MIT Libraries | DSpace@MIT

MIT Open Access Articles

Stochastic mobility prediction of ground vehicles over large spatial regions: a geostatistical approach

Massachusetts Institute of Technology

Stochastic Mobility Prediction of Ground Vehicles over Large Spatial Regions. A Geostatistical Approach.

Ramón González · Paramsothy Jayakumar · Karl Iagnemma

Abstract This paper describes a stochastic approach to vehicle mobility prediction over large spatial regions ($> 5 \times 5\ [Km^2]$). The main source of uncertainty considered in this work derives from uncertainty in terrain elevation, which arises from sampling (at a finer resolution) a Digital Elevation Model (DEM). In order to account for such uncertainty, Monte Carlo simulation is employed, leading to a stochastic analysis of vehicle mobility properties. Experiments performed on two real data sets (namely, the Death Valley region and Sahara desert) demonstrate the advantage of stochastic analysis compared to classical deterministic mobility prediction. These results show the computational efficiency of the proposed methodology. The robotic simulator ANVEL has also been used to validate the proposed methodology.

Keywords NATO Reference Mobility Model (NRMM) · Mission planning · Geographical Information Systems (GIS) · D* path planner · Statistical sampling

R. González
Massachusetts Institute of Technology
77 Massachusetts Ave., Room 35-135a,
Cambridge, MA 02139 USA, Tel: +1 617 324-2113
E-mail: ramong@mit.edu

P. Jayakumar
US Army RDECOM TARDEC
6501 E. 11 Mile Road, MS 157, Bldg. 215,
Warren, MI 48397-5000, Tel: +1 586 282-4896
E-mail: paramsothy.jayakumar.civ@mail.mil

K. Iagnemma
77 Massachusetts Ave., Room 35-237a
Cambridge, MA 02139 USA, Tel: +1 617 452-3262
E-mail: kdi@mit.edu

1 Introduction

It is well-known that before attempting a mission involving a ground vehicle in off-road conditions a reliable and comprehensive analysis of the mobility capabilities of such a vehicle is desired. This goal can be solved by means of computer simulation, where terrain profile plays a key role. Traditionally, this analysis considers nominal values for the key variables involved in the simulation (e.g. elevation). This leads to unreliable and limited results due to the uncertainty present in those variables (Lessem et al 1996; Peynot et al 2014).

Terrain geometry information typically comes from remote sensory sources (e.g. radar technology, imagery methods, etc.). Those techniques lead to models of the terrain with uncertainty associated with the spatial position of data points. Thus, any elevation model of the terrain is corrupted by uncertainty. Digital Elevation Models (DEMs) produced by the US Geological Survey agency are a good example of this issue (Zhou and Liu 2004).

Here, we describe a stochastic approach for vehicle mobility prediction over large regions ($> 5 \times 5[Km^2]$). In this framework, a model of the terrain is created using geostatistical methods. The performance of a vehicle is then evaluated while considering the terrain profile. In order to account for uncertainty, Monte Carlo simulations are performed, leading to a statistical analysis. Uncertainty in elevation is due to the new interpolated terrain model to a higher spatial resolution than the original DEM (through a geostatistical method called Ordinary Kriging (Chiles and Delfiner 2012)). The ultimate goal is to compare different routes using different cost functions in the path planning algorithm and various performance indices.

Notice that this work is framed in the context of the next-generation NRMM military software (the current software is described in (Haley et al 1979)). One requirement is that such software can be executed on general-purpose computers available in current military vehicles or military camps in battle fields. In this regards, the proposed approach means an efficient solution in terms of computation burden, and demonstrates its suitability on a standard-performance computer.

The rest of the paper is organized as follows. Related work is reviewed in Section 2. Section 3 gives a general view of the proposed methodology. Section 4 reviews some geostatistical tools used in this research and a subsampling approach is also proposed to reduce the complexity of a large Digital Elevation Model (DEM). The mobility prediction problem is explained in Section 5. Illustrative examples are given in Section 6. Finally, conclusions and future work are discussed in Section 7.

2 Related Work

This paper addresses the problem of manned and unmanned vehicles mobility prediction in off-road conditions. In this context, a comprehensive survey of the state-of-the-art is found in (Papadakis 2013). As in that survey, many publications cope with 3D path planning in the vicinity of the robot (Goldberg et al 2002), (Thrun et al 2006). Those approaches are generally not appropriate for planning longer routes over large environments (e.g. Curiosity rover over a large Martian region or a military vehicle over a battle field) because they are based on sensors that perceive only the surrounding environment (e.g. stereovision, LIDAR).

Regarding the issue of path planning over longer routes, an important research effort has been made in the field of combining remote data with ground data. The focus of that work has been mainly in the fields of obstacle avoidance (Stentz et al 2003) and robot location (Vandapel et al 2006). In this way, some ideas proposed in this paper were inspired by the high-fidelity traversability analysis proposed in (Helmick et al 2009) and the research related to combining remote and ground data in (Stentz et al 2002; Vandapel et al 2006). In (Helmick et al 2009) mobility prediction is carried out online (every 20-50 meters) in order to find an optimal path between the current robot position and a desired target. In this way, many paths are randomly generated between the robot position and the next waypoint; the robot motion is thus analyzed for each path, and eventually the route with the lowest cost is followed.

In our case, mobility prediction is planned for the entire route before the vehicle moves (planning stage). In (Vandapel et al 2006) a 3D elevation map is used to test the performance of a vehicle traversing between two given points. In this case, a kinematic-model-based analysis is performed. But only a binary answer is given for each cell in the map: traversable or non-traversable. After that, a path planner is used to determine the path of least cost in terms of a performance index. In (Stentz et al 2002) an unmanned air vehicle that flies ahead of the vehicle is used to detect holes and other hazards before the onboard vehicle sensors would be able to detect them.

The research in (Lessem et al 1996) addressed the problem of stochastic mobility prediction in military vehicles. However, uncertainty is arbitrarily selected depending on "expert opinions". The output is a speed map based on the maximum and minimum speeds that a military vehicle can achieve in each map cell. In contrast, in our methodology uncertainty comes from a rigorous mathematical formulation (ordinary kriging). Furthermore, it results in an optimal route between two desired points in terms of several performance indices.

Another similar research to that described here is found in (Peynot et al 2014). Here, the authors define mobility prediction as the problem of estimating the likely behavior of a planetary exploration rover in response to given control actions on a given terrain. In this sense, uncertainty is in the control rather than in the DEM/terrain map.

Mobility prediction has also been considered in terms of the vehicle-terrain interactions (Ishigami et al 2009; Willoughby et al 2006). In (Ishigami et al 2009), a statistical method for mobility prediction considering uncertainty in terrain physical properties (cohesion and internal friction angle) is proposed. In this case, both a vehicle dynamic model and a wheel-terrain contact model are employed for simulating the robot motion considering uncertainty in terrain parameters. (Karumanchi et al. 2010) a non-parametric learning technique is presented to generate a mobility map based on terrain elevation and wheel slip. Path planning takes advantage of such map improving vehicle heading and velocity in off-road slopes.

On the other hand, some research projects focus on the reconstruction of a 3D surface (digital elevation map) from sparse data obtained from a remote sensor (Hadsell et al 2009; Kweon and Kanade 1992). However, these works do not consider the second goal of our research, that is, stochastic mobility prediction.

3 General Framework

The first problem addressed in this paper is framed within the context of estimating the values of a regionalized variable (elevation) at places where it has not been

measured and, after that, analyzing the performance of a vehicle over such terrain.

Figure 1 shows the methodology followed in this research. Initially, a DEM[1] is selected related to the environment where the vehicle is going to operate. After that, a reduced-order representation of the DEM points is obtained via a subsampling approach. This is required in order to enable an affordable computation of the variogram and kriging method. Once a set of representative points, in terms of the variogram and elevation profile, are selected, the ordinary kriging method is applied. This procedure yields a model of the terrain at a finer resolution (see Assumption 1). This model can be applied to statistical simulation because each interpolated point has an associated uncertainty (kriging variance). Next, the D* algorithm is applied in order to obtain multiple routes between two points (i.e. starting point and desired goal or waypoint). Finally, the optimal route is given in terms of some performance indices (e.g. the shortest path, the path with the lowest uncertainty, the flattest route, etc.) (see Assumption 2).

Assumption 1 *We assume that there is no uncertainty or error associated to the original DEM of the terrain (positional errors, data precision, etc.). In this paper, the uncertainty is only considered after obtaining a model of the terrain at a finer spatial resolution (new interpolated points), that is, after applying kriging.*

Assumption 2 *It is assumed that the method proposed here will be used for the mobility of car-like vehicles. In particular, the testbed considered in this research is the known Humvee military vehicle (AM General 2015). The reasoning behind the spatial resolution of the model of the terrain and the use of a car-like kinematic model in the simulations are both based on the dimensions and features of such vehicle.*

The main features and limitations of this approach can be summarized as follows:

- Global path planning is considered rather than local path planning (planning in the vicinity of the vehicle). The latter problem is mainly related to detection and avoidance of local obstacles using onboard sensors. In this sense, this research does not ensure local obstacle avoidance. It is assumed that a lower layer to the one proposed here will be available in the real vehicle.
- This solution does not result in a binary answer, i.e. the path is traversable or not; instead statistical data supporting each decision is given.

---

[1] A DEM is defined as a set of elevation values which are recorded on a regular grid (i.e. square form) (Fisher and Tate 2006).
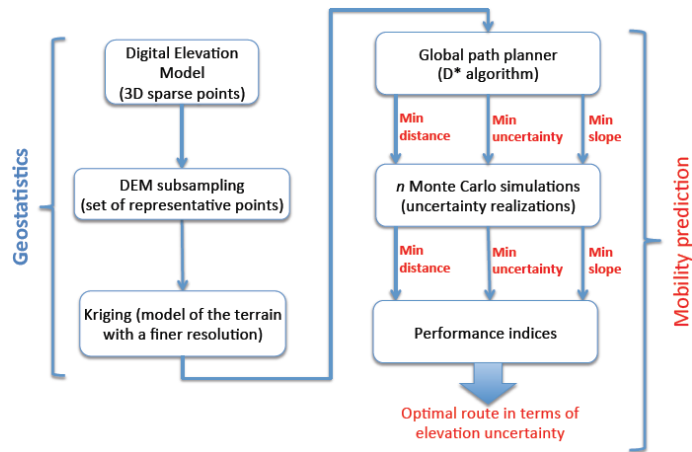

Fig. 1 Schematic view of the steps carried out in this research for predicting the mobility of a ground vehicle over a large region ($> 5 \times 5[Km^2]$).

4 Reconstruction of a 3D Terrain Surface from Remote Data. Spatial Prediction Problem

This section introduces the geostatistical tools used in this research (Subsection 4.1). The use of geostatistical tools is motivated because the original DEM has a resolution that is too coarse for vehicle mobility prediction (> 30 meters). Thus, an interpolation strategy is required in order to obtain a model of the terrain at a finer resolution. In our case, the desired resolution is the double of the vehicle length, that is, 10 meters (Humvee's length is 5 meters, see Section 6 for details, and Remark 1 for a discussion about this reasoning). The most common approach is ordinary kriging, which estimates the elevation of points (at a finer resolution) depending on the elevation and spatial arrangement of measured points (variogram). However, before attempting kriging, a subsampling step is required in order to reduce the number of points of the original DEM (Subsection 4.2). This step is necessary to perform efficient geostatistics-related computations on a standard-performance computer (as remarked in the introductory section).

