DOT/FAA/TC-25/36

Federal Aviation Administration
William J. Hughes Technical Center
Aviation Research Division
Atlantic City International Airport
New Jersey 08405

Development of Serviceability Level (SL) Index Model for Extended Airport Pavement Life

January 2026

Final report

U.S. Department of Transportation
Federal Aviation Administration

NOTICE

This document is disseminated under the sponsorship of the U.S. Department of Transportation in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof. The U.S. Government does not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the objective of this report. The findings and conclusions in this report are those of the author(s) and do not necessarily represent the views of the funding agency. This document does not constitute FAA policy. Consult the FAA sponsoring organization listed on the Technical Documentation page as to its use.

This report is available at the Federal Aviation Administration William J. Hughes Technical Center's Full-Text Technical Reports page: actlibrary.tc.faa.gov in Adobe Acrobat portable document format (PDF).

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

1 Report No. DOT/FAA/TC-25/36	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle DEVELOPMENT OF SERVICEABILITY LEVEL (SL) INDEX MODEL FOR EXTENDED AIRPORT PAVEMENT LIFE		5. Report Date January 2026 6. Performing Organization Code
7. Author(s) Ali Z. Ashtiani, Adam Amos-Binks, Scott Murrell, Ebenezer Duah, and Aaron Williams		8. Performing Organization Report No.
9. Performing Organization Name and Address Applied Research Associates, Inc. 2628 Fire Rd Egg Harbor Township, NJ 08234		10. Work Unit No. (TRAIS) 11. Contract or Grant No. 692M15-22-T-00028
12. Sponsoring Agency Name and Address U.S. Department of Transportation Federal Aviation Administration Airport Engineering Division 800 Independence Ave., SW Washington, D.C. 20591		13. Type of Report and Period Covered Final Report 14. Sponsoring Agency Code
15. Supplementary Notes The Federal Aviation Administration Program Office/Technical Officer is Dr. David Brill, ANG-E262.		

16. Abstract

The Federal Aviation Administration (FAA) initiated the Extended Airport Pavement Life (EAPL) program to evaluate and enhance the long-term performance of airport pavements, aiming to extend their service life beyond the 20-year current standard considered in pavement thickness design. The FAA collected extensive pavement performance data - including surface groove geometry, longitudinal profile roughness, and surface distresses - from 22 major U.S. airports. The data included both flexible and rigid pavements. Data analysis shows that pavements designed to fail structurally in 20 years often remain structurally intact but exhibit functional failure sooner than the intended design life. It was found that with effective routine and preventive maintenance, pavements can remain serviceable well past their original design life. To quantify this extended serviceability, the FAA introduced the serviceability level (SL) index, a combined measure of structural integrity and functional condition that indicates a pavement's suitability for aircraft operations. Supplementary data from each airport was gathered by the FAA to support the analysis, including material characterizations, pavement cores, maintenance histories, runway usage, and weather data. All data have been consolidated in a dedicated database, PA40.

This study documents the analysis of the PA40 data to calculate key pavement condition indexes contributing to overall serviceability and to explore their relationships with factors like weather, traffic, material properties, and pavement structure. Machine learning (ML) models were developed to predict the SL index for both flexible and rigid airport pavements. The SL model is structured as a classification ML task, aiming to classify the runway pavement as either serviceable or unserviceable, based on a set of pavement condition indexes including pavement condition (structural and non-structural), runway roughness, and groove condition. In addition, individual ML models were created to predict each of these indexes as part of the SL model, incorporating predictors such as pavement age, environmental factors, and traffic.

The ML models were developed in Python and integrated into a Visual Basic (VB) library in anticipation of future use in standard FAA pavement design and management programs, including FAARFIELD.

17. Key Words Pavement performance, Airport pavement, Pavement life, Serviceability, Machine learning, performance index, design life, PA40	18. Distribution Statement This document is available to the U.S. public through the National Technical Information Service (NTIS), Springfield, Virginia 22161. This document is also available from the Federal Aviation Administration William J. Hughes Technical Center at actlibrary.tc.faa.gov.

19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 122	22. Price

Contents

Figures

Tables

Acronyms

Acronym	Definition
AC	Advisory Circular
ADASYN	Adaptive synthetic sampling
Anti-SCI	Anti-structural condition index (Non-load-related component of PCI)
ATL	Hartsfield-Jackson Atlanta International Airport
AUC	Area under the curve
Avg Temp	Average daily temperature
BBI	Boeing Bump Index
BOS	Boston Logan International Airport
BWI	Baltimore/Washington International Thurgood Marshall Airport
CMH	John Glenn Columbus International Airport
DCA	Ronald Reagan Washington National Airport
DEN	Denver International Airport
DFW	Dallas Fort Worth International Airport
DPrec	Days precipitation
EAPL	Extended airport pavement life
FAA	Federal Aviation Administration
FAARFIELD	FAA Rigid and Flexible Iterative Elastic Layered Design
FDD	Freeze degree days
FLL	Fort Lauderdale International Airport
FOD	Foreign object damage
FPD	Freeze precipitation days
FPR	False positive rate
FThC	Freeze-thaw cycles
GSO	Piedmont Triad International Airport
HDBSCAN	Hierarchal density-based spatial clustering of applications with noise
HWD	Heavyweight deflectometer
HydD	Hydration days
IAD	Dulles International Airport
IAH	George Bush Intercontinental Airport
IND	Indianapolis International Airport
LAX	Los Angeles International Airport
LGA	LaGuardia Airport

MCI	Kansas City International Airport
MIA	Miami International Airport
ML	Machine learning
MST	Model search triple approach
ORD	Chicago O'Hare International Airport
PCC	Portland cement concrete
PCI	Pavement condition index
PMS	Pavement Management System
PMML	Predictive Model Markup Language
RH	Relative humidity
RF	Random forest
RMSE	Root mean square error
ROC	Receiver operating characteristic
RRI	Runway Roughness Index
SCI	Structural Condition Index (load-related component of PCI)
SEA	Seattle-Tacoma International Airport
SFO	San Francisco International Airport
SL	Serviceability level
SLC	Salt Lake City International Airport
SHAP	SHapley Additive exPlanations
SMOTE	Synthetic minority oversampling technique
Temp90	Days with temperature over 90°F
SVR	Support vector regression
TPrec	Total precipitation
TPR	True positive rate
TUS	Tucson International Airport
UMAP	Uniform manifold approximation and projection
VB	Visual Basic

Executive summary

The current Federal Aviation Administration (FAA) pavement design methodology is based on structural failure and does not directly address functional failures caused by issues such as excessive profile roughness, foreign object damage (FOD) potential, non-load-related material distresses or loss of tractive surface. As a result, airport pavements designed to the 20-year structural life standard may experience functional failures before reaching their anticipated design life. However, with proper regular maintenance pavement can remain serviceable beyond their intended design life.

The FAA initiated the Extended Airport Pavement Life (EAPL) program with the goal of developing methods to prolong the serviceable life of airport pavements beyond 20 years. To support this initiative, the FAA has collected data from 28 runways inclusive of both flexible and rigid pavements at 22 major U.S. airports, encompassing pavement distress, surface groove geometry, longitudinal profile roughness, material characterization, pavement cores, maintenance work histories, runway usage, and weather data. The data are stored in a dedicated database, designated PA40.

To formulate the EAPL concept, the FAA proposed the establishment of the Serviceability Level index (SL), a performance metric designed to assess pavement suitability for aircraft operations by providing smooth ride, tractive surface and low FOD potential calculations. The SL incorporates various load-related and non-load-related pavement condition indexes. The end of pavement life is reached when SL falls below a set threshold that maintenance cannot restore. This study uses machine learning (ML) techniques to develop predictive models for SL and underlying condition indexes, using inputs such as pavement age, weather variables, and traffic parameters.

The report examines the EAPL database to identify key variables influencing pavement longevity and key pavement condition indexes that contribute to decisions on maintenance or replacement. Five pavement performance indexes are calculated for each runway section. Distress survey data are used to calculate Pavement Condition Index (PCI), Structural Condition Index (SCI), and anti-Structural Condition Index (anti-SCI). Longitudinal profile data are used to calculate Runway Roughness Index (RRI). FAA inertial profiler data are used to calculate the Groove Index (GI).

1 Introduction

1.1 Airport pavement serviceable life

The FAA pavement thickness design procedure considers 20 years as the standard design life for new pavements and overlays on federally funded airport projects. This design life corresponds to the structural life during which the pavement must withstand specified design traffic. The design process anticipates that the pavement will reach failure at the end of its design life. Pavement failure, as defined by the FAA for various pavement categories and design scenarios, occurs when the pavement loses its structural integrity and can no longer support aircraft loads safely. The current version of the FAA Rigid and Flexible Iterative Elastic Layered Design (FAARFIELD) software (version 2.1) employs structural failure models based on full-scale tests to determine pavement layer thicknesses that can support the design traffic over the targeted design life (FAA, 2023). For flexible pavements, the structural failure model is characterized by excessive permanent shear deformation of the subgrade that leads to rutting at the pavement surface, or excessive alligator cracking where the asphalt surface can no longer provide waterproofing. For rigid pavements, the structural failure model is characterized by extensive structural cracking, i.e., 50% of slabs exhibiting linear cracking. However, the existing design method does not directly address functional failures caused by issues such as excessive profile roughness, foreign object damage (FOD) potential, non-load related material distresses, or loss of tractive surface. These functional failures are primarily associated with material or construction problems.

In practice, airport pavements often exhibit longer-than-intended structural life, primarily due to conservative design methods and overestimation of forecast design traffic. This can result in pavements designed to meet failure criteria remaining serviceable beyond the intended 20-year design life with proper maintenance. Conversely, some pavements designed to the current 20-year structural life standard experience functional failure before reaching their expected design life; while the underlying structure is intact, the pavement can be rendered unserviceable before structural failure.

Routine and preventive maintenance (e.g., crack sealing, rubber removal, grooving), along with rehabilitation measures (e.g., milling and overlaying an asphalt surface) to correct functional failures, can be applied to extend the serviceable life of pavement beyond 20 years. Prolonging the service life of airfield pavement reduces the frequency of reconstruction, which minimizes associated delays and costs to benefit both taxpayers and travelers.

1.2 Extended airport pavement life

The FAA Airport R&D Branch initiated the Extended Airport Pavement Life (EAPL) program with the goal of developing models for predicting long-term airport pavement performance. Extended-life design requires an expanded definition of failure, encompassing both structural and functional failure modes, the latter triggered by factors such as FOD potential, high profile roughness, and/or reduced skid resistance. To formulate this concept, the FAA proposed a performance index called Serviceability Level (SL), a metric reflecting the pavement's suitability for aircraft use based on load-related and non-load-related components. The end of pavement life condition can be defined in terms of an SL threshold. This definition aligns with concepts used in life-cycle cost analysis (LCCA), where the end of pavement life is considered to occur when the annual cost of maintaining continued service at a safe operational level surpasses the annualized cost of reconstruction. For flexible pavements the FAA defines full reconstruction as the replacement of at least the asphalt surface and base layers, including the subbase if a stabilized base design is required (Brill & Parsons, 2017). The subbase and subgrade material that is deemed structurally intact may remain in place. Full reconstruction for rigid pavement involves the replacement of Portland cement concrete (PCC) slabs at a minimum, with the base and subgrade retained if intact. When a limited number of slabs (20 percent or less) are replaced, the project is considered rehabilitation rather than reconstruction.

Developing long-term pavement performance models for EAPL design necessitates collecting accurate and recurring pavement condition data over an extended period, generally 15 to 20 years. These models seek to identify correlations between pavement condition indexes and a range of variables that influence pavement performance. These variables include environmental condition, runway usage, material properties, and pavement structure, all of which need to be quantified. Since 2012, the FAA has been gathering various airfield pavement data under the EAPL program from 28 runways at 22 major airports in the continental U.S., with 14 flexible pavements and 14 rigid pavements. The data includes pavement performance measures, such as surface groove geometry, profile roughness and surface distresses, as well as maintenance work histories, runway usage, and weather data.

1.3 Background on FAA EAPL modeling

In 2018, the FAA introduced an initial framework for an age-based SL index based on the data obtained from the EAPL (Ashtiani, Shirazi, Murrell, Spier, & Brill, 2019). This SL model used logistic regression methods to predict the probability of a pavement section being serviceable based on PCI, a roughness index (RRI) and a GI. The model regarded pavement age as the sole

explanatory variable. Regression models were developed to predict each component of that SL index, using pavement age as the sole predictor.

While the age-based SL model was informative, it assumed that age serves as a surrogate for both traffic and weather events. It did not account explicitly for various environmental conditions and traffic levels, which could potentially lead to inaccurate prediction of rehabilitation intervention timing (i.e., model incorrectly predicts equal service life for two sections of the same age but located in different climate zones or supporting widely different levels of traffic). In 2021, the FAA extended its research to enhance the SL model by incorporating additional design factors, including weather and traffic, to improve accuracy. Conventional regression analysis was found to be insufficient, prompting the exploration of machine learning (ML) techniques to handle multiple design features effectively. The 2021 study investigated various ML methods to determine which applications are best suited for analysis of the EAPL data (Ashtiani, 2021). The initial method validation involved developing univariate models for three pavement condition indexes namely, PCI, structural condition index (SCI) and anti-structural condition index (anti-SCI), using pavement age as the sole predictor. While similar pavement deterioration curves were observed in the regression models, the 2021 study emphasized the need for feature selection to avoid over-specifying models. Ashtiani et al. (2021) performed feature selection analysis, evaluating the impact of environmental variables on pavement performance. Several feature ranking and subset selection algorithms were implemented using anti-SCI, a derivative of the PCI containing only those distresses that are not directly caused by aircraft loads, as the pavement condition index. Clustering and Principal Component Analysis (PCA) were applied for dimensionality reduction.

Subsequent research by Ashtiani et al. (2021) (2022) enhanced the age-based anti-SCI model by incorporating environmental variables, resulting in improved model accuracy. This updated model is based on an autoregressive approach and used random forest as the learning algorithm. These advancements signify a move towards more robust ML-based models for predicting pavement performance, considering multiple influential factors.

2 Research objectives

The objective of the research described in this report was to develop models based on ML techniques capable of predicting SL and underlying condition indexes using features such as pavement age, weather variables, and traffic parameters. The result of this research will support the FAA's goal of extended-life airport pavement design. The model development has five major components:

1. EAPL data processing
2. Feature engineering
3. Development of ML models for different pavement condition indexes.
4. Development of model for SL index.
5. Implementation of the models in a .NET application.

This report reviews the EAPL database encompassing pavement condition data, environmental parameters, traffic data, design, construction, and rehabilitation related data from various airports. The intent is, first, to identify the key variables in the EAPL data that most influence pavement longevity and, second, to identify key pavement condition indexes or combinations of indexes tied to pavement failure or a decision to rehabilitate/reconstruct/replace. The findings are used to create a dataset for training the ML prediction models. The report also incorporates various feature engineering methods to identify trends/correlations in the EAPL data, focus on the key variables that most influence pavement performance, and eliminate variables that have little or no influence. Results from feature engineering are used to develop ML prediction models for different pavement condition indexes. Finally, an ML model is developed for the SL index. The developed models are implemented in a format suitable for integration in a .NET application.

3 Extended airport pavement life data

3.1 Data collection

Since 2012, the FAA has collected a diverse set of data from airports participating in the EAPL program. Some of these data were provided by the airports, including PCI, construction/maintenance, material/geotechnical, design, and traffic data. The FAA conducted field visits to 19 runways -10 with asphalt surface and 9 with PCC surface - to augment and enhance the dataset. Field data collection efforts typically include distress survey, heavy weight deflectometer (HWD) testing, longitudinal profile roughness measurement, groove measurement, material sampling, and pavement coring. To ensure thorough monitoring of the pavement condition over time, some airports have been visited multiple times, typically 5 to 8 years apart.

3.2 Extended airport pavement life database

The FAA stores data collected under the EAPL program in a dedicated database, designated as PA40. The PA40 database is built on the FAA PAVEAIR 3.0 platform, a web-based application designed to fulfill the requirements of an Airport Pavement Management System (PMS), as identified in Advisory Circular (AC) 150/5380-7B (FAA, 2014; FAA, 2024). PA40 contains tables and functions not currently implemented in the public version of FAA PAVEAIR. Structured similarly to other PMS databases, PA40 stores its data in a Microsoft SQL Server database. The data organization follows the standard pavement management inventory hierarchy, as illustrated in . Networks within PA40 can comprise various branches, such as runways, shoulders, and overruns. Each runway is typically divided transversely and longitudinally into sections. The determination of the number and configuration of sections is usually conducted by the airports and PA40 follows the sectioning laid out in the airport's PMS.

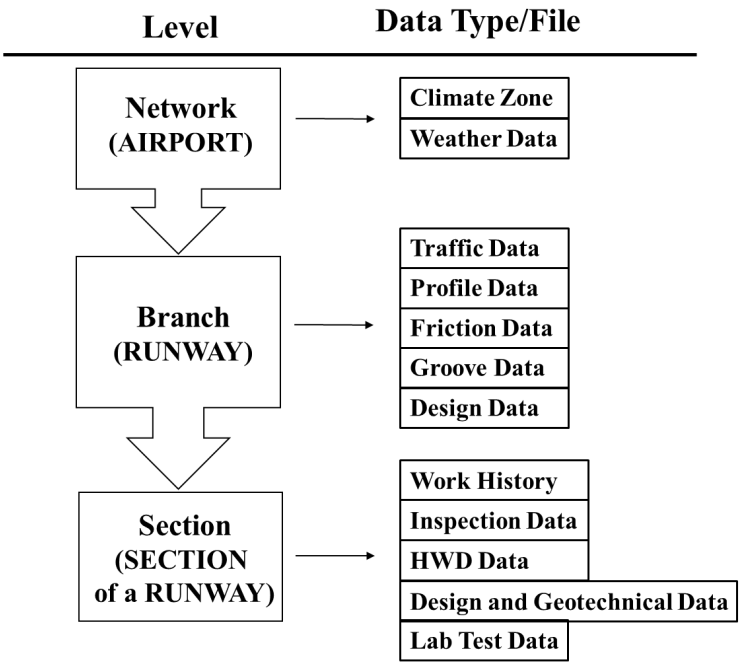


Figure 1. Data structure in PA40 database

PA40 includes data from 27 runways at 22 U.S. airports, as shown in Figure 2. These include 17 of the 31 designated as large hub airports in the United States in 2024, as well as three medium-hub and two small-hub airports. The dataset includes pavement performance-related data and additional fields, as listed below:

- Pavement performance data
  - Distress inspection
  - Longitudinal profile roughness
  - Groove condition
  - HWD testing
- Weather data
- Traffic data (historical runway usage)
- Structural design and as-built pavement structure data
- Field core and laboratory test data

Data were collected from three primary sources: directly from the airports (e.g., PCI data, design/construction/maintenance data), field visits conducted by the FAA (e.g., performance data, material sampling data), and other third-party resources (traffic and weather data). The combination of these diverse data sources contributes to a comprehensive understanding of different aspects affecting airport pavement performance.

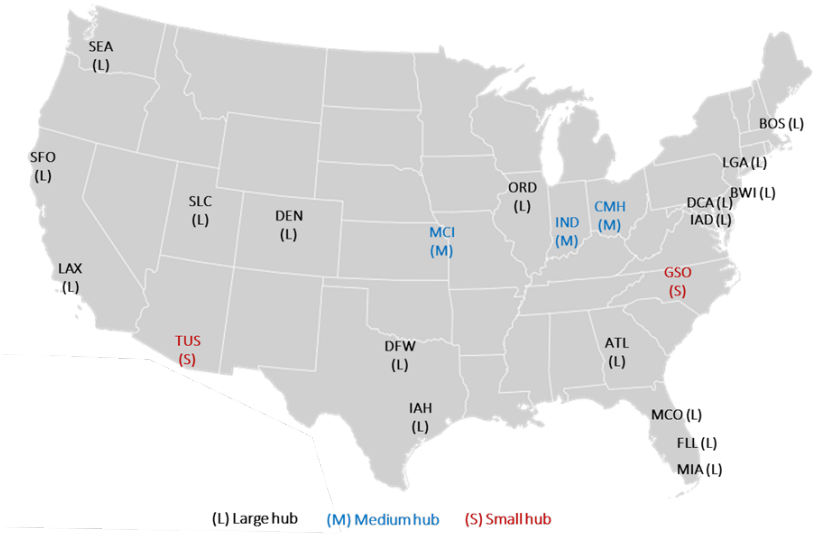


Figure 2. Twenty-two airports in PA40 database

3.3 Airport pavement performance data

The PA40 database stores key performance-related data metrics that significantly influence assessment of runway serviceability. This study specifically concentrates on three of the four pavement performance metrics listed in section 3.2: surface distress, pavement profile roughness

and surface skid resistance. PA40 includes tools for the calculation of indexes associated with each of these data types. The study also considers input data from the following data sources: historical weather data, historical runway usage (traffic) data, pavement work history, and design traffic (which influences the pavement structure but is distinct from the actual traffic history).

3.3.1 Surface distress

PA40 calculates three indexes using the distress inspection data: PCI, SCI, and anti-SCI. These indexes are calculated at the section level. The PCI, as per ASTM D5340 (ASTM International, 2020), serves as a widely accepted indicator of overall airport pavement condition, commonly employed in the development of pavement management plans. The PCI amalgamates structural and non-structural distresses in a single index on a 0 to 100 scale, with 100 indicating a non-distressed surface. SCI contains the load-related components, while anti-SCI contains the non-load-related components of PCI, with each scaled from 0 to 100. For rigid pavements in FAA design, the distresses included in SCI are corner break, shattered slab, shrinkage cracks, longitudinal/transverse/diagonal cracking, joint spalling, and corner spalling (FAA, 2021). An SCI value of 80 or below aligns with the FAA's definition of structural failure for rigid pavement, signifying that 50 percent of slabs in the traffic area exhibit a structural crack. For flexible pavements, SCI is not defined in any current standard, but the calculation in PA40 incorporates alligator cracking and rutting. Currently, there is no FAA-established standard for the minimum value PCI index that corresponds to serviceability failure for either rigid or flexible pavements.

The PCI data within the PA40 database includes both field data collected by the FAA and the data supplied from the airport PMS. Some runways surveyed were relatively new and therefore have no historic PCI. The surveys performed by the FAA are based on a 100 percent sampling of the runway pavements, while those provided by the airports are typically based on statistical sampling, i.e., less than 100 percent, as prescribed by ASTM D5340.

3.3.2 Pavement roughness

The FAA classifies airfield pavement roughness into two categories based on the dimensions and frequency of surface deviations:

1. Single Event Bump. AC-150/5380-9 (FAA, 2009), page 1, defines single event bumps as "isolated events where changes in pavement elevation occur over a relatively short distance of 100 meters (328 feet) or less." The Boeing Co. (Boeing, 2002) developed a method to assess single event roughness, called the Boeing Bump Index (BBI), which was incorporated into AC 150/5380-9.

2. Profile Roughness. AC-150/5380-9 (FAA, 2009), page 1, defines profile roughness ss "surface profile deviations present over a portion of the runway that cause airplanes to respond in ways that can increase fatigue on airplane components, reduce braking action, impair cockpit operations, and/or cause discomfort to passengers." While passengers may not perceive surface roughness, it can impact aircraft components and compromise operational safety.

In 2009 as part of AC 150/5380-9, the FAA introduced ProFAA, a computer program designed to simulate an aircraft's response to a pavement profile, including models of aircraft such as the McDonnell Douglas DC-9 and DC-10 and the Boeing 727 and 747. This program calculates various pavement profile roughness indexes, including BBI, International Roughness Index (IRI), Straightedge, California Profilograph, and root mean square bandpass profile. Notably, there is currently no industry standard defining the acceptable limits for these roughness indexes concerning airport pavements.

The FAA developed a new roughness index for airports (RRI), using data from the B737 aircraft simulator at the FAA Mike Monroney Aeronautical Center in Oklahoma City, OK (Kuncas, 2021). RRI is defined as the scaled ISO 2361-1 weighted root mean square (WtRMS) value of vertical acceleration, computed for the model B737 at the pilot station using the ProFAA software, as shown in Equation 1.

$$RMS\ RRI = 0.378 \times WtRMS \tag{1}$$

The RRI is designed for runway pavements, assuming aircraft speed of 100 knots. RRI uses a threshold value system similar to BBI (Boeing 2002). The scale factor of 0.378 s^2/m is applied to align the RRI scale with that of the BBI when evaluated on a runway. The development of the RRI involved statistical analyses using the average BBI, pilot subjective rating, and acceleration data from the Mike Monroney Aeronautical Center simulator.

According to Kuncas (2021), three threshold categories are proposed based on single event BBI and the average RRI of the entire runway: acceptable, excessive, and unacceptable. Values falling into the "excessive" category indicate that corrective action is necessary, but the runway can still be operational. If a runway surpasses the "unacceptable" threshold, operations must cease until corrective measures are implemented to prevent potential damage to aircraft. Table 1 shows the thresholds for each category. Table 2 shows the index ranges corresponding to the pilot ratings.

Table 1. Runway roughness index threshold (Kuncas, 2021)

Index	**Maximum acceptable threshold**	**Comment**
RMS RRI	> 0.75	Excessive
	0.37-0.75	Scheduled monitoring
	< 0.37	No action required

Table 2. Runway roughness index limits corresponding to pilot rating (Kuncas, 2021)

Percent of pilots rate the runway as unacceptable	**RMS RRI**
5%	0.13
10%	0.18
50%	0.34
95%	0.67

The FAA collected longitudinal profile data for specific runways within the PA40 database using two devices: a commercial walk-behind profiler (SurPro) and an internally developed FAA Inertial Profiler (FIP) device (Figure 2a). The measurements are conducted along the runway length at typically four or five offsets from the runway centerline, covering both the keel and outside-the-keel areas. However, for most runways, the measurements were only conducted on the keel area.

(a) (b)

Figure 2a. Profile data collection using (a) the SurPro walk-behind profiler and (b) the FIP (visible attached to the passenger-side cab door).

PA40 has the capability to calculate various roughness indexes, including the RRI. The calculated indexes represent average values over the length of the measured pavement profile as included in the profile file. Illustrated in Figure 3 is a sample table PA40 output featuring calculated roughness indexes for a designated network (a large hub airport) and branch (runway). The table provides complementary information, such as the type of the profiler and the testing date, and provides a profile file link for downloading.

Current Database: ATL

Network Large Hub

Branch Runway

Name	Type	Date	Database	Band Pass	Boeing Bump	Boeing Bump Max	Straight Edge	CA Profilograph	IRI	RRI	Creation Date	Creation User	Download File	File Size
21FE0110 - 1.pro	SurPro	2/21/2017		0.041456	0.248400		0.182100	27.545500	148.051900	0.320972	9/19/2017	faa	21FE0110 - 1.pro	604572.00
21FE0228 - 1.pro	SurPro	2/21/2017		0.046944	0.281600		0.188700	32.859100	156.384800	0.348464	9/19/2017	faa	21FE0228 - 1.pro	604612.00
21FE0341 - 1.pro	SurPro	2/21/2017		0.044119	0.256200		0.182600	30.648600	147.275900	0.316920	9/19/2017	faa	21FE0341 - 1.pro	604412.00
21FE0454 - 1.pro	SurPro	2/21/2017		0.057036	0.317900		0.230400	45.149800	192.523000	0.381882	9/19/2017	faa	21FE0454 - 1.pro	604852.00
22FE0102 - 1.pro	SurPro	2/22/2017		0.055538	0.303800		0.226200	44.168100	183.655800	0.360698	9/19/2017	faa	22FE0102 - 1.pro	604212.00
22FE0333 - 1.pro	SurPro	2/22/2017		0.048592	0.290000		0.189000	33.110000	154.820600	0.338278	9/19/2017	faa	22FE0333 - 1.pro	604412.00

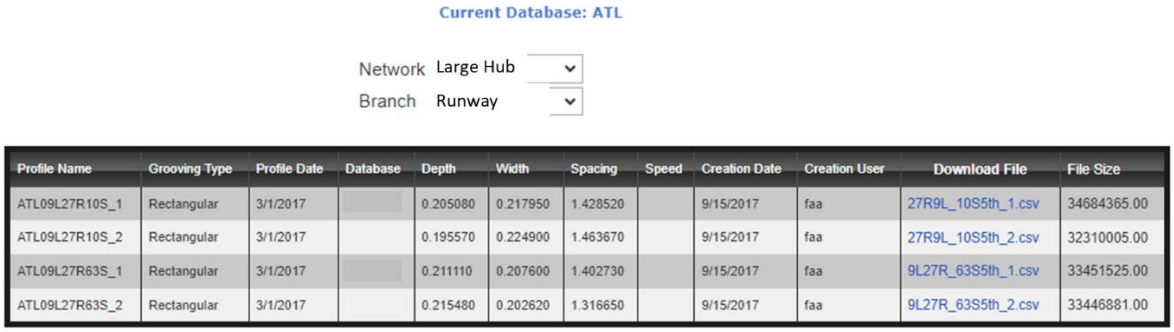
Figure 3. Example of computed roughness indexes in PA40

3.3.3 Surface skid resistance

The skid resistance of runway pavement deteriorates over time due to various factors such as wear and tear from aircraft tire loads, the accumulation of rubber deposits, jet fuel, and oil slippage. These factors can significantly alter or completely cover the pavement surface texture, leading to a reduction in aircraft braking capability and directional control, especially under wet runway conditions.

Two major performance measures associated with surface skid resistance are groove shape measurement for selected runway locations and friction parameters obtained from continuous friction measuring equipment (CFME). PA40 includes only groove shape measurement data. The FAA collects pavement groove data using a laser displacement sensor affixed to the FAA Inertial Profiler (FIP) vehicle equipped with a fifth wheel distance measuring device. The FIP travels at a constant speed of 30 mph. Groove data are collected along the runway length at typically two to three offsets from the runway centerline, primarily covering the keel area, as the groove shape outside the loaded area is typically unaffected. The groove geometry is computed using the FAA software program ProGroove (FAA, 2016b) and PA40 computes the average groove depth, width, and spacing across the measured pavement length. Figure 4 illustrates a PA40 sample table output featuring groove geometry.

As outlined in AC 150-5320-12D (FAA, 2016), the groove performance criteria are defined by geometric limits. Specifically, when 40 percent of the grooves in the runway measure equal to or less than 1/8 inch (3 mm) in depth and/or width for a distance of 1,500 feet (457 m), the effectiveness of the grooves in preventing hydroplaning is significantly reduced. In such cases, the airport operator is advised to promptly implement corrective action to reinstate a groove depth and/or width of ¼ inch (6 mm).



Profile Name	Grooving Type	Profile Date	Database	Depth	Width	Spacing	Speed	Creation Date	Creation User	Download File	File Size
ATL09L27R10S_1	Rectangular	3/1/2017		0.205080	0.217950	1.428520		9/15/2017	faa	27R9L_10S5th_1.csv	34684365.00
ATL09L27R10S_2	Rectangular	3/1/2017		0.195570	0.224900	1.463670		9/15/2017	faa	27R9L_10S5th_2.csv	32310005.00
ATL09L27R63S_1	Rectangular	3/1/2017		0.211110	0.207600	1.402730		9/15/2017	faa	9L27R_63S5th_1.csv	33451525.00
ATL09L27R63S_2	Rectangular	3/1/2017		0.215480	0.202620	1.316650		9/15/2017	faa	9L27R_63S5th_2.csv	33446881.00

Figure 4. Example of computed groove geometry in PA40

3.3.4 Weather data

PA40 obtains weather records for each airport as hourly readings from the National Oceanic and Atmospheric Administration (NOAA) measured by Automated Weather Observing Systems (AWOS) and Automated Surface Observing Systems (ASOS). The weather records include six basic elements: temperature, precipitation, dew point, solar radiation, sky cover, and wind speed. PA40 incorporates a method to compute seventeen weather variables represented as the summation or average of the basic data. The parameters can be calculated for a particular airport (network-level) and for a selected date range.

