9th AIAA Aviation Technology, Integration, and Operations Conference (ATIO)
and
Air
21 - 23 September 2009, Hilton Head, South Carolina

AIAA 2009-6960

A Robust Approach for Predicting Dynamic Density

Michael Bloem[*]
NASA Ames Research Center, Moffett Field, CA 94035, USA

Chris Brinton[†], Jim Hinkey[‡], and Kenneth Leiden[§]
Mosaic ATM, Inc., Leesburg, VA 20175, USA

Kapil Sheth[¶]
NASA Ames Research Center, Moffett Field, CA 94035, USA

This research proposes and analyzes an approach for predicting controller workload by predicting dynamic density. Most dynamic density formulations estimate workload with a linear combination of a set of dynamic density factors that describe the traffic situation in a sector. The robust approach proposed here uses this linear structure and the available data to explicitly consider the relative levels of uncertainty in dynamic density factor predictions when predicting dynamic density. The benefits of the robust approach are analyzed by using predicted and actual dynamic density factor data collected while playing back traffic data in the Future ATM Concepts Evaluation Tool. Results indicate that the robust approach produces errors that are more than an order of magnitude smaller than those produced by a simple approach that ignores factor prediction uncertainties. However, other approaches achieve lower prediction errors than the proposed robust approach.

I. Introduction

The problem addressed in this research is the prediction of controller workload. Controller workload sets the capacity of airspace sectors. Improved predictions of workload are useful when balancing sector capacity and demand by altering air traffic or by changing airspace. Altering air traffic and changing airspace are more efficient if executed a few hours before a potential demand-capacity imbalance is realized, assuming that the imbalance is accurately predicted. For example, actions like delaying the departure of flights may not impact the workload in a sector for several hours and actions like changing sector boundaries are operationally expensive and cannot be implemented arbitrarily often. Currently, operational workload predictions are not trustworthy beyond about 45 minutes due to errors in the prediction of factors that impact flight trajectories, such as departure times, traffic flow management restrictions, air traffic control actions, flight routes and speeds, and weather.[1] This poor predictability limits the efficiency of actions that alleviate demand-capacity imbalances.

There is a significant body of research concerned with developing estimators and predictors of controller workload. Some approaches involve non-linear estimators or predictors based on approaches like neural networks.[2] Others are multi-dimensional workload estimates with factors selected based on enabling supervisor decisions.[3] Many of the proposed workload estimates are referred to as dynamic density, and they typically generate an estimate from a linear combination of a set of dynamic density factors.[4–8] Dynamic density factors attempt to quantify workload-inducing characteristics of traffic in a sector, such as aircraft density, the number of ascending or descending aircraft, the variance of aircraft headings, or the number of aircraft

[*]Research Aerospace Engineer, Systems Modeling and Optimization Branch, MS 210-10, NASA Ames Research Center, Moffett Field, CA 94035.

[†]President and Principal Analyst, Mosaic ATM, Inc., 801 Sycolin Road, Suite #212, Leesburg, VA, 20175 and AIAA Member.

[‡]Senior Software Engineer, Mosaic ATM, Inc., 801 Sycolin Road, Suite #212, Leesburg, VA, 20175 and AIAA Member.

[§]Principal Analyst, Mosaic ATM, Inc., 801 Sycolin Road, Suite #212, Leesburg, VA, 20175 and AIAA Senior Member.

[¶]Research Aerospace Engineer, Systems Modeling and Optimization Branch, MS 210-15, NASA Ames Research Center, Moffett Field, CA 94035 and AIAA Associate Fellow.

near sector boundaries. Much of this dynamic density research considers predicting controller workload with predictions of dynamic density.[3,5–7] A variety of methods for predicting dynamic density up to 2 hours in advance were investigated and compared in Ref. 5. These methods used predictions of each of the factors, and a single dynamic density formulation was derived for all look-ahead times. In Ref. 7, predicted factors were used to predict dynamic density. However, this work was only concerned with a prediction time horizon of 20 minutes and only tried to improve dynamic density predictions by improving aircraft trajectory predictions. In Ref. 3, factors were selected for predictions of dynamic density while considering their predictability, defined by the number of minutes for which the correlation between the prediction and realized value exceeds a threshold.

In this work, a robust approach for predicting controller workload by predicting dynamic density is proposed. This approach explicitly takes into account the varying degrees of accuracy in factor predictions at various prediction time horizons. Different coefficients are used in the linear combination of factors for different look-ahead times, thereby accounting for variations in the relative predictability of factors at different prediction time horizons. This work does not propose a new set of factors and does not endorse any particular set of factors and coefficients. Rather, it proposes and analyzes a way to make any linear dynamic density formulation robust to factor prediction uncertainties when the dynamic density formulation is used to predict controller workload.

Some further discussion of previous research in workload estimation is provided in Section II. In Section III, the robust approach for predicting dynamic density will be derived and discussed. The data collection method will be explained, assumptions will be investigated, and predictor performance results will be presented in Section IV. Conclusions and future research finish the paper in Section V.

II. Background

When workload is estimated with a linear combination of a set of factors, the coefficients in the linear combination must be estimated. This has been done in previous research using controller workload estimates, exact measurements of factor values, and regression techniques.[4,9] Air traffic controllers are asked to control an air traffic situation in a simulator and estimate their workload during the simulation. Factor values are computed from the exact measurements of aircraft states during the simulation. The samples of workload estimates and instantaneous measurements of factor values are then used by statistical regression techniques to estimate appropriate coefficients for the factors in a linear workload estimator. These coefficients are appropriate for the instantaneous estimation of current workload from exact measurements of factors.

A simple approach for predicting dynamic density (DD) is to use these coefficients in a linear combination of predictions of factor values. However, this approach ignores factor prediction errors. Using factors that are hard to predict accurately when predicting DD often leads to significant errors. These errors may be reduced by choosing the coefficients in a way that considers the relative reliability of the factor predictions at various prediction time horizons.

Choosing the coefficients in this way does not imply that factors are more or less important for estimating workload from exact factor measurements than has been demonstrated in previous research. It only implies that due to their varying levels of predictability, factors vary in their usefulness for predicting workload.

III. Robust Prediction of Workload

III.A. Derivation of Robust Predictor

Let w denote the scalar-valued workload in a sector at a particular time. Furthermore, let f_0 be a $p \times 1$ vector of factors measured at that same time. Similarly, let f_t be a $p \times 1$ vector of predictions of the factors that will occur in a given sector t time units in the future. The objective of the estimation problem is to find a DD function, $d_0(f_0)$, that estimates w given f_0. The objective of the prediction problem is to find a DD function, $d_t(f_t)$, that predicts w given f_t.