It bears mentioning that in this research DEM of the terrain are employed. Since the dimensions of the grid are known and the number of observations in each row is known, the implicit spatial relationships between elevation values can be determined. There are mainly two types of DEM depending on the coverage and on the resolution. The US Geological Survey (USGS) agency produced a DEM with coverage of the US with a resolution of 30 meters. A worldwide coverage is offered by the SRTM NASA and NGA agencies. The SRTM DEM have a resolution of 90 meters. More information about those formats can be found in (Webgis 2015).

Remark 1 There is not a clear answer to what is the most appropriate spatial resolution in order to perform a reliable stochastic mobility prediction analysis. It is not known whether any detailed study on this issue has been performed. It appears that spatial resolution of data for 3D terrain models should be dependent on the size of the vehicle, the variability of the terrain, and on the nature of any natural or man-made obstacles that the vehicle must negotiate. In this research, a resolution of 10 meters has been considered for the interpolated model of the terrain (originally such model was 30-m or 90-m resolution) because the length of the intended vehicle (i.e. Humvee military vehicle) has a length of 5 meters. Smaller resolutions to 10 meters led to unfeasible solutions (impractical computation time). This key point will be considered in the coming research.

4.1 Fundamentals of Geostatistics

Geostatistics aims at providing quantitative descriptions of natural variables distributed in space time (Bivand et al 2013; Chiles and Delfiner 2012), (Isaaks and Srivastava 1989; van der Meer 2012), (Webster and Oliver 2007). Furthermore, it deals with a methodology to quantify spatial uncertainty. Geostatistical modeling is based on the statistical relationships among the measured points. It then produces an interpolation function based on a covariance or variogram model derived from the data, rather than an *a priori* model of the interpolation function (as it is the case of deterministic interpolation) (Bohling 2005). For that reason, geostatistics is data-driven rather than model-driven.

4.1.1 Variogram

A variogram describes the statistical relationship between points as a function of point separation. That is, it shows how the dissimilarity between $z(s)$ and $z(s+h)$ evolves with the separation, or lag distance, h (Chiles and Delfiner 2012; Isaaks and Srivastava 1989), (Lakhankar et al 2010). In this sense, the variogram represents the spatial correlation between available samples, and it is calculated according to:

$$\gamma(h) = \frac{1}{2N} \sum_{i=1}^{N} \big(z(s_i) - z(s_i + h) \big)^2, \qquad (1)$$

where N is the number of data points, $z(s_i)$ is the elevation at point s_i, and h represents the lag, that is, the vector separating the locations of two variables (ASTM Standards 1996). For instance, $h = 1 \times d, \ldots,$

d being the distance between points. If points are sampled at 10 meters, then $d = 10$, and the variogram is calculated for the "lag": $[10, 20, 30, \ldots]$.

An empirical variogram provides information on the spatial autocorrelation considering the raw data, but does not provide information for all possible directions and distances. For that reason, it is necessary to fit a model (i.e. a continuous function) to the empirical variogram. There are several variogram models, with the most common ones: spherical, exponential and gaussian. More information about theoretical variogram models can be found in (Chiles and Delfiner 2012).

Some important features of a variogram model are:

– It starts at zero (for $h = 0, z(s_i + h) - z(s_i) = 0$). Except when there is a "nugget effect", for further detail (Chiles and Delfiner 2012).
– It generally increases with h.
– *Range* reflects the degree of dissimilarity of ever more distant samples. If the variogram stabilizes at a certain level, such level is called *sill*. In other words, the range represents the value of the curve in the x axis when the variogram reaches steady state, and the sill is the value of such curve but in the y axis.
– Various ways can be considered to fit a variogram model, including (Chiles and Delfiner 2012), (Gorsich and Genton 2000): manual fitting, least squares (automatic fitting), and maximum likelihood. In this research, an automatic least squares approach has been used.

Notice that it is expected the points that are close assume close values because these values were generated under similar physical conditions ("same geological environment"). In contrast, at long distances the geologic conditions are different, and greater variations are to be expected.

4.1.2 Ordinary Kriging Formulation

Once a variogram model has been fitted to sparse data, predictions can be made using the Kriging formulation. There are mainly three formulations of kriging depending on whether the mean of the data is known (simple kriging) or not (ordinary kriging and universal kriging). In the field of geostatistics, the most common approach is ordinary kriging (OK) (Li and Heap 2011), (Srivastava 2013; Tsui et al 2013), (Webster and Oliver 2007) because the mean is not known a priori and because only spatial variability is meaningful. Following the OK formulation, at every point where the elevation is unknown, it is interpolated following a weighted linear combination of the available samples

(Isaaks and Srivastava 1989)

$$\hat{Z}(s_0) = \sum_{i=1}^{N} w_i Z(s_i), \qquad (2)$$

where $\hat{Z}(s_0)$ is the estimated elevation of the sample located at s_0 (new interpolated point), N is the number of samples, and w_i are the weights for the linear combination of known samples $Z(s_i)$. The index i represents the samples available in the dataset.

Ordinary kriging minimizes the mean squared error of prediction

$$\min \sigma_e^2 = \mathbb{E}\left[Z(s_0) - \sum_{i=1}^{N} w_i Z(s_i)\right]^2 \; s.t. \sum_{i=1}^{N} w_i = 1. (3)$$

As shown in (Isaaks and Srivastava 1989), the mean squared error can be written as

$$\sigma_e^2 = 2\sum_{i=1}^{N} w_i \gamma(s_0 - s_i) - \sum_{i=1}^{N}\sum_{j=1}^{N} w_i w_j \gamma(s_i - s_j), \quad (4)$$

recall that $\gamma(h)$ is the variogram model.

After differentiating the previous equation with respect to the weights and setting the derivatives equal to zero, it leads to

$$\sum_{j=1}^{N} w_j \gamma(s_i - s_j) + \lambda = \gamma(s_0 - s_i), \qquad (5)$$

where λ is the Lagrange parameter used for ensuring $\sum_{i=1}^{N} w_i = 1$ (Isaaks and Srivastava 1989). Using matrix notation the previous system of equations can be written as

$$W = C^{-1}D, \qquad (6)$$
$$W = [w_1, w_2, \ldots, w_N, \lambda],$$
$$C = \begin{cases} \gamma(s_i - s_j), & i = 1, \ldots, N; \, j = 1, \ldots, N, \\ 1, & i = N+1; \, j = 1, \ldots, N, \\ 1, & i = 1, \ldots, N; \, j = N+1, \\ 0, & i = N+1; \, j = N+1. \end{cases}$$
$$D = [\gamma(s_0 - s_1), \gamma(s_0 - s_2), \ldots, \gamma(s_0 - s_N), 1]^T.$$

Once the weights are calculated, the predicted elevations at the unknown locations can be estimated using Eq. (2). Additionally, the error variance (or kriging variance) for each predicted point is obtained using Eq. (4). This kriging variance represents the uncertainty associated with each *new* point. Stochastic analysis of the possible realizations of those points can then be performed.

4.2 Subsampling a DEM

The reliability of an experimental variogram is affected by the number of samples or its inverse, the density of data. Evidently when all the points of the DEM are considered, the variogram is obtained precisely. However, by increasing the number of samples the computation time required for calculating the variogram (and the subsequent Kriging-based interpolation) increases as well. In this sense, this section proposes a methodology that reduces dramatically the computation time for calculating a variogram despite a high correlation with respect to the variogram obtained if all the point of the DEM were employed (see Remark 2).

Remark 2 As remarked in the introductory section, one key requirement of the proposed approach deals with an efficient computation. Efficiency is understood in terms of being able to run the proposed methodology in a general-purpose computer in a reasonable computation time (around 1-2 hours depending on the size of the environment considered). The reason why the subsampling step has been added to this methodology is because when the variogram and ordinary kriging functions were run over the entire datasets (e.g. the two scenarios introduced in Section 6), the computation time increased dramatically (several hours), even the computer ran out of RAM memory (Intel Core i7 3GHz, 16 MB RAM). After adding this subsampling step, before running kriging, we were able to successfully apply kriging and at a reasonable computation time.

4.2.1 Geostatistical Subsampling-Literature Review

The standard way to estimate the variogram is through the Method of Moments (MoM) (Chiles and Delfiner 2012), Eq. (1). In (Webster and Oliver 2007), the authors point out that in order to estimate a variogram following this method more than 100-150 samples are required. The MoM is followed in numerous papers where a set of discrete samples (100-300 samples) are used for estimating the variogram (Basaran et al 2011; Bechler et at 2013), (Hosseini et al 2014).

The work (Brus and Gruijter 1994) suggests a procedure based on sampling theory in order to determine the samples used for calculating the variogram. This procedure is based on repeatedly selecting pairs of points; the first point is randomly chosen, and the second point is chosen at distance h from the first one. Notice that many publications follow a similar approach (simple or systematic random sampling), see for example (Anderson et al 2006; Lakhankar et al 2010).

As shown in this paper, systematic sampling is effective for variogram estimation, but shows a poor performance in terms of fixing the elevation profile of the entire DEM. Recall that this issue is fundamental in case of subsequent kriging (i.e. obtaining a new DEM at a finer spatial resolution). Additionally, we demonstrate

that the precision of a variogram does not only depend on the number of samples. In fact, it depends on the sampling method. Additionally, those papers neither present an analysis of the computation time required nor statistical evidence of how well the samples fit the variogram of the entire dataset.

On the other hand, the residual maximum likelihood method (REML) constitutes a methodology for obtaining a variogram (Kerry et al 2010), (Webster and Oliver 2007). In this case, the three parameters that define a variogram (range, sill, and nugget) are solved by means of an optimization problem (maximum likelihood problem). Specifically the model parameters are obtained directly from the generalized increments of a covariance matrix of the full data. However, as explained in (Webster and Oliver 2007): it is computationally intensive; it is only valid for three variogram functions; and it too readily converges to a local optimum. For those reasons, we focus on the known method of moments in order to estimate the variogram. Additionally, it bears mentioning that REML has been applied to identify the optimal sampling interval in order to get a map or set of samples before attempting kriging (Kerry et al 2010). This issue is not exactly applied to the subsampling problem that we attempt in this research.

4.2.2 Systematic Sampling

The systematic sampling method arranges the population according to some ordering scheme and then selecting elements at regular intervals (Thompson 2012). This strategy matches properly with the structure of a DEM, and with the research addressed here. Recall that in a DEM information is recorded at regularly spaced intervals (raster form). For example, if a USGS DEM sampled with a resolution of 30 meters is considered, a sample would be selected every 30 meters.

### 4.2.3 Stratified Sampling via K-Means Algorithm

The stratified sampling approach is based on partitioning a given population into homogeneous subgroups or "strata". Each stratum is then sampled as an independent sub-population, out of which individual elements can be randomly selected. The principle of stratification is to partition the population in such a way that the units within a stratum are as similar as possible (Pengelly 2002; Rubinstein and Kroese 2007).

