

# COMMON AIRCRAFT PERFORMANCE MODELING EVALUATION TOOLS AND EXPERIMENT RESULTS

*Carlos Garcia Avello, EUROCONTROL DAS-ATS, Brussels, Belgium*  
*Husni Idris PhD, Robert A. Vivona, TITAN Corporation, Billerica, MA*  
*Steve Green, AIAA Associate Fellow NASA-AMES, Moffet Field, CA*

## Abstract

This paper describes a set of tools, methodology, and metrics of key performance indicators developed for the analysis of different aircraft performance models approaches. These tools, methods, and metrics, developed collaboratively by Eurocontrol, the FAA and NASA, are proposed for common use throughout the Air Traffic Management community. The paper also describes a study that analyzed two different aircraft performance modeling methods for a common medium twin-jet transport. Aircraft performance modeling errors are presented in terms of rate of climb and descent over the aircraft's nominal performance envelope. The results indicated that performance model errors may vary significantly over an aircraft's flight envelope. The study also demonstrated the ability of this methodology to identify if and where performance model errors are significant.

## Introduction

Models of aircraft behavior are used within nearly every automation element within the Air Traffic Management (ATM) system including airborne components: Decision Support Tool (DST) capabilities for air traffic control and traffic flow management, Flight Management Systems, automation for flight planning/processing and flight simulators. The performance of such automation systems depends on the performance of the underlining Trajectory Predictor (TP). TP performance itself is limited by two major factors, the quality of input data and the underlying modeling technique used to emulate aircraft behavior.

A broad range of aircraft performance modeling methods exists, and many have been implemented for various TP applications. However, there is very little in the way of objective

comparisons of alternative modeling methods (e.g. kinetic vs. kinematic) to determine and document the relative advantages and disadvantages. Such information, if made available, would benefit the developers and users of automation. For example, the information would enable stakeholders to perform trade offs studies for choosing the best modeling method for their application.

This paper proposes methods, metrics of Key Performance Indicators (KPIs) for objective analysis of Aircraft Performance Model (APM) approaches.<sup>1</sup> The use of these methods and metrics is illustrated through an initial study.

The paper begins with a high level description of various techniques employed for modeling aircraft performance. A description of tools and proposed methodology follows. The paper then presents the analysis and results of an initial study, including a proposed set of metrics to support the methodology. The paper closes with a set of conclusions and next steps.

This paper is part of a set of papers produced by the Core Team of Action Plan 16 of the FAA-Eurocontrol Coordination Committee. Other papers in this set address the proposed structure of a common Trajectory Predictor [1], the Validation methodology [2], the validation metrics to apply [3] and a common language to describe aircraft intent [4].

---

<sup>1</sup> These methods, metrics and tools, developed collaboratively by Eurocontrol, the FAA and NASA, are proposed for common use throughout the Air Traffic Management community.

## Aircraft Performance Modeling Approaches

A survey of existing APM approaches was conducted to identify specific characteristics to discriminate one APM approach from another [5]. The scope of this study was limited to TP applications for ATM. We did not consider full six degrees of freedom APMs or the level of sophistication of the ones used in airline training simulators. Instead, this study analyzes two models intended to support TPs for ATM decision support. One model was based on a purely kinematic approach, while the other was based on a force-based kinetic approach.

While some differences in model performance may be directly due to differences in the characteristics of an APM model (i.e., the difference between a kinetic vs. purely kinematic approach), many differences in performance may be due to other attributes such as the quality of the input data used to create the performance data for either approach.

A compelling reason for studying two or more model approaches under a set of controlled and repeatable test conditions is to begin to understand and document the sensitivities of APM performance as a function of APM approach. The goal is to develop an objective set of advantages and disadvantages across the range of potential APM approaches and implementations. Such data will provide TP designers with the information necessary to optimize the design of their APM to meet the specific requirements demanded of their TP.

