A multi-objective optimization approach to the location of road weather information system in New York State

Julie Fetzer, Hernan Caceres, Qing He & Rajan Batta

To cite this article: Julie Fetzer, Hernan Caceres, Qing He & Rajan Batta (2018): A multi-objective optimization approach to the location of road weather information system in New York State, Journal of Intelligent Transportation Systems, DOI: 10.1080/15472450.2018.1439389

To link to this article: https://doi.org/10.1080/15472450.2018.1439389

Accepted author version posted online: 12 Feb 2018.
Published online: 01 Mar 2018.

Submit your article to this journal

Article views: 28

View related articles

View Crossmark data
A multi-objective optimization approach to the location of road weather information system in New York State

Julie Fetzer\textsuperscript{a,b}, Hernan Caceres\textsuperscript{b,c}, Qing He\textsuperscript{a,b}, and Rajan Batta\textsuperscript{b}

\textsuperscript{a}Department of Civil, Structural and Environmental Engineering, The State University of New York, Buffalo, NY, USA; \textsuperscript{b}Department of Industrial and Systems Engineering, The State University of New York, Buffalo, NY, USA; \textsuperscript{c}Department of Industrial Engineering, Universidad Catolica del Norte, Antofagasta, Chile

\section*{ABSTRACT}
In the paper, we present a practical and robust approach to solving the location problem of road weather information systems (RWIS) for New York State. We use a multi-objective optimization methodology to find the optimal deployment of environmental sensor stations (ESS) in the existing road network. This approach takes into account various factors such as vehicular accident data, vehicle-mile traveled, area coverage, accessibility for power and maintenance, and existing ESS. The study produces a Pareto set of multiple efficient solutions that can guide the decision-making process for RWIS deployment. We also conduct a sensitivity analysis to examine the effects of different parameters and propose a non-preference solution.

\section*{Introduction}
A road weather information system (RWIS) is a web of meteorological sensors located along the highway network. An RWIS is comprised of strategically located environmental sensor stations (ESS) that collect a range of real-time data about weather and pavement conditions. This data can include visibility, precipitation, water level, humidity, temperature, and status of the pavement (Branke, Deb, & Miettinen, 2008). A central station then models and interprets the road and weather data collected by the ESS and transmits the information to road maintenance facilities, emergency operation centers, other highway authorities, and in some cases the general public. This road weather data can then be used to support the decision making of highway personnel in times of inclement weather, including snowstorms, rainstorms, and hurricanes. The information collected can indicate when and where to deploy snowplows, deposit road salt, close highways, and warn motorists of dangerous road conditions. For highway personnel, data provided by an RWIS provides a more timely and efficient system of decision making in critical times. This leads to minimal traffic inconveniences and delays caused by severe weather and safer roadway conditions for the traveling public.

The locations of the ESS matter greatly when implementing an RWIS. ESS should be located at sites that optimize their potential for collecting meaningful meteorological and pavement data about the network. In a modern urban transportation network, weather stations and systems of decision making by highway authorities during extreme weather already exist. Because of this, the locations of the ESS should focus on what is lacking in the current system in order to optimize the investment of implementing an RWIS. For example, ESS should be placed by areas in the network most affected by traffic collisions during severe weather. However, several other factors must also be taken into account when choosing the locations of ESS, such as the sites of pre-existing meteorological stations, accessibility for power and maintenance, and the physical characteristics of the sites. Additionally, different

\section*{ARTICLE HISTORY}
Received 14 November 2016 Revised 7 February 2018 Accepted 7 February 2018

\section*{KEYWORDS}
inclement weather; road weather information system; environmental sensor stations; location problem; multi-objective optimization
configurations of ESS exist, which vary their cost and the range of their data detection. These elements contribute to making the strategic location of ESS within a network difficult for highway authorities, but essential in maximizing the benefit of implementing an RWIS. The key contributions of this paper lie on:

- Introduce a multi-objective optimization approach for the location problem of ESS within an RWIS network and implement the model across the entire state of New York.
- Consider previously installed ESS stations that already exist in the RWIS network in the optimization model.
- Take into account different types of configurations for the ESS in the RWIS network.
- Consider access to power and maintenance when selecting candidate ESS sites.
- Develop an exact solution method to solve the optimization formulation and derive the Pareto frontier for three competing objectives.
- Investigate the effect of budget upon different objectives so that the implementation strategy can be determined accordingly by transportation agencies.