The major concern regarding the stratified sampling approach is how to separate the original DEM into different strata. In our context, one reasonable approach is to find such strata depending in elevation. However,

this issue is not simple because of the continuous nature of the elevation, especially in uneven terrain. To solve this problem, K-means clustering has been considered (Wu 2012). Notice that K-means has already been applied in geostatistics, see for example (Arieira et al 2011; Kumar et al 2011). In such publications, K-means is mainly used for stratifying a large region into similar areas mainly in terms of habitat type and soil type. Here, the K-means algorithm is used to split the DEM in terms of elevation. After that, one or two points are randomly selected within each strata. Figure 2 shows the application of the proposed stratified sampling method to the two scenarios considered in this research (those scenarios are introduced in Section 6).
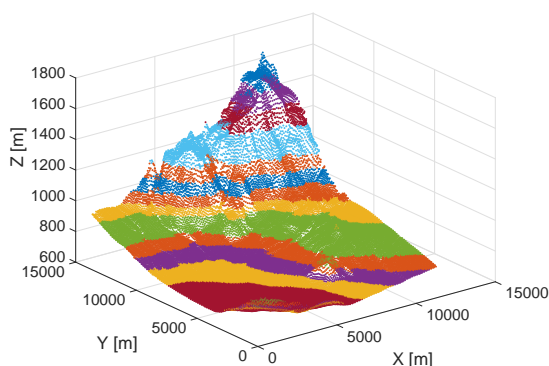
(a) Death Valley (CA, USA). No. strata = 15
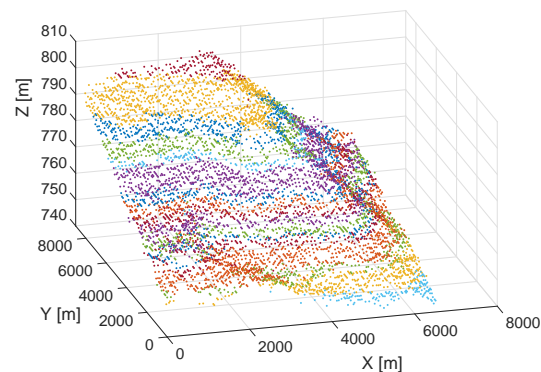
(b) Sahara desert (Chad, Africa). No. strata = 20

Fig. 2 Stratification of two types of DEM. The first one deals with a 30-m resolution DEM and the second example is related to a 90-m resolution DEM. Colors represent different layers or strata.

The key issue of this approach deals with the number of centers used by the K-means algorithm. This is a major issue related to any kind of algorithm based on K-means (Wu 2012). In this paper, the optimal value is

obtained according to the theoretical variogram fitted to the DEM (explained subsequently).

Notice that a method called DP-Means has been recently proposed in order to solve this issue (Kulis and Jordan 2012). Specifically, DP-Means uses a scale/penalty parameter to control the creation of new clusters during clustering rather than requiring a fixed K as in K-means. In this way, a new cluster is only formed when a point is sufficiently far away from all existing cluster centroids. It bears mentioning that this solution is not applicable to our problem because we do not need to cluster the elevation profile according to a fixed number of strata. In fact, we are interested in testing different strata/samples in order to reduce the error in fixing a given variogram model (i.e. the variogram model of the entire DEM).

4.2.4 Proposed Methodology: Combining Systematic and Stratified Sampling

Recall that the goal of the subsampling approach is to find the minimum number of samples required such that the variogram of this reduced DEM fits the theoretical variogram of the raw DEM. In order to solve this issue the properties of the variogram and the sampling methods must be taken into account.

First, the variogram is calculated by Eq. (1). To obtain a meaningful set of samples the reduced data set should be as sparse as possible in order to cover the whole range of distances in the original DEM. If this fact is not ensured, the variogram would not fit properly the sill. On the other hand, the set of samples should comprise the range of elevations presents in the original DEM as much as possible. This issue has an effect on the range of the variogram. Additionally, we are interested in fitting the elevation profile in order to obtain a meaningful model after applying the Kriging process.

One proposal of combining systematic and stratified sampling approaches is motivated by the following fact. Notice that the systematic sampling approach leads to a sparse set of samples, which cover the entire dataset (Figure 3a). This feature is not found using the stratified sampling approach (Figure 3b). In fact, many samples are clustered at the top right corner. This is explained because of the hill in the environment (this experiment is related to the Airport Lake region, which is introduced in detail in Section 6). Recall that the stratified sampling approach deals with getting samples regarding the different strata (elevations) in the environment, those strata are mainly present in the hill. On the other hand, the stratified sampling approach captures properly the elevations in the original DEM

as observed in Figure 2a, but the systematic sampling approach has a worse result in terms of representing the terrain profile (i.e. elevations); see Section 6.2.3 for a quantitative comparison of these methods. A filter has been implemented to remove those samples that are very close, and hence, to avoid redundancy in the samples.
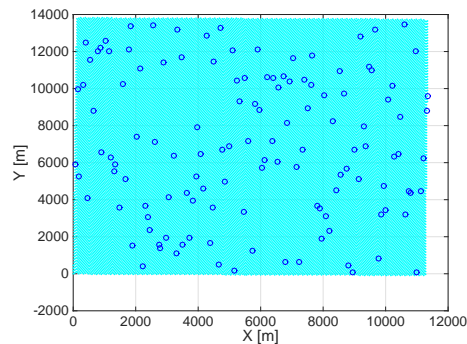
(a) Systematic sampling, 127 samples
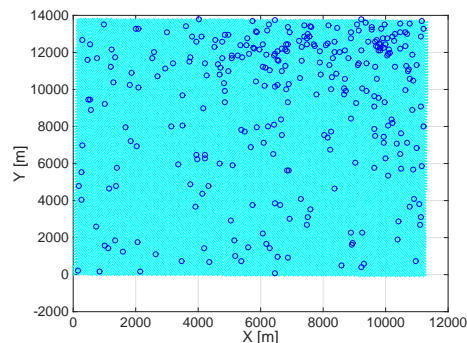
(b) Stratified sampling, 300 samples

Fig. 3 Samples obtained after applying the systematic and stratified sampling approaches to one of the scenarios considered in this work (Death Valley, CA, USA), see Section 6 for further details.

5 Mobility Prediction based on Elevation Uncertainty

As already mentioned, the kriging procedure not only produces a model of the environment at a finer resolution than the original but also, and most importantly, the uncertainty associated with each new point (error estimate), see Section 4.1.2.

The next step deals with analyzing the performance of a vehicle moving between two given points, starting point and goal point, considering this new model of the terrain. The D* algorithm has been employed in order to obtain the optimal route between the starting and goal points (Corke 2011; Stentz 1995). In this

research, three different cost functions has been considered in order to obtain the optimal route. The first cost finds the shortest route between the starting point and the goal point. For that purpose, the D* cost function considers the Euclidean distance between points (x-y plane). Notice that, in this case, uncertainty is not considered. This route is labeled as "min-distance". The second route is obtained as the shortest distance between the starting and goal points but also minimizing the uncertainty. That is, the variance associated to each point is also considered in the D* cost function. From now on, this route is called "min-uncertainty". Finally, the route labeled as "min-slope" represents the shortest route between the starting and goal points but also minimizing the slope between points. It bears mentioning that such slope is obtained estimating the elevation among the 8 neighbors around one point (i.e. the angle). After that, the point with the maximum angle is saved in an array paired with the sampled point. It means that when the D* algorithm is employed those angles will be minimized and it will result in the flattest route.

The following issue deals with checking the performance of the vehicle over such routes. In this sense, the kinematic model of a car-like vehicle is considered (LaValle 2006). Furthermore, a motion controller is required, in this case, the well-known pure pursuit path follower has been taken into account (Amidi 1990). This step will allow us to estimate the energy required to actually traverse each route.

Finally, n Monte Carlo simulations are performed in order to analyze the effect of uncertainty in the energy required to traverse each route. This idea is inspired by the following references in the context of soil and Earth sciences: (Davis and Keller 1997; Fisher 1991), (Hunter and Goodchild 1997). Specifically, a realization for each sample within the desired route is generated according to its associated uncertainty (Kriging estimate and Kriging variance).

5.1 Performance Indices

One key aspect in order to perform the stochastic mobility analysis of each route is related to the performance indices considered. In this sense, five indices have been selected: (i) length of the route, (ii) average value of the uncertainty, (iii) variation of the elevation, (iv) energy required to traverse each route, (v) average value of the slope considering the kriging variance. The last index has been obtained simulating different realizations of the elevation of each point (according to its associated uncertainty). Thus, it represents the average slope between points after n Monte Carlo simulations

(the absolute value of the difference has been considered in order to avoid problems with the sign).

The energy is calculated according to the following equation

$$E_j = \sum_{m=1}^{2} \sum_{k=1}^{N_j} \tau_m(k)v_m(k), \tag{7}$$

where $j = 1, 2, 3$ means min-distance, min-uncertain, min-slope route, respectively; m represents the steering or the traction motor, N is the number of samples of each route it is related to the length of the route), τ is the torque, and v is the linear velocity. Notice that the torque required for each motor has been obtained from the following equation

$$\tau_m = k_t \frac{V_s - k_i\omega_m}{R}, \quad \forall m = 1, 2, \tag{8}$$

where k_t is the torque constant of the motor, V_s is the nominal voltage, k_i is the inverse of the speed constant, ω is the angular velocity, and R is the terminal resistance. Practical details about those constants are given in Table 2, see Section 6.4.

6 Illustrative Examples

6.1 Environments

In order to validate the methods proposed in this research, two different scenarios have been analyzed[2]. First, a region of the Death Valley called Airport Lake (LAT: 35 52 30 N; LONG: -117 37 30 W), see Figure 4. The data were obtained in the 7.5-Min USGS format. This region covers an area of $11.34 \times 13.65\,[Km^2]$. It is important to mention that this DEM is represented in projected coordinates (i.e. UTM system). On the other hand, a region of the Sahara desert in Chad (Africa) has also been considered (LAT: 15 01 67 N; LONG: 21 16 50 E), see Figure 5. The data were obtained in the SRTM format. This region covers an area of $6.55 \times 8.18\,[Km^2]$. As in the previous case, this DEM is represented in projected coordinates (i.e. UTM system).

Figure 4c and Figure 5c show the raw DEM related to the Airport Lake and the Sahara desert. The difference between them is that in the former the points are sampled every 30 meters and in the latter are sampled every 90 meters. For that reason, the first DEM is denser than the second one.

To calculate the variograms and applying Ordinary Kriging the Matlab *mGstat toolbox* has been employed (mGstat 2015). Additionally, the SRTM DEM was first
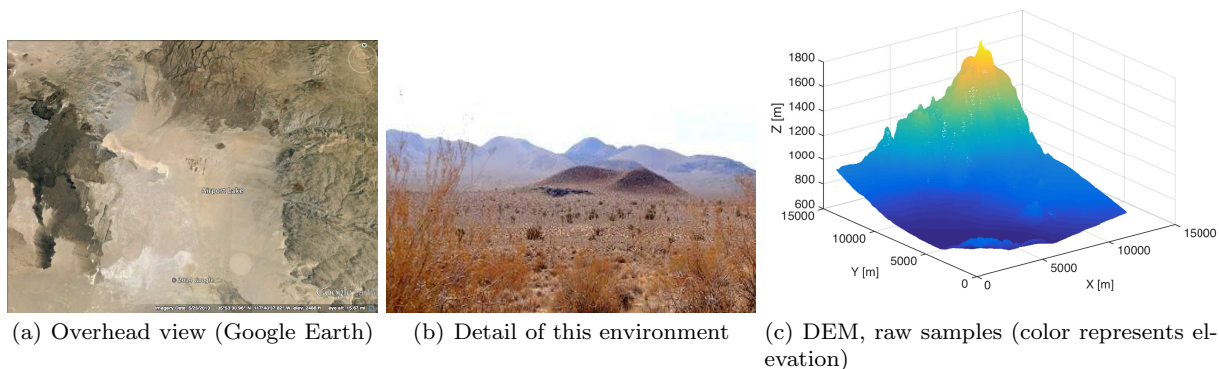
____
[2] Both DEMs were downloaded from (Webgis 2015)

(a) Overhead view (Google Earth)   (b) Detail of this environment   (c) DEM, raw samples (color represents elevation)

Fig. 4 Airport Lake, Death Valley (CA, USA). USGS DEM (30-m resolution).

