Decision Support Systems 96 (2017) 67–82

Contents lists available at ScienceDirect

Decision Support Systems

journal homepage: www.elsevier.com/locate/dss

Aircraft re-routing optimization and performance assessment under uncertainty

Xiaoge Zhang, Sankaran Mahadevan*

Department of Civil and Environmental Engineering, Vanderbilt University, Nashville, TN 37235, USA

ARTICLE INFO

Article history:
Received 8 May 2016
Received in revised form 14 December 2016
Accepted 13 February 2017
Available online 17 February 2017

Keywords:
Air traffic control
Uncertainty quantification
Support vector regression
Censored data
Reliability analysis
Performance assessment

ABSTRACT

The need for aircraft re-routing arises when there is disruption in the system, such as when an airport is closed due to extreme weather. In this paper, we investigate a simulation-based approach to optimize the aircraft re-routing process, by considering multiple sources of uncertainty. The proposed approach has four main components: system simulation, uncertainty representation, aircraft re-routing algorithm, and system performance assessment. Several sources of uncertainty are accounted for in this approach, related to incoming aircraft, space availability in neighboring airports, radar performance, and communication delays. An aircraft re-routing optimization model is formulated to make periodic re-routing decisions with the objective of minimizing the overall distance travelled by all the aircraft, subject to the system resources. We analyze the performance of this aircraft re-routing system using system failure time as the metric. Since the simulation time is limited, right-censored data arises with respect to system failure time. A novel methodology is developed to compute the lower bound of system failure time in the presence of right-censored data, and to analyze the sensitivity of the system performance metric to the uncertain variables relating to the aircraft, radars, nearby airports, and communication system. Since the simulation is time-consuming, we build a Support Vector Regression (SVR) surrogate model to efficiently construct the system failure time distribution.

1. Introduction

Airport operations face many uncertain factors [1–3] related to bad weather, departure delay, arrival delay, ground delay, mechanical problems, etc. For example, in February 2014, a heavy storm disrupted air travel in the New York region [4]. A large number of outgoing flights were delayed or cancelled due to severe weather, while many incoming flights were diverted to nearby airports due to airport closure. In the same year, in December 2014, rain, wind and snow caused flight delays and cancellations throughout the day at major airports including Atlanta, Boston, and Chicago [5]. In addition to weather, another important concern in the airspace system is delay, either on the ground or en route. Such delays might result in significant disruptions at upstream or downstream airports. To mitigate their impact on subsequent airline operations, some modifications to the original flight schedule may be needed by redistributing the existing slack [6].

Air traffic in USA and Europe has experienced spectacular growth and a 50% increase traffic is expected by 2018 over the traffic in 1999 [7]. The increasing traffic challenges the capacity of the airspace system in satisfying the growing demand, leading to congestion and delay. As reported by the U.S. Department of Transportation, in 2007, approximately 26% of the flights were delayed by more than 15 min, whereas another 3% were cancelled [8]. As indicated in the report released by Federal Aviation Administration (FAA) in 2015 [9], with steady traffic growth as forecasted, more hub airports will become congested and delays will continue to grow in 2030. Typical Air Traffic Flow Management (ATFM) actions, such as rerouting aircraft due to convective weather [10], delaying flights already en route by restricting miles-in-trail or other means, and Ground Delay Programs (GDPs) [11], have been studied to alleviate the congestion. Among these actions, aircraft re-routing is one of the most commonly used strategies because it imposes less controller workload, and the risk is very low. Currently, many ATFM operators, such as the Air Traffic Command Center (ATCC) coordinating flow management in US, are called upon to reschedule and re-route the aircraft so as to minimize the delay costs caused by congestion. Specifically, in the context of airline operations, prior to the disruption, the aircraft are scheduled to fly a set of routes as planned. Upon the occurrence of a disruption, the initial routes become infeasible, which calls for aircraft re-routing, that is to identify a new feasible route for each aircraft. In this paper, we consider the case when an airport is closed, thus we

* Corresponding author.
E-mail address: sankaran.mahadevan@vanderbilt.edu (S. Mahadevan).

http://dx.doi.org/10.1016/j.dss.2017.02.005
0167-9236/

need to seek feasible routes for all the aircraft that are flying towards this airport based on the characteristics of each aircraft and available space at nearby airports.

There is extensive literature on flight re-routing to minimize congestion cost [12–16]. In particular, Odoni presented seminal work [17] to model the flow management problem (FMP) as a discretized representation of flows, which is fundamental to the approaches developed subsequently. Bertsimas and Patterson [18] extended the Air Traffic Flow Management (ATFM) model to account for the re-routing of soon-to-depart or airborne flights. Subsequently, they developed a dynamic network flow approach to address the problem of determining how to reroute aircraft when faced with dynamically changing weather conditions [19]. de Matosand and Ormerod [20] discussed the issues in European ATFM and identified the needs in terms of Decision Support Systems (DSS) for re-routing flights. In response to de Matos and Ormerod, de Matos and Powell [21] attempted to build a re-routing decision support system in Europe by identifying participants and investigating the requirements for re-routing systems. To respond to aircraft grounding and delays experienced over the course of the day, Bard and Arguello [22] proposed a time-band optimization model for reconstructing aircraft routing, in which the time horizon transformed from the routing problem is discretized. van Balen et al. [23] analyzed an optimal aircraft re-routing plan in a free flight environment in the case of airspace closure and developed a software to re-route the aircraft around the closed airspace. Recently, an integer programming model has been proposed in [8] for large-scale instances of the ATFM problem covering all the stages of each flight, i.e., take off, cruise, and landing. Subsequently, a method is proposed in [24] based on [8], which accounts for the airspace volume. Seminal work is presented in [25] on ATFM rerouting *under uncertainty* in airport arrival and departure capacity, air sector capacity and flight demand. Rosenberger et al. [26] developed an optimization model for aircraft recovery (ARO) that reschedules legs and reroutes aircraft by minimizing an objective function involving rerouting and cancellation costs.

Some of the above studies have considered uncertainty, but mostly related to the external environment, e.g., uncertainty in weather prediction [20,27,28]. Even when some studies focus on the inherent uncertainty in the air traffic control (ATC) system, only one or two system components are taken into consideration, such as ground and departure delays [29]. Aircraft re-routing is a complicated process, in which several subsystems are involved, such as radar performance, message communication, and airport resource management. Visit https://www.faa.gov/air_traffic/flight_across_america/ for details on the process of a flight across the USA. At each phase, several sources of variability and uncertainty emerge in different subsystems. These uncertainty sources have a significant effect on the air traffic, which challenges the management techniques and decision support systems classically used by the researchers.

In the past years, several studies have made significant effort in the development of management policy or decision support systems in the presence of uncertainty because the problem under uncertainty is completely different from the deterministic case. For example, to cope with the uncertainty in the travel demands and traffic conditions, Balbo and Pinson [30] developed an adaptive solution to the bus network management problem by utilizing multi-agent approach to model the various bus network activities. To design the optimal policy for the transportation networks, Pathak et al. [31] proposed an agent-based model, in which they accounted for both aleatory uncertainties (e.g., random demand and randomness in speed) and epistemic uncertainty (e.g., the incomplete information of user behavior). Yoon et al. [32] developed a computer-based training prototype to improve the emergency response and recovery effectiveness under the pressures of incomplete and erroneous information. As mentioned earlier, due to the uncertainty arisen in the ATC system, there is also an increasing demand for developing a decision support system with the consideration of the multiple sources of uncertainties emerging in the re-routing process [33]. Prior to developing the decision support system, we need to answer several questions: where the uncertainties come from, how to characterize these uncertainties, how to propagate the uncertainties through the system, and how to quantify their influence on the system performance. Answering these questions will lay a firm foundation for the development of ATC decision support system under uncertainty.

In this paper, we consider re-routing aircraft given the occurrence of an abnormal weather condition. Thus, the weather condition is given and there is no uncertainty, especially considering the short duration of the re-routing simulation. As indicated in the paper later, there are 10 simulation cycles in our analysis, and each simulation cycle lasts for 5 min, which means the total simulation duration is only 50 min. We assume that there is no uncertainty in forecasting the weather condition for the 50 min; this is reasonable given the current technology for monitoring and forecasting the weather. On the other hand, our focus is on the inherent uncertainty sources within the re-routing system, e.g., message communication, radar, aircraft, and airport. By investigating their effect on the system performance, we aim to establish the foundation for the design optimization of the re-routing system. Obviously, weather is an external condition and is not a design variable.

Since aircraft re-routing is a complicated process, there are many interactions among the components, e.g. radar and decision maker, decision maker and airport, and airport and aircraft. Moreover, the status of the aircraft (e.g., detected or not, registered or not, and assigned or not) and radar (e.g., beam orientation) are varying over time. Simulation methods [31,34] are capable of capturing the time-varying updates and the information exchange between the multiple components of such systems. As a result, we construct a simulation-based approach to mimic the aircraft re-routing process and account for the aforementioned multiple sources of uncertainties. Compared to the current state of the art, we have made the following contributions:

1. We develop a simulation framework to model the aircraft re-routing process and incorporate a formulation for optimal re-routing assignments. Our simulation framework considers the aircraft re-routing process as a system of systems, and enables us to have a comprehensive understanding of the system work flow and the interactions among various systems. Such a simulation-based approach facilitates quantitative evaluation of the re-routing methodology.
2. We analyze various uncertainty sources in the systems involved in the re-routing process (aircraft, radars, communication systems, and neighboring airports), and perform the simulation and re-routing optimization by incorporating the uncertain sources. Incorporation of uncertainty sources is essential in ensuring a robust and reliable re-routing system.
3. We develop a methodology to quantitatively evaluate the performance of the re-routing system by considering the effect of the various uncertainty sources on the system performance, and incorporate reliability analysis techniques for this purpose.
4. We incorporate novel stochastic sensitivity analysis techniques to quantify the relative contributions of the different uncertainty sources on the system performance; this facilitates a quantitative basis for optimal allocation of limited resources to improve the system performance.
5. We incorporate surrogate modeling to improve the computational efficiency and enable fast analysis of system performance. This helps to quantitatively evaluate and compare

several different competing strategies for re-routing optimization and system improvement.

In our simulation system, there are three core modules: radar system, communication system, and the assignment algorithm, which will be introduced in Section 2. The following uncertainty sources have been considered in this paper:

- Radar performance: Radar performance depends on three parameters, namely: initial orientation, detection radius, and beam rotation speed. All three parameters are stochastic and follow different distributions.
- Neighboring airports: Space availability (gate, runway, and airspace) at each airport is uncertain, and is varying with time.
- Message communication system: Delay is commonly encountered in the communication system, and the amount of delay is stochastic.
- Aircraft: Aircraft have four attributes: speed, size, entry time and location. All of them are stochastic variables when analyzing or designing the overall system.

In the presence of the aforementioned uncertainty sources, some challenges emerge: (1) How to propagate the uncertainty through the re-routing system, (2) how to evaluate the system performance with the consideration of the aforementioned uncertainties, and (3) how to quantify the effect of each uncertain variable on the overall system performance.

In this study, we address these challenges by developing a dynamic simulation-based approach that accounts for multiple uncertainty sources and their impact on the aircraft re-routing decision. We do not consider the uncertainty from the external environment, such as the uncertainty in weather prediction; that is, we consider re-routing for a given weather condition. We focus on the aleatory uncertainty in the re-routing system related to message transmission, radar performance, demand from incoming aircraft, and space availability at nearby airports.

An optimization model is constructed to make periodic assignments of the registered aircraft with the objective of minimizing the overall travel distance. We use the system failure time, which is defined as the time instant that at least one aircraft in the simulation runs out of fuel, as a metric to measure the re-routing system performance. Since the duration of each simulation is limited, the system performance is not available in all simulations. An effective strategy is developed to extract the lower bound of the system failure time from right-censored data. The above procedures are represented as Step 1 in Fig. 1.

As stated previously, we are interested in evaluating the re-routing system performance in terms of the system failure time. To achieve this objective, a large number of simulations is required to build the probability distribution of system failure time. However, each simulation is very time consuming. Hence, we build a surrogate model to replace the original simulation. Since there are too many random input variables, we implement a data-driven sensitivity analysis method to identify the dominant variables, then build the surrogate model using support vector regression with the identified dominant variables. Then we construct the system failure time distribution using the surrogate model. The aforementioned processes are denoted as Steps 2 and 3 in Fig. 1.