A separate linear model is assumed for each prediction look-ahead time t:

$$w = d_t(f_t) + \varepsilon_t = f_t^T \beta_t + \varepsilon_t. \tag{1}$$

Here β_t is a $p \times 1$ vector of coefficients that must be estimated from data. The ε_t term captures the fact that pairs of f_t and w are not actually governed by the linear relationship assumed by $d_t(f_t)$. It is assumed to be a random error with expected value 0 ($\mathbf{E}\,\varepsilon_t = 0$) and $\mathbf{var}\,\varepsilon_t = \sigma_t^2$. Also, ε_0 is assumed to be independent of the errors in f_t.

A few assumptions will be made about the uncertainty in predictions of future factors. Assume that the predictions of factors are unbiased. Mathematically, this means that $\mathbf{E}\,f_t = f_0$, where f_0 is the set of actual factors that are realized t time units after the predictions f_t are made. This may or may not be true, but its truth will be checked and it could be prescribed as a prerequisite for factor predictions. Also, let Σ_{f_t} denote the error covariance matrix for the predictions f_t. As mentioned previously, assume that ε_0 is statistically independent of the errors in the factor predictions f_t.

The robust approach will create a linear predictor $d_t(\cdot)$ for each lookahead time t by estimating the coefficients β_t. The data used for this estimation are N corresponding sets of realized factor vectors $(f_{0,1}, f_{0,2}, \ldots, f_{0,N})$, factor prediction vectors $(f_{t,1}, f_{t,2}, \ldots, f_{t,N})$, and measurements of workload $(w_1, w_2, \ldots, w_N)$. Let F_0 be an $N \times p$ matrix in which the i^{th} row is the i^{th} set of factors $f_{0,i}^T$ for $i = 1, \ldots, N$. F_t is the corresponding matrix containing the factor predictions. Similarly, let v denote an $N \times 1$ vector containing the measurements of workload.

What is unique about the robust approach is that the uncertainty in v and F_t will be modeled explicitly. This explicit treatment of uncertainty and the assumptions mentioned previously lead to a special case of stochastic robust approximation in which it is possible to write a closed-form expression for the objective.[10] Since $\mathbf{E}\,f_t = f_0$, F_t can be modeled as $F_t = F_0 + E_t$, where E_t is a random $N \times p$ prediction error matrix with $\mathbf{E}\,E_t = 0$. Similarly, v can be modeled as $v = F_0\hat{\beta}_0 + e_0$, where e_0 is a random $N \times 1$ vector of the errors between the DD values produced by the estimator $d_0(f_0) = f_0^T\hat{\beta}_0$ and the actual workload in the sector (w) for each of the N measurements. The coefficient estimates in $\hat{\beta}_0$ are used to estimate current workload with exact measurements actual factor data. They have been estimated in the previous research described in Section II and are considered given for this research.

The objective of traditional least squares regression is to minimize the residual sum of squares (RSS) in the training data. With uncertainties modeled explicitly, the expectation of the RSS in the training data is the natural choice for the objective:

$$\hat{\beta}_t = \underset{\beta_t}{\operatorname{argmin}}\,\mathbf{E}\,\mathrm{RSS}(\beta_t) = \underset{\beta_t}{\operatorname{argmin}}\,\mathbf{E}\,\|v - F_t\beta_t\|_2^2 = \underset{\beta_t}{\operatorname{argmin}}\,\mathbf{E}\,\|(F_0\hat{\beta}_0 + e_0) - (F_0 + E_t)\beta_t\|_2^2. \tag{2}$$

In order to simplify this objective into a closed-form expression that can be minimized, it will be assumed that the training data provides a good estimate of the factor prediction error covariance matrix:

$$\Sigma_{f_t} \approx \frac{1}{N}(F_t - F_0)^T(F_t - F_0). \tag{3}$$

Assuming that each of the N observations of $(f_t - f_0)$ are independent and normally distributed, this is the maximum likelihood estimator of Σ_{f_t}. An unbiased estimator of Σ_{f_t} for a generally distributed $(f_t - f_0)$ is $\frac{1}{N-1}(F_t - F_0)^T(F_t - F_0)$, which will be nearly equal to $\frac{1}{N}(F_t - F_0)^T(F_t - F_0)$ for large N. The insight that the available data (F_0 and F_t) enables the calculation of this estimate is what allows the robust approach to explicitly consider the relative uncertainty of factor predictions. While approximating Σ_{f_t}, a $p \times p$ matrix, involves approximating $\sim p^2/2$ terms because Σ_{f_t} is symmetric, each term is an expected value that is approximated by an average of N samples. Furthermore, N can be made arbitrarily large by running more fast-time simulations, as described in sub-section IV.B. The quality of this estimate will be evaluated in Section IV.D.1.

After using this estimate, the optimization problem that uses the data to find $\hat{\beta}_t$, the estimate of β_t, is

$$\hat{\beta}_t = \underset{\beta_t}{\operatorname{argmin}}(F_0\hat{\beta}_0 - F_0\beta_t)^T(F_0\hat{\beta}_0 - F_0\beta_t) + \beta_t^T(F_t - F_0)^T(F_t - F_0)\beta_t + N\sigma_0^2. \tag{4}$$

The derivation of the objective in this problem is in Appendix A.

This optimization problem can be solved by readily available convex optimization software (). However, the optimal $\hat{\beta}_t$ can also be computed by a closed-form expression (see Appendix A):

$$\hat{\beta}_t = (F_0^TF_0 + (F_t - F_0)^T(F_t - F_0))^{-1}F_0^TF_0\hat{\beta}_0. \tag{5}$$

III.B. Discussion of Robust Predictor

The objective of Problem (4) has three terms, and each quantifies a particular type of error:

$$\underbrace{(F_0\hat{\beta}_0 - F_0\beta_t)^T(F_0\hat{\beta}_0 - F_0\beta_t)}_{\text{Bias Error}} + \underbrace{\beta_t^T(F_t - F_0)^T(F_t - F_0)\beta_t}_{\text{Factor Prediction Error}} + \underbrace{N\sigma_0^2}_{\text{Model Error}} \quad . \tag{6}$$

Understanding these terms illustrates the tradeoffs that are made when the robust predictor solves Problem (4) to find $\hat{\beta}_t$.

The first term captures the bias error. This term can easily be set to zero by setting $\beta_t = \hat{\beta}_0$ (which implies that $d(\cdot) = d_t(\cdot)$). Doing so ignores uncertainties in the problem, but since $\mathbf{E}\, f_t = f_0$, it also ensures that $\mathbf{E}\, d_t(f_t) = \mathbf{E}\, d_0(f_t) = d_0(f_0)$. By estimating β_t as anything but $\hat{\beta}_0$, bias error is introduced into the RSS because $\mathbf{E}\, d_t(f_t) \neq d_0(f_0)$. This bias error leads to the first term in Eq. (6). This term accounts for the errors in predicting workload that occur when the coefficients are changed from the values derived for the instantaneous estimation of workload.

