Evolutionary algorithms under noise and uncertainty

Digital Object Identifier (DOI):
10.1109/SSCI.2016.7849959

Published In:
2016 IEEE Symposium Series on Computational Intelligence (SSCI)

Evolutionary Algorithms under Noise and Uncertainty: a location-allocation case study

Marta Vallejo
Department of Computer Science
Heriot-Watt University
Edinburgh, UK

Email: mv59@hw.ac.uk

David W. Corne
Department of Computer Science
Heriot-Watt University
Edinburgh, UK

Email: d.w.corne@hw.ac.uk

***Abstract*—Evolutionary approaches are metaheuristics that can deal with the effect of noise and uncertainty in data using different strategies. In this paper is depicted the method used to cope with these elements in a dynamical location-allocation problem. The use of Monte Carlo sampling and statistical historical data that can be applied to a single and multi-objective problems and within an online and offline scenario is tested and evaluated.**

I. INTRODUCTION

Evolutionary Algorithms (EAs) are biologically-inspired algorithms for search and optimization that have received much attention as a tool to be applied in a wide range of problems. The technique requires to repetitively evaluate a fitness function for evolving its population of solutions. However, in real-world problems these evaluations can be challenging due to different causes including technological limitations, the presence of randomness, the existence of unknown model parameters or when objectives are based on estimations and combinatorial relationships that because of their high level of complexity may be hard to be effectively modelled analytically.

Furthermore, the nature of the fitness may be challenged even for conventional fitness evaluations. This can be the case when the fitness function is available but very expensive to compute like in structural design optimisation problems [1]. The goal here is to use the real fitness function, whose efficiency is maximised, along with an approximative fitness function much simpler to solve than the original one. Built from a small number of samples of the original function, the combination of this model-based fitness is normally called *evolution control*. In other cases, the explicit fitness function may not exist like in art design or music composition. The approach followed here is to focus on the use of interactive methods that allow the collection of different human opinions [2].

In the concrete context of urban planning, uncertainties presented in objective functions and constraints are very important factors to take into account specially when reliability of predictions and robustness of solutions are considered. Hence, if the application of EA techniques to solve urban planning problems is considered, further mechanisms should be taken into consideration in order to generate robust solutions.

However, not all types of uncertainties should be approached in the same way. Jin & Branke [3] differentiate among four types of uncertainty that can affect the performance of EA techniques. These variants are named noise, robustness, fitness approximation and time varying fitness function. In this paper, these concepts will be discussed. A special focus will be done on important forms of noise and uncertainty that affect the objective function evaluations within an Agent-Based System scenario. From the multiple strategies to cope with these factors, the developed approach will uniquely analyse and apply an approximative fitness function method which is the result of transforming a time varying fitness function. This process is performed for simplification purposes.

In this regard, an approach that makes use of a statistical model of the agent-based system's behaviour to collect this missing data and inform a rapid approximation of the fitness function is investigated. This approach requires a limited number of prior simulations of the objective function that are averaged and used as an estimate of the real objective value. Then, this procedure allows the application of an evolutionary algorithm to optimise urban growth policies, where the quality of a policy is evaluated within a highly noisy and uncertain environment.

The paper is organised as follows: Section II describes the problem of applying an evolutionary approach to a problem with noise and/or uncertainty in both versions single and multi-objective. Section III is focused on different methodologies that can be followed to calculate the fitness function within an evolutionary algorithm scenario and, how it is applied to the problem and a short study of the validity of the method. The section is also devoted to illustrate the use of Monte Carlo sampling techniques in offline learning and different strategies that can be pursued to perform an efficient generation of samples according to the particular characteristics of the problem. Section IV explains the particular problem that the EA algorithm has to cope with, how the unknown information is generated and how the approach is tested. Finally, Section V summarise the main elements included in the paper.

August 15, 2016

II. EVOLUTIONARY ALGORITHMS UNDER NOISE AND UNCERTAINTY

As it was previously mentioned, EA algorithms can face some issues of applicability when they are confronted with real-world problems. In the field of optimisation, this strategy in both, single and multiple objective versions, can be characterised as a significant robust method when it has to deal with noisy environments [4], [5]. This advantage over other methods is mainly caused by its intrinsic use of a population of solutions to solve the problem under consideration that acts as a filter for noise when the average performance is computed [6].

Sources of noise can be varied. It can be caused by aleatory uncertainty related to the representation of sensors and actuators, by measurement limitations, due to the inherent stochasticity of some techniques such as multiagent simulations, by the propagation of the uncertainty resulted from the input data or because of the aggregate behaviour of different factors.

A. Single-objective Optimisation

In a noisy environment with a certain degree of randomness, typical of stochastic simulation models, predictability is challenged by the fact that under the same initial conditions and input parameters, results may vary every time they are generated. In Fig. 1 this effect is graphically shown. In systems where these variations are inherent and irreducible, data can be represented as a probabilistic distribution.

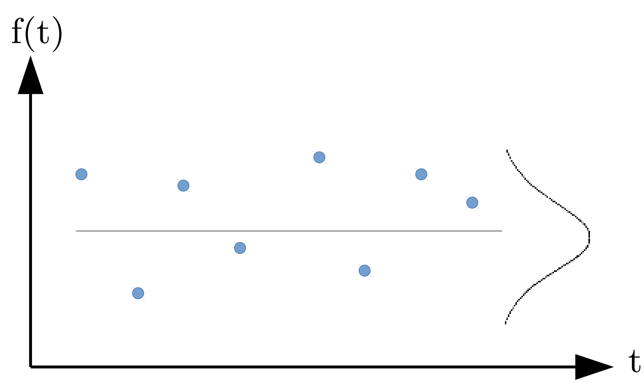


Fig. 1: Illustration of the variation in fitness values due to noise. For repeated measurements of the same specific problem, the objective fitness function f changes. In this case, these perturbations are considered to be ruled by a normal distribution.

However, if this randomness can affect the system after the evaluation is performed because the current solution is disturbed, then this specific type of noise is denoted as *Robustness*. This can be caused by manufacturing tolerances and directly affects to structural design problems [7].

In scenarios where noise is present, the selection operator within the EA can deliver unstable results and the convergence of the solutions may be adversely affected propagating inferior solutions [8]. In these cases it is convenient to quantify the probability that the operator generates wrong decisions [5]. This fact occurs when the fitnesses of the solutions A and B are $f(A) < f(B)$ but their expected distribution values are the contrary $dv(B) < dv(A)$. These distributions can be constructed by performing multiple evaluations for each chromosome which is very expensive in terms of computational costs. A less demanding approach would be to perform different evaluations in a single random chromosome to estimate the entire distribution assuming that it can be extended for all the population of solutions.

