A Stochastic Spatiotemporal Weather-Impact Simulator: Representative Scenario Selection

Article · September 2011

DOI: 10.2514/6.2011-6812

Sandip Roy
Washington State University

Yan Wan
University of Texas at Arlington

Christine Taylor
MITRE

Craig R Wanke
MITRE

11th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference, including the AIA
20 - 22 September 2011, Virginia Beach, VA

AIAA 2011-6812

A Stochastic Spatiotemporal Weather-Impact Simulator: Representative Scenario Selection

Mengran Xue[1], **Sandip Roy**[2]
School of EECS, Washington State University, Pullman, WA, 99163

Stephen Zobell[3],
Center for Advanced Aviation Systems Development, The MITRE Corporation, McLean, VA, 22102

Yan Wan[4],
Department of Electrical Engineering, University of North Texas, Denton, TX, 76207

Christine Taylor[5], **Craig Wanke**[6]
Center for Advanced Aviation Systems Development, The MITRE Corporation, McLean, VA, 22102

I. Introduction

The Next Generation Air Transportation System (NextGen) concept includes a new strategic traffic flow management function called Flow Contingency Management (FCM). In FCM, automation will predict a set of possible situations requiring traffic management action, and classify them such that contingency plans can be developed to manage them. These situations most frequently involve severe weather systems, and severe weather forecasts contain a great deal of uncertainty. Thus, we have been working on effective ways to identify the range of weather-induced ATM situations that could occur given a forecast.

In our recent work Ref. 22, we introduced a promising tool for simulating weather impact on air traffic, that our group is developing as part of a comprehensive solution for FCM for NextGen. This simulator can rapidly and simply generate stochastic scenarios of weather impact that statistically match probabilistic weather forecasts at snapshot times, and seems capable of capturing the rich, uncertain dynamics observed in weather over a one-day time horizon. As such, we believe that the simulator holds promise to serve as a key component of the weather-forecasting capabilities envisioned for NextGen, and more broadly to inform new decision-support capabilities for the NAS that can achieve coordinated flow management at a strategic scale (NAS-wide spatial extent, 2-24 hour temporal horizon).

Our ongoing development of the weather-impact simulator has been documented previously (Ref. 15, 22). These introductory studies draw on foundational research in the air traffic management community (Ref. 7, 11, 18), ongoing probabilistic weather-forecast research and development (Ref. 4, 23), and stochastic network modeling efforts (Ref. 2, 20, 21). Specifically, the concept of modeling spatial weather-impact rather than weather originates from two significant foundational works Ref. 11 and Ref. 16. In our first effort Ref. 15, we motivated the need for a stochastic spatial weather-impact *simulator* (a tool that generates *time-trajectories* of weather-impact in airspace regions), introduced in concept an *influence-model*-based methodology for weather-impact simulation, and developed several illustrative simulation examples. In recent work Ref. 22, we made explicit the sense in which the weather-impact simulator extends current ensemble/probabilistic weather-forecasting capabilities toward meeting air traffic needs, and showed how to parameterize the weather-impact simulator to match probabilistic-forecast statistics at particular times while tuning spatial correlation. This parameterization effort builds on our ongoing research on

---

[1] Graduate Student, School of EECS, Washington State University, Pullman, WA, AIAA Student Member.
[2] Associated Professor, School of EECS, Washington State University, Pullman, WA, AIAA Member.
[3] Lead Engineer, CAASD, The MITRE Corporation, 7515 Colshire Drive, McLean, VA, 22102, AIAA Senior Member.
[4] Assistant Professor, Department of EE, University of North Texas, Denton, TX, 76207, AIAA Member.
[5] Lead Simulation Modeling Engineer, CAASD, The MITRE Corporation, 7515 Colshire Drive, McLean, VA, 22102, AIAA Member.
[6] Senior Principal Engineer, CAASD, The MITRE Corporation, 7515 Colshire Drive, McLean, VA, 22102, AIAA Senior Member.

design/inference in influence models and other stochastic-network models (Ref. 20, 21). An outcome of Ref. 22 is an example weather-impact simulation of a particular bad weather day (September 26, 2010) in the Southeastern United States, that is matched with the convection probability product of the *Short Range Ensemble Forecast (SREF)* for that day (Ref. 3, 5). Using the parameterized simulator, several potential scenarios of weather impact were obtained. In our two previous works Ref. 15 and Ref. 22, we also enumerated numerous tractabilities afforded by the weather-impact model, including the fast computation of weather-impact statistics (including spatial correlations) and the development of lower-order models for critical airspace locations.

While the weather-impact simulator is promising as a weather-modeling capability for strategic traffic management, and more broadly for the NextGen system, several further avenues of research must be pursued to permit these applications of the model. We view two research tasks as being of particular importance:

1) To facilitate the simulator's use in decision making, we require a procedure for selecting a few representative scenarios among the large number of generated scenarios, and assigning probabilities (chances) to each representative scenario. Essentially, to capture the wide range of possible weather outcomes in a way that is useful for decision-making (manually or automatically), these potential outcomes need to be distilled into evocative representative cases with associated probabilities. These representative cases highlight potential weather features for which flow contingencies need to be designed, and hence facilitate planning.
2) The weather-impact simulator is structured to permit simple representation of weather-impact, and its uncertainty, at the time-scale and resolution needed for FCM. Since such a simple abstraction is used to represent an incredibly complex process, it is important to evaluate whether or not the model is in fact predictive of weather-impact and its uncertainty. We thus need to pursue comprehensive validation of the weather-impact simulator.

This purpose of this article is to address the first of these two outstanding needs, namely representative scenario selection. Specifically, we describe research challenges in representative scenario selection in some detail, and so motivate a signal-clustering-based viewpoint on the problem. We advocate for methods that cluster scenarios and identify representative ones according to NAS-performance-related metrics or comparisons, and introduce two specific approaches that show promise. The methodologies are used to find representative weather-impact scenarios for a particular bad-weather event, namely for a tropically-driven convective weather event that strongly impact traffic in the airspace managed by the Atlanta Air Route Traffic Control Center (henceforth simply called Atlanta Center or ZTL) on September 26, 2010.

The remainder of the article is organized as follows:

- We will briefly review the formulation and parameterization of the weather-impact simulator (Section II), and review products of the weather-impact simulator generated for the ZTL weather event.
- We will provide a comprehensive treatment of the representative scenario selection problem, and describe two different methods – namely, a Gaussian quadrature-based method and a signal comparison-based method – for solving the problem (Section III).
- In Section IV, we will apply these two methods to the ZTL weather example.
- Finally, Section V provides some brief conclusions about representative scenario selection.

II. Weather-Impact Simulation: A Brief Review

Strategic flow management for the NAS critically requires modeling of weather uncertainty and its impact on NAS constraints. While available weather forecast tools provide some probabilistic information about weather (e.g., convective-weather probability forecasts), several challenges remain in modeling weather-impact in the NAS, including:

1. Translating weather characteristics to quantitative impacts on NAS parameters, including Sector capacities and Airport Arrival Rates (AARs).
2. Characterizing *trajectories* (temporal dynamics) of weather and weather-impact, rather than only snapshots.
3. Capturing the inherent uncertainty in weather propagation at a multi-hour time horizon.
4. Developing tools that permit fast simulation of weather- and weather-impact, and fast computation of spatial and temporal weather-impact statistics.