Details of the weather parameters and their implementation in PA40 can be found in the technical note, “Development of Historical Climate and Weather Data Links Between PA40 and Existing FAA Databases,” which is included in this report as Appendix A.

3.3.5 Traffic data

Threaded track data (TTD) are a product of the MITRE Corporation (Eckstein, Kurcz, & Silva, 2012) that combines data from various sources, including Airport Surface Detection Equipment, Model X (ASDE-X) reports, to create synthetic end-to-end aircraft flight trajectories. PA40 incorporates traffic data obtained from threaded track data (TTD) reports supplied to the FAA by the MITRE Corporation. Additionally, PA40 has the functionality to store runway usage data from other sources, including data provided by the airports. PA40 data includes arrivals and departure operations of each aircraft type on each runway end. The traffic data are searchable based on the specified date range and runway end.

3.3.6 Work history

PA40 includes information about the history of work conducted on the runway, detailing pavement construction, maintenance, and rehabilitation activities for each Section within each

Branch of each Network. This information includes key details such as the date of the work, the type of work conducted, the thickness of layers applied, and the material types used.

3.3.7 Design traffic

Design traffic information is available for nearly half of the runways within the PA40 airports. Design traffic refers to estimates of future traffic used by the design engineer to size the pavement layers and generally deviates from actual traffic (section 3.3.6). The key information includes the aircraft type, the aircraft weight, the anticipated annual departures, and the annual growth.

3.4 Availability of data for pavement performance modeling

Availability of each type of data varies from runway to runway in PA40. The availability of PCI data depends on pavement age and the completeness of each airport's PMS database. Regarding longitudinal profile and groove measurements, most airports either do not regularly perform these measurements or do not adhere to FAA standards. Therefore, the available data were limited to what was collected by the FAA during field visits as described in Section 3.1. The PMS data provided by airports typically include maintenance and rehabilitation activities, although the records may be incomplete or not up to date. The FAA made efforts to augment the PMS data by gathering detailed maintenance and rehabilitation information during field visits and through discussions with airport managers. The TTD-based runway usage data provided by MITRE Corp. (section 3.3.6) covered all runways from 2014 to 2022, with additional data provided by specific airports. Of the data sets included in PA40, the most comprehensive and versatile is the weather data; from the 1960s onwards it is available for all airports. Table 3 and Table 4 summarize general information about the 27 runways in PA40, including initial construction, pavement type, numbers of sections, and major rehabilitation events. Table 5 and Table 6 summarize field-collected and traffic data for the runways, including the number of field visits (if any), the number of surveys performed of each data type, and the availability of traffic data.

Each airport adopts its own method to divide runways into sections for pavement management purposes. Therefore, the sectioning among the PA40 runways lacks a standardized approach. With the exceptions of Runway 10-28 at Baltimore/Washington International Thurgood Marshall Airport (BWI) and Runway 4-22 at LaGuardia Airport (LGA), runway sectioning plans distinguish between the keel (trafficked) area and the outside the keel (non-trafficked) area. Keel sections typically represent the center 50-60 feet width of a runway, divided equally on the left and right sides of the centerline. Outside keel sections are situated on the left and right sides of

the keel sections where aircraft rarely travel, except at taxiway intersections. The number of sections along the runway length also varies among the PA40 runways. For instance, Runway 9L-27R at Hartsfield-Jackson Atlanta International Airport (ATL) divides the keel area into 16 sections, while Runway 18L-36R at Dallas Fort Worth International Airport (DFW) has only two sections along its keel area.

Table 3. Information for 14 runways with flexible pavement in PA40

Airport	Network	Runway	Construction date	Surface type	No. of keel sections	No. of outkeel sections	No. of major rehabilitation events
Boston Logan International Airport	BOS	4L-22R	1947	AC	14	23	5
Baltimore/Washington International Airport	BWI	10-28	1950	AC	9	0	4
John Glenn Columbus International	CMH	10L-28R	1997	AC	3	6	1
		10R-28L	2013	AC	10	20	0
Ronald Reagan Washington National Airport	DCA	1-19	2011	AC	4	8	0
Piedmont Triad International Airport	GSO	5L-23R	2010	AC	3	6	0
LaGuardia Airport	LGA	4-22	1994	AC	18	36	3
Kansas City International Airport	MCI	9-27	1968	AC	9	18	3
Miami International Airport	MIA	12-30	1985	AC	8	16	1
San Francisco International Airport	SFO	10L-28R	1959	AC	7	14	>2
		10R-28L	1954	AC	4	8	>2
Salt Lake City International Airport	SLC	16L-34R	1977	AC	3	6	4
Tucson International Airport	TUS	11L-29R	1943	AC	4	8	4
		3-21	1970	AC	5	10	5-6

Table 4. Information for 13 runways with rigid pavement in PA40

Airport	Network	Runway	Construction date	Surface type	No. of keel sections	No. of outkeel sections	No. of major rehabilitation events
Hartsfield-Jackson Atlanta International Airport	ATL	9L-27R	1974	PCC	16	32	0
Denver International Airport	DEN	17L-35R	1992	PCC	5	10	0
Dallas Fort Worth International Airport	DFW	18L-36R	1975	PCC	2	4	0
Fort Lauderdale International Airport	FLL	10R-28L	2014	PCC	3	6	0
Dulles International Airport	IAD	1R-19L	1959	PCC	3	6	1
		1C-19C	1959	PCC	4	8	1
George Bush Intercontinental Airport	IAH	9-27	1998	PCC	3	6	0
Indianapolis International Airport	IND	5R-23L	1989	PCC	3	6	0
Los Angeles International Airport	LAX	6R-24L	1987	PCC	4	8	0
Chicago O'Hare International Airport	ORD	10C-28C	2013	PCC	12	24	0
Seattle-Tacoma International Airport	SEA	16R-34L	2008	PCC	7	14	0
		16C-34C	1969	PCC	8	16	1
Salt Lake City International Airport	SLC	16R-34L	1995	PCC	3	6	0

Table 5. Data collected at 14 runways with flexible pavement in PA40

Network	Runway	FAA field visits	No. of PCI survey	No. of profile measurements	No. of groove measurements	No. of HWD measurements	Traffic data availability
BOS	4L-22R	1	7	NA	1	1	Jun 2014 - Oct 2020
BWI	10-28	2	8	2	2	2	Jan 2011 - Oct 2020
CMH	10L-28R	2	7	2	2	2	Jun 2014 - Oct 2020
	10R-28L	2	1	2	1	NA	Jun 2014 - Oct 2020
DCA	1-19	0	3	NA	NA	NA	Jun 2014 - Oct 2020
GSO	5L-23R	1	1	1	1	1	Jun 2014 - Oct 2020
LGA	4-22	0	8	1	NA	NA	Jun 2014 - Oct 2020
MCI	9-27	1	6	1	1	1	Jun 2014 - Oct 2020
MIA	12-30	2	2	2	1	2	Jun 2014 - Oct 2020
SFO	10L-28R	0	4	NA	NA	NA	Jan 1999 - Oct 2020
	10R-28L	1	5	1	1	1	Jan 1999 - Oct 2020
SLC	16L-34R	2	1	2	2	2	Jun 2014 - Oct 2020
TUS	11L-29R	1	6	1	1	1	Jun 2014 - Oct 2020
	3-21	0	4	NA	NA	NA	Jun 2014 - Oct 2020

Table 6. Data collected at 13 runways with rigid pavement in PA40

Network	Runway	FAA field visits	No. of PCI surveys	No. of profile measurements	No. of groove measurements	No. of HWD measurements	Traffic data availability
ATL	9L-27R	1	8	1	1	1	Jun 2014 - Oct 2020
DEN	17L-35R	1	2	1	NA	NA	Jun 2014 - Oct 2020
DFW	18L-36R	1	2	1	1	1	Jun 2014 - Oct 2020
FLL	10R-28L	1	1	1	1	2	Jun 2014 - Oct 2020
IAD	1R-19L	1	8	1	1	1	Jun 2014 - Oct 2020
	1C-19C	0	5	NA	NA	NA	Jun 2014 - Oct 2020
IAH	9-27	0	1	NA	NA	1	Jun 2014 - Oct 2020
IND	5R-23L	1	4	1	1	1	Jan 2010 - Oct 2020
LAX	6R-24L	0	1	NA	NA	NA	Apr 2010 - Oct 2020
ORD	10C-28C	2	2	2	2	2	Jun 2014 - Oct 2020
SEA	16R-34L	1	4	1	1	1	Jun 2014 - Oct 2020
	16C-34C	0	8	NA	NA	NA	Jun 2014 - Oct 2020
SLC	16R-34L	2	5	2	2	2	Jun 2014 - Oct 2020

4 Identifying features influencing serviceability

A fundamental step in constructing data-driven prediction models involves identifying the targets (values or indexes) the models seek to predict and the essential features (also referred to as predictors, explanatory variables, or independent variables) affecting these targets. The objective is to discern correlations among the features and to understand the relationships between individual features, or groups of features, and the targets. Features must be carefully selected to effectively represent causal effects on the targets. Models that incorporate insufficient features may lack robustness. Conversely, in models with an abundance of specified features, there is a high risk of overfitting, especially when the set of incorporated features is extensive compared to the number of samples available.

In the context of pavement performance models, the targets encompass sets of pavement condition indexes affecting serviceability of an airfield pavement for safe operation. These indexes are influenced by a variety of features, including environmental conditions, traffic loads, pavement material aging, and routine preventive maintenance, either independently or in combination.

This section identifies the pavement condition indexes used in developing the SL model and the key features influencing these indexes. All data are obtained from airports within the PA40 database. To implement reliable models, it is essential to use high-quality, relevant data for training that are free from noise and redundancy. A comprehensive review of the PA40 data is conducted to meet this goal, identify outliers, and address missing data. The pavement performance data are then processed, and the condition indexes calculated.

4.1 Model targets: Pavement Condition Indexes

The most significant pavement performance indexes that are predictive of runway serviceability are those related to surface distress, loss of smoothness (roughness), and groove deterioration.

4.1.1 Surface distress—PCI, SCI, and anti-SCI

Researchers evaluated PCI, SCI, and anti-SCI data from the fourteen flexible pavement and thirteen rigid pavement runways in the PA40 database listed in Table 3 and Table 4. These indexes are computed in PA40 for all sections within these runways, with each section having at least two and up to eight inspections. The number of PCI samples are not uniform among the runways, i.e., the number of sections per runway, the pavement age, and the frequency of inspections may vary. For example, Runway 4-22 at LGA has 115 samples, but Runway 5L-23R

at Piedmont Triad International Airport (GSO) has only 3 samples. Figure 5 and Figure 6 display the number of PCI samples in both keel and outkeel areas of runways with flexible and rigid pavements, respectively.

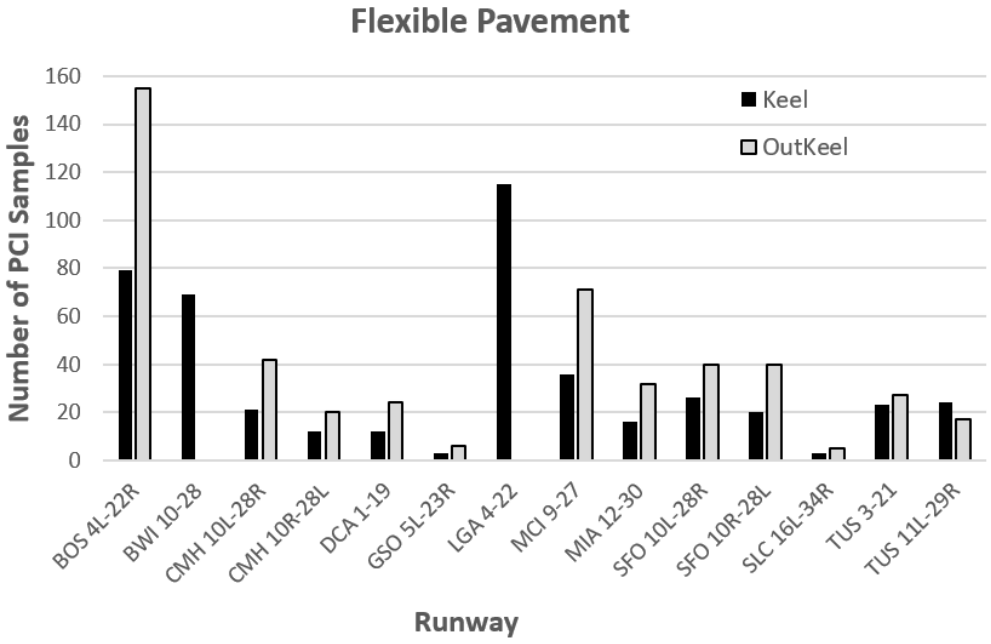


Figure 5. Number of PCI samples in runways with flexible pavement

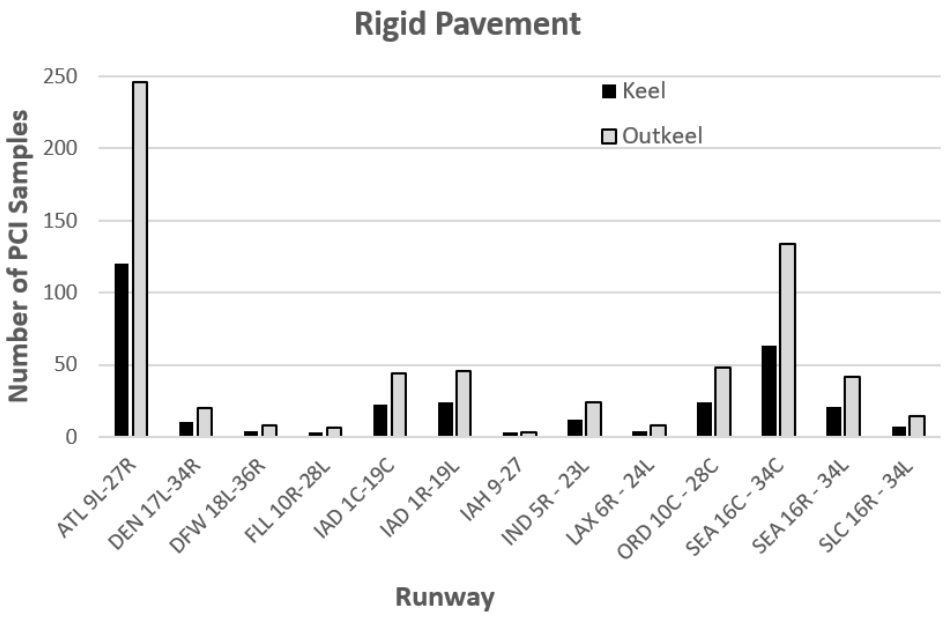


Figure 6. Number of PCI samples in runways with rigid pavement

Figure 7 and Figure 8 illustrate the range and distribution of values for each index for the flexible and rigid runways. For both flexible and rigid pavements, PCI and anti-SCI exhibit a wide range of values, although the majority are above 60. This is likely because airports generally maintain runways above a PCI range of 50 to 60. SCI values for flexible pavements are typically above 90, indicating that in flexible airport pavements, load-related distresses do not commonly occur before reaching the design life or before rehabilitation is warranted due to functional failure. For rigid pavements, SCI values are typically above 80, aligning with the FAA threshold for structurally failed pavement.

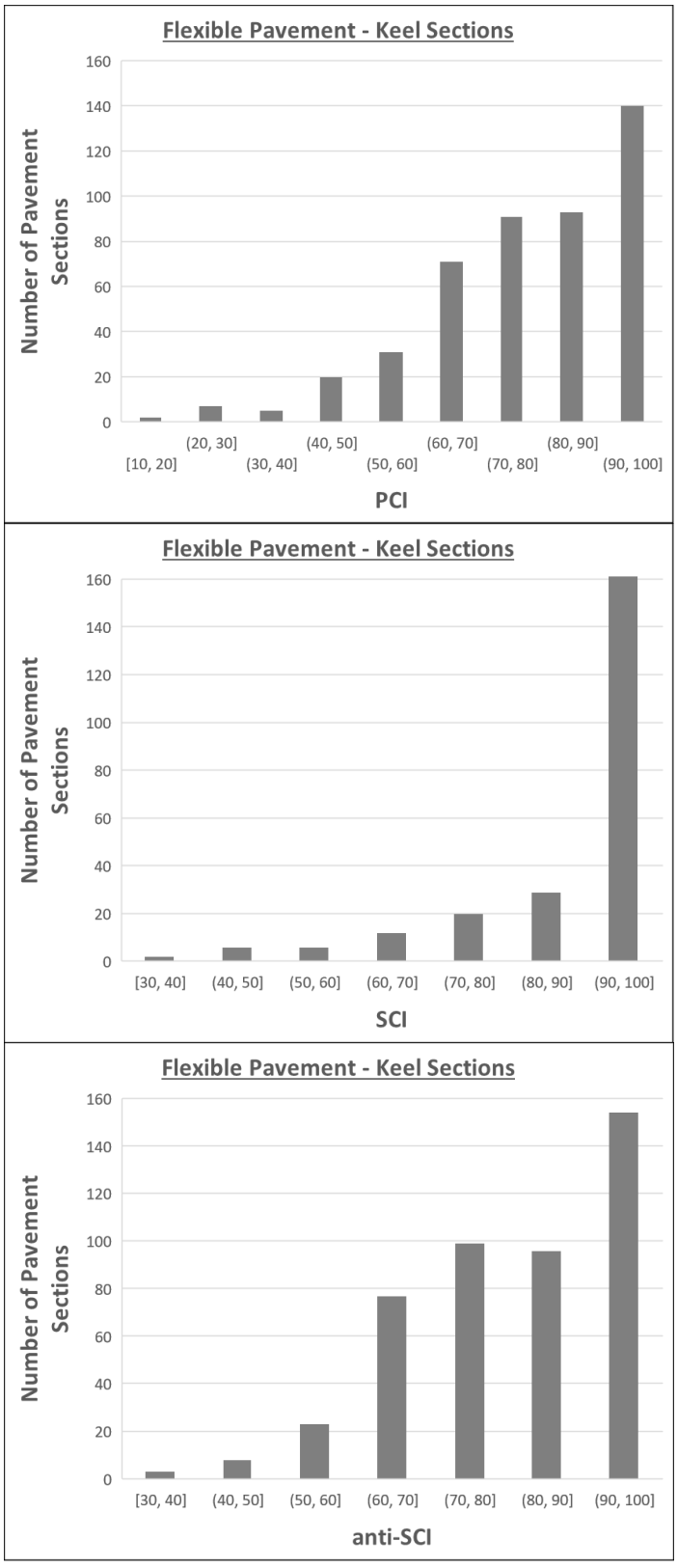


Figure 7. Range of flexible pavement indexes (PCI, SCI, anti-SCI) in PA40

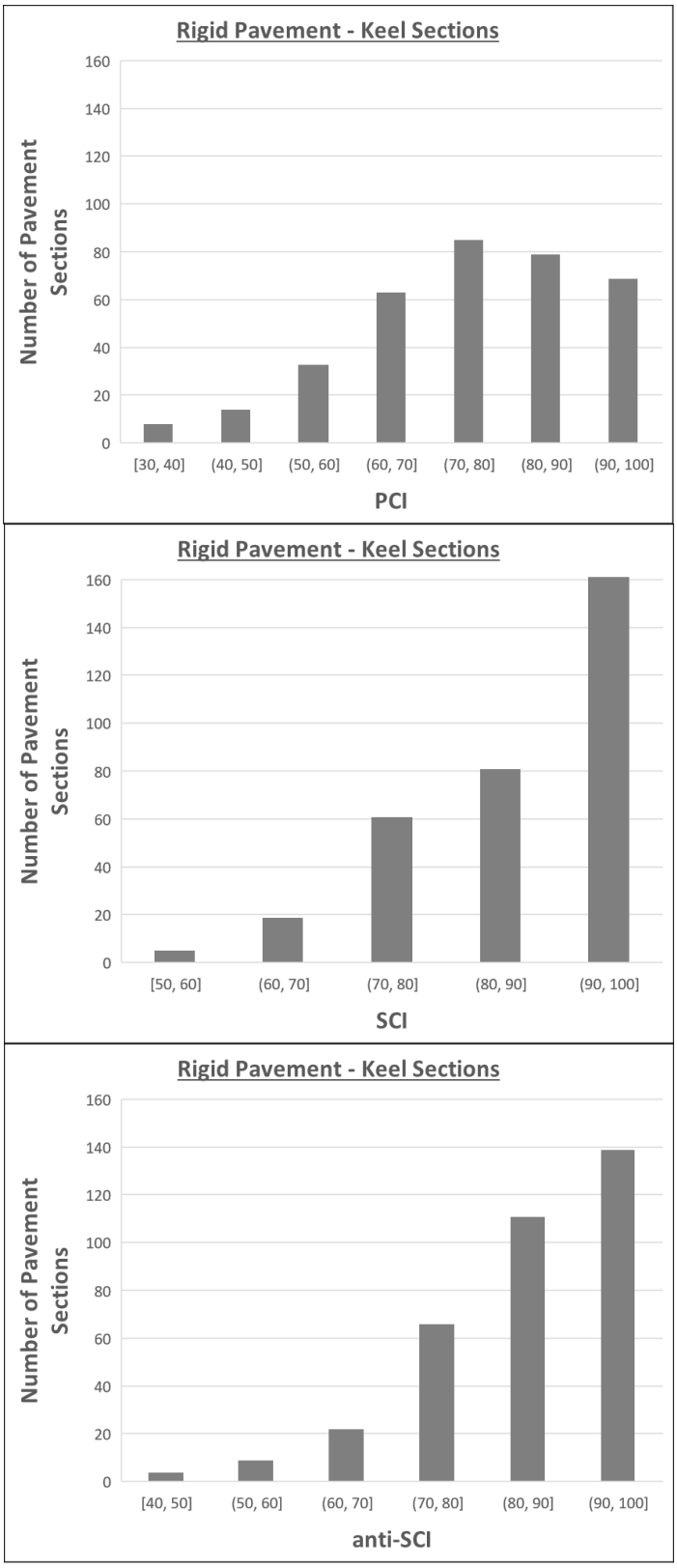


Figure 8. Range of rigid pavement indexes (PCI, SCI, anti-SCI) in PA40

4.1.2 Loss of smoothness (roughness) - RRI

The FAA has collected longitudinal profile data using a vehicle-mounted inertial profilometer and a SurPro walking profiler at some of the airports within the PA40 database (as noted in Table 5 and Table 6. Since PA40 includes computed average RRI values over the entire length of the runway and does not provide RRI per section, researchers calculated average RRI values for each runway section to correlate the RRI values with the PCI indexes. It was determined that data from the Inertial Profiler are not reliable for calculating the roughness indexes. Therefore, only data from the SurPro are used for RRI calculation.

SurPro data are available for 10 flexible pavement runways, where 5 were visited twice and five were visited once. SurPro data are available for nine rigid pavement runways, where two were visited twice and 7 were visited once. This means that no more than two data points are available over the pavement life of a given runway, posing a challenge for investigating roughness progression over time. Figure 9 and Figure 10 illustrate the range and distribution of RRI values for the flexible and rigid runways. A lower RRI represents smoother surface. For the majority of runway sections, the RRI is below 0.34 (value marked in the figures), the threshold corresponding to 50% of pilots rating the runway as unacceptable.

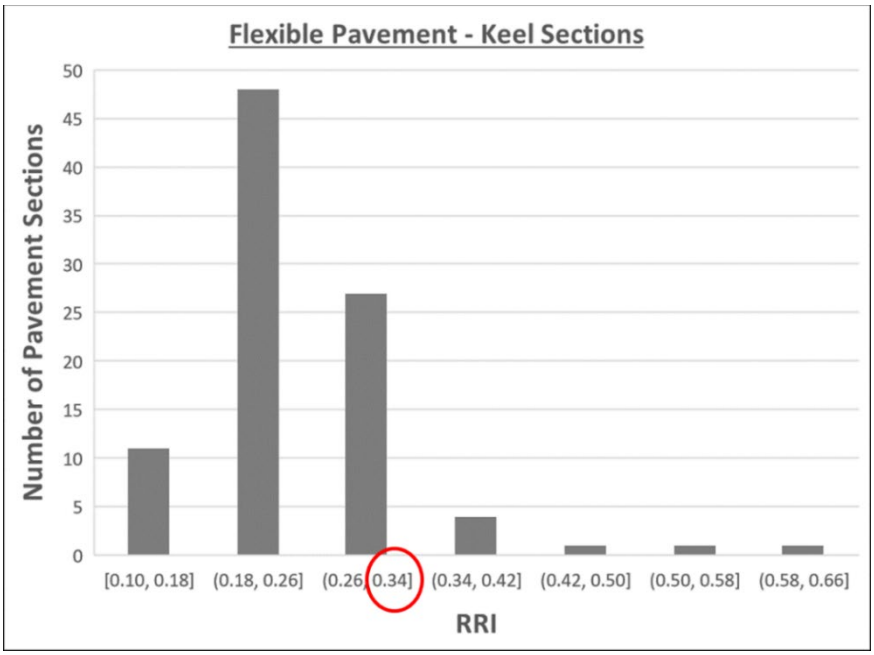


Figure 9. Range of flexible pavement RRI in PA40 sections

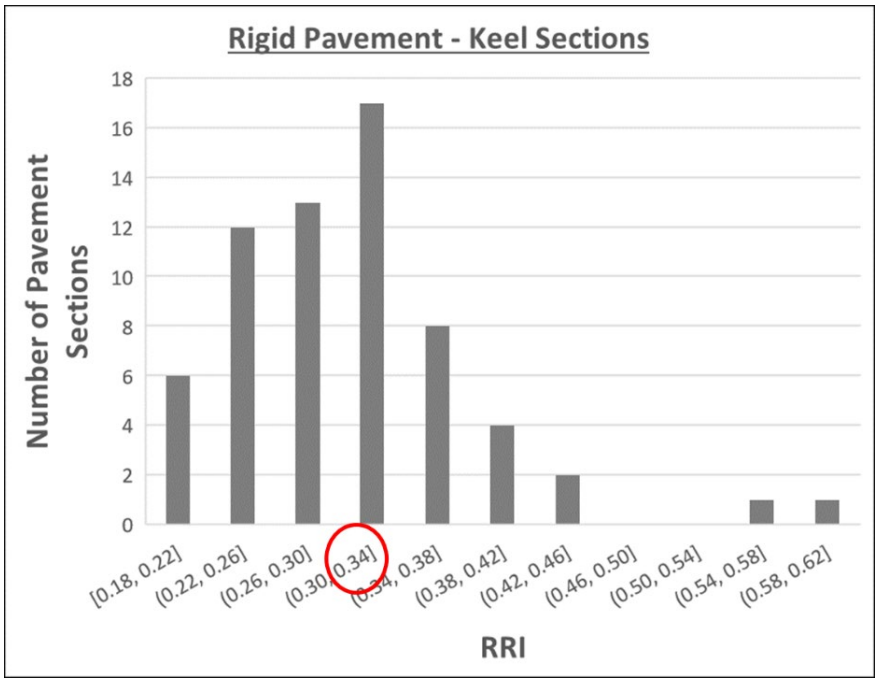


Figure 10. Range of rigid pavement RRI in PA40 sections

4.1.3 Groove deterioration - Groove Index

The FAA has collected pavement grooving data for 10 flexible pavement runways and 8 rigid pavement runways within the PA40 database. Most runways have only one measurement, with some including two measurements.

Similar to RRI, PA40 provides computed groove geometry over the entire runway length, not per section, so researchers processed the groove data for each runway section to correlate the data to the other indexes. Using the ProGroove program, the data are filtered once with a decimation filter based on the vehicle speed and the sample rate of the distance measuring unit, resulting in a final groove sample rate of 0.003363 feet. A GNU Octave script is then used on ProGroove's comma-separated value (CSV) output files to split the data based on the PCI sections. The average groove depth and width per section is used to calculate the GI, defined as the percentage of grooves with depth or width below 0.125 inch. A lower GI denotes better surface traction.

Figure 11 and Figure 12 illustrate the range and distribution of GI values for the flexible and rigid runways. For most runway sections, the GI is below 40% (value marked in the figures), the FAA threshold mandating immediate corrective action.

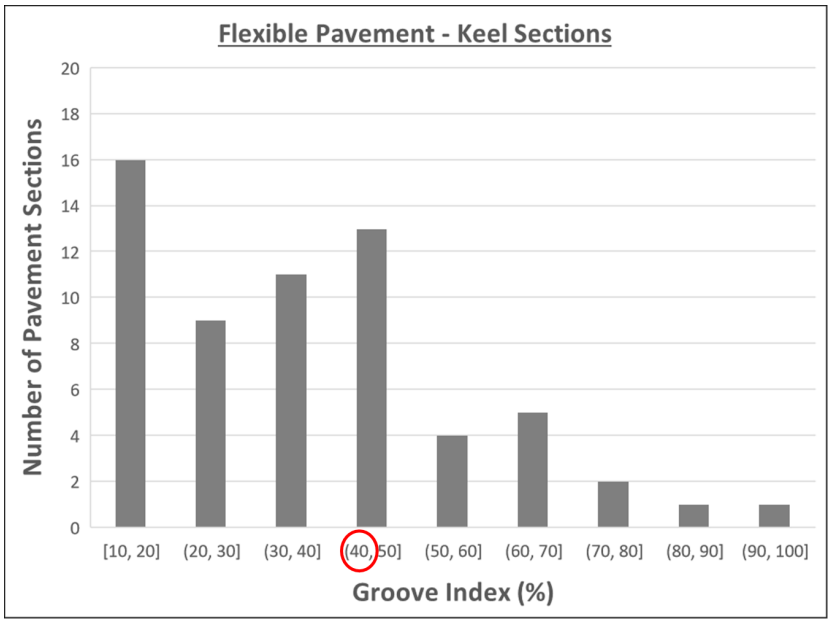


Figure 11. Range of flexible pavement GI in PA40 sections

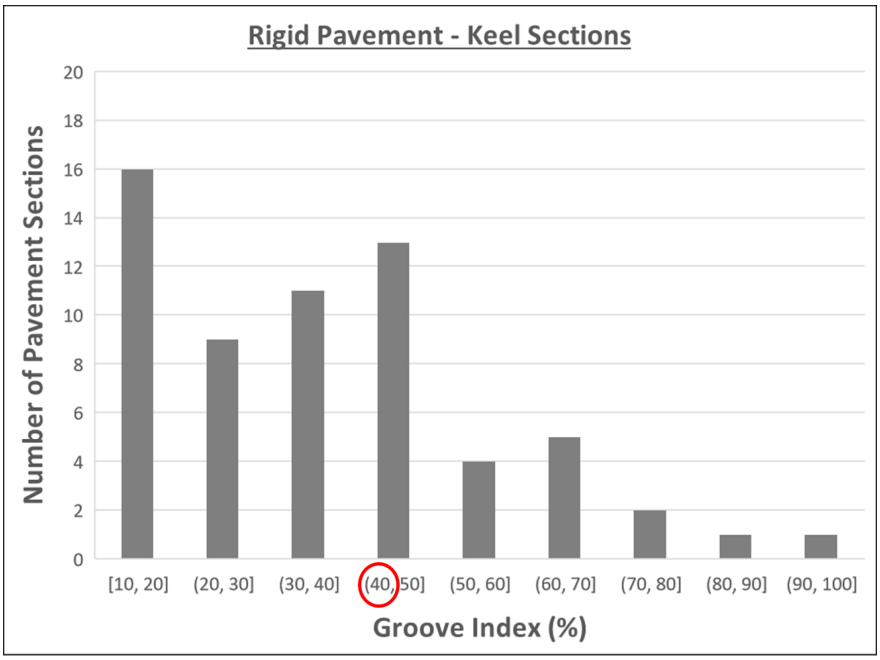


Figure 12. Range of rigid pavement GI in PA40 sections

4.2 Model inputs: Features influencing pavement performance

The features are the independent variables used as inputs to the ML models to predict the pavement performance indexes and the SL index. This section identifies features that have the potential to affect the variation in pavement performance.

