ABSTRACT: The reconstructed flexible pavement test sections at the National Airport Pavement Test Facility (NAPTF) consisted of similar Asphalt Concrete (AC) and granular base thicknesses, but variable granular subbase thicknesses. These sections were simultaneously trafficked by 6-wheel and 4-wheel landing gears. Rut depth measurements were made at frequent intervals. The development of rut depths under trafficking were characterized using the Power model and the Pavement Surface Rutting Rate (RR) model. For similar number of load repetitions, test sections with reduced subbase thicknesses yielded larger rut depths. The Power model coefficient “A” is helpful in quantifying the rutting potential and early-life rut depth development. The final rut depths varied between 4 to 8 inches. The NAPTF 1-inch surface upheaval failure criteria did not yield consistent rut depths. A functional failure criterion is suggested.

KEY WORDS: NAPTF, rutting, dual tridem gear, ILLI-PAVE.

1 INTRODUCTION

The National Airport Pavement Test Facility (NAPTF) located at the Federal Aviation Administration (FAA) William J. Hughes Technical Center, Atlantic City International Airport, New Jersey, USA was constructed to generate full-scale testing data to support the investigation of the performance of airport pavements subjected to complex gear loading configurations of New Generation Aircraft (NGA) such as the Boeing 777.

The NAPTF test pavement area is 274.3 m (900 feet) long and 18.3 m (60 feet) wide. The first set of test pavements included a total of nine test sections (six flexible and three rigid) built on three different subgrade materials: low-strength (target CBR of 4), medium-strength (target CBR of 8), and high-strength (target CBR of 20). Two different base sections were used: conventional (granular) and stabilized (asphalt concrete). Trafficking tests were conducted using Boeing 777 (B777) gear in one lane and Boeing 747 (B747) gear in the other lane.

The primary objective of the trafficking tests conducted on the first set of NAPTF flexible test sections was to determine the number of load applications to cause shear failure in the subgrade. Per NAPTF “failure” criterion, this is reflected as 1-inch surface upheaval adjacent to the traffic lane. While the medium-strength subgrade test sections “failed” at the subgrade level, the low-strength subgrade test sections showed significant distress in the surface layers, signifying tire pressure or other upper layer failure effects, but not subgrade level failure (Gervais et al. 2003).
Subsequently, the pavement structures from the low-strength subgrade were removed. Four pavement sections (LFC1, LFC2, LFC3, and LFC4) with variable granular subbase thicknesses were reconstructed for retesting (Fig. 1). LFC refers to a conventional granular-base pavement built over a low-strength (CBR4) subgrade. The P-209 base layer was graded crushed rock and the P-154 subbase layer consisted of stone screenings. The LFC1 and LFC2 sections were designed to “fail” early and the LFC3 and LFC4 sections were designed to last longer. These sections were simultaneously trafficked using 6-wheel and 4-wheel landing gears. Transverse Surface Profile (TSP) measurements as well as straightedge rut depth measurements were made at frequent intervals to monitor rutting.

The distinctive feature of these test sections is that the subbase thickness is the only structural variable. Thus, the exclusive effect of Subbase thickness on response and performance of the test sections can be studied. In this paper, the effect of variable subbase thickness on the rutting performance of airport flexible pavement tests sections subjected to 6-wheel and 4-wheel NGA gear loading is presented.

![Figure 1: Cross-Sectional views of NAPTF reconstructed test sections (not to scale).](image)

2 NAPTF TRAFFIC TESTING

The reconstructed test sections were trafficked by a 6-wheel gear configuration on the North side traffic path and by a 4-wheel gear configuration on the South side traffic path (Fig. 2). Note that the 6-wheel gear configuration is very similar to the B777 gear configuration used during the first series of tests. The 4-wheel gear configuration, obtained by removing one of the axles from the 6-wheel gear configuration, is similar to the B747 gear configuration except for the dual-wheel spacing, which is 1372 mm [54 inches] in 4-wheel gear and 1118 mm [44 inches] in B747 gear. The wheel load was set to 25 tonnes (55,000 lb), and the cold inflation pressure was 1688 Kpa (245 psi). The traffic speed was set to 8 km/h (5 mph). The traffic tests were conducted during September and October of 2002.