**Literature review**

Although factors such as visibility, precipitation intensity, and wind speed all contribute to weather-related accidents, road surface conditions—namely surface friction—is found to be the most influential factor affecting safety. Usman et al. (2012) found that a 1% improvement in road surface conditions from the mean value, as interpreted by a measure likened to surface friction, would cause almost a 2% reduction in the mean number of accidents. Because of this, it is vital that winter road maintenance operations preserve pavement conditions during and after inclement weather in order to maintain safety in the network (Abdel-Aty, Pande, Lee, Gayah, & Santos, 2007). Many studies show weather has an influence on the number of crashes (Walker, 2012, Wierzbicki, 1999), especially in areas with recurrent inclement weather such as thunderstorms and snowfall (Yu, Hu, & Chang, 2015, Zhang & Chen, 2017, Zolfaghari, Jaber, & Azizi, 2002), which is the case of New York State (Zhao, Chien, Meegoda, Luo, & Liu, 2015). RWIS are useful in this way, providing agencies localized information in order to more quickly and cost-effectively indicate when and where to deploy snowplows, deposit road salt, close highways, and warn motorists of dangerous road conditions. Because of this, RWIS have become a part of many transportation agencies’ efforts to maintain safety and mobility despite adverse weather (Abounacer, Rekik, & Renaud, 2014). Currently, the most widely used resource for the placement of ESS is a guide published by Federal Highway Administration (Andrey, 2010). Although the guidelines take into account several physical criteria in order to choose an ESS site (close proximity to the roadway, avoid standing water, etc.), they do not adequately describe how to choose optimal locations beyond the physical criteria of the sites.

Eriksson and Normman (2001) adjusted a linear regression model to predict the number of occasions where slipperiness in the road is observed to aid in the location problem of RWIS (Caceres, Hwang, & He, 2016). For this, they classified weather observations (precipitation, temperature, humidity, etc.) from existing stations into types of slipperiness and fitted the coordinates, elevation, distance from water and topography of the location of the stations. The model is then used to predict the number of occurrences of slipperiness per month of potential sites for new stations along a section of a road. They performed a case study of their work on Sweden. However, their model does not consider traffic or accident data, and the task of selecting sites for new stations is not proposed in their work. Future work can use their results as input data, along with other such traffic and cost, for an optimization model (Caceres et al., 2016). Similarly, Vlahogianni et al. (2012) proposed a neural network approach to modeling the secondary accident likelihood where they consider traffic and weather data, but rely on given locations of the detectors of the incidents (Manfredi et al., 2008). In 2012, a thesis published out of University of Texas at Austin was able to address vehicular crash data and existing meteorological information sites when optimizing ESS site locations. This model, however, did not take into account factors such as crash severity or different configurations of ESS. The case study that the model was applied to also only considered one corridor of a network in Austin, Texas (Eisenberg, & Warner, 2005). Jin et al. (2014) expanded on this methodology and also focused on the Austin District in Texas for their RWIS location model, which used a discrete network representation of the roads. Each link in the network was divided into equal length segments, and a formulation of the optimization problem for choosing the location (segment in the network) of new stations was proposed. However, their model only had one objective function: to minimize the sum of safety concern index (which is defined for each segment as the product of crash rate and a reduction factor). Their solution method was based on a greedy heuristic, and no optimality gap was offered. Additionally, the approach did not have consideration for cost or existing stations (Eriksson, & Norrman, 2001). A two-stage sequential model was developed by Ewan, Al-Kaisy, and Veneziano, (2013). The Stage-1 model maximizes spatial coverage with the least overlapping area and Stage-2 model maximizes the life cycle profit subject to limited
budget. However, this model did not consider several factors such as traffic operations, traffic safety data, and different configurations of ESS.

Recently, Singh et al. 2016 proposed a methodology to isolate the effect of the weather from the crash rate by subtracting the rate corresponding to accidents that occurred during good weather conditions, also using Austin as a case study (Farahani, SteadieSeifi, & Asgari, 2010). However, the proposed optimization method contained only one objective and did not support multiple types of ESS. It contained nonlinear expressions in the objective and constraints. The model was solved approximately with a heuristic and no information about the optimality gap was provided. No details were provided regarding the implementation of the heuristic solution. Kwon and Fu (2013) proposed a framework to choose the location of stations based on a scoring method for the providence of Ontario, Canada (Gopalakrishna, Martin, & Neuner, 2016). They divided the geographical region into a grid and assigned a score to each cell based on a combination of weather and traffic information. In order to determine the factor to be included, those that are correlated are discarded in order to avoid bias to those factors. A way to determine weight when combining the factors was not provided, which reflects the difficulties of trying to develop methods that are based on one objective. Trying to set weight in factor before the optimization based on preference or importance can lead to inferior solutions due to the difficulty of translating subjective preference of decision maker into numeric value to assign weights (Gopalakrishna et al., 2016). They expanded on their work with Melles using a case study in Minnesota. Here their methodology aims to minimize the hazardous road surface conditions while maximizing the coverage of accident-prone areas with ESS locations. However, their work does not consider variables such as annual average daily traffic (AADT) and offers room for expansion with regards to a multi-objective solution. Additionally, limited access to power and maintenance is addressed as a restriction, but not included in the case study (Jin, Walker, Cebelak, & Walton, 2014).