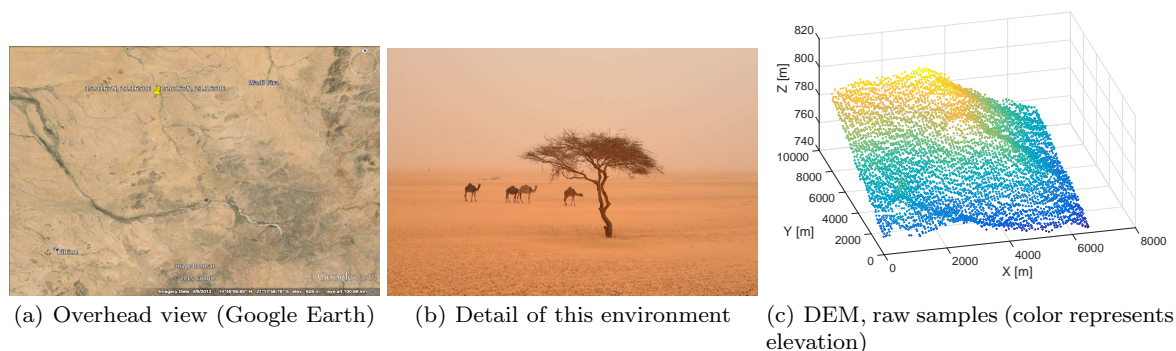(a) Overhead view (Google Earth)   (b) Detail of this environment   (c) DEM, raw samples (color represents elevation)

Fig. 5 Sahara desert (Chad, Africa). SRTM DEM (90-m resolution).

processed in ArcGIS suite (Esri) before using it in Matlab (spherical coordinates to UTM). The stratified sampling approach has been implemented taking advantage of the *VL-Feat toolbox* (Vedaldi and Fulkerson 2015). The D* algorithm has been implemented using the *Robotics toolbox* (Corke 2011).

6.2 Subsampling Algorithms

In this section the performance of the subsampling approaches before attempting Kriging and the mobility prediction analysis is examined. In particular, the following sampling approaches are compared: systematic sampling, stratified sampling, and the method based on combining systematic and stratified approaches. These methods have been applied to the Airport Lake and the Sahara desert regions.

It bears mentioning that the tuning parameter (i.e. the number of points or strata) of each subsampling approach has been selected in order to maximize the following metrics: minimum number of samples, minimum computation time, minimum error in range and sill, maximum correlation with the variogram model (Pearson's correlation coefficient), and maximum corre-

lation with the elevation profile (Pearson's correlation coefficient).

6.2.1 Experiment 1. Airport Lake (Death Valley, CA, USA)

This section details the performance of the subsampling approaches considering the Airport Lake. The stratified sampling has been configured with 300 strata and one random sample has been selected within each strata (K-Means = 300). The systematic approach has been configured in order to obtain a random sample every 90 meters leading to 127 samples. Additionally, two configurations of the proposed method are tested. The first one considered 15 strata and two random points for each strata plus 127 samples from the systematic sampling (156 points, 1 point was filtered). The second configuration deals with 15 strata and one random sample per strata plus 127 samples from the systematic sampling (142 points).

Firstly, we estimate the theoretical variogram of the entire DEM. In this case, the raw DEM is composed of 34741 samples leading to a computation time of 44 minutes (Intel Core i7 3GHz, 16 MB RAM). This vari-

ogram is automatically fitted, through the least squares method, by a Gaussian model (5th-order polynomial).

Figures 6a, 6b show the first experiments carried out in order to check the performance of the systematic sampling approach (127 samples, 0.1 seconds) and the stratified sampling approach (300 samples, 0.2 seconds), respectively. The empirical variograms do not fit the Gaussian model that represents the entire DEM (black curve). An interesting conclusion from these experiments is that the stratified sampling approach fits the range of the Gaussian model (5167 versus 5273), but not the sill (81164 versus 29814). In contrast, the systematic sampling method has a smaller error in the sill (21371 versus 29814), but there is a deviation in the range (4841 versus 5273). Notice that the units of the sill and range are $[m^4]$ and $[m]$, respectively.

Figures 6c, 6d displays the result obtained with the proposed subsampling approach (156 and 142 samples). Observe that the empirical variograms fit better the variogram model obtained with the entire DEM. Especial mention is for the case where two random samples are selected within each strata (Figure 6c). Here, as the range as the sill fit the theoretical variogram model (5089 versus 5273; 28529 versus 29814). Notice that a deep quantitative analysis is addressed in Section 6.2.3.

6.2.2 Experiment 2. Sahara Desert (Chad, Africa)

This section focuses on the performance of the subsampling approaches considering the Sahara desert. The stratified sampling has been experimentally configured with 65 strata and three random samples per strata leading to 195 samples. The systematic approach has been configured in order to obtain 73 samples. The two configurations of the proposed subsampling approach are: 23 strata with one random point for each strata plus 73 samples from the systematic sampling. The second configuration is 23 strata with two random samples per strata plus 73 samples from the systematic sampling.

As in the previous experiments, the theoretical variogram of the entire DEM is first calculated. In this case, the raw DEM is composed of 6660 samples leading to a computation time of 1 minute and 20 seconds (Intel Core i7 3GHz, 16 MB RAM). Figures 8a, 8b show the experiments carried out with the systematic sampling and the stratified sampling approach, respectively. The main result is that the empirical variograms do not fit the Gaussian model that represents the entire DEM. A similar behavior to the previous scenario is observed, that is, the stratified sampling approach fits the range of the Gaussian model (5167 versus 5273), but not the sill (81164 versus 29814). In contrast, the systematic

sampling has a smaller error in the sill (21371 versus 29814), but there is a deviation in the range (4841 versus 5273).

Figures 8c, 8d display the result obtained with the proposed subsampling approach (156 and 142 samples). Observe that the empirical variograms fit the variogram model obtained with the entire DEM. The best case is obtained when two random samples are selected within each strata (Figure 8c). Here, as the range as the sill are similar to the theoretical variogram model (5089 versus 5273; 28529 versus 29814).

6.2.3 Discussion about the Subsampling Experiments

In this section we analyze the results obtained with the subsampling approaches in terms of the following metrics: number of samples, computation time, range and sill of the variograms, Pearson's correlation coefficient between the original variogram and the variogram obtained for each sampling approach, Pearson's correlation coefficient between each elevation profile, and the mean of the squared residuals between the experimental values and the fitted variogram model. In order to calculate the Pearson's correlation coefficients a 5th-order polynomial has been fitted to the empirical variogram and to the elevation profile in each case. Recall that those experiments were run in a general-purpose computer (Intel Core i7 3GHz, 16 MB RAM).

Table 1 shows that the subsampling proposed methodology is worthy in terms of computation time and without reducing the precision of the empirical variogram. Notice that the best result is obtained with the two configurations highlighted in the table (156 samples for the Airport Lake and 96 samples for the Sahara desert). The deviation from the original DEM is the smallest one considering all the metrics. It bears mentioning that the maximum similarity in the elevation profile is usually achieved by the stratified sampling approach. However, this approach yields a poor correlation in terms of the variogram (Pearson's correlation coefficient). This research also draws the following points:

- The number of samples is not directly related to the confidence of the empirical variogram. We demonstrate that 156 samples (Airport Lake) and 96 samples (Sahara desert) are better than 300 samples and 195 samples, respectively (combined subsampling approach versus stratified subsampling).
- Regarding the sampling step, it is important to take into account not only the elevation of the samples (stratified sampling), but also the distance between points (systematic sampling). The solution that leads to the best metrics is in fact a combination of both approaches (Pearson's correlation coefficient).
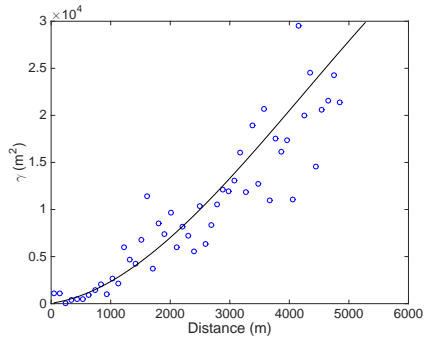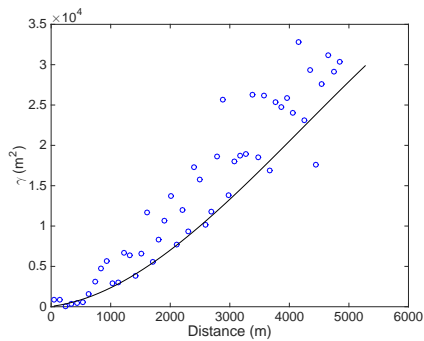
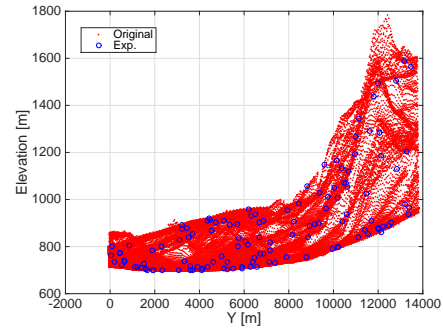(a) Systematic sampling, 127 points
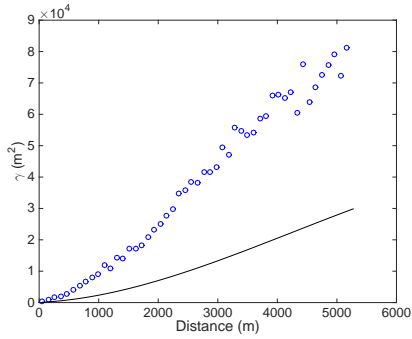
(b) Stratified sampling, 300 points
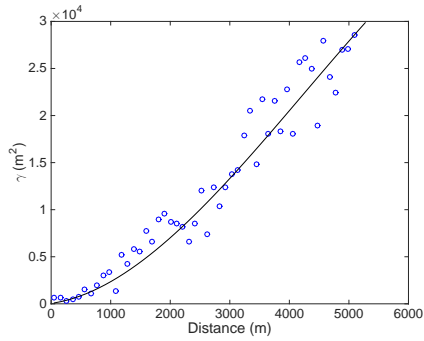
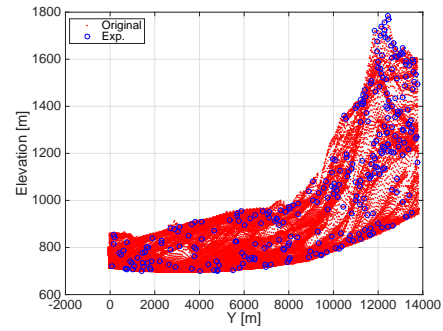
(c) Combined sampling, 156 points
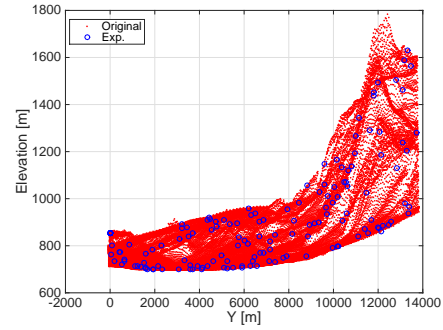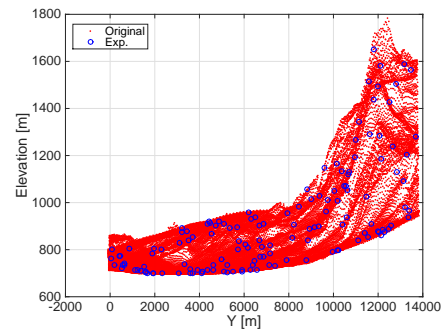

(d) Combined sampling, 142 points

Fig. 6 Variograms (Airport Lake). The continuous line is the theoretical variogram considering all the raw samples, the blue circles mean the empirical variogram of each method