4.2.1 Pavement functional age

Pavement age has historically served as the primary predictor in many pavement performance models, assuming it acts as a surrogate for other variables such as runway usage and weather events. For each runway pavement section, pavement age is defined as functional age, i.e., the time since construction or the last major rehabilitation.

It is assumed that a major rehabilitation restores the pavement to "new" functional condition. As such, the PCI is reset to 100 with a functional age of 0 years.

4.2.2 Number of rehabilitation events

The construction dates of the runways within the PA40 database vary widely. Some runways are relatively new or recently reconstructed, while others have been in service for decades. Some airports have implemented corrective rehabilitation at regular intervals over the years to maintain the runways in service. For example, the flexible pavement on Runway 4L-22R at BOS has

undergone five major rehabilitation projects since its original construction in 1947, and the runway is still in service. In this case, the airport managers likely determined that the underlying structure is still intact after nearly 75 years, and the pavement has remaining life. The rehabilitation efforts involved mill and overlay of the asphalt surface (up to 6 inches) roughly every 15 years. Each time the pavement receives an asphalt overlay, the pavement's functional condition is restored, and the airport's APMS treats it as effectively a brand-new pavement. However, the remaining life of the restored pavement may not be similar to that of a newly constructed pavement as structural life or fatigue life may have diminished over the years. Pavements that have undergone multiple rehabilitation over their life (i.e., having higher structural life) are expected to exhibit structural-related distresses at a higher rate than a relatively new pavement with the same functional age. The number of major rehabilitations for each pavement section can serve as a proxy for structural life.

For flexible pavements, surface reconstruction or any overlay equal to or thicker than 2 inches is considered a major rehabilitation. For rigid pavements, the replacement of more than 20% of slabs is considered a major rehabilitation. Maintenance measures, such as crack sealing in flexible pavement or joint sealing in rigid pavement, are not classified as major rehabilitation. Isolated patching (in flexible pavement) or slab replacement (less than 20%), while affecting surface condition, is not considered a major rehabilitation.

4.2.3 Weather variables

When examining the effect of environmental factors on pavement performance, it is important to understand whether pavement deterioration results from general climate conditions, specific weather events, or a combination of both. Although climate and weather are interconnected, they pertain to distinct aspects of environmental and atmospheric conditions. Weather refers to the short-term environmental conditions at a specific location and time, while climate represents the long-term average of weather variations in a particular region. Weather is expressed in temporal scales such as hourly, daily, monthly, seasonally, or even yearly events, while climate is the typical condition over a long span (e.g., 30-50 years).

It is also important to note that the selection of airports within the PA40 database was not random across states. Consequently, the airport locations do not encompass all geographic and climatic regions, and the data may not represent a complete range of potential weather events.

It is crucial to differentiate between the overall pavement condition (e.g., as expressed by PCI), and specific distress types when assessing the effect of climate or weather. Certain distress types that manifest in specific climate conditions may not be as common in other climates. For

instance, asphalt thermal transverse cracking is more common in freezing climates, while block cracking is likely to occur in areas with high solar radiation and high evaporation rate. However, when these distress types contribute to PCI calculations, they may result in the same overall pavement condition deterioration. This implies that changes in the overall pavement condition index value may not be significantly influenced by the climate region. At the same time, weather events have the potential to affect specific distress types. A region that is in a freezing climate may undergo a varying number of freezing degree days (FDD) events in consecutive years, implying that in some years, pavements may be exposed to much more severe weather conditions than in others. While extreme events have the potential to accelerate pavement deterioration, they may not have an immediate impact on overall pavement deterioration or specific distress types.

This study focuses on predicting pavement condition indexes rather than individual distress types. The aim is to establish correlations between pavement condition indexes and weather events. As discussed in Section 3.3.5, PA40 computes up to 17 different weather variables for each airport within a specified date range. Given the current availability of pavement performance data, 10 airports with flexible pavement runway and 10 airports with rigid pavement runway, there are more weather variables than the number of airport locations for each pavement category. This may pose challenges in assessing the environmental impact. An FAA study (Ashtiani, 2021) demonstrated that some of these weather variables exhibit moderate to high correlations, indicating that they often represent different manifestations of the same principal environmental parameter. The study concluded that highly correlated weather variables can be omitted, as including them adds no extra information, and removing them may enhance prediction performance.

As an initial step, researchers streamlined the weather variables to 12, as listed in Table 7. Seven of the weather variables can be represented as either average or cumulative values, while five can only be expressed as average values. For each PCI data record for a flexible pavement, weather events are characterized in two ways: 1) average annual values of weather variables since the last inspection or, if there is no prior inspection, since the last rehabilitation/reconstruction, and 2) cumulative values of weather variables since the last rehabilitation/reconstruction. For each PCI data record for a rigid pavement, weather events are calculated as average annual values of weather variables since the last inspection or since the last rehabilitation/reconstruction if there is no prior inspection. For every RRI and GI data record, weather events are determined as the annual average values of weather variables since the last rehabilitation/reconstruction. Box plots in Figure 13 and Figure 14 show the ranges of average annual values of eight weather variables across all PCI data records for flexible and rigid pavements, respectively. Plots show minimum,

first quartile (25th percentile), median, third quartile (75th percentile), and maximum values of each weather variables for each runway.

Table 7. Weather variables considered for modeling

Average Environmental Variables	**Cumulative Environmental Variables**	**Units**
FDD	FDD	°F-days
FThC	FThC	cycles
Temp90	Temp90	days
DPrec	DPrec	days
TPrec	TPrec	inches
FPD	FPD	days
HydD	HydD	days
Avg Temp		°F
RH		%
Average wind speed (Wind)		mph
Thornthwaite index		%
Sky cover		oktas

Avg Temp = Average daily temperature
DPrec = Days precipitation
FThC = Freeze-thaw cycles
HydD = Hydration days
RH = Relative humidity
Temp90 = Days temperature over 90°F

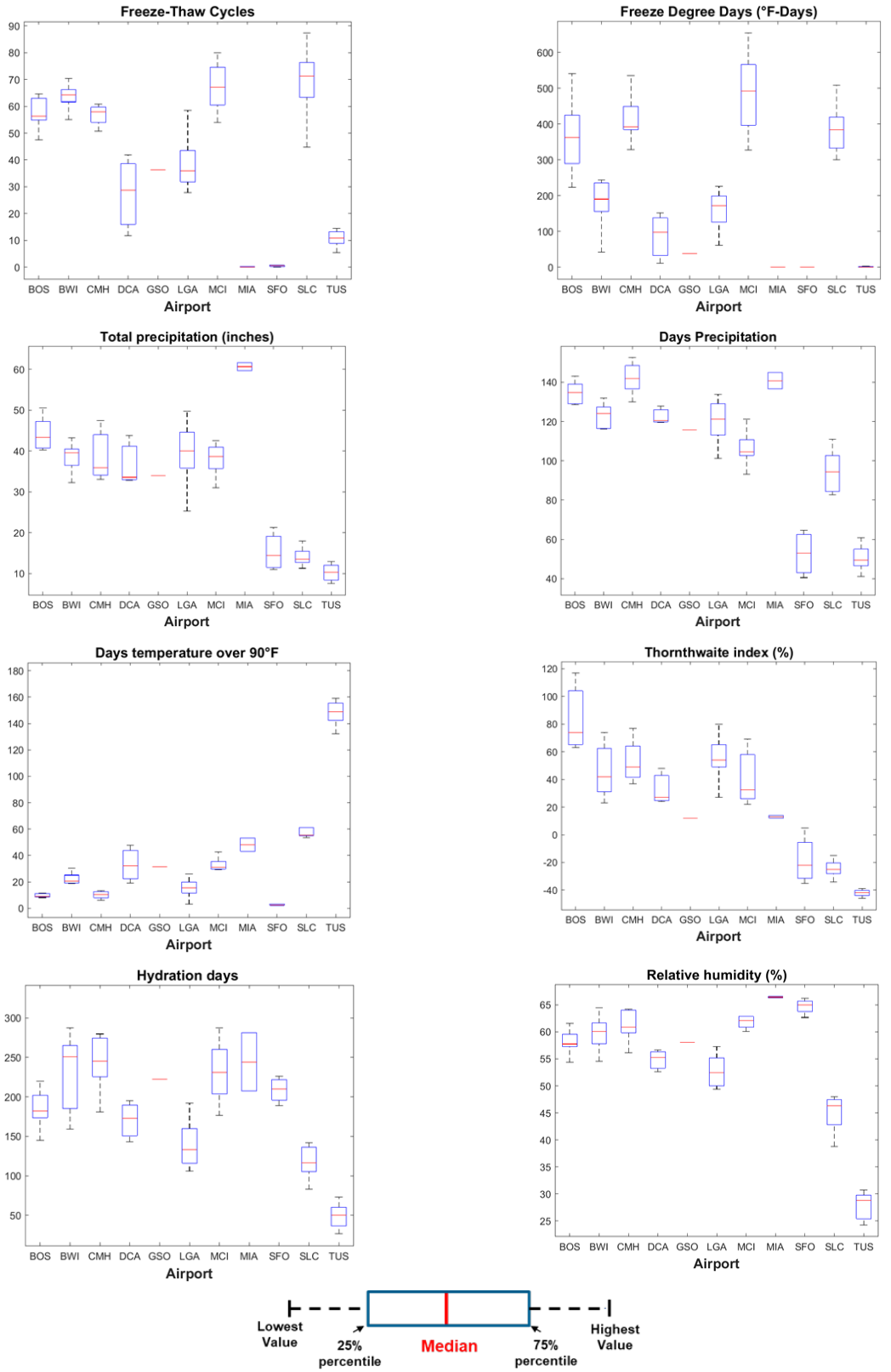


Figure 13. Variation in annual average environmental variables for airports with flexible pavements during pavement condition index inspection period

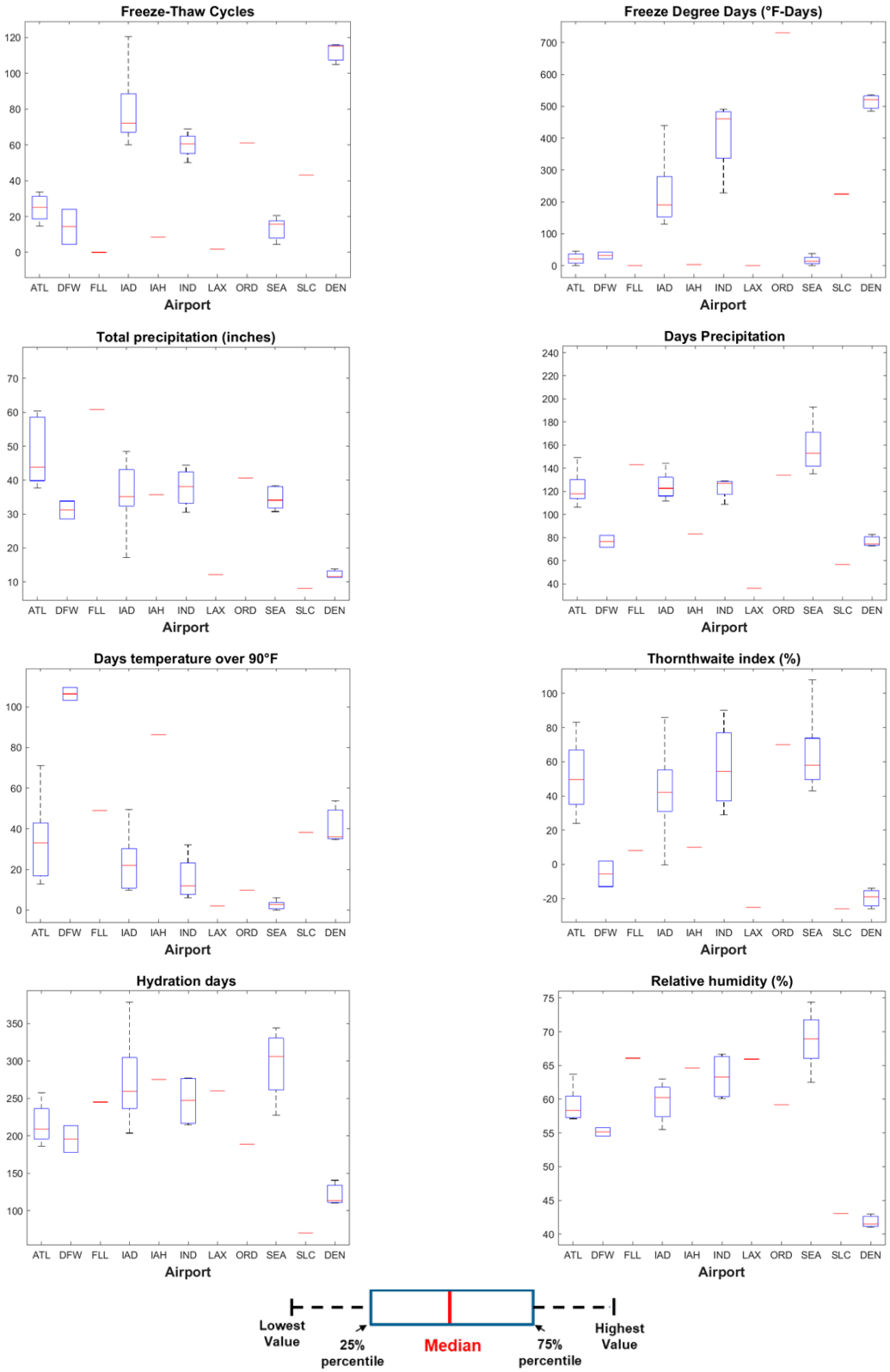


Figure 14. Variation in annual average environmental variables for airports with rigid pavements during PCI inspection period

4.2.4 Traffic variables

The assessment of the effect of traffic on overall pavement performance or specific distresses poses challenges for several reasons:

1. **Uncertainty in traffic forecast**. Runway pavements are subject to diverse and sometimes unforeseen loading patterns, including variations of aircraft weight and gear configurations, throughout their lifespan. Fleet mixes may also change due to changes in operational plans or environmental constraints. Moreover, the limited data available for most runways in PA40 (only 5 to 6 full years) makes it challenging to reasonably determine a growth rate for traffic volume or forecast traffic patterns.

2. **Conservative pavement design**. The PA40 data indicates that load-related distresses are relatively uncommon in flexible airport pavements. Conservative FAA pavement design methods contribute to runways sustaining applied traffic loads for longer than their designed life. Additionally, forecasted traffic considered in the design tends to be highly conservative, typically exceeding the actual traffic. It is crucial to account for the design traffic when assessing the effect of traffic on performance, but this information may not be readily available.

3. **Relevance of traffic data**. For most runways studied, PCI survey data are largely for years prior to 2014, with no available traffic data. This makes it challenging to directly correlate the traffic with changes in pavement conditions.

4. **Lack of standard traffic index**. Traffic information in PA40 is presented as the number of departures and arrivals of aircraft types using a runway end. Incorporating hundreds of different aircraft types into a usable model input is a challenge. Using traffic counts alone is not explanatory, as it disregards under-design or overdesign of pavement structure. Alternative approaches, such as classifying aircraft into categories based on maximum takeoff weights, face challenges in arbitrary weight range determination and unclear actual aircraft weight data. Another possible approach is to use detailed aircraft information and incorporate FAA pavement design procedures to determine traffic induced damage. However, this approach is primarily suitable for deterministic analysis of pavements under traffic loads and is not practical for integration into data-driven pavement performance prediction models.

Given these challenges, it is essential to identify simple and comprehensive parameters to reasonably standardize traffic data characteristics. This study employed a method similar to the "design aircraft" concept from cancelled Advisor Circular 150/5320-6D (FAA, 1995). Aircraft

are classified based on standard generic gear type categories S-75, D-200, 2D-400, B777-300, B747-400, and A380-800, providing a consistent way to compare design with actual traffic. In this approach, the actual aircraft and the design aircraft are converted to equivalent aircraft traffic without reference to the pavement structure. The steps to calculate are following:

1. For each runway, identify the aircraft type and the actual traffic (departure and arrivals separately) during period *t* when traffic data is available. Due to variations in data sources, there is inconsistency in data presentation, but the aircraft codes in the PA40 data are typically based on ICAO's designations. Convert these codes to common names following FAARFIELD conventions. Since there is a large variety of aircraft types using each runway, exclude small aircraft with less than 10 operations for simplicity.
2. For each aircraft type, assign a reasonable departure weight, possibly using the FAARFIELD library.
3. For each aircraft type, assign a generic aircraft type (S-75, D-200, 2D-400, B777-300, B747-400, or A380-800) based on gear configuration or aircraft type.
4. For each aircraft type, calculate equivalent departures of the generic aircraft using equation below based on Advisory Circular 150/5320-6D:

$$logR_1 = logR_2 \times \left(\frac{W_2}{W_1}\right)^{1/2} \quad (2)$$

   where,

   R_1 = equivalent departures of the generic aircraft

   R_2 = departures of actual or design aircraft

   W_1 = wheel load of the generic aircraft

   W_2 = wheel load of the design or actual aircraft
5. Sum up the equivalent departures of each generic aircraft during time t.
6. Perform similar conversions for the design traffic. For each gear type, multiply the equivalent annual design traffic by time *t* in years to obtain the equivalent design traffic.

Table 8 shows the assumed aircraft weight for each generic group.

Table 8. Weight assigned to generic aircraft groups

Generic aircraft group	Weight (lb)
S-75	75,000
D-200	200,000
2D-400	400,000
A380-800	1,238,998
B747-400	890,000
B777-300	777,000

Considering both the actual and design traffic represented on the same standard scale, the research team defined a new index called XSDepartures. This index is equal to the numerical difference between the actual equivalent traffic and the design equivalent traffic for a given generic aircraft. Unlike a ratio-based index, XSDepartures can be used to determine whether the runway is overused or underused. A positive XSDepartures index may correlate with pavement structural damage, even if the actual departures are not large. Conversely, a negative XSDepartures index may indicate that the pavement is overdesigned, or the design traffic is conservative. This new index also allows the determination of whether one type of aircraft group (e.g., B777) is significantly more demanding than other groups.

A challenge encountered with the XSDepartures metric is the absence of design traffic for seven flexible runway pavements and four rigid runway pavements. To address this, a reverse-engineering approach was employed to estimate the equivalent design traffic based on the pavement structure at the time of design. Using actual traffic data, the percentage composition of each generic aircraft category in the traffic mix was determined, assuming that the aircraft types observed in the actual data were representative of the design traffic. With these percentages and the known design pavement structure, a series of FAARFIELD "Life" analyses was conducted, incorporating various combinations of annual departures for each generic aircraft type. The resulting aircraft mix that yields a pavement life equal to 20 years was considered as the equivalent design traffic. Employing this approach, the design traffic was determined for the following runways: 11L-29R at Tucson International Airport (TUS), 12-30 at Miami International Airport (MIA), 9L-27R at ATL, and 16C-34C at Seattle-Tacoma International Airport (SEA). However, for runways where design traffic data are unavailable due to lack of information regarding the pavement structure, estimating the design traffic proved unfeasible.

Table 9 and Table 10 show the calculated XSDepartures for flexible and rigid runway pavements, respectively. Results indicate that most runway designs overestimate the heavier aircraft (D-200, 2D-400, B777, B747, and A380) and almost all runways underestimate the

lighter aircraft (S-75). The only exception is Runway 10-28 at BWI which had 772 B747 departures during a 10-year period, although there was no B747 in the design traffic. Runway 5R-23L at Indianapolis International Airport (IND) underestimated 2D aircraft. Runway 16R-34L at SEA has been largely used for arrivals, meaning the number of actual departures is considerably smaller than the number of design departures. Therefore, XSDepartures were calculated for both departures and arrivals for this runway.

Table 9. XSDepartures for flexible runway pavements

Runway	Duration (years)	XSDepartures of Generic Aircraft Groups					
		S-75	D-200	2D-400	B777-300	B747-400	A380-800
BWI 10-28	9.8	90	-104,289	-4,938	-2,531	772	-
CMH 10R-28L	6.4	514	-16,084	-1,315	-	-	-
CMH 10L-28R	6.4	222	-1,235	-31	-	-	-
DCA 1-19	6.4	-	-370,354	-41,843	-	-	-
GSO 05L-23R	6.4	138	-79,361	-23,797	-	-	-
MCI 09-27	6.4	28	-26,495	-5,251	-	-	-
TUS 11L-29R*	6.4	-	-19,177	1,736	-	45	-

* Estimated design traffic

Table 10. XSDepartures for rigid runway pavements

Runway	Duration (years)	XSDepartures of Generic Aircraft Groups					
		S-75	D-200	2D-400	B777-300	B747-400	A380-800
IAD 1C-19C	6.4	180	-70,945	-158,098	-22,686	-11,572	-6,234
IAD 1R-19L	6.4	337	-53,336	-145,989	-22,262	-11,350	-5,723
IND 5R-23L	7.3	271	-34,083	12,556	2,234	-492	0
ORD 10C-28C	6.4	22	-148,861	-740,601	-15,464	-7,132	265
SLC 16R-34L	6.4	18	-85,082	-76,075	~~0~~	~~0~~	~~0~~
SEA 16R-34L	6.4	22	-25,492	-9,053	-463	-1,311	0
ATL 9L-27R*	6.4	-	-4,640	97,733	4,679	-1,052	-4,319
SEA 16C-34C*	6.4	-	7,517	-3,210	38	-17	-949

* Estimated design traffic

Since the design traffic cannot be determined and XSDepartures cannot be calculated for nearly half of the runways, this index is excluded as an input parameter for the machine learning (ML) models. Instead, the actual traffic counts were converted into equivalent operations of appropriate generic aircraft categories. Since pavement distresses are mainly caused by

departures of heavier aircraft, the models for PCI and SCI consider equivalent departures of D-200, 2D-400, and the summation of A380-800, B747-400, and B777-300 (defined as Heavy Aircraft), as traffic inputs. On the other hand, since roughness and grooving may be affected by both departures and arrivals of all aircraft, regardless of their weight, the RRI and Groove Index models consider total departures and arrivals counts as inputs, in addition to the equivalent departures of generic aircraft. Ideally, all these traffic parameters must be calculated between the last rehabilitation or reconstruction and the inspection dates. Since most PCI survey data or rehabilitation dates are for years prior to 2014 when no traffic data is available, it is assumed that the traffic data from the available years is representative of the traffic levels for prior years. Therefore, for each traffic parameter identified above, the average annual values for the years the data are available are used as model inputs. Figure 15 and Figure 16 show the average annual departures of the three generic aircraft groups for runways with flexible and rigid pavements, respectively. Figure 17 and Figure 18 show the average annual departures and arrivals of all aircraft for runways with flexible and rigid pavements, respectively. Table 11 summarizes the traffic parameters used for each pavement condition index.

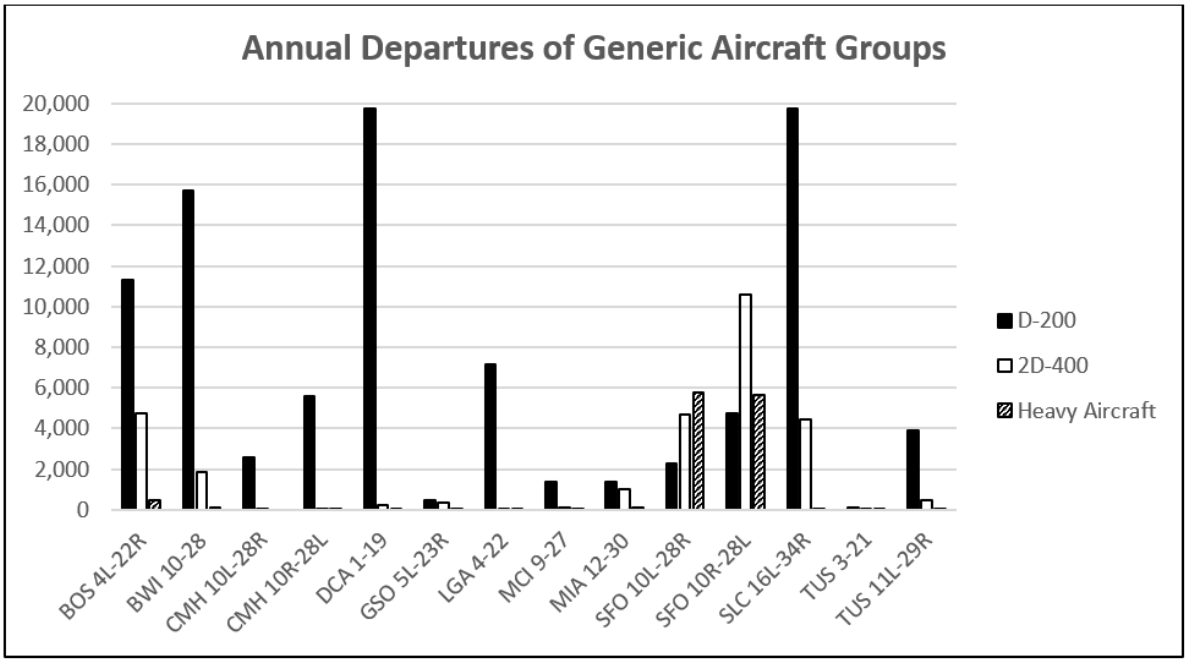


Figure 15. Average annual departures of generic aircraft on runways with flexible pavement

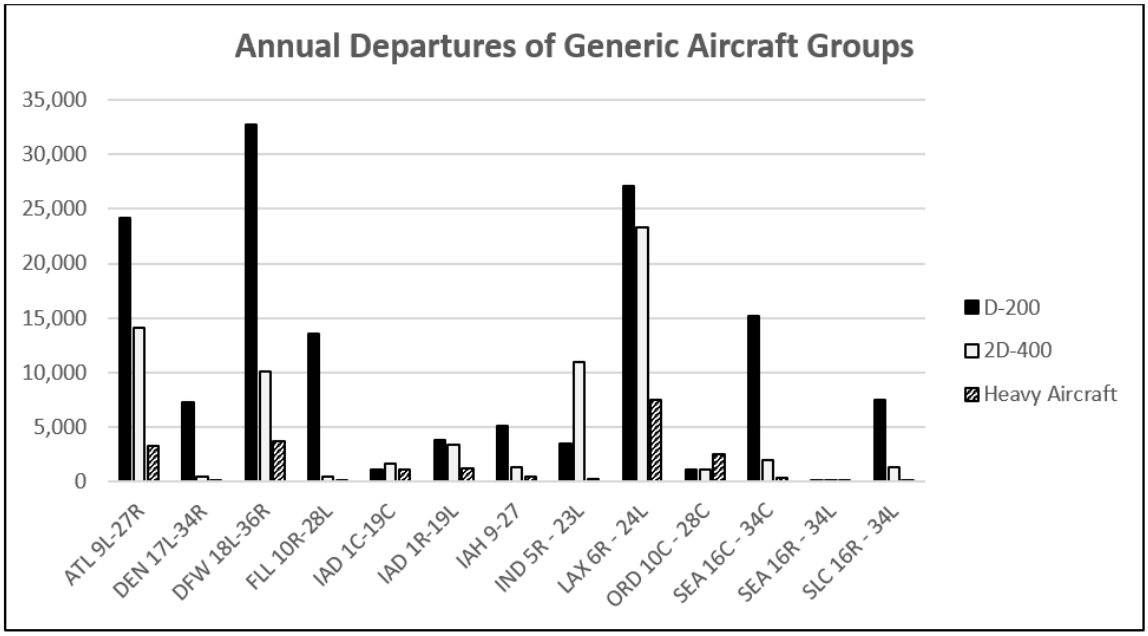


Figure 16. Average annual departures of generic aircraft on runways with rigid pavement

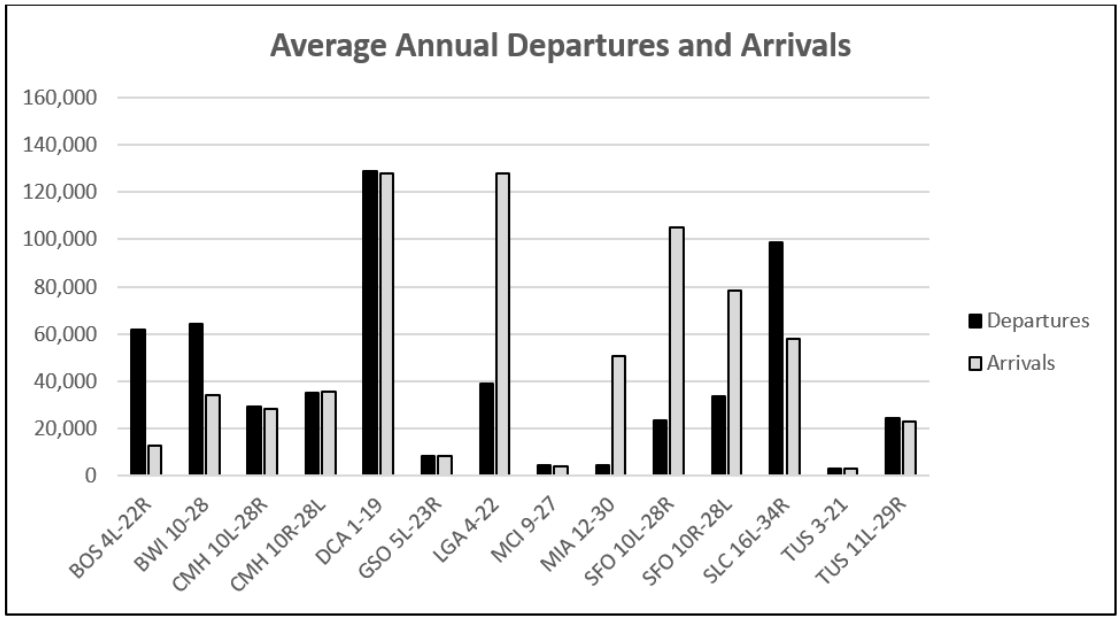


Figure 17. Average annual departures of all aircraft on runways with flexible pavement

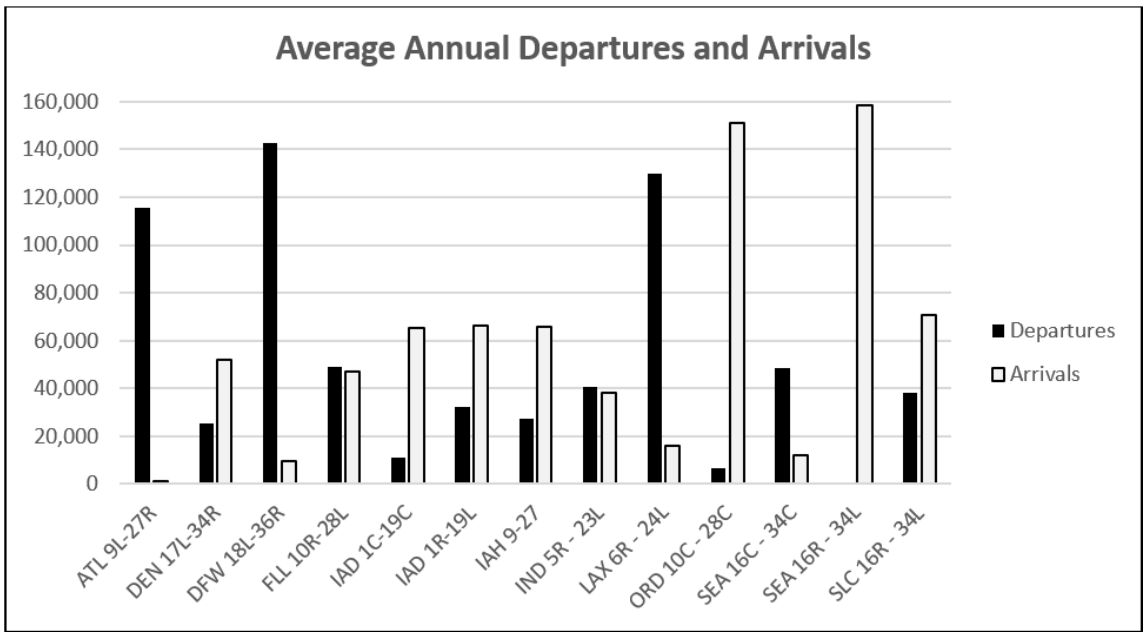


Figure 18. Average annual departures of all aircraft on runways with rigid pavement

Table 11. Traffic index used for pavement condition prediction models

Performance/condition index	**Traffic index**
PCI and SCI	-Average Annual Departures of D-200, 2D-400, and Heavy Aircraft
RRI	-Average Annual Departures -Average Annual Arrivals
Groove Index	-Average Annual Departures -Average Annual Arrivals

5 Development of pavement performance prediction models

This section presents the results of the ML models developed for predicting critical pavement condition indexes within the SL model. These models are used to estimate missing condition indexes necessary to construct the training database for the SL model. Additionally, the models for each individual index provide valuable insights for assessing long-term pavement performance and can serve as stand-alone tools.

