

SUMMER
2021



JOURNAL OF
MANAGEMENT
&
ENGINEERING
INTEGRATION

AIEMS
VOL 14
NO 1

ISSN: 1939-7984

JOURNAL OF MANAGEMENT AND ENGINEERING INTEGRATION

Editor-in-Chief

Edwin Sawan, Ph.D., P.E.
Professor Emeritus
Wichita State University
Edwin.sawan@wichita.edu

AIEMS President

Gamal Weheba, Ph.D.,
Professor and ASQ Fellow
Wichita State University
gamal.weheba@wichita.edu

Scope: The Journal of Management and Engineering Integration (JMEI) is a double-blind refereed journal dedicated to exploring the nexus of management and engineering issues of the day. JMEI publishes two issues per year, one in the Summer and another in Winter. The Journal's scope is to provide a forum where engineering and management professionals can share and exchange their ideas for the collaboration and integration of Management and Engineering research and publications. The journal will aim on targeting publications and research that emphasizes the integrative nature of business, management, computers and engineering within a global context.

Editorial Review Board

Sue Abdinnour Ph.D.
Wichita State University
sue.abdinnour@wichita.edu

Scott D. Swain, Ph.D.
Clemson University
sdswain@clemson.edu

Gordon Arbogast, Ph.D.
Jacksonville University
garboga@ju.edu

Alexandra Schönning, Ph.D.
University of North Florida
aschonni@unf.edu

Deborah Carstens, Ph.D.
Florida Institute of Technology
carstens@fit.edu

John Wang, Ph.D.
Montclair State University
wangj@montclair.edu

Hossein Cheraghi, Ph.D.
West New England University
cheraghi@wne.edu

Gamal Weheba, Ph.D.
Wichita State University,
gamal.weheba@wichita.edu

Nabin Sapkota, Ph.D.
Northwestern State University, LA
sapkotan@nsula.edu

Wei Zhan, D.Sc., PE
Texas A&M University
wei.zhan@tamu.edu

Reviewers

The Journal Editorial Team would like to thank the reviewers for their time and effort. The comments that we received were very constructive and detailed. They have been very helpful in our effort to continue to produce a top-quality journal. Your participation and timely response are very important for the success in providing a distinguished outlet for original articles.

Edwin Sawan, Ph.D., P.E.
Editor-in-Chief

Mohammed Ali

Ed Hernandez

Sherisse Pierre

Abdulaziz Abdulaziz

Byul Hur

Navya Ramakrishnan

Gordon Arbogast

Xiaochun JIANG

Dennis Ridley

Deborah Carstens

Chris Kluse

Kellie Schneider

Tristan Davison

Jingyun Li

Lauren Schumacher

Roelof de Vries

Huitian Lu

Kaustubh Sonawane

Isaac Gang

Arun Manoharan

Muhammad Rahman

Andrzej Gapinski

Holger Mauch

Brooke Wheeler

Thomas Gasque

Roger Merriman

Xun Xu Xu

Hongsheng He

Narasimha Nagaiah

Bayram Yildirim

Table of Contents

Collegiate Flight Students' Willingness to Pilot in Different Aircraft Sanitization Scenarios	1
A Novel Use of the K-mean Clustering Technique for Operational Efficiency in Electricity Distribution	9
Increasing General Aviation (GA) Pilot Reports (PIREPs) through Reducing Errors	17
Robust Cardiovascular Disease Prediction Using Logistic Regression	26
Comparing the Frequency of Reported Wildlife Strikes by Region in the United States	37
Impact of Organizational Slack on Innovation Outcomes	46
Leadership Styles and Lean Six Sigma Effectiveness within the Aerospace and Defense Manufacturing Industry	55
A study on Relocation of Home Field and Improvement of Game Schedules for Nippon Professional Baseball	67
Cost-Effective Attendance Management System Using Cloud Computing and Face Recognition	81
Review of Supply Chain Metrics to Support Performance Excellence	90

Building Design, Resident's Behavior, and Indoor Radon in the Commonwealth of Kentucky	105
Studying the Properties of Polymeric Composites of Metal Hydrides and Carbon Particles for Hydrogen Storage	119
3D Printing of Portland Cement Using a Binder Jetting System	128

Collegiate Flight Students' Willingness to Pilot in Different Aircraft Sanitization Scenarios

Brooke Wheeler¹

Jesse Shacknai¹

Nicholas Valluzzi¹

Tianhua Li¹

¹ *Florida Institute of Technology, College of Aeronautics*

bwheeler@fit.edu; jshacknai2017@my.fit.edu; nvalluzzi2017@my.fit.edu; tli2017@my.fit.edu

Abstract

The COVID-19 pandemic has added an entirely new dimension to flight education and training, thus highlighting the importance of understanding student perceptions of training in the close quarters of small, piston aircraft. Schools and institutions across the world have developed cleaning procedures and policies to mitigate the spread of COVID-19 while continuing collegiate programs. Flight schools face a greater challenge relating to the sanitization and the scheduling of an aircraft because social distancing in a light training aircraft is impossible, and an aircraft is particularly challenging to clean because of sensitive equipment and numerous touch surfaces.

We used a stratified cluster sample of Part 141 collegiate flight students from the core Aeronautics courses to recruit participants at all levels of training for a questionnaire. Ninety-seven responses were collected. The within-subjects design presented four aircraft sanitization scenarios in a random order to each participant to determine willingness to pilot (Rice et al., 2020) under various sanitization procedures: (1) assignment to an aircraft at random with no sanitization or scheduling procedures, (2) assignment to an aircraft at random with an Instructor who is responsible for the cleaning of the aircraft with alcohol wipes, (3) assignment to an aircraft that is specifically assigned to the same Instructor and his/her students, and (4) assignment to an aircraft that is specifically assigned to the same Instructor and his/her students every day and Instructor who is responsible for the cleaning of the aircraft with alcohol wipes. Sanitization scenarios with more precautions resulted in a higher willingness to pilot measurements.

1. Introduction

During the Spring 2020 semester, many universities transitioned to remote instruction due to the COVID-19 pandemic. Some courses, notably lab courses and flight training courses, require hands-on, in-person instruction. With the return to some face-to-face courses, it is important to monitor and understand student perceptions and comfort levels. Flight training is particularly concerning because it occurs in smaller, piston aircraft where the student and flight instructor are in very close quarters by necessity, the sensitive avionics are high-touch surfaces that are challenging to sanitize, and students

need to be comfortable and focused to maintain a high level of safety. Thus, it is important to increase our understanding of how students view pandemic precautions in order for flight schools to tailor their approach to promote both safety and student willingness to continue flight training.

The purpose of this study was to examine flight students' willingness to pilot aircraft in different sanitization scenarios as introduced during the COVID-19 pandemic. Willingness to Pilot (WTP) was measured using the scale developed by Rice et al. (2020). Participants at Part 141 collegiate flight program in Florida rated their willingness to pilot in four scenarios: (1) aircraft assignment at random with no sanitization or alternative scheduling procedures in place, (2) aircraft assignment at random with an Instructor who is provided with alcohol wipes and is responsible for the cleaning of the aircraft before and after all flight operations, (3) aircraft specifically assigned to the same Instructor and his/her students every day to minimize the number of people using each training aircraft, and (4) aircraft specifically assigned to the same Instructor and his/her students every day to minimize the number of people using each training aircraft, and the instructor is provided with alcohol wipes and is responsible for the cleaning of the aircraft before and after all flight operations.

The research question was as follows: Is there a difference in flight student WTP between the different aircraft sanitization scenarios? We hypothesized that students would have the highest WTP in the scenario with both enhanced cleaning and scheduling procedures, and the control scenario with no sanitization or scheduling procedures would have the lowest WTP.

Flight students across the United States are continuing their flight training throughout the COVID-19 pandemic, and pilot training is particularly critical now, especially because prior to the pandemic we were in a pilot shortage. Flight training must happen in close quarters, in simulators and smaller piston training aircraft. Thus, the future of aircraft sanitization and precautionary measures at flight schools will be of great importance in not only stemming the spread of typical viruses and bacteria but also in safeguarding against community spread in the on-going pandemic and in the event of another pandemic. In addition to minimizing the risk of spreading of disease, it is important to understand pilot comfort level operating under various precautions to assist flight schools in maintaining a positive training environment. The results of this study should be generalizable to all collegiate Part 141 flight schools because Part 141 training follows similar universal guidelines put forth by the FAA for Part 141 flight schools, and collegiate programs typically have similar student characteristics, though during a pandemic the local government health department messaging and case numbers differ. Therefore, the results should be applied at the host school, and then replicated in other flight schools.

2. Background

Flight schools, often a vital program to the universities and institutions that have them, are critical for economic development and the building of flight hours to enable students to advance their careers. Therefore, it is important that flight schools across the country and the world implement alternative measures to ensure the safety of instructors and students while flying during pandemics. This is challenging because avionics are sensitive and can only be cleaned with certain chemicals (e.g., 75-100% isopropyl alcohol), and training aircraft have both close quarters and a large number of high-touch surfaces in the cockpit. These conditions make it difficult to minimize the risk of contracting COVID-19.

To properly flight train, students must be comfortable with the situation they are in while flying the aircraft, so that they can focus on learning the skills they are practicing in a given lesson. Therefore, it is important to identify the ways to ensure aircraft are as clean as possible and to make sure that the students are able to focus on flight training rather than worrying about their health.

SARS-Cov-2, the virus that causes COVID-19 is primarily spread through droplets that are expelled by infected individuals when they cough or sneeze, or do anything that expels air, including singing, speaking, and breathing. However, the disease can also be spread by touching the face after contact with contaminated surfaces (CDC, 2021; WHO, 2021). There is also evidence of airborne spread in spaces with poor air circulation (CDC, 2021). The secondary mode of spreading SARS-Cov-2 is through common contact with an infected person's respiratory droplets, which can be projected onto surfaces through sneezing, coughing, speaking, or simply breathing (Public Health Agency of Canada, 2020). When within six feet of another person, the risk of contracting the coronavirus is much greater than if they were appropriately distanced (CDC, 2021). However, in-flight training, students and instructors typically sit in very close physical proximity, within three feet in a small training aircraft, for extended periods of time, and with low to no air circulation until the plane is in motion.

Additionally, once the virus is contracted, students may not realize they are sick and spread it before symptoms develop because of the relatively long incubation period of up to 14 days. Even if a student contracts the virus outside of the flight school, they may be able to spread it to other pilots and instructors affiliated with their flight program up to two to fourteen days later. The average incubation period is about 5.5 days (Lauer, et al., 2020). In the event someone is carrying the virus, respiratory droplets (containing SARS-Cov-2) can also be spread among several different surfaces, including training aircraft that are used by several instructors and students.

The amount of time COVID remains on a surface is directly affected by the type of surface. For example, COVID infected respiratory droplets on copper can last up to four hours whereas on cardboard, it can last up to twenty-four hours (van Doremalen, et al. 2020). The droplets on some metals, glass, and plastics have been shown to even last up to nine days (Kampf, et al. 2020). Many of these materials can be found within training aircraft. To combat the spread of the virus, Flight schools should be aware of the most vulnerable surfaces and materials. Plastic, glass, and metal surfaces, which can withhold the virus for upwards of nine days, can be found all over training aircraft.

Aviation as an industry has been working to adapt to the pandemic to clean and disinfect as effectively as possible. Proper cleaning and appropriate disinfection products should be considered in order to kill the virus on surfaces. ABM Industries, known for their worldwide janitorial and cleaning service, implemented a new program known as EnhancedClean. This program, designed to combat the COVID-19 pandemic, consists of three major steps: a comprehensive site assessment, site-specific cleaning suggestions, then the implementation of large-scale disinfection strategies (Baratti, 2020). This is a holistic approach to addressing the increased cleaning needs during the pandemic that can work for larger companies, particularly airports, airlines, and hospitals.

For flight schools, which are typically smaller aviation operations, there are several cleaning challenges. Many cleaning agents approved for use against SARS-Cov-2 are not safe to use in aircraft. One of the cleaners, recommended by both the aircraft and avionics manufacturers, is 70% Isopropyl

alcohol, which was incredibly difficult to source at the start of the pandemic. Additional sanitization and safety measures were also considered by flight schools. Scheduling aircraft to only be used by the same instructor and their students is a planning strategy to minimize cross-contamination of aircraft (I. Silver, personal communication, April 30, 2021). Several Part 141 flight schools considered and implemented these and similar plans, making these scenarios ideal for study.

The science of coronavirus mitigation is clear: when social distancing is not possible, mask mandates and distancing should be used, as well as medical and industrial products to sanitize surfaces and lessen the risk of contracting the disease. Flight schools, like several other companies and organizations, have considered increasing sanitization and changing procedures to adapt to the global pandemic. There is likely a level of concern regarding aircraft sanitation among flight students, but to date, we have not measured student perceptions of safety under any of the proposed or implemented sanitization options. Student willingness to pilot during the pandemic and under different aircraft sanitization scenarios has not been examined.

3. Methodology

The research employed a within-subjects design using a questionnaire. We presented all aircraft sanitization scenarios to each participant to examine differences in WTP.

3.1. Participants, sampling, and instrument

The target population was collegiate Part 141 flight students in the United States. The accessible population was 275 students flight students enrolled in core aeronautics courses within a Part 141 collegiate flight program at a university in Florida. Respondents were recruited from Aeronautics 1-4 and Instructional Techniques courses. The IRB exemption was approved (20-096), and all data were collected electronically with remote recruitment to minimize the risk of COVID-19 transmission. Participation was both anonymous and voluntary, and each participant must have been at least 18 years of age. We asked for student's participation during a Zoom conference within their allotted class time.

The sampling method was a stratified cluster sample; participants were recruited from existing core aeronautics course sections to provide a cross-section flight training experience from student pilots through Certified Flight Instructor ground school. After receiving permission from the professors of the core aeronautics classes, participants were recruited from each section during their regular class Zoom meeting. Data were collected using Qualtrics to administer a questionnaire, which included the WTP scale (Rice et al., 2020) for each of the four aircraft sanitization scenarios, as well as demographic questions. The scenarios were designed based on what the collegiate flight program and other Part 141 programs initially implemented as a part of their pandemic response. Scenario 1 represents a control scenario with no precautions, essentially the equivalent of normal operations before the start of the pandemic. Scenario 2 includes the Certified Flight Instructor sanitizing the aircraft with wipes before and after each lesson. Scenario 3 assigns each aircraft to a specific instructor and their students to minimize crossover contamination between students using the aircraft. Scenario 4 combines the measures in scenarios 2 and 3. The link to the questionnaire was distributed via chat in the Zoom meeting. The survey was set to counterbalance the order of the four

scenarios to each participant to eliminate any order effect. The scale was validated for use in a pilot population.

3.2. Data Analysis

The data were collected using a questionnaire and were then downloaded from Qualtrics into Excel. Responses to the WTP scale were converted from Likert Scale to numerical values as follows: Strongly disagree -2, Disagree -1, Neutral 0, Agree 1, and Strongly Agree 2. Cronbach's alpha was calculated to determine the internal consistency of the scale. High internal consistency allowed calculation of WTP for each participant by averaging the responses to all items on the WTP scale to produce one value for participant WTP. Descriptive statistics were calculated in Excel for each of the four scenarios. Inferential statistics, including an ANOVA, Tukey's pairwise comparison, and η^2 , were calculated in RStudio version 1.3.

4. Results

We recruited from a total of 156 students currently enrolled in the Aeronautics 1-4 classes and Instructional Techniques class. Of the students we recruited, 118 of them participated in the study. However, there was some mortality, with 21 participants stopping before completing the questionnaire, resulting in a total of 97 completed questionnaires ($N = 97$). This yielded a response rate of 63%. Age varied but most of the participants fell within the 18-23 range. In terms of gender, there were more male participants than female participants. Of the participants, the numbers of the students who were taking (a) Aeronautics 1, (b) Aeronautics 2, (c) Aeronautics 3, (d) Aeronautics 4, and (e) Instructional Techniques were 68, 27, 31, 24, and 6, respectively. This distribution captured both the range of experience of collegiate flight students and was also representative of the number of students at each stage of flight training. Cronbach's alpha was 0.97 for the first scenario, 0.98 for the second and third, and 0.96 for the fourth. This showed a very high degree of internal consistency, thus we proceeded with analyses using the average of the seven items on the scale as the WTP measurement for each scenario.

Table 1 shows the descriptive statistics for WTP by aircraft sanitization scenario. WTP ranged from strongly disagree (-2) to strongly agree (2) in all four scenarios. Figure 1 also illustrates the differences in average WTP across scenarios. The scenario with the lowest average WTP was scenario 1, whereas the highest WTP was in scenario 4. There was a noticeable difference between the WTP in scenarios 1 and 2 and 1 and 3; however, scenarios 2 and 3 had less of a difference.

The repeated measures one-way ANOVA was significant, indicating an effect of aircraft sanitization scenario on WTP: $F(3, 187) = 62.33, p < .001$. Therefore, a post hoc analysis was conducted to determine which scenarios differed. Tukey's pairwise comparison showed that most aircraft sanitization scenario pairs were significantly different at $p < .05$, except for scenarios 2 and 3, the random aircraft with specific instructor cleaning and the specific aircraft assignment with no cleaning, respectively (see Table 2). Finally, the eta squared ($\eta^2 = 0.34$) indicated a medium to large effect size.

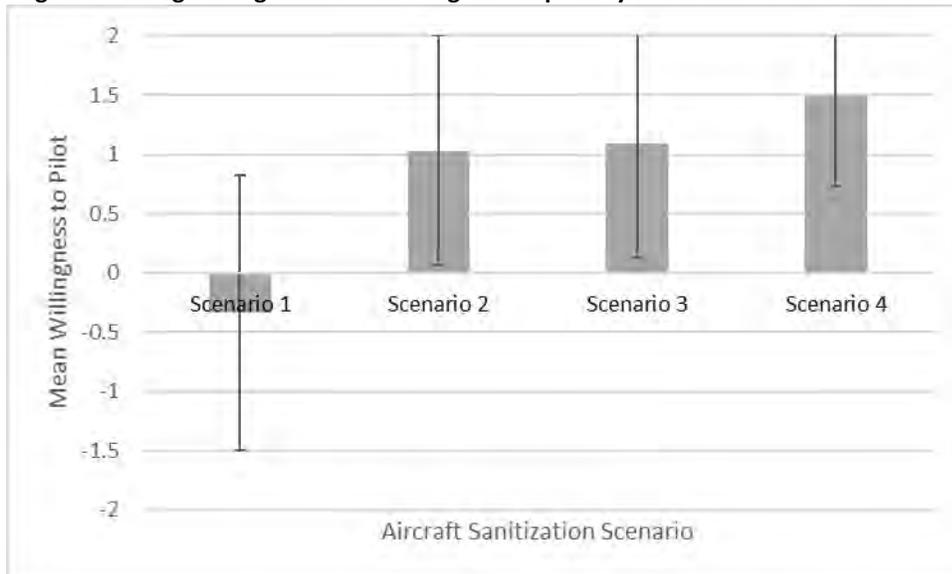
Table 1. Descriptive statistics for willingness to pilot by aircraft sanitization scenario during COVID-19.

Aircraft Sanitization Scenario	Mean	Median	Mode	Standard Deviation
Scenario 1. No sanitization or alternative scheduling procedures	-0.34	-0.71	-1	1.16
Scenario 2. Random aircraft assignment with an instructor who is provided personal sanitizing equipment	1.03	1	2	0.97
Scenario 3. Specifically assigned aircraft to limit the number of people in and out of an airplane, with an instructor who is not provided personal sanitizing equipment	1.09	1	2	0.96
Scenario 4. Specifically assigned aircraft to limit the number of people in and out of an airplane, with an instructor provided with personal sanitizing equipment	1.49	2	2	0.76

Table 2. Tukey pairwise comparison of Aircraft Sanitization Scenarios

Scenario Pair	<i>p</i>
1-2	<0.001
1-3	<0.001
1-4	<0.001
2-3	0.97
2-4	0.01
3-4	0.02

Figure 1. Collegiate flight student willingness to pilot by aircraft sanitization scenario



Note: Error bars represent one standard deviation. Scenario 1 was the only scenario in which students were not willing to pilot. Willingness to pilot (Rice, 2020) is on a Likert scale from strongly agree (2) to strongly disagree (-2).

5. Discussion

With 97 responses and a response rate of 63%, the participation exceeded the a priori power calculation and anticipated participation. The distribution of responses across core aeronautics classes was also representative of the flight student population at Part 141 collegiate program. Thus, while the surveys were anonymous, the sample was representative of the student body across flight experience and training levels.

The ANOVA demonstrated that there is a statistically significant difference in WTP by aircraft sanitization scenario ($p < .001$), and the effect size indicates that this is indeed a substantive difference in student WTP ($\eta^2 = 0.34$). As expected, Scenario 1, the control scenario with no COVID-19 precautions, had a significantly lower WTP than all the other scenarios. It was also notably the only negative WTP, indicating that student pilots were not comfortable flying under the control aircraft sanitization scenario, which aligned with the typical practices of all Part 141 flight schools before the pandemic. Scenario 4 (both cleaning and specific aircraft assignment) had a higher WTP than all other scenarios, though scenarios 2, 3, and 4 all had a positive WTP. However, there was no significant difference in WTP between scenarios 2 and 3 (Table 2).

In practice, these results affirm the importance of instituting precautions during a pandemic. Flight training occurs in small piston aircraft where social distancing is impossible. Airflow is dependent on the aircraft being in motion. To maintain a training environment where flight students feel comfortable, it is advisable to take multiple precautions. Scenarios 2 and 3 each implemented one measure: sanitization of the aircraft by the instructor or specifically assigning the aircraft to only one instructor (and his or her students). While each of these measures shifted the WTP from negative to positive, scenario 4, which implemented both measures, resulted in the highest WTP. If feasible, we recommend implementing multiple measures to promote student WTP. Measures should be informed by science and adapted as new information becomes available. In the case of COVID-19, multiple measures including specific aircraft scheduling and use of masks whenever possible would be advisable.

If this research is replicated later, the results may vary, particularly as local case numbers and risk levels increase and decrease. Individual participants' views will change over time due to their personal experiences with the pandemic (e.g., if they, a family member, or a friend contracts COVID-19, after vaccination) and to our evolving understanding of COVID-19. The responses are self-report measures; however, anonymity should allow participants to feel comfortable answering honestly. Also, the WTP scale has been previously validated (Rice et al., 2020) and had an extremely high internal validity, indicating the consistency of participants' responses to items on the scale.

This is a relevant research topic because the COVID-19 pandemic is a dynamic situation. Flight schools have responded following government guidelines as well as the recommendations and requirements of their institutions as the pandemic progresses. Every flight school nationwide is attempting to determine the best way to navigate through this while continuing to train pilots as safely as possible. Promoting a training environment where students feel as comfortable as possible is critical to allowing for training to continue. This study demonstrated that with two aircraft sanitization procedures in place, student WTP was higher than either scenario with just one measure implemented. The control situation with no COVID-19 precautions generated a negative WTP, which points to the need to adapt and actively take measures to minimize risk during a pandemic.

Future research should repeat this study at this flight school later in the pandemic and at other Part 141 flight schools to determine the findings' generalizability. Studies could also build on this work by including other pilots, such as flight instructors, airline pilots, and corporate pilots, to determine if there are similar WTP trends in these more experienced pilot groups, keeping in mind the differences between

training in piston aircraft and jets. We also need to continue research on how to minimize disease transmission in aircraft for both flight training and commercial transportation so that the most effective options can be selected. This research will give flight departments in the United States a better understanding of how to create a training environment that is as safe and comfortable as possible for students so that they can focus on learning to pilot aircraft, even during a pandemic.

It is evident that implementing both sanitization and scheduling procedures result in a higher WTP. We recommend that flight schools consider what practices may be appropriate for them and be prepared to quickly implement multiple measures in the event of a future pandemic, or even a serious flu season. If flight schools can implement and advertise sanitization measures consistent with a higher student WTP, then these schools will benefit in both student retention and recruitment as people are drawn to safety, especially in the aviation field.

6. References

- Baratti, L. (2020, August 26). The EnhancedClean Program Ensuring Airports and Airplanes Are Safe Amid COVID-19 Pandemic. TravelPulse. <https://www.travelpulse.com/news/features/the-enhancedclean-program-ensuring-airports-and-airplanes-are-safe-amid-covid-19-pandemic.html>
- CDC. (2021). How COVID-19 Spreads. Center for Disease Control. Accessed on April 30, 2021, at <https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/how-covid-spreads.html>
- Kampf, G., Todt, D., Pfaender, S., & Steinmann, E. (2020). Persistence of coronaviruses on inanimate surfaces and their inactivation with biocidal agents. *Journal of Hospital Infection*, 104(3), 246–251. <https://doi.org/10.1016/j.jhin.2020.01.022>
- Lauer, S. A., Grantz, K. H., Bi, Q., Jones, F. K., Zheng, Q., Meredith, H. R., ... Lessler, J. (2020). The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Annals of Internal Medicine*, 172(9), 577–582. <https://doi.org/10.7326/m20-0504>
- Public Health Agency of Canada. (2020, April 4). Coronavirus disease (COVID-19): Prevention and risks. Government of Canada. <https://www.canada.ca/en/public-health/services/diseases/2019-novel-coronavirus-infection/prevention-risks.html>
- Rice, S., Winter, S., Capps, J., Trombley, J., Robbins, J., Milner, M., & Lamb, T. (2020). Creation of Two Valid Scales: Willingness to Fly in an Aircraft and Willingness to Pilot an Aircraft. *International Journal of Aviation, Aeronautics, and Aerospace*. <https://doi.org/10.15394/ijaaa.2020.1440>
- van Doremalen, N., Bushmaker, T., Morris, D. H., Holbrook, M. G., Gamble, A., Williamson, B. N., ... Munster, V. J. (2020). Aerosol and surface stability of HCoV-19 (SARS-CoV-2) compared to SARS-CoV-1. *The New England Journal of Medicine*. <https://doi.org/10.1101/2020.03.09.20033217>
- WHO. (2021). Coronavirus disease (COVID-19): How is it transmitted? World Health Organization. Accessed on April 30, 2021, at <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/question-and-answers-hub/q-a-detail/coronavirus-disease-covid-19-how-is-it-transmitted>

A Novel Use of the K-mean Clustering Technique for Operational Efficiency in Electricity Distribution

Anyama Hanson Tetley¹
Kim Hensley, PE
Ann Schrimpscher
Dongsheng Wu, PhD

¹anyama.tetley@hsvutil.org; Kim.Hensley@hsvutil.org; ann.schrimpscher@gmail.com; dw0001@uah.edu

Abstract

Observed data shows that extreme temperature is the main driver of electricity demand in locations of the world where seasonal variations in the weather are a yearly occurrence. Although utility planners and management are aware of this fact, setting expectations and load projections during extreme temperature events have been a challenge. Extreme temperatures in the winter, usually have a lot of variations with very infrequent occurrences for certain events. Despite these uncertainties, it is imperative to design the electrical systems to be ready for all uncertain events that may come up. This brings about the consideration of economical, safe, and efficient alternatives to shave off or manage the peak demand during extreme temperature conditions. Modeling the behavior of temperature and demand becomes very important for decision-making in this regard. We show that the first step needed to develop such a model of temperature and demand is to split days into groups that share similar temperature trends. The K-mean clustering technique is used in a novel way to do this. The paper gives insight into some interesting behavioral trends and patterns observed in historical temperature and demand data and sets a stage for effective forecasting of electric demand in the industry

1. Introduction

The effect of the weather, notably the extreme states affects so many human systems. (Bloesch & Gourio, 2015; Lesk et al., 2016; Rezaee et al., 2016) discuss some effects on Agriculture. In a health-related direction, It is estimated that the number of annual deaths attributable to cold temperatures is 0.8% of the average annual deaths of the sample of deaths used for the study (Deschênes & Moretti, 2009). Carder et al use Poisson regression models to show that mortality increases steeply at temperatures below 11°C in Scotland (Carder et al., 2005).

The effect on the transportation sector cannot go unmentioned. A time-series study on the effects of cold temperature on road traffic injuries shows that there is a sharp increase below freezing temperatures (Lee et al., 2014).

The power system industry is another area plagued with so many complexities during extreme weather. This paper focuses on the effect of extreme temperatures on the demand trends of electricity for a utility company. In a previous publication by the author, there is a discussion on how extreme temperatures place great demands on electrical energy sources mostly for heating and cooling in the

winter and summer temperatures respectively. There is also an adverse effect on the operational efficiency of generator turbines and engines. As if that is not enough, manufacturers of heat pumps and service providers of air conditioning services have expressed that heat pumps work best when temperatures are above 40°F. Once outdoor temperatures drop to 40 degrees, heat pumps start losing efficiency, and they consume more energy to do their jobs. Figure 1. below confirms this assertion by the manufacturers using data for a distribution circuit in Huntsville Utilities. A plot of demand in Amps taken at 15-minute intervals against temperature during the winter of 2021, shows a negatively correlated nonlinear trend with an upward and slightly steeper curve as the temperature decreases.

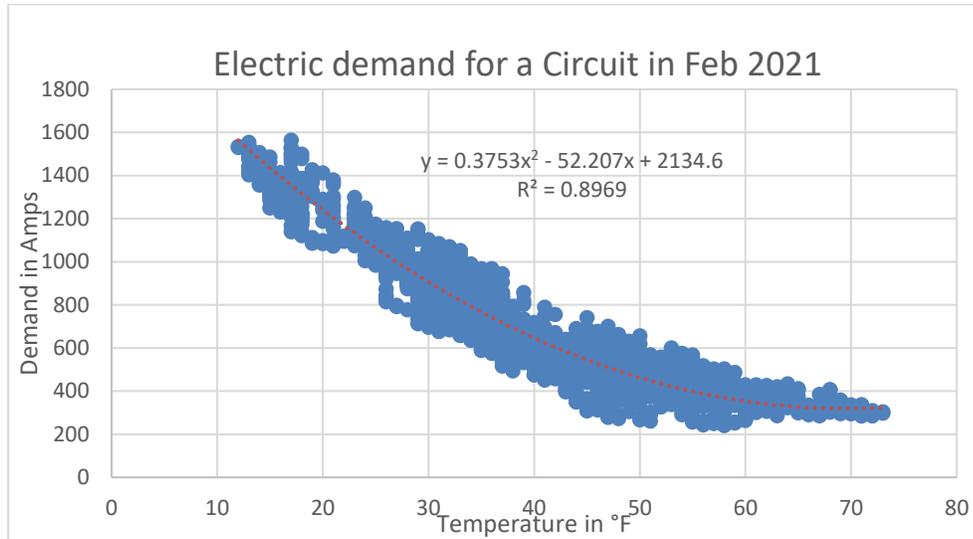


Figure 1. Trend in Electric Demand for a circuit in February 2021

The severe cold in Texas this year was not new (Nazir, 2021). Documentation shows that such severe cold temperatures were experienced most notably in 1899, 1951, 1983, 1989, and 2011; however, what was seen in February 2021 was the most intense in 30 years (Doss-Gollin et al., 2021). The recent Texas experience resonates the fact that during such cold temperatures, lack of electrical energy is very costly and can even result in loss of lives. The management of electric utility companies wants to avoid any outage or power disruption during such periods.

Fortunately, the weather forecast predicts the expected daily temperatures well enough. Can such information be made useful to the utility planner? Observed data shows major differences in trends between temperature and electricity demand in the winter and summer seasons. We show that even within the winter, different temperature ranges in a day attract different behavioral patterns. However, in most studies that attempt to model such relationships, these differences are rarely acknowledged.

The discussion above shows the need for an in-depth exploration of trends in demand and temperature during extreme weather. We use data from a utility company to point out some interesting trends. The discussion will lead to the need to group days according to a similarity in temperature for further modeling and forecasting if desired.

2. Exploratory analysis

For the Power System planner, It is relatively easy to plan for summer occurrences because variability in maximum temperatures is not much. Table 1. shows summer maximum and winter minimum temperatures in Huntsville over the past 6 years. There are very consistent values with a range of 4°F in the summer. The same cannot be said for the winter temperatures. Not surprisingly a big range in the peak demand for the winter dataset compared to the summer is also observed.

Table 1. Variations in Summer max and Winter min temperatures

Season	Year						Variance
	2015	2016	2017	2018	2019	2020	
Summer maximums in °F	99	98	96	95	98	97	2.2
Winter minimums in °F	8	17	12	7	22	22	44.7

Trends within the day are also very different in the summer compared to the winter. Even though not shown and discussed in detail in this paper, individual plots of demand against temperature for each of the different clusters during the winter gave good linear fits with higher beta coefficients in the extreme cold temperatures.

