The influence of uncertainties on optimization of vaccinations on a network of animal movements

Noname manuscript No.
(will be inserted by the editor)

The Influence of Uncertainties on Optimization of Vaccinations on a Network of Animal Movements

Krzysztof Michalak* · **Mario Giacobini**

Received: date / Accepted: date

Abstract In this article multiobjective optimization of vaccinations is studied using graph-based modelling and simulations of the spreading of the disease. Real-life dataset of animal movements between farms and pastures in the Piedmont region of Italy is used, from which a dynamic network of contacts is reconstructed. Evolutionary multiobjective optimization of vaccinations is compared with vaccination strategies based on degrees or strengths of graph nodes, number of animals in the farms as well as with the ring vaccination strategy. In the article the influence of uncertainties represented by the lack of knowledge of initial disease cases and the change of the contacts network by a rewiring process on the vaccination optimization is studied.

Results of experiments show that evolutionary optimization of vaccinations can outperform vaccination strategies when enough information is provided. When many disease cases remain unknown or when the changes in the contacts network are large, the performance of the optimization algorithm is adversely affected. Obtained results motivate further research on modelling changes in animal movement patterns, as well as hybrid methods combining evolutionary optimization with vaccination strategies.

Keywords Disease Prevention · Epidemics Control · DPEC · Graph-based problems · Combinatorial optimization · Multiobjective optimization

K. Michalak
Department of Information Technologies, Wroclaw University of Economics, Wroclaw, Poland
E-mail: krzysztof.michalak@ue.wroc.pl

M. Giacobini
Department of Veterinary Science, University of Torino, Torino, Italy
E-mail: mario.giacobini@unito.it

1 Introduction

The optimization problem tackled in this article is optimization of a vaccination scheme intended to stop an epidemic spreading among animal farms through animal movements. In this section key concepts are introduced, such as multiobjective metaheuristic optimization and graph-based epidemic modelling. This section concludes by presenting highlights and structure of the article.

1.1 Multiobjective metaheuristic optimization

Metaheuristic optimization methods are commonly used for solving complex optimization problems with large search spaces and high computational complexity. Application domains range from engineering applications (Biswas and Pal (2019)), through transportation optimization (Barma et al. (2019); Roy et al. (2019)) to project scheduling (Pellerin et al. (2020)). Population-based metaheuristics, such as evolutionary algorithms are often used for solving multiobjective optimization problems (Talbi et al. (2012); Zavala et al. (2014)). Without the loss of generality a multiobjective optimization problem (MOP) can be formalized in the following way:

$$\begin{aligned} &\text{minimize } F(x) = (f_1(x), \ldots, f_m(x)) \\ &\text{subject to } x \in \Omega, \end{aligned} \tag{1}$$

where:
Ω - the decision space,
m - the number of objectives.

Because in MOPs conflicting objectives very often exist it is usually not possible to find one, the best,

solution. Therefore, a concept of Pareto dominance was introduced (Deb (2001); Miettinen (1999)).

Let $F(x_1)$ and $F(x_2)$ be the objectives of two solutions $x_1, x_2 \in \Omega$:

$$\begin{aligned} F(x_1) &= (f_1(x_1), \ldots, f_m(x_1)) \\ F(x_2) &= (f_1(x_2), \ldots, f_m(x_2)) \end{aligned} \tag{2}$$

We say that x_1 *Pareto-dominates* x_2 ($x_1 \succ x_2$) iff:

$$\begin{aligned} &\forall i \in \{1, \ldots, m\} : f_i(x_1) \leq f_i(x_2) \\ &\exists i \in \{1, \ldots, m\} : f_i(x_1) < f_i(x_2) \end{aligned} \tag{3}$$

A solution x is said to be *nondominated* (*Pareto optimal*) iff:

$$\neg \exists x' \in \Omega : x' \succ x. \tag{4}$$

Typically, an algorithm solving a multiobjective optimization problem searches for a good approximation of the *Pareto set* which consists of nondominated solutions in the decision space Ω. Vectors of objectives of solutions in the Pareto set form the *Pareto front* in the objective space $\mathbb{R}^m$. As mentioned before, in MOPs it is often impossible to select a better one from two solutions because one solution can be better with respect to one objective and the other solution can be better with respect to some other objective. On the other hand, optimization methods need to select preferred solutions, for example in order to perform mating pool selection in evolutionary algorithms, or solution acceptance in local search methods. In Pareto-based methods, such as the Non-dominated Sorting Genetic Algorithm, NSGA-II (Deb et al. (2002)) and the Strength Pareto Evolutionary Algorithm, SPEA2 (Zitzler et al. (2002a)) the Pareto dominance relation (equation (3)) is used do decide which solutions to select. In decomposition-based methods, such as the Multiobjective Evolutionary Algorithm Based on Decomposition, MOEA/D (Li and Zhang (2009); Zhang and Li (2007)) the multiobjective optimization problem is decomposed to a set of single-objective optimization subproblems. In these subproblems the objectives $f_1, \ldots, f_m$ are combined to a single value, for example using weights assigned to each of the objectives. Through such scalarization a single objective is obtained, which can be used for selection in evolutionary algorithms, or for solution acceptance in local search methods.

1.2 Graph-based epidemic modelling

Numerous real-life systems, such as computer networks, social relationships and business contacts, can be studied using a graph model in which nodes represent entities, for example computers, people, and businesses, and edges represent network connections, personal contacts, etc. In networked systems various threats may emerge and spread, such as computer viruses, epidemics, and waves of bankruptcies. Obviously, when the system is affected by a threat, various actions are taken to limit the damage. Such scenarios naturally give rise to optimization problems in which, for example, the cost of protective actions has to be minimized while maximization of protection effectiveness is expected. In the area of combinatorial optimization the Firefighter Problem (FFP) (Hartnell (1995)) is often studied in which fire spreads from node to node on a graph in discrete time steps and the goal of the optimization is to decide which nodes to protect in order to save as large portion of the graph as possible. Various metaheuristic methods were used to solve this problem, such as Ant-Colony Optimization (Blum et al. (2014)), Evolutionary Algorithms (Michalak (2014a)), Estimation of Distribution Algorithms (Lipinski (2017)) and Variable Neighbourhood Search (Hu et al. (2015)). Apart from the classical single-objective, deterministic FFP, different variants were studied, such as the multiobjective (Michalak (2019)) and non-deterministic version (Michalak and Knowles (2016)).

Vaccination optimization (Parousis-Orthodoxou and Vlachos (2014)) is an example of a real-life optimization problem in which a spreading threat has to be stopped. This problem can be solved as an optimal control problem (Witbooi et al. (2015)) using compartmental models (Brauer (2008)), such as the Susceptible-Vaccinated-Infected-Recovered, SVIR model (Tornatore et al. (2014)) in which four classes (compartments) are used, instead of individuals, to represent the population. In the case of the SVIR model these classes contain all the susceptible ('S'), vaccinated ('V'), infected ('I') and recovered ('R') individuals. In more complex, hybrid compartmental models (Yu et al. (2016)), compartments can represent more properties of the individuals than just the state (for example geographical location). Another approach to vaccination optimization is to use individual-based modelling (Grimm and Railsback (2005)) which can be combined with metaheuristic algorithms in order to find optimal vaccination schemes (da Cruz et al. (2017)). In this approach graph-based representation is often used, with graph nodes representing various entities (people, animal farms, cities, etc.) and edges representing contacts which allow the pathogen to spread. In order to evaluate a given solution x to the tackled optimization problem, simulations of the spreading of the disease are performed (Juan et al. (2015)). Counter-epidemic actions encoded by the solution x are used to limit the spreading, and solution

evaluation usually involves counting infected individuals, and calculating costs of counter-epidemic actions, such as vaccinations. An obvious advantage of this approach is that complex scenarios can be evaluated, taking into account information regarding individuals or contacts much more detailed than that represented by a compartmental model. Also, there exist studies showing the superiority of individual-based optimization of vaccine allocation strategies in comparison to deriving such strategies using the compartmental model (Dalgıç et al. (2017)). On the other hand, simulations are usually lengthy, necessitating the use of considerable computational resources.

As mentioned above, the graph-based approach to epidemic modelling has the advantage in that it allows incorporating detailed information about individuals and/or contacts. For example, in the case of animal diseases, data describing animal movements are often available allowing detailed modelling of the spreading of the pathogen (Dubé et al. (2009)). In this article a real-life dataset of animal movements between farms and pastures in the Piedmont region of Italy is used to reconstruct a dynamic graph of contacts over which the disease spreads. Numerous works stress the importance of dynamical modelling of the contacts network (Bajardi et al. (2011)), because temporal correlations between contacts cannot be adequately reflected using weighted static graphs, leading in some cases to a greatly underestimated final epidemic size (Vernon and Keeling (2009)). A problem that arises in epidemics control on the contacts network using movements data is that prediction of epidemic spreading requires knowledge of future movements, while, naturally, only past movements are known.