2.1 Traffic Wander Pattern

To realistically simulate transverse aircraft movements, a wander pattern consisting of a fixed sequence of 66 vehicle passes (33 traveling in the East direction and 33 traveling in the West direction), arranged in nine equally spaced wander positions (or tracks) at intervals of 260 mm (10.25 inches), was used during traffic testing. This wander pattern simulates a normal
distribution of aircraft traffic with a standard deviation ($\sigma$) of 775 mm (30.5 inches) that is typical of multiple gear passes in airport taxiways.

![Image of landing gear configurations](image_url)

**Figure 2:** Landing gear configurations used in traffic tests.

### 2.2 Failure Criterion

The NAPTF “failure” criterion, based on the criterion used by the US Corps of Engineers’ Multiple Wheel Heavy Gear Load (MWHGL) Tests (Ahlvin et al 1971), is “at least 25.4-mm (1-inch) surface upheaval adjacent to the traffic lane”. This is considered to reflect structural or shearing failure in the subgrade.

It is important to note that, in the 25.4-mm (1-inch) surface upheaval “failure” criterion, there is no limit on the maximum rut depth. Thus, a surface upheaval of 25.4 mm (1 inch) may be accompanied by a 13-mm (0.5-inch) rut depth or rut depths in excess of 50 mm (2 inches) to 75 mm (3 inches) with no limit on the maximum allowable rut depth. However, according to the Unified Facilities Criteria (UFC) (US COE 2001), a rut depth in excess of 25.4 mm (1 inch) is considered as “High” severity rutting and constitutes a significant functional failure requiring major maintenance activities. It is recognized that ultimately the magnitude of surface rut depths will dictate performance irrespective of surface upheaval.

### 3 RUTTING MEASUREMENTS

A non-contact laser profiling device with a span of 6.6 m (21.5 feet) and a range of 203 mm (8 inches) was used to measure transverse surface elevation profiles at frequent intervals during trafficking. Transverse Surface Profiles (TSPs) were measured along three profile lines (one in East side, one in West side and one in Center) marked across the pavement. The maximum rut depths and maximum uplifts were extracted from the TSPs. Apart from TSP measurements, rut depths were also periodically measured using a 3.66-m (12-ft) straightedge.

The LFC1 and LFC2 sections “failed” at 200 passes and 3,200 passes, respectively. Both the LFC3 and LFC4 sections showed little distress, other than accumulation of some rutting (less than 89 mm [3.5 inches]) at 4,000 passes and therefore the wheel loads were increased to 29.5 tonnes (65,000-lb) for the remaining duration of traffic testing. The 6-wheel North traffic lane of LFC3 “failed” at 21,000 passes. The LFC4 section did not show signs of “failure” at 21,000 passes. The trafficking was terminated on LFC3 and LFC4 sections at 23,286 passes.

TSP maximum rut depths were obtained as a function of number of load repetitions (N) and the rutting data were characterized using some of the well-known pavement rutting models such as the Power model and the Pavement Surface Rutting Rate (RR) model.
DEVELOPMENT OF RUT DEPTHS

The progression of TSP rutting in NAPTF test sections with trafficking (N) are summarized in Fig 3. Similar results are shown in Fig 4 based on Straightedge rut depth measurements. The straight-edge rut depths were consistently lower than the TSP rut depths, especially as N increased. This was also observed during the first series of traffic tests (Gopalakrishnan and Thompson 2003). This may be because, as the rut width increases above 3.66 m (12 feet) due to the effect of trafficking, the position of the 3.66-m (12-ft) straight edge goes down, thus recording smaller rut depths. It is also noted that the accuracy of straightedge measurements depends mostly on the ability of the operator to correctly place the straightedge to measure the maximum ruts in the profile (Gramling et al 1991).

The LFC1 section showed rapid accumulation of rutting and “failed” at 90 passes on the 6-wheel traffic path and at 122 passes on the 4-wheel traffic path. In LFC2, rutting accumulated at a rapid rate with rut depths reaching 4 inches in both the traffic lanes within 400 passes. After 500 passes, the rut depth accumulations in the 6-wheel side and 4-wheel side diverged with the 6-wheel side increasing more rapidly and reaching “failure” at 1,100 passes. The 4-wheel traffic path reached “failure” at around 3,000 passes. The same trend is seen in the LFC3 and LFC4 sections with the divergence in rutting curves occurring around 7,000 passes.

