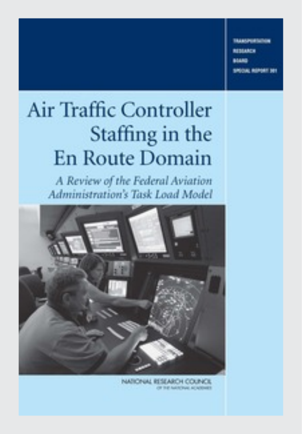
TRANSPORTATION RESEARCH BOARD SPECIAL REPORT 301

Air Traffic Controller Staffing in the En Route Domain

A Review of the Federal Aviation Administration's Task Load Model

NATIONAL RESEARCH COUNCIL OF THE NATIONAL ACADEMIES

TRB Special Report 301: Traffic Controller Staffing in the En Route Domain: A Review of the Federal Aviation Administration's Task Load Model (2010)

DETAILS

84 pages | 8.5 x 11 | PAPERBACK
ISBN 978-0-309-16069-8 | DOI 10.17226/13022

CONTRIBUTORS

Committee for a Review of the En Route Air Traffic Control Complexity and Workload Model, National Research Council

SUGGESTED CITATION

Transportation Research Board. 2010. *TRB Special Report 301: Traffic Controller Staffing in the En Route Domain: A Review of the Federal Aviation Administration's Task Load Model*. Washington, DC: The National Academies Press. https://doi.org/10.17226/13022.

TRANSPORTATION RESEARCH BOARD

SPECIAL REPORT 301

Air Traffic Controller Staffing in the En Route Domain

A Review of the Federal Aviation Administration's Task Load Model

Committee for a Review of the En Route
Air Traffic Control Complexity and Workload Model

NATIONAL RESEARCH COUNCIL
OF THE NATIONAL ACADEMIES

Transportation Research Board
Washington, D.C.
2010
www.TRB.org

Transportation Research Board Special Report 301

Subscriber Category
Aviation

Transportation Research Board publications are available by ordering individual publications directly from the TRB Business Office, through the Internet at www.TRB.org or nationalacademies.org/trb, or by annual subscription through organizational or individual affiliation with TRB. Affiliates and library subscribers are eligible for substantial discounts. For further information, contact the Transportation Research Board Business Office, 500 Fifth Street, NW, Washington, DC 20001 (telephone 202-334-3213; fax 202-334-2519; or e-mail TRBsales@nas.edu).

Printed in the United States of America.

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competencies and with regard for appropriate balance.

This report has been reviewed by a group other than the authors according to the procedures approved by a Report Review Committee consisting of members of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. This report was sponsored by the Federal Aviation Administration of the U.S. Department of Transportation.

Typesetting by Circle Graphics.
Cover photo courtesy of the National Air Traffic Controllers Association.

Library of Congress Cataloging-in-Publication Data

National Research Council (U.S.). Committee for a Review of the En Route Air Traffic Control Complexity and Workload Model.
Air traffic controller staffing in the en route domain : a review of the Federal Aviation Administration's task load model / Committee for a Review of the En Route Air Traffic Control Complexity and Workload Model, Division on Behavioral and Social Sciences and Education, National Research Council of the National Academies.
p. cm.
Includes bibliographical references.
1. Air traffic capacity—United States—Mathematical models. 2. United States. Federal Aviation Administration.—Officials and employees—Workload—Mathematical models. 3. Manpower planning—United States—Statistical methods. I. Title.
TL725.3.T7N3685 2010
387.7'404260683—dc22

2010042255

ISBN 978-0-309-16069-8

THE NATIONAL ACADEMIES

Advisers to the Nation on Science, Engineering, and Medicine

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. On the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Ralph J. Cicerone is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. Charles M. Vest is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, on its own initiative, to identify issues of medical care, research, and education. Dr. Harvey V. Fineberg is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both the Academies and the Institute of Medicine. Dr. Ralph J. Cicerone and Dr. Charles M. Vest are chair and vice chair, respectively, of the National Research Council.

The **Transportation Research Board** is one of six major divisions of the National Research Council. The mission of the Transportation Research Board is to provide leadership in transportation innovation and progress through research and information exchange, conducted within a setting that is objective, interdisciplinary, and multimodal. The Board's varied activities annually engage about 7,000 engineers, scientists, and other transportation researchers and practitioners from the public and private sectors and academia, all of whom contribute their expertise in the public interest. The program is supported by state transportation departments, federal agencies including the component administrations of the U.S. Department of Transportation, and other organizations and individuals interested in the development of transportation. **www.TRB.org**

www.national-academies.org

Committee for a Review of the En Route Air Traffic Control Complexity and Workload Model

R. John Hansman, Jr., Massachusetts Institute of Technology, Cambridge, *Chair*
Monica S. Alcabin, Boeing Commercial Airplanes, Seattle, Washington
Michael O. Ball, University of Maryland, College Park
Mary L. Cummings, Massachusetts Institute of Technology, Cambridge
William J. Dunlay, Jacobs Consultancy, Burlingame, California
Antonio L. Elias, Orbital Sciences Corporation, Dulles, Virginia
John J. Fearnsides, MJF Strategies, McLean, Virginia
J. Victor Lebacqz, National Aeronautics and Space Administration (retired), Aptos, California
Michael J. Powderly, Airspace Solutions, Marietta, Georgia
Philip J. Smith, Ohio State University, Columbus
Antonio A. Trani, Virginia Polytechnic Institute and State University, Blacksburg
Roger Wall, Federal Express Corporation (retired), Kent, Washington
Greg L. Zacharias, Charles River Analytics, Cambridge, Massachusetts

National Research Council Staff
Thomas R. Menzies, Jr., Study Director, Transportation Research Board
Susan Van Hemel, Senior Program Officer (retired), Division on Behavioral and Social Sciences and Education

Preface

For the past decade, the Federal Aviation Administration (FAA) has sponsored the development of modeling capabilities for the analysis of en route sector complexity, controller workload, and sector capacity. These capabilities have been developed by the agency's federally funded research and development center, MITRE Corporation's Center for Advanced Aviation System Development (CAASD). Upon FAA's request, the Transportation Research Board (TRB), in conjunction with the Division on Behavioral and Social Sciences and Education (DBASSE), agreed to provide an expert review of the model for use in informing the agency's workforce planning. The details of the request are provided in the study statement of task contained in Box 1-2 (page 14).

To conduct the independent review, TRB and DBASSE assembled a committee of experts in human factors, modeling, and air traffic control research, planning, operations, and management. R. John Hansman, Jr., Professor of Aeronautics and Astronautics at the Massachusetts Institute of Technology, chaired the committee, whose 13 members served in the public interest without compensation. Over the course of seven months, the committee met three times. During its first meeting in December 2009, the committee received overview briefings from FAA and CAASD about the model and its current and potential uses. During the second meeting, in March 2010, the committee visited the Washington Air Route Traffic Control Center (ARTCC) in Leesburg, Virginia, and received more detailed briefings from FAA and CAASD on the model and its use to inform workforce planning. The committee's final meeting, in June 2010, consisted mainly of committee deliberations to produce this report.

The committee thanks all of the individuals from FAA and MITRE who made presentations during the meetings and otherwise assisted the

committee during the course of the study, especially Dan Williams, FAA and Diane E. Boone, MITRE Corporation. The committee also wishes to thank Larry Bogner and Bill Holtzman from the Washington ARTCC for assisting in making the arrangements for and hosting the committee's February visit. Thomas R. Menzies, Jr., managed the study and assisted the committee in drafting the final report under the supervision of Stephen R. Godwin, Director of Studies and Special Programs, TRB and Barbara Wanchisen, Interim Director, Committee on Human–Systems Integration, DBASSE.

Suzanne Schneider, Associate Executive Director of TRB, managed the report review process. The report was edited by Naomi Kassabian; Jennifer J. Weeks prepared the manuscript for web posting; and Juanita L. Green managed the design and production, under the supervision of Javy Awan, Director of Publications, TRB. Special appreciation is expressed to Amelia Mathis for assistance with meeting arrangements and communications with the committee.

The report has been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the National Research Council's Report Review Committee. The purpose of this independent review is to provide candid and critical comments that will assist the institution in making its published report as sound as possible and to ensure that the report meets institutional standards for objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process.

Thanks go to the following individuals for their review of the report: John B. Hayhurst, Boeing Company, Kirkland, Washington; Brian Hilburn, Northrop Grumman Corporation, Atlantic City, New Jersey; William C. Howell, Arizona State University, Mesa, and Rice University, Houston, Texas; Bill F. Jeffers, Newnan, Georgia; Waldemar Karwowski, University of Central Florida, Orlando; Amy R. Pritchett, Georgia Institute of Technology, Atlanta; Christopher D. Wickens, University of Illinois (Emeritus), Urbana–Champaign, and Alion Science and Technology, Boulder, Colorado.

Although these seven reviewers provided many constructive comments and suggestions, they were not asked to endorse the committee's findings

or recommendations, nor did they see the final draft of the report before its release. The review was overseen by Adib K. Kanafani, University of California, Berkeley, and C. Michael Walton, University of Texas, Austin. Appointed by the National Research Council, they were responsible for making certain that an independent examination of this report was carried out in accordance with institutional procedures and that all review comments were carefully considered. Responsibility for the final content of this report rests entirely with the authoring committee and institution.

Contents

Summary

The Federal Aviation Administration (FAA) is seeking to improve a mathematical model that estimates the time spent by controllers performing tasks in working the air traffic in each of the more than 750 sectors of the nation's en route airspace. FAA has been using the model's estimates of task time expenditure, or "task load," to assess the number of controllers required to work each sector's traffic. The model simulates the traffic activity experienced in each sector and then associates task times with this activity to compute task load. While the task load values do not portray the total workload on controllers—since workload is driven by other factors such as stress, fatigue, and expertise—they can provide a consistent and objective source of information for controller staffing. It is for this reason that an earlier TRB report[1] urged FAA to pursue task-based modeling for workforce planning.

FAA's task load model is currently being used as one of several inputs in the agency's annual controller workforce plan (CWP). The modeled task loads are used to estimate the number of controllers required in position in each sector to perform the traffic-driven tasks, which FAA refers to as "positions to traffic," or PTT. When a sector is open to traffic, it has at least one controller in position, the lead controller. Depending on traffic demand and other factors, the lead controller may be accompanied by an associate controller. Thus, PTT values are usually 1 or 2. When traffic is exceptionally heavy a third controller may be added to the team, although this setup is seldom a planned staffing configuration.

Having used the PTT estimates from the model to inform the CWP for the past few years, FAA sought an independent review of the modeling

---

[1] TRB. 1997. *Special Report 250: Air Traffic Control Facilities—Improving Methods to Determine Staffing Requirements.* National Research Council, Washington, D.C.

process to assess its utility and validity going forward. Specifically, FAA asked the National Academies to convene an expert committee to examine and offer advice where appropriate for improving (*a*) the overall technical approach of task-based modeling, (*b*) input data and processes used for modeling traffic activity, (*c*) tasks and methods used to assign task times, and (*d*) means for validating model assumptions, parameters, and output. In addressing this charge, the committee was asked to be cognizant of the "overall tradeoffs made due to data availability" and to consider the "adaptability of the approach to reflect changes in the tasks of controllers as their roles evolve over time." Key study findings with respect to each of these elements of the study charge are given next, followed by recommendations.

KEY FINDINGS

Task-Based Approach

The results of task-based modeling can be a valuable source of objective information for workforce planning, and FAA's current model is a marked improvement over previous models. Earlier models measured the number of aircraft flying through a sector without accounting for the variability in the complexity of this traffic, and thus the variability in controller tasks and time demands. For example, aircraft changing headings and altitude create more traffic complexity than aircraft cruising straight through a sector. FAA's current model accounts for traffic complexity by simulating the traffic flows and patterns experienced in the en route sectors and relating them to the time-varying tasks that controllers perform. The basic model structure, in which traffic activity is simulated and controller tasks and task times are associated with traffic, represents a logical approach to estimating task load. The methods used to derive model parameters and values and to convert the modeled task load into PTT are the subject of most of the criticism and advice in this report.

Simulations of Traffic Activity

By using available traffic operations and flight-planning data, flight plans, and trajectory modeling, the task load model simulates past sector traffic flows and patterns. The traffic activity is modeled in sufficient depth and resolution to enable reasonable approximations of traffic complexity and associated controller tasks. Because the simulated traffic can be checked against records of actual traffic activity, there is ample opportunity to use empirical data to validate output accuracy and guide the development and calibration of the traffic modeling methods and parameters. Model developers have been taking advantage of these opportunities to make periodic improvements to the traffic modeling process.

Task Coverage

The task load model does not analyze all of the tasks performed by controllers but only certain ones performed by the lead controller in communicating with aircraft, monitoring flights on the radar screen, and communicating with controllers from other sectors and centers. The modeling of these lead controller tasks is essential for analyzing the traffic throughput capacity of individual sectors, which was the original purpose of the model. Task coverage for this purpose appears to be adequate. Yet in order to know when the demands of traffic necessitate more than one controller—that is, in order to estimate PTT—it is necessary to know the *total* task load on controllers, including the task load on the associate controller. By omitting all of the tasks performed by the associate controller, the model's task load output alone is not adequate for estimating PTT.

FAA and model developers have sought to compensate for this significant gap in task coverage by employing various processes that infer the missing task load to enable conversions of model output into PTT. All of the PTT conversion methods used, including the current one using fuzzy logic modeling, exhibit the same fundamental flaw—they imply an ability to estimate total task load without ever identifying the unmodeled tasks, much less measuring the time it takes to perform them. The PTT conversions using fuzzy logic modeling rely on experts to assign complexity weightings to the unidentified and unmodeled tasks. These weightings are not validated, nor can they be in the absence of any empirical data on task performance. On the whole, the use of fuzzy logic modeling to infer task load adds little more than spurious precision to the

PTT estimates while complicating and reducing the transparency of the modeling process.

Derivation of Task Times

Since task load output is the sum of the time spent by controllers performing tasks, the task completion times are critical model parameters. Many of the task times in the model are derived from a separate modeling process known as Goals, Operators, Methods, and Selection Rules (GOMS). The GOMS-derived times are based largely on expert judgment and are only loosely validated against a limited set of task performance data obtained from human-in-the-loop (HITL) experiments conducted for other purposes. GOMS modeling is typically used where conditions do not permit the observation and analysis of task performance in operational or experimental settings. The committee believes that such conditions do not exist in the air traffic control domain to the extent that warrants such heavy reliance on GOMS. The use of GOMS to derive many task times, coupled with reliance on expert judgment for validating these modeled times and for estimating many others, raises serious questions about the accuracy of the model's task load values.

Validation

Modeled traffic activity can be checked for accuracy through comparisons with records of actual traffic. In contrast, validating PTT estimates is more challenging since there is no external measure of staffing requirements against which the accuracy of the estimates can be judged. Analyzing staffing records is of limited value since the main purpose of PTT modeling is to find out when staffing levels can be better aligned with traffic demand. A main means by which model developers have sought to assess PTT estimates is by presenting them to groups of experts, often consisting of individuals who manage and staff the en route centers. Yet such checks can suffer from the same shortcoming that limits the value of comparisons with staffing records—the potential for bias toward existing staffing practice.

Because PTT estimates cannot be assessed through direct observation, all of the model's key assumptions, processes, and parameters must be well justified and validated. A lack of data on task performance precludes validation of the task times constructed from GOMS and the task complexity weightings used in the fuzzy logic conversion method. The deficiencies of these two modeling processes go well beyond parameter validation, as explained earlier. Yet the lack of empirical data on task performance has hindered validation throughout the modeling process, from assessing key assumptions about tasks being performed sequentially and at a fixed pace to characterizing the tasks handled by the associate controller.

Data Availability and Model Adaptability

In the study charge, FAA asked the committee to be cognizant of trade-offs that must be made because of limited data availability, which presumably refers to the cost and complications of obtaining task performance data. FAA also asked for advice on the model's adaptability to reflect changes in controller roles and tasks over time.

