Case study: Influences of Uncertainties and Traffic Scenario Difficulties in a Human-In-The-Loop Simulation

Nancy Bienert, Joey Mercer, Jeffrey Homola, and Susan Morey
Airspace Operations Laboratory
San Jose State University/ NASA Ames Research Center
Moffett Field, CA, USA

Thomas Prevôt
Airspace Operations Laboratory
NASA Ames Research Center
Moffett Field, CA, USA

***Abstract*— This paper presents a case study of how factors such as wind prediction errors and metering delays can influence controller performance and workload in Human-In-The-Loop simulations. Retired air traffic controllers worked two arrival sectors adjacent to the Atlanta terminal area. The main tasks were to provide safe air traffic operations and deliver the aircraft to the metering fix within ±25 seconds of the scheduled arrival time with the help of provided decision support tools. Analyses explore the potential impact of metering delays and system uncertainties on controller workload and performance. The results suggest that trajectory prediction uncertainties impact safety performance, while metering fix accuracy and workload appear subject to the scenario difficulty.**

Keywords— trajectory uncertainties, difficulty, metering delay, decision support tools

I. INTRODUCTION

Air Traffic Management (ATM) systems are complex, dynamic, information-driven systems operated by humans. Traffic demand is predicted to grow steadily in the next 20 years and consequently, air traffic controllers will need to handle increasingly complex traffic situations [1]. To combat these problems, proponents of various ATM modernization programs (such as SESAR and NextGen) are developing new automation-based technologies to support controllers in maintaining a safe and efficient flow of air traffic, while doing so within reasonable levels of cognitive workload.

Trajectory Based Operations (TBO) will likely be a central element of future ATM systems. As such, it requires enhanced precision in predicting the individual flight trajectory, especially when controllers are tasked to reduce delays in the arrival time. This paper describes a study which investigates the impact of errors in the input data (wind forecast data, aircraft performance models and air speed profile), and the resulting trajectory prediction uncertainties' impact on controller workload. Air traffic controllers were asked to meter aircraft to a fix within ±25 seconds, without compromising the separation between aircraft. The count of loss of separation events, the metering fix delivery accuracy and the controller workload represent performance indicators to evaluate the controller ability to work with imprecise decision support tools. The goal was to find a point at which trajectory prediction uncertainties underlying the decision support tools would become unacceptable for the controllers [2]. This paper will investigate the question: Do the traffic scenarios have a bigger impact on performance and workload than error conditions?

The aim of this paper is not only to discuss the terms 'difficulty' and 'complexity' in context of controller workload and performance, but also to examine the influence of trajectory prediction uncertainties in this context.

II. BACKGROUND

A. Scenario difficulty

The term 'difficulty' appears repeatedly in the context of complexity definitions [3], whereas Sousa identifies a difference between complexity and difficulty [4]. An example in the air traffic environment shows the following case in Figure 1: The baseline traffic scenario (A) has 3 aircraft; the controller task is to monitor these aircraft for separation. A higher level of difficulty (B) can be reached by increasing the amount of traffic, whereas complexity (C) can be for instance achieved by a conflict situation. Instead of simply monitoring aircraft (as in B), this requires a more careful assessment of the situation to monitor for conflicts. Difficulty refers to the subject's memory, whereas the complexity task requires analyzing, evaluating and solving the situation [4].

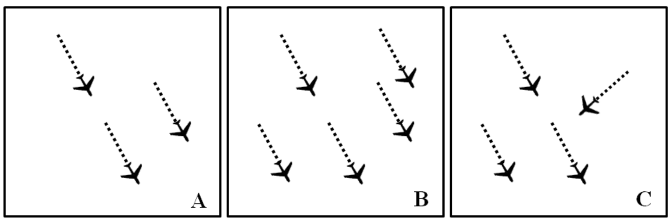

Figure 1. Baseline (A), Increased Difficulty (B), Increased Complexity (C).

