40-Year Pavement Life

Machine Learning and PA40 Data

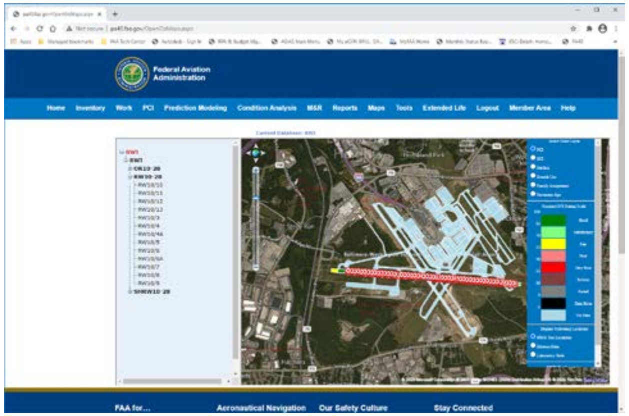
Presented to: REDAC Briefing to Sub-committee on Airports

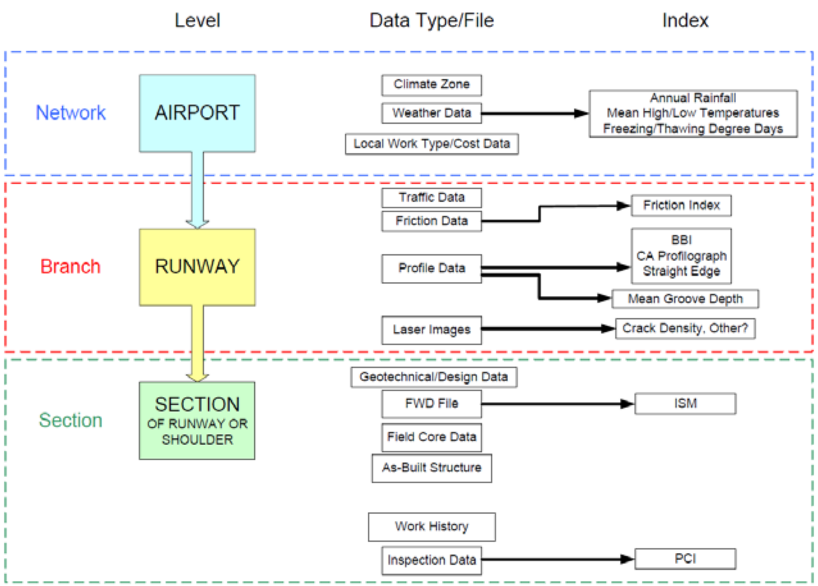
By: David R. Brill, P.E., Ph.D.

Date: March 3, 2021

Extended Airport Pavement Life Database (PA40)

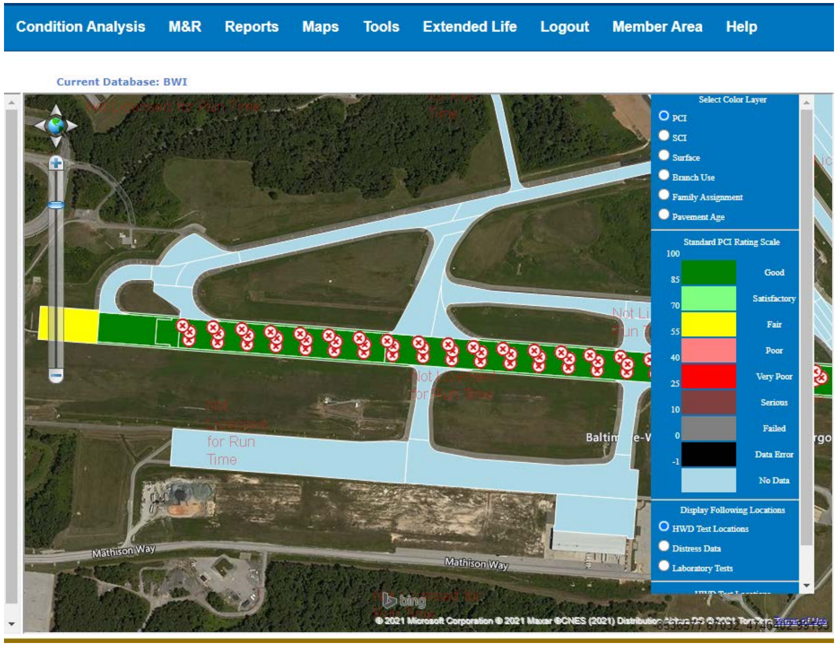
- **Database of construction and performance data on 28 runways at 22 large- and medium-hub airports in the U.S.**
- **Built on FAA PAVEAIR 3.0 code base.**
- **Structured like PMS, but with additional fields for:**
  - Surface friction, profile roughness, groove condition, HWD data.
  - Historical runway usage and weather data.
  - Structural design & as-built section data.
  - Field core test data.