In real-life epidemic modelling various kinds of uncertainties are encountered that influence the accuracy of predictions produced by computational methods. For example, parameters of compartmental models cannot be determined with certainty (Danila et al. (2014)). Individual-based methods are also affected by uncertainties, such as limited knowledge of initial outbreaks (Bozzette et al. (2003)) and difficult to predict outcomes of various preventive actions (Li et al. (2017)).

1.3 Highlights of this article

Key points summarizing the contents of this article are:

- Vaccination optimization is studied using **simulations of the spreading of the disease on the graph of animal movements**.
- A **real-life dataset** of animal movements between farms and pastures in the Piedmont region of Italy is used, from which a dynamic network of contacts is reconstructed. Note, that it is assumed that animal movements cannot be stopped, because they help satisfy vital business needs of the farms, such as animal trading and moving animals to pastures for grazing. Therefore, the article focuses on vaccinations as a measure to stop the epidemic without disrupting animal movements.
- **Evolutionary multiobjective optimization of vaccinations** is studied, with three optimization algorithms compared to each other and with typical vaccination strategies, for example based on degrees or strengths of graph nodes.
- To improve the working of the evolutionary algorithm **a local search procedure dedicated to the solved problem** is used in this article and shown to achieve a statistically significant improvement of the optimization results with respect to algorithms without this local search.
- **Two kinds of uncertainties** are considered: limited knowledge of initial disease cases, and differences between known past animal movements used for simulations and unknown future movements driving the real-life epidemic. In this article **a method of modelling and parameterizing changes in the dynamic network of contacts** over which the disease spreads is proposed.
- The analysis of the influence of uncertainties shows both **strengths and limitations of evolutionary optimization as well as vaccination strategies**. This knowledge can be **useful for making decisions** and determining situations when general vaccination strategies are better suited, and when optimization of the vaccination scheme can be performed.

The article is structured as follows. Section 2 describes the epidemic scenario considered in this article. Section 3 defines the optimization problem. Section 4 describes the experiments. Section 5 discusses the results and section 6 concludes the article.

2 Epidemic Scenario

This section describes the details of the epidemic modelling: the movements dataset and contacts network; epidemic parameters, the initial state of the system, and disease spreading dynamic; and uncertainties.

2.1 The movements dataset and contacts network

The epidemic studied in this article is modelled on a graph in which nodes are farms and pastures in the

Piedmont region of Italy, and edges represent movements of animals between these locations in the year 2017. There are 9313 farms housing 560812 animals, and 573 pastures. Movements can occur between farms and other farms, as well as pastures, and for each movement the date of the movement and the number of animals are known. The time period considered in this article is from 2017.01.01 to 2017.12.31. The number of farm-to-farm movements is 36293 involving 118146 animals and 50289 animals are moved between farms and pastures in 1998 farm-to-pasture movements and 2037 pasture-to-farm movements. In order to reduce noise, a practice proposed by other authors was followed (Bajardi et al. (2011); Rautureau et al. (2011)) and movements were aggregated in 7-day periods thereby forming 53 time periods.

2.2 Epidemic parameters and disease spreading dynamic

The epidemic is modelled on a graph $G = \langle V, E \rangle$, with $|V| = 9886$ nodes (farms and pastures) in $N_t = 53$ time steps, following the Susceptible-Vaccinated-Infected-Recovered, SVIR model (Tornatore et al. (2014)). For each node in the graph the number of animals in each of the 'S', 'V', 'I', and 'R' states is recorded. The size of the initial outbreak was set to $\alpha_{inf} = 1\%$ of the farms, so 93 farms are randomly selected with uniform probability and all the animals in these farms are set to the 'I' state. Within nodes the spreading of the disease is simulated assuming a full network of contacts between animals in that node. Disease transmission probability per a time step is assumed to be $\beta_s = 0.5$ for susceptible animals and $\beta_v = 0.01$ for vaccinated animals. Spreading of the disease between nodes is only possible when animals are moved. When a movement occurs, a required number of animals is selected from the source node with the states selected in proportion to those present in that node when the movement occurs. For example if 20 animals are to be moved from a node that contains 40 'S', 20 'V', 10 'I' and 30 'R' animals (100 animals in total) the movement will contain 8 'S', 4 'V', 2 'I' and 6 'R' animals. Counters in the source and destination nodes are updated accordingly. Figure 1 presents spreading of the disease with $\alpha_{inf} = 0.01$, $\beta_s = 0.5$ and $\beta_v = 0.01$ during one year (53 time periods). Node color is assigned proportionally to the fraction of animals that are infected in each node with blue representing a node with no infected animals and red representing a node with all infected animals. Node radius is proportional to the square root of the number of animals.

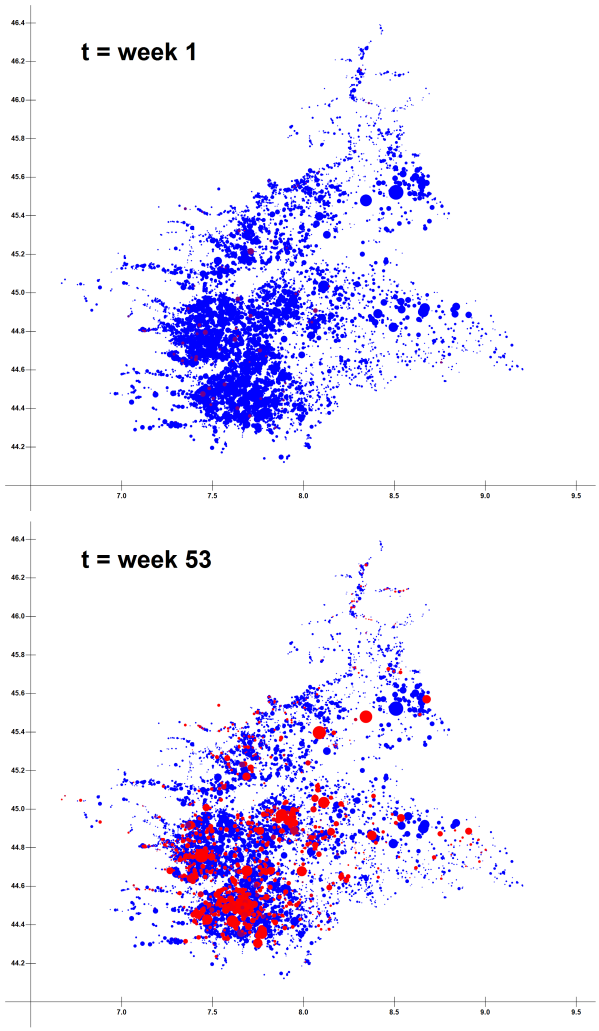


Fig. 1 Spreading of the disease with $\alpha_{inf} = 0.01$, $\beta_s = 0.5$ and $\beta_v = 0.01$ during one year (53 time periods). Axes show geographic coordinates.

2.3 Uncertainties

In this article two kinds of uncertainties are taken into account. First, it is assumed that not all initially infected farms are known, but only a fraction α_{known}. To compensate for this lack of knowledge, additional, randomly selected farms can be treated as infected. The actual initial state of the epidemic is a binary vector $S_0 \in \{0,1\}^{N_{farms}}$ in which N_{inf} randomly selected farms are infected (elements of S_0 set to 1). However, vaccination optimization or selection of farms to vaccinate using the strategies is not performed using the actual initial state S_0, but rather using a different initial state S_0' in which only $\alpha_{known} \cdot N_{inf}$ randomly selected farms from S_0 are known to be infected. A multiplication factor R_a is used to control the addition of randomly selected infected farms to S_0'. For $R_a = 0$ no

artificial infected farms are added, for $R_a = 1$ there are as many additional as the known infected farms, etc. The motivation for studying this kind of uncertainties is that in real life not all infections have to be initially detected and there may exist additional initial cases of the disease at locations which are not easy to determine.

The second kind of uncertainties regards the movements of animals. As mentioned before, a database of animal movements is available, but these are, obviously, movements recorded in the past. When an epidemic breaks out, the spreading of the disease depends on future movements which may, or may not, resemble those from the past. In order to study the influence of differences in animal movements on the vaccination schemes the network of movements is rewired in the following manner. Movements (edges of the graph G) are rewired by taking pairs of movements of the same type (farm-to-farm, farm-to-pasture, or pasture-to-farm), keeping the number of moved animals unchanged, but swapping the destinations. So, if two edges $\langle v_1, v_2\rangle$ and $\langle v_3, v_4\rangle$ are selected, the rewiring removes these two edges and inserts two new ones: $\langle v_1, v_4\rangle$ and $\langle v_3, v_2\rangle$. The parameter α_{rewire} is used to determine how many movements will be changed. The number of rewired pairs of movements is $\alpha_{rewire} \cdot |E|/2$, where $|E|$ is the number of edges in the graph G. In the later part of the article we will denote the contacts network obtained using the rewiring mechanism as $G'(\alpha_{rewire})$. The motivation for rewiring movements in this manner is that animal movements are performed to satisfy certain business needs, such as selling animals to another farm or moving animals to pastures for grazing. Here, it is assumed that if a farm v_1 sells animals to another farm v_2 the same number animals will be sold in the modified set of movements, but possibly to a different farm. Note, that, while motivated by realistic assumptions, this rewiring mechanism does not change the degrees nor strengths of the nodes and thus may favor degree- and strength-based vaccination strategies.