The second term quantifies the factor prediction error. Factor predictions are not perfect ($f_t \neq f_0$), so when a factor prediction is used to predict workload, the corresponding factor prediction errors will introduce errors into the workload predictions. The magnitude of the factor prediction error term can be reduced by reducing the magnitude of β_t. In fact, this error term is zero when $\beta_t = 0$. However, setting $\beta_t = 0$ would induce a large bias error. When selecting β_t, the optimization process balances the bias error resulting from choosing coefficients that differ from those derived for the instantaneous estimation of workload and the prediction error resulting from factor prediction errors. The simple approach mentioned in Section II ignores the factor prediction error and only minimizes the bias error.

The third term captures the model error. This term results from the error in the choice of linear workload estimator. This part of the error cannot be reduced by the robust prediction method proposed here. Mathematically this is clear because the variable β_t is not part of this term and cannot be chosen in any way to reduce the value of the model error term. In effect, previous research has already investigated ways to reduce this error by choosing an appropriate estimator function $d_0(\cdot)$ and by finding the right estimates of its coefficients β_0.

These three errors are illustrated with a simple one-dimensional example. Consider a dynamic density formulation that has just one factor with the scalar coefficient $\hat{\beta}_0 = 1$. Let the model error for this formulation be such that $\sigma_0^2 = 1$. Suppose that there are only two sets of factor predictions and corresponding factor values in the training data. Let the predicted factor values be $F_t = [0\ 2]^T$ and the actual factor values be $F_0 = [1\ 1]^T$, which implies that $\Sigma_{f_t} \approx \frac{1}{N}(F_t - F_0)^T(F_t - F_0) = 1$. In figure 1 the three components of the error, measured in expected RSS per sample, are plotted as a function of β_t.

The model error is constant and cannot be impacted by modifying β_t. The bias error increases as β_t varies away from $\hat{\beta}_0$ and could be minimized by selecting $\beta_t = \hat{\beta}_0$. The factor prediction error grows as β_t increases in magnitude because of the difficulty in predicting the factor. The robust estimate of the coefficient would be $\hat{\beta}_t = 1/2$ because that value minimizes the total RSS per sample (and therefore also the expected total RSS of the training data). The robust coefficient leads to an expected RSS per sample of 3/2 while a simple approach would select $\hat{\beta}_t = \hat{\beta}_0 = 1$, leading to a larger expected RSS per sample of 2.

IV. Analysis and Results

IV.A. Selection of Baseline Dynamic Density Formulation

The robust approach presented here is indifferent to the particular set of factors under consideration. The robust approach should improve the quality of workload predictions for any model that is formulated as a linear combination of factor predictions.

While the approach proposed in this research applies to any such formulation, one particular formulation had to be chosen for the analysis. This work makes use of a formulation that is similar to Simplified Dynamic Density (SDD),[8] created using factor values that are computed in the Future ATM Concepts Evaluation Tool (FACET).[11] This formulation is chosen because it was developed for the purpose of dynamically designing airspace, and thus it performs relatively well for a wide variety of sector types. The SDD-like formulation used here includes ten factors. SDD coefficients are non-negative in part because every SDD factor should

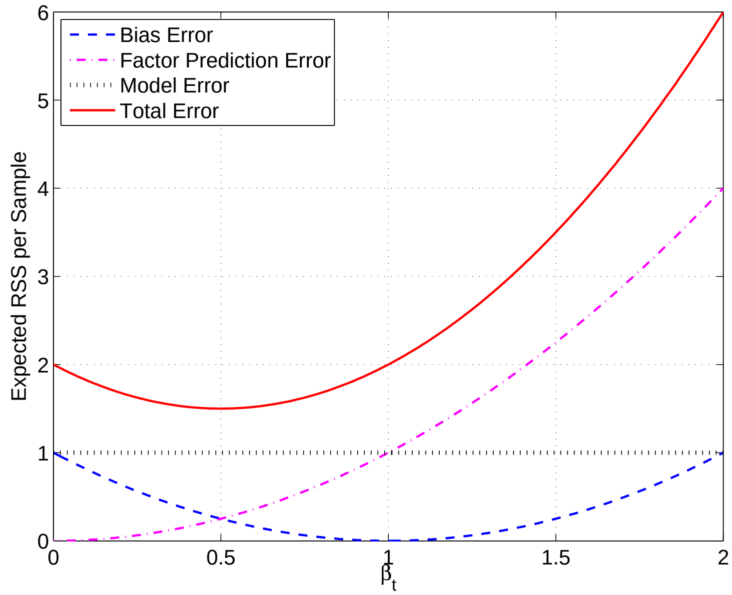


Figure 1. Notional example of expected RSS per sample as a function of β_t.

increase workload. Therefore, the approaches for estimating coefficients evaluated here are constrained to select only non-negative coefficient values. The baseline factors and coefficients are shown in table 1.

Table 1. Baseline Factor Descriptions and Coefficients

Factor	Description	SDD-like $\hat{\beta}_0$ Estimate
AC	Aircraft count	2.2
AD2	Number of aircraft divided by sector volume	4723
C2	Number of climbing aircraft	0.2
C4	Number of descending aircraft	0.2
NumHoriz	Number of aircraft with horizontal separation <8 nmi	0.3
S5	Number of aircraft with 3D Euclidean distance between 0–5 nmi excluding violations	1.2
S10	Number of aircraft with 3D Euclidean distance between 5–10 nmi excluding violations	0.6
WBPROX	Number of aircraft within 10 miles of a sector boundary	0.4
C14	Variance of aircraft speeds	0.0005
HDGVARI	Variance of aircraft headings	0.0005

A histogram of the realized DD values considered for the analysis is shown in figure 2. The SDD-like formulation leads to DD values mostly between 0 and 70.

IV.B. Data Collection and Validation

FACET was used to generate the predicted factor data used in the study.[11] Through custom modifications, the FACET planning application generated predicted trajectories for aircraft based on historical or flight plan information. These predictions use only information that was available at the time that the predictions would have to be made in a real-time system. Predicted trajectories were generated for both airborne flights and flights scheduled for departure.

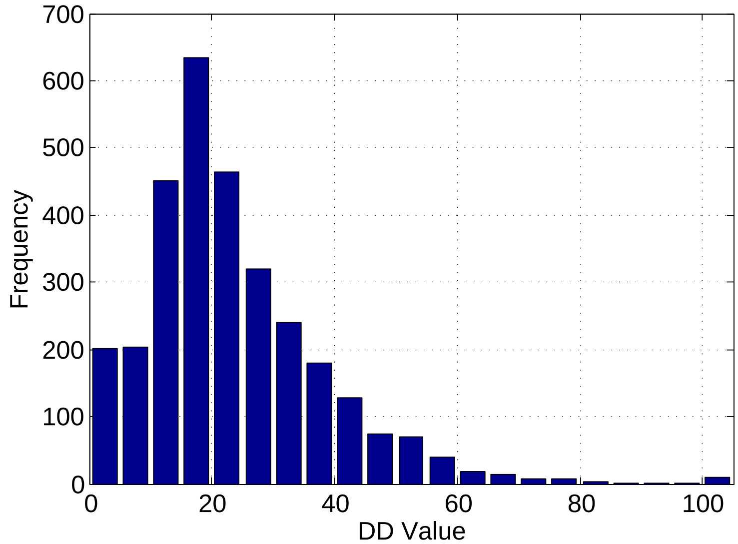


Figure 2. Histogram of actual DD values in the data set. The right-most bar counts any DD value larger than 100.