For example, one school of thought argues that kinetic performance models provide a potentially higher level of fidelity (i.e., accuracy over the flight envelope) because the basis of the modeling stems directly from the underlying physics governing the flight through the forces of thrust, drag, lift and weight. The kinetic approach also lends itself to the approximation of the limits of the flight envelope. On the other hand, purely kinematic approaches are often considered by some to be a “lower fidelity” method. Instead of directly modeling the actual physical equations of motion (representing the physics of forces acting on a point mass), kinematic approaches directly model key

performance elements, such as Rate Of Climb (ROC) or Rate Of Descent (ROD), as a function of the external parameters affecting it.

Since most TP applications are interested more in performance parameters like ROC, as opposed to Thrust or Drag, kinematic approaches can often provide sufficient accuracy while offering other advantages in the areas of simplicity and computation load. The performance model evaluation process presented here offers a method to objectively assess such considerations and trade offs that TP designers must consider when choosing an APM approach to implement. Assuming a well designed set of KPIs and test cases, the evaluation process provides the TP designer with a set of quantitative and qualitative KPIs representing the unique “capabilities” (relative advantages and disadvantages) of alternative modeling approaches.

As extensive literature is available to describe various kinetic and kinematic approaches, this paper will concentrate on methods for analyzing their relative performance, benefits and limitations.

## Analysis Tools

A Platform for Aircraft Modeling Performance Analysis (PAMPA) has been developed as a common framework for analyzing one or more aircraft performance models and to validate their accuracy against reference (“truth”) data.<sup>2</sup> The following sections describe the main components of the PAMPA platform.

## Management of Reference Data

A major challenge for the performance analysis of an APM approach stems from the significant amount of data needed to adequately describe the performance of today’s transport aircraft over its complete operational envelope. The extent of this challenge is clearly illustrated by the scope of the data manufacturers use today in their performance model programs.

For most of the developers of APMs, the analysis and validation process has been limited to

---

<sup>2</sup> The commonality of the PAMPA framework has been verified using several APMs from across the US and Europe.

the amount of reference data used to produce the APM. This does not guarantee the correct behavior of the model over the whole flight envelope.

The handling of these large amounts of reference data has been streamlined with PAMPA through the development of interfaces to the performance programs of several major manufacturers, and storing the data into PAMPA's tabular reference data base.

### ***Management of the Comparison Process***

The comparison process consists of three elementary steps.

- Selection of performance models and aircraft types to be analyzed (among those already interfaced to PAMPA).
- Selection of performance parameters to be analyzed (such as Rate of Climb/Descent (ROC/D)).
- Selection of the desired range and granularity of the independent variable(s) to be analyzed. (e.g. selecting aircraft weight between 60 and 80 tons, in increments of 5 tons)

The range of input parameters may be selected as a subset of the reference data available (for cases where the analysis is only desired over a part of the flight envelope), or as a superset when limited reference data is available but the user still wants to analyze the behavior of the model over the entire flight envelope.

### ***Comparison Results Data Base***

The results of a comparison scenario, executed by the comparison engine, are fed into a comparison data base which serves as input for the analysis and display of results. The database includes, for each performance parameter and aircraft type studied, the input parameters for the ranges defined, the reference (truth) data, the output of each APM and KPI (e.g. average signed error) selected for evaluation.

This is a stand alone data base, with fully documented contents that can be readily used by other analysts, through the R&D community, whether they use PAMPA's analysis tools or others such as Matlab macros.

### ***Comparison Results HMI***

Although not used for the study described in this paper, a comparison results Human Machine Interface (HMI) was developed to help the user/analyst digest the potentially large amount of results that would be generated by a typical exercise, and present the information in a manner that is clear and concise to the user. The comparison HMI consists of several components that enable the user to progressively delve into the analysis and focus on areas of interest within and across the flight envelope.

### ***Methodology for Aircraft Performance Models Evaluation***

The work on metrics and methodology presented hereafter is a key component of a broader set of TP-related metrics and analysis.<sup>3</sup> [3, 6]

The methodology applied within this study proposes a three step evaluation of the performance of any APM.

The first step consists of evaluating any model performance parameter (e.g. rate of climb/descent) over the complete flight envelope. This step is illustrated in the *Accuracy and Goodness of Fit over the Entire Flight Envelope* subsection within Analysis and Results.