The data shows that temperatures may go from very cold in the morning to warmer states in the afternoon and vice versa in the winter. On the other hand, hardly do we have maximum temperatures in the morning being higher than the afternoon equivalent in the summer.

We observed an interesting pattern in the winter dataset, where the morning minimum temperatures and not the daily minimums were better predictors of the daily peak. In the summer it was vice versa, the afternoon maximum temperatures (usually the daily maximums) were better predictors of the peak. Table 2. shows the R-squared values for the different scenarios. Due to space constraints, the graph is shown only for the winter scenario in Figure 2.

Table 2. Predicting demand with morning/ afternoon temperatures in the summer and winter

Graph axes		R ²
Summer (Linear)	Peak demand against morning max temps	0.27
	Peak demand against afternoon max temps	0.69
Winter (Nonlinear)	Peak demand against morning min temps	0.93
	Peak demand against afternoon temps	0.49

In their study to predict the response of electricity load due to climate change, the range of temperatures between 16°F and 20°F were classified as noisy and therefore excluded in the analysis (Sullivan et al., 2015). Their study and our discussion above suggest the need to cluster days based on a range of temperature similarity within groups if it is desired to forecast electric load demand accurately. This is done using the k-mean clustering algorithm in the next section.

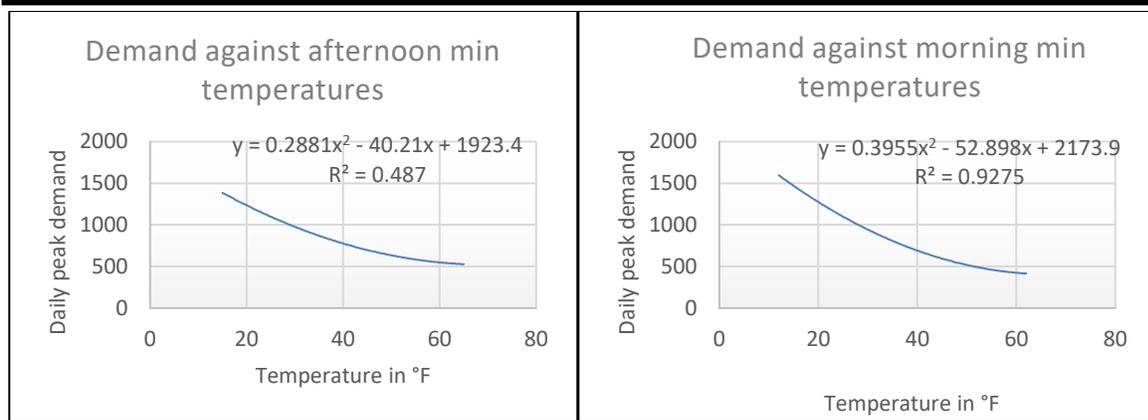


Figure 2. A comparison of demand against morning and afternoon minimum winter temperatures

3. K-mean Clustering

K-means is the most widely used and studied clustering algorithm. Given a set of n data points in real d -dimensional space (R_d), and an integer k , the clustering problem is to determine a set of k points in R_d , the set of points is called cluster centers, the set of n data points are divided into k groups based on the distance between them and cluster centers (Vora & Oza, 2013). The technique minimizes within-cluster variance.

The approach is a nonhierarchical clustering method that is designed to group items rather than variables into a collection of K clusters. The term k -means was suggested by MacQueen, who used it to describe an algorithm he was working on that assigns each item to the cluster having the nearest centroid (MacQueen, 1967). (Johnson & Wichern, 2002) summarizes the k -mean procedure into 3 main steps.

1. Partition the items into K initial clusters.
2. Going through each of the items, reassign items to clusters whose centroids are nearest using Euclidean distance as a measure of distance. After a reassignment, recalculate the centroid for the cluster receiving the new item and the one losing the item.
3. Repeat step 2 until no more reassignments can be made.

In this work, the technique is used to group days based on 2-dimensional temperature data.

3.1. Hypotheses Testing

To show that there are differences in demand that can be attributed to temperature, we hypothesize that the mean daily demand is the same irrespective of the cluster. In other words, the temperature has no significant impact on the demand. We use the single-sample t -test to determine just how certain we can be that our hypothesis is true. The populations are approximately normal, and so we use the student's t distribution.

In situations where normality cannot be ascertained, non-parametric tests can be done.

The null and alternate hypotheses are below.

- $H_o: \mu_1 = \mu_2 = \mu_3 = \mu_0$
- $H_a: \text{Mean demand is not equal to } \mu_i \text{ (overall mean) where,}$

μ_1 , μ_2 , and μ_3 , are the mean demands for clusters 1, 2, and 3 respectively and μ_0 is the overall mean.

4. Methodology

The first step is to identify periods of extreme temperatures of interest to the distribution system planner. Though other weather variables such as wind speed, insolation, and humidity invariably influence electric demand, we block their effect by doing our analysis with the same month in the season, in consecutive years. The idea is akin to blocking in experimental design.

For example, instead of using 3 months of summer data during June, July, and August in a season, we choose to use data in August over a 3-year period just to block out the effect of other temperature variables that may be different for different months. In the future, a maximum of 5-years of consecutive data will be used because growth and other behavioral trends often cause differences in demand after about 5 years period.

Once the historical data is decided on, historical temperature data of interest can be collected. For our purposes, we are not just interested in the daily max/ mins. Two daily peaks/ mins are of interest.

- The peak during the first half of the day: 12:00 AM – 11:59 AM and the peak in the second half of the day: 12:00 noon to 11:59 PM.

This gives us a two-dimensional vector that results in a better characterization of days based on temperature. As was seen in Figure 2, the minimum temperature in the day is not necessarily the best predictor of demand.

5. Analysis and Results

K-mean clustering in Matlab, with K=3, resulted in the clustering groups shown in Figure 3. for February 2020 and 2021 temperatures.

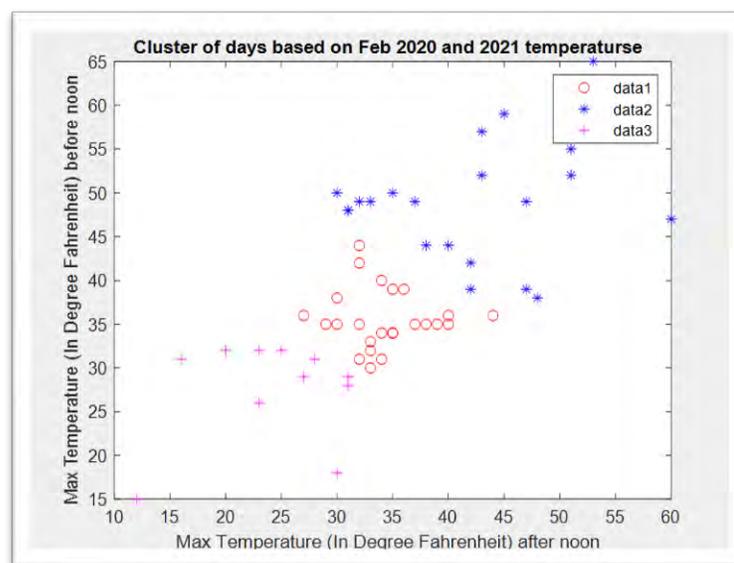


Figure 3. k-mean clustering for February 2020 and 2021 temperature data

Thus, when looking for trends and patterns in the temperature data, we should bear in mind that days shown as * will most likely show different relationships compared to those shown with the + sign. A linear plot of demand against temperature for February 2020 gave a beta coefficient of -25.4 for the extreme lower temperatures in the lower left corner of Figure 3., and -18.8 for temperatures in the middle cluster.

The algorithm was also used to cluster 2019 August temperature data. The results were three days in cluster 1, where maximum temperatures before noon and afternoon hovered around the 70's (°F) with a maximum of 85° F in the afternoon. Demand values less than 1000A were observed. In clusters 2 and 3, we observed higher temperatures and corresponding higher demands, (mostly above 1000A)

The single sample t-test was used to ascertain how certain we can be that the mean demand (for all clusters) is not the same and equal to 1190A (The overall mean).

Null and alternate hypotheses:

$$H_0: \mu_1 = \mu_2 = \mu_3 = 1190A$$

H_a: Mean demand is different for at least one of the clusters

The results in Table 3. below, show a non-significant difference in demand during days in cluster 2 but significant differences for clusters 1 and 3.

Table 3. Results of Hypotheses Test

	Test Value = 1189.6					
	T-statistic	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Cluster 1	-5.3	2	0.0	-295.7	-536.2	-55.1
Cluster 2	-1.0	19	0.3	-21.2	-64.7	22.3
Cluster 3	7.6	7	0.0	163.9	113.1	214.7

The results show the need to adjust for different temperature ranges when one wants to forecast electric demand using temperature as a predictor.

6. Discussion and Conclusion

This paper used data from a utility company to study the effect of temperature on electricity demand. Differences in trends for the summer and winter seasons as well as differences across different temperature ranges in a day led to a conclusion that it will be necessary to cluster days according to temperature similarities within groups.

K-mean clustering of daily extreme temperature is used in a novel way to group days based on

similarity in temperature within groups. This was done for August 2019 and February 2020/ 2021. For the August data, the algorithm grouped 3 days as cluster 1, 20 as cluster 2, and 8 days as cluster 3. Giving probabilities of 0.1, 0.6, and 0.3 respectively that such events will occur. The results show that there is a 0.3 probability that temperatures reach some high bounds.

Differences in mean demand values were observed for the 3 clusters and the single sample hypotheses test was used to ascertain the significance of the difference. There was no significant difference between the overall mean and the mean for cluster 2. The implication is that if we have technical issues at such temperatures associated with cluster 2, our system will be very unreliable. The probability of a day in August being in cluster 2 is close to 0.65. Therefore, the base design of the system should at least comfortably handle demands on such days.

Significantly higher electric demand is observed for cluster 3. Cluster 1 is not too big a deal. If the system can handle cluster 2 then we should be fine during days of lower demands. During days in cluster 3, peak temperature rises to or above 90°F before noon and stays that way in the afternoon. These are the temperatures during which cooling our homes, offices, etc. cannot be a luxury or an option. The same reasoning goes for heating in the winter temperatures.

Thus, for the System planner in Huntsville, when a forecast of a maximum temperature of 90°F or above is made for a day in August, preparations should be made to manage unusually higher demands.

Recommendations may include encouraging lower electric energy use (for customers who have options) and using gas or other available sources instead during such extreme temperatures as observed for cluster 3. A practical implication may be to hold off using an electric laundry dryer at such a time. This results in increased utilization and reliability of the installed electrical equipment, which are essential attributes to give the customer value for his or her money.

Finally, the clustering algorithm establishes the need to consider looking at trends based on temperature similarity to achieve accurate results in any forecasting study.

7. Limitations and Future work

One limitation of this work is the sample size of data used for analysis. Future work will look at increasing it. The caveat is to limit the number of consecutive years to about five (5) maxima since loading patterns usually change within a 5-year time frame and this will significantly affect the comparison of mean demand for the years under study.

An attempt will also be made to come up with a model to predict peak demand based on forecasted temperature. Even though Huntsville is growing at a fast rate we have not seen single-digit temperatures since 2018 and it is very important to be able to predict current demand accurately in such temperatures to avoid situations such as the harsh impact of the cold temperatures on the power system in Texas in February 2021.

Future work will also attempt to account for the impact of extreme temperature on consecutive

days when establishing trends and patterns. If peak temperatures are going to be above 90 for three (3) consecutive days, we will most likely expect higher demand on the third day compared to the first day. Last but not least the authors will consider the evaluation for an optimal K as part of future work.

8. References

- Bloesch, J., & Gourio, F. (2015). The effect of winter weather on US economic activity. *Economic Perspectives*, 39(1).
- Carder, M., McNamee, R., Beverland, I., Elton, R., Cohen, G. R., Boyd, J., & Agius, R. M. (2005). The lagged effect of cold temperature and wind chill on cardiorespiratory mortality in Scotland. *Occupational and Environmental Medicine*, 62(10), 702–710. <https://doi.org/10.1136/oem.2004.016394>.
- Deschênes, O., & Moretti, E. (2009). Extreme Weather Events, Mortality, and Migration. *The Review of Economics and Statistics*, 91(4), 659–681. <https://doi.org/10.1162/rest.91.4.659>.
- Doss-Gollin, J., Farnham, D. J., Lall, U., & Modi, V. (2021). How unprecedented was the February 2021 Texas cold snap? *Environmental Research Letters*, Volume 16, Number 6. <https://iopscience.iop.org/article/10.1088/1748-9326/ac0278>
- Johnson, R. A., & Wichern, D. W. (2002). *Applied multivariate statistical analysis* (Vol. 5, Issue 8). Prentice hall Upper Saddle River, NJ.
- Lee, W.-K., Lee, H.-A., Hwang, S., Kim, H., Lim, Y.-H., Hong, Y.-C., Ha, E.-H., & Park, H. (2014). A time-series study on the effects of cold temperature on road traffic injuries in Seoul, Korea. *Environmental Research*, 132, 290–296. <https://doi.org/10.1016/j.envres.2014.04.019>
- Lesk, C., Rowhani, P., & Ramankutty, N. (2016). Influence of extreme weather disasters on global crop production. *Nature*, 529(7584), 84–87. <https://doi.org/10.1038/nature16467>
- Nazir, H. (2021). Lessons Learned from the February 2021 Texas Power Outage. Canadian Energy Research Institute (CERI). [https://ceri.ca/assets/files/Electricity%20Commodity%20Update April%202021.pdf](https://ceri.ca/assets/files/Electricity%20Commodity%20Update%20April%202021.pdf)
- Rezaee, S., Pelot, R., & Ghasemi, A. (2016). The effect of extreme weather conditions on commercial fishing activities and vessel incidents in Atlantic Canada. *Ocean & Coastal Management*, 130, 115–127. <https://doi.org/10.1016/j.ocecoaman.2016.05.011>
- Sullivan, P. [National R. E. L. (NREL), Golden, CO (United States)], Colman, J. [National R. E. L. (NREL), Golden, CO (United States)], & Kalendra, E. [National R. E. L. (NREL), Golden, CO (United States)]. (2015). Predicting the Response of Electricity Load to Climate Change. <https://doi.org/10.2172/1215283>
- Vora, P., & Oza, B. (2013). A survey on k-mean clustering and particle swarm optimization. *International Journal of Science and Modern Engineering*, 1(3), 24–26.

Increasing General Aviation (GA) Pilot Reports (PIREPs) through Reducing Errors

Deborah S. Carstens, Ph.D., PMP¹ (Corresponding Author)

Michael S. Harwin, J.D., M.S.¹

Tianhua Li, Ph.D.¹

¹ Florida Institute of Technology, Melbourne, FL. 32901

carstens@fit.edu (Corresponding Author E-mail)

Abstract

The purpose of this study was to increase General Aviation (GA) pilot reports (PIREPs) through reducing errors, trust issues, and inaccuracies in PIREP submittals that prevent inclusion in the PIREP system. PIREPs are an essential source for pilots because they contain the most current weather conditions and help pilots avoid flying into unforeseen hazardous weather. PIREPs can also provide vital information to meteorologists who develop aviation forecasts. A Failure Modes and Effects Analysis (FMEA) was conducted to identify failure modes for two PIREP processes consisting of air traffic control (ATC) processes failures and airmen induced issues contributing to the failure to submit PIREPs. After the failure modes were identified, recommended actions to eliminate or reduce the error or error effects were determined. The results, discussion, and future research are also addressed.

Acknowledgment: This research was sponsored by the Federal Aviation Administration (FAA), NextGen Weather Technology in the Cockpit (WTIC) Research Program.

1. Introduction

The purpose of the study was to increase GA PIREPs by reducing errors, trust issues, and inaccuracies in PIREP submittals that prevent inclusion in the National Airspace (NAS). A FMEA was conducted to identify failure modes for two PIREP processes consisting of ATC process failures and airmen-induced issues contributing to a failure to submit PIREPs. An FMEA is a systematic method to evaluate processes and to determine where failures may occur. The FMEA resulted in recommended actions to eliminate or reduce the error or error effects to increase the volume of PIREPs being submitted by air traffic controllers and GA pilots. PIREPs are pilot reports from pilots describing their actual in-flight weather conditions. PIREPs can increase the accuracy of forecasted weather and provide pilots with real-time weather conditions. Submission of PIREPs can allow aircraft to avoid hazardous weather and prevent weather-related incidents and accidents. An accurate weather forecast and current weather are critical for all aircraft operating in the NAS (NTSB, 2017a). PIREPs are an essential source for pilots because they contain the most current weather conditions and help pilots avoid flying into unpredictable hazardous weather. PIREPs can also provide vital information to meteorologists who develop aviation forecasts. Aircraft Owners and Pilots Association (AOPA) identified accidents caused by pilots flying under visual

flight rules (VFR) and continuing into instrument meteorological conditions (IMC) that can be fatal (see Table 1) (2015; 2016a, 2017, 2018, 2021a, 2021b, 2021c).

Table 1

VFR into IMC accidents and fatalities from 2012 to 2018 documented by the Nall Reports

Report	Year	Involved	Fatalities
24th Nall Report (AOPA Air Safety Institute, 2015)	2012	23	22 (95.6%)
25th Nall Report (AOPA Air Safety Institute, 2016a)	2013	33	17 (73.9%)
26th Nall Report (AOPA Air Safety Institute, 2017)	2014	22	20 (90.9%)
27th Nall Report (AOPA Air Safety Institute, 2018)	2015	21	20 (95.2%)
28th Nall Report (AOPA Air Safety Institute, 2021a)	2016	13	7 (53.8%)
29th Nall Report (AOPA Air Safety Institute, 2021b)	2017	28	22 (78.5%)
30th Nall Report (AOPA Air Safety Institute, 2021c)	2018	14	13 (92.8%)

Note. The values in parentheses show the percentage of the fatalities.

The failure of pilots to submit PIREPs falls into seven categories: Lack of awareness of the importance of PIREPs, lack of confidence in PIREP format, lack of weather assessment skills, cockpit workload, fear of enforcement action, a cumbersome and time-consuming reporting process, and prior experience with PIREPs not being disseminated (NTSB, 2017a). The failure of air traffic controllers to enter PIREPs into the system is classified into four areas: noncompliance with solicitation requirements, inadequate dissemination of weather information, data entry mistakes, and consolidating multiple PIREPs.

2. Background Information

The NTSB (2017a) discussed the importance of entering PIREPs into the system in their special investigation report: Improving Pilot Weather Report Submissions and Dissemination to Benefit Safety in the NAS. The report identified deficiencies with air traffic controllers and airmen as contributing to the failure of PIREPs to be entered into the system.

2.1 Air Traffic Control Issues

If pilots submit a PIREP to an air traffic controller, the controller must enter the PIREP into the system. Air traffic controllers are also required to advise local facilities of reported PIREPs within their control area (NTSB, 2017a). Additionally, air traffic controllers must solicit PIREPs when conditions are reported below specific weather standards (FAA, 2019a; NTSB, 2017a). However, separating and issuing safety alerts to aircraft is ATC's primary duty and takes precedence over the requirement to solicit PIREPs (FAA, 2019b; NTSB, 2017a). The NTSB (2017a) classified ATC PIREP failures into four categories: (1) noncompliance with solicitation requirements, (2) inadequate dissemination of both urgent and routine weather information, (3) data entry errors, and (4) inappropriate consolidation of multiple reports.

When certain weather conditions exist, air traffic controllers must solicit PIREPs (FAA, 2019a; NTSB, 2017a). In February of 2015, in Andrews, Texas, an airplane on approach impacted the ground short of the runway because of inflight icing. The air traffic controller did not solicit PIREPs from aircraft arriving and departing the airport when weather conditions met the minimum solicitation requirements. The NTSB

found air traffic controllers frequently cannot solicit PIREPs because of the high workload (NTSB, 2015; NTSB, 2017a).

The failure of air traffic controllers to disseminate PIREPs after their receipt may be a causal factor in an incident or accident because a pilot may make a poor adverse weather avoidance decision (NTSB, 2017a). A Denver Air Route Traffic Control Centers (ARTCC) received a PIREP from a Boeing 737 reporting hail potential but failed to respond to an Airbus 320 crew's repeated request for information regarding hazardous weather (NTSB, 2017b). The Airbus 320 encountered hail that resulted in a shattered windshield and airframe damage. The NTSB determined that a contributing factor was the controller's failure to provide the flight crew with the PIREP.

In March 2012, during an approach to land in Anchorage, Alaska, a Learjet 35A encountered severe in-flight icing conditions (NTSB, 2012a). The conditions were so severe that the windshield became obscured with ice, and the airplane veered off the runway during landing and into a snowbank. Another aircraft 7 miles away, while on approach to another airport, encountered severe icing conditions. Although the tower reported the information to the Anchorage approach controller, the controller did not relay the information to the Learjet flight crew (NTSB, 2017a; NTSB, 2012a).

In June of 2012, in Aspen, Colorado, a pilot submitted a PIREP reporting low-level wind shear on final approach (NTSB, 2012b). Although ATC disseminated the PIREP to local traffic, including a Learjet 60 involved in an accident, air traffic controllers did not enter the PIREP in the NAS. The PIREP should have been given priority handling because the PIREP included a report of a 15-knot loss of airspeed on a short final approach that met the classification of an urgent PIREP (NTSB, 2017a; NTSB, 2012b).

To facilitate and save time, air traffic controllers often consolidate multiple PIREPs and enter one PIREP into the NAS (NTSB, 2017a). By consolidating PIREPs, important information is lost. For example, a controller might consolidate several urgent PIREPs reporting moderate-to-severe turbulence from numerous types of aircraft. For example, moderate turbulence for a Cessna 172 might not be considered the same for a Boeing 737. Additionally, by consolidating PIREPs, information such as the time and location of the PIREP might be lost. Meteorologists need this information to update forecasts and are also needed by pilots who are inflight and seeking updated weather.

2.2 Pilot Issues

During the NTSB's (2017a) June 2016 forum on pilot weather reports, an AOPA manager reported that a preliminary review of the comments of their PIREP survey indicates that pilots had to leave an ATC frequency to file a PIREP, finding the correct flight service station (FSS) frequency, and then communicating with FSS was too time-consuming. A review of the AOPA (2016b) survey confirms this impediment. Pilots further described reporting PIREPs with an FSS as inefficient because it takes too long for the FSS to read back the PIREP report for accuracy (AOPA, 2016b; NTSB, 2017a).

The NTSB (2017a) noted that pilots often cannot submit PIREPs because of cockpit workload. This was confirmed by the comments in the AOPA (2016b) report. In a survey by Casner (2010), 159 (58%) GA pilots reported they were interested in a method to submit PIREPs that was quicker and more convenient. In a second survey by Casner (2014), GA pilots said they would be more apt to report PIREPs if they had a cockpit interface that automatically reported parameters such as aircraft location, time altitude, aircraft type, wind, and temperature and also provided menus for selecting the other elements of a PIREP.

3. Methods

An FMEA was conducted to identify failure modes within the two identified PIREP processes, ATC process failures, and airmen-induced issues contributing to the failure to submit PIREPs (NTSB, 2017a; AOPA, 2016b). After the failure modes were identified, recommended actions to eliminate or reduce the

error or error effects were developed (Carstens, 2005; Carstens, 2006; Carstens et al., 2014). Failure modes were identified based on the literature and researchers' expertise to identify what can go wrong with a process to include errors, trust issues, and inaccuracies occurring in PIREP submittals. The identified failure modes were assessed by determining the cause of the failure and the effect, and the consequence of each failure. Usually, an FMEA includes three scales to calculate the risk priority number (RPN) value, including severity, occurrence, and detection. Due to trust being identified within the PIREP literature, the FMEA was modified to include a trust rating resulting in the FMEA having four scales. A high RPN symbolizes a process that needs to be changed to reduce the risks presented by the failure mode. Part of the FMEA process includes the identification of recommended actions from a proactive standpoint that, if implemented, could reduce the RPN value. Therefore, the RPN is recalculated based on the recommended action, referred to the post-RPN, which helps assess proposed changes to a process in advance of the recommended action being implemented.

The FMEA spreadsheet was placed into Excel. Each column of the FMEA assisted the research team in analyzing each failure mode through trust, severity, occurrence, and detection scales. Researchers moved through the FMEA spreadsheet by answering the different column questions and assessing the failure modes based on four scales. The FMEA spreadsheet consists of: (a) process, (b) potential failure mode, (c) trust description, (d) trust rating, (e) potential effects, (f) severity rating, (g) potential causes, (h) existing prevention control, (i) occurrence rating, (j) existing detection controls, (k) detection rating, (l) RPN, (m) RPN ranking, (n) recommended actions, (o) new trust rating, (p) new severity rating, (q) new occurrence rating, (r) new detection rating, (s) new RPN, (t) new RPN ranking, (u) effectiveness, and (v) effectiveness ranking.

The process column describes the processes affiliated with a failure mode. The potential failure mode is a description of the failure that can occur. The trust reasoning describes the identified trust issues, on which the trust rating value is assigned. The potential effect refers to the probable outcomes of the failure mode, and the severity of the failure mode is rated accordingly. The potential cause is the reason why the failure can occur. The prevention design, if available, is described based on the identified potential cause of the failure mode. According to the potential cause and prevention design, the likelihood of occurrence of the failure can be determined. The detection control describes the methods used to detect the failure if it happens, and the likelihood that the failure can be detected is assigned a rating. The RPN is the product of the four ratings (i.e., trust, severity, occurrence, detection). Based on the rating provided for each of the four scales, the RPN can range from 1 to 10,000, where the failure mode with the highest risk will have the largest number. The recommended actions describe the proposed actions that may be effective in reducing the risk of the failure mode. Then, a post recommended action RPN rating of trust, severity, occurrence, and detection is calculated based on how well the recommended action can eliminate or reduce the error and error effects affiliated with the failure mode. The post-RPN and the associated ranking uses a stoplight approach in identifying failure modes with a low risk of failure by coding these cells in green. Failure modes with a medium risk are coded in yellow, and those with a high risk are coded in red. If the recommended action is not implemented, then the first RPN should be used in assessing the risks presented by the failure mode within the current PIREP process. There is also a column dedicated to effectiveness. This is an internal tool for the researchers to assess how well the recommended action lowers the risk. Therefore, it is calculated based on how effective the recommended action was at reducing the RPN value. The effectiveness result is presented as a percentage, where a high value indicates high effectiveness. It is calculated by $[1 - (\text{post RPN}/\text{first RPN})]$ within each row. The ranking of the effectiveness is also provided. The RPNs, effectiveness, and rankings are each automatically calculated within the FMEA spreadsheet.

The four scales use 1, 3, 5, 7, and 10 as the available ratings with the descriptions provided in Table 2. For the trust scale, 1 means calibrated trust, and 3, 5, 7, and 10 are low, moderate, high, and extreme overtrust or undertrust. Calibrated trust symbolizes that the subjective trust in the process

agrees with the objective trust measures such as the perceived number of errors match the actual number of errors (Wickens, Hollands, Banbury & Parasuraman, 2016). Therefore, if a process for instance despite an overtrust or undertrust rating, either extreme, is a concern for safety then the process aligns with the purpose of conducting a FMEA. For the severity scale; 1, 3, 5, 7, and 10 refer to none, minor, moderate, major, and catastrophic consequences respectively. The occurrence ratings indicate rare, low, moderate, high, and very high occurrences of the failure depicted moderate, low, and none, respectively.

Table 2

FMEA Rating Description

Ratings	Trust	Severity	Occurrence	Detection
10	Pilots have extreme undertrust/overtrust	Catastrophic consequence	Very high occurrence	Cannot detect failure (0-24%)
7	Pilots have high undertrust/overtrust	Major consequence	High occurrence	Low detect rate (25-49%)
5	Pilots have moderate undertrust/overtrust	Moderate consequence	Moderate occurrence	Moderate detect rate (50-79%)
3	Pilots have low undertrust/overtrust	Minor consequence	Low occurrence	High detect rate (80-89%)
1	Pilots have calibrated trust	No (or positive) consequence	Rare occurrence	Very High detect rate (90-100%)

4. Results and Discussion

4.1 ATC submission PIREP processes failures

The first identified process was ATC failing to forward PIREPs to other facilities in their control area (NTSB, 2017a). This process was assigned ratings for trust, severity, occurrence, and detection of 3, 3, 5, and 10 that corresponded to an RPN of 450. With the recommended action to require air traffic controllers to enter all PIREPs into the NAS system, each rating was reduced to 1, which corresponded to an RPN of 1 with an effectiveness rating of 99.78%.

The second identified process was ATC consolidating PIREPs before entering them into the NAS (NTSB, 2017a). This process was deemed to have ratings for trust, severity, occurrence, and detection of 3, 3, 5, and 10 that corresponded to an RPN of 450. With the recommended action to require ATC to enter all PIREPs into the NAS system, the ratings were all reduced to 1 which corresponded to an RPN of 1 with an effectiveness rating of 99.78%.

The third and final identified process identified by NTSB (2017a), ATC failing to solicit PIREPs, the assigned ratings for trust, severity, occurrence, and detection were 7, 5, 5, and 10 that corresponded to an RPN of 1750. With the recommended action to require ATC to enter all received PIREPs into the NAS, the ratings were all reduced to 1, corresponding to an RPN of 1 with an effectiveness rating of 99.78%.

These recommendations were consistent with the NTSB’s recommendations made to the FAA. The NTSB suggested that the FAA develop a best practice guide for disseminating PIREPs for each type of ATC facility (NTSB, 2017a). These actions would guide controllers in the dissemination process. The NTSB also recommended that automation technology capture data elements from the air traffic controller’s displays, including aircraft type, time, location, and altitude, to automatically populate these data into the PIREP. The controller would then have to enter the remaining PIREP elements and disseminate the PIREPs. This recommendation would shorten the time it takes for the controller to submit the PIREP so they can

focus on traffic separation and safety alerts to aircraft.

4.2 Airmen submission PIREP processes failures

The first identified process was that airmen were not comfortable with what weather to report in a PIREP (NTSB, 2017a). The process was deemed to have trust, severity, occurrence, and detection ratings of 5, 5, 5, and 10 corresponding with an RPN 1250. By providing more extensive PIREP training, the ratings could be reduced to 1, 3, 3, and 10, corresponding to an RPN of 90 and an effectiveness rating of 92.8%.

The second identified process was that airmen had poor experience assessing weather, which discouraged them from filing PIREPs (NTSB, 2017a). This process was assigned trust, severity, occurrence, and detection ratings of 5, 5, 5, and 10, corresponding with an RPN of 1250. By providing more extensive PIREP training, the ratings could be reduced to 1, 3, 3, and 10 for an RPN of 90 and an effectiveness rating of 92.8%.

The third identified process was that airmen perceived it took too long for FSS to accept a PIREP and too long to read back a PIREP (AOPA, 2016b). This process was assigned trust, severity, occurrence, and detection ratings of 5, 5, 5, and 1, corresponding with an RPN of 125. NTSB (2017a) concluded the following:

some flight circumstances place demands on the pilot such that he or she must minimize the time spent filing a report. Pilots can reduce time spent filing a report by concisely providing thorough and accurate information, and specialists can help by prioritizing how critical information is obtained from the pilot. The NTSB concludes that the process by which FSS specialists receive and verify verbal PIREP information is having an unintended deterrent effect on reporting because some pilots find the process too time-consuming and, therefore, choose not to submit PIREPs. Therefore, the NTSB recommends that the FAA review the process by which federal and contract FSS specialists receive verbal PIREPs and then simplify procedures to reduce the amount of time the specialists take to obtain the necessary information from pilots. (p. 12)

By limiting the readback information required by pilots, pilots may be more apt to submit PIREPs thereby resulting in new ratings reduced to 1, 1, 1, 1 for an RPN of 1, and an effectiveness rating of 99.2%.

The fourth identified process was that airmen were fearful of an enforcement action for flying into poor weather without having the minimum weather requirements for IMC (NTSB, 2017a). This process was assigned trust, severity, occurrence, and detection ratings of 3, 3, 3, and 10, corresponding with an RPN of 270. By amending the rules to provide immunity from violations for negligently flying into poor weather, the ratings and RPN were unchanged. It was reasoned that unless full immunity was provided, airmen would be reluctant to submit PIREPs if they violated the Federal Aviation Regulations.

The fifth identified process was position and time reporting errors (NTSB, 2017a). This process was assigned trust, severity, occurrence, and detection ratings of 3, 3, 3, and 10, corresponding with an RPN of 270. By providing a pre-call to ATC, the controller can mark the time and location until the full PIREP could be reported; the ratings were reduced to 1, 1, 1, and 1 for an RPN of 1 and an effectiveness rating of 99.63%.