The predicted trajectory data was transferred to software that computes factors based on the predicted trajectories. In addition, the actual track data was used to compute the realized factors that occurred in each sector. These realized factor values were the truth data when computing the predictor coefficient estimates.

Enhanced Traffic Management System (ETMS) data is recorded from the live operation of the National Airspace System and was used as the input for this work. ETMS data contains both actual track and flight plan data for flights. The data was recorded from 1:00–2:30 pm Eastern Daylight Time on Thursday 09-06-2007. There were roughly 4,000 airborne aircraft during this period. For airborne flights, actual track data and current flight plan routes were assembled into the TRX format. TRX files contain track and flight plan data at one minute intervals for each aircraft in flight. The TRX file was used as input to FACET's sector planning application. For scheduled flights, the estimated times of departure (ETDs) and estimated take-off trajectories were integrated into auxiliary files used by FACET. These estimated take-off trajectories included airspeeds and altitudes that were either historical or filed by airlines. The latitude and longitude of the origin airports were used as the take-off trajectory coordinates.

The FACET planning application was used to generate predictions for prediction time horizons of 15, 30, and 60 minutes. Actual and predicted trajectory data for each airborne and scheduled aircraft were extracted from FACET and written to respective flight data files. The scheduled aircraft predicted trajectories were used from the estimated time of departure until either the aircraft actually departed or the schedule timed out. The following scenarios were reflected in the resulting flight data:

- the flight departs before the ETD,
- the flight is delayed, then eventually departs, and
- the flight reaches schedule "time-out" and is considered cancelled.

Extensive verification of both trajectories and factor predictions was conducted. For example, horizontal (great circle) and vertical prediction errors were calculated based on the difference between the actual and predicted trajectories to determine if reasonable uncertainties were being included in the trajectory predictions. Furthermore, the flight trajectory files were visually inspected and compared with original track and schedule ETMS data to verify that the correct sequence of pre-departure and airborne (i.e., post-departure) predictions were being made. This was particularly important when a flight was delayed since additional pre-departure predictions (to be made in 5 minute increments) were required. Once the trajectory prediction process was verified, the factor prediction verification process focused on understanding the cause of large outlier factor prediction errors and addressing them as needed.

Some of the data was filtered out before the statistical analysis was conducted. Only sectors with altitude floors above 24,000 feet were considered. Also, sectors with relatively low traffic volumes were not considered so that the workload predictions would not be biased by empty sectors. More specifically, any set of factor predictions for which neither the predicted number of aircraft nor the realized number of aircraft was at least five were removed from the data.

IV.C. Analysis Technique

The first step in the analysis is to verify some of the assumptions made in the motivation for and derivation of the robust approach to predicting workload, as explained in Section III. For example, the robust approach makes sense when some factors can be predicted more reliably than others. This non-uniform predictability of factors is demonstrated from the data. Furthermore, the assumption that $\mathbf{E}\, f_t = f_0$ is investigated. Finally, the estimation of Σ_{f_t} with $\frac{1}{N}(F_t - F_0)^T(F_t - F_0)$ will be studied by comparing the value of $\frac{1}{N}(F_t - F_0)^T(F_t - F_0)$ in the training and test data.

To be precise, the ability of an approach to predict DD will be quantified by the residual sum of squares (RSS) in the test data. The data is iteratively divided into training and test sections, allowing for cross-validation.[12] The truth data for the error calculations is the DD as computed with the realized factor values and $\hat{\beta}_0$ from table 1.

IV.D. Results

IV.D.1. Investigation of Assumptions

The data collection approach (see sub-section IV.B) captures uncertainty in the take-off times of not-yet-departed flights. This uncertainty is viewed as being important to the factor prediction errors and therefore also the DD prediction errors. To verify that take-off time uncertainty is important, horizontal prediction error statistics were computed for airborne and pre-departure flights at a 30-minute prediction horizon. The cumulative percentages of flight position predictions with less than or equal to various horizontal prediction errors are shown in figure 3. If a trajectory prediction has large horizontal prediction error, many DD factor computations will also be erroneous.

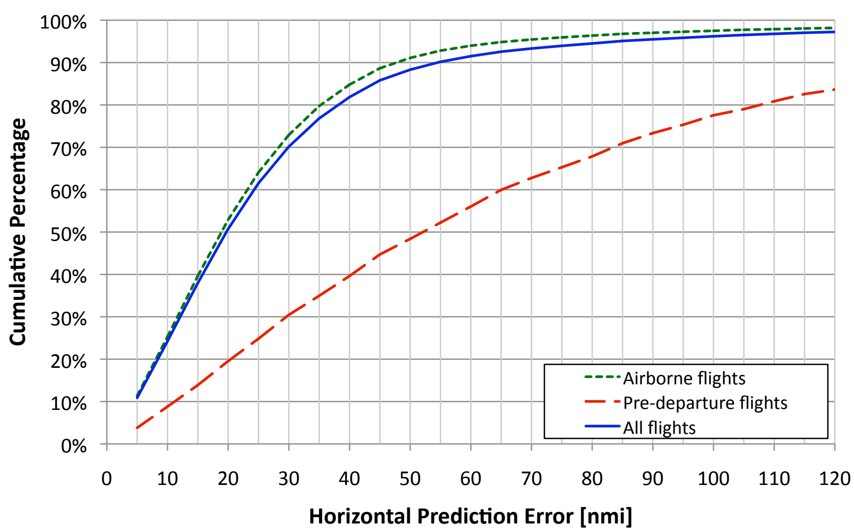


Figure 3. Cumulative percentages of 30-minute flight position predictions with less than or equal to various horizontal prediction errors.

The horizontal prediction errors are considerably larger for pre-departure flights. For example, only about half of all 30-minute position predictions of pre-departure flights have horizontal prediction errors less than 50 nmi. However, more than 90% of all 30-minute predictions that consider only airborne flights have horizontal prediction errors less than 50 nmi. When all flights involved in 30-minute factor predictions are considered, the distribution of horizontal prediction errors is nearly the same as for airborne flights only. This is because most flights involved in 30-minute factor predictions are airborne at the time the prediction is made. Predictions of factors further in the future would be influenced more by the uncertainty in the predictions of the pre-departure flights. Take-off time uncertainty is important and should be considered when studying the predictability of sector workload.

The robust approach suggested in Section III makes sense when different factors have differing levels of prediction accuracy that also vary as a function of prediction time horizon. To investigate whether this is true, the errors in the predictions of various factors are compared.