Since the introduction of multi-criteria decision making to management sciences, this concept has been implemented in location and transportation problems (Maze, Agarwai, & Burchett, 2006, Olia, Abdelgawad, Abdulhai, & Razavi, 2017, Singh, Li, Murphy, & Walton, 2016, Ulungu & Teghem, 1994, Usman, Fu, & Miranda-Moreno, 2012, Vlahogianni, Karlaftis, & Orfanou, 2012). In an application related to ours, Olia et al. (2017) and Mavrotas (2009) proposed a multi-objective formulation to the problem of deciding the number and location of roadside equipment (RSE) unit for travel time estimation in vehicle-to-infrastructure and vehicle-to-vehicle communication environments. Their model minimizes two conflicting objectives: travel time estimation error rate and the number of RSE units. To find the Pareto frontier, their solution strategy relied on the use of genetic algorithms. The benefits of installing RSE units depends on the market penetration of vehicles with the technology to communicate with one another and with the infrastructure, and to understand its influence, Olia et al. performed a sensitivity analysis to find a Pareto solution set for each of 6 levels of market penetration.

The methodology proposed in this paper is applied to a case study of the implementation of an expanded road weather information system in New York State (NYS). Therefore, an effective approach for the deployment of RWIS that considers several important factors such as vehicular accidents, traffic volume, area coverage, existing sites, and access to power is still needed.

**Methodology**

**Model description**

To model this problem, the geographical region in question was divided into a square grid with an arbitrary cell size based on the scope of the region and data available. This grid divided the network into cells where each cell was associated with the information of any roadway segments it contained, namely the summation of the vehicle miles traveled (VMT) of the road segments, the amount of vehicular crashes occurred within the cell during a set time period, whether or not any existing ESS was present, and any access points to power. The factors, discussed with state agencies, were chosen related to the availability of data needed to perform realistic numerical experiments. Further, the proposed mathematical model can easily accommodate other factors after minor modifications. Only the limited number of cells in the network that contained access to power and maintenance due to an existing power source were considered potential locations for ESS to be installed in this model. Any cell that was suitable for the installation of a new station due to their proximity to power is called a potential location going forward. Such a grid over a mock region is shown in Figure 1, the gray area represents the region, the darker square represents potential locations, and the numbered location shows the location of 4 existing stations.

The cell size was determined in part by the proportion of cells that contained key information relating to the model. For example, a very high percentage of cells that did not contain any pertinent information (vehicular crashes, access to power and communications, etc.) would add excessive computation time to the model. Additionally, it would needlessly divide the network into cells that
have a smaller area than is of real world importance to statewide agencies. In other words, when looking at a large state-wide network like in the case study presented, it is not of practical importance to narrow down the potential solution locations to an area as small as a square half mile. Rather, if the model can suggest cells of a reasonable but larger area for locations for ESS, agency officials can decide exactly where within that cell it is practically feasible to place the ESS. On the opposite end of the spectrum, the cell size should not be so large that key location information is muddled together geographically. For example, cells should not be so large that many contain multiple access points to power and communications, or that agencies would be overwhelmed with the cell size of the potential solutions.

The model aims to solve the location problem of which potential locations should the ESS be deployed and which ESS configuration should be chosen for each location given a certain budget and the existing traffic and crash conditions of the region.

The multi-objective model aims to concurrently maximize three objectives: Eq. 1, the VMT within the stations’ ranges, Eq. 2, the geographic area covered by the station ranges, and Eq. 3, the potential increase in safety, namely the crash rate reduction. These objectives were constrained by ESS only being able to be placed at a potential location (those locations that have access to power), a limited budget intended for the acquisition of new stations, the requirement that only one new station can be installed per location and at least a certain distance away from another station, and existing ESS in the network.

**Mathematical formulation**

Let $A$ be the set of all cells in the grid, then:

- $L \subseteq A$, the set of cells corresponding to both potential locations and existing stations;
- $E \subseteq L$, the set of cells with existing stations; and
- $S$, the set of alternative configuration types for new stations.

In the model, the following attributes of each cell are used:

- $V_i$, number of weather related vehicular accidents occurred within cell $i \in A$;
- $A_i$, the maximum AADT of links within cell $i \in A$; and
- $l_i$, the sum of link lengths (ft).

