ARTIFICIAL INTELLIGENCE-BASED PREDICTION OF PERMEABLE PAVEMENT SURFACE INFILTRATION RATES

ARHAM MALIK

A THESIS SUBMITTED TO
THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
MASTER OF APPLIED SCIENCE

GRADUATE PROGRAM IN CIVIL ENGINEERING
YORK UNIVERSITY

TORONTO, ONTARIO

DECEMBER 2021

ABSTRACT

Permeable pavements are a type of low impact development technology that is an alternative to conventional asphalt pavements. These pavements are used to address urban stormwater runoff concerns through infiltration and storage. Overtime, sediments carried by stormwater runoff degrade the performance of these pavements and can eventually diminish the infiltration capacity to the point where no infiltration takes place. Maintenance procedures have been developed for permeable pavements and these procedures are necessary for sufficient long-term hydraulic performance. However, these procedures are expensive and are thus performed infrequently or not at all, leading to many permeable pavement systems that no longer perform at their designed infiltration capacity.

The objective of this research is to develop a data-driven model to predict the infiltration rate of permeable pavements. Four permeable concrete lab specimens were constructed and subjected to clogging cycles while obtaining surface images and infiltration data. An artificial neural network was created to investigate the relationship between the images of the pavement surface and its associated surface infiltration rate. Images of the surface were converted to grayscale and parameters of the grayscale image, namely the mean, variance, and skewness, were used as inputs to the model.

Modelling results presented in this thesis demonstrate that the use of images of the surface of a pavement were adequate in predicting the surface infiltration rate. Image parameters were seen to have significant trends as test cycles progressed. In general, images of specimens that yielded lower SIRs were whiter (larger grayscale mean), more compact (smaller grayscale variance), and had a larger positive skewness. These models can be used to estimate the surface infiltration rates of permeable concrete pavements, leading to more widespread maintenance and thus, ensure the designed hydraulic performance level is maintained.

ACKNOWLEDGEMENTS

To begin with, I would like to acknowledge my supervisors Dr. Usman T. Khan and Dr. Liam J. Butler for their continued support, expertise, and friendship these past two years. It has been a great experience working under their supervision and I learned so much from them throughout my graduate program.

I would also like to thank York University and Lassonde School of Engineering for funding the research project as well as the Department of Civil Engineering. I would also like to thank everyone who helped with the laboratory component of the research including Riad Rajab, Melessa Salem, and Kunjan Rupakheti. I am also very grateful to research assistants Karen Abogadil and Apostolos Vasileiou who were extremely helpful for the data collection phase of this thesis.

Lastly, I am very grateful for my family and friends for their support over the last two years and for their continuous support as I advance in my career.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS

LIST OF ABBREVIATIONS

ANN	Artificial Neural Network
BMP	Best Management Practices
CNN	Convolutional Neural Network
GTA	Greater Toronto Area
KSDT	Kolmogorov-Smirnov Distribution Test
LID	Low Impact Development
LOGSIG	Log-Sigmoid Transfer Function
MGL	Mean Gray Level
ML	Machine Learning
MLP	Multi-Layer Perceptron
MLR	Multiple Linear Regression
MSE	Mean Squared Error
PA	Pervious Asphalt
PCPC	Portland Cement Pervious Concrete
PICP	Permeable Interlocking Concrete Pavers
POSLIN	Positive Linear transfer function
PSD	Particle Size Distribution
PURELIN	Pure Linear transfer function
RMSE	Root Mean Squared Error
RPN	Region Proposal Network
SIR	Surface Infiltration Rate
TANSIG	Hyperbolic Tangent Sigmoidal transfer function
TRCA	Toronto and Region Conservation Authority

TRAINBR	Bayesian Regularization Training Algorithm
TRAINLM	Levenberg-Marquardt Training Algorithm
TRAINSCG	Scaled-Conjugate-Gradient Training Algorithm

LIST OF NOTATIONS

Symbol

English alphabet

A	Percent absorption of aggregates (%)
A_{slab}	Area of constructed slab (m^2)
A/B	Aggregate to binder ratio
GL	Gray level (0 – 255)
I	Surface infiltration rate (mm/hr)
I/P	Impervious to pervious ratio
L	Average length of specimen (mm)
M_a	Mass of aggregate plus measure (kg)
M_{cyl}	Mass of dried cylinder (g)
M_i	Mass of infiltrated water (kg)
M_m	Mass of measure (kg)
M_o	Mass of oven dry sample (g)
$M_{sediment}$	Mass of sediment (g/year)
M_{SSD}	Mass of saturated surface dry sample (g)
M_{sub}	Mass of submerged cylinder (g)
N_p	Total number of pixels in the image
N_o	Total number of observations in dataset
P	Number of pixels at each gray level

R	Amount of precipitation in one year (m/year)
R^2	Coefficient of determination
S	Bulk specific gravity (kg)
TSS	Average total suspended solids concentration found in Toronto
V_m	Volume of measure (m^3)
V_{runoff}	Volume of runoff calculated for one year (m^3/year)
w/cm	Water to cementitious material
y_i	Observed SIR in dataset
$\hat{y}_i$	Predicted SIR in dataset
$\bar{y}$	Mean of observed SIR

Greek alphabet

ρ_b	Bulk density of the aggregate (kg/m^3)
ρ_w	Density of water (kg/m^3)
φ_{agg}	Void content of aggregates (%)
φ_{cyl}	Void content of dried cylinders (%)
$\emptyset_{cyl}$	Average diameter of cylinder (mm)
$\emptyset_{ring}$	Diameter of infiltration ring (mm)
σ	Standard deviation of gray level

CHAPTER 1: INTRODUCTION

1.1 Introduction and Problem Statement

Urban stormwater runoff is defined as the surface flow of water that occurs after a precipitation event or after snowmelt. As cities become more urbanized, natural landscapes are replaced with impermeable surfaces such as roads, walkways, rooftops, and other structures. The presence of these surfaces disrupts the natural, pre-developed, hydrology of a site by inhibiting the surface's capacity to absorb, store, attenuate, and infiltrate runoff. This results in larger amounts of stormwater runoff being produced with higher flow rates with the potential of carrying large amounts of contaminants into receiving water bodies.

In urban landscapes, this higher amount of volume and flow rate of stormwater runoff gives rise to multiple environmental concerns. This includes an increased risk of urban flooding, eroding of downstream water ways, and the increased risk of carrying a wide array of pollutants originating from urban areas into receiving water bodies. Traditional stormwater management practices involved 'end of pipe' treatment methods which include a network of pipes that are designed to carry stormwater runoff to a receiving waterbody such as a stormwater pond or lake. These practices focused on reducing the peak flow of urban runoff by conveying water away from the urban city as fast as possible towards the receiving waterbody. Although these methods proved to be capable of conveying these higher volumes and flowrates, they provided negligible improvements to the quality of the runoff. Recently, stormwater professionals are moving towards the use of Low Impact Development (LID) technologies as an alternative to traditional stormwater management practices to address urban stormwater runoff concerns. These technologies mimic a site's natural hydrology by capturing and treating stormwater runoff at the source, as it would in pre-developed conditions. LIDs work by using natural processes such as infiltration, evapo-transpiration, and storage to capture and treat stormwater runoff. These technologies have been proven to reduce the total volume of runoff generated, reduce the total volume of runoff leaving a site, reduce the peak flow rate during storm events, reduce the mass and concentration of contaminants

present in stormwater runoff, contribute to groundwater recharge, and decrease traditional stormwater infrastructure and maintenance costs.

Permeable pavements are one type of LID technology that are typically used in low-traffic applications such as parking lots, sidewalks, walkways, shoulders, bike lanes, and around buildings. These pavements are made with asphalt or concrete where fine aggregates are omitted to create voids to capture and infiltrate stormwater runoff. These pavements are underlain by aggregate base layers that facilitate the storage of stormwater before percolating into subsurface soils. As stormwater passes through the pavement system, some contaminants present in the runoff are filtered. The use of a geotextile membrane layer is sometimes used to further enhance this filtering process. These pavement systems have been extensively researched and are now being implemented in various applications for their ability of capturing runoff and filtering contaminants. Early on in their lifespan, these pavements have been observed to capture large volumes of runoff; however, over time, the infiltration capacity of these pavements degrade due to sediment present in runoff clogging the voids responsible for infiltration. This degradation eventually results in a reduction of permeability and of infiltration capacity of these pavements. Consequently, many installed permeable pavement systems no longer operate at their designed performance level and their ability to mitigate urban floods and remove contaminants is not fully realised. Maintenance of these pavement systems has been researched and shown to restore their permeable nature however, these maintenance procedures are not performed optimally or not performed at all.

To perform these tests optimally, the infiltration capacity of the pavement must be known. The surface infiltration rate (SIR) of a permeable pavement system is the speed at which water enters the pavement system. Current methods of obtaining the SIR of a permeable pavement system involve on-site testing procedures that are time consuming, labour intensive, and require special equipment. Moreover, asset managers and owners do not know when and how to conduct these tests. This highlights the need for a quick, reliable, and easy method of obtaining the SIR of a permeable pavement system which will ensure that these systems continue to operate at their designed performance level.

1.2 Thesis Objectives

The general objectives of this thesis are to:

1. Investigate the relationship between the surface of permeable pavement systems and infiltration capacity.
2. Develop a predictive model that uses image data to predict the SIR of permeable pavements.

1.3 Thesis Layout

The thesis consists of six chapters. Chapter 2 provides an in-depth literature review on urbanisation and stormwater runoff, permeable pavements performance and their need for maintenance, and a review of artificial neural networks. Chapter 3 provides the gaps that were found from the preceding literature review and lists the specific objectives of this thesis. Chapter 4 provides the methods and materials for conducting the laboratory component of the research. Chapter 5 provides the methods for the data collection and modelling aspect of the research. Chapter 6 discusses the results for the image analysis, laboratory analysis, and modelling results. Finally, Chapter 7 provides the major conclusions of the thesis and ends with recommendations for future work.

CHAPTER 2: LITERATURE REVIEW

2.1 Urban Stormwater Runoff

As population continues to grow around the world, it is expected that a larger fraction will live in urban cities. As a result of higher populations, cities are forced to expand and densify, leading to urban sprawl which can adversely affect the environment and more specifically, the hydrological cycle (Booth & Jackson, 1997; Wang et al., 2001). As cities around the world become more urbanized, the number of impervious surfaces such as roads, rooftops and other structures increases. These surfaces disrupt the natural hydrological cycle by preventing rainfall from infiltrating below ground, as it should in predeveloped conditions and thus increasing stormwater runoff.

Stormwater is defined as water that originates from a precipitation event that can infiltrate into soil, evaporate, be stored as on the surface as lakes or puddles due to natural depressions, or contribute as surface runoff. Urban stormwater runoff is defined as the surface flow of water after a precipitation event or snow melt that occurs in urban cities. In traditional water management practices, stormwater was considered equivalent to wastewater but the potential for stormwater as a resource of water is increasing in popularity (Ahiablame et al., 2012). The proper management of urban stormwater runoff is vital for the protection of health, infrastructure and ecosystems (Roy et al., 2008).

Natural channels that convey stormwater are replaced with stormwater sewers, drains and gutters which are designed to carry the resulting large volume of runoff as quickly as possible. This results in higher peak flow rates and higher surface runoff volume when compared to predeveloped conditions (Berry, 2008). This increase in peak flow rate and surface runoff volume carries two major consequences. First, it increases the risk of local flooding in urban areas (Suriya & Mudgal, 2012). Secondly, the increased surface runoff volume and velocity is able to carry large amounts of pollutants to receiving water bodies and degrade aquatic ecosystems and diminish the quality of potable water supplies. Pollutants such as sediments, nutrients, and heavy metals that are generated in urban cities accumulate on these surfaces and are carried by runoff to receiving water bodies such as lakes, rivers and stormwater ponds (Pratt, 1995; Legret & Pagotto,1999; Winter & Duthie, 1998; Niemczynowicz, 1999).

2.1.1 Urban Runoff Management and Low Impact Development

Traditional urban stormwater management techniques were designed with the goal of managing stormwater to minimize the risk of flooding. The basic concept involves constructing a network of pipes and conduits that carry the expected peak stormwater flow between storm events. During low intensity storm events, these systems have the capacity to convey runoff as they are designed for higher flow rates. These expected high flow rates or design flow rates are obtained from a statistical storm event and the network is designed such that it minimizes surface flooding to an acceptable level. (Packman & Kidd, 1980).

Research has shown that these systems are detrimental to downstream aquatic ecosystems and with increasing urbanization, 'end-of-pipe' systems such as detention and retention ponds were implemented to control flooding (Niemczynowicz, 1999; Dhakal & Chevalier, 2017). These ponds provide an artificial means of storing runoff and therefore, reducing the risk of flooding. Additionally, these ponds provide some level of water quality improvements such as reduction in Total Suspended Solids (USA EPA, 1999). Though these storage systems are able to collect large volumes of surface runoff and provide some form of water quality improvement, these centralized systems are still responsible in reducing the health of downstream aquatic ecosystems (USA EPA, 2000). Moreover, as cities expand, new developments and their stormwater systems are usually connected to pre-existing catchment systems, reducing the existing systems overall capacity, and eventually requiring additional investment for upgrading and rehabilitation. Traditional methods of stormwater management are costly, inefficient and due to the large amount of land required, are impractical in many urban settings (USA EPA, 2000; Cettner et al., 2014). Furthermore, these methods waste an abundant and typically overlooked domestic and industrial water resource (Niemczynowicz, 1999).

In recent years, source control and natural treatment methods such as Low Impact Development (LID) technologies have been gaining popularity as a way of capturing and treating urban stormwater runoff. The basic principle of LIDs is to maintain a site's post-development hydrology to the natural hydrology of the site before development (Ahiablame et al., 2012; Fletcher et al., 2015; Khan et al., 2012; USA EPA, 2000). The

main goal of implementing LIDs is to reduce the amount of impervious area, thereby reducing the amount of runoff generated. Several other benefits of implementing LID include the ability to treat stormwater on-site through natural hydrological processes, decrease in air pollution, and other socioeconomic benefits (Kaykhosravi et al., 2019, Li et al., 2017). LID technologies promote natural hydrological processes such as infiltration, evapo-transpiration, on-site storage, absorption, adsorption, and percolation. By encouraging these processes to occur, the increased peak flow rates and high surface runoff volumes observed in urban land use types will decrease. Examples of LID infrastructure include bioretention cells, green roofs, rain gardens, grass swales, rain barrels, and porous pavements (Ahiablame et al., 2012).

Another benefit of implementing LIDs instead of traditional stormwater management practices is their potential to reduce costs in terms of construction as well as long term maintenance. Additionally, LIDs have the added benefits of using less land and can be retrofitted urban development downstream to a receiving water body (Ahiablame et al., 2012; USA EPA, 2000). This is especially important in dense urban areas where land is scarce and existing stormwater infrastructure is reaching design capacity. Although their benefits improve the issues of traditional stormwater management practices, LIDs still have room for improvement.

2.1.2 Limitations of Low Impact Development Technologies

The main benefit of LIDs is their ability to manage and treat stormwater on-site. This means the implementation of LIDs and their design are site-specific as there will not be standardized solutions that can be adopted for different locations. Factors including soil permeability, site gradation, water table depth, and stormwater runoff characteristics will dictate where and which type of LID to implement (Eckart et al., 2017). Additionally, the ability for LIDs to infiltrate stormwater directly on-site brings further concerns. Throughout the history of modern civilization, stormwater management techniques revolved around the idea of moving water away from buildings and infrastructure. The idea of allowing water to infiltrate on-site is relatively novel and many developers are hesitant due to concerns surrounding structural damage (Podolsky, 2012). Moreover,

implementing LIDs is further complicated with the presence of brownfield sites (e.g. gas stations, industrial sites, etc.) as infiltration of stormwater on these sites may lead to groundwater contamination (Eckart et al., 2017).

In addition to site specific design constraints, LIDs performance in different storm conditions is also questionable. The performance of different types of LIDs vary significantly according to the time of peak intensity, rainfall volume, and duration of storm event. Grass swales, permeable pavements and green roofs perform best with an early peak, middle peak, and late peak, respectively (Qin et al., 2013). Furthermore, LIDs performance may become less effective in high duration and high intensity storm events (Eckart et al., 2017; Hu et al., 2019). A study that compared the performance of LIDs to traditional Best Management Practices (BMP) such as the aforementioned detention and retention ponds found that infiltration-based LIDs were more effective than storage based BMPs for smaller storms while BMPs were more effective for larger storms (Damodaram et al., 2010). This highlights the fact that hybrid systems that combine the use of LIDs and traditional stormwater practices may be the answer to sustainable stormwater management that is still effective in mitigating flood events.

Maintenance and long-term performance of LIDs are also major concerns for these technologies. Some LIDs, such as permeable pavements and bioretention cells, function by filtering pollutant loads from runoff. As these LIDs filter pollutants in stormwater runoff, their performance in capturing and treating runoff decreases (Davis et al., 2009; Valeo & Gupta, 2018; Huang et al., 2016). This highlights the fact that LIDs need regular maintenance to preserve their hydrological benefits. Long term monitoring of LIDs is essential in obtaining information of degrading performance, particularly due to the fact that these technologies are relatively new and long-term performances and long-term cost and benefits are not fully understood (Eckart et al., 2017). Furthermore, since LIDs are designed for site-specific conditions, maintenance procedures may need to be optimized accordingly as well. Due to the nature of urban cities being owned by different sectors (e.g. private and public) it is hard to determine who should take responsibility for maintenance costs (Roy et al., 2008).

In a survey conducted to stormwater professionals regarding what barriers LIDs faced, it was found that a lack of regulatory and operating environment (76%), limited

quantitative data on long-term performance and best practices (75%), insufficient information on operation, maintenance and structural best practices (70%), and institutional fragmentation of responsibilities (67%) were the top four ranking barriers that LIDs face in widespread adoption (Eckart et al., 2017).

The performance of LIDs is typically divided into two categories: quantity and quality. Hydraulic performance of LIDs is acquired by investigating runoff reduction and peak flow reduction. Research surrounding various types of LIDs has been conducted to determine their performance. Field performance of bioretention cells were investigated and it was found that 18% of events were small enough to be entirely captured by the bioretention cell and peak flow reductions ranged from 44-63% (Davis, 2008). Another study in Seattle, Washington found bioretention cells to reduce runoff by 48-74% (Chapman & Horner, 2010). Additionally, one study investigated the hydraulic performance of bioretention cells in warm and cold climates and found that 93.5% and 87.5% of runoff was reduced in warm and cold climates respectively in addition to the reduction of peak flow by 96.2% and 93.5% in warm and cold climates respectively (Khan et al., 2012). The hydraulic performance of green roofs was investigated in Vancouver, British Columbia where it was found that green roofs have the potential to reduce 29 – 58% of runoff (Roehr & Kong, 2010). The hydraulic performance of permeable pavements will be discussed in greater detail in a forthcoming section.

In addition to hydraulic performance, pollutant removal efficiencies were extensively researched to determine water quality performance of LIDs. Davis in 2001 investigated pollutant removal efficiencies of bioretention cells in a laboratory setting. The results showed excellent removal efficiencies for metal (>90%) and moderate reductions of ammonium and phosphorus (60 – 80%). Chapman & Horner in 2010 investigated pollutant removal efficiencies of bioretention cells for various types of pollutants such as TSS, nitrate, phosphorus, copper, lead, zinc, and oil. Khan et al. in 2010 found bioretention cells to remove nearly all suspended solids and significantly reduce BOD_5/COD, total nitrogen, and total phosphorus. Grass swales are another type of LID that are typically used to improve water quality. Barrett et al. in 1998 investigated the performance of vegetative controls for treating highway runoff. Results show removal efficiencies of 85%, 31 – 61%, 31 – 61%, and 68 – 93% for sediments, phosphorus,

nitrogen, and zinc respectively. **Table 1** shows a summary of removal efficiencies for different types of LIDs. It can be seen that these technologies are adequate in removing pollutant loads present in stormwater runoff. A more detailed summary of permeable pavement's water quality performance will be discussed in a forthcoming section.

Table 1: Removal efficiencies of LIDs found in literature

Reference	Type of LID	TSS	P/TP	TKN/TN	Cu	Pb	Zn
Davis, 2001	Bioretention	-	60–80	60–80	>90	>90	>90
Chapman & Horner, 2010	Bioretention	87–93	67–83	-	80–90	86–93	80–90
Khan et al., 2010	Bioretention	99	70 - 99	52 - 99	-	-	-
Barrett et al., 1998	Grass Swale	85-90	31-61	31-61	-	-	68-93

2.1.3 Runoff Characteristics in Toronto, ON

Due to the fact that implementing LIDs and their performance are site specific, it is important to understand the characteristics of runoff at the location of research. The focus of this research will be in the city of Toronto, Ontario, thus, runoff characteristics in Toronto need to be investigated. The City of Toronto reported runoff quality from various site locations in their Wet Weather Flow Management Plan (WWFMP) in 2003. **Table 2** shows average event mean concentrations from various types of catchment areas in Toronto. **Table 3** shows typical stormwater management practices and their removal efficiencies for these pollutants. The City of Toronto recommends a long-term average removal of 80% of Total Suspended Solids (TSS) on an annual loading basis from all runoff leaving a development. The table shows that only few management practices can achieve this. Comparing Table 1 to Table 3, it can be seen that LID technologies typically perform better than standard best management practices in terms of water quality

performance. Thus, ideal stormwater management approaches may include the use of both LID and standard management practices where both water quality and quantity needs are met.

Table 2: Average event mean concentrations of pollutants in City of Toronto runoff. Modified from (City of Toronto, 2003).

Land Use	TSS (mg/l)	Nitrate + Nitrite (mg/l)	TKN (mg/l)	Lead (mg/l)	Total Phosphorus (mg/l)	Benzo (G, H, I) Perylene (ng/l)
Residential	91	1.75	1.92	0.011	0.36	370
Commercial	70	0.67	0.71	0.009	0.25	260
Industrial	67	1.16	1.06	0.016	0.30	435
Highway	331	0.76	2.00	0.027	0.39	-
Airport	25	0.54	0.97	0.014	0.12	-

Table 3: Removal Efficiencies for Stormwater Management Facilities. Modified from (City of Toronto, 2006)

Pollutant Group	Pollutant	% Removal									
		Source Controls	Conveyance controls			Dry Ponds	Wet Ponds Level II / I	Constructed Wetlands	Hybrid Wet Ponds/ Wetlands Level II/I	End-of-Pipe Sand Filters	Dunkers Flow Balancing System
			In-line Filter System	Roadside Ditches	Grass Swales						
General	TSS	50	40	30	80	60	70 / 80	70	70 / 80	70	70
Nutrients	Total Phosphorus	30	30	0	34	20	57 / 65	40	57 / 65	33	57
	Nitrates + Nitrite	0	0	20	31	0	31 / 36	40	31 / 36	0	31
	TKN	35	20	0	34	30	22 / 26	30	22 / 26	45	22
Heavy Metals	Copper	50	35	10	50	30	57 / 65	40	57 / 65	45	57
	Zinc	50	35	10	70	30	63 / 72	40	63 / 72	45	63
	Lead	50	35	10	70	50	54 / 62	60	54 / 62	45	54
PAH	Benzo (G,H,I) Perylene	60	40	30	80	60	70 / 80	70	70 / 80	70	70
Pesticide	Dieldrin	0	0	20	30	0	31 / 36	40	31 / 36	0	31

2.2 Overview of Permeable Pavements

Permeable pavements are a type of LID technology that are an alternative to conventional impervious pavement surfaces such as asphalt and concrete. Permeable pavements in general have been shown to provide stormwater management by encouraging groundwater recharge, reducing surface runoff, and treating runoff pollutant loads through infiltration and storage (Dietz, 2007; Ahiablame et al., 2012).

There are several types of permeable pavements but their overall goal is the same: to infiltrate stormwater (Dietz, 2007). Permeable interlocking concrete pavers (PICP) are one type of permeable pavement that uses precast concrete blocks with open void spaces. These spaces are typically filled with stone or gravel to allow for infiltration. Pervious asphalt (PA) is a variant of typical hot mix asphalt that is commonly used for roads and parking lots. PA mixes omit fine aggregates to introduce pores for the infiltration of stormwater. Similarly, Portland Cement Pervious Concrete (PCPC) omit or use very little amounts of fine aggregates in their mix design. In terms of infiltration capacity and water quality treatment, PA and PCPC pavements seem to perform similarly. One study (Welker et al., 2012) compared PA and PCPC performance in water quality for 19 different storm events for pH, total suspended solids, chlorides, nutrients, and metals and found the only parameter to be significantly different was pH. Another study investigated the permeability of PA and PCPC mixes at different void contents and found PCPC mixes gave higher initial permeability than PA mixes and PCPC showed better clogging resistance when subject to clogging cycles (Fwa et al., 2015). In reality, the adoption of PA versus PCPC mixes likely depend on local availability of materials and construction differences.

The focus of this research will pertain to Portland Cement Pervious Concrete. The nomenclature of this type of technology is not unique. No-fines concrete, porous concrete, permeable concrete, or enhanced porosity concrete are other names found in literature for this type of technology. Portland Cement Pervious Concrete (PCPC) will be used for this thesis. PCPC have been used in various applications including low-traffic pavements such as parking lots, sidewalks and highway shoulders and medians, as well as walkways around buildings.

2.2.1 Design

The design of PCPCs typically considers two parameters: structural adequacy and hydraulic performance. In PCPCs, these two characteristics are highly correlated. Typically, as the strength of PCPCs increases, the hydraulic performance (typically permeability or infiltration rate) decreases. PCPCs used for pavements need to be designed to support the intended traffic loads and maximize the capture of site-specific stormwater runoff. The entire system of the PCPCs is considered in design. This includes the porous concrete layer itself as well as the base, subbase and subgrade (soil) layers. The use of these layers as well as their thicknesses usually depend on site-specific conditions. The base and subbase layers are layers with aggregates that are installed

below the PCPC layer. These layers provide structural support and additional storage and filtering ability (ACI, 2010). In some cases, the subgrade layer may be adequate for the structural support and drainage and the subbase becomes optional. If the subgrade is inadequate in support, drainage or filtering or in areas where freeze thaw cycles are a concern, base and subbase layers are recommended. Often, geotextile and choker layers that have smaller size aggregates are used to provide additional filtration of runoff. Figure 1 illustrates a typical porous concrete system cross section.

A PCPC thickness of 5 – 6 inches (127 – 152.4 mm) has been seen to adequately handle expected traffic loads and thicker pavements of 8 – 12 inches (203.2 – 304.8 mm) have been used for higher traffic conditions (Obla, 2010; Tennis et al., 2004). It should be noted that PCPCs beyond a thickness of 8 inches (203.2 mm) may be difficult to adequately compact (ACI, 2010).

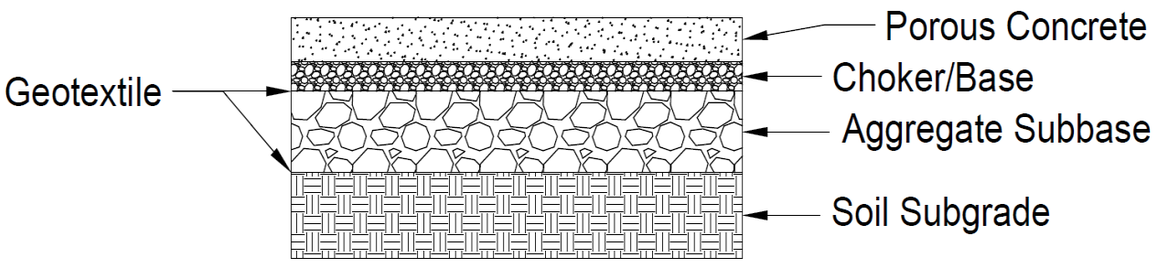


Figure 1: Typical cross section of a Porous Concrete system (not to scale).

2.2.1.1 Mix Design

The mix design of PCPC pavements is usually done on a trial-and-error basis. The process begins with identifying target strengths and hydrologic performance measures. Typically, compressive strength and porosity are used. The philosophy of a proper mix design is to find the correct balance of strength, voids, paste content and workability (ACI, 2010). Trial batches are prepared and tested to ensure the mix reaches the target performance. In Ontario, a recommended compressive strength and porosity of 15 MPa and 20% are recommended for low traffic applications (OPSS, 2018).

The constituents of PCPC are identical to regular Portland cement concrete mixes. Water, cement, and coarse aggregates are used whereas fine aggregates are typically omitted, although some mix designs do use small portions of fine aggregates to improve strength (Xu et al., 2018). The type of coarse aggregate used is an important characteristic that influences the physical properties of PCPCs. Coarse aggregates can either be single sized or graded between 9.5 and 19 mm (Sonebi et al., 2016). Single sized aggregates typically have lower strength and higher porosity whereas graded aggregates have higher strength. Single sized aggregates are preferred for maximizing infiltration rates due to maximizing void space. The most common gradings of coarse aggregate used in PCPCs meet the requirements for ASTM C33/C33M aggregate sizes of 7 (1/2 in. to No. 4), 8 (3/8 in. to No. 8), 67 (3/4 in. to No. 4), and 89 (3/8 in. to No. 16) (ACI, 2010). Aggregates can be rounded or angular. Typically, rounded aggregates achieve higher strengths but angular aggregates can be used (Tennis et al., 2004).

Another important parameter to consider when designing a PCPC mix is the water to cementitious material (*w/cm*) ratio. A high *w/cm* ratio reduces the adhesion of the paste to the aggregate and causes the paste to flow and fill voids when compacted while a low *w/cm* ratio prevents good mixing and leads to balling in the mixer, prevents even distribution of paste and therefore reduces the strength and durability of the concrete (ACI, 2010). A *w/cm* ratio of 0.26 up to 0.40 have been used and trial batches are usually conducted to select an optimal ratio (ACI, 2010). A *w/cm* of 0.27 – 0.34 is recommended (Tennis et al., 2004; ACI, 2010).

The use of admixtures has been investigated to improve typical low strengths of PCPCs while still maintaining their hydraulic performance. Water reducing agents are used to reduce the water to cement ratio and therefore increase workability. Viscosity modifying agents encourage the paste to bind to the aggregate and therefore can improve strength. Retarders are used to reduce setting time to aid in placement of PCPCs. Supplementary cementitious material such as fly ash and silica fume as well as superplasticizers have been used and shown to improve strength (Chen et al., 2013). One study (Joshaghani et al., 2015) uses the Taguchi method to obtain an optimal mix design and found that aggregates graded from 4.75 – 9.5 mm, a *w/cm* ratio of 0.35 and paste to aggregate ratio of 28% was most optimal for maximizing strength. **Table 4** provides a summary of different mix designs for porous concrete pavements found in literature. Typically, there are three main parameters that researchers adjust to find an optimal mix design. These include the type of aggregate, the amount of cementitious material, typically Portland Cement, and the amount of water. More sophisticated mixes utilize different types cementitious material and admixtures to obtain higher strengths while still maintaining their void content (Chen et al., 2013). However, as seen in **Table 4**, majority of the mixes have strengths ranging from 8 – 20 MPa while having a porosity of 15 – 30%.

Table 4: Summary of PCPC mix designs in literature

Reference	Aggregate Bulk Density (kg/m^3)	Water (kg/m^3)	Cementitious Material (kg/m^3)	W/CM	Agg./CM Ratio	28-D f'c (MPa)	Porosity (%)	Comments
Costa et al., 2018	1483.75	96.44	370.94	0.26	4	8 – 12	25 – 30	Neoprene caps for f'c found to be more conservative
Yao et al., 2018	1711	32-99	123 – 380	0.26	4.5 – 13.9	2 – 12	13.9 – 32.8	Freeze-thaw cycles also evaluated
Khankhaje, et al., 2016	1460	108.6	339.5	0.32	4.3	12 – 14	22 – 24	143 kg/m^3 of sand used
Chandrappa & Biligiri, 2016	1462.35 – 1691.67	84.28 – 160.63	321.42 – 487.45	0.25, 0.3, 0.35	3, 4 & 5	-	13.48 – 37.48	w/c, agg/cm & gradation all contributed significantly to permeability
Joshaghani et al., 2015	1565 – 1672	61.94 – 136.50	68.29 – 147.35	0.25, 0.30, 0.35	-	5.5 – 9.4	18 – 40	Taguchi method for optimal mix
Chen et al., 2013	1450 - 1712	-	320 – 440	0.28 – 0.34	-	32.1 – 46.7	15.2 – 25	Admixutres (polymers & superplasticizers), fly ash & silica fume used
Bhutta et al., 2012	1620 - 1800	70, 85, 89	260, 300, 320	0.27, 0.30	-	10 – 20	15 – 30	High performance mix reached 35MPa

2.2.3 Hydrological and Water Quality Performance

Similar to other LIDs, the performance of permeable pavement systems is evaluated based on stormwater runoff quantity and quality. Permeable pavements have been shown to be successful in reducing runoff and pollutant loads in a variety of locations (Pratt et al., 1995; Dierkes et al, 2002; Barttebo & Booth, 2003; TRCA, 2006; Coupe et al., 2008; Tota-Maharaj & Scholz, 2010;). In a 6-year study of different types of permeable pavements in Renton, Washington, found the systems capable of infiltrating virtually all precipitation including intense (7.4 mm/hr) storm events (Brattebo & Booth, 2003). The systems also showed to have lower effluent concentrations of zinc, copper, and motor oil than in the runoff of asphalt pavements. Similarly, permeable pavements have been used to infiltrate 93% of runoff and have potential for controlling small storm events (< 2 cm) and for treating 'first flush' runoff for larger events (Dreelin et al., 2006). The Toronto and Region Conservation Authority investigated a pervious pavement parking lot in King, Ontario and found the pavement to produce no runoff even in large events (34 mm) (TRCA, 2006).

Permeable pavement systems have also been shown to remove fecal coliform and the use of geotextile layers have been shown to significantly improve the removal of nitrates and ammonium (Tota-Maharaj & Scholz, 2010; Tota-Maharaj et al, 2012). Permeable pavements also have been shown to reduce metal loadings from runoff. 50 years of runoff were simulated and permeable pavements were shown to retain 96 to 98% of copper, lead and zinc (Dierkes et al., 2002). Additionally, the authors showed that groundwater contamination due to metals should not be a concern in permeable pavement systems, even after 50 years. This result was verified in the field which showed very low potential of heavy metal and hydrocarbons entry to soil and groundwater in a 15-year-old permeable pavement system. Rushton et al. in 2001 investigated a parking lot in Tampa Florida where permeable pavements parking lots with swales reduced pollutant loads by over 75% for metals and total suspended solids when compared with asphalt pavements with no swales. **Table 5** provides a summary of performances of permeable pavement systems.

Table 5: Summary of Permeable Pavement performance (% Removal) found in literature

Reference	Location	Runoff	TSS	P/TP	NH_4-N	Cu	Pb	Zn	Fecal Coliform
Tota-Maharaj & Scholz, 2010	Edinburgh, Scotland	-	-	78	85	-	-	-	99
Brattebo & Booth, 2003	Renton, Washington	97	-	-	-	-	-	-	-
Dreelin et al., 2006	Athens, Georgia	93	-	10	-	-	-	80	-
Dierkes et al., 2002	Lab experiment, Germany	-	-	-	-	89 - 96	89 - 98	72 - 97	-
TRCA, 2006	King, Ontario	100	-	20	-	-10	Not detectable	50 - 60	-
Crookes et al., 2017	St. Catherines, Ontario	21-100	90	88	-	73	-	70-90	-
Rushton, 2001	Tampa, Florida	50	>75	76	80	94	93	89	-

2.2.4 Maintenance Concerns

The previous section highlighted the ability of permeable pavements in reducing stormwater runoff and pollutant loads. Although newly installed permeable pavements can have sufficient infiltration rates for even rare storm events (Brattebo & Booth, 2003; TRCA, 2006) over time, accumulation of pollutants (sediments, organic matter, and debris) within the voids of the concrete and on the surface causes the SIR of the pavement to decrease, reducing its performance in terms of both runoff reduction and pollutant load reduction. The accumulation of fine sediments in particular has been shown to severely affect SIR (Bean et al., 2007). Factors such as traffic, near by presence of fine sediments, sediment size distribution in runoff, presence of vegetation, the need for winter maintenance (salt/sanding), and age have indicated to influence clogging (Bean et al., 2007; Fassman & Blackbourn, 2010; Kayhanian et al., 2012).

Clogging of permeable pavement systems has been shown to greatly reduce the SIR. Bean et al. (2007) conducted an investigation of SIR of different types of permeable pavements. The authors found SIRs in PCPC pavements as high as 7000 cm/hr where there was no visible clogging. Other locations with visible buildup of fine sediments had SIRs as low as 11 cm/hr. PCPC pavements specifically were found to range between 13 cm/hr to 4000 cm/hr (130 mm/hr to 40,000 mm/hr). Winston et al. (2016) found that newly installed PA pavements yielded SIRs up to 3000 cm/hr whereas pavements 25 – 30 years old had SIRs as low as 0.3 cm/hr. Valeo & Gupta (2018) conducted infiltration tests using the ASTM C1701 method on PA and found SIRs as high as 302.4 cm/hr and as low as 0 cm/hr in clogged locations.

Due to this degradation of SIR, maintenance has been recommended to restore clogged pavements so that their long term hydraulic and environmental benefits can be fully realized. It has been suggested that clogging, in most cases, occurs near the pavement surface. ‘Simulated Maintenance’ was conducted on concrete grid pavers where the first 13 – 19 mm of buildup was removed from the void spaces between the pavers resulting in significant improvements of SIR (Bean et al., 2007). Cored samples were collected from various pervious concrete pavement systems and Computed Tomography scans were conducted to investigate clogging (Kayhanian et al., 2012).

Results show that in general, the porosity of the top layer (25 mm) of the cored samples were generally lower, suggesting that clogging takes place near the top of the pavement system. The authors also found that age plays a significant role in pavement performance and in general, older pavements had poorer performance than newer ones.

A four-year study (Kumar et al., 2016) in Chicago investigated the performance of three types of porous pavement systems including porous asphalt, porous concrete and permeable pavers. The pavements were subject to winter maintenance involving snow removal via snowplows and de-icing salts composed of sodium, calcium and magnesium. These pavements underwent preventative maintenance that involved sweeping with a general street sweeper twice a year. After the first year, the infiltration capacity of the parking lots decreased slightly (2 – 10%). The following 2 – 3 years showed substantial decrease in performance in all three types of pavements where the infiltration capacity decreased 80 – 90% and the most significant decline occurred after year 3. Out of the three types of pavements, porous concrete performed the best followed by permeable pavers and porous asphalt. The authors suspect this may be due to the higher traffic present in the porous asphalt parking lot. The authors conclude that more robust and targeted maintenance procedures are necessary for these pavement systems and vacuum suction and power washing are recommended.

Different pavement types in North America and Europe and the effect of different maintenance techniques were investigated (Winston et al., 2016). Simulated maintenance was conducted on grid pavers by removing the first 2 cm of fill material. Grid pavers also underwent mechanical sweeping, air sweeping, and vacuuming. Porous asphalt pavements were subjected to pressure washing, vacuuming, and a combination of the two. Milling, which involves removing some amount of the top layer of pavement, was also conducted on these pavements at three different depths (0.5 cm, 1.5 cm, and 2.5 cm). Results show that removing the 2 cm of fill material for grid pavers did not significantly change the SIR. The authors suspect that this is due to poor construction as fine aggregates were observed throughout the underlaying pavement. This highlights the fact that proper construction procedures are vital for these technologies. Additionally, vacuuming, pressure washing and their combination showed significant improvement of SIR. Milling of the porous asphalt pavements at a depth of 0.5 cm showed to increase

SIR substantially with diminishing returns with subsequent depths. The study highlights the fact that different pavements require different types of maintenance techniques and non-traditional maintenance such as vacuuming, pressure-washing and milling appear to be more impactful in restoring SIR. Regular maintenance and the right type of maintenance is essential in ensuring long-term hydraulic performance of permeable pavement systems (Winston et al., 2016; Sehgal et al., 2018).

It is becoming increasingly understood that maintenance for these technologies is needed for sufficient hydraulic performance (Bean et al., 2007). The aforementioned methods to conduct maintenance are expensive and thus are either performed infrequently or not performed at all leading to permeable pavement systems that no longer perform at their designed infiltration capacity (Liu et al., 2020). Moreover, maintenance cannot be done optimally unless the clogging status of the pavement is known. Thus, determining the status of the pavement is a way of encouraging maintenance and therefore enabling the pavement to perform as it is designed.

2.3 Determining Permeable Pavement Hydraulic Performance

2.3.1 Contact Methods

The most common performance metric for permeable pavement systems is the Surface Infiltration Rate (SIR) which is inversely proportional to the amount of clogging and is likely dependant on the quality of the first 1” of pavement depth (Kayhanian et al., 2012). Several methods of determining the SIR have been investigated. Early methods involved using permeameters in the field to determine the permeability of the pavement using either a constant head or falling head instead of destructive approaches that involved collecting cored samples (Cooley & Brown, 2000; Fwa et al., 2001). Other methods used double-ring infiltrometers that were adopted from the standard method of determining the infiltration rate of field-soils (Bean et al., 2007; Fassman & Blackbourn 2010). Brown & Borst (2014) used single ring infiltrometers with a constant head on various types of permeable pavement systems. Custom methods of determining pavement performance have also been conducted. Lucke & Beechman (2011) used double ring infiltrometers but in locations where the infiltration rate was too high, an inundation method where a larger volume of water was used. Lucke et al. (2015) used a

custom infiltrometer to measure the performance of interlocking concrete pavers that was shown to provide accurate results while being much more efficient than the standard method. Other methods of rainfall simulation and large-scale surface inundation have also been investigated and shown to have promising results, although likely impractical for widescale adoption (Boogaard et al., 2014; Lucke et al., 2014).

The various methods and variability of results in these different methods led to the establishment of standard methods for determining the SIR for two distinct types of permeable pavement systems. ASTM C1701 and C1781 were developed for porous concrete and interlocking concrete pavers, respectively. It has been shown that C1701 can be utilized for porous asphalt pavements as well (Valeo & Gupta, 2018). Both methods utilize a single-ring infiltrometer whereby water is poured at a constant head. Pre-wetting of the surface is done where 3.60 kg of water is initially poured on the location. If the amount of water required for pre-wetting takes less than 30 seconds to infiltrate, 18 kg of water is used for the actual infiltration test. However, if the prewetting time is greater than 30 seconds, 3.60 kg of water is used instead. The infiltration rate is then calculated with the following equation:

$$I = \frac{KM}{\emptyset_{ring}{}^2 t} \qquad \textbf{(Equation 1)}$$

where I is the surface infiltration rate (mm/min), K is a constant with units (mm^3*s/(kg*min), M is the mass of infiltrated (not pre-wet) water (kg), $\emptyset_{ring}$ is the inside diameter of the infiltration ring (mm), and t is the time for the water to infiltrate the surface (s). Typically, a large diameter single infiltration ring is used as opposed to a double ring infiltrometers to reduce variability and improve repeatability of results where larger ring sizes are recommended for pavements that yield high infiltration rates (Valeo & Gupta 2018). In general, this method was intended to provide some measure of surface clogging in permeable pavement systems and to reduce variability between different methods of determining SIR. Have

Since these standards were introduced, several studies investigated their accuracy by comparing them to other methods of determining SIR and identified several issues. Brown & Borst (2014) used a custom infiltrometer that consisted of a Polyvinyl

chloride pipe compressed onto a Neoprene sheet to prevent lateral flow rather than the ASTM C1701/C1781 standard recommendation of sealing the edges with putty. The authors concluded that the standard's method of pre-wetting before the actual test is redundant for PCPC and interlocking concrete pavers and that a more efficient use of time would be to conduct only the pre-wet test on several locations. This would eliminate the burden of carrying large quantities of water to the site and increase the amount of data collected. Furthermore, the authors recommended strategically selecting locations as opposed to randomly conducting` tests. Unfortunately, neither the research findings of Brown & Borst (2014) nor ASTM provided a method for choosing optimal locations. Winston et al. (2016) used a custom 'Simple Infiltration Test' that involved a rectangular infiltrometer which allowed less lateral flow of water and produced higher SIR values with lower variability than the ASTM method. Valeo & Gupta (2018) investigated a simple falling head test using a custom-built apparatus and found less variability in infiltration rates when compared to standard methods. They also found that the standard test can be applicable on porous asphalt pavements and confirmed that it can be used as an indicator of surface clogging.

These various contact-methods of determining permeable pavement performance can take hours to complete if a pavement is severely clogged (Valeo & Gupta, 2018). Additionally, the logistics of bringing large amounts of water to a pavement site can be troublesome. Moreover, conducting multiple tests at different locations on-site for an accurate representation of the status of the pavement system is vital, however can be labor intensive and costly (Brown & Borst, 2014). For these reasons, and the fact that maintenance is often overlooked, many installed permeable pavement systems no longer perform at their designed infiltration rate. Other, non-contact, methods for determining SIR need to be explored that aim to remove some of these drawbacks.

2.3.2 Non-Contact Methods

Non-Contact methods are those that do not need physical tests to be conducted for determining the infiltration rate of permeable pavement systems. These types of methods aim to decrease the cost of labor and optimize as well as encourage maintenance. Extensive research has not been conducted for these types of methods

however, developing a method for determining the infiltration rate of these pavements without having to conduct physical tests may encourage widespread adoption of permeable pavement systems by overcoming some of their aforementioned limitations. In essence, the problem is defined as finding a method of predicting the infiltration rate of a pavement system without conducting a physical test.

Haselbach et al. (2006) developed a theoretical relationship for predicting the permeability of sand-clogged PCPC pavements by investigating the permeability of sand and the porosity of an unclogged system. The method involved measuring the permeability of PCPC systems which were fully covered with sand using simulated rainfall. Their method was able to correctly predict a pavement's permeability if it were overlain with a layer of sand and had pores that were clogged. The authors concluded that their method provided a representation of the system when fully covered in sand and thus a lower limit of permeability of PCPC systems. Although the results are promising, the method only provides a theoretical lower limit of permeability and does not provide predictions of actual clogging. Additionally, the paper only focuses on sand as the primary sediment and thus is only feasible at coastal locations where sand is more prevalent. Finally, the paper uses an experimental PCPC system with sand as the subbase and the results may not be transferrable to more typical PCPC designs.

Kayhanian et al. (2012) investigated the effect of hydrologic and physical characteristics of 20 PCPC pavements by performing a detailed statistical analysis to determine which factors influenced a pavement's permeability. The authors found factors such as age, mass of particles under 38 μm, number of days over 38°C, and the presence of surrounding vegetation to have significant impact on pavement infiltration performance. The study provided useful insight towards the factors affecting clogging but did not provide a means of determining the actual status of clogging.

Image-based analysis has been increasingly used for investigating physical properties of pavements. Neithalath et al. (2010) reasoned that due to the non-linear relationship of porosity and permeability, other parameters are needed to correctly characterize permeability prediction of PCPC pavements. They extracted cored samples from permeable pavement parking lots and used image analysis and electrical property techniques were to extract physical parameters that were further used to predict

permeability. The images were used to extract data for pore size and pore area fractions and the electrical conductivity of the specimen was determined through an impedance measurement apparatus. Through these parameters, the authors successfully derived a theoretical expression that relates permeability to pore structure features such as porosity, pore connectivity and pore sizes. Although this method was successful in predicting the permeability of a pavement, it involved collecting core samples of pavements which can be costly and is destructive. Furthermore, comprehensive lab analysis of the samples is required. Obaidat et al. (2017) used cellphone images of hot mix asphalt samples to successfully determine their volumetric design properties. Parameters were extracted from the image and used to generate equations that successfully predicted a pavement sample's mix properties including voids in mineral aggregate, voids in total mix, and voids filled with asphalt. This shows the potential of using images for determining physical properties of pavements.

Valeo & Gupta (2018) investigated the use of cellphone cameras for capturing images of the surface of PA pavements and a proprietary pavement 'Porous Pave'. The authors extracted parameters of the image, namely the mean gray level (MGL) and performed a linear regression to evaluate the relationship of the MGL with the SIR. Since the sediment that clogs the pores are lighter in colour compared to the pores themselves, as a pavement becomes more clogged, the overall gray level of the image should be whiter. In other words, as the MGL increases, the SIR of the pavement should decrease. To investigate a positive relationship, the authors extracted the inverse of the MGL instead. The analysis proved to show promising results and the linear regression models showed that the inverse of the MGL has high correlation with the obtained SIR even with the proprietary pavement which had a non-traditional surface colour of red. The study was limited by a low number of collected data points and the authors highly recommend investigating other image parameters. Furthermore, due to the non-linear relationship of porosity and permeability (Neithalath et al., 2010), it may be beneficial to explore non-linear modelling approaches.

Radfar & Rockway (2016) used time domain reflectometers installed beneath the pavements to measure volumetric water content as water flows following a storm event. The authors used the collected data and rainfall characteristics such as rainfall depth,

peak intensity and duration with an Artificial Neural Network (ANN) (discussed in section 2.4) to predict a pavement's volumetric water content as a performance indicator. The authors concluded that the method was successful however, several limitations were identified. Installing devices under the pavement to monitor a pavement's status may be expensive. Furthermore, a single reflectometer produced its own model, thus installing multiple devices on one site would mean the construction and optimization of several different models. Additionally, the authors used a non-traditional performance indicator (volumetric water content) for measuring infiltration rate and its correlation to clogging is not entirely clear. Similarly, Razzaghmanesh & Borst (2018) used water content reflectometers as well as buried tipping buckets to investigate clogging progression of a permeable interlocking concrete pavement via statistical analysis. The results indicated that both methods were consistent in predicting clogging progression, however, most of the limitations identified for Radfar & Rockway (2016) can be applied to this study as well.

Abogadil et. al. in 2018 conducted a preliminary field investigation on a permeable pavement parking lot on York University's campus in Toronto, Canada where SIR tests were conducted after taking a photograph of the surface. Results indicated that many locations of the pavement were already subject to high levels of clogging. Due to this clogging, some SIR tests took over an hour to complete which proved how time consuming conducting these tests can be. A brief analysis on image properties and their relationship with the SIR was conducted and it was found that with this particular dataset, there was not a strong relationship. This was likely due to a low number of collected data points and a biased dataset in general. The data collected represented only clogged pavements and a proper analysis to determine the relationship between properties of the surface image and the SIR would require data to be collected over the full lifespan of a pavement. Detailed method and results of the field investigation can be found in Appendix A.

Non-contact methods have been identified ranging from mostly image analysis to prediction of clogging using collected data. These approaches have several limitations of their own and it is evident that their common objective of reducing maintenance cost has not been achieved. The success of image analysis in identifying physical parameters of pavement surfaces (Neithalath et al., 2010; Obaidat et al., 2017; Valeo & Gupta, 2018).

and the use of other modelling approaches such as artificial neural networks (Radfar & Rockway, 2016) needs to be further investigated.

2.4 Modelling and Machine Learning Technology

In basic terms, a numerical model is a simplified representation of a real-world system that is constructed to gain insight into select characteristics of a physical, biological, economic or social system. The best models are those that are able to give results close to real-world observations with the least amount of input parameters and model complexity (Devi et al., 2015). In hydrology, a runoff model is able to estimate runoff as a function of various parameters of watershed characteristics. Input data such as the amount of rainfall, drainage area along with watershed characteristics such as soil properties, vegetation, and ground water characteristics are used to build a relationship that can correctly simulate real-world observations. There are three major classifications of models: physical, conceptual, and empirical. **Table 6** provides an overview of these models and the focus of this thesis will be on empirical models. In summary, a physical model is an idealized representation of a real-world phenomenon that includes the underlying physical relationships between variables, a conceptual model uses semi-empirical equations and field data for calculations, and empirical models use mathematical equations and input data to carry out predictions. These models use state variables which are measurable and change with time and space (Devi et al., 2015). In hydrology, water movement processes are represented by finite difference equations. These models require a large number of parameters (e.g., soil moisture, groundwater table, topography, dimensions of downstream river, etc.) and their morphology. Due to their scale, the development of these models requires large amounts of computation and in-depth knowledge and expertise regarding input parameters, which is reported to be a major challenge (Mosavi et al., 2018).

Empirical models, or data-driven models, are observation-based models that use existing data as inputs with no consideration of mathematical relationships between parameters within the modelled system. Statistically based methods use regression and correlation models to find the functional relationship between inputs and outputs (Devi et al., 2015). Multiple linear regression (MLR), autoregressive integrated moving average

and Bayesian forecasting models are examples of common statistical data-driven methods for modelling flood prediction (Mosavi et al., 2018). These models were seen to struggle in short-term prediction and, in the absence of a large (10+ years) dataset, struggle to provide quantitative flood predictions. To that end, machine learning technologies, or advanced data-driven models, are being widely investigated as an alternative due to their potential in improving accuracy and prediction power. These models can numerically model non-linearity between variables solely based on historic data without requiring the knowledge of underlying physical processes (Mekanik et al., 2013).

Machine learning (ML) is a field of artificial intelligence that is able to easily find patterns and regularities in large datasets with relatively low computation cost while maintaining fast training, validation, and testing performance with much less complexity when compared to physical models (Mosavi et al., 2018; Carleo et al., 2019). ML can be classified into supervised learning, unsupervised learning, semi-supervised learning, reinforcement learning, transduction, and learning to learn (Ayodele, 2010). Among those, supervised learning and unsupervised learning are the most common. Supervised learning methods use pre-defined labelled training datasets for model optimization and prediction. Support vector machines, ANNs (Mekanik et al., 2013; Radfar & Rockaway, 2016), and Decision Trees are some common machine learning algorithms that use supervised learning. Unsupervised learning are typically classification or clustering algorithms that learn intrinsic properties of the dataset without being given any labelled datasets prior to training (Ayodele, 2010). Unsupervised learning is often related to how learning occurs in the real world; where humans and animals mostly learn from observation.

Table 6: Three main categories of models, in the context of hydrology (Devi et al., 2015)

Physical Model	Conceptual Model	Empirical Model
Mechanistic or white box model	Parametric or grey box model	Data based or metric or black box model
Based on spatial distribution, Evaluation of parameters describing physical characteristics	Based on modeling of reservoirs and include semi empirical equations with a physical basis	Involve mathematical equations, derive value from available time series
Require data about initial state of model and morphology of catchment	Parameters are derived from field data and calibration	Little consideration of features and processes of system
Complex model. Require human expertise and computation capability.	Simple and can be easily implemented in computer code	High predictive power, low explanatory depth
Suffer from scale related problems	Require large hydrological and meteorological data	Cannot be generated to other catchments
SHE or MIKESHE model, SWAT	HBV model, TOPMODEL	MLR, ANN, unit hydrograph
Valid for wide range of situations.	Calibration involves curve fitting make difficult physical interpretation	Valid within the boundary of given domain

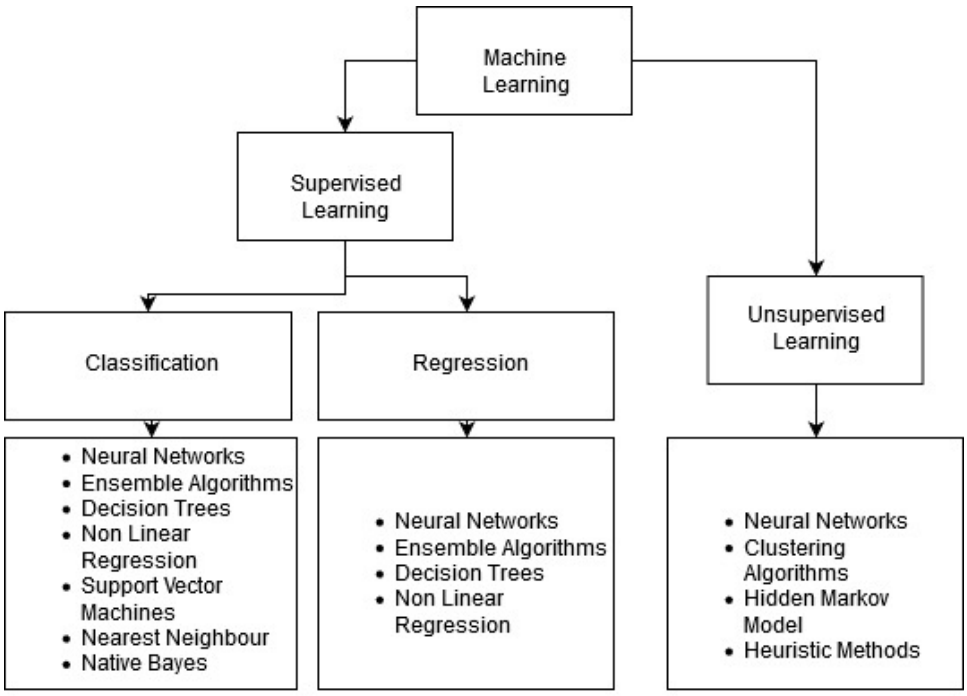


Figure 2: Breakdown of Machine Learning, specifically for Supervised and Unsupervised Learning. Revised from Ayodele, 2010.

Supervised learning models are trained based on given output values. In the training database, the inputs are accompanied by their corresponding ground truth outputs. Objective functions are built to calculate the differences between the input and output and the training process aims to minimize the objective function. Supervised learning deals with two types of problems: classification and regression. If the ground truth variables are discrete variables, the task is regarded as a classification problem. Otherwise, if output variables are continuous, the task is regarded as a regression problem. ANNs are a popular type of ML technique that is able to deal with both classification and regression problems and in both supervised and unsupervised learning (Zador, 2019). **Figure 2** shows a breakdown of ML and common algorithms and techniques used for the different types of learning problems.

2.4.1 Artificial Neural Networks

Conventional ML technologies (such as Support Vector Machines, Decision Trees, Clustering, etc.) are limited in their ability to process natural data in their raw form (LeCun et al., 2015). These ML technologies, also known as 'Shallow Learning' methods, require intense and careful engineering to extract features from raw data (such as an image) to a processed vector that the model could use for solving either regression or classification problems. 'Deep' neural networks or ANNs have made tremendous advances in recent years and were used across many fields and for a multitude of applications such as image classification, speech recognition, rainfall forecasting (Mekanik et al., 2013), flood prediction (Mosavi et al., 2018), and image feature extraction. Compared to traditional shallow learning methods, ANNs are able to extract higher-level features from large data inputs (e.g. an image), can model non-linearity between inputs and outputs, and are able to be used in multiclass classifications.

The idea of ANNs is inspired by the biological neuron. A biological neuron consists of three main components: (i) dendrites that receive input signals, (ii), a cell body, which gathers and processes the weighted input signals, and (iii) an axon that transmits the output signal to other neurons. Similarly, in an ANN, networks are able to learn knowledge of the inputs through the learning process, and the weights that interconnect the neurons store knowledge (Samarasinghe, 2016). **Figure 3** shows the comparison between a biological neuron and ANN.

ANNs are able to perform a variety of tasks including prediction, pattern classification, clustering and forecasting. They are powerful tools used to fit functions to data and are able to fit arbitrarily complex nonlinear models to multidimensional data with high accuracy. Often called universal approximators, ANNs can be thought of as the next step to multivariate techniques such as MLR and nonlinear regression (Samarasinghe, 2016).

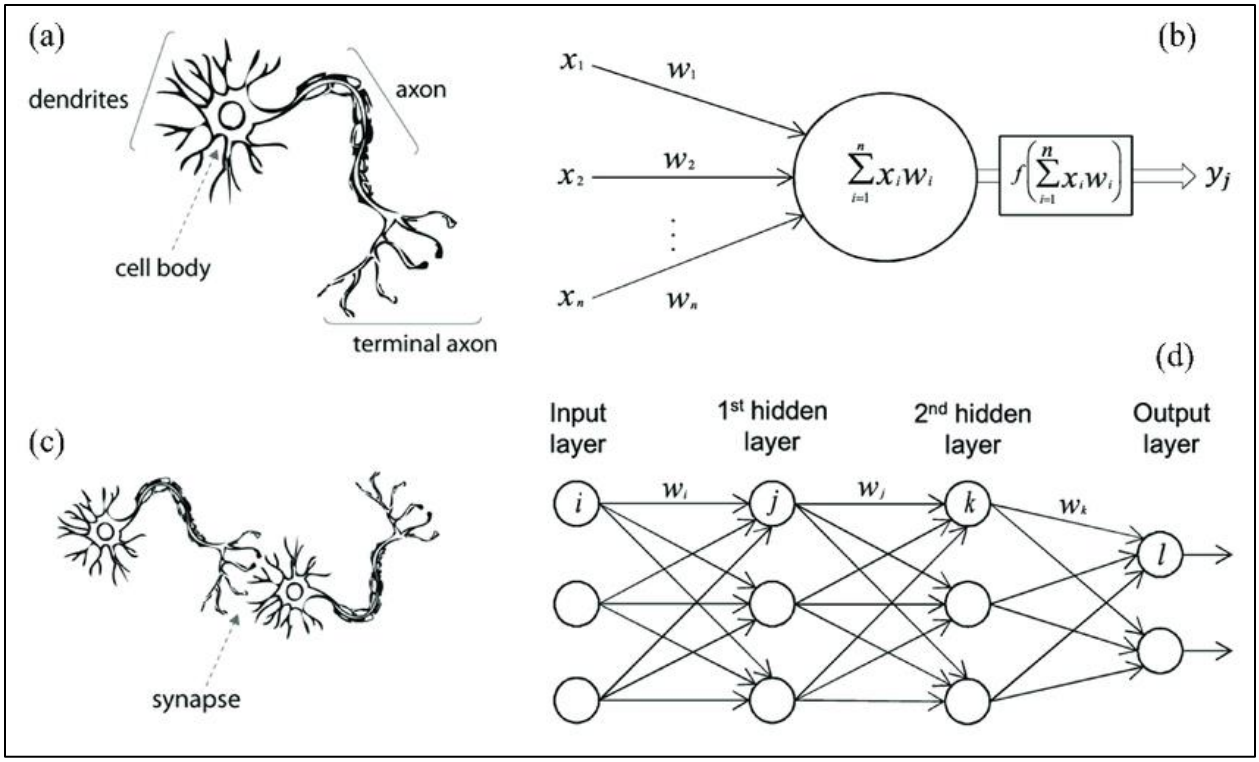


Figure 3: A comparison of a biological neuron and an artificial neural network: (a) human neuron; (b) artificial neuron; (c) biological synapse; and (d) Artificial Neural Network (Multi-layer perceptron) (Meng et al., 2020).

A complete ANN consists of several layers and a number of neurons that connect the input layer to the output via weights, biases, and activation functions. The following equation describes the operations conducted at each neuron

$$y_i = f_i(\sum_{j=1}^{n} w_{ij}x_j + b_i) \qquad \textbf{(Equation 2)}$$

where y_i represents the output of the i^{th} neuron which is used as the input of the next layer; f_i represents the activation function of the i^{th} neuron; w_{ij} represents the weight between the i^{th} and j^{th} neurons; x_j represents the j^{th} input to the i^{th} neuron; and b_i represents the bias of the i^{th} neuron.

The design of an ANN has three key aspects that need to be taken into account to ensure the accuracy of the model: (i) constructing an appropriate architecture (number of

hidden layers, neurons, etc.) (ii) The selection of a proper activation function (**Table 7**), (iii) selection of training process (backpropagation, stopping criteria) (Huyan, 2019).

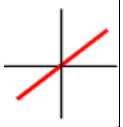
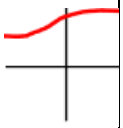
Feed-forward neural networks are the simplest type of ANN where information moves in only one direction; from the input layer to the output layer. The simplest feed-forward neural network, often called a single-layer perceptron, has one input and output layer that is connected through weights and activation functions. Recurrent Neural Networks are another class of ANNs where input data is propagated in two directions, forward and backwards. This allows the network to exhibit temporal dynamic behavior. The most common type of ANN is the multi-layer perceptron.

2.4.1.1 Multi-layer Perceptron

The multi-layer perceptron (MLP) is a type of feed-forward neural network that is praised for its ability to model non-linear functions and its performance in handling new, unseen datasets. Similar to the single-layer perceptron, the MLP has input and output layers, however, the main difference is that the weights and activation functions are controlled by 'hidden' layers. These hidden layers contain neurons that augment the input data using non-linear activation functions. The outputs of the neurons are scaled by the connecting weights and fed forward to the next layer of the network. Early in its development, the multi-layer perceptron was not feasible due to the practice of hand-engineering features of the network (Gardner & Dorling, 1998). It was discovered that multi-layer architectures can be trained using backpropagation.

Training MLPs involves changing the values of the weights within the network such that accurate predictions can be made. Backpropagation was introduced for training where weights of the hidden layers are optimized by minimizing the error objective function through gradient descent (Gardner & Dorling, 1998). In the 1990s, there was concern among the machine learning and computer vision communities that using this method could trap the model in local minima (LeCun et al., 2015). The backpropagation method has two adjustable parameters that help in avoiding this problem: learning rate and momentum.

Table 7: Activation functions and their characteristics. (Huyan, 2019)

Name	Plot	Equation	Derivative	Monotonic	Derivative Monotonic	$f(x) \approx 0$ at $x \approx 0$
Linear		$f(x) = x$	$f'(x) = 1$	Yes	Yes	Yes
Binary Step		$f(x) = \begin{cases} 0 \; for \; x < 0 \\ 1 \; for \; x \geq 0 \end{cases}$	$f'(x) = \begin{cases} ? \; for \; x < 0 \\ 0 \; for \; x \geq 0 \end{cases}$	Yes	No	No
Logistic Sigmoid		$f(x) = \frac{1}{1 + e^{1-x}}$	$f'(x) = f(x)(1 - f(x))$	Yes	No	No
Tanh		$f(x) = \tanh(x)$ $\frac{2}{1 + e^{-2x}} - 1$	$f'(x) = 1 - f(x)^2$	Yes	No	Yes
Rectified Linear Unit (ReLU)		$f(x) = \begin{cases} 0 \; for \; x < 0 \\ x \; for \; x \geq 0 \end{cases}$	$f(x) = \begin{cases} 0 \; for \; x < 0 \\ 1 \; for \; x \geq 0 \end{cases}$	Yes	Yes	No

After training, hidden layers learn to represent the network's inputs such that the network is able to predict target outputs with high accuracy. The application of MLPs can be categorized into prediction (regression) problems, function approximation, and pattern classification (Gardner & Dorling, 1998). In general, the main objective of MLPs is to find an unknown function, f, (or a set of functions) by finding the relationship between input vectors in X and output vectors in Y. The function f is optimised during training such that the network's outputs for the inputs in X are as close as possible to the target values in Y. As mentioned earlier, the training process involves adjusting the weights in the architecture.

The use of adjustable weights and activation functions allows MLPs to find non-linear relationships in X and Y while not needing prior knowledge of the physical relationship between them. This is one the main benefits of using MLPs over traditional linear regression analysis. If the relationship between X and Y is clearly non-linear, using linear regression for analysis would be inappropriate. Non-linear regression analysis may be a viable option but would only be applicable if the non-linear relationship that is observed in the physical world is understood (Motulsky, 1987). There are extreme non-linear relationships in the real world and attempting to understand the physical relationship is often not ideal. Thus, the use of MLPs to investigate these non-linear relationships is useful.

One of the drawbacks of MLPs is when dealing with large input data with multiple dimensions, for example an RGB image. Since MLPs are fully connected, the number of parameters in the network becomes unmanageable for large images (LeCun et al., 2015). Convolutional Neural Networks (CNNs), a type of ANN that is able to take dimensional arrays as inputs are able to solve the limitations of MLPs with respect to these types of inputs. A brief literature review on CNNs can be found in Appendix D. The use of CNNs may be able to provide additional insights due to the ability to use full image arrays as inputs. ANNs, on the other hand, would need extracted parameters representing the images such as the mean, variance, and skewness of grayscale image histograms.

2.5 ANN applications in Civil Engineering

Some examples of ANNs in civil engineering applications were already discussed above. In general, MLPs are used mostly for some type of regression analysis or prediction problem whereas CNNs are typically used for object detection or classification. This mostly holds true in the field of civil engineering. The use of MLPs has been used extensively in the field of hydrology for rainfall prediction and flood forecasting (Mekanik et al., 2013; Mosavi et al., 2018; Snieder et al., 2020). These approaches typically involve collecting local rainfall and climate data as inputs to predict a downstream water level. Mekanik et al. (2013) compared an MLP approach to multiple linear regression for predicting rainfall and found that the MLP performed better at predicting peak rainfall and had generally lower error. Abbot & Marohasy (2014) predicted rainfall with an optimized MLP by carefully selecting input variables for the model and compared the results with traditional physical based models. The MLP model proved to be superior with lower errors than the physical model. Snieder et al. (2020) highlighted the importance of selecting adequate input variables for flood forecasting and ANNs in general and investigated different input variable selection methods. Moon & Kim (2011) used features from images and properties of cracks as inputs to an MLP for predicting the presence of cracks in concrete. Radfar and Rockaway (2016) used MLPs to predict clogging progression of porous concrete pavements using rainfall data as inputs and volumetric water content as outputs.

The application of CNNs in civil engineering has mostly focused on rehabilitation and damage inspection. Most authors share a common objective in wanting to automate the inspection process and thereby reducing costs and encouraging repair and rehabilitation of aging infrastructure. Zhang et al. (2016) was the first to use a deep learning approach for crack detection of roadway pavements. The authors used a CNN and compared their method to traditional methods that used Support Vector Machines for classification and found their method to be significantly superior. These traditional methods involve making assumptions of the geometry of cracks but the use of CNNs allow raw images as inputs while still yielding more accurate classification. Traditional methods that use images or features extracted from images are sensitive to noise and

different lighting conditions. Cha et al. (2017) used CNNs to detect cracks in concrete in various different lighting conditions, shadows and noise. The model proved to remain accurate in various different conditions. Gopalakrishnan et al. (2017) implemented transfer learning to a CNN architecture trained on the ImageNet database and used 5 different classifiers (Neural Network, Random Forest, Extremely Randomized Trees, Support Vector Machines, Logistic Regression) instead of the traditional SoftMax layer for classification to predict the presence of cracks. The results indicated that the CNN for feature extraction coupled with a neural network for classification resulted in the most accurate prediction, however the network did not solve their objective of distinguishing joint lines from cracks. Similarly, Nhat-Duc et al. (2018) used CNNs for feature extraction and a separate MLP for classification and found it to be superior than traditional 'Sobel' and 'Canny' algorithms for edge detection. CNNs have also been used to investigate infrastructure distresses at different magnitudes. Hoskere et al. (2018) were able to use images of post-disaster structures collected through an unmanned aircraft vehicle for building-scale distress detection whereas Dung & Anh (2019) used a CNN for pixel-level crack detection as well as a decoder for semantic segmentation.

These methods of detecting cracks focus on only one type of distress (crack) and the problem is a binary classification problem (surface has crack or does not have crack). This led to investigating methods for detecting multiple types of distresses and the severity of the damage. Gao & Mosalam (2018) used four CNN models with transfer learning to identify object types, classify different distresses and their damage level. For object detection, models have incorporated the use of Region Proposal Networks that use anchors that are assigned labels of potential objects and are able to predict their location on the image using bounded boxes. Cha et al. (2018) used the Faster-RCNN model which is composed of a Region Proposal Network and the Fast-CNN to detect multiple types of distresses and their location on the image. Similarly, Liang (2019) used three models for post-disaster inspection of concrete bridges at the system, component, and damage level. **Table 8** provides a summary table of research in civil engineering that utilize CNNs.

Table 8: Summary of ANN usage in Civil Engineering

Reference	Research Objective	Method
Mekanik et al., 2013	Rainfall prediction	Linear Regression & MLP
Abbot & Marohasy,2014	Rainfall prediction	MLP
Moon & Kim, 2011	Crack detection	Image + MLP
Radfar and Rockaway, 2016	Clogging prediction	MLP
Zhang et al., 2016	Crack detection	CNN
Cha et al., 2017	Crack detection with different lighting conditions	CNN
Gopalakrishnan et al., 2017	Crack detection with different classifiers	CNN + MLP
Nhat-Duc et al., 2018	Crack detection	CNN + MLP
Hoskere et al., 2018	Crack detection (building scale)	CNN
Dung & Anh, 2019	Semantic Segmentation Crack detection (pixel scale)	CNN
Gao & Mosalam, 2018	Damage detection	CNN (transfer learning)
Cha et al., 2018	Damage detection	Faster-CNN
Liang, 2019	Object detection & Damage detection	Faster-CNN
Veeraragavan et al., 2018	Mix design moisture damage prediction	CNN
Jiang et al., 2018	Mix design prediction	CNN
Hoong et al., 2020	Recycle aggregate detection and sorting	CNN

Although these methods are able to achieve their objective of object detection and segmentation, their goal of automating the inspection process of buildings is still questionable. For a fully functioning process of automating distress detection, a functional and usable application in which end-users can take images and process them without having to build the network architecture and evaluate its performance is necessary. Moreover, the problems that have been discussed mostly deal with binary classification or object detection and segmentation. There is yet to be a method in civil engineering that uses images and CNNs to evaluate the *performance* of the object that it is detecting.

LID technologies have been shown to improve water quality while also having a positive impact on water quantity needs. Although these technologies have been shown to be beneficial, there is still room for improvement. It has been discussed that the performance of LIDs such as permeable pavements degrades due to clogging overtime and eventually maintenance is required to restore the pavement's infiltration capacity (Winston et al., 2016; Lucke et al., 2015). Methods of investigating clogging progression and SIRs of in-situ permeable pavements are typically on-site methods that are labour intensive and time consuming (Brown & Borst, 2014; Valeo & Gupta, 2018). Non-contact methods of determining a pavement's surface infiltration rate were investigated but the objective of lowering maintenance costs has not yet been achieved. Machine learning, more specifically artificial neural networks, seem promising in being able to find relationships in large amounts of data such as images and physical characteristics such as the infiltration rate of the pavement. This relationship needs further investigation.

CHAPTER 3: THESIS OBJECTIVES

3.1 Gaps in Knowledge

There are still aspects of permeable pavements that need to be further investigated for their optimization and widespread adoption. The literature review identified limitations of the technology and possible approaches for their optimization. These are summarized below.

3.1.1 Permeable Pavement Performance

The literature review has shown how permeable pavements have been effective in providing benefits for water quantity and water quality. Newly installed pavements are capable of achieving suitable infiltration during most storm events. Specifically, it has been shown that these pavements perform well in low-intensity storm events (< 2cm) where over 93% of the runoff is captured (Dreelin et al., 2006) while still being effective in more intense storms (7.3 mm/hr) (Brattebo & Booth, 2003). At the same time, these pavements can provide adequate source control water quality treatment for various pollutants that can be more effective than traditional best management practices. It has been shown that over time the infiltration capacity of these pavements greatly reduces. Sediments carried by runoff eventually clog the pores of these pavements and reduce their infiltration capacity. This clogging is suspected to mostly be a surface phenomenon and maintenance procedures have been employed for the restoration of these pavements. Without routine and proper maintenance procedures, the environmental benefits of these pavements are not fully realized and many pavements no longer perform at their designed infiltration capacity. Maintenance procedures differ based on the type of pavement, are expensive, are often overlooked, and cannot be conducted optimally until the status of the pavement is known.

3.1.2 Determining Pavement Performance

The standard methods of determining the performance of permeable pavement systems involve on-site testing of the infiltration capacity of the pavement. Other methods

of determining pavement performance have been investigated that involve different types of infiltrometers, permeameters, rainfall simulators, and large-scale inundation methods. These methods require special equipment and can take hours to complete, especially if the pavement is severely clogged. Consequently, these tests are rarely conducted and the clogging status of installed pavements is unknown. Alternative, less-labour intensive methods of determining the clogging status of these pavements can optimize maintenance and may encourage their adoption.

3.1.3 Predicting Performance

There have been initial investigations of predicting the infiltration performance of permeable pavements using rainfall data, however they require the installation of data collection devices and building separate models for each location which may increase the cost of installation and will require monitoring of the system (Radfar & Rockway, 2016; Razzaghmanesh & Borst, 2018). The use of images to extract physical properties of the pavement surface has shown promising results in predicting infiltration performance (Valeo & Gupta, 2018). With surface clogging suspected to be mostly a surface phenomenon, images of the surface may be highly correlated to the infiltration capacity of the pavement and warrants further investigation.

Artificial Neural Networks were explored and found to be accurate in finding relationships between variables and even able to remain accurate while maintaining non-linearity. These networks have been used extensively in many fields including in hydrology to predict rainfall, and flooding. Their application of predicting the performance of permeable pavements needs further investigation. The literature review identified some methods of predicting the hydraulic performance of permeable pavement systems, however, they required special equipment and were unintuitive. It was also discussed that the mean of the grayscale image may increase due to the accumulation of sediment on the surface and within the pores (Valeo & Gupta, 2018). Further investigation on how images change as SIR decreases is necessary. The change in images as PCPC slabs are clogged may provide meaningful insight for ANNs to predict the SIR.

3.2 Thesis Objectives and Tasks

3.2.1 Investigate Clogging Progression of Portland Cement Pervious Concrete Pavement Specimen

The majority of the research that investigated the performance of permeable pavement systems did field testing of actual pavements. This process is not ideal due to many pavements no longer operating as they are intended. Consequently, this leads to a lower number of collected data points due to long testing times or low-quality data where the actual performance of the pavement is not being represented. There is a need of investigating the performance of permeable pavements that involves subjecting these systems to controlled levels of clogging and determining their infiltration capacity as clogging progresses. This will be done by constructing lab-scale PCPC systems and conducting infiltration tests while simulating clogging. Images of the surface of the pavement will be taken and properties of the images will be extracted to investigate the relationship of the image and its associated surface infiltration rate.

3.2.2 Developing Models as a Non-Contact Method for Predicting Clogging

As identified in the review, non-contact methods of determining infiltration rates need to be further investigated. Current methods of predicting the clogging of these pavements have been shown to require special equipment, unintuitive, and require special monitoring. Much of the research that involved field data found that many currently installed pavement systems need maintenance to restore infiltration performance. Current methods of determining the infiltration performance of these systems have been shown to be labour intensive and time consuming. Due to this, maintenance of these pavements is rarely conducted. Therefore, there is a need of developing a non-contact method that is easily able to predict the infiltration performance of permeable pavement systems. An artificial neural network will be created that will use properties of the image as inputs to predict a pavement’s infiltration rate.

CHAPTER 4: DEVELOPMENT OF PERMEABLE CONCRETE TEST SPECIMENS

To investigate typical permeable pavement systems, a preliminary field investigation was conducted (see APPENDIX A) where SIR tests were performed on a permeable pavement parking lot on York University's campus. Prior to each SIR test, an image of the surface of the pavement was acquired using a cellphone camera. A brief analysis was conducted where the correlation between properties of the image data and the SIR was investigated. Results from the preliminary field investigation revealed that many locations of the in-service parking lot were already subject to high levels of clogging (SIR near 0) which highlights the fact that current in-service permeable pavement parking lots are not performing to their designed infiltration capacity. Due to this, the field investigation was inadequate and did not represent the full lifespan of a permeable pavement parking lot. Thus, to investigate the clogging progression of PCPC systems, laboratory scale pavements were constructed.

4.1 Laboratory Experiment and Setup

Due to the clogged status of pavements in the preliminary field investigation, collecting a full set of data that represents a pavement's full life span proved to be very difficult. Thus, lab scale pavements were constructed with the objective of artificially clogging them while simultaneously conducting infiltration tests to obtain a large dataset that represents the pavements full lifespan, from construction to becoming completely clogged. Four lab-scale pavement specimens were constructed in the High Bay Laboratory in the Bergeron Center of Engineering Excellence at York University.

Four identical lab-scale permeable concrete pavement specimens were constructed and included three layers: a permeable concrete slab with a thickness of 152 mm [6 in.], a 50.8 mm [2 in.] choker layer with 19 mm [0.75 in.] crushed stone, and a 152 mm [6 in.] aggregate base layer consisting of 50.8 mm [2 in.] crushed stone. In total, the specimen measured 355.6 mm [14 in.] deep and had a square cross-section with length of 350 mm [13.78 in.]. **Figure 4** shows the schematic of the lab-scale specimen in an acrylic container attached to a mobile wooden stand. The wooden stand was made with 38 x 89 mm [1.5 x 3.5 in.] pieces of wood and wheels were attached to for easy transport.

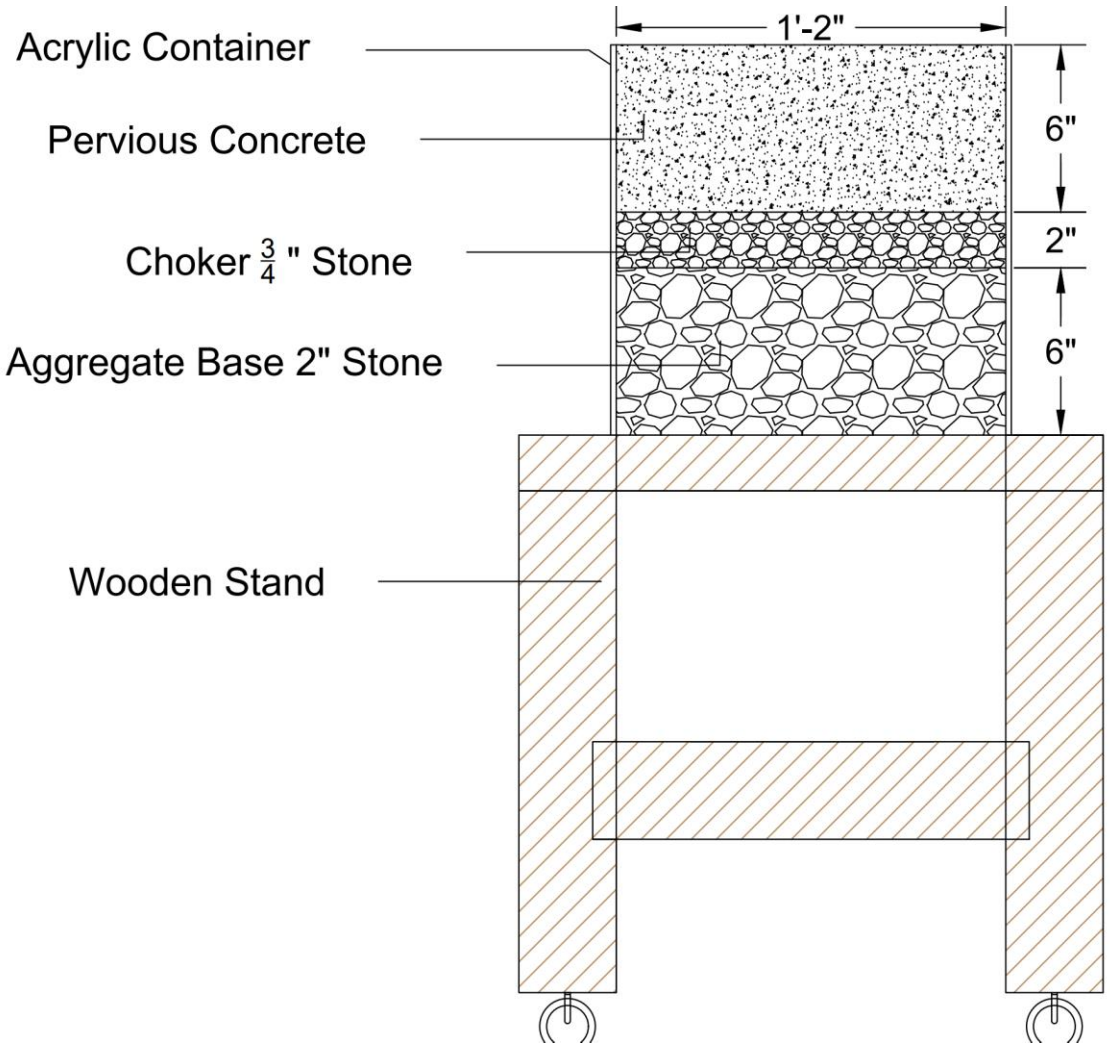


Figure 4: Schematic of experimental setup for conducting infiltration tests on PCPC specimen.

Twelve holes that measured 3 mm [0.125 in.] in diameter were drilled into each acrylic container to facilitate the drainage of water. A container was placed at the bottom of the stand, underneath the drainage holes, to capture the infiltrated water. **Figure 5** shows the acrylic container after the drainage holes were drilled. Once prepared, the layers were placed into the acrylic containers and were ready for SIR tests to be conducted.

Figure 5: Drainage holes in acrylic container

4.2 Iterative Mix Design for the Portland Cement Permeable Concrete Slab

Prior to completing the lab specimen, an iterative mix design was undertaken to obtain a PCPC slab with similar physical characteristics to typical in-service permeable concrete pavements. Namely, a mixture with compressive strength of 15 MPa and porosity of 20% was the target (OPSS, 2018). As identified in the literature review, permeable concrete omits the use of fine aggregates to create void spaces that are used for infiltration. Once the type of aggregate was chosen, the components that were iterated included the cement content, the amount of water, and the the use of chemical admixtures.

A 9.5 mm [3/8 in.] nominal size pea-stone coarse aggregate (ASTM C33 aggregate size no. 67) was used to ensure higher infiltration rates instead of using a well graded aggregate which typically have higher strengths. **Figure 6** shows the particle size distribution of the pea-stone coarse aggregate. Due to its nominal size of 9.5 mm [3/8 in.], most of the aggregate is shown to pass until the 9.5 mm sieve where the percent passing drastically reduces and quickly reaches zero. This means that most of the particles are around 9.5 mm [3/8 in.] in diameter. This aggregate was round rather than angular to ensure higher porosity. (Tennis et al., 2004).

The specific gravity and absorption of the aggregate were measured using ASTM C127 (ASTM, 2001). A sample of the coarse aggregate was taken and submerged in water for approximately 24 hours to fill the pores. The aggregates were then removed from the water and put on an absorbent cloth and patted dry to remove all visible surface water. The sample was then weighed and the mass was recorded and labelled as the saturated surface-dry condition. The sample was then submerged and disturbed in water to remove any entrapped air. The submerged mass was recorded. Finally, the sample was placed in an oven at a temperature of 110 +/- 5 °C [230 +/- 9°F] for 24 hours and then weighed, and recorded as the oven dry mass. The absorption of the aggregates was then calculated using **(Equation 3**:

$$A = \frac{[M_{SSD} - M_o]}{M_o} X\ 100 \qquad \textbf{(Equation 3)}$$

A = Percent Absorption

M_o = mass of oven-dry test sample in air (g)

M_{SSD} = mass of saturated-surface-dry test sample in air (g)

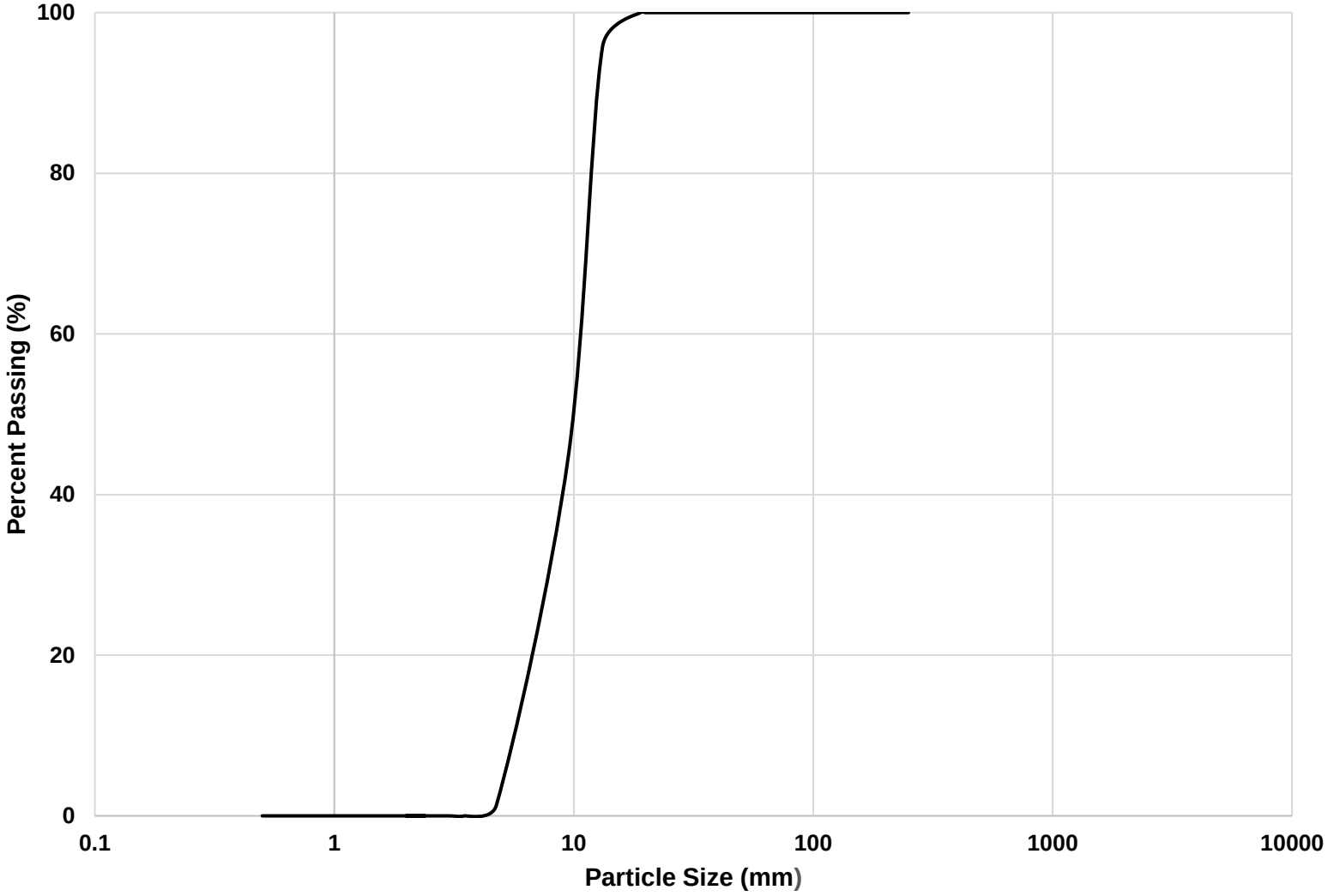


Figure 6: Particle Size Distribution of the pea-stone aggregate used for the PCPC mix

The bulk density and void content of the aggregate was then measured using ASTM C29 (ASTM, 2009). The rodding method was used since the nominal maximum aggregate size was less than 37.5 mm. A 2.8 L metal container was used and one third of the container was filled with the aggregate and leveled with fingers. The layer was rodded 25 times with a tamping rod. This step was repeated again with a second and third layer while ensuring the tamping rod did not penetrate the previous layer. The layers were leveled by pressing loose aggregates with fingers before proceeded to the next layer. Once the final layer was rodded, the mass of the sample was taken. The bulk density and void content were then calculated using **(Equation 4** and **Equation 5** respectively.

$$\rho_b = \frac{[M_a - M_m]}{V_m} \quad \textbf{(Equation 4)}$$

ρ_b = bulk density of the aggregate (kg/m^3)

M_a = mass of aggregate, plus the measure (kg)

M_m = mass of measure (kg)

V_m = volume of measure (m^3)

$$\varphi_{agg} = 100 * \frac{[(S * W) - \rho_b]}{(S * \rho_w)} \quad \textbf{(Equation 5)}$$

φ_{agg} = Void content of aggregates (%)

ρ_b = bulk density of the aggregate (kg/m^3)

S = bulk specific gravity as determined with test method ASTM C127 (kg)

ρ_w = density of water (kg/m^3)

The absorption of the coarse aggregate was found to be 1.48%. The dry bulk density and the void content of the pea-stone coarse aggregate was found to be 1655 kg/m^3 and 39% respectively. The measured aggregate properties were used as a starting point for the mix design procedure. With the bulk density and a known batching volume, the mass of the aggregate was calculated. Aggregate to cement ratios (*a/c*) and water to cement ratios (*w/c*) were iterated and ranged from 3.5 – 5.0 and 0.22 – 0.34 respectively. The aggregates were stored outdoors and thus, their moisture condition was unknown due to changes in weather. Consequently, the aggregates were oven-dried prior to each mix. Due to the aggregate being oven dried, the absorption percentage of the aggregate was taken in to account for when determining the total water content to be used in the mix design. Thus, the *w/c* ratio was calculated as:

$$w/c = \frac{\left(M_w - \left(A * M_{agg}\right)\right)}{M_c} \qquad \textbf{(Equation 6)}$$

A = absorption percent of aggregate, from **(Equation 3** (%)

M_c = mass of cement used in mix design (kg)

M_{agg} = mass of aggregates used in mix design (kg)

M_w = total mass of water in mix design (kg)

An 85 L [5.5 ft.^{3}] drum mixer was used for batching all trial mixes while a 190 L [6.7 ft.^{3}] pan mixer was used when casting the final concrete mixes for the four PCPC pavement slabs (see **Figure 7**). Dry ingredients were first mixed together for two minutes, or at least until the cement was well distributed and no clumps were visible. If a water reducing admixture was used, it was mixed with the water prior to adding it to the mixer. The wet ingredients were then added to the mixture and mixed for another two minutes or until a uniform mix was produced where the concrete was easily able to be formed into a ball and hold its shape.

(a) (b)

Figure 7: Rotating drum mixer shown in (a) used for iterating mix designs; revolving pan mixer shown in (b) used for casting slabs.

For each mix, six large (150 mm x 300 mm [6 in. x 12 in.]) cylinders were cast where three cylinders were used for 14-day and 28-day testing. The 14-day testing was conducted in lieu of 7-day tests due to limitations on access to lab spaces due to COVID-19. A standard proctor hammer was used for compaction instead of a tamping rod as in traditional concrete mixes. Three even layers were used and 5 to 10 blows of the standard proctor hammer were done for each layer. For the last layer, after being compacted with the proctor hammer, the concrete was rolled with a 305 mm [12 in.] long tamping rod with an end diameter of 9.5 mm [3/8 in.] to create a smooth and even finish. This procedure was adopted and slightly modified from Washington State Department of Transportation (Nassiri & AlShareedah, 2017). A 177.8 mm [7 in.] square steel plate was used to distribute the force from the proctor hammer to cast the slabs. The fresh masses of the cylinders were recorded and used to obtain a theoretical density and porosity. **Figure 8** shows the tools used for compacting the freshly cast PCPC cylinders. It should be noted that these theoretical porosities were only able to be calculated once the fresh mass of the cylinders were obtained.

Figure 8: Compaction tools used for casting permeable concrete cylinders and slabs. A large tamping rod, and standard proctor hammer were used for cylinders and a square steel plate with the proctor hammer was used for casting the permeable concrete slabs

Table 9 shows the trial mixture design proportions and their expected porosities. Mixes with the letter 'A' indicate the use of a high-range water-reducing admixture. The mix designs that are indicated with 'slab' were batched using the revolving pan mixer. The *w/c* ratio needed to be adjusted due to the differences in mixers and a slab was cast for investigating the infiltration of the mix. The aggregate-to-binder ratio (*A/B*) was also reported which is the ratio of aggregates to the cement, water, and admixture binder by weight. The cylinders were demoulded the day after they were cast. Due to pores, pressurized air was not able to be used to demould these cylinders. Instead, the cylinders were demoulded using a concrete cylinder stripping tool and a hammer to break the moulds (see **Figure 9**). Thus, the moulds that were used were plastic single time use moulds. The cylinders were then stored and allowed to cure under a polyurethane sheet until they were ready for testing after 14-days and 28-days.

Figure 9: Tools for removing concrete cylinder from their mould. Top: Cylinder mould stripping tool; Bottom: Hammer

For each mixture, a series of hardened concrete tests were conducted. First, a test to obtain the porosity of the cylinders was completed. The porosity test was conducted in accordance with ASTM C1754 (ASTM, 2012) with slight modifications. **Figure 10** shows the test setup for obtaining the porosity of the permeable concrete cylinders. A hole was cut into a wooden plank that was long enough for a weigh scale to hang over the side of a countertop. The plank was clamped to the countertop and the scale was placed on top and leveled. A wire basket was hung from a chain that attached to the scale and the scale was zeroed. While the apparatus was set up, the cylinders were prepped for testing.

Table 9: Iterations of PCPC mix designs

Mix ID#	Aggregate (kg/m^3)	Cement (kg/m^3)	Water (kg/m^3)	A/C	W/C	A/B	Expected Porosity (%)	Admixture (ml) Polyheed 980 - 780ml/100kg Cement
1111	1655	331	99.3	5.00	0.22	3.85	25.91	None
1122	1655	326	117.4	5.00	0.29	3.68	22.47	None
1133	1655	465.7	167.7	3.50	0.31	2.57	14.96	None
2233	1655	439	136	3.77	0.25	2.88	21.36	None
2234	1655	465.7	144.4	3.50	0.26	2.67	20.79	None
2235	1655	465.7	182.8	3.50	0.34	2.51	6.71	None
2236	1655	407.5	146.7	4.00	0.30	2.94	17.28	None
2237	1655	407.5	155	4.00	0.32	2.90	15.00	None
2237A (slab)	1655	407.5	155	4.00	0.32	2.88	5.04	117
2238A (slab)	1655	407.5	138.5	4.00	0.28	2.97	13.51	368
2240A (slab)	1655	407.5	134.5	4.00	0.27	2.99	14.37	368

The cylinders were lightly brushed with a wire brush to remove any loose particles. The dimensions of the cylinders were measured using jaw calipers. Two dimensions of the height of the specimen were taken. Due to the size of the cylinders, the calipers were unable to measure the diameter from the midspan. Instead, diameter measurements were taken on the ends of the cylinder. Four diameter measurements were taken for each cylinder, two at each end of the cylinder. The cylinders were then dried using an oven. The specimen was dried at 38 +/- 3 °C [100 +/- 5 °F] for 24 hours. According to ASTM C1754, this step is to be repeated until the weight change of the cylinders is less than 0.5%. This was not done due to restricted access to the laboratory space thus, only one day of drying was completed. This means that there was still potentially some water present in the void spaces; however, this also means that the results would be more conservative as the obtained void content would be less than the actual void content if the cylinders were dried according to the standard. The dry mass of the cylinders was recorded. The cylinders were then submerged in water using a hanging wire basket. Due to the size of the cylinders, instead of being placed upright as specified in ASTM C1754, the cylinders were placed sideways. While submerged, the cylinders were tapped with a rubber mallet 10 – 15 times to release air bubbles. The cylinders were submerged for 30 minutes while being simultaneously rotated and tapped with the mallet. Once air bubbles were no longer visible, the void content of each cylinder was calculated:

$$\varphi_{cyl} = \left[1 - \frac{K * \left(M_{cyl} - M_{sub}\right)}{\rho_w * \emptyset_{cyl}{}^2 * L}\right] * 100 \qquad \textbf{(Equation 7)}$$

φ_{cyl} = Void content of dry cylinder (%)

M_{cyl} = dry mass of the specimen (g [lb])

M_{sub} = submerged mass of the specimen (g [lb])

$\emptyset_{cyl}$ = average diameter of the specimen (mm [in.])

L = average length of the specimen (mm [in.])

K = 1 273 240 ($mm^3 \cdot kg)/(m^3 \cdot g)$ or 2 200 $(in.^3)/(ft^3)$

ρ_w = density of water at temperature of the water (kg/m^3 [lb/ft^3])

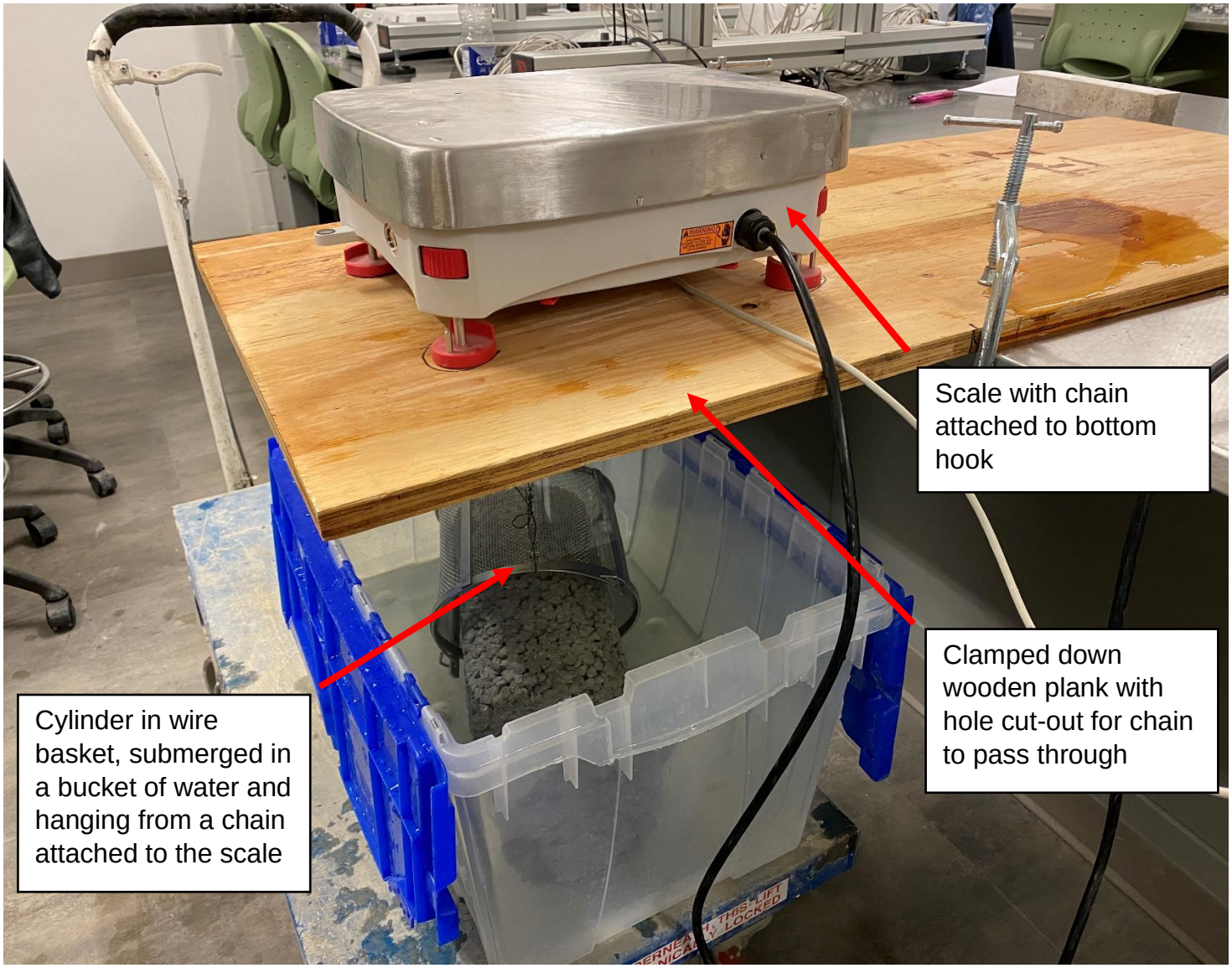


Figure 10: Test setup for obtaining void content of permeable concrete cylinders

The above void content test was conducted for every cylinder one day prior to conducting a compression strength test. The compressive strength was measured in accordance with ASTM C39 (ASTM, 2014). **Figure 11** shows a cylinder ready to be tested. Neoprene caps were used at the ends of the cylinders due to the porous nature of the cylinders and to ensure a more even distribution of loading. A displacement-based loading rate was used instead of a stress-based loading rate. Trial specimen were cast and the loading rate was adjusted until a suitable loading rate was found that provided a reasonable test duration of 5 – 8 minutes with a steady, no-shock loading and failure. A loading rate of 0.02 mm/s [0.0008 in./s] was chosen.

Figure 11: Test setup for measuring compressive strength of a PCPC cylinder with neoprene caps at the ends of the specimen

4.3 Results of the Iterative Mix Design

A total of eleven mix designs were developed and 6 cylinders each measuring 150 mm x 300 mm [6 in. x 12 in.] were cast per mix. Compressive strengths were measured for these mixes at 14 and 28 days after casting. Three cylinders were used on each test day and porosity was tested one day prior to compressive strength test and were only conducted for the 28-day cylinders, with the exception of mixes 2237AS and 2240AS. As mentioned previously, the recommended porosity and compressive strength for permeable concrete pavements in light traffic applications is 20% and 15 MPa respectively (OPSS, 2018). The results of the porosity test and the compressive strength test are shown in **Figure 12**.

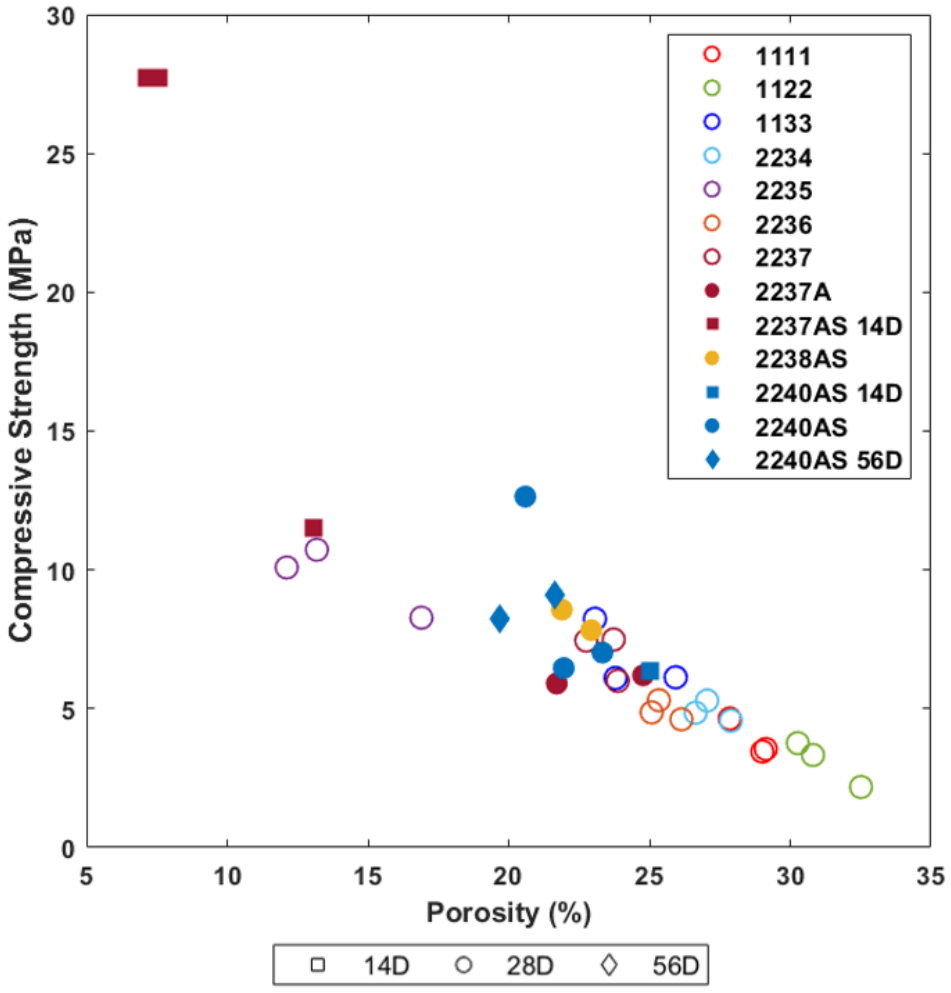


Figure 12: Results of the porosity and compressive strength tests for all mix designs

As expected, porosity is inversely proportional to compressive strength where a higher porosity would mean more void content, less concrete, and thus less compressive strength. A limited number of data points were collected in the 5 – 15 % porosity range since the target was to obtain a mix of around 20% porosity. In the figure, circles represent 28-day strengths where as squares and diamonds represent 14-day and 56-day respectively. Two mixes were tested for porosity at 14 days and only the final mix was tested at 56-days to provide insight on long term strength development. The ‘S’ in the mix designs indicate the use of the pan mixer where a sample PCPC slab was cast in addition to the 6 cylinders. Mix number 2237 was cast once with the drum mixer, then mixed again with the use of an admixture (2237A) and then once more with the pan mixer (2237AS). The use of the water-reducing admixture is seen in the figure with a filled marker. As seen in the figure, the pan mixer likely allowed for more uniform mixing to take place and the porosity decreased from around 22% (2237A) to around 8% (2237AS). Only 14-day tests were conducted on mix 2237A due to the initial results yielding properties that were very far from the target porosity and strength combination. Thus, to accommodate the use of the pan mixer, the *w/c* ratio was systematically reduced until the mix design procedure was finalized with mix 2240AS where the 28-day cylinders had an average porosity and compressive strength of 21.96% and 8.71 MPa respectively and the 56-day cylinders had an average porosity and compressive strength of 20.66% and 8.65 MPa respectively.

The 14-day compressive strengths were plotted with 28-day compressive strengths which can be seen in **Figure 13**. The results indicate that the compressive strengths of the permeable concrete mixes were generally higher at 14-days compared to 28-days. The 28-day cylinders were subject to a porosity test prior to undergoing the compression test. This involved placing the cylinders in the oven until their mass was consistent as well as submerging them in water. The results indicate that subjecting cylinders to porosity tests prior to conducting the compression test may have led to a decrease in measured compressive strength. It should be noted that not every mix was cast with 14-day cylinders, however; the figure indicates that the mixes tested, had higher 14-day compressive strengths compared to 28-day compressive strengths.

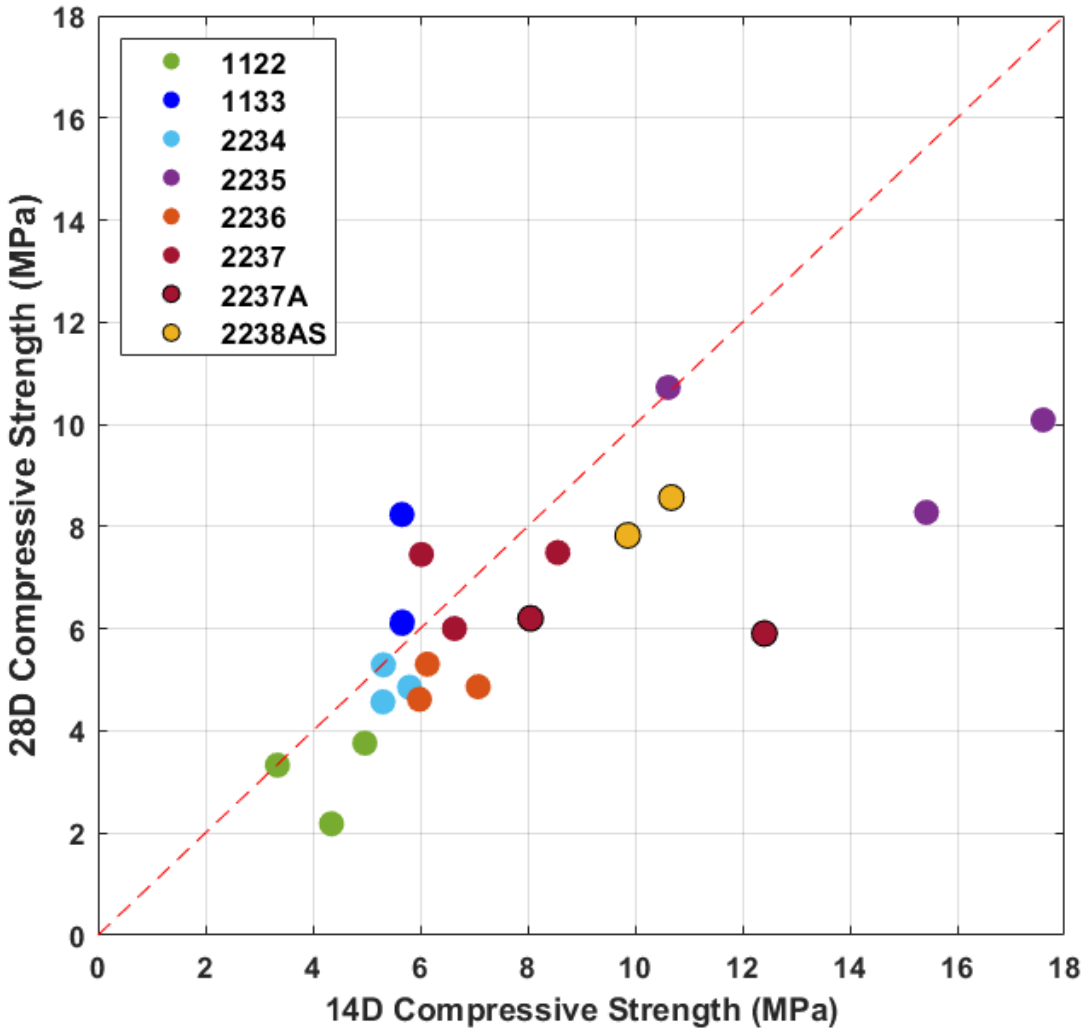


Figure 13: 14-day compressive strengths compared to 28-day compressive strengths for various PC mixes

4.4 Construction of Laboratory Permeable Concrete Specimen

Once the mix design was finalized, four slabs and six cylinders were cast. Wooden formwork for the slabs were constructed in advance of the cast date (see **Figure 14**). The side forms were made of spruce-pine-fir lumber with cross section 38.1 x 152.4 mm [1.5 x 6 in.] and were cut at 412.75 mm [16.25 in.] to allow for a slab side length of 336 mm [13.25 in.]. Plywood was used as the formed face of the slabs. Two applications of form oil were used on the wood surface prior to casting. The first application was done one day prior to casting and the second application was done an hour before casting the slabs. This was to ensure the wood did not absorb any moisture from the concrete as well as allow an easier stripping process.

Figure 14: Wooden forms for PCPC slabs.

Once the slabs were demoulded, they were cured using a similar regime as the cylinders where they were encased in a polyurethane sheet and left to cure for 28 days. After 28 days, the slabs were placed into acrylic containers that were already filled with a drainage layer stone (50 mm [2 in.]) and a bedding layer stone (19 mm [0.75 in.]) (see **Figure 15**). Four slabs were constructed with the intention of clogging three specimens while the fourth would act as a control where no clogging would take place. This would allow variances in the collected SIR (without clogging) to be captured and thus would allow for the calibration of the clogging regime for the other slabs. Wooden covers were used to cover the slab so that particles from the laboratory did not contaminate the surfaces.

Figure 15: Four lab-scale pavement specimen and their three distinct layers: PCPC slab, base layer, and drainage layer

4.5 Clogging of Laboratory Specimens

The literature review as well as the preliminary field investigation identified the issue of many in-service permeable pavements being subject to high levels of clogging. Thus, the lab specimens were created with the purpose of subjecting them to artificial clogging to obtain a full dataset that encapsulates a pavements lifespan from construction to becoming fully clogged. The laboratory specimens were subjected to artificial clogging where yearly equivalent amounts of sediment were deposited onto the slab surfaces. To calculate the amount of sediment required for clogging the specimens, yearly precipitation data, the amount of sediment loading, and tributary area of the pavement is required.

The experiment took place in Toronto, Ontario where the yearly precipitation was found to be 831 mm/year (Government of Canada, 2021). The average event mean concentration for total suspended solids originating from parking lots in Toronto was found to be 91 mg/L (City of Toronto, 2003). This value was rounded up to 100 mg/L. The tributary area for each pavement specimen was found using impervious to pervious ratios (I/P) to characterize the lifespan of the pavement as per the Urban Water Resources Centre (2013). The I/P ratio is the ratio of impervious to pervious areas that the pavement surface infiltrates water from. In practice, the pavement may receive runoff from nearby areas such as rooftops, sidewalks, downspouts, or other impermeable surfaces. Areas upstream from the pavement as well as the area of the pavement itself must be taken into account when calculating the amount of runoff received by a pavement. Permeable pavements with higher I/P ratios receive more runoff and higher pollutant loads than those pavements with lower I/P ratios. The Urban Water Resources Centre (2013) provides an outline of the expected lifespan of permeable pavements with differing I/P ratios and catchment descriptions. These lifespans were obtained by reviewing results reported in the literature and I/P curves were created by plotting laboratory tested lifespans with hydraulic conductivity. For the purposes of this research, an I/P ratio of four was taken which, according to the Urban Water Resources Centre, would result in a lifespan of four years if the pavement was located in an average suburb with a moderate number of trees. Each pavement specimen had a surface area of 0.1225 m^2 [0.0187 m^2/m^3]. With the I/P ratio and precipitation amount, the volume of runoff was calculated:

$$V_{runoff} = A_{slab} * \frac{I}{P} * R$$ **(Equation 8)**

$$V_{runoff} = 0.1225\ m^2 * 4 * 0.831\ m/year$$
$$V_{runoff} = 0.4072\ m^3/year$$

With the average suspended solid concentration of 100 mg/L, the yearly sediment loading was then calculated as:

$$M_{sediment} = V_{runoff} * TSS$$ **(Equation 9)**

$$M_{sediment} = 0.4072\ m^3/year * 100\frac{mg}{L} * 1000\frac{L}{m^3} * \frac{1}{1000}\frac{g}{mg}$$
$$M_{sediment} = 40.72\frac{g}{year} for\ each\ slab$$

Therefore, the total mass of sediment needed for clogging each slab was calculated to be 40.72 grams per slab per year. To accurately represent typical sediment from parking lot runoff, a representative particle size distribution was needed. Selbig and Bannerman (2011) conducted an investigation of the particle size distribution of sediment in stormwater runoff. **Table 10** summarizes the study's particle size distribution findings in parking lot locations. The table shows the percentage of sediment passing through each particle size sieve. **Table 11** shows the percent mass of the sediment in each particle size category. This was approximated with Figure 3 provided in Selbig and Bannerman (2011).

The Toronto and Region Conservation (TRCA) conducted an investigation where sediment with a known particle size distribution (PSD) was sent to various laboratories to determine the accuracy of various methods of obtaining the PSD (Goncalves & Seters, 2012). Comparison of PSD results showed good correlation among the labs but some differences were as high as 80% in certain size ranges. Multiple samples were taken across the Greater Toronto Area (GTA) including parking lots as well as mixed residential and commercial land use types. Results showed that the GTA has finer particle PSDs

than those collected in other cold climate jurisdictions. **Figure 16** shows the average PSD of runoff sediment from various parking lots in the GTA. Comparing these results with the results from the study in (Selbig & Bannerman, 2011), it can be seen that, in general, parking lots in Toronto have a finer particle distribution.

Table 10: Summary of particle size distribution found in parking lots represented as percent finer than (Selbig & Bannerman, 2011)

Statistic	Particle size, in micrometers								
	500	**250**	**125**	**63**	**32**	**14**	**8**	**5**	**2**
No. of Observations	94	94	94	94	94	94	94	94	94
Minimum	43	12	4	2	1	1	1	<1	<1
Maximum	100	100	99	97	95	93	85	71	37
Median	95	80	69	54	38	26	22	18	10
Mean	87	74	63	51	38	30	25	21	11
Standard Deviation	16	23	25	24	22	19	17	14	8

Table 11: Approximate percent contribution of total mass for particles in sediment of parking lot runoff. (Selbig & Bannerman, 2011)

Particle Size, in micrometers	Percent of Cumulative Mass
125 – 500	25
32 – <125	35
8 – <32	12
2 – <8	16
<2	12

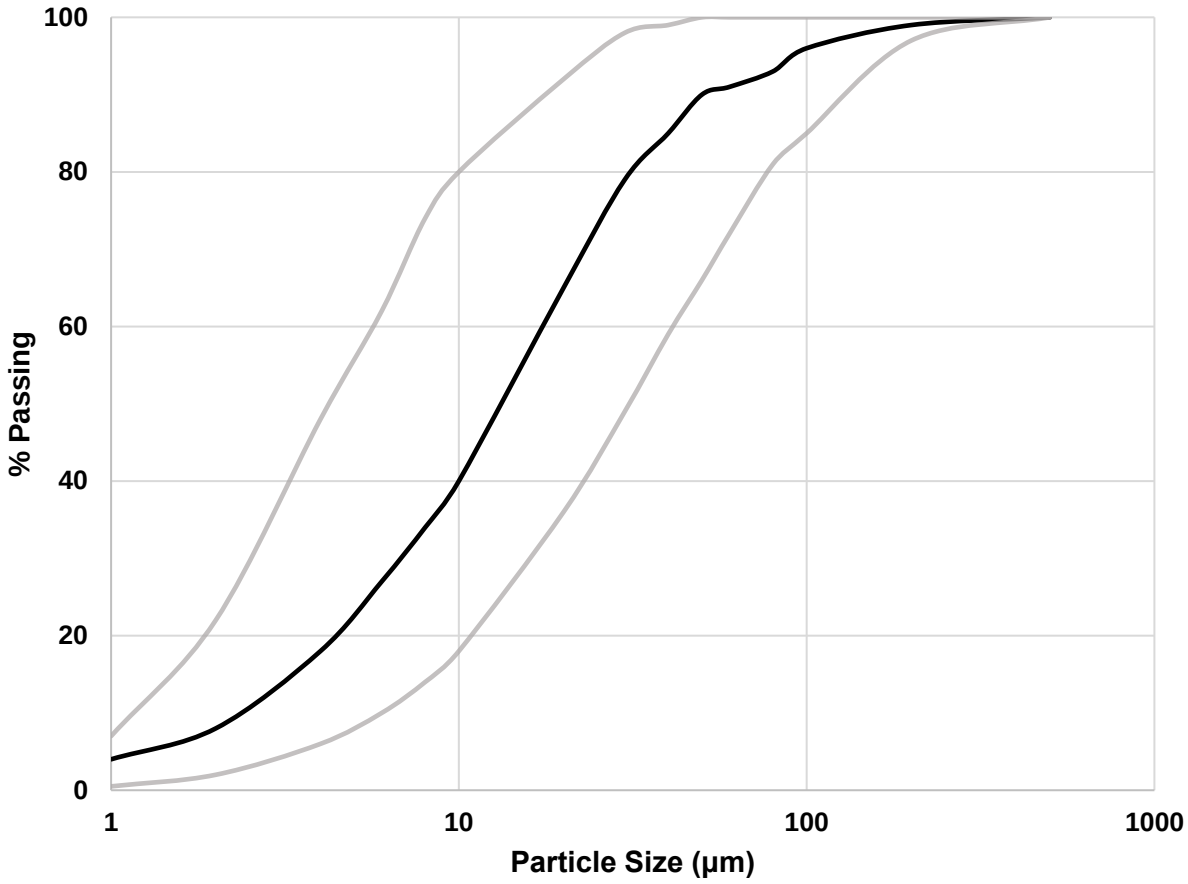


Figure 16: Particle Size Distribution of parking lot sites in the Greater Toronto Area. The black line represents the average of 13 parking lots while the gray lines are the extreme boundaries of the analyzed samples (Goncalves & Seters, 2012)

The particle size distribution of several clay and limestone mixtures were investigated with the purpose of obtaining a composition that resembles the distribution reported in Goncalves & Seters (2012). A limestone sediment was passed through a 600 µm sieve. This limestone was then mixed with Kaolin clay. Four different mixtures were tested. Namely, the clay-limestone mixtures included compositions of (by percent mass) 50-50, 60-40, 80-20, and 90-10. Three samples of each mixture were taken, and a Laser Diffraction Particle Size Analyzer was used to obtain each sample's particle size distribution (see **Figure 17**). With four mixtures with three samples each, a total of 12 particle size distributions were obtained. The distributions were inspected to ensure the data collected was adequate. For example, there should be no particles higher than 600 µm due to the sediment being sieved at 600 µm.

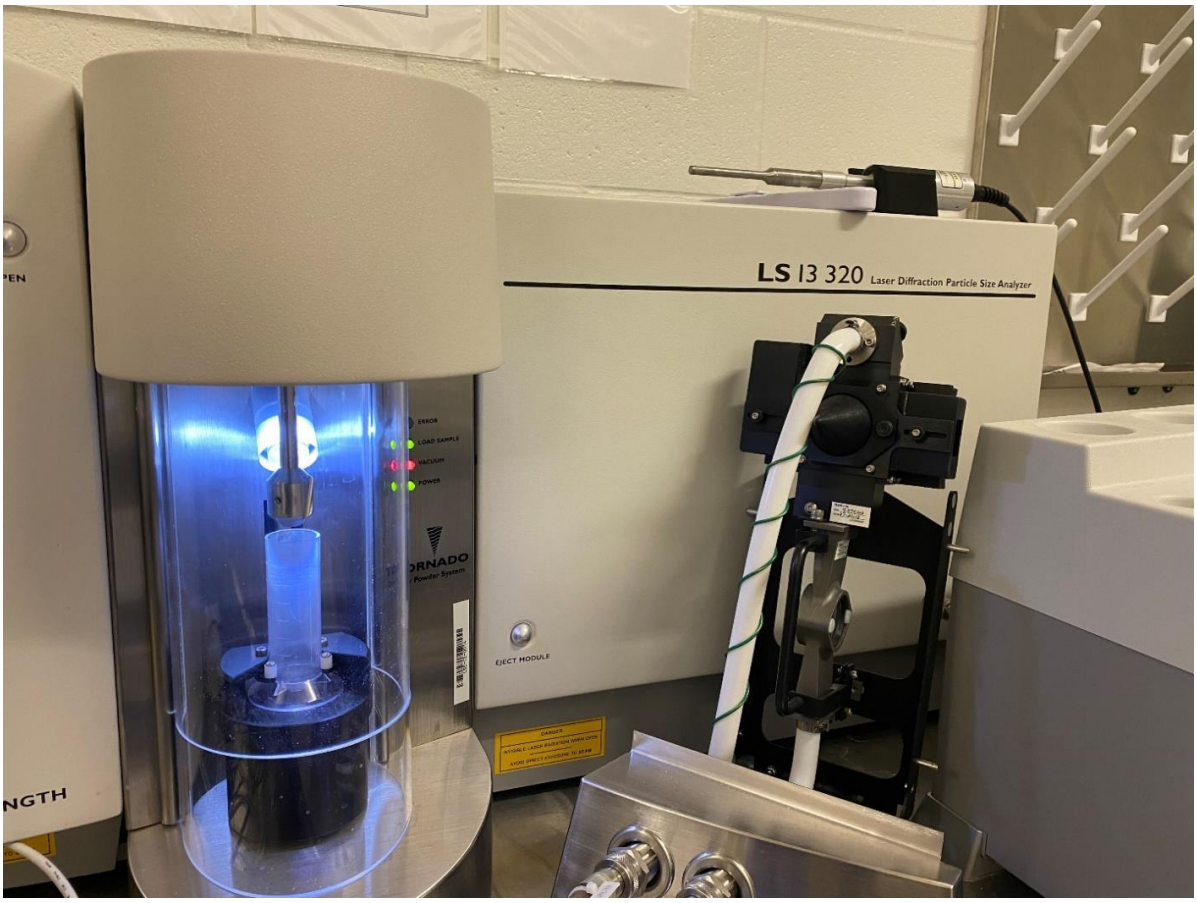


Figure 17: Laser Diffraction Particle Size Analyzer used for obtaining a sample's particle size distribution

The averages of the three samples were plotted along with the already obtained PSD of stormwater runoff sediment found in the GTA by Goncalves & Seters in 2012 (**Figure 18**). The mixture containing equal amounts of clay and limestone was seen to have many more particles above 60 µm. The other mixtures were similar in distribution with slight variations. As expected, as the amount of clay was increased, the distribution tended to shift towards the smaller sized particles. As seen near the 40 – 80 µm particles, the mixes with higher amounts of clay were slightly rounding their upper tail early. Although all mixes were generally within the boundaries found in Goncalves & Seters (2012), the mixture with a clay-limestone composition of 60-40 was chosen to be used as the clogging sediment for the lab experiment. It should also be noted that the clay-limestone mixtures were found to have a higher representation of very fine particles (less than 1 µm) compared to the results found by Goncalves & Seters (2012).

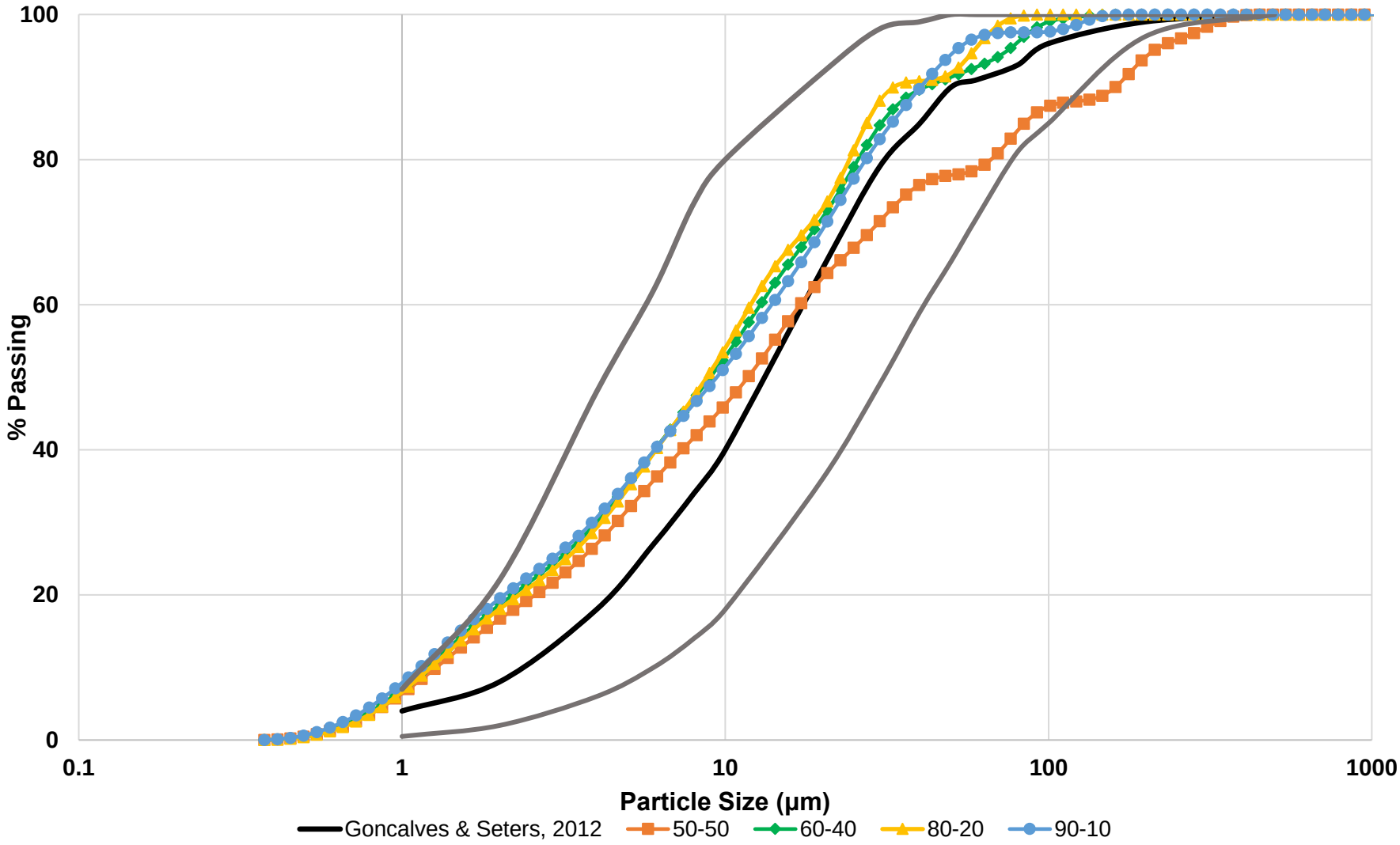


Figure 18: PSD of clay-limestone composition mixes compared to observed PSD of runoff sediment found in the GTA where the black line is the average and gray lines are the extreme observed PSDs found by Goncalves & Seters 2012

Once the clogging mixture was finalized, a total of 7.5 kg of clogging sediment was mixed consisting of 4.5 kg of Kaolin clay and 3 kg of the limestone that was passed through a 600 µm sieve and were mixed together. This mixture was then used for slowly clogging the lab-scale permeable concrete specimens. As calculated previously, a total of 40.72 g of sediment is needed to simulate one year of equivalent clogging for one slab. Three slabs underwent 58 years of simulated clogging while one slab acted as a control where no clogging was conducted. Thus, a total of 7 kg of the clay-limestone mixture was needed to clog three slabs to simulate 58 years of clogging. It should be noted that this is a simulated clogging procedure and that it is not representative of how long real world PCPC slabs would take to clog. Instead of years of clogging, the data was reported using test cycles. A dry application of sediment was used where the appropriate amount of sediment was sprinkled on top of the pavement and a hose on the 'mist' setting was used to simulate rain. This allowed the sediment to travel through the system, eventually blocking pores and reducing SIR.

An accelerated clogging schedule was needed due to time constraints in laboratory access. In total, 20 test cycles were conducted and sediment was applied at the end of each test cycle. The first four clogging cycles used 1 year of simulated clogging every cycle. Two years of simulated clogging was applied for the next three test cycles. All subsequent test cycles received an application of an equivalent of 4 years of clogging sediment. Literature found that PPs did clog much earlier than observed in this thesis. Razzaghmanesh & Beecham (2018) found that generally in literature the infiltration rate of permeable pavements drastically decreased from 1000 – 10 000 cm/hr to around 100 cm/hr after an age of 4 years. Schaefer & Kevern (2011) found that clogging occurred just after a few (around 4 years equivalent) clogging cycles for silty sand and clay sediments. Tong et al. in 2011 found that permeability of pervious concrete specimen reaches a clogged status depending on the clogging material. In general, this was found to be after 4 – 6 years however the presence of finer particles such as clay significantly reduced the time to clog to 2 – 4 years. It should be noted that the yearly clogging application in this experiment does not represent an accurate real-world pavement and should not be used to extrapolate the lifespan of these pavement systems. The clogging procedure was done as an exhaustive procedure to slowly degrade the hydraulic performance of the system.

CHAPTER 5: DATA COLLECTION AND MODELLING METHOD

5.1 Image Analysis

5.1.1 Greyscale Image Analysis

Images of the surface of the pavement were taken using an iPhone 11 Pro Max. The phone is equipped with a 12-megapixel telephoto lens with a 2.0-aperture and auto image stabilization. The images were captured in JPEG format by changing the camera capture option to 'Most Compatible' in 'Settings'. A 'Square' aspect ratio option was chosen in the 'Camera' application and the camera was positioned so that the surface being investigated aligned with the borders of the 'Square'. This resulted in holding the phone at a distance of about 0.5 m [1.6 ft.] above the surface. Images were collected with and without the use of flash. This resulted in two datasets that were tested and the impact of using flash was analyzed. MATLAB was used to convert the images to 8-bit grayscale. This was done by using the 'rgb2gray' function in MATLAB. Once converted to grayscale, image histograms were graphed by extracting the pixel count at each gray level (0 – 255). The pixel counts were normalized by the total number of pixels (resolution) of the image. Image properties including the mean, variance, and skewness were extracted from the images:

$$Mean = \frac{\sum GL * P}{N}$$

$$Variance = \frac{\sum[(GL - Mean)^2 * P]}{N - 1} \qquad \textbf{(Equation 10)}$$

$$Skew = \frac{\sum[(GL - Mean)^3 * P}{(N - 1) * \sigma^3}$$

GL = Gray level (0 – 255)

P = Number of pixels at each gray level

σ = Standard Deviation of the gray level

N = Total number of pixels in the image

5.1.2 Lighting Condition Analysis

In practice, the research could potentially allow a user to take a photograph of a pavement surface and determine the associated infiltration rate. There are many variables that need to be accounted for when dealing with real-world images. These include controlling for sources of error such as the presence of shadows, debris on the pavement, and taking good quality photographs (focus, angle of camera, etc.). The above-mentioned sources of error can be controlled by taking extra care prior to taking the photograph. However, there are still some issues that cannot be controlled for such as outdoor lighting conditions. The sun is responsible for illuminating the outside world. Due to Earth's rotation, lighting is not uniform throughout the day. During sunrise and sunset, there are hues of orange and dark blue prior to nightfall whereas during the day the sun radiates bright white over the surface of the earth. The variability is further exacerbated through different weather conditions. This variability in lighting condition may be problematic in terms of using outdoor image parameters as model inputs. The affects of outdoor lighting conditions have on pavement surfaces was investigated.

Five locations were chosen on an asphalt pavement and the different lighting conditions were tested for. Photographs were taken using an iPhone 11 Pro Max at different times of the day. Namely, at noon, sunrise, and sunset. Furthermore, the use of flash was investigated by taking photographs at noon with flash on as well as at night with flash on. The angle at which the camera was held was also investigated by taking the photograph while holding the cellphone at a slight angle at noon. Three images for each of the five locations were taken in each of the six different scenarios for a total of 90 images. To ensure consistent photographs of the pavement, a 350 mm by 350 mm wooden frame was constructed and lines were marked on the pavement to ensure the framing of the photograph was consistent between the photographs (see **Figure 19a**). A wooden plank was used with a box as a platform for the cellphone to sit on to ensure steady photographs were taken (see **Figure 19b**). Grayscale histograms and image properties were extracted from these images and Kolmogorov-Smirnov Distribution Tests (KSDT) were conducted to determine if the affects of the aforementioned variables had a statistical impact on the image.

(a) **(b)**

Figure 19: (a) Asphalt pavement and sliding frame to obtain consistent images of surface; (b) Wooden plank and box used as a stand for steady images

5.2 Laboratory Analysis

5.2.1 Surface Infiltration Rate Tests

SIR tests were conducted on the laboratory specimens using the standard ASTM C1701 method for determining the surface infiltration rate of in-place permeable concrete pavement. Although not required from the standard, a photograph of the test location was taken using a smartphone camera at around 0.5 m [1.5 feet] from the surface, ensuring that the smartphone is parallel to the ground and that there are no shadows in the image. This photograph was later used for image analysis.

A single ring 300 +/- 10 mm [12.0 +/- 0.5 in.] with a minimum height of 50 mm [2.0 in.] infiltrometer was procured by cutting a 19L Home Depot bucket. The inner surface of the ring was marked with two lines at 10 mm and 15 mm [0.40 and 0.60 in.] from the bottom of the ring. The infiltration ring was installed on the surface of the pavement by pressing the ring into the pavement and using plumber's putty on the outside edge of the ring to create a watertight seal with the pavement. **Figure 20** shows the ring after installation with a water tight seal that prevents leakage of water during the test. It is recommended that the test be performed when the surface of the pavement is less than 38°C [100 °F] which was not a concern for the laboratory specimen.

Figure 20: Installed infiltration ring

Prewetting was done as per ASTM C1701 where 3.60 +/- 0.05 kg [8.0 +/- 0.1 lb] of water was poured into the ring at a constant head such that the water level is maintained between the two marked lines in the ring. A stopwatch was used to measure the time from when the water initially contacts the surface until there is no longer any free water present on the surface. The time was recorded to the nearest 0.1 s.

The actual test was conducted within two minutes of prewetting. If the elapsed time during the prewetting step was less than 30 seconds, 18.0 +/- 0.25 kg [40.0 +/- 0.5 lb] of water was used. If the time for infiltrating the prewetting amount was greater than 30 seconds, then 3.60 +/- 0.05 kg [8.0 +/- 0.1 lb] of water was used. The water was poured at a sufficient rate to maintain a water level between the two marked lines. Time was recorded from when the water makes contact with the surface until there was no longer any free water present on the surface.

If a test was repeated, the repeat test did not undergo prewetting if conducted within 5 minutes of the first test as per ASTM C1701 (ASTM, 2017). If more than one test was conducted at the same location on the same day, the infiltration rate was the average of the two tests. The standard also recommends to not repeat this test more than twice at the same location on a given day as well as to not conduct any tests within 24 hours of the last precipitation, however this was omitted for the purpose of collecting data. The infiltration rate was then calculated with:

$$I = \frac{KM}{\emptyset_{ring}{}^2 t}$$ **((Equation 1)**

I = surface infiltration rate (mm/h [in./h])

M = mass of infiltrated water (kg [lb])

$\emptyset_{ring}$ = inside diameter of infiltration ring (mm [in])

t = time required for measured amount of water to infiltrate the surface (s)

K = 4,583,666,000 (mm^3·s)/(kg·h) or 126,870 (in.^{3}·s)/(lb·h)

Infiltration tests were conducted on the four specimens with minor adjustments to the aforementioned standard ASTM C1701. The standard specifies that no more than two SIR tests shall be conducted at the same location in 24 hours and that the SIR is the average of the two tests. For the purposes of data collection, three tests were conducted on each specimen per day and were reported individually instead of taking the average. SIR tests were done in the sequence of slabs meaning that SIR tests were conducted on each slab before repeat SIR tests were conducted. This allowed the water to drain through the systems while SIR tests were being conducted for the other slabs. For each level of clogging, six SIR tests were conducted on each slab.

Photographs of the laboratory specimens were taken prior to conducting surface infiltration tests. The laboratory specimens were placed on a waterproof tarp that was placed in a 0.60 m x 2.1 m [2 ft. x 7 ft.] wooden container. The waterproof container collected the water that drained from the specimens and a submersible pump was used to pump the water out of the container. **Figure 21** shows the setup. Throughout the clogging procedure, image and SIR data were continuously collected which resulted in a dataset that consisted of 472 images that represented images of freshly cast to fully clogged PCPC slabs where each image had its own associated SIR that.

Figure 21: PCPC Laboratory Specimen inside waterproof container with submersible pump

Since the SIR tests were done sequentially, it was more efficient to leave the infiltration ring on the slab rather than taking the ring off and reapplying it after taking another photograph. Thus, three photographs were taken of each slab prior to applying the infiltration ring at the beginning of each day. These photographs would then be matched with each measured SIR. The photographs were taken using an iPhone 11 Pro Max using the default camera application and the 'Square' option for the aspect ratio. The edges of the slab were aligned with the frame displayed on the phone screen. This yielded a distance of about 0.5 m [1.6 ft.] from the surface of the pavement to the phone. Extra caution was used to ensure the phone was horizontal prior to taking the photograph. Two sets of photos were taken; with flash and without. This resulted in two datasets which were analyzed to assess the impact of using flash.

5.3 Creating Empirical Models

The laboratory component of the thesis consisted of PCPC specimens being slowly subjected to clogging while SIR and image data were simultaneously obtained. This produced a dataset composed of images of the pavement surface and that surface's SIR rate. The images were converted to grayscale and image properties such as the mean, variance, and skewness were extracted. These image properties, along with SIR data, were used as inputs to artificial neural network models that can use image data inputs to predict a value of SIR.

5.3.1 Optimizing Artificial Neural Networks

Artificial Neural Networks are a type of empirical model which explore the relationship between input and output variables. Unlike linear regression models, ANNs have been shown to accurately capture the non-linearity between the input and output variables. Being able to predict non-linear relationships is a major benefit and one of the reasons ANNs are popular throughout a wide range of disciplines. ANNs, more specifically MLPs, have three distinct layers: an input layer, a hidden layer and an output layer. The input and the output layer are connected by the hidden layer through weights, biases, and activation functions which allow MLPs to capture the non-linearity between the input and output layers. Creating a MLP model requires training the MLP on a dataset and adjusting the weights and biases such that accurate predictions can be made. A single or multiple cost functions are used to determine whether the predictions are accurate.

To train MLPs, it is necessary to fine tune hyperparameters such as the number of hidden layers, training function, transfer function, number of epochs, and data split. As mentioned previously, the hidden layers connect the input and output layers. The number of hidden layers dictates how many neurons there will be in the network. The higher number of hidden layers, the more complex the network. However, a large number of hidden layers does not mean that the accuracy of the network improves. Keeping the network simple by having a lower number of hidden layers relieves computation effort and time and also prevents overfitting the model to the training data. Therefore, finding the optimal number of hidden layers is important for model accuracy and computational

optimization. The size of the hidden layer was iterated through by using one to 60 hidden layers. The optimal number of the hidden layer was found to be the one that provided the lowest cost function. In this case, the Root Mean Squared Error (RMSE) was used.

The data split dictates the ratios of the dataset that are used for training, validation, and testing. Having a large and diverse dataset is important for the robustness of the model. The training and validation data is used to train the model by updating the weights and biases. The final performance of the updated model is then evaluated by using the testing data. The sampling of the dataset is also important. The dataset was randomly divided and sampled in MATLAB. The training-validation-testing percentages were iterated through from 50-25-25 to 90-5-5 and the ratio combination that resulted in the lowest RMSE was used in the final optimized model.

Transfer functions, also known as activation functions, are a mathematical representation of the relationship between the input and output. These functions are responsible for generating outputs and are usually sigmoid shaped but are also sometimes seen as non-linear, piecewise linear, or step functions (Agatonovic-Kustrin & Beresford, 2000). These functions are also monotonic, continuous, and differentiable. The most common transfer function used is the Log-sigmoid transfer function which takes input values and transforms them into values of 0 or 1. Four transfer functions were tested with the dataset: log-sigmoidal (LOGSIG), hyperbolic-tangent-sigmoidal (TANSIG), pure-linear (PURELIN), and positive-linear (POSLIN). The optimal transfer function was used in the optimized network and the same transfer function was used before and after the hidden layer.

There are different training functions that are used depending on the application. Also known as an optimisation algorithm, these functions often use backpropagation in order to update the weights and biases of the network according to their algorithm by minimizing a cost function. Common cost functions include mean squared error, sum of squared error, and mean absolute error. Three training functions were tested for the dataset including Levenberg-Marquardt (TRAINLM), Scaled Conjugate Gradient (TRAINSCG), and Bayesian Regularization backpropagation (TRAINBR). The best performing training function was selected for the optimized network.

Selecting the number of epochs is the stopping criteria for the network and is responsible for preventing underfitting and overfitting. A single epoch represents a full pass of the training data through the network. With each pass of the dataset, the weights of the network are updated. If enough passes have not been done, the model will be undertrained and will not be able to make accurate predictions for the training dataset. If too many passes occur during training, the model will be overtrained and will make accurate predictions for the training dataset but will not be able to make accurate predictions for the testing dataset. The maximum number of epochs was set to 1000, however a maximum failure parameter was used which stops the learning process after a certain number of consecutive performance decline. The optimal number of how many performance declines that is needed to stop the training process was found by iterating this early-stopping criteria.

An ensemble approach for the ANN was used where the model was run 100 times to create a 'prediction band'. With each ensemble run, the weights and biases of the model is updated, essentially creating a new model. The data passes through each model in the ensemble and generates a single set of outputs. Once the dataset passes through each model within the ensemble, a total of 100 output vectors are obtained as well as 100 values for the performance functions. This approach allows for a prediction band instead of a single output vector and encapsulates the range of possible predictions. The mean of the performance functions was taken to obtain the performance of the ensemble.

ANNs with different inputs were selected for the analysis. These models used image parameters as inputs to predict the SIR of the slab in the image. Input parameters to these models included the grayscale mean, variance, and skewness and normalized variants of these parameters. These normalized parameters were found by dividing the grayscale images of the slabs that received clogging by the mean of the grayscale image of the control slab. This normalized image was then used to find the normalized mean, variance, and skewness. The models with normalized input variables did not use the data collected for the control slab. **Table 12** summarizes the different input variables used and **Table 13** shows the different combinations of the input variables for each of the four ANN models. ANN3 used raw image data similar to ANN1, however, image and SIR data from the control slab was removed. Doing so allowed ANNs 2 and 3 to be directly compared

since they then had the same dataset size with the only difference being ANN2 used normalized image data whereas ANN3 used raw image data. The use of normalized data in predictive models limits the model for real world applications, however it will give insight on the effect of this normalization technique.

Table 12: List of potential independent variables for the ANN model

Input Variables	**Description**
X1	Mean
X2	Variance
X3	Skewness
XN1	Normalized Mean
XN2	Normalized Variance
XN3	Normalized Skewness

Table 13: ANN Models with different combinations of independent variables

Model ID	**Independent Variables**
ANN1	X1, X2, X3
ANN2	XN1, XN2, XN3
ANN3	X1, X2, X3 (No control)

5.3.2 Performance of Artificial Neural Networks

To evaluate the performance of the different ANN models, three performance metrics were used. Namely, these included the coefficient of determination (R^2), the root-mean-square error (RMSE), and the percent capture. The R^2 was calculated by first finding the coefficient of correlation (R) between the predicted and the observed SIR for each model. In general, R^2 explains how much variation is captured and explained by the model. A larger value of R^2 indicates that the model captures more variability predicted data around its mean. With higher R^2 values, the model accounts for more variance and the closer the predicted data will be towards the fitted line. In conjunction to the R^2, the RMSE of each model was also calculated. The RMSE is a frequently used parameter in measuring the performance of predictive models. The RMSE was calculated by squaring the difference between observed and predicted SIR for each ensemble. The RMSE is a performance metric that incorporates the residuals of the model. In general, the lower the RMSE, the better the predictive model. The percent capture was the last parameter that was used to evaluate the performance of each model. The percent capture is defined as the percent of observed values that are captured within the maximum and minimum predicted value for each ensemble. This parameter gives insight on the robustness of the model. Each performance metric was found separately for the training, validation, and testing stages of the model as well as with the full dataset. With an ensemble approach, 100 ensembles were run for each ANN. Thus, the performance metrics were found for each ensemble and then averaged. In addition to reporting these performance functions, true and predicted plots were generated as well as graphing input variables with the predicted and observed SIR. The percent capture was calculated by counting the number of times the observed value fell within the maximum and minimum prediction for each observation for every ensemble. The R^2 and RMSE were calculated with the following equations:

$$R^2 = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2} \qquad \textbf{(Equation 11)}$$

$$RMSE = \sqrt{\frac{\sum(y_i - \hat{y}_i)^2}{N}} \quad \textbf{(Equation 12)}$$

y_i = ith SIR observation of the dataset

$\hat{y}_i$ = ith SIR prediction of each ensemble of the dataset

$\bar{y}$ = mean of the observed SIR

N = total number of observations in dataset

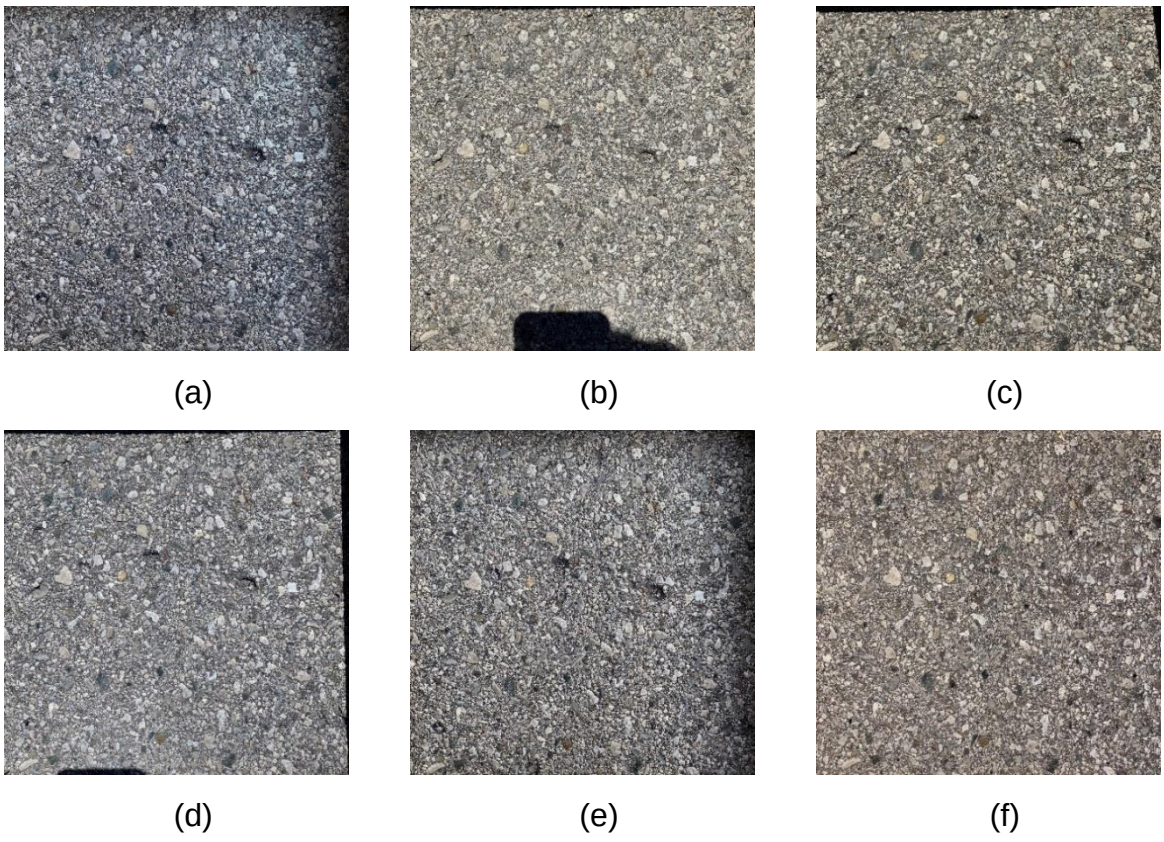
Finally, a brief sensitivity analysis was conducted for one model. Each input variable was given a range of values while holding the other input variables constant. For example, the mean was given a range of values while the variance and skewness were kept constant at the maximum or minimum value found in the dataset. Different combinations were used to understand their affect on the model predictions. The ranges of values chosen for each of the input variables was determined by choosing a range of values that was similar to the observed values for those input variables. The mean was selected to have a range of values from 100 to 140. The variance was selected to range from 3000 to 5000. The skewness was selected to range from -0.2 to 0.4. While each input variable varied, the remaining inputs remained constant at either minimum or maximum values which were obtained from the observed data. A sensitivity parameter was calculated for each input parameter and scenario with $S = \frac{\partial Y}{\partial P_i}$ where Y is the predicted SIR of the model and P is each input parameter (Jørgensen, 2001).

CHAPTER 6: RESULTS AND DISCUSSION

6.1 Image Analysis

6.1.1 Lighting Condition Analysis

The use of image data in real world applications is complicated due to the inconsistencies in lighting condition of outdoor images. To investigate the variability of outdoor images on pavement surfaces, images were taken of five different locations of an asphalt pavement during six different lighting conditions. Namely, these lighting conditions included images taken at sunrise, noon, sunset, as well as at noon with flash, at night with flash, and at noon with the camera held at an angle. **Figure 22** shows 'Location 1' with the six different lighting conditions. At first glance, there are obvious differences between some of the images, however, some images seem to be similar. Different locations and their histograms can be found in APPENDIX B.



(a) (b) (c)

(d) (e) (f)

Figure 22: Asphalt pavement 'Location 1' at (a) sunrise, (b) noon, (c) noon at an angle, (d) noon with flash, (e) sunrise, (f) night with flash.

Figure 23 shows the effect of different lighting conditions on a pavement surface. The overall shapes of the distributions are similar in nature; however, each lighting condition has its own distinct characteristics. For example, the 'Night with Flash' curve shows a much smoother curve with no spikes at the extreme grey pixel values. The images taken at 'Noon' and 'Noon with an Angle' show large spikes at the white pixel values. This could be due to high amounts of reflection of lighter colour aggregates at that time. The 'Sunrise' and 'Sunset' images show a shift towards the darker pixel values. This is likely due to less direct reflection of light and larger amounts of shadows causing an overall darker tinge to the images. Although these plots are visually easier to look at, to understand if there are any underlying similarities or differences between these curves, hypothesis tests to compare the distributions is necessary.

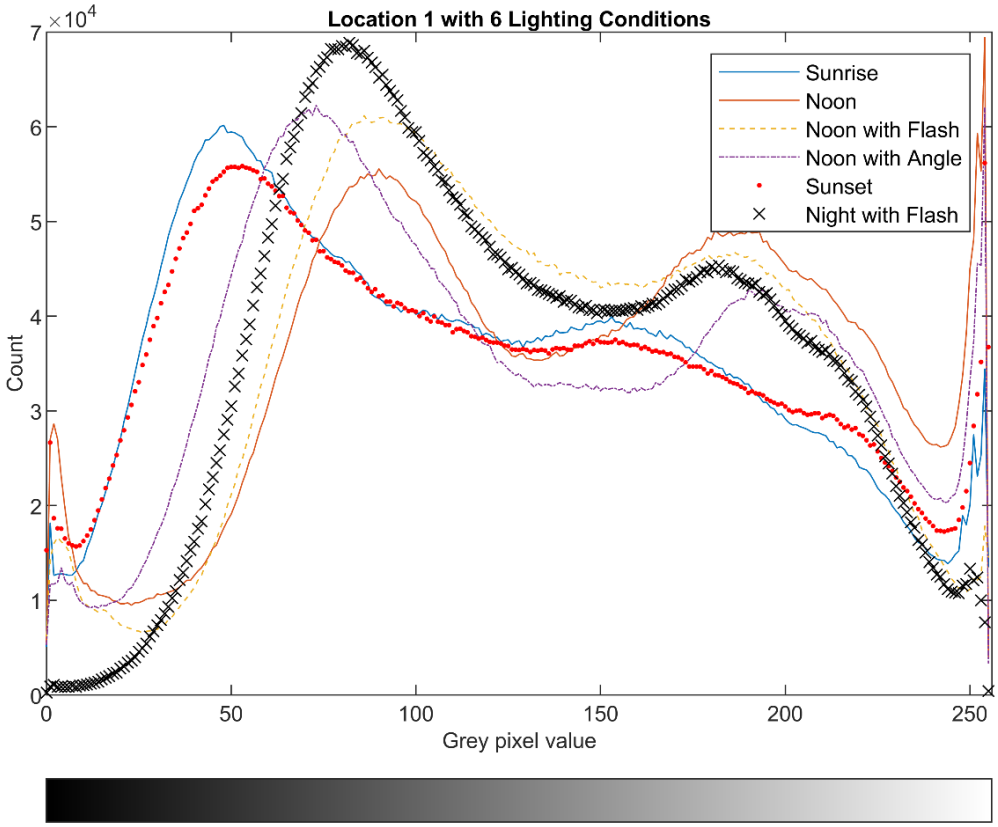


Figure 23: Image histograms of different lighting conditions for 'Location 1'

To test the differences between these images, a statistical test on the distribution of the image histograms was conducted. Three images were taken per location and condition resulting in a total image dataset of 90 images. These images were converted to grayscale and their image histograms were used to conduct a series of Kolmogorov-Smirnov Distribution Tests (KSDT) to investigate if the different locations and lighting conditions had a significant effect on their image histogram distributions. An alpha value of 0.05 was used for the KSDTs with the null hypothesis being the image histograms are from the same distribution and the alternative hypothesis being they are not. One lighting condition was first examined at the 5 different locations. **Table 14** compares image distributions from different locations during sunrise and the p-value results for KSDTs are shown. Green boxes indicate the null hypothesis was accepted, meaning the two images being compared have statistically the same distribution, whereas, red boxes indicate the null hypothesis was rejected. Additional tables and lighting conditions can be found in APPENDIX B.2.

Table 14: P-value results from KSDTs for different locations of pavement during Sunrise where green accepts the null hypothesis that the images are from the same distribution

Testing Location/Condition	IMG ID	P-Values from KSDTs for "Sunrise'				
Location 1 Sunrise	1	1				
Location 2 Sunrise	2	1.12E-05	1			
Location 3 Sunrise	3	0.0062	0.1999	1		
Location 4 Sunrise	4	0.3415	6.12E-04	0.0062	1	
Location 5 Sunrise	5	0.0258	0.0033	0.3415	0.003	1
		1	2	3	4	5

The results from **Table 14** and the tables in APPENDIX B.2 indicate that generally, the distribution of the image histograms do not change spatially. That is, when looking at different locations of a pavement surface with the same lighting condition, the images tend to follow similar distributions within the same lighting condition. To understand both spatial and temporal effects on the dataset, a 25% random sample of the image dataset was used to conduct 253 KSDTs to compare the distributions of images pertaining to different location and lighting conditions from the dataset. Each image from the random sample was compared with 22 other images.

The results from the KSDTs of the 25% random sample image dataset showed that the distributions of the images of the pavement surfaces were, in general, statistically different. 30 out of the 253 KSDTs (11.9%) showed that the image histograms followed the same distribution. Out of these 30, only 8 comparisons pertained to images of two different lighting conditions. 22 out of the 30 comparisons that accepted the null hypothesis were spatially different (different locations on the pavement), but had the same lighting condition, which suggests that images of asphalt are spatially similar. The remaining 223 KSDTs rejected the null hypothesis, meaning the histograms of the images being compared were significantly different. 12 out of those 223 KSDTs were comparisons of images with the same lighting condition. The remaining 211 KSDTs had combinations of different location and lighting conditions. In total, 88.2% of comparisons rejected the null hypothesis, meaning the image histograms being compared were from statistically different distributions. In summary, 11.9% of comparisons were statistically similar, however only 3.15% of comparisons were statistically similar with different lighting conditions.

Common image processing techniques were used to assess if processing of the images would increase the similarities between the images of different locations and lighting conditions. Processing techniques such as median filter noise removal, Gaussian filtering, and a white balance algorithm were used. The median filtering of an image creates an output pixel which contains the median value of a 3 by 3 neighbourhood around the corresponding pixel in the input image thereby removing noise. Gaussian filtering creates a blending effect and smooths the histograms of the images depending on a kernel value where the larger kernel value increases the amount of blending. The illuminate gray white balance algorithm was used which assumes the average colour of the image is gray and estimates the illumination. The processing techniques were evaluated both individually and together with different combinations by extracting and inspecting the output image histograms. A process of noise removal into white balancing and then Gaussian filtering with a kernel of 0.5 was used for the image dataset. Conducting the processing techniques in this order allowed the histograms to remain smooth and keep their overall shape.

This preprocessing procedure was used on the same images in the 25% random sample and KSDTs were conducted to evaluate if preprocessing of the images can remove the discrepancies between the images in terms of location and lighting condition. After preprocessing the images, 54 out of 253 (21.3%) of the comparisons showed the images followed a similar distribution which is up from 11.9% when no processing was done. 10 images out of those 54 were images with the same lighting condition and 44 out of those 54 images were from comparisons with different lighting conditions. 17.4% of comparisons showed a similar distribution where the lighting condition was different (up from 3.6% with no processing). Although the number of comparisons with similar distributions increased for those comparisons with different lighting conditions, the opposite was true for scenarios with the same lighting condition but different location where only 10 images out of 253 comparisons were statistically similar versus 22 before processing.

The analysis showed that pavement images acquired in the field can be statistically different depending on the location and lighting condition. In general, images taken with different locations but with the same lighting condition were more likely to be statistically similar. Even when processing the images, only 21% of the comparisons were similar. Comparisons of images with both spatial and lighting differences, which accounted for a total of 220 of the 253 KSDTs that were conducted, yielded 96.4% of comparisons being rejected. Comparing images with different lighting conditions yielded only 3.6% with similarities (17.4% after processing). This means that when collecting surface image data outdoors, the time of day the images were captured and general lighting conditions should be accounted for. Consistent lighting for an image dataset may be critical in maintaining the integrity of the results, especially if statistical parameters of the images are to be extracted for the analysis. Differences in lighting conditions between images may significantly impact the distributions of the histograms and thus have an effect on statistical parameters such as the grayscale mean, variance, and skewness. It should be noted that there may be other processing techniques which may improve the results and that processing of images may not even be necessary with sophisticated modelling techniques such as specifically trained ANNs and CNNs.

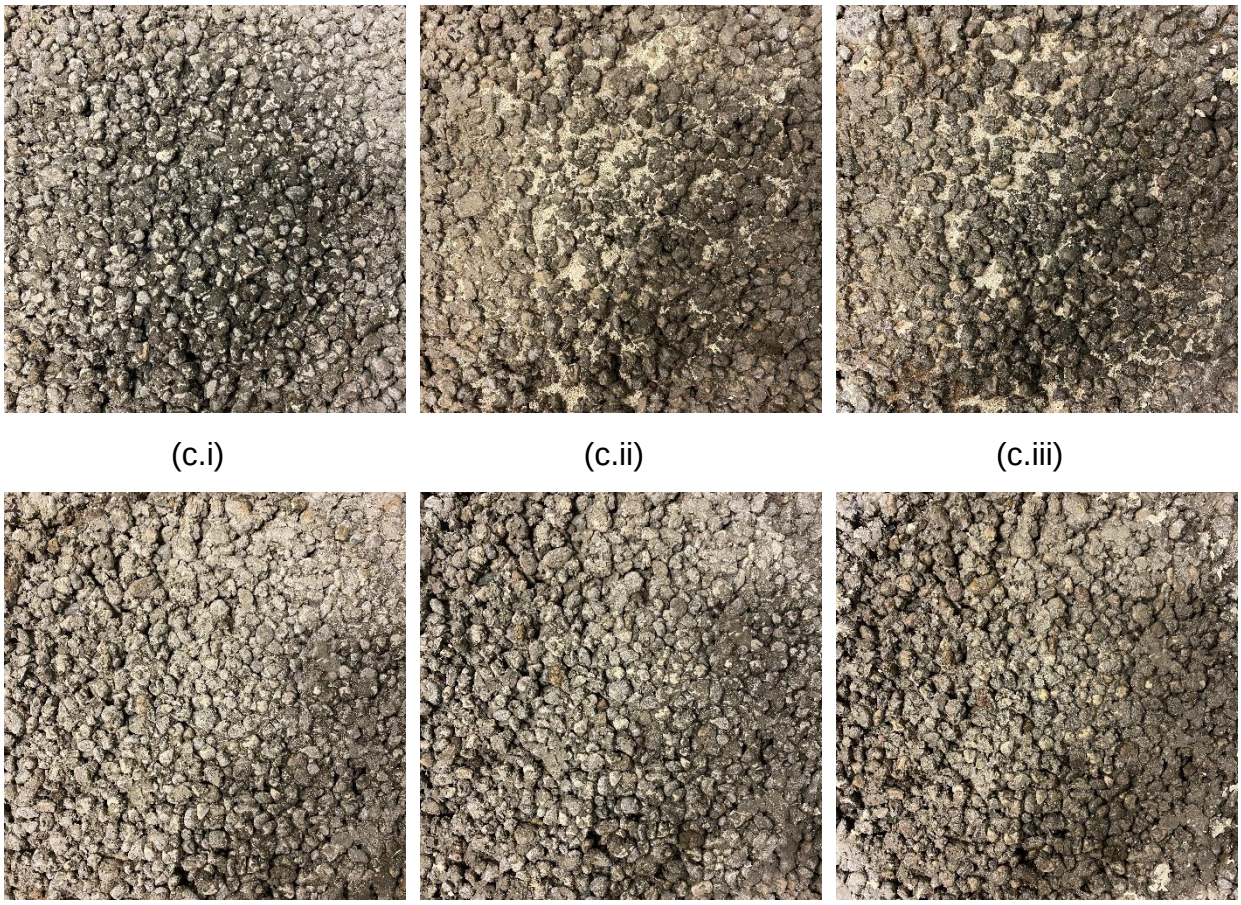
6.2 Laboratory Analysis

6.2.1 Laboratory Image Dataset and Analysis

SIR tests of four PCPC slabs were conducted over the course of 20 test cycles. Each test cycle consisted of 6 SIR tests per slab which was then followed up by clogging. Three of the four slabs were subject to clogging where the fourth slab acted as a control and received no clogging. **Figure 24** shows images collected for all four slabs as clogging progressed. It can be observed that the pores become clogged over time. For each SIR test, two images of the surface of the pavement were collected, one with flash and one without flash. Images were taken using an iPhone 11 Pro Max. In total, 944 images of the four slabs were collected where 472 images were taken for each flash scenario. This resulted in two image datasets – one with flash and one without flash.

(a.i) (a.ii) (a.iii)

(b.i) (b.ii) (b.iii)



(c.i) (c.ii) (c.iii)

(d.i) (d.ii) (d.iii)

Figure 24: Images of slabs as clogging progressed. Test cycle 1 (i), test cycle 10 (ii), and test cycle 20 (iii) for slabs 1 (a) 2 (b) 3 (c) and 4 (d) - 'Flash Off' dataset.

The images were converted to grayscale and statistical parameters including the mean, variance, and skewness were extracted. **Figure 25** – **Figure 27** show the results of the extracted mean, variance, and skewness for the flash off dataset. The 20 test cycles are seen on the x-axis and the lines represent the application of clogging sediment. Solid, dashed, and dotted lines represented 1 year, 2 years, and 4 years of equivalent clogging. An initial visual inspection of the results from the 'Flash Off' dataset revealed that the variance and skewness of the images seemed to decrease and increase respectively as clogging progressed while the control remained steady. It is hard to visually detect any significant trend of the mean but the control seems to decrease which suggests that there were some variables unaccounted for while obtaining image data. The results of the remaining image datasets can be found in APPENDIX C.

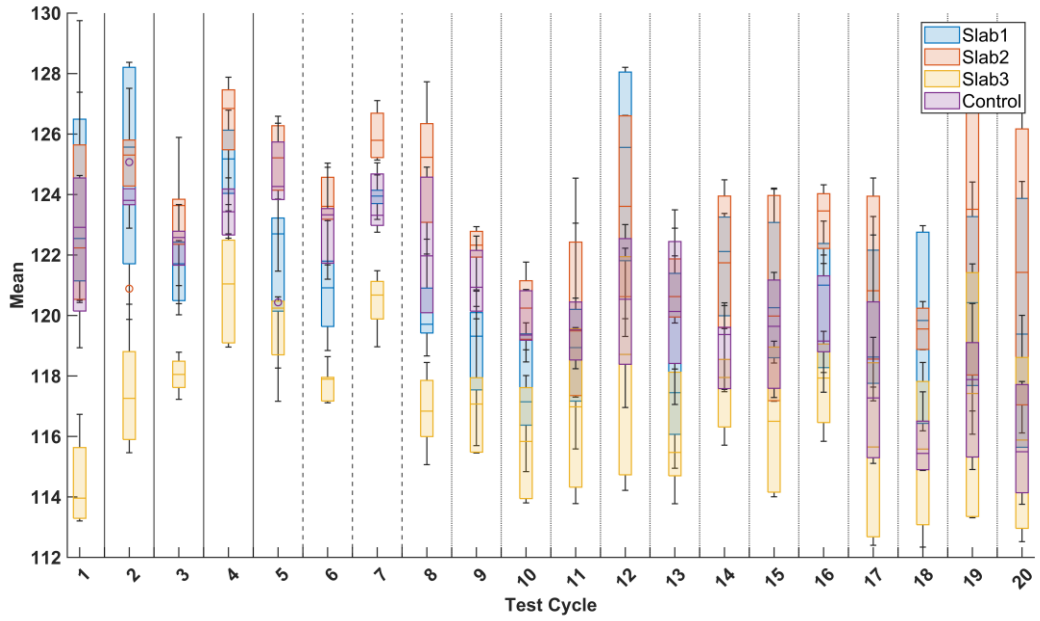


Figure 25: Mean of the images for PCPC slabs as clogging progressed (Flash Off Dataset)

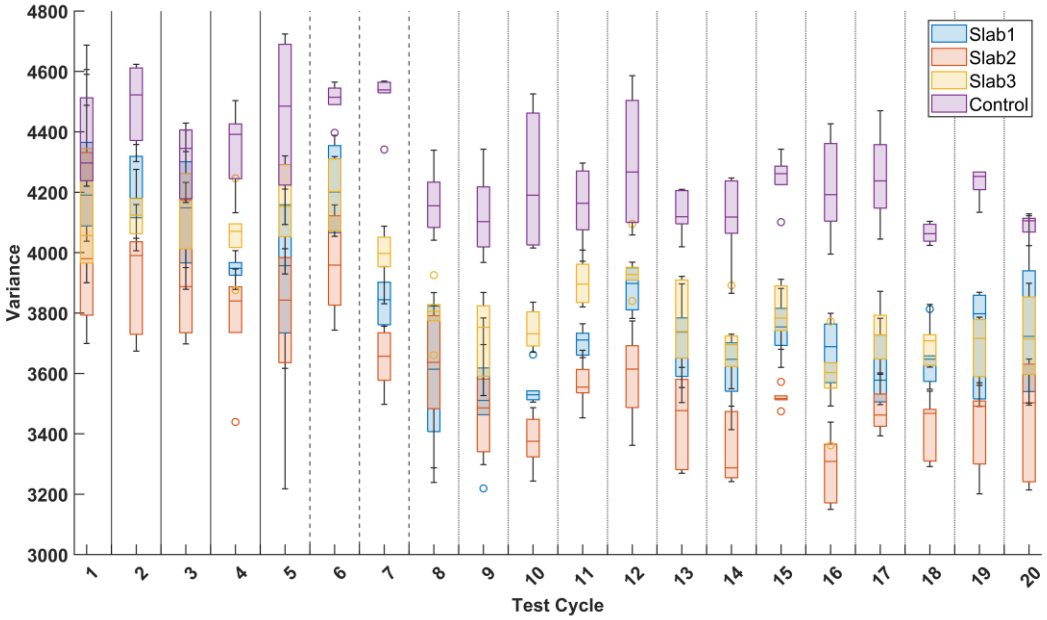


Figure 26: Variance of images for PCPC slabs as clogging progressed (Flash Off Dataset)

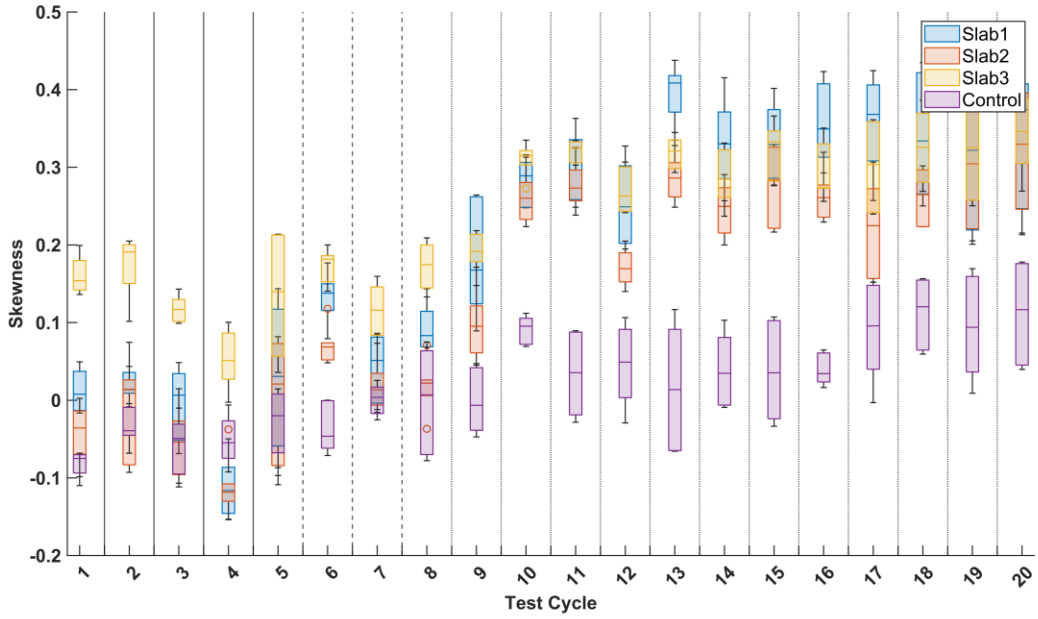


Figure 27: Skewness of images for PCPC slabs as clogging progressed (Flash Off Dataset)

From initial inspection, it was seen that the variance and skewness for the clogged slabs seemed to have a significant trend when compared to the control for the flash off dataset. To confirm these suspected trends, a Man-Kendall trend test was conducted for each extracted parameter for the 'Flash Off' and 'Flash On' datasets. **Table 15** to **Table 18** show the p-value and magnitude of the trend test where a p-value less than 0.05 shows a significant trend and magnitude is represented by Sen's slope. The trend test for the 'Flash Off' dataset revealed that all of the extracted image parameters have significant trends for all four slabs. The control slab exhibiting trends was unexpected and indicates that there were potentially some variables that were unaccounted for.

The underlined values in the tables indicate that there was no significant trend for that parameter. It can be seen that all parameters, even for the control slab, were seen to have significant trends in the 'Flash Off' dataset. Although no changes were made to the control slab, differences of the pavement surface overtime may have resulted in these significant trends. These may include curing of the concrete, wet-dry cycles from conducting SIR tests and putty residue left behind after SIR tests. The lighting condition was kept consistent by keeping the pavement systems indoors and in the same location

throughout the testing period, however, discrepancies from capturing images such as framing of the photograph. The 'Flash On' dataset did not have a significant trend for the variance in Slab 1 and Slab 2. This suggests that using flash decreases the overall variance of images and using parameters from the 'Flash On' dataset as inputs to a model may decrease its overall accuracy.

Table 15: Trend test p-value results of image parameters for the 'Flash Off' dataset (p-value =< 0.05 is significant)

SLAB	Mean	Variance	Skewness
1	1.88e-04	1.5e-09	2.29e-22
2	5.15e-06	3.68e-14	1.03e-21
3	0.017	5.41e-17	3.00e-19
Control	2.72e-20	6.74e-07	8.5e-16

Table 16: Trend test Sen's slope results of image parameters for the 'Flash Off' dataset

SLAB	Mean	Variance	Skewness
1	-3.73	-6.04	9.73
2	-4.56	-7.57	9.57
3	-2.39	-8.38	8.97
Control	-9.23	-4.97	8.05

Table 17: Trend test p-value results of image parameters for the 'Flash On' dataset (p-value =< 0.05 is significant)

SLAB	Mean	Variance	Skewness
1	1.22e-09	0.40	4.35e-23
2	4.81e-09	0.70	2.87e-21
3	8.22e-13	2.23e-06	8.65e-14
Control	0.021	0.0081	1.91e-04

Table 18: Trend test Sen's slope results of image parameters for the 'Flash On' dataset

SLAB	Mean	Variance	Skewness
1	-6.08	0.84	9.90
2	-5.85	0.39	9.47
3	-7.16	4.73	7.46
Control	-2.31	-2.65	3.73

Due to the control exhibiting significant trends with increasing test cycles, preprocessing of the images was explored where the image dataset was normalized to by removing the trends seen in the control. To do this, each image was converted from RGB to grayscale. The grayscale image was then converted to an array. Each image was then divided by the mean grayscale control image for each of its respective test cycle. This centered the trends of the control image parameters while simultaneously removing the causes of these trends from the images of the other slabs. This resulted in a normalized image array which was then used to calculate the normalized versions of the image parameters (mean, variance, and skewness). Due to the inadequate trend test results of the 'Flash On' dataset, the normalization process was only done to the 'Flash Off' dataset. **Table 19** shows the results of the trend test analysis for the 'Normalized Flash Off' dataset. The results show that the control had no significant trend for all three

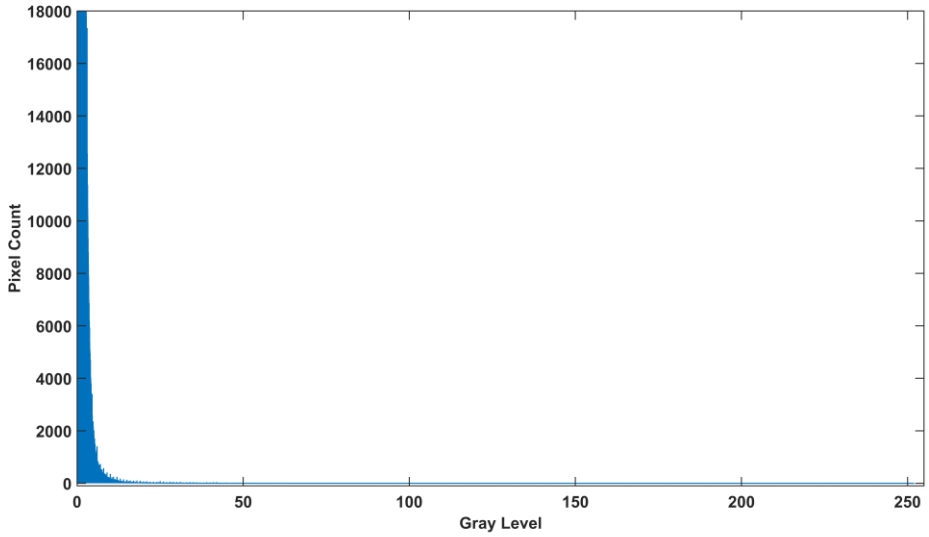
parameters meaning that the normalization technique succeeded in removing the trends seen in the control in the 'Flash Off' dataset. It should be noted that the normalized skewness for slab 1 had no significant trend. Moreover, the strength of the trends themselves decreased substantially for most parameters, which suggests that there were variables in the image set that were not controlled for that affected the image of all of the slabs. The normalization procedure needed further investigation on how it affected the images and their histograms.

Table 19: Trend test p-value results of image parameters for the 'Normalized Flash Off' dataset (p-value =< 0.05 is significant)

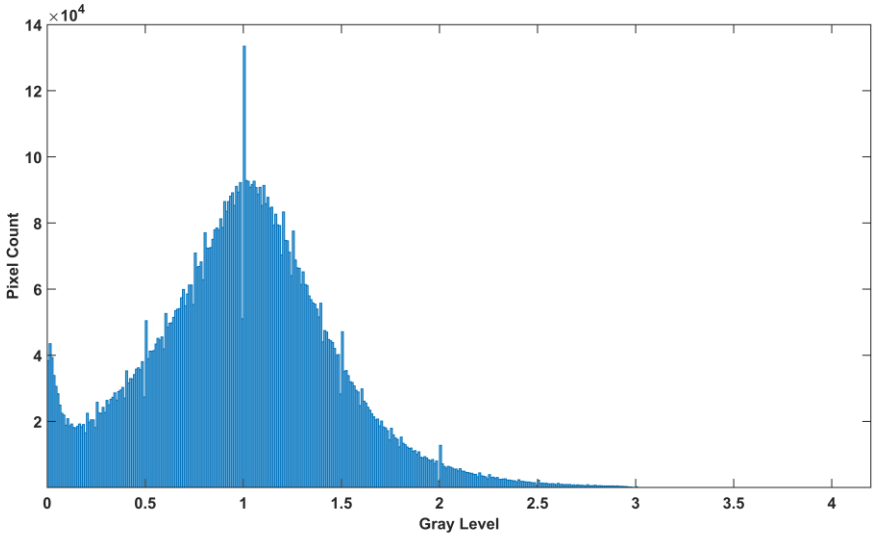
SLAB	Normalized MGL	Normalized VARIANCE	Normalized SKEWNESS
1	6.5e-09	0.002	0.908
2	5.17e-10	0.002	0.003
3	1.72e-11	0.005	0.032
Control	0.659	0.057	0.725

To understand this normalization procedure and its effect on the image, the histograms of the normalized slabs were generated for visual inspection. These histograms revealed that this normalization technique was not working as intended. Dividing Slabs 1, 2, and 3 with the mean of the control slab resulted in highly positively skewed histograms where a large amount of data points was present in the 0 – 1 range but there were some values in the histogram with values from 1 – 255. This was happening due to the presence of whiter pixels (pixels with gray level > 100) being divided by darker pixels (pixels with gray level <100). Since this combination was not common when conducting the image array division, the result was a very highly skewed histogram for slabs 1, 2, and 3 (see **Figure 28(a)**). Conversely, when dividing a control image with the mean image of the control images in a test cycle, there were no combinations in the image division that resulted in high gray levels (see **Figure 28(b)**). This means that the normalization technique that was used where each image was divided by a mean control

image did not work as intended as it transformed the normalized histograms in a manner that removed insight from slabs 1, 2 and 3. Using the mean, variance, and skewness from histograms that are similar to **Figure 28(a)** would not be helpful as inputs to a predictive model and a new normalization method was explored.



(a)



(b)

Figure 28: Normalized histograms where (a) image of Slab 1 and (b) image of control was divided by the mean of a set of Control Images

The previous normalization method consisted of an image array division where each image was divided by the mean control image of each test cycle. An updated normalization method was tested where each image was divided by the mean of the gray level of the mean control image of each test cycle instead of the full image. The original normalization method did somewhat succeed in the objective of removing the trends in the data, however, further investigation of the image histograms revealed that affect of the method on the image histogram was not accurate. This updated normalization method should remove the trends of the control seen in the mean of the grayscale image without impacting the shape of the original histogram. Once the normalized image was calculated, the variance and skewness of the histograms were then determined. The results of this updated normalized method can be found in **APPENDIX C: IMAGE AND LABORATORY ANALYSIS** The boxplots found in the appendix show the mean of the control images centered at 1, whereas slabs 1, 2, and 3 show the mean of the images increasing as the test cycles progress.

Table 20 shows the trend test results for the updated normalized parameters for the four slabs and **Table 19** shows the Sen's slope which indicates the magnitude of the trend. The trend in the mean for the control slab was successfully removed and was shown to increase as test cycles progressed. Unexpectedly, the trends for the variance were completely removed for slabs 1 – 3. Additionally, this updated normalization method had no effect on the trends seen in the skewness which is expected. Although the updated normalized mean of the control slab showed no significant trend, the control also had significant trends for both variance and skewness. Although this may not be ideal and that the trends for variance is not significant, the difference between the variance and skewness of the three slabs versus the control may still be significantly different and still worth using as inputs to models.

Table 20: Trend Test p-value results for the Updated Normalized Mean, Variance, and Skewness (p-value =< 0.05 is significant)

SLAB	Updated Normalized Mean	Updated Normalized Variance	Updated Normalized Skewness
1	1.14e-07	0.71	5.23e-22
2	4.30e-08	0.33	3.41e-21
3	4.09e-14	0.92	3.42e-19
Control	0.72	1.05e-07	3.3e-15

Table 21: Trend Test Sen's slope results for the Updated Normalized Mean, Variance, and Skewness

SLAB	Updated Normalized Mean	Updated Normalized Variance	Updated Normalized Skewness
1	5.30	0.37	9.64
2	5.45	-0.98	9.45
3	7.56	-0.10	8.95
Control	-0.36	5.32	7.88

The raw (not normalized) image parameters as well as the updated normalized image parameters were used as inputs to ANN models. The updated normalized parameters will now be known as the normalized parameters. **Table *22*** summarizes the Sen's slope for the raw image parameters and normalized image parameters for the four slabs. The normalization procedure not only removed the trends observed in the control but increased the magnitude of the trend and reverse the direction from a negative trend

to a positive one. It was initially hypothesized that adding depositing sediment on the surface of the slab would create an overall whiter image in grayscale due to dark pores being replaced by lighter coloured sediment. It seems that normalizing the image by removing the trends from the control does exhibit this behaviour. Although promising, using a normalized dataset as inputs to a predictive model would mean that all future predictions made by that model would require normalized data. This is not ideal from an application standpoint as field data would need to be normalized with laboratory control data. Although not ideal, it may still be worth using both the raw and normalized parameters in separate ANNs to evaluate their predictive capabilities and then evaluate their feasibility in real world applications.

Table 22: Summary table of Sen's slope for raw and normalized parameters

Parameter	**Slab 1**	**Slab 2**	**Slab 3**	**Control**
Mean	-3.73	-4.56	-2.39	-9.23
Variance	-6.04	-7.57	-8.38	-4.97
Skewness	9.73	9.57	8.97	8.05
Normalized Mean	5.3	5.45	7.56	-0.36
Normalized Variance	0.37	-0.98	-0.1	5.32
Normalized Skewness	9.64	9.45	8.95	7.88

It was observed that initial images of the slabs exhibited some locations with bright and white coloured aggregates. It is suspected that this occurred due to the curing process of the slabs and that these locations may have caused unexpectedly brighter images in the earlier testing cycles. The slabs were covered in wet burlap for 28 days for curing. **Figure *29*** shows an image of slab 1 in test cycle 1. It can be seen that the centre of the slab has dark aggregates due to the area being saturated with water. The edges of the slab can be seen to be much brighter than the centre. This difference in colour

between the edges and centre of the slab may be due to the slab drying unevenly after the curing process or even a consequence of the curing process itself.

Figure 29: Slab 1 Image from Test Cycle 1

To investigate the effect of the images in the earlier test cycles, a separate trend analysis was conducted where the first 5 test cycles were removed and only the last 15 test cycles were considered. Namely, test cycles 6 – 20 were used and a trend analysis was conducted for the raw and normalized data and the magnitude of the trends compared to the full test cycle trends. Differences in the magnitude of the trends would reveal the impact the first 5 test cycles to the overall dataset. **Table 23** and **Table *24*** shows the results of the trend analysis for test cycles 6 – 20. Comparing the magnitude of the trends found in **Table 22** with the results of the condensed test cycles in **Table *24*** one can see a general decrease in the magnitude of the trend for the raw image parameters. This suggests that the effect of the suspected discolouration of the aggregates and the curing process did not affect the raw image parameters in a way that

improves the magnitude of the trends seen in the full dataset. The first 5 test cycles provided the dataset with significant trends for the raw image parameters and the full test cycle dataset would be better as inputs to data-driven models.

The normalized mean and variance showed increases in the magnitude of the trend for test cycles 6 – 20 but a decrease in the skewness. This suggests that after normalizing for the control, some image parameters had a stronger trend when looking at the last 15 test cycles rather than the full dataset and that the first 5 test cycles did distort the strength of the trend. This was contrary to what was found with the raw image parameters which showed a general decrease in the trend with test cycles 6 – 20. This suggests that there may have been issues with the images of the first 5 test cycles which is enhanced when normalizing for the control. Using data from the last 15 test cycles for the normalized dataset may be beneficial as inputs to a data-driven model, however, removing 5 test cycles worth of data may result in less accurate models due to fewer data points. The effect that removing the first 5 test cycles had on both the raw and normalized data suggests that there should be extra care taken when collecting such images. It is recommended that images be taken with consideration to as many variables as possible and perhaps surface aggregates should be saturated with water to remove any white aggregates caused by the cement drying.

Table 23: Trend analysis p-values for raw and normalized image parameters for four slabs. Test Cycles 6 – 20. (p-value =< 0.05 is significant)

Parameter	Slab 1	Slab 2	Slab 3	Control
Mean	0.33	0.0022	0.047	3.9e-14
Variance	0.11	1.51e-05	1.64e-07	0.0028
Skewness	7.66e-11	2.31e-10	2.86e-11	5.01e-07
Normalized Mean	1.19e-09	1.62e-06	6.36e-11	0.90
Normalized Variance	0.0038	0.19	0.15	1.27e-06
Normalized Skewness	7.66e-11	2.31e-10	2.86e-11	5.01e-07

Table 24: Trend analysis Sen's slope for raw and normalized image parameters for four slabs (Test Cycles 6 - 20)

Parameter	Slab 1	Slab 2	Slab 3	Control
Mean	-0.98	-3.06	-1.98	-7.56
Variance	-1.61	-4.33	-5.24	-2.99
Skewness	6.51	6.34	6.65	5.03
Normalized Mean	6.08	4.79	6.53	-0.13
Normalized Variance	2.90	1.30	1.46	4.84
Normalized Skewness	6.51	6.34	6.65	5.03

The use of images and their parameters as inputs to data-driven models is complicated due to the number of variables needed to be controlled for. Normalizing the images was seen to successfully remove the trends in the mean of the images while maintaining information in the variance and skewness. It was observed that the first 5 test cycles did affect the magnitude of the trends of the dataset which suggests that there may be residual curing affects that impacted the early images of the slabs. When photographs are taken, it is recommended that all images be taken at the same condition. This means, that images for the entire dataset should be taken while the slabs are either completely dry or when the surface is wet after applying some amount of water to saturate the cement and aggregate. Image parameters, both raw and normalized, were seen to have significant trends over the testing period and may be suitable as inputs to ANNs to predict a slabs SIR.

6.2.2 Laboratory SIR Analysis

Four PCPC slabs were constructed with the intention of conducting SIR tests. Three of these slabs were subjected to clogging while the fourth slab received no clogging and acted as a control. Each test cycle consisted of 6 SIR tests and 1 round of clogging. ASTM C1701 was followed for the SIR tests where water was passed through an infiltration ring at a constant head and the time for a certain amount of water to infiltrate the surface was recorded. The amount of water used is dictated by the 'prewetting time' which is the time for 3.6 kg of water to infiltrate the surface of the pavement. If the prewetting time was less than 30 seconds, 18 kg of water was used for the SIR test. If the prewetting time was greater than 30 seconds, 3.6 kg of water was used for the SIR test. The prewetting was done prior to each SIR test. After 6 SIR tests were conducted per slab, a clogging sediment was loaded onto three of the slabs using a dry deposition where the sediment was evenly sprinkled on the surface of the pavements. The clogging sediment was composed of 60% kaolin clay and 40% crushed limestone sand that was sieved at 600 µm. This testing cycle was repeated for twenty test cycles. Both prewet and SIR times were recorded for each test cycle. **Figure 30** shows the results of the prewet times of the four PCPC slabs and **Figure 31** shows the results of the calculated SIR using the infiltration time. The three different lines represent different amounts of clogging sediment that were used. For the solid lines from test cycles 1 – 5, an equivalent of one year of clogging was deposited on the slabs per test cycle. For the dashed lines from test cycle 5 – 8, an equivalent of two years of clogging sediment was used per test cycle. The dotted lines from test cycles 8 – 20 represents 4 years of clogging sediment per test cycle. A table with the results can be found in APPENDIX C.4.

From the figures, we can visually see that the slabs that received clogging show a significant increase in prewet times and decrease in SIR as the test cycles progressed. The control slab stayed relatively constant throughout the test cycles. The four slabs had an initial SIR of around 35,000 to 55,000 mm/hr. Although the slabs were cast with the same mix design, on the same day with the same batch, and with the same casting procedure, variabilities during the construction and compaction procedure may have caused the differences in the initial SIR. Due to all four slabs being constructed from the same batch, the concrete began drying out by the time the third and fourth slabs were

constructed. This required additional effort in compacting the concrete and may have resulted in some slabs experiencing more compaction than others. This highlights the difficulties in constructing PCPC pavement systems and the importance in keeping the construction procedure consistent. Although their initial SIRs were different, a minimum infiltration rate of 35,000 mm/hr is still high when compared to some in-service permeable concrete and asphalt systems. Bean et al. (2007) found in-service PCPC systems to range from 130 mm/hr to 40,000 mm/hr, suggesting that the starting infiltration rates of these pavements is typical and adequate.

Table *25* shows the percent decrease in SIR that each slab experienced, comparing the average of test cycle 1 to the average of test cycle 20. The results indicate that all slabs experienced a decrease in SIR and the slabs that received the clogging sediment experienced a larger decrease in SIR. The control exhibiting a decrease in SIR was not expected but it is thought to be associated with the general sources of error when conducting the SIR tests as well as changes in personnel conducting the tests.

Table 25: Percent decrease in SIR for four PCPC slabs

Slab	**% Decrease in SIR**
1	79%
2	91%
3	72%
Control	14%

A Mann-Kendall trend test was conducted to determine if there were significant trends of the prewet time and SIR as clogging occurred for the different slabs. **Table *26*** and **Table *27*** shows the results of the trend tests for the prewet time and SIR as the test cycles progressed for the four slabs. From the tables, we can see that there are significant trends for all of the slabs for both prewet time and SIR. Unexpectedly, the control was seen to experience trends for both prewet time and SIR as test cycles progressed. It is suspected that these trends are occurring due the slight changes in prewet time and SIR found in the control after test cycle 15. At this test cycle, a change in lab assistants took place who were responsible for adhering the infiltration rings and operating the stopwatch.

These results show that there are variabilities in the testing procedure depending on the personnel conducting the test. These include different techniques in sealing the infiltration ring and differences in when the water is considered ‘infiltrated’ to stop the timing of the test.

Table 26: Trend test p-value results for prewet time and SIR for the four slabs (p-value =< 0.05 is significant)

SLAB	PREWET TIME	SIR
1	2.16e-37	1.87e-50
2	7.18e-49	6.77e-53
3	2.21e-40	1.45e-47
Control	2.32e-21	2.24e-20

To confirm if the significant trend for the control is associated with having different personnel assisting in conducting the test, another trend test was conducted for the control slab where test cycles 1 and 15 as well as 16 and 20 were compared. Conducting a trend test for the prewet times of the control slab for test cycles 1 and 15 reveal there is no significant trend with a p-value of 0.36. Similarly, there is no significant trend for test cycles 16 and 20 with a p-value of 0.29. Comparing the control SIR of test cycle 1 to test cycle 15 shows that the trend is not significant with a p-value of 0.11. Likewise, the control SIR for test cycle 16 and test cycle 20 shows no significant trend with a p-value of 0.94. This confirms that for the test cycles where the person assisting in conducting the SIR tests was the same, there were no significant differences in prewet time or SIR for the control slab and the trends found in the control indicate the variability between the individuals conducting the test.

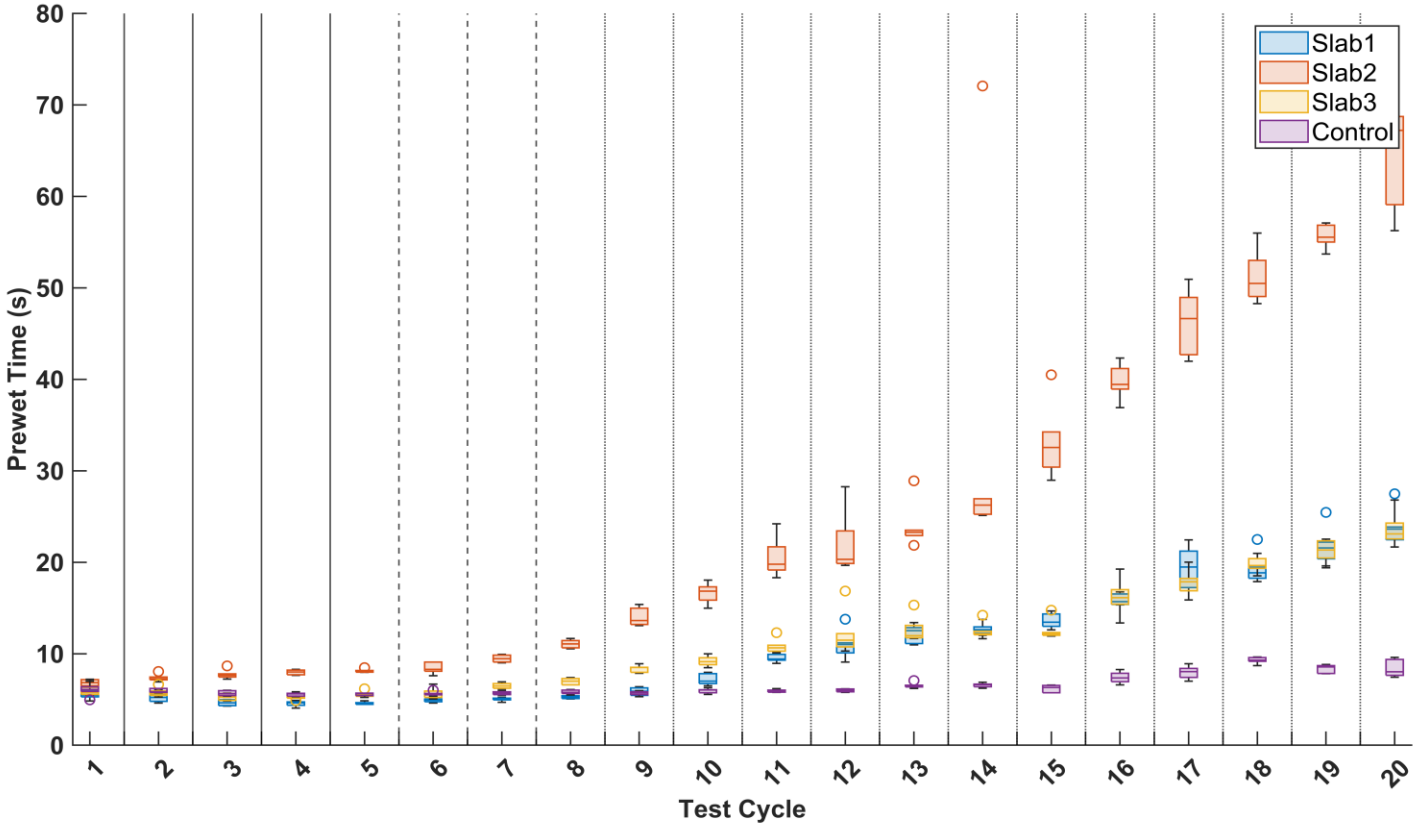


Figure 30: Prewet times of four PCPC slabs where the control slab received no clogging

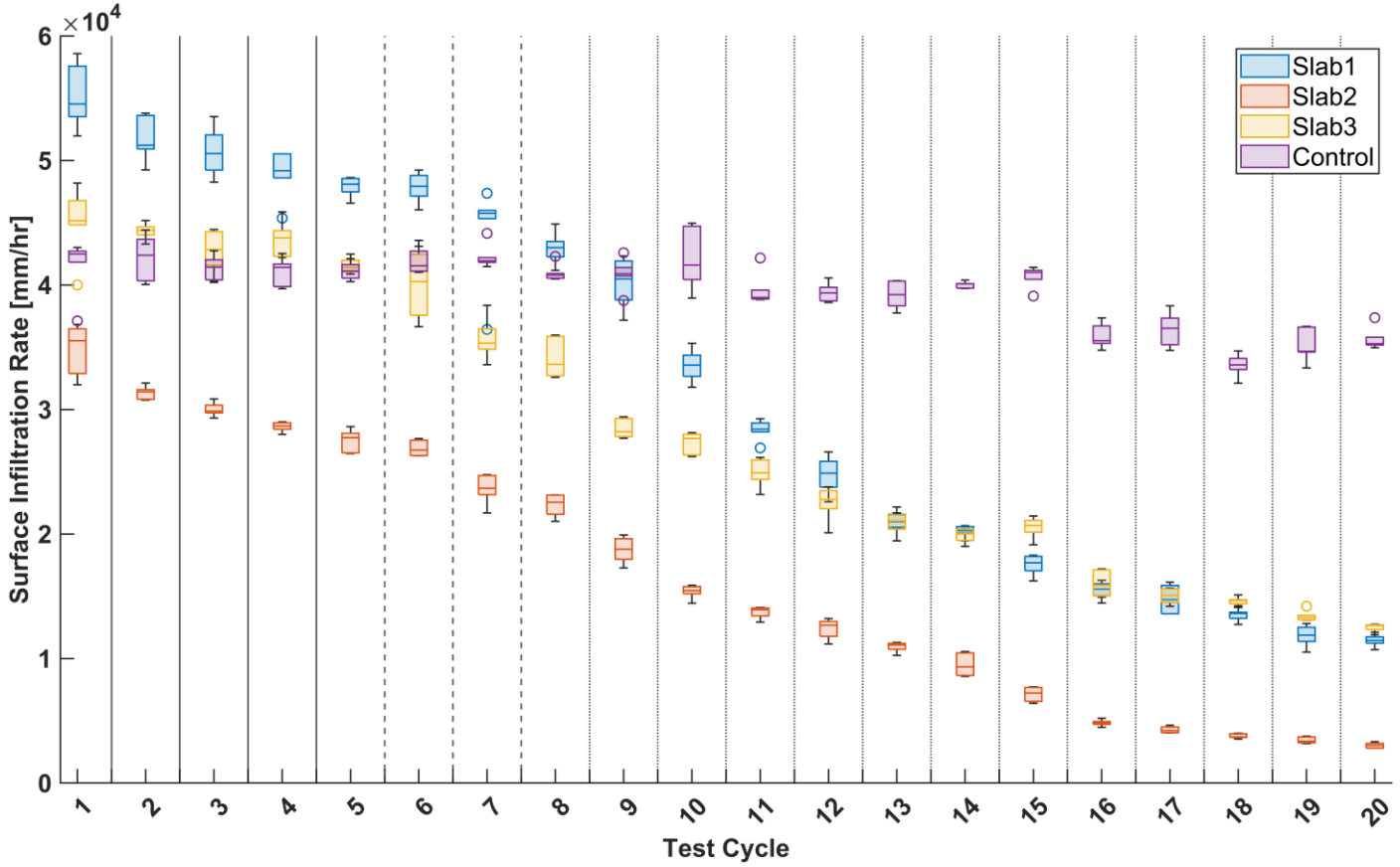


Figure 31: SIR of four PCPC slabs where the control slab received no clogging

Table 27: Trend test Sen's Slope results for prewet time and SIR for the four slabs

SLAB	PREWET TIME	SIR
1	12.78	-14.94
2	14.69	-15.31
3	13.30	-14.49
Control	9.49	-9.25

The clogging mechanism and procedure that was used successfully decreased the SIR of the pavements over time. The sediment used to clog these slabs was composed of fine particles including clay and fine sand/silt. Although it was found that the use of fine particle sizes had negligible permeability reduction (Tong et al., 2011), the results from this experiment reveal that sediment composed of fine particles do significantly reduce the SIR of PC slabs over time. After twenty cycles, the experiment was stopped to leave adequate time for the analysis. Slab 2 ended with an SIR of about 3000 mm/hr after twenty test cycles and slabs 1 and 3 converged to around 12,000 mm/hr. The rate of clogging and the shape of the overall curves for all three slabs look similar. It is expected that if the test cycles were to have continued, all three slabs would have converged to a similar ‘fully clogged’ infiltration rate. This infiltration rate would likely be around 0 – 100 mm/hr which would be similar to the SIR found in the preliminary field investigation (see APPENDIX A). Although the pavements did not reach this eventual end condition, the significant trends seen in the image and SIR data may be sufficient in using them in data-driven models.

6.3 Artificial Neural Network Results

6.3.1 Models and Input Variables

A stepwise approach was used where analysis of different combinations of inputs were used in multiple ANNs. In total, three main ANNs were evaluated, each with different inputs. From these three main ANNs, further investigation of different inputs was done such as only using data from test cycles 6 – 20 to conduct the analysis. This was done due to suspicion that the curing process affected the initial images of the slabs causing the aggregate to be white in colour which may have affected the initial images from earlier test cycles. **Table 28** summarizes the ANN models and their input variables. ANN1 used raw (not normalized) image parameters (mean, variance, and skewness) of the slabs including images and data obtained from the control slab. ANN2 used normalized image parameters where the images were normalized by dividing each image in a test cycle by the mean gray level of the mean control image of each test cycle. Since ANNs 1 and 2 were different sizes due to the use of the control slab, ANN3 was evaluated where no control data was used so that ANN2 and ANN3 could be directly compared. Additionally, ANN2 and ANN3 used the same testing indices so their direct comparison is even more accurate. Each ANN was analyzed with variations of input parameters such as only considering test cycles 6 – 20 as well as including the prewet time as an input variable.

The prewet time was used as a fourth input to the ANNs in addition to the three image parameters. This was done to evaluate the effect of having a variable with an obviously strong relationship to the SIR. Using prewet time as an input would mean that models would require a prewet time to make predictions. This goes against the objective of the research which was to create a model that could use solely image parameters to predict a pavements SIR, however, using the prewet as an input may improve accuracy and simplify the ASTM procedure to still meet the objective of improving and optimizing maintenance procedures for permeable pavements.

Table 28 ANN models evaluated and their input variables

Model	Input Variables
ANN1	Mean, Variance, Skewness (with Control)
$ANN1_{15Cycles}$	Mean, Variance, Skewness (Test Cycles 6-20, with Control)
$ANN1_{Prewet}$	Mean, Variance, Skewness, Prewet Time (with Control)
ANN2	Normalized Mean, Variance, Skewness (no Control)
$ANN2_{15Cycles}$	Normalized Mean, Variance, Skewness (Test Cycles 6-20 no Control)
$ANN2_{Prewet}$	Normalized Mean, Variance, Skewness, Prewet Time (no Control)
ANN3	Mean, Variance, Skewness (no Control)
$ANN3_{15Cycles}$	Mean, Variance, Skewness (Test Cycles 6-20, no Control)
$ANN3_{Prewet}$	Mean, Variance, Skewness, Prewet Time (no Control)

6.3.2 Optimizing Hyperparameters

Nine ANN models were evaluated, each with different input variables. For each model, hyperparameters were carefully selected. A step-by-step approach was used where one hyperparameter was evaluated at a time before moving to the next hyperparameter. Each hyperparameter with the resulting lowest cost function (RMSE) was selected to be used in the optimized model. The number of hidden layers was first evaluated followed by the training/validation/testing data split. The training function was then selected followed but the input and output transfer functions. Finally, the early stopping criteria was selected and the chosen hyperparameters were used in the final optimized model. **Figure *32*** - **Figure *37*** shows different hyperparameters and their effect on a cost function for ANN1. The selected value for each hyperparameter is noted in the text box in each figure. In general, the value of each hyperparameter was selected with one that minimized the cost function. For hyperparameters that were not obvious such as the data split, input transfer function, and early stopping criteria, best judgement was used for selecting their values. For example, a typical training-validation-testing data split of 60-20-20 was used since the effect of different data splits on RMSE was not obvious.

The training function was difficult to decide as the RMSE plot was not as obvious as the other hyperparameters (see **Figure 34**). The TRAINLM algorithm was seen to have a lower mean RMSE but had more overall variability when compared to the TRAINSCG algorithm. It was observed that the Levenberg-Marquardt algorithm did train the model quicker and had less instances of training taking a high number of epochs before stopping. To that end, the Levenberg-Marquardt training algorithm was used. This process of choosing the optimal value for each hyperparameter was repeated for all models. **Table 29** gives a summary of the hyperparameters used for each of the four ANN models. In general, most of the hyperparameters were similar with exception to the number of hidden layers.

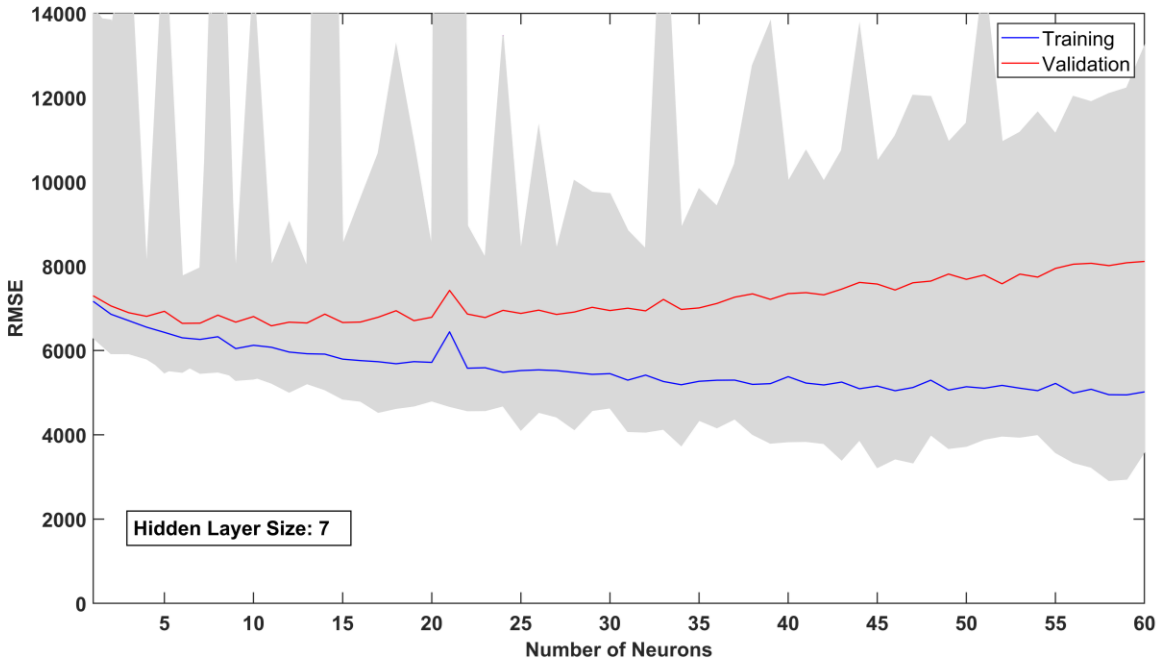


Figure 32: ANN1 RMSE with different Hidden Layer Sizes

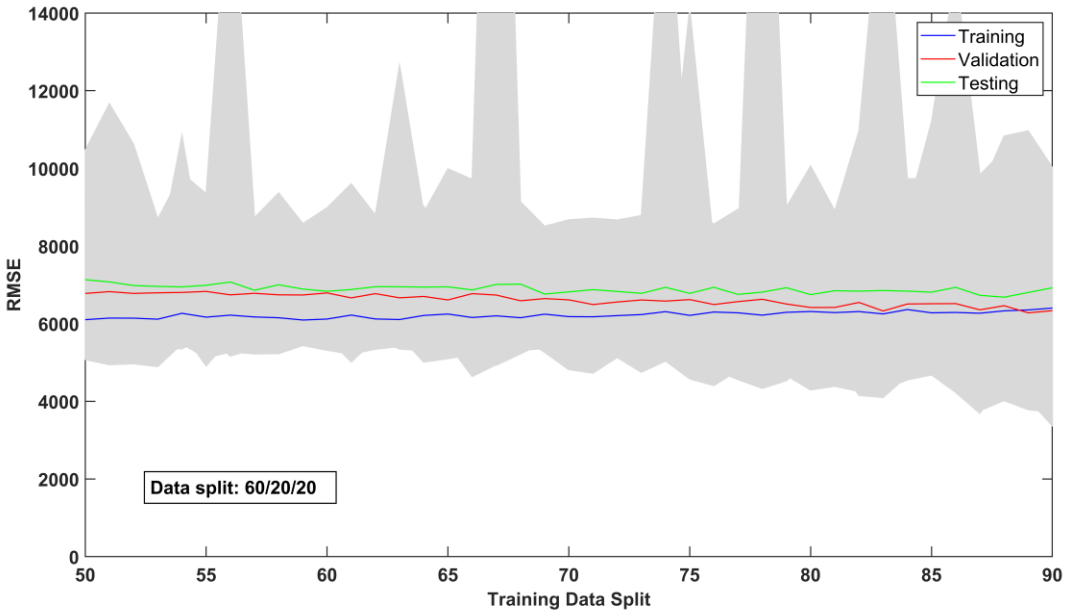


Figure 33: ANN1 RMSE with different Data Splits

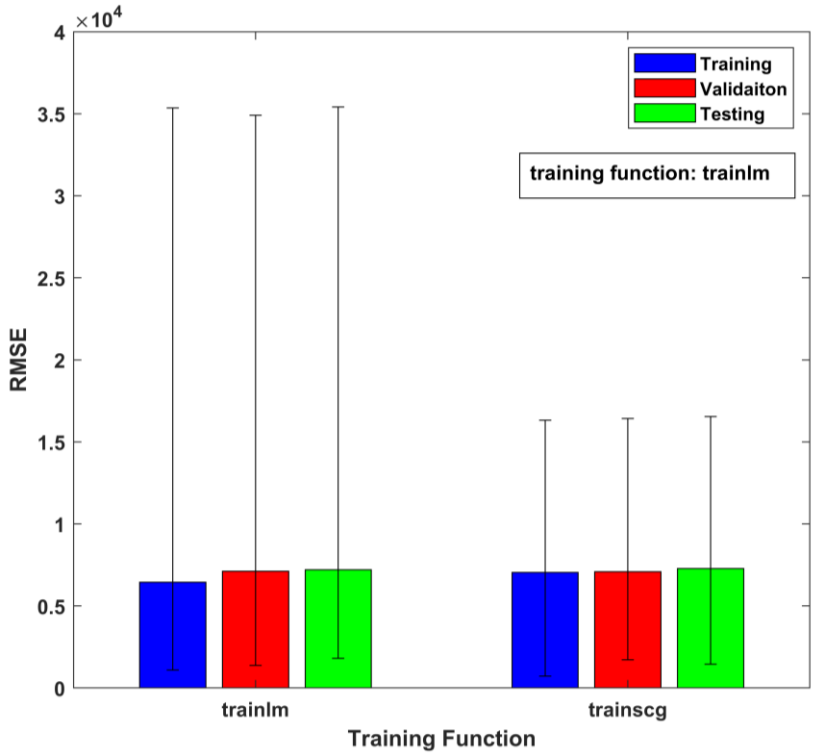


Figure 34: ANN1 RMSE with different Training Functions

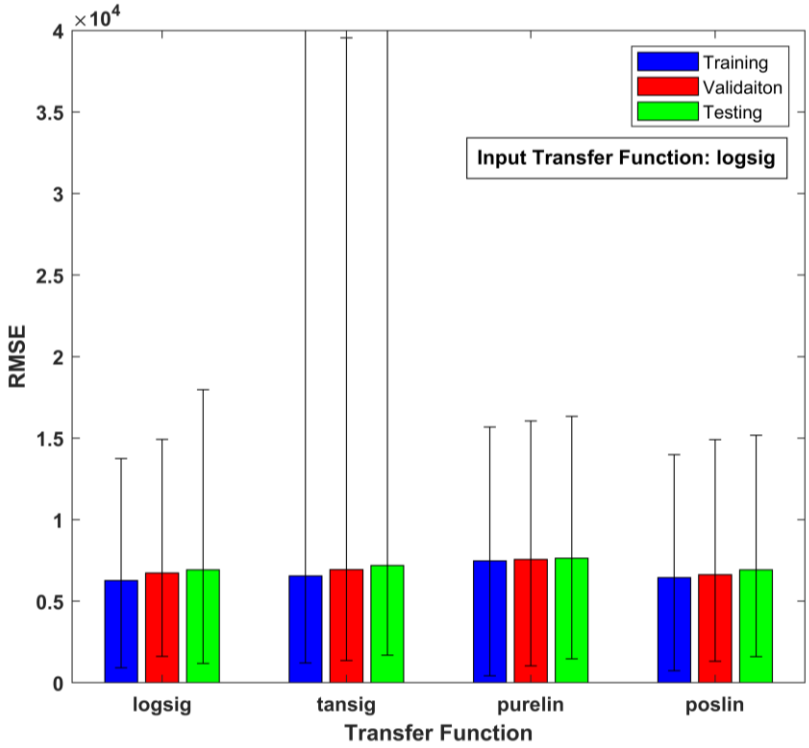


Figure 35: ANN1 RMSE with different Input Transfer Functions

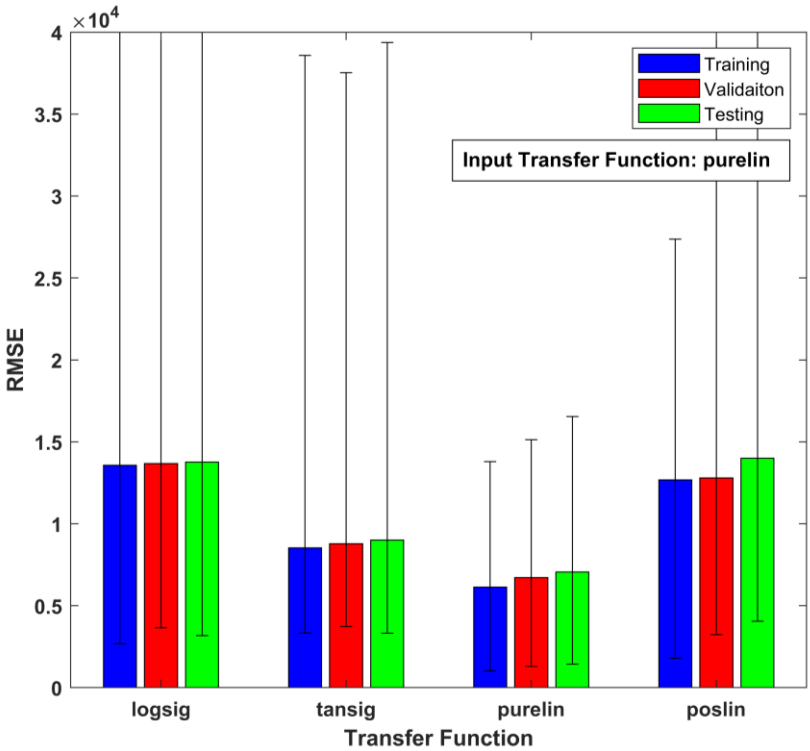


Figure 36: ANN1 RMSE with different Output Transfer Functions

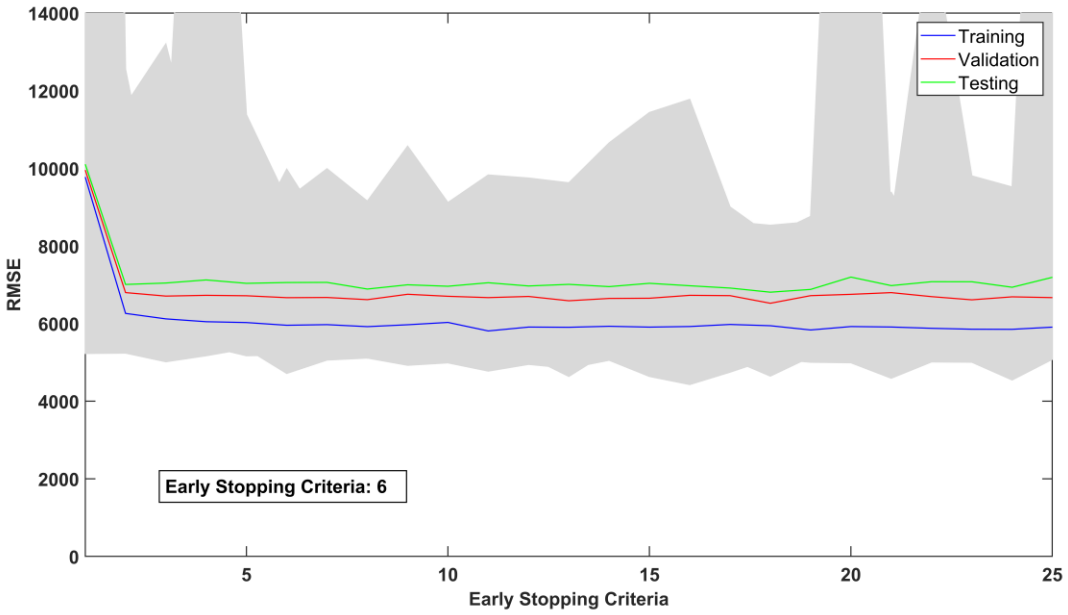


Figure 37: ANN1 RMSE with different Early Stopping Criteria

Table 29: Summary of hyperparameters used for each ANN1 model

Hyperparameter	**ANN1**	**ANN1$_{15Cycles}$**	**ANN1$_{Prewet}$**
Hidden Layer Size	7	10	8
Data Split	60/20/20	60/20/20	60/20/20
Training Function	TRAINLM	TRAINLM	TRAINLM
Input Transfer Function	LOGSIG	LOGSIG	LOGSIG
Output Transfer Function	PURELIN	PURELIN	PURELIN
Early Stopping Criteria	6	6	6

Table 30: Summary of hyperparameters used for each ANN2 model

Hyperparameter	**ANN2**	**ANN2$_{15Cycles}$**	**ANN2$_{Prewet}$**
Hidden Layer Size	7	7	7
Data Split	60/20/20	60/20/20	60/20/20
Training Function	TRAINLM	TRAINLM	TRAINLM
Input Transfer Function	LOGSIG	LOGSIG	LOGSIG
Output Transfer Function	PURELIN	PURELIN	PURELIN
Early Stopping Criteria	6	6	6

Table 31: Summary of hyperparameters used for each ANN3 model

Hyperparameter	**ANN3**	**ANN3$_{15Cycles}$**	**ANN3$_{Prewet}$**
Hidden Layer Size	8	9	9
Data Split	60/20/20	60/20/20	60/20/20
Training Function	TRAINLM	TRAINLM	TRAINLM
Input Transfer Function	LOGSIG	LOGSIG	LOGSIG
Output Transfer Function	PURELIN	PURELIN	PURELIN
Early Stopping Criteria	6	6	6

6.3.3 Model Performance

Three metrics were used to evaluate the performance of the nine ANNs. These included the coefficient of determination (R^2), the percent capture, and the root-mean-square error (RMSE). Each performance metric gives an indication of different aspects of the model. The R^2 captures the variance, the percent capture measures the robustness and fitting, and the RMSE gives insight on the residuals. Each performance metric was obtained for each of the stages of developing the model including training, validation, and testing as well as using the full dataset. These results are tabulated in **Table 32** – **Table 35**. In general, it is seen that the training and validation performances are the highest which is expected since the model is being calibrated with this data. The testing dataset is new data that the model is not calibrated with and thus is seen to have the lowest performance. The testing data was consistent throughout the model calibration process and each ensemble used the same testing data within each model. ANN1 and ANN2 had different sizes due to ANN1 using the control data as inputs to the model. ANN3 did not use control data as inputs to the model and thus had the- same size as ANN2. Thus, each model derivation between ANN2 and ANN3 used the same testing data (ex. ANN2 and ANN3 used the same testing data; ANN2$_{15Cycles}$ and ANN3$_{15Cycles}$ used the same testing data, etc.).

ANN1 uses the raw inputs of the entire dataset including the control and was seen to have a lower full dataset performance compared to a model that used data from the last 15 test cycles. Even though the model with 15 test cycles used 25% less data overall,

it was still seen to perform the same or even better. This suggests that the use of the first 5 test cycles may be adding noise to the overall dataset due to unaccounted variables that were present in the images of the slabs for the earlier test cycles. Although, it should be noted that this is not the case in ANN2 and ANN3 where the use of just the last 15 test cycles decreased the overall performance. This may be due to insufficient data since ANN2 and ANN3 already had low number of datapoints and further removing data by only using the last 15 test cycles may have caused a reduction of model performance.

Table 32: Training Performance

Model	**R^2**	**Percent Capture**	**RMSE (mm/hr)**
ANN1	0.80	92.82	6197
$ANN1_{15Cycles}$	0.80	94.14	5906
$ANN1_{Prewet}$	0.97	96.62	2182
ANN2	0.78	92.44	6389
$ANN2_{15Cycles}$	0.71	92.77	6126
$ANN2_{Prewet}$	0.97	98.45	2354
ANN3	0.76	96.25	6709
$ANN3_{15Cycles}$	0.66	96.91	6554
$ANN3_{Prewet}$	0.97	97.77	2252

Table 33: Validation Performance

Model	**R^2**	**Percent Capture**	**RMSE (mm/hr)**
ANN1	0.77	77.66	6643
$ANN1_{15Cycles}$	0.75	79.28	6486
$ANN1_{Prewet}$	0.97	80.60	2408
ANN2	0.74	81.27	6949
$ANN2_{15Cycles}$	0.65	90.50	6686
$ANN2_{Prewet}$	0.96	81.16	2538
ANN3	0.71	87.94	7455
$ANN3_{15Cycles}$	0.58	93.28	7316
$ANN3_{Prewet}$	0.96	88.14	2655

Table 34: Testing Performance

Model	R^2	Percent Capture	RMSE (mm/hr)
ANN1	0.73	65.96	7115
$ANN1_{15Cycles}$	0.73	80.56	6775
$ANN1_{Prewet}$	0.96	68.09	2715
ANN2	0.76	61.43	7441
$ANN2_{15Cycles}$	0.71	75.93	6605
$ANN2_{Prewet}$	0.97	87.14	2594
ANN3	0.74	62.86	7670
$ANN3_{15Cycles}$	0.65	72.22	7440
$ANN3_{Prewet}$	0.97	82.86	2671

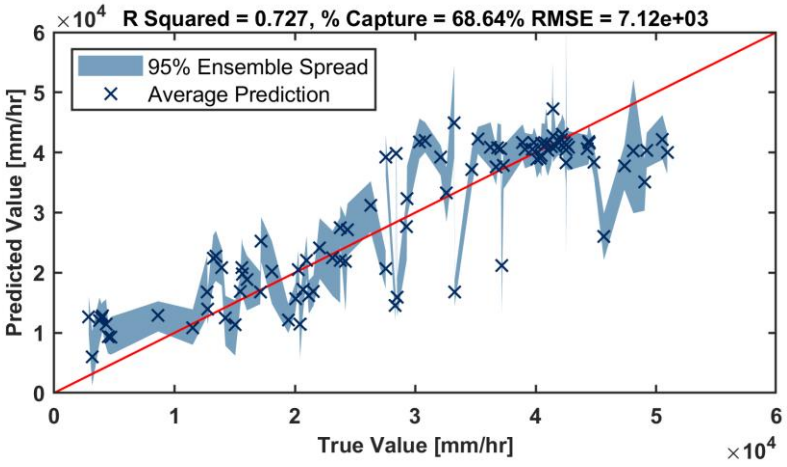
Table 35: Full Data Performance

Model	R^2	Percent Capture	RMSE (mm/hr)
ANN1	0.78	68.64	6489
$ANN1_{15Cycles}$	0.77	76.94	6286
$ANN1_{Prewet}$	0.97	70.76	2350
ANN2	0.77	66.38	6735
$ANN2_{15Cycles}$	0.70	73.70	6347
$ANN2_{Prewet}$	0.97	87.75	2450
ANN3	0.75	68.09	7072
$ANN3_{15Cycles}$	0.64	74.44	6913
$ANN3_{Prewet}$	0.97	83.76	2455

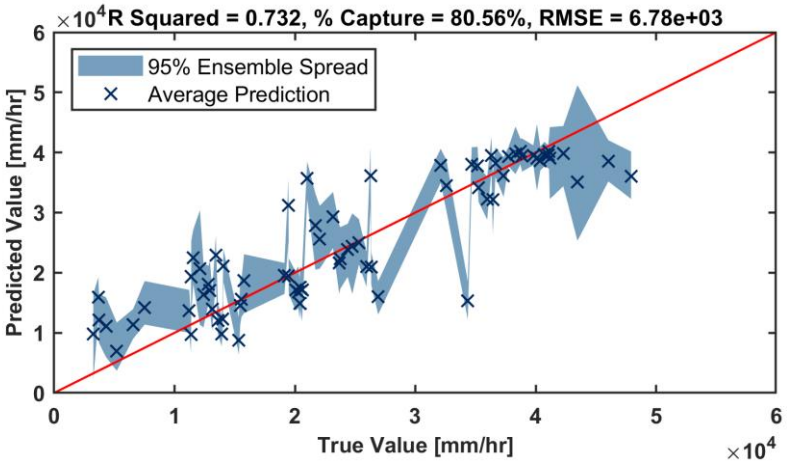
Using the prewet time as an input variable was seen to generate the most accurate models, as expected. This is due to the high correlation of the prewet time and the SIR. Although the use of just image parameters is capable of generated adequate predictive models for this dataset, using the prewet time as an input variable shows there is some variance that is not being accounted for when using just image data. This can even be explained by looking at the SIR of each slab. All four slabs were cast with the same mix

and on the same day with the same procedure, however, all four slabs were seen to have different starting SIRs. The images of the surfaces of these pavements would not indicate the differences of SIR between the slabs in their early conditions. It is the change of the surface conditions as SIR decreases that provides insight to the model for it to predict the SIR. Although the image parameters were seen to change over the duration of the testing period, there is no ability of the model to discern the SIR between different slabs for the same test cycle. This is the variability that the prewet time provides and thus enhances the predictive capabilities of the model. Although this may be true, using the prewet time as an input to the model is unintuitive since using prewet time as an input to predict SIR is similar to needing SIR to predict SIR. With that being said, using a model with prewet time would still be beneficial since the ASTM standard only required 3.6 kg of water for determining the prewet time whereas it required the use of 18 kg of water for determining the SIR. From that perspective, using a model that uses prewet time as an input variable would provide an efficient and less labour-intensive way of determining the SIR of a permeable pavement.

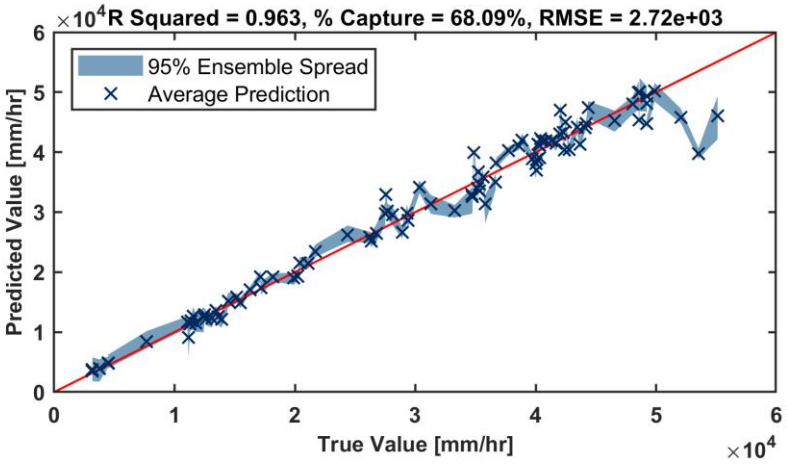
Comparing the models that used normalized data (ANN2) with the raw images (ANN3) we see that ANN2, in general, had higher performance than ANN3. This suggests that normalizing the data by removing the trends seen in the control slab improved model accuracy. ANN2 and ANN3 had the same testing data so a direct comparison between the two models can be made. Comparing ANN1 and ANN3 we see that removing the control data decreased the overall performance of the model. This is likely due to the fact that ANNs benefit from having more input data. Using the control data in ANN1 provided meaningful insight to the model for the data the control represented. This included more image parameters for the SIR ranges of the control slab. This allowed the model to make more accurate predictions for those SIR values due to the extra insight provided by the control data. To visualize the performance of these models, the observed SIRs were graphed with the predicted SIRs from the testing set for each model. **Figure 38** to **Figure 40** show the true vs predicted plots for the testing dataset results for the three different ANNs and their subsets where the average prediction is the median of the ensemble.



(a)

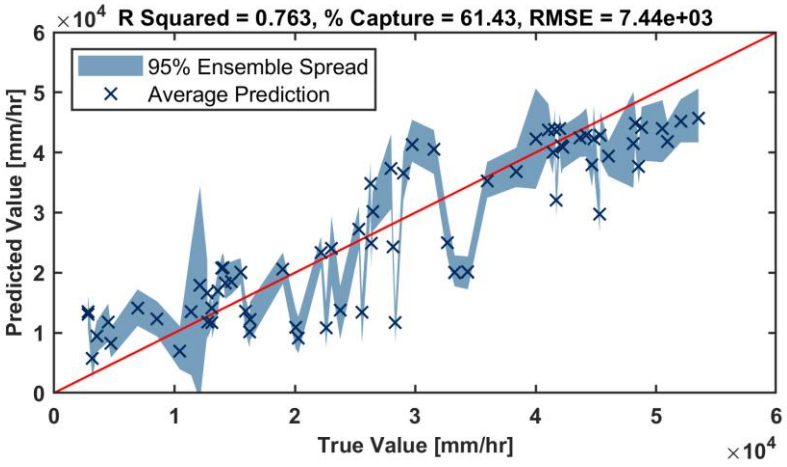


(b)

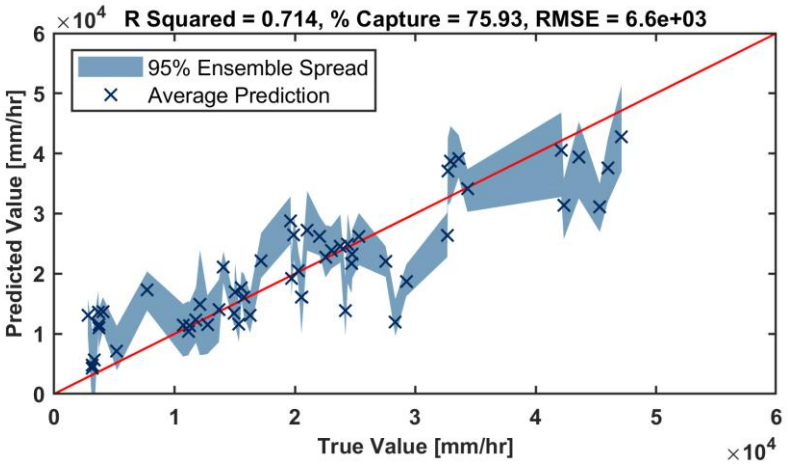


(c)

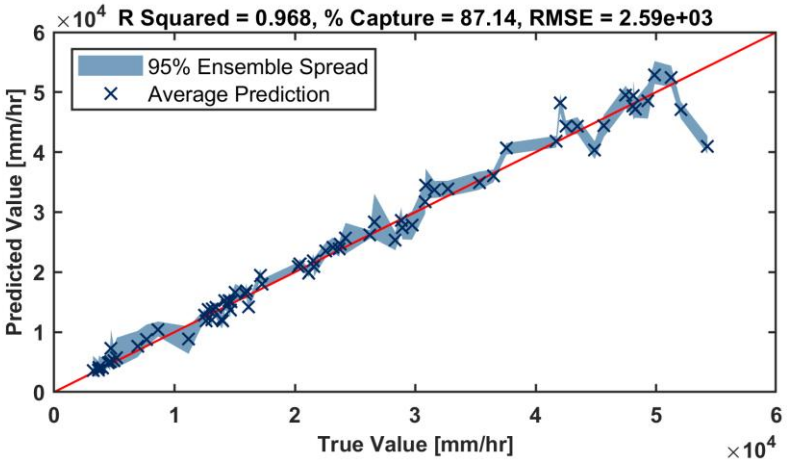
Figure 38: ANN1 True vs Predicted plots for Testing data for (a) ANN1, (b) $ANN1_{15Cycles}$, and (c) $ANN1_{Prewet}$



(a)

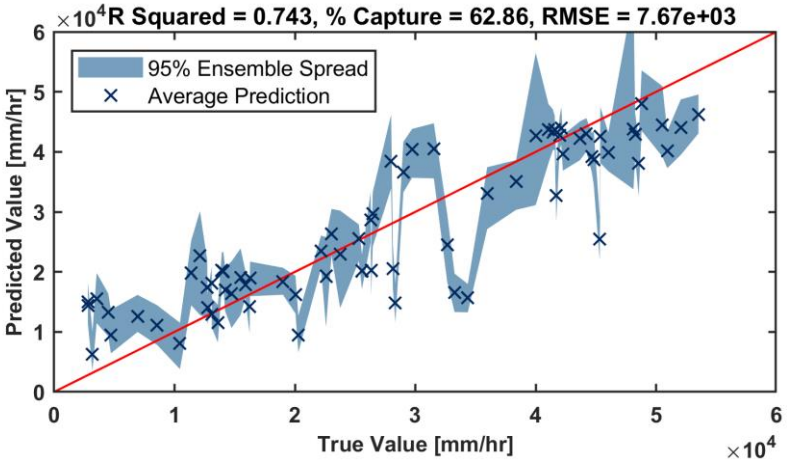


(b)

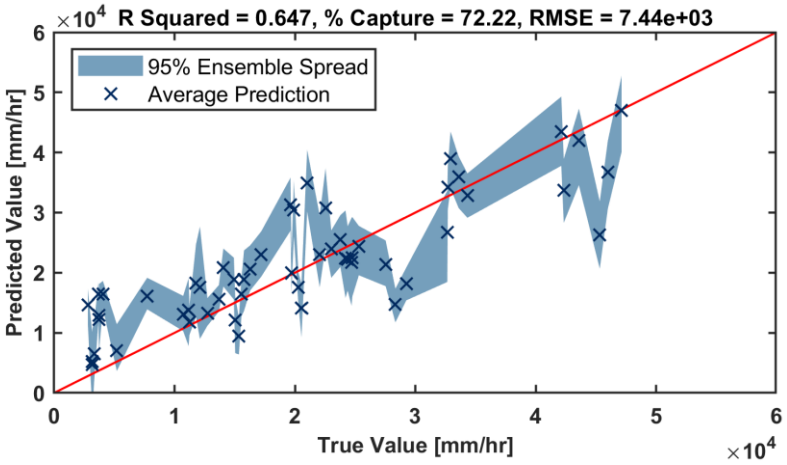


(c)

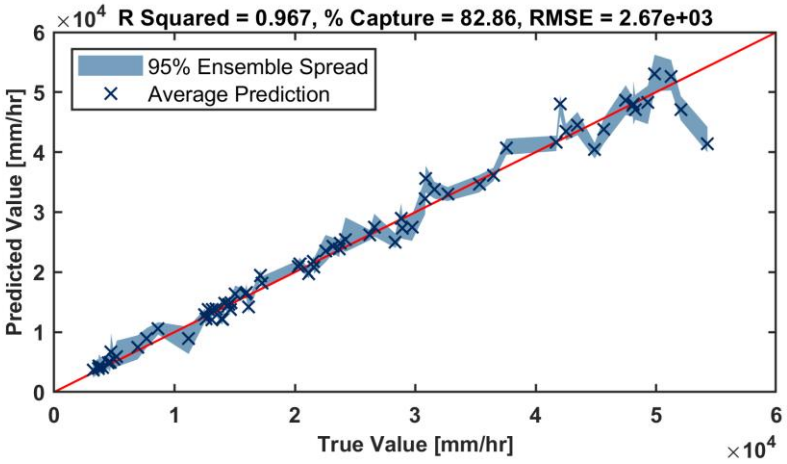
Figure 39: ANN2 True vs Predicted plots for Testing data for (a) ANN2, (b) $ANN2_{15Cycles}$, and (c) $ANN2_{Prewet}$



(a)



(b)



(c)

Figure 40: ANN3 True vs Predicted plots for Testing data for (a) ANN3, (b) $ANN3_{15Cycles}$, and (c) $ANN3_{Prewet}$

Plotting the true SIRs with the predicted SIRs from each model provides insight on how the model is performing as well as where the model is overpredicting and underpredicting. In general, it was observed that the models were less accurate at higher SIRs. This could be due to a lack of data at those SIRs as well as unaccounted for variances in the input data. Higher SIRs indicate earlier test cycles where the slabs were not clogged. There likely was some sources of error that was present in the images of the earlier test cycles. Additionally, in ANNs that did not use prewet as an input, SIRs with a true value of around 30,000 – 40,000 mm/hr were seen to have a high range of predicted values. This may be due to a large amount of data within that SIR range. In addition to the control slab, all slabs experienced SIRs within this range. Additionally, the slabs that started with higher SIRs eventually degraded to this SIR range with sediment on the surface of the slab. This may have caused some variation in the model since some image data consisted of clean slabs whereas other images had sediment already present at that SIR. Thus, this difference in clean and clogged images may have caused the large range in predictions at that SIR range.

To understand where in the input data range the model is underperforming, 3D plots with different combinations of input variables was graphed with the observed and predicted SIR. **Figure 41** - **Figure 46** show 3D plots for different combinations of inputs for ANN1 to show where the model underperforms. The mean did not have a high correlation with the SIR whereas the variance and especially the skewness were seen to have high correlation. At high mean values as well as extreme values of variance and low values of skewness the model seemed to be less accurate. This is showcased in the figures where at those locations, there are large variations in the predicted SIR. A high mean, variance, and low skewness was seen with earlier test cycles. This highlights that the images may have had unexpected variables that were not accounted for in earlier test cycles and were not present in later test cycles. This variance may also be due to not having substantial data at the extreme values of these input variables. It is also important to understand the effect changing the value of the inputs has on the overall model. To do this, a sensitivity analysis was conducted where particular values (e.g., high and low) of input variables were used with the model to test the robustness of the output.

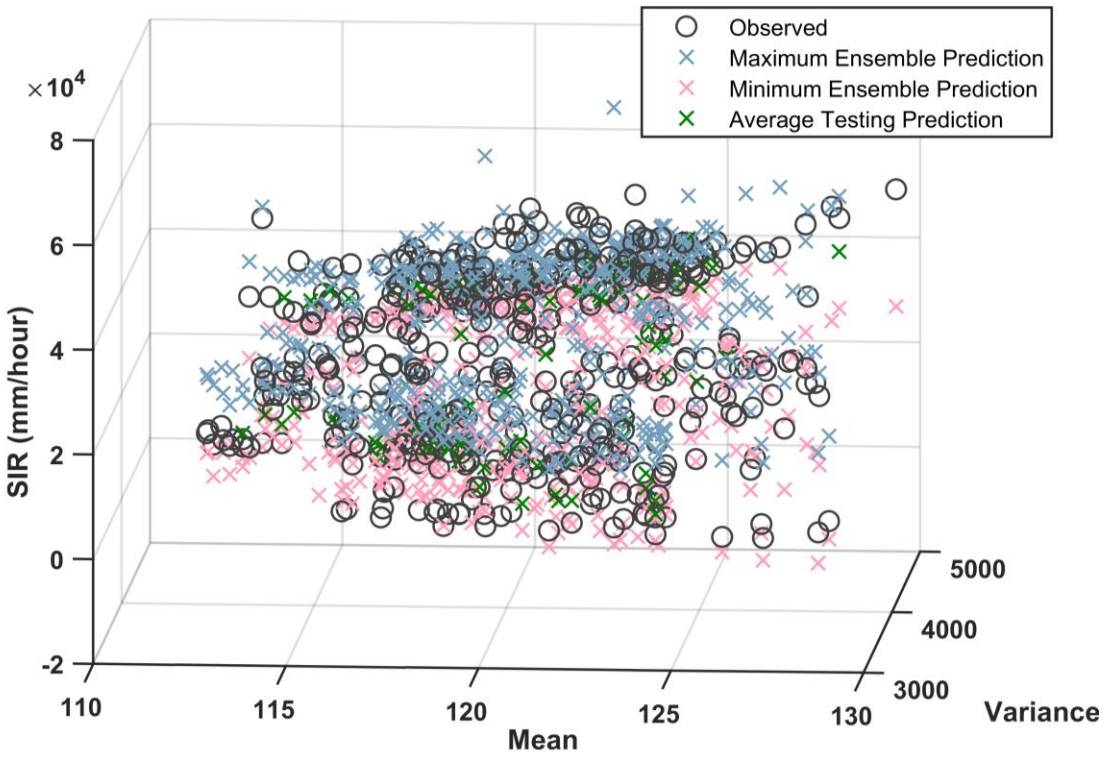


Figure 41: ANN1 Mean vs Variance vs SIR (Observed and Predicted)

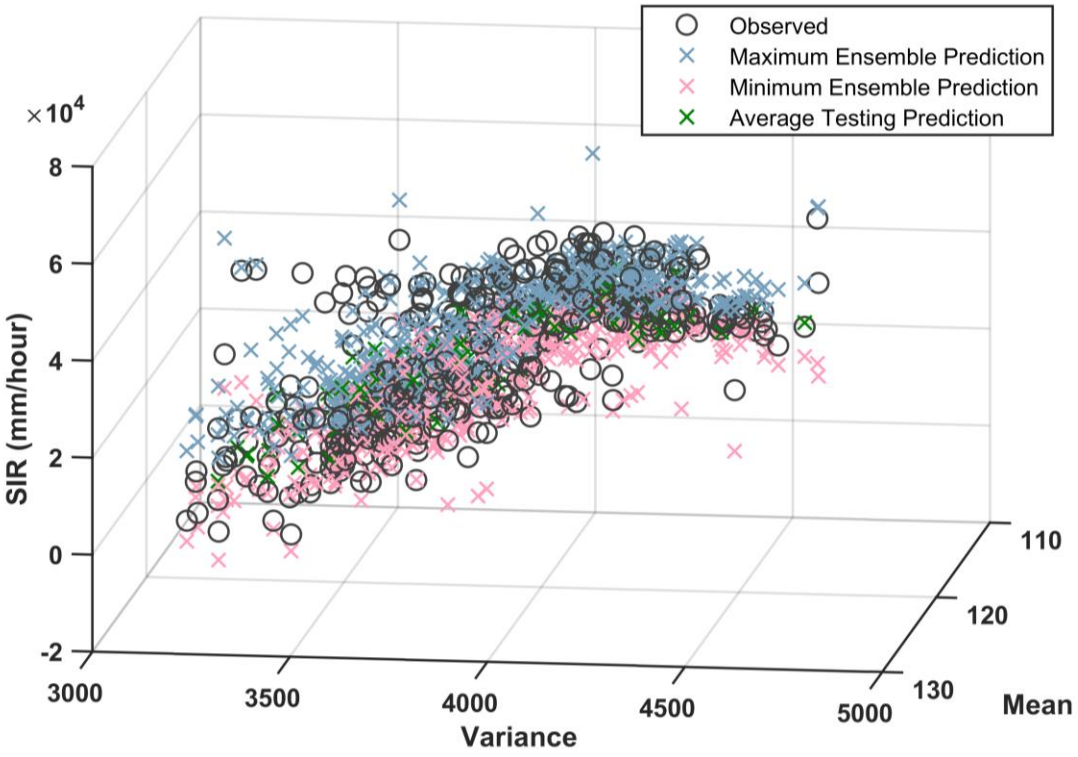


Figure 42: ANN1 Variance vs Mean vs SIR (Observed and Predicted)

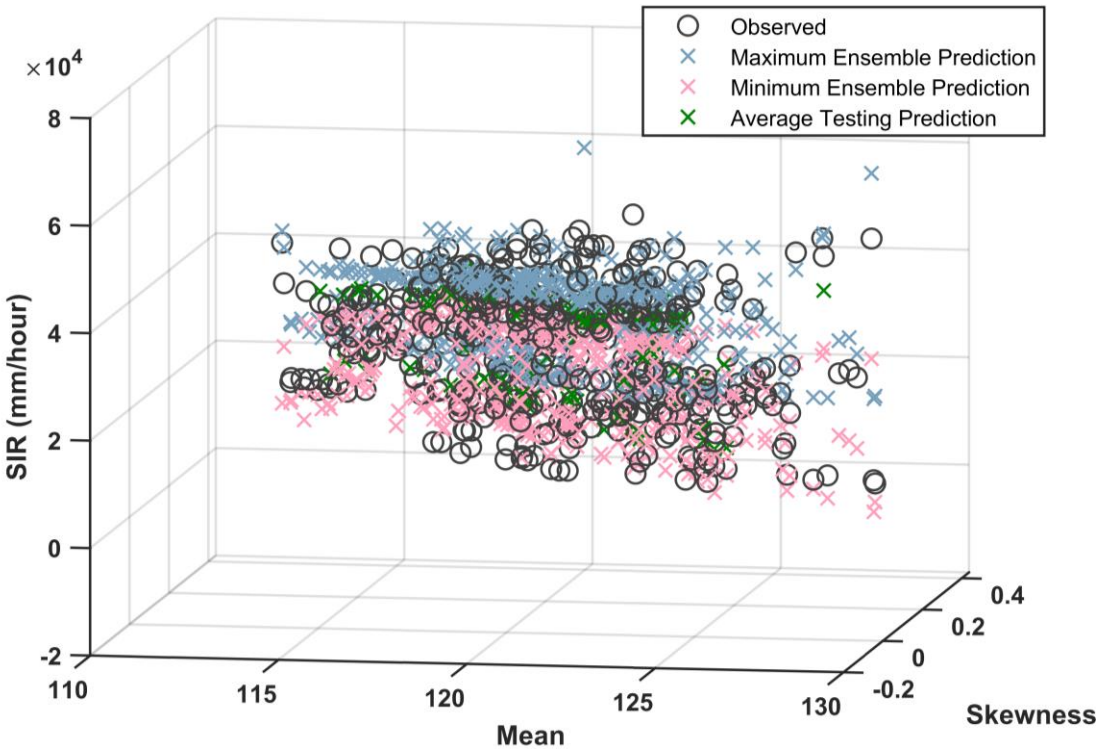


Figure 43: ANN1 Mean vs Skewness vs SIR (Observed and Predicted)

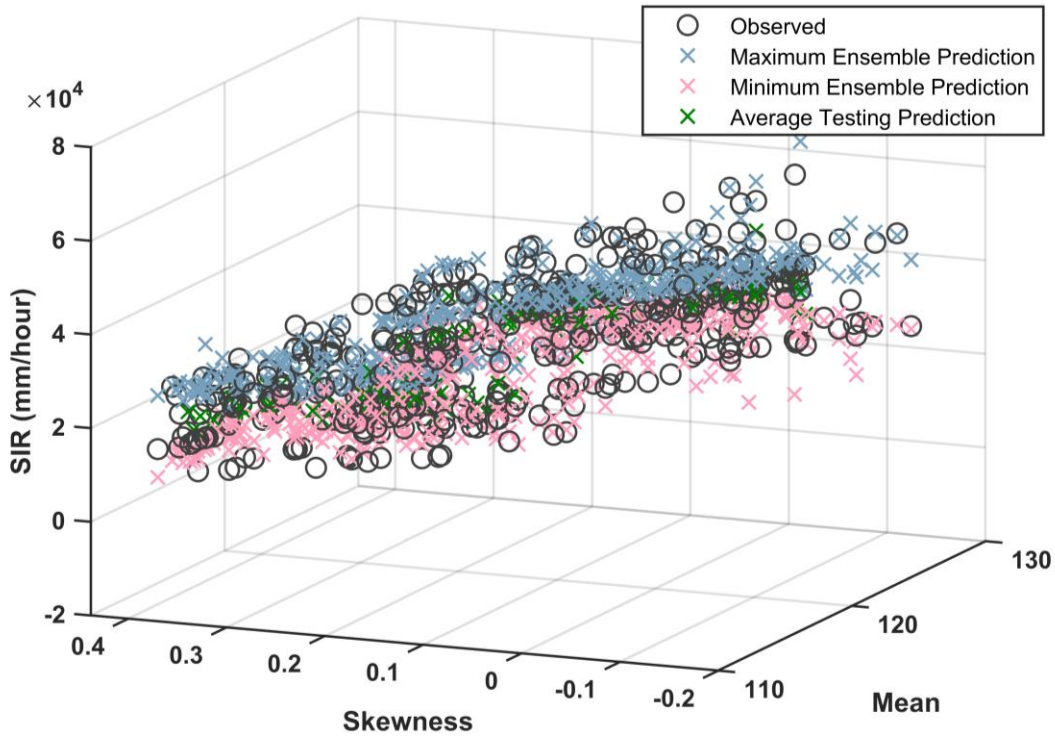


Figure 44: ANN1 Skewness vs Mean vs SIR (Observed and Predicted)

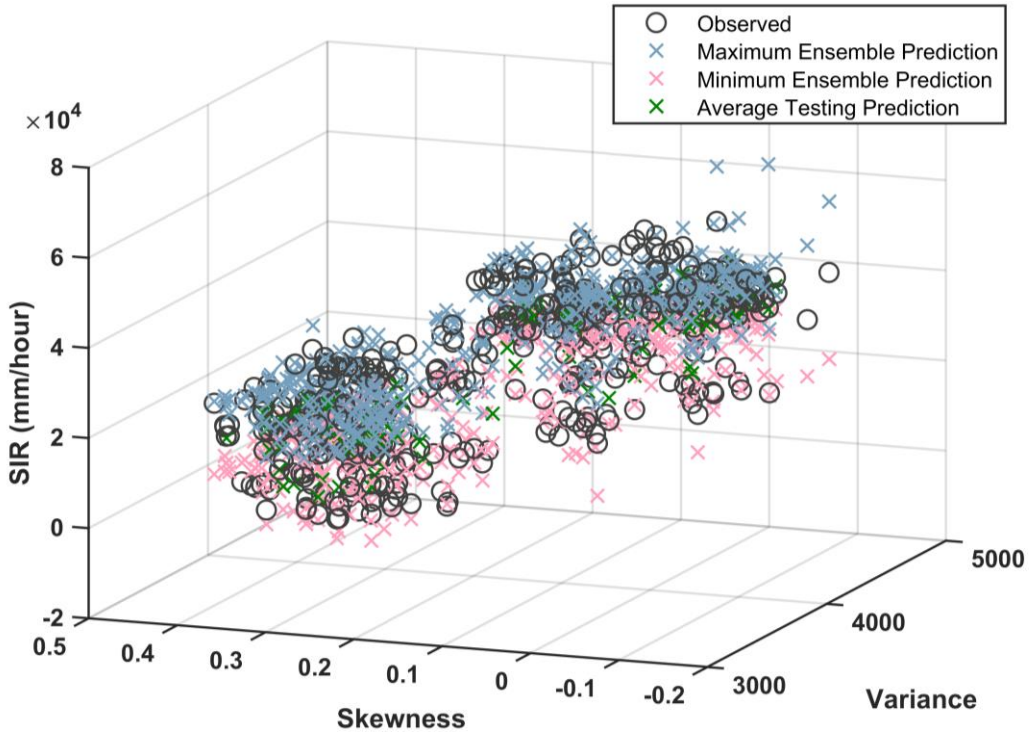


Figure 45: ANN1 Skewness vs Variance vs SIR (Observed and Predicted)

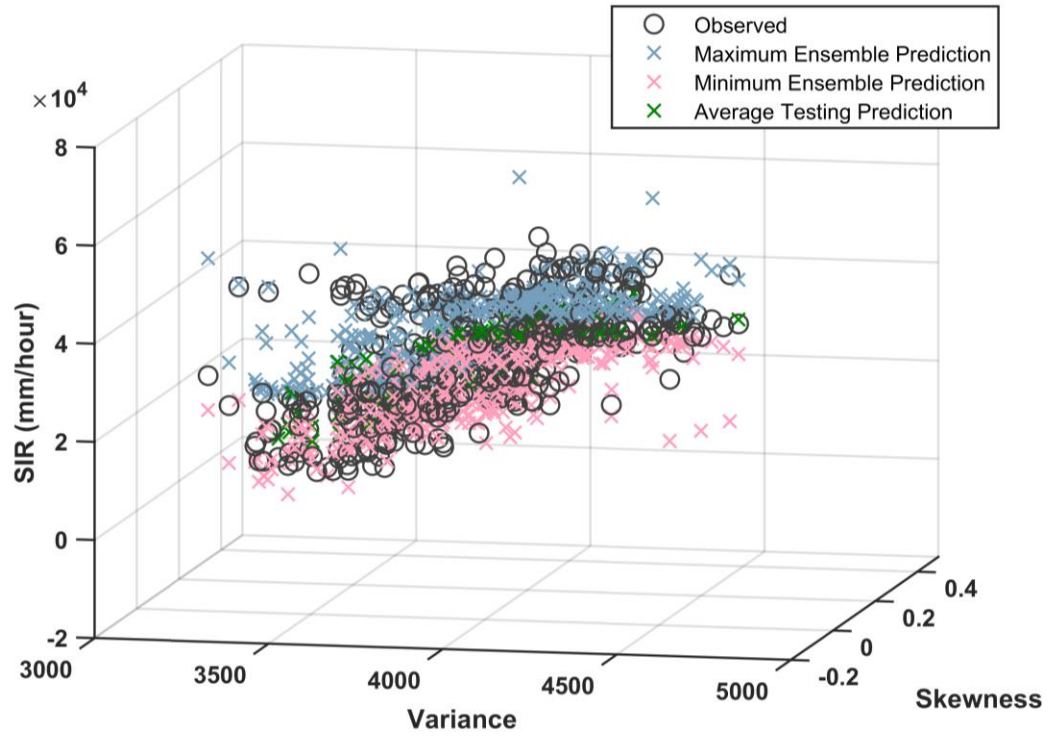


Figure 46: ANN1 Variance vs Skewness vs SIR (Observed and Predicted)

6.3.3 Sensitivity Analysis

A brief sensitivity analysis was conducted where each input used in ANN1 were given a range of values while holding the other inputs constant. The given range of values correspond to the observed range seen for that input variable. The constant values were chosen as either the minimum or the maximum values found in the observed dataset. For example, a range of mean values were given from 100 – 140 while the variance and skewness were kept constant along the range of mean values and corresponded to the lowest and highest variance and skewness values found in the dataset. Different combinations of these low and high variance and skewness values were used with the range of given mean values. These inputs were given to the optimized ANN1 model to assess the model's capability in predicting SIR given these different ranges of input values and to visually assess the impact of each input on the predictions.

An ensemble approach was used where the inputs were given to the ANN 100 times. The median of the predictions for each input case was reported. **Figure 47** shows the predicted SIR values for a range of mean values while the variance and skewness were given different combination of low and high values. With a low variance and low skewness, an increasing mean gray level yields a lower predicted SIR. This means that a compact histogram with a mean that shifts towards being whiter was shown to decrease the SIR. Physically, this means that as an image were to get whiter, and the histogram was compact, the SIR would decrease. The other three combinations of variance and skewness have constant predictions over the range of different mean values. This means that the signal from the extreme values of either the variance or skewness are causing the predictions to be constant. In general, a higher variance was seen to cause the model to predict higher SIRs. Additionally, higher skewness typically meant lower SIR predictions. This suggests that a grayscale histogram with a lower variance, larger skewness, and higher mean would yield lower SIRs.

Input values for the variance were given values ranging from 3000 – 5000 with the mean and skewness held at different combinations of low and high values. **Figure 48** shows the results of the median predictions for each combination. The model with low mean accompanied by low skewness consistently predicted high SIR values for the entire

range of given variance values. This means that a low mean and low skewness (darker image with a smaller skew) is weighed heavily in the model towards higher SIR values. The last three combinations show a general increase in SIR as variance increases. High variance values with a high mean and a low skewness caused the model to predict higher SIRs. This suggests that even if an image were to have a high mean (whiter), the combination of having a higher variance with low skewness is causing the model to predict higher SIRs. This suggests that not only is the average colour of the image important, but the shape of the overall histogram is also key in predicting SIRs. High skewness values are seen to have consistently low predictions despite low or high mean values when changing the variance.

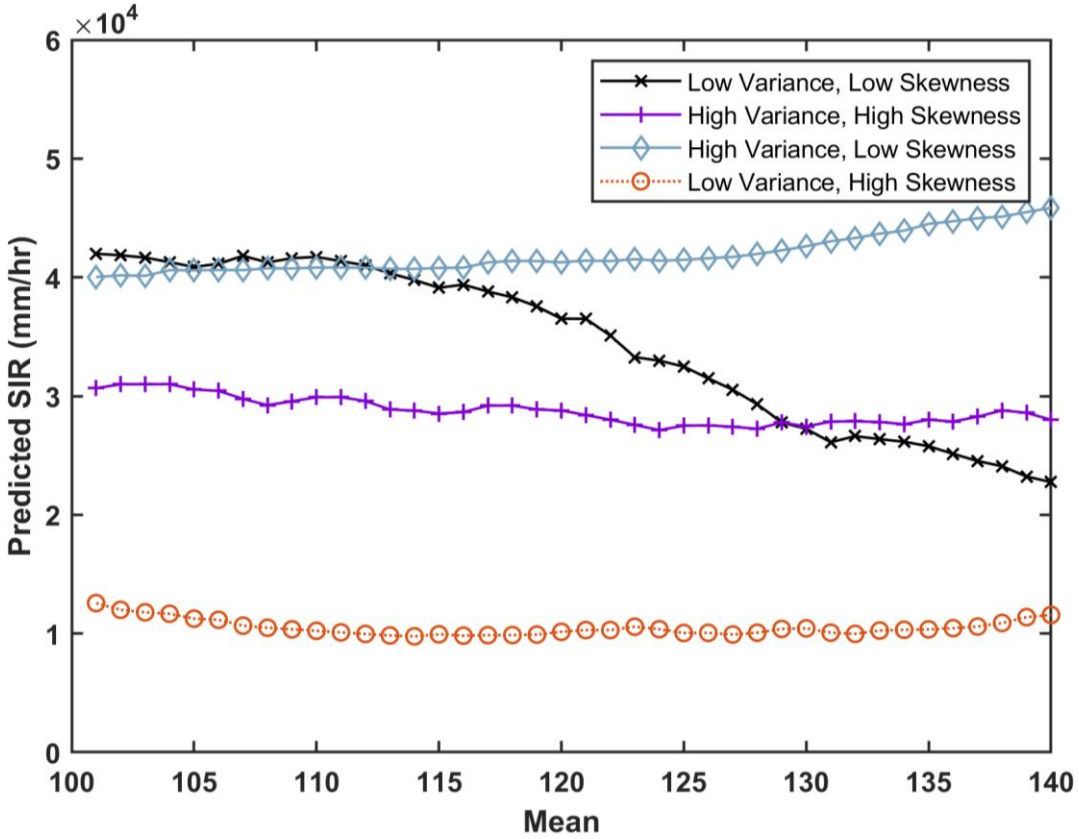


Figure 47: Median prediction for different values of Mean where Variance and Skewness were either low or high values (ANN1)

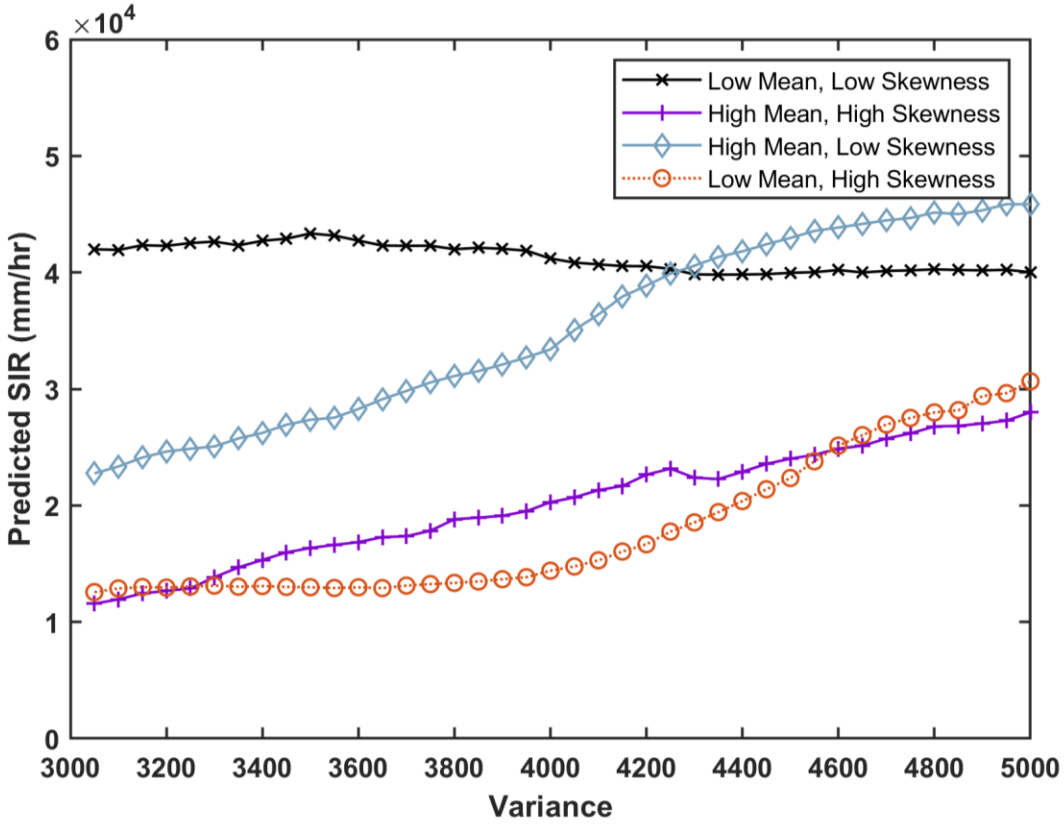


Figure 48: Median prediction for different values of Variance where Mean and Skewness were either low or high values (ANN1)

Different values of skewness were used as inputs while the mean and variance were kept as either the minimum or maximum values observed in the dataset. **Figure 49** summarizes the predictions of SIR of ANN1 when skewness changes with different low and high values of the mean and variance. In general, all combinations of the mean and variance see a general decrease in SIR as skewness increased. A high mean with low variance was seen to give the lowest predictions. This means a whiter image with a lower spread of pixels. A high mean with high variance was seen to give larger predictions in SIR. This means that a histogram that is whiter does not necessarily mean a smaller SIR and that the spread of the histograms also dictates the SIR. The change of the skewness was seen to have the most change in the predicted SIR. All four low and high scenarios of the mean and variance were seen to have a decrease in SIR as skewness increased. A high-low mean and variance was seen to produce low predicted SIR. This means that a compact histogram centered at whiter pixels resulted in lower SIR predictions.

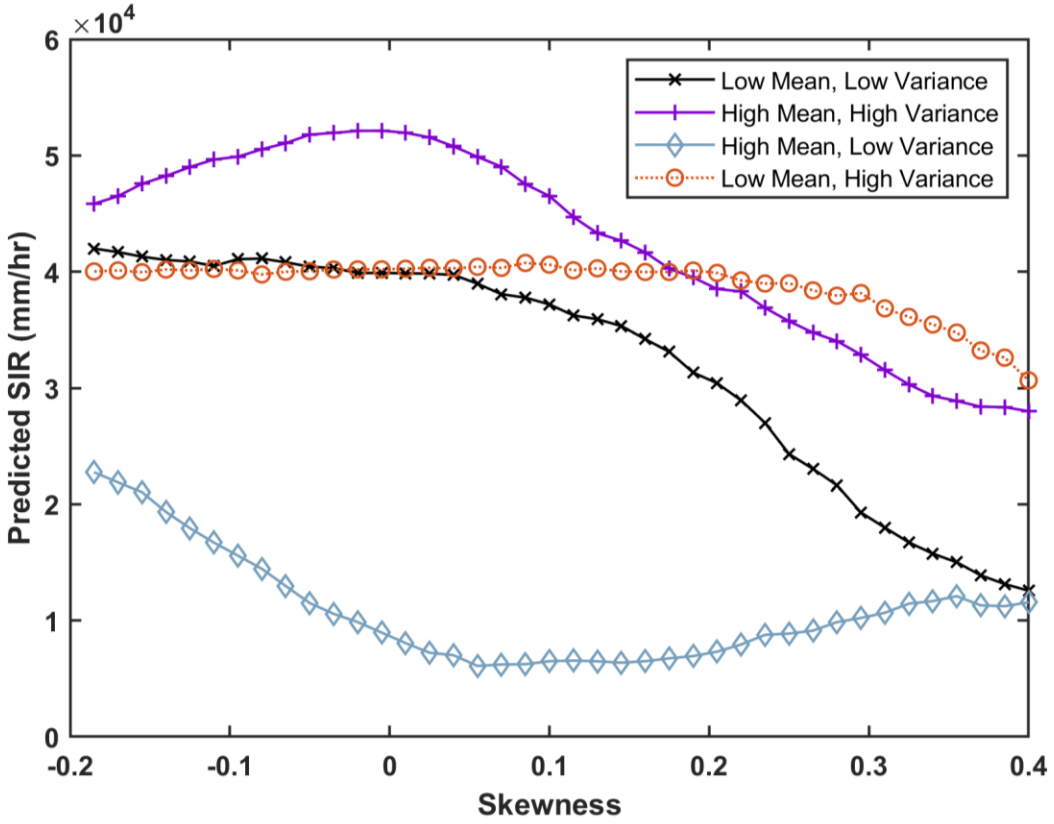


Figure 49: Median prediction for different values of Skewness where Mean and Variance were either low or high values (ANN1)

A sensitivity parameter (S) was calculated for each input parameter and each high-low scenario of the remaining inputs with $S = \frac{\partial Y}{\partial P_i}$ where Y is the predicted SIR of the model and P is each input parameter (Jørgensen, 2001). **Table 36** summarizes the calculated sensitivity parameters for each of the input parameters and the different high-low combinations for the remaining parameter. The input parameters were normalized on a scale from 0 – 1 so that they can be compared easily. From the table, the skewness parameter was shown to have the largest sensitivity, meaning that the predictions were seen to change the most when skewness changed, despite holding the other variables constant. The mean was seen to have a large response when the variance and skewness were both low. This means a compact histogram with a higher mean resulted in lower SIRs. Larger SIR predictions were made as variance increased whereas smaller variances were seen to yield lower SIR predictions. This means that an image with less spread would yield lower SIRs.

Table 36 Sensitivity (x10^4) of each input for High (H) and Low (L) values of the remaining inputs

Parameter	LL	HH	HL	LH
Mean	-1.92	-0.27	0.58	-0.10
Variance	-0.20	1.64	2.31	1.81
Skewness	-2.94	-1.78	-1.12	-0.94

The analysis showed that the skewness was weighed higher than the other input variables. As skewness increased, the predicted SIR decreased for all scenarios. Additionally, a lower variance showed that, in general, the model predicted a lower SIR. As the mean changed, a general trend was not seen in terms of the predicted SIR, however, coupled with a low variance and skewness, higher mean values (whiter images) were shown to have a lower SIRs. This is all to say that, in general, grayscale pavement images with lower SIRs were found to have larger means, are more compact (smaller variance) and have a larger skewness.

6.3.4 Summary of Modelling Results

From an application perspective, the modelling results demonstrated the potential of using images of the surface of a pavement to predict its SIR. Image parameters were seen to change significantly as the slabs clogged and using these parameters as model inputs resulted in satisfactory data-driven models. These models were capable of predicting the SIR; however, the variance of predictions was seen to be high. Some of this variance was suspected to be caused by errors in collecting data such as the change in colour of the aggregate of the concrete after curing and wet-dry cycles after conducting SIR tests. It is suspected that another reason that the models experienced high variance in predicted SIR is due to the high variance of SIRs between the slabs that image data is not capable in capturing which is especially prevalent in earlier test cycles. Overall, the research showed that it is possible to predict the SIR of a pavement using images that represent the change of the surface as the pavement clogs, however, there is some underlying variance in the models that images were not able to fully capture.

CHAPTER 7: CONCLUSIONS AND RECOMMENDATIONS

This research investigated maintenance concerns of permeable concrete pavements and addressed current practices of determining their hydrologic performance. The objective of this research was to investigate the ability to use image properties as inputs to an ANN to predict the SIR of laboratory constructed pervious concrete pavement specimens to obtain the hydraulic performance of these systems and thus inform maintenance procedures and frequency. The research showed that data-driven models that used PCPC image parameters as inputs were capable of adequately predicting the SIR, however, the models demonstrated high variances in predictions and the data collected was from a controlled laboratory setting where many variables, including lighting condition, were controlled for. The research highlighted the potential of using image parameters as a method of predicting the SIR of PCPC slabs, but a more robust method needs to be evaluated for real-world applications.

7.1 Lighting Condition Analysis

A brief investigation was conducted where images of an asphalt pavement were taken from different locations and under different lighting conditions. These varying conditions included capturing images at different times of the day as well as using a camera flash and holding the camera at an angle. The distributions of the images were compared to evaluate if the images taken from different locations and under varying lighting conditions were statistically similar. In total, 253 KSDTs were conducted to compare the distributions of the grayscale histograms of these images. It was found that 88.2% of comparisons rejected the null hypothesis meaning the image histograms were from statistically different distributions. Furthermore, 11.8% of the comparisons were statistically similar, however only 3.15% of comparisons were statistically similar with different lighting conditions. Preprocessing of the images was done to try and improve the results of the analysis. After preprocessing, 21.3% of comparisons showed images followed a similar distribution which was an increase over 11.9% when no preprocessing was done. Lastly, 17.4% of comparisons showed a similar distribution when the lighting condition was different which was an increase from 3.6% with no preprocessing.

The analysis showed that pavement images acquired in the field were statistically different depending on the location and lighting condition. Images taken with different locations but with the same lighting condition were more statistically similar than those taken under different lighting conditions. Preprocessing the images did show to improve this difference however, 79% of the images were still statistically different. This analysis showed that when collecting surface image data outdoors, accounting for the variation in lighting conditions between the images is important. Consistent lighting for an image dataset may be critical in maintaining the integrity of the results, especially if statistical parameters such as the mean, variance, and skewness are being extracted from the images. Although this may be the case, properly trained models such as ANNs and CNNs may be able to overcome this obstacle.

7.2 Laboratory Analysis

Laboratory PC specimen were constructed and subjected to clogging while SIR and image data were collected. A clogging sediment made of clay and fine sand/silt was used to slowly clog the pavements over time. The slabs that received clogging were found to have a decrease in SIR of between 72 – 91% over 20 test cycles. The control slab was shown to have a decrease in SIR by 14%. This decrease in the control slab was found to be associated with sources of error when conducting the SIR test such as different personnel conducting the test and setting up the infiltration rings. The results show that even fine particle sediments are capable of clogging PCPC systems, contrary to the findings reported in the literature.

Images of the slabs were converted to grayscale and parameters including the mean, variance, and skewness were extracted from their histograms. A trend analysis was conducted on these image parameters which revealed that the trends for the image parameters for all slabs were significant and that the skewness parameter had the strongest trend. The control slab, which did not receive any clogging sediment, unexpectedly demonstrated that for all image parameters, a significant trend was identified after 20 testing cycles. This may have been due to unaccounted variables such as wet-dry cycles of the slabs. To account for the trends seen in the control slab, a normalization procedure was completed whereby the images of the slabs that received

clogging were divided by the mean histogram value of the average control slab image in each test cycle. This normalization technique centered the mean of the control slabs while the slabs that received clogging showed an increasing trend of the mean of the images as clogging progressed. This suggests that if variables are properly accounted for, the mean of the images increases as SIR decreases. Normalizing the images in this way did not significantly reduce the strength of the trends that were seen in the raw image parameters which confirms that adding sediment to the surface of a pavement does significantly affect the images of the surface.

7.3 Modelling Performance

ANN models were created that used grayscale image parameters as inputs to predict PC slab SIRs. The model with the raw images as well as data from the control slab (ANN1) yielded a testing performance of R^2, percent capture, and RMSE of 0.73, 65.96, and 7115 respectively. Normalized inputs were used in ANN2 without the control slab to investigate if normalizing the parameters increased model performance. This yielded a model with performance metrics of 0.76, 61.43, and 7441. The results of this model cannot be directly compared with ANN1 due to the fact that they use different input data sizes due to the lack of using the control slab in ANN2. Thus, the data from the control slab was removed in ANN3 which yielded a model with testing performance metrics of 0.74, 62.86, and 7670. Comparing ANN2 and ANN3, the use of normalized parameters showed to improve overall model performance with a slight increase in R^2 and decrease in RMSE, although percent capture did slightly decrease. This suggests that normalizing the image parameters by removing the trends in the control slab increased the overall trend of the parameters for the slabs that received clogging and thus increased model performance. Although normalizing the parameters were shown to highlight the change of the image specifically from adding sediment, the use of normalized parameters as inputs to an empirical model means that all future inputs to that model to create predictions need to also be normalized. If images of field permeable pavement systems were to be evaluated, a model that uses normalized parameters would not be effective. Thus, to create a robust model that uses raw image data, it should take into consideration all possible variables that may influence the raw image data or the model

should be trained specifically for different types of scenarios such as lighting condition, types of pavement surface, and different clogging sediments.

The models were seen to be adequate in predicting SIR but there was some unexplained variance. When using the prewet time as an input to the model, the performance of the models increased drastically. This is due to the high correlation between the prewet time and the SIR. Although the image parameters were seen to change as the pavement surfaces were clogged, the variances in the initial SIR of the slabs are not captured using image data. The slabs were constructed on the same day with the same mix of concrete, but had different initial SIRs. This highlights the importance of proper construction procedures to ensure adequate hydraulic performance of these pavement systems. The use of prewet time as an input variable is not ideal since it still requires on-site measurements to be taken, but would still save time, effort and water since it would not require the full SIR test to be conducted. To account for the differences in starting SIRs, models may need to be calibrated by conducting an initial SIR test which would provide insight to the model that is not captured with image data.

A sensitivity analysis was conducted where inputs were given to ANN1 to evaluate how the predictions change with different values of input parameters. Each input parameter changed according to observed values of that parameter while holding the other input parameters to high or low values seen in the dataset. The results showed that not only is the mean of the grayscale image an important parameter as pavements clogged, but the shape of the histogram including the variance and skewness also affect model predictions. Skewness, in general, was shown to have the largest sensitivity and that a higher image skewness yielded lower SIRs. Larger variances (more spread histograms) yielded larger SIRs. Images of slabs that yielded lower SIRs, in general, were whiter (larger mean), more compact (smaller variance), and had a larger positive skewness.

7.4 Future Research and Recommendations

One of the major limitations of this research is that the data collected was based on measurements obtained under controlled laboratory conditions. The models were trained and tested on laboratory data but to further assess the robustness of the models,

field data should also be tested. However, using grayscale image parameters of field data may prove to be more complicated due to the addition of several variables such as different lighting conditions, pavement types, sediment types, and condition of the surface. Thus, a model should be created that is trained on field data while also accounting for these variables. This may be done by creating a large dataset that includes images captured for different weather conditions, slab types, and sediment loading types.

Surface moisture conditions of the pavement may be one variable that was not accounted for during the laboratory analysis. Different moisture conditions of the pavement caused earlier images of the slabs to be brighter. It is recommended that a standardized approach is used when capturing images. This includes camera type, position of the camera, the use of flash, as well as the moisture condition of the slab. Images should be captured when the cement-aggregate particles are either completely dry or saturated with water. This issue may not be prevalent in other permeable pavement systems such as permeable asphalt but should also be tested. Permeable asphalt pavements may have yielded more robust models due to the uniformity of the material and the difference between pavement and sediment may have been more prevalent.

Grayscale image parameters such as the mean, variance, and skewness were used as inputs to the ANN models. More robust models should include the full RGB image as inputs to predict the clogging status of pavement slabs. CNNs are one type of model that may be able to achieve this (APPENDIX D gives a brief literature review on CNNs). Although CNNs are capable in terms of using full images as inputs, their ability in regression-based tasks has not yet been widely studied in the literature, thus, a classification approach may be more suitable.

Target SIRs to conduct maintenance procedures need to be identified and standardized. Future research could use these target SIRs as benchmarks to predict the performance of permeable pavement systems as a classification problem (eg., clogged or not clogged) which may simplify the modelling procedure. Lastly, clogging sediments with different colour and compositions should be evaluated.

REFERENCES

Abbot, J., & Marohasy, J. (2014). Input selection and optimisation for monthly rainfall forecasting in Queensland, Australia, using artificial neural networks. *Atmospheric Research*, *138*, 166-178.

ACI Committee 522. (2010). 522R-10: Report on Pervious Concrete. In N. Neithalath & (Eds.). N.p.: American Concrete Institute.

Agatonovic-Kustrin, S., & Beresford, R. (2000). Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. *Journal of pharmaceutical and biomedical analysis*, *22*(5), 717-727.

Ahiablame, L. M., Engel, B. A., & Chaubey, I. (2012). Effectiveness of low impact development practices: literature review and suggestions for future research. *Water, Air, & Soil Pollution*, *223*(7), 4253-4273.

Allen Cooley Jr, L., & Ray Brown, E. (2000). Selection and evaluation of field permeability device for asphalt pavements. *Transportation Research Record*, 1723(1), 73-82.

ASTM International. (2001). C29/C29M-09 *Standard Test Method for Specific Gravity and Absorption of Coarse Aggregate*; West Conshohocken, PA, USA.

ASTM International. (2009). C29/C29M-09 *Standard Test Method for Bulk Density ("Unit Weight") and Voids in Aggregate*; West Conshohocken, PA, USA.

ASTM International. (2012). C1754/C1754M-12 *Standard Test Method for Density and Void Content of Hardened Pervious Concrete*; West Conshohocken, PA, USA.

ASTM International. (2014). C39/C39M-14 *Standard Test Method for Compressive Strength of Cylindrical Concrete Specimens*; West Conshohocken, PA, USA.

ASTM International. (2017). C1701/C 1701M-09, *Standard Test Method for Infiltration Rate of in Place Pervious Concrete*; West Conshohocken, PA, USA.

ASTM International. (2018) C1781 / C1781M-18e1, *Standard Test Method for Surface Infiltration Rate of Permeable Unit Pavement Systems*; West Conshohocken, PA, USA.

Ayodele, T. O. (2010). *New Advances in Machine Learning* (pp. 19-46). Rijeka: IntechOpen.

Bean, E. Z., Hunt, W. F., & Bidelspach, D. A. (2007). Field survey of permeable pavement surface infiltration rates. *Journal of Irrigation and Drainage Engineering*, 133(3), 249-255.

Berry, B. J. (2008). Urbanization. In *Urban ecology* (pp. 25-48). Springer, Boston, MA.

Bhutta, M. A. R., Tsuruta, K., & Mirza, J. (2012). Evaluation of high-performance porous concrete properties*. Construction and Building Materials*, 31, 67-73.

Boogaard, F., Lucke, T., Van de Giesen, N., & Van de Ven, F. (2014). Evaluating the infiltration performance of eight Dutch permeable pavements using a new full-scale infiltration testing method. *Water*, 6(7), 2070-2083.

Booth, D. B., & Jackson, C. R. (1997). Urbanization of aquatic systems: Degradation thresholds, stormwater detection, and the limits of mitigation 1. JAWRA Journal of the American Water Resources Association, 33(5), 1077-1090.

Brattebo, B. O., & Booth, D. B. (2003). Long-term stormwater quantity and quality performance of permeable pavement systems. *Water research*, 37(18), 4369-4376.

Brown, R. A., & Borst, M. (2014). Evaluation of surface infiltration testing procedures in permeable pavement systems. Journal of Environmental Engineering, 140(3), 04014001.

Carleo, G., Cirac, I., Cranmer, K., Daudet, L., Schuld, M., Tishby, N., ... & Zdeborová, L. (2019). Machine learning and the physical sciences. Reviews of Modern Physics, 91(4), 045002.

Cettner, A., Ashley, R., Hedström, A., & Viklander, M. (2014). Sustainable development and urban stormwater practice. *Urban Water Journal*, *11*(3), 185-197.

Cha, Y. J., Choi, W., & Büyüköztürk, O. (2017). Deep learning-based crack damage detection using convolutional neural networks. *Computer-Aided Civil and Infrastructure Engineering*, 32(5), 361-378.

Cha, Y. J., Choi, W., Suh, G., Mahmoudkhani, S., & Büyüköztürk, O. (2018). Autonomous structural visual inspection using region-based deep learning for detecting multiple damage types. *Computer-Aided Civil and Infrastructure Engineering*, *33*(9), 731-747.

Chandrappa, A. K., & Biligiri, K. P. (2016). Comprehensive investigation of permeability characteristics of pervious concrete: A hydrodynamic approach. *Construction and Building Materials*, 123, 627-637.

Chen, Y., Wang, K., Wang, X., & Zhou, W. (2013). Strength, fracture and fatigue of pervious concrete. *Construction and Building Materials*, 42, 97-104.

City of Toronto. (2003). *Toronto Wet Weather Flow Management Master Plan*.

City of Toronto. (2006). *Toronto Wet Weather Flow Management Guidelines*.

Costa, F. B. P., Lorenzi, A., Haselbach, L., & Silva Filho, L. C. P. (2018). Best practices for pervious concrete mix design and laboratory tests. *Revista IBRACON de Estruturas e Materiais*, 11(5), 1151-1159.

Crookes, A. J., Drake, J. A., & Green, M. (2017). Hydrologic and quality control performance of zero-exfiltration pervious concrete pavement in Ontario. *Journal of Sustainable Water in the Built Environment*, 3(3), 06017001.

Damodaram, C., Giacomoni, M. H., Prakash Khedun, C., Holmes, H., Ryan, A., Saour, W., & Zechman, E. M. (2010). Simulation of combined best management practices

and low impact development for sustainable stormwater management 1. *JAWRA Journal of the American Water Resources Association*, *46*(5), 907-918.

Davis, A. P., Shokouhian, M., Sharma, H., & Minami, C. (2001). Laboratory study of biological retention for urban stormwater management. Water Environment Research, 73(1), 5-14.

Davis, A. P. (2008). Field performance of bioretention: Hydrology impacts. Journal of Hydrologic Engineering, 13(2), 90-95.

Davis, A. P., Hunt, W. F., Traver, R. G., & Clar, M. (2009). Bioretention technology: Overview of current practice and future needs. *Journal of environmental engineering*, *135*(3), 109-117.

Devi, G. K., Ganasri, B. P., & Dwarakish, G. S. (2015). A review on hydrological models. *Aquatic Procedia*, *4*, 1001-1007.

Dhakal, K. P., & Chevalier, L. R. (2017). Managing urban stormwater for urban sustainability: Barriers and policy solutions for green infrastructure application. Journal of environmental management, 203, 171-181.

Dierkes, C., Kuhlmann, L., Kandasamy, J., & Angelis, G. (2002). Pollution retention capability and maintenance of permeable pavements. In *Global Solutions for Urban Drainage* (pp. 1-13).

Dietz, M. E. (2007). Low impact development practices: A review of current research and recommendations for future directions. *Water, air, and soil pollution*, 186(1-4), 351-363.

Dreelin, E. A., Fowler, L., & Carroll, C. R. (2006). A test of porous pavement effectiveness on clay soils during natural storm events*. Water Research*, 40(4), 799-805.

Dung, C. V., Anh, L. D., (2019). Autonomous concrete crack detection using deep fully convolutional neural network. *Automation in Construction*, *99*, 52-58.

Eckart, K., McPhee, Z., & Bolisetti, T. (2017). Performance and implementation of low impact development–A review. *Science of the Total Environment*, *607*, 413-432.

Fassman, E. A., & Blackbourn, S. (2010). Urban runoff mitigation by a permeable pavement system over impermeable soils. *Journal of Hydrologic Engineering*, 15(6), 475-485.

Fletcher, T.D.; Shuster, W.; Hunt, W.F.; Ashley, R.; Butler, D.; Arthur, S.; Trowsdale, S.; Barraud, S.; Semadeni-Davies, A.; Bertrand-Krajewski, J.-L., Mikkelsen, P.S., Rivard, G., Uhl, M., Dagenais, D., & Viklander, M. (2015). SUDS, LID, BMPs, WSUD and more—The Evolution and Application of Terminology Surrounding Urban Drainage. *Urban Water J*. 12, 525–542.

Fwa, T. F., Lim, E., & Tan, K. H. (2015). Comparison of permeability and clogging characteristics of porous asphalt and pervious concrete pavement materials. *Transportation Research Record*, 2511(1), 72-80.

Fwa, T. F., Tan, S. A., Chuai, C. T., & Guwe, Y. K. (2001). Expedient permeability measurement for porous pavement surface. *International Journal of Pavement Engineering*, 2(4), 259-270.

Gardner, M. W., & Dorling, S. R. (1998). Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. *Atmospheric environment*, 32(14-15), 2627-2636.

Gao, Y., & Mosalam, K. M. (2018). Deep transfer learning for image-based structural damage recognition. *Computer-Aided Civil and Infrastructure Engineering*, 33(9), 748-768.

Goncalves, C., & Seters, T. V. (2012). Characterization of Particle Size Distributions of Runoff from High Impervious Urban Catchments in the Greater Toronto Area. Toronto, Canada: Toronto and Region Conservation Authority.

Gopalakrishnan, K., Khaitan, S. K., Choudhary, A., & Agrawal, A. (2017). Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection. *Construction and Building Materials*, *157*, 322-330.

Government of Canada. (2021, March 25). In Canadian Climate Normals 1981-2010 Station Data. Retrieved from https://climate.weather.gc.ca/climate_normals/results_1981_2010_e.html?stnID=5051&autofwd=1

Haselbach, L. M., Valavala, S., & Montes, F. (2006). Permeability predictions for sand-clogged Portland cement pervious concrete pavement systems. *Journal of environmental management*, *81*(1), 42-49.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).

Hoong, J. D. L. H., Lux, J., Mahieux, P. Y., Turcry, P., & Aït-Mokhtar, A. (2020). Determination of the composition of recycled aggregates using a deep learning-based image analysis. *Automation in Construction*, *116*, 103204.

Hoskere, V., Narazaki, Y., Hoang, T., & Spencer Jr, B. (2018). Vision-based structural inspection using multiscale deep convolutional neural networks*:1805.01055*.

Hu, M., Zhang, X., Li, Y., Yang, H., & Tanaka, K. (2019). Flood mitigation performance of low impact development technologies under different storms for retrofitting an urbanized area. *Journal of Cleaner Production*, 222, 373-380.

Huang, J., Valeo, C., He, J., & Chu, A. (2016). Three types of permeable pavements in cold climates: hydraulic and environmental performance. *Journal of Environmental Engineering*, 142(6), 04016025.

Huyan, J. (2019). *Development of Machine Learning Based Analytical Tools for Pavement Performance Assessment and Crack Detection* (Doctoral dissertation, UWSpace).

Jiang, J., Zhang, Z., Dong, Q., & Ni, F. (2018). Characterization and identification of asphalt mixtures based on Convolutional Neural Network methods using X-ray scanning images. *Construction and Building Materials*, *174*, 72-80.

Jørgensen, S. E., & Bendoricchio, G. (2001). Fundamentals of ecological modelling (Vol. 21). *Elsevier*.

Joshaghani, A., Ramezanianpour, A. A., Ataei, O., & Golroo, A. (2015). Optimizing pervious concrete pavement mixture design by using the Taguchi method. *Construction and Building Materials*, *101*, 317-325.

Kayalibay, B., Jensen, G., & van der Smagt, P. (2017). CNN-based segmentation of medical imaging data. 1701.03056.

Kayhanian, M., Anderson, D., Harvey, J. T., Jones, D., & Muhunthan, B. (2012). Permeability measurement and scan imaging to assess clogging of pervious concrete pavements in parking lots. *Journal of Environmental management*. 95(1), 114-123

Kaykhosravi, S., Abogadil, K., Khan, U. T., & Jadidi, M. A. (2019). The Low-Impact Development Demand Index: A New Approach to Identifying Locations for LID. *Water*, 11(11), 2341.

Khan, U. T., Valeo, C., Chu, A., & van Duin, B. (2012). Bioretention cell efficacy in cold climates: Part 1 — hydrologic performance. *Canadian Journal of Civil Engineering*, 39(11), 1210–1221.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097-1105).

Kumar, K., Kozak, J., Hundal, L., Cox, A., Zhang, H., & Granato, T. (2016). In-situ infiltration performance of different permeable pavements in an employee used parking lot–A four-year study. *Journal of Environmental Management*, 167, 8-14.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, *521*(7553), 436-444.

LeCun, Y., Boser, B. E., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W. E., & Jackel, L. D. (1990). Handwritten digit recognition with a back-propagation network. In *Advances in neural information processing systems* (pp. 396-404).

Legret, M., & Pagotto, C. (1999). Evaluation of pollutant loadings in the runoff waters from a major rural highway. *Science of the Total Environment*, 235(1-3), 143-150.

Liang, X. (2019). Image-based post-disaster inspection of reinforced concrete bridge systems using deep learning with Bayesian optimization. *Computer-Aided Civil and Infrastructure Engineering*, 34(5), 415-430.

Li, J., Deng, C., Li, Y., Li, Y., & Song, J. (2017). Comprehensive benefit evaluation system for low-impact development of urban stormwater management measures. *Water Resources Management*, *31*(15), 4745-4758.

Liu, J., Li, H., Wang, Y., & Zhang, H. (2020). Integrated life cycle assessment of permeable pavement: Model development and case study. *Transportation Research Part D: Transport and Environment*, 85, 102381.

Lucke, T., & Beecham, S. (2011). Field investigation of clogging in a permeable pavement system. *Building Research & Information*, 39(6), 603-615.

Lucke, T., Boogaard, F., & van de Ven, F. (2014). Evaluation of a new experimental test procedure to more accurately determine the surface infiltration rate of permeable pavement systems. *Urban, Planning and Transport Research*, 2(1), 22-35.

Lucke, T., White, R., Nichols, P., & Borgwardt, S. (2015). A simple field test to evaluate the maintenance requirements of permeable interlocking concrete pavements. *Water*, 7(6), 2542-2554.

Mekanik, F., Imteaz, M. A., Gato-Trinidad, S., & Elmahdi, A. (2013). Multiple regression and Artificial Neural Network for long-term rainfall forecasting using large scale climate modes. *Journal of Hydrology*, *503*, 11-21.

Meng, Z., Hu, Y., & Ancey, C. (2020). Using a Data Driven Approach to Predict Waves Generated by Gravity Driven Mass Flows. *Water*, *12*(2), 600.

Moon, H. G., & Kim, J. H. (2011). Intelligent crack detecting algorithm on the concrete crack image using neural network. *Proceedings of the 28th ISARC*, 1461-1467.

Mosavi, A., Ozturk, P., & Chau, K. W. (2018). Flood prediction using machine learning models: Literature review. *Water*, *10*(11), 1536.

Motulsky, H. J., & Ransnas, L. A. (1987). Fitting curves to data using nonlinear regression: a practical and nonmathematical review*. The FASEB journal*, 1(5), 365-374.

Nassiri, S., & AlShareedah, O. (2017). Preliminary procedure for structural design of pervious concrete pavements (No. WA-RD 868.2). Washington (State). Dept. of Transportation. Research Office.

Neithalath, N., Sumanasooriya, M. S., & Deo, O. (2010). Characterizing pore volume, sizes, and connectivity in pervious concretes for permeability prediction. *Materials characterization*, *61*(8), 802-813.

Nhat-Duc, H., Nguyen, Q. L., & Tran, V. D. (2018). Automatic recognition of asphalt pavement cracks using metaheuristic optimized edge detection algorithms and convolution neural network. *Automation in Construction*, *94*, 203-213.

Niemczynowicz, J. (1999). Urban hydrology and water management–present and future challenges. *Urban water*, *1*(1), 1-14.

Niu, Z., Zhou, M., Wang, L., Gao, X., & Hua, G. (2016). Ordinal regression with multiple output cnn for age estimation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4920-4928).

Obaidat, M. T., Ghuzlan, K. A., & Alawneh, M. M. (2017). Analysis of volumetric properties of bituminous mixtures using cellular phones and image processing techniques. *Canadian Journal of Civil Engineering*, *44*(9), 715-726.

Obla, K. H. (2010). Pervious concrete—An overview. *Indian Concrete Journal*, 84(8), 9.

Ontario Provincial Standard Specification (2018). Construction Specification for Pervious Concrete Pavement for Low-Volume Traffic Applications. (OPSS.MUNI #356).

Packman, J. C., & Kidd, C. H. R. (1980). A logical approach to the design storm concept. *Water Resources Research*, *16*(6), 994-1000.

Podolsky, L. (2012). Barriers to Low Impact Development. Sacramento, CA: Local Government Commission.

Pratt, C. (1995). “A review of source control of urban stormwater runoff.” *Water Environ. Manage*., 9(2), 132–139.

Pratt, C. J., Mantle, J. D. G., and Schofield, P. A. (1995). UK research into the performance of permeable pavement, reservoir structures in controlling stormwater discharge quantity and quality. *Water Sci. Technol*., 32(1), 63–69.

Qin, H. P., Li, Z. X., & Fu, G. (2013). The effects of low impact development on urban flooding under different rainfall characteristics. *Journal of environmental management*, *129*, 577-585.

Radfar, A., & Rockaway, T. D. (2016). Clogging prediction of permeable pavement. *Journal of Irrigation and Drainage Engineering*, 142(4), 04015069.

Rawat, W., & Wang, Z. (2017). Deep convolutional neural networks for image classification: A comprehensive review*. Neural computation*, 29(9), 2352-2449.

Razzaghmanesh, M., & Beecham, S. (2018). A review of permeable pavement clogging investigations and recommended maintenance regimes. *Water*, 10(3), 337.

Razzaghmanesh, M., & Borst, M. (2018). Investigation clogging dynamic of permeable pavement systems using embedded sensors. *Journal of hydrology*, 557, 887-896.

Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. *In Advances in neural information processing systems* (pp. 91-99).

Roehr, D., & Kong, Y. (2010). Runoff reduction effects of green roofs in Vancouver, BC, Kelowna, BC, and Shanghai, PR China. *Canadian Water Resources Journal*, 35(1), 53-68.

Roy, A. H., Wenger, S. J., Fletcher, T. D., Walsh, C. J., Ladson, A. R., Shuster, W. D., ... & Brown, R. R. (2008). Impediments and solutions to sustainable, watershed-scale

urban stormwater management: lessons from Australia and the United States. *Environmental management*, *42*(2), 344-359.

Rushton, B. T. (2001). Low-impact parking lot design reduces runoff and pollutant loads. *Journal of Water Resources Planning and Management*, 127(3), 172-179.

Samarasinghe, S. (2016). *Neural networks for applied sciences and engineering: from fundamentals to complex pattern recognition*. Crc Press.

Schaefer, V. R., & Kevern, J. T. (2011). An integrated study of pervious concrete mixture design for wearing course applications.

Sehgal, K., Drake, J., Seters, T. V., & Vander Linden, W. K. (2018). Improving restorative maintenance practices for mature permeable interlocking concrete pavements. *Water*, 10(11), 1588.

Selbig, W. R., & Bannerman, R. T. (2011). Characterizing the Size Distribution of Particles in Urban Stormwater by Use of Fixed-Point Sample-Collection Methods. Reston, VA: US Geological Survey.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition*.* 1409.1556.

Snieder, E., Shakir, R., & Khan, U. T. (2020). A comprehensive comparison of four input variable selection methods for artificial neural network flow forecasting models. *Journal of Hydrology*, *583*, 124299.

Sonebi, M., Bassuoni, M., & Yahia, A. (2016). Pervious concrete: mix design, properties and applications. *RILEM Technical Letters*, 1, 109-115.

Suriya, S., & Mudgal, B. V. (2012). Impact of urbanization on flooding: The Thirusoolam sub watershed–A case study. *Journal of hydrology*, *412*, 210-219.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. In Proceedings of the *IEEE conference on computer vision and pattern recognition* (pp. 1-9).

Tennis, P. D., Leming, M. L., & Akers, D. J. (2004). Pervious concrete pavements (No. PCA Serial No. 2828). Skokie, IL: Portland Cement Association.

Tong, B., Schaefer, V., Wang, K. (2011). Clogging effects of portland cement pervious concrete. Graduate Theses and Dissertations Iowa State University. 12048.

Tota-Maharaj K, Scholz M. (2010). Efficiency of permeable pavement systems for the removal of urban runoff pollutants under varying environmental conditions*. Environ Prog Sustain Energy*, 29(3):358–369.

Tota-Maharaj, K., Grabowiecki, P., Babatunde, A., & Coupe, S. J. (2012, May). The performance and effectiveness of geotextiles within permeable pavements for treating concentrated stormwater. *In Sixteenth International Water Technology Conference*, *IWTC*. Vol. 16.

TRCA Toronto and Region Conservation Authority. *Performance evaluation of permeable pavement and a bioretention swale*. Seneca College, King City, Ontario. Toronto and Region Conservation Authority Interim Report #2; 2006.

Urban Water Resources Centre. (2013). *Water Sensitive Urban Design: Basic Procedures for 'Source Control' of Stormwater*. In J. R. Argue &. (Eds.). N.p.: University of South Australia.

US EPA. (1999). *Stormwater technology fact sheet: Wet detention ponds* (# EPA 832-F-99-048). Washington, DC: Environmental Protection Agency, Office of Water.

US EPA. (2000). *Low impact development (LID): A literature review* (# EPA-841-B-00-005). Washington, DC: Environmental Protection Agency, Office of Water.

Valeo, C., & Gupta, R. (2018). Determining surface infiltration rate of permeable pavements with digital imaging. *Water*, 10(2), 133.

Veeraragavan, R. K., Kottayi, N. M., Mallick, R. B., Nirala, M. K., & Sarkar, S. (2018, July). Application of artificial intelligence to predict moisture damage of hot-mix asphalt mixes. In *Proceedings of the Institution of Civil Engineers-Transport* (pp. 1-10).

Wang, L., Lyons, J., Kanehl, P., & Bannerman, R. (2001). Impacts of urbanization on stream habitat and fish across multiple spatial scales. Environmental management, 28(2), 255-266.

Winston, R. J., Al-Rubaei, A. M., Blecken, G. T., & Hunt, W. F. (2016). A simple infiltration test for determination of permeable pavement maintenance needs. *Journal of Environmental Engineering*, 142(10), 06016005.

Winston, R. J., Al-Rubaei, A. M., Blecken, G. T., Viklander, M., & Hunt, W. F. (2016). Maintenance measures for preservation and recovery of permeable pavement surface infiltration rate–The effects of street sweeping, vacuum cleaning, high pressure washing, and milling. *Journal of Environmental Management*, 169, 132-144.

Winter, J., and Duthie, H. (1998). "Effects of urbanization on water quality, periphyton and invertebrate communities in a southern Ontario stream." *Can. Water Resour. J*., 23(3), 245–257.

Welker, A. L., Barbis, J. D., & Jeffers, P. A. (2012). A Side-by-Side Comparison of Pervious Concrete and Porous Asphalt 1. *JAWRA Journal of the American Water Resources Association*, 48(4), 809-819.

Wu, B., Iandola, F., Jin, P. H., & Keutzer, K. (2017). Squeezedet: Unified, small, low power fully convolutional neural networks for real-time object detection for autonomous driving. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (pp. 129-137).

Xu, G., Shen, W., Huo, X., Yang, Z., Wang, J., Zhang, W., & Ji, X. (2018). Investigation on the properties of porous concrete as road base material. *Construction and Building Materials*, 158, 141-148.

Yao, A., Ding, H., Zhang, X., Hu, Z., Hao, R., & Yang, T. (2018). Optimum design and performance of porous concrete for heavy-load traffic pavement in cold and heavy rainfall region of NE China. *Advances in Materials Science and Engineering, 2018*.

Zador, A. M. (2019). A critique of pure learning and what artificial neural networks can learn from animal brains. *Nature communications*, *10*(1), 1-7.

Zeiler, M. D., & Fergus, R. (2014, September). Visualizing and understanding convolutional networks. In *European conference on computer vision* (pp. 818-833). Springer, Cham.

Zhang, L., Yang, F., Zhang, Y. D., & Zhu, Y. J. (2016, September). Road crack detection using deep convolutional neural network. In *2016 IEEE international conference on image processing (ICIP)* (pp. 3708-3712). IEEE.

APPENDICES

APPENDIX A: PRELIMINARY FIELD INVESTIGATION AND DATA ANALYSIS

A.1 Methodology

A.1.1 Preliminary Field Investigation

Valeo & Gupta (2018) conducted a brief regression analysis to investigate the relationship between the mean gray level of an image of the surface of a permeable pavement and the acquired infiltration rate of the pavement. Results from the study indicate that clogged pavements result in images with larger mean gray levels, or overall whiter images. We investigated this relationship further by conducting a preliminary field investigation of permeable pavement parking lots located on and near York University's Keele Campus.

The preliminary investigation took place on York University's Keele Campus at the Osgoode Hall Law School parking lot as well as a parking lot at the Kortright Centre for Conservation which is a 20-minute drive North-West from York University. The Osgoode parking lot is made of permeable asphalt and has a drainage system leading to Stong Pond. **Figure 50** shows the site location at York University with the specific locations that were chosen for conducting the tests. The Kortright Centre had a permeable concrete pavement. Similarly, locations at the Kortright parking lot were chosen; however, the pavement there was severely clogged and only a handful of tests were able to be conducted due time constraints. For both sites, the locations for conducting tests were chosen strategically based on Borst and Brown (2014) as opposed to randomly. The chosen locations were suspected of being susceptible to clogging. For example, locations with higher levels of vehicle traffic such as in parking spots and locations where water collected during rainfall events were chosen. A total of 25 data points were collected between the two sites. Some points were later discarded due to suspicions of human error when conducting the surface infiltration test.

Due to extreme clogging in various locations of the parking lots, infiltration tests were time-consuming. Due to this and project time constraints, multiple tests were conducted where only the prewetting was done. Some tests were conducted using 3.6

+/- 0.05 kg [8.0 +/- 0.1 lb] of water which is the amount specified in ASTM C1701. For some tests, a lower mass of 1.8 +/- 0.05 kg [4.0 +/- 0.1 lb] of water was used and only one test was conducted. The standard states that tests that took longer than one hour to infiltrate should be omitted. Some areas were so clogged that the water for prewetting the pavement took over an hour to infiltrate. For these locations, infiltration rates were reported using the pre-wetting amount.

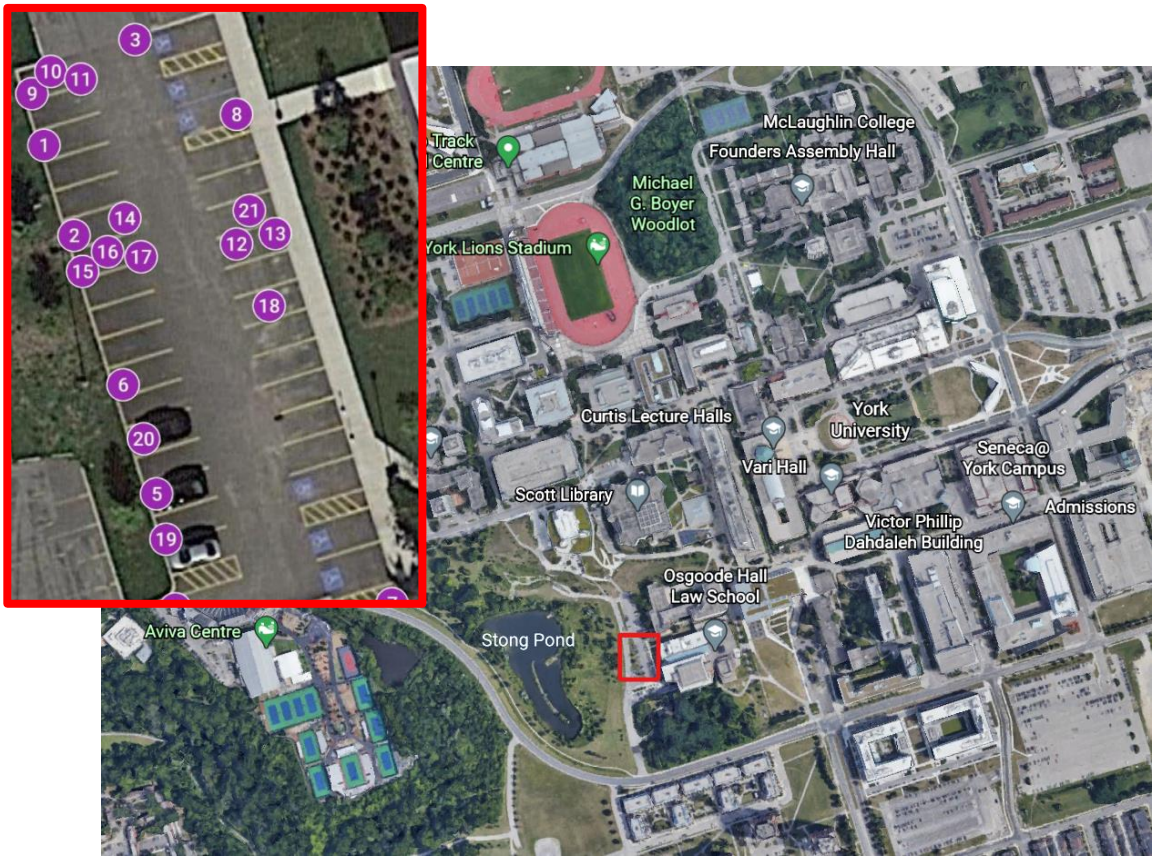


Figure 50: Location of Preliminary Investigation outlined in red; Insert: Expanded view of the parking lot with experimental locations

A.1.2 Data Analysis Techniques

MATLAB was used to extract properties of the images acquired during the field investigation. The images were first converted to greyscale and their individual histograms of each grey level was acquired. In greyscale images, black pixels are represented as 0 whereas white pixels are represented as 255. In between are different shades of grey with increasing whiteness as the grey level increases to 255. The mean

grey level was extracted from these images and plotted with surface infiltration rate to understand their correlation. The mean grey level was also used as an input variable to a preliminary linear regression model to investigate the potential of using these images to predict clogging. A data split of 60% was used where 60% of the data was used for training and 40% was used for testing the model.

It was difficult to obtain an adequate dataset due to the parking lot being severely clogged due to many of the collected data points representing the pavement when it was clogged. Thus, it was difficult to obtain a dataset that represented the full lifespan of the pavement system, from freshly cast to fully clogged. This informed the decision to create laboratory-scale specimens where permeable concrete systems were constructed with the intention to slowly subject them to clogging while simultaneously acquiring image and SIR data. Due to COVID-19, access to the laboratory was restricted for extended periods of time. Consequently, MATLAB was used to generate artificial pavement images with specified mean grey levels. **Figure 51** shows samples of the generated images where the darker image represents unclogged pavements. The field investigation did not yield a large enough dataset that ANNs require for accurate predictions to take place. Instead, these generated ‘pavement’ images were used as inputs to linear regression and artificial neural network models to build a proof of concept, model calibration, and compare the potential of both type of models.

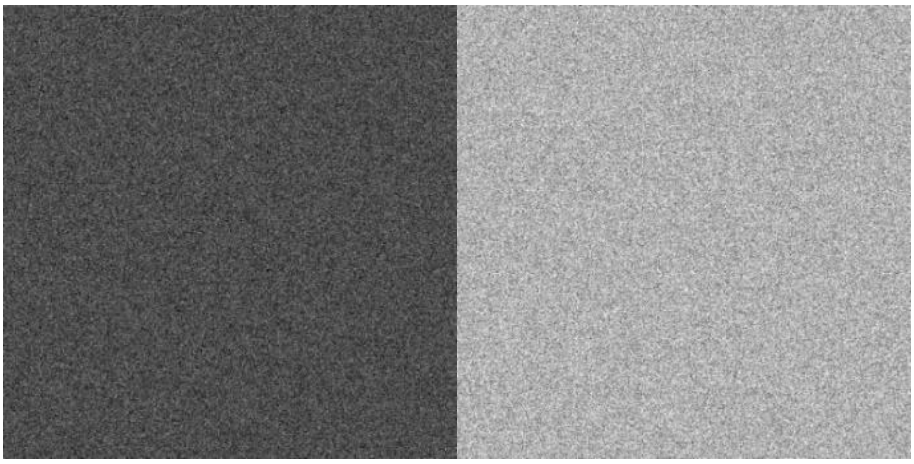


Figure 51: Generated ‘pavement’ images where the left image represents an unclogged pavement with high infiltration and the right represents one that is clogged with low infiltration

A.2 Results and Discussion

A.2.1 Preliminary Field Investigation

The mean grey level was extracted from the images and was plotted with their associated SIR to investigate their relationship. **Figure 52** shows a plot of the inverse mean grey level with the associated infiltration rate. The mean grey level is suspected to be inversely proportional to the infiltration rate, meaning the larger the mean grey level, or the whiter the image, the smaller the pavement's infiltration rate. Thus, the inverse of the mean grey level was used to maintain the hypothesized positive relationship, meaning the larger the inverse mean grey level, or the blacker the image, the higher the infiltration rate. The plot shows data collected from pervious asphalt and pervious concrete pavements. As seen in the figure, many locations that were tested were subject to high levels of clogging, thus yielding infiltration rates close to 0 mm/min. This means that many 'pervious' pavements have similar infiltration capacity to ordinary asphalt pavement where most stormwater events results in runoff. The correlation between the inverse mean grey level and the SIR resulted in a value of -0.1259. This means that the data is not very well correlated and the relationship between the inverse mean grey level and the SIR is very weak with this dataset.

Variability in collecting field data may have affected the results. Some sources of this variability include inconsistencies in lighting condition when taking photos and variations in conducting infiltration tests. This calls for conducting a similar lab-scale experiment where these variables can be controlled and the pavement system is slowly clogged while simultaneously conducting infiltration tests and capturing images. This method should provide the dataset necessary to thoroughly investigate the relationship between the infiltration rate of the pavement and the properties of the image.

A simple linear regression was conducted with the field images as seen in **Figure 53**. Only the pervious asphalt data points were used and some outliers were omitted due to suspicions of leakage during the SIR test. The R^2 of the model was 0.55. Due to the high number of clogged observations and low number of observations in general, the model does not provide adequate predictive power. If the hypothesis that as the pavements get clogged, the overall image shifts to a whiter image is to be thoroughly

investigated, a larger, full-range dataset that captures the full lifecycle of the pavement as it clogs is required.

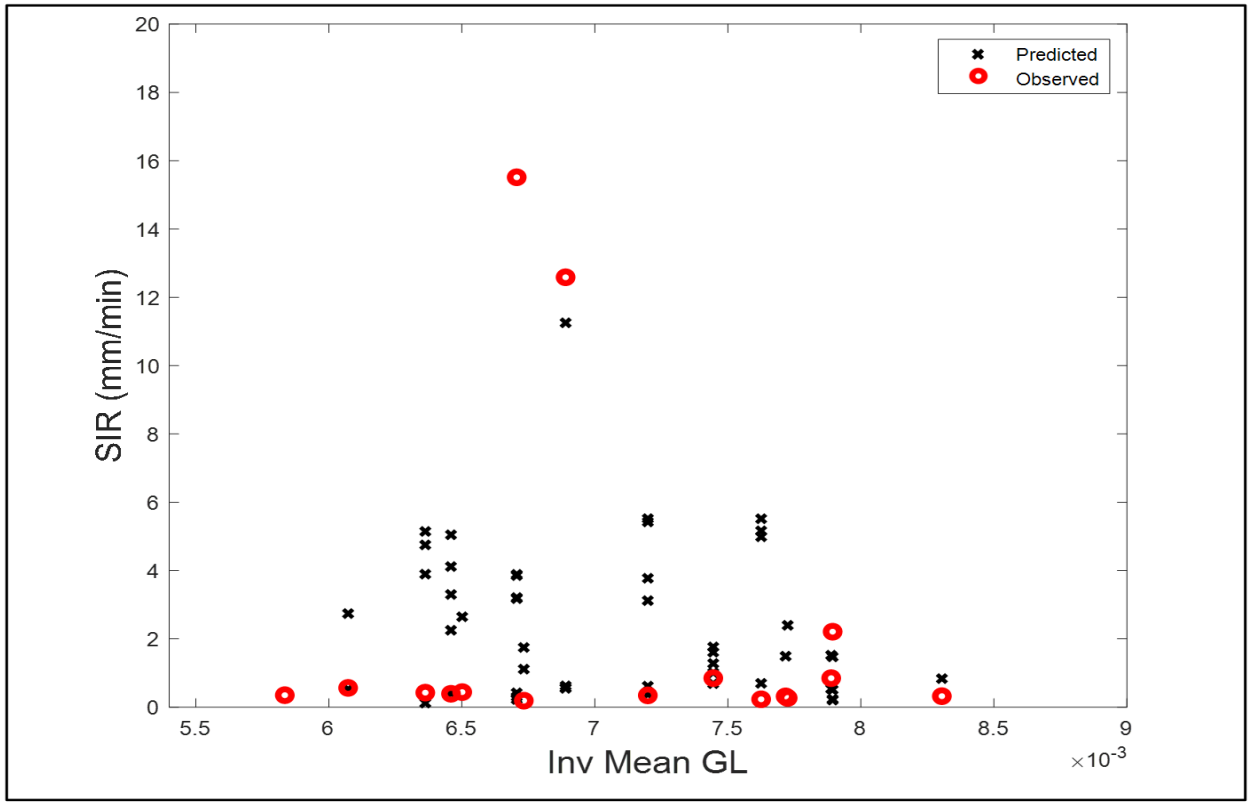
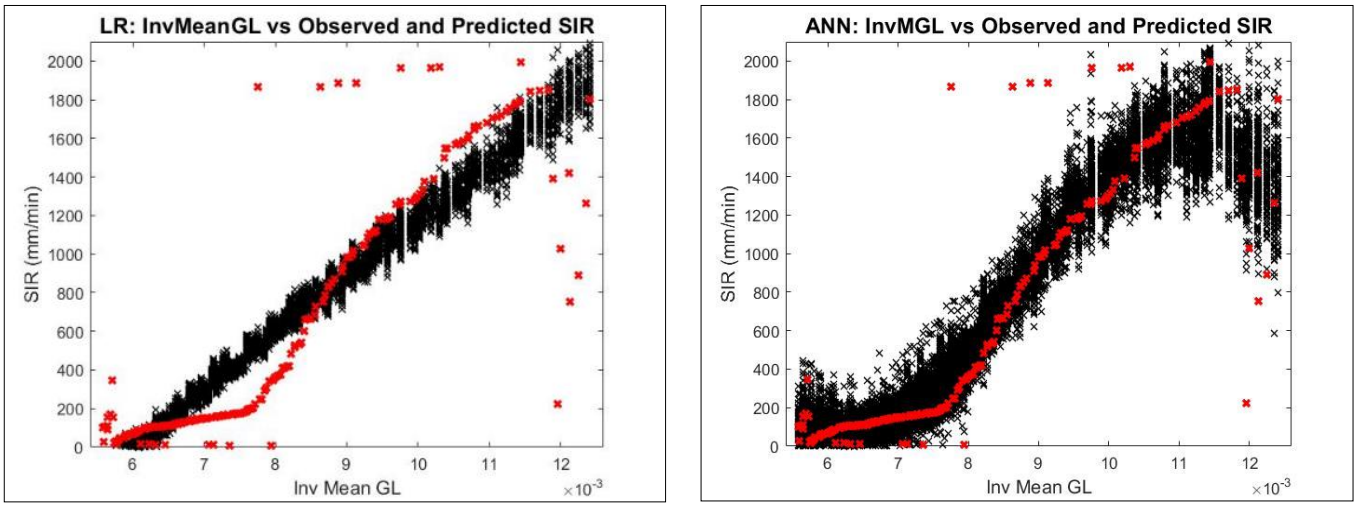


Figure 53: Inverse Mean Grey Level vs. Observed SIR where observed data was obtained through the preliminary field investigation. The predicted points were obtained through linear regression

A.2.2 Artificial Pavement Images

An artificial dataset was used where 'pavement' images were generated using MATLAB to investigate the relationship between the properties of the image and theoretical SIRs. The properties (inverse mean grey level, standard deviation, and skewness) of these images were used as inputs to calibrate linear regression and artificial neural network (ANN) models. The ANN used a training-validation-testing data split of 55/22.5/22.5, nine neurons, an early stopping criterion of 6, the Levenberg-Marquardt training function, and a positive-linear transfer function. **Figure 54** highlights the predictive power of ANNs as they are able to capture the non-linearity of the data set as well as more of the extreme data points.



(a) (b)

Figure 54: Model results where black 'X' are the predicted values from the ensemble (100 runs) and red are the observed values. Where (a) shows the linear regression results and (b) shows the artificial neural network results

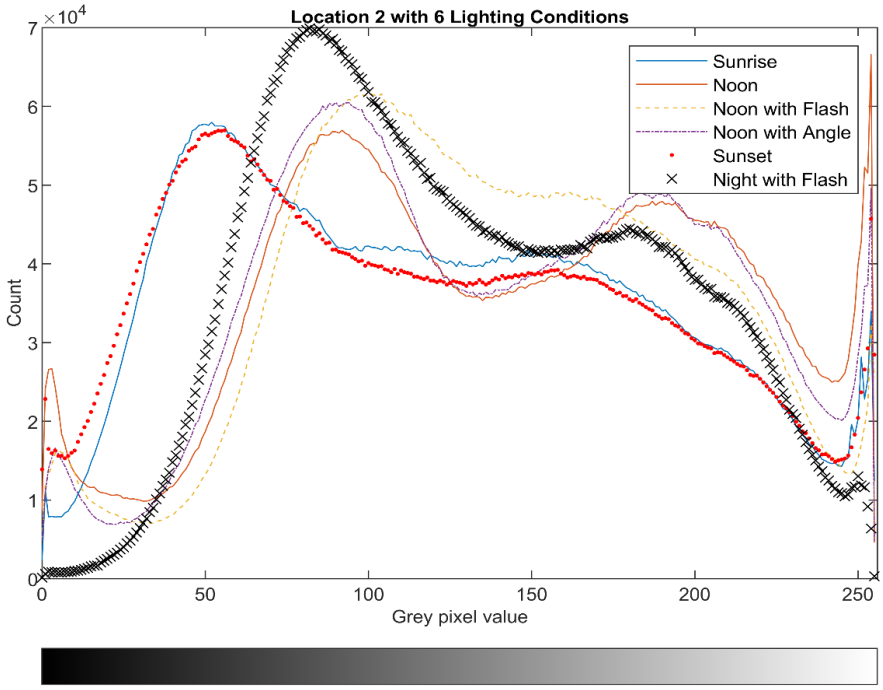
A.3 Conclusion

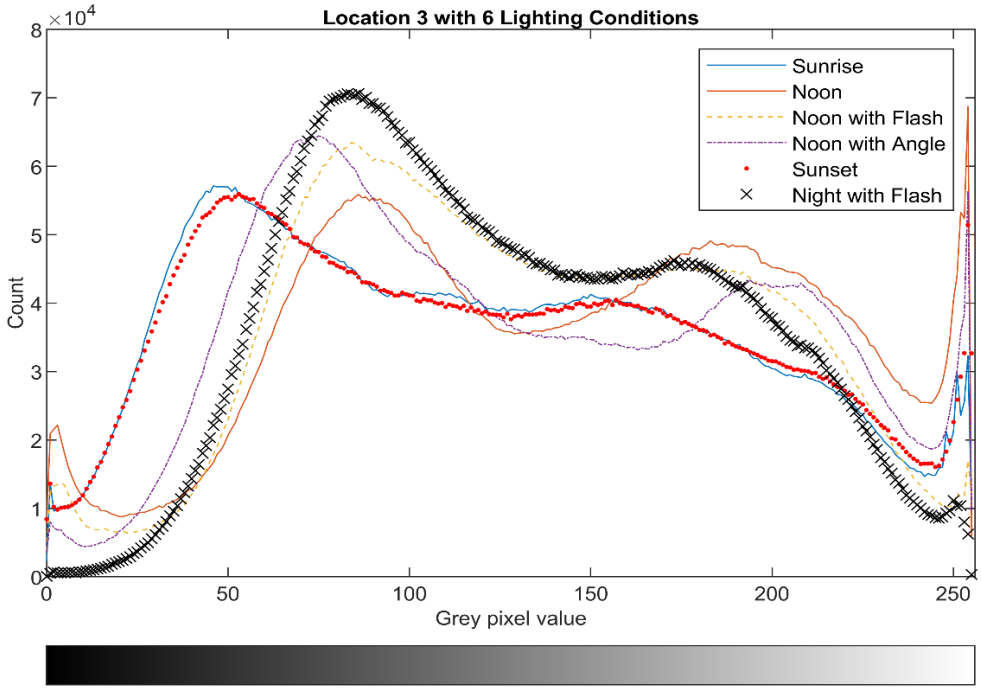
The preliminary field investigation revealed that many locations of different parking lots containing permeable pavements were already subject to high levels of clogging. Due to this, obtaining a dataset of images and their associated infiltration rates proved difficult. Moreover, these locations required a large amount of time for water to infiltrate which added more logistical challenges in obtaining a large dataset due to time constraints and further reinforces the fact that there is a need of a method to quickly and easily obtain the performance of these pavement systems. Although some images demonstrated the hypothesized relationship between the mean grey level and infiltration rate, the data collected in this dataset cannot be used to statistically confirm it. This is due to the low quantity of collected data points, the low number of conducted tests that were able to actually infiltrate water, and variability in conducting field tests. The variability in the dataset includes differences in lighting conditions, the angle at which the phone was held when taking photographs of the pavement, and human errors in conducting infiltration tests such as improper sealing of the infiltration ring or inconsistencies of pouring water for different tests. These challenges were propagated in the linear regression model as

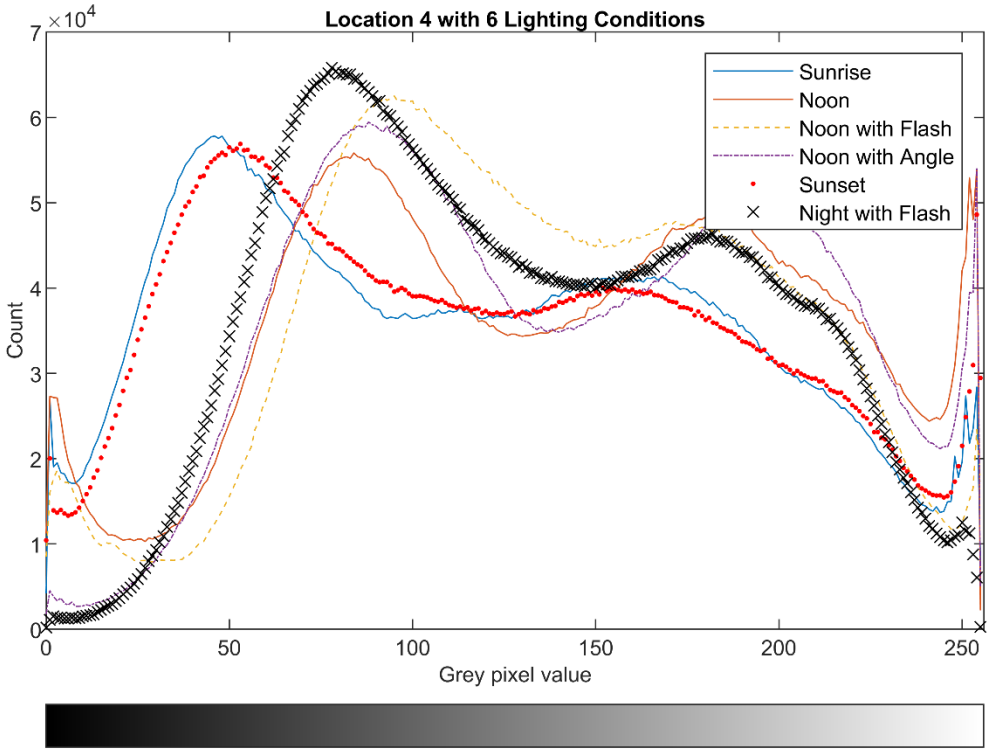
inadequate data resulted in an inaccurate model. The R2 of the linear regression model was 0.55. Due to COVID-19, lab work was put on hold and artificial data was generated using MATLAB to calibrate models. This dataset was input to linear regression and artificial neural network models which resulted in R2 values of 0.78 and 0.86, respectively. The ANN was found to produce a more accurate predictive model where it was able to capture the non-linearity in the dataset as well as predict more extreme points. The field investigation did not yield conclusive results between the relationship of the properties of the image of the pavement and its associated infiltration rate. A controlled lab investigation where pervious concrete slabs were slowly clogged as their images and infiltration rate were obtained was conducted to eliminate sources of error and to obtain a full dataset that encapsulated the entire infiltration lifespan of the pavement slab.

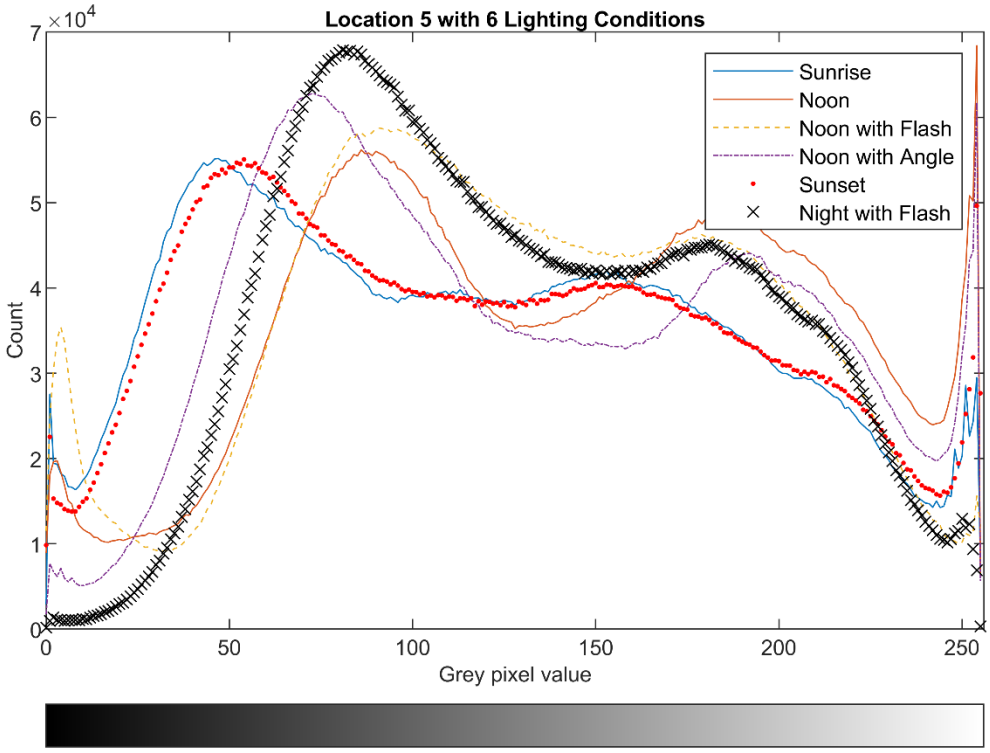
APPENDIX B: FIELD INVESTIGATION OF DIFFERENT LIGHTING CONDITIONS

B.1 Plots of Different Locations and Lighting Conditions






Location 4 with 6 Lighting Conditions
×10^4
Count
Grey pixel value
Sunrise
Noon
Noon with Flash
Noon with Angle
Sunset
Night with Flash


Location 5 with 6 Lighting Conditions
×10^4
Count
Grey pixel value
Sunrise
Noon
Noon with Flash
Noon with Angle
Sunset
Night with Flash

B.2 KSDT Tables of Different Locations with Single Lighting Condition

Testing Location/Condition	IMG ID	P-Values from KSDT for 'Noon'				
Location 1 Noon	1	1				
Location 2 Noon	4	0.8284	1			
Location 3 Noon	7	0.969	0.7596	1		
Location 4 Noon	10	0.6858	0.6858	0.401	1	
Location 5 Noon	13	0.969	0.9883	0.6105	0.7596	1
		1	4	7	10	13

Testing Location/Condition	IMG ID	P-Values from KSDT for 'Noon w/ F'				
Location 1 Noon with Flash	31	1				
Location 2 Noon with Flash	34	0.0033	1			
Location 3 Noon with Flash	37	0.5367	0.0046	1		
Location 4 Noon with Flash	40	0.2882	0.0258	0.015	1	
Location 5 Noon with Flash	43	0.1645	0.0062	0.1086	0.1342	1
		31	34	37	40	43

Testing Location/Condition	IMG ID	P-Values from KSDT for 'Noon w Angle'				
Location 1 Noon with Angle	16	1				
Location 2 Noon with Angle	19	0.0033	1			
Location 3 Noon with Angle	22	0.015	0.0197	1		
Location 4 Noon with Angle	25	0.0033	0.1086	0.015	1	
Location 5 Noon with Angle	28	0.0334	0.0062	0.5367	0.0113	1
		16	19	22	25	28

Testing Location/Condition	IMG ID	P-Values from KSDT for ''Sunset'				
Location 1 Sunset	46	1				
Location 2 Sunset	49	1.36E-04	1			
Location 3 Sunset	52	1.12E-05	0.0548	1		
Location 4 Sunset	55	0.0033	0.6858	0.1342	1	
Location 5 Sunset	58	1.12E-05	0.1086	0.969	0.1645	1
		46	49	52	55	58

Testing Location/Condition	IMG ID	P-Values from KSDT for 'Night w F'				
Location 1 Night with Flash	61	1				
Location 2 Night with Flash	64	0.401	1			
Location 3 Night with Flash	67	0.0084	0.0113	1		
Location 4 Night with Flash	70	0.6105	0.401	0.0548	1	
Location 5 Night with Flash	73	0.2411	0.8284	0.0548	0.2411	1
		61	64	67	70	73

APPENDIX C: IMAGE AND LABORATORY ANALYSIS

C.1 Plots of Extracted Image Parameters for Flash On Dataset

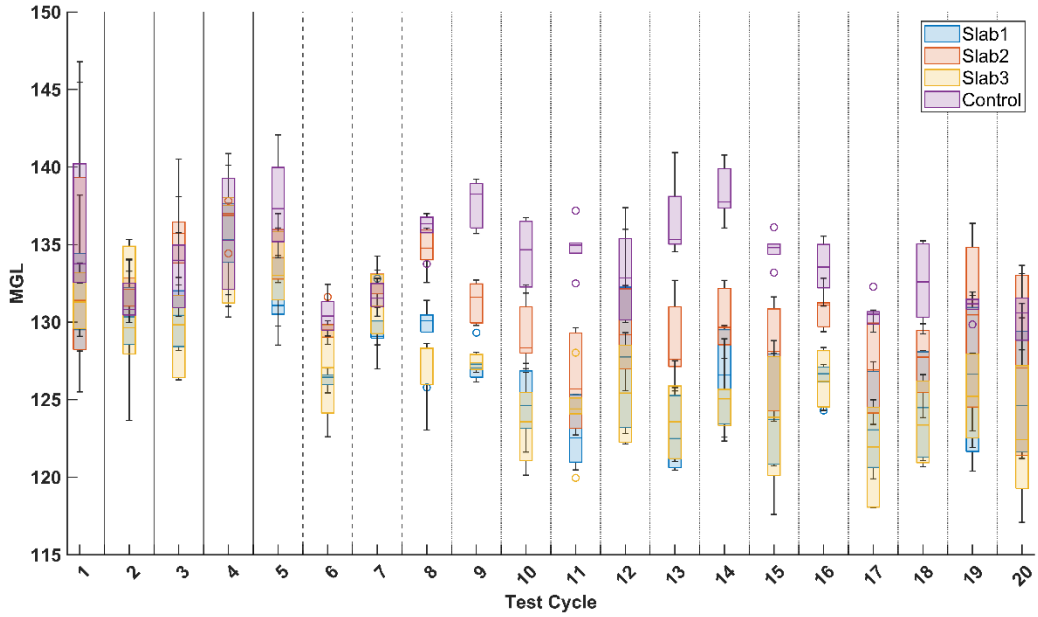


Figure 55: MGL of images for PC slabs as clogging progresses (Flash On Dataset)

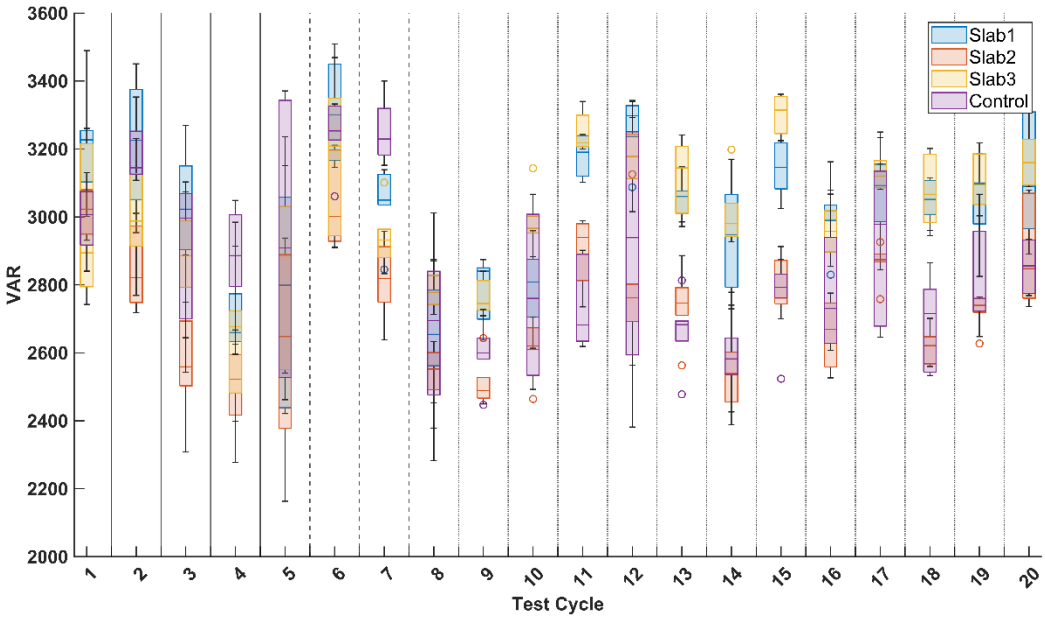


Figure 56: Variance of images for PC slabs as clogging progresses (Flash On Dataset)

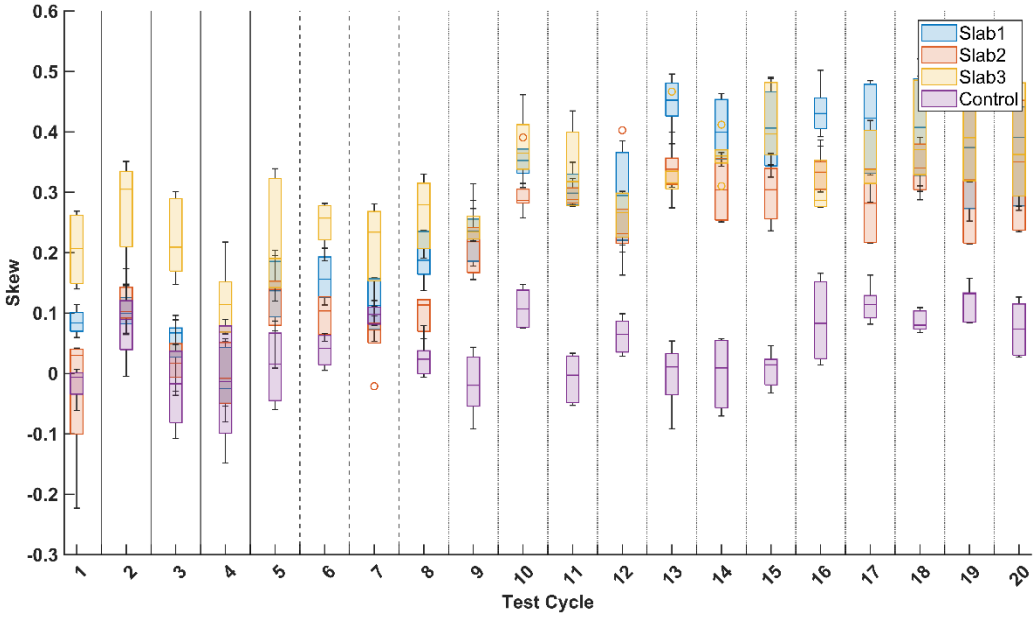


Figure 57: Skewness of images for permeable concrete slabs as clogging progresses (Flash On Dataset)

C.2 Plots of Extracted Image Parameters for Normalized Dataset

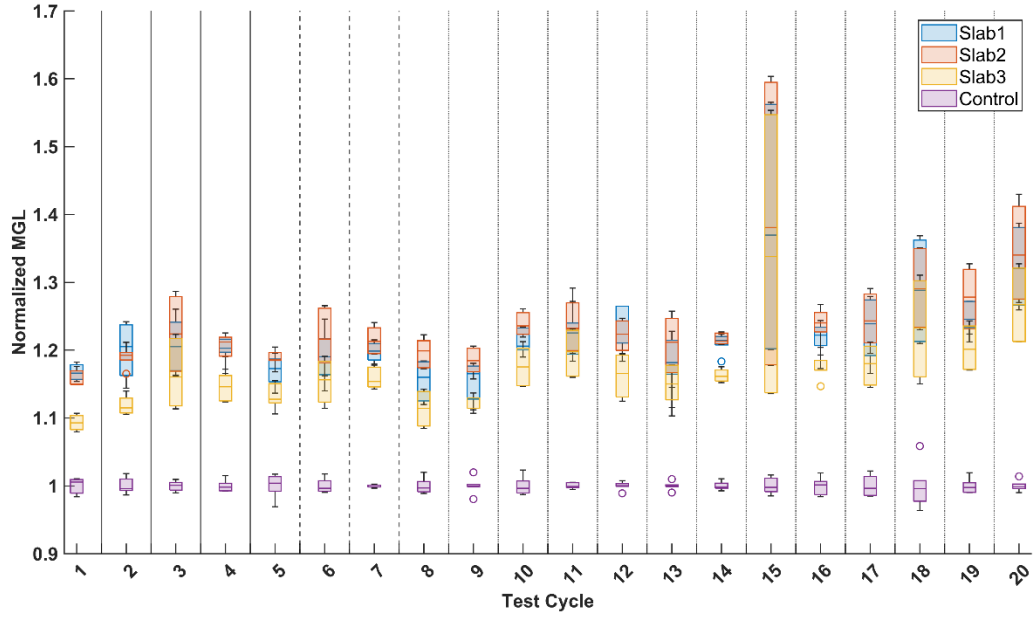


Figure 58: Normalized MGL of images for permeable concrete slabs as clogging progresses (Flash Off Dataset)

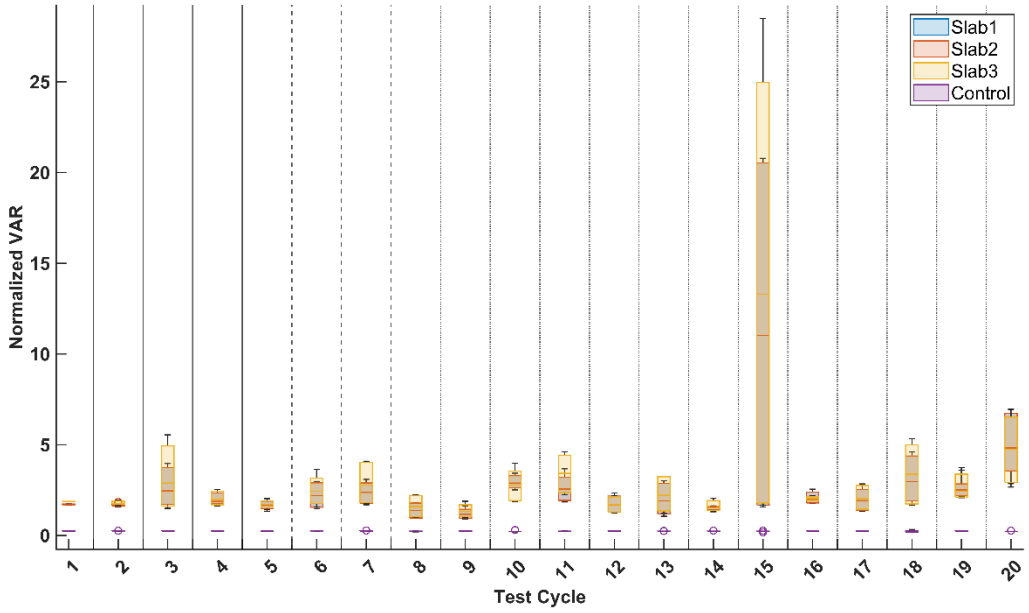


Figure 59: Normalized variance of images for permeable concrete slabs as clogging progresses (Flash Off Dataset)

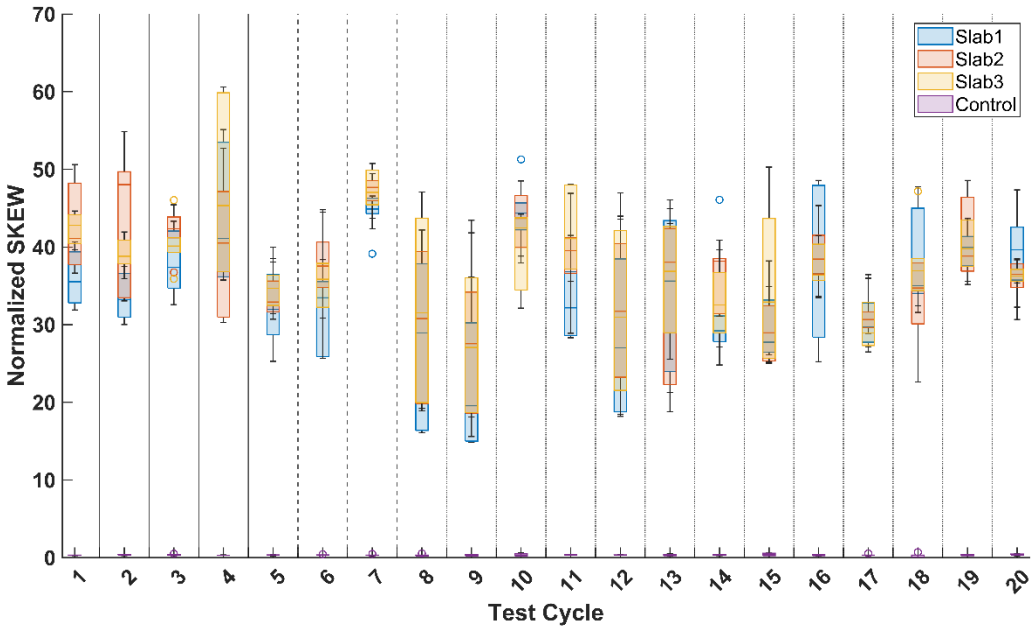


Figure 60: Normalized skewness of images for permeable concrete slabs as clogging progresses (Flash Off Dataset)

C.3 Plots of Extracted Image Parameters for Updated Normalized Dataset

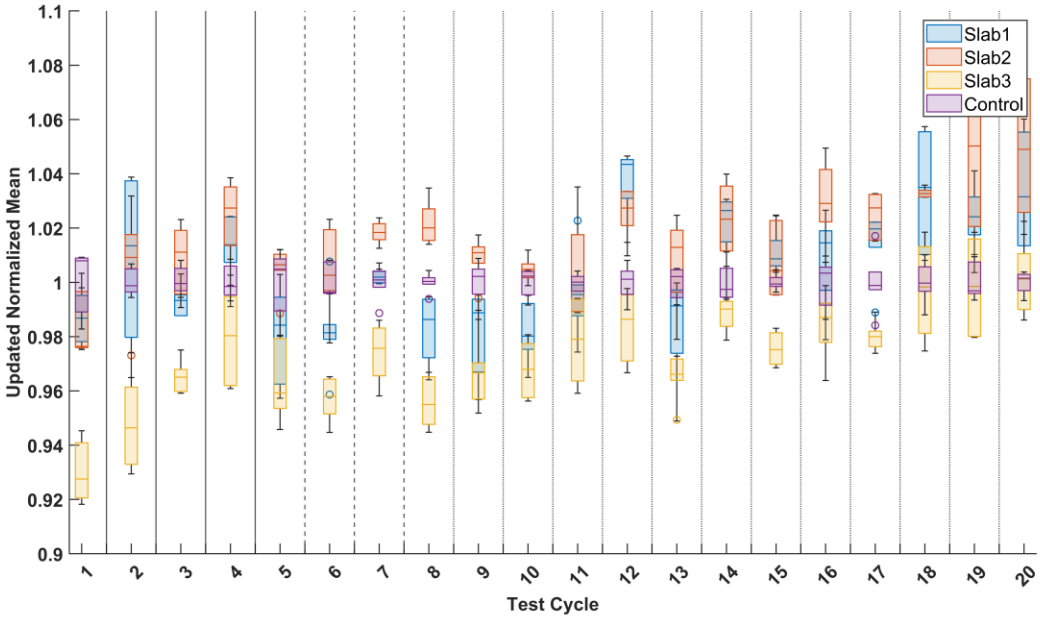


Figure 61: Updated normalized mean of images for permeable concrete slabs as clogging progresses (Flash Off Dataset)

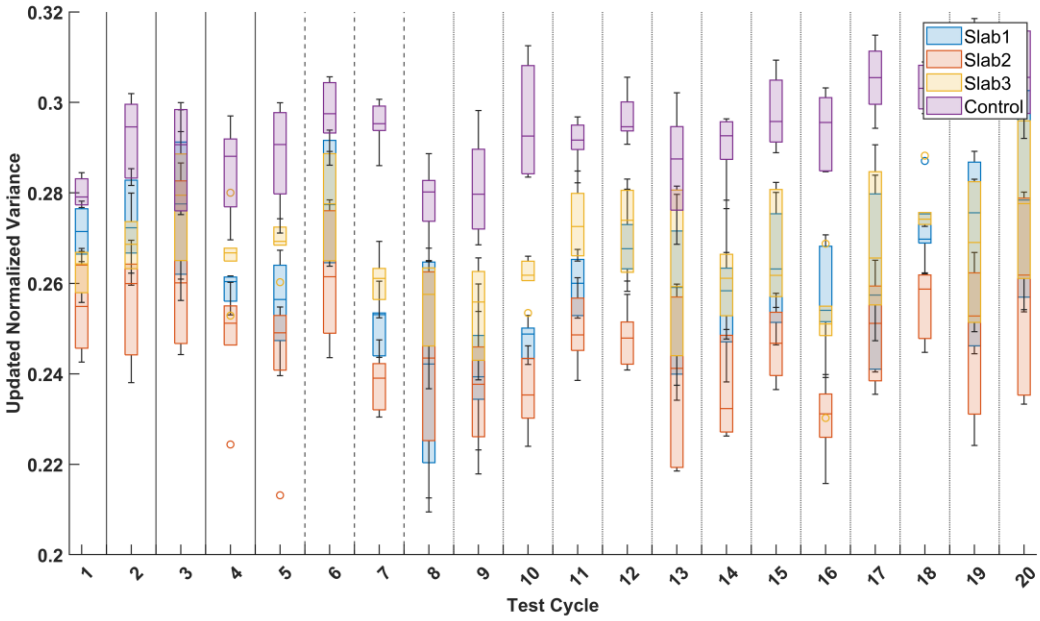


Figure 62: Updated normalized variance of images for permeable concrete slabs as clogging progresses (Flash Off Dataset)

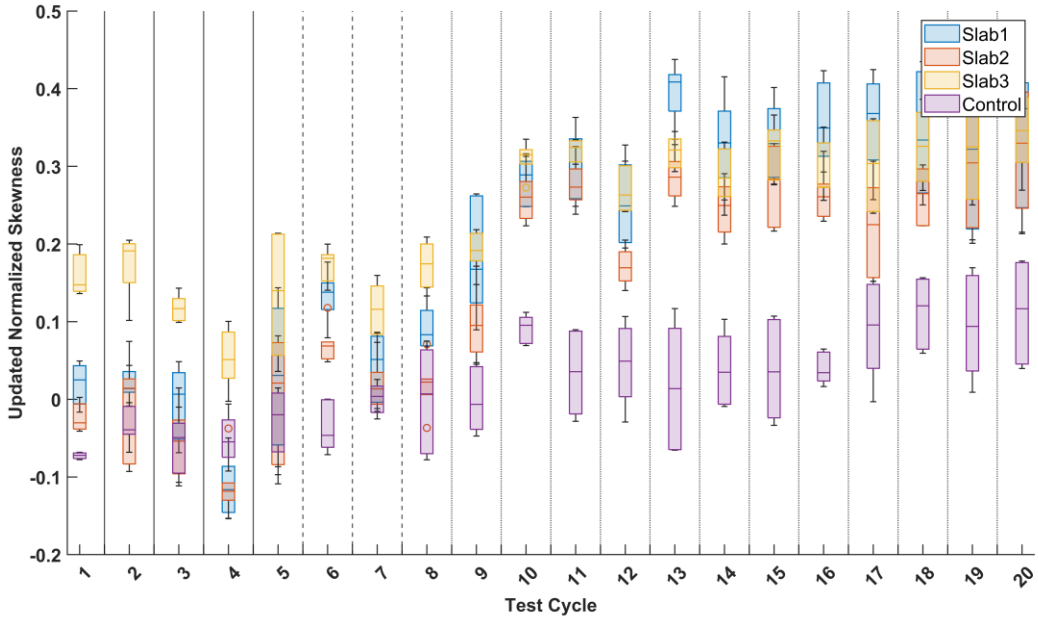


Figure 63: Updated normalized skewness of images for permeable concrete slabs as clogging progresses (Flash Off Dataset)

C.4 Infiltration Data Collecting During Laboratory Analysis

Date	Slab	Prewet Time (s)	Infiltration Test Time (s)	SIR (mm/hr)
21-Jun	1	5.62	18.62	51968.49
	2	7.16	30.25	31988.54
	3	5.58	20.69	46769.13
	Control	5.87	22.65	42722.00
	1	5.29	17.66	54793.50
	2	7.08	29.42	32891.00
	3	5.56	21.28	45472.43
	Control	6.17	22.5	43006.81
	1	5.31	16.52	58574.65
	2	7.2	27.8	34807.67
	3	5.49	20.08	48189.90
	Control	4.96	23.13	41835.42
22-Jun	1	5.98	17.83	54271.07
	2	6.44	26.52	36487.68
	3	7.16	24.19	40002.20
	Control	6.99	26.07	37117.50
	1	6.32	18.08	53520.65
	2	6.6	26.69	36255.27
	3	6.33	21.58	44840.28
	Control	6	22.82	42403.74

	1	4.83	16.81	57564.14
	2	6.49	26.27	36834.92
	3	5.95	21.59	44819.51
	Control	6.4	22.73	42571.64
25-Jun	1	5.25	19.65	49244.44
	2	7.45	30.92	31295.38
	3	6.65	21.99	44004.24
	Control	6.2	23.35	41441.25
	1	5.62	19	50929.12
	2	6.93	30.68	31540.20
	3	5.27	21.89	44205.27
	Control	6.2	23.99	40335.69
	1	5.21	17.99	53788.40
	2	8.06	31.4	30816.98
	3	5.93	21.72	44551.26
	Control	5.86	24.17	40035.30
05-Jul	1	6.06	18.81	51443.55
	2	7.35	31.46	30758.21
	3	5.46	22.35	43295.45
	Control	5.73	21.79	44408.14
	1	4.8	18.05	53609.60
	2	7.14	30.12	32126.60
	3	5.49	21.42	45175.22
	Control	5.79	22.33	43334.23
	1	4.61	18.97	51009.66
	2	7.19	30.64	31581.37
	3	5.79	21.66	44674.67
	Control	5.77	22.16	43666.66
07-Jul	1	4.67	18.08	53520.65
	2	8.67	32.55	29728.21
	3	5.65	24.07	40201.63
	Control	5.37	23.03	42017.08
	1	4.99	18.59	52052.35
	2	7.67	32.55	29728.21
	3	5.7	22.16	43666.66
	Control	5.96	22.63	42759.76
	1	4.97	20.05	48262.01
	2	7.8	33.01	29313.94
	3	4.83	23.27	41583.72
	Control	5.64	23.13	41835.42
08-Jul	1	4.29	19.4	49879.03
	2	7.59	31.89	30343.47

	3	5.27	21.85	44286.19
	Control	5.42	23.59	41019.64
	1	4.33	18.88	51252.82
	2	7.23	31.36	30856.29
	3	5.27	21.77	44448.93
	Control	6	24.01	40302.09
	1	4.6	19.66	49219.39
	2	7.46	32.23	30023.37
	3	4.9	23.02	42035.33
	Control	5.62	23.95	40403.06
09-Jul	1	4.72	19.15	50530.20
	2	8.21	34.05	28418.60
	3	5.15	22.93	42200.32
	Control	5.84	24.37	39706.74
	1	4.55	19.62	49319.74
	2	8.06	33.36	29006.39
	3	5.33	21.94	44104.52
	Control	5.64	22.76	42515.52
	1	4.9	19.91	48601.37
	2	7.94	33.56	28833.53
	3	5.3	22.88	42292.54
	Control	5.3	23.3	41530.18
14-Jul	1	4.37	19.15	50530.20
	2	7.69	34.57	27991.13
	3	5.28	22.26	43470.50
	Control	5.36	24.27	39870.34
	1	4.74	21.33	45365.84
	2	8.29	33.45	28928.35
	3	5.25	21.1	45860.34
	Control	5.68	23.43	41299.76
	1	4.06	19.73	49044.77
	2	7.63	33.93	28519.11
	3	4.87	21.81	44367.41
	Control	5.59	23.22	41673.27
19-Jul	1	4.66	20.13	48070.21
	2	8.02	36.51	26503.79
	3	5.67	23.67	40881.00
	Control	5.43	22.98	42108.50
	1	4.52	19.9	48625.79
	2	8.48	36.56	26467.54
	3	6.17	23.36	41423.51
	Control	5.7	23.36	41423.51

	1	4.46	19.93	48552.60
	2	8.16	35.19	27497.96
	3	5.35	22.78	42478.19
	Control	5.23	23.24	41637.40
20-Jul	1	4.6	20.11	48118.01
	2	7.96	33.8	28628.79
	3	5.24	23.57	41054.44
	Control	5.63	24.03	40268.55
	1	4.86	20.78	46566.57
	2	8.13	34.58	27983.03
	3	5.61	23.37	41405.79
	Control	5.68	23.7	40829.25
	1	4.47	20.38	47480.53
	2	7.99	34.45	28088.63
	3	5.37	23.05	41980.62
	Control	5.7	23.87	40538.47
22-Jul	1	5.02	21.02	46034.88
	2	8.5	36.8	26294.93
	3	6.67	26.4	36653.53
	Control	5.36	22.45	43102.60
	1	5.05	20.2	47903.63
	2	9.09	36.12	26789.96
	3	5.92	25.76	37564.18
	Control	5.52	23.22	41673.27
	1	4.87	20.19	47927.35
	2	8.1	34.98	27663.04
	3	5.66	25.2	38398.94
	Control	5.57	22.65	42722.00
23-Jul	1	4.75	20.53	47133.62
	2	7.6	35.13	27544.93
	3	5.06	22.2	43587.98
	Control	5.66	23.54	41106.77
	1	4.62	19.66	49219.39
	2	8.08	36.82	26280.64
	3	5.56	22.97	42126.83
	Control	6.01	23.59	41019.64
	1	5.17	19.83	48797.44
	2	9.09	36.25	26693.88
	3	5.23	22.79	42459.56
	Control	5.61	23.37	41405.79
28-Jul	1	5.15	26.56	36432.73
	2	9.84	40.86	23682.17

	3	6.23	25.22	38368.49
	Control	5.54	23.1	41889.75
	1	4.95	21.19	45665.56
	2	9.08	39.17	24703.94
	3	6.63	27.28	35471.16
	Control	5.47	23.05	41980.62
	1	5.24	21.35	45323.34
	2	9	39.07	24767.17
	3	6.73	26.52	36487.68
	Control	5.73	21.92	44144.77
03-Aug	1	5.02	21.11	45838.62
	2	9.92	44.6	21696.26
	3	6.92	28.8	33599.07
	Control	5.67	23.33	41476.78
	1	5.13	21.04	45991.12
	2	9.5	41.77	23166.23
	3	6.12	27.77	34845.27
	Control	5.84	22.93	42200.32
	1	4.69	20.43	47364.33
	2	9.41	40.87	23676.37
	3	6.31	27.49	35200.19
	Control	5.94	23.14	41817.34
05-Aug	1	5.07	22.26	43470.50
	2	10.55	41.82	23138.53
	3	7.22	29.69	32591.89
	Control	5.65	22.88	42292.54
	1	5.5	21.55	44902.70
	2	10.83	41.84	23127.47
	3	7.4	29.59	32702.04
	Control	5.69	23.64	40932.88
	1	5.11	22.25	43490.03
	2	11.44	43.87	22057.29
	3	7.3	29.39	32924.58
	Control	6.03	23.9	40487.58
05-Aug	1	5.3	23.5	41176.73
	2	10.64	42	23039.36
	3	6.78	26.9	35972.24
	Control	5.49	23.74	40760.46
	1	5.4	22.9	42255.60
	2	11.31	44.84	21580.14
	3	6.59	28.19	34326.12
	Control	6.09	23.87	40538.47

	1	5.19	22.76	42515.52
	2	11.67	46.04	21017.66
	3	6.59	26.98	35865.58
	Control	5.88	23.7	40829.25
06-Aug	1	6.39	26.03	37174.54
	2	14.97	56.03	17270.27
	3	8.91	34.77	27830.12
	Control	5.53	23.75	40743.30
	1	6.26	24.65	39255.71
	2	15.38	53.49	18090.36
	3	8.51	32.9	29411.95
	Control	5.83	24.97	38752.63
	1	5.75	24.94	38799.25
	2	13.92	53.87	17962.75
	3	7.96	33.04	29287.33
	Control	5.31	22.72	42590.37
06-Aug	1	6.07	23.08	41926.05
	2	13.34	49.29	19631.84
	3	8.51	33.97	28485.52
	Control	5.96	23.37	41405.79
	1	5.45	22.86	42329.54
	2	13.21	49.71	19465.97
	3	7.86	34.65	27926.50
	Control	5.59	23.76	40726.15
	1	5.41	23.2	41709.19
	2	13.07	48.6	19910.56
	3	8.49	34.94	27694.71
	Control	5.83	23.58	41037.03
12-Aug	1	6.5	30.44	31788.87
	2	18.04	67	14442.59
	3	9.57	36.9	26223.67
	Control	6.29	24.84	38955.45
	1	7.27	28.57	33869.56
	2	15.85	63.05	15347.40
	3	9.06	34.77	27830.12
	Control	5.73	23.85	40572.46
	1	6.73	28.16	34362.69
	2	14.97	63.6	15214.67
	3	8.48	35.14	27537.09
	Control	6.03	23.93	40436.83
12-Aug	1	7.82	29.63	32657.89
	2	16.87	61.35	15772.67

	3	9.99	36.74	26337.87
	Control	5.55	21.52	44965.30
	1	7.96	29.09	33264.12
	2	16.81	60.91	15886.61
	3	8.8	34.39	28137.63
	Control	6.05	21.64	44715.95
	1	6.75	27.4	35315.81
	2	17.32	62.2	15557.13
	3	9.22	34.54	28015.44
	Control	6.06	22.69	42646.68
13-Aug	1	10.03	35.95	26916.64
	2	24.21	74.86	12926.17
	3	12.3	41.75	23177.32
	Control	6.19	24.75	39097.10
	1	9.92	34.14	28343.68
	2	21.68	69.55	13913.06
	3	10.94	39.45	24528.60
	Control	5.79	24.89	38877.19
	1	9.38	33.08	29251.91
	2	18.31	68.63	14099.57
	3	10.66	37.29	25949.40
	Control	6.03	24.44	39593.01
13-Aug	1	9.45	33.96	28493.91
	2	19.16	69.48	13927.08
	3	10.61	37.01	26145.72
	Control	5.91	22.95	42163.54
	1	9.29	33.47	28911.06
	2	19.83	72.06	13428.44
	3	10.16	39.69	24380.28
	Control	5.78	24.92	38830.39
	1	8.95	34.31	28203.24
	2	19.74	68.85	14054.51
	3	10.27	38.23	25311.36
	Control	5.97	24.88	38892.82
16-Aug	1	13.77	42.85	22582.34
	2	28.26	86.64	11168.67
	3	16.85	48.14	20100.82
	Control	5.85	24.99	38721.62
	1	10.89	40.73	23757.75
	2	23.43	82.12	11783.41
	3	12.19	43.91	22037.20
	Control	6.16	24.31	39804.74

	1	10.09	39.98	24203.43
	2	20.7	78	12405.81
	3	10.74	42.93	22540.26
	Control	6.15	24.86	38924.11
16-Aug	1	11.13	37.46	25831.64
	2	19.66	73.22	13215.70
	3	11.01	42.04	23017.44
	Control	5.84	23.86	40555.46
	1	11.19	37.85	25565.48
	2	19.87	74.79	12938.27
	3	11.94	41.2	23486.73
	Control	6.13	25.07	38598.06
	1	9.1	36.39	26591.19
	2	19.92	74.49	12990.38
	3	10.29	40.7	23775.26
	Control	5.78	24.31	39804.74
18-Aug	1	10.97	47.37	20427.55
	2	28.9	94.35	10256.00
	3	15.32	49.73	19458.14
	Control	6.55	24	40318.89
	1	12.85	46.45	20832.15
	2	23.22	85.71	11289.85
	3	13.1	47.44	20397.41
	Control	6.47	25.25	38322.90
	1	12.26	44.85	21575.32
	2	23.5	90.03	10748.12
	3	11.74	44.94	21532.12
	Control	7.07	24.69	39192.11
18-Aug	1	13.4	47.1	20544.66
	2	22.92	86.89	11136.53
	3	12.09	46.15	20967.57
	Control	6.44	23.98	40352.51
	1	12.81	45.75	21150.89
	2	23.35	87.47	11062.69
	3	11.67	45.77	21141.65
	Control	6.2	25.63	37754.71
	1	11.12	44.59	21701.13
	2	21.85	86.24	11220.47
	3	11.86	43.65	22168.46
	Control	6.37	24.65	39255.71
19-Aug	1	13.76	49.75	19450.32
	2	26.94	98.34	9839.87

	3	14.21	50.92	19003.40
	Control	6.56	24.35	39739.35
	1	12.93	47.71	20281.98
	2	25.25	92.63	10446.43
	3	12.47	47.7	20286.23
	Control	6.23	24.36	39723.04
	1	12.24	48.27	20046.68
	2	25.12	91.64	10559.29
	3	12.02	47.37	20427.55
	Control	6.68	24.13	40101.67
19-Aug	1	12.86	47.8	20243.79
	2	72.06	113.06	8558.76
	3	12.18	49.67	19481.64
	Control	6.39	24.13	40101.67
	1	11.64	46.83	20663.11
	2	26.78	109.76	8816.08
	3	12.08	47.43	20401.71
	Control	6.66	23.96	40386.20
	1	12.41	46.99	20592.75
	2	25.73	111.96	8642.85
	3	12.17	49.08	19715.84
	Control	6.88	24.16	40051.87
20-Aug	1	14.67	59.6	16235.79
	2	40.49	151.26	6397.28
	3	14.76	50.58	19131.14
	Control	5.74	23.63	40950.20
	1	14.35	55.97	17288.78
	2	34.24	147.64	6554.14
	3	12.31	48.04	20142.66
	Control	6.34	23.56	41071.87
	1	12.61	56.78	17042.15
	2	32.96	139.56	6933.60
	3	12.32	47.7	20286.23
	Control	6.27	24.74	39112.90
20-Aug	1	13.48	53.19	18192.39
	2	32.14	128.67	7520.43
	3	12.02	45.87	21095.56
	Control	5.73	23.36	41423.51
	1	13.4	53.52	18080.22
	2	30.4	126.09	7674.31
	3	12.05	45.89	21086.36
	Control	6.53	23.92	40453.73

	1	12.99	52.89	18295.58
	2	28.96	125.23	7727.01
	3	11.93	45.12	21446.22
	Control	6.58	23.5	41176.73
26-Aug	1	15.62	64.35	15037.35
	2	36.9	40.94	4727.18
	3	17	66.94	14455.53
	Control	7.87	27.84	34757.66
	1	15.79	64.82	14928.31
	2	42.32	43.23	4476.77
	3	16.59	62.58	15462.66
	Control	6.6	26.36	36709.15
	1	16.76	63.82	15162.23
	2	39.36	39.2	4937.01
	3	19.26	64.44	15016.34
	Control	7.19	27.32	35419.23
27-Aug	1	15.34	60.57	15975.78
	2	38.93	37.19	5203.84
	3	13.37	56.51	17123.58
	Control	6.95	25.91	37346.71
	1	15.39	60.45	16007.50
	2	41.18	41.06	4713.36
	3	15.41	59.41	16287.72
	Control	7.51	27.16	35627.88
	1	16.53	59.45	16276.76
	2	39.53	39.08	4952.17
	3	15.67	56.28	17193.55
	Control	8.27	27.4	35315.81
30-Aug	1	20.41	71.15	13600.19
	2	44.46	44.87	4313.14
	3	18.1	68.19	14190.55
	Control	8.23	27.49	35200.19
	1	22.44	69.13	13997.59
	2	41.97	43.05	4495.49
	3	18.23	66.74	14498.85
	Control	8.39	26.64	36323.32
	1	21.22	71.15	13600.19
	2	42.69	41.65	4646.59
	3	20.02	66.04	14652.53
	Control	8.91	27.85	34745.18
30-Aug	1	18.54	59.99	16130.24
	2	48.82	47.8	4048.76

	3	17.62	62.09	15584.69
	Control	7.42	25.92	37332.30
	1	17.23	62.61	15455.25
	2	50.94	47.99	4032.73
	3	16.88	61.81	15655.29
	Control	7.01	26.35	36723.08
	1	15.88	60.97	15870.97
	2	48.95	47.02	4115.92
	3	16.88	62.48	15487.41
	Control	7.88	25.25	38322.90
31-Aug	1	22.49	76	12732.28
	2	55.99	54.82	3530.29
	3	19.47	64	15119.58
	Control	8.71	28.64	33786.78
	1	19.41	70.7	13686.75
	2	52.99	52.75	3668.83
	3	20.4	68.58	14109.85
	Control	9.65	30.14	32105.28
	1	18.28	70.57	13711.96
	2	51.58	52.66	3675.10
	3	19.82	66.15	14628.17
	Control	9.58	28.99	33378.86
31-Aug	1	19.46	73.18	13222.92
	2	49.4	48.36	4001.87
	3	19.33	67.19	14401.75
	Control	9.24	28.38	34096.31
	1	17.88	71.25	13581.10
	2	49.04	51.98	3723.18
	3	20.97	65.68	14732.85
	Control	9.17	27.89	34695.35
	1	18.25	68.01	14228.10
	2	48.27	48.97	3952.02
	3	18.51	65.93	14676.98
	Control	9.34	29.14	33207.04
02-Sep	1	25.45	92.09	10507.69
	2	57.09	60.98	3173.67
	3	19.4	73.67	13134.97
	Control	7.83	26.38	36681.32
	1	22.19	85.11	11369.44
	2	55.41	59.96	3227.66
	3	21.26	73.85	13102.96
	Control	8.7	27.9	34682.91

	1	22.11	83.75	11554.07
	2	56.85	59.52	3251.52
	3	22.52	73.91	13092.32
	Control	8.83	27.96	34608.49
02-Sep	1	20.37	79.23	12213.22
	2	53.69	56.13	3447.90
	3	22.35	71.8	13477.06
	Control	7.84	27.91	34670.49
	1	21	77.34	12511.68
	2	55.65	52.24	3704.64
	3	20.46	72.03	13434.03
	Control	8.42	29.03	33332.87
	1	19.6	75.54	12809.81
	2	55.01	51.48	3759.34
	3	21.4	68.16	14196.79
	Control	8.69	26.43	36611.93
03-Sep	1	27.48	90.3	10715.98
	2	66.79	60.82	3182.02
	3	26.8	81.21	11915.44
	Control	9.37	27.68	34958.57
	1	23.4	86.12	11236.10
	2	59.08	61.33	3155.56
	3	24.28	78.54	12320.52
	Control	9.6	27.02	35812.48
	1	21.66	84.04	11514.20
	2	56.25	58.31	3319.00
	3	22.95	75.91	12747.38
	Control	8.37	27.51	35174.60
03-Sep	1	23.86	84.98	11386.84
	2	72.01	68.97	2806.01
	3	23.23	76.56	12639.15
	Control	7.64	27.45	35251.49
	1	23.84	82.4	11743.36
	2	67.62	69.13	2799.52
	3	22.54	76.41	12663.96
	Control	7.7	25.89	37375.56
	1	22.47	79.91	12109.29
	2	68.73	67.44	2869.67
	3	21.66	76.11	12713.88
	Control	7.44	27.39	35328.71

APPENDIX D: CONVOLUTIONAL NEURAL NETWORKS

Convolutional Neural Networks (CNNs) are feedforward neural networks that are able to take large, multi-dimensional arrays as inputs. These types of networks did display very early success (LeCun et al., 1990), but it was not until the advancement in computer processing power, the advent of the Graphics Processing Unit (GPU), and better algorithms and architecture that contributed to these network's recent popularity in computer vision applications (Raway & Wang, 2017).

One of the most notable achievements of CNNs in computer vision was the success of the AlexNet architecture in the Image Large Scale Visual Recognition Challenge (ILSVRC) competition in 2012 where a database containing 1.2 million of high-resolution images (ImageNet) was used for classification (Krizhevsky et al., 2012). Significant characteristics of this network are: its innovative use of a Rectified Linear Unit (ReLU) activation function, its use of dropout to prevent overfitting, and its efficient use of GPUs during the convolution process. In 2013, ZFNet won the ILSVRC competition by improving the AlexNet architecture (Zeiler & Fergus, 2014). The most significant feature of ZFNet was its ability to visualize the architecture and thus, improving details in the architecture by adjusting hyperparameters through visual inspection was made possible. In 2014, Google joined and won the ILSVRC competition with their GoogleNet architecture (also known as Inception V1). GoogleNet has 22 layers with 4 million parameters compared to AlexNet's 7 layers and 60 million parameters. This optimization in architecture allowed significant improvement for classification and detection tasks (Szegedy et al., 2015). Also, in 2014, the runner up to the ILSVRC competition is also worth mentioning. Visual Geometry Group (VGG) developed the VGGNet architecture which uses very small (3x3) convolution filters and 16 – 19 layers (Simonyan & Zisserman, 2014). The overall architecture of VGGNet is straightforward and its use of small convolution kernels and the fact that it is open source allowed it to be used in a variety of applications. The drawback of the VGGNet architecture is its 138 million parameters that may be difficult to work with. Finally, in 2015, Residual Neural Network (ResNet) won the ILSVRC by introducing a novel architecture approach that uses shortcut (residual) connections to skip layers (He et al., 2016). This technique allowed for a much deeper network while maintaining relatively low complexity. The ResNet architecture

consists of 152 layers with 60 million parameters. **Table 37** provides a summary of the different CNN architectures and their performance for image classification with the ImageNet database.

Table 37: Summary of different CNN architectures used for Image Classification and their performance with ImageNet database

Developed by & Year Developed	Architecture	Number of Layers	Number of Parameters	Top-5 error rate	ILSVRC placing
Krizhevsky et al., 2012	AlexNet	7	60 million	15.3%	1st
Zeiler & Fergus, 2013	ZFNet	7	60 million	14.8%	1st
Simonyan & Zisserman, 2014	VGGNet	16 - 19	138 million	7.3%	2nd
Szegedy et al., 2015 (Google)	GoogleNet (Inception V1)	19	4 million	6.67%	1st
He et al., 2015	ResNet	152	60 million	3.6%	1st

The training procedure of CNNs is similar to MLPs. Initial weights in the network are chosen randomly and deviations between the actual and predicted class are calculated. This deviation is minimized by adjusting weights and biases throughout the network. Backpropagation is the most popular approach for CNNs where the gradient of an objective function is computed with respect to the weights and biases within the network to obtain desired outputs (LeCun et al., 2015; Rawat & Wang, 2017). There are several methods for this but the most popular is the use of stochastic gradient descent (SGD) using backpropagation as it is the most efficient and simplest way of minimizing

the deviations (Cha et al., 2017). SGD consists of showing some examples of input vectors, computing the outputs and the errors, computing the gradient for those examples and adjusting the weights accordingly. This process repeats for many small sets of examples from the training set until the objective function stops decreasing (LeCun et al., 2015). Backpropagation is applied to propagate gradients through the different layers starting from the output to the input. Learning rate and momentum are hyperparameters that are used to optimize the learning process. Typically, a lower learning rate is recommended for training efficiency as large learning rates tend to converge slowly and even diverge (Cha et al., 2017). Momentum accelerates the gradients and leads to faster convergence. **Figure 64** shows an example of gradient descent with two different learning rates. The figure illustrates how a loss function is minimized by adjusting two weights W1 and W2. If the partial derivatives with respect to W1 and W2 at point *a* in Figure 54(a) are calculated, a vector (gradient) is obtained. The projection of this vector on the W1-W2 plane always tends to head to the steepest gradient, which is towards the minimum of the loss function (descent) (Cha et al., 2017).

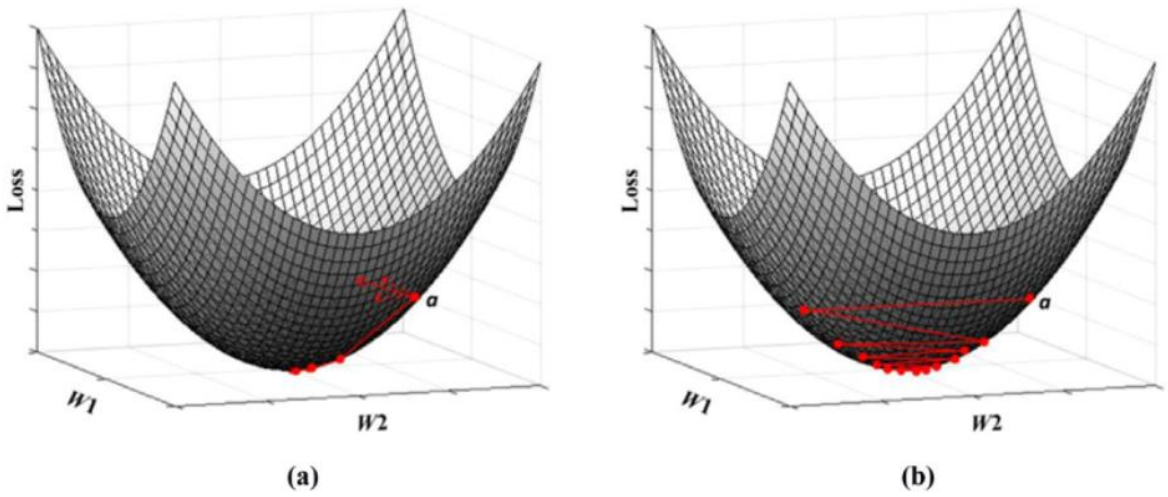


Figure 64: Example of gradient descent for (a) small learning rate and (b) large learning rate (Cha et al., 2017).

When training a CNN, it is recommended that a large input dataset be used for effective training. Small datasets could lead to overfitting or inaccurate networks. Collecting and labelling large datasets is time consuming and often expensive. Thus, transfer learning became an effective tool when large datasets are unavailable. Given a source domain and its learning task and a target domain and its target task, the objective of transfer learning is to help improve the prediction function in learning target tasks using the knowledge from the source domain with source targets (Gao & Mosalam, 2018). In CNNs, parameters of a well-trained model are able to be 'transferred' to another model. Typically, parameters of the shallower layers that represent low-level features are transferred to another model for different classification tasks. Gao & Mosalam in 2018 implemented transfer learning by using the VGG-16 network architecture trained on the ImageNet database for structural damage recognition. Liang in 2019 implemented transfer learning from the VGG-16 architecture as well for structural object detection, damage detection and segmentation. Liang compared the architectures of AlexNet, Google Net and VGG-16 and found VGG-16 to perform the best for his classification task in terms of testing accuracy. Although VGG-16 lost to Google Net in the ILSVRC competition in 2014, it seems the simplicity of the model allows for easier transferability.

The general process of CNNs involve several stages of convolution and pooling layers that extract features of the input while reducing complexity followed by fully connected layers which are similar in architecture to MLPs. Typically, an image is directly, or after some initial pre-processing, input to the network and is followed by convolution and pooling layers to extract a feature vector. This feature vector is fed into the fully connected layers until the last layer gives an output depending on the task at hand.

Convolution layers extract features, also known as feature maps, of the input through the following procedure. Element-by-element multiplication is done with a subarray of the input and a receptive field (LeCun et al., 2015; Rawat & Wang, 2017). The receptive field, also known as a filter or kernel, contain trainable weights that are optimized through the training procedure. Inputs are convolved with the receptive field and the results are sent to a non-linear activation function to give nonlinearity (Rawat & Wang, 2017). The following equation describes the convolution process where the k^{th}

output feature map Y_k can be computed. The input image is denoted as x; the filter related to

$$Y_k = f(W_k * x)$$

the k^{th} feature map is represented as W_k; the multiplication sign represents the convolution operation; and the function f represents the non-linear activation function. One important parameter of the convolution process is the ability to choose stride size. The stride defines how many of the receptive field's columns and rows (pixels in the case of an image input) slide at a time across the input array's width and height (Cha et al., 2017). The advantage of this is to reduce computation by skipping some of the input. The reason this is able to be done is due to the fact that images are usually spatially dependant and using stride typically does not reduce accuracy. The main benefit of the convolution process in general is the reduction in input size while extracting features, as seen in **Figure 65**.

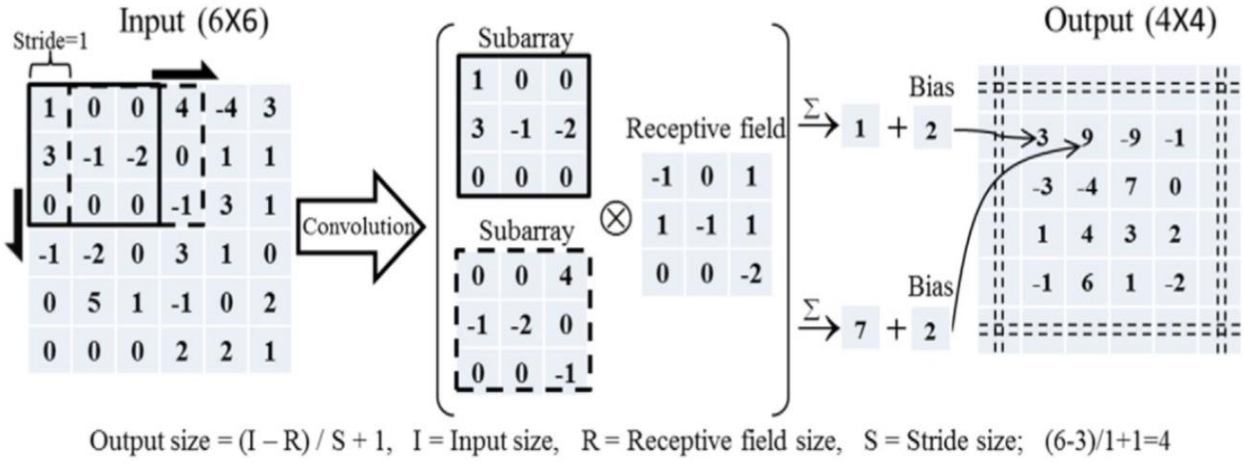


Figure 65: Example of the Convolution process. The output is the feature map after the convolution process (Cha et al., 2017).

The objective of the pooling layer is to reduce the spatial size of an input and is done to achieve spatial invariance to input distortion and translation (Rawat & Wang, 2017; Cha et al., 2017). There are two types of pooling options. Initially, average pooling was done where the pooling outputs the average rounded integer value of a subarray. Recent models (Krizhevsky et al., 2012; Simonyan & Zisserman, 2014; Szegedy et al., 2015) use max pooling where the output is the maximum value in the subarray. Equation 2.3 describes the max pooling operation where the output of pooling operation of the k^{th} feature map is denoted by Y_{kij}, x_{kpq} represents the element at location (p, q) in the pooling region R_{ij}, which incorporates the receptive field around the position (i, j) (Rawat & Wang, 2017). Similar to convolution layers, choosing a stride length is an option for

$$Y_{kij} = \max_{(p,q) \in R_{ij}} x_{kpq} \tag{2.3}$$

computational efficiency but at the cost of features of the input data. **Figure 66** shows a general pooling procedure and the reduction in size of the input is seen.

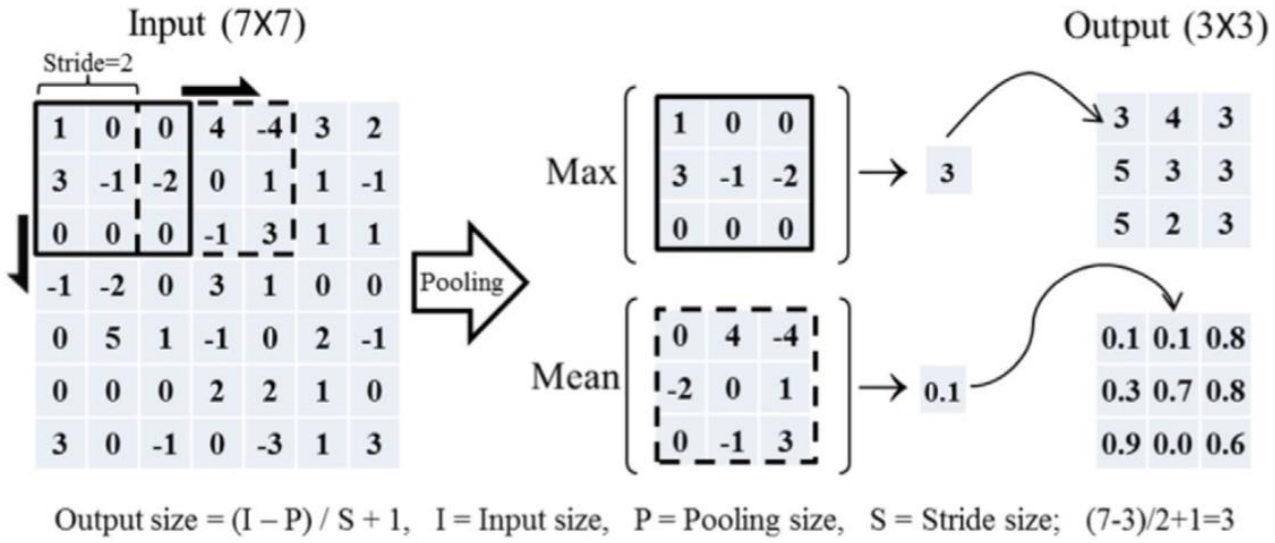


Figure 66: Example of the Pooling process and the difference between Max Pooling and Mean Pooling (Cha et al., 2017).

The convolution and pooling layers are used in CNNs for feature extraction. Multiple layers of convolution and pooling are used sequentially to extract features and reduce the size of the input data as it moves through the network. Fully connected layers follow these layers and interpret the extracted feature data to perform the function of the required task, for example classification. A ‘Softmax’ layer is often employed as the final layer of a CNN for the task of classification (Krizhevsky et al., 2012; Simonyan & Zisserman, 2014; Szegedy et al., 2015). Other approaches involve replacing the final connected layers and using the outputs of the feature extraction segment of the CNN as inputs to a separate classification network (Gopalakrishnan et al., 2017; Nhat-Duc et al., 2018). Other than classification, regression problems have been investigated with the use of CNNs for end-to-end classification and ordinal regression. Niu et al. (2016) developed a multiple output CNN algorithm for age prediction that showed successful results when compared to traditional hand-crafted techniques.

In general, CNNs in computer vision are used in the following applications. Classification is the most common application where an image input is given to the network for object classification. For example, if an image of a cat is given to a network that is trained to detect animals, the network should be able to predict with high accuracy that the given image contains a cat. From classification, other uses of CNNs in computer vision are object detection and semantic segmentation. Object detection applications typically use ‘Regional Proposal Networks’ (RPN) that use anchors to zone in on potential objects (Ren et al., 2015). The output of the RPN is then used in a classification network or SoftMax layer for the task of classifying those object proposals. Thus, multiple objects can be classified in the same image (Ren et al., 2015; Cha et al., 2018). Other approaches use a ‘Fast-CNN’ which uses the image as a direct input instead of feature maps to identify the region of proposal by using a region of interest pooling layer that identifies potential objects in a bounded box which was found to greatly speed up the process. The latest innovation of these models is the Faster-RCNN which combines the RPN and Fast-CNN and uses the RPN within its model to predict region proposals.

Another common application of CNNs is in semantic segmentation. This application uses a similar approach in multi-object detection such as the use of an RPN

but also utilizes an encoder-decoder convolution network where the encoder deconstructs the input to extract features and the decoder reconstructs the segmented image as an output (Dung et al., 2019; Liang, 2019). Object detection and semantic segmentation are used in the field of autonomous vehicles (Wu et al., 2017), medical imagery (Kayalibay et al., 2017) and various others.