3 Optimization Problem

This section defines the vaccination optimization problem tackled in this article. We assume that vaccinations are performed for entire farms, that is, when a farm is vaccinated, all animals change their state to 'V'. Vaccinations are performed in a reactive manner, that is, the outbreak of the disease happens first, and then the vaccine is administered. This assumption is motivated by the fact that in the case of some animal diseases, such as the Foot-and-Mouth Disease (FMD), preemptive mass vaccinations are prohibited in certain countries. For example, in the European Union only reactive FMD vaccinations are allowed. If animals at a given farm are initially infected it is not possible to vaccinate such farm, because vaccinations would be ineffective. Also, it is not possible to vaccinate animals in pastures, only in farms. Therefore, the search space in this optimization problem is $\Omega = \{0,1\}^{N_{farms}}$ where: $N_{farms} = |V_f|$ and V_f denotes the set of those nodes in graph G which are farms (therefore, $N_{farms} = 9313$ in this article). Farms for which the corresponding element in a solution $x \in \Omega$ equals 1 are vaccinated.

The optimization problem studied in this article is a biobjective one, because it is desirable to reduce vaccination costs and the number of infected animals at the same time. For a given solution $x \in \Omega$ two objectives $f_1(x)$ and $f_2(x)$ are calculated, both of which are to be minimized. Objective f_1 is the number of animals in the vaccinated farms. It is calculated by simply adding the number of animals for those farms for which a corresponding element in x is set to one, excluding those which are initially infected:

$$f_1(x) = \sum_{i=1}^{N_{farms}} N_{anim}[i]x[i](1 - S_0[i]) \ , \tag{5}$$

where:

$N_{anim}[i]$ - the number of animals in the i-th farm,

S_0 - the initial state in which the elements set to 1 correspond to initially infected farms.

Objective f_2 is the number of animals infected in a simulation starting with farms infected according to the initial state S_0 and with vaccinations performed according to the solution x. After vaccinations are performed, the spreading of the disease is simulated using the pathogen spreading mechanism described in section 2. Since the spreading of the disease is non-deterministic, $N_{sim} > 1$ simulations are performed for a given S_0 and x in order to reduce random variation. The value of the f_2 objective is calculated by averaging the results of these N_{sim} simulations.

Note, that because of uncertainties, evaluations performed when optimizing vaccinations use different data than those performed when solutions found using different methods are compared. Evaluations performed by the optimization algorithm start from the initial state S'_0 in which only a fraction α_{known} of farms infected in the true initial state S_0 is known, but additional number $R_a \cdot \alpha_{known} \cdot N_{inf}$ of randomly selected farms may be assumed to be infected. Also, these evaluations use the rewired contact network $G'(\alpha_{rewire})$, instead of the original graph G, for simulating the spreading of the disease. When comparing solutions generated by various methods the true initial state S_0 is used and the

spreading of the disease is simulated on the original graph G. This way the optimization algorithm works without the full knowledge of the outbreak and the movements, but the solutions it finds are evaluated as if applied to a real scenario.

4 Experiments

The experiments described in this article were aimed at the following goals:

- Comparing three multiobjective optimization algorithms: MOEA/D, NSGA-II and SPEA2 and selecting the best one for the vaccination optimization problem.
- Testing the local search procedure (Algorithm 2), dedicated to the vaccination optimization problem.
- Determining the effectiveness of vaccination optimization, in particular comparing solutions generated by the optimization algorithm and those produced using various vaccination strategies.
- Understanding the impact of uncertainties on the optimization process as well as the effectiveness of the strategies. A related goal was to study the possibility of using randomly selected 'infected' farms in order to compensate for the missing information about really infected farms.

4.1 Vaccination Strategies

In epidemiology a number of strategies are known which aim to detect good targets for vaccination, based on the attributes of individuals (e.g. age, occupation) or properties of the contacts network. Among the latter, vaccinating high-degree (or high-strength) nodes is especially popular, because estimating the node degree and strength using local information is relatively easy and is computationally cheap. In this article several *order-based strategies* were used in which a selected criterion determines which farms are preferred for vaccination. These strategies do not determine how many nodes or animals to vaccinate, but only order the farms in a certain way and so they do not produce any predetermined value for the objective f_1 (the number of vaccinated animals). In order to make these strategies comparable to the results of a multiobjective optimization algorithm a set of solutions was generated using each strategy by iteratively setting f_1 to 10000, 20000, ..., 560000 and using the strategy to determine which farms to vaccinate (Algorithm 1). Note, that, similarly as for optimization, the farms that were known to be infected at the time of vaccinations were omitted.

Algorithm 1: Generating solutions using order-based strategies.

Inputs:

N_{farms}	-	the number of farms
N_{anim}	-	a vector containing the number of animals for each farm
$Attr$	-	farm attributes (e.g. node degrees)
S'_0	-	the initial (known) state of the epidemic

Output:

$P \subset \Omega$	-	a set of generated solutions

```
// Farms preferred by the strategy are placed first
R := {1, ..., N_farms}
R := StrategySort(R, Attr)

// Vaccinate an increasing number of animals
P := ∅
for Max_vacc := 10000, 20000, ..., 560000 do
    // A vector of 9313 zeros
    x := 0⃗

    N_vacc := 0
    for i := 1, ..., N_farms do
        // Select the i-th best farm according
        // to the strategy
        j := R[i]

        // Vaccinate if not initially infected and not
        // exceeding the maximum number of animals
        if S'_0[j] = 0 and N_vacc + N_anim[j] ≤ Max_vacc
         then
            x[j] := 1
            N_vacc := N_vacc + N_anim[j]

    P := P ∪ {x}
```

The selection of farms for vaccination using the strategies was always performed using the known initial state S'_0, not the true one S_0, and similarly graph-based attributes were calculated using the rewired graph $G'(\alpha_{rewire})$. The *StrategySort* procedure in Algorithm 1 sorted the farms according to the adopted strategy. The following order-based strategies were used in the experiments:

- *Strength-based* that prefers nodes of the graph with a higher strength. In the experiments the strength was calculated using either the number of animals transported (*NumAnimalsInMovements*), the number of movements in which a given node was involved (*NumMovements*) or the number of locations to which the node was connected (*NumLocations*, which is equal to the node degree). Each measure was calculated in the *In*, *Out* and *Both* directions, so the total number of strength based strategies was

nine: (*NumAnimalsInMovements*, *NumMovements*, *NumLocations*) × (*In*, *Out*, *Both*).
- *Farm size* that prefers farms with more animals.
- *Random* that shuffles the farms in a random order.

Apart from order-based strategies a strategy called *Rings* was used which combined the idea of ring vaccination (Toma et al. (2002)) and acquaintance vaccination (Ball and Sirl (2013)). Ring vaccination strategy calls for vaccinating all nodes adjacent to the infected ones, with an obvious intent to cut off the infected nodes from the rest of the graph. Acquaintance vaccination strategy suggests vaccinating nodes adjacent to previously vaccinated ones. The observation supporting the use of this strategy is that those "adjacent nodes" have higher probability of being selected from high-degree nodes. The *Rings* strategy generates a set of solutions $P = \{x_0, x_1, ..., \}$. The solution x_0 is a vector in $\Omega = \{0,1\}^{N_{farms}}$ in which farms adjacent to those infected in S'_0 are marked by the value 1 (selected for vaccination). For $n = 1, \ldots$ the solution x_{n+1} is constructed by selecting for vaccination all farms from x_n and farms adjacent to them. As with all previously discussed strategies, this strategy uses the known initial state S'_0 and the rewired graph $G'(\alpha_{rewire})$ to decide which farms to vaccinate. Only after the decision which farms to vaccinate is made, the solutions are evaluated using the true initial state S_0 and the original graph G as if applied to a real scenario.

4.2 Evolutionary Algorithms

Apart from the strategies, three multiobjective evolutionary algorithms were used for generating solutions to the vaccination optimization problem. These algorithms belong to two main groups of multiobjective evolutionary algorithms: based on decomposition (MOEA/D) and based on Pareto dominance (NSGA-II and SPEA2). The MOEA/D algorithm (Li and Zhang (2009); Zhang and Li (2007)) is one of the most popular decomposition-based algorithms used in the literature for solving multiobjective optimization problems. In this article Tchebycheff decomposition was used for scalarizing multiple objectives in the MOEA/D algorithm.

As the Pareto-based methods NSGA-II (Deb et al. (2002)) and SPEA2 (Zitzler et al. (2002a)) algorithms were used. Both algorithms were, to date, used for solving numerous optimization problems such as stock market portfolio optimization (Kaucic et al. (2019)), inventory optimization (Cholodowicz and Orlowski (2017)) and engineering problems (Gadhvi et al. (2016)).

