An-Najah National University

Faculty of Graduate Studies

An Empirical Study to Estimate PCI Using Mobile Sensing and Artificial Intelligence

By
Saeed Riad Ali

Supervisor
Dr. Amjad Issa

Co-supervisor
Dr. Eehab Hijazi

This Thesis is Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Roads and Transportation Engineering, Faculty of Graduate Studies, An-Najah National University, Nablus, Palestine.
2021

An Empirical Study to Estimate PCI Using Mobile Sensing and Artificial Intelligence

By

Saeed Riad Ali

This Thesis was defended successfully on 25/2/2021 and approved by:

Defense Committee Member	**Signature**
– **Dr. Amjad Issa / Supervisor**	....
– **Dr. Eehab Hijazi / Co-Supervisor**	....
– **Dr. Khaled Abaza / External Examiner**	....
– **Dr. Anas Toma / Internal Examiner**	....

Dedication

This thesis is dedicated to my dear parents, brothers and wife who have encouraged and supported me. Also, it is dedicated to my great family and friends.

Acknowledgment

First of all, I am thankful to the Almighty God for granting me good health, strength and peace throughout the research period.

I would like to thank everyone who has contributed to accomplishing this thesis, especially the surveying teams.

I also express my sincere gratitude to my academic advisors Dr. Amjad Issa and Dr. Eehab Hijazi. I am grateful for their continuous support, patience, motivation, enthusiasm, and knowledge. Their guidance has helped me all the time of the research and writing of this thesis that I would never reach on my own.

Also, I would like to thank the discussion committee members who have a great effect in achieving the benefits.

Finally, I would like to thank my family, especially, my parents and brothers for supporting me spiritually through my life, and I offer my gratitude for my wife who has helped me in this research.

الإقرار

أنا الموقع أدناه مقدم الرسالة التي تحمل العنوان :

An Empirical Study to Estimate PCI Using Mobile Sensing and Artificial Intelligence

أقر بأنّ ما اشتملت عليه هذه الرسالة إنما هي نتاج جهدي الخاص ، باستثناء ما تمّت الإشارة إليه حيثما ورد ، و أنّ هذه الرسالة ككل ، أو أي جزء لم يُقدّم لنيل أي درجة أو لقب علمي أو بحثي لدى أي مؤسسة تعليمية أو بحثية أخرى.

Declaration

The work provided in this thesis, unless otherwise referenced, is the researcher's own work and has not been submitted elsewhere for any other degree or qualification.

Student's name: **اسم الطالب: سعيد رياض عبد الحفيظ علي**

Signature: **التوقيع :**

Date: **التاريخ: 25 / 2 / 2021**

Table of Content

List of Figure

List of Tables

List of Abbreviations

AC	Asphalt Concrete
APWA	American Public Works Association
ASTM	American Society for Testing and Materials
DV	Deduct Values
DFT	Discrete Fourier Transform
FFT	Fast Fourier Transform
M&R	Maintenance and Rehabilitation
PCI	Pavement Condition Index
PMS	Pavement Management System
PI	Priority Index
VB	Visual Basic Language

An Empirical Study to Estimate PCI Using Mobile Sensing and Artificial Intelligence

By
Saeed Riad Ali
Supervisor
Dr. Amjad Issa
Co-Supervisor
Dr. Eehab Hijazi

Abstract

Palestinian road and transportation agencies are facing a monumental challenge in dealing with aging infrastructure. For pavements in particular, it is found that many streets were built 20 or 30 years ago and they are near the end of their economic life. Other streets have been deteriorated because of misuse, overuse and mismanagement. In addition, recent and future threats affect the hoped mission of these pavements for rapid, safe and comfort movements of people and goods. Moreover, the current management reveals that the system used is not flexible enough to reflect the changing conditions and poor to assist in making decisions since it needs a huge number of experts to reflect the condition in appropriate time.

This study aims to build an integrated pavement management system using smartphones built in sensors to produce the Auto-phone surveying method. This study provides a model to recognize the type of distress automatically and also implements a formula to calculate the value of PCI by making field data collection using a vehicle containing a smartphone. The system also comprises a set of tools to facilitate a more flexible approach that will enable the users to perform tasks more economically, effectively and of higher quality.

The Auto-Phone method is presented to facilitate the decision-making process. It is based on the direct integration between pavement management system (PMS) database accommodated in Microsoft Excel and Orange software in order to fully exploit the capabilities of each individual package.

The Auto-Phone model was generated and tested by a case study in which the system was fed by 60 km of paved roads (40 Km for developing the formula and 20 for testing) from Nablus City. Finally, condition and prioritization analyses were successfully performed to determine types of distresses and PCI values for streets to be used in maintenance needs.

Chapter 1

INTRODUCTION

Pavement represents a critical component of the highway transportation infrastructure. Asphalt is the most important part of the highway structural layers as it deals directly with the road users. Municipalities and ministries dedicate huge budgets annually on maintenance and rehabilitation road projects. The rapid increase in the number of vehicles year by year leads to increase pavement deterioration. For example, in the West-Bank in 2017, the number of vehicles that were permitted and working is 325,638 vehicles which forms about 10.1% greater than the number of vehicles permitted and working in 2016 (Palestinian Ministry of transportation, 2017). Moreover, the limited fund specified for pavement maintenance in Palestine represents a huge challenge to achieve the needed pavement maintenance and rehabilitation activities.

Pavement Management System (PMS) is a tool or a systematic method that can provide an inclusive inventory for pavement network and organize the work with saving time and effort. The system also provides the data that refers to the current condition of the pavement network with the ability to store the historical data which helps to predict the future pavement condition. In addition, the system can evaluate the pavements and find out a desirable maintenance needs with priorities under the available funds (Shahin, 2005).

1.1 Background

A perfect maintenance scheme on road networks is one that keeps all sections at a sufficiently high level of functional and structural condition (Agarwal, et al., 2004). Due to the increasing traffic on roads, a timely repair that is often critical is constrained by time, budget and other resource availability such as manpower and equipment. This makes a priority ranking scheme for the selection and scheduling of pavement sections for maintenance an essential dimension for study, and an integral part of pavement maintenance management systems (Fwa & Chan, 1993).

According to Kulkarni and Miller (2002), in the early 1970s, pavement maintenance management systems (PMMSs) were introduced and they have evolved continuously in terms of their scope, methodology, and application. The authors described these systems by evaluating the past and current practices and identified future directions for the key elements (Kulkarni & Miller, 2002).

In the early 1980s, the first system that considered the network perspective was developed for the Arizona Department of Transportation (Kulkarni and Miller, 2002). However, systems which were developed again in the 1990s utilize integrated techniques of performance prediction, network-level and project-level optimization, and Geographic Information Systems (GIS) (Kulkarni & Miller, 2002).

PMS has lots of definitions and the following paragraphs illustrate some of these definitions:

A PMS is defined as a set of activities including the planning and programming of investments, maintenance, design, construction, and the periodic assessment of the performance (Hudson & Fernando, 1983). The PMS performs management at all levels and involves comparing alternatives, coordinating activities, making decisions and seeing that they are implemented in an efficient and economical manner to achieve goals. Pavement management is a systematic method for routinely collecting, storing, and retrieving the kind of decision-making information needed to make maximum use of limited maintenance (and construction) dollars (APWA, 1993).

A PMS can also be defined as a set of tools or methods that assist decision-makers in finding optimum strategies for providing, evaluating, and maintaining pavements in a serviceable condition over a period of time. In addition, the system can produce a priority sequence for the evaluated pavements according to maintenance needs with the available funds (AASHTO, 1993).

(SMEC, 2011) gives a third definition for PMS as a planning tool that is able to model pavement and surface deterioration due to the effects of traffic and environmental ageing. It can be used to determine long-term maintenance funding requirements and to examine the consequences on network condition if insufficient funding is available.

In Palestine, there is, generally, no scientific and unified PMS, in which the road conditions are documented and evaluated by the engineers in a systematic way based on scientific mechanisms. The available pavement management practices in Palestine are based on the individual engineering experience only. Also, the maintenance process is carried without defining maintenance priorities in terms of a list of criteria such as road or pavement conditions, citizens' complaints, road classification or road importance.

Pavement Condition Index (PCI) is a numerical indicator that rates the surface condition of the pavement. The PCI provides a measure of the present condition of the pavement based on the distress observed on the surface of the pavement, which also indicates the structural integrity and surface operational condition (localized roughness and safety) (Shahin, 2005).

One of the most recent and modern technologies used in identifying pavement distresses is the use of smartphone applications. Modern smartphones have several kinds of sensors. The most popular sensors which most smartphones have are accelerometer, gyroscope, magnetometer, microphone, and camera (Song, 2013).

An accelerometer measures proper acceleration, which is the acceleration it experiences relative to free fall and is the acceleration felt by people and objects. In other words, at any point in spacetime the equivalence principle guarantees the existence of a local inertial frame, and an accelerometer measures the acceleration relative to that frame. Such accelerations are popularly measured in terms of g-force (Tinder, 2007).

Gyroscope is a very sensitive device which is good at detecting the spin movement. Same as accelerometer, gyroscope returns three-dimensional values. The value gyroscope return is angular velocity which indicates how fast the device rotates around the axes (Song, 2013).

A magnetometer is a measuring instrument used to measure the strength and perhaps the direction of magnetic fields (Jain, 2012). Accelerometer and gyroscope are able to detect the direction of a movement; however, the direction is a relative direction; it obeys the coordinate system that a smartphone uses. Sometimes, different smartphones need to synchronize their directions; therefore, a magnetometer is needed to get an absolute direction (the direction obeys the coordinate system of earth) (Song, 2013).

When drivers are sitting in a vehicle, their smartphones are able to measure the acceleration, velocity, and turns through sensors embedded. Because the smartphone is very popular, using smartphone is an easy way to implement mobile sensing (Song, 2013).

Sensors are the key factors of developing more and more interesting applications on the smartphones, and the sensors make the smartphone different from conventional computing devices like computers. Most applications used accelerometer and gyroscope because they are somehow the most accurate sensors. However, the vision contains huge information. It is believed that camera and pattern recognition will be used more and more in the future (Song, 2013).

Although efforts have been done to detect road conditions, a number of methods/systems have been developed using Smartphone sensors and some of these are reliable, real time road distresses detection is so challenging that none of the methods using smartphone sensors can address it alone completely.

1.2 Background on Pavement Management System (PMS)

In the past, pavements were maintained, but not managed. The pavement engineer's experience tended to dictate the selection of Maintenance and Repair (M&R) techniques with little regard given to life cycle costing nor to priority as compared to other pavement requirements in the network. In today's economic environment, as the pavement infrastructure has aged, a more systematic approach to determining M&R is needed and priorities is necessary. Now, pavement networks must be managed, not simply maintained. (Shahin, 2005)

Recent developments in microcomputers and pavement management technology have provided the tools needed to manage pavements economically. A Pavement Management System (PMS) provides a systematic, consistent method for selecting M&R needs and determining priorities and the optimal time of repair by predicting future pavement condition. The consequences of poor maintenance timing are illustrated in Figure 11 which mean that deterioration of pavement is low until its rating is fair then it deteriorates rabidly and if M&R is performed during the early stages of deterioration, before the sharp decline in pavement condition, over

50% of repair costs can be avoided. In addition to cost avoidance, long periods of closure to traffic and detours can also be avoided. A PMS is a valuable tool that alerts the pavement manager to the critical point in a pavement's life cycle. (Shahin, 2005)

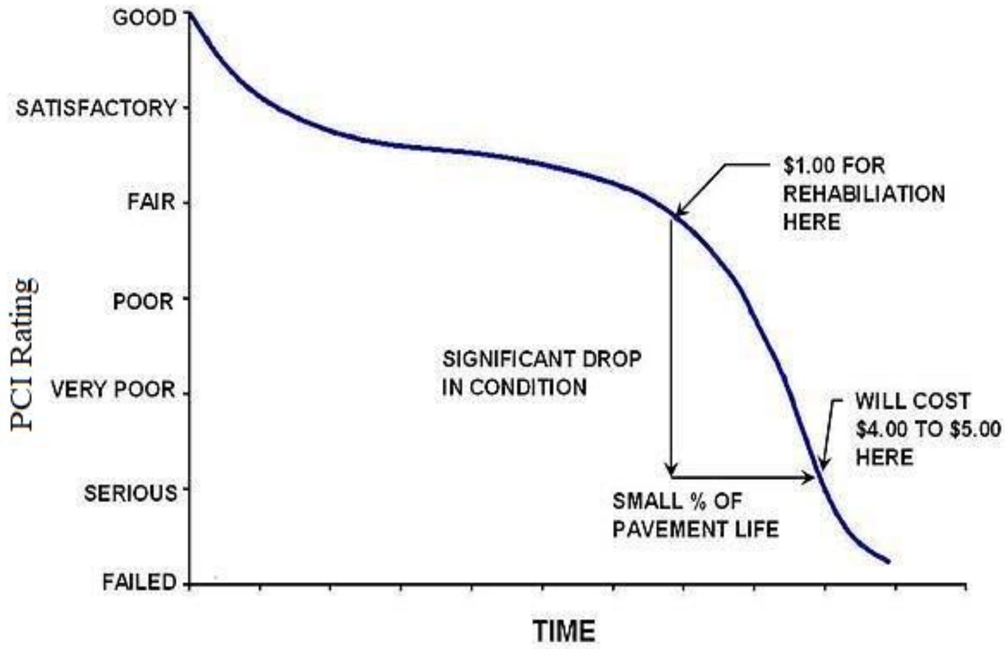


Figure 1.1: Conceptual illustration of a pavement life cycle (AASHTO, 2001)

- **Pavement Management: Levels and Process**

The purpose of road management is to improve and maintain the existing road network to enable its continuous use by traffic in an efficient and safe manner. Road management can be seen as a process that it attempts to optimize the overall performance of the road network over time.

1-Pavement Management Levels

Pavement management has two main levels: project and network levels. The project level focuses on creating the competent use of budgetary resources for a selected pavement segment that has been recognized for rehabilitation. On the other hand, network level focuses on the whole pavement network. Network level includes the assessment of all pavements below an agency jurisdiction. The analysis of this level is best utilized for complete budget estimates and projected considerations. The network level requires aggregated information. Thus, this level has more interest to be used by the manager.

Project level focuses on a particular pavement segment and normally comes afterward network level in local agencies. This level is a sequence of steps used to find out the cause, extent of pavement deterioration and analyzing life cycle cost.

2- PMS Process

The implementation of a PMS for a selected pavement network is carried out through a systematic operation which contains several tasks on a periodic basis. This system is used internationally with a very little difference, as mentioned in the following steps:

- **Network Definition**

The primary step in establishing a PMS is the network identification. A network is a consistent combination of pavements for M&R management. The pavement manager could be responsible for managing the pavement. The pavement network firstly, must be divided into branches and then into a unique section. A section can be defined as a smallest management unit while considering the selection of M&R treatments. Several factors should be taken into account as dividing branches into sections; these factors are pavement structure, traffic, construction history, surface type, and pavement condition (Shahin, 2005).

In this research, the road network is defined and divided according to the method used in the operation and maintenance manual for the roads of the West Bank, which will be explained in detail at a later stage.

- **Pavement Inventory**

Pavement inventory is the basis step of any PMS. usually contains the physical characteristics of the pavement sections and normally these data do not change maintenance actions. The main intention of the inventory is to provide data for identifying the pavement's physical features. The minimum information needed for establishing pavement inventory are listed below (WSDT, 1994):

- Pavement section ID and name.
- Starting and ending location for each pavement section.
- Functional classification.
- Number of lanes.
- Pavement rank.
- Pavement surface.
- Pavement thickness.
- Pavement width.

- Pavement length.
- Pavement surface area.
- Construction date (last surface).
- Average Daily Traffic (ADT).
- Weather conditions.

It is important to mention that the precise type of inventory data needed relies on the agency and the PMS software necessities. Sometimes in inventory data collection, more information is being collected such as: drainage condition, sidewalk condition, and number of traffic signs which may be used at project level that usually comes after network level. The following attached form from ASTM D6433 is used in this research for distress data collection as illustrated in Figure 12.

ASPHALT SURFACED ROADS AND PARKING LOTS
CONDITION SURVEY DATA SHEET
FOR SAMPLE UNIT

SKETCH:

BRANCH__________ SECTION________ SAMPLE UNIT________
SURVEYED BY__________ DATE________ SAMPLE AREA________

1. Alligator Cracking
2. Bleeding
3. Block Cracking
4. Bumps and Sags
5. Corrugation
6. Depression
7. Edge Cracking
8. Jt. Reflection Cracking
9. Lane/Shoulder Drop Off
10. Long & Trans Cracking
11. Patching & Util Cut Patching
12. Polished Aggregate
13. Potholes
14. Railroad Crossing
15. Rutting
16. Shoving
17. Slippage Cracking
18. Swell
19. Weathering/Raveling

DISTRESS SEVERITY	QUANTITY										TOTAL	DENSITY %	DEDUCT VALUE

Figure : 1.2 Inventory she et (ASTM D6433)

- **Pavement Condition Assessment**

After preparing the pavement inventory for the whole network pavement condition, evaluation can be set out. Pavement inspection is one of the vital steps in PMS that encompasses distresses survey. The inspection can be carried out manually or utilizing automated data collection vehicles. The vehicle may comprise cameras, profiling devices, and laser sensors. The collected data are changed to a tape for more processing, either by a software program or by individuals (WSDT, 1994).

Visual inspections are usually carried out manually by an engineer walking through the whole sections or by driving a vehicle at slow speeds and stopping from time to time. Data collection by walking is more accurate than driving, but it has more costs and needs longer times.

- **Pavement Performance Definition:**

The pavement performance has been defined as the serviceability history of the pavement surface and this requires a determination of the types and causes of distress, as well as the extent of pavement deterioration (AASHTO, 2001). Recent literature classified pavement performance as; Functional Performance, Structural Performance and Overall Performance. Those three types of performance were defined as follows (Easa. S and Kikuchi. S, 1989):

A. **Structural Performance:** This is a measure of pavement condition in terms of the appearance of various forms of distresses and their relative importance in triggering pavement preservation actions. This type of performance reflects the owner (or agency) concern. Measures used for these types of performance such as: Surface Distress Index (SDI), Distress Rating, Pavement Condition Rating, and Pavement Condition Index (PCI), which would be adopted later in this research (Easa. S and Kikuchi. S, 1989).

B. **Functional Performance:** This is measuring the quality of finding conditions which is an assessment of how the pavement will serve the public travelling. Mainly the index used for measuring the performance is a roughness index, and it goes under different names such as: Present Serviceability Index (PSI), Pavement Rideability Index (PRI), Riding

Comfort Index (RCl) AND International Rough Index (IRI) (Easa. S and Kikuchi. S, 1989).

C. **Overall Performance:** That is a composite measure of structural and functional performance Indicators. Overall performance Indicators have been designated by several names including Pavement Serviceability Index (PSI) and Pavement Condition Rating (PCR). Each of them is a composite Indicators of roughness and distress. Another type of Index is Pavement Quality Index (PQI) which is a composite index of roughness, distress, and deflection which have been also introduced (Easa. S and Kikuchi. S, 1989).

The evaluation of pavement condition includes consideration of specific problems that exist in the pavement. This requires a determination of the types of causes and distresses, as well as the extent of pavement deterioration. It is also important that condition surveys would be conducted after new construction or rehabilitation work. Such monitoring is a tool for network assessment and it provides information regarding the rate of distress buildup. These survey results are major input when determining whether to undertake a major rehabilitation project (AASHTO, 2001).

From the discussion above, it can be concluded that the necessity of pavement performance indicators is to monitor and control the performance of pavement in each step of its life period. On the other hand, this pavement performance indicator has to be developed in order to match the nature of each country and sometimes the nature of each city, as pavement deterioration can happen due to environmental factors. Moreover, the type of traffic that is using the network whether this city is consider.

- **Using (PCI) as a Pavement performance indicator**

In the 1950s, pavement condition ratings were carried out by a panel of raters who led along the pavement and personally rate the condition of the pavement on a numerical scale or a verbal description. This form of rating, developed by the American Association of State Highway Officials (AASHO), used a 0-5 scale and it was known as the Present Serviceability Rating (PSR). Despite the fact that this was simple, the evaluations did not give sufficient designing premise to endorsing the sort and degree of repair and rehabilitation work to be done on harmed pavements. To deal with this issue, researchers formed scientific that were able to give the condition of pavement sections based on the type, severity and extent of distresses. This led to the development of a more objective means of condition rating in the late 1950s. This index, known as the Present Serviceability Index (PSI) was based on the relationship between panel ratings and measurements such as rutting and roughness (Lee, 1998). The PSI is calculated as shown in $^{PSI}=5.03 - \text{Log}(1+SV) - 1.38(RD)^2 - 0.01(C+P)^{1/2}$ Equation 1. This provided an index that can be calculated from objective measurements of roughness, cracking, patching and the slope variance of the pavement section under consideration.

$$PSI=5.03 - \text{Log}(1+SV) - 1.38(RD)^2 - 0.01(C+P)^{1/2} \qquad \text{Equation 1}$$

Where, PSI= Present Serviceability Index.
SV: Slope Variance of section obtained using CHLOE Profilometer.
RD: Mean Rut Depth (cm).
C: Cracking (m/1000 sq. m).
P: Patching (sq. m/1000 sq. m).

The PSR and PSI were widely accepted among several states. However, during the late 1960s, pavement condition was investigated in 1981 by US Army Corps of Engineers (Shahin, 2005). The study resulted in the development of a single rating number called Pavement Condition Index (PCI) using the PAVER method to represent the condition of the pavement. PAVER is one of the most detailed distress evaluation method implemented to data. This method depends on detailed visual inspection of up to 19 different pavements distresses for flexible pavement. The PCI is a numerical index, ranging from 0 for a failed pavement to 100 for a pavement in perfect condition. It measures pavement structural integrity and surface operational condition. The essential concept behind PCI is to consider both distress extent and value as a negative deduct on the pavement condition The PCI index uses only one pavement condition parameter which is distress types in determining the pavement condition index. The PCI calculation is illustrated in Figure 13 and Equation 2 (Shahin & Kohn, 1981).

$$PCI = C - \sum \left(\sum a (Ti, Sj, Dij) \times Dij \right) x F \qquad \text{Equation 3}$$

Where, Ti: Distress type.
Sj: Severity level.

Dij: Density of distress.

C: Constant (usually 100).

a: Weighing factor.

F: Adjustment factor for multiple distress.

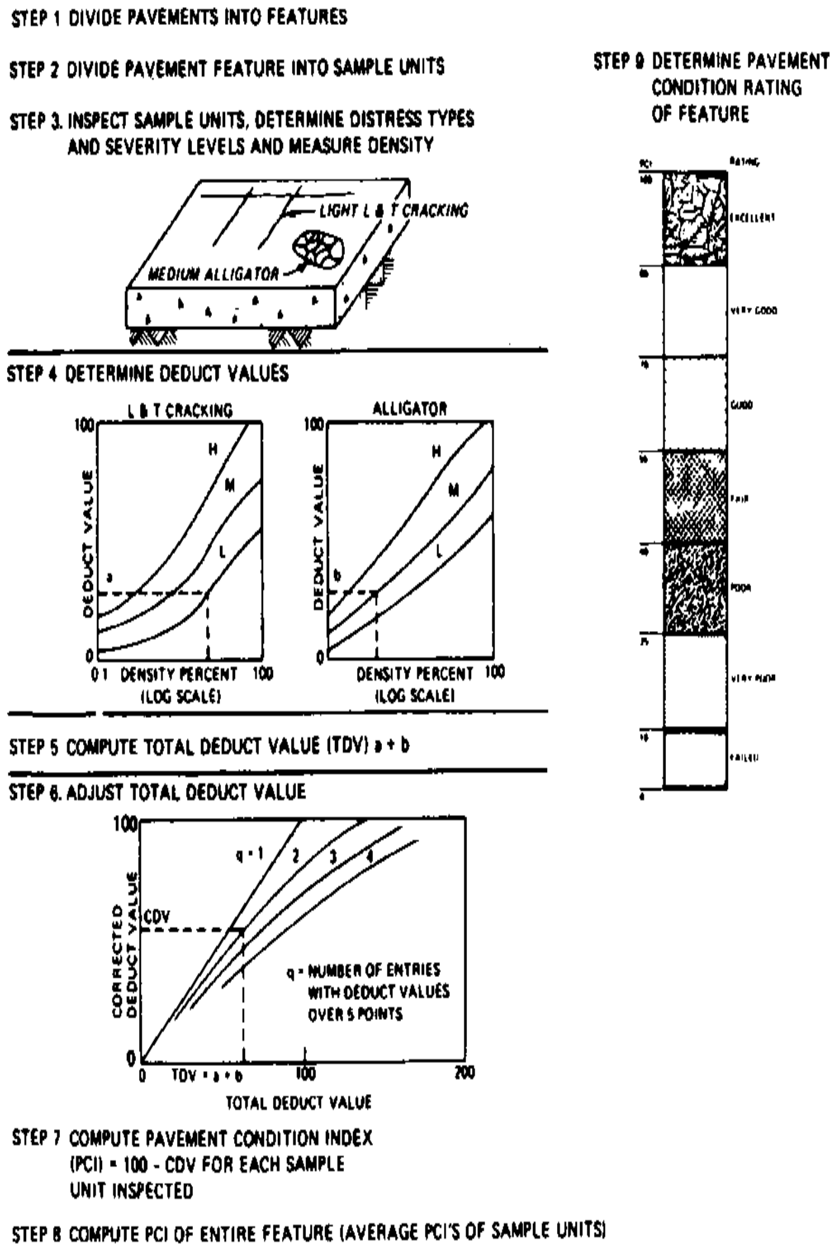


Figure1.3: Steps for determining PCI of a pavement section (Shahin & Kohn, 1981(

- **Pavement Distresses**

The PCI method considers a total of 19 distress types, each with three levels of severity and density. Following is a brief description of each distress type (Shahin, 2005):

1. **Alligator Cracking:** it is a series of interconnecting cracks caused by fatigue failure of the asphalt concrete surface under repeated traffic loading. It occurs only in areas subjected to repeated traffic loading, such as wheel paths. It is considered a major structural distress and is usually accompanied by rutting.
2. **Bleeding:** it is a film of bituminous materials on the pavement surface which creates a shiny, glasslike, reflecting surface that usually becomes sticky. It occurs when asphalt fills the voids of the mix during hot weather and then expands onto the pavement surface.
3. **Block Cracking:** block cracks are interconnected cracks that divide the pavement into approximately rectangular pieces. It is normally occurring over a large proportion of pavement area. But sometimes it will occur only in non-traffic areas.
4. **Bumps & Sags:** bumps are small localized, upward displacements of the pavement surface. Sags are small abrupt, downward displacement of the pavement surface.
5. **Corrugation:** it is a series of closely spaced ridges and valleys occurring at fairly regular intervals usually less than 3m along the

pavement. It is usually caused by traffic action combined with an unstable pavement surface or base.

