

Industrial

Electronics

Biomedical

Civil

Aerospace

Computer

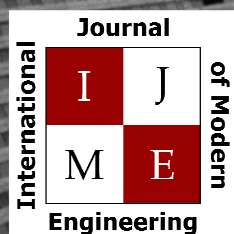
Electrical

Chemical

Mechanical



ENGINEERING



www.ijme.us

Print ISSN: 2157-8052
Online ISSN: 1930-6628



www.iajc.org

INTERNATIONAL JOURNAL OF MODERN ENGINEERING

ABOUT IJME:

- IJME was established in 2000 and is the first and official flagship journal of the International Association of Journal and Conferences (IAJC).
- IJME is a high-quality, independent journal steered by a distinguished board of directors and supported by an international review board representing many well-known universities, colleges and corporations in the U.S. and abroad.
- IJME has an impact factor of **3.00**, placing it among the top 100 engineering journals worldwide, and is the #1 visited engineering journal website (according to the National Science Digital Library).

OTHER IAJC JOURNALS:

- The International Journal of Engineering Research and Innovation (IJERI)
For more information visit www.ijeri.org
- The Technology Interface International Journal (TIIJ).
For more information visit www.tiij.org

IJME SUBMISSIONS:

- Manuscripts should be sent electronically to the manuscript editor, Dr. Philip Weinsier, at philipw@bgsu.edu.

For submission guidelines visit
www.ijme.us/submissions

TO JOIN THE REVIEW BOARD:

- Contact the chair of the International Review Board, Dr. Philip Weinsier, at philipw@bgsu.edu.

For more information visit
www.ijme.us/ijme_editorial.htm

INDEXING ORGANIZATIONS:

- IJME is currently indexed by 22 agencies.
For a complete listing, please visit us at www.ijme.us.

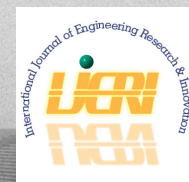
Contact us:

Mark Rajai, Ph.D.

Editor-in-Chief
California State University-Northridge
College of Engineering and Computer Science
Room: JD 4510
Northridge, CA 91330
Office: (818) 677-5003
Email: mrjai@csun.edu



www.tiij.org



www.ijeri.org

INTERNATIONAL JOURNAL OF MODERN ENGINEERING

The INTERNATIONAL JOURNAL OF MODERN ENGINEERING (IJME) is an independent, not-for-profit publication, which aims to provide the engineering community with a resource and forum for scholarly expression and reflection.

IJME is published twice annually (fall and spring issues) and includes peer-reviewed research articles, editorials, and commentary that contribute to our understanding of the issues, problems, and research associated with engineering and related fields. The journal encourages the submission of manuscripts from private, public, and academic sectors. The views expressed are those of the authors and do not necessarily reflect the opinions of the IJME editors.

EDITORIAL OFFICE:

Mark Rajai, Ph.D.
Editor-in-Chief
Office: (818) 677-2167
Email: ijmeeditor@iajc.org
Dept. of Manufacturing Systems
Engineering & Management
California State University-
Northridge
18111 Nordhoff Street
Northridge, CA 91330-8332

THE INTERNATIONAL JOURNAL OF MODERN ENGINEERING EDITORS

Editor-in-Chief:

Mark Rajai
California State University-Northridge

Associate Editors:

Alok Verma
Old Dominion University

Li Tan
Purdue University North Central

Production Editor:

Philip Weinsier
Bowling Green State University-Firelands

Subscription Editor:

Morteza Sadat-Hossieny
Northern Kentucky University

Web Administrator:

Saeed Namyar
Advanced Information Systems

Executive Editor:

Paul Wilder
Vincennes University

Manuscript Editor:

Philip Weinsier
Bowling Green State University-Firelands

Copy Editor:

Li Tan
Purdue University North Central

Technical Editors:

Michelle Brodke
Bowling Green State University-Firelands

Marilyn Dyrud
Oregon Institute of Technology

Barry Hoy
St. Leo University

Publisher:

Bowling Green State University Firelands

TABLE OF CONTENTS

<i>Editor's Note: IJME Celebrates 15 Years of Service with this Special Anniversary Issue</i>	3
Philip Weinsier, IJME Manuscript Editor	
<i>Analysis of Past Storm Events and Their Contribution to Flooding in the Clark County Wetlands Park</i>	5
Anil Acharya, Alabama A&M University; Thomas C. Piechota, University of Nevada, Las Vegas	
<i>Nonlinear and Buckling Analysis of Steel Transmission Towers</i>	17
Mustafa Mahamid, University of Illinois at Chicago; Cenk Tort, Miteng, TURKEY; Mutaz Al Hijaj, University of Illinois at Chicago; Adeeb Rahman, University of Wisconsin-Milwaukee; Faris Malhas, Central Connecticut State University	
<i>Self-Sustainability Study of Energy Harvesting Wireless Sensor Networks for Infrastructure Monitoring</i>	29
Jin Zhu, University of Northern Iowa	
<i>Advancing Motorsports Driver Protection using Modeling and Test Correlation</i>	36
Grant Hankins, Indiana University Purdue University Indianapolis; Pete Hylton, Indiana University Purdue University Indianapolis; Ryan Hoover, Center for Advanced Product Evaluation; Kevin Tribbett, Center for Advanced Product Evaluation	
<i>Gasoline Vapor Liquid Equilibrium Using the Peng Robinson Equation of State and Rachford-Rice Equation</i>	49
Mebougna L. Drabo, Alabama A&M University; Marcus D. Ashford, University of Alabama	
<i>Spiral Sensing and Probability Map of Impact (PMOI) for Impact Characterization</i>	56
Rowshan A. Rima, University of South Carolina; Agbasi Chijioke, Cummins, Inc.; Paramita Chakraborty, University of South Carolina; Sourav Banerjee, University of South Carolina	
<i>A Multiprocessor Version of Modular Neural Networks for Fault Diagnosis of Analog Electronic Circuits</i>	69
Mehran Aminian, St. Mary's University; Farzan Aminian, Trinity University	
<i>Models for Predicting the Compression Index of Soils Using Fractal Scaling of Particle Size Distributions</i>	78
Fred Boadu, Duke University	
<i>Characterization of Weathered Limestone using Thermal Imaging and Split Tension Testing</i>	89
John Nuskowski, University of North Florida; Mason McGough, University of Florida; Lexi Kimes, University of North Florida; Alan Harris, University of North Florida; Nick Hudyma, University of North Florida; Faris Malhas, Central Connecticut State University	
<i>Predicting Solar Radiation for Renewable Energy Technologies: A Random Forest Approach</i>	100
Frank E. Yeboah, North Carolina A&T State University; Robert Pyle, North Carolina A&T State University; Christian A. Bock Hyeng, North Carolina A&T State University	
<i>Instructions for Authors: Manuscript Submission Guidelines and Requirements</i>	108

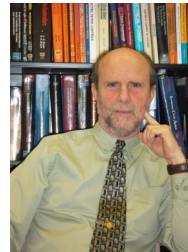
IJME Celebrates 15 Years of Service with this Special Anniversary Issue

We are pleased to present to you, our readers, this special anniversary issue, in celebration of our first 15 years of service. IJME has come a long way since its days as an online-only journal in the spring of 2000, and has become one of the most respected and well-known engineering journals in academia. We hope that you enjoy reading and learning from our articles as much as I do as editor. Throughout the years, the diversity of authors springs from our dedication to personal service and support, and word-of-mouth advertising by you, our loyal authors and readers. I would also like to thank all of the members of our International Advisory Board for their service in the review of the great works that we publish every year.

For our upcoming 2016 IAJC conference, I would like to invite you to join us in Orlando, Florida. For this fifth IAJC conference, we will be partnering again with the International Society of Agile Manufacturing (ISAM). This event will be held at the new Embassy Suites hotel, November 6-8, 2016, and is sponsored by IAJC, IEEE, ASEE, and the LEAN Institute. The IAJC/ISAM Executive Board is pleased to invite faculty, students, researchers, engineers, and practitioners to present their latest accomplishments and innovations in all areas of engineering, engineering technology, math, science, and related technologies.

Selected papers from the conference will be published in the three IAJC-owned journals. Oftentimes, these papers, along with manuscripts submitted at-large, are reviewed and published in less than half the time of other journals. Publishing guidelines are available at www.iajc.org, where you can read any of our previously published journal issues, as well as obtain information on chapters, membership, and benefits.

I am pleased to report that, based on the most recent impact factor (IF) calculations (Google Scholar method), the International Journal of Modern Engineering (IJME) has a remarkable IF = 3.0 and continues its



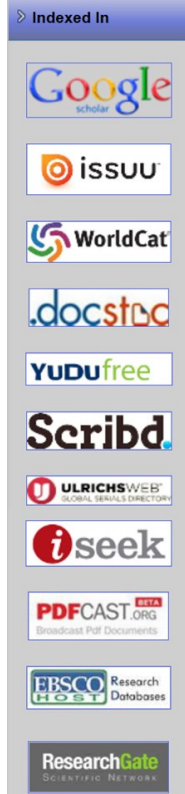
EDITOR'S NOTE

Philip Weinsier, IJME Manuscript Editor

march toward the top 20 engineering journals. IJME's sister journal, the International Journal of Engineering Research and Innovation (IJERI) also has a strong IF = 1.58, which is noteworthy, as it is a relatively young journal (in publication since 2009). Any IF above 1.0 is considered high, based on the requirements of many top universities, and places the journals among an elite group.

Currently, there is no official ranking system for journals that publish engineering-related topics the way that IJME and IJERI do, but both IJME and IJERI are indexed in most well-known indexing databases including DOAJ, which is the most prestigious and comprehensive database for open-access journals worldwide. Both journals are also indexed in the libraries of all 10 campuses of the University of California system, the 23 campuses of the California State University system, and all of the top 10 universities in the world:

- #1 California Institute of Technology
- #2 Harvard
- #3 Stanford
- #4 University of Oxford
- #5 Princeton University
- #6 University of Cambridge
- #7 Massachusetts Institute of Technology
- #8 Imperial College London
- #9 University of Chicago
- #10 University of California, Berkeley



Editorial Review Board Members

Mohammed Abdallah	State University of New York (NY)	Soo-Yen Lee	Central Michigan University (MI)
Nasser Alaraje	Michigan Tech (MI)	Chao Li	Florida A&M University (FL)
Aly Mousaad Aly	Louisiana State University (LA)	Jimmy Linn	Eastern Carolina University (NC)
Jahangir Ansari	Virginia State University (VA)	Dale Litwhiler	Penn State University (PA)
Kevin Berisso	Ohio University (OH)	Guoxiang Liu	University of North Dakota (ND)
Salah Badjou	Wentworth Institute of Technology (MA)	Louis Liu	University of New Orleans (LA)
Pankaj Bhambri	Guru Nanak Dev Engineering (INDIA)	Mani Manivannan	ARUP Corporation
Water Buchanan	Texas A&M University (TX)	G.H. Massiha	University of Louisiana (LA)
Jessica Buck Murphy	Jackson State University (MS)	Thomas McDonald	University of Southern Indiana (IN)
John Burningham	Clayton State University (GA)	David Melton	Eastern Illinois University (IL)
Shaobiao Cai	Penn State University (PA)	Shokoufeh Mirzaei	Cal State Poly Pomona (CA)
Vigyan Chandra	Eastern Kentucky University (KY)	Bashir Morshed	University of Memphis (TN)
Isaac Chang	Cal Poly State University SLO (CA)	Sam Mryyan	Excelsior College (NY)
Bin Chen	Purdue University Calumet (IN)	Wilson Naik	University of Hyderabad (INDIA)
Wei-Yin Chen	University of Mississippi (MS)	Arun Nambiar	California State University Fresno (CA)
Hans Chapman	Morehead State University (KY)	Ramesh Narang	Indiana University-Purdue University (IN)
Rigoberto Chinchilla	Eastern Illinois University (IL)	Anand Nayyar	Institute Management and Tech (INDIA)
Phil Cochrane	Indiana State University (IN)	Stephanie Nelson	Cal State LA (CA)
Michael Coffman	Southern Illinois University-Carbondale (IL)	Hamed Niroumand	Universiti Teknologi (MALAYSIA)
Emily Crawford	Southern Wesleyan University (SC)	Aurenice Oliveira	Michigan Tech (MI)
Brad Deken	Southeast Missouri State University (MO)	Troy Ollison	University of Central Missouri (MO)
Z.T. Deng	Alabama A&M University (AL)	Reynaldo Pablo	Indiana University-Purdue University (IN)
Sagar Deshpande	Ferris State University (MI)	Basile Panoutsopoulos	Community College of Rhode Island (RI)
David Domermuth	Appalachian State University (NC)	Shahera Patel	Sardar Patel University (INDIA)
Ryan Dupont	Utah State University (UT)	Jose Pena	Purdue University Calumet (IN)
Marilyn Dyrud	Oregon Institute of Technology (OR)	Karl Perusich	Purdue University (IN)
Mehran Elahi	Elizabeth City State University (NC)	Thongchai Phairoh	Virginia State University (VA)
Ahmed Elsayy	Tennessee Technological University (TN)	Huyu Qu	Honeywell Corporation
Rasoul Esfahani	DeVry University (OH)	John Rajadas	Arizona State University (AZ)
Dominick Fazarro	Sam Houston State University (TX)	Desire Rasolomampionona	Warsaw University of Tech (POLAND)
Rod Flanigan	University of Nebraska-Kearney (NE)	Mulchand Rathod	Wayne State University (MI)
Ignatius Fomunung	University of Tennessee Chattanooga (TN)	Mohammad Razani	New York City College of Tech (NY)
Ahmed Gawad	Zagazig University EGYPT)	Sangram Redkar	Arizona State University-Poly (AZ)
Daba Gedafa	University of North Dakota (ND)	Michael Reynolds	University of Arkansas Fort Smith (AR)
Ralph Gibbs	Eastern Kentucky University (KY)	Marla Rogers	Wireless Systems Engineer
Mohsen Hamidi	Utah Valley University (UT)	Dale Rowe	Brigham Young University (UT)
Mamoon Hammad	Abu Dhabi University (UAE)	Anca Sala	Baker College (MI)
Youssef Himri	Safety Engineer in Sonelgaz (ALGERIA)	Mehdi Shabaninejad	Zagros Oil & Gas Company (IRAN)
Xiaobing Hou	Central Connecticut State University (CT)	Ehsan Sheybani	Virginia State University (VA)
Shelton Houston	University of Louisiana Lafayette (LA)	Musibau Shofoluwe	North Carolina State University (NC)
Barry Hoy	St. Leo University (VA)	Siles Singh	St. Joseph University Tanzania (AFRICA)
Ying Huang	North Dakota State University (ND)	Ahmad Sleiti	University of North Carolina Charlotte (NC)
Charles Hunt	Norfolk State University (VA)	Jiahui Song	Wentworth Institute of Technology (MA)
Dave Hunter	Western Illinois University (IL)	Yuyang Song	Toyota Corporation
Christian Hyeng	North Carolina A&T University (NC)	Carl Spezia	Southern Illinois University (IL)
Pete Hylton	Indiana University Purdue (IN)	Michelle Surerus	Ohio University (OH)
Ghassan Ibrahim	Bloomsburg University (PA)	Vassilios Tzouanas	University of Houston Downtown (TX)
John Irwin	Michigan Tech (MI)	Jeff Ulmer	University of Central Missouri (MO)
Sudershan Jetley	Bowling Green State University (OH)	Mihaela Vorvoreanu	Purdue University (IN)
Rex Kanu	Ball State University (IN)	Phillip Waldrop	Georgia Southern University (GA)
Reza Karim	North Dakota State University (ND)	Abraham Walton	Purdue University (IN)
Tolga Kaya	Central Michigan University (MI)	Liangmo Wang	Nanjing University of Science/Tech (CHINA)
Satish Ketkar	Wayne State University (MI)	Jonathan Williams	Lake Erie College (OH)
Manish Kewalramani	Abu Dhabi University (UAE)	Boonsap Witchayangkoon	Thammasat University (THAILAND)
Tae-Hoon Kim	Purdue University Calumet (IN)	Alex Wong	Digilent Inc.
Doug Koch	Southeast Missouri State University (MO)	Shuju Wu	Central Connecticut State University (CT)
Sally Krijestorac	Daytona State College (FL)	Baijian Yang	Ball State University (IN)
Ognjen Kuljaca	Brodarski Institute (CROATIA)	Mijia Yang	North Dakota State University (ND)
Chakresh Kumar	Uttar Pradesh Tech University (INDIA)	Faruk Yildiz	Sam Houston State University (TX)
Zaki Kuruppalil	Ohio University (OH)	Yuqiu You	Morehead State University (KY)
Edward Land	Johns Hopkins Medical Institute	Jinwen Zhu	Missouri Western State University (MO)
Ronald Land	Penn State University (PA)		
Jane LeClair	Excelsior College (NY)		
Shiyoung Lee	Penn State University Berks (PA)		

ANALYSIS OF PAST STORM EVENTS AND THEIR CONTRIBUTION TO FLOODING IN THE CLARK COUNTY WETLANDS PARK

Anil Acharya, Alabama A&M University; Thomas C. Piechota, University of Nevada, Las Vegas

Abstract

In this study, the authors analyzed historical rainfall patterns and their impact on runoff generation from three major watersheds—Duck Creek (DC), Pittman Wash (PM) and C1 Channel (C1)—that contributed to flooding in the Clark County Wetlands Park in the Las Vegas Valley. This data-driven analysis assessed major changes in rainfall and runoff over time (1990-2010) and identified threshold values of water depth for a significant occurrence of flooding. The higher variation in rainfall signified the general nature of rainfall in Las Vegas. DC possesses a higher percentage of events (>50 mm) lasting more than 12 hours, compared to C1 and PM. Ninety percent of total storm events have rainfall depths of 18 mm or less (11 to 14 mm/hr.). The median water depths for DC, PM, and C1 during the 18 mm rain event were 0.48, 0.24, and 0.24 m, respectively. During low flows, the runoff generation for C1, DC, and PM occurred at water depths of 0.06, 0.075, and 0.04 m, respectively, for 2, 4, and 1 mm of total rainfall. A higher percentage of storm events was observed in the lower water depths, which do not produce significant runoff. This analysis can be utilized by planning agencies to determine the area contributing more for floods and design flood control facilities and management measures.

Introduction

Clark County, Nevada, is considered one of the fastest growing Counties in the U.S. Based on Clark County's official population estimate, the population has reached almost 2.5 million—an increase of almost two and a half times since 1990 [1]. The County is located in the arid environment of the Mojave Desert. Las Vegas is the most populous city in the state of Nevada located at $36^{\circ}11'39''\text{N}$ and $115^{\circ}13'19''\text{W}$. Rapid population growth along with urban, commercial, and industrial development has brought about an increase of physical infrastructures (e.g., roads, structural drainage, concrete channels, buildings, etc.). According to the U.S. Census Bureau, it possesses a total area of 340.06 km^2 (131.3 sq. miles) with 339.8 km^2 (131.2 sq. miles) not covered with water and 0.26 km^2 (0.1 sq. miles $\sim 0.04\%$) covered with water. The city is located at an elevation of around 620 m (2030 feet) above mean sea level, and

the landscape mostly features rocky and dusty terrain.

The mean annual precipitation varies from 10 to 50 cm (4 to 20 inches) between southeast Clark County and the Spring Mountains [2]. Clark County has already experienced more than 11 floods since 1993. The July 8, 1999, event produced over 3.8 cm (1.5 inches) of rainfall within 60-90 minutes, and several gauges had over 7.6 cm (3 inches) of rainfall in the same time period [3]. The August 19, 2003, event was limited to the northwest portion of the Las Vegas Valley and produced over 5.1 cm (2 inches) of rainfall in a 30-90 minute time period. The total damage to property and roadways resulting from flooding was estimated to be over \$2 million [3]. Another rainfall event of January 10-14, 2005, resulted in over \$20 million in damages to property in the Mesquite area of Clark County [3].

Practically, the storm events are considered very rare in arid regions. Las Vegas Valley has experienced major flood events, due to heavy local thunderstorms [2]. The major cause of flash flooding includes intense localized precipitation as a result of the prevailing convective mechanisms over the watershed. The major, intense thunderstorms occur during the summer months and are localized, whereas the mild storms occur during winter months with longer duration and cover a larger area. The most damaging storms mainly occur from July to September. The average rainfall in the Las Vegas Valley is 11.40 cm (4.49 inches) per year with half of the total rainfall generated by summer thunderstorms. Although rainfall events with magnitudes greater than a tenth of an inch occur less than 10 days a year on average in Las Vegas [4], flash floods have become a major concern, due to their devastating damage. They are also difficult to predict, due to lack of long-term precipitation records.

Las Vegas Wash receives runoff from the entire Las Vegas Valley and is situated adjacent to the Clark County wetlands park. The Wash drains approximately 2351 km^2 (1460 sq. miles) of Clark County starting from the North Western portion of the valley and discharging into Lake Mead over towards the East. The nine major watersheds (North, Gowan, Range, Central, Flamingo Tropicana, Duck Creek, Pittman Wash, Lower, and C1) that contribute to Las Vegas Wash are shown in Figure 1. The Wetlands Park (the poly-

gon inside Lower and in the boundary of Pittman and C1 in Figure 1) is an invaluable environmental resource with various walking trails and the facility is open to the public. However, it carries a potentiality of inundation with flood waters during significant rainfall events. This research project analyzed the historical rainfall pattern and its impact on the generation of runoff from major watersheds that are contributing to flooding in the Clark County Wetlands Park. The analysis was data driven and incorporated these specific objectives: a) observe seasonal variation in rainfall patterns based on past storm events; b) identify the threshold values of water depths for significant occurrence of flooding; and, c) analyze regional variation in rainfall and water depth for major watersheds. This comprehensive study for this region utilized long-term, real-time data from many local/regional weather stations and the authors performed threshold value analyses (on runoff) to gain additional insights into flood management.

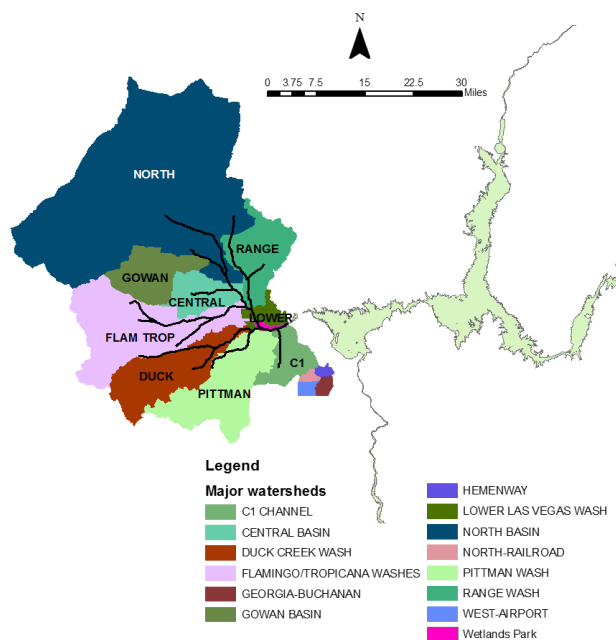


Figure 1. Location of the Nine Major Watersheds, Wetlands Park, the River Network of Lake Mead, and the Colorado River

Clark County Regional Flood Control District

The Clark County Regional Flood Control District (CCRFCDD), established in 1985, plays a major role in regulating the land use in flood hazard areas and developing a comprehensive master plan to solve flood-related problems in Clark County. The Flood Threat Recognition System (FTRS) was implemented by CCRFCDD in 1987 in coopera-

tion with the U.S. Geological Survey (USGS) and the National Weather Service (NWS). Out of the total 355 real-time meteorological sensors, almost 75 percent are present in Las Vegas Valley, with most of them in urban areas. Only about 150 operational weather stations provided by FTRS can inform the decision makers of flooded areas in the Las Vegas Valley. The flood control situation has slightly improved in some watersheds, due to construction of flood-control facilities like detention basins. However, the valley is under continuous development and not enough flood control facilities—such as planned/designed in MPU—are available to control extreme flash floods. Therefore, it still needs special concern in order to reduce losses and damage caused by flood events similar to earlier storm events [3].

The Hydrologic Criteria and Drainage Design Manual published by CCRFCDD is used for drainage design purposes inside Clark County. A major design storm is considered for a storm having a return period of 100 years and a duration of six hours. Based on the Regional Flood Control Master Plan Update [5], the regional flood control facilities planned by CCRFCDD in most areas are considered adequate for reducing heavy losses, since the existing facilities in place appear to be functioning as designed. However, the development of the active terrain causes a high probability of change in runoff over the entire watershed.

Data and Methodology

Among the nine major watersheds contributing to the Las Vegas Wash, the watersheds considered for this study included Duck Creek (DC), Pittman Wash (PM), and C1 Channel (C1). The selection was based on the fact that their runoff was assumed to significantly impact the Wetlands Park. Each watershed possessed a different number of weather stations, most of them installed by CCRFCDD and some by USGS. There are three types of weather stations established by FTRS: full weather, precipitation, and water level. Full weather stations collect temperature, humidity, and wind data in addition to rainfall data. Water level stations record real-time rainfall and water level at the same location, and as separate station IDs. Precipitation type only collects rainfall data. The real-time sensor data (in essence, the rainfall and water level data for each station)—in addition to details for each station regarding the type, station ID, location and their installation date—can be obtained from the CCRFCDD FTRS website [3].

This study utilized instrument data to determine trends in rainfall and water depth over the study period (1990-2010). Altogether, 28 weather stations (DC: 9, PM: 11, C1: 8), which were distributed throughout the watershed, were active in the three watersheds. The position of weather sta-

tions and their station IDs are shown in Figure 2. A more widely distributed pattern of rain gauges represents the watershed in a better scenario and avoids any exaggerated predictions [6]. GIS data showing the watershed boundaries, location of gauge stations, and Wash alignments were downloaded from the Clark County Geographic Information System Management Office (GISMO) [7]. Only one flow measuring station installed by the USGS was available in C1 (represented by a star symbol in Figure 2). The rain gauge network possessed both rainfall and water depth measurements at almost all stations. All circular dots in Figure 2 represent available gaging stations, while filled triangular stations were considered for water depth analyses. The polygon at the boundary of PM and C1 is the wetlands park.

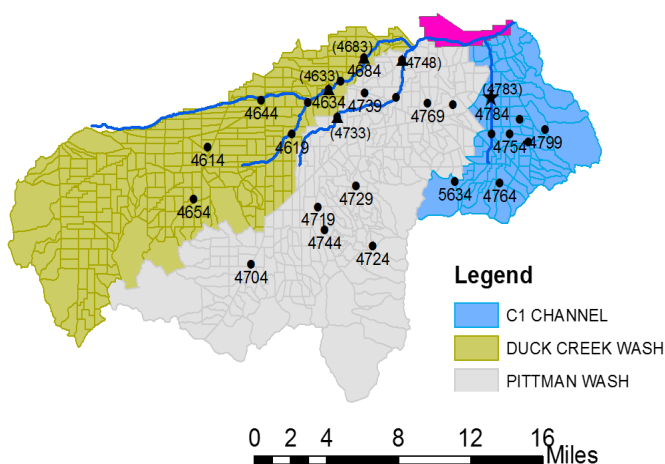


Figure 2. Location of Weather Stations and USGS Streamflow Gauge in three Major Watersheds

Rainfall Data

For each station, the real-time rainfall data were recorded at irregular time intervals and did not follow a fixed pattern. The recorded data incorporated longer time intervals (8 to 12 hours) in case of no precipitation; the recorded time interval varied from seconds to minutes during a rainfall event. The gauge data provided a continuous record of rainfall at increments of 1 mm, which was the smallest measurement of rainfall that the sensor could detect. The stations established earlier possessed data for longer periods, continuing from 1990, while the most recent had data records since 2000. This study included all events with cumulative rainfall greater than 1 mm; the duration of each event (in effect, the time between start and end of each rainfall event) was computed for each station. The separate events mainly represented the maximum slope for a plot of rainfall and duration. The number of total rainfall events for each station and the available records of rainfall data are summarized in Table 1. A total of 990 events for C1, 937 events for DC, and 1267 events for PM were considered for this study. Events in Table 1 represent only the events greater than 0.20 inches (50 mm) of rainfall.

Water Level Data

There is a lack of stations measuring streamflow in these watersheds. The water level data were used to fulfill the research objective and find the peak water surface elevation in the Wash. The water level stations considered here for analysis were located downstream of all rain gauge stations for each watershed. The most downstream water level sta-

Table 1. Total Rainfall Events (above 0.04 inches/1mm) for Weather Stations and Watersheds

C1 channel											
St. ID	4754	4764	4779	4784		4789	4794	4799	5634		
From	3/97	8/94	9/02	9/89		6/93	1/93	6/00	4/00		
To	5/08	5/08	5/08	5/08		5/08	5/08	5/08	5/08		
Events*	66	73	42	108		89	82	35	48		
Total Events	141	173	65	185		157	133	58	78		
DC channel											
St. ID	4614	4619	4624	4634		4644	4654	4674	4684	4694	
From	6/89	9/01	6/91	8/89		3/97	7/97	1/02	7/01	11/05	
To	5/08	5/08	5/08	5/08		5/08	5/08	5/08	5/08	05/08	
Events*	103	35	113	98		54	67	45	39	5	
Total Events	169	46	178	170		99	116	70	71	12	
PM Wash											
St. ID	4704	4719	4724	4729	4734	4739	4744	4749	4759	4769	4774
From	8/89	1/03	6/01	8/02	10/88	8/95	2/93	3/00	9/99	5/02	4/92
To	5/08	5/08	5/08	5/08	5/08	5/08	5/08	5/08	5/08	5/08	5/08
Events*	121	33	44	36	140	68	74	38	50	37	85
Total Events	205	55	70	66	199	127	139	83	85	66	172

tions in DC and PM did not possess sufficient data; a second water level station was also selected at an upstream part (Figure 2) that possessed a larger number of data points to compare to the behavioral change of the complete watershed. Station 4633 was located upstream of station 4683 and station 4733 was located upstream of station 4748. The existing data for each water level station are summarized in Table 2. The real-time water depth and rainfall data both follow a similar pattern. Maximum water depth was considered for each rainfall event selected for this analysis. Only some of the stations showing major rainfall events also had data showing water level changes. The effect rainfall had on increasing or decreasing the water level was analyzed for two water level stations along the DC and PM, and only one station along the C1 channel.

Results

Rainfall Analysis: Total Rainfall versus Number of Occurrences

The rainfall analysis was performed by categorizing the total events into two major seasons: December through May—representing the events in winter—and June through November representing the summer season. Table 3 summarizes the calculations of cumulative percentage of occurrence of total significant rainfall events with respect to total rainfall ranging from 1 mm to more than 50 mm for several ranges of rainfall depths for each season. The duration of

Table 2. Water Level Data Available for Selected Water Depth Stations in Each Watershed

Name of watersheds	C1 Channel	Duck Creek		Pittman Wash	
Stn ID	4783	4633	4683	4733	4748
From	02/1993	07/1991	07/2001	05/1991	01/2004
To	05/2008	05/2008	05/2008	05/2008	05/2008

Table 3. Cumulative Percentage of Occurrence of Rainfall Events for a Range of Rainfall Depths

Rainfall depths (mm)	December-May								
	Duck Creek			Pittman Wash			C1 Channel		
	No. of events	% of occurrence	Cum %	No. of events	% of occurrence	Cum %	No. of events	% of occurrence	Cum %
1-5	202	39.38	39.4	298	44.21	44.2	254	45.77	45.8
5-12.5	211	41.13	80.5	253	37.54	81.8	195	35.14	80.9
12.5-25	88	17.15	97.7	95	14.09	95.8	90	16.22	97.1
25-37.5	7	1.36	99.0	20	2.97	98.8	5	0.90	98.0
37.5-50	3	0.58	99.6	6	0.89	99.7	8	1.44	99.5
> 50	2	0.39	100.0	2	0.30	100.0	3	0.54	100.0
Total events	513			674			555		
Rainfall depths (mm)	June-November								
	Duck Creek			Pittman Wash			C1 Channel		
	No. of events	% of occurrence	Cum %	No. of events	% of occurrence	Cum %	No. of events	% of occurrence	Cum %
1-5	170	40.67	40.7	243	40.98	41.0	193	44.37	44.4
5-12.5	157	37.56	78.2	223	37.61	78.6	159	36.55	80.9
12.5-25	63	15.07	93.3	93	15.68	94.3	60	13.79	94.7
25-37.5	22	5.26	98.6	25	4.22	98.5	19	4.37	99.1
37.5-50	3	0.72	99.3	9	1.52	100.0	2	0.46	99.5
> 50	3	0.72	100.0	0	0.00		2	0.46	100.0
Total events	418			593			435		

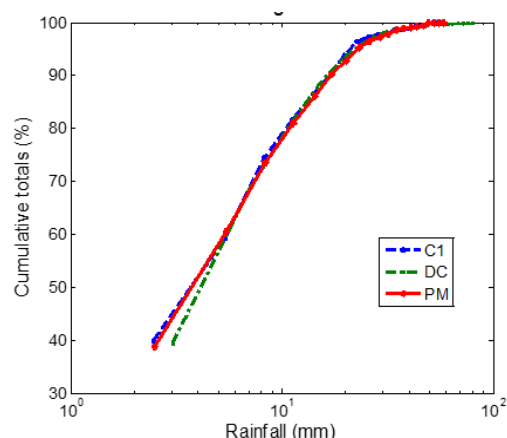
rainfall considered varied from a few minutes up to 24 hours. Based on seasonal calculations, more than 90 percent of total events were concentrated in the range of 25 mm of rainfall, and less than 1 percent of events with rainfall more than 50 mm. DC possessed a comparatively higher percentage of rainfall events above 50 mm. The rainfall events within 37.5 mm of rainfall for all three watersheds were slightly higher for the winter season, and almost equal among the three above that range. There was a high number of small rainfall events and a low number of larger storms observed. This indicated that the smaller events were more frequent with small return periods, whereas larger storms possessed larger return periods and occurred less frequently; a higher percentage of storm events were concentrated in the region, which did not produce any significant changes in water depth.

The resulting rainfall depth corresponding to 90 percent cumulative probability of occurrence was considered a design storm for management of storm water quality best management practices (BMPs). This type of storm possesses rainfall depths less than 1.25 inches (31.75 mm) and a recurrence interval of three to four months [8]. As mentioned earlier, more than 90 percent of rainfall events occurred within the depth range of 25 mm for all watersheds and for both seasons. The normalized distribution of the total number of events showed that about 90 percent of the total cumulative events were 18 mm or less (see Figure 3). For this case, the rainfall intensity ranged from 11 to 14 mm/hr with higher values for PM and lower for DC. Therefore, this depth could be utilized to determine water quality volume (WQV) for all watersheds for each season. The heavier storms create larger runoff, which makes storm water quality sampling very difficult. Also, due to the very infrequent nature of these storms, most of the storm water quality sampling was done for smaller storm events that were regular and reliable. Since this was the case, the minimum rainfall depth that could possibly produce surface runoff was more important in this respect.

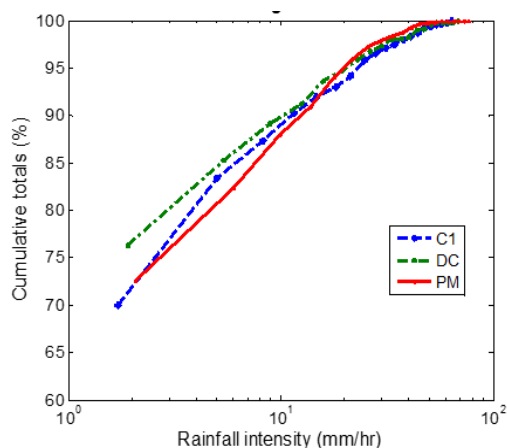
Total Rainfall Duration versus Number of Occurrences

Figure 4 shows the percentage of occurrence for all events with respect to total rainfall duration for both seasons. The event duration is categorized in terms of two hours or less, up to 15 hours or more of rainfall. The percentage of occurrence of rainfall events below two hours was almost double for June-November in compared to December-May, which might be due to the general nature of occurrence of rainfall in Las Vegas. In general, the summer rain events were assumed to last up to six hours, while the winter events extended up to 24 hours. The higher occurrences might also be

the result of the storm passing to other rain gauge stations after a certain lag time in the watershed. For rainfall durations more than two hours, the number of events was higher during the period of December-May. The number of events showed a decrease between the intervals of 2 to 12 hours after which they increased slightly for events greater than 12 hours in duration. A similar pattern was observed for all watersheds and for both seasons. Almost 12 percent of the events occurred beyond 12 hours in the wet season and 4-8 percent in dry season. C1 and DC had almost equal numbers of events in the interval of 2 to 8 hours, whereas DC showed comparatively higher events for durations of more than eight hours. Most of the rainfall events occurring for a longer duration and with higher intensities could become a major storm event. Thus, for rainfall of higher intensity, C1 and DC (in particular, DC with a dominating effect) could have higher effects on flooding when compared to PM Wash.



(a) Rainfall (mm)



(b) Rainfall Intensity (mm/hr)

Figure 3. Normalized Distribution of Cumulative Number of Rainfall Events versus Rainfall and Rainfall Intensity

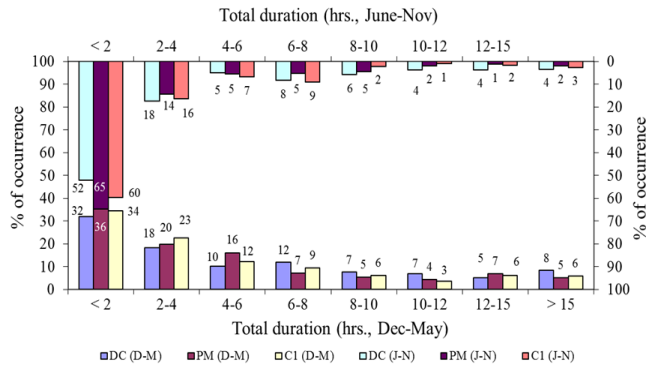
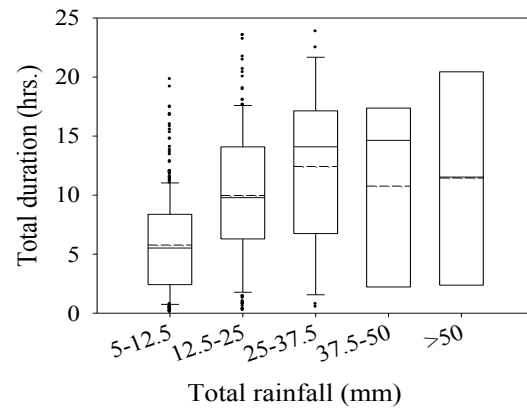


Figure 4. Total Rainfall Duration versus Percentage of Occurrence of Total Events for DC, PM, and C1

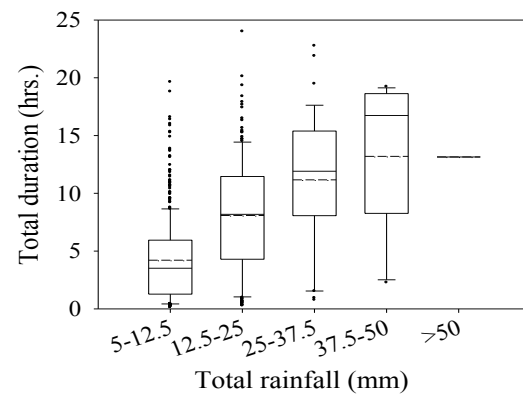
Total Rainfall versus Total Duration

The box plots in Figure 5 relate the range of rainfall depths with total rainfall duration for each watershed. Box plots show the means (indicated by dashed lines in the box), for first quartile ($x_{.25}$), median ($x_{.50}$), third quartile ($x_{.75}$), and interquartile (IQR, $x_{.75} - x_{.25}$) ranges. Any data observation which lay more than 1.5 times IQR lower than the first quartile, or 1.5 times IQR higher than the third quartile, was considered an outlier [9]. The outlier indicated the point to be of some different value than the remaining set, faulty data, or erroneous procedure. Also, the median value represented the data better than the mean during higher data variation. The median value showed an increase in rainfall duration with respect to increase in rainfall depth up to 50 mm of rainfall for all watersheds. The outliers within the range of 5 mm to 25 mm of rainfall showed larger variation in the total rainfall duration. The longer duration events (acting as outliers) resulting in a low total rainfall depth implied that the intensity of these events was very low and, thus, represented winter storms.

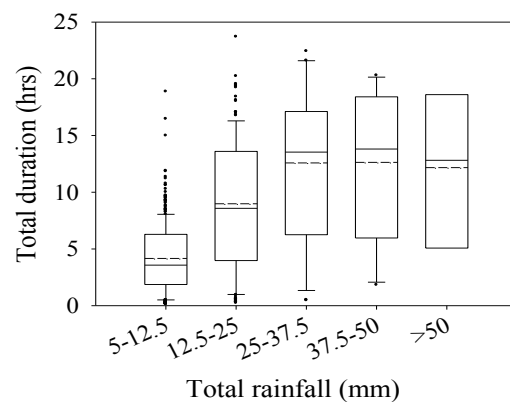
As noted earlier, the contribution of smaller events was higher in both seasons for all watersheds. The rainfall duration was higher where DC was within a range of 50 mm of rainfall. However, the higher interquartile range for DC showed a larger variation in the observed data when compared with PM and C1. The events above 25 mm of rainfall showed almost equal duration for DC and C1. The higher intensity events occurring for a longer duration resulted in higher runoff. This implied that the runoff produced was higher for DC and C1 than it was for PM. The comparatively higher number of events observed in PM between 37.5 mm and 50 mm might be due to the fact that there were simply larger numbers of major storm events occurring in the watershed. For PM, the lack of observations of more than 50 mm of rainfall indicated very few events occurring with more than 50 mm of rainfall depth.



(a) Duck Creek (DC)



(b) Pittman Wash (PM)



(c) C1 Channel

Figure 5. Box Plots of Total Rainfall versus Total Duration

The box plots in Figure 6 (a-f) showed the seasonal variation in total rainfall with total duration for DC, PM, and C1. The median duration for storm events during the winter period was approximately the same for C1 and DC, which implied that there was an equal contribution of higher runoff from both watersheds. A larger variation in the event duration for higher rainfalls in DC and C1 might be due to different nature and intensity of storms occurring in summer and winter periods. The difference between the mean and median values for a range of rainfall depths in Figure 6 also shows the higher variation in rainfall duration. The median duration was found to be considerably higher for winter periods than summer periods for the same range of rainfall in each watershed. Very few observations were seen in the higher range of rainfall (>37.5 mm) during summer months; these were associated with very short duration events (3 to 8 hours). This implies that summer storms are typically local thunderstorms and heavy rainfall lasting for shorter periods, whereas the winter rainfall lasts for longer periods.

Water Level Analysis: Total Rainfall versus Total Water Depth

The box plots shown in Figures 7-9 display the change in water depth with respect to total rainfall depth. Similar to rainfall analysis, the observed rainfall in the range of 5 mm to more than 50 mm were categorized into five classes: 5-12.5, 12.5-25, 25-37.5, 37.5-50, and >50 mm. The rainfall events below 5 mm were not considered in this analysis as they were assumed to play a very small role in runoff generation. The calculation was based on total rainfall events throughout the year. A higher number of outliers was observed in DC (see Figure 7), due to a larger variation in the observed storm events; these outliers may also represent major storm events that occurred in the watershed. The water depth was higher for station 4683, as compared to station 4633, for all ranges of rainfall; this was mainly due to the location of stations in the watershed. Station 4683 was lo-

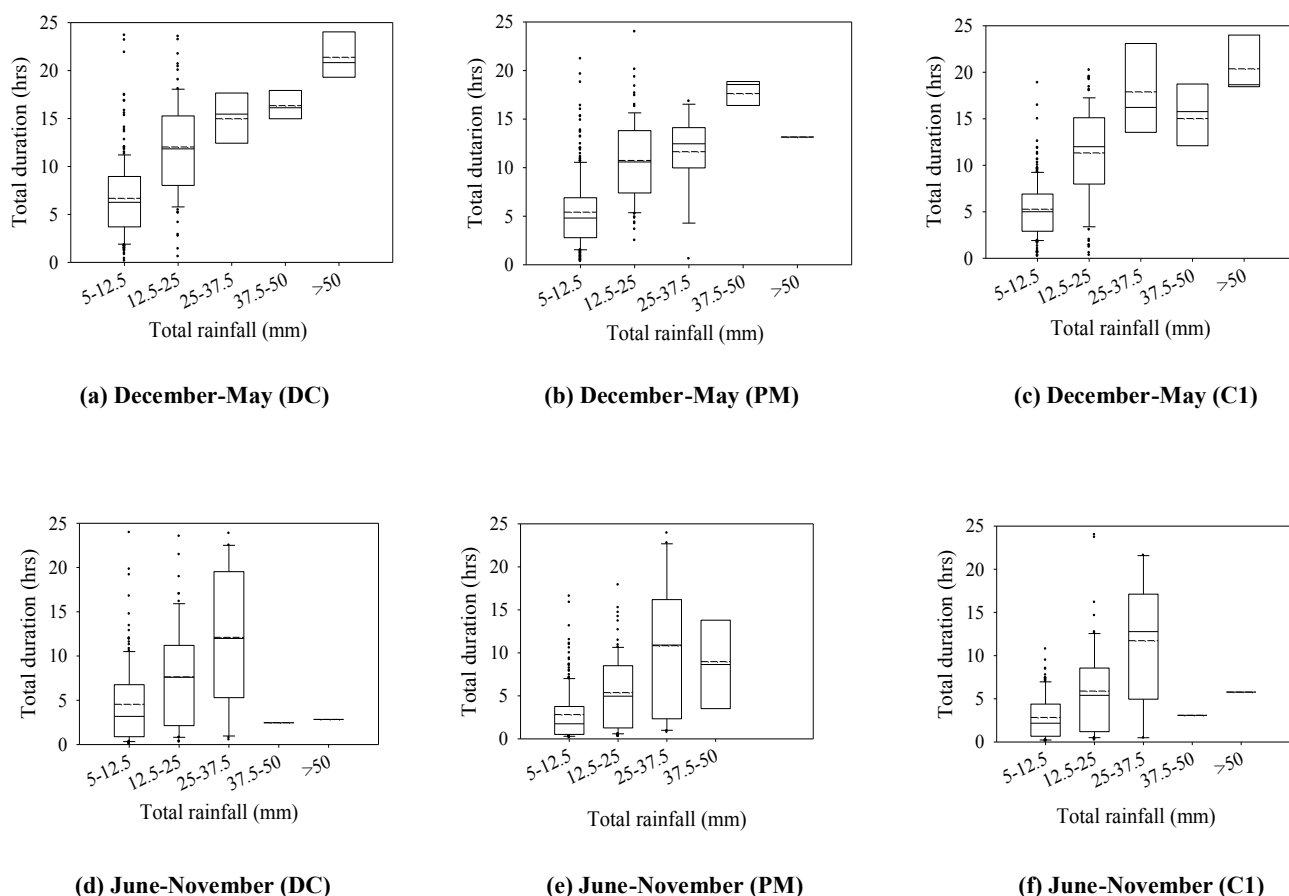
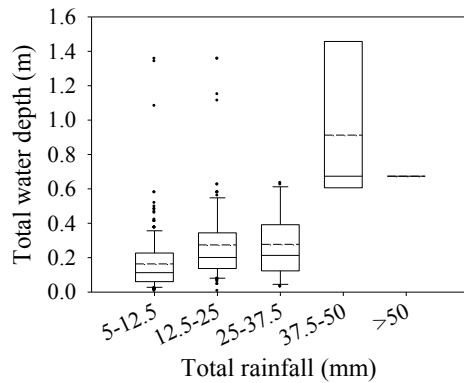
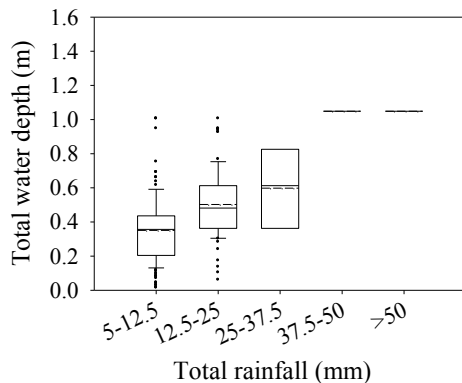


Figure 6. Seasonal Variation in Total Rainfall versus Total Duration for DC, PM, and C1

cated far downstream, while station 4633 was further upstream, as shown in Figure 2. Thus, the higher catchment area was contributing more flow at station 4683 than station 4633, thus increasing water level for the same range of rainfall.



(a) Station ID 4683

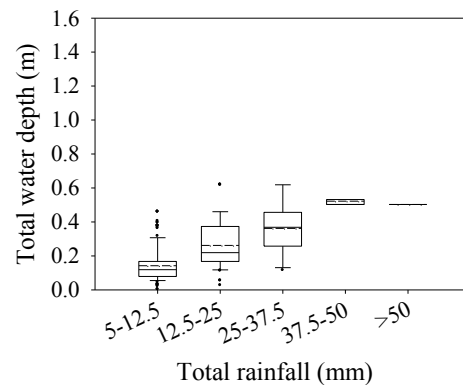


(b) Station ID 4683

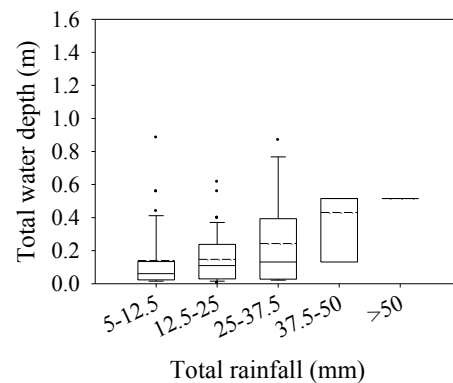
Figure 7. Total Rainfall versus Total Water Depth for Water Level Stations in DC

The water depth rainfall plot shows a linear increase of rainfall up to a range of 25.0-37.5 mm, which then increased drastically for intervals greater than 37.5 mm at station 4633. This also indicated the occurrence of higher storm events in the watershed. For station 4683, a similar water depth of more than 37.5 mm of rainfall also demonstrated a lack of available water level data in that range of rainfall. Both stations followed a similar pattern of increasing water depth with increasing total rainfall. Thus, the observed higher water depth produced a higher runoff when the rainfall event was greater than 37.5 mm in the DC watershed. The nature of PM shown in Figure 8 followed a pattern similar

to DC. Station 4748, located downstream, displayed a higher water depth (both mean and median) as compared to station 4733 located upstream over the same range of rainfall. The additional catchment area was contributing flow towards station 4748, causing an increase in water depth. As in DC, very few water level data were present for more than 50 mm of rainfall. The median water depth in PM was lower than DC for the same range of rainfall; this implied that the runoff of PM was comparatively lower than DC. A higher variation in water level was observed for C1 (see Figure 9). The median water depth was lower than DC and almost similar to PM up to 37.5 mm of rainfall. For higher rainfall, the water depths in C1 were similar to DC, but more than PM. This leads one to believe that the storm runoff in C1 was similar to PM but smaller than DC for a rainfall less than 37.5 mm; however, some storm events showed similarities between DC and C1 in events greater than 37.5 mm.



(a) Station ID 4748



(b) Station ID 4748

Figure 8. Total Rainfall versus Total Water Depth for Water Level Stations in PM

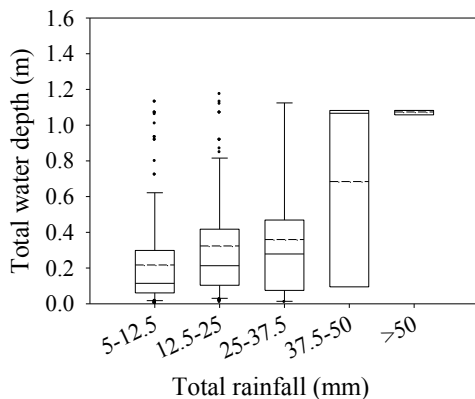


Figure 9. Total Rainfall versus Total Water Depth for Water Level Station ID 4783 in C1

The higher water depth observed as outliers for low rainfall suggested an additional effect of rainfall upstream in the watershed. The increase in velocity, runoff volume, and the time of concentration increased peak discharge flow rates and decreased mean lag time. In short, this generated sudden floods occurring for a very short duration [10]. Low rainfall coupled with a high flow depth observed at a single station might have been surcharged by an event coming from upstream in the basin. This phenomenon included the rare flood events that greatly increase the peak discharge and water levels in arid watersheds, as discussed by Caraco [4].

The difference in land use between these three watersheds also played a significant role in the contribution of runoff. Urbanization increases storm water runoff generation as well as peak flow, while decreasing base flow from subsurface water [11]. Based on land use, DC is highly urbanized—possessing various residential and commercial districts. Most of the area of PM consists of special districts, which are defined as the broad range of developments to serve citizens and that include recreational vehicle parks, public facilities, general highway frontage, urban village, limited resorts and apartments, and open spaces [12]. Only a very small area was developed for commercial and residential uses. C1 possessed fewer catchment areas than other watersheds and had no commercial districts. Half of the C1 basin was covered by special districts and the other half by rural open land and residential districts.

Although the direct calculations of Curve Number (CN) are not shown here, the average weighted curve number (WCN) for each of these watersheds was calculated as 84.373 for DC (min 57.610; max 95.690), 86.219 for PM (min 66.959; max 95.959), and 79.100 for C1 (min 0.000;

max 93.599). The higher the WCN, the higher the percentage of impervious surface and the higher the quantity of surface runoff. Therefore, it can be expected that DC and PM generate comparatively higher runoff in a short period of time for the same rainfall event occurring over all three watersheds.

Water Depths for High and Low Rainfalls

As previously mentioned, 90 percent of rainfall events were found to be less than 18 mm for all three watersheds. The runoff depth and storm durations were highly variable between storm events. Thus, similar to rainfall analyses, the median depth could be considered significant from a design point of view. As shown in earlier box plots, for 18 mm of rainfall, the median water depth corresponded to an approximate value of 0.24 m, 0.48 m, and 0.24 m for C1, DC, and PM, respectively. This depth represents the higher threshold value of runoff depth that needs to be considered during the design of drainage channels. The storm events occurring in each watershed with water depths greater than the depths indicated above correspond to storm events that may cause devastating damages.

The nature of surface runoff is small at the beginning of the storm event. As it surpasses the initial abstraction, the total rainfall exceeds the total losses that cause an increase in the runoff. This can be considered the lower threshold value of rainfall for runoff generation. The threshold value is usually a little bit higher than the mean rate of rainfall occurring in a particular area [13]. Knowledge of threshold value of runoff for low flows is also important during dry seasons with very low rainfall. The threshold values for low flows can be determined at a point where there is a sudden change in water depth from previous rainfall events; this point indicates the starting point of runoff generation in the watershed.

A step change analysis was performed using TREND software [14] to find the threshold value of rainfall for low flows. The rank sum test and student t-test were also applied during the step change analysis of rainfall and water depth to test the statistical significance of observed rainfall for low water depths. From the test, the threshold values of low rainfall events for C1, DC, and PM were found to be 2 mm, 4 mm, and 1 mm, respectively. Figure 10 shows the plot of water depth for lower threshold values of rainfall for each watershed. The corresponding water depth for low rainfall after which the runoff might start were found to be 0.06 m, 0.08 m, and 0.04 m for C1, DC, and PM, respectively. The test data showed a statistically significant increasing step at the 99-percent significance level for the rank-sum test and the 95-percent significance level from the student t-test.

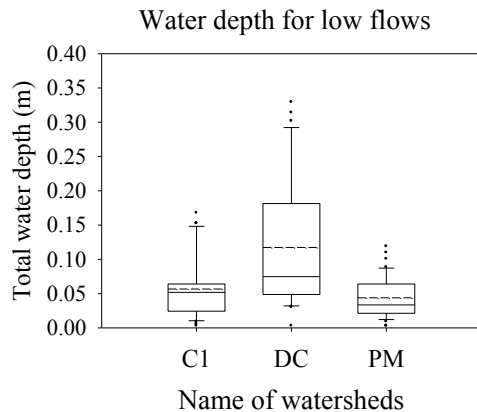


Figure 10. Determination of Water Depth for Low Rainfall Events (low flows) in Each Watershed

Water Depth for Total Rainfall and Total Duration

The contour plots of water depth with respect to total rainfall and total rainfall duration are shown in Figure 11. The observed maximum water depth for PM was very low as compared to DC and C1. The highest magnitude of water level was in the range of 1.1 m for C1, 1.2 m for DC, and 0.6 m for PM. In general, higher runoff was generated by higher rainfall and vice versa; a study performed in three urban watersheds in Minnesota showed an increase in runoff as a result of an increase in annual precipitation, the size of the maximum storm, and a decrease in the number of days between the storms [15]. However, there was no fixed pattern showing a significant increase in water depth for DC. C1 and PM somewhat followed theoretical concepts of higher runoff generation, due to higher rainfall events for longer duration.

The spatial and temporal variability of surface runoff depend on many factors such as soil type and its imperviousness, catchment area, rainfall intensity and duration, and other physical characteristics. In addition to these factors, human activities and ongoing urbanization also increase runoff volume and peak flow rates [16]. The maximum water depth observed for DC was higher than other watersheds, which signifies the occurrence of few major storm events. The higher flow depth coupled with lower rainfall depth and duration in all watersheds indicated the variation of rainfall patterns within the watershed. This variation implied that the cumulative rainfall over longer durations in the upper portions of the watershed was influencing the change in water depth downstream, causing higher water depth for lower rainfall. This might also indicate that the rainfall measured near the water depth station had a greater

influence on increasing water depth, as compared with other rain gauge stations recording the same rainfall but which were located further upstream in the watershed, thus causing higher water depth in a lower duration.

