Optimizing Remediation of Spatially Dispersed Contaminated Parcels under an Annual Budget Constraint

Floris Abrams[1,2], Lieve Sweeck[1], Johan Camps[1]
[1]Belgian Nuclear Research Centre (SCK CEN),
Mol, Belgium
e-mail: {Floris.Abrams, Lieve.Sweeck,
Johan.Camps}@sckcen.be

Dirk Cattrysse[2], Jos Van Orshoven[2]
[2]Katholieke Universiteit Leuven (KU Leuven)
Leuven, Belgium
e-mail: {Dirk.Cattrysse, Jos.Vanorshoven}@kuleuven.be

***Abstract*—In environmental disaster management, due to the large impacted area or limited availability of labor and financial resources, setting priorities of where, how and when to act are indispensable. When prioritized interventions on spatially dispersed entities are costly and technically challenging to perform, clustering of individual entities in larger homogeneous actionable units can improve feasibility and reduce cost of the remediation. In this article, a spatio-temporal clustering approach under a budget constraint is presented to determine homogenous clusters of polygons and interventions to reduce cost while still attaining an overall optimal distribution of interventions. We demonstrate the effectiveness of this clustering algorithm with a hypothetical case study of contaminated agricultural land in Belgium. Finally, we demonstrate the capabilities of the proposed cluster algorithm to provide decision makers with a multi-period action plan, reducing the cost of intervention while still prioritizing resources for the most important sites.**

Keywords-Spatio-temporal clustering; Budget constraint; Disaster management; Multi-Attribute Decision Making; MADM.

I. INTRODUCTION

This paper extends a previous paper that was originally presented at the Fourteenth International Conference on Advanced Geographic Information Systems, Applications, and Services (GEOProcessing) [1].

When dealing with large natural or man-made disasters, decision makers are confronted with setting priorities of where, how and when to act because of the limited availability of labor and financial resources. This priority setting is particularly applicable when the impact of remedial actions is costly and has long-lasting influences. For spatially distributed sites with variable characteristics, priority setting among the sites and the determination of the most adequate remedial action per site are of major importance. The United States Environmental Protection Agency (US EPA) identified the following benefits of these optimization efforts: more cost-effective expenditure, lower energy use, reduced carbon footprint, improved remedy protectiveness, improved project and site decision making, and acceleration of project and site completion [2].

Addressing the questions of where and how to act consecutively results in a nested ranking of sites and interventions per site. From those rankings, a spatio-temporal action plan can be determined.

To assist decision makers in setting such priorities, spatial Decision Support Systems (sDSS) become of importance [3]. The effectiveness of related decisions is typically conditioned by multiple and often contradicting criteria of economic, social, technical, environmental, and human health-related nature [4]. These characteristics of the decision problem make it suitable for the application of a spatially discrete Multi-Attribute Decision Making (GIS-MADM) approach [5]. 'GIS' points to the spatial aspect of the decision problem, while MADM encompasses a subset of Multi-Criteria Decision Analysis (MCDA) methods. MADM supports the decision-maker by describing and evaluating the performance of a finite number of decision alternatives with respect to multiple criteria expressed as attributes of the alternatives, representing several points of view. The MADM results in a ranking of the alternatives based on the selected criteria and their relative importance [3]. The MADM framework is often applied because it supports a structured and inclusive decision process, addressing a plurality of preferences and socio-technical dimensions that cannot always be brought to a common monetary scale [6].

This paper presents a GIS-MADM approach that provides actionable support to decision makers by proposing a coherent action plan in space and time for decontamination of the agricultural domain in a region affected by the deposition of radionuclides. It uses a spatio-temporal approach to deal with the clustering of spatially scattered polygon-based parcels, whereby a budget constraint limits the extent and/or type of interventions that can be performed in each time step, i.e., in one year. The paper elaborates on the classic region-growing principles, adapted to polygon-based data structures, and explicitly takes into account the attributes of the individual polygons to find the optimal compromise attribute for the whole cluster. Because a spatial and temporal clustering of sites and actions is likely to create "economies of scale" [7], the cost of remediation interventions will be lowered, resulting in an overall cheaper and faster remediation process.

The rest of this paper is organized as follows: Section II presents related work where MADM is used to support prioritization of resources in environmental remediation.

Section III provides an in depth explanation of the spatio-temporal cluster approach. In Section IV, the approach is illustrated with a case study for an agricultural region in Belgium, contaminated after a hypothetical accidental release of Caesium-137 from a nuclear power plant. Section V discusses the applicability of the algorithm to help improve decision making, while Section VI draws the most pertinent conclusions.

II. Related Work

The use of MADM approaches for supporting remediation on a regional scale by prioritization contaminated sites for decontamination (‘Where to act?’) was reported by several authors [8]–[10]. In addition, the support on a local scale by prioritization of the remedial technologies for a given site (‘How to act?’) was also addressed in several publications [11]–[13]. However, no reports were found, where MADM was used for simultaneously prioritizing of where and how to act decisions into a coherent spatio-temporal action plan. When both prioritizations are done separately, the procedure typically yields a geographically distributed set of priority sites as well as neighbouring sites with different interventions. Different propositions were made to improve MADM on a regional scale, to reduce the scattered priorities. For example, by incorporating a compactness measure to ensure sites were big enough to ensure a feasible intervention [14].

MADM approaches have been used with raster as well as polygon-based datasets. For this application, it was chosen to use polygon-based data because they provide a natural representation for many types of geospatial entities, such as agricultural parcels, buildings, or polluted sites. In addition, these entities form the smallest units used in real-world decision making. Therefore, it is interesting to provide actionable support to decision makers based on polygon-based representations. By addressing the problem with polygon-based data, the adaptations using compactness measures and clustering of entities are more complicated compared to raster-based datasets. Because the topology of spatially dispersed polygons is less straightforward when dealing with unlinked features [15].

Further, due to the limited availability of resources, a budget constraint limits the extent of interventions possible for each period; thus, a multiple-period action plan is required. A multi-period decision problem requires a Dynamic Multi-Attribute Decision Making (dMADM) methodology [17]–[19]. However, the majority of documented MADM applications only address a decision problem for a specific time period [16]. In contrast, dMADM determines the criteria scores and relative relevance for each time period to accurately reflect the decision variables at that time.

Some spatial DSS tools were developed for supporting decisions with similar spatio-temporal aspects. These authors included a temporal dimension in their approach to determine how a set of land use types should be distributed over space and time in order to optimize the multi-dimensional land performance of a region over a period of 30 years [20]. However, they found that their approach, which was based on integer programming (IP), resulted in land use plans that were too spatially and temporally fragmented for real-world application and recommended that a clustering strategy could be a suitable next step.

III. Proposed Method

The spatio-temporal clustering approach combines a site priority score (PPS) and an action priority score (APS), as discussed in Section A. The iterative and dynamic cluster growing algorithm is discussed in Sections B and C.