To address these challenges, our group has pursued development of a weather-impact (WI) simulator, as one key component in our comprehensive flow contingency management solution (please see our companion paper Ref. 24). The key purpose of this weather-impact simulator is the generation of many spatiotemporal trajectories (scenarios) of weather impact, for use in characterizing and designing flow-management capabilities that effectively use NAS resources. The conceptual development, parameterization, and use of the model (for scenario generation and statistical analysis) has been detailed in a series of papers (Ref. 15, 20, 22), and is briefly summarized in our overview of FCM Ref. 24. The outputs of the weather-impact simulator –in particular, an ensemble of weather-impact scenarios – serve as the starting point for selecting representative scenarios, as described in Section III. It is worth noting that our representative scenario selection methods do not draw on the internal workings of the WI simulator (i.e., the simulator is a black box which simply generates scenarios from which representative ones are selected); hence, details about the weather-impact simulator's internal workings are not necessary for understanding the representative scenario selection methods described here. However, for the reader's convenience and to give a better perspective on the research presented here, we include verbatim the brief summary of the WI simulator's development given in our companion paper Ref. 24.

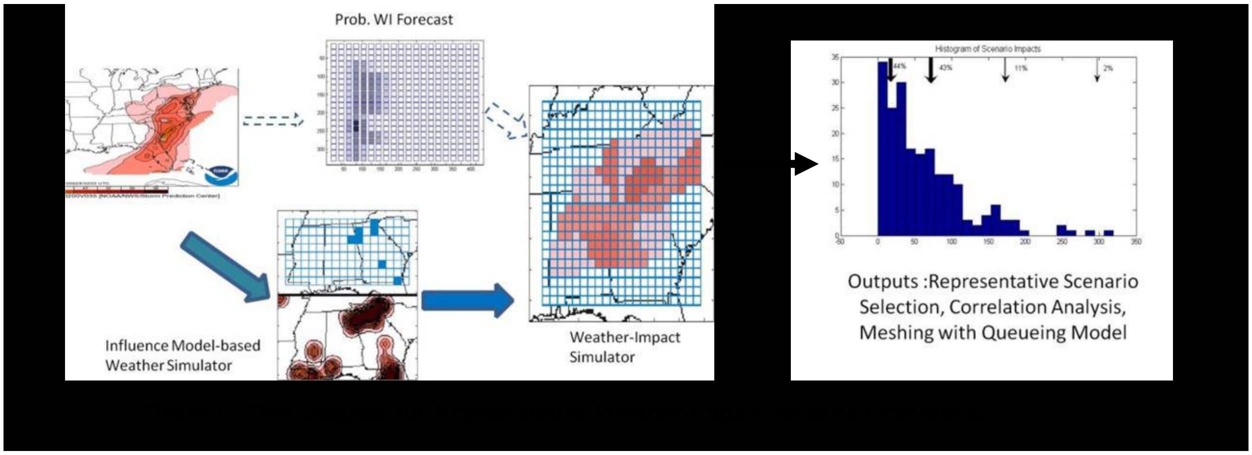
The focus of our earlier efforts (Ref. 15, 22) has been to 1) propose a promising architecture for the WI simulator, 2) develop a methodology for building (parameterizing) the simulator for a particular day and region from existing forecast products. In particular, we have advanced the influence model, a stochastic network model that describes propagation of discrete-valued quantities on a graph, as a means for capturing propagation of convective and other weather on a grid. The influence model is promising for representing weather and weather-impact propagation, in that it can naturally represent complex temporal propagations of statuses in a way that permits both fast simulation and significant analysis.

The parameters of the influence model - which govern initiation, movement, and decline of convective weather events - can be selected so that probabilities of convective weather in individual grid squares match probabilistic weather forecasts at snapshot times. The parameterization of the model requires data from current probabilistic weather forecasts for the desired simulation duration (the planning horizon). Such probabilistic forecasts (including ensemble forecasts) typically use a combination of physics-based models and extrapolation techniques to generate probabilities of weather in individual grid squares at snapshot time, but do not themselves provide probabilistic spatiotemporal trajectories of weather. The influence model can be parameterized (as described in Ref. 22) so that generated weather trajectories statistically match the probabilistic forecasts at the snapshot times, but also have realistic spatial and temporal correlation and interpolate between the snapshot times. In our initial efforts, we have used the hourly convective-weather probability forecast associated with the Short Range Ensemble Forecast (SREF) tool (which derives from the Rapid Update Cycle or RUC forecasting suite) for parameterization.

Upon parameterization, the influence model can be used to rapidly generate trajectories (scenarios) of weather propagation. In turn, each influence-model-generated trajectory can be translated to a weather-impact scenario capturing evolution of NAS constraints (Sector capacities, AARs, etc.), using existing research on the impact of weather on these constraints. Once built, the WI simulator permits 1) fast generation of a large number of weather-impact scenarios; 2) generation of a few representative scenarios with associated probabilities; 3) analysis of temporal and spatial correlations in weather impact among other statistics; and 4) interface with the queuing model for traffic flow.

For FCM, the generation of representative scenarios with associated probabilities using the WI simulator is of particular importance, because these special scenarios can potentially guide the design of contingencies. With this motivation in mind, we have put significant effort into developing solutions for the representative-scenario-selection problem. A fruitful approach is to first pursue clustering of scenarios according to one or metrics that that are germane to NAS performance/management (e.g., capacity reduction at critical airports). Then, a representative element of each cluster can be selected as a representative scenario, with associated probability given by the fraction of all scenarios that are in the cluster. A point-selection technique that is traditionally used for numerical integration provides a particularly useful tool for clustering and finding representative scenarios in clusters. In addition to studying representative scenario selection, we have pursued temporal/spatial analysis of weather impact using moment-linearity properties of the influence model, and have begun to develop reduced-order models for critical weather impacts for interfacing with the queuing model.

The WI simulator is diagrammed in Fig. 1. It is worth noting that, although we have here advocated generating a weather simulator from available forecasts and then translating weather trajectories to weather-impact ones, the alternate approach of translating weather forecasts to weather impact forecasts first and then directly generating a weather-impact forecast is also feasible. Both approaches are diagrammed in the figure.



III. Selection of Representative Scenarios: Preliminary Results

The weather-impact simulator generates a very large number of weather-impact scenarios or trajectories. The NextGen strategic traffic management function, termed Flow Contingency Management (FCM), is premised on the ability to predict a range of contingencies that must be prepared for. Thus, for FCM automation design and for display to NAS operators, it is desirable to distill the possible outcomes of the weather-impact simulator to a small number of *representative scenarios* (with associated chances or probabilities), that span the possible weather-impact scenarios. Here, we introduce and apply two promising methodologies for generating representative scenarios and associated probabilities from a sample-set of trajectories (scenarios). The first methodology that we propose is a metric-value-grouping-based scheme that draws on a mathematical notion known as *Gaussian quadrature*. This methodology allows us to select representative scenarios that 1) faithfully capture statistics of weather-impact metric(s) for the whole ensemble and 2) are ideally suited for exposing the dependence of NAS performance metrics on weather-impact metrics. Meanwhile, the second methodology that we propose is based on signal-comparison-based clustering of scenarios. This approach is well-suited for allowing complex comparisons of scenarios in determining clusters and representative scenarios, including comparisons based on multiple or multi-faceted measures. Let us here first motivate both approaches broadly by viewing representative-scenario-selection as a clustering problem (Section III-A), and then describe the two approaches in some detail (Sections III-B and III-C).

A. **Motivation: A Clustering Viewpoint**

To motivate the representative-scenario-selection methodologies, let us envision the desired outcomes of a selection tool. We would like a selection tool to take in (or generate) a moderate number of weather-impact scenarios, and output 1) a short list of representative scenarios and 2) a probability associated with each scenario. Ideally, these outputs of the selection tool should have the following properties:

1) Each representative scenario should capture a significantly-different weather outcome with respect to one or more weather-impact performance metrics.
2) The representative scenarios together should adequately span the possible scenarios generated by the simulator (again according to the performance metrics). That is, we should be able to *cluster* all (or almost all) of the scenarios with the representative scenarios, such that each cluster associated with each scenario is well-described by the representative one.
3) The probability associated with each representative scenario should indicate the chance that the generated weather-impact scenario (and hence the actual weather-impact trajectory) is close to the representative one, i.e. in its cluster.
4) Each representative scenario should be an actual scenario generated by the simulator.
5) The selected representative scenarios should be predictive, in the sense that their use in traffic-flow analysis and management design should give adequate predictions of flow and of relationships between weather-impact and traffic metrics.

