



**Department of Transportation
Federal Aviation Administration**

Finding of No Significant Impact (FONSI)

and

Record of Decision (ROD)

**for the Houston Optimization of
Airspace and Procedures
in the Metroplex (OAPM)**

June 2013

TABLE OF CONTENTS

<u>Section</u>	<u>Page</u>
Introduction.....	3
Background.....	4
Proposed Action.....	5
Purpose and Need.....	6
Alternatives.....	6
Study Areas and Affected Environment.....	8
Environmental Impact Categories and Consequences.....	9
Cumulative Impacts.....	14
Agency and Public Involvement.....	14
Comments and FAA Responses.....	15
Mitigation.....	15
Other Considerations.....	15
Finding of No Significant Impact.....	16
Decision.....	16
Right of Appeal.....	16
Attachments:	
- Houston Optimization of Airspace and Procedures in the Metroplex, Houston Texas (EA), January 2013	
- Concurrence Letters, Comments Received, and Responses to Comments	
- Errata Sheet	

**DEPARTMENT OF TRANSPORTATION
FEDERAL AVIATION ADMINISTRATION**

**Finding of No Significant Impact
and Record of Decision for the
Houston Optimization of Airspace and Procedures in the Metroplex**

Introduction

This Finding of No Significant Impact and Record of Decision (FONSI/ROD) is being issued for the *Environmental Assessment for Houston Optimization of Airspace and Procedures in the Metroplex*, January 2013, attached hereto and incorporated by reference. This FONSI/ROD has been prepared in compliance with the National Environmental Policy Act of 1969, as amended; implementing regulations issued by the Council on Environmental Quality (40 Code of Federal Regulations (CFR), parts 1500-1508); and FAA Order 1050.1E, *Environmental Impacts: Policies and Procedures*, effective March 20, 2006. This FONSI/ROD:

- Documents the FAA's compliance with the several procedural and substantive requirements of aeronautical, environmental, programmatic, and other statutes and regulations that apply to FAA decisions and actions which are based on the environmental analysis and findings in the attached EA and all other applicable documents which were available and considered, and which constitute the administrative record.
- Documents the FAA's finding that the Houston Optimization of Airspace and Procedures in the Metroplex (OAPM) will not have significant environmental impacts and explains the basis for that finding.
- Approves the actions necessary to implement the Houston OAPM. The approved actions do not include any airport-related development, land acquisition, construction, or other ground disturbance activities.

In approving the Houston OAPM, the FAA has considered 49 U.S.C. 40101(d)(4), which gives the FAA various responsibilities and holds it accountable for controlling the use of navigable airspace and regulating civil and military operations in that airspace in the interest of safety and efficiency of both of these operations. Additionally, consideration has been given to 49 U.S.C. 40103(b)(2) which authorizes and directs the FAA Administrator to prescribe air traffic regulations governing the flight of aircraft (including regulations on safe altitudes) for the navigation, protection, and identification of aircraft; the protection of persons and property on the ground; the efficient utilization of the navigable airspace; and the prevention of collisions between aircraft, between aircraft and land or water vehicles, and between aircraft and airborne objects.

Furthermore, FAA has given careful consideration to: the aviation safety and operational objectives of the Houston OAPM in light of the various aeronautical factors and judgments presented; the need to enhance the efficiency of the national air transportation system; and the anticipated environmental impacts of the project.

The Houston OAPM centers on the area's two busiest airports:

- George Bush Intercontinental/Houston, Houston, TX (IAH)
- William P. Hobby, Houston, TX (HOU)

The Houston OAPM Study Area also includes the following satellite airports:

- David Wayne Hooks Memorial, Houston, TX (DWH)
- Ellington Field, Houston, TX (EFD)
- Lone Star Executive, Houston, TX (CXO)
- Sugar Land Regional, Houston, TX (SGR)
- Scholes International at Galveston, Galveston, TX (GLS)
- West Houston, Houston, TX (IWS)
- Houston Executive, Houston, TX (TME)
- Houston-Southwest, Houston, TX (AXH)
- Texas Gulf Coast Regional, Angleton/Lake Jackson, TX (LBX)
- Pearland Regional, Houston, TX (LVJ)
- Chambers County, Anahuac, TX (T00)
- La Porte Municipal, La Porte, TX (T41)
- RWJ Airpark, Baytown, TX (54T)
- Weiser Air Park, Houston, TX (EYQ)
- Baytown, Baytown, TX (HPY)

The EA refers to the two primary and 15 satellite airports collectively as the Houston OAPM Airports.

Background

The FAA is in the process of implementing the Next Generation Air Transportation System (NextGen), the FAA's plan to modernize the National Airspace System (NAS) through 2025. NextGen is a complex program intended to develop and implement new technologies, while integrating existing technologies and adapting the air traffic management system to a new way of operating. NextGen represents an evolution from an air traffic control system that is primarily ground-based to a system that is satellite-based, and will allow the FAA to guide and track air traffic more precisely and efficiently. To achieve NextGen goals, the FAA is implementing new Area Navigation (RNAV) and Required Navigation Performance (RNP) air traffic routes and instrument procedures (RNAV Standard Instrument Departures (SIDs) and Standard Terminal Arrival Routes (STARs), and Standard Instrument Approach Procedures (SIAPs)) around the country that leverage emerging technologies and aircraft navigation capabilities. The implementation of RNAV and RNP enable the use of other Performance Based Navigation (PBN) technology in the National Airspace System (NAS), and would facilitate more efficient procedures such as Optimized Profile Descents (OPD).

The Department of Transportation selected the Houston OAPM as a high priority infrastructure project for inclusion on the Federal Infrastructure Projects Dashboard ("Dashboard").¹ Established pursuant to

¹ The Federal Infrastructure Projects Dashboard is publicly available at <http://permits.performance.gov>

an August 2011 Presidential Memorandum,² the Dashboard is part of an inter-agency initiative, spearheaded by the Office of Management and Budget, to institutionalize best practices to reduce the amount of time required to make permitting and review decisions, and improve environmental and community outcomes. The Federal government is enhancing and expanding the Dashboard to serve as a government-wide tool to enable and support collaboration within and among Federal agencies, as well as to provide increased public transparency regarding the schedules and status of nationally or regionally significant projects, permitting timelines, and overall Federal infrastructure project permitting and review processes. The Houston OAPM is the first aviation project selected for inclusion on the Dashboard; other projects include green infrastructure, surface transportation, renewable energy, community development, and electricity transmission.

Proposed Action

The Proposed Action consists of development of air traffic control and airspace management procedures to establish and maintain safe and efficient handling and movement of traffic into and out of the Houston, Texas, Metroplex airspace. A Metroplex is one or more busy airports surrounded by complex airspace. This action would:

- establish 20 new RNAV-SIDs (departure routes) from the region
- establish 20 new RNAV STARs (arrival routes)
- establish 5 new conventional (i.e. non-RNAV) STARs³
- modify 4 existing conventional STARs
- establish 4 new RNP Authorization Required (AR) approaches
- modify 2 existing RNP AR approaches for George Bush Intercontinental Airport (IAH)
- modify 6 existing Instrument Landing System (ILS) approaches by adding RNAV transitions

Nineteen (19) existing procedures (SIDs and STARs) would be cancelled, and eleven (11) existing SIDs would be retained unchanged.

The Proposed Action would: (1) improve operational efficiency through use of new PBN procedures; (2) increase flight path predictability; and (3) decrease required pilot-controller voice communication. In some cases, PBN routes that mirror the existing flight paths over the ground would replace standard routings achieved currently through radar vectoring. This would typically result in shorter and more predictable routes as compared to current published routes. The new PBN procedures would also provide vertical navigation, allowing the aircraft to descend from cruise altitude into the airport area with reduced pilot-controller communications and fewer inefficient level flight segments. Additionally, modifications to routes that interact with adjacent Fort Worth Air Route Traffic Control Center (ZFW ARTCC) would improve integration with ZFW procedures. Finally, certain procedures would change in order to better align routes and profiles for international flights to Mexico and South America. Chapter

² White House, “Presidential Memorandum – Speeding Infrastructure Development through More Efficient and Effective Permitting and Environmental Review,” August 31, 2011.

³ In the EA and this FONSI/ROD, the terms “non-RNAV” and “conventional” are used interchangeably unless otherwise noted.

3 and Appendix F of the EA provide detail on the proposed alterations, deletions, or additions to each procedure associated with the Proposed Action.

Implementation of the Proposed Action would not require any ground disturbance or development of facilities, nor would it require local or state action. The Proposed Action consists only of procedural changes intended to improve operational efficiency, increase flight path predictability, and reduce required controller-pilot voice communication. Therefore, it would not increase the number of aircraft operations within southeast Texas airspace when compared to the No Action Alternative.

The target date for publication of the Houston Metroplex optimized procedures is December 12, 2013.

Purpose and Need

The Houston OAPM project consisted of a Study Team phase, which analyzed Metroplex operational challenges and situations and explored opportunities to optimize airspace, followed by a Design and Implementation (D&I) Team phase, which developed the Proposed Action. The Study Team concluded that existing published (charted) air traffic procedures in the Houston Metroplex are inefficient, inflexible, and overly complex compared to what recent advances in technology would allow.

The Study Team materials reflect three key factors as causes of inefficiencies and complexities in the Houston Metroplex:

- Limitations of the conventional, ground-based navigation system and existing RNAV procedures
- Limited flight path predictability and flexibility, particularly during adverse weather conditions
- High occurrence of voice communications among controllers and pilots, leading to excessive workload, and increased hear-back and read-back errors

These three factors demonstrate the need for the Proposed Action.

The purpose of the Proposed Action is to address the three components of the need, as described above. The FAA's primary drivers are improved efficiency of airspace operations, increased flight path predictability and flexibility, and decreased errors in controller/pilot voice communication, along with preserved or improved air traffic safety. In order to address the need, the FAA intends to implement readily available NextGen technologies designed to support these types of improvements.

Alternatives

Identification and Evaluation of Potential Alternatives - The Houston OAPM Study Team and Design & Implementation (D&I) Team each identified and evaluated potential alternatives to individual procedures. Collectively, the final set of proposed changes to instrument flight procedures (IFPs), became the Proposed Action. Implementation of the Proposed Action would (1) improve operational efficiency through use of PBN procedures, (2) increase flight path predictability, and (3) decrease required pilot-controller voice communication.

The Study Team convened in May 2011 to define operational issues in the Houston Metroplex and identify potential corresponding solutions. Work by the Study Team served to guide later detailed design efforts and inform the FAA decision-making processes related to these efforts. During three sets of outreach meetings, the Study Team obtained input from air traffic control experts, airspace

users, and industry representatives. These meetings helped identify existing operational challenges, enhancement opportunities, and evaluation metrics. Initially, 105 issues were identified, after which similar issues were grouped to determine potential solution sets. Considerations included reviewing level flight segments and flight profiles. During a descent or climb profile, level flight segments are less efficient. Therefore, the team emphasized optimization of aircraft climb and descent profiles for the various procedures. The Study Team identified several potential modifications to each arrival/departure procedure that addressed issues identified in the outreach meetings. The Study Team rejected or modified several of the initial proposals because, on further analysis, they would not address the need for the project or would adversely affect existing operations.

The Study Team also worked closely with environmental specialists to consider whether any of the proposed solutions might create an environmental impact and to adjust those proposed solutions as necessary. The Study Team considered the Alabama-Coushatta Tribe of Texas Reservation, as well as resources protected under Section 4(f) of the Department of Transportation (DOT) Act (including the Big Thicket National Preserve, the Sam Houston National Forest, and the Anahuac and Trinity River National Wildlife Refuges). The result of the initial screening by the Study Team indicated minimal risk for significant environmental impacts to these lands. The Study Team recommendations became the basis for the initial set of procedures evaluated by the D&I Team.

The Houston OAPM Design and Implementation (D&I) Team, comprised of FAA and industry personnel, convened in January 2012 to review the procedures recommended by the Study Team. The D&I Team carefully considered the Study Team recommendations and made numerous modifications and improvements to those recommendations. The D&I Team adopted, refined, rejected, and added to the proposal elements recommended by the Study Team. In some instances, design concerns or other issues precluded the development of procedures as originally envisioned by the Study Team. As the D&I Team analyzed changes to individual procedures, and their associated interactions between procedures, it elected not to carry some changes forward because they did not meet FAA design or safety criteria, and/or the purpose and need of this project. This evaluation was an iterative process, as modifying one procedure had potential to affect one or more other procedures.

The D&I Team engaged airspace users and environmental specialists regularly for feedback throughout their deliberations. During this complex iterative process, the D&I Team considered various environmental factors based on use and location of routes. The following are some examples of consideration of environmental factors by the D&I Team during the development process:

- Modification of initial Study Team recommendations that would have increased runway use on IAH Runway 26R. With the modification, arrival use of Runways 26L, 26R, and 27 would not be expected to change.
- Change of the initial design of a proposed airway to direct aircraft away from Big Thicket National Preserve.
- Revision of proposed departure procedures for IAH Runways 15L/R to minimize changes in noise exposure.

The Study Team recommended use of procedural deconfliction⁴ (or procedural separation) where practical. The D&I Team looked at each procedure individually and considered the benefits both gained and lost due to the use of procedural deconfliction in developing PBN procedures. When an operational advantage seemed likely, the D&I team employed procedural deconfliction. In some cases, though, burdens imposed on airspace users would have outweighed the operational advantage of procedural deconfliction. For example, where two routes intersect at the same altitude, procedural deconfliction would require one aircraft to take a less efficient routing or altitude in order to maintain adequate separation.

In addition to standard design considerations (e.g., aircraft performance capabilities, airfield layout and runway geometry, and locations of satellite airports), other factors specific to the Houston Metroplex also influenced the design. Two such considerations included the proximity of the two primary airports in the region, IAH and HOU, and the presence and location of Special Use Airspace (SUA) for military operations. The D&I Team also accounted for other operational factors, including preferred runways during fluctuating wind conditions.

Alternatives Analyzed in the EA - In addition to the Proposed Action (described above), the EA also analyzed the No Action Alternative. Under the No Action Alternative, the FAA would maintain the existing arrival and departure procedures in the Houston Metroplex. There are currently 34 published SIDs and STARs in the Houston Metroplex, serving the Houston OAPM Airports.

- 4 RNAV SIDs
- 11 RNAV STARs
- 12 conventional (i.e., non-RNAV) SIDs
- 7 conventional (i.e., non-RNAV) STARs

The No Action Alternative would include expected future actions that are independent of the Houston OAPM process. The existing conventional and RNAV arrival and departure procedures would remain as is, subject to minor, periodic reviews and revisions in response to changes in the operational environment. The No Action Alternative would not implement the specific procedures designed as part of the Houston OAPM Project.

Study Areas and Affected Environment

The Primary Study Area (PSA) was defined to allow for a reasonable evaluation of potential impacts associated with jet aircraft noise under the Proposed Action. To determine the study area boundaries for this EA, the FAA considered the geographic areas where new or revised aircraft routings under the Proposed Action would differ from the No Action Alternative. The FAA evaluated the existing flight paths in the southeast Texas region as a basis to determine where Proposed Action changes are likely to occur. Initially, the FAA collected radar data for arrival and departure operations from airports in the southeast Texas region for periods during 2010-2011, focusing on aircraft traffic controlled by the Houston ARTCC (ZHU) and the Houston Terminal Radar Approach Control Facility (I90 TRACON). The ZHU ARTCC is located at IAH and controls airspace in southern Texas, Louisiana, southern Mississippi,

⁴ “Procedural deconfliction” means defining mandatory altitude or lateral restrictions as part of a procedure to keep aircraft from conflicting with others on routes in close proximity.

southwestern Alabama, and areas in the Gulf of Mexico. The I90 TRACON provides approach control for airports within the Houston Metroplex.

To determine the Study Area, the FAA analyzed radar data for approaching and departing Instrument Flight Rule (IFR) Houston Metroplex jet aircraft. A review of radar data showed that approximately 95 percent of Houston Metroplex jet aircraft operations below 10,000 ft. above ground level (AGL) occur within 50 nautical miles (NM) of a point midway between IAH and HOU. This became the PSA.⁵ Development of the Supplemental Study Area (SSA) applied the same methodology as was used for the PSA. Approximately 95 percent of IFR Houston Metroplex jet aircraft operations below 18,000 ft. AGL occur within 85 NM of a point midway between IAH and HOU.⁶ The PSA encompasses approximately 7,850 square nautical miles (NM²) and 19 counties in part or whole (all within Texas). The SSA, including the PSA, encompasses approximately 22,700 NM² spanning 16 additional counties and 2 Louisiana parishes in part or whole.

The geographic area for each impact category can vary, sometimes coinciding with the study area, but more often constituting a smaller area within the study area. The relevant geographic area can, and often does, differ from one impact category to another (e.g., Flood Plain area can be different from Air Quality area). The effects on some environmental impact categories are highly localized and confined, whereas others cover a broader geographic area. Focusing on specific geographic areas for each environmental impact category allows for more complete analysis of the potential effects of the Proposed Action, especially given the vast size and geographical diversity of the study area. None of the impact categories areas of analysis extend beyond the noise study area.

Detailed information regarding the affected environment with respect to each relevant impact category is presented in Chapter 4 of the EA.

Environmental Impact Categories and Consequences

The FAA analyzed the potential environmental impacts that could result from implementation of the Proposed Action and the No Action Alternative on all relevant environmental impact categories specified in FAA Order 1050.1E. The FAA evaluated both alternatives for conditions in 2014, the first full year aircraft would use the optimized air traffic routes under the Proposed Action, and 2019, five years after expected implementation of the Proposed Action.

The Proposed Action would not involve land acquisition, physical disturbance, or construction activities and, therefore, would not affect certain environmental impact categories. Both the Proposed Action and the No Action Alternative have the same number and type of aircraft operations. Those environmental impact categories that could potentially be affected by the Proposed Action and their corresponding consequences, reported for the years 2014 and 2019, are:

⁵ The altitude ceiling for environmental consideration regarding “air traffic airspace actions” is generally 10,000 ft. above ground level (AGL). FAA Order 1050.1E, Appendix A, Paragraph 14.5e.

⁶ Analysis between 10,000 ft. and 18,000 ft. AGL may be considered when the proposed changes are over a National Park or Wildlife Refuge. FAA Order JO 7400.2J, Chapter 32, Section 2, paragraph 32-2-1(b)(2)(e). The SSA contains a unit of the National Park System (the Big Thicket National Preserve), as well as several national wildlife refuges.

- **Noise:** As required by FAA Order 1050.1E, noise calculations were completed using the FAA’s Noise Integrated Routing System (NIRS) model, and recorded using the standard yearly day/night average sound level (DNL) metric. The NIRS model computed DNL values for 2014 and 2019 conditions at three sets of data points throughout the PSA:
 1. Population census block centroids
 2. Unique points representing certain specific cultural resources and areas potentially protected under the Department of Transportation Act, Section 4(f)
 3. A uniform grid throughout the PSA (using 3,000 ft. spacing) and SSA (using 6,000 ft. spacing) to document aircraft noise levels at potential noise sensitive locations that were not otherwise identified

The results identified the differences in noise exposure between the two alternatives to determine if implementing the Proposed Action would result in significant noise impacts. A significant noise impact is a DNL increase of 1.5 dB or more in noise sensitive areas exposed to aircraft noise at or above 65 DNL dB. The analysis also looked to identify any DNL increase of 3 dB or more between DNL 60 dB and 65 dB and any DNL increase of 5 dB or more between DNL 45 dB and 60 dB. While the EA refers to such increases as a “Reportable Increase,” they are not significant. The results of the NIRS modeling showed that:

1. The Proposed Action would not result in a DNL 1.5 dB increase in areas exposed to aircraft noise at or above DNL 65 dB.
2. The Proposed Action would not result in DNL increases of 3 dB or more between DNL 60 dB and 65 dB.
3. The Proposed Action would result in a DNL increase of 5 dB between DNL 45 dB and 60 dB.
 - a) Twenty-two people, represented by a single population centroid would experience such an increase in 2014 because of the Proposed Action.⁷ This location is in Liberty County between Route 146 and Farm-to-Market (FM) 1409. As noted above, these increases, while reportable, are not considered significant.
 - b) Four hundred five people would experience a DNL 5 dB or greater increase between DNL 45 to 60 dB in 2019 because of the Proposed Action.⁸ This location is 6-10 miles south of the city of Liberty, Texas, extending from FM 1409 in the southwest to FM 770 in the southeast. Most of the affected population centroids are slightly north of the intersection of FM 1409 and County Road (CR) 450. As noted above, these increases, while reportable, are not considered significant.

⁷ This location would be exposed to DNL 40.7 dB in the 2014 No Action and DNL 45.7 dB in the 2014 Proposed Action. This location is depicted in the EA, Figure 24, Inset 2. The coordinates of this population centroid are: Latitude 29.972719 N, Longitude 94.884928 W

⁸ The locations of these increases are depicted in the EA, Figure 27, Inset 2.

Thus, the Proposed Action would not result in significant noise impacts.

- **Compatible Land Use:** The Proposed Action would not directly affect land uses within the PSA in 2014 or 2019, nor would it result in increased aircraft noise exposure exceeding the FAA's significance threshold for noise impacts on people in 2014 or 2019. The Proposed Action would result in some population centroids being added to the area exposed to DNL 65 dB and higher, but none would experience a significant noise increase (DNL 1.5 dB or greater). Thus, the Proposed Action would not result in significant impacts with regard to Compatible Land Use.
- **Department of Transportation Act, Section 4(f):** As noted under "Noise" above, the FAA's noise modeling included areas potentially protected under Section 4(f). The results of the modeling showed that the Proposed Action would not result in a significant noise increase (i.e., a DNL increase of 1.5 dB or more at or above DNL 65 dB) at any potential Section 4(f) resource. In addition, there are no potential Section 4(f) resources that would experience reportable noise increases (i.e., a DNL increase of 5 dB or more between DNL 45 dB and 60 dB, or a DNL increase of 3 dB or more between DNL 60 dB and 65 dB).

The FAA also examined the general altitudes at which aircraft route changes would occur beyond the immediate environs of the Houston OAPM Airports to determine the potential for light emissions and other visual impacts on potential Section 4(f) resources. The Proposed Action would not involve changes to ground-based light sources and the potential visual effects would be substantially the same as any aircraft overflight, i.e., visual sight of aircraft, contrails, or aircraft lights at night. These effects would not materially differ from those occurring under the No Action Alternative.

Thus, the Proposed Action would not result in substantial impairment of any Section 4(f) resources that would constitute "constructive use" of those resources. Nor would the Proposed Action involve any physical use of a Section 4(f) resource. Under FAA Order 1050.1E, a significant impact would occur when a proposed action either involves more than a minimal physical use of a Section 4(f) resource or would result in a "constructive use" of such a resource. Since the Proposed Action would not result in either a physical or constructive use of Section 4(f) resources, there would not be a significant impact on those resources.

- **Historical, Architectural, Archeological, and Cultural Resources:** The Proposed Action would not adversely affect the historical, architectural, or cultural characteristics of Tribal Lands or historic resources. There are no historically, architecturally, or culturally significant properties that would experience a significant noise increase (i.e., a DNL change of 1.5 dB resulting in a noise exposure level greater than or equal to DNL 65 dB). The Proposed Action would not cause reportable increases (i.e., DNL increase of 3 dB or more between DNL 60 dB and 65 dB or DNL increase of 5 dB or more between DNL 45 dB and 60 dB) in noise at any of the resources studied. Moreover, the Proposed Action would not have an adverse effect on historical, architectural, or cultural resources through introduction of a visual feature to the area that would diminish the integrity of the setting for those properties where setting contributes to the property's historic, architectural, or cultural significance. Accordingly, the FAA has determined that the Proposed Action would not have an "adverse effect" on historic resources under the meaning of 36 CFR, section 800.5(a). The Executive Director of the Texas Historic

Commission (THC) is the appointed State Historic Preservation Officer (SHPO). The FAA consulted with the THC during the EA process. The THC reviewed the EA and, in a letter dated February 21, 2013, concurred with FAA's determination that the Proposed Action would have no adverse effect on historic resources. (See Attachment, "Concurrence Letters, Comments Received, and Responses to Comments".)

In a letter dated February 11, 2013, the Department of the Interior, Bureau of Indian Affairs, stated that there would be "no significant impacts to tribal or Individual Indian trust lands under the jurisdiction of the Southern Plains Region if the proposed changes in aircraft flight paths and altitudes are implemented." (See Attachment, "Concurrence Letters, Comments Received, and Responses to Comments".)

- **Air Quality:** Pursuant to the Federal Clean Air Act of 1970 (CAA), the U.S. Environmental Protection Agency (EPA) established National Ambient Air Quality Standards (NAAQS) for major pollutants, called "criteria pollutants." Currently there are six criteria pollutants: ozone (O₃), carbon monoxide (CO), nitrogen dioxide (NO₂), sulfur dioxide (SO₂), particulate matter (PM), and lead (Pb). PM includes particles with a diameter less than 10 micrometers (PM₁₀) and with a diameter of less than 2.5 micrometers (PM_{2.5}). The EPA has designated parts of the PSA as non-attainment areas for the 8-hour ozone NAAQS. The counties in the PSA are in attainment of the NAAQS for the remaining criteria pollutants.

Implementation of the Proposed Action would result in slightly more fuel burned compared to the No Action. However, the Proposed Action is presumed to conform with the State Implementation Plan (SIP). Accordingly, implementation would not cause or contribute to a new violation of the NAAQS. Therefore, implementation would not have a significant impact on air quality and a conformity determination under Section 176(c) of the CAA is not required.

In a letter dated February 19, 2013, EPA stated that it had "closely considered" the effects of the Proposed Action on air quality and had "no objection to the implementation of the proposed project." (See Attachment, "Concurrence Letters, Comments Received, and Responses to Comments".)

- **Natural Resources and Energy Supply (Aircraft Fuel):** The FAA's NIRS model calculates aircraft-related fuel burn as an output along with calculating aircraft noise exposure. The inputs to NIRS to estimate aircraft-related fuel burn are the same as those used in the noise analysis, such as the average annual day flight schedules, flight tracks, and runway use. The results of the fuel burn analysis for the Proposed Action when compared with the No Action Alternative show that the Proposed Action would result in 1,183 MT (0.21 percent) more fuel burned in 2014 and 2,013 MT (0.28 percent) more fuel burned in 2019. Given these relatively small increases, the FAA expects that the Proposed Action would not adversely affect local fuel supplies when compared with the No Action Alternative. Therefore, the effects of the Proposed Action on natural resources and energy supply would not be significant.
- **Greenhouse Gases and Climate:** Although fuel burn would increase slightly with the Proposed Action compared to the No Action Alternative, no significant project-related effects on climate are expected.

- **Fish, Wildlife, and Plants:** The Proposed Action would not involve any ground disturbance and would not increase the probability of aircraft strikes to migratory birds, nor would it result in an increase in noise that would have the potential to adversely affect the long term survival of any species. Accordingly, the FAA has determined that the Proposed Action is not likely to adversely affect any federally-listed species. In a letter dated March 11, 2013, the U.S. Fish and Wildlife Service stated that it concurs with the FAA’s determination. (See Attachment, “Concurrence Letters, Comments Received, and Responses to Comments”.) Similarly, in a letter dated February 19, 2013, EPA stated that it had “closely considered” the Proposed Action’s effects on endangered and threatened species and had “no objection to the implementation of the proposed project”. (See Attachment, “Concurrence Letters, Comments Received, and Responses to Comments”.) Therefore, the Proposed Action would not have a significant impact on fish, wildlife, or plants.
- **Light Emissions and Visual Impacts:** The Proposed Action would not cause aircraft on the revised routes to be visually intrusive to normal activities on the ground surface. Thus, the Proposed Action would not result in significant impacts with regard to Light Emissions and Visual Impacts.
- **Environmental Justice:** The Proposed Action would not affect low income or minority populations at a disproportionately higher level than other population segments, nor would it disproportionately affect children. In a letter dated February 19, 2013, EPA stated that it had “closely considered” environmental justice and had no objection to implementation of the Proposed Action. (See Attachment, “Concurrence Letters, Comments Received, and Responses to Comments”.) Thus, the Proposed Action would not result in significant impacts with regard to Environmental Justice.

The following environmental impact categories were not analyzed in detail in the EA because there is no potential for the Proposed Action to affect them for the reasons noted below:

- **Coastal Resources** – Without construction or land-disturbing activities, there is no potential for the Proposed Action to affect coastal resources or barrier islands.
- **Construction Impacts** – The Proposed Action does not involve any construction activities.
- **Farmlands** – The Proposed Action has no potential to convert existing prime farmland to a non-agricultural use and the agricultural economy of the area will not be affected.
- **Floodplains** – Without construction or land-disturbing activities, there is no potential for the Proposed Action to affect floodplains.
- **Hazardous Materials, Pollution Prevention, and Solid Waste** – The Proposed Action would not generate, disturb, transport, or treat hazardous materials.
- **Natural Resources and Energy Supply (other than aircraft fuel)** – The Proposed Action would not require unusual natural resources or materials, or those in short supply.

- **Socioeconomic Impacts and Children’s Environmental Health and Safety Risks:**
 - **Socioeconomic Impacts** – The Proposed Action would not involve acquisition of real estate, relocation of residents or community businesses, disruption of local traffic patterns, loss in community tax base, or changes to the fabric of the community.
 - **Children’s Environmental Health** – The Proposed Action would not affect products or substances that a child is likely to come into contact with, ingest, use, or be exposed to, and would not result in environmental health and safety risks that could disproportionately affect children.

- **Secondary (Induced) Impacts** – The Proposed Action would not have the potential for induced or secondary impacts on surrounding communities. It would not cause changes in patterns of population movement or growth, public service demands, or business and economic activity. Furthermore, the Proposed Action does not involve construction activities, so it would not involve the relocation of people or businesses.

- **Water Quality** – Without construction or land-disturbing activities, there is no potential for the Proposed Action to increase impervious surfaces or affect water quality or ground water.

- **Wetlands** – Absent construction or land-disturbing activities, there is no potential for the Proposed Action to affect wetlands.