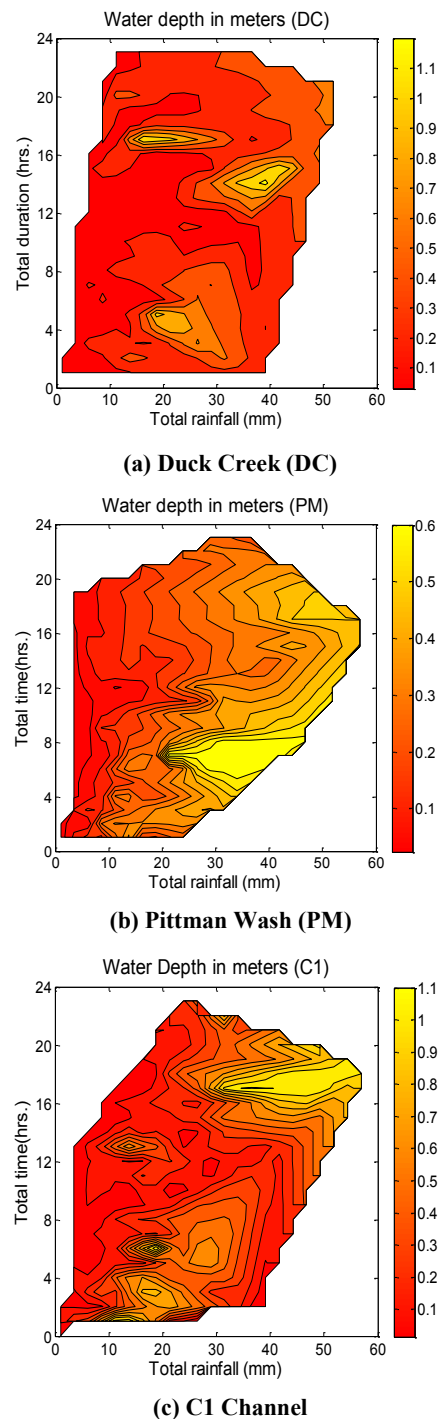


Figure 11. Contour Plots of Water Level (depth in meters) With Respect to Total Rainfall and Total Duration

Although the results are not shown and discussed separately, while relating total rainfall between the rain gauges for the same duration, some rain gauges showed a higher correlation with the most downstream rain gauge station. The most downstream station was taken as a reference, due to the fact that the water depth utilized for this analysis was also measured at the same location. The case is somewhat similar to the results presented by Smith et al. [6]; a greater correlation between the rainfall amounts measured at different gauges does not necessarily mean a higher correlation between the runoff volumes.

Conclusions

This study followed a data-driven analysis of rainfall and water depth despite the lack of streamflow data in each watershed. A water balance analysis is usually performed to determine the threshold value of rainfall for a specific region to get more accurate results; practically, though, it was not possible in this study, as the gauge stations measured only rainfall and water depth. Based on this study, the total number of rainfall events, classified into two seasons, showed a higher percentage of storm events at lower water depths, which did not produce discernable runoff.

Based on rainfall and water depth, DC showed higher impact on runoff depth compared to other watersheds. Including rainfall duration, DC and C1 possessed higher magnitudes of water depth; DC was slightly higher than C1 and almost double the water depth observed in PM. From this result, it was concluded that DC had a higher effect on flood-related damages during larger storm events compared to other watersheds. This implies that higher attention should be paid to this watershed during the planning, design, and construction of flood control facilities.

The concepts adopted in this study can be transferred and implemented elsewhere with various data constraints, since the runoff generation was mainly dependent on rainfall. It can be suggested that if there is availability of streamflow data and more statistical relationships can be built in between the observed parameters, it would increase the accuracy of the obtained results and conclusions.

Acknowledgements

The results presented in this paper are a part of the WAITS project funded by the USGS for Clark County, Nevada. The author gratefully acknowledges the help provided by various staffs from Clark County Regional Flood Control District.

References

- [1] Clark County (2011). Living in Clark County. Retrieved July, 2011, from www.accessclarkcounty.com.
- [2] Reilly, J. A., & Piechota, T. C. (2005). Actual Storm Events Outperformed Synthetic Design Storm: A Review of SCS Curve Number Applicability. *ASCE Conference Proceedings of World Water and Environmental Resources Congress*. doi: 10.1061/40792(173)95.
- [3] Clark County Regional Flood Control District. (2006). History of Flooding in Clark County, Flood Threat Recognition System (FTRS). Retrieved Nov. 1, 2007, from <http://www.ccrfcd.org/03-history.htm>.
- [4] Caraco, D. (2000). Stormwater Strategies for Arid and Semi-Arid Watersheds. *Watershed Protection Techniques*, 3(3), 695-706.
- [5] PBS&J, G. C. Wallace, Inc., & Louis Berger Group. (2002). Las Vegas Valley Flood Control Master Plan Update. (Master Plan Update).
- [6] Smith D., Li, J., & Banting, D. (2005). A PCSWMM/ GIS based water balance model for the Reesor Creek Watershed. *Atmospheric Research*, 77, 388-406.
- [7] Geographic Information System Management Office. (2007). Clark County Regional Flood Control GIS Database. Retrieved July 1, 2008, from <http://gisgate.co.clark.nv.us/gismo/gismo.htm>
- [8] Iowa Department of Natural Resources. (2007). 2C-2 Rainfall and Runoff Analysis. In *Iowa Stormwater Management Manual* (Version 1; Feb. 19, 2007).
- [9] Moore, D. S., & McCabe, G. P. (1999). *Introduction to the Practice of Statistics*. (3rd ed.). New York, U.S.A.: W. H. Freeman.
- [10] Schonwiese, C. D., & Rapp, J. (1997). *Climate Trend Atlas of Europe Based on Observations 1891–1990*. Dordrecht: Kluwer Academic Publishers.
- [11] Olivera, F., & Defee, B. B. (2007). Urbanization and Its Effect on Runoff in the Whiteoak Bayou Watershed, Texas. *Journal of American Water Resources Association*, 43(1), 170-182.
- [12] Title 30 Development Code (2014). Chapter 30.36/30.40: Zoning Base Districts and Maps. Retrieved Nov. 5, 2014, from http://www.clarkcountynv.gov/depts/comprehensive_planning/zoning/pages/title30.aspx.
- [13] Short, D. A., Wolff, D. B., Rosenfield, D., & Atlas, D. (1993). A Study of the Threshold Method Utilizing Raingage Data. *Journal of Applied Meteorology*, 32, 1379-1387.
- [14] TREND software. Retrieved Feb. 19, 2008, from <http://www.toolkit.net.au/Tools/TREND>.

-
- [15] Brown, R. G. (1988). Effects of Precipitation and Land Use on Storm Runoff. *Water Resources Bulletin*, 24(2), 421-426.
- [16] Brilly, M., Rusjan, S., & Vidmar, A. (2006). Monitoring the Impact of Urbanization on the Glinscica Stream. *Physics and Chemistry of the Earth*, 31, 1089-1096.

Biographies

ANIL ACHARYA is an assistant professor in the Civil and Mechanical Engineering department at Alabama A&M University (AAMU). He earned his B.S. degree (Civil Engineering, 2003) from Tribhuvan University, Nepal; his M.S. degree (Water Resources Engineering, 2006) from the Asian Institute of Technology, Thailand; and Ph.D. (Civil and Environmental Engineering, 2011) from the University of Nevada Las Vegas, Nevada. Dr. Acharya is currently teaching and conducting research at AAMU. His major interests are in modeling of water resources systems, streamflow forecasting, climate change studies (hydroclimatology), streamflow disaggregation, water quality, and GIS applications in water resources. Dr. Acharya may be reached at anil.acharya@aamu.edu

THOMAS C. PIECHOTA is the Vice President for Research and Economic Development, and a professor in the Civil and Environmental Engineering department at the University of Nevada Las Vegas (UNLV). Dr. Piechota is currently teaching and conducting research at UNLV. His major interests are interdisciplinary in the areas of surface water hydrology, storm water quality in urban environments, hydroclimatology, water resources planning, and Geographic Information Systems. Dr. Piechota may be reached at thomas.piechota@unlv.edu

NONLINEAR AND BUCKLING ANALYSIS OF STEEL TRANSMISSION TOWERS

Mustafa Mahamid, University of Illinois at Chicago; Cenk Tort, Miteng, TURKEY; Mutaz Al Hijaj, University of Illinois at Chicago; Adeeb Rahman, University of Wisconsin-Milwaukee; Faris Malhas, Central Connecticut State University

Abstract

Transmission line structures are often constructed from steel lattice structures, due to their ease of assembly and due to their light weight, which results in relatively small foundations. These structures are typically constructed from standard single or built-up steel angles all of which are assembled and connected together by a steel gusset plate. To better understand the true behavior of the tower, full-scale tests are usually performed under design loads. These tests are required to investigate the overall behavior of the structure and the behavior of the steel connections that connect multiple members together.

This practice is essential before producing multiple transmission towers for a single project. However, full-scale tests are expensive to build and time consuming, which may cause delays in project schedules. As an alternative to full-scale testing, and as an approach to prevent certain unexpected behavior of these structures, the authors of this current research project proposed detailed finite element models to determine the capacity of these structures and to investigate the global and local buckling behaviors, which are hard to predict using full-scale testing. Previous experimental results were used to validate the finite element models. The finite element modeling of the members used beam elements for the steel angles with pin connections for the joints; the commercial finite element package ABAQUS was used for the modeling.

Introduction

Transmission line towers are important structures, and their design requires special techniques and criteria. The safety of these structures is essential in order to provide continuous energy transfer from power plants to communities. Due to electrical clearances and attachments, transmission line towers often have complex geometries. A typical tower structure has five main parts comprised of the leg, body, neck, cross-arms, and grounding tower. The parts that make up a typical tower are illustrated in Figure 1. The leg and body parts are often designed in a modular pattern with different heights so that, depending on the topology of the region, a family of towers can be generated. This variation of geometry makes the design even more complicated.

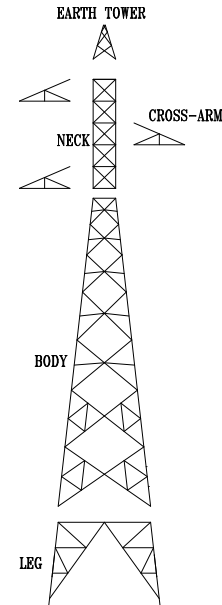


Figure 1. Geometry of the Transmission Tower

In a typical electrical line, transmission tower types also vary based on their usage. Terminal towers are placed at the end of the line before the conductors are extended to the substation. Angle or tension towers are used where a transmission line changes direction at an angle. Suspension towers serve to carry the conductors and are located between the angle towers. The weight of terminal and angle towers is often large, due to the high level of forces to which they are subjected. The role of the suspension towers is to carry its own weight and that of the conductors. Though their weight is relatively small, they constitute the major portion of the electrical line. Therefore, it is the suspension tower that is often subjected to design verification with load testing.

Background

The design of transmission towers is performed according to special design specifications. In the U.S. and around the world, ASCE 10-97 is considered to be the major design document; a new edition of ASCE 10 has been published recently; ASCE 10-15. The structural elements constituting the tower serve as truss elements and are assumed to carry axial loading only. Based on their slenderness, the elements are checked for allowable buckling and tensile stresses.

However, these assumptions may not always be valid for the following reasons: 1) the structural elements are often connected through more than single bolts and 2) if the clear distance of the angle flanges is not enough, the connections may have larger gusset plates, as can be seen in Figure 2.

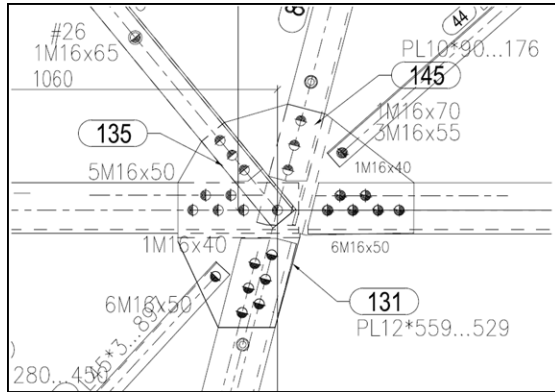


Figure 2. Bolts and Gusset Plates in Transmission Towers

Due to the uncertainty in their behavior, transmission towers should be analyzed using detailed finite element (FE) modeling. There exist several studies in the literature on the behavior of transmission towers. Battista et al. [1] conducted a study to investigate the stability of transmission towers under wind loading. The transmission tower and two adjacent spans were modeled using frame elements. Frame elements were preferred in order to study the effect of bending stresses. A dynamic modal analysis of the tower was performed under wind loading. In the models, the isolators were all modeled as double-hinged suspension rods and found to have significant effects on the tower behavior.

Rao et al. [2] studied the nonlinear behavior of transmission tower structures. A total of five tower structures were analyzed for both geometric and material nonlinearities. The structural elements were modeled using beam elements but, for some of the towers, plate elements were also employed. The tower structures were selected from among those subjected to full-scale experimental testing. Good correlation was obtained with the test and computational results. Al-Bermani and Kitipornchai [3] developed a beam-column element that included both geometric and material nonlinearities. Four types of towers used in practical applications were analyzed and a good correlation was obtained. This study did not present any local strain or stress data.

Matheson and Holmes [4] studied the response of transmission line cables under turbulent wind loading. A series of wind velocities were generated using Monte Carlo simulation techniques. The analysis of the cable was performed using the finite difference method. The loads exerted on the support and line deflections were compared to the available

methods in the literature and identical results were obtained. Rao et al. [5] investigated the elasto-plastic response of transmission towers. Tested towers were modeled by beam elements and plate elements. The effect of a bracing pattern and redundant members on the tower response was studied. Numerous software packages were identified in the literature for conducting finite element analysis of truss structures. ABAQUS [6] offers advanced beam and plate elements to perform both geometric and material nonlinear analysis. The graphical tools and user interface allows fast generation of the finite element mesh of transmission towers.

Weilian et al. [7] presented a two-step analytical study to predict the vertical member instability in transmission towers. The study utilized wavelets to determine buckling-critical regions in the first step. In the second step, the buckling-critical members were detected based on a strain-energy rate of change. The analysis method was tested using a refined finite element model of a previously tested tower. In addition, the method was applied to a small-scale experimental tower. The tower was instrumented with accelerometers. For both cases, the damage was accurately predicted. Alberman et al. [8] proposed a finite element formulation to conduct both geometric and material nonlinear behavior of transmission towers. Geometric nonlinearity was simulated using an Updated-Lagrangian approach. For material nonlinearity, a concentrated plasticity technique was used where inelasticity was monitored at the element ends. A full-scale, 275 kV tower was modeled and analyzed under test load cases. The buckling response of the tower during the test was predicted with good accuracy.

Rao and Kalyanaraman [9] proposed modeling of transmission lines structures with MSC-NASTRAN software. A model was developed to account for eccentricity, joint rigidity, and nonlinear effects. Design recommendations were provided for redundant members. Alberman et al. [10] conducted both experimental and analytical studies on improving the structural performance of tower structures. Specimens of tower substructures were tested under lateral loads with different bracing patterns. The bracing type with superior performance was applied to an existing tower structure. The tower structure was analyzed nonlinearly using AK-Tower software and the performance of the bracing system was verified and a larger stiffness was obtained.

Yasui et al. [11] investigated the response of transmission lines under wind action. The transmission line was modeled as a whole including both towers and conductors. Response characteristics of the line were evaluated with both time-history analysis and power-spectrum density. Jamaledine et al. [12] studied the ice-shedding phenomena of transmis-

sion lines and its effect on the towers. Small-scale specimens were tested by dropping dead weights from the conductors to simulate ice falling. The specimens were also modeled using ADINA software. The experimental response was accurately predicted by conducting both static and dynamic analyses. McClure and Lapointe [13] examined the response of tower structures under dynamic effects such as ice-shedding and conductor breakage. Both towers and conductors were modeled using ADINA software with linear elastic materials. A dynamic analysis was performed and response characteristics such as cable tensions and tower element forces were reported.

Nonlinear Finite Element Modeling of the Transmission Tower

Before performing a buckling analysis for the transmission towers designated as 1L (1st Longitudinal tower) and DS (Direct current Suspension tower) below, the load tests were simulated in the finite element software ABAQUS v. 6.13-2, and the loads applied to the tower during the load test were applied to the finite element model for comparison and verification. Similar to the load test, five load cases were studied. A three-dimensional model of the tower was created with beam elements type B31 with the degrees of freedom shown in Figure 3. This element type was formulated with the approximation that it can support large axial strains, large rotations, and transverse shear strain. This can be seen in the virtual work equation, Equation (1), to derive element stiffness in ABAQUS [6].

$$\delta\Pi = \int_l dl \int_A \left(\sigma_{11} \delta e_{11} + \tau_{(\alpha+1)1} \delta \gamma_{(\alpha+1)1} \right) dA \quad (1)$$

where,

- $\delta\Pi$ = virtual work
- dl = infinitesimal length
- dA = infinitesimal area
- σ = axial stress
- e = axial strain
- τ = shear stress
- γ = shear strain

The element type used for modeling was a beam element with a linear and quadratic interpolation. The formulation of this element allowed for transverse shear strain, which means that cross-section does not necessarily remain normal to the beam axis.

Structures of this type have connections at every joint; the connections at the tower investigated in this study were bolted connections that consisted of a gusset plate and bolts. The high number of connections added significantly to the

total weight of the tower. Therefore, it was important to consider the weight of the connection in the finite element model to reflect the actual behavior of the tower. This is also an important parameter for buckling analysis and the determination of buckling loads. The buckling load was obtained from the eigenvalue equation, Equation (2), provided in ABAQUS [6].

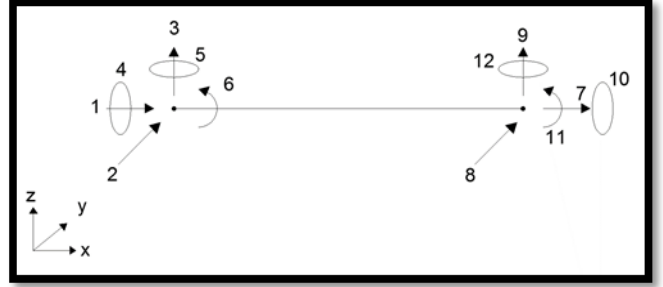


Figure 3. Degrees of Freedom for Beam Elements

$$\left(K_0^{NM} + \lambda K_{\Delta}^{NM} \right) v^M = 0 \quad (2)$$

where,

- K_0^{NM} = base state stiffness
- K_{Δ}^{NM} = differential stiffness
- λ = Eigen value
- v^M = mode shape

The weight of the gusset plates and bolts was calculated and applied as a mass at the connection locations. Additionally, an appropriate mesh was used after several trials to obtain accurate results. The overall finite element model is shown in Figure 4. The material properties used in the finite element model are the material properties used in the actual tests. Coupon tests were performed in order to obtain the stress-strain curves for the different member sizes used in the tower and the actual stress-strain curves were inputted into the model to simulate the actual behavior. The stress-strain curves for the different members are shown in Figure 5.

Verification of Experimental Results

The nonlinear finite element model was analyzed using the different load cases during the load test. Due to space limitations, the loading tree for one of the load cases is illustrated in Figure 6. In these tests, strain gauges were placed at critical members to monitor the stress-strain response during the tests. Figure 7 shows an example of strain gauges mounted on the structural members. Strain gauge results from the load tests were compared with the finite element results. Comparative results for some of the strain gauge results are shown in Figure 8. Two strain gauge results are

presented here: first, where the strain gauge was attached to the leg and, second, where the strain gauge was attached to the top cross-arm. The results showed good correlations between the load test and the finite element results. The results also showed that the tower behavior was elastic until the maximum load applied during the load test; this was observed from the load-test results and from the finite element results. It was obvious from the load-test results that the trend of the stresses was elastic at some location. Since transmission tower structures have bolted connections, significant slip was observed during the tests. Considering the gusset plate connections and scale of the towers, it was a challenge to reproduce the stress-strain response. The close correlation obtained from the analysis clearly showed that the finite element model successfully simulated the load distribution observed during the test.

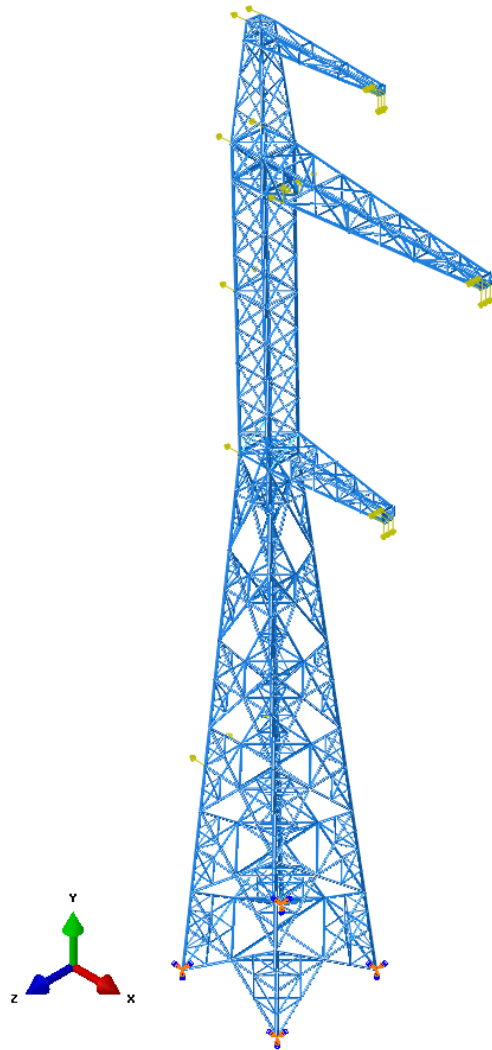


Figure 4. Finite Element Model of the Transmission Tower

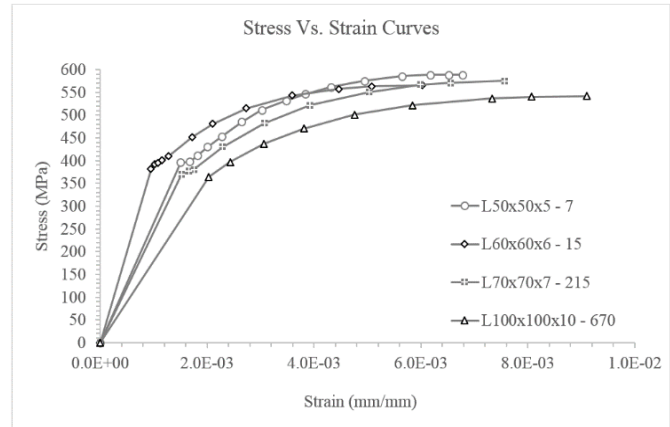


Figure 5. Material Properties Used in the Finite Element Models

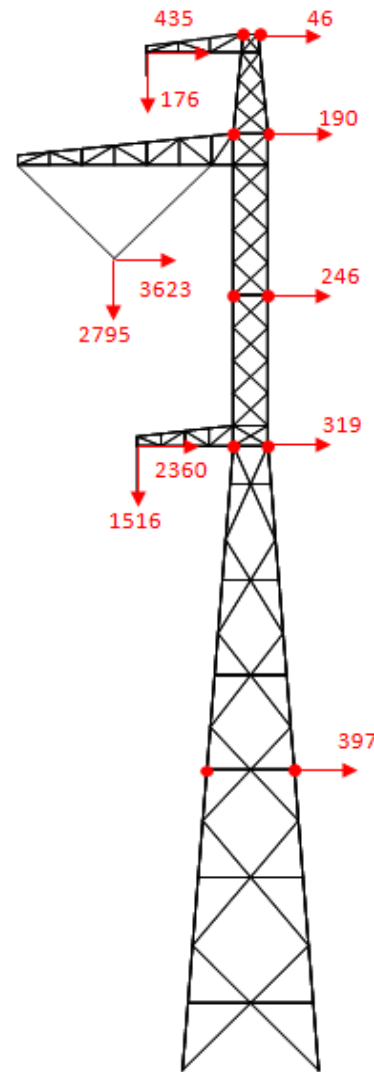


Figure 6. Load Case 1

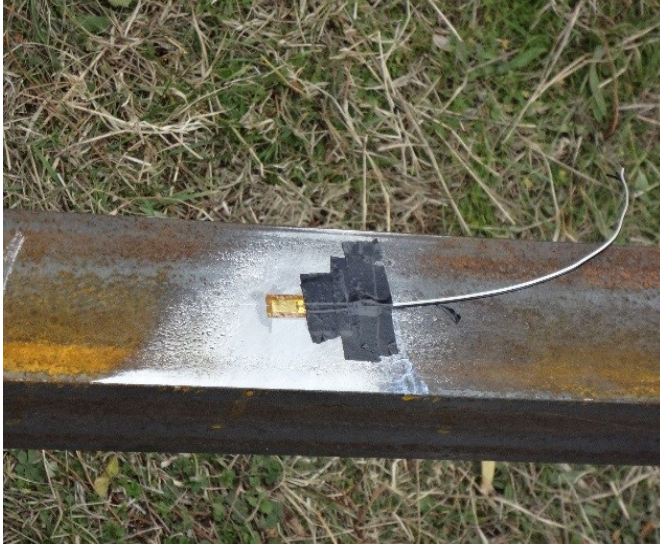
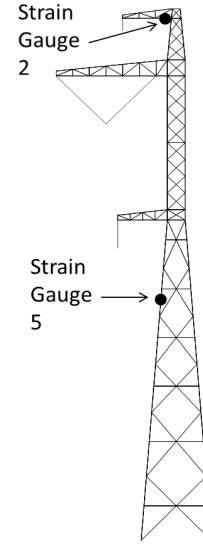


Figure 7. Strain Gauge Mounted on a Structural Member

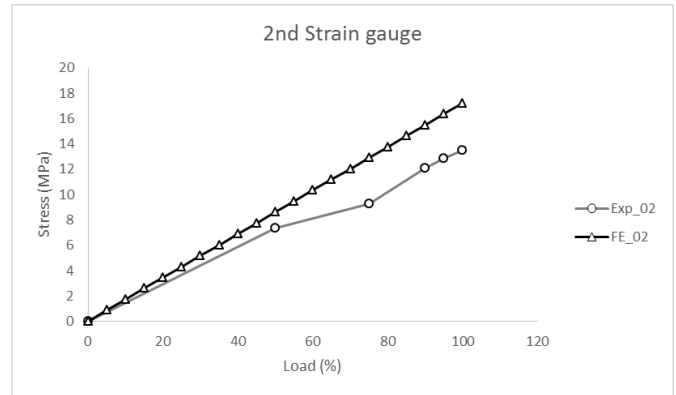
Figures 9(a) and 9(b), respectively, show a sample of stresses and deflections in the tower 1L for load case 1. It was observed that the model was adequate for performing buckling analyses for the transmission tower. The maximum stress was below the yield stress, which was also consistent with the test results. A visual inspection performed after the test indicated no visual sign of yielding in the members. The maximum deflection of the tower was determined to be 158 mm, which matched the design objective of maximum lateral deflection that was around $H/200 = 156$ mm.

The strain gauge locations and results of other strain gauges are presented in Figures 10-20, respectively. While some figures showed good agreement, others showed partial agreement, and some showed no agreement. Figures 11, 18, and 19 showed no agreement; however, it was obvious that these strain gauges at the specified load cases showed no known or expected behavior, where the finite element analysis results showed linear elastic behavior, which was consistent with the overall behavior of the tower or members during the test, though it is possible that the strain gauges might have been damaged or that readings were not recorded correctly.

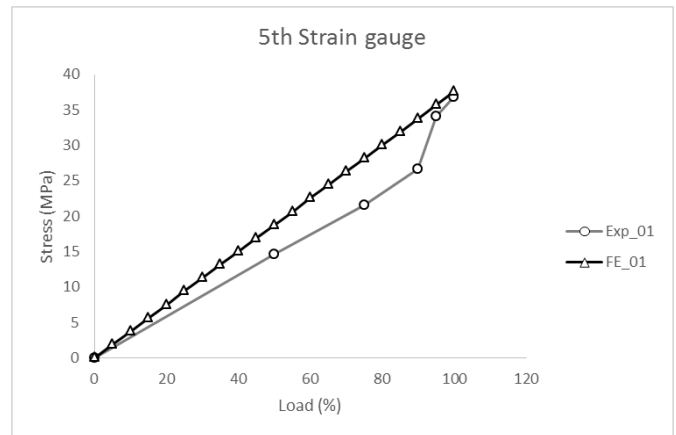
Figures 13 and 14 showed good agreement until a certain load during the test, then the strain gauge readings deviated from the expected behavior of the tower or members. It is worth mentioning that the members, where the strain gauges were mounted, were monitored during and after the load test and it was observed that there was no damage or plastic deformation at these locations, which is an indication that the strain gauges were damaged during the tests.



(a) Locations of Strain Gauges 2 and 5

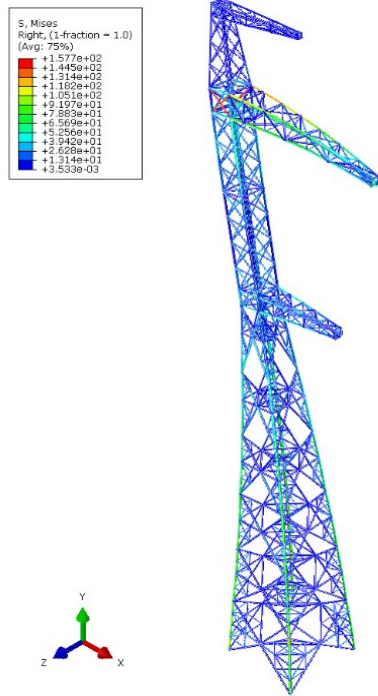


(b) Results for Strain Gauge 2, Load Case 2

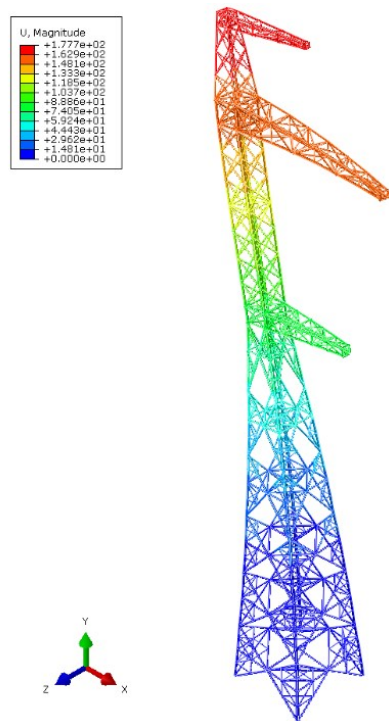


(c) Results for Strain Gauge 5, Load Case 1

Figure 8. Strain Gauge Locations and Results



(a) Stress Results



(b) Deflection Results

Figure 9. Stresses and Deflections in the Tower for Load Case 1

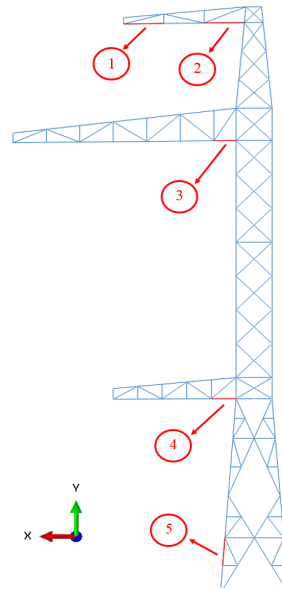


Figure 10. Strain Gauge Locations

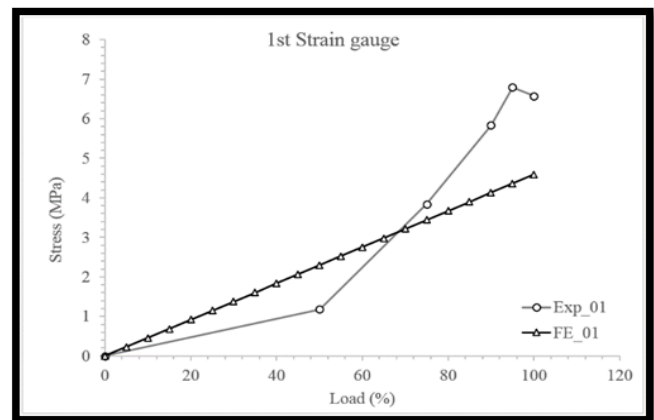


Figure 11. First Strain Gauge Results for Load Case 1

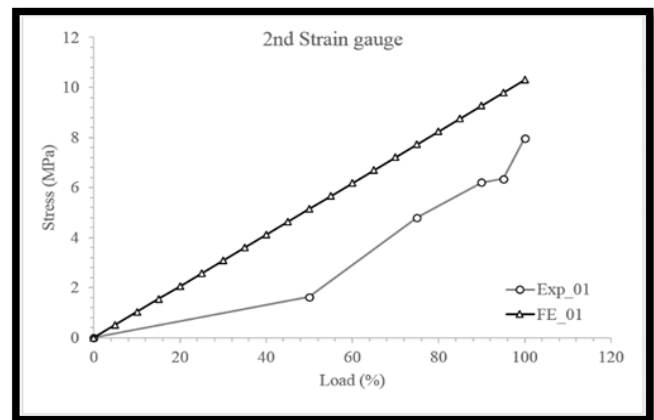


Figure 12. Second Strain Gauge Results for Load Case 1

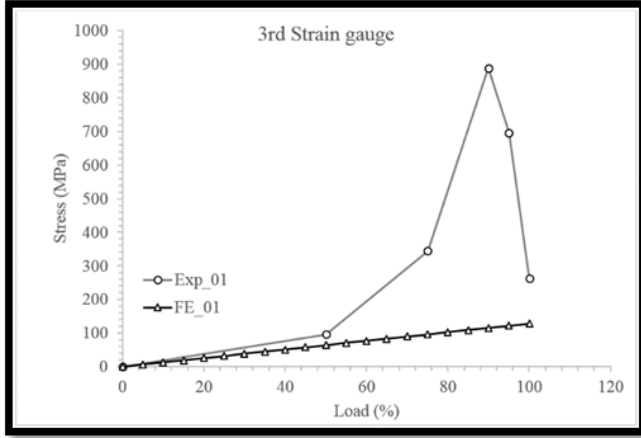


Figure 13. Third Strain Gauge Results for Load Case 1

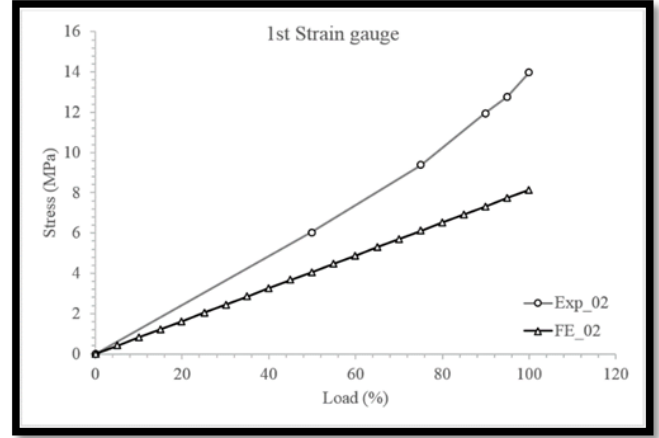


Figure 16. First Strain Gauge Results for Load Case 2

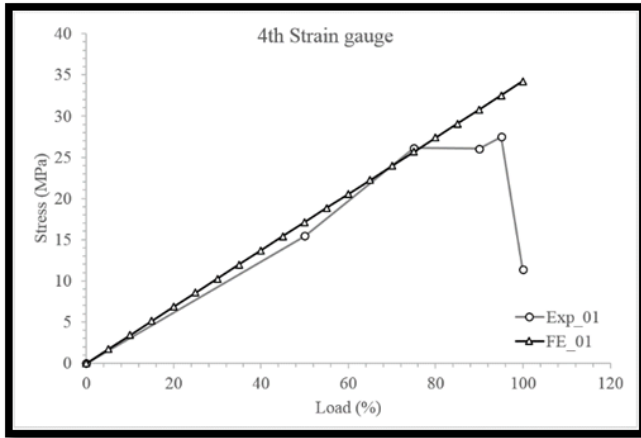


Figure 14. Fourth Strain Gauge Results for Load Case 1

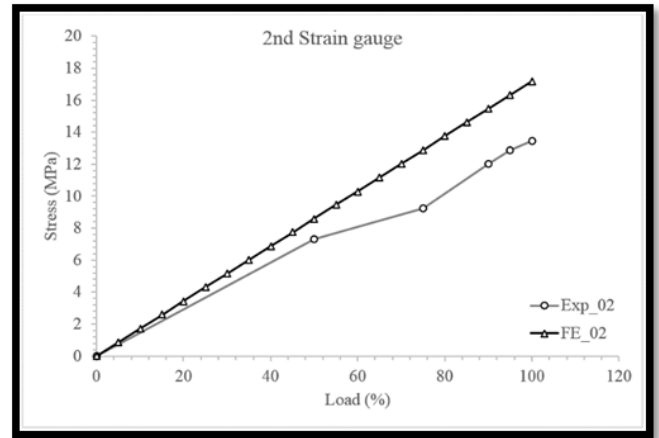


Figure 17. Second Strain Gauge Results for Load Case 2

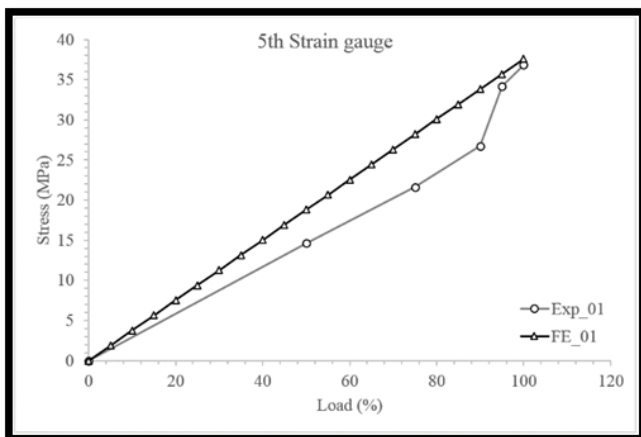


Figure 15. Fifth Strain Gauge Results for Load Case 1

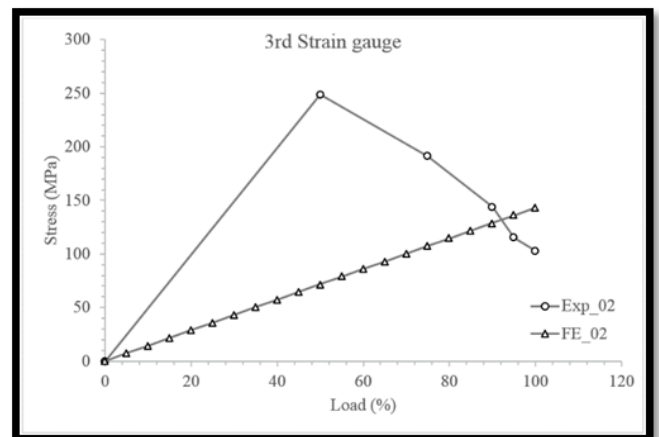


Figure 18. Third Strain Gauge Results for Load Case 2

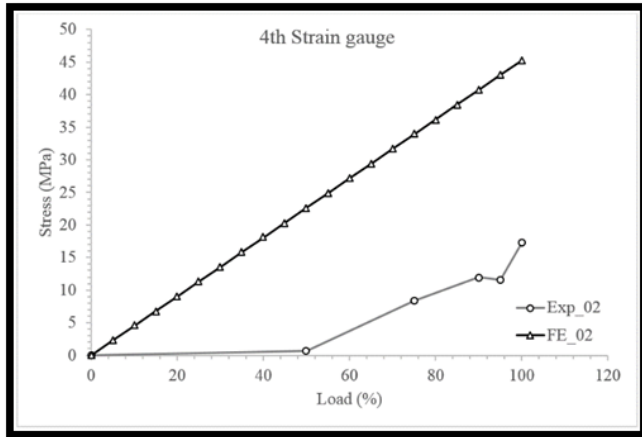


Figure 19. Fourth Strain Gauge Results for Load Case 2

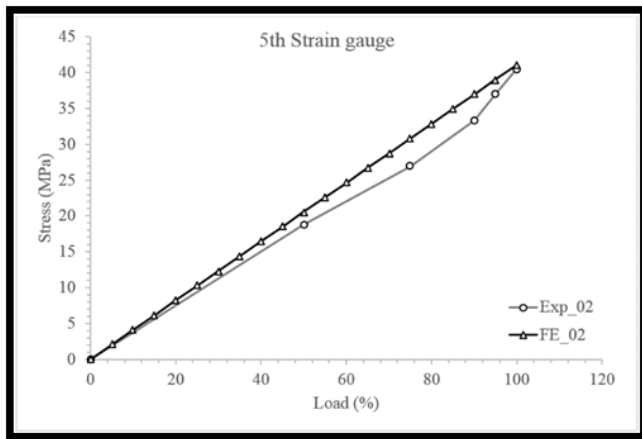


Figure 20. Fifth Strain Gauge Results for Load Case 2

Buckling Analysis of the 1L Transmission Tower

The load test is considered to be acceptable if the tower resists 100 percent of the load cases for a duration of one minute without any sign of failure. Failure of the transmission towers may be buckling globally or locally at certain sections; therefore, the study of buckling behavior is essential for such structures. The load-test results in this current study showed that the tower did not buckle globally during the load test, and local buckling at specific locations was not observed. Therefore, the finite element model enabled the determination of the buckling load and the buckling modes. After verifying the load-test results, a finite element buckling analysis was performed in order to determine the buckling load at different modes. The buckling analysis was performed by applying a unit load at the tip of the arm, as shown in Figure 21; the deflection of the tower in the first

mode is shown in Figure 22. The buckling results at different modes are shown in Table 1. The buckling results are presented as the ratio between the applied vertical load and the buckling load. It was observed from the buckling results that the ratio of the applied vertical load to the buckling load at the first mode ranged from 0.53 to 24.34 percent for load cases 1 to 5b. Similar results at the different modes for the different loads cases are shown in Table 2. The finite element results showed that the tower would buckle at much higher loads than the applied loads during the load test.

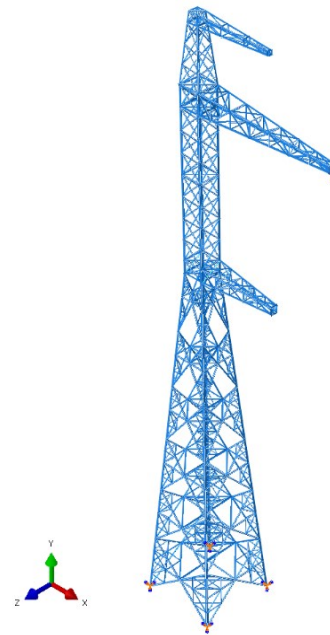


Figure 21. Applied Unit Load to Perform Buckling Analysis

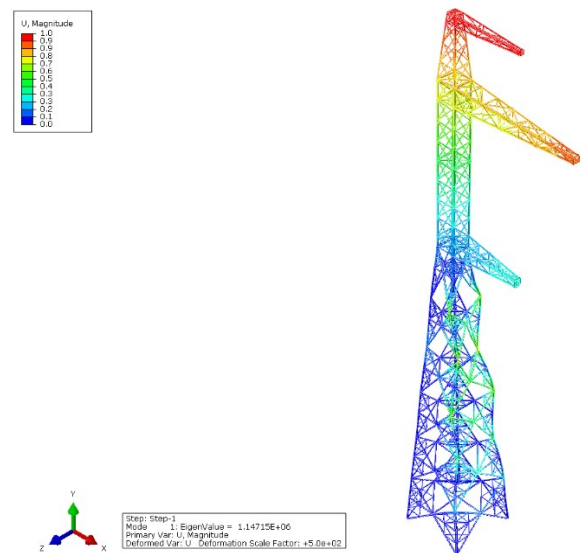


Figure 22. Deflections for the First Buckling Mode

Table 1. Vertical Load-to-Buckling Load Ratios for the First Five Buckling Modes

		Eigenvalue * 10 ⁶ (N)				
		1	2	3	4	5
		1.1472	1.2228	-1.4092	1.4382	-1.5167
Load Case	Σ Vertical Loads (N)	(Σ Vertical Loads/Eigenvalue)*100 (%)				
1	6071	0.53	0.5	-0.43	0.42	-0.4
2	55640	4.85	4.55	-3.95	3.87	-3.67
3	61022.5	5.32	4.99	-4.33	4.24	-4.02
4	54934.3	4.79	4.49	-3.9	3.82	-3.62
5.a	201015	17.52	16.44	-14.26	13.98	-13.25
5.b	279187.5	24.34	22.83	-19.81	19.41	-18.41

Table 2. Vertical Load-to-Buckling Load Ratios for the Last Five Buckling Modes

		Eigenvalue * 10 ⁶ (N)				
		6	7	8	9	10
		-1.5441	1.558	-1.603	-1.6378	-1.6706
Load Case	Σ Vertical Loads (N)	(Σ Vertical Loads/Eigenvalue)*100 (%)				
1	6071	-0.39	0.39	-0.38	-0.37	-0.36
2	55640	-3.6	3.57	-3.47	-3.4	-3.33
3	61022.5	-3.95	3.92	-3.81	-3.73	-3.65
4	54934.3	-3.56	3.53	-3.43	-3.35	-3.29
5.a	201015	-13.02	12.9	-12.54	-12.27	-12.03
5.b	279187.5	-18.08	17.92	-17.42	-17.05	-16.71

Buckling Analysis of the DS Transmission Tower

A modeling approach similar to the 1L tower was employed for the DS tower. The structural elements were modeled in ABAQUS with beam elements. The analysis was conducted nonlinearly for the five loading cases. As an example, the loading tree for one of the load cases is provided in Figure 23. The computational results were compared with the test results in terms of deflections and also in terms of buckling loads. During the test, the deflections were measured at 7 different locations around the tower, as defined in Figure 24. The deflections at the same loads were also monitored during the analysis. In Figure 25, the deflection response of the tower can be seen for the testing of load case 1. A promising agreement can be seen between the test and the computational model. The first buckling mode of the tower is presented in Figure 26. During the test for the load combination 5, a local failure of one of the legs of the tower was observed; this type of behavior was also captured computationally, as can be seen in Figure 27.

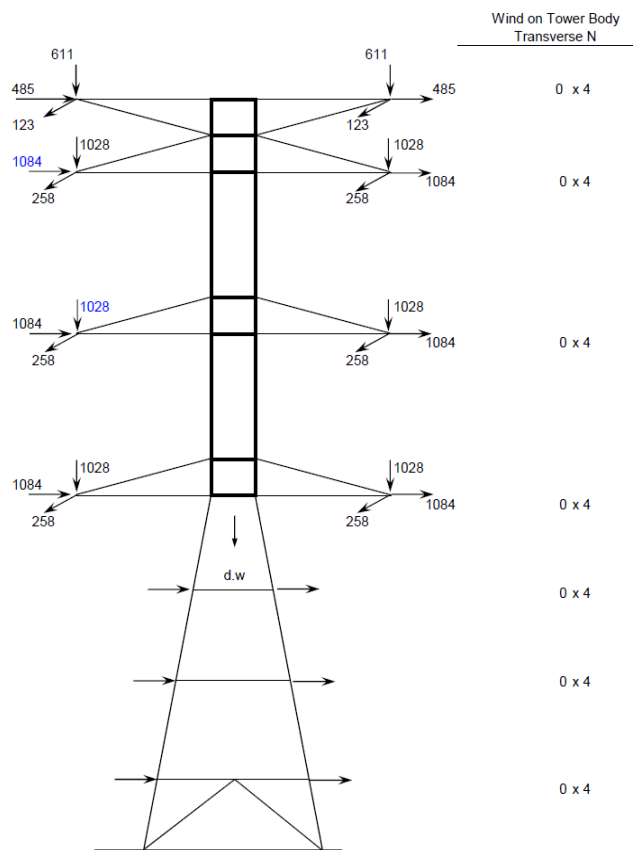


Figure 23. Load Combination 1

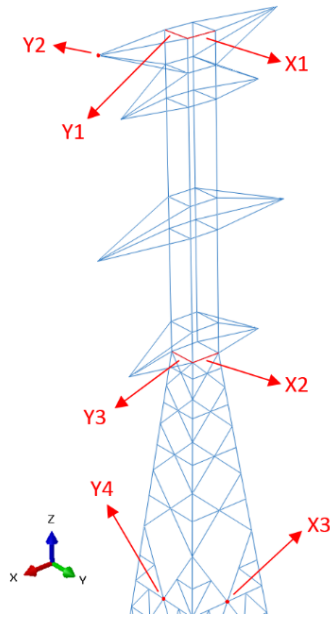


Figure 24. Location of Points where the Deflection was Obtained

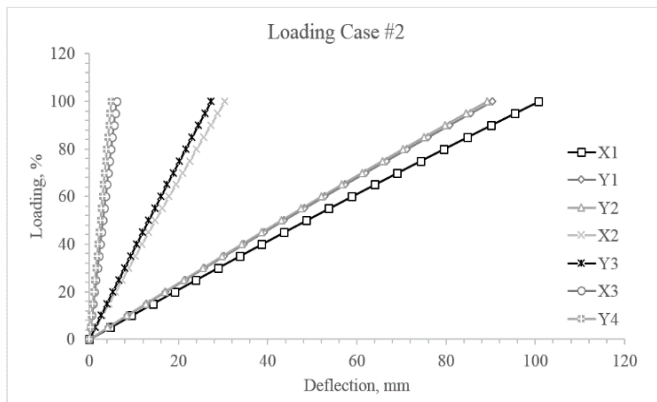


Figure 25. Load-Deflection Curves for Load Case 2

Conclusions

Nonlinear 3D finite element analysis and buckling analysis were performed in this study. It was observed that the finite element results were in agreement with the load-test results, except for some locations where strain gauges were possibly damaged. For such structures with many structural elements, it was challenging to determine buckling loads experimentally, especially at the local levels. It was concluded that the finite element model can be used to determine the buckling load and to capture the buckling behavior for transmission towers and its different components. Buckling loads at different modes were evaluated and compared with the applied vertical loads applied during the load test.

It was observed that the loads applied were far from causing the tower to fail in buckling for the 1L tower. The DS tower showed buckling at one of the main legs; the same behavior was captured during the first mode of buckling.

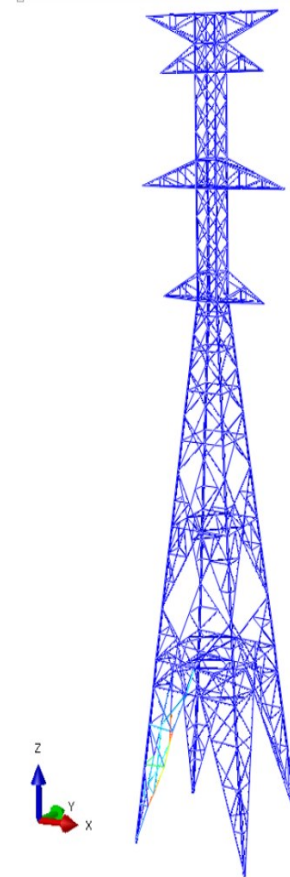


Figure 26. Deflections for the First Buckling Mode



Figure 27. Comparison between Experimental and FE Failure Modes

Currently, different transmission towers with different geometry are under investigation for which the load test has been already performed and a failure was observed at the base of the tower; nonlinear analysis and buckling analysis will be performed to investigate the failure of the towers to further justify using finite element buckling analysis as a tool to predict buckling of transmission towers.

References

- [1] Battistaa, C., Rodriguesa, S., & Pfeila, S. (2003). Dynamic Behavior and Stability of Transmission Line Towers under Wind Forces. *Journal of Wind Engineering and Industrial Aerodynamics*, 91, 1051–1067.
- [2] Prasad Rao, N., Samuel Knight, G. M., Lakshmana N., Iyer, R., & Nagesh R. (2010). Investigation of Transmission Line Tower Failures. *Engineering Failure Analysis*, 17, 1127–1141.
- [3] Al-Bermani, F. G. A., & Kitipornchai, S. (1993). Nonlinear Finite Element Analysis of Latticed Transmission Towers. *Engineering Structures*, 15(4), 259–269.
- [4] Matheson, M. J., & Holmes, J. D. (1981). Simulation of the Dynamic Response of Transmission Lines in Strong Winds. *Engineering Structures*, 3, 105–110.
- [5] Prasad Rao, N., Samuel-Knight, G. M., Mohan, S. J., & Lakshmanan, N. (1981). Studies on Failure of Transmission Towers in Testing. *Engineering Structures*, 3, 55–70.
- [6] ABAQUS Version 6.13-2 User's Manual. 2013. SIMULIA, Pawtucket, RI.
- [7] Weilian, Q., Wanjun, S., Yong, X., Youlin, X., Wenke, Q., & Zhicai, J. (2013). Two-Step Method for Instability Damage Detection in Tower Body of Transmission Structures. *Advances in Structural Engineering*, 16(1), 219–232.
- [8] Albermani, F., Kitipornchai, S., & Chan, R. W. K. (2009). Failure Analysis of Transmission Towers. *Engineering Failure Analysis*, 16(6), 1922–1928.
- [9] Prasad, R., & Kalyanaraman, V. (2001). Non-Linear Behaviour of Lattice Panel of Angle Towers. *Journal of Constructional Steel Research*, 57, 1337–1357.
- [10] Albermani, F., Mahendran, M., & Kitipornchai, S. (2004). Upgrading of Transmission Towers Using a Diagram Bracing System. *Engineering Structures*, 26, 735–744.
- [11] Yasui, H., Marukawa, Y., & Ohkuma, T. (1999). Analytical Study of Wind-induced Vibration of Power Transmission Towers. *Journal of Wind Engineering and Industrial Aerodynamics*, 83, 431–441.
- [12] Jamaledine, A., McClure, G., Rousselet, R., &

Beauchemi, R. (1993). Simulation of Ice-shedding on Electrical Transmission Lines Using Adina, *Computers & Structures*, 47(4/5), 523–536.

- [13] McClure, G., & Lapointe, M. (2003). Modeling the Structural Dynamic Response of Overhead Transmission Lines. *Computers & Structures*, 81, 825–834.

Biographies

MUSTAFA MAHAMID is a clinical associate professor at the University of Illinois at Chicago and an independent consultant working for several engineering firms. He earned a B.S. in Civil Engineering from the Eastern Mediterranean University in Turkey; an M.S. in Civil Engineering from Bradley University; and, a Ph.D. in Civil Engineering from the University of Wisconsin–Milwaukee. Dr. Mahamid has extensive practical experience in industry and has worked on projects in the U.S., Europe, the Middle East, and China. His interests include earthquake engineering, structural steel building behavior and connection design, concrete structures, and computational mechanics. Dr. Mahamid may be reached at mmahamid@uic.edu

CENK TORT holds a B.S. (METU 2000), an M.S. (University of Minnesota 2003), and a Ph.D. (University of Minnesota 2007) in civil engineering. He was a staff structural engineer at Skidmore, Owings, and Merrill (SOM) in Chicago from 2007 to 2010. Since 2010, he is an R&D director at Miteng in Ankara/Turkey working on design and optimization of transmission tower structures. Dr. Tort may be reached at cenk.tort@gmail.com

MUTAZ AL HIJAJ is a Ph.D. student in the Civil and Materials Engineering Department at the University of Illinois at Chicago (UIC) since January 2013. He graduated from Jordan University of Science and Technology with M.S. and B.S. degrees in structural engineering. Al Hijaj is a member of several engineering associations similar to the American Society of Civil Engineers (ASCE) and the American Institute of Steel Construction (AISC). Furthermore, his research focuses on investigating special types of steel connections and steel tower performance. Mutaz's passion is a deeper knowledge leading to development of better and safer steel structures. Mr. Al Hijaj may be reached at malhij2@uic.edu

ADEEB RAHMAN is an associate professor of engineering mechanics in the Department of Civil Engineering and Mechanics at the University of Wisconsin–Milwaukee. He received his Ph.D. from the University of Wisconsin–Madison in the area of solid mechanics. His area of expertise includes computational structural mechanics, the response of steel connections to fire loading, and computa-

tional modeling of structural frames in seismic loading. He has more than fifty journal and conference publications in highly specialized publications. He is currently active in research focusing on the performance of steel connections in elevated temperatures. He was twice the recipient of the prestigious Fulbright Award for the U.S. Fulbright Scholarship program. Dr. Rahman may be reached at adeeb@uwm.edu

FARIS MALHAS is Dean of the School of Engineering & Technology at Central Connecticut State University. He earned his B.S.E. and M.S.E. degrees in civil engineering from the University of Michigan-Ann Arbor, and his M.S. degree in engineering mechanics and Ph.D. in civil engineering from the University of Wisconsin-Madison. His research interests include the behavior of concrete structures and structural steel design codes and provisions. Dr. Malhas may be reached at fmalhas@ccsu.edu

SELF-SUSTAINABILITY STUDY OF ENERGY HARVESTING WIRELESS SENSOR NETWORKS FOR INFRASTRUCTURE MONITORING

Jin Zhu, University of Northern Iowa

Abstract

Wireless sensor networks may be used to improve the efficiency and effectiveness of infrastructure monitoring. The power supply or battery maintenance of a large-scale wireless sensor network in an outdoor environment is a challenge to be addressed before a system can be widely deployed in the field. In this current research project, the energy harvesting and self-sustainability issue of a wireless sensor network was studied. An energy flow model of a self-powered wireless sensor system for energy sustainable is presented here. A self-powered prototype system that performs energy harvesting from its ambient environment and stores energy in supercapacitors was implemented. The system was tested in the field for a bridge monitoring application in order to demonstrate its ability to achieve self-sustainability.

Introduction

Monitoring of transportation infrastructure, such as bridges and overpasses, is traditionally performed through periodic visual inspections. According to data from the Federal Highway Administration (FHWA), over 20 percent of all bridges in the U.S. are deficient, as of December 2012 [1], including 28,800 bridges that were structurally deficient and 71,888 bridges that were functionally obsolete. Traditional inspection is typically performed every quarter or six months by sending a staff member to the bridge site to manually check the status of a bridge, depending on the number of staff workers and bridges in each county. Manual inspections are not cost-efficient and not very reliable. Problems may not be caught during one single inspection and, if the structure problem develops quickly, one may not be able to catch the problem before a catastrophe occurs. Therefore, the development of an automatic and low-cost bridge structural health monitoring (SHM) system is in high demand and crucial for reducing the costs associated with manual inspection in order to effectively monitor the status of bridges and, therefore, prevent bridge collapses.

Within the last decade, there has been much interest in developing wireless structure monitoring for easy and fast installation and improved flexibility, because a wireless

sensor network (WSN) does not need to run cables all over the bridge and because a significant cost of any bridge monitoring system is in the cost of cabling and its installation. However, wireless sensors are typically battery powered. The major drawback is the need to change or charge the batteries periodically [2]. Therefore, a wireless monitoring system that can obtain energy from the ambient environment is very attractive to power the sensing, processing, and communications operations to achieve complete self-sustainability. The process that converts energy in the ambient environment into usable electrical power is called energy harvesting or energy scavenging. In this study, the self-sustainability problem of wireless sensor networks was investigated. A general energy flow model was developed to represent the system operation to achieve self-sustainability. Then, a wireless sensor network used for infrastructure monitoring was studied as an example. Energy harvesting design issues of a wireless sensor network were evaluated and a prototype for bridge monitoring was implemented. Its performance on self-sustainability in field tests was presented and studied.