Another possible risk is the existence of epistemic uncertainty in the system. Galbraith1973 [9] defines this type of uncertainty in terms of the difference between the amount of information necessary to perform a given task and the amount of information already known. The sources of this type of uncertainty come from scarcity in the amount of experimental data collected, lack of accuracy in the approximations and assumptions selected to simplify the system, significant missing factors not included into the model or even a poor understanding of the processes involved in [10]. To avoid any kind of confusion, from this point irreducible and random uncertainty will be named as noise and epistemic uncertainty will be denoted simply uncertainty.

In systems characterised by the presence of this type of uncertainty, the definition of the problem under consideration has a lack of accuracy in objectives values or in the parameters that describe the system. Under these circumstances, a pair of successive evaluations of the same individual solution will retrieve the same objective values and not different ones like in the previous case. However, these values may not be totally accurate. The complexity in this case arises when two chromosomes are compared. Due to the inaccurate evaluations caused by the uncertainty, solutions can be also misclassified. Generally this uncertainty can be reduced by increasing the knowledge within the system.

Apart from the problem of prioritising the best solutions, the presence of noise or uncertainty in the objectives causes a slower rhythm in the evolution of the population of solutions. Hence, taking into account all of these circumstances, for a classical implementation of the evolutionary algorithm, its use and performance have been questioned [11], [12]. From a computational point of view, it is important to mention that epistemic uncertainty is more challenging to cope with than random noise [10].

In order to capacitate EA to solve this kind of problems it is necessary to include external tools and mechanisms to support the process. Existing methods that can be applied to evolutionary systems which work in uncertain and noisy environments include approximation techniques such as fully and simplified computational simulations and metamodels [13]. However, even if these techniques can aid the EA to be considered suitable for this purpose, the number of studies focused on applying EA techniques to Dynamical Planning (DP) or to a Sequential Decision Making (SDM) problem under uncertainty is almost insignificant. Instead they have been approached generally using decision trees [14], [15], influence diagrams [16] and Partially Observable Markov Decision Processes (POMDPs) [17], [18]. However, such strategies cannot be generally scaled to large problems, since to properly represent these kind of problems, the number of required states grows exponentially [17], [19].

B. Multi-objective Optimisation

In many practical applications where noise is present, a multi-objective algorithm requires not only to be able to cope with multiple optimisation objectives that can be complex and non-linear, but also with the stochastic noise that is generated as a consequence of uncontrollable variations in the system [3].

Concretely in a multi-objective scenario, the system no longer generates two possible outcomes from the comparison between two solutions. Instead there is a triple possible composition that is, $f(A) < f(B)$ and $f(A) > f(B)$ and a non-dominate option $f(A) \equiv f(B)$ where any of the solutions can be considered better than the other. This extension makes the filtering of noise a harder task. One reason of this increment of the complexity is that uncertainty and noise in multi-objective systems change the nature of the solutions within the Pareto front which are transformed from points in the search space to hypercubes, see Fig. 2.

Noise may alter the dominance relationship between different solutions in such a way that it could be possible that dominated solutions may become non-dominated or vice-versa [20]. Consequently, the application of the selection operator may be also misled, eliminating good solutions or reproducing inferior ones. This effect may produce a reduction in the convergence rate and a poor quality set of final solutions [21]–[23].

Apart from this aspect, the noise in the fitness calculation may produce outlier solutions whose values are placed at an abnormal distance from the rest of solutions in the search space. In this case, the optimization algorithm might get stuck in one of the solutions which dominates all present solutions [24]. The appearance of outliers can be caused by insufficient sampling or by the disparity in the distance to the Pareto front among objectives [25].

Different approaches have been investigated for such multi-objective scenarios. In this regard, a modified Pareto ranking scheme adapted from Goldberg [26] has been proposed to deal with the presence of noise. There are two major ranking scheme versions that have been studied: one which focuses on probability techniques and another based on clustering methods. The probability-based Pareto ranking schema of Hughes2001 [5] uses a probabilistic ranking process to take noise into consideration by defining probabilities of dominance between noisy solutions [27]. The standard deviation of each evaluation for the entire population of solutions can be used to correct the noise. In this technique, the probabilistic rank of an individual is calculated by the sum of the probabilities of those solutions that this chromosome dominates. Finally, in the clustering variant [25], the Pareto front is formed by the best found solutions plus solutions that belong to their neighbourhood. The neighbourhood calculation takes into account a user-defined restriction factor and the standard deviation for each objective.

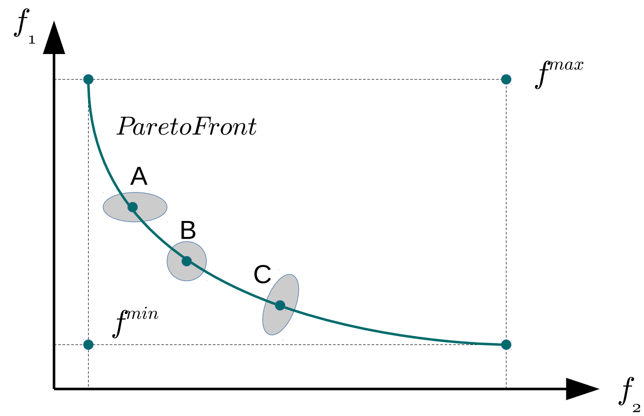


Fig. 2: Graphical representation of the difference between the representation of a solution within the Pareto front in scenarios with and without noise. The normal point representation in a standard search space (solutions A, B and C) is transformed in an uncertain environment into a hypercube which is represented by grey areas surrounding the point solutions.

Additionally Büche et al. [24] proposed a modification of the (μ, κ, λ) algorithm [28] to minimise the effect of noise and outliers. The called Domination Dependent Lifetime (DDL) assigns a maximal lifetime κ to each individual based on the number of solutions it dominates, in such a way that the lifetime value k will be shorter if the number of chromosomes in the population is large. This feature contrasts with the effect of elitism which may preserve solutions for an infinite amount of time by limiting the impact of inferior new individual solutions. However, to prevent the elimination of good solutions, the approach is complemented with a mechanism that allows the re-evaluation of the lifetime of the expired solutions. If these solutions are good enough they will be added again to the pool of solutions with new objective values resulted from a new reevaluation. This will change previous good solutions with other noisy samples.

III. Fitness Approximation

In an uncertain and noisy context the fitness function, which is evaluated by means of statistical, conceptual or physical simulations, is normally the most computationally intensive element of the given application [29]. This high computational requirement has caused the development of approximative alternatives to alleviate the corresponding cost. This is the case of the use of Artificial Neural Networks (ANN) as a modelling tool for function approximation [30], [31], which can be aimed at replacing computationally intensive models. The critical issue in this approach is to find a good quality approximative strategy in such a way that the behaviour of this approximation is similar enough to the original model.

Otherwise the final system could experience a severe negative impact when errors are evaluated [32].

The manner noise influences the fitness value is varied. In additive noise, additional values are randomly added to or subtracted from the real fitness value. This type of additive fitness function can be defined as such: if ρ_i is the fitness function that is defined based on a determined configuration of the problem for a determined chromosome i, then the noisy fitness function ρ'_i can be formalised as follows:

$$\rho'_i = \rho_i + rnd[N(0, \sigma^2)] \tag{1}$$

where ρ is the noisefree fitness function and $rnd[N(0, \sigma^2)]$ denotes the assumption that the noise can be approximated to a normal distributed noise component added in each evaluation. The uncertainty set can be defined as:

$$U(\rho) = \{\xi \in \mathbb{R}^n : \rho - \triangle \leq \xi \leq \rho + \triangle\} \tag{2}$$

where $\triangle = (\triangle_1, \triangle_2 \cdot \triangle_n)^T \in \mathbb{R}^n$ is the aggregate uncertainty and n is the dimension of the decision space.

To reduce the level of noise within the function, a sampling process can be used based on the central limit theorem, solving n times the fitness function and averaging these values:

$$\rho^*_{i,n} = \frac{1}{n} \sum_{j=1}^{n} \rho'_{i,j} \tag{3}$$

where $\rho'_{i,j}$ is the sampling evaluation number j of the individual i and $\rho^*_{i,n}$ is the distribution of ρ inferred from the mean of n samples of ρ'. As a larger value of n is defined, the standard deviation is decreased.

The use of a normal distribution is very common in the literature, however the nature of the source of noise can be characterised by other types of distributions [4] that, in general, are largely unexplored. In this regard, Beyer et al. [33] transformed a non-Gaussian noise distribution in a nearly Gaussian type in order to deal with this type of problems. However, this procedure cannot be generalised for all situations. There are numerous cases where this transformation is only possible at a disproportionate error cost. Sendhoff et al. [34] introduced another approach where multiple types of functions beyond the simple additive noise model were analysed.

A. Types of fitness by approximation

In order to generate the analytical values required to define the fitness by approximation, four basic strategies have been introduced namely explicit averaging, implicit averaging, fitness inheritance and selection modification [3]. All methods assume that the search space is characterised by a known and homogeneous noise distribution most commonly a uniform or a normal distribution type. It is also considered that an estimation of its magnitude is possible to calculate [35]. However, these assumptions limit the effectiveness of the selected approach due to the fact that, in general, the effect of noise is not spread homogeneously over the search space and the absence of knowledge regarding the level of noise are the most common characteristics of real-world problems.

The explicit averaging, also called static resampling, was introduced by Miller [36] and it is the most commonly used method for coping with noise. The strategy consists of generating a determined number of times the sampling of the objectives, followed by the averaging of the generated values [37]. In a sample size of n, this operation allows a proportional reduction of the variance by a factor of $\sqrt{n}$. Additionally it also implies the increment in the computational effort used by a factor of n [3]. To avoid extra evaluations, the fitness from the neighbourhood can be used [38], [39].

Another possible approach is to apply a statistical model constructed beforehand with historical data to model the fitness using techniques such as local regression and adaptation [40]. If the approximate model is generated by an offline training process before the optimisation is run, it is common the use of Monte Carlo techniques to generate these samples. Sampling is a popular method to reduce noise and estimate unknown information.

In the implicit averaging, on the other hand, sample size is defined as an inverse function of the population size [37]. The idea behind this interpretation consists of the fact that in systems defined with a large population of solutions, it is very common the existence of numerous chromosomes that are very similar to each other. The frequent evaluation of these related areas of the search space reduces the noise.

Bui et al. [35] introduced a technique to solve this problem based on the idea of fitness inheritance. They proposed that the offspring created in each generation additionally inherits two variables from its parents: μ that represents the mean of the objective value and σ that corresponds to the standard deviation. These variables will control whether a new resampling is required or not. The resampling operation consists of calculating the new fitness by performing a predefined number of evaluations where the final values μ and σ correspond to the mean and the standard deviation of these evaluations. When a new child is evaluated, a resampling is only required if its objective values fall outside the confidence interval. Otherwise, the inherited fitness is assigned. Consequently the evaluation of solutions characterised with higher noise will result in larger standard deviation values which facilitate the fitness acceptance in its children.

Finally, the modification of the selection operator is another method investigated to cope with the noise when the fitness reevaluation is too costly. Teich [27] defined a selection and ranking procedure that take into account some conditions like the probability of dominance to compensate the noise. Another similar strategy uses a threshold value when finesses are evaluated to overtake the effect of the hypercubes in a multi-objective scenario [41]. Refer to Jin & Branke [3] for more information about this topic and also to Qian et al. [42] for a brief update on the state-of-the-art.

B. Monte Carlo

Monte Carlo simulation is a technique that was developed in the 40's by Metropolis & Ulam [43]. Since then it became a widely used and effective tool for those problems whose analytic solutions do not exist or have a high level of complexity to be easily obtained. By means of random sampling, the strategy allows the study of the properties of random-nature systems when analytic solutions are not easily available. To recreate properly the desired dynamics and patterns of the studied system into the model, it is normally used real information gathered from this objective system. However, in some cases the information collected in this way has not enough quality or cannot be easily measured and structured as a probabilistic distribution. It can be also possible that even if, in fact this information exists, its application in a large stochastic model could be a very challenge task [44]–[46].

The number of draws used within the sampling should be defined according to the level of noise and uncertainty which characterises the search space of the problem. In general, very noisy scenarios will require extra samples to come up with the same level of robustness than in more deterministic search spaces [8]. However, each new sample will increase the computational effort required for generating a single evaluation. If Monte Carlo techniques are used along with an EA approach, an alternative option to manage the noise is to increase the number of individuals that forms the population of solutions [36], [47]. However, it could be hard to know a priori, the most efficient size for the population, because this aspect depends on several factors including the level of noise, formulation and problem-specific parameters [36], [47]. At this point, there is controversy surrounding the trade-off created between the role that these two factors plays in decreasing the level of noise [48]. Fitzpatrick & Grefenstette [37] and Arnold & Beyer [49] highlighted the size of the population to increase the robustness against uncertainty over the sample size, meanwhile Beyer [50] and Hammel & Bäck [51] favoured the sample size instead. However, these conclusions strongly depend on the definition of the problem. In this regard these authors state that for the $\mu/\mu, \lambda$-ES an increment in the population of solutions is preferable when the parameter called truncate ratio μ/λ is calibrated appropriately. However for $(1, \mu)$-ES averaging over multiple samples is the best option.

Under these circumstances, it is challenging to exactly know before the algorithm is empirically tested, the reliability or level of robustness of a determined formulation. For single objective problems, Miller & Goldberg [52] inferred a lower bound of the optimal sample size and suggested that, in a system with uncertain parameters, the EA solutions only require the generation of a small number of samples. They stated that a limited number of Monte Carlo draws, that can range from 5 to 20 per population member, should be enough to compute their average fitness. This assumption is based on the idea that in EA new samples are included into the population in each generation by the application of elitist operators that highlight good solutions. As a direct consequence of this mechanism, this process will have the side effect of implicitly increasing the number of Monte Carlo realizations.

A proposed extensions of this approach is the introduction of an operator that limits the age of the solution, which control the survival of fit members. By the use of this element it is possible to further reduce the number of Monte Carlo draws in the fitness evaluation [12], [53], [54].

IV. DESCRIPTION OF THE PROBLEM

In a dynamical location-allocation problem where a set of urban green areas have to be allocated during a determined period of time subject to some constraints, the major objective to achieve can be defined by the fulfilment of the population satisfaction. This satisfaction can be depicted in terms of the distance to the household to these areas. The availability of parks at a close distance provides a varied types of services and amenities that these green facilities offer to the population from different perspectives such as aesthetic, physical, social and environmental [55]–[57]. The search can be extended to cover other objectives like environmental protectionism, level of connection between areas and profitability among others. This conflictive set of goals, more typical of real-world problems needs multi-objective techniques to be appropriately solved.

A 'policy', in this context, amounts to the city authorities' planned schedule for protection of a specific set of green spaces maximising the objectives selected in both, short and long-term. However due to the fact that governmental purchase decisions are subject to a budget that normally limits its capacity of giving a full and continuous provision of green spaces during the construction of new urban developments in cities, a careful planning should be carried out in advance. Computational optimisation techniques can be applied to the search of this maximum. Noteworthy to mention is that this budget may normally quantitatively much lower than current prices of the patches of land that are significant for the new areas under construction. Additionally since land prices generally increase with the time because of multiple factors including the rise in the demand of these spaces, scarcity of available land and other related economical factors, current acquisition policies should take into consideration not only the present status of the objective system but also a reliable projection of future necessities. However, dealing with future conditions implies irrevocably to cope with epistemic uncertainty due to a lack of knowledge that this future entails.

In this regard, there is much active research in designing long-term feasible public open space plans, whereby researchers interested in urban planning and sustainability have investigated a range of agent-based systems and similar mechanisms to explore the consequences of different green-space allocation strategies [58]–[60].

In general, the application of modelling techniques is another element that aggregates epistemic uncertainty, mostly inherited from the selection of the model and the subsequent structural changes required to adequate the system to the considered problem. Apart from that, due to the use of a Cellular Automata and an Agent-Based framework that this work applies as a modelling technique, it should be added the fact that these concrete technologies implicitly add noise to the system into consideration. Consequently, the applied EA algorithm should be robust enough to be able to cope with both factors at the same time that provide significant results.

In such a context a method that effectively obtains a model-driven approximation of the simulation to lead the evolutionary algorithm towards policies that yield vastly better satisfaction levels than unoptimised policies is investigated. By wrapping the optimisation over the agent-based simulation process, it can be used a rapidly accelerated model of the agent-based simulation in place of the real knowledge. This requires a limited number of prior simulations of the agent-based urban growth system, and then allows the use of an evolutionary algorithm to optimise urban growth policies. This strategy is based on the fact that even if approximative models do not have the capability of creating new information, they can gather useful information from the history of the optimisation and prevent its lost [61].

A. Problem Definition

The general objective of the problem consists of designing an offline planning process which leads us to find the optimal subset of green spaces out to a set $V = \{1...n\}$ of locations along with the corresponding time schedule of each of the purchase decisions. The offline nature of the selected planning approach means that the policy construction step will be done entirely before the plan is executed.

Each element within the set of different purchasing plans P includes a set of parcels of land $T \subset V$ located within a given geographic area close to or within a city that are intended to be acquired for conservation and/or social purposes. For the sake of simplicity each patch of land has a homogeneous size and shape. Each of them is considered an independent unit and no clustering techniques to group them are implemented in the model. We also assume that once an open area is selected and transformed into a park, it remains protected from urbanisation until the end of the planning horizon.

Formally a certain purchasing plan $p \in P$ is depicted by a set of parcels of land $\tau \in T$ and a purchase schedule $\psi \in \Psi$ which can be defined as a mapping from the parcels contained in T to a series of purchase times in $\psi = \{0, 1, \cdots, H\}$, where H is the maximum time horizon considered in the plan.

Commonly each candidate patch of land τ_i has associated its own cost $c_t(\tau_i)$ which covers the acquisition and the restoring/transforming process from a rural patch of land into a green space. This cost is calculated in time t which is the moment when the area is purchased and it is defined for this specific problem as a monetary cost. The value of this cost, which is always strictly positive, is not static and could vary over the time. After that step, no further maintenance costs over the area are considered. Under these circumstances, every purchase schedule included in Ψ is linked with a corresponding non-decreasing cost function C_Ψ for the entire conservation plan. This tuple describes the purchase history in relation to the accumulated cost of the land in such a way that:

$$C_\Psi(H) = \sum_{t=0}^{H} c_t(\tau) \tag{4}$$

where t_0 is the starting point in time in which acquisitions can be done and H is the maximum time covered in the planning process. Since budgets that can cope with single purchasing transactions involving large extensions of land are normally rather unlikely in real-world scenarios, to afford these purchasing investments, financial resources in terms of individual budgets $b_i \in B$ are available periodically. As a result, the acquisition process is restricted by this financial constraint that has to be respected always on the total cost.

In summary, the goal of an individual static acquisition problem is to select an optimal plan $\hat{p}$ at the same time that the budget constraint is respected on the total cost $C_\Psi(H)$. This could formally be expressed as:

$$\hat{p} = \underset{p \in P}{argmax}\{\mathscr{F}(p_i), C_\Psi(H) \leqslant B\} \tag{5}$$

Since a plan is comprised by a schedule and a subset of cells, the final goal can be also defined as the finding of a schedule $\psi \in \Psi$ and a subset of green areas $\tau \in T \subseteq V$ that effectively use the budget $b \in B$ in order to maximise a predefined objective function $\mathscr{F}()$ which assesses the utility of the corresponding plan p_i. This function $\mathscr{F}()$ quantifies in which way the pursued goal is accomplished after the completion of the prediction horizon H.

From the provision of green services perspective this function can be aimed at solving a covering problem also called Maximum Service Distance (MSD) [62], where a set of elements needs to be covered with a minimum number of subsets, subject to some constraints. An element is considered covered if it is located within a specified distance from one of these subsets. In this regard, a given scenario can be configured by the set of facilities of a certain type, concretely green areas located at a given distance to a central point. This central location is represented in this case by the CBD that attracts most of the dwellers since typically in an urban scenario population decays with distance. The final goal consists of maximising the number of inhabitants who are located relatively close to this type of service, in this case green areas.

Regarding the level of complexity of this type of problem it can be stated that by applying a reduction of the MCP [63], even for a unique time step of the problem, selecting the set of patches of land that maximise the acquisition probability is an NP-hard optimisation problem [64].