- **Wild and Scenic Rivers** – The Proposed Action would not foreclose or downgrade Wild, Scenic, or Recreational river status of a river or river segment included in the Wild and Scenic River System.

Cumulative Impacts

To evaluate the potential for cumulative impacts, the FAA considered the incremental impacts of the Proposed Action in conjunction with the impacts of other past, present, and reasonably foreseeable future actions in the vicinity of the Houston OAPM airports, including other projects at the Houston OAPM Airports, other regional airspace projects, and surface transportation projects. Reasonably foreseeable actions were defined as those expected to begin within five years of the Proposed Action.

Due to the nature of the Proposed Action (i.e., the lack of land disturbing or construction activities), the FAA considered potential cumulative impacts in four categories within the Study Area: (1) Noise (including potential impacts on populations in the PSA, Tribal Lands, compatible land use, potential Section 4(f) resources, and historic properties); (2) Air Quality; (3) Natural Resources and Energy Supply; and (4) Climate. Detailed discussion of the cumulative impact analysis with respect to each of these impact categories is presented in Section 5.12 of the EA. Based on that analysis, the FAA does not expect the Proposed Action to result in significant cumulative impacts.

Agency and Public Involvement

The FAA conducted an early consultation process in July and August 2012. The process included letters to the Alabama-Coushatta Tribe of Texas; 20 Federal, state, and local agencies; and 44 Federal and state elected officials. In addition, public notices were placed in 19 area newspapers and a website was

developed (www.oapmenvironmental.com). The FAA provided the web address in the public notices as well as the letters to elected officials. The tribe, agencies, elected officials, and the public were invited to comment on the information available at that time. This early consultation process, including copies of materials made available by the FAA and comments received, is discussed in Chapter 6 and Appendix J of the EA.

The EA was released on January 18, 2013. It was made available in three local libraries (names and addresses listed in Appendix K of the EA) and copies were sent to the Alabama-Coushatta Tribe of Texas, USFWS, EPA, and the THC. The FAA updated the project website to reflect the release of the EA, including making the entire EA available electronically. The FAA published notices of availability of the EA in two area newspapers. All notices and the website solicited comments on the EA. In addition, the FAA sent letters to the previous recipients of the early coordination letters to update them on the status of the project, advise them of the release of EA (including the project's web address), and solicit comments. Letters were sent to the entities listed in Appendix K of the EA. Additionally, after publication of the EA, letters were also sent to the Houston OAPM Airports, the National Aeronautics and Space Administration (NASA), other Department of Defense representatives, and the Ellington Field Air Traffic Control Site Manager. The comment period ended on February 19, 2013.

Comments and FAA Responses

The FAA received comments and/or concurrence letters from six commenters (four agencies and two individuals). The FAA carefully considered all comments received and none warranted revision of the EA. The comments and the FAA's responses are attached to this FONSI/ROD. (See Attachment, "Concurrence Letters, Comments Received, and Responses to Comments".)

Although the comments received resulted in no revisions to the EA, an errata sheet was prepared to correct errors identified after the EA's January 18, 2013, release. The errata sheet is attached to this FONSI/ROD. (See Attachment, "Errata Sheet".)

Mitigation

Because the Proposed Action would not have any significant environmental impacts, no mitigation is being proposed as part of this project.

Other Considerations

The Proposed Action involves air traffic control routing changes for airborne aircraft only. The United States Government has exclusive sovereignty of airspace in the United States [49 U.S.C. §40103(a)]. Congress has provided extensive and plenary authority to the FAA concerning the efficient use and management of the navigable airspace, air traffic control, air navigation facilities, and the safety of aircraft and persons and property on the ground [49 U.S.C. Section 40103(b)(1) and (2)]. Therefore, any applicable community planning initiatives may be preempted by Federal law. To the extent applicable, and as there are no significant impacts under noise or compatible land use, the Proposed Action is consistent with the plans, goals, and policies for the area and with the applicable regulations and policies of Federal, State, and local agencies.

Finding of No Significant Impact

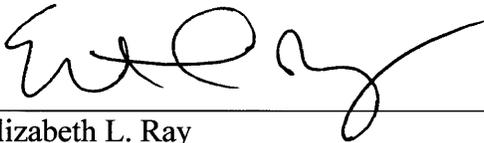
After careful and thorough consideration of the EA and the facts contained herein, I find that the Proposed Action is consistent with existing national environmental policies and objectives as set forth in Section 101 of NEPA and other applicable environmental requirements and will not significantly affect the quality of the human environment or otherwise include any condition requiring consultation pursuant to Section 102(2)(C) of NEPA. Therefore, an environmental impact statement will not be prepared.

Decision

I have carefully considered the FAA's statutory mandate to ensure the safe and efficient use of the national airspace system as well as the other aeronautical goals and objectives discussed in the EA. I find that the Proposed Action is reasonably supported.

Accordingly, under the authority delegated to me by the Administrator of the FAA, I approve the Proposed Action and direct that the necessary actions be taken to implement it.

APPROVED



Elizabeth L. Ray
Vice President, Mission Support Services
Air Traffic Organization
Federal Aviation Administration

June 13, 2013
Date

Right of Appeal

This FONSI/ROD represents the FAA's final decision and approval for the actions identified in the EA and constitutes a final order of the FAA Administrator subject to review by the Courts of Appeal of the United States for the District of Columbia Circuit or in the Court of Appeals of the United States for the circuit in which the person resides or has its principal place of business. The petition must be filed not later than 60 days after the order is issued in accordance with the provisions of 49 U.S.C. 46110. Any party seeking to stay implementation of the FONSI/ROD must file an application with the FAA prior to seeking judicial relief as provided in Rule 18(a) of the Federal Rules of Appellate Procedure.

Concurrence Letters, Comments
Received, and
Responses to Comments
to the Houston Optimization of
Airspace and Procedures in the
Metroplex, Houston Texas (EA),
January 2013

Concurrence Letters and Comments
Received (Annotated)
to the Houston Optimization of Airspace
and Procedures in the Metroplex,
Houston Texas (EA), January 2013

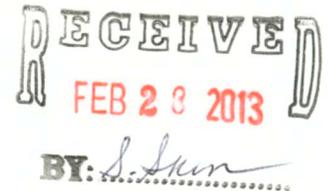
Table of Contents

Concurrence Letters	2
Texas Historical Commission - SHPO	3
U.S. Fish and Wildlife Service	5
Comments Received	6
Department of the Interior- Bureau of Indian Affairs	7
Environmental Protection Agency	8
Michael Kroposki	9
Supplemental Documents - M. Kroposki	11
M. Kroposki, Attachment - Aircraft Noise Measurement and Modeling	12
M. Kroposki, Attachment - Sensitivity of FAA Integrated Noise Model to input parameters	23
M. Kroposki, Attachment - Statistical Concepts in Metrology	37
Tim Honeycutt	84

Concurrence Letters

TEXAS HISTORICAL COMMISSION

real places telling real stories



February 21, 2013

Roger McGrath
Airspace and Environmental Specialist
Operations Support Group
2601 Meacham Boulevard
Fort Worth, Texas 76137

Re: *Project review under Section 106 of the National Historic Preservation Act of 1966, Houston Optimization of Airspace and Procedures in the Metroplex (OAPM) Environmental Assessment – Notice of Availability and Request for Concurrence, Houston, Harris County, Texas 106/FAA (THC Track #201304074, see also 201212674 and 201212152)*

Dear Mr. McGrath,

Thank you for your correspondence providing additional information regarding the above referenced project. This letter serves as comment on the proposed undertaking from the State Historic Preservation Officer (SHPO), the Executive Director of the Texas Historical Commission.

The review staff has completed its review of the material submitted, including the Environmental Assessment (EA), received on January 22, 2013. It is our understanding that the Houston OAPM project seeks to improve the efficiency of the national airspace system in the Houston metroplex by optimizing aircraft arrival and departure procedures at a number of airports, including George Bush Intercontinental (IAH) and William P. Hobby Airport (HOUS). This will involve changes in aircraft flight paths and altitudes in certain areas. Specifically, the FAA proposes to publish and implement optimized standard arrival and departure instrument procedures serving air traffic flows into and out of airports in the Houston Metroplex.

The proposed undertaking will not require any ground disturbance, construction, or land acquisition, and will not increase the number of aircraft operations within southeast Texas airspace. As requested in our agency's correspondence of September 2012, the FAA has assessed noise levels at historic properties in the APE to determine if the undertaking would result in any noise increases that would meet the criteria of adverse effect specified in 36 CFR 800.5. Of the historic resources in the APE that will have a change in noise exposure, all but two sites will have a Day Night Average Sound Level change of one decibel or less. The greatest predicted change in noise exposure will affect the San Jacinto Battlefield and the U.S.S. Texas, both of which are historic resources that are designated National Historic Landmarks. The battlefield in particular is a cultural landscape whose setting greatly contributes to its historical significance. That said, the increase in noise exposure is in the range of 4-4.2 decibels; we find that this change will not diminish the integrity of the property's setting, feeling, or association.

SHPO The FAA has found that the undertaking will have **no adverse effect** on historic resources. Based on the information provided in the EA, the SHPO **concurs** with this determination.

We look forward to further consultation with your office and hope to maintain a partnership that will foster effective historic preservation. Thank you for your cooperation in this federal review process, and for your efforts to preserve the irreplaceable heritage of Texas. **If you have any questions**



Roger McGrath
February 21, 2013
Page 2 of 2

concerning our review or if we can be of further assistance, please call Kelly Little at 512/463-7687.

Sincerely,



Kelly Little, Project Reviewer

For: Mark Wolfe, State Historic Preservation Officer

MW/kl

cc: Janet Wagner, Chair, Harris County Historical Commission



United States Department of the Interior

FISH AND WILDLIFE SERVICE

Division of Ecological Services
17629 El Camino Real #211
Houston, Texas 77058-3051
281/286-8282 / (FAX) 281/488-5882



March 11, 2013

Roger McGrath
Operations Support Group
2601 Meacham Boulevard
Fort Worth, Texas 76137

Dear Mr. McGrath:

Thank you for your letter requesting concurrence with the Federal Aviation Administration's (FAA) determination that the proposed implementation of the Houston Optimization of Airspace and Procedures in the Metroplex in Harris County, Texas, is not likely to adversely affect any federally listed endangered and threatened species in the project area.

FWS1

The United States Fish and Wildlife Service concurs with FAA's determination that the proposed changes in flight paths and altitudes are not likely to adversely affect any federally listed species under our jurisdiction. This concurrence is based upon a review of our project files, and is contingent upon implementation of the avoidance, minimization, and conservation measures proposed by FAA.

In the event the project changes or additional information on the distribution of listed or proposed species or designated critical habitat becomes available, the project should be reevaluated for effects not previously considered.

Our comments are provided in accordance with the provisions of the Endangered Species Act of 1973 (16 U.S.C. 1531 et seq.). Please contact Ms. Kelsey Gocke, staff biologist, at 281/286-8282 ext. 224, if you have any questions or if we can be of further assistance.

Sincerely,

Edith Erfling
Field Supervisor

Comments Received

FEB 1 2013

BY: *S. Sun*



UNITED STATES DEPARTMENT OF THE INTERIOR

**BUREAU OF INDIAN AFFAIRS
SOUTHERN PLAINS REGION
BRANCH OF NATURAL RESOURCES
P.O. BOX 368
ANADARKO, OKLAHOMA 73005**

FEB 1 2013

IN REPLY REFER TO:
NATURAL RESOURCES (405) 247-6673

Roger McGrath, Environmental Specialist
Operations Support Group
2601 Meacham Boulevard
Fort Worth, TX 76137

Dear Mr. McGrath:

We have reviewed the Environmental Assessment (EA) that was prepared to assess the potential environmental impacts of the implementation of the Houston Optimization of Airspace and Procedures in the Metropolis (Houston OAPM) project. **A review of the EA and the associated maps indicate that there would be no significant impacts to tribal or Individual Indian trust lands under the jurisdiction of the Southern Plains Region if the proposed changes in aircraft flight paths and altitudes are implemented.**

DOI 1

We appreciate the opportunity to review and comment on the EA for the Houston OAPM project and have no objections to the proposed changes.

If any additional information is required, please contact David Anderson, Regional Environmental Scientist at 405-247-1532.

Sincerely,

ACTING
David Anderson

Regional Director



UNITED STATES ENVIRONMENTAL PROTECTION AGENCY

REGION 6

1445 ROSS AVENUE, SUITE 1200

DALLAS, TX 75202-2733

February 19, 2013

Roger McGrath
Environmental Specialist
Operations Support Group
Federal Aviation Administration
2601 Meacham Boulevard
Fort Worth, TX 76137

Dear Mr. McGrath:

The Environmental Protection Agency (EPA), Region 6, has completed its review of the Draft Environmental Assessment (DEA) for the implementation of the Houston Optimization of Airspace and Procedure in the Metroplex (Houston OAPM) project prepared by the Federal Aviation Administration. Our review and comments are in accordance with Section 309 of the Clean Air Act and the National Environmental Policy Act (NEPA).

The Houston OAPM project would improve the efficiency of the national airspace system in the Houston metroplex by optimizing aircraft arrival and departure procedures at a number of airports, including George Bush Intercontinental Airport (IAH) and William P. Hobby Airport (HOU). The project would involve changes in aircraft flight paths and altitudes in certain areas, but would not require any ground disturbance or increase the number of aircraft operations within southeast Texas airspace.

Based upon the environmental assessment information and related correspondence of State and other Federal resource agencies, EPA has no objection to the implementation of the proposed project. Factors closely considered include the effects upon air quality, endangered and threatened species, and environmental justice.

EPA 1

Thank you for this opportunity to comment. If you have any questions, please contact Michael Jansky of my staff at 214/655-7451 or by e-mail at jansky.michael@epa.gov for assistance.

Sincerely,

A handwritten signature in black ink, appearing to read "Debra A. Griffin".

Debra A. Griffin
Associate Director
Compliance Assurance and
Enforcement Division

David A. Crandall

From: 9-ASW-HoustonOAPMcomment@faa.gov
Sent: Friday, February 22, 2013 15:45
To: David A. Crandall
Cc: Kirk C. Harris
Subject: Fw: Houston EA comments
Attachments: Modeling Aircraft Noise.htm; AppliedAcoustics2005.pdf; error analysis NBSSpecialPublication747.pdf

Importance: High

-----Forwarded by 9-ASW-HoustonOAPMcomment/ASW/FAA on 02/22/2013 02:35PM -----

To: 9-ASW-HoustonOAPMcomment/ASW/FAA, 9-ASW-HoustonOAPMcomment/ASW/FAA
From: MICHAEL G KROPOSKI <mkroposki@sbcglobal.net>
Date: 02/13/2013 10:40AM
Subject: Houston EA comments

Dear Sirs,

Uncertainty Estimation

MK1 Being somewhat acquainted with the INM software and the mathematics behind it I was very surprised to see in the draft EA reference to calculated DNL differences of 0.1 and 0.4 dB. These numbers are meaningless and therefore misleading because it is my understanding that the precision and accuracy of INM output is at best about +/- 1.0 dB [1,]. The NIRS software uses the INM computation engine for its output. So the NIRS output can have no better precision and accuracy and it is most likely less due to errors introduced by truncation, round off and averaging in NIRS.

MK2 The final EA should state the uncertainty in the output by indicating the validated range of the output DNL numbers, for example 63.45 +/- 0.05dB. This range is commonly referred to as the confidence interval. For the INM successor, AEDT, the FAA's Office of Environment and Energy has undertaken an evaluation of uncertainty for this new software, see *AEDT Version 2a Uncertainty Quantification Report*. Since AEDT uses modeling calculations very similar to INM it is likely the levels of uncertainty found in the AEDT sensitivity analysis are likely to also be present in the NIRS output. The recently studied AEDT levels of uncertainty should be cited as a guide to uncertainty in the similar NIRS software.

MK3 The FAA NEPA regulations use round numbers, for example 65 dB, with one exception 1.5 dB (Order 1050.1E section 14), however table 23 in the draft EA on page 134 cites the noise limits as 3.0 and 5.0 dB. These are different numbers. A change of 4.6 dB would satisfy the 5 dB limit under standard numerical nomenclature [3] but would not be valid according to table 23. Table 23 should be revised in accordance with Order 1050.1E

Aircraft Take Off Weight Estimation

MK4 In the EA draft at section 5.2.2 Methodology, on page 134 it is stated that the " FAA assembled detailed information on IFR aircraft operations for the Analyzed Airports for input into FAA's noise model NIRS." In Appendix G it is stated on page G-6 that

MK5 "G.3.1 Aircraft Noise Performance Specific noise and performance data must be entered into the NIRS for each aircraft type operating at the airport. Noise data are included in the form of sound exposure levels (SELs) at a range of distances from a particular aircraft with engines at a specific thrust level. Performance data

include thrust, speed and altitude profiles for takeoff and landing operations. The NIRS database contains standard noise and performance data for over one hundred different fixed wing aircraft types, most of which are civilian

aircraft. The NIRS automatically accesses the noise and performance data for takeoff and landing operations by those aircraft."

There is no mention in the draft EA of how aircraft take off weights were determined. INM has a default setting for take off weight estimation outlined in the INM technical manual on page 170 **Table G-4-14: Guidance for Determining Departure Takeoff Weights**. This method use trip length to estimate fuel load and adds a factor of 65% payload to estimate the take off weight. The INM User manual warns however on page 13, Section 2.1.3 that the user should " Make every effort to develop accurate average values for input data. In particular, flight profiles and ground tracks must be modeled realistically. and **If feasible, obtain actual takeoff weights and use average weight to choose profile stage numbers instead of using trip length.**" The EA should state specifically how take off weights were determined.

MK6

MK7

MK8

MK9

MK10

While use of the default settings 65% payload may have been realistic in 1970, the current Load Factors clearly show it is not so today. A more realistic average weight is most likely near 100% payload. INM noise calculations are especially sensitive to variations in take off weight. One study of input sensitivities has shown that a 10% variation in take off weight leads to an error of 3-7 dB [2]. Also since large jet aircraft are most likely the largest contributors of noise energy, an error in the largest contributors to DNL will predominate since noise as measured by DNL is aggregated logarithmically. Assuming unrealistically low take off weights have been used in the draft EA, it may be assumed that the calculated DNL's are significantly underestimated! The consequence of this conclusion has direct impact on the overall environmental impact determination because even the underestimated DNLs were extremely close to the 65 dB level of regulatory significance. If they in fact exceed this level, a finding of significant impacts is warranted instead of the finding of no significant impact stated in the EA draft.

Submitted by

Michael Kroposki Esq.

NOTES:

[1] Aircraft Noise Measurement and Modeling at A 4.7.2.5 (copy attached)

[2] "Sensitivity of FAA Integrated Noise Model to Input Parameters" , Applied Acoustics, Vol 66, Issue 3, March 2005, pages 263-276, (copy attached)

[3] "Statistical Concepts in Metrology, NBS Publication 747 at pages 1-2. (copy attached)

Supplemental Documents provided by Michael Kroposki

1. Aircraft Noise Measurement and Modeling
2. Sensitivity of the FAA Integrated Noise Model to Input Parameters
3. Statistical Concepts in Metrology

Environmental Assessment for Houston Optimization of Airspace and Procedures in the Metroplex

Aircraft Noise Measurement and Modeling

The following is an excerpt from Appendix A of *Findings of the Low-Frequency Noise Expert Panel of the Richfield-MAC Noise Mitigation Agreement of 17 December 1998, Annotated*, Sept. 30, 2000. The *red italic type* is text added after the final report was issued May 9, 2000, and was intended by the authors to indicate whether consensus had been reached on significant points. The illustrations are not yet available for reposting here (June 1, 2002).

A.4 HOW AIRCRAFT NOISE IS MEASURED AND MODELED

Standardized procedures have evolved for both measuring and modeling aircraft noise for common purposes. These are outlined in the following subsections.

A.4.1 Frequency-Related Measurement Conventions

The human ear is capable in principle of detecting sounds within a ten octave range extending from about 20 Hz to 20 kHz. It has been well understood since the early 1920s, however, that sensitivity to sounds varies greatly over frequencies within this range. The greatest sensitivity is concentrated within a two octave range extending from roughly 1000 to 4000 Hz that includes many important speech sounds. At extremely low and extremely high frequencies, the ear is thousands of times less sensitive than in the speech range.

When systematic measurements of urban noise were first made in the late 1920s, it was quickly realized that an adjustment of some sort was needed to represent measurements of sounds of differing frequency content in terms meaningful for assessing effects of such noise on people. The simplest solution available at the time was to apply a "frequency weighting network" to measurements of environmental sounds. Three such networks were standardized initially during the 1930s: the A-weighting network for sounds of relatively low absolute sound pressure level, the B-weighting network for sounds of intermediate level, and the C-weighting network for relatively high level sounds. These weighting networks were intended as approximations to the inverse of human hearing sensitivity at increasing sound levels.

The A-weighting network eventually gained acceptance as the default weighting network for general environmental noise measurement purposes. When FAA was charged with regulating aircraft noise emissions, however, it adopted a different measurement procedure for the 1969 Part 36 of the Federal Aviation Regulations -- Perceived Noise Level, or PNL. PNL is a more complex frequency weighting network than the A-weighting network, that is slightly more sensitive than the A-weighting network to low-frequency sounds, and also to sounds in the vicinity of 1 to 3 kHz.

Most references to FAR Part 36 cite the standard in terms of the Effective Perceived Noise Level (EPNL). While an instantaneous level is given in terms of PNL, the level from an event (i.e., a takeoff or a landing) is given in terms of the EPNL. This is analogous to the instantaneous level being cited as an A-weighted level and the sound from an event as the Sound Exposure Level (SEL).

Environmental Assessment for Houston Optimization of Airspace and Procedures in the Metroplex When the Office of Noise Abatement and Control of the Environmental Protection Agency recommended adoption of the Day-Night Average Sound Level for general assessments of environmental noise levels in 1974, readily available instrumentation could not conveniently measure PNL values. The A-weighting network was therefore retained as the basis for routine environmental noise measurements, such as monitoring of aircraft noise levels near airports.

A.4.2 Duration-Related Conventions

A.4.2.1 The "Equal Energy Hypothesis"

As a matter of regulatory policy, it is commonly assumed that people are indifferent between the annoyance of small numbers of very high-level noise events of short duration and the annoyance of large numbers of compensatingly lower level and/or longer duration noise events. In other words, it is conventionally assumed that the number, level, and duration of noise events are fully interchangeable determinants of annoyance, as long as their product (energy sum) remains constant. Thus, a small number of noisy aircraft operations is considered to create the same impact as that of a compensatingly greater number of operations by less noisy aircraft.

It is misleading to attribute the equal energy hypothesis to "regulatory policy." As part of its responsibilities under the mandates of the Noise Control Act of 1972, the EPA recommended adoption of DNL, based on A-weighted levels. As is clear from the report containing that recommendation, the "Levels Document," the EPA base its decision on previous research and experience in other countries, mainly in Europe, and in California, not regulatory policy. ²⁸

[28] Anon. "Information on Levels of Environmental Noise Requisite to Protect Public Health and Welfare with an Adequate Margin of Safety," EPA 550/9-74-004, March 1974.

The assumption of linearity of acoustic effects underlies reliance on the equal energy hypothesis for purposes such as predicting the prevalence of annoyance from long-term, time-weighted average sound levels (such as Day-Night Average Sound Level). This assumption is untenable for present purposes, since the occurrence of noise-induced rattle is a threshold-like phenomenon. In residential settings, people hear rattle when outdoor noise levels exceed some structure-specific and frequency-specific sound level. Furthermore, sound levels of rattling objects do not necessarily increase in direct proportion to the amount by which sound levels exceed a rattle threshold (cf. Schomer et al., 1987a). [P.D. Schomer et al., "Expedient Methods for Rattle-proofing Certain Housing Components," U.S. Army Corps of Engineers, CERL Report N-87/24, 1987]

Under these circumstances, time-integrated noise exposure cannot be expected to predict the annoyance of rattle as well as quantities such as the number or temporal density of noise events in excess of a threshold of rattle.

A.4.2.2 Family of "equivalent level" noise metrics

Figure 76 shows the characteristic form of a time history of sound levels produced during an aircraft overflight of a fixed point on the ground. The sound pressure level at the measurement point initially rises to a maximum, after which it decreases. Since the sound pressure levels vary throughout the overflight, and since the durations of different overflights also vary, no single number can usefully characterize the moment-to-moment changes in sound levels. The usual method for representing the sound energy produced during the entire overflight is therefore to "normalize" the measurement to a standard time period (one second). This measure, "sound

Environmental Assessment for Houston Optimization of Airspace and Procedures in the Metroplex exposure level," simplifies the comparison of noise events of varying duration and maximum level by compressing the acoustic energy of the entire noise event into a standard time period.

Figure 76 (not available)

Relationship of sound exposure level (SEL) to time history of an aircraft overflight.

The concept of a sound exposure level can be generalized to an "equivalent level" of time periods longer than one second. For example, a full day's worth of sound exposure can be expressed as a 24-hour equivalent level, symbolized as Leq_{24} . If a different weighting factor is assigned to the equivalent level of day time (0700 - 2200 hours) and night time (2200-0700 hours), the noise metric becomes a time-weighted 24-hour metric. When the nighttime weighting of the time average is ten times greater than the daytime weighting, the noise measure is known as Day-Night Average Sound Level, abbreviated DNL and symbolized as L_{dn} .

A.4.3 Field Measurement of Aircraft Noise

Part 36 of the Federal Aviation Regulations specifies levels of noise emissions of commercial aircraft offered for sale or otherwise operating in the United States. Regulatory language indicates in great detail the conditions of measurements and analysis of sound level measurements made for purposes of certifying that air craft types are in compliance with Part 36. These include constraints on aircraft operating procedures, atmospheric conditions, multiple microphone positions, half-second sampling of one-third octave band levels from 50 to 10,000 Hz, calculation of variant forms of Perceived Noise Levels, and so forth.

Although Part 36 does not apply to aircraft noise measurements made for purposes other than certification, half-second sampling of one-third octave band sound levels in the 24 bands from 50 to 10,000 Hz are commonplace in field measurements made under less controlled circumstances as well. However, adventitious measurements of aircraft noise (those made under circumstances in which aircraft movements are unconstrained) are much more likely to be influenced by factors such as variability in aircraft operating conditions (thrust settings, flight profiles, etc.), weather conditions, and the presence of extraneous noise sources. These uncontrollable sources of error limit the precision of most field measurements of aircraft noise, and often contribute to the sort of scatter seen in Figures [].

Another obvious limitation of field measurement of aircraft noise is that it is applicable only to existing circumstances of noise exposure. Noise that has not yet been made cannot be measured, but only modeled.

A.4.4 Standard Approach to Modeling Aircraft Noise Exposure Near Airports

Aircraft noise can be modeled in as many ways as there are purposes for modeling. The standard approach to aircraft noise modeling in the immediate vicinity of civil airfields answers the question "How much noise does an airplane flying here make there?" To answer this question, mathematical models of atmospheric propagation of sound are applied to standard sets of aircraft noise levels, to propagate noise emissions away from aircraft (whether in flight or on the ground) in all directions. These calculations are summarized graphically as sets of

Environmental Assessment for Houston Optimization of Airspace and Procedures in the Metroplex source-based emission contours, or sometimes as point values. The goal of this form of aircraft noise modeling is protection of public investment in an airport.

The results of contouring exercises are usually summarized in terms of a time-weighted daily average exposure index devised by the Environmental Protection Agency (EPA, 1974), known as Day-Night Average Sound Level (DNL). DNL provides a convenient means for combining all of the noise energy created in the course of daily flight operations into a single number, for which interpretive criteria and regulatory policy have evolved. Airports routinely produce aircraft noise exposure contours in units of DNL for NEPA disclosure purposes; for purposes related to federal aviation regulations; for land use planning purposes; and for various other purposes.

FAA's preferred aircraft noise prediction software, INM, can produce not only noise exposure (i.e., DNL or CNEL) contours, but with equal facility, contours of maximum noise levels and contours of duration of aircraft noise in excess of a user-specified threshold level ("time-above" contours). INM can also produce spot estimates (rather than entire contour sets) for various noise metrics.

For reasons discussed in Section 2.3 of Volume II, DNL contours are of no direct value as predictors of low-frequency sound level.

A.4.5 Overview of Airfield-Vicinity Noise Exposure Modeling

Computer-based aircraft noise exposure modeling began in the 1970s with the creation of early versions of the U.S. Air Force's NOISEMAP software. FAA began construction of an "Integrated Noise Model" (Olmstead et al., 1997) several years later. [J. Olmstead et al., "INM Version 5.0 User's Guide," FAA Report FAA-AEE-95-01, 1997] Both noise modeling programs have been released in versions for different computing platforms and operating systems. Variants on both programs have also been produced by various government and commercial organizations worldwide.²⁹

[29] For example, ARTS MAP is a commercial software package intended for retrospective use only. At airports with access to information produced by FAA's ARTS III surveillance radars, ARTS MAP replaces assumptions about aircraft operating conditions with information developed from position reports made by aircraft transponders during actual operations.

Although the Air Force and FAA noise models were initially developed separately, recent versions share some algorithms and software modules. NOISEMAP and INM may both be used for retrospective and prospective purposes: to produce noise contours for an historical set of operating conditions, or to predict the noise exposure resulting from alternate hypothetical operating conditions. FAA accepts contours produced by either INM or NOISEMAP as equivalent for regulatory purposes.

INM remains under active development, with Version 6.0 recently released. Differences in DNL contours from release to release for the same input specifications can be sizable. It is expected, for example, that sideline noise contours will be notably wider in Version 6.0 than in current versions of INM. Version 6.0 can also produce C-weighted noise exposure estimates in addition to the A-weighted metrics to which earlier versions of INM were limited.

A.4.6 General Properties of Aircraft Noise Exposure Contours

Environmental Assessment for Houston Optimization of Airspace and Procedures in the Metroplex
As a generality, aircraft noise exposure contours about an individual runway are elliptical, with the major axis oriented along the runway centerline and the minor axis perpendicular to the runway heading. Contours produced by aircraft arriving at an airport are usually straighter and narrower than departure contours, which often show bulges or lobes corresponding to turns away from the runway heading shortly after takeoff. At an airport with intersecting or multiple runways and operating patterns, the number, complexity and variability in aircraft flight paths tend to obscure the basic shapes of noise contours for individual runways. In such cases, noise exposure contours for the airport as a whole tend toward broader shapes.