5.1 Machine learning methods

ML operates within two fundamental paradigms: supervised and unsupervised learning. In supervised learning, the model is trained on a labeled dataset, where each input data point is paired with a corresponding output target. The objective is to make predictions or classifications based on the mapping functions derived from the input and output target labeled training set. In unsupervised learning, the model tries to learn patterns and structure from the data without explicit guidance on input-output associations. The goal is to discover hidden patterns, cluster data objects, or reduce the dimensionality of the data. Both supervised and unsupervised learning methods are used to construct the prediction models.

5.2 Machine learning modeling approaches for pavement performance

Researchers used two approaches to develop ML prediction models for pavement performance indexes:

1. Continuous Function Approximation. In this approach, the condition index at time *t* is estimated based on a set of features at the same time *t*.
2. Time-Series or Autoregressive Modeling. This modeling approach is used when there is time dependency between data, indicating relationships between lagged observations over

time. Time-series models can be formulated as one-step discrete forecasting, with data points indexed sequentially over time.

The continuous function approach is applied to the RRI and Groove Index prediction models, where "time *t*" denotes the pavement functional age at the time of pavement roughness or groove prediction. Time-series models are not suitable for these indexes, as data are not available for more than one measurement for most of the runways.

The time-series approach is applied to the PCI, SCI, and anti-SCI prediction models. These data are presented as longitudinal time-series, where the same samples (runway sections) are repeatedly measured at different times. In these models, the pavement condition from the previous time step is used, along with other features, to predict the condition at the next time step.

5.3 Machine learning model implementation

The model search triple (MST) approach (Kumar, McCann, Naughton, & Patel, 2016), which is an iterative three-element process, is used to design the ML models. The first element of MST involves feature selection and engineering, where features are identified, developed, and evaluated for their relevance to predictions. The second element is model and algorithm selection, an iterative process to evaluate ML models that are suitable for the modeling goals. The best candidate ML models are determined based on modeling goals (prediction or classification), data size and type (e.g., time dependent), the nature of underlying physical phenomena, and data sparsity (infrequent, missing, irregular data). The third element is hyper-parameter tuning, the process of determining the range of hyper-parameters for selected ML algorithms. Hyper-parameters are unique parametric configurations for each ML model. Hyper-parameter values are set initially and then fine-tuned iteratively during training to improve model performance. The MST approach is implemented to ensure a systematic and thorough exploration of key influencing variables, linear and non-linear models, and their associated hyper-parameters.

This section describes various feature selection methods used to achieve two primary objectives:

1. Evaluate and identify the features that most significantly influence pavement performance.
2. Reduce the dimensionality of the data by identifying and eliminating a subset of irrelevant or redundant variables that may otherwise decrease the accuracy and quality of the model.

A specific goal of this analysis was to assess the variability of weather features across the set of airports. The variability of these features is crucial as it affects the potential for developing pavement performance prediction models incorporating weather variables. If the variability across airports was found to be small, then including weather variables in these models would be unlikely to capture the effects of environmental conditions on pavement performance.

5.3.1 Clustering

Cluster analysis, an unsupervised learning technique, aims to explore homogenous subsets of data objects within a dataset by grouping data objects with the highest similarity to one another into clusters. Researchers implemented density-based clustering to group runway performance data samples based on the similarity between their weather features. Specifically, the hierarchical density-based spatial clustering of applications with noise (HDBSCAN) (McInnes, Healy, & Astels, 2017) method is employed to cluster data, where proximity is measured using cosine distance (Deza & Deza, 2009) across the standardized-scaled weather variables. The data objects for the cluster analysis are the averages of each weather variable over the time between two pavement condition inspections or between the inspection and the last rehabilitation, for each performance record in the dataset. In Figure 19, areas of red along the diagonal indicate similarity or high-density between data from the same airport with a flexible pavement as measured by cosine similarity. These areas of high density will likely lead to the same airports being in the same clusters, while airports in different climate regions (e.g., BOS and TUS) are blue, indicating larger distances between their weather variables and unlikely to be in the same cluster.

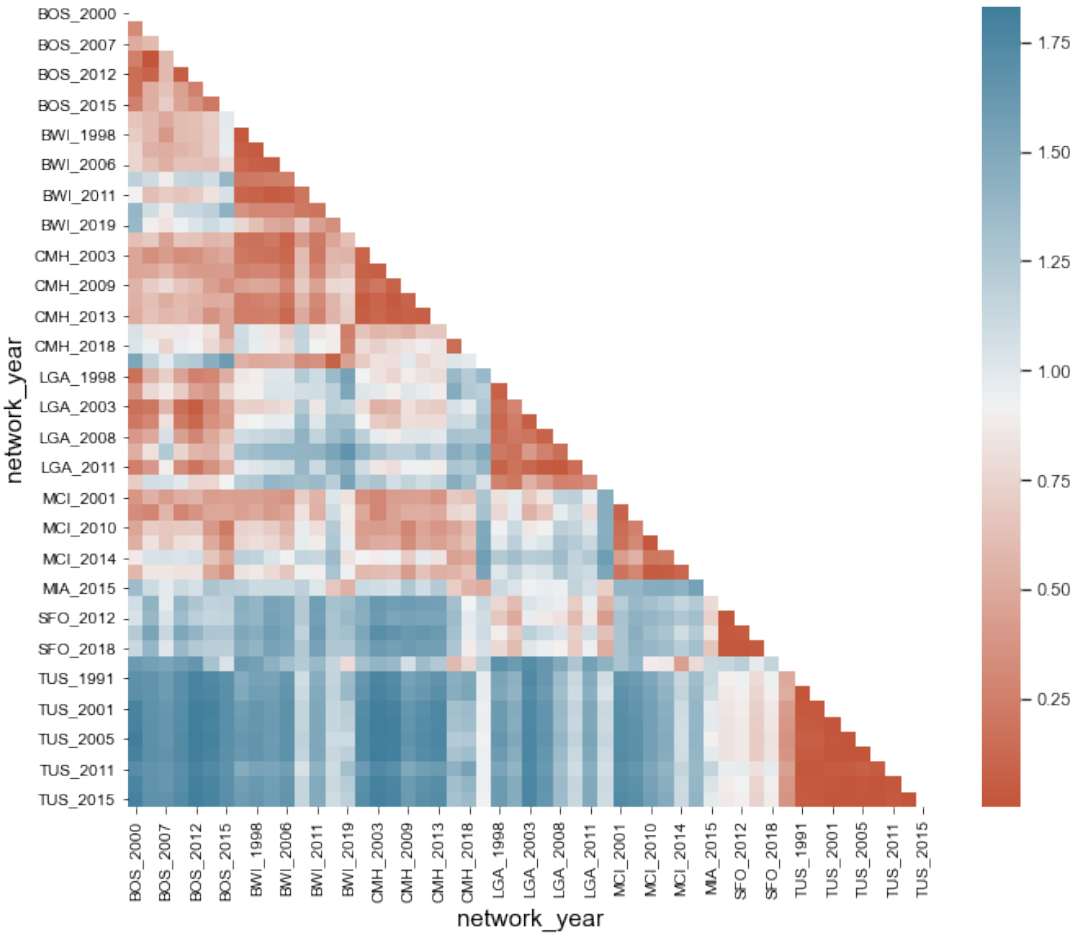


Figure 19. Similarities among weather variables in runways with flexible pavement

Figure 20 shows the results of HBDSCAN for airports with flexible pavements. The diagonal indicates how many correctly clustered data samples HDBSCAN grouped together (36 out of 50 total). The small number of erroneously grouped data samples (14 non-diagonal entries) indicate that these weather variables do contain variability necessary for incorporating weather features into pavement performance models. Notably, 12 of the 13 samples that were clustered into the incorrect network were all from the same climate region. This is a function of the sample quantity, i.e., there are too few to uniquely characterize an airport, but enough to identify climate regions.

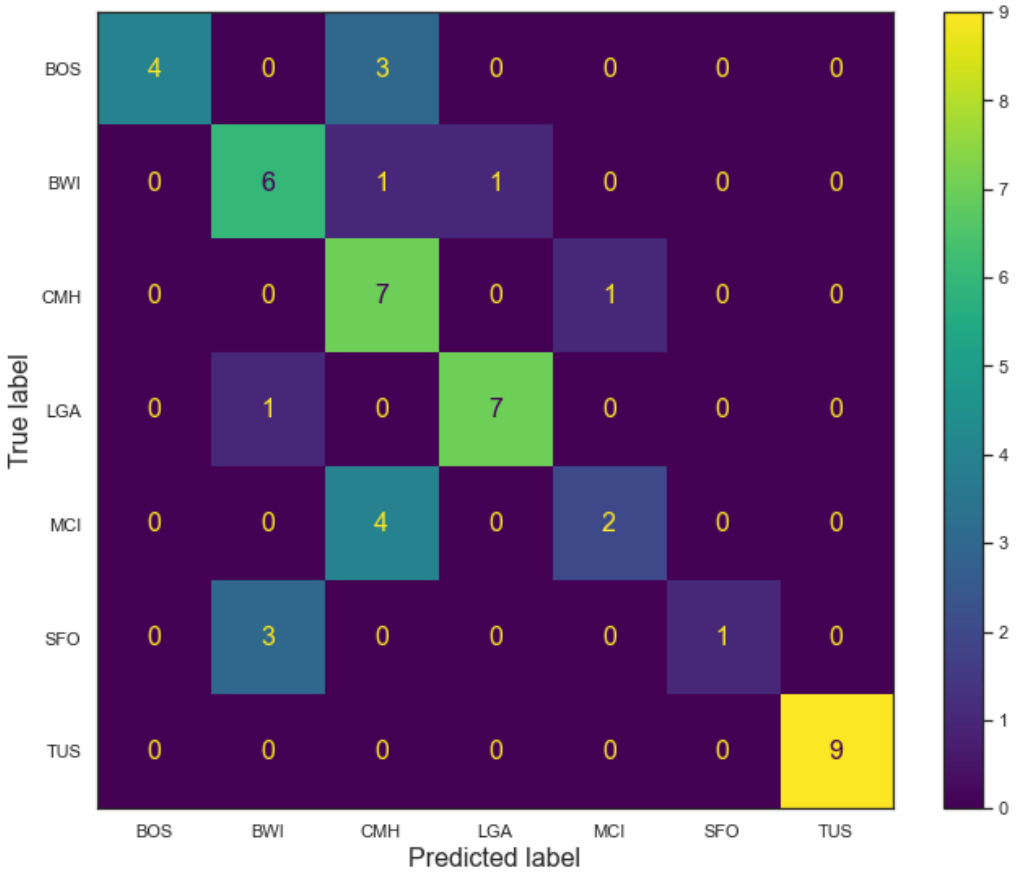
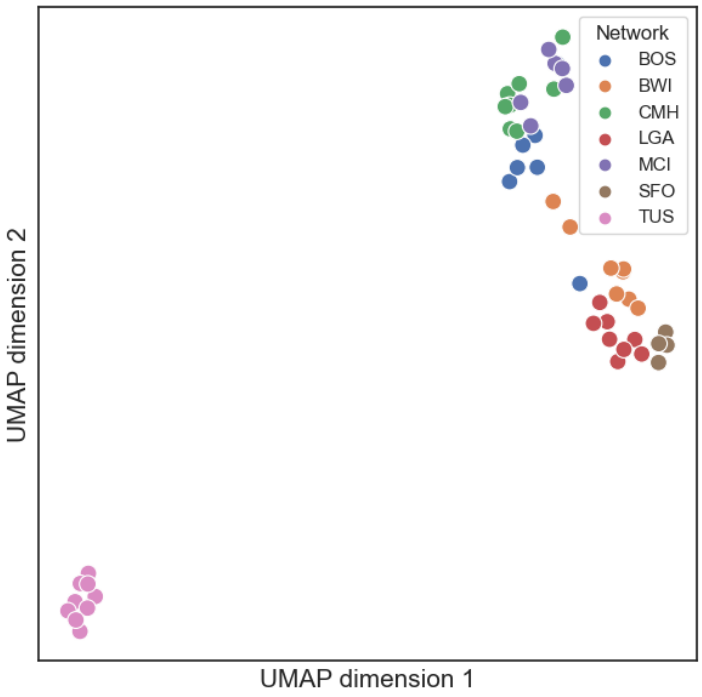


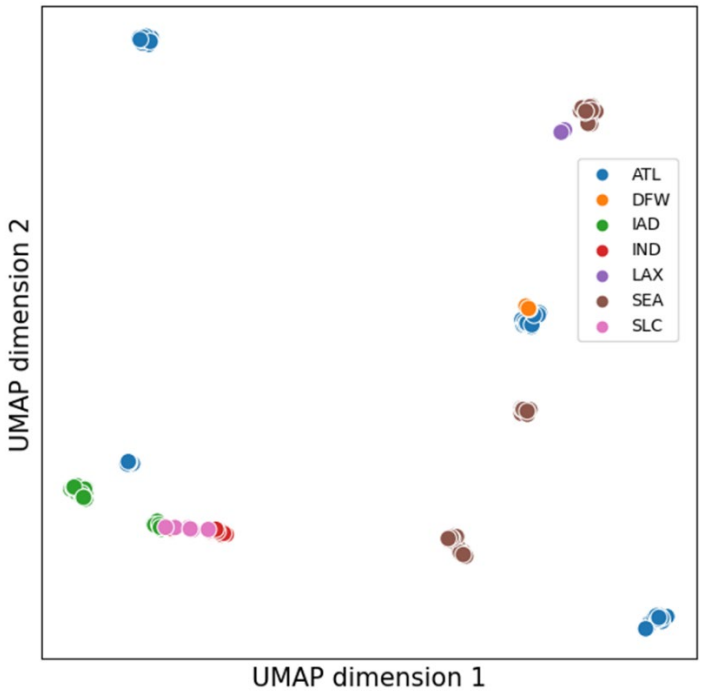
Figure 20. Hierarchical density-based spatial clustering of applications with noise flexible pavement performance data samples based on weather variables

5.3.2 Dimensionality reduction

Dimensionality reduction, also an unsupervised learning method, can assess the variability of weather features necessary to train a high-performing model. This technique compresses the feature set into a smaller set while maintaining the statistical relationships between and within each feature. The uniform manifold approximation and projection (UMAP) (McInnes, Healy, & Melville, 2018) is used to reduce the twelve weather features to two for easy visualization. Figure 21 illustrates the two new UMAP dimensions on the X and Y axes, with points colored by airport. The similar location of individual airport data and their proximity to data from other airports in the same climate region indicate that the weather features uniquely characterize airports and regions.



(a) Airports with flexible pavement



(b) Airports with rigid pavement

Figure 21. UMAP dimensionality reduction of weather variables showing that airports from similar climate regions have similar reduced dimensions

5.3.3 Dimensionality reduction

Embedding is another unsupervised learning technique that can assess features. The objective is to identify and reduce the effect of correlated weather variables on model predictions. Highly correlated features indicate measurements of the same climate property and may lead to models that are overly sensitive to these features. Computed correlations between all climate variables identify potential features for targeted dimensionality reduction into an embedded feature. Figure 22 shows the correlation values for all the weather features. Freeze-thaw cycles, FDD, days precipitation, total precipitation, and hydration days exhibit high correlations (0.75 and above), making them candidates for UMAP reduction to a single feature, also known as an embedding. Trial embedding of this feature set into one and two features with UMAP yielded a small performance improvement. This improvement was not deemed significant enough to justify the loss of interpretability however, as embedded feature values are not directly relatable to real-world phenomena and are solely dependent on the data used to create them.

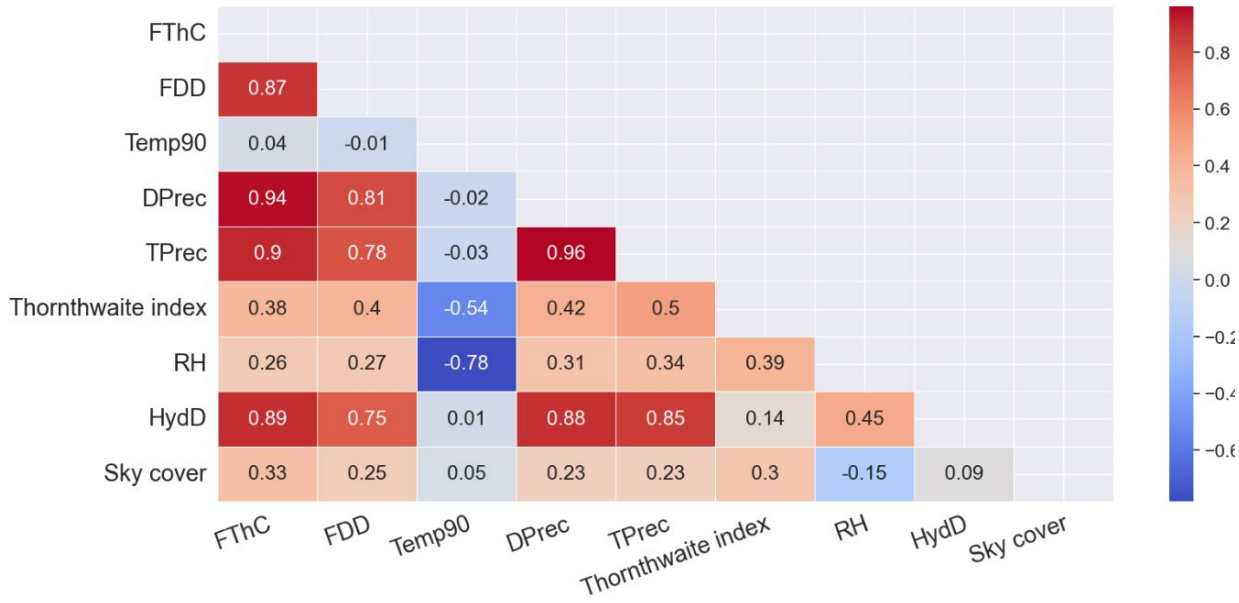


Figure 22. Correlation matrix between weather features

5.4 Re-sectioning of runway pavements

As discussed in section 3.4, the runways in the PA40 database have varying numbers of sections along their length. Airports divide the runways into more or fewer sections depending on their PMS practice. This variability could introduce a sampling bias when developing prediction

models, as the predictions may disproportionately reflect runways with more sections (i.e., more samples). To mitigate the potential bias, the researchers consolidated sections in runways that, in their original configuration, had many relatively short sections (less than 1000 feet long). The original sections chosen for consolidation all had similar characteristics (structure, age, condition/PCI, and location in relation to the runway end). The condition indexes for these newly combined sections were averaged from the contributing sections. The resulting average indexes were then taken as the representative condition indexes. For example, using this process the number of sections in BOS Runway 4L-22R was reduced from fourteen to six.

5.5 Prediction models for PCI, SCI, and anti-SCI

This section presents the results of ML prediction models for PCI, SCI, and anti-SCI. Using findings from the feature engineering analyses, features with the potential to influence pavement performance are incorporated into the prediction models. Weather variables showing poor predictive value were eliminated, leaving eight remaining variables for model development. An autoregressive approach was used for all indexes (i.e., the previous index value and the time since the previous inspection were predictors of the next index value). Predicting the current index is formulated as a regression task where it predicts the annual rate of change per year and multiplies this rate by the number of years since the previous inspection to obtain the final prediction. This approach allows prediction improvement on those variables whose effect is independent of the previous measurement of the index in question.

Table 12 provides a summary of the input features considered for each condition index. Traffic features include the annual average departures of three generic aircraft groups: D-200, 2D-400, and Heavy aircraft. In addition, the number of major rehabilitation events for each runway section is included as a feature. The location of the section with respect to the runway end (End or Center) is also considered a feature. The “End” label refers to the sections within approximately 1000 feet from either runway end, while the “Center” label refers to all other sections (between ends).

Table 12. Feature sets used for PCI, SCI, and anti-SCI prediction models

Feature Set	Feature Name	Descriptions
Age	Age	Time since last major rehabilitation (year)
Prev	Previous Index	Previous PCI, SCI, or anti-SCI measurement
	Delta Age	Time since previous PCI, SCI, or anti-SCI measurement (year)
Env	AvgTemp, FDD, FThC, Temp90, Dprec, Tprec, Thornthwaite, RH, HydD	Average annual weather variables between two PCI, SCI, or anti-SCI measurements or since last rehabilitation/reconstruction
SelEnv	FDD, Temp90, Dprec, HydD	
Const	Const	Number of previous rehabilitation projects
Location	CENTER	Binary variable indicating section location (0 End or 1 Center)
UMAP	UMAP 1 and 2	UMAP variables: FDD, FThC, Dprec, Tprec, HydD
Traffic	D-200	Average annual departures of generic D-200
	2D-400	Average annual departures of generic 2D-400
	Heavy Aircraft	Average annual departures of generic Heavy aircraft

Both linear (Linear, ElasticNet, Support Vector Regression) and gradient-boosting (Random Forest and CatBoost) methods were employed in the models. Emerging evidence suggests that these models typically have higher performance than their non-linear counterparts (e.g., neural networks) when working with tabular data (Borisov, et al., 2022). Linear regression models seek parameters for each variable that, when summed, minimize a loss function (typically and in this case, root mean squared error, RMSE) with respect to the ground truth. Gradient-boosting methods train an ensemble of (a collection of independent) tree-based models that all "vote" on the prediction. This ensemble approach often enhances performance compared to using a single model.

For each pavement performance index, the models and hyperparameters in Table 13 are used for the Model Selection and Hyperparameter Tuning steps of the MST. Hyperparameters represent the configuration options for each model and are not comparable across models. Support vector regression (SVR) models specify two hyperparameters for the decision function kernel as a categorical value indicating the function's complexity. By contrast, the three random forest (RF)

regressor hyperparameters are based on the number of features and data samples to ensure the tree-based structure is not overfit and doesn't simply memorize the data. Model training separates data into training and test sets, with 4-fold cross-validation where each of the 4 folds functions as the test set while the remaining three are used to train the model.

Table 13. Machine learning models and the hyperparameters used for prediction models

Linear regression	
N/A	No hyper-parameters are used/needed for linear regression
SVR (Cortes & Vapnik, 1995)	
Kernel	Linear, poly, rbf, sigmoid
Gamma	Auto, scale
RF (Breiman, 2001)	
Max depth	3, 6, 9
Min Leaf Samples	5, 10, 15
Max Leaf Nodes	10, 20, 30
ElasticNet regression (Zou & Hastie, 2005)	
alpha	0.8, 0.9, 1.0
L1 ratio	0.3, 0.4, 0.5, 0.6, 0.7, 1.0
CatBoost (Prokhorenkova, Gusev, Vorobev, Dorogush, & Gulin)	
Iterations	10, 20, 30
Learning rate	0.25, 0.5, 0.75
Depth	3, 6, 9

Figure 23 summarizes the highest performing models for PCI prediction across different combinations of feature sets. Model performance is evaluated using RMSE between the true and predicted values from the test dataset. RMSE is the standard deviation of prediction error and indicates how concentrated the predictions are around the fitted prediction curve. Lower RMSE indicates a better prediction.

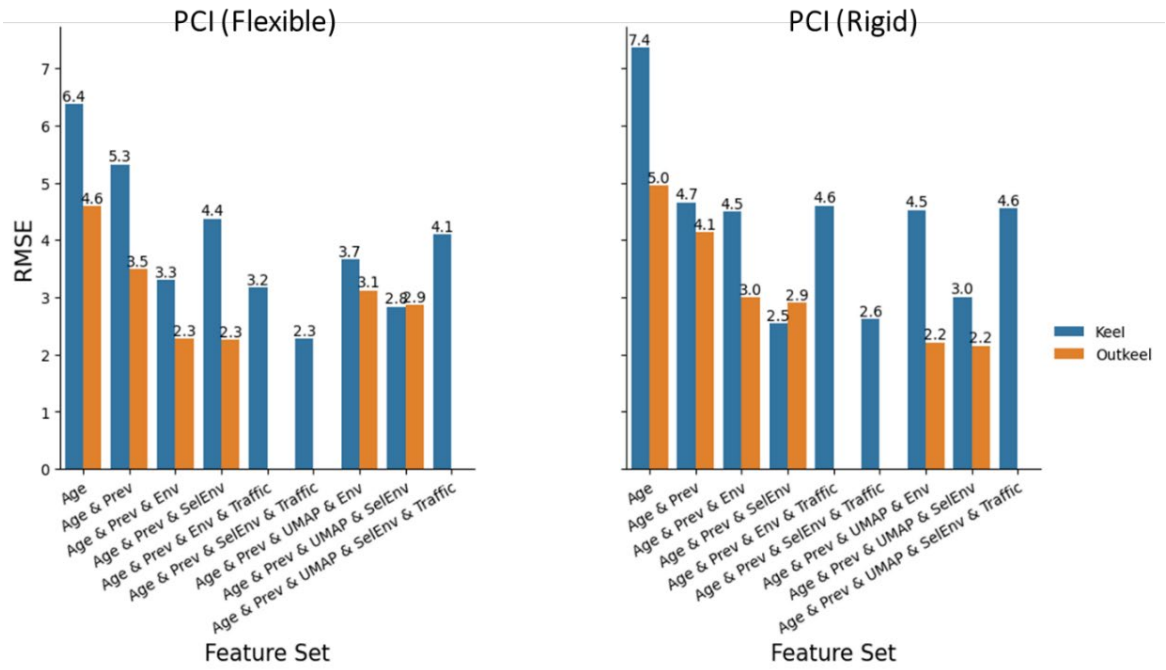


Figure 23. Pavement condition index using CatBoost model across different features for flexible and rigid runway pavements

Models incorporating only pavement age had the highest RMSE for both flexible and rigid pavements. Incorporating previous measurements improves performance. For flexible keel pavement sections, a trained CatBoost model achieved the highest performance (RMSE 2.3) using a feature set with age, previous measurements, selected weather variables and traffic. For flexible outkeel pavement sections, the best performing CatBoost model (also RMSE 2.3) included age, previous measurements, and selected weather variables in the feature set.

UMAP embedded features generally increased performance but not significantly enough to be the highest performing, indicating that while the original five features are correlated, their collective influence on the model was not strongly negative.

High model performance is crucial but understanding how the model makes predictions is also important for interpretability. SHAP (SHapley Additive exPlanations) values (Lundberg & Lee, 2017) are a method for attributing input feature values to the prediction outputs, aiding in verifying a model's predictions with engineering judgment. Figure 24 shows the SHAP values for the highest performing PCI flexible keel model. In this plot, color indicates the feature input value and its corresponding SHAP value, which refers to a difference from the average predicted value, not a measure of performance. Three properties of the features displayed in the SHAP values plot are of interest: dispersion, separation, and symmetry. Dispersion refers to how

densely features are clustered around similar prediction values, indicating the sensitivity of the model to changes in a feature. For example, the "Annual Avg Days Temperature over 90" feature spans almost the entire SHAP range, suggesting its impact on predictions is widely dispersed. This observation aligns with the fact that some climate zones (such as the cool, wet region including BOS) experience relatively few days with temperatures over 90°F (Temp90), while such high temperatures are very common in the hot, dry zone including TUS. Therefore, while the effect of Temp90 on pavement is minimal in BOS, it may be significant in TUS. Separation refers to how high or low feature values affect predictions, indicating a correlation or anti-correlation relationship. For instance, the "Previous PCI" feature's lowest values (blue) have a predominantly negative impact on the SHAP value, whereas higher values (red) exhibit a range across negative and positive SHAP values, indicative of a small correlative effect. By way of contrast, the number of previous rehabilitation or "Const" feature's low (blue) and high (red) values are densely clustered around a narrow range of prediction effect, but without any separation between high and low feature values, indicating the feature's contribution to predictions is complex without direct correlation. Symmetry examines whether feature contributions are evenly distributed around a SHAP value of 0 (on the x-axis), indicating a linear relationship. The "Delta Age" feature exhibits a symmetrical distribution, though with limited separation. Its lack of strong correlation and weak linearity suggests a minimal and indirect influence on the predicted value.

