into play for determining the best interpolation procedure. One can mention, among other things, the total area of the domain and its topology, the number of

known points and their distribution in space, the presence of a trend and/or of one or more anisotropies, the number of locations to interpolate, and the physical complexity of the studied phenomenon.

Geostatistical methods and interpolation procedures based on inverse distance weighting allow for extrapolation, which is not the case for methods based on a triangulation.

As for interpolating noise values around airports, it is clear that geostatistical methods are particularly suitable because they allow for the consideration of important covariates to obtain credible noise maps. The distance from NMTs to the airport is taken into account as well as runway axes. Linear interpolation is too dependent on the chosen triangulation and does not appear to be suitable in this case.

Regarding the interpolation of meteorological values over multiple pressure layers, it appears that the trend and the presence of very complex anisotropies do not allow for good results with standard geostatistical methods. To account for the complex physics of the phenomena at play, the SR-PDE method seems much more suitable and flexible. Trilinear interpolation is excellent when data is provided on a regular grid, with a fine spatial and temporal resolution and without outliers. It is typically the case for ERA5 reanalysis data. If data are scattered, the SR-PDE method can be seen as a statistical extension of trilinear interpolation.

It is possible to correct the humidity values from ERA5 data based on a ground-truth provided by IAGOS flights. Yet, it is not possible to quantify the agreement between the two humidity fields for a given area at a given hour because the spatiotemporal coverage of IAGOS data is limited. At the very least, one can ensure that the empirical distributions of humidity values agree. A proper parameterization for the correction of humidity values is essential.

References

Anderson, T. W. (1962) On the Distribution of the Two-Sample Cramer-von Mises Criterion. *The Annals of Mathematical Statistics*, **33**, 1148–1159.

Arnone, E., Negri, L., Panzica, F. and Sangalli, L. M. (2023) Analyzing data in com-

plicated 3D domains: Smoothing, semiparametric regression, and functional principal component analysis. *Biometrics*.

Aumond, P., Can, A., Mallet, V., De Coensel, B., Ribeiro, C., Botteldooren, D. and Lavandier, C. (2018) Kriging-based spatial interpolation from measurements for sound level mapping in urban areas. *The Journal of the Acoustical Society of America*, **143**, 2847–2857. URL: `https://doi.org/10.1121/1.5034799`.

Azzimonti, L., Nobile, F., Sangalli, L. M. and Secchi, P. (2014) Mixed Finite Elements for Spatial Regression with PDE Penalization. *SIAM/ASA Journal on Uncertainty Quantification*, **2**, 305–335.

Bendarkar, M. V., Bhanpato, J., Puranik, T. G., Kirby, M. and Mavris, D. N. (2022) Comparative Assessment of AEDT Noise Modeling Assumptions Using Real-World Data. In *AIAA AVIATION 2022 Forum*. American Institute of Aeronautics and Astronautics. URL: `https://arc.aiaa.org/doi/abs/10.2514/6.2022-3917`. _eprint: https://arc.aiaa.org/doi/pdf/10.2514/6.2022-3917.

Boulanger, D., Blot, R., Bundke, U., Gerbig, C., Hermann, M., Nédélec, P., Rohs, S. and Ziereis, H. (2018) IAGOS final quality controlled Observational Data L2 – Time series.

Can, A., Dekoninck, L. and Botteldooren, D. (2014) Measurement network for urban noise assessment: Comparison of mobile measurements and spatial interpolation approaches. *Applied Acoustics*, **83**, 32–39.

Chevallier, R., Shapiro, M., Engberg, Z., Soler, M. and Delahaye, D. (2023) Linear Contrails Detection, Tracking and Matching with Aircraft Using Geostationary Satellite and Air Traffic Data. *Aerospace*, **10**, 578.

Chilès, J.-P. and Delfiner, P. (2012) *Geostatistics, modeling spatial uncertainty*, vol. 713 of *Wiley Series in Probability and Statistics*.

Cohen, J. P. and Coughlin, C. C. (2008) Spatial Hedonic Models of Airport Noise, Proximity, and Housing Prices. *Journal of Regional Science*, **48**, 859–878.

Craven, P. and Wahba, G. (1978) Smoothing noisy data with spline functions. *Numerische Mathematik*, **31**, 377–403.

Duda, D. P., Minnis, P., Nguyen, L. and Palikonda, R. (2004) A Case Study of the Development of Contrail Clusters over the Great Lakes. *Journal of the Atmospheric Sciences*, **61**, 1132–1146.

Filippone, A. (2014) Aircraft noise prediction. *Progress in Aerospace Sciences*, **68**, 27–63.

Franssen, E. A. M. (2004) Aircraft noise around a large international airport and its impact on general health and medication use. *Occupational and Environmental Medicine*, **61**, 405–413.

Gasco, L., Asensio, C. and de Arcas, G. (2017) Communicating airport noise emission data to the general public. *Science of The Total Environment*, **586**, 836–848.

Genescà, M., Romeu, J., Arcos, R. and Martín, S. (2013) Measurement of aircraft noise in a high background noise environment using a microphone array. *Transportation Research Part D: Transport and Environment*, **18**, 70–77.

Gibbons, J. D. and Chakraborti, S. (2010) *Nonparametric Statistical Inference*. Boca Raton: Chapman and Hall/CRC, 5th edition edn.

Gierens, K., Matthes, S. and Rohs, S. (2020) How Well Can Persistent Contrails Be Predicted? *Aerospace*, **7**, 169.

Gierens, K., Spichtinger, P. and Schumann, U. (2012) Ice Supersaturation. In *Atmospheric Physics: Background – Methods – Trends* (ed. U. Schumann), Research Topics in Aerospace, 135–150. Berlin, Heidelberg: Springer. URL: `https://doi.org/10.1007/978-3-642-30183-4`$_9$.

Gierens, K. M. and Vazquez-Navarro, M. (2018) Statistical analysis of contrail lifetimes from a satellite perspective. *Meteorologische Zeitschrift*.

Harman, B. I., Koseoglu, H. and Yigit, C. O. (2016) Performance evaluation of IDW, Kriging and multiquadric interpolation methods in producing noise mapping: A case study at the city of Isparta, Turkey. *Applied Acoustics*, **112**, 147–157.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., Rosnay, P., Rozum, I., Vamborg, F., Villaume, S. and Thépaut, J. (2020) The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, **146**, 1999–2049. URL: `https://onlinelibrary.wiley.com/doi/10.1002/qj.3803`.

Hristopulos, D. T. (2020) *Random Fields for Spatial Data Modeling: A Primer for Scientists and Engineers*. Advances in Geographic Information Science. Dordrecht: Springer Netherlands. URL: `http://link.springer.com/10.1007/978-94-024-1918-4`.

Huynh, J. L., Mahseredjian, A. and John Hansman, R. (2022) Delayed Deceleration Approach Noise Impact and Modeling Validation. *Journal of Aircraft*, **59**, 992–1004. URL: `https://arc.aiaa.org/doi/10.2514/1.C036631`. Publisher: American Institute of Aeronautics and Astronautics.

Journel, A. G. and Huijbregts, C. J. (2004) *Mining Geostatistics*. Caldwell, NJ: The Blackburn Press.

Jäger, D., Zellmann, C., Schlatter, F. and Wunderli, J. M. (2021) Validation of the sonAIR aircraft noise simulation model. *Noise Mapping*, **8**, 95–107. URL: `https://www.degruyter.com/document/doi/10.1515/noise-2021-0007/html`. Publisher: De Gruyter Open Access.

Kärcher, B. (2018) Formation and radiative forcing of contrail cirrus. *Nature Communications*, **9**, 1824.

Lee, D. S., Fahey, D. W., Forster, P. M., Newton, P. J., Wit, R. C. N., Lim, L. L., Owen, B. and Sausen, R. (2009) Aviation and global climate change in the 21st century. *Atmospheric Environment*, **43**, 3520–3537. URL: `https://www.sciencedirect.com/science/article/pii/S1352231009003574`.

Lee, D. S., Fahey, D. W., Skowron, A., Allen, M. R., Burkhardt, U., Chen, Q., Doherty, S. J., Freeman, S., Forster, P. M., Fuglestvedt, J., Gettelman, A., De León, R. R., Lim, L. L., Lund, M. T., Millar, R. J., Owen, B., Penner, J. E., Pitari, G., Prather, M. J., Sausen, R. and Wilcox, L. J. (2021) The contribution of global aviation to anthropogenic climate forcing for 2000 to 2018. *Atmospheric Environment*, **244**, 117834.

Li, J. and Heap, A. D. (2014) Spatial interpolation methods applied in the environmental sciences: A review. *Environmental Modelling & Software*, **53**, 173–189.

Marenco, A., Thouret, V., Nédélec, P., Smit, H., Helten, M., Kley, D., Karcher, F., Simon, P., Law, K., Pyle, J., Poschmann, G., Von Wrede, R., Hume, C. and Cook, T. (1998) Measurement of ozone and water vapor by Airbus in-service aircraft: The MOZAIC airborne program, an overview. *Journal of Geophysical Research: Atmospheres*, **103**, 25631–25642. URL: `http://doi.wiley.com/10.1029/98JD00977`.

Myers, D. E. (1994) Spatial interpolation: an overview. *Geoderma*, **62**, 17–28.

Neis, P., Smit, H. G. J., Rohs, S., Bundke, U., Krämer, M., Spelten, N., Ebert, V., Buchholz, B., Thomas, K. and Petzold, A. (2015) Quality assessment of MOZAIC and IAGOS capacitive hygrometers: insights from airborne field studies. *Tellus B: Chemical and Physical Meteorology*, **67**, 28320.

Paoli, R. and Shariff, K. (2016) Contrail Modeling and Simulation. *Annual Review of Fluid Mechanics*, **48**, 393–427.

Penner, J. E., Lister, D. H., Griggs, D. J., Dokken, D. J. and McFarland, M. (1999) Aviation and the Global Atmosphere. *Tech. rep.*, Intergovernmental Panel on Climate Change.

Petzold, A., Thouret, V., Gerbig, C., Zahn, A., Brenninkmeijer, C. A. M., Gallagher, M., Hermann, M., Pontaud, M., Ziereis, H., Boulanger, D., Marshall, J., Nédélec, P., Smit, H. G. J., Friess, U., Flaud, J.-M., Wahner, A., Cammas, J.-P., Volz-Thomas, A. and Team, I. (2015) Global-scale atmosphere monitoring by in-service

aircraft – current achievements and future prospects of the European Research Infrastructure IAGOS. *Tellus B: Chemical and Physical Meteorology*, **67**, 28452. URL: `https://b.tellusjournals.se/article/10.3402/tellusb.v67.28452/`.

Pretto, M., Giannattasio, P. and De Gennaro, M. (2022) Mixed analysis-synthesis approach for estimating airport noise from civil air traffic. *Transportation Research Part D: Transport and Environment*, **106**, 103248.

Reutter, P., Neis, P., Rohs, S. and Sauvage, B. (2020) Ice supersaturated regions: properties and validation of ERA-Interim reanalysis with IAGOS in situ water vapour measurements. *Atmospheric Chemistry and Physics*, **20**, 787–804. URL: `https://acp.copernicus.org/articles/20/787/2020/`.

Sabatini, R. and Gardi, A. (2023) *Sustainable Aviation Technology and Operations: Research and Innovation Perspectives*. Hoboken, NJ, USA: John Wiley & Sons Inc.

Salvi, M. (2008) Spatial Estimation of the Impact of Airport Noise on Residential Housing Prices. *Swiss Journal of Economics and Statistics*, **144**, 577–606.

Sangalli, L. M. (2021) Spatial Regression With Partial Differential Equation Regularisation. *International Statistical Review*, **89**, 505–531.

Sangalli, L. M., Ramsay, J. O. and Ramsay, T. O. (2013) Spatial spline regression models. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, **75**, 681–703.

Schumann, U. (1996) On conditions for contrail formation from aircraft exhausts. *Meteorologische Zeitschrift*, 4–23.

— (2012) A contrail cirrus prediction model. *Geoscientific Model Development*, **5**, 543–580.

Simons, D. G., Besnea, I., Mohammadloo, T. H., Melkert, J. A. and Snellen, M. (2022) Comparative assessment of measured and modelled aircraft noise around Amsterdam Airport Schiphol. *Transportation Research Part D: Transport and Environment*, **105**, 103216.

Teoh, R., Schumann, U., Gryspeerdt, E., Shapiro, M., Molloy, J., Koudis, G., Voigt, C. and Stettler, M. E. J. (2022) Aviation contrail climate effects in the North Atlantic from 2016 to 2021. *Atmospheric Chemistry and Physics*, **22**, 10919–10935. URL: `https://acp.copernicus.org/articles/22/10919/2022/`. Publisher: Copernicus GmbH.

Tsai, K.-T., Lin, M.-D. and Chen, Y.-H. (2009) Noise mapping in urban environments: A Taiwan study. *Applied Acoustics*, **70**, 964–972.

Vazquez-Navarro, M., Mannstein, H. and Mayer, B. (2010) An automatic contrail tracking algorithm. *Atmospheric Measurement Techniques*, **3**, 1089–1101.

Webster, R. and Oliver, M. A. (2007) *Geostatistics for Environmental Scientists*. Wiley, 1 edn. URL: `https://onlinelibrary.wiley.com/doi/book/10.1002/9780470517277`.

Wilhelm, L., Gierens, K. and Rohs, S. (2021) Weather Variability Induced Uncertainty of Contrail Radiative Forcing. *Aerospace*, **8**, 332.

— (2022) Meteorological Conditions That Promote Persistent Contrails. *Applied Sciences*, **12**, 4450.

Zheng, X., Peng, W. and Hu, M. (2020) Airport noise and house prices: A quasi-experimental design study. *Land Use Policy*, **90**, 104287.