Noise exposure gradients (rates of change of noise exposure with distance from runway ends) on the order of a thousand feet per decibel are common at large airports. In such cases, uncertainties of fractions of decibels in predicted noise levels may lead to mis-classification of the noise exposure of many city blocks.

A.4.7 Sensitivity of Contour Size And Shape to Modeling Assumptions

A.4.7.1 Major factors affecting noise contour shapes

The orientations of an airport's runways have a major but not necessarily dominant effect on the shape of aircraft noise exposure contours. At an airport with a complex runway layout, assumed departure and arrival tracks can also have pronounced effects on contour shapes, depending on how they are populated with different types of aircraft at different times of day.

A.4.7.2 Major factors affecting contour size

The size of a set of aircraft noise exposure contours is sensitive to more factors than their shape. Two major operational factors affecting contour size are aircraft type and relative proportion of nighttime use. Numbers of operations, especially at large airports, may have a relatively minor effect on relative contour size as compared with flight profiles, stage length, and other factors. Under most conditions, aircraft ground operations do not greatly affect the size of A-weighted noise exposure contours more than a mile or two away from the airport.

A.4.7.2.1 Aircraft type

The proportion of airport operations flown by older (Stage II) aircraft has a major effect on the size of DNL contours. The increasing proportion of Stage III aircraft operations in recent years has been a main factor in shrinking departure contours at many airports. Approach contours are less sensitive to the proportion of Stage II aircraft operating at an airport, since airframe noise may contribute substantially to an aircraft's total A-weighted emissions during approach. Low-frequency noise produced by jet aircraft is more closely related to engine power than to the classification of an aircraft as Stage II or Stage III.

A.4.7.2.2 Fleet mix

All other things being equal, greater proportions of larger (three- and four-engine) jet transports in the fleet serving an airport will lead to larger noise contours. Greater numbers of operations of smaller commuter aircraft (both turboprop and jet) do not generally compensate for their lower noise levels on departures, so that

Environmental Assessment for Houston Optimization of Airspace and Procedures in the Metroplex increasing representation of smaller aircraft in an airport's fleet mix does not necessarily expand an airport's noise contours.

A.4.7.2.3 *Time of day*

The 10 dB nighttime "penalty" incorporated into DNL treats a single nighttime operation as the equivalent of ten daytime operations by the same aircraft. Thus, the 10% of operations that often occur at night at large airports have an effect on contour size equivalent to the 90% of daytime operations. Even small changes in the proportion of nighttime operations can thus have a substantial effect on the size of a set of noise exposure contours.

A.4.7.2.4 *Indirect factors*

Certain assumptions made in creating a noise model can also affect contour size substantially through their indirect influences on operational factors. These include assumptions about wind speed and direction and air temperature, which affect engine power settings, and hence, noise levels.

A.4.7.2.5 *Propagation assumptions*

FAA has not published figures on the fundamental precision of the acoustic propagation algorithms of INM. It is unlikely, however, that INM's air-to-ground acoustic propagation algorithms are much more precise than about ± 1 dB directly beneath an airplane's flight path. Algorithms in past and current versions of INM that are intended to account for "lateral attenuation" -- the absorption of noise in passage over the ground to the side of an aircraft flight track -- are considerably less precise. Bias or random errors in these algorithms can lead to mis-prediction of contour size and shape under some conditions.

A.4.8 Manner of Use of INM

INM is a sufficiently complex program that operates on so many variables that it is possible to use the software in more than one way to accomplish the same end. In particular, a program parameter intended by INM developers to model a particular phenomenon may be used as a de facto means for modeling a different phenomenon, often for reasons of convenience. Rather than creating a custom flight profile for a particular aircraft type as flown from a particular runway, for example, a user might intentionally instruct the program that the destination of a particular flight was closer or farther than is actually the case. This might provide a conveniently simple method for taking into consideration air traffic constraints that prevent a departure stream from gaining altitude as rapidly as might otherwise be the case.

Likewise, rather than creating a unique noise-power-distance curve to describe the manner of operation of a certain class of aircraft at a particular airport, a user might instruct INM to achieve the same effect by treating the approach and departure noise of a particular aircraft type as though it were created by two different aircraft: one for approaches, and a different one for departures. From the perspective of engineering expedience, use of INM parameters in ways unintended by its developers may be viewed as no more than a harmless tactic to save time, effort, and cost in creating an aircraft noise exposure model. Such expedients might also permit a complex noise model to execute on an available computing platform.

Environmental Assessment for Houston Optimization of Airspace and Procedures in the Metroplex

From other perspectives, however, such uses of INM carry certain disadvantages. Perhaps the most basic of these is directness of application. If there is reason to believe that INM does not operate appropriately on some particular information, is it preferable to correct the information or the algorithm that operates on it, or to manipulate the program into producing a modified prediction by other means? From the perspective of improving INM, it is clear that the only way to make progress in correcting potential deficiencies in the program is by addressing them directly rather than working around them. This is also the case from the longer term perspective of recurring uses of INM at the same airport.

Ultimately, the issue is whether INM is viewed as a means for inferring the size and shape of noise exposure contours from first principles -- as intended by its developers -- or whether it is simply an elaborate tool for drawing arbitrary shapes resembling aircraft noise contours. In practice, both the imperfections of modeling and measurement of aircraft noise, as well as differing short- and long-term perspectives on modeling purposes, create a gray area in which professional opinions may differ about the appropriateness of various uses of INM.

A.4.9 Limitations of Interpretations of Aircraft Noise Contours

Aircraft noise contours are often presented in the form of sets of detailed concentric closed form curves overlaid on street grids. This creates the impression that the contours are as fixed, precise, and real as the underlying mapping of streets. In reality, aircraft noise contours are mathematical constructs whose size, shape, and position depend wholly on computational algorithms and assumptions. A given set of assumptions will lead to one set of contours, while a slightly different set of assumptions (about numbers, and types and times of day of aircraft operations from particular runways, on varying flight paths, with different stage lengths and flight profiles, under various meteorological conditions) can lead to very different sets of noise contours. Since there are no facts about the future, any set of prospective noise contours is necessarily speculative and arbitrary to some extent.

All interpretations of aircraft noise contours made for purposes of prospective land use planning must take into consideration the uncertainties inherent in modeling aircraft noise that has not yet occurred.

A.5 UNCERTAINTY IN MEASUREMENT AND MODELING OF AIRCRAFT NOISE

All measurement and modeling is intrinsically imperfect, in that no real world measurement can be absolutely accurate, precise, and reliable, and no modeling is free of simplifying assumptions and approximations. Some of the factors that lead to imperfections of measurement and modeling are manageable, while others are not. Factors that introduce uncertainty into field measurements of aircraft noise include the vagaries of atmospheric propagation of sound (e.g., atmospheric gradients of wind, temperature, humidity, and surface impedance in various propagation paths between the noise source and its measurement), calibration of instrumentation, operational variability in noise sources, and many other "nuisance" variables. Factors that can affect the credibility of aircraft noise modeling include the representativeness of a large number of unverifiable modeling assumptions (e.g., numbers, types, flight paths, and stage lengths of future aircraft operations) and the adequacy of propagation calculations. Factors that can affect measurements of attitudes (such as annoyance) include representativeness and size of samples, as well as wording of questionnaire items.

In the best of circumstances, the inevitable uncertainties of measurement and modeling lead to random errors of specifiable size in estimates of quantities such as sound levels in one-third octave bands, noise reductions of

Environmental Assessment for Houston Optimization of Airspace and Procedures in the Metroplex structures, positions of aircraft noise contours, percentages of survey respondents highly annoyed, and so forth. Under less benign circumstances, these uncertainties can lead to systematic errors of unknown size. As a rule of thumb, it may be assumed that errors of estimation and measurement of acoustic quantities described in this report are generally on the order of ± 2.5 dB, and that errors of measurement of the prevalence of annoyance are generally on the order of $\pm 5\%$.

A.6 FREQUENTLY ASKED QUESTIONS ABOUT THE NATURE OF MEASUREMENT AND MODELING

The following subsections answer frequently asked questions about errors of aircraft noise measurement and modeling.

A.6.1 What is Measurement?

Measurement is a means of associating numbers with quantities such that the ordinary mathematical properties of numbers apply to the quantities of interest. The length of a hanging spring, for example, increases as the weight suspended from it increases. The deflection of a pointer attached to the spring measures weight by pointing to increasingly larger numbers as the weight attached to the spring increases.

A.6.2 What is Modeling?

In the present sense, "modeling" is the process of creating a computer simulation of real world phenomena for purposes of efficiently characterizing the effects of varying assumptions on model predictions. The basic rationale for modeling is cost-effectiveness: since the real world phenomena of interest are too expensive or otherwise inconvenient to characterize directly, a computer-based model of the phenomena is studied instead. The gross behavior of the model — its treatment of major influences on the phenomena of interest, its sensitivity to factors affecting the modeled real world phenomena, and so forth — is intended to resemble the phenomena of interest at a level of detail adequate to provide useful insights.

A.6.3 What is Error?

In the context of the present discussion, error is a technical term that describes a difference between one or more estimates of the numeric value of a quantity. The term does not carry any connotation of intentional or unintentional fault or mistake.

A.6.4 What is Error of Measurement?

Error of measurement is inescapable. No form of measurement, whether of length, weight, economic activity, political preferences, or aircraft noise, is ever error-free. Although more elaborate and costly measurement procedures may produce smaller errors, no amount of money can purchase perfectly error-free measurements. For most practical purposes, what matters is not whether a measurement system is perfect or imperfect, but whether the measurements it produces are adequate to support whatever decisions are made on their basis. It is

.. Environmental Assessment for Houston Optimization of Airspace and Procedures in the Metroplex.. therefore helpful to understand not only the nature of errors of measurement, but also the purposes for which the measurements are made in the first place.

A.6.5 What is Error of Estimation?

"Error of estimation" is a statistical term that refers to the probability that a given estimate lies within a certain interval about a true (but unknowable) exact value. Just as no measurement can ever be perfect, no prediction produced by a software model of long-term aircraft noise levels can be perfect. The statistical term "error of estimation" is sometimes borrowed to describe the inevitable discrepancies between modeled and actual quantities.

Each of the acoustic propagation effects modeled by INM has some associated error, ranging from fractions of a decibel to several decibels under differing conditions. For example, predictions of sound exposure levels at points on the ground directly beneath and relatively close to flight tracks can often be made to agree within a decibel of physical measurements, whereas prediction of sound exposure levels to the sides of flight tracks can be considerably greater.

A.6.6 What is a Confidence Interval?

A confidence interval is a range of values that has a high probability of encompassing a true ("population") value of some parameter. Different sets of measurements ("samples") of the same quantities virtually always differ from one another to some degree for various reasons. For example, average aircraft noise levels observed at the same point near a runway will almost certainly differ from one day to the next. A 90% confidence interval on the mean of a large set of such daily observations encompasses 90% of the daily values. To say that the 90% confidence interval about a mean noise level of 80 dB is 5 dB wide is thus to say that the means of 90% of all sets of measurements of this average noise level will lie between 75 and 85 dB.

The width of a confidence interval depends in large part on (the square root of) the number of observations on which it is based. All other things being equal, small numbers of observations will produce wide confidence intervals, while large numbers of observations will produce narrow confidence intervals. By itself, a wide confidence interval about a data point suggests only that relatively few measurements have been made of its value, not that the underlying variable is somehow incapable of supporting informed decision making.

A.6.7 What are Error Bars?

Error bars attached to data points in charts and graphs are visual indications of the extent of some measure of uncertainty. Plotting a data point with associated error bars serves as a reminder that the point is not the result of a measurement of infinite precision. Figure 77 illustrates error bars plotted for both the independent variable and the dependent variable for a hypothetical data point.³⁰ The ends of the error bars are often used to indicate the upper and lower bounds of confidence intervals. The interval between the upper and lower bounds of error bars need not necessarily be a well defined confidence interval. Charts and graphs are sometimes marked with upper and lower bounds of the envelope of all observations within a data set, or with even less formal ranges of values (such as a range of typical values).

[30] When it is desirable to emphasize errors of measurement on both the abscissa and ordinate, data points are sometimes plotted as ellipses of varying size. The area within an ellipse then serves as a graphic reminder of the uncertainties of measurement associated with each observation. The dashed lines outlining a rectangle in Figure 71 define a region of joint uncertainty of measurement of both the independent and dependent variables.

Figure 77 (not available)

Illustration of the use of error bars to indicate measures of uncertainty for both independent and dependent variables.

A.6.8 Why Are Simplifying Assumptions Necessary for Modeling?

Computer models of real-world phenomena are necessarily simpler than the phenomena themselves. This simplification is necessary both for tractability of calculation, and also because a software model as complex as the modeled phenomena would be both unwieldy and uneconomical. A good software model seeks a balance between excessive and insufficient complexity in its algorithms; between the cost of its construction and use and the savings it yields in study of model rather than real-world behavior; and between accuracy and precision of prediction and the burden it imposes on users for detailed input information.

A.6.9 What is the Difference between Accuracy and Precision?

Errors of estimation may occur either systematically or randomly. Systematic errors (bias errors) affect the accuracy of a measurement or model prediction, while random errors affect its precision. A pattern of target shots is a common metaphor useful for illustrating the two kinds of errors. The bull's eye represents the "true" value of a measurement. The pattern of shots illustrates the accuracy and precision of the measurement. The shot patterns in the four bull's eyes in Figure 78 represent (from top to bottom and left to right) measurements (or predictions) of low accuracy and low precision, low accuracy but high precision, high accuracy and low precision, and high accuracy and high precision.

Figure 78 (not available)

Shot patterns representing four combinations of low and high precision and accuracy in errors of measurement.

In statistical terms, accuracy reflects the difference between the mean of a sample of (say) aircraft noise measurements and the "true" (but unknowable) central tendency. Precision is a measure of the dispersal (variance) of a distribution of measurements. Both the accuracy and precision of measurement of a quantity can be improved by making repeated measurements, as long as the errors of successive measurement are not systematically related to one another. Accuracy and precision of modeling are generally improvable only through more sophisticated algorithms or more comprehensive input information.



ELSEVIER

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Applied Acoustics 66 (2005) 263–276

**applied
acoustics**

www.elsevier.com/locate/apacoust

Sensitivity of the FAA Integrated Noise Model to input parameters

J. Clemente ^{a,*}, E. Gaja ^b, G. Clemente ^c, A. Reig ^b

^a *GARCÍA-BBM Acústica, S.L Valencia, Spain*

^b *Industrial Acoustics Laboratory, Higher Technical School of Industrial Engineering, Polytechnic University of Valencia, Valencia, Spain*

^c *Department of Applied Statistics and Operational Research and Quality, Higher Technical School of Industrial Engineering, Polytechnic University of Valencia, Valencia, Spain*

Received 17 June 2004; received in revised form 16 August 2004; accepted 20 August 2004

Available online 25 September 2004

Abstract

The standard method for computing noise contours around civil airports is SAE-AIR-1845 (FAA Integrated Noise Model, European ECAC-CEAC Doc. 29). It is subject to the inaccuracies implicit in the model as well as those caused by erroneous or imprecise input data. Regarding the latter, the existing errors and/or uncertainties, may be amplified in the output results, to a greater or lower extent, in some cases offering unreliable predictions.

In order to study this phenomenon, the Institute of Sound and Vibration Research (ISVR – Southampton) carried out a theoretical sensitivity analysis based on the segmentation technique, regarding the input parameters of the SAE-AIR-1845, and obtaining the input variables of the model, the variation of which implied greater changes in the output variables. The results were validated by using the FAA Integrated Noise Model Version 6.0 software, based on the aforementioned document. It has been revealed that the model has a greater sensitivity to factors that modify the flight path, and a lower sensitivity to the other parameters. Thus, an error greater than 10% in the variable “gross weight” offers an additional error of between 3 and 7 dB. However, parameters such as the ID of the flaps hardly modify the results obtained for the least favourable case by 1 dB.

* Corresponding author.

E-mail addresses: fraclehe@lycos.es (J. Clemente), egaja@fis.upv.es (E. Gaja), gclemente@eio.upv.es (G. Clemente).

As a result of this research, the sensitivity of the model was quantified for each of the input parameters (taken alone and also in interaction with other parameters), and criteria for the minimisation of global error, resulting from uncertainties in the input parameters, were stated. © 2004 Elsevier Ltd. All rights reserved.

Keywords: Integrated noise model; Sensitivity analysis; Noise modelling; Aircraft noise

1. Introduction

Noise pollution caused by aerial traffic is considered as one of the main environmental problems derived from the normal operation of airports. The control and evaluation of the effects of aerial traffic in the vicinity of airports can be carried out from two different approaches: monitoring and modelling using predictive techniques. The second one is more economical and practical, but it may offer erroneous results unless they get continuously validated and optimised by using real measurements. Therefore, the most efficient solution is a combination of predictive techniques and a system of environmental monitoring that allows the real control of noise in the most sensitive or conflicting areas, as well as a continuous validation and optimisation of the noise model chosen.

At present, there are two kind of predictive mathematical models: those based on simulation and the integrated models. The first ones [1] permit a close representation of a particular scenario (meteorology, terrain, barriers, aircraft, etc.), considering certain effects such as refraction or directivity changes. This offers a profile of sound levels, similar to that which an adequately programmed environmental monitor would offer. However, these types of models involve considerable computational expense, which makes them an inadequate option for the modelling of the aerial traffic of an entire airport.

The integrated models are based on the use of integrated noise data (the data characteristic of each aircraft), obtained from the certifications of aircraft or from real measurements at airports. This permits the modelling of real or hypothetical situations at long time intervals for a great amount of operations, with reasonable computational needs, through the modelling of flights and the use of “Noise-Power-Distance” curves. For this reason, the integrated approach has been mainly selected for the development of aircraft noise predictive models, such as the “FAA Integrated Noise Model” (USA) and the “European ECAC-CEAC Doc. 29” [2] (EU) (based on the standard released by the SAE Committee A-21, SAE-AIR-1845 [3]), or others, such as ANCON2 (United Kingdom), DANSIM (Denmark) or the Kosten Model (Germany).

Among them, the mathematical model chosen for the study was the FAA Integrated Noise Model, in particular its basic document SAE-AIR-1845 as the most used and accepted model worldwide. The software allows the calculation of sound levels in different ways, depending on the nature of the data available [4]. On the one hand, if the data of the instantaneous position, speed and weather, as well as operational conditions of the engine are known, the flight trajectories and the calculation of the sound levels can be calculated as a function of power, distance and

speed, when dealing with the exposure-based metrics. The second option (segmentation technique), indicated for calculations of sound levels over long periods of time, uses standard flight procedures, defined by the user, in order to calculate sound levels, through the creation of a virtual path made of straight finite segments, defined from the values of position, speed and power at their ends. Once all the segments are fully defined, a similar situation to that explained in the first case is obtained, and the calculation of sound levels by interpolation of the NPD curves is identical.

It must be pointed out, that while the first is particularly useful in the evaluation of simple or single cases, it is not feasible for a high number of flights, due to the great amount of data required. In addition, since they are based on real trajectories, the reliable prediction of levels in hypothetical scenarios is not possible, unless a complete and sufficiently broad statistic of totally modelled flights for each sort of aircraft, flight path and procedure, is available. For this reason, the most widely used technique in environmental noise management of all types of airports is segmentation, which allows for the use of average procedures, statistically representative of the others. In addition, this technique is the most complex one, and therefore it is more likely to generate errors within the model, since apart from sound calculations, trajectory and power must be worked out from a relatively small amount of data, according to specific guidelines and hypothesis.

The particular internal architecture of the model allows the errors experienced in the input data to significantly contribute to the global error of the model, largely due to the simplifications and calculation hypothesis used for the derivation of their expressions and procedures.

To evaluate possible uncertainties due to imprecise input, and therefore the reliability and stability of the model, a validation was carried out against a practical case [5]. Consideration was given to the two main existing problems when this model and other mathematical models are applied and the sensitivity of the model explored with regard to the variables that affect its application and the validation of the results obtained.

The results of the research on sensitivity are stated below, together with the methodology used to carry out the sensitivity analysis. It identifies the most influential parameters in the final response of the model, understood as a system with input variables (given by the flight procedures, the characteristics of the aircraft and meteorological conditions) and output variables (sound levels at specific locations). This analysis is carried out both theoretically (SAE-AIR-1845) and using the software (FAA Integrated Noise Model 6.0), and results are processed and interpreted through the use of statistical tools. It must be noted that the FAA Integrated Noise Model includes the entire standard SAE-AIR-1845, as well as some later additions aimed at completing the existing model.

2. Methodology

The sensitivity analysis of the input parameters of the SAE-AIR-1845 leads to the knowledge about the behaviour of these parameters, affecting (to a greater or lesser extent) the output variables, that is, noise levels. It is also possible to analyse the

sensitivity of the fixed coefficients participating in the model equations [6] (the function of the type of aircraft, thrust setting, etc.). However, a sensitivity analysis of the coefficients offers the output variations as a function of hypothetical variations in certain fixed parameters existing within the model, not the output variations as a function of the input data. Thus, this kind of analysis provides model developers useful information about coefficients (obtained from tests and technical essays) and their performance within the model, whereas a sensitivity analysis of the input variables offers information to the entire set of users of the model.

Although this procedure hardly implies a theoretical complexity, it is not straightforward, and it is necessary to follow a sequence of stages in order to obtain precise and reliable results for any analysis conditions.

For this reason, a complete methodology, based on stages has been developed to allow for all aspects for each case under consideration. Since the analysis has been developed from a theoretical and a practical point of view, the procedure has been designed to adjust to both cases, in such a way that in the stage sequence, manual implementation is common to both, with the exception of some tasks (application of equations, interpolation of curves, etc.), which are developed by the INM, as specified in the theoretical case, with some slight modifications.

The basis of the methodology is the calculation of the power developed and the sound levels as a function of the type of aircraft, flight conditions/procedures, the calculated power and distance, respectively. These variables are calculated for all different flight steps, which in commercial aviation are usually takeoff roll (initial acceleration), initial climb, acceleration and flaps retraction and descent. All of these imply an estimation of the power developed by the aircraft. There are additional phases such as continued climb, level flight/cruise climb, landing and deceleration, which are not included, since they are defined by other means or specified by the user. The latter do not consider the calculation of power.

The previous step, common to both cases, is to select all the parameters or variables for analysis among the input parameters of the SAE-AIR-1845/FAA Integrated Noise Model. The following table describes all the parameters involved in the different steps in which a calculation of the power occurs, and whether they are input ones in INM/SAE-AIR-1845, as well as included in the sensitivity analysis (see [Table 1](#)).

In each segment (takeoff, climb, etc.), only one variable among weight, speed (calibrated air speed), flaps ID, climb rate, altitude at closest point of approach (CPA) and descent angle can be considered, regarding the rest of them as constant, and defining their values according to a simplified flight procedure which summarises a standard flight procedure, which consists of a series of segments arranged to describe the entire departure or arrival flight operation. The FAA Integrated Noise Model considers only three types of flight segments for departures (takeoff, climb and acceleration) and three for arrivals (descent, landing and deceleration), and therefore a flight procedure based only on each of the different segments greatly simplifies the process. In this case, standard procedures for a Boeing 747–400 (intermediate weight) included in INM database (weighing 301 ton for departures and 257 ton for arrivals, that is, a typical aircraft covering an intermediate distance) were chosen due to its relative simplicity, and then simplified and adjusted to fit both standard

Table 1
Parameters considered in the sensitivity analysis

Flight step	Parameter	Input parameter	Analysed
Takeoff	Weight	✓	✓
	Speed (CAS)		
	Flaps ID	✓	✓
Initial climb	Weight	✓	✓
	Speed (CAS)		
	Flaps ID	✓	✓
	Climb rate		
Acceleration	Altitude at CPA	✓	✓
	Weight	✓	✓
	Speed (CAS)	✓	✓
	Flaps ID	✓	✓
	Climb rate	✓	✓
Descent	Altitude at CPA		✓
	Weight	✓	✓
	Speed (CAS)	✓	✓
	Flaps ID	✓	✓
	Descent angle	✓	✓
	Altitude at CPA	✓	✓

and simplified altitude, speed and thrust vs. distance plots as much accurately as possible. Considered weights are 310 ton for departures and 250 kg for arrivals. The following tables show both the standard flight procedure and the simplified procedure for departures and arrivals (see Table 2).

Table 2
Standard INM flight procedure for a Boeing 747–400 (departures): Weight 301 ton

	Flight step	Flaps ID	Thrust setting	Final altitude (ft)	Climb speed (ft/min)	Final speed (kt)
1	Takeoff	10	MaxTakeoff	–	–	–
2	Climb	10	MaxTakeoff	1000	–	–
3	Acceleration	10	MaxTakeoff	–	2602.03	190.98
4	Acceleration	5	MaxClimb	–	1000	250
5	Climb	ZERO	MaxClimb	3000	–	–
6	Acceleration	ZERO	MaxClimb	–	1000	269.97
7	Climb	ZERO	MaxClimb	5500	–	–
8	Climb	ZERO	MaxClimb	7500	–	–
9	Climb	ZERO	MaxClimb	10,000	–	–

Table 3
Simplified flight procedure for a Boeing 747-400 (departures): Weight 310 ton

	Flight step	Flaps ID	Thrust setting	Final altitude (ft)	Climb speed (ft/min)	Final speed (kt)
1	Takeoff	10	MaxTakeoff	–	–	–
2	Climb	10	MaxTakeoff	1000	–	–
3	Acceleration	5	MaxClimb	–	1377.95	261.98
4	Climb	ZERO	MaxClimb	10,000	–	–

Table 4
Standard INM flight procedure for a Boeing 747-400 (arrivals): Weight 257 ton

	Flight step	Flaps ID	Initial altitude (ft)	Initial speed (kt)	Touch-down roll (ft)	Distance (ft)	% Reverse Thrust
1	Descent	5	6000	250	–	–	–
2	Descent	10	3000	175.37	–	–	–
3	Descent	D-25	1500	161.39	–	–	–
4	Descent	D-30	1000	155.40	–	–	–
5	Landing	D-30	–	–	533.46	–	–
6	Deceleration	–	–	147.51	–	4802.49	60
7	Deceleration	–	–	30.02	–	0	10

Table 5
Simplified flight procedure for a Boeing 747-400 (arrivals): Weight 250 ton

	Flight step	Flaps ID	Initial altitude (ft)	Initial speed (kt)	Touch-down roll (ft)	Distance (ft)	% Reverse thrust
1	Descent	D-25	6000	250	–	–	–
2	Landing	D-25	–	–	533.46	–	–
3	Deceleration	–	–	147.51	–	4802.49	60

From this point, in order to carry out the sensitivity analysis, the four possible segments (takeoff, climb, acceleration and descent) are used and each variable is considered separately to develop the analysis, according to the different phases outlined below.

- (1) Exhaustive description of the related equations to be considered in each case, for each flight segment, and identification of each variable in every step (as well as the variable analysed) and the conditions of application.
- (2) Assignment of a rank of values within which the study variable can differ in normal flight conditions. The aim of this phase is to be able to represent corrected net thrust (a measure of power) versus the parameter analysed (taking into account the rest of the fixed parameters), and to determine the consistency of the values used in the analysis. Note that all values considered in this analysis are preferred to be within the real operational limits in order to get a close-to-reality approach throughout all the process. However, some configurations might not be operational (let us say, a light aircraft reaching a 400 kt final speed after acceleration segment, or a light aircraft at takeoff flaps ID greater than 10, etc.), but provided it is a sensitivity analysis to know the behaviour of a inter-related set of equations (the model) the application of intermediate theoretical assumptions and states is not important mathematically.
- (3) Selection of a proper hypothetical receiver (located under the flight path), to which the output parameters are calculated later (obviously, they vary throughout the flight, and it is necessary to fix a constant point for their calculation). These receivers are different for each kind of segment, and are selected to

Table 6

Coordinates of receiver points measured from the start of takeoff roll (departure segments) or the touchdown point (arrival segments)

Flight step	X (ft)	Y (ft)	Z (ft)
Takeoff	2500	6000	0
Initial climb	12,000	0	0
Acceleration	25,000	0	0
Descent	–10,000	0	0

represent real observers at the vicinity of any airport. X axis follows the takeoff ground path, Y is perpendicular to this one and Z is the elevation above ground level. Since no ground reflection or absorption effects are considered the elevation is taken as 0 ft.

- (4) Assignment of fixed values to those parameters omitted from the analysis. These values are obtained directly from the simplified flight models shown above (for departures and arrivals), or for the closest point between the receiver and the flight path (CPA) when necessary, by linear interpolating between the initial and final values of the segment or iterations when the first choice is not possible directly (eg. for final altitude calculation when analysing SEL vs. flaps ID in the acceleration segment, in which final altitude depends on a fixed coefficient R depending on flaps ID, the analysed parameter).
- (5) Derivation of the expression relating the corrected net thrust per engine to the parameter considered, giving a graphic plot where possible.
- (6) Calculation of the net thrust (as a numerical value) for the standard value of the parameter analysed. This standard value is obtained from the simplified flight procedure as shown in Tables 3 and 4, as well as the fixed parameters.
- (7) Calculation of the net thrust for the values of the parameter 1%, 10% and 50% greater than the standard value, considering small, medium sized and large variations respectively, in the parameter with regard to its standard value. The standard values have been chosen to be all intermediate-low normal operation ones, so there is a certain margin to increase them by percentages greater than 50% without falling out of normal operation ranks. Nevertheless, if any increased value exceeds the established rank of normal values for the specified aircraft the maximum value of it will be taken instead (and then the % adjusted). For example, the maximum recommended gross weight for a departing Boeing 747–400 may be around 395 ton. An increase of 50% from the standard 310 ton (683,433 lbs, in the analysis considered 700,000 lbs for simplicity reasons) value yields to 465 ton, much greater than the upper bound of the operation rank. Instead, a 395 ton weight, that is, 869,218 lb (as specified in the B747-400 technical specifications at the INM) is considered, which represents a 24.17% variation.

For the case of flaps ID discrete value, where only one higher or lower ID exist, no percentage is considered, though in the tables those variations appear at the 50% variation column (indicating the highest variation) or at the 1% column (indicating a lower value), placing the standard ID

- at the 10% column when both higher and lower values exist at a time (only for presentation purposes).
- (8) Calculation of noise levels through a direct reading of the NPD curves (appearing in the chart with distance and power, and obtaining L_{AE} (dBA)), as internally developed by the INM 6.0 (and, therefore, as specified by standard SAE-AIR-1845), and calculation of the variations in the output variables due to variations in the input variables (conveniently considering each parameter). The calculations will be carried out starting from the net thrust data obtained during steps 6 and 7.

The validation of the theoretically obtained results is driven through an applied sensitivity analysis, in which the same parameters as used for the theoretical analysis (the same variations of 1%, 10% and 50%, receivers, etc.) are considered, using the FAA Integrated Noise Model Version 6.0 software, and using the segmentation technique. The same output data is registered (net thrust and sound exposure level on the receivers) as in the theoretical case, and afterwards, these are used to verify and validate the results obtained theoretically, thereby improving the analysis (see [Tables 5 and 6](#)).

Thus, the applied sensitivity analysis starts from step 1, and all of steps 2, 3, and 4 are developed in the same way as in the theoretical case. Steps 6–8 are computer-aided. In this case, step 5 and the remainder of step 1, which imply the direct use of expressions, are not carried out.

It must be noted that the aim of the analysis is not the calculation of the greatest contributors to the total noise in absolute terms (these would be the engine and the aerodynamic noise [7] as the emitting source, and the distance as the attenuating parameter during propagation), but the calculation of those parameters whose variations or errors affect the output parameter the most, that is, those that can most influence the differences or errors observed in the output data (differences from the theoretical point of view of a sensitivity analysis and errors from the real application of INM point of view).

3. Results of the sensitivity analysis

The sensitivity analysis is carried out both theoretically and practically, by using the same methodology in both cases, as previously explained, and the same parameters (with the exception of the entry “height”, which cannot be specified on the INM 6.0 software, since this application does not permit the analysis of individual segments, which can be carried out theoretically). Instead, the values for the height calculated by the application for the rest of the parameters are used.

The calculations are developed for power and sound exposure level (SEL), the basic noise indicator on the NPD curves, direct consequence of power. However, the present report will only show those corresponding to sound levels, since they are the ultimate result of the model application and for reasons of simplicity.

Once the SEL variations are obtained as a function of variations in the input parameters, a multiple regression analysis is carried out for each segment, with the

different parameters analysed as input variables (with their respective standard values at the simplified flight procedures, and variations of 1%, 10% and 50%, excepting a 24.3% for weight case) the rest of the parameters remain fixed) and the corresponding calculated SEL to determine which variables are significant in the model, offering an idea of which variables are the most important within each segment. This analysis allows to set the variability of a dependent variable (in this case SEL), as a function of other independent or dependent parameters. The accuracy of the regression model is expressed by coefficient R^2 , in such a way that the nearer it is to 1, the greater will be the reliability of the total model. Individually, those variables which coefficient in the regression is significantly different from zero with $p < 0.05$, are considered significant. Among the significant variables, those that, as a whole, explain a 95% of the variability of the SEL will be chosen, in order to filter and select only the significant parameters.

The results obtained for the Sound Exposure Level are given below (see Tables 7 and 8). %Ex and %Acu columns explain the percentage of explanation of the model for each variable and the accumulated percentage, respectively.

3.1. Interpretation of the theoretical sensitivity analysis

The parameter that most directly affects the errors in the estimated levels is the minimum distance between source and receiver. Thus, any parameter that during

Table 7
Results of the theoretical sensitivity analysis (SEL)

	Parameter	SEL variation for increases (%) on the parameters (dBA)			Regression model			
		1%	10%	50%	Signific	%Ex	%Acu	R^2
Takeoff	Weight^a	0.05	0.48	1.03	✓	99.33	99.33	0.9933
	Flaps ID ^b	–	–	–0.09		–	–	
Initial climb	Weight^a	1.60	2.57	7.01	✓	99.67	99.67	0.9991
	Altitude at CPA	0.00	–0.98	–2.85	✓	–	–	
	Flaps ID ^b	–	–	0.29		–	–	
Acceleration	Altitude at CPA	0.01	–0.61	–2.92	✓	90.86	90.86	0.9940
	Weight^a	0.13	1.07	2.30	✓	3.99	94.85	
	Speed (CAS)	0.04	0.51	1.52	✓	2.85	97.70	
	Flaps ID ^c	–0.36	–	0.43	✓	–	–	
	Climb rate	0.03	–0.21	–0.84	✓	–	–	
Descent	Angle	–0.56	–0.89	–3.81	✓	54.30	54.30	0.9867
	Altitude at CPA	0.00	–0.74	–3.13	✓	41.84	96.14	
	Flaps ID ^b	–	–	0.71	✓	–	–	
	Speed (CAS)	–0.01	–0.06	–0.30		–	–	
	Weight ^a	0.00	–0.02	0.03		–	–	

^a The actual percentage for high variations is 24.17%.

^b The highest Flaps ID was considered to represent 50% of variation.

^c ZERO Flaps ID, the lowest value, was considered lowest variation (1%), the standard (5°) the 10%, and the highest (10°) the 50%.

Table 8

Results of the practical sensitivity analysis (SEL), worked out with the INM

	Parameter	SEL variation for increases (%) on the parameters (dBA)			Regression model			
		1%	10%	50%	Signific	%Ex	%Acu	R ²
Takeoff	Weight^a	0.10	0.70	1.60	✓	99.33	99.33	0.9933
	Flaps ID ^b	–	–	0.10		–	–	
Initial climb	Weight^a	0.10	2.20	7.30	✓	99.63	99.63	0.9999
	Altitude at CPA	–	–	–	✓	–	–	
	Flaps ID ^b	–	–	–0.40	✓	–	–	
Acceleration	Altitude at CPA	–	–	–	✓	81.55	81.55	0.9999
	Climb rate	–0.10	–0.20	–0.50	✓	14.32	95.87	
	Flaps ID^c	–0.40	–	0.50	✓	3.52	99.39	
	Speed (CAS)	0.00	0.30	1.50	✓	–	–	
	Weight ^a	0.10	1.00	2.50		–	–	
Descent	Altitude at CPA	–	–	–	✓	95.24	95.24	0.9989
	Flaps ID^b	–	–	0.10	✓	3.93	99.17	
	Speed (CAS)	0.00	0.00	–0.30	✓	–	–	
	Weight ^a	0.00	0.00	0.10	✓	–	–	
	Angle	–0.30	–0.80	–3.50		–	–	

^a The actual percentage for high variations is 24.17%.

^b The highest Flaps ID was considered to represent 50% of variation.

^c ZERO Flaps ID, the lowest value, was considered lowest variation (1%), the standard (5°) the 10%, and the highest (10°) the 50%.

the process of path construction, carried out using the SAE-AIR-1845, greatly contributes to the aircraft position will be a significant parameter to be considered.

Thereby, the main parameter is weight, although it must be pointed out that weight is not only an element that increases the required power, but also an element that modifies the flight path, and therefore, the distance to the receiver. According to the individual segments, weight is considered a key parameter in the takeoff and initial climb steps, although, in the case of takeoff, it affects power, and in the initial climb it affects distance (it must be recalled that power and distance define the sound level in the NPD curves).

In the acceleration step, the main parameter is altitude at CPA, as well as weight and speed. This altitude is basically defined by the final state of the previous step, and, therefore, there is an important interaction with the “weight” variable, which is the actual determining feature. It must be pointed out that, the fact that a significant parameter does not lie in an analysis within the main parameters can be because its variability is explained by other parameters of greater significance, and therefore, it is removed from the regression model. For instance, the “Climb rate” variable is a direct function of altitude and speed. It is, therefore, reasonable that it is not included among the main variables, since both the “Altitude at CPA” and the “Speed (CAS)” explain almost the entire variability due to “Climb rate” in the model.

With regard to the descents, the angle and altitude are the main parameters, provided that altitude, an internal parameter (not an input parameter) is defined by the angle, and, therefore, it could be stated that almost the entire variability of the model is caused by this parameter.

3.2. Interpretation of the practical sensitivity analysis

In general, as happens in the theoretical case, weight and altitude are observed as the main elements that most influence the SEL variations offered by the INM. In fact, both are present in all segments, sharing almost the entire variability of the model.

In particular, weight is the main factor affecting the SEL variability in the takeoff and initial climb segments (although there are other significant parameters, such as altitude and the flaps ID, however, these are of lesser importance). In the acceleration segment, the altitude at the CPA is the most important parameter, followed by the ascent speed CAS or the climb rate and the flaps ID. It must be noted, that in this case, the weight variability is fully explained by altitude, corroborating the phenomenon explained in the theoretical case.

In the descent segment, the main parameter is altitude, as well as the flaps ID. The angle, considered significant in the theoretical case, is now completely omitted from the model (resulting of no importance). This is due to the fact that altitude and angle are so closely related, that the former, being the most powerful parameter within the regression model, includes its own variability and that of the angle, which is left with almost no percentage of explanation in the full model. However, to all intents and purposes, the angle will be considered the most important parameter, since it is a variable directly specified as an input in the model.

3.3. Validation of the theoretical study with the practical study

Comparing the theoretical results with those obtained using the INM 6.0, it is observed that results are very similar (theoretically, they should be the same), which offers practically equal results in the identification of the parameters. In fact, from the aforementioned, with the exception of small differences caused by the interactions among variables, it is deduced that results obtained are virtually the same. A simple regression analysis, putting theoretically and practically calculated thrust into relation regarding SEL, has been carried out, achieving linear equations slope resulting in almost 1 (which implies that $Y = X$), proving the validity of both studies.

However, it must be pointed out that, even if differences existed, these would be perfectly explainable since the number of data used for the multiple regression analysis is quite small, and even smaller in the case of the practical study, where altitude at CPA (an internal parameter, not an actual input) cannot be entered into the model, removing four additional points from the regression analysis (corresponding to its standard value, and the three percentage variations).

A comparative table, with the most important considered parameters both in the theoretical and in the practical study, is shown below (see [Table 9](#)). The existing

Table 9
Most relevant parameters affecting SEL in the INM

Flight step	Theoretical	INM
Takeoff	Weight	Weight
Initial climb	Weight	Weight
Acceleration	Altitude at CPA Weight Speed (CAS)	Altitude at CPA Speed (CAS) Flaps ID
Descent	Angle Altitude at CPA	Altitude at CPA Flaps ID

differences are explained from the point of view of the interactions between the parameters.

3.4. The thrust setting parameter

Although thrust setting parameter was not considered for the analysis, mainly due to its direct and obvious influence on noise in NPD curves, an accurate estimation of the exact cutback point, in which the thrust setting is changed from Takeoff thrust to Climb thrust, has a noticeable influence on predicted noise levels. In fact, for a Boeing 747–400 series aircraft the accurate consideration of the ‘cutback point’ may imply a reduction of over 10,000 lbs. per engine, which translates into nearly 5 dB L_{Amax} reduction. Current Noise Abatement Policies settled in major worldwide airports deal with cutback points closer and closer to the runway, enhancing noise emissions and comfort on areas nearby due to the lower flyover altitude in which this change is performed. Therefore a precise estimation of this point might have a great importance in noise predictions and it would improve the global model accuracy with independence from the behaviour of the input parameters.

4. Conclusions

Once the analysis was carried out, some conclusions could be drawn from the results for the segmentation technique used. However, these results can be extrapolated to the fixed-point technique, since some parameters of this one are also included in the segmentation. Taking this into account, it can be stated that:

- (a) The most influential input parameters on the total predicted noise at a receiver are the aircraft type and the direct distance to the receiver. The rest of parameters are secondary, and their importance depends on the aircraft’s position in the flight procedure.
- (b) For a specific type of aircraft, the INM input parameters that most influence the errors and uncertainties in the estimations are the gross weight of the aircraft, the calibrated airspeed and to a lesser extent the position of flaps.

- (c) Although some variations have been carried out in the analysis (simulating input “errors”) of up to 50%, the errors incurred from data used in real life are in general no greater than 10%, which reduces the associated uncertainty, since it is small by reason of the results obtained. This leads us to consider that the INM has a robust set of equations, slightly sensitive to small perturbations in the input data and even slighter the less they affect aircraft flight path.

As a result of this research project, a series of recommendations with regard to the optimal use of the INM have been proposed. In particular, those obtained from the results of the sensitivity analysis are as follows:

- (a) For the modelling of the airport noise with the Integrated Noise Model the use of the fixed-point technique would be preferable, since it uses precise values for distances source-receiver, whenever the data availability and the computational cost permit it. However, an additional effort must be made in order to obtain paths that are as exact as possible, both using the fixed-point technique or the segmentation technique. When necessary, different flight procedures should be used for the same aircraft and even for similar weight values, in order to show different flight groups (for instance, the different procedures of different airlines), and as such, the least possible dispersion in the standard path in comparison with the real ones. As the number of the different flight procedures increases, a lower dispersion and a greater accuracy will be obtained.
- (b) With regard to the input parameters, a special effort should be made to obtain the precise weight data, as well as the approximate configuration of the aircraft, whenever possible. It is advisable to establish the greatest possible number of different weight groups, since each will have its own flight procedures. On the other hand small errors (that may be inevitable) are permitted, since the INM is, in general, little sensitive to small variations in the input parameters. However, the model is more sensitive to variations of a certain entity (above 5%) in distance and thrust respectively, which can significantly affect the calculations.
- (c) Finally, the consideration of a “Cutback point”, will result in a significant improvement of predictions, mainly for points close to the runway, due to the influence of thrust on noise emitted.

Acknowledgements

We would like to thank the ISVR, in particular Dr. Ian H. Flindell, for the support given in the development of this project. We would also like to thank the International Erasmus Office of the Higher Technical School of Industrial Engineering at the UPV, for the steps taken in order to enable the visit of UPV students to the ISVR.

References

- [1] Hullah P, Cavadini L. Aircraft noise modelling validation through the use of full 4-D flight trajectories including thrust calculation. In: Fourth FAA/Eurocontrol R&D Conference.
- [2] European Civil Aviation Conference. Report on Standard Method of Computing Noise Contours around Civil Airports, doc. 29.
- [3] Society of Automotive Engineers, A21 Committee. Procedure for the Calculation of Aircraft Noise in the Vicinity of Airports, SAE-AIR-1845.
- [4] Federal Aviation Administration. Integrated Noise Model 5.1 Technical Manual.
- [5] Clemente FJ. Uncertainties in Aircraft Noise Modelling. Institute of Sound and Vibration Research, Southampton, United Kingdom, 2003.
- [6] EUROCONTROL Experimental Centre. Error Sensitivity Analysis of the Integrated Noise Model, doc, EEC/ENV/2002/006, 2002.
- [7] Singer BA, Lockard DP, Brentner KS. Computational aeroacoustic analysis of slat trailing-edge flow.

NBS Special Publication 747

***Statistical Concepts in Metrology – With a Postscript
on Statistical Graphics***

Harry H. Ku

Statistical Engineering Division
Center for Computing and Applied Mathematics
National Engineering Laboratory
National Bureau of Standards
Gaithersburg, MD 20899

August 1988



U.S. Department of Commerce
C. William Verity, Secretary

National Bureau of Standards
Ernest Ambler, Director

Library of Congress
Catalog Card Number: 88-600569
National Bureau of Standards
Special Publication 747
Natl. Bur. Stand. (U.S.),
Spec. Publ. 747,
48 pages (Aug. 1988)
CODEN: XNBSAV

U.S. Government Printing Office
Washington: 1988

For sale by the Superintendent
of Documents,
U.S. Government Printing Office,
Washington, DC 20402

Contents

Statistical Concepts of a Measurement Process	1
Arithmetic Numbers and Measurement Numbers.....	1
Computation and Reporting of Results	2
Properties of Measurement Numbers	3
The Limiting Mean	3
Range, Variance, and Standard Deviation.....	4
Population and the Frequency Curve.....	4
The Normal Distribution	6
Estimates of Population Characteristics.....	8
Interpretation and Computation of Confidence Interval and Limits...	9
Precision and Accuracy	11
Index of Precision	11
Interpretation of Precision	12
Accuracy.....	13
Statistical Analysis of Measurement Data	13
Algebra for the Manipulation of Limiting Means and Variances	14
Basic Formulas	14
Propagation of Error Formulas.....	16
Pooling Estimates of Variances.....	18
Component of Variance Between Groups.....	19
Comparison of Means and Variances	20
Comparison of a Mean with a Standard Value.....	20
Comparison Among Two or More Means.....	21
Comparison of Variances or Ranges.....	23
Control Charts Technique for Maintaining Stability and Precision....	24
Control Chart for Averages.....	24
Control Chart for Standard Deviations	25
Linear Relationship and Fitting of Constants by Least Squares	28
References.....	29
Postscript on Statistical Graphics	31
Plots for Summary and Display of Data	31
Stem and Leaf.....	31
Box Plot.....	33
Plots for Checking on Models and Assumptions	35
Residuals	36
Adequacy of Model.....	36
Testing of Underlying Assumptions	38
Stability of a Measurement Sequence.....	40
Concluding Remarks.....	42
References.....	42

List of Figures

2-1	A symmetrical distribution	5
2-2	(A) The uniform distribution. (B) The log-normal distribution.....	5
2-3	Uniform and normal distribution of individual measurements having the same mean and standard deviation, and the correspond- ing distribution(s) of arithmetic means of four independent measurements.....	7
2-4	Computed 90% confidence intervals for 100 samples of size 4 drawn at random from a normal population with $m = 10$, $\sigma = 1$...	11
2-5	Control chart on \bar{x} for NB'10 gram.....	25
2-6	Control chart on s for the calibration of standard cells.....	26
1	Stem and leaf plot. 48 values of isotopic ratios, bromine (79/81)...	32
2	Box plot of isotopic ratio, bromine (79/91)	34
3	Magnesium content of specimens taken	35
4	Plot of deflection vs load.....	37
5	Plot of residuals after linear fit.....	37
6	Plot of residuals after quadratic fit.....	38
7	Plot of residuals after linear fit. Measured depth of weld defects vs true depth	39
8	Normal probability plot of residuals after quadratic fit	39
9	Differences of linewidth measurements from NBS values. Measure- ments on day 5 inconsistent with others—Lab A	40
10	Trend with increasing linewidths—Lab B	41
11	Significant isolated outliers—Lab C	41
12	Measurements (% reg) on the power standard at 1-year and 3-month intervals	42

List of Tables

2-1	Area under normal curve between $m - k\sigma$ and $m + k\sigma$	6
2-2	A brief table of values of t	10
2-3	Propagation of error formulas for some simple functions	17
2-4	Estimate of σ from the range.....	19
2-5	Computation of confidence limits for observed corrections, NB'10 gm.....	21
2-6	Calibration data for six standard cells.....	27
1	Y—Ratios 79/81 for reference sample.....	32

Statistical Concepts in Metrology – With a Postscript on Statistical Graphics

Harry H. Ku

Statistical Engineering Division, National Bureau of Standards, Gaithersburg, MD 20899

“Statistical Concepts in Metrology” was originally written as Chapter 2 for the Handbook of Industrial Metrology published by the American Society of Tool and Manufacturing Engineers, 1967. It was reprinted as one of 40 papers in NBS Special Publication 300, Volume I, Precision Measurement and Calibration; Statistical Concepts and Procedures, 1969. Since then this chapter has been used as basic text in statistics in Bureau-sponsored courses and seminars, including those for Electricity, Electronics, and Analytical Chemistry.

While concepts and techniques introduced in the original chapter remain valid and appropriate, some additions on recent development of graphical methods for the treatment of data would be useful. Graphical methods can be used effectively to “explore” information in data sets prior to the application of classical statistical procedures. For this reason additional sections on statistical graphics are added as a postscript.

Key words: graphics; measurement; metrology; plots; statistics; uncertainty.

STATISTICAL CONCEPTS OF A MEASUREMENT PROCESS

Arithmetic Numbers and Measurement Numbers

In metrological work, digital numbers are used for different purposes and consequently these numbers have different interpretations. It is therefore important to differentiate the two types of numbers which will be encountered.

Arithmetic numbers are exact numbers. 3, $\sqrt{2}$, $\frac{1}{3}$, e , or π are all exact numbers by definition, although in expressing some of these numbers in digital form, approximation may have to be used. Thus, π may be written as 3.14 or 3.1416, depending on our judgment of which is the proper one to use from the combined point of view of accuracy and convenience. By the

usual rules of rounding, the approximations do not differ from the exact values by more than ± 0.5 units of the last recorded digit. The accuracy of the result can always be extended if necessary.

Measurement numbers, on the other hand, are not approximations to exact numbers, but numbers obtained by operation under approximately the same conditions. For example, three measurements on the diameter of a steel shaft with a micrometer may yield the following results:

<i>No.</i>	<i>Diameter in cm</i>	<i>General notation</i>
1	0.396	x_1
2	0.392	x_2
3	0.401	x_3
	Sum 1.189	$\sum_{i=1}^n x_i$
	Average 0.3963	$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$
	Range 0.009	$R = x_{\max} - x_{\min}$

There is no rounding off here. The last digit in the measured value depends on the instrument used and our ability to read it. If we had used a coarser instrument, we might have obtained 0.4, 0.4, and 0.4; if a finer instrument, we might have been able to record to the fifth digit after the decimal point. In all cases, however, the last digit given certainly does not imply that the measured value differs from the diameter D by less than ± 0.5 unit of the last digit.

Thus we see that measurement numbers differ by their very nature from arithmetic numbers. In fact, the phrase “significant figures” has little meaning in the manipulation of numbers resulting from measurements. Reflection on the simple example above will help to convince one of this fact.

Computation and Reporting of Results. By experience, the metrologist can usually select an instrument to give him results adequate for his needs, as illustrated in the example above. Unfortunately, in the process of computation, both arithmetic numbers and measurement numbers are present, and frequently confusion reigns over the number of digits to be kept in successive arithmetic operations.

No general rule can be given for all types of arithmetic operations. If the instrument is well-chosen, severe rounding would result in loss of information. One suggestion, therefore, is to treat all measurement numbers as exact numbers in the operations and to round off the final result only. Another recommended procedure is to carry two or three extra figures throughout the computation, and then to round off the final reported value to an appropriate number of digits.

The “appropriate” number of digits to be retained in the final result depends on the “uncertainties” attached to this reported value. The term “uncertainty” will be treated later under “Precision and Accuracy”; our only concern here is the number of digits in the expression for uncertainty.

A recommended rule is that the uncertainty should be stated to no more than two significant figures, and the reported value itself should be stated

Environmental Assessment for Houston Optimization of Airspace and Procedures in the Metroplex to the last place affected by the qualification given by the uncertainty statement. An example is:

“The apparent mass correction for the nominal 10 g weight is +0.0420 mg with an overall uncertainty of ± 0.0087 mg using three standard deviations as a limit to the effect of random errors of measurement, the magnitude of systematic errors from known sources being negligible.”

The sentence form is preferred since then the burden is on the reporter to specify exactly the meaning of the term uncertainty, and to spell out its components. Abbreviated forms such as $a \pm b$, where a is the reported value and b a measure of uncertainty in some vague sense, should always be avoided.

Properties of Measurement Numbers

The study of the properties of measurement numbers, or the Theory of Errors, formally began with Thomas Simpson more than two hundred years ago, and attained its full development in the hands of Laplace and Gauss. In the next subsections some of the important properties of measurement numbers will be discussed and summarized, thus providing a basis for the statistical treatment and analysis of these numbers in the following major section.

The Limiting Mean. As shown in the micrometer example above, the results of *repeated measurements of a single physical quantity under essentially the same conditions* yield a set of measurement numbers. Each member of this set is an estimate of the quantity being measured, and has equal claims on its value. By convention, the numerical values of these n measurements are denoted by x_1, x_2, \dots, x_n , the arithmetic mean by \bar{x} , and the range by R , i.e., the difference between the largest value and the smallest value obtained in the n measurements.

If the results of measurements are to make any sense for the purpose at hand, we must require these numbers, though different, to behave as a group in a certain predictable manner. Experience has shown that this is indeed the case under the conditions stated in italics above. In fact, let us adopt as the postulate of measurement a statement due to N. Ernest Dorsey (reference 2)*:

“The mean of a family of measurements—of a number of measurements for a given quantity carried out by the same apparatus, procedure, and observer—approaches a definite value as the number of measurements is indefinitely increased. Otherwise, they could not properly be called measurements of a given quantity. In the theory of errors, this limiting mean is frequently called the ‘true’ value, although it bears no necessary relation to the true quaesitum, to the actual value of the quantity that the observer desires to measure. This has often confused the unwary. Let us call it the limiting mean.”

Thus, according to this postulate, there exists a limiting mean m to which \bar{x} approaches as the number of measurements increases indefinitely, or, in symbols $\bar{x} \rightarrow m$ as $n \rightarrow \infty$. Furthermore, if the true value is τ , there is usually a difference between m and τ , or $\Delta = m - \tau$, where Δ is defined as the bias or systematic error of the measurements.

*References are listed at the end of this chapter.

In practice, however, we will run into difficulties. The value of m cannot be obtained since one cannot make an infinite number of measurements. Even for a large number of measurements, the conditions will not remain constant, since changes occur from hour to hour, and from day to day. The value of τ is unknown and usually unknowable, hence also the bias. Nevertheless, this seemingly simple postulate does provide a sound foundation to build on toward a mathematical model, from which estimates can be made and inference drawn, as will be seen later on.

Range, Variance, and Standard Deviation. The range of n measurements, on the other hand, does not enjoy this desirable property of the arithmetic mean. With one more measurement, the range may increase but cannot decrease. Since only the largest and the smallest numbers enter into its calculation, obviously the additional information provided by the measurements in between is lost. It will be desirable to look for another measure of the dispersion (spread, or scattering) of our measurements which will utilize each measurement made with equal weight, and which will approach a definite number as the number of measurements is indefinitely increased.

A number of such measures can be constructed; the most frequently used are the variance and the standard deviation. The choice of the variance as the measure of dispersion is based upon its mathematical convenience and maneuverability. Variance is defined as the value approached by the average of the sum of squares of the deviations of individual measurements from the limiting mean as the number of measurements is indefinitely increased, or in symbols:

$$\frac{1}{n} \sum (x_i - m)^2 \rightarrow \sigma^2 = \text{variance, as } n \rightarrow \infty$$

The positive square root of the variance, σ , is called the standard deviation (of a single measurement); the standard deviation is of the same dimensionality as the limiting mean.

There are other measures of dispersion, such as average deviation and probable error. The relationships between these measures and the standard deviation can be found in reference 1.

Population and the Frequency Curve. We shall call the limiting mean m the location parameter and the standard deviation σ the scale parameter of the population of measurement numbers generated by a particular measurement process. By population is meant the conceptually infinite number of measurements that can be generated. The two numbers m and σ describe this population of measurements to a large extent, and specify it completely in one important special case.

Our model of a measurement process consists then of a defined population of measurement numbers with a limiting mean m and a standard deviation σ . The result of a single measurement X^* can take randomly any of the values belonging to this population. The probability that a particular measurement yields a value of X which is less than or equal to x' is the proportion of the population that is less than or equal to x' , in symbols

$$P\{X \leq x'\} = \text{proportion of population less than or equal to } x'$$

*Convention is followed in using the capital X to represent the value that might be produced by employing the measurement process to obtain a measurement (i.e., a random variable), and the lower case x to represent a particular value of X observed.

Similar statements can be made for the probability that X will be greater than or equal to x'' , or for X between x' and x'' as follows: $P\{X \geq x''\}$, or $P\{x' \leq X \leq x''\}$.

For a measurement process that yields numbers on a continuous scale, the distribution of values of X for the population can be represented by a smooth curve, for example, curve C in Fig. 2-1. C is called a frequency curve. The area between C and the abscissa bounded by any two values (x_1 and x_2) is the proportion of the population that takes values between the two values, or the probability that X will assume values between x_1 and x_2 . For example, the probability that $X \leq x'$, can be represented by the shaded area to the left of x' ; the total area between the frequency curve and the abscissa being one by definition.

Note that the shape of C is not determined by m and σ alone. Any curve C' enclosing an area of unity with the abscissa defines the distribution of a particular population. Two examples, the uniform distribution and the log-normal distribution are given in Figs. 2-2A and 2-2B. These and other distributions are useful in describing certain populations.

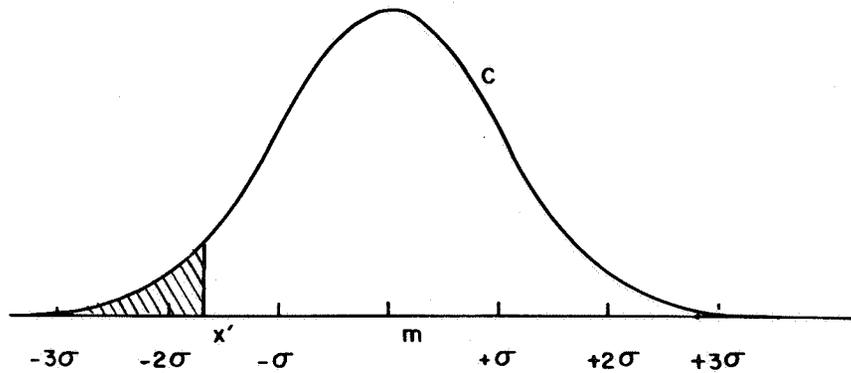


Fig. 2-1. A symmetrical distribution.

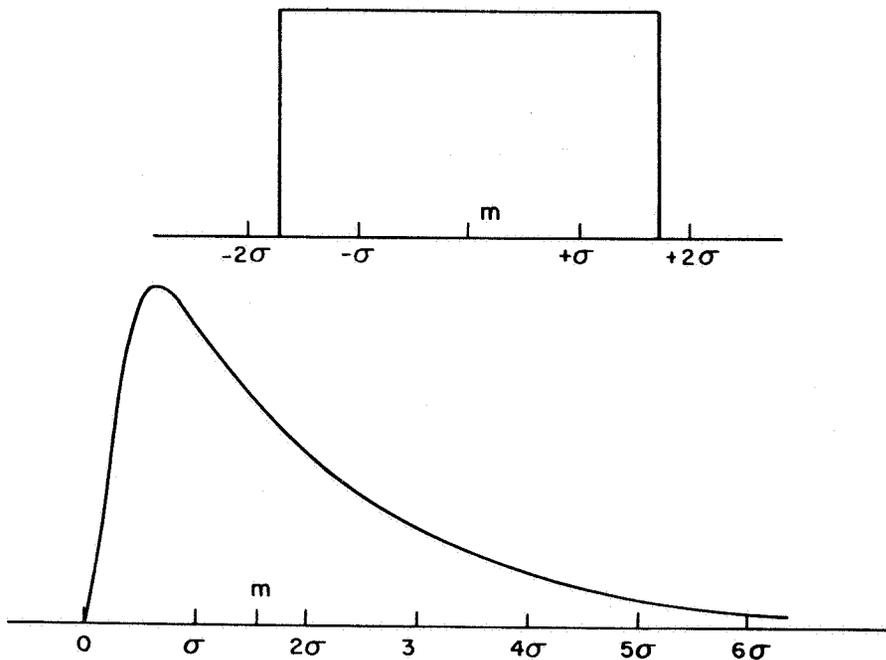


Fig. 2-2. (A) The uniform distribution (B) The log-normal distribution.

The Normal Distribution. For data generated by a measurement process, the following properties are usually observed:

1. The results spread roughly symmetrically about a central value.
2. Small deviations from this central value are more frequently found than large deviations.

A measurement process having these two properties would generate a frequency curve similar to that shown in Fig. 2-1 which is symmetrical and bunched together about m . The study of a particular theoretical representation of a frequency curve of this type leads to the celebrated bell-shaped normal curve (Gauss error curve.). Measurements having such a normal frequency curve are said to be normally distributed, or distributed in accordance with the normal law of error.

The normal curve can be represented exactly by the mathematical expression

$$y = \frac{1}{\sqrt{2\pi} \sigma} e^{-1/2[(x-m)^2/\sigma^2]} \tag{2-0}$$

where y is the ordinate and x the abscissa and $e \doteq 2.71828$ is the base of natural logarithms.

Some of the important features of the normal curve are:

1. It is symmetrical about m .
2. The area under the curve is one, as required.
3. If σ is used as unit on the abscissa, then the area under the curve between constant multiples of σ can be computed from tabulated values of the normal distribution. In particular, areas under the curve for some useful intervals between $m - k\sigma$ and $m + k\sigma$ are given in Table 2-1. Thus about two-thirds of the area lies within one σ of m , more than 95 percent within 2σ of m , and less than 0.3 percent beyond 3σ from m .

Table 2-1. Area under normal curve between $m - k\sigma$ and $m + k\sigma$

k :	0.6745	1.00	1.96	2.00	2.58	3.00
Percent area under curve (approx.):	50.0	68.3	95.0	95.5	99.0	99.7

4. From Eq. (2-0), it is evident that the frequency curve is completely determined by the two parameters m and σ .

The normal distribution has been studied intensively during the past century. Consequently, if the measurements follow a normal distribution, we can say a great deal about the measurement process. The question remains: How do we know that this is so from the limited number of repeated measurements on hand?

The answer is that we don't! However, in most instances the metrologist may be willing

1. to assume that the measurement process generates numbers that follow a normal distribution approximately, and act as if this were so,
2. to rely on the so-called Central Limit Theorem, one version of which is the following*: "If a population has a finite variance σ^2 and mean m , then the distribution of the sample mean (of n independent

*From Chapter 7, *Introduction to the Theory of Statistics*, by A. M. Mood, McGraw-Hill Book Company, New York, 1950.

measurements) approaches the normal distribution with variance σ^2/n and mean m as the sample size n increases." This remarkable and powerful theorem is indeed tailored for measurement processes. First, every measurement process must by definition have a finite mean and variance. Second, the sample mean \bar{x} is the quantity of interest which, according to the theorem, will be approximately normally distributed for large sample sizes. Third, the measure of dispersion, i.e., the standard deviation of the sample mean, is reduced by a factor of $1/\sqrt{n}$! This last statement is true in general for all measurement processes in which the measurements are "independent" and for all n . It is therefore not a consequence of the Central Limit Theorem. The theorem guarantees, however, that the distribution of sample means of *independent* measurements will be *approximately* normal with the specified limiting mean and standard deviation σ/\sqrt{n} for large n .

In fact, for a measurement process with a frequency curve that is symmetrical about the mean, and with small deviations from the mean as compared to the magnitude of the quantity measured, the normal approximation to the distribution of \bar{x} becomes very good even for n as small as 3 or 4. Figure 2-3 shows the uniform and normal distribution having the same mean and standard deviation. The peaked curve is actually two curves, representing the distribution of arithmetic means of four independent measurements from the respective distributions. These curves are indistinguishable to this scale.

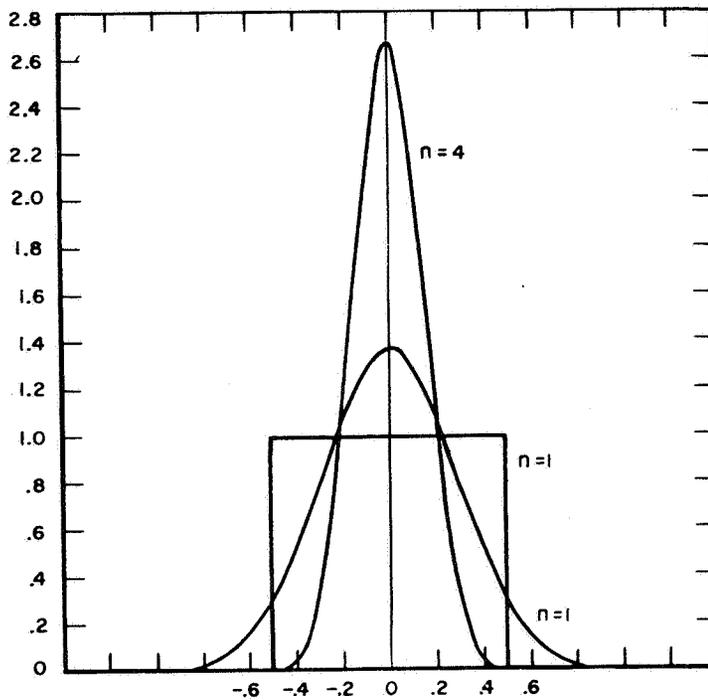


Fig. 2-3. Uniform and normal distribution of individual measurements having the same mean and standard deviation, and the corresponding distribution(s) of arithmetic means of four independent measurements.

A formal definition of the concept of "independence" is out of the scope here. Intuitively, we may say that n normally distributed measurements are independent if these measurements are not correlated or associated in any

Environmental Assessment for Houston Optimization of Airspace and Procedures in the Metroplex way. Thus, a sequence of measurements showing a trend or pattern are not independent measurements.

There are many ways by which dependence or correlation creeps into a set of measurement data; several of the common causes are the following:

1. Measurements are correlated through a factor that has not been considered, or has been considered to be of no appreciable effect on the results.
2. A standard correction constant has been used for a factor, e.g., temperature, but the constant may overcorrect or undercorrect for particular samples.
3. Measurements are correlated through time of the day, between days, weeks, or seasons.
4. Measurements are correlated through rejection of valid data, when the rejection is based on the size of the number in relation to others of the group.

The traditional way of plotting the data in the sequence they are taken, or in some rational grouping, is perhaps still the most effective way of detecting trends or correlation.

Estimates of Population Characteristics. In the above section it is shown that the limiting mean m and the variance σ^2 completely specify a measurement process that follows the normal distribution. In practice, m and σ^2 are not known and cannot be computed from a finite number of measurements. This leads to the use of the sample mean \bar{x} as an estimate of the limiting mean m and s^2 , the square of the computed standard deviation of the sample, as an estimate of the variance. The standard deviation of the average of n measurements, σ/\sqrt{n} , is sometimes referred to as the standard error of the mean, and is estimated by s/\sqrt{n} .

We note that the making of n independent measurements is equivalent to drawing a sample of size n at random from the population of measurements. Two concepts are of importance here:

1. The measurement process is established and under control, meaning that the limiting mean and the standard deviation do possess definite values which will not change over a reasonable period of time.
2. The measurements are randomly drawn from this population, implying that the values are of equal weights, and there is no prejudice in the method of selection. Suppose out of three measurements the one which is far apart from the other two is rejected, then the result will not be a random sample.

For a random sample we can say that \bar{x} is an unbiased estimate of m , and s^2 is an unbiased estimate of σ^2 , i.e., the limiting mean of \bar{x} is equal to m and of s^2 to σ^2 , where

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

and

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n-1} \left[\sum x_i^2 - \frac{(\sum x_i)^2}{n} \right]$$

In addition, we define

$$s = \sqrt{s^2} = \text{computed standard deviation}$$

Examples of numerical calculations of \bar{x} and s^2 and s are shown in Tables 2-5 and 2-6.

Environmental Assessment for Houston Optimization of Airspace and Procedures in the Metroplex
**Interpretation and Computation of
 Confidence Interval and Limits**

By making k sets of n measurements each, we can compute and arrange k , \bar{x} 's, and s 's in a tabular form as follows:

Set	Sample mean	Sample standard deviation
1	\bar{x}_1	s_1
2	\bar{x}_2	s_2
.	.	.
.	.	.
j	\bar{x}_j	s_j
.	.	.
.	.	.
k	\bar{x}_k	s_k

In the array of \bar{x} 's, no two will be likely to have exactly the same value. From the Central Limit Theorem it can be deduced that the \bar{x} 's will be approximately normally distributed with standard deviation σ/\sqrt{n} . The frequency curve of \bar{x} will be centered about the limiting mean m and will have the scale factor σ/\sqrt{n} . In other words, $\bar{x} - m$ will be centered about zero, and the quantity

$$z = \frac{\bar{x} - m}{\sigma/\sqrt{n}}$$

has the properties of a single observation from the "standardized" normal distribution which has a mean of zero and a standard deviation of one.

From tabulated values of the standardized normal distribution it is known that 95 percent of z values will be bounded between -1.96 and $+1.96$. Hence the statement

$$-1.96 < \frac{\bar{x} - m}{\sigma/\sqrt{n}} < +1.96$$

or its equivalent,

$$\bar{x} - 1.96 \frac{\sigma}{\sqrt{n}} < m < \bar{x} + 1.96 \frac{\sigma}{\sqrt{n}}$$

will be correct 95 percent of the time in the long run. The interval $\bar{x} - 1.96(\sigma/\sqrt{n})$ to $\bar{x} + 1.96(\sigma/\sqrt{n})$ is called a *confidence interval* for m . The probability that the confidence interval will cover the limiting mean, 0.95 in this case, is called the confidence level or confidence coefficient. The values of the end points of a confidence interval are called confidence limits. It is to be borne in mind that \bar{x} will fluctuate from set to set, and the interval calculated for a particular \bar{x} , may or may not cover m .

In the above discussion we have selected a two-sided interval symmetrical about \bar{x} . For such intervals the confidence coefficient is usually denoted by $1 - \alpha$, where $\alpha/2$ is the percent of the area under the frequency curve of z that is cut off from each tail.

In most cases, σ is not known and an estimate of σ is computed from the same set of measurements we use to calculate \bar{x} . Nevertheless, let us form a quantity similar to z , which is

$$t = \frac{\bar{x} - m}{s/\sqrt{n}}$$

Environmental Assessment for Houston Optimization of Airspace and Procedures in the Metroplex and if we know the distribution of t , we could make the same type of statement as before. In fact the distribution of t is known for the case of normally distributed measurements.

The distribution of t was obtained mathematically by William S. Gosset under the pen name of "Student," hence the distribution of t is called the Student's distribution. In the expression for t , both \bar{x} and s fluctuate from set to set of measurements. Intuitively we will expect the value of t to be larger than that of z for a statement with the same probability of being correct. This is indeed the case. The values of t are listed in Table 2-2.

Table 2-2. A brief table of values of t

Degrees of freedom ν	Confidence Level: $1 - \alpha$			
	0.500	0.900	0.950	0.990
1	1.000	6.314	12.706	63.657
2	.816	2.920	4.303	9.925
3	.765	2.353	3.182	5.841
4	.741	2.132	2.776	4.604
5	.727	2.015	2.571	4.032
6	.718	1.943	2.447	3.707
7	.711	1.895	2.365	3.499
10	.700	1.812	2.228	3.169
15	.691	1.753	2.131	2.947
20	.687	1.725	2.086	2.845
30	.683	1.697	2.042	2.750
60	.679	1.671	2.000	2.660
∞	.674	1.645	1.960	2.576

*Adapted from *Biometrika Tables for Statisticians*, Vol. I, edited by E. S. Pearson and H. O. Hartley, The University Press, Cambridge, 1958.

To find a value for t , we need to know the "degrees of freedom" (ν) associated with the computed standard deviation s . Since \bar{x} is calculated from the same n numbers and has a fixed value, the n th value of x_i is completely determined by \bar{x} and the other $(n - 1)x$ values. Hence the degrees of freedom here are $n - 1$.

Having the table for the distribution of t , and using the same reasoning as before, we can make the statement that

$$\bar{x} - t \frac{s}{\sqrt{n}} < m < \bar{x} + t \frac{s}{\sqrt{n}}$$

and our statement will be correct $100(1 - \alpha)$ percent of the time in the long run. The value of t depends on the degrees of freedom ν and the probability level. From the table, we get for a confidence level of 0.95, the following lower and upper confidence limits:

ν	$L_l = \bar{x} - t(s/\sqrt{n})$	$L_u = \bar{x} + t(s/\sqrt{n})$
1	$\bar{x} - 12.706(s/\sqrt{n})$	$\bar{x} + 12.706(s/\sqrt{n})$
2	$\bar{x} - 4.303(s/\sqrt{n})$	$\bar{x} + 4.303(s/\sqrt{n})$
3	$\bar{x} - 3.182(s/\sqrt{n})$	$\bar{x} + 3.182(s/\sqrt{n})$

The value of t for $\nu = \infty$ is 1.96, the same as for the case of known σ . Notice that very little can be said about m with two measurements. However, for n larger than 2, the interval predicted to contain m narrows down steadily, due to both the smaller value of t and the divisor \sqrt{n} .

It is probably worthwhile to emphasize again that each particular confidence interval computed as a result of n measurements will either include m or fail to include m . The probability statement refers to the fact that if we make a long series of sets of n measurements, and if we compute a confidence interval for m from each set by the prescribed method, we would expect 95 percent of such intervals to include m .

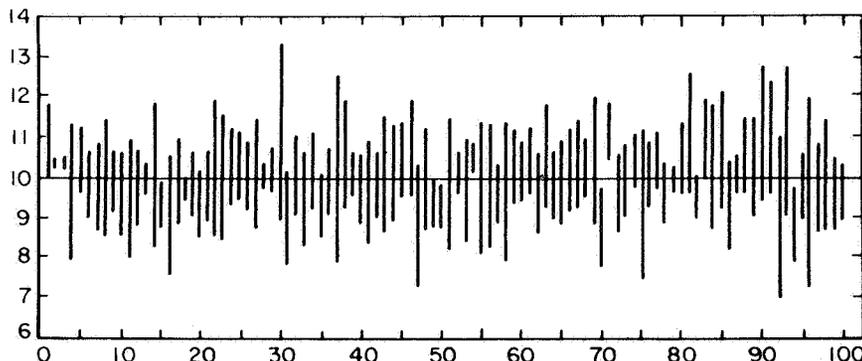


Fig. 2-4. Computed 90% confidence intervals for 100 samples of size 4 drawn at random from a normal population with $m = 10$, $\sigma = 1$.

Figure 2-4 shows the 90 percent confidence intervals ($P = 0.90$) computed from 100 samples of $n = 4$ from a normal population with $m = 10$, and $\sigma = 1$. Three interesting features are to be noted:

1. The number of intervals that include m actually turns out to be 90, the expected number.
2. The surprising variation of the sizes of these intervals.
3. The closeness of the mid-points of these intervals to the line for the mean does not seem to be related to the spread. In samples No. 2 and No. 3, the four values must have been very close together, but both of these intervals failed to include the line for the mean.

From the widths of computed confidence intervals, one may get an intuitive feeling whether the number of measurements n is reasonable and sufficient for the purpose on hand. It is true that, even for small n , the confidence intervals will cover the limiting mean with the specified probability, yet the limits may be so far apart as to be of no practical significance. For detecting a specified magnitude of interest, e.g., the difference between two means, the approximate number of measurements required can be solved by equating the half-width of the confidence interval to this difference and solving for n , using σ when known, or using s by trial and error if σ is not known. Tables of sample sizes required for certain prescribed conditions are given in reference 4.

Precision and Accuracy

Index of Precision. Since σ is a measure of the spread of the frequency curve about the limiting mean, σ may be defined as an index of precision. Thus a measurement process with a standard deviation σ_1 is said to be more precise than another with a standard deviation σ_2 if σ_1 is smaller than σ_2 . (In fact, σ is really a measure of imprecision since the imprecision is directly proportional to σ .)

Consider the means of sets of n independent measurements as a new derived measurement process. The standard deviation of the new process is σ/\sqrt{n} . It is therefore possible to derive from a less precise measurement process a new process which has a standard deviation equal to that of a more precise process. This is accomplished by making more measurements.

Suppose $m_1 = m_2$, but $\sigma_1 = 2\sigma_2$. Then for a derived process to have $\sigma'_1 = \sigma_2$, we need

$$\sigma'_1 = \frac{\sigma_1}{\sqrt{n}} = \frac{2\sigma_2}{\sqrt{4}}$$

or we need to use the average of four measurements as a single measurement. Thus for a required degree of precision, the number of measurements, n_1 and n_2 , needed for measurement processes I and II is proportional to the squares of their respective standard deviations (variances), or in symbols

$$\frac{n_1}{n_2} = \frac{\sigma_1^2}{\sigma_2^2}$$

If σ is not known, and the best estimate we have of σ is a computed standard deviation s based on n measurements, then s could be used as an estimate of the index of precision. The value of s , however, may vary considerably from sample to sample in the case of a small number of measurements as was shown in Fig. 2-4, where the lengths of the intervals are constant multiples of s computed from the samples. The number n or the degrees of freedom ν must be considered along with s in indicating how reliable an estimate s is of σ . In what follows, whenever the terms standard deviation about the limiting mean (σ), or standard error of the mean ($\sigma_{\bar{x}}$), are used, the respective estimates s and s/\sqrt{n} may be substituted, by taking into consideration the above reservation.

In metrology or calibration work, the precision of the reported value is an integral part of the result. In fact, precision is the main criterion by which the quality of the work is judged. Hence, the laboratory reporting the value must be prepared to give evidence of the precision claimed. Obviously an estimate of the standard deviation of the measurement process based only on a small number of measurements cannot be considered as convincing evidence. By the use of the control chart method for standard deviation and by the calibration of one's own standard at frequent intervals, as subsequently described, the laboratory may eventually claim that the standard deviation is in fact known and the measurement process is stable, with readily available evidence to support these claims.

Interpretation of Precision. Since a measurement process generates numbers as the results of repeated measurements of a single physical quantity under essentially the same conditions, the method and procedure in obtaining these numbers must be specified in detail. However, no amount of detail would cover all the contingencies that may arise, or cover all the factors that may affect the results of measurement. Thus a single operator in a single day with a single instrument may generate a process with a precision index measured by σ . Many operators measuring the same quantity over a period of time with a number of instruments will yield a precision index measured by σ' . Logically σ' must be larger than σ , and in practice it is usually considerably larger. Consequently, modifiers of the words "precision" are recommended by ASTM* to qualify in an unambiguous manner what

*"Use of the Terms Precision and Accuracy as Applied to the Measurement of a Property of a Material," ASTM Designation, E177-61T, 1961.

Environmental Assessment for Houston Optimization of Airspace and Procedures in the Metroplex is meant. Examples are "single-operator-machine," "multi-laboratory," "single-operator-day," etc. The same publication warns against the use of the terms "repeatability" and "reproducibility" if the interpretation of these terms is not clear from the context.

The standard deviation σ or the standard error σ/\sqrt{n} can be considered as a yardstick with which we can gage the difference between two results obtained as measurements of the same physical quantity. If our interest is to compare the results of one operator against another, the single-operator precision is probably appropriate, and if the two results differ by an amount considered to be large as measured by the standard errors, we may conclude that the evidence is predominantly against the two results being truly equal. In comparing the results of two laboratories, the single-operator precision is obviously an inadequate measure to use, since the precision of each laboratory must include factors such as multi-operator-day-instruments.

Hence the selection of an index of precision depends strongly on the purposes for which the results are to be used or might be used. It is common experience that three measurements made within the hour are closer together than three measurements made on, say, three separate days. However, an index of precision based on the former is generally not a justifiable indicator of the quality of the reported value. For a thorough discussion on the *realistic* evaluation of precision see Section 4 of reference 2.

Accuracy. The term "accuracy" usually denotes in some sense the closeness of the measured values to the true value, taking into consideration both precision and bias. Bias, defined as the difference between the limiting mean and the true value, is a constant, and does not behave in the same way as the index of precision, the standard deviation. In many instances, the possible sources of biases are known but their magnitudes and directions are not known. The overall bias is of necessity reported in terms of estimated bounds that reasonably include the combined effect of all the elemental biases. Since there are no accepted ways to estimate bounds for elemental biases, or to combine them, these should be reported and discussed in sufficient detail to enable others to use their own judgment on the matter.

It is recommended that an index of accuracy be expressed as a pair of numbers, one the credible bounds for bias, and the other an index of precision, usually in the form of a multiple of the standard deviation (or estimated standard deviation). The terms "uncertainty" and "limits of error" are sometimes used to express the sum of these two components, and their meanings are ambiguous unless the components are spelled out in detail.

STATISTICAL ANALYSIS OF MEASUREMENT DATA

In the last section the basic concepts of a measurement process were given in an expository manner. These concepts, necessary to the statistical analysis to be presented in this section, are summarized and reviewed below. By making a measurement we obtain a number intended to express quantitatively a measure of "the property of a thing." Measurement numbers differ from ordinary arithmetic numbers, and the usual "significant figure" treatment is not appropriate. Repeated measurement of a single physical

quantity under essentially the same conditions generates a sequence of numbers x_1, x_2, \dots, x_n . A measurement process is established if this conceptually infinite sequence has a limiting mean m and a standard deviation σ .

For many measurement processes encountered in metrology, the sequence of numbers generated follows approximately the normal distribution, specified completely by the two quantities m and σ . Moreover, averages of n independent measurement numbers tend to be normally distributed with the limiting mean m and the standard deviation σ/\sqrt{n} , regardless of the distribution of the original numbers. Normally distributed measurements are independent if they are not correlated or associated in any way. A sequence of measurements showing a trend or pattern are not independent measurements. Since m and σ are usually not known, these quantities are estimated by calculating \bar{x} and s from n measurements, where

$$\bar{x} = \frac{1}{n} \sum_1^n x_i$$

and

$$s = \sqrt{\frac{1}{n-1} \sum_1^n (x_i - \bar{x})^2} = \sqrt{\frac{1}{n-1} \left[\sum_1^n x_i^2 - \frac{(\sum x_i)^2}{n} \right]}$$

The distribution of the quantity $t = (\bar{x} - m)/(s/\sqrt{n})$ (for x normally distributed) is known. From the tabulated values of t (see Table 2-2), confidence intervals can be constructed to bracket m for a given confidence coefficient $1 - \alpha$ (probability of being correct in the long run).

The confidence limits are the end points of confidence intervals defined by

$$L_l = \bar{x} - t \frac{s}{\sqrt{n}}$$

$$L_u = \bar{x} + t \frac{s}{\sqrt{n}}$$

where the value of t is determined by two parameters, namely, the degrees of freedom ν associated with s and the confidence coefficient $1 - \alpha$.

The width of a confidence interval gives an intuitive measure of the uncertainty of the evidence given by the data. Too wide an interval may merely indicate that more measurements need to be made for the objective desired.

Algebra for the Manipulation of Limiting Means and Variances

Basic Formulas. A number of basic formulas are extremely useful in dealing with a quantity which is a combination of other measured quantities.

1. Let m_x and m_y be the respective limiting means of two measured quantities X and Y , and a, b be constants, then

$$\left. \begin{aligned} m_{x+y} &= m_x + m_y \\ m_{x-y} &= m_x - m_y \\ m_{ax+by} &= am_x + bm_y \end{aligned} \right\} \quad (2-1)$$

2. If, in addition, X and Y are independent, then it is also true that

$$m_{xy} = m_x m_y \quad (2-2)$$

For paired values of X and Y , we can form the quantity Z , with

$$Z = (X - m_x)(Y - m_y) \quad (2-3)$$

Then by formula (2-2) for independent variables,

$$\begin{aligned} m_z &= m_{(x-m_x)}m_{(y-m_y)} \\ &= (m_x - m_x)(m_y - m_y) = 0 \end{aligned}$$

Thus $m_z = 0$ when X and Y are independent.

3. The limiting mean of Z in (2-3) is defined as the covariance of X and Y and is usually denoted by $\text{cov}(X, Y)$, or σ_{xy} . The covariance, similar to the variance, is estimated by

$$s_{xy} = \frac{1}{n-1} \sum (x_i - \bar{x})(y_i - \bar{y}) \quad (2-4)$$

Thus if X and Y are correlated in such a way that paired values are likely to be both higher or lower than their respective means, then s_{xy} tends to be positive. If a high x value is likely to be paired with a low y value, and vice versa, then s_{xy} tends to be negative. If X and Y are not correlated, s_{xy} tends to zero (for large n).

4. The correlation coefficient ρ is defined as:

$$\rho = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \quad (2-5)$$

and is estimated by

$$r = \frac{s_{xy}}{s_x s_y} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \quad (2-6)$$

Both ρ and r lie between -1 and $+1$.

5. Let σ_x^2 and σ_y^2 be the respective variances of X and Y , and σ_{xy} the covariance of X and Y , then

$$\begin{aligned} \sigma_{x+y}^2 &= \sigma_x^2 + \sigma_y^2 + 2\sigma_{xy} \\ \sigma_{x-y}^2 &= \sigma_x^2 + \sigma_y^2 - 2\sigma_{xy} \end{aligned} \quad (2-7)$$

If X and Y are independent, $\sigma_{xy} = 0$, then

$$\sigma_{x+y}^2 = \sigma_x^2 + \sigma_y^2 = \sigma_{x-y}^2 \quad (2-8)$$

Since the variance of a constant is zero, we have

$$\begin{aligned} \sigma_{ax+b}^2 &= a^2 \sigma_x^2 \\ \sigma_{ax+by}^2 &= a^2 \sigma_x^2 + b^2 \sigma_y^2 + 2ab\sigma_{xy} \end{aligned} \quad (2-9)$$

In particular, if X and Y are independent and normally distributed, then $aX + bY$ is normally distributed with limiting mean $am_x + bm_y$ and variance $a^2\sigma_x^2 + b^2\sigma_y^2$.

For measurement situations in general, metrologists usually strive to get measurements that are independent, or can be assumed to be independent. The case when two quantities are dependent because both are functions of other measured quantities will be treated under propagation of error formulas (see Eq. 2-13).

6. Standard errors of the sample mean and the weighted means (of independent measurements) are special cases of the above. Since $\bar{x} = (1/n) \sum x_i$ and the x_i 's are independent with variance σ_x^2 , it follows, by (2-9), that

$$\sigma_{\bar{x}}^2 = \left(\frac{1}{n}\right)^2 \sigma_{x_1}^2 + \left(\frac{1}{n}\right)^2 \sigma_{x_2}^2 + \dots + \left(\frac{1}{n}\right)^2 \sigma_{x_n}^2 = \frac{\sigma_x^2}{n} \quad (2-10)$$

as previously stated.

If \bar{x}_1 is an average of k values, and \bar{x}_2 is an average of n values, then for the over-all average, \bar{x} , it is logical to compute

$$\bar{x} = \frac{x_1 + \cdots + x_k + x_{k+1} + \cdots + x_{k+n}}{k + n}$$

and $\sigma_{\bar{x}}^2 = \sigma_x^2 / (k + n)$. However, this is equivalent to a weighted mean of \bar{x}_1 and \bar{x}_2 , where the weights are proportional to the number of measurements in each average, i.e.,

$$w_1 = k, \quad w_2 = n$$

and

$$\begin{aligned} \bar{x} &= \left(\frac{w_1}{w_1 + w_2} \right) \bar{x}_1 + \left(\frac{w_2}{w_1 + w_2} \right) \bar{x}_2 \\ &= \frac{k}{n + k} \bar{x}_1 + \frac{n}{n + k} \bar{x}_2 \end{aligned}$$

Since

$$\frac{\sigma_{\bar{x}_1}^2}{\sigma_{\bar{x}_2}^2} = \frac{\sigma^2/k}{\sigma^2/n} = \frac{n}{k} = \frac{w_2}{w_1}$$

the weighting factors w_1 and w_2 are therefore also inversely proportional to the respective variances of the averages. This principle can be extended to more than two variables in the following manner.

Let $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_k$ be a set of averages estimating the same quantity. The over-all average may be computed to be

$$\bar{x} = \frac{1}{w_1 + w_2 + \cdots + w_k} (w_1 \bar{x}_1 + w_2 \bar{x}_2 + \cdots + w_k \bar{x}_k)$$

where

$$w_1 = \frac{1}{\sigma_{\bar{x}_1}^2}, \quad w_2 = \frac{1}{\sigma_{\bar{x}_2}^2}, \quad \dots, \quad w_k = \frac{1}{\sigma_{\bar{x}_k}^2}$$

The variance of \bar{x} is, by (2-9),

$$\sigma_{\bar{x}}^2 = \frac{1}{w_1 + w_2 + \cdots + w_k} \quad (2-11)$$

In practice, the estimated variances $s_{\bar{x}}^2$ will have to be used in the above formulas, and consequently the equations hold only as approximations.

Propagation of error formulas. The results of a measurement process can usually be expressed by a number of averages \bar{x}, \bar{y}, \dots , and the standard errors of these averages $s_{\bar{x}} = s_x / \sqrt{n}, s_{\bar{y}} = s_y / \sqrt{k}$, etc. These results, however, may not be of direct interest; the quantity of interest is in the functional relationship $m_w = f(m_x, m_y)$. It is desired to estimate m_w by $\bar{w} = f(\bar{x}, \bar{y})$ and to compute $s_{\bar{w}}$ as an estimate of $\sigma_{\bar{w}}$.

If the errors of measurements of these quantities are small in comparison with the values measured, the propagation of error formulas usually work surprisingly well. The $\sigma_{\bar{w}}^2, \sigma_{\bar{x}}^2$, and $\sigma_{\bar{y}}^2$ that are used in the following formulas will often be replaced in practice by the computed values $s_{\bar{w}}^2, s_{\bar{x}}^2$, and $s_{\bar{y}}^2$.

The general formula for $\sigma_{\bar{w}}^2$ is given by

$$\sigma_{\bar{w}}^2 = \left[\frac{\partial f}{\partial x} \right]^2 \sigma_{\bar{x}}^2 + \left[\frac{\partial f}{\partial y} \right]^2 \sigma_{\bar{y}}^2 + 2 \left[\frac{\partial f}{\partial x} \right] \left[\frac{\partial f}{\partial y} \right] \rho_{\bar{x}\bar{y}} \sigma_{\bar{x}} \sigma_{\bar{y}} \quad (2-12)$$

where the partial derivatives in square brackets are to be evaluated at the averages of x and y . If X and Y are independent, $\rho = 0$ and therefore the last term equals zero. If X and Y are measured in pairs, $s_{\bar{x}\bar{y}}$ (Eq. 2-4) can be used as an estimate of $\rho_{\bar{x}\bar{y}} \sigma_{\bar{x}} \sigma_{\bar{y}}$.

If W is functionally related to U and V by

$$m_w = f(m_u, m_v)$$

and both U and V are functionally related to X and Y by

$$m_u = g(m_x, m_y)$$

$$m_v = h(m_x, m_y)$$

then U and V are functionally related. We will need the covariance $\sigma_{uv} = \rho_{uv}\sigma_u\sigma_v$ to calculate σ_w^2 . The covariance σ_{uv} is given approximately by

$$\begin{aligned} \sigma_{uv} = & \left[\frac{\partial g}{\partial x} \cdot \frac{\partial h}{\partial x} \right] \sigma_x^2 + \left[\frac{\partial g}{\partial y} \cdot \frac{\partial h}{\partial y} \right] \sigma_y^2 \\ & + \left\{ \left[\frac{\partial g}{\partial x} \cdot \frac{\partial h}{\partial y} \right] + \left[\frac{\partial g}{\partial y} \cdot \frac{\partial h}{\partial x} \right] \right\} \rho_{xy} \sigma_x \sigma_y \end{aligned} \quad (2-13)$$

The square brackets mean, as before, that the partial derivatives are to be evaluated at \bar{x} and \bar{y} . If X and Y are independent, the last term again vanishes.

These formulas can be extended to three or more variables if necessary. For convenience, a few special formulas for commonly encountered functions are listed in Table 2-3 with X, Y assumed to be independent. These may be derived from the above formulas as exercises.

Table 2-3. Propagation of error formulas for some simple functions

(X and Y are assumed to be independent.)

Function form	Approximate formula for s_w^2
$m_w = Am_x + Bm_y$	$A^2s_x^2 + B^2s_y^2$
$m_w = \frac{m_x}{m_y}$	$\left(\frac{\bar{x}}{\bar{y}}\right)^2 \left(\frac{s_x^2}{\bar{x}^2} + \frac{s_y^2}{\bar{y}^2}\right)$
$m_w = \frac{1}{m_y}$	$\frac{s_y^2}{\bar{y}^4}$
$m_w = \frac{m_x}{m_x + m_y}$	$\left(\frac{\bar{w}}{\bar{x}}\right)^4 (\bar{y}^2 s_x^2 + \bar{x}^2 s_y^2)$
$m_w = \frac{m_x}{1 + m_x}$	$\frac{s_x^2}{(1 + \bar{x})^4}$
$*m_w = m_x m_y$	$(\bar{x}\bar{y})^2 \left(\frac{s_x^2}{\bar{x}^2} + \frac{s_y^2}{\bar{y}^2}\right)$
$*m_w = m_x^2$	$4\bar{x}^2 s_x^2$
$m_w = \sqrt{m_x}$	$\frac{1}{4} \frac{s_x^2}{\bar{x}}$
$*m_w = \ln m_x$	$\frac{s_x^2}{\bar{x}^2}$
$*m_w = km_x^a m_y^b$	$\bar{w}^2 \left(a^2 \frac{s_x^2}{\bar{x}^2} + b^2 \frac{s_y^2}{\bar{y}^2} \right)$
$*m_w = e^{m_x}$	$e^{2\bar{x}} s_x^2$
$W = 100 \frac{s_x}{\bar{x}}$ (=coefficient of variation)	$\frac{\bar{w}^2}{2(n-1)}$ (not directly derived from the formulas)†

*Distribution of \bar{w} is highly skewed and normal approximation could be seriously in error for small n .

†See, for example, *Statistical Theory with Engineering Applications*, by A. Hald, John Wiley & Sons, Inc., New York, 1952, p. 301.

In these formulas, if

- (a) the partial derivatives when evaluated at the averages are small, and
- (b) σ_x, σ_y are small compared to \bar{x}, \bar{y} ,

then the approximations are good and \bar{w} tends to be distributed normally (the ones marked by asterisks are highly skewed and normal approximation could be seriously in error for small n).

Pooling Estimates of Variances. The problem often arises that there are several estimates of a common variance σ^2 which we wish to combine into a single estimate. For example, a gage block may be compared with the master block n_1 times, resulting in an estimate of the variance s_1^2 . Another gage block compared with the master block n_2 times, giving rise to s_2^2 , etc. As long as the nominal thicknesses of these blocks are within a certain range, the precision of calibration can be expected to remain the same. To get a better evaluation of the precision of the calibration process, we would wish to combine these estimates. The rule is to combine the computed variances weighted by their respective degrees of freedom, or

$$s_p^2 = \frac{\nu_1 s_1^2 + \nu_2 s_2^2 + \dots + \nu_k s_k^2}{\nu_1 + \nu_2 + \dots + \nu_k} \quad (2-14)$$

The pooled estimate of the standard deviation, of course, is $\sqrt{s_p^2} = s_p$. In the example, $\nu_1 = n_1 - 1$, $\nu_2 = n_2 - 1, \dots, \nu_k = n_k - 1$, thus the expression reduces to

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \dots + (n_k - 1)s_k^2}{n_1 + n_2 + \dots + n_k - k} \quad (2-15)$$

The degrees of freedom for the pooled estimate is the sum of the degrees of freedom of individual estimates, or $\nu_1 + \nu_2 + \dots + \nu_k = n_1 + n_2 + \dots + n_k - k$. With the increased number of degrees of freedom, s_p is a more dependable estimate of σ than an individual s . Eventually, we may consider the value of s_p to be equal to that of σ and claim that we know the precision of the measuring process.

For the special case where k sets of duplicate measurements are available, the above formula reduces to:

$$s_p^2 = \frac{1}{2k} \sum_1^k d_i^2 \quad (2-16)$$

where d_i = difference of duplicate readings. The pooled standard deviation s_p has k degrees of freedom.

For sets of normally distributed measurements where the number of measurements in each set is small, say less than ten, an estimate of the standard deviation can be obtained by multiplying the range of these measurements by a constant. Table 2-4 lists these constants corresponding to the number n of measurements in the set. For large n , considerable information is lost and this procedure is not recommended.

If there are k sets of n measurements each, the average range \bar{R} can be computed. The standard deviation can be estimated by multiplying the average range by the factor for n .

n	Multiplying factor
2	0.886
3	0.591
4	0.486
5	0.430
6	0.395
7	0.370
8	0.351
9	0.337
10	0.325

*Adapted from *Biometrika Tables for Statisticians*, Vol. I, edited by E. S. Pearson and H. O. Hartley, The University Press, Cambridge, 1958.

Component of Variance Between Groups. In pooling estimates of variances from a number of subgroups, we have increased confidence in the value of the estimate obtained. Let us call this estimate the within-group standard deviation, σ_w . The within-group standard deviation σ_w is a proper measure of dispersions of values within the same group, but not necessarily the proper one for dispersions of values belonging to different groups.

If in making calibrations there is a difference between groups, say from day to day, or from set to set, then the limiting means of the groups are not equal. These limiting means may be thought of as individual measurements; thus, it could be assumed that the average of these limiting means will approach a limit which can be called the limiting mean for all the groups. In estimating σ_w^2 , the differences of individuals from the respective group means are used. Obviously σ_w does not include the differences between groups. Let us use σ_b^2 to denote the variance corresponding to the differences between groups, i.e., the measure of dispersion of the limiting means of the respective groups about the limiting mean for all groups.

Thus for each individual measurement x , the variance of X has two components, and

$$\sigma^2 = \sigma_b^2 + \sigma_w^2$$

For the group mean \bar{x} with n measurements in the group,

$$\sigma_{\bar{x}}^2 = \sigma_b^2 + \frac{\sigma_w^2}{n}$$

If k groups of n measurements are available giving averages $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_k$, then an estimate of $\sigma_{\bar{x}}^2$ is

$$s_{\bar{x}}^2 = \frac{1}{k-1} \sum_{i=1}^k (\bar{x}_i - \bar{\bar{x}})^2$$

with $k - 1$ degrees of freedom, where $\bar{\bar{x}}$ is the average of all nk measurements.

The resolution of the total variance into components attributable to identifiable causes or factors and the estimation of such components of variances are topics treated under analysis of variance and experimental design. For selected treatments and examples see references 5, 6, and 8.

Comparison of means is perhaps one of the most frequently used techniques in metrology. The mean obtained from one measurement process may be compared with a standard value; two series of measurements on the same quantity may be compared; or sets of measurements on more than two quantities may be compared to determine homogeneity of the group of means.

It is to be borne in mind in all of the comparisons discussed below, that we are interested in comparing the limiting means. The sample means and the computed standard errors are used to calculate confidence limits on the difference between two means. The "t" statistic derived from normal distribution theory is used in this procedure since we are assuming either the measurement process is normal, or the sample averages are approximately normally distributed.

Comparison of a Mean with a Standard Value. In calibration of weights at the National Bureau of Standards, the weights to be calibrated are intercompared with sets of standard weights having "accepted" corrections. Accepted corrections are based on years of experience and considered to be exact to the accuracy required. For instance, the accepted correction for the NB'10 gram weight is -0.4040 mg.

The NB'10 is treated as an unknown and calibrated with each set of weights tested using an intercomparison scheme based on a 100-gm standard weight. Hence the observed correction for NB'10 can be computed for each particular calibration. Table 2-5 lists eleven observed corrections of NB'10 during May 1963.

Calculated 95 percent confidence limits from the eleven observed corrections are -0.4041 and -0.3995 . These values include the accepted value of -0.4040 , and we conclude that the observed corrections agree with the accepted value.

What if the computed confidence limits for the observed correction do not cover the accepted value? Three explanations may be suggested:

1. The accepted value is correct. However, in choosing $\alpha = 0.05$, we know that 5 percent of the time in the long run we will make an error in our statement. By chance alone, it is possible that this particular set of limits would not cover the accepted value.
2. The average of the observed corrections does not agree with the accepted value because of certain systematic error, temporary or seasonal, particular to one or several members of this set of data for which no adjustment has been made.
3. The accepted value is incorrect, e.g., the mass of the standard has changed.

In our example, we would be extremely reluctant to agree to the third explanation since we have much more confidence in the accepted value than the value based only on eleven calibrations. We are warned that something may have gone wrong, but not unduly alarmed since such an event will happen purely by chance about once every twenty times.

The control chart for mean with known value, to be discussed in a following section, would be the proper tool to use to monitor the constancy of the correction of the standard mass.

Table 2-5. Computation of confidence limits for observed corrections, NB'10 gm *

Date	<i>i</i>	X_i Observed Corrections to standard 10 gm wt in mg
5-1-63	1	-0.4008
5-1-63	2	-0.4053
5-1-63	3	-0.4022
5-2-63	4	-0.4075
5-2-63	5	-0.3994
5-3-63	6	-0.3986
5-6-63	7	-0.4015
5-6-63	8	-0.3992
5-6-63	9	-0.3973
5-7-63	10	-0.4071
5-7-63	11	-0.4012
		$\sum x_i = -4.4201$
		$\bar{x} = -0.40183 \text{ mg}$
		$\sum x_i^2 = 1.77623417$
		$\frac{(\sum x_i)^2}{n} = 1.77611673$
		difference = 0.00011744

$$s^2 = \frac{1}{n-1}(0.00011744) = 0.000011744$$

$s = 0.00343 =$ computed standard deviation of an observed correction about the mean.

$$\frac{s}{\sqrt{n}} = 0.00103 = \text{computed standard deviation of the mean of eleven corrections.}$$

$=$ computed standard error of the mean.

For a two-sided 95 percent confidence interval for the mean of the above sample of size 11, $\alpha/2 = 0.025$, $\nu = 10$, and the corresponding value of t is equal to 2.228 in the table of t distribution. Therefore,

$$L_l = \bar{x} - t \frac{s}{\sqrt{n}} = -0.40183 - 2.228 \times 0.00103 = -0.40412$$

and

$$L_u = \bar{x} + t \frac{s}{\sqrt{n}} = -0.40183 + 2.228 \times 0.00103 = -0.39954$$

*Data supplied by Robert Raybold, Metrology Division, National Bureau of Standards.

Comparison Among Two or More Means. The difference between two quantities X and Y to be measured is the quantity

$$m_{x-y} = m_x - m_y$$

and is estimated by $\bar{x} - \bar{y}$, where \bar{x} and \bar{y} are averages of a number of measurements of X and Y respectively.

Suppose we are interested in knowing whether the difference m_{x-y} could be zero. This problem can be solved by a technique previously introduced, i.e., the confidence limits can be computed for m_{x-y} , and if the upper and lower limits include zero, we could conclude that m_{x-y} may take the value zero; otherwise, we conclude that the evidence is against $m_{x-y} = 0$.

Let us assume that measurements of X and Y are independent with known variances σ_x^2 and σ_y^2 respectively.

By Eq. (2.10)

$$\sigma_{\bar{x}}^2 = \frac{\sigma_x^2}{n} \text{ for } \bar{x} \text{ of } n \text{ measurements}$$

$$\sigma_{\bar{y}}^2 = \frac{\sigma_y^2}{k} \text{ for } \bar{y} \text{ of } k \text{ measurements}$$

then by (2.8),

$$\sigma_{\bar{x}-\bar{y}}^2 = \frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{k}$$

Therefore, the quantity

$$z = \frac{(\bar{x} - \bar{y}) - 0}{\sqrt{\frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{k}}} \quad (2-17)$$

is approximately normally distributed with mean zero and a standard deviation of one under the assumption $m_{x-y} = 0$.

If σ_x and σ_y are not known, but the two can be assumed to be approximately equal, e.g., \bar{x} and \bar{y} are measured by the same process, then s_x^2 and s_y^2 can be pooled by Eq. (2-15), or

$$s_p^2 = \frac{(n-1)s_x^2 + (k-1)s_y^2}{n+k-2}$$

This pooled computed variance estimates

$$\sigma^2 = \sigma_x^2 = \sigma_y^2$$

so that

$$\sigma_{\bar{x}-\bar{y}}^2 = \frac{\sigma_x^2}{n} + \frac{\sigma_y^2}{k} = \frac{n+k}{nk} \sigma^2$$

Thus, the quantity

$$t = \frac{(\bar{x} - \bar{y}) - 0}{\sqrt{\frac{n+k}{nk} s_p^2}} \quad (2-18)$$

is distributed as Student's "t", and a confidence interval can be set about m_{x-y} with $\nu = n + k - 2$ and $p = 1 - \alpha$. If this interval does not include zero, we may conclude that the evidence is strongly against the hypothesis $m_x = m_y$.

As an example, we continue with the calibration of weights with NB'10 gm. For 11 subsequent observed corrections during September and October, the confidence interval (computed in the same manner as in the preceding example) has been found to be

$$L_t = -0.40782$$

$$L_u = -0.40126$$

Also,

$$\bar{Y} = -0.40454 \quad \text{and} \quad \frac{s}{\sqrt{k}} = 0.00147$$

It is desired to compare the means of observed corrections for the two sets of data. Here

$$n = k = 11$$

$$\bar{x} = -0.40183, \quad \bar{y} = -0.40454$$

$$s_x^2 = 0.000011669, \quad s_y^2 = 0.000023813$$

$$s_p^2 = \frac{1}{2}(0.000035482) = 0.000017741$$

$$\frac{n+k}{nk} = \frac{11+11}{121} = \frac{2}{11}$$

$$\sqrt{\frac{n+k}{nk} s_p^2} = \sqrt{\frac{2}{11} \times 0.000017741} = 0.00180$$

For $\alpha/2 = 0.025$, $1 - \alpha = 0.95$, and $\nu = 20$, $t = 2.086$. Therefore,

$$L_u = (\bar{x} - \bar{y}) + t\sqrt{\frac{n+k}{nk}} s_p = 0.00271 + 2.086 \times 0.00180 = 0.00646$$

$$L_l = (\bar{x} - \bar{y}) - t\sqrt{\frac{n+k}{nk}} s_p = -0.00104$$

Since $L_l < 0 < L_u$ shows that the confidence interval includes zero, we conclude that there is no evidence against the hypothesis that the two observed average corrections are the same, or $m_x = m_y$. Note, however, that we would reach a conclusion of no difference wherever the magnitude of $\bar{x} - \bar{y}$ (0.00271 mg) is less than the half-width of the confidence interval ($2.086 \times 0.00180 = 0.00375$ mg) calculated for the particular case. When the true difference m_{x-y} is large, the above situation is not likely to happen; but when the true difference is small, say about 0.003 mg, then it is highly probable that a conclusion of no difference will still be reached. If a detection of difference of this magnitude is of interest, more measurements will be needed.

The following additional topics are treated in reference 4.

1. Sample sizes required under certain specified conditions—Tables A-8 and A-9.
2. σ_x^2 cannot be assumed to be equal to σ_y^2 —Section 3-3.1.2.
3. Comparison of several means by Studentized range—Sections 3-4 and 15-4.

Comparison of variances or ranges. As we have seen, the precision of a measurement process can be expressed in terms of the computed standard deviation, the variance, or the range. To compare the precision of two processes a and b , any of the three measures can be used, depending on the preference and convenience of the user.

Let s_a^2 be the estimate of σ_a^2 with ν_a degrees of freedom, and s_b^2 be the estimate of σ_b^2 with ν_b degrees of freedom. The ratio $F = s_a^2/s_b^2$ has a distribution depending on ν_a and ν_b . Tables of upper percentage points of F are given in most statistical textbooks, e.g., reference 4, Table A-5 and Section 4-2.

In the comparison of means, we were interested in finding out if the absolute difference between m_a and m_b could reasonably be zero; similarly, here we may be interested in whether $\sigma_a^2 = \sigma_b^2$, or $\sigma_a^2/\sigma_b^2 = 1$. In practice, however, we are usually concerned with whether the imprecision of one process exceeds that of another process. We could, therefore, compute the ratio of s_a^2 to s_b^2 , and the question arises: If in fact $\sigma_a^2 = \sigma_b^2$, what is the probability of getting a value of the ratio as large as the one observed? For each pair of values of ν_a and ν_b , the tables list the values of F which are exceeded with probability α , the upper percentage point of the distribution of F . If the computed value of F exceeds this tabulated value of F_{α, ν_a, ν_b} , then we conclude that the evidence is against the hypothesis $\sigma_a^2 = \sigma_b^2$; if it is less, we conclude that σ_a^2 could be equal to σ_b^2 .

For example, we could compute the ratio of s_y^2 to s_x^2 in the preceding two examples.

Here the degrees of freedom $\nu_y = \nu_x = 10$, the tabulated value of F which is exceeded 5 percent of the time for these degrees of freedom is 2.98, and

$$\frac{s_y^2}{s_x^2} = \frac{0.000023813}{0.000011669} = 2.041$$

Since 2.04 is less than 2.98, we conclude that there is no reason to believe that the precision of the calibration process in September and October is poorer than that of May.

For small degrees of freedom, the critical value of F is rather large, e.g., for $\nu_a = \nu_b = 3$, and $\alpha' = 0.05$, the value of F is 9.28. It follows that a small difference between σ_a^2 and σ_b^2 is not likely to be detected with a small number of measurements from each process. The table below gives the approximate number of measurements required to have a four-out-of-five chance of detecting whether σ_a is the indicated multiple of σ_b (while maintaining at 0.05 the probability of incorrectly concluding that $\sigma_a > \sigma_b$, when in fact $\sigma_a = \sigma_b$).

<i>Multiple</i>	<i>No. of measurements</i>
1.5	39
2.0	15
2.5	9
3.0	7
3.5	6
4.0	5

Table A-11 in reference 4 gives the critical values of the ratios of ranges, and Tables A-20 and A-21 give confidence limits on the standard deviation of the process based on computed standard deviation.

Control Charts Technique for Maintaining Stability and Precision

A laboratory which performs routine measurement or calibration operations yields, as its daily product, numbers—averages, standard deviations, and ranges. The control chart techniques therefore could be applied to these numbers as products of a manufacturing process to furnish graphical evidence on whether the measurement process is in statistical control or out of statistical control. If it is out of control, these charts usually also indicate where and when the trouble occurred.

Control Chart for Averages. The basic concept of a control chart is in accord with what has been discussed thus far. A measurement process with limiting mean m and standard deviation σ is assumed. The sequence of numbers produced is divided into “rational” subgroups, e.g., by day, by a set of calibrations, etc. The averages of these subgroups are computed. These averages will have a mean m and a standard deviation σ/\sqrt{n} where n is the number of measurements within each subgroup. These averages are approximately normally distributed.

In the construction of the control chart for averages, m is plotted as the center line, $m + k(\sigma/\sqrt{n})$ and $m - k(\sigma/\sqrt{n})$ are plotted as control limits, and the averages are plotted in an orderly sequence. If k is taken to be 3, we know that the chance of a plotted point falling outside of the limits, if the process is in control, is very small. Therefore, if a plotted point falls outside these limits, a warning is sounded and investigative action to locate the “assignable” cause that produced the departure, or corrective measures, are called for.

The above reasoning would be applicable to actual cases only if we have chosen the proper standard deviation σ . If the standard deviation is estimated by pooling the estimates computed from each subgroup and denoted by σ_w (within group), obviously differences, if any, between group averages have

not been taken into consideration. Where there are between-group differences the variance of the individual \bar{x} is not σ_w^2/n , but, as we have seen before, $\sigma_b^2 + (\sigma_w^2/n)$, where σ_b^2 represents the variance due to differences between groups. If σ_b^2 is of any consequence as compared to σ_w^2 , many of the \bar{x} values would exceed the limits constructed by using σ_w alone.

Two alternatives are open to us: (1) remove the cause of the between-group variation; or, (2) if such variation is a proper component of error, take it into account as has been previously discussed.

As an illustration of the use of a control chart on averages, we use again the NB'10 gram data. One hundred observed corrections for NB'10 are plotted in Fig. 2-5, including the two sets of data given under comparison of means (points 18 through 28, and points 60 through 71). A three-sigma limit of $8.6 \mu\text{g}$ was used based on the "accepted" value of standard deviation.

We note that all the averages are within the control limits, excepting numbers 36, 47, 63, 85, and 87. Five in a hundred falling outside of the three-sigma limits is more than predicted by the theory. No particular reasons, however, could be found for these departures.

Since the accepted value of the standard deviation was obtained by pooling a large number of computed standard deviations for within-sets of calibrations, the graph indicates that a "between-set" component may be present. A slight shift upwards is also noted between the first 30 points and the remainder.

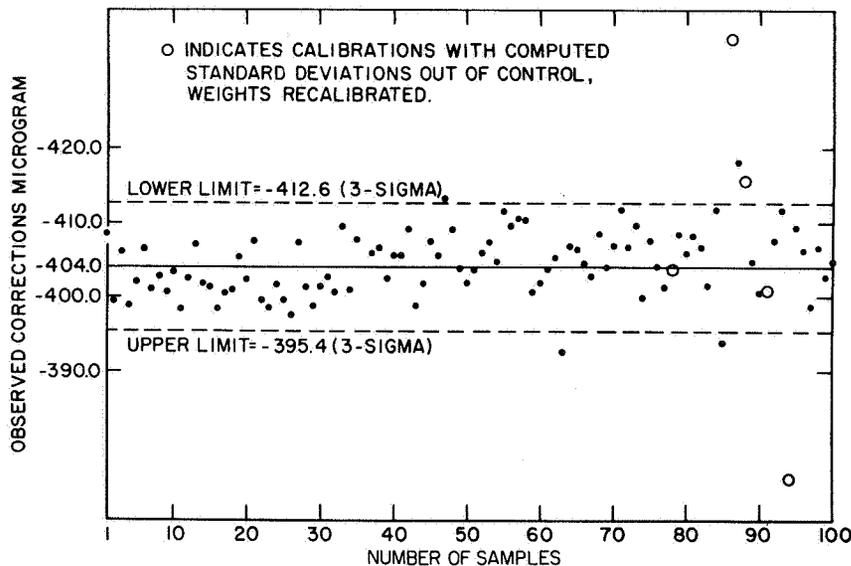


Fig. 2-5. Control chart on \bar{x} for NB'10 gram.

Control Chart for Standard Deviations. The computed standard deviation, as previously stated, is a measure of imprecision. For a set of calibrations, however, the number of measurements is usually small, and consequently also the degrees of freedom. These computed standard deviations with few degrees of freedom can vary considerably by chance alone, even though the precision of the process remains unchanged. The control chart on the computed standard deviations (or ranges) is therefore an indispensable tool.

The distribution of s depends on the degrees of freedom associated with it, and is not symmetrical about m_s . The frequency curve of s is limited on the left side by zero, and has a long "tail" to the right. The limits, therefore,

are not symmetrical about m_s . Furthermore, if the standard deviation of the process is known to be σ , m_s is not equal to σ , but is equal to $c_2\sigma$, where c_2 is a constant associated with the degrees of freedom in s .

The constants necessary for the construction of three-sigma control limits for averages, computed standard deviations, and ranges, are given in most textbooks on quality control. Section 18-3 of reference 4 gives such a table. A more comprehensive treatment on control charts is given in ASTM "Manual on Quality Control of Materials," Special Technical Publication 15-C.

Unfortunately, the notation employed in quality control work differs in some respect from what is now standard in statistics, and correction factors have to be applied to some of these constants when the computed standard deviation is calculated by the definition given in this chapter. These corrections are explained in the footnote under the table.

As an example of the use of control charts on the precision of a calibration process, we will use data from NBS calibration of standard cells.* Standard cells in groups of four or six are usually compared with an NBS standard cell on ten separate days. A typical data sheet for a group of six cells, after all the necessary corrections, appears in Table 2-6. The standard deviation of a comparison is calculated from the ten comparisons for each cell and the standard deviation for the average value of the ten comparisons is listed in the line marked SDA. These values were plotted as points 6 through 11 in Fig. 2-6.

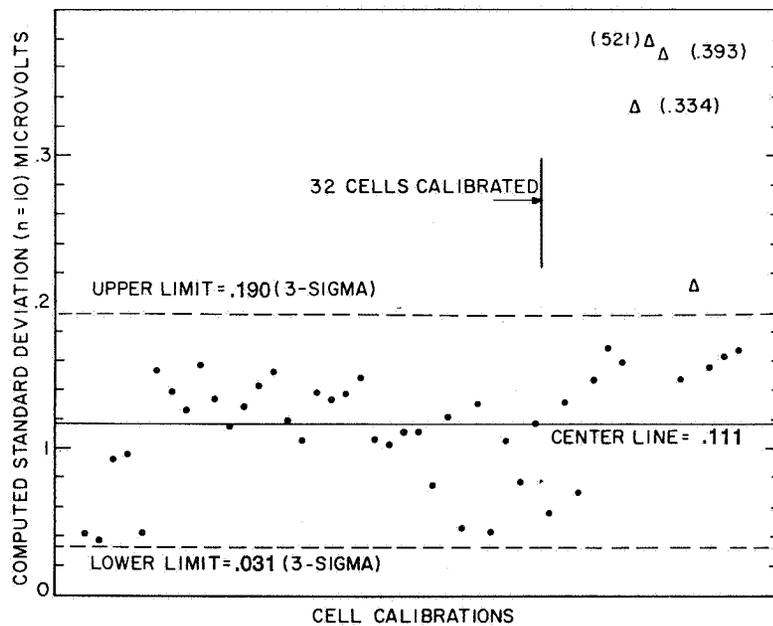


Fig. 2-6. Control chart on s for the calibration of standard cells.

Let us assume that the precision of the calibration process remains the same. We can therefore pool the standard deviations computed for each cell (with nine degrees of freedom) over a number of cells and take this value as the current value of the standard deviation of a comparison, σ . The corresponding current value of standard deviation of the average of ten comparisons will be denoted by $\sigma' = \sigma/\sqrt{10}$. The control chart will be made on $s' = s/\sqrt{10}$.

*Illustrative data supplied by Miss Catherine Law, Electricity Division, National Bureau of Standards.

Environmental Assessment for Houston Optimization of Airspace and Procedures in the Metroplex

For example, the SDA's for 32 cells calibrated between June 29 and August 8, 1962, are plotted as the first 32 points in Fig. 2-6. The pooled standard deviation of the average is 0.114 with 288 degrees of freedom. The between-group component is assumed to be negligible.