![Summary of TSP Rut Depth Measurements](image)

In general, after a certain number of passes, rutting on the 6-wheel side increased rapidly compared to the 4-wheel side. This was more pronounced in the LFC1 and LFC2 sections. As indicated earlier, the LFC1 and LFC2 sections reached “failure” while the LFC3 and LFC4 sections did not “fail” (i.e., 1-inch surface upheaval) even after reaching 5 to 8 inches of rut depth. The magnitudes of rut depths at “failure” in LFC1 and LFC2 sections were approximately 4 inches and 8 inches respectively. At the termination of traffic testing, the rut depths in the LFC3 and LFC4 sections were approximately 8 inches and 6 inches respectively.
Figure 4: Summary of Straightedge Rut Depth Measurements

The LFC3 section is similar in cross-section (thickness and material composition) to the LFC section (a conventional granular base pavement built over low-strength subgrade) built during the first series of test pavements. The only difference was that the P-154 unbound granular subbase in LFC was 2 inches thicker. The B777 traffic gear configuration used during first series of traffic tests is identical to the 6-wheel gear configuration used for trafficking the reconstructed test sections.

The LFC section was trafficked under 20.4-tonnes (45-kip) wheel loading for 20,000 passes and the wheel load was subsequently increased to 29.5-tonnes (65-kip). In 20,000 passes of (20.4-tonnes) 45-kip wheel load, the LFC section accumulated just 19 mm (0.75 inch) of rut depth. After 30,000 passes of 29.5-tonnes (65-kip) wheel load trafficking, the final rut depth was 76.2 mm (3 inches) at the termination of test trafficking (Gopalakrishnan and Thompson 2003). The LFC3 section was initially subjected to 4,000 passes of 25-tonnes (55-kip) wheel load and the wheel load was subsequently increased to 29.5-tonnes (65-kip). In 4,000 passes of 55-kip wheel load, the rut accumulation in LFC3 was 76.2 mm (3 inches).

Thus, although the LFC section was subjected to 30,000 passes of 29.5-tonnes (65-kip) wheel loading after initial 20,000 passes of 20.4-tonnes (45-kip) wheel loading, the final rut depth was 76.2 mm (3 inches), whereas under 4,000 passes of 25 tonnes (55-kip) wheel load, the rut accumulation was 76.2 mm (3 inches) in LFC3. This clearly indicates the effect of stress sequence on rutting. Unpublished laboratory results (U of I) for a lean clay showed that a stress history of gradually increasing stress level may cause less permanent deformation in the subgrade than a sequence where the high stress level is initially applied (Bejarano and Thompson 1999).
5 RUTTING MODELS

A NCHRP 1-26 (1990) study considered several material permanent strain accumulation models and pavement system rutting models and concluded that the predominant flexible pavement rutting model is the Power model. The Pavement Surface Rutting Rate (RR) model is based on the Power model and both these models provide a logical and consistent link between model parameters and available engineering properties affecting permanent deformation.

5.1 Power Model

The Power model is:

\[ \log RD = a + b \log N \]
\[ \text{or} \]
\[ RD = AN^b \]

where \( RD \) = rut depth (mm or mils); \( N \) = number of load repetitions; \( a \) and \( b \) = experimentally determined factors; and \( A \) = antilog of \( a \).

The A term is quite variable and is strongly influenced by material/soil type, repeated stress state, and factors influencing material shear strength (Thompson and Nauman 1993). Stress state is generally expressed in terms of (1) repeated deviator stress \( (\sigma_D = \sigma_1 - \sigma_3) \), (2) principal stress ratio \( (\sigma_1/\sigma_3) \), and (3) deviator stress ratio \( (\sigma_D/\sigma_3) \). Most soils exhibit a “threshold stress level” defined as that stress level above which permanent deformation accumulates rapidly under repeated loading and below which the rate of cumulative deformation from additional load applications is very small and acceptable. In most cases, the threshold stress level is 50 to 60 percent of the ultimate strength of the soil (Thompson and Nauman 1993). The stress ratio (repeated deviator stress/ultimate strength) was found to relate to the A term and is a valid indicator of the rutting potential. In relative terms, low A values are noted for reduced stress ratios and large A values are associated with increased stress ratios.