The second step of the methodology aims at evaluating any model performance at the nominal operating conditions. This is a key step in the evaluation methodology proposed as the ATM system objective is to move aircraft operators as little as possible away of the nominal profile they want to fly. However due to various commercial policies, this nominal operating conditions are operator dependent and care needs to be taken when selecting them. This is illustrated in the *Accuracy at and around Nominal Operating Conditions* subsection within Analysis and Results.

Finally the third step of the evaluation methodology completes the process by performing a sensitivity analysis of any model performance parameter and this over the entire range of the relevant input parameter. This step is illustrated in

---

<sup>3</sup> This work was performed under a combination of US/Europe Action Plan 16 and CARE-TP activities [3,8].

the *Sensitivity to Input Parameters* subsection within Analysis and Results.

## Initial Study

This initial study was performed to illustrate the use of the proposed tools, methodology, and metrics. Whereas the tools and methods were presented above, the metrics are introduced within the Analysis and Results section that follows.

With consideration for resource and schedule constraints, the scope of this initial study was limited to the analysis of two different models that are differentiated by characteristics such as their technique for modeling physics, the reference data, and process used to derive and validate each models data prior to this study. The scope was also limited to one aircraft type, a popular medium size twin turbofan commercial transport.

- Model 1 is a kinetic APM based on tables of thrust and drag reference data for the dynamic calculation of thrust and drag needed to evaluate the kinetic equations of motion. The drag data was derived directly from a manufacturer’s performance program. The thrust data however, was adapted from a surrogate model to meet the deadline of this conference publication.
- Model 2 is a kinematic APM based on polynomial representation of kinematical performance parameters (e.g. rates of climb and descent as a function of aircraft state). This model was derived from the same source as model 1.

## Analysis and Results

The following sections present the KPI and analysis results for each of the three steps of the methodology.

### *Accuracy and Goodness of Fit over the Entire Flight Envelope*

The following set of KPIs has been selected for the metrics to capture the accuracy of an APM over the entire flight envelope:

- Absolute error magnitude averaged over the entire envelope of input parameters

- Signed error magnitude averaged over the entire envelope of input parameters
- Error standard deviation over the entire envelope of input parameters
- Maximum error over the entire envelope of input parameters

For the goodness of fit, it has been decided to use R-squared, applied to distances (or error value) between the reference data and the model data as a KPI.

Table 1 and 2 summarize these KPIs for the 2 APMs evaluated and show the average absolute error, the average signed error, the standard deviation of the average signed error, the maximum absolute error and the R-squared goodness of fit previously defined..

**Table 1: Model 1**

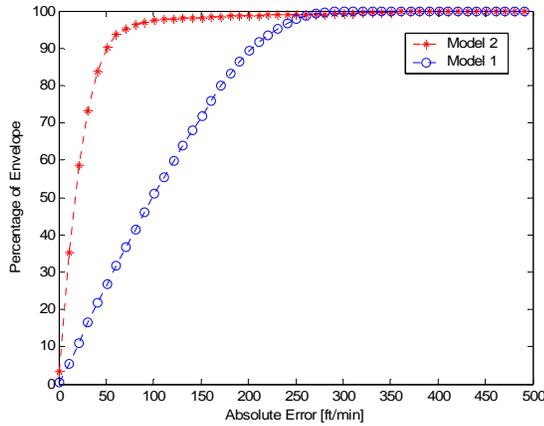
<b>Model 1 Error [ft/min]</b>	<b>ROC CAS</b>	<b>ROC Mach</b>	<b>ROD CAS</b>	<b>ROD Mach</b>
Average absolute	106.6	101.0	132.1	173.0
Average signed	67.7	71.8	-33.4	-13.6
STD signed	107.4	105.3	161.6	212.4
Maximum absolute	301	389	594	608
R-squared	0.96	0.91	0.93	0.84