Several studies used the variation of aircraft count as a regulator to evaluate controller performance and workload [5]. Schmidt proposed that the amount of time a controller needs to process a given event is a measure of complexity. Conflicts top the list with the highest process time. This notion of an event's complexity, combined with the event's frequency, is believed to lead to higher workload [6].

Other research studied different approaches and contributing factors to models of difficulty and complexity [7, 8], but, to the best of our knowledge, none of those utilized uncertainties and metering tasks to influence controller workload and performance in a Human-In-The-Loop simulation.

B. Trajectory Prediction Uncertainties

The term 'uncertainty' describes a state of doubt about the future [9] and can be caused by different factors. In "Common Methodology and Resources for the Validation and Improvement of Trajectory Prediction Capabilities," the uncertainty management is described as critical [10]. There is always a probability that a system's predictions can be unreliable, and would present controllers with incorrect recommendations from the decision support tools. Morey [11] examined how controllers cope with trajectory prediction uncertainties focusing on sector performance and automation usage. The results show that controllers developed strategies to compensate for the different levels of uncertainties while managing aircraft delay.

III. METHOD

A. Overview of study

A Human-In-The-Loop simulation, called Trajectory Prediction Uncertainty (TPU), took place during January 2013. Twelve retired air traffic controllers were separated into two test groups to work identical, independent simulations. Two controller of each team operated one high-altitude sector and one low-altitude sector, which fed the northwest corner of the Atlanta Terminal Radar Approach Control (TRACON). The main tasks were to provide safe air traffic operations and deliver the aircraft to the metering fix within ±25 seconds of the scheduled arrival time with the help of the following decision support tools [2, 11]:

- Delay display in data block
- Meter list at ERLIN
- Conflict list with time to conflict in data block
- Trial planning tools

Controllers worked traffic scenarios under different uncertainty conditions, where errors in forecast wind data and aircraft performance models were varied. The first part of the one-week study used a balanced test matrix with all conditions tested in two different traffic scenarios; a third scenario was introduced later, in the exploratory portion of the study.

Data collected during the simulation included subjective feedback from the participants, in form of the real-time workload ratings measured with the Air Traffic Workload Input Technique (ATWIT) [5], and objective performance indicators, such as delivery accuracy at the metering fix and the count of loss of separation events. More details regarding the experiment appear in [2].

B. Sources of uncertainties

The TPU study varied uncertainties inherent in the system's trajectory predictions to investigate the impact of errors on controller performance and workload.

Previous studies have noted that wind, aircraft performance model and weight are major sources for trajectory prediction errors [12, 13]. Table 1 shows an overview of the error sources used during the TPU simulation, and their influence on the trajectory prediction.

Table 1. Manipulated error sources in the TPU simulation.

Error source	**Implementation**	**Effect**
Wind forecast	Variation between forecast wind data and real environment	Ground speed, track, displacement of Top of Descent
Aircraft performance model	Variations in the aircraft weight factor	Different descent profiles caused by displacement of Top of Descent

The simulation investigated the impact of wind forecast errors, and their effect on trajectory prediction, across five levels. This paper will focus on three of those levels. The *no wind error* case (perfect prediction) signified an impossible case at present, but provided controller performance data in an uncertainty-free environment as a baseline reference. Current-day operations were assumed to have forecast errors of approximately 10 knots, which was represented by the *realistic wind error* case. In the *large wind error* case, the forecast winds differed from the actual winds by an average of 30 knots. Additionally, the wind forecast errors were configured as either positive or negative wind biases, representing wind over-predictions or wind under-predictions, as shown in Table 2.

Table 2. Variation of wind velocity forecast errors and error biases [kts].