Since the search space in the vaccination optimization problem is $\Omega = \{0,1\}^{N_{farms}}$ the solutions were represented using binary vectors of length $N_{farms} = 9313$ in all the algorithms. Three crossover operators (*SinglePoint*, *TwoPoint* and *Uniform*) were used as well as six mutation operators (*BitFlip*, *Displacement*, *Insertion*, *Inversion*, *Scramble* and *Transpose*). To determine the probabilities for applying the operators an auto-adaptation mechanism described in (Michalak (2014a)) was used.

4.2.1 Local Search

Solutions to the vaccination optimization problem are evaluated using simulations. When simulations are performed, it is possible to store the number of times each node in the graph was infected in a vector $F_{infected} \in \mathbb{N}^{N_{farms}}$. This information can be used for improving the search capabilities of the evolutionary algorithm with a dedicated local search.

The local search starts from a solution $x_0 \in \Omega$ and attempts to construct a solution with fewer vaccinated farms x^- by setting to 0 those elements of the vector x_0 which are equal to 1 and correspond to the lowest values in the $F_{infected}$ vector (indicating rarely infected farms). Conversely, a solution with more vaccinated farms x^+ can be constructed by setting to 1 those elements of the vector x_0 which are equal to 0 and correspond to the highest values in the $F_{infected}$ vector (indicating frequently infected farms). In each run of the local search a fraction α_{LS} of the elements of x_0 can be changed.

Algorithm 2 presents the working of the local search used in this article. In this algorithm the *Evaluate* procedure is used for evaluating a solution in the same way as in the evolutionary algorithm. The notation $x_1 \succ x_2$ signifies that the solution x_1 is considered better than x_2 and thus should be accepted by the local search. Of course, in multiobjective optimization problems solutions cannot be ordered linearly because one solution may be better with respect to some objectives and the other solution with respect to some other objectives. This issue can be addressed by scalarizing multiple objectives or by using the concept of Pareto dominance (equation (3)). Therefore, the working of the $\succ$ operator used in the local search procedure is different for the MOEA/D, which uses scalarized objectives and for the Pareto-based algorithms (NSGA-II and SPEA2).

Since the MOEA/D algorithm uses weight vectors for multiobjective problem decomposition, solutions tested by the local search starting from the solution x_0 are compared by scalarization using the weight vector λ_0 assigned by the MOEA/D to x_0. The local search

procedure uses the utopia point $x^{(U)}$ (best values of the objectives found so far) and the nadir point $x^{(N)}$ (worst values of the objectives found so far) for objective values normalization. When the $\succ$ operator compares two solutions x_1 and x_2 the vectors of objectives:

$$\begin{aligned} F(x_1) &= (f_1(x_1), \ldots, f_m(x_1)) \\ F(x_2) &= (f_1(x_2), \ldots, f_m(x_2)) \end{aligned} \tag{6}$$

are normalized using the positions of the utopia point $x^{(U)}$ and the nadir point $x^{(N)}$:

$$\forall j = 1, \ldots, m: \quad f'_j(x_1) = \frac{f_j(x_1)}{\left| x_j^{(N)} - x_j^{(U)} \right|} \ , \quad f'_j(x_2) = \frac{f_j(x_2)}{\left| x_j^{(N)} - x_j^{(U)} \right|} \ . \tag{7}$$

The normalized vectors of objectives:

$$\begin{aligned} F'(x_1) &= (f'_1(x_1), \ldots, f'_m(x_1)) \\ F'(x_2) &= (f'_1(x_2), \ldots, f'_m(x_2)) \end{aligned} \tag{8}$$

are scalarized using the weight vector λ_0 and the resulting scalar values are used for comparing the solutions:

$$x_1 \succ x_2 \equiv \lambda_0 \cdot F'(x_1) < \lambda_0 \cdot F'(x_2) \ , \tag{9}$$

where $\cdot$ denotes a dot product of the weight vector λ_0 and the corresponding vector of objectives.

When the local search is used with the NSGA-II and SPEA2 the $x_1 \succ x_2$ condition signifies that x_1 Pareto-dominates x_2 (equation (3)), which is in line with the main idea used in these algorithms, that is, Pareto-dominance.

In the experiments the local search described in this section was run for each solution found by the evolutionary algorithm with a probability P_{LS}.

Algorithm 2: Local search used in this article.

Inputs:
- x_0 - the initial solution found by the evolutionary algorithm
- $F(x_0)$ - vector of objectives for x_0
- $F_{infected}$ - infection counters for farms

Inputs used only with MOEA/D:
- λ_0 - the weight vector corresponding to x_0
- $x^{(U)}$ - the utopia point
- $x^{(N)}$ - the nadir point

Output:
- x'_0 - an improved solution (or x_0 if none found)

```
// Number of elements to change
N_LS := α_LS · N_farms

// Ordering of the farms in the order of ascending
// infection counters
R := {1, ..., N_farms}
R := Sort(R, F_infected, 'ascending')

// A solution with fewer vaccinated farms
x⁻ := x_0
N_changed := 0
for i := 1, ..., N_farms do
    if x⁻[R[i]] = 1 then
        x⁻[R[i]] := 0
        N_changed := N_changed + 1
        if N_changed ≥ N_LS then
            break

// Replace the original solution by a better one
if N_changed > 0 then
    F(x⁻) := Evaluate(x⁻)
    x^(U) := min(x^(U), F(x⁻))
    x^(N) := max(x^(N), F(x⁻))
    if x⁻ ≻ x_0 then
        x_0 := x⁻
        F(x_0) := F(x⁻)

// A solution x⁺ with more vaccinated farms
// constructed and tested in the same way as x⁻
...

// Return the updated solution
x'_0 := x_0
```

4.3 Experimental Setup

In the experiments the evolutionary algorithms described in section 4.2 were tested and compared to twelve strategies described in section 4.1: nine *strength-based* strategies with edge types (*NumAnimalsInMovements*, *NumMovements*, *NumLocations*) and edge directions (*In*, *Out*, *Both*), and the strategies *FarmSize*, *Random* and *Rings*.

As discussed in previous sections test scenarios considered in this article are parameterized with several parameters. Two parameters that describe how much (or little) is known about the actual outbreak and animal movements are: the fraction of known disease cases α_{known} and the fraction of rewired edges in the contact network α_{rewire}. In the experiments these parameters were set to $\alpha_{known} = 0.1$, 0.2, 0.5, 0.8, 0.9 and 1.0, and $\alpha_{rewire} = 0.00$, 0.02, 0.04, 0.06, 0.08, 0.10, 0.20, 0.30, 0.40 and 0.50. These parameters are intended to represent the difference between information available when

vaccination scheme is decided upon and when vaccinations are actually performed. Obviously, their values cannot be adjusted by any of the tested methods, but have to be treated as given. The R_a parameter controlling the generation of additional disease cases was set to three values: $R_a = 0$, 1 and 2 and was not modified by the tested methods. Therefore, each method was tested on each triple of values $\langle \alpha_{rewire}, \alpha_{known}, R_a \rangle$ (180 different triples, overall). For each triple of the parameter values 30 tests were performed using initial states S_0 and S_0' for initializing simulations, and the original graph G and the rewired graph $G(\alpha_{rewire})$ for simulating the spread of the disease. The initial states and graphs were independently, randomly generated for each of the parameter triples and 30 test instances. In order to obtain comparable results, each method tested in one of the 30 tests for a given triple of values $\langle \alpha_{rewire}, \alpha_{known}, R_a \rangle$ was run with the same S_0, S_0', G, and $G(\alpha_{rewire})$.

In each test each method generated a set of solutions using the known initial state S_0' and the rewired graph $G(\alpha_{rewire})$. For example, the *Rings* strategy selected for vaccination the nodes adjacent to those infected in S_0'. Subsequently nodes connected to the vaccinated ones with edges of the graph $G(\alpha_{rewire})$ were selected. Similarly, the evolutionary algorithm performed evaluations of solutions by simulating the epidemic spreading on the graph $G(\alpha_{rewire})$ starting from the known initial state S_0'. After the methods finished generating solutions, these solutions were evaluated again using simulations starting from the true initial state S_0 and using the original graph G in order to simulate using the discovered solutions in a real epidemic. The results of this final evaluation were used for comparing methods tested in the experiments.

Since the vaccination optimization problem is multiobjective, the tested methods produced Pareto fronts of non-dominated solutions. In order to compare these Pareto fronts the hypervolume indicator (Zitzler et al. (2002b)) was used. The hypervolume is calculated for a set of points P as the Lebesgue measure of the portion of the objective space that is dominated by solutions in P collectively and bounded by a reference point $r^{(N)}$:

$$HV(P) = L\left(\bigcup_{x \in P} \left[f_1(x), r_1^{(N)}\right] \times \ldots \times \left[f_m(x), r_m^{(N)}\right]\right), \tag{10}$$

where:

m - the dimensionality of the objective space,
$f_i(\cdot)$, $i = 1, \ldots m$ - the objective functions,
$r^{(N)} = (r_1^{(N)}, \ldots, r_m^{(N)})$ - a reference point,
$L(\cdot)$ - the Lebesgue measure on $\mathbb{R}^m$.