6. **Depression:** is a localized pavement surface area with elevations slightly lower than those of the surrounding pavement. It is created by settlement of the foundation soil or it is a result of improper construction.
7. **Edge Cracking**: Edge cracks are parallel to and usually 0.3 to 0.6m of the outer edge of the pavement. It can be caused by frost weakened base or sub-grade near the edge of the pavement.
8. **Joint Reflection Cracking:** It occurs only on asphalt-surfaced pavements, which have been laid over a Portland concrete cement PCC slab. It is not load related and mainly caused by the thermal or moisture induced movement of the PCC slab beneath the AC surface.
9. **Lane/Shoulder Drop Off:** It is the difference in elevation between the pavement edge and the shoulder. It is caused by shoulder erosion, shoulder settlement or by building up the roadway without adjusting the shoulder level.
10. **Longitudinal and Transverse Cracking**: Longitudinal cracks are parallel to the pavement’s centerline or laydown direction. It may be caused by (1) a poorly constructed paving lane joint, (2) Shrinkage of the AC surface due to low temperatures or hardening of the asphalt and/or (3) daily temperature cycling, or by cracking beneath the surface course. Transverse cracks extend across the pavement at approximately

right angles to the pavement centerline or direction of laydown. Conditions (2) or (3) above may cause it.

11. **Patching and Utility Cut Patching:** A patch is an area of pavement, which has been replaced with new material to repair the existing pavement. It is considered a defect no matter how well it is performed (a patch area or adjacent area usually does not perform as well as an original pavement section.) Generally, some roughness is associated with this distress.
12. **Polished Aggregate:** When the aggregate in the surface becomes smooth to the touch, adhesion with vehicle tires is considerably reduced. It is indicated when the number on a skid resistance test is low. It is caused by repeated traffic applications.
13. **Potholes:** Potholes are small (usually less 0.9m in diameter), Bowl-shaped depression in the pavement surface. They generally have sharp edges and vertical sides near the top of the hole. Potholes are generally structurally related.
14. **Railroad Crossing:** They are depression or bumps around and/or between tracks. If the crossing does not affect ride quality, it should not be counted.
15. **Rutting:** It is a surface depression in the wheel paths. It is usually caused by consolidated or lateral movement of the materials due to traffic loads.

16. **Shoving:** It is a permanent, longitudinal displacement of a localized area of the pavement surface caused by traffic loading. This distress normally occurs in unstable liquid asphalt mix pavements.
17. **Slippage Cracking:** They are crescent or half-moon shaped cracks having two ends pointing away from the direction of traffic. It occurs when there is a low strength surface mix or poor bond between the surface and the next layer of the pavement surface.
18. **Swell:** It is characterized by an upward bulge in the pavement surface. It is usually caused by frost action in the sub-grade or by swelling soil.
19. **Weathering and Raveling:** They are the wearing away of the pavement surface caused by the loss of asphalt or tar binder and dislodged aggregated particles. These distresses indicate that either the asphalt binder has hardened appreciably or that a poor-quality mixture is present. **Appendix (A)** shows distress details in a more comprehensive way.

- **Prioritization**

Priority ranking are methods used to give a rank for all the pavement sections in a series for maintenance and repair. Prioritization is one of the main steps in PMS used in order to take the decision on the repair and maintenance program.

Priority setting techniques as used in the PMS cover a wide spectrum of methods and approaches ranging from simple priority lists based on engineering judgment to complex network optimization models as shown in Table 11 (Haas, et al., 1994). These prioritization methods can be further divided as:

1. Ranking Methods
2. Optimization Methods
3. Artificial Intelligence Techniques
4. Analytical Hierarchy Process Method

Table 1.1: Different classes of priority methods (Haas et al., 1994)

	Class of Method	Advantages	Disadvantages
1	Simple subjective ranking of projects based on judgment, overall condition index, or decreasing first cost	Quick, simple;	Subject to bias and inconsistency; may be far from optimal
2	Ranking based on parameters, such as serviceability or distress; can be weighted by traffic	Simple and easy to be used ;	May be far from optimal, particularly if traffic weighting not used
3	Ranking based on condition analysis and traffic, with economic analysis	Reasonably simple;	May be closer to optimal
4	Annual optimization by mathematical programming model for year-by-year basis over analysis period	Less simple;	May be close to optimal; effect of timing not considered
5	Near optimization using heuristics including benefit-cost ratio and marginal cost effectiveness	Reasonably simple; Suitable for microcomputer environment;	Close to optimal results
6	Comprehensive optimization by mathematical programming model taking into account the effects of M, R&R timing	Can give optimal program (maximization of benefits)	Most complex and Computationally demanding;

A short description for each ranking method is introduced in this research, a composite priority index will be used because it is compatible with the available data and it is the simpler practical approach. Then, from the following methods, the method that gives a closest result to the optimal will be chosen.

- **Ranking Methods**

1- Composite Index Ranking Method

The ranking of pavement sections for maintenance is performed on the basis of the priority index calculated by combining different pavement indices. These indices are estimated by considering parameters like pavement distresses, riding quality, traffic conditions, economic analysis, functional class, accident details, geometric deficiencies, structural capacity, skid resistance, pavement age, engineering judgment... etc. The prioritization program should not only be based on current pavement condition, for making it efficient future pavement condition should also be considered. As generally, the treatments are applied at least one year after the condition surveys are carried out, giving the time to require for relative prioritization and organizing for funds. (Butt, 1995).

2- Economic–Based Methods

The prioritization methods based on economic analysis can be of two types: (1) using optimal benefit/cost ratio (2) using incremental benefit/cost ratio. In the first method, prioritization process uses the optimal M&R recommendations and corresponding benefit/cost ratios (or effectiveness/cost (E/C) ratio) for each pavement section of the network produced from the dynamic programming. The higher the E/C ratio of a section, the higher the priority of that section for repair. The available budget is allocated to the pavement sections as per the priority list till the budget is completely exhausted. The second method is a heuristic method for budget optimization. In this method all feasible M&R alternatives of a section are identified and the corresponding inflated initial cost, present-worth costs, and weighted benefits are obtained. This information is then used in the program to produce optimal M&R recommendations for each pavement section, including initial cost and type of treatment. The budget optimization also gives the total network-weighted benefits corresponding to optimal M&R recommendations (Butt, 1995).

the Composite Index Ranking Method is adopted in the research in order to arrange pavement section priority in terms of maintenance. The selected method is simple in calculation, easy to be applied, data base needed is available and the results are easity explained to decision. This approach was used in operation and maintenance O&M manual prepared for the benefit of the municipal Development and lending fund (MDLF). The MDLF distributed the manual among the municipalities to rank the pavement sections and arranging them in descending order based on PI value.

1.3 Statement of Problem

Road quality is an important factor in the economic growth of any country (Acquah & Fosu, 2017). Pavement surface is one of the main components of the transportation infrastructure and greatly affects the comfort, costs and safety of road users. Deterioration and catastrophic failure of these roads may occur because of aging, overuse, misuse and/or mismanagement. Therefore, their maintenance and preservation should have a great national interest.

According to the Ministry of Public Works and Housing (PCBS, 2018), the road network in the West Bank has a total length of 3,847.1km, **Table 12** shows the road classification and length in West Bank (PCBS, 2018).

Table 1.2: Road Network in the West Bank (PCBS,2018)

West Bank	**Paved Roads**				**Un Paved Roads**	**Total**
	Main	Regional	Local	Total		
	676.0	1148.6	1575.6	3400.2	446.9	3,847.1

Where, the main rural roads are found within main cities. They connect nearly all Palestinian communities and are generally of poor quality. The regional roads are narrow roads mainly within and around built-up areas. They have a total length of 1,148.6 km, with usually 2-lane roads and are rarely wider than 6 meters. Some of the main and regional roads within the West Bank are Israeli bypass roads. These roads are built for Israeli settlers and soldiers so that they can easily bypass Palestinian communities (MPWH, 2015)

In Palestine, transportation agencies are facing a great challenge in dealing with an aging road infrastructure. For pavements in particular, it is known that many streets were built 20 or 30 years ago and they are near the end of their economic life (Jendia & Al Hallaq, 2005). That means the age of these roads are 30-40 years right now. According to MPWH report, 65% of the road network is in fair to bad condition and they cost the national economy between 2% and 5% of the annual Gross Domestic Product GDP (MPWH, 2015).

According to (Jendia & Al Hallaq, 2005), the conventional pavement management system that is currently in use in Palestine reveals that:

- There is a lack of documentation.
- There is no use of database programs in storing and processing system data in road maintenance management.
- The system is not flexible enough to modify work plans and schedules to reflect changing circumstances.
- The system is poor to assist in making decisions.

In Palestine, maintenance and rehabilitation projects are not enough with respect to the number of projects needed since they depend mainly on funding by external agencies. Due to lack of enough maintenance and rehabilitation funds, priority sequence for the roads in the region becomes very important. In order to rank the pavement sections based on existing distresses, an indicator “priority index” (PI) can be used. The main component of the PI equation is the pavement condition surface represented by the Pavement Condition Index (PCI). The PCI is defined and calculated for each road section at regular distance (100 m length). Moreover, the huge amount of road sections powered by the municipality needs a greater number of experts to make field surveys in order to have a continuously flow of data for all roads.

Transportation engineers who are responsible for conducting both short and long-term transportation planning are facing the following important questions: What is the most effective method to collect data from roads all over the country with lower costs and continuous data gathering? what is the ability to use road users' mobile applications to gather data automatically from the pavement surface and use it to implement special criteria to help in PMS?

According to the above-mentioned, there is a need for a comprehensive pavement management system (PMS) that involves data gathering, storing, processing, analyzing, testing and modification.

1.4 Importance of the Study

The importance of this study lies in the establishment of a new methods of data collection related to use mobile applications that is connected with mobile sensors model called PMS-data mining model to manage the pavement simply based on a unified scientific basis and solve the problems found in the pavement management system in Palestine, which were mentioned earlier in this research.

Using smart phone apps. in defining signature for pavement distresses is expected to develop the current condition in PMS applications at both the local, regional and international levels as illustrated hereafter.

- Locally: Encourage municipalities in Palestine to use PMS tools in their maintenance plans for development.
- Regionally: Replacement of expert's field work and replace it by low-cost data collecting method.
- Internationally: expected to have models and applications to deal with much huge data recorded with a respectively small time.

1.5 Objectives and Scope

Up to now, there are no specific studies related to use mobile applications to give data to be used in PMS in Palestine. This study includes two primary research objectives:

- The first objective is to find and recommend a mobile application to reflect the condition of the pavement surface.
- The second objective is to establish a method to convert the waves created by the application to the reference PCI to simplify calculating the priority index (PI) and giving a rank to the studied roads.
- The third objective is to implement a model to predict the type of distresses or some of them.

Based on the above-mentioned objectives, there is a need to implement a new generalized modified criterion to be used widely in order to detect mainly the four selected distresses: patching, potholes, transverse cracking and alligator cracking, etc.

1.6 Research Organization

The chapters in this dissertation are arranged carefully in the order or sequence of steps to make it clear and understandable. While, Chapter One provides detailed information about the nature of this study, Chapters Two is oriented as a literature review and previous case studies about the main topic of this study, Chapter Three is the methodology and model design. This chapter describes the research methodology adopted in this study. Chapter Four is the application; a case study from Nablus city. This Chapter presents the development of a model to predict PCI value. Finally, Chapter Five summarizes the findings and conclusions of this research as well as the suggested recommendations.

The flowchart in Figure 14 describes the research organization presented in this thesis.

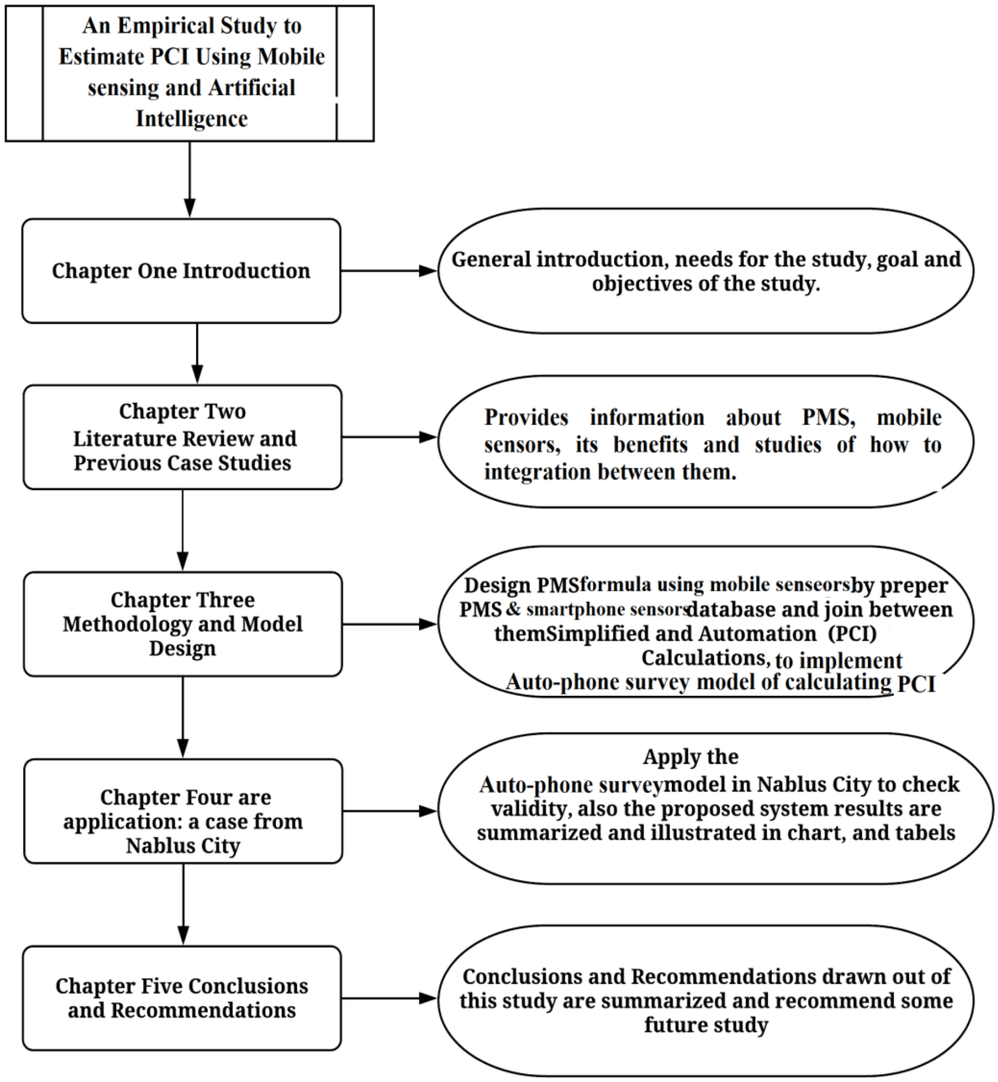


Figure 1.4: Thesis Organization.

Chapter 2

Literature Review

2.1 Introduction

In order to understand the concept of pavement condition and investigating existing pavement conditions, an overview of the related previous and current studies as well as the pavement distresses measuring approaches are introduced in this chapter. The significance of related topics and the potential study topics expected are also presented.

The quality of road pavement significantly affects traveling, as poor conditions which put people’s safety at risk. Over time, the reduction in pavement strength and quality interns of deterioration is mainly due to the impact of wheel pressure on surfaces caused by heavy loads, weather conditions and weakness in subgrade. For these reasons, pavement maintenance has become an important aspect of highway management and infrastructure development in the field of transportation. This chapter covers the related literature and some of the previous case studies about mobile applications and sensors in PMS as well as the suitability of mobile sensors to be used as a tool for detecting pavement distresses to be used in pavement management. Chapter two is divided into three main sections based on the topics that are presented in the research.

2.2 Using mobile sensors in pavement management

In light of rapid development in technology, features and applications in the world of smart phones are increasing dramatically. Most of the smartphones have sensors with an acceptable accuracy. Due to the widespread presence of these smartphones, they have entered into many applications in order to reduce costs and achieve the largest spread. The idea of this research is that smartphone sensors will be used for vehicle users in order to obtain the necessary data related to pavement distresses considering medium-high severity for the process of applying pavement management system.

Although efforts have been done to detect road conditions, a number of methods/systems have been developed using Smartphone sensors and some of these are reliable, real time road distresses detection is so challenging that none of the methods using smartphone sensors can address it alone completely.

Silva, et al.(2018) described a road anomalies detection system based on collaborative mobile sensing that detects and automatically classifies road anomalies using data-mining approaches on data collected by smartphone sensors, more specifically, accelerometer data. The authors also presented a service-oriented architecture that supports data acquisition, data processing, data storage, a road anomalies service and the development of end-user applications in the context of road condition domains. The authors deployed the road anomalies detection system in a real-world environment, in which driving activity was performed without any constraints of speed or road design.

Shahidul Islam (2015) concluded that smartphone application appears to be a viable option to measure pavement roughness. With this application, it would be possible to get most updated pavement roughness values for a roadway network for maintenance and rehabilitation planning.

Chugh, et al., (2014) declared that roads needed to be monitored continuously for roughness and other anomalies to avoid inconvenience to the road users. Road monitoring can also help to predict the estimated arrival time from one place to another. The authors presented a detailed survey of methods for detecting road conditions. From the survey, it was noted that the most commonly used are sensors accelerometer and GPS. Smartphone sensors are gaining importance in this field as they are cost effective and also increase scalability. From analyzing the research activities, it is certain that this area will gain more importance in recent future. There are several research issues that can be explored for improvement in existing methods and develop a highly reliable method.

Schlotjes, et al., (2014) Stated that the remoteness of the Pacific Island Countries (PICs), similar to parts of Africa, creates difficulties, both logistically and economically, to undertake detailed in-country investigations on the road networks. Therefore, rapid assessments of the condition of the existing road pavements are required to determine the level of required donor investments to maintain the integrity of the road network.

This paper explored the use of Roadroid, a simple android application, as a low cost solution to evaluating road roughness in the Pacific region. The case study presented in this paper demonstrated the use of the Roadroid application on the road network in Kiribati, one of the smaller and debatably the most remote PIC. The results from the study discussed the performance and practicability of the android application, primarily as an Information Quality Level-3/4 information device, in the Pacific region. The results from the field surveys supported the delivery of a 21 Information Quality Level 3/4 device. The large variation reported in the surveys between the International Roughness Index collected was attributed to the small sampling intervals embedded in the device. Post-processing of the data which averaged the unfiltered data across one-kilometer sub-sections along the main road in South Tarawa, reduced the variability reported across the road network and provided results consistent with what experienced evaluators expected. Field surveys were conducted with the smartphone device and the data was analyzed post survey. However, the statistical reliability of the device was less satisfactory when the roughness measurements were compared across various speeds. However, within the accuracy limits of an Information Quality Level-3/4 device, considered to be ±20 % of the International Roughness Index, the equipment more than satisfied the need. Roadroid can assist the asset management of road networks by offering a low-cost solution to monitoring and reporting on the roughness condition of pavements in the Pacific region, as well as in other developing regions. Although this paper reported on the performance of the device, further comparison was recommended to confirm the reported International Roughness Index values accurately reflected the condition of the road pavement. To do so, it was recommended to comparatively study the results from the Roadroid android application with those from specialized instrumented vehicles, such as a laser profilometer.

HANS & LARS (2013) Presented that road condition is an important variable not only to measure in order to decrease road and vehicle operating/maintenance costs, but also to increase ride comfort and traffic safety. By using the built-in vibration sensor in smartphones, it was possible to collect road roughness data which could be an indicator of road condition up to a level of class 2 or 3 in a simple and cost efficient way. Since data collection therefore was possible to be performed more frequently, one can better monitor roughness changes over time. The continuous data collection can also give early warnings of changes and damage, enable new ways to work in the operational road maintenance management, and can serve as a guide for more accurate surveys for strategic asset management and pavement planning. Data collection with smartphones will not directly compete with class 1 precision profiles measurements, but instead complement them in a powerful way. As class 1 data is very expensive to collect it cannot be done often beside this advanced data collection system also demand complex data analysis and takes long time to deliver the result. With smartphone based data collection, it is possible to meet both these challenges. A smartphone based system is also an alternative to class 4 – subjective rating, on roads where heavy, complex and expensive equipment is impossible to use, and for bicycle roads. The technology is objective, highly portable, and is simple to be used. This gives a powerful support to road inventories, inception reports, tactical planning, program analysis and support maintenance project evaluation.

The Roadroid smartphone solution has two options for roughness data calculation:

1. Estimated IRI (eIRI) - based on a Peak and Root Mean Square (RMS) vibration analysis – which is correlated to Swedish laser measurements on paved roads. The setup is fixed but made for three types of cars and is thought to compensate for speed between 20-100 km/h. eIRI is the base for the Roadroid Index (RI) classification of single points and stretches (road links) of the road.
2. Calculated IRI (cIRI) - based on the quartercar simulation (QCS) for sampling during a narrow speed range such as 60-80 km/h. When measuring cIRI, the sensitivity of the device can be calibrated by the operator to a known reference.

Collected data were wirelessly transferred by the operator when needed via a web service to an internet mapping server with spatial filtering functions. The measured data can be aggregated in preferred sections (default 100m), as well as exported to other Geographical Information Systems (GIS) or road management system.

By broadcasting road condition warnings through standards for Intelligent Transportation Systems (ITS) the information could provide new kinds of dynamic and valuable input to automotive navigation systems and digital route guides for special traffic etc.

Johnston (2013) detected that over the past decade New Zealand Transport Authority (NZTA) has noticed that road users have been complaining about high levels of ride discomfort despite reports indicating low levels of roughness. This is mainly due to the fact that NZTA is assessing the quality of their roads based on a system developed in the 1980’s. Roadroid is a new roughness measurement application designed to provide cost effective measurements that also monitor the roughness felt by a road user. This research aims to determine whether the Roadroid system can represent the roughness felt by a road user in the Auckland network. Testing was conducted by surveying 20 roads of variable characteristics. The results were compared with industry accepted measurement systems to determine accuracy and wavelength energy to determine response. Results showed Roadroid had an 81% similarity to Laser data and could represent the roughness felt by a road user to a ‘good’ level.

Douangphachanh & Oneyama (2013) carried out an experiment to obtain data. The data was checked and with referenced data. The matched data then cut into sections representing many 100-meter road sections. Analysis was carried out in frequency domain to calculate magnitudes of acceleration data. An adopted group of road condition indexes had been proposed. From the analysis, it had been found that acceleration data from smartphones had linear relationship with road roughness condition. However, the significant of relationship depended on speed in which it was considerably significant when speed was less than 60kph. Furthermore, the relationship also partly depended on vehicle type and device. Based on the condition indexes, similar tendency of the classification of the sum of magnitudes of acceleration vibration was observed.

Bhoraskar, et al., (2012) stated that road traffic is increasing day by day. Monitoring it in an effective way had been challenge to researchers. Since smart phones are penetrating into common people's lives very fast, utilizing the sensors available in them for traffic monitoring is good idea. The data processed by the mobile can be sent to a central server, which can use the information received to annotate maps accessed by the users through this application. This annotation can contain a lot of information like the intensity of traffic at a junction, the bumpy nature of the road, …etc. All this could be done in an energy efficient manner by using low energy consuming components of the mobile like accelerometer and magnetometer, and occasionally using GPS for localization and finding the bearing of the road. Also, applying machine learning techniques in classifying data can help the system to adapt to changing factors like nature of the road and vehicle type the users use.

Johnson & Trivedi (2011) created a system which detected and recorded events that characterize a driver's style, thereby increasing the awareness of potentially-aggressive actions, and further promoting driver safety. The authors also showed that the sensors available in smartphones could detect movement with similar quality to a vehicle Controller Area Network (CAN) bus, and therefore made it a viable and inexpensive utility for vehicle instrumentation.