	No error	**Realistic error**	**Large error**
Wind bias positive	0	10	30
(environment/forecast)	70/70	70/80	70/100
Wind bias negative	0	-10	-30
(environment/forecast)	90/90	90/80	90/60

The second simulated uncertainty is related to the system's aircraft performance models. Manipulations to the weight of individual aircraft served to impact their actual descent and climb profiles, thereby differing from the nominal descent and climb profiles assumed by the ground system. Translated as Top of Descent (TOD) or Top of Climb (TOC) errors (i.e., location of assumed TOD/TOC vs. location of actual TOD/TOC), a *no aircraft performance error* case served as a baseline (although unrealistically perfect) condition, errors around ±5% represented a *realistic aircraft performance error* case approximating current-day operations, and errors around ±25% a represent a *large aircraft performance error* case. The realistic and large errors were implemented as separate, normal distributions over arrival and non-arrival aircraft.

The TPU simulation also included flight technical errors (FTE), modeled as errors in the airspeed guidance accuracy. For all aircraft in the simulation, the Vertical Navigation

Guidance (VNAV) corrected speed deviations only when current speeds differed from the target speed by more than 10 knots. The presence of these errors was constant across all conditions.

C. *Design of traffic scenarios*

The basic scenario was created according to Kupfer [14]. An hour of recorded live traffic formed a base scenario, which required several adjustments in preparation for the simulation environment [15]. Since the study’s focus was on aircraft arriving from the northwest into Atlanta, the scenarios only included arrivals, departures and overflights directly concerning the test airspace and the surrounding sectors.

The resulting file was divided into two separate files; one containing only the arrivals into Atlanta, and the other including the overflights and departure traffic. Modifications to the initial conditions of the arrival aircraft allowed researchers to manipulate the delays against scheduled arrival times. Three scenarios were created with different amount of delay, in other words difficulty levels: Low, moderate, and high. The same overflight and departure traffic were added to the different arrival scenario files, ensuring that the circumambient traffic was consistent across all scenarios.

The method of building the three scenarios used only the no wind error environment. Other factors potentially affecting scenario complexity were similarly held constant across the three scenarios airspace factors (sector dimensions, standard flows etc.), traffic factors (density of traffic, ranges of aircraft performance etc.) and operational constrains (procedural restrictions, communication limitations etc.) [16], besides the previously mentioned uncertainty conditions and the amount of metering delay.

As Table 3 shows, the amount of delay at the metering fix ERLIN reveals noticeable differences between the scenarios. These differences are an indication of the amount of space between the arrival aircraft and not the number of aircraft. Larger delays would be expected when aircraft spaced closely together conform to the schedule. The same schedule then, in the presence of aircraft spaced farther apart, would result in less delay. During the TPU simulation, aircraft were scheduled two minutes apart at ERLIN.

Table 3. Delays at ERLIN in no error condition (no controller involvement).

	Scenario Difficulty		
	Low	**Moderate**	**High**
Average delay at ERLIN [min:s]	0:20	0:38	0:45

Figure 2 shows the airspace characteristics and the wind field (orange wind vectors in background) used in this study. Two major arrival traffic routes merged at the ERLIN metering fix: A western route over CALCO, and the northern route over NEUTO. The aircraft entered the high-altitude sector at or below FL350 and began their descent into Atlanta. Working below FL240, the low-altitude sector merged the flows at RMG, before delivering the aircraft to the TRACON at 12000 ft and 280 kts over ERLIN.