For a minimization problem the reference point $r^{(N)}$ should be selected in such a way that all the points in the compared Pareto fronts have coordinates smaller than $r^{(N)}$. Thus, most often the reference point $r^{(N)}$ is obtained by calculating the global nadir point (containing the *worst* coordinates) from all solutions found by the tested methods. In the biobjective case, the hypervolume is equal to the area of the subset of the objective space that is dominated by a given set of solutions. For each of the 30 tests the hypervolume was calculated and the median of these results was used as the final result attained by a given method.

5 Results

In this section the results of the experiments are discussed. Because evolutionary algorithms and the local search require setting some parameters, parameter tuning of these methods was performed first, then a comparison with vaccination strategies was performed, and finally the addition of artificial disease cases was tested as a mechanism for substituting the unknown ones. Since the vaccination optimization problem is multiobjective, numerical comparison of the results produced by various methods was performed using the hypervolume indicator. The reference point $r^{(N)} = [560623, 240724.4]$ used for hypervolume calculation (cf. equation (10)) was obtained from all solutions found by the tested methods.

5.1 Parameter Tuning

The optimization algorithms used in the experiments, as well as the local search procedure, require setting some parameters, which can influence their performance. The MOEA/D algorithm requires, apart from the population size and operator probabilities, also some parameters determining how large the neighbourhood of each solution is and how it is used in the algorithm. Based on values available in the literature, in particular in the original articles on MOEA/D (Li and Zhang (2009); Zhang and Li (2007)) the values of the maximum number of neighbours replaced $n_r = 2$ and the probability of selecting parents from the neighbourhood $\delta = 0.9$ were selected. According to the findings in the article (Michalak (2014b)) the number of neighbours was set to an odd number $T = 21$ rather than an even one. The population size N_{pop}, the crossover probability P_{cross}, and the mutation probability P_{mut} for the evolutionary algorithms were tuned using the grid

search approach in preliminary experiments. Tested values were in the range: $N_{pop} \in \{ 50, 100, 200, 500 \}$, $P_{cross} \in \{ 0.2, 0.4, 0.6, 0.8, 1.0 \}$ and $P_{mut} \in \{ 0.02, 0.04, 0.06, 0.08, 0.10 \}$. The stopping criterion was set to $max_{FE} = 10000$ solution evaluations and tests were performed using all possible triples of the N_{pop}, P_{cross}, and P_{mut} parameter values. The values of these parameters for which the highest median hypervolume was produced were selected as the best ones. After tuning the parameters for evolutionary algorithms, the parameters of the local search were tuned in the same manner. The tested ranges for the local search parameters were $P_{LS} \in \{ 0.2, 0.4, 0.6, 0.8, 1.0 \}$ and $\alpha_{LS} \in \{ 0.02, 0.04, 0.06, 0.08, 0.10 \}$. Table 1 lists the values of the parameters obtained in the parameter-tuning phase of experiments for each algorithm separately.

Table 1 The values of the parameters obtained using the grid search approach.

Parameter name	MOEA/D	NSGA-II	SPEA2
Population size N_{pop}	200	100	100
Crossover probability P_{cross}	1.0	1.0	1.0
Mutation probability P_{mut}	0.08	0.02	0.02
LS probability (per solution) P_{LS}	0.6	0.4	0.6
Fraction of positions changed α_{LS}	0.10	0.04	0.04

It can be observed that MOEA/D works better with a larger population ($N_{pop} = 200$) and Pareto-based algorithms (NSGA-II and SPEA2) require smaller populations ($N_{pop} = 100$). Also, MOEA/D works better with a higher mutation rate ($P_{mut} = 0.08$ vs. $P_{mut} = 0.02$). The best crossover probability for all three algorithms turned out to be 1.0. Clearly, it is beneficial to perform the local search on about half the solutions in the population as the best probability of applying the local search is $P_{LS} = 0.4$ or 0.6 depending on the evolutionary algorithm used. Interestingly, the best parameter settings for both Pareto-based algorithms are the same with the MOEA/D requiring a larger population, a higher mutation rate and a more intensive local search (higher values of P_{LS} and α_{LS}).

Figure 2 shows combined Pareto fronts obtained from 30 runs using the three evolutionary algorithms with and without local search with the best parameter settings. Table 2 shows median hypervolume values obtained from 30 runs of these algorithms. The same table contains p-values obtained using the Wilcoxon statistical test (Wilcoxon (1945)) with the null hypothesis stating the equality of the medians. Each p-value shows if the difference between median hypervolume attained by two algorithms (corresponding to the row and column) is statistically significant. Only the difference between NSGA-II and SPEA2 without local search cannot be considered statistically significant (p-value of 0.504). In all the other cases p-values are lower than 0.001 indicating statistically significant differences between median hypervolume values produced by different algorithms. In the "Hypervolume" column it can be seen that the best performing method is the MOEA/D with local search. This observation corresponds to Pareto fronts shown in Figure 2, because the Pareto front for MOEA/D with local search contains the most diversified solutions with the lowest values of both objectives.

Pareto fronts presented in Figure 2 and numerical results shown in Table 2 confirm that the local search used in this article improves the search capabilities of all the algorithms. Hypervolume values in Table 2 are higher for algorithms with local search than without it and differences between algorithms with and without local search are statistically significant. Pareto fronts obtained using local search extend much wider, offering a greater range of possibilities for a decision maker to choose from including scenarios with very few vaccinations (good when the vaccine is in short supply) and those with numerous vaccinations reducing the number of infections. Overall, the best algorithm is the MOEA/D with the local search with the parameters shown in Table 1.

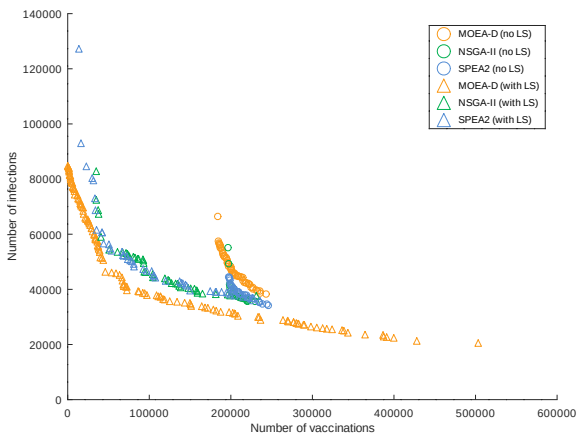


Fig. 2 Comparison of Pareto fronts obtained from 30 runs of the three tested evolutionary algorithms with and without local search with the best parameter settings.

Table 2 Median hypervolume values obtained from 30 runs of all three evolutionary algorithms with and without local search with the best parameter settings. The p-values were obtained using the Wilcoxon statistical test with the null hypothesis stating the equality of the medians.

Algorithm name	Hyper-volume	p-value					
		MOEA/D (no LS)	NSGA-II (no LS)	SPEA2 (no LS)	MOEA/D (with LS)	NSGA-II (with LS)	SPEA2 (with LS)
MOEA/D (no LS)	$7.11 \cdot 10^{10}$	—	$1.74 \cdot 10^{-4}$	$1.89 \cdot 10^{-4}$	$1.73 \cdot 10^{-6}$	$1.73 \cdot 10^{-6}$	$1.73 \cdot 10^{-6}$
NSGA-II (no LS)	$6.98 \cdot 10^{10}$	$1.74 \cdot 10^{-4}$	—	$5.04 \cdot 10^{-1}$	$1.73 \cdot 10^{-6}$	$1.73 \cdot 10^{-6}$	$1.73 \cdot 10^{-6}$
SPEA2 (no LS)	$7.01 \cdot 10^{10}$	$1.89 \cdot 10^{-4}$	$5.04 \cdot 10^{-1}$	—	$1.73 \cdot 10^{-6}$	$1.73 \cdot 10^{-6}$	$1.73 \cdot 10^{-6}$
MOEA/D (with LS)	$\underline{1.13 \cdot 10^{11}}$	$1.73 \cdot 10^{-6}$	$1.73 \cdot 10^{-6}$	$1.73 \cdot 10^{-6}$	—	$1.73 \cdot 10^{-6}$	$1.73 \cdot 10^{-6}$
NSGA-II (with LS)	$9.07 \cdot 10^{10}$	$1.73 \cdot 10^{-6}$	$1.73 \cdot 10^{-6}$	$1.73 \cdot 10^{-6}$	$1.73 \cdot 10^{-6}$	—	$1.89 \cdot 10^{-4}$
SPEA2 (with LS)	$9.72 \cdot 10^{10}$	$1.73 \cdot 10^{-6}$	$1.73 \cdot 10^{-6}$	$1.73 \cdot 10^{-6}$	$1.73 \cdot 10^{-6}$	$1.89 \cdot 10^{-4}$	—

5.2 Comparison of evolutionary optimization and vaccination strategies

In this round of experiments MOEA/D, NSGA-II and SPEA2 were compared with vaccination strategies described in Section 4.1.

