

# Quantifying uncertainty in Pareto fronts arising from spatial data

Moritz Hildemann<sup>a</sup>, Judith A. Versteegen<sup>b,\*</sup>

<sup>a</sup> University of Münster, Heisenbergstraße 2, 48149, Münster, Germany

<sup>b</sup> Wageningen University, Droevendaalsesteeg 4, 6708, Wageningen, the Netherlands

## ARTICLE INFO

### Keywords:

Spatial optimization  
Land use allocation  
Uncertain spatial data  
Uncertain Pareto fronts  
Seeding

## ABSTRACT

Multi-objective spatial optimization problems require spatial data input that can contain uncertainties. Via the validation of constraints and the computation of objective values this uncertainty propagates to the Pareto fronts. Here, we develop a method to quantify the uncertainty in Pareto fronts by finding the extreme lower and upper bound of the range of optimal values in the objective space, i.e. the Pareto interval. The method is demonstrated on a land use allocation problem with initial land use (for objectives and constraints) and soil fertility (for one objective) as uncertain input data. Pareto intervals resulting from uncertain land use data were wide and irregularly shaped, whereas the ones from uncertain soil data were narrow and regularly shaped. Furthermore, in some objective-space regions, optimal land use patterns remained relatively stable under uncertainty, while elsewhere they were clouded. This information can be used to select solutions robust to spatial input data uncertainty.

## 1. Introduction

Spatial optimization covers a range of approaches to find configurations of space that are optimal given one or more decision variables, one or more objectives, and one or more constraints (Cao, 2018). Most spatial optimization problems involve multiple conflicting objectives, such as minimizing production costs while minimizing green house gas emissions (Versteegen et al., 2017), or minimizing the travel time to hospitals while minimizing the costs involved in building new hospitals (Luo et al., 2017). In such cases, the spatial optimization does not result in a single optimal spatial configuration, but instead in a set of spatial configurations, called the non-dominated solutions. These configurations can all be considered optimal, but represent trade-offs between the conflicting objectives, e.g. high crop yield with high water usage or the other way around. When plotted in the objective space, i.e. plotting the values of the objective functions for these non-dominated solutions, they together form the Pareto front.

The spatial input data to an optimization problem may be uncertain. The sources of uncertainty are diverse; the data can contain errors, vagueness or ambiguity (Fisher et al., 2006). These uncertainties lead to arbitrariness in whether or not a solution is non-dominated, for two reasons. Firstly, the validation of the feasibility of solutions with the defined constraints may become uncertain. Secondly, the computed objective values become uncertain.

As such, the uncertainties from spatial input data propagate to the output of the optimization: the Pareto fronts (objective space) and the corresponding optimal spatial configurations (solution space). A quantification of the uncertainty in these two output dimensions would serve decision makers to assess the likelihood that their objectives are met for a point on the Pareto front as well as how the optimal spatial configuration may vary at this point (Autuori et al., 2016). Wide ranges in the Pareto fronts with low similarity in the optimal spatial configurations could serve as a warning, while narrow ranges and high similarity in the spatial configurations bring confidence in the selection process for decision-makers.

The question arises how the desired information of uncertainty in the optimization results can be obtained. A method to analyze the propagation of uncertainties and errors from model inputs to model outputs is Monte Carlo simulation (Heuvelink, 1998). In Monte Carlo simulation, samples are randomly drawn from probability distributions of the input variables and for each sample the model is run to obtain an estimate of the uncertainty in the model output (Anderson, 1976). Monte Carlo simulation can be applied to analyze the uncertainty in optimization outputs; for example, Villa et al. (2013) evaluate the objectives in a non-spatial optimization with uncertain data with a Monte Carlo simulation to approximate lower and upper bounds of the uncertain Pareto fronts. Here, the Monte Carlo is implemented within the optimization, i.e. interior sampling. However, when the constraint data are uncertain,

\* Corresponding author.

E-mail address: [judith.versteegen@wur.nl](mailto:judith.versteegen@wur.nl) (J.A. Versteegen).

the lower and upper bounds can be hard or impossible to estimate (Homem-de Mello and Bayraksan, 2014), because the feasible regions in objective space become non-convex (Ahmed and Shapiro, 2008). An exterior sampling method (Shapiro, 2003) can be used to simulate both uncertain objective and constraint data, where one sample is generated before the optimization starts and the optimization is performed with these samples. By repeating the sampling and the optimization execution, the probability distributions of the optimization outputs can be derived.

A limitation of the exterior sampling method is the high computational effort as it typically requires a high number of optimization executions. And, in multi-objective spatial optimization, the computational effort of one execution is often high. Uncertainty assessment in Pareto fronts from multi-objective spatial optimizations with a sampling procedure of uncertain constraint and objective evaluation data has not been researched yet, possibly because of this high computational effort. We aim to overcome this research gap with a method in which we first assess the uncertainty in computationally cheaper single objective evaluations and then use that information to produce Pareto fronts of the multi-objective optimization problem.

Hereto, we adapt an approach proposed by Guariso and Sangiorgio (2020), who solved single-objective optima first and used those solutions as elite members in a multi-objective optimization. Given that the single-objective optimal solutions of the multi-objective problem can be derived with a much lower computational effort, we can cost-effectively apply the exterior sampling method on the single-objective optima to compute the uncertainty in the single points at the outer ends of the Pareto fronts. To extend the available information from the single objectives to the uncertainty in the Pareto front, we execute the multi-objective optimization including the single objective optima, with a method called seeding (Friedrich and Wagner, 2015). Seeding entails that a part of the random initial solutions of the optimization is replaced by better solutions, in our case the single-objective optima. In our proposed approach, we seed the two samples from the uncertain input data that lead to the lower and upper bound extremes of the single objectives, consecutively. The resulting two Pareto fronts are estimations of the

lower and upper extremes of attainable Pareto fronts in the objective space, i.e. a Pareto interval. The total number of optimization executions is thus limited to the number objectives multiplied by two. We use a multi-objective land use allocation as a case study for the uncertainty assessment of multi-objective spatial optimizations.

The following research questions are answered in this work: 1) What is the effect of seeding single-objective optima into the initial set of solutions on the multi-objective land use allocation optimization? 2) What is the effect of uncertain spatial input data on the width and shape of the Pareto interval? 3) What is effect of uncertain spatial input data on the optimal spatial configurations?

## 2. Methods

### 2.1. Overview

In this work, we extend an existing multi-objective land use allocation optimization under four objectives (Strauch et al., 2019) with input data containing quantified uncertainties and methods to propagate this uncertainty to the objective and solution space (Fig. 1). The first step is to construct the probability distributions of the spatial input data, and then draw samples from these distributions (Section 2.3, Fig. 1a). The second step is to use the two input data samples resulting in the lower and upper bound of the single objectives as input data for the multi-objective spatial optimization. Furthermore, the single-objective optimal solutions are seeded into the first population of the optimization (Section 2.4.1, Fig. 1b). Finally, the effect of uncertainty in the spatial input data on the width and shape of the Pareto interval (Fig. 1c) and the similarity of solutions in the solution space is quantified and visualized (Section 2.4.2).

### 2.2. Optimization problem and algorithm

In this study, we build upon an existing land use allocation optimization, CoMOLA, that optimizes a land use raster under four objectives,

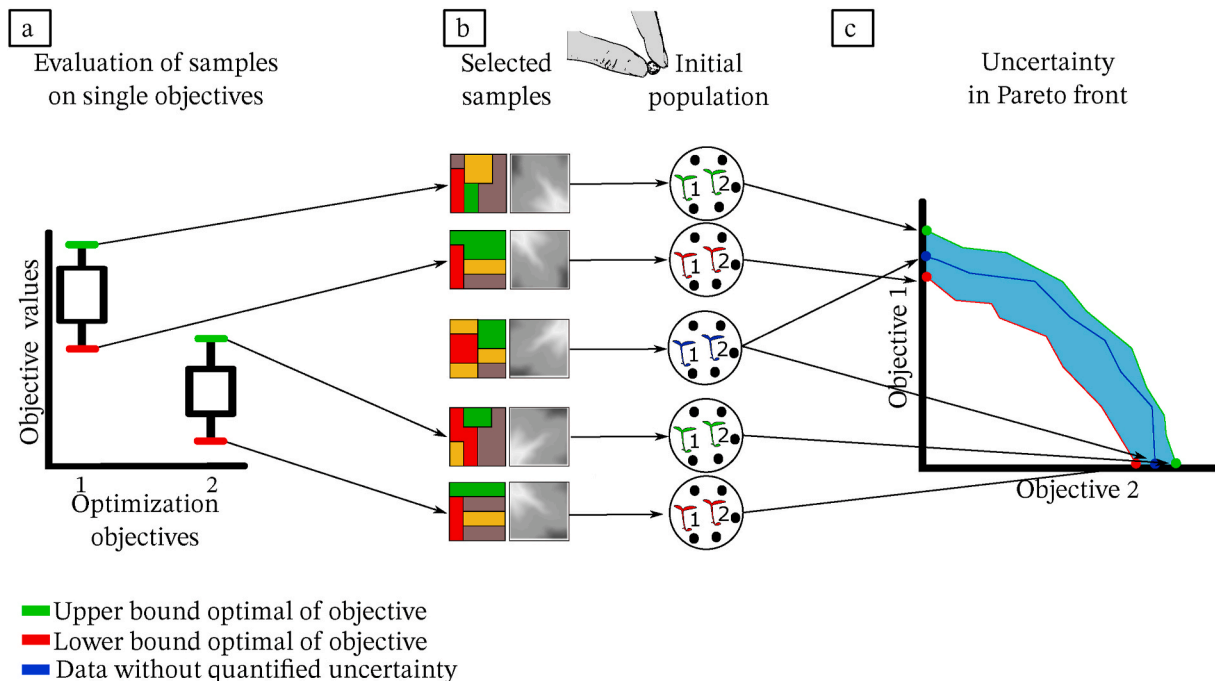


Fig. 1. Workflow to approximate extreme Pareto intervals from uncertain input data. a) Sample spatial data and compute objective values for each sample, b) Seed extreme lower and upper bound samples of land use (land use maps with colours indicating land use) and soil fertility data (soil fertility map), as well as the reference data (middle row) into initial population of Genetic Algorithm, c) Run Genetic Algorithm twice for each objective (upper bound and lower bound) and quantify and visualize uncertainty in Pareto fronts.