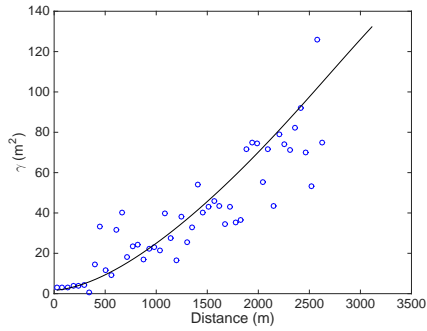(a) Systematic sampling, 127 points

(b) Stratified sampling, 300 points

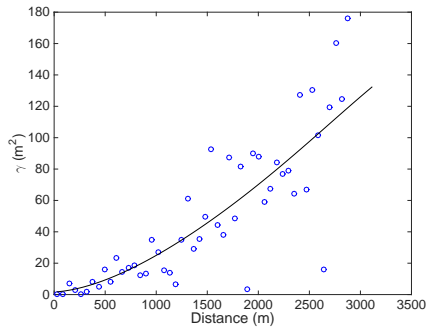
(c) Combined sampling, 156 points

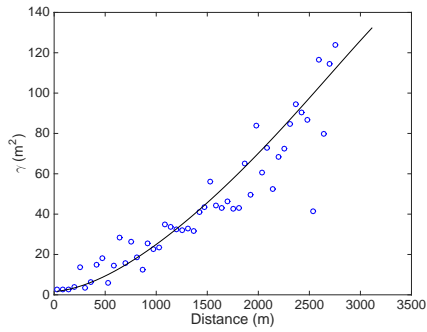(d) Combined sampling, 142 points

Fig. 7 Elevations (Airport Lake). Notice that the circles represents the subsampled points and the red profile the entire data set (raw samples).

11

(a) Systematic sampling, 73 points

(b) Stratified sampling, 195 points

(c) Combined sampling, 96 points

(d) Combined sampling, 117 points

Fig. 8 Variograms (Sahara desert). The continuous line is the theoretical variogram considering all the raw samples, the blue circles mean the empirical variogram of each method
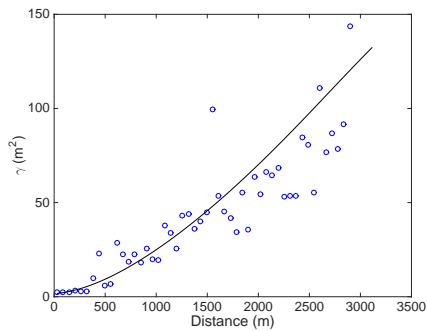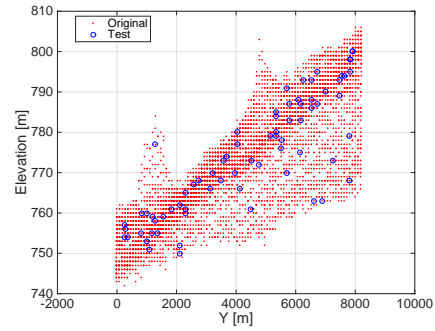

(a) Systematic sampling, 73 points
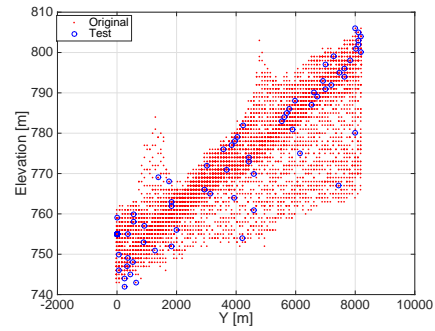
(b) Stratified sampling, 195 points

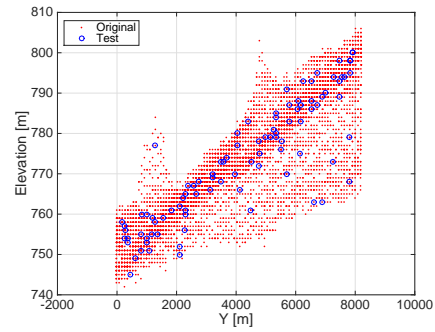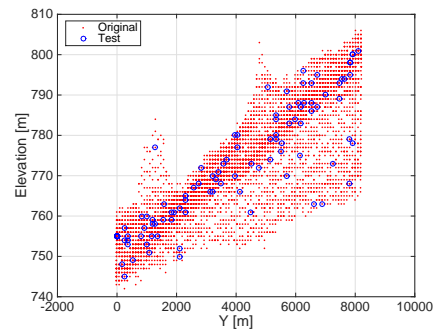(c) Combined sampling, 96 points

(d) Combined sampling, 117 points

Fig. 9 Elevations (Sahara desert). Notice that the circles represents the subsampled points and the red profile the entire data set (raw samples).

12

- The computation time can be dramatically reduced by selecting the appropriate samples, instead of considering the entire DEM (44 minutes versus 0.1 seconds, in case of the Airport Lake environment).
- The suitability of the combined sampling approach is demonstrated in terms of the range and sill of the variogram and the elevation profile.

It is important to remark that a slight variation in the shape of the variogram (small deviation in the range or sill) has no effect on the ordinary kriging weights and the kriging estimates (elevations). It only affects the kriging variance (Isaaks and Srivastava 1989). For that reason, considering the reduced-order datasets will not affect the kriging performance.

6.3 Ordinary Kriging and Global Path Planning

Once a set of samples have been selected representing the raw DEM, the next step deals with applying the ordinary kriging algorithm in order to obtain a new model of the terrain at a finer resolution. In this case, 10-m resolution models are desired, which leads to an affordable computation burden (see Remark 2 and discussion in Section 7). After that, the D* algorithm is applied in order to obtain the optimal route between two desired points. Recall that three cost functions are evaluated: min-distance (the shortest route), min-uncertainty (the route with the minimum uncertainty), and min-slope (the flattest route).

Figure 10a details the result of applying ordinary kriging to the Airport Lake environment, and Figure 10b deals with the Sahara desert. Notice that the generated models fit properly the original DEMs, see Figure 4 and Figure 5, respectively. Furthermore, it is important to highlight the small standard deviation (uncertainty) obtained. In both environments it is smaller than 0.2 [m] (blue color). Brighter colors mean larger uncertainty, notice that they appear near the boundaries of the surfaces because no samples are available for those regions (from the subsampling step). The range in the standard deviation associated to the Death Valley model is (0, 15.67) [m], the mean value is 0.2 [m] with standard deviation 0.81 [m]. For the Sahara scenario the range is (0, 4.48) [m], the mean value is 0.05 [m] with standard deviation 0.18 [m].

After obtaining the new models of the terrain with kriging, the D* path planning algorithm is used to obtain the optimal route according to a cost function. Figure 11 shows the min-distance and the min-uncertainty routes for the Airport Lake. As expected, the shortest route (straight-line) corresponds to the min-distance line (red line). The min-uncertainty route considers the
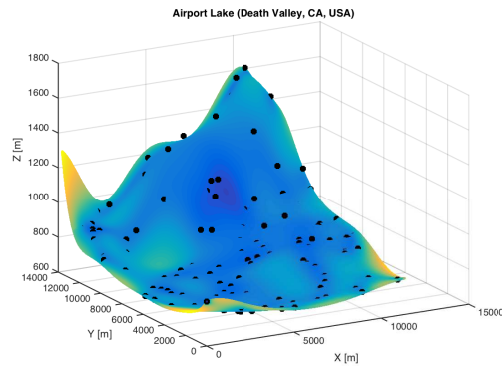
(a) Airport Lake (Death Valley, CA, USA)
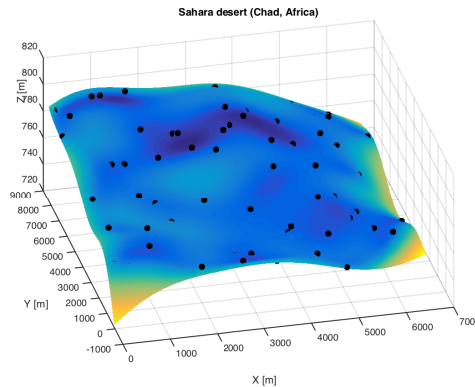
(b) Sahara desert (Chad, Africa)

Fig. 10 Surface reconstruction using ordinary kriging (stochastic model of the terrain). The original resolutions were 30-m and 90-m, respectively. The new digital elevation models have a resolution of 10 meters. Notice that blue color means low uncertainty, < 0.2 [m] for Death Valley and < 0.05 [m] for Sahara desert.

variance of the elevation (uncertainty obtained from kriging). For that reason this route passes as close as possible to the original sampled points (black dots).

Figure 12 displays a deterministic terrain profile illustrating the minimum slope between points (8-neighbors to each point). In this sense, a path going through a brighter region (yellow) would mean a flatter route (small variation in the elevation between one point and its neighbor). On the other hand, hazards such as high slopes are represented by blue or red color, that is, the difference in elevation between one point and its neighbors is larger than in a brighter region. Notice that positive values (red color) mean positive slopes (the vehicle would pitch up), and negative values (blue color) represent negative slopes (the vehicle would pitch down).

Figure 13 shows the three routes obtained with the D* algorithm in the x-y plane. Observing these plots is even easier to understand the difference between the three cost functions. For example, the min-distance route

Table 1 Summary of the performance of the subsampling approaches

Airport Lake	No. points	Comp. time	Range (m)	Sill (m²)	ρ (variogram)	ρ (elev. profile)	Mean residual (m⁴)
Original DEM	34741	44 min.	5273	29814	-	-	8.5e3
Systematic sampling	127	0.1 sec.	4840	21370	0.9988	0.9809	1.13e7
Stratified sampling	300	0.2 sec.	5167	81164	0.9916	0.9984	9.41e8
Combined sampling (2 points/strata)	156	0.1 sec.	5089	28529	0.9991	0.9862	4.73e6
Combined sampling (1 point/strata)	142	0.1 sec.	4841	30332	0.9927	0.9882	2.22e7
Sahara desert	No. points	Comp. time	Range (m)	Sill (m²)	ρ (variogram)	ρ (elev. profile)	Mean residual (m⁴)
Original DEM	6660	1 min. 20 sec.	3112	132	-	-	1.97e-1
Systematic sampling	73	0.1 sec.	2626	75	0.9942	0.9868	2.14e4
Stratified sampling	195	0.1 sec.	2876	176	0.9757	0.9887	5.70e2
Combined sampling (1 point/strata)	96	0.1 sec.	2756	124	0.9942	0.9862	1.54e2
Combined sampling (2 points/strata)	117	0.1 sec.	2897	143	0.9738	0.9944	3.16e5

follows a straight line between the starting and the goal points, which is expected. However, the min-uncertainty route follows a different pattern. Although a longer path is generated, the ground vehicle would move over a safer route where more "certain points" are traversed (black points). The min-slope route also takes a different path.

Figure 14 shows the min-distance and the min-uncertainty routes regarding the Sahara desert scenario. Again the shortest route (straight-line) corresponds to the min-distance line. As expected, the min-uncertainty route passes fairly close to the sampled points, although it means a longer path.

Figure 15 displays the deterministic terrain profile dealing with the minimum slope between points. Notice that compared to the previous scenario the Sahara desert is a flatter scenario, the elevation range is smaller than 100 meters. For that reason, the slope angles are smaller than in the previous environment (range 2.5 to

-1.5 [deg]). This fact explains why the min-slope route is quite similar to the min-distance route.