Many of the findings cited earlier point to a need for a firmer empirical basis both for evaluating the structure of the model and for estimating the values of the parameters used in it. By and large, the model was developed and has been evaluated with heavy reliance on the insights and opinions obtained from subject matter experts and facility personnel. More objective and quantitative task performance data are clearly needed, not only for developing the model parameters and evaluating the task load output but also for including more controller tasks for the modeling of PTT. The committee recognizes that gathering such data from operational and experimental settings will require more resources and access to controllers, which may present budget and labor relations issues. Although such cost implications were not examined in this study, it must be pointed out that there is a cost in model credibility from not obtaining such data. This cost is manifested in many ways throughout the current model, from the added opaqueness caused by fuzzy logic modeling to the excessive reliance on expert opinion and judgment for model development and validation.

Whether FAA is committed to taking this data-gathering step will presumably depend on its assessment of the cost trade-offs and its plans for using the model for a long time and for other possible purposes. Not knowing these plans, the committee nevertheless believes that FAA would not have asked for this review absent a strong interest in improving its modeling capabilities. It is in this context that the committee wishes to express its strong view that the current model falls short in its ability to estimate PTT and that continuing to iterate on it in the same manner as in the past while not incorporating more complete and representative task performance data will not overcome the deficiency.

Looking farther out, the durability of the task load model for PTT analysis and for other possible applications, such as to inform traffic flow planning, will depend not only on the successful gathering and use of task performance data but also on the nature and pace of change in the air traffic control enterprise. Developments anticipated for the planned Next Generation Air Transportation System (NextGen), such as increased automation and many more decision-support tools, could substantially alter controller roles and responsibilities in ways that are highly relevant to the modeling of PTT. Without more knowledge about the nature and timing of these NextGen changes, it is not possible to predict how the model will hold up structurally, much less how changes in traffic data, task coverage, and task times might make it more adaptable.

RECOMMENDATIONS

In commencing its review, the committee expected to find—but did not—strong documentation explaining the logic and structure of the model and evidence of its having been the subject of statistical tests and other scientific methods for establishing and validating model parameters, assumptions, and output. More rigorous documentation and peer review during earlier stages of model development would likely have exposed many of the problems identified in this report, providing earlier opportunities to avoid or correct them. Nevertheless, as a preface to offering advice on ways to improve the modeling process going forward, it is important to restate the finding that the current model framework, despite the data shortcomings, represents a major improvement over past modeling methods to inform workforce planning. In the following recommendations it is presumed that FAA will elect to retain the core model and invest meaningfully in its improvement.

Observe and Measure Controller Task Performance

Through more systematic and carefully designed observation and analysis of controller performance, model developers should gain a better understanding of the tasks that controllers perform in working en route traffic, how they perform them, and the time required to do so. The gathering and analysis of data on controllers working alone and interacting in teams, whether through field observations or HITL experiments, should be the primary method to identify and elicit information on controller tasks.

Model All Controller Tasks

Modeling all tasks that contribute significantly to total controller task load is fundamental for estimating PTT. FAA should use the information gained from observing, measuring, and analyzing controller task performance to quantify the task load associated with the services provided by both the lead and associate controllers. The modeling of all controller tasks will eliminate the need to infer task load to derive estimates of PTT. Using a single model for estimating task load rather than separate ones for each controller is the preferred approach, since it will facilitate both PTT conversion and model validation.

Validate Model Elements

Task performance data should be used also to assess the validity and impact of all key modeling processes, relationships, and assumptions. Because it is not possible to validate PTT estimates against actual staffing levels, ensuring that the model elements are well justified and viewed as credible is vitally important. Examples of modeling assumptions that would seem to warrant early attention are those that concern task performance by the controllers when working alone and in teams, whether tasks are performed sequentially or concurrently, and how total task load affects the pace of task performance.

1

Study Charge and Background

The Federal Aviation Administration (FAA) is responsible for designing and operating the national airspace system (NAS). The NAS consists of terminal and en route airspace and a complex network of navigation, surveillance, and communications systems used to guide and control aircraft traffic within this airspace and on the ground at airports. FAA's workforce of about 15,700 air traffic controllers, working in more than 300 facilities across the country, directs the more than 50,000 aircraft operations that occur each day in the NAS. A key component of this workforce is the approximately 5,000 controllers who work in the 20 en route facilities that separate and direct aircraft operating in the air routes not assigned to towers, terminal facilities, and the military.

FAA is expected to provide safe, orderly, and efficient air traffic control services while meeting resource and budgetary constraints. To ensure efficient provision of these services, the agency needs good information to support decisions on the hiring, training, and deployment of controllers across its many air traffic control facilities. For many years, FAA's decisions on controller staffing have been informed by an array of data sources and methods that have at times come under scrutiny. In particular, FAA has been asked to explain and justify the variability in controller staffing levels found across its facilities. In 1995, Congress sought an independent assessment of these methods for use in planning controller staffing in the individual control facilities. A special committee of the Transportation Research Board (TRB) was formed to conduct that earlier study, whose recommendations prompted FAA to develop further the model that is the subject of this follow-on study.[1]

---

[1] TRB. 1997. *Special Report 250: Air Traffic Control Facilities—Improving Methods to Determine Staffing Requirements.* National Research Council, Washington, D.C.

During the 1990s, FAA's staffing models were used mainly to develop national workforce targets for general budgetary decisions. Regional offices generated their own estimates of the number of controllers required per facility. FAA recognized that its national staffing models were too generalized and imprecise to predict staffing needs at each facility. Its models were built on highly aggregated data derived from a small number of sampled sectors and control centers. While considered adequate for generating systemwide estimates of workforce needs, the models lacked the detail needed to inform facility-level planning.

Acknowledging that FAA's staffing models were not designed to inform facility-level staffing plans, the 1997 TRB study committee nevertheless questioned the models' use of simple counts of the number of flights traversing a block of airspace as the main indicator of traffic demand on controlling capacity. FAA found that these volume-based measures generated staffing values that were much higher than facility managers believed were reasonable on the basis of demonstrated experience. In particular, the measures did not take into account how the complexity of traffic activity, in addition to its quantity, affects controlling capacity. The committee observed, for instance, that the models employed data on various controller actions that could be readily observed and timed, such as scanning a radar screen, typing on a keyboard, and radioing a pilot. These identified controller actions, however, were not linked to the specific tasks performed by controllers working different types of aircraft activity; for instance, a flight entering the airspace requires the controller to accept the hand-off from another controller, whereas a flight that is changing heading or altitude requires the controller to perform various checks and clearances. Establishing these connections between traffic activity and the actions that controllers must take in response was viewed as critical to developing more accurate estimates of the time demands on controllers.

In particular, the TRB committee recommended that FAA try to quantify the time controllers spend performing specific tasks as they monitor, inform, and direct the actions of aircraft. By modeling traffic activity and coupling this activity with time-varying controller tasks that must be performed in response, FAA could then estimate the total time spent executing tasks, referred to here as "task load." The TRB committee observed that such task load estimates could be used for assessing the number of controllers needed to work given levels and patterns of traffic demand.

In the early 2000s, MITRE Corporation's Center for Advanced Aviation System Development (CAASD), a federally funded research and development center, happened to be developing a task-based model for analyzing the traffic capacity of en route sectors. Understanding controller time demands is important for assessing traffic capacity, since the maximum number of aircraft that can safely traverse a sector in a given time period is constrained by controlling capacity, or the total time available to controllers to work the traffic. The CAASD model used historical flight operations and planning data to simulate the traffic levels and patterns of activity experienced in individual sectors. Sectors experiencing more complex traffic patterns, which require more time-intensive controller tasks per flight, were assumed to reach their maximum traffic capacity with fewer total flights than those sectors experiencing more straightforward traffic patterns. In this way, the model could estimate the real, controller-constrained traffic capacity of each en route sector at different points in time.

In light of the recommendations of the TRB study, FAA asked CAASD to investigate whether its capacity-oriented task load model could also be used to estimate the number of controllers that were needed in position to work each sector's experienced traffic activity. By generating such retrospective estimates of "positions to traffic" (PTT), FAA believed the model could help inform staffing plans for each en route center.

How the model was adapted to meeting these needs is explained and reviewed in this report. Generally satisfied with the early results of the model's adaptations for PTT estimation, FAA has been using the modeling results since 2007 to inform the en route portion of its annual controller workforce plan (CWP). The CWP projects the total controller workforce needed for the next decade.[2] It also provides annual staffing ranges for each traffic control facility. These facility-level ranges are developed through a multistep process, as explained in Box 1-1, that considers past facility productivity, the performance of other peer facilities,

---

[2] Air Traffic Control Workforce Plan. http://www.faa.gov/air_traffic/publications/controller_staffing.

BOX 1-1

FAA Process for Projecting Facility Staffing Ranges (from 2010 Controller Workforce Plan)

The FAA uses the following four information sources to estimate facility staffing ranges:

1. **Past productivity:** controller head count required to match the historical best productivity for the facility. Productivity is defined as operations per controller. Facility productivity is calculated by using operations and controller data from 1999 to 2009. If any annual point falls outside ±5 percent of the 1999 to 2009 average, it is discarded. From the remaining data points, the year with the highest productivity is used as the benchmark.
2. **Peers:** head count required to match peer group productivity. Comparable facilities are grouped by type and level and their corresponding productivity is calculated. If the data for the facility being considered are consistently above or below that of the peer group, the peer group data are not used in the overall average and analysis.
3. **Service unit input:** consultations with field managers.
4. **Staffing standards:** mathematical models used to relate controller task load and air traffic activity.

The average of these data is calculated, rounded to the nearest whole number, multiplied by 0.9 and 1.1, and then rounded again to determine the high and low points in the facility staffing range. Exceptional situations, or outliers, are removed from the averages (for example, if a change in the type or level of a facility occurred over the period of evaluation). By analyzing the remaining data points, staffing ranges are generated for each facility. The agency's hiring and staffing plans consider all of these inputs as well as other considerations such as time on position and overtime. All of these data points are reviewed collectively and adjustments are made to facility staffing plans during the year as necessary.

consultations with field managers, and a set of quantitative models that are referred to collectively as “staffing standards.”

The CAASD model is one of the three quantitative models that make up FAA’s staffing standards. The two others forecast future traffic levels and calculate staffing needs after making allowances for vacation, training, and work rules (e.g., required work breaks). Although FAA has been using the CAASD model to inform CWP projections for several years, the agency would like to continue doing so with even higher confidence. The agency therefore asked TRB for this follow-on study of the model and its utility for estimating PTT.

As detailed in Box 1-2, the study’s charge calls for a review of the model’s technical approach, to include the assumptions, methods, and data used for profiling traffic activity, triggering tasks, assigning task execution times, and calibrating and evaluating model parameters and results. The charge also asks for advice on how the model can be adapted to an evolving air traffic control system in a next-generation air transportation environment—one in which air traffic controller technologies and procedures may change significantly to create new controller roles, tasks, and performance capabilities.

REPORT ORGANIZATION

As background for the study, the next section describes the en route domain and role of the controllers in managing the traffic in each en route sector. Chapter 2 provides an overview of the modeling effort, including the main assumptions and basic structure of the CAASD task load model and the methods used to convert its output into estimates of PTT. Chapter 3 describes in more detail the individual elements of the model, including the modeled tasks, events that trigger them, and methods used to derive task execution times. It concludes with the committee’s assessment of these model elements and efforts to check their accuracy and that of the modeled task load. Chapter 4 provides a detailed description of the methods used to convert task load estimates into PTT. It concludes with the committee’s assessment of these conversion methods and checks on their validity. On the basis of these assessments, Chapter 5 summarizes the main study findings and conclusions relevant to the

BOX 1-2

Study Statement of Task

The study will provide an expert review of methodologies and modeling capabilities for post facto analysis of en route sector capacity and positions to traffic (PTT) developed by the Federal Aviation Administration's (FAA) federally funded research and development center, MITRE Corporation's Center for Advanced Aviation System Development (CAASD). Specifically, the study will offer suggestions where applicable for strengthening the following areas:

1. Technical approach
    - Task-based approach for post facto analysis of en route sector capacity and PTT,
    - Adaptability of the approach to reflect changes in the tasks of controllers as their roles evolve over time (e.g., Next Gen–related changes),
    - Tasks and methods for developing task times.
2. Input data
    - Task triggers and overall trade-offs made because of limited data availability
3. Model output
    - Processes for evaluating and calibrating the models given the availability of objective real-world data on capacity and PTT.

statement of task. It concludes with the committee's recommendations for improving the modeling effort.

BACKGROUND ON EN ROUTE AIR TRAFFIC CONTROL

From the towers of approximately 450 airports, local air traffic controllers direct takeoff and landing clearances as well as surface movements between gates, taxiways, and runways. These controllers manage traffic within approximately 5 miles of the airport up to an altitude of about 3,000 feet. In larger metropolitan areas with multiple busy airports, controllers at terminal radar approach control (TRACON) facilities sequence departing aircraft from takeoff to transition altitude and arriving aircraft during descent.[3]

Once at higher altitude, transiting aircraft are monitored with radar and directed by controllers located in the 20 air route traffic control centers (ARTCCs) located across the continental United States. Each center is responsible for traffic operating in the airspace over a specific region of the country, and some also control aircraft operating over the ocean. The airspace managed by each of these centers usually covers portions of several states, typically covering 60,000 to 350,000 square miles (see Figure 1-1).

En route controllers strive to maintain a safe separation between transiting aircraft as they accept traffic from and pass traffic to other controllers at centers or terminal facilities. They communicate over voice and data channels with pilots and other controllers. Through radar surveillance and radio communications, they also provide pilots with traffic and weather advisory information. En route controllers also direct approaches to airports that do not have operational towers.

Each en route center employs between 100 and 350 controllers, with most employing 200 to 275. A controller is certified to direct traffic only within defined areas of specialization. Most centers consist of four to eight areas. Each area is typically responsible for a slice of airspace that is divided into five to nine sectors of low, high, and ultrahigh airspace (see Figure 1-2). The sectors vary in size from 500 to more than 30,000 cubic miles. There are more than 750 sectors of airspace over the continental United States.

---

[3] There are multiple air traffic control domains: tower, terminal, en route, ocean, and traffic flow management. A flight will typically be controlled by a tower on departure from an airport. Shortly after departure, the flight will be handed off to a terminal radar control facility for control in the airspace near the airport. As a flight reaches a higher altitude, it will be handed off to an en route facility for control until it nears its arrival airport. A flight may be controlled by multiple en route facilities until upon its arrival the process is reversed with an en route facility handing off control to a terminal facility as the flight descends. The tower at the arrival airport will then take control of the flight on its final approach to the airport. If a flight's path takes it over airspace far from the continental United States, it will also be under the direction of controllers responsible for ocean operations. Traffic flow management has the responsibility of coordinating operations across all the air traffic control domains.

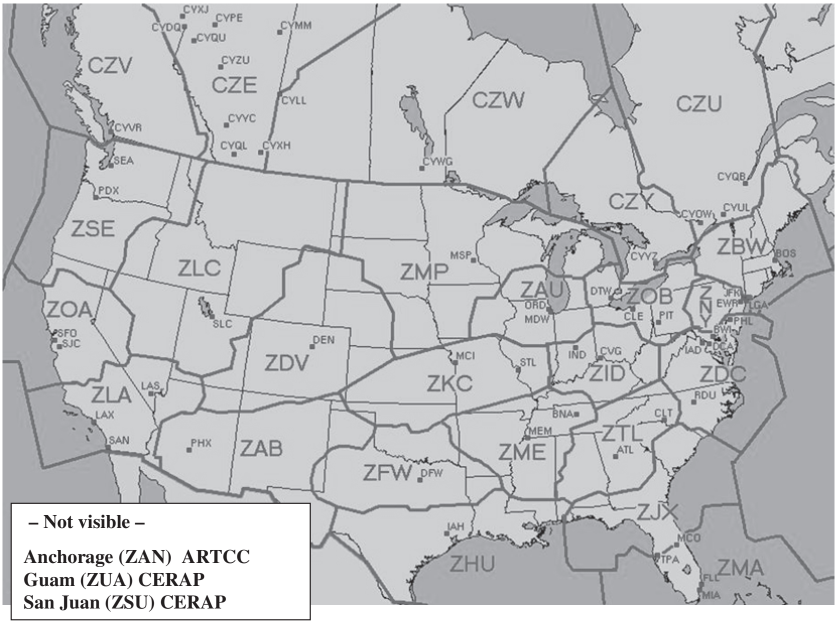

FIGURE 1-1 **Boundaries of ARTCCs in the continental United States and parts of Canada.** (CERAP = central radar approach.)