Energy Self-sustainability Model

Self-sustainability of a wireless sensor network is defined here as if a wireless sensor network system with energy harvesting ability can achieve a balance between energy needed for desired operations and the energy harvested from the ambient environment. If a system can maintain energy self-sustainability, theoretically it can maintain its operation without external interventions such as battery changes. A wireless sensor node can choose to operate on different power consumption levels by adjusting either its duty cycle of sensing or sampling rate. Furthermore, based on the condition of the available ambient energy source, it can adjust its work mode to achieve sustainability. Given:

- E_{av} : available energy in the energy storage device (ESD)
- E_{max} : energy capacity of the ESD
- E_{th+} : positive going energy threshold—a hibernate node may resume normal operation when its energy level exceeds E_{th+}
- E_{th-} : negative going energy threshold—a node will enter hibernate when its energy level falls below E_{th-}

- P_{hav} : output power from the energy harvester
- P_{sn} : power consumption of a normal operating wireless sensor node

A general state machine model of the self-powered energy harvesting sensing system based on energy flow is shown in Figure 1. In order to avoid bouncing effects between the negative energy flow state and hibernate state, the concept of hysteresis switching was introduced.

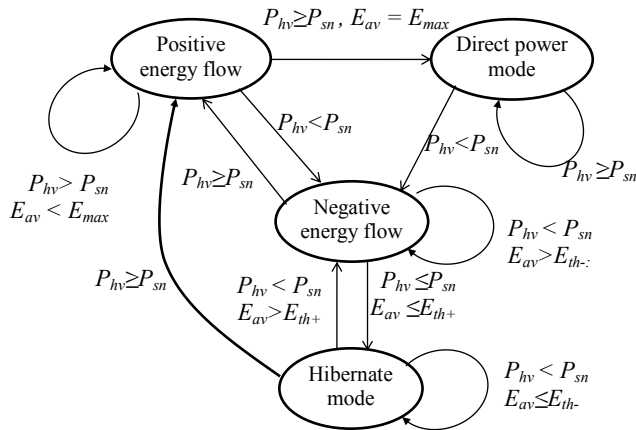


Figure 1. Energy Flow Model of a Self-Powered Wireless Sensor System

Micro Energy Harvesting for Outdoor Wireless Sensor Networks

Although renewable energy technologies such as solar panels and wind turbines are relatively mature, they are generally for large-scale systems and not suitable to low-cost, small-sized wireless sensor nodes. Some pioneer projects have been undertaken to investigate the possibilities of harvesting energy from the ambient environment for low-cost, micro wireless sensors for structure monitoring [3-6]. The potential ambient energy sources that may be used for bridge monitoring include vibration, light, air flow, temperature variations, and ambient RF energy. Ambient energy sources that may be used for wireless sensor networks are summarized and compared in Table 1 [7].

An energy harvesting module consists of charger circuits, energy storage components, and voltage regulators. The energy harvested from the ambient environment needs to be stored in energy storage components such as:

- Electrochemical batteries
- Micro-fuel cells
- Ultra-capacitors or supercapacitors
- Micro-heat engines

Table 1. Energy Source Comparison

Energy Source Type	Harvesting Performance (Power Density)
Vibration/Motion Human motion Industry machines Bridge vibration due to passing traffic	4 $\mu\text{W}/\text{cm}^2$ 100 to 800 $\mu\text{W}/\text{cm}^2$ < 1 $\mu\text{W}/\text{cm}^3$
Acoustic Noise	0.003 $\mu\text{W}/\text{cm}^3$ at 75 dB 0.96 $\mu\text{W}/\text{cm}^3$ at 100 dB
Temperature Difference Human Industry	25 $\mu\text{W}/\text{cm}^2$ 1 to 10 mW/cm^2
Wind/Air Flow	380 $\mu\text{W}/\text{cm}^2$ (assumes air velocity of 5 m/s, i.e., 11 mph)
Solar/Light Indoor Outdoor	7.2 $\mu\text{W}/\text{cm}^2$ (office light) 100 mW/cm^2 (directed toward bright sun) 10 mW/cm^2 (under sun) 150 $\mu\text{W}/\text{cm}^2$ (cloudy)
RF GSM Wi-Fi Broadband (1 - 3.5GHz)	0.1 $\mu\text{W}/\text{cm}^2$ 0.001 $\mu\text{W}/\text{cm}^2$ 0.0063 $\mu\text{W}/\text{cm}^2$

Batteries are the most common energy storage devices. Rechargeable batteries can be used with energy harvesters in WSNs. Lithium-ion batteries are popular among others due to its high energy density. However, the charging temperature for Lithium-ion batteries is limited to 0°C to 40°C. The operating temperature for supercapacitors is wider than that for Lithium-ion batteries, especially for charging, which is -40°C to 65°C. This wider operating temperature range is a good fit to the outdoor environment in hot/cold weather. In addition, one of the issues of WSNs with rechargeable batteries for long-term monitoring is the limited cycle life, typically around 500, while the cycle life of supercapacitors can be from 300,000 to 1 million cycles. Therefore, supercapacitors could be an attractive option in some wireless sensor node applications because of their increased lifetimes, short charging times, and wide operating and charging temperature ranges. In addition, harvesting components and energy storage elements usually have different voltage-current characteristics. Therefore, efficient charging and power management circuits must be integrated with the system to minimize the conversion loss and translate the scavenged energy into increased system lifetime.

Energy Harvesting Prototype Implementation

A prototype of a wireless sensor network with an energy harvesting module, named EHSuperCap, was designed and implemented. According to the comparison shown in Table

1, wind and solar energy are still the most appealing options based on the energy density for harvesting. In this prototype, a small film solar panel was adopted as the energy source. Before the wireless sensor node platform was selected, the specific requirements for the bridge monitoring application were identified. According to objectives of the project, the sensing elements used were traditional strain gauges, as the system needed to monitor the bridge strain peaks due to live passing traffic.

Based on research by Lu [8], it is suggested that the 125 Hz data acquisition frequency is adequate to capture strain peaks produced by highway-speed trucks. In addition, synchronization of sampling data from multiple sensor nodes is necessary to ensure the accuracy of the engineering analyses performed on the response data. Considering that the error introduced due to synchronization error should be no greater than 1 percent, the synchronization error should be less than 80 μ s.

Given a total bridge span of 320 ft. (97.5 m), the transmission range of the wireless sensor nodes must be 320 ft. or more. The platform should also support point-to-multipoint or mesh communication, given that multiple strain gauges are needed for synchronized strain monitoring. Assuming a measurement range of ± 500 μ strain with 1 μ strain resolution, at least a 10-bit analog-to-digital converter (ADC) is needed. In addition, the sensor nodes should be able to provide excitation to the Wheatstone bridge for strain gauges. Given that an energy harvester is needed to power the sensor node, the platform should provide an interface for the external power supply. The platform needs to be operated outdoors and, therefore, a wide operating temperature range is expected.

Based on the requirements, after comparing multiple nodes, SG-Link nodes were selected from LORD MicroStrain Sensing Systems [9]. SG-Link-LXRS node has a differential input channel (strain channel), a single-ended input channel (analog channel), and an on-board temperature sensor channel. The differential channel is excited with 3 volts and the input is first passed through two-stage amplification and then into a 12-bit analog-to-digital converter (ADC). The data can be sampled in four different modes: data logging, low-duty cycle (LDC), streaming, and synchronous. The synchronous mode was used in this application since multiple nodes were needed to collect the strain data synchronously for later analysis. In the synchronous mode, the base sends beacons to synchronize all the nodes within the network.

In this project, based on application requirements, the synchronous mode was used with a sample rate of 128 Hz. According to the power profile of the SG nodes in the syn-

chronous sampling mode, the total energy needed for different operation can be determined by Equation (1):

$$E_{total} = V_{in} \left(\frac{N_s}{f_s} \times I_{sample} + t_{tx} \times I_{tx} + t_{sleep} \times I_{sleep} \right) \times \frac{24 \times 3600}{t_{interval}} \quad (1)$$

where, V_{in} is the power supply voltage; N_s is the number of sweeps (samples) per cycle; f_s is the sample frequency; I_{sample} is the average current used during sampling; t_{tx} and I_{tx} are the time and average current when the node continues to upload the data, respectively; t_{sleep} and I_{sleep} are the time and average current, respectively, when the node is in a sleep stage.

The duty ratio for synchronous sampling burst mode was defined via Equation (2):

$$D_{burst} = \frac{N_s / f_s}{t_{interval}} \times 100\% \quad (2)$$

The estimated energy needed for one-day operation (24 hours) of synchronous sampling with the frequency of 128 Hz is 1385 J for $D_{burst} = 0.2$, 2020 J for $D_{burst} = 0.33$, 2723 J for $D_{burst} = 0.5$, and 3631 J for $D_{burst} = 0.67$. According to this energy requirement analysis, a thin-film solar panel (270 mm $\times$ 100 mm $\times$ 1 mm, 31 gram) was selected as the energy generator and two 350 F supercapacitors as the energy storage devices.

The primary reason for using supercapacitors for storage was their wide temperature range of operation and charging. Since the winter temperature in Iowa could easily be below 0°C (32°F), which is already out of the charging temperature range for a typical rechargeable Lithium-ion battery. Another consideration was the number of cycles that a device could be discharged and recharged before losing more than 80 percent of its capacity. The supercapacitor has an unlimited shelf life when stored in a discharged state. When referring to lifecycle, the manufacturer data sheets usually reflect the change in performance, typically a decrease in capacitance and increase in equivalent serial resistance (ESR). For example, a 15-percent reduction in rated capacitance and a 40-percent increase in rated resistance may occur for a supercapacitor held at 2.5 V after 88,000 hr. (10 years) at 25°C [10], while the lifecycle for Lithium-ion rechargeable batteries is much shorter under similar operating conditions. The lifecycle of a rechargeable Lithium-ion battery is typically around 500 hr. [11].

Design goals of the energy harvesting and storage circuit included low-cost, low-power consumption, and a simple design. Three Zener diodes were used to protect the super-

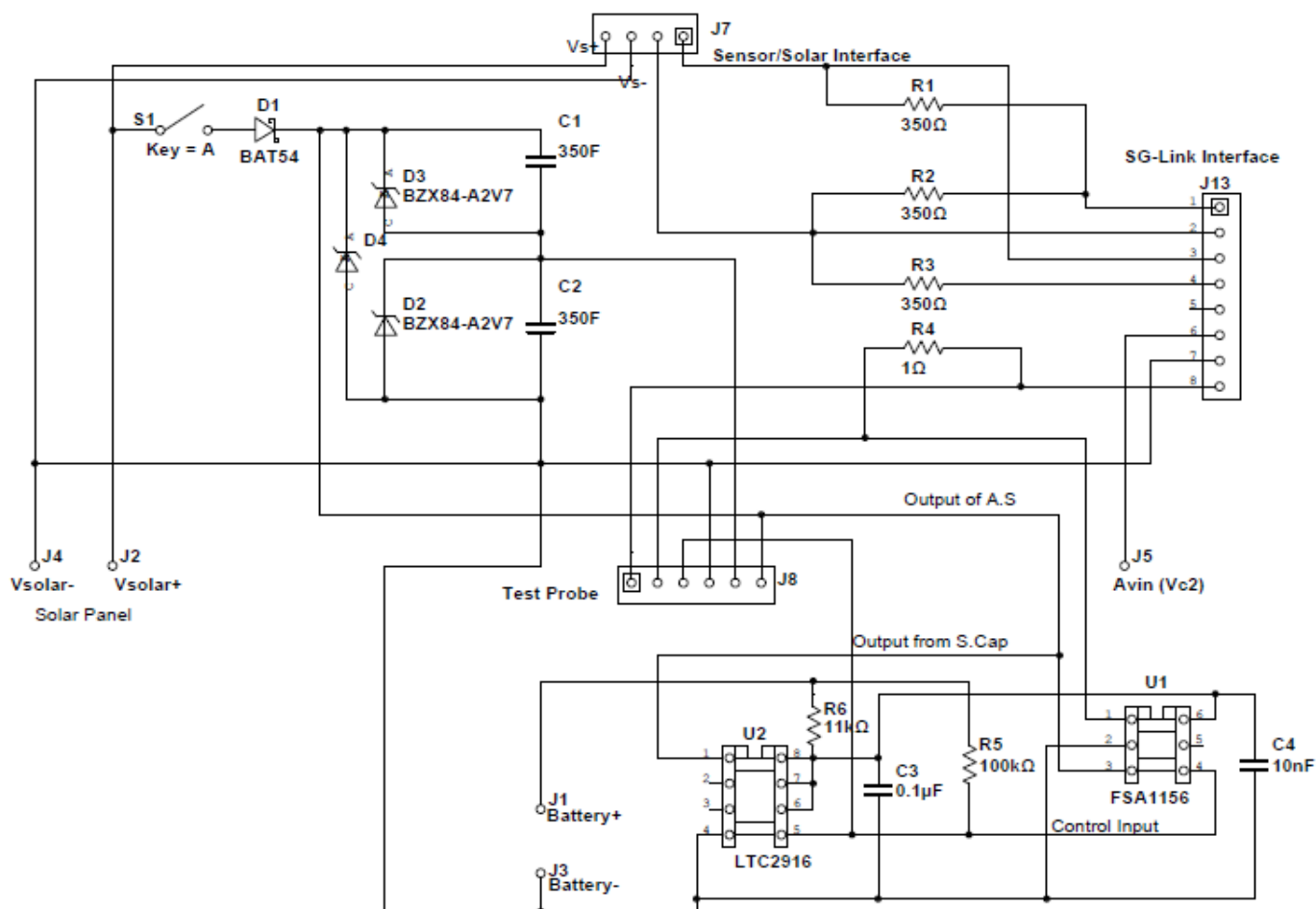
capacitors from over charges exceeding their rated voltage (2.7 V) and the total voltage not exceeding 5 V. The minimum operation voltage of SG-Link nodes is 3 V. After tests, it was found that an SG-Link node stopped working until the supply voltage dropped below 2.8 V. When it stops working, the SG-Link still draws the current, which is uncontrolled and larger than with normal operation, depleting the supercapacitors completely in a short period. Therefore, it was desired to shut down or disconnect the supercapacitor from the SG-Link when the power supply falls below 3 V to prevent wasting energy. A relatively high charging voltage of 3 V or more also helps the solar cells to work on the regions close to the maximum power point (MPP).

For this purpose, a low on-resistance, low-voltage single-pole/single-throw (SPST) analog switch, FSA1156 from Fairchild Semiconductor, and a selectable voltage supervisor, LTC2915 from Linear Technology, were introduced.

The voltage supervisor was selected to monitor the voltage level of supercapacitors. If the voltage is lower than 2.9 V, it sends a control signal to the analog switch and the switch will open and shut down the pass between the supercapacitors to the SG-Link node. When the voltage increases back to 2.95 V from energy harvesting, the voltage supervisor will again close the analog switch so that the node will resume work. Accordingly, $E_{th+} = 762 \text{ J}$ and $E_{th-} = 736 \text{ J}$. The schematic of the EHSuperCap board is shown in Figure 2.

SG-Link nodes were configured to enter the synchronous mode at start-up, so a node may rejoin the network and continue collecting data whenever it scavenges enough energy from the ambient environment. Using ICs, a reliable power supply for the ICs was required no matter what level of energy the supercapacitors possessed. A primary battery was used; it was not ideal, but given that both ICs were ultra-low

Figure 2. Schematic of the Prototype



power, with a quiescent supply currents of $1\ \mu\text{A}$ and $30\ \mu\text{A}$, a small-sized battery may last a long time. Assuming a button battery capacity of 1000 mAh (for example, CR2477, operation temperature -30°C to 60°C) and an average supply current of $35\ \mu\text{A}$, a single button battery can last more than three years. The implemented system with an energy harvesting module is shown in Figure 3.



Figure 3. Energy Harvesting Module for Wireless Sensor Networks

Field Test Results

The system was deployed on a bridge with a span of 320 ft. in a field test to monitor the bridge strain peaks due to live passing traffic. Strain data were collected with a burst mode in the frequency of 128 Hz and could be used to indicate the health or deformation of the bridges. Four nodes were deployed in the field tests. The base was connected to a laptop, which was locked in the cabinet close to the north-west bridge abutment. As discussed previously, the sample rate was set to 128 Hz and synchronous sampling burst mode was used. For the burst mode, it was set to 38,400 sweeps every 30 minutes (taking samples at 128 Hz for 5 minutes every 30 minutes). So the duty cycle was $D_{\text{burst}} = 1/6$ or 16.7 percent.

Strain gauges were connected to the sensor node platform using a Wheatstone bridge, where a regulated 3 V excitation was provided from the sensor node during sampling. A Weldable $350\ \Omega$ strain gauge from Vishay was adopted for strain measurement. Strain samples were then collected via the differential input. Besides the strain, all nodes also measured the supercapacitor voltage via the single-ended input. In addition, one node (Node 649) also measured the temperature using the internal temperature sensor. The configurations of the wireless sensor nodes are listed in Table 2. Voltages of the supercapacitors were also collected to monitor operations, in addition to the strain and temperature data.

Table 2. Configurations of Synchronized SG-Link-LXRS Sampling Nodes

Node (#)	# Active Channels	% of Bandwidth	TDMA Address	Location
639	2	1.56%	1	North girder 2
649	3	1.56%	9	South girder 1
651	2	1.56%	17	South girder 2
712	2	1.56%	25	North girder 2

Some sample strain data and their corresponding frequency responses are shown in Figure 4. Comparing the data with previous results using wired sensors [8], it was concluded that the data collected using wireless sensors were valid.

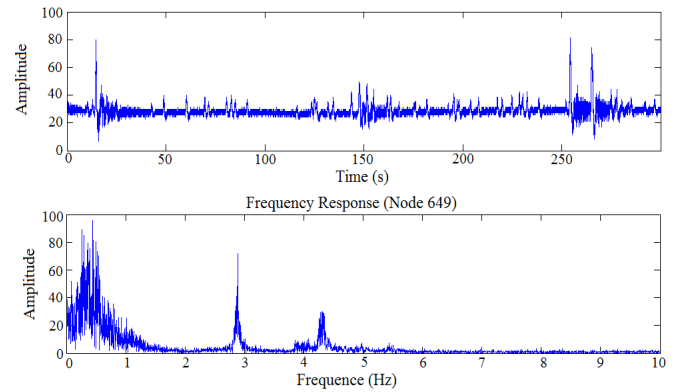


Figure 4. Sample Raw Strain Data and Frequency Response

Although potential energy resources are renewable for harvesting, they are still limited in magnitude and each node must be as energy-efficient as possible to achieve self-sustainability. The energy sources, including the solar panels used in this project, are typically irregular and intermittent. In this study, the wireless sensor nodes were evaluated for their ability to achieve self-sustainability and operate with or without minimal maintenance. For this purpose, supercapacitor voltage was monitored during the test periods. The capacitor voltages for the four wireless sensor nodes for one 6-day test are discussed here. The initial supercapacitor voltages before the 6-day test were measured as follows:

Node 649: 2.38V Node 651: 1.68V
Node 712: 2.14V Node 639: 1.70V

The supercapacitor voltage results for the 6-day test are presented in Figure 5. Note that the voltage shown is the voltage of one supercapacitor. According to the test results

in the lab, the capacitance was not exactly the same for the two supercapacitors in series, but the difference was small enough that it could be assumed that the power supply voltage was approximately twice that of one supercapacitor.

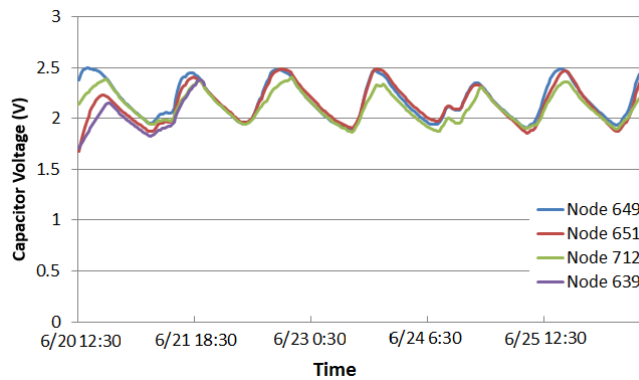


Figure 5. Supercapacitor Voltage Results for a 6-Day Test

As shown, the nodes were working properly except for Node 639. Node 639 worked during the first 1.5 days and then lost its connection. After testing the EHSuperCap, it was found to be working. Even though the supercapacitor voltage was still above 2V, this SG-Link node stopped transmitting data back to the base station and needed to be reset. Figure 5 shows that the EHSuperCap board is able to provide enough energy to support the operation of wireless sensors. Taking Node 651 as an example, it started at 1.68 V and, after the first day, it reached 2.23 V at 7:00 p.m. The voltage kept dropping during the night until it reached its low point at 7:00 a.m. at 1.87 V, before the voltage started to increase again during the daytime. The lowest voltage during the test was 1.85 V, well above the required 1.5 V. Overall, it can be seen that the nodes can be self-powered and maintain self-sustainability using a solar panel with a 270 cm² area.

The available solar energy depends highly on the weather. The light level was between 800 and 2000 FC around noon, as measured on-site, but the radiant level changed quickly and significantly when clouds moved in. The energy gain or loss did not depend directly on temperature, but the temperature data gave a general idea of daylight time. The temperature data measured and weather conditions are presented for the 6-day test duration in Figure 6. The top blue line is the temperature over six days, while the red bottom line is a single capacitor voltage (at Node 651). It can be seen that the peak voltage followed daily temperature peaks in general, sometimes with a delay.

The energy gained or lost for storage (besides the energy that was consumed directly from the solar panel during the daytime) in Node 651 is presented in Figure 7. The positive

value indicate the energy gain from the bottom to the peak while the negative value is the energy loss from peak to bottom during each day. It can be seen that energy gain and loss fluctuated but overall trend is balanced.

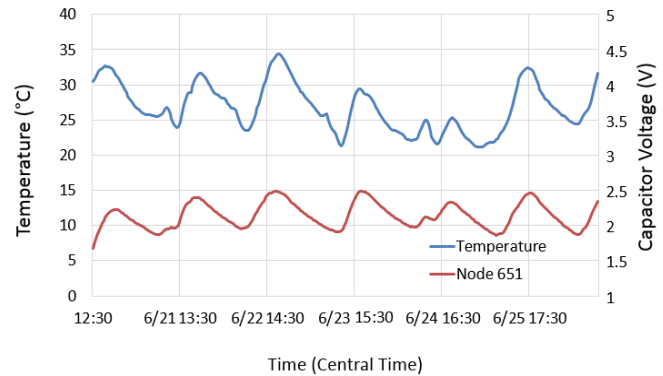


Figure 6. Temperature Data versus Supercapacitor Voltages

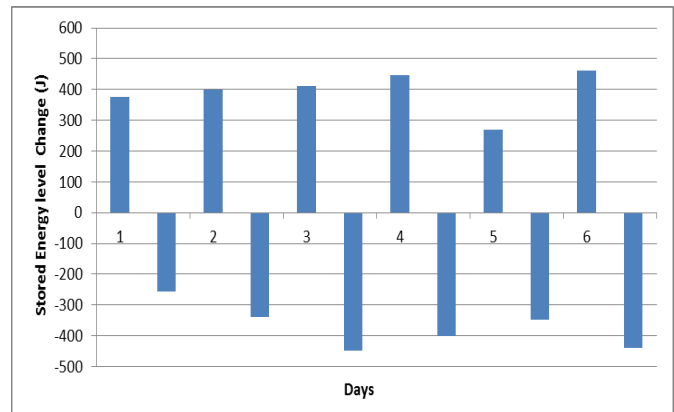


Figure 7. Energy Level Gain/Loss in Node 651

Conclusions

Most wireless sensors are powered using batteries and can last a limited time, especially for high sample frequency demanding applications. With this project, a minimum 125 Hz sample rate was needed. With a sample rate of 128 Hz, three high-capacity D-cell batteries (assuming 1.5V, 18,000 mA each) can provide only a two-month power supply to SG-Link nodes. It is not convenient or cost-efficient to replace batteries. Therefore, a self-powered wireless sensor network is a very attractive option. A self-sustainable wireless sensor node can support its operation by harvesting energy from the environment and, thus, its lifetime is limited only by the lifetime of the electronics. An implemented EHSuperCap board with a film solar panel and supercapacitors was used to supply power to wireless sensors for bridge monitoring.

Voltages of the supercapacitors were collected to monitor operations besides the strain and temperature data. The test results showed that the EHSuperCap board was able to provide enough energy to support the operation of wireless sensors. When a node works in the synchronous burst mode, the collected data were not continuous but still provided the strain information resulting from ambient traffic. Given that the strain data give more usable information during traffic-dense periods such as rush hours, the nodes may be set to take more samples during the traffic-dense hours and save energy during other periods. Overall, energy harvesting-based wireless sensor networks are an attractive option for various infrastructure monitoring applications. Future work will include the incorporation of the energy information into the protocol design and operation of the sensor nodes, and use a simulation to estimate the lifetime and evaluate the sustainability of energy harvesting-based wireless sensor systems using an energy flow model.

References

- [1] FHWA Bridge Programs NBI Data: Bridges by Posting Status (n.d.) Retrieved May 23, 2014, from <https://www.fhwa.dot.gov/bridge/nbi/posting.cfm>
- [2] Zhu, J., & Salim, M. D. (2013). Implementation and Field Test of a Tireless Sensor Network for Infrastructure Monitoring. *Journal of Engineering Technology*, 30(1), 18-27.
- [3] Musiani, D., Lin, K., & Rosing, T. S. (2007). *Proceedings of ACM 6th International Symposium on Information Processing in Sensor Networks (IPSN)*, (pp. 390-399), Cambridge, MA, USA.
- [4] Wu, J., & Zhou, G. (2011). A New Ultra-low Power Wireless Sensor Network with Integrated Energy Harvesting, Data Sensing, and Wireless Communication. *Proceedings of IEEE International Conference on Communications (ICC)*, (pp. 1-5). Kyoto, Japan.
- [5] Sazonov, E., Li, H., Curry, D., & Pillay, P. (2009). Self-Powered Sensors for Monitoring of Highway Bridges. *IEEE Sensors Journal*, 9(11), 1422-1429.
- [6] Park, G., Rosing, T., Todd, M. D., Farrar, C. R., & Hodgkiss, W. (2008). Energy Harvesting for Structural Health Monitoring Sensor Networks. *ASCE Journal of Infrastructure Systems, Special Issue: New Sensors, Instrumentation, and Signal Interpretation*, 14(1), 64-79.
- [7] Zhu, J. (2013). *Wireless Bridge Monitoring Hardware, TR-636 Technical Report* (Vol. 3). Iowa Department of Transportation.
- [8] Lu, P. (2008) *A Statistical Based Damage Detection Approach for Highway Bridge Structural Health Monitoring* (Doctoral dissertation). Retrieved from <http://lib.dr.iastate.edu/cgi/viewcontent.cgi?article=2655&context=etd>.
- [9] Microstrain SG-link-LXRS Wireless Strain Nodes. (n.d.). Retrieved May 10, 2014, from <http://www.microstrain.com/wireless/sg-link>
- [10] Maxwell Technologies Design Considerations for Ultracapacitors (n.d.). Retrieved May 10, 2014 from http://maxwell.com/products/ultracapacitors/docs/200904_whitepaper_designinguide
- [11] Lithium-Ion battery datasheets (n.d.). Retrieved on May 20, 2014, from <http://www.panasonic.com/industrial/batteries-oem/oem/lithium-ion.aspx>

Biography

JIN ZHU is an associate professor of electrical engineering technology at the University of Northern Iowa. She earned her Ph.D. (electrical engineering, 2005) from the New Jersey Institute of Technology. Her research interests include wireless networks, ambient energy harvesting, embedded systems, and robotics. Dr. Zhu may be reached at jin.zhu@uni.edu

ADVANCING MOTORSPORTS DRIVER PROTECTION USING MODELING AND TEST CORRELATION

Grant Hankins, Indiana University Purdue University Indianapolis; Pete Hylton, Indiana University Purdue University Indianapolis;
Ryan Hoover, Center for Advanced Product Evaluation; Kevin Tribbett, Center for Advanced Product Evaluation

Abstract

Crash testing of a Sprint Car racing chassis was performed using an inverted drop test, which was carried out for the purpose of determining if a correlation existed between empirical test data and finite element analysis results. The analysis of a finite element model was correlated using data logged during the drop test to within a mean error of 9.11% of the maximum impact acceleration and velocity. Once a correlation between the finite element model and the drop test data had been established, three variant chassis designs were evaluated using the same metrics at two different simulated crash speeds. The results of the finite element studies showed that a significant reduction in peak impact acceleration of 72.8% could be achieved using a crash attenuator mounted to the top of the car's roll cage. This level of reduction in G loads in a potential crash scenario could result in a significant decrease in the risk of death or serious injury to the driver.

Introduction

Motorsports sanctioning bodies first mandated roll cages for Sprint Cars during the 1970 season, after driver Kenny Gritz died during the 1969 season at an International Motor Contest Association (IMCA) race event. Ironically, Gritz died after showing up with a roll cage fitted to his car. Officials at the race forced him to remove it, as it was considered a performance enhancing addition to the car [1]. Following this, sanctioning bodies such as IMCA and the United States Auto Club (USAC) began to mandate roll cages. The National Association of Stock Car Auto Racing (NASCAR) had first mandated this safety device in 1953, and the Federation Internationale de l'Automobile (FIA) had followed suit in 1961, mandating rollover structures in Formula 1 [2], [3]. Roll-hoops as a safety structure can actually trace their roots to the early 1900s [4], [5]. Today, Sprint Cars and all racecars of similar construction (e.g. Midgets, Mini-Sprints, Silver Crown cars, etc.) utilize a roll cage, but its form is mostly unchanged from the 1970s.

At first glance, it is rather surprising that roll structures were not commonplace on these race cars before 1969 [6]. However, speed was historically placed in front of safety in the early years, and many drivers paid for it with their lives

[7]. Many, like famed NASCAR driver Dale Earnhardt, could have been saved by earlier implementation of improved driver safety equipment [8]. The lower echelons of motorsports also have their stories of tragedy such as Shane Hmiel, whose crash in a Silver Crown car left him a quadriplegic [9]. It is arguable that if additional driver protection had been in place, the severity of these drivers' injuries might well have been avoided. Promoters, fans, and racers alike are beginning to question the lack of safety innovation in Sprint Car racing since the initial implementation of the roll cage [7]. The Charlotte Observer reported in 2014 that since 1990 there had been 523 motorsports deaths with 53 occurring at short tracks, where the majority of Sprint Car drivers compete [10]. With an average of 15 drivers dying every year in grassroots dirt-track racing, there is growing concern that improvements in safety devices are desperately needed [7].

A collaboration between the Center for Advanced Product Evaluation (CAPE), Spike Chassis Inc., AED Motorsport Products, the MotorSport Engineering Conference, and Indiana University Purdue University Indianapolis (IUPUI) has enabled the generation of research into roll cage design, which is discussed in this paper. Previous empirical testing and research conducted into the dynamics of Sprint Car roll cages undergoing a top plane impact have relied on the use of sled testing [11-13], whereby the Sprint Car chassis was fixed to a sled and propelled towards a solid barrier at speed. Due to the constraints required to attach the chassis to the sled, the degrees of freedom available to the chassis during the crash event were severely compromised. This presents a distinct problem when attempting to understand Sprint Car crash dynamics in a top plane impact, as the limitations imposed by the sled can never truly be removed from the empirical testing results.

Restrictions on chassis crash characteristics that use a cable-driven sled were a driving factor in the decision to use drop testing as a basis for this current correlation of empirical model to finite element analysis results. This type of testing was utilized by Patalak and Gideon [3] in the analysis of NASCAR's sixth generation of stock car chassis. Additionally, neither the Society of Automotive Engineers (SAE) [14] nor the National Highway Traffic Safety Administration (NHTSA) [15] mandate or provide standards for sled-testing impacts to simulate rollovers. In a notable

study conducted by Forrest et al. [16], the authors sought to determine real-world roof strength through a comparison of vehicle styles that had passed Federal Motor Vehicle Safety Standard 216 [15] and subjected them to an inverted drop test. Based on their work, along with the proven use of the SAE J996 inverted drop-test standard [17] as a procedure for testing motorsports vehicles in Patalak and Gideon's work, there is strong evidence to suggest that use of inverted drop testing to determine Sprint Car roll cage crash characteristics is a viable solution.

Currently there exists no known model for the rollover characteristics of Sprint Cars. The research presented here provides basic insight into understanding the load paths and deformation characteristics for a Sprint Car undergoing a top-plane impact. Furthermore, this research may have the potential to improve the crashworthiness of Sprint Cars and other similar racecar chassis with the possibility of reducing the incidence of severe driver injury brought about by rollover events resulting in top-plane impacts.

Test and Analysis Procedure

In this current study, the authors sought to correlate test data from an inverted drop test of a Sprint Car at CAPE to a finite element model (FEM) and finite element analysis (FEA). The correlation of the test vehicle to the FEM of the vehicle was accomplished through a combination of comparisons between logged data and visual inspections of the high-speed video footage. The correlations, FEA, and all subsequent analyses were conducted using ANSYS Explicit Dynamics [18-20]. Once a correlation was established between the CAPE inverted drop test and the FEM, previous work conducted at IUPUI [21] was used as basis for exploring the possibility of reducing chassis crash accelerations and frame deformation through use of a crash attenuator. On December 9, 2014, during the MotorSport Engineering Conference (MSEC) in Indianapolis, IN, the CAPE staff instrumented a Sprint Car chassis, which had been prepared by IUPUI, for the inverted drop test, as shown in Figure 1. The vehicle was outfitted with hardware necessary to approximate the weight distribution of a Sprint Car as it would be raced. A Hybrid II dummy (non-biofidelic) was utilized along with accelerometers on the vehicle and load cells in the floor. The test dummy was seated in a non-containment seat and restrained with a 5-point safety harness that met Safety Foundation, Inc. (SFI) standards [22].

For the purpose of this test, the car assembly was comprised of limited or otherwise non-functioning components beyond the seat, restraints, and chassis. The main purpose of component inclusion within the test chassis was to attain the minimum weight requirement for USAC Sprint Cars in

competition of 1325 lb. However, this goal was not quite met and the final weight on the day of the test was 1220 lbs. The suspension arms, dampers, and torsion bars allowed for limited suspension movement and varied from semi-used to vintage condition. This was due to the need to compile and assemble the parts within a limited schedule and budget. The test chassis was in used condition and featured (prior to the test) several damaged areas within the front section of the vehicle. These portions of the chassis were areas not directly associated with the roll cage and were considered to have negligible influence on the test.



Figure 1. Test Vehicle Prepared for Drop Test

Once the test vehicle was inverted, it was positioned so that the chassis' bottom frame rails were parallel to the platform. The test vehicle, once hoisted to the maximum allowable height for the test site, was allowed to rest so as to reach the desired equilibrium state of zero degrees roll, pitch, or yaw. The drop height was measured from the top surface of the platform to the rear of the chassis' roll cage. This height was recorded as 145.25 in. (3.690 m). The data logger used for the test had a logging frequency of 20,000 Hz and began logging as the top plane of the car passed through a laser gate forming a horizontal plane located 21.89 in. (0.57 m) from the top surface of the platform. All results from the drop test were filtered in accordance with SAE J211-1 using Channel Frequency Class (CFC) Butterworth filters, where needed [17]. The data system contained a single 3-axis accelerometer and four load cells located underneath the platform onto which the chassis was dropped. The Sprint Car chassis was reverse engineered using a combination of hand measurement tools and interviews with Sprint Car chassis manufacturers. The opportunity to review chassis drawings used by Spike Chassis, Inc., a Sprint Car chassis manufacturer located in Brownsburg, Indiana, also proved to be very useful when finalizing the chassis model, which is shown in Figure 2.

In order to correctly export the reverse-engineered chassis model to ANSYS Design Modeler, the 2D sketch data were copied into a 3D sketch within the model. This step was

needed as 3D sketch geometry was the only line geometry that can be imported into ANSYS Design Modeler. Once the 2D lines were copied into a 3D sketch, any lines that formed a T-joint comprised of two bodies were trimmed and redrawn to create three bodies.

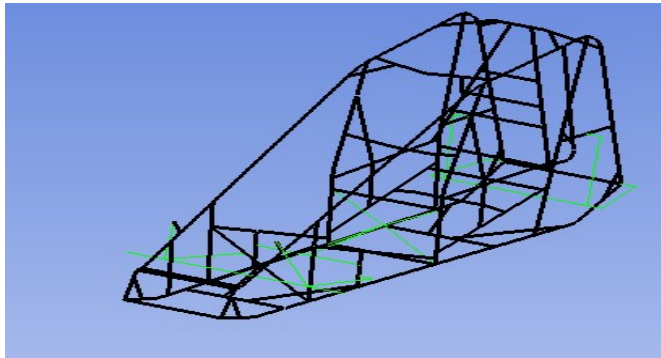


Figure 2. FEM Model of the Chassis

This is important as ANSYS Design Modeler is only able to bond the intersection of two or more line bodies at their respective endpoints. Failure to ensure that all line intersections contained a number of vertices and end points equal to the number of intersecting lines resulted in failure to bond bodies when meshed in ANSYS Mechanical. Due to the characteristic length of line body elements being equivalent to the element's specified length within the solver properties, time was taken to ensure that no single sketch line or line body was less than 1 inch in length, in order to keep the minimum time step for the analysis at a reasonable size. In order to reduce the chance of a non-bonded line body chassis model being solved in ANSYS Explicit Dynamics, the chassis model once imported into the ANSYS environment was first run through a six-mode modal analysis in order to identify where line body intersections had been missed. The modal analysis is substantially less computationally intensive than an explicit dynamic analysis and allows for relatively quick identification of problem areas, which typically appear as floating tubes or disconnected joints.

Finally, before the geometry was imported into the ANSYS Mechanical environment, correct cross-section attributes were applied to all line bodies; also, the individual line bodies were combined into a multi-body part, which allowed the topology of the model's mesh once generated to be shared and create a continuous mesh across common regions [19]. Table 1 shows the various densities that were applied to each body in order to attain proper mass distribution for the FEM. Attempting to include free-moving suspension members with realistic contact sets including bolts proved to be too complex to mesh correctly for an explicit dynamics test, due to the many small surfaces created when using bolts.

Table 1. Mass Distribution

Location	Weight lb.
Front Axle	127.58
Front Arms	18.23
Front Engine Plate	3.17
Engine	590.00
Rear Engine Plate	6.83
Seat	164.00
Torque Tube	12.50
Rear End	132.22

The engine plates were machined from aluminum and were reversed engineered as the overall dimensions were what was important and required for the study. The seat weight used the Hybrid II adult male dummy in the CAPE test [23]. The original model of the chassis, in drop-test configurations, was shown relative to the base plate in Figure 3. All analyses were run with ANSYS's material failure model activated to allow for erosion of elements that exceeded the yield stress of the 4130 normalized steel from which the Sprint Car chassis and cage were constructed.

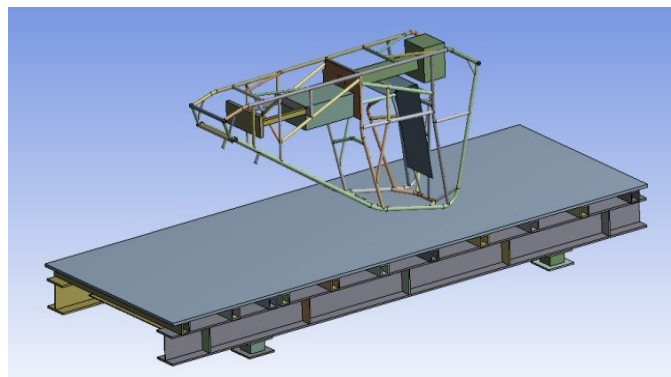


Figure 3. ANSYS Initial FEA Configuration

Results and Correlation

Figure 4 shows the first impact event with the platform. The images were taken from frames of the high-speed footage captured during the test. The blue mark and puff of chalk dust that appears in the photo indicate that the dummy's helmet struck the platform during impact. Based on the initial height of the vehicle at the start of the test, 145.25 in. (3.690 m), an estimated impact time for the rear of the roll cage was calculated as 19.03 mph (8.51 m/s). Dividing this rate by the gravitational constant gave a time for the drop from release to platform contact of $T_c = 0.8674$ s.

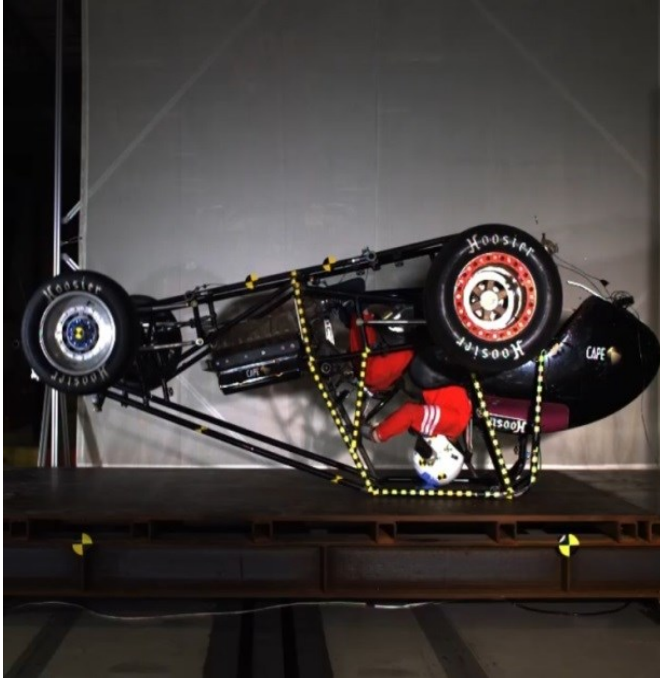


Figure 4. Initial Impact during Drop Test

Based on a frame-by-frame review of the video footage at a recorded frame rate of 30 frames per second, it appeared that it took 27 frames for the vehicle to travel the drop distance, resulting in a calculated time of $T_v = 0.9009$ s. Examining the raw CAPE data in Excel showed that there was a correlation of the calculation and visual review of the video footage at $T_x = 0.0996$ s into the CAPE data. As a final check, the height of the laser gate was calculated from the speed at $T_x = 0.0996$ s in the CAPE data and found to be 105.50 in. (2.67 m). This distance is the gate height relative to the chassis' initial position. To find the height from the ground, $d = 15.09$ in. (0.383 m), the height of the platform was subtracted from $z = 145.25$ in. (3.690 m) and found to be slightly more than the known height of the laser gate at 54.07 in. (1.39 m).

It should be noted that the impact speed of the chassis in this test proved to be roughly half of the impact velocity of 37 mph (16.54 m/s), which was found in the previously reported sled-test data. This was significant as the dummy driver's head still made contact with the barrier, just as it had in the sled tests. The peak velocity of the vehicle during the primary impact was approximately 25.38 mph (11.35 m/s), based on the CFC 180 Butterworth filtered data with a cutoff of 300 Hz; however, this conflicted with test data shown in Figure 5, which reported a peak primary impact velocity of approximately 23.50 mph (10.5 m/s). This was believed to have occurred as the result of incorrect filter settings for the CAPE output, as all other filtered signals

were reproduced from the raw data. The primary impact speed was still more than 10 mph slower than either of the prior sled tests.

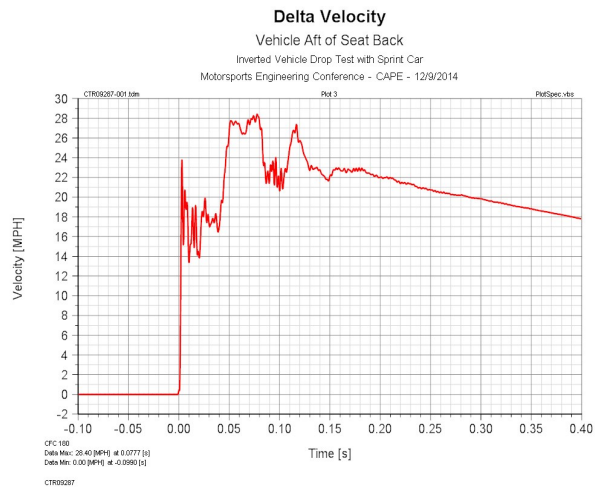


Figure 5. Velocity Data

The large spike in velocity at time 0.05 s was attributed to the front wheels striking the platform (Figure 6) when the test vehicle pitched forward, due to the shape and trajectory of impact. The rebound force from the tires caused the whole of the test vehicle to catapult back into the air, as shown in Figure 7. Figure 6 shows the vehicle just prior to the rebound action at $T \approx 0.35$ s; Figure 7 shows the vehicle just after rebound at $T \approx 0.5$ s.

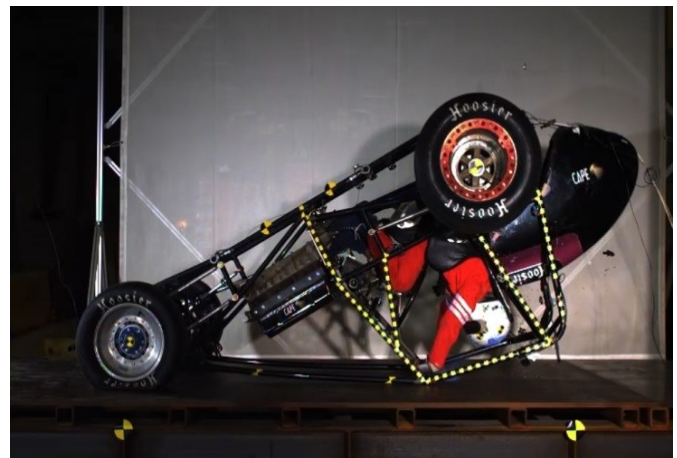


Figure 6. The Front Wheels Impact as the Car Pitches Forward

The unfiltered accelerometer data that were logged during the crash were used in conjunction with the high-speed video motion capture to generate a plot of the test vehicle's velocity during the crash event and is shown Figure 8. The crash event had a peak G load in the vertical direction (Z-

Axis) of 243.77 G's at $T = 0.0017$ s. Figure 9 shows the impact force logged by the load cells located underneath the platform used in the test. The peak impact force was 39,436.37 lbf and occurred at $T=0.0083$ s, which, as expected, correlated to the time shortly after the peak acceleration logged for the Z-Axis at $T = 0.0017$ s. Due to a processing error, raw data output for the load cells were unfortunately not saved. Thus, it was not possible to generate filtered plots for direct chart comparison to the FEM results later.

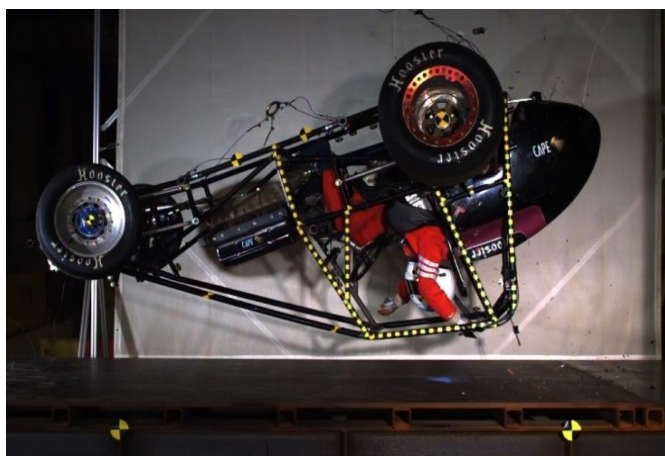


Figure 7. Rebound Prior to Second Impact

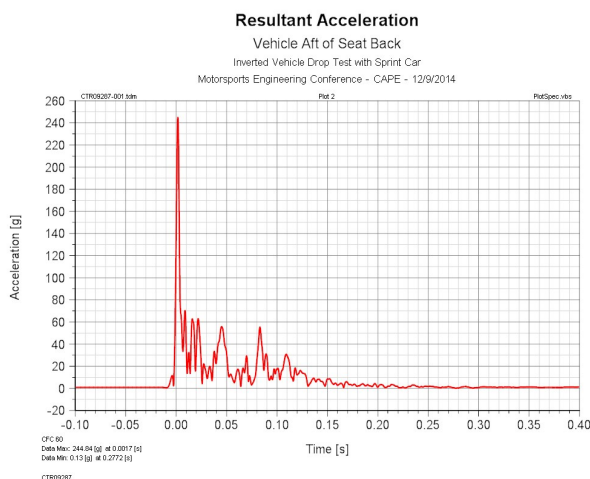


Figure 8. Accelerometer Data

The chassis damage sustained from the initial impact was minimal. While warping of the frame rails forward of the driver's cell did occur, the cause of that deformation was the forward rolling motion of the vehicle after the initial impact. This forward roll brought the tires into contact with the platform, as shown in Figure 6. The vehicle's forward roll was stopped when the front suspension mounting arms made contact with the platform. This tertiary impact caused the

front suspension supports to fail. Damage to the roll hoop, caused by the primary and secondary (i.e., rebound) impacts during the crash are shown in Figure 10. The damage to the front side of the driver's cell appears to be superficial. No other areas of the frame were observed to have sustained significant damage.

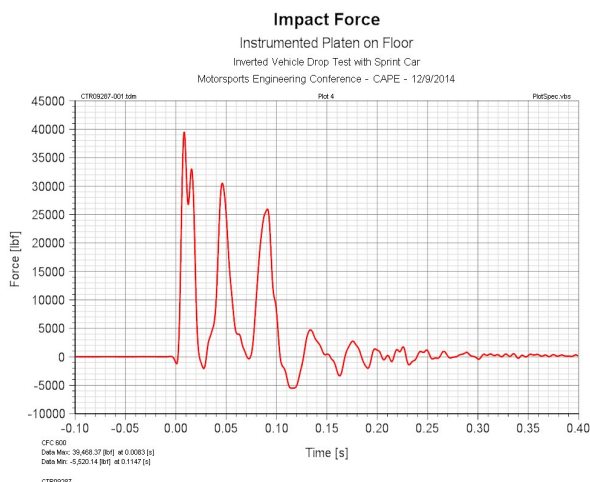


Figure 9. Load Data



Figure 10. Roll Hoop Damage

One last area of discussion regarding the CAPE test concerns the positioning of the Hybrid II dummy's head. Immediately prior to the start of the test, with the test vehicle resting inverted on the ground, a measurement was taken to determine the distance from the dummy's head to a plane tangent to the top plane of the tubes that form the Sprint Car driver's cell. This is of principal interest due to the wide variance in safety rules dictating the distance between the top of a Sprint Car driver's helmet and the roll cage. The measurement taken showed approximately 4.75 in. after inversion. This gives a distance of 3.25 in. between the top of the helmet and the inside top plane surface of the roll cage. USAC mandates that the roll cage should extend "4 in. above the driver's helmet when seated in the driving position" [24]. Compliance with this rule prior to this drop test was verified.

The empirical results from the CAPE drop test and the predictions for the FEM model were compared for maximum acceleration during the test and simulation. The results, as shown in Figure 11 and summarized in Table 2, indicated a good correlation.

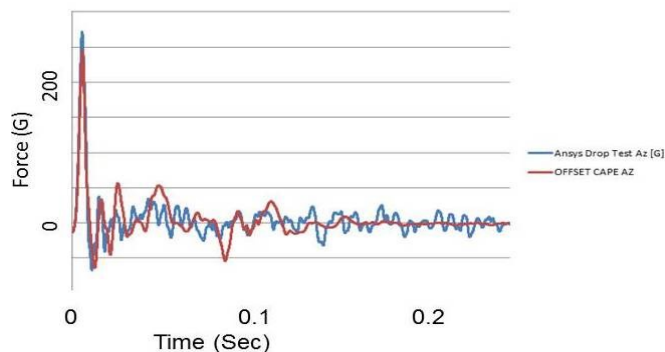


Figure 11. Comparison of Acceleration from Test Results and Prediction

Table 2. Comparison of Acceleration Test Results and Prediction

ANSYS Drop Test		
Signal	Az	% Error
Maximum G	271.2103624	9.72%
Minimum G	-67.73351743	6.19%

Figure 12 shows a comparison of the filtered and velocity data from both FEM and CAPE filters using a CFC 180 Butterworth with a cutoff of 300 Hz. The CAPE data were slightly out of shift, due to the difference in data start points along with the necessity of locating the ANSYS a small distance away from the platform. The error between the peak velocities was expected based on the acceleration results in Figure 11. The FEM calculated a peak velocity of 27.74 mph (12.40 m/s), which gave an error of 8.50% relative to the CAPE peak velocity of 25.38 mph (11.35 m/s) for the CFC 180 filtered signal. The error-to-peak impact velocity reported earlier was significantly larger at 15.30%.

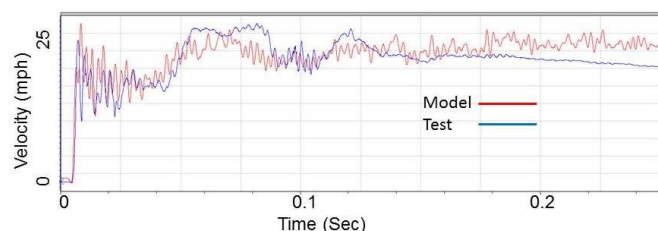


Figure 12. Comparison of Velocity from Test Results and Predictions

The plot in Figure 13 shows the contact loads calculated during the ANSYS drop test filtered using a CFC 600 Butterworth filter with a cutoff at 1000 Hz. The peak load for the duration of the analysis time was recorded as 45,664.85 psi. This high load was believed to be caused by the front suspension members making contact with the platform without the added area of the front wheels from the CAPE test to distribute that load. The primary and secondary impact loads show correlation to the recorded test data. The primary impact load for the CAPE test was recorded as 39,468.37 psi and the ANSYS results output a calculated result of 40,715.33 psi for the correlated event. This gave an error of 5.52% for the primary contact load. There was significant error in the FEM's calculated minimum contact load directly before the crash event of 79.19%. The reason for this is unknown, but was likely due to the material density difference between the FEM platform and the one used in the CAPE test. It is also possible that the mesh control size imposed on the platform for the study contributed heavily to the inaccuracy in the calculation of the minimum contact loads.

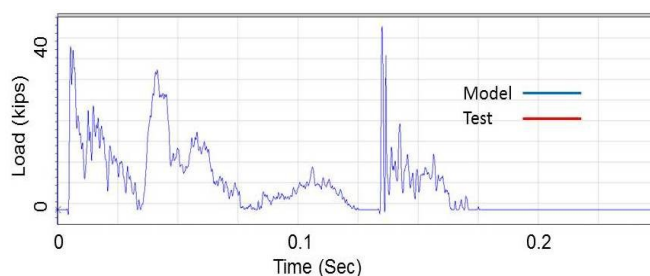


Figure 13. Comparison of Load from Test Results and Prediction

The signal outputs discussed in the previous results sections were compared to the available CAPE baseline signal through Power Spectral Density (PSD), as it was believed that this would provide the best indication that signal correlation existed, if at all. Figure 14 shows the PSD plot for the CFC 60 Butterworth filtered accelerometer channels with a cutoff set at 100 Hz. The plot axis has been reduced to the 0-250 Hz range, as this was where most of the significant frequency data lay. The output from ANSYS is represented by the red trace and CAPE by the blue. It appears that there is a strong correlation between the 20 Hz and 50 Hz ranges of the two signals. This lends strength to the argument that the ANSYS FEA does represent an accurate model for the CAPE drop test. Further review of the velocity frequency data strengthens this argument, as the following plots show even stronger correlations between the signal frequencies (Figure 15). These data were generated by integrating the CFC 180 Butterworth filtered accelerometer data per SAE J 121-1 with a cutoff at 300 Hz.

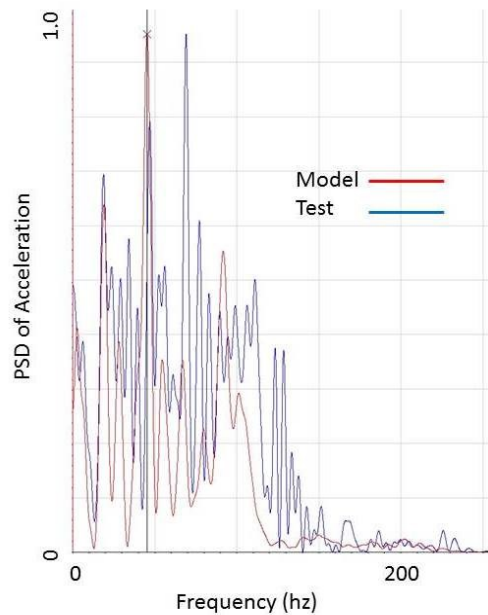


Figure 14. PSD Plot of CAPE and FEM Accelerations

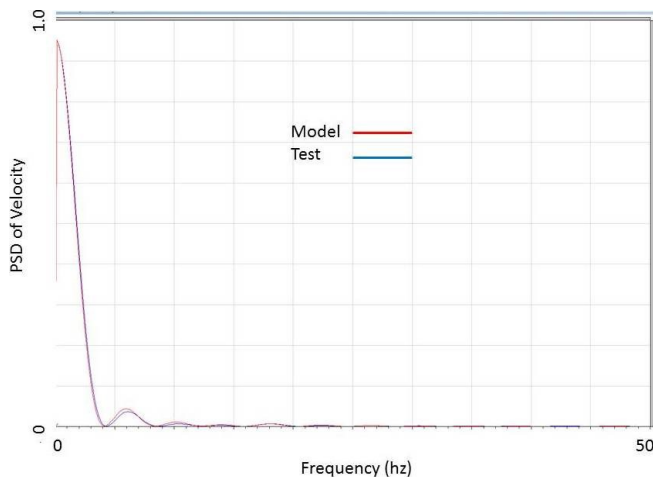


Figure 15. PSD Plot of CAPE and ANSYS Velocities

The comparison of empirical results and FEA modeling indicates that the FEM model constructed and the analysis approach taken correlated well with the test data from the CAPE drop test. This builds a foundation for using the model to examine possible modifications to the cage design, which will now be investigated.

Additional Analysis

After the baseline chassis model had been correlated against the CAPE drop-test results, which had a 19.03 mph

impact, a second FEA of the same model was carried out using a 37 mph (16.54 m/s) initial velocity. This velocity was chosen because it had been the impact velocity on the previous CAPE sled tests, and had constituted the velocity used in a previous IUPUI research project examining a Midget racecar chassis [21].

Non-Modified Chassis, 37 mph: Figure 16 shows the FEA-predicted deformation for the non-modified Sprint Car chassis undergoing a drop test with a 37 mph (16.540 m/s) velocity before impact. The plots in each frame of the image show total deformation for all nodes and elements. The second and third frames show that the front roll cage support buckled under the impact forces. There was also a considerable bending moment being placed on the horizontal rear suspension damper mounting tubes located aft of the driver's cell. The data were filtered using a CFC 60 Butterworth filter with a cutoff of 100 Hz. The peak impact acceleration was 523.41 G's. This is on the order of, but not as large as, either of the two CAPE sled tests, which recorded maximum G-loads of 716.20 [12] and 604.5 [13], respectively. Due to time constraints, the analysis time was limited to 0.100 s. With this analysis of the baseline Sprint Car chassis, it was possible to compare the data to a standard Sprint Car chassis with modified designs. Two further studies were then conducted using modified chassis designs to evaluate the possible influence that a crash attenuator would have on acceleration and velocities during the crash event. A modified roll cage design, utilizing an attenuator structure, had been developed previously by the IUPUI research team [21] for a Midget chassis utilizing data from the two prior CAPE sled tests.