The results indicate that that factors do have varying levels of prediction accuracy and that these levels of accuracy do vary depending on the prediction time horizon. To illustrate this, the accuracy of predictions of sector aircraft count will be compared with the prediction accuracy of the number of aircraft with horizontal separation under 8 nmi ("NumHoriz" from Ref. 9). Prediction accuracy will be quantified by the prediction root mean squared error (RMSE), normalized by the mean of the realized factor values. The results are shown in the figure 4.

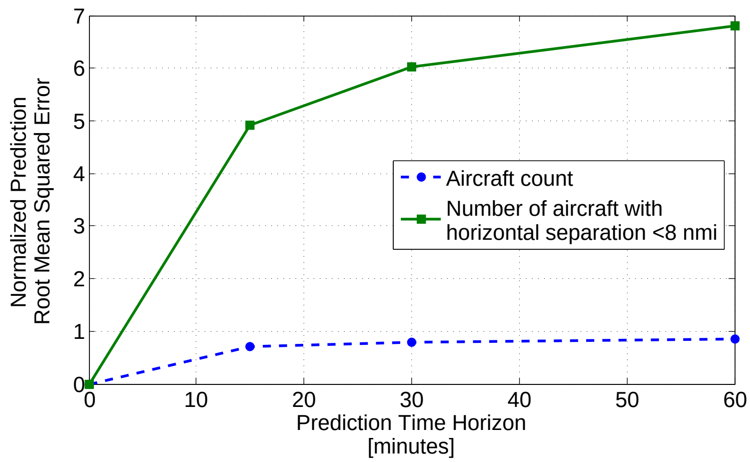


Figure 4. Normalized prediction root mean squared error for two factors.

The normalized RMSE is several times larger for the horizontal separation factor than for the aircraft count factor. For most other factors in the baseline SDD-like formulation (see table 1), the normalized RMSE varies between around 1 and 10. The speed variance factor normalized RMSE for a 60-minute prediction time horizon is larger than 700, indicating that it is particularly difficult to predict. Overall, the factors exhibit different levels of prediction accuracy.

The normalized RMSE for both factors in figure 4 increases with prediction time horizon, as expected. The magnitude of the increase varies from factor to factor. For example, the horizontal proximity factor experiences a greater reduction in predictability with prediction time horizon than the aircraft count factor.

An assumption used to derive the robust controller is that the mean factor prediction error for each factor is zero. Of the 10 factors used here to approximate SDD, the normalized prediction mean absolute error is less than one for only about half of the factors at each prediction time horizon. This means that for many factors, the bias in the factor prediction is larger than the average realized factor value. The robust predictor performance is hindered by the incorrect assumption that factor prediction errors have mean zero. With enough data, the bias in the prediction error for each factor could be estimated and used to correct factor predictions so that they are mean zero.

Another assumption made by the robust predictor is that the error covariance of the factor predictions can be approximated by the factor predictions and actual factor values in the training data. The accuracy of this assumption depends on the number of samples and factors, how well the samples represent the space of possible factor predictions and corresponding actual factor values, and whether or not the samples are independent.

There are 3,076 samples of 15-minute factor predictions, 2,538 samples of 30-minute factor predictions, and 1,210 samples of 60-minute factor predictions available to estimate the 10×10 error covariance matrix Σ_{f_t}. These samples are from a single 90 minute period of time and they were repeatedly broken into training and test data sets as part of the cross-validation process. Many are from the same sector at nearby times or from neighboring sectors at the same time, so it is unlikely that they are all independent. More independent samples would be needed to generate a better estimate of the error covariance matrix. The number of factors can be reduced by using factor-selection methods, which will improve the error covariance matrix estimate.

Accuracy of the error covariance matrix estimation is analyzed by comparing the variance of each factor as estimated by each training data set to the variance estimated by each test data set. The estimate in Eq. (3) is used. If the training set is a large set of independent samples and the performance of the predictor is evaluated over a large set of independent test data samples, these two variance estimates would be nearly identical. The variance in the test and training samples are not equal, and they diverge more for longer prediction time horizons. When measured as a percentage of the variance of the factor prediction errors in the entire data set, the average absolute difference between the training data set estimate of variance and the test data set estimate of variance is typically between 10% and 20% for the 15-minute prediction time horizon and between 20% and 50% for the 60-minute prediction time horizon. This degree of inequality in the error covariance matrix estimates from the training and test data sets will also hinder the performance of the proposed robust prediction approach. This issue could be overcome with a larger data set.

IV.D.2. A Simple Example

This example demonstrates some of the characteristics of the robust predictor. Consider a dynamic density formulation that is based on only two factors:

$$d(f) = \beta_1 f_1 + \beta_2 f_2. \tag{7}$$

For this example, f_1 is the number of aircraft in the sector and f_2 is the number of aircraft with horizontal separation under 8 nmi ("NumHoriz" from Ref. 9). Both coefficients are set to 1 for this DD.

The mean of the realized values of f_1 is 7.0 and the mean for f_2 is 2.2. However, it is more difficult to predict f_2 than f_1. The error covariance matrix for the prediction errors of these two factors at a 30-minute prediction time horizon is estimated by Eq. (3) as

$$\Sigma_{f_{30}} \approx \begin{bmatrix} 26.5 & 27.5 \\ 27.5 & 108 \end{bmatrix}.$$

The variance for f_1 prediction errors is four times smaller than the error variance for the predictions of f_2, indicating that f_1 is considerably more predictable than f_2. For the 30-minute prediction of this DD, the weight on f_2 should be relatively small because it is so difficult to predict 30 minutes in advance. Similarly, the more predictable f_1 should be given a relatively large weight.

The robust approach adjusts the weights so as to minimize the expected value of the RSS error in the DD predictions. It does so by taking advantage of the estimate of the factor prediction error covariance matrix that can be constructed from training data. In this case, the robust approach sets the coefficients as indicated in table 2.

Table 2. Coefficients for Simple Example

Factor	Coefficient for Instantaneous Estimation (β_0)	Coefficient for 30-minute Prediction ($\hat{\beta}_{30}$)
f_1	1.00	0.901
f_2	1.00	2.82×10^{-8}

As expected, the robust predictor computes a relatively large coefficient for the prediction of f_1 and reduces the magnitude of the coefficient for the prediction of f_2. The robust predictor chooses the coefficients to balance the prediction error and the bias error. As it does so, the relatively large prediction error covariance for f_2 translates into a relatively large factor prediction error when the magnitude of the coefficient on f_2 is increased (see Eq. (6)). Therefore, the robust predictor chooses a small magnitude for the f_2 coefficient. This does not indicate that f_2 is unimportant when estimating workload from exact measurements of factor values. Nor does it change the meaning of workload as estimated by the choice of $d(f)$ in Eq. (7). Rather, it indicates that f_2 is so difficult to predict that it is essentially useless for making predictions of workload.