First, the tested methods were compared separately for each of the 180 triples of values $\langle \alpha_{rewire}, \alpha_{known}, R_a \rangle$ and the number of times each method attained the best median hypervolume calculated from 30 repetitions of the tests was counted. In the count only results statistically significant at the level $\alpha = 0.01$ were considered. When no statistically significant difference was found for a given triple of parameters $\langle \alpha_{rewire}, \alpha_{known}, R_a \rangle$ such results were not included in the count. Table 3 presents the results of this comparison. There are three methods that attained the best hypervolume for at least one of the 180 triples of parameters: the MOEA/D optimization algorithm and two strength-based strategies using the number of locations to which a node is connected (*NumLocations*) and the number of animal movements (*NumMovements*). In both cases the strength was based on incoming as well as outgoing movements (the direction is *Both*).

Clearly, when node strength was used in the vaccination strategy, the most important characteristics turned out to be the number of locations to which the node was connected and the number of movements which went from, or to the node. The *NumLocations* parameter is the node degree and determines how many locations the disease may spread to, so is understandably important. The *NumMovements* parameter is important, because of the temporal dependencies in the system. The disease can only spread from a node with movements that occur after that node is infected. Therefore, if a node is involved in multiple movements the chances that the disease will spread from this node increase.

In order to better understand the influence of the uncertainties quantified by the parameters α_{known} and α_{rewire} the number of times evolutionary optimization produced the best result was compared to the number of times vaccination strategies produced the best result, separately for each value of each of these parameters. In the comparison the methods that obtained the best result at least one time were included, that is MOEA/D and two strategies: Strength(*NumLocations*, *Both*) and Strength(*NumMovements*, *Both*). Similarly as with the values presented in Table 3 only results statistically significant at the level $\alpha = 0.01$ were considered. Results for the known cases fraction α_{known} are presented in Figure 3. It can be observed that vaccination strategies perform well when only a small fraction of initial disease cases are known. The MOEA/D outperforms the

Table 3 Comparison of performance of the tested methods with respect to the median hypervolume calculated from 30 repetitions of the tests for each of the 180 triples of values $\langle \alpha_{rewire}, \alpha_{known}, R_a \rangle$. #best is the number of times a method performed better than all the other methods for a given $\langle \alpha_{rewire}, \alpha_{known}, R_a \rangle$ triple. In the count only results statistically significant at the level $\alpha = 0.01$ were considered.

Method	#best
MOEA/D	61
NSGA-II	0
SPEA2	0
Strength(*NumLocations*, *Both*)	40
Strength(*NumLocations*, *In*)	0
Strength(*NumLocations*, *Out*)	0
Strength(*NumAnimalsInMovements*, *Both*)	0
Strength(*NumAnimalsInMovements*, *In*)	0
Strength(*NumAnimalsInMovements*, *Out*)	0
Strength(*NumMovements*, *Both*)	7
Strength(*NumMovements*, *In*)	0
Strength(*NumMovements*, *Out*)	0
Farm size	0
Random	0
Ring	0

strategies when at least 80% of initial disease cases are known when optimization is performed.

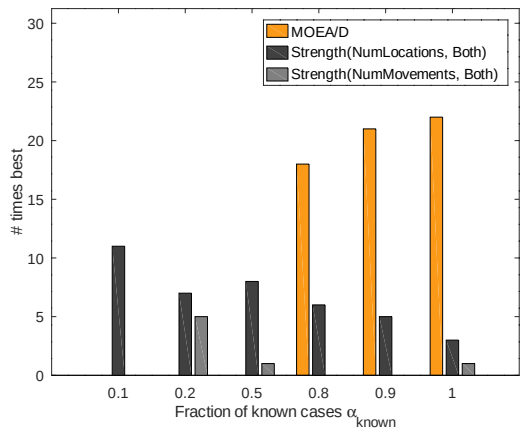


Fig. 3 The number of times evolutionary optimization and vaccination strategies attained the best result. Counted separately for each value of the known fraction of initial disease cases α_{known}.

Results for the rewired movements fraction α_{rewire} are presented in Figure 4. It can be observed that vaccination strategies are the most effective when a large fraction of movements is rewired. The MOEA/D outperforms vaccination strategies when at most 20% of movements are rewired. For 30% of movements changed by rewiring or more, evolutionary optimization performs worse than vaccination strategies. It should be noted, however, that the adopted rewiring scheme, while based on realistic assumptions, does not change the degrees and strengths of the nodes and thus favors the strategies based on these node attributes.

The influence of unknown disease cases on the Pareto fronts can be assessed by comparing Figure 5 a) and b). When only 10% of the disease cases are known ($\alpha_{known} = 0.1$) the performance of all three optimization algorithms is deteriorated, with MOEA/D producing the most diversified Pareto front of the three evolutionary algorithms, but worse than the vaccination strategies. The influence of changes in the network of contacts on the Pareto fronts can be assessed by comparing Figure 5 a) and c). In the presence of intensive rewiring ($\alpha_{rewire} = 0.50$) the performance of evolutionary algorithms is deteriorated, with MOEA/D being less impacted that NSGA-II and SPEA2.

5.3 Influence of artificial disease cases

From the results presented in Section 5.2 it can be concluded, that optimization is able to outperform vaccination strategies if enough initial cases of the disease are known ($\alpha_{known} \geq 0.8$) and when the change in the contacts network is not large ($\alpha_{rewire} \leq 0.2$). In the experiments artificial disease cases were generated in order to compensate for the lack of knowledge of the initial cases of the disease. This section presents results obtained when optimization was performed with and without artificial disease cases. In the comparison the results obtained using MOEA/D were used, because its performance was the best from the three tested evolutionary algorithms. The performance of the evolutionary optimizer was compared to the results obtained using vaccination strategies. Figure 6 presents values of rewired movements fraction α_{rewire} and known disease cases fraction α_{known} for which a better hypervolume was attained by the MOEA/D (orange) and those for which a better hypervolume was attained by one of the strength-based strategies (gray). Subplots in this figure present comparison of MOEA/D and strategies for R_a = 0, 1, 2 and, in the subplot entitled "R_a = best", a situation when the best of results obtained for R_a = 0, 1, 2 are compared. It can be seen that generating artificial disease cases ($R_a = 2$) can improve optimization results when only a half of the initial disease cases are known ($\alpha_{known} = 0.5$). However, this improvement is obtained at the cost of deteriorating the results when most disease cases are actually known ($R_a = 2$, α_{known} = 0.9 and 1.0). In an idealized case (R_a = all) the MOEA/D is able to outperform vaccination strategies up to $\alpha_{known} = 0.5$ and, although not simultaneously, $\alpha_{rewired} = 0.3$. The setting of R_a = all corresponds to selecting the best possible value of R_a for a given value of α_{known}.

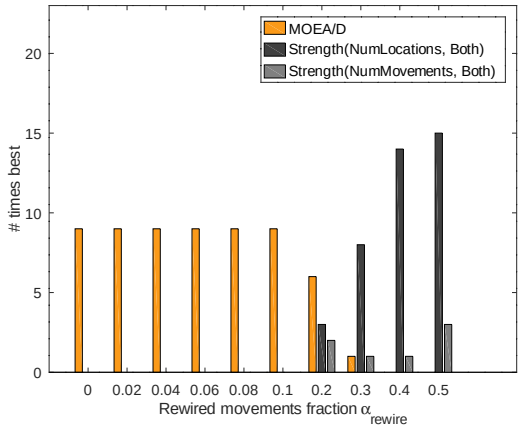
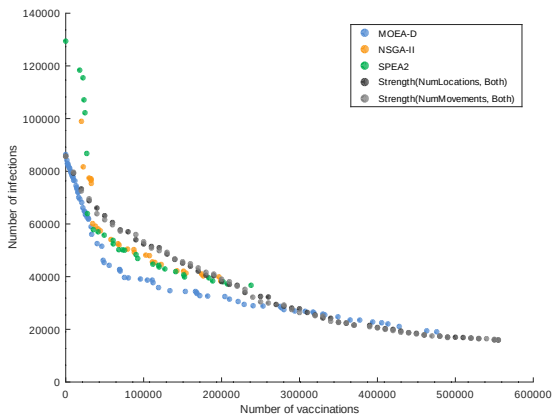
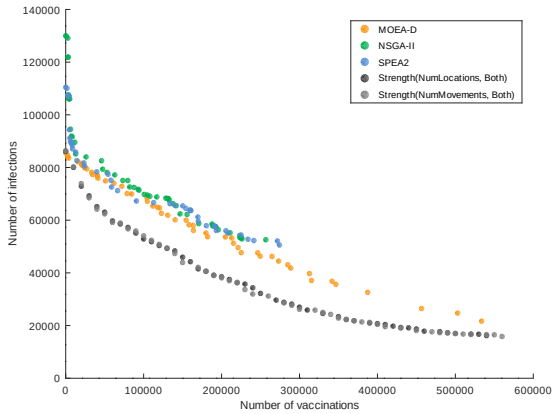