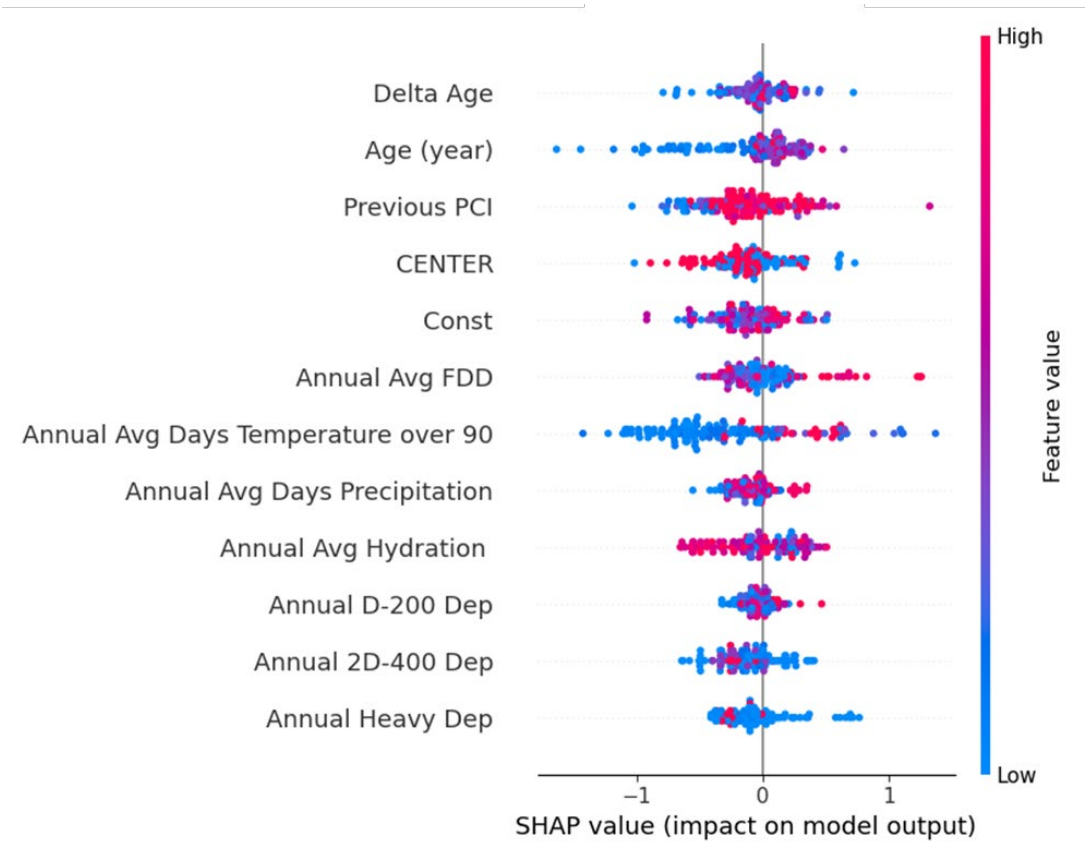


Figure 24. SHapley Additive exPlanation values for the highest performing PCI model for flexible keel

Returning to Figure 23, for rigid pavement keel sections, the feature set for the highest-performing (RMSE 2.5) trained CatBoost model included age, previous measurements, and selected weather variables. A close second (RMSE 2.6) added traffic to the feature set. Traffic level has practically no influence on the prediction, as expected, because the traffic indexes used in the analysis neglect overdesign or (less commonly) under-design of the pavement structure.

As the next validation step, researchers evaluated the accuracy of the PCI models for different prediction horizons. For the best performing flexible keel and rigid keel models, separate models were trained using these following values of "Delta Age": < 5, 10, 15, and >15 years. Figure 25 shows the RMSE for each horizon. While the performance was relatively consistent until 15 years, the error for the rigid model started to increase beyond this point, suggesting that the model should be used with caution for pavement age beyond 15 years. Figure 26 shows the measured vs predicted PCI for flexible and rigid keel sections.

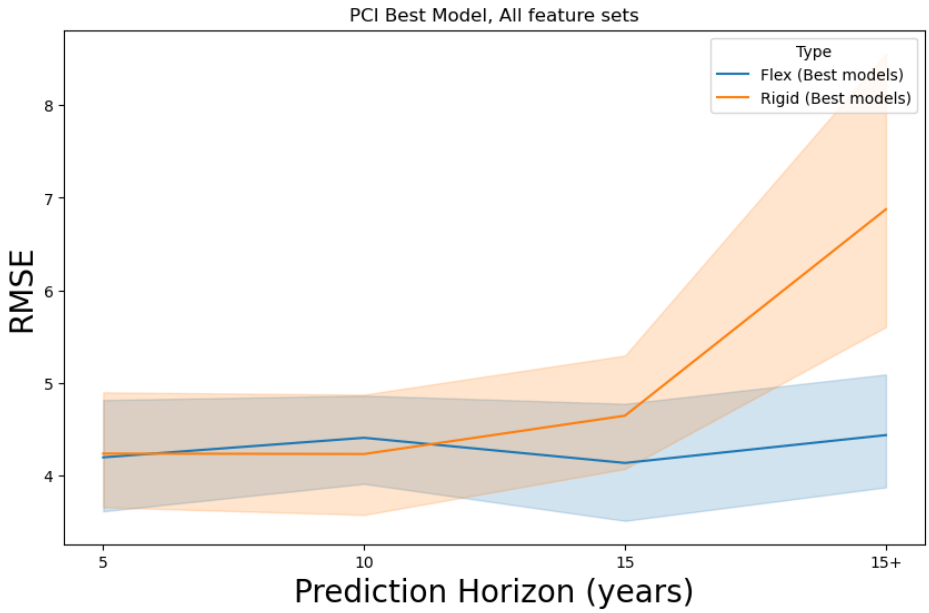
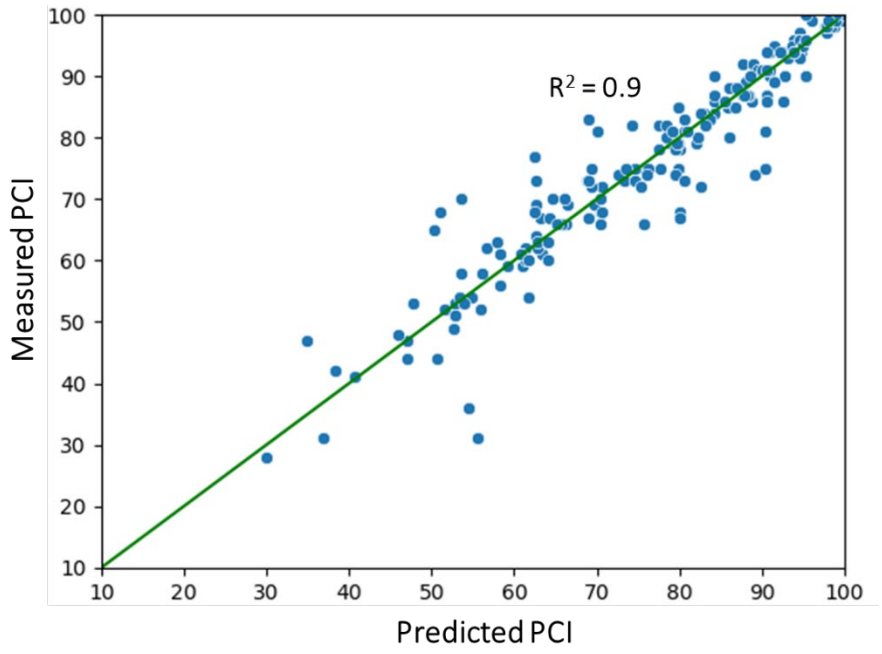
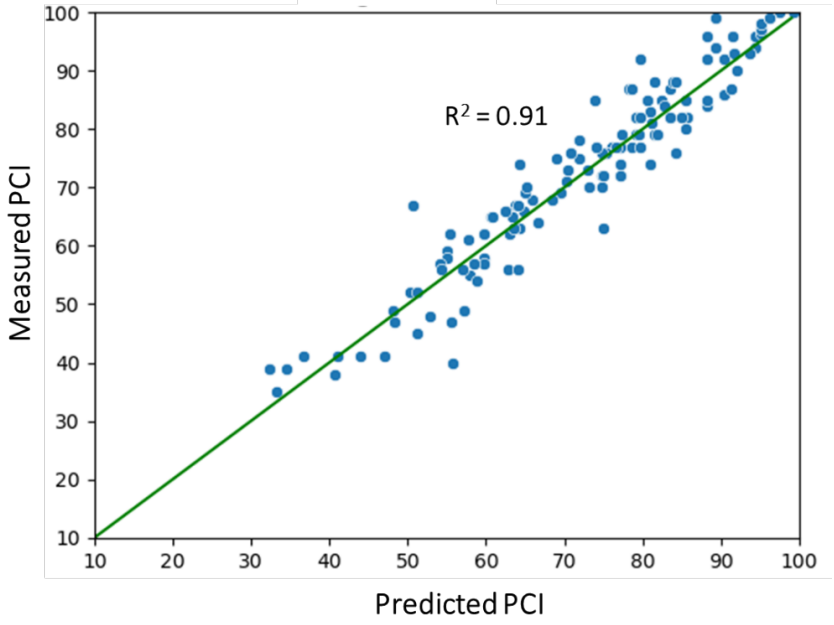


Figure 25. Pavement condition index model prediction horizon



(a) Flexible keel



(b) Rigid keel

Figure 26. Measured vs predicted for the best performing (a) flexible and (b) rigid keel models

Figure 27 presents a summary of the highest performing models for anti-SCI prediction across different combinations of feature sets. Incorporating the weather variables improved the model performance compared to the purely age-based model. UMAP embedded features generally increased performance, but again, not enough to justify the lack of interpretability. The anti-SCI

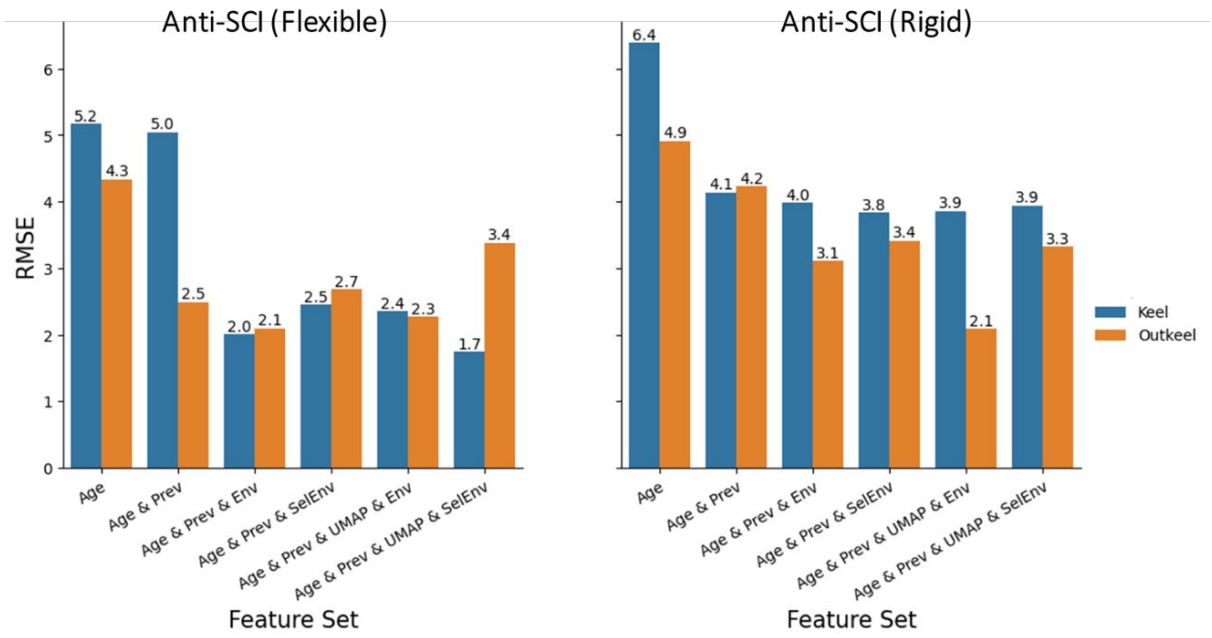


Figure 27. Anti-structural condition index prediction performance models for flexible rigid runway pavements

models for rigid pavement had generally higher errors compared to the models for flexible pavement.

Figure 28 shows the SHAP values for the highest performing anti-SCI flexible keel model. The plot represents how the values (color) of each feature (y-axis) affected a model's average prediction value (x-axis), the SHAP value. The “Delta Age” feature had an outlier, meaning it was widely dispersed and the impact on predictions varies more than the other features. The "Previous anti-SCI" feature's lowest values (blue) had a predominantly negative impact on the SHAP value, whereas higher values (red) had a range across negative and positive SHAP values, indicating a small correlative effect. The “CENTER” feature had a symmetrical property with high separation, indicating a linear and correlated relationship with SHAP values. However, its effect on predictions (RMSE) was in a relatively small range [-0.5, +0.5].

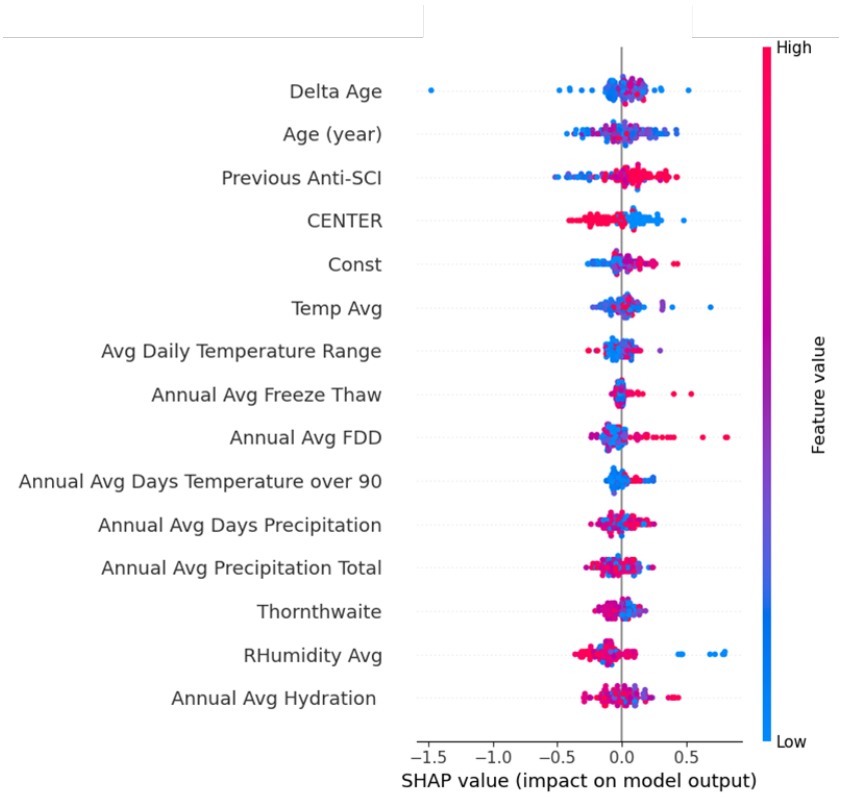


Figure 28. SHapley Additive exPlanation values for the highest performing anti-SCI model for flexible keel

For the best performing flexible keel and rigid keel models, separate models were trained using delta age less than 5, 10, 15, and beyond 15 years. Figure 29 shows the RMSE of anti-SCI models. Model performance generally decreases as the prediction horizon increases, but the RMSE for predictions beyond 15 years increased more significantly for the rigid models compared to the flexible models.

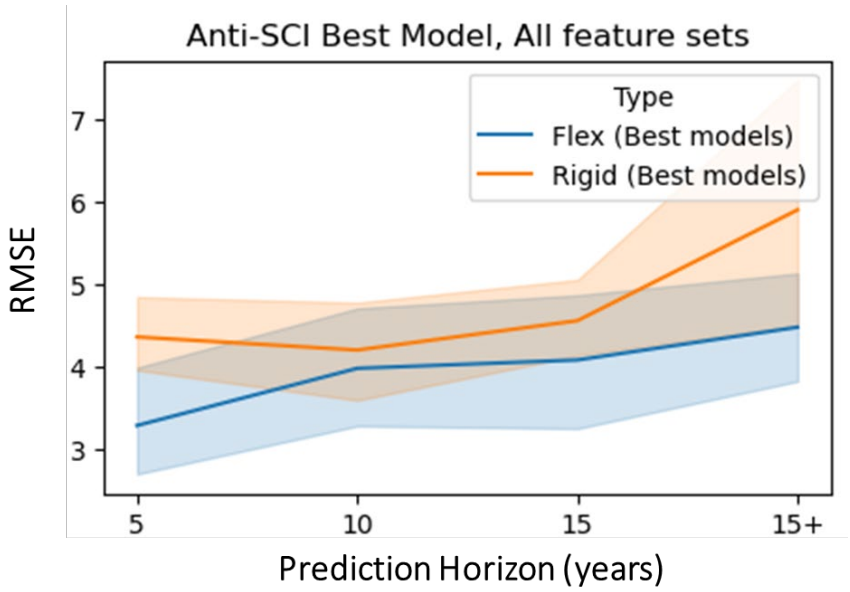


Figure 29. Anti-structural condition index model prediction horizon

Figure 30 summarizes the highest performing models for SCI prediction across different combinations of feature sets. Incorporating weather variables improved performance compared to the solely age-based model. While incorporating the traffic variables had no effect on the performance of flexible pavement models, it improved the performance of the SCI model for rigid pavement. The UMAP embedded features slightly reduced RMS in flexible pavement models but had no meaningful effect on rigid pavement performance.

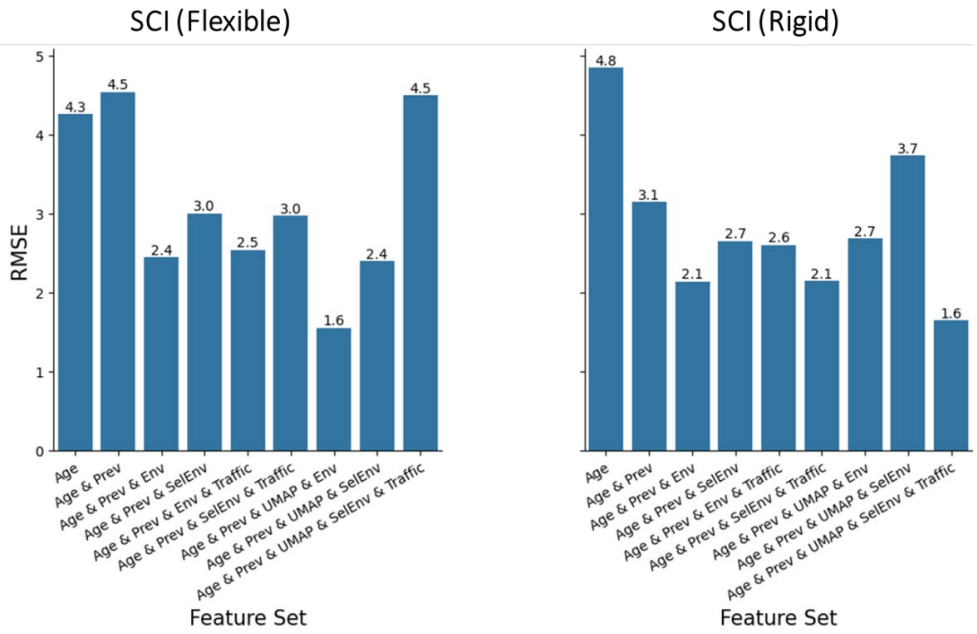


Figure 30. Structural condition index prediction performance across models for flexible and rigid keel sections

Figure 31 shows the SHAP values for the highest performing SCI flexible keel model. The “Previous SCI” feature spans a wide SHAP range, meaning it was widely dispersed and the impact on predictions varied more than other features. The “Age” feature's lowest values (blue) had predominantly negative SHAP values (i.e., low values of this feature are associated with lower-than-average decreases in SCI), while higher values (red) had positive SHAP values (a small positive correlation). The "Center" feature has a symmetrical property, but without good separation.

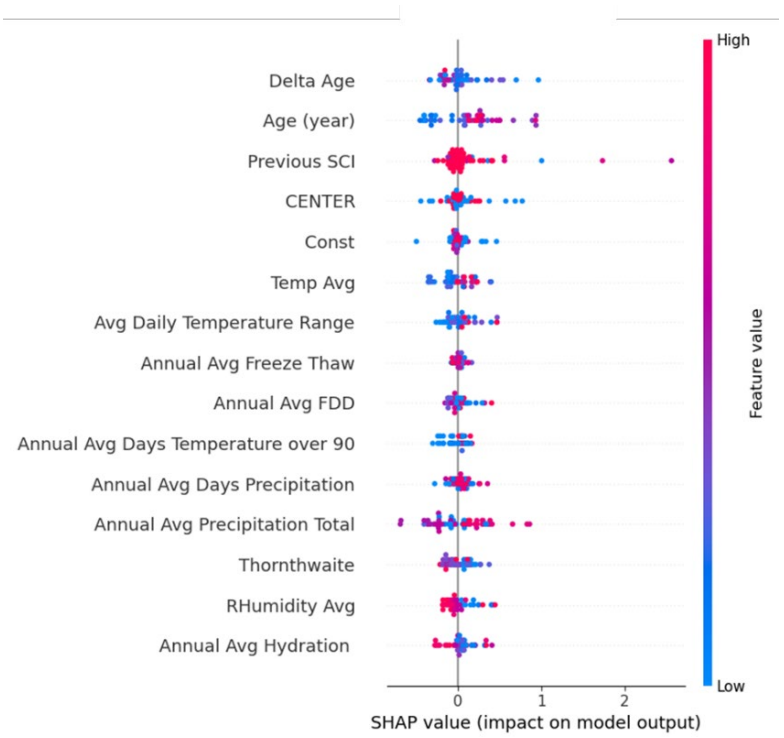


Figure 31. SHapley Additive exPlanation values for the highest performing SCI model for flexible keel

For the best performing flexible keel and rigid keel models, researchers trained separate SCI models using prediction horizon ("delta age") values: < 5, 10, 15, and >15 years. Figure 32 shows the RMSE of the SCI model prediction for the various time horizons. While the performance is relatively consistent until 15 years, the error for the rigid model increases after 15 years.

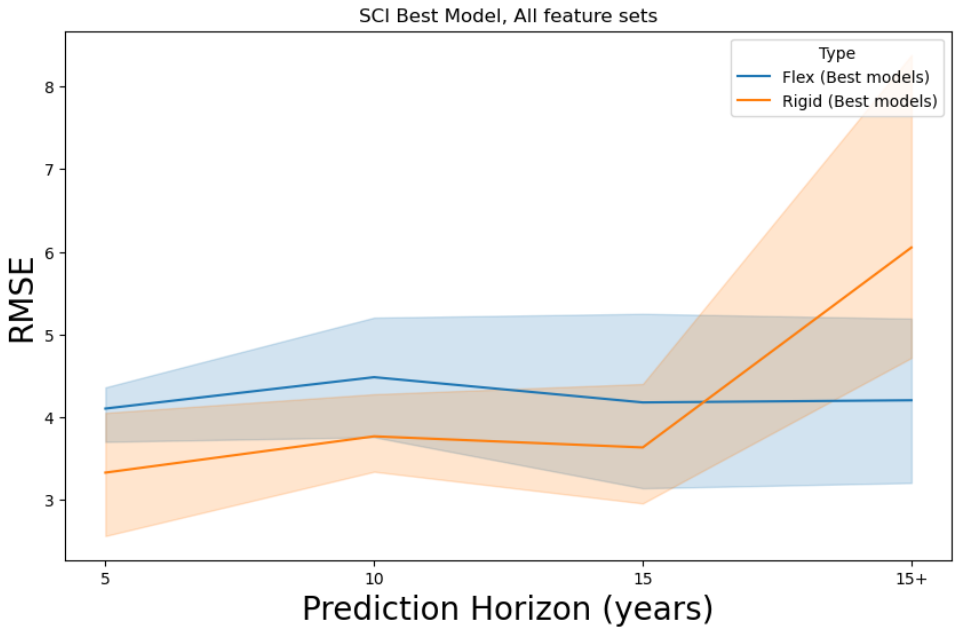


Figure 32. SCI model prediction horizon

5.6 Prediction model for runway roughness index (RRI)

This section covers ML models that predict RRI for keel sections of flexible and rigid runway pavements. Compared to the PCI indexes, there are fewer available data samples for RRI (50 flexible and 46 rigid samples), and most samples do not contain a previous measurement. Therefore, the ML model for RRI uses continuous regression (unlike the autoregressive model used for PCI-derived indexes). Table 14 summarizes the input features considered for RRI. The model incorporates five traffic features, as follows: average total annual departures, average total annual arrivals, and annual average departures of three generic aircraft groups.

Table 14. Feature sets used for RRI and groove index prediction models

Feature set	Feature name	Descriptions
Age	Age	Time since last major rehabilitation (year)
Env	Avg Temp, FDD, FThC, Temp90, Dprec, Tprec, Thornthwaite, RH, HydD	Average annual weather variables since last rehabilitation/reconstruction
SelEnv	FDD, Temp90, Dprec, HydD	
Location	CENTER	Binary variable indicating section location (0 End or 1 Center)
Traffic	D-200	Average annual departures of generic D-200
	2D-400	Average annual departures of generic 2D-400
	Heavy aircraft	Average annual departures of generic Heavy aircraft
	Total departure	Average annual departures of all aircraft
	Total arrivals	Average annual arrivals of all aircraft

The RRI model uses a simple linear regression method due to the small dataset size, instead of the models/hyperparameters used for the PCI index models. Figure 33 summarizes the model performance (RMSE) across the relevant feature sets. The RMSE must be interpreted within the range of the measured RRI, which falls between 0.2 to 0.43. For the flexible pavement, the model based solely on pavement age exhibits the best performance, indicating that weather and traffic variables have no influence on RRI. However, incorporating these variables decreases the error for rigid pavements. These error rates suggest that the model is likely overfit to the data; in other words, the trained model simply memorizes the training data. This is caused by the lack of variability in the test data (due to the small number of samples) which may give an incorrect impression of acceptable performance.

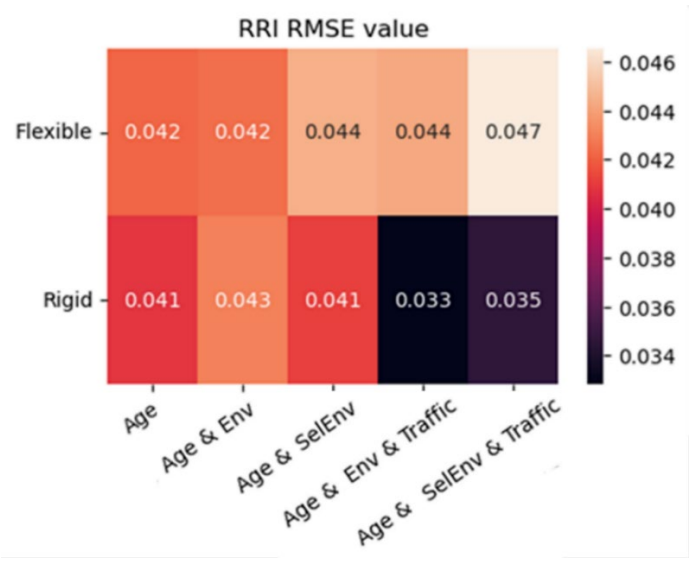


Figure 33. Runway roughness index model results from linear regression

Another reason for using linear regression as a preliminary approach is to facilitate regularization. Regularization is a technique to deliberately shape the range of learned weights for each feature, constraining the solution space of trained models with the goal of obtaining a less overfit model. The goal was to determine what, if any, feature sets are more influential and hold potential for higher predictive performance. The degree of regularization applied to the model is represented by *a*, with larger *a*-values corresponding to stronger regularization (less overfit). As shown in Figure 34, researchers adjusted the weight allocated (*x*-axis) to feature sets (*y*-axis) and trained a series of RRI models for flexible pavement. Results indicated that regularization across features did not yield any substantive performance improvement or trend. The same results were observed for rigid pavement models. This suggests that substantially more data samples are needed to identify any relationship between RRI and the feature sets.

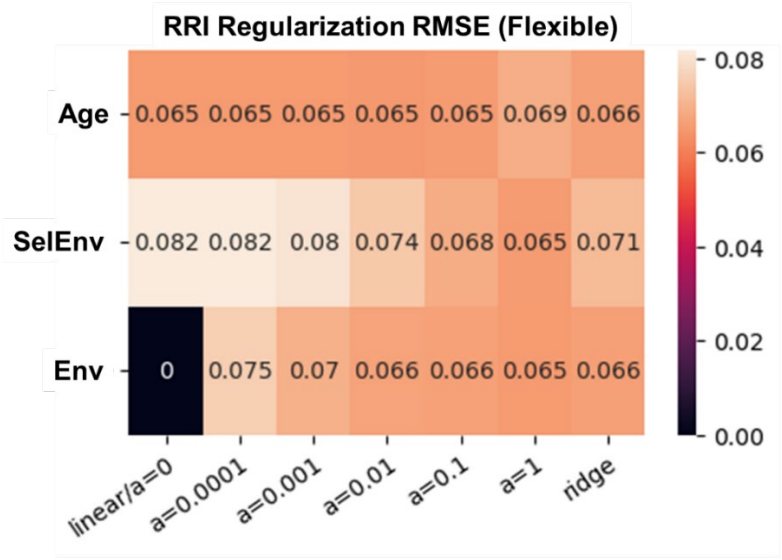


Figure 34. Regularization term (*a*-value on *x*-axis) applied to different feature sets in RRI model

5.7 Prediction model for Groove Index (GI)

This section covers ML models that predict GI for flexible and rigid keel runway pavements. Similar to RRI (section 5.6) the GI model is based on a continuous supervised learning regression model. Also, like RRI, the data samples for training were small (40 flexible and 35 rigid samples) and most samples did not contain a previous measurement. The feature sets used for training were the same as those for the RRI model (Table 14).

The GI model also uses a simple linear regression method due to the small dataset size. Figure 35 summarizes the model performance (RMSE) across the relevant feature sets. The RMSE must be interpreted within the range of the measured GI, which falls between 10 to 85 percent. For the flexible pavement, incorporating weather variables slightly decreased the error, while for the rigid pavement, including traffic features reduced the error. While the errors were not large, the model is likely overfit to the data. This was caused by the lack of variability in the test data due to the small number of samples.

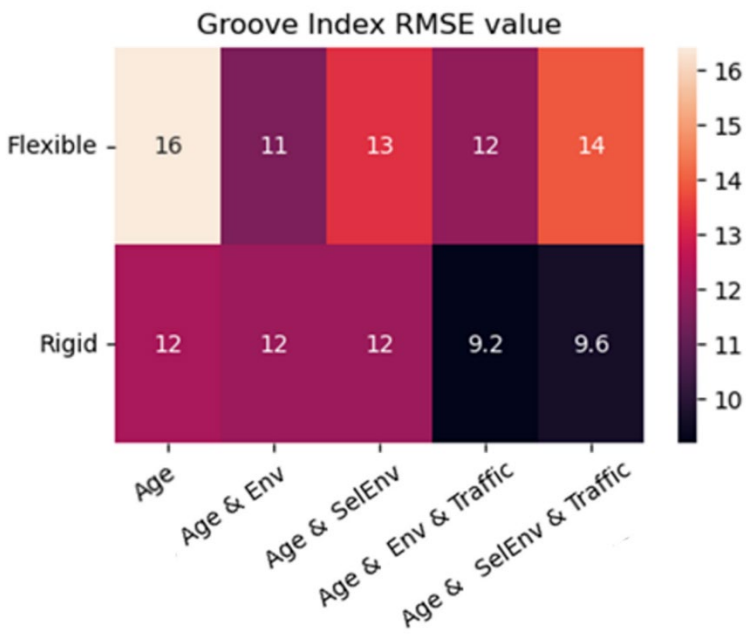


Figure 35. GI prediction model results from linear regression

A regularization method was used to identify influential features that might lead to higher predictive performance. As in section 5.6, the regularization term is represented by the *a*-value on the *x*-axis in Figure 36. In this approach, the regularization term was increased by orders of magnitude and applied to different feature sets (*y*-axis). The GI models were retrained for each regularization term and feature set. Figure 36 shows that as the weight assigned to the Env feature set (all weather variables) increased, the RMSE in both flexible and rigid models gradually decreased (to 9.37 for flexible pavements and 11.7 for rigid pavements). This finding is encouraging and suggests that an increase in data samples may reinforce this relationship, leading to a higher-performing and more generalizable model for predicting GI in both flexible and rigid runways.

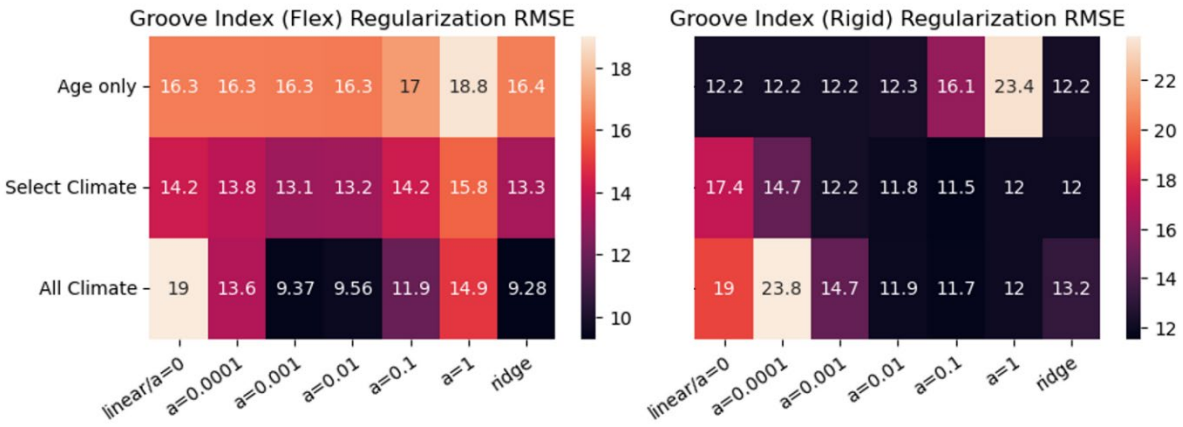


Figure 36. Regularization term (*a*-value on *x*-axis) applied to different feature sets in GI models

6 Serviceability Level (SL) model

The serviceability level (SL) model is formulated as a classification task, aiming to determine whether the runway pavement is serviceable, predicated on a set of pavement condition indexes. Classification, a supervised machine learning technique, entails predicting discrete-valued classes using a given set of attributes. As detailed in the preceding sections, the SL model incorporated five pavement condition indexes (PCI, SCI, anti-SCI, RRI, and GI) as the attributes. The model's output is binary, classifying pavements into serviceable or unserviceable categories.

Distinct SL models were formulated based on two categories: pavement type (flexible or rigid), and section location (keel or outkeel). The MST approach was applied for SL model development. Given the insights obtained from the feature engineering analysis in Section 5, a separate analysis was unnecessary, as the previous analysis provides sufficient information on the contributing features to the SL model.

6.1 Construction of ML training database for SL model

The training database consists of labeled data points, each associated with a set of attributes and a corresponding class label. It is essential that all data points contain values for all contributing attributes (pavement indexes). The following four steps were undertaken to construct the training database for each model category (flexible/rigid and keel/outkeel).

Step 1 – Identify Classes

All inspection and rehabilitation/reconstruction records from all runway sections were integrated into a unified dataset to form the data points for training. Any data point with a rehabilitation/reconstruction record is classified as “Unserviceable,” since it may be assumed that pavements undergoing major rehabilitation are at or below the lower serviceability thresholds. Conversely, a data point with an inspection record is classified as “Serviceable,” as the runways are likely in serviceable condition during the inspection.

Step 2 – Determine Missing Attributes for Unserviceable Pavement

Pavements are often not surveyed immediately prior to undergoing rehabilitation. Consequently, many data points classified as "Unserviceable" lack associated pavement condition indexes, which poses a challenge for maintaining a complete set of attributes for each data point in the training database. To address this, the ML models developed for individual indexes were leveraged to estimate the missing pavement condition indexes at the time of runway rehabilitation or reconstruction. This approach involves computing all contributing features, such as pavement age, weather variables, and traffic variables, for each rehabilitation record and inputting them into the models to calculate the estimated index. Pavement condition indexes associated with pavement unserviceability are critical attributes, informing decisions regarding rehabilitation, reconstruction, or replacement. Figure 37 shows all PCI, SCI, and anti-SCI data samples from flexible keel sections, including the estimated indexes at the rehabilitation time (marked as unserviceable).

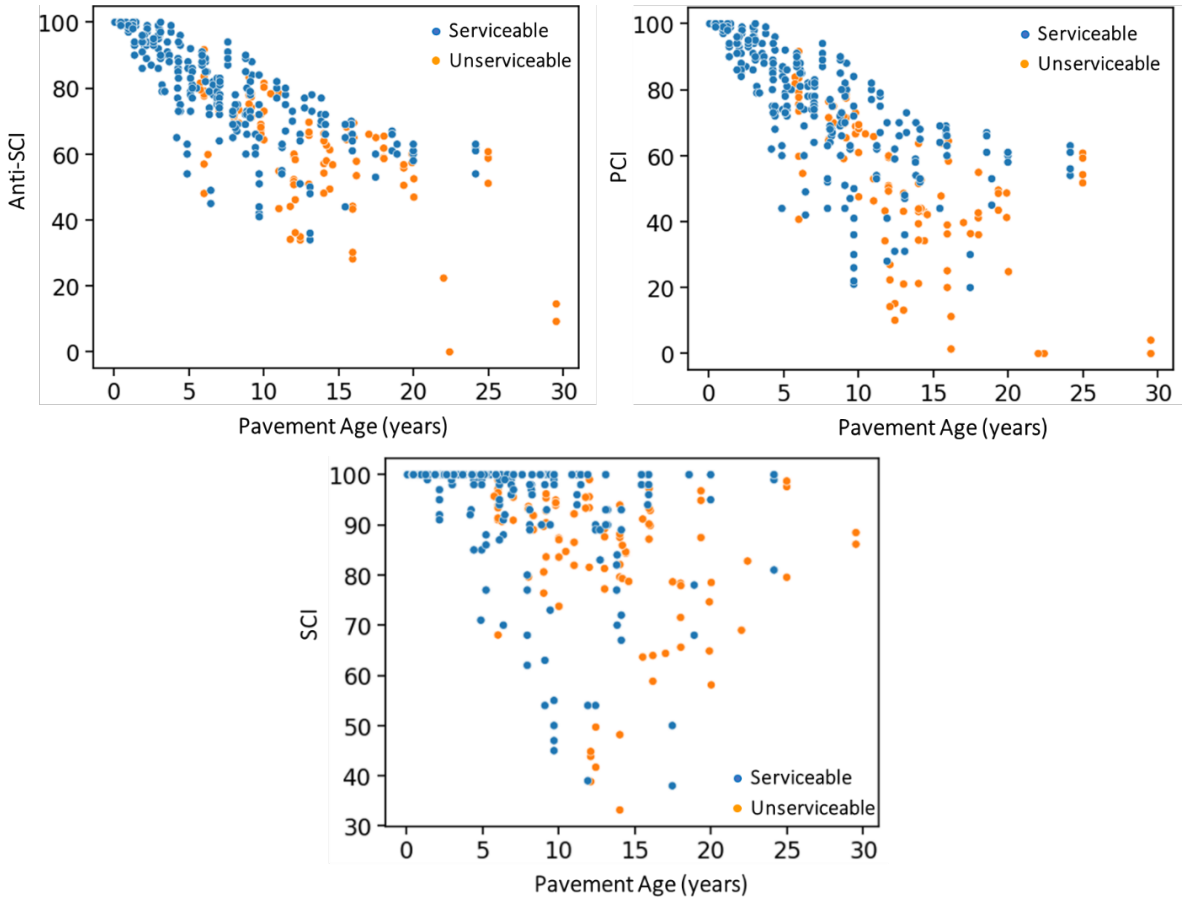


Figure 37. Measured (serviceable) and predicted (unserviceable), PCI, SCI, and anti-SCI data from flexible keel sections

Step 3 – Determine Missing Attributes for Serviceable Pavement

Since roughness and grooving data are only available for runways surveyed by the FAA, these indexes are absent for other runways and from earlier surveys where historic PCI data are available. To bridge this gap, the ML models were used to estimate the RRI and GI for every inspection survey date.

Step 4 – Generate Synthetic Data

There is a significant imbalance within the data samples, with a much larger proportion of serviceable pavements than unserviceable ones. This class imbalance can result in models with misleading performance due to underfitting, where the signal may be mistaken for noise. To ensure the validity of the resulting model performance, researchers generated synthetic unserviceable data samples using adaptive synthetic data (ADASYN) (He, Bai, Garcia, & Li, 2008) and synthetic minority oversampling technique (SMOTE) (Chawla, Bowyer, Hall, & Kegelmeyer, 2002). Figure 38 illustrates the class imbalance in the original data and the balanced

classes achieved by adding synthetic data from ADASYN and SMOTE to the original dataset to create new datasets. Figure 39 shows the distribution of generated synthetic data along with the original PCI, SCI, and anti-SCI.

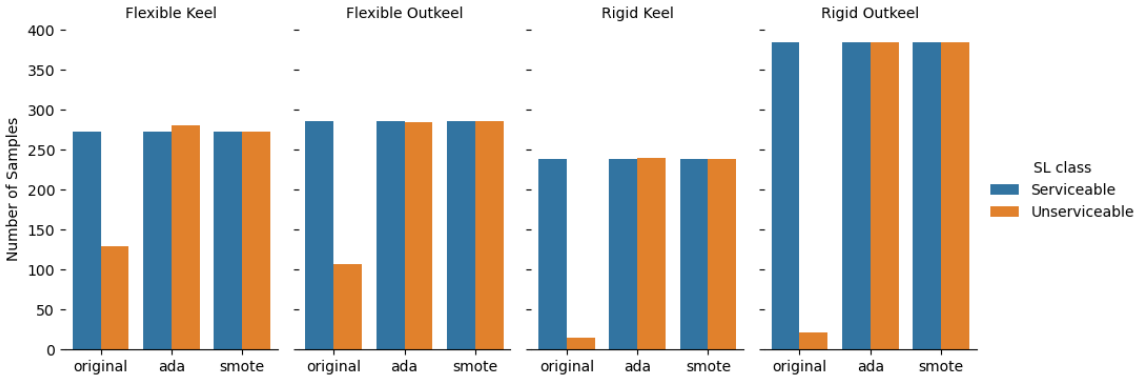


Figure 38. Synthetic data samples from ADASYN and SMOTE reduce the class imbalance between serviceable/unserviceable across pavement types and locations

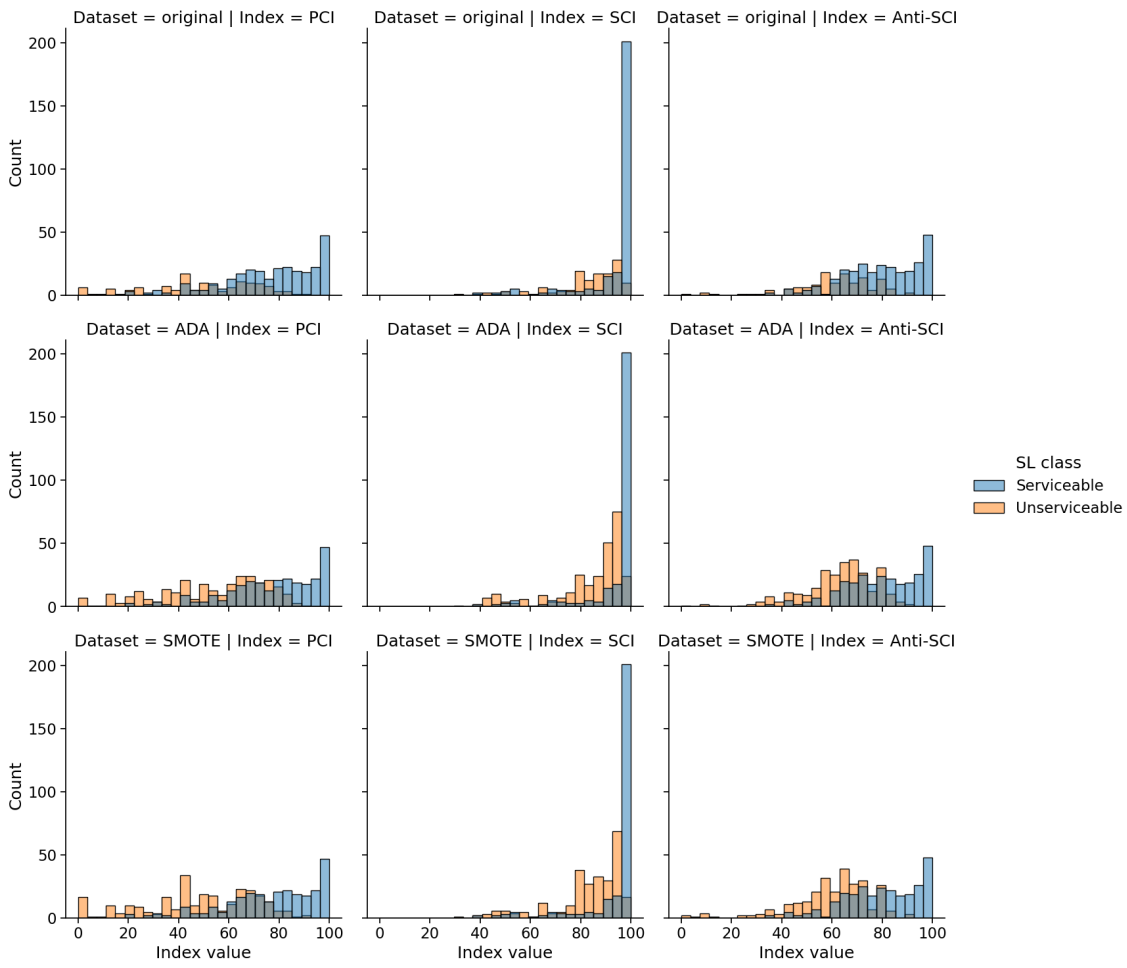


Figure 39. Original and synthetic (ADASYN) PCI, SCI, and anti- SCI data from flexible keel sections

6.2 Classification performance measures

Formulating SL index prediction as a classification task requires different performance measures than regression-based tasks. Performance is initially represented by a confusion matrix, which allows for the comparison of predictions and actual quantities of positive and negative classes. From the confusion matrix, various measures can be derived, all of which emphasize the ratios of correct and incorrect predictions of positive and negative classes. Different measures are applicable depending on specific risks associated with the application. For example, in cancer screenings, it may be preferable to have false positives rather than false negatives. In this context, the F1 score was used as it provides a balanced assessment of these risks within a single calculation. The F1 score is defined by Equation 3:

$$F1 = \frac{2tp}{2tp+fp+fn} \quad (3)$$

where:

tp are true positives (unserviceable predicted as unserviceable),

fp are false positives (serviceable predicted as unserviceable), and

fn are false negatives (unserviceable predicted as serviceable).

6.3 Classification model selection

In a classification task, adjustments to the models and hyperparameters used in the MST are necessary. Researchers conducted experiments to adapt the prediction models listed in Table 13 to their classification versions while keeping the hyperparameters consistent. For example, instead of employing linear regression, logistic regression was used. While the underlying mathematical principles remain the same, the prediction output differs. That is, rather than a scalar prediction, logistic regression provides a likelihood estimate indicating the probability that the input features belong to the positive class. This alteration to the prediction is the sole distinction across all the models.

6.4 Serviceability Level (SL) model results

Researchers explored various scenarios for the SL model by incorporating a combination of five attributes (i.e., pavement performance indexes), providing airport managers with flexibility in determining serviceability based on available performance metrics. The models were built using three datasets: the original dataset, the original dataset with ADASYN synthetic data, and the original dataset with SMOTE synthetic data. Table 15 illustrates the contributing attributes for each pavement type and location model.

Table 15. Attribute sets used in the MST approach for the SL model development

Attributes set	**Flexible keel**	**Flexible outkeel**	**Rigid keel**	**Rigid outkeel**
PCI + RRI + GI	X		X	
Anti-SCI + SCI + RRI + GI	X		X	
Anti-SCI + SCI	X	X	X	X
Anti-SCI + RRI + GI	X		X	
Anti-SCI + SCI + RRI	X		X	
Anti-SCI + SCI + GI	X		X	
PCI + RRI	X		X	
PCI + GI	X		X	
PCI	X	X	X	X
Anti-SCI		X	X	X
Anti-SCI + RRI			X	
Anti-SCI + GI			X	

The fast, lightweight auto-ML (FLAML) (Wang, Wu, Weimer, & Zhu, 2021) package was used to automate model-hyperparameter training across gradient-boosting and linear model types. Figure 40 summarizes the F1 scores for the SL models trained over the original dataset across the attribute sets for flexible keel and outkeel runway sections. The highest F1 score for the flexible keel SL model was 0.95, incorporating PCI and RRI as attributes to the random forest as the classification method. The SL model for flexible outkeel had an F1 score of 0.94, incorporating SCI and anti-SCI as attributes. Although both scores are high performing, the model with synthetic data was used to further investigate these scores to ensure they are not a function of the class imbalance.

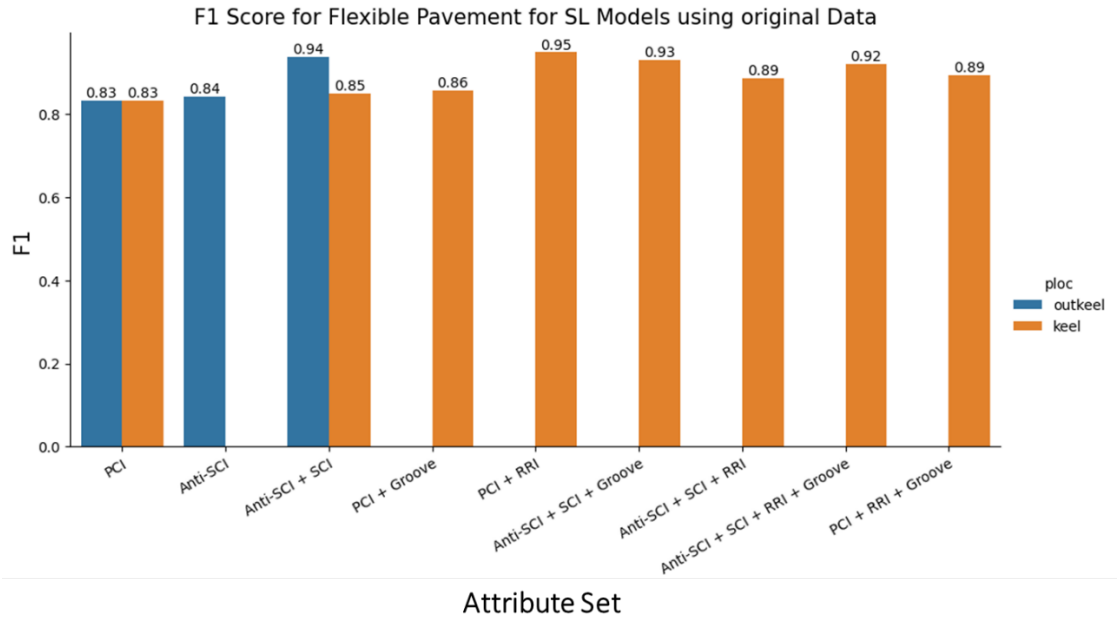


Figure 40. Serviceability level F1 performance for flexible keel and outkeel sections for different attribute sets

The left-hand side of Figure 41 illustrates the confusion matrix for the flexible keel SL model trained on the original dataset. The diagonal contains the correctly classified data samples (127 total) while the anti-diagonal contains the incorrectly classified data samples (only 6). This low disparity is excellent and indicative of high-performance; however, the imbalance on the diagonal (82 vs. 46) suggests that the high F1 performance may be unrealistic. The right-hand side of Figure 41 shows the confusion matrix of the same model with the same attributes but trained using the original dataset plus class balancing synthetic data from SMOTE. The additional number of unserviceable samples in both the correctly and incorrectly classified predictions resulted in a reduced overall F1 performance of 0.91. Despite this reduction, the score still indicated a valid prediction. The synthetic samples generated from a reasonably sized minority class gave a more reliable performance estimate than solely relying on the F1 score of the original dataset.

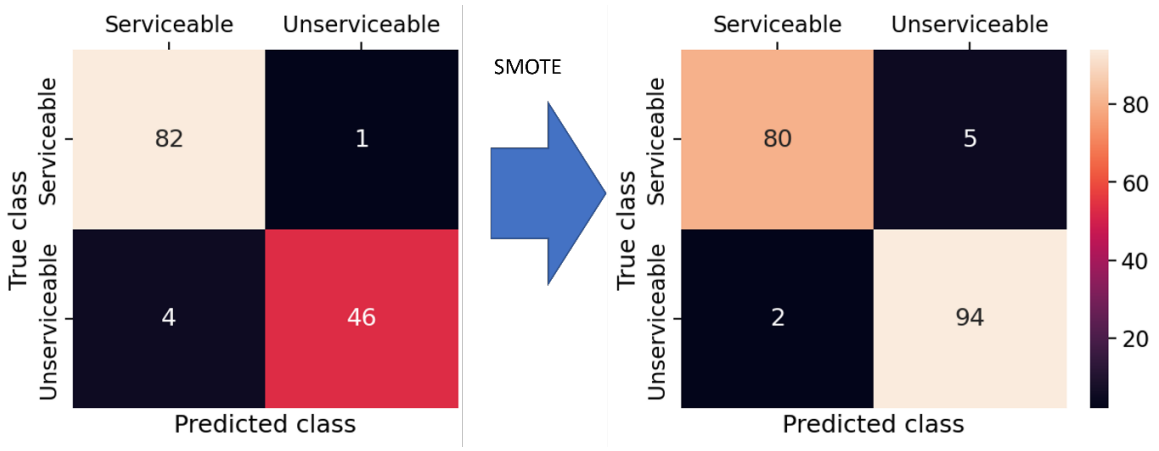


Figure 41. Confusion matrix of best performing serviceability model for flexible keel sections on original dataset(left) and augmented synthetic dataset (right)

Figure 42 shows the results of rigid keel and outkeel SL prediction models for various attribute sets trained on the original dataset. The highest F1 score for rigid keel model was 1.0 using a Random Forest model with SCI, anti-SCI and RRI as attributes. The highest performing SL model for the rigid outkeel had an F1 score of 0.31.

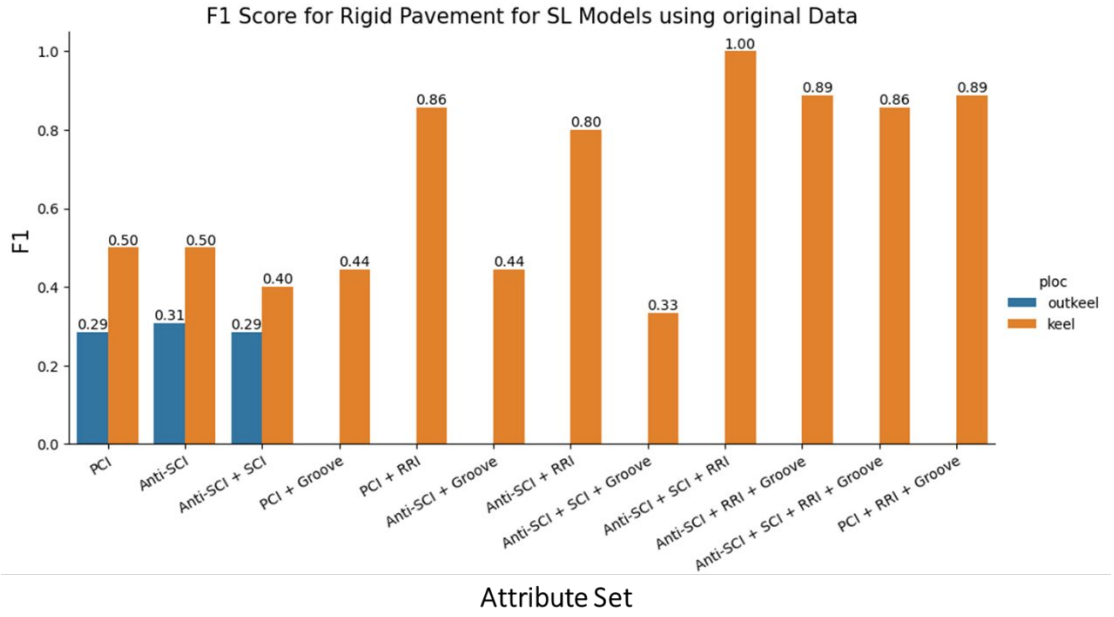


Figure 42. Serviceability level model F1 performance for rigid keel and outkeel sections for different attribute sets

The left-hand side of Figure 43 illustrates the confusion matrix for the rigid keel SL model trained on the original dataset. The diagonal contains the correctly classified data samples (84 total) while the anti-diagonal has none. The imbalance on the diagonal (80 vs. 4) makes the high

F1 performance unreliable for predicting serviceability. This imbalance between the serviceable and unserviceable records stems from the scarcity of rehabilitation records in rigid pavements. In rigid pavements, rehabilitation efforts often consist of isolated slab replacements, which frequently remain undocumented in airport PMS. Moreover, due to the greater durability and longer service life of rigid pavements, the need for rehabilitation is less warranted when compared to flexible pavements.

The right-hand side of Figure 43 is the confusion matrix of the same model with the same attributes but trained using the original data plus class-balancing synthetic data from SMOTE. The additional number of unserviceable classes in both the correctly and incorrectly classified predictions resulted in a reduced overall F1 performance of 0.96. While the score remained high, the additional synthetic samples did not significantly enhance the reliability of the performance estimate, contrary to what was observed in the flexible keel SL model. Given that the minority class comprises only 4 samples, it is challenging for any synthetic samples to deviate significantly from existing ones. Thus, the effectiveness of synthetic data augmentation in this scenario is limited.

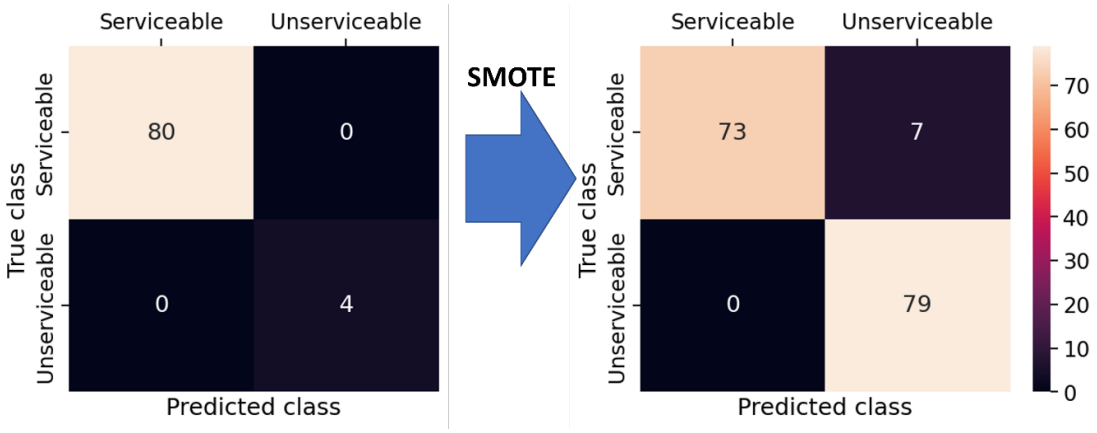


Figure 43. Confusion matrix of best performing SL model for rigid keel sections on original dataset (left) and augmented synthetic dataset (right)

To provide a more practical representation of the SL models, the SL is expressed as the probability that a section is “unserviceable.” The receiver operating characteristic (ROC) curve is used to determine the classification threshold of the SL models based on this probability. The ROC curve provides a graphical representation for assessing the performance of binary classification models. By selecting the pareto optimal point, which represents the most concave point on the curve, a balanced trade-off between the true positive rate (TPR), also known as

sensitivity, and the false positive rate (FPR), also known as specificity, is achieved across various threshold values. This resulted in a decision threshold of 0.30 for the flexible keel model and 0.42 for the rigid keel model. Figure 44 and Figure 45 depict the ROC curve and the probability distribution of SL predictions for the flexible keel and rigid keel models, respectively. The area under the curve (AUC) is 0.99 for flexible keel models, indicating acceptable performance for the models given that a perfect model would result in area under the curve equal to 1.00. The AUC for the rigid keel model was 1.0, suggesting potential overfitting. Another way of expressing these findings is that, when the SL probability is below 0.30 (flexible keel), or 0.42 (rigid keel), the model classifies as "serviceable," and when it is above these thresholds, the model classifies as "unserviceable." The original SL model, which solely used pavement age as a predictor (Ashtiani, Shirazi, Murrell, Spier, & Brill, 2019), had a decision threshold of 0.23 for the flexible pavement, resulting in 88% and 80% correct classifications for serviceable and unserviceable pavements, respectively. The AUC for that model was 0.89, so the current model is a significant improvement for flexible pavements.

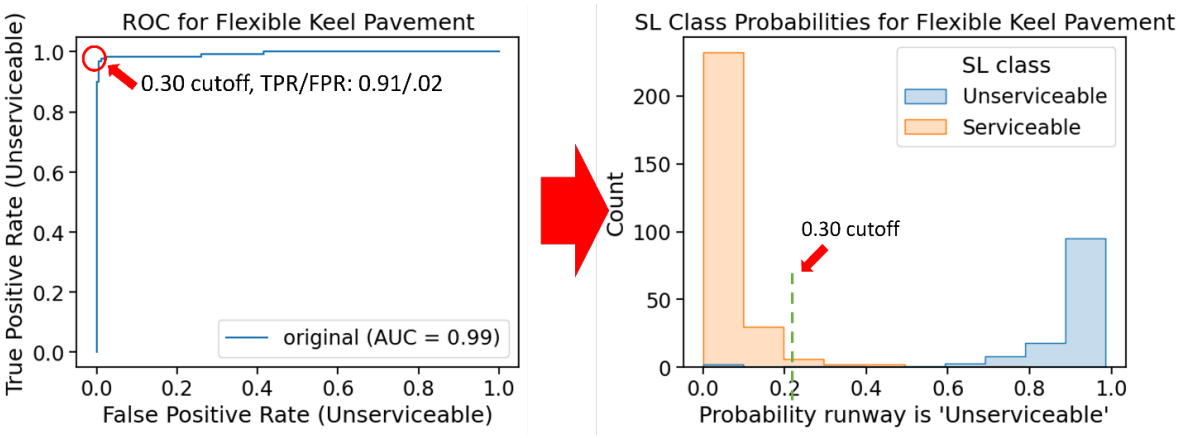


Figure 44. Establishing trade-off between TPR and FPR for flexible keel serviceability level model

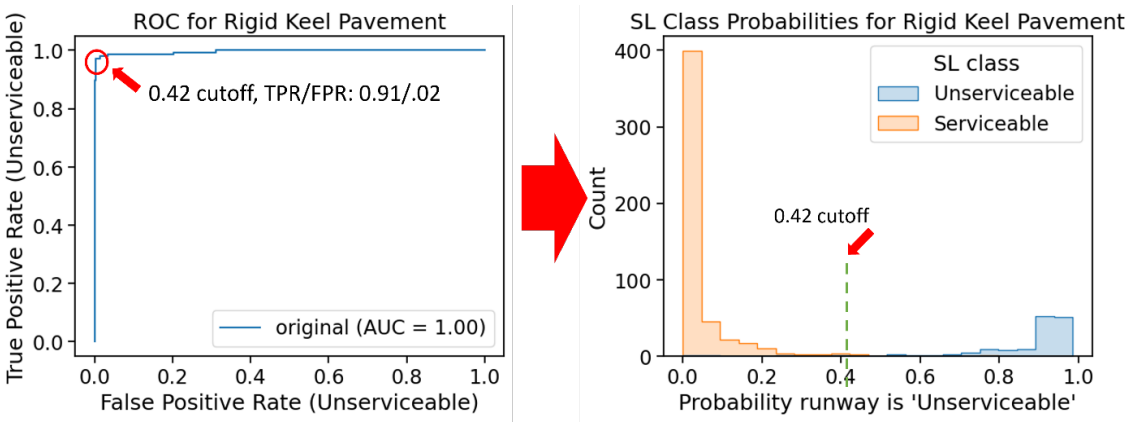


Figure 45. Establishing trade-off between TPR and FPR for rigid keel SL model

Figure 46 is a plot of the probability values assigned to each prediction of the flexible keel section as a function of pavement age. Probability refers to the likelihood of a section being unserviceable. The plot also indicates which samples were correctly or incorrectly classified. For all sections with pavement age of less than 6 years, and 79 percent of the sections with pavement age less than 12 years, the probability of being unserviceable was below the threshold (< 0.3). Additionally, 61 percent of sections more than 12 years old had greater than 80 percent probability of being “unserviceable.” While these classifications are generally reasonable, there are some instances that may not align with common expectations. For example, many sections less than 10 years old have a greater than 80 percent probability of being unserviceable. Conversely, there are some older pavement sections (>15 years) that fall below the unserviceable threshold. This disparity in classifications can be attributed to the fact that this plot shows the SL values solely as a function of pavement age, without directly showing the effect of environmental conditions and traffic level. Another reason is that in the training dataset for the SL model, the decision that classified the pavement into serviceable and unserviceable categories was essentially based on the implementation of major rehabilitation undertaken on the pavement. However, there are factors other than the pavement conditions that may trigger the decision to implement a major rehabilitation. For example, pavements categorized as unserviceable may receive overlays simply because the budget is available, or a change in airport operations may necessitate a thicker pavement. It is also possible that a pavement section in serviceable condition may be rehabilitated because other sections within the runway require treatment. Additionally, some sections in the database marked as serviceable may be in dire substandard condition, as indicated by their surveyed indices, but are kept in service due to a lack of budget for rehabilitation.