The coefficients found by the robust approach lead to much smaller errors when predicting this DD than the coefficients for the instantaneous estimation of workload. The mean absolute error in the test data is 11.0 when using the coefficients for the instantaneous estimation of workload, but only 4.24 when using the robust coefficients.

IV.D.3. *Comparison of Approaches for Predicting Workload*

In this section, the performance of the robust prediction method is compared with the performance of several other methods to determine if it provides any benefits over the other methods. The baseline method is a simple linear predictor in which the coefficients used for the prediction of workload are those found for the instantaneous estimation of workload (inspired by SDD and reproduced in Appendix B). Another approach uses least squares regression to estimate the coefficients, and produces different coefficient estimates for each prediction horizon time. More precisely, it solves

$$\hat{\beta}_t = \underset{\beta_t}{\operatorname{argmin}}\, \mathrm{RSS}(\beta_t) = \underset{\beta_t}{\operatorname{argmin}}\, \|F_0\hat{\beta}_0 - F_t\beta_t\|_2^2 \tag{8}$$

to estimate β_t for each prediction time horizon t. This approach can be motivated in several statistical frameworks, but these frameworks do not explicitly consider uncertainties in the factor prediction data.[10,12] It is referred to as "least squares." A similar approach used in Refs. 5 and 6 is to estimate coefficients using least squares regression but with the prediction time horizon as an additional factor. This approach is referred to as "least squares time factor." For this approach, a single set of coefficients is estimated and used for all prediction time horizons. Finally, an approach that performs factor subset selection is also implemented and evaluated. This approach is based on the robust approach and it reduces the number of factors in the predictor by setting some coefficients to zero. It uses a technique known as the lasso[12] or ℓ_1 regularization.[10] It is referred to as "robust lasso" and described in Appendix C.

The coefficients for each approach are specified as the solution of a convex optimization problem. These problems were solved using the `CVX` package for specifying and solving convex programs in Matlab.[13]

The metric by which predictors are judged in this research is the RSS in the test data. RSS is directly proportional to the RSS per test sample and monotonic in RMSE (a larger RMSE always corresponds to a larger RSS and vice versa). Table 3 shows the RSS per test sample for these approaches for various prediction time horizons. The mean absolute error (MAE) for each approach and each prediction time horizon is shown in table 4.

Table 3. Residual Sum of Squares per Sample in Test Data

Prediction Time Horizon [minutes]	Simple	Robust	Least Squares	Least Squares Time Factor	Robust Lasso
15	17,810	7,143	6,017	5,996	7,142
30	126,800	3,001	2,941	2,913	3,010
60	10,210,000	4,909	4,914	4,687	4,924

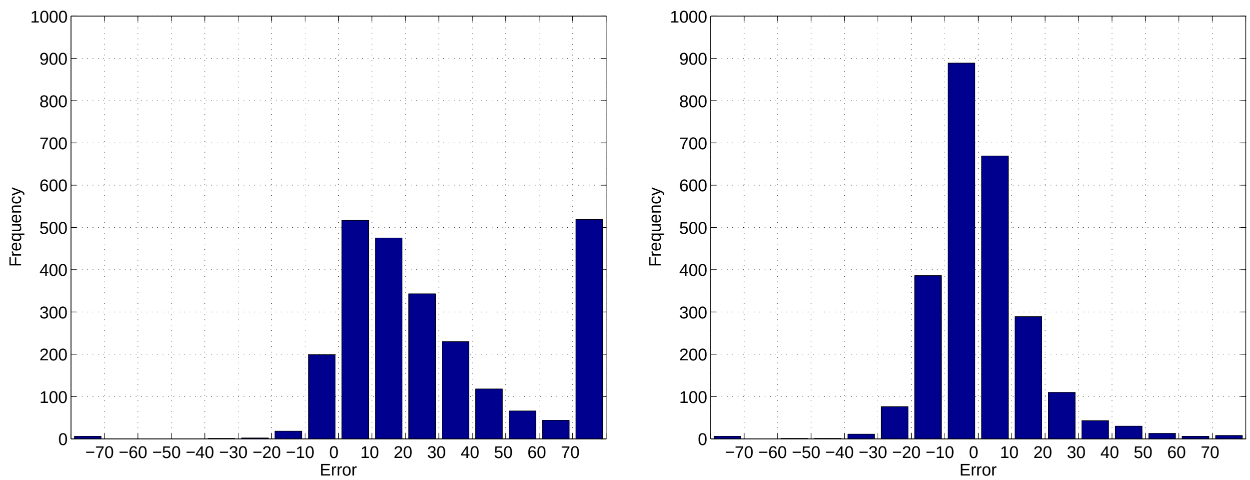
Table 4. Mean Absolute Error in Test Data

Prediction Time Horizon [minutes]	Simple	Robust	Least Squares	Least Squares Time Factor	Robust Lasso
15	43.91	16.40	12.09	11.53	16.38
30	138.8	12.62	10.76	11.05	12.79
60	582.2	16.67	16.25	13.37	17.22

The most important conclusion from these results is that the simple approach performs very poorly. The MAE of this approach ranges between 40 and 600 for DD values that generally range between 0 and 70. Using the coefficients derived for the instantaneous estimation of workload to also predict workload will not work due to factor prediction errors.

Any of the other approaches perform significantly better than the simple approach. At the 30- and 60-minute prediction time horizons, they all provide more than an order of magnitude reduction in prediction errors over the simple approach. The variation in performance among the other approaches is relatively small when measured in terms of RSS per test sample. The difference between the largest and smallest RSS per test sample is less than 20% of the smallest RSS per test sample at each prediction time horizon, and for two of the prediction time horizons it is around 5%. Among the other approaches, the least squares or least squares time factor approaches perform best for each prediction time horizon and for both mean RSS and MAE. The strong performance of the least squares time factor approach is interesting because it uses a single set of coefficients for all prediction time horizons. This implies that it is sufficient to derive coefficients that do not vary with prediction time horizon. The errors are larger for the robust approach and the robust lasso approach, probably because the assumptions that were made during the derivation of the robust predictor were shown in sub-section IV.D.1 to not hold in this data set. The robust lasso approach performs about as well as the robust approach.

The error histograms for the simple and robust predictors are shown in figure 5. Many of the predictions from the simple predictor yield errors larger than 70, and there is a positive bias in the predictions. An error of 70 when predicting a workload that is usually between 0 and 70 is likely unacceptable for most applications. The error histogram for the robust predictor shows more estimates with small errors.



(a) Simple predictor.

(b) Robust predictor.

Figure 5. Error histograms for the 30-minute look-ahead predictors. The left-most and right-most bars represent any errors less than -70 or greater than $+70$, respectively.