**Table 2: Model 2**

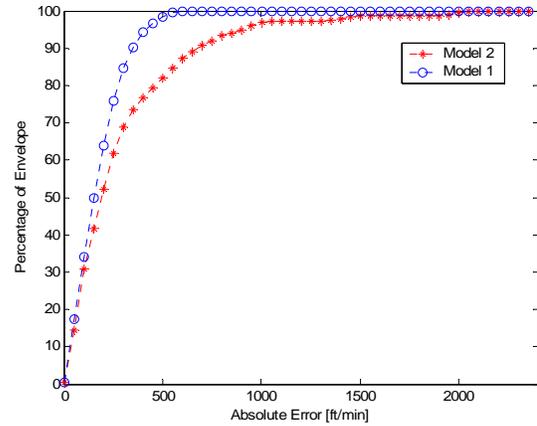
<b>Model 2 Error [ft/min]</b>	<b>ROC CAS</b>	<b>ROC Mach</b>	<b>ROD CAS</b>	<b>ROD Mach</b>
Average absolute	26.6	50.0	50.3	294.2
Average signed	2.58	23.4	-17.7	36.2
STD signed	48.8	62.6	66.6	442.9
Maximum absolute	513	272	504	2054
R-squared	0.99	0.97	0.99	0.28

To give an idea of the magnitude of the errors presented in the tables, a typical ROC for this aircraft type is in the range 2000-3000 ft/min and a typical ROD may be in the range 3000-4000 ft/min.

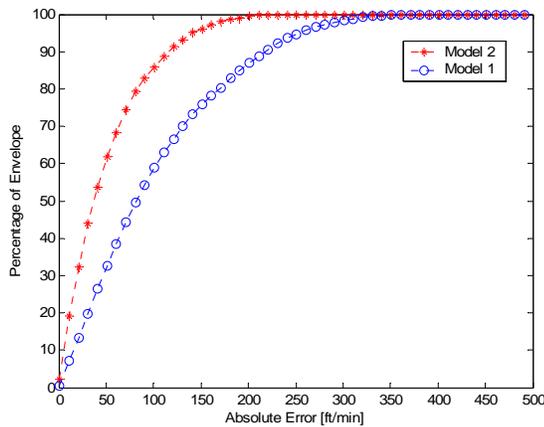
Figures 1 through 4 provide the evolution of the average absolute error of the two models as a function of the percentage of coverage of the whole flight envelope for the performance parameters ROC\_CAS, ROC\_M, ROD\_CAS and ROD\_M.



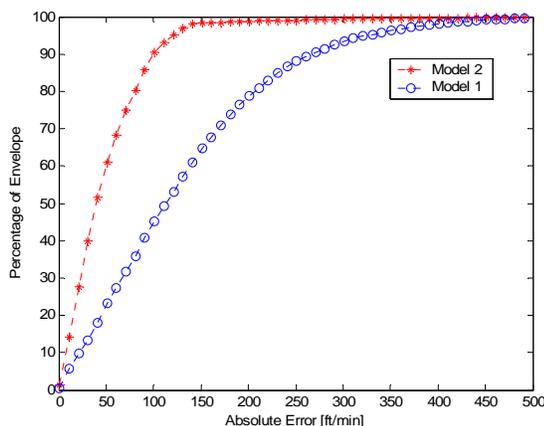
**Figure 1: Absolute Error versus Percentage of Flight Envelope for ROC\_CAS**



**Figure 4: Absolute Error versus Percentage of Flight Envelope for ROD\_Mach.**



**Figure 2: Absolute Error versus Percentage of Flight Envelope for ROC\_Mach**



**Figure 3: Absolute error versus percentage of flight envelope for ROD\_CAS**

With these results the user can decide what the acceptable error limit for its application is and therefore identify the model which covers the wider range of the flight envelope for the performance parameters selected.

***Accuracy at and around Nominal Operating Conditions***

Nominal conditions for the phases of flight covered in the study (Climb and Descent in clean configuration) are the ones provided by the manufacturer for the specific aircraft type analyzed.

This step of the methodology and the associated metrics is key to the evaluation of the performance of any APM as it covers the area of the flight envelope most frequently used. However it is clear that the accuracy requirement at extreme values for some input parameters like temperature may not be as important as the one for weight variations.

In the case of the medium twin-jet transport used in this initial study, the following nominal operating conditions have been derived from manufacturer data:

	<b>CAS</b>	<b>Mach</b>	<b>Weight</b>	<b>Delta ISA</b>
Climb	320 kt	.8	75 tons	0 degree
Descent	320 kt	.8	65 tons	0 degree

And for the input parameters affecting the performance parameters the following ranges have been used (Different Altitude and weight ranges

have been used for the different performance parameters):

	CAS	Mach	Delta ISA
All parameters	220/350 kt	.6/.8	-20/+20 degree

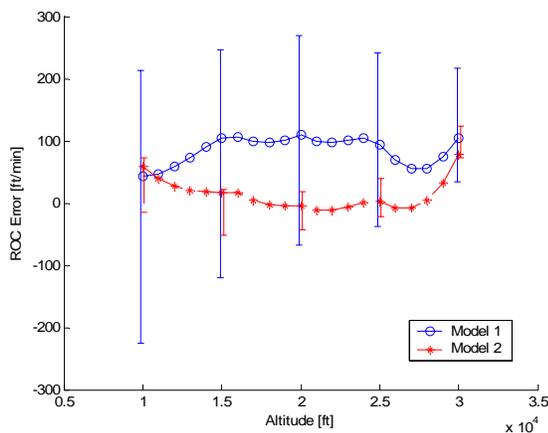
	FL	Weight tons
Climb CAS	100/300	60/80
Descent CAS	300/100	60/70
Climb Mach	270/390	60/80
Descent Mach	390/270	60/70

These ranges for input parameters are used every time any KPI for any specific performance parameter is averaged over the envelope of one of the input parameters whilst ensuring that the combination always stays within the flight envelope.

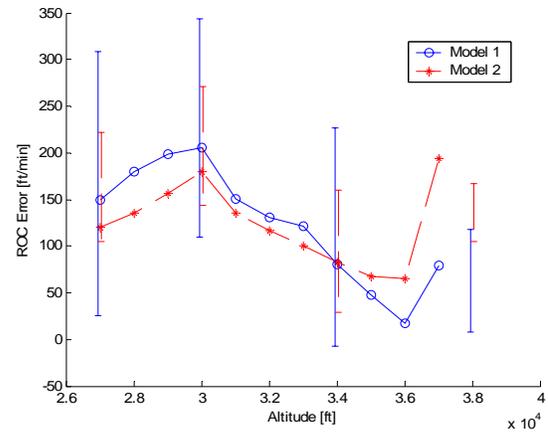
We have identified the following KPIs to be considered in the metrics applied for the accuracy of an APM:

- Signed error magnitude at nominal values
- Error standard deviation in a window around the nominal values

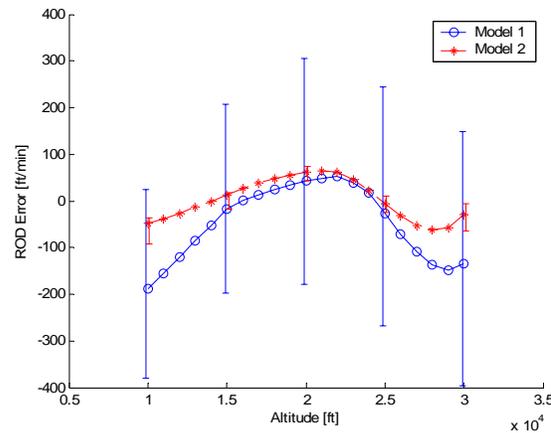
Figures 5-8 present the evolution of the error for ROC and ROD as a function of altitude. For convenience, the vertical lines present the standard deviation over the range of input parameters at different altitude intervals for the CAS and MACH phases.



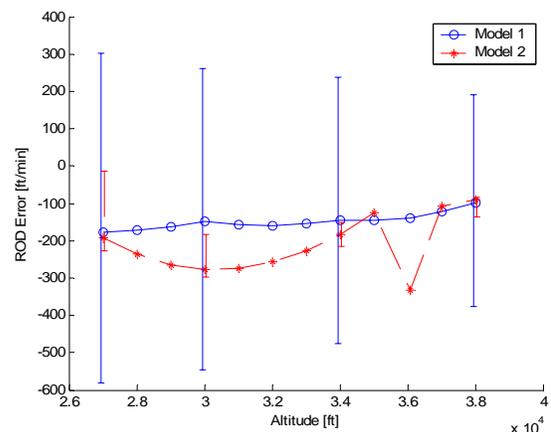
**Figure 5: ROC Signed Error vs Altitude for CAS 300 kt at Nominal Climb Operating Conditions**



**Figure 6: ROC Signed Error vs Altitude for Mach .8 at Nominal Climb Operating Conditions**



**Figure 7: ROD Signed Error vs Altitude for CAS 300 kt at Nominal Descent Operating Conditions**



**Figure 8: ROD Signed Error vs Altitude for Mach .8 at Nominal Descent Operating Conditions**

Further analysis can be performed to differentiate the impact of Delta ISA and weight on the standard deviation around the nominal conditions and therefore identify which of these two parameters requires further study.

This last figure shows that the Model 2 error suddenly increases as the altitude passes the tropopause and returns to its normal pattern. Although the reason was not validated in time for publication, it is believed that this may be due to a modeling assumption that is not in line with the reference data (with respect to accounting for the hysteresis discontinuity coming from below or from above the tropopause).

These four figures show that whilst the fidelity of the different APMs at nominal conditions does not vary significantly, the error standard deviation due to the range of the input parameters could vary significantly from a model to another.

### Sensitivity to Input Parameters

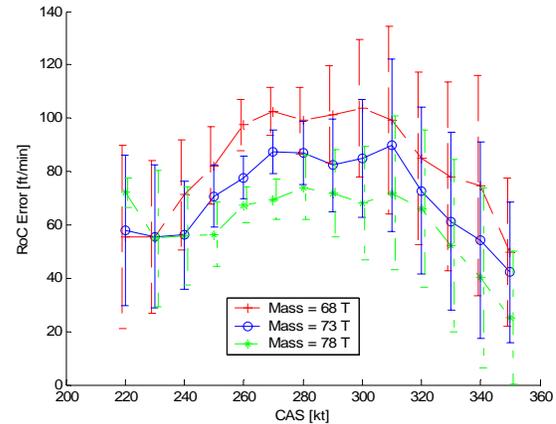
This section of the metrics is not directly related to the intrinsic accuracy of the APMs, but aims at capturing how sensitive the accuracy of the APM output is as a function of several key input parameters. Some APM input parameters may have poor accuracy or granularity that may contribute as much error, or more, compared to errors in the model itself. Related KPIs will help the potential user of the APM to identify which parameters affect the selected APM performance the most and develop its application accordingly. The following KPIs have been selected for this sensitivity analysis:

- Signed error magnitude
- Standard deviation

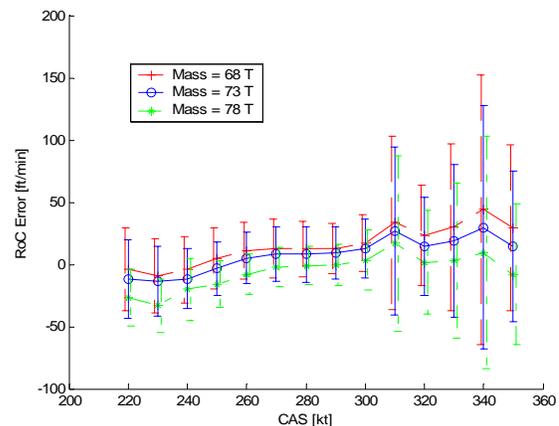
To show the impact of all the input parameters affecting any performance parameter in a concise manner, the following figures show, for each of the models, the evolution of the error as a function of the speed, while the variation of the error due to other input parameters is shown as vertical shifts in the curve or as vertical error bars.