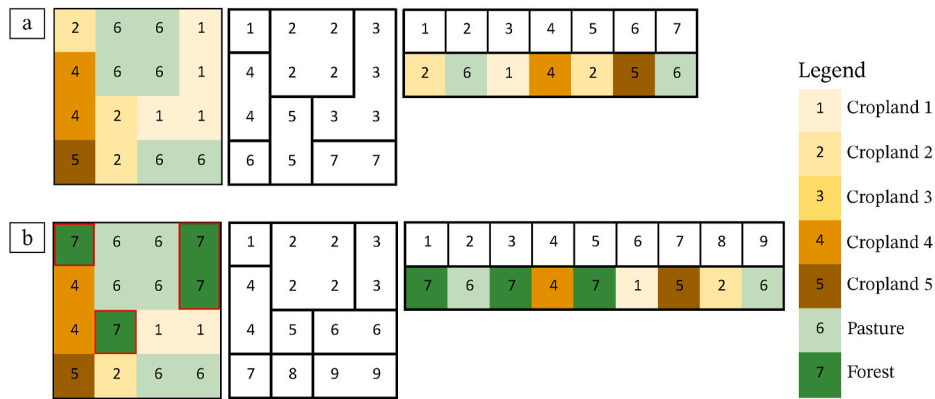


Fig. 2. Land use map, patch ID map and array of input land use map, a) Reference input b) Sample from uncertain land use data with changed grid cells in land use map highlighted in red.

optionally under constraints (Strauch et al., 2019). The optimization uses a multi-objective genetic algorithm called Non-dominated Sorting Genetic Algorithm II (NSGA II) (Deb et al., 2002). Genetic algorithms mimic the natural competition within a population consisting of individuals by means of reproduction, crossover and mutation of the individuals Holland (1984).

The decision variable of the optimization is the land use type of each grid cell, where the possible land use types are pasture, forest, urban, and cropland 1–5, representing five different levels of agricultural productivity. To be in line with CoMOLA, we use a synthetic initial land use raster of 10x10 cells. The representation of an individual in the NSGA II is an array with land use patches (Fig. 2). Herein, a patch is a set of contiguous cells with the same land use type. Each patch is assigned an identifier (ID).

Two types of constraints are used, land use transition constraints and area proportion constraints. The land use transition constraints are: urban areas can neither be extended nor removed, forest can only be converted to pasture, and pasture cannot be converted. All other conversions are allowed. The area proportion constraint was set to permitted ranges of 10–25% for forest, and of 10–30% for pasture. All other land use types have no area proportion constraints.

The four objectives are the maximization of forest species richness, habitat heterogeneity, water yield, and crop yield. They are explained below and their equations can be found in Appendix A.

The species richness (SR) objective function bases on empirical relationships between habitat area and species richness from MacArthur and Wilson (1967). The number of grid cells of land use forest defines the objective value, and the objective value is the sum of 5 times the forest area to the power of 0.2. An optimal valid solution for this objective is a map where the maximum area constraint of 25% forest is reached, the worst solution has the minimum permitted area of 10% forest.

The habitat heterogeneity (HH) is the sum over edges between different land use types, i.e. between patches. Edges have different weights: a higher land use intensity leads to a lower habitat heterogeneity. Edges with forest, pasture and cropland 1 have the lowest intensity and get a weight of 1, edges with cropland 2–5 have corresponding intensity weights of 2–5. Edges with urban are ignored and therefore do not contribute to a higher habitat heterogeneity. The optimal valid solution for this objective has the highest possible number of edges between forest, pasture and cropland 1 obtainable under the maximum area constraints of 25% forest and 30% pasture. The worst solution has 10% forest, 10% pasture and 80% urban in three large patches.

The water yield (WY) objective function is based on the relative differences in evapotranspiration rates between land use types. The total water yield is computed by summing all land use areas divided by the land-use specific evapotranspiration rate. The evapotranspiration rates are, in increasing order: cropland 1 = 0.900, cropland 2 = 0.925, cropland 3 = 0.950, pasture = 0.960, cropland 4 = 0.975, cropland 5 = 1.00, forest = 1.14. The optimal valid solution for this objective is found when the minimum area constraint of 10% forest and pasture are reached and the other 80% of the study area is cropland 1. The worst solution is composed by 25% forest, 30% pasture and 45% cropland 5.

Crop yield (CY) is the sum of all logarithmic products of cropland intensity and soil fertility over all cells. It is the only objective function for which a second spatial input data set is required; a soil fertility map with values ranging from 0.1 to 1. The optimal valid solution for this objective is found when all permitted land use is transitioned into cropland 5 on the cells with the highest soil fertility value while the transition constraints are not violated (10% forest, 10% pasture, 80% cropland 5). The worst solution has no cropland at all.

### 2.3. Quantify uncertainty in input data and create samples and seeds

#### 2.3.1. Uncertainty in data for constraints

The synthetic input land use map in CoMOLA (Strauch et al., 2019) is the reference input land use map in this work. Given that no information about its accuracy is available, we use overall land use class errors to construct the uncertainty in this map, as explained in the following.

A measure to quantify uncertainty in land use classifications are confusion matrices (Fang et al., 2006). Confusion matrices indicate for every class how often a class was correctly classified as such, how often a class was incorrectly assigned to another class (commission error), and how often a class was not classified as such (omission error) (Foody, 2002). We assign omission errors to every cell as the probability to be reclassified, on the basis of the land use type in the reference input land use map. If the classification accuracy of land use A is high, then the reclassification probability is low. But, if land use A was often incorrectly classified as land use B, then the probability of reclassifying a cell of class B to A is accordingly high.

The confusion matrix of land cover data from GlobCover (Bicheron et al., 2008) is used to derive the reclassification probabilities (Table 1). The optimization handles land use instead of land cover data, which makes a mapping between land cover and land use classes necessary. The mapping is displayed in Appendix C; Table A1. The probabilities were computed by dividing the number of classifications by the row sum.

**Table 1**

Relative class omission errors in percent derived from the confusion matrix with correct classifications highlighted in the diagonally. Columns represent observations, rows represent the map.

Class	Cropland 1	Cropland 2,3	Cropland 4	Cropland 5	forest	Pasture	Urban
Cropland 1	12.5	0	60.4	6.3	15.6	31.25	2
Cropland 2,3	0	7.4	55.6	7.4	14.8	2.2	5.2
Cropland 4	0	0	76.2	15.7	4.3	1	2.7
Cropland 5	0	0	17.4	82.6	0	0	0
Forest	0	0	8.4	0.5	90	0.5	0.5
Pasture	0	0	18.7	0.7	22.3	53.2	5
Urban	0	0	8.3	0.8	0.8	0	75

The reclassification of the grid cells on the basis the omission errors does not only influence the land use map but may also affect the patch ID map (Fig. 2). This happens when a cell in a patch is transitioned while the other cells in the patch are not or are transitioned to another land use type. If the patch ID map changes, the array length changes accordingly.

### 2.3.2. Uncertainty in data for objective evaluations

Besides land use as an input for the objective evaluations (see previous section for uncertainty estimation), the objective crop yield is computed with additional spatial data: a soil fertility map. Soil fertility or soil quality maps are typically derived by interpolating field samples (point data) of the physical, chemical, and biological properties of the soil (Klimkowitz-Pawlas et al., 2019). Soil quality estimations from interpolating points are uncertain and the spatial uncertainty depends on the spatial arrangement of samples (Hunsaker et al., 2013).

We simulate the uncertainty by synthetically positioning field samples on the original soil fertility map of CoMOLA and interpolating between them (Fig. 3a). The geostatistical interpolation method kriging is used for interpolating the samples (Ismaili et al., 2014) on a resolution 50 times finer than the original raster. Along with the expected values (Fig. 3b), kriging generates estimated variance map (Fig. 3c) (Wackernagel, 1995). Lastly, we derive the average variance for each cell of the 10x10 cell extent of the study area (Fig. 3d).

