Ecological Complexity 31 (2017) 149–164

Contents lists available at ScienceDirect

Ecological Complexity

journal homepage: www.elsevier.com/locate/ecocom

Original Research Article

On dynamics underlying variance of mass balance estimation in Chilean glaciers

M. Stehlík[a,b,*], P. Hermann[a], S. Torres[c], J. Kiseľák[a,d], A. Rivera[e,f]

[a] Department of Applied Statistics and Linz Institute of Technology, Johannes Kepler University Linz, Altenbergerstraße 69, 4040 Linz, Austria
[b] Institute of Statistics, University of Valparaíso, Blanco 951, Valparaíso, Region de Valparaíso, Chile
[c] Departamento de Matemática, Universidad Técnica Federico Santa María, Casilla 110-V, Valparaíso, Chile
[d] Institute of Mathematics, P.J. Šafárik Univesity in Košice, Slovakia
[e] Laboratorio de Glaciología, Centro de Estudios Científicos (CECs), Arturo Prat 514, Valdivia, Chile
[f] Departamento de Geografía, Universidad de Chile, Marcoleta 250, Santiago, Chile

ARTICLE INFO

Article history:
Received 31 March 2017
Received in revised form 16 June 2017
Accepted 22 June 2017
Available online 7 July 2017

MSC:
Primary 62K05
62K15
Secondary 62M30

Keywords:
D-optimality
Variance
Dynamical system
Blow-up
Glaciers
Mass balance

ABSTRACT

Mass balance of a glacier is an accepted measure of how much mass a glacier gains or loses. In theory, it is typically computed by integral functional and empirically, it is approximated by arithmetic mean. However, the variability of such an approach was not studied satisfactory yet. In this paper we provide a dynamical system of mass balance measurements under the constrains of 2nd order model with exponentially decreasing covariance. We also provide locations of optimal measurements, so called designs. We study Ornstein–Uhlenbeck (OU) processes and sheets with linear drifts and introduce K optimal designs in the correlated processes setup. We provide a thorough comparison of equidistant, Latin Hypercube Samples (LHS), and factorial designs for D- and K-optimality as well as the variance. We show differences between these criteria and discuss the role of equidistant designs for the correlated process. In particular, applications to estimation of mass balance of Olivares Alfa and Beta glaciers in Chile is investigated showing that simple application of full raster design and kriging based on inter- and extrapolation of points can lead to increased variance. We also show how the removal of certain measurement points may increase the quality of the melting assessment while decreasing costs. Blow-ups of solutions of dynamical systems underline the empirically observed fact that in a homogenous glaciers around 11 well-positioned stakes suffices for mass balance measurement.

1. Introduction

Glacier mass balance is an important measure of the glacier health, since accounts for the mass gains and losses during a specific period of time, normally a hydrological or calendar year (Rivera et al., 2016). We will focus in the glaciological method (Cogley et al., 2012), here the mass balance is measured at stakes or poles' networks installed on the glacier surface, whose distribution depends on altitude, slope, topography, and other parameters. In many cases, the distribution is skewed or seriously limited by accessibility (mainly due to crevasses) or logistical constrains. These measurements are ideally done at monthly frequency, but accessibility or budget limitations reduce the number of surveys, sometimes to a very minimum of two per year, one in the accumulation season peak and another in the ablation season maximum. The stakes height above the snow/ice surface at the beginning of the mass balance year (normally at the end of the ablation season), is therefore compared to successive measurements along the year, when snow/ice density must be also determined in order to convert vertical heights into water equivalent volumes (Cuffey and Paterson, 2010). The discrete mass balance data must be integrated over the entire glacier surface by applying geo-spatial interpolation methods or simply by computing the arithmetic mean of measurements (Cogley et al., 2012).