Finally, Figure 16 shows the three routes obtained with the D* algorithm in the x-y plane. Notice the similarity between the min-distance and the min-slope routes. The main difference still takes place with the min-uncertainty route. As in the previous scenario, a longer path is generated, but with the benefit of following a safer route (smallest uncertainty regarding the true elevation of the terrain).

6.4 Mobility Prediction

This section deals with analyzing the mobility of a vehicle over the routes discussed in the previous section. Notice that a maximum linear velocity of 1 [m/s] and a maximum steering angle of 75 [deg] have been assumed for the testbed considering in these experiments. The lookahead distance of the pure pursuit algorithm is 4 [m], and the sampling period is 1 [s]. More technical
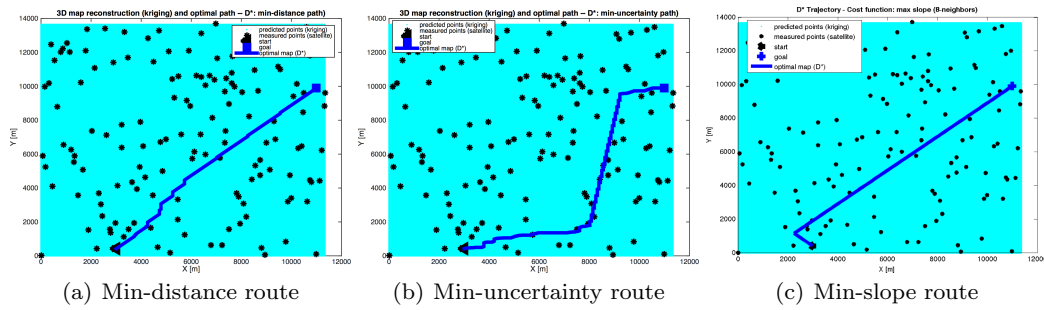
(a) Min-distance route   (b) Min-uncertainty route   (c) Min-slope route

Fig. 13 Airport Lake (USA). Comparison of the three routes obtained the D* algorithm considering the variance (uncertainty) of each point as the cost function (min-uncertainty path), considering the elevation (min-distance path), or the terrain profile (min-slope path).

(a) Min-distance route   (b) Min-uncertainty route   (c) Min-slope route

Fig. 16 Sahara desert (Chad). Comparison of the three routes obtained the D* algorithm considering the variance (uncertainty) of each point as the cost function (min-uncertainty path), considering the elevation (min-distance path), or the terrain profile (min-slope path).
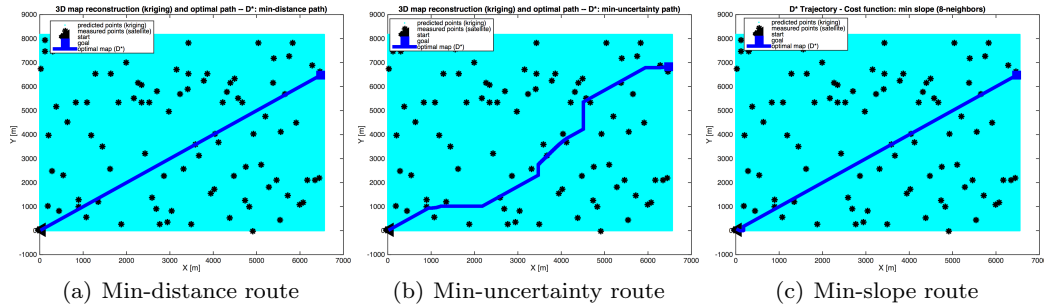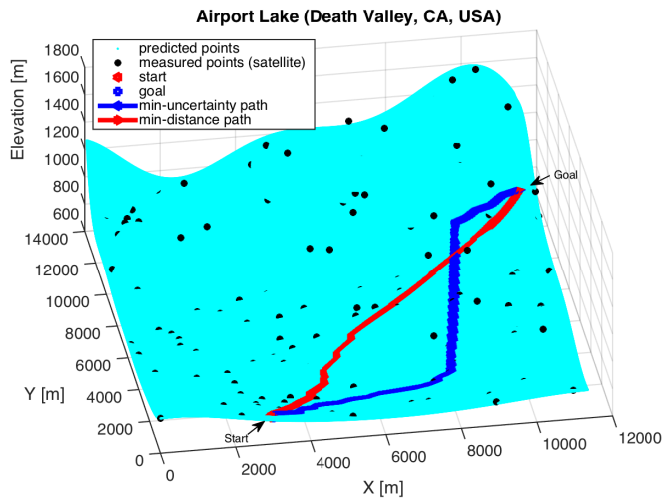

Fig. 11 Routes obtained using the D* path planner. Notice that the min-uncertainty path goes closer to the actual sampled points. In contrast, the min-distance path follows a straight line between the starting point and the goal. The mesh represents the interpolated model considering the nominal elevations (kriging estimations).

Fig. 12 Deterministic terrain profile: representation of the 3D map in terms of the maximum slope between the 8-neighbors around each point and D* optimal path minimizing the slope between points. Notice that in order to estimate this map the nominal value of the elevation of each point has been considered.

Mason DC motors, model 253298 (one motor for traction and another motor for steering).

It bears mentioning that 150 Monte Carlo simulations have been carried out for each reference route and

data about the vehicle configuration is summarized in Table 2. Furthermore, the vehicle is equipped with two

Fig. 14 Routes obtained using the D* path planner. The mesh represents the interpolated model considering the nominal elevations (kriging estimations).
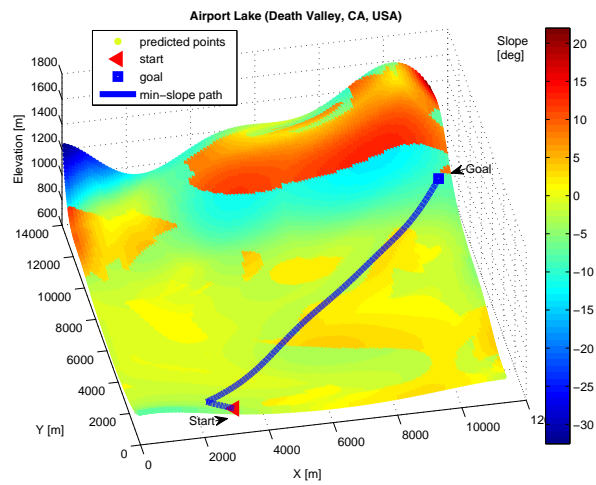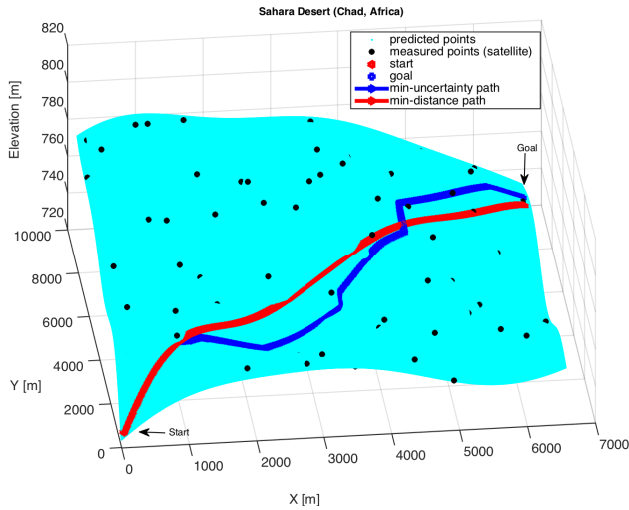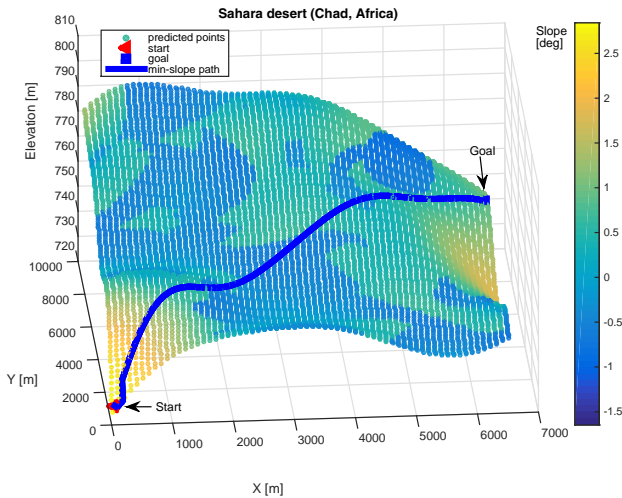

Fig. 15 Deterministic terrain profile: representation of the 3D map in terms of the maximum slope between the 8-neighbors around each point and D* optimal path minimizing the slope between points.

Table 2 Features of the robot considered in this research

Feature	Value
Wheel radius [m]	0.25
Wheelbase [m]	2.5
Terminal resistance [Ω]	0.568
Torque constant [mNm/A]	158
Speed constant [rpm/V]	60.4
Speed/Torque gradient [rpm/mNm]	0.217
Rotor inertia [gcm^2]	1320

for each scenario. It means that a random value for the elevation of each reference point has been generated

according to the mean value of the elevation and the Kriging variance.

Table 3 summarizes the performance of each route in terms of the cost functions previously introduced (see Section 5.1). As expected, the shortest route is obtained employing the min-distance cost in the D* algorithm. The longest route is obtained using the min-uncertainty cost because it was shown it pursues to follow as much "true" samples as possible.

The route with the lowest uncertainty results from the min-uncertainty cost. Recall that this value has been obtained as the mean value of the Kriging variance. For that reason, there is a clear difference between the values obtained in the first scenario and those obtained for the second environment. Uncertainty is larger in the Sahara desert because the original DEM had a resolution of 90 meters and the new model has a resolution of 10 meters. In contrast, the resolutions of the two models of the Death Valley region are much more similar, 30 meters versus 10 meters. Those results show the importance of uncertainty in elevation and its relation with the spatial resolution.

The trajectory with the lowest mean value for the elevation is given by the min-slope cost. Again, this cost function is given in terms of the average value of the mean elevation. For that reason, when the deterministic terrain profile related to the maximum slope angles between points is considered by the D* algorithm, it is not surprising that it leads to the flattest route.

One interesting conclusion is drawn from the last two columns. First, notice that the energy spent for the two motors of the vehicle (steering motor and traction motor) is fairly similar regarding the min-distance and the min-slope costs because the length of the trajectories is similar in both cases. Even in the first scenario the energy required for the min-slope route is smaller than the min-distance despite the fact it is slightly longer. This is explained because, according to the current configuration of the vehicle (maximum steering angle, maximum linear velocity, etc.), the min-distance route requires more energy to be followed. On the other hand, it bears mentioning that although the min-slope route achieves the minimum value in terms of the mean elevation it does not ensure the smallest value in the slope column. This is not altogether unexpected because the min-slope route has a larger mean uncertain value than the min-uncertainty route. It means that after 150 runs the average value of the slope is smaller when the uncertainty in elevation is smaller. Recall that the index dealing with elevation is obtained as the average of the mean elevation for each route (Kriging mean), and the column related to the slope results from

Table 3 Summary of the three cost functions considered in this research

Airport Lake (Death Valley, CA, USA)	Length [m]	Uncertainty [m]	Elevation [m]	Energy [Kw]	Slope [deg]
Min-distance	12794	0.05	830	241	2.46
Min-uncertainty	16617	0.04	853	306	2.17
Min-slope	13407	0.05	827	221	2.38
Sahara desert (Chad, Africa)	**Length [m]**	**Uncertainty [m]**	**Elevation [m]**	**Energy [Kw]**	**Slope [deg]**
Min-distance	9079	0.12	766	172	1.33
Min-uncertainty	10410	0.10	767	196	1.31
Min-slope	9216	0.19	765	174	1.59

simulating the Kriging variance associated to each point during 150 runs.