To obtain a sample of the soil fertility map, we draw independent values for each cell from a Gaussian distribution with the mean value from the reference soil fertility map, and the standard deviation according to the variance map from the kriging (its square root). Values above 1 are set to 1 and values below 0 are set to 0 to maintain the original value scale.

### 2.3.3. Sampling and seeding procedure

In total, 1000 samples are realized from the uncertain land use data (Section 2.3.1), and 1000 samples are realized from the uncertain soil fertility data (Section 2.3.1). The number 1000 was selected by iteratively assessing the maximum single objective values of the samples. The optimal single-objective value did not longer increase for any of the objectives after 410 iterations (Appendix B; Fig. A1). That indicates that 1000 samples suffices to estimate the single-objective extreme optimal solutions.

We compute the single-objective optima for all samples. Hereto, we use knowledge about the objective functions to compute the single-objective optima deterministically. The optimal land use configurations for every objective are derived by replacing land uses with the optimal land use per objective while meeting the area and transition constraints (see Section 2.2).

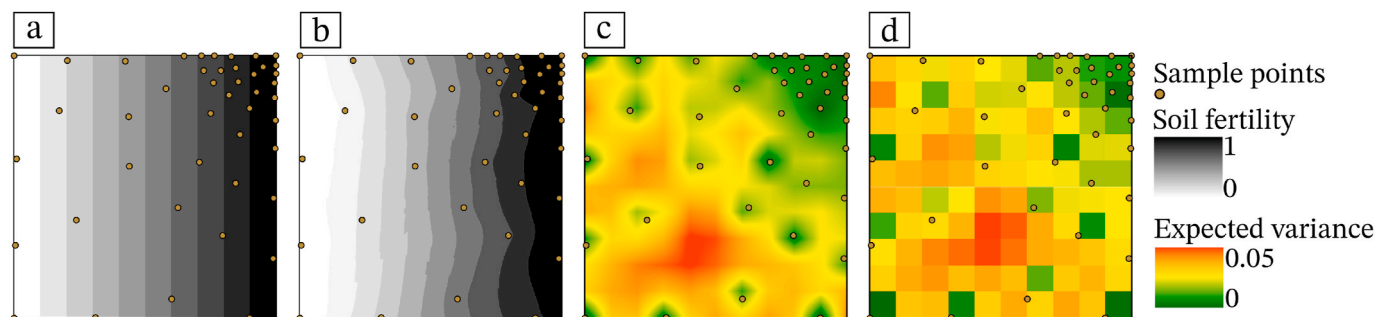
The first step of the exterior sampling method is finished with the detection of the samples leading to the extreme lower and upper bounds of the single-objectives. The next step the execution of the optimization with the extreme lower and upper bound samples as spatial data inputs. In addition, for each execution of the optimization, the four optimal single-objective optimal solutions belonging to these inputs are seeded into the initial population of the multi-objective Genetic Algorithm, following the method of Guariso and Sangiorgio (2020).

### 2.4. Uncertainty assessment of Pareto fronts from extreme samples

#### 2.4.1. Experimental design

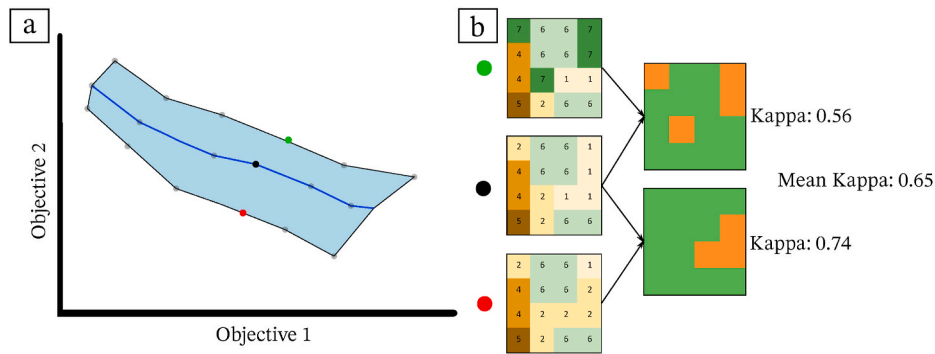
We execute the experiments with following different input data:

1. With the reference data from Strauch et al. (2019) without quantified uncertainty and excluding the seeding procedure.
2. With the reference data from Strauch et al. (2019) without quantified uncertainty and including the seeding procedure. This experiment is meant to demonstrate the effect of the seeding only.
3. With the eight extreme lower and upper bound samples of the uncertain land use data and the original soil data and including the seeding procedure. Therefore, the reference soil fertility map of CoMOLA and eight different land use maps are used as data input in eight optimizations. This experiment is meant to demonstrate the effect of uncertainty in the initial land use data on the objective and solution space.
4. With the extreme lower and upper bound samples of the soil data for objective crop yield and the original land use data and including the seeding procedure. That means, two different soil fertility maps and the reference land use map of CoMOLA are used as data inputs in two optimizations. This experiment is meant to demonstrate the effect of uncertainty in the soil fertility data on the objective and solution space.



**Fig. 3.** Uncertainty in soil fertility map, a) original soil fertility map with samples, b) interpolated surface from with the samples, c) expected variance from interpolation method ordinary kriging, d) expected variance resized to grid cell size of original soil fertility map (10x10).





**Fig. 4.** Quantification and visualization of uncertainty in Pareto fronts. a) The convex hulls of the extreme Pareto fronts b) The Kappa statistics as metric for agreement between the resulting land use maps belonging to the closest solutions in objective space. The higher kappa, the higher the similarity between maps.

5. With the extreme lower and upper bound samples of the combined uncertain soil and land use data for objective crop yield and including the seeding procedure. Each land use data sample is combined with each soil data sample. Two different soil fertility maps and two different land use map that lead to the extreme lower and upper bound crop yield values are used as data inputs in two optimizations. This experiment is meant to demonstrate the effect of uncertainty in the initial land use data and soil fertility data together on the objective and solution space.

In total, 13 different optimizations are executed. Each of these optimizations is executed 10 times to account for the stochasticity in NSGA II. In line with CoMOLA, we use a population size of 300 over 300 generations, a crossover rate of 0.9, and a mutation rate of 1 divided by the number of spatial units of the individual, which is the number of patches. Runs are performed in parallel on a high-performance Linux cluster (MEGWARE cluster with 15.120 cores, 412 nodes and Intel Xeon Gold 6140 18C 2.30 GHz processors).

2.4.2. Quantification and visualization of uncertainty in the Pareto fronts

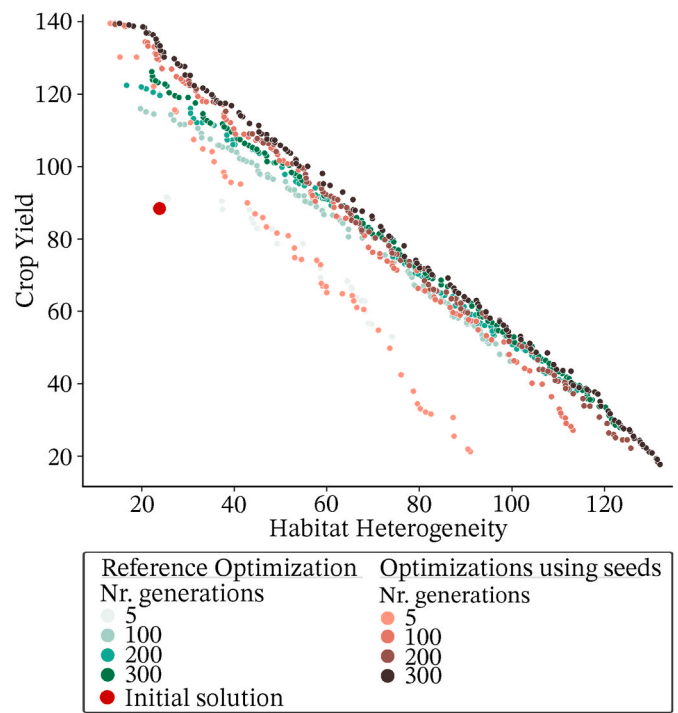
The resulting Pareto front from each optimization has four dimensions, as four objectives are optimized. We use a scatter plot matrix, which illustrates the position of the objective values in a 2D-scatter plot for each combination of objectives (Ibrahim et al., 2016). In the plots, we show the lower bound front, the upper bound front, and the Pareto front from reference data (reference front). To illustrate the Pareto interval resulting from the uncertain data, we plot convex hulls between the three (Fig. 4a). The plot’s axes are normalized to the range from 0 to 1 to show the relative influence of the uncertainty for each objective. Herein, where 0 is the worst objective value and 1 is the best objective value of the reference Pareto front.