The effective sample size has been addressed previously in (Cogley, 1999), where the analysis of multiple time series of point mass balance measurements have shown that correlation decreases along differences in elevation between the points. A

* Corresponding author at: Department of Applied Statistics and Linz Institute of Technology, Johannes Kepler University Linz, Altenbergerstraße 69, 4040 Linz, Austria.
E-mail addresses: mlnstehlik@gmail.com (M. Stehlík), philipp.hermann@jku.at (P. Hermann), sebastian.torresle@alumnos.usm.cl (S. Torres), jozef.kiselak@gmail.com (J. Kiseľák), arivera@cecs.cl (A. Rivera).

http://dx.doi.org/10.1016/j.ecocom.2017.06.008
1476-945X/

similar conclusion is obtained in Fountain and Vecchia (1999), where the dominant effect of the gradient of mass balance with altitude was shown to be more relevant than transverse variations. Hence, the number of mass balance measurements required to determine the mass balance appears to be scale invariant for small glaciers ($<10\,km^2$) with five to ten stakes being enough.

In this paper we define an underlying dynamical system, thus putting measurement locations and underlying parameters into dynamical relation. Such dynamical systems can appear naturally because of various modelling modalities (see Stehlík et al., 2016, 2017). Based on this background we discuss the statistical optimality of stakes distributed over the glacier. A problem which arises is that neither locations nor measurements themselves are independent. This is why we provide a comparison between monotonic and space-filling designs, introduced in the case of Ornstein–Uhlenbeck (OU) sheets in Baran et al. (2013); Baran and Stehlík (2015). We consider both, D- and K-optimality and minimization of variance of arithmetic mean. Practically, we study Smit's paradox (Smit, 1961) for planar Ornstein Uhlenbeck process, so called OU sheet.

In the next section we show different designs strategies, then in Section 3, we consider OU processes and OU sheet. We continue with Section 4 where we provide a motivating example on mass balance for glaciers Olivares Alpha and Beta in Chile. Some results follow in Section 5 where the simulation on defined optimal design criteria allows us to compare different design strategies. Section 6 contains the unpleasant increment of variance after using kriging methods for the estimation of the mass balance. Moreover, the variance changes after the removal of stakes, even showing a decrease of it under a particular covariance structure. Finally, Section 7 provides a dynamical system and blow-ups in order to technically explain interrelations between parameters and design points. We show, that under some generic circumstances variance of mass balance estimator can grow with the number of stakes.

2. Optimal design under correlated errors

In many situations we can meet problems of unavoidable increase of variance (this can be related to Smit's work Smit, 1961) when using additional interpolation or extrapolation by simple kriging. Hence, more sophisticated designs than usage of full rasterization of the grid, namely equidistant, factorial, Latin Hypercube Samples under S-optimality (LHS^*) and Latin Hypercube Samples optimal with respect to Euclidean distances (LHS^+) designs are compared with respect to D- and K-optimality as well as the variance. Therefore, mass balance estimation of glaciers Olivares Alpha and Beta, important meltwater contributors to the Maipo River in Santiago, requires proper data sampling techniques. To construct such techniques, we need to deal with optimal design strategies.

The determination of optimal designs for models with correlated errors is substantially more difficult and for this reason not sufficiently developed. A stochastic process with parameterized mean and covariance is observed on a compact set. The information obtained from observations is measured through the information functional (defined on the Fisher information matrix (FIM)). We focus on efficient designs for parameters of correlated processes and discuss the role of equidistant designs for correlated processes. Such designs have been proven to be optimal for parameters of trend of stationary Ornstein–Uhlenbeck process (see Kiselák and Stehlík, 2008). For such a process a study of small samples and asymptotical comparisons of the efficiencies of equidistant designs was provided whilst taking both the parameters of trend as well as the parameters of covariance into account. If only trend parameters are of interest, the designs covering more or less uniformly the whole design space will rather be efficient when correlation decreases exponentially (see Kiselák and Stehlík, 2008). Some other issues on designs for spatial processes, i.e. identifiablity and existence of optimal designs, are given in Dette et al. (2008); Müller and Stehlík (2009); Stehlík et al. (2008). The role of heteroscedasticity is studied in Boukouvalas et al. (2014).