Figure 17 shows where the attenuator tubes were attached to the top plane of the baseline Sprint Car chassis in a vertical orientation. Each portion of the attenuator was the same height relative to the top plane of the chassis. The tubes shared the same wall outside diameter and thickness of 1.50 in. by 0.095 in. as the exterior roll cage tubes.

Chassis with an Attenuator, 19.03 mph: A comparative analysis of a baseline Sprint Car chassis with the same original constraints and initial velocity of 19.03 mph was completed in order to assess the differences in peak acceleration and velocities. Figure 18 shows the deformation response of the Sprint Car chassis with the attenuator attached. Due to the high angle of the chassis top plane, the attenuator, when attached, extended from the top of the chassis at two different heights. The shortness of the analysis time prevented the chassis from completing its roll forward onto the platform.

The crash acceleration comparisons of the baseline Sprint Car chassis without an attenuator attached at the same 19.03

mph (8.51 m/s) initial velocity indicated a maximum acceleration for the impact of 166.14 G's for the attenuator model. That is a difference of 38.74% (94 G's) for peak impact acceleration between the baseline chassis and the attenuator-equipped chassis.

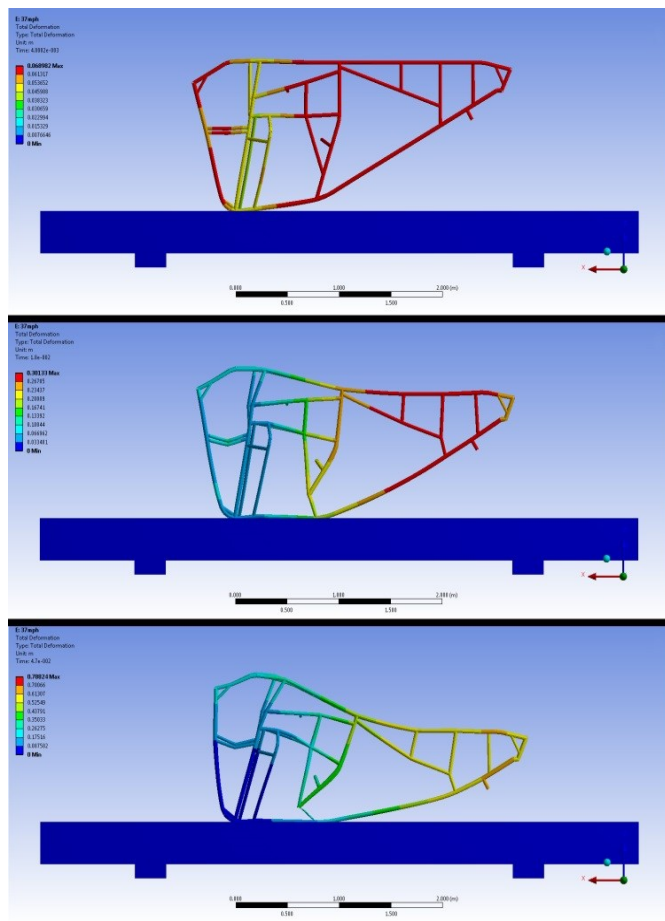


Figure 16. Deflections for Baseline Chassis under 37 mph Impact

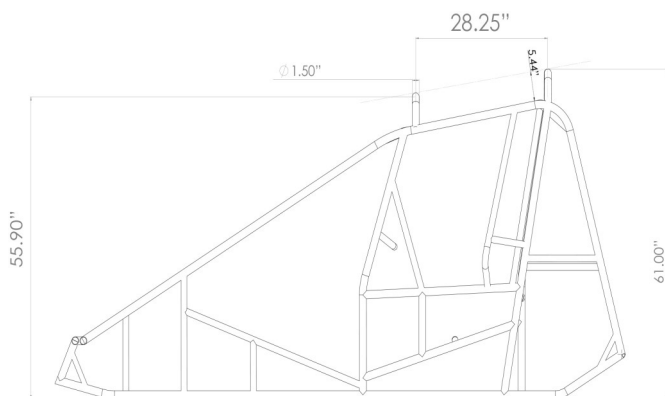


Figure 17. Attenuator Drawing

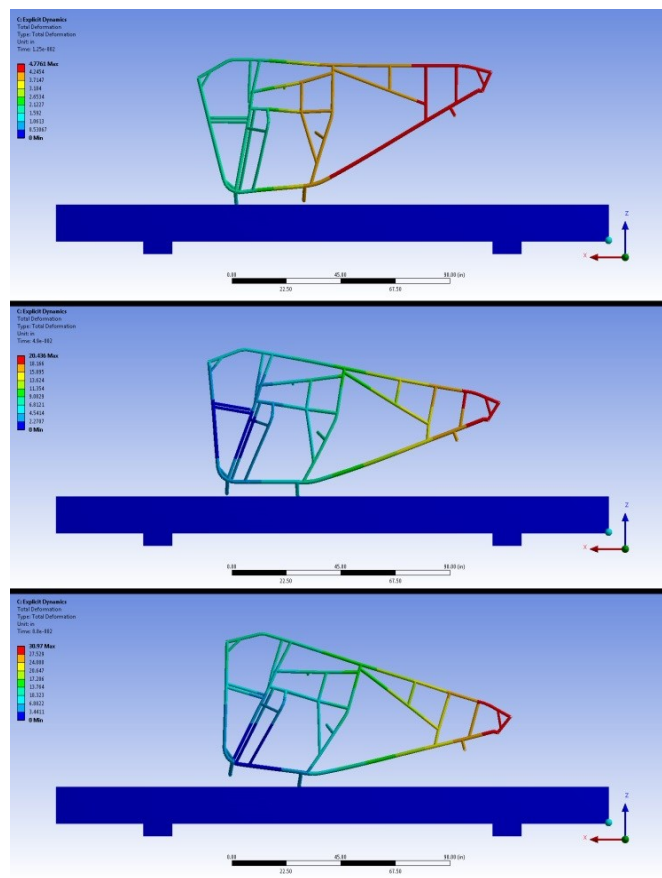


Figure 18. Deflections for Chassis with an Attenuator, 19.03 mph Impact

Chassis with an Attenuator, 37 mph: Figure 19 shows the deformation of the chassis equipped with an attenuator undergoing a crash with an initial velocity of 37 mph (16.54 m/s). The third frame shows that the front roll cage supports buckled under the impact forces after the complete deformation of the attenuator in frame two. This failure mode was similar to the one shown for the baseline chassis without an attenuator. The peak acceleration for the attenuator-equipped chassis at 37 mph was 177.16 G's, which was 345.65 G's less than the baseline chassis (523.41 G's) without an attenuator. This is a 66.15% decrease in peak crash acceleration over the baseline chassis. Further review of the velocity data revealed that the attenuated chassis' peak velocity for the crash occurred significantly later than the chassis which was not equipped with an attenuator. This rapid increase in velocity is believed to be due to the chassis rebounding from the platform. Even with the vast reduction in crash acceleration, the chassis' front roll cage supports would apparently fail. As a result, an additional series of FEAs were run in order to determine if slight modifications to the driver's cell had the potential to prevent the failure mode without a giant leap in crash acceleration.

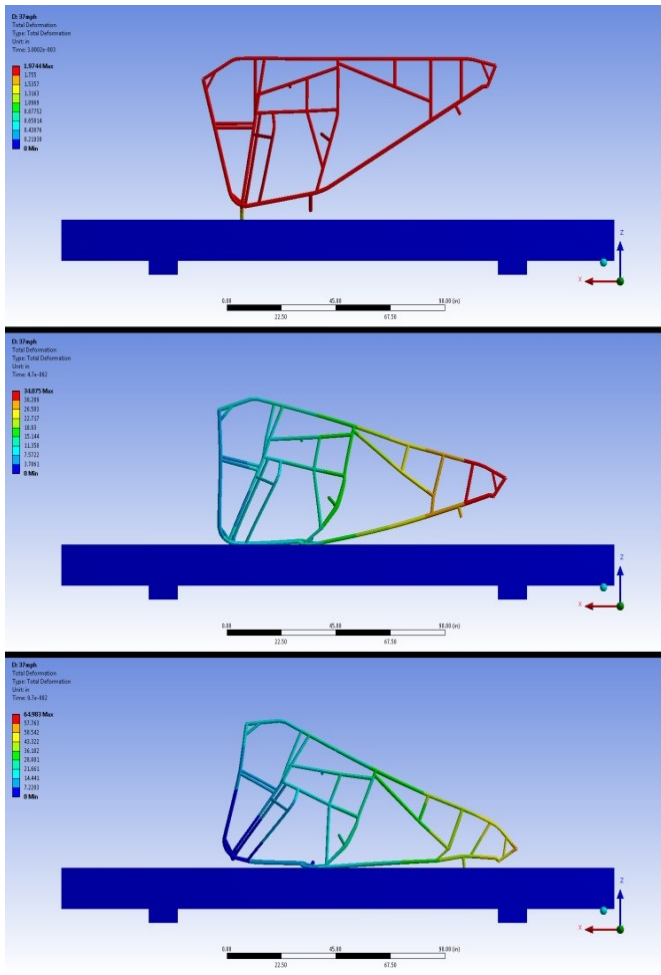


Figure 19. Deflections for Chassis with an Attenuator, 37 mph Impact

Modified Chassis A with an Attenuator, 19.03 mph: Figure 20 shows the modified chassis with added tubes highlighted. The modification included the addition of a 1.375 inch x 0.083 inch tube that connected the horizontal side impact structure of the chassis to the seat rails and then onward to the bottom frame rails at the rear-most transverse cross tube. The figure shows its placement in detail highlighted with yellow lines.

Figure 21 shows the deformation response for this further modified chassis. There appears to be a marked improvement in attenuator deflection and, thus, energy absorption. Notable results from this study are the further decrease in peak crash acceleration for the top-plane impact and an increase in rebound velocity, as more energy is transferred to the attenuator with the addition of the new support tube. With a peak of 147.48 G's, this configuration achieved a 45.62% reduction in peak crash acceleration.

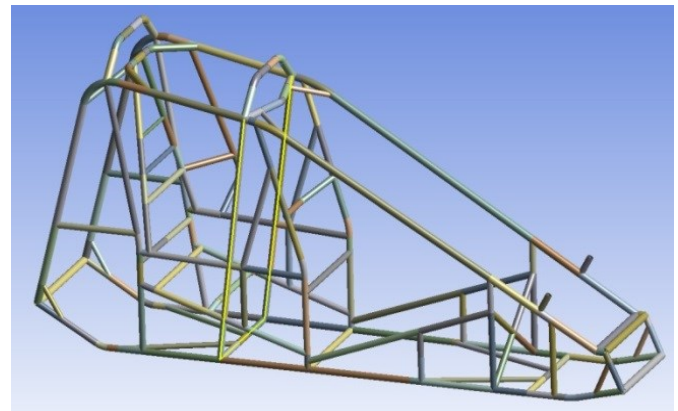


Figure 20. Modified Chassis A with New Tubes in Yellow

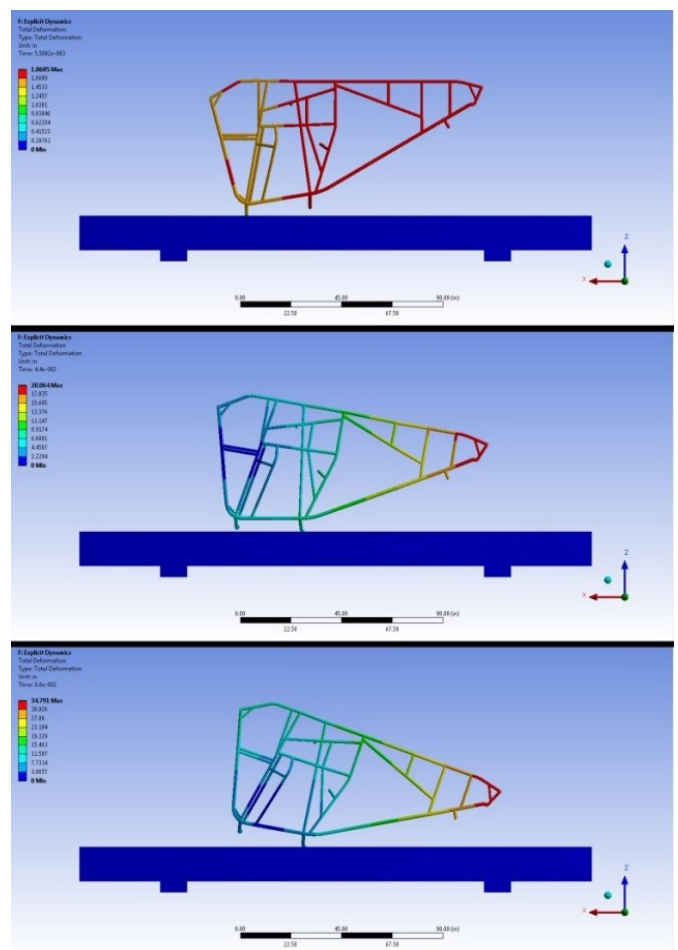


Figure 21. Deflections for Modification A, 19.03 mph Impact

Modified Chassis A with an Attenuator, 37 mph: Figure 22 shows the modified chassis undergoing a crash with a 37 mph (16.50 m/s) initial velocity. Frame two of this figure shows that full deformation of the attenuators occurred similarly to the baseline chassis equipped with an attenuator.

The key difference in this version from the previous cases is visible in frame three, which shows that the addition of the support tube connecting the attenuator to the bottom frame rails through the side impact structure allowed this chassis configuration to survive the crash without as much deformation to the top plane of the roll structure. There was, however, significant bending and warping of the lower rails of the driver's cell and rear end, as seen in frame three. The peak acceleration for this crash event was 175.33 G's, a 1.36% drop in peak acceleration as compared to the case with the attenuator alone.

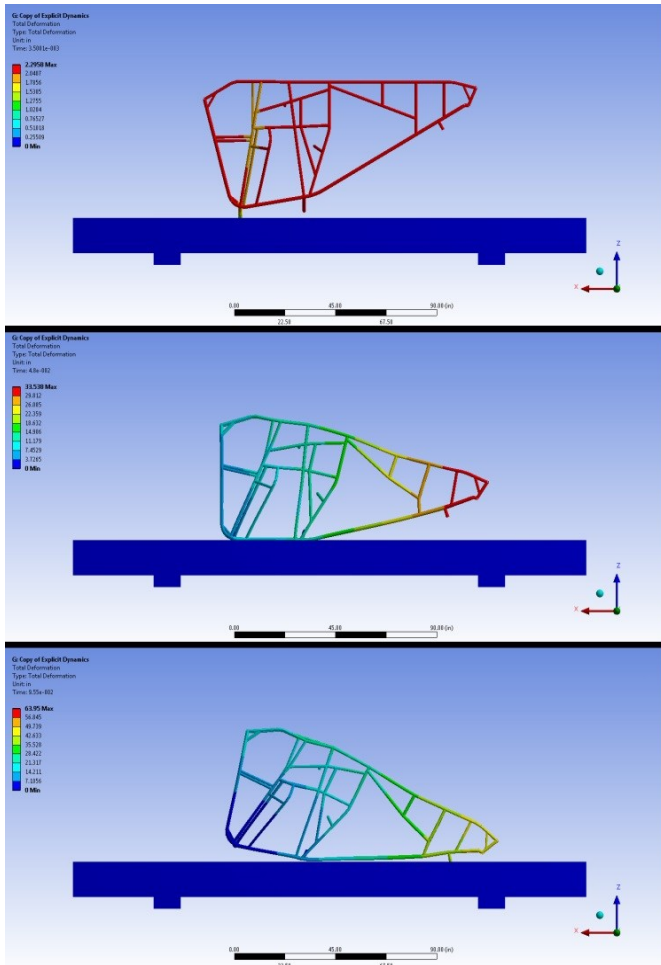


Figure 22. Deflections for Modification A, 37 mph Impact

Chassis Modification B, 19.03 mph: The final configuration tested attempted to further increase the triangulation of the Sprint Car roll cage by introducing an additional two 1.375 inch x 0.083 inch tubes to support and triangulate the side impact structure. As a Sprint Car chassis is asymmetrical across its longitudinal axis by design, the resulting chassis modifications highlighted in yellow in Figure 23 reflect this. Figure 24 shows the mode of deformation that the Modifica-

tion B chassis underwent for the crash study initiated with a starting velocity of 19.03 mph (8.54 m/s).

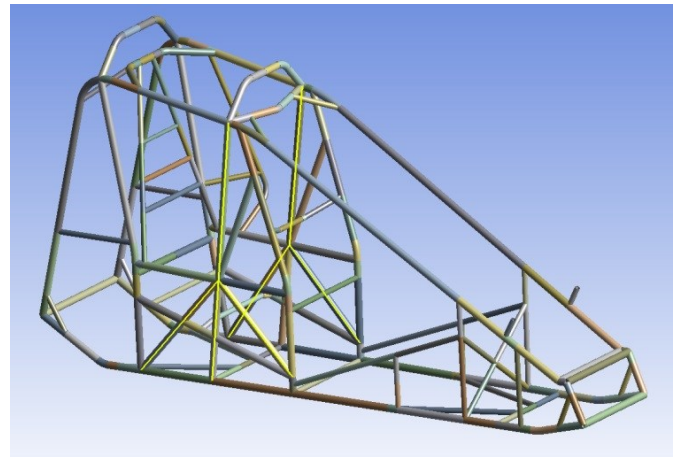


Figure 23. Modification B Chassis with Alternate Tube Positions Highlighted in Yellow

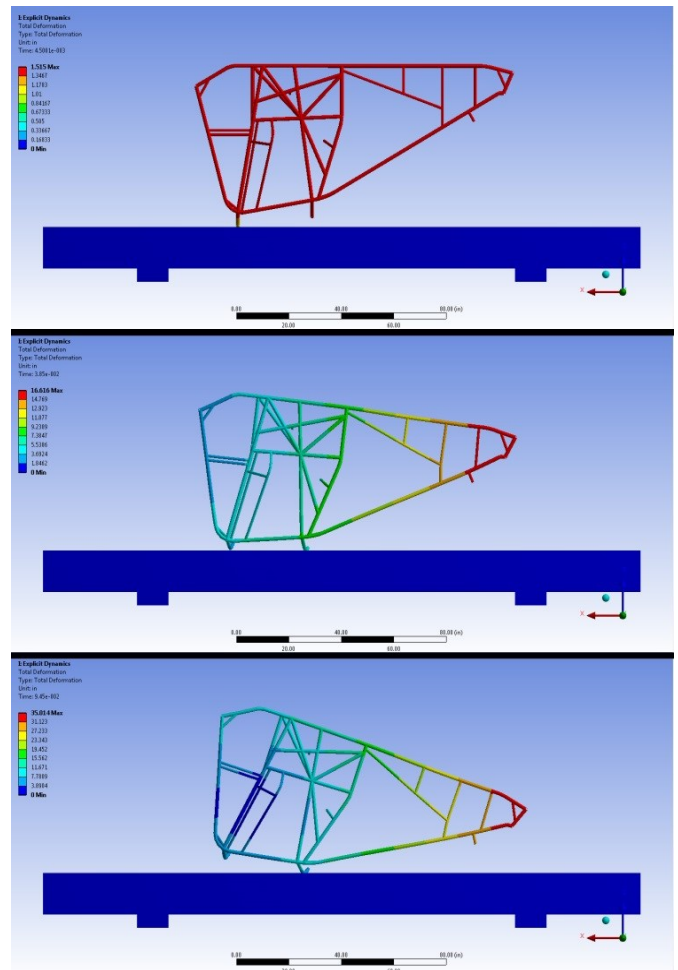


Figure 24. Deflections for Modification B, 19.03 mph Impact

Frame two of this figure shows the forward attenuator structure compressing towards the front of the vehicle under forward weight transfer. This motion continues into frame three and the final time step of the study, leaving the forward structure of the attenuator half crushed and the chassis model halfway through a forward rotation. It became apparent that the increase in triangulation present in Mod B prevented the chassis rails from warping and, as a result, maximized the energy transfer to the rear attenuator structure before the pitching moment occurred during the transition to frame three. The resulting peak acceleration for the Mod B chassis was the lowest of all of the studies at a maximum G-load of 60.70 G's, a 76.62% reduction in peak acceleration over the non-attenuated baseline chassis. The results of this study are significant, so much so that this study was reevaluated twice to ensure accuracy before moving forward.

Chassis Modification B, 37 mph: The final study completed in this series was the evaluation of the Mod B chassis during a crash with an initial velocity of 37 mph (16.54 m/s). Based on the results in previous section, the expectation for this study was a reduction in peak crash acceleration; however, the opposite occurred. Figure 25 shows Mod B's chassis deformation during the crash study. After an intense review of the deformation footage in ANSYS Mechanical, it appeared that, due to a significant increase in initial velocity, the rearward attenuator structure collapsed before forward weight transfer caused by the weight distribution of the vehicle could pitch the chassis forward, as was observed in frames two and three of Figure 24 in the previous case. This caused the momentum of the engine, whose mass was connected to the front roll cage tubes, to carry its momentum down and directly towards the platform. This was unlike the previous tests in which the rear attenuator structure collapsed at a slower rate and, thus, allowed for the momentum of the vehicle as a whole to pitch forward. The result was a complete collapse of the forward attenuator structure and a large increase in peak crash acceleration to 350.43 G's. While this is still 33.05% less than the baseline chassis, it is a 49.28% increase over the attenuator-equipped chassis and a 49.96% increase when compared to Mod A. It should be noted that the previous IUPUI study [21] utilizing the Midget chassis and the sled-test data [11-13] had arrived at a conclusion that stiffening the chassis could actually cause an increase in the G levels experienced in the crash scenario, despite use of an attenuator. That study concluded that the most beneficial modification examined was the addition of an attenuator to the baseline chassis.

Discussion

Based on the results of the 19.03 mph (8.54 m/s) correlation study between the ANSYS FEM and the empirical

CAPE drop-test results, there is a strong indication that the FEM can accurately predict the contact loads, peak velocities, and crash accelerations with a mean error of 9.11% for maximum acceleration and velocities. This is significantly better than originally expected when the use of line body elements was dictated due to project constraints. Based on ANSYS documentation and industry standard practices with regards to mesh size and model accuracy, it is conceivable that this error could be further reduced through an increase in mesh density or conversion to shell and solid elements [25].

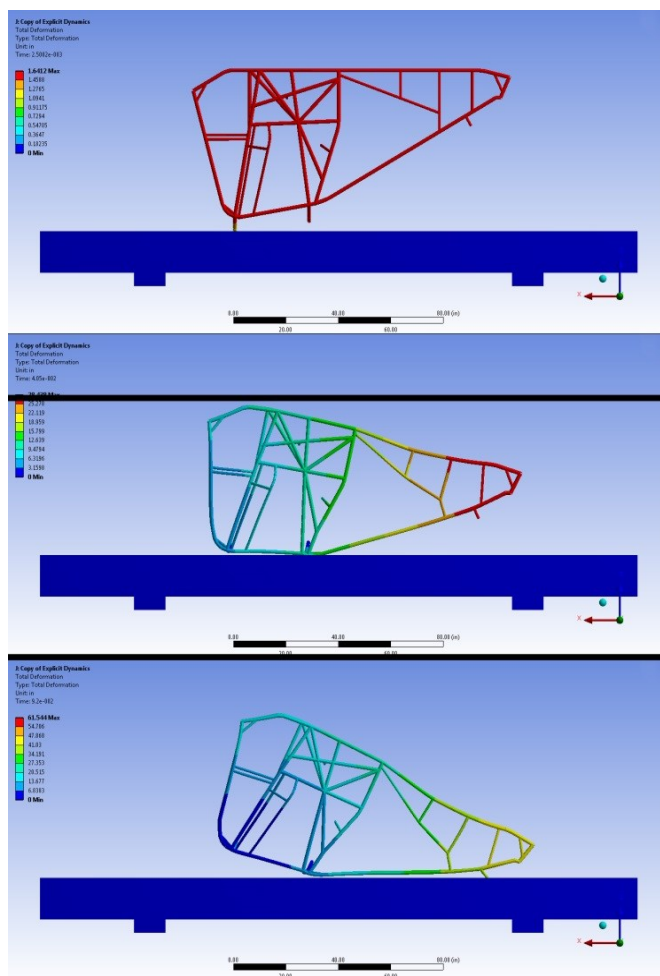


Figure 25. Deflections for Modification B, 37 mph Impact

Based on the results of using the correlated FEM to examine alternate chassis configurations, there is a significant case for equipping Sprint Cars with attenuator structures, possibly in conjunction with some chassis stiffening. This point is illustrated in Table 3, which summarizes the calculated percent change in the experienced G load used to evaluate all alternate chassis designs against the baseline-correlated Sprint Car chassis.

Table 3. Summary of G Loads and Percent Reduction

Chassis Configuration	Max G Load Experienced			
	19.03 mph		37 mph	
	G's	%Red	G's	%Red
Baseline	271	0%	523	0%
Baseline w/Attenuator	166	-38%	177	-66%
Mod A w/Attenuator	148	-46%	175	-67%
Mod B w/Attenuator	61	-77%	350	-33%

The table shows that the simple addition of an attenuator to the pre-existing chassis design resulted in a significant reduction in maximum G load experienced. However, when the correlation model was extended to compare the FEAs completed for the 37 mph (16.54 m/s) impacts, it appeared necessary to forego the use of the baseline chassis and attenuator, as the baseline chassis in either configuration failed to survive the 37 mph (16.54 m/s) study without the buckling of the front roll cage supports. Therefore, when considering a recommendation for the best chassis configuration, Modification A appears to be the best choice. It is worth noting that the results of the 37 mph (16.54 m/s) studies showed correlation to previous research completed by the authors and others on Midget race cars. That research found that the addition of an attenuator to a baseline (standard) Midget chassis reduced overall crash accelerations by 68%, as compared to a non-attenuator-equipped Midget chassis [21].

Modification B showed potential, based on its reduction in maximum G load at lower velocity impacts. However, further analysis is needed to determine if the dynamic behavior of the Modification B chassis configuration is consistent as initial velocity is varied from low speed (< 19.03 mph) to much greater than the 37 mph initial velocity. If the dynamic response of the Mod B chassis, as previously discussed, does indeed persist beyond the 37 mph collision speed, research should focus on the effect that an increase in front attenuator stiffness has on the peak crash acceleration. The hypothesis is that an increase in front attenuator stiffness will net a reduction in the peak crash acceleration and velocities for high-speed crashes; however, this may result in an increase in peak acceleration for low-speed impacts.

Conclusions

- Correlation between drop-test empirical data and Finite Element Model predictions was strong.
- Using a correlated model of a Sprint Car chassis to examine alternate configurations indicated that an attenuator added to the top of the roll cage significantly

reduced the G loads experienced in a top-plane impact crash.

- Performance of an attenuator design correlated well with the results determined previously, which utilized a different chassis and analysis approach, further supporting the viability of an attenuator design.
- Stiffening of the chassis affected the ability of the attenuator to perform the job of reducing G loads and, therefore, any additional chassis stiffening must be carefully considered.

Acknowledgements

The authors would like to acknowledge the support of the Center for Advanced Product Evaluation (CAPE), Spike Chassis Inc., AED Motorsport Products, the MotorSport Engineering Conference, and the Motorsports Engineering program of Indiana University Purdue University Indianapolis.

References

- [1] Hungness, C., Fox, J. C., & Blackwell, J. (1981). *USAC Sprint History: Twenty-Five Years of United States Sprint Car History*. Indianapolis, IN: Hungness-Whitt.
- [2] Donaldson, G. (2002). *Formula 1: The Autobiography*. New York, NY: Weidenfield & Nicolson.
- [3] Patalak, J., & Gideon, T. (2013). Occupant Rollover Protection in Motorsports, *SAE International Journal of Transportation Safety*, 1(2), 386-398.
- [4] Morrison, R. H. (1913). Playing Polo In Autos. *Technical World Magazine*, 4, 103-104.
- [5] Perry, R. (1924, July 3). Miami's New Sport Will Provide More Thrills For Fans. *Miami Daily News and Metropolis*, p. 10.
- [6] Berger, M. L. (2001). *The Automobile in American History and Culture: A Reference Guide*. Westport, CT: Greenwood Press.
- [7] Anderson, L. (2013, May 25). Life and Death and the Heart of American Racing. *Sports Illustrated*, p. 20.
- [8] Hinton, E. (2011, February 7). Earnhardt's Death a Watershed Moment. *ESPN*. Retrieved July 1, 2015 from <http://sports.espn.go.com/rpm/nascar/cup/columns/story?id=6116145>.
- [9] Smith, M. (2012, November 24). Broken body, Repaired Soul. *ESPN*. Retrieved July 1, 2015 from http://espn.go.com/racing/nascar/cup/story/_/id/8662362/nascar-shane-hmiel-broken-body-spirit-soars.
- [10] Schwab, G., Alexander, A., & Scott, D. (2014, August 18). More than 520 people have died in U.S.

auto racing in past 25 years. *Charlotte Observer*, Retrieved July 1, 2015 from <http://www.charlotteobserver.com/sports/nascar-auto-racing/thatsracin/article9152399.html>.

- [11] Porter, J. (2011). *Report CTR04921*. Center for Advanced Product Evaluation, Westfield, IN.
- [12] Hoover, R. (2012). *Report CTR07288*. Center for Advanced Product Evaluation, Westfield, IN.
- [13] Porter, J. (2011). *Report CTR05154*. Center for Advanced Product Evaluation, Westfield, IN.
- [14] Society of Automotive Engineers International. (2011). *Rollover Testing Methods* (SAE J2926_201108). Detroit, MI.
- [15] National Highway Traffic Safety Administration. (2015). *Roof Crush Resistance* (NHTSA-1999-5572). Detroit, MI.
- [16] Forrest, S. M., Herbst, B., Meyer, S. E., & Orton, T. (2003). Evaluating 'Real World' Roof Strength Through Inverted Drop Testing. *Proceedings of the Summer Bioengineering Conference* (pp. 1-2). Key Biscayne, FL.
- [17] Society of Automotive Engineers International. (2015). *Instrumentation for Impact Test Part 1 Electronic Instrumentation* (SAE J211-1). Warrendale, PA, USA.
- [18] ANSYS Inc. (2015, September 01). High Performance Computing. *ANSYS.com*. Retrieved July 1, 2015 from <http://www.ansys.com/staticassets/ANSYS/staticassets/resourcelibrary/brochure/ansys-hpc-brochure.pdf>.
- [19] ANSYS, Inc. (2015, September 01). *ANSYS Mechanical Users Guide*. Canonsburg, PA.
- [20] ANSYS, Inc. (2015). *Configuring High Performance Computing*. Canonsburg, PA.
- [21] Hankins, G. T., Krajnik, K., Galedrige, B., Sakha, S., Hylton, P., & Otoupal, W. (2014). Improving Safety Structures on Sprint and Midget Race Cars. *SAE Paper 2014-01-0561* (pp 1-8).
- [22] Safety Foundation Inc. (2015, September 01). Driver Restraint Assemblies. *SFI Foundation*. Retrieved July 1, 2015 from http://www.sfifoundation.com/wp-content/pdfs/specs/Spec_16.1_022614.pdf.
- [23] Humanetics Innovative Solutions. (2015, September 01). Hybrid II 50th Male Dummy. *Humanetics Innovative Solutions* Retrieved July 1, 2015 from <http://www.humaneticsatd.com/crash-test-dummies/frontal-impact/hybrid-ii-50th>.
- [24] United States Auto Club. (2014). *USAC Rule Book*. Indianapolis, IN.
- [25] Lee, H. (2012). *Finite Element Simulations with ANSYS Workbench 14 : Theory, Applications, Case Studies*. Mission, KS: Ansys, Inc.

Biographies

GRANT HANKINS is a graduate research assistant for the Motorsports Engineering Program at Indiana University Purdue University Indianapolis. He has been extensively involved in research dealing with the creation and correlation of models representing motorsport chassis for vehicles ranging from Top Fuel dragsters to short-track Midget and Sprint cars. Mr. Hankins may be reached at ghankins@iupui.edu

PETE HYLTON is an associate professor and director of motorsports engineering at Indiana University Purdue University Indianapolis. He has industry experience in both aerospace and motorsports and is responsible for the creation of the first Motorsports Engineering B.S. degree program in the U.S. He is the author of several books on motorsports history; is a member of the Sports Car Club of America's Hall of Fame; and, is a Fulbright Scholar. Dr. Hylton may be reached at phylton@iupui.edu

RYAN HOOVER is Director of the Center for Advanced Product Evaluation, a facility conducting advanced testing of automotive products. He was in charge of the data acquisition of the three barrier crash tests discussed in this paper and used to correlate the model. He may be reached at rhoover@imminet.com

KEVIN TRIBBETT is General Manager for the Center for Advanced Product Evaluation, a facility conducting advanced testing of automotive products. He was in charge of the data acquisition of the drop test discussed in this paper and used to correlate the model. He may be reached at ktribbett@imminet.com

GASOLINE VAPOR LIQUID EQUILIBRIUM USING THE PENG ROBINSON EQUATION OF STATE AND RACHFORD-RICE EQUATION

Mebougna L. Drabo, Alabama A&M University; Marcus D. Ashford, University of Alabama

Abstract

By utilizing an equation of state (EOS), both the liquid and vapor phases of a fluid or mixture can be described. This can be accomplished by using the Peng-Robinson EOS. Such an application of state is presented here. Also included is the computation of vapor-liquid equilibrium (VLE) as well as an estimation of pure component properties. In this study, a flash distillation problem for a mixture of hydrocarbons (n-butane, isopentane, toluene, iso-octane, and decane) was solved using the Rachford-Rice equation. The phi-phi technique, utilizing the Peng Robinson EOS as a basis, was used to calculate the equilibrium constants. All results were in agreement with those given by the National Institute of Standards and Technology (NIST) Thermophysical Properties of Hydrocarbon Mixtures database (SUPERTRAPP Software).

Introduction

The automobile is one of the most depended on machines of the 21st century; it has affected how we live in both positive and negative ways and has changed human society as a result. With an ever-increasing number of automobiles being used on American roads, it has been found that automobiles are significant contributors to the emission of hydrocarbons (HC) nationwide. The U.S. Environmental Protection Agency (EPA) in 2012 found that volatile organic compounds (VOC) emitted by highway vehicles contributed 20% to the total national emission, second only to industrial processes [1]. High hydrocarbon emissions occur when automobiles undergo transient operation, in particular when cold-starting. During a cold-start, there is a rapid change in pressure and temperature which causes HC to be emitted. It has also been found that 60-95% of HC emissions are tail-pipe emissions [2-4].

The geographical origin of gasoline is an important factor when it comes to its volatility; this is also affected by the seasonal blend of the gasoline. Dependent on the fuel type, the amount of fuel vaporization varies accordingly and may lead to the need to over fuel in order to generate a combustible mix. Most engine controllers are not programmed to recognize the type of fuel in use and, therefore, cannot auto-

matically compensate for the difference in volatility of the gasoline. These onboard computers are usually programmed to assume certain scenarios like hesitation fuel. Hesitation fuel is the lowest volatility fuel typically expected to be found in the fuel supply. Consequently, fueling can be 10-20 times the stoichiometric amount [2-4].

The problem of fueling inaccuracy is not limited to the cold-start transient, but also can occur in engines at operating temperatures, where some of the injected fuel joins intake port and in-cylinder fuel films. The fuel film's mass and mass rate of change are functions of pressure, temperature, and fuel composition. When throttle-opening transients occur there is an increase in pressure; this causes the film mass to grow and a lean excursion to be created unless additional fuel is injected to compensate for this. However, throttle-closing transients cause rich excursions, due to the decrease in fuel film mass.

Fuel Major Component Model

To carefully design a multi-component fuel that would model California's Phase II gasoline, researchers from the Massachusetts Institute of Technology (MIT) divided hundreds of gasoline components into various groups based upon the carbon number and the type of organic compound (alkanes, alkenes, aromatics, etc.). After being separated, the components in the various groups were characterized by species with the greatest abundance of the group. That particular species carries the weight percentage for the whole group. The distillation curve that resulted from 15 component fuels was compatible with California's Phase II reformulated gasoline actual distillation curve [4].

In a study conducted at Doshisha University to model the evaporation process for multi-component fuels, Senda et al. [5] discovered that hydrocarbons having lower boiling points evaporate at temperatures greater than their own boiling point and hydrocarbons with higher boiling points actually evaporate at temperatures lower than their own boiling points. It is apparent that heavier hydrocarbons inhibit the evaporation of lighter hydrocarbons. For temperatures lower than 60°C, the relatively small percentage of light hydrocarbons was constricted in the mixture by heavier hydrocar-

bons and, thus, vapor condensate was not able to form [5]. The major component fuel was formulated for modeling the evaporative process of multi-component fuels and was not formulated to copy the composition of standard commercial gasoline.

A major component fuel model for the Active Vapor Utilization System (AVUS) testing was formulated to contain 25% isopentane, 22.5% iso-octane, 17.5% toluene, 17.5% heptanes, and 17.5% hexane by volume. It was tested with AVUS and the production rates of the condensate were similar to that of commercial gasoline. This fuel contained 45.5% light ends, while MIT fuel contained about 43% light ends. The mix of light hydrocarbons was also similar; however, the MIT fuel contained a very small percentage of heavier hydrocarbons that the UA (University of Alabama) fuel did not. The UA five-component fuel should simplify analysis, while reasonably mirroring commercial gasoline. Isopentane, n-hexane, n-heptane, 2, 2, 4-trimethylpentane, toluene, and ethanol were used to produce the UA five-component fuel and its E85 variant [6].

A comparison of various fuel compositions [4], [5], [7] is shown in Table 1. The components are categorized by carbon numbers and aromatics. Researchers from MIT added 13% methyl tertiary butyl ether (MTBE) to the fuel to increase octane rating. MTBE is listed under C5 for fuel comparisons.

Table 1. Comparisons of Major Component Fuels by Carbon Number

Component	Santoso-MIT (% mol)	Senda-Doshisha University (% mol)	UA Fuel (% mol)
C4	2.0	4.0	
C5+MTBE	29.6	35.0	28.2
C6	10.6	12.0	17.3
C7	11.9	6.0	15.5
C8	16.1	12.0	17.6
C9	2.6		
C10	0.7		
C11	0.4		
C12	0.3		
MTBE	13		
Aromatics	25.8	30.0	21.4

Problem Statement

Several techniques for estimating fuel volatility at engine operating conditions have been developed. The majority of these models utilize surrogate fuels with a restricted number of components (<20) to represent gasoline [4], [8], [9]. The principal purpose of this current study was to develop an algorithm that can compute fuel volatility via the principles of phase equilibrium. More specifically, if the liquid fuel composition is known, the algorithm will calculate the composition of liquid and vapor fuel at any given temperature and pressure. The problem statement is illustrated schematically in Figure 1.

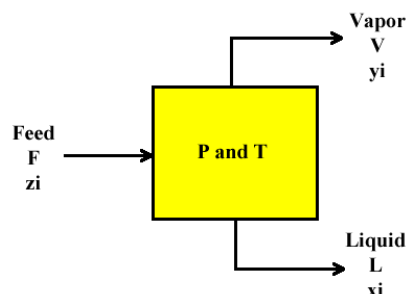


Figure 1. Depiction of the Problem Statement

The basic algorithm is comprised of two modules: a cubic equation of state for predicting phase equilibrium and the Rachford-Rice equation to determine fuel composition. The 1978 version of the Peng-Robinson Equation of State (EOS) (hereafter referred to as PR78) was employed for phase equilibrium calculations [7], [10].

Estimation of Pure Component Properties

The Peng-Robinson equation of state (PR EOS) originally developed in 1976 [7] was proposed in order to satisfy the following objectives:

- Parameters had to be “expressed in terms of the critical properties and the acentric factor” [11].
- Provide reasonable accuracy near “the critical point, particularly for calculations of the compressibility factor and liquid density” [11].
- A single binary interaction parameter is required for the mixing rule that is independent of composition and temperature pressure.

The PR78 Model

For a pure component, PR78 is given by Equations (1)-(5b).

$$P = \frac{RT}{v - b_i} - \frac{a_i(T)}{v(v + b_i) + b_i(v - b_i)} \quad (1)$$

where,

$$R = 8.314475 \frac{J}{mol * K} \quad (2a)$$

$$b_i = 0.07780 \frac{RT_{c,i}}{P_{c,i}} \quad (2b)$$

$$a_i = 0.45724 \frac{(RT_{c,i})^2}{P_{c,i}} \left[1 + m_i \left(1 - \sqrt{T_r} \right) \right]^2 \quad (3)$$

$$T_r = \frac{T}{T_{c,i}} \quad (4)$$

and,

if $w_i > 0.491$:

$$m_i = 0.379642 + 1.48503w_i - 0.164423w_i^2 + 0.01666w_i^3 \quad (5a)$$

if $w_i \leq 0.491$:

$$m_i = 0.37464 + 1.54226w_i - 0.26992w_i^2 \quad (5b)$$

where, P is the pressure; T the temperature; v is the molar volume; R is the ideal gas constant; w is the acentric factor; P_c is the critical pressure; and, T_c is the critical temperature [12].

PR78 was used in this study. To apply this EOS to mixtures case, classical mixing rules were utilized in order to compute the values of interest in the mixture (a and b). The mixture (a and b) is given by Equation (6):

$$a = \sum_{i=1}^N \sum_{j=1}^N z_i z_j \sqrt{a_i a_j} [1 - k_{ij}(T)] \quad (6)$$

$$b = \sum_{i=1}^N z_i b_i$$

$$k_{ij} = \frac{-\frac{1}{2} \left[\sum_{k=1}^{N_g} \sum_{l=1}^{N_g} (\alpha_{ik} - \alpha_{jk})(\alpha_{il} - \alpha_{jl}) \left(\frac{298.15}{T} \right)^{\left(\frac{B_{kl}}{A_{kl}} - 1 \right)} \right] - \left(\frac{\sqrt{a_i(T)}}{b_i} - \frac{\sqrt{a_j(T)}}{b_j} \right)^2}{2 \sqrt{a_i(T) \cdot a_j(T)}} \quad (7)$$

where, N represents the number of components in the mixture and z_k the mole fraction of component “ k ” in a mixture. More information about Equation (6) can be found in the study by Vitu et al. [13]. Also, $k_{ij}(T)$ is the binary interaction parameter describing molecular interactions among molecules “ i ” and “ j ” [13], when $i = j$, $k_{ij} = 0$.

In order to get a predictive model and to describe PR78, k_{ij} , which relies on temperature, was determined by the contribution method of Equation (7) [12], [13]. The binary interaction parameter is given in Equation (7) [13-26]. The PR78 model was defined by Equations (1)-(7). In this study, the elementary groups considered are shown in Table 2.

Table 2. Elementary Groups

Group	Molecule
1	CH ₃
2	CH ₂
3	CH
4	C
5	CH ₄
6	C ₂ H ₆
7	CH _{aro}
8	C _{aro}
9	C _{fusedAro}
10	CH _{2cyc}
11	CH _{cyc} , C _{cyc}
12	CO ₂
13	N ₂

An example for n-butane divided into groups can be found in the study by Drabo [27]. More details on α_{ik} , $A_{kl}=A_{lk}$, and $B_{kl}=B_{lk}$ can be found in the study by Privat et al. [12]. Further descriptions of obtaining A_{kl} and B_{kl} can be found in studies by Jaubert & Coniglio [14], Jaubert & Mutelet [15], Jaubert et al. [16-18], Privat et al. [20-23], Privat & Jaubert [24], and Qian et al. [25].

Vapor-liquid Equilibrium

The vapor-liquid equilibrium of a mixture can be computed using PR EOS. The liquid and vapor mole fractions are related by Equation (8):

$$y_i = K_i x_i \quad (8)$$

where, K_i is the equilibrium constant.

The PR EOS is part of a family of equations referred to as cubic because the compressibility factor (Z), given by Equations (9)-(13), which is a key in the cubic equation of Equation (9), written for a multi-component mixture, where the mixing and combining rules have been used:

$$Z^3 + (1 - B)Z^2 + (A - 3B^2 - 2B)Z + (-AB + B^2 + B^3) = 0 \quad (9)$$

where,

$$A = \sum_{i=1}^C \sum_{j=1}^C y_i y_j A_{ij} \text{ or } \sum_{i=1}^C \sum_{j=1}^C x_i x_j A_{ij} = 0 \quad (10)$$

$$A_{ij} = (A_i A_j)^{0.5} (1 - k_{ij}) = 0 \quad (11)$$

$$B = \sum_{i=1}^C y_i B_i \text{ or } \sum_{i=1}^C x_i B_i = 0 \quad (12)$$

$$A_i = 0.45724 * \left[1 + m_i (1 - \sqrt{T_r}) \right]^2 * \frac{PT_{ci}^2}{P_{ci} T^2} \text{ and } B_i = 0.07780 \frac{PT_{ci}}{P_{ci} T} \quad (13)$$

and where, m_i can be obtained from Equations (5a) and (5b) and the binary interaction parameter, k_{ij} from Equation (7).

Equation (14) describes the equilibrium constant. The equilibrium constants are obtained using the phi-phi method:

$$K_i = \frac{\phi_{li}}{\phi_{vi}} \text{ for } i = 1 \text{ to } C \quad (14)$$

where, ϕ_{vi} is given by Equation (15)

A related expression was obtained for the liquid phase fugacity coefficient, ϕ_{li} , by changing the gas phase compressibility factor, Z_{vi} , with Z_{li} . These compressibility factors are the largest and smallest roots of Equation (9). As a final point, a flash calculation was performed in order to obtain the vapor fraction by using the Rachford-Rice equation. This famous equation is given by Equation (16):

$$\sum_i \frac{z_i (K_i - 1)}{1 + \phi (K_i - 1)} = 0 \quad (16)$$

where, z_i is the mole fraction of component i in the feed.

Results and Discussion

In order to test the precision of the algorithm, its predicted fuel mixtures were inputted into the NIST SUPERTRAPP [28] vapor-liquid equilibrium program in order to compare calculated HC concentrations. Tests were conducted with different mixture blends. The proposed mixture ingredients were butane, isopentane (methylbutane), toluene, iso-octane (2, 2, 4-trimethylpentane), and decane, chosen because they collectively account for the majority the content of commercially available pump gasoline. In addition, these components cover the majority of the boiling point range of gasoline. The fuel blend was tested at a temperature of 300 K. The pressure was held constant at 100 kPa throughout all the tests. This was done because the initial pressure for engine starts is usually atmospheric pressure. The SUPERTRAPP computations were considered the reference standard, as its precision has been verified in studies by Santosa & Cheng [4] and Cowert & Hamilton [29]. The results of the simulation can be found in Tables 3 [27] and 4 [27]. Tables 3 and 4 show the SUPERTRAPP and the algorithm simulation results for the mixture.

The vapor composition for n-decane showed an order of magnitude of 10^{-5} , which was very small compared to other fuels, which can be explained by the fact that a temperature of 300 K is much lower than the n-decane boiling point. Therefore, absolute error, even though small, would result in a significantly higher percentage of error. Tables 5 [27] and 6 [27] show 1-K_{ij} values for different hydrocarbons. Table 5 was obtained using the algorithm and Table 6 was obtained using SUPPERTRAP software.

$$\phi_{vi} = \exp \left[(Z_v - 1) \frac{B_i}{B} - \ln(Z_v - B) - \frac{A}{2\sqrt{2}B} \left(\frac{2\sum_j y_j A_{ij}}{A} - \frac{B_i}{B} \right) \ln \left(\frac{Z_v + (1 + \sqrt{2})B}{Z_v + (1 - \sqrt{2})B} \right) \right] \quad (15)$$

Table 3. SUPERTAPP Simulation Results for the Mixture (n-butane, isopentane, toluene, 2, 2, 4-trimethylpentane 224TMP, and n-decane) at 300 K and 1 bar

Component	Feed	Liquid (mole fraction)	Vapor (mole fraction)
n-butane	0.33	0.306405	0.023594
isopentane	0.165	0.160221	0.00478
toluene	0.165	0.164692	0.000307
2,2,4-trimethylpent	0.175	0.174555	0.000444
n-decane	0.165	0.164989	1.09551*10 ⁻⁵

Table 4. Algorithm Simulation Results for the Mixture (n-butane, isopentane, toluene, 2, 2, 4-trimethylpentane 224TM, and n-decane) at 300 K and 1 bar

Component	Feed	Liquid (mole fraction)	Vapor (mole fraction)
n-butane	0.33	3.263	0.041
isopentane	0.165	6.241	0.212
toluene	0.165	6.061	3.357
2,2,4-trimethylpent	0.175	5.728	2.164
n-decane	0.165	6.607	91.281

Values are given as Percentage Error Compared to SUPPERTRAPP values.

Table 5. 1-Kij Values for Different Hydrocarbon Values Obtained using the Algorithm

	n-butane	Iso-pentane	Toluene	224TMP	n-decane
n-butane	1	1.1826	-3.8528	0.8137	0.7363
Iso-pentane	1.1826	1	0.674	1.1223	0.9865
Toluene	-3.8528	0.674	1	1.008	1.0166
224TMP	0.8137	1.1223	1.008	1	0.9948
n-decane	0.7363	0.9865	1.0166	0.9948	1

Conclusions

A simple flash calculation of a mixture of hydrocarbon using the Rachford-Rice equation with the capability of predicting liquid-fuel composition was established in this study. The algorithm has proven to be very effective and relatively precise.

Table 6. 1-Kij Values for Hydrocarbon Values Obtained using SUPERTRAPP Software

	n-butane	Iso-pentane	Toluene	224TMP	n-decane
n-butane	1	0.997	0.995	0.986	1.003
Iso-pentane	0.997	1	0.996	0.992	1.010
Toluene	0.995	0.996	1	0.979	1.012
224TMP	0.986	0.992	0.979	1	1.018
n-decane	1.003	1.010	1.012	1.018	1

References

- [1] Our Nation's Air. (n.d.). Retrieved from www.epa.gov/airtrends.
- [2] Ashford, M. D., & Matthews, R. (2006). On-Board Generation of a Highly Volatile Starting Fuel to Reduce Automobile Cold-Start Emissions. *Environmental Science and Technology*, 40(18), 5770-5777.
- [3] Kane, E., Metha, D., & Frey, C. (2001). Refinement of a Dedicated E85 1999 Silverado with Emphasis on Cold Start and Cold Drivability. *SAE Technical Paper*, doi: 10.4271/2001-01-0679.
- [4] Santosa, H., & Cheng, W. K. (2002). Mixture Preparation and Hydrocarbon Emissions Behaviors in the First Cycle of SI Engine Cranking. *SAE Technical Paper*, doi: 10.4271/2002-02-2805.
- [5] Senda, J., Higaki, T., Sagane, Y., & Fujimoto, H. (2000). Modeling and Measurement on Evaporative Process of Multicomponent Fuels. *SAE Paper*, doi: 10.4271/2000-01-0280.
- [6] Crawford, J. (2009). *Study of a Cold Start Fuel Produced By An Active Vapor Utilization System For Use in Gasoline Powered Vehicles*. Master thesis, The University of Alabama in Tuscaloosa.
- [7] Peng, D., & Robinson, D. B. (1976). A New Two Constant Equation of State. *Industrial & Engineering Chemistry Fundamentals*, 15, 59-64.
- [8] Chen, K., DeWitte, K., & Cheng, W. (1994). A Species-Based Multi-Component Volatility Model for Gasoline. *SAE Technical Paper*, doi: 10.4271/941877.
- [9] Luan, Y., Henein, N. A., & Tagomori, M. K. (2006). Port Fuel Injection Gasoline Engine Cold Start Fuel Calibration. *SAE Technical Paper*, doi: 10.4271/2006-01-1052.
- [10] Robinson, D. B., & Peng, D. Y. (1978). The Characterization of the Heptanes and Heavier Fraction for the GPA Peng-Robinson Programs. *Gas Processors Association, Research Report*, RR-28.

- [11] Peng Robinson Cubic Equation of State. (n.d.). Retrieved from <http://www.eq-comp.com/peng-robinson-cubic-equation-of-state/>
- [12] Privat, R., Garcia, F., & Jaubert, J. N. (2009). Ethanol-Hydrocarbon Blend Vapor Prediction. *Proceedings of ASME Turbo Expo: Power for Land, Sea and Air, Paper # GT2009-59336*, (pp.1-11). Orlando, FL.
- [13] Vitu, S., Privat, R., Jaubert, J. N., & Mutelet, F. (2008). Predicting the phase equilibria of CO₂ + hydrocarbon systems with the PPR78 Model (PR EOS and kij calculated through a group contribution method). *J. Supercrit. Fluids*, 45(1), 1-26.
- [14] Jaubert, J. N., & Coniglio, L. (1999). The Group Contribution Concept: A Useful Tool to Correlate Binary Systems and to predict the Phase Behavior of Multicomponent Systems Involving Supercritical CO₂ and Fatty Acids. *Ind. Eng. Chem. Res.*, 38(12), 5011-5018.
- [15] Jaubert, J. N., & Mutelet, F. (2004). VLE predictions with the Peng-Robinson equation of State and temperature Dependent kij calculated through a group contribution method. *Fluid Phase Equilib.*, 224(2), 285-304.
- [16] Jaubert, J. N., Borg, P., Coniglio, L., & Barth, D. (2001). Phase Equilibria Measurements and Modeling of EPA and DHA Ethyl Esters in Supercritical Carbon Dioxide. *J. Supercrit. Fluids*, 20(2), 145-155.
- [17] Jaubert, J. N., Coniglio, L., & Denet, F. (1999). From the Correlation of Binary Systems Involving Supercritical CO₂ and Fatty Acid Esters to the Prediction of (CO₂-Fish Oils) Phase Behavior. *Ind. Eng. Chem. Res.*, 38(8), 3162-3171.
- [18] Jaubert, J. N., Vitu, S., Mutelet, F., & Corriou, J. P. (2005). Extension of the PPR78 Model Predictive 1978, Peng Robinson EOS With Temperature Dependent kij Calculated Through a Group Contribution Method to Systems Containing Aromatic Compounds. *Fluids Phase Equilib.*, 237(1-2), 193-211.
- [19] Mutelet, F., Vitu, S., Privat, R., & Jaubert, J. N. (2005). Solubility of CO₂ in Branched Alkanes in Order to Extend the PPR78 Model Predictive 1978 Peng Robinson EOS With Temperature Dependent kij Calculated Through a Group Contribution Method to Such Systems. *Fluids Phase Equilib.*, 238(2), 157-168.
- [20] Privat, R., Jaubert, J. N., & Mutelet, F. (2008). Addition of the Nitrogen Group to the PPR78 Model (Predictive 1978, Peng Robinson EOS With Temperature Dependent kij Calculated Through a Group Contribution Method). *Ind. Eng. Chem. Res.*, 47(6), 2033-2048.
- [21] Privat, R., Jaubert, J. N., & Mutelet, F. (2008). Addition of the Sulfhydryl Group (-SH) to the PPR78 Model (Predictive 1978, Peng Robinson EOS With Temperature Dependent kij Calculated Through a Group Contribution Method). *J. Chem. Thermodyn.*, 40(9), 1331-1341.
- [22] Privat, R., Jaubert, J. N., & Mutelet, F. (2008). Use of the PPR78 Model to Predict New Equilibrium Data of Binary Systems Involving Hydrocarbons and Nitrogen. Comparison With Other GCEOS. *Ind. Eng. Chem. Res.*, 47(19), 7483-7489.
- [23] Privat, R., Mutelet, F., & Jaubert, J. N. (2008). Addition of the Hydrogen Sulfide Group to the PPR78 Model (Predictive 1978, Peng Robinson EOS With Temperature Dependent kij Calculated Through a Group Contribution Method). *Ind. Eng. Chem. Res.*, 47(24), 10041-10052.
- [24] Privat, S., & Jaubert, J. N. (2012). Addition of the sulfhydryl group (SH) to the PPR78 model: Estimation of missing group-interaction parameters for systems containing mercaptans and carbon dioxide or nitrogen or methane from newly published data. *Fluid Phase Equilib.*, 334, 197-203.
- [25] Qian, J. W., Jaubert, J. N., & Privat, R. (2013). Phase equilibria in hydrogen-containing binary systems modeled with the Peng-Robinson equation of state and temperature-dependent binary interaction parameters calculated through a group – contribution method. *J. Supercrit. Fluids*, 75, 58-71.
- [26] Vitu, S., Jaubert, J. N., & Mutelet, F. (2006). Extension of the PPR78 Model Predictive 1978, Peng Robinson EOS With Temperature Dependent kij Calculated Through a Group Contribution Method to Systems Containing Naphtenic Compounds. *Fluid Phase Equilib.*, 243, 9-28.
- [27] Drabo, M. (2011). *Fuel modification, and sensing to improve cold-start and transient performance*. Doctoral dissertation, The University of Alabama in Tuscaloosa.
- [28] Huber, M. L. (2007). NIST Thermophysical Properties of Hydrocarbon Mixtures Database (SUPPERTRAPP), Version 3.2. Retrieved January 2007.
- [29] Cowert, J. S., & Hamilton, L. J. (2006). An Experimental and Modeling-Based Investigation into Post Fuel Injection Vapor Generation. *SAE Technical Paper*, doi: 10.4271/2006-01-0492.

Biographies

MEBOUGNA L. DRABO is an assistant professor of mechanical engineering at Alabama A&M University. His research interests include flash atomization, laser diagnostics in IC engines, IDLTS, alternative fuels, advanced combustion IC engine concepts, IC engine emission reduction

and efficiency improvement. Dr. Drabo may be reached at mebougna.drabo@aamu.edu

MARCUS D. ASHFORD is an associate professor of mechanical engineering at the University of Alabama in Tuscaloosa. His research interests include alternative fuels and hydrogen engine/vehicle applications, advanced combustion IC engine concepts, internal combustion engine control schemes, IC engine emission reduction and efficiency improvement, and combustion and fire science. Dr. Ashford may be reached at marcus.ashford@eng.ua.edu

List of Abbreviations and Symbols

AVUS	Active Vapor Utilization System
$A_{kl} = A_{lk}$	Constant parameter
$B_{kl} = B_{lk}$	Constant parameter
C4	Butane
C5	Isopentane
C6	2-Methylpentane
C8	Meta-Xylene
C9	1, 2, 4-Trimethylbenzene
C10	1, 2-Dimethyl-4-Ethylbenzene
C11	1-Tert Butyl-2-Methylbenzene
C12	n-dodecane
K_i	Equilibrium constant
k_{ij}	Binary interaction parameter
MTBE	Methyl Tertiary Butyl Ether
N_g	Number of different groups defined by the method
v	Molar volume
VLE	Vapor/Liquid Equilibrium
x	Liquid mole fraction
y	Vapor mole fraction
Z	Compressibility factor
Z_i	Mole fraction of component I in the feed
Z_{vi}	Gas phase compressibility factor of i component
Z_l	Liquid phase compressibility factor

SPIRAL SENSING AND PROBABILITY MAP OF IMPACT (PMOI) FOR IMPACT CHARACTERIZATION

Rowshan A. Rima, University of South Carolina; Agbasi Chijioke, Cummins, Inc.;
Paramita Chakraborty, University of South Carolina; Sourav Banerjee, University of South Carolina

Abstract

In this study, the authors developed a method for characterizing and localizing foreign body impacts using a spiral sensing technique followed by the Probability Map of Impact (PMOI). Localized placement of sensors, using the geometric configuration of the Theodorus spiral (TS), was proposed; through active comparisons, the authors report that TS was found to be superior to conventional sensor clusters under velocity-free impact detection. A statistical classifier was also proposed to enumerate PMOI for impact characterization. Impact events were detected by minimizing an error function from all possible combinations of sensor clusters from TS. All possible combinations of impact data were then plugged in to a statistical classifier that utilized Mahalanobis distance to obtain the PMOI. PMOI provides a map of probable impact on the entire structure, which improves the inspection methodologies that are essentially carried out by the maintenance personnel. If a low-velocity impact causes incipient damage in the structure, prior knowledge of the type of impact will reinforce the finding from an active SHM system. This process will lead to reliable verification of an active system and will generate confidence on the combined passive-active systems, without any additional effort.