Known parameters are:

- $t$, time period for the crash data collection (years);
- $r_{ik}$, information radius reached at location $i \in L$ by configuration type $k \in S$ (ft);
- $C_{ik}$, cost of installing at location $i \in L$ a station type $k \in S$; and
- $\delta$, the minimum separation distance for new stations.

The following is calculated using the parameters above:

- $d_{ij}$, VMT in potential location $i \in L$ (miles);
- $Q_k = \sum_{j \in L, d_{ij} \leq \delta} q_j$, VMT reached by station with configuration type $k \in S$ installed potential location $i \in L$ (miles);
- $O_{ij}$, the overlapping area of sites $i \in L$ and $j \in L$ (ft$^2$);
- $\alpha_{ij} = 1 - e^{-\lambda d_{ij}}$, a reduction factor based on the distance from location $i \in L$ to cell $j \in A$, where $\lambda$ is constant and $d_{ij}$ is distance from the centroid of potential location $i \in L$ to the centroid of cell $j \in A$;
- $\beta_{ij} = \begin{cases} 1 & d_{ij} < \tau \vspace{3pt} \\ 0 & d_{ij} \geq \tau \end{cases}$, the truncated complement of $\alpha_{ij}$, where $\tau$ is the radius up to which the reduction factor applies; and
- $R_i = \frac{V_i \times 10^5}{365 \times 4 \times 1 \times 365 \times 1}$, the crash rate of cell $i \in A$.

These binary variables are used in the model:

- $\delta_{ik}$, binary variable that equals 1 when at a potential location $i \in L$ a sensor type $k \in S$ is installed;
- $\gamma_{ij}$, binary variable that equals 1 when locations $i \in L$ and $j \in L$ are used with sensor type $k_i \in S$ and $k_j \in S$ respectively; and
- $\omega_{ij}$, binary variable that equals 1 when cell $i \in A$ has the maximum reduction in crash rate due to sensing installed at location $j \in L$.

The multi-objective model used to solve this problem aims to concurrently maximize the following three objectives:

- $f_1 \Delta$ VMT within the stations’ ranges;
- $f_2 \Delta$ geographic area covered by the station ranges; and
- $f_3 \Delta$ the potential increase in safety (crash rate reduction).

These objectives were constrained by existing ESS in the network, access to power and maintenance, and:

- $B$, a limited budget intended for the acquisition of new stations;
• \( \sum_{k \in S} x_{ik} \leq 1 \), only one new station being installed per potential location; and
• \( \sum_{k \in S} (x_{ik} + x_{jk}) \leq 1 \), each station being at least a certain distance away from another station where \((i, j) \in L^2: (i, j) \notin E^2\) and \(j < i, d_{ij} < \delta\).

Then, the mathematical formulation is as follows:

\[
\text{max } f_1 = \sum_{i \in L} \sum_{k \in S} x_{ik} Q_{ik} \quad (1)
\]

\[
\text{max } f_2 = \sum_{i \in L} \sum_{k \in S} x_{ik} \pi r_{ik}^2 - \sum_{i \in L} \sum_{k \in S} \sum_{j \in L: k \in S} z_{ij} C_{ijk} \quad (2)
\]

\[
\text{max } f_3 = \sum_{i \in L} R_i \sum_{j \in L} \beta_{ij} w_{ij} \quad (3)
\]

\[
\sum_{i \in L} C_{ik} x_{ik} \leq B \quad (4)
\]

\[
\sum_{i \in L} x_{ik} \leq 1 \quad i \in L \quad (5)
\]

\[
z_{ij} \geq x_{ik} + x_{jk} - 1 \quad i \in L, k_i \in S, j \in L, k_j \in S, j < i, d_{ij} \geq \delta \quad (6)
\]

\[
z_{ij} \leq \frac{x_{ik} + x_{jk}}{2} \quad i \in L, k_i \in S, j \in L, k_j \in S, j < i, d_{ij} \geq \delta \quad (7)
\]

\[
\sum_{i \in L} w_{ij} \leq |A| \sum_{k \in S} x_{jk} \quad j \in L \quad (8)
\]

\[
\sum_{j \in L} w_{ij} \leq 1 \quad i \in A \quad (9)
\]

\[
\sum_{k \in S} (x_{ik} + x_{jk}) \leq 1 \quad (i, j) \in L^2: (i, j) \notin E^2, j < i, d_{ij} < \delta \quad (10)
\]

\[
x_0 = 1 \quad i \in E \quad (11)
\]

\[
x_{ik} \in \{0, 1\} \quad i \in L, k \in S \quad (12)
\]

\[
w_{ij} \in \{0, 1\} \quad i \in A, j \in L \quad (13)
\]

\[
z_{ij} \in \{0, 1\} \quad i \in L, k_i \in S, j \in L, k_j \in S \quad (14)
\]