The remainder of this paper is organized as follows. Sections 2 and 3 describe the two elements of the proposed methodology: aircraft re-routing and performance assessment. In Section 2, we present the airspace system, aircraft re-routing process, and the assignment algorithm. In Section 3, we define the system performance metric, discuss how to extract it from the simulation data, incorporate variance-based sensitivity analysis, and construct the SVR surrogate model with both precise and imprecise data. Section 4 demonstrates

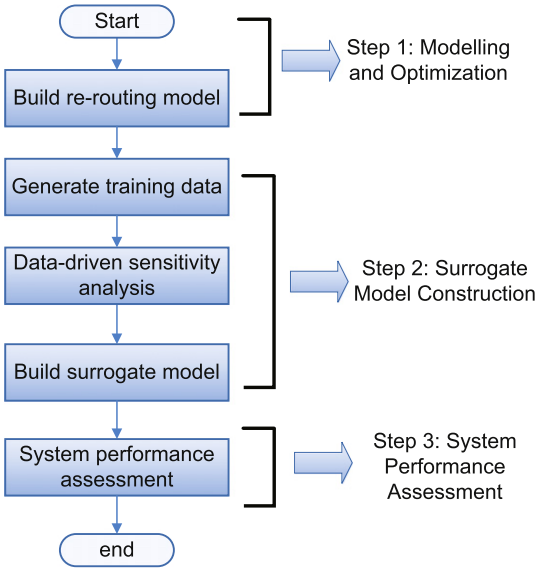


Fig. 1. Framework of aircraft re-routing optimization and performance assessment.

proposed re-routing model and assesses its performance using a numerical example. Section 5 provides concluding remarks.

2. Proposed methodology for aircraft re-routing

In this section, we introduce the components of the proposed re-routing methodology – problem description, airspace system model, uncertainties considered and the assignment algorithm.

2.1. Problem description

In the United States, the control of air traffic is carried out at 22 regional centers that receive information from aircraft and ground-based radars on location, altitude, and speed of the aircraft. The Air Traffic Command Center (ATCC) aggregates all the information collected by the 22 regional centers to reschedule and reroute flights when a hazardous event (e.g., bad weather) occurs [19]. In this paper, as shown in Fig. 2, we characterize this process in a simplified manner using a five-step procedure:

(1) In the first phase, the aircraft enters the region at a randomly generated location and time instant, and flies towards its destination airport at a constant speed. We assume that the flight path follows a straight line in a 2-D space that connects the entry location and the destination airport.
(2) The time when the aircraft enters the region of interest is Entry Time. Once the aircraft enters the region, it is detected after some time by either a regional radar or a local radar. The detection time is uncertain depending on the radar's orientation, detection angle, and detection radius. The time instant the aircraft is detected by the radar is Detection Time. Once a radar detects an incoming aircraft, it sends a message to the decision maker (DM). This message might be delayed. When DM receives the message from the radar, the aircraft is registered in the system. The time instant when the aircraft is registered in the system is called Registration Time.
(3) The DM periodically makes re-routing assignments for all the aircraft that are registered in the system and waiting to be re-routed. The re-routing decision for a particular aircraft requires the corresponding resources to be available and compatible, i.e., the aircraft of a particular size can only

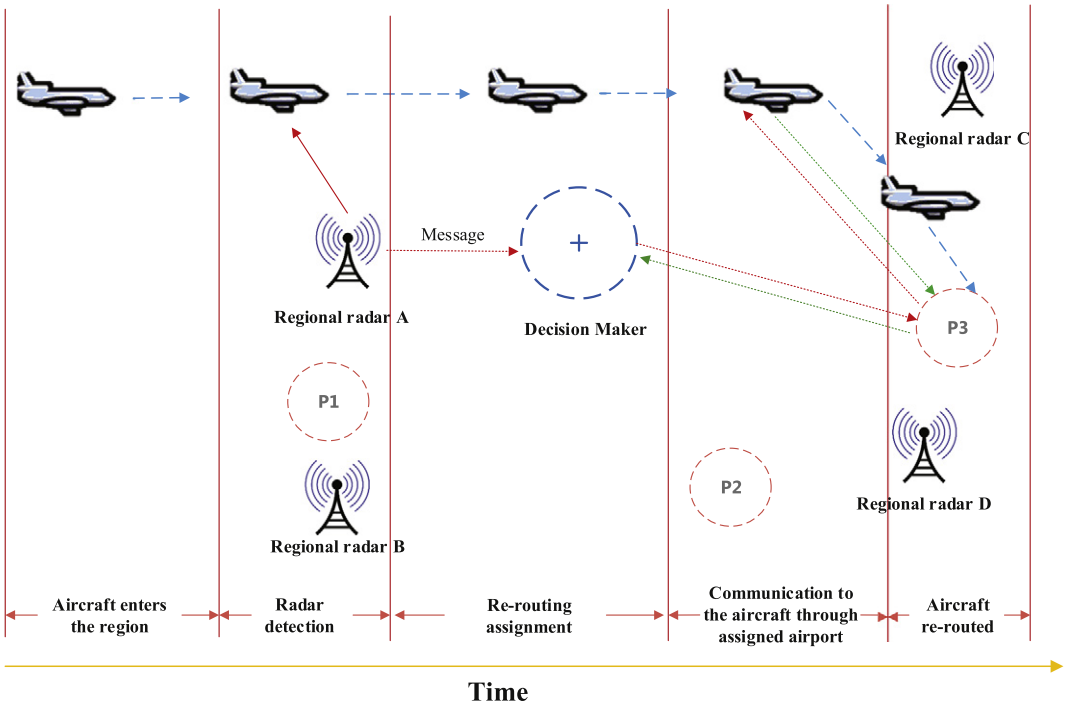


Fig. 2. Aircraft re-routing process.

land at the airport with available airspace, runway, and gate that are suitable for that aircraft size. The specific assignment algorithm will be developed in Section 2.4. The time instant when the aircraft is re-assigned is called Assignment Time.

(4) After the assignment decision is made, DM sends a message to the assigned airport. When the assigned airport receives the message, it sends an acknowledgement message to DM. It also sends a message to the assigned aircraft to establish further communication for landing. All three messages can be delayed.

(5) When the aircraft receives the message from the assigned airport, it sends an acknowledgement message to the assigned airport and starts flying towards the assigned airport. After the aircraft arrives at the assigned airport, the corresponding resources, e.g., gate, runway, and airspace, are released for future assignment. We name the time when the aircraft lands safely as Leaving System Time.

Obviously, aircraft re-routing is a complicated process, where multiple subsystems and components are involved, and the uncertainty related to each component further intensifies the complexity of this problem. In this paper, the system performance we are interested in is measured by the system failure time metric, defined as follows:

Definition 2.1 (System failure time). When at least one aircraft runs out of fuel, the re-routing system fails at that time instant. This time instant is defined as the system failure time.

Since several uncertain variables contribute to the variability of system performance, we are motivated to quantify the contribution of each uncertain variable, thus facilitating our decision making process for the purpose of enhancing system performance. To achieve these objectives, a simulation-based approach is used to characterize the re-routing process, by which we are able to capture the complex interactions among various subsystems. In the following sections, we introduce the modeling of each component and subsystem in the re-routing process.

2.2. System model

In this section, we introduce the simulation model we built to account for the various uncertainties in the process, related to airports, aircraft, radar, and communication. Before formulating the problem, we need to introduce the fundamental assumptions with respect to each component, in order to facilitate the discussion.

2.2.1. Region

Region refers to the specific area where we perform the simulation; all the components (e.g. radar, aircraft, and airport), are operated within the region. For the sake of illustration, in the numerical example, we use a 500 × 500 square area as shown in Fig. 3 as our simulation space. In reality, the region can be any shape, e.g., polygons, triangle, or any other shape.

2.2.2. Aircraft

In the re-routing problem considered, the aircraft are assumed to be flying towards a single destination airport at a constant speed. For each aircraft, we need to generate its entry time, entry location, aircraft size, speed, and remaining miles. Each of these variables is associated with uncertainty; we will discuss these uncertainties in Section 2.3. For the sake of simplicity, in the re-routing simulation, we only consider the aircraft flying into the region and heading towards the closed airport. As shown in Fig. 3, the airport indicated by a dashed circle in the center of the square is closed due to extreme weather, hence all the aircraft heading towards this airport will be re-routed to other alternate airports.

2.2.3. Nearby airports

A very important component in the re-routing decision is space availability at the nearby airports. In this paper, we choose 16 airports for the sake of illustration as shown in Fig. 3, but our simulation

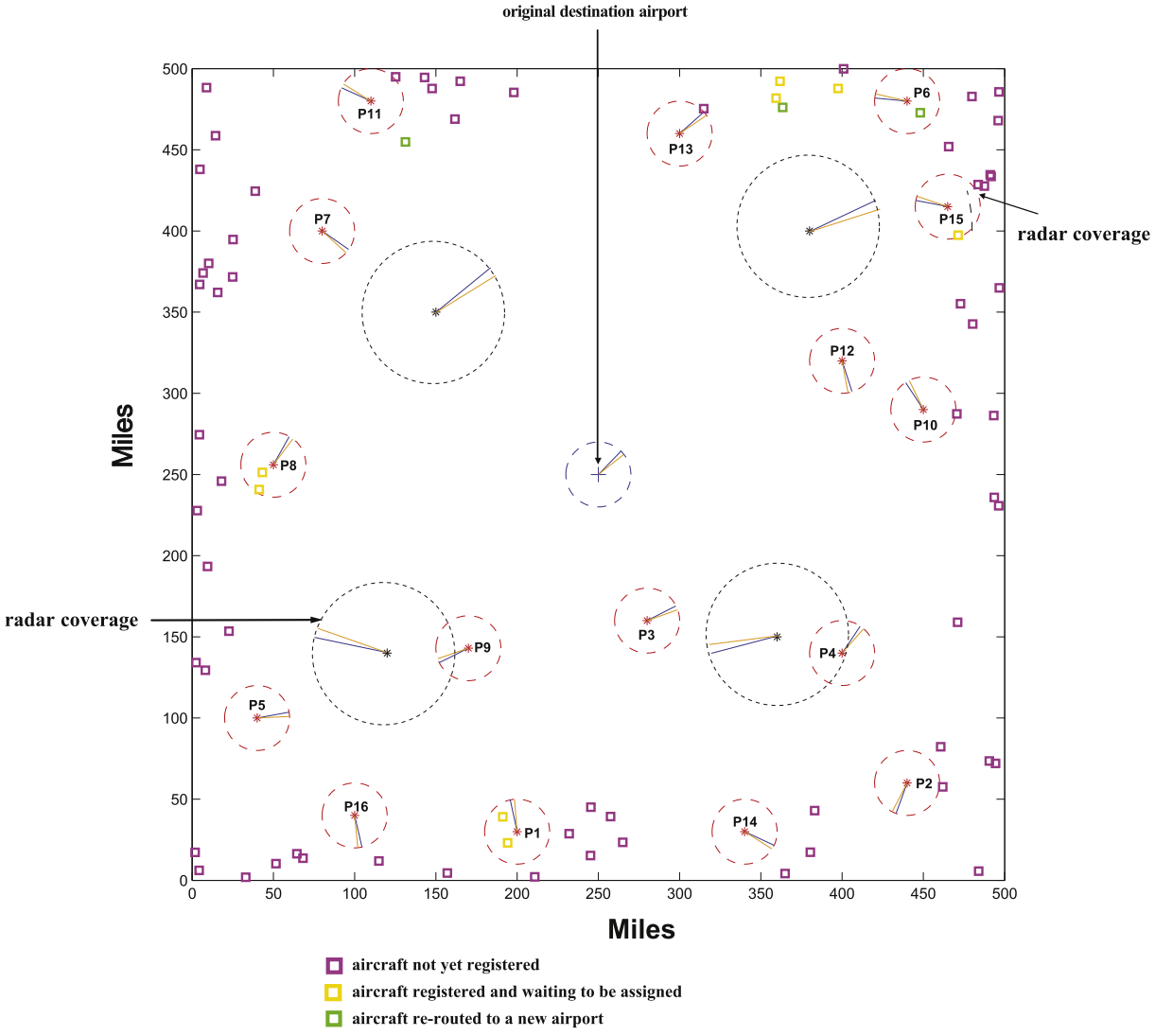


Fig. 3. A snapshot of the simulation system.

can handle any number of available airports for re-routing the aircraft. The availability of runways, gates, and airspace in each airport is considered by the assignment algorithm. In order to guarantee the safe landing of the aircraft, all the necessary facilities need to be available. There are several uncertain variables related to the space availability, and these will be introduced in Section 2.3.