PA40 Recent Updates

- **Runway traffic data updated through October 2020.**
- **Added GIS map labels for pavement data.**
- **Added ability to search/display distress data on the sample unit level.**
- **Completed advanced query tool enhancements. Now displays traffic/weather totals between any two dates.**
- **New modeling features: PCI/SCI/Anti-SCI vs. Age/Traffic**



Current Database: BWI

Back | Export as csv

Database	Network	Branch	Section	Inspection	Sample Number	Sample Type	Sample Size	Sample SizeUnit	No Distress	Distress	Description	Severity	Quantity	Quantity Unit
BWI	BWI	RW10-28	RW10/10	8/15/2011 12:00:00 AM	05	R	5000.0	ft²	0	48	L & T CR	L	50.0	ft
BWI	BWI	RW10-28	RW10/10	8/15/2011 12:00:00 AM	05	R	5000.0	ft²	0	48	L & T CR	M	75.0	ft
BWI	BWI	RW10-28	RW10/10	8/15/2011 12:00:00 AM	05	R	5000.0	ft²	0	57	WEATHERING	M	5000.0	ft²

Performance Models

- **Serviceability Index (*SL*)**
  - Combination of multiple performance indexes.
    - Structural: SCI, FWD-derived index.
    - Functional: Runway roughness (RRI), friction, “anti-SCI” (non-structural components of PCI).
  - Predictors may include: Pavement age, traffic, environmental cycles.
- **Initial FAA study found that conventional regression analysis is inadequate to predict pavement performance as a function of multiple predictors.**
- **Decision to use machine learning (ML) methods to predict *SL* from multiple predictors.**

Machine Learning Goals

(from BAA ARAP0004, Unsupervised Learning and Database Analysis)

- **Identify the key variables that most influence pavement longevity, and eliminate variables that have little or no influence.**
- **Identify key performance indexes or combinations of indexes tied to pavement failure or a decision to rehabilitate/reconstruct/replace.**
- **Perform clustering, or find trends/correlations in the EAPL data that may not be obvious.**
- **Develop data-based models for predicting long-term pavement condition, employing a mix of inputs such as traffic cycles, weather cycles, age, maintenance data, and structural or material properties.**

Flexible Runways Studied

Airport	Runway
Boston Logan Airport (BOS)	4L-22R
Columbus International Airport (CMH)	10L-28R; 10R-28L
Greensboro International Airport (GSO)	5L-23R
Kansas City International Airport (MCI)	9-27
LaGuardia Airport (LGA)	4-22
Miami International Airport (MIA)	12-30
San Francisco International Airport (SFO)	10R-28L
Tucson International Airport (TUS)	11L-29R; 3-21

Feature Selection

- **Use ML to identify the key environmental variables that affect runway pavement performance.**
- **Problem is to assess a set of candidate predictors against a target value (in this case, anti-SCI).**
- **There are more independent variables than the number of climate conditions in the database.**
- **Potential issues:**
  - Collinearity. Variables that are highly correlated are redundant and can affect the prediction performance negatively.
  - On the other hand, the PA40 database does not cover all geographic/climate scenarios. Models may be underspecified, or significant features could be wrongly eliminated.

Weather Variables Considered

Environmental Variables	Unit
Freezing Degree Days (FDD)	ºF
Freeze Thaw Cycles (FThC)	cycles
Days Temperature Over 90˚F (Temp90)	days
Days Precipitation (DPrec)	days
Total Precipitation (TPrec)	inches
Freeze Precipitation Days (FPD)	days
Hydration Days (HD)	days
Average Daily Temperature (Avg Temp)	ºF
Average Daily Temperature Difference (Temp Diff)	ºF
RHumidity Avg	%
Avg Wind Speed	mph
Thornthwaite Index	%
Sky Cover	oktas

- **Approach 1 – Weather variables are treated as cumulative values from date of construction/rehabilitation to date of inspection.**
  - Target is measured anti-SCI.
  - Include pavement age and previous measured anti-SCI as predictors (auto-regressive approach)
- **Approach 2 – Weather variables are treated as average values.**

Feature Selection using Ranking Algorithms in Approach 1

Filter Methods				Wrapper Methods			
Pearson Correlation		RReliefF		SVM (Gaussian Kernel)		Linear Regression	
Input Variable	*R*	Input Variable	Score	Input Variable	Score	Input Variable	Score
Previous anti-SCI	0.77	Previous anti-SCI	0.040	Age	12.3	Previous anti-SCI	5.7
Age	-0.77	Temp90	0.022	Previous anti-SCI	10.5	Age	5.6
DPrec	-0.63	Age	0.016	Temp90	9.3	DPrec	3.5
Temp90	-0.59	TPrec	0.010	DPrec	9.1	Temp90	2.9
TPrec	-0.57	HD	0.009	HD	8.0	TPrec	2.8
HD	-0.57	DPrec	0.009	TPrec	7.6	HD	2.7
FThC	-0.54	FDD	0.006	FDD	7.1	FThC	2.4
FDD	-0.50	FThC	0.005	FThC	6.7	FDD	2.0