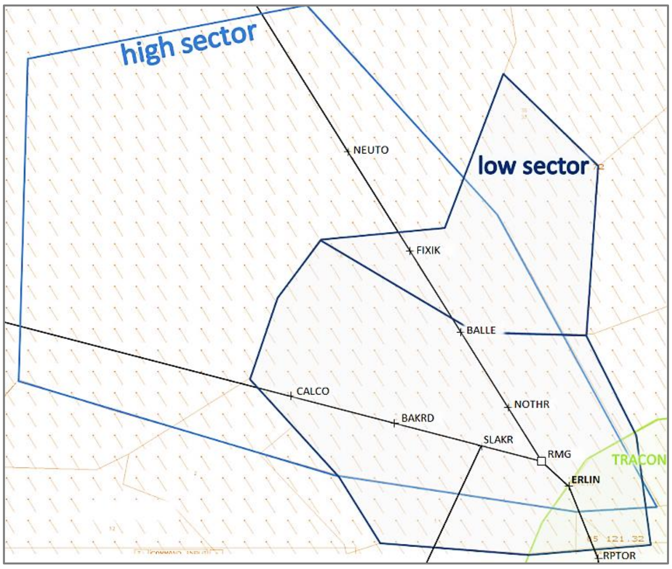

Figure 2. Depiction of the sectors for the arrival flow into Atlanta with northwest wind (low sector reaches from FL 100 to FL 239; high sector above FL 240).

D. *Configuration of study*

The conducted study used four different error conditions [2]. This paper examines a subset of three of the study’s conditions, identified as the pairings of the three wind forecast error cases with the matching aircraft performance error cases.

These three conditions were examined with the three traffic scenarios at 55 minutes each. It is worth noting that the traffic scenario designed with high amounts of delay (i.e., high difficulty), was not simulated in the no error condition since it was used in the exploratory part of the study. Additionally, the wind forecast error’s bias direction was distributed across the traffic scenarios and uncertainty conditions as shown Table 4.

Table 4. Combination of the errors and scenario difficulties.

Error / **Difficulty**	**No error**	**Realistic error**	**Large error**
Low	0kts	-10kts	-30kts
Moderate	0kts	10kts	30kts
High	-	-10kts	-30kts

IV. RESULTS

This section describes analyses of how the uncertainty conditions and the characteristics of the exercised traffic scenarios impacted controller performance and workload. Data from open-loop runs (no controller involvement) provided reference measurements, helping to isolate the natural characteristics of each traffic scenario, and are presented first. Followed by results from the high sector of the Human-In-The-Loop simulation, which offer insights regarding how the controllers reacted to the varying conditions.

A. *Open-loop runs*

The open-loop runs provided the opportunity to understand how the wind impacted the amount of metering delay, and across the different arrival flows. Table 5 shows how the schedule delay evolved to the two wind bias directions. As the uncertainty conditions progressed from 'no errors' to 'large errors,' the negative wind bias produced decreasing amounts of delay (as measured when the aircraft entered the high-altitude sector). Used with the low- and high-difficulty scenarios, the negative wind bias presented the delay as smaller than the reality, due to the system's underestimation of the winds. Similarly, the positive biased winds, paired with the moderate-difficulty scenario, presented the metering delays as being larger than they actually were, caused by the system's overestimation of the winds.

Table 5. Metering fix delay time, measured at the sector entry [min:s].

Error / Difficulty	No error	Realistic error	Large error
Low	1:08	1:04	0:47
Moderate	3:26	3:23	3:35
High	-	3:50	3:08

In addition to the wind's speed, its direction influenced the metering delay values, and did so differently across the two arrival flows. The winds had a greater effect on the northern flow over NEUTO, where they were nearly a direct tail wind. The wind's direction relative to the western flow over CALCO had a larger cross-wind component, and therefore had less impact. Although wind contributed to trajectory prediction errors on both arrival flows, the effect was stronger on the northern flow [17].

These analyses confirmed that the direction and the magnitude of the wind forecast error substantially influenced how the traffic scenarios unfolded. Consequently, the interpretation of the Human-In-The-Loop results warrant consideration of this finding.

B. *Human-In-The-Loop simulation*

1) *Controller Performance*

Figure 3 and 4 illustrate the two performance indicators (metering fix delivery accuracy and loss of separation events) in relation to scenario difficulty (Fig. 3), and level of uncertainty (Fig. 4). The metering fix accuracy metric describes an aircraft's successful delivery at ERLIN. Aircraft crossing the metering fix counted as successful if the following criteria were met:

- The aircraft arrived within +/-25 seconds,
- met the altitude restriction at ERLIN (+/-300 ft), and
- flew directly over ERLIN.