5.2 Pavement Surface Rutting Rate Model

Thompson and Nauman (1993) proposed and evaluated a phenomenological pavement surface RR model based on the Power model and the Ohio State University model (Majidzadeh et al. 1981) assuming that all of the paving materials and the subgrade soils generally follow the Power model:

\[ RR = \frac{RD}{N} = A/N^B \]

where \( RR \) = rutting rate; \( RD \) = rut depth (mm or mils); \( N \) = number of load repetitions; and \( A \) and \( B \) = terms developed from field calibration testing data. Note that in Equation 2, \( B = b - 1 \), where \( b \) is the slope of the Power model (Equation 1).

The validity of the RR concept was evaluated by Thompson and Nauman (1993) using AASHO Road Test (1962) data and field performance data from the AASHO Test Site’s (I-80) rehabilitated and new flexible pavement sections. The results indicated that stable pavement rutting trends were related to estimated structural responses, particularly the Subgrade Stress Ratio (SSR). Thompson and Nauman (1993) concluded that the RR approach can be effectively used in a priori pavement analysis and design and pavement management system activities.
6 RUTTING ANALYSIS

Using the maximum TSP surface rut depths, the rutting results were characterized using the Power model and the Pavement Surface RR model.

In the LFC3 and LFC4 sections, wheel loading was increased from 25 tonnes (55-kip) to 29.5 tonnes (65-kip) around 4,000 passes. The rutting patterns were different depending on traffic wheel load magnitude. Therefore, separate models were applied to characterize 25-tonnes (55-kip) and 29.5-tonnes (65-kip) rutting data. In modeling the rutting data obtained under the 29.5-tonnes (65-kip) wheel loading in the LFC3 and LFC4 sections, the pre-accumulated rutting under the 25-tonnes (55-kip) loading was neglected (i.e., rutting at the initiation of [29.5-tonnes] 65-kip traffic loading was set to zero). The results of rutting analyses using the Power model are summarized in Table 1 for 55-kip wheel load and in Table 2 for 65-kip wheel load. Only the LFC3 and LFC4 sections were subjected to 65-kip wheel loading.

The Rutting Rates (RRs) are plotted against N on a Log-Log scale for NAPTF reconstructed test sections in Fig. 5. The RRs show linear relation with N. It is seen that RR plots for LFC1 and LFC2 are distinctly separate whereas the RR plots for LFC3 and LFC4 overlap under 25-tonnes (55-kip) load and start diverging after the introduction of 29.5-tonnes (65-kip) wheel load. This trend is similar to the maximum surface deflections ($D_0$) obtained during 16.4-tonnes (36-kip) Heavy Weight Deflectometer (HWD) testing conducted prior to the initiation of trafficking (Gopalakrishnan and Thompson, 2004). The $D_0$ values were distinctly different for the LFC1 and LFC2 sections (difference of 0.635 mm [25 mils]) whereas they only differed by 0.203 mm (8 mils) for the LFC3 and LFC4 sections.

Table 1: Power model rutting analysis results (20.4-tonnes [45-kip] wheel loading)

<table>
<thead>
<tr>
<th>Test Section</th>
<th>Final N</th>
<th>Final Rut Depth (mils)</th>
<th>A</th>
<th>b</th>
<th>b-1</th>
<th>$R^2$</th>
<th>SEE</th>
</tr>
</thead>
<tbody>
<tr>
<td>LFC1</td>
<td>90</td>
<td>3990</td>
<td>316</td>
<td>0.55</td>
<td>-0.45</td>
<td>0.976</td>
<td>0.086</td>
</tr>
<tr>
<td>LFC2</td>
<td>1452</td>
<td>7370</td>
<td>227</td>
<td>0.48</td>
<td>-0.52</td>
<td>0.994</td>
<td>0.020</td>
</tr>
<tr>
<td>LFC3</td>
<td>3228</td>
<td>3210</td>
<td>125</td>
<td>0.41</td>
<td>-0.59</td>
<td>0.992</td>
<td>0.037</td>
</tr>
<tr>
<td>LFC4</td>
<td>3228</td>
<td>2800</td>
<td>214</td>
<td>0.32</td>
<td>-0.68</td>
<td>0.998</td>
<td>0.014</td>
</tr>
</tbody>
</table>