Introduction

Guided ultrasonic waves are frequently used to detect damage in structures [1-3] by employing piezoelectric sensors. The defects in structures may originate from mechanical and/or thermal fatigue loading during operation [4], [5] or due to foreign body impacts [6-8]. To reliably verify the damage state, prior knowledge of the impact location and the degree of impact are valuable for active-health monitoring systems. Thus, combined passive-active monitoring systems were proposed for this study [9]. Current state-of-the-art passive monitoring systems tend to report only the exact location of impact. Passive monitoring systems are not designed to report the type of impact (e.g., size and/or material properties of the impactors). Thus, the argument can be made that having access to a suitable read-out tool that could reliably estimate the impact events and the impact type would be beneficial for developing a combined passive-active system. It is important to judge the severity of dam-

age by an impactor; thus, it was reasoned, knowledge is needed on the impact type. The proposed PMOI method used in this study would be a suitable tool to use in the pursuit of the proposed goal. Probability of detection (POD) of an active-health monitoring system is well studied in the literature [10-12]; whereas, estimation of the probability of an impact event on a structural component is scarce.

In most cases, accurate and pin-point coordinates of the exact damage location are imperative, while an approximate damage location would suffice and eventually would provide tentative guidelines to the maintenance personnel for further investigations. Hence, in this study, a Probability Map of Impact (PMOI) was proposed. The proposed method would discretize the target structure into small pixels and each pixel would obtain a representative PMOI value after analysis. The contour map of the Probability of Impact (POI) demonstrated the most probable location of an impact. Simultaneously, above a pre-determined threshold value, geometric parameters associated with the PMOI map revealed the information on the impactor. According to the proposed method, PMOI was introduced not only to quantify the accuracy of a passive system but also the accuracy of the estimated impact location. Once the PMOI was obtained from the proposed approach, the region of the structure that manifests $POI > x\%$ can be further investigated with the same set of sensors in the active mode. Here, $x\%$ was the threshold value of the POI and can be problem dependent (e.g., $> 95\%$ or 98%) similar to the probability of detection in the active mode. Hence, unlike current passive systems that provide impact location without any probability of estimation, on the entire structure, PMOI automatically provided the landscape of probability distribution, revealing the impact location and possible size of the impactor, if pre-calibrated.

Generally, a network of piezoelectric (PZT) transducers is used to develop a real-time SHM system using a combined active-passive system [9] or simply a passive system [6], [7], [13-17]. In traditional active-passive mode, few PZTs act as actuators and few act as receivers/sensors. In passive mode, the wave signals received by the sensors are analyzed to find the time of flight (TOF) of the first wave packet. Next, using a triangulation technique, the event location is estimated [15]. But it has been found that the velocity profile in the structure is the fundamental barrier for accu-

rately locating the impact events. Over the last forty years, many algorithms have been proposed for locating acoustic sources. While it was impossible to review all of the primary principles, they were divided into three major techniques: In the first group are hyperbolic positioning algorithms, where the impact events are located by calculating the time differences between different arrival times of the wave packets captured by the sparse array sensors.

The most common approach to finding the impact location is to use the triangulation technique (also known as the Tobias algorithm). In the Tobias algorithm, the impact points are identified from three intersecting circles, whose centers are located at the sensor position. This approach is strongly limited by the assumption that the wave velocity must be known and remains constant in all directions. Due to the impact, flexural wave modes are actuated in the structure. The wave velocities in the structure are not constant along all directions. Velocities of the wave modes are also dispersive; i.e., they depend on the frequency of the propagating waves. However, in passive systems, frequencies are the function of three factors: the size of the object, the non-linear coupling between the object and the plate, and the amount of energy imparted by the object and the actual energy transmitted into the plate.

In the second group is an inverse problem based on a system identification technique by using the transfer function method [18]. Neural network- and genetic algorithm-based approaches were proposed [19], [20] for impact detection. Unfortunately, from a computational and data-storage point of view, such procedures require an extensive number of training data, which makes the method slower. In the third group are the direct strategies, which allow the location of the wave source to be detected by capturing the wave direction of propagation without the prior knowledge on the TOF. This method requires anisotropic transducers capable of revealing the direction of the incoming waves. Hence, transducer rosettes comprised of Macro-Fiber Composite (MFC) transducers were proposed [13], [16] for impact classification.

To locate the impact and also the velocity-free estimation of impact location, some researchers have proposed a dispersion-compensation method for the guided waves generated in the plate. A passive monitoring technique, based on dispersion compensation [17], [21], [22], showed that the velocity profile of the guided wave in a plate was not necessary when a sensor cluster of three sensors was used. It was concluded, nonetheless, that the accurate selection of the time of flight is fundamental and has been a manual process. A rapid and velocity-independent damage-localization approach (for active systems) was proposed [23]. A new con-

cept borrowed from the vibrational analysis was also proposed [24], where an order-analysis algorithm was used to locate the damage in the plate without using the wave velocity profile. In this current study, the authors adopted an error minimization approach and generalized the approach by introducing a PMOI. In addition, in order to avoid multiple issues related to symmetricity of the sensor cluster, a localized sensor cluster in a spiral pattern was proposed.

Previously, impact location was obtained by optimizing an error function formulated by taking the differences between theoretical and experimental sensor signals [21]. However, this proposed objective function had inherent drawbacks with multiple singularities, which were later overcome [22]. The optimization technique was further improved [25] and implemented, where the time differences between two wave packets could be found using peak detection techniques [6], [7] and cross-correlation techniques [26]. The precision of the methods discussed thus far are limited by the dispersion of wave modes and noisy acquisition of sensor signals, which complicate the detection of the exact time of arrival. In this study, the authors proposed a method for automatic and accurate detection of the TOF, based on the cross-correlation technique and short-term Fourier transform algorithm. Only recently were sparse clusters of sensors [27] proposed, where clusters of three closely spaced sensors were distributed on the structure at different locations. Each cluster of three sensors placed at different locations on the structure was capable of bypassing the necessity of the velocity profiles in the structure. Although novel, this method was bound to provide error, due to fewer numbers of voting sensors. Hence, in this study, the authors proposed spirally oriented PZTs for detecting the impact signal.

To the best of the authors' knowledge, none of the passive techniques presented in the literature take advantage of the special orientation of the PZTs. However, a few examples of spiral sensing can be found in the literature that primarily discuss active-health monitoring systems [28], [29]. Localized clusters of multiple combinations of three sensors with variable arm-lengths produce Theodorus Spiral (TS) patterns of sensors at close proximity such that the distant impact event can be monitored with increased reliability. This feature of spiraling was very helpful for the impact point location in both isotropic and anisotropic plates.

Theodorus Spiral (TS) Sensing

Also proposed in this study was a local sensor network, as opposed to a distributed sensor network, which could provide PMOI with quantified reliability for a passive monitoring system. The primary reason for proposing TS was to

improve the number of unbiased voting elements. TS geometry starts with a right-angle triangle at the origin, O , and propagates in a spiral network of contiguous right-angled triangles that are constrained by the origin and its unit projected normal. Figure 1 shows how the projected normal from the origin and the angles between two vertices of the spiral with respect to the origin varies. The angle can be obtained from Equation (1),

$$\theta_n = \sum_{k=1}^{n-1} \arctan\left(\frac{1}{\sqrt{k}}\right), n = 0, 1, 2, \quad (1)$$

and the change in angle viz from Equation (2):

$$\theta_{n+1} - \theta_n = \arctan\left(\frac{1}{\sqrt{n}}\right) \quad (2)$$

The layout of the TS can be written using Equation (3):

$$T(e) = \prod_{k=1}^{\infty} \frac{1 + \frac{i}{\sqrt{k}}}{1 + \frac{i}{\sqrt{k+e-1}}}, e > -1 \quad (3)$$

where, e is the n th of the vertices of the spiral (i.e., the n th sensor) of the TS sensor cluster. This function is called the Theodorus function in honor of Theodorus. The function maps the vertices of the spiral towards the origin as e decreases from ∞ to 0, and away from the origin as e increases.

The importance of TS geometry for passive sensing is the distinct location of its vertices. No two vertices are symmetric in any way, because vertex location is determined by the angle function θ_n in Equation (1) and radius function $T(e)$ in Equation (2). As shown in Figure 2(a), the projected dashed lines along each hypotenuse of the right-angled triangles in TS do not overlap, confirming the distinctness of the vertex positions. The vertices are useful for strategic discrete sensor placement in the design of unbiased voting elements of sensor clusters. Signals obtained from such sensor clusters are mutually exclusive, thereby assuring the efficient use of discrete sensors as well as robust characterization of an event.

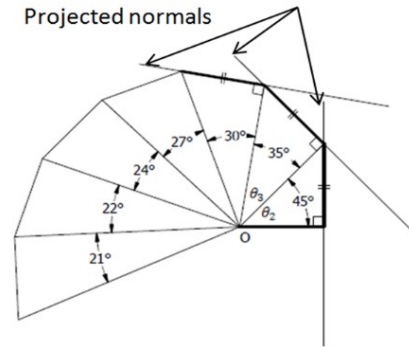
Although mutually exclusive signals can also be acquired by random placement of sensors, TS facilitates an organized and clustered layout which, in turn, helps wiring and connection of the sensors to hardware for practical applications. TS confirmed its usefulness in characterizing impactor's material property [30]. TS is also applicable for beam-

forming techniques that use pitch catch and differential arrays to detect damage location. While comparing TS to conventional sensor clusters with geometries like circles or squares, it is easy to see from Figure 2 (b) that inherent symmetry exists among the discrete sensors in circle or square clusters with respect to an incident wave front. This is the same case for other linear sensor arrays. Opposite mode vectors would have symmetric components defined by Equations (4) and (5):

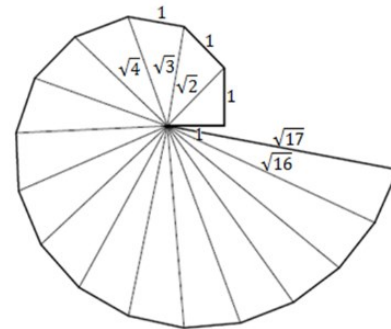
$$k_1 \sin \theta = k_2 \cos \theta \quad (4)$$

$$k_1 \cos \theta = k_2 \sin \theta \quad (5)$$

where, k_1 and k_2 are wave vectors passing through the centers of the top left and top right sensors. And k_n is considered as the n th mode vector that passes through the center line with respect to the cluster.



(a) Projected Tangents and Changing Angle θ

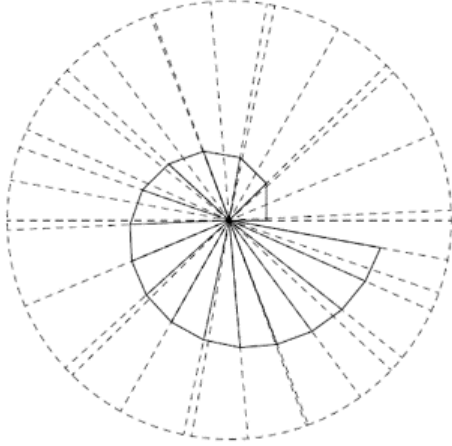


(b) The Spiral of Theodorus Displaying the "Construction of Anderhub"

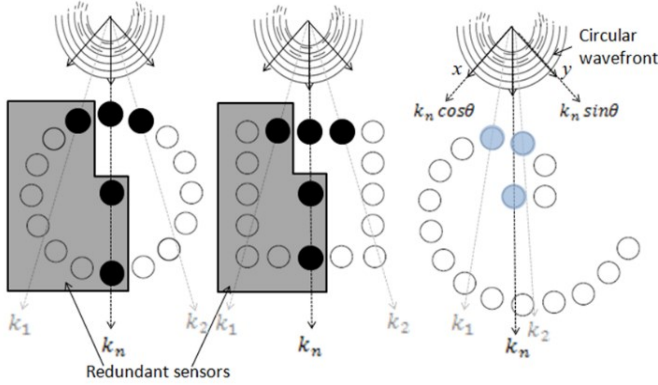
Figure 1. Constructing a Theodorus Spiral (TS)

The mode vectors displayed in Figure 2(b) pass through the center of the sensors marked in black. The chirality of the TS has no symmetry in sensing a particular mode. For wave vector k_n , a circle or a square cluster will have three

sensors collinear with this same wave vector, thereby making two of the three sensors redundant under any circumstances. However, the TS sensor cluster has only one sensor collinear with the wave vector. The TS can overcome the situation of redundancy because all of the discrete sensor positions are non-symmetric and unbiased.



(a) Projected Non-Overlapping Hypotenuse Lines



(b) Redundancy due to Symmetry and Inline Sensors of a Circle and Square Sensor Clusters Compared with Nonsymmetry and Nonoverlapping Hypotenuse Networks of the TS

Figure 2. The Distinctness in the Vertices of the TS

Equations (4) and (5) show how two symmetric mode vectors in a circular (CSC) or square sensor cluster (SSC) can be decomposed into their orthogonal components to confirm that they are mere images of one another. This means that symmetric sensors will be unable to capture mutually exclusive data sets. For CSC and SSC, more than half of the sensors acquire biased information. Such inefficiency can be overlooked in active-health monitoring where beam forming, pulse-echoes are implemented for damage inspection. However, in passive monitoring, signals from an

acoustic emission, such as an arbitrary impact event, could travel from any direction with any dominant mode and also be mixed with noise or other impact events. It becomes necessary to implement a device that can sense a range of wave modes through unbiased voting. This, in return, makes the solution cheaper because some of the hardware channels that are dedicated in circle and square geometries are found to be redundant. However, the TS makes all sensors valuable by using the same number of sensors in a cluster.

In this study, the authors verified their proposed method of using an isotropic Aluminum 5xxx plate. Eighteen passive sensors were arranged in a TS pattern on an Aluminum 5083 plate. To accurately capture the time of arrival, a cross-correlation technique was used. Multiple three-sensor combinations were employed and an optimization technique via minimization of error between the theoretical and the experimental time of arrival was performed to find the impact location. It was found that the impact points obtained from multiple combinations of sensors were quite diverse and a single impact point could not be obtained. Therefore, a statistical analysis was performed in order to find the probable direction of the impact point. The most significant feature of this approach was that it showed a directional range of probable locations of the impact. This feature will be very helpful for the critical area of the structure, where placement of sensors might not be possible.

Experimental Setup

The experiment was conducted on an isotropic aluminum plate, 609 mm $\times$ 609 mm. Eighteen piezo-electric sensors were adhesively bonded with the surface in the TS orientation, as shown in Figure. 3.

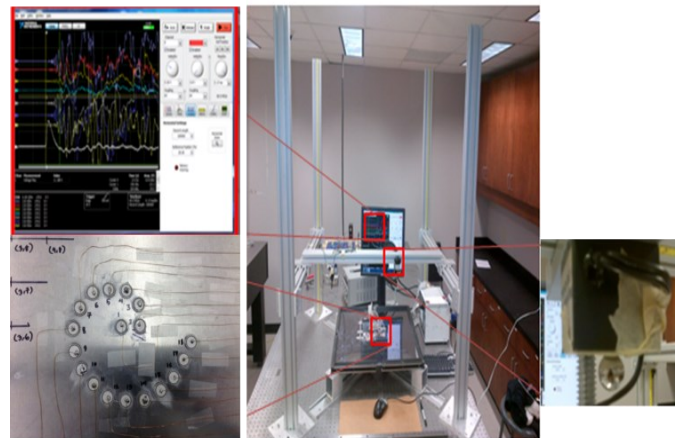


Figure 3. Experimental Setup and Piezoelectric Sensors Mounted on an Aluminum Plate in the TS Pattern

Multiple objects were used to carry out the experiments, as shown in Table 1. The diameters of the objects were 12.7 mm. Additionally, a steel ball (~185 GPa) with a diameter of 5 mm < 12.7 mm and a golf ball (~0.3 GPa) with a diameter of 43 mm > 12.7 mm were used. These balls were dropped on the plate from different heights to keep the impact energy constant. The ball drop height was 18.5 inches for the steel ball and 11 inches for the golf ball.

Table 1. Different Impact Objects

Material	E. Modulus (GPa)	Density (Kg/m ³)	Poisson's ratio
Stainless Steel	200	7750	0.29
Titanium	103	4521	0.37
Aluminum 2024-T3	72	2780	0.33
Marble	30	2685	0.23
Polytetrafluoroethylene (PTFE) aka Teflon	0.5	2214	0.46

The arrangement (see Figure 3) was made such that it can control the ball drop height accurately. To avoid multiple bounces of the ball during its vertical line of drop, the plate was kept at a slightly inclined angle. In this paper, the authors report the results of the PMOI from the 5 mm diameter steel ball and the 43 mm diameter golf ball, whereas the classification of impact was performed using all of the objects used during the experiments. Sensor 1 was located at the center of the plate (152.4mm x 152.4 mm in the redefined coordinate system, where 0,0 was located at 152.4mm x 152.4 mm away from the bottom left corner of the plate). Repeated experiments were carried out to verify the consistency of the signal and achieve reliability of the sensor/hardware system. Multiple locations were tested with impacts. Impact experiments, with each impact object at each impact location, was repeated 10 times (see Figure 3).

Due to an abundance of data, an analysis was performed on the average signals obtained from two different impact locations—location A (254.0 mm x 254.0 mm) and location B (50.8 mm x 254.0 mm), using 5 mm steel balls and 43 mm golf balls. Sensors 1-18 (denoted as S1-S18) represent the locations of the piezoelectric sensors. The signals at the sensor points were received using an NI PXI-5105, 3 8-Ch (3x8<18), 60 MS/sec digitizer. Simultaneously, in order to re-verify the signals (see Figures 4 and 5) a Tektronix TDS 2004C oscilloscope was used. Sensors were connected with the digitizer using an SMB-type cable at the digitizer end and an SSMA-type cable at the sensor end. A splitter (SSMA cable splitter and then SSMA to BNC) setup was used to split sensor signals in order to send to the digitizer and the oscilloscope simultaneously. This arrangement was used for

four sensors at a time and all of the sensor signals were verified.

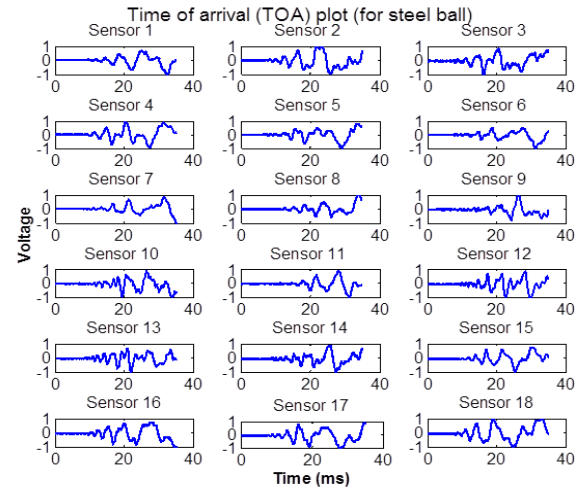


Figure 4. Passive Signals from all of the Sensors Organized in the TS when a Steel Ball (shown in Figure 2) Impacted the Plate at a Distance of 5.6 in. from the Center of the Spiral

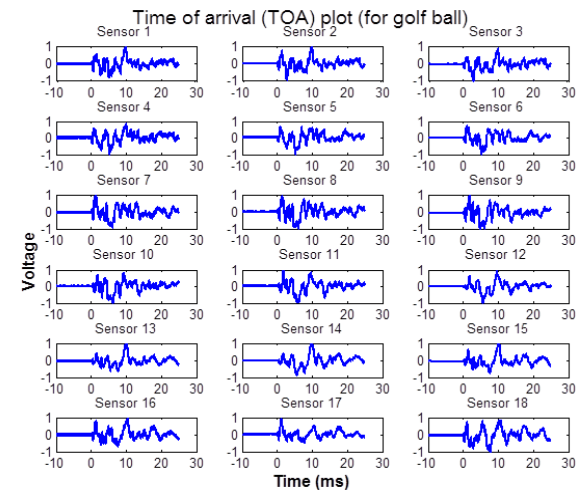


Figure 5. Passive Signals from all of the Sensors Organized in the TS when a Golf Ball (shown in Figure 2) Impacted the Plate at a Distance of 5.6 in. from the Center of the Spiral

From the sensor signals, the time of arrival (TOA / TOF) could be obtained through signal processing. It was quite difficult to measure the exact TOA from the time history signals because there were ambiguities in selecting exact TOAs, as arrival of the first symmetric mode (S_0) was hidden inside the noise. Therefore, the exact TOA of the S_0 mode could not be determined. However, the TOA of the relatively stronger flexural mode, A_0 , could be obtained from the plots by employing the cross-correlation technique.

First, arrival of the ultrasonic energy was recorded using the threshold from the first noticeable peak greater than the noise level. The dominant frequency in the signals received during first 150 μsec was obtained at different time windows using the Short-Term Fourier Transform (STFT) technique. Next, composite tone burst signals made of dominant frequencies were created containing three cycles for each frequency at a sequential time step (time windows) used in the STFT.

A cross-correlated signal was then obtained by cross correlating original experimental signals and the composite tone burst signals. First, predominant peaks observed from the cross-correlated signals were selected and designated as the TOA. Using the threshold technique, the TOAs were also obtained and it was found that they were quite comparable within $\sim 5\%$ error. The automated cross-correlation approach is depicted in Figure 6.

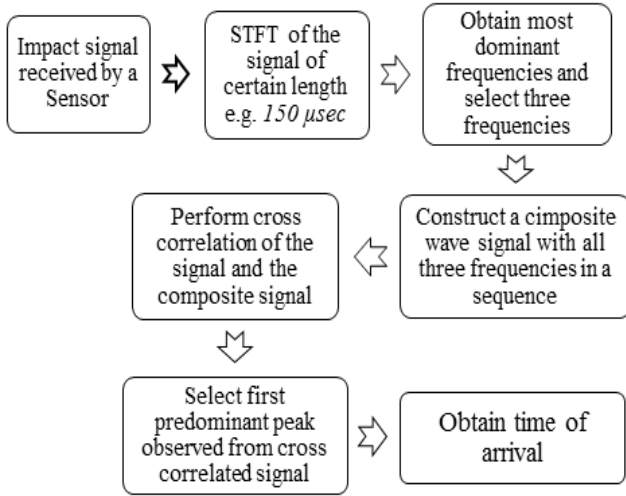


Figure 6. Process for Describing the Selection of Appropriate TOAs from the Impact Signal

In a complicated structure, threshold techniques were bound to give error and, thus, it was recommended to use the cross-correlation approach. Further, to improve the accuracy of impact detection, a PMOI was introduced.

Objective Error Function

Sensors arranged in the TS were divided into multiple combinations, keeping at least three sensors. All possible combinations of three sensors could produce multiple error functions and were minimized. The structure was discretized into small pixels with each pixel point receiving a vote from all of the combinations. These votes were used to calculate the PMOI over the entire structure. Let us first focus

on one combination of three sensors out of $18P_3$ combinations. The time of arrival of the signal at the i^{th} sensor can be, viz, t_{it} . If the time of impact in Equation (6) is t_{imp} then the propagation time for the wave starting from the impact point to the sensor is given as:

$$t_i = t_{it} - t_{imp} \quad (6)$$

To define both t_{it} and t_{imp} , the same time of reference was used. If the coordinate of the sensor is (x_i, y_i) and the coordinate of the impact point is (x_0, y_0) then the scalar distance of the sensors from the point of impact can be written as shown by Equation (7):

$$d_i = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2} \quad (7)$$

The time of travel of the wave to the sensor locations is denoted as t_i ; and, if the velocity is c along the direction of travel, Equation (7) can be modified and written as shown in Equation (8):

$$d_i = ct_i = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2} \quad (8)$$

In structural health-monitoring systems, the sensor location (x_i, y_i) is known, but the impact point location (x_0, y_0) is unknown. If the assumed coordinate values (x_0, y_0) are different from the impact point, an error function $E(x_0, y_0)$ with a positive value may appear. If the assumption is correct then the error function will result in a value of zero. Since the exact time of arrival is often hidden in the noise, the values obtained from the error function will have relatively high error. The relative time of arrival was used, as shown in Equation (9), instead of the actual time of arrival [22, 27]:

$$t_{ij} = t_{it} - t_{jt} = (t_i + t_{imp}) - (t_j + t_{imp}) = t_i - t_j \quad (9)$$

Here, i and j take on values for number of sensors. Equation (10), then, defines the relative TOA between the i^{th} and j^{th} sensors and their ratios.

$$\begin{aligned} t_{ij} &= \frac{d_i}{c} - \frac{d_j}{c} = \frac{d_i - d_j}{c} \\ \Rightarrow t_{ij} &= \frac{1}{c} \left\{ (x_i - x_0)^2 + (y_i - y_0)^2 - (x_j - x_0)^2 + (x_j - x_0)^2 \right\} \end{aligned} \quad (10)$$

In a three-sensor system, equal importance is given to all TOA measurements. The error function is defined after few steps of mathematical simplification, as defined by Equation (11). According to the formulation, the processes to estimate the location of the impact are not influenced by the experimental error in any particular set of measurement.

$$E(x_0, y_0) = \left(\frac{\left\{ \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2} - \sqrt{(x_2 - x_0)^2 + (y_2 - y_0)^2} \right\}}{\left\{ \sqrt{(x_2 - x_0)^2 + (y_2 - y_0)^2} - \sqrt{(x_3 - x_0)^2 + (y_3 - y_0)^2} \right\}} - \frac{t_{12}}{t_{23}} \right)^2 + \left(\frac{\left\{ \sqrt{(x_2 - x_0)^2 + (y_2 - y_0)^2} - \sqrt{(x_3 - x_0)^2 + (y_3 - y_0)^2} \right\}}{\left\{ \sqrt{(x_3 - x_0)^2 + (y_3 - y_0)^2} - \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2} \right\}} - \frac{t_{23}}{t_{31}} \right)^2 + \left(\frac{\left\{ \sqrt{(x_3 - x_0)^2 + (y_3 - y_0)^2} - \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2} \right\}}{\left\{ \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2} - \sqrt{(x_2 - x_0)^2 + (y_2 - y_0)^2} \right\}} - \frac{t_{31}}{t_{12}} \right)^2 \quad (11)$$

For a few unknown values of (x_0, y_0) , the denominator

$$\left(\sqrt{(x_j - x_0)^2 + (y_j - y_0)^2} - \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2} \right)$$

[subscripts i and j take values 1, 2, and 3] can vanish. For those special values of (x_0, y_0) , the objective function becomes infinity. So, a special process was adopted during the computation to avoid this singularity. The problem was avoided by changing the expression of the error function from Equation (12), to that described in Equation (12) [22]:

$$E(x_0, y_0) = \{t_{23}(d_1 - d_2) - t_{12}(d_2 - d_3)\}^2 + \{t_{31}(d_2 - d_3) - t_{23}(d_3 - d_1)\}^2 + \{t_{12}(d_3 - d_1) - t_{31}(d_1 - d_2)\}^2 \quad (12)$$

Minimization of Equation (12) yields the pixel points where the error was minimum and, therefore, the location of the impact point was identified. It was expected that all combinations would provide a very close estimation of the points of impacts; however, a sparse distribution of impact points was observed from all the combinations in the TS sensor cluster. Sparseness of data was found to be impact-type dependent, and it was consistently found that the lower frequency (large impact object) impacts had wider sparseness in data than the higher frequency (smaller impact object) impacts.

Probability Map of Impact

To measure a particular impact point, different sensor combinations were used. Using the methods described earlier, the estimated true impact point from each combination could be obtained. Let the estimated impact point from the i^{th} sensor combination be $(x(i), y(i))$, which can be derived

using Equation (12). Using the data from the varying sensor combinations, a probability model was constructed which could provide the most probable region of the true impact and the equally probable regions of the impact.

The sensor combination estimated from the same true impact showed a distinctive pattern. The data had a high concentration near a central coordinate and a gradually decreasing frequency away from that central location. This pattern was expected because for each experiment different sensor combinations are supposed to measure the true impact point. Naturally, most of the estimated points would cluster around the true impact point that is being measured with fairly equal amounts of over and under estimations due to natural randomness (sampling error), which is inherent in any experimental measurement. This distinct pattern is a classic characteristic of a Gaussian probability distribution, which is most commonly used to model sampling error. Therefore, the authors proposed to fit a Gaussian probability distribution to the sensor data with the assumption that the center of the distribution would represent the true impact point.

A bivariate Gaussian distribution was fitted to the impact point data $(x(i), y(i))$ obtained from multiple sensor combinations. The parameters of the fitted distribution were estimated from the sensor data. The mean coordinate provided a point estimate of the true impact point. The scatter plot of the sensor data in this study showed a significant level of noise, hence the robust estimation approach for model finding was engaged.

As explained, the variations in the sensor data that were collected due to the impact at the true impact location were due to the sampling errors. The most common probability distribution for modeling such sampling errors is the Gaussian distribution, hence a bivariate Gaussian distribution to model the sensor data coordinates $(x(i), y(i))$ was used. A bivariate normal distribution was determined by the following parameters: the center (μ_x, μ_y) , the spread or scatter (standard deviations) (σ_x, σ_y) , and the correlation between the two variables, ρ . Specifically, if random variables (X, Y) follow a bivariate Gaussian distribution then the joint probability distribution is given by the density function, f viz, as shown in Equation (13):

For and $-\infty < a < b < \infty$ and $-\infty < c < d < \infty$

$$P(a < X < b, c < Y < d) = \int \int_{ac}^{bd} f(x, y) dx dy \quad (13)$$

where, the bivariate Gaussian distribution (μ_x, μ_y) can be written as shown in Equation (14):

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \cdot \exp \left[-\frac{1}{2(1-\rho^2)} \left\{ \left(\frac{x-\mu_x}{\sigma_x} \right)^2 + \left(\frac{y-\mu_y}{\sigma_y} \right)^2 - 2\rho \left(\frac{x-\mu_x}{\sigma_x} \right) \left(\frac{y-\mu_y}{\sigma_y} \right) \right\} \right] \quad (14)$$

For this current study, the center of the bivariate Gaussian distribution (μ_x, μ_y) , from Equation (14) represented the true impact point. Precisely, it was assumed that the observed impact point coordinates (x, y) followed a bivariate Gaussian distribution with the center or location parameter (μ_x, μ_y) . In practice, the center will be unknown and can be estimated from the observed data. In fact, to find an appropriate bivariate distribution, it is required to estimate all of the parameters $(\mu_x, \mu_y, \sigma_x, \sigma_y, \rho)$ from the TS sensor data. The classical way of acquiring this is to use the mean-square estimation method that minimizes an error function consisting of the mean squared distance of the sensor coordinates from the mean coordinates (comparable to the objective error function described earlier. This method, however, is heavily influenced by the outliers.

Considering the variables (x, y) with n observations, where the impact points obtained from each combination of three sensors from the TS data points, the result is (Equation (15):

$$X = \begin{bmatrix} x(1) & y(1) \\ \vdots & \vdots \\ x(n) & y(n) \end{bmatrix} \quad (15)$$

where, $x(k)$ is the k^{th} observation on x and $y(k)$ is the k^{th} observation on y . The classical mean-square estimation estimates the center or the location viz, as defined by Equations (16a) and (16b):

$T(X) = \{\text{arithmetic means},$

$$\bar{x} = \frac{1}{n} \sum_{k=1}^n x(k) \quad (16a)$$

$$\bar{y} = \frac{1}{n} \sum_{k=1}^n y(k) \quad (16b)$$

And, classical mean-square estimation estimates the scattering in the data the sample covariance matrix given by Equation (17):

$$\hat{\sigma}_{xy} = \frac{1}{n-1} \sum_{k=1}^n [x(k) - \bar{x}][y(k) - \bar{y}] \quad (17)$$

Both measures—Equations (16) and (17)—are significantly affected by the outliers and, hence, for data with high noise, the classical method does not provide appropriate measures of either location or scatter (standard deviations).

Some robust estimation methods with high break points were incorporated for such special cases. The scatter plot of the sensor data in this current study showed some outliers and a significant amount of noise. Hence, the classical mean-square estimation method may not have been the best approach for parameter estimation for the model. Thus, as follows, a robust estimation approach for PMOI modeling was used to incorporate the noise into the model. In order to determine the “good” points and provide the final estimation of the location and the scatter from the filtered data, further minimum volume estimation [31], [32] and the minimum covariance determinants methods [33] were used along with the classical methods.

In the minimum volume estimation method, a subsets of h points, which could be taken to be equal to $[(n+p+1)/2]$ for a p -variate Gaussian distribution, were chosen and the volume of the ellipsoid which contained the aforementioned points were considered. The subset with the minimum volume was selected, the location $T(X)$ was estimated by the center of the ellipsoid, and the scatter (standard deviations) were measured by the covariance matrix of those h points (adjusted with a constant scale factor in order to achieve the consistency in the multivariate normal distribution). Similarly, in the minimum covariance determinants method, a subset of h data points was selected and the sample covariance matrix was calculated for each subset. The subset of the covariance matrix with a minimum determinant was selected and the arithmetic means of the points in the selected subset provided the location estimator $T(X)$. The sample covariance matrix (with adjusted scale) of the subset provided the estimation of the scatter.

The fit procedure [34] inherently employs the classical minimum volume estimation and minimum covariance determinants methods to get estimated $T(X)$ and $C(X)$. Next, the Mahalanobis distance for each observed data vector from the estimated location was computed. The points with Mahalanobis distances less than the 97.5th percentile of a chi-square distribution with p degrees of freedom (for a p -variate Gaussian distribution) were considered to be the “good” point [34]. The chi-square distribution is justified by the underlying multivariate normal distribution since the sum of the squares of the independent normal variables follow a chi-square distribution. The final estimated measure of the location and the scatter was obtained from these good points and was used in the actual fitting process.

The actual model-finding process, on top of the core methods explained above, was improved and adjusted. A finite sample correction [35] was incorporated for the first estimate. A fast-estimate algorithm based on minimum covariance determinants was used for initial estimation [32]. Further, a weighted estimation step was employed for the final covariance estimate. In this step, the covariance estimate was weighted and the weights were based on the raw minimum volume estimation estimates (described above) and the raw minimum covariance determinant estimates [36], which eventually resulted in greater efficiency.

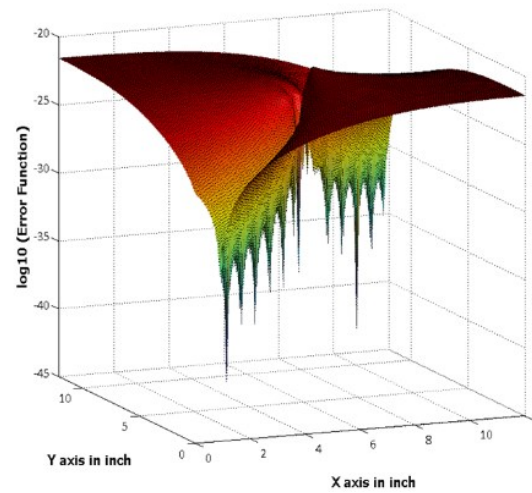
The objective of the PMOI (probability map of impact) was to estimate the probability of an impact on the entire structure. Hypothetically, each pixel point (as divided) on the structure had a possible chance of impact. The sensor data should show a higher probability of impact near the actual impact site and a lower probability away from the impact site. A 3D probability distribution over the entire structure was useful in visualizing the highest probability zones with grades. Further, a contour plot of the projection of the probability distribution revealed the equal probability zones. For example, if one is interested in investigating a structural region with 98% probability of impact, the contour plot can be used to identify the 98% equal probability area that can be constructed from the sensor data. The results from the PMOI analysis are presented in Figures 8-10. The point estimates of the true impact points from the fitted distributions are shown in Table 2.

Table 2. Impact Identification

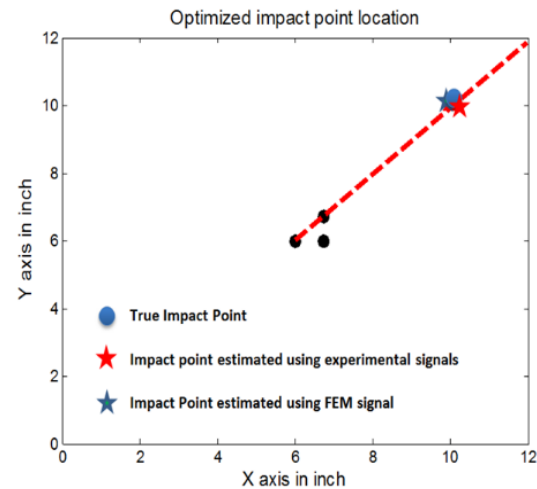
Impact Material	True Impact Point	Estimated Impact Point from Sensor Data (center of the equal probability ellipses: $T(X)$)
Steel Ball	50.8mm×254 mm	59.2 mm × 250.16 mm
Golf Ball	50.8mm×254 mm	54.90mm × 265.45mm
Steel Ball	254 mm × 254 mm	252.39mm×253.35mm
Golf Ball	254 mm × 254 mm	253.6mm × 254.37mm

Results and Discussions

A typical objective function—using Equation (12) for the 1st set of sensors, combination 1, i.e., sensor numbers 1, 2, and 3—is plotted in Figure 7(a). The corresponding impact location is shown in Figure 7(b) and was obtained from the experiment and the simulation, though this is not discussed in this paper, as it is not the objective of this study. It can be seen that the impact locations was in very close agreement and, thus, the method was validated.



(a) Objective Function or Error Function Variation for a Three-Sensor Set Plotted in Log Scale



(b) Impact Point Identification using Three Sensors

Figure 7. Objective Function and Impact Point Identification

The minimum value of the error function can be obtained by virtually meshing the entire plate. The coordinate of the pixel point was selected at the global minimum of the error function and was identified as a probable impact point. Even after rigorous selection of the TOA and compensation for an erroneous TOA, estimation of the impact point could be away from the true impact points if the voting were taken from a single cluster. However, according to the proposed technique, such a dilemma can be avoided. Next, the statistical analysis was performed. All figures presented here were 304.5 mm x 304.5 mm square at the middle of the plate, considering that sensor 1 was at 152.4 mm x 152.4 mm.

The proposed statistical analysis was performed on the data set obtained from the impact localization by virtue of minimizing the error function. Figure 8 shows the typical PMOIs that depict the impact on the plate with a steel ball and a golf ball at locations A and B, respectively. The figures in the columns are presented in order to characterize the impact type, and the figures in the rows are presented to classify the impact location. Projection of 3D plots of the PMOI at equal values of probability produced equal probability lines. The projections of the 3D plots on an x-y plane were elliptical. Each ellipse represents an equal probability region. Hence, a region of the plate was exposed to a certain probability of impact. The center most ellipse is the most probable region of the impact event. In Figure 8, the gray ball signifies the actual impact point and the black star signifies the estimated impact point.

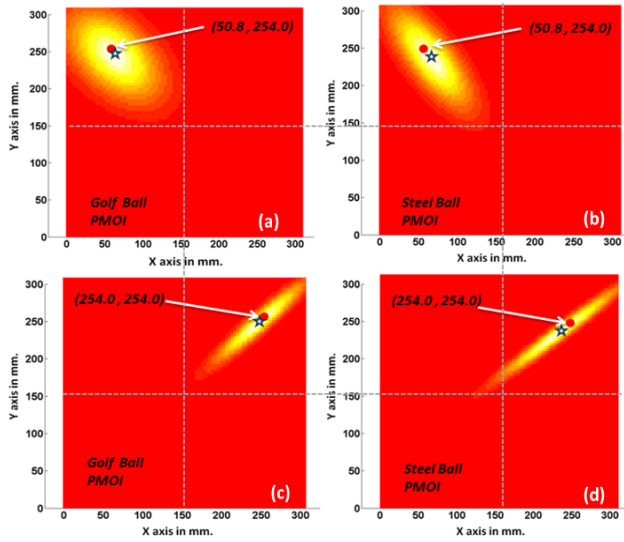


Figure 8. Contour Projection Calculation of the PMOI Map

Figure 9 shows the equal probability lines. It can be seen that the true impact point for all cases falls within the 95% probability lines. From Figures 8 and 9, it is immediately apparent that the major and minor axes of the ellipses representing the equal probability lines are impact-type dependent. Irrespective of impact type, the direction of the incoming wave with respect to the center of the spiral sensors is inherent to the plot; this was achieved only due to the spiral configuration of the sensors. Hence, using a single plot, many comments can be made—e.g., the direction of impact, most probable impact location, a region with high probability of impact, and the type of impact with a higher or lower frequency.

High-frequency waves cause narrower distributions and supposedly have an impact from a smaller object with smaller contact area. Low-frequency waves cause wider

distributions and supposedly have an impact from a larger object with larger contact area. A calibration curve can be obtained from the FEM model, the experiment, and the statistical analyses in order to characterize the impact type.

Figure 8 shows that the center-most ellipse (lightest region) is the most probable region of the true impact point, whereas the darker color represents a decreasing probability of the true impact point. Figure 9 shows the projections of 3D PMOIs on the x-y plane, where the gray ball signifies the actual impact point.

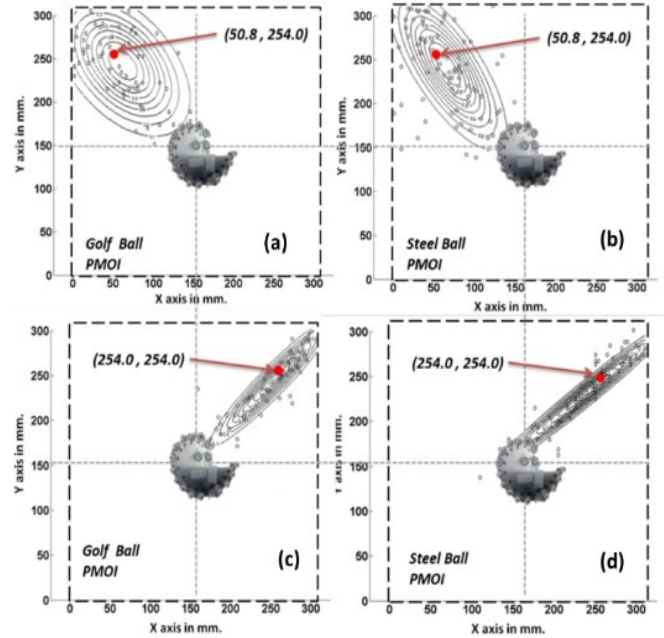


Figure 9. Projection of the PMOI Maps in at Every 10% Change in Probability Calculated

From Figures 8 and 9, it can be seen that this special orientation can locate the direction of the impact point, as most of the impact point locations can be viewed in a singular direction. Furthermore, the probability density was significantly higher near the actual impact point. If the TS dimension is much smaller than the impact distance from the center of the TS, using the proposed technique, a unique direction of impact can be predicted without having a velocity profile. Based on the PMOI map and the dominant frequencies captured (the third process in Figure 6) from individual sensor signals in a cluster, the objects mentioned in Table 1 were successfully classified. The object type was categorized based on the elastic modulus and the most responsive sensors in the cluster were identified for any particular object shown in Figure 10. By virtue of the TS sensor cluster (TSSC in Figure 10), both impact localization and classification can now be achieved.

Conclusion

The special orientation of the passive sensors improvised the directional and frequency sensitivity, thus minimizing the prior knowledge of velocity profile of the structure. As the proposed process required low computational work using real-time information, the algorithm proposed can be especially useful for the critical components of a structure, where positioning of scattered passive sensors might not be possible. Experiments were carried out with five different impacting objects of the same size (diameter) and two additional objects with smaller (a steel ball) and larger diameters (a golf ball). It was shown that a directional range of probable impact area can be predicted without considering the velocity profile. This is the primary validation of this approach and further work will be conducted using an anisotropic plate with this orientation. The five objects with different elastic moduli were classified using the proposed method.

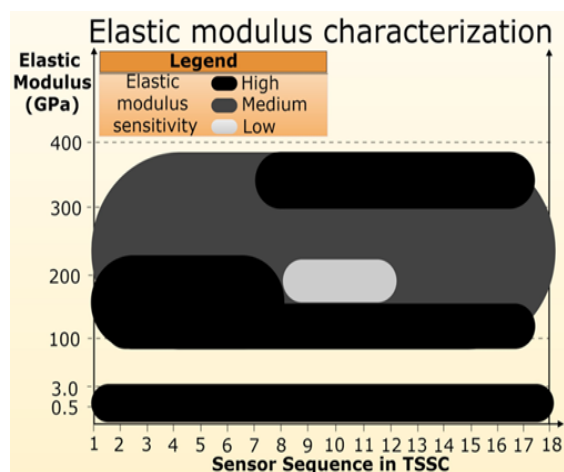


Figure 10. Characterization of Elastic Modulus of Different Impact Objects Based on the TSSC Data

Acknowledgement

The authors would like to thank the Office of the Vice President of Research, at the University of South Carolina, for partially supporting this study. The authors would also like to thank the Skolkovo Technical Institute, Moscow, Russia, for providing partial support for this study. Agbasi Chijioke would like to thank the undergraduate best student award committee at the University of South Carolina.

References

- [1] Giurgiutiu, V. (2008), *Structural Health Monitoring with Piezoelectric Wafer Active Sensors*. Elsevier

Academic Press, New York, USA.

- [2] Giurgiutiu, V., (2011), Piezoelectric Wafer Active Sensors for Structural Health Monitoring of Composite Structures Using Tuned Guided Waves. *Journal of Engineering Materials and Technology*, 133(4), 041012.1-6.
- [3] *Proceedings of the fifth European Workshop on Structural Health Monitoring*. 2010. Sorrento, Naples, Italy: DEStech Publications, Lancaster, Pa.
- [4] Banerjee, S., Qing, X. P., Beard, S., Chang, F. K. (2009). Prediction of Progressive Damage States at The Hot Spots Using Statistical Estimation. *Journal of Intelligent Materials Systems and Structures*, 21(6), 595-605.
- [5] Jata, K. V., Kundu, T., & Parthasarathy, T. A. (2010), An Introduction to Failure Mechanisms and Ultrasonic Inspection. In T. Kundu (Eds) *Advanced Ultrasonic Methods for Material and Structure Inspection* (pp 1-42). ISTE, London, UK.
- [6] Tracy, M., & Chang, F. K. (1998), Identifying Impacts in Composite Plates with Piezoelectric Strain Sensors, *Journal of Intelligent Material Systems and Structures*, 9(11), 920-928.
- [7] Seydel, R., & Chang, F. K. (2001). Impact Identification of Stiffened Composite Panels: I. System Development. *Smart Materials and Structures*, 10(2), 354.
- [8] Mal, A. K., Shih, F., & Banerjee, S. (2003), Acoustic Emission Waveforms in Composite Laminates Under Low Velocity Impact. In T. Kundu (Eds) *Proceedings of the SPIE Health Monitoring and Smart Nondestructive Evaluation of Structural and Biological Systems*, SPIE.
- [9] Qing, P. X., Wilson, H., Baines, A. G., Bordick, N., Banerjee, S., Chung, H., et.al. (2011). Hot-Spot Fatigue and Impact Damage Detection on a Helicopter Tailboom. *International Workshop on Structural Health Monitoring*, (pp 1907-1914). DEStech Publication: Stanford, California.
- [10] Cobb, A. C., Michaels, J. E., & Michaels, T. E. (2009). Ultrasonic Structural Health Monitoring: A Probability of Detection Case Study. *AIP Conference Proceedings*, 1096(1), 1800-1807.
- [11] MIL-HDBK-1823A. (2009), *Department of Defense Handbook: Nondestructive Evaluation System Reliability Assessment*.
- [12] Muller, I., Janapati, V., Banerjee, S., Lonkar, K., Roy, S., & Chang, F. K. (2011). On the Performance Quantification of Active Sensing SHM Systems Using Model-Assisted POD Methods, *8th International Workshop of Structural Health Monitoring*. DEStech Publication, Stanford University.
- [13] Salamone, S., Bartoli, I., Di Leo, P., Lanza Di Scalea,

- F., Ajovalasit, A., et. al. (2010). High-Velocity Impact Location On Aircraft Panels Using Macro-Fiber Composite Piezoelectric Rosettes. *Journal of Intelligent Material Systems and Structures*, 21(9), 887-896.
- [14] Gaul, L., & Hurlebaus, S. (1998). Identification of The Impact Location On a Plate Using Wavelets. *Mechanical Systems and Signal Processing*, 12(6), 783-795.
- [15] Coverley, P. T., & Staszewski, W. J. (2003). Impact Damage Location in Composite Structures Using Optimized Sensor Triangulation Procedure. *Smart Materials and Structures*, 12(5), 795.
- [16] Matt, H. M., & Lanza Di Scalea, F. (2007). Macro-Fiber Composite Piezoelectric Rosettes for Acoustic Source Location in Complex Structures. *Smart Materials and Structures*, 16(4), 1489.
- [17] Marchi, L. D., Marzani, A., Speciale, N., & Viola, E. (2011). A Passive Monitoring Technique Based On Dispersion Compensation to Locate Impacts in Plate-Like Structures. *Smart Materials and Structures*, 20(3), 035021.
- [18] Park, J., Ha, S., & Chang, F. K. (2009). Monitoring Impact Events Using a System-Identification Method. *AIAA Journal*, 47(9), 2011-2021.
- [19] Staszewski, W. J., Boller, C., & Tomlinson, G. R. (2004). *Health Monitoring and Aerospace Structures*. New York Wiley.
- [20] Staszewski, W. J., Worden, K., Wardle, R., & Tomlinson, G. R. (2000). Fail-Safe Sensor Distributions for Impact Detection in Composite Materials. *Smart Materials and Structures*, 9(3), 298.
- [21] Kundu, T., Das, S., & Jata, K. V. (2007). Point of Impact Prediction in Isotropic and Anisotropic Plates from the Acoustic Emission Data. *The Journal of the Acoustical Society of America*, 122(4), 2057-66.
- [22] Kundu, T., Das, S., & Jata, K. V. (2009). Detection of the Point of Impact on a Stiffened Plate by the Acoustic Emission Technique. *Smart Materials and Structures*, 18(3), 035006.
- [23] Moll, J., Heftrich, C., & Fritzen, C. P. (2012). A Rapid and Velocity-Independent Damage Localization Approach for Ultrasonic Structural Health Monitoring. *IEEE Trans Ultrason Ferroelectr Freq Control*, 59(6), 1309-1315.
- [24] Liu, X., Veidt, M., Bo, L., & Peng, C. (2013). Order Analysis for Measurement of Lamb Wave Propagation Distance. *Measurement Science and Technology*, 24(3), 10.
- [25] Hajzargerbashi, T., Kundu, T., & Bland, S. (2011). An Improved Algorithm for Detecting Point of Impact in Isotropic Inhomogeneous Plates. *Ultrasonics*, 51(3), 317-24.
- [26] Kosel, T., Grabec, I., & Mužič, P. (2000). Location of Acoustic Emission Sources Generated by Air Flow. *Ultrasonics*, 38(1-8), 824-826.
- [27] Kundu, T., Nakatani, H., & Takeda, N. (2012). Acoustic Source Localization in Anisotropic Plates. *Ultrasonics*, 52(6), 740-6.
- [28] Yoo, B., Purekar, A., & Pines, D. J. (2011). Piezoceramic-Based 2D Spiral Array and Multiple Actuators for Structural Health Monitoring: Thin Isotropic Panel with Straight Boundaries. *Journal of Intelligent Material Systems and Structures*.
- [29] Baravelli, E., Senesi, M., Ruzzene, M., & De Marchi, L. Low Velocity of Impact and Impactor Using Spiral Sensing Technique. Unpublished manuscript, University of South Carolina, Columbia, South Carolina, USA.
- [31] Rousseeuw, P. J., & Van Zomeren, B. C. (1990). Unmasking Multivariate Outliers and Leverage Points (with discussion). *Journal of the American Statistical Association*, 85, 633-651.
- [32] Rousseeuw, P. J., & Driessen, K. V. (1999). A Fast Algorithm for the Minimum Covariance Determinant Estimator. *Technometrics*, 41(3), 212-223.
- [33] Rousseeuw, P. J., & Leroy, A. M. (1987). *Robust Regression and Outlier Detection*. Wiley.
- [34] Hahne, F. (2012). "prada" R-package. Reference for "prada" package in R: <http://www.bioconductor.org/packages//2.10/bioc/manuals/prada/man/prada.pdf>.
- [35] Marazzi, A. (1993). *Algorithms, Routines and S Functions for Robust Statistics*. Belmont, CA: Wadsworth and Brooks/Cole
- [36] Pison, G. S., Van Aelst, & Willems, G. (2002). Small Sample Corrections for LTS and MCD, *Metrika*, 55(1-2), 111-123.

Biographies

ROWSHAN A. RIMA received her M.S. degree in mechanical engineering from the University of South Carolina in 2013. She earned her B.S. degree (2009) from Bangladesh University of Engineering and Technology, Bangladesh. Her areas of research are related to structural health monitoring and computational modeling of ultrasonic waves. Rima may be reached at [rima@email.sc.edu](mailto:rима@email.sc.edu)

CHIJOKE AGBASI is currently working at Cummins Inc. as an applied mechanics engineer. He received his M.S. degree (2014) from the University of South Carolina, Columbia, B.S. degree (2011) from the University of South Carolina, Aiken. His research interests are in signal processing. Agbasi may be reached at agbasi@email.sc.edu

PARAMITA CHAKRABORTY is an assistant professor of statistics at the University of South Carolina, Columbia. She earned her B.Sc. degree (2002) from St. Xavier's College, Kolkata, India, with a major in statistics. She received her master's degree (M.Stat.) from the Indian Statistical Institute (2004), specializing in mathematical statistics. She received her Ph.D. (probability and statistics) in 2009 from Michigan State University. Dr. Chakraborty was a faculty member at California State University in the Department of Mathematics from 2009 to 2012. She is currently teaching at the University of South Carolina. Her research interests include applied probability, stochastic differential equations, and stochastic modeling. Dr. Chakraborty may be reached at chakrabp@stat.sc.edu

SOURAV BANERJEE is an assistant professor of mechanical engineering at the University of South Carolina. He earned his B.S. degree (2000) from Bengal Engineering College, West Bengal, India, in civil engineering. He received his M.Tech., with a specialization in structural engineering, from the Indian Institute of Technology, Bombay, India, in 2002, and his Ph.D. (engineering mechanics) in 2005 from the University of Arizona. Dr. Banerjee is currently teaching at the University of South Carolina. His research interests include theoretical and computational modeling of ultrasonic waves and phonons in homogeneous and heterogeneous media (e.g., metal, composites, functionally graded materials etc.). He has published 22 peer-reviewed research papers and about 40 papers in conference proceedings. He has earned numerous awards and recognitions in the field of structural health monitoring such as the prestigious Achenbach Medal in 2010. Dr. Banerjee may be reached at banerjes@cec.sc.edu

A MULTIPROCESSOR VERSION OF MODULAR NEURAL NETWORKS FOR FAULT DIAGNOSIS OF ANALOG ELECTRONIC CIRCUITS

Mehran Aminian, St. Mary's University; Farzan Aminian, Trinity University

Abstract

A multiprocessor version of modular neural networks for fault diagnosis of real and simulated analog circuits was developed. In this approach, at the first level, the circuit faults were placed in two groups called level-one groups. At level two, the faults in each previous group were divided into two or three subgroups. This process continued until, at the last level, each group contained two or, at most, three faults. At each level, separate neural network modules were trained to identify the group to which the fault belonged. It has been shown previously that, when applied to simulated circuits, the uniprocessor version of this method performed accurately in diagnosing faults, even if they had similar features. This stems from the unique feature of this method by which the difficult task of fault classification is divided into smaller tasks assigned to several neural network modules whose sizes are considerably smaller than the neural network used in the conventional method. In this current study, the multiprocessor and uniprocessor versions of modular and conventional neural networks were used to diagnose faults in both simulated and real two-stage, four-op-amp biquad low-pass filters. It was shown that, while maintaining its high classification accuracy, the training time for the multiprocessor version of modular neural network was significantly less when compared to its uniprocessor counterpart. To create optimal features for neural network training, wavelet decomposition, principal component analysis, and data normalization were used on the raw data obtained from the impulse response of the circuits.

Introduction

Issues such as poor fault models, component tolerances, and non-linearity made fault diagnosis of electronic circuits very complicated. Neural networks can effectively handle these complexities, making the application of neural networks to fault diagnosis very appealing. The preliminary work in this area [1-6] suffers from a range of issues, the most important of which is using non-optimal features for neural network training. To address this issue, a combination of wavelet decomposition, principal component analysis, and normalization [7-11] were used to extract optimal features for neural network training. Other studies in this

area included the use of kurtosis and entropy [12] and S-Transform [13] to preprocess the output of the electronic circuits for optimal feature extraction. Different types of classifiers have also been applied to fault diagnosis of electronic circuits. This included application of polynomial curve fitting [14], radial basis neural networks [15], genetic algorithms and virtual instrumentation (LabVIEW) [16], [17], Particle Swarm neural networks [18], and distributed neural networks [19].

One of the biggest challenges in this area, however, is to effectively train neural networks in the presence of faults with extensive overlapping features. To address this issue, a modular neural network, in which the faults associated with the circuit are successively divided into several groups, was developed [20]. At the first level, the faults in the circuit were divided into two groups. At each subsequent level, each group at the previous level was divided into smaller groups until each group at the last level contained two or, at most, three fault classes. An appropriate number of neural networks were trained at different levels to identify the group to which the faults belonged. For example, one neural network was needed for fault classification at level one since there were only two groups to which the faults may belong. At level two, the number of fault groups increased to four, with each pair requiring one neural network for fault classification leading to a total of two neural networks. This process is shown in Tables 1 and 2, where $\uparrow$ and $\downarrow$ signify faulty components whose resistances or capacitances were higher and lower than nominal values, respectively. In the study by Aminian and Aminian [20], it was clearly demonstrated that modular neural networks lead to more efficient and robust training. Another group also reported success in distinguishing faults with similar features using S-Transform [13]. That study, however, was limited to a small number of faults.