The coefficients computed by some approaches for the 30-minute prediction horizon are shown in Appendix B. While the lasso approach leads to larger errors, it uses significantly fewer factors to make its predictions. Remarkably, when trained on the entire data set, the robust lasso approach uses only one nonzero coefficient instead of the ten used by all of the other approaches. This coefficient corresponds to the number of aircraft within a threshold distance of a sector boundary ("WBPROX" in Ref. 9). It is the factor with the second smallest normalized prediction RMSE in the data, indicating that it is a relatively predictable factor, which partially explains its inclusion in the robust lasso predictor.

V. Conclusions and Future Work

A robust approach for predicting controller workload with a linear combination of DD factors has been presented. This approach leverages the available data and the structure of this problem to explicitly consider the relative uncertainties in factor predictions. Results indicate that ignoring factor prediction errors by using factor coefficients derived for instantaneous workload estimation leads to very large workload prediction errors. Any of the other investigated approaches lead to reductions in workload prediction errors of more than an order of magnitude at the 30- and 60-minute prediction time horizons. The robust approach does not perform as well as a least squares regression approach. The robust approach combined with a subset selection method indicates that relatively good DD prediction quality can be achieved with as little as one factor.

There is much future work to be done on the important topic of workload prediction. The robust approach considered here could be analyzed with a larger data set and with other DD formulations. For applications where workload is predicted for existing static sectors rather than for dynamically adapting sectors, sector-specific coefficients for workload prediction could be derived using this robust approach or another approach. In this case the robust approach could possibly be extended to something like a Kalman filter in which sector-specific factor prediction error covariance matrices are updated in real time. New factors that help estimate workload and are also predictable could be proposed and analyzed. Finally, it would be useful to determine the probability distribution of the future workload in a sector.

A. Detailed Derivation of Robust Predictor

This appendix contains the details of the derivation of the robust predictor. The expression to be simplified is the objective in Problem (2). When the terms are rearranged and the squared norm is written out as a product of matrices, this objective becomes

$$\mathbf{E}(F_0\hat{\beta}_0 - F_0\beta_t - E_t\beta_t + e_0)^T(F_0\hat{\beta}_0 - F_0\beta_t - E_t\beta_t + e_0).$$

As this multiplication is carried out, the first two of these four terms will be kept together as a single unit. This means that there will be three squared or nine terms after the multiplication is carried out and before any like terms are grouped. These nine terms are

$$\begin{aligned}\mathbf{E}[&(F_0\hat{\beta}_0 - F_0\beta_t)^T(F_0\hat{\beta}_0 - F_0\beta_t) - (F_0\hat{\beta}_0 - F_0\beta_t)^T E_t\beta_t + (F_0\hat{\beta}_0 - F_0\beta_t)^T e_0 \\ &- \beta_t^T E_t^T(F_0\hat{\beta}_0 - F_0\beta_t) + \beta_t^T E_t^T E_t\beta_t - \beta_t^T E_t^T e_0 \\ &+ e_0^T(F_0\hat{\beta}_0 - F_0\beta_t) - e_0^T E_t\beta_t + e_0^T e_0].\end{aligned}$$

There are several like terms that can be grouped because $(F_0\hat{\beta}_0 - F_0\beta_t)^T E_t\beta_t = \beta_t^T E_t^T(F_0\hat{\beta}_0 - F_0\beta_t)$, $(F_0\hat{\beta}_0 - F_0\beta_t)^T e_0 = e_0^T(F_0\hat{\beta}_0 - F_0\beta_t)$, and $\beta_t^T E_t^T e_0 = e_0^T E_t\beta_t$. Once these terms are grouped, the result is

$$\begin{aligned}\mathbf{E}[&(F_0\hat{\beta}_0 - F_0\beta_t)^T(F_0\hat{\beta}_0 - F_0\beta_t) - 2(F_0\hat{\beta}_0 - F_0\beta_t)^T E_t\beta_t + 2(F_0\hat{\beta}_0 - F_0\beta_t)^T e_0 \\ &+ \beta_t^T E_t^T E_t\beta_t - 2\beta_t^T E_t^T e_0 + e_0^T e_0].\end{aligned} \tag{9}$$

Note that because only E_t and e_0 are random, and because both have expected value of zero, the expected value of the terms $2(F_0\hat{\beta}_0 - F\beta_t)^T E_t\beta_t$ and $2(F_0\hat{\beta}_0 - F_0\beta_t)^T e_0$ are both zero. Also, the expected value of $2\beta_t^T E_t^T e_0$ is zero because it was assumed that E_t was independent of e_0. Noting that the first term in Eq. (9) is deterministic and that $\mathbf{E}\, e_0^T e_0 = N\sigma_0^2$, the objective simplifies to

$$(F_0\hat{\beta}_0 - F_0\beta_t)^T(F_0\hat{\beta}_0 - F_0\beta_t) + \beta_t^T\, \mathbf{E}[E_t^T E_t]\beta_t + N\sigma_0^2. \tag{10}$$

The final remaining expectation can be written as $\mathbf{E}[E_t^T E_t] = N\Sigma_{f_t}$, but Σ_{f_t} is not known. It is estimated with Eq. (3) as described previously. After substituting this into Eq. (10), the objective is approximated by the objective in Problem (4).

The closed-form expression in Eq. (5) for the robust estimator coefficients $\hat{\beta}_t$ will also be derived here. The optimization problem that is solved to find these coefficients is

$$\hat{\beta}_t = \underset{\beta_t}{\operatorname{argmin}}(F_0\hat{\beta}_0 - F_0\beta_t)^T(F_0\hat{\beta}_0 - F_0\beta_t) + \beta_t^T(F_t - F_0)^T(F_t - F_0)\beta_t + N\sigma_0^2.$$

The objective is quadratic in β_t, so the minimizing $\hat{\beta}_t$ can be found by differentiating with respect to β_t, setting the resulting linear expression equal to zero, and then solving for $\hat{\beta}_t$. Before differentiating, the expression to be minimized can be rewritten as

$$\beta_t^T(F_0^T F_0 + (F_t - F_0)^T(F_t - F_0))\beta_t - 2\beta_t^T F_0^T F_0\hat{\beta}_0 + \hat{\beta}_0^T F_0^T F_0\hat{\beta}_0 + N\sigma_0^2.$$

If this is differentiated with respect to the vector β_t, the result is

$$2(F_0^T F_0 + (F_t - F_0)^T(F_t - F_0))\beta_t - 2F_0^T F_0\hat{\beta}_0.$$

By setting this expression to zero and solving for β_t, the closed form expression for $\hat{\beta}_t$ is found to be

$$\hat{\beta}_t = (F_0^T F_0 + (F_t - F_0)^T(F_t - F_0))^{-1} F_0^T F_0\hat{\beta}_0.$$

B. Coefficient Descriptions and Estimates

This appendix contains SDD-like estimates of coefficients for coefficients from Ref. 9, as well as estimates for 30-minute prediction horizon coefficients computed by some of the approaches discussed in this paper. A description of each factor is in table 1. These descriptions are from Ref. 9. The coefficients for various prediction approaches are shown in table 5. The coefficients shown here are computed by using all of the available data as training data (which was never done while generating the results in the paper).

Table 5. Sample Coefficient Estimates

Coefficient Name	SDD-like $\hat{\beta}_0$ Estimate	Robust $\hat{\beta}_{30}$ Estimate	Least Squares $\hat{\beta}_{30}$ Estimate	Robust Lasso $\hat{\beta}_{30}$ Estimate
AC	2.2	1.855	1.295	0
AD2	4723	0.05078	0.0	0
C2	0.2	3.943	2.673	0
C4	0.2	9.914	5.214×10^{-7}	0
NumHoriz	0.3	4.070×10^{-8}	5.772×10^{-11}	0
S5	1.2	6.530×10^{-8}	3.303×10^{-10}	0
S10	0.6	9.807×10^{-8}	9.974×10^{-11}	0
WBPROX	0.4	0.2287	1.808×10^{-8}	3.887
C14	0.0005	9.557×10^{-6}	7.612×10^{-7}	0
HDGVARI	0.0005	0.02546	0.1004	0

C. Lasso Coefficient Shrinkage

This section briefly describes the coefficient shrinkage method known as the lasso or ℓ_1 regularization.[10,12] The main idea is to add a constraint to the minimization problem that is solved to find the coefficients $\hat{\beta}_t$. This constraint puts an upper bound on the sum of the absolute value of the coefficients, which is the ℓ_1 norm of the coefficient vector. The factor values and $\hat{\beta}$ coefficients are normalized by the largest corresponding factor value that occurs in the data set so that each coefficient has roughly the same magnitude. The optimization problem (4) then becomes

$$\begin{array}{ll} \text{minimize} & (F_0\hat{\beta}_0 - F_0\beta_t)^T(F_0\hat{\beta}_0 - F_0\beta_t) + \beta_t^T(F_t - F_0)^T(F_t - F_0)\beta_t \\ \text{subject to} & \|\beta_t\|_1 \le s. \end{array} \tag{11}$$

A non-negativity constraint on the coefficients can also be added to this problem. This is done when solving for the coefficients for the prediction of the SDD-like formulation in this paper. This is a convex optimization problem that is implemented using the `CVX` package for specifying and solving convex programs in Matlab.[13] Here s is a parameter that can be tuned to control how small the sum of the absolute values of the coefficients should be. The effect of the ℓ_1 bound is to reduce the magnitude of the coefficients and also to encourage sparsity in the coefficient vector, thereby selecting factors to be used by the predictor.

Optimization problem (11) is used in an algorithm that finds a reasonable value for the parameter s, selects which factors to include in the predictor, and solves for the final robust coefficient estimates. The selection of s is based on a procedure described in Ref. 12. An estimate of the largest meaningful s value is determined by solving problem (4) with the training data and calculating the sum of the absolute values of the coefficients. This estimate of the largest meaningful bound s is denoted by $\bar{s}$. Several s values between 0 and $\bar{s}$ are selected to be evaluated. Next, cross-validation is used within the training data to evaluate the performance of each s value possibility. Problem (11) is solved for each s using a subset of the training data. Then, the RSS of the subset of the training data being used for testing is computed with predictors based on the resulting coefficients. The average and standard deviation of the RSS over all test data sets for each s are computed. Typically the average RSS decreases as s becomes larger. The selected s value is the smallest s with an average RSS that is within one standard deviation of the s value that produces the smallest average RSS.

Once the s value is selected, the problem (11) is solved with all of the training data. The solution typically contains several coefficients that are set to zero. The final step is to solve problem (4) using only those factor coefficients that were not set to zero by the previous lasso problem. All other coefficients are forced to be zero, so only a subset of the original factors are used by the robust lasso predictor.

Acknowledgments

The authors thank Jim Aldous of Mosaic ATM for his contributions to generating and validating the factor predictions. M. Bloem thanks Stefan Schuet of NASA Ames Research Center for his significant help with the theoretical aspects of this research.

References

[1] Krozel, J., Rosman, D., and Grabbe, S. R., "Analysis of En Route Sector Demand Error Sources," *Proc. of AIAA Guidance, Navigation, and Control Conference*, Monterey, CA, August 2002.

[2] Chatterji, G. B. and Sridhar, B., "Measures for Air Traffic Controller Workload Prediction," *Proc. of AIAA Aircraft, Technology, Integration, and Operations Forum*, Los Angeles, CA, October 2001.

[3] Massalonis, A., Callaham, M. B., and Wanke, C. R., "Dynamic Density and Complexity Metrics for Realtime Traffic Flow Management," *Proc. of 5th USA/Europe Air Traffic Management Research & Development Seminar*, Budapest, Hungary, December 2003.

[4] Laudeman, I. V., Shelden, S. G., Branstrom, R., and Brasil, C. L., "Dynamic Density: An Air Traffic Management Metric," Technical Report NASA/TM–1998-112226, NASA Ames Research Center, 1998.

[5] Kopardekar, P. and Magyarits, S., "Measurement and Prediction of Dynamic Density," *Proc. of USA/Europe Air Traffic Management R & D Seminar*, Budapest, Hungary, 2003.

[6] Kopardekar, P. and Magyarits, S., "Dynamic Density: Measuring and Predicting Sector Complexity," *Proc. of AIAA/IEEE Digital Avionics Systems Conference*, Irvine, CA, October 2002.

[7] Sridhar, B., Sheth, K. S., and Grabbe, S., "Airspace Complexity and its Application in Air Traffic Management," *Proc. of 2nd USA/Europe Air Traffic Management Research & Development Seminar*, Orlando, FL, December 1998.

[8] Klein, A., Rogers, M. D., and Leiden, K., "Simplified Dynamic Density: a Metric for Dynamic Airspace Configuration and NEXTGEN Analysis," *Proc. of AIAA/IEEE Digital Avionics Systems Conference*, St. Paul, MN, October 2008.

[9] Kopardekar, P., Schwartz, A., Magyarits, S., and Rhodes, J., "Airspace Complexity Measurement: An Air Traffic Control Simulation Analysis," *Proc. of 7th USA/Europe ATM Seminar*, Barcelona, Spain, July 2007.

[10] Boyd, S. and Vandenberghe, L., *Convex Optimization*, Cambridge University Press, Cambridge, UK, 2004.

[11] Bilimoria, K., Sridhar, B., Chatterji, G., Sheth, K., and Grabbe, S., "FACET: Future ATM Concepts Evaluation Tool," *Air Traffic Control Quarterly*, Vol. 9, No. 1, 2001, pp. 1–20.

[12] Hastie, T., Tibshirani, R., and Friedman, J., *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer-Verlag, New York, 2001.

[13] Grant, M. and Boyd, S., "CVX: Matlab software for disciplined convex programming," Web page and software, `http://stanford.edu/~boyd/cvx`, January 2009.