Where Eq. 1—3 are the objectives functions previously defined. Constraint 4 enforces the budget for the acquisition of new stations. Constraint 5 limits the installation of new stations to one per potential location. Constraints 6–7 are the linear inequalities that enforce the nonlinear relation \( z_{ij} = x_{ik}, x_{jk} \). Constraint 8 limits the association of cells with only those locations with installed stations. Constraint 9 limits the association of a cell to up to one location. Constraint 10 ensures that the distance from a new station to any other station is at least \( \delta \) units. Constraint 11 represents the existing stations, and 12–14 are the binary constraints.

Solution method: Modified \( \epsilon \)-constraint method

Consider the multi-objective optimization program

\[
\text{max}_{x \in S} \{ f_1(x), \ldots, f_p(x) \},
\]

where \( x \) is the vector of decision variables, \( f_i(x) \) are the objective functions with \( k = 1, \ldots, p \), and \( S \) the feasible region. Then, the modified \( \epsilon \)-constraint method solves \( \text{max } f_1(x) : f_2(x) \geq e_2, \ldots, f_p(x) \geq e_p, x \in S \) iteratively by parametrical variation of \( e_k \), obtaining efficient solutions for the problem at hand.

Mavrotas (2009) (Khattak, Kantor, & Council, 1998) recognizes several advantages of the \( \epsilon \)-constraint method over the weighting method for solving multi-objective optimization problems. However, this work also identifies three areas where the implementation of the \( \epsilon \)-constraint method can be improved.

First, to better calculate the range of the objective functions over the efficient set, Mavrotas (2009) proposes a lexicographic optimization for building the payoff table. Second, to guarantee obtaining only efficient solutions, he proposes a modified formulation of the modified \( \epsilon \)-constraint that reads as follows:

\[
\text{max } f_1(x) + \epsilon (s_2 + \cdots + s_p) : f_2(x) - x_2 = e_2, \ldots, f_p(x) - s_p = e_p, s_k \geq 0, x \in S \quad (15)
\]

Finally, Mavrotas added to the algorithm an early exit from the nested loop when the problem becomes infeasible. This addition significantly accelerates the algorithm especially in the case of several (more than three) objective functions, since it avoids the expiration of solution in infeasible regions.

To implement problem 4–14 using Eq. 15 results in a mixed integer program, which needs to be solved with branch and bound. Given its computational complexity, to find an optimal solution iteratively is time-consuming. Then, to make the method time efficient, the requirement of solving it optimally needs to be relaxed. However, no optimal solutions of Eq. 15 may divert significantly from the defined grid points.

For our implementation, we modified the \( \epsilon \)-constraint method taking elements from the work of Mavrotas. We first obtain the ranges of the objective functions with the lexicographic optimization for building the payoff table. The value for the \( e_k \) parameters is assigned by dividing the ranges of the each objective function \( f_k(x) \) \( k = 2, \ldots, p \) into \( g_k \) equal intervals, respectively. As a result, we obtain a grid consisting of \( (g_2 + 1) \times \cdots \times (g_p + 1) \) points. For each of the grid points we solve a modification of Eq. 15 defined as follows:

\[
\text{Problem P max } 100 \times \frac{f_1(x) - lb_1}{r_1} - \frac{1}{p - 1} \sum_{k \in Z} s_k \quad (16)
\]

\[ r_k = \frac{\sum_{k \in S} \sum_{j \in L} z_{ij} C_{ijk}}{\sum_{i \in L} \sum_{k \in S} x_{ik} Q_{ik}} \]
Recall that we are not necessarily looking for optimal solutions. The benefits of this modification are twofold. On the one hand, allowing the exploration in the left side of \( e_k \) makes it easier to find feasible solutions given the discrete nature of the solution space of our application. On the other hand, limiting the exploration of solutions around \( e_k \) to a fraction of the range of the interval prevents finding solutions that significantly depart from the grid points.

**Case study: New York State**

**Data description**

A dataset from the NYSDOT was utilized containing vehicular crash information from all incidents recorded in NYS during the two-year period of June 2012 until May 2014. This dataset included, among other data, the coordinate location of the crash and weather at the time of the incident. In our model, all types of vehicular accidents were included (collisions with other vehicles, pedestrians, fixed objects, etc.). However, only accidents occurring during inclement weather were considered, that being rain, snow, sleet/hail/freezing rain, or fog/smog/smoke. In total, 122,200 crashes occurring during inclement weather in NYS over a two-year period were considered.

Figure 3 displays the vehicular crash data as the amount of crashes due to a given type of weather throughout the two-year period. This figure shows the increase in vehicular crashes occurring during snow and sleet, freezing rain, or hail in the winter months. The geographical distribution of accidents across the state is shown as a gradient in Figure 4a.

In this case study, 186 sites of continuous count stations (with known access to power and maintenance) were considered potential locations. 32 existing ESS were also considered. These are shown in Figure 4b.
We considered two configuration types of stations, having their cost $c_k$ and effective radius $r_k$ of the type-1 station is $50,000 and 5 miles, respectively, and for the type-2 is $100,000 and 15 miles, respectively. The total budget for the acquisition of stations $B$ is one-and-a-half million dollars. The minimum distance between stations $\delta$ is 10 miles. The parameter $\lambda$ in the crash rate reduction factor equals 0.2.

A cell size around $2.0 \times 2.0$ miles was chosen for this particular case study because it balanced the percentage of cells containing information with the number containing more than one access point, while maintaining a practical implementation size for agencies. For this case, approximately 70% of the cells contained information while 30% were “empty,” and only three potential locations cells contained more than one communication and power access point, while 170 cells only contained one.

**Model implementation**

The multi-objective problem contains a total of 78,378 binary variables and 11,079 constraints. For the implementation, we coded the algorithm for the modified $e$-constraint method with Java 8 using CPLEX 12.6.3 to solve the mixed integer program, for which we considered an optimality gap of 5%. We further deployed the code in a university cluster. Each computing node had 12 CPUs (cores) with 128 GB of RAM, and the CPUs were Intel Xeon E5-2620 at 2.40 GHz. The operating system used was Linux CentOS 7. With this configuration, an instance with budget of 500 k took 48 minutes to solve. The parameters for the modified $e$-constraint method consider the maximization of $p = 3$ objectives, having the corresponding grid points $g_2 = g_3 = 5$, and $\rho = 5\%$.

There are 36 grid points from where the method yielded 29 efficient solutions, having 7 infeasible points. **Figure 5** shows the value of the objectives functions for each of the 29 solutions found. In the horizontal axis is objective 3, the reduction in crash rate, in the vertical axis is objective 2, the covered area, and represented as the radius of the bubbles is objective 1, the VMT. For illustrative purposes, we selected 5 solutions from the set for display of the solutions in the map. Among these five solutions we selected three extreme solutions, labeled in **Figure 5** as E1, E2, and E3, having each the maximum value found for the corresponding objective functions: VMT, covered area, and reduction in crash rate. The other two solutions were selected in the middle ground among the Pareto set, and they are labeled as S1 and S2 in **Figure 5**.

**Figure 6** shows the selected solutions. In each map, we see the potential locations, the locations of the existing stations, and the selected locations for the new stations. Notable or unintuitive parts of the results include the smaller radius ESS configuration being produced for all selected locations for the crash rate objective. This is because the reduction factor of the crash rate is a function of the distance to the closest ESS, and it does not depend...
on the radius associated with such station. Additionally, in some instances, two points were chosen around an area with a high crash rate density. This can be explained by recalling that the reduction factor of the crash rate is a function of the distance to the closest ESS. Then, the reduction factor will not change by the presence of another ESS that is further away. Therefore, by putting two points around an area of a high crash rate density instead of one, we can cover not only that area but also more surrounding territory. Additionally, it should be noted that in the results for the area objective, stations in the Adirondacks (located in northern NY) were not chosen because the solution obtained from maximizing the covered area is not unique. Provided that the area covered by a station does not overlap with one from any other, their locations can be at any place. However, the method we use to solve the multi-objective problem considers lexicographic optimization, in which the objectives can be ranked in the order of importance. Therefore, the other objective still influences the result of the selected locations.

**Sensitivity analysis**

The available budget is a constraint in our multi-objective model. Then, in order to find the effect of its variation over the objectives, we need to run the algorithm for each level of the budget we wish to study. We set four levels for the budget: 0.5, 1.0, 1.5, and $2.0 million.

Figure 7 shows the collective result for each budget level. To compare the effect of the budget, all four graphs...
Figure 7. Solution set per budget level: (a) Objectives vs. budget of 500 k, (b) objectives vs. budget of $1 million, (c) objectives vs. budget of $1.5 million, and (d) objectives vs. budget of $2 million.

have the same scale. Each graph displays objective 1, the reduction of crash rate, in the horizontal axis; objective 2, the area covered by all sensors, in the vertical axis; and objective 3, the VMT, represented proportionally by the size of each bubble. The same information is portrayed differently in Figure 8. Here, each graph shows the possible values of the objective function against four budget levels, and the blue horizontal line represents the baseline value of each objective considering only the existing RWIS sites in NY State. It can be seen how the budget significantly affects the covered area. However, the magnitude of the change in the other two objectives is not as substantial. For example, for the solutions that maximize the reduction in crash rate, those to the right of the solution set, we can see that the marginal increment of going from a budget of 500 k to 1 m is significantly less to that of going from 1.5 m to 2 m. The same can be said for objective 1, maximizing VMT. The VMT is

Figure 8. Individual objectives vs. Budget: (a) $f_1$: vehicle miles traveled, (b) $f_2$: covered area, and (c) $f_3$: reduction in crash rate
Figure 9. Individual objectives vs. minimum distance between stations: (a) $f_1$: vehicle miles traveled, (b) $f_2$: covered area, and (c) $f_3$: reduction in crash rate.

represented as the radius of the bubble, and we can see that their marginal growth decreases when the budget increases.

Aside from the budget, we need to understand the effect on the objective the other parameters of the model have. The minimum distance between new stations $\delta$ was set to 10 miles. We now vary this value to see how it affects the solution generated. We choose four level combinations for this parameter $\delta \in \{5, 10, 15, 20\} \forall i$. The results are illustrated in Figure 9. The parameters $r_{ik}$, the information radius of a type-$k$ station at location $i$, were set as follows: $r_{i1} = 5$ miles $\forall i$ and $r_{i2} = 15$ miles $\forall i$. We now vary these values to see how they affect the solution generated. We again choose four level combinations for these parameters $(r_{i1}, r_{i2}) \in \{(5, 10), (5, 15), (10, 15), (10, 20)\} \forall i$. Results are shown in Figure 10. The parameters $C_{ik}$, the cost of installing a type-$k$ station at location $i$, were set as follows: $C_{i1} = $50,000 $\forall i$ and $C_{i2} = $100,000 $\forall i$. We now vary these values to see how they affect the solution generated. We choose four level combinations for these parameters $(C_{i1}, C_{i2}) \in \{(25, 75), (50, 100), (75, 125), (100, 150)\} \forall i$. Results are depicted in Figure 11.

In the implementation of our modified $\varepsilon$-constraint algorithm, we vary three parameters that affect computational time and the number of Pareto solutions that can be found: the grid point ($g$), the permissible deviation from any grid point ($\rho$), and the relative optimality gap in CPLEX. We performed a full factorial experiment where for each factor we arbitrarily chose three levels: $g = 3, 5, 7$; $\rho = 0.5\%, 1\%, 5\%$; and $\text{gap} = 1\%, 5\%, 10\%$.

In Figure 12 we observe the results of the factorial experiment. We studied how the computational time and the number of Pareto solution are affected by the parameters of our modified $\varepsilon$-constraint implementation. For the computational time, we considered the time needed to create the payoff table (CPU Time 1) and the time needed
to solve the problem at all the grid points (CPU Time 2). We can see the results for these in the first two rows of Figure 1. The number of Pareto solutions is shown in the third row of the same figure. The level of $g$ is displayed in the legend; the levels of $\rho$ correspond to the horizontal axis of each individual chart; and the three levels of the optimality gap are represented on each of the three columns of the figure as the panel variable.

The biggest effect is for the value of $g$. More grid intervals result in higher CPU Time 2 and number of solutions. Also, smaller optimality gap leads to higher CPU Time 2. Notethe higher grid size always results in higher CPU Time 2. In addition, when $\rho$ increases the CPU Time 2 decreases and the number of solutions found increases.

It is found that The CPU Time 1 increases when the optimality gap increases. It is due to the way how the lexicographic optimization performs when building the payoff table. Prior to maximize an objective, we add the result from the previous one as a constraint; if the previous objective is not solved optimally (with a gap), then the feasible region of the remainder objectives is bigger, which consequently increases the time it takes to explore the solution set. However, note that CPU Time 1 is very small in relation to CPU Time 2.

A no-preference method

Since no explicit preferences from the decision maker (DM) are available we developed a procedure to generate an approximation of the Pareto Frontier. Thus, the DM can later choose the most preferred solution from the Pareto set. Alternatively, we can generate a single solution without the involvement of the DM by using the Neutral Compromise Solution suggested by Wierzbicki (1999)

Figure 11. Individual objectives vs. cost combination: (a) $f_1$: vehicle miles traveled, (b) $f_2$: covered area, and (c) $f_3$: reduction in crash rate.