For quantifying and visualizing the difference between the resulting spatial configurations of the non-dominated solutions, the Kappa statistic is used as a metric of similarity between the land use maps. Kappa has a range from -1 (all cell values differ) to 1 (all cells values cohere) (Monserud and Leemans, 1992). For each non-dominated solution obtained from the optimization without quantified uncertainty (Fig. 4a, black point), the closest non-dominated solution in the objective space is selected from each of the two Pareto fronts obtained with the extreme lower and upper bound samples, respectively (Figs. 4a, 5-6, red and green point). For the two pairs, the Kappa statistic is calculated, and averaged over the pairs (Fig. 4b). The non-dominated solutions from the optimization without quantified uncertainty are then colored according these average Kappa values to visualize the level of agreement in spatial configuration (the solution space) for each part of the Pareto front.

3. Results and discussion

3.1. Reference data input with and without seeding

The experiments 1 and 2 (Sec. 2.4.1) are compared to assess the effect of the seeding. The Pareto front of habitat heterogeneity against crop yield obtained with the seeding procedure has a wider spread of solutions in the objective space than the Pareto front obtained without seeding (Fig. 5). Furthermore, the optimal solutions of the optimization with seeding dominate the optimal solutions obtained without seeding after 300 generations (Fig. 5). The extreme values are 7% better for habitat heterogeneity and 10% better for crop yield. For the other two objectives, not shown in Fig. 5, the improvements +9% (water yield) or the same optimal objective values were obtained (species richness). Finally, the optimization with seeding leads to faster convergence. This is, for example, illustrated by the fact that the Pareto front obtained with seeding in the 200th generation already dominates the Pareto front in the 300th generation without seeding for a crop yield of 85 (Fig. 5).



**Fig. 5.** Comparison of the convergence of optimization with and without seeding for the objectives habitat heterogeneity and crop yield.

It can be concluded that the seeding has positive effects on the quality of solutions in the Pareto fronts as was also concluded by Friedrich and Wagner (2015) for a non-spatial optimization. The high quality of the Pareto front after 200 generations is in line with their finding that seeding lowers the computational demand of an optimization. Our observation of the wider spread of solutions in the objective space also support the finding from Guariso and Sangiorgio (2020) that injecting single-objective optimal solutions as seeds into the initial population leads to a better covered objective space. In sum, our results support the application of the seeding for uncertainty analysis in the Pareto fronts.

### 3.2. Probability distributions of the sampled single-objective optima

The probability distributions of the optimal single-objective values (Fig. 6a-d) of habitat heterogeneity (Fig. 6a), water yield (Fig. 6c) and crop yield (Fig. 6d) from the samples of the uncertain land use data resemble normal distributions. The optimal objective values of habitat heterogeneity and water yield from the reference land use data are close to the mean of the probability distributions, so roughly half of the samples lead to better objective values and the other half to worse objective values. In contrast, 97% objective values of crop yield from the samples are worse than the objective value from the reference data. The unbalance is caused by the misclassification probabilities in the confusion matrix: on average, the land use types contributing to the objective crop yield (cropland 1-5) were more frequently wrongly classified as such than land use types not contributing to the objective crop yield, especially cropland 1,2 and 3 (Table 1). As a consequence, fewer cropland cells than non-cropland cells are expected in the samples, because they are more often reclassified to forest, pasture and urban than vice versa.

The objective values for species richness are, in contrast to those of other objectives, not uncertain. This is directly linked to the maximum allowed amount of land use type forest. The optimal species richness for a land use map with 25% forest cells is 9.51 and this 25% can not be exceeded due to the predefined area constraints. The reason for no values below 9.51 is again related to the confusion matrix (Table 1). The species richness can only be lower if more than 75% of all cells were reclassified to land use types that cannot be converted to forest (defined in transition constraints). The only land use types that cannot be converted to forest are pasture and urban. Samples with more than 75% of either pasture or urban are theoretically possible but highly unlikely given the probabilities in the confusion matrix.

The distribution of crop yield values with the reference land use data but with the uncertain soil data (Fig. 6e) is, with a range 4.70, narrower than the distribution from uncertain land use data, with a range of 45.3. Most samples are equal or close to the crop yield objective value obtained with the original soil data without quantified uncertainty. The combined uncertain data for the objective crop yield result in the widest range of 52.4. The combination of 1000 uncertain land use and soil data samples also results in a smoother frequency distribution. The reason for the smoother distribution is the higher sample size of 1 million (1000 \* 1000 combinations).

The ranges indicate that the uncertain land use data have a higher impact on the objective value crop yield than the uncertain soil data. This is because a reclassification from cropland to non-cropland reduces the crop yield to zero. On the other hand, the errors in the soil fertility map are likely to still allow crop production even when the soil fertility is reduced. In a comparison of set-ups with different strengths of the constraints Strauch et al. (2019) found that constraints limited the attainable optimal solutions, for example higher area constraints lead to a narrower Pareto front. In our case, the area constraints remain the

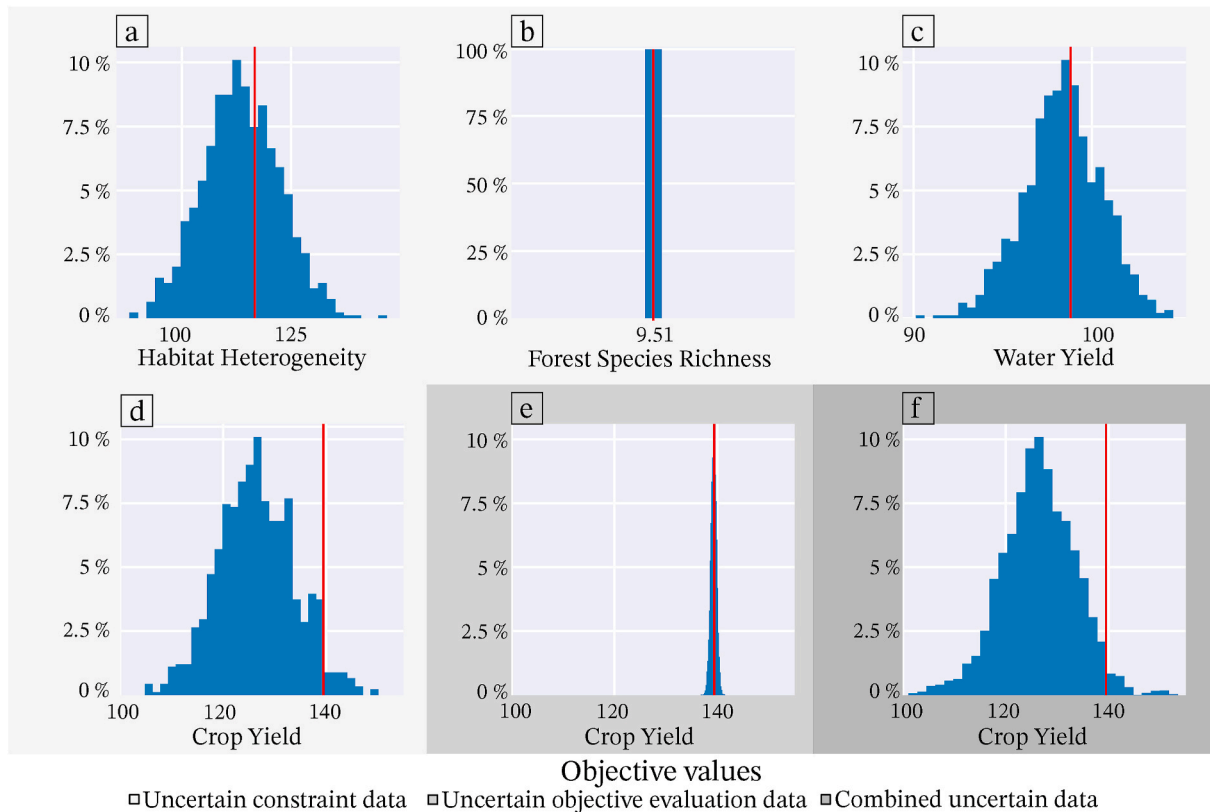


Fig. 6. Histograms of objective values and reference lines (red) displaying the values obtained with the original data input. a)-d) show the distributions of the objective values attained in the sampling procedure with uncertain land use data input, e) shows the distribution of the uncertain data for objective evaluation water yield and f) shows the distribution of combined uncertain constrain data and uncertain data for objective evaluation water yield.

same, but the strength of the transition constraints varies with every sample because the amount and position of cells that underlay the transition constraints differ. Our results show that uncertain spatial input data affects the objective space limitation in land use allocation optimizations when transition constraints are defined. The objective space limitations are not only affected by the definition of what land use types underlay transition constraints, but also by the quantified uncertainty in the land use data inputs. Furthermore, the probability distribution of the optimal single-objective solutions for the objective species richness lead to the conclusion that a strong constraint can cancel out the effect of other sources of uncertainty.