The representative-scenario-selection task detailed above is an example of a *clustering* problem. In fact, a very wide range of methods for clustering data, including time-series data, have been developed in the engineering and computer science literature, see Ref. 13, 1, 17 for just a few representative publications. Although clustering has been quite exhaustively studied, our problem presents some new features which differ from the typical efforts in the time-series clustering literature. First, we are interested in clustering spatiotemporal (rather than only temporal) signals. Second, the focus in our clustering effort is selecting the representative scenarios (and associated probabilities), in addition to finding the clusters themselves: of particular note, the representative scenario must itself be a simulation output, not simply a measure of the center of the cluster group. Third, it is worth stressing that we are clustering simulation outputs, not actual data: more scenarios can be generated as desired, albeit with some computational cost; from this viewpoint, it is also worth noting that clustering could potentially be achieved from the simulation tool itself (rather than from considering its outputs), though we do not take this approach here. Fourth, from the perspective of FCM as a whole, the representative scenarios are valuable only so far as they aid in evaluating and designing flow contingencies. As such, the representative scenario selection should be based on metrics that are at least roughly indicative of NAS performance, and further scenarios selection should be structured to facilitate prediction of overall NAS performance under weather uncertainty.

B. A Gaussian Quadrature-Based Tool for Representative Scenario Selection

Given the distinct features of the representative-scenario-selection task, we propose a new approach to scenario-selection based on Gaussian quadrature. Very broadly, quadrature methods are tools for sampling continuous functions so as to permit accurate computation of integrals defined from these functions (Ref. 9). Here, we will apply quadrature techniques to probability distributions (Ref. 10, 19). In this scope, quadrature methods can be viewed as sampling or discretizing continuous random variable(s), and hence coming up with representative outcomes and associated probabilities. These representative outcomes (or discretizations) have the following properties: 1) their statistics match those of the original random variables and 2) complex mappings between the random variable values (in our case, the weather-impact metrics) and dependent variables (e.g., delays) can be obtained with high fidelity through evaluation at the discrete points.

Gaussian quadrature is a special selection method among various quadrature methods that is optimal in the sense that the highest-possible order of statistics are faithfully captured for a given number of points (e.g., Ref. 19). Specifically, statistics up to order 2n-1 are captured when n points are used. We believe that Gaussian quadrature is apt in selecting representative instances of random variables and associated probabilities in the senses that we desire, namely to faithfully span the set of possible outcomes accurately enough to match the random variable's statistics, and to permit prediction of consequent outcomes like traffic delays. We thus propose a representative-scenario-selection methodology, that is based on applying Gaussian quadrature point selection to probability distributions of important weather-impact metrics. The following is a detailed description of the quadrature-based methodology for selecting representative scenarios and associating probabilities with them:

Step 1: Choose one or more weather-impact metrics that will be used for comparison and clustering of scenarios. Broadly, we will consider metrics of two types. First, we will consider metrics that are aggregate statistics, e.g., the total capacity lost during the simulation duration, the total time that a set of routes are all blocked, or the excess nominal demand compared to capacity in the airspace. Second, we will consider metrics that measure impact at critical airspace locations, for example total capacity reduction at an airport or in a critical airspace region during a busy period. We envision that both aggregate and local measures must be used in conjunction to appropriately distinguish scenarios and hence select representative ones.

In general, metric selection may be guided by experience, or the metric(s) may be chosen systematically in such a way that generated representative scenarios are well-suited for evaluating and designing flow-management contingencies. In the example developed here, we will simply choose a couple of plausible metrics based on our experience. However, a brief discussion of how metrics can be chosen systematically/automatically to optimize scenario-selection is worthwhile. Recalling that selected representative scenarios will be used for evaluation and design of flow contingencies, we contend that appropriate metrics for scenario selection are ones that, together, are *predictive* of NAS performance upon application of flow management. If the metrics are predictive of NAS performance, then the Gaussian quadrature-based methodology can aptly capture the mapping between the uncertain weather and NAS performance over the range of plausible uncertain weather, and hence should facilitate analysis/design of flow contingencies. Unfortunately, choosing metrics that are good predictors of NAS performance is complicated, since we wish to avoid real-time simulation of complex NAS dynamics – in fact, one of the primary reasons for using representative scenarios is to avoid simulation of NAS dynamics for a large number of possible weather scenarios. Instead, we envision choosing metrics that are good predictors in two ways: 1) based on historical precedent regarding the metric's predictive capability; or 2) based on quick, approximate sensitivity and/or uncertainty analyses of appropriate traffic flow models.

Step 2: Generate many scenarios using the weather-impact simulator, and compute the value(s) of the metric(s) for each scenario. From these generated scenarios and corresponding metrics, compute *sample statistics* and a *sample (joint) probability distribution function* for the metrics. We note that the particular sample statistics needed for Gaussian quadrature-based point selection depends on the desired number of points (representative scenarios), and the dimension of the metric space. The required statistics are tabulated in the quadrature literature; we kindly ask the reader to see e.g. Ref. 9, 19 for details.

We also refer the reader to Section II, and references described therein, for background on the scenario-generation methodology. Computationally, scenario generation requires a one-time parameterization of the influence model-based simulator, followed by weather-impact trajectory generation using the parameterized simulator. The simple structure of the influence model permits fast trajectory generation. We note that the number of scenarios needed for accurate computation of sample statistics and probability distributions depends on 1) the dimension of the metric space and 2) the number of statistics required (which depends on the desired number of representative scenarios). Classical statistical techniques can be used to find the required number of scenarios; we omit the details.

Step 3: Based on the sample statistics, apply Gaussian quadrature-based point selection to determine representative value sets for the metric(s), and probabilities associated with each representative metric value set. The selection of Gaussian quadrature points, and the association of probabilities with each point (assuming that the sampled function is a probability distribution), is extensively discussed in the quadrature and simulation literatures for a scalar metric (e.g., Ref. 9). For vector metrics, several alternative point-selection approaches have been proposed, depending on the desired number of points and on whether or not the different metrics are statistically independent (e.g., Ref. 25). However, even for the vector-metric case, the computation of the point locations and probabilities is well-understood. Thus, we thus exclude the details. Briefly, point selection is achieved by first finding polynomials of successive degrees that are orthogonal with respect to the sample probability distribution: this computation requires using the sample statistics obtained in Step 2. It turns out that the roots of these orthogonal polynomials are the quadrature points, and their associated probabilities can then be selected so that the discretized distribution's statistics match the original sample distribution's statistics. The analyses in Ref. 19 further verify that simulation at the quadrature points allows accurate identification of the relationship between the metric(s) and dependent simulation outputs (in our case, NAS performance outputs).

Step 4: Choose the scenario that most closely achieves each representative metric value set as a representative scenario, and associate the probability for the representative metric set with this representative scenario.

Several notes are necessary regarding the quadrature-based algorithm for representative scenario selection. First, we again stress that the success of the algorithm in matching metric statistics, and in finding the mapping between the metrics and dependent variables, can be mathematically justified: we kindly ask the reader to see Ref. 19 for the relevant analyses. Second, let us note that, to use the algorithm, the user must decide on the number of representative scenarios to generate. This desired number of scenarios may be guided by the user's needs in terms of planning flow contingencies. Alternatively, the methodology admits a natural automated ``stopping rule" for choosing the proper number of representative scenarios, based on information-theoretic constructs. This automated criterion determines whether or not additional representative scenarios provide a more refined representation of the metrics' probability distribution (according to an information-theoretic measure), and decides on the number of representative scenarios based on such comparison. We refer the reader to the thesis Ref. 25 for the details.