A. *Distance based priority scores*

Different implementations of MADM exist, each with their own strengths and weaknesses. We opted for a distance-based MADM, called Compromise Programming (CP), to rank the considered set of feasible alternatives [21]–[24]. CP uses the distance in the feature space to the so-called ideal point of each alternative to rank them. The feature space is constructed from independent, operational, non-redundant, and continuous attributes [25]. The criteria used vary significantly between different case studies, depending on the problem, the site characteristics and the available data. For each criterion, a weight reflecting the importance of the criterion is set by the stakeholders, preferably through a collaborative process [26]. This weight takes into account the relative importance of the criterion, where its value can be understood as a trade-off value between criteria. For this set of criteria and corresponding weights, the CP methodology determines the optimal point, a vector of performance attribute values corresponding to an alternative with the best observed performance on each criterion separately. This ideal point is mostly hypothetical, because multi-criteria decision problems involve conflicting criteria. The ideal point does however allow to determine a ranking of the alternatives based on each alternative’s distance to the ideal point, whereby the alternative that comes ‘closest’ to the ideal point is the most preferred. The definition of ‘closeness’ requires the formulation of a distance metric (1), where a larger distance equals a less optimal alternative [3]. Distances based on (1) fall within the range [0-1], with a distance of 0 being the best alternative that requires no compromise because it outperforms all other alternatives on all criteria. In contrast, a distance of 1 reflects an alternative that scores the lowest on all criteria.

$$L = \left[\sum_{i}^{n} \left[\frac{f_i^+ - f_i(x)}{f_i^+ - f_i^-} \right]^p \right]^{1/p} \qquad (1)$$

- n is the number of criteria under consideration;
- w_i is the relative importance (weight) assigned to performance attribute i;

- p is a parameter that determines the type of distance function, where 2 represents the Euclidian distance;
- f_i^+ is the optimal value for performance criterion i;
- $f_i(x)$ is the value of the i^{th} performance criterion expressed as a function of the decision variables x;
- f_i^- is the anti-ideal corresponding to the i^{th} attribute that is the "worst" value for this attribute.

To determine the optimal remediation plan for a territory of interest two important questions need to be answered. The first question is "Where are the sites situated for which intervention is most urgent?". The CP methodology returns a distance score for each polygon, representing the priority/urgency of a polygon to be intervened on. From these scores a ranking of the polygons from high priority (small distance) to low priority (large distance) can be made. For the case study in this paper, this score is referred to as Parcel Priority Score (PPS). The second question is "What is the most optimal action for each site?". Therefore, for each polygon, the feasible intervention actions need to be ranked. In our proposed approach, the ranking of the alternative interventions is similarly based on a distance score, computed by CP. In the following case study, for each alternative intervention on a specific site, the Action Priority Score (APS) is calculated. The further clustering of parcels is based on the combination of PPS and APS.

B. *Temporal dynamics in MADM*

When actions are postponed in time, the initial decision variables (criteria scores for the alternatives and criteria weights) may alter, and the decision problem needs to be redefined, resulting in a multi-period MADM. The number and extent of polygons that can be acted on in each time period depends on the budget available in each period, which is set to one year in our case study. While performing actions on the most urgent polygons first, each of the actions comes at a cost. For each intervention, the cost can be calculated based on the cost per unit of area and the size of the polygon. Interventions can be done in one period until the total cost of remediation exceeds the period's budget. When the budget is reached, the remaining polygons become candidates for the next period, where changes in the criteria scores and weights may occur and should be taken into account.

C. *Spatio-temporal clustering algorithm*

The algorithm operates in a similar fashion as a region-growing algorithm, where it consecutively checks whether one of the neighbouring polygons can be added to the cluster, taking the similarity between the priority scores of the seed polygon and the neighbouring candidate into account. The clustering algorithm is iterative and consists of two phases: The cluster initialization phase is followed by the cluster growing phase, which ends as soon as one of the stopping criteria is met. The procedure is illustrated in Figure 1 and the pseudo code is given in Figure 2.

1) *Cluster initialisation*

To optimally allocate resources, the most urgent sites should be treated first. Therefore, the seed parcel is the one with the lowest PPS (smallest distance to the ideal point).

2) *Cluster growing procedure*

After the seed parcel has been determined, the cluster-growing procedure attempts to find neighbouring parcels that can be added to the seed parcel or the growing cluster, where parcels in a cluster have the same intervention action to be performed in the same period.

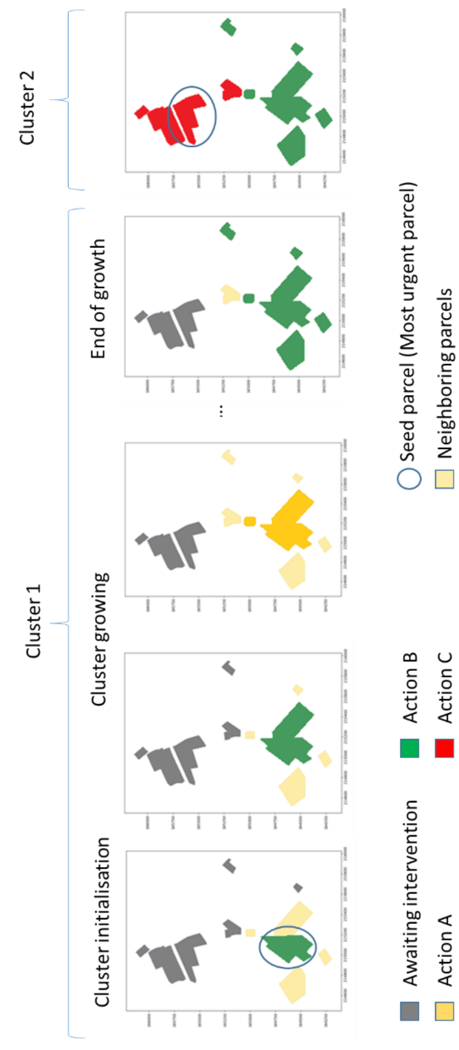

FIGURE 1. THE CLUSTER GROWING PROCEDURE APPLIED ON 12 PARCELS, CONSIDERING 3 POSSIBLE ACTIONS, RESULTING IN 2 CLUSTERS EACH WITH 1 ACTION.

```
Algorithm: Spatio-temporal cluster approach

Input: collection of polygon-based parcels, yearly budget (budget), similarity threshold (ST)

Create data structure R to store parcels
Add parcels in need for remediation to R

Create data structure S to store remediated parcels
Create data structure RC to store remediation clusters
Create data structure CC to store cluster candidates

Set t to 1
Set BT to budget
Compute PPS for each parcel

While size of R > 0 do

    Select parcel with lowest PPS from R as Seed Parcel (SP)
    Compute APS values for each remedial action
    Select all feasible remediation actions for period t
    Determine the optimal action for seed parcel as action_SP
    Determine neighboring parcels of SP as candidates
    Add candidates to CC

    IF BT - remediation cost action_SP > 0 do
        Set BT to BT - remediation cost action_SP
        Add period to SP
        Add SP to RC
        While CC > 0 do
            Compute composite score (PPS + APS) for all candidate-action combinations
            Select candidate-action_opt with lowest composite score for the whole cluster as candidate parcel (CP)
            If composite score_cp for action_opt – composite score_SP for action_SP < ST do
                If BT - remediation cost CP > 0 do
                    Add period to CP
                    Add CP to RC
                    Determine neighboring parcels of CP as new_candidates
                    Add new_candidates to CC
                Else do
                    Set t to t +1
                    Set BT to budget + BT
                    End while
                Endif
            Else do
                End while
            Endif

        End while
        add RC to S
        Remove RC from initial set R

    Else
        Set t to t +1
        Set BT to budget + BT
    Endif
End while

Output: solution set (S)
```

FIGURE 2. PSEUDO CODE OF THE SPATIO-TEMPORAL CLUSTER APPROACH, DETERMINING THE REMEDIAL TECHNIQUE AND TIMING OF THE CLUSTERS.

Adding more parcels to the cluster enlarges the cluster, therefore creating larger actionable units, which are preferred from the perspective of reducing the complexity and operational cost of the intervention. But since the parcels added to the cluster potentially have a different optimal action, it is important to find a compromise remediation action that minimizes the deviation in performance with the parcels considered individually. The cluster growing can be subdivided into three consecutive steps that are repeated until the constraints for the end of cluster growth are met.

a) *Determination of the parcel neighbours*

Compared to a raster dataset, where pixels are spatially arranged in a systematic way and neighbours are easily defined, in a data set of spatially distributed polygons, determining the neighbours is more challenging. To define neighbouring polygons, which are not necessarily sharing a border but are rather separated by irrelevant space, a technique called morphologic tessellation (MT) is used. At the core of MT is the Voronoi tessellation (VT), a method of geometric partitioning of the 2D space, where a planar set of “seed points” generates a series of polygons known as Voronoi polygons (VP). Each VP encloses the portion of the plane that is closer to its seed than to any other polygon [27]. From the partitioned space, the neighbours of a VP can be determined by examining the VPs sharing borders. An example of the portioning by VPs is given in Figure 3.

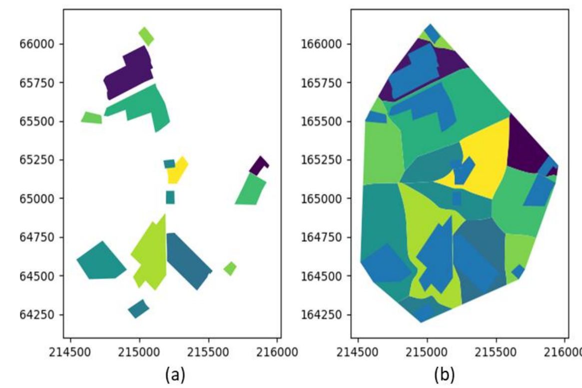

FIGURE 3. INITIAL SET OF DISTRIBUTED PARCELS (A) AND VP COMPUTED BY THE EMT, RESULTING IN A PARTITIONED COVERAGE (B).

To deal with the distributed nature of the polygons, use is made of an enclosed tessellation based on the enhanced morphological tessellation algorithm (EMT). EMT allows for setting limits to the expansion of the MT, limiting the allowed distance between polygons that can be considered to be neighbours. Furthermore, it allows for the establishment of break lines (e.g., rivers or administrative boundaries) beyond which the VPs are not permitted to trespass. The VP constructed by the EMT algorithm captures the spatial configuration of all parcels, from which the neighbouring parcels of each parcel can be determined. The EMT algorithm is accessible from an open-source Python package (http://docs.momepy.org). Fleischmann (2019, 2020) provides more information regarding the EMT methodology.

b) *Determining the optimal neighbour*

To determine the neighbouring polygon that is best suited for growing the cluster, the sum of the PPS and APS scores of each neighbour is considered. Whereby the neighbour leading to the lowest increase in the composite score of the cluster is added. From this, it follows that adding a parcel to the cluster can change the remediation action to be applied to all the parcels in the cluster. Moreover, when the best candidate is found, it is verified whether the candidate neighbour is similar enough to the seed pixel to be added. If the similarity threshold is not exceeded, the parcel is added to the cluster, and this procedure is repeated; otherwise, the end of the cluster growing phase is reached. To highlight the process of finding a compromise between all parcels on the cluster level, five iterations of the growing procedure are shown in Table 1 and Table 2. The similarity threshold applied is 0.31 for Table 1 and 0.15 for Table 2. The cells with the same color show the current parcels in the cluster, and the remedial action of the cluster is shown with a subscript on the APS. The APS values in bold show the optimal action per parcel. Table 1 illustrates that while a cluster grows iteratively, the optimal remediation action for all parcels combined within the cluster changes. In iteration III, the optimal remediation on the cluster level is the worst-performing action for the seed parcel (A) and the second-best action for parcel B. Nevertheless, from the perspective of the cluster, action 3 is the best compromise solution. In addition, Table 2 shows the impact of the similarity threshold: In iteration V, parcel E is not added to the growing cluster due to a difference larger than the similarity threshold between it and the seed parcel (parcel A). Parcel E will then be selected as the next seed parcel. The different cluster configuration (1 vs. 2) in iteration V for both tables highlights that a lower similarity threshold will result in an overall lower (better) composite score for the solution.

TABLE 1. THE GROWING PROCEDURE OF A CLUSTER FOR 5 ITERATIONS FOR A SIMILARITY THRESHOLD OF 0.3, RESULTING IN ONE CLUSTER

Parcels / Iterations	A PPS = 0.15	B PPS= 0.17	C PPS= 0.28	D PPS= 0.35	E PPS= 0.41
Iteration I (Seed parcel)	**APS_1: 0.17** APS_2: 0.22 APS_3: 0.33	**APS_1: 0.18** APS_2: 0.23 APS_3: 0.20	APS_1: 0.30 APS_2: 0.24 **APS_3: 0.11**	APS_1: 0.20 **APS_2: 0.11** APS_3: 0.17	**APS_1: 0.15** APS_2: 0.22 APS_3: 0.26
Iteration II (A+B)	**APS_1: 0.35 PPS: 0.32 Composite score$_1$: 0.67**				
Iteration III (A+B+C)	**APS_3: 0.64 PPS: 0.60 Composite score$_3$: 1.24**				
Iteration IV (A+B+C+D)	**APS_2: 0.80 PPS: 0.95 Composite score$_2$: 1.75**				
Iteration V (A+B+C+D+E)	**APS_2: 1.02 PPS: 1.36 Composite score$_2$: 2.38**				

TABLE 2. THE GROWING PROCEDURE OF A CLUSTER FOR 5 ITERATIONS FOR A SIMILARITY THRESHOLD OF 0.15, RESULTING IN TWO CLUSTERS.