2.2.4. Radars

Once an aircraft enters the region of interest, it might be detected by the radars in the system. Two different types of radar are considered: regional radars and airport (local) radars. The only difference between them is the detection radius. Regional radars have larger detection radius than local ones. The locations of all the radars are fixed. In this paper, we consider three uncertain variables related to radar: radar's initial orientation, radar beam rotation speed, and radar range, which will be described in Section 2.3 in detail.

2.2.5. Communication

There are many communications among the various components, for example, radar needs to send message to the decision maker to inform the newly detected aircraft. Currently, controller-pilot data link communication (CPDLC) is a commonly used means of communication between controller and pilot [35]. There is a sequence of messages between the pilot and the controller, which is referred to as dialogue. The duration of the dialogue is variable, depending on several factors, such as clarity of the vocal phraseology, controller workload, and the deviation of the revised route from the original one. All these factors increase our uncertainty in estimating the time of the communication.

Such delay has significant influence on the controller-pilot communication, manifested in increased number of "step-ons" (instances when a pilot or a controller begins his/her transmission before the previous transmission is over, resulting in interference blocking both transmissions), increased controller and pilot workload, and increased risk of violating aircraft separation [36]. Typically, controllers control traffic by issuing specific instructions and commands to pilots, then the pilots, after an unavoidable delay, execute the instructed maneuvers. If the time delays are long or have significant variability, accurate prediction of their consequences becomes difficult, thus substantially increasing controllers' mental workload and the probability of errors [37]. Therefore, understanding the effects of communication delay on system performance is important. In Ref. [36], several sources of uncertainty: audio delay (AD), pilot delay

(PD), and controller delay (CD), were identified, and two experiments were performed to examine the impact of systemic delay in the pilot-controller communication loop. The cumulative effect of communication delays may be of significance depending on the specific context. In particular, in congested airspaces, the controller often has a limited number of alternatives to choose from, and the success of his or her control actions will depend on the appropriate timing of the issued command and the pilot's response.

2.3. Stochastic variables considered

Note that in our case, initially, we do not know the relative contribution of each stochastic variable (aircraft, nearby airports, radar performance, and communication delays) on the system performance. Therefore, it is reasonable to first include all of them in our model; later, sensitivity analysis is used to identify the significant variables to be included in the model. In the following sections, we introduce the modeling of the uncertainty associated with each variable one by one.

2.3.1. Aircraft

With respect to the aircraft, we consider five different stochastic variables:

- **Entry Time**: Entry time determines when the aircraft enters the simulation region. Since the Poisson distribution is frequently used to express the probability of a given number of events occurring in a given amount of time if these events occur with a known average rate [38], we use it to generate the sequence of aircraft to enter the region in each simulation cycle. Thus, the inter-arrival time is modeled by an exponential distribution.
- **Entry Location**: The exact entry location of each aircraft could be anywhere on the boundary of the region. Since we use a square to represent the region of interest, two uniform random variables are used to characterize this uncertainty. The first variable determines one of four edges indicating the approach direction, while the other one specifies the location on the randomly selected edge. Using two independent random variables allows us to adjust the proportion of aircraft entries from different directions; for example, more aircraft could enter the region from the bottom edge if we assign a larger weight to it.
- **Size**: Aircraft of similar size have approximately the same passenger capacities, ranges, and velocities. Once an aircraft's arrival time and location are simulated, its size is randomly selected in the simulation. In this paper, aircraft are assumed to have three different sizes: large, medium, and small. We assume aircraft size is randomly selected among three options – large, medium, and small – with equal probabilities.
- **Speed**: Aircraft speed is related to its size. Typically, aircraft of larger size have higher cruise speed. In addition, even among the aircraft of the same size, the speed might vary from aircraft to aircraft. In this paper, we assume three different lognormal distributions to represent the speed for aircraft of large, medium, and small sizes, respectively.
- **Remaining Miles**: The number of remaining miles depends on the amount of fuel remaining in the aircraft. It is difficult to know the exact number of remaining miles of each aircraft. Thus, we use a lognormal distribution to characterize this uncertainty. The standard deviation is assumed to be 10 miles, and the mean value is assumed to be the product of a factor β and the distance from the aircraft entry location to the farthest airport in the region. This is only for the sake of illustration; the uncertainty regarding remaining miles can be represented in different ways. For example, the mean value could also be the distance from the entry location to the destination airport, multiplied by a factor greater than unity.

2.3.2. Nearby airports

In reality, each nearby airport might have different numbers of runways, gates, and airspace. To characterize the variability across the airports, we represent their mean values using variables κ, ξ, and η, which follow three different uniform distributions. In addition to that, for a given airport, there are aircraft arrivals and departures over time, which will occupy or release the resources over time. Two types of sampling are implemented in this paper to represent this variability: at the beginning of each simulation, the mean values κ, ξ, and η are randomly generated; in each assignment cycle, we make specific realizations with respect to the available runways, gates, and airspace, by which we account for the dynamic change of available space at each nearby airport. The above three types of spaces are classified by three different sizes of aircraft: large, medium, and small. In order to guarantee the safe landing of an aircraft in an airport, all three types of spaces should be available and suitable for that aircraft size. Then we account for the physical relationship between aircraft size and airport availability. For example, if the airport cannot accommodate large size aircraft, the variables κ, ξ, and η corresponding to large size aircraft will be zero.

2.3.3. Radar performance uncertainty

In reality, the radar performance is affected by many factors, such as the power density at the target position, reflected power, radar cross section, and antenna gain . The target detection not only depends on the power density at the target position, but also on how much power is reflected in the direction of the radar. A commonly used equation to compute the maximum radar detection range is [39]

$$R = \sqrt[4]{\frac{P_S \cdot G^2 \cdot \lambda^2 \cdot \rho}{P_E \cdot (4\pi)^3}} \tag{1}$$

where P_S is the transmit power, G is the antenna gain, λ is the transmit wavelength, ρ is the radar cross section, and P_E is the power received by radar antenna, expressed as

$$P_E = \frac{P_S \cdot G \cdot \rho}{(4\pi)^2 \cdot R^4} \tag{2}$$

As shown in Eq. (2), the power received by the radar P_E is inversely proportional to R^4, while the strength of the received power will influence whether the target can be detected accurately or not. Suppose $P_{E_{min}}$ is the smallest power that can be detected by the radar; if the power is less than $P_{E_{min}}$, then it is not usable because it is lost in the noise of the receiver. In other words, the probability of detecting a given target is associated with the strength of the received power, while the received power varies with the distance between the target and the center of the antenna beam. In addition to distance, when electromagnetic waves cross airlayers at different density, the occurrence of refraction results in the dispersion of the transmitter energy, which affects the energy received by the receiving antenna. Except refraction and dispersion, some other phenomena also have significant effect on radar's performance, e.g., interference with other signals and noise in the radar signal. Another important factor is the radar cross section ρ, and it is related to the angle formed by the aircraft flight path and the beam orientation.

These uncertainties are important for safety in aircraft operations. As indicated in Ref. [40], airborne radar can be utilized to sense the surrounding environment, e.g., wind speed and meteorological condition, thus aiding the aircraft Inertial Navigation System (INS). In this process, the uncertainty inherent in radar antenna and radar pointing angles, if not corrected, will seriously degrade the dual-Doppler winds. Besides, radar is also used to monitor aircraft location

and aid aircraft navigation [41,42]. If the uncertainty associated with the radar is too large, it increases the difficulty in estimating the aircraft location. In this case, to maintain the safety of each aircraft, we must make the separation distance larger. On the contrary, if we know the precise location of each aircraft, we can make the separation distance smaller, thus enabling us to accommodate more aircraft in the same airspace. This argument has been reflected in the development of NextGen technologies. One proposal to overcome radar uncertainty is the transition from radar surveillance to ADS-B (Automatic Dependent Surveillance-Broadcast) to track airplanes in flight and on the ground more accurately and reliably [43]. Such a proposal clearly demonstrates the significance of radar uncertainty in air traffic system. See Ref. [44] for the description of ADS-B.

To account for the above uncertainties related to radar, we model it in a simplified manner by characterizing its detection radius as a lognormal distribution. Specifically speaking, we consider the following uncertainties related to the radar:

- The detection radius varies from radar to radar. In the numerical example, regional radars are assumed to have a detection radius following a lognormal distribution $LN(R_r, 10)$ whereas local radars have a detection radius of $LN(R_l, 5)$. In the simulation, we implement different realizations for each radar. When the distance d is less than the specific radar range realization R, where d is the distance from the aircraft to the center of the radar (see Fig. 4 (a)), then the detection event occurs.
- The exact time when the aircraft is detected also depends on the initial orientation and radar beam rotation speed. The initial orientation of the radar beam varies from radar to radar. In this paper, the initial orientation is assumed to follow a uniform distribution $U(0, 2\pi)$. The beam rotation speed determines the sector that a radar can cross in a second. Suppose the radar antenna has 2 revolutions per minute (RPM), then it covers a 12^0 sector in 1 s. The beam rotation speed also varies from radar to radar. A lognormal distribution is used to characterize the variability of the beam rotation speed. As shown in Fig. 4 (b), an aircraft can be detected only when it is within the detection sector.

2.3.4. Communication delays

When disruption or inclement weather happens, the Air Traffic Command Center (ATCC) contacts each airline's Airline Operations Center (AOC) to inform about the necessity of re-routing. In this process, the regional radar and associated operators need to send messages to ATCC on the aircraft status, e.g., its altitude, location, and original landing airport. Based on the information collected from the ground-based radars, ATCC re-routes the corresponding flight to the available airport [19]. Multiple communications emerge in this process. Once the flight plan is filed, the communication starts between

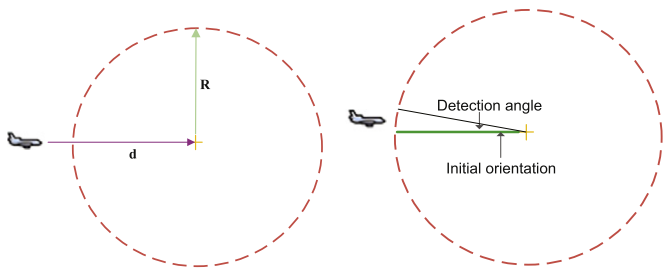


Fig. 4. Radar performance uncertainty.

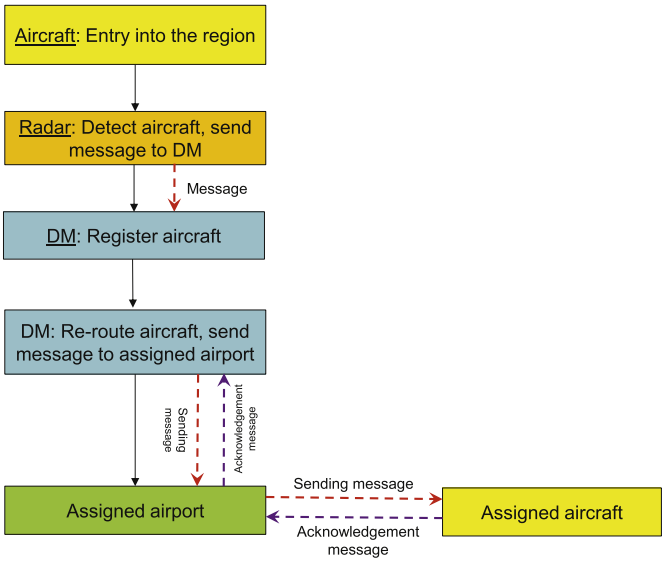


Fig. 5. System work flow.

aircraft (pilot) and assigned airport (controller). Prior to landing, all Instrument Flight Rule (IFR) flights are conducted with controller-pilot communication regarding the status of runway clearance, gate availability, and taxiways [45]. Each time the communication occurs, message delay or loss arises with a given probability.

In this paper, we simulate the above process in a simplified manner, in which the DM plays the role of ATCC. Fig. 5 illustrates the specific system work flow. There are three types of communication: registration, message communication between DM and assigned airport, and communication between assigned airport and assigned aircraft. Message delay can happen in any of three communications shown in Fig. 5. For the sake of simplicity, all delays are simulated using independent but identical lognormal distributions $LN(m_d, 4)$ seconds. If the message delay exceeds a predefined threshold, the message is treated as lost. To avoid message loss, each message sender periodically checks for acknowledgement (every 40 s for the sake of illustration), and sends a repeat message if acknowledgement is not received.

2.4. Mathematical model of aircraft re-routing optimization

Nomenclature

α_{ij}	A variable used to denote the size of aircraft i. If $j = 1$, the size is large; if $j = 2$, the aircraft's size is medium; if $j = 3$, its size is small
AS_i	Speed of aircraft i
$d_t(i,k)$	Distance between the ith aircraft and the kth airport at time t
LF_{it}	The remaining miles for ith airplane at time t
M	The number of airports in the system
P_t	The number of aircrafts at time t
SG_{jt}^k	The number of available gates for airplanes of size j in the kth airport at time t
SR_{jt}^k	The number of available runways for airplanes of size j in the kth airport at time t
SS_{jt}^k	The number of available airspaces for airplanes of size j in the kth airport at time t.
δ_{ikt}	A binary variable. $\delta_{ikt} = 1$, if ith aircraft is assigned to the kth airport at time t. Otherwise, $\delta_{ikt} = 0$

In the aircraft re-routing problem, suppose we have P_t aircraft in the region at time t and M airports. The DM makes periodic re-routing assignments. The objective is to the total distance travelled by all the airplanes, for safety and economic reasons [46,47], subject to system resources. The optimization problem can be formulated as

$$Minimize \sum_{i=1}^{P_t}\sum_{k=1}^{M} d_t(i,k)\delta_{ikt} \tag{3}$$

s.t.

$$\delta_{ikt} = 0 \quad or \quad 1, \tag{4}$$

$$\sum_{k=1}^{M} \delta_{ikt} = 1, \quad i = 1,\ldots,P_t \tag{5}$$

$$\sum_{j=1}^{3} \alpha_{ij} = 1, \quad i = 1,\ldots,P_t \tag{6}$$

$$\sum_{k=1}^{M}\sum_{i=1}^{P_t} \delta_{ikt} = P_t, \tag{7}$$

$$P\left(LF_{it} - d_t(i,k)\right) \geq 0.9, \quad i = 1,\ldots,P_t,\ k = 1,\ldots,M \tag{8}$$

$$\sum_{i=1}^{P_t} \alpha_{ij}\delta_{ikt} \leq \min\left(SR_{jt}^k, SG_{jt}^k, SS_{jt}^k\right) \quad j = 1,\ldots,3,\ k = 1,\ldots,M \tag{9}$$

$$LF_{it} = LF_{i0} - AS_i \times t, \quad i = 1,\ldots,P_t \tag{10}$$

where the decision variable δ_{ikt} is binary, defined as

$$\delta_{ikt} = \begin{cases} 1, & \text{if the } i_{th} \text{ aircraft is assigned to the } k_{th} \text{ airport at time } t \\ 0, & \text{otherwise.} \end{cases}$$

The constraint in Eq. (5) requires that each aircraft registered in the system in the current cycle must be assigned to one of the available airports. Eq. (6) imposes aircraft size restrictions. Eq. (7) requires that all the aircraft registered before time t should be assigned at time t. Eq. (8) is a reliability constraint. Both LF_{it} and $d_t(i,k)$ are random variables. This constraint requires that the remaining miles in any aircraft is larger than the distance between the aircraft and the assigned airport with a probability of 0.9. Eq. (9) imposes the constraint that the number of aircraft of size j assigned to airport k should be less than the number of available resources in size j in airport k (runways, gates, and airspace). Constraint (10) denotes the relationship between the remaining miles and time, where LF_{i0} indicates the initial remaining miles for ith aircraft.

This is an integer linear programming problem, in which only δ_{ikt} is the decision variable. In this paper, we use the built-in algorithm – **intlinprog** – in Matlab to approach the solution of this problem.

When the assignment model formulated in Eq. (3) cannot make feasible assignment of all the registered aircraft, i.e., space is not available at any of the nearby airports for some of the registered aircraft, we do not consider the system to have failed immediately. Instead, we build a waiting queue for the registered but yet unassigned aircraft. According to the available space, e.g., runways, gates, and airspaces, at all the airports, we select the same number of aircraft from the waiting queue to make the assignment. There are many ways to rank the aircraft in the waiting queue, e.g., by registration time, or by remaining miles. For the sake of illustration, we carry out a "first-come first-served" (FCFS) strategy to order the aircraft in the waiting queue according to their registration time.

The optimization objective formulated in Eq. (3) is only for the sake of illustration, there are many other alternatives in real-world applications. In some programs, the airline company or ATCC aims to minimize the delay costs, e.g., passengers' delay, by re-routing certain flights. When some unexpected disruptions occur, passengers are inconvenienced and various costs are incurred. From the airline's perspective, their goal is to minimize the influence caused by disruption, e.g., minimizing deviations from the original schedule. In this paper, we focus more on the uncertainty involved in the re-routing process rather than the optimization objective. Hence, for the sake of illustration, we minimize the total distance travelled by all the assigned aircraft to ensure safe landing within available remaining miles for as many aircraft as possible.

3. Proposed methodology for performance assessment

In this section, we describe the extraction of the system failure time from the re-routing simulation, conduct variance-based sensitivity analysis to identify significant variables, and construct an SVR surrogate model with precise and imprecise data in order to construct the probability distribution of system failure time.

3.1. System failure time

As mentioned in Section 2, system failure time (see Definition 2.1) is the metric of interest in analyzing the re-routing system performance. To build the system failure time distribution efficiently, the following analyses need to be performed:

- Given a specific input variable setting, how to efficiently predict the system failure time distribution without running the entire simulation?
- How to quantify the contribution of each uncertain input variable towards the variance of the output of interest, i.e. system failure time?

In each simulation, we have a fixed duration T to run the simulation to see if the system survives or fails. However, in some cases, the system survives longer than time T, which results in right-censored data defined as follows:

Definition 3.1 (Type 1 censored data). Consider reliability testing of n (non-repairable) units taken randomly from a population. During T hours, we observe $r(0 \leq r \leq n)$ failures. The exact failure times are $t_1, t_2, \ldots, t_r$, and the other $n - r$ units survive the entire T-hour test without failing. This type of censoring is named "right censored" data since the times of failure to the right (i.e., larger than T) are not available.

Since the times of failure to the right (i.e., larger than T) are missing for some simulations, the challenge is how to estimate the system failure time. But when the system fails, we have the information of all the aircraft waiting to be assigned, and we can identify the aircraft with the least remaining miles. Here, we apply a conservative strategy by using the lower bound of the system failure time as the actual system failure time. Suppose the least remaining miles is LF_{min}, the following theorem can be utilized to estimate the system failure time.

Theorem 3.1. *If LF_{min} is less than the minimum distance between the aircraft current location and all the airports, given its mileage consumption rate r, the time to failure is*

$$S_{FT} = T + \frac{LF_{\min}}{r} \tag{11}$$

where T represents the duration of the simulation.

Since LF_{min} is less than the minimum distance between its current location and all the nearby airports, in the subsequent assignments, it is impossible to make a feasible assignment. As a result, its remaining survival time determines how long the system can survive further. If LF_{min} is larger than the minimum distance, then we only know that the system survives longer than the S_{FT} in Eq. (11), thus S_{FT} captures the lower bound of the system failure time. In this case, the system failure time can only be expressed as an interval: $[S_{FT}, +\infty]$.

3.2. Variance-based sensitivity analysis

The sensitivity of a model output to the stochastic inputs can be measured through the relative contribution of each stochastic input to the variance of the output; the well known Sobol' indices (defined in Appendix C) are commonly used for this purpose. As mentioned in Appendix C, to compute Sobol' indices, a double-loop Mente Carlo simulation is required. There are 24 stochastic variables in our system. To calculate x_i's sensitivity, in the inner loop, we need to fix x_i and change all the other variables x_{-i}. Whereas, the outer loop requires fixing x_i at different values. Suppose both inner loop and outer loop have 100 cycles, then 10,000 simulations are needed, which in turn takes a large amount of time to collect the data. This is only sensitivity analysis for one single variable, not to mention sensitivity analysis for the remaining 23 variables. Assume that only 1000 simulations are affordable.

Fortunately, an efficient data-driven sensitivity analysis method has been recently developed [48], based on the concept of stratified sampling [49]. Suppose we divide the range of an input variable x_i into equally probable intervals $\boldsymbol{\phi} = \{\phi^1, \cdots, \phi^M\}$, then the first-order Sobol' index can be computed as

$$S_i = 1 - \frac{E_\phi\left(V_{\phi^l}(y)\right)}{V(y)}, \quad l = 1, \cdots M. \tag{12}$$

where $V_{\phi^l}(y)$ represents the variance of y when x_i is in the subspace ϕ^l, $V(y)$ denotes the variance of the system response y. See Ref. [48] for details of this approach.

3.3. SVR surrogate model with precise and imprecise data

Generally speaking, in reliability analysis, the failure rate is quite small [50], e.g., $10^{-4}$and 10^{-5}. If we aim to use simulation-based approaches to quantify the system failure probability, a large number of samples is required to estimate the failure probability. Suppose the failure probability is 10^{-4}, then we expect to obtain 1 failure data point out of 10,000 samples on average. In our re-routing simulation, each simulation lasts a duration of 2 to 10 min; depending on initial input variable settings, suppose we wish to collect 10,000 samples from the system, this will take 20,000 to 100,000 min, which is prohibitive. As a result, we need to build a surrogate model to replace the original simulation in order to efficiently build the system failure time distribution.

As mentioned earlier, system failure is not necessarily observed in all simulations. In such cases, we represent its survival time as an interval: $[S_{FT}, +\infty]$, where S_{FT} denotes the lower bound of system failure time. Suppose the dataset we have collected from the re-routing simulation is represented by $\mathbb{U}$, which we divide it into two separate groups: right-censored data, denoted by $\mathbb{U}_1$, and uncensored point targets, denoted by $\mathbb{U}_2$.

Consider a censored data point $\boldsymbol{x_i} \in \mathbb{U}_1$, in this setting, we want the predicted survival time for x_i to be within the range $[S^i_{FT}, +\infty]$, where S^i_{FT} denotes the lower bound of system failure time for x_i. As long as the predicted value is larger than S^i_{FT}, there is no penalty. Otherwise, we penalize if the predicted survival time is less than S^i_{FT}. Hence, the robust ϵ-insensitive cost function L_ϵ for x_i can be updated as follows:

$$L_\epsilon(f(\boldsymbol{x}), y) = \begin{cases} S^i_{FT} - f(\boldsymbol{x}) - \epsilon & \text{if } S^i_{FT} - f(\boldsymbol{x}) > \epsilon, \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$

Obviously, the loss function for the right-censored data $\mathbb{U}_1$ becomes one sided, that is

$$S^i_{FT} - \boldsymbol{w}\cdot\phi(\boldsymbol{x_i}) - b \le \epsilon + \xi_i, \quad \boldsymbol{x_i} \in \mathbb{U}_1, \tag{14}$$

subject to,

$$\xi_i \ge 0.$$

For the point target $\boldsymbol{x_i} \in \mathbb{U}_2$, its loss function is still two-sided, that is

$$\begin{cases} y_i - (\boldsymbol{w}\cdot\phi(\boldsymbol{x_i}) + b) \le \epsilon + \xi_i, \\ \boldsymbol{w}\cdot\phi(\boldsymbol{x_i}) + b - y_i \le \epsilon + \xi^*_i, \\ \xi_i, \xi^*_i \ge 0. \end{cases} \tag{15}$$

By combining the above equations, we now propose the following formula for SVR regression in the re-routing system:

$$\min \quad C\frac{1}{N}\sum_{i=1}^{N}(\xi_i + \xi^*_i) + \frac{1}{2}\|\boldsymbol{w}\|^2 \tag{16}$$

subject to,

$$\begin{aligned} &S^i_{FT} - \boldsymbol{w}\cdot\phi(\boldsymbol{x_i}) - b \le \epsilon + \xi_i, \quad \text{if } \boldsymbol{x_1} \in \mathbb{U}_1, \\ &y_i - (\boldsymbol{w}\cdot\phi(\boldsymbol{x_i}) + b) \le \epsilon + \xi_i, \quad \text{if } \boldsymbol{x_1} \in \mathbb{U}_2, \end{aligned} \tag{17a}$$

$$\boldsymbol{w}\cdot\phi(\boldsymbol{x_i}) + b - y_i \le \epsilon + \xi^*_i, \quad \text{if } \boldsymbol{x_1} \in \mathbb{U}_2, \tag{17b}$$

$$\xi_i, \xi^*_i \ge 0, \quad i = 1, \ldots, N. \tag{17c}$$

As can be noted from the above equations, we take advantage of all the useful information available in the data. By introducing Lagrange multipliers α_i for the inequalities (Eq. (17a)), and α^*_i for the inequalities (Eq. (17b)), the dual of the above problem can be formulated as

$$\begin{aligned} \max \quad & -\epsilon\sum_{i=1}^{N}\left(\alpha_i + \alpha^*_i\right) + \sum_{\boldsymbol{x_i}\in\mathbb{U}_2}\left(\alpha^*_i - \alpha_i\right)y_i - \sum_{\boldsymbol{x_i}\in\mathbb{U}_1}\alpha_i S^i_{FT} \\ & - \frac{1}{2}\sum_{ij=1}^{N}\left(\alpha^*_i - \alpha_i\right)\left(\alpha^*_j - \alpha_j\right)\boldsymbol{K}\left(\boldsymbol{x_i}, \boldsymbol{x_j}\right) \end{aligned} \tag{18}$$

subject to

$$\begin{aligned} &\sum_i \alpha_i - \alpha^*_i = 0, \\ &0 \le \alpha, \alpha^* \le C. \end{aligned} \tag{19}$$

where $\alpha^*_i = 0$, $\forall \boldsymbol{x_i} \in \mathbb{U}_1$. By maximizing the objective function (Eq. (18)), for a new variable $\boldsymbol{x}$, its function value is represented by $f(x) = \sum_{i=1}^{N}(\alpha_i - \alpha^*_i)\boldsymbol{K}(\boldsymbol{x}, \boldsymbol{x_i}) + b$.

The performance of ϵ-SVR is mainly influenced by three parameters: kernel function $\boldsymbol{K}$, C, and ϵ. Currently, there are no unified rules for determining these parameters [51]. A common approach is to

select their values by trial and error. In this paper, we implement a grid search cross-validation approach to choose their values.

The prediction performance of the SVR model can be evaluated using the following performance measure, namely relative mean errors (RMEs):

$$RME = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - y_i^*}{y_i}\right| \tag{20}$$

where y_i is the actual value, and y_i^* is the forecast value. For censored data, if $y_i^* \geq y_i$, then RME is 0. RME measures the deviation between actual and predicted values. The smaller its value is, the closer are the predicted system failure times to the values from the original simulation model.

4. Numerical examples

In this section, two numerical examples are given to demonstrate the proposed method for the aircraft re-routing and system performance assessment.

4.1. Input variable settings

Table 1 lists the random variables, their distributions and parameters. These variables can be grouped into two classes: system variables (e.g., regional radar range, local radar range, regional radar detection angle, local radar detection angle, message delay, and space availability), and aircraft variables (remaining miles ratio, speed, and arrival rate). Besides, each radar's initial orientation follows a uniform distribution $U(0, 2\pi)$. In this paper, we assume all these stochastic variables, e.g., message delay and radar beam rotation speed, are independent from each other. If there are correlations between the stochastic variables, it can be easily included in MCS simulations. Well-established methods [38] are available in the literature to generate MCS samples with correlated factors.

4.2. Data collection

In our analysis, each simulation lasts for 10 cycles, each of duration 5 min. Every 5 min, the re-routing system makes feasible assignment decisions, pairing registered aircraft with airports based on available space. Based on the probability distributions for each variable described in Table 1, we run 2000 simulations and collect 2000 data points. Among them, 71 data points are found to be right-censored.

As mentioned previously, the space availability at each nearby airport is associated with three resources: runway, gate and airspace. In Table 1, the availabilities are represented by three variables with their means κ, ξ, and η. As shown in Eq. (9), the assignment algorithm is constrained by the minimum available resources among runway, gate, and airspace. Thus, $\min(\kappa, \xi, \eta)$ is used to characterize the space availability at each nearby airport. In this way, 16 variables are employed to represent the space availability at the nearby airports. Thus, a total of 25 input variables are used to characterize the system inputs, and the system failure time is the only output.

4.3. Sensitivity analysis

Due to the appearance of right-censored data, we cannot perform global sensitivity analysis directly because we do not know the exact system failure time for some of the simulations. Since we could only estimate the lower bound of the system failure time for right-censored data, to perform sensitivity analysis, we need to use a crisp number to represent the system failure time. There are several ways to handle censored data. We can perform sensitivity analysis on the dataset after discarding the censored data, but the system failure time characterized by the censored data is lost. Another way is to use a vary large number to represent the system failure time. But which specific number to use is an issue. In this case, we use a conservative strategy to conduct the sensitivity analysis the lower bound is used as the actual system failure time of the censored data, then we carry out global sensitivity analysis. Based on Eq. (12), we compute the first-order sensitivity using 2000 samples from the simulation. Fig. 6 shows the first-order sensitivity index for each input variable. Obviously, the remaining miles has the highest sensitivity.

Another observation is that the sensitivity of space availability varies with the location of each airport. Among all the airports, the space availabilities at airports P14 and P15 have higher sensitivity indices whereas the sensitivities at airports 4, 10, 11 and 12, are negligible. The sensitivity of space availability at each airport is greatly influenced by the complicated interactions in the system. In the assignment model, the overall travelling distance is affected by the location of the registered aircraft. But the time that an aircraft is registered is affected by other factors, e.g., its entry time and location, and the time it is detected by the radar. When the aircraft is registered determines its location, which in turn influences our assignment decision.

4.4. Surrogate model construction

Based on the sensitivity analysis result shown in Fig. 6, we eliminate the variables with low contribution to the overall system variance, thus reducing the problem dimension and accelerating the surrogate model construction. Among the 25 random variables considered in this paper, we remove ten variables with low sensitivity indices. The 15 variables that are retained in the surrogate model are regional radar range, local radar range, radar beam rate, message delay, remaining miles, aircraft speed (large), aircraft speed (medium), aircraft speed (small), and the space availability at airports 5, 6, 8, 9, 13, 14 and 15.

Table 1
Assumed probability distributions for the input variables.

Variables	Numbers	Distributions and parameters	Parameter ranges	Unit
Regional radar range	1	Lognormal: mean R_r, stdev 10	$R_r \sim U[40, 60]$	Miles
Local radar range	1	Lognormal: mean R_l, stdev 5	$R_l \sim U[14, 24]$	Miles
Radar beam rotation speed	1	Lognormal: mean B_s, stdev 0.3	$B_s \sim U[1.8, 2.2]$	RPM (revolution per minute)
Initial orientation	16	Uniform	$U[0, 2 * \pi]$	
Message delay	1	Lognormal: mean m_d, stdev 4	$m_d \sim U[25, 50]$	Seconds
Space availability at airport k	16	Runway – Lognormal: mean κ, stdev 1	$\kappa \sim U[1, 3]$	
		Gate – Lognormal: mean ξ, stdev 5	$\xi \sim U[0, 24]$	
		Airspace – Lognormal: mean η, stdev 3	$\eta \sim U[0, 9]$	
Arrival rate	1	Poisson: λ	$\lambda \sim U[20, 30]$	5 min
Remaining miles ratio β	1	Uniform	$\beta \sim U[0.6, 1]$	
Aircraft speed s	3	Large – Lognormal: mean ν, stdev 65	$\nu \sim U[546, 609]$	Miles per hour
		Medium – Lognormal: mean ς, stdev 40	$\varsigma \sim U[345, 414]$	Miles per hour
		Small – Lognormal: mean χ, stdev 28	$\chi \sim U[226, 292]$	Miles per hour

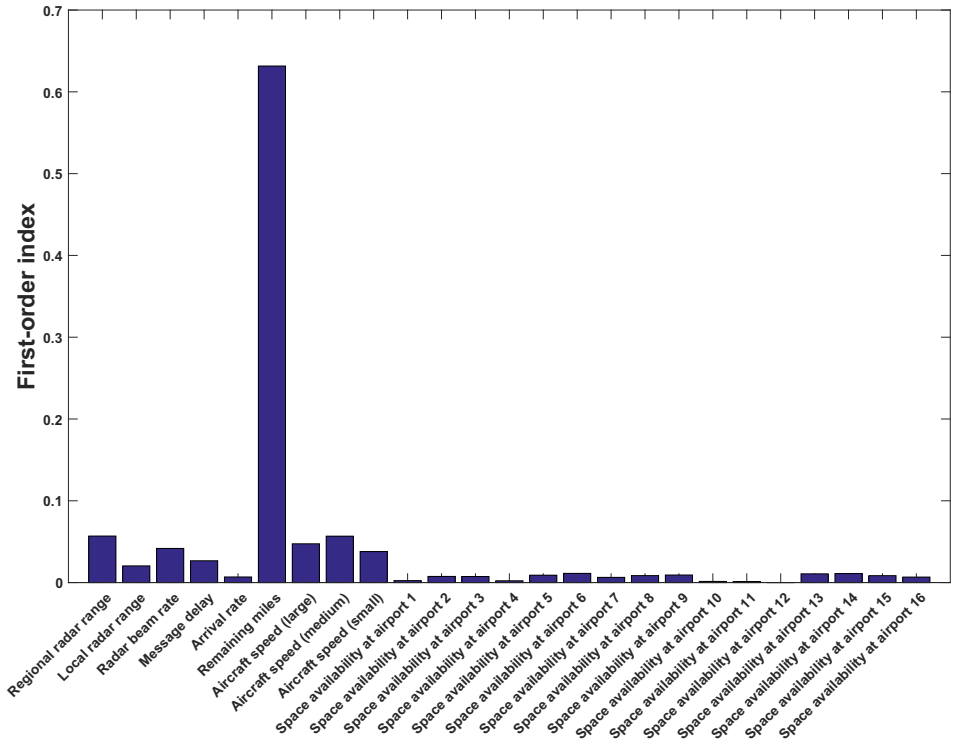


Fig. 6. First-order Sobol' indices.

Next, cross validation is used to determine parameter values. The data is split into ten groups: nine out of ten are used to train the model, and the remaining one is used to validate the performance of the trained model. Then grid search is used to search the space of these variables using exponentially growing sequences of C and ϵ to identify good values of the parameters (for example, $C = 2^{-4}, 2^{-2}, 2^{0}, 2^{2}, 2^{4}$). We try three kernel functions: linear kernel, Gaussian kernel, and polynomial kernel. The kernel function $\boldsymbol{K}$, with parameter set of C and ϵ, which yields the minimum RME, is selected. It is found that the parameter set $C = 2^{-3}, \epsilon = 2^{-10}$, and $\gamma = 2^{2}$ gives the best prediction result (minimizing the test RME).

Fig. 7 shows the comparison results between the actual simulations and the SVR surrogate model predictions for 200 data points. The SVR predictions (with polynomial kernel) show good agreement

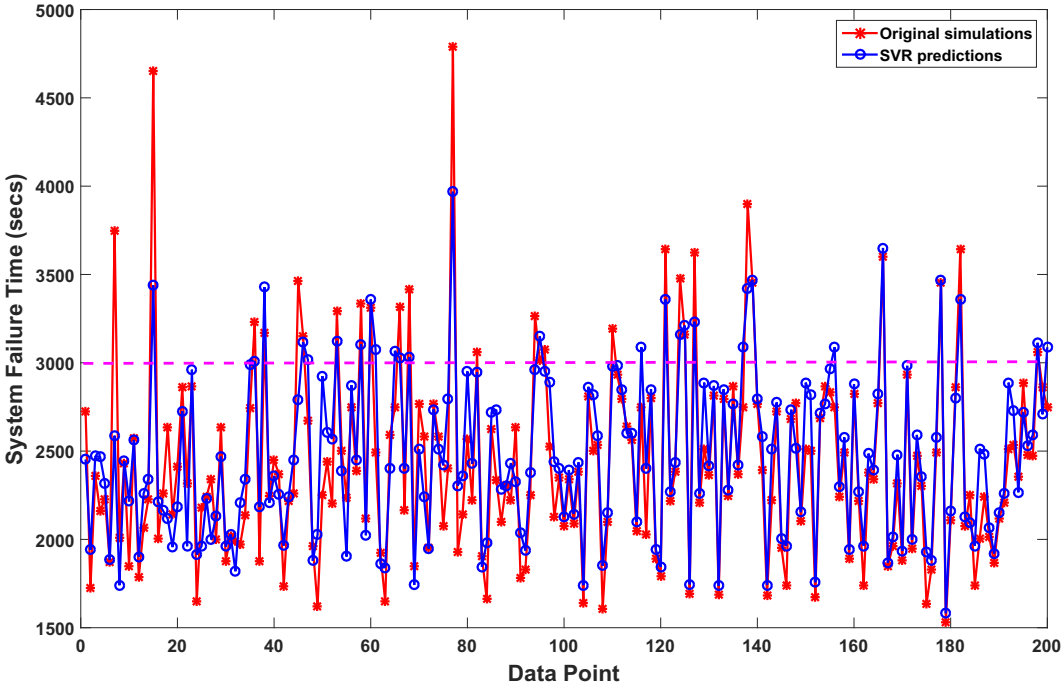


Fig. 7. System failure time prediction with ϵ-SVR model.

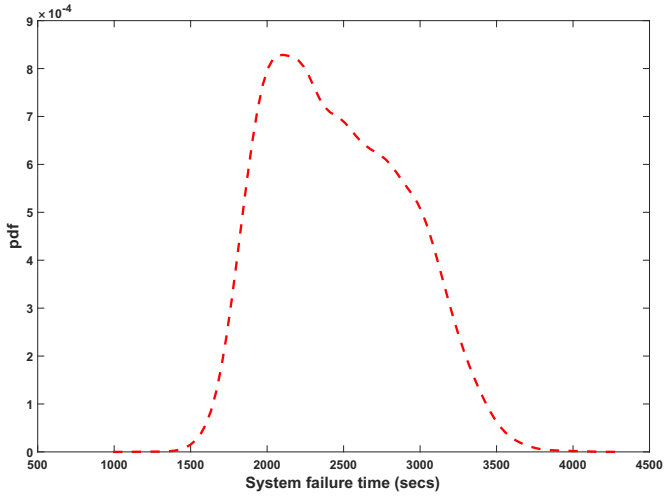


Fig. 8. First case: System performance assessment.

with the observed system failure time. The relative mean error is 7.3%, which indicates that the SVR model is reasonably accurate. If higher accuracy is desired, more runs of the original re-routing simulation are needed.

4.5. Case I

With the surrogate model constructed in the previous section, we evaluate the system performance by randomly sampling the 15 variables within their ranges. Fig. 8 shows the system failure time distribution based on 20,000 samples.

Given the 20,000 samples, we fit the data using five candidate distributions: Weibull, lognormal, normal, gamma, and exponential. The parameters related to these distributions are estimated by maximizing the corresponding likelihood function. Two goodness-of-fit plots are shown in Fig. 9 to demonstrate the performance of each candidate distribution. Fig. 9 (a) compares the density functions of the fitted distributions along with the histogram of the empirical distribution, and Fig. 9 (b) denotes the comparisons between the CDF plot of both the empirical distribution and the fitted distributions. As can be observed, Gamma and Lognormal distributions give reasonable agreement with the data.

In addition, three well-known quantitative tests are employed to quantify the goodness-of-fitness of each candidate distribution: Kolmogorov-Smirnov, Cramer-von Mises, and Anderson-Darling, which are defined in Table 2. The statistical test results are shown in Table 3. Among the candidate distributions considered, the lognormal distribution has the lowest statistical error in all three tests. In other words, the lognormal distribution fits the data best. The parameters of the best-fitted lognormal distribution are $\mu = 7.79$and $\sigma = 0.1786$. Suppose we want to estimate the probability that the system survives longer than 2600s; the results are shown in Table 4. As can be noted, there is a 6.3% deviation of the system reliability estimation between the fitted distribution and the actual simulation. But considering the two approximations in estimating the system reliability, our prediction is very precise. Specifically, when we build the surrogate model from the 2000 simulation data, there is approximation error between the simulation data and the surrogate model, which is about 7.3%. When we fit the distributions according to the samples generated by the surrogate model, the second approximation arises. Thus, the 6.3% prediction discrepancy is within the range of the deviation caused by the two approximations.

In addition, the Kaplan-Meier estimator fits the data very well. Since the simulation lasts for 3000 s, when we estimate $P(S_{FT} > 2600)$, there is no right-censored data involved. In this case, the Kaplan-Meier estimator is equivalent to the empirical cdf. In contrast, when we evaluate the probability of $P(S_{FT} > 4200)$, censored data arises. It can be observed that the Kaplan-Meier estimator overestimates the probability of $P(S_{FT} > 4200)$. However, the fitted distribution gives a more conservative prediction for this example when compared with the Kaplan-Meier estimator.

4.6. Case 2

In this section, we broaden the range of the mean for the space availability at the nearby airports: $\kappa \sim U[1,6]$, $\xi \sim U[0,48]$, and

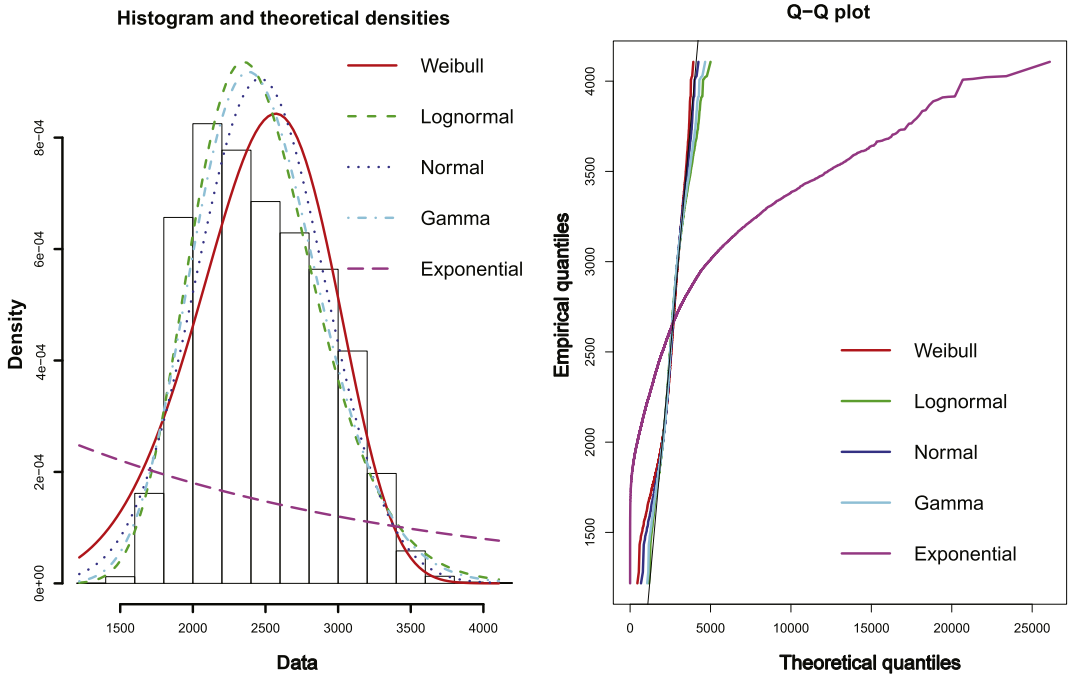


Fig. 9. Two goodness-of-fit plots for various distributions fitted to continuous data (Weibull, Lognormal, normal, gamma, and exponential distributions fitted to 20,000 samples).

Table 2
Goodness-of-fitness statistics as defined by Stephens [52].

Statistic	General formula	Computational formula
Kolmogorov-Smirnov (KS)	$\sup \lvert F_n(x) - F(x)\rvert$	$\max(D^+, D^-)$ $D^+ = \max_{i=1,\ldots,n}\left(\frac{i}{n} - F_i\right)$ $D^- = \max_{i=1,\ldots,n}\left(F_i - \frac{i-1}{n}\right)$
Cramer-von Mises (CvM)	$n\int_{-\infty}^{+\infty}(F_n(x) - F(x))^2 dx$	$\frac{1}{12n} + \sum_{i=1}^{n}\left(F_i - \frac{2i-1}{n}\right)^2$
Anderson-Darling (AD)	$n\int_{-\infty}^{+\infty}\frac{(F_n(x)-F(x))^2}{F(x)(1-F(x))}dx$	$-n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\log(F_i(1-F_{n+1-i}))$

Table 3
Goodness-of-fit statistics.

	Weibull	Lognormal	Normal	Gamma	Exponential
Kolmogorov-Smirnov statistic	0.06790564	**0.04280514**	0.06185114	0.04557395	0.489977
Cramer-von Mises statistic	29.82418846	**15.17619515**	23.01538695	15.60813899	1322.880084
Anderson-Darling statistic	199.44501732	**95.18102060**	143.27729802	98.24155121	6221.194507

The entries in bold indicates the goodness–of–fit statistics of the best–fitted distribution.

Table 4
System reliability analysis.

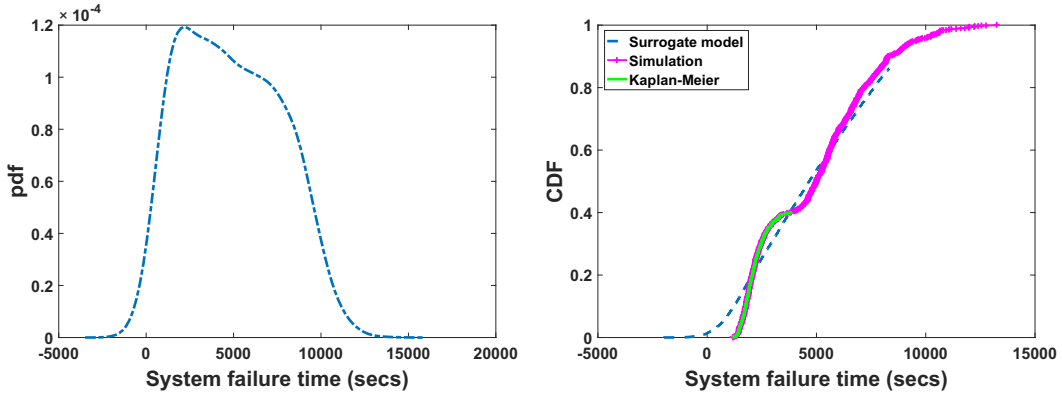
	Original simulation	Lognormal distribution	Kaplan-Meier estimator
$P(S_{FT} > 2600)$	0.3255	0.3408	0.3255
$P(S_{FT} > 4200)$	0.023	0.012	0.0337

$\eta \sim U[0, 18]$. Since the upper bound in the mean values of available runways, gates, and airspace has increased, the system survives a longer time than Case I. Following the same procedure, we collect 1000 data points, of which 600 are found to be right-censored. The system failure time of the right-censored data is represented by an interval $[S_{FT}, \infty]$. As long as the predictions of our surrogate model fall within the range, there is no penalty. In other words, the prediction can be any number in the range $[S_{FT}, \infty]$. Thus, the right-censored data contributes to the uncertainty in the surrogate model prediction. The performance of the surrogate model will deteriorate if there are more right-censored data due to the increase in uncertainty.

Contrary to the first case, the average error of the surrogate model has risen to 18% in the second case due to the increase of right-censored data. Fig. 10 (a) shows the system failure time distribution. Since some of the system failure time predictions in the surrogate model are negative, we cannot fit it using the four candidate distributions. Table 5 shows the comparison results among original simulation, surrogate model, and Kaplan-Meier estimator. Fig. 10 (b) shows the comparisons of cumulative distribution function. When the time is less than 3844 s, there are only 2 right-censored data out of 402 points. The Kaplan-Meier estimator is equivalent to the empirical CDF if there is no right-censored data. As a result, the Kaplan-Meier estimator nearly coincides with the empirical cdf. When the time is larger than 3844 s, all the system responses are right-censored. The Kaplan-Meier estimate of system reliability then becomes a constant 0.4012. However, the surrogate model still fits the cumulative distribution of system failure time very well, as shown in Fig. 10 (b) and Table 5.

5. Conclusion

This paper proposes an aircraft re-routing methodology in the presence of uncertainty arising from several different sources (radar performance, communication system, aircraft characteristics, and space availability in nearby airports) and develops an approach to assess the performance of the proposed methodology. A simulation-based approach is used to characterize the actual re-routing process and account for the uncertainty contributed by several heterogeneous subsystems. Based on the simulation data, we employ a data-driven sensitivity analysis method to quantify the contribution of each uncertain variable towards the overall variance of system response. The variables with the lowest importance are removed to reduce the problem dimension. A support vector regression surrogate model is built for predicting the system failure time and constructing the probability distribution for the failure time of the re-routing system.



(a) Case 2: System performance assessment. (b) Case 2: Comparisons of cumulative distribution function.

Fig. 10. Performance assessment.

Table 5
Case 2: system reliability analysis.

	Original simulation	Surrogate model	Kaplan-Meier estimator
$P(S_{FT} > 2600)$	0.6766	0.7311	0.6747
$P(S_{FT} > 3844)$	0.5978	0.5885	0.4012

Our contributions are several fold. First of all, since the current aircraft re-routing is operated manually, it is hard to assess the performance of aircraft re-routing operations. In this paper, we develop a simulation-based model to accommodate the various uncertainties arising in this process, and it allows us to evaluate the performance of the proposed re-routing algorithm quantitatively. Secondly, based on the simulation, we perform sensitivity analysis to quantify the contribution of each stochastic variable to the system reliability. The sensitivity analysis results enable us to identify which factor affects the system performance most, thus providing a quantitative basis for decisions such as component design optimization and specific operational measures in order to improve system performance. Given limited resources to improve the system performance, the sensitivity analysis result provides the guidance to optimally allocate the limited resources. Thirdly, our study models the aircraft re-routing from a system of systems perspective considering the interactions of multiple subsystems, and integrates the uncertainties arising from each subsystem (radar detection, re-routing assignment, and communication system). In this way, we are able to view this problem from a systemic point of view, have a comprehensive understanding of the system work flow, and characterize the relationship among the subsystems. At last, the construction of surrogate model improves the computational efficiency and facilitates fast analysis of system performance, thus supporting real-time decision-making.

Our work brings several benefits to the practitioners (airline dispatcher and air traffic controller) when extreme weather affects the original aircraft routes. With the aforementioned uncertainties considered, we provide quantitative analysis on the performance of aircraft re-routing operation. In this sense, the airline dispatcher is able to make risk-averse decisions given the system failure time distribution given a specific set of system parameters. In addition, since as the technology is moving towards the Next Generation Air Transportation System (NextGen), there is a huge demand for automatic operations in order to reduce the amount of information the air crew must process at one time [53,54]. This paper provides a case study of evaluating the performance of a re-routing algorithm under uncertainty, and the simulation platform we have built is extendable to test the performance of other re-routing algorithms with additional constraints imposed, e.g., separation assurance.

Future work can be carried out in several directions. The present work is based on several simplifications in the modeling of the simulation region, aircraft characteristics, space availability in neighboring airports, radar and the communication systems. However, the proposed methodology is quite general, and it is easy to replace the simple models with more sophisticated models of the system components. For example, with respect to the space availability in nearby airports, we model its dynamic update by using a lognormal distribution. A more realistic way is to look up the official flight time table of arrivals and departures at each nearby airport, then build a model to capture the dynamic change of space availability. Besides, different distributions are assumed for the random variables related to aircraft, airports, radar and communication. In the future, data mining techniques can be implemented to construct these distributions from available data sources, e.g., Airline On-Time Statistics and Delay Causes reported by the US Department of Transportation [55]. Finally, the reliability estimation carried out in this paper can be used towards reliability-based optimization of the system variables (e.g., airport capacity and message delay) in order to improve the system performance.

Appendix A. Censored data

In supervised learning, a set of labeled instances are available, where each instance consists of a data vector, e.g., features, explanatory variables, and a target variable (response variable). Depending on the specific type of the response variable, we can formulate different problems [56]:

- **Point Targets:** This is the classical regression problem, where each input vector $x_i \in \mathbb{R}^d$ has a point target $y_i \in \mathbb{R}$.
- **Binary Classification:** The binary class labels are usually denoted by $y_i \in \{\pm 1\}$, while the input vector $x_i \in \mathbb{R}$ is the same $x_i \in \mathbb{R}^d$.
- **Interval Targets:** An interval number is used to represent an upper bound and a lower bound on the target. For a given input vector $x_i \in \mathbb{R}^m$, $[l_i, u_i]$ denotes an interval target, where $l_i \in \mathbb{R}$, $u_i \in \mathbb{R}$, and $l_i < u_i$.
- **Survival Times:** In the context of failure testing, right-censored data arises when we do not observe a test subject to fail during the duration of the study; we know how long the subject survived, but we do not know what its failure time might be. For an observation x_i whose survival time is greater than $l_i \in \mathbb{R}$, we can denote its survival time as $[l_i, +\infty]$. The Kaplan-Meier estimator is a commonly used, non-parametric technique to handle such right-censored data [57]. In this estimator, at any given time t_i, the right-censored observation x_i is included in the system survival probability calculation if $t_i < l_i$, and is excluded if $t_i > l_i$. The system survival probability at any time instant t is computed as [58]

$$S(t) = \prod_{t_i \le t} \left(1 - \frac{d_i}{n_i}\right) \tag{A.1}$$

where d_i is the number of failures up to time t_i, and n_i is the number of observations included at time t_i.

Appendix B. Support vector regression

Support vector regression (SVR) pioneered by Vapnik [59,60], is a powerful tool for regression problems. Since SVR has greater generalization ability and guarantees global minima for given training data, it has drawn the attention of researchers and has been applied in many practical applications, e.g., financial time series forecasting [61], and travel time prediction [62]. Different from neural networks, SVR formulates the learning process as a quadratic programming optimization problem with linear constraints.

Consider a typical regression problem with a set of data (x_i, y_i), $i = 1, \ldots, N$, where x_i is a vector of the model inputs, y_i is the actual value. The objective of regression analysis is to determine a function $f(\boldsymbol{x})$ so as to predict the desired targets accurately. To resolve the regression problem, we define a cost function to measure the degree of discrepancy between the prediction and the actual value. A commonly used cost function in SVR is the robust ϵ-insensitive loss function L_ϵ, which is defined as [63]

$$L_\epsilon(f(\boldsymbol{x}), y) = \begin{cases} |f(\boldsymbol{x}) - y| - \epsilon & \text{if } |f(\boldsymbol{x}) - y| > \epsilon, \\ 0 & \text{otherwise.} \end{cases} \tag{B.1}$$

where ϵ is a user-prescribed parameter, which quantifies the approximation accuracy placed on the training data points. Fig. B.11 illustrates the concept of ϵ-insensitive SVR. As long as the deviation between prediction $f(x_i)$ and actual value y_i is less than ϵ, L_ϵ is 0. Otherwise, $L_\epsilon = |f(x_i) - y_i| - \epsilon$.

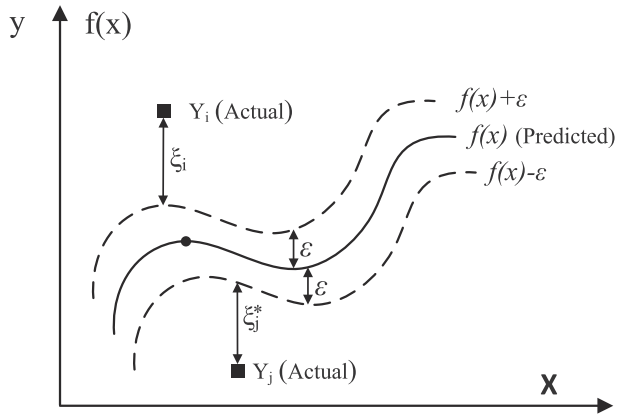


Fig. B.11. ϵ-insensitive SVR.

By minimizing the regularized risk function, we can estimate the unknown parameters $\boldsymbol{w}$ and b.

$$\min \quad C\frac{1}{N}\sum_{i=1}^{N} L_\epsilon(f(\boldsymbol{x_i}), y_i) + \frac{1}{2} \parallel \boldsymbol{w}\parallel^2 \tag{B.2}$$

where C is a user-determined parameter to control the trade-off between the prediction error and the complexity of the regression model.

Two positive slack variables, $\xi_i$andξ_i^*, can be used to quantify the derivation ($|f(\boldsymbol{x}) - y|$) from the boundaries of the ϵ-insensitive zone. Due to the slack variables, we can update Eq. (B.2) as follows:

$$\min \quad C\frac{1}{N}\sum_{i=1}^{N} (\xi_i + \xi_i^*) + \frac{1}{2}\parallel \boldsymbol{w} \parallel^2, \tag{B.3}$$

subject to

$$\begin{cases} y_i - (\boldsymbol{w}\cdot\phi(\boldsymbol{x_i}) + b) \le \epsilon + \xi_i, \\ \boldsymbol{w}\cdot\phi(\boldsymbol{x_i}) + b - y_i \le \epsilon + \xi_i^*, \\ \xi_i, \xi_i^* \ge 0, \quad i = 1, \ldots, N. \end{cases}$$

With the help of Lagrange multipliers and KKT conditions, Eq. (B.3) yields the following dual Lagrangian form [63]:

$$\begin{aligned} \max \quad & -\epsilon\sum_{i=1}^{N}(\alpha_i + \alpha_i^*) + \sum_{i=1}^{N}(\alpha_i^* - \alpha_i) y_i \\ & - \frac{1}{2}\sum_{i,j=1}^{N}(\alpha_i^* - \alpha_i)\left(\alpha_j^* - \alpha_j\right)\boldsymbol{K}(\boldsymbol{x_i}, \boldsymbol{x_j}), \end{aligned} \tag{B.4}$$

subject to

$$\begin{cases} \sum_{i=1}^{N}(\alpha_i - \alpha_i^*) = 0, \\ 0 \le \alpha_i, \alpha_i^* \le C, \quad i = 1, \ldots, N. \end{cases}$$

where $\alpha_i$andα_i^* are the optimum Lagrange multipliers, and $\boldsymbol{K}$ is the kernel function defined as

$$\boldsymbol{K}(\boldsymbol{x_1}, \boldsymbol{x_2}) = \phi(\boldsymbol{x_1})\,\phi(\boldsymbol{x_2}). \tag{B.5}$$

The optimal weight vector of the regression hyperplane is $\boldsymbol{w}^* = \sum_{i=1}^{N}(\alpha_i - \alpha_i^*)\boldsymbol{x_i}$. Hence, the general form of the SVR-based regression function can be written as

$$f(\boldsymbol{x}, \boldsymbol{w}) = f(\boldsymbol{x}, \alpha, \alpha^*) = \sum_{i=1}^{N}(\alpha_i - \alpha_i^*)\boldsymbol{K}(\boldsymbol{x}, \boldsymbol{x_i}) + b. \tag{B.6}$$

Appendix C. Global sensitivity analysis

Global sensitivity analysis is used to quantify the influence of stochastic model inputs on the output variability of a physical or mathematical model [64]. Among the abundant literature on sensitivity measures, Sobol' indices based on variance decomposition have received much attention. Consider the model

$$y = f(\boldsymbol{x}) \tag{C.1}$$

where y is the output, $\boldsymbol{x} = (x_1, \cdots, x_p)$ are p independent input variables, f is the model function. These contributions of the variance of a single model input or a group of model inputs to the output variance of f are quantified using the following sensitivity indices:

$$S_i = \frac{V_{x_i}\left(\mathbb{E}_{x_{-i}}(y|x_i)\right)}{V(y)}, \; S_{ij} = \frac{V_{x_i x_j}\left(\mathbb{E}_{x_{-i}x_{-j}}(y|x_i x_j)\right)}{V(y)} - S_i - S_j \tag{C.2}$$

Here, S_i is the first order Sobol' index quantifying the sensitivity of output variance to an individual variable x_i by itself (individual effect). The second order index S_{ij} quantifies the sensitivity to the interaction between variable x_i and x_j. The larger an index value is, the greater is the importance of the corresponding variable or group of variables.

According to the theorem of variance decomposition, we have

$$V(y) = \mathbb{E}_{x_i}\left(V_{x_{-i}}(y|x_i)\right) + V_{x_i}\left(\mathbb{E}_{x_{-i}}(y|x_i)\right) \tag{C.3}$$

which implies that

$$S_i = 1 - \frac{\mathbb{E}_{x_i}\left(V_{x_{-i}}(y|x_i)\right)}{V(y)} \tag{C.4}$$

A double-loop Monte Carlo simulation is required to estimate Sobol' indices S_i. The inner loop $V_{x_{-i}}(y|x_i)$ in Eq. (C.4) requires fixing x_i and changing all the other variables x_{-i}. The outer loop $\mathbb{E}_{x_i}(\bullet)$ requires fixing x_i at different locations, and these selected locations are sampled from the distribution of x_i.

References

[1] V.A. Marchau, W.E. Walker, G. Van Wee, Dynamic adaptive transport policies for handling deep uncertainty, Technol. Forecast. Soc. Chang. 77 (6) (2010) 940–950.
[2] A.I. Czerny, Airport congestion management under uncertainty, Transp. Res. B Methodol. 44 (3) (2010) 371–380.
[3] S.J. Landry, X.W. Chen, S.Y. Nof, A decision support methodology for dynamic taxiway and runway conflict prevention, Decis. Support. Syst. 55 (1) (2013) 165–174.
[4] Deadly winter storm hits Northeast: the latest, http://nypost.com/2014/02/13/live-updates-on-nycs-winter-storm/. Accessed: 2015-04-14.
[5] Rain, wind and snow causing delays at several airports around the country, http://fox6now.com/2014/12/24/ugly-weather-hits-holiday-travelers/. Accessed: 2015-04-14.
[6] S. Ahmadbeygi, A. Cohn, M. Lapp, Decreasing airline delay propagation by re-allocating scheduled slack, IIE Trans. 42 (7) (2010) 478–489.

[7] A. Agustı, A. Alonso-Ayuso, L.F. Escudero, C. Pizarro, et al. On air traffic flow management with rerouting. Part II: stochastic case, Eur. J. Oper. Res. 219 (1) (2012) 167–177.
[8] D. Bertsimas, G. Lulli, A. Odoni, An integer optimization approach to large-scale air traffic flow management, Oper. Res. 59 (1) (2011) 211–227.
[9] Airport capacity needs in the National Airspace System, https://www.faa.gov/airports/planning_capacity/media/FACT3-Airport-Capacity-Needs-in-the-NAS.pdf. Accessed: 2016-12-05.
[10] A. Mukherjee, M. Hansen, A dynamic rerouting model for air traffic flow management, Transp. Res. B Methodol. 43 (1) (2009) 159–171.
[11] J. Xiong, M. Hansen, Value of flight cancellation and cancellation decision modeling: ground delay program postoperation study, Transp. Res. Rec. J. Transp. Res. Board (2106) (2009) 83–89.
[12] B.G. Thengvall, J.F. Bard, G. Yu, Balancing user preferences for aircraft schedule recovery during irregular operations, IIE Trans. 32 (3) (2000) 181–193.
[13] M.E. Berge, M.L. Carter, A. Haraldsdottir, B. Repetto, L. Kang, Airline schedule recovery in flow management: an application for departure re-routing, 25th Digital Avionics Systems Conference, 2006 IEEE/AIAA, IEEE. 2006, pp. 1–9.
[14] V. Chiraphadhanakul, C. Barnhart, Robust flight schedules through slack re-allocation, EURO J. Transp. Logist. 2 (4) (2013) 277–306.
[15] A. D'Ariano, M. Pistelli, D. Pacciarelli, Aircraft retiming and rerouting in vicinity of airports, IET Intell. Transp. Syst. 6 (4) (2012) 433–443.
[16] J. Clausen, A. Larsen, J. Larsen, N.J. Rezanova, Disruption management in the airline industry concepts, models and methods, Comput. Oper. Res. 37 (5) (2010) 809–821.
[17] A.R. Odoni, The flow management problem in air traffic control, Flow Control of Congested Networks, Springer. 1987, pp. 269–288.
[18] D. Bertsimas, S.S. Patterson, The air traffic flow management problem with enroute capacities, Oper. Res. 46 (3) (1998) 406–422.
[19] D. Bertsimas, S.S. Patterson, The traffic flow management rerouting problem in air traffic control: a dynamic network flow approach, Transp. Sci. 34 (3) (2000) 239–255.
[20] P.L. de Matos, R. Ormerod, The application of operational research to European air traffic flow management-understanding the context, Eur. J. Oper. Res. 123 (1) (2000) 125–144.
[21] P.L. de Matos, P. Powell, Decision support for flight re-routing in Europe, Decis. Support. Syst. 34 (4) (2003) 397–412.
[22] J.F. Bard, G. Yu, M.F. Arguello, Optimizing aircraft routings in response to groundings and delays, IIE Trans. 33 (10) (2001) 931–947.
[23] C. van Balen, C. Bil, Optimal re-routing of aircraft around closed airspace in free flight, 26th American Institute of Aeronautics and Astronautics Applied Aerodynamics Conference, American Institute of Aeronautics and Astronautics. 2008, pp. 1–10.
[24] A.M. Churchill, D.J. Lovell, M.O. Ball, Evaluating a new formulation for large-scale traffic flow management, Proceedings of the 8th USA/Europe Air Traffic Seminar (ATM09), 310, 2009.
[25] A. Agustı, A. Alonso-Ayuso, L.F. Escudero, C. Pizarro, On air traffic flow management with rerouting. Part II: stochastic case, Eur. J. Oper. Res. 219 (1) (2012) 167–177.
[26] J.M. Rosenberger, E.L. Johnson, G.L. Nemhauser, Rerouting aircraft for airline recovery, Transplant. Sci. 37 (4) (2003) 408–421.
[27] M. Kamgarpour, V. Dadok, C. Tomlin, Trajectory generation for aircraft subject to dynamic weather uncertainty, Decision and Control (CDC), 2010 49th IEEE Conference on, IEEE. 2010, pp. 2063–2068.
[28] G. Clare, A. Richards, Air traffic flow management under uncertainty: application of chance constraints, Proceedings of the 2nd International Conference on Application and Theory of Automation in Command and Control Systems, IRIT Press. 2012, pp. 20–26.
[29] N. Pyrgiotis, K.M. Malone, A. Odoni, Modelling delay propagation within an airport network, Transp. Res. C Emerg. Technol. 27 (2013) 60–75.
[30] F. Balbo, S. Pinson, Dynamic modeling of a disturbance in a multi-agent system for traffic regulation, Decis. Support. Syst. 41 (1) (2005) 131–146.
[31] S. Pathak, M. McDonald, S. Mahadevan, A framework for designing policies for networked systems with uncertainty, Decis. Support. Syst. 49 (2) (2010) 121–131.
[32] S.W. Yoon, J.D. Velasquez, B. Partridge, S.Y. Nof, Transportation security decision support system for emergency response: a training prototype, Decis. Support. Syst. 46 (1) (2008) 139–148.
[33] M. Prandini, L. Piroddi, S. Puechmorel, S.L. Brázdilová, Toward air traffic complexity assessment in new generation air traffic management systems, IEEE Trans. Intell. Transp. Syst. 12 (3) (2011) 809–818.
[34] B.P. Zeigler, H. Praehofer, T.G. Kim, Theory of Modeling and Simulation: Integrating Discrete Event and Continuous Complex Dynamic Systems, Academic Press. 2000.
[35] Controller-pilot data link communications, https://en.wikipedia.org/wiki/Controller-pilot_data_link_communications. Accessed: 2016-09-24.
[36] E.M. Rantanen, J.S. McCarley, X. Xu, Time delays in air traffic control communication loop: effect on controller performance and workload, Int. J. Aviat. Psychol. 14 (4) (2004) 369–394.
[37] F.T. Durso, C.A. Manning, Air traffic control, Rev. Hum. Factors Ergon. 4 (1) (2008) 195–244.
[38] A. Haldar, S. Mahadevan, Probability, Reliability and Statistical Methods in Engineering Design, John Wiley & Sons, New York, 2000.
[39] S.W. Henriksen, Radar-range equation, IEEE Proceedings, 63, 1975. pp. 813.
[40] B.L. Bosart, W.-C. Lee, R.M. Wakimoto, Procedures to improve the accuracy of airborne Doppler radar data, J. Atmos. Ocean. Technol. 19 (3) (2002) 322–339.
[41] J. Farrell, Integrated Aircraft Navigation, Elsevier. 2012.
[42] F. Gustafsson, F. Gunnarsson, N. Bergman, U. Forssell, J. Jansson, R. Karlsson, P.-J. Nordlund, Particle filters for positioning, navigation, and tracking, IEEE Trans. Signal Process. 50 (2) (2002) 425–437.
[43] K. Sampigethaya, R. Poovendran, L. Bushnell, A framework for securing future e-enabled aircraft navigation and surveillance, AIAA Proceedings, 2009. pp. 1–10.
[44] ADSB over satellite, https://directory.eoportal.org/web/eoportal/satellite-missions/a/ads-b. Accessed: 2016-12-13.
[45] Controller-pilot communications, http://virtualskies.arc.nasa.gov/communication/11.html. Accessed: 2016-02-16.
[46] M.C. Bartholomew-Biggs, S.C. Parkhurst, S.P. Wilson, Global optimization approaches to an aircraft routing problem, Eur. J. Oper. Res. 146 (2) (2003) 417–431.
[47] M. Haouari, N. Aissaoui, F.Z. Mansour, Network flow-based approaches for integrated aircraft fleeting and routing, Eur. J. Oper. Res. 193 (2) (2009) 591–599.
[48] C. Li, S. Mahadevan, An efficient modularized Monte Carlo method to estimate the first order Sobol' index, Reliab. Eng. Syst. Saf. 153 (2016) 110–121.
[49] B. Iooss, M. Ribatet, Global sensitivity analysis of computer models with functional inputs, Reliab. Eng. Syst. Saf. 94 (7) (2009) 1194–1204.
[50] R. Rackwitz, Reliability analysis — a review and some perspectives, Struct. Saf. 23 (4) (2001) 365–395.
[51] V. Cherkassky, Y. Ma, Practical selection of SVM parameters and noise estimation for SVM regression, Neural Netw. 17 (1) (2004) 113–126.
[52] R.B. D'Agostino, Goodness-of-fit-techniques, 68. CRC Press. 1986.
[53] T. Prevot, J. Homola, J. Mercer, Initial study of controller/automation integration for NextGen separation assurance, AIAA Guidance, Navigation, and Control (GNC) Conference and Exhibit, 2008.
[54] J. Mercer, J. Homola, C. Cabrall, L. Martin, S. Morey, A. Gomez, T. Prevôt, Human-automation cooperation for separation assurance in future NextGen environments, Proceedings of the International Conference on Human-Computer Interaction in Aerospace, ACM. 2014, pp. 1.
[55] Airline on-time statistics and delay causes, http://www.transtats.bts.gov/ot_delay/ot_delaycause1.asp. Accessed: 2016-03-24.
[56] P.K. Shivaswamy, W. Chu, M. Jansche, A support vector approach to censored targets, Data Mining, 2007. ICDM 2007. Seventh IEEE International Conference on, IEEE, 2007. pp. 655–660.
[57] M. Goel, P. Khanna, J. Kishore, Understanding survival analysis: Kaplan-Meier estimate, Int. J. Ayurveda Res. 1 (4) (2010) 274.
[58] E.L. Kaplan, P. Meier, Nonparametric estimation from incomplete observations, J. Am. Stat. Assoc. 53 (282) (1958) 457–481.
[59] V.N. Vapnik, V. Vapnik, Statistical Learning Theory, 1. Wiley New York. 1998.
[60] V.N. Vapnik, An overview of statistical learning theory, IEEE Trans. Neural Netw. 10 (5) (1999) 988–999.
[61] C.-J. Lu, T.-S. Lee, C.-C. Chiu, Financial time series forecasting using independent component analysis and support vector regression, Decis. Support. Syst. 47 (2) (2009) 115–125.
[62] C.-H. Wu, J.-M. Ho, D.-T. Lee, Travel-time prediction with support vector regression, IEEE Trans. Intell. Transp. Syst. 5 (4) (2004) 276–281.
[63] V. Vapnik, The Nature of Statistical Learning Theory, Springer Science & Business Media. 2013.
[64] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, S. Tarantola, Global Sensitivity Analysis: The Primer, John Wiley & Sons. 2008.

Xiaoge Zhang received the B.S. degree from Chongqing University of Posts and Telecommunications in 2011, and the M.S. degree from Southwest University, Chongqing, China, in 2014. Right now, he is pursuing towards Ph.D. degree in Vanderbilt University. He has published more than 10 papers in international peer-reviewed journals, such as Annals of Operations Research, Reliability Engineering and System Safety, International Journal of Production Research, Scientific Reports, Expert Systems with Applications, Applied Soft Computing, and Applied Intelligence .

Sankaran Mahadevan is a John R. Murray Sr. Professor of Engineering, and Professor of Civil and Environmental Engineering at Vanderbilt University, Nashville, Tennessee, where he has served since 1988. His research interests are in the areas of uncertainty quantification, model verification and validation, reliability and risk analysis, design optimization, system health monitoring, and integrated multi-physics simulation, with applications to civil, mechanical and aerospace systems. His research has been extensively funded by NSF, NASA, FAA, DOE, DOD, DOT, NIST, GE, GM, Chrysler, Union Pacific, American Railroad Association, and Sandia, Idaho, Los Alamos and Oak Ridge National Laboratories. His research contributions are documented in more than 500 publications, including two textbooks on reliability methods and 220 journal papers. He has directed 40 Ph.D. dissertations and 24 M. S. theses, and has taught several industry short courses on reliability and risk analysis methods. His awards include the NASA Next Generation Design Tools award (NASA), the SAE Distinguished Probabilistic Methods Educator Award, SEC Faculty Award, and best paper awards in the MORS Journal and the SDM and IMAC conferences. Professor Mahadevan obtained his B.S. from Indian Institute of Technology, Kanpur, M.S. from Rensselaer Polytechnic Institute, Troy, NY, and Ph.D. from Georgia Institute of Technology, Atlanta, GA.