Alternative Metric Value-Based Approaches:

The Gaussian quadrature-based approach to representative-scenario selection described above can be viewed as an example of a metric *value*-based clustering approach. That is, the quadrature-based approaches uses valuation(s) of each weather-impact scenario (which are meant to be predictors of consequent NAS performance or management) to cluster scenarios, and hence choose representative ones and associate probabilities with them. In other words, scenarios are clustered based on whether they have similar metric values, rather than based on pairwise distances. The quadrature-based approach that we have developed above is one of several possible metric value-based approaches. The approach is appealing for our purposes since it is tailored for effective performance evaluation of flow contingencies, however alternative approaches can certainly be used instead. Alternatives may be especially valuable when many metrics are being used together for clustering, since the quadrature-based approach typically does not permit sparse selection of representative scenarios in this case.

One natural alternative to the quadrature-based approach is to use classical K-means clustering, which aims to group samples (in our case, scenarios) by their metric values into a specified number of groups so as to minimize the total squared deviation of the samples' metric values from their respective cluster means. K-means clustering, which is usually achieved using an iterative optimization algorithm, can be directly implemented from a table of metric value(s) for each scenario. K-means clustering is an appealing alternative to metric value-based clustering, in the case where many metrics are considered.

C. A Signal Comparison-Based Approach

From a flow management viewpoint, two different weather-impact scenarios may yield completely different management strategies, even if global measures like the total capacity reduction are of similar value for the two scenarios. On the other hand, two scenarios that look very "similar" to each other are likely to have the same impact on traffic flows, and hence require similar management strategies. For example, in managing traffic to an airport or a critical Sector, the precise timing of weather impacts in nearby areas may be significant (while details of distant regions are less important). In this specific example, it is natural that we group weather-impact scenarios by fully comparing the weather-impact trajectories in these important regions. With this motivation in mind, the second clustering methodology we consider here is a signal comparison-based approach. Specifically, we aim to cluster scenarios based on *differences* between weather-impact signals (or metrics derived from these signals), rather than grouping based on metric values.

There are two tasks that we are concerned with in this comparison-based approach. The first task is to properly quantify distances (or differences) among different scenario signals. One possible approach is to directly compute the distance between the spatio-temporal weather- or weather-impact scenarios according to a norm of the algebraic difference between the signals. A very different approach is to abstract each scenario signal to a single or small number of points (metric values), and quantify pairwise distances among all the points. More broadly, distances can be viewed as normed differences between metrics/statistics of the spatio-temporal weather-impact signals, which may range from the signal itself to highly aggregated metrics (such as the ones used in the Gaussian quadrature-based methodology). From this viewpoint, we see that the signal-comparison-based approach permits use of many possible comparison metrics, with comparisons between the global metrics used in the quadrature-based methodology representing only a special sub-case.

The second task is to cluster all the scenarios based on the obtained pair-wise signal distances. Actually, clustering based on such pairwise distances has been well studied in the data mining literature (e.g. Ref. 6, 8, 12, 14). Classical algorithms (e.g., *spectral clustering algorithms*) can be directly applied here to obtain groups of points (scenarios) that are close to each other, i.e. whose pairwise distances are small. In our development, we choose the spectral clustering algorithms as our clustering method, because they are simple to implement, and also comparatively efficient when treating a massive ensemble of scenarios (Ref. 12, 14). The following is the detailed procedure for the signal comparison-based methodology:

Step 1: Choose proper distance metrics as our signal difference measures, and compute all the pairwise scenario distances. In analogy with metrics chosen for the quadrature-based method, choosing proper signal-distance measures depends on application requirements. In general, quantifying scenario signal distances requires two decisions: where to measure, and what to compare. With regard to the measurement location, we need to figure out which aspects of the scenario signals differentiate performance/control and hence need to be compared (e.g., global vs. local measures, time horizons for comparison). What to compare can also be different for different application requirements. For example, as discussed above, we can directly measure distances between pairs of weather or weather-impact scenarios. We can also measure distances between certain local or aggregate metrics or statistics (e.g., sector capacity reductions, excess demand relative to capacity) derived from the original scenarios. After deciding on the measurement location and comparison specifics, we need to compute distances between each pair. Noting that each scenario signal (or derived metrics/statics) can simply be viewed as a data vector, any norm of the algebraic distance between the signals can be used as the distance measure. How to properly quantify distances among scenario signals is an important factor influencing the quality of representative scenario selection.

Step 2: Define a graph based on scenario-signal distances, obtain a proper adjacency matrix, and apply standard spectral clustering algorithms. Classical spectral clustering algorithms identify strongly connected subgroups of a graph. To apply such algorithms, we need to define a graph first. Specifically, we consider a weighted, undirected graph, where each vertex in the graph represents a single scenario. We note that the distances we obtain from the above step reflect how close the scenario signals are, i.e. a big distance means that the two signals are far away from each other while a small distance means that they are similar to each other. However, in many spectral clustering algorithms, small values in distance matrices usually indicate weak connections among vertices, while big values indicated strong connections. Thus, for each edge, we use the inverse of each signal distance as the corresponding edge weight, and construct an adjacency matrix based on the edge weights. After we obtain the adjacency matrix of the graph, several standard spectral clustering algorithms can be directly applied to obtain a desired number of strongly connected groups. We omit discussion of the clustering procedures here, please see e.g. Ref. 12, 14 for a description of the procedures.

Step 3: Choose a representative scenario from each cluster, and associate each selected scenario a probability. Once we obtain a clustering result, we can choose a "most central" scenario in each cluster (e.g., the one with the smallest total sum of the distances to other nodes) as the representative scenario. The fraction of the scenarios in each cluster can be chosen as the corresponding probability.

We have thus provided detailed descriptions regarding two methodologies for selecting representative scenarios. Let us now discuss some of the differences between these two approaches:

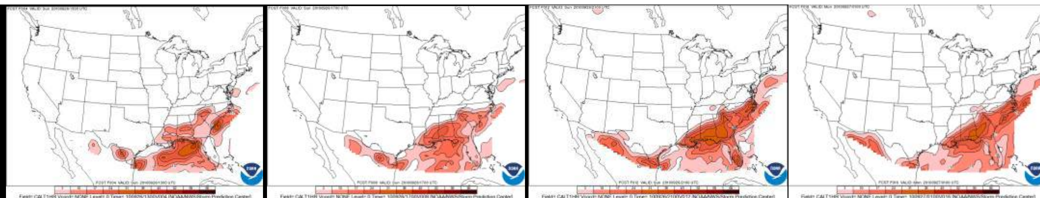
- The metrics that we consider for the two approaches evaluate scenarios in two different ways. The metrics that we use in the Gaussian quadrature-based method capture properties of individual scenarios (e.g., aggregated statistics), while we measure pairwise differences in the signal comparison-based method (e.g., norm-based distances between the full signals or other metrics defined thereof).
- The representative scenarios selected by the two methods also have different properties. The quadrature-based method finds representative scenarios that can statistically span the sample distribution for metrics, without explicitly providing clusters. The scenarios selected through the quadrature-based method match statistics up to certain order. The comparison-based method actually generates clusters, whose members tend to have similar characteristics.

IV. Example: Convective Weather in ZTL

We have applied both the quadrature-based and signal-comparison-graph-based methodologies for representative scenario selection, to a weather-simulation example from September 26, 2010. The weather event being modeled is a long-duration tropically-driven convective weather outbreak, which caused significant air traffic delay on both September 26 and September 27, 2010. In this example, a simulator was built to generate stochastic scenarios of weather and weather-impact between 5AM and midnight on September 26, 2010, as part of an example study of NAS-wide flow contingency management for that day. Specifically, the simulator tracks convective weather presence/absence in lattice squares of dimension 30mi x 30mi covering ZTL, as well as ZTL high-altitude (24K-35K feet) Sector capacities, at 15 minute intervals over the 19-hour duration. In Ref 22, we demonstrated parameterization of the simulator from probabilistic weather forecast data, and we use this parameterized simulator here. Snapshots of the SREF probabilistic forecasts used for parameterization are shown in Fig. 2.

We have applied the quadrature-based method for selecting representative scenarios. The selection was done using a metric that is defined from the weather scenario as well as nominal sector demands (Sector traffic counts computed assuming no weather restriction of traffic), that we postulate is predictive of NAS performance. Specifically, the metric used for representative scenario selection was the excess nominal demand relative to Sector capacity, totaled over the Sectors in ZTL and over the 19-hour duration of the simulation. Using the quadrature-based approach, we have selected four representative scenarios according to this excess demand metric, from an ensemble of 200 generated scenarios. We note that our choice of a 200-member ensemble is for convenience: the number is small enough to permit rapid experimentation with scenario-selection methodologies, while also appearing to capture the essential features of the metric distribution. In the future, some statistical analysis is needed to decide on the number of scenarios used to select representative ones: a natural means is to determine how many scenarios are needed to characterize the cumulative distribution of the metric to within an error bound.

The scenario-selection process and resulting representative scenarios/probabilities are illustrated in Fig. 3 and Fig. 4 below. To begin, the distribution of metric values over the ensemble of 200 scenarios is displayed (Fig. 3). We apply the quadrature-based methodology to find four representative scenarios and associate probabilities of occurrence with each scenario: we note that these representative scenarios and associated probabilities faithfully capture the first seven moments of the metric-value distribution, permit efficient characterization of consequent NAS performance metrics, and appear to span the range of possible metric values. The metric values for the representative scenarios and the associated probabilities are also shown on Fig. 3. The four representative scenarios can be viewed as very low impact, low impact, high impact, and very high impact scenarios, respectively. Snapshots of excess demand and capacity reduction are also shown for the four scenarios, at four hour intervals over the duration of the simulation (Fig. 4 and Fig. 5). In order of increasing metric values (VL impact through VH impact), the four representative scenarios display increasing capacity reductions in several critical sectors. These critical sectors suffer from significant excess demand upon capacity reduction due to weather, and hence we expect increasing capacity reduction to yield increasing NAS delays.



[redacted]

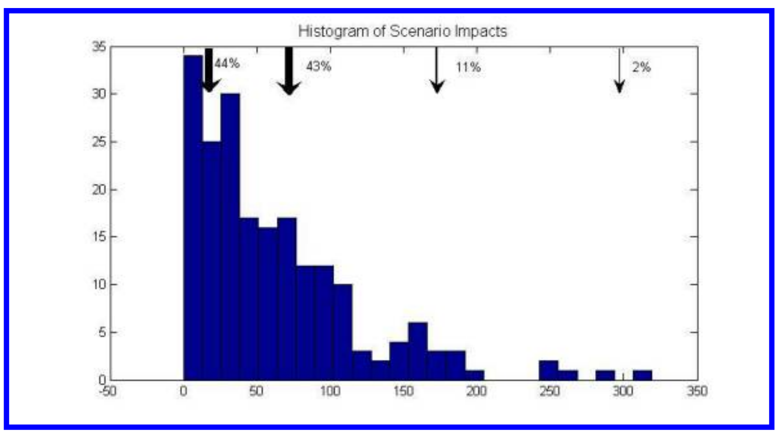


Figure 3. Histogram of Total Excess Demand for 200 Scenarios. *Metric values and probabilities for the representative scenarios, as obtained through the Gaussian quadrature-based approach, are shown at the top of the plot.*

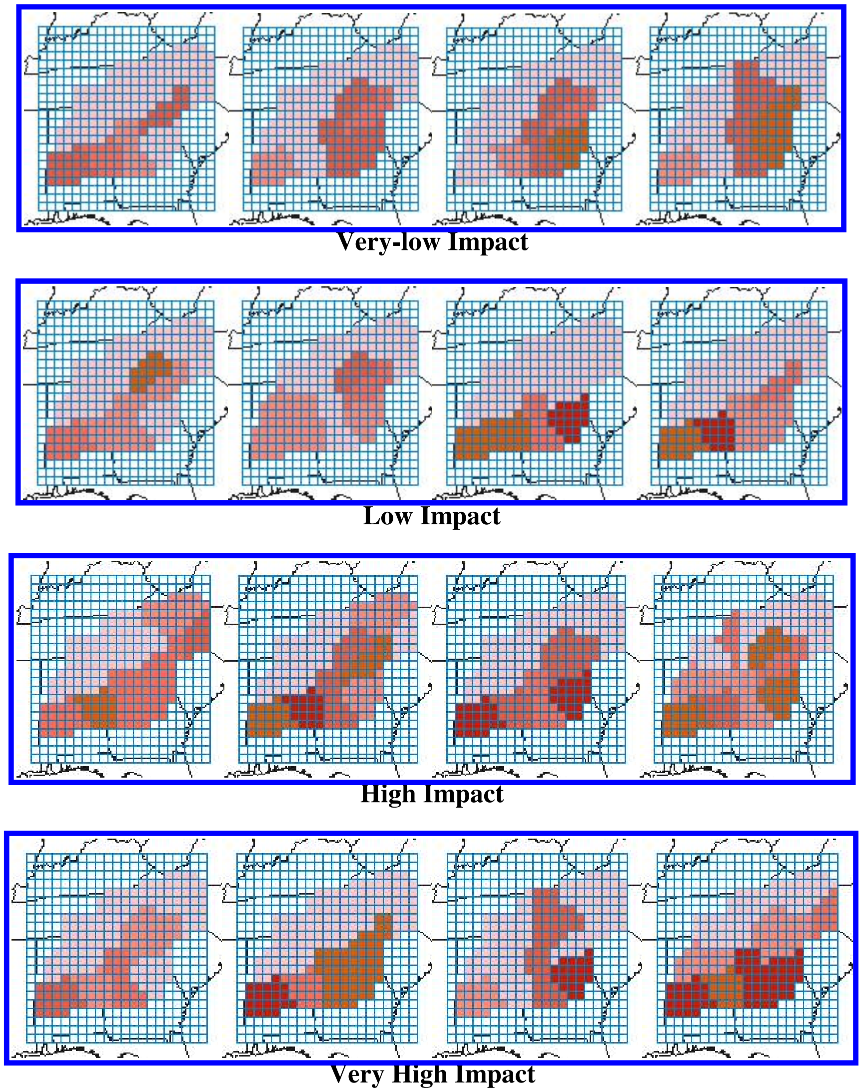


Figure 4. Snapshots of Sector capacity reductions for the four representative weather-impact scenarios obtained through the Gaussian quadrature methodology. *Darker shades of red represent higher percentage capacity reductions. The snapshots are at look-ahead times of 4, 8, 12, and 16 hours from the 5AM EDT forecast time. We note here that the shaded regions provide gridded approximations of the actual Sector boundaries; these grid squares represent regions in which weather is tracked in the underlying influence model, and the Sectors are drawn from the grid squares for convenience.*