Before this study, two problems remained unresolved in modular neural networks. First, the fact that different modules were completely independent of each other and may be trained in parallel was not exploited. This is an important feature of this architecture, which can significantly reduce the training time. Second, the modular neural networks were only applied to SPICE simulation, as opposed to real circuits [20].

Table 1. Fault Distribution among Neural Network Modules at Different Levels for the Simulated Circuit

Level	Groups	Fault distribution among groups	Neural network
1	1.1	C ₄ ↑, R ₁₇ ↓, R ₈ ↓, R ₁₆ ↑, R ₂₁ ↓, R ₇ ↑, C ₂ ↓, R ₁₀ ↑, R ₃ ↓, R ₆ ↑, R ₆ ↓, R ₉ ↑, C ₄ ↓, R ₁₅ ↑, R ₁₉ ↑, C ₁ ↑, R ₂₁ ↑, R ₃ ↑	1
	1.2	R ₇ ↓, R ₈ ↑, R ₉ ↓, R ₁₆ ↓, R ₁₉ ↓, R ₄ ↑, R ₁₀ ↓, R ₁₅ ↓, R ₂₅ ↑, R ₂₂ ↑, R ₂₅ ↓, C ₂ ↑, C ₃ ↓, R ₁₇ ↑, C ₄ ↓, R ₂₂ ↓, C ₃ ↑, C ₁ ↓	
2	2.1.1	C ₄ ↑, R ₁₇ ↓, R ₈ ↓, R ₁₆ ↑, R ₂₁ ↓, R ₇ ↑, C ₂ ↓, R ₁₀ ↑, R ₃ ↓	2
	2.1.2	R ₆ ↑, R ₆ ↓, R ₉ ↑, C ₄ ↓, R ₁₅ ↑, R ₁₉ ↑, C ₁ ↑, R ₂₁ ↑, R ₃ ↑	
	2.2.1	R ₇ ↓, R ₈ ↑, R ₉ ↓, R ₁₆ ↓, R ₁₉ ↓, R ₄ ↑, R ₁₀ ↓, R ₁₅ ↓, R ₂₅ ↑	3
	2.2.2	R ₂₂ ↑, R ₂₅ ↓, C ₂ ↑, C ₃ ↓, R ₁₇ ↑, C ₄ ↓, R ₂₂ ↓, C ₃ ↑, C ₁ ↓	
3	3.1.1.1	C ₄ ↑, R ₁₇ ↓, R ₈ ↓	4
	3.1.1.2	R ₁₆ ↑, R ₂₁ ↓, R ₇ ↑	
	3.1.1.3	C ₂ ↓, R ₁₀ ↑, R ₃ ↓	
	3.1.2.1	R ₆ ↑, R ₆ ↓, R ₉ ↑	5
	3.1.2.2	C ₄ ↓, R ₁₅ ↑, R ₁₉ ↑	
	3.1.2.3	C ₁ ↑, R ₂₁ ↑, R ₃ ↑	
	3.2.1.1	R ₇ ↓, R ₈ ↑, R ₉ ↑	6
	3.2.1.2	R ₁₆ ↓, R ₁₉ ↓, R ₄ ↑	
	3.2.1.3	R ₁₀ ↓, R ₁₅ ↓, R ₂₅ ↑	
	3.2.2.1	R ₂₂ ↑, R ₂₅ ↓, C ₂ ↑	7
	3.2.2.2	C ₃ ↓, R ₁₇ ↑, C ₄ ↓	
	3.2.2.3	R ₂₂ ↓, C ₃ ↑, C ₁ ↓	
4	4.1.1.1.1-3	C ₄ ↑ R ₁₇ ↓ R ₈ ↓	8
	4.1.1.2.1-3	R ₁₆ ↑ R ₂₁ ↓ R ₇ ↑	9
	4.1.1.3.1-3	C ₂ ↓ R ₁₀ ↑ R ₃ ↓	10
	4.1.2.1.1-3	R ₆ ↑ R ₆ ↓ R ₉ ↑	11
	4.1.2.2.1-3	C ₄ ↓ R ₁₅ ↑ R ₁₉ ↑	12
	4.1.2.3.1-3	C ₁ ↑ R ₂₁ ↑ R ₃ ↑	13
	4.2.1.1.1-3	R ₇ ↓ R ₈ ↑ R ₉ ↓	14
	4.2.1.2.1-3	R ₁₆ ↓ R ₁₉ ↓ R ₄ ↑	15
	4.2.1.3.1-3	R ₁₀ ↓ R ₁₅ ↓ R ₂₅ ↑	16
	4.2.2.1.1-3	R ₂₂ ↑ R ₂₅ ↓ C ₂ ↑	17
	4.2.2.2.1-3	C ₃ ↓ R ₁₇ ↑ C ₄ ↓	18
	4.2.2.3.1-3	R ₂₂ ↓ C ₃ ↑ C ₁ ↓	19

Table 2. Fault Distribution among Neural Network Modules at Different Levels for the Real Circuit

Level	Groups	Fault distribution among groups	Neural network
1	1.1	R ₂₆ S, R ₁₇ ↓, R ₁₀ ↑, R ₂₂ ↑, R ₁₆ ↓, R ₈ ↓, R ₂₁ ↑, R ₇ ↓, R ₂₅ ↑, R ₆ ↓, R ₁₅ ↓, R ₂₁ ↓, R ₈ ↑, R ₁₅ ↑, R ₄ ↓, R ₂₂ ↓, R ₃ ↓, R ₂₆ ↑	1
	1.2	R ₂₅ ↓, R ₄ ↑, R ₉ ↓, R ₂₁ S, R ₃ ↑, R ₃ S, R ₁₉ ↓, R ₁₉ ↑, R ₁₆ ↑, R ₁₇ ↑, R ₇ ↑, R ₁₆ S, R ₉ ↑, R ₁₀ ↓, R ₇ S, R ₆ ↑, R ₁₁ S, R ₂₆ ↓	
2	2.1.1	R ₂₆ S, R ₁₇ ↓, R ₁₀ ↑, R ₂₂ ↑, R ₁₆ ↓, R ₈ ↓, R ₂₁ ↑, R ₇ ↓, R ₂₅ ↑	2
	2.1.2	R ₆ ↓, R ₁₅ ↓, R ₂₁ ↓, R ₈ ↑, R ₁₅ ↑, R ₄ ↓, R ₂₂ ↓, R ₃ ↓, R ₂₆ ↑	
	2.2.1	R ₂₅ ↓, R ₄ ↑, R ₉ ↓, R ₂₁ S, R ₃ ↑, R ₃ S, R ₁₉ ↓, R ₁₉ ↑, R ₁₆ ↑	3
	2.2.2	R ₁₇ ↑, R ₇ ↑, R ₁₆ S, R ₉ ↑, R ₁₀ ↓, R ₇ S, R ₆ ↑, R ₁₁ S, R ₂₆ ↓	
	3.1.1.1	R ₂₆ S, R ₁₇ ↓, R ₁₀ ↑	4
	3.1.1.2	R ₂₂ ↑, R ₁₆ ↓, R ₈ ↓	
	3.1.1.3	R ₂₁ ↑, R ₇ ↓, R ₂₅ ↑	
3	3.1.2.1	R ₆ ↓, R ₁₅ ↓, R ₂₁ ↓	5
	3.1.2.2	R ₈ ↑, R ₁₅ ↑, R ₄ ↓	
	3.1.2.3	R ₂₂ ↓, R ₃ ↓, R ₂₆ ↑	
	3.2.1.1	R ₂₅ ↓, R ₄ ↑, R ₉ ↓	6
	3.2.1.2	R ₂₁ S, R ₃ ↑, R ₃ S	
	3.2.1.3	R ₁₉ ↓, R ₁₉ ↑, R ₁₆ ↑	
	3.2.2.1	R ₁₇ ↑, R ₇ ↑, R ₁₆ S	7
	3.2.2.2	R ₉ ↑, R ₁₀ ↓, R ₇ S	
	3.2.2.3	R ₆ ↑, R ₁₁ S, R ₂₆ ↓	
4	4.1.1.1.1-3	R ₂₆ S R ₁₇ ↓ R ₁₀ ↑	8
	4.1.1.2.1-3	R ₂₂ ↑ R ₁₆ ↓ R ₈ ↓	9
	4.1.1.3.1-3	R ₂₁ ↑ R ₇ ↓ R ₂₅ ↑	10
	4.1.2.1.1-3	R ₆ ↓ R ₁₅ ↓ R ₂₁ ↓	11
	4.1.2.2.1-3	R ₈ ↑ R ₁₅ ↑ R ₄ ↓	12
	4.1.2.3.1-3	R ₂₂ ↓ R ₃ ↓ R ₂₆ ↑	13
	4.2.1.1.1-3	R ₂₅ ↓ R ₄ ↑ R ₉ ↓	14
	4.2.1.2.1-3	R ₂₁ S R ₃ ↑ R ₃ S	15
	4.2.1.3.1-3	R ₁₉ ↓ R ₁₉ ↑ R ₁₆ ↑	16
	4.2.2.1.1-3	R ₁₇ ↑ R ₇ ↑ R ₁₆ S	17
	4.2.2.2.1-3	R ₉ ↑ R ₁₀ ↓ R ₇ S	18
	4.2.2.3.1-3	R ₆ ↑ R ₁₁ S R ₂₆ ↓	19

These two issues were addressed in this current study, where the multiprocessor implementation of modular neural networks in which the task of training each module was assigned to a separate processor. This method was applied to both real and SPICE-simulated two-stage, four-op-amp biquad low-pass filters. As shown here, the results show significant reduction in training time for multiprocessor implementation of modular neural networks compared to its single processor version, while maintaining high accuracy in fault diagnosis.

Generating Raw Data from the Circuit

In this current study, a modular neural network was applied to both real and simulated two-stage, four-op-amp biquad low-pass filters [21], as shown in Figure 1. Version 23 of the circuit simulation software NGSPICE, installed on Ubuntu 11.10, was used to obtain the raw simulated data from the circuit output. To generate the data from the real circuit, it was first constructed on a National Instruments ELVIS Prototype Board after which LabVIEW was utilized to obtain its raw output. In both cases, a pulse with the characteristics shown in Table 3 was used to excite the system. This input pulse was chosen because it had a frequency spectrum in the range 0-100 kHz. Since the cutoff frequency of the two-stage, four-op-amp biquad low-pass filter was at 10 kHz, the applied pulse exposed the complete frequency behavior of the circuit. Therefore, features extracted from the response of the circuit to such an input pulse carried the necessary information for its fault diagnosis.

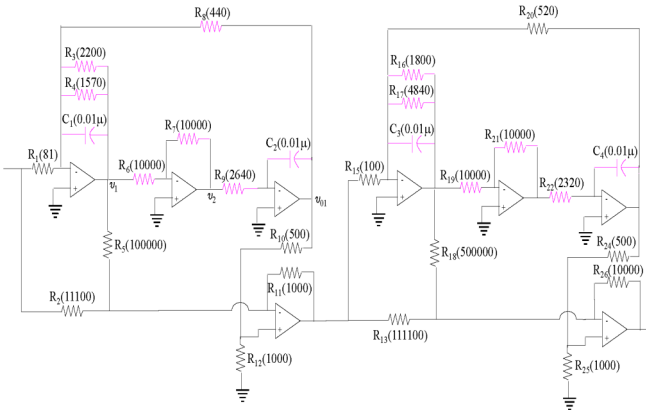


Figure 1. Circuit Diagram for the Two-Stage, Four-Op-Amp Biquad Low-Pass Filter Used in This Study

The nominal values for the components, which resulted in a non-faulty behavior, are shown in Tables 4 and 5 for simulated and real circuits, respectively. To introduce faults into the circuit, the value of one of the resistors or capacitors was allowed to fall outside of its tolerance, 5% for re-

sistors and 10% for capacitors, while other components varied within their allowed range. To generate sample data, five resistors and capacitors were allowed to change values within their tolerances in the simulated circuit. For the real circuit, however, the five elements whose values changed within their tolerances were limited to resistors. This is because variable resistors are more readily available than variable capacitors. The faulty values of different resistors and capacitors used in this study are shown in Tables 4 and 5 for simulated and real circuits, respectively. Note that the absence of the hard faults—replacing R_3 , R_7 , R_{11} , R_{16} , R_{21} , and R_{26} with short circuits—in the simulated circuit was due to numerical instabilities encountered in simulating the circuit using SPICE.

Table 3. Descriptions of the Pulse Used to Excite the Real Circuit

Amplitude	1V
Period	1ms
Delay	5 μ s
Pulse Width	10 μ s
Rise/Fall Time	1ns

Preprocessing

To prepare optimal features for neural network training, the raw output obtained from the circuit were processed using wavelet decomposition, principal component analysis, and normalization [8-10]. Previously, Discrete Wavelet Transform was used to decompose the signal into low- and high-frequency components [7], [11], [20], [22], [23]. This was accomplished by computing the approximation (low-frequency) and detail (high-frequency) coefficients using Equation (1):

$$C(a, b) = \frac{1}{\sqrt{a}} \int I(x) \psi\left(\frac{x-b}{a}\right) dx \quad (1)$$

where, $I(x)$ is the raw output signal of the circuit; $\psi\left(\frac{x-b}{a}\right)$ is the shifted and scaled version of the mother wavelet; $a = 2^j$, $b = k2^j = ka$; and, $(j, k) \in Z^2$ [8], [9].

This process was repeated on approximation coefficients to further decompose them into high- and low-frequency components. This method was very effective for extracting useful features for neural network training. In this study, however, Wavelet Packet Decomposition (WPD) in which detail as well as approximation coefficients are decomposed into high- and low-frequency bands at different levels was utilized. This provided a richer accumulation of wavelet

coefficients. A small number of these coefficients were then collected for the next stage of processing, namely principal component analysis.

Table 4. Fault Classes Used for the Simulated Circuit

Fault Class	Nominal Value	Faulty Value
C ₁ ↑	0.01μF	0.05μF
C ₁ ↓	0.01μF	0.002μF
C ₂ ↑	0.01μF	0.05μF
C ₂ ↓	0.01μF	0.002μF
C ₃ ↑	0.01μF	0.05μF
C ₃ ↓	0.01μF	0.002μF
C ₄ ↑	0.01μF	0.05μF
C ₄ ↓	0.01μF	0.002μF
R ₃ ↑	2200Ω	5000Ω
R ₃ ↓	2200Ω	500Ω
R ₄ ↑	1570Ω	5000Ω
R ₄ ↓	1570Ω	600Ω
R ₆ ↑	10000Ω	16500Ω
R ₆ ↓	10000Ω	3000Ω
R ₇ ↑	10000Ω	50000Ω
R ₇ ↓	10000Ω	5500Ω
R ₈ ↑	440Ω	2200Ω
R ₈ ↓	440Ω	50Ω
R ₉ ↑	2640Ω	10000Ω
R ₉ ↓	2640Ω	300Ω
R ₁₀ ↑	500Ω	800Ω
R ₁₀ ↓	500Ω	100Ω
R ₁₅ ↑	100Ω	200Ω
R ₁₅ ↓	100Ω	50Ω
R ₁₆ ↑	1800Ω	7800Ω
R ₁₆ ↓	1800Ω	100Ω
R ₁₇ ↑	4840Ω	15000Ω
R ₁₇ ↓	4840Ω	1600Ω
R ₁₉ ↑	10000Ω	30000Ω
R ₁₉ ↓	10000Ω	3000Ω
R ₂₁ ↑	10000Ω	50000Ω
R ₂₁ ↓	10000Ω	3750Ω
R ₂₂ ↑	2320Ω	7000Ω
R ₂₂ ↓	2320Ω	100Ω
R ₂₅ ↑	1000Ω	3000Ω
R ₂₅ ↓	1000Ω	500Ω

Applying principal component analysis to selected wavelet coefficients further reduced the dimensionality of the input feature space, while keeping much of the relevant information for fault classification intact [10]. Principal component analysis achieved this goal by mapping vectors x in a d -dimensional input space $(x_1, \dots, x_d)$ onto vectors z in an M -dimensional space $(z_1, \dots, z_M)$ in such a way that loss of essential information and data variation needed for classification was minimized while keeping $M < d$. In the last preprocessing stage, the data were normalized to have

zero mean and unit standard deviation. This minimized large variations in feature sizes, which can dominate important but smaller trends in the data. As shown in the study by Aminian and Aminian [7], normalizing features in this fashion can make the neural network training phase very efficient.

Table 5. Fault Classes Used for the Real Circuit

Fault Class	Nominal Value	Faulty Value
R ₃ ↑	2200Ω	5000Ω
R ₃ ↓	2200Ω	500Ω
R ₄ ↑	1570Ω	5000Ω
R ₄ ↓	1570Ω	600Ω
R ₆ ↑	10000Ω	16500Ω
R ₆ ↓	10000Ω	3000Ω
R ₇ ↑	10000Ω	50000Ω
R ₇ ↓	10000Ω	5500Ω
R ₈ ↑	440Ω	2200Ω
R ₈ ↓	440Ω	50Ω
R ₉ ↑	2640Ω	10000Ω
R ₉ ↓	2640Ω	300Ω
R ₁₀ ↑	500Ω	800Ω
R ₁₀ ↓	500Ω	100Ω
R ₁₅ ↑	100Ω	200Ω
R ₁₅ ↓	100Ω	50Ω
R ₁₆ ↑	1800Ω	7800Ω
R ₁₆ ↓	1800Ω	100Ω
R ₁₇ ↑	4840Ω	15000Ω
R ₁₇ ↓	4840Ω	1600Ω
R ₁₉ ↑	10000Ω	30000Ω
R ₁₉ ↓	10000Ω	3000Ω
R ₂₁ ↑	10000Ω	50000Ω
R ₂₁ ↓	10000Ω	3750Ω
R ₂₂ ↑	2320Ω	7000Ω
R ₂₂ ↓	2320Ω	100Ω
R ₂₅ ↑	1000Ω	3000Ω
R ₂₅ ↓	1000Ω	500Ω
R ₂₆ ↑	10000Ω	15000Ω
R ₂₆ ↓	10000Ω	5000Ω
R ₃ Short	2200Ω	0.5Ω
R ₇ Short	10000Ω	0.5Ω
R ₁₁ Short	1000Ω	0.5Ω
R ₁₆ Short	1800Ω	0.5Ω
R ₂₁ Short	10000Ω	0.5Ω
R ₂₆ Short	1000Ω	0.5Ω

Applying Modular Neural Networks to Analog Fault Diagnosis

The application of modular neural networks to fault diagnosis of electronic circuits started with organizing all of the faults into two groups at level one. At level two, each fault

group in the previous level was divided into two or three subgroups such that the faults were as evenly distributed among these subgroups as possible. This operation continued at the higher levels, dividing each subgroup at level n into smaller groups at level $n + 1$, as shown in Figure 2. The highest level was reached when each subgroup at that level contained two or, at most, three fault classes. The implementation of this method in the circuit resulted in dividing the two-stage, four-op-amp biquad low-pass filter into 2, 2, 3, and 3 groups at levels one through four, respectively. This is presented in Tables 1 and 2 for the simulated and real circuits, respectively. As shown in columns two of these tables, a series of numbers were used to label the subgroups at different levels. The first number indicates the level and all of the subsequent numbers represent the way in which the subgroups are divided at different levels. For example, the faults at level one were divided into the two groups labeled 1.1 and 1.2. At level two, subgroup 1.1 was divided into two smaller groups, namely 2.1.1 and 2.1.2. The corresponding division of subgroup 1.2 led to subgroups 2.2.1 and 2.2.2, etc.

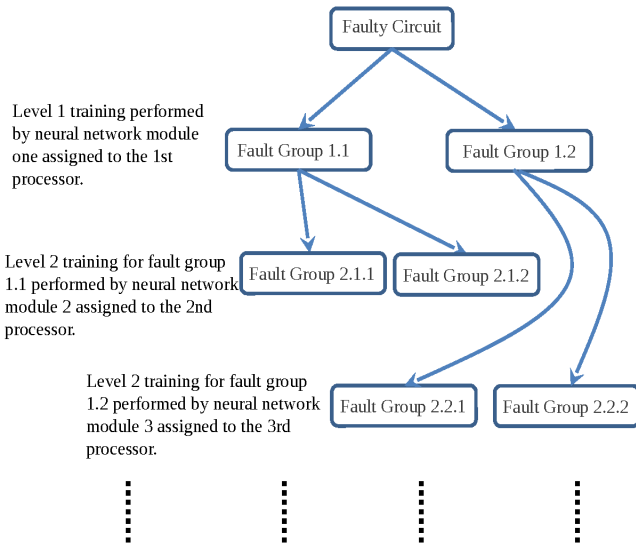


Figure 2. Fault Distribution among Groups at Different Levels

Next, at each level, a collection of very simple neural network modules were trained using a backpropagation algorithm [10], [24] with sigmoid transfer functions to classify the subgroups to which a given fault belonged. Applying this method to two-stage, four-op-amp biquad low-pass filters resulted in training one neural network module at level one to distinguish among faults in groups 1.1 and 1.2. This neural network module was labeled 1 in column four of Tables 1 and 2. At level two, two separate neural network modules labeled 2 and 3 in Tables 1 and 2 were used. One of these networks was for fault classification between 2.1.1

and 2.1.2 and the other for groups 2.2.1 and 2.2.2. This process continued until the last level, level four, was reached. As shown in column four of Tables 1 and 2, a total of nineteen neural network modules were needed to perform the fault classification for the two-stage, four-op-amp biquad low-pass filter. The targets for the neural network modules were constructed using 1-of-K architecture [7], [10], [11], [20], [23]. For example, the targets of

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

were used to train the neural network at level one to classify fault groups in 1.1 or 1.2. At level three, however, the targets

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

were used to train the neural network to distinguish among groups 3.1.1.1, 3.1.1.2, and 3.1.1.3. The output of the neural network with 1-of-K architecture had the advantage of representing the probability that a sample test data belonged to different fault classes [10]. To identify the group to which the fault belonged, the outputs of all of the neural networks at different levels were used. For example, the outputs of

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix},$$

indicated that the faulty component belonged to groups 1.1, 2.1.2, 3.1.2.3, and 4.1.2.3.2 at levels 1 through 4, respectively. This indicated that $R_3\downarrow$ was the faulty component in the circuit.

It is important to emphasize that each of these networks had a very simple task of identifying the group to which a particular fault belonged at different levels. The modular implementation of neural networks is considerably more efficient than the conventional approach. This efficiency stems from two important features of the modular approach to neural network training. First, in the conventional approach, the partitioning of the entire feature space into different fault classes was performed by one neural network. This can make the training of the neural network especially difficult when there is extensive overlap among fault classes in the feature space. This problem becomes much less pronounced in the modular approach since the input space is divided into progressively smaller subspaces or partitions at lower levels. Using simple techniques such as boxplot to determine the degree of overlap among features (see Figures 3 and 4), one can distribute fault classes among partitions such that classes with similar features are distributed among

different partitions in a particular level as evenly as possible. In this scheme, once one of the partitions is identified as containing the faulty component by a neural network, the remaining partitions and, hence, the similar faults assigned to them, are eliminated from further investigation. The training phase at the next level is now easier to conduct, as similar faults make the training of neural networks more challenging. Second, the modular approach allowed the training of the neural networks to be performed by separate processors in parallel, since the training of each module was completely independent of any other module at all levels. This can significantly reduce the training time, as will be shown in the Results section.

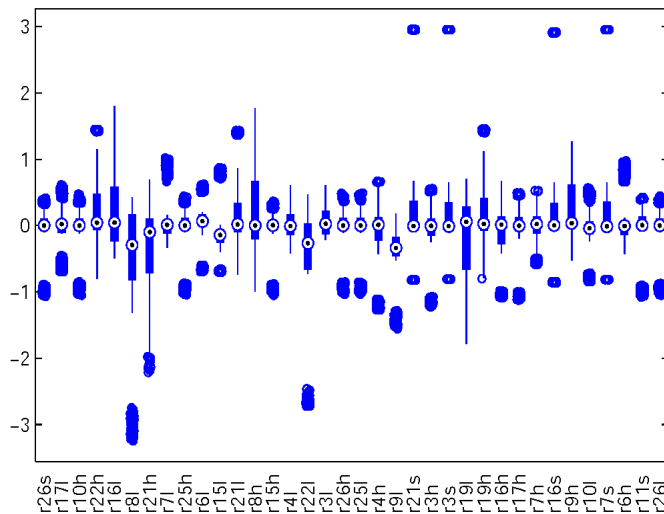


Figure 3. Boxplot of the Features Extracted From the Data Obtained From the Real Circuit

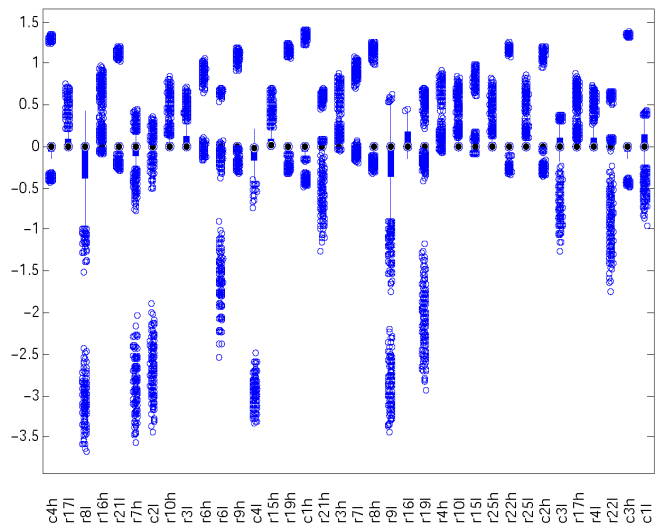


Figure 4. Boxplot of the Features Extracted From the Data Obtained From the Simulated Circuit

Results

In this study, modular neural network architecture was used to perform fault diagnosis on both real and simulated two-stage, four-op-amp biquad low-pass filters. This required 19 different neural network modules to fully classify the 36 faults introduced in the simulated and real circuits, as shown in Tables 1 and 2, respectively. The training of these modules was performed in parallel using 19 processors. For the purpose of comparison, the same modular neural network using a single processor [20] and a conventional neural network were also trained to perform the classification of all 36 faults in the circuit.

The architectures of the modular and conventional neural networks used in this study are shown in Tables 6 and 7 for the simulated and real circuits, respectively. Note the drastic reduction in the number of hidden layer neurons in the modular architecture compared to the conventional approach. This led to superior performance and a dramatic reduction in training time, as shown in Tables 8 and 9 for the two circuits. Note that the average training times taken over 15 trials for the simulated and real circuits were reduced by a factor of 120 and 23, respectively, when a uniprocessor modular neural network was used for fault classification compared to its conventional counterpart. This stems from distributing the difficult task of fault classification into smaller tasks performed by several simple neural network modules.

Table 6. Number of Hidden Neurons in the Neural Network used to Classify Faults in the Simulated Circuit

Number of hidden layer neurons used for simulated circuit		
Conventional	Modular	
	Module number	Number of neurons
16	1	9
	2	5
	3	5
	4	3
	5	3
	6	3
	7	3
	8	3
	9	3
	10	3
	11	3
	12	3
	13	3
	14	3
	15	3
	16	3
	17	3
	18	3
	19	3

Table 7. Number of Hidden Neurons in the Neural Network used to Classify Faults in the Real Circuit

Number of hidden layer neurons used for simulated circuit		
Conventional	Modular	
	Module number	Number of neurons
20	1	9
	2	6
	3	4
	4	3
	5	2
	6	2
	7	3
	8	5
	9	3
	10	4
	11	2
	12	4
	13	3
	14	4
	15	3
	16	3
	17	4
	18	2
	19	3

Table 8. Performance Accuracy and Average Training Times for the Conventional, Uniprocessor Version, and Multiprocessor Version of the Modular Neural Networks Applied to the Simulated Circuit

Simulated Circuit			
Number of processors	Neural network architecture	Performance accuracy (percent)	Training time (minutes)
1	Conventional	91	124.64
	Modular (uniprocessor)	100	1.04
19	Modular (multiprocessor)	100	0.15

Table 9. Performance Accuracies and Average Training Times for the Conventional, Uniprocessor Version, and Multiprocessor Version of the Modular Neural Networks Applied to the Real Circuit

Actual Circuit			
Number of processors	Neural network architecture	Performance accuracy (percent)	Training time (minutes)
1	Conventional	98	110.93
	Modular (uniprocessor)	98 (level 1) 100 (levels 2-4)	4.89
19	Modular (multiprocessor)	97 (level 1) 100 (levels 2-4)	0.51

As shown in Tables 8 and 9, the training time of parallel modular neural networks using 19 processors resulted in 7- and 10-fold reductions in training time compared to the uniprocessor version for the simulated and real circuits, respectively. This significant reduction stemmed from two main features of modular neural networks. First, modular architecture divided the complex task of classifying all of the faults performed by one neural network into smaller and much simpler tasks performed by each module. Second, the training of each neural network module was performed in parallel by different processors, because the neural network modules were completely independent of each other. It is important to note that the efficiency of parallel modular neural networks increases with the number of fault classes.

An examination of Tables 8 and 9 showed that the modular neural network performs more accurately compared to the conventional method. To measure the fault classification accuracy against other published works, the study by Tan et al. [13] was particularly useful. In that work, the authors applied S-Transform and Particle Swarm Optimization to diagnose three faults with similar features in the same circuit as the one used in this current study. Tan et al. reported accuracies ranging from 92.76% to 96.82%. The system in this current study could classify 36 faults, three of which were similar to the ones used by Tan et al., with 100% accuracy as shown in Tables 8 and 9. In the studies by Ke et al. [25] and Guo et al. [26], diagnostic accuracies ranging from 83% to 100% were reported in classifying 13 faults associated with a single-stage, four-op-amp biquad high-pass filter. This makes direct comparison between the two results difficult, since those other studies used a smaller circuit with fewer faults.

Conclusion

In this current study, a fault diagnostic system for analog electronic circuits using a multiprocessor version of modular neural networks was developed. In this approach, the faults associated with the circuit were successively divided into groups. At each level, where a fault group was subdivided into smaller groups, a set of neural network modules was trained in order to determine the groups to which the faults of interest belonged. The modular fault diagnostic system provided a very efficient and reliable approach for diagnosis of circuits of various sizes. This approach is especially suited for larger circuits with a significant number of faults. This technique also performs reliably in the presence of faults with similar features, where the traditional approach can fail. The advantage of the modular system lies in its ability to break down a difficult classification problem in the traditional approach into several simpler problems to which much smaller neural networks may be applied. This

significantly enhances the efficiency of the training phase and the performance of the fault diagnostic system. The results presented here clearly demonstrated the superiority of the modular system, which led to higher accuracy in the classification of fault classes, while reducing the training time. Comparing the multiprocessor version of modular neural networks with the traditional approach, the training time was reduced by a factor of 831 and 218 for the simulated and real two-stage, four-op-amp biquad low-pass filters, respectively.

References

- [1] Spence, H. (1996). Automatic Analog Fault Simulation. *Proceedings of the AUTOTESTCON'96 Conference*, (pp. 17–22).
- [2] Liu, R. W. (1991). *Testing and Diagnosis of Analog Circuits and Systems*. Van Nostrand Reinhold Publishing Company.
- [3] Spina, R., & Upadhyaya, S. (1997). Linear Circuit Fault Diagnosis Using Neuromorphic Analyzer. *IEEE Transactions on Circuits and Systems II*, 44(3).
- [4] (1987). *Selected Papers on Analog Fault Diagnosis*. IEEE Press.
- [5] Bandler, J. W., & Salama, A. (1985). Fault Diagnosis of Analog Circuits. *Proceedings of IEEE*, 73, (pp. 1279–1325).
- [6] Catelani, M., & Gori, M. (1996). On the Application of Neural Networks to Fault Diagnosis of Electronic Analog Circuits. *Measurement*, 17, 73–80.
- [7] Aminian, M., & Aminian, F. (2000). Neural-Network Based Analog Circuit Fault Diagnosis Using Wavelet Transform as Preprocessor. *IEEE Transactions on Circuits and Systems II*, 47(2), 151–156.
- [8] Hubbard, B. B. (1996). *The World According to Wavelets*. A. K. Peters Ltd.
- [9] Strang, G., & Nguyen, T. (1996). *Wavelet and Filter Banks*. Wellesley-Cambridge Press.
- [10] Bishop, C. M. (1995). *Neural Networks for Pattern Recognition*. Oxford University Press.
- [11] Aminian, F., Aminian, M., & Collins, B. (2002). Analog Fault Diagnosis of Actual Circuits Using Neural Networks. *IEEE Transactions on Instrumentation and Measurement*, 51(3), 544–550.
- [12] Yuan, L., Yigang, H., Huang, J. Y., & Sun, Y. (2010). A New Neural-Network-Based Fault Diagnosis Approach for Analog Circuits by Using Kurtosis and Entropy as a Preprocessor. *IEEE Transactions on Instrumentation and Measurement*, 59(3), 586–595.
- [13] Tan, A. Y., Sun, B. Y., He, Y., & Zhang, D. G. (2012, January). Analog Fault Diagnosis based on S-Transform and PSO. In *Proceedings of the International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA)* (p. 1). The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp).
- [14] Kumar, A., & Singh, A. P. (2013). Electronic Circuit Using Polynomial Curve Fitting. *International Journal of Computer Applications*, 61(16).
- [15] Wuming, H., & Peliang, W. (2010). Analog Circuit Fault Diagnosis Based on RBF Neural Network Optimized by PSO Algorithm. *International Conference on Intelligent Computation Technology and Automation (ICICTA)*, (pp. 628–631).
- [16] Yang, L. Z. Wang, S., & Zhai, G. (2012, August). *Analog Circuits Fault Diagnosis by GA-RBF Neural Network and Virtual Instruments*. Symposium conducted at the International conference on Measurement, Sensor Network and Automation (IMSNA).
- [17] Xiang, L., Yang, Z., Shujuan, W., & Guofo, Z. (2011). A Method for Analog Circuits Fault Diagnosis by Neural Network and Virtual Instruments. *3rd International Workshop on Intelligent Systems and Applications (ISA)*. (pp. 1–5).
- [18] Han, B., Wu, H., & Liu, S. (2013). Analog Circuit Fault Diagnosis Method Based on Particle Swarm Neural Network. *Journal of Theoretical and Applied Information Technology*, 49(3).
- [19] Wang, L., Liu, Y., Li, X., Guan, J., & Song, Q. (2010). Analog Circuit Fault Diagnosis Based on Distributed Neural Network. *Journal of Computers*, 5 (11).
- [20] Aminian, M., & Aminian, F. (2007). A Modular Fault Diagnosis System for Analog Electronic Circuits Using Neural Networks With Wavelet Transform as Preprocessor. *IEEE Transactions on Instrumentation and measurement*, 56(5), 1546–1554.
- [21] Valkenburg, M. V. (1982). *Analog Filter Design*. Oxford University Press.
- [22] Aminian, F., & Aminian, M. (2001). Fault Diagnosis of Analog Circuits Using Bayesian Neural Networks With Wavelet Transform as Preprocessor. *Journal of Electronic Testing: Theory and Applications*, (17), 29–36.
- [23] Aminian, F., & Aminian, M. (2001). Fault Diagnosis of Nonlinear Circuits Using Neural Networks With Wavelet and Fourier Transforms as Preprocessors. *Journal of Electronic Testing: Theory and Applications*, 17, 471–481.
- [24] Hagan, M. T., Demuth, H. B., & Beale, M. (1996). *Neural Network Design*. PWS Publishing.
- [25] Ke, G., Ye, S., & Huishu, S. (2013). An Improved DAGSVM Approach for Analog Circuit Fault Diagnosis with KPCA plus LDA as Preprocessor. *International Journal of Advancements in Computing Technology*, 5(6).

-
- [26] Guo, K., Zhu, Y., & San, Y. (2012). Analog circuit intelligent fault diagnosis based on greedy kpca and one-against-rest SVM approach. *Journal of Theoretical and Applied Information Technology*, 46(1), 147-157.

Biographies

MEHRAN AMINIAN received his M.S. and Ph.D. degrees in electrical engineering from the University of Oklahoma, Norman, in 1982 and 1989, respectively. Currently, he is a professor of electrical engineering at St. Mary's University in San Antonio, Texas. His research interests include neural networks and quantum collision theory. Dr Mehran Aminian may be reached at maminian@stmarytx.edu

FARZAN AMINIAN received his M.S. and Ph.D. degrees in electrical engineering from The Ohio State University, Columbus, in 1983 and 1989, respectively. He joined the Engineering Science Department at Trinity University, San Antonio, in 1989, where he is currently a professor and chair. His areas of interest include application of Monte Carlo techniques, solid state devices, and neural networks. Dr. Aminian served as the chairman of the Central Texas Section of IEEE. Dr Farzan Aminian may be reached at faminian@trinity.edu

MODELS FOR PREDICTING THE COMPRESSION INDEX OF SOILS USING FRACTAL SCALING OF PARTICLE SIZE DISTRIBUTIONS

Fred Boadu, Duke University

Abstract

In this study, the author improved empirical equations to predict the compression index of soil incorporating fractal parameters describing their particle size distributions (PSD) and standard index properties. A fractal model was utilized to describe the PSD of soils in two fractal scaling domains, each characterizing different ranges of particle sizes. Incorporating the fractal parameters (fractal dimensions) as predictor variables in regression models significantly improved the prediction of the compression index. Models to predict the compression index of soils using 1) the fractal dimension solely as input predictor variables and 2) a combination of fractal parameters and index properties as input predictor variables were also developed.

Models developed in other studies were compared with models developed in this current study and their predictive performances were assessed. For single and multiple predictor variables, the predictive empirical models developed in this study performed better than those in other studies. Using the fractal dimension as a single input predictor variable provided a better predictive model than using both initial void ratio, e_0 , and liquid limit, w_L , as input variables as in other studies. The empirical models developed in this study are suggested as useful alternatives for estimating the compression index of soils. In certain circumstances it may also be useful in giving initial information about soil compressibility in a field environment based on analysis of particle size distribution.

Introduction

The mechanical and hydraulic behavior of granular materials such as soils is affected by their particle-size distributions (PSD), and these distributions develop as a result of complex geological and geophysical processes. The three most important engineering properties of soils, which are influenced by their PSD, include shear strength, compressibility, and hydraulic conductivity [1-3]. The PSD of soils have been shown to influence the strength and permeability of materials used in compacted dams [4]. Experimental and modeling studies [5-8] indicated that the mechanical behavior of soils is influenced by their PSDs.

Changes in a soil's volume results in changes in its strength and deformation properties, which, in turn, affects its stability [1]. Compressibility is a measure of the volume change behavior of a soil mass subjected to different amounts of external pressure and is quantified using either the coefficient of volume compression, M_v , or the compression index, C_c . Compression index is required for computation of the settlements of structures founded on soils and can be obtained by performing laboratory tests on soil samples using the oedometer test. However, an oedometer test can be relatively expensive and time consuming in comparison to standard index tests. Thus, several attempts have been made by various researchers to estimate the compression index of soils from their index properties that are easily measured in the laboratory or field environment, such as liquid limit, w_L , natural moisture content, w_n , and initial void ratio, e_0 [9-14].

Volume changes or settlements in clayey sediments depend also on the relative proportions of clay, silt, and sand inherent in them, but most studies have not given much attention to the influence of particle size distribution. However, particle size and shape and distribution are the most important factors influencing both the void ratio at any stress level and the effects that physico-chemical and other mechanical factors have on consolidation and swelling of soils [1], [15]. Also, studies on the compressional behavior of clayey sand indicate that soil texture (fines and sand content) evidently influence the compressional behavior of soils with findings that, as the fines content increases, the compression index increases [16-18]. Experimental studies indicate that the effective size, D_{10} , the mean grain size, D_{50} , and the coefficient of uniformity, C_u , are pertinent grading parameters that can be used to characterize the undrained shear strength (liquefaction resistance) of sand-silt mixtures [19].

One-dimensional compression experiments on sand [20] show that porosity or void ratio at any given effective stress level is highly dependent on grain size, grain-size distribution, and mineralogy. The discrete-element method (DEM) has been utilized to investigate effects of particle-size distribution on the compressive behavior of granular soils [21]. The challenge in fully utilizing PSD appears to be how to uniquely quantify the distribution of particle sizes in a soil

beyond the summary statistics, such as D_{10} , D_{50} , and C_u . In recent times, researchers have turned to fractal analysis in search of more fundamental and representative parameters to characterize PSD of soils. Soil micro-structure, as reflected in the inter-connectivity among the soil grains, has been quantified using fractal principles [22]. An improved model for predicting hydraulic conductivity of soils has been developed using unique representations of their particle-size distribution and fractal dimensions [3]. Size distributions and stabilities of soil aggregates have been related to their fractal dimensions [23]. However, the existence of a single unique fractal dimension in describing the PSD of soils has been questioned [24], [25] on evident deviations from the linear behavior of log-log plots in the estimation of the fractal dimensions. Fractal scaling of soils over six orders of magnitude associated with the distribution of their particle sizes have been determined [26].

The above discussions indicate that the compressibility of soils depends on their particle-size distributions. Presently, though, a general empirical equation for compressibility and soil gradation incorporating parameters that uniquely represent or quantify the PSDs is not yet available. In this study, the author investigated the influence of particle-size distribution of soils on their compressive behavior using information from laboratory experiments. The distribution of the particle sizes was quantified by their segmented fractal dimensions obtained by partitioning PSDs of field soil samples into two scaling domains using a piecewise fractal model. Relationships between the segmented fractal dimensions and compression index of soils obtained in the field were analyzed. Comparative assessments of compressive index predictions using regression models were performed by which predictions of compressive index using fractal parameters (this study) were compared with models, which use index properties of soils as prediction variables (other studies). Improvements in the predictions of compressive index by inclusion of fractal dimensions as unique representative parameters for particle size distribution of soils was then assessed.

Soil Sampling and Analysis

Undisturbed and disturbed soil samples were obtained at 32 sampling sites, mainly test pits and fresh highway cuts, with varying subsurface geology consisting of sedimentary, metamorphic, and intrusive igneous rocks in North Carolina and South Carolina (U.S.). The collected soils samples were weathered products from the underlying bedrocks of different rock types with inherent wide variations in composition and texture. The undisturbed samples were obtained with Shelby tubes and used for consolidation tests, while the disturbed samples were used for determination of index prop-

erties and particle-size distribution. All of the sampled soils were sealed and taken directly to the laboratory and their natural moisture contents [27], liquid limits [28], and particle-size distributions were determined [29]. The distribution of particle sizes larger than 0.075 mm (retained on the No. 200 sieve) was determined by sieving, while that of particle sizes smaller than 0.075 mm was determined by a sedimentation process using a hydrometer. The data from the sieve and hydrometer analyses were then combined to obtain a distribution of soil particle sizes ranging from 0.0011 mm to 4.75 mm. Figure 1 illustrates the grain-size distribution of a few selected soils used in this current study.

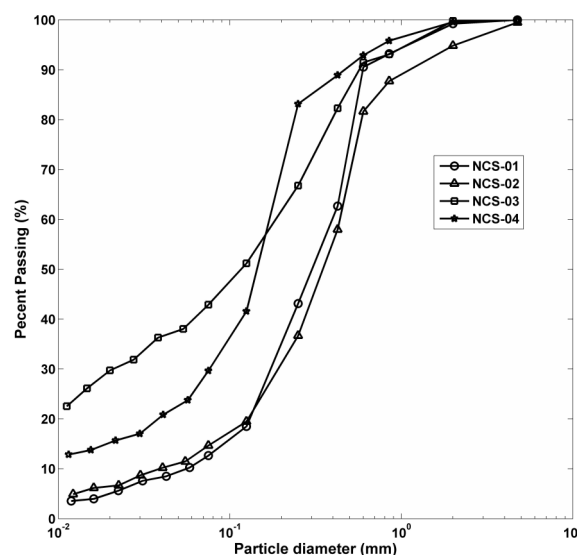


Figure 1. Plot of the Particle Size Distribution (PSD) Curves for Representative Soil Samples

Fractal Analysis and Scaling of PSD

Fractal concepts have been utilized to describe the physical properties of soils such as bulk density [30], pore-size distribution [31], [32], pore surface area [33], and particle-size distribution [26], [34]. The concept of fractal scaling as it applies to PSD suggests that, across a wide range of scales, the solid phase of a soil appears self-similar. The idea was originated by Mandelbrot [35], who showed that observations of patterns or distributions at different scales could be related to each other by a power law function with an exponent termed the fractal dimension, D .

Characterization of particle-size distribution using fractal concepts inherently assumes a fragmentation process [36]. In this approach, the mass or weight of soil grains whose sizes range between an upper and lower bound defined by

diameters obtained from sieve and hydrometer analysis was used. A mass-based relation [34], [36], [37] contained the conventional expression describing particle-size distribution data; that is, “percentage of mass less than” or equivalently $M(d < d_i)$ given by Equation (1):

$$M(d < d_i) = cd_i^{3-D} \quad (1)$$

where, $M(d < d_i)$ is the cumulative mass of soil grains below an upper limit d_i ; D is the fractal dimension of the PSD; and, c is a composite scaling constant.

In practice, D is estimated from Equation (1) by log-transforming both sides of the equation and employing a linear regression fit [38], [39].

Segmented Fractal Model

The assumption made in capturing the basic features observed in the PSDs of the studied soils was that there could be two different scaling domains in different ranges of particle sizes over which values of D could be different [40], [41]. In this respect, there would also be a critical or break point value, d_c , which could be characterized as the upper limit of the first domain and the lower cutoff of the second domain. These characteristic features necessitated the consideration of a piecewise approximation of the dependence $M(d \leq d_i)$ on d_i in such a way that it could be used to describe the possible fractal segmentation of PSDs and given by Equation (2):

$$M(d \leq d_i) = c_1 d_i^{3-D_1} (d_i \leq d_c) + c_2 d_i^{3-D_2} (d_i > d_c) \quad (2)$$

where, D_1 is the fractal dimension associated with the first domain; D_2 is that of the second domain; and, d_c is the cut-off size separating the two particle-size domains.

Each term in parentheses represents a logical operation, which partitions the whole domain into two subdomains, $0 < d_i \leq d_c$, and $d_c < d_i \leq 4.75$ mm. The premise in using Equation (2) is that the cumulative percentage of the soil mass can be represented by two geometric structures in terms of their respective D values [40], [41]. In the second stage, the two phases were joined smoothly by a transition function, which depended on the transition point chosen by the grid search [42-44].

In solving Equation (2), a two-phase non-linear regression method with smooth transitions between phases [42-44] was employed. In this method, the estimation of the parameters (c_1 , D_1 , c_2 , D_2 , and d_c) was performed in two stages. In the first stage, each phase was fitted separately with the transition point determined simultaneously. The method eliminat-

ed non-uniqueness of solutions and ensured smoothness of the resulting regression function. Double logarithm plots of PSDs for four representative soils are shown in Figures 2(a-d) illustrating that a single power law cannot be used to describe the data across the measured particle sizes and all identify two scaling domains. The data up to the critical size (break point) were fitted with a line with a slope related to D_1 , and beyond the critical size data fitted to a line with slope related to D_2 . For Figure 2(a), D_1 , the fractal dimension of the first domain for this soil, was 2.29; D_2 , the fractal dimension of the second domain, was equal to 2.93; d_c , the critical size related to the cut-off domain, was equal to 0.83 mm; and, R^2 , for the overall fit, was equal to 0.98. For the soil shown in Figure 2(b), D_1 , D_2 , d_c , and R^2 values were 2.66, 2.96, 0.66 mm and 0.99, respectively.

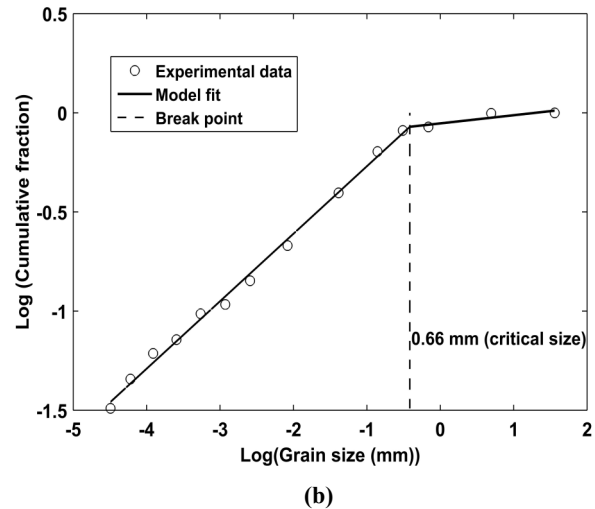
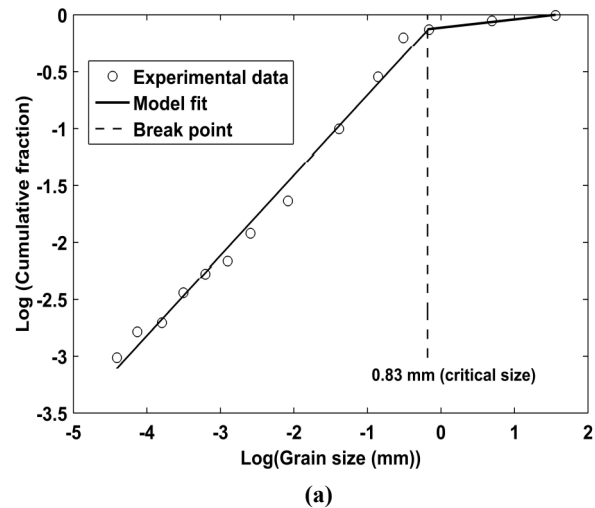


Figure 2. Double Logarithm Plot of Representative Soil Samples Identifying Two Scaling Domains for Each Soil PSD

Consolidation Tests

One-dimensional consolidation tests were conducted according to the standard test methods for one-dimensional consolidation properties of soils [45]. Tests were conducted using rear-loading oedometers equipped with linear variable differential transducers, which were connected to a data acquisition system. An oedometer ring with a diameter of 63.5 mm and height of 20 mm was used for all the tests. The inside of the rings was lubricated with silicone grease to minimize side friction between the ring and the soil specimen during deformation. The consolidation tests were conducted at a room temperature of 20°C. Soil samples were trimmed and inserted into the consolidation rings between two porous stones. The porous stones allowed water to drain from the top to the bottom of the soil sample when vertical stress was applied. Filter papers were placed on the top and bottom of each soil sample to prevent particles from being forced into the pores of the porous stones placed on both sides of the specimen. The ring with the sample was placed in a consolidation cell and screwed tightly to the close-fitting metal jacket on top of the cell. The cell was filled with water and then mounted on a loading frame with an applied seating pressure of 5 kPa and left for 24 hours to allow saturation.

After equilibrium was attained conventional oedometer tests were carried out. A load or stress-increment ratio of unity (12.3, 24.5, 49, 98, 196, 392, and 794 kPa) was adopted. As each stress increment was applied, readings of vertical compression were taken at regular time intervals until movement ceased and then the next stress was then applied. At the end of each test, the soil sample was taken out of the cell and oven dried to obtain its final dry weight, water content, void ratio, and, subsequently, void ratio at each stress increment or level. The variation of void ratio with effective stress was obtained for each sample and plotted on a semi-logarithmic scale. Figure 3 shows a plot of void ratio (e) - log stress curves of representative soil samples used in this study. The compression index C_c for each sample was computed in the stress range 98-794 kPa.

Results and Discussion

Fractal Scaling of PSDs

As already mentioned, the plots of Figures 2(a-d) clearly indicate that a single power law cannot be used to describe the data across the range of measured particle sizes. On the contrary, two power-law functions fit the data points extremely well, indicating two scaling domains for the ranges of particle sizes analyzed. Such multiple linear trends on log

-transformed PSD of soil samples have been observed by other authors [26], [46]. The fractal dimensions for the two scaling regimes were estimated using Equation (2) for each sample and the coefficient of determination of the linear regression, R^2 , was computed. The two power-law domains separated the particle sizes into two class domains, smaller and larger particle sizes, and their respective fractal dimensions were denoted as D_1 and D_2 . Thus, the PSDs can be partitioned into two subsets such that each subset could be a fractal with its own scaling exponent.

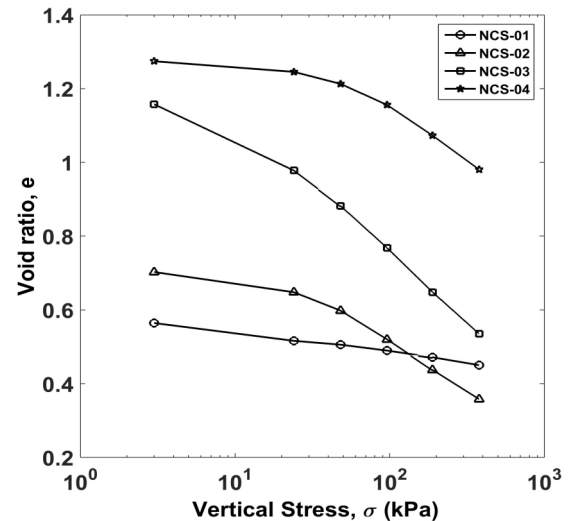


Figure 3. Plot of Void Ratio-Vertical Stress Curves for Representative Soil Samples Obtained from Oedometer Tests

The influence of soil texture in the development of the two power-law scaling domains in PSD was investigated. Figure 4 displays the variation of fractal parameters, D_1 , D_2 , and d_c with the textural parameters (e.g., clay and sand content) of the studied soils. From Figures 4(b,c) it was recognized that both D_1 and D_2 increased with an increase in clay content, while they both decreased weakly with an increase in sand content, as shown in Figures 4(e,f). On the other hand, the values critical size parameter, d_c , which separated the two scaling domains within each soil, decreased with an increase in clay content [Figure 4(a)] and increased with an increase in sand content [Figure 4(d)].

These relationships indicated that the power-law domains characterizing the fractal scaling properties of soil particles analyzed in this study were directly related to important soil textural parameters. This indication provided the opportunity to relate values of D_1 and D_2 to some engineering and physical attributes of soils that are influenced by their texture (PSD), and allowing the use of fractal statistics as a robust technique to characterize PSD and, hence, to estimate these attributes.

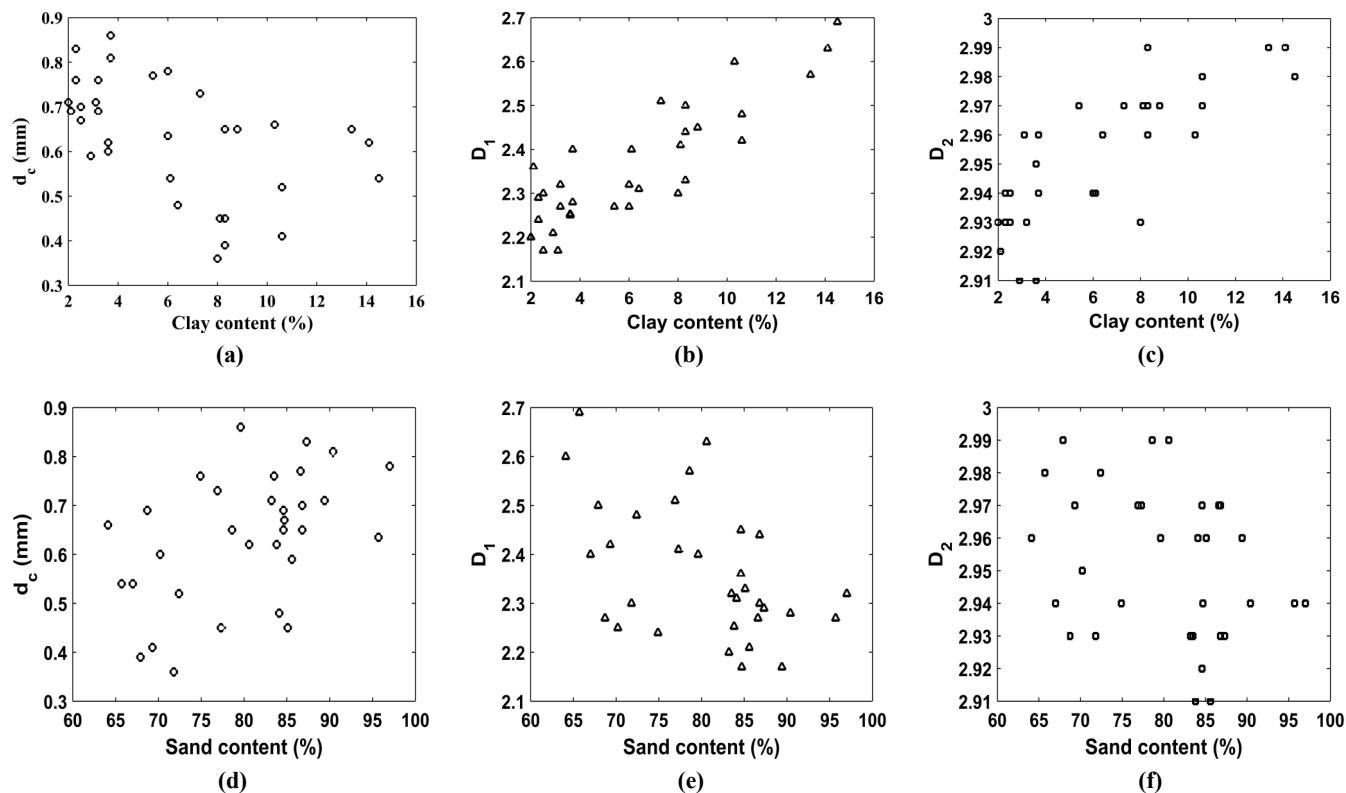


Figure 4. Variation of Fractal Parameters D_1 , D_2 , and d_c with Some Textural Features (sand and clay contents) of the Studied Soils

Figure 5 illustrates the variation of the fractal parameters with the ratio of sand content to clay content (SCR) in each soil sample. While weak variations of D_2 and d_c were recognized with SCR [Figures 5(a,b)], there was an indication of strong variation of D_1 and the ratio D_1/D_2 with SCR [Figures 5(c,d)]. These figures illustrate that D_1 decreased sharply with an increase in SCR within a range of low values of SCR (0-10) up to a value of SCR roughly equal to 10 [Figure 5(c)] after which the D_1 values decreased slowly or were virtually insensitive to increases in SCR values.

The low SCR range is characterized by increasing numbers of smaller clay particles in the sample and gradual filling of the void spaces (space filling) and separating the bigger sand particles. The PSD of the soil in this case had relatively higher values of D_1 . The variation of D_1 with SCR [Figure 5(c)] exhibited a power law-behavior and a model describing this behavior is given by Equation (3):

$$D_1 = 2.2 + 3.2SCR^{-1.25} \quad (3)$$

where, coefficient of determination, R^2 , was equal to 0.79 and the ratio of D_1/D_2 exhibited a similar trend in variation with SCR [Figure 5(d)].

Compressibility and Fractal Parameters

The relationships between compressive index, C_c , which describes the compressibility of the soils, and the fractal parameters D_1 and D_2 , which uniquely describe the distribution of grain sizes within the soils, as well as the critical size, d_c , were analyzed. Figure 6 shows the relationships between the fractal parameters and compression index. C_c showed little or no variation with d_c [Figure 6(a)] and D_2 [Figure 6(b)].

However, the compression index bears a stronger relationship with D_1 [Figure 6(c)], whereby the compression index increases with an increased in the fractal dimension, D_1 . Again, an increase in D_1 was the result of an increase in the amount of smaller clay particles and gradually filled the void spaces separating the relatively bigger sand particles and becoming the load-bearing unit or component of the soil. This enhanced the compressibility of the soil and yielded a higher value of C_c . At low values of D_1 , characterized by dominance of larger sand particles, grain-to-grain contacts were potentially enhanced [21] and, thus, load was transferred mainly through the sand skeleton and yielded reduced compressibility.

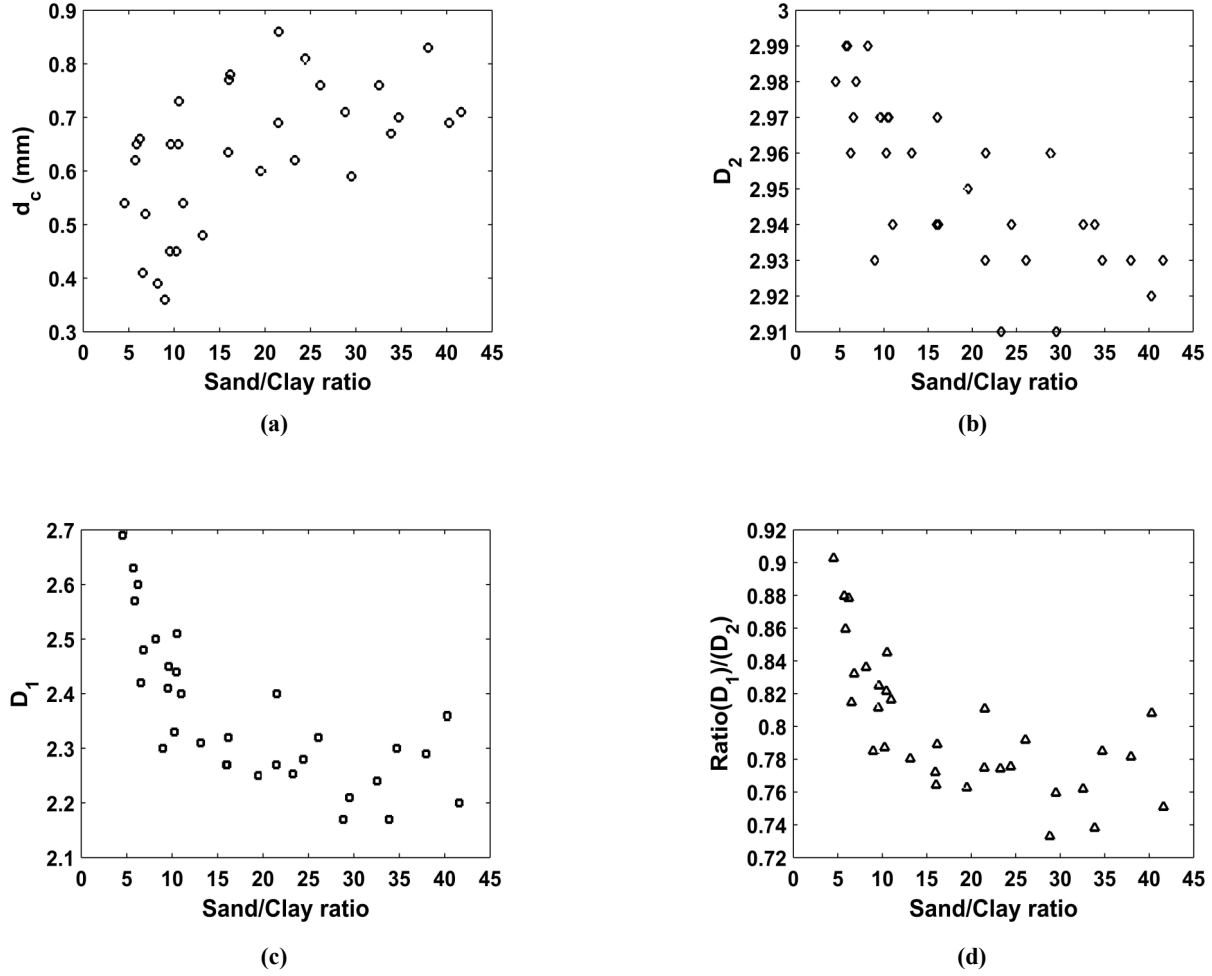


Figure 5. Variation of Fractal Parameters D_1 , D_2 , d_c and D_1/D_2 with the Ratio of Sand to Clay Fractions of the Studied Soils

Model Prediction and Comparisons

Regression models were developed to predict the compression index of soils using only the fractal parameters as input predictor variables and a combination of fractal parameters and index properties as input predictor variables. The predicted C_c values were compared with measured values and their performances were assessed. In addition, outputs of existing regression models, which used standard index properties of soils as input variables, were also compared and assessed. The question of how well the predicted values fit the measured data was addressed using the following criteria: 1) the coefficient of multiple determination, R^2 (equivalent to the square of the correlation coefficient) and 2) the standard error of estimate, S_e (the square root of error variance). The R^2 value was computed using Equation (4) [3], [10]

$$R^2 = \frac{\sum_{i=1}^n (Y_i^{pre} - Y_{ave})^2}{\sum_{i=1}^n (Y_i - Y_{ave})^2} \quad (4)$$

and the standard error of the estimate was computed using Equation (5):

$$S_e = \left[\frac{1}{v} \sum_{i=1}^n (Y_i^{pre} - Y_{ave})^2 \right]^{0.5} \quad (5)$$

where, Y_i is the i th measured or observed value; Y^{pre} is predicted value; Y_{ave} is the mean of the observations; and, v is the degree of freedom (sample size n minus the number of descriptors variables or unknowns).

The coefficient of determination was interpreted as the proportion of variability in the observed response variables that was accounted for by the regression model and, thus,

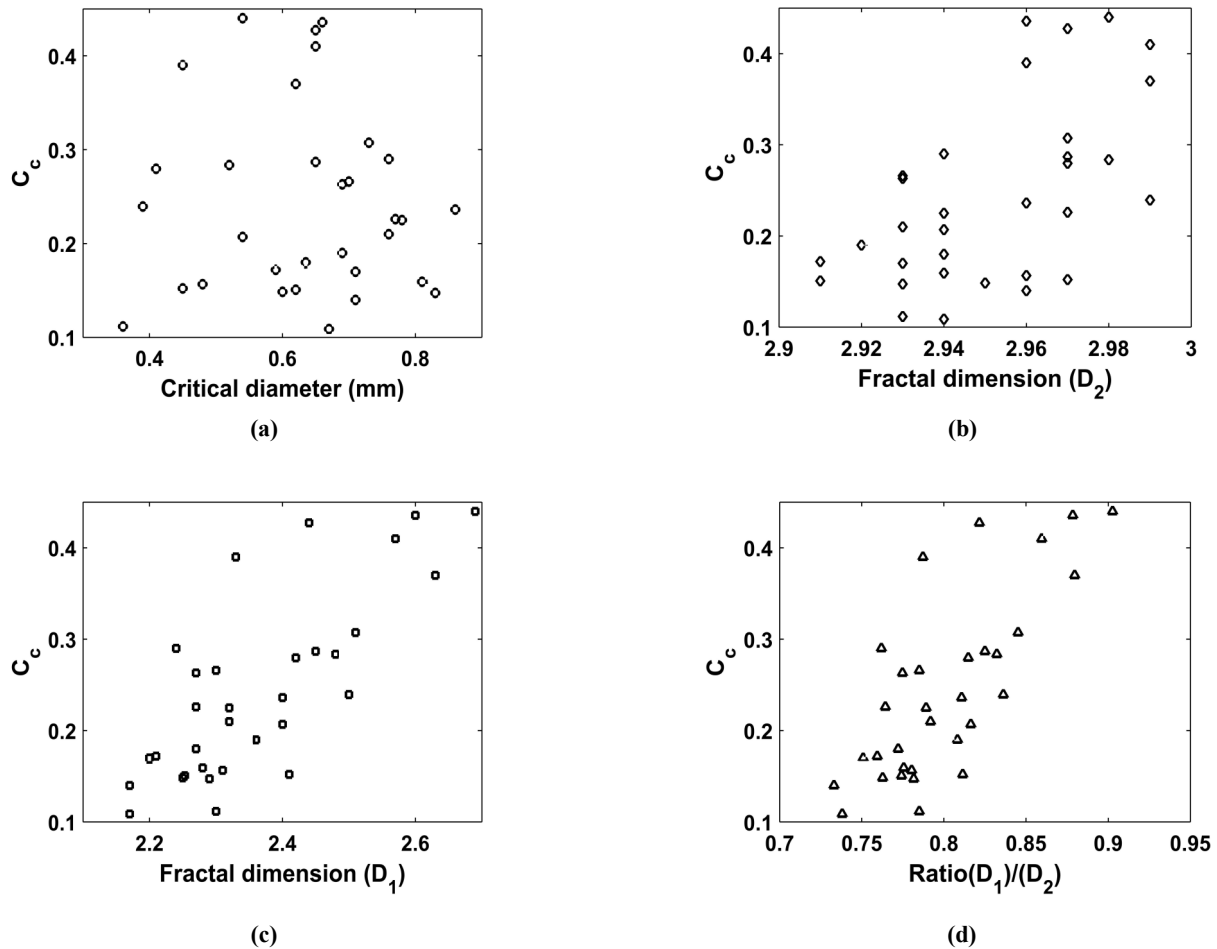


Figure 6. Variation of Fractal Parameters D_1 , D_2 , d_c , and D_1/D_2 with the Compression Index Values of the Studied Soils

served as an indicator of the accuracy of predictions. The magnitude of the standard error of the estimate or the error variance (square root of S_e) on the other hand was a physical indicator of the error between the measured and predicted values.

Figures 7(a-f) illustrate comparisons between the predicted and measured values of C_c for the soils used in this study.

One model group [Figures 7(a,b,c)] used a single soil parameter as input predictor variable; for the other group [Figures 7(d,e,f)], two soil parameters were used as the input. Figure 7(a) shows the relationship between the predicted and measured C_c values using the fractal dimension D_1 as the input. The R^2 value was 0.66 and the error variance, S_e , was equal to 0.05, indicating that D_1 as a predictor variable accounted for 66% of the variability in the measured compressibility values and, thus, provided a reasonably accurate prediction of C_c .

This result demonstrates that accounting for the distribution of particle sizes in soils can be helpful in describing their compressive behavior. Using the liquid limits, w_L , of the soils as a single input parameter, Figure 7(b) shows comparison between predicted and measured C_c values with an R^2 value of 0.32 and an error variance equal to 0.0049. Obviously, this prediction was weaker than the previous prediction using the fractal dimension as the input variable. Further, using the initial void ratio, e_o , of the soils as the sole input variable, the predicted C_c value compared to the measured C_c value is illustrated in Figure 7(c) with R^2 and error variance values of 0.30 and 0.005, respectively. This indicated a weaker prediction model compared to the models with w_L and D_1 as sole input parameters. Thus, in the overall comparative assessment of models using a single predictor variable, the model incorporating D_1 performed better in predicting C_c than those using the index properties e_o or w_L .

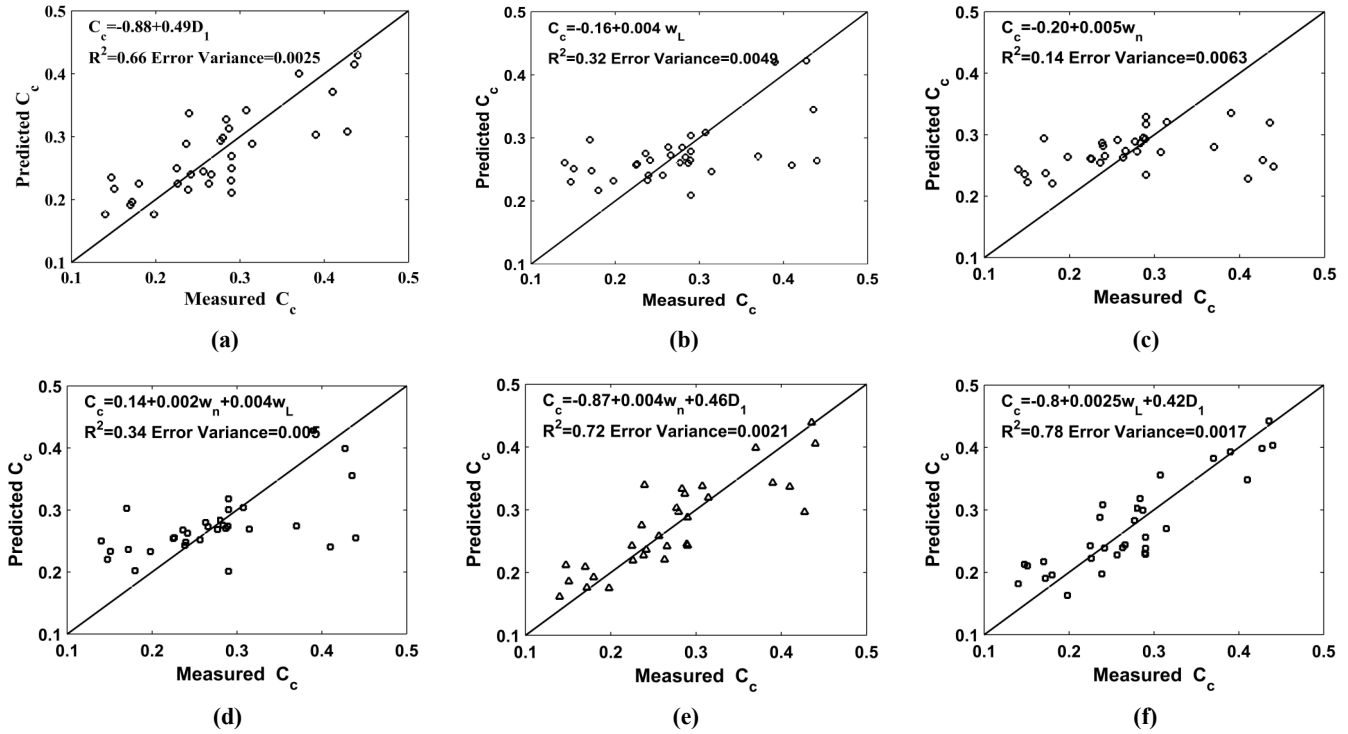


Figure 7. Plot of the Measured Values of the Compression Index (K_{exp}) versus the Computed Values using the Regression Models [the straight lines are lines of perfect equality]

Relationships between predicted and measured C_c values, using two soil index parameters as input predictor variables, are shown in Figures 7(d,e,f). Using e_o and w_L as input variables, Figure 7(d) shows a comparison between predicted and measured C_c values, which gives an R^2 value of 0.44 and an error variance equal to 0.002. As expected, there is an indication of a slight improvement in the prediction performance using the two input parameters compared to when the parameters were used as variables in isolation.

Using e_o and D_1 as input predictor variables significantly improved the prediction of C_c with R^2 equal to 0.75 and an error variance value of 0.002. A slight improvement in performance in prediction of C_c was recognized using w_L and D_1 as input variables with R^2 and error variance values of 0.78 and 0.0017, respectively. The above results also indicated that using D_1 as a single variable provided a better model performance than using e_o and w_L , underscoring the relevance of the fractal characteristics of the PSD in assessing the compressibility of soils. Comparative assessment of other published models are described in Table 1 and their respective performances are illustrated in Figure 8.

It can be inferred that Model 1, Model 5, and Model 6, all of which incorporate the fractal dimension D_1 as input variable in the regression models, provided better prediction

performance than those with standard soil index properties as input variables.

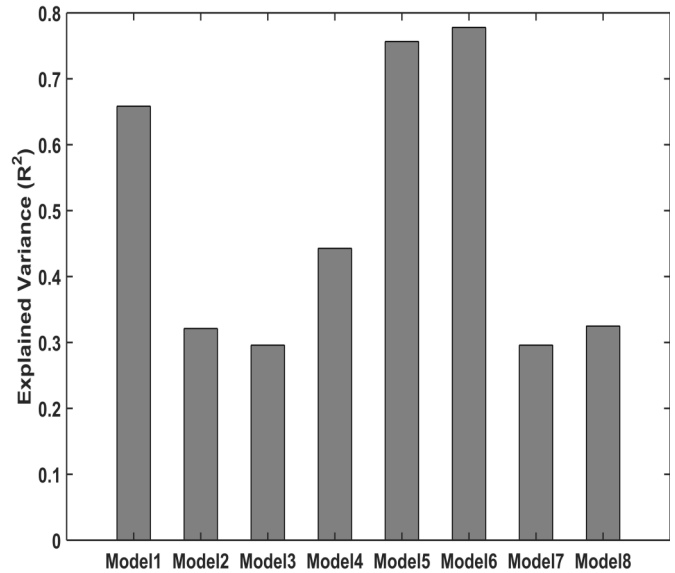


Figure 8. A Graphical Illustration Comparing the Performance of Models used in This Study using the Coefficient of Determination (R^2) as the Performance Criterion [models are referenced in Table 1]

Table 1. Compression Index Equation used in This Study and Other Sources

Function type	Equation	Reference (Model)
$C_c = a.w_n + b$	$C_c = 0.01.w_n - 0.05$ $C_c = -0.2 + 0.005w_n$	Azzouz et al. [10] (Model 7) This current study (Model 2)
$C_c = a.w_L + b$	$C_c = -0.16 + 0.004w_L$	This current study (Model 3) Mayne [47]
$C_c = a.D_1 + b$	$C_c = -0.88 + 0.49D_1$	This current study (Model 1)
$C_c = a.w_n + b.w_L + c$	$C_c = 0.14 + 0.002w_n + 0.004w_L$ $C_c = 0.009w_n + 0.005w_L$	This current study (Model 4) Koppula [48] (Model 8)
$C_c = a.w_n + b.D_1 + c$	$C_c = -0.87 + 0.004w_n + 0.46D_1$	This current study (Model 5)
$C_c = a.w_L + b.D_1 + c$	$C_c = -0.8 + 0.025w_L + 0.42D_1$	This current study (Model 6)

Conclusions

Empirical equations were developed to predict the compression index of soils obtained from varying geological terrains using regression models. The predictor variables used in the regression models to predict the compression index of the soils were easily obtained from particle-size distribution analyses and standard index tests. Particle-size distributions of the soils were analyzed using fractal scaling methods and uniquely characterized by their fractal dimensions. Two fractal scaling domains (upper and lower) with two fractal dimensions and a cut-off or critical size were determined for each soil. The fractal dimensions of the lower scaling domain of the soils bear strong relationships with the measured compression index values; an indication that soils with higher fractal dimensions had higher compressibility. Incorporating the fractal dimensions as predictor variables in the regression models significantly improved the prediction of the compression index.

Models to predict the compression index of soils using only the fractal dimension as input predictor variables and a combination of fractal dimensions and standard index properties as input predictor variables were also developed. Models developed in other studies were compared with models developed in this study and their predictive performances were assessed. For single and multiple predictor variables, predictive empirical models developed in this study performed better than those in other studies. Using the fractal dimension as a single predictor variable provided a better predictive model than using both initial void ratio, e_o , and liquid limit, w_L , as input variables, as in other studies. The empirical models developed in this study were suggested as useful alternatives for estimating the compression index of soils. In certain circumstances, it may also be useful in giving initial information about soil compressibility in a field environment based on an analysis of particle-size distribution.

References

- [1] Mitchell, J. K., & Soga, K. (2005). *Fundamentals of soil behavior*. (3rd ed.). Wiley.
- [2] Trani, L.D.O., & Indraratna, B. (2010). The use of particle size distribution by surface area method in predicting the saturated hydraulic conductivity of graded granular soils. *Géotechnique*, 60(12), 957-962.
- [3] Boadu, F. K. (2000). Hydraulic conductivity of soils from grain size distribution: New models. *Journal of Geotechnical and Geoenvironmental Engineering*, 126(8), 739-745.
- [4] Casagli, N., Ermini, L., & Rosati, G. (2003). Determining grain size distribution of material composing landslide dams in the Northern Apennine: sampling and processing methods. *Engineering Geology*, 69, 83-97.
- [5] Lade, P. V., & Yamamuro, J. A. (1997). Effects of Non-Plastic Fines on Static Liquefaction of Sands, *Canadian Geotechnical Journal*, 34(6), 918-928.
- [6] Thevanayagam, S., Ravishankar, K., & Mohan, S. (1997). Effects of Fines on Monotonic Undrained Shear Strength of Sandy Soils. *Geotechnical Testing Journal*, 20(1), 394-406.
- [7] Yilmaz, Y., Mollamahmutoglu, M., Ozaydin, V., & Kayabali, K. (2008). Experimental Investigation of the Effect of Grading Characteristics on the Liquefaction Resistance of Various Graded Sands. *Engineering Geology Journal*, 100, 91-100.
- [8] Belkhatir, M., Arab, A., Della, N., Missoum, H., & Schanz, T. (2012). Experimental Study of Undrained Shear Strength of Silty Sand: Effect of Fines and Gradation. *Geotechnical and Geological Engineering*,

- 30(8), 1103-1118
- [9] Skempton, A. W. (1944). Notes on the compressibility of clays. *Quarterly Journal of the Geological Society of London*, 100(2), 119-135.
- [10] Azzouz, A. S., Krizek, R. J., & Corotis, R. B. (1976). Regression analysis of soil compressibility. *Soils and Foundations*, 16(2), 19-29.
- [11] Rendon-Herrero, O. (1980). Universal compression index equation. *Journal of the Geotechnical Engineering Division*, 106(11), 1179-1199.
- [12] Di Maio, C., Santoli, L., & Schiavone, P. (2004). Volume change behavior of clays: the influence of mineral composition, pore fluid composition and stress state. *Mechanics of Materials*, 36(6), 435-451.
- [13] Burland, J. B. (1990). On the compressibility and shear strength of natural Clays. *Geotechnique*, 40(3), 329-378.
- [14] Nagaraj, T. S., & Srinivasa Murthy, B. R. (1983). Rationalization of Skempton's compressibility equation. *Géotechnique*, 33(4), 433-443.
- [15] Meade, R. H. (1964). Removal of water and rearrangement of particles during the compaction of clayey sediments - Review USGS Professional Paper: 497-B
- [16] Fukue, M., Okusa, S., & Nakamura, T. (1986). Consolidation of sand-clay mixtures. Consolidation of Soils: Testing and Evaluation, ASTM STP892, R. N. Yong and F. C. Townsend, eds., American Society for Testing and Materials, Philadelphia, 627-641.
- [17] Martins, F. B., Bressani, L. A., Coop, M. R., & Bica, A. V. D. (2001). Some aspects of the compressibility behaviour of a clayey sand. *Canadian Geotechnical Journal*, 38(6), 1177-1186.
- [18] Monkul, M. M., & Ozden, G. (2007). Compressional behavior of clayey sand and transition fines content. *Engineering Geology*, 89, 195-205.
- [19] Belkhatir, M., Arab, A., Schanz, T., Missoum, H., & Della, N. (2011). Laboratory study on the liquefaction resistance of sand-silt mixtures: effect of grading characteristics. *Granular Matter*, 13(5), 599-609
- [20] Chuhan, F. A., Kjeldstad, A., Bjørlykke, K., & Høeg, K. (2003) Experimental compression of loose sands: relevance to porosity reduction during burial in sedimentary basins. *Canadian Geotechnical Journal*, 40 (5), 995-1011.
- [21] Minh, N. H., & Cheng, Y. P. (2013). A DEM investigation of the effect of particle-size distribution on one-dimensional compression. *Géotechnique*, 63(1), 44-53
- [22] Moore, C. A., & Donaldson, C. F. (1995). Quantifying soil micro-structure using fractals, *Geotechnique*, 45, 105-106.
- [23] Ahmadi, A., Neyshabouri, M., Rouhipour, H. A., & Asadi, H. (2011). Fractal dimension of soil aggregates as an index of soil erodibility. *Journal of Hydrology*, 400, 305-311.
- [24] Korvin, G. (1992). *Fractal Models in the Earth Sciences*. (2nd ed.) Elsevier.
- [25] Crovelli, R. A., & Barton, C. C. (1995). Fractals and the Pareto distribution applied to petroleum accumulation-size distributions. In Barton, C.C. & La Pointe, P.R. (Eds.) *Fractals in Petroleum Geology and Earth Processes*. (1st eds., pp. 59-72). New York, Plenum.
- [26] Wu, Q., Borkovec, M., & Sticher, H. (1993). On the particle-size distributions in soils. *Soil Science Society of America Journal*, 57, 883-890.
- [27] ASTM D2216-10 (2010) Standard test methods for laboratory determination of water (Moisture) content of soil and rock by Mass. doi:10.1520/D2216-1
- [28] ASTM D4318-10 (2010) Standard test methods for liquid limit, plastic limit, and plasticity index of soils. doi:10.1520/D4318-1
- [29] ASTM D422 2007 Standard Test Method for Particle-Size Analysis of Soils, doi: 10.1520/D0422-63R07E02
- [30] Peyton, R. L., Gantzer, C. J., Anderson, S. H., Haeflner, B. A., & Pfeifer, P. (1994). Fractal dimension to describe soil macropore structure using X-ray computed tomography. *Water Resources Research*, 30,691-700
- [31] Tyler, S. W., & Wheatcraft, S. W. (1989). Application of fractal mathematics to soil water retention estimation. *Soil Science Society of America Journal*, 53, 987-996.
- [32] Brakensiek, D. L., Rawls, W. J., Logsdon, S. D., & Edwards, W. M. (1992). Fractal description of microporosity. *Soil Science Society of America Journal*, 56: 1721-1723.
- [33] Friesen, W. I., & Mikula, R. J. (1987). Fractal dimensions of coal particles. *Journal of Colloid and Interface Science*, 120, 263-271.
- [34] Tyler, S. W., & Wheatcraft, S. W. (1992). Fractal scaling of soil particle size distribution: analysis and limitations. *Soil Science Society of America Journal*, 56,362-369.
- [35] Mandelbrot, B. B. (1982). *The Fractal Geometry of Nature*. (2nd ed). Freeman.
- [36] Turcotte, D. L. (1986). Fractals and fragmentation. *Journal of Geophysical Research*, 91, 1921-1926
- [37] Turcotte, D.L., (1992). *Fractals and Chaos in Geology and Geophysics*. (2nd ed.) Cambridge University Press.
- [38] Boadu, F. K., & Long, L. T. (1994a). Statistical Distribution of Natural Fractures and the Possible Physical Generating Mechanism. *Pure and Applied Geophysics*, 142, 273-293.

-
- [39] Boadu, F. K., & Long, L. T. (1994b). The Fractal Character of Fracture Spacing and RQD. *International Journal of Rock Mechanics and Mining Science and Geomechanics Abstracts*, 31, 127-134.
 - [40] Millán, H., González-Posada, M., Aguilar, M., Domínguez, J., & Céspedes, L. (2003). On the fractal scaling of soil data. Particle-size distributions. *Geoderma*, 117, 117-128.
 - [41] Prosperini, N., & Perugini, D. (2008). Particle size distributions of some soils from the Umbria Region (Italy): Fractal analysis and numerical modeling. *Geoderma*, 145, 185-195.
 - [42] Seber, G. A. F., & Wild, C. J. (1989). *Nonlinear Regression*, Wiley, (1st ed.). New York.
 - [43] Walkowiak, R., & Kala, R. (2000). Two-phase Non-linear Regression with Smooth Transmission. Communications in Statistics. *Simulations and Computation*, 29(2), 385-397.
 - [44] Lerman, P. M. (1974). Fitting segmented egression models by grid search, *Journal of Applied Statistics*, 29(1), 77-84.
 - [45] ASTM D2435/D2435 M-11 (2011). Standard test methods for one- dimensional consolidation properties of soils using incremental loading. doi:10.1520/D2435_D2435M-1
 - [46] Bittelli, M., Campbell, G., & Flury, M. (1999). [Characterization of particle size distribution in soils using a fragmentation model. *Soil Science Society of America Journal*, 63, 782-788.
 - [47] Mayne, P. W. (1980). Cam-clay predictions of undrained strength. *Journal of the Geotechnical Engineering Division, ASCE*, 106(11), 1219-1242.
 - [48] Koppula, S. D. (1981). Statistical estimation of compression index. *Geotechnical Testing Journal*, 4 (2), 68-73.

Biography

FRED KOFI BOADU is an associate professor in civil engineering at Duke University. His research interests focus on geo-environmental engineering, engineering, and petrophysical characterization of soils and rocks using seismic and electrical techniques, environmental and engineering geophysics and applications of fractal and neural network concepts to geophysical and engineering problems. Dr. Boadu may be reached at boadu@duke.edu

CHARACTERIZATION OF WEATHERED LIMESTONE USING THERMAL IMAGING AND SPLIT TENSION TESTING

John Nuzzkowski, University of North Florida; Mason McGough, University of Florida; Lexi Kimes, University of North Florida; Alan Harris, University of North Florida; Nick Hudyma, University of North Florida; Faris Malhas, Central Connecticut State University

Abstract

Weathering of rock has a negative influence on both physical and engineering properties. For the design of infrastructure assets within weathered rock, it is imperative to determine the influence of weathering. Numerous techniques, both destructive and non-destructive, have been employed to assess the influence of weathering on rock properties. In this study, thermal characterization of weathered limestone specimens was conducted and results were evaluated in light of split tension tests. Thermal characterization consisted of heating the specimens to 200 degrees Celsius and recording the cooling of the specimens for 45 minutes with a thermal camera. A thermal time constant was determined for each specimen using the lumped capacitance method. The specimens were categorized into one of three weathering classes. Fresh or un-weathered specimens had the highest unit weight, highest average thermal time constant, and highest tensile strength.

Differentially weathered specimens had intermediate unit weights, intermediate average thermal time constants, and, on average, intermediate tensile strengths. The highly weathered specimens had the lowest unit weights and lowest thermal time constants. The fresh and highly weathered specimens exhibited classic tensile fracture patterns and the differentially weathered specimens exhibited surface activated fractures, which reflected their internal structures. There was a clear distinction between the average specimen thermal time constant and weathering classification. There was no clear distinction between tensile strength and weathering state. Results indicate that thermal characterization is a promising technique for the assessment of the weathering state of limestone laboratory specimens.

Introduction

Weathering, or the breakdown of rock into soil through biological, chemical, and physical processes, has a detrimental effect on the physical and engineering properties of rock. Weathered rock profiles may contain materials of a wide range of strengths [1]. This presents difficulties for design engineers, who must assess strength and stiffness

properties of weathered materials, which support large and important infrastructure assets.

Given the potential variability in physical and engineering properties of weathered rock, it would be advantageous to perform strength tests on all specimens recovered from sampling activities conducted within weathered rock masses. It would be cost prohibitive to test all specimens recovered. Additionally, as shown by Hudyma and Hiltunen [2], performing careful sampling procedures on weathered rock may not provide specimens of appropriate size for testing. These constraints have forced engineers to rely on qualitative descriptions, physical measures, and index tests to provide an indication of the weathering state of rock, which may then be correlated to strength and stiffness. However, relying on qualitative visual descriptions to assess the weathering state of rock specimens has been shown to be troublesome. For the most part, individuals can successfully identify fresh (relatively un-weathered) specimens and highly weathered specimens, but have difficulty assessing intermediate stages of weathering in rock specimens [3], [4]. This demonstrates that assessment of the weathering state of rock should be based on a quantitative measure. Weathering manifests itself in rock specimens as an increase in porosity [5] and potentially an increase in the abundance of microfractures [1].

Techniques for assessing weathering grades of specimens are expected to be simple, inexpensive, and efficient [6]. Ideally, porosity measurements would be used to assess weathering, but such measurements are not easy to perform and require specialized equipment, which is not readily available in most commercial laboratories. Numerous studies [7-11] have been dedicated to identifying laboratory techniques that are sensitive to variations in porosity and assess weathering. The techniques can be categorized as either destructive or non-destructive. Non-destructive techniques include unit weight, dynamic characterization, Schmidt hammer testing, and textural characterization. Destructive techniques include indirect tensile testing, point load testing, and unconfined compression testing. The noted studies have shown that there is some degree of overlap between weathering grades and measured properties.

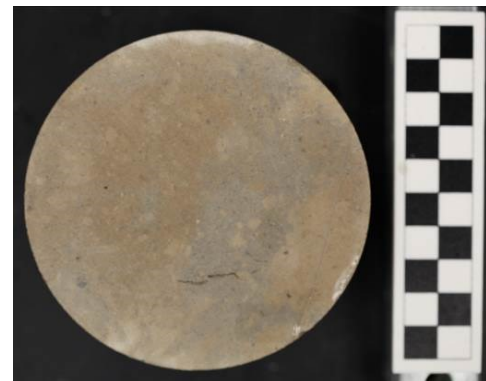
One yet unexplored technique to assess weathering is thermal imaging. Thermal imaging is a common assessment tool for detecting objects or areas with different heat signatures. This technology has been applied to laboratory investigations of soil thermal diffusivity [12] and mechanical damage in granite specimens during loading [13], but the technique is not commonly used. In this paper, the authors present results from thermal decay measurements and split tension testing conducted on weathered limestone specimens obtained from various locations in Florida. The thermal decay measurements were used to assess the weathering of the specimens. Split tension specimens are ideal for thermal imaging applications because they are relatively thin; thickness is approximately equal to the radius of the specimens. More importantly, split tension tests were used for the design of drilled shafts in Florida.

Specimen Selection and Preparation

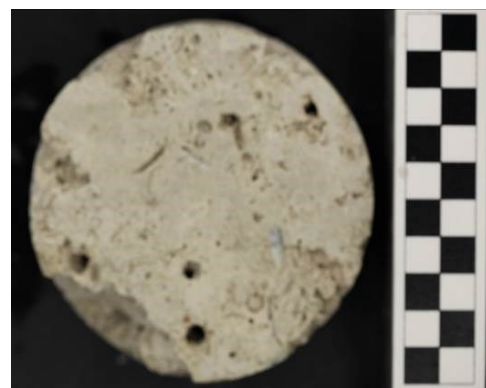
Core samples were obtained from a number of locations in southern Florida. As a general rule, when dealing with sampling of weathered rock, using large diameter core barrels and using either double or triple core barrels increases core recovery. For transportation-based projects in Florida, it is recommended to obtain 101 mm diameter cores using a minimum of a double core barrel [14]. The core samples for this study had diameters on the order of 101 mm; un-weathered or fresh specimens had full diameters, whereas highly weathered specimens had diameters less than 101 mm. Once sections of core samples were identified, specimens were cut using a wet diamond saw. The specimens were cut to meet the specifications of split tension (Brazilian test) requirements, based in standards set forth by the International Society for Rock Mechanics [15]. Ten different limestone specimens were selected, based solely on surface appearance, for this study. The specimens were then identified as fresh, highly weathered, or differentially weathered. Photographs of specimens from the three weathering states are shown in Figures 1(a-c). Fresh or un-weathered specimens had smooth saw-cut surfaces, did not have any defects along their circumference, contained very few surface cracks or pits, and had the appearance of concrete. Figure 1(a) is a photograph of the only un-weathered specimen used in this study.

Highly weathered specimens had almost the opposite appearance of un-weathered specimens. These specimens had rough saw-cut surfaces, had uneven circumferential surfaces, and contained surface pitting. An example of a highly weathered specimen used in this study is shown in Figure 1(b). The third specimen category included differentially weathered specimens. The presence of differential weathering for geotechnical characterization of limestone speci-

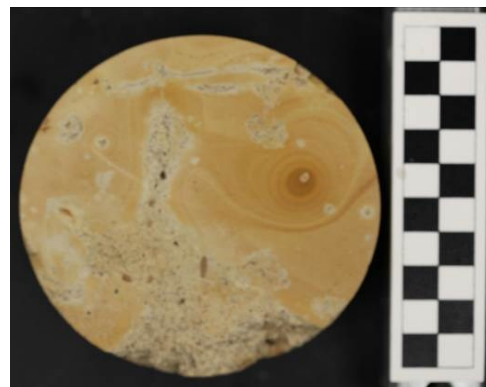
mens from Florida was first described by Sarno [16]. These specimens contained surface features (or lack thereof) which were not as pronounced as those in either the fresh or highly weathered state. What made these specimens difficult to assess was the presence of distinct zones of materials with different weathering states. This is clearly shown in Figure 1(c).



(a) Fresh Specimen H3



(b) Highly Weathered Specimen H20



(c) Differentially Weathered Specimen H18

Figure 1. Photographs of Fresh, Highly Weathered, and Differentially Weathered Specimens

Thermal Characterization

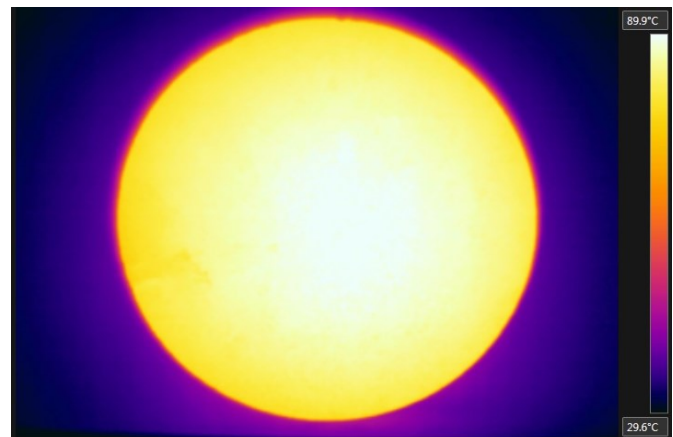
The first characterization activity conducted was the non-destructive thermal characterization. Specimens were initially heated to 200 degrees Celsius for 24 hours in a standard laboratory oven. After 24 hours, the specimens were removed from the oven, placed on a laboratory bench, and allowed to cool in a room-temperature laboratory. During cooling, the thermal decay was monitored for 45 minutes using a thermal imaging camera. Thermal images, with a resolution of 240 x 320 pixels, were captured at a rate of one image every twelve seconds. Each pixel within the thermal image can be thought of as an individual infrared thermometer. Thermal images from fresh, differentially weathered, and highly weathered specimens are shown in Figures 2(a-c). The thermal image of the fresh specimen was as expected. The specimen was composed of material in a single weathering state and could be considered somewhat homogeneous. As the specimen cooled, heat loss was highest at the edges because of exposure to cool temperatures. The heat signature during cooling was similar to a “bulls-eye” with the coolest temperatures on the outside and warmest temperatures in the center. The heat signature from a fresh specimen is shown in Figure 2(a). The highly weathered specimen was comprised of a relatively homogeneous material and the same “bulls-eye” heat signature was present. The relatively large pits on the specimen surface allowed heat to escape from the interior of the specimen so they appeared as localized hot spots. A thermal image of a highly weathered specimen is shown in Figure 2(b).

The differentially weathered specimens, which contained zones of differentially weathered materials, showed a unique temperature signature. The highly weathered zones released heat faster than less weathered zones. Specimen cooling did not take place at a consistent rate throughout the specimen, which resulted in an irregular heat signature. As shown in Figure 2(c), the heat signature of the differentially weathered specimen contained an irregularly shaped hotspot, which was different than the roughly circular hot spot shown in the thermal image of the fresh and highly weathered specimens.

Heat Transfer Model Theory

To assess the thermal characteristics of the specimens, the lumped capacitance method of Equation (1) was used [17]. This method incorporated the surface area, A_s , mass density, ρ , volume, V , specific heat, c , and convection coefficient, h , of the surrounding environment on the solid. The lumped capacitance method assumes the solid was at a uniform temperature throughout and that radiation heat transfer was neg-

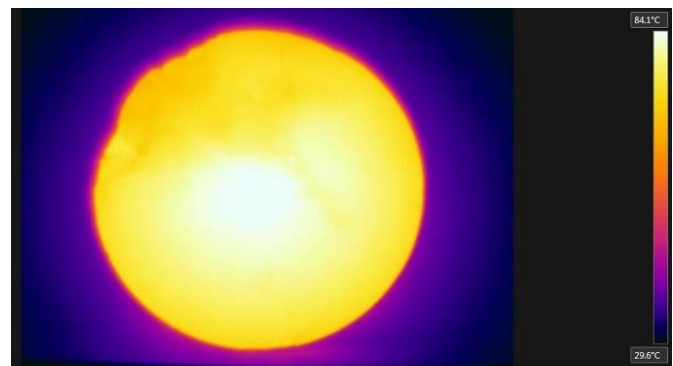
ligible due to the convection heat transfer being much greater than the radiation heat transfer.



(a) Fresh Specimen H3



(b) Highly Weathered Specimen H20



(c) Differentially Weathered Specimen H18

Figure 2. Video Capture of Thermal Images for Fresh, Highly Weathered, and Differentially Weathered Specimens during Cooling

Substitution of $dT = d(T - T_\infty)$ into Equation (1), followed by rearrangement and integration, resulted in Equation (2). Taking the exponential of both sides of Equation (2) and subsequent rearrangement, resulted in Equation (3). Defining θ from Equation (4) as the dimensionless temperature response, and τ from Equation (5) as the thermal time constant, and substitution of those equations into Equation (3), resulted in Equation (6). The thermal time constant is a measure of how fast the temperature of an object reacts to its thermal environment. A solid with a larger thermal time constant, τ , will take longer to reach thermal equilibrium than a solid with a smaller thermal time constant.

$$hA_s(T - T_\infty) = \rho Vc \frac{dT}{dt} \quad (1)$$

$$-\frac{hA_s}{\rho Vc} t = \ln \left(\frac{T(t) - T_\infty}{T_i - T_\infty} \right) \quad (2)$$

$$\frac{T(t) - T_\infty}{T_i - T_\infty} = e^{-\frac{hA_s}{\rho Vc} t} \quad (3)$$

$$\theta(t) = \frac{T(t) - T_\infty}{T_i - T_\infty} \quad (4)$$

$$\tau = \frac{\rho Vc}{hA_s} \quad (5)$$

$$\theta(t) = e^{-t/\tau} \quad (6)$$

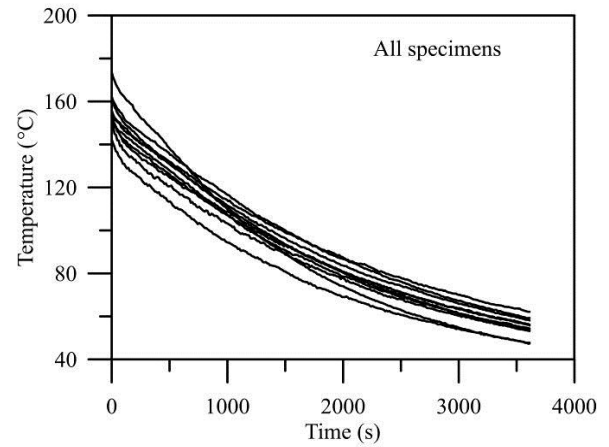
The validity of the lumped capacitance model occurs when the dimensionless Biot, Bi , number given in Equation (7) is less than 0.1 [17]. The Biot number is the ratio of the resistance to conduction heat transfer to the resistance of convection heat transfer. When $Bi \ll 1$, the resistance to conduction heat transfer across the solid is much less than the resistance to convection heat transfer at the fluid boundary. The investigated solid can then be assumed to be at a uniform temperature at each time instance.

$$Bi = \frac{R_{cond}}{R_{conv}} = \frac{L/k}{1/h} \quad (7)$$

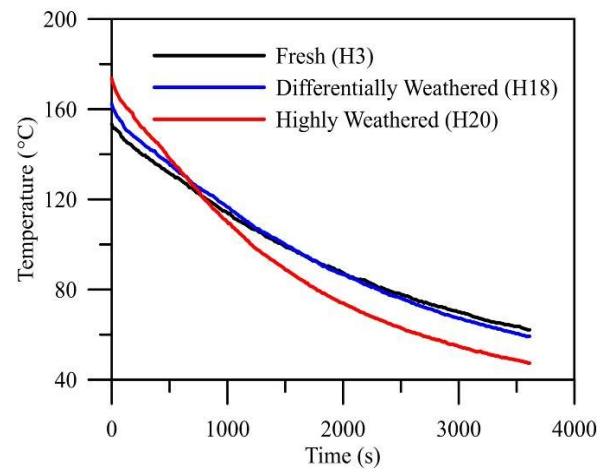
Results

Figure 3 shows the thermal decay profile of the center pixel for all ten specimens. All thermal decay profiles exhibited the expected exponential decay of temperature with time. Looking closely at the thermal decay profiles for the selected fresh specimen, H3, the differentially weathered specimen, H18, and the highly weathered specimen, H20, several interesting observations were made. The temperature at time zero (y-intercept) was misleading. The time zero temperature must be less than 200 degrees Celsius be-

cause the specimens were removed from the oven in order to record the thermal decay. The fresh specimen should have the highest starting temperature and the highly weathered specimen should have the lowest starting temperature. Unfortunately, the time it took to remove the specimens from the oven and center them beneath the thermal camera, focus the camera, and begin recording was not the same.



(a) Thermal Decay Profiles for all Specimens



(b) Thermal Decay Profiles for Select Specimens

Figure 3. Thermal Decay Profiles

The rate of thermal decay was as expected. The highly weathered specimen lost heat at the fastest rate and had the lowest temperature at the end of the measurement time. The fresh specimen lost heat at the slowest rate and had the highest temperature at the end of the measurement time. The heat loss rate and end temperature of the differentially weathered specimen was between the fresh and highly weathered specimens. The differences in the decay profiles indicated that further investigation was warranted.

The heat transfer model of Equation (6) was fitted to the temperature response of Equation (4) during cooling of each pixel using non-linear least squares to iterate for the thermal time constant, τ , of each pixel. Figure 4 shows the center pixel's dimensionless temperature profile and the fitted heat transfer model curve. The curve fit of the heat transfer model represented the collected temperature data well. It should be noted that the Biot number for each specimen, as a whole, was approximately 0.6, but the Biot number for each pixel was approximately 0.005. Therefore, the lumped capacitance method was only valid for each pixel, since the $Bi < 0.1$.

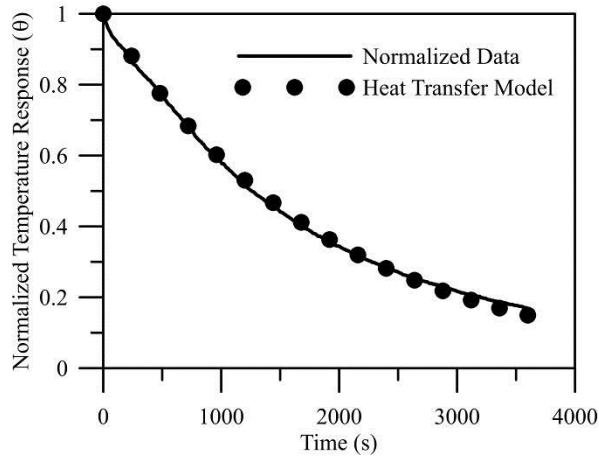
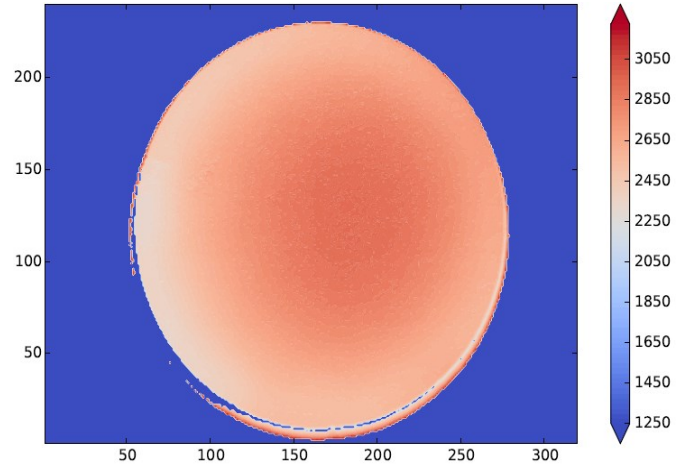


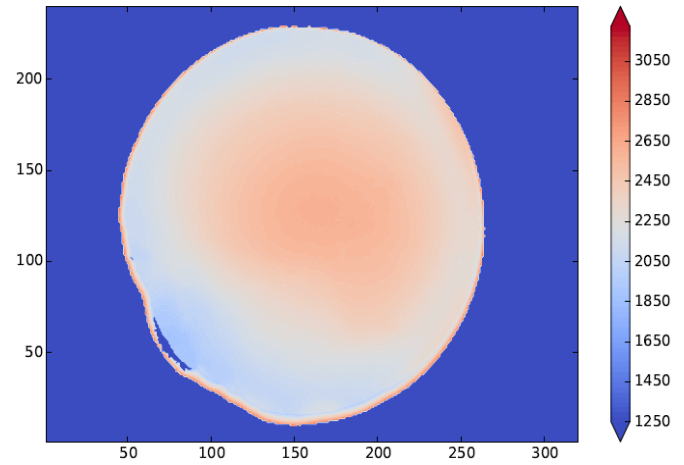
Figure 4. Heat Transfer Model and Temperature Profile for Specimen H20

Figure 5 shows a series of contour plots of the thermal time constants computed for each pixel of the fresh specimen, H3, the differentially weathered specimen, H18, and the highly weathered specimen, H20. To remove the influence of the pixel locations outside the specimen, the average and standard deviation of the thermal time constant was calculated only from the thermal time constants with values below 3000 s. Any pixel location with a thermal time constant less than or greater than three standard deviations from the average thermal time constant of that specimen was given a thermal time constant of zero. The highly weathered specimen had the lowest thermal time constants and the fresh specimen had the highest thermal time constants. Visually, the differentially weathered specimen had thermal time constants between the fresh and highly weathered specimens.

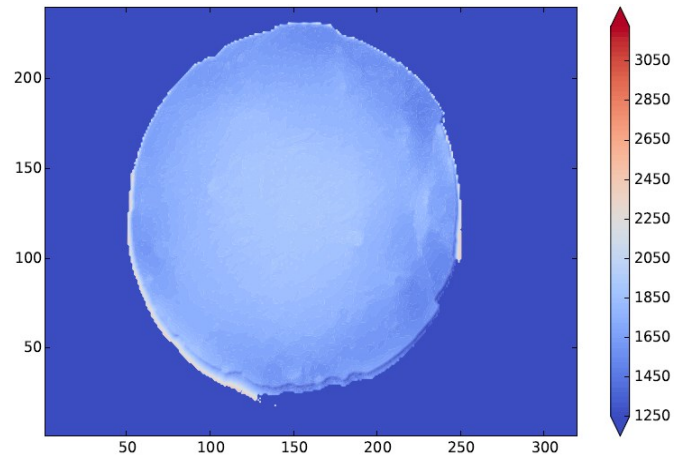
The frequency distribution of the thermal time constants for each specimen were analyzed using histograms. Figure 6 shows the frequency distribution of thermal time constants for specimens H3, H18, and H20. The summation of all the frequencies did not result in 100%, because the pixels outside the rock specimen were removed from the figure.



(a) Fresh Specimen H3



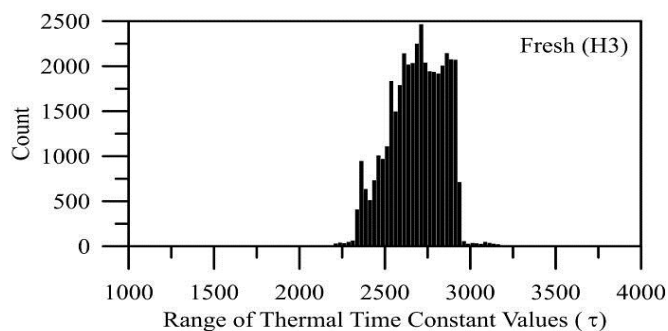
(b) Differentially Weathered Specimen H18



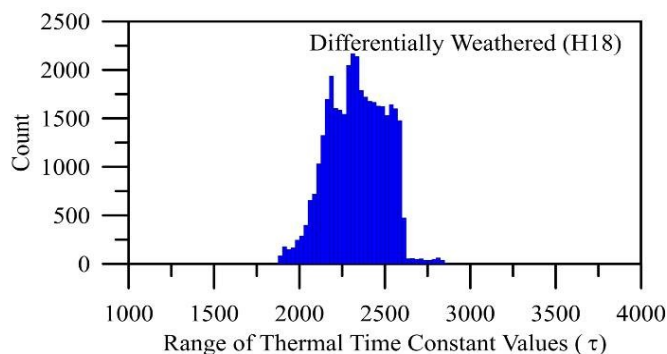
(c) Highly Weathered Specimen H20

Figure 5. Contour Plot of Thermal Time Constants from the Heat Transfer Model

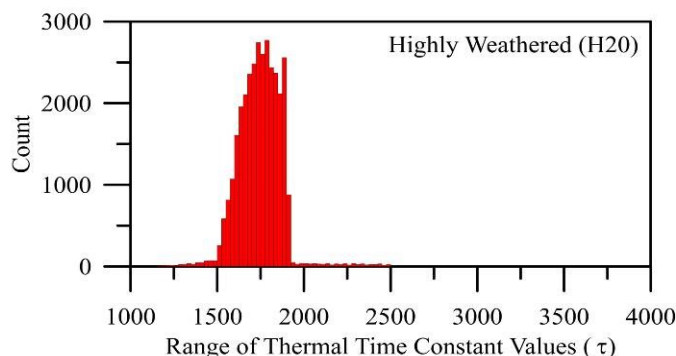
The figure shows that the thermal time constants ranged from ~1500-1800 seconds, 2000-2600 seconds, and 2300-3000 seconds for specimens H20, H18, and H3, respectively. There was overlap within the ranges of thermal time constants. The most frequent thermal time constant was highest for the fresh specimen and lowest for the highly weathered specimen. The histograms showed that the smallest range of thermal time constants was for the highly weathered specimen, which indicated that it was composed of the most homogeneous material.



(a) Fresh Specimen (H3)



(b) Differentially Weathered Specimen H18



(c) Highly Weathered Specimen H20

Figure 6. Frequency of Thermal Time Constants from the Heat Transfer Model

A lower thermal time constant should indicate greater porosity and, therefore, greater weathering. Table 1 shows the number of analyzed pixels, the average, standard deviation, and coefficient of variation (CV) for thermal time constants calculated from the frequency distribution for each specimen. The order of the least to the most weathered specimens, as indicated by the average thermal time constants from the heat transfer model, was H20, H13, H19, H15, H7, H18, H14, H5, H2, and H3. As expected, and shown in Table 1, the specimens with a lower average thermal time constant had a higher CV. It should be expected that specimens with greater weathering should have a larger range of thermal time constants over the surface of the specimen.

Table 1. Heat Transfer Model Data

Weathering classification	Specimen	Number of pixels	Thermal time constant		
			Average (s)	Std. Dev. (s)	CV (%)
Fresh	H3	40245	2679	163	6.1
Differentially weathered	H2	40318	2396	136	5.7
	H5	37865	2363	162	6.8
	H14	40277	2358	127	5.4
	H18	37614	2337	173	7.4
Highly weathered	H7	40374	2135	184	8.6
	H13	34101	1962	168	8.6
	H15	37987	2125	175	8.3
	H19	37432	2076	177	8.5
	H20	35154	1793	235	13.1

Since every pixel in the thermal image provided data to the average thermal time constant calculation, descriptive statistics provided some added benefit for characterizing weathered limestone specimens. For the limited data set, there was no overlap between the coefficient of variation (COV) of the thermal time constant for highly weathered specimens and differentially weathered specimens. The average coefficient of variation for the highly weathered specimens was 9.4 and the average coefficient of variation for the differentially weathered specimens was 6.3. However, there was overlap between the coefficient of variation of the thermal time constant values for fresh and differentially weathered specimens. There was also overlap between all three categories of specimens if the standard deviation of the thermal time constant values were used.

Split Tension Characterization

Since thermal characterization to assess weathering states is a new technique, it was decided to incorporate split tension testing to validate the results. Split tension (Brazilian) testing is a common characterization test in Florida, because drilled shaft design in limestone requires both unconfined

compression testing and split tension testing. Split tension characterization consists not only of tensile strength assessment but also qualitative assessment of failure modes. Tensile strength is a function of weathering. The greater the degree of weathering, the lower the tensile strength. The reduction in strength can be attributed to the effects of weathering, most notably the increase in porosity. The split tensile strength of a specimen was calculated using Equation (8):

$$\sigma_t = \frac{0.636P}{Dt} \text{ (MPa)} \quad (8)$$

where, P is the load at failure, N; D is the diameter of the specimen (mm); and, t is the thickness of the specimen measured at the center [15].

The tensile strengths of the specimens are presented in Table 2. As expected, there was some overlap within the tensile strength values between the highly weathered and differentially weathered specimens. One issue with using only tensile strength as an indicator of weathering is that there is often overlap in strength values between weathering states. As demonstrated in Table 2, there was overlap between the tensile strength of highly weathered specimens and differentially weathered specimens.

Table 2. Unit Weight and Tensile Strength of Specimens

Weathering Classification	Specimen	Unit weight (kN/m ³)	Tensile strength (MPa)
Fresh	H3	22.74	4.61
Differentially Weathered	H2	21.43	2.53
	H5	21.56	1.44
	H14	21.82	2.62
	H18	21.31	1.25
Highly weathered	H7	17.68	1.07
	H13	17.65	0.40
	H15	19.06	1.24
	H19	16.75	0.08
	H20	17.89	1.56

The failure mode (or fracture pattern) exhibited in specimens tested under split tension conditions can provide insight into the internal structure of the specimen. For a non-structured (massive or isotropic) specimen, the typical failure mode of a split tensile strength specimen [18] is shown in Figure 7. Plastic wedges at the top and bottom of the specimen are the first macroscopic features to develop. Subsequently a tensile fracture develops [19]. Using the terminology presented by Tavallali and Verboort [20], this type of fracture is termed a central fracture if it is contained within 10% of the diameter on both sides of the vertical centerline of the specimen.

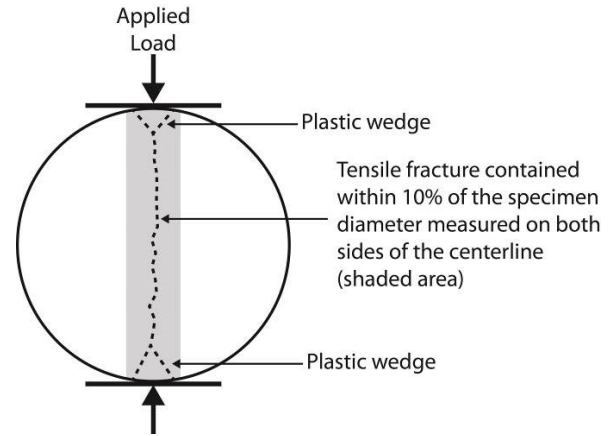


Figure 7. Typical Failure Mode for an Isotropic Specimen under Split Tension Conditions

For transversely isotropic specimens, such as layered sedimentary rock or foliated metamorphic rock, the tensile strength and failure mode depends on both of the structures within the specimen and their orientation with respect to the applied load. Recent studies on layered sandstone have demonstrated that specimens tested with layering perpendicular to the applied load have the highest tensile strength, and specimens tested with layering parallel to the applied load have the lowest tensile strength. Specimen tensile strengths decreased as layering orientation deviated from perpendicular towards parallel [20], [21].

Recent research has also shown that both the orientation of layering and the number of layers in the specimen influence the failure mode. Tavallali and Verboort [20] identified three distinct failure modes: layer activation (fracture occurs along the interface between layers), central fracture (previously described), and non-central fracture (fractures occurring outside the 10% diameter and not classified as layer activation fractures). They also noted that two or three fracture types occurred in the same test. Researchers have noted that when the failure mode deviates from the typical failure mode (see Figure 6), the true tensile strength may not have been achieved [22] and the strength should be considered a failure strength and not necessarily a tensile strength [20].

Unlike the transversely isotropic specimens tested in the previous studies, the weathered limestone specimens tested in this current study were considered to be either homogeneous or zoned. Homogeneous specimens contain a consistent weathering state throughout the entire specimen. These specimens were the fresh or highly weathered specimens and should exhibit the classic central fracture associated with isotropic specimens. The differentially weathered specimens contain zones of materials with different weathering

states and should exhibit layer activation and non-central fracture failure modes. However, the failure planes exhibited in the transversely isotropic specimens would actually be failure surfaces, which would follow the interface between differentially weathered zones. The failure modes for the specimens are listed in Table 3 and the corresponding images showing the failed specimens is shown in Figure 8. As expected, the fresh and highly weathered specimens exhibited the classic failure mode of a central fracture. The differentially weathered specimens exhibited a central fracture with zone activation. With this failure mode, secondary fractures form along the interface between zones of materials with different weathering states.

Table 3. Failure Modes of Specimens

Weathering Classification	Specimen	Failure mode
Fresh	H3	Central fracture
Differentially weathered	H2	Non-central fracture
	H5	Central fracture with zone activation
	H14	Central fracture with zone activation
	H18	Central fracture with zone activation
Highly weathered	H7	Central fracture
	H13	Central fracture
	H15	Central fracture
	H19	Central fracture
	H20	Central fracture

Characterization Results

The most meaningful analysis for the characterization of the weathered limestone specimens is assessment of split tensile strength and the thermal time constant, τ , as a function of unit weight (see Figure 9), since unit weight is often considered one of the best indicators of weathered rock. As expected, the highly weathered specimens had the lowest unit weight and the fresh specimen had the highest unit weight. The unit weights of the differentially weathered specimens fell between those of the fresh and highly weathered specimens. In terms of tensile strength [see Figure 9 (a)], the specimen with the largest unit weight, H3, had the largest tensile strength, and the specimen with the smallest unit weight, H19, had the smallest tensile strength. However, there was considerable overlap between the tensile strengths of the highly weathered specimens and the differentially weathered specimens. The tensile strength for the differentially weathered specimens ranged between 2.62 and 1.25 MPa, and the tensile strength for the highly weathered specimens ranged between 1.56 and 0.08 MPa.



(a) Fresh Specimen (H3) with Central Fracture



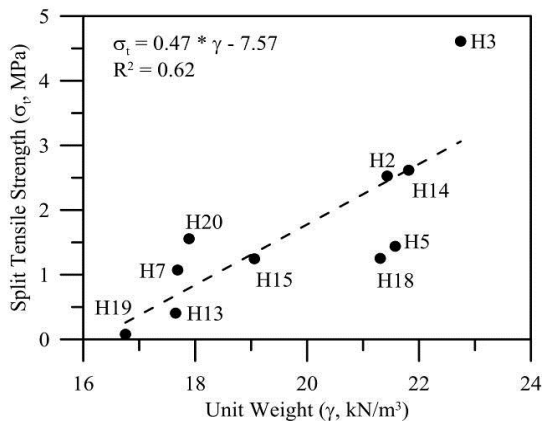
(b) Differentially Weathered Specimen (H18) with Central Fracture and Zone Activation



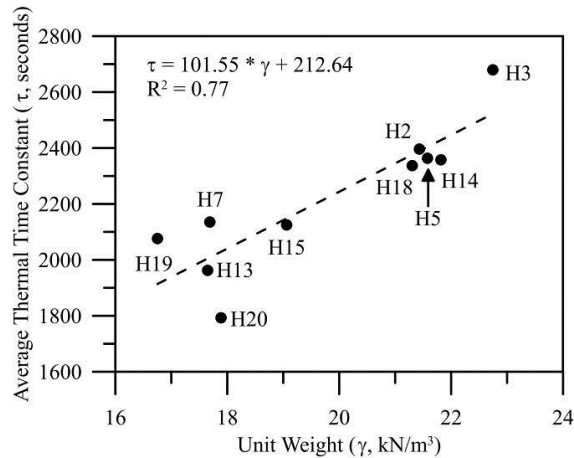
(c) Weathered Specimen (H20) with Central Fracture

Figure 8. Failure Modes for Specimens

Another interesting feature of tensile strength versus unit weight plot was the maximum tensile strength value. The dashed linear regression line shown in the plot makes the maximum tensile strength value appear to be an outlier. However, based on the failure modes exhibited in the specimens, the highly weathered and fresh specimens provided the best estimate of the true split tensile strength variation with unit weight. The differentially weathered specimens, which had a central fracture with zone activation failure modes, under-predict the tensile strengths. If the differentially weathered specimen tensile strength data are removed from the analysis, the R^2 value becomes 0.93; however, the dataset is very limited.



(a) Split Tensile Strength as a Function of Unit Weight



(b) Thermal Time Constant as a Function of Unit Weight

Figure 9. Specimen Properties as a Function of Unit Weight

In terms of the thermal time constant, τ , as a function of unit weight [see Figure 9(b)], there was considerable scatter within the highly weathered specimens. Surprisingly, the specimen with the lowest unit weight did not have the lowest thermal time constant value. The average thermal time

constant value for the highly weathered specimens was 2018.2 seconds. The thermal time constant values for the differentially weathered specimens were very similar and the data points were highly clustered and had an average value of 2363.5 seconds. Importantly, there was no overlap between the thermal time constant values of the highly weathered and differentially weathered specimens, which makes this technique a good candidate for the assessment of weathering states.

There was a positive linear relationship between the thermal time constant and unit weight. The R^2 value (0.77) was greater than the R^2 value for tensile strength as a function of unit weight (0.62). The thermal time constant was a poor predictor of tensile strength, as shown in Figure 10. There was a fairly clear distinction between the thermal time constants and the highly weathered, differentially weathered, and fresh specimens. However, the high variability in tensile strengths had a significant influence on the relationship between thermal time constant and tensile strength.

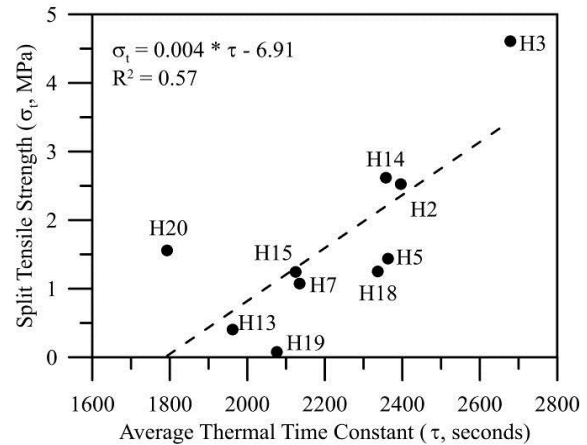


Figure 10. Relationship between Thermal Time Constant and Tensile Strength

Relevance to Existing Standards

As with any potentially new characterization technique, relevance to prevailing standards of practice should be made. Existing standards do exist for the assessment of weathering states of rock. Perhaps the most widely used standard is that of the International Society for Rock Mechanics (ISRM) [15]. This assessment is a qualitative visual description that can be applied to any rock type. The rock material is described based on the state of alteration/discoloration/disintegration. Based on the visual description, the rock material is assessed a weathering grade of I (fresh rock) through VI (completely weathered residual soil). The difficulties with using visual descriptions to assess weather-

ing states have been reported in the literature [3], [4]. ISRM [15] also recommends using a Schmidt hammer to assess weathered rock. This test is considered to be non-destructive. However, for weathered limestone specimens from Florida (Ocala limestone), the Schmidt hammer (L hammer) proved to be a destructive test. The Schmidt hammer caused significant indentations in the specimens or cracked the specimens [10].

There are several advantages for using thermal characterization to assess the weathering state of limestone. The test is truly non-destructive; the specimen is heated and allowed to cool, which will not affect the engineering properties of the specimen. The equipment needed to perform the characterization are found in almost every geotechnical/materials testing laboratory. Finally, the results of the thermal characterization are quantitative and not qualitative. With additional research studies, this technique may become a widely used assessment tool.

Conclusion

Thermal characterization is a rarely used procedure for the assessment of weathering in laboratory rock specimens. Ten specimens were characterized in this current study using temperature decay measurements and split tension testing to assess effects of weathering. The specimens were first visually categorized into one of three weathering classes: fresh (un-weathered), differentially weathered, or highly weathered. The fresh and highly weathered specimens consisted of material of a single weathering state. The differentially weathered specimens contained multiple weathering states within the same specimen. Thermal characterization consisted of heating the specimens to 200 degrees Celsius then recording their cooling using a thermal camera. Using the lumped capacitance method, the thermal time constant was determined for each pixel in the image of the cooling specimen. The fresh specimen had the highest average thermal time constant and the highly weathered specimens had the lowest average thermal time constant. The differentially weathered specimens had average thermal time constants between the fresh and highly weathered specimens.

Tensile testing was used to investigate both the tensile strength and failure modes of specimens as a function of weathering state. The fresh and highly weathered specimens exhibited the classic indirect tension failure mode. The differentially weathered specimens exhibited surface activation failure modes, which indicated their varied internal structure. There was considerable overlap of tensile strengths between weathering states, so tensile strength alone is not a valid predictor of weathering state. There was no overlap of specimen average thermal time constants and weathering

state. There was a poor relationship between specimen average thermal time constants and split tensile strength, which was attributed to the highly variable tensile strength values. Thermal properties have the potential to be a valuable tool in characterizing the weathering state and potentially the geotechnical properties of weathered limestone specimens.

References

- [1] Hencher, S. R., & McNicholl, D. P. (1995). Engineering in Weathered Rock. *Quarterly Journal of Engineering Geology*, 28(3), 253-266.
- [2] Hudyma, N., & Hiltunen, D. R. (2014). Variability of Weathered Limestone Properties in Light of Drilled Shaft Design in Florida. *Geotechnical Special Publication No. 234 – Proceedings of the 2014 Geo-Congress*, (pp. 989-999). Atlanta, GA.
- [3] Cragg, D. J., & Ingman, J. (1995). Rock Weathering Descriptions: Current Difficulties. *Quarterly Journal of Engineering Geology*, 28(3), 277-286.
- [4] Pinho, A., Rodrigues-Carvalho, J. A., Gomes, C., & Duarte, I. M. (2006). Overview of the Evaluation of the State of Rock Weathering by Visual Inspection. *Proceedings of the 10th IAEG International Congress, Paper #260*. Nottingham, UK.
- [5] Tuğrul, A. (2004). The Effect of Weathering on Pore Geometry and Compressive Strength of Selected Rock Types from Turkey. *Engineering Geology*, 75(3-4), 215-227.
- [6] Ündül, Ö., & Tuğrul, A. (2012). The Influence of Weathering on the Engineering Properties of Dunites. *Rock Mechanics and Rock Engineering*, 45(2), 225-239.
- [7] Tuğrul, A., & Gurbinar, O. (1997). A Proposed Weathering Classification for Basalts and Their Engineering Properties (Turkey). *Bulletin of the International Association of Engineering Geology*, 55(1), 139-149.
- [8] Gupta, A. S., & Rao, K. S. (1998). Index Properties of Weathered Rocks: Inter-Relationships and Applicability. *Bulletin of Engineering Geology and the Environment*, 57(2), 161-172.
- [9] Tuğrul, A., & Zarif, I. H. (2000). Engineering Aspects of Limestone Weathering in Istanbul, Turkey. *Bulletin of Engineering Geology and the Environment*, 58(3), 191-206.
- [10] Sarno, A., Farah, R., Hudyma, N., & Hiltunen, D. R. (2009). Relationships between Index and Physical Properties of Weathered Ocala Limestone. *Proceedings of the 43rd U.S. Rock Mechanics Symposium and 4th Canada US Rock Mechanics Symposium, Paper #09-105*, Asheville, NC.

-
- [11] Sarno, A., Farah, R., Hudyma, N., & Hiltunen, D. R. (2010). Relationships between Dynamic and Compression Testing for Weathered Ocala Limestone. *Geotechnical Special Publication No. 199 – Proceedings of GeoFlorida 2010: Advances in Analysis, Modeling & Design*, (pp. 950-959), West Palm Beach, FL.
 - [12] Kodikara, J., Rajeev, P., & Rhoden, N. J. (2011). Determination of Thermal Diffusivity of Soil Using Infrared Thermal Imaging. *Canadian Geotechnical Journal*, 48(8), 1295-1302.
 - [13] Luong, M., & Emami, M. (2014). Characterization of Mechanical Damage in Granite. *Frattura ed Integrità Strutturale* 27(January), pp. 38-42.
 - [14] Florida Department of Transportation. (2015). *Soils and foundations handbook*. Gainesville, FL.
 - [15] Ulusay, R., & Hudson, J. A. (Eds.). (2007). *The complete ISRM suggested methods for rock characterization, testing and monitoring*. Ankara, Turkey: Kozen Ofse Matbaacilik San. Ve Tic. Sti.
 - [16] Sarno, A. I. (2010). *Correlations of static, dynamic, and physical properties to the weathering state of Ocala limestone*. Unpublished Master's thesis, University of North Florida.
 - [17] Incropera, F., & Dewitt, D. (2001). *Fundamentals of Heat and Mass Transfer*. (5th ed.). Wiley.
 - [18] Hobbs, D. W. (1963). The Tensile Strength of Rocks. *International Journal of Rock Mechanics and Mining Sciences & Geomechanics Abstracts*, 1(3), 385-396.
 - [19] Coviello, A., Lagioia, R., & Nova, R. (2005). On the Measure of the Tensile Strength of Soft Rocks. *Rock Mechanics and Rock Engineering*, 38(4), 251-273.
 - [20] Tavallali, A., & Vervoort, A. (2013). Behaviour of Layered Sandstone under Brazilian Test Conditions: Layer Orientation and Shape Effects. *Journal of Rock Mechanics and Geotechnical Engineering*, 5(5), 366-377.
 - [21] Khanlari, G., Rafiei, B., & Abdilor, Y. (2015). An Experimental Investigation of the Brazilian Tensile Strength and Failure Patterns of Laminated Sandstones. *Rock Mechanics and Rock Engineering*, 48(2), 843-852.
 - [22] Dan, D. Q., Konietzky, H., & Herbst, M. (2013). Brazilian Tensile Strength Tests on Some Anisotropic Rocks. *International Journal of Rock Mechanics and Mining Sciences*, 58(February), 1-7.

Biographies

JOHN NUSZKOWSKI is an assistant professor of mechanical engineering at the University of North Florida (UNF). He completed his B.S. in mechanical engineering at the Rochester Institute of Technology (RIT) and earned his

M.S. and Ph.D. degrees in mechanical engineering at West Virginia University (WVU). His research includes thermodynamic systems, vehicle powertrain optimization, in-use vehicle emissions, advanced combustion, alternative fuels, and large-bore diesel emissions reduction. Dr. Nuszowski may be reached at john.nuszowski@unf.edu

MASON MCGOUGH is a master's student of biomedical engineering at the University of Florida. He received his B.S. degree from the University of North Florida (Electrical Engineering, 2014). His research interests include bio-imaging informatics, image processing, machine learning, and feature selection. Mason McGough may be reached at masonmcgough@ufl.edu

LEXI KIMES is a junior civil engineering student at the University of North Florida. She is interested in both geotechnical and structural engineering. In her spare time, she enjoys rock climbing and riding motocross. Lexi Kimes may be reached at lexikimes@gmail.com

ALAN HARRIS is an associate professor of electrical engineering at the University of North Florida. He earned his B.S. degree from the University of Central Oklahoma (Physics, 1999), M.S. degree from the University of Oklahoma (Electrical Engineering, 2002), and his Ph.D. from the University of Oklahoma (Electrical and Computer Engineering, 2005). His interests include image and signal processing for geotechnical engineering, optical communications, and service learning in undergraduate engineering education. Dr. Harris may be reached at alan.harris@unf.edu

NICK HUDYMA is a professor of civil engineering at the University of North Florida. He earned his B.Sc. degree from the University of Manitoba (Geological Engineering), M.S. (Civil and Environmental Engineering, 1994) from the University of Nevada Las Vegas, and Ph.D. (Engineering, 1999) from the University of Nevada Las Vegas. His interests include pre-test characterization of rock, including ultrasonics, laser scanning, and photogrammetry; laboratory characterization of soil and rock; and, ground-penetrating radar in infrastructure and geotechnical site characterization. Dr. Hudyma may be reached at nhudyma@unf.edu

FARIS MALHAS is Dean of the School of Engineering and Technology at Central Connecticut State University. He earned his B.S.E. and M.S.E. (Civil Engineering) degrees from the University of Michigan – Ann Arbor, and his M.Sc. (Engineering Mechanics) and Ph.D. (Civil Engineering) degrees from the University of Wisconsin – Madison. His research interests include structural concrete. Dr. Malhas may be reached at fmalhas@ccsu.edu

PREDICTING SOLAR RADIATION FOR RENEWABLE ENERGY TECHNOLOGIES: A RANDOM FOREST APPROACH

Frank E. Yeboah, North Carolina A&T State University; Robert Pyle, North Carolina A&T State University;
Christian A. Bock Hyeng, North Carolina A&T State University

Abstract

The recent, historically warmer summers recorded in several geographic regions has once again brought back the limelight on global climate change. Climate change is thought to be caused by greenhouse gases (GHGs), most notable among them, carbon dioxide (CO₂). In most economies, CO₂ is generated from the combustion of fossil fuels—coal, natural gas, and oil. Renewable energy technologies such as wind and solar have been known to be environmentally more benign than their fossil-fuel counterparts. However, unlike the latter, the former is much more difficult to predict and to quantify for planning purposes. Their estimations are mostly affected by seasonality, geographic, and other meteorological factors.

All of these make predicting solar insolation an hour, day, or week ahead very complex. Traditional modeling techniques such as regression and time series are not robust enough to handle the complexity and the nonlinearity of the predictive variables involved in solar resource modeling. In this paper, the authors introduce Random Forests (RF), an ensemble technique, as an example of a supervised machine learning (SML) technique that is capable of incorporating the complexities and the nonlinearity for predicting solar resources.

Introduction

Electrical energy systems are very complex and delicate, because supply normally has to match demand in order to maintain the stability of the power grid. Most power plants are of the conventional type: fossil-fueled and nuclear power plants. Fossil-fueled power plants use coal, natural gas, and oil to generate electrical energy. Such plants are, however, associated with the emission of greenhouse gases, most notable among them, CO₂. Greenhouse gases are known to effect global climate change, which could lead to severe weather conditions, coastal flooding, etc. Consequently, policy-makers have been proposing that a certain amount of renewable energy, solar and wind, be made part of the electrical energy generation portfolio. One of the biggest challenges of renewable energy resources is predicting

their availability and quantity. They are highly variable and subject to daily and annual variations and, therefore, can be difficult to estimate and lead to grid instability. Being able to predict their availability, especially in a day-ahead energy market, would be crucial. Traditional forecasting techniques such as regression, autoregressive integrated moving average (ARIMA), and other time-series models are not robust enough for predicting renewable energy resources. These techniques are referred to as data models by Breiman [1].

According to De'arth [2], traditional statistical methods, while good at explaining data relationships, are relatively poor predictors. Hence, predicting complex systems are better done using machine learning (ML) techniques [2]. Evans et al. [3] note that traditional statistical methods have difficulty in addressing models with complex non-linear interactions, which also exhibit autocorrelation, non-stationarity, and high dimensionality. Sharma et al. [4] posit that machine learning techniques give better prediction accuracy than traditional methods. They are of the view that predicting solar intensity cannot be done accurately using just a single weather variable.

Several variables have to be incorporated for better prediction to be achieved, thereby increasing the dimensionality and making the predictive model much more complex [4]. They also note that solar intensity exhibits complex relationships with weather variables, and that the variables themselves depict complex relationships among each other. All of these complexities in predicting solar radiation, therefore necessitate the use of machine learning techniques [4].

Machine learning techniques are suitable for predicting complex systems, especially where large quantities of data are available. The methods automatically detect patterns in the data and then use the discovered pattern to predict future data [2]. Furthermore, complex relationships and spatial patterns can be discovered more readily, compared with the traditional probability data model that assumes normality [3]. ML techniques, also referred to as algorithmic models, are much more capable of producing reliable information about the structure of the relationship between predictor variables and response variables than data models [1].

Machine learning techniques may be categorized into two main groups; supervised and unsupervised techniques [5]. Supervised machine learning (SML) techniques are used in predictive modeling, where features or attributes or input variables are mapped onto an output or response variable. If the output or response variable is categorical (e.g., yes/no, male/female, etc.) then it is a classification or pattern recognition problem. For an output or response variable that is continuous (nominal, ordinal, or real-valued) then it is a regression problem. SML techniques are the most widely used ML techniques. Unsupervised machine learning (UML) techniques, on the other hand, are used for information or knowledge discovery. Clustering is a typical example of the UML techniques. The focus of this current study was on SML techniques.

There are several methods for solving SML problems. These include neural networks (NN), support vector machines (SVM), ensemble methods (EM) including random forest (RF) classification and regression, which build a series of trees using a random selection of variables. A review of the use of ML techniques for predicting renewable energy generation and integration was carried out by Perera et al. [6]. Artificial neural networks have been used in modeling renewable energy resources such as solar and wind [7–11]. Yona et al. [12] also compared the performance of different neural networks in predicting short-term solar radiation. A comparison of different forecasting methods for predicting wind speed was conducted by Jursa [13] and Sfetsos [14]. Furthermore, SVMs have been used to predict wind speed [15]. In this current study, discussions were limited to the EM methods. Since this study was a regression problem, meaning that the output was a continuous variable, only the ensemble regression methods—specifically, random forest (RF) and bagged decision trees (BDT)—were used.

Ensemble Methods

The starting point for discussions on most ensemble methods are classification and regression trees (CART). CART is a nonparametric model that can be used on datasets with large numbers of cases and large numbers of variables. It is also extremely resistant to outliers [3], [16]. It does not assume normality [16] of the underlying distribution and is also not based on user-defined model statements, unlike ordinary least squares regression and discriminant analysis. CART can account for non-stationarity, can identify and incorporate complex variable interaction, and is able to handle high-dimensional data [1], [3], [17]. Unfortunately, CART is also subject to severe over-fit [3], [16], [18]; the final tree may not provide the optimal solution; high variance and small changes in the data can lead to different

splits, making interpretation unstable; and, interpretation can be difficult if there exist a collinearity between independent variables.

To address these limitations of CART, ensemble learning methods—bagging [19], boosting [2], [20], [21], and random forest [22], were developed by Evans et al. [3]. In bagging (**bootstrap aggregating**), multiple bootstrap samples were drawn from the original data and some prediction method was applied to each bootstrap sample; finally, the results were aggregated in order to obtain the overall prediction [16]. When predicting a numerical outcome, as in regression, averaging is used for aggregation; and, when predicting a class, as in classification, plurality or simple vote is used. Bootstrap replicates of the learning set used to form multiple versions by making replicates of bagging can reduce the variance in trees, thereby improving predictions [19].

Methods such as CART and neural networks are very unstable and, therefore, should benefit from bagging [16]. Random forest (RF) is a technique developed from CART and bagging approaches [3]. It is a collection of tree-structured weak learners (i.e., have high prediction errors) that are comprised of identically distributed random vectors, where each tree contributes to prediction. A random subset of predictors was selected to determine the split for each tree of collection. RF is robust to noise, even for a very large number of independent variables [2], [3], [18], [22]. The number of randomly selected predictors is fixed in RF and, by injecting randomness, the correlations between the predictions made by each tree are reduced, leading to a reduction in the variance of the prediction error [2].

The technique allows one to use out-of-bag (OOB) data to get an unbiased estimate of the model error. Furthermore, it is easy to calculate a measure of variable importance [22]. In addition, RF is not subject to over-fit [1], [23], albeit caution should be taken to avoid over-correlation or inflation of the variance of the RF ensemble [3]. Bagging is a special case of random forest [23], where the number of randomly sampled predictors is equal to the number of predictors of the original sample.

In other words, RF uses $p^{1/2}$ (where p is number of predictors) for classification and $p/3$ for regression. Thus, when RF uses p predictors for the analysis then the bagging technique is being used. Boosting is another ensemble technique, which combines the predictions from weak classifiers such that average predictions create a strong classifier (i.e., one with a low prediction error) [2]. This study focused on bagging (BDT) and random forest (RF) for the ensemble model prediction.

Data Preparation and Modeling

The data for modeling came from 5-minute interval, 1-year (January 1 – December 31, 2014) solar data acquired from a weather station in Greensboro, NC, USA (Lat. 36.08° N, Lon. 79.82° W) installed by the authors. A total of 105,120 datasets of 5-minute data were aggregated into 8760 hourly datasets for training and testing of the model. The predictor (input) variables included weather data (temperature: average, minimum, maximum; dew point; and, relative humidity, RH. In addition, seasonal variables (month of year; day of month; and, hour of day) were created from the timestamp of the data in order to capture seasonality. The response (target/output) variable was global solar radiation (GSR in kWh/m²) that were able to be predicted. A separate dataset (January 1 – 4, 2015) that was not used in the modeling was used in predicting hourly and daily GSR four days ahead.

Two different types of ensemble methods—random forests/bagged decision tree (RF/BDT) and bagging techniques (BDT)—were developed for predicting global solar radiation (GSR). Typical steps for modeling using random forests, as illustrated in Figures 1-4, included:

- Step 1: Determine the optimal leaf size to grow on a tree.
- Step 2: Select variables of importance.
- Step 3: Determine the optimal number of trees to grow in the forest.
- Step 4: Determine prediction accuracy (as measured by MSE of OOB test samples).
- Step 5: Use the model for prediction.

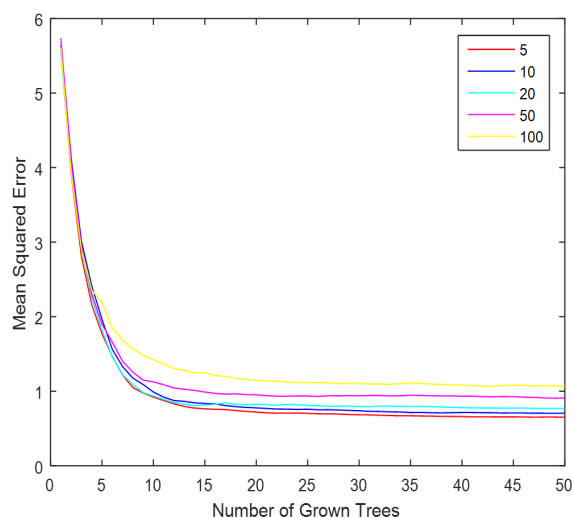


Figure 1. Optimal Leaf Size

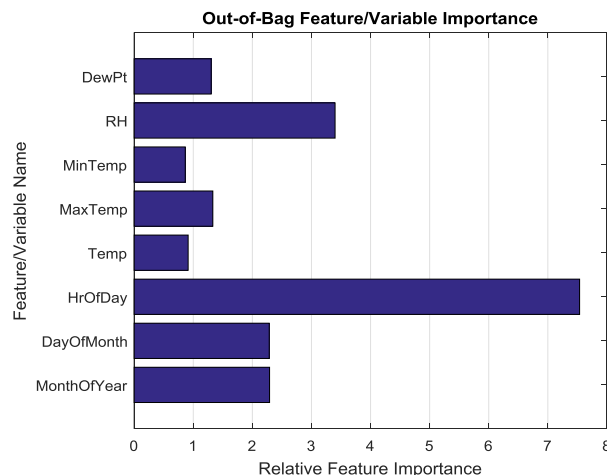


Figure 2. Variable Selection/Importance

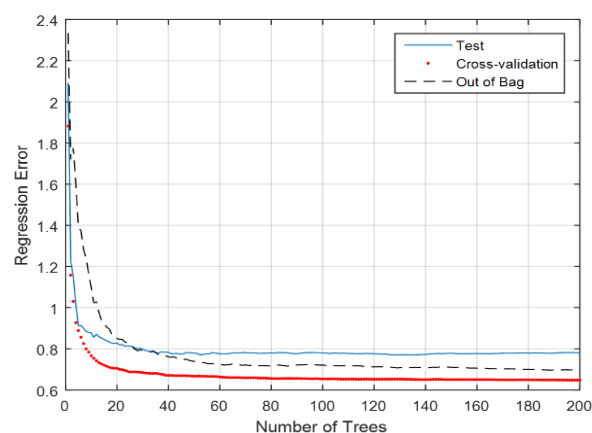


Figure 3. Determining the Number of Trees to Grow

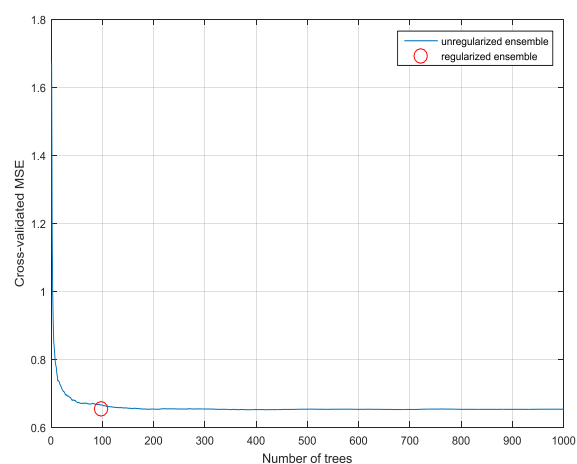


Figure 4. Regularizing the Random Forest Ensemble

MATLAB Tool was used in modeling the ensembles. From Figure 1, it can be shown that the optimal leaf size to grow on a tree in the forest is five. This corresponds to the curve with the lowest out-of-bag (OOB) mean square error (MSE). While HrOfDay is deemed much more important (according to Figure 2), the other variables were also included as being important and, therefore, were kept in the model. An RF model was developed with the number of trees varied up to 200 trees in the forest. Regression error tests were developed using independent test sets, cross-validation, and out-of-bag (OOB) techniques; refer to Figure 3, as it indicates that beyond 100 trees there were no improvements in the model's MSE. Since having more trees is synonymous with more computational space and time, finding the optimal number of trees that would not significantly increase the MSE of the model was sought. This could be done using regularization as an additional modeling option. Based on the regularization of the model, a 5-fold, cross-validation, as shown in Figure 4, using no more than 100 trees, and removing any tree beyond this limit did not affect the overall MSE of the cross-validated model for prediction to any significant degree. Therefore, 100 trees were selected to be grown in the forest.

A feed-forward neural network (FFNN) technique was developed by Yeboah [24] for predicting solar radiation in Greensboro, NC. The architecture of the FFNN model was comprised of a 5-layer neural network, an input layer, three hidden layers, and an output layer. The input layer (IP) consisted of eight elements, corresponding to the input variables used in the model. The three hidden layers (HLs) were created with 65 neurons each. The network design, therefore, was an 8-65-65-65-1, which was the optimal model. The hyperbolic tangent-sigmoid transfer function was used to convert the output from the input layer to the first HL, and also from the first to the second, and from the second to the third HLs. The output of the third HL, which was connected to the output layer, was converted using a linear transfer function. The Levenberg-Marquardt (LM) algorithm was employed in the backpropagation stage. The results of this FFNN model were used to compare the efficacy of the BDT and BDT/RF models developed in this study.

Results and Discussion

In this current study, two main algorithmic models were used to predict the hourly day-ahead global solar radiation, in $\text{kWh}\cdot\text{m}^{-2}$, four days ahead: BDT and RF. A feed-forward neural network (FFNN) was developed by Yeboah [24] for predicting global solar radiation at the same geographical location. The results in this current study were compared with the FFNN model. Figures 5-7 depict plots of the actual and the predicted solar energy for each model. These plots

illustrate how well each model was able to predict GSR using the original data (January 1 – December 31, 2014) that were used in developing the models.

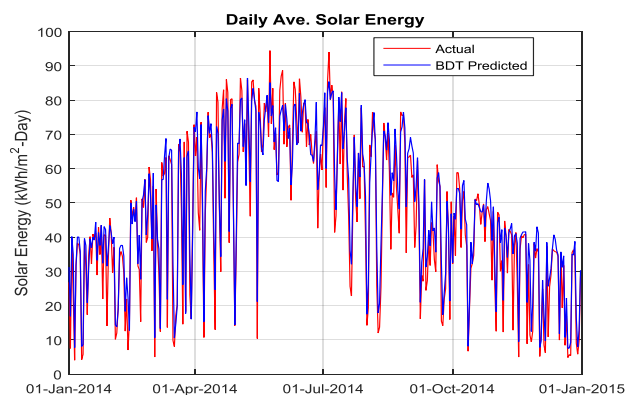


Figure 5. Predicting Global Solar Radiation using the Bagging Technique (BDT)

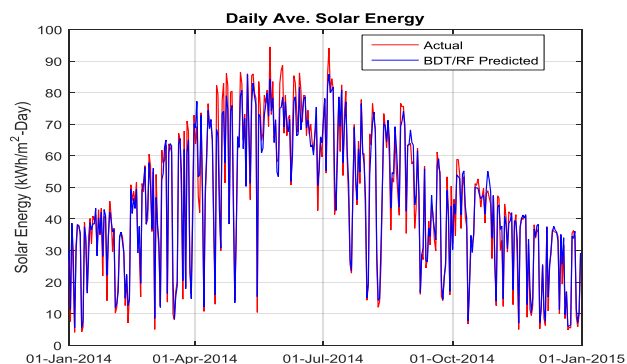


Figure 6. Predicting Global Solar Radiation using the Bagged Decision Tree (BDT)/Random Forest (RF)

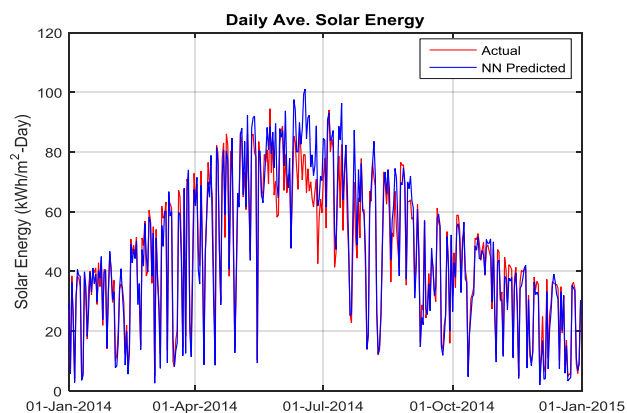


Figure 7. Predicting Global Solar Radiation using the Feed-Forward Neural Network (FFNN)

Tables 1-4 illustrate the forecasts of solar energy against the actual values and the corresponding prediction errors for each of the models considered in the study. The efficacy of the models and their prediction accuracy were measured using MSE (mean square error), RMSE (root mean square error), and MAE (mean absolute error), and may be expressed mathematically by Equations (1)-(3):

$$MSE = \frac{1}{N} \sum_{t=1}^N (Y_t - \hat{Y}_t)^2 \quad (1)$$

$$RMSE = \left[\frac{1}{N} \sum_{t=1}^N (Y_t - \hat{Y}_t)^2 \right]^{0.5} \quad (2)$$

$$MAE = \frac{1}{N} \sum_{t=1}^N |Y_t - \hat{Y}_t| \quad (3)$$

where, N is the forecasting horizon; Y_t is the actual GSR (global solar radiation) at time period t measured at the given geographical location; and, $\hat{Y}_t$ is the predicted GSR.

Table 1 depicts the generalization accuracy of the developed models.

Table 1. Accuracy of the BDT, RF, and FFNN Models

MEASURE	MODEL		
	BDT	BDT/RF	FFNN
MAE	0.3199	0.2641	0.3235
MSE	0.3935	0.2303	0.3420
RMSE	0.6273	0.4799	0.5848

Based on MSE, MAE, and RMSE, the model with the best forecasting accuracy was the random forest (RF) model. While the FFNN was slightly better than the bagging technique (BDT), based on MSE (and certainly RMSE), it may not be as robust as the ensemble methods when it comes to predicting ahead of time; that is, when new data are introduced for prediction, as will be manifested in short-term prediction.

Short-Term Prediction of Solar Radiation

The models were used to predict solar radiation (hourly and average daily) four days ahead, from January 1-4, 2015. Figures 8-13 depict the plots of the actual and the predicted solar radiation, for the bagging (BDT), random forest (BDT/RF) and feed-forward neural network (FFNN) techniques, respectively.

The plots of the actual and the predicted hourly 4-day ahead global solar radiation for the BDT, BDT/RF and FFNN models are shown in Figures 8, 9 and 10, respectively.

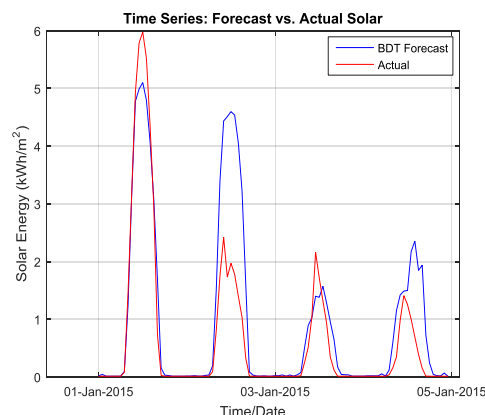


Figure 8. BDT Prediction of Hourly 4-Day-Ahead Solar Radiation

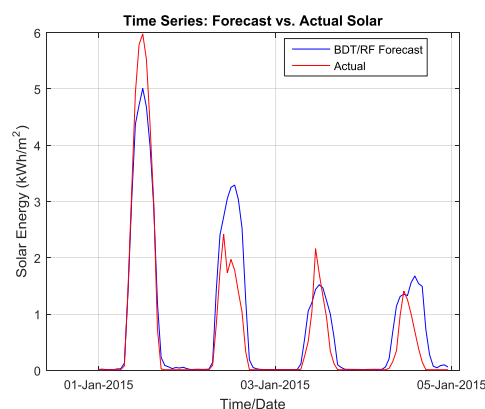


Figure 9. BDT/RF Prediction of Hourly 4-Day-Ahead Solar Radiation

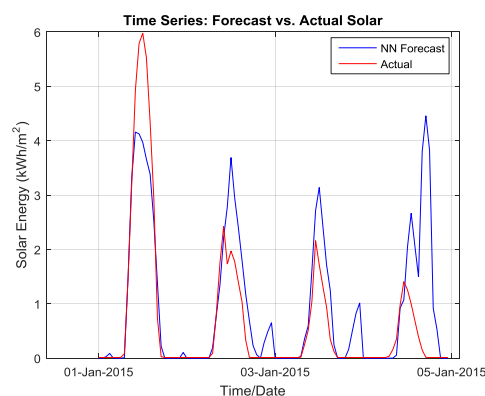


Figure 10. FFNN Prediction of Hourly 4-Day-Ahead Solar Radiation

Figures 11, 12 and 13, depict the actual and the predicted 4-day ahead daily average global solar radiation for the BDT, BDT/RF and FFNN models, respectively.

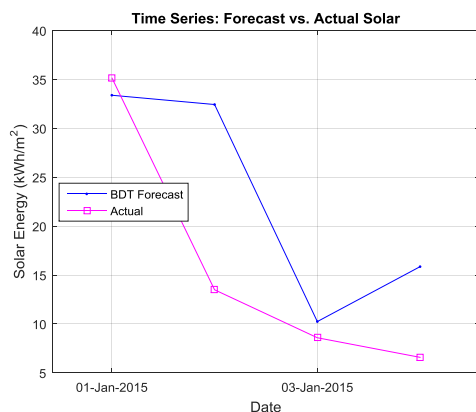


Figure 11. BDT Prediction of 4-Day-Ahead Average Daily Solar Radiation

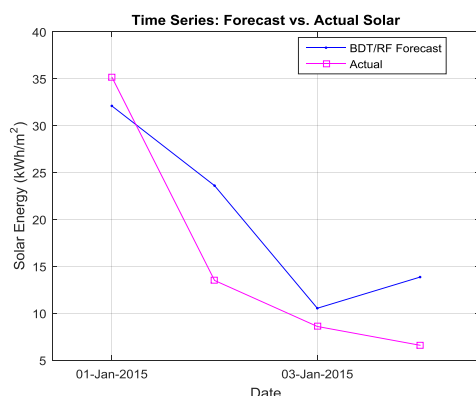


Figure 12. BDT/RF Prediction of 4-Day-Ahead Average Daily Solar Radiation

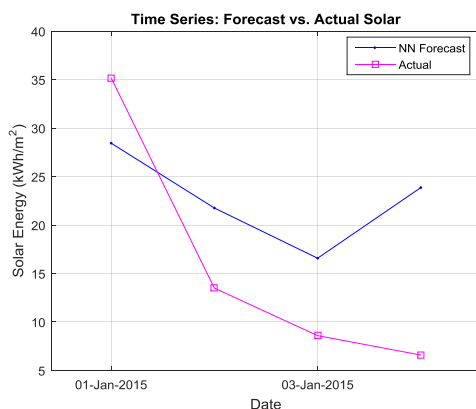


Figure 13. FFNN Prediction of 4-Day-Ahead Average Daily Solar Radiation

The prediction of global solar radiation using the bagging technique (BDT) is given in Table 2.

Table 2. Prediction of Global Solar Insolation using the Bagging Technique (BDT)

Day	Actual (kWh.m ⁻²)	Forecast (kWh.m ⁻²)	Error (%)
1	35.18	33.386	5.10
2	13.493	32.437	-53.85
3	8.589	10.242	-4.70
4	6.585	15.858	-26.36

Table 3 illustrates the prediction of the global solar radiation using the random forest (BDT/RF) model. The prediction of the global solar radiation using the feed-forward neural network (FFNN) is depicted in Table 4.

Table 3. Prediction of Global Solar Radiation using the Bagged Decision Tree/Random Forest (BDT/RF)

Day	Actual (kWh.m ⁻²)	Forecast (kWh.m ⁻²)	Error (%)
1	35.18	32.096	8.77
2	13.493	23.596	-28.72
3	8.589	10.53	-5.52
4	6.585	13.851	-20.66

Table 4. Prediction of Global Solar Radiation using the FFNN

Day	Actual (kWh.m ⁻²)	Forecast (kWh.m ⁻²)	Error (%)
1	35.18	28.447	19.14
2	13.493	21.756	-23.49
3	8.589	16.576	-22.70
4	6.585	23.865	-49.12

The prediction errors for the 4-day-ahead hourly GSR, as measured by MAE, MSE, and RMSE, for each of the models, are given in Table 5. From Table 5, it is quite clear that, based on MAE, MSE, and RMSE, the ensemble methods—BDT and BDT/RF—have better prediction accuracy than the FFNN model.

Table 5. 4-Day Hourly Prediction Errors of the Predictive Models

MEASURE	MODEL		
	BDT	BDT/RF	FFNN
MAE	0.3801	0.2752	0.6700
MSE	0.6144	0.2482	1.0094
RMSE	0.7838	0.4982	1.0047

While the FFNN model seems to have a lower prediction error, when the prediction data come from the original data used in developing the model (see Figure 7), its generalization error for the model has never been seen before (see Figures 10 and 13) and tends to be higher than both the BDT and BDT/RF models. In other words, the ensemble methods are better predictors (as depicted in Table 5) than the FFNN, because of the better generalization errors. Another point to note is that all of the models seemed to predict very well on the first day (or 24-hour period), but their accuracy deteriorated as the forecasting horizon increased, with the FFNN model deteriorating much faster than the BDT and BDT/RF ensemble techniques.

Conclusions

Predicting solar energy ahead of time is a sure and efficient means of managing power demand and supply, especially where renewable energy technologies are part of the portfolio. Since electric power systems consist of many complex factors with nonlinear attributes or characteristics, using traditional data modeling techniques such as regression and time-series would not lead to appropriate results. Algorithmic (machine learning) models, however, can handle such complexities very easily. The ensemble method—random forest, a supervised machine learning technique—was used for predicting 4-day-ahead global solar radiation.

In addition, a special case of the RF model, the bagging technique (BDT) and a feed-forward neural network (FFNN), were also employed in order to compare the efficacy of the RF model. While all three models yielded comparable predictions, the ensemble methods (BDT and BDT/RF) produced better predictions (based on prediction error) for short-term predictions than the FFNN model. Other supervised machine learning techniques such as support vector machines (SVM) for regression, and other ensemble methods such as boosting, and also radial basis function (RBF) and recurrent neural networks (RNN), should be explored further for predicting global solar radiation in the short-term.

Acknowledgements

Sincere thanks and gratitude to Christian Lotte and Duc Nguyen, former construction management students, who helped with the installation of the weather station that was used in gathering the meteorological data for this study. Financial assistance from the Office of Success, School of Technology, for the acquisition of the weather station is also acknowledged.

References

- [1] Breiman, L. (2001b). Statistical Modeling: The Two Cultures (with comments and a rejoinder by the author). 199-231. doi: 10.1214/ss/1009213726
- [2] De'ath, G. (2007). Boosted Trees for Ecological Modeling and Prediction. *Ecology*, 88(1), 243-251.
- [3] Evans, J. S., Murphy, M. A., Holden, Z. A., & Cushman, S. A. (2001). Modeling Species Distribution and Change Using Random Forest. In C. A. Drew, et al. (Ed.), *Predictive Species and Habitat Modeling in Landscape Ecology: Concepts and Applications* (pp. 139-159): Springer Science + Business Media, LLC.
- [4] Sharma, N., Sharma, P., Irwin, D., & Shenoy, P. (2011). *Predicting solar generation from weather forecasts using machine learning*. Paper presented at the Smart Grid Communications (SmartGridComm), 2011 IEEE International Conference on.
- [5] Murphy, K. P. (2012). *Machine learning: a probabilistic perspective*. Cambridge, MA: MIT Press.
- [6] Perera, K. S., Aung, Z., & Woon, W. L. (2014). Machine Learning Techniques for Supporting Renewable Energy Generation and Integration: A Survey *Data Analytics for Renewable Energy Integration* (pp. 81-96): Springer.
- [7] Al Shamisi, M. H., Assi, A. H., & Hejase, H. A. N. (2011). Using MATLAB to Develop Artificial Neural Network Models for Predicting Global Solar Radiation in Al Ain City – UAE. In A. Assi (Ed.), *Engineering Education and Research Using MATLAB: In Tech*.
- [8] Fadare, D. A. (2010). The application of artificial neural networks to mapping of wind speed profile for energy application in Nigeria. *Applied Energy*, 87(3), 934-942. doi: <http://dx.doi.org/10.1016/j.apenergy.2009.09.00>
- [9] Rehman, S., & Mohandes, M. (2008). Artificial neural network estimation of global solar radiation using air temperature and relative humidity. *Energy Policy*, 36(2), 571-576. doi: <http://dx.doi.org/10.1016/j.enpol.2007.09.03>
- [10] Sözen, A., Arcaklioğlu, E., Özalp, M., & Kanit, E. G. (2004). Use of artificial neural networks for mapping of solar potential in Turkey. *Applied Energy*, 77(3), 273-286. doi: [http://dx.doi.org/10.1016/S0306-2619\(03\)00137-](http://dx.doi.org/10.1016/S0306-2619(03)00137-)
- [11] Stathopoulos, C., Kaperoni, A., Galanis, G., & Kallos, G. (2013). Wind power prediction based on numerical and statistical models. *Journal of Wind Engineering and Industrial Aerodynamics*, 112(0), 25-38. doi: <http://dx.doi.org/10.1016/j.jweia.2012.09.00>
- [12] Yona, A., Senjyu, T., Funabashi, T., Mandal, P., & Kim, C. (2013). Decision Technique of Solar Radia-

- tion Prediction Applying Recurrent Neural Network for Short-Term Ahead Power Output of Photovoltaic System. *Smart Grid and Renewable Energy*, 4(6A), 32-38.
- [13] Jursa, R. (2007). *Wind power prediction with different artificial intelligence models*. Paper presented at the Proceedings of the European Wind Energy Conference EWEC'07.
 - [14] Sfetsos, A. (2000). A comparison of various forecasting techniques applied to mean hourly wind speed time series. *Renewable Energy*, 21(1), 23-35. doi: [http://dx.doi.org/10.1016/S0960-1481\(99\)00125-](http://dx.doi.org/10.1016/S0960-1481(99)00125-)
 - [15] Mohandes, M. A., Halawani, T. O., Rehman, S., & Hussain, A. A. (2004). Support vector machines for wind speed prediction. *Renewable energy*, 29(6), 939-947. doi: <http://dx.doi.org/10.1016/j.renene.2003.11.00>
 - [16] Sutton, C. D. (2005). Classification and Regression Trees, Bagging, and Boosting. In C. R. Rao, E. J. Wegman & J. L. Solka (Eds.), *Handbook of Statistics: Data Mining and Data Visualization* (Vol. 24). Amsterdam, Netherlands: Elsevier.
 - [17] Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*: CRC press.
 - [18] Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: data mining, inference, and prediction* (2nd ed.). New York, NY: Springer.
 - [19] Breiman, L. (1996). Bagging Predictors. *Machine Learning*, 24(2), 123-140. doi: 10.1023/A:1018054314350
 - [20] Freund, Y., & Schapire, R. E. (1996). *Experiments with a new boosting algorithm*. Paper presented at the ICML.
 - [21] Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 1189-1232.
 - [22] Breiman, L. (2001a). Random Forest. *Machine Learning*, 45(1), 5-32.
 - [23] Liaw, A., & Wiener, M. (2002). Classification and Regression by randomForest. *R News*, 2/3, 18-22.
 - [24] Yeboah, F. E. (2015). *A Supervised Machine Learning Technique for Predicting Solar Radiation*. Unpublished Manuscript, North Carolina Agricultural and Technical State University, Greensboro, NC.
- ana State University Consortium Ph.D. Program in Technology Management, and a visiting professor of energy economics at the Kwame Nkrumah University of Science and Technology in Ghana. He earned his M.E. (Dipl.-Ing.) from the Technical University of Clausthal, Germany (Mining and Mineral Economics, 1997) and D.Eng.Sc. degree (Earth and Environmental Engineering, with a focus on energy economics, 2004) from Columbia University in the City of New York. His research interests include environmental sustainability, energy security and independence, numerical and Monte Carlo simulation of energy systems, applied financial management, risk management, modeling, and forecasting using machine learning techniques such as random forest, bagged decision trees, support vector machines, and artificial neural networks. Dr. Yeboah may be reached at feyeboah@ncat.edu
- ROBERT PYLE** is a professor of construction management at North Carolina A&T State University. He earned his B.A. degree (Industrial Art/Construction Concentration, 1966), M.A. degree (Industrial Education/Construction Concentration, 1968) from the College of New Jersey, and his Ph.D. (Vocational Technical Education/Building Maintenance, 1975) from the University of Pittsburgh. Dr. Pyle is currently teaching at North Carolina A&T State University. His interests include solar energy and residential construction. Dr. Pyle may be reached at pyle@ncat.edu
- CHRISTIAN A. BOCK HYENG** is an assistant professor of construction management. He earned his B.S. degree (Building Construction, 2000), M.S. degree (Bridge Construction, 2002), and Ph.D. (Strength of Material, 2006) from the Russian Peoples Friendship University of Moscow. Dr. Bock Hyeng is currently teaching at North Carolina Agricultural and Technical State University. His interests include sustainable construction material, water scarcity, and thin-shell construction. Dr. Bock Hyeng may be reached at cbhyeng@ncat.edu

Biographies

FRANK E. YEBOAH is an assistant professor of construction management in the Department of Built Environment, North Carolina Agricultural and Technical State University (NCATSU). He is also a faculty member in the Indi-

INSTRUCTIONS FOR AUTHORS: MANUSCRIPT REQUIREMENTS

The INTERNATIONAL JOURNAL OF MODERN ENGINEERING is an online/print publication designed for Engineering, Engineering Technology, and Industrial Technology professionals. All submissions to this journal, submission of manuscripts, peer-reviews of submitted documents, requested editing changes, notification of acceptance or rejection, and final publication of accepted manuscripts will be handled electronically. The only exception is the submission of separate high-quality image files that are too large to send electronically.

All manuscript submissions must be prepared in Microsoft Word (.doc or .docx) and contain all figures, images, and/or pictures embedded where you want them and appropriately captioned. Also, for all accepted manuscripts, each figure, image, or picture that was imported into your Word document must be saved individually as a **300dpi or higher JPEG (.jpg)** file and submitted separately; your manuscript and figure numbers must be used in the title of the file (e.g., **M12-S-18 Figure 4**); that means one additional file for each image imported into your manuscript. These 300dpi images do NOT need to be embedded in your manuscript. For tables or graphs created directly in Word, you do not need to submit them as separate files.

Included below is a summary of the formatting instructions. You should, however, review the [sample Word document](#) on our website for details on how to correctly format your manuscript. The editorial staff reserves the right to edit and reformat any submitted document in order to meet publication standards of the journal.

The references included in the References section of your manuscript must follow APA-formatting guidelines. In order to help you, the sample Word document also includes numerous examples of how to format a variety of scenarios. Keep in mind that an incorrectly formatted manuscript will be returned to you, a delay that may cause it (if accepted) to be moved to a subsequent issue of the journal.

1. **Word Document Page Setup:** Two columns with 1/4" spacing between columns; Top of page = 3/4"; Bottom of page = 1" (from the top of the footer to bottom of page); Left margin = 3/4"; Right margin = 3/4".
2. **Paper Title:** Centered at the top of the first page with a 22-point Times New Roman (Bold), Small-Caps font.

3. **Page Breaks:** Do not use page breaks.
4. **Body Fonts:** Use 10-point Times New Roman (TNR) for body text throughout (1/8" paragraph indentation); 9-point TNR for author names/affiliations under the paper title; 16-point TNR for major section titles; 14-point TNR for minor section titles; 9-point TNR BOLD for caption titles for tables and figures; other font sizes as noted in the sample document.
5. **In-text Referencing:** List and number each reference when referring to them in the body of your document (e.g., [1]). The first entry must be [1] followed by [2], [3], etc., continuing in numerical order to the final entry in your References section. Again, see the sample Word document for specifics. Do not use the End-Page Reference utility in Microsoft Word. You must manually place references in the body of the text.
6. **Tables and Figures:** Center all tables and figures. Captions for tables must be above the table, while captions for figures are below; all captions are left-justified.
7. **Page Limit:** Manuscripts should not be more than 15 pages (single-spaced, 2-column format).
8. **Page Numbering:** Do not use page numbers.
9. **Publication Charges:** Manuscripts accepted for publication are subject to mandatory publication charges.
10. **Copyright Agreement:** Review the [copyright transfer agreement](#).
11. **Submissions:** All manuscripts and associated files must be submitted electronically; the only exception would be if you wish to submit a CD with the individual files for your high-quality images, as noted above.

MANUSCRIPTS should be submitted to Dr. Philip D. Weinsier, manuscript editor, at philipw@bgsu.edu along with a copy to editor@ijme.us.

FILES containing your high-quality images should **ONLY** be submitted to philipw@bgsu.edu.



www.ijeri.org

Print ISSN: 2152-4157
Online ISSN: 2152-4165



www.iajc.org

INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH AND INNOVATION

ABOUT IJERI:

- IJERI is the second official journal of the International Association of Journals and Conferences (IAJC).
- IJERI is a high-quality, independent journal steered by a distinguished board of directors and supported by an international review board representing many well-known universities, colleges, and corporations in the U.S. and abroad.
- IJERI has an impact factor of **1.58**, placing it among an elite group of most-cited engineering journals worldwide.

IJERI SUBMISSIONS:

- Manuscripts should be sent electronically to the manuscript editor, Dr. Philip Weinsier, at philipw@bgsu.edu.

For submission guidelines visit
www.ijeri.org/submissions

TO JOIN THE REVIEW BOARD:

- Contact the chair of the International Review Board, Dr. Philip Weinsier, at philipw@bgsu.edu.

For more information visit
www.ijeri.org/editorial

INDEXING ORGANIZATIONS:

- IJERI is currently indexed by 16 agencies. For a complete listing, please visit us at www.ijeri.org.

OTHER IAJC JOURNALS:

- The International Journal of Modern Engineering (IJME)
For more information visit www.ijme.us
- The Technology Interface International Journal (TIIJ)
For more information visit www.tiij.org

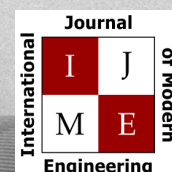
Contact us:

Mark Rajai, Ph.D.

Editor-in-Chief
California State University-Northridge
College of Engineering and Computer Science
Room: JD 4510
Northridge, CA 91330
Office: (818) 677-5003
Email: mrajai@csun.edu



www.tiij.org



www.ijme.us

THE LEADING JOURNAL OF ENGINEERING, APPLIED SCIENCE AND TECHNOLOGY

The latest impact factor (IF) calculation (Google Scholar method) for IJME of 3.0 moves it even higher in its march towards the top 10 engineering journals.

**IJME IS THE OFFICAL AND FLAGSHIP JOURNAL OF THE
INTERNATIONAL ASSOCIATION OF JOURNALS AND CONFERENCE (IAJC)**

www.iajc.org



The International Journal of Modern Engineering (IJME) is a highly-selective, peer-reviewed journal covering topics that appeal to a broad readership of various branches of engineering and related technologies.

IJME is steered by the IAJC distinguished board of directors and is supported by an international review board consisting of prominent individuals representing many well-known universities, colleges, and corporations in the United States and abroad.

IJME Contact Information

General questions or inquiries about sponsorship of the journal should be directed to:

Mark Rajai, Ph.D.

Editor-in-Chief

Office: (818) 677-5003

Email: editor@ijme.us

Department of Manufacturing Systems Engineering & Management

California State University-Northridge

18111 Nordhoff St.

Northridge, CA 91330