Table 2: Power model rutting analysis results (29.5-tonnes [65-kip] wheel loading)

<table>
<thead>
<tr>
<th>Test Section</th>
<th>Final N</th>
<th>Final Rut Depth (mils)</th>
<th>A</th>
<th>b</th>
<th>b-1</th>
<th>$R^2$</th>
<th>SEE</th>
</tr>
</thead>
<tbody>
<tr>
<td>6-Wheel</td>
<td>LFC3</td>
<td>19752</td>
<td>4480</td>
<td>13</td>
<td>0.61</td>
<td>-0.39</td>
<td>0.994</td>
</tr>
<tr>
<td>LFC4</td>
<td>19752</td>
<td>2880</td>
<td>10</td>
<td>0.58</td>
<td>-0.42</td>
<td>0.984</td>
<td>0.026</td>
</tr>
<tr>
<td>4-Wheel</td>
<td>LFC3</td>
<td>19752</td>
<td>3750</td>
<td>19</td>
<td>0.55</td>
<td>-0.45</td>
<td>0.991</td>
</tr>
<tr>
<td>LFC4</td>
<td>19752</td>
<td>2560</td>
<td>12</td>
<td>0.56</td>
<td>-0.44</td>
<td>0.988</td>
<td>0.021</td>
</tr>
</tbody>
</table>
7 EFFECT OF P-154 GRANULAR SUBBASE THICKNESS ON RUTTING PERFORMANCE

The effect of subbase thickness on A and b parameters of the Power model are shown in Figure 6 and Figure 7, respectively. Note that the LFC3 and LFC4 sections were subjected to 29.5-tonnes (65-kip) wheel loading as they did not show “failure” under 25-tonnes (55-kip) wheel loading. Large scatter is observed for LFC3 and LFC4 sections under 29.5-tonnes (65-kip) loading. This is due to the pre-conditioning experienced by these sections under 25-tonnes (55-kip) loading and the unstable rutting behavior under 29.5-tonnes (65-kip) wheel loading. Under these conditions, the applicability of a simple rutting model like the Power model to characterize rutting may not be appropriate.

Figure 6: Effect of P-154 unbound granular subbase thickness on Power mode parameter A
The A term in the Power model is a function of the Subgrade Stress Ratio (SSR) (Thompson & Bejarano 1999, Thompson & Nauman 1993). Low A values indicate reduced SSRs, and large A values are noted for increased SSRs. Plots of A versus P-154 subbase thickness indicate that subbase thickness has a significant influence on the rutting potential. Test sections with thinner subbases showed higher rutting potential than those with thicker subbases.

Using the ILLI-PAVE flexible pavement finite element structural model, the critical structural responses such as subgrade deviator stress (σ_d) and vertical compressive subgrade strain (ε_v) were computed (Gopalakrishnan and Thompson 2004). The reconstructed test sections were modeled in ILLI-PAVE and typical material properties corresponding to the traffic test conditions were used. Regression models were developed to predict Power model parameter A as a function of ILLI-PAVE computed structural responses. As expected, the results showed that A is significantly related to SSR.

8 PROPOSED FUNCTIONAL FAILURE CRITERIA

The 25.4-mm (1-inch) surface upheaval NAPTF “failure” criteria did not yield consistent rut depths. The rut depths varied between 50.8 mm (2 inches) to 127 mm (5 inches) in the first series of NAPTF test sections and between 99.1 mm (3.9 inches) to 190.5 mm (7.5 inches) in the reconstructed test sections. It is obvious that an airport pavement would be considered hazardous to aircraft long before such a large rut occurred. According to the Unified Facilities Criteria (UFC), a rut depth in excess of 25.4 mm (1 inch) is considered as “High” severity rutting and it constitutes a significant functional failure requiring major maintenance activities. Similarly, the ASTM Standard Test Method D 5340-98 suggests functional failure of the pavement if the rut depth exceeds 25.4 mm (1 inch).