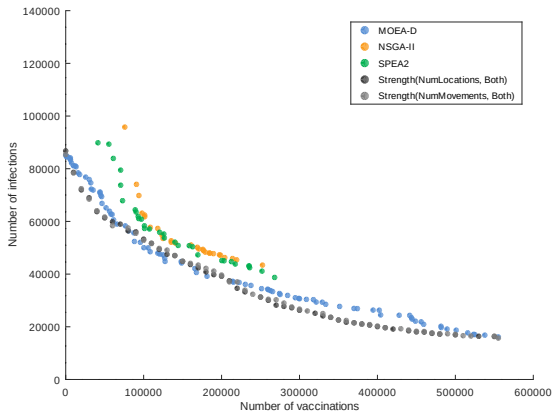
Fig. 4 The number of times evolutionary optimization and vaccination strategies attained the best result. Counted separately for each value of the rewired movements fraction α_{rewire}.



a) $\alpha_{known} = 1.0$, $\alpha_{rewire} = 0.00$



b) $\alpha_{known} = 0.1$, $\alpha_{rewire} = 0.00$



c) $\alpha_{known} = 1.0$, $\alpha_{rewire} = 0.50$

Fig. 5 Pareto fronts obtained in optimization scenarios with a) no unknown disease cases and no rewiring ($\alpha_{known} = 1.0$, $\alpha_{rewire} = 0.00$); b) most of the disease cases unknown and no rewiring ($\alpha_{known} = 0.1$, $\alpha_{rewire} = 0.00$); c) no unknown disease cases and intensive rewiring ($\alpha_{known} = 1.0$, $\alpha_{rewire} = 0.50$).

Similar conclusions can be drawn from Figure 7 which presents hypervolume values attained by MOEA/D (orange) and the best of the strength-based strategies (gray) depending on the value of rewired movements fraction α_{rewire} and known disease cases fraction α_{known}. In this figure the decrease of the hypervolume caused by unknown disease cases is larger than the decrease cause by rewiring, an effect that can also be seen in Figure 5 showing Pareto fronts. The Pareto front attained by MOEA/D is close to the one produced by vaccination strategies when $\alpha_{rewire} = 0.50$, but is much worse for $\alpha_{known} = 0.1$. In Figure 7 it can be seen that adding artificial disease cases improves optimization results when only a few initial disease cases are known (low values of α_{known}), but deteriorates the results when most of the initial cases are known (high values of α_{known}).

6 Conclusions

In this article evolutionary optimization of vaccinations was compared with vaccination strategies such as those prioritizing vaccination of nodes with a high strength or degree. In the experiments the influence of uncertainties was tested, represented by the lack of knowledge of initial disease cases and the change of the contacts network by a rewiring process. The experiments shown that the MOEA/D algorithm used for optimization can outperform vaccination strategies, however, its performance is degraded when unknown initial cases of the disease are present ($\alpha_{known} < 0.8$). Also, changes in the contacts network have a negative impact on optimization as evidenced by the fact that the evolutionary algorithm performed worse than the strategies for $\alpha_{rewire} > 0.3$. Naturally, it is important to note, that the adopted rewiring process rewards the degree- and strength-based vaccination strategies, because degrees and strengths of the graph nodes remain unchanged. Nevertheless, the performance of the optimization algorithm is somewhat degraded also in absolute values.

Presented research shows strengths and weaknesses of both the optimization of vaccinations and vaccination strategies discussed in this article. Evolutionary optimization produces good solutions, but when there are many unknown disease cases the performance is degraded. This effect can be mitigated by adding artificial disease cases to the simulations used when evaluating solutions in the evolutionary algorithm. However, this approach has to be applied carefully, because when too many artificial disease cases are added the performance can be degraded. Solutions provided by vaccination strategies are inferior to the optimized ones when most of the disease cases are known, but vaccination

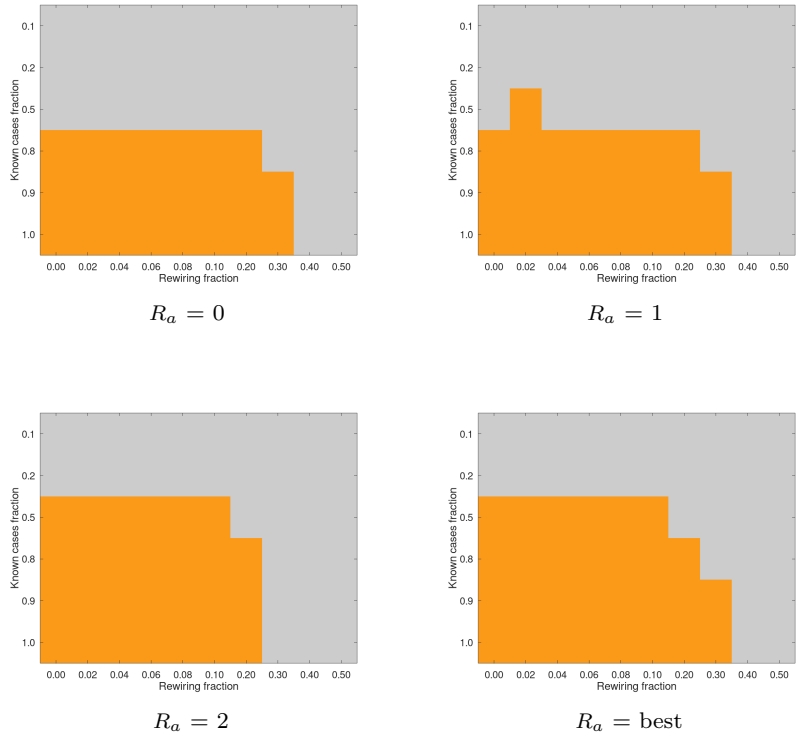


Fig. 6 Values of rewired movements fraction α_{rewire} and known disease cases fraction α_{known} for which a better hypervolume was attained by the MOEA/D (orange) and those for which a better hypervolume was attained by one of the strength-based strategies (gray).

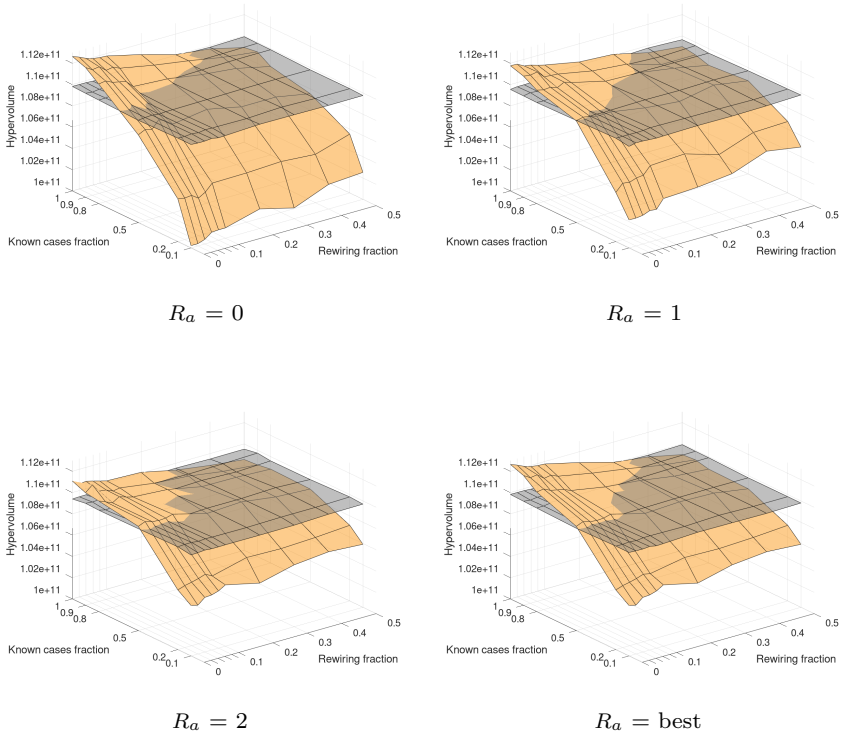


Fig. 7 Hypervolume values attained by the MOEA/D (orange) and the best of the strength-based strategies (gray) depending on the value of rewired movements fraction α_{rewire} and known disease cases fraction α_{known}.

strategies are less affected by uncertainties as evidenced in Figure 7 by a smaller decrease of hypervolume for larger values of α_{rewire} and smaller values of α_{known}.

Obtained results motivate further work on optimization methods for the vaccination optimization problem described in this article aimed at reducing the impact of uncertainties on optimization and combining the strengths of both evolutionary optimization and vaccination strategies. Because of the large impact of uncertainties on the performance of the optimization algorithms it seems worthwhile to attempt modelling and predicting changes in the patterns of animal movements. Another approach that could be attempted would be to better model the unknown initial cases of the disease. The comparison performed in an idealized case (selecting the value of R_a producing the best results) suggests that optimization methods may benefit from correct identification of the fraction of unknown disease cases. The fact that different types of methods show good performance under different conditions suggests that hybrid methods combining evolutionary optimization and vaccination strategies could be effective in solving the vaccination optimization problem described in this article.