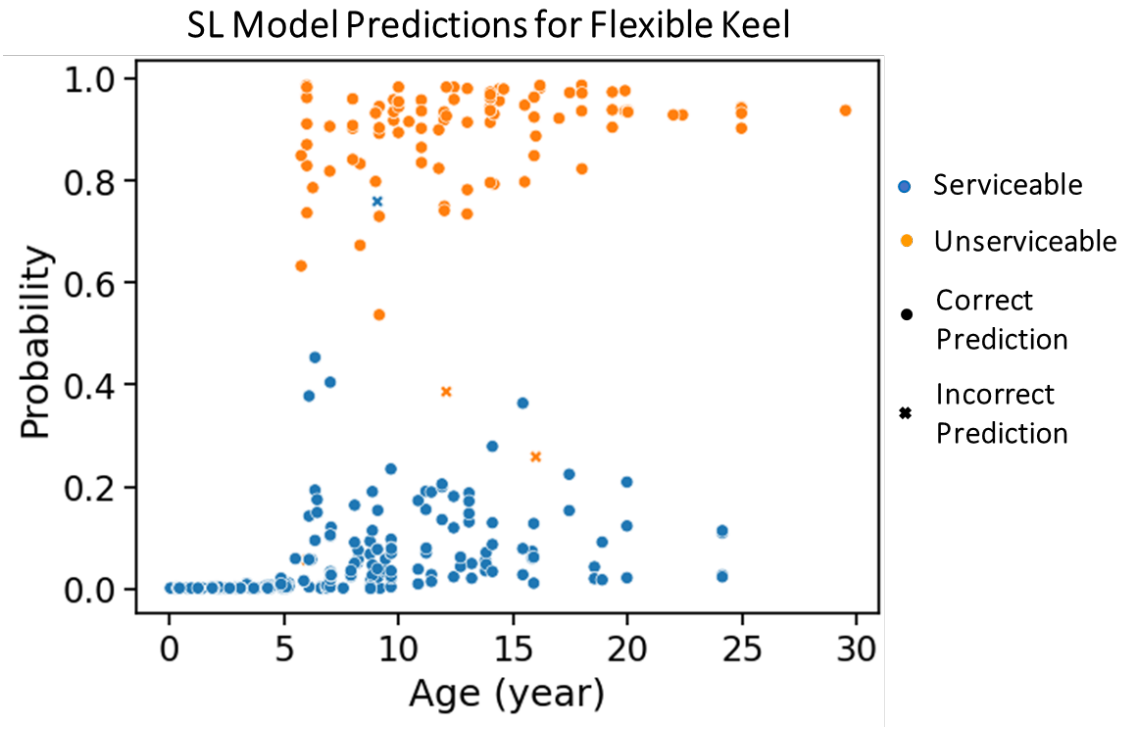


Figure 46. Serviceability level probabilities of being unserviceable for flexible keel sections as a function of pavement age

7 Implementation of ML models as .NET library

The models developed for the pavement condition indexes and SL are designed for integration into the FAA's pavement design and management programs, such as PAVEAIR, PA40 and FAARFIELD, which are built on the .NET Framework. To achieve this integration, the ML models were implemented into a Visual Basic (VB) library that can be referenced by .NET applications.

The management of trained models and performance results was facilitated through ML-Flow (www.mlflow.org), a platform for model versioning and persistence. Trained models are stored in PMML (Predictive Model Markup Language) for its platform independence, as well as in joblib (native Python) formats. The choice of joblib format was made to avoid licensing charges associated with .NET PMML libraries, such as Syncfusion (https://www.syncfusion.com/products/data-science/predictive-analytics). Integration of joblib models with .NET is achieved through Python.NET, which facilitates the integration of the CPython engine within a .NET runtime environment. This capability enables the seamless functioning of common Python packages, some of which are essential for running the models

developed in Python. Python.NET also integrates smoothly with VB, allowing for the easy reference of variables and functions between these two distinct programming languages. The use of Python.NET requires only the installation of the Python programming language on the FAA install location, with a single Python dependency downloaded (scikit-learn). Figure 47 shows a flowchart of model integration process.

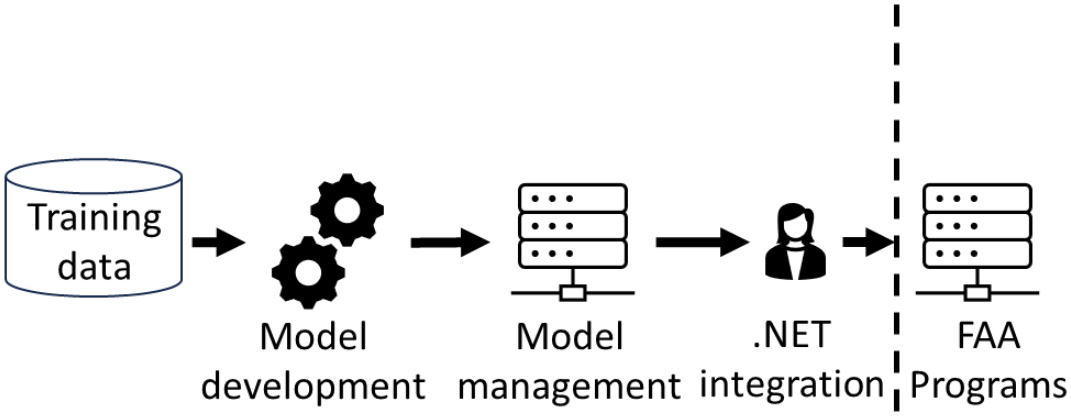


Figure 47. Process of model implementation

Researchers created a VB code that facilitates user interaction by allowing users to input features and select from various models, such as the PCI prediction model. Upon startup, the application checks for the presence of required Python packages. To execute the application, users must have Python 3.8 installed, and any additional necessary packages will be installed automatically during startup. Upon running the VB code, it initializes variables that will later be utilized by the Python code based on the selected model within the application. Subsequently, the VB code employs Python.NET to retrieve the Python output variable and converts it back to a format usable within other VB.NET applications. Finally, the VB code displays the output to the user. The VB library consists of six classes:

1. InputFeatures: This class organizes the pavement, weather, and traffic input data for the ML models related to five pavement condition indexes: anti-SCI, PCI, SCI, RRI, and Groove Index.
2. IndexPredictModel: Using InputFeatures as arguments, this class predicts anti-SCI, PCI, SCI, RRI, and Groove Index for both flexible and rigid pavements.
3. InferenceResult: This class provides a single value of predicted index based on the input data.
4. Attributes: Designed to organize the input data for the SL model.

5. SLPredictModel: Using Attributes as input arguments, this class predicts SL for both flexible and rigid pavements.
6. InferenceSLResult: Returns a single value of predicted SL based on the input data.

Objects, methods, and properties of each class in the compiled VB library are transparently available to .NET programmers in a .NET environment. Appendix B provides a schema of library classes.

8 Summary of findings

This research focused on developing models to predict the airport pavement SL index using data from the FAA's PA40 database. SL is a measure that assesses the suitability of pavement for aircraft use based on its condition. In response to the FAA's EAPL study, which aims to identify factors that prolong the serviceable life of airport pavement, the definition of pavement serviceable life has been expanded to include both structural and non-structural (or functional) failures triggered by factors like FOD potential, high roughness, and reduced skid resistance. The EAPL program has collected comprehensive airfield pavement data from 28 runways at 22 major airports in the continental U.S. This dataset includes various pavement performance measures, such as surface groove geometry, profile roughness, and surface distresses, as well as maintenance work histories, runway usage, and weather data. Some data are provided by the airports, while other data are collected by the FAA during field visits.

The initial predictive models for SL and other pavement condition indexes relied on regression analysis with pavement age as the sole explanatory variable. However, these age-based models did not capture all variations affecting pavement performance, as they overlooked factors such as environmental conditions and traffic levels. To address this limitation, this study employed ML techniques to enhance the accuracy of SL models by incorporating other variables that have the potential to affect pavement performance.

A comprehensive review of the PA40 database, encompassing data from 14 flexible pavement runways and 13 rigid pavement runways, identified the key variables influencing pavement longevity and condition indexes associated with pavement failure. Each runway is typically segmented transversely and longitudinally into sections. Five pavement performance indexes are calculated for each section of each runway. Distress survey data are used to calculate PCI, SCI, and anti-SCI. Longitudinal profile data from the SurPro device are used to calculate RRI. FAA inertial profiler data is used for to calculate the Groove Index. While distress survey data are readily available in the PA40 database, roughness and groove data are only available for runways

visited by the FAA. The availability of PCI data varies among runways based on factors such as pavement age and the completeness of each airport's PMS database.

To develop the ML models for the pavement condition indexes, researchers identified five feature sets with the potential to explain variations in pavement performance: number of major rehabilitations, pavement functional age (defined as the time since construction or the last major rehabilitation), weather variables, traffic variables, and location of runway sections with respect to runway ends. Researchers considered twelve weather variables with the potential to affect pavement performance. For each inspection record in the database, these variables were calculated as annual averages since (a) the last inspection or, (b) if there was no prior inspection, since the last rehabilitation/reconstruction. Aircraft were classified into standard generic gear type categories, S-75, D-200, 2D-400, B777-300, B747-400, and A380-800, to standardize the traffic features. The intent was to determine whether the runway is under-designed or overdesigned by comparing the converted design and actual traffic to these generic groups. However, this comparison is not possible for all runways due to the lack of design traffic data for almost half of them. Consequently, the equivalent departures of generic aircraft groups D-200, 2D-400, and the summation of A380-800, B747-400, and B777-300 (defined as Heavy Aircraft) were considered as traffic features. Since traffic data are only available for limited years and do not cover the same period as other performance data, average annual values for the available years were used as model inputs.

Separately for flexible and rigid pavements, researchers developed ML models to predict critical pavement performance indexes within the SL model, specifically PCI, SCI, anti-SCI, RRI, and GI. An MST approach was employed for designing ML models. The MST involves implementing various feature selection methods for weather features to assess their variability and identify those that most significantly influence pavement performance. Density-based clustering grouped runway performance data samples based on similarities between weather features. Linear and gradient-boosting models were employed for model training, using a 4-fold cross-validation approach.

8.1 Key findings: PCI, SCI and anti-SCI models

Time-series (autoregressive) models are used for predicting PCI, SCI, and anti-SCI. These models consider previous inspections and the time between two successive inspections as predictors. The developed models incorporate various feature sets, including pavement age, previous values of the index, number of previous rehabilitation projects, section location with respect to runway end, weather variables (all or selected), and traffic variables. UMAP-

embedded features, which reduced the twelve weather features to two, were also considered. Separate models were developed for the keel and outkeel sections. Key findings of the developed models are:

- Models incorporating only pavement age exhibit the lowest performance. Performance improves when previous index values are included.
- Incorporating weather variables improves model performance compared to solely age-based models.
- For flexible keel pavement PCI models, a trained CatBoost model achieves the highest performance, with an RMSE of 2.3 using age, previous measurements, selected weather variables, and traffic feature sets.
- For rigid keel pavement PCI models, a trained CatBoost model also demonstrates the highest performance, with an RMSE of 2.5 using the age, previous measurements, and selected weather variables.
- The anti-SCI models exhibit generally higher errors for rigid pavement compared to flexible pavement models.
- While traffic variables have no effect on flexible pavement models' performance metrics, they do enhance the performance of the SCI model for rigid pavement.
- UMAP embedded features slightly improve performance, but not to a degree significant enough to justify their lack of interpretability.
- Performance is sensitive to the time horizon considered. Performance remains consistent until 15 years for both flexible and rigid models, but the error increases beyond this point for rigid models, suggesting caution is warranted when using the model for predicting performance beyond 15 years.

8.2 Key findings: RRI and GI models

A continuous function approximation method was used for the RRI and GI models, as data for more than one measurement were not available for most of the runways. With significantly fewer data samples for RRI (50 for flexible, 46 for rigid) and GI (40 for flexible, 35 for rigid), and the majority lacking a previous measurement, a simple linear regression method was applied due to the small dataset size. Key findings of the developed models are as follows:

- For most runway sections, RRI remains below the unacceptable threshold (0.35), indicating that runway roughness is not a critical factor leading to pavement unserviceability.
- The model based solely on pavement age performs best for the flexible pavement RRI, indicating that weather and traffic variables have no significant influence on RRI. However, incorporating these variables decreases error for rigid pavements.
- For the flexible pavement GI model, incorporating weather variables slightly reduces the error. For the rigid pavement model, including traffic features reduces the error.
- Both the RRI and GI models are prone to overfitting, with the trained model tending to memorize the training data due to the low variability in the test data caused by the small number of samples.
- Regularization across features does not significantly improve performance or exhibit any discernible trend for the RRI models.
- In contrast, for the GI model, regularization suggests that as the weight assigned to weather variables increases, the RMSE in both flexible and rigid models gradually decreases. This finding implies that an increase in data samples may lead to a more robust and accurate GI prediction model.

8.3 Key findings: Developed SL models

The SL model is structured as a classification task, i.e., it determines whether a runway pavement is serviceable given a set of pavement condition indexes. A labeled dataset was created for SL model training, where each data point consists of a set of attributes (PCI, SCI, anti-SCI, RRI, and GI) and a corresponding class (serviceable or unserviceable). Instances with rehabilitation/reconstruction records were classified as "Unserviceable," while those with inspection records were labeled as "Serviceable". Since pavement condition indexes (attributes) were typically unavailable at the time of runway rehabilitation, their values were estimated using the individual ML models. To counter class imbalance and ensure model performance validity, synthetic unserviceable data samples were generated using SMOTE and ADA methods. Employing the MST approach, SL models were developed by incorporating a combination of the five attributes. Models were trained on the original dataset and the original dataset plus the synthetic database for validation. Key findings from the developed SL models are as follows:

- The highest F1 score for the flexible keel SL model is 0.94, incorporating SCI, anti-SCI, RRI, and GI as attributes and using RF as the classification method. Even with the

incorporation of synthetic data, the F1 score remains at 0.91, affirming the model's reliability.

- Synthetic data effectively addresses class imbalance in flexible pavement samples.
- Class imbalance in rigid pavements persists due to the scarcity of rehabilitation records, rendering synthetic data insufficient to supplement unserviceable samples adequately.
- The rigid keel model achieves a high F1 score of 0.94; however, due to significant class imbalance, this score is deemed unreliable.
- ROC curve analysis establishes cut-off points of 0.23 for the flexible pavement SL model and 0.42 for the rigid pavement SL model. Probabilities below these thresholds predict "Serviceable" conditions, while those above predict "Unserviceable" conditions.
- Despite the acceptable F1 score for the flexible SL model, disparities in classification exist. This is attributed to the fact that other factors not captured by pavement index data, such as budget availability, may influence major rehabilitation decisions. Additionally, entire runway overlays sometimes occur for reasons unrelated to existing poor condition, masking sections that may still be in good condition.

9 Recommendations for future research

Based on the findings of this study, the following additional efforts are recommended:

1. Collect more pavement profile data to supplement and complement the RRI and GI data, enhancing the performance of prediction models. It is recommended to collect more data from representative runways in suitable time intervals over the next five to ten years.
2. Obtain design traffic data from all runways evaluated in this study to facilitate the calculation of XSDepartures. This will enable a more accurate analysis of the effect of traffic on pavement performance.
3. Retrain the SL model using a more comprehensive database including additional RRI, G and traffic data.
4. Obtain data from taxiways and develop a dedicated SL model for taxiway pavement.

10 References

Ashtiani, A. Z. (2021). Application of machine learning techniques to pavement performance modeling. *(DOT/FAA/TC-21/44)*. U.S. Department of Transportation, Federal Aviation Administration. Retrieved from https://www.airporttech.tc.faa.gov/Products/Airport-Pavement-Papers-Publications/Airport-Pavement-Detail/application-of-machine-learning-techniques-to-pavement-performance-modeling

Ashtiani, A. Z., Murrell, S., Speir, R., & Brill, D. R. (2022). Machine learning solutions for development of performance deterioration models of flexible airfield pavements. *Eleventh International Conference on the Bearing Capacity of Roads, Railroads, and Airfields.* Trondheim, Norway. doi:10.1201/9781003222910-16

Ashtiani, A. Z., Shirazi, H., Murrell, S., Spier, R., & Brill, D. R. (2019). Performance model development for extended airport pavement life. *International Airfield and Highway Pavements Conference.* Chicago, IL. doi:10.1061/9780784482476.025

ASTM International. (2020). Standard test method for airport pavement condition index surveys. *(ASTM D5340-20)*. West Conshohocken, PA: ASTM International. Retrieved from www.astm.org

Boeing. (2002). *Airport Operations -- Frequently Asked Questions*. Retrieved from http://www.boeing.com/commercial/airports

Borisov, V., Leeman, T., Sessler, K., Haug, J., Pawelczyk, M., & Kasneci, G. (2022). Deep neural networks and tabular data: A survey. *IEEE Transactions on Neural Networks and Learning Systems*, 1–21. doi:10.1109/TNNLS.2022.3229161

Breiman, L. (2001). Random forests. *Machine Learning, 45*, pp. 5–32. doi:10.1023/A:1010933404324

Brill, D. R., & Parsons, T. A. (2017). Development of new design procedures for extended airport pavement life. In I. A.-Q. A. Loizos (Ed.), *10th International Conference on the Bearing Capacity of Roads, Railroads, and Airfields.* Athens, Greece. doi:10.1201/97813151003333

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research, 16*, 321–357. doi:10.1613/iair.953

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning, 20*, 273–297. doi:10.1007/BF00994018

Deza, M., & Deza, E. (2009). *Encyclopedia of distances* (Vol. 2006). Springer. doi:10.1007/978-3-642-00234-2

Eckstein, A., Kurcz, C., & Silva, M. (2012). *Threaded Track: Geospatial Data Fusion for Aircraft Flight Trajectories.* Mitre.

FAA. (1995). Advisory Circular (AC) 150/5320-6D (cancelled). *Airport pavement design and evaluation*. Washington, D.C.: U.S. Department of Transportation, Federal Aviation Administration. Retrieved from https://www.faa.gov/regulations_policies/advisory_circulars/index.cfm/go/document.information/documentID/22105

FAA. (2009). AC 150/5380-9. *Guidelines and procedures for measuring airfield pavement roughness*. Washington, D.C.: U.S. Department of Transportation, Federal Aviation Administration. Retrieved from https://www.faa.gov/regulations_policies/advisory_circulars/index.cfm/go/document.information/documentID/99720

FAA. (2014). AC 150/5380-7B. *Airport pavement management program (PMP)*. Washington, D.C.: U.S. Department of Transportation, Federal Aviation Administration. Retrieved from https://www.faa.gov/airports/resources/advisory_circulars/index.cfm/go/document.current/documentNumber/150_5380-7

FAA. (2016). AC 150/5320-12D. *Measurement and maintenance of skid-resistant airport pavement surfaces*. Washington D.C.: U.S. Department of Transportation, Federal Aviation Administration. Retrieved from https://www.faa.gov/regulations_policies/advisory_circulars/index.cfm/go/document.information/documentID/1029469

FAA. (2016b). ProGroove computer program. Retrieved from https://www.airporttech.tc.faa.gov/Products/Airport-Safety-Papers-Publications/Airport-Safety-Detail/progroove-40

FAA. (2021). AC 150/5320-6G. *Airport pavement design and evaluation*. Washington, D.C.: U.S. Department of Transportation, Federal Aviation Administration. Retrieved from https://www.faa.gov/airports/resources/advisory_circulars/index.cfm/go/document.current/documentNumber/150_5320-6

FAA. (2023, December 22). FAARFIELD 2.1.1. Retrieved from https://www.airporttech.tc.faa.gov/Products/Airport-Safety-Papers-Publications/Airport-Safety-Detail/ArtMID/3682/ArticleID/2841/FAARFIELD-20

FAA. (2024, June 19). FAA PAVEAIR 3.7.4. Retrieved from https://www.airporttech.tc.faa.gov/Products/Airport-Pavement-Papers-Publications/Airport-Pavement-Detail/faa-paveair-374

He, H., Bai, Y., Garcia, E. A., & Li, S. (2008). ADASYN: Adaptive synthetic sampling approach for imbalanced learning. *2008 IEEE Proceedings of the International Joint Conference on Neural Networks* (pp. 1322–1328). IEEE. doi:10.1109/IJCNN.2008.4633969

Karunarathne, A., Gad, E. F., Disfani, M. M., Sivanerupan, S., & Wilson, J. L. (2016, March). Review of calculation procedures of Thornthwaite Moisture Index and its impact on footing design. *Australian Geomechanics Journal, 51*(1), 85-95. Retrieved from https://geomechanics.org.au/papers/review-of-calculation-procedures-of-thornthwaite-moisture-index-and-its-impact-on-footing-design/

Kumar, A., McCann, R., Naughton, J., & Patel, J. M. (2016). Model selection management systems: the next frontier of advanced analytics. *ACM SIGMOD Record, 44*(4), 17–22. doi:10.1145/2935694.2935698

Kuncas, A. (2021). New index testing and verification—Runway roughness index. *(DOT/FAA/TC-21/32)*. Washington, D.C.: U.S. Department of Transportation, Federal Aviation Administration. Retrieved from https://www.airporttech.tc.faa.gov/products/airport-pavement-papers-publications/airport-pavement-detail/new-index-testing-and-verificationrunway-roughness-index

Lundberg, S., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In U. vonLuxburg, I. Guyon, S. Bengio, H. Wallach, & R. Fergus (Ed.), *NIPS '17: Proceedings of the 31st International Conference on Neural Information Processing Systems* (pp. 4768–4777). Red Hook, NY: Curran Associates. doi:10.5555/32952222.32955230

McInnes, L., Healy, J., & Astels, S. (2017). hdbscan: Hierarchal density based clustering. *Journal of Open Source Software, 2*(11), 205. doi:10.21105/joss.00205

McInnes, L., Healy, J., & Melville, J. (2018). UMAP: Uniform manifold approximation and projection for dimension reduction. Retrieved from http://arxiv.org/abs/1802.03426

Mehta, P. K., & Monteiro, P. J. (2006). *Concrete: microstructure, properties, and materials* (3rd ed.). McGraw-Hill.

National Oceanic and Atmospheric Administration, National Centers for Environmental Information & U.S. Air Force-14th Weather Squadron. (2018, January 12). Federal

climate complex data documentation for integrated surface data (ISD). Asheville, NC. Retrieved from https://www.ncei.noaa.gov/data/global-hourly/doc/isd-format-document.pdf

New York City Panel on Climate Change. (2013, June). Climate Risk Information 2013: Observations, Climate Change Projections, and Maps. (C. Rosenzweig, & W. Solecki, Eds.) New York, New York: City of New York Special Initiative on Rebuilding and Resiliency. Retrieved from https://www.nyc.gov/html/planyc2030/downloads/pdf/npcc_climate_risk_information_2013_report.pdf

Philp, M., & Taylor, M. (2012). *Beyond Agriculture: Exploring the application of the Thornthwaite Moisture Index to infrastructure and possibilities for climate change adaptation.* University of South Australia, Barbara Hardy Institute. Australian Climate Change Adaptation Research Network for Settlements and Infrastructure. Retrieved from https://www.unsw.edu.au/content/dam/pdfs/engineering/civil-environmental/water-research-laboratory/accarnsi/beyond-agriculture---exploring-the-application-of-the-thornthwaie-moisture-index.pdf

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018, December 3–8). CatBoost: Unbiased boosting with categorical features. In S. Bengio, K. Grauman, H. Wallach, N. Cesa-Bianchi, H. Larochella, & R. Garnett (Ed.), *32nd Conference on Neural Information Processing Systems*, *31*, pp. 6638–6649. Montreal, Canada. doi:10.5555/3327757.3327770

Thornthwaite, C. W. (1948). An approach twoard a rational classification of climate. *Geographical Review*(38), 55-94.

Wang, C., Wu, Q., Weimer, M., & Zhu, E. (2021). Flaml: A fast and lightweight automl library. *Proceedings of Machine Learning Systems, 3*, 434–447.

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B: Statistical Methodology, 67*(2), 301–320. doi:10.1111/j.1467-9868.2005.00527.x

A Development of historical climate and weather data links in PA40

A.1 Introduction

This appendix presents the content of an unpublished technical note authored by Russell Gorman, Scott Murrell, and Timothy Parsons of Applied Research Associates, Inc. The technical note describes the research effort that determined how weather events were summarized and implemented in the software tool, PA40.

Background

The Federal Aviation Administration (FAA) Office of Airport Safety and Standards (AAS) has requested that methodologies be developed to extend the expected life of large hub runway construction from the current standard of 20 to 40 years. The development of life extension methodologies is being executed using a three-pronged approach consisting of: (a) comprehensive data collection for a selection of runways at large- and medium-hub U.S. airports, including construction, performance, material property, traffic and environmental data; (b) the development of advanced performance models based on the collected data relating various performance measures to traffic, structural, and environmental inputs; and (c) the development of a new definition of airport pavement life that is consistent with the assumptions of life cycle cost analysis (LCCA), e.g., that the end of pavement life occurs when the annual cost of maintaining continued service at a safe operational level exceeds the annualized cost of replacement.

Data collection in support of extended airport pavement life began in 2012. To date, data have been collected on 28 runways at 22 U.S. airports. Field visits were made to a subset of these airports and additional field data collected on 13 of these runways. To facilitate analysis of the collected data for the development of advanced performance models the FAA has developed a project specific software tool—PA40.

PA40 is a standalone implementation of the FAA PAVEAIR pavement management system. The PA40 data warehouse contains tables and functions not currently implemented in the public version of FAA PAVEAIR. PA40 is not accessible through the FAA PAVEAIR website. To facilitate the future study of the influence of climate/weather on the long-term performance of airfield pavements, the FAA directed that new remote database access and calculation functionality be added to PA40.

Scope

This document details the weather and climate data search functions in PA40, including the justifications for and planned use of the particular data elements selected and the source data for each element.

The FAA databases contain National Oceanic and Atmospheric Administration (NOAA) maintained Automated Weather Observing System (AWOS) and Automated Surface Observing Systems (ASOS) information. The scope of this research includes working with these data sets from within PA40. Specific required weather data include temperature, wind speed, precipitation, relative humidity (or dew point), solar radiation (or sky cover). PA40 was also modified to compute derived weather fields from the basic data. The weather data were considered network-level (airport-level) data in PA40 and made available as hourly readings for each airport in the database. Climate is distinct from weather and refers to long-term environmental trends expressed by statistics, such as average annual precipitation or average annual high/low temperatures. The scope also includes creating new functions as needed within the PA40 graphical user interface to query, return, and display the weather and climate data for all airports in the study.

A.2 Approach

First, climate/weather data elements to link to PA40 were assessed to determine which derived climate/weather fields should be computed and displayed. Then, the available climate/weather information in the FAA databases was determined to communicate what was needed and the format and to assist the FAA in the development of the web service. The FAA climate/weather databases contain NOAA reported data measured by AWOS and ASOS at the Extended Airport Pavement Life Study (EAPL) airports. The solar radiation data originates from the National Solar Radiation Database maintained by the National Renewable Energy Laboratory (NREL) of the Department of Energy (DOE). While the web service was being implemented by others, the web page was developed, and server-side processing calculations were coded. The relationship of the three main components of the Climate/Weather Link are shown in Figure A- 1.

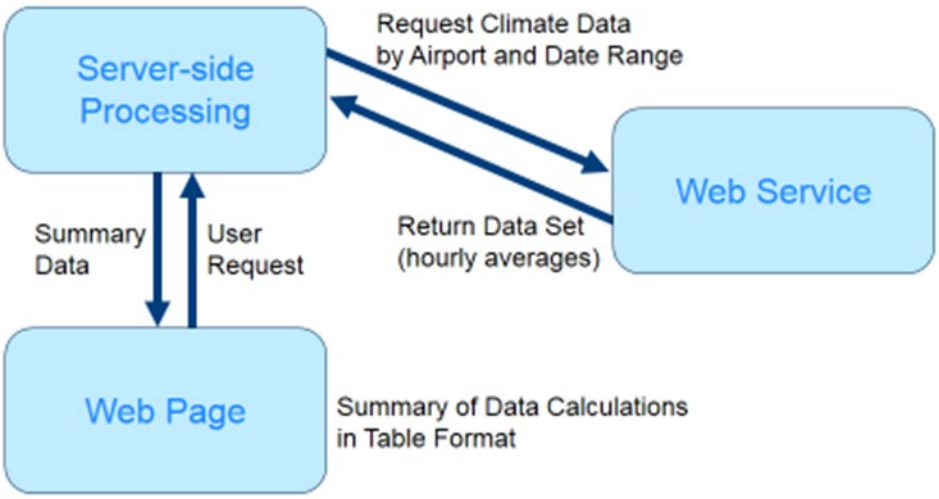


Figure A- 1. Climate weather link components and their relationships

The user can select a particular site and date/time range on the web page user interface. Then, the data for that date/time range and site are retrieved from the web service. Those data are then received by the server-side processing component, and the report calculations are processed. After all processing is complete, the results are fed back to the web page and displayed to the user in tabular format.

The following sections describe the calculations performed by the server-side software component on the raw data that result in the report that is presented to the user.

High temperature (degrees Fahrenheit)

Each hourly average temperature received by the web service is reviewed. The highest hourly average temperature in the data is retrieved.

Low temperature (degrees Fahrenheit)

Each hourly average temperature received by the web service is reviewed. The lowest hourly average temperature in the data is retrieved.

Average temperature (degrees Fahrenheit)

Each hourly average temperature received by the web service is stored. The totaled result is used to calculate the average hourly average temperature in the entire data set (i.e., every hourly average temperature contained in the user selected date/time range).

Average daily temperature range (degrees Fahrenheit)

The temperature range is calculated daily by calculating the difference between the high and low values found each day. The average of all daily temperature ranges is then calculated and reported to the user in degrees Fahrenheit.