The main conclusion of these experiments is that this research demonstrates from a stochastic point of view that the min-uncertainty route is more appropriate than the min-distance and the min-slope routes if uncertainty in elevation is considered. It means that a minimum value in the slope (5th column in Table 3) would eventually lead to a lower consumption. Recall that the fourth column, called Energy, is only calculated for the x-y plane, it does not then account for all the energy required to traverse the 3D world. This issue is in fact shown through a realistic animation discussed in the following section.

6.5 Realistic animation using ANVEL

The results obtained in this section validate the main conclusion drawn from the previous experiments[3]. In particular, the DEM of the Sahara desert environment has been loaded in the robotic simulator ANVEL (Quantum Signal, http://anvelsim.com), see Figure 17a. After that, two different routes have been defined: the first route constitutes a flat path, however in the second one the vehicle moves gradually through a slope until the goal, see Figures 17a, 17b. This scenario validates the importance of slope variable analyzed in previous section and why the min-uncertainty trajectory results in the most efficient route in terms of energy saving.

It bears mentioning that the original routes have been shortened in order to perform an attractive animation ($\sim$55 [m] instead of the original $\sim$10 [Km] routes). A similar vehicle to the one used previously has also been employed in ANVEL. In particular, a Jeep vehicle with a mass of 850 [Kg] and a maximum velocity of 1 [m/s] has been employed.

---

[3] Videos about this realistic simulation are available at: http://www.ual.es/personal/rgonzalez/videos.html

As observed in Figure 17c, the power required by the flattest trajectory ("route 1") is smaller than that required by the route where the vehicle has to face a slope ("route 2") despite the fact both routes traverse the same distance, 55 meters. In particular, the power of the first trajectory is 7300 [HP], and the power required in order to traverse the second route is 10480 [HP]. This result demonstrates the conclusions drawn in the previous section, that is, a flatter terrain is preferred to a steepest route despite the traversed distance is the same. Finally, Figure 17d shows the forward speed of the vehicle during the traverses. Notice that the speed achieved during the first experiment is slightly higher than in the second case. This is not altogether unexpected because in the second route a higher slip occurs.

7 Conclusions and Future Work

This paper presents a mobility prediction strategy for manned and unmanned ground vehicles planned to operate over large regions ($> 5 \times 5 [Km^2]$). An important contribution of this work is that we have demonstrated the importance of considering uncertainty in elevation in the cost function of the path planning algorithm. It not only leads to a safer route but also it could mean a lower consumption. However, depending on the criteria or objective selected, it might not mean the optimal solution (e.g. when the shortest route is desired). Specifically, two different environments with different elevation profiles have been tested. The first one deals with a $\sim 13 \times 13 [Km^2]$ region; the second environment has a dimensions of $\sim 8 \times 8 [Km^2]$. The reason why those two environments have been selected is because they have two different DEM formats and two different resolutions. The Airport Lake DEM has a USGS format with 30-m resolution. The Sahara desert has an SRTM format with 90-m resolution, which means points are more sparse in the second DEM than in the first one. Notice that 30-m resolution DEMs are only available for the US territory so far, worldwide coverage is only
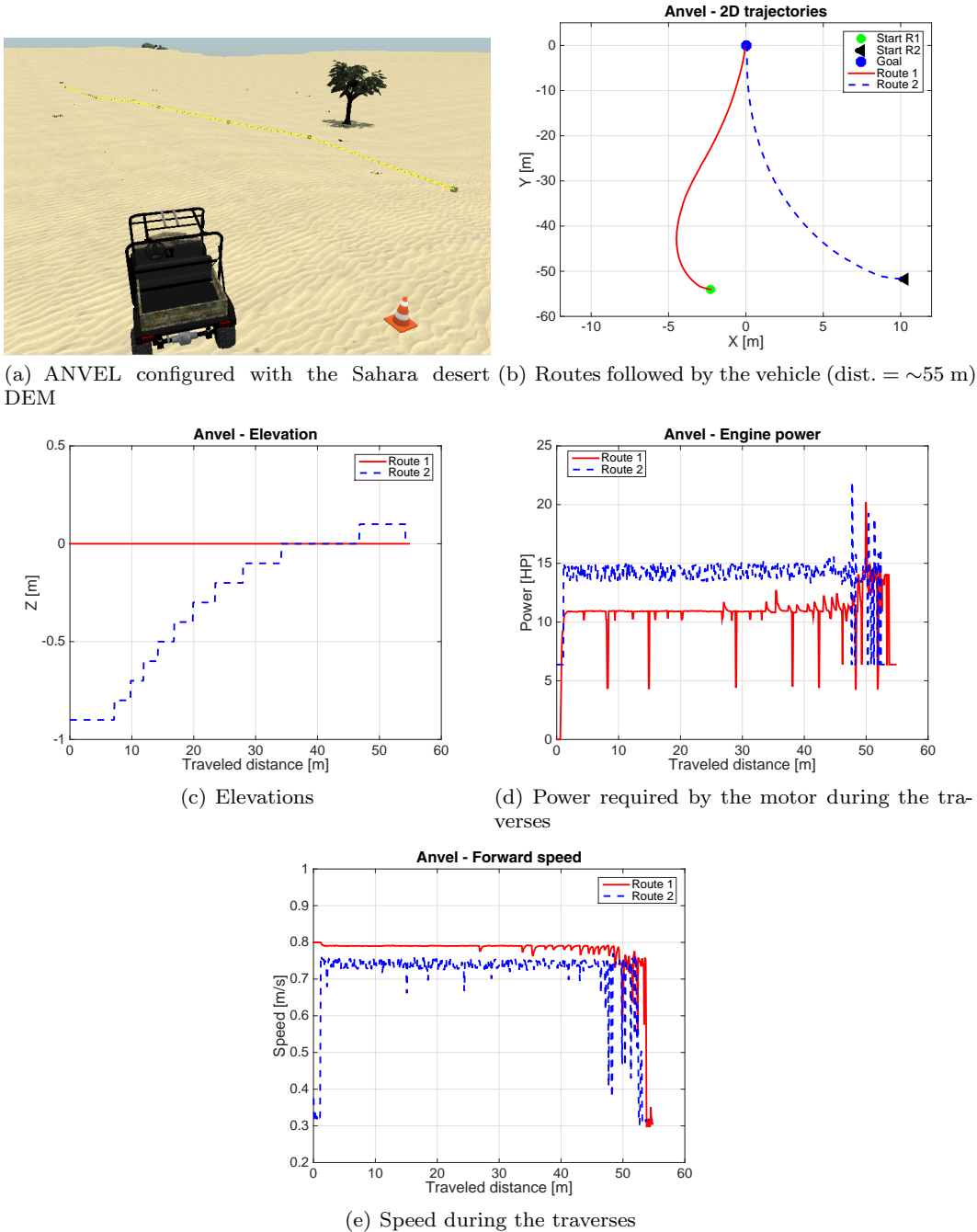
(a) ANVEL configured with the Sahara desert DEM

(b) Routes followed by the vehicle (dist. = $\sim$55 m)

(c) Elevations

(d) Power required by the motor during the traverses

(e) Speed during the traverses

Fig. 17 Comparing the performance of the proposed methodology through the realistic simulator ANVEL. Notice that the same DEM than used before has been loaded in ANVEL.

available through the SRTM format and 90-m resolution.

The computation burden is a major issue working with actual digital elevation models of regions larger than $5 \times 5 [Km^2]$ because large datasets have to be handled (e.g. Airport Lake DEM has almost 35000 samples and the model obtained with kriging has almost 2 million points). For this reason, identifying methods for obtaining a reduced-order representation of that DEM constitutes a key point in this field. In this research, a subsampling approach is proposed. It certainly reduces computation time in the variogram calculation and in the kriging step without reducing the precision of the geostatistics-related metrics. Due to this contribution, we have been able to perform kriging over a reduced-order representation of the DEMs. When we tried kriging over those scenarios with no subsampling

algorithm the computation time increased dramatically (Intel Core i7 3GHz, 16 MB RAM).

Future efforts will focus on combining uncertainty in elevation with information from the terrain itself (i.e. terrain trafficability). This issue will lead to more reliable and safer routes. Firstly, taking into account uncertainty in elevation will lead to more efficient routes, and secondly, considering terrain information will reduce the risk of vehicle entrapment.

References

AM General 2015. AM General LLC specialized vehicles for commercial and military customers. URL: `http://www.amgeneral.com/vehicles/hmmwv`; Visited: November, 2015.

ASTM Standards 1996. American Society for Testing and Materials (1996) Annual Book of ASTM Standards. Volume 04.08, Chapter Standard Guide for Analysis of Spatial Variation in Geostatistical Site Investigations.

Amidi 1990. Amidi O (1990) Integrated Mobile Robot Control. Tech. Rep. CMU–RI–TR–90–17, Robotics Institute, Carnegie Mellon University, USA.

Anderson et al 2006. Anderson A, Wang G, Gertner G (2006) Local Variability based Sampling for Mapping a Soil Erosion Cover Factor by Co-simulation with Landsat TM Images. Int J of Remote Sensing 27(12):2423 – 2447.

Arieira et al 2011. Arieira J, Karssenberg D, de Jong S, Addink E, Couto E, da Cunha CN, Skoien J (2011) Integrating Field Sampling, Geostatistics and Remote Sensing to Map Wetland Vegetation in the Pantanal, Brazil. Biogeosciences 8:667 – 686.

Basaran et al 2011. Basaran M, Erpul G, Ozcan A, Saygin D, Kibar M, Bayramin I, Yilman F (2011) Spatial Information of Soil Hydraulic Conductivity and Performance of Cokriging over Kriging in a Semi-arid Basin Scale. Environ Earth Sci 63:827 – 838.

Bechler et at 2013. Bechler A, Romary T, Jeannee N, Desnoyers Y (2013) Geostatistical Sampling Optimization of Contaminated Facilities. Stochastic Environmental Research and Risk Assessment 27(8):1967 – 1974.

Bivand et al 2013. Bivand R, Pebesma E, Gomez-Rubio V (2013) Applied Spatial Data Analysis with R, Second ed. Use R, Springer.

Bohling 2005. Bohling G (2005) Kriging. URL: `http://people.ku.edu/~gbohling/cpe940`; Visited: November, 2015.

Brus and Gruijter 1994. Brus D, Gruijter J (1994) Estimation of Non-ergodic Variograms and their Sampling Variance by Design-based Sampling Strategies. Mathematical Geology 26(4):437 – 454.

Chiles and Delfiner 2012. Chiles J, Delfiner P (2012) Geostatistics. Modeling Spatial Uncertainty, Second edn. Probabily and Statistics, Wiley.

Corke 2011. Corke P (2011) Robotics, Vision and Control. Fundamental algorithms in Matlab. Springer Tracts in Advanced Robotics, Springer, Germany.

Davis and Keller 1997. Davis T, Keller C (1997) Modelling Uncertainty in Natural Resource Analysis using Fuzzy Sets and Monte Carlo Simulation: Slope Stability Prediction. International Journal of Geographical Information Science 11(5):409 – 434.

Fisher 1991. Fisher P (1991) Modelling Soil Map-Unit Inclusions by Monte Carlo Simulation. Int. Jornal of Geographical Information Systems 5(2):193 – 208.

Fisher and Tate 2006. Fisher P, Tate N (2006) Causes and Consequences of Error in Digital Elevation Models. Progress in Physical Geography 30(4): 467 – 489.

Goldberg et al 2002. Goldberg S, Maimone M, Matthies L (2002) Stereo Vision and Rover Navigation Software for Planetary Exploration. In: IEEE Aerospace Conference, IEEE, vol 5, pp 2025 – 2036.