### 3.3. Uncertainty in the objective space

The third experiment with the extreme lower and upper extreme samples of uncertain land use data (Sec. 2.4.1) resulted in Pareto

intervals with diverse shapes, rotations and areas. Those geometric properties of the Pareto intervals allow to assess both information about the trade-offs between objectives and the uncertainty. They could be categorized into:

1. Non-conflicting objectives: Pareto fronts containing one single point become horizontal or vertical lines when one objective is uncertain (or rectangular Pareto intervals when both are uncertain, but this situation is not observed in our results)
2. Conflicting objectives: non-horizontal and non-vertical Pareto fronts become non-horizontal and non-vertical Pareto intervals

The two objectives habitat heterogeneity and species richness represent category 1: non-conflicting objectives. The synergy between the objectives is explained by the common property that maximum number of forest cells is optimal. With uncertain spatial input data the

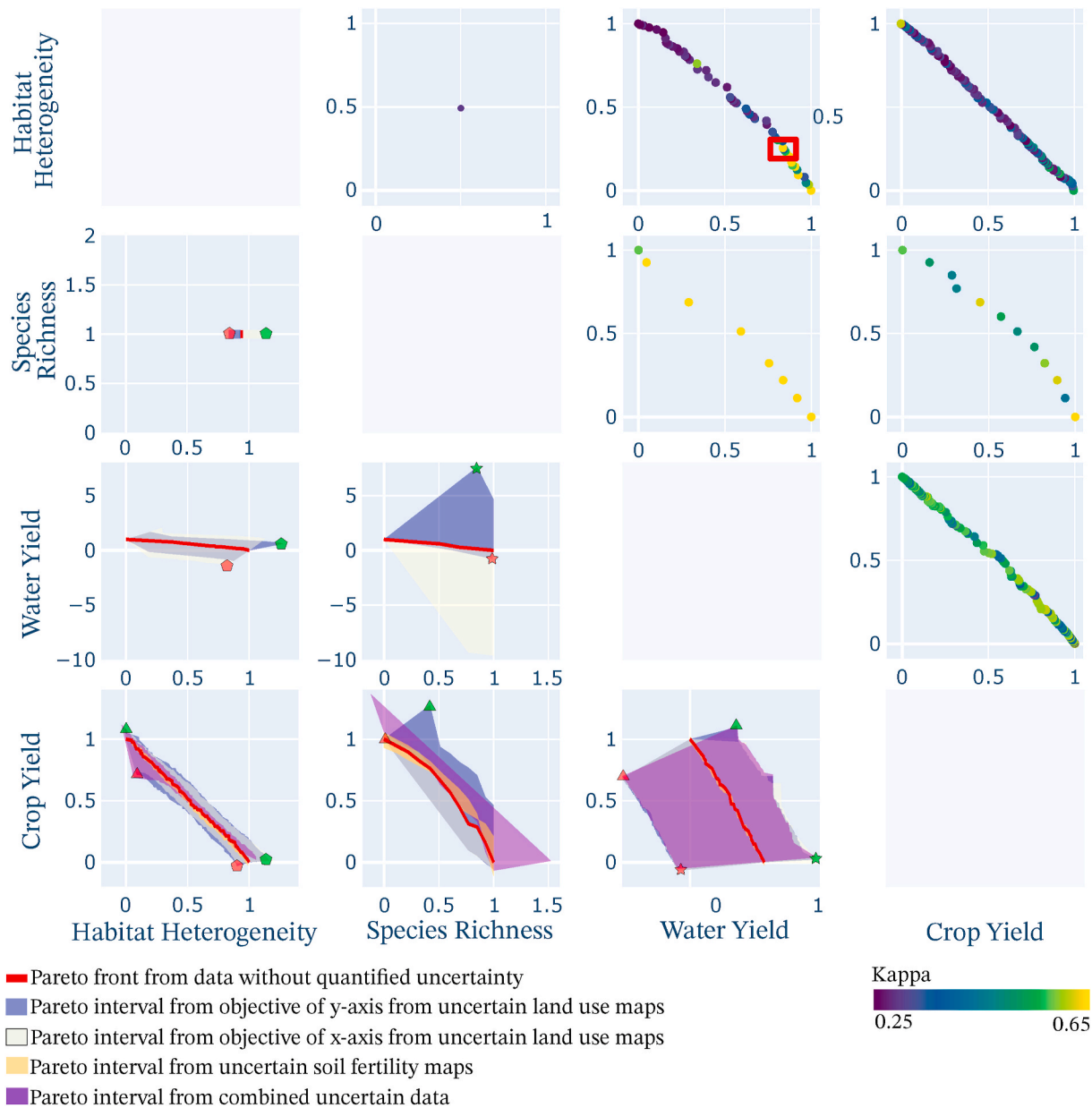


Fig. 7. Pareto intervals resulting from uncertain input data on the lower-left side of the matrix and the reference Pareto front with the Kappa statistic, representing the similarity in solution space, on the upper-right side of the matrix. The symbols (star, triangle and pentagon) represent three selected solutions for which the corresponding spatial configurations are illustrated in Fig. 8. The highlighted region in row one, column two shows the two solutions for which the corresponding spatial configurations, and those of the closest solutions on the other two fronts, are illustrated in Fig. 9.

species richness remains constant over the experiments, while the habitat heterogeneity varies (Fig. 7 row 2, column 1). It was possible in all cases to produce a solution with a maximum habitat heterogeneity that also has the maximum allowed 25% forest cells and thus maximum species richness.

All other objective pairs fall into the second category, conflicting objectives. With uncertain spatial input data the lines become intervals, where the width of the interval increases with the uncertainty. The angle of the Pareto fronts with conflicting objectives how much the values of one objective decrease with a certain increase in another objective, i.e. the trade-off between the two objectives. The area of the Pareto interval is therefore affected by the uncertainty in the objective values and the trade-off.

In some cases, the angles of the Pareto fronts, i.e. the trade-offs between the objectives, remain equal, regardless of the uncertain input data, e.g. for the pairs water yield - habitat heterogeneity (Fig. 7 row 3, column 1) and crop yield - habitat heterogeneity (Fig. 7 row 4, column 1). Here, all extreme lower and upper bound samples resulted in similarly shaped Pareto fronts where mainly the positions of the fronts change.

For other objective pairs the trade-offs change because of the uncertainty in the spatial input data, e.g. for the objective pair water yield - species richness (Fig. 7 row 3, column 2). The Pareto fronts obtained with the lower and upper extreme samples consist of two optima, while the reference Pareto front consists of eight optima. The reason for the different number of feasible optimal solutions in the Pareto fronts is the different number of transition-constrained cells in the uncertain land use data. The more transition-constrained cells exist, the lower the possible number of optimal land use allocations.

The fourth experiment with uncertain soil data only (Sec. 2.4.1) results in Pareto intervals with a width of only  $\sim 7.1\%$  in the x-dimension and  $\sim 5.5\%$  in the y-dimension of the Pareto intervals from uncertain land use data only (averaged over all three objective pairs in row 4 in Fig. 7). In the objective pair crop yield - habitat heterogeneity, and crop yield - water yield (Fig. 7 row 4, column 1), the Pareto intervals constantly enclose the reference Pareto fronts over the whole range of

the objective space, in line with the assessment of uncertain objective evaluation data by Bassi et al. (2018). Only in the objective pair crop yield - species richness (Fig. 7 row 4, column 2), a small part of the reference Pareto front is not enclosed by the Pareto interval; we believe that this is an artifact of the stochasticity in the optimization algorithm.

The Pareto intervals from the fifth experiment with the combined uncertain data inputs (Sec. 2.4.1) are displayed in the last row of Fig. 7. For the objective pair water yield - crop yield (Fig. 7 row 4, column 3), the Pareto interval has a similar shape and the interval is  $\sim 2\%$  wider than the Pareto interval from the uncertain land use data only. In the objective pair species richness - crop yield (Fig. 7 row 4, column 2), the combined uncertain data inputs lead to a  $\sim 24\%$  wider range between the extreme points compared to Pareto interval from the uncertain land use data only. The objective pair habitat heterogeneity - crop yield (Fig. 7 row 4, column 1), resulted in a Pareto interval with a range that increases with increasing crop yield values. The cone-shape can be explained by the increasing influence of soil fertility (and thus its uncertainty) with an increasing crop production. On the other side of the Pareto front (low crop yield), a high amount of forest and pasture is present, because these land uses improve habitat heterogeneity. These land use types can not be converted into cropland, which diminishes the effect of the soil fertility on the Pareto front.