Acknowledgment

The authors would like to acknowledge Informatica Area Prevenzione of the ASL CN1 of Piedmont for the data and useful discussion.

Compliance with ethical standards

Funding This study was funded by the Polish National Agency for Academic Exchange (NAWA) within the Bekker programme (grant number PPN/BEK/2018/1/00430). Calculations have been carried out using resources provided by Wroclaw Centre for Networking and Supercomputing (http://wcss.pl), grant No. 407.

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

Bajardi P, Barrat A, Natale F, Savini L, Colizza V (2011) Dynamical patterns of cattle trade movements. PLOS ONE 6(5):1–19

Ball F, Sirl D (2013) Acquaintance vaccination in an epidemic on a random graph with specified degree distribution. Journal of Applied Probability 50(4):1147–1168

Barma P, Dutta J, Mukherjee A (2019) A 2-opt guided discrete antlion optimization algorithm for multi-depot vehicle routing problem. Decision Making: Applications in Management and Engineering 2

Biswas P, Pal BB (2019) A fuzzy goal programming method to solve congestion management problem using genetic algorithm. Decision Making: Applications in Management and Engineering 2(2):36–53

Blum C, Blesa MJ, García-Martínez C, Rodríguez FJ, Lozano M (2014) The firefighter problem: Application of hybrid ant colony optimization algorithms. In: Blum C, Ochoa G (eds) Evolutionary Computation in Combinatorial Optimisation, LNCS, vol 8600, Springer Berlin Heidelberg, pp 218–229

Bozzette SA, Boer R, Bhatnagar V, Brower JL, Keeler EB, Morton SC, Stoto MA (2003) A model for a smallpox-vaccination policy. New England Journal of Medicine 348(5):416–425

Brauer F (2008) Compartmental Models in Epidemiology, Springer Berlin Heidelberg, Berlin, Heidelberg, pp 19–79

Cholodowicz E, Orlowski P (2017) Comparison of SPEA2 and NSGA-II applied to automatic inventory control system using hypervolume indicator. Studies in Informatics and Control 26:67–74

da Cruz AR, Cardoso RT, Takahashi RH (2017) Multiobjective synthesis of robust vaccination policies. Applied Soft Computing 50:34 – 47

Dalgıç ÖO, Özaltın OY, Ciccotelli WA, Erenay FS (2017) Deriving effective vaccine allocation strategies for pandemic influenza: Comparison of an agent-based simulation and a compartmental model. PLOS ONE 12(2):1–19

Danila R, Nika M, Wilding T, Knottenbelt WJ (2014) Uncertainty in on-the-fly epidemic fitting. In: Horváth A, Wolter K (eds) Computer Performance Engineering, Springer International Publishing, Cham, pp 135–148

Deb K (2001) Multi-Objective Optimization Using Evolutionary Algorithms. John Wiley & Sons, Inc., New York, NY, USA

Deb K, Pratap A, Agarwal S, Meyarivan T (2002) A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation 6:182–197

Dubé C, Ribble C, Kelton D, McNab B (2009) A review of network analysis terminology and its application to foot-and-mouth disease modelling and policy development. Transboundary and Emerging Diseases

56(3):73–85

Gadhvi B, Savsani V, Patel V (2016) Multi-objective optimization of vehicle passive suspension system using NSGA-II, SPEA2 and PESA-II. Procedia Technology 23:361–368

Grimm V, Railsback S (2005) Individual-based Modeling and Ecology. EBSCO ebook academic collection, Princeton University Press

Hartnell B (1995) Firefighter! an application of domination. In: 20th Conference on Numerical Mathematics and Computing

Hu B, Windbichler A, Raidl GR (2015) A new solution representation for the firefighter problem. In: Ochoa G, Chicano F (eds) Evolutionary Computation in Combinatorial Optimization, LNCS, vol 9026, Springer, pp 25–35

Juan AA, Faulin J, Grasman SE, Rabe M, Figueira G (2015) A review of simheuristics: Extending metaheuristics to deal with stochastic combinatorial optimization problems. Operations Research Perspectives 2:62–72

Kaucic M, Moradi M, Mirzazadeh M (2019) Portfolio optimization by improved NSGA-II and SPEA 2 based on different risk measures. Financial Innovation 5(1):26

Li H, Zhang Q (2009) Multiobjective optimization problems with complicated pareto sets, MOEA/D and NSGA-II. IEEE Trans Evol Comput 13(2):284–302

Li SL, Bjørnstad ON, Ferrari MJ, Mummah R, Runge MC, Fonnesbeck CJ, Tildesley MJ, Probert WJM, Shea K (2017) Essential information: Uncertainty and optimal control of Ebola outbreaks. Proceedings of the National Academy of Sciences 114(22):5659–5664

Lipinski P (2017) Solving the firefighter problem with two elements using a multi-modal estimation of distribution algorithm. In: 2017 IEEE Congress on Evolutionary Computation (CEC), pp 2161–2168

Michalak K (2014a) Auto-adaptation of genetic operators for multi-objective optimization in the firefighter problem. In: Corchado E, Lozano JA, Quintián H, Yin H (eds) Intelligent Data Engineering and Automated Learning - IDEAL 2014, LNCS, vol 8669, Springer, pp 484–491

Michalak K (2014b) The effects of asymmetric neighborhood assignment in the MOEA/D algorithm. Applied Soft Computing 25:97–106

Michalak K (2019) Solving the parameterless firefighter problem using multiobjective evolutionary algorithms. In: Proceedings of the Genetic and Evolutionary Computation Conference Companion, ACM, New York, NY, USA, GECCO '19, pp 1321–1328

Michalak K, Knowles JD (2016) Simheuristics for the multiobjective nondeterministic firefighter problem in a time-constrained setting. In: Squillero G, Burelli P (eds) Applications of Evolutionary Computation: 19th European Conference, EvoApplications 2016, Porto, Portugal, March 30 – April 1, 2016, Proceedings, Part II, Springer International Publishing, pp 248–265

Miettinen K (1999) Nonlinear Multiobjective Optimization, International Series in Operations Research and Management Science, vol 12. Kluwer Academic Publishers, Dordrecht

Parousis-Orthodoxou KJ, Vlachos DS (2014) Evolutionary algorithm for optimal vaccination scheme. Journal of Physics: Conference Series 490:012027

Pellerin R, Perrier N, Berthaut F (2020) A survey of hybrid metaheuristics for the resource-constrained project scheduling problem. European Journal of Operational Research 280(2):395–416

Rautureau S, Dufour B, Durand B (2011) Vulnerability of animal trade networks to the spread of infectious diseases: A methodological approach applied to evaluation and emergency control strategies in cattle, france, 2005. Transboundary and Emerging Diseases 58(2):110–120

Roy A, Manna A, Maity S (2019) A novel memetic genetic algorithm for solving traveling salesman problem based on multi-parent crossover technique. Decision Making: Applications in Management and Engineering 2

Talbi EG, Basseur M, Nebro A, Alba E (2012) Multi-objective optimization using metaheuristics: Non-standard algorithms. International Transactions in Operational Research 19

Toma B, Moutou F, Dufour B, Durand B (2002) Ring vaccination against foot-and-mouth disease. Comparative Immunology, Microbiology and Infectious Diseases 25(5):365 – 372

Tornatore E, Vetro P, Buccellato SM (2014) SIVR epidemic model with stochastic perturbation. Neural Computing and Applications 24(2):309–315

Vernon MC, Keeling MJ (2009) Representing the uk's cattle herd as static and dynamic networks. Proceedings of the Royal Society B: Biological Sciences 276(1656):469–476

Wilcoxon F (1945) Individual comparisons by ranking methods. Biometrics Bulletin 1(6):80–83

Witbooi PJ, Muller GE, Van Schalkwyk GJ (2015) Vaccination Control in a Stochastic SVIR Epidemic Model. Computational and Mathematical Methods in Medicine 2015:271654

Yu Z, Liu J, Wang X, Zhu X, Wang D, Han G (2016) Efficient Vaccine Distribution Based on a Hybrid

Compartmental Model. PloS one 11(5):e0155416–e0155416

Zavala GR, Nebro AJ, Luna F, Coello Coello CA (2014) A survey of multi-objective metaheuristics applied to structural optimization. Structural and Multidisciplinary Optimization 49(4):537–558

Zhang Q, Li H (2007) MOEA/D: A multiobjective evolutionary algorithm based on decomposition. IEEE Transactions on Evolutionary Computation 11(6):712–731

Zitzler E, Laumanns M, Thiele L (2002a) SPEA2: Improving the strength pareto evolutionary algorithm for multiobjective optimization. In: Giannakoglou K, et al. (eds) Evolutionary Methods for Design, Optimisation and Control with Application to Industrial Problems (EUROGEN 2001), International Center for Numerical Methods in Engineering (CIMNE), pp 95–100

Zitzler E, Thiele L, Laumanns M, Fonseca CM, da Fonseca VG (2002b) Performance assessment of multiobjective optimizers: An analysis and review. IEEE Transactions on Evolutionary Computation 7:117–132