Math Geosci (2020) 52:527–542
https://doi.org/10.1007/s11004-019-09802-y

Surface Estimation for Multiple Misaligned Point Sets

Ashton Wiens[1] · **William Kleiber**[1] · **Katherine R. Barnhart**[2] · **Dylan Sain**[1]

Received: 16 December 2018 / Accepted: 5 April 2019 / Published online: 22 April 2019

Abstract Two common tasks when processing point cloud data sets are surface estimation and point cloud registration. In this paper, a statistical approach is developed to solve both of these problems simultaneously. In particular, a surface is estimated from a pair of unregistered three-dimensional scans of the same spatial region. In this method, one point cloud defines the fixed coordinate system, and a rigid transformation is applied to the second cloud. Observations from both scans are considered a single realization of a Gaussian process. The registration problem is solved by jointly optimizing the likelihood over the parameters specifying the domain transformation and the mean and covariance functions. Given parameter estimates, surface estimation follows using the spatial stochastic model. While other existent approaches do not account for registration uncertainty, the likelihood-based approach to solving the registration and surface estimation problems jointly allows uncertainty in registration to be propagated to the surface prediction variance. The new method is motivated and illustrated using a digital elevation model estimation problem near the Chalk Cliffs in Colorado. The method developed is compared against the popular iterative closest point method. The results of a simulation study show significant improvement in transformation parameter estimates using the statistical approach. In a cross-validation experiment with the Chalk Cliffs data, there is an 18% reduction in predictive mean squared error using the likelihood method over iterative closest point.

Keywords Gaussian process · Point cloud registration · Point cloud · Structure from motion · Digital elevation model · LiDAR

✉ Ashton Wiens
ashton.wiens@colorado.edu

[1] Department of Applied Mathematics, University of Colorado, Boulder, CO, USA

[2] Department of Geological Sciences and Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, USA

1 Introduction

Point clouds are datasets consisting of points located in space, most often $\mathbb{R}^3$. These datasets are also sometimes informally called point clouds, scenes, or scans. Such data are common in a variety of applications in geology, computer vision, agricultural science, remote sensing and medical imaging. There are several methods for collecting point cloud data, such as LiDAR and structure from motion (SfM) algorithms. Most of these data collection methods record a large number of points at very high resolution.

Point cloud data can often be viewed as a possibly noisy partial observation of some continuous underlying surface, and surface estimation is typically the primary goal. A prime example of point cloud data are those used to develop digital elevation models (DEMs) that are widely used in geological, environmental, ecological and geophysical sciences. Conversion of point cloud data to a regular grid can follow from many methods, including linearly interpolated triangulation, inverse distance weighting, and kriging (Lloyd and Atkinson 2002; Schwendel et al. 2012). Arun (2013) reviews and compares several methods for estimating DEMs, finding kriging to outperform most other approaches but at an added computational cost (Sunila and Virrantaus 2011).

Most surface estimation methods rely on having a single partial observation of the surface. However, especially for producing DEMs, multiple scans, flyovers or datasets may be available that are representative of the same geographical region. A potential issue with multiple scans of the same region is that they may not be a priori aligned in the same coordinate system. An example case, which is explored in the data example, is the construction of an SfM point cloud at a site that is too steep for the placement of ground control points (GCPs). Point cloud registration is a process involving finding a spatial transformation which approximately aligns two point clouds.

Point cloud registration is a well-established field in remote sensing and engineering. Registration algorithms generally fall into two classes—registration of the raw point clouds or a surface representation step before registration. Methods which do not convert the point clouds to a continuous representation rely on the concept of point matching, in which the algorithm attempts to match each point in the first data set to its closest neighbor in the second data set (either a point or another geometric object). A transformation is then estimated to minimize a cost function, which is often the sum of the squared differences between each pair of points. The popular iterative closest point (ICP) method (Besl and McKay 1992) and the robust point matching algorithm (Gold et al. 1998) fall into this category. The first step in the ICP algorithm is discrete point matching, where every point in the source (or moving) point cloud is put in correspondence with a point in the reference (or fixed) point cloud, usually the closest point according to some distance measure. Then, the optimal spatial transformation applied to the moving point cloud is found by minimizing the sum of squared distances between each set of two points in correspondence. After a suitable transformation has been estimated, the transformation is applied to the moving point cloud. This process is repeated iteratively by updating the point correspondences and re-estimating and applying a spatial transformation. The algorithm terminates after a fixed number of iterations or after the error becomes smaller than a set tolerance. One

disadvantage of point matching methods is that finding nearest neighbors can incur a large computational burden when trying to solve the discrete point matching problem.

The second class of methods circumvent point matching by using one or more continuous representations of the point clouds. The kernel correlation method (Tsin and Kanade 2004), a similar method based on Gaussian mixtures (Jian and Vemuri 2005), and the coherent point drift method (Myronenko and Song 2010) model the point clouds using kernel density estimators or Gaussian mixture models and then find the optimal spatial transformation between these new continuous representations. This conversion strategy can reduce the computational burden of the optimization. See Tam et al. (2013) and Maiseli et al. (2017) for recent comprehensive reviews of point cloud registration algorithms. Although registration algorithms have been well explored, the problem of registration uncertainty is not often addressed in the spatial statistics literature: usually uncertainty quantification is limited to the response variable and not extended to the positional variables.

Several authors have addressed inference and prediction of spatial models in the presence of location error (Cressie and Kornak 2003; Fanshawe and Diggle 2011; Fronterré et al. 2018; Kotsakis 2019). Cervone and Pillai (2015) provide a review, and then they develop an MCMC approach to prediction. However, the main focus of these approaches are cases where an individual dataset has positional data with independent or uncorrelated errors. This is fundamentally a different problem than multiple dataset registration where each position is internally consistent (i.e., without location error), but there is inconsistency in the coordinate systems between datasets. As a working hypothesis, we assume there is no location error. This is justifiable when the location errors are much smaller than the translation parameters in the registration, and we expect the location errors to be close to the order of magnitude of the nugget effect, which is very small in the applications in this paper. In other words, the bias in the positional variables applies to only one dataset and is common for all locations within that dataset. Thus, the goal is to estimate and quantify uncertainty in the transformation parameters and not the observation locations themselves.

In this paper a statistical approach is proposed to simultaneously solve the two problems of point cloud registration and surface estimation. The solution was motivated by a dataset consisting of two scans performed on the same spatial region near the Chalk Cliffs in Colorado. An SfM algorithm generates point clouds over the same domain, but whose location data are not registered. The approach developed here is relatively simple: a joint Gaussian process specification for the observational model, where one scan's statistical model includes a rigid transformation of its coordinate system. Observations from both scans are considered a single realization of the process. The registration problem is solved by jointly optimizing the likelihood over the transformation parameters and the other parameters specifying the mean and covariance functions. Model parameters are estimated in a maximum likelihood framework and use standard asymptotic arguments to obtain approximate parameter uncertainties. Given model parameters, surface estimation follows directly by kriging the registered datasets. This method is different compared to the current point cloud registration methods: the task of discrete point matching is avoided, and the optimization in the continuous domain doesn't rely on kernels, mixture models, or representing the point clouds as realizations of a point process before registration. A novel feature of the

method developed in this paper is that the uncertainty in the surface estimate reflects both the uncertainty that results from the stochastic surface model and the uncertainty of the registration, implemented using Monte Carlo techniques.

2 A Surface Estimation Model with Registration

This section briefly describes the dataset and sets up some notation, followed by a detailed description of the statistical model, and the registration and surface estimation procedures.

2.1 A Drone-Based Elevation Dataset

Our data come from two drone-based flyovers of the Chalk Cliffs debris flow basin in central Colorado. The Chalk Cliffs site is a 0.37 km^2 watershed at the base of Mount Princeton, seven miles west of Nathrop, Colorado in the Arkansas River Valley. Debris flows occur at the site between one and four times per year, typically in response to rain events (Coe et al. 2008; McCoy et al. 2010; Kean et al. 2009). The frequency of debris flows and the uniformity of the underlying rock make the Chalk Cliffs an ideal study catchment for developing a mechanistic understanding of natural debris flows. Observational data relevant to a mechanistic understanding of sediment creation and transport processes at Chalk Cliffs include maps of topographic change.

SfM is an image-based photogrammetric method that uses multiple images of a static scene taken from many different locations to jointly solve for the camera lens calibration parameters, the location and orientation of each photograph, and a sparse cloud of prominent features in the scene (James and Robson 2012; Westoby et al. 2012; Fonstad et al. 2013). Finally, a multi-view stereo algorithm is used to construct depth maps of each image and create a dense point cloud. A common step in geological applications is the placement of GCPs within the study scene that are used to place the scene in an absolute reference frame (Reitman et al. 2015; Warrick et al. 2017).

For the data application, point clouds from two survey flights on October 3, 2017 are available. A DJI Phantom 4 Pro drone was used, programmed using Universal Ground Control Software (v2.11.250). The drone was programmed to survey a portion of the Chalk Cliffs basin at 76 m above ground level with 70% photo overlap and 30% sidelap. Each flight resulted in around forty photographs. Processing the photographs into point clouds was done in Agisoft Photoscan (v1.4.2). Each survey resulted in a sparse point cloud of about 80–100 thousand points and a dense cloud of about 60 million points (about 250 points per square meter). The Chalk Cliffs site is extremely steep—over half of the basin is exposed bedrock cliff. For this reason, the surveys do not have ground control; instead, the coordinates in the point clouds rely on the less precise GPS on board the drone.

Several authors have attempted error quantification in producing point cloud SfM data products, either in the context of georegistration (Harwin and Lucieer 2012; Nilosek et al. 2014), or comparing against other data sources such as LiDAR or a TLS, e.g. Mancini et al. (2013). See Ozyesil et al. (2017) for a survey of SfM research, and Seitz et al. (2006) for a comparison of several multi-view stereo reconstruction

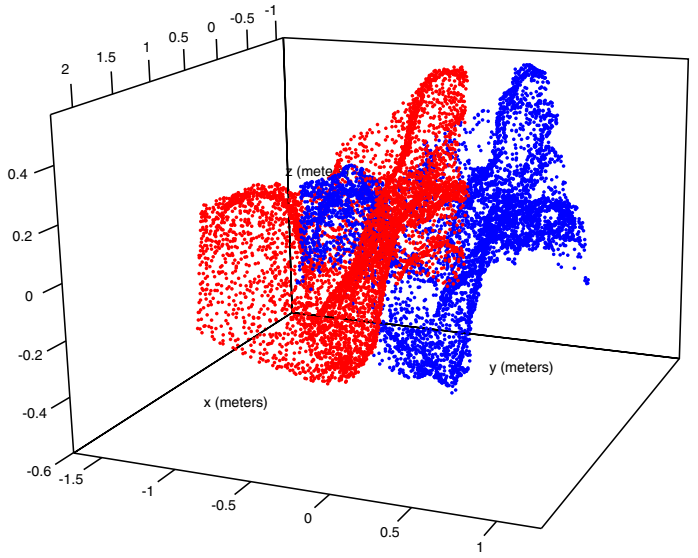

Fig. 1 The two unregistered point clouds partially overlapping in red and blue. Axes are projected and centered easting, northing and elevation coordinates

algorithms. Truong et al. (2017) explored registration of SfM point clouds corresponding to RGB and thermal images using rigid body transformations and a variant of ICP, but attempts at statistical uncertainty quantification have not been made.

The data are two point clouds in $\mathbb{R}^3$, but the approach developed in this paper can easily be extended with minor modifications to account for higher dimensions and more than two datasets. As the interest in this paper is in creating a DEM by joining two datasets, denote the first dataset as $Y_1(\mathbf{s}_{1j})$ for $j = 1, \ldots, n_1$ and $Y_2(\mathbf{s}_{2j})$ for $j = 1, \ldots, n_2$. Here, $\mathbf{s}_{ij} \in \mathbb{R}^2$ denotes easting and northing positional coordinates while $Y_i(\cdot)$ denotes the measured elevation. Figure 1 shows both datasets in their unregistered coordinate systems. There are clear coordinate shifts required for registering, and there is also potentially rotational misalignment.

2.2 The Statistical Model

The following data generating model is proposed

$$\begin{aligned} Y_1(\mathbf{s}) &= Z(\mathbf{s}) + \varepsilon_1(\mathbf{s}) \\ Y_2(\mathbf{s}) &= \mu + Z(T(\mathbf{s})) + \varepsilon_2(T(\mathbf{s})), \end{aligned}$$

where the transformation is applied to the second dataset. Both datasets represent the same underlying continuous surface, $Z(\cdot)$, which is modeled as a mean zero Gaussian process. Data are subject to observational noise or microscale variations $\varepsilon_i(\cdot)$ for $i = 1, 2$, which is assumed to be mean zero Gaussian white noise processes with variance τ^2. Due to the unregistered nature of the two datasets, a bias parameter $\mu \in \mathbb{R}$

which is a translation in the elevation coordinate and T which is a transformation in the positional coordinates, are introduced.

Note that the identically distributed assumption on both datasets' error processes is natural, as both derive from the same observational platform. The independence assumption (i.e., rather than specifying $\varepsilon_1 = \varepsilon_2$) is necessary, as any co-located points between the two datasets are subject to independent errors as they were captured at two different time points, not in the same flyover. If nonstationary error measures are provided in the data set, the distribution of the residual process can be generalized to $\varepsilon_i(s) \sim N(0, \tau_i^2(s))$.

In this work $T : \mathbb{R}^2 \to \mathbb{R}^2$ is restricted to be a rigid transformation that includes both a translation and rotation of the positional coordinates but not a reflection. The transformation can be written with a translation vector $\mathbf{r} \in \mathbb{R}^2$ and a 2×2 orthogonal rotation matrix R parameterized by an angle ϕ as follows.

$$\begin{aligned} T(\mathbf{s}) &= R\mathbf{s} + \mathbf{r} \\ &= \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} s_x \\ s_y \end{bmatrix} + \begin{bmatrix} r_x \\ r_y \end{bmatrix}, \end{aligned}$$

where $\mathbf{s} = [s_x\, s_y]^{\mathrm{T}}$ and $\mathbf{r} = [r_x\, r_y]^{\mathrm{T}}$. The transformation depends on parameters r_x, r_y and ϕ. By definition, a proper rigid transformation preserves Euclidean distance among the points to which it is applied.

While the rigid transformation provides positional translations and rotations, a more complicated model could include three rotation parameters (the roll, pitch, and yaw). Based on exploratory data analysis, this more complicated rotation does not appear to be necessary for the Chalk Cliffs dataset. Moreover, such a choice preserves the additive error structure of the model: if the z-axis is allowed to be rotated, the error structure in the transformed data set will not be additive with the error in the first data set, and it is not clear how to incorporate such errors into the model. Other work has addressed how nonlinear transformations using splines and Gaussian processes can be used to model the mapping function describing the transformation (Zhu et al. 2009). However, it is unclear how such a framework could preserve the structure of the additive model. Thus, one limitation of this approach is that it assumes that the points are a single-valued function in some reference frame.

All that remains to be specified is the structure of the continuous spatial variation process $Z(\cdot)$. Z is modeled as an isotropic random field with Matérn covariance function. In particular

$$\mathrm{Cov}(Z(\mathbf{s}), Z(\mathbf{s}')) = C(\mathbf{s}, \mathbf{s}') = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{d}{a}\right)^{\nu} K_\nu\left(\frac{d}{a}\right),$$

where $d = \|\mathbf{s} - \mathbf{s}'\|$ is the Euclidean distance between points, σ^2 is the marginal variance, $a > 0$ a spatial range parameter and $\nu > 0$ the smoothness. Γ is the gamma function, and K_ν is the modified Bessel function of the second kind, of order ν. The Matérn covariance function is frequently used in spatial statistics due to its flexibility and the interpretability of its parameters. Stein (1999) recommended general adoption

of the Matérn family of models because spatial fields with any degree of differentiability can be modeled by using different ν parameters.

2.3 Parameter Estimation

Data are available as $\mathbf{Y} = [\mathbf{Y}_1^{\mathrm{T}}, \mathbf{Y}_2^{\mathrm{T}}]^{\mathrm{T}}$, where $\mathbf{Y}_1 = (Y_1(\mathbf{s}_{11}), \ldots, Y_1(\mathbf{s}_{1n_1}))^{\mathrm{T}}$ and $\mathbf{Y}_2 = (Y_2(\mathbf{s}_{21}), \ldots, Y_2(\mathbf{s}_{2n_2}))^{\mathrm{T}}$. Based on the statistical model, one can decompose $\mathbf{Y}_1 = \mathbf{Z}_1 + \boldsymbol{\varepsilon}_1$ and $\mathbf{Y}_2 = \mu\mathbf{1}_{n_2} + \mathbf{Z}_2 + \boldsymbol{\varepsilon}_2$, where $\mathbf{Z}_1 = (Z(\mathbf{s}_{11}), \ \ldots \ , Z(\mathbf{s}_{1n_1}))^{\mathrm{T}}$ and $\mathbf{Z}_2 = (Z(T(\mathbf{s}_{21})), \ \ldots \ , Z(T(\mathbf{s}_{2n_2})))^{\mathrm{T}}$ with analogous definitions for $\boldsymbol{\varepsilon}_1$ and $\boldsymbol{\varepsilon}_2$.

The mean of $\mathbf{Y}$ is $\mathbb{E}\mathbf{Y} = \mathbf{m} = [\mathbf{0}_{n_1}^{\mathrm{T}} \ \mu\mathbf{1}_{n_2}^{\mathrm{T}}]^{\mathrm{T}}$ where $\mathbf{0}_{n_1}$ and $\mathbf{1}_{n_2}$ are zero and unit vectors of lengths n_1 and n_2, respectively. The covariance matrix for $\mathbf{Y}$ is

$$\Sigma + \tau^2 I = \begin{bmatrix} \mathrm{Cov}(\mathbf{Z}_1, \mathbf{Z}_1) & \mathrm{Cov}(\mathbf{Z}_1, \mathbf{Z}_2) \\ \mathrm{Cov}(\mathbf{Z}_2, \mathbf{Z}_1) & \mathrm{Cov}(\mathbf{Z}_2, \mathbf{Z}_2) \end{bmatrix} + \tau^2 I \,,$$

where I is an identity matrix of dimension $n_1 + n_2$. Here, as an example, the (i, j)th element of $\mathrm{Cov}(\mathbf{Z}_1, \mathbf{Z}_2)$ is $C(\mathbf{s}_{1i}, T(\mathbf{s}_{2j}))$. In the event that, for some pair $\mathbf{s}_{1i}$ and $\mathbf{s}_{2j}$, it is true that $\mathbf{s}_{1i} = T(\mathbf{s}_{2j})$, the covariance matrix is still strictly positive definite due to the independent error processes ε_1 and ε_2. If the error processes were equal, $\varepsilon_1 = \varepsilon_2$ then $\mathrm{Var}\,\mathbf{Y}$ would still be nonnegative definite, but would possess at least one zero eigenvalue.

The log-likelihood function for $\mathbf{Y}$ is

$$\ell(\mathbf{Y}) = -\frac{n_1 + n_2}{2} \log 2\pi - \frac{1}{2} \log |\Sigma + \tau^2 I| - \frac{1}{2} (\mathbf{Y} - \mathbf{m})^{\mathrm{T}} (\Sigma + \tau^2 I)^{-1} (\mathbf{Y} - \mathbf{m}).$$

The development of the likelihood function can be found in the "Appendix". The maximum likelihood estimators for σ^2, τ^2, a, μ, r_x, r_y and ϕ are those that maximize $\ell(\mathbf{Y})$. For this work the smoothness parameter is fixed at $\nu = 1$, corresponding to the Whittle covariance model. For computational purposes, an iterative procedure is implemented in the maximization of the likelihood, alternating fixing the covariance parameters σ^2, τ^2 and a and the transformation parameters μ, r_x, r_y and ϕ. As a preliminary step to help satisfy the zero-mean assumption of the model, the empirical mean response of each point cloud individually is subtracted from each dataset. This step forces the two point clouds to overlap near the origin of the fixed cloud at the very least. Forcing the data sets to overlap is an important aspect in recovering the optimal registration parameters using this model. A practical issue arises when the optimization over the transformation parameters is left unconstrained. If the surfaces are allowed to be completely non-overlapping, there is no incentive to intersect the convex hulls of the two datasets—thus, the translation parameters must be constrained to force the two point clouds to overlap.

The first step in the iteration is to optimize over the transformation parameters while keeping the other covariance function parameters fixed at some reasonable values. The initial values for the covariance function parameters can be roughly estimated by fitting a Gaussian process to one or both of the point clouds individually with the

same covariance function (excluding the transformation part of the algorithm). In simulations it was observed that the likelihood function of the translation parameters is well-behaved, even if the parameters in the covariance function are misspecified. This first step should yield transformation parameters close to the optimal solution, and can be verified with a visual check. In the second step, the transformation parameters are kept fixed and the optimization is performed over the other covariance function parameters. This process is then iterated until changes in parameter estimates become negligible. Then, a final joint optimization can be performed over all parameters simultaneously.

2.4 Predictive Distribution of the Surface

The goal in surface estimation is to estimate $Z(\mathbf{s}_0)$ at an arbitrary spatial location $\mathbf{s}_0$, given data $\mathbf{Y}$. Considering the MLEs as fixed, the conditional expectation $\mathbb{E}(Z(\mathbf{s}_0)|\mathbf{Y})$ is used, which is the optimal predictor for minimizing predictive mean squared error. Due to the Gaussian process specification, this coincides with the standard simple kriging estimator, with minor modifications due to the positional translation function. Specifically, let

$$\Sigma_0 = \mathrm{Cov}(Z(\mathbf{s}_0), \mathbf{Y}) = [C(\mathbf{s}_0, \mathbf{s}_{11}), \ldots, C(\mathbf{s}_0, \mathbf{s}_{1n_1}), C(\mathbf{s}_0, T(\mathbf{s}_{21})), \ldots, C(\mathbf{s}_0, T(\mathbf{s}_{2n_2}))],$$

denote the row vector of covariances between the process at the location of interest and the observational process. Then $\mathbb{E}(Z(\mathbf{s}_0)|\mathbf{Y}) = \Sigma_0(\Sigma + \tau^2 I)^{-1}(\mathbf{Y} - \mathbf{m})$ and $\mathrm{Var}(\hat{Z}(\mathbf{s}_0) - Z(\mathbf{s}_0)|\mathbf{Y}) = \sigma^2 - \Sigma_0(\Sigma + \tau^2 I)^{-1}\Sigma_0^{\mathrm{T}}$.

An advantage of taking a statistical approach to point set registration is that uncertainty in the registration parameters can be estimated. Indeed, most other point cloud registration methods provide only a single point estimate of the transformation parameters, either because they are not statistical in nature or are unconcerned with uncertainty quantification. Uncertainty in the parameter estimates is quantified by relying on the asymptotic distribution of the MLEs using an increasing domain framework (Mardia and Marshall 1984). In particular, the Fisher information matrix is approximated using a numerical approximation of the Hessian of the log-likelihood function. Similar to a Bayesian approach, the predictive distribution of $Z(\mathbf{s}_0)$ is the quantity of interest. The predictive distribution is $[Z(\mathbf{s}_0)|\mathbf{Y}] = [Z(\mathbf{s}_0)|\boldsymbol{\theta}, \mathbf{Y}][\boldsymbol{\theta}|\mathbf{Y}]$, by concatenating all statistical parameters into a vector $\boldsymbol{\theta} = (\sigma^2, \tau^2, a, \mu, r_x, r_y, \phi)^{\mathrm{T}}$. Instead of relying on the posterior distribution of $\boldsymbol{\theta}$, one can generate samples from the asymptotic distribution of the MLEs. To generate a prediction of the surface, a sample of each MLE is drawn from its asymptotic distribution, which then defines a conditional normal predictive distribution for $Z(\mathbf{s}_0)$, given $\mathbf{Y}$ and the sample of $\boldsymbol{\theta}$.

3 Data Examples

This section is split into a simulation study and the data analysis of the Chalk Cliffs point clouds. In both cases the method developed in this paper is compared against one state-of-the-art method in point set registration, ICP.

The simulation study allows for comparison of parameter estimation quality, while the drone-based data example is validated by cross-validation where the fully executed algorithm including surface estimation is considered.

3.1 A Simulation Study

The simulation study is designed to mimic the ensuing data example to suggest a realistic measure of quality of registration ability from both methods. The simulation study begins with sampling 1200 pairs of x- and y-coordinates independently from a random uniform distribution on $[0, 6] \times [0, 6]$. Next, the z-coordinates are simulated as a spatially correlated field at these locations using the fixed Matérn covariance parameters $a = 0.6$, $\sigma^2 = 1$, and $\tau^2 = 0.01$. These parameters are similar to estimates from the next section, and the estimated fields indeed appear qualitatively similar to the geologic data sets analyzed in this paper. Next, the dataset is split in half by randomly assigning each point to one of two subsets, and then one subset is de-registered using randomly sampled translation and rotation parameters. The translation parameters are sampled uniformly on $[0, 1]$ and the two-dimensional rotation parameter is sampled uniformly on $[0, \pi/4]$. The two de-registered datasets are inspected to make sure they were not completely non-overlapping. Finally, both the proposed likelihood-based method and ICP are used to recover the registration parameters. The experiment is repeated thirty times, recording both true registration parameters and the parameters estimated by each method.

ICP was performed using the implementation in the open source CloudCompare software (Girardeau-Montaut 2018).

The optimization in the likelihood-based algorithm was performed with the `optim` function in R using the `L-BFGS-B` option. Box constraints were included on the parameter space: the translation parameters were confined to an interval of width 0.8 centered around the true values and, similarly, the rotation angle (in radians) was confined to an interval of width 0.4 centered around the true value. The initial guess for the optimization for all parameters was a random uniform drawn from the constrained sample space.

In two of the thirty simulations, the likelihood-based algorithm returned one or more parameters on the boundary of the search space. This indicates the optimization got stuck in a local minimum instead of the global minimum near the true parameter values. In these two cases a different initial guess was generated and the algorithm was performed again, which resulted in near recovery of the true parameters, not on the boundary of the search space. In general, to avoid finding local minima, the likelihood-based algorithm can be run several times with different initial guesses. In the applications of this paper, it was observed that different initial guesses returned the same results, unless one or more returned a solution on the boundary of the search space. Since there were no constraints put on the registration parameters in the ICP algorithm, each simulation was visually inspected after registration. Only one time, ICP returned results in which the two data sets were not registered by eye. In this case, a manual transformation was applied to the unregistered dataset to bring it closer

Table 1 RSMEs of transformation parameter estimates over 30 simulations for the proposed likelihood-based method and ICP

Parameter	Root mean squared error	
	Likelihood-based	ICP
r_x	0.005	0.308
r_y	0.009	0.187
μ	0.010	0.060
ϕ	0.002	0.009

Table 2 Maximum likelihood estimates and standard errors as well as ICP estimates of the transformation and covariance parameters for the full model using both point clouds

Parameter	MLE	MLE SE	ICP est.
r_x	− 0.350	1.8×10^{-6}	− 0.406
r_y	0.505	3.3×10^{-7}	0.486
μ	− 0.063	7.3×10^{-7}	− 0.049
ϕ	− 0.008	1.4×10^{-6}	0.037
σ^2	0.145	7.9×10^{-3}	–
τ^2	0.0009	2.3×10^{-10}	–
a	0.831	6.6×10^{-2}	–

to alignment with the fixed dataset. These adjustments realistically reflect how these algorithms would be used in practice.

Based on the thirty sets of true and estimated parameters, the root mean squared error (RMSE) in the estimates was calculated for both methods, displayed in Table 1. Overall, the likelihood-based method performs substantially better than ICP in estimation, indicating anywhere between an 81–98% improvement depending on the parameter. Table 2 shows that ICP can perform an order of magnitude worse than the likelihood-based method when estimating the translation parameters. It should be noted that both algorithms tend to fail if the data sets are not aligned well enough before they are registered. Even when the two data sets have $> 50\%$ overlap, the algorithm might not successfully register the data sets, especially if there is a large rotation needed to realign them. In such cases an initial rough alignment by the practitioner is required, but is also a realistic expectation in actual applications.

3.2 Drone-Based Point Clouds from the Chalk Cliffs

The proposed method is applied to a subset of the drone-based elevation data described in Sect. 2.1. The original data consist of two scans of the same spatial region, taken on the same day, in which no large scale erosion was occurring. The full scans contain around 60 million data points each. For the purposes of this manuscript, the same

geologic feature is identified in both scenes and each dataset is manually cropped around this feature; see Fig. 1. Each cropped scene contains 6000 points. The goal is to register the two point clouds and produce a surface estimate of the elevation (a DEM) with parametric uncertainty incorporated. Note that the spatial boundaries of the two cropped regions are not identical. Following the convention of this paper, denote the two (subsetted) point clouds in this application $\mathbf{Y}_1$ and $\mathbf{Y}_2$. Finally, for predictive validation, the data were split into training and testing subsets—randomly holding out 25% of both point clouds as the testing sets, and the remaining 75% were used for parameter estimation and spatial prediction.

For comparison, the ICP registration algorithm was also applied to the data. To match the restrictions of the likelihood-based method, in the ICP algorithm the rotation was restricted to the xy-plane, along with the standard three-dimensional translation. The open source CloudCompare 2.9.1 software package was used to perform ICP on the two unregistered training sets (Girardeau-Montaut 2018).

To determine the maximum likelihood estimates, optimization of the likelihood was performed on the training data in an iterative fashion as detailed at the end of Sect. 2.3. The changes in the parameter estimates from the first to the second iteration were negligible, likely due to the large, spatially dense nature of the study data set. The final joint optimization over all parameters resulted in small changes compared to the initial estimates. The ICP algorithm's four registration parameters are the same as those in the statistical model, and are thus directly comparable. All parameter estimates with associated standard errors are displayed in Table 2. The transformation parameters are significantly nonzero in all three cardinal directions, while the rotation parameter shows that very slight rotation is necessary for proper registration of these two point clouds. Note the noise-to-signal ratio for these data is estimated to be under 1%, indicating very accurate elevation estimates. Harwin and Lucieer (2012) estimate elevation accuracy up to 25–40 mm from a point cloud generated via SfM with $< 1-3$ cm point spacing. The Chalk Cliffs point cloud is 0.1389 points/cm^2 and, remarkably, the estimated residual standard deviation is $\tau = 0.03$ m, or 30 mm accuracy, which conforms with the estimated accuracy reported by Harwin and Lucieer (2012). Figure 2 shows the registered datasets using the optimal transformation estimated by maximum likelihood. Visually the point clouds appear to be aligned, with the common valleys and hills coinciding.

To quantify the quality of the alignment, a cross-validation experiment was also conducted. In particular, after training the model on the training data, the predictive performance is compared using standard simple kriging to predict values in the testing data. As a statistical method, the method developed here produces point estimates, but also a predictive distribution for the testing data that incorporates predictive and parametric uncertainty, as described in Sect. 2.4. Thus, predictive performance is quantified in terms of RMSE, validating the kriging predictor, and also the continuous ranked probability score (CRPS) for assessing the predictive distributions (Gneiting and Raftery 2007; Gneiting et al. 2007). Given a forecast distribution F and realizing observation x, the CRPS can be written

$$\mathrm{CRPS}(F, x) = \mathbb{E}_F|X - x| - \frac{1}{2}\mathbb{E}|X - X'|\,, \tag{1}$$

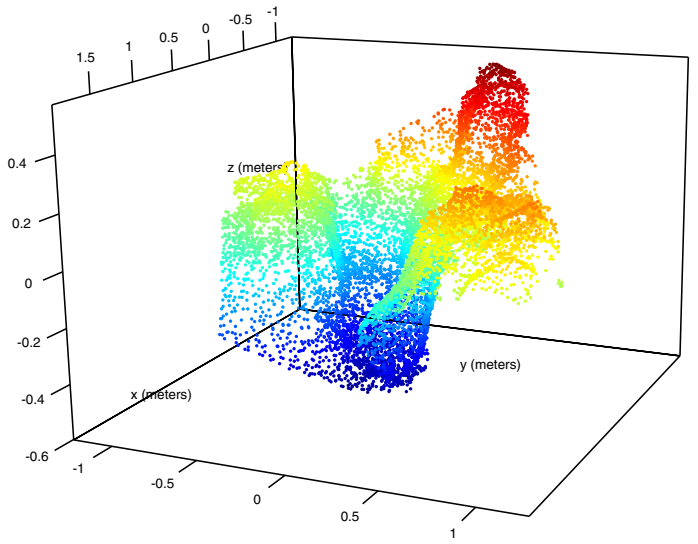

Fig. 2 The two point clouds are aligned by eye after the estimated transformation has been applied, colored by elevation (Z). Alignment was performed by the likelihood algorithm with a three-dimensional translation and a two-dimensional rotation, meaning there are 4 transformation parameters r_x, r_y, μ, and ϕ

where X and X' are independent random variables with distribution F. For Gaussian predictive distributions, this formula can be simplified in terms of the standard normal probability density and cumulative distribution functions (Gneiting and Raftery 2007) as follows

$$\mathrm{CRPS}(\mathcal{N}(\mu, \sigma^2), x) = \sigma \left[\frac{1}{\sqrt{\pi}} - 2\varphi\left(\frac{x-\mu}{\sigma}\right) - \frac{x-\mu}{\sigma}\left(2\Phi\left(\frac{x-\mu}{\sigma}\right) - 1\right)\right],$$

where φ and Φ are the probability density and cumulative distributions functions, respectively. The predictive distribution of Sect. 2.4 is only normal conditional on parameter estimates; the posterior predictive distribution implemented here accounts for uncertainty in parameter estimates and is thus no longer exactly normal. To estimate CRPS, Monte Carlo approximations of Eq. (1) are used.

To assess the method, the cross-validation experiments are split up into predicting from each individual dataset, as well as the aligned dataset pair. The goal is to predict the held-out 25% of data from each point cloud. The four prediction methods employed are [A] predict held-out data using only $\mathbf{Y}_1$, [B] predict held-out data using only $\mathbf{Y}_2$, [C] predict held-out data using $\mathbf{Y} = (\mathbf{Y}_1^{\mathrm{T}}, \mathbf{Y}_2^{\mathrm{T}})^{\mathrm{T}}$ aligned by the proposed statistical model, and [D] predict held-out data using $\mathbf{Y} = (\mathbf{Y}_1^{\mathrm{T}}, \mathbf{Y}_2^{\mathrm{T}})^{\mathrm{T}}$ aligned by ICP. In all four cases, predictions are made at the locations of all held out data from both point clouds. In experiments A, B and C the transformation parameters estimated jointly are applied; the major difference between these three experiments is the data on which kriging is conditioned. In experiment D, the transformation is estimated using ICP, but

Table 3 RMSE and CRPS of the testing data for kriging using only individual point clouds or both point clouds with alignment estimated statistically or using ICP

Data used for kriging	Alignment method	RMSE	CRPS
$\mathbf{Y}_1$	Statistical	2.75	0.023
$\mathbf{Y}_2$	Statistical	2.14	0.020
$\mathbf{Y}_1$ and $\mathbf{Y}_2$	ICP	2.06	0.019
$\mathbf{Y}_1$ and $\mathbf{Y}_2$	Statistical	1.69	0.016

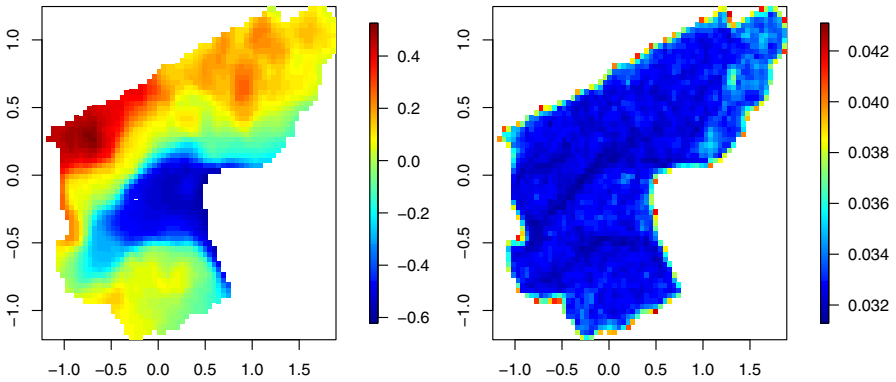

Fig. 3 **a** Mean Kriging surface predictions evaluated on a 100 × 100 grid. **b** Kriging standard errors on the same grid. Predictions are valid only within the area encompassed by the data, so the grids are constructed to cover the data but not to extrapolate. Units for x, y, z (elevation), and the standard errors are all in meters

the kriging is performed on the full testing dataset using spatial parameters estimated by maximum likelihood after the ICP transformation.

Table 3 shows summary statistics for all four experiments. Considering only point prediction accuracy, kriging based on datasets aligned by ICP is slightly better than based on one subset of data. However, accuracy of the point predictions is 18% better using statistical alignment over ICP. The results are striking for the predictive distributions as well: there is a 15% reduction in CRPS using the statistically aligned data over the data aligned by ICP. The superior performance of the statistical model using both $\mathbf{Y}_1$ and $\mathbf{Y}_2$ is due to the better-aligned data that have a substantially higher sampling density than either individual point cloud alone.

To conclude the validation, the final DEM product estimated by the statistical model is presented. The left panel of Fig. 3 shows the mean surface prediction evaluated on a grid, conditional on both point clouds and the maximum likelihood estimates. The right panel of Fig. 3 shows predictive kriging standard errors using fixed covariance and registration function parameters at their MLEs. Greater uncertainty is apparent in regions without changing elevation, due to the SfM's inclination to produce fewer points in these regions.

4 Conclusions

In this paper, a method was developed based on a Gaussian process model which can simultaneously perform point cloud registration and surface estimation of two point clouds. The model is statistically straightforward, which allows for standard likelihood-based estimation methods. Furthermore, the parsimonious approach performs well against non-statistical competitors. The model was compared against ICP in both a simulation study and on a digital elevation dataset. In both cases the likelihood-based model significantly outperformed ICP in recovering registration parameters, and thus also produced a better surface prediction based on these parameter estimates. Moreover, by using a statistical approach, uncertainty is readily quantified in parameter estimates and ensuing surface estimates using either Bayesian methods or asymptotic distributions of maximum likelihood estimators. Another advantage of the model is that it is easily extendable for multiple unregistered point clouds to be used in the surface estimation, whereas most existent registration algorithms are constrained to pairwise comparisons.

The computational efficiency for the straightforward likelihood-based algorithm is $\mathcal{O}(n^3)$. For ICP, a naive nearest neighbor search is $\mathcal{O}(dn)$ for n points in d dimensions, while the number of iterations has been bounded by $\mathcal{O}(n/d)^{d+1}$ (Ezra et al. 2008; Arthur and Vassilvitskii 2009). Clearly there is a large computational savings in using ICP, possibly at the expense of accuracy as shown in this work. Future work could use techniques in spatial statistics to reduce the computational burden of the likelihood-based approach (Gelfand et al. 2010). Using sparse matrix methods or other techniques would also make this approach capable of handling larger data sets.

Future work may address fitting spatio-temporal models to multiple scans of the same spatial region that are temporally correlated. In the context of the drone-based elevation application in this paper, a spatio-temporal model could be a way to model the erosion that occurs over time in the Chalk Cliffs. However, before this task can be attempted, it is necessary to have an accurate spatial model for multiple temporally-indexed point clouds. Another possible extension of this model would be to tackle the large sample size problem without subsetting the point clouds. Fitting Gaussian process models to massive elevation point clouds has been implemented by Vasudevan et al. (2009) using a combination of nearest neighbor prediction and k-d trees for the storage and retrieval of the point clouds. The authors also account for discontinuities in the complex terrain by employing a nonstationary, neural network covariance function.

Data and R scripts are available at https://github.com/awiens11/GPregistration.

Acknowledgements Funding was provided by Division of Mathematical Sciences (Grant No. 1811294).

Appendix

Let $n = n_1 + n_2$. The data generating process is $\mathbf{Y} \sim \text{MVN}_n(\mathbf{m},\ \Sigma + \tau^2 I)$. The density for $\mathbf{Y}$ is

$$\pi(\mathbf{Y}) = \frac{1}{(2\pi)^{n/2}\sqrt{|\Sigma + \tau^2 I|}} \exp\left(-\frac{1}{2}(\mathbf{Y}-\mathbf{m})^{\mathrm{T}}(\Sigma + \tau^2 I)^{-1}(\mathbf{Y}-\mathbf{m})\right).$$

The log-likelihood function for $\mathbf{Y}$ is then

$$\begin{aligned}\ell(\mathbf{Y}) &= \log\left(\frac{1}{(2\pi)^{n/2}\sqrt{|\Sigma + \tau^2 I|}} \exp\left(-\frac{1}{2}(\mathbf{Y}-\mathbf{m})^{\mathrm{T}}(\Sigma + \tau^2 I)^{-1}(\mathbf{Y}-\mathbf{m})\right)\right) \\ &= -\frac{n}{2}\log 2\pi - \frac{1}{2}\log|\Sigma + \tau^2 I| - \frac{1}{2}(\mathbf{Y}-\mathbf{m})^{\mathrm{T}}(\Sigma + \tau^2 I)^{-1}(\mathbf{Y}-\mathbf{m}).\end{aligned}$$

References

Arthur D, Vassilvitskii S (2009) Worst-case and smoothed analysis of the ICP algorithm, with an application to the k-means method. SIAM J Comput 39(2):766–782

Arun PV (2013) A comparative analysis of different DEM interpolation methods. Egypt J Remote Sens Space Sci 16(2):133–139

Besl PJ, McKay ND (1992) A method for registration of 3-d shapes. IEEE Trans Pattern Anal Mach Intell 14(2):239–256

Cervone D, Pillai NS (2015) Gaussian process regression with location errors. arXiv:1506.08256 [stat.ME]

Coe JA, Kinner DA, Godt JW (2008) Initiation conditions for debris flows generated by runoff at Chalk Cliffs, central Colorado. Geomorphology 96(3–4):270–297

Cressie N, Kornak J (2003) Spatial statistics in the presence of location error with an application to remote sensing of the environment. Stat Sci 18(4):436–456

Ezra E, Sharir M, Efrat A (2008) On the performance of the ICP algorithm. Comput Geom 41(1):77–93

Fanshawe TR, Diggle PJ (2011) Spatial prediction in the presence of positional error. Environmetrics 22(2):109–122

Fonstad MA, Dietrich JT, Courville BC, Jensen JL, Carbonneau PE (2013) Topographic structure from motion: a new development in photogrammetric measurement. Earth Surf Process Landf 38(4):421–430

Fronterrè C, Giorgi E, Diggle P (2018) Geostatistical inference in the presence of geomasking: a composite-likelihood approach. Spat Stat 28:319–330

Gelfand EA, Diggle P, Fuentes M, Guttorp P (2010) Handbook of spatial statistics. CRC Press, Boca Raton

Girardeau-Montaut D (2018) Cloudcompare version 2.9.1. [GPL software] http://www.cloudcompare.org/. Accessed 15 May 2018

Gneiting T, Raftery AE (2007) Strictly proper scoring rules, prediction, and estimation. J Am Stat Assoc 102(477):359–378

Gneiting T, Balabdaoui F, Raftery AE (2007) Probabilistic forecasts, calibration and sharpness. J R Stat Soc: Ser B (Statistical Methodology) 69(2):243–268

Gold S, Rangarajan A, Lu CP, Pappu S, Mjolsness E (1998) New algorithms for 2d and 3d point matching: pose estimation and correspondence. Pattern Recognit 31(8):1019–1031

Harwin S, Lucieer A (2012) Assessing the accuracy of georeferenced point clouds produced via multi-view stereopsis from unmanned aerial vehicle (UAV) imagery. Remote Sens 4(6):1573–1599

James MR, Robson S (2012) Straightforward reconstruction of 3D surfaces and topography with a camera: accuracy and geoscience application. J Geophys Res Earth Surf 117:F03017

Jian B, Vemuri BC (2005) A robust algorithm for point set registration using mixture of gaussians. In: Tenth IEEE international conference on computer vision (ICCV'05) volume 1, vol 2, pp 1246–1251. ISSN: 1550-5499

Kean JW, McCoy SW, Tucker GE, Staley DM, Coe JA (2013) Runoff-generated debris flows: observations and modeling of surge initiation, magnitude, and frequency. J Geophys Res Earth Surf 118(4):2190–2207

Kotsakis C (2019) Nonlinear geospatial frame transformations in the presence of noisy data. Math Geosci 51(4):437–461. https://doi.org/10.1007/s11004-018-9742-1

Lloyd CD, Atkinson PM (2002) Deriving DSMs from LiDAR data with kriging. Int J Remote Sens 23(12):2519–2524
Maiseli B, Gu Y, Gao H (2017) Recent developments and trends in point set registration methods. J Vis Commun Image Represent 46:95–106
Mancini F, Dubbini M, Gattelli M, Stecchi F, Fabbri S, Gabbianelli G (2013) Using unmanned aerial vehicles (UAV) for high-resolution reconstruction of topography: the structure from motion approach on coastal environments. Remote Sens 5(12):6880–6898
Mardia KV, Marshall RJ (1984) Maximum likelihood estimation of models for residual covariance in spatial regression. Biometrika 71(1):135–146
McCoy SW, Kean JW, Coe JA, Staley DM, Wasklewicz TA, Tucker GE (2010) Evolution of a natural debris flow: in situ measurements of flow dynamics, video imagery, and terrestrial laser scanning. Geology 38(8):735–738
Myronenko A, Song X (2010) Point set registration: coherent point drift. IEEE Trans Pattern Anal Mach Intell 32(12):2262–2275
Nilosek D, Walvoord DJ, Salvaggio C (2014) Assessing geoaccuracy of structure from motion point clouds from long-range image collections. Opt Eng 53(11):113112
Ozyesil O, Voroninski V, Basri R, Singer A (2017) A survey of structure from motion. Acta Numer 26:305–364
Reitman NG, Bennett SEK, Gold RD, Briggs RW, DuRoss CB (2015) High-resolution trench photomosaics from image-based modeling: workflow and error analysis. Bull Seismol Soc Am 105(5):2354–2366
Schwendel AC, Fuller IC, Death RG (2012) Assessing DEM interpolation methods for effective representation of upland stream morphology for rapid appraisal of bed stability. River Res Appl 28(5):567–584
Seitz SM, Curless B, Diebel J, Scharstein D, Szeliski R (2006) A comparison and evaluation of multi-view stereo reconstruction algorithms. In: 2006 IEEE computer society conference on computer vision and pattern recognition (CVPR'06), vol 1, pp 519–528. ISSN: 1063-6919
Stein ML (1999) Interpolation of spatial data: some theory for kriging, chapter 1. Springer, New York, pp 1–14
Sunila R, Virrantaus K (2011) Digital elevation model construction using geostatistics and geological expert knowledge—a case study in oitti area in southern finland. Nordic J Surv Real Estate Res 8(1). https://journal.fi/njs/article/view/6420
Tam GKL, Cheng Z, Lai Y, Langbein FC, Liu Y, Marshall D, Martin RR, Sun X, Rosin PL (2013) Registration of 3d point clouds and meshes: a survey from rigid to nonrigid. IEEE Trans Vis Comput Graph 19(7):1199–1217
Truong TP, Yamaguchi M, Mori S, Nozick V, Saito H (2017) Registration of RGB and thermal point clouds generated by structure from motion. In: 2017 IEEE international conference on computer vision workshops (ICCVW), pp 419–427. ISSN: 2473-9944
Tsin Y, Kanade T (2004) A correlation-based approach to robust point set registration. In: Pajdla T, Matas J (eds) Computer vision—ECCV 2004. Springer, Berlin. ISBN: 978-3-540-24672-5, 558–569
Vasudevan S, Ramos F, Nettleton E, Durrant-Whyte H (2009) Gaussian process modeling of large-scale terrain. J Field Robot 26(10):812–840
Warrick JA, Ritchie AC, Adelman G, Adelman K, Limber PW (2017) New techniques to measure cliff change from historical oblique aerial photographs and structure-from-motion photogrammetry. J Coast Res 331:39–55
Westoby MJ, Brasington J, Glasser NF, Hambrey MJ, Reynolds JM (2012) 'Structure-from-Motion' photogrammetry: a low-cost, effective tool for geoscience applications. Geomorphology 179:300–314
Zhu J, Hoi SCH, Lyu MR (2009) Nonrigid shape recovery by gaussian process regression. In: 2009 IEEE conference on computer vision and pattern recognition, pp 1319–1326. ISSN: 1063-6919