In our results, not all shapes of the Pareto fronts from uncertain land use and combined uncertain land use and soil data resemble each other, which is in contrast with the findings about uncertain objective evaluation data by Bassi et al. (2018). This is caused by non-convex feasible objective space (Homem-de Mello and Bayraksan, 2014), that is introduced by the uncertain land use maps in combination with the land use transition constraints. The selected maps in Fig. 8 (highlighted in Fig. 7) illustrate the reason for the non-convex feasible space. One common factor that determines the interval between the Pareto fronts is the number of urban cells, because urban cells do not contribute to any of the objectives and can not be transformed into another land use, so more urban cells lead to worse objective values. E.g., the three examples from the upper bound Pareto fronts (green in Fig. 8), have a low amount of urban land with 5%, 3%, and 4%, while the lower bound Pareto front

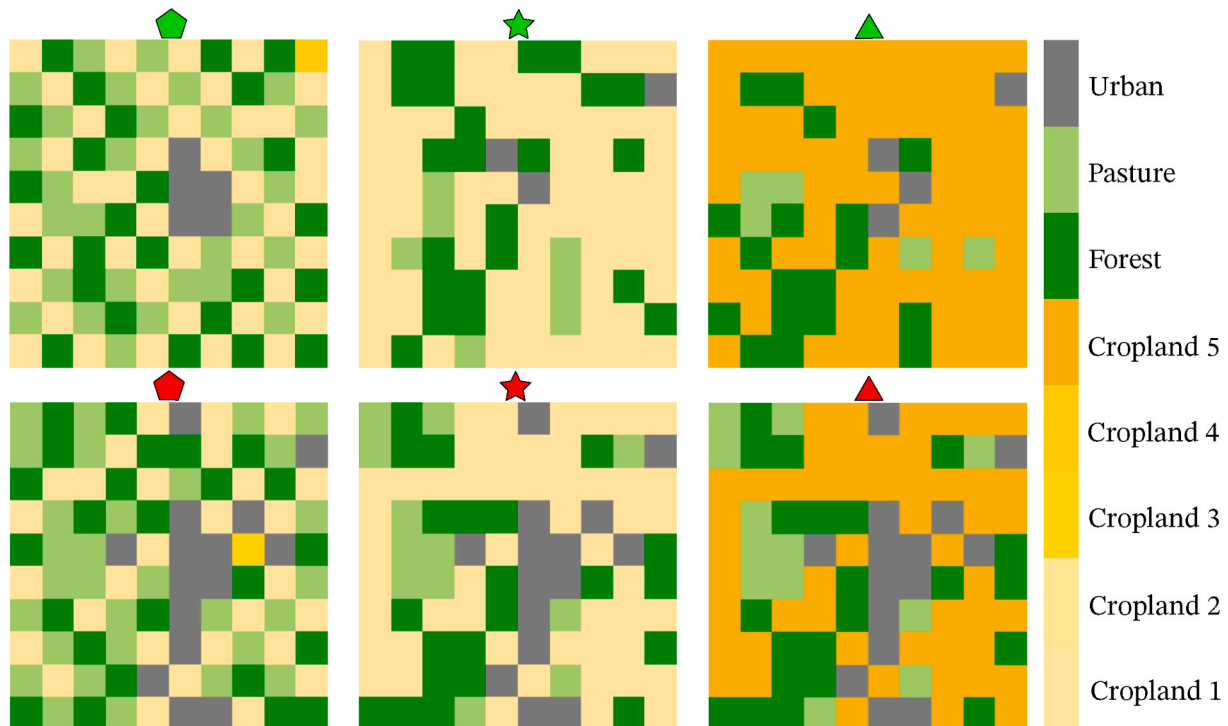


Fig. 8. Selected spatial configurations of solutions in Pareto fronts obtained with uncertain data, green symbols represent solutions in Pareto fronts obtained with the upper bound extreme samples and red symbols represent solutions in Pareto fronts obtained with the lower bound extreme samples.



(red) all have 15% urban land. The reason for the different number of cells in the upper bound Pareto fronts is that not only the amount of land use is of importance, but also the location. The uncertainty in the uncertain land use data led to samples, in which both the amount and the location of cells benefiting the objective limit the feasible objective space, leading to the non-convex feasible objective space under uncertainty.

The uncertainty in the amount and location of urban cells is thus a strong determinant of output uncertainty. Validating urban areas can therefore help to reduce uncertainty in all pairwise Pareto fronts. Ancillary data can be used to increase the accuracy of 75% in urban area classifications (Table 1), for example existing maps defining the outlines of urban areas (Corbane et al., 2021).

### 3.4. Uncertainty in the solution space

The average Kappa values as well as the variation in Kappa values differ per objective pair (Fig. 7). All Pareto fronts containing habitat heterogeneity have relatively low Kappa values (Fig. 7). In addition, the solutions with a higher habitat heterogeneity have increasingly low Kappa values. This is because habitat heterogeneity is computed by summing up heterogeneous land use edges (independent of the specific land use type), while the Kappa statistic compares the land use type in the cells. For example, if Kappa was computed between a land use map and the transposed version of this map, the Kappa value would likely be low, but the habitat heterogeneity would remain the same, as only the number of edges counts.

In the Pareto front between species richness and water yield, all Kappa values are similar with a range of only 0.08, while the Kappa values of the Pareto fronts between species richness and crop yield, and water yield and crop yield have ranges of 0.21 and 0.3, respectively. High Kappa values indicate a high robustness in the solution space (Fig. 9). The solutions with a high Kappa (Fig. 9, upper row) have more spatial characteristics in common between the uncertain solutions than the ones with a low Kappa (Fig. 9, upper row). For example, in the high Kappa spatial configurations the North and East are dominated by Cropland 1 with few patches of forest or pasture, while the South and East contain more forest. The spatial configurations with a low Kappa value (Fig. 9, upper row) have fewer characteristics in common. Even though the objective values vary less than 0.5%, this variation in the

solution space is crucial for decision makers as land use is their decision variable. Input data uncertainty becomes most problematic when it causes unclarity in the characteristics of the optimal spatial configuration for the selected point at the Pareto front. The solution space uncertainty computed here can serve as an incentive to select a different, more robust compromise solution.

### 3.5. Limitations and future work

We were able to approximate the uncertainty in the Pareto fronts with highly reduced computational effort compared to a full Monte Carlo. A full Monte Carlo with the selected sample size of 1000, would require 1000 executions of the optimization. The computational effort was reduced by 98.7% to 13 executions. The additional computation time to compute the objective value is only ~ 10% of a single optimization execution time. The main limitation of our work is that the distribution and shape of Pareto fronts between the two extremes remain unknown. Hereby, we miss information about how uncertain land use data that affects the land use transition constraints may lead to irregular infeasible regions (Homem-de Mello and Bayraksan, 2014). That missing information is the trade-off between the highly reduced computation time and available information about the propagation of input data uncertainty. Another trade-off from not performing an optimization for all samples is that there is no guarantee that the data samples resulting in the extreme single-objective values lead to the widest Pareto interval. It is possible that other samples result in more extreme Pareto front sections in the center, i.e. away from the single objective optima.

Moreover, the amount of uncertainty in our case study was based on empirical data but fixed. We did not investigate the effect of varying amounts of uncertainty in the spatial input data on the Pareto fronts. In future work, this may be assessed, for example by means of a sensitivity analysis.

Furthermore, our visualization of uncertainty in the Pareto fronts is shown in separate scatter plots in a matrix containing all pairwise Pareto fronts. As such, the uncertainty distribution over the whole objective space is not directly observable. Part of possible future work is incorporating both the extreme intervals and the similarity metric with all objectives in one single visualization, for example with the multi-objective visualization method of 3D-RadVis (Ibrahim et al., 2016).

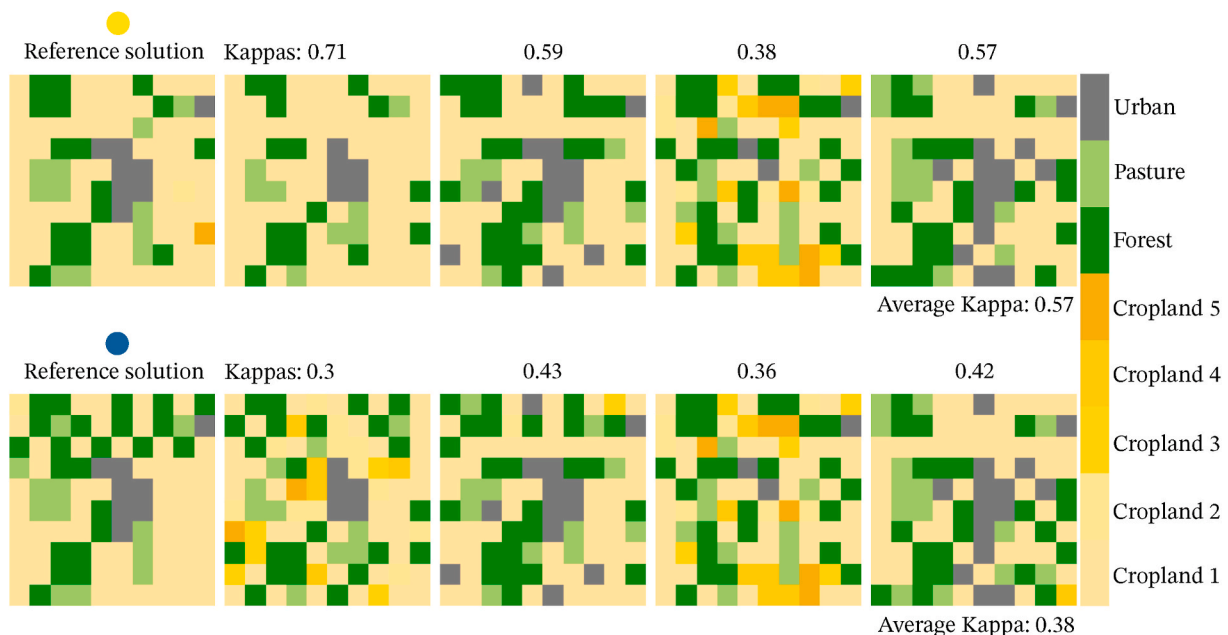


Fig. 9. Selected solutions in Fig. 7 (highlighted red area in right of diagonal) with one high and one low Kappa value with the maps used for calculating the Kappa.