Loss of separation events were recorded when the separation between two aircraft was less than 4.5 nm laterally and 800 ft vertically. A loss of separation event was only considered if its duration was longer than 12 seconds.

When averaged across all three uncertainty conditions, the metering fix delivery accuracy was generally high across all three traffic scenarios, and the highest during the moderate-difficulty traffic scenario (see Figure 3). This data challenges the hypothesis that metering fix delivery accuracy would worsen as the traffic scenario's difficulty level increased, and instead suggests that controllers were able to accurately deliver aircraft to the metering fix regardless of the automation's estimates of how much delay was present.

Interestingly, unique to the moderate-difficulty scenario was the fact that it used the wind forecasts with positive bias errors, while the low- and high-difficulty traffic scenarios used the negative wind bias. The negative bias wind errors caused the automation to present the controller with less delay than in reality. An example will demonstrate the effect: The controller receives an aircraft predicted to arrive at ERLIN 90 seconds early. To meet the scheduled time, the automation recommends a slower speed. However, incorporated into the automation's predicted arrival time is the forecasted tail wind, which is weaker than the actual tail wind. As a consequence of this incorrect assumption, the automation believes the problem to be easier than it actually is, and thus suggests a speed reduction that is likely too small. As the situation unfolds, the automation's predicted arrival time will update, and the controller may need to make another speed reduction. The positive wind error bias produces the opposite situation, in which the automation predicts an aircraft's arrival time based on stronger-than-actual winds. Here, the automation, believing the problem to be worse than it actually is, suggests a speed that may be too much of a reduction, effectively overcorrecting.

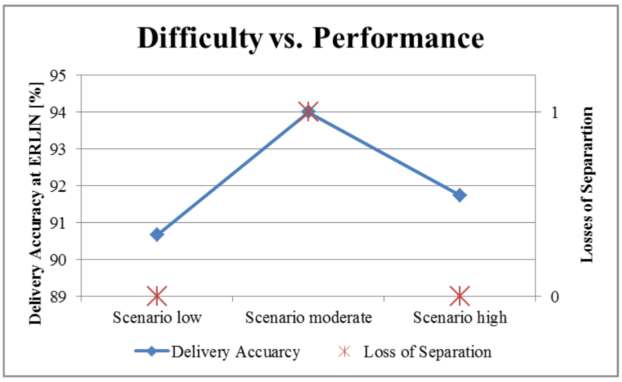

Figure 3. Influence of difficulty on controller performance.

When compared across all three uncertainty conditions, data averaged across all three traffic scenarios showed similar trends in the delivery accuracy data (see Figure 4). However, the simulation's only recorded separation violation occurred in a moderate-difficulty scenario used in conjunction with the large uncertainty condition. It is possible that either (or both) the scenario difficulty or the uncertainty condition contributed to the occurrence of the loss of separation events. Yet, the data does not appear to support the hypothesis that safety performance would worsen as the traffic scenario's difficulty level increased; instead, the data suggests safety performance worsened as the system's prediction uncertainty level increased. The evaluation of the post-run questionnaire in [2] supports this idea. Controllers rated the high-difficulty scenario in the large error condition as unsafe and unmanageable.

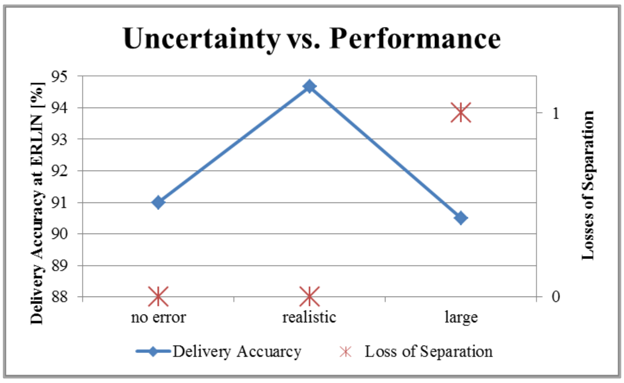

Figure 4. Influence of uncertainty on controller performance.

2) Controller Workload

The low, moderate and high levels of scenario difficulty, and the metering tasks associated with those scenarios, are reflected in the workload measurements. The ATWIT uses Workload Assessment Keypads, which inquire the controller to rate workload based on a modified six-point scale (1 as low workload, 6 as high workload) in a three-minute interval during simulations.

Both the average and peak workload ratings increase with the difficulty of the scenario (see upper portion of Figure 5). The trend was different when analyzed in relation to the uncertainty conditions, suggesting that the controllers felt they did not work harder in conditions with higher levels of system uncertainty (see lower portion of Figure 5). Rather, the larger metering delays appear to have impacted the controllers' workload.

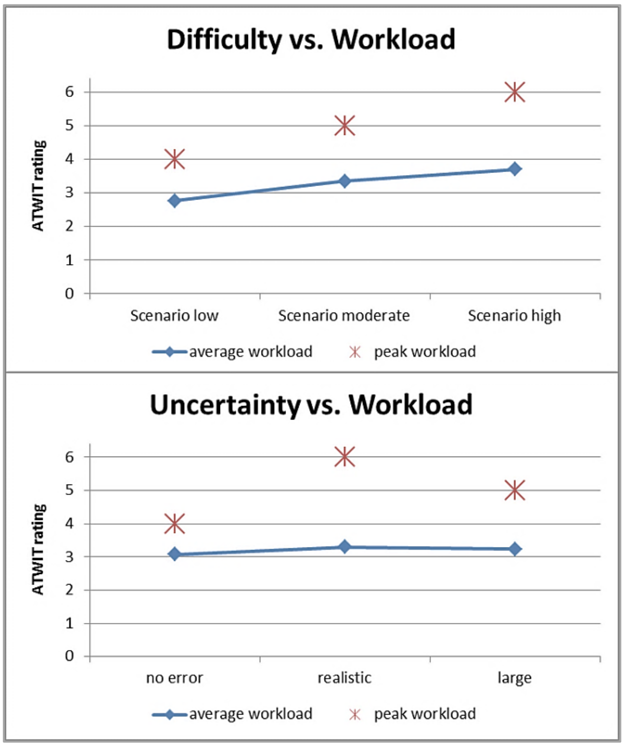

Figure 5. Influence of difficulty and uncertainty on controller workload

Another aspect to explain the workload – performance relationship is that there could have been an under-load in the low-difficulty scenario, and an overload in the high-difficulty scenario. Consideration of a theoretical 'U-shaped curve' plotting performance against workload, offers a classical explanation of performance decrements in 'too little' and 'too much' environments, however the controllers' workload data (a value of 3, on average) suggests their workload was within an acceptable range.

3) Interaction of difficulty and complexity

Figure 6 adapts the principle from Figure 1, but applied to the route structure used in the simulation. A generic traffic example shows an aircraft flow with low delays (A), in which the controller can achieve the scheduled times with only small speed adjustments. In (B), a higher delay scenario produces increased delays. Delays cannot be absorbed with speed changes alone, and require the controller to find alternatives to achieve the scheduled times. This typically results in vectoring the aircraft (C). Such operations are not only workload-intensive for both controllers and pilots, but also change the nominal traffic flow patterns, increasing the number of converging and crossing trajectories. This example illustrates the close relationship between difficulty and complexity, suggesting that difficult situations can easily lead to complex situations when aircraft are being vectored to remain on time.

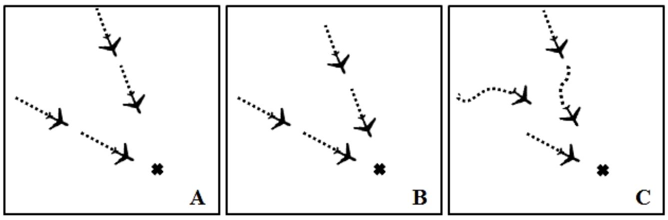

Figure 6. The progression of nominal (A), to difficult (B), and then complex (C) traffic situations, in the context of metering.