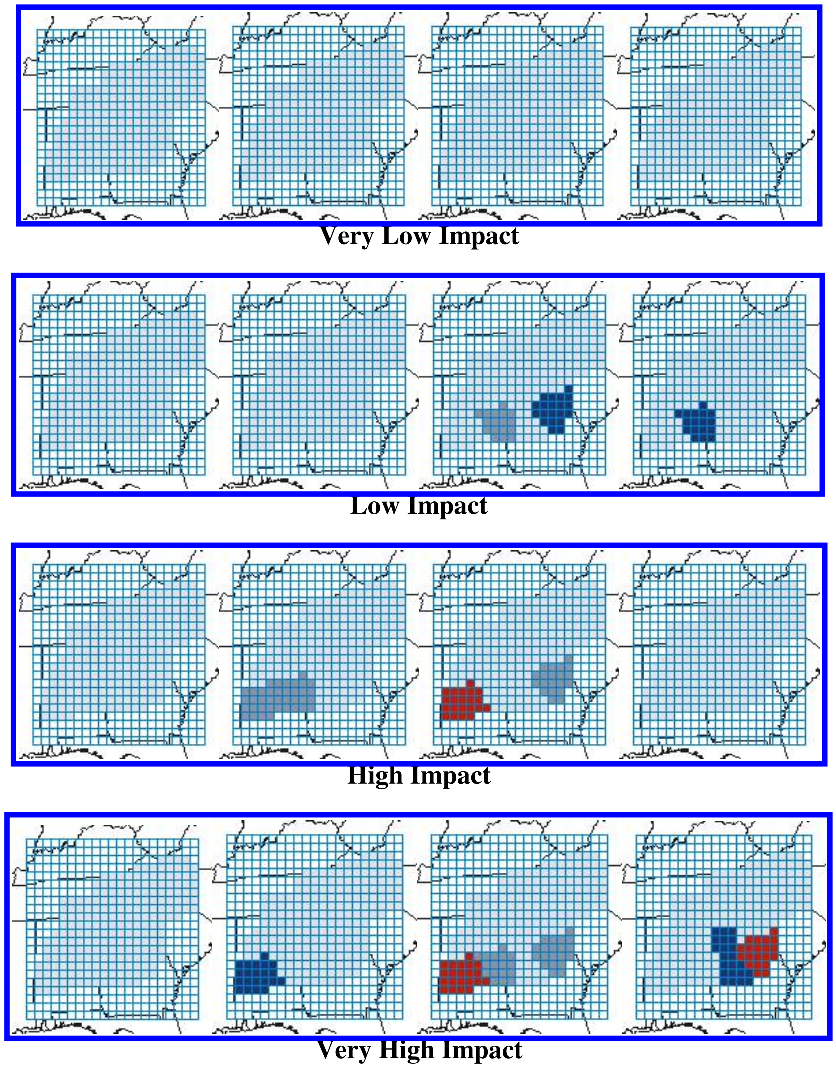
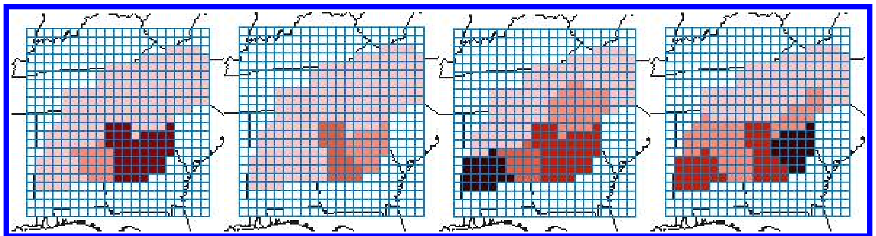


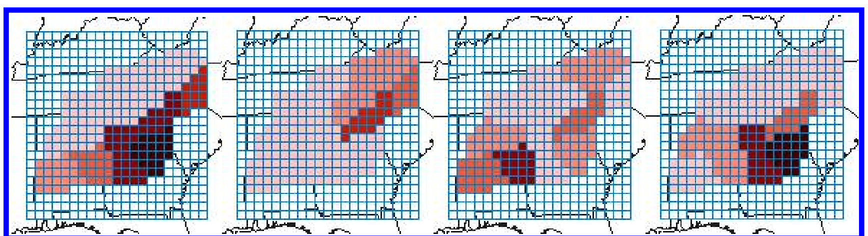
Figure 5. Snapshots of the excess Sector demand relative to capacity for the four representative weather-impact scenarios obtained through the Gaussian quadrature methodology. *The blue/grey shading represents a slight demand excess, the dark blue shading represents a moderate excess, and the red shading represents a significant excess. The snapshots are at look-ahead times of 4, 8, 12, and 16 hours from the 5AM EDT forecast time. We note that demand excesses are concentrated in only a few Sectors and time periods, although there are also widespread capacity reductions in sectors and time periods which have relatively low demand.*

We have also applied the signal comparison-based approach to select another set of representative scenarios for the September 26, 2010 example. In this case, we use an ensemble of 1000 weather-impact scenarios generated by the simulator, and aim to select four representative scenarios from the ensemble. Specifically, we choose three different signal-distance metrics as our comparison measures, and apply a normalized spectral clustering algorithm in each case. The algorithm can be found in Ref. 12 and Ref. 14, so we omit the details here. In the remaining of this section, we introduce each comparison measure and present the representative scenarios for each case.

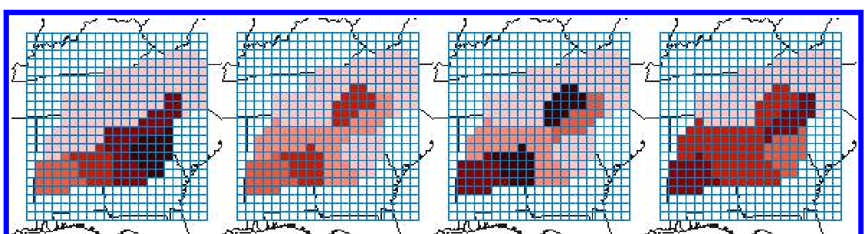
- *Total capacity metric:*
  For each scenario signal, we first compute the available sector capacity, totaled over all time intervals and sectors. Then, for each pair of scenarios, we use the absolute difference between these scenarios' total capacities as the signal distance. We then apply the signal-comparison-based method to select four representative scenarios. Although the total capacity metric does not differentiate between the capacity reductions in different sectors, it reflects the intensity of the whole weather event. To illustrate the four selected representative scenarios, we track the fraction of nominal capacity available for each sector at different time steps. In Fig. 6, we show snapshots at 4hr, 8hr, 12hr, and 16hr from each scenario, and give the corresponding probability associated with each scenario.



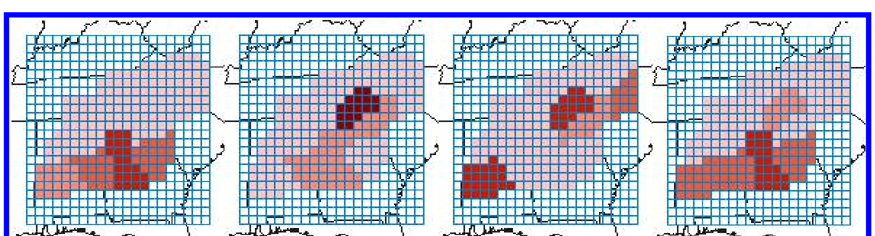
(a) Scenario 1 with probability 0.46



(b) Scenario 2 with probability 0.28



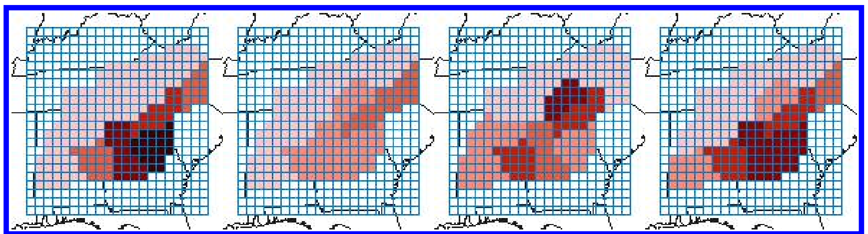
(c) Scenario 3 with probability 0.24



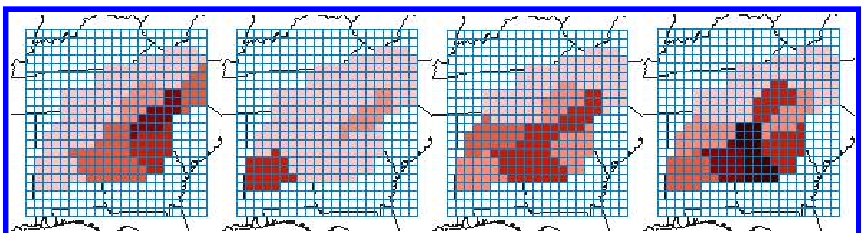
(d) Scenario 4 with probability 0.02

Figure 6: Snapshots of the available capacity fraction, for the four representative scenarios generated by the signal-comparison-based method. *Darker red shading represents less available capacity, and lighter red shading represents more available capacity. We use the total capacity difference as our metric to measure all pairwise scenario distances, and cluster all scenarios into four groups. We choose one representative scenario from each cluster and associate a probability with it. The snapshots are at look-ahead times of 4, 8, 12, and 16 hours from the 5AM EDT forecast time.*

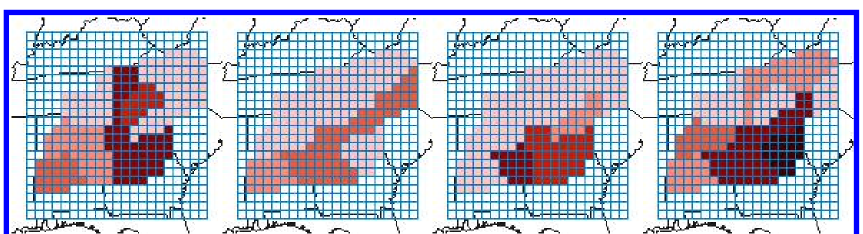
- *Time-aggregated sector capacity metric:*
  For each scenario, we add each sector's capacities over all time steps, and assemble these time-aggregated sector capacities in a vector. Then, we use the 2-norm difference between two different time-aggregated sector capacity vectors to represent the distance between the corresponding two scenarios. This metric partially differentiates between capacity reductions at different locations, which may be especially significant in designing flow-management capabilities. To show the selected representative scenarios, we again track the fraction of nominal capacity available for each sector at different time steps. In Fig. 7, we show snapshots at every four hours from each scenario, and give the associated probabilities.



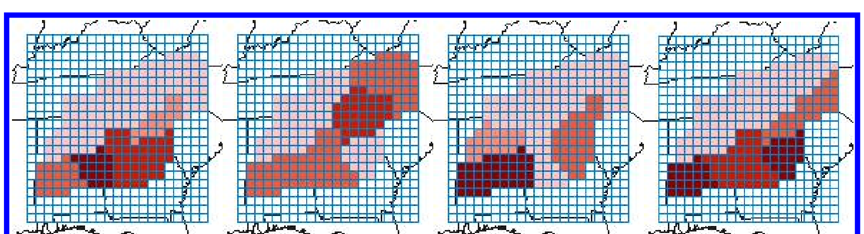
(a) Scenario 1 with probability 0.26



(b) Scenario 2 with probability 0.25



(c) Scenario 3 with probability 0.25

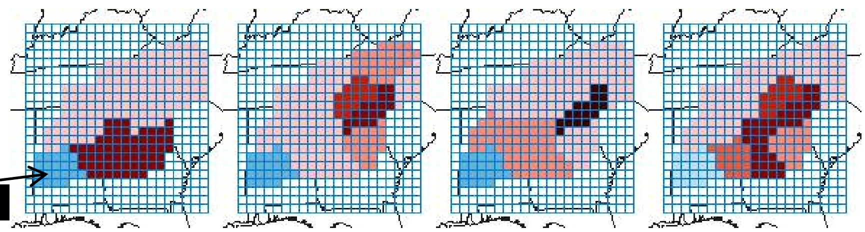


(d) Scenario 4 with probability 0.24

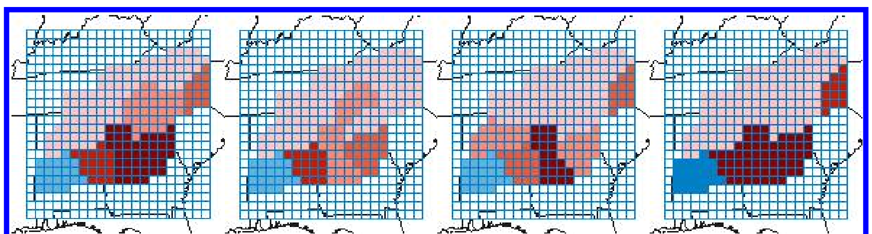
Figure 7: Snapshots of available capacity fractions, for the four representative scenarios generated by the signal comparison-based method. *Darker red shading represents less available capacity, and lighter red shading represents more available capacity. We use the 2-norm of the summed sector-capacity difference as our metric to measure all pairwise scenario distances, and cluster all scenarios into four groups. We choose one representative scenario from each cluster and associate a probability with it. The snapshots are at look-ahead times of 4, 8, 12, and 16 hours from the 5AM EDT forecast time.*

- *Single sector capacity metric:*
  For the single sector capacity metric, we only track the capacity reduction of a single sector (ZTL Sector 11 in the example). For each scenario, we consider a capacity vector where each entry represents the capacity of this single sector at each time step. We also use the 2-norm distance between two capacity vectors to represent the corresponding scenarios' distance. The metric can help in decision making especially for congested regions (as Sector 11 is in this case). In Fig. 8, we show snapshots of available capacity fraction at every four hours from each representative scenario, and give the associated probabilities. To highlight the single sector capacity reductions, we use a different set of color to represent the fraction of nominal capacity available for that sector.

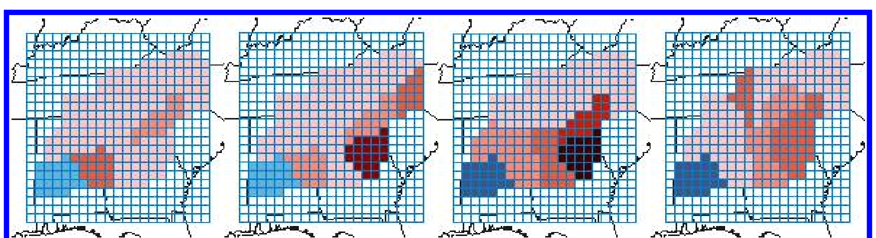
[redacted]



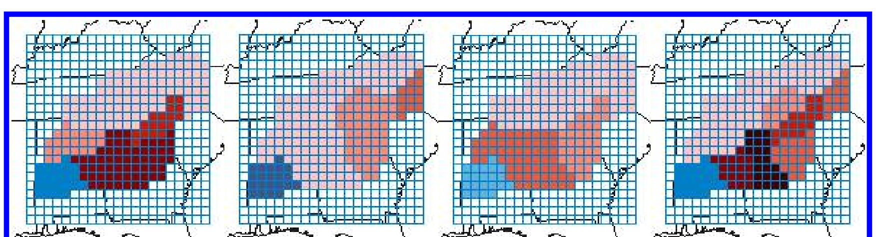
(a) Scenario 1 with probability 0.28



(b) Scenario 2 with probability 0.26



(c) Scenario 3 with probability 0.24



(d) Scenario 4 with probability 0.23

Figure 8: Snapshots of available capacity fractions, for the four representative scenarios generated by the signal comparison-based method. *Darker red/blue shading represents less available capacity, and lighter red/blue shading represents more available capacity. We use the absolute capacity difference for Sector 11 as our metric to measure all pairwise scenario distances, and cluster all scenarios into four groups. We choose one representative scenario from each cluster and associate a probability with it. The snapshots are at look-ahead times of 4, 8, 12, and 16 hours from the 5AM EDT forecast time. We note that Sector 11 is in the bottom left corner of ZTL: shades of blue rather than red are used to capture capacity reductions for this sector, so as to distinguish it from the others.*

V. Conclusions

In this article, we have introduced two approaches for extracting representative spatio-temporal weather- or weather-impact scenarios and corresponding probabilities of occurrence, from a large stochastically-generated ensemble of possible scenarios. These approaches have been applied to find representative weather-impact scenarios from simulations of a convective weather event in ZTL on September 26, 2010 that significantly disrupted air traffic. This research is very much at a preliminary stage, and we are far from validating or even comparing these approaches. Nevertheless, we believe that the results obtained through these two approaches are promising, and it is worth our while to draw several preliminary conclusions about the approaches:

1) In the ZTL example, the two approaches identify rather different scenarios as representative ones. We believe that this variation reflects the rich diversity in weather-impact trajectories for this particular example, which permits clustering of scenarios in several different ways. From this example, we conjecture that the quadrature-based methodology is effective at finding representative scenarios that together span the range of possible total weather impact on NAS performance, but the scenarios within each cluster may not look particularly similar. On the other hand, the signal comparison-based approach appears to find clusters of scenarios whose members look similar, but these representative cases do not immediately give a valuation of NAS performance.
2) For both approaches, the metric used for representative scenario selection critically impacts the results. Much more study is needed to determine metrics that are predictive (of NAS performance for the quadrature-based approach, and of difference in impact for the signal-comparison-based approach). In cases where the space of possible scenarios is rich (such as this example), we conjecture that using multiple metrics may be more effective.
3) Although our focus has been on selecting representative scenarios among samples generated from a particular weather-impact simulator, the methodologies developed here could be adapted to select representative members from other ensembles of spatiotemporal trajectories. It is easy to imagine a range of applications for such methods in transportation-related and other infrastructure-related domains, for instance in quantifying “similar-looking” historical dynamics in these networks or in modeling environmental uncertainties other than weather.

Notice:

References

[1]Aldenderfer, M. S., and Blashfield, R. K., *Cluster Analysis*, Newbury Park (CA): Sage, 1984.

[2]Asavathiratham, C., Roy, S., Lesieutre, B., Verghese, G., “The influence model,” *IEEE Control Systems Magazine*, vol. 21, no. 6, pp. 52-64.

[3]Bright, D. R., and Grams, J. S., “Short range ensemble forecast (SREF) calibrated thunderstorm probability forecasts: 2007-2008 verification and recent enhancements,” Preprints, *3rd Conf. on Meteorological Applications of Lightning Data*, Phoenix, AZ, Amer. Meteor. Soc., 6.3., 2008.

[4]Charles, M. E., and Colle, B. A., “Verification of Extratropical Cyclones within the NCEP Operational Models. Part II: The Short-Range Ensemble Forecast System,” *Weather and Forecasting* 24:5, 1191-1214, Online publication date: 1-Oct-2009.

[5]Du, J., DiMego, G., Tracton, M. S., and Zhou, B., “NCEP short-range ensemble forecasting (SREF) system: multi-IC, multi-model and multi-physics approach,” Research Activities in *Atmospheric and Oceanic Modelling* (edited by J. Cote), Report 33, CAS/JSC Working Group Numerical Experimentation (WGNE), WMO/TD-No. 1161, 5.09-5.10.

[6]Dubes, R. C., and Jain, A. K., *Algorithms for Clustering Data*, Prentice Hall (1988).

[7]FAA ATO NextGen & Operations Planning, “NextGen Mid-Term Concept of Operations for the National Airspace System,” June 30, 2009.

[8]Hartigan, J. A., and Wong, M. A., “A K-Means Clustering Algorithm,’’ *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, vol. 28, no. 1, pp. 100-108, 1979.

[9]Davis, P. J., and Rabinowitz, P., *Methods of Numerical Integration*, Harcourt Brace Jovanovich, 1984.

[10]Hockenberry, J. R., and Lesieutre, B. C., “Evaluation of uncertainty in dynamic simulations of power system models: the probabilistic collocation method,” *IEEE Transactions on Power Systems*, vol. 19, no.3, pp. 1483-1491, 2004.

[11]Krozel, J. A., Mitchell, J., Polishchuk, V., and Prete, J., “Capacity Estimation for Airspaces with Convective Weather Constraints,” *AIAA Guidance, Navigation, and control Conference,* Hilton Head, SC, August 2007.

[12]Luxburg, U. von, “A tutorial on spectral clustering,” *Statistics and Computing*, vol. 17, no. 4, pp. 395-416, August 2007.

[13]MacQueen, J. B., “Some Methods for classification and Analysis of Multivariate Observations,” in *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, University of California Press, 1:281-297, 1967.

[14]Ng, A., Jordan, M., and Weiss, Y., “On spectral clustering: analysis and an algorithm,” *Advances in Neural Information Processing Systems 14* (pp. 849 – 856), edited by T. Dietterich, S. Becker and Z. Ghahramani , MIT Press, 2002.

[15]Roy, S., Wan, Y., Taylor, C., and Wanke, C., “A stochastic network model for uncertain spatiotemporal weather impact at the strategic time horizon,” in *Proceedings of the 2010 Aviation Technology, Integration, and Operations Conference*, Fort Worth, TX, Sep. 2010.

[16]Steiner, M., Mueller, C. K., Davidson, G., and Krozel, J. A., “Integration of Probabilistic Weather Information with Air Traffic Management Decision Support Tools: A Conceptual Vision for the Future,” *13th Conference on Aviation, Range and Aerospace Meteorology*, New Orleans, Louisiana, January 2008.

[17]Tan, P., Steinbach, M., Kumar, V., *Introduction to Data Mining,* 1st ed., Addison-Wesley Longman Publishing Co., Inc., Boston, MA, 2005, Chapter 8.

[18]Taylor, C., Wanke, C., Zobell, S. M., Holland, R., Wang, L., Wan, Y., Roy, S., and Xue, M., “Flow Contingency Management Decision Framework,” The MITRE Corporation MP100309, Sep. 2010.

[19]Wan, Y., and Roy, S., “Uncertainty evaluation through mapping identification in intensive dynamic simulations,” *IEEE Transactions on Systems, Man, and Cybernetics*, Part A, vol. 40, no. 5, pp. 1094-1104, 2010

[20]Xue, M., Roy, S., Wan, Y., and Saberi, A., “Designing Asymptotics and Transients of Linear Stochastic Automata Networks,” In *Proceeding of the 2010 AIAA Guidance, Navigation, and Control Conference*, Toronto, Canada, August 2010.

[21]Xue, M., Roy, S., Saberi, A., and Lesieutre, B., “On Generating Sets of Binary Random Variables with Specified First- and Second- Moments,” In the *Proceedings of the 2010 American Control Conference*, Baltimore, MD, June 30-July 2 2010.

[22]Xue, M., Zobell, S. M., Roy, S., Taylor, C., Wan, Y., Wanke, C., “sing Stochastic, Dynamic Weather-Impact Models in Strategic Traffic Flow Management,” in *Proceedings of 91st American Meteorological Society Annual Meeting*, Seattle, WA, January 2011.

[23]Zhou, B., “Introduction to the NCEP Very Short Range Ensemble Forecast System (VSREF),” *14th Conference on Aviation, Range and Aerospace Meteorology*, Atlanta, GA, January 2010.

[24]Taylor, C., Wanke, C., Wan, Y., and Roy, S., “A Framework for Flow Contingency Management,” *2011 AIAA Aircraft Technology, Integration, and Operations Conference*, Virginia Beach, Sep. 2011.

[25]Ramamurthy, D., *Smart Simulation Techniques for the Evaluation of Parametric Uncertainties in Black Box Systems*, Master’s Thesis, School of Electrical Engineering and Computer Science, Washington State Unversity, June 2005.