Yet, it should then be tested whether this amount of information in one plot is still understandable (Senaratne et al., 2012). Lastly, a real-world case study of a spatial optimization is important future work. Such application will also allow comparing the propagation of uncertainty from different data sources to the optimization outputs.

#### 4. Conclusion

In this work, we used seeding of optimal solutions in combination with an exterior sampling method to estimate the propagation of uncertainty from spatial input data to the results of a multi-objective spatial optimization. This approach allowed to reduce the computation time by 98.7% compared to a full Monte Carlo. The approach was demonstrated on a land use allocation optimization with uncertain constraint and objective evaluation data.

Our first research question was: What is the effect of seeding single-objective optima into the initial set of solutions on the multi-objective land use allocation optimization? The optimization with the seeding procedure resulted in Pareto fronts with +6.25% better objective values (averaged over all four objectives) compared to the optimization without seeding. Furthermore, the optimization with seeding produced solutions with the same quality in shorter time: After 200 generations, the Pareto front from the optimization with seeding dominated the Pareto front without the seeding procedure after 300 generations.

Our second research question was: What is the effect of uncertain spatial input data on the width and shape of the Pareto interval? We found two types of shapes: 1) A horizontal or vertical line, which was obtained between two non-conflicting objective of which one was affected by the uncertainty in input data, and 2) A non-horizontal and non-vertical Pareto intervals, which was obtained between two conflicting objective of which at least one was affected by the uncertainty in input data. In our case study, the uncertain initial land use data had by far the highest effect on the width and shape of the Pareto intervals. The uncertain soil fertility data led to ranges in the Pareto intervals that were ~ 6% of the ranges from uncertain land use data. In the two objective pairs affected by both inputs, the combined Pareto intervals were ~ 2% and ~ 24% wider compared to the Pareto intervals from uncertain land use data only.

Furthermore, in line with previous research on uncertainty in non-spatial optimizations, we found that uncertain soil fertility data used for computing objective values resulted in regular Pareto intervals, while uncertain land use data that affects the transition constraints resulted in irregular Pareto intervals.

Our third research question was: What is the effect of uncertain spatial input data on the optimal spatial configurations? Given that the number and location of urban cells in the input land use map affected the

ranges of all pairwise Pareto fronts, we conclude that increasing the classification accuracy of urban land is an option to reduce uncertainty, for example by using auxiliary data. Furthermore, trends of the Kappa values could be observed within the single Pareto fronts. The Pareto front of the objective-pair species richness and water yield had comparatively high Kappa values, whereas the Pareto fronts containing habitat heterogeneity had comparatively low Kappa values respective to the other objective pairs. The solution space uncertainty computed here can serve as an incentive to select compromise solutions that are robust to uncertainty in the spatial input data.

#### Software and data availability

The dataset ‘‘Sampling procedure of land use and soil fertility map under uncertainty’’ is available at Hildemann (2021b). The dataset contains the input data, the Python code and the output data along with a description and steps to reproduce. Developer: M. Hildemann, first author, see contact details at first page. Year first available: 2021. Running the sampling procedure requires Python version 3.8 with NumPy 1.20.1, pandas 1.2.3. Plotting requires seaborn 0.11.1, scipy 1.6.2 and Matplotlib 3.3.4. The software is free.

The dataset ‘‘Algorithm for constrained multi-objective land use allocation optimization under uncertainty’’ is available at Hildemann (2021a). The dataset contains a description, the program files and the steps to reproduce. We reused and extended the Python software CoMOLA developed by Strauch et al. (2019), which is available at <https://github.com/michstrauch/CoMOLA>. We included a seeding procedure and adapted the program for the uncertainty propagation analysis. Developer: M. Hildemann, first author, see contact details at first page. Year first available: 2021. The optimization was executed on a MEGWARE cluster with 15.120 cores, 412 nodes and Intel Xeon Gold 6140 18C 2.30 GHz processors, with the freely available Python 3.7.2 software in a Linux environment with the modules NumPy 1.19.1, pandas 1.2.3 and pickle 12.1.

#### CRedit authorship contribution statement

**Moritz Hildemann:** Conceptualization, Methodology, Software.  
**Judith A. Versteegen:** Conceptualization, Writing – review & editing, Supervision, Validation.

#### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Appendix A. Objective functions

$$SR = 5 * A_F^{0.2} \quad (1)$$

where  $A_F$  = area of forest

$$HH = \sum_{i=1}^5 HH_{E_i} = \frac{E_i}{I} \quad (2)$$

where  $E$  = number of edges,  $I$  = edge intensity

$$WY = \sum_{c=1}^7 WY_{A_c} = \frac{A_{K_c}}{K_c} \quad (3)$$

where  $K_c$  = evaporation rate,  $A_{K_c}$  = area of land use with evatranspiration rate  $K_c$

$$CY = \sum_{i=1}^I CY_i = \log(P_i(1 + F_i)) \tag{4}$$

where  $P_i$  = crop production intensity,  $F$  = soil fertility.

**Appendix B. Progression of objective values in Monte Carlo simulation**

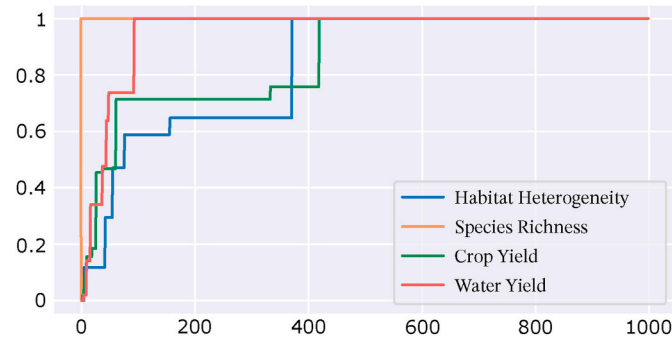


Fig. A1. Progression of best achieved objective values (normalized) in the sampling process of the uncertain land use data.

**Appendix C. Mapping between land cover and land use**

In general, land covers dominated by cropland or vegetation is assumed to be used as cropland. Forest dominated land covers are mapped to the land use forest. Shrubland and grassland dominated land cover are mapped to pasture, and artificial surfaces and bare areas as urban area. 19 land covers are mapped to 7 land uses, therefore some closes of the original confusion matrix from GlobCover (Bicheron et al., 2008) are aggregated. For example, GlobCover land cover classes 40–110 correspond to the land use forest. After the class aggregation in mapping land covers to land uses, it is not a misclassification, if e.g. land cover class 40 was classified as land cover class 100. It counts as an error, if a land use forest (land cover classes 40–100) was misclassified as urban (land cover classes 190–200) or one of the other land uses.

**Table A1**  
Mapping land cover to land use classes

GC class	GC label	Land use label	ID
11	Post-flooding or irrigated croplands	Cropland, intensity 5	5
14	Rainfed croplands	Cropland, intensity 4	4
20	Mosaic cropland (50–70%)/Vegetation (grassland, shrubland, Forest) (20–50%)	Cropland, intensity 3	3
20	Mosaic cropland (50–70%)/Vegetation (grassland, shrubland, Forest) (20–50%)	Cropland, intensity 2	2
30	Mosaic vegetation (grassland, shrubland, Forest) (50–70%)/Cropland (20–50%)	Cropland, intensity 1	1
30	Natural and semi-natural primarily terrestrial vegetation/Cultivated and managed terrestrial areas	Cropland, intensity 1	1
40	Closed to open (>15%) broadleaved evergreen and/or semi-deciduous Forest (>5 m)	Forest	6
50	Closed (>40%) broadleaved deciduous Forest (>5 m)	Forest	6
60	Open (15–40%) broadleaved deciduous Forest (>5 m)	Forest	6
70	Closed (>40%) needle-leaved evergreen Forest (>5 m)	Forest	6
90	Open (15–40%) needle-leaved deciduous or evergreen Forest (>5 m)	Forest	6
100	Closed to open (>15%) mixed broadleaved and needle-leaved Forest (>5 m)	Forest	6
110	Mosaic Forest/shrubland (50–70%)/grassland (20–50%)	Pasture	7
120	Mosaic grassland (50–70%)/Forest/shrubland (20–50%)	Pasture	7
130	Closed to open (>15%) shrubland (<5 m)	Pasture	7
140	Closed to open (>15%) grassland	Pasture	7
150	Sparse (>15%) vegetation (woody vegetation, shrubs, grassland)	Pasture	7
190	Artificial surfaces and associated areas (Urban areas >50%)	Urban area	8
200	Bare areas	Urban area	8

**References**

Ahmed, S., Shapiro, A., 2008. Solving chance-constrained stochastic programs via sampling and integer programming. In: Chen, Z.L., Raghavan, S., Gray, P., Greenberg, H.J. (Eds.), *State-of-the-Art Decision-Making Tools in the Information-Intensive Age*. INFORMS, pp. 261–269. <https://doi.org/10.1287/educ.1080.0048>.