The sixth identified process was that pilots do not believe PIREPs are important (NTSB, 2017a). This process was assigned trust, severity, occurrence, and detection ratings of 3, 3, 1, and 3, corresponding with an RPN of 27. NTSB (2017a) concluded the following:

One reason that some pilots do not routinely file PIREPs may be that they are unaware of how important all PIREPs—including those of fair weather conditions or conditions consistent with the forecast—are for improving aviation weather products. One survey of 189 GA pilots found a likely

relationship between pilots' perceived importance of PIREPs and their willingness to submit them (Casner 2010). The survey results also indicated that the pilots surveyed were more likely to submit PIREPs for severe or unexpected weather phenomena than for moderate or as-forecasted conditions. The survey results strongly suggested that pilots believed that the primary purpose of PIREPs was for reporting bad weather. AOPA's review of the responses to its 2016 survey revealed a similar trend: 81% of the pilots who responded said they would rarely to never file a PIREP for as-forecast conditions, and 76% said that they would rarely to never report benign conditions (NTSB 2016b, 21). (p. 6)

By providing pilots training, the ratings were reduced to 3, 1, 1, and 1 for an RPN of 3 and an effectiveness rating of 88.89%.

The seventh identified process was pilots do not believe that air traffic controllers want to accept PIREPs (NTSB, 2017a). This process was assigned trust, severity, occurrence, and detection ratings of 3, 3, 3, and 3, corresponding with an RPN of 81. By having ATC solicit more PIREPs, GA pilots will realize ATC is amenable to accepting PIREPs. Thus, the ratings were reduced to 1, 1, 1, and 1 for an RPN of 1, and an effectiveness rating of 98.77%.

The above recommendations are consistent with the recommendations that the NTSB (2017a) made to AOPA and flight instructors. Those recommendations include for AOPA to update their online PIREP weather course to explain the value of PIREPs and how meteorologists use PIREPs to revise weather. Recommendations to flight instructors through the National Association of Flight Instructors and Society of Aviation Flight Educators include using real-world examples to file PIREPs and practice assessing real-time weather conditions.

5. Conclusion

Many aviation accidents and incidents are caused by pilots flying into hazardous weather (AOPA, 2015). Accidents caused by GA pilots who are not instrument rated and fly into IMC can be fatal. Unforecasted turbulence is primarily responsible for injuries to passengers and crew onboard air carriers (NTSB, 2017a). The best source of PIREPs comes from pilots and air traffic controllers. However, pilots are often not comfortable with the PIREP format or what weather to report in the PIREP. They have poor experience assessing weather, so they are discouraged from submitting a PIREP. Pilots believe the reporting process is cumbersome and time-consuming and that air traffic controllers do not want to accept PIREPs. Some pilots fear an enforcement action if they inadvertently fly into poor weather and therefore will not submit a PIREP. Flight deck workload often prevents pilots from submitting PIREPs. Pilots often do not want to leave the ATC frequency to contact FSS. Air traffic controllers also contribute to the lack of PIREP submission because of an imperfect dissemination process, the consolidation of PIREPs, and the lack of time to solicit PIREPs. By creating solutions identified through conducting a FMEA, potential failures can be identified, and by implementing the solutions these failures can be proactively and reactively mitigated.

The purpose of the study was to increase GA PIREPs by reducing errors, trust issues, and inaccuracies in PIREP submittals that prevent inclusion in the NAS. The methods for this study included a literature review and conducting an FMEA to identify failure modes with the PIREP process that includes recommended actions to eliminate or reduce the error or error effects. The FMEA research resulted in recommended actions that were consistent with NTSB recommendations to increase PIREPs in eliminating or reducing the error or error effects with the failure modes identified. Thereby, the RPN was reduced in the majority of processes and affiliated subprocesses for ATC process failures and airmen-induced issues contributing to a failure to submit PIREPs.

Several future research areas have been identified. First, research should be conducted to

determine if the recommendations made by the NTSB (2017a) to AOPA have been implemented. The recommendations include whether AOPA updated their online PIREP course content to include scenario-based training that illustrates the value of PIREPs, explain how meteorologists use PIREPs, provide guidance on how to assess the weather, and demonstrate various ways to submit a PIREP. Second, research should be conducted to determine if the recommendations made by the NTSB to flight instructors have been implemented. Recommendations to flight instructors include providing students with real-world examples of PIREPs, explanations on how PIREPs are used and benefit the NAS, examples of various ways to submit PIREPs, and assessing the weather. Third, research should be conducted to further determine if other recommendations made by the NTSB have been implemented. Those recommendations include whether procedures have been implemented to reduce the amount of time to submit a PIREP. Whether automated pilot weather report data-collection tools have been implemented and whether best practice procedures have been implemented regarding the dissemination of PIREPs. A fourth research area identified would be to assess the impact of pilot trust of speech recognition capabilities in general software on perceived reliability and validity of speech recognition and “hands minimized” PIREP submittal tools. A fifth future research area would be developing and testing a speech recognition prototype PIREP submission tool for pilots that mitigates PIREP submission errors. This research would involve calculating word error rates and word information loss with regard to the tool. Overall, the researchers identified recommended actions and will continue to explore how to further increase GA PIREPs through reducing errors, trust issues, and inaccuracies in PIREP submittals that prevent inclusion in the NAS.

7. References

- AOPA Air Safety Institute (2015). 24th Joseph T. Nall Report: General Aviation Accidents in 2012.
<https://www.aopa.org/-/media/Files/AOPA/Home/Pilot-Resources/Safety-and-Proficiency/Accident-Analysis/Nall-Report/15-FN-0022-1-24th-Nall-V6.pdf>
- AOPA Air Safety Institute (2016a). 25th Joseph T. Nall Report: General Aviation Accidents in 2013.
<https://www.aopa.org/-/media/Files/AOPA/Home/Training-and-Safety/Nall-Report/25thNallReport.pdf>
- AOPA (2016b). Pilot Report Survey.
http://download.aopa.org/advocacy/0417_2016_pilot_report_survey_final_report.pdf
- AOPA Air Safety Institute (2017). 26th Joseph T. Nall Report: General Aviation Accidents in 2014.
<https://www.aopa.org/-/media/Files/AOPA/Home/Training-and-Safety/Nall-Report/26thNallReport2017.PDF>
- AOPA Air Safety Institute (2018). 27th Joseph T. Nall Report: General Aviation Accidents in 2015.
<https://www.aopa.org/-/media/Files/AOPA/Home/Training-and-Safety/Nall-Report/27thNallReport2018.pdf>
- AOPA Air Safety Institute (2021a). 28th Joseph T. Nall Report: General Aviation Accidents in 2016.
<https://www.aopa.org/training-and-safety/air-safety-institute/accident-analysis/joseph-t-nall-report/28th-nall-report>
- AOPA Air Safety Institute (2021b). 29th Joseph T. Nall Report: General Aviation Accidents in 2017.
<https://www.aopa.org/training-and-safety/air-safety-institute/accident-analysis/joseph-t-nall-report/nall-report-figure-view?category=all&year=2017&condition=all&report=true>

- AOPA Air Safety Institute (2021c). 30th Joseph T. Nall Report: General Aviation Accidents in 2018. <https://www.aopa.org/training-and-safety/air-safety-institute/accident-analysis/joseph-t-nall-report/nall-report-figure-view?category=all&year=2018&condition=all&report=true>
- Carstens, D. S. (2005, July) Human Factors in Medicine: Task and Error Analysis, Proceedings of the Human-Computer Interaction International 2005, Las Vegas, NV. CD-ROM [invited paper].
- Carstens, D. S. (2006) Error analysis leading technology development, Theoretical Issues in Ergonomics Science, 7(6), 525-549.
- Carstens, D.S., Rodriguez, W. & Wood, M.B. (2014). Task and Error Analysis Balancing Benefits over Business of Electronic Medical Records, International Journal of Electronic Healthcare, 7(4), 331-348.
- Casner, S., M. (2014). Increasing participation in the pilot weather reporting (PIREP) system through user interface design. Journal of Aviation Technology and Engineering, 4:1, 11-19. <https://docs.lib.purdue.edu/jate/vol4/iss1/2/>
- Casner, S. M. (2010). Why don't pilots submit more pilot weather reports (PIREPs)? The International Journal of Aviation Psychology, 20(4), 347-374. <https://www.tandfonline.com/doi/abs/10.1080/10508414.2010.487015>
- Federal Aviation Administration (FAA). (2019a). FAA Order Number JO 7110.65Y, Change 3, Paragraph 2-6-2. <https://www.faa.gov/documentlibrary/media/order/atc.pdf>
- Federal Aviation Administration (FAA). (2019b). FAA Order Number JO 7110.65Y, Change 3, Paragraph 2-1-1. <https://www.faa.gov/documentlibrary/media/order/atc.pdf>
- National Transportation Safety Board (NTSB). (2015). Accident Investigation, CEN15LA137. <https://data.nts.gov/Docket?ProjectID=90694>
- National Transportation Safety Board (NTSB). (2017a). Special Investigation Report: Improving Pilot Weather Report and Dissemination to Benefit Safety in the National Airspace System. <https://www.nts.gov/safety/safety-studies/Documents/SIR1702.pdf>
- National Transportation Safety Board (NTSB). (2017b). Accident Investigation, OPS15IA020. [https://data.nts.gov/Docket?ProjectID=91756accident report](https://data.nts.gov/Docket?ProjectID=91756accident%20report)
- National Transportation Safety Board (NTSB). (2016b). PIREPs: Pay It Forward ... Because Weather for One is Weather for None. Transcript of Forum, Day 2 (June 22). Washington, DC: NTSB.
- National Transportation Safety Board (NTSB). (2012a). Accident Investigation, ANC12IA024. <https://data.nts.gov/Docket?ProjectID=83048>
- National Transportation Safety Board (NTSB). (2012b). Accident Investigation, CEN12LA345. <https://data.nts.gov/Docket?ProjectID=83888>
- Wickens, C., Hollands, J., Parasuraman, R. & Banbury, S. (2016). Engineering Psychology & Human Performance. 4th Edition, New York, NY: John Wiley & Sons. (ISBN-13: 978-0205021987 or ISBN-10: 0205021980).

Robust Cardiovascular Disease Prediction Using Logistic Regression

Snigdha Datta¹

Isaac K. Gang²

¹*Colorado State University Global*

²*George Mason University*

snigdha.datta@csuglobal.edu; iqang@gmu.edu;

Abstract

Cardiovascular disease, commonly known as heart disease, is one of the leading causes of death in the United States and worldwide. Early detection of the disease can save thousands of lives and billions of dollars in healthcare costs. A statistical model with the ability to accurately predict heart disease could be of immense help to the patients, their families, the medical community, and the healthcare system. Hospitals and providers collect many patient health metrics during screening and routine lab tests, which could be used to build such a statistical model. A robust heart disease prediction model is built using a sample dataset from the University of California, Irvine Machine Learning repository. Initial Hypotheses are formulated, and the most significant predictor variables are identified using the Wald test. The statistical significance of the proposed model is tested using the Likelihood-Ratio test. A repeated 10-fold cross-validation technique is used to evaluate the model's prediction power on previously unseen data. Keeping in mind the simplicity, usability, and explainability of results to the medical community, a Logistic Regression model that predicts the heart disease class with a high degree of accuracy is presented in this paper.

1. Introduction

According to the CDC, "heart disease" refers to several heart conditions, the most common of which is coronary artery disease, a significant cause of heart attacks. Other related conditions include the inability of the heart to pump blood efficiently due to malfunction of the heart valves resulting in heart failure. While some people are born with heart disease, it is essential to point out that anyone, including children, can develop heart disease. The condition occurs when plaque builds up in the arteries causing them to narrow over time, reducing blood flow to the heart. Habits such as smoking, eating an unhealthy diet, and not getting enough exercise increase the risk of having heart disease. The recent CDC data suggests that heart disease causes 1 in 4 deaths in the United States, equating to about 655,000 lives and \$219 billion each year, including healthcare services, medicine, and loss in productivity. Essentially, a person dies every 36 seconds due to heart disease in the United States (CDC, 2020).

The symptoms of underlying heart disease are not always visible. It may go undetected for a prolonged period causing irreversible damage to the human body. One of the biggest challenges while dealing with heart diseases is the associated medical costs exacerbated by late detection. Since early detection reduces medical costs and saves lives, efforts to promote it are critical and one of the key goals for this paper. Many important metrics collected during screening and routine lab tests can be

extracted and utilized to find hidden patterns in the data and lead to early detection of the disease. The emergence of Machine Learning (ML) and Artificial Intelligence (AI) coupled with the availability of quality datasets has presented us with an opportunity to make early detection possible algorithmically. To that end, we surveyed similar efforts in Section 2. We discuss our methodology in detail in Section 3 and present our results with relevant analysis in Section 4. We conclude our work with future direction in Section 5 while providing references in Section 6.

2. Background

Modeling and algorithmically predicting cardiovascular disease may prove very helpful in the early detection of heart disease, saving thousands of lives and billions of dollars in healthcare costs. Leveraging statistical and ML methods have been deemed valuable in this regard by many authors before us, and they applied several such techniques to address the problem.

Fredrick David & Belcy (2018) used data mining techniques to build multiple classification models for predicting the heart disease class. The UCI heart disease dataset was used to compare the results of three algorithms- Random Forest (RF), Decision Tree (DT), and Naïve Bayes (NB). At 81% accuracy, the RF result was the best among the three.

Mythili et al. (2013) proposed using the Cleveland Heart Disease dataset and comparing four algorithms- Support Vector Machine (SVM), Logistic Regression, DTs, and Rule-based approach on their ability to predict the heart disease class. Khanna et al. (2015) also utilized the Cleveland Heart Disease dataset to compare SVM, Logistic Regression, and Artificial Neural Network (ANN) models. According to their results, the Logistic Regression with 10-fold cross-validation and the linear SVM with 5-fold cross-validation achieved the best testing accuracy of 88.2% and 87.6% respectively. Hassani et al. (2020) applied Hybrid ANN and DT using the UCI datasets and achieved 87.3 precision. Khemphila & Boonjing (2010) used a dataset containing 303 observations to compare the prediction power of three algorithms - Logistic Regression, DTs, and ANNs. Compared to the 81.2 % accuracy rate of Logistic Regression, ANNs achieved a slightly lower accuracy rate of 80.2% but had the least error rate among the three algorithms.

Mufudza & Erol (2016) used model-based clustering techniques to predict the heart disease class. The results from two different Poisson Regression models- standard and concomitant variable mixture regression models, were compared. The prediction results of a two-component concomitant variable Poisson mixture regression model were better than the standard model. However, the zero-inflated Poisson mixture regression model was the best performing model with the lowest Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) values. Khateeb & Usman (2017) achieved 80% accuracy with their prediction model using K-Nearest Neighbor and 14 data attributes.

Singh et al. (2018) used the WEKA Data Mining tool to implement an ANN-based heart disease prediction system. A Multi-layer Perceptron NN model with backpropagation was trained and tested using a 60-40 train-test split heart disease dataset containing 303 records. The experimental results show that the system could predict heart disease with almost 100% accuracy. Jan et al. (2018) also used the WEKA Data Mining tool to create an intelligent heart disease prediction system. An ensemble model combining SVM, ANN, NB, Logistic Regression, and RF classifiers was created to predict cardiovascular disease recurrence. The UCI Cleveland and Hungarian datasets were used in the analysis. The RF

ensembled model achieved the highest accuracy of 98.17%.

Ibrahim et al. (2019) leveraged adversarial learning to design a “fair” model to equitably distribute therapies across race and gender groups. Fernandez-Lozano et al. (2018) and Sajeev & Maeder (2019) used a Generalized Linear Model to predict complications in peritoneal dialysis patients. They surveyed PubMed for relevant prediction models finding 229 articles in total.

Usman (2018) highlight the inherent problem with feature selection while building a heart disease prediction model. Two slightly different cuckoo-inspired algorithms – the cuckoo search algorithm (CSA) and the cuckoo optimization algorithm (COA), were used for feature selection on multiple datasets. The reduced features were used to build, train, and test four classification algorithms: NB, RF, Multi-layer Perceptron, and SVM. The experimental results show that CSA performed better than COA for minimal feature selection. SVM achieved the highest classification accuracy rate on the Eric, Hungarian, and Stat log datasets.

3. Methodology

Due to its simplicity and interpretability of results, our ML algorithm of choice is Logistic Regression (LR). LR isn’t regression at all. In fact, it is a parametric classification technique, and thus it perfectly suits this use case. Instead of a standard linear function of a straight line, it models using a sigmoid function to shrink the domain of possible predicted values between 0 and 1 compared to a standard linear function that allows values from $-\infty$ to $+\infty$ (James et al., 2013). For multiple predictors, an LR can be mathematically represented as:

$$p(X) = \frac{e^{\beta_0 + \beta_1 x_1 + \dots + \beta_n x_n}}{1 + e^{\beta_0 + \beta_1 x_1 + \dots + \beta_n x_n}}; \text{ where } X = (x_1, x_2, \dots, x_n) \text{ for } n \text{ predictors.} \quad (1)$$

It can also be represented as a *log-odds* or *logit* function:

$$\log\left(\frac{p(X)}{1 - p(X)}\right) = \beta_0 + \beta_1 x_1 + \dots + \beta_n x_n; \beta_0 \dots \beta_n \text{ are the parameters of the function.} \quad (2)$$

The parameters of the equation are estimated using the Maximum Likelihood function. The goal is to estimate the values for parameters $\widehat{\beta}_0, \widehat{\beta}_1 \dots \widehat{\beta}_n$ such that the $p(X)$ is close to 1 for the positive class and close to 0 for the negative class (James et al., 2013). Mathematically, the Maximum Likelihood function can be represented as:

$$\ell(\beta_0, \beta_1, \dots, \beta_n) = \prod_{i: y_i=1} p(x_i) \prod_{i': y_{i'}=0} (1 - p(x_{i'})) \quad (3)$$

Unlike complex ANN-based connectionist models, the results of an LR model are easier to explain. It also takes a lot less data to train such models. Since the purpose of building such a model is to aid the medical community in the early detection of heart disease, the explainability of results is particularly crucial. To trust the model results, the medical community expects to clearly understand why the model predicted what it predicted.

Our work is divided into five distinct phases and follows a standard Data Analytics Lifecycle (EMC Education Services, 2015). The high-level process flow and critical subtasks for each phase are shown in Figure 1. The Identification of data source, programming resources, preliminary data analysis, framing the correct problem statement, formulating a set of Initial Hypotheses (IHs), and defining the success and failure criteria for the project are performed in the Data Discovery phase. The data cleansing and conditioning, summary, and visual statistical analysis are performed in the Data Preparation phase to get additional insight into the nature of the data. A more detailed analysis of the potential predictor variables is performed in the Model Planning phase using various descriptive statistical analysis techniques. Moreover, the hypotheses testing and model selection are also performed in this phase. In the Model Building phase, the initial dataset is split into a training and a testing dataset. The training dataset is used to train the selected statistical model, and the testing dataset is used to test its ability to accurately predict the outcome class. Finally, all the findings and inferences from the experimental results are communicated.

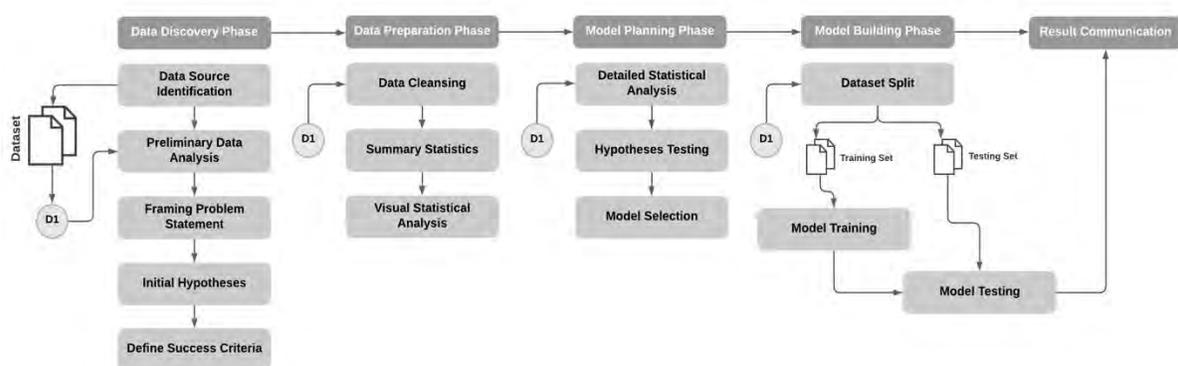


Figure 1: High-Level Analytics Process Flow

4. Analysis and Results

4.1. Data Discovery

The Cleveland Heart Disease dataset (Detrano, 1988), from the University of California Irvine Machine Learning Repository's publicly available large datasets, is used for this experiment. The dataset contains sample data collected from VA Medical Center, Long Beach, and Cleveland Clinic Foundation. Though the original dataset contains 76 data elements, the processed dataset which is being used contains only 14 selected and widely used data elements. The variable names, variable data types, and descriptions have been consolidated in a tabular format below.

It is essential to perform a preliminary data analysis to become familiar with the data content, quality, limitations and understand any interdependencies among the data elements. The initial analysis results show that the dataset contains 303 rows and 14 columns. 8 out of the 13 predictor variables are categorical, and 5 are continuous. The outcome variable is categorical.

Table 1: Data element description

Variable Name	Variable Type	Description
Age	Integer	Age in years.
Gender	Integer	Gender (Values: 1 = male; 0 = female).
CP	Integer	Chest pain type (Values: 1 = typical angina; 2 = atypical angina; 3 = non-anginal pain; 4 = asymptomatic).
Trestbps	Integer	Resting blood pressure (Measured in mm Hg on admission to the hospital).
Chol	Integer	Serum cholesterol measured in mg/dl.
FBS	Integer	Fasting blood sugar > 120 mg/dl (Values: 1 = true; 0 = false).
RestECG	Integer	Resting electrocardiographic results (Values: 0 = normal; 1 = having ST-T; 2 = hypertrophy)
Thalach	Integer	Maximum heart rate achieved.
Exang	Integer	Exercise induced angina (Values: 1 = yes; 0 = no).
Oldpeak	Floating Point Number	ST depression induced by exercise relative to rest.
Slope	Integer	The slope of the peak exercise ST segment (Values: 1 = up-sloping; 2 = flat; 3 = down-sloping).
CA	Factor	The number of major vessels (Values: 0-3) colored by fluoroscopy.
Thal	Factor	Thallium stress test result (Values: 3 = normal; 6 = fixed defect; 7 = reversible defect).
Num	Integer	The predicted attribute - diagnosis of heart disease (angiographic disease status) (Values: 0 = < 50% diameter narrowing; 1 = > 50% diameter narrowing).

4.2. Problem Statement and Initial Hypotheses

Formulating a set of IHs is crucial during the Data Discovery phase of the Data Analytics Lifecycle. Learning about the data sources, their domain, and framing the right problem the analytical system is trying to solve are essential before the hypotheses can be formulated. Understanding the data domain provides the proper context to comprehend the data's characteristics and a meaningful way to interpret it. The proposed analytical system is attempting to predict the heart disease class using one or more predictor variables. The results of this experiment must be able to answer the following questions:

1. Is there a statistically significant relationship between one or more predictor variables and the outcome class?
2. Can an acceptably accurate heart disease prediction model be designed using the selected predictor variables?

Based on the preliminary data analysis, it is likely that more than one predictor is significantly contributing to the increased risk of heart diseases. Under the assumption of independence, the below IHs are formulated:

- Null Hypothesis(H_0): There is no statistically significant relationship between the individual predictor variables – *Gender*, *CP*, *CA*, and *Thal* and the *heart disease* outcome class.
- Alternate Hypothesis (H_a): There is a statistically significant relationship between the individual predictor variables – *Gender*, *CP*, *CA*, and *Thal* and the *heart disease* outcome class.

For the purposes of this experiment, an acceptable model is required to achieve at least 80% accuracy in correctly predicting the positive class. The success criteria of this experiment are accordingly set.

4.3. Data Preparation

A detailed data exploration, data cleansing, conditioning, and visual statistical analysis on the dataset is performed. The statistical analysis results help identify the right set of predictor variables for building an accurate statistical model. A working data frame is created using the original dataset. A new categorical variable- *HeartDisease* is created from the original outcome variable – *Num*. The levels: 0,1,2,3, and 4 of the original variable are re-coded to 0 and 1. Meaningful labels are added to the re-coded values- *Healthy* for 0 and *Heart Disease* for 1. As the original outcome variable – *Num* is no longer required, it is dropped from the data frame. The variables *Gender*, *CP*, *FBS*, *RestECG*, *Exang*, and *Slope* with discrete numerical values are re-coded as categorical and meaningful labels are added. The variables *CA* and *Thal* are already categorical. However, they have missing values in specific categories. The missing values are removed from the data frame. The dimensions of the cleansed dataset are validated. It is observed that 6 rows having NULL values are successfully removed from the data frame. The cleansed dataset has 14 variables, including 1 outcome variable, *HeartDisease*. 6 variables, which were initially quantitative and discrete, are successfully re-coded as categorical variables. The data frame now contains 9 categorical variables, including the outcome variable, *HeartDisease*. As expected, 5 variables remain quantitative and continuous.

A detailed visual statistical analysis on all the dataset variables is performed. *Age*, *Trestbps*, *Chol*, *Thalach*, and *Oldpeak* are continuous numerical variables; Boxplot is used for analyzing the critical features of the distribution, such as overall centrality, spread, and skewness. Bar plot, a common visualization technique for qualitative data, is used to visualize the characteristics of *HeartDisease* and the rest of the categorical predictor variables.

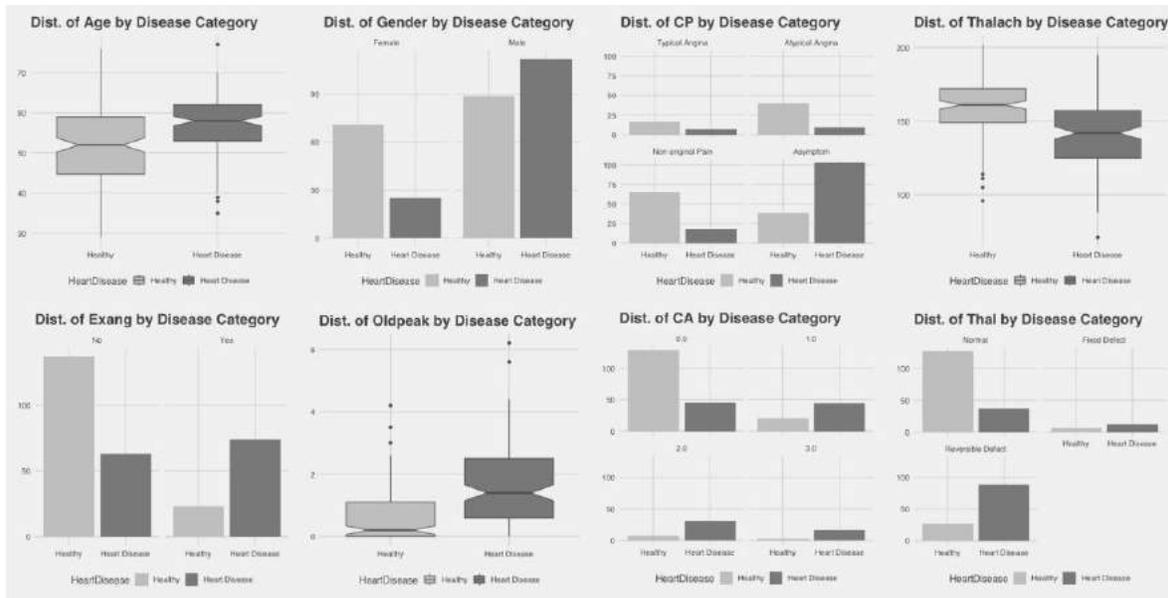


Figure 2: Visualization results for Heart Disease predictors

The dataset contains more data in the healthy category than in the heart disease category, indicating a slight imbalance. Outliers are detected for the variables- *Age*, *Tresbps*, *Chol*, *Thalach*, and *Oldpeak*. The following are the inferences drawn based on data exploration and visual statistical analysis results:

1. The predictor variables, *Tresbps*, *Chol*, *FBS*, and *RestECG*, do not seem to have any significant relationship with the outcome variable- *HeartDisease*. So, these variables are not contributing to the increased risk of heart diseases and can be excluded from further analysis.
2. Further analysis needs to be performed on the remaining predictor variables- *Age*, *Gender*, *CP*, *Thalach*, *Exang*, *Oldpeak*, *Slope*, *CA*, and *Thal*.

4.4. Model Planning

A more detailed statistical analysis on all the potential predictor variables is performed during the Model Planning phase. Understanding the data properties, relationships with other variables, and their significance in influencing the outcome variable class will determine the final list of significant predictors. The Wald test is used to test the statistical significance of individual predictors. The goal is to be able to reject the Null Hypothesis and thus accept the Alternate Hypothesis. The Likelihood-Ratio (L-R) test is used to test the significance of the Logistic Regression model. The most significant predictor variables and an appropriate statistical model will be used in the next phase for model building.

The full model parameter estimates are generated using the Generalized Linear Model (GLM) in R. Since the outcome variable is categorical and has only two levels, the binomial family function, *logit* is specified. Important statistics, such as the p-values of each predictor's parameter estimates, are recorded and analyzed. Based on the p-values, it can be inferred that the variables, *Age*, *Gender*, *CP*, *Tresbps*, *Thalach*, *Slope*, *CA*, and *Thal* are probable significant variables. The relative importance of individual predictor variables is assessed using the *varimp* function of R.

The Wald test for predictor variables is performed to evaluate the statistical significance of each predictor variable in its ability to influence the outcome variable of the model. Based on the results from variables of importance and the Wald test, only the most significant variables whose p values are ≤ 0.05 are being considered as predictor variables. These results will also be used during hypothesis testing.

The correct statistical testing techniques and modeling methods are determined by the characteristics of the predictor and outcome variables. There are 4 predictor variables, and they are all categorical. The outcome variable, *HeartDisease*, is also categorical. There are two distinct groups of data in the outcome variable – *Healthy* and *Heart Disease*. So, the outcome variable is also binomial. LR is a suitable modeling method as both the predictor variables and the outcome variable are categorical.

4.5. Hypotheses Testing

The Wald test results for individual predictors are used to select the most significant predictor variables for the LR model. It is observed that regression coefficients of only the variables- *Gender*, *CP*, *CA*, and *Thal* have p-value ≤ 0.05 . All other predictor variables can be removed from the model.

```

Wald test for Gender
  in glm(formula = HeartDisease ~ ., family = binomial(link = "logit"),
    data = df_clean)
F = 9.138371 on 1 and 276 df: p= 0.0027383
Wald test for CP
  in glm(formula = HeartDisease ~ ., family = binomial(link = "logit"),
    data = df_clean)
F = 6.186999 on 3 and 276 df: p= 0.00044072
Wald test for CA
  in glm(formula = HeartDisease ~ ., family = binomial(link = "logit"),
    data = df_clean)
F = 9.805762 on 3 and 276 df: p= 3.627e-06
Wald test for Thal
  in glm(formula = HeartDisease ~ ., family = binomial(link = "logit"),
    data = df_clean)
F = 6.094114 on 2 and 276 df: p= 0.0025712
    
```

Figure 3: Wald test results for Gender, CP, CA, and Thal.

L-R test is a suitable hypothesis test for LR models. It compares the likelihood of fit of the proposed model compared to a full model. In R, either the *lrtest* or the Analysis of Variance (ANOVA) function can be used for comparing the two models. The full model represents the Null Hypothesis, H_0 , and the proposed model represents the Alternate Hypothesis, H_a . It is necessary to test if the observed difference in model fit is statistically significant. Since the p-value corresponding to the likelihood ratio chi-square statistic of the proposed model is ≤ 0.05 , it can be safely assumed that the proposed model is statistically significant and fits the data better than a full model. The result is used to reject the Null Hypothesis (H_0) and accept the Alternate Hypothesis (H_a).

Likelihood ratio test

```

Model 1: HeartDisease ~ Age + Gender + CP + Trestbps + Chol + FBS + RestECG +
  Thalach + Exang + Oldpeak + Slope + CA + Thal
Model 2: HeartDisease ~ Gender + CP + CA + Thal
#Df  LogLik  Df  Chisq Pr(>Chisq)
1   21  -91.551
2   10 -113.369 -11  43.636  8.412e-06 ***
    
```

Figure 4: LR test comparing the full model reduced model.