(a) (b)

FIGURE 1-2 **En route sectors over southeastern United States: (*a*) low altitude and (*b*) high altitude.** (Note: Low sectors are surface to 23,000 feet. High sectors are 24,000 to 33,000 feet. Very high—sometimes called superhigh—sectors are 35,000 feet and higher.)

Controllers certified to direct traffic in an area can work traffic in any one of its sectors. If traffic demand decreases—such as during nighttime hours—a controller may be assigned to handle traffic in contiguous sectors that are combined or the entire area may be combined into one large sector. As traffic demand increases, the sectors will be uncombined and additional controllers may be added to each sector. Repeatable traffic patterns due to scheduled commercial flights aid in the planning of such sector assignments on a daily and seasonal basis.

Each sector is typically positioned with one or two controllers. A radar controller, or "R-side" controller, is the lead, responsible for radio communications with aircraft, monitoring the radar screen to maintain safe separation, and communicating with other controllers. All open sectors are staffed with one R-side controller. When two controllers work a sector, the second is an associate controller, known as a data, or D-side, controller. The D-side controller typically receives flight-plan information and helps plan and organize the flow of traffic within the sector. In the absence of a D-side controller, the R-side controller must handle these D-side responsibilities along with R-side responsibilities. During exceptionally busy periods, a third controller (T-side) may be assigned to the team, although three-member teams are not typically planned for.

2

Model Overview

The key assumptions and basic structure of the task load model are outlined in this chapter as well as methods used to convert its task load output into estimates of positions to traffic (PTT). The chapter concludes with committee observations about the model's assumptions and structure and about the methods employed for PTT conversion. These observations are explored in more depth in the following chapters.

MODELING R-SIDE TASKS ONLY

Aircraft typically pass through multiple en route sectors during a flight. The more aircraft transiting the sector, the more work is created for controllers, thereby consuming controlling capacity. The volume of traffic alone, however, is not the only factor demanding controller time. The complexity of the traffic activity along with the associated controller procedures and technologies are important factors. The amount of time from when an aircraft enters a sector to when it exits a sector is typically less than 15 minutes. During this time, some aircraft may cruise directly through the sector, requiring the controller to monitor the flight to ensure safe vertical and horizontal separation from other aircraft. Other activity may require additional actions by the controller, such as performing clearances for aircraft transitioning to lower or higher altitudes and adjusting headings in response to weather, traffic congestion, or crossing traffic. Thus, both the volume and the complexity of the aircraft affect the time demands on controllers.

As explained in Chapter 1, an en route sector may be staffed by a lone R-side (lead) controller or an R-side and a D-side (associate) controller working as a team. When traffic is very heavy, a third controller (T-side) may be in position as well, although this setup is not a normal (planned-for) configuration. The lead controller is in charge of communicating with pilots, monitoring the radar screen to maintain safe separation, coordinating with other controllers, and other services only provided by the R-side controller. When working alone, the lead controller also has the responsibility to receive and process flight-plan information and to plan and organize the flow of traffic within the sector, which are considered D-side services. When an associate controller is present, the D-side services are no longer the responsibility of the lead controller. The addition of the second controller therefore frees up more time for the lead controller to focus attention on R-side services. This division of responsibilities allows for more traffic to be handled by a two-controller team than by a single controller.

The Center for Advanced Aviation System Development (CAASD) task load model was originally developed for the purpose of assessing the maximum number of flights that can safely traverse a sector during a time period. For the reasons explained earlier, a two-controller team can handle more traffic than a single controller because the lead controller can devote all of his or her controlling time to the R-side tasks that accompany all flights. Thus for assessing the maximum throughput capacity of a sector, it is necessary to assume that two controllers are in position—one handling exclusively R-side tasks and the other handling exclusively D-side tasks. Given this assumption, it is not necessary to model the D-side task load to assess sector traffic capacity. Traffic capacity is simply a function of the controlling time available to the lead controller to perform more R-side tasks. Once the lead controller's time is fully occupied with R-side tasks, the sector will have reached its maximum traffic capacity.

The assumption that two controllers are in position and that traffic capacity is solely a function of the controlling time available to the lead controller caused CAASD to structure the task load model so that it only estimates R-side task load. This modeling limitation, as will become evident later, has important implications for the model's ability to predict PTT.

OVERVIEW OF MODEL STRUCTURE

Figure 2-1 shows the basic structure of the task load model. Box 1 in the figure lists eight of the nine major R-side tasks in the model. To determine when these tasks must be performed—or when they are triggered—the

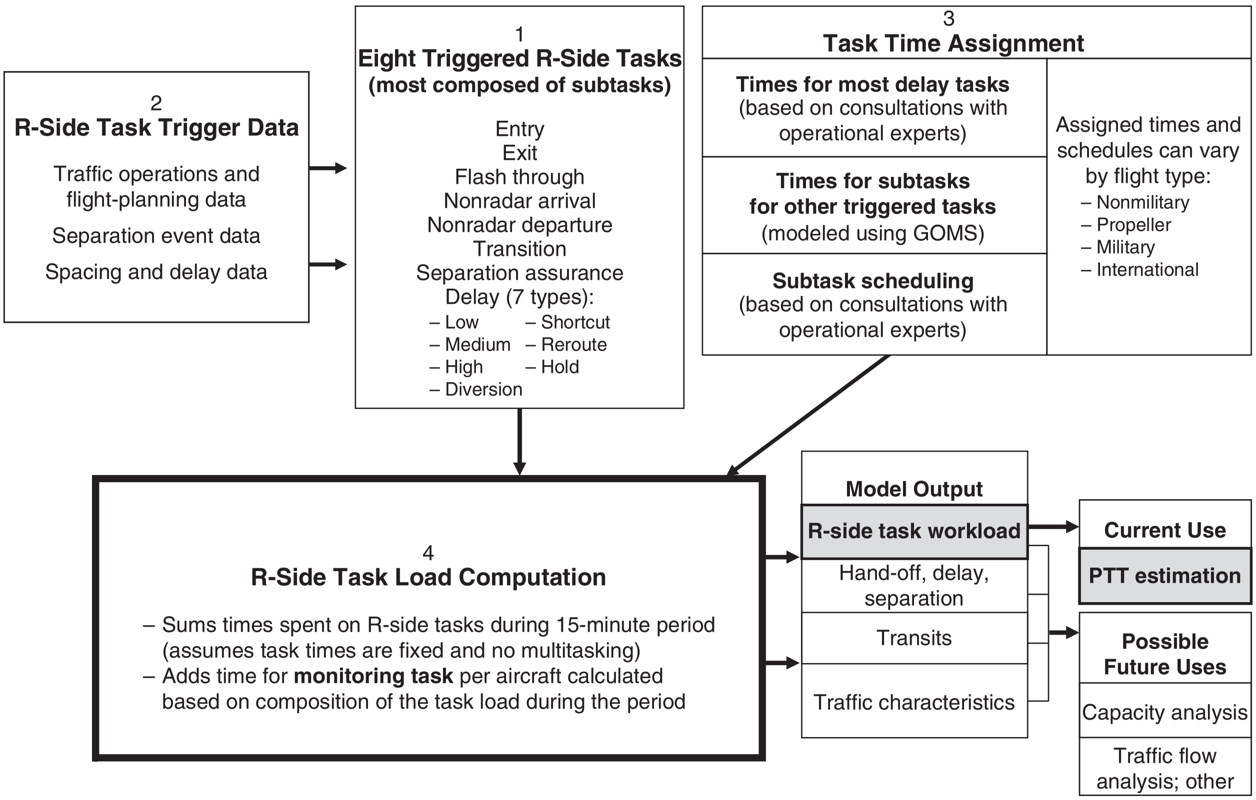

FIGURE 2-1 General structure of task load model.

model requires information on the volume and type of traffic in a sector. As shown in Box 2, the model relies on traffic operations and flight-planning data to simulate—or re-create—the traffic activity that occurred in the sector during past time periods. The traffic operations and flight-planning data, for instance, are used to indicate when a flight entered a sector, an event that triggers an entry task.

When a task is triggered, the model associates it with it a series of actions—or subtasks—that the lead controller is presumed to have performed, such as identifying the aircraft, establishing a clearance plan, and accepting the hand-off. For most of the triggered tasks, the task execution times are derived from a modeling process using Goals, Operators, Methods, and Selection Rules (GOMS) procedures (Box 3 in Figure 2-1). By using GOMS, each subtask is divided into its most basic operators, such as entering a keystroke or scanning a radar screen. Each operator is assigned an estimated execution time. The operator times are summed to calculate the total time required to perform the subtask. The subtasks are then scheduled across the total period of time it takes to complete the task. For some tasks involving certain types of flights—military, propeller, and international flights—the model increases the computed task times by 25 percent to reflect the assumed additional complexity of this traffic.

Of the eight triggered tasks, only the delay tasks (with the exception of the shortcut) are not divided into subtasks that use GOMS-derived times. The traffic operations and flight-planning data that trigger a delay event are used to separate the delay task into the following seven types: shortcut, low delay, medium delay, high delay, reroute, diversion, and hold. The shortcut task time of 11 seconds is derived from GOMS. Task times ranging from 25 to 75 seconds are assigned to each of the other six types of delay. According to CAASD, these delay task times were developed through consultations with operational experts.

Finally, task times are assigned for monitoring, which is the ninth modeled R-side task (Box 4). It is assumed that all flights are monitored by the controller using the radar screen, and thus the model assigns a monitoring task time to each flight that transits a sector. The assigned time can differ for each flight, depending on how long the aircraft stays in the sector and the other tasks it triggers. Monitoring times range from 0.2 to 0.8 seconds for every minute an aircraft is in the sector. The lower per-minute times are assigned to those flights that also trigger the delay tasks, under the assumption that a certain amount of screen monitoring is already included in the calculated delay task time (and thus avoid double counting of monitoring time).

For modeling ease, the model assumes that all of these task times are independent of one another and that tasks are performed sequentially by controllers rather than concurrently via multitasking. Total R-side task load is thus estimated by simply summing all of these times spent on tasks to generate 1-minute task loads averaged over 15-minute periods on a rolling basis.

CONVERTING R-SIDE TASK LOAD INTO PTT

Estimating PTT requires information on when the total R-side and D-side task load fully occupies the controlling time available to the lead controller. At that point a second controller is presumably needed to handle any more traffic. If both R-side and D-side task load is modeled, converting the task load output into PTT is a fairly straightforward process—when total modeled task load exceeds the assumed controlling time available to the lead controller, a second controller is required and PTT increases from 1 to 2. Yet as explained earlier, the CAASD model only computes R-side task load. Consequently, the total task load is not known, and therefore it is not possible to know when the R-side and D-side task load together are occupying all of the controlling time available to the lead controller. Thus CAASD had to develop methods for converting this limited R-side task load into estimates of PTT. How these conversions were originally made and how they are currently being made are summarized next.

Single Threshold for Conversion

The original method used to convert the modeled R-side task load into PTT was to assume that when a specific amount of task load is reached, an associate controller is needed to assist the lead controller. Specifically, the modelers assumed that if total R-side task load exceeded 600 seconds during a 15-minute period, the second controller was needed if the sector experienced any more traffic. According to CAASD, this 600-second threshold was initially developed after consultations with operational experts in which lower time thresholds (e.g., 500, 540, and 550 seconds) were considered. These lower thresholds imply that more D-side task load is associated with traffic than is implied by the 600-second cutoff.

For various reasons explained further in this report, CAASD was not satisfied with the PTT estimates produced by applying a single (600-second) threshold across all en route sectors. In particular, managers and controllers consulted in several of the en route centers expressed concern that using a single threshold neglected the variability in D-side task load that occurs across sectors and centers because of variability in traffic complexity. The staff in centers having more nonradar and international traffic pointed out that the D-side controller has a significant role in managing this traffic, for instance, by having to perform manual hand-offs. They claimed that in these cases the D-side task load is higher in relation to the modeled R-side task load than is implied by a single 600-second threshold.

Responding to these specific concerns, the model developers added new triggers for international and nonradar tasks and created rules that led to lower conversion thresholds when these tasks contributed a certain amount of the R-side task load in a sector. These adjustments led to the addition of a second controller at a lower R-side task load when the traffic consisted of a significant amount of international and nonradar flights.

Fuzzy Logic Modeling for PTT Conversion

Even after making these rule-based adjustments to the task load conversion process, model developers were concerned that the generated PTT values still did not fully account for how variability in traffic complexity affects total (R-side and D-side) controller task load, and thus PTT. The rule-based adjustments only accounted for the effect of two types of aircraft traffic activity—international and nonradar flights. The model developers, however, were interested in estimating how all variability in traffic complexity affects total task load and resultant PTT.

To account for more of this traffic complexity, CAASD turned to fuzzy logic modeling as another way to infer total task load from the modeled R-side tasks. On the basis of consultations with operational experts, the model developers characterized each modeled task according to its perceived complexity, that is, the extent to which it generates more or less D-side work in addition to the modeled R-side task load. Although the D-side tasks were never identified explicitly, three of the R-side tasks—entry, exit, and monitoring—were characterized by the experts as basic tasks that are accompanied by relatively little D-side activity. Three other tasks—transition, separation, and delay—were characterized as complex, generating a moderate amount of D-side work. Several other tasks, including those involving international and nonradar flights, were characterized as "other" and assumed to have the greatest amount of D-side task load. By characterizing and weighting the complexity of each modeled task in this manner, and then using fuzzy logic rules and algorithms to imply a total task load for different combinations of R-side tasks, this conversion method was viewed as providing a more traffic-dependent and realistic set of PTT values.

KEY POINTS FROM OVERVIEW

The specific elements of the task load model and PTT conversion methods are elaborated on in the remainder of this report: the following is a summary of the basic model structure and assumptions as outlined in this overview chapter.

- The task load model uses an array of traffic operations and flight-planning data to simulate traffic activity experienced in each of the en route sectors. By quantifying both the volume and type of traffic activity in each sector, the simulations provide a more complete picture of the traffic demand on controller time than is possible through simple counts of traffic alone.
- The task load model was originally developed to estimate the throughput capacity of en route sectors, believed to be a function of the lead controller's available time to perform R-side tasks. Because two controllers are needed to maximize throughput in a sector, the model assumes that two controllers are in position at all times, with the lead controller performing all R-side tasks and the associate controller performing all D-side tasks. The assumption that the lead controller is dedicated to R-side work, coupled with sector traffic capacity as a func-

tion of R-side task performance, has led to a model that only estimates R-side task load.

- To compute task load, the times spent on individual tasks are summed, and thus assumed to be independent of one another and performed sequentially rather than concurrently. Model task times are generated primarily through GOMS modeling and from information obtained from consultations with subject matter experts.
- To model PTT, which is an estimate of whether one or more controllers are required to work the traffic, requires information on total controller task load. Because the model only estimates R-side task load, various methods are employed to infer total task load. The method now being used, which employs fuzzy logic modeling, weights each R-side task according to its perceived D-side complexity. Although the D-side tasks are neither measured nor identified, these complexity weightings are treated as valid representations of D-side task loads because they were developed by using the judgment of operational experts.

3

Task Load Model

As shown in Figure 3-1, the task load model consists of four basic data and modeling components:

1. A defined set of R-side tasks performed by a controller when working traffic in a sector;
2. Analyses of traffic operations and flight-planning data to simulate the traffic activity in a sector during a time period and to indicate traffic events that trigger the occurrence of one or more of the defined R-side tasks;
3. The calculation and assignment of time spent by the R-side controller in performing each triggered task; and
4. Computation of total task load by summing the time spent by the R-side controller on these triggered tasks plus some additional time spent by the controller in monitoring the radar screen for all flights transiting the sector.

Explanations of each of these components are provided in this chapter along with a summary of the description by the Center for Advanced Aviation System Development (CAASD) of its efforts to evaluate and validate the traffic simulations, key model parameters such as task times, and the task load output from the model. The chapter concludes with the committee's assessment of each modeling element, including evaluation and validation efforts.

TASKS IN MODEL

Eight Triggered Tasks

Each of the triggered tasks is associated with the action of an aircraft under the control of a sector. In most cases, "tasks" are actually aircraft

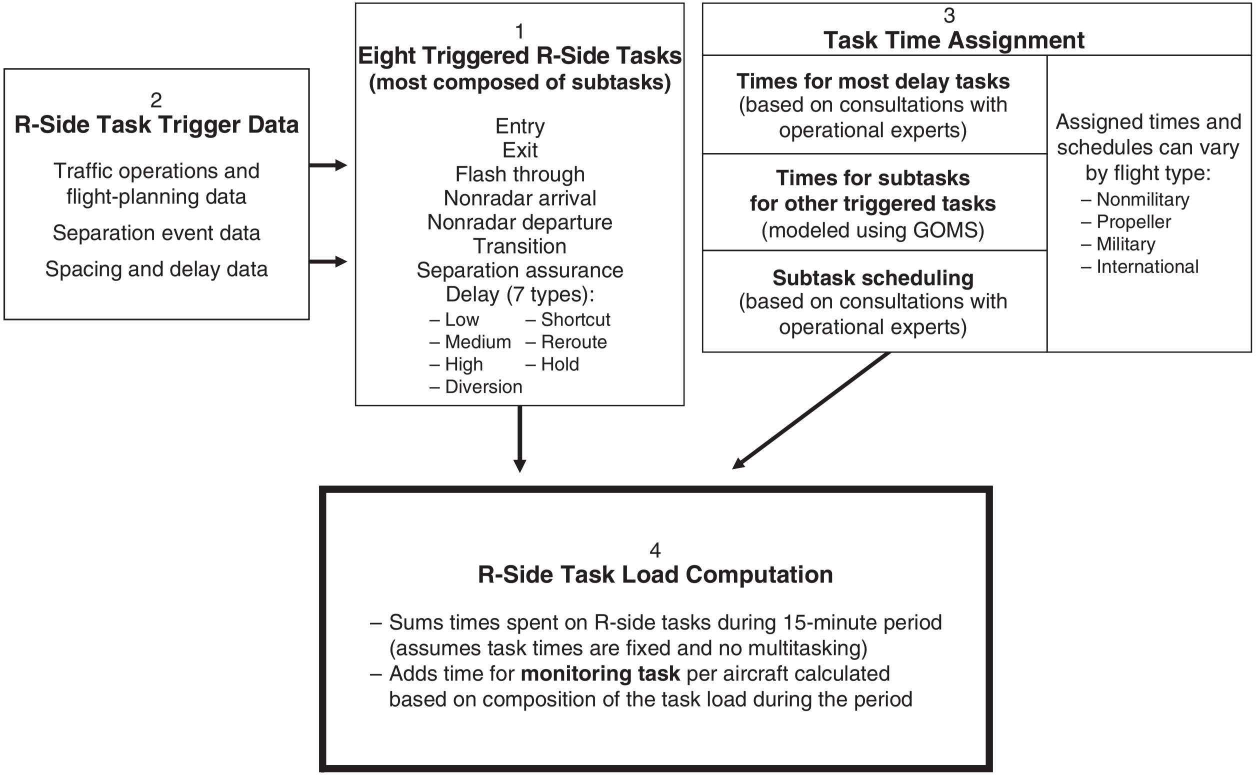

FIGURE 3-1 CAASD task load model.

actions, or events, that require the lead controller to execute a series of procedures. For example, an aircraft entering a sector is referred to as an entry task. When an entry occurs, the lead controller must take a number of actions that involve verbal communication with the pilot and other controllers, such as identifying the aircraft, establishing a clearance plan, and accepting the hand-off. Only the R-side procedures handled by the lead controller are included in the model. For reasons explained in Chapter 2, the CAASD model does not estimate D-side task load.

The eight tasks in the model that are triggered by traffic activity are the following:

- **Entry,** which encompasses the actions that the R-side controller must take in accepting the hand-off of an aircraft from another sector. Three types of hand-offs are distinguished in the model: those involving aircraft entering from international airspace, from a sector controlled by a different en route center, and from a sector controlled by the same center. Hand-offs are differentiated in this way in order to associate different task times with each.
- **Exit,** which encompasses the actions that the R-side controller takes in handing off an aircraft to a downstream sector. Hand-offs are differentiated in the same manner as for the entry task.
- **Flash through** is an entry in which the R-side controller handles the aircraft for less than 2 minutes. If after accepting the hand-off of an aircraft from an upstream sector, the controller determines that communication with the aircraft is not needed (e.g., the aircraft will be in the sector for a very short time and is separated from other aircraft), the controller will hand off the aircraft to the downstream sector without ever communicating with the pilot. The upstream sector controller will instruct the pilot to tune to the downstream sector's radio frequency. Because the flash through task entails less work than the full entry task, the assigned task time is less than that for the entry task.
- **Nonradar arrival** includes the actions taken by the R-side controller to provide services to aircraft arriving at an airport not offered the full range of radar services.
- **Nonradar departure** includes the actions taken by the R-side controller to provide services to aircraft departing from an airport not offered the full range of radar services.

- **Transition** consists of the actions performed by the R-side controller in clearing an aircraft to climb or descend and monitoring the aircraft to ensure that it is separated from other aircraft during the altitude transition.
- **Separation assurance** requires the R-side controller to identify aircraft pairs that are projected to lose lateral and vertical separation, determine a maneuver to ensure that separation will not be lost, and issue the maneuver clearance. Identification of this task does not mean that the aircraft pair has lost separation, only that a controller's attention was likely drawn to the pair to ensure separation.
- **Delay** involves the vectoring performed by the R-side controller for traffic separation and coordination such as merging aircraft in a stream. It also involves other controller actions that increase (delay) or decrease (cut short) the amount of time that aircraft are controlled by the sector such as holding, rerouting, and diversion.

All eight of these tasks are triggered in the model by a specific aircraft action, or event, that can be identified through the flight data that are used to simulate the traffic experienced in each sector for given time periods. The one modeled R-side task that is omitted from this list is a more generalized task referred to as "monitoring." This task entails scanning the radar display to maintain situational awareness and ensure that aircraft are following their clearances. All aircraft that transit the sector are assumed to require some amount of monitoring by the R-side controller. The modeling of the monitoring task is explained later in the discussion of task time derivation.

According to CAASD, these nine tasks are the main contributors to the lead controller's R-side task load. The committee was told that they represent about 90 percent of the R-side tasks performed based on CAASD's review of the literature and consultations with controller personnel and other subject matter experts.

TRAFFIC SIMULATIONS AND TASK TRIGGERS

Data Sources

The two sources of flight data used to simulate traffic and trigger tasks are the National Offload Program (NOP) and the CAASD Analysis Platform for En Route (CAPER). NOP is a collection of messages generated by each en route center's host computer system as traffic is worked through the center. It provides aircraft and sector identification, hand-off time, altitude, equipment type, origin, and destination. These data are used to trigger entries, exits, flash throughs, transitions, and nonradar arrivals and departures. NOP messages also provide the flight and equipment information needed to determine whether the flight is a military operation, which are allowed higher times for some tasks (as explained later).

Because FAA does not currently maintain an operations data set that contains information to trigger separation and delay tasks, CAASD uses its own analytic model to develop this information. By analyzing flight-plan and tracking data from FAA's Traffic Flow Management System, CAPER produces a four-dimensional projection of the path of an aircraft flying through the airspace from departure to arrival with associated transit times. Separation events are identified for aircraft pairs when the CAPER-modeled positions of two aircraft are predicted to come within a set of lateral and vertical threshold parameters. These parameters have been set to identify situations that would likely have drawn the attention of the R-side controller to ensure separation.[1]

The CAPER model is also used to trigger the reroute and hold delay tasks. It does this by capturing the change in the estimated time of arrival (ETA) as a flight enters and exits a sector. CAPER records relevant data for all active flights each time their modeled trajectories are updated. This process creates a historical record of the change in ETA for each flight including the time stamp for when the trajectory was updated and the reason for the update. For instance, a flight may enter a sector with an ETA of 1:00 p.m.; however, while it transits the sector it is vectored to provide the spacing needed to manage the traffic flows in a downstream sector. The extra time taken to vector is 5 minutes, and thus the ETA indicated upon exit from the sector will show 1:05 p.m. The model infers

---

[1] According to CAASD, the algorithms used by CAPER to detect separation events are similar to the conflict probe algorithms in the User Request Evaluation Tool (URET) automation, which is used operationally by en route controllers. The URET prototype was developed by CAASD to assist controllers with timely detection and resolution of predicted separation problems. URET is designed for the D-side controller, who typically has a more strategic look-ahead for potential separation. The prototype was used operationally for over 5 years at en route centers in Memphis, Tennessee, and Indianapolis, Indiana, to develop requirements for a capability to be deployed nationwide.

the type of action taken by assessing the size of the change in ETA. For instance, a large change in ETA is inferred to be caused by holding rather than a simple rerouting. As explained later, each of these types of delay is assigned a different task time. Finally, changes in flight-plan information are also assessed by CAPER to identify reroutes and diversions.

Evaluation and Calibration of Data Sources

CAASD performs various checks on the simulated traffic and its data sources, particularly the CAPER tool used for separation and delay triggers. The CAPER output, for instance, is routinely examined against recorded weather data to ensure that during known periods of severe weather the sector traffic simulations show an increase in the occurrence of delay tasks from holds, reroutes, and diversions. Likewise, checks are made to ensure that during periods of heavy traffic involving high volumes of crossing and transitioning aircraft the model detects higher levels of the separation-related task load.

When the use of CAPER to trigger separation events was first considered, CAASD compared the separations identified by CAPER for all Indianapolis en route sectors with actual aircraft conflicts identified in the sectors by a prototype of FAA's URET. Consistency in results led CAASD to conclude that CAPER would generate accurate separation event data for the task load model. Further analyses were performed to calibrate the parameters used by CAPER in detecting separation events. These analyses included the random sampling of URET conflicts and the checking of flight-track histories, flight-plan amendments, and controller–pilot voice recordings. In addition to the CAPER–URET comparisons, CAASD performed sensitivity analyses to evaluate the use of alternative parameter thresholds for lateral and vertical separation minimums as well as the maximum look-ahead time for CAPER to probe flights along their trajectories.

On the basis of these analyses and through an iterative process of calibrating the CAPER trajectory modeler and air-to-air separation probe algorithms, CAASD believes there is a strong correspondence between URET conflicts and CAPER separation events. Similar checks have apparently been performed on CAPER results used for triggering the delay events. In addition, CAASD continues to sample the modeled flights to verify whether the calculated delay is consistent with observed flight operations. These comparisons with actual traffic observations are used to adjust the CAPER parameters when warranted.

TASK TIMES AND SCHEDULES

Derivation of Task Times

The time assigned to each triggered task is a critical element of the task load model, since the task load output is a time summation. In briefings to the committee, CAASD maintained that it encountered difficulty establishing task times from the literature. Model developers apparently could find little documentation on the task times used in other relevant efforts to model controller task performance, and in the few cases in which task times could be identified, the tasks did not represent the same set of actions or assumptions as those used in the CAASD model. Hence, whereas early versions of the model tried to use literature-based task times—which had to be modified and supplemented by estimates from operational experts—CAASD later concluded that another process was needed to derive these times.

CAASD's investigation of options for deriving task times led it to select the modeling process known as Goals, Operators, Methods, and Selection Rules (GOMS) to develop many of the task times in the model. GOMS is a type of human information processor model that is used to predict user performance for a given task and to provide an estimate of how much time it takes to accomplish the task. In short, the model assumes that humans pursue tasks according to goals. Each goal is accomplished by employing various "operators" consisting of cognitive processes, perceptions, and motor actions. These operators are sequenced into "methods" that relate how the operators are used to accomplish the goal. Because there can be more than one method for accomplishing a goal, various selection rules (e.g., "if–then" statements) are employed to describe when a user would choose one method over another. In this way, GOMS modeling predicts the time it takes for a person to accomplish a task by associating specific times with each operator and then sequencing them according to the selected method.

Specifically, GOMS-generated task times are used in the CAASD model for the entry, exit, flash through, transition, separation, nonradar, and shortcut (delay-related) tasks. Accordingly, CAASD had to decompose each of these tasks into constituent subtasks and then further convert each subtask into its execution method and the specific operators involved (e.g., uttering words, pressing keys) as modeled by GOMS. To identify the subtasks and their execution methods and operators, model developers consulted with subject matter experts. The nature of these consultations and the methods used were not documented or explained to the committee other than to describe them as involving an iterative process until experts were satisfied with the decomposition of subtasks into operators. CAASD then simulated the actions for each operator to develop execution times. Documentation on the method of simulation was not provided, although apparently these simulations did not attempt to capture all of the perceptual or cognitive processes that affect execution time, but only the time required to perform motor actions. While CAASD informed the committee that the GOMS estimates do include subtasks such as identification of the problem, priority-ranking of the problem, and generation of a problem resolution, the potential impact on task time estimates from not including the time required for perceptual and cognitive processes was never addressed in depth.

The subtasks associated with each of the eight triggered tasks identified above, and their GOMS-generated execution times, are shown in Tables 3-1 and 3-2. It merits noting that GOMS is not used for most of the delay tasks because model developers could not identify constituent subtasks. Accordingly, CAASD consulted with subject matter experts to establish the times assigned to the delay-related actions of rerouting, holding, and diverting. To establish these times, traffic replays were presented to the experts, who estimated the task completion time. These delay task times range from 25 to 75 seconds.

The consultations with subject matter experts also led CAASD to conclude that the longer task times were needed for certain types of traffic, particularly flights involving propeller and military aircraft. For this particular traffic, assigned task times are increased by 25 percent for entry, exit, and nonradar arrivals and departures, under the assumption that additional communication and coordination are required. Although the validity of this adjustment factor was not researched by the committee, CAASD maintains that it was derived from information in FAA's Position Classification Standard for Air Traffic Control Series.

TABLE 3-1 Times Assigned to R-Side Tasks and Subtasks

Event or Main Task	Subtask	Subtask Time (s)	Task Time (s)
Entry	Identify aircraft	1.5	17.4
	Establish clearance plan	6.4	
	Pilot call-in, hand-off accepted	9.5	
Exit	Identify aircraft	1.5	14.0
	Automated hand-off	1.3	
	Change frequency	11.2	
Flash through	Identify aircraft	1.5	12.3
	Verify flight path exit sector	1.3	
	Contact center controller	9.5	
Transition	Identify aircraft	1.5	14.6
	Determine altitude	1.6	
	Determine clear path	1.4	
	Issue clearance	5.4	
	Listen to readback	4.7	
Nonradar arrival	Identify aircraft	1.5	52.3
	Altitude assignment/restriction	9.3	
	Traffic advisory	11.8	
	Issue approach clearance	9.7	
	Weather issuance	9.3	
	Change frequency	10.7	
Nonradar departure	Identify aircraft	1.5	21.0
	Altitude verification	1.9	
	Radar identification	8.3	
	Altitude assignment/restriction	9.3	
Separation	Identify aircraft	1.5	27.6
	Determine Vector 1	1.3	
	Ensure clear path	1.4	
	Issue Vector 1 clearance	9.1	
	Determine Vector 2	1.3	
	Ensure clear path	1.4	
	Issue Vector 2 clearance	7.5	
	Listen to readback	4.1	
Delay (shortcut)	Identify aircraft	1.5	11.0
	Determine clear path	1.4	
	Issue direct clearance	5.0	
	Listen to readback	3.1	
Delay (low)		None	25.0
Delay (medium)		None	35.0
Delay (high)		None	45.0
Delay (reroute)		None	60.0
Delay (diversion)		None	75.0
Delay (hold)		None	3 s/min in holding
Monitoring		None	Variable

s = seconds; min = minutes.

NOTE: When aircraft is military or propeller, all task times for entry, exit, and nonradar arrivals and departures increase 25 percent.

TABLE 3-2 Task Scheduling Distributions

Task	Distribution Minutes	Distribution Type
Entry	E to $E + 3$	Quasi-uniform
Exit	$X - 2$ to $X + 2$	Custom
Flash through	E to $E + 1$	Uniform
Transition	$\frac{1}{2}(E + X)$ to $\frac{1}{2}(E + X) + 1$	Uniform
Nonradar arrival	$E + 4$ to X	Uniform
Nonradar departure	$E - 2$ to $E + 3$	Uniform
Separation	$S - 6$ to $S + 1$	Quasi-uniform
Delay	E to X	Uniform

E = minute of hand-off from upstream sector
X = minute of hand-off to downstream sector
S = minute separation event begins

To assign times to the monitoring task, CAASD evaluated several alternative approaches. At first, the model assumed that a fixed amount of controller time is spent monitoring each aircraft transiting the sector. However, because some of the task times already presumed a certain amount of monitoring—especially in the case of the time-consuming delay tasks—CAASD was concerned that such a fixed time would lead to double counting and an overestimation of total monitoring task time for some flights. Model developers therefore created an algorithm to produce a monitoring task time for each aircraft transiting a sector. The algorithm assigns a monitoring time per minute a flight is in the sector. The assigned time varies depending on the composition of the R-side task load generated by the flight. Hence, if the flight's R-side task load consists of many tasks identified as being complex (such as delays that already include a large amount of monitoring time), the algorithm selects a lower monitoring rate per minute because it is assumed that a large amount of monitoring time is already included in the flight's R-side task load.

Task Scheduling

As discussed earlier, many of the tasks used in model are made up of subtasks. These subtasks are presumed to be performed by the R-side controller in a defined order but interspersed with subtasks being performed for other tasks. Because the model computes task load for 1-minute intervals to obtain rolling 15-minute task load estimates, CAASD needed a method for distributing the subtasks over the entire span it takes to complete the full task. The model therefore sequences subtasks in a pattern that is thought to be typical for the task and over a completion period that is considered typical.[2] These sequencing and scheduling profiles, which are shown in Tables 3-1 and 3-2, were determined through consultations with subject matter experts.

As shown in Table 3-2, the entry, exit, and separation tasks are assumed to have nonuniform distributions, since the majority of the task time is scheduled at the point when the task is triggered. Time for transition, flash through, and delay tasks is uniformly distributed over the life of the flight in the sector. The time for nonradar arrivals and departures is uniformly distributed on the basis of the entry time identified by the trigger.

Evaluation of Task Times

CAASD selected GOMS to model task times because it was viewed as providing efficiency and flexibility, since the time estimate can be calculated with relatively little effort if the operators and methods are known and accurate operator time data are available. In discussion with model developers, committee members noted that GOMS models are more commonly used to assess user performance across various prototype products and systems in which there are few practical opportunities for direct observation of human performance in operational or experimental settings, such as evaluating alternative workstation layouts and computer interfaces. When questioned about the applicability of GOMS to the air traffic control environment—which is an observable operational setting—the model developers restated their belief that GOMS offers the needed efficiency and flexibility and provided the 12 literature sources identified in Box 3-1

---

[2] For example, a sector entry is identified from a hand-off message contained in NOP data, with the time of the message being noted. The entry task time is then distributed around the time of the hand-off message relative to a typical sequence of actions, or subtasks, that a controller performs for a hand-off. The larger portion of the task time is scheduled at the minute that the hand-off is accepted in consideration of the actions a controller typically performs when accepting a hand-off: determining that the aircraft is not in conflict with other aircraft and that it is following its recorded flight plan. The remaining time is spread over the few minutes following the hand-off message representing the time the controller would be monitoring the aircraft until it actually crosses the boundary into the controller's airspace.

TABLE 3-3 Comparison of GOMS and HITL Operator Times by Data Source and Operator

	GOMS	HITL Average	HITL Standard Deviation	GOMS Value as a Percent of HITL Average
Syllable utterance	150 ms	187 ms	12.3 ms	80.2
Keystroke	280 ms	247 ms	25.7 ms	88.2
Fixation	500 ms	542 ms	117.3 ms	92.3

ms = milliseconds.
SOURCE: CAASD submission to committee.

as containing examples of previous research supporting the use of GOMS for modeling task times in situations similar the air traffic control domain. While the committee did not review each of these sources, it notes that only a few (4 of the 12) appear to involve air traffic control tasks.

Questioned on how the GOMS times were validated, CAASD pointed to the limited comparisons that have been made with experimental data from human-in-the-loop (HITL) experiments conducted in 2008 by FAA. As shown in Tables 3-3 and 3-4, when the GOMS operator times are compared with the HITL times, the former are found to be 10 to 20 percent lower than the latter. According to CAASD, these GOMS error rates are comparable with GOMS error rates generally, as observed from the literature cited in Box 3-1.

However, because the FAA HITL experiments were not conducted for the specific purpose of developing model times, they could only be used to assess some of the task times. CAASD has therefore consulted with

TABLE 3-4 Comparison of GOMS and HITL Times: Times Aggregated to Task Level with Typical Usage*

	GOMS	HITL Average	GOMS Value as a Percent of HITL Average
17 syllables	2,550 ms	3,181 ms	80.2
5 keystrokes	1,400 ms	1,237 ms	88.4
2 fixations	1,000 ms	1,083 ms	92.3

*A typical usage would be a pilot readback consisting of 17 syllables.
ms = milliseconds.
SOURCE: CAASD submission to committee.

BOX 3-1

Sources Cited by CAASD in Support of Using GOMS Modeling for Deriving Controller Task Times*

Card, S., T. P. Moran, and A. Newell. 1983. *The Psychology of Human–Computer Interaction.* Lawrence Erlbaum Associates, Mahway, N.J.

Endestad, T., and P. Meyer. 1993. *GOMS Analysis as an Evaluation Tool in Process Control: An Evaluation of the ISACS-1 Prototype and the COPMA System.* Technical Report HWR-349. Organization for Economic Cooperation and Development Halden Reactor Project. Institute for Energiteknikk, Halden, Norway.

Estes, S., C. Bonaceto, K. Long, S. Mills, and F. Sogandares. 2009. Carbon Copy: The Benefits of Autonomous Cognitive Models of Air Traffic Controllers in Large-Scale Simulations. In *Proceedings of the 8th USA/Europe Air Traffic Management Research and Development Seminar,* Napa, Calif.

Gong, R. 1993. *Validating and Refining the GOMS Model Methodology for Software User Interface Design and Evaluation.* PhD dissertation. University of Michigan, Ann Arbor.

Gray, W. D., B. E. John, and M. E. Atwood. 1993. Project Ernestine: A Validation of GOMS for Prediction. *Human–Computer Interaction,* Vol. 8, No. 3, pp. 237–309.

Irving, S., P. Polson, and J. E. Irving. 1994. A GOMS Analysis of the Advanced Automated Cockpit. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Celebrating Interdependence,* April 24–28, Boston, Mass., pp. 344–350.

Kieras, D. E., S. Wood, K. Abotel, and A. Hornof. 1995. GLEAN: A Computer-Based Tool for Rapid GOMS Model Usability Evaluation of User Interface Designs. In *Proceedings of the 8th Annual Association for Computing Machinery (ACM) Symposium on User Interface and Software Technology,* Nov. 15–17, Pittsburgh, Pa., pp. 91–100.

Lee, A. 1992. Accuracy of MHP/GOMS Predictions for the Task of Issuing Recurrent Commands. In *ACM–Special Interest Group on Human–Computer Interaction (SIGCHI) Conference on Human Factors in Computing Systems,* Monterey, Calif., pp. 105–106.

(continued)

BOX 3-1 *(continued)*

Sources Cited by CAASD in Support of Using GOMS Modeling for Deriving Controller Task Times*

Lee, S. M., U. Ravinder, and J. C. Johnston. 2005. Developing an Agent Model of Human Performance in Air Traffic Control Using APEX Cognitive Architecture. In *Proceedings of the 2005 Winter Simulation Conference,* Orlando, Fla., Vols. 1–4, pp. 979–987.

Nesbitt, K., D. Gorton, and J. Rantanen. 1994. *A Case Study of GOMS Analysis: Extension of User Interfaces.* Technical Report BHPR/ETR/R/94/048. Newcastle Laboratories, Wallsend, New South Wales, Australia.

Ravinder, U., R. W. Remington, and S. Lee. 2005. A Reactive Computational Model of En-Route Controller. In *Proceedings of the 2005 IEEE International Conference on Systems, Man and Cybernetics,* Oct. 10–12, Waikoloa, Hawaii, pp. 1628–1633.

Smith, E. C. 2008. Flight Management System Execution Task Time Modeling for Loading Terminal Area Navigation Procedure Changes. In *Proceedings of the Human Factors and Ergonomics Society 52nd Annual Meeting,* Vol. 52, No. 13, pp. 912–916.

---

* GOMS Modeling for the MITRE En Route Workload Model, a briefing presented by MITRE to the TRB Committee for a Review of the En Route Air Traffic Control Complexity and Workload Model, March 2010.

operational experts to obtain their opinions on the validity of the GOMS times. Although the nature of these consultations was not explained, they apparently led CAASD to conclude that the GOMS times are generally reasonable but require some adjustment to represent the times associated with speaking, which the experts thought were too high. The experts consulted also raised questions about the model's basic assumption that tasks are performed sequentially rather than in parallel in some instances.

When the committee asked about the possibility of performing dedicated HITL experiments, CAASD agreed that doing so could yield a rich array of information but restated the concern that the experiments can be time-consuming and expensive. Model developers reported that they are continuing to evaluate the 2008 HITL data to assess the prevalence of multitasking and to research the GOMS speech operator to make it produce times that are closer to those indicated by the HITL data and the judgment of consulted experts. In general, CAASD believes the future use of HITL experiments will be confined to the development and validation of a select number of task performance times because of the perceived time and expense of conducting such experiments.

TASK LOAD COMPUTATION

Rollup to Task Load

As previously discussed, the task times are scheduled in 1-minute intervals. The model processes the 1-minute intervals with the output rolled up to a larger time interval, typically 15 minutes. The processing that the model performs to roll up 1-minute task load is summarized by the following equation:

$$W_n = \sum_{i=n-7}^{n+7} \sum_j \sum_k w_{ijk}$$

where

W_n = 15-minute workload at minute n, and
w_{ijk} = 1-minute workload at minute i due to task j being performed in service of aircraft k.

The task load output is computed by summing all of the time spent on R-side tasks during the measured period. In theory, the highest value for R-side task load for a 15-minute period is 900 seconds (15 minutes times 60 seconds/minute), assuming (unrealistically) that a controller can effectively use all 900 seconds of available time and that a second controller is handling the D-side task load. As discussed in Chapter 4, when the task load rollup exceeds a certain threshold (around 600 seconds), it is assumed that two controllers are working the traffic.

Evaluation of Task Load Rollups

At various stages in the development of the model, CAASD has undertaken evaluations of its task load output for accuracy. Early evaluations, including the one described in Box 3-2, suggested that the results of the model were a major improvement over the volume-based (aircraft count) metrics that had been used previously to inform controller staffing requirements. These initial evaluations caused FAA to favor the task-based approach over the earlier methods.

In more recent evaluations, CAASD once again turned to experts for their opinions on the task load output. In 2006, CAASD assembled a group of front-line managers from 10 en route centers.[3] Each manager was briefed on the background, objectives, and outputs of the task load model. Before the model results were presented, the managers were asked to rank their respective sectors by traffic complexity. The rankings were then compared with the rankings of the same sectors based on the traffic simulations and task loads generated by the model. The participants were asked if the model's results were accurate in characterizing the individual sectors in terms of typical traffic volume and types of activity (e.g., prevalence of separations, delays, transitions). According to CAASD, for most of the sectors the managers responded that the model results closely matched their own perception of sector traffic complexity.

Although based on perceptions, these assessments were used by CAASD as guidance in making further refinements to elements of the model, particularly the delay task. The evaluations were also one of the factors that caused CAASD to seek additional information to represent international flights and flights to and from airports with no radar services.

COMMITTEE ASSESSMENT

FAA asked the committee to examine the input data and processes used for modeling traffic activity, the tasks and methods used to assign task times, and the means for validating model assumptions, parameters, and output. An assessment of each is offered next.

---

[3] Atlanta, Georgia; Boston, Massachusetts; Dallas, Texas; Denver, Colorado; Houston, Texas; Memphis, Tennessee; Minneapolis, Minnesota; New York City; Salt Lake City, Utah; and Seattle, Washington.

BOX 3-2

CAASD Comparisons of Task Load Output with Results of Dynamic Density Experiments

A concept known as dynamic density was critical to the free flight paradigm guiding the planning of the NAS in the late 1990s and early 2000s. The dynamic density concept is built on two of the same basic principles as the CAASD model: (*a*) complexity affects the capacity of a sector, and (*b*) complexity is dynamic and changes over the course of a day for a sector. In 2002, the FAA requested that CAASD evaluate the effectiveness of four sets of dynamic density metrics developed by various research organizations for predicting air traffic complexity as perceived by controllers.* To conduct the study, traffic scenarios were evaluated by controllers in HITL experiments at the FAA Technical Center using a rating scale from the National Aeronautics and Space Administration (NASA) known as the Air Traffic Workload Input Technique (ATWIT). Using ATWIT as the basis, controllers were asked to rate their subjective assessment of the complexity level they experienced on a scale of 1 to 7.

CAASD leveraged the results from those dynamic density experiments in a 2004 analysis of the task load model, comparing the scores provided in that study with the values generated by running the model with scenario data obtained from the dynamic density HITL experiments. The results of the study indicated that as estimated task load increased, the controller-perceived complexity rating tended to increase as well. While the actual predicted amount of task load was not validated, CAASD believes that the analysis demonstrated consistency between increasing task load and increasing complexity. In addition, CAASD concluded that the results indicated that the output of the model outperformed aircraft count as a predictor of both perceived complexity and the number of required controllers as rated by operational experts involved in the original dynamic density experiments.

---

* Holly, K., Y. Cabeza, M. Callaham, D. Greenbaum, A. Masalonis, and C. Wanke. 2002. Feasibility of Using Air Traffic Complexity Metrics for TFM Decision Support. MTR 02W0000055. MITRE Corporation, McLean, Va.

Traffic Modeling

Compared with simple traffic counts, the simulations of traffic in the CAASD model provide a more complete picture of both the volume and nature of traffic activity in the en route sectors. The simulations are developed through an array of traffic operations and flight-planning data that represent opportunistic use of many existing traffic data and modeling tools. The traffic activity is modeled in sufficient depth and resolution to enable reasonable approximations of traffic complexity and associated controller tasks. Because the simulated traffic can be checked against records of actual traffic activity, there is ample opportunity to validate the output accuracy and to guide the development and recalibration of modeling processes and parameters. CAASD appears to have taken advantage of these opportunities to improve the traffic modeling capabilities.

Task Coverage

The nine tasks in the model appear to be representative R-side services that must be performed in response to traffic. However, CAASD's assertion that the model covers 90 percent of the R-side tasks is not well established. To be sure, all R-side responsibilities are not modeled; for instance, the committee observes that there are no tasks associated with issuing weather and traffic advisories, which is an R-side service. While such unmodeled tasks may or may not have a significant effect on task load, the rationale for their absence and the potential impact on task load need to be addressed.

Compared with the other modeled tasks, monitoring is the most confusing and difficult to connect to traffic activity. Monitoring involves scanning of the radar display by the controller to maintain situational awareness of flights under sector control. The model assumes that monitoring is performed by the R-side controller for all traffic, which is a reasonable assumption. It is assumed further that a certain (but undefined) amount of monitoring is already included in other task times, particularly in the time-consuming delay tasks. CAASD nevertheless added a separate monitoring task so as not to underestimate monitoring, particularly for the most straightforward traffic transiting a sector. While monitoring is an important task, its treatment in the model is confusing and unconvincing. Since the concern about overestimating monitoring time relates mainly to the delay tasks, a simpler and more transparent treatment would be to define how much monitoring time is already included in these task times.

It is important to keep in mind that the nine tasks in the model represent only the R-side tasks. In considering the scope of R-side services only, the modeled tasks may be adequate in coverage. From the standpoint of estimating PTT, however, the model provides an incomplete picture of controller task load because the modeled tasks are not linked to D-side services. More consideration is given to this shortcoming in the following chapter.

Task Time Derivation

For seven of the nine modeled tasks, GOMS is used to derive task times. The other task times are developed through consultations with subject matter experts. None of the task times is derived from the observation and analysis of controllers performing tasks in the field or in experiments.

CAASD's comparison of some GOMS and HITL times indicate that the former are 10 to 20 percent lower than the latter. These limited comparisons, however, are the only means by which task times have been evaluated, apart from asking subject matter experts to assess them. CAASD claims the literature lacks relevant task times, prompting it to use GOMS and other means for estimating times. CAASD selected GOMS as a primary method believing it to be an efficient and inexpensive approach, particularly when compared with gathering and analyzing data from operational and experimental settings, such as those for HITL experiments. CAASD believes that GOMS modeling will allow for continued updating of the task times even as controller procedures and capabilities change.

The committee questions the extensive use of GOMS for task time derivation and the complete absence of task times developed through field observation or HITL experiments. GOMS modeling is typically used where there is limited opportunity to observe and analyze task performance in operational or experimental settings. These conditions do not exist in the air traffic control domain. The GOMS-derived times are based largely on expert judgment, and only loosely validated against task performance data obtained from HITL experiments conducted for other purposes. Given these circumstances, there is no way to know whether the task times used in the model are at all valid.

Computation of Task Load

The addition of task times to calculate R-side task load may be the simplest approach to computing task load while still being reasonable. However, adding task times does not account for the possibility—and real-world probability—that some tasks are performed concurrently and that the time it takes to performs tasks can change depending on the total task load or the number of controllers working the sector. These tenuous assumptions may or may not be critical to the task load results. Examining their potential impact on task load, however, is important for making a convincing case that the assumptions do not represent serious modeling deficiencies. This case has not been made.

4

Converting Task Load into Positions to Traffic

Converting task load estimates into positions to traffic (PTT) requires knowing when the total time spent by the lead controller on both R- and D-side tasks reaches a point where assistance from an associate controller is needed. As explained in the previous chapter, the CAASD model produces estimates of the R-side task load only. The challenge, therefore, is in finding ways to convert this limited task load into values for PTT.

The first part of this chapter describes the methods employed by CAASD to make these conversions. The second part consists of CAASD's own evaluations of the conversion results, and the last part gives the committee's assessment of the methods.

CONVERSION METHODS

Time Threshold

CAASD has used two basic methods for converting the modeled R-side task load into PTT. The first, which is no longer being used, presumes that once R-side task load reaches a given threshold, then an associate (D-side) controller is needed to help work the traffic in the sector. The time threshold originally used by modelers—600 seconds during a 900-second (15-minute) period—presumes that at this point the combination of modeled R-side tasks and unmodeled D-side tasks fully occupies the controlling time available to the lead controller. The specific 600-second threshold was identified by model developers on the basis of consultations with facility managers, who found that the resulting PTT values were closer to their expectations of staffing levels associated with the modeled traffic than those generated by other cutoff points (550, 580, etc.). Figure 4-1

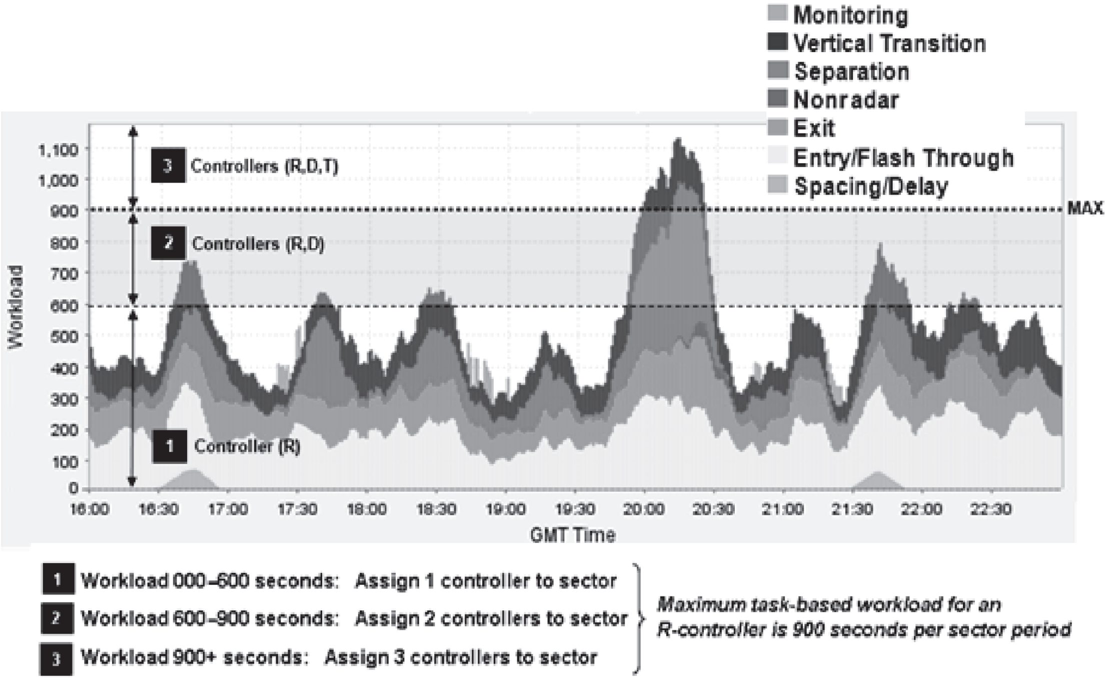

FIGURE 4-1 Example use of 600-second (R-side) task load threshold to estimate PTT over 8-hour period.

provides a graphic representation of the R-side task load converted to PTT in this manner for a single sector for an 8-hour block of time.

Nevertheless some of the facility managers questioned whether a single threshold was appropriate for predicting PTT across sectors that encountered wide variability in traffic patterns and complexity. They observed, for instance, that the D-side task load tends to be higher for some types of traffic activity than for others, which would imply the need for a second controller at a threshold lower than 600 seconds for R-side task loads generated by such traffic. For instance, in some sectors the complexity of traffic may be relatively straightforward, consisting of mostly entries and exits as aircraft transit a sector, which generates minimal D-side work. In other sectors (or even in the same sector at a different time of day), traffic patterns may be more complex, consisting of more delays, international entries and exits, and nonradar arrivals and departures, which creates much more D-side work.

In response to these concerns, CAASD added new triggers to its model for R-side tasks associated with international and nonradar traffic activity. In addition, new rules for nonradar and international tasks were added to adjust the 600-second conversion. Specifically, if the total R-side task load were less than 600 seconds but consisted of at least 100 seconds of time spent on nonradar tasks or if more than five aircraft were in the sector when any amount of time was spent on nonradar tasks, then two controllers were assumed to be needed. Likewise, if total R-side task load were less than 600 seconds, but international task load exceeded 40 seconds or if more than 10 aircraft were in the sector along with any amount of time spent on international traffic, then two controllers were assumed to be needed. CAASD referred to this rule-adjusted conversion as the “600-plus” method. In general, the PTT estimates from this adjusted method were found to be more in line with the expectations of operational personnel consulted from facilities having significant international and nonradar activity.

Fuzzy Logic Modeling

Even after rules for international and nonradar tasks were applied, CAASD worried that other variability in sector traffic patterns was creating even more variability in total task load than could be accounted for by the modified 600-second cutoff. Accordingly, CAASD considered developing even more rule sets, basically extending its "600-plus" method. To do so, however, required more information on the D-side task load associated with different patterns and volumes of traffic activity. Yet CAASD only had the qualitative judgments of operational experts to assess total task load—that is, their judgments about whether one traffic profile, or R-side task load, tended to create more or less D-side task load. CAASD concluded that the inference rules used in fuzzy logic modeling might be well suited to inferring total task load from these qualitative judgments.

In an explanatory document submitted to the committee,[1] CAASD described the fuzzy modeling process and its purpose as follows:

> Fuzzy logic involves setting multiple thresholds for each input variable, and then creating rules of interaction. The technique has three distinct steps: fuzzification, inference, and defuzzification. Each of these steps is discussed below relative to the fuzzy model for PTT. Fuzzification is the process by which a degree of membership is determined for each of the eight task workload inputs (entry, monitor, exit, transition, separation, delay, international, and nonradar). Three overlapping fuzzy terms were used for all task workloads: "low," "medium," and "high." These terms are referred to as linguistic variables and represent the relative degrees of difficulty, i.e., total team workload for increasing degrees of specific R-controller task workload. The membership function for each of the three terms, which ranges from 0 to 1, was calibrated separately for each workload task. For each of the tasks, there is an inflection point where membership is equal to 1. Figure 4-2 shows an example of how the Entry input variable looks in the software interface.
>
> The second step in fuzzy modeling is to apply inference rules. Once the level of membership has been determined for each task workload relative to each linguistic variable (i.e., fuzzification), this level of membership is combined with similar information for other task workload in the same grouping. Three task groups are used in the model: basic tasks, complex tasks, and other tasks. These task groups were chosen based on their staffing impact. Basic workload tasks alone (entry, exit and monitor), will not require a second or third controller, unless they are relatively elevated due to high traffic levels. However, if there is a significant amount of complex task workload present (spacing, delay, and transition), an R-side will be more likely to need

---

[1] The committee asked CAASD to draft a paper explaining the task model load and processes used to convert its output into PTT. The quoted sections that follow are derived from this submitted paper, which can be obtained from TRB.

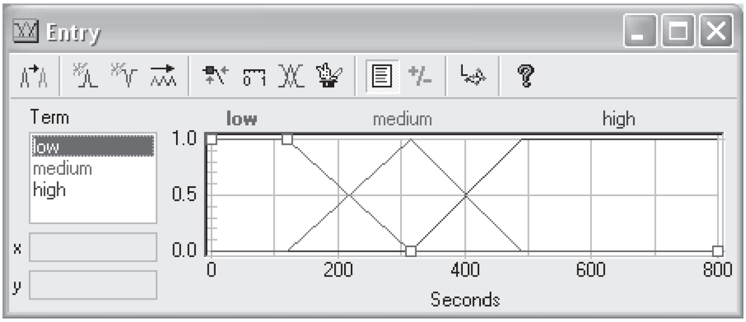

FIGURE 4-2 Entry input variable in fuzzy logic model software interface.

assistance. Also, as mentioned earlier, if other task workload (international and nonradar) is present, then it is highly likely that a D-side is present to assist with the nonautomated hand-offs.

Each of the three task groupings uses a system of “if–then” statements to translate individual task weights into task group weights, and then a final task group weighting is translated into an estimate of the number of controllers needed. Varying degrees of these rules are simultaneously activated or “fired” based on the individual contributions of the input linguistic variables (e.g., a medium will contribute more than a low). Figure 4-3 illustrates the framework used in the PTT fuzzy model.

The final step in the process is defuzzification. It involves applying all of the inference rules, which are weighted, to obtain a definitive solution value. The model produces both a discrete number (0, 1, 2, or 3) and a value that can be a fractional value between 0 and 3. The discrete value is generated by a process known as the Mean of Maximum (MoM) method and is the method used to translate the final degree of membership into the discrete PTT variable (0, 1, 2 or 3). This methodology is typically used for classification problems and produces the most plausible or likely result. The other method that produces the fractional value between 0 and 3 is known as Center of Maximum (CoM). Although the fractional value is not used for the PTT data provided to the FAA for the CWP staffing models, it has been useful in calibration of the model. It indicates whether the model was close to producing a different value for the discrete method used for the PTT data. For example, a value of 2.4 indicates that the workload is trending towards the need of a third (T-side) controller.

Essentially, the translation produced by the PTT fuzzy model reflects how operational experts characterize position needs: low degree of workload is equivalent to one controller, medium degree of workload is equivalent to

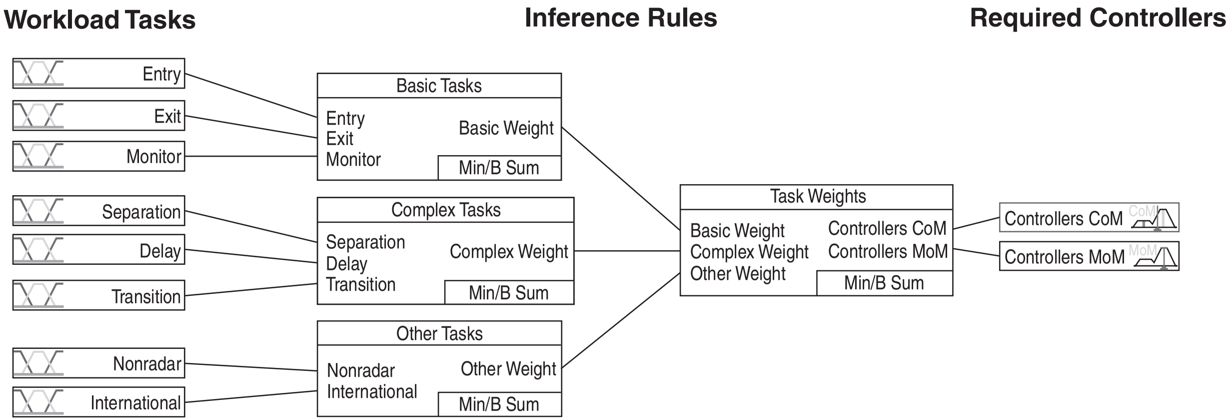

FIGURE 4-3 Framework for fuzzy logic modeling process to infer PTT.

> two controllers, and high degree of workload is equivalent to three controllers. The translation is performed without having to define "low," "medium," and "high."

The last paragraph in CAASD's description explains what the fuzzy logic model is doing to generate PTT values. In essence, it is assigning additional task time to each of the modeled R-side tasks based on expert opinions on the associated D-side task load. However, the implied D-side tasks are never identified, nor are the times assigned to them by the experts made explicit. Nevertheless, the D-side task loads must be determined in order to generate the PTT values associated with the various combinations of modeled R-side task load.

That such D-side task load estimates are being made, however vaguely and opaquely, raises questions about the validity of this modeling process and whether the characterizations of the operational experts are indeed accurate. The committee sought, but was not was presented with, the total task loads that are implied by the different combinations of R-side tasks that generate specific PTT values. Making explicit these implied D-side task loads so that they can be assessed is essential to judging the validity of the PTT estimates produced through fuzzy logic modeling.

Although not checked in this most fundamental manner, the fuzzy logic modeling process is being used by CAASD for its PTT conversions, and model developers have indicated satisfaction with the results. In the next section, the methods used by model developers to assess the conversions are examined.

CAASD EVALUATIONS OF PTT CONVERSIONS

The PTT conversions were evaluated by CAASD primarily on the basis of (*a*) consultations with operational experts asking them to judge whether the results seem reasonable and (*b*) comparisons of the model output with available operational data and staffing records.

A dilemma in consulting facility personnel and staffing records to validate PTT estimates is that a central purpose of PTT modeling is to assess whether staffing levels are aligned with experienced traffic demand. A problem with asking facility personnel to assess PTT estimates is that their response may be skewed by a view that existing staffing levels are optimal. Likewise, consulting staffing records can be problematic. On the one hand, if the records show that one controller successfully worked the traffic in a sector when the model estimated a need for two controllers, then such a comparison could be helpful in identifying modeling problems that tend to overestimate PTT. Indeed, it is through such comparisons that volume-based methods of generating PTT were found to be lacking. On the other hand, if the PTT model indicates a need for one controller when staffing records show that two were in position, it is much more difficult to ascertain whether the model underestimated the need for a second controller or whether actual staffing levels were too high for the experienced traffic activity.

Review of PTT Estimates by Facility Personnel

To assess its PTT estimates, CAASD presented the results to managers and controllers at 13 centers spread across the country.[2] At each evaluation session, participants were given a general overview of the modeling process. Controllers then analyzed the task load and PTT output for a high-traffic day as well as for overall seasonal staffing averages for their area. Specific feedback was sought on the accuracy of the model in capturing the type and quantity of task load as well as in producing reasonable PTT estimates. The results of these evaluations were apparently used to modify the task load model and the conversion methods, although the specific adjustments made in response to each facility visit were not explained to the committee. Nevertheless, according to CAASD, the recommended number of changes to the model declined with each center visit. CAASD took these developments as indicative of a model that was providing an increasingly accurate portrayal of PTT.

Comparison with Staffing Records

In addition to these center visits, model developers examined sector staffing records as a point of reference for evaluating the PTT estimates. FAA's controller time and attendance system, known as Cru-ART, contains

---

[2] Albuquerque, New Mexico; Boston, Massachusetts; Chicago, Illinois; Indianapolis, Indiana; Jacksonville, Florida; Kansas City, Missouri; Memphis, Tennessee; Miami, Florida; Minneapolis, Minnesota; New York City; Oakland, California; Salt Lake City, Utah; and Washington, D.C.

reports of the number of controllers who timed into a sector during a given time period. CAASD noted to the committee, however, that the Cru-ART data do not always show the number of controllers actually required to work the traffic because some controllers are timed in for training purposes. CAASD therefore looked for opportunities to use the Cru-ART data[3] in ways that would minimize the influence of some of its shortcomings.

Table 4-1 shows the results of a Cru-ART analysis presented to the committee. In this case, CAASD compared the number of controllers recorded on position with the number that would have been estimated using the earlier traffic-volume method and using the task load model's output converted to PTT using the 600-second and fuzzy logic methods. Traffic operation counts were evaluated for each of the nation's 20 en route centers to identify the 90th-percentile traffic days, that is, those days in which traffic volumes were higher than those experienced in 90 percent of the other days during the year. In particular, two days close to the 90th percentile for each center were chosen for the analysis, focusing on the staffing shifts between 7:00 a.m. and 11:00 p.m.[4] Using the Cru-ART records, CAASD calculated the total number of controllers working the traffic during these periods for each center. These totals were then compared with the PTT estimates from the models and task load conversion methods cited earlier.

The results from this analysis show that the previous volume-based method of PTT estimation, which does not consider the complexity of the traffic, consistently overstated the number of controllers required to work the traffic when compared with the Cru-ART records of actual staffing levels. Indeed, in most of the centers, the volume-based methods led to estimates of PTT that were 17 to 70 percent higher than the Cru-ART numbers. By comparison, the 600-second threshold yielded results much closer to the staffing levels indicated by Cru-ART, although the values were mostly lower. The fuzzy logic method came closest to the staffing levels indicated by the Cru-ART records. Because the fuzzy logic model tries to take into account differences in sector traffic complexity and the

---

[3] The Cru-ART data were processed to isolate the "PTT-like" information for each en route center.

[4] Staffing for the midnight shift is often based on factors in addition to PTT needs; thus only 7:00 a.m. to 11:00 p.m. local time was used for the comparative analysis.

TABLE 4-1 Percent Difference in PTT Estimates: Earlier Volume-Based Method for Assessing Staffing Compared with Results from Task Load Model Using Two Alternative Conversion Methods

ARTCC	Traffic Volume-Based Method[a]	CAASD Task Load Model Results Using 600-second Conversion[b]	CAASD Task Load Model Results Using Fuzzy Logic Conversion
ZAB	70	2	7
ZAU	27	–14	–1
ZBW	49	2	10
ZDC	70	17	26
ZDV	68	5	10
ZFW	34	–14	–6
ZHU	34	–15	–5
ZID	36	–6	3
ZJX	47	–10	–1
ZKC	50	–5	0
ZLA	40	–5	7
ZLC	54	–10	–6
ZMA	29	–15	–5
ZME	41	–11	–1
ZMP	34	–18	–8
ZNY	41	–6	9
ZOA	17	–21	–11
ZOB	32	–8	1
ZSE	19	–21	–10
ZTL	38	–10	1

[a] The volume-based method uses simple traffic counts in sectors as the basis for calculating controller workload.
[b] CAASD did not assess the "600-second plus" conversion method.
NOTE: The numbers in the table indicate the percentage difference in positions estimated by each model and conversion method when compared with historical Cru-ART staffing records for the same time period (e.g., a value of 50 means that the model and its conversion method led to a PTT estimate that is 50 percent higher than the number of positions indicated by Cru-ART recorded staffing levels).

resultant impacts on total controller task load, CAASD believes that this is why the latter conversion method yields PTT values that are closer to the Cru-ART numbers.

It merits reiterating, however, that comparing PTT estimates with staffing records is problematic because these records are not necessarily indicative of the staffing that was required to work the experienced traffic. Thus, it not possible to conclude on the basis of this particular analysis that the conversion method using fuzzy logic modeling yields any more accurate predictions of PTT than those yielded by the 600-second conversion method. Indeed, if actual staffing levels (as indicated by Cru-ART) were much too high, the lower PTT values generated by the 600-second conversion may have been more reflective of actual staffing needs.

To be sure, however, the analysis in Table 4-1 brings into question the accuracy of the simple volume-based method for estimating PTT. The PTT values produced through this method are much higher than the staffing numbers in Cru-ART. If one assumes that the controllers staffing the sectors were able to meet their traffic management responsibilities, then clearly these volume-based staffing levels were far too high.

COMMITTEE ASSESSMENT

The CAASD task load model examines only one set of controller tasks: the R-side tasks performed by the lead controller. Because of this limitation, use of the model results to estimate PTT requires either supplemental measures of D-side task load or a creative means of converting the model output into measures of total task load. CAASD decided against obtaining information on D-side task load. Instead, model developers pursued an alternative approach that has relied heavily on expert judgment rather than data. The following recap of this process makes very clear the weakness of this approach.

In its initial efforts to convert the task load output to PTT, model developers consulted with operational experts to estimate when total R-side task load equated to an accompanying amount of D-side task load to warrant a second controller. These consultations, which did not involve any measurement of D-side tasks, apparently led CAASD to conclude that 600 seconds of R-side task load during a 15-minute period was the appropriate threshold. To validate this expert-derived threshold, CAASD again consulted with operational experts to assess the PTT estimates that resulted. The advice from these experts caused CAASD to make further adjustments to the threshold to account for the additional D-side work that accompanies certain kinds of traffic, such as international and nonradar operations. These iterations produced results closer to the expectations of consulted facility managers and controllers.

When other facility managers reviewed the PTT results, they concluded that further modifications were needed to account for even more of the D-side work that the simple conversion thresholds neglected. Accordingly, CAASD introduced the fuzzy logic modeling process. Experts were once again consulted to assign complexity weightings to the different R-side tasks and their combinations. These weightings are intended to characterize the complexity of the D-side task load, even though none of the experts consulted was asked to identify explicitly the D-side tasks involved or to estimate the time it takes to perform each. The PTT values generated from this conversion process were again presented to facility managers for feedback. Their advice led to further adjustments to the fuzzy logic inference rules and complexity weightings until the PTT values satisfied the expectations of the managers consulted.

Both the conversion and validation processes involve repeated consultation with subject matter experts and facility managers and no evidence that data on the performance of D-side tasks were obtained and analyzed to assess their judgments. The heavy reliance on the experience and expectations of facility manager to evaluate the PTT estimation techniques and results is at odds with the purpose of PTT modeling; presumably this purpose is to provide independent quantitative estimates of staffing requirements. All of the PTT conversion methods applied, including the current method of fuzzy logic modeling, exhibit the same fundamental flaw—they imply an estimation of total task load without ever identifying the unmodeled tasks, much less measuring the time it takes to perform them. The conversions rely almost exclusively on experts to determine thresholds and to assign complexity weightings to the unidentified and unmodeled tasks. The D-side task loads implied by these thresholds and weightings are not validated, nor can they be in the absence of any empirical data on task performance.

To adjust these conversion methods further would be insufficient and would risk making the modeling process even less transparent and less convincing. Indeed, it is by no means apparent that past adjustments have led to more accurate PTT predictions—only that they have produced values closer to the expectations of facility managers. In the case of fuzzy logic modeling, this outcome has been achieved at the cost of model transparency and credibility.

5

Findings and Recommendations

The Federal Aviation Administration (FAA) has supported the development of a quantitative model that estimates the task load on controllers created by air traffic activity in each of the more than 750 sectors of the nation's en route airspace. The model uses traffic operations and flight-planning data to simulate the traffic activity in each sector. It then associates with this traffic the specific controller tasks that must be performed, computes and assigns a time to perform each task, and calculates the total time spent by controllers on the modeled tasks. FAA has been using the model's task load output to estimate the number of controllers required to work the traffic in each sector, an estimate known as "positions to traffic" (PTT).

FAA asked the National Academies to convene an expert committee to examine and offer advice where appropriate for improving (*a*) the overall technical approach of task-based modeling, (*b*) input data and processes used for modeling traffic activity, (*c*) tasks and methods used to assign task times, and (*d*) means for validating model assumptions, parameters, and output. In addressing this charge, the committee was asked to be cognizant of the "overall tradeoffs made due to data availability" and to consider the "adaptability of the approach to reflect changes in the tasks of controllers as their roles evolve over time."

Findings from the assessments in the previous chapters are provided next, including those relevant to the use of task load output for estimating PTT. These findings are followed by recommendations for model improvements.

FINDINGS

Task-Based Approach

The results of task-based modeling can be a valuable source of objective information for workforce planning, and FAA's current model is a marked improvement over previous models that did not account for traffic complexity. The basic structure of the CAASD model, in which traffic activity is simulated and controller tasks and task times are associated with traffic, represents a logical approach to estimating task load.

Traffic Modeling

Compared with simple traffic counts, the simulations of traffic in the CAASD model provide a more complete picture of both the volume and nature of traffic activity in the en route domain. The traffic activity is modeled in sufficient depth and resolution to enable reasonable approximations of traffic complexity and associated controller tasks. CAASD can check the model results against records of actual traffic activity to improve the traffic modeling capabilities.

Task Coverage

The nine tasks in the model appear to be representative of the main R-side services that must be performed to work traffic—although whether the specific claim that 90 percent of R-side tasks are covered has not been well established. Compared with the other eight modeled tasks, the monitoring task is treated in the most confusing manner and is difficult to connect with traffic activity. A simpler and more transparent means of estimating monitoring time deserves consideration. While the model's coverage of R-side tasks may be sufficient for traffic capacity analysis, the omission of all tasks performed by the associate controller makes its task load output inadequate for estimating PTT.

Task Time Derivation

Many of the task times in the model are derived from a separate modeling process known as Goals, Operators, Methods, and Selection Rules (GOMS). For seven of the nine modeled tasks, GOMS is used to derive task times. The other task times are developed through consultations with subject matter experts. None of the task times is derived from the observation and analysis of controllers performing tasks in the field or in experiments. The GOMS-derived times are based largely on expert judgment and are only loosely validated against a limited set of task performance data obtained from human-in-the-loop (HITL) experiments conducted for other purposes. The use of GOMS to derive many task times, coupled with reliance on expert judgment for validating these modeled times and for estimating many others, raises serious questions about the accuracy of the model's task load output.

Computation of Task Load

Summing all of the time spent on tasks may be the most practical approach for computing total task load. However, adding one task time to another does not account for the possibility—and real-world probability—that some tasks are performed concurrently. The additive approach also does not account for the possibility that the time it takes to perform a specific task may vary depending on the level of traffic activity and the number of controllers working the sector. Taking these possibilities into account may not have meaningful effects on the modeled task load. Examining their potential effects, however, is important for making this case.

Conversion of Task Load to PTT

FAA and model developers have sought to compensate for the absence of D-side task load by employing various processes that infer total task load to facilitate conversion to PTT. All of the PTT conversion methods used, including the current method of fuzzy logic modeling, exhibit the same fundamental flaw—they imply an understanding of total task load without ever identifying the unmodeled tasks, much less measuring the time it takes to perform them. The conversions rely on experts to determine thresholds and to assign complexity weightings to the unidentified and unmodeled tasks. The D-side task loads implied by these thresholds and weightings are not validated, nor can they be in the absence of any empirical data on task performance. On the whole, the use of the fuzzy logic modeling to infer task load adds little more than spurious precision to the PTT estimates while complicating and reducing the transparency of the modeling process.

Validation

Modeled traffic activity can be checked for accuracy through comparisons with records of actual traffic. In contrast, validating PTT estimates is more challenging since there is no external measure of staffing requirements against which the accuracy of the estimates can be judged. Analyzing staffing records is of limited value since the main purpose of PTT modeling is to find out when staffing levels can be better aligned with traffic demand. The main method by which model developers have sought to assess PTT estimates is by presenting them to groups of experts, often consisting of individuals who manage and staff the en route centers. Yet such checks can suffer from the same shortcoming that limits the value of comparisons with staffing records—the potential for bias toward existing staffing practice.

Because PTT estimates cannot be assessed through direct observation, all of the model's key assumptions, processes, and parameters must be well justified and validated. A lack of data on task performance precludes validation of the task times constructed from GOMS and the task complexity weightings used in the fuzzy logic conversion method. The deficiencies of these two modeling processes go well beyond parameter validation, as explained earlier. Yet the lack of empirical data on task performance has hindered validation throughout the modeling process, from assessing key assumptions about tasks being performed sequentially and at a fixed pace to characterizing the tasks handled by the associate controller.

Data Availability and Model Adaptability

In the study charge, FAA asked the committee to be cognizant of tradeoffs that must be made because of limited data availability, which presumably refers to the cost and complications of obtaining task performance data. FAA also asked for advice on the model's adaptability to reflect changes in controller roles and tasks over time.

Many of the findings cited earlier point to a need for a firmer empirical basis both for evaluating the structure of the model and for estimating the values of the parameters used in it. By and large, the model was developed and has been evaluated with heavy reliance on the insights and opinions obtained from subject matter experts and facility personnel. More objective and quantitative task performance data are clearly needed, not only for developing the model parameters and evaluating the task load output but also for including more controller tasks in the modeling of PTT. The committee recognizes that gathering such data from operational and experimental settings will require more resources and access to controllers, which may present budget and labor relations issues. Although such cost implications were not examined in this study, it must be pointed out that there is a cost in model credibility from not obtaining such data. This cost is manifested in many ways throughout the current model, from the added opaqueness caused by fuzzy logic modeling to the excessive reliance on expert opinion and judgment for model development and validation.

Whether FAA is committed to taking this data-gathering step will presumably depend on its assessment of the cost trade-offs and its plans for using the model for a long time and for other possible purposes. Not knowing these plans, the committee nevertheless believes FAA would not have asked for this review absent a strong interest in improving its modeling capabilities. It is in this context that the committee wishes to express its strong view that the current model is deficient for estimating PTT and that continuing to iterate on it in the same manner as in the past while not incorporating more complete and representative task performance data will do little to improve this situation.

Looking farther out, the durability of the task load model for PTT analysis and for other possible applications, such as to inform traffic flow planning, will depend not only on the successful gathering and use of task performance data but also on the nature and pace of change in the air traffic control enterprise. Developments anticipated for the planned Next Generation Air Transportation System (NextGen), such as increased automation and many more decision-support tools, could substantially alter controller roles and responsibilities in ways that are highly relevant to the modeling of PTT. Without more knowledge about the nature and timing of these NextGen changes, it is not possible to predict how the model will hold up structurally, much less how changes in traffic data, task coverage, and task times might make it more adaptable.

RECOMMENDATIONS

In commencing its review, the committee expected to find—but did not—strong documentation explaining the logic and structure of the model and evidence of its having been the subject of statistical tests and other scientific methods for establishing and validating model parameters, assumptions, and output. More rigorous documentation and peer review during earlier stages of model development would likely have exposed many of the problems identified in this report, providing earlier opportunities to avoid or correct them. Nevertheless, as preface to offering advice on ways to improve the modeling process going forward, it is important to restate the finding that the current model framework, despite the data shortcomings, represents a major improvement over past modeling methods to inform workforce planning. In the following recommendations it is presumed that FAA will elect to retain the core model and invest meaningfully in its improvement.

Observe and Measure Controller Task Performance

Through more systematic and carefully designed observation and analysis of controller performance, model developers should gain a better understanding of the tasks that controllers perform in working en route traffic, how they perform them, and the time required to do so. The gathering and analysis of data on controllers working alone and interacting in teams, whether through field observations or HITL experiments, should be the primary method to identify and elicit information on controller tasks.

Model All Controller Tasks

Modeling all tasks that contribute significantly to total controller task load is fundamental for estimating PTT. FAA should use the information gained from observing, measuring, and analyzing controller task performance to quantify the task load associated with the services provided by both the lead and associate controllers. The modeling of all controller tasks will eliminate the need to infer task load to derive estimates of PTT. Using a single model for estimating task load rather than separate ones for each controller is the preferred approach, since it will facilitate both PTT conversion and model validation.

Validate Model Elements

Task performance data should be used also to assess the validity and impact of all key modeling processes, relationships, and assumptions. Because it is not possible to validate PTT estimates against actual staffing levels, ensuring that the model elements are well justified and viewed as credible is vitally important. Examples of modeling assumptions that would seem to warrant early attention are those that concern task performance by the controllers when working alone and in teams, whether tasks are performed sequentially or concurrently, and how total task load affects the pace of task performance.

Study Committee Biographical Information

R. John Hansman, Jr., *Chair,* is Professor of Aeronautics and Astronautics, Head of the Division of Humans and Automation, and Director of the International Center for Air Transportation at Massachusetts Institute of Technology (MIT). In addition to teaching, he conducts research in several areas related to air transportation, flight vehicle operations, and safety. His current research activities focus on information technology applied to air transportation systems, air traffic control, integrated human–automation systems, advanced vehicles, and advanced cockpit information systems. He is also an internationally recognized expert on aviation meteorological hazards such as icing and wind shear. He is a member of the Aeronautics and Space Engineering Board (ASEB) of the National Research Council (NRC) and serves on the ASEB Committee for the Review of NASA's Aviation Safety-Related Programs. He has previously served on the NRC Committee to Identify Potential Breakthrough Technologies and Assess Long-Term R&D Goals in Aeronautics and Space Transportation Technology and the Committee on the Effects of Aircraft–Pilot Coupling on Flight Safety. He holds a doctorate in physics, meteorology, and aeronautics from MIT, an SM in physics from MIT, and an AB from Cornell University.

Monica S. Alcabin is Associate Technical Fellow in the Air Traffic Management Unit of Boeing Commercial Airplanes, where she studies the costs and benefits of new air traffic control technologies and the challenges of integrating new technologies into today's air traffic control system. She has 25 years of experience analyzing a variety of aviation problems with particular emphasis on the benefit assessment of airport, airspace, and air traffic management operational enhancements.

She has been assisting the Joint Planning and Development Office (JPDO) Evaluation and Analysis Division by assessing the airport capacity benefits of the Next Generation Air Transportation System. Prior to joining the Boeing Company in 1997, she spent five years in airport consulting at KPMG Peat Marwick and TRA/Black & Veatch, four years at The MITRE Corporation supporting the FAA Office of System Capacity and Requirements, and five years at NASA Ames Research Center doing research in air traffic control. She is currently working on developing avionics and ground infrastructure requirements for conducting independent parallel approaches to closely spaced parallel runways. She earned a BS in aeronautical engineering from MIT and an MS in engineering–economic systems from Stanford University.

Michael O. Ball is Orkand Corporation Professor of Management Science in the Robert H. Smith School of Business at the University of Maryland. He also holds a joint appointment within the Institute for Systems Research in the Clark School of Engineering. He is currently Director of Research for the Smith School and is former chair of the Department of Decision, Operations, and Information Technologies. His research interests are in network optimization and integer programming, particularly as applied to problems in transportation systems and supply chain management. He is Co-Director of the National Center of Excellence for Aviation Operations Research (NEXTOR), and he leads the NEXTOR Collaborative Decision Making Project. He is, or has been, associate editor for *Operations Research, Transportation Science, IIE Transactions, IEEE Transactions on Reliability,* and *Operations Research Letters and Networks.* He is currently area editor for transportation for *Operations Research.* He received a doctorate in operations research from Cornell University.

Mary L. Cummings is the Boeing Associate Professor and Director of the Humans and Automation Laboratory in the Department of Aeronautics and Astronautics of MIT. She performs research in collaborative human-computer decision making for command and control domains and is a recognized expert in the area of human supervisory control. She served as a naval officer from 1988 to 1999 and was among the first female fighter pilots in the U.S. Navy. She is a member of the NRC Committee on

Human Systems Integration, and has also served as a member of the NRC Committee on Opportunities in Neuroscience for Future Army Applications. She is an Associate Editor for the *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans,* and is on the editorial board for the *Human Factors* journal. She earned a BS from the U.S. Naval Academy, an MS from the U.S. Naval Postgraduate School, and a PhD from the University of Virginia.

William J. Dunlay is retired Director of Jacobs Consultancy, formerly Leigh Fisher Associates. He has more than 40 years of experience in transportation engineering and airport planning, having directed airfield and airspace studies for more than 40 airports in the United States and overseas. He managed the airfield–airspace elements of airport master plans for many large airports around the country, including Cleveland Hopkins, Dallas–Fort Worth, Las Vegas McCarran, Orlando, Lambert–St. Louis, and Washington Dulles International Airports. In 2003 and 2004, he divided his time between Jacobs Consultancy and the University of California (UC), Berkeley, returning to full-time status with the company in 2005. As a UC Berkeley research engineer, he played a key role on the Virtual Airspace Modeling and Simulation (VAMS) Project at the National Aeronautics and Space Administration (NASA), investigating various concepts for the Next Generation Air Transportation System. He served on ASEB's Panel on Airspace Systems Program Committee for the Review of NASA's Revolutionize Aviation Program. He holds a PhD in civil engineering from UC Berkeley and an MS and BS in civil engineering from Pennsylvania State University.

Antonio L. Elias (NAE 2001) is Executive Vice President and General Manager of Orbital Sciences Corporation. His recent professional activity has centered on the design, development, and manufacture of orbital launch vehicles and the overall architecture of a space transportation system. He joined Orbital Sciences Corporation in 1986 as Chief Engineer and later became Vice President of Engineering, Senior Vice President of the Space Systems Division, Senior Vice President for Advanced Projects, and Senior Vice President and Chief Technical Officer. From 1980 to 1986, he was an Assistant Professor of Aeronautics and Astronautics at MIT. He is a private pilot and has maintained an interest in air transportation and air traffic control automation. Dr. Elias was elected to the National Academy of Engineering in 2001. He holds a PhD in flight transportation, an MS in flight dynamics and control, and a BS in aeronautical engineering, all from MIT.

John J. Fearnsides is Partner and Chief Strategist, MJF Strategies, LLC. Until 1999, he was Vice President and General Manager of the MITRE Corporation and Director of its Center for Advanced Aviation System Development. He worked at the U.S. Department of Transportation from 1972 to 1980, serving as Deputy Under Secretary and Chief Scientist, Executive Assistant to the Secretary, and Acting Assistant Secretary for Policy and International Affairs. He was a National Science Foundation Fellow and is a Fellow of the Institute of Electrical and Electronics Engineers and the National Academy of Public Administration. He has served on numerous NRC and TRB committees, including the Committee for a Review of the National Automated Highway System Consortium Research Program and Committee for a Study on Air Passenger Service and Safety Since Deregulation. Dr. Fearnsides holds a BS and a PhD in electrical engineering from the University of Maryland.

J. Victor Lebacqz was the Associate Administrator for Aeronautics Research at NASA from 2002 to 2005. In this position, he had overall technical, programmatic, and personnel management responsibility for all aeronautics technology research and development within the agency. During 2006, he was a Research Fellow at the University of California, Santa Cruz. Soon after, he founded VICC Associates, which specializes in executive consulting for aviation and other technology organizations. Earlier in his career, he worked at the NASA Ames Research Center, specializing in avionics, stability and control, handling qualities, and human factors. He held management positions of increasing responsibility as Branch Chief, Division Chief, Program Manager, Deputy Director of Aerospace, and Associate Center Director. He is currently a member of FAA's Research Engineering and Development Advisory Committee (REDAC), chairing a subcommittee on the National Airspace System (NAS) Operations. He is a member of the Editorial Board of *Air Traffic Control Quarterly* and was a member of a National Academy of Public Administration Panel to Assess FAA Program Management Capabilities. He holds a PhD in aeronautical engineering from Princeton University.

Michael J. Powderly is President of Airspace Solutions, a private consulting firm specializing in airport, airspace, and air traffic capacity and efficiency projects. From 1995 to 2000, he was General Manager of Airspace Capacity and Efficiency for Delta Air Lines. In this position, he directed staff responsible for the airline's global communications, navigation, surveillance, and air traffic management activities. These duties required a strong understanding of FAA air traffic control procedures and policy positions. From 1967 to 1995, he was employed by FAA, starting his career as a controller. From 1976 to 1986 he was Chief Controller of the Atlanta Air Traffic Control Tower. From 1986 to 1995 he managed the Southern Region Air Traffic Division Branches, including procedures, operations, and evaluations. He oversaw the administration and operations of 113 control towers, 5 en route facilities, 8 automated flight service stations, and nearly 6,000 air traffic controllers. During his career he received numerous honors and awards, including the Air Traffic Control Association Quasada Award for the Advancement of Air Traffic Control, the U.S. Department of Transportation's Administrator Silver Medal for Excellence, and an RTCA, Inc., citation for leadership on surveillance and air traffic management modernization.

Philip J. Smith is Co-Director of the Institute for Ergonomics and Professor in the Department of Industrial and Systems Engineering and a Professor with the Industrial and Systems Engineering Program, Biomedical Engineering, and the Advanced Computing Center for Arts and Design at Ohio State University. He teaches courses in areas of cognitive systems engineering, artificial intelligence, human–computer interaction and the design of cooperative problem-solving systems, intelligent information retrieval systems, and intelligent tutoring systems. His research focuses on issues concerned with design of cooperative problem-solving systems to support people in performing complex tasks such as information retrieval, planning, database exploration, and diagnosis while using fields such as aviation, medicine, library systems, and education as test beds. He is a member of the NRC Committee on Human-Systems Integration and served on the Committee to Study the FAA's Methodologies for Estimating Air Traffic Controller Staffing Standards. He earned a PhD in cognitive psychology and industrial and operations engineering from the University of Michigan.

Antonio A. Trani is Professor of Civil and Environmental Engineering at the Virginia Polytechnic Institute and State University, where he specializes in air transportation simulation and modeling and airport engineering systems. His current research centers on aviation demand modeling for the Next Generation Air Transportation System and airspace and airfield simulation modeling. He holds a BS in aeronautical engineering from Embry-Riddle Aeronautical University and an MS in systems engineering and a PhD in transportation engineering from the Virginia Polytechnic Institute and State University.

Roger Wall is retired FAA Coordinator and ATM Projects Manager for Federal Express Corporation. Before joining FedEx, he was Director of Air Traffic Operations for FAA, having risen from air traffic controller. At FAA, he held management positions at air traffic control facilities, FAA regional offices, and FAA headquarters. He began his career as an air traffic controller for the U.S. Navy in 1959. From 1996 to 2008, he served as chairman of the Free Flight Select Committee of the Radio Technical Commission for Aviation (RTCA) and the Requirements and Planning Working Group of the Air Traffic Management and Airport System. He was honored with RTCA's Lifetime Achievement Award in 2008.

Greg L. Zacharias is President and Senior Principal Scientist of Charles River Analytics. In this position he provides strategic direction for the Government Services and Commercial Solutions Divisions while contributing to efforts in cognitive systems engineering and advanced decision-support systems. Before founding Charles River Analytics, he was a Senior Scientist at BBN Technologies, a Research Engineer at CS Draper Labs, and a U.S. Air Force attaché for the Space Shuttle program at NASA's Johnson Space Center. He serves on the U.S. Department of Defense Human Systems Area Review and Assessment Panel. He has been a member of the NRC Committee on Human Factors and co-chaired the Committee on Organizational Models: from Individuals to Societies. He has a PhD in aeronautics and astronautics from MIT.

TRANSPORTATION RESEARCH BOARD
2010 EXECUTIVE COMMITTEE*

Chair: Michael R. Morris, Director of Transportation, North Central Texas Council of Governments, Arlington
Vice Chair: Neil J. Pedersen, Administrator, Maryland State Highway Administration, Baltimore
Executive Director: Robert E. Skinner, Jr., Transportation Research Board

J. Barry Barker, Executive Director, Transit Authority of River City, Louisville, Kentucky
Allen D. Biehler, Secretary, Pennsylvania Department of Transportation, Harrisburg
Larry L. Brown, Sr., Executive Director, Mississippi Department of Transportation, Jackson
Deborah H. Butler, Executive Vice President, Planning, and CIO, Norfolk Southern Corporation, Norfolk, Virginia
William A. V. Clark, Professor, Department of Geography, University of California, Los Angeles
Eugene A. Conti, Jr., Secretary of Transportation, North Carolina Department of Transportation, Raleigh
Nicholas J. Garber, Henry L. Kinnier Professor, Department of Civil Engineering, and Director, Center for Transportation Studies, University of Virginia, Charlottesville
Jeffrey W. Hamiel, Executive Director, Metropolitan Airports Commission, Minneapolis, Minnesota
Paula J. Hammond, Secretary, Washington State Department of Transportation, Olympia
Edward A. (Ned) Helme, President, Center for Clean Air Policy, Washington, D.C.
Adib K. Kanafani, Cahill Professor of Civil Engineering, University of California, Berkeley (Past Chair, 2009)
Susan Martinovich, Director, Nevada Department of Transportation, Carson City
Debra L. Miller, Secretary, Kansas Department of Transportation, Topeka (Past Chair, 2008)
Sandra Rosenbloom, Professor of Planning, University of Arizona, Tucson
Tracy L. Rosser, Vice President, Corporate Traffic, Wal-Mart Stores, Inc., Mandeville, Louisiana
Steven T. Scalzo, Chief Operating Officer, Marine Resources Group, Seattle, Washington
Henry G. (Gerry) Schwartz, Jr., Chairman (retired), Jacobs/Sverdrup Civil, Inc., St. Louis, Missouri
Beverly A. Scott, General Manager and Chief Executive Officer, Metropolitan Atlanta Rapid Transit Authority, Atlanta, Georgia
David Seltzer, Principal, Mercator Advisors LLC, Philadelphia, Pennsylvania
Daniel Sperling, Professor of Civil Engineering and Environmental Science and Policy; Director, Institute of Transportation Studies; and Interim Director, Energy Efficiency Center, University of California, Davis
Kirk T. Steudle, Director, Michigan Department of Transportation, Lansing
Douglas W. Stotlar, President and Chief Executive Officer, Con-Way, Inc., Ann Arbor, Michigan
C. Michael Walton, Ernest H. Cockrell Centennial Chair in Engineering, University of Texas, Austin (Past Chair, 1991)

Peter H. Appel, Administrator, Research and Innovative Technology Administration, U.S. Department of Transportation (ex officio)
J. Randolph Babbitt, Administrator, Federal Aviation Administration, U.S. Department of Transportation (ex officio)
Rebecca M. Brewster, President and COO, American Transportation Research Institute, Smyrna, Georgia (ex officio)
George Bugliarello, President Emeritus and University Professor, Polytechnic Institute of New York University, Brooklyn; Foreign Secretary, National Academy of Engineering, Washington, D.C. (ex officio)
Anne S. Ferro, Administrator, Federal Motor Carrier Safety Administration, U.S. Department of Transportation (ex officio)
LeRoy Gishi, Chief, Division of Transportation, Bureau of Indian Affairs, U.S. Department of the Interior, Washington, D.C. (ex officio)
Edward R. Hamberger, President and CEO, Association of American Railroads, Washington, D.C. (ex officio)
John C. Horsley, Executive Director, American Association of State Highway and Transportation Officials, Washington, D.C. (ex officio)
David T. Matsuda, Deputy Administrator, Maritime Administration, U.S. Department of Transportation (ex officio)
Victor M. Mendez, Administrator, Federal Highway Administration, U.S. Department of Transportation (ex officio)
William W. Millar, President, American Public Transportation Association, Washington, D.C. (ex officio) (Past Chair, 1992)
Tara O'Toole, Under Secretary for Science and Technology, U.S. Department of Homeland Security (ex officio)
Robert J. Papp (Adm., U.S. Coast Guard), Commandant, U.S. Coast Guard, U.S. Department of Homeland Security (ex officio)
Cynthia L. Quarterman, Administrator, Pipeline and Hazardous Materials Safety Administration, U.S. Department of Transportation (ex officio)
Peter M. Rogoff, Administrator, Federal Transit Administration, U.S. Department of Transportation (ex officio)
David L. Strickland, Administrator, National Highway Traffic Safety Administration, U.S. Department of Transportation (ex officio)
Joseph C. Szabo, Administrator, Federal Railroad Administration, U.S. Department of Transportation (ex officio)
Polly Trottenberg, Assistant Secretary for Transportation Policy, U.S. Department of Transportation (ex officio)
Robert L. Van Antwerp (Lt. General, U.S. Army), Chief of Engineers and Commanding General, U.S. Army Corps of Engineers, Washington, D.C. (ex officio)

---

*Membership as of November 2010.

Copyright National Academy of Sciences. All rights reserved.