Parcels / Iterations	A PPS = 0.15	B PPS= 0.17	C PPS= 0.28	D PPS= 0.35	E PPS= 0.41
Iteration I (Seed parcel)	**APS_1: 0.17** APS_2: 0.22 APS_3: 0.33	**APS_1: 0.18** APS_2: 0.23 APS_3: 0.20	APS_1: 0.30 APS_2: 0.24 **APS_3: 0.11**	APS_1: 0.20 **APS_2: 0.11** APS_3: 0.17	**APS_1: 0.15** APS_2: 0.22 APS_3: 0.26
Iteration II (A+B)	**APS_1: 0.35 PPS: 0.32 Composite score$_1$: 0.67**				
Iteration III (A+B+C)	**APS_3: 0.64 PPS: 0.60 Composite score$_3$: 1.24**				
Iteration IV (A+B+C+D)	**APS_2: 0.80 PPS: 0.95 Composite score$_2$: 1.75**				
Iteration V (A+B+C+D+E)	**APS_2: 0.80 PPS: 0.95 Composite score$_2$: 1.75**				**APS_1: 0.15 PPS= 0.41 Composite score$_1$: 0.56**

c) *Cost calculation*

Every intervention has a corresponding cost, determined by the intervention type and size of the parcel. Discounts can be taken into account when a cluster reaches a certain size (e.g., a 20% cost reduction for the whole cluster if a cluster reaches a size of 5 ha). Before a parcel is added to the cluster it is confirmed if there is still enough budget left for performing the intervention. If the budget constraint is exceeded when the parcel would be added to the cluster, the cluster growing is stopped and the remaining budget is transferred to the next year's budget.

3) *End of growth*

The end of growth phase is reached when one of the two constraints is not met.

a) *Similarity threshold*

The similarity threshold determines the variability of parcels that is allowed within the cluster. By lowering the threshold, only parcels with a similar composite score will be allowed to enter the cluster, resulting in a more homogenous cluster. As a consequence, the growth of clusters is more rapidly stopped, and the clusters tend to remain smaller, possibly not achieving a large enough size to be entitled to a discounted remedial cost. Therefore, the threshold should be chosen according to a tradeoff between the homogeneity of the clusters on the one hand and the ease and cost of implementing the remediation strategy on the other. The reasoning behind the threshold setting is that when the difference in performance between seed and candidate parcels is large, resources will be used for less urgent parcels or for suboptimal intervention. When the similarity threshold is not met, the cluster growing is stopped and a new seed polygon is found for building the next cluster.

b) *Budget constraint*

The budget constraint limits the amount of resources that can be allocated to interventions in each period. The implementation of a budget constraint in the spatial clustering algorithm ensures that cluster growth cannot lead to exceedi,g the budget for the given period. Once the budget is reached, the attributes of the remaining (unclustered) polygons are adapted to reflect their status for the new period. Next, the clustering can be started for the new period.

IV. CASE STUDY

To demonstrate the capabilities of the proposed spatio-temporal clustering model, it is applied to a case study addressing the remediation of contaminated agricultural parcels. The case study deals with a hypothetical deposition of radioactive Cesium-137 on 1257 agricultural parcels situated in the Maarkebeek Valley in Flanders, Belgium. A remediation plan must be designed for a budget of 500 000 euros per year to ensure that all parcels are remediated so that food can be produced in accordance with the legally set contamination limits. In this case study, five possible remedial interventions are considered: potassium fertilizers, shallow ploughing, deep ploughing, skim and burial ploughing and topsoil removal (Table 5).

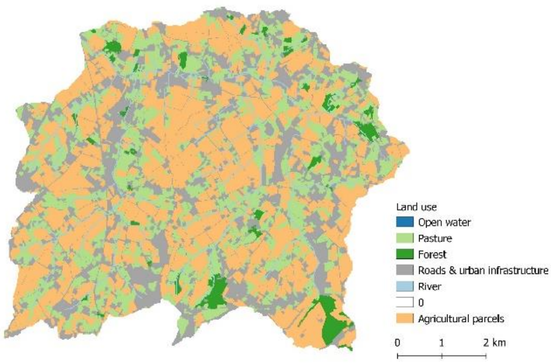

FIGURE 4. LAND USE MAP OF THE MAARKEBEEK WATERSHED IN FLANDERS.

A. *Determination of the Parcel Priority Score*

A parcel is characterized by a set of attributes such as geographic location, environmental characteristics, and agricultural practices. These attributes form the basis for the decision criteria used for determining the PPS (Table 3). The criteria for assessing the priority for remediation of sites with polluted soils were determined from a literature review [28]. Furthermore, each of the criteria was assigned a relative weight based on expert assessment of its importance. The weight is expressed by a linguistic score, which corresponds to a triangular fuzzy number (TFN). TFN are then converted to a quantitative value using the center of gravity method [29].

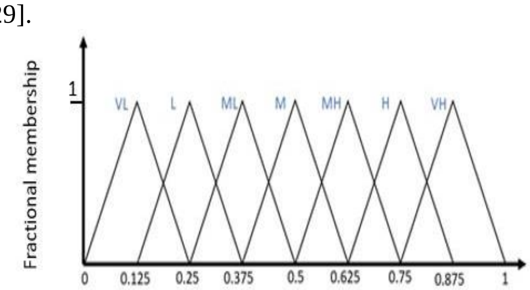

FIGURE 5. MEMBERSHIP FUNCTIONS OF THE LINGUISTIC EXPERT RATINGS USED FOR QUANTIFYING THE CRITERIA WEIGHTS, WITH ABBREVIATIONS VL : VERY LOW, L:LOW, ML: MEDIUM LOW, M: MEDIUM, MH: MEDIUM HIGH, H: HIGH AND VH: VERY HIGH.

The seven criteria and corresponding weights, shown in Table 3, are then used by the CP methodology to determine the feature distance of each parcel to the hypothetical parcel with the highest societal burden and therefore the need for remediation. In Figure 6, the CP methodology, limited to three alternatives and two criteria, is illustrated. The priorities based on this distance for each parcel are shown in Figure 7. Parcels with a low PPS are identified as the most urgent to remediate.

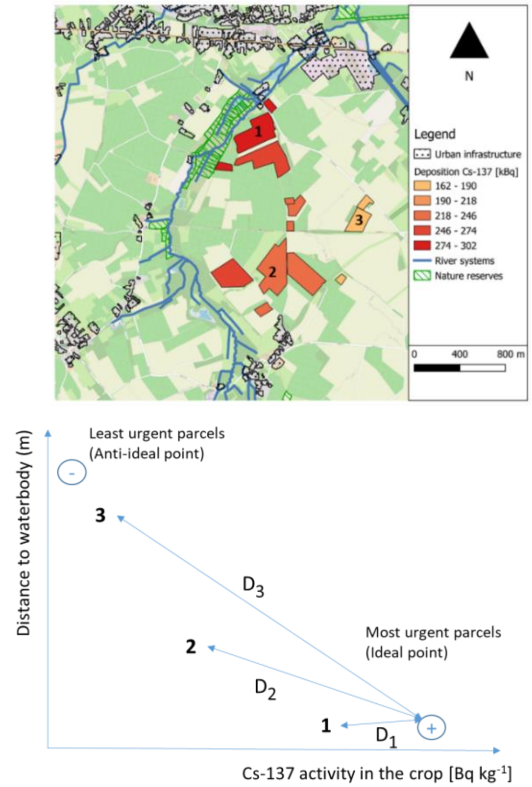

FIGURE 6. REPRESENTATION OF A 2 DIMENSIONAL COMPROMISE PROGRAMMING DISTANCE FOR 3 PARCELS (BOTTOM) AND ITS GEOGRAPHIC REPRESENTATION (TOP).

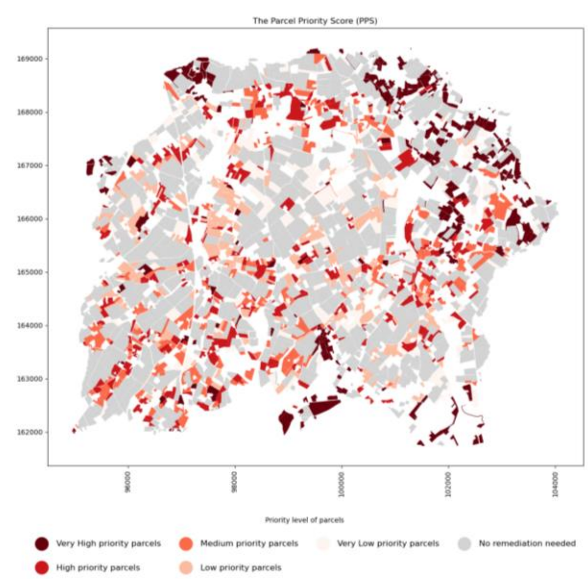

FIGURE 7. PARCEL PRIORITY SCORE (PPS) FOR THE AFFECTED AGRICULTURAL PARCELS, THE LOWER THE PPS THE MORE URGENT THE REMEDIATION.

TABLE 3. CRITERIA USED TO DETERMINE THE PARCEL PRIORITY SCORES (PPS), WITH THE CORRESPONDING WEIGHTS DETERMINED BY EXPERTS

Criterion	Description	Weight
Activity in the food products	The activity of Cs-137 found in the crop after harvest from this field [Bq/kg]	VH
Importance of the food in the local diet	The amount consumed of this product on yearly basis [kg/year]	M
Distance to the urban infrastructure	Distance to the closest urban infrastructure (houses and gardens) [meter]	H
Distance to nature reserves	Distance to the closest nature reserve [meter]	L
Distance to surface water	Distance to the closest surface water (lake/river) [meter]	M
Population density	Population density of the municipality [pp/km2]	H
Erodibility of the parcel	The erosion sensitivity of the field [scale (0 : None - 0.5 : medium - 1: very high)]	L

B. *Determination of the Action Prority Score*

For the determination of the remedial intervention among the five potential remedial actions, six criteria have been selected (Table 4). The applicability of the intervention depends on the parcel's contamination level and the crop type, because some remedial actions are unsuitable for specific agricultural crops or inadequate to reduce the contamination levels below the legal permissible levels. For example, ploughing actions are unfeasible for parcels with perennial crops. The criteria to assess remedial actions can vary largely based on the geographical region, contamination type, stakeholders, and data availability [28].

TABLE 4. CRITERIA USED TO DETERMINE THE ACTION PRIORITY SCORE (APS) OF EACH REMEDIAL INTERVENTION, THE WEIGHTS ARE BASED ON EXPERT JUDGMENT.

Criteria	Description	Weight
Feasibility	The probability that the remediation strategy is implemented successfully.	MH
Incremental Dose	Exposure dose to the workers that need to implement the remediation technique.	MH
Environmental Impact	Risk or actual impact on the living and or non-living environment due to the remediation.	M
Local Impact	Changes to the landscape/ way of life of the population.	MH
The cost of remediation	The total implementation cost of remediation minus the otherwise paid compensation to the farmer. The full remediation cycle is included from investigation to monitoring and waste treatment. [€/ha]	H
Reduction Effectivity	Reduction in activity of agricultural product (compared to doing nothing). [%]	VH

The five remedial alternatives are scored on the six criteria that produce the alternative-criterion matrix (Table 5), which is the basis for the distance calculations by the CP. More information on the determination of the criteria scores in the alternative-criterion matrix can be found in [30].

TABLE 5. ALTERNATIVE-CRITERION MATRIX FOR THE FIVE REMEDIAL ALTERNATIVES, SCORING THEM ON SIX CRITERIA.

	Feasibility	Incremental dose	Environmental impact	Local impact	Direct cost of application	Reduction Effectivity
Potassium fertilizers	H	L	H	L	66 (Yearly)	69
Shallow ploughing	VH	MH	VL	L	39 (Single time)	50
Deep ploughing	M	H	M	MH	53 (Single time)	70
Skim and burial ploughing	L	H	M	MH	95 (Single time)	87.5
Topsoil removal	L	H	VH	VH	24490 (Single time)	93.5

The incorporation of the temporal dynamics in this case study is necessary since the values of certain decision variables change through time. Because of natural attenuation, which causes the mass, toxicity, volume or concentration of contaminants in the soil or groundwater to reduce over time. This implies that the contamination decreases over time without the interference of specific remedial actions. For radioactive contaminations the reduction of the contaminant is strongly determined by the radioactive decay, the radionuclide's half-live. For a remedial action to be considered feasible, it should be able to reduce the contamination levels below the legally allowed limits. From the dynamic nature of the contamination, it follows that, after a certain period of time other remedial options can become more effective and outperform the previously selected option. Consequently, the remedial actions for each parcel should be revised to ensure they are still optimal for this time period. For this case study, the weights are not considered to change between periods.

C. *Individual per parcel solution*

For each individual parcel and for each time period, an APS score for each feasible remediation technique can be calculated. This is illustrated in Figure 8 for a cereal parcel. For this specific field, only four remedial actions are feasible, and deep ploughing is considered the most optimal since it has the lowest value. Topsoil removal is the second-most optimal remedial technology.

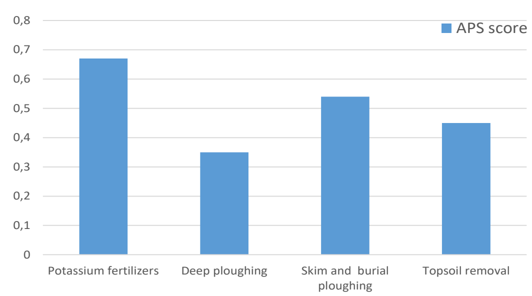

FIGURE 8. ACTION PRIORITY SCORE (APS) FOR THE DIFFERENT CANDIDATE REMEDIAL ACTIONS ON AN AGRICULTURAL PARCEL WITH CEREAL CULTIVATION.

In Figure 9, the optimal remediation technique for each parcel, based on the technique with the lowest APS, is shown.

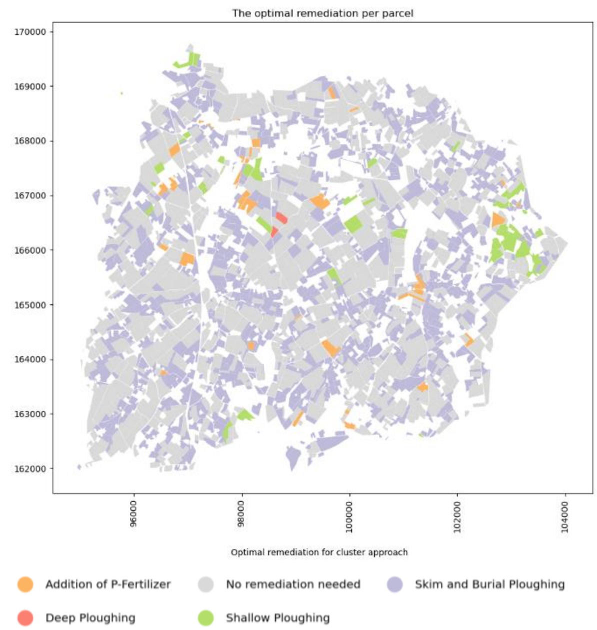

FIGURE 9. PROPOSED REMEDIATION PLAN BASED ON THE OPTIMAL REMEDIAL ACTION FOR EACH PARCEL.

D. *Spatio-temporal cluster solution for the affected region*

With the spatio-temporal cluster approach, a multi-period action plan can be designed, taking into account when and how to remediate the parcels. For the same area, the model proposes a remedial technique and timing. Both can be found in Figures 10 and 11, respectively.

The difference in remedial technologies between Figures 9 and 10 can be explained by the clustering of parcels and the changing of some of the parcel characteristics due to the delayed remediation. The remedial action "food restriction" found in Figure 10 is for agricultural parcels where, due to the physical decay process described above, the food crops can be produced with radioactivity below the permissible levels without the need for a remedial action given the time elapsed since the deposition of the radionuclides. It is clear that the model will seek optimal homogenous clusters, where the solution is optimal overall and not for each individual parcel.

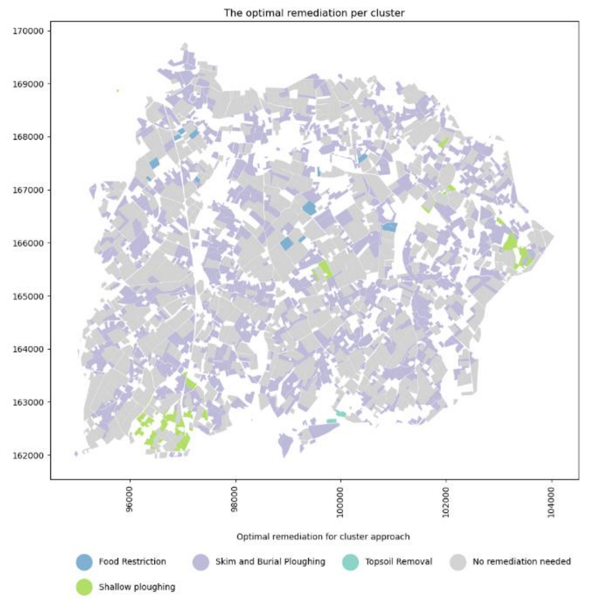

FIGURE 10: THE REMEDIAL TECHNOLOGIES PROPOSED BY THE SPATIO-TEMPORAL CLUSTERING ALGORITHM WITH A SIMILARITY THRESHOLD OF 0.025.

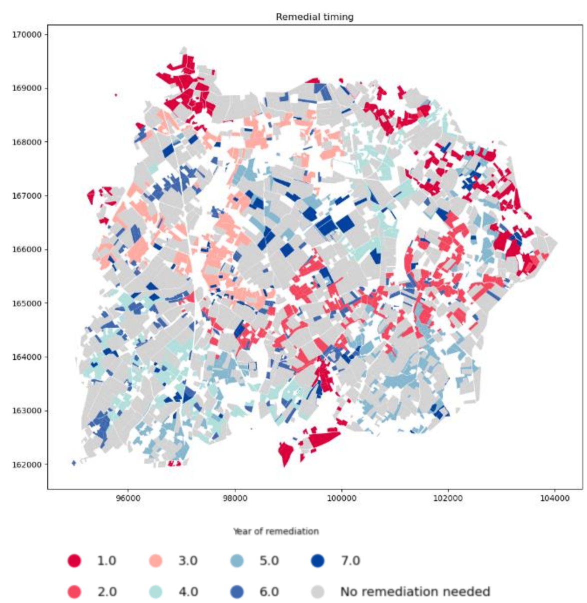

FIGURE 11. THE TIMING OF REMEDIATION PROPOSED BY THE SPATIO-TEMPORAL CLUSTERING ALGORITHM WHEN THE SIMILARITY THRESHOLD IS SET TO 0.025.

E. *Intracluster variability*

The variability of the PPS score within the cluster should be as low as possible to make sure that resources are used for the most urgent parcels. When the similarity threshold is set to 0, as in Figure 12, the clusters consists of only the seed parcel. It is clear that as cluster rank increases, so does the value of the cluster's PPS score, demonstrating the prioritization of resources for the most important parcels.

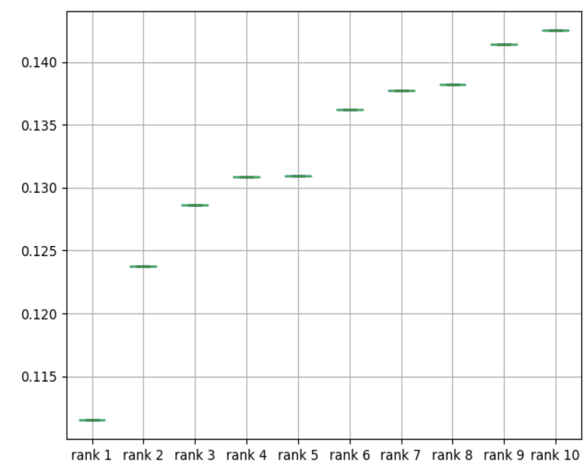

FIGURE 12. PPS SCORE FOR THE 10 HIGHEST RANKED CLUSTERS, WHICH ARE EQUAL TO THE 10 HIGHEST RANKED SEED PARCELS FOR A SIMILARITY THRESHOLD OF 0.

With an increasing similarity threshold (see Figures 13 and 14), the variability of the PPS within the cluster is allowed to increase. Furthermore, it is important to observe the increased presence of outliers due to the higher similarity threshold. This can be important when the seed parcel is the outlier, because then resources are potentially used on a less important site first.

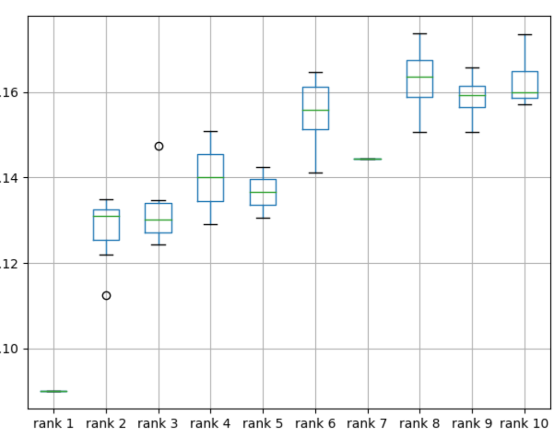

FIGURE 13. BOXPLOTS REPRESENTING THE VARIABILITY OF THE PPS SCORE WITHIN THE 10 HIGHEST RANKED CLUSTERS, WHEN THE SIMILARITY THRESHOLD IS 0.025.

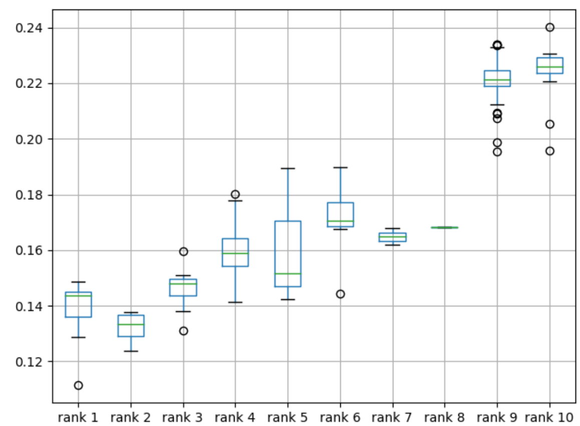

FIGURE 14. BOXPLOTS REPRESENTING THE VARIABILITY OF THE PPS SCORE WITHIN THE 10 HIGHEST RANKED CLUSTERS, WHEN THE SIMILARITY THRESHOLD IS 0.05.

It is clear that a higher similarity threshold results in more resources going to less important parcels, but on the other hand, it results in larger clusters and therefore lower operational costs. When lowering the similarity threshold for more optimal decision making, the overall cost of remediation will increase, resulting in more time needed for the remediation of the affected region. This effect can be seen in Figure 15, where the remediation will take nine years instead of seven, increasing the budget by around 1 million euros.

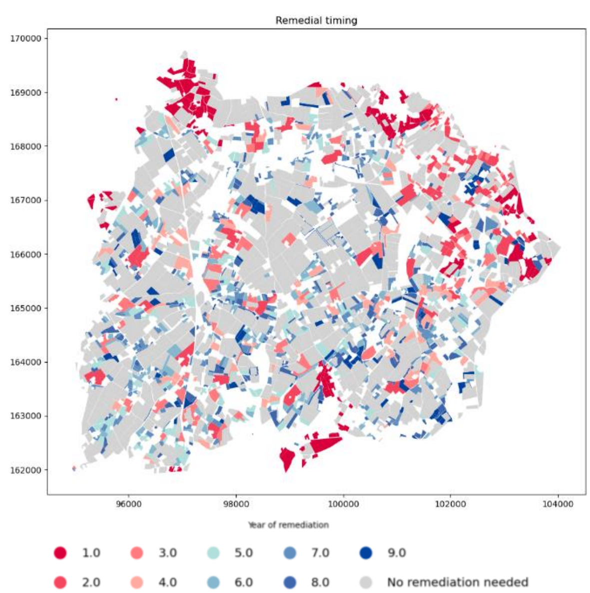

FIGURE 15. THE TIMING OF REMEDIATION PROPOSED BY THE SPATIO-TEMPORAL CLUSTERING ALGORITHM WHEN THE SIMILARITY THRESHOLD IS SET TO 0.01.

V. DISCUSSION

This case study shows the complexity of designing spatio-temporal remedial schemes. Therefore, the use of a GIS-MADM based DSS, as proposed in this paper, could help decision makers find clarity and see the impacts of certain decisions. A major benefit of these tools is the ability to do scenario analysis and uncertainty analysis. The impact of varying degrees of uncertainty in this decision context is described in [31]. Further the use of these dynamic MADM approaches allows for a shift to a more adaptive management paradigm [32].

The spatio-temporal MADM relies heavily on the PPS and APS of a parcel; therefore, the determination of these scores should be done with great care. The determination of the specific applicable criteria and weights is not only the work of experts, but it is highly suggested to take into account all stakeholders to ensure a solution supported by society is proposed [28].

For the purpose of this research, CP was used with a Euclidean distance measure, but other distance metrics are possible (e.g., Manhattan distance). Because of the use of two distance-based metrics with similar range, the composite distance score still has a physical meaning (distance to the ideal or anti-ideal situation).

Figures 12 to 14 show the effect of the increasing similarity threshold on the variability of PSS scores within clusters. A larger similarity threshold allows more variation within the cluster; therefore, less optimal clusters are formed and more deviation from the optimal per-parcel-solution is allowed. However, larger clusters give rise to lower operational costs, resulting in cheaper and faster remediation. Decision makers can decide what is the best setting for their own specific case, but a rule of thumb to determine the initial similarity threshold is half of the range of the APS values. The budget constraint limits the amount of interventions per year, therefore, a lower budget will spread the remediation over more years. This increase in remediation time could potentially change the remedial actions for parcels because of delayed remediation.

The reduced cost of remediation for larger units is the main driver for the introduction of remedial management clusters. For this specific case study, expert-based estimations for the discounts were used because empirical data for these large-scale remedial actions is not widely available. Nevertheless, they should be determined with great care and potentially adapted during the remedial process to improve the model estimations.

The proposed technical implementation of the budget constraint stops the remediation if the most optimal neighbor of the cluster with the specific remedial action exceeds the available budget, whereby the remaining budget is transferred to the next year. This transfer has a low impact, when the yearly budget exceeds largely the remediation cost of a single cluster.

When working with polygon-based datasets, topological errors, such as gaps, may occur. Relying solely on these topological relationships can have major impacts on determining the neighbours. Our EMT approach is less impacted by these errors.

Other cases could benefit from a similar approach. For example, when afforesting a large region, not all sites can be afforested at the same time because it is a very costly and labor intensive intervention. Additional, every plot has a certain suitability and urgency to be afforested. In addition, afforesting connected parcels with a similar tree composition would severely reduce the cost of planting and also improve the ecological connectivity of the landscape. Therefore, finding optimal clusters of parcels to be afforested with similar tree compositions could be facilitated with our approach. A similar approach for raster datasets was already reported by [14].

VI. Conclusions and future work

With the proposed spatio-temporal clustering approach, dispersed polygons can be clustered in space and time and be assigned the most optimal intervention type under a budget constraint. This allows decision makers to form multi-period remedial schemes to address the environmental disaster. The approach also gives decision makers the possibility to do scenario analysis and uncertainty analysis to better understand the impact of the different parameters in the model. In addition, the approach shows promise for other fields of application. More research on the impact of the similarity threshold is needed. In addition, the introduction of off-site impacts (e.g., transport and re-deposition of contaminated sediment) should be incorporated in the MADM criteria [33] to better mimic the contamination behavior. Future research should consider multiple consecutive remedial actions rather than single ones [20], to be more in line with the reality of remediation.

Acknowledgment

This work was supported by a PhD grant for Floris Abrams from the Belgian Nuclear Research Centre (SCK CEN).

References

[1] F. Abrams, L. Sweeck, J. Camps, D. Cattrysse, and J. Van Orshoven, “Spatio-Temporal Clustering of Polygon Objects and per Object Interventions Optimizing Remediation of Spatially Dispersed Contaminated Parcels Under an Annual Budget Constraint,” in *The Fourteenth International Conference on Advanced Geographic Information Systems, Applications, and Services (GEOProcessing 2022)*, 2022, pp. 1–6.

[2] U.S. EPA, “Superfund Optimization Progress Report,” 2020.

[3] J. Malczewski and C. Rinner, *Multicriteria Decision Analysis in Geographic Information Science*. 2015.

[4] R. Anderson, J. Norrman, P. E. Back, T. Söderqvist, and L. Rosén, “What’s the point? The contribution of a sustainability view in contaminated site remediation,” *Sci. Total Environ.*, vol. 630, pp. 103–116, 2018.

[5] J. Malczewski, *GIS and Multicriteria Decision Analysis*. 1999.

[6] C. D. Gamper and C. Turcanu, “Multi-criteria analysis: A tool for going beyond monetization?,” *Tools Policy Formul. Actors, Capacit. Venues Eff.*, pp. 121–141, 2015.

[7] J. Kingscott and R. J. Weisman, “Cost Evaluation for Selected Remediation Technologies,” *Remediat. J.*, vol. 12, no. 2, pp. 99–116, Mar. 2002.

[8] S. Polat, A. Aksoy, and K. Unlu, “A Fuzzy Rule Based Remedial Priority Ranking System for Contaminated Sites,” *GROUNDWATER*, vol. 53, no. 2, pp. 317–327, 2015.

[9] K. Zhang, G. Achari, and Y. Pei, “Incorporating linguistic, probabilistic, and possibilistic information in a risk-based approach for ranking contaminated sites,” *Integr. Environ. Assess. Manag.*, vol. 6, no. 4, pp. 711–724, 2010.

[10] Y. Lin, J. Hoover, D. Beene, E. Erdei, and Z. Liu, “Environmental risk mapping of potential abandoned uranium mine contamination on the Navajo Nation, USA, using a GIS-based multi-criteria decision analysis approach,” *Environ. Sci. Pollut. Res.*, vol. 27, no. 24, pp. 30542–30557, 2020.

[11] M. A. B. Promentilla, T. Furuichi, K. Ishii, and N. Tanikawa, “Evaluation of remedial countermeasures using the analytic network process,” *Waste Manag.*, vol. 26, no. 12, pp. 1410–1421, 2006.

[12] I. Linkov, A. Varghese, S. Jamil, T. P. Seager, G. Kiker, and T. Bridges, “Multi-criteria decision analysis: A framework for structuring remedial decisions at contaminated sites,” in *Comparative Risk Assessment and Environmental Decision Making*, 2004, vol. 38, pp. 15–54.

[13] L. Rosen *et al.*, “SCORE: A novel multi-criteria decision analysis approach to assessing the sustainability of contaminated land remediation,” *Sci. Total Environ.*, vol. 511, pp. 621–638, Apr. 2015.

[14] P. Vanegas, D. Cattrysse, A. Wijffels, and J. Van Orshoven, “Finding sites meeting compactness and on- and off-site suitability criteria in raster maps,” in *2nd International Conference on Advanced Geographic Information Systems, Applications, and Services, GEOProcessing 2010*, 2010, pp. 15–20.

[15] M. Fleischmann, “momepy: Urban Morphology Measuring Toolkit,” *J. Open Source Softw.*, vol. 4, no. 43, p. 1807, 2019.

[16] Z. Xu, “On multi-period multi-attribute decision making,” *Knowledge-Based Syst.*, vol. 21, no. 2, pp. 164–171, 2008.

[17] Y. Chen and B. Li, “Dynamic multi-attribute decision making model based on triangular intuitionistic fuzzy numbers,” *Sci. Iran.*, vol. 18, no. 2 B, pp. 268–274, 2011.

[18] M. Karatas, “Multiattribute Decision Making Using

Multiperiod Probabilistic Weighted Fuzzy Axiomatic Design," *Syst. Eng.*, vol. 20, no. 4, pp. 318–334, 2017.

[19] Q. Dong and Y. Guo, "Multiperiod multiattribute decision-making method based on trend incentive coefficient," *Int. Trans. Oper. Res.*, vol. 20, no. 1, pp. 141–152, 2013.

[20] R. Estrella, D. Cattrysse, and J. Van Orshoven, "An Integer Programming Model to Determine Land Use Trajectories for Optimizing Regionally Integrated Ecosystem Services Delivery," pp. 1–26, 2016.

[21] J. Tian, Z. Huo, F. Ma, X. Gao, and Y. Wu, "Application and Selection of Remediation Technology for OCPs-Contaminated Sites by Decision-Making Methods," *Int. J. Environ. Res. Public Health*, vol. 16, no. 11, Jun. 2019.

[22] C. A. Salt and M. C. Dunsmore, "Development of a spatial decision support system for post-emergency management of radioactively contaminated land," *J. Environ. Manage.*, vol. 58, no. 3, pp. 169–178, 2000.

[23] A. L. Yang, G. H. Huang, X. S. Qin, and Y. R. Fan, "Evaluation of remedial options for a benzene-contaminated site through a simulation-based fuzzy-MCDA approach," *J. Hazard. Mater.*, vol. 213, pp. 421–433, Apr. 2012.

[24] L. Bai *et al.*, "TOPSIS-Based Screening Method of Soil Remediation Technology for Contaminated Sites and its Application," *Soil Sediment Contam.*, vol. 24, no. 4, pp. 386–397, 2015.

[25] V. Belton and T. J. Stewart, *Multiple criteria decision analysis: An integrated approach*. 2002.

[26] N. H. Zardari, A. Kamal, S. Sharif Monirussaman, and Y. bin Zulkifli, *Weighting Methods and their Effects on Multi-Criteria Decision Making Model Outcomes in Water Resources Management*. 2015.

[27] M. Fleischmann, A. Feliciotti, O. Romice, and S. Porta, "Morphological tessellation as a way of partitioning space: Improving consistency in urban morphology at the plot scale," *Comput. Environ. Urban Syst.*, vol. 80, no. May 2019, p. 101441, 2020.

[28] F. Abrams, L. Sweeck, L. Hendrickx, C. Turcanuc, R. Estrella, and J. Van Orshoven, "Multi-criteria Decision Analysis to Support the Remediation of Polluted Soils: a Review," *J. Environ. Plan. Manag.*, 2022.

[29] W. J. Wang and L. Luoh, "Simple computation for the defuzzifications of center of sum and center of gravity," *J. Intell. Fuzzy Syst.*, 2000.

[30] L. Hendrickx, F. Abrams, L. Sweeck, and J. Van Orshoven, "Supporting Remediation of Agricultural Land Contaminated With Cesium-137: A Multi-criteria Decision Workflow Accounting for Uncertainties," KULeuven, 2021.

[31] F. Abrams, L. Hendrickx, L. Sweeck, J. Camps, D. Cattrysse, and J. Van Orshoven, "Accounting for Uncertainty and Disagreement in Multi-criteria Decision Making Using Triangular Fuzzy Numbers and Monte Carlo Simulation: A Case Study About Selecting Measures for Remediation of Agricultural Land After Radioactive Contamination," in *Real Life Applications of Multiple Criteria Decision Making Techniques in Fuzzy Domain*, Springer Nature, 2022.

[32] I. Linkov, F. K. Satterstrom, G. Kiker, C. Batchelor, T. Bridges, and E. Ferguson, "From comparative risk assessment to multi-criteria decision analysis and adaptive management: Recent developments and applications," *Environ. Int.*, vol. 32, no. 8, pp. 1072–1093, 2006.

[33] R. Estrella, P. Vanegas, D. Cattrysse, and J. Van Orshoven, "Trading off Accuracy and Computational Efficiency of an Afforestation Site Location Method for Minimizing Sediment Yield in a River Catchment," 2014, no. c, pp. 94–100.