Anderson, G., 1976. Error propagation by the Monte Carlo method in geochemical calculations. *Geochem. Cosmochim. Acta* 40, 1533–1538. [https://doi.org/10.1016/0016-7037\(76\)90092-2](https://doi.org/10.1016/0016-7037(76)90092-2).

Autuori, J., Hnaïen, F., Yalaoui, F., 2016. A mapping technique for better solution exploration: Nsga-ii adaptation. *J. Heuristics* 22, 89–123. <https://doi.org/10.1007/s10732-015-9303-4>.

Bassi, M., de Cursi, E.S., Pagnacco, E., Ellaia, R., 2018. Statistics of the pareto front in multi-objective optimization under uncertainties. *Lat. Am. J. Solid. Struct.* 15 <https://doi.org/10.1590/1679-78255018>.

Bicheron, P., Defourny, P., Brockmann, C., Schouten, L., Vancutsem, C., Huc, M., Bontemps, S., Leroy, M., Frédéric, A., Herold, M., Ranera, F., Arino, O., 2008. *Globcover: products description and validation report*.

Cao, K., 2018. Spatial optimization for sustainable land use planning. In: *Comprehensive Geographic Information Systems*. Elsevier, pp. 244–252. <https://doi.org/10.1016/B978-0-12-409548-9.09603-2>.

Corbane, C., Sabo, F., Politis, P., Syrris, V., 2021. Ghs-built-s2 r2020a: Ghs built-up grid, derived from sentinel-2 global image composite for reference year 2018 using convolutional neural networks (ghs-s2net). <https://ghsl.jrc.ec.europa.eu/download.php?ds=buS2>.

- Deb, K., Pratap, A., Agarwal, S., Meyarivan, T., 2002. A fast and elitist multiobjective genetic algorithm: Nsga-ii. *IEEE Trans. Evol. Comput.* 6, 182–197. <https://doi.org/10.1109/4235.996017>.
- Fang, S., Gertner, G., Wang, G., Anderson, A., 2006. The impact of misclassification in land use maps in the prediction of landscape dynamics. *Landsc. Ecol.* 21, 233–242. <https://doi.org/10.1007/s10980-005-1051-7>.
- Fisher, P., Comber, A., Wadsworth, R., 2006. Approaches to uncertainty in spatial data. In: Devillers, R., Jeansoulin, R. (Eds.), *Fundamentals of Spatial Data Quality*. ISTE, London, UK, pp. 43–59. <https://doi.org/10.1002/9780470612156.ch3>.
- Foody, G.M., 2002. Status of land cover classification accuracy assessment. *Rem. Sens. Environ.* 80, 185–201. [https://doi.org/10.1016/S0034-4257\(01\)00295-4](https://doi.org/10.1016/S0034-4257(01)00295-4).
- Friedrich, T., Wagner, M., 2015. Seeding the initial population of multi-objective evolutionary algorithms: a computational study. *Appl. Soft Comput.* 33, 223–230. <https://doi.org/10.1016/j.asoc.2015.04.043>.
- Guariso, G., Sangiorgio, M., 2020. Improving the performance of multiobjective genetic algorithms: an elitism-based approach. *Information* 11, 587. <https://doi.org/10.3390/info11120587>.
- Heuvelink, G.B.M., 1998. Uncertainty analysis in environmental modelling under a change of spatial scale. In: Finke, P.A., Bouma, J., Hoosbeek, M.R. (Eds.), *Soil and Water Quality at Different Scales*. Springer Netherlands, Dordrecht, pp. 255–264. [https://doi.org/10.1007/978-94-017-3021-1\\_24](https://doi.org/10.1007/978-94-017-3021-1_24).
- Hildemann, M., 2021a. Algorithm for constrained multi-objective land use allocation optimization under uncertainty. <https://doi.org/10.17632/pwkdwprv3j.3>.
- Hildemann, M., 2021b. Sampling Procedure of Land Use and Soil Fertility Map under Uncertainty. <https://doi.org/10.17632/6x6ccfc4x.3>.
- Holland, J.H., 1984. Genetic algorithms and adaptation. In: Selfridge, O.G., Rissland, E. L., Arbib, M.A. (Eds.), *Adaptive Control of Ill-Defined Systems*. Springer US, Boston, MA, pp. 317–333. [https://doi.org/10.1007/978-1-4684-8941-5\\_21](https://doi.org/10.1007/978-1-4684-8941-5_21).
- Homem-de Mello, T., Bayraksan, G., 2014. Monte Carlo sampling-based methods for stochastic optimization. *Surv. Oper. Res. Manag. Sci.* 19, 56–85. <https://doi.org/10.1016/j.sorms.2014.05.001>.
- Hunsaker, C.T., Goodchild, M.F., Friedl, M.A., Case, T.J., 2013. *Spatial Uncertainty in Ecology: Implications for Remote Sensing and GIS Applications*. Springer, New York, NY.
- Ibrahim, A., Rahnamayan, S., Martin, M.V., Deb, K., 2016. 3d-radvis: visualization of pareto front in many-objective optimization. In: 2016 IEEE Congress on Evolutionary Computation (CEC). IEEE, pp. 736–745. <https://doi.org/10.1109/CEC.2016.7743865>.
- Ismaili, S., Ahmed, D., Lhoussaine, M., 2014. Soil fertility mapping: comparison of three spatial interpolation techniques: (ordinary kriging, inverse distance weighted, and spline). *Int. J. Eng. Res. Technol.* 3. <https://www.ijert.org/soil-fertility-mapping-comparison-of-three-spatial-interpolation-techniques>.
- Klimkowicz-Pawlas, A., Ukalska-Jaruga, A., Smreczak, B., 2019. Soil quality index for agricultural areas under different levels of anthropopressure. *Int. Agrophys.* 33, 455–462. <https://doi.org/10.31545/intagr/113349>.
- Luo, J., Tian, L., Luo, L., Yi, H., Wang, F., 2017. Two-step optimization for spatial accessibility improvement: a case study of health care planning in rural China. *BioMed Res. Int.* 2017, 2094654. <https://doi.org/10.1155/2017/2094654>.
- MacArthur, R.H., Wilson, E.O., 1967. *The theory of island biogeography*. In: *Of Monographs in Population Biology*, ume 1. Princeton University Press, Princeton, N.J.
- Monserud, R.A., Leemans, R., 1992. Comparing global vegetation maps with the kappa statistic. *Ecol. Model.* 62, 275–293. [https://doi.org/10.1016/0304-3800\(92\)90003-W](https://doi.org/10.1016/0304-3800(92)90003-W).
- Senaratne, H., Gerharz, L., Pebesma, E., Schwing, A., 2012. Usability of spatio-temporal uncertainty visualisation methods. In: Gensel, J., Josselin, D., Vandenbroucke, D. (Eds.), *Bridging the Geographic Information Sciences*. Springer Berlin Heidelberg. Lecture Notes in Geoinformation and Cartography, Berlin, Heidelberg, pp. 3–23. [https://doi.org/10.1007/978-3-642-29063-3\\_1](https://doi.org/10.1007/978-3-642-29063-3_1).
- Shapiro, A., 2003. Monte Carlo sampling methods. In: *Stochastic Programming*. Elsevier, pp. 353–425. [https://doi.org/10.1016/S0927-0507\(03\)10006-0](https://doi.org/10.1016/S0927-0507(03)10006-0) volume 10 of *Handbooks in Operations Research and Management Science*.
- Strauch, M., Cord, A.F., Pätzold, C., Lautenbach, S., Kaim, A., Schweitzer, C., Seppelt, R., Volk, M., 2019. Constraints in multi-objective optimization of land use allocation – repair or penalize? *Environ. Model. Software* 118, 241–251. <https://doi.org/10.1016/j.envsoft.2019.05.003>.
- Versteegen, J.A., Jonker, J.G.G., Karssen, D., van der Hilst, F., Schmitz, O., de Jong, S. M., Faaij, A.P., 2017. How a pareto frontier complements scenario projections in land use change impact assessment. *Environ. Model. Software* 97, 287–302. <https://doi.org/10.1016/j.envsoft.2017.08.006>.
- Villa, C., Lozquez, E., Labayrade, R., 2013. Multi-objective optimization under uncertain objectives: application to engineering design problem. In: Hutchison, D., Kanade, T., Kittler, J., Kleinberg, J.M., Mattern, F., Mitchell, J.C., Naor, M., Nierstrasz, O., Pandu Rangan, C., Steffen, B., Sudan, M., Terzopoulos, D., Tygar, D., Vardi, M.Y., Weikum, G., Purshouse, R.C., Fleming, P.J., Fonseca, C.M., Greco, S., Shaw, J. (Eds.), *Evolutionary Multi-Criterion Optimization*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 796–810. [https://doi.org/10.1007/978-3-642-37140-0\\_59](https://doi.org/10.1007/978-3-642-37140-0_59) volume 7811 of *Lecture Notes in Computer Science*.
- Wackernagel, H., 1995. Ordinary kriging. In: Wackernagel, H. (Ed.), *Multivariate Geostatistics*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 74–81. [https://doi.org/10.1007/978-3-662-03098-1\\_11](https://doi.org/10.1007/978-3-662-03098-1_11).