Table 2-6. Calibration data for six standard cells

Day	Corrected Emf's and standard deviations, Microvolts					
1	27.10	24.30	31.30	33.30	32.30	23.20
2	25.96	24.06	31.06	34.16	33.26	23.76
3	26.02	24.22	31.92	33.82	33.22	24.02
4	26.26	24.96	31.26	33.96	33.26	24.16
5	27.23	25.23	31.53	34.73	33.33	24.43
6	25.90	24.40	31.80	33.90	32.90	24.10
7	26.79	24.99	32.19	34.39	33.39	24.39
8	26.18	24.98	32.18	35.08	33.98	24.38
9	26.17	25.07	31.97	34.27	33.07	23.97
10	26.16	25.16	31.96	34.06	32.96	24.16
R	1.331	1.169	1.127	1.777	1.677	1.233
AVG	26.378	24.738	31.718	34.168	33.168	24.058
SD	0.482	0.439	0.402	0.495	0.425	0.366
SDA	0.153	0.139	0.127	0.157	0.134	0.116

Position	Emf, volts	Position	Emf, volts
1	1.0182264	4	1.0182342
2	1.0182247	5	1.0182332
3	1.0182317	6	1.0182240

Since $n = 10$, we find our constants for three-sigma control limits on s' in Section 18-3 of reference 4 and apply the corrections as follows:

$$\text{Center line} = \sqrt{\frac{n}{n-1}} c_2 \sigma' = 1.054 \times 0.9227 \times 0.114 = 0.111$$

$$\text{Lower limit} = \sqrt{\frac{n}{n-1}} B_1 \sigma' = 1.054 \times 0.262 \times 0.114 = 0.031$$

$$\text{Upper limit} = \sqrt{\frac{n}{n-1}} B_2 \sigma' = 1.054 \times 1.584 \times 0.114 = 0.190$$

The control chart (Fig. 2-6) was constructed using these values of center line and control limits computed from the 32 calibrations. The standard deviations of the averages of subsequent calibrations are then plotted.

Three points in Fig. 2-6 far exceed the upper control limit. All three cells, which were from the same source, showed drifts during the period of calibration. A fourth point barely exceeded the limit. It is to be noted that the data here were selected to include these three points for purposes of illustration only, and do not represent the normal sequence of calibrations.

The main function of the chart is to justify the precision statement on the report of calibration, which is based on a value of σ estimated with perhaps thousands of degrees of freedom and which is shown to be in control. The report of calibration for these cells ($\sigma = 0.117 \pm 0.12$) could read:

“Each value is the mean of ten observations made between ____ and _____. Based on a standard deviation of 0.12 microvolts for the means, these values are correct to 0.36 microvolts relative to the volt as maintained by the national reference group.”

Environmental Assessment for Houston Optimization of Airspace and Procedures in the Metroplex
Linear Relationship and Fitting of Constants by Least Squares

In using the arithmetic mean of n measurements as an estimate of the limiting mean, we have, knowingly or unknowingly, fitted a constant to the data by the method of least squares, i.e., we have selected a value \hat{m} for m such that

$$\sum_1^n (y_i - \hat{m})^2 = \sum_1^n d_i^2$$

is a minimum. The solution is $\hat{m} = \bar{y}$. The deviations $d_i = y_i - \hat{m} = y_i - \bar{y}$ are called residuals.

Here we can express our measurements in the form of a mathematical model

$$Y = m + \epsilon \quad (2-19)$$

where Y stands for the observed values, m the limiting mean (a constant), and ϵ the random error (normal) of measurement with a limiting mean zero and a standard deviation σ . By (2-1) and (2-9), it follows that

$$m_y = m + m_\epsilon = m$$

and

$$\sigma_y^2 = \sigma^2$$

The method of least squares requires us to use that estimator \hat{m} for m such that the sum of squares of the residuals is a minimum (among all possible estimators). As a corollary, the method also states that the sum of squares of residuals divided by the number of measurements n less the number of estimated constants p will give us an estimate of σ^2 , i.e.,

$$s^2 = \frac{\sum (y_i - \hat{m})^2}{n - p} = \frac{\sum (y_i - \bar{y})^2}{n - 1} \quad (2-20)$$

It is seen that the above agrees with our definition of s^2 .

Suppose Y , the quantity measured, exhibits a linear functional relationship with a variable which can be controlled accurately; then a model can be written as

$$Y = a + bX + \epsilon \quad (2-21)$$

where, as before, Y is the quantity measured, a (the intercept) and b (the slope) are two constants to be estimated, and ϵ the random error with limiting mean zero and variance σ^2 . We set X at x_i , and observe y_i . For example, y_i might be the change in length of a gage block steel observed for n equally spaced temperatures x_i within a certain range. The quantity of interest is the coefficient of thermal expansion b .

For any estimates of a and b , say \hat{a} and \hat{b} , we can compute a value \hat{y}_i for each x_i , or

$$\hat{y}_i = \hat{a} + \hat{b}x_i$$

If we require the sum of squares of the residuals

$$\sum_{i=1}^n (y_i - \hat{y}_i)^2$$

to be a minimum, then it can be shown that

$$\hat{b} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (2-22)$$

and

$$\hat{a} = \bar{y} - \hat{b}\bar{x} \quad (2-23)$$

The variance of Y can be estimated by

$$s^2 = \frac{\sum (y_i - \hat{y}_i)^2}{n - 2} \quad (2-24)$$

with $n - 2$ degrees of freedom since two constants have been estimated from the data.

The standard errors of \hat{b} and \hat{a} are respectively estimated by s_b and $s_{\hat{a}}$, where

$$s_b^2 = \frac{s^2}{\sum (x_i - \bar{x})^2} \quad (2-25)$$

$$s_{\hat{a}}^2 = s^2 \left[\frac{1}{n} + \frac{\bar{x}^2}{\sum (x_i - \bar{x})^2} \right] \quad (2-26)$$

With these estimates and the degrees of freedom associated with s^2 , confidence limits can be computed for \hat{a} and \hat{b} for the confidence coefficient selected if we assume that errors are normally distributed.

Thus, the lower and upper limits of a and b , respectively, are:

$$\hat{a} - ts_{\hat{a}}, \quad \hat{a} + ts_{\hat{a}}$$

$$\hat{b} - ts_b, \quad \hat{b} + ts_b$$

for the value of t corresponding to the degree of freedom and the selected confidence coefficient.

The following problems relating to a linear relationship between two variables are treated in reference 4, Section 5-4.

1. Confidence intervals for a point on the fitted line.
2. Confidence band for the line as a whole.
3. Confidence interval for a single predicted value of Y for a given X .

Polynomial and multivariate relationships are treated in Chapter 6 of the same reference.

REFERENCES

The following references are recommended for topics introduced in the first section of this chapter:

1. Wilson, Jr., E. B., *An Introduction to Scientific Research*, McGraw-Hill Book Company, New York, 1952, Chapters 7, 8, and 9.
2. Eisenhart, Churchill, "Realistic Evaluation of the Precision and Accuracy of Instrument Calibration System," *Journal of Research of the National Bureau of Standards*, Vol. 67C, No. 2, 1963.
3. Youden, W. J., *Experimentation and Measurement*, National Science Teacher Association Vista of Science Series No. 2, Scholastic Book Series, New York.

In addition to the three general references given above, the following are selected with special emphasis on their ease of understanding and applicability in the measurement science:

Statistical Methods

4. Natrella, M. G., *Experimental Statistics*, NBS Handbook 91, U.S. Government Printing Office, Washington, D.C., 1963.
5. Youden, W. J., *Statistical Methods for Chemists*, John Wiley & Sons, Inc., New York, 1951.
6. Davies, O. L., *Statistical Method in Research and Production* (3rd ed.), Hafner Publishing Co., Inc., New York, 1957.

Textbooks

7. Dixon, W. J. and F. J. Massey, *Introduction to Statistical Analysis* (2nd ed.), McGraw-Hill Book Company, New York, 1957.
8. Brownlee, K. A., *Statistical Theory and Methodology in Science and Engineering*, John Wiley & Sons, Inc., New York, 1960.
9. Areley, N. and K. R. Buch, *Introduction to the Theory of Probability and Statistics*, John Wiley & Sons, Inc., New York, 1950.

Additional Books on Treatment of Data

10. American Society for Testing and Materials, *Manual on Presentation of Data and Control Chart Analysis* (STP 15D), 1976, 162 p.
11. American Society for Testing and Materials, *ASTM Standard on Precision and Accuracy for Various Applications*, 1977, 249 p.
12. Box, G. E. P., Hunter, W. G., and Hunter, J. S., *Statistics for Experimenters, an Introduction to Design, Data Analysis, and Model Building*, John Wiley and Sons, Inc., New York, 1978, 575 p.
13. Cox, D. R., *Planning of Experiments*, John Wiley and Sons, Inc., New York, 1958, 308 p.
14. Daniel, C. and Wood, F. S., *Fitting Equations to Data, Computer Analysis of Multifactor Data*, 2d ed., John Wiley and Sons, Inc., New York, 1979, 343 p.
15. Deming, W. E., *Statistical Adjustment of Data*, John Wiley and Sons, Inc., New York, 1943, 261 p.
16. Draper, N. R. and Smith, H., *Applied Regression Analysis*, John Wiley and Sons, Inc., New York, 1966, 407 p.
17. Himmelblau, D. M., *Process Analysis by Statistical Methods*, John Wiley and Sons, Inc., New York, 1970, 464 p.
18. Ku, H. H., ed., *Precision Measurement and Calibration: Statistical Concepts and Procedures*, Natl. Bur. Stand. (U.S.) Spec. Publ. 300, 1969, v.p.
19. Mandel, J., *The Statistical Analysis of Experimental Data*, Interscience, New York, 1964, 410 p.
20. Mosteller, F. and Tukey, J. W., *Data Analysis and Regression, a Second Course in Statistics*, Addison-Wesley, Reading, Massachusetts, 1977, 588 p.

Environmental Assessment for Houston Optimization of Airspace and Procedures in the Metroplex
POSTSCRIPT
ON
STATISTICAL GRAPHICS

Over the years since the publication of the above article, it has become apparent that some additions on recent developments for the treatment of data may be useful. It is equally apparent that the concepts and techniques introduced in the original article remain as valid and appropriate as when first written. For this reason, a few additional sections on statistical graphics are added as a postscript.

The power of small computers and the associated sophisticated software have pushed graphics into the forefront. Plots and graphs have always been popular with engineers and scientists, but their use has been limited by the time and work involved. Graphics packages now-a-days allow the user to do plots and graphs with ease, and a good statistical package will also automatically present a number of pertinent plots for examination. As John Tukey said, "the greatest value of a picture is when it forces us to notice what we never expected to see." [1] An outlier? Skewed distribution of values? Poor modelling? What is the data trying to say? Answers to all these come naturally through inspection of plots and graphs, whereas columns of numbers reveal little, if anything.

Control charts for the mean (Fig. 2-5) and standard deviation (Fig. 2-6) are classical examples of graphical methods. Control charts were introduced by Walter Shewhart some 60 years ago, yet the technique remains a popular and most useful tool in business and industry. Simplicity (once constructed), self-explanatory nature, and robustness (not depending on assumptions) are, and should be, the main desirable attributes of all graphs and plots.

Since statistical graphics is a huge subject, only a few basic techniques that are particularly useful to the treatment of measurement data will be discussed, together with references for further reading.

Plots for Summary and Display of Data

Stem and Leaf. The stem and leaf plot is a close relative of the histogram, but it uses digits of data values themselves to show features of the data set instead of areas of rectangles. First proposed by John W. Tukey, a stem and leaf plot retains more information from the data than the histogram and is particularly suited for the display of small to moderate-sized data sets.

Fig. 1 is a stem and leaf plot of 48 measurements of the isotopic ratio of ⁷⁹Bromine to ⁸¹Bromine. Values of these 48 data points, listed in Table 1, range from 1.0261 to 1.0305, or 261 to 305 after coding. The leaves are the last digits of the data values, 0 to 9. The stems are 26, 27, 28, 29, and 30. Thus 261 is split into two parts, plotted as 26 | 1. In this case, because of the heavy concentration of values in stems 28 and 29, two lines are given to each stem, with leaves 0 to 4 on the first line, and 5 to 9 on the second. Stems are shown on the left side of the vertical line and individual leaves on the right side. There is no need for a separate table of data values - they are all shown in the plot!

The plot shows a slight skew towards lower values. The smallest value separates from the next by 0.7 units. Is that an outlier? These data will be examined again later.

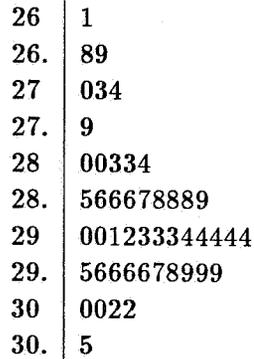


Fig. 1. Stem and leaf plot. 48 values of isotopic ratios, bromine (79/81). Unit = $(Y - 1.0) \times 10^4$, thus 26|1 = 1.0261.

Table 1. Y-Ratios 79/81 for reference sample

	DETERMINATION I		DETERMINATION II	
	Instrument #4	Instrument #1	Instrument #4	Instrument #1
	1.0292	1.0289	1.0296	1.0284
	1.0294	1.0285	1.0293	1.0270
	1.0298	1.0287	1.0302	1.0279
	1.0302	1.0297	1.0305	1.0269
	1.0294	1.0290	1.0288	1.0273
	1.0296	1.0286	1.0294	1.0261
	1.0293	1.0291	1.0299	1.0286
	1.0295	1.0293	1.0290	1.0286
	1.0300	1.0288	1.0296	1.0293
	1.0297	1.0298	1.0299	1.0283
	1.0296	1.0274	1.0299	1.0263
	1.0294	1.0280	1.0300	1.0280
Ave.	1.029502	1.028792	1.029675	1.027683
S ²	.00000086	.00000041	.00000024	.00000069
S	.00029	.00064	.00049	.00083
S _y	.00008	.00018	.00014	.00024

Box Plot. Customarily, a batch of data is summarized by its average and standard deviation. These two numerical values characterize a normal distribution, as explained in expression (2-0). Certain features of the data, e.g., skewness and extreme values, are not reflected in the average and standard deviation. The box plot (due also to Tukey) presents graphically a five-number summary which, in many cases, shows more of the original features of the batch of data than the two number summary.

To construct a box plot, the sample of numbers are first ordered from the smallest to the largest, resulting in

$$x_{(1)}, x_{(2)}, \dots, x_{(n)}.$$

Using a set of rules, the median, m , the lower fourth, F_l , and the upper fourth, F_u , are calculated. By definition, the interval $(F_u - F_l)$ contains half of all data points. We note that m , F_u , and F_l are not disturbed by outliers.

The interval $(F_u - F_l)$ is called the fourth spread. The lower cutoff limit is

$$F_l - 1.5(F_u - F_l)$$

and the upper cutoff limit is

$$F_u + 1.5(F_u - F_l).$$

A "box" is then constructed between F_l and F_u , with the median line dividing the box into two parts. Two tails from the ends of the box extend to $x_{(1)}$ and $x_{(n)}$ respectively. If the tails exceed the cutoff limits, the cutoff limits are also marked.

From a box plot one can see certain prominent features of a batch of data:

1. Location - the median, and whether it is in the middle of the box.
2. Spread - The fourth spread (50 percent of data): - lower and upper cut off limits (99.3 percent of the data will be in the interval if the distribution is normal and the data set is large).
3. Symmetry/skewness - equal or different tail lengths.
4. Outlying data points - suspected outliers.

The 48 measurements of isotopic ratio bromine ($79/81$) shown in Fig. 1 were actually made on two instruments, with 24 measurements each. Box plots for instrument I, instrument II, and for both instruments are shown in Fig. 2.

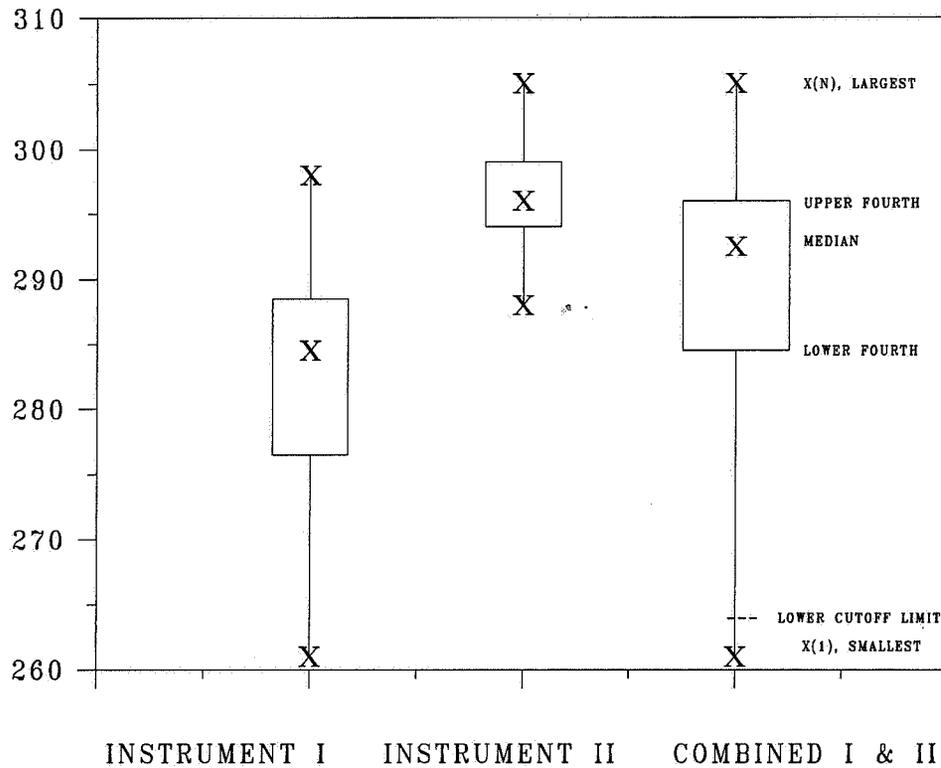


Fig. 2. Box plot of isotopic ratio, bromine (79/91).

The five number summary for the 48 data point is, for the combined data:

Smallest:	$X(1) = 261$
Median X_m :	$m = (n + 1)/2 = (48 + 1)/2 = 24.5$ $X_m = x_{(m)}$ if m is an integer; $= [x_{(M)} + x_{(M+1)}]/2$ if not; where M is the largest integer not exceeding m . $X_m = (291 + 292)/2 = 291.5$
Lower Fourth X_l :	$l = (M + 1)/2 = (24 + 1)/2 = 12.5$ $X_l = x_{(l)}$ if l is an integer; $= [x_{(L)} + x_{(L+1)}]/2$ if not, where L is the largest integer not exceeding l . $X_l = (284 + 285)/2 = 284.5$
Upper Fourth X_u :	$u = n + 1 - l = 49 - 12.5 = 36.5$ $X_u = x_{(u)}$ if u is an integer; $= [x_{(U)} + x_{(U+1)}]/2$ if not, where U is the largest integer not exceeding u . $X_u = (296 + 296)/2 = 296$
Largest:	$X_{(n)} = 305$

Box plots for instruments I and II are similarly constructed. It seems apparent from these two plots that (a) there was a difference between the results for these two instruments, and (b) the precision of instrument II is better than that of instrument I. The lowest value of instrument I, 261, is less than the lower cutoff for the plot of the combined data, but it does not fall below the lower cutoff for instrument I alone. As an exercise, think of why this is the case.

Box plots can be used to compare several batches of data effectively and easily. Fig. 3 is a box plot of the amount of magnesium in different parts of a long alloy rod. The specimen number represents the distance, in meters, from the edge of the 100 meter rod to the place where the specimen was taken. Ten determinations were made at the selected locations for each specimen. One outlier appears obvious; there is also a mild indication of decreasing content of magnesium along the rod.

Variations of box plots are given in [3] and [4].

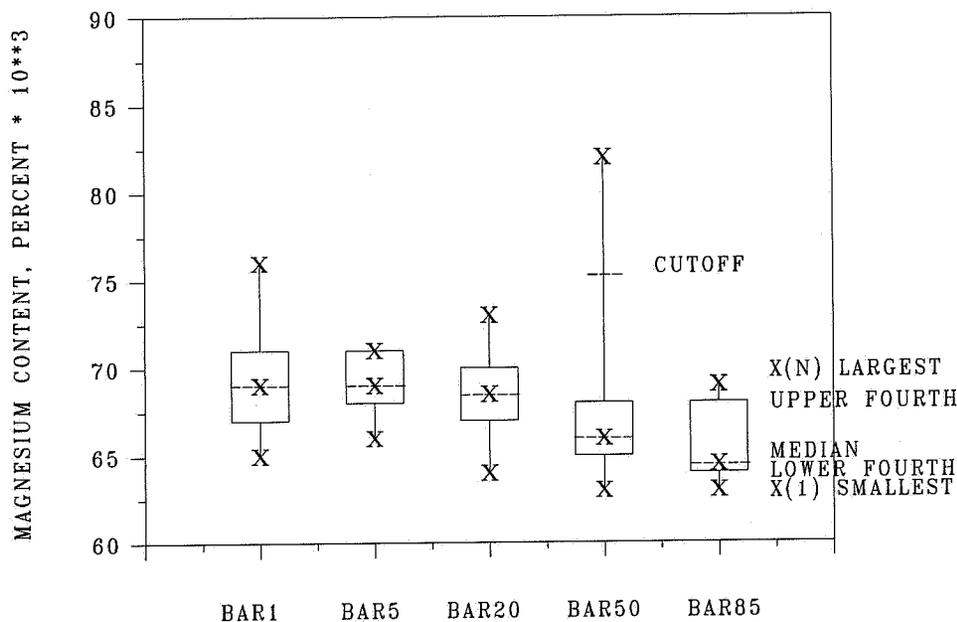


Fig. 3. Magnesium content of specimens taken.

Plots for Checking on Models and Assumptions

In making measurements, we may consider that each measurement is made up of two parts, one fixed and one variable, i.e.,

$$\text{Measurement} = \text{fixed part} + \text{variable part},$$

or, in other words,

$$\text{Data} = \text{model} + \text{error}.$$

We use measured data to estimate the fixed part, (the Mean, for example), and use the variable part (perhaps summarized by the standard deviation) to assess the goodness of our estimate.

Residuals. Let the i th data point be denoted by y_i , let the fixed part be a constant M , and let the random error be ϵ_i as used in equation (2-19). Then

$$y_i = M + \epsilon_i, \quad i = 1, 2, \dots, n.$$

If we use the method of least squares to estimate m , the resulting estimate is

$$m = \bar{y} = \sum_i y_i / n,$$

or the average of all measurements.

The i th residual, r_i , is defined as the difference between the i th data point and the fitted constant, i.e.

$$r_i = y_i - \bar{y}.$$

In general, the fixed part can be a function of another variable X (or more than one variable). Then the model is

$$y_i = F(x_i) + \epsilon_i$$

and the i th residual is defined as

$$r_i = y_i - F(x_i),$$

where $F(x_i)$ is the value of the function computed with the fitted parameters. If the relationship between Y and X is linear as in (2-21), then $r_i = y_i - (a + bx_i)$ where a and b are the intercept and the slope of the fitted straight line, respectively.

When, as in calibration work, the values of $F(x_i)$ are frequently considered to be known, the differences between measured values and known values will be denoted d_i , the i th deviation, and can be used for plots instead of residuals.

Adequacy of Model. Following is a discussion of some of the issues involved in checking the adequacy of models and assumptions. For each issue, pertinent graphical techniques involving residuals or deviations are presented.

In calibrating a load cell, known deadweights are added in sequence and the deflections are read after each additional load. The deflections are plotted against loads in Fig. 4. A straight line model looks plausible, i.e.,

$$(\text{deflection } i) = b_0 + b_1(\text{load}_i).$$

A line is fitted by the method of least squares and the residuals from the fit are plotted in Fig. 5. The parabolic curve suggests that this model is inadequate, and that a second degree equation might fit better:

$$(\text{deflection}_i) = b_0 + b_1(\text{load}_i) + b_2(\text{load}_i)^2.$$

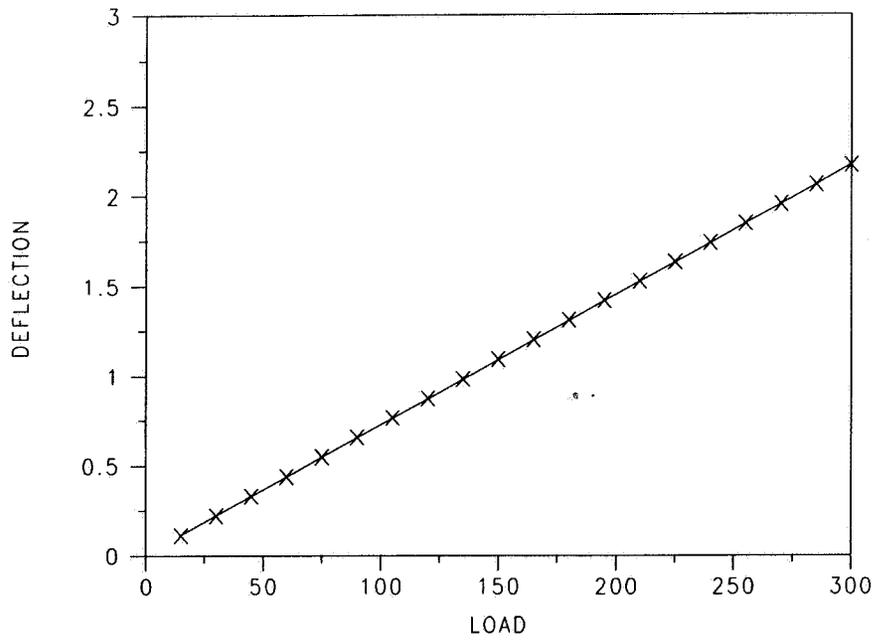


Fig. 4. Plot of deflection vs load.

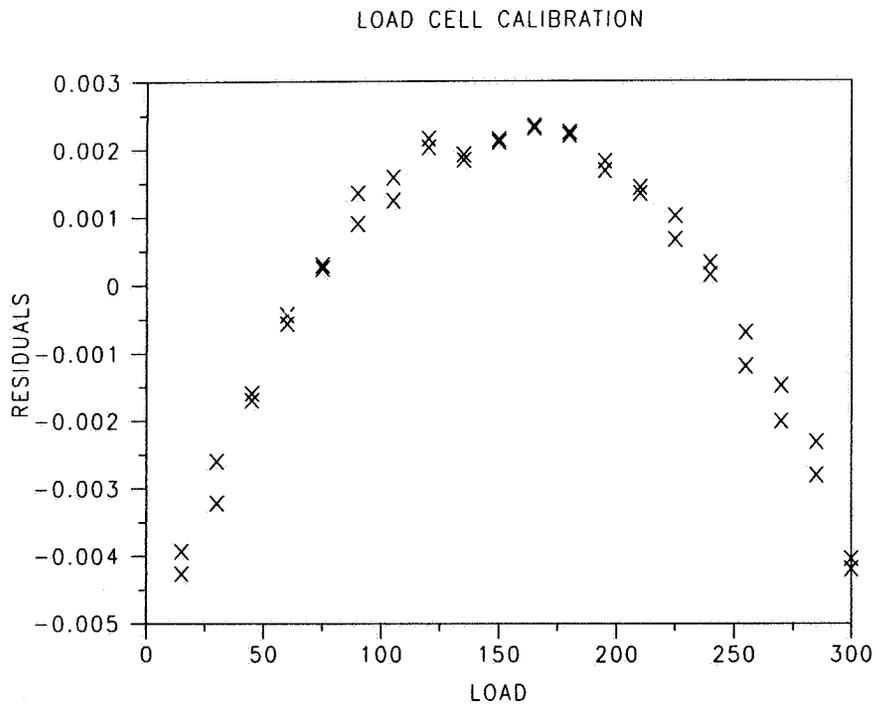


Fig. 5. Plot of residuals after linear fit.

This is done and the residuals from this second degree model are plotted against loads, resulting in Fig. 6. These residuals look random, yet a pattern may still be discerned upon close inspection. These patterns can be investigated to see if they are peculiar to this individual load cell, or are common to all load cells of similar design, or to all load cells.

Uncertainties based on residuals resulting from an inadequate model could be incorrect and misleading.

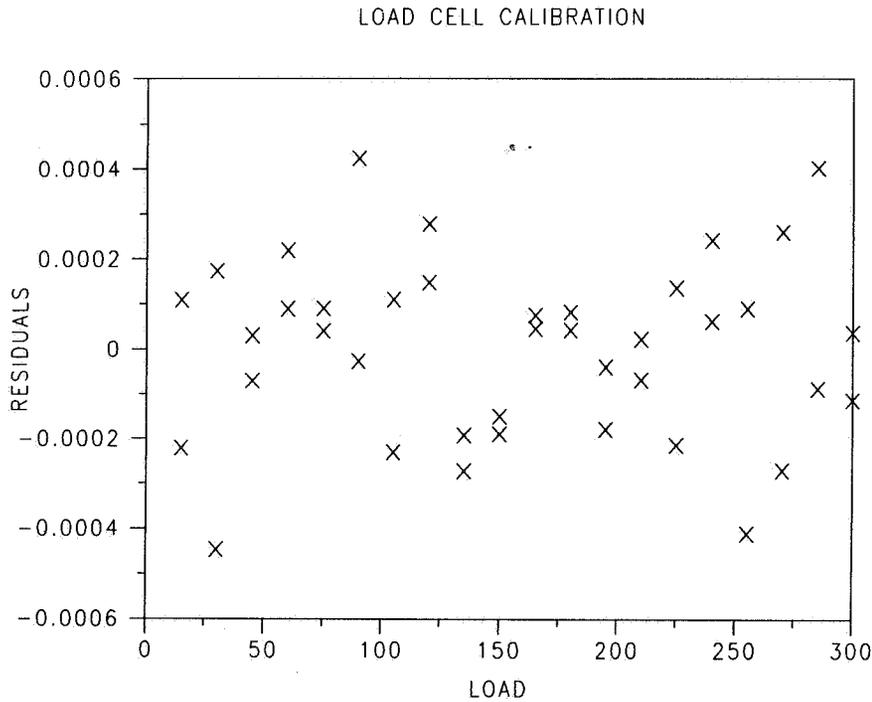


Fig. 6. Plot of residuals after quadratic fit.