Kelly and Thompson (1988) studied the MWHGL performance test data and concluded that the 25.4-mm (1-inch) surface upheaval “failure” criteria appeared to be inconsistent. Some sections were considered failed with a 12.7-mm (0.5-inch) rut while another was not considered failed until a 73-mm (2.88-inch) rut had been achieved and the pavement was severely cracked. It was proposed that transfer functions in which failure is defined as dropping to a specific Pavement Condition Index (PCI) level would be useful.

Ultimately, the surface rut depths will dictate the performance of the pavement and not the surface upheaval. Therefore, appropriate criteria should be considered in addressing pavement performance. It is proposed that the performance measure could be defined in terms of number of load repetitions to reach rut depth failure of 12.7 mm (0.5 inch), 25.4 mm (1 inch),
SUMMARY AND FINDINGS

The NAPTF reconstructed test sections consisted of similar pavement structures (AC + granular base) with variable granular subbase thicknesses (406 mm [16 inches] in LFC1, 610 mm [24 inches] in LFC2, 864 mm [34 inches] in LFC3, and 1092 mm [43 inches] in LFC4). The test sections were trafficked by a 6-wheel gear configuration on one side and by a 4-wheel gear configuration on the other side. During trafficking, rutting was monitored periodically using a 3.66-m (12-ft) straight edge and by measuring the Transverse Surface Profiles (TSPs). The rutting results were characterized using the Power model and the Pavement surface Rutting Rate (RR) model.

The following are the findings of this study:

1. The test section with the smallest subbase thickness (LFC1) showed rapid accumulation of rutting and “failed” at 90 passes on the 6-wheel traffic path and at 122 passes on the 4-wheel traffic path. In the LFC2 section with 203 mm (8 inches) more subbase than the LFC1 section, the 6-wheel side reached “failure” at 1,100 passes and the 4-wheel traffic path reached “failure” around 3,000 passes.

2. The magnitudes of rut depths at “failure” in LFC1 and LFC2 sections are approximately 101.6 mm (4 inches) and 203 mm (8 inches) respectively. The LFC3 and LFC4 sections did not “fail” even after reaching 127 mm (5 inches) to 203 mm (8 inches) of rut depth. At the termination of traffic testing, the rut depths in the LFC3 and LFC4 sections were approximately 203 mm (8 inches) and 152 mm (6 inches) respectively.

3. Plots of A (Power model coefficient) versus P-154 subbase thickness indicate that subbase thickness has a significant influence on the early-life rutting potential as quantified by A. Lower subbase thicknesses showed higher levels of early life rut depth development.

4. The Rutting Rates (RRs) show linear relation with N. The RR plots for LFC1 and LFC2 are distinctly separate whereas the RR plots for LFC3 and LFC4 overlap under 25-tonnes (55-kip) load and start diverging after the introduction of 29.5-tonnes (65-kip) wheel load.

5. The 25.4-mm (1-inch) surface upheaval NAPTF failure criteria did not yield consistent rut depths. It is recognized that ultimately, the magnitude of surface rut depths will dictate pavement performance and not the surface upheaval. Functional failure criteria defined in terms of number of load repetitions to reach rut depth failure of 12.7 mm (0.5 inch), 25.4 mm (1 inch), 38.1 mm (1.5 inches), etc. or rut depths at 5,000 passes, 10,000 passes, etc. are proposed.

ACKNOWLEDGMENTS/DISCIAMER

This paper was prepared from a study conducted in the Center of Excellence (COE) for Airport Technology. Funding for the COE is provided in part by the Federal Aviation Administration (FAA). The COE is maintained at the University of Illinois at Urbana-Champaign who works in partnership with Northwestern University and the FAA. Ms. Patricia Watts is the FAA Program Manager for Air Transportation COE and Dr. Satish Agrawal is the Manager of the FAA Airport Technology R&D Branch. The contents of this paper reflect the views of the authors who are responsible for the facts and accuracy of the
data presented within. The contents do not necessarily reflect the official views and policies of the FAA. This paper does not constitute a standard, specification, or regulation.

REFERENCES