Gorsich and Genton 2000. Gorsich D, Genton M (2000) Variogram Model Selection via Nonparametric Derivative Estimation. Mathematical Geology 32(3):249 – 270.

Haley et al 1979. Haley PW, Jurkat MP, Brady PM (1979) NATO Reference Mobility Model, Edition I. Tech. Report 12503. US Army TARDEC, Warren, Michigan.

Hadsell et al 2009. Hadsell R, Sermanet P, Ben J, Erkan A, Scoffier M, Kavukcuoglu K, Muller U, , Hadsell YR, Sermanet P, Ben J, Erkan A, Scoffier M, Kavukcuoglu K, Muller U, LeCun Y (2009) Learning Long-Range Vision for Autonomous Off-Road Driving. Journal of Field Robotics 26(2):120 – 144.

Helmick et al 2009. Helmick D, Angelova A, Matthies L (2009) Terrain Adaptive Navigation for Planetary Rovers. Journal of Field Robotics 26(4):391 – 410.

Hosseini et al 2014. Hosseini S, Kappas M, Bodaghabadi M, Chahouki M, Khojasteh E (2014) Comparison of Different Geostatistical Methods for Soil Mapping using Remote Sensing and Environment Variables in Pshtkouh Rangelands, Iran. Polish

Journal of Environmental Studies 23(3):737 – 751.

Hunter and Goodchild 1997. Hunter G, Goodchild M (1997) Modeling the Uncertainty of Slope and Aspect Estimates Derived from Spatial Databases. Geographical Analysis 29(1):35 – 49.

Isaaks and Srivastava 1989. Isaaks E, Srivastava R (1989) An Introduction to Applied Geostatistics. Oxford University Press.

Ishigami et al 2009. Ishigami G, Nagatani K, Yoshida K (2009) Slope Traversal Controls for Planetary Exploration Rover on Sandy Terrain. Journal of Field Robotics 26(3):264–286.

Karumanchi et al. 2010. Karumachi S, Allen T, Bailey T, Scheding, S (2010) Non-parametric Learning to Aid Path Planning over Slopes. The International Journal of Robotics Research 29(8): 997 – 1018.

Kweon and Kanade 1992. Kweon IS, Kanade T (1992) High-Resolution Terrain Map from Multiple Sensor Data. IEEE Transactions on Pattern Analysis and Machine Intelligence 14(2):278 – 292.

Kerry et al 2010. Kerry R, Oliver M, Frogbrook Z (2010) Geostatistical Applications for Precision Agriculture, Springer, Germany, chap Sampling in Precision Agriculture, pp 35 – 63.

Kulis and Jordan 2012. Kulis B, Jordan M (2012) Revisiting K-Means: New Algorithms via Bayesian Nonparametrics. In: Int. Conf. on Machine Learning, International Machine Learning Society (IMLS), pp 513 – 520.

Kumar et al 2011. Kumar J, Mills R, Hoffman F, Hargrove W (2011) Parallel K-Means Clustering for Quantitative Ecoregion Delineation using Large Data Sets. Procedia Computer Science 4:1602 – 1611.

Lakhankar et al 2010. Lakhankar T, Jones A, Combs C, Sengupta M, Haar TV, Khanbilvardi R (2010) Analysis of Large Scale Spatial Variability of Soil Moisture using a Geostatistical Method. Sensors 10:913 – 932.

LaValle 2006. LaValle SM (2006) Planning Algorithms. Cambridge University Press, United Kingdom. URL: http://planning.cs.uiuc.edu; Visited: November, 2015.

Lessem et al 1996. Lessem A, Mason G, Ahlvin R (1996) Stochastic Vehicle Mobility Forecasts using the NATO Reference Mobility Model. Journal of Terramechanics 33(6):273 – 280.

Li and Heap 2011. Li J, Heap A (2011) A Review of Comparative Studies of Spatial Interpolation Methods in Environmental Sciences: Performance and Impact Factors. Ecological Informatics 6(3-4):228 – 241.

mGstat 2015. mGstat: A Geostatistical Matlab toolbox. URL: http://mgstat.sourceforge.net; Visited: November, 2015.

Papadakis 2013. Papadakis P (2013) Terrain Traversability Analysis Methods for Unmanned Ground Vehicles: A Survey. Engineering Applicatons of Artificial Intelligence 26(4):1373 – 1385.

Pengelly 2002. Pengelly J (2002) Monte Carlo Methods. http://www.cs.otago.ac.nz/cosc453; Visited: November, 2015.

Peynot et al 2014. Peynot T, Lui S, McAllister R, Fitch R, Sukkarieh S (2014) Learned Stochastic Mobility Prediction for Planning with Control Uncertainty on Unstructured Terrain. Journal of Field Robotics 31(6):969 – 995.

Rubinstein and Kroese 2007. Rubinstein R, Kroese D (2007) Simulation and the Monte Carlo Method. Applied Probability and Statistics. Wiley, Second Ed.

Stentz 1995. Stentz A (1995) The Focussed D* Algorithm for Real-Time Replanning. In: Proc. of the INt. Joint Conference on Artificial Intelligence.

Stentz et al 2003. Stentz A, Kelly A, Rander P, Herman H, Amidi O, Mandelbaum R, Salgian G, Pedersen J (2003) Real-time, Multi-perspective Perception for Unmanned Ground Vehicles. In Proc. of AUVSI Unmanned Systems Symposium.

Stentz et al 2002. Stentz T, Kelly A, Herman H, Rander P, Amidi O (2002) Integrated Air/Ground Vehicle System for Semi-Autonomous Off-Road Navigation. In: AUVSI Symposium, pp 1 – 15.

Srivastava 2013. Srivastava R (2013) Geostatistics: A Toolkit for Data Analysis, Spatial Prediction and Risk Management in the Coal Industry. Int. J. of Coal Geology 112:2 – 13.

Thompson 2012. Thompson S (2012) Sampling, Third Ed. Probability and Statistics, Wiley.

Thrun et al 2006. Thrun S, Montemerlo M et al. (2006) STANLEY, the robot that won the DARPA Gran Challenge. Journal of Field Robotics 23(9): 661–692.

Tsui et al 2013. Tsui O, Coops N, Wulder M, Marshall P (2013) Integrating Airborne LIDAR and Spaceborne Radar via Multivariable Kriging to Estimate Above-ground Biomass. Remote Sensing of Environment 139:340 – 352.

Vandapel et al 2006. Vandapel N, Donamukkala R, Hebert M (2006) Unmanned Ground Vehicle Navigation using Aerial Ladar Data. Int Journal of Robotics Research 25(1):31 – 51.

van der Meer 2012. van der Meer F (2012) Remote-sensing Image Analysis and Geostatistics. Int Journal of Remote Sensing 33(18):5644 – 5676.

Vedaldi and Fulkerson 2015. Vedaldi A, Fulkerson B (2008) VLFeat: An Open and Portable Library of Computer Vision Algorithms. URL: http://www.vlfeat.org; Visited: November, 2015.

Webgis 2015. WebGIS. Geographic Information Systems Resource. URL: `http://www.webgis.com`; Visited: November, 2015.

Webster and Oliver 2007. Webster R, Oliver M (2007) Geostatistics for Environmental Scientists, Second edn. Statistics in Practice, Wiley.

Willoughby et al 2006. Willoughby W, Jones R, Mason G, Shoop S, Lever J (2006) Application of Historical Mobility Testing to Sensor-based Robotic Performance. In: Proc. SPIE 6230, Unmanned Systems Technology VIII, pp 1 – 8.

Wu 2012. Wu J (2012) Advances in K-means Clustering. Springer Theses, Springer.

Zhou and Liu 2004. Zhou Q, Liu X (2004 Analysis of Errors of Derived Slope and Aspect related to DEM Data Properties. Journal of Computer & Geosciences 30: 369 – 378.

RAMÓN GONZÁLEZ

Ramon is currently working in the Robotic Mobility Group at the Massachusetts Institute of Technology as a Postdoctoral Associate. He holds a Computer Science Engineering degree and a Ph.D. (with European mention) in mobile robotics from the University of Almeria (Spain; 2006, 2011). He worked at the University of Zaragoza (Spain) as a PhD Assistant from January to September 2014. From 2006 to 2014, he worked as a researcher and a lecturer at the University of Almeria. In 2009, he was a visiting researcher at ASL-ETH (Zurich, Switzerland; 3.5 months). He is author of the monograph *Autonomous Tracked Mobile Robots in Planar Off-Road Conditions* (Springer, 2014). His main research interests include: modelling, localization and motion control of mobile robots in off-road conditions; terrain classification; computer vision and machine learning.; adaptive, predictive, and robust control; and geostatistics applied to mobile robotics. Ramon is a member of the IEEE Robotics and Automation Society.

PARAMSOTHY JAYAKUMAR

Dr. Paramsothy Jayakumar is a Senior Research Scientist, SAE Fellow, and a member of the Analytics Team at the U.S. Army Tank Automotive Research, Development, & Engineering Center (TARDEC) in Warren, Michigan. Prior to joining U.S. Army TARDEC, he worked for BAE Systems, Ford Motor Company, Altair Engineering, and Engineering Mechanics Research Corporation in the areas of multibody dynamics software development, vehicle dynamics modeling & simulation consulting, simulation technology development, durability load simulation, vehicle instrumentation & loads measurement, and road load engineering.

Jayakumar has written more than 100 technical publications including journal articles and conference papers. His research in terramechanics and multibody dynamics won the best paper awards at the National Defense Industrial Association's Ground Vehicle Systems Engineering and Technology Symposium in 2011 and 2012. He holds a U.S. patent for a system for virtual prediction of road loads and tire modeling. He was also instrumental in developing seven SAE standards for tire testing for the purpose of tire modeling for which he received the SAE 2014 James M. Crawford Technical Standards Board Outstanding Achievement Award and Arch T. Colwell Cooperative Engineering Medal.

Jayakumar is a member of the U.S. Army Acquisition Corps, an Honorary Fellow of the Department of Mechanical Engineering at the University of Wisconsin – Madison, and an Associate Editor for the ASME Journal of Computational and Nonlinear Dynamics. He received his M.S. and Ph.D. degrees in structural dynamics from Caltech, and B.Sc. Eng. (Hons, First Class) from the University of Peradeniya, Sri Lanka.

KARL IAGNEMMA

Karl Iagnemma is a principal research scientist in the Mechanical Engineering

department at the Massachusetts Institute of Technology. He holds a B.S. from the University of Michigan, and an M.S. and Ph.D. from MIT, where he was a National Science Foundation Graduate Fellow. He has performed postdoctoral research at MIT, and has been a visiting researcher at the NASA Jet Propulsion Laboratory and the National Technical University of Athens (Greece). He currently serves on the editorial board of the IEEE Transactions on Robotics and the Journal of Field Robotics. Dr. Iagnemma's primary research interests are in the areas of design, sensing, motion planning, and control of mobile robots in outdoor terrain, including modeling and analysis of robot-terrain interaction. He is author of the monograph *Mobile Robots in Rough Terrain: Estimation, Planning and Control with Application to Planetary Rovers* (Springer, 2004). He has recently led research programs for agencies including the U.S. Army Tank-Automotive and Armaments Command, the Army Research Office, DARPA, the NASA Mars Program Office, Ford Motor Company, and the NASA Institute for Advanced Concepts, among others. He has authored or co-authored many conference and journal papers on a wide range of robotic topics, and has consulted for various private companies and government agencies. Dr. Iagnemma was chosen as one of 16 people who redefined science in 2003 by SEED magazine. He was a 2005 nominee for a MacArthur Foundation fellowship.