Figure 12. Results for the factorial experiment on the parameters of the e-constraint algorithm.
and Kwon and Fu (2013). To that end, we solve the following problem

$$\min \max_{i=1,...,k} \left[ f_i(x) - \frac{z^*_i + z^*_n}{2} \right]: x \in S$$

where $z^*_i$ is the ideal objective vector, obtained by maximizing each objective individually, $z^*_n$ is the nadir objective vector, obtained by taken the minimum value of the corresponding objective function from the payoff table, and $z^*_u = z^*_i + \varepsilon$ is the utopian objective vector with $\varepsilon$ being a small positive scalar.

Figure 13 displays the Pareto solution and also includes the no-preference solution marked with “NP.” The no-preference solution here represents a projection of the ideal point (the maximum value of each objective) onto the Pareto frontier, and as such it is not located far from the selected reasonable solutions S1 and S2. Therefore, “NP” corresponds to a rather equivalent compromise for all objectives, appropriate when the decision maker’s preferences are not present.

**Discussion of the practical implementation**

The analysis of the effect of the budget over the objectives shows that for the objectives the returns to scale over the budget are not necessarily constant. For the cases of the objective vehicle miles traveled and reduction in crash rate, the return to scale are decreasing, whereas for the covered area the return to the scale of the budget appears to be constant. This means that the addition of budget does not generate the same benefits at higher levels as when it does at lower levels. Thus, there is a level after which the addition of budget, namely a new station, won’t have a significant effect on the objectives that we aim to optimize.

In our paper, we present a model and solution method that yields multiple sets of Pareto solutions from which the decision maker can choose based on their preference. Hence, up to this point, we have assumed the preferences of the decision maker are unknown. Moreover, we only know what the objectives should be. We have not placed any weight nor tried to unify the objectives into one. This prevents making an assumption on the unrealistic preference function of the decision maker. Actually, the preference function is in fact many times even unknown to the decision maker. Recall multi-objective optimization has a main task: optimization for finding Pareto optimal solutions and a second task for the decision maker of choosing a single-most preferred solution. For both tasks, the literature offers different approaches, and in this paper focus is on the first task by modifying the $\epsilon$-constraint method and applying it to the RWIS location problem.

Also, when implementing the proposed method, one has to be aware of local implementation constraints, including equipment, human resource, power, communication, etc. Those practical constraints could be easily incorporated into the proposed formulation. Depending on the needs of the local region, transportation agencies may consider using different configurations of RWIS systems, which can be addressed with minor modifications in proposed model.

**Conclusions**

Selecting optimal locations to implement ESS as part of an RWIS is crucial in maintaining safety and efficiency in times of inclement weather. The locations in our application should be chosen such that the coverage of high traffic roads, the geographic area covered, and the increase in safety are the preferred criteria. Usually, problems of this nature are solved using a single objective optimization model. However, the different objectives are noncommensurable, so it is difficult to aggregate them into one synthetic objective. For example, in a sum of weighted objectives we need to make a choice of weights that reflect the importance of each objective, but it may happen that some Pareto optimal solutions cannot be found no matter how the weights are selected (Kwon, Fu, & Melles, 2016). In this study, we treated the problem as a true multi-objective problem, and we did not make any prior assumptions about the preferences of the decision maker. Additionally, since such preferences might change based on the implementation scenario, there is an underlying need to have a flexible model at hand. The use of a multi-objective model yields a set of efficient solutions from where, depending on the preferences, the decision maker can choose the most suitable. This is of particular importance because the decision maker is often required to pick the best scenario based on a set of variable circumstances.

This study introduces unified multi-objective optimization model for locating RWIS sites in a state level
region. The proposed model is able to optimize multi-objectives concurrently. Also, the proposed formulation can provide alternative Pareto solutions for multi-criteria decision making of the RWIS locations. This paper provides transportation agencies a tool to optimally determine the location of RWIS sites. An exact solution method was developed to solve the model formulation. The set of Pareto solutions can be retrieved within an hour of computation time. A case study of New York State shows that the proposed model can optimally add new RWIS sites into the existing network.

Future research should focus on developing a model that considers a multi-period budget, where the solution is a plan implantation over a time horizon. The multi-period model should also include projected traffic flows and the number of accident due to planned road closing for maintenance, availability of new road, or change in drivers’ behavior. Additionally, the severity of the crash was not considered in the model presented, and it is unknown if the vehicle crashes were directly triggered by the inclement weather. There should also be further research on measuring the effective radius of RWIS stations, and what is the appropriate rate of depletion of such effectiveness (i.e., how to estimate the reduction factor).

Acknowledgement

The authors would like to thank Charlie Morgante, Michael Lashmet, and Joe Doherty, from New York State Department of Transportation, who provided the data and useful discussions for this study.

References


Walker, A. J. (2012). Optimization of site locations for a road weather information system in Austin, Texas based on inclement weather crashes (MS Thesis). University of Texas at Austin.