Figure 9 shows for each model the sensitivity of ROC\_CAS error to speed and weight for nominal ISA and Figure 10 shows the sensitivity of ROD\_CAS error to speed and ISA for nominal

weight. In both figures the error at each speed value is averaged over the altitude range (FL 100 to 300) and the error bars show the standard deviation of the error over the altitude range. The results indicate that the accuracy of the models studied degrades at the edges of the flight envelope (e.g. CAS close to the maximum).

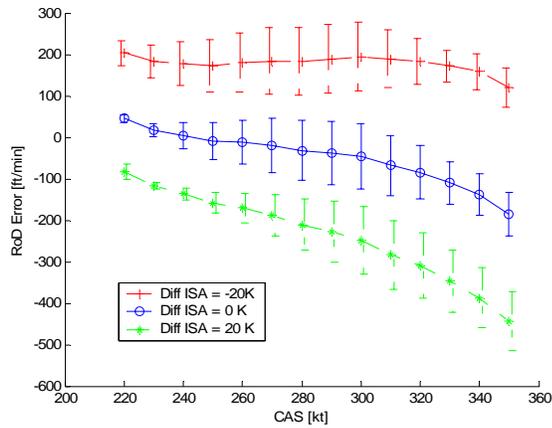


(a)

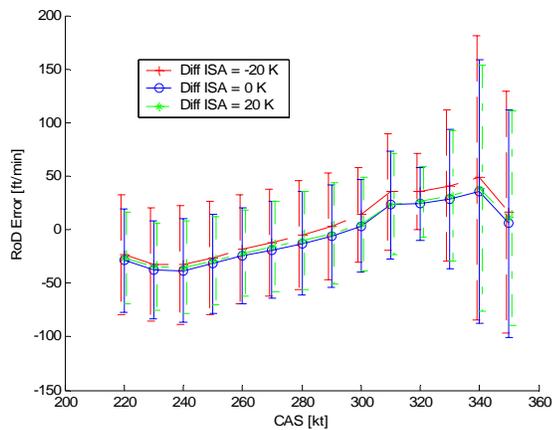


(b)

**Figure 9: Signed Error vs Speed (CAS) for Diff ISA = 0 K and 3 Masses, respectively for the a) Model 1 and b) Model 2 ROC Models**



(a)



(b)

**Figure 10: Signed Error vs Speed (CAS) for Mass = 68 T and 3 Diff ISA values, respectively for the a) Model 1 and b) Model 2 ROD Models**

## Experiment Comments and Limitations

The results presented here aim at illustrating the evaluation methodology and analysis, however, they are not suitable for a quantitative analysis. In the case of the APM representing a kinetic approach, the thrust model was replaced by an inappropriate surrogate as time did not permit the thrust model to be generated from the source data available to both APMs studied. If the kinetic APM had been generated from the common source data, the results would have been valid for quantitative comparison.

In addition when performing a study that compares the performance of different APMs it must be ensured that the reference data used and the meteorological model implemented within the evaluation process are consistent. Special care has been taken to ensure this.

Given that the scope of this initial study was limited to one type of aircraft operating in the clean configuration over typical climb/descent profiles, we have not yet incorporated any “aggregate” metrics. Aggregated metrics, such as those described in reference 6, provide the analyst with “higher level” KPIs, i.e., characterizing one APM for all the phases of flight and for all the aircraft types analyzed. It is clear that an external user of an APM may be more interested in this type of KPI, a topic that will require a broader follow-on study.

## Conclusions and Next Steps

Future ATM concepts, such as trajectory based ATM, rely on the assumption that the fidelity of the trajectory prediction predictions will support the accuracy, look-ahead times, and computational speed required to achieve operational feasibility and the desired benefits. The level of performance required from an aircraft performance modeling approach, suitable to support the TP in meeting these requirements, will be a tradeoff between fidelity, complexity and speed of calculation.

The methodology applied within this study proposes a three steps evaluation of the accuracy of any aircraft performance model.

The first step consists of evaluating the performance of each model of interest over the flight envelope. This study clearly demonstrated that the accuracy of APMs varies significantly over the flight envelope. Conversely, given a required level of performance, this identifies the portion of the envelope for which the APM is acceptable.

The second step of the methodology aims at evaluating the accuracy of any APM at and around the nominal operating conditions. This is a critical check for two reasons: operators often choose to operate at nominal conditions, and manufacturers typically provide actual data essential for validating an APM. In addition, the results from this step may suggest to the user that further investigation is warranted, such as sensitivity analysis.

Finally the third step of the evaluation methodology completes the process by performing a sensitivity analysis of any model performance parameter. This information enables the user to identify input parameters with significant impact on APM performance, and the location of the errors within the flight envelope. For example, the results indicated that the accuracy of the models studied degrades at the edges of the flight envelope. This third step is critical for ensuring that the APM meets the required level of accuracy at the conditions needed.

Together, these steps provide the user with a complete picture including both the accuracy of an APM approach, and an indication of where an APM may need further improvement.

This study also verified that the aircraft performance modeling component of an application can be isolated from the host application independent of whether it is a kinetic or kinematic modeling approach.

Looking forward, it is desirable to extend this work to provide quantitative results and conclusions directly useable by APM clients. Such work will need to include, among others, a representative set of multiple aircraft types; inclusion of additional APMs (including those used for FMS and fast time simulations); and the use of aggregated metrics to synthesize results over many KPIs of interest.

Such a study will provide three important aspects: a thorough analysis of the relative merits of kinetic and kinematic APM approaches; a complete analysis of existing APMs; and reference data to support the APM user in the selection of the APM which suits best its application with respect to tradeoffs in factors such as accuracy, model calculation speed, and the intrinsic complexity of the APM in terms of database size and ease to diagnose, modify, and add models for new aircraft types.

## Acknowledgements

*This work was jointly funded by the EUROCONTROL Agency AP-16 TP activities and NASA Airspace Systems Program, and coordinated between Eurocontrol, NASA, and the FAA.*

## References

- [1] Mondoloni S., S. Swierstra, 2005, Commonality in Disparate Trajectory Predictors for Air Traffic Management Applications, 24<sup>th</sup> DASC, Washington, 2005
- [2] Paglione M., C. Garcia-Avello, S. Swierstra, R. Vivona, 2005, A Collaborative Approach to Trajectory Modeling Validation, 24<sup>th</sup> DASC, Washington, 2005
- [3] Mondoloni S., S. Swierstra, 2005, Assessing Trajectory Prediction Performance - Metrics Definition, 24<sup>th</sup> DASC, Washington, 2005
- [4] Miguel A. Vilaplana, Eduardo Gallo, Francisco A. Navarro, Sip Swierstra, 2005, Towards a formal language for the common description of aircraft intent, 24<sup>th</sup> DASC, Washington, 2005
- [5] One Sky Team "Common TP initiative" [www.eurocontrol.int](http://www.eurocontrol.int)
- [6] CARE (Collaborative ATM research in Eurocontrol) Key performance indicators for aircraft performance models.

## Abbreviations

AP16	Action plan 16: Common TP capability
API	Applications Program Interface
APM	Aircraft Performance Model
CCOM	Eurocontrol-FAA R&D Coordination committee
DST	Decision support tool
ISA	International Standard Atmosphere
KPI	Key Performance Indicator
PAMPA	Platform for Aircraft performance Modeling Performance Analysis
ROC	Rate Of Climb
ROD	Rate Of Descent
CAS	Calibrated air speed in knots
TP	Trajectory predictor
TS	Trajectory synthesizer

## Key words

DST, aircraft performance modeling, trajectory prediction, simulation, Key Performance Indicator, PAMPA, kinetic, Kinematic

## **Biographies**

**C. Garcia Avello**, Mechanical engineer, R&D activities in aircraft performance for TP and ATC simulation and in DST developments.

**H. Idris**, PhD. Human Factors and Automation, M.S. Operations Research, activities in cost/benefit assessment and design of DSTs.

**R. A. Vivona**, M.S. Aeronautics and Astronautics, activities in design and development of DSTs.

**S. Green**, activities in NASA Trajectory Synthesizer and CTAS DSTs suite development

24<sup>th</sup> Digital Avionics Systems Conference

*October 30, 2005*