The following figure shows the flown trajectories included these analyses, for each scenario difficulty level across each uncertainty condition. The color gradient represents the Indicated Airspeed, associating warmer colors with faster speeds (e.g., red = 400 kts), and cooler colors with slower speeds (e.g., green = 225 kts, cyan = 160 kts). Differences are clearly noticeable between the three difficulty levels. The speed profiles and flight paths support the finding that workload increased as a function of scenario difficulty level. As described in the example scenario illustrated by Figure 6, higher metering delay values were associated with more vectoring of aircraft. Such maneuvering increases the variability of the traffic flow patterns, which in turn increases the number of trajectories in potential conflicts, and as a result, increases complexity.

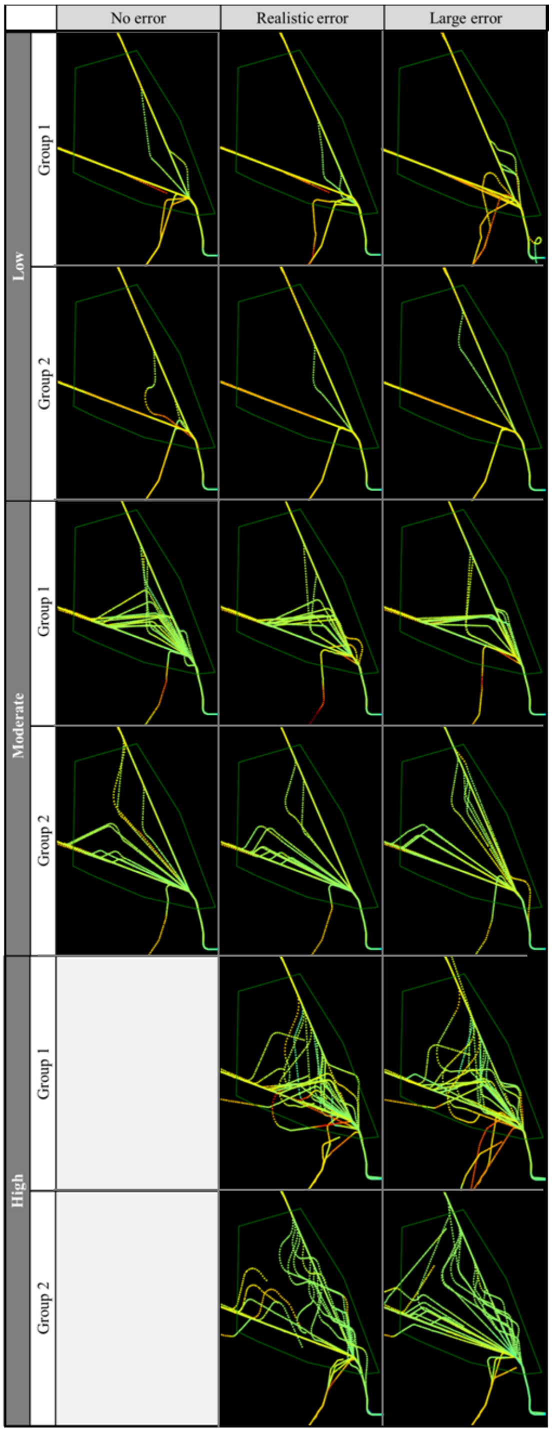

Figure 7. Comparison of flown trajectories under different conditions

V. CONCLUDING REMARKS

The goal of this case study was to assess whether the inherent characteristics of traffic scenarios have a bigger impact on performance and workload than uncertainties in the underlying system's prediction capabilities. It appears that the error conditions have an impact on the safety performance, while the metering fix delivery accuracy and workload seem more dependent on the scenario difficulty. Finally, this study acknowledges the converging theories on difficulty and complexity. Both terms can be specified individually, but difficulty has a major influence on complexity. This paper is a case study; hence the generalized conclusions originate from limited data points from a single simulation. Further studies should investigate more in into this topic to validate the results.