Mednis, et al., (2011) proposed a system which used Android OS based smart-phones having accelerometer sensor for detection of potholes in real time. This system detected events in real time and also collected the data for off-line post-processing. The data was collected using 3-axis accelerometer sensor presented in smart-phones.

SankarPandi (2020) defined the accelerometer as a built-in component for measuring the acceleration of any mobile device. Motions like swinging, tilting, rotating, shaking is detected using accelerometer. The values of XYZ are used to calculate and detect the motions. Besides mobile, the accelerometer is used to measure vibration on cars, machines, buildings, process control systems and safety installations.

According to SankarPandi (2020), accelerometer in the mobile device provides the XYZ coordinate values, which are used to measure the position and the acceleration of the device. The XYZ coordinate represents direction and position of the device at which acceleration occurred. The rotation direction and position are measured using gyroscope sensors that are found in the Android devices. The mobile device rest in the Earth includes the acceleration due to gravity (g = 9.81m/s2) and the acceleration value. The accelerometer values provided by the device normally includes the gravity as well. Accelerometer along with the linear acceleration and gyroscope will provide results close to accuracy. Linear acceleration does not include the gravity. The accelerometer value is passed into the low/high pass filters for refining the result, based on the application that has been used.

The XYZ values change for every acceleration (for every 20 ms). So, the iteration of values passed into a low/high pass filter will provide the range for walking, running, jogging, etc.

Strazdins, et al., (2011) used an Android smartphone device with accelerometer to detect location of potholes. Their approach included many simple algorithms to detect events in the acceleration vibration data.

Perttunen, et al., (2011) and Tai, et al., (2010) analyzed data obtained by smartphone accelerometers in frequency domain to extract features that were corresponding to road bumps.

Chang, et al., (2009) authorities tested whether an autonomous robot could be used to measure the International Roughness Index (IRI), a description of pavement ride quality in terms of its longitudinal profile. A ready-made robot, the Pioneer P3-AT, was equipped with odometers, a laptop computer, CCD laser, and a SICK laser ranger finder to autonomously perform the collection of longitudinal profiles. ProVAL (Profile Viewing and AnaLysis) software was used to compute the IRI. The preliminary test was conducted indoors on an extremely smooth and uniform 50 m length of pavement. The average IRI (1.09 m/km) found using the P3AT was robustly comparable to that of the commercial ARRB walking profilometer. This work was an initial step toward autonomous robotic pavement inspections. The authers also discussed the future integration of inertial navigation systems and global positioning systems (INS and GPS) in conjunction with the P3-AT for practical pavement inspections. An integrated set of vertical displacement sensors (CCD laser), odometers, SICK laser ranger finder, and control laptop were mounted on the P3-AT, which was manufactured by MobileRobots Inc (2008). The P3-AT, which could move up to 3 km/h, was capable of measuring longitudinal profiles using a CCD laser at 15 cm or smaller sampling 19 intervals, from which the IRI could be simultaneously computed using laptop-based ProVAL software.

González, et al., (2008) used a standalone accelerometer in order to fit in a simulation vehicle and used it to assess road roughness condition. Their simulations concluded that roughness of the road could be estimated from acceleration data obtained from the sensor.

Eriksson, et al., (2008) studied an application of mobile sensing: detecting and reporting the surface conditions of roads. They described the P2 system and associated algorithms to monitor this important civil infrastructure using a collection of sensor-equipped vehicles. P2 used the inherent mobility of the participating vehicles, opportunistically gathering data from vibration and GPS sensors, and processing the data to assess road surface conditions. The researchers deployed P2 on 7 taxis running in the Boston area. They used a signal processing and machine-learning based approach and show that P2 is well suited for detecting adverse road conditions. Via careful selection of training data and features, the P2 detector misidentified road features as having potholes less than 0.2% of the time in controlled experiments. They also evaluated the system on data from thousands of kilometers of “uncontrolled” taxi drives, and found that out of reported detections, 90% contain road anomalies in need for repair.

Mohan, et al., (2008) defined Nericell as a system for rich monitoring of road and traffic conditions using mobile smartphones equipped with an array of sensors (GPS, accelerometer, microphone) and communication radios. In this paper, the authors had focused on the sensing component of Nericell, specifically on how these sensors and radios were used to detect bumps and potholes, braking, honking, and to localize the phone in an energy-efficient manner. The authors had presented techniques to virtually reorient a disoriented accelerometer and to use multiple sensors in tandem, with one triggering the other, to save energy. Their evaluation on extensive drive data gathered in Bangalore has yielded promising results. They have a prototype of Nericell, minus GSM-based localization, running on Windows Mobile 5.0 Pocket PCs.

According to Woodstrom (1990), modern inertial profilometers require four basic sub-systems:

- Accelerometers to determine the height of the vehicle relative to an inertial frame of reference look at Figure .21
- Height sensors to measure the instantaneous riding height of the vehicle relative to a location on the road below the sensor

- Distance or a speed sensor to determine of the position of the vehicle along the length of the road (nowadays combined with GPS) as described in Figure .21
- Computer hardware and software for computation of the road profile

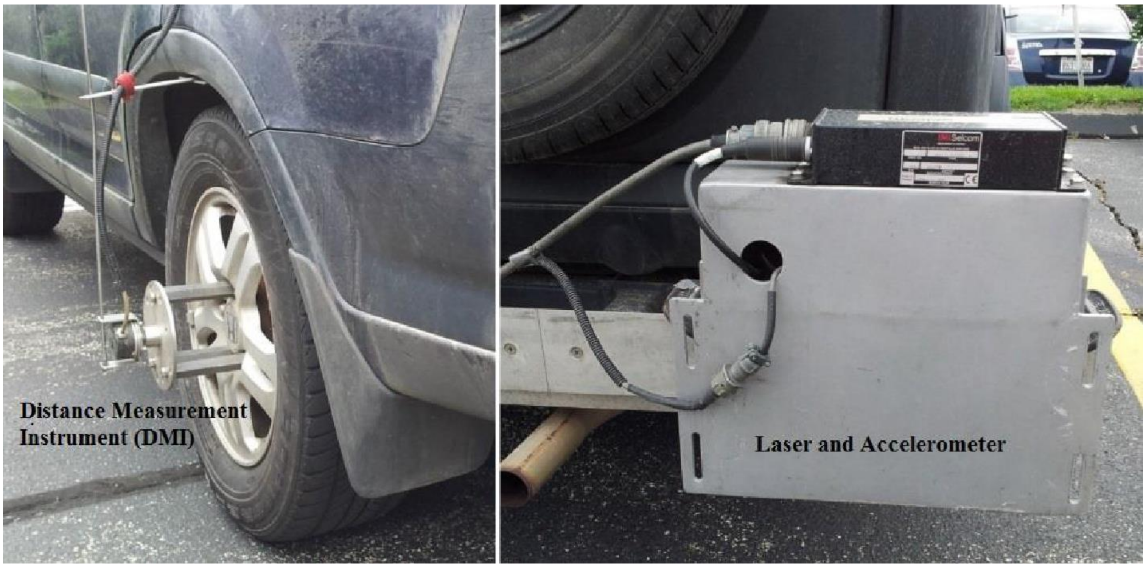


Figure 2.1: Automated Pavement Profiler and Equipment on a Typical Data Collection Van

Perera, et al. (2005) calibration and construction control devices were generally used to check the profile of the new constructed layer which included California profilegraphs “which are shown in Figure 22”, dipsticks, Ames profilographs, and Rainhart profilograph. Profilographs were generally used for construction inspection, quality control, and acceptance of smoothness of concrete pavement. The California profilograph was widely used for over fifty years. There was a rigid beam or frame in profilographs which was supported by wheels at both ends and center. The midpoint wheel moved vertically and a strip chart recorder for capturing the movement of recording wheel. Pavement roughness measured by the profilographs was expressed as Profile Index (PI) in inch/mile. However, the California profilograph can only evaluate 1.9 to 3.1 miles of pavement per hour. It had been reported that profilographs tend to amplify or attenuate true pavement profile.

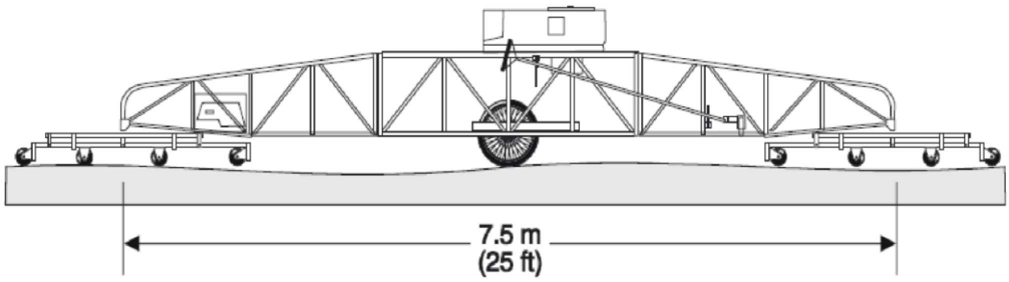


Figure 2.2: California Profilograph

Spangler & Kelley(1964) Stated that the Rolling Straightedge as presented in Figure was one of first methods developed to measure pavement 23 smoothness. A rolling straightedge consisted of a rigid beam having a fixed wheel on each end and a third wheel capable of vertical movement located at the middle of the straightedge. An indicator was attached to the middle wheel which records the deviation of the pavement at the center wheel relative to the plane of the rolling straightedge. However, it had been found that rolling straightedge was not able to capture and reflect longer features that could attribute significant roughness. Moreover, its slow operating speed and manual operation made it incompatible for testing under traffic and safety issues were also concerned. With these operating problems, straightedges were quickly becoming out-of-dated and impractical for general use due to inefficiency and inaccuracy.

Figure 2.3: Rolling Straightedge

2.3 Summary

This chapter explores pavement maintenance management in different agencies in the world. Various aspects of pavement maintenance management were investigated, including pavement maintenance in general, pavement inspection and monitoring practices, and pavement maintenance. Other topics explored in the research include pavement maintenance management practices among local road authorities, particularly their pavement condition assessment, pavement maintenance management systems and major issues affecting pavement maintenance management among local authorities.

Through literature review, all relevant studies were presented in detail considering the main related issues with their merits and demerits. These studies provided a summary of the current design practices concerning established design guidelines, and safety aspects that have been identified. The Major work of most of the authors is pertains to the development of the pavement quality, serviceability and economical to local authorities. Least work has been done on the Smartphone applications and their use in analyzing roads.

The main contribution of this research is to propose a new practical method aimed at measuring pavement distresses through using mobile app and evaluating the effect of both the economical and easier way other than costlier equipped methods. While the contribution over other studies can cleared that there is no previous studies on using mobile sensors to calculate the value of PCI, since all of previous studies related to measure the effect of one destress on the quality of the surface. Another important paradox point that previous studies can detect is only one type of destress, but in this research can define mostly all types of destresses.

Based on the outcomes of the literature review, pavement maintenance management refers to various practices that help in improving and prolonging pavement surfaces in order to achieve both structural and functional performance to avoid accidents and risks to safety as well as traffic and experiences? on the road. Pavement maintenance may be corrective or conducted for emergency purposes. Nevertheless, the most important action during pavement maintenance is prevention. Preventive measures help local road authorities assess road conditions and make plans to increase the life span of pavement and, thereby, reduce spending on pavement maintenance as well as disruptions to traffic. For this reason, local road authorities are recommended to adopt preventive strategy as an aspect of pavement maintenance management.

In Palestine the responsibilities of road maintenance activities are divided between the Ministry of Public Works and Housing (MOPWH) which in charge of all the roads outside the municipal boundaries, and the municipalities in charge of all the roads inside their boundaries. The municipalities operate under the guidance and support of the Ministry of Local Government (MoLG). (Issa & Abu-Eisheh, 2017).

The current road maintenance practices in Palestine are not based on scientific methodological methods. Most of the Municipalities use one criterion for determining maintenance priorities, the "worst first" criterion. And therefore, current practices in determining road maintenance plans do not deal with the generally known stages including those relating to the establishment of road inventories, clear the road paving situation, the overall assessment of pavement conditions, identification of appropriate maintenance rehabilitation measures and priority setting (Issa & Abu-Eisheh, 2017).

Based on the above, there has been a need to develop and adopt appropriate methods to assist in the decision-making process related to maintaining and upgrading the pavement structures of the roadway network in Palestine. This research aims on proposing a systematic approach for road maintenance and identifying a decision-making model that helps the local agency in Palestine in creating their operational and maintenance plans, and in prioritizing their annual maintenance activities based on scientific base by using accessories of the smart phones to help municipalities and ministries in identifying specific pavement distresses using smart phones.

Chapter 3

Methodology

3.1 Introduction

Based on the findings from the literature review, it is essential to develop a method aims to estimate the pavement condition using smart phone applications to manage pavement maintenance effectively. Assuming that different road surface conditions cause vehicles to vibrate differently, therefore, by placing smartphones that come with acceleration sensors, the variation of the vibration is believed to be captured.

Methodology processes pass throw steps. These steps shorten in Figure 31 are clarified briefly in the coming sections.

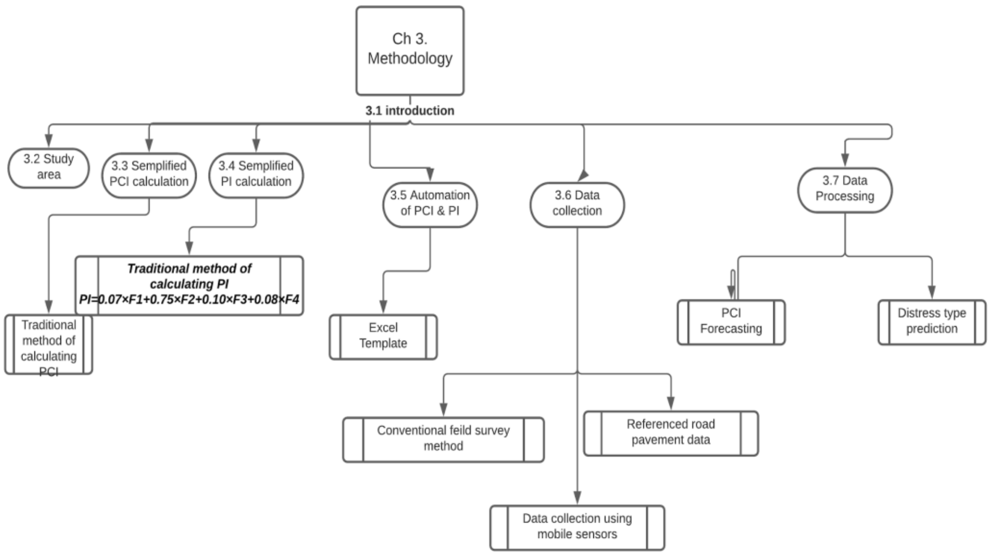


Figure 3.1: Methodology framework chart

3.2 Study area discussion

The mobile sensing is applied in Nablus city and at specific selected roads in Nablus city. The length of the road network adopted in the case study is about 30 km. Mainly; urban arterial roads are to be included. About 600 sections of 100-meter length included in the study. The 600 sections will be divided into two main categories; 400 sections for learning, training and developing the models, formulas and signatures, and the remaining 200 sections for assessing and testing the validation of the output model.

- **Developing the formula study area**

Field investigations and measures will be performed for the selected 400 sections in order to find the PCI for targeted sections taking into consideration the four selected distresses as mentioned before.

The second step is to use the selected vehicle, and to conduct the field driving session on the targeted 400 selected sections in order to collect the needed data using the mobile application and representing the data collected graphically. In order to have a view for the results expected to be produced from this research, preliminary experiments and field tests were conducted. The preliminary experiment conducted using the “AndroSensor” mobile application installed on Samsung Note 8 smart phone fixed in front of the cabin of a Peugeot Partner 2013 car. Accelerometer is the main mobile sensor used in these experiments mainly in X & Z directions. Figure 32 shows the location of the study area. Figure illustrates the boundary of 33 .the study area

Also, Figure shows the developing and testing sections as clarified 33 .below

Roads used to develop the formula are:

1. The ring road; Mohammad Ibn Rashed All Maktoum
2. Tunes street.
3. The part of Nablus - Tulkarm street. Haifa street (between Tunes and Ring road).
4. The connection between Tunes and Ring road that goes parallel to Tulkarm road (Yaser Arafat street).
5. The municipal football court road that begins from the Tunis street and ends to Jamal Abdunnasir park (Jaffa street).
6. Alquds street
7. Amman street
8. Central business district (CBD) area main streets.

- **Testing study area**

Roads for assessing and testing the validation and configuration of the output model are:

1. Rojeeb street.
2. Tulkarm street (Starts from ring road and ends at Qosin intersection).

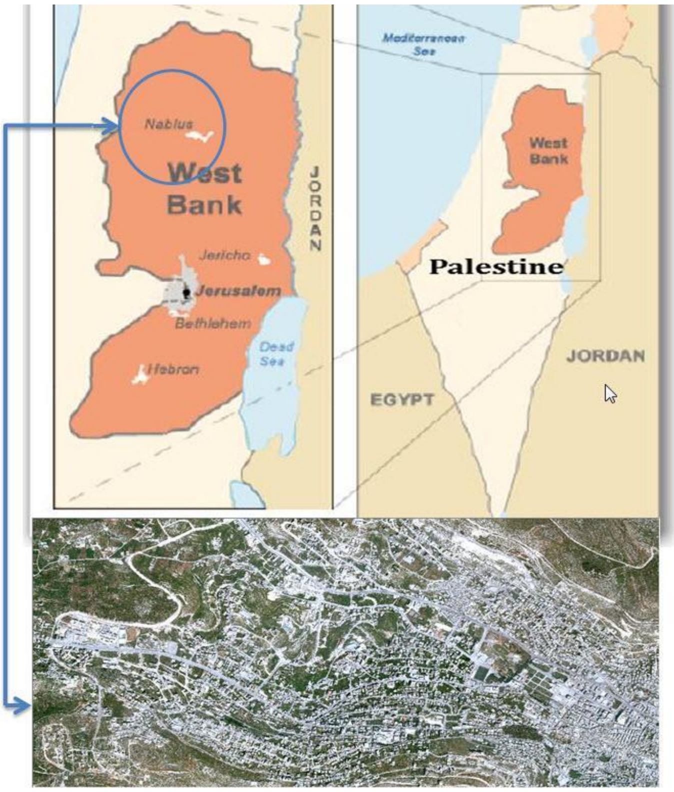


Figure 3.2: Study area (Nablus city)

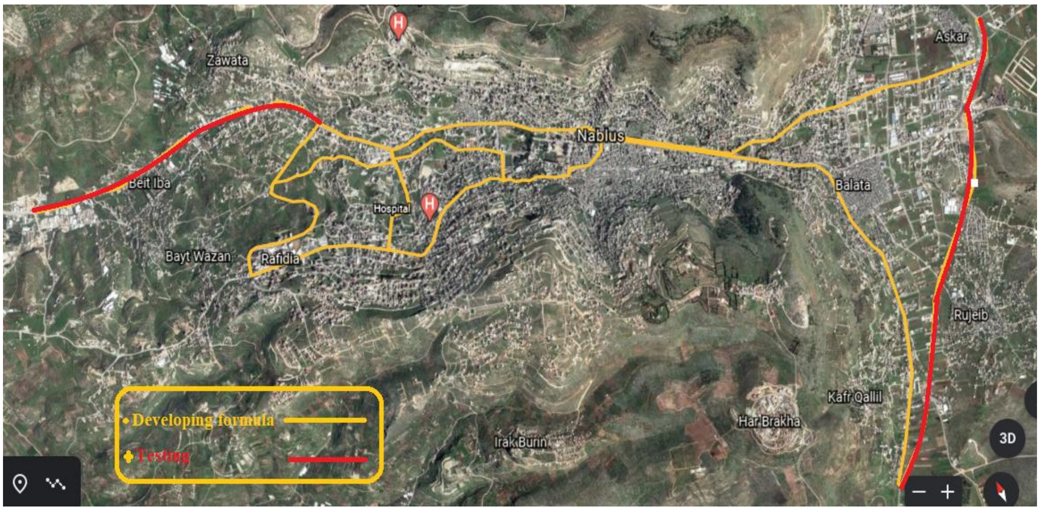


Figure 3.3: Site plan (boundary of the study area)

3.3 Simplified Pavement Condition Index (PCI) Calculation

As discussed earlier, PCI is a numerical index between 0 and 100 values that is used to indicate the general condition of the surface of a pavement section, with 100 value representing the best possible condition and 0 value representing the worst possible condition. The PCI survey procedure and calculation method has been standardized by American Society for Testing and Materials (ASTM) for roads and parking lots pavements (ASTM, 2007). The terminologies defined by ASTM standard are also used in development of automated PCI calculation templates. The coming paragraphs provide some basic definitions for relevant PCI terminologies for calculation used in the ASTM procedure.

Pavement Distress: External indicators of pavement condition deterioration caused by loading, environmental factors, or a combination thereof. Typical distresses are cracks, rutting, and weathering of the pavement surface. Each distress, based upon its effect on pavement performance and riding quality, is classified into three severity levels: Low (L), Moderate (M), and High (H). A completed distress identification manual was provided by Federal Highway Administration (FHWA) in 2003 (FHWA, 2014).

Depending on the type of distress density and severity, the amount of distress within the pavement section is measured either in square meters (square feet), linear meters (feet) or the number of occurrences. For instance, fatigue and block cracking are measured in square feet or square meters, whereas longitudinal and transverse cracking are measured in linear unit.

Pavement Section: A continuous pavement area with a uniform structure, maintenance, usage history, and condition. A section should have similar traffic volume, structure and geometric characteristics.

Distress Density: The percentage to indicate the ratio of distress within an area. It is obtained by dividing the total quantity of each type of distress at each level of the total area of the pavement section.

Deduct Value (DV): Statistical weight number of distresses to determine a combined condition index for pavement sections. According to **ASTM 6433-07**, for each distress type and severity level, there is a distress deduct value curves for deduct value determination (ASTM, 2007).

Corrected Deduct Value (CDV): Adjustment of the cumulative deduct value or the total deduct value (TDV). The CDV adjusts the TDV to fit for a range of 0-100 by using a set of CDV-TDV adjustment curves. The maximum of CDV (maxCDV) is used to calculate PCI (PCI=100-maxCDV). If there is only one deduct value, then the TDV is used in place of the maxCDV in determining the PCI (ASTM, 2007).

3.4 Simplified Priority Index (PI) Calculation

In prioritizing maintenance work for the road network, the Operation and Maintenance Manual identified Priority Indicator (PI) of the four declared indicators of different weights. The dominant indicator of weight is the pavement condition, set to 0.75. The estimated weight in Palestine is 0.75 as a result of the accumulation of road maintenance, which is considered one of the main assets in Palestine. All of the target municipalities have adopted this weight. For each indicator, the remaining weights are less as shown in equation 3.1 shown below (Issa & Abu-Eisheh, 2017).

The PI is calculated using the following equation, as defined in the O&M Manual:

$$PI=0.07\times F1+0.75\times F2+0.10\times F3+0.08\times F4 \qquad \text{Equation 4}$$

Where:

F1, F2, F3, and F4 are factors related to functional classification of roads, pavement condition, Importance of Road to Community, and Citizen’s Complaints respectively.

3.5 Automation of PCI and PI Calculations

The existing ASTM PCI method provides an objective assessment of the pavement condition. In addition, the O&M Manual for Palestinian Municipalities provides a detailed explanation of the method of calculating the Priority Index (PI) and clarifies all the factors that affect the process of prioritizing the maintenance of the road. However, a large labor-intensive work is needed. Because there are a lot of calculations needed to be completed, even for the small pavement network. It is therefore, useful to develop a tool for automating the PCI and PI calculation of road at the project level and network level.

The following sections describe the development of mathematical formulas based upon the available DV curves found in the ASTM 6433-07 procedure and Priority Index (PI) according to the O&M Manual for Palestinian Municipalities; this is followed by describing how these equations are used in an automated PCI and PI calculation Excel template.

- **DV Curves Nonlinear Math Functions**

The family of DV curves provides a reference for manually determining the deduct values. However, in this research, mathematical equations derived through two international studies are used. As previously mentioned in the literature review. A total of 91 nonlinear (multinomial) functions and plots were developed to be used for the determination of DVs. The same approach was used to determine the CDVs, the regression analysis shows the polynomial function between DV and logarithm of density with high degree of accuracy. The plots, regression analysis and nonlinear equations for all of the distresses can be found in Appendix (B).

- **Excel Template for PCI&PI Calculation**

According to ASTM, the procedure used to determine PCI for a pavement section can be divided into the following four steps:

(1) Converting raw data to distress density (%) using area of surveyed section as denominator;

(2) Finding deduct value (DV) using DV-Density graph;

(3) Summing the largest 7 DVs resulting in total deduct value (TDV);

(4) Finding corrected deduct value (CDV) using CDV-TDV graph and PCI equal to 100-CDV.

According to Rabaya (2018) An Excel templet was implemented. The template provides user-friendly transformation of regions to PCI and PI values for each road section data. The format developed was made compatible with the dataset available.

The excel templet uses the field data which is collected manually and entered to the first sheet “field sheet” manually. The field data also programed to be saved automatically. The templet calculates the PCI for each road section and then the overall PCI of the road. PI also is obtained for each road with respect to its area then the priority table is done.

Appendix C shows figures taken from the templet for the processes to obtain PI for the roads specified from A to Z.

PCI and PI values for the roads in general obtained using the weighted mean method. Where, in this method, the contribution in PCI value from each section in the road are calculated based on the area of sections not length because the width of sections not equal for all sections. All sections are calculated for the small sample size to achieve the highest accuracy. Shahin (2005) calculated PCI for the hole road using the PCI vlues of each section in equation 5.

$$PCIroad = (\mathrm{Area1} * \mathrm{PCI1} + \mathrm{Area2} * \mathrm{PCI2} + \mathrm{Area3} * \mathrm{PCI3})/(\mathrm{TotalAreaforroad})$$

$$PCIroad = (\mathrm{Area1} * \mathrm{PCI1} + \mathrm{Area2} * \mathrm{PCI2} + \mathrm{Area3} * \mathrm{PCI3})/(\mathrm{TotalAreaforroad})$$

Equation 5

Where,

Area(i)=width(i)*length(i)

The length almost equal 100 m upon the PCI standard calculation, so the width is the main variable in PCI calculation and this value is taken into consideration in data collection and build PMS database stages.

Also, PI value for the road in general is calculated using the same method as in PCI according to (Shahin, 2005) as in equation 6.

$$PIroad = (Area1 * PI1 + Area2 * PI2 + Area3 * PI3)/(TotalAreaforroad)$$

$$PIroad = (Area1 * PI1 + Area2 * PI2 + Area3 * PI3)/(TotalAreaforroad)$$

Equation 6

3.6 Data collection

Data collection is divided into two sub-titles in order to be more specific. The first one is about the conventional method of collecting data that is being followed to reflect the real pavement condition to be used in order to derive the equations. The second sub-title is about the new contribution of this research using smart phone applications.

- **Conventional field survey method**

Field surveys were conducted for the above-mentioned roads as presented in Figure 33 for about 400 sections. The surveyed roads are:

1. The ring road; Mohammad Ibn Rashed All Maktoum
2. Tunes street.
3. The part of Nablus - Tulkarm street. Haifa street (between Tunes and Ring road).
4. The Connection between Tunes and Ring road that goes parallel to Tulkarm road (Yaser Arafat street).
5. The municipal football court road that begins from the Tunis street and eds to Jamal Abdunnasir park (Jaffa street).

6. Alquds street
7. Amman street
8. Central business district (CBD) area main streets.

While the roads surveyed for testing and configuration are:

1. Rojeeb street.
2. Tulkarm street (starts from ring road and ends at Qosin intersection).

- **Data collection using mobile sensors**

This method basically depends on sensing the pavement surface condition by fixing a smart phone in a moving vehicle. Using a mobile application which measures the acceleration in the three directions while the vehicle still moving and record the data. For this experiment and studies, only acceleration data (x, y, z) from accelerometer; location data (longitude, latitude, speed…) from GPS are needed. Data recording is done at an interval of 0.01 second or at a frequency rate of 100Hz.

The following technical requirements were chosen as a basis for distress detection system:

1. The system should be able to detect events on road surface in real time.
2. The system should use a generic Android OS based smart-phone with accelerometer sensors as the hardware/ software platform.
3. The system should be able to run on different smartphone models with different parameters. During the system implementation process the

set of minimal smartphone parameters should be determined and described.

4. The system running on a smart-phone should be able to perform its native communication tasks at an adequate quality level.
5. The system should be able to detect events while driving in different four-wheel vehicle types such as passenger cars, minivans and buses. Two-wheel vehicles such as motorcycles and scooters are not considered.
6. The system should have a calibration or self-calibration functionality, as different vehicles are likely to yield different sensor data when encountering a pothole for example. This functionality should be based on signal patterns specific to the certain vehicle types.

- **Equipment and software**

The used equipments and applications are as follows:

1. The vehicle used in this research Peugeot partner 2013 car.
2. The smart phone is Samsung note 8 using android system.
3. Mobile app. used is AndroSensor. AndroSensor is an application that can collect data from almost all of the sensors available on the handsets and it is available for free download in Google Play store. This app. can capture the reading of the sensors in the mobile each 0.1 second for the seriously accurate needs. In this step one of the main objectives of this research is achieved since “AndroSensor” is the

recommended mobile application in measuring the values of the sensors.

4. Mobile Handel.

This four equipment’s are illustrated in Figure to Figure 34.37

Figure 3.4: Mobile Handel

Figure 3.5: Peugeot Partner

Figure 3.6: Smart phone used I

Figure 3.7: AndroSensor screen the research

➢ **Equipment Setting**

The smartphone must be fixed using the handle at about the center of the vehicle. Before the movement starts, the GPS must be turned on, the mobile app. AndroSensor also be turned on and clicking on the recording button. Figure Presents the fixation of 38 the smart phone in the vehicle using the previously mentioned equipment's.

Figure 3.8: Fixation of the smartphone in the vehicle

- **Referenced Road Pavement Condition Data**

According to section Conventional field survey method concerning the conventional way of collecting data from field by measuring pavement destresses by hand, the data collected and analyzed in the conventional way are used to be the referenced road pavement condition data in order to use it to build the model.

3.7 Data processing

Data from the AndroSensor which is saved in the internal memory of the smart phone will be uploaded into a computer and and also converted to excel spread sheets. The data in the spread sheets are carefully manually checked to select the complete data sets for road sections.

The data imported from AndroSensor consists of: GPs coordinates, acceleration in X,Y & Z directions, linear speed, and other sensors readings which are not needed in this research.

According to the objectives of the research, data will be in two separated processing lines, the first is for predicting the PCI and the other is to describe the type of distresses appears from the data.

- **PCI forecasting**

The flow chart summarizes the steps to use the row data “which are collected throw the Auto-mobile method” compared with the manually measured PCI values, in predicting the value of PCI. Figure presents the followed flow 39 .chart used in forecasting the values of PCI

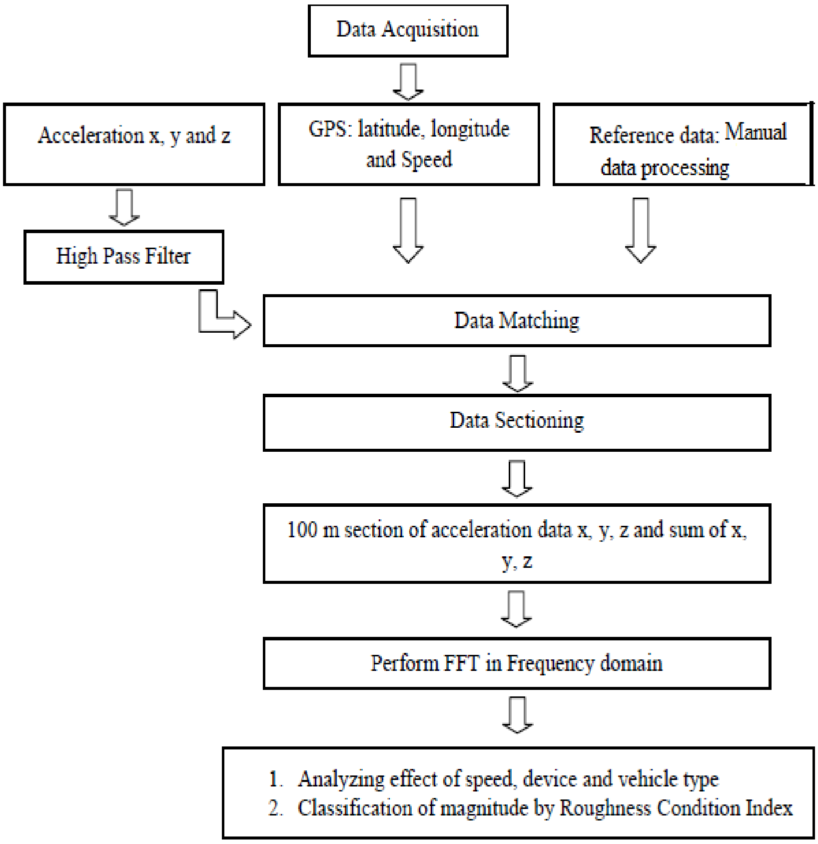


Figure 3.9: Data processing flow chart

The steps of the flow chart are described briefly as follows:

1. A simple high pass filter is applied in order to remove unrelated low frequency signals, which is usually produced by the effect of vehicle maneuver like turning and changing speed as well as the contribution of the force of gravity from all axes (x, y and z) of the acceleration data (Android Developer Official, 2012).

2. Data matching by GPS coordinates is carried out. This process merges PCI data from conventional field method with the acceleration data from smartphones for the same road sections.

3. Data sectioning, the merged data files are cut into small 100-meter sections based on GPS coordinates. A 100-meter length of acceleration data is chosen as a unit for road surface estimation in this study. The reasons are

   For calculation, the PCI based on existing pavement density and severity, the pavement section length is chosen to be 100 m because Road Management System in lows requires road surface condition to be estimated for every 100-meter section, therefore, it would be more convenient for us to select the same unit so that it is compatible for future application.

   a. There is a concern on the accuracy of GPS position data, thus choosing a shorter section unit may cause some issues for data matching between the real Position and smartphone GPS data. In the sectioning process, road sections where experiment vehicles have stopped (checking from speed) are excluded since data at these sections cannot be used to estimate road roughness condition.

4. All selected data sections are converted to frequency domain and perform Fast Fourier Transform (FFT). Fast Fourier transform is one of the most useful tools and is widely used in the signal processing. A FFT is an algorithm that computes the discrete Fourier transform (DFT) of a sequence, or its inverse (IDFT). Fourier analysis converts a signal from its original domain (often time or space) to a representation in the frequency domain and vice versa. The DFT is obtained by

decomposing a sequence of values into components of different frequencies. This operation is useful in many fields, but computing it directly from the definition is often too slow to be practical (Heckbert, Feb. 1995).

The importance of converting acceleration values to frequencies is to minimize the effect of outliers and also minimize the effect of vehicle motor vibrations on the targeted values.

5. Magnitude from FFT is the amplitude or strength of the associated frequency component. For a specific frequency window, it is d that the average of magnitudes represents the strength of the vibration at that frequency window.
6. Therefore, average of magnitudes from FFT is studied to find out features, effect and relationship that the acceleration data might have in connection with road surface condition or pavement condition index.

- **Distress type prediction**

After the acquisition of the first test data set, a search for potential event related features was performed. The emphasis was applied on features that did not require resource-intensive signal processing techniques and therefore were suitable for implementation for detection using devices with limited hardware and software resources.

To predict the types of distresses from the selected distresses (Pot-holes, alligator cracking, Patching and transverse cracking), a series of steps must be followed in order to predict type of pavement distress:

1. Collecting data using the Auto-Phone surveying method for different sections, each section contains a specific type of distress to form a typical data base of the approximate acceleration for each direction axes.

2. The data collected in the previous step is assigned to a data mining program in order to use it as a reference data for each type of distress. Orange 3.24.1 data mining program used in this research. Orange is an open-source data mining program. Orange mainly is a model building program consist of tasks and links all together forms a model (Demsar, et al., 2013). Orange has three main steps which may have other secondary steps or tasks. The main steps are:

   A. Assigning the typical data to the program.

   B. Applying the needed mathematical and statistical operations, and.

   C. Monitoring the results using lots of methods.

3. Statistical prediction tests take place in model building in order to reach the best model to predict type of distress. Also, these statistical tests used to compare the field data for a street with the typical data stored in the program to choose the type of distress. Then it is important to test

the precision of prediction for each test type. Finally, the resulted types of distresses can be represented in a drawing named “Geo-map”.

Orange has many statistical Prediction tests. In this research, many types of prediction methods are tested, but this research has chosen the highest four main types of prediction methods in precision of predictions after examen them practically and discussed them briefly and compared them each other to choose the most accurate prediction method. These statistical tests are as follows:

a. **Logistic Regression**

Logistic Regression is a statistical model that in its basic form uses a logistic function to model a binary dependent variable, although many more complex extensions exist. In regression analysis, logistic regression (or logit regression) is estimating the parameters of a logistic model (a form of binary regression).

In statistics, the logistic model (or logit model) is used to model the probability of a certain class or event existing such as pass/fail, win/lose, alive/dead or healthy/sick. This can be extended to model several classes of events such as determining whether an image contains a cat, dog, lion, etc. Each object being detected in the image would be assigned a probability between 0 and 1 and the sum adding to one. (Tolles & Meurer, 2016)

Logistic Regression can be Binomial, Ordinal or Multinomial. Binomial or Binary Logistic Regression deals with situations in which the observed outcome for a dependent variable can have only two possible types, "0" and

"1" (which may represent, for example, "dead" vs. "alive" or "win" vs. "loss"). Multinomial Logistic Regression deals with situations where the outcome can have three or more possible types (e.g., "disease A" vs. "disease B" vs. "disease C") that are not ordered. Ordinal logistic regression deals with dependent variables that are ordered.

In Binary Logistic Regression, the outcome is usually coded as "0" or "1", as this leads to the most straightforward interpretation (Hosmer & Lemeshow, 2000). If a particular observed outcome for the dependent variable is the noteworthy possible outcome (referred to as a "success" or a "case") it is usually coded as "1" and the contrary outcome (referred to as a "failure" or a "noncase") as "0". Binary logistic regression is used to predict the odds of being a case based on the values of the independent variables (predictors). The odds are defined as the probability that a particular outcome is a case divided by the probability that it is a “noncase”.

Like other forms of regression analysis, Logistic Regression makes use of one or more predictor variables that may be either continuous or categorical. Unlike ordinary linear regression, however, logistic regression is used for predicting dependent variables that take membership in one of a limited number of categories (treating the dependent variable in the binomial case as the outcome of a Bernoulli trial) rather than a continuous outcome. Given this difference, the assumptions of linear regression are violated. In particular, the residuals cannot be normally distributed. In addition, linear regression may make nonsensical predictions for a binary dependent variable. What is needed is a way to convert a binary variable into a continuous one that can take on any real value (negative or positive). To do that, binomial logistic regression first calculates the odds of the event happening for different levels of each independent variable, and then takes its logarithm to create a continuous criterion as a transformed version of the dependent variable. The logarithm of the odds is the logit of the probability, the logit is defined as follows:

$$\operatorname{logit} p = \ln \frac{p}{1-p} \quad \text{for } 0 < p < 1. \qquad \text{Equation 7}$$

Where, P is the probability.

Although the dependent variable in logistic regression is Bernoulli, the logit is on an unrestricted scale (Hosmer & Lemeshow, 2000). The logit function is the link function in this kind of generalized linear model, i.e.

$$\operatorname{logit} \mathrm{E}(Y) = \alpha + \beta x \qquad \text{Equation 8}$$

Y is the Bernoulli-distributed response variable and x is the predictor variable.

The logit of the probability of success is then fitted to the predictors. The predicted value of the logit is converted back into predicted odds via the inverse of the natural logarithm, namely the exponential function. Thus, although the observed dependent variable in binary logistic regression is a 0-or-1 variable, the logistic regression estimates the odds, as a continuous variable, that the dependent variable is a success (a case). In some applications, the odds are all that is needed. In others, a specific yes-or-no prediction is needed for whether the dependent variable is or is not a case; this categorical prediction can be based on the computed odds of success, with predicted odds above some chosen cutoff value being translated into a prediction of success.

The assumption of linear predictor effects can easily be relaxed using techniques such as spline functions.

An explanation of Logistic Regression can begin with an explanation of the standard logistic function. The logistic function is a sigmoid function, which takes any real input ***t***, (***t*** ϵ ℝ), and outputs a value between zero and one; (Hosmer & Lemeshow, 2000) for the logit, this is interpreted as taking input log-odds and having output probability. The standard logistic function σ: ℝ → (0,1) is defined as follows:

$$\sigma(t) = \frac{e^t}{e^t + 1} = \frac{1}{1 + e^{-t}}$$ Equation 9

A graph of the logistic function on the t-interval (−6,6) is shown in Figure .39

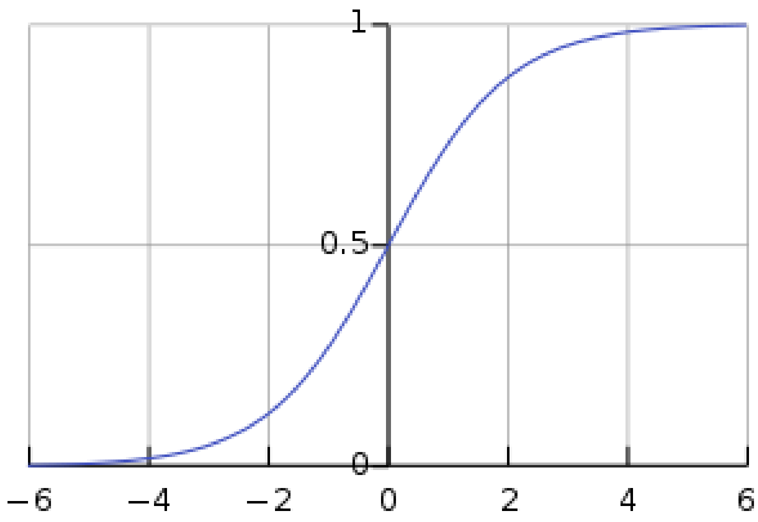


Figure 3.9: Logistic function on the t-interval (−6,6)

Let us assume that ***t*** is a linear function of a single explanatory variable x .(the case where ***t*** is a linear combination of multiple explanatory variables is treated similarly). ***t*** can then be expressed as follows:

$t = \beta_0 + \beta_1 x$ Equation 10

And the general logistic function p: $\mathbb{R} \rightarrow (0,1)$ can now be written as:

$$p(x) = \sigma(t) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x)}}$$ Equation 11

In the logistic model, $p(x)$ is interpreted as the probability of the dependent variable Y equaling a success/case rather than a failure/non-case. It's clear that the response variables Y_i are not identically distributed: $P(Y_i = 1 \mid X)$ differs from one data point X_i to another, though they are independent given design matrix X and shared parameters β (Freedman, 2009).

If there are multiple explanatory variables, the above expression $\boldsymbol{\beta_0} + \boldsymbol{\beta_1 x}$ can be revised to:

$$\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_m x_m = \beta_0 + \sum_{i=1}^{m} \beta_i x_i$$ Equation 12

Then when this is used in the equation relating the log odds of a success to the values of the predictors, the linear regression will be a multiple regression with m explanators; the parameters βj for all j = 0, 1, 2, ..., m are all estimated.

Again, the more traditional equations are:

$$\log \frac{p}{1-p} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_m x_m$$ Equation 13

and

$$p = \frac{1}{1 + b^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_m x_m)}}$$ Equation 14

where usually $b = e$.

Logistic Regression Model is a generalized form of Linear Regression Model. It is a very good discrimination tool. Following are the advantages and disadvantage of Logistic Regression:

Advantages of Logistic Regression

1. Logistic Regression performs well when the **dataset is linearly separable**.
2. Logistic Regression is less prone to over-fitting but it can overfit in high dimensional datasets. You should consider Regularization (L1 and L2) techniques to avoid over-fitting in these scenarios.

3. Logistic Regression not only gives a measure of how relevant a predictor (coefficient size) is, but also its direction of association (positive or negative).
4. Logistic Regression is easier to implement, interpret and very efficient to train.

Disadvantages of Logistic Regression

1. Main limitation of Logistic Regression is the **assumption of linearity** between the dependent variable and the independent variables. In the real world, the data is rarely linearly separable. Most of the time data would be a jumbled mess.
2. If the number of observations is lesser than the number of features, Logistic Regression should not be used, otherwise it may lead to overfit.
3. Logistic Regression can only be **used to predict discrete functions**. Therefore, the dependent variable of Logistic Regression is restricted to the discrete number set. This restriction itself is problematic, as it is prohibitive to the prediction of continuous data.

Advantages and disadvantages of Logistic Regression are coted according to (Tu, 1996).

b. **SVM support vector machines map**

Support vector machine (SVM) is a machine learning technique that separates the attribute space with a hyperplane, thus maximizing the margin between the instances of different classes or class values. The technique often yields supreme predictive performance results. Orange embeds a popular implementation of SVM from the LIBSVM package. This widget is its graphical user interface. For regression tasks, **SVM** performs linear regression in a high dimension feature space using an ε-insensitive loss. Its estimation accuracy depends on a good setting of C, ε and kernel parameters. Where the constant *C* is the box constraint, a positive numeric value that controls the penalty imposed on observations that lie outside the epsilon margin (ε) and helps to prevent overfitting (regularization).

➢ Mathematical Formulation of SVM Regression

Support vector machine (SVM) analysis is a popular machine learning tool for classification and regression, first identified by Vladimir Vapnik and his colleagues in 1992 (Vapnik, 1995). SVM regression is considered a nonparametric technique because it relies on kernel functions.

Statistics and Machine Learning implements linear epsilon-insensitive SVM (ε-SVM) regression, which is also known as $L1$ loss. In ε-SVM regression, the set of training data includes predictor variables and observed response values. The goal is to find a function $f(x)$ that deviates from yn by a value no greater than ε for each training point x, and at the same time is as flat as possible.

➢ Linear SVM Regression: Primal Formula

A set of training data were supposed to be taken where xn is a multivariate set of N observations with observed response values yn.

To find the linear function $f(x)=x'\beta+b$, and ensure that it is as flat as possible, find $f(x)$ with the minimal norm value ($\beta'\beta$). This is formulated as a convex optimization problem to minimize

$J(\beta)=0.5\beta'\beta$ Equation 15

subject to all residuals having a value less than ε; or, in equation form:

$\forall n: |yn-(xn'\beta+b)| \leq \varepsilon.$ Equation 16

It is possible that no such function $f(x)$ exists to satisfy these constraints for all points. To deal with otherwise infeasible constraints, introduce slack variables ξn and $\xi^{*}n$ for each point. This approach is similar to the "soft margin" concept in SVM classification, because the slack variables allow regression errors to exist up to the value of ξn and $\xi^{*}n$, yet still satisfy the required conditions.

Including slack variables leads to the objective function, also known as the primal formula (Vapnik, 1995):

$J(\beta)=0.5\beta'\beta+C\sum_{n=1}^{N}\xi n+\xi_{*}n\sum_{n=1}^{N}\xi n+\xi_{*}n$ Equation 17

subject to:

$\forall n: yn-(xn'\beta+b) \leq \varepsilon+\xi n$

$\forall n: (xn'\beta+b) - yn \leq \varepsilon+\xi{*}n$

$\forall n: \xi{*}n \geq 0$

$\forall n: \xi n \geq 0.$

The constant *C* is the box constraint, a positive numeric value that controls the penalty imposed on observations that lie outside the epsilon margin (ε) and helps to prevent overfitting (regularization). This value determines the trade-off between the flatness of $f(x)$ and the amount up to which deviations larger than ε are tolerated.

The linear ε-insensitive loss function ignores errors that are within ε distance of the observed value by treating them as equal to zero. The loss is measured based on the distance between observed value y and the ε boundary. This is formally described by

$$L_\varepsilon = \begin{cases} 0 & , \text{if} |\boldsymbol{y} - \boldsymbol{f}(\boldsymbol{x})| \leq \boldsymbol{\varepsilon} \\ |\mathrm{y} - \mathrm{f}(\mathrm{x})| - \boldsymbol{\varepsilon} & , \text{otherwise} \end{cases}$$

$$L_\varepsilon = \begin{cases} 0 & , \text{if} |\boldsymbol{y} - \boldsymbol{f}(\boldsymbol{x})| \leq \boldsymbol{\varepsilon} \\ |\mathrm{y} - \mathrm{f}(\mathrm{x})| - \boldsymbol{\varepsilon} & , \text{otherwise} \end{cases}$$

Equation 18

According to Dhiraj(2019), the advantages and disadvantages of SVM method are listed below.

Advantages:

1. SVM works relatively well when there is clear margin of separation between classes.

2. SVM is more effective in high dimensional spaces.

3. SVM is effective in cases where number of dimensions is greater than the number of samples.

4. SVM is relatively memory efficient

Disadvantages:

1. SVM algorithm is not suitable for large data sets.
2. SVM does not perform very well, when the data set has more noise i.e., target classes are overlapping.
3. In cases where number of features for each data point exceeds the number of training data sample, the SVM will underperform.
4. As the support vector classifier works by putting data points, above and below the classifying hyper plane there is no probabilistic explanation for the classification.

Advantages and disadvantages of Advantages and disadvantages of SVM all are according to (Anguita, et al., 2010).

c. **AdaBoost**

The AdaBoost (short for “Adaptive boosting”) widget is a machine-learning algorithm, formulated by Yoav Freund and Robert Schapire. It can be used with other learning algorithms to boost their performance. It does so by tweaking the weak learners. It can also be used for both classification and regression tasks.

The Mathematics Behind AdaBoost

Here comes the hair-tugging part. Let's AdaBoost be broken down, step-by-step and equation-by-equation so that it's easier to comprehend.

Let's start by considering a dataset with N points, or rows, in our dataset.

$$x_i \in \mathbb{R}^n, y_i \in \{-1, 1\}$$

In this case,

- *n* is the dimension of real numbers, or the number of attributes in our dataset
- *x* is the set of data points
- *y* is the target variable which is either -1 or 1 as it is a binary classification problem, denoting the first or the second class (e.g. Fit vs Not Fit)

The weighted samples for each data point is being calculated. AdaBoost assigns weight to each training example to determine its significance in the training dataset. When the assigned weights are high, set of training data points are likely to have a larger say in the training set. Similarly, when the assigned weights are low, they have a minimal influence in the training dataset.

Initially, all the data points will have the same weighted sample *w*:

$$w = 1/N \in [0, 1]$$

where N is the total number of data points.

The weighted samples always sum to 1, so the value of each individual weight will always lie between 0 and 1. After this, the actual influence for this classifier in classifying the data points is calculated using the formula:

$$\alpha_t = \frac{1}{2} ln \frac{(1 - TotalError)}{TotalError}$$ Equation 19

Alpha is how much influence this stump will have in the final classification. *Total Error* is nothing but the total number of misclassifications for that training set divided by the training set size. To plot a graph for *Alpha* by plugging in various values of *Total Error* ranging from 0 to 1. Figure illustrates Alpha Vs Error rate according 30to (McCormick, 2013).

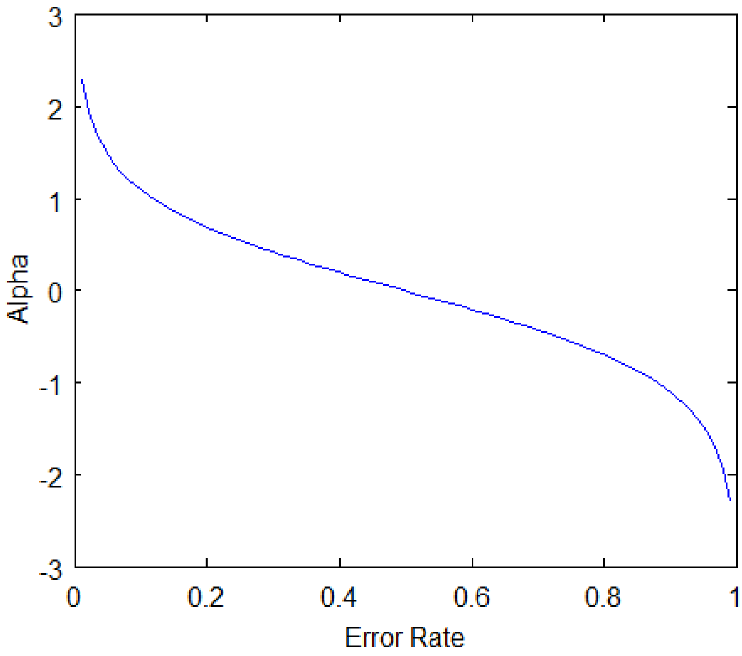


Figure 310: Alpha vs Error Rate

Notice that when a Decision Stump does well, or has no misclassifications (a perfect stump!) this results in an error rate of 0 and a relatively large, positive alpha value.

If the stump just classifies half correctly and half incorrectly (an error rate of 0.5, no better than random guessing!) then the alpha value will be 0. Finally, when the stump ceaselessly gives misclassified results (just do the opposite of what the stump says!) then the alpha would be a large negative value.

After plugging in the actual values of *Total Error* for each stump, it's time for us to update the sample weights which we had initially taken as *1/N* for every data point. We'll do this using the following formula:

$$w_i = w_{i-1} * e^{\pm\alpha} \qquad \text{Equation 20}$$

In other words, the new sample weight is equal to the old sample weight multiplied by Euler's number, raised to plus or minus alpha (which we just calculated in the previous step).

The two cases for alpha (positive or negative) indicate:

- Alpha is positive when the predicted and the actual output agree (the sample was classified correctly). In this case the sample weight should be decreased from what it was before, since it is already performing well.
- Alpha is negative when the predicted output does not agree with the actual class (i.e., the sample is misclassified). In this case there is a need to increase the sample weight so that the same misclassification does not repeat in the next stump. This is how the stumps are dependent on their predecessors.

➢ **Advantages and Disadvantages of AdaBoost**

AdaBoost has a lot of advantages such as:

- Mainly, it is easier to use with less need for tweaking parameters unlike algorithms like SVM.
- Theoretically, AdaBoost is not prone to overfitting though there is no concrete proof for this. It could be because of the reason that parameters are not jointly optimized — stage-wise estimation slows down the learning process.
- AdaBoost can be used to improve the accuracy of your weak classifiers hence making it flexible. It has now been extended beyond binary classification and has found use cases in text and image classification as well.
- The main Disadvantages of AdaBoost are:
- Boosting technique learns progressively, it is important to ensure that you have quality data.
- AdaBoost is also extremely sensitive to Noisy data and outliers so if AdaBoost is used then it is highly recommended to eliminate them.
- Advantages and disadvantages of AdaBoost all are according to (Kurama, 2020)

d. **Decision Tree**

Tree is a simple algorithm that splits the data into nodes by class purity. It is a precursor to Random Forest. Tree in Orange is designed in-house and can handle both discrete and continuous datasets. It can also be used for both classification and regression tasks. Random Forest builds a set of decision trees. Each tree is developed from a bootstrap sample from the training data. When developing individual trees, an arbitrary subset of attributes is drawn (hence the term “Random”), from which the best attribute for the split is selected. The final model is based on the majority vote from individually developed trees in the forest. (Breiman & L., 2001).

Random forests or random decision forests are an ensemble learning method for classification, regression and other tasks that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees (Ho, 1995). Random decision forests correct for decision trees' habit of overfitting to their training set (Hastie, et al., 2008).

Decision trees are a popular method for various machine learning tasks. Tree learning "come’s closest to meet the requirements for serving as an off-the-shelf procedure for data mining", Hastie et al., (2008) says : "Because it is invariant under scaling and various other transformations of feature values, is robust to inclusion of irrelevant features, and produces inspectable models. However, they are seldom accurate” (Hastie, et al., 2008).

In particular, trees that are grown very deep tend to learn highly irregular patterns: they overfit their training sets, i.e. have low bias, but very high variance. Random forests are a way of averaging multiple deep decision trees, trained on different parts of the same training set, with the goal of reducing the variance (Hastie, et al., 2008). This comes at the expense of a small increase in the bias and some loss of interpretability, but generally greatly boosts the performance in the final model.

Forests are like the pulling together of decision Decision Tree efforts. Taking the teamwork of many trees thus improving the performance of a single random tree. Though not quite similar, forests give the effects of a K-fold cross validation.

Like any other machine learning algorithm, Decision Decision Tree has both advantages and disadvantages. The following discussed advantages and disadvantages are according to (Bark, 2019).

Advantages of Decision Decision Tree

a) When using Decision Decision Tree, it is not necessary to normalize the data.

b) Decision Decision Tree implementation can be done without scaling the data as well.

c) When using Decision Decision Tree, it is not necessary to impute the missing values.

d) The data pre-processing step for decision trees requires less code and analysis.

e) The data pre-processing step for decision trees requires less time.

f) The concept behind decision tree is more familiar to programmers and comparatively easier to understand than other similar algorithms.

Disadvantages of Decision Decision Tree

a) The mathematical calculation of decision tree mostly requires more memory.

b) The mathematical calculation of decision tree mostly requires more time.

c) The reproducibility of decision tree model is highly sensitive as small change in the data can result in large change in the tree structure.

d) The space and time complexity of decision tree model is relatively higher.

e) Decision tree model training time is relatively more as complexity is high.

f) Single Decision tree is often a weak learner so we require a bunch of decision tree for called random forest for better prediction.

Hence, before implementing decision prediction method, brainstorm is needed, which of the previously discussed methods is suitable for the needed problem statement or not.

Table 31 summarizes the related important comparison points between the statistical tests.

In comparison between the four methods of algorithms according to (Barstuğan & R.Ceylan, 2014) SVM is the most appropriate method when simpler data is used like if the results of prediction one of two decisions. But when the prediction becomes more difficult (three or more decisions) **AdaBoost** algorithm become the best of the four algorithms. Theoretically, Adaboost expected to be give the most precise prediction. in the next chapter, the four statistical tests are applied to predict the results and also the results are discussed to choose the best prediction method.

Table 3.1: Statistical test comparison table

Comparison	**Logistic regression**	**AdaBoost**	**SVM**	**Tree**
Expected results Complicity (number of results)	Precise when choosing one of two choices	No mater the number of expected results	The most precise when choosing one of two choices	Can deal with more than two choices but consumes more difficulty and time to build
Type of data	Consumes linearity of data	No mater the type of data is	No mater the type of data is	No mater the type of data is
Amount of inputted data	Becomes more difficult when dealing with huge data	No mater with the inputted data amount is.	Not good when dealing with large dataset	Becomes more difficult when dealing with huge data

Chapter 4

Analysis and Discussion

4.1 Introduction

In this chapter, the previously described methods are applied in order to achieve the goals of the research. This chapter is divided into two Parts depending on calculating PCI and type of distress in both conventional and Auto-phone survey methods. Each method of survey is used to calculate PCI; the first is calculating PCI and the second is to estimate the type of distresses appeared in the road section.

4.2 Conventional field survey

First of all, a field survey for finding PCI using the conventional method of measuring each type of distresses for each 100 m section length of the road was conducted and applied to the excel templet as discussed in chapter three and also revised manually. The sample results of Conventional method for calculating PCI results are presented in Table .41

It is must noted that this conventional method of calculating PCI consumes about 2 months of hard working to collect and process the data and not forgetting to take into consideration the difficulty of weather condition for some durations of the year.

Table 4.1: Conventional method PCI results

section #	PCI	lane direction		Section #	PCI	Lane direction	
1	30	L	**Tunis street**	1	63	R	Tunes - Ring road connection street
2	32	L		2	57		
3	27	L		3	25		
4	60			4	20		
5	88			5	28		
6	76			6	67		
7	74			7	32		
8	100			8	54		
9	100			9	24		
10	82			10	36		
11	93			11	51		
12	72			12	48		
13	96	R		13	32		
14	89			14	22		
15	93			15	28		
16	89			16	38	L	
17	88			17	53		
18	89			18	13		
19	89			19	16		
20	85			20	21		
21	92			21	35		
22	80			22	12		
23	92			23	11		
24	91			24	20		
				25	18		
1	24	R	Ring road	26	55		
2	37			27	0		
3	34			28	15		
4	22			29	28		
5	69			30	44		
6	55						
7	36			1	83	R	Football court street
8	80			2	85		
9	41			3	84		
10	50			4	76		
11	80			5	100		
12	55			6	80		
13	73			7	90		
14	42			8	25		
15	63			9	57		

4.3 Auto-Phone survey

The data was collected using AndroSensor by riding the specified vehicle with a speed more than 20 KMPH. The produced data was saved automatically in an excel file including the acceleration in the three directions, linear speed, location coordinates and many other sensors readings.

The results from the collected data are described in the following two topics:

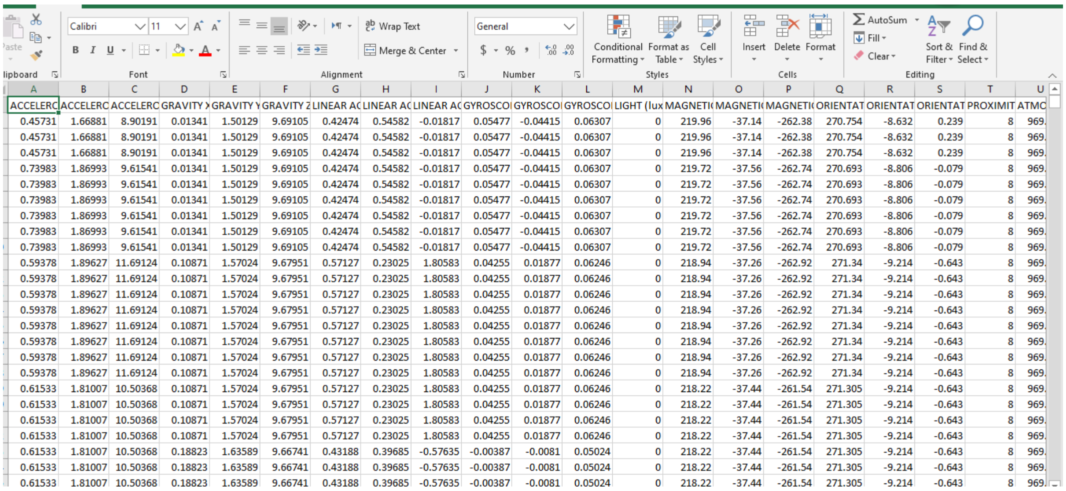
- **PCI forecasting**

First of all, in order to describe the data collected using smartphone associated with AndroSensor application, the raw data is extracted from the application and produced from the field survey using auto-phone survey method as presented in Figure The aforementioned prosses is considered .41 .to be the first step in forecasting PCI value

Each road has its separated survey excel sheet in order to be compared with PCI values collected manually by using the conventional method. Moreover, unrelated results (like for example light and sound intensity, gravity, Gyroscope, linear acceleration …, etc) are omitted in order to organize the data and reduce confusion through giving a column with number for each raw of data.

The following steps were followed to deal with the data in order to predict a formula for calculating PCI in the end:

1. A simple high pass filter is applied in order to remove unrelated low frequency signals. Spots with speed less than 20 KMH are neglected. From location speed (KMH) column, each raw with speed less than 20 is deleted. Because at low speeds, the vibrations from the vehicle motors affects greatly the Z acceleration values.
2. Data matching by GPS coordinates is carried out. This process merges PCI data from conventional field method with the acceleration data from smartphones for the same road sections.



ACCELERO	ACCELERO	ACCELERO	GRAVITY X	GRAVITY Y	GRAVITY Z	LINEAR AC	LINEAR AC	LINEAR AC	GYROSCOI	GYROSCOI	GYROSCOI	LIGHT (lux	MAGNETI(	MAGNETI(	MAGNETI(	ORIENTAT	ORIENTAT	ORIENTAT	PROXIMIT	ATMO
0.45731	1.66881	8.90191	0.01341	1.50129	9.69105	0.42474	0.54582	-0.01817	0.05477	-0.04415	0.06307	0	219.96	-37.14	-262.38	270.754	-8.632	0.239	8	969.
0.45731	1.66881	8.90191	0.01341	1.50129	9.69105	0.42474	0.54582	-0.01817	0.05477	-0.04415	0.06307	0	219.96	-37.14	-262.38	270.754	-8.632	0.239	8	969.
0.45731	1.66881	8.90191	0.01341	1.50129	9.69105	0.42474	0.54582	-0.01817	0.05477	-0.04415	0.06307	0	219.96	-37.14	-262.38	270.754	-8.632	0.239	8	969.
0.73983	1.86993	9.61541	0.01341	1.50129	9.69105	0.42474	0.54582	-0.01817	0.05477	-0.04415	0.06307	0	219.72	-37.56	-262.74	270.693	-8.806	-0.079	8	969.
0.73983	1.86993	9.61541	0.01341	1.50129	9.69105	0.42474	0.54582	-0.01817	0.05477	-0.04415	0.06307	0	219.72	-37.56	-262.74	270.693	-8.806	-0.079	8	969.
0.73983	1.86993	9.61541	0.01341	1.50129	9.69105	0.42474	0.54582	-0.01817	0.05477	-0.04415	0.06307	0	219.72	-37.56	-262.74	270.693	-8.806	-0.079	8	969.
0.73983	1.86993	9.61541	0.01341	1.50129	9.69105	0.42474	0.54582	-0.01817	0.05477	-0.04415	0.06307	0	219.72	-37.56	-262.74	270.693	-8.806	-0.079	8	969.
0.73983	1.86993	9.61541	0.01341	1.50129	9.69105	0.42474	0.54582	-0.01817	0.05477	-0.04415	0.06307	0	219.72	-37.56	-262.74	270.693	-8.806	-0.079	8	969.
0.73983	1.86993	9.61541	0.01341	1.50129	9.69105	0.42474	0.54582	-0.01817	0.05477	-0.04415	0.06307	0	219.72	-37.56	-262.74	270.693	-8.806	-0.079	8	969.
0.59378	1.89627	11.69124	0.10871	1.57024	9.67951	0.57127	0.23025	1.80583	0.04255	0.01877	0.06246	0	218.94	-37.26	-262.92	271.34	-9.214	-0.643	8	969.
0.59378	1.89627	11.69124	0.10871	1.57024	9.67951	0.57127	0.23025	1.80583	0.04255	0.01877	0.06246	0	218.94	-37.26	-262.92	271.34	-9.214	-0.643	8	969.
0.59378	1.89627	11.69124	0.10871	1.57024	9.67951	0.57127	0.23025	1.80583	0.04255	0.01877	0.06246	0	218.94	-37.26	-262.92	271.34	-9.214	-0.643	8	969.
0.59378	1.89627	11.69124	0.10871	1.57024	9.67951	0.57127	0.23025	1.80583	0.04255	0.01877	0.06246	0	218.94	-37.26	-262.92	271.34	-9.214	-0.643	8	969.
0.59378	1.89627	11.69124	0.10871	1.57024	9.67951	0.57127	0.23025	1.80583	0.04255	0.01877	0.06246	0	218.94	-37.26	-262.92	271.34	-9.214	-0.643	8	969.
0.59378	1.89627	11.69124	0.10871	1.57024	9.67951	0.57127	0.23025	1.80583	0.04255	0.01877	0.06246	0	218.94	-37.26	-262.92	271.34	-9.214	-0.643	8	969.
0.59378	1.89627	11.69124	0.10871	1.57024	9.67951	0.57127	0.23025	1.80583	0.04255	0.01877	0.06246	0	218.94	-37.26	-262.92	271.34	-9.214	-0.643	8	969.
0.59378	1.89627	11.69124	0.10871	1.57024	9.67951	0.57127	0.23025	1.80583	0.04255	0.01877	0.06246	0	218.94	-37.26	-262.92	271.34	-9.214	-0.643	8	969.
0.61533	1.81007	10.50368	0.10871	1.57024	9.67951	0.57127	0.23025	1.80583	0.04255	0.01877	0.06246	0	218.22	-37.44	-261.54	271.305	-9.214	-0.643	8	969.
0.61533	1.81007	10.50368	0.10871	1.57024	9.67951	0.57127	0.23025	1.80583	0.04255	0.01877	0.06246	0	218.22	-37.44	-261.54	271.305	-9.214	-0.643	8	969.
0.61533	1.81007	10.50368	0.10871	1.57024	9.67951	0.57127	0.23025	1.80583	0.04255	0.01877	0.06246	0	218.22	-37.44	-261.54	271.305	-9.214	-0.643	8	969.
0.61533	1.81007	10.50368	0.10871	1.57024	9.67951	0.57127	0.23025	1.80583	0.04255	0.01877	0.06246	0	218.22	-37.44	-261.54	271.305	-9.214	-0.643	8	969.
0.61533	1.81007	10.50368	0.18823	1.63589	9.66741	0.43188	0.39685	-0.57635	-0.00387	-0.0081	0.05024	0	218.22	-37.44	-261.54	271.305	-9.214	-0.643	8	969.
0.61533	1.81007	10.50368	0.18823	1.63589	9.66741	0.43188	0.39685	-0.57635	-0.00387	-0.0081	0.05024	0	218.22	-37.44	-261.54	271.305	-9.214	-0.643	8	969.
0.61533	1.81007	10.50368	0.18823	1.63589	9.66741	0.43188	0.39685	-0.57635	-0.00387	-0.0081	0.05024	0	218.22	-37.44	-261.54	271.305	-9.214	-0.643	8	969.

Figure 4.1: Raw data produced from the application during the field survey.

1. Data sectioning, the merged data files are cut into small 100-meter sections based on GPS coordinates. A 100-meter length of acceleration data is chosen as a unit for road surface estimation in this study. By adding a column named distance for each sheet, two methods were used to separate each 100 meters alone.

The first method is by allocating data in its location using Autodesk civil-3D or GIS program with respect to its location coordinates. Coordinates are derived from results of GPS sensor readings (longitude and latitude). The points located in the same 100-meter section are grouped together.

The second method is using excel worksheet by accumulating distance between each two consecutive points until reach a sum of 100 m distances. Each of these points within a 100 m distance are joined together in a group to be used later.

Since the research uses excel in developing the work firstly and due to easy in selecting the points throw each section, the researcher uses the second method which using excel to accumulate the distance.

As the WGS84 coordinate system is used by the application, then a conversion equation to be used to find the distance between two points. This equation known as Haversine equation which is illustrated in

$$d = 1000 * 6378.8 * \arccos[(\sin(lat1) * \sin(lat2)) + \cos(lat1) * \cos(lat2) * \cos(long2 - long1)]$$

Equation 21.

$$d = 1000 * 6378.8 * \arccos[(\sin(lat1) * \sin(lat2)) + \cos(lat1) * \cos(lat2) * \cos(long2 - long1)]$$

$$d = 1000 * 6378.8 * \arccos[(\sin(lat1) * \sin(lat2)) + \cos(lat1) * \cos(lat2) * \cos(long2 - long1)]$$

Equation 21

Where,

d: Distance in meter

lat1: Latitude coordinate of point number 1.

lat2: Latitude coordinate of point number 2.

long1: Longitudinal coordinate of point number 1.

long2: Longitudinal coordinate of point number 2.

The distance could be also calculated by programming a visual basic function called get distance which is clarified in Figure .42

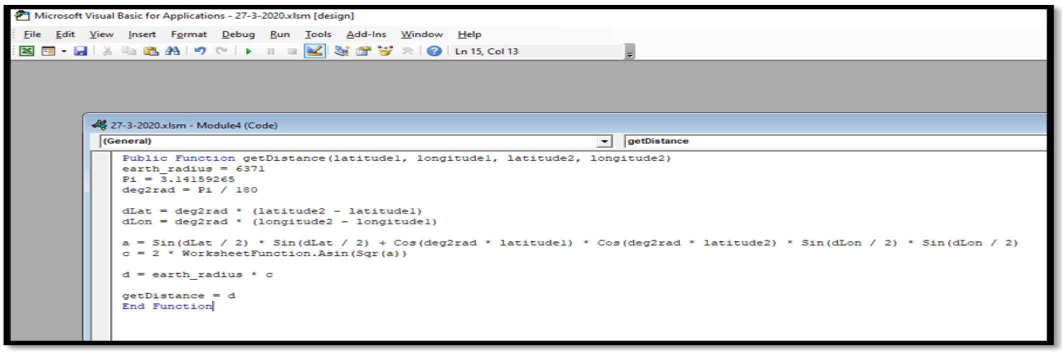
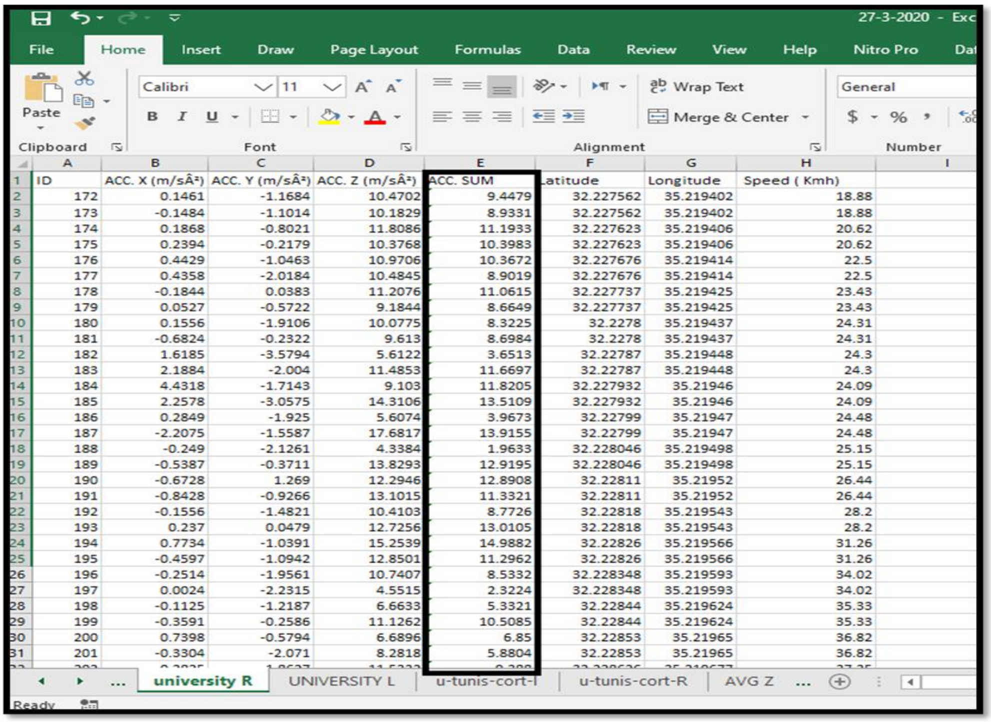


Figure 4.2: Getdistance command

2. All selected data sections are converted to frequency domain and perform Fast Fourier Transform (FFT) to smoothen the waves frequencies produced from the relations between PCI values and acceleration values. This step moved throw many sub-steps described below:

a. A new column is added which will include the values of summing XY&Z acceleration values. Since each variation in the pavement surface will affect each values of accelerations in X, Y &Z directions. Figure illustrates the summation of the accelera 43tion values.



ID	ACC. X (m/sÂ²)	ACC. Y (m/sÂ²)	ACC. Z (m/sÂ²)	ACC. SUM	Latitude	Longitude	Speed (Kmh)
172	0.1461	-1.1684	10.4702	9.4479	32.227562	35.219402	18.88
173	-0.1484	-1.1014	10.1829	8.9331	32.227562	35.219402	18.88
174	0.1868	-0.8021	11.8086	11.1933	32.227623	35.219406	20.62
175	0.2394	-0.2179	10.3768	10.3983	32.227623	35.219406	20.62
176	0.4429	-1.0463	10.9706	10.3672	32.227676	35.219414	22.5
177	0.4358	-2.0184	10.4845	8.9019	32.227676	35.219414	22.5
178	-0.1844	0.0383	11.2076	11.0615	32.227737	35.219425	23.43
179	0.0527	-0.5722	9.1844	8.6649	32.227737	35.219425	23.43
180	0.1556	-1.9106	10.0775	8.3225	32.2278	35.219437	24.31
181	-0.6824	-0.2322	9.613	8.6984	32.2278	35.219437	24.31
182	1.6185	-3.5794	5.6122	3.6513	32.22787	35.219448	24.3
183	2.1884	-2.004	11.4853	11.6697	32.22787	35.219448	24.3
184	4.4318	-1.7143	9.103	11.8205	32.227932	35.21946	24.09
185	2.2578	-3.0575	14.3106	13.5109	32.227932	35.21946	24.09
186	0.2849	-1.925	5.6074	3.9673	32.22799	35.21947	24.48
187	-2.2075	-1.5587	17.6817	13.9155	32.22799	35.21947	24.48
188	-0.249	-2.1261	4.3384	1.9633	32.228046	35.219498	25.15
189	-0.5387	-0.3711	13.8293	12.9195	32.228046	35.219498	25.15
190	-0.6728	1.269	12.2946	12.8908	32.22811	35.21952	26.44
191	-0.8428	-0.9266	13.1015	11.3321	32.22811	35.21952	26.44
192	-0.1556	-1.4821	10.4103	8.7726	32.22818	35.219543	28.2
193	0.237	0.0479	12.7256	13.0105	32.22818	35.219543	28.2
194	0.7734	-1.0391	15.2539	14.9882	32.22826	35.219566	31.26
195	-0.4597	-1.0942	12.8501	11.2962	32.22826	35.219566	31.26
196	-0.2514	-1.9561	10.7407	8.5332	32.228348	35.219593	34.02
197	0.0024	-2.2315	4.5515	2.3224	32.228348	35.219593	34.02
198	-0.1125	-1.2187	6.6633	5.3321	32.22844	35.219624	35.33
199	-0.3591	-0.2586	11.1262	10.5085	32.22844	35.219624	35.33
200	0.7398	-0.5794	6.6896	6.85	32.22853	35.21965	36.82
201	-0.3304	-2.071	8.2818	5.8804	32.22853	35.21965	36.82

Figure 4.3: The summation of the Acceleration values

b. Magnitude from FFT

Since FFT can’t be calculated except for 2^n (number of points to be transformed using FFT method is a multiple of 2), another way to compute FFT but with more complicity is Discrete Fourier Transform (DFT) since FFT is a fast method to compute DFT but for limited data domain.

Equation 22

$$A_k = \sum_{n=0}^{N-1} e^{-i\frac{2\pi}{N}kn} a_n$$

Where,

A_k: is the value of magnitude.

a_n: is the original data signal for n = 0,1, 2, ……N-1.

kn: number of components (number of converted deformations).

also, there is a ribbon named NumXL added to excel as described in Figure From this ribbon FFT can be performed automatically for the dat .44a. Figure shows the number of components (kn) which stands for maximum 45 .number of wave ranges for the raw data to be smoothen

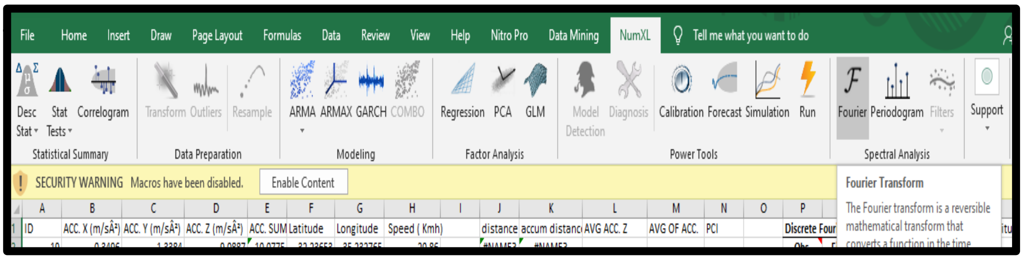


Figure 4.4: NumXL Ribbon

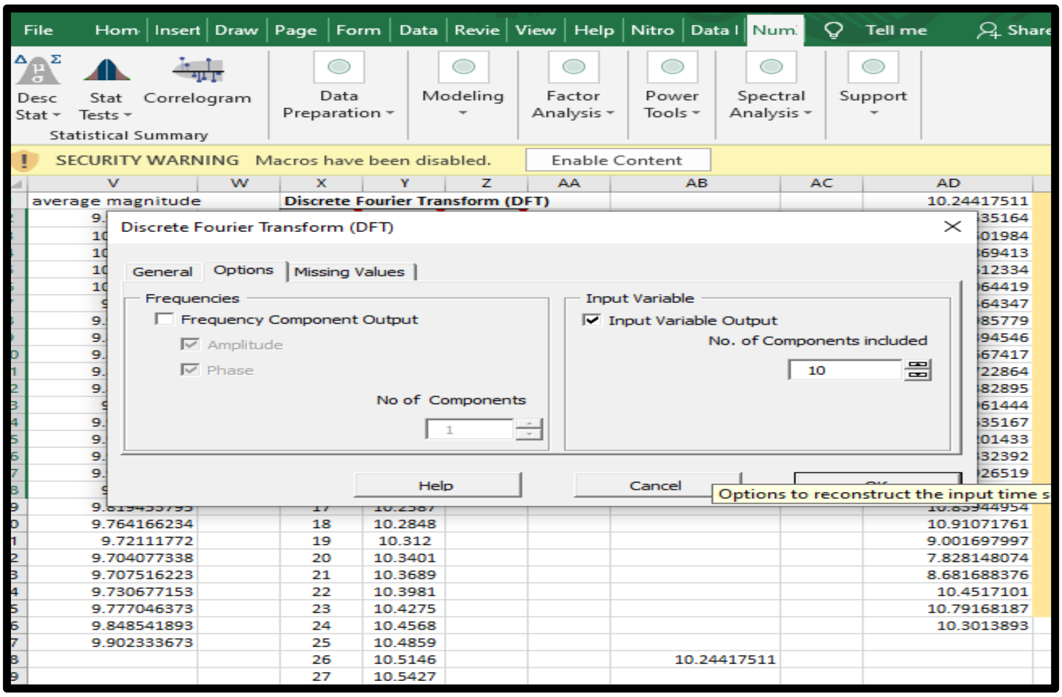
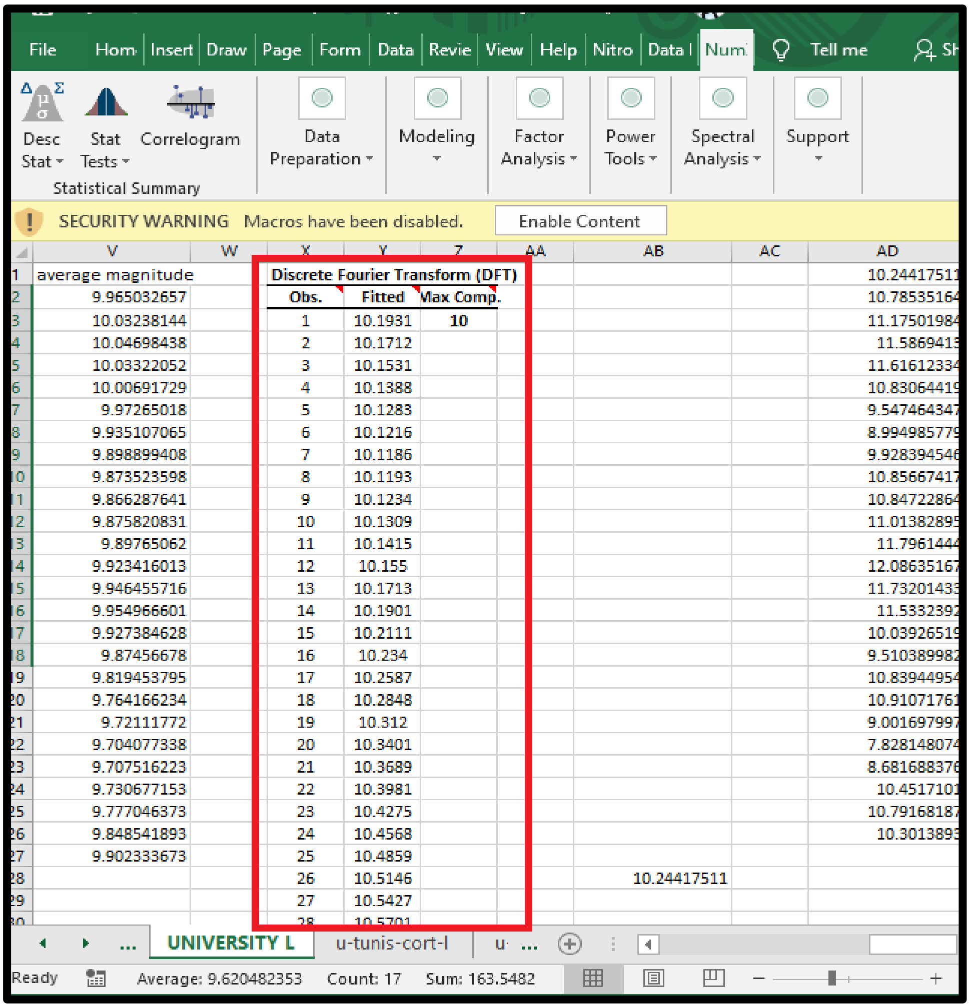


Figure 4.5: No. of components kn

Assigning the data of acceleration to the tool then assigning the number of Fourier components like in Figure In this research, the researcher .45 assumed the number of Fourier components to be 10 as each number of the ten expresses 10 points of PCI as PCI ranges from 0 to 100. Then the magnitudes transformed from the original data are appeared in a table like in Figure .46



	V	W	X	Y	Z	AA	AB	AC	AD
1	average magnitude		Discrete Fourier Transform (DFT)						10.24417511
2	9.965032657		Obs.	Fitted	Max Comp.				10.78535164
3	10.03238144		1	10.1931	10				11.17501984
4	10.04698438		2	10.1712					11.5869413
5	10.03322052		3	10.1531					11.61612334
6	10.00691729		4	10.1388					10.83064419
7	9.97265018		5	10.1283					9.547464347
8	9.935107065		6	10.1216					8.994985779
9	9.898899408		7	10.1186					9.928394546
10	9.873523598		8	10.1193					10.85667417
11	9.866287641		9	10.1234					10.84722864
12	9.875820831		10	10.1309					11.01382895
13	9.89765062		11	10.1415					11.7961444
14	9.923416013		12	10.155					12.08635167
15	9.946455716		13	10.1713					11.73201433
16	9.954966601		14	10.1901					11.5332392
17	9.927384628		15	10.2111					10.03926519
18	9.87456678		16	10.234					9.510389982
19	9.819453795		17	10.2587					10.83944954
20	9.764166234		18	10.2848					10.91071761
21	9.72111772		19	10.312					9.001697997
22	9.704077338		20	10.3401					7.828148074
23	9.707516223		21	10.3689					8.681688376
24	9.730677153		22	10.3981					10.4517101
25	9.777046373		23	10.4275					10.79168187
26	9.848541893		24	10.4568					10.3013893
27	9.902333673		25	10.4859					
28			26	10.5146			10.24417511		
29			27	10.5427					
30			28	10.5701					

Figure 4.6: Discrete Fourier Transform (DFT)

c. Therefore, average of magnitudes from FFT is studied to find out features, effect and relationship that the acceleration data might have in connection with road surface condition or pavement condition index.

Then computing the average of magnitudes for each 100-meter section from the step of sectioning (step 3).

d. A new table containing two separated columns titled as (average of magnitudes and PCI values) respectively is generated for the targeted sections. A sample of the results concerning average magnitude Vs PCI is illustrated in Table .42

Figure and equation 4 present the 47 resulting formula.

Table 4.2: Average magnitude Vs PCI results

avg. magnitude sum	PCI	avg. magnitude sum	PCI
13.525	10	10.51572	89
13.541	14	10.97819	93
13.557	20	11.6085	89
13.573	14	11.20517	88
13.583	0	11.32555	89
13.589	9	11.71478	89
13.605	0	10.74472	85
13.621	3	10.62335	92
7.586	55	12.05623	80
8.235	73	12.37843	92
9.012	88	9.525957	91
8.365	76	11.17263	73
8.956	87	10.87286	81
8.589	81	8.48393	59
10.265	94	9.879138	55
10.384	93	10.90292	44
10.023	94	9.736449	20
9.025	88	8.93718	52
9.354	92	9.840697	90
9.654	93	10.24418	9
9.989	94	10.78535	44
8.214	73	11.17502	17
7.958	66	11.58694	67
7.654	57	11.61612	61
7.853	63	10.83064	66

e. Logicly, each offset from the mean of the data"which is the practical value of acceleration that gives a value of PCI equals 100" in the positive direction due to a suden action occures on the pavement surface, will have a value confront the positive value in the negative direction. So, if the average is done for the acceleration or the magnitude values, the result will be nearly the mean value. Then the hole of the model and formulas will be useless. Due to the previously deduced note, the mean value must be defined clearly from the field.

In order to find the mean value, it is important to understand its definition and the procedure used to measure it. The proposed mean value is the value of magnitude that gives a value of 100 for PCI. That means, to know the mean value, the Auto-phone method must be applied on a tpical pavement section have a value of PCI aquale to 100 many times. The deficulty of this method and also the vehicle may face a reactions on the setion of driving according to the emotions of th edriver and these reactions may vary from person to anothor, leads to use another more practical method. This method depending on applying the Auto-phone method on a theoritically excelant surface with no deficts. When operating the vehicle and the application on the smartphone and strat recording the value of acceleration with never moving the vehicle ever from its place. By repeating this method many times and in different conditions and also by giving more and distinct power to the vehicle engine, the mean value is the average of magnitudes at each condition. Applying the previous methode in this research and at the same conditions, the ranges of the mean values are [9.89,9.95]. Consedering the mean value to be average of the ranges which is 9.92 m/s^2.

By comparing the data collected for each 100-meter section with the mean value, the following steps are followed to give a PCI value its corresponding magnitude:

1. In excell sheet separateing all the magnitudes greater than the mean in a group, and the values less than the mean in another group.
2. Then, calculating the average of each group alone.
3. The results of each average value of the two groups are the same value of PCI of the pavement section.

The resulted table of average values of magnitudes and its corresponding PCI values are presented in Figure The resulting formula from the .47 expected trendline is finalized in **$Y = -6.4402X^2 + 128.09X - 546.05$** Equation 23**Error! Reference source not found.**

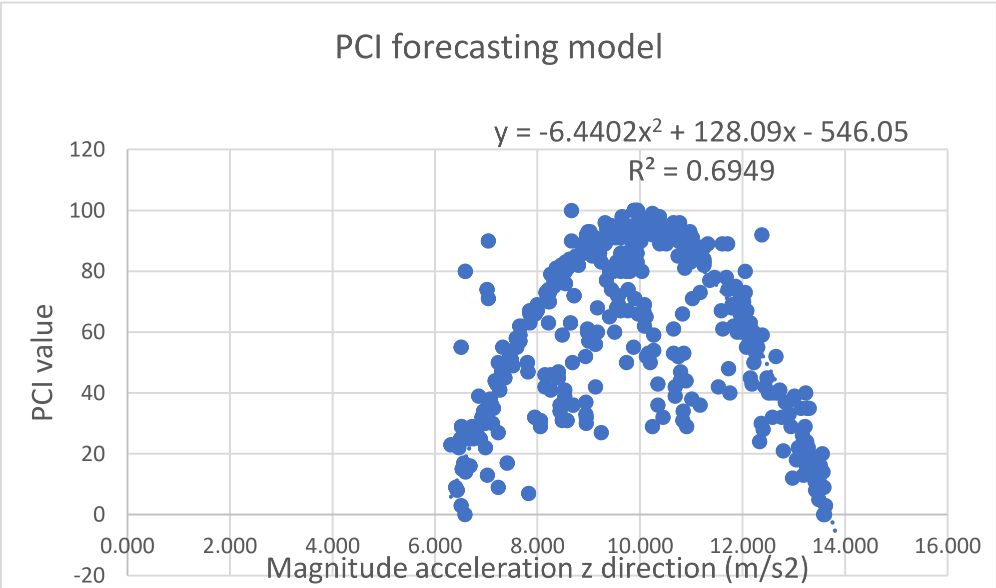


Figure 4.7: PCI forecasting

$$\mathbf{Y= -6.4402X^{2} + 128.09X -546.05} \qquad \text{Equation 23}$$

Where,

Y: is the resulted PCI value (dependent variable).

X: is the input value of magnitude driven from FFT of the acceleration values (independent parameter).

For the purpose of classifying the road pavement condition index into groups according to its values and in order to reduce the error resulting from calculation due to the lower accuracy of the smartphone sensors with respect to specialized machines, two ways are used in order to categorize these groups. The first is dividing the results of PCI into five categories and the second is by dividing it into three categories. The second method "which is dividing the results into 3 categories" is selected to neglect most of the acceptable proposed error in data collection and analyzing. According to (ASTM, 2003) these two ways are presented in Figure .48

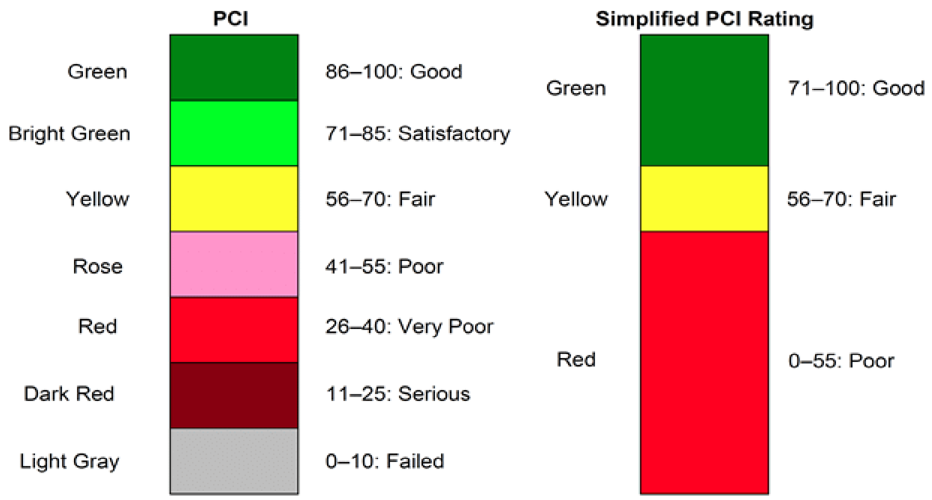
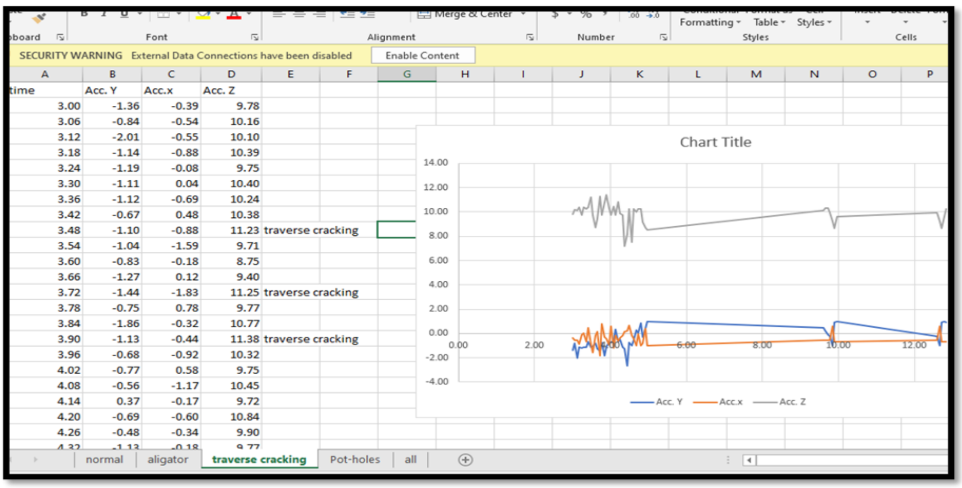


Figure 4.8: Pavement condition index (PCI) Rating

4.3.1 Distress type prediction

To predict the types of distresses from the selected distresses (Patching, Pot-holes, alligator cracking and transverse cracking), a series of steps could be followed:

1. Collecting data using the Auto-Phone surveying method for different sections, each section contains a specific type of distress to form a typical data base of the approximate acceleration for each direction axes. In this research four typical driving tests were applied on different surface conditions in order to be assigned to Orange program to provide the program with a data base describing the four needed types of surfaces (normal surface, alligator cracking, Pot-holes and transverse cracking) by adding a description column as shown in Figure to 4-12. Finally, 49 all four types are gathered in one file as described in Figure .413



time	Acc. Y	Acc.x	Acc. Z	
3.00	-1.36	-0.39	9.78	
3.06	-0.84	-0.54	10.16	
3.12	-2.01	-0.55	10.10	
3.18	-1.14	-0.88	10.39	
3.24	-1.19	-0.08	9.75	
3.30	-1.11	0.04	10.40	
3.36	-1.12	-0.69	10.24	
3.42	-0.67	0.48	10.38	
3.48	-1.10	-0.88	11.23	traverse cracking
3.54	-1.04	-1.59	9.71	
3.60	-0.83	-0.18	8.75	
3.66	-1.27	0.12	9.40	
3.72	-1.44	-1.83	11.25	traverse cracking
3.78	-0.75	0.78	9.77	
3.84	-1.86	-0.32	10.77	
3.90	-1.13	-0.44	11.38	traverse cracking
3.96	-0.68	-0.92	10.32	
4.02	-0.77	0.58	9.75	
4.08	-0.56	-1.17	10.45	
4.14	0.37	-0.17	9.72	
4.20	-0.69	-0.60	10.84	
4.26	-0.48	-0.34	9.90	

Figure 4.9: Simulation of transverse cracking

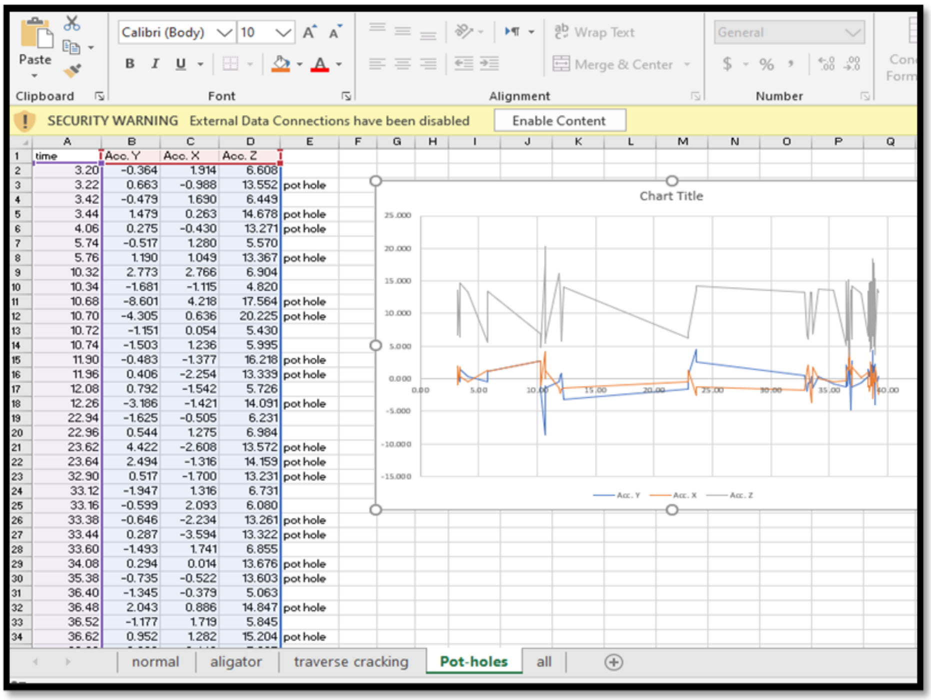


Figure 4.10: Simulation of pot-hole

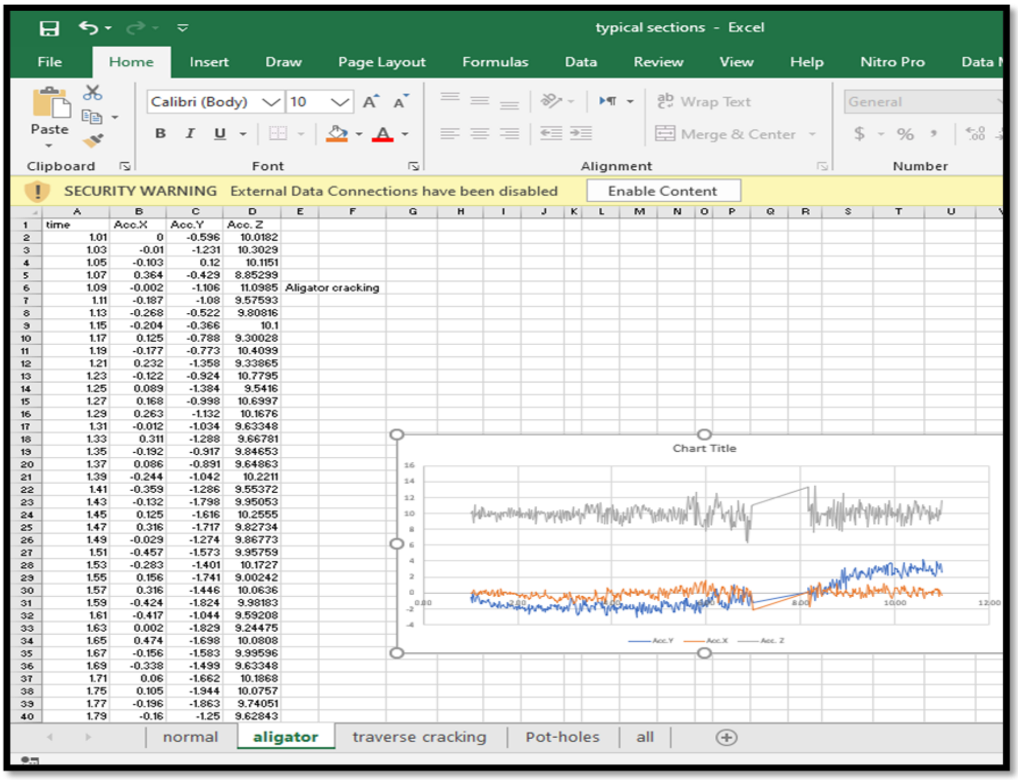


Figure 4.11: Simulation of alligator cracking

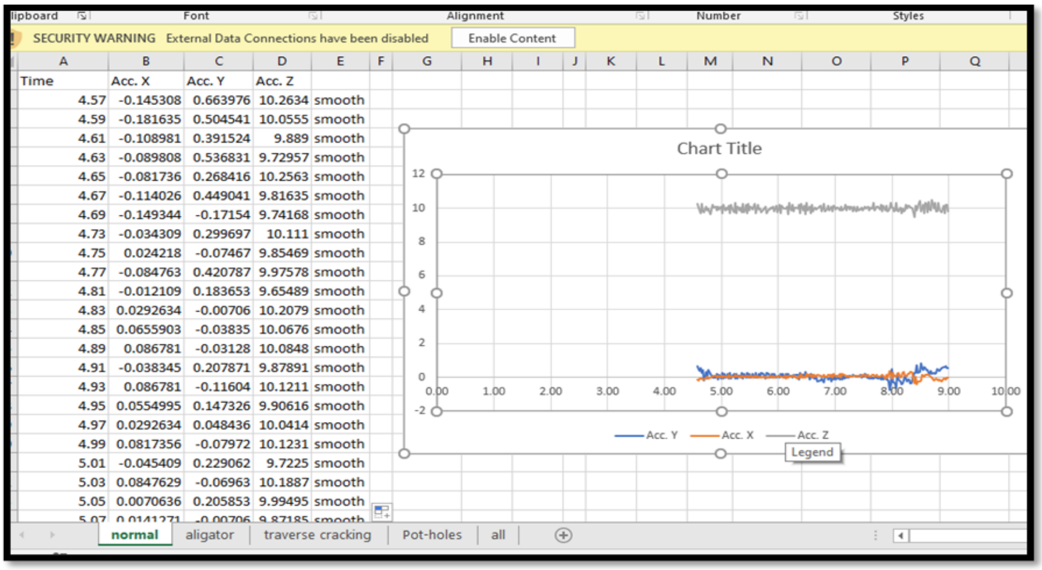


Figure 4.12: Simulation of smooth surface

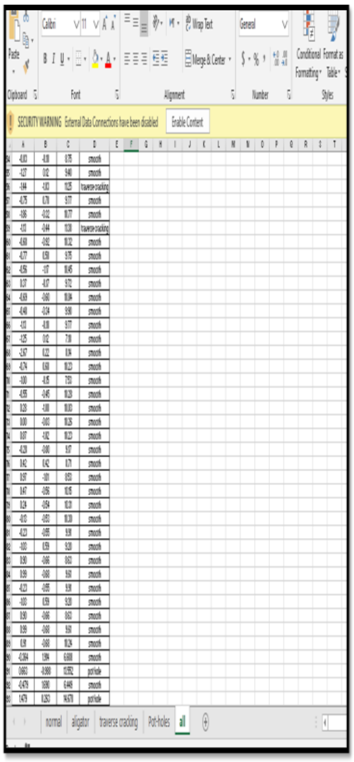


Figure 4.13: All defects with its corresponding acceleration in three directions.

2. The collected data in the previous step is integrated into Orange data mining program in order to use it as a reference data for each type of distress to learn the program. Orange 3.24.1 data mining program is

used in this research as mentioned earlier. The steps of adding the typical data file in Orange are clarified in four steps as illustrated in Figure .414

3. Statistical prediction tests are applied in model building in order to obtain the best model in predicting the different types of distresses. The applied statistical prediction tests are discussed hereafter.

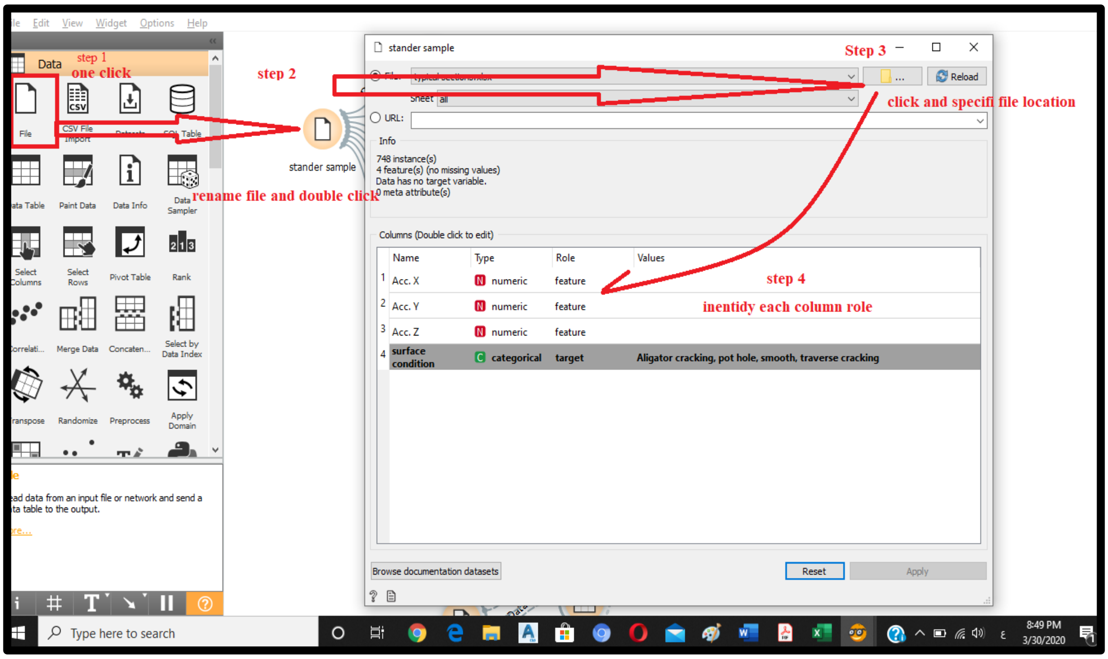


Figure 4.14: Identifying typical distresses to Orange

a. **Logistic Regression**

Logistic Regression is a statistical model that in its basic form uses a logistic function to model a binary dependent variable although many more complex extensions exist. In regression analysis, logistic regression (or logit regression) is estimating the parameters of a logistic model (a form of binary regression).

b. **SVM Support Vector Machines Map**

Support Vector Machine (SVM) is a machine learning technique that separates the attribute space with a hyperplane, thus maximizing the margin between the instances of different classes or class values.

c. **AdaBoost**

The AdaBoost (short for “Adaptive boosting”) widget is a machine-learning algorithm, formulated by Yoav Freund and Robert Schapire.

d. **Decision Tree**

Tree is a simple algorithm that splits the data into nodes by class purity. It is a precursor to Random Forest.

Accordingly, the applied statistical tests in Orange are through inserting the corresponding widget from the ribbon named model as described in Figure .415

4. From the main menu “Evaluate” choose the “Predictions” tool and assign it to the model and connect it by links to the statistical prediction tests as illustrated in the previous step. Figure illustrates the 415 .evaluation and prediction step

   Steps from 1 to 4 are related to the training processes which are all connected to standard sample file. Steps 5&6 are related to the testing stage which are connected to the testing program.

5. The data needed to be tested are added to the model using the same format and the same columns order, same headings and the column of

surface condition description must be added. Moreover, the unneeded data must be deleted. Figure illustrates the process of inserting the 415 data for testing as well as connecting to predict widgets. From the figure below, many statistical tests are tested but the the four previously described statistical testes have been used because they give the greatest precision in prediction when testing them and also when comparing between the types of tests theoretically. The most convenience methods are the selected methods.

6. Then the new data file is connected to "Predictions" as shown in Figure .415

7. By double clicking on "Predictions", the results will appear in a table showing the predicted type of distress (if there is a defect). The table contains all the predicted types of distresses according to the four types of the statistical tests which were illustrated previously in step number three. The results describing the different types of distresses will appear in a table having four additional columns, each one shows the type of distresses according to one statistical prediction method as shown in Figure .416

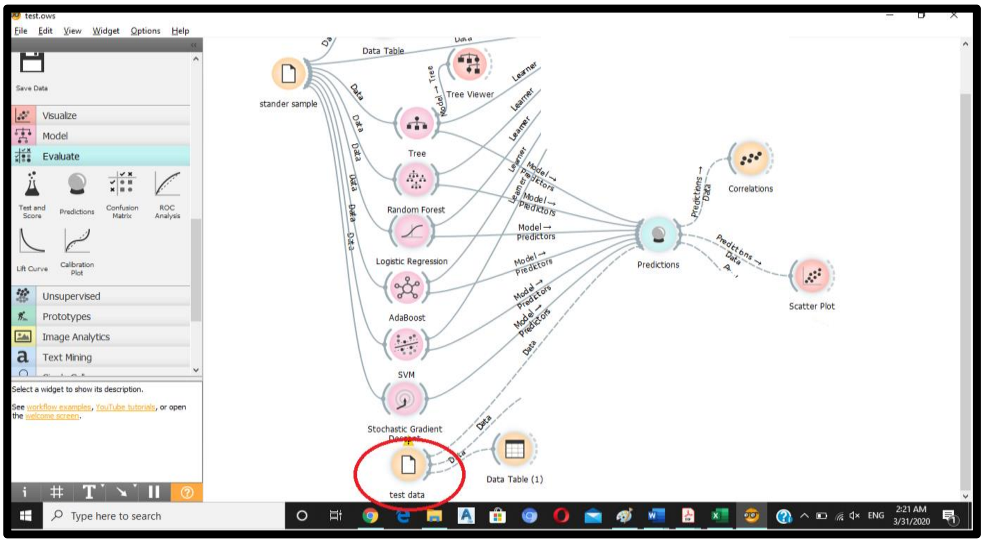


Figure 4.15: Inserting data for testing.

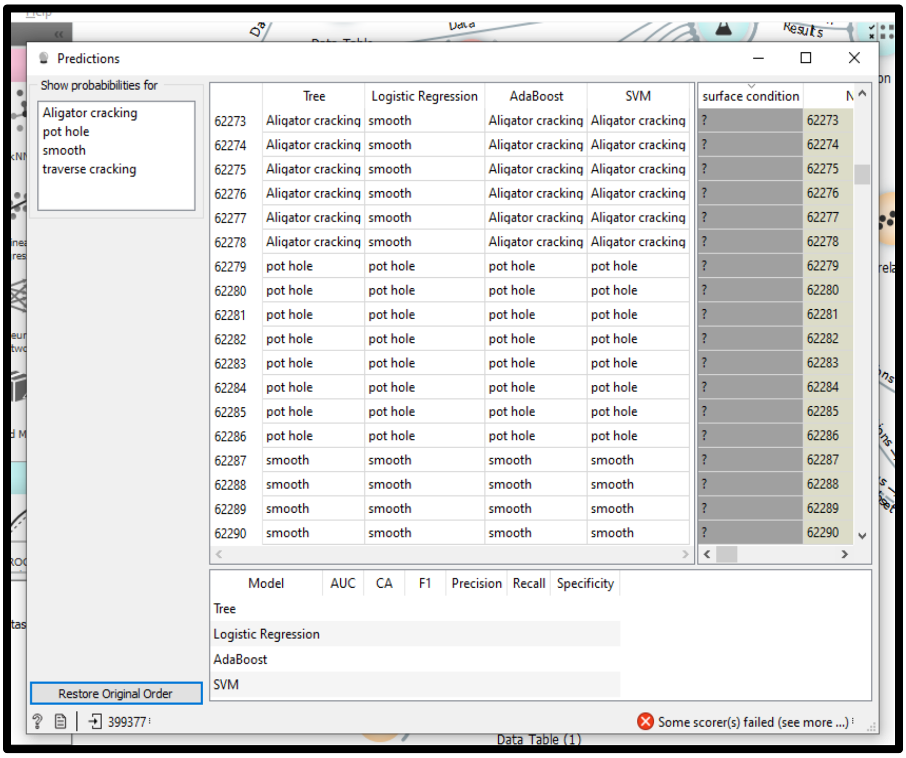


Figure 4.16: Predicted distresses types

8. A tool for choosing the best statistical tests named "test and score" can be added separately and connecting it by links with the statistical tests. By double clicking on "test and score", the precision of prediction for

each method are presented in Figure 417. This step is related to the training stage which is connected to the training file.

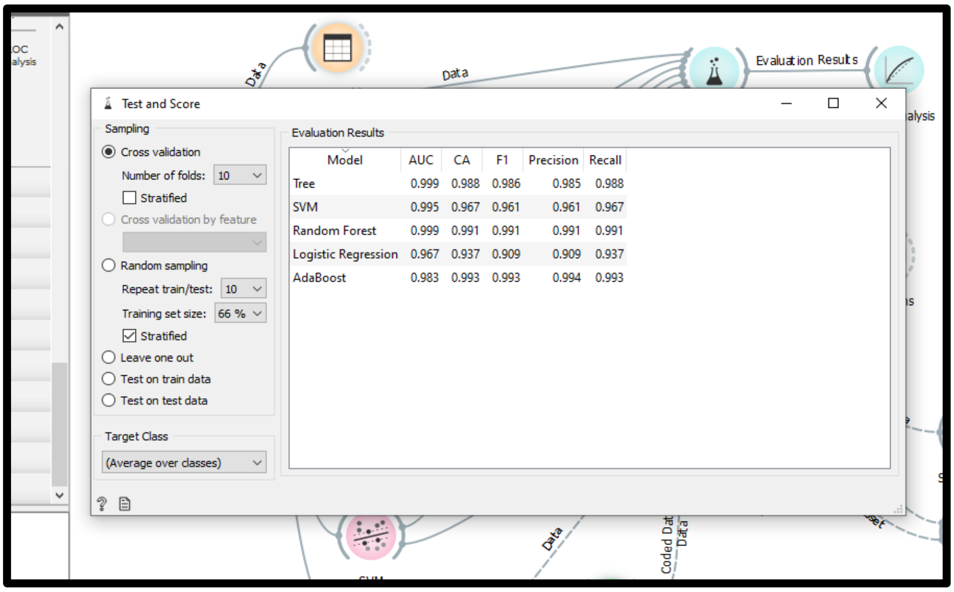


Figure 4.17: Testing and score for the precision of the prediction.

Where,

AUC: Area Under ROC is the area under the receiver-operating curve.

CA: Classification Accuracy is the proportion of correctly classified examples.

F-1: is a weighted harmonic mean of precision and recall.

Precision: is the proportion of true positives among instances classified as positive.

Recall: is the proportion of true positives among all positive instances in the data.

9. The best prediction method has the best precision. from the results obtained in Figure the best ,417of the four methods is AdaBoost because it has the best precision value.

10. Finally, the resulted types of distresses can be represented using “Geo-map” tool. Figure presents the two sub-steps to get 418 GeoMap which shows type and location of distresses on a map. The first sub-step is to let Orange know location coordinate system from the tested file by identifying the longitude and latitude columns from the map. This step named Geocoding as it is explained in Figure .419 The second step adding GeoMap widget to the model. The resulted map shown in Figure have different shapes and colors for each 420 .distress types

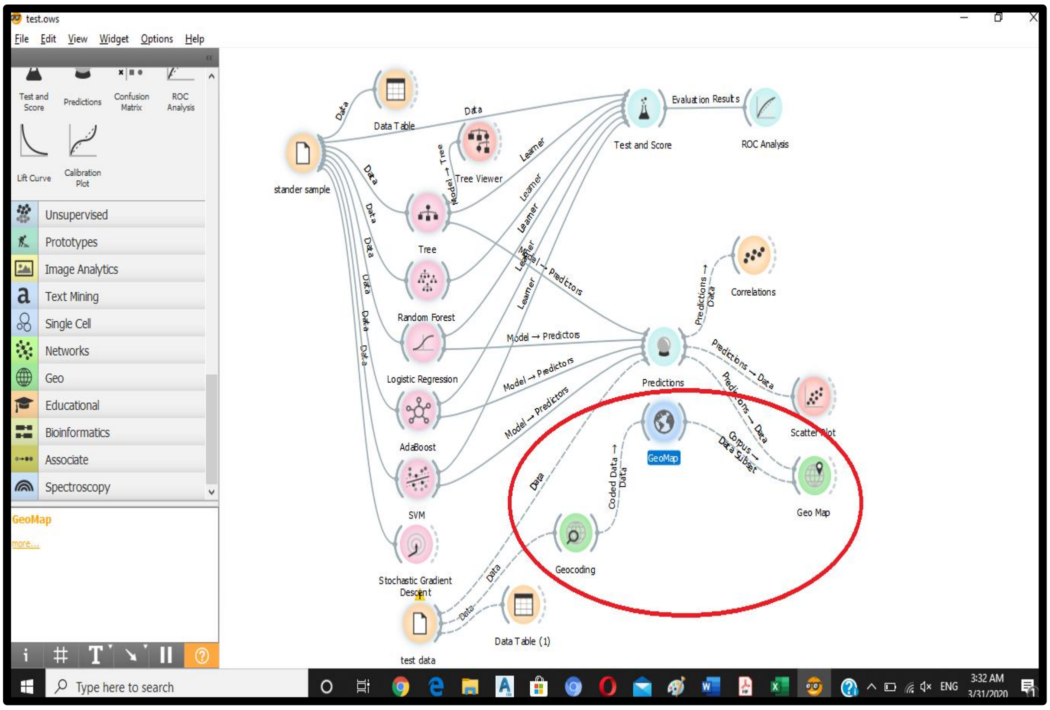


Figure 4.18: GeoMap model

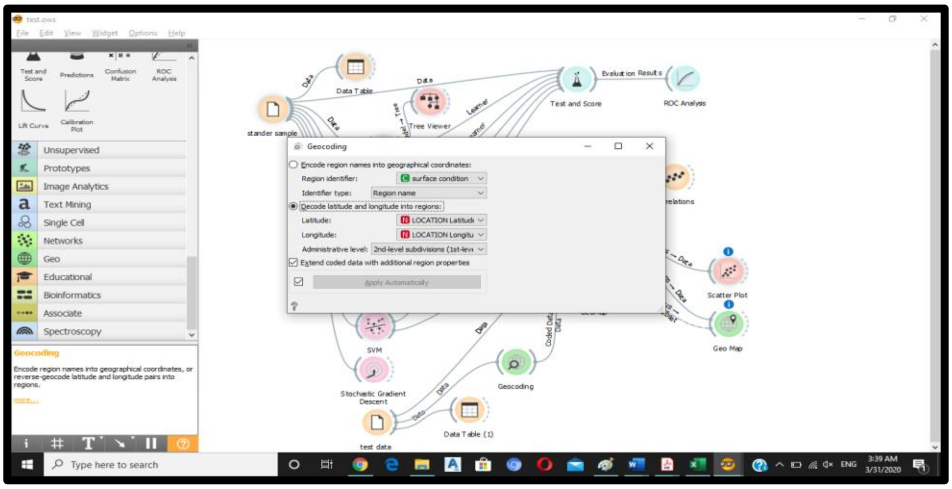


Figure 4.19: Geocoding

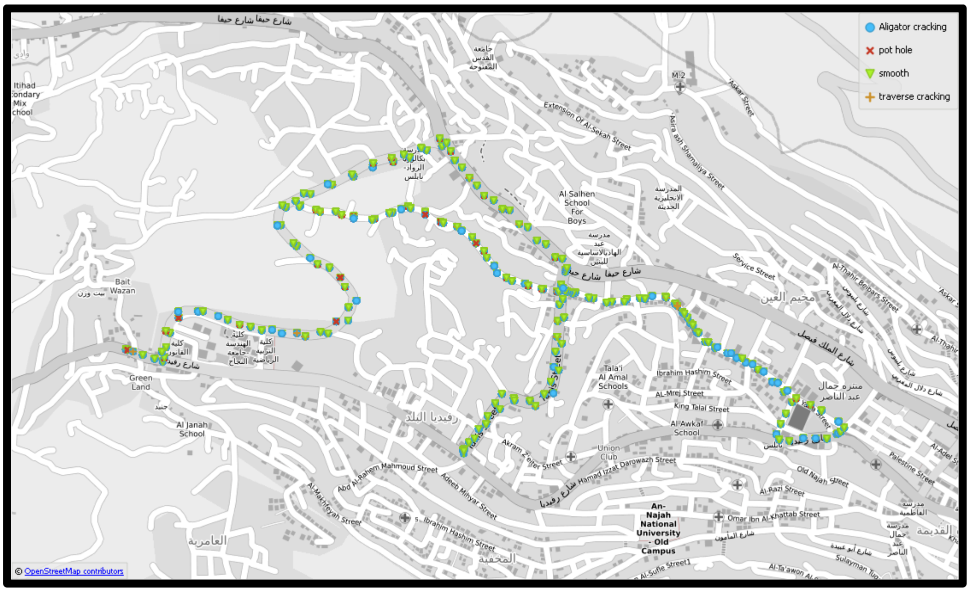


Figure 4.20: Geo-Mapping

The final model for predicting types of distresses is shown in Figure .421

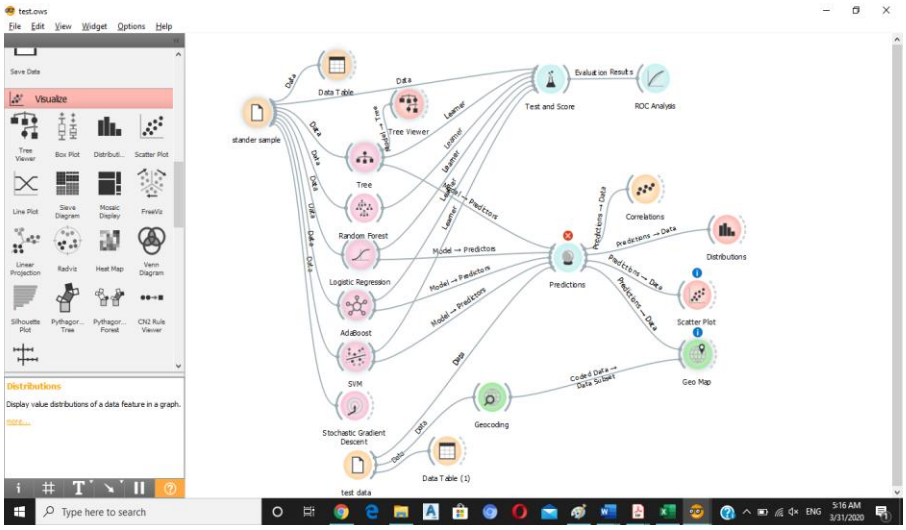


Figure 4.21: Model of predicting distresses type

4.4 Results and Discussion

To discuss the results, it is important to compare it with the objectives of the research. The final results of this research are as follows:

1. A formula to predict PCI illustrated in **$Y= -6.4402X^2 + 128.09X - 546.05$** Equation 23.

2. A model to predict the type of distress.

While the objectives were

- Developing a mobile application in order to reflect the condition of the pavement surface. it is found that there is a mobile application collects the surface data and stores it. The application measures the acceleration in the three axes, velocity, sound, gravity, vibration and many other readings from mobile sensors. This application is “Andro-Sensor”; an open-source application which is precise enough for the objectives of this research.

- Predicting a criterion in order to convert the waves created by the application to the reference PCI to simplify calculating the PI and give a rank to the studied roads. Since the step of calculating PCI is one of the main reactors in calculating PI. This step was achieved by deriving a formula introduced in Figure and 47**$Y= -6.4402X^2 + 128.09X - 546.05$** Equation 23. Figure shows a value of R 47^2 0.69 which is precise enough for a method reducing costs and labor. Which means that about 70% of the data fits the regression model while 'pure science' studies require R-squared values to be over 60% according to (Ozili, 2016). So, it is mathematically acceptable.

- Implementing a model in order to predict the type of distresses or some of them. The model presented in Figure using Orange program and 421 threw many steps depending on defining original learning file containing the values of acceleration in X,Y &Z for each type of distress, and then comparing the test data with the original stored file to predict the type of distress using many types of statistical tests. This model is precise enough to predict the type of distress since the value of error in predicting type of distress does not exceed 6% when comparing between this method and the conventional and modern technological methods in cost, time and effort. Precision of the predictions depends mainly on collecting accurate data. Figure 422 presents the distresses distribution in terms of frequency.

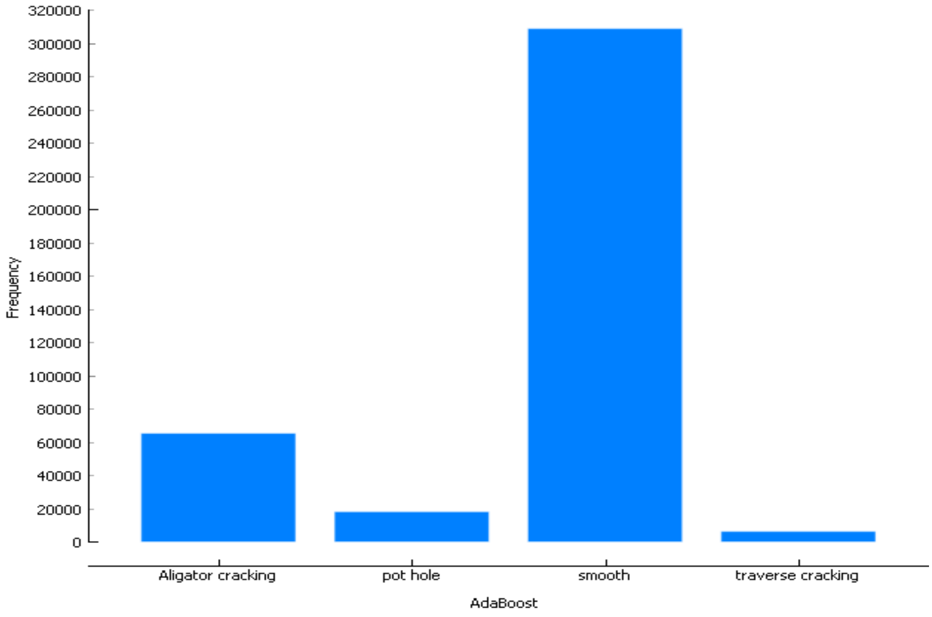


Figure 4.22: Distresses distribution.

The percentage of each defect helps decision-makers in the process of determining the causes of the defect and thus identifying the proper type of maintenance to be conducted in each road section.

Figure presents the distribution of the type of distr 422esses that were surveyed on the road surface. The most frequent distress on the surveyed roads according to AdaBoost method is the alligator cracking which forms about one seventh the road surfaces surveyed since it forms about 60000 points out of 320000 points were produced at normal sections. Also, the summation of the readings in the distress's sections are 85000 point which means that about 21% of the road surfaces are defected and deteriorated, 15% are alligator cracking, 5% potholes and 1% longitudinal cracking as illustrated in Figure .423

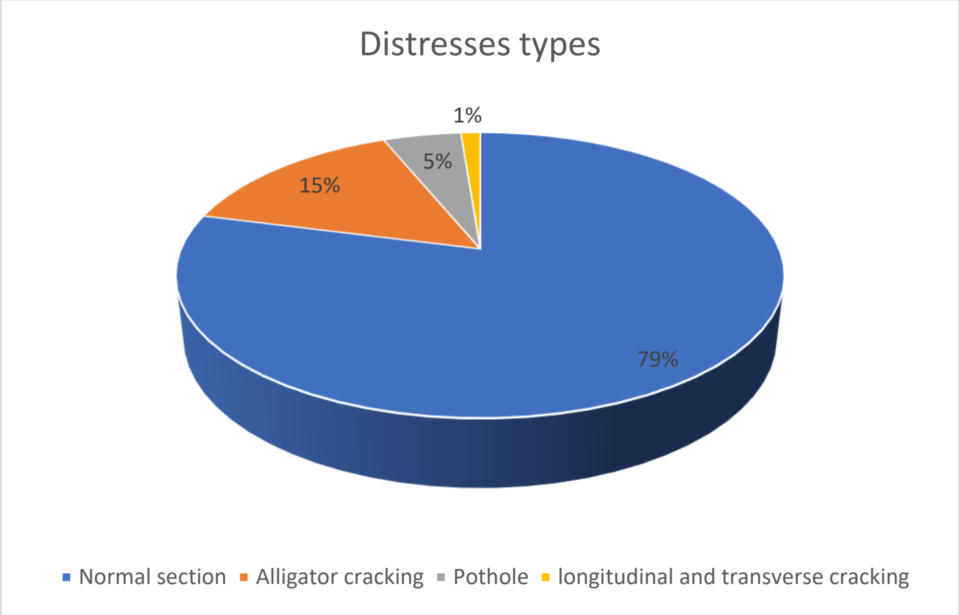


Figure 4.23: Distress Types Distribution Bichat.

To test the validation of the predicted models, it is important to divide the the test validation into two ways as discussed below:

- **PCI Forecasting**

To test the validation of the previous indicated formula $\mathbf{Y= -6.4402X^2 + 128.09X -546.05}$ Equation 23, about 200 section of length 100 meter were surveyed manually and using Auto-phone method. The following steps were followed to test the validation of the formula:

1. Field surveys were conducted using the traditional method and the new method on two new streets (Tulkarm and part of Rojeeb streets).
2. Actual PCI values were calculated using the traditional method.
3. The magnitudes were extracted using the same steps in previous sections. Also, forecasted PCI values were calculated using Equation23.

4. A comparison between the two values of PCI was made. Appendix D shows all the comparison table. The following Table 43 shows part of the results.

According to Appendix D and Table 43, most of the error values does not exceed 5% and some of them are greater than 5%. Taking into consideration that the values in Appendix D are for pavement sections of 100-meter length, and in calculating PCI for a street, the average of the PCI values for sections represents the PCI of the street. Then PCI of the entire pavement surface becomes with less error than that from a section. Moreover, the effect of the outlier values of error disappears when comparing with entire pavement surface expectations.

Table 4.3: Comparison table between PCI expected and actual

Magnitude (DFT value)	PCI based on conventional method	PCI from equation	Actual Rating Scale	Rating scale based on model	Error
7.586	55	55	fair	fair	0%
8.235	73	72	good	good	-1%
9.012	88	85	good	good	-3%
9.989	94	91	good	good	-3%
8.214	73	72	good	good	-1%
7.958	66	65	fair	fair	-1%
7.654	57	57	fair	fair	0%
7.853	63	63	fair	fair	0%
9.873	94	91	good	good	-3%
9.587	93	90	good	good	-3%
11.257	83	80	good	good	-3%
12.001	66	64	fair	fair	-2%
8.753	84	82	good	good	-2%
8.357	76	75	good	good	-1%
9.324	91	88	good	good	-3%
7.598	55	55	fair	fair	0%

- **Distress Type Prediction**

To test the validation of the produced model, the type and location of actual distresses are compared with the expected geographical location and types of the defects "which are presented in Figure The testing area is the .420 main roads surrounding An-Najah National University new campus between Qalqeelia and Tulkarm streets as described in Figure The results .420 indicated that all of the distresses are mentioned the table of distresses, and resulted from the built model in Orange are briefly located in its location as described in Figure .420

To conclude the results from the previously mentioned field test in this topic, the following points are derived:

566455576. Some of distresses from the field are not located in the model. The number of these distresses was 8 defects with respect to 127 defects are located correctly. The error percentage in distress type prediction is about 6%.

566455577. Due to the variety of distresses, types which are 19 distresses compared with 4 distresses are defined in the case study. The distresses which were not assigned to the model makes unreliability to the model in these undefined distresses.

Chapter 5

Conclusion and Recommendations

5.1 Introduction

Applying a PMS is a necessary step that most agencies around the world apply to help the decision makers in identifying the proper treatment for the proper section at the proper time. The weakness of the pavement management system in Palestine and the need for optimum utilization of resources by making the appropriate decision for proper maintenance actions were the ideas behind this research. Smartphones are reliable simple data collector and processor in the coming future according to its continuously developed sensors and applications. Using these applications based on PMS is the future in managing pavement in developing countries.

5.2 Conclusion

The early discussed results and analysis shows that the smart-phone application is reliable, precise enough and can be used for an effective monitoring and evaluation of the road network. The results from the case study show that the percentages of paved surface area are classified based on the present 2020-year conditions. The results from the case study show that the most repeated type of distresses is alligator cracking then becomes the pot-hole. At the project level, it can be seen that 75 % of the pavement surfaced area are in (smooth) condition, 18 % are having (alligator cracking), 6 % are ranked as (pot-hole), 1 % are ranked as (traverse and longitudinal cracking),

Finally, the resulted model to detect distresses and the formula to find the value of PCI are reliable and precise enough since it produces a reliable data and can be benefit to municipalities and ministry of public work and housing with a much lower costs and in small time.

These models are carrying out a huge work which needs huge budget and a large number of experts or a very expensive machine. Also, a milestone benefit from these models that it builds the foundation to use surface users in collecting data without the need of experts in data collecting. If implemented, it will help decision-makers in taking decisions interactively with the current continuously developed pavement condition with a lower specialized technical skill and consequently less cost.

5.3 Recommendations

The main recommendations for this research are:

1. The transportation agencies concerned with the pavement management in Palestine are recommended to use the PCI prediction model. This research proved that it is a practical system. Which, enables decision makers to apply the maintenance strategies for them making the right decision in prioritization of maintenance activities.
2. It’s recommended to establish a specialized unit in each municipality to manage the pavement maintenance process in all relevant

Palestinian agencies responsible for the management and maintenance of the roads.

3. It’s recommended to specify typical surfaces to define its values learn the model. Each surface having one distress only to give more precise prediction for the type and location of distresses.
4. Additionally, more researches and works should be done on using smartphone sensors and applications to explore additional uses of smartphones in the field of PMS. Some examples of possible studies include; defining all types of distresses using Auto-phone surveying method. Also, the severity of each distress can be defined accurately. Another possible studies, to use smartphones in transportation counting and also can be used to trip generation studies. Future research should be considered in studying how these types of accessories could be useful to transportation engineers.

References

1. AASHTO, 1993. ***AASHTO Guide for Design of Pavement Structures 1993.*** **Washington: American association of state highways and transportation officials**.

2. AASHTO, 2001. ***Pavement management guide,*** **Washington DC: American Association of State Highway and Transportation Officials**..

3. Acquah, P. C. & Fosu, C., 2017. **Implementation of Geographic Information System Application in the Maintenance Management of Roads in Ghana : A Case Study of Roads in Kumasi Metropolis**. pp. 90-102.

4. Agarwal, P. K., Animesh, D. & Chakroborty, P., 2004. ***A Rational Approach for Prioritisation of Highway Sections for Maintenance.*** **India, s.n.,** pp. 1-12.

5. Ali, A. et al., 2018. ***Field Inspection and Classefication of Pavement Distresses of St John's City in Newfoundland Canada,*** **JOHN'S CITY: Fredericton**, Canada.

6. Android Developer Official, 2012. ***Android Developer Reference (2012) SensorEvent.*** [Online] Available at: http://developer.android.com/reference/android/hardware/SensorEvent.html

7. Anguita, D. et al., 2010. ***Model selection for support vector machines: Advantages and disadvantages of the Machine Learning Theory.*** **Barcelona, Spain, IEEE.**

8. APWA, 1993. s.l.: **The American Public Works Association (APWA).**

9. ASTM, 2003. ***Standard Practice for Roads and Parking Lots Pavement Condition Index Surveys,*** **West Conshohocken: ASTM International**.

10. ASTM, 2007. ***Standard Practice for Roads and Parking Lots Pavement Condition Index Survay.*** **ASTM D 6433 – 07 ed. West Conshohocken: ASTM.**

11. Bark, B., 2019. ***Top 6 Advantages and Disadvantages of Decision Decision Tree.*** [Online] Available at: https://botbark.com/2019/12/19/top-6-advantages-and-disadvantages-of-decision-tree-algorithm/ [Accessed 16 may 2020].

12. Barstuğan, M. & R.Ceylan, 2014**.** ***Comparison of Decision Tree and SVM Based AdaBoost Algorithms on Biomedical Benchmark Datasets,*** **Konya, Turkey : Selcuk University**.

13. Bhoraskar, R., Vankadhara, N. & Rama, B., 2012. ***Wolverine : Traffic and Road Condition Estimation using Smartphone Sensors.*** **s.l., IEEE**.

14. Breiman & L., 2001. **Random Forests In Machine Learning**. Volume 45(1), pp. 5-32.

15. Chang, J.-R.et al., 2009. ***Measurement of The International Roughness Index (Iri) Using an Autonomous Robot.*** **26th ISARC,** pp. 325-331.

16. Chugh, G., Bansal, D. & Sofat, S., 2014. ***Road Condition Detection Using Smartphone Sensors: A Survey.*** **International Journal of Electronic and Electrical Engineering,** pp. 595-602.

17. Demsar, J. et al., 2013. **Orange: Data Mining Toolbox in Python.** ***Journal of Machine Learning Research,*** 14(Aug), p. 2349−2353.

18. Deng, H., Runger, G. C. & Tuv, E., 2011. ***Bias of Importance Measures for Multi-valued Attributes and Solutions.*** **Espoo, Finland, IEEE**.

19. Dhiraj, K., 2019. ***Top 4 advantages and disadvantages of Support Vector Machine or SVM.*** [Online] Available at: https://medium.com/@dhiraj8899/top-4-advantages-and-disadvantages-of-support-vector-machine-or-svm-a3c06a2b107 [Accessed 21 may 2020].

20. Douangphachanh, V. & Oneyama, H., 2013. **A Study on the Use of Smartphones for Road Roughness Condition Estimation**. 9(Proceedings of the Eastern Asia Society for Transportation Studies).

21. Easa, S. & Kikuchi, S., 1989. ***Pavement Performance Prediction Models:Review and Evaluation,*** **s.l.: Delaware Transportation Center**.

22. Eriksson, J. et al., 2008. ***The Pothole Patrol: Using a Mobile Sensor Network for.*** **6th international conference on Mobile systems, applications, and services,** pp. 29-39.

23. FHWA, 2014. **Distress Identification Manual for the Long-Term Pavement Performance Program (Fifth Revised Edition),** Georgetown: FHWA.

24. Freedman, D. A., 2009. ***Statistical Models: Theory and Practice. In:*** **Cambridge University Press.** s.l.:s.n., p. 128.

25. Fwa, T. & Chan, W., 1993. ***Priority Rating of Highway Maintenance Need By Neural Networks.*** **Journal of Transportation Engineering, 119 (3), pp. 419 -432.**

26. González, A., O'brien, E. J., Li, Y. Y. & Cashell, K., 2008. ***The use of vehicle acceleration measurements to estimate road roughness..*** **Vehicle System Dynamics,** *46(6),* p. 483–499.

27. Haas, R., Hudson, W. & Zaniewski, J., 1994. **Modern Pavement Management,** Florida: Krieger Publishing Company..

28. HANS, J. & LARS, F., 2013. ***ROADROID CONTINUOUS ROAD CONDITION MONITORING WITH SMART PHONES.***

29. ***Hastie, T., Tibshirani, R. & Friedman, J., 2008.*** **The Elements of Statistical Learning (2nd ed.)**, s.l.: Springer. ISBN 0-387-95284-5.

30. Heckbert, P., Feb. 1995. ***Fourier Transforms and the Fast Fourier Transform (FFT) Algorithm.*** **Computer Graphics,** Volume 2, pp. 15-463.

31. Hosmer, D. W. & Lemeshow, S., 2000. **Applied Logistic Regression (2nd ed.),** Wiley: ISBN 978-0-471-35632-5.

32. Ho, T. K., 1995. Random Decision Forests (RDF). ***Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC,*** **14-16 august**, pp. 278-282.

33. Hudson, W. R. & Fernando, E. G., 1983. ***development of prioritization procedure for the network level pavement management system,*** **Texas: Center for Transportation Research,Bureau of Engineering Research ,The University of Texas at Austin.**

34. Issa, A. & Abu-Eisheh, S., 2017. ***Evaluation of implementation of municipal roads' maintenance plans in Palestine: A pilot case study.*** **International Journal of Pavement Research and Technology,** 10(5), p. 454–463.

35. Jain, P., 2012. ***Engineers Garage.*** [Online] Available at:
https://www.engineersgarage.com/articles/magnetometer

36. Jendia, S. & Al Hallaq, M., 2005. ***Development of a pavement maintenance management system (PMMS) for Gaza City.*** **Journal of the Islamic University of Gaza, Volume 13, Number** *1*, pp. 119-138.

37. Johnson, D. A. & Trivedi, M. M., 2011. **Driving style recognition using a smartphone as a sensor platform.** ***Washington, IEEE, p.*** 1609–1615.

38. Johnston, M., 2013. **Using Cell-Phones to Monitor Road Roughness,** s.l.: s.n.

39. Kulkarni, R. B. & Miller, R. W., 2002. ***Pavement Management Systems: Past, Present, and Future.*** **Transportation Research Record Journal of the Transportation Research Board** ·.

40. Kurama, V., 2020. ***A Guide to AdaBoost: Boosting To Save The Day.*** [Online] Available at: https://blog.paperspace.com/adaboost-optimizer/ [Accessed21may 2020].

41. McCormick, C., 2013. **AdaBoost Tutorial**, s.l.: s.n.

42. Mednis, A. et al., 2011. **Real Time Pothole Detection using Android Smartphones with Accelerometers. .. s.l.,** ***International Conference on Distributed Computing in Sensor Systems and Workshops (DCOSS),IEEE..***

43. Mednis, A. et al., June 2011. **Real Time Pothole Detection using Android Smartphones with Accelerometers. .. s.l.,** ***International***

Conference on Distributed Computing in Sensor Systems and Workshops (DCOSS),IEEE..

44. Mohan, P., N.Padmanabhan, V. & Ramjee, R., 2008. ***Nericell: Rich Monitoring of Road and Traffic Conditions using Mobile Smartphones.*** **6th ACM conference on Embedded network sensor systems, SenSys,** pp. 323-336.

45. MPWH, 2015. **annual report 2015**, s.l.: MPWH.

46. MPWH, 2015. **Road Sector In Palestine Achievements, Needs, and Challenges,** s.l.: MPWH.

47. Ozili, P. K., 2016. ***Re: What is the acceptable r-squared value?. Retrieved from:.*** [Online] Available at: https://www.researchgate.net/post/what_is_the_acceptable_r-squared_value/57cfcb0d3d7f4be2fb47f515/citation/download [Accessed 13 abril 2020].

48. Palestinian Ministry of transportation, 2017. **Annual Statistical report, s.l.:** ***Palestinian Ministry of transportation.***

49. PCBS, 2018. ***Transportation and Communications Statistics in Palestine:*** **Annual Report, 2018**, s.l.: PEBS.

50. Perera, R. W., Kohn, S. D. & Tayabji., S., 2005. ***Achieving a High Level of Smoothness in Concrete Pavements without Sacrificing Long-Term Performance.*** **Washington, D.C.:** FHWA-HRT-05-068.

51. Perttunen, M. et al., 2011. ***Distributed road surface condition monitoring using mobile phones..*** **Ubiquitous Intelligence and Computing**, pp. 64-78.

52. Rabaya, R. H., 2018. ***Design of an Integrated Pavement Management Systems with Geographic Information Systems,*** Nablus: An-Najah National University.

53. SankarPandi, S. K., 2020. ***How to Measure Acceleration in Smartphones using Accelerometer?.*** [Online] Available at: https://blog.contus.com/how-to-measure-acceleration-in-smartphones-using-accelerometer/ [Accessed 18 Nov 2020].

54. Schlotjes, M. R., Visser, A. & Bennet, C., 2014. **Evaluation of a smart phone roughness meter**, , s.l.: University of Pretoria.

55. Shahidul Islam, M., 2015. ***DEVELOPMENT OF A SMARTPHONE APPLICATION TO MEASURE PAVEMENT ROUGHNESS AND TO IDENTIFY SURFACE IRREGULARITIES,*** Urbana, Illinois: University of Illinois at Urbana-Champaign.

56. Shahin, M., 2005. **Pavement for airports, roads, and parking lots**, New York, USA.: Springer.

57. Shahin, M. & Kohn, S., 1981. **Pavement maintenance management for roads and parking lots. , Technical Report M-294**., s.l.: United States Army Corps of Engineers.

58. Shahin, M. Y., n.d. **Pavement Management for Airports, Roads, and Parking lots**, s.l.: ASTM D 6433 – 07.

59. Silva, N., Shan, V., Soares, J. & Rodrigues, H., 2018. ***Road Anomalies Detection System Evaluation. :*** , Guimaraes: University of Minho.

60. SMEC, April 2011. **Pavement Management Systems**, s.l.: SMEC.

61. Song, C., 2013. ***A Study of Mobile Sensing Using Smartphones.*** **International Journal of Distributed Sensor Networks,** pp. Volume 2013, Article ID 272916, 11 pages.

62. Spangler, E. B. & Kelley, W. J., 1964. ***GMR Road Profilometer - A Method for Measuring Road Profile. In*****: Research Publication GMR - 452,. Warren, Michigan: General Motor Research Laboratory**.

63. Strazdins, G. et al., 2011. ***Towards Vehicular Sensor Networks with Android Smartphones for Road Surface Monitoring.*** USA, April 11, **Paper presented at the 2nd International Workshop on Networks of Cooperating Objects, Chicago.**

64. Tai, Y., Chan, C. & Hsu, J. Y., 2010. ***Automatic road anomaly detection using smart mobile device.*** Hsinchu, Taiwan, , **Paper presented at the 2010 Conference on Technologies and Applications of Artificial Intelligence,,** pp. 18-20.

65. Tinder, R. F., 2007. **Relativistic Flight Mechanics and Space Travel: A Primer for Students, Engineers and Scientists**, s.l.: Morgan & Claypool Publishers.

66. Tolles, J. & Meurer, W. J., 2016. ***Logistic Regression Relating Patient Characteristics to Outcomes.*** **PMID** *27483067*.

67. Tu, J. V., 1996. ***Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes.*** **Journal of Clinical Epidemiology,** 01 NOVEMBER, VOLUME 49(11), pp. P1225-1231.

68. Vapnik, V., 1995. **The Nature of Statistical Learning Theory**, New York: Springer.

69. Woodstrom, J. H., 1990. ***Measurements, Specifications, and Achievement of Smoothness for Pavement Construction..*** Washington, D.C.: **National Cooperative Highway Research Program: Transportation Research Board.**

70. WSDT, 1994. ***A guide for local agency pavement managers***, Olympia: **Washington State Department of Transportation, Trans Aid Service Center**.

جامعة النجاح الوطنية

كلية الدارسات العليا

دراسة تجريبية للتنبؤ بقيمة PCI باستخدام نظام استشعار الهاتف المحمول و الذكاء الاصطناعي

إعداد

سعيد رياض عبد الحفيظ علي

إشراف

د. امجد عيسى

د. ايهاب حجازي

قدمت هذه الأطروحة استكمالا لمتطلبات الحصول على درجة الماجستير في هندسة الطرق والمواصلات بكلية الدارسات العليا في جامعة النجاح الوطنية في نابلس، فلسطين.

2021

دراسة تجريبية للتنبؤ بقيمة PCI باستخدام نظام استشعار الهاتف المحمول و الذكاء الاصطناعي

إعداد

سعيد رياض عبد الحفيظ علي

إشراف

د. أمجد عيسى

د. ايهاب حجازي

الملخص

تواجه هيئات المواصلات في فلسطين تحديًا كبيرا في التعامل مع البنية التحتية القديمة و خصوصا البنية التحتية الخاصة بالرصفات , حيث ان اعمار اغلب الطرق في فلسطين تتراوح ما بين 20–30 سنة وهذا العمر يعتبر العمر الافتراضي للطرق اما باقي الطرق فهي تعاني من نقص في المتابعة والإدارة لعمليات الصيانة .ان التوجه القائم والمستقبلي هو الحفاظ على الرصفات لكي توفر الأمان والراحة لنقل الناس والبضائع علاوة على ذلك، تكشف الإدارة الحالية أن النظام المستخدم غير مرن بما يكفي ليعكس الظروف المتغيرة و يبين أيضا ضعف المساعدة في اتخاذ قرارات الصيانة المناسبة.

تهدف هذه الدراسة إلى بناء نظام متكامل لإدارة الأرصفة باستخدام اجهزة الاستشعار المدمجة في الهواتف الذكية لإنتاج طريقة ال Auto-phone لمسح الواقع و التنبؤ بنتائج PCI و انواع العيوب لهذا المقطع. تقدم هذه الدراسة نموذجًا للتنبؤ بنوع عيوب الاسفلت تلقائيًا ، كما تنفذ صيغة للتنبؤ بقيمة PCI من خلال جمع البيانات الميدانية باستخدام سيارة تحتوي على هاتف ذكي. ويضم النظام أيضا مجموعة من الأدوات لتيسير اتباع نهج أكثر مرونة يمكّن المستعملين من أداء المهام على نحو أكثر اقتصاديا وفعالية وذات نوعية أعلى.

يتم عرض طريقة Auto-Phone لتسهيل اتخاذ القرار المناسب بالنسبةلصناع القرار. ويستند على التكامل المباشر بين قاعدة بيانات نظام إدارة الرصفات (PMS) المدمجة في برنامج مايكروسوفت إكسل و برنامج أورانج للتحليل الاحصائي من أجل استغلال قدرات كل حزمة على حدة بشكل كامل.

تم اختبار نموذج PMS (Auto–phone–) بواسطة حالة دراسية تم فيها تغذية النظام بطرق معبدة بمجموع طول 40 كم للدراسة و 20 كم للاختبار من شبكة الطرق في مدينة نابلس.

و اخيرا، تم اجراء الاختبارات اللازمة للتحقق من النماذج و المعادلات الناتجة و تم التاكد من صحتها و مطابقتها لحدود المقبول.