Testing of Underlying Assumptions. In equation (2-19),

$$Y = m + \epsilon,$$

the assumptions are made that ϵ represents the random error (normal) and has a limiting mean zero and a standard deviation σ . In many measurement situations, these assumptions are approximately true. Departures from these assumptions, however, would invalidate our model and our assessment of uncertainties. Residual plots help in detecting any unacceptable departures from these assumptions.

Residuals from a straight line fit of measured depths of weld defects (radiographic method) to known depths (actually measured) are plotted against the known depths in Fig. 7. The increase in variability with depths of defects is apparent from the figure. Hence the assumption of constant σ over the range of $F(x)$ is violated. If the variability of residuals is proportional to depth, fitting of $\ln(y_i)$ against known depths is suggested by this plot.

The assumption that errors are normally distributed may be checked by doing a normal probability plot of the residuals. If the distribution is approximately normal, the plot should show a linear relationship. Curvature in the plot provides evidence that the distribution of errors is other than

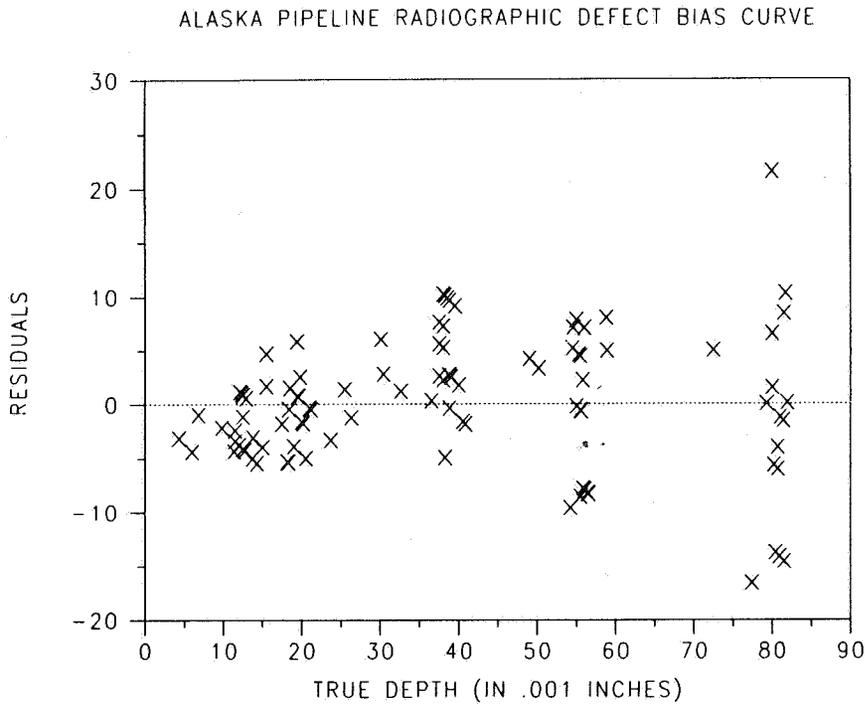


Fig. 7. Plot of residuals after linear fit. Measured depth of weld defects vs true depth.

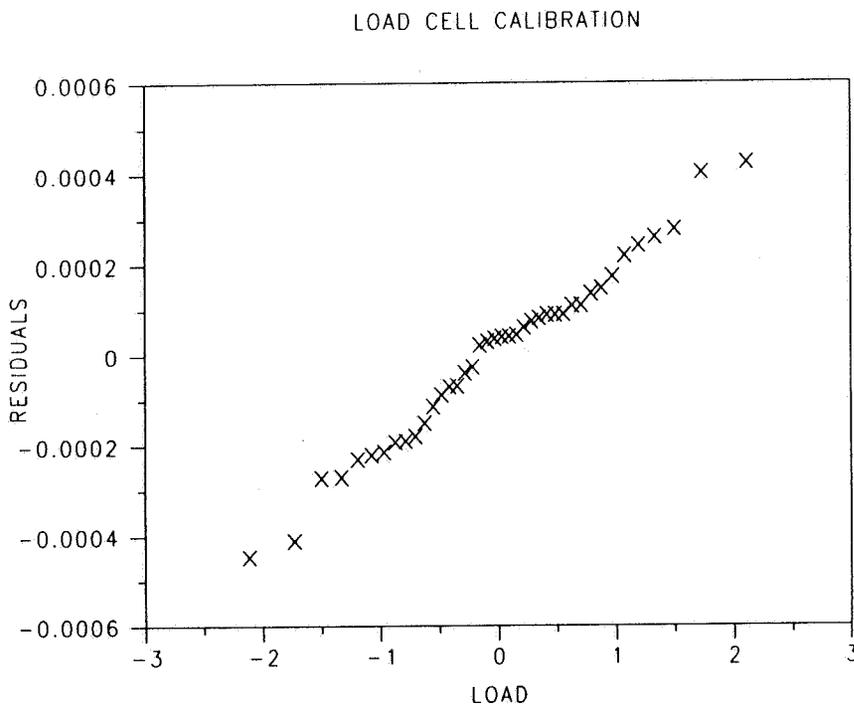


Fig. 8. Normal probability plot of residuals after quadratic fit.

normal. Fig. 8 is a normal probability plot of the residuals in Fig. 6, showing some evidence of departure from normality. Note the change in slope in the middle range.

Inspection of normal probability plots is not an easy job, however, unless the curvature is substantial. Frequently symmetry of the distribution of

errors is of main concern. Then a stem and leaf plot of data or residuals serves the purpose just as well as, if not better than, a normal probability plot. See, for example, Fig. 1.

Stability of a Measurement Sequence. It is a practice of most experimenters to plot the results of each run in sequence to check whether the measurements are stable over runs. The run- sequence plot differs from control charts in that no formal rules are used for action. The stability of a measurement process depends on many factors that are recorded but are not considered in the model because their effects are thought to be negligible.

Plots of residuals versus days, sets, instruments, operators, temperatures, humidities, etc., may be used to check whether effects of these factors are indeed negligible. Shifts in levels between days or instruments (see Fig. 2), trends over time, and dependence on environmental conditions are easily seen from a plot of residuals versus such factors.

In calibration work, frequently the values of standards are considered to be known. The differences between measured values and known values may be used for a plot instead of residuals.

Figs. 9, 10, and 11 are multi-trace plots of results from three laboratories of measuring linewidth standards using different optical imaging methods. The difference of 10 measured line widths from NBS values are plotted against NBS values for 7 days. It is apparent that measurements made on day 5 were out of control in Fig. 9. Fig. 10 shows a downward trend of differences with increasing line widths; Fig. 11 shows three significant outliers. These plots could be of help to those laboratories in locating and correcting causes of these anomalies. Fig. 12 plots the results of calibration of standard watt- hour meters from 1978 to 1982. It is evident that the variability of results at one time, represented by σ_w (discussed under Component of Variance Between Groups, p. 19), does not reflect the variability over a period of time, represented by σ_b (discussed in the same section). Hence, three measurements every three months would yield better variability information than, say, twelve measurements a year apart.

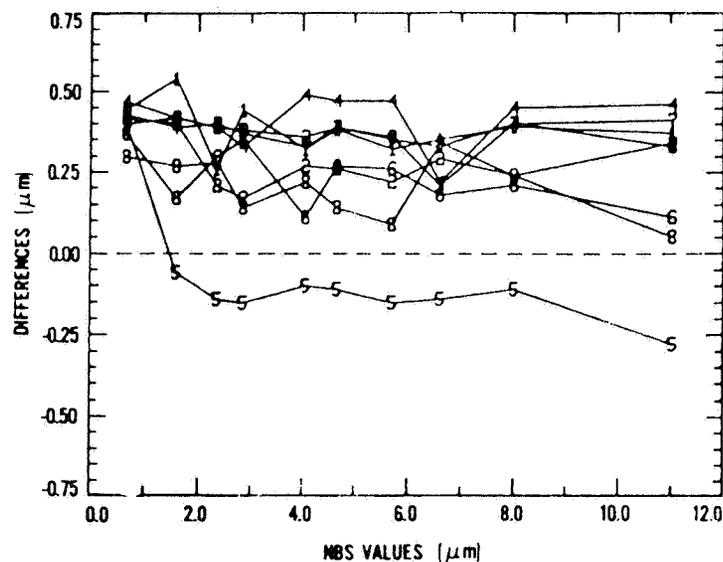


Fig. 9. Differences of linewidth measurements from NBS values. Measurements on day 5 inconsistent with others—Lab A.

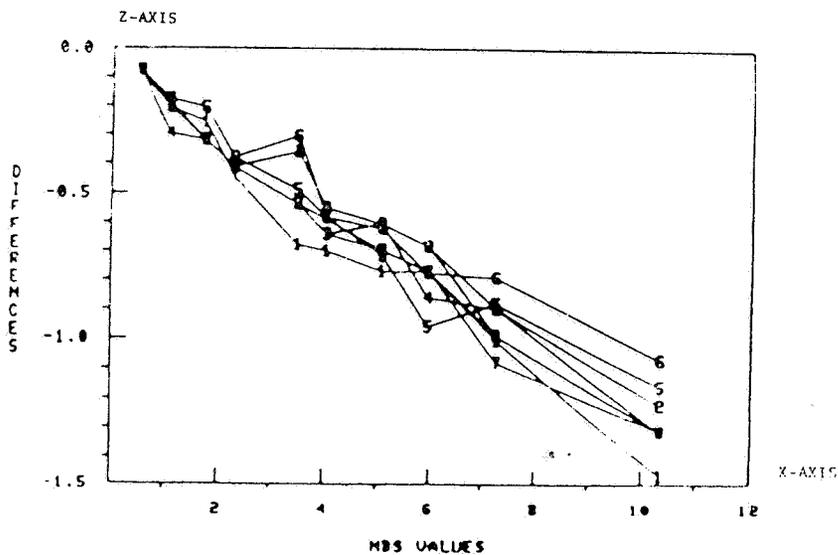


Fig. 10. Trend with increasing linewidths—Lab B.

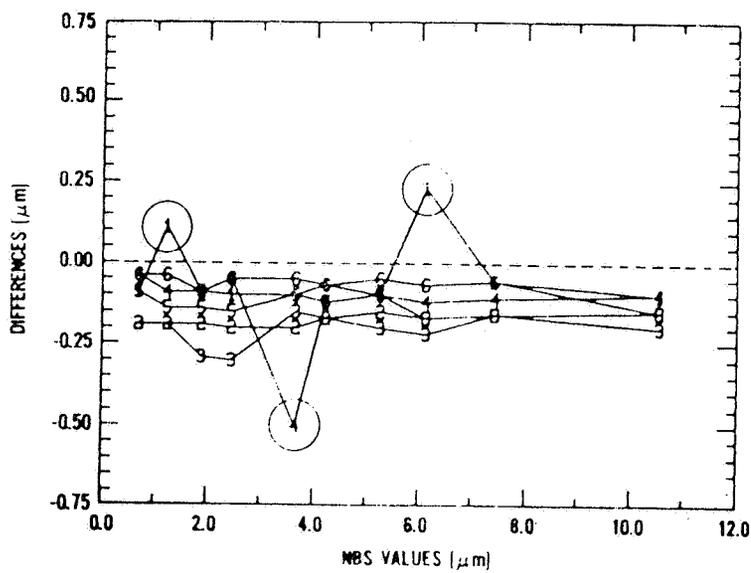


Fig. 11. Significant isolated outliers—Lab C.

- [4] Hoaglin, David C., Mosteller, Frederick, and Tukey, John W., *Understanding Robust and Exploratory Data Analysis*, John Wiley & Sons, 1983.
- [5] Velleman, Paul F., and Hoaglin, David C., *Applications, Basics, and Computing of Exploratory Data Analysis*, Duxbury Press, 1981.
- [6] Filliben, James J., 'DATAPLOT - An Interactive High-level Language for Graphics, Nonlinear Fitting, Data Analysis and Mathematics,' *Computer Graphics, Vol. 15, No. 3*, August, 1981.
- [7] Cleveland, William S., et al., 'Research in Statistical Graphics,' *Journal of the American Statistical Association, Vol. 82, No. 398*, June 1987, pp. 419-423.

David A. Crandall

From: 9-ASW-HoustonOAPMcomment@faa.gov
Sent: Friday, February 22, 2013 15:47
To: David A. Crandall
Cc: Kirk C. Harris
Subject: Fw: Keep the Jets Out of High Meadow Ranch!

Importance: High

-----Forwarded by 9-ASW-HoustonOAPMcomment/ASW/FAA on 02/22/2013 02:45PM -----

To: 9-ASW-HoustonOAPMcomment/ASW/FAA
From: Tim Honeycutt <timhoneycutt@earthlink.net>
Date: 02/19/2013 03:13PM
Subject: Keep the Jets Out of High Meadow Ranch!

Dear Sir:

I, along with many others, own property in the High Meadow Ranch neighborhood in Magnolia, Texas. I sincerely hope that the proposed flight paths for DRLLR STAR and AGGEE STAR do not track through High Meadow Ranch. The property owners in High Meadow Ranch are affluent and chose the area for its serenity. None would appreciate a steady stream of descending aircraft on an unwavering GPS-dictated ground track, even if in some sort of continuous descent profile.

--
Regards,

Tim Honeycutt
16607 Saint Johns Wood Dr.
Tomball, Texas 77377
281.379.7999 voice
832.565.1761 fax

TH1

David A. Crandall

From: Roger.McGrath@faa.gov
Sent: Friday, February 22, 2013 13:31
To: David A. Crandall
Cc: Kirk C. Harris
Subject: Fw: Houston OAPM Mailbox for comments

----- Forwarded by Roger McGrath/ASW/FAA on 02/22/2013 12:25 PM -----

From: Tim Honeycutt <timhoneycutt@earthlink.net>
To: Roger McGrath/ASW/FAA@FAA
Date: 02/19/2013 03:43 PM
Subject: Re: Houston OAPM Mailbox for comments

Dear Mr. McGrath:

TH2 Will the proposed ground track for DRLLR STAR and/or AGGEE STAR pass through High Meadow Ranch in Magnolia, Texas? Figure 3 on page 79 of appendix F of the Houston OAPM EA is devilishly hard to decipher without access to really detailed aeronautical charts. Thanks.

Regards,



Tim Honeycutt
Attorney at Law
16607 Saint Johns Wood Dr.
Tomball, Texas 77377
281.379.7999 voice
832.565.1761 fax

On 5/30/2012 10:00 AM, Roger.McGrath@faa.gov wrote:

Mr Honeycutt,

Thanks for your two previous inquiries about the FAA's Houston OAPM initiative. To bring you up to date we have a dedicated email address for those who wish to comment on the project. The address is:

9-ASW-HoustonOAPMcomment@faa.gov

This email is being sent to that mailbox as a reminder to us that you wanted more information.

We are developing a website that will give much information on the project and will be available very soon. I will notify you when the website is active.

Thanks again for your interest.

Environmental Assessment for Houston Optimization of Airspace and Procedures in the Metroplex

Roger

Roger McGrath
Airspace and Environmental Specialist
FAA Central Service Area
Operations Support Group
817-321-7735
roger.mcgrath@faa.gov

Responses to Comments
to the Houston Optimization of Airspace
and Procedures in the Metroplex,
Houston Texas (EA), January 2013

**Comments and Responses
to the Houston Optimization of Airspace and Procedures in the Metroplex, Houston Texas (EA), January 2013**

Agency/Organization	Name	Letter Date	Comment Number	Comment	FAA Response
Concurrence Letters					
Texas Historical Commission	Kelly Little (for Mark Wolfe)	2/21/2013	SHPO	The FAA has found that the undertaking will have no adverse effect on historic resources. Based on the information provided in the EA, the SHPO concurs with this determination.	Comment Noted
U.S. Fish and Wildlife Service	Edith Erfling	3/11/2013	FWS1	The United States Fish and Wildlife Service concurs with FAA's determination that the proposed changes in flight paths and altitudes are not likely to adversely affect any federally listed species under our jurisdiction. This concurrence is based upon a review of our project files, and is contingent upon implementation of the avoidance, minimization, and conservation measures proposed by FAA.	Comment Noted
Comments Received					
U.S. Dept of Interior - Bureau of Indian Affairs	David Anderson	2/11/2013	DOI1	A review of the EA and the associated maps indicate that there would be no significant impacts to tribal or Individual Indian trust lands under the jurisdiction of the Southern Plains Region if the proposed changes in aircraft flight paths and altitudes are implemented.	Comment Noted
U.S. Env. Protection Agency	Debra Griffin	2/19/2013	EPA1	Based upon the environmental assessment information and related correspondence of State and other Federal resource agencies, EPA has no objection to the implementation of the proposed project. Factors closely considered include the effects upon air quality, endangered and threatened species, and environmental justice.	Comment Noted

**Comments and Responses
to the Houston Optimization of Airspace and Procedures in the Metroplex, Houston Texas (EA), January 2013**

Agency/Organization	Name	Letter Date	Comment Number	Comment	FAA Response
Self	Michael Kroposki Esq.	2/13/2013	MK1	<p>I was very surprised to see in the draft EA reference to calculated DNL differences of 0.1 and 0.4 dB. These numbers are meaningless and therefore misleading because it is my understanding that the precision and accuracy of INM output is at best about +/- 1.0 dB [1]. The NIRS software uses the INM computation engine for its output. So the NIRS output can have no better precision and accuracy and it is most likely less due to errors introduced by truncation, round off and averaging in NIRS.</p> <p>[1] "Aircraft Noise Measurement and Modeling" at A 4.7.2.5 [FAA note: copy attached to Mr. Kroposki's comment]</p>	As required by FAA Order 1050.1E, Appendix A, Section 14.5e, the FAA conducted noise modeling using NIRS. The noise values presented in the EA are those reported by NIRS. Reporting NIRS results to 0.1 dB is consistent with FAA Order 1050.1E, under which the criterion for significant noise impact is expressed to units of 0.1 dB (see FAA Order 1050.1E, Appendix A, Section 14.3).
Self	Michael Kroposki Esq.	2/13/2013	MK2	<p>The final EA should state the uncertainty in the output by indicating the validated range of the output DNL numbers, for example 63.45 +/- 0.05dB. This range is commonly referred to as the confidence interval.</p>	FAA noise models do not provide confidence interval information. Please note that the FAA has not revised the EA, except as noted on the errata sheet attached to the FONSI/ROD.
Self	Michael Kroposki Esq.	2/13/2013	MK3	<p>Since AEDT uses modeling calculations very similar to INM it is likely the levels of uncertainty found in the AEDT sensitivity analysis are likely to also be present in the NIRS output. The recently studied AEDT levels of uncertainty should be cited as a guide to uncertainty in the similar NIRS software.</p>	The AEDT uncertainty quantification, still under preparation, is not an applicable guide for previous noise models. The AEDT uncertainty quantification will not provide confidence interval information for noise.
Self	Michael Kroposki Esq.	2/13/2013	MK4	<p>The FAA NEPA regulations use round numbers, for example 65 dB, with one exception 1.5 dB Order 1050.1E section 14), however table 23 in the draft EA on page 134 cites the noise limits as 3.0 and 5.0 dB. These are different numbers. A change of 4.6 dB would satisfy the 5 dB limit under standard numerical nomenclature [3] but would not be valid according to table 23. Table 23 should be revised in accordance with Order 1050.1E</p> <p>[3] "Statistical Concepts in Metrology, NBS Publication 747 at pages 1-2 [FAA note: copy attached to Mr. Kroposki's comment]</p>	<p>It is not necessary to revise Table 23 because the presentation of the significant impact threshold and the reportable increase criteria in Table 23 is consistent with FAA Order 1050.1E, Appendix A, Section 14. The fact that some numbers in the Order are presented as whole numbers is a matter of formatting and stylistic convention.</p> <p>The EA presents the noise values as reported by NIRS, which rounds to the nearest 1/10th of a dB. For comparison to noise criteria, FAA applied the criteria to the nearest 1/10th of a dB (consistent with the above reference Appendix A, section 14.3). FAA consistently applied this methodology throughout the analysis.</p>

**Comments and Responses
to the Houston Optimization of Airspace and Procedures in the Metroplex, Houston Texas (EA), January 2013**

Agency/Organization	Name	Letter Date	Comment Number	Comment	FAA Response
Self	Michael Kroposki Esq.	2/13/2013	MK5	In Appendix G it is stated on page G-6 that " G.3.1 Aircraft Noise Performance Specific noise and performance data must be entered into the NIRS for each aircraft type operating at the airport. Noise data are included in the form of sound exposure levels (SELs) at a range of distances from a particular aircraft with engines at a specific thrust level. Performance data include thrust, speed and altitude profiles for takeoff and landing operations. The NIRS database contains standard noise and performance data for over one hundred different fixed wing aircraft types, most of which are civilian aircraft. The NIRS automatically accesses the noise and performance data for takeoff and landing operations by those aircraft."	The quoted text is not from the Houston OAPM EA.
Self	Michael Kroposki Esq.	2/13/2013	MK6	The EA should state specifically how take off weights were determined.	The EA, in Appendix G.2 at page G-38, Section 4.7 "Aircraft Stage Length," provides a discussion of the departure weight estimates.
Self	Michael Kroposki Esq.	2/13/2013	MK7	While use of the default settings 65% payload may have been realistic in 1970, the current Load Factors clearly show it is not so today. A more realistic average weight is most likely near 100% payload.	A load factor of near 100% is not realistic when computing average annual day conditions. NIRS has a Total Payload factor built into the model. Therefore, the average weight calculation includes more than passenger load factor. It also includes the weight of the aircraft, cargo, and fuel.
Self	Michael Kroposki Esq.	2/13/2013	MK8	INM noise calculations are especially sensitive to variations in take off weight. One study of input sensitivities has shown that a 10% variation in take off weight leads to an error of 3-7 dB [2]. [2] "Sensitivity of FAA Integrated Noise Model to Input Parameters" , Applied Acoustics, Vol 66, Issue 3, March 2005, pages 263-276 [FAA note: copy attached to Mr. Kroposki's comment]	Noise calculations are sensitive to many noise modeling input variables. It is not technically sound to look at one variable, e.g., takeoff weight, in isolation. The commenter has misstated the data in the referenced study. The study reports in Table 8 that a 10 percent weight increase can result in a SEL (not DNL) variation of +0.70 decibels to +2.20 decibels.

**Comments and Responses
to the Houston Optimization of Airspace and Procedures in the Metroplex, Houston Texas (EA), January 2013**

Agency/Organization	Name	Letter Date	Comment Number	Comment	FAA Response
Self	Michael Kroposki Esq.	2/13/2013	MK9	...since large jet aircraft are most likely the largest contributors of noise energy, an error in the largest contributors to DNL will predominate since noise as measured by DNL is aggregated logarithmically. Assuming unrealistically low take off weights have been used in the draft EA, it may be assumed that the calculated DNL's are significantly underestimated!	The commenter's assumption that calculated DNLs are significantly underestimated is not accurate and appears to be based on his assumption that the passenger load factor is the prevailing variable in the noise model. Noise calculations are sensitive to many noise modeling input variables. For example, the noise model uses a conservative value of 100% thrust for departure procedures, although airlines typically do not use 100% power in takeoff. Thrust reduction at takeoff varies. Therefore, the 100% thrust assumption will result in higher noise calculations than may occur for particular departures. The goal of the noise analysis is to capture the average annual conditions at the airport.
Self	Michael Kroposki Esq.	2/13/2013	MK10	The consequence of this conclusion [see MK9] has direct impact on the overall environmental impact determination because even the underestimated DNLs were extremely close to the 65 dB level of regulatory significance. If they in fact exceed this level, a finding of significant impacts is warranted instead of the finding of no significant impact stated in the EA draft.	See response to MK9 above. Noise calculations are sensitive to many noise modeling input variables. The goal of the noise analysis is to capture the average annual conditions at the airport. The FAA has determined that the DNL results do not exceed the FAA's threshold for a significant noise impact. Please note that the FAA has not revised the EA, except as noted on the errata sheet attached to the FONSI-ROD.

**Comments and Responses
to the Houston Optimization of Airspace and Procedures in the Metroplex, Houston Texas (EA), January 2013**

Agency/Organization	Name	Letter Date	Comment Number	Comment	FAA Response
Self	Tim Honeycutt	2/19/2013	TH1	I sincerely hope that the proposed flight paths for DRLLR STAR and AGGEE STAR do <u>not</u> track through High Meadow Ranch. The property owners in High Meadow Ranch are affluent and chose the area for its serenity. None would appreciate a steady stream of descending aircraft on an unwavering GPS-dictated ground track, even if in some sort of continuous descent profile.	<p>The FAA understands the High Meadow Ranch area to be south of Magnolia, Texas, and located in an area roughly described as east of Magnolia Waller Road, north of Butera Road/Walter Creek Rd./Stagecoach Rd./Decker Prairie Rd., and southwest of Magnolia Blvd/Farm-to-Market 1774 Road/Rt. 249.</p> <p>Operations that use the existing AGGEE STAR currently pass over the High Meadow Ranch area. Under the Proposed Action, the AGGEE STAR will be cancelled and these operations will use the proposed MSCOT STAR during west flow arrivals to IAH. The proposed MSCOT STAR would be 2 to 3 statute miles northeast of AGGEE and approximately 1 statute mile northeast of Magnolia Blvd/Farm-to-Market 1774 Road. (Distances are approximate and can vary depending on the measurement location as these paths are not parallel.)</p> <p>Operations that use the existing BAZBL STAR would use the proposed DRLLR STAR during west flow arrivals to IAH. The BAZBL STAR is currently east of Magnolia, Texas, and the proposed DRLLR STAR would be further east of the BAZBL STAR and Magnolia. The Proposed Action would cancel the existing BAZBL STAR.</p> <p>In summary, neither the proposed MSCOT nor DRLLR STARs would overlie the High Meadow Ranch area. Rather, both the proposed MSCOT and DRLLR STARs would be further east than the existing procedures that they would replace.</p>
Self	Tim Honeycutt	2/19/2013	TH2	Will the proposed ground track for DRLLR STAR and/or AGGEE STAR pass through High Meadow Ranch in Magnolia, Texas? Figure 3 on page 79 of appendix F of the Houston OAPM EA is devilishly hard to decipher without access to really detailed aeronautical charts.	See response to TH1 above.

Errata Sheet
to the Houston Optimization of
Airspace and Procedures in the
Metroplex, Houston Texas (EA),
January 2013

Errata
June 2013

This errata sheet corrects errors or omissions that were identified after printing of the *Environmental Assessment for Houston Optimization of Airspace and Procedures in the Metroplex* in January 2013. This errata sheet must be attached to the EA to comprise a full and complete record of the environmental analysis for the project. The EA will not be reprinted.

Section	Page	Correction	Note
3.4	76	Table 11: “AGEEE” should be “AGGEE”	
5.1	132	Table 22; Row for Compatible Land Use: "Proposed Action would result in one population centroid being added to the area exposed to DNL 65 dB and higher, but it would experience a significant noise increase (DNL 1.5 dB or greater). No significant impact." should be "Proposed Action would result in some locations being added to the area exposed to DNL 65 dB and higher, but it would not experience a significant noise increase (DNL 1.5 dB or greater). No significant impact."	The phrase “some locations” is used to better summarize the text on pages 144 and 152. We could also use “some population centroids” to match the text on page 153. Also, the word “not” was omitted from the sentence.
5.6.1	158	“State Implementation Plna” should be “State Implementation Plan”	
Appendix G	G-39	Table 15 should be replaced with the attached revision.	The revised Table 15 corrects inadvertent documentation errors discovered after release of the EA. As revised, the table accurately reflects the inputs used in the noise modeling. Therefore, this revision to Table 15 does not affect the EA’s analysis, results, or conclusions

Table 15 Average Annual Day Departure Stage Length Assignments for Noise Modeling

Category	HOU			IAH		
	2012	2014	2019	2012	2014	2019
D-1	172.07	175.15	202.98	282.76	306.45	353.08
D-2	55.05	55.97	77.08	300.71	324.42	359.25
D-3	20.65	21.17	29.49	132.32	147.29	169.62
D-4	0.00	0.00	0.00	28.07	31.81	35.76
D-5	0.00	0.00	0.00	3.56	3.67	3.80
D-6	0.00	0.00	0.00	12.08	12.74	40.37
D-7	0.00	0.00	0.00	2.50	2.71	11.43
D-8	0.00	0.00	0.00	1.21	1.39	9.14
D-9	0.00	0.00	0.00	0.46	0.46	0.46
Total	247.77	252.29	309.55	763.66	830.94	982.91

Category	DWH			EFD			IWS		
	2012	2014	2019	2012	2014	2019	2012	2014	2019
D-1	30.65	31.17	32.50	28.84	28.84	28.84	19.23	19.48	20.90
D-2	0.08	0.08	0.08	0.44	0.44	0.44	0.00	0.00	0.00
D-3	0.03	0.03	0.03	0.27	0.27	0.27	0.00	0.00	0.00
D-4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
D-5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
D-6	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
D-7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
D-8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
D-9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Total	30.77	31.28	32.61	29.55	29.55	29.55	19.23	19.48	20.90

Category	LBX			SGR		
	2012	2014	2019	2012	2014	2019
D-1	5.33	5.43	6.03	34.33	35.07	37.02
D-2	0.00	0.00	0.00	0.35	0.35	0.35
D-3	1.90	1.93	2.16	0.04	0.04	0.04
D-4	0.00	0.00	0.00	0.00	0.00	0.00
D-5	0.00	0.00	0.00	0.00	0.00	0.00
D-6	0.00	0.00	0.00	0.00	0.00	0.00
D-7	0.00	0.00	0.00	0.00	0.00	0.00
D-8	0.00	0.00	0.00	0.00	0.00	0.00
D-9	0.00	0.00	0.00	0.00	0.00	0.00
Total	7.23	7.36	8.19	34.72	35.46	37.40

Note: Totals and subtotals may not match exactly due to rounding

Source: Flight Plan Data and Radar Data from National Offload Program, (Mitre Corp.); FAA ATADS (2011); FAA TAF (2012); HMMH Analysis (2012)