ACKNOWLEDGMENT

This work was conducted as part of the Functional Allocation and Separation Assurance research focus area, funded by the Concept and Technology Development Project within NASA's Airspace Program. The authors acknowledge the contributions of many individuals in the Airspace Operations Laboratory and the Human-Systems Integration Division's system support group.

REFERENCES

[1] Federal Aviation Administration, "FAA Aerospace Forecast, Fiscal Years 2013-2033," Washington D.C., 2013.

[2] J. Mercer, S. Green, T. Prevot, N. Bienert, A. N. Gomez, J. M. Kraut, and S. E. Morey, "The Impact of Trajectory Prediction Uncertainty on Air Traffic Controller Performance and Acceptability," 2013, [AIAA Aviation 2013, Los Angeles, California].

[3] C. Meckiff, R. Chone, and J.-P. Nicolaon, "The Tactical Load Smoother for Multi-Sector Planning," 1998, [2nd USA/Europe Air Traffic Management R&D Seminar].

[4] D. A. Sousa, How the Brain Learns, 3nd ed., 2006.

[5] E. S. Stein, "Air Traffic Controller Workload: An Examination of Workload Probe," 1985.

[6] D. K. Schmidt, "On Modeling ATC Work Load and Sector Capacity," 1976.

[7] R. H. Mogford, J. A. Guttman, S. L. Morrow, and P. Kopardekar, "The Complexity Construct in Air Traffic Control: A Review and Synthesis of the Literature," 1995.

[8] B. Hilburn, "Cognitive Complexity in Air Traffic Control: A Literature Review," 2004.

[9] http://www.collinsdictionary.com/dictionary/english/uncertainty, Collins, 2012, [Cited: August 13, 2013].

[10] FAA/Eurocontrol, "White Paper – Common Methodology and Resources for the Validation and Improvement of Trajectory Prediction Capabilities – A Proposed Plan for Action," 2004.

[11] S. E. Morey, T. Prevot, N. Bienert, J. M. Kraut, L. H. Martin, & J. Mercer, "Controller Strategies for Automation Tool Use under Varying Levels of Trajectory Prediction Uncertainty," 2013, [AIAA Aviation 2013, Los Angeles, California].

[12] W. Chan, R. Bach, and J. Walton, "Improving and Validating CTAS Performance Models," 2000, [AIAA Guidance, Navigation, and Control Conference and Exhibit 2000, Denver, Colorado].

[13] S. Mondoloni, M. Paglione, S. Green, "Trajectory Modeling Accuracy for Air Traffic Managment Decision Support Tools," 2002, [ICAS 2002, Toronto, Canada].

[14] M. Kupfer, C. D. Cabrall, T. J. Callantine, J. R. Homola, and J. Mercer, "Designing Scenarios for Controller-in-the-Loop Air Traffic Simulations," 2013, [AIAA Modeling and Simulation Technologies Conference 2013, Boston, Massachusetts].

[15] T. Prevot, "Exploring the Many Perspectives of Distributed Air Traffic Management: The Multi Aircraft Control System MACS," 2002.

[16] J. Histon, R. J. Hansman, B. Gottlieb, H. Kleinwakr, S. Yenson, "Structural Considerations and Cognitive Complexity in Air Traffic Control," 2002, [DASC 2002, Irvine, California].

[17] J. Mercer, T. Prevot, S. Green, N. Bienert, S. Hunt, J. Kraut, et al., "Differing Air Traffic Controller Responses to similar Trajectory Prediction Errors," 2014, in press, [ICAS 2014, St. Petersburg, Russia].