Exact K-optimal designs have been firstly introduced by Ye and Zhou (2013) in the setup of polynomial regression models and were later extended in Rempel and Zhou (2014). Both of these setups consider cases having independent errors. K-optimality is a new design criterion for the construction of regression designs, based on the condition number of the information matrix. Thus, K-optimal design minimizes the condition number $\kappa(M)$ of Fisher information matrix M, i.e.

$$\kappa(M) = \frac{\lambda_1(M)}{\lambda_p(M)},$$

if $\lambda_p(M) > 0$ and ∞ otherwise. Here, $\lambda_1(M)$ and $\lambda_p(M)$ are the largest and the smallest eigenvalues, respectively.

Multicollinearity is a common problem when estimating linear or generalized linear models. It occurs when there are high correlations among the predictor variables, leading to unreliable and unstable estimates of regression coefficients. However, many data analysts do not realize that there are several situations in which multicollinearity can be safely ignored, and we hope that K-optimal design is a helpful tool in this direction (see Baran, S., K-optimal designs for parameters of shifted Ornstein-Uhlenbeck processes andsheets. J. Stat. Plan. Inference 186 (2017), 28–41).

3. Ornstein–Uhlenbeck process and sheet

Consider the stochastic process

$$Y(s) = \alpha_1 + \alpha_2 s + \varepsilon(s), \tag{1}$$

where $C(\varepsilon(s), \varepsilon(t)) = \exp(-r|s-t|)$. For model (1) the Fisher information matrix $M_\theta(n)$ on the unknown parameter vector $\theta = (\alpha_1, \alpha_2)$ based on observations $\{Y(s_i),\ i = 1, 2, \ldots, n\}$, $n \geq 2$, equals

$$M_\theta(n) = H(n)C(n)^{-1}H(n)^\top, \quad \text{where} \quad H(n) : = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ s_1 & s_2 & \cdots & s_n \end{pmatrix},$$

and $C(n)$ is the covariance matrix of the observations (see Pázman, 2007; Xia et al., 2006). On the other hand, consider now the stationary process

$$Y(s,t) = \theta + \varepsilon(s,t) \tag{2}$$

with design points taken from a compact design space $\mathcal{X} = [a_1, b_1] \times [a_2, b_2]$, where $b_1 > a_1$ and $b_2 > a_2$ and $\varepsilon(s,t), s,t \in \mathbb{R}$, is a stationary Ornstein–Uhlenbeck sheet, that is a zero mean Gaussian process with covariance structure

$$\mathrm{E}\,\varepsilon(s_1,t_1)\varepsilon(s_2,t_2) = \frac{\tilde{\sigma}^2}{4\alpha\beta}\exp(-\alpha|t_1 - t_2| - \beta|s_1 - s_2|), \tag{3}$$

where $\alpha > 0, \beta > 0, \tilde{\sigma} > 0$. We remark that $\varepsilon(s,\ t)$ can also be represented as

$$\varepsilon(s,t) = \frac{\tilde{\sigma}}{2\sqrt{\alpha\beta}}\mathrm{e}^{-\alpha t - \beta s}\mathcal{W}(\mathrm{e}^{2\alpha t}, \mathrm{e}^{2\beta s}),$$

where $\mathcal{W}(s,t), s,t \in \mathbb{R}$, is a standard Brownian sheet (Baran et al., 2013). Under this setup we compute Fisher information (FIM) for four designs defined in Section 2. Eq. (4) computes FIM for equidistant design points on the diagonal of a square $[0,1] \times [0,1]$ (also equidistant on diagonal line (EDL) hereafter). Eq. (5) calculates FIM for factorial design with points in $[0,\ 1] \times [0,\ 1]$. Fisher information matrix calculation for the LHS designs (both