Number of freeze thaw cycles (number of cycles)

Each hourly average temperature received by the web service is reviewed. The number of times the temperature goes from a value of less than 32 degrees Fahrenheit to a value greater than 32 degrees Fahrenheit is recorded.

Freeze degree days (degrees Fahrenheit)

Each hourly average temperature received by the web service is reviewed and the average daily temperature is calculated for each day that falls within the date/time range selected by the user. Any day where the average daily temperature is less than 32° F is considered a freeze degree day.

For each freeze degree day that is found, an overall running total of the number of degrees the daily average temperature of each freeze degree day falls below 32° F is maintained. The value of the running total of these temperature deltas is the freeze degree days value that is presented to the user.

Days temperature over 90 degrees Fahrenheit (days)

For the contiguous United States, temperatures above 90° F are typically considered hot, and NOAA and other agencies keep track of the occurrences across the U.S. The New York City Panel on Climate Change defined an extreme temperature event as a day when the temperature reached 90° F or hotter (New York City Panel on Climate Change, 2013). Higher than normal temperatures are known to have a negative impact on pavement performance, which warrants further study under the EAPL. Although 90° F does not have the same inherent physical significance as 32° F (the freezing point of water), it nevertheless serves as a convenient high-temperature threshold.

This calculation is similar to the Freeze-Degree-Day calculation, except degree-days above 90° F are summed instead of degree-days below freezing. Each hourly average temperature received by the web service is reviewed. The number of days in which the temperature exceeded 90° F is recorded.

Days precipitation (days)

Each hourly average precipitation amount received by the web service is reviewed. The number of days in which the total daily precipitation exceeded 0 in. is recorded.

Duration precipitation (hours)

Each hourly average precipitation amount received by the web service is reviewed. Using this data, each precipitation event will be calculated (in hours) and recorded. A precipitation event starts when precipitation is first detected (i.e., precipitation value is greater than 0 in.) in a given hour of data. The precipitation event continues as long as precipitation is also detected in the next contiguous hour or hours. The precipitation event stops upon the first contiguous hour in which there is no precipitation recorded (i.e., precipitation value equals 0 in.). A single precipitation event may span one or more days as long as the consecutive hours all have experienced precipitation.

The number of precipitation events is recorded, and the average precipitation event duration in hours is calculated.

Precipitation total (inches)

Each hourly average precipitation amount received by the web service is summed. The total number of inches of precipitation is recorded.

Freeze precipitation days (days)

Each hourly average precipitation amount and each hourly average temperature received by the web service is reviewed. If at least one average hourly temperature in a particular day is below 32° F and at least one average hourly precipitation value in the same day is greater than 0 in., then this is considered a freeze-precipitation day. Note that the “freeze hour” and the “precipitation hours” are not necessarily the same hour. That is, one hour can contain precipitation, and a different hour can have an average freezing temperature. However, it is also acceptable if they are the same hour. The total count of days meeting these criteria is reported.

Wind speed average (miles per hour)

Each hourly average wind speed value received by the web service is summed. The totaled result is used to calculate the average hourly wind speed in the entire data set (i.e., every hourly average wind speed contained in the user selected date/time range).

Sky cover (oktas)

Each hourly average sky cover value received by the web service is summed. The totaled result is used to calculate the average hourly sky cover in the entire data set (i.e., every hourly average sky cover contained in the user selected date/time range).

Table A- 1 shows the list of possible integer values for sky coverage layer code provided by the web service and each corresponding text description. Values and meanings are defined in the document written by NOAA/National Centers for Environmental Information, titled "Federal Climate Complex Data Documentation for Integrated Source Data," dated January 12, 2018 (National Oceanic and Atmospheric Administration, National Centers for Environmental Information & U.S. Air Force-14th Weather Squadron, 2018).

Table A- 1. Sky coverage layer codes

Sky Coverage Code	Description	Amount of Sky Covered (oktas)
0	CLR (clear)—no obscuration	0/0
1-2	FEW	$0<X \leq 2$
3-4	SCT (scattered)	$2< X \leq 4$
4-7	BKN (broken)	$4< X \leq 8$
8	OVC (overcast)	8
9	Obscured	8
10	Partial obscuration	NA

Thornthwaite Moisture Index

The Thornthwaite Moisture Index (Thornthwaite, 1948) (TMI) is a measure of moisture available in the soil. It can be generally described as reflecting the aridity or humidity of the soil and climate, calculated from the collective effects of precipitation, evapotranspiration, soil water storage, moisture deficit and run off. Although it was developed as an agricultural tool, it has also been used for engineering applications in which soil moisture is a consideration, including foundation design and pavement performance (Philp & Taylor, 2012). In TMI calculations, precipitation and evapotranspiration are equally important climatic factors. Thornthwaite described the climatic cycle by balancing the rainfall, potential evapotranspiration, and soil water holding capacity. Computations of the TMI are based upon monthly rainfall, mean temperature

data, an estimate of the water holding capacity of the soil and the site location. Moist climates have positive values of TI and dry climates have negative values. Table A- 2 shows the climate types corresponding to TMI ranges.

Table A- 2. Climate types corresponding to TMI ranges

Thornthwaite Climate Type		**Thornthwaite Moisture Index (%)**
A	Perhumid	>100
B4	Humid	80 to 100
B3	Humid	60 to 80
B2	Humid	40 to 60
B1	Humid	20 to 40
C2	Moist subhumid	0 to 20
C1	Dry subhumid	-20 to 0
D	Semi-arid	-40 to -20
E	Arid	-60 to -40

The publication titled "Review of Calculation Procedures of Thornthwaite Moisture Index and its Impact on Footing Design" (Karunarathne, Gad, Disfani, Sivanerupan, & Wilson, 2016) was used as a guide for the approach on how to calculate the TMI. Method 2 documented in this paper was used as the basis of the calculations and is described below.

There are 3 major steps in this calculation:

Step 1:

Cycle through each day's worth of data and calculate the following:

- total precipitation per month (P)
- the average temperature per month (t)
- the monthly heat index value (i)

$$i = (0.2 \times t)^{1.514}$$

Step 2:

Calculate the following:

$$a = 6.75 \times 10^{-7} \times I^3 - 7.771 \times 10^{-5} \times I^2 + 0.01792 \times I + 0.49239$$

- Determine the Total Heat Index (I) by calculating the summation of all monthly heat index values that were calculated in Step 1.
- Determine the value of (a), which will be used in upcoming calculations in Step 3.

Step 3:

Cycle through the monthly calculations from Step 1 and use the values calculated in Step 2 to calculate the following below:

- Calculate potential evaporation (PE) for each month (assume 12-hour daylight each day)

$$PE = 1.6 \times (\frac{10 \times t}{1})^a$$

- Calculate soil moisture storage (S)

Note that P is the total monthly precipitation.

Assume that the initial soil moisture value and the maximum soil moisture store are both 10cm.

S = MIN (MAX(Si-1 + (P – PE), 0), Smax)

- Runoff ®

R = MAX((Si-1 + (P – PE) – Smax), 0)

- Moisture Deficit (D)

D = -MIN (Si-1 + (P – PE), 0)

Then, after completing the cycle through each month of data, calculate the following:

- Total Runoff is the summation of all monthly runoff values.
- Total Deficit is the summation of all monthly moisture deficit values.
- Total potential evaporation is the summation of all monthly potential evaporation values.
- Calculate Ia:

$$Ia = 100 \times \frac{D}{PE}$$

- Calculate Ih:

$$I_h = 100 \times \frac{R}{PE}$$

- Caculate TMI:

$$TMI = I_h - 0.6 \times I_a$$

A minimum of 12 months of data is required to calculate TMI. TMI is calculated for each year in the data set and then the annual TMI values are averaged to determine the value reported. When a query is for a number of months not simply divisible by 12, the remaining fraction of a year is disregarded.

Relative humidity high (%)

The following calculation determines the relative humidity. This calculation is used when processing the daily data to determine the overall high and low humidity values and the average humidity.

$$Relative\ Humidity = 100 \times \frac{e^{\left(\frac{17.625 * \text{dewpoint}}{243.04 + \text{dewpoint}}\right)}}{e^{\left(\frac{17.625 * \text{temperature}}{243.04 + \text{temperature}}\right)}}$$

Each hourly average dewpoint and temperature value received by the web service is reviewed. The humidity is calculated for each hour, and the highest value found across the entire data set is reported.

Relative humidity low (%)

Each hourly average dewpoint and temperature value received by the web service is reviewed. The humidity is calculated for each hour, and the lowest value found across the entire data set is reported.

Relative humidity average (%)

Each hourly average dewpoint and temperature value received by the web service is reviewed. The humidity is calculated for each hour, and the average humidity across the entire data set is calculated and reported.

Hydration days (days)

Each hourly average dewpoint and temperature value received by the web service is reviewed. The humidity is calculated for each hour, and then a count of the number of days where at least one hour's relative humidity reading exceeds 80% is determined. The 80% value was selected based on the relative humidity in the pores required for the Portland Cement Hydration reaction to occur, as reported in Mehta and Monteiro (2006).

Solar radiation (watts per meter squared)

Each hourly solar radiation value received by the web service is reviewed. The total radiation for every hour is stored and reported to the user.

A.3 Retrieving and displaying climate/weather data using PA40

The software flow is as follows: first, a user selects a particular site and date/time range on the web page user interface and then selects the "Climate Report" button, as shown in Figure A- 2. This action results in a request sent by the server-side processing component to the web service. The request to the web service contains the site name and the start and end date range. Note that the "Filter" button and field are not functional in the current version of the software but are placeholders for future use.

The following data are returned by the web service for each hour in the requested date/time range:

- The respective hour
- The average temperature for that hour in degrees Fahrenheit
- The average wind speed in miles per hour
- The dew point average in degrees Fahrenheit
- The total precipitation in inches and hundredths of an inch
- The solar radiation average in watts per meter squared
- Average sky cover in oktas

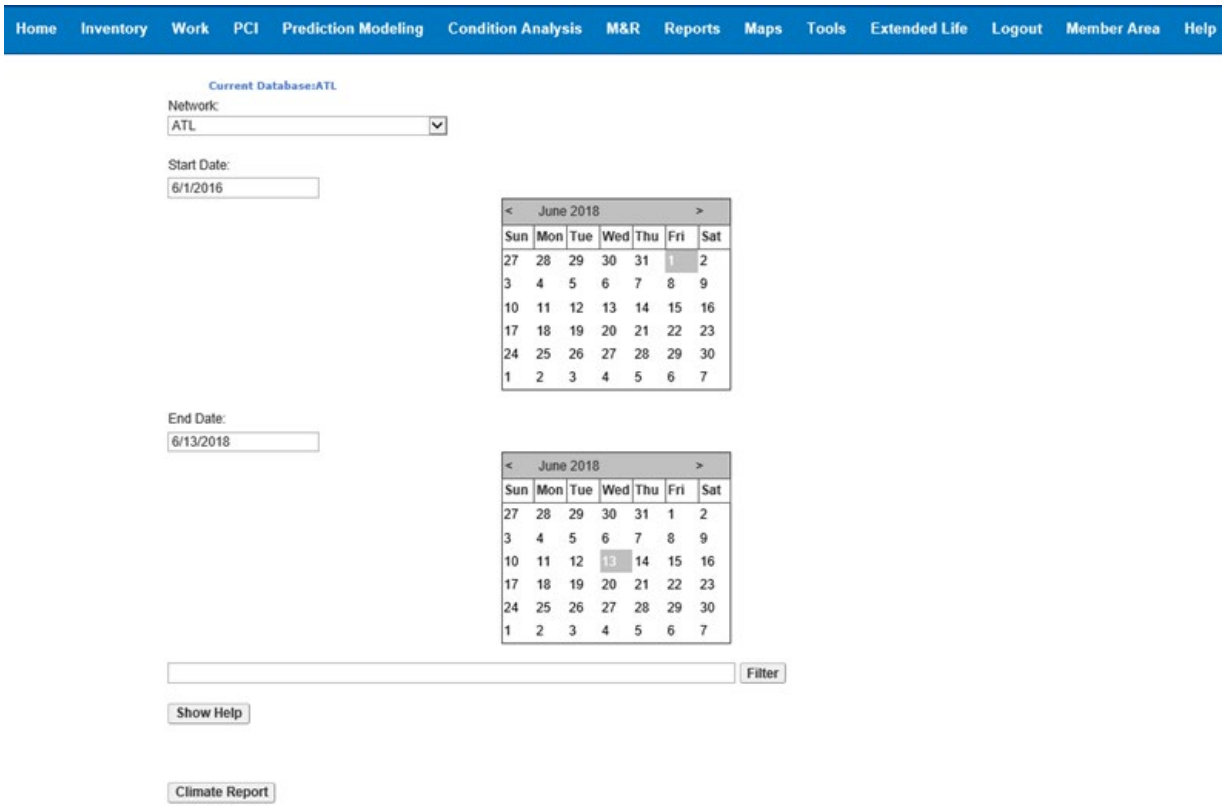


Figure A- 2. Web page user interface example

After the server-side component receives this data back from the web service, it then processes all the data and returns it back to the web page, where it is displayed to the user, as illustrated in Figure A- 3.

Network Name	ATL
Start Date	Friday, January 1, 1999
End Date	Thursday, January 1, 2004
Completeness	36,286 of 61,343 hours contained data (i.e., completeness percentage - 59.15%)
Temp High	97.00 F
Temp Low	8.00 F
Temp Average	61.68 F
Average Daily Temperature Range	17.00 F
Freeze Thaw	109 freeze-thaw cycles
Freeze Degree Days	85.35 F
Days Temperature over 90	47 days
Days Precipitation	462 days
Average Duration of Precipitation Event	3.22 hours
Precipitation Total	274.41 inches
Thornthwaite	65 %
RHumidity High	85.40 %
RHumidity Low	4.98 %
RHumidity Avg	62.78 %
Hydration Days	1025 days
Sky Cover	6.19 oktas
Solar Radiation	12,657,832.00 watts/m^2
Wind Average	8.79 mph
Freeze Precipitation Days	12 days

Figure A- 3. Results table displayed on web page

In addition to the climate/weather data processing results, missing data are reported as expected hours versus actual hours for which data were returned, and the percentage of completeness.

A.4 Planned use of climate and weather data

An approach to determine climate and/or weather's influence on pavement performance is described in the following steps:

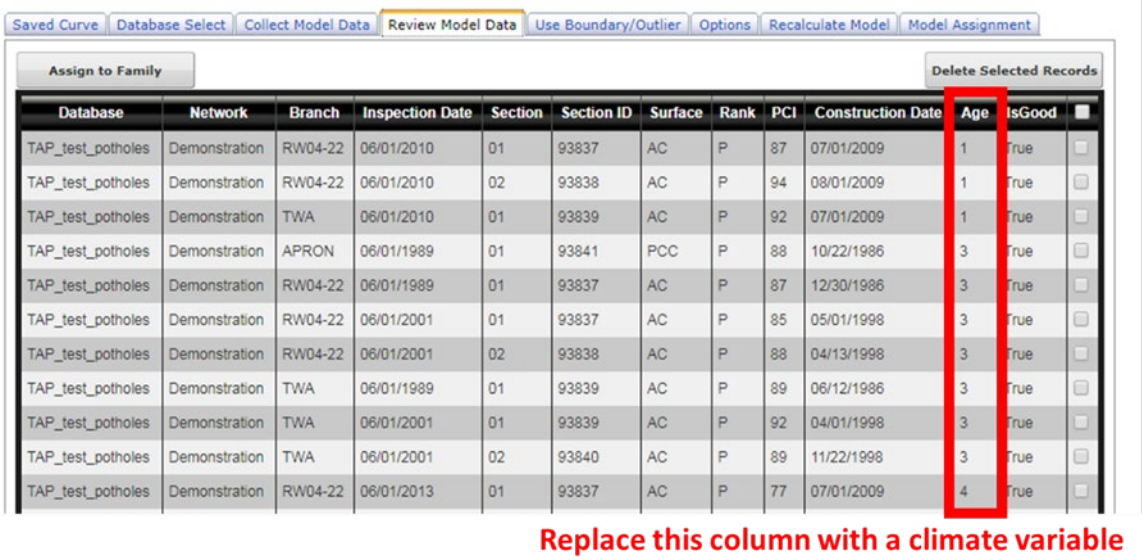
- A review of the available distress data stored in PA40 will be performed with the goal of identifying airports with relatively complete data sets covering an extended period. A number of sections will exhibit distresses, and data will be available from trafficked (keel) and lightly trafficked (outside keel) sections. Sections with multiple inspections will be reviewed to determine if the distresses exhibit progressive behavior, that is, a distress that increases in quantity and/or severity over time.
- For the airports with complete distress data sets, climate and weather data will be retrieved from the FAA web service using PA40. Data will be retrieved and calculations performed for the periods coinciding with the available distress data.
- This step includes performing various analyses (analysis of variance [ANOVA], main effect, liner regression, t-Tests or others) of climate and weather against the distress data, to look for correlation between the variables. Some previously reported relationships that can be further investigated include:
    - PCC pavements
        - Average temperature/humidity/wind speed/solar radiation and shrinkage cracking/map cracking
        - Freeze precipitation days and durability cracking
        - Precipitation/humidity and Alkali Silica Reactivity (ASR)
        - Precipitation and faulting/corner breaks
        - Temperature and joint spall/corner spall
        - Freeze/thaw cycles and joint spall/corner spall
        - Days above 90° F and blowups
        - Average temperature and Longitudinal, Transverse, and Diagonal (LTD) cracking
    - Asphalt pavements
        - Average temperature and rutting
        - Average temperature and shoving

        - Average temperature and Longitudinal &Transverse (L&T) cracking
        - Freeze/thaw cycles and L&T cracking
        - Freeze/thaw cycles and block cracking
        - Freeze /thaw cycles and swell
        - Solar radiation and weathering/raveling
        - Precipitation and depressions
        - Precipitation and L&T cracking
- In addition to a detailed analysis of individual distresses an assessment of climate/weather data versus anti-structural condition index (Anti-SCI) (a non-load-related component of pavement condition index [PCI]), foreign object debris (FOD) index, roughness and PCI can also be calculated to determine if a relationship exists.

A.5 Summary and discussion

Previous research indicates that the presence and severity of some distresses defined in ASTM D5340 (ASTM International, 2020) correlate to the climate zone in which a pavement is located (dry-freeze, dry-non-freeze, wet-freeze, wet non-freeze). PA40 has been modified to further explore this relationship by providing the capability to associate specific weather events experienced by a pavement with the distresses and condition of that pavement at the time of inspection. The system focuses on the historical moisture and temperature conditions of a pavement as most climate-related distresses are believed to be affected by these two factors. Solar radiation is also included in the system in order to determine if a correlation can be established between weathering/raveling and solar radiation input.

The expected use of the system is to produce data sets similar to those used for generating family deterioration curves in a standard pavement management system, but allowing the use of data other than age as the independent variable. Once the existence of a correlation between a condition and a climate variable is established, future work can add the capability to generate these model types in PA40, as shown in Figure A- 4 and Figure A- 5.



Database	Network	Branch	Inspection Date	Section	Section ID	Surface	Rank	PCI	Construction Date	Age	IsGood	
TAP_test_potholes	Demonstration	RW04-22	06/01/2010	01	93837	AC	P	87	07/01/2009	1	True	
TAP_test_potholes	Demonstration	RW04-22	06/01/2010	02	93838	AC	P	94	08/01/2009	1	True	
TAP_test_potholes	Demonstration	TWA	06/01/2010	01	93839	AC	P	92	07/01/2009	1	True	
TAP_test_potholes	Demonstration	APRON	06/01/1989	01	93841	PCC	P	88	10/22/1986	3	True	
TAP_test_potholes	Demonstration	RW04-22	06/01/1989	01	93837	AC	P	87	12/30/1986	3	True	
TAP_test_potholes	Demonstration	RW04-22	06/01/2001	01	93837	AC	P	85	05/01/1998	3	True	
TAP_test_potholes	Demonstration	RW04-22	06/01/2001	02	93838	AC	P	88	04/13/1998	3	True	
TAP_test_potholes	Demonstration	TWA	06/01/1989	01	93839	AC	P	89	06/12/1986	3	True	
TAP_test_potholes	Demonstration	TWA	06/01/2001	01	93839	AC	P	92	04/01/1998	3	True	
TAP_test_potholes	Demonstration	TWA	06/01/2001	02	93840	AC	P	89	11/22/1998	3	True	
TAP_test_potholes	Demonstration	RW04-22	06/01/2013	01	93837	AC	P	77	07/01/2009	4	True	

Figure A- 4. Recommended method to integrate a climate variable into the PA40 modeling system

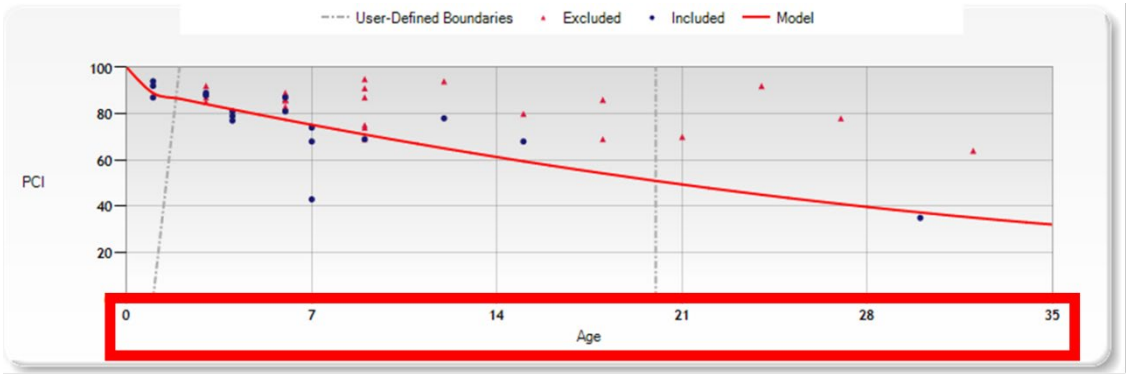


Figure A- 5. Recommended climate modeling capability for PA40 after the existence of a correlation is established

B .NET Library Classes

B.1 InputFeatures

InputFeatures is a class to organize the pavement, weather, and traffic input data for the machine learning (ML) models for five pavement condition indexes, including anti-structural condition index (anti-SCI), pavement condition index (PCI), structural condition index (SCI), runway roughness index (RRI) and Groove Index. Each InputFeatures object describes each input parameter for the ML models. Various models were developed for each pavement condition index by incorporating different sets of InputFeatures.

Constructor

InputFeatures ()	**Initializes input features objects**

Table B- 1. Properties of anti-SCI, PCI, and SCI models for flexible pavement

Property	**Type**	**Range**	**Description**
Age	Double	0.5 to 20	Time since last major rehabilitation (year)
PreviousIndex	Double	45 to 100 (anti-SCI) 30 to 100 (PCI) 40 to 100 (SCI)	Previous anti-SCI, PCI, or SCI measurement
DeltaAge	Double	0.5 to 15	Time since previous anti-SCI, PCI, or SCI measurement (year)

Property	Type	Range	Description
Center	Integer	1 if Center, 0 if End	Section location with respect to runway end (0 if End or 1 if Center)
Const	Double	1 to 5	Number of previous rehabilitation projects
AvgTemp	Double	52 to 78 (°F)	Average annual weather variables between two anti-SCI, PCI, or SCI measurements or since last rehabilitation/reconstruction
AnnualAvgFDD	Double	1 to 655 (°F-days)	
AnnualAvgFThC	Double	1 to 88 (cycles)	
AnnualAvgTemperatureOver90	Double	2 to 180 (days)	
AnnualAvgDaysPrecipitation	Double	40 to 152 (days)	
AnnualAvgTotalPrecipitation	Double	7 to 62 (inches)	
Thornthwaite	Double	-45 to 120 (%)	
AnnualAvgRelativeHumidity	Double	24 to 65 (%)	
AnnualAvgHydration	Double	26 to 288 (days)	
AnnualD200Dep	Double	120 to 19000	Average annual departures of generic D-200
Annual2D400Dep	Double	7 to 10600	Average annual departures of generic 2D-400
AnnualHeavyDep	Double	0 to 5600	Average annual departures of generic Heavy aircraft

Table B- 2. Properties of anti-SCI, PCI, and SCI models for rigid pavement

Property	Type	Range	Description
Age	Double	0.5 to 50	Time since last major rehabilitation (year)
PreviousIndex	Double	54 to 100 (anti-SCI) 40 to 100 (PCI)	Previous anti-SCI, PCI, or SCI measurement

Property	Type	Range	Description
		50 to 100 (SCI)	
DeltaAge	Double	0.5 to 40	Time since previous anti-SCI, PCI, or SCI measurement (year)
Center	Integer	1 if Center, 0 if End	Section location with respect to runway end (0 if End or 1 if Center)
AvgTemp	Double	44 to 77 (°F)	Average annual weather variables between two anti-SCI, PCI, or SCI measurements or since last rehabilitation/reconstruction
AnnualAvgFDD	Double	1 to 730 (°F-days)	
AnnualAvgFThC	Double	1 to 121 (cycles)	
AnnualAvgTemperatureOver90	Double	1 to 109 (days)	
AnnualAvgDaysPrecipitation	Double	36 to 236 (days)	
AnnualAvgTotalPrecipitation	Double	8 to 76 (inches)	
Thornthwaite	Double	-26 to 108 (%)	
AnnualAvgRelativeHumidity	Double	41 to 74 (%)	
AnnualAvgHydration	Double	70 to 343 (days)	
AnnualD200Dep	Double	45 to 32000	Average annual departures of generic D-200
Annual2D400Dep	Double	9 to 23000	Average annual departures of generic 2D-400
AnnualHeavyDep	Double	4 to 7500	Average annual departures of generic Heavy aircraft

Table B- 3. Properties of RRI and groove index models for flexible and rigid pavement

Property	Type	Range	Description
Age	Double	1.5 to 18	Time since last major rehabilitation (year)

Property	Type	Range	Description
Center	Integer	1 if Center, 0 if End	Section location with respect to runway end (End or Center)
AvgTemp	Double	52 to 78 (°F)	Average annual weather variables since last rehabilitation/reconstruction
AnnualAvgFDD	Double	1 to 500 (°F-days)	
AnnualAvgFThC	Double	1 to 75 (cycles)	
AnnualAvgTemperatureOver90	Double	2 to 160 (days)	
AnnualAvgDaysPrecipitation	Double	45 to 158 (days)	
AnnualAvgTotalPrecipitation	Double	3 to 75 (inches)	
Thornthwaite	Double	-45 to 85 (%)	
AnnualAvgRelativeHumidity	Double	26 to 66 (%)	
AnnualAvgHydration	Double	30 to 270 (days)	
AnnualD200Dep	Double	300 to 35000	Average annual departures of generic D-200
Annual2D400Dep	Double	12 to 10600	Average annual departures of generic 2D-400
AnnualHeavyDep	Double	0 to 5600	Average annual departures of generic Heavy aircraft
TotalAnnualDepart	Double	3000 to 98500	Average annual departures of all aircraft
TotalAnnualArrive	Double	2600 to 127000	Average annual arrivals of all aircraft

B.2 IndexPredictModel

Predict pavement condition indexes including anti-SCI, PCI, SCI, RRI, and Groove Index for both flexible and rigid pavements. The InputFeatures are input arguments. The PredictResults return predicted indexes.

Constructor

IndexPredictModel ()	Initializes Prediction Model for Pavement Condition Indexes

Properties

antiSCIFlexModel ()	Initializes Prediction Model for anti-SCI Flexible Pavement
PCIFlexModel ()	Initializes Prediction Model for PCI Flexible Pavement
SCIFlexModel ()	Initializes Prediction Model for SCI Flexible Pavement
RRIFlexModel ()	Initializes Prediction Model for RRI Flexible Pavement
GrooveIndexFlexModel ()	Initializes Prediction Model for Groove Index Flexible Pavement
antiSCIRigidModel ()	Initializes Prediction Model for anti-SCI Rigid Pavement
PCIRigidModel ()	Initializes Prediction Model for PCI Rigid Pavement
SCIRigidModel ()	Initializes Prediction Model for SCI Rigid Pavement
RRIRigidModel ()	Initializes Prediction Model for RRI Rigid Pavement
GrooveIndexRigidModel ()	Initializes Prediction Model for Groove Index Rigid Pavement

Methods

Method	**Arguments**	**Returns**
predict ()	List of InputFeatures	A single value of PredictResult

B.3 PredictResult

Results of pavement condition index prediction. It provides a single value of predicted anti-SCI, PCI, SCI, RRI, or Groove Index for both flexible and rigid pavements.

Constructor

PredictResults ()	**Initializes a result data object for pavement condition indexes**

Table B- 4. Properties

Property	**Type**	**Description**
AntiSCIFlexResults	Single	Predicted anti-SCI for flexible pavement
PCIFlexResults	Single	Predicted PCI for flexible pavement
SCIFlexResults	Single	Predicted SCI for flexible pavement

RRIFlexResults	Single	Predicted RRI for flexible pavement
GrooveIndexFlexResults	Single	Predicted Groove Index for flexible pavement
AntiSCIRigidResults	Single	Predicted anti-SCI for rigid pavement
PCIRigidResults	Single	Predicted PCI for rigid pavement
SCIRigidResults	Single	Predicted SCI for rigid pavement
RRIRigidResults	Single	Predicted RRI for rigid pavement
GrooveIndexRigidResults	Single	Predicted Groove Index for rigid pavement

B.4 Attributes

Attributes is a class to organize the input data for the serviceability level (SL) model. Each Attribute represents a contributed pavement condition index object for the SL models. Various models were developed for SL incorporating different sets of Attribute.

Constructor

Attributes ()	**Initializes attribute objects**

Table B- 5. Properties of attributes

Property	Type	Range	Description
antiSCI	Double	34 to 100 (flexible pavement) 50 to 100 (Rigid pavement)	Measured or Predicted anti-SCI
PCI	Double	20 to 100 (flexible pavement) 38 to 100 (Rigid pavement)	Measured or Predicted PCI
SCI	Double	40 to 100 (flexible pavement) 58 to 100 (Rigid pavement)	Measured or Predicted SCI
RRI	Double	0.11 to 0.43 (flexible pavement) 0.2 to 0.5 (Rigid pavement)	Measured or Predicted RRI
GrooveIndex	Double	10% to 80% (flexible pavement) 1% to 65% (Rigid pavement)	Measured or Predicted Groove Index

B.5 SLPredictModel

Predict SL for flexible pavement using Attributes as input arguments.

Constructor

SLPredictModel ()	**Initializes Prediction Model for SL Flexible Pavement**

Table B- 6. Methods

Method	**Arguments**	**Returns**
Predict ()	List of Attributes	A single value of PredictSLResults

B.6 PredictSLResults

Results of SL predictions. It provides a single value of predicted SL for flexible pavement.

Constructor

PredictSLResults ()	**Initializes a data object for SL flexible pavement**

Table B- 7. Methods

Method	**Type**	**Description**
antiSCISCIRRIGrooveIndex ()	Single	Initializes Prediction SL Model for flexible pavement using antiSCI, SCI, RRI and GrooveIndex
PCIRRIGrooveIndex ()	Single	Initializes Prediction SL Model for flexible pavement using PCI, RRI and GrooveIndex
PCIRRI	Single	Initializes Prediction SL Model for flexible pavement using PCI and RRI
PCIGrooveIndex	Single	Initializes Prediction SL Model for flexible pavement using PCI and GrooveIndex

Table B- 8. Properties

Property	**Type**	**Description**
SL	Single	Predicted SL for flexible pavement: 0 = Serviceable 1 = Unserviceable