Furthermore, due to the fact that the consideration of a single optimisation step can hardly accomplish the final long-time objectives of such plans, the problem should be formulated instead by a sequential set of the previously defined static problems. The management of the financial resources between time steps can be defined in a way that the unused budget assigned in a given time step $t-1$, denoted by b_r, is added over the following period t to the corresponding budget b_0 as $b(t) = b_0(t) + b_r(t-1)$.

Apart from that, since land acquisition costs may change over time and additionally urban dynamics could transform areas in the fringe of the city into new developments, which made them inappropriate to be included within any acquisition plan, different patches of land can be available in each time step t. Consequently this problem cannot be solved statically in advance without basing the new decisions on previous actions and the forecast of future tendencies.

Under these circumstances, let $P_t \subseteq V$ be the set of patches of land that is available to be purchased in a given time step t. For each new time step t' the amount of available resources both new $b_0(t')$ and old $b_r(t'-1)$, the cost dynamics and the amount of non-urbanised areas refine the set $P_{t'}$ taking also into account the areas already selected, in such a way that $P_{t'} = P_{t'-1} \cap V$.

B. Statistical Data Generation for Sampling

The developed EA algorithm requires to receive as input parameters the concrete information that characterised the scenario where the algorithm has to work on. However, if some of these elements are totally or partially unknown, external mechanisms should generate this information. The way in which this process is performed, will have a significant impact on the feasibility of the proposed solution.

In this concrete problem, this lack of knowledge is due to the complex interactions generated among the different processes involved in the stochastic growth of the urban model. These dynamics cannot be foreseen beforehand, since different relationships can lead to the development of a significant and varied range of future scenarios.

In this regard, to assign values to these uncertain parameters, a Monte Carlo sampling strategy is implemented to generate a sample set that captures the spatial dynamics of these factors over the time [65]. The source of these realisations is a surrogate model based on a the same urban model. This external model keeps most of the characteristics of the actual site without including green externalities. This implies to eliminate possible non-linearities resulted from the relationship between urban prices and green areas [66] and residential preferences and parks [67]. These perturbations are particular of each individual realisation.

The same gathering method was successfully applied in dialogue systems [68], environmental studies [69] or emulators for managing uncertainty in urban climate models, such as the Multilayer Urban Canocopy Model (MUCM) [70], [71]. By means of such a method, these systems are capable of gathering the required data by an offline sampling mechanism.

Data sampling techniques could be applied in both online and offline planning. In an online approach, the system collects information of the environment during the entire optimisation process, meanwhile in an offline scenario, the optimisation algorithm needs to perform a training process beforehand, incorporating prior knowledge. One of the advantages of using offline techniques is that, since sample size describes a function that is pareto-optimal [72] in relation with the size of the sampling and its accuracy, the offline procedure permits the system to perform the desirable amount of sampling beforehand, focusing then only on the accuracy factor.

If the sample set is generated by an offline process, the number of samples collected will be decided in advance and will remain constant for the entire optimisation process. Other approaches consider a variable number of samples for different individuals or for different phases of the optimisation process. Aizawa & Wah [73] focus on minimising the expected estimation error, sampling using the best individuals of the population and Branke & Schmidt [23] take samples only from individuals included in the mating pool selected using a Tournament Selection technique.

After that, the required information that the EA will use is sampled 20 times, each for every factor analysed (prices of the land, population distribution and urbanisation cells), to form an initial estimate of the amount of noise in each of them. The size of the sample was decided based on empirical observations of the amount of new information added to the distribution in each new realisation. Afterwards, the mean and standard deviation were calculated in single and multiple-objective scenarios. In this case, the difference is the total number of variables sampled, since in the multi-objective version, the ecological value of the land is also taken into consideration.

The sample technique aids the EA to generate reliable solutions with a reduced number of samples [52], which is a computational advantage compared to other approaches that require the generation of a much larger number of realisations to achieve good results [74].

C. Model Application

In the model data generation can fulfil two different roles: as a gathering method to inform about constraints and as part of the calculation of the fitness. A graphical representation of this is shown in Figure 3.

As it was previously mentioned, the selection of future green location is performed sequentially for a predefined period of time. This allocation process is normally limited by different constraints, two of them are studied in this work. Firstly, the configuration of the area into consideration, its land-use type, in the precise moment the acquisition is made is an important factor to analyse. The transformation of a patch of land into a park is not permitted in areas that are already urbanised. Secondly, the subset of affordable areas that can be acquired with the current financial resources and the subset of available land generally have different cardinality. These values are linked with the fact that purchasing budgets are significant lower than the prices of the land. In an online planning, this information is available at any time. However, if the learning process is done offline, the generation of a probable evolution of these factors, budget and prices, is required to be able to take the correct decisions beforehand.

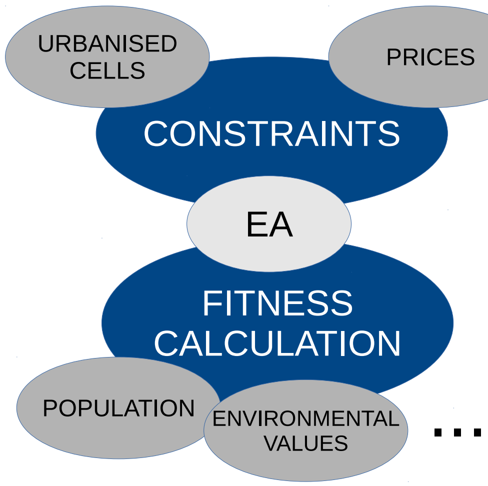
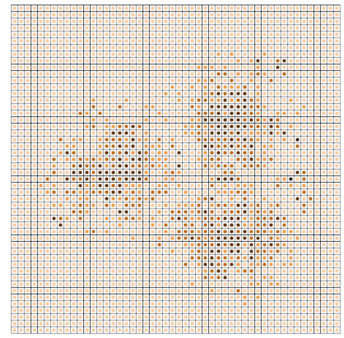


Fig. 3: Different sources of data collected to support the EA optimisation process. The data can be divided into two groups according to the role they play that within the model. As such sampling can be collected to underpin the generation of constraints or be involved in the calculation of the fitness. These sources of data are related to data needed to calculate the fitness of different objectives like the population density and its distribution or the environmental value of a patch of land. Prices of the non-urban areas that are available for being purchased and the urbanisation spread over the grid are two factors that could be used to constraint the problem.

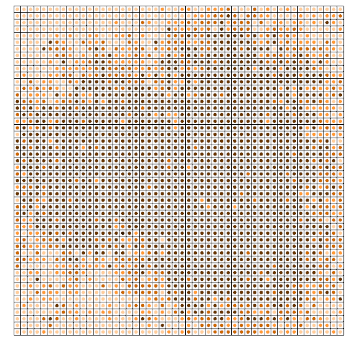
The value of the fitness function will assess the suitability of transforming a given area according to a determine criterion/a, depending on the number of factors taken into account. Regarding its calculation, independently of the number of objectives defined within the problem, if the satisfaction of the population is included in its valuation, its expected distribution needs to be inferred for the entire period covered by the planning process. Considering that it is impossible to know exactly this spatial distribution since a city is a system with complex spatial and temporal dynamics [75] then, if the fitness measures the distance from each agent to the closest green area, external tools are necessary to incorporate into the system. This process can be done in different ways: if only the current necessities of the population are taken into account, the information related to the satisfaction can be collected identically than it was defined for the constraints. However, if future long-term conditions are factors to include in the development of the present policy, a different approach should be followed.

This process can be seen as a way of generating an offline sampling fitness function [76]. As such, the fitness function is estimated for each considered time step in our discrete system by Monte Carlo sampling and the noise of each chromosome X^* is reduced by calculating the fitness function of individuals which belong to a similar search space which was previously evaluated in an offline process.

A graphical representation of the fitness in two different time-steps of the simulation is depicted in Figures 4a and 4b.



(a) Lattice with the fitness approximation values for the time step 300.



(b) Lattice with the fitness approximation values for the time step 600.

Fig. 4: Visual representation of the approximative fitness function for two different time-steps of the simulation. The fitness function is represented in a range of red colours where darker tones represent the most crowded areas of the system in relative terms.

D. Significance of Results

To assess the robustness of the use of this type of surrogate model, one aspect of the model is considered and analysed. As such, the generated population distribution was compared with the real one which was gathered once the optimisation terminated. The process was repeated 20 times and finally the values were averaged. These results mainly depend on the intrinsic effects of the random noise within the system and on the variability of the factors considered in the study. To measure this variability, the degree of similarity between distributions is analysed by the use of correlation techniques. By means of canonical correlation strategies, it is possible to estimate a symmetric measurement of the congruence of two matrices [77]. In this concrete case, it is analysed the Pearson's linear correlation matrix resulted from comparing the Monte Carlo pregenerated matrix M' with the matrix composed by the real population distribution M. Both matrices are identical in dimensions (600×2500), resulting from vectorising the grid of (50×50) values that depicts the population distribution within the city in each time step (from 0 to $T = 600$). The values of the linear dependency of the final correlation matrix are shown in the following table:

	M	**M'**
M	1	0.7634
M'	0.7634	1

TABLE I: Correlation matrix calculated from the real population distribution (M) and the simulated sampling distribution (M').

These values show a strong correlation between both sources of data. To validate this conclusion, the matrix of p-values for testing the hypothesis of no correlation from the Pearson correlation coefficients was calculated.

	M	**M'**
M	1	0
M'	0	1

TABLE II: p-values results from testing the hypothesis of no correlation between both matrices.

From these p-values, it can be rejected the assumption that the correlation is due to random sampling. Hence, the use of this type of approximative fitness function within the EA algorithm is reasonable robust. However, this is based on the general assumption that the method is supported by a consistent urban model, where the generated surrogate system has enough power of mimic the reality.

V. CONCLUSIONS

In this paper, a review of varied types of problems that EA techniques may face when it is used in real-world problems under noise and epistemic uncertainty is performed. The multiple types of mechanisms and tools along with their concrete uses with single and multiple objective variants are also commented. Afterwards the problem of estimating data within a dynamical location-allocation problem is introduced along with the selected method applied to cope with the uncertainty of the model. By means of Monte Carlo sampling, the EA optimisation process is able to calculate the fitness and generate the necessary information related to the constraints presented in the system under consideration and avoid that the selection process may become unstable.

The proposed gathering technique can be applied and tested using along with different configurations of the EA on a typical urban growth simulation. In its single version, in which the overall goal is to find policies that maximise the 'satisfaction' of the residents, an offline EA methodology was applied within a set of different scenarios where multiple levels of complexity are considered [78]. The application of the same techniques using an online planning and an offline/online multi-objective version of the problem, where other conflicting objectives are included, can be also considered.

REFERENCES

[1] Y. S. Ong, P. B. Nair, and A. J. Keane, "Evolutionary optimization of computationally expensive problems via surrogate modeling," *AIAA journal*, vol. 41, no. 4, pp. 687–696, 2003.

[2] J. Biles, "Genjam: A genetic algorithm for generating jazz solos," in *Proceedings of the International Computer Music Conference*. INTERNATIONAL COMPUTER MUSIC ACCOCIATION, 1994, pp. 131–131.

[3] Y. Jin and J. Branke, "Evolutionary optimization in uncertain environments-a survey," *Evolutionary Computation, IEEE Transactions on*, vol. 9, no. 3, pp. 303–317, 2005.

[4] D. Büche, P. Stoll, R. Dornberger, and P. Koumoutsakos, "Multiobjective evolutionary algorithm for the optimization of noisy combustion processes," *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, vol. 32, no. 4, pp. 460–473, 2002.

[5] E. J. Hughes, "Evolutionary multi-objective ranking with uncertainty and noise," in *Evolutionary multi-criterion optimization*. Springer, 2001, pp. 329–343.

[6] D. V. Arnold and H.-G. Beyer, "A comparison of evolution strategies with other direct search methods in the presence of noise," *Computational Optimization and Applications*, vol. 24, no. 1, pp. 135–159, 2003.

[7] J.-F. Barthelemy and R. T. Haftka, "Approximation concepts for optimum structural designa review," *Structural and Multidisciplinary Optimization*, vol. 5, no. 3, pp. 129–144, 1993.

[8] A. Syberfeldt, A. Ng, R. I. John, and P. Moore, "Evolutionary optimisation of noisy multi-objective problems using confidence-based dynamic resampling," *European Journal of Operational Research*, vol. 204, no. 3, pp. 533–544, 2010.

[9] J. R. Galbraith, *Designing complex organizations*. Addison-Wesley Longman Publishing Co., Inc., 1973.

[10] W. L. Oberkampf, J. C. Helton, C. A. Joslyn, S. F. Wojtkiewicz, and S. Ferson, "Challenge problems: uncertainty in system response given uncertain parameters," *Reliability Engineering & System Safety*, vol. 85, no. 1, pp. 11–19, 2004.

[11] V. Rieser, D. T. Robinson, D. Murray-Rust, and M. Rounsevell, "A comparison of genetic algorithms and reinforcement learning for optimising sustainable forest management," in *GeoComputation*, 2011.

[12] J. Wu, C. Zheng, C. C. Chien, and L. Zheng, "A comparative study of monte carlo simple genetic algorithm and noisy genetic algorithm for cost-effective sampling network design under uncertainty," *Advances in Water Resources*, vol. 29, no. 6, pp. 899–911, 2006. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0309170805002071

[13] Y. Jin, "A comprehensive survey of fitness approximation in evolutionary computation," *Soft computing*, vol. 9, no. 1, pp. 3–12, 2005.

[14] L. Garcia and R. Sabbadin, "Possibilistic influence diagrams," *Frontiers in Artificial Intelligence and Applications*, vol. 141, p. 372, 2006.

[15] G. Jeantet, P. Perny, and O. Spanjaard, "Sequential decision making with rank dependent utility: A minimax regret approach." in *AAAI*, 2012.

[16] W. Guezguez, N. B. Amor, and K. Mellouli, "Qualitative possibilistic influence diagrams based on qualitative possibilistic utilities," *European Journal of Operational Research*, vol. 195, no. 1, pp. 223–238, 2009.

[17] J. Pineau, G. Gordon, and S. Thrun, "Anytime point-based approximations for large pomdps," *Journal of Artificial Intelligence Research*, pp. 335–380, 2006.

[18] R. Sabbadin, "A possibilistic model for qualitative sequential decision problems under uncertainty in partially observable environments," in *Proceedings of the Fifteenth conference on Uncertainty in artificial intelligence*. Morgan Kaufmann Publishers Inc., 1999, pp. 567–574.

[19] R. D. Smallwood and E. J. Sondik, "The optimal control of partially observable markov processes over a finite horizon," *Operations Research*, vol. 21, no. 5, pp. 1071–1088, 1973.

[20] K. C. Tan and C. K. Goh, "Handling uncertainties in evolutionary multi-objective optimization," in *Computational Intelligence: Research Frontiers*. Springer, 2008, pp. 262–292.

[21] H.-G. Beyer, "Evolutionary algorithms in noisy environments: theoretical issues and guidelines for practice," *Computer methods in applied mechanics and engineering*, vol. 186, no. 2, pp. 239–267, 2000.

[22] D. V. Arnold, *Noisy optimization with evolution strategies*. Springer Science & Business Media, 2002, vol. 8.

[23] J. Branke and C. Schmidt, "Selection in the presence of noise," in *Genetic and Evolutionary ComputationGECCO 2003*. Springer, 2003, pp. 766–777.

[24] D. Büche, P. Stoll, and P. Koumoutsakos, "An evolutionary algorithm for multi-objective optimization of combustion processes," *Center for Turbulence Research Annual Research Briefs*, vol. 2001, pp. 231–239, 2001.

[25] M. Babbar, A. Lakshmikantha, and D. E. Goldberg, "A modified nsga-ii to solve noisy multiobjective problems," in *Genetic and Evolutionary Computation Conference. Late-Breaking Papers (GECCO 2003)*, vol. 2723, 2003, pp. 21–27.

[26] D. Goldberg, "Genetic algorithms in search, optimisation, and machine learning," 1989.

[27] J. Teich, "Pareto-front exploration with uncertain objectives," in *Evolutionary multi-criterion optimization*. Springer, 2001, pp. 314–328.

[28] T. Bäck, F. Hoffmeister, and H. Schwefel, "A survey of evolution strategies," in *Proceedings of the 4th international conference on genetic algorithms*, 1991, pp. 2–9.

[29] J. Nicklow, P. Reed, D. Savic, T. Dessalegne, L. Harrell, A. Chan-Hilton, M. Karamouz, B. Minsker, A. Ostfeld, A. Singh *et al.*, "State of the art for genetic algorithms and beyond in water resources planning and management," *Journal of Water Resources Planning and Management*, vol. 136, no. 4, pp. 412–432, 2009.

[30] K.-I. Funahashi, "On the approximate realization of continuous mappings by neural networks," *Neural networks*, vol. 2, no. 3, pp. 183–192, 1989.

[31] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural networks*, vol. 2, no. 5, pp. 359–366, 1989.
[32] Y. Jin, M. Olhofer, and B. Sendhoff, "A framework for evolutionary optimization with approximate fitness functions," *Evolutionary Computation, IEEE Transactions on*, vol. 6, no. 5, pp. 481–494, 2002.
[33] H.-G. Beyer, M. Olhofer, and B. Sendhoff, "On the behavior of (μ/μ1, λ)-es optimizing functions disturbed by generalized noise." in *FOGA*, 2002, pp. 307–328.
[34] B. Sendhoff, H.-G. Beyer, and M. Olhofer, "The influence of stochastic quality functions on evolutionary search," *Recent advances in simulated evolution and learning*, vol. 2, pp. 152–172, 2004.
[35] L. T. Bui, H. A. Abbass, and D. Essam, "Fitness inheritance for noisy evolutionary multi-objective optimization," in *Proceedings of the 7th annual conference on Genetic and evolutionary computation*. ACM, 2005, pp. 779–785.
[36] B. L. Miller, "Noise, sampling, and efficient genetic algorithms," *Doctoral Dissertation. University of Illinois at Urbana-Champaign*, 1997.
[37] J. M. Fitzpatrick and J. J. Grefenstette, "Genetic algorithms in noisy environments," *Machine learning*, vol. 3, no. 2-3, pp. 101–120, 1988.
[38] J. Branke, "Creating robust solutions by means of evolutionary algorithms," in *Parallel Problem Solving from NaturePPSN V*. Springer, 1998, pp. 119–128.
[39] Y. Sano and H. Kita, "Optimization of noisy fitness functions by means of genetic algorithms using history of search with test of estimation," in *Evolutionary Computation, 2002. CEC'02. Proceedings of the 2002 Congress on*, vol. 1. IEEE, 2002, pp. 360–365.
[40] J. Branke, C. Schmidt, and H. Schmec, "Efficient fitness estimation in noisy environments," in *Proceedings of genetic and evolutionary computation*. Citeseer, 2001.
[41] S. Markon, D. V. Arnold, T. Back, T. Beielstein, and H.-G. Beyer, "Thresholding-a selection operator for noisy es," in *Evolutionary Computation, 2001. Proceedings of the 2001 Congress on*, vol. 1. IEEE, 2001, pp. 465–472.
[42] C. Qian, Y. Yu, and Z.-H. Zhou, "Analyzing evolutionary optimization in noisy environments," *arXiv preprint arXiv:1311.4987*, 2013.
[43] N. Metropolis and S. Ulam, "The monte carlo method," *Journal of the American statistical association*, vol. 44, no. 247, pp. 335–341, 1949.
[44] G. Huang, B. W. Baetz, and G. G. Patry, "A grey linear programming approach for municipal solid waste management planning under uncertainty," *Civil Engineering Systems*, vol. 9, no. 4, pp. 319–335, 1992.
[45] Y. Li and G. Huang, "Interval-parameter robust optimization for environmental management under uncertainty," *Canadian Journal of Civil Engineering*, vol. 36, no. 4, pp. 592–606, 2009.
[46] Y. Lv, G. Huang, Y. Li, Z. Yang, Y. Liu, and G. Cheng, "Planning regional water resources system using an interval fuzzy bi-level programming method," *J Environ Inform*, vol. 16, no. 2, pp. 43–56, 2010.
[47] G. Harik, E. Cantú-Paz, D. E. Goldberg, and B. L. Miller, "The gambler's ruin problem, genetic algorithms, and the sizing of populations," *Evolutionary Computation*, vol. 7, no. 3, pp. 231–253, 1999.
[48] E. Cantú-Paz, "Adaptive sampling for noisy problems," in *Genetic and Evolutionary Computation–GECCO 2004*. Springer, 2004, pp. 947–958.
[49] D. V. Arnold and H.-G. Beyer, "Local performance of the (μ/μi, λ)-es in a noisy environment," *Foundations of Genetic Algorithms*, vol. 6, pp. 127–141, 2001.
[50] H.-G. Beyer, "Toward a theory of evolution strategies: On the benefits of sexthe (μ/μ, λ) theory," *Evolutionary Computation*, vol. 3, no. 1, pp. 81–111, 1995.
[51] U. Hammel and T. Bäck, "Evolution strategies on noisy functions how to improve convergence properties," in *Parallel Problem Solving from NaturePPSN III*. Springer, 1994, pp. 159–168.
[52] B. L. Miller and D. E. Goldberg, "Optimal sampling for genetic algorithms," *Urbana*, vol. 51, p. 61801, 1996.
[53] A. B. C. Hilton and T. B. Culver, "Groundwater remediation design under uncertainty using genetic algorithms," *Journal of water resources planning and management*, 2005.
[54] Z. S. Kapelan, D. A. Savic, and G. A. Walters, "Multiobjective design of water distribution systems under uncertainty," *Water Resources Research*, vol. 41, no. 11, 2005.
[55] A. Chiesura, "The role of urban parks for the sustainable city," *Landscape and urban planning*, vol. 68, no. 1, pp. 129–138, 2004.
[56] D. E. Bowler, L. M. Buyung-Ali, T. M. Knight, and A. S. Pullin, "A systematic review of evidence for the added benefits to health of exposure to natural environments," *BMC Public Health*, vol. 10, no. 1, p. 456, 2010.
[57] D. E. Bowler, L. Buyung-Ali, T. M. Knight, and A. S. Pullin, "Urban greening to cool towns and cities: A systematic review of the empirical evidence," *Landscape and urban planning*, vol. 97, no. 3, pp. 147–155, 2010.
[58] D. C. Parker, M. J. Hoffmann, P. Deadman, D. C. Parker, S. M. Manson, S. M. Manson, M. A. Janssen, and M. A. Janssen, "Multi-agent systems for the simulation of land-use and land-cover change: a review," *Annals of the association of American Geographers*, vol. 93, no. 2, pp. 314–337, 2003. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.92.2240
[59] Y. Sasaki and P. Box, "Agent-based verification of von thünen's location theory," *Journal of Artificial Societies and Social Simulation*, vol. 6, 2003.
[60] L. Sanders, D. Pumain, H. Mathian, F. Guérin-Pace, and S. Bura, "Simpop: a multiagent system for the study of urbanism," *Environment and Planning B: Planning and Design*, vol. 24, no. 2, pp. 287–305, 1997.
[61] A. Ratle, "Optimal sampling strategies for learning a fitness model," in *Evolutionary Computation, 1999. CEC 99. Proceedings of the 1999 Congress on*, vol. 3. IEEE, 1999.
[62] C. Toregas, R. Swain, C. ReVelle, and L. Bergman, "The location of emergency service facilities," *Operations Research*, vol. 19, no. 6, pp. 1363–1373, 1971.
[63] A. A. Ageev and M. I. Sviridenko, "Approximation algorithms for maximum coverage and max cut with given sizes of parts," in *International Conference on Integer Programming and Combinatorial Optimization*. Springer, 1999, pp. 17–30.
[64] A. Krause, D. Golovin, and S. Converse, "Sequential decision making in computational sustainability through adaptive submodularity," *Ai Magazine*, vol. 35, no. 2, 2014.
[65] J. Rada-Vilela, M. Johnston, and M. Zhang, "Population statistics for particle swarm optimization: Resampling methods in noisy optimization problems," *Swarm and Evolutionary Computation*, vol. 17, pp. 37–59, 2014.
[66] J. Wu and A. J. Plantinga, "The influence of public open space on urban spatial structure," *Journal of Environmental Economics and Management*, vol. 46, no. 2, pp. 288–309, 2003.
[67] S. del Saz and L. García, "Estimating the non-market benefits of an urban park: Does proximity matter?" *Land Use Policy*, vol. 24, no. 1, pp. 296–305, 2007.
[68] V. Rieser and O. Lemon, *Reinforcement Learning for Adaptive Dialogue Systems: A Data-driven Methodology for Dialogue Management and Natural Language Generation*, ser. Theory and Applications of Natural Language Processing. Springer, 2011.
[69] O. A. A. C. W. L. M. W. F. I. H. A. Kennedy, M. C. and J. P. Gosling, "Quantifying uncertainty in the biospheric carbon flux for england and wales," in *Journal of the Royal Statistical Society A*, vol. 171, 2006, pp. 109–135.
[70] H. Kondo and F.-H. Liu, "A study on the urban thermal environment obtained through one-dimensional urban canopy model," *Journal of Japan Society for Atmospheric Environment/Taiki Kankyo Gakkaishi*, vol. 33, no. 3, pp. 179–192, 1998.
[71] D. Cornford, "Mucm: Managing uncertainty in complex models," 2010.
[72] N. Srinivas and K. Deb, "Muiltiobjective optimization using nondominated sorting in genetic algorithms," *Evolutionary computation*, vol. 2, no. 3, pp. 221–248, 1994.
[73] A. N. Aizawa and B. W. Wah, "Scheduling of genetic algorithms in a noisy environment," *Evolutionary Computation*, vol. 2, no. 2, pp. 97–122, 1994.
[74] A. T. Murray and R. L. Church, "Heuristic solution approaches to operational forest planning problems," *Operations-Research-Spektrum*, vol. 17, no. 2-3, pp. 193–203, 1995.
[75] J. Jacobs, *The death and life of great American cities*. Vintage, 1961.
[76] J. B. Smalley, B. S. Minsker, and D. E. Goldberg, "Risk-based in situ bioremediation design using a noisy genetic algorithm," *Water Resources Research*, vol. 36, no. 10, pp. 3043–3052, 2000.
[77] J. Ramsay, J. ten Berge, and G. Styan, "Matrix correlation," *Psychometrika*, vol. 49, no. 3, pp. 403–423, 1984.
[78] M. Vallejo, D. W. Corne, and V. Rieser, "Genetic algorithm evaluation of green search allocation policies in multilevel complex urban scenarios," *Journal of Computational Science JOCS*, 2015.