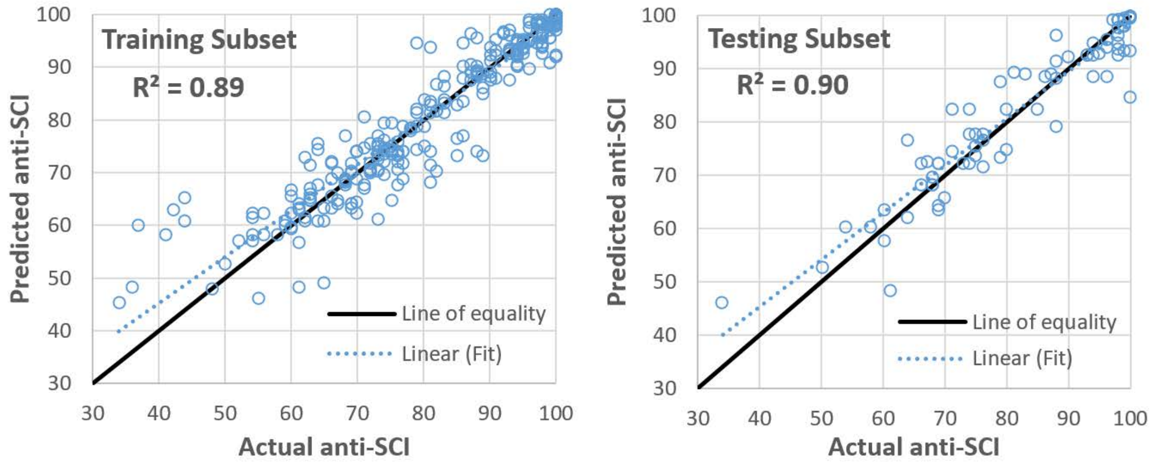
Modeling Approaches

- **Three modeling approaches were used for initial model development:**
  - Subset of variables
  - Principal component analysis (PCA)
  - Features from *k*-means cluster analysis
- **All approaches based on autoregressive random forest (RF) learning algorithm.**
- **Used "Weka" (freeware) for data analysis and model implementation.**
- **Used 278 data records from 10 flexible runways to train the RF model.**

Principal Component Analysis (PCA)

- **Unsupervised method for reducing dimensionality of feature space.**
  - Map data onto a set of new uncorrelated variables (the PCs).
  - Unlike original features, PCs possess no physical meaning.
- **Only 4 PCs explain more than 95% of the variance.**

Comparison of predicted and actual anti-SCI - Modeling Approach 2: PCA



Performance of Trained RF Model

Performance Measure	1. Subset of Variables		2. PCA		3. Cluster Analysis		[No Climate]	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing
R^2	0.89	0.9	0.89	0.9	0.88	0.93	0.8	0.84
RMSE	5.16	4.55	5.11	4.81	5.41	3.86	6.8	5.9
RRSE	9.5%	9.0%	9.0%	7.3%	10.2%	6.3%	13.4%	8.1%
Accuracy (5% error)	69%	68%	74%	67%	69%	68%	67%	68%
Accuracy (10% error)	90%	95%	88%	90%	90%	92%	87%	89%

- **Take-away: Pavement age and previous anti-SCI remain the most significant predictors of current anti-SCI.**
- **Upcoming research will explore ML methods for predicting other components of *SL* (related to FOD, roughness and low friction).**

Technical Products

- **Technical Report (in editing):** *Application of Machine Learning Techniques to Pavement Performance Modeling*, Sept. 2020
- **Two papers accepted for presentation:**
  - BCRRA 2021: "*Machine Learning Solutions for Development of Performance Deterioration Models of Flexible Airfield Pavements*" (Conference delayed until 2022 due to COVID-19)
  - ASCE T&DI Pavements 2021: "*Machine Learning Approach to Identifying Key Environmental Factors for Airfield Asphalt Pavement Performance*"

Thank You!

http://www.airporttech.tc.faa.gov/

david.brill@faa.gov

Acknowledgments:

FAA Airport Technology R&D Branch:
Dr. Michel Hovan, Branch Manager;
Jeff Gagnon, Airport Pavement Section Manager;

FAA Airport Engineering Division:
Doug Johnson (retired); Harold Honey

ARA:
Dr. Ali Z. Ashtiani; Scott Murrell; Timothy Parsons; Rich Speir