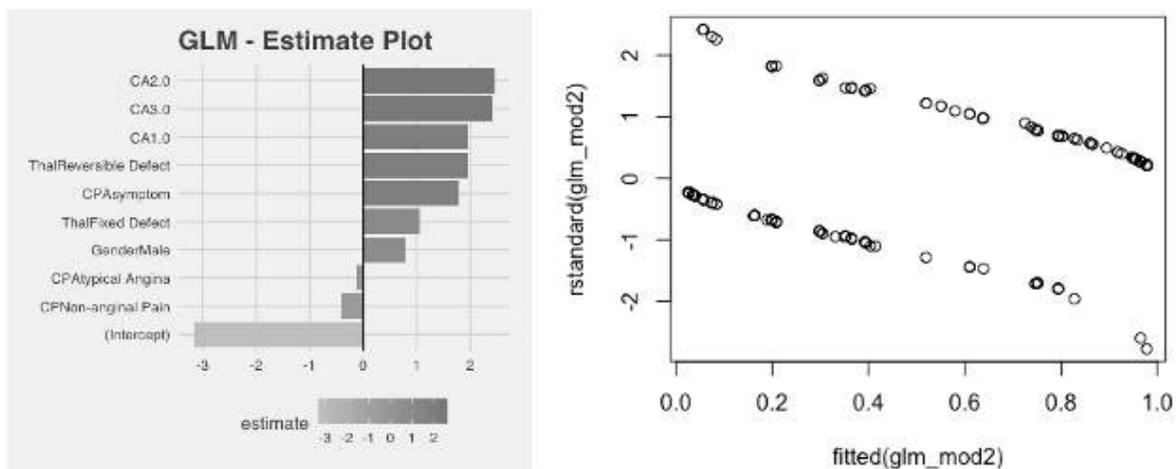


Figure 5: GLM estimate plot and standard vs. fitted model plot

4.6. Model Building, Training, and Testing

The working dataset is first split into a training and a testing dataset. The training dataset has 75% of the data (223 rows), and the testing dataset has the rest 25% of the data (74 rows). The training dataset is used to train the proposed LR model. A repeated K-fold cross-validation technique is used to customize the training control parameters, and the training data is partitioned into K equal-sized partitions called “folds.” The K-1 folds are used to train the model, while one-fold is used for evaluating the accuracy of the training. For this experimental setup, a 10-fold cross-validation technique is used. The Receiver Operating Characteristic (ROC), Sensitivity, and Specificity values of the trained model are recorded. The testing dataset is used to evaluate the ability of the model to correctly predict the outcome when new, never-seen data is presented to the model. The statistical results such as Accuracy, Kappa, p-value, and the confusion matrix results are recorded as shown in Figure 6.

4.7. Results

Based on the Wald test results for individual predictors and the L-R test for the LR model, the Alternate Hypothesis (H_a) was accepted by successfully rejecting the Null Hypothesis. A Multiple Logistic Regression (MLR) model with 4 predictor variables, *Gender*, *CP*, *CA*, and *Thal*, predicted the *HeartDisease* category with 81% accuracy.

The complete R code for the experiment has been uploaded to GitHub and publicly shared. The code can be accessed via the URL- <https://github.com/dattasd/analytics-ai-ml.git>

```
Confusion Matrix and Statistics

hd_pred      Healthy Heart Disease
Healthy      33              7
Heart Disease 7              27

      Accuracy : 0.8108
      95% CI   : (0.703, 0.8925)
      No Information Rate : 0.5405
      P-Value [Acc > NIR] : 1.077e-06

      Kappa : 0.6191

      Mcnemar's Test P-Value : 1

      Sensitivity : 0.7941
      Specificity : 0.8250
      Pos Pred Value : 0.7941
      Neg Pred Value : 0.8250
      Prevalence : 0.4595
      Detection Rate : 0.3649
      Detection Prevalence : 0.4595
      Balanced Accuracy : 0.8096

      'Positive' Class : Heart Disease
```

Figure 6: Confusion Matrix and other statistical results of the final model.

5. Conclusion

In this paper, we presented a Logistic Regression model that successfully predicts cardiovascular disease with 81% accuracy on test data. It is comparable to and, in most cases, outperforms models that

use similar setup and data splits. Moreover, compared to the black-box nature of complex ANN-based models, the results of LR are far more explainable. It also takes a lot less data to train such models. The purpose of this model is to help the medical community make accurate and early heart disease predictions. Though model accuracy is critical, the simplicity, usability, and explainability aspects should not be neglected. Only then can AI and ML-based solutions such as this be trusted, accepted, and widely used. In the future, we plan to investigate hybrid techniques and models with the hope of achieving greater than 90% accuracy on the same dataset and using similar data splits.

6. References

- CDC. (2020, September 8). *Heart disease facts*. Centers for Disease Control and Prevention. <https://www.cdc.gov/heartdisease/facts.htm>
- Davies, T. M. (2016). *The book of R: A first course in programming and statistics*. No Starch Press.
- Detrano, R. (1988). *UCI machine learning repository: Heart disease data set*. <https://archive.ics.uci.edu/ml/datasets/Heart+Disease>
- EMC Education Services. (2015). *Data science and big data analytics: Discovering, analyzing, visualizing, and presenting data*. John Wiley & Sons.
- Fernandez-Lozano, C., Valente, R. A., Díaz, M. F., & Pazos, A. (2018). A generalized linear model for cardiovascular complications prediction in PD patients. *Proceedings of the First International Conference on Data Science, E-learning and Information Systems*. <https://doi.org/10.1145/3279996.3280039>
- Fredrick David, H. B., & Belcy, S. A. (2018). Heart disease prediction using data mining techniques. *ICTACT Journal on Soft Computing*,9(01), 1817-1823. <https://doi.org/10.21917/ijsc.2018.0253>
- Hassani, M. A., Tao, R., Kamyab, M., & Mohammadi, M. H. (2020). An approach of predicting heart disease using a hybrid neural network and decision tree. *Proceedings of the 2020 5th International Conference on Big Data and Computing*. <https://doi.org/10.1145/3404687.3404704>
- Ibrahim, M., Louie, M., Modarres, C., & Paisley, J. (2019). Global explanations of neural networks. *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*. <https://doi.org/10.1145/3306618.3314230>
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning: With applications in R*. Springer Science & Business Media.
- Jan, M., Awan, A. A., Khalid, M. S., & Nisar, S. (2018). Ensemble approach for developing a smart heart disease prediction system using classification algorithms. *Research Reports in Clinical Cardiology*, 9, 33-45. <https://doi.org/10.2147/rrcc.s172035>
- Khanna, D., Sahu, R., Baths, V., & Deshpande, B. (2015). Comparative study of classification techniques (SVM, logistic regression, and neural networks) to predict the prevalence of heart disease. *International Journal of Machine Learning and Computing*, 5(5), 414-419. <https://doi.org/10.7763/ijmlc.2015.v5.544>
- Khateeb, N., & Usman, M. (2017). Efficient heart disease prediction system using k-nearest neighbor classification technique. *Proceedings of the International Conference on Big Data and Internet of Thing - BDIOT2017*. <https://doi.org/10.1145/3175684.3175703>

- Khemphila, A., & Boonjing, V. (2010). Comparing performances of logistic regression, decision trees, and neural networks for classifying heart disease patients. *2010 International Conference on Computer Information Systems and Industrial Management Applications (CISIM)*.
<https://doi.org/10.1109/cisim.2010.5643666>
- Mufudza, C., & Erol, H. (2016). Poisson mixture regression models for heart disease prediction. *Computational and Mathematical Methods in Medicine*, 2016, 1-10.
<https://doi.org/10.1155/2016/4083089>
- Mythili, T., Mukherji, D., Padalia, N., & Naidu, A. (2013). A heart disease prediction model using SVM- decision trees-logistic regression (SDL). *International Journal of Computer Applications*, 68(16), 11-15.
<https://doi.org/10.5120/11662-7250>
- Sajeev, S., & Maeder, A. (2019). Cardiovascular risk prediction models. *Proceedings of the Australasian Computer Science Week Multiconference*. <https://doi.org/10.1145/3290688.3290725>
- Singh, P., Singh, S., & Pandi-Jain, G. S. (2018). Effective heart disease prediction system using data mining techniques. *International Journal of Nanomedicine*, 13, 121-124. <https://doi.org/10.2147/ijn.s124998>
- Usman, A. M., Yusof, U. K., & Naim, S. (2018). Cuckoo inspired algorithms for feature selection in heart disease prediction. *International Journal of Advances in Intelligent Informatics*, 4(2), 95.
<https://doi.org/10.26555/ijain.v4i2.245>

Comparing the Frequency of Reported Wildlife Strikes by Region in the United States

Gabriella Ferra¹

Hamzah Alghamdi¹

Brooke E. Wheeler¹

Tianhua Li¹

¹*Florida Institute of Technology, College of Aeronautics*

bwheeler@fit.edu; gferra2018@my.fit.edu; halqhamdi2017@my.fit.edu;

tli2017@my.fit.edu

Abstract

Wildlife strikes are a threat to the safety of flight and a financial burden to aircraft operators. Although wildlife strikes are an unavoidable hazard in aviation, they can be mitigated through various methods, including airport wildlife management plans, deterring animal inhabitation around airports, and a national reporting system to bring awareness to the issue. This ex-post-facto research determined the difference in frequency of wildlife strikes at non-military class B airports in the contiguous United States by geographical region. Regions were defined as follows: (1) the Pacific coast, (2) the western mountainous areas and Texas, (3) the Midwest, (4) the Southeast, and (5) the Northeast. A total of 27,036 reports ranging from 1/1/2015 to 12/31/2019 were collected from the publicly available Federal Aviation Administration (FAA) database. The data analysis of regions showed that Region 2 had the greatest frequency of wildlife strikes, while Region 1 had the least frequency of strikes. Moreover, a between-subjects ANOVA suggested that there was a significant difference in the frequency of strikes between regions in the contiguous US. The airports in Region 1 underwent significantly fewer wildlife strikes than those in Regions 2, 3, and 4. Region 2 also had a significantly greater frequency of strikes than Region 5.

1. Introduction

Wildlife strikes are an unavoidable issue between nature and machines. Aviation history has witnessed remarkable situations where wildlife strikes threatened the safety of flight operations. The iconic story of US Airways Flight 1549 was one of those cases after a series of strikes caused dual engine failures during a critical phase of flight. A more thorough understanding of the behavior of wildlife by geographical regions can be the next step towards improving aviation safety in the United States and worldwide.

The purpose of this study was to determine the difference in the frequency of wildlife strikes from January 1, 2015 to December 31, 2019, at class B airports based on the geographical region the airports were within. A wildlife strike was defined as a collision reported in the FAA National Wildlife Strike Database (NWSD) (Federal Aviation Administration [FAA], 2020a) between an animal and an aircraft in movement within 250 feet from the runway centerline or 1,000 feet from the runway end (FAA, 2013). All civil class B airports were categorized into five regions of the contiguous US for the purpose of this study: Region 1, Region 2, Region 3, Region 4, and Region 5 (see Figure 1). The separation of airports into

also due to the financial burden of repairing an aircraft after a collision. On average, the FAA NWSD receives 39 reports per day (FAA, 2020b). According to a study by Bowie et al. (2013), the number of reports made has increased substantially in recent decades due to the use of faster, quieter aircraft, increased air traffic, and more species of wildlife in urban areas. This shows that as an aircraft operator in the United States, wildlife strikes are an unavoidable issue.

If an aircraft hits an animal, the operator must pay for damage repair while also losing money due to downtime operations. If the strike was substantial enough to result in a delay or flight cancellation, the aircraft operator might also be responsible for passenger and crew hotel accommodations and reallocation expenses. Using reports from the FAA NWSD, Dolbeer et al. (2015) found that the average aircraft downtime after a strike was 119.4 hours per incident. Moreover, the average repair cost per incident was \$169,349, and other monetary losses from passenger accommodation and flight cancellations averaged \$28,596 per incident. A limiting factor to these statistics was the lack of reported information about monetary expenses. Only 35.6% of reports made from 1990 to 2014 provided an estimated aircraft downtime; 16% provided an estimated cost of repair; 12% provided an estimated amount of other expenses from hotel accommodation, and flight rescheduling or cancellation (Dolbeer et al., 2015). Overall, the FAA (2013) found that wildlife strikes create about \$718 million in damage and 567,000 hours of downtime for civil aircraft alone.

These findings express the financial importance of moderating wildlife strikes. Roca-González et al. (2019) recommended more involvement from aircraft operators in supporting mitigation programs. This research suggested that airlines collaborate to more actively assist with the completion of Wildlife Management Plans (WMP) at airports and implement mandatory reporting rules within their organizations. With more awareness and proactive actions, aircraft operators may be less burdened by the costly expense of wildlife strikes.

2.2. Wildlife Strikes and Aircraft Airworthiness

Wildlife strikes create serious aerodynamic concerns for pilots. The smoothness of the wing's surface determines the effectiveness of an airplane's performance; therefore, the wings of an aircraft would not be able to efficiently produce lift if the leading edge was subject to a dent, fracture, or other forms of structural damage (FAA, 2016). The impairment differs in severity depending on several factors, such as the species of wildlife, size of the animal, and velocity on impact.

In several reported cases of wildlife strikes, the affected aircraft lost the ability to operate normally, sometimes leading to an accident. This is more probable after a high-velocity impact. A study by Goraj and Kustron (2018) indicated that high-velocity collisions cause the airframe of the aircraft to seem more delicate because, at faster speeds, there is not enough time for vibrations to develop during the collision. Since jet aircraft often operate at class B airports, many reports occurred at higher speeds due to the increased takeoff and landing velocities used by larger and heavier aircraft.

2.3. Airport Prevention

In the past few decades, wildlife strikes have received greater attention from governing agencies. After serious accidents resulting from a collision with an animal occur and are investigated, new regulations are put into effect to correct deficiencies in the safety of the aviation industry. Airports have been required to make several design changes to adjust for these regulations and make operations safer for aircraft.

Per 14 CFR Part §139.337, airports that experience multiple wildlife strikes or substantial damage from one must undergo a wildlife hazard management assessment (Wildlife Hazard Management, 2020). WMPs are developed by airport personnel in response to this assessment to provide a safer environment

for aircraft operators. An effective WMP is developed by first observing the wildlife that inhabits the area to then determine the most effective tools that can be implemented to deter the animals from nesting in a location. Some examples of these tools include non-fatal electrical fences, wire coils, and lighting systems. This study focuses on class B airport operations, some of the busiest airports in the United States. Since most flight operations are commercial air transport, wildlife strikes possess a greater risk to aviation safety at these airports. Therefore, a reliable WMP that reduces the likelihood of experiencing a collision with wildlife must be adhered to.

Several things attract animals to airports, including water, food, and shelter (DeVault, 2017). Water sources can include natural or man-made bodies of water that provide a habitat to species of animal. Many certificated airports have storm draining systems where water is collected away from airport surface areas like runways and taxiways to ensure safer aircraft operations. However, this conflicts with wildlife management as this temporary pooling of water can attract animals. Grassy, weedy zones around airport movement areas are also attractive to rodents as well as many vegetarian and carnivorous bird species hunting for food. Lastly, trees and shrubs, airport buildings, and grasslands are all forms of shelter for different species (DeVault, 2017). As an example, deer and small mammals find cover in tall grasslands, resultingly attracting predators like coyotes and hawks. For these reasons, it is vital that an airport takes action to mitigate the risk of wildlife strikes.

Jackson Hole Airport in Wyoming commonly experiences wildlife strikes because of the surrounding environment that attracts animals, leading to interference with flight operations. National parks, mountainous areas, bodies of water, and animal refuge areas are all in the vicinity of this airport (Harmon & Hirschert, 2016). Although Jackson Hole provides an extreme example of the impact of wildlife attractors in the airport vicinity, this applies at a smaller scale to other airports in the United States. In different geographical areas, different factors provoke wildlife. For example, in Florida - which is part of Region 4 in this study - various bodies of water like swamps, rivers, lakes, and oceans all attract different species of wildlife. This can be hazardous for airport operations in this region.

An additional consideration is climate differences and the impact this has on migration. Migratory birds that nest in northern regions travel to warmer temperatures during the early months of Fall and make their return in the early months of Spring. As a result, more bird strikes are reported during the months of migration (Dolbeer, et al., 2015). Moreover, migrant birds use two primary routes to travel between northern and southern states in the United States: they travel north and south between the Yucatan peninsula and the Mississippi River or take an east and west track over the Gulf of Mexico towards Florida (Lafleur, 2016). Airports located near these migratory paths should expect more strikes during migration months, and aircraft flying near these paths are at greater risk of experiencing a bird strike. Additionally, airports in regions of warmer climates should expect and plan for an influx in bird populations around the field between the Fall and Spring seasons.

2.4. FAA Reporting System

The FAA NWSD is a public database containing reports from 1990 to the present year. Reports are voluntary but highly encouraged by the FAA. The number of reports made in the last two decades has steadily increased, due in part to the amplified emphasis the FAA and other agencies have placed on the importance of reporting. Although reports were being collected by the FAA since 1990, the government did not disclose this information to the public until 2009 (FAA, 2013). The FAA's decision to give unrestricted access to this data most likely played a substantial role in increasing the number of received reports. Moreover, several organizations such as flight schools and aircraft operators require pilots to report a wildlife strike if encountered. On the other hand, privately-owned operators and general aviation pilots are usually not required to report by an organization, leading to more undocumented incidents. Since this research study focuses solely on class B airports - which have more commercial flights - many

collected reports were required to be made by the aircraft operator. Nevertheless, the FAA NWSD is a credible database and includes a majority of wildlife strikes.

2.5. Summary

Wildlife strikes are a concern to the aviation industry for a variety of reasons. Not only are they a significant expense, but they also threaten the safety of individuals. As previously mentioned, it is impossible to predict the location or severity of a wildlife strike, which demonstrates the importance of constant preparedness. Several groups play a role in ensuring this preparation, including the FAA, pilots, aircraft operators and manufacturers, and airports. Further knowledge and understanding can be gathered by analyzing how wildlife strikes differ by geographical region of the contiguous US. More strikes can be prevented by considering regional trends, and aviation can become an even safer industry.

3. Methodology

This study was ex-post-facto research. The investigation used events that have already occurred and data that has been previously collected to compare the reported wildlife strikes in various geographical regions in the United States.

The sampled airports were all the civil, class B airports in the contiguous United States; a total of 35 airports were analyzed in this study. This allowed for more consistent amounts of traffic and similar flight operations across airports. The definition of regions was based on the USGS official model (Vandegrift, 2013) with some adjustments. Regarding this model, the airports were separated into five regions: Region 1, Region 2, Region 3, Region 4, and Region 5 (see Figure 1). The class B airports have been divided into these groups based on geographical location. The sole source of data was the FAA NWSD (FAA, 2020a). The database provided several variables about reported strikes, such as the date and time, airport, operator (if applicable), aircraft, engine type, severity, and species of the animal. The primary data of interest was the airport and how frequently strikes were reported. All wildlife strikes from January 1, 2015 to December 31, 2019, reported to the database at the selected airports were included.

First, the data from the FAA NWSD (FAA, 2020a) was assessed and organized by region. Then, the regions were compared based on the frequency of reported strikes. The data were analyzed using descriptive and inferential statistics. The descriptive statistics included the mean, median, mode, range, and standard deviation of the number of wildlife strikes by region. Inferential statistics were calculated using a one-way between-subjects ANOVA. For a significant main effect, Tukey's pairwise comparison was conducted to determine the significant difference between each pair.

4. Results

From the beginning of 2015 to the end of 2019, there were 27,036 wildlife strike reports made at the sampled airports. This included 2,132 from Region 1, 7,170 from Region 2, 5,145 from Region 3, 6,606 from Region 4, and 5,983 from Region 5. KDEN was the airport with the greatest number of strikes within the given time period (2,230), while KSAN had the least number of strikes (146). No data was excluded from the results.

Figure 2 and Table 1 show the mean monthly frequency of reported wildlife strikes per region. Region 2 had the highest mean frequency of the five regions at approximately 86 strikes per month. KDEN had the highest mean frequency of all class B airports at 185.8 strikes per month, and KDFW, which was also part of Region 2, had the second-highest mean frequency at 160.9 (Figure 3).

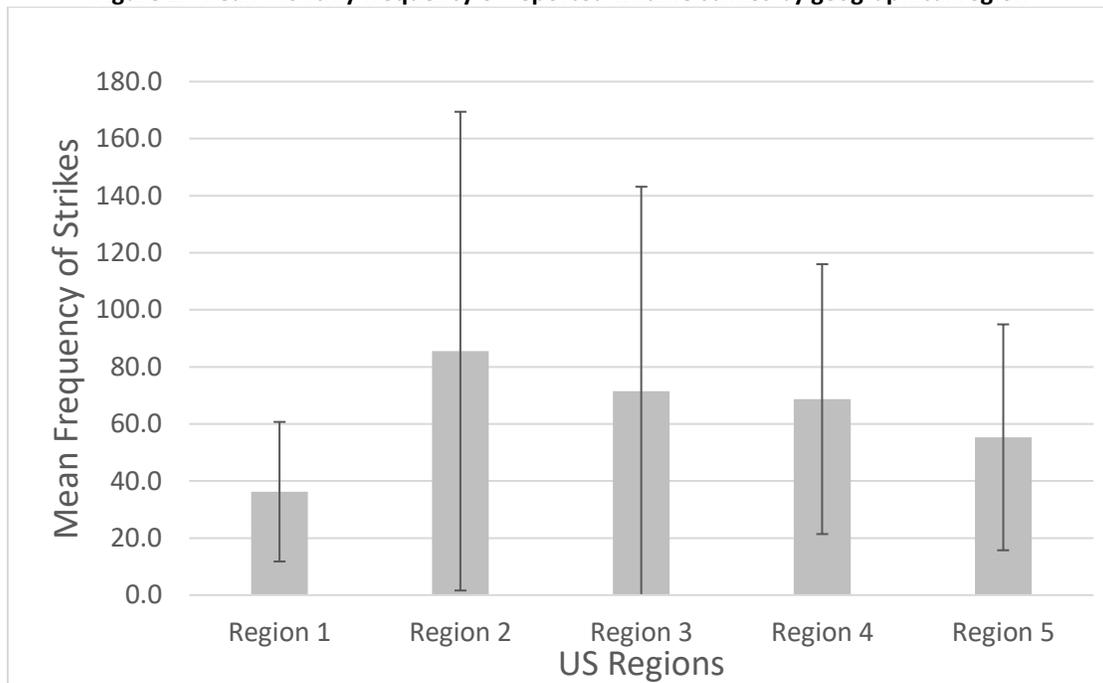
Region 3 had the second-highest mean frequency per region with 71.5 strikes per month, and Region 4 and Region 5 fell closely behind (Table 1 and Figure 2). Regions 2 through 5 had similar wildlife strike mean frequencies, but Region 1 appeared to have more varied results. The mean frequency of wildlife strikes in Region 1 was 36.3 strikes per month (Table 1). The largest range was Region 2 at 409 strikes, and

the smallest range was Region 1 with 119 strikes (Table 1).

Table 1. Descriptive statistics for monthly wildlife strikes at class B airports by region

	Mean	Median	Mode	Maximum	Minimum	Range	Standard Deviation
Region 1	36.3	29	15	120	1	119	24.5
Region 2	85.5	59	8	415	6	409	83.9
Region 3	71.5	45	38	338	1	337	71.7
Region 4	68.7	57	31	207	6	201	47.3
Region 5	55.3	47	12	191	3	188	39.6

Figure 2. Mean monthly frequency of reported wildlife strikes by geographical region



Note. Error bars indicate the standard deviation of the mean frequency.

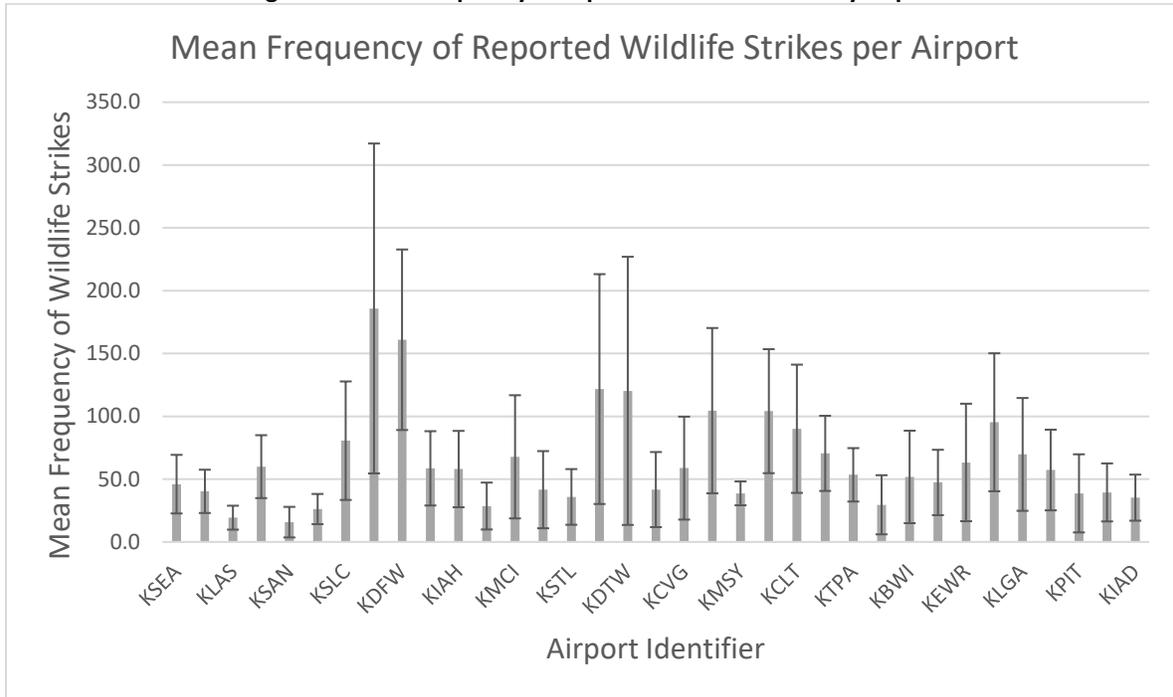
The one-way analysis of variance (ANOVA) indicated an effect of geographical region on monthly wildlife strikes, $F(4, 415) = 7.55, p < 0.001$. Tukey's Multiple Comparison of Means analysis found four region pairs to be significantly different: 1-2, 1-3, 1-4, 2-5 (Table 2). This showed that Region 1 was different from regions 2, 3, and 4. Additionally, Region 2 was significantly different from Region 5. The effect size was medium ($\eta^2 = .07$).

5. Discussion

The data supported part, but not all, of the hypothesis that the regions in order of greatest frequency to least frequency would be Region 2, Region 4, Region 1, Region 5, and lastly Region 3. Region 2 had the greatest frequency of wildlife strikes on average (Table 1) with KDEN and KDFW accounting for most of the strikes in that region (Figure 3), and this was in accordance with the hypothesis. However, Region 3 had a larger average frequency than predicted with only 14 strikes fewer than Region 2 (Table 1). Moreover, the results from Region 1 were much lower than expected and had the smallest average frequency compared to the other regions. This is potentially a result of differing average traffic flows managed daily by each of these airports. Also, Region 1 consisted of five airports – the fewest out of all

regions. Airports in Region 1, such as KLAS and KSAN, manage fewer flights compared to other airports in this study, such as KJFK and KATL.

Figure 3. Mean frequency of reported wildlife strikes by airport



Note. Error bars indicate the standard deviation of the mean frequency.

Table 2. Tukey's pairwise analysis of regions

Region Pairs	P-value
1-2	< 0.001
1-3	0.004
1-4	0.006
1-5	0.237
2-3	0.545
2-4	0.286
2-5	0.003
3-4	0.998
3-5	0.346
4-5	0.458

KDEN had the highest wildlife strike frequency out of all other airports considered in this study (Figure 3). As a result of this high strike frequency, KDEN wildlife risk management has set measures in place to help reduce the threat of wildlife strikes on aviation safety (Scheideman et al., 2017). The FAA NWSD shows that one of the most common species involved in strikes around this area was the horned lark, which is a small bird that often does not result in serious damage to the aircraft. Moreover, KDEN's airport elevation is the highest of all airports considered in this study. This is an appealing environment for many bird species to nest.

The one-way ANOVA indicated a significant effect of geographical region on the frequency of strikes. The results of Tukey's pairwise test (Table 2) determined the regions that were significantly different. Region 1's difference from almost all other regions coincides with having the fewest reports and a much

smaller mean frequency of wildlife strikes. This small number of wildlife strikes in Region 1 may be due to the deserted environment that covers most of this area in addition to fewer daily airport operations at many of the airports within this region. The pairwise comparison also determined a significant difference between Regions 2 and 5. As mentioned previously, the location and surrounding environment of an airport can play a major role on the number of wildlife strikes. Region 5, representing the northeastern United States, encompassed a highly-populated area of the country with fewer habitats for animals to occupy. This was potentially a reason why Region 5 had a small mean monthly frequency.

Overall, the results of this data analysis show that Region 2 stands out as the region with the highest wildlife strike count and highest mean frequency even though it did not have the greatest number of airports. Region 3, Region 4, and Region 5 had similar frequency counts, while Region 1 had the lowest frequency, with some months in that group having only one reported wildlife strike.

5.1. Future Research

A limitation of this research study was the FAA reporting system and the fact that some wildlife strikes are not reported into the database, which likely impacted the results of this study. Many organizations, such as airlines and Part 141 flight training programs, mandate reporting; however, general aviation flights undergo voluntary reporting by the pilot. Future research could explore the reporting rate more thoroughly to determine what operators are less inclined to report wildlife strikes and potential reasons why. Moreover, the results of this study are largely dependent on how the regions of the United States were defined. Other viable classifications of airports into regions will result in different outcomes that still provide an accurate understanding of the relationship between regions and wildlife strikes.

Other future research could explore the reasons for differences between frequencies of strikes at airports with similar surrounding environments and animal species, as well as control for differences in average air traffic values managed by each airport. The data collected can be used to determine expected wildlife behavior during specific months. This research could be the next step in understanding wildlife behavior in airport environments and potential causes of wildlife strikes.

5.2. Conclusion

The frequency of reported wildlife strikes differed depending on the geographical region that the airport was located in. This study researched all wildlife strikes reported in the FAA NWSD from the years 2015 to 2020 at all civil class B airports in the contiguous United States. The results showed that there was a significant difference in the frequency of strikes between certain regions of the United States. While the findings of this study provide noteworthy preliminary results, future research should investigate the accuracy and reliability of the NWSD and how it may encourage aviation safety as well as species' behavior around airports.

6. References

Bird Strike Damage, 14 C.F.R. § 25.631 (2020).

Bowie, R., Harmon, L., & Hirschert, D. (2013). Reducing wildlife hazards to aviation at BWI Thurgood Marshall Airport through agency and community coordination. *Journal of Airport Management*, 7(3), 265–288.
<https://portal.lib.fit.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=88997557&site=eds-live>

Crain, A.R., Belant, J.L., & DeVault, T.L. (2015). Carnivore incidents with U.S. civil aircraft. *Transportation Research Part D*, 36, 160-166. <http://dx.doi.org/10.1016/j.trd.2014.12.001>

- DeVault, T.L., Blackwell, B.F., Belant, J.L., & Begier, M.J. (2017). Wildlife at airports. United States Department of Agriculture, Animal & Plant Health Inspection Service.
https://www.aphis.usda.gov/wildlife_damage/reports/Wildlife%20Damage%20Management%20Technical%20Series/Wildlife-at-Airports-WDM-Technical-Series.pdf
- Dolbeer, R., Wright, S.E., Weller J.R., Anderson, A.L., Begier, M.J. (2015). Wildlife strikes to civil aircraft in the United States 1990-2014. United States Department of Transportation. (7548).
<https://rosap.ntl.bts.gov/view/dot/7548>
- Federal Aviation Administration (FAA). (2020a). FAA wildlife strike database. U.S. Department of Transportation.
<https://wildlife.faa.gov/search>
- FAA. (2020b). Fact Sheet – The Federal Aviation Administration’s wildlife hazard mitigation program.
https://www.faa.gov/news/fact_sheets/news_story.cfm?newsId=14393
- FAA. (2016). Pilot’s handbook of aeronautical knowledge. United States Department of Transportation. FAA-H-8083-25B.
- FAA. (2013). Advisory Circular 150/5200-32B, Reporting Wildlife Aircraft Strikes.
https://www.faa.gov/documentLibrary/media/Advisory_Circular/150_5200_32b.pdf
- Frank, R. (2013). Covering Captain Cool: The “Miracle on the Hudson” as a Hero Tale. *Western Folklore*, 72(1), 59.
https://www-jstor-org.portal.lib.fit.edu/stable/24550907?seq=3#metadata_info_tab_contents
- Google Maps, (2020).
<https://www.google.com/maps/d/u/0/viewer?mid=1umUey5pQlgkWDpgQHnfOJe38I5hJ4z5u&ll=35.20579004477989%2C-113.93987460000001&z=3>
- Goraj, Z. J., & Kustron, K. (2018). Review of current research trends in bird strike and hail impact simulations on wing leading edge. *Aircraft Engineering and Aerospace Technology*, 90(4), 602-612.
<http://dx.doi.org.portal.lib.fit.edu/10.1108/AEAT-02-2017-0053>
- Harmon, L., & Hirschert, D. (2016). Managing safety and wildlife at Jackson Hole Airport. *Journal of Airport Management*, 10(4), 381–401.
<https://portal.lib.fit.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=119586342&site=eds-live>
- Lafleur, J.M., Buler, J.J., & Moore, F.R. (2016) Geographic position and landscape composition explain regional patterns of migrating landbird distributions during spring stopover along the northern coast of the Gulf of Mexico. *Landscape Ecol* 31,1697–1709. <https://doi.org/10.1007/s10980-016-0354-1>
- Roca-González, J.L., Vera-Lopez, J.A., & Rodriguez-Bermudez, G. (2019). Organisational and costing aspects to prevent wildlife strikes on airports: A case study of Spanish airport security managers. *Journal of Safety Science*, 122(104520). <https://doi-org.portal.lib.fit.edu/10.1016/j.ssci.2019.104520>
- Scheideman, M., Rea, R., Hesse, G., Soong, L., Green, C., Sample, C., & Booth, A. (2017). Use of wildlife camera traps to aid in wildlife management planning at airports. *Journal of Airport Management*, 11(4), 408–419.
<https://portal.lib.fit.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=125398657&site=eds-live>
- Vandegraft, D.L. (2013). Cartographic and GIS activities in the U.S. Fish and wildlife service. U.S. Geological Survey.
<https://pubs.usgs.gov/of/2005/1428/vandegraft/index.html>
- Wildlife Hazard Management, 14 C.F.R. § 139.337 (2020).

Impact of Organizational Slack on Innovation Outcomes

John Hill¹

Abram Walton, Ph.D.¹

¹Nathan M. Bisk College of Business, Florida Institute of Technology

Johnhill1@gmail.com; awalton@fit.edu

Abstract

This research analyzes the relationship between the impact of organizational slack on patent productivity as a proxy for innovation and the resulting firm performance by longitudinally investigating 114 U.S. firms. The role of slack as a moderated measure of innovation was also explored, with its impact on technology and non-technology industry firms compared. In study one, the evidence concluded that slack is positively correlated with innovation productivity, but the company type was not statistically significant. In addition, the study did not find an inversion in the relationship between slack and innovation that would produce a negative correlation as the level of slack in firms reaches a tipping point. Study two demonstrated a positively correlated relationship between the rate of innovation and firm growth. The presence of slack positively moderates this relationship, and this moderation effect increases for technology companies.

1. Introduction

While definitive research has been conducted on the positive relationship between slack and innovation (Singh, 1986; Damanpour, 1987; Marlin & Geiger, 2015), there remain some studies (Nohria & Gulati, 1996; Geiger & Cashen, 2002) that indicate there is an inverted-U type relationship between slack and innovation. When slack and innovation exist at lower levels in the firm, their relationship is positive. However, as the level of slack in a firm increases, a point of diminishing returns is created. Eventually, a tipping point is reached where increases in slack decrease the amount of innovation in the firm. However, these studies used R&D intensity (Nohria & Gulati, 1996; Geiger & Cashen, 2002) to measure innovation which is defined as R&D expense divided by total sales. As a result, these studies reflect merely an inverted U-shaped relationship between innovation effort and slack. This research will seek to test the relationship between slack and innovation using patents, instead of R&D intensity, as the measure of innovation to evaluate if evidence of the supposed curvilinear relationship exists. It is expected that this inverted U-shape relationship will not be present, and instead, a positive correlation between patent productivity and slack will exist.

Much of the research into the relationship between innovation and firm performance has been focused on profitability-type measures. A meta-analysis of 66 studies by Daniel et al. (2004) demonstrated a positive effect of slack on ROA, ROE, or ROI, while none of the studies cited measured the impact on firm growth. The relationship between innovation and firm growth has not received the same level of inquiry and has mixed results, although much of the research indicates a positive correlation between innovation and firm growth (Crepon et al., 1998; Coad, 2009; Ortega-Argiles et al., 2011). However, no study has

been conducted to determine the role of slack in moderating the relationship between innovation and firm growth.

This research set out to re-examine the relationship between slack and patent productivity as a proxy for innovation and to understand the potential as a moderator between innovation and firm growth. The two studies in this research involved a longitudinal analysis of 114 U.S.-based firms that were granted 40 or more patents per year over five years. A further decision was made to analyze the differences between technology firms and non-technology firms across the research questions. Thus, the research aimed to investigate the following hypotheses:

- H1: The level of available and recoverable slack in a firm is positively correlated with innovation productivity.
 - H1a: Being a technology company increases this correlation
 - H1b: There is a positive relationship between slack and innovation
- H2: Increases in the rate of innovation productivity within a firm are positively correlated with changes in revenue.
 - H2a: Being a technology company increases this correlation
- H3: Available slack positively moderates the relationship between innovation rate changes and revenue changes.
 - H3a: Being a technology company increases this correlation

2. Literature Review

2.1 Innovation and Slack

There have been varying approaches to the definition and measurement of innovation in a firm. One popular method uses R&D intensity as the dependent variable (Hansen & Hill, 1991; Hitt et al., 1996; Hitt et al., 1997; Geiger & Cashen, 2002). However, R&D intensity is simply a measure of the innovation effort, as it reflects the sum of firm expenditures on research and development activities, whereas innovation is better thought of as an output of the efforts of teams and individuals (International Standards Organization, 2020). Other research has used patent quality (Herold et al., 2006) or product count (Sorescu & Spanjol, 2008) as the dependent variable, but issues remain with these measures as well. Patent quality does not reflect all innovation even within the patent space as it focuses only on patents receiving citations in research. Product count is a problematic proxy for innovation as it only captures new product launches and misses out on expansions to current products that have benefited from innovation, the improvements of which often consume the majority of R&D investments. Finally, another proxy for innovation is the use of patent counts (Acs & Audretsch 1989; Acs et al., 2002). While one could argue that not all patents are worthwhile and that not all innovations are patented, patents do represent a solid proxy of innovation outputs (International Standards Organization, 2019 & 2020), particularly for technology companies.

A firm's ability to innovate can be adversely impacted due to its resources being fully allocated to operational tasks or other firm initiatives (Voss et al., 2008). The complete absorption of resources in a firm can create an obstacle for firms when they need to react to changing market conditions or innovation opportunities. To enable this adaptive capability, the firm is faced with the need to create slack. Bourgeois defines slack as "that cushion of actual or potential resources which allows an organization to adapt successfully to internal pressures for adjustment or to external pressures for change in policy, as well as to invite changes in strategy with respect to the external environment" (1981, p 30). Another definition of slack defines it as the excess delta between what resources a firm needs to meet a specific objective and those available to them (Geiger & Cashen, 2002). A key characteristic of slack is that it provides the firm a level of discretion on the utilization of the resources and still meets the desired outcomes (Cheng

& Kesner, 1997). The creation of slack within a firm can be achieved in several ways; executives can achieve slack by bundling their resources (Sirmon et al., 2011) or deliberately creating a buffer or cancellation of planned activities (Voss et al., 2008).

Slack can be broken down into several types defined by the degree that slack is immediately available. The three types of slack range from those that are readily available (or unabsorbed slack), recoverable (or absorbed slack), and potential slack (Bourgeois, 1981; Cheng & Kessner, 1997; Palmer & Wiseman, 1999; Geiger & Cashman, 2002). Available slack, or unabsorbed slack, refers to resources readily available to the firm but not deployed. Recoverable slack, or absorbed slack, refers to those resources currently deployed within the firm but which could be cut to create financial capacity for the firm in the event of financial distress (Bourgeois & Singh, 1983). It is expected that available and recoverable slack is expected to impact innovation productivity rates given their more immediately deployable (or repurposable) nature. However, for the moderating impact between innovation and firm performance, only available slack is expected to have an impact as the slack must be immediately available during the year. To measure these constructs of slack various firm financial ratios can be utilized, such as the quick ratio as a proxy for available slack, the ratio of general and administrative expense to sales as a proxy for recoverable slack, and the firm's debt to equity ratio as a proxy for potential slack (Bourgeois, 1981; Cheng & Kessner, 1997; Palmer & Wiseman, 1999; Geiger & Cashman, 2002).

The role of slack in improving innovation outcomes has received significant research attention. Firms can use slack as a method of driving experimentation and innovation (Levinthal & March 1981). An organization can pursue innovation opportunities as the existence of slack protects the organization from adverse impacts from those innovation endeavors (Bourgeois, 1981). However, the correlation between slack and innovation outcomes has produced mixed innovation results. Slack was demonstrated to positively affect innovation in several empirical studies (Damanpour, 1987; Singh, 1986), while inconclusive in determining innovation in others (Zajac, Golden, & Shortell, 1991).

There are contrarian views of slack, and some researchers have argued that its existence is a measure of managerial incompetence (Leibenstein, 1969), a driver of managerial risk-taking (Palmer & Wiseman, 1999), or a precursor to wasteful spending on ill-conceived R&D efforts (Jensen, 1993). There is even an argument that excess slack can lead to poor oversight on portfolio management (Jensen, 1993, Leibenstein, 1969). Low levels of slack can result in firms embarking on riskier paths (Bromiley, 1991; Wiseman and Bromiley, 1996; Palmer & Wiseman, 1999). Several empirical studies indicate that slack and innovation have an inverted U-shaped relationship (Nohria & Gulati, 1996; Geiger & Cashen, 2002), such that too little slack discourages innovation, while too much slack breeds complacency and, as a result, increases the risk of poor projects being pursued; Tan and Peng (2003) found a similar relationship between slack and firm performance.

2.2 Innovation and Firm Growth

The relationship between innovation and firm growth has generally been found to be a positive one (Crepon et al., 1998; Coad, 2009; Ortega-Argiles et al., 2011). However, some research has found that the positive relationship between innovation and firm growth is primarily concentrated in fast-growing firms (Coad & Rao, 2008). Further research has suggested that firm characteristics such as size and approach to patenting are key to the relationship between innovation and firm growth (Demirel & Mazzucato, 2012).

2.3 Slack as a Modifier between Innovation and Firm Growth

The creation of innovative products or services by a firm must be commercialized to extract value. It is expected that the ability to commercialize an opportunity quickly is moderated by the ability of the firm

to pivot and re-align resources. The use of slack resources can be instrumental in providing an opportunity for the firm to react to contingencies or to drive institutional change or innovation (Chandy & Tellis, 1998; Gatignon & Xuereb, 1997). Slack's specific role in commercialization processes has also been explored. For example, Danneels (2008) demonstrated that slack was vital in building second-order marketing competencies. The presence of slack gives firm management the capacity to react to constraints and redirect lower-order resources to create the higher-order capabilities that are necessary to drive the implementation of a new product or service. The firm must create slack to respond to changes in project priorities and market conditions effectively.

3. Methodology & Results

This research consists of two studies that utilize that same data set of U.S.-based companies on the NASDAQ or NYSE with 40 or more patents each year from 2015 to 2019; excluding those in the medical field. Firms in the medical industry suffer from exceptionally long cycle times on their patent-based innovations, and including those firms would have required a substantially longer longitudinal study-time period, while their results would have been statistically incompatible with firms with shorter patent cycles. While 124 firms met the threshold of >40 patents, which was ultimately reduced to 114 firms after adjusting for mergers and companies that were not publicly traded for all five years. Of the 114 firms, 71 of them were classified as technology firms and 43 as non-technology firms. A technology firm was defined as a firm whose primary business was software development, social networks, telecommunications, I.T. hardware development, I.T. services, or I.T. component-level manufacturing. The patent information was collected from the United States Patent Office website (USPTO, 2021a), and the financial information from the companies' income and balance sheets was retrieved from Yahoo Finance.

3.1 Study One: Relationship Between Slack and Innovation Productivity

Study one is focused on the relationship between slack and innovation, with:

- H1: The level of available and recoverable slack in a firm is positively correlated with innovation productivity.
 - H1a: Being a technology company increases this correlation
 - H1b: There is a positive relationship between slack and innovation

The dependent variable is patent productivity, which is defined as the number of patents granted per \$1B in revenue. The independent variables are the lagged measures of available slack and recoverable slack. The lag was done on a two-year basis as the average time total pendency time is 22.9 months from the time it is filed with U.S. Patent Office (USPTO, 2021b). Total pendency time is defined as the average time from filing to grant or abandonment by the USPTO. Given the difficulty in assigning the variables precisely to a time period, an average over the three years of 2017 to 2019 was used for each of the dependent and independent variables. A quadratic version of the independent variables was used to test the curvilinear relationship. The control variable is the type of company, as either technology or non-technology.

Linear regression was used to predict the impact of available and recoverable slack on the patent productivity rate for a firm. A visual inspection suggested a negative to positive curve for recoverable slack. Therefore, a quadratic version of recoverable slack was added to the model. The results indicate that available slack is positively correlated with patent productivity, while recoverable slack has a negative to a positive curvilinear relationship with patent productivity (See Table 1). The p-value for available slack is .001, and for recoverable slack and the quadratic version of recoverable slack, the p-values are less than .001. However, there was no statistically significant relationship between the type of company and the patent productivity rate. An alpha level of .05 was used for all statistical tests.

Table 1
Regression Analysis Summary for Slack Predicting Patent Productivity

	B	95% CI	β	t	p
(Constant)	211.358	[85.838, 336.877]		3.337	0.001
Available Slack	40.817	[16.833, 64.801]	0.158	3.373	0.001
Recoverable Slack	-2563.405	[-3239.986, -1886.825]	-0.869	-7.509	<.001
Recoverable Slack Squared	4688.146	[4007.991, 5368.301]	1.577	13.661	<.001
Company Type	-7.042	[-111.473, 97.388]	-0.006	-0.134	0.894

Notes: $R^2 = .791$, CI = Confidence interval for B

Therefore, H1 is accepted, H1b is partially accepted, and hypothesis H1a is rejected.

3.2 Study 2: Relationship between innovation production, firm growth, and slack

Study 2 focuses on understanding the moderating effect of slack between the rate of innovation and firm growth, with:

- H2: Increases in the rate of innovation productivity within a firm is positively correlated with changes in revenue.
 - H2a: Being a technology company increases this correlation
- H3: Available slack positively moderates the relationship between innovation rate changes and revenue changes.
 - H3a: Being a technology company increases this correlation

In this model, the dependent variable is the firm revenue growth rate, and the independent variable is the growth rate in the number of patents per \$1B in revenue. The moderating variable is the available slack growth rate, and the control variable is the status as a technology company or not. Two statistical tests were conducted to assess this relationship. The process macro in SPSS was used to test the regression impact from the addition of the moderating variable. However, as this model is quite simplistic and its use is limited to testing the relationship between the three variables, the regression model would not produce a high R^2 . Therefore, the Pearson correlation coefficient was calculated with and without the addition of slack as an interaction.

The results indicate a weak but statistically significant positive correlation between patent growth and revenue growth (See tables 2 and 3). The Pearson r was calculated as .132 with a p -value of .005. The addition of available slack as a moderator increased the correlation to .185 with a p -value of <.001. The same calculations were run within the data sets comprised of only technology and only non-technology firms. The increase in correlation associated with the moderation effect was more significant in the technology firm sample. The Pearson r was calculated as .147 with a p -value of .013. The addition of available slack as a moderator increased the correlation to .295 with a p -value of <.001. However, in non-technology firms, the relationship between patent growth and revenue growth was not statistically significant.

Table 2

Correlations of Revenue Growth, Patent Growth, and Slack Moderating Effect (All Firms)

Variable	Revenue Growth	Patent Growth	Slack Interaction
Revenue Growth	1	.132**	.185**
Patent Growth	.132**	1	-.095*
Slack Interaction	.185**	-.095*	1

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

Table 3

Correlations of Revenue Growth, Patent Growth, and Slack Moderating Effect (Technology Firms)

Variable	Revenue Growth	Patent Growth	Slack Interaction
Revenue Growth	1	.147*	.295**
Patent Growth	.147*	1	-.182**
Slack Interaction	.295**	-.182**	1

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

The positive effect of slack between innovation and firm growth rate in technology firms was examined in a two-step regression using the PROCESS macros within SPSS. Patent growth rate and revenue growth rate were entered in the first step of the regression analysis. In the second step of the regression analysis, the interaction term between available slack growth rate and patent growth rate was entered, and it explained a significant increase in variance in revenue growth rate, $\Delta R^2 = .1167$, $F(3, 279) = 38.34$, $p < .001$. Thus, available slack was a significant moderator of the relationship between patent growth and revenue growth. Therefore, hypotheses H2, H2a, H3, and H3a are accepted.

4. Conclusions & Future Research

The research conducted reaffirmed the role of slack on innovation productivity. However, contrary to previous research, a positive to negative curvilinear relationship was not found between slack and innovation. Available slack demonstrated a positive relationship, and recoverable slack showed a negative to positive relationship. Further research is needed to understand this phenomenon better; as it may be that initial expenses in G&A crowd out innovation activity until the organization reaches a tipping point, at which point the increases in recoverable slack produce increases in innovation. The research also did not indicate any increased effect of slack on innovation for technology firms. The implication is that the innovation process is not as different as anticipated in technology firms with respect to the ability to deploy slack more efficiently.

The research identified a positive relationship between the rate of innovation and the rate of revenue growth. Slack significantly and positively moderates the relationship between the changes in the growth rate of innovation and revenue. This moderating effect is substantially higher in technology firms;

where the nature of the product(s) made by technology firms may be more receptive to immediate deployment through changes in slack rates.

There are several implications of this research for firm executives. They should understand the positive impact of slack on innovation and the commercialization of innovation in their firms. Their organizations must create capacity in their teams that allow them to easily set aside tasks that are not critically time-dependent and of lesser value. This does not mean that firms must have resources on the bench that are not deployed. The firm needs to avoid optimizing resources to 100% utilization that are tied to critical projects and operational tasks, which render it too difficult for the resources to adjust and reprioritize on opportunities that will drive innovation or deploy innovation.

5. Limitations and Future Research

There are a couple of limitations to the research conducted. First, the sample was limited to publicly traded U.S.-based firms awarded more than 40 patents. The nature of smaller, start-up organizations with a heavy emphasis on rapidly scalable innovation may not reflect the same dimensions. Second, the definition of slack utilizes financial ratios as proxies to estimate its rates. Efficiency differences between firms that may produce more slack in one of them over another would not be represented. Finally, the regression model for understanding the impact of slack on the relationship between innovation rate and revenue growth is limited, as it does not capture all factors that impact revenue growth.

This research aimed to understand the relationship of slack as a moderator first. Therefore, future research would expand to include non-U.S.-based firms and private firms while building a more complete regression model to understand the relationship between innovation rate and firm growth. A more thorough examination into understanding the negative to positive curvilinear relationship between recoverable slack and innovation productivity is also needed.

6. Author Names and Affiliations

Mr. John Hill is the Chief Information Officer and Senior Vice President, Business Planning for Carhartt. He has more than 20 years of executive experience in technology across several firms and industries. He is a DBA candidate at the Nathan M. Bisk College of Business at Florida Institute of Technology, and his dissertation research is focused on digitalization inefficiencies in firms.

Dr. Walton is the Director of the Center for Innovation Management and Business Analytics and Professor of Management and Innovation at Florida Institute of Technology. He is a U.S. Delegate and Working Group Chair on ISO's Technical Advisory Committee on Innovation Management Standards (ISO 56000) and is the Co-Founder and Former Deputy Editor-in-Chief of the International Journal of Innovation Science and the International Association of Innovation Professionals.

7. References

- Acs, Z., D. Audretsch, 1989, Patents as a measure of innovative activity, *Kyklos* 42, 171-180
- Acs, Z. J., Anselin, L., & Varga, A. (2002). Patents and innovation counts as measures of regional production of new knowledge. *Research Policy*, 31(7), 1069-1085.
- Bourgeois, L. J. (1981). On the measurement of organizational slack. *The Academy of Management Review*, 6(1), 29-39. doi:10.5465/AMR.1981.4287985
- Bromiley, P. (1991). Testing a causal model of corporate risk-taking and performance. *Academy of Management Journal*, 34, 37-59.

- Chandy, R. K., & Tellis, G. J. (1998). Organizing for radical product innovation: The overlooked role of willingness to cannibalize. *Journal of Marketing Research*, 35(4), 474-487. doi:10.1177/002224379803500406
- Cheng, J. L. C., & Kesner, I. F. (1997). Organizational slack and response to environmental shifts: The impact of resource allocation patterns. *Journal of Management*, 23(1), 1-18. doi:10.1016/S0149-2063(97)90003-9
- Coad, A., & Rao, R. (2008). Innovation and firm growth in high tech sectors: A quantile regression approach. *Research Policy*, 37, 633-648.
- Coad, A. (2009). *The growth of firms: A survey of theories and empirical evidence*. Cheltenham: Edward Elgar
- Crepon, B., Duguet, E., & Mairesse, J. (1998). Research, innovation, and productivity: An econometric analysis at the firm level. *Economics of Innovation and New Technologies*, 7, 115-158.
- Cyert, R. M., & March, J. G. 1963. *A behavioral theory of the firm*. Englewood Cliffs, NJ: Prentice-Hall.
- Damanpour, F. (1987). The adoption of technological, administrative, and ancillary innovations: Impact of organizational factors. *Journal of Management*, 13(4), 675-688. doi:10.1177/014920638701300408
- Danneels, E. (2008). Organizational antecedents of second-order competences. *Strategic Management Journal*, 29(5), 519-543.
- Daniel, F., Lohrke, F. T., Fornaciari, C. J., & Turner Jr, R. A. (2004). Slack resources and firm performance: a meta-analysis. *Journal of Business Research*, 57(6), 565-574.
- Demirel, P., & Mazzucato, M. (2012). Innovation and firm growth: Is R&D worth it? *Industry and Innovation*, 12(1), 45-62.
- Gatignon, H., & Xuereb, J. (1997). Strategic orientation of the firm and new product performance. *Journal of Marketing Research*, 34(1), 77-90. doi:10.1177/002224379703400107
- Geiger, S. W., & Cashen, L. H. (2002). A multidimensional examination of slack and its impact on innovation. *Journal of Managerial Issues*, 14(1), 68-84.
- Jensen, M. C. 1993. The modern industrial revolution, exit, and the failure of internal control systems. *Journal of Finance*, 48: 831-880.
- Hansen, G. S., & Hill, C. W. (1991). Are institutional investors myopic? A time-series study of four technology-driven industries. *Strategic Management Journal*, 12(1), 1-16.
- Hitt, M. A., Hoskisson, R. E., Johnson, R. A., & Moesel, D. D. (1996). The market for corporate control and firm innovation. *Academy of Management Journal*, 39(5), 1084-1119.
- Hitt, M. A., Hoskisson, R. E., & Kim, H. (1997). International diversification: Effects on innovation and firm performance in product-diversified firms. *Academy of Management Journal*, 40(4), 767-798.
- Herold, D. M., Jayaraman, N., & Narayanaswamy, C. R. (2006). What is the relationship between organizational slack and innovation?. *Journal of Managerial Issues*, 372-392.
- International Standards Organization. (2020, February). *ISO 56000:2020. Innovation management — Fundamentals and vocabulary*. Geneva, Switzerland.
- International Standards Organization. (2019, February). *ISO 56004:2019. Innovation Management Assessment — Guidance*. Geneva, Switzerland.

- Leibenstein, H. (1969). Organizational or frictional equilibria, X-efficiency, and the rate of innovation. *The Quarterly Journal of Economics*, 83(4), 600-623. doi:10.2307/1885452
- Levinthal, D., & March, J.G. (1981). A model of adaptive organizational search. *Journal of Economic Behavior and Organization*, 2: 307-333.
- Marlin, D., & Geiger, S. W. (2015). A reexamination of the organizational slack and innovation relationship. *Journal of Business Research*, 68(12), 2683-2690.
- Nohria, N., & Gulati, R. (1996). Is slack good or bad for innovation?. *Academy of Management Journal*, 39(5), 1245-1264.
- Ortega-Argiles, R., Potters, L., & Vivarelli, M. (2011). R&D and productivity: Testing sectoral peculiarities using microdata. *Empirical Economics*, 41, 817–839.
- Palmer, T. B., & Wiseman, R. M. (1999). Decoupling risk-taking from income stream uncertainty: A holistic model of risk. *Strategic Management Journal*, 20(11), 1037-1062.
- Sirmon, D. G., Hitt, M. A., Ireland, R. D., & Gilbert, B. A. (2011). Resource orchestration to create competitive advantage: Breadth, depth, and life cycle effects. *Journal of Management*, 37 (5), 1390-1412.
- Sorescu, A. B., & Spanjol, J. (2008). Innovation's effect on firm value and risk: Insights from consumer packaged goods. *Journal of Marketing*, 72(2), 114-132.
- USPTO (2021a). Retrieved from: https://www.uspto.gov/web/offices/ac/ido/oeip/taf/reports_topo.htm on January 11, 2021
- USPTO (2021b). Retrieved from: <https://www.uspto.gov/dashboard/patents/> on March 20, 2021.
- Voss, G. B., Sirdeshmukh, D., & Voss, Z. G. (2008). The effects of slack resources and environmental threat on product exploration and exploitation. *The Academy of Management Journal*, 51(1), 147-164. doi:10.5465/AMJ.2008.30767373
- Wiseman, R.M., & Bromiley P. (1996). Toward a model of risk in declining organizations: An empirical examination of risk, performance, and decline. *Organization Science*, 7, 524-543.
- Zajac, E. J., Golden, B. R., & Shortell, S. M. (1991). New organizational forms for enhancing innovation: The case of internal corporate joint ventures. *Management Science*, 37(2), 170-184. doi:10.1287/mnsc.37.2.170

Leadership Styles and Lean Six Sigma Effectiveness within the Aerospace and Defense Manufacturing Industry

Corey Gellis¹

Ahmad Elshennawy¹

¹*Department of Industrial Engineering and Management Systems, University of Central Florida*
cgellis@knights.ucf.edu; ahmad.elshennawy@ucf.edu

Abstract

The aerospace and defense (A&D) industry has shifted into a globally competitive market that is prioritizing innovative advancements in technological capabilities. Corporations are now having to further develop customer-focused strategies grounded in adding value while reducing costs. Large corporations often embrace continuous improvement methodologies, such as Lean Six Sigma (LSS), to execute these strategies. Though numerous benefits for continuous improvement programs have been realized, there are several factors that contribute to whether it will be successful or not. The type of leadership styles found within organizations heavily influence the outcomes of implementing the LSS methodology. This paper aims to provide a literature review summary of leadership styles and LSS within the A&D manufacturing industry.

1. Introduction

The A&D industry is a global infrastructure that supports the manufacturing of advanced aerospace and military products. Globalization has created a fast-paced competitive market that requires organizations to meet rapid changes in customer demand (Jonsdottir et al., 2014). Recent surges of innovations and advancements in technology have created customer-focused strategies of adding value and reducing costs to remain globally competitive against other organizations (Wang et al., 2018). The A&D industry has always been prone to budget cuts, thus spurring the need to compete as the technology matures and costs increase (Papin & Kleiner, 1998). U.S. Defense contractors are facing multiple challenges when addressing innovation including limited budgets for development and foreign threats from low-cost competition (Steinbock, 2014). The primary approach for addressing these limitations is by adding value through prioritizing and implementing continuous improvement methodologies.

Two common continuous improvement methodologies are Six Sigma and Lean. Six Sigma is a process improvement methodology that enables organizations to understand and improve their processes through higher rates of quality and lower operating costs (Antony, 2008; Suresh et al., 2012). Lean manufacturing is a method that aims to reduce waste or “non-value added” variables from processes without compromising productivity. Together LSS is a systematic approach that utilizes statistical analysis to minimize defects per million opportunities to 3.4 while simultaneously removing waste from production processes (Spedding & Pepper, 2010).

The technical nature of organizations within the A&D industry requires leaders who can adapt to shifting circumstances. Leadership is collectively defined as modeling values and beliefs that will empower and motivate people to unite to achieve a shared common goal (Emmerling et al., 2015; Yukl, 2011). Organizational goals remain rooted in providing value to customers while simultaneously driving out inefficiencies. Dating three decades, Hull (1990) argued that to survive in a global

economy the United States must continuously develop technology, shift focus to a global management perspective, and improve upon current work practices. The evolving digital environment has triggered higher customer demands that must be addressed through customization and agility within manufacturing (Sousa & Rocha, 2019).

Leadership theory has been heavily researched over the last century and has observed multiple theories. The theories range from behavioral approaches that focus on internal behaviors to inspirational vision-based approaches (Emmerling et al., 2015). A prominent leadership theory model proposed by Bass and Avolio (2004) provides a comprehensive multifactor leadership questionnaire (MLQ) that measures five transformational factors, 2 transactional factors, and 2 laissez-faire factors. The LSS methodology requires culture change, customer focus, process management, and statistical analysis of data (Antony, 2004). A common reason for the organizational failure of Lean Six Sigma effectiveness is due to leadership's lack of commitment and focus on the culture (Testani & Ramakrishnan, 2011). Leading a culture change to create an innovative environment through transformational leadership is one of the primary components for success (Chen & Zhang, 2011). This paper reviews the literature on different leadership styles and Lean Six Sigma to identify opportunities for the A&D manufacturing industry.

2. Background

Originally introduced in the 1980s by Motorola, Six Sigma has become one of the leading approaches for continuous improvement because it generated a global standard for measuring quality in relation to performance and cost (Stankalla et al., 2018). Though comparable to previous quality management techniques, leading organizations have touted that Six Sigma transformed their respective organization (Schroeder et al., 2012). Snee (2010) articulated that General Electric, Honeywell, Du Pont, and American Standard used the LSS methodology to spur leadership growth. Key findings from Laureani and Antony's (2017) systematic review exemplified the necessity for leadership when sustaining Lean Six Sigma improvements. In their study which focused on Six Sigma and leadership, Suresh et al. (2012) proposed future research on needing to validate leadership variables that would enable successful Six Sigma deployment.

A systematic review of continuous improvement failures in manufacturing environments by McLean and Antony (2014) identified a lack of management leadership as a core theme. Direct leadership styles favor process-focused continuous improvement while supportive leadership styles favor cultural improvement (Brown et al., 2008). Inability to identify processes for improvement through LSS create leadership impediments concerning project success and employee involvement (Pamfilie et al., 2012). Lack of successful projects or engagement from the team further muddles the leadership traits that are necessary to lead and facilitate the LSS methodology. Swain et al. (2018) research provided multiple newer leadership theories that still required understanding how leadership traits and characteristics impact LSS success. The continuous piecemeal contributions to leadership theory in relation to LSS have not yet collectively replaced the comprehensive model proposed by Bass and Avolio (2004).

3. Transactional and Transformational Leadership Styles

Transactional leadership and transformational leadership are comprised of two contrasting views: transactional focuses on task orientation and transformational focuses on relationship-oriented (Tyssen et al., 2014). The leader's role is to provide an infrastructure of policies and goals to facilitate the employee being able to execute tasks (Halaychik, 2016). Transactional leadership motivates employees to complete tasks through rewards or punishments. Three dimensions of transactional

leadership are contingent reinforcement, active management-by-exception, and passive management-by-exception.

Northouse (2018) defines contingent reinforcement as followers who subscribe to their leader's agenda of tasks for rewards or punishment. In management-by-exception, active leaders take initiative before goal departures occur while passive leaders do not take initiative until after the fact (Den Hartog et al., 1997). The key difference being that the active leaders are ahead of their problems while passive leaders are behind theirs. Though transactional leadership does have a purpose in some scenarios, the effectivity is often challenged. The reason for this is primarily due to a leader and employee relationship that is built on transactions that aim to reward or punish in each situation. Lack of consideration for other factors a leader or organization may face has led this theory to be highly criticized amongst scholars (McCleskey, 2014).

Transformational leadership is modeled through the four I's which are idealized influence, inspirational motivation, intellectual stimulation, and individualized consideration. Bass and Avolio (1993) summarized the four I's as follows: Idealized influence explains how well followers view their leader as a role model and someone they want to emulate in the work environment. Inspirational motivation is how well the leader can "paint the picture" of important goals and easily motivate followers to participate. Intellectual stimulation involves the degree to which the leader can creatively challenge their followers to problem solve and create a new baseline of standards. Individualized consideration is the leader's ability to cater to individual differences and personalities found within the followers of their team. Please reference Figure 1 below.

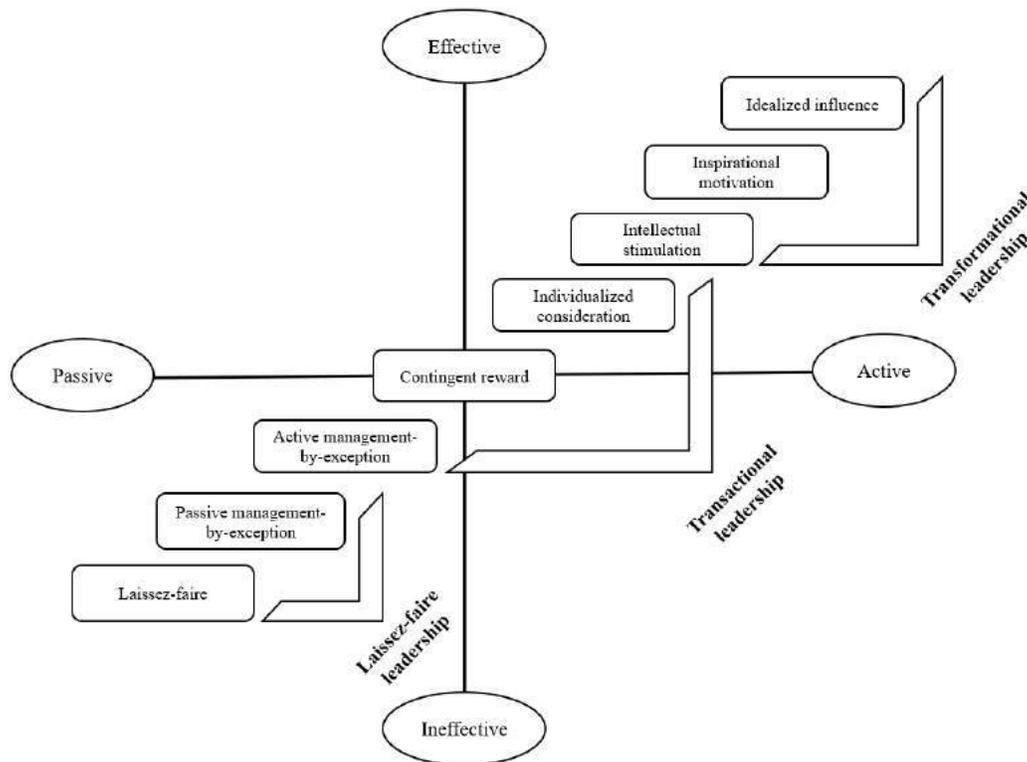


Figure 1. Full range leadership model diagram (Bass and Avolio, 1995)

Research investigating leadership styles and innovation in manufacturing companies found a significant relationship between transformational leadership and exploratory innovation (Ebrahimi

et al., 2016). The author’s population for the study included approximately 5000 manufacturing companies and utilized the MLQ proposed by Bass and Avolio. A study concerning CEO leadership styles and innovation found that transformational leadership styles were more effective when compared to transactional leadership in dynamic organizations (Prasad & Junni, 2016). Strang’s (2005) case study found a positive correlation between leaders displaying transformational behaviors and organizational output (deliverables, metrics, customer satisfaction). Xie et al. (2018) postulated transformational leadership is more conducive for innovative environments but transactional leadership provides value for teams in other situations. These positive findings for transformational leadership provide key insights for A&D corporations wanting to pursue exploratory innovation.

4. Lean Six Sigma

LSS encompasses the complementary benefits offered by both the Lean and Six Sigma continuous improvement methodologies. Implementing Lean in isolation narrows available tools for improvement while implementing Six Sigma in isolation results in a loss of strategic vision (Spedding & Pepper, 2010). Khaled’s (2013) analysis of the A&D industry found that the use of Six Sigma techniques was gaining more prominence due to reductions in costs and time. Zhang et al.’s (2012) literature review of Lean Six Sigma found that the military industry uses the methodology to focus on process improvement and root cause investigation. Arnheiter and Maleyeff’s (2005) comparative study on Lean and Six Sigma integration identified six primary tenets that are summarized in Figure 2. The Six Sigma approach provides the lowest cost for the producer while the Lean approach adds the highest value to the customer. A Combination of both disciplines results in an optimal trajectory of higher customer value and lower cost for the organization.

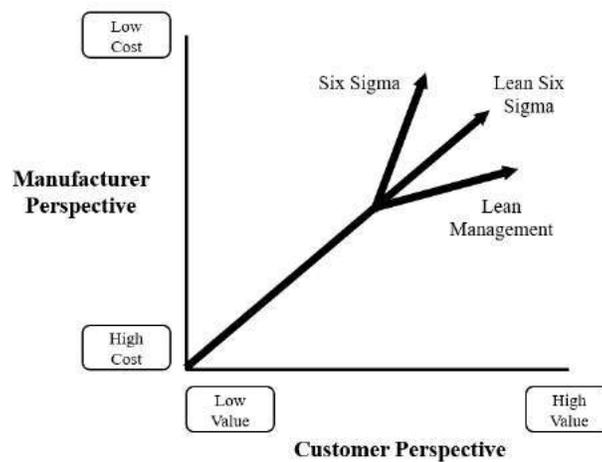


Figure 2. The Advantage of Lean Six Sigma (Arnheiter and Maleyeff, 2005)

Sreedharan and Raju’s (2016) literature review of LSS in multiple industries annotated gaps in deployment methodologies and how to apply tools within DMAIC. A systematic review of the manufacturing industry found a lack of implementation guidelines and not understanding how to use the tools within the top five limitations for LSS (Albliwi et al., 2015). Raval and Kant’s (2017) exhaustive study on 58 LSS frameworks observed numerous inconsistencies and concluded that only 1 framework was comprehensive. The authors discussed academic “conceptual” frameworks that lack practicality and urged researchers to utilize corporate practitioner input. Singh and Rathi’s

(2019) review of LSS implementation found the manufacturing industry still needing further research despite the overall growth of the philosophy within the sector.

Laureani and Antony (2012) highlight the evolution of quality management knowledge and tools occurring independently from the business realm. The parallel yet delayed application between academia and organizations support Raval and Kant's (2017) and Singh and Rathi's (2019) findings. Nonetheless, organizations can identify and align with critical success factors to influence successful implementation. Rungasamy et al. (2002) state "Critical Success Factors are those which are essential to the success of any program or technique, in the sense that, if objectives associated with the factors are not achieved, the application of the technique will perhaps fail catastrophically" (p. 218). Snee (2010) articulates that for LSS to be successful the organization must have the following eight items:

1. Financial results
2. Involved top management leadership
3. DMAIC methodology
4. Project completions within six months
5. Defined goals and objectives
6. Certified practitioners
7. Voice of customer and variation reduction
8. Statistical analysis

A case study of 40 large manufacturing organizations that implemented LSS experienced positive financial results, satisfied customers, and multiple types of reductions within the manufacturing processes (Antony et al., 2017). The current LSS literature identifies multiple key critical success factors including top management commitment, project selection, and training (Abu Bakar et al., 2015; Frinsdorf et al., 2014; Manville et al., 2012; Muraliraj et al., 2018; Näslund, 2013; Raja Sreedharan et al., 2018; Setijono et al., 2012; Walter et al., 2019). Albliwi et al. (2014) literature review of critical failure factors posited lack of management support, lack of training, and poor project selection as the primary three causes of unsuccessful LSS deployment. The success factors identified by these authors all share common themes of management, project selection, and training

5. Leadership Styles and Lean Six Sigma

Laureani and Antony's (2019) review of leadership and LSS found a symbiotic relationship supporting continuous improvement and overall success. Leadership's role is to guide cultural transformation through vision, influence, and measurable results (Suresh et al., 2012). Albliwi et al. (2014) review of critical LSS failures identified insufficient vision and lack of supportive leadership as contributing factors. McLean and Antony (2014) proposed a current state assessment of motivations, organizational culture, and management leadership to remedy failures associated with continuous improvement efforts in manufacturing. These findings suggest LSS has a strong dependency on leadership involvement to mitigate failures that are often observed from using the methodology. Figure 3 summarizes the relationship between leadership and LSS.

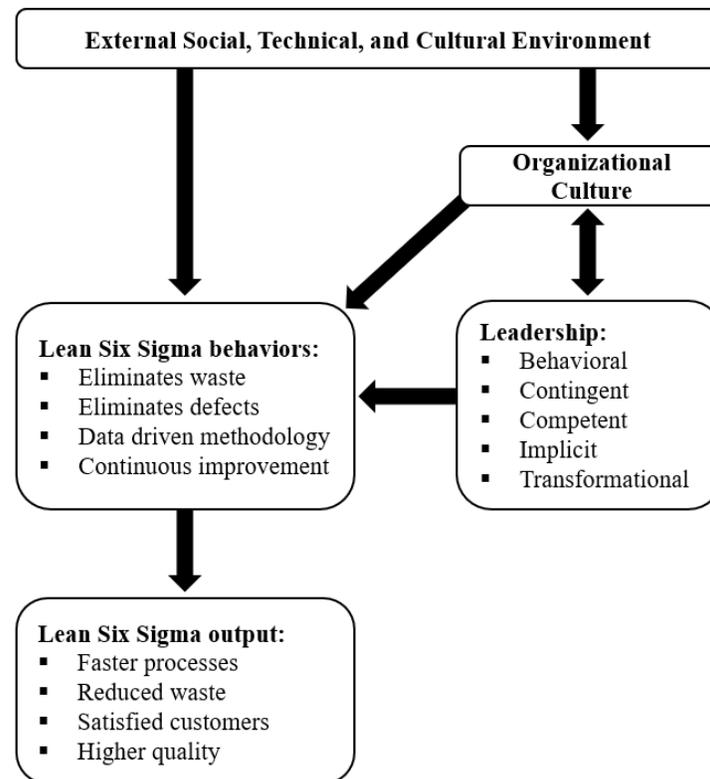


Figure 3. Model of Leadership, Culture, and Lean Six Sigma (Laureani and Antony, 2018)

Organizational transformation for continuous improvement includes reduction of bureaucratic layers, openness to creative risk, and leadership commitment through the ‘Do what you say and say what you do’ motto (Pyzdek & Keller, 2003). Cultural change requires a motivated workforce sharing goals and values that are a direct product of focused leadership commitment (Pamfilie et al., 2012). Knapp (2015) explains the active role of leadership in teaching and mentoring the culture to mitigate resistance barriers during implementation. Top-down management commitment must be matched with bottom-up leadership along all levels of the organizational hierarchy (Antony & Gupta, 2019). Manville et al. (2012) case study discussed empowering middle management with strategic leverage in choosing projects to maximize operational return. Leadership must occur through top-management strategy and through middle management project execution (Antony et al., 2018).

Laureani and Antony (2019) review of emerging themes since 2000 observed “new” leadership styles that lack uniqueness and share commonality with established leadership styles. Setijono et al. (2012) empirical study identified the growing importance of leadership styles to practitioners during implementation that was not as pronounced in the literature. A study performing qualitative analysis of effective leadership traits for Lean Six Sigma highlighted visibility, communication, consistency, and the three C’s (connection, competence, character) as critical traits (Laureani & Antony, 2017). Alexander et al. (2019) identify that the main challenge of implementing Lean Six Sigma is a lack of strong leadership at every level. In relation to continuous improvement efforts within manufacturing, transaction and transformational leadership were found to have positive impacts on quality management practices (Laohavichien et al., 2009). These findings recognize the importance that leadership plays in LSS implementation.

Transactional leadership is most effective in chain-of-command organizations with established business practices while transformational leadership seeks to disrupt those environments through

innovation and synchronization of tasks and relationships (Halaychik, 2016). Kassotaki's (2019) research of A&D organizations found that transactional leadership was predominantly used by management due to environmental constraints. One example is the compliance and export regulations imposed by the U.S. government due to the nature of manufacturing military weapons for domestic and international customers (Nielsen, 2005). This may explain why the A&D industry may lean more towards favoring transactional styles of leadership.

Within an innovation context, Oke et al. (2009) drew distinctions that transactional leadership is suited for implementation while transformational leadership is suited for cultivating post-implementation activities. Both leadership styles embrace innovation but transactional drives results at any cost while transformational focuses on empowering the culture (Chen & Zhang, 2011). Knapp (2015) found that transformational leadership coupled with innovative developmental cultures resulted in successful LSS implementation. To remain competitive, the A&D industry will face challenges when transitioning classical top-down transactional structures into more open transformational structures.

6. Conclusions

Though LSS has provided impressive gains for key organizations, there is a portion of organizations that have not been able to reap any benefits. It is often hard to isolate the exact causes that may have contributed to organizations receiving no value from the LSS methodology. The literature identifies multiple critical failure factors contributing to implementation and sustainment efforts. Antony and Gupta (2019) summarized the following regarding LSS process improvement project failures:

The top ten reasons in our opinion include lack of commitment and support from top management; poor communication practices; incompetent team; inadequate training and learning; faulty selection of process improvement methodology and its associated tools/techniques; inappropriate rewards and recognition system/culture; scope creepiness; sub-optimal team size and composition; inconsistent monitoring and control; and resistance to change (p. 367).

It is not surprising that top management commitment is first on the list of top ten reasons for project failures. Leadership must begin at the top and allow itself to flow down throughout the organization. The other nine reasons for failures all include aspects in which leadership would have significant influence. Thompson's (2005) study of a military organization seeking continuous improvements summarized that combining leadership and LSS provided a high probability for maximum benefit. Reed (2020) studied sixteen aerospace manufacturing business leaders to analyze the criteria they used to make LSS projects successful. Four common themes emerged from the study results which were planning, objectives, training, and collaboration. It can therefore be observed that A&D corporations should ensure leadership is an integral part of their LSS initiatives.

A recurring theme identified in the literature is that LSS is a powerful methodology subject to human alignment and integration. The three critical success factors of management commitment, project selection, and training correlate more with human-based interactions than the structure of the methodology itself. Laureani and Antony (2018) expressed this relationship through Figure 3, which depicts the interrelatedness of leadership, culture, and LSS. An A&D corporation pursuing LSS independent of leadership would struggle to realize maximum benefits derived from the methodology. These literature review findings conclude that utilizing LSS short of organizational investment in the right leaders will hinder successful LSS implementation and sustainment within the A&D industry.

7. References

- Abu Bakar, F. A., Subari, K., & Mohd Daril, M. A. (2015). Critical success factors of Lean Six Sigma deployment: a current review. *International Journal of Lean Six Sigma*, 6(4), 339-348.
- Albliwi, S., Antony, J., Abdul Halim Lim, S., & van der Wiele, T. (2014). Critical failure factors of Lean Six Sigma: a systematic literature review. *International Journal of Quality & Reliability Management*, 31(9), 1012-1030.
- Albliwi, S. A., Antony, J., & Lim, S. A. H. (2015). A systematic review of Lean Six Sigma for the manufacturing industry. *Business Process Management Journal*, 21(3), 665-690. doi:10.1108/BPMJ-03-2014-0019
- Alexander, P., Antony, J., & Rodgers, B. (2019). Lean Six Sigma for small-and medium-sized manufacturing enterprises: a systematic review. *International Journal of Quality & Reliability Management*, 36(3), 378-397.
- Antony, J. (2004). Some pros and cons of Six Sigma: An academic perspective. *TQM Magazine*, 16(4), 303-306. doi:10.1108/09544780410541945
- Antony, J. (2008). Can Six Sigma be effectively implemented in SMEs? *International Journal of Productivity and Performance Management*, 57(5), 420-423. doi:10.1108/17410400810881863
- Antony, J., & Gupta, S. (2019). Top ten reasons for process improvement project failures. *International Journal of Lean Six Sigma*, 10(1), 367-374.
- Antony, J., Gupta, S., Sunder M, V., & Gijo, E. V. (2018). Ten commandments of Lean Six Sigma: a practitioners' perspective. *International Journal of Productivity and Performance Management*, 67(6), 1033-1044.
- Antony, J., Snee, R., & Hoerl, R. (2017). Lean Six Sigma: yesterday, today, and tomorrow. *International Journal of Quality & Reliability Management*, 34(7), 1073-1093.
- Arnheiter, E. D., & Maleyeff, J. (2005). The integration of Lean management and Six Sigma. *The TQM Magazine*, 17(1), 5-18.
- Avolio, B. J. B. M. (2004). Multifactor leadership questionnaire: Manual & review copy. In. California: Mind Garden.
- Bass, B. M., & Avolio, B. J. (1993). Transformational leadership and organizational culture. *Public administration quarterly*, 112-121.
- Bass, B. M., & Avolio, B. J. (1994). Improving organizational effectiveness through transformational leadership: Sage.
- Brown, A., Eatock, J., Dixon, D., Meenan, B. J., & Anderson, J. (2008). Quality and continuous improvement in medical device manufacturing. *TQM Journal*, 20(6), 541-555. doi:10.1108/17542730810909329
- Den Hartog, D. N., Van Muijen, J. J., & Koopman, P. L. (1997). Transactional versus transformational leadership: An analysis of the MLQ. *Journal of Occupational and Organizational Psychology*, 70(1), 19-34.

- Ebrahimi, P., Moosavi, S. M., & Chirani, E. (2016). Relationship between Leadership Styles and Organizational Performance by Considering Innovation in Manufacturing Companies of Guilan Province. In (Vol. 230, pp. 351-358).
- Emmerling, R., Canboy, B., Serlavos, R., & Batista-Foguet, J. (2015). Leadership Education: Theory and Practice. In J. D. Wright (Ed.), *International Encyclopedia of the Social & Behavioral Sciences* (Second Edition) (pp. 655-663). Oxford: Elsevier.
- Frinsdorf, O., Zuo, J., & Xia, B. (2014). Critical factors for project efficiency in a defence environment. *International Journal of Project Management*, 32(5), 803-814.
doi://doi.org/10.1016/j.ijproman.2013.10.008
- Guo-yi, C., & Jian-sheng, Z. (2011, 2011). Analysis on transformational leadership in post-crisis Era: Evidence from Jack Welch in GE. Paper presented at the 2011 International Conference on Electronics, Communications, and Control, Piscataway, NJ, USA.
- Halaychik, C. S. (2016). Chapter 2 - Leadership Styles. In C. S. Halaychik (Ed.), (pp. 57-96): Chandos Publishing.
- Hull, C. W. (1990). Engineering in a global economy: A United States perspective. *Technology in Society*, 12(2), 107-120. doi://doi.org/10.1016/0160-791X(90)90003-U
- Jonsdottir, S., Ingason, H. T., & Jonasson, H. I. (2014). Continuous Improvement Projects in Certified Organizations in Iceland: Traditional Projects or not? *Procedia - Social and Behavioral Sciences; Selected papers from the 27th IPMA (International Project Management Association), World Congress, Dubrovnik, Croatia, 2013*, 119, 142-151. doi://doi.org/10.1016/j.sbspro.2014.03.018
- Kassotaki, O. (2019). Explaining ambidextrous leadership in the aerospace and defense organizations. *European Management Journal*, 37(5), 552-563. doi:https://doi.org/10.1016/j.emj.2019.04.001
- Khaled, M. M. (2013). Analysis of Six Sigma in the aerospace industry. *International Journal of Mechanical and Mechatronics Engineering*, 7(12), 3072-3075.
- Knapp, S. (2015). Lean Six Sigma implementation and organizational culture. *International journal of health care quality assurance*, 28(8), 855-863.
- Laohavichien, T., Fredendall, L. D., & Cantrell, R. S. (2009). The effects of transformational and transactional leadership on quality improvement. *Quality Management Journal*, 16(2), 7-24.
- Laureani, A., & Antony, J. (2012). Standards for Lean Six Sigma certification. *International Journal of Productivity and Performance Management*.
- Laureani, A., & Antony, J. (2017). Leadership characteristics for Lean Six Sigma. *Total Quality Management & Business Excellence*, 28(3-4), 405-426.
- Laureani, A., & Antony, J. (2019). Leadership and Lean Six Sigma: a systematic literature review. *Total Quality Management & Business Excellence*, 30(1-2), 53-81.
- Manville, G., Greatbanks, R., Krishnasamy, R., & Parker, D. W. (2012). Critical success factors for Lean Six Sigma programmes: a view from middle management. *International Journal of Quality & Reliability Management*, 29(1), 7-20.
- McCleskey, J. A. (2014). Situational, transformational, and transactional leadership and leadership development. *Journal of Business Studies Quarterly*, 5(4), 117.

- McLean, R., & Antony, J. (2014). Why continuous improvement initiatives fail in manufacturing environments? A systematic review of the evidence. *International Journal of Productivity & Performance Management*, 63(3), 370-376. doi:10.1108/IJPPM-07-2013-0124
- Muraliraj, J., Zailani, S., Kuppusamy, S., & Santha, C. (2018). Annotated methodological review of Lean Six Sigma. *International Journal of Lean Six Sigma*, 9(1), 2-49.
- Näslund, D. (2013). Lean and Six Sigma—critical success factors revisited. *International Journal of Quality and Service Sciences*, 5(1), 86-100.
- Nielsen, C. (2005). Competition within the U.S. National Security Regime: A study of the U.S. Aerospace Defense Sector. *Journal of International Management; Global Security Risks and International Competitiveness*, 11(4), 497-517. doi://doi.org/10.1016/j.intman.2005.09.007
- Northouse, P. G. (2018). *Leadership: Theory and practice*: Sage publications.
- Oke, A., Munshi, N., & Walumbwa, F. O. (2009). The Influence of Leadership on Innovation Processes and Activities. In (Vol. 38, pp. 64-72).
- Pamfilie, R., Andreea, J. P., & Draghici, M. (2012). The Importance of Leadership in Driving a Strategic Lean Six Sigma Management. *Procedia - Social and Behavioral Sciences; 8th International Strategic Management Conference*, 58, 187-196. doi://doi.org/10.1016/j.sbspro.2012.09.992
- Papin, M. W., & Kleiner, B. H. (1998). Effective strategic management in the aerospace industry. *Aircraft Engineering and Aerospace Technology*, 70(1), 38-44.
- Prasad, B., & Junni, P. (2016). CEO transformational and transactional leadership and organizational innovation: The moderating role of environmental dynamism. *Management Decision*, 54(7), 1542-1568. doi:10.1108/MD-11-2014-0651
- Pyzdek, T., & Keller, P. A. (2003). *A complete guide for green belts, black belts, and managers at all levels*.
- Raja Sreedharan, V., Vijaya Sunder, M., & Raju, R. (2018). Critical success factors of TQM, Six Sigma, Lean and Lean Six Sigma: A literature review and key findings. *Benchmarking: An International Journal*, 25(9), 3479-3504. doi:10.1108/BIJ-08-2017-0223
- Raval, S. J., & Kant, R. (2017). Study on Lean Six Sigma frameworks: a critical literature review. *International Journal of Lean Six Sigma*, 8(3), 275-334.
- Reed, V. (2020). *Strategies for Improving the Success Rates of Lean Six Sigma Projects*. Walden University,
- Rungasamy, S., Antony, J., & Ghosh, S. (2002). Critical success factors for SPC implementation in UK small and medium enterprises: some key findings from a survey. *The TQM Magazine*, 14(4), 217-224.
- Schroeder, R. G., Linderman, K., Liedtke, C., & Choo, A. S. (2008). Six Sigma: Definition and underlying theory. In (Vol. 26, pp. 536-554).
- Setijono, D., Laureani, A., & Antony, J. (2012). Critical success factors for the effective implementation of Lean Six Sigma. *International Journal of Lean Six Sigma*.
- Shafer, S. M., & Moeller, S. B. (2012). The effects of Six Sigma on corporate performance: An empirical investigation. In (Vol. 30, pp. 521-532).
- Singh, M., & Rathi, R. (2019). A structured review of Lean Six Sigma in various industrial sectors. *International Journal of Lean Six Sigma*, 10(2), 622-664.

- Snee, R. D. (2010). Lean Six Sigma - getting better all the time. *International Journal of Lean Six Sigma*, 1(1), 9-29. doi://dx.doi.org/10.1108/20401461011033130
- Sousa, M. J., & Rocha, Á. (2019). Digital learning: Developing skills for digital transformation of organizations. In (Vol. 91, pp. 327-334).
- Spedding, T. A., & Pepper, M. P. J. (2010). The evolution of Lean Six Sigma. *Int J Qual & Reliability Mgmt*, 27(2), 138-155. doi:10.1108/02656711011014276; 1410.1108/02656711011014276
- Sreedharan, V. R., & Raju, R. (2016). A systematic literature review of Lean Six Sigma in different industries. *International Journal of Lean Six Sigma*, 7(4), 430-466.
- Stankalla, R., Koval, O., & Chromjakova, F. (2018). A review of critical success factors for the successful implementation of Lean Six Sigma and Six Sigma in manufacturing small and medium-sized enterprises. 1-16. doi:10.1080/08982112.2018.1448933
- Steinbock, D. (2014). The challenges for America's defense innovation. *The Information Technology & Innovation Foundation*, 20-21.
- Strang, K. D. (2005). Examining effective and ineffective transformational project leadership. *Team Performance Management*, 11(3), 68-103. doi:10.1108/13527590510606299; 25 10.1108/13527590510606299
- Suresh, S., Antony, J., Kumar, M., & Douglas, A. (2012). Six Sigma and leadership: some observations and agenda for future research. *The TQM Journal*, 24(3), 231-247. doi:10.1108/17542731211226754
- Swain, A. K., Cao, Q. R., & Gardner, W. L. (2018). Six Sigma success: Looking through authentic leadership and behavioral integrity theoretical lenses. In (Vol. 5, pp. 120-132).
- Testani, M. V., & Ramakrishnan, S. (2011, 2011). Lean transformation leadership model: Leadership's role in creating Lean culture. Paper presented at the 61st Annual Conference and Expo of the Institute of Industrial Engineers, May 21, 2011 - May 25, Reno, NV, United States.
- Thompson, J. A. (2005). Gaining greater benefit from Lean Six Sigma and leadership initiatives within the military. Paper presented at the IEMC 2005: 2005 IEEE International Engineering Management Conference, September 11, 2005 - September 13, St. John's, Newfoundland, Canada.
- Tyssen, A. K., Wald, A., & Spieth, P. (2014). The challenge of transactional and transformational leadership in projects. *International Journal of Project Management*, 32(3), 365-375. doi://doi.org/10.1016/j.ijproman.2013.05.010
- Waldman, D. A. (1993). A theoretical consideration of leadership and total quality management. *The Leadership Quarterly*, 4(1), 65-79.
- Walter, O. M. F. C., & Paladini, E. P. (2019). Lean Six Sigma in Brazil: a literature review. *International Journal of Lean Six Sigma*, 10(1), 435-472.
- Wang, C.-N., Nguyen, X.-T., Le, T.-D., & Hsueh, M.-H. (2018). A partner selection approach for strategic alliance in the global aerospace and defense industry. *Journal of Air Transport Management*, 69, 190-204. doi://doi.org/10.1016/j.jairtraman.2018.03.003
- Xie, Y., Xue, W., Li, L., Wang, A., Chen, Y., Zheng, Q., . . . Li, X. (2018). Leadership style and innovation atmosphere in enterprises: An empirical study. In (Vol. 135, pp. 257-265).

Yukl, G. (2011). Contingency theories of effective leadership. *The SAGE handbook of leadership*, 24(1), 286-298.

Zhang, Q., Irfan, M., Khattak, M. A. O., Zhu, X., & Hassan, M. (2012). Lean Six Sigma: a literature review. *Interdisciplinary Journal of Contemporary research in business*, 3(10), 599-605.

A study on Relocation of Home Field and Improvement of Game Schedules for Nippon Professional Baseball

Yutaka Shirai¹

Hiroyuki Ono¹

¹*Chiba Institute of Technology, Japan*

yutaka.shirai@it-chiba.ac.jp

Abstract

Professional sports are regarded as a business in Europe and the U.S. lately. Sport scheduling problems have been studied for increased attendance and improvement in game schedules. By contrast, Nippon Professional Baseball (NPB) regards ball games as a performance. Therefore, the improvement of game schedules has not been considered seriously. However, it has been revealed that sufficient attendance for sustainment of baseball league management is not securable only at the present home fields. For that reason, it is fundamentally important to acquire spectators in various regions other than the present home fields for increased attendance. In addition to such management issues, trading of a ball club might also require the relocation of a team's home field. Given that background, it is important to design a game schedule that lessens variation in the travel distance of each team. This study analyzes factors that affect attendance by multiple regression analysis and investigates ball club relocation based on anticipated attendance. Furthermore, we determine a game schedule that minimizes the total travel distance using Tabu Search and propose a method of determining candidate destination cities for the ball clubs to be relocated. Finally, we verify the effectiveness of the proposed method using a numerical experiment.

1. Introduction

Professional sports are regarded as a business in Europe and the U.S. lately. Sporting event scheduling problems have been studied for increased attendance and improvement in game schedules [Ball & Webster, 1977; Bartsch, Drexl, & Kroger, 2006; Cruz-Chavez, Flores-Pichardo, Martinez & Cruz- Rosales, 2016; Easton, Nemhauser & Trick, 2001; Nemhauser & Trick, 1998; Ribeiro, 2012; Ribeiro & Urrutia, 2004]. The improvement of game schedules has even reduced the total travel distance by one-third in some professional sports leagues. Nippon Professional Baseball (NPB) regards ball games as performances, improvement in game schedules has not been considered severely. When the number of spectators was changed from an estimated value to a real value in 2005, it became clear that the number of spectators was not as large as the conventional assumption, so the management risk and the separation from baseball became obvious.

For that reason, it is fundamentally important not to compete for spectators in the same area but to acquire spectators in various regions other than the present home fields for increased attendance. In addition to such management issues, trading of a ball club might also require the relocation of a home field. Given such a background, it has become important to design a game schedule that reduces variation in the travel distance of each team.

We use multiple regression analysis in this study to assess factors that affect attendance for the improvement of game schedules of the present scheme of the NPB. We investigate ball clubs to be relocated based on the anticipated attendance. Subsequently, we investigate a game schedule that minimizes the total travel distance by solving a Travelling Tournament Problem (TTP) [Easton, Nemhauser & Trick, 2001], an NP-hard problem, using Tabu Search [Munakata, 2008]. Furthermore, based on these two experimentally obtained results (ball clubs to be relocated and the game schedule computed), we ascertain the most feasible candidate city for a ball club to be relocated. Finally, we verify the effectiveness of the proposed method by numerical experimentation using data for 2015–2018.

2. Sports Scheduling Problem

2.1 Outline of Sports Scheduling Problem

A sports scheduling problem determines the optimal game schedule for improvement in cost aspects such as travel, improvement in popularity, and increased revenue. The oldest literature related to sport scheduling problems [Ball & Webster, 1977] is made in the 1970s. Various studies have been conducted thereafter [Miyashiro, Imahori & Matsui, 2012]. They include various cases involving large-scale sporting event organizations, for example, the Brazilian professional soccer league [Ribeiro & Urrutia, 2004], the college basketball tournament of the Atlantic Coast Conference in the U.S. [Nemhauser & Trick, 2001], and the professional soccer leagues of Austria and Germany [Bartsch, Drexl, & Kroger, 2006].

Analyses of the factors affecting attendance include examinations of the attendance boost effect in the interleague play of the Major League Baseball (MLB) [Bruggink & Roosma, 2003] and the influence of TV live broadcasting of the National Football League (Premier League) on attendance [Carmichael, Millington & Simmons, 1999].

2.2 Travelling Tournament Problem (TTP)

Typical sport scheduling problems include a Travelling Tournament Problem (TTP) [Easton, Nemhauser & Trick, 2001]. A TTP discussed herein establishes a game schedule that minimizes the total moving distance of all teams in a double round-robin league competition under the following conditions, where n is an even number of teams:

- Each team has a different home field.
- One team meets all others once at home and once away.
- Each team plays a game once in each slot (day); a league competition (season) consists of $2(n-1)$ slots.
- Each game is played at the home of either of the two playing teams.

A TTP includes elements of two problems: a Home–Away Table (HAT) feasibility problem and a travel distance minimization problem. A HAT feasibility problem prepares an executable schedule, whereas a travel distance minimization problem investigates a game schedule that minimizes the total travel distance.

(1) Home–Away Table (HAT) Feasibility Problem

A problem is assumed in which only the dates of home games and away games are fixed in advance; opponent teams are assigned afterward. A game schedule table in which every game is played either at home or away is designated as a Home–Away Table (HAT). It is not that a feasible game schedule can always be generated from any given HAT, but there might be cases in which any combination of opponent teams generates only an infeasible game schedule. Requirements for a HAT to generate a feasible game schedule include the following two points:

- The numbers of home and away games are the same for each slot.
 - Each team obeys the game schedule of a different pattern of home and away games.
- Nevertheless, there might be a HAT that cannot generate a feasible game schedule even if it satisfies these requirements. A HAT feasibility problem judges whether a HAT can be generated as a feasible game schedule.

(2) Travel distance minimization problem

A travel distance minimization problem aims at finding a game schedule that minimizes the total travel distance of all teams participating in a round-robin league. Therefore, the reduction effect of transportation costs by shortening travel distance is considerable. Accordingly, a TTP is divided into two problems in this study, a HAT feasibility problem, and a travel distance minimization problem.

2.3 Formulation of TTP

The formulation of a problem to ascertain a game schedule that minimizes the total travel distance with a fixed HAT is described below.

$$\min. \sum_{i,j \in T} \sum_{s \in S} d_{ij} x_{ijs}^t \quad (1)$$

subject to

$$\sum_{t \in AT(s)} y_{is}^t = 1 \quad \forall i \in HT(s), \forall s \in S \quad (2)$$

$$\sum_{s \in HD(t)} y_{is}^t = 1 \quad \forall \tau \in T \setminus \{t\}, \forall t \in T \quad (3)$$

$$\sum_{s \in AD(t)} y_{is}^t = 1 \quad \forall \tau \in T \setminus \{t\}, \forall t \in T \quad (4)$$

$$\sum_{j \in T} x_{ijs}^t = y_{is}^t \quad (5)$$

$$\sum_{i \in T} x_{ijs}^t = y_{j,s+1}^t \quad (6)$$

$$y_{is}^t + y_{i,s+1}^t \leq 1 \quad \forall s \in AD(t), \forall i, t \in T (i \neq t) \quad (7)$$

$$y_{is}^t = \begin{cases} 0 & i \neq t \\ 1 & i = t \end{cases} \quad \forall s \in HD(t), \forall t \in T \quad (8)$$

$$y_{is}^t = 0 \quad i = t, \forall s \in AD(t), \forall t \in T \quad (9)$$

$$\sum_{i \in HT(s)} y_{is}^t = 1 \quad \forall s \in AD(t), \forall t \in T \quad (10)$$

In those equations, n represents the number of teams (even number),

T denotes a set of teams $T = \{1, 2, \dots, n\}$,

S expresses a set of slots $S = \{1, 2, \dots, (2n-1)\}$,

d_{ij} stands for the distance between stadiums i and j ,

$HD(t)$ signifies the set of slots of home games of team t ,

$AD(t)$ signifies set of slots of away games of team t ,

$HT(s)$ signifies set of teams that play a home game in slot s ,

$AT(s)$ signifies set of teams that play an away game in slot s ,

x_{ijs}^t is a decision variable that takes 1 when team t is at stadium i in slot s and moves to stadium j in the next slot $s+1$, and 0 otherwise,

y_{is}^t is a decision variable that takes 1 when team t is at stadium i in slot s , and 0 otherwise, and

τ represents the opponent team of team t in slot s .

Equation (1) is an objective function that minimizes the sum of the travel distance of all the teams. Equation (2) expresses that a team that plays an away game in slot s is present in the home stadiums of an opposing team. Equations (3) and (4) represent that team t plays one home game and one away game against every other team. Equations (5) and (6) show the travel of team t (including staying at home). Equation (7) prohibits the same opposing team to play games successively. Equation (8) indicates that team t is in the home stadium i in slot s . Finally, Equations (9) and (10) show that, in slot s where team t plays an away game, team t is present at either of the homes of teams present at home in the same slot s , but not at the home of team t .

3. Proposed Method

3.1 Procedures of Proposed Method

This study was conducted using the following procedures.

[Step 1] Factors affecting attendance are analyzed by the multiple regression analysis. Ball clubs to be relocated are investigated.

[Step 2] A game schedule that minimizes a total travel distance is found using a Tabu Search based on the result of [Step 1]. Candidate cities for the ball clubs to be relocated are investigated.

3.2 Investigation on Factors Affecting Attendance and Ball Clubs to Be Relocated

First, factors affecting attendance are analyzed using multiple regression analysis. Data of the following items are used for response variables and explanatory variables.

– Response variables

Each ball club's attendance is used as a response variable.

– Explanatory variables

The following data of seven kinds are adopted as explanatory variables for analysis. The groups of data are quoted from the websites of the following organizations: The Statistics Bureau of the Ministry of Public Management, Home Affairs, Posts and Telecommunications of Japan [Statistic Bureau, 2019], Nippon Professional Baseball Organization [NPB, 2019], the Japan Meteorological Agency of the Ministry of Land, Infrastructure, and Transport of Japan [JMA, 2019], Google Maps [Google Maps, 2019], and East Japan Railway [JR East, 2019]

(1) Travel time

Travel time from the largest city near each ball club to its home stadium is used. Longer traveling time is regarded as depressing spectators' interest in visiting the stadium to watch a game. Therefore, travel time is presumed to exert a negative influence on attendance.

(2) Revenue

The financial capability index of a region of each ball club's hometown is used. A region with a high financial capability index has a large fiscal space in which many people can afford amusement. Consequently, it is presumed to make a positive influence on attendance because people have more chances to observe professional sports.

(3) Stadium capacity

The maximum accommodation of each ball club's home field (stadium) is used. It is surmised to make a positive influence on attendance because a greater capacity accommodates more spectators per game.

(4) Regional population

The population of the region of each ball club's hometown is used. It is presumed to make a positive influence on attendance because a larger population might yield more opportunities for people to visit the stadium.

(5) Population density

The Population density of the region of each ball club's hometown is used. It is presumed to make a positive influence on attendance because a larger Population density might yield more opportunities for people to visit the stadium.

(6) Number of days with rain

The number of days with rain of the game days at the home field of each ball club is used. It is expected to exert a negative influence on attendance because, in the case of rainy weather, it is thought that audience will be few. In the case of the dome stadium, it shall be 0 days.

(7) Number of station users

The number of users of the near station with each ball club's home field is used. It is regarded as exerting a positive influence on attendance because many attendances might gather at a station with many users.

The expected attendance to each home field is found based on the obtained affecting factors by a multiple regression expression. Ball clubs with lower attendance compared with the current attendance are selected as candidates for relocation.

3.3 Generation of Best Game Schedule by TTP

The best game schedule as calculated using TTP is generated using the following procedures.

[Step 1] A HAT is prepared.

[Step 2] Opponent teams are assigned.

[Step 3] A total travel distance is computed.

[Step 4] A game schedule (approximate solution) is determined using a Tabu Search based on [Step 1] – [Step 3].

(1) Preparation of HAT [Step 1]

A fundamental HAT is prepared first for solving a TTP. Requirements for preparing a HAT are the following:

- The numbers of home games and away games of each team are the same.
- The numbers of teams at home and away in slots are the same.
- More than three consecutive home games (or away games) might not be scheduled.

For example, when the number of teams is set to 4, the number of slots (game days) becomes 6 ($2 \times (4 \text{ teams} - 1) = 6$). Numbers from 1 to 24 [$4 \text{ teams} \times 6 \text{ slots}$] are stored in array α at random (see Figure 1). A HAT is prepared using this array α . Home games (h) are allocated to the right-hand side table of Figure 1 at the location designated by the number of the left-hand side table corresponding to the element of array α , one by one from the head of α so that the requirements described above might be satisfied. The allocation is skipped when the requirements are not satisfied. The preparation of a HAT is completed when the tail of array α is reached. Blanks left on the table at this time express away games. When the HAT generated does not satisfy the requirements, array α is re-generated at random. Then preparation of a HAT is attempted again.

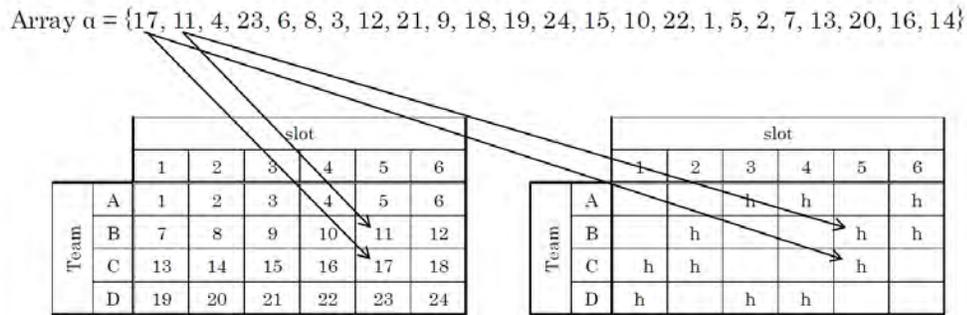


Figure 1. Example of HAT preparation.

(2) Assignment of opponent teams [Step 2]

Opponent teams are assigned after a HAT is completed so that the following requirements are satisfied.

- Any two teams meet once at home and once away.
- A team playing a home game meets another team playing an away game.

When multiple assignable teams exist, one team is selected at random and assigned.

(3) Computation of total travel distance [Step 3]

If opponent teams are fixed, then a total travel distance is calculated based on distance data between the playing teams.

(4) The search of game schedule (approximate solution) [Step 4]

When a total travel distance is calculated, the numbers allocated to the home (h) of the HAT are stored in a tabu list. Then a new HAT is generated by swap neighborhood search. Then a game schedule corresponding to the HAT (approximate solution) is obtained. A game schedule that minimizes a total travel distance is determined by repeating these processes.

Swap neighborhood search is a method to obtain an approximate solution with the best objective function by swapping elements.

4. Numerical Experiment

4.1 Outline of Numerical Experiment

The NPB consists of a total of 12 ball clubs: 6 ball clubs in the Pacific League and 6 ball clubs in the Central League. Table 1 shows the team name, the region of the hometown of the 12 ball clubs of the NPB, and candidate cities for relocation. Figures 2 and 3 respectively portray the locations of hometowns of ball clubs of the Pacific League and the Central League. Figure 4 depicts the locations of candidate cities for hometown relocation.

The candidate cities for relocation of hometowns are selected among regions (stadiums) where a regular-season game was actually held in the past, and where there is no franchise of the 12 ball clubs of the NPB.

Table 1. Club name, team name, region of hometown, and candidate cities for relocation

	Club Name	Team Name	Hometown (Region)	Candidate City
Pacific League	Hokkaido Nippon Ham Fighters	Fighters	Sapporo (Hokkaido)	Shizuoka
	Tohoku Rakuten Golden Eagles	Eagles	Sendai (Miyagi)	Nagano
	Saitama Seibu Lions	Lions	Tokorozawa (Saitama)	Niigata
	Chiba Lotte Marines	Marines	Chiba	Toyama
	ORIX Buffaloes	Buffaloes	Osaka	Matsuyama (Ehime)
	Fukuoka SoftBank Hawks	Hawks	Fukuoka	Miyazaki
Central League	Yomiuri Giants	Giants	Tokyo	Kumamoto
	Tokyo Yakult Swallows	Swallows	Tokyo	Kagoshima
	Yokohama DeNA Baystars	BayStars	Yokohama (Kanagawa)	Naha (Okinawa)
	Chunichi Dragons	Dragons	Nagoya (Aichi)	
	Hanshin Tigers	Tigers	Nishinomiya (Hyogo)	
	Hiroshima Toyo Carp	Carp	Hiroshima	

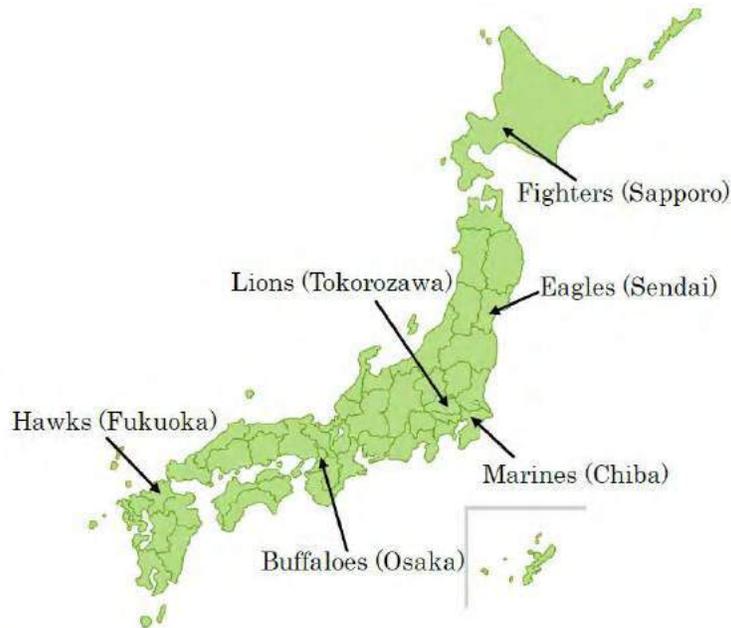


Figure 2. Current hometown locations of ball clubs of Pacific League.



Figure 3. Current hometown locations of ball clubs of Central League.

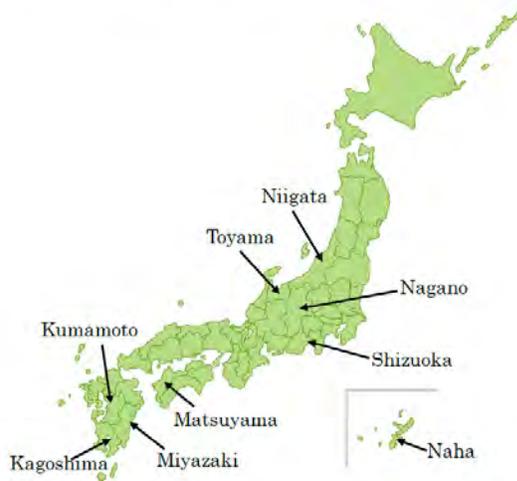


Figure 4. Candidate cities for relocation of hometown.

A numerical experiment is conducted in the procedure described in Section 3.1. Data for four years in 2015–2018 are employed. Furthermore, the popularity level of each team is not considered.

4.2 Results of Investigation on Factors Affecting Attendance and Ball Club to be Relocated

Table 2 presents the result of the multiple regression analysis using the average data for four years of 2015–2018 (See Section 3.2 for the items of response variables and explanatory variables). We also confirmed that the Independence, Homoscedasticity, and Normality of the residuals were satisfied. The analysis results reveal that the stadium capacity, Regional population, and population density affect attendance.

Multiple regression analysis suggests that a larger capacity of a stadium would encourage attendance. However, a stadium of a large capacity but with many vacant seats might engender depression in attendance. For that reason, it is preferable to find a stadium of the size suitable for the region and to extend it when the number of spectators is expected to increase by business efforts, etc. However, when extending a stadium, it is important that the ball club has ownership of the stadium.

Table 2. Result of multiple regression analysis
(Average data for 2015–2018)

	Coefficient	Standard Error	P-Value	Significance Level
Intercept	-650792.123	697365.275	0.404	
Traveling Time	-10126.054	8635.151	0.306	
Revenue Amount	-0.585	0.375	0.193	
Stadium's Maximum Capacity	102.318	19.648	0.006	**
Prefecture Population	-2672.492	1190.818	0.038	*
Population density	4036.483	1873.569	0.043	*
Rainy Days	-4702.691	13968.077	0.753	
Number of Passengers at the Station	0.440	5.038	0.935	

*p < 0.05 **p < 0.01

Table 3 presents the expected attendance meeting at the significance level of 5% at the home field of each ball club and candidate cities for relocation based on the result of Table 2.

This result suggests that Eagles and Marines of the Pacific League and Swallows and BayStars of the Central League are ball clubs of less expected attendance and are therefore candidates for relocation. However, the home field of the Eagles has such a small capacity that a remedy to enlarge the stadium capacity is regarded as effective. The home field of Eagles has a high occupancy rate so that the increase of seating anticipates much greater attendance. BayStars achieved significant increases in attendance through business efforts undertaken after the parent company changed in 2012. For these reasons, we nominate the Marines in the Pacific League and the Swallows in the Central League as candidate ball clubs for relocation.

Table 3. Expected attendance at the home field of each ball club and candidate cities for relocation

Team Name	Average Attendance for Four Years	Expected Attendance	Candidate Site	Expected Attendance
Fighters	1,923,404	2,182,130	Shizuoka	928,127
Eagles	1,237,018	1,170,432	Nagano	1,775,631
Lions	1,553,872	1,632,641	Niigata	1,862,362
Marines	1,344,605	1,124,348	Toyama	2,141,080
Buffaloes	1,433,884	1,390,450	Ehime	2,037,284
Hawks	2,338,206	2,153,996	Miyazaki	2,113,749
Giants	2,924,733	2,804,468	Kumamoto	1,339,095
Swallows	1,367,950	1,486,375	Kagoshima	1,018,113
BayStars	1,285,828	1,350,680	Okinawa	2,319,806
Dragons	2,119,187	2,147,769		
Tigers	2,850,021	2,793,943		
Carp	1,685,950	1,827,425		

Table 4 presents the expected improvement of attendance of Marines and Swallows by relocation, defined as the ratio of attendance after relocation to the present based on the data of Table 3. This result suggests that Shizuoka and Kagoshima are unsuitable as destination hometowns (home fields) of relocation because the ratio of expected attendance after relocation to the present attendance is less than 1 and because improvement is not expected. Kumamoto shows a similar trend so that Shizuoka,

Kumamoto and Kagoshima are excluded from destination cities for relocation this time.

Table 4. Expected improvement in attendance by relocation

Candidate City	Marines	Swallows
Shizuoka	0.69	0.68
Nagano	1.32	1.30
Niigata	1.39	1.36
Toyama	1.59	1.57
Matsuyama	1.52	1.49
Miyazaki	1.57	1.55
Kumamoto	1.00	0.98
Kagoshima	0.76	0.74
Naha	1.73	1.70

4.3 Result of Investigation on Game Schedule and Candidate Destination City for Relocation of Ball Clubs

A game schedule obtained using a Tabu Search is evaluated by comparison with the actually executed game schedule.

(1) Preconditions

A game schedule is determined under the following preconditions in this study.

- The actual game schedule in 2018 is adopted for comparison.
- A total of 125 games, excluding the interleague games (3 games × 6 teams = 18 games) from 143 games for one season, undergo scheduling.
- Sets of three games in a row at the same stadium are defined as a card: Tuesday–Thursday and Friday–Sunday.
- Each team travels or stays at its home field after each card.
- One card shall be one slot.
- A set of 40 cards ((3 consecutive games × 35 cards) + (4 consecutive games × 5 cards) = 125 games) are subjected to the computation of the travel distance.
- No team plays home games (or away games) for more than three consecutive cards.
- Cancellation due to rain (replay) is not considered.

(2) Comparison between the present conditions (2018) and the results obtained using proposed methods

Table 5 shows the travel distances obtained from numerical experiments. This result indicates that travel distances are shortened significantly by the relocation of the two ball clubs to cities other than Naha (Okinawa) and Miyazaki, both for the Pacific and Central Leagues. Especially, comparison of the game schedule of the present conditions (2018) and that obtained using the proposed method (with no relocation) demonstrates that the latter can shorten the total travel distance by about 30,000 km (about 20%) for the Pacific League and about 20,000 km (about 30%) for the Central League. Additionally, this result indicates a markedly improved standard deviation compared with the present game schedule (2018), so that an impartial game schedule can be obtained to some degree.

Table 5. Comparison of game schedule (travel distance) (km)
(a) Pacific League

	Fighters	Eagles	Lions	Marines	Buffaloes	Hawks	Total Travel Distance (km)	Standard Deviation
	(Sapporo)	(Sendai)	(Tokorozawa)	(Chiba)	(Osaka)	(Fukuoka)		
Current Schedule	29,290	19,980	19,624	19,281	21,781	28,937	138,843	4,297.92
Proposed Method (No Relocation)	22,998	15,783	20,412	15,455	15,290	19,580	109,518	2,933.84
Proposed Method (Marines → Nagano)	21,067	16,067	15,494	14,646 (Nagano)	16,428	19,040	102,741	2,221.88
Proposed Method (Marines → Niigata)	19,879	16,334	16,994	15,770 (Niigata)	16,630	19,067	104,674	1,497.88
Proposed Method (Marines → Toyama)	21,546	14,952	15,508	15,899 (Toyama)	16,122	21,194	105,220	2,736.48
Proposed Method (Marines → Matsuyama)	21,610	16,665	18,573	16,988 (Matsuyama)	17,065	18,725	109,624	1,689.42
Proposed Method (Marines → Miyazaki)	24,794	18,794	17,105	19,888 (Miyazaki)	19,658	20,002	120,239	2,342.94
Proposed Method (Marines → Naha)	29,695	23,558	22,123	31,317 (Naha)	22,428	26,797	155,917	3,566.44

(b) Central League

	Giants	Swallows	BayStars	Dragons	Tigers	Carp	Total Travel Distance (km)	Standard Deviation
	(Tokyo)	(Tokyo)	(Yokohama)	(Nagoya)	(Nishinomiya)	(Hiroshima)		
Current Schedule	10,837	11,335	9,900	10,340	12,132	16,249	70,793	2,113.02
Proposed Method (No Relocation)	7,107	7,844	7,914	7,056	8,046	10,533	48,499	1,161.61
Proposed Method (Swallows → Nagano)	8,133	9,149 (Nagano)	8,102	7,517	8,369	11,035	52,305	1,143.09
Proposed Method (Swallows → Niigata)	8,932	10,395 (Niigata)	8,926	9,267	8,925	11,996	58,442	1,134.45
Proposed Method (Swallows → Toyama)	8,094	8,498 (Toyama)	8,981	7,590	7,972	10,796	51,931	1,051.10
Proposed Method (Swallows → Matsuyama)	8,597	8,877 (Matsuyama)	10,725	9,293	7,733	10,018	55,241	968.15
Proposed Method (Swallows → Miyazaki)	11,418	12,675 (Miyazaki)	10,239	10,174	11,513	10,896	66,913	853.91
Proposed Method (Swallows → Naha)	16,713	25,317 (Naha)	14,342	15,651	15,880	16,238	104,140	3,633.49

As for the Pacific League, the travel distance of the Lions alone is slightly extended by generating a game schedule according to the proposed method (with no relocation). However, that of other teams is shortened, so that the variation in the travel distance between teams is regarded as reduced to some extent.

Moreover, in cases where the Marines are relocated to Nagano, Niigata, Toyama, or Matsuyama (Ehime), the total travel distance is shortened slightly compared to the proposed method (with no relocation). This result is obtained because the candidate cities for relocation are located close to a straight line between the hometowns of Fighters and Hawks. Therefore, they are effective as relocation cities that would implement efficient travel.

As for the Central League, compared with the present game schedule (2018), the travel distance of each team is shortened by not less than 2,000 km according to the proposed method (with no relocation). Moreover, the difference in travel distance between teams is reduced.

The total travel distance is not shortened wherever the Swallows are relocated compared with the proposed method (with no relocation). This occurs probably because the travel distance between the home fields of Giants and Swallows is very short (about 6 km). Although the travel distance is extended, variation in the travel distance between teams is reduced if the Swallows relocate to Matsuyama or Toyama.

(3) Results of investigation of candidate cities for relocation

Table 6 presents the improvement of the attendance and the travel distance based on the results described above. The respective improvement is given as the ratio to the corresponding datum by the present conditions (2018).

Table 6 suggests that relocation to Naha apparently engenders an undesirable result in which the total travel distance is elongated although more attendance is expected. Consequently, Naha is removed from the candidate cities for relocation this time. The overall improvement, defined as the sum of improvement in attendance and improvement in travel distance, is high when Marines of the Pacific League and Swallows of the Central League are relocated to Toyama. Accordingly, both attendance and travel distance have been improved by relocating Swallows to Toyama with the highest overall improvement and Marines to Matsuyama with the second-highest overall improvement, so that the relocation of hometowns is verified as effective.

Table 6. Improvement of attendance and travel distance for candidate cities for relocation
(a) Pacific League

Movement Candidate (Marines)			
Candidate City	Improvement in Attendance	Improvement in Travel Distance	Total
Nagano	1.32	1.35	2.67
Niigata	1.39	1.33	2.72
Toyama	1.59	1.32	2.91
Matsuyama	1.52	1.27	2.79
Miyazaki	1.57	1.15	2.72
Naha	1.73	0.89	2.62

(b) Central League

Movement Candidate (Swallows)			
Candidate City	Improvement in Attendance	Improvement in Travel Distance	Total
Nagano	1.30	1.35	2.65
Niigata	1.36	1.21	2.57
Toyama	1.57	1.36	2.93
Matsuyama	1.49	1.28	2.77
Miyazaki	1.55	1.06	2.61
Naha	1.70	0.68	2.38

5. Conclusion

We investigated factors that might encourage increased attendance for NPB games in this study. In addition, we selected ball clubs to be relocated. Then we used Tabu Search with a TTP to determine a game schedule that minimizes the total travel distance. We investigated candidate cities for relocation of the ball clubs. The obtained knowledge is the following:

- (1) Factors such as stadium capacity, prefectural population, and producing population affect the change of attendance.
- (2) Marines of the Pacific League and Swallows of the Central League are nominated as ball clubs to be relocated based on the anticipated attendance.

- (3) Comparison of the game schedule of the present conditions (2018) and that of the proposed method (with no relocation) demonstrates a reduction of the total travel distance by about 20–30%.
- (4) The expected improvement of the attendance and anticipated travel distance suggest that relocating the hometowns of Marines and Swallows to Matsuyama and Toyama, respectively, would reduce the difference in the travel distance of each team and implement an impartial game schedule to some extent.

Our future subjects include elucidation of the validity of the ball clubs to be relocated and relocation destinations considering the variation of the attendance and travel distance of each team by hometown relocation.

6. References

- Ball, B.C. & Webster, D.B. (1977). Optimal Scheduling for Even-Numbered Team Athletic Conferences, *AIIE Transactions*, 9(2), 161-169.
- Bartsch, T., Drexl, A. & Kröger, S. (2006). Scheduling the professional soccer leagues of Austria and Germany, *Computers & Operations Research*, 33(7), 1907-1937.
- Bruggink, T.H. & Roosma, C. (2003). The Attendance Boost is Overrated for Interleague Baseball, and the Big Mac Attack is a Hit on the Road: All this and more from the Within-Season Demand Model for Major League Baseball, *Journal of Business and Economic Research*, 1(3), 55-62.
- Carmichael, F., Millington, J. & Simmons, R. (1999). Elasticity of demand for Rugby League attendance and the impact of BskyB, *Journal Applied Economics Letters*, 6(12), 797-800.
- Cruz-Chávez, M. A., Flores-Pichardo, M., Martinez, A. & Cruz-Rosales, M. H. (2016). Solving a Real Constraint Satisfaction Model for the University Course Timetabling Problem: A Case Study, *Mathematical Problems in Engineering*, 3, 1-14.
- East Japan Railway Company (JR East). (2019, October). <http://www.jreast.co.jp/e/>
- Easton, K., Nemhauser, G. & Trick, M. (2001). The Traveling Tournament Problem Description and Benchmarks, *Lecture Notes in Computer Science book series*, 2239, 580-585, Springer.
- Google Maps. (2016, March). <https://www.google.com/maps/>
- Japan Meteorological Agency (JMA). (2019, October). <http://www.jma.go.jp/jma/indexe.html> Miyashiro, R.,
- Imahori, S. & Matsui, T. (2012). An approximation algorithm for the traveling tournament problem, *Annals of Operations Research*, 194(1), 317-324.
- Munakata, T. (2008). *Fundamentals of the New Artificial Intelligence: Neural, Evolutionary, Fuzzy and More*, Springer.
- Nemhauser, G.L. & Trick, M.A. (1998). Scheduling A Major College Basketball Conference, *Operations Research*, 46(1), 1-8.
- Nippon Professional Baseball (NPB) Organization in Japan. (2019, October). <http://npb.jp/eng/>
- Ribeiro, C.C. (2012). Sports scheduling: Problems and applications, *International Transactions in Operational Research*, 19(1-2), 201-226.

Ribeiro, C.C. & Urrutia, S. (2004). OR on the Ball Applications in Sports Scheduling and Management, *OR/MS Today*, 31, 50-54.

Statistic Bureau, Ministry of Internal Affairs and Communications in Japan. (2019, October).
<http://www.stat.go.jp/english/index.htm>

Cost-Effective Attendance Management System Using Cloud Computing and Face Recognition

Xiaoli Huan¹

Hong Zhou²

¹*Troy University*

²*University of Saint Joseph*

xhuan@troy.edu; hzhou@usj.edu

Abstract

Attendance systems are essential for all institutions, including business and educational organizations. Time tracking ensures that all team members correctly report their presence of participation and time spent on work. The resulting data can be transferred and imported into payroll or grading systems. However, maintaining the attendance records is a demanding task for the administrators due to its time-consuming efforts and resources. In this paper, a cost-effective attendance management system using cloud computing and face recognition technology is introduced. The described attendance system is straightforward to implement and generates accurate recognition results and timestamps saved in CSV records without requiring any bulk program installation.

1. Introduction

Attendance management is used in all organizations to record their employees/students beginning and ending work timings. Time tracking ensures that all team members correctly report their presence of participation and time spent on work. The resulting data can be transferred and imported into payroll or grading systems, which is crucial for every business and school.

In a traditional attendance system, employee/student hours were tracked manually, using punch cards or timesheets computed by hand. Administrators/teachers must pay more physical and mental attention to monitor sheets/documents, record information, and maintain accurate details. Mistakes could easily happen during this process and cause business loss and complaints of incorrect grades. Corrections to the errors were needed when there was inconsistency in data entry.

An automated attendance system can save organizations time and money. A sound automated attendance system should be simple to implement and manage. The overall cost can be saved by removing the time-consuming and tedious manual handling process involved in marking attendance and leaving entry and manually calculating attended hours. With an automatic attendance system, administrators/teachers can more accurately and quickly track employee/students' time in the

organizations and devote more time to other essential tasks.

A demo of a cost-effective attendance management system using cloud computing and face recognition technology is introduced in this paper.

2. Cloud computing

Cloud computing refers to storing the data, accessing it, running applications, and retrieving the results over the internet. It provides a great deal of versatility that everything is entirely internet-based. The data can be accessed anywhere if there is an internet connection. One of the most significant benefits of using cloud computing is cost saving. The infrastructure is invested and managed by cloud providers. The substantial capital cost can be saved for the cloud computing subscribers since no physical hardware investments, no software installations and purchases, and no physical warehouse with power and cooling maintenance are on-site. No IT professional is needed to manage the infrastructure. Security is another advantage of cloud computing. Data can be stored and mirrored at multiple redundant sites on the provider's network. Cloud computing provides reliability on data backup and disaster recovery.

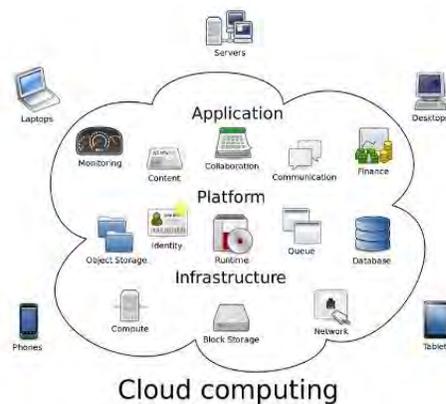


Figure 1. Cloud computing (Wikipedia, 2020)

Our project is implemented on Google Colab, which runs on the Google Cloud Platform. A folder is created to contain known face images with names (training images). Face images to be identified (testing images) are captured daily. The face recognition algorithm on Colab is run to generate the attendance results.

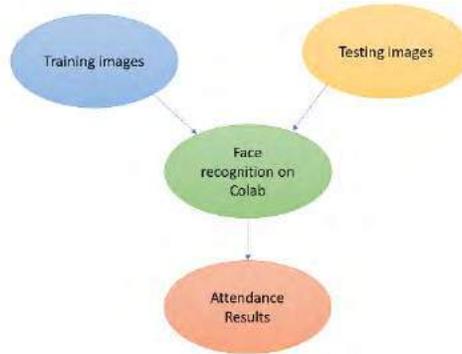


Figure 2. Face recognition process

Colab is a powerful and flexible application based on the open-source Jupyter notebook. Machine learning programming requires lots of computing resources to process extensive input data, train the deep learning model, store the trained model, and deploy it. Colab allows the users to write and execute codes, save and share analyses, and use powerful computing hardware entirely from the browsers with no software/hardware installations and free of charge. CPU stands for Central Processing Unit. GPU stands for Graphics Processing Unit, and TPU stands for Tensor Processing Unit. The user can choose to use GPU or TPU, which costs a lot in the market but is free to use in Colab. A CPU printing speed is like character-by-character, while a GPU can print a few words simultaneously. For a TPU, it can print a whole page at a time. They are used for different applications. The high-performance hardware in Colab is readily accessed by choosing the Notebook settings. For example, in our application, GPU is required for image processing.

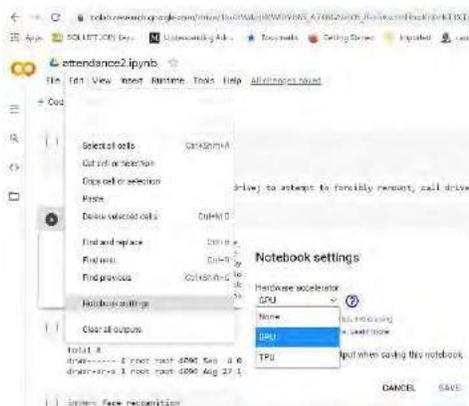


Figure 3. GPU/TPU support in Colab

The known face database and the daily captured images for identification can be imported to Colab in multiple methods. The `files.upload()` function in the Colab notebook gives the users a conventional upload button in a Colab notebook to upload files from the local computer file system into the Colab environment. The disadvantages of the method are that it might take longer for larger files, and manual upload is required whenever the notebook is restarted for any reason. Another approach is to upload

local files to Google Drive, which is integrated with Colab. The users then mount Google Drive onto the Colab environment so the notebook can directly access files in Google Drive. Manual reloading is not mandatory when the notebook restarts. For larger projects with higher architecture standards, Cloud storage such as Google Cloud Storage or AWS S3 for file storage and SQL for relational database storage provide external storage and better scalability and security options. In our project, Google Drive stores the face database and the captured images' folders.

```
[ ]  
from google.colab import drive  
drive.mount('/content/drive')
```

Figure 4. Google Drive Mounting

Colab has many popular pre-built machine learning libraries such as PyTorch, TensorFlow, Keras, and OpenCV. The notebook in Colab can be saved and published on the GitHub repository shared with the public community. A GitHub repository can be fetched and cloned into Colab. Installing libraries in Colab is simple.

3. Face recognition

Facial recognition as a biometric method has become easier to achieve with improved camera equipment and machine learning development. Many algorithms have been developed that differ in efficiency, requirements, and processing time (Li, Mu, Li, & Peng, 2020) (Taskiran, Kahraman, & Erdem, 2020) (Masi, Wu, Hassner, & Natarajan, 2018) (Insaf, Ouahabi, Benzaoui, & Taleb-ahmed, 2020). Image-based face recognition (FR) methods can be classified into four main groups:

- a. Holistic methods. Statistical tools such as SVM (Support Vector Machines), PCA (Principal Component Analysis), LDA (Linear Discriminant Analysis), independent component analysis(ICA) are the most known linear techniques employed. For example, the Eigenfaces works (Sirovich & Kirby, 1987) (Kirby & Sirovich, 1990) present using PCA to efficiently extract essential features from the training face images and determine the weights of an individual face by projecting the face image into a standard facial image called eigenpicture (Femmam, M'Sirdi, & Ouahabi, 2001). Those weights can be used to decode and reconstruct any face image. This method reduces computation and space complexity by extracting the essential elements from the training data and rejecting the less critical segments. However, recognition accuracy is low in an unconstrained environment with light variability and facial expression change.
- b. Geometric-based methods. Active Appearance Model (Edwards, Taylor, & Cootes, 1998) produces a combined statistical photo-realistic model of shape and grey-level appearance. Landmarks on a face that localize facial key structures, including the nose, the eyes, the mouth, the eyebrows, and the jaw, are matched to the target image in an optimization process. The least-square techniques are used to quickly evaluate the difference between the current estimate of appearance and the target image to drive an iterative matching scheme. High memory usage and a large amount of clean and labeled training data are required for geometric-based methods.

- c. Texture (local appearance) based methods. The work (Ahonen, Pietikainen, Hadid, & Maenpaa, 2004) proposed a facial representation based on Local Binary Pattern (LBP) features. A face descriptor is built by dividing face images into several regions. The LBP features from those regions are extracted and concatenated into an enhanced feature vector. The approach achieved high efficiency and accuracy for recognition rate but did not perform well for other factors such as low resolution, variances of posture, and different illumination conditions.
- d. Deep learning-based face recognition methods. Deep learning applies successive hidden layers to process information and extracts features to form a data abstraction hierarchy. The approach shows strong tolerance to the facial pose, lighting, and expression changes. DeepFace (Taigman, Yang, Ranzato, & Wolf, 2014) and DeepID (Sun, Chen, Wang, & Tang, 2014) first showed to surpass the human-level performance to recognize faces in variant circumstances. Although the deep learning approaches are robust, there are some challenges. Improvements are needed to reduce the large number of training models. Advanced hardware for iterative model optimization is required. An optimal global solution may be missing in the final stage.

Our recognition system is based on Adam Geitgey's face recognition module (Geitgey, 2016). The algorithm works in the following steps:

- 1) Detecting the faces in an image: Histogram of Oriented Gradients (HOG) (Dalal & Triggs, 2005) is used in this step. Color images are first converted to black and white. The distributions of gradients' directions are set as features. The magnitude of gradients is usually larger around regions with sharp intensity changes such as edges and corners where more information about object shape is filled than flat areas. Then the image part with the most similar pattern to a known HOG training face pattern is extracted.

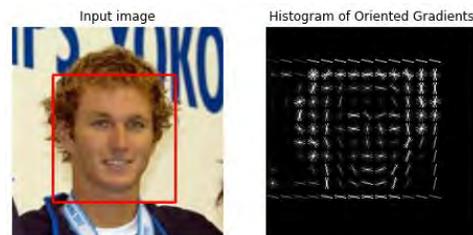


Figure 5. A HOG representation of an image

- 2) Face positioning: There are 68 specific points/landmarks in a human face to locate eyebrows, eyes, nose, mouth, etc. Those landmarks help to reposition faces in different directions, so the eyes and mouth are about in the same position in the image to compare. The face alignment technique (Kazemi & Sullivan, 2014) is employed in this step.

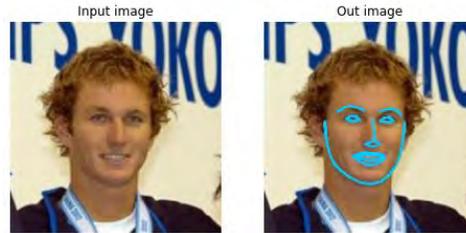


Figure 6. Landmarks on a facial image

- 3) Face encoding: After the faces are detected in a given image, distinctive facial features are extracted to identify faces. There are 128 key measurements to determine a face’s facial features (Schroff, Kalenichenko, & Philbin, 2015). A deep convolutional network has been trained to optimize the embedding process directly (Amos, Ludwiczuk, & Satyanarayanan, 2016). Similarity detection is based on the numbers generated by the network. Two different pictures of the same face will generate nearly the same numbers of 128 key measurements.

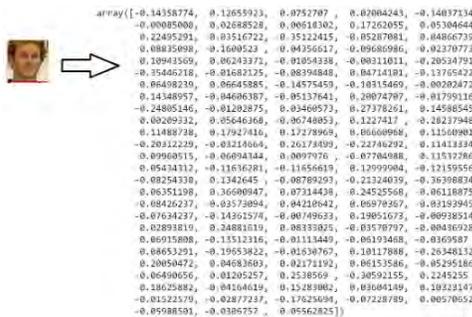


Figure 7. An example of encodings on a facial image

- 4) Matching names in the known images: A simple linear SVM classifier is trained for known face measurements. Then new test images can be classified to the closest match of known people or found not exist. Missing names are also reported.
- 5) Recording attendance results: After the facial recognition process, the system will generate an attendance CSV sheet named by the date that includes the identified people’s names and times. Another CSV sheet will consist of missing people or image names with nameless faces.

4. Results

The described attendance system is easy to implement and efficiently detect and label faces. All images of known faces are contained in the “known” folder, and testing images are included in the “unknown” folder. After the data folders are uploaded to Google Drive, the program can detect faces with different angles and generate plausible recognition results and timestamps saved in CSV records. Testing images are either recognized or unidentified (including images from people not in the roster). Missing people are also reported. Adam Geitgey’s face recognition module (Geitgey, 2016) uses the state-of-the-art face recognition dlib built with deep learning models. It has an accuracy of 99.38% (Face Recognition - Dlib, 2021) on the Labeled Faces in the Wild benchmark (Huang, Ramesh, Berg, & Learned-Miller, 2007).

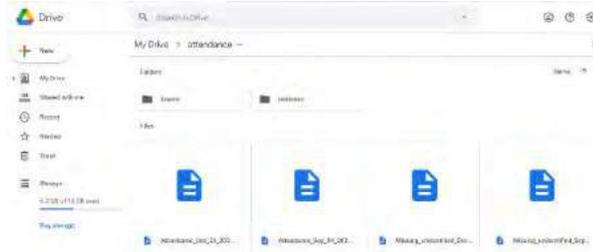
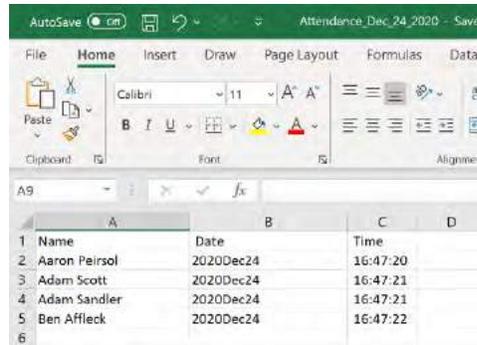


Figure 8. CSV records and image folders in Google Drive

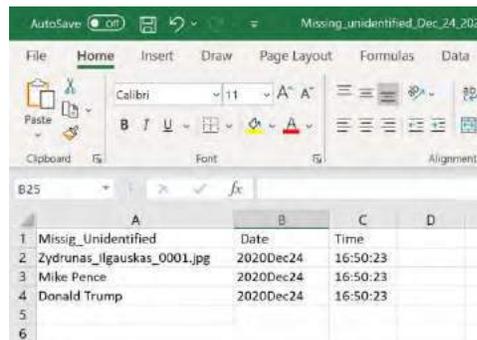


Figure 9. Examples of faces (Huang, Ramesh, Berg, & Learned-Miller, 2007) detected and recognized



	A	B	C	D
1	Name	Date	Time	
2	Aaron Peirsol	2020Dec24	16:47:20	
3	Adam Scott	2020Dec24	16:47:21	
4	Adam Sandler	2020Dec24	16:47:21	
5	Ben Affleck	2020Dec24	16:47:22	
6				

Figure 10. Attendance_Dec_24_2020.CSV



	A	B	C	D
1	Missig_Unidentified	Date	Time	
2	Zydrunas Ilgauskas_0001.jpg	2020Dec24	16:50:23	
3	Mike Pence	2020Dec24	16:50:23	
4	Donald Trump	2020Dec24	16:50:23	
5				
6				

Figure 11. Missing_unidentified_Dec_24_2020.CSV

5. Conclusion and future research

In this paper, a demo of a cost-effective attendance management system is presented. The face recognition module run on the cloud environment is used to detect faces and mark attendance. The system achieves satisfactory results for recognizing faces from different angles and expressions without requiring any bulk program installation or hardware. Furthermore, the method can extend to recognize multiple faces in the same picture or faces in a video. Partial occlusion that a face is covered with sunglasses, a scarf, a mustache, etc., is a face recognition problem (Fu, Chiu, & Wang, 2017) (Anwarul & Dahiya, 2019). In the covid-19 pandemic, masks have been required in most areas, and partial occlusion becomes a significant challenge for face recognition. Future research can focus on overcoming this kind of issue.

6. References

- Ahonen, T., Pietikainen, M., Hadid, A., & Maenpaa, T. (2004). Face recognition based on the appearance of local regions. Proceedings of the 17th International Conference on Pattern Recognition, 2004, Vol 3, pp. 153-156. Cambridge, UK.
- Amos, B., Ludwiczuk, B., & Satyanarayanan, M. (2016). OpenFace. Retrieved December 2020, from <https://cmusatyalab.github.io/openface/>
- Anwarul, S., & Dahiya, S. (2019). A Comprehensive Review on Face Recognition Methods and Factors Affecting Facial Recognition Accuracy. Proceedings of Recent Innovations in Computing(ICRIC), 2019, pp 495-514.

- Dalal, N., & Triggs, B. (2005). Histograms of oriented gradients for human detection. Conference on Computer Vision and Pattern Recognition (CVPR'05), 2005, pp. 886-893.
- Edwards, G. J., Taylor, C. J., & Cootes, T. F. (1998). Interpreting face images using active appearance models. Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition, 1998, pp. 300.
- Face Recognition - Dlib. (2021). Retrieved July 2021 from http://dlib.net/face_recognition.py.html
- Femmam, S., M'Sirdi, N., & Ouahabi, A. (2001). Perception and characterization of materials using signal processing. IEEE Transactions on Instrumentation and Measurement, Vol 50, pp. 1203–1211.
- Fu, T., Chiu, W., & Wang, Y. (2017). Learning guided convolutional neural networks for cross resolution. 2017 IEEE 27th International Workshop on Machine Learning or Signal Processing (MLSP), pp. 1-5.
- Geitgey, A. (2016). ageitgey/face_recognition. Retrieved December 2020, from https://github.com/ageitgey/face_recognition
- Huang, G., Ramesh, M., Berg, T., & Learned-Miller, E. (2007). Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments. Amherst: University of Massachusetts.
- Insaf, A., Ouahabi, A., Benzaoui, A., & Taleb-ahmed, A. (2020). Past, Present, and Future of Face Recognition: A Review. Electronics, 9(8):1188.
- Kazemi, V., & Sullivan, J. (2014). One millisecond face alignment with an ensemble of regression trees. IEEE Conference on Computer Vision and Pattern Recognition. 2014, pp. 1867-1874.
- Kirby, M., & Sirovich, L. (1990). Application of the Karhunen-Loève procedure for the characterization of human faces. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1990, Vol 12, pp. 103-108.
- Li, L., Mu, X., Li, S., & Peng, H. (2020). A Review of Face Recognition Technology. IEEE Access, 2020, Vol 8, pp. 139110-139120.
- Masi, I., Wu, Y., Hassner, T., & Natarajan, P. (2018). Deep Face Recognition: A Survey. Conference on Graphics, Patterns, and Images(SIBGRAPI), 2018, pp. 471-478.
- Schroff, F., Kalenichenko, D., & Philbin, J. (2015). FaceNet: A unified embedding for face recognition and clustering. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 815-823.
- Sirovich, L., & Kirby, M. (1987). Low-Dimensional procedure for the characterization of human faces. Journal of the Optical Society of America, Vol 4, 519–524.
- Sun, Y., Chen, Y., Wang, X., & Tang, X. (2014). Deep learning face representation by joint identification-verification. Proceedings of the 27th international conference on neural information processing systems(NIPS), 2014, pp. 1988-1996.
- Taigman, Y., Yang, M., Ranzato, M., & Wolf, L. (2014). Deepface: Closing the gap to human-level performance in face verification. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014, pp. 1701–1708.
- Taskiran, M., Kahraman, N., & Erdem, C. (2020). Face recognition: Past, present, and future (a review). Digital Signal Processing, 2020, pp. 106.
- Wikipedia 2020. Retrieved December 2020: https://en.wikipedia.org/wiki/Cloud_computing

Review of Supply Chain Metrics to Support Performance Excellence

Alper Senol¹

Ahmed A. Bakhsh²

Ahmad K. Elshennawy³

¹Department of Industrial Engineering & Management Systems, University of Central Florida, Orlando, FL 32816 USA, +1(757)814-4949, senolalper@knight.ucf.edu

²Department of Industrial Engineering, King Abdulaziz University, Jeddah, 21589 Saudi Arabia, aabakhsh@kau.edu.sa

³Department of Industrial Engineering & Management Systems, University of Central Florida, Orlando, FL 32816 USA, ahmad.elshennawy@ucf.edu

Abstract

The supply chain industry is one of the fast pace and technology-driven industries. The dynamic supply chain concept expresses how significant is the adaption to immediate changes in the industry. Many organizations strive to manage their operations with high visibility to adopt changes for performance excellence.

This study identifies the metrics of supply chain performance that affect the performance excellence of supply chain operations. It investigates categories of performance metrics in terms of time, cost, and quality. Likewise, it discovers the characteristics of performance metrics through measuring working capital in the supply chain. In addition, it defines the key performance indicators (KPIs) and its core calculations for the end-to-end supply chain to measure the practicality of operations' efficiency. The supply chain operations with respect to critical KPIs are determined as plan, source, make, delivery, and return.

The proposed performance metrics- KPIs presented as the milestones of dynamic supply chain models. Moreover, the change in the supply chain was reviewed due to the Covid-19 pandemic along with performance metrics.

Keywords: Supply chain management, Performance metrics, Key performance indicators, Quality metrics, Covid-19.

1. Introduction

The concept of global supply chain management (GSCM) is largely dependent on the size of the organizations and the types of products and services provided. According to Wu (2016), the main purpose of a supply chain is to make products available within planned time, quantity, place, and condition for the customer (Wu et al.2016). However, organizations face common challenges due to the complexity of the supply chain systems. Controlling and monitoring supply chain operations become imperative in the industry regardless of the size of the companies as well as

the types of products and services they provide. So that, supply chain management (SCM) performance metrics as known as key performance indicators (KPIs) are needed for evaluating and monitoring the current progress of the operations. According to Gunasekran et al. (2001), performance metrics and measurement for the supply chain are needed to monitor the material, information, and cash flow, as well as providing feedback for the decision-making process, and help with non-value items elimination in the supply chain operation.

GSCM metrics and measurements can give management a better understanding of how the organization is operating within a certain period. Along with the help of technology, tracking the metrics would be much easier and accurate. In addition, it boosts performance excellence in the areas of sourcing, manufacturing, warehousing, inventory control, transportation & distribution, and customer service. As Ananda and Grover stated in their article; even if supply chain performance metrics have significant importance, implementation, and integration of these metrics in the organization are always a very challenging subject. It is mentioned to add certain KPIs in place could be a complicated process and very challenging for companies, which do not have incentive programs, support from the upper management and company culture that is willing to adopt any kind of implementation for a performance measurement system (Anand & Grover, 2015). Nevertheless, this paper aims to bring practical solutions for the organizations to set up essential operational metrics in terms of time, cost, and quality perspective.

1.1. Fundamentals of Global Supply Chain Management

Global Supply Chains Management (GSCM) incorporates the suppliers, the manufacturers, the vendors, the logistics and transportation channels, and most importantly the customers. Management of the operations covers product or raw material procurement, manufacturing of the products, and distribution of final product to the end customer effectively. Supply chain organization manages the planning, evaluation, improvement, and optimization of their operations to be more efficient and proactive within a certain period (A. Jayaram, 2016).

Managing the supply chain involves many activities such as material planning & product development, manufacturing procurement, transportation, warehouse & fulfillment, and distribution. Organizations perform the entire physical and informational supply chain activities to deliver the products to end customers. Handfield & Nichols (2002) summarized the definition of the supply chain as "activities within the supply chain operation includes mainly physical and informational flow that cover procurement of the raw material to the delivery of the final product to the end-customer." To be able to manage all the supply chain operations, companies focus on various systems integration channels to share the information. It is essential to share correct and accurate information among the related parties within the supply chain to make business and operational decisions correctly and effectively. Shuangqin Liu & Bo Wu (2010) proposed two core principles of supply chain management that organizations follow: system principles and information-sharing principles. They stated that the supply chain is considered as a system that includes subjects related to different functions and their integration with each other. The core of the organization aims to integrate and connect the functions such as the supplier, manufacturer, distributor, retailer, and end-user by managing and monitoring information and capital flow. Also, in terms of information sharing, the accuracy of the operational information helps the enterprise make more effective and correct decisions. (Shuangqin Liu & Bo Wu, 2010). However, companies struggle to understand how to determine and use the information gathered to develop supply chain metrics. Choosing the essential KPIs is quite tedious to study. It is important to evaluate operational processes in the supply chain to set the correct number of the most effective metrics to quantify

its performance. Based on the proposed framework, GSCM consists of planning, sourcing, making, delivering, and accepting returns. Figure 1 shows a diagram of GSCM flow. On the other hand, integration of the set of correct KPIs is necessary to monitor such operations. Most importantly, companies would have the opportunity to target their strategic goals in terms of time, cost, and quality perspective by having the selected KPIs. In Figure 2, the paper proposes a framework in the theoretical aspect for the relationship between supply chain operations and related KPIs to be able to achieve the equilibrium point of time, cost, and quality constraints.



Figure 1. GSCM flow



2. Supply Chain Performance Metrics

The success of supply chain management is measured by its performance in three main areas such as time, cost, and quality. These metrics help organizations to recognize strengths and weaknesses in operations and can provide a simplistic approach for business decisions. Hausman (2004) points out that to be able to see success in changing supply chain environments, continuous improvement methodologies should be adapted. To accomplish the integrated continuous improvement culture, performance measures, or "metrics," should be used as global Supply Chain Performance improvements tools instead of narrow company-specific or function-specific (silo) metrics that potentially prevent improvements in the organization. In addition, Chae (2009) stated in his research paper that, supply chain companies could evaluate potential issues and areas for continuous improvement by monitoring performance metrics. Measuring metrics allows organizations to eliminate any gap between the planning to execution stages of the operations (Chae, 2009). On the other hand, Angerhofer and Angelides (2006) confirmed that the primary goal of performance excellence in the supply chain is achieved by the collective effort of each member of the organization (Angerhofer and Angelides, 2006). Moreover, performance metrics along with robust supply chain dashboards provide visibility to businesses in the supply chain industry to make timely and effective decision-making in such a highly volatile environment. Organizations rely on visual presentations of their supply chain operational performance throughout the chain of operational locations, business units, warehouses, production factories, product categories, and brands. Standardized KPIs can be compared to actual operational performance systematically. Supply chain dashboards reflect a precise view of the operations that help management to discern the problematic areas in the operation and act promptly before these problems turn out to be major issues for the operations. According to Sandra Durcevic (2021), existing software models in use for the supply chain management process are Supply Chain management dashboards and Supply Chain Costs Dashboard. Supply Chain Management Dashboard targets to pinpoint inventory metrics such as inventory-to-sales ratio, inventory turnover time, inventory accuracy, to reduce the unnecessary cost along with increasing customer retention and business stability. On the other hand, Supply Chain Costs dashboards reflect the metrics related to the cost of the operations such as average cash- to- cash cycle, net sales, inventory carrying cost, warehouse operational cost, transportation cost. (Durcevic, 2021)

Overall, three main constraints cover the supply chain operational metrics. Also, the collaboration and the individual performance of time, cost, and quality metrics measure operational efficiency and risk/return ratio.

2.1 Time

Supply chain management heavily relies on managing the time concept effectively from procurement to delivery of the goods to the customer. According to researchers in the supply chain subject matter; time performance has a direct impact on embedded strategic processes within value delivery systems and has introduced various measurements to assess the time-related supply chain performance (Tersine and Hummingbird, 1995; Naylor et al., 1999; Mason-Jones and Towill, 1999). Companies practically monitor their operations from sourcing to shipping the final product to the end customer. Time has always significant impact on cost and quality. For example, outsource inventory should be delivered on a planned date to keep the operations and inventory management costs minimal. In addition, it increases customer satisfaction when products are delivered to the end customer. So that, time-related KPIs play an essential part in performance excellence to stay highly competitive. The most useful metrics are Supplier Lead Time in sourcing perspective, On-Time

Production in terms of making/production aspects, and finally On-Time Shipping for monitoring delivery efficiency.

2.2 Cost

Supply chain organizations strive to increase operational performance and customer satisfaction by reducing overall costs. According to Randall and Theodore, the focus on supply chain management has limitations because of intra-firm collaboration in the classical logistics functions such as transportation, warehouse, and inventory management. However, the success started when a collaborative approach initiated on these functions between trading partners in an inter-firm aspect (Randall & Theodore, 2009). Consequently, nowadays supply chain management shifts to controlling the flow of the processes such as products, information, and cash.

Due to the highly complex structure of the global supply chain operation, the concept of managing operational costs is complicated. Major supply chain cost drivers are listed as start-up costs such as warehouse space, and material handling equipment is also known as initial investment cost as well as inventory cost, production/manufacturing cost, transportation cost, and procurement cost. Most importantly, companies need to know how and what has to be measured to run overall effective supply chain operations. Therefore, it is essential to monitor certain KPIs such as Forecasting accuracy rate, and Inventory turnover rate that have a direct impact on procurement cost. In addition, the distribution cost rate over the product sold affects the transportation cost. Effective supply chain network planning, and routing; and getting better shipping rates and packaging utilization are practical approaches to manage distribution costs.

2.3 Quality

Measuring and managing success in customer service practice and customer satisfaction are the essential parts of supply chain management for competitive advantage in the global market. Organizations in the supply chain industry use various quality metrics to assure that they are doing the most possible interactions with customers and manage the operation effectively and satisfy their customers at a manageable and suitable cost. Carmignani said the customer expectations increase tremendously for the product with level customer service and delivery time. This creates quite a bit of concern in the industry to be able to become competitive in the global market. (Carmignani, 2009). On the other hand, the quality discussed in Guangshu Chang's article for supply chain management focuses on a customer that has strong leadership, involving people, managing processes and systems, initiating continuous improvement, having a factual approach for the decision-making, and considering long term relationship with the supplier. It is also considered the eighth modern quality application in global supply chain management (Chang, 2009).

Defining the measurement metrics is the initial step to making sure a company's strategic goal aligns with the action taking with the progress. The quality factor has a significant impact on the relationship between suppliers and customers. Performance metrics such as supplier defect rate help to assess the strength of that relationship. Customers may lose interest in suppliers that have higher defect rates. Due to the defective products have a direct impact on production, they potentially have negative effects upon process cycle time, order fill rate, on-time delivery, rework, and return. Consequently, results in low satisfaction levels over the customer end. Order fill rate, product return rate, and product damage rate are the key KPIs to monitor the synchronization of the operations to meet the quality standards.

3. GSCM operations with proposed KPIs

Based on the literature review for the existing SCM KPIs, there are around twenty KPIs effectively used in SCM. To support the validations of the most effective KPIs, Bressoles and Lang (2020) prepared a survey questionnaire and sent it out to hundred companies to identify the most relevant SCM KPIs. There was a predetermined 30 KPIs in the questionnaire. As a result of the survey, 13 most relevant KPIs are identified within different criteria. In terms of the SCM operational aspect, inventory cycle time, safety stock, rate of obsolete inventory, order to delivery time, and out of stock rate are highly used by the companies to monitor stock and inventory efficiency. In addition, shipping accuracy, cost per shipment, invoice accuracy are used to evaluate picking and order efficiency. Lastly, the percentage of shipment arrived in good condition, and the number of damage claims is used to understand the causes of return items (Bressoles & Lang, 2020). In this paper, SCM operations are grouped into five categories; plan, source, make, deliver, and return. Proposed KPIs are distributed and organized within relevant categories to present a logical and convincing approach.

3.1. Plan

Planning is the process performed by the company that acquires the final product. The company decides the type and quantity of the product for manufacturing. It also includes the activities of demand forecasting, planning, and scheduling manufacturing operations, material requirement planning, and distribution requirement planning. Software programs such as Oracle, SAP, and i2 can compute large amounts of data used for predictive modeling and decision-making.

3.1.1. Proposed KPIs and related Constraint

Forecasting accuracy. Accurate forecasting data plays a very important role in achieving successful management goals in the supply chain. It supports reducing the cost of inventory that does not need it within a given period. Lee and Aviv summarized the concept of forecasting accuracy that collaborative forecasting methods have been studying with the majority of the companies in the supply chain industry to eliminate excess inventory and avoid the bullwhip effect (Lee et al., 1997; Aviv, 2007). In a larger perspective, forecasting accuracy lowers the supply chain operation cost. The main constraints for performance excellence in the supply chain are cost-related which helps shorten lead times and enhances customer satisfaction. Moreover, accurate data and transactions between manufacturer and buyer reduce the error for the forecasting model. The collaborated supply chain performance metrics demand is informational support from the buyer's-ends than the supplier's that may improve the results of predictive analysis and forecasting accuracy. It can eliminate excessive on-hand inventory and achieve better operational performance in the supply chain (Lorentz et al., 2011).

Forecasting accuracy formulated as follows;

Forecasting Accuracy = Predictive Analysis / Actual Data from Operations (sales & inventory)

Inventory Turnover Rate. It is a numerical output of the inventory cycles in the company each year. Inventory management is considered one of the core operating performance metrics in the supply chain industry. As Ronald H. Ballou stated, the cost of carrying inventory is the top priority

of the upper management since inventories consume a noticeable investment by many companies. Usually, the cost of carrying inventory changes between 20% - 40% of the inventory value itself (Ballou, 2000). Inventory turn around rate is an essential indicator for production companies to compete in the market according to Bahl and Ritzman (1984). On the other hand, this indicator is affected by manufacturing companies' lot sizes and capacity decisions, as recognized by Bahl and Ritzman (1984) and Bishop (1979). A higher and faster inventory turnover increases the profitability of the enterprise. Consequently, the main constraints for the performance excellence of inventory turnover rate would be cost-related and it formulated as follow;

$$\text{Inventory Turnover Rate} = \text{Total Cost of Goods Sold} / \text{Average On-Hand Inventory}$$

3.2. Source

Sourcing is a procurement process controlled by an institution that performs purchasing activities. according to Ben-Daya et al. statement sourcing process is where products or services are acquired by the company. Accurate forecasting and planning help sourcing for the supply chain organizations Making-insourcing versus buying outsourcing, selecting the supplier, and cost of purchase is the strategic decisions of the enterprises in the industry (M. Ben-Daya et. Al, 2017). Effective sourcing process rests on two pillars of supplier delivery performance and product quality. Verdouw, Beulens, and van der Vorst (2013) argue that quality control and planning are enhanced by virtual control systems. It allows buyers and suppliers to manage and track the progress of the operation in the supply chain such as tracing goods as they move through in each stage. To support the importance of the communication between supplier and buyer, organizations integrating various technological systems. For example, an electronic data interchange (EDI) system was built to increase the speed and accuracy of delivering and processing business documents. In terms of vendor-customer relationship, EDI adds so many values to any organization such as improving customer service (Tan, Kannan, & Handfield, 1998).

3.2.1. Proposed KPIs and related Constraint

Supplier Lead Time. It is the total amount of time between the received order from the buyer and the time the supplier ships the requested order quantities. Ideally, suppliers aim to provide shorter lead times, but it is quite challenging to get shorter lead-time due to the variability in the process. As Bandaly stated, variability in lead-time is extremely important in supply chain management and is controlled by managing uncertainty. Various risk management factors are used to manage this uncertainty since lead-time is considered an internal factor to the supply chain. (Bandaly et al., 2012). In addition, longer lead times cause and higher on-hand and in-transit inventory that increases operational cost. Main constraints for the performance excellence of supplier lead-time would be time-related and affect the cost indirectly and it formulated as follow;

$$\text{Supplier lead time} = \text{Time for delivery of the goods to destination} - \text{Time when PO created}$$

Supplier Defect Rate. Supplier defect rate is the measurement that checks the final product quality received from the manufacturer or supplier. It helps to measure a supplier's production quality and calculates the percentage of the received product that does not comply with the requested quality specification of the buyer. Shivasankari B. points out the procurement KPIs that enterprises cannot ignore. He mentioned that to measure defect rates of suppliers,

defect types need to be broken down with reason codes. This provides insight information about your supplier's strengths and weaknesses. Defect rates can be measured in defects per million (Bhuvaneshwaran, 2019). The main constraints for the performance excellence of supplier defect rate would be quality-related and it formulated as follow;

$$\text{Supplier defect rate} = \text{Number of defective items received} / \text{Total number of items tested}$$

3.3. Make

Sourced raw materials process by the manufacturer to be able to make the final product. Jaraham re-phrases the "make" component of GSCM in his article; enterprises such as suppliers and manufacturers work with raw materials to be able to produce finished goods. Usual operation includes; logistics flow, marketing, finance, R&D, manufacturing, and sourcing. Every department and team members communicate among them during the manufacturing process (Jaraham, 2016). There are several production-related KPIs to monitor how well manufacturing processes are performed.

3.3.1. Proposed KPIs and related Constraint

On-Time Production Rate. This metric evaluates planned production versus actual production times on a daily, weekly, and monthly basis. Weak On-Time Production performance affects directly the end customers, and it indicates poor production efficiency and materials handling procedures. The term "takt time" calculates planned product time. In theory, this time determines what quantity of the items the manufacturer should produce within the day to meet the planned throughput (Duanmu, J., & Taaffe, K., 2007). The main constraints for the performance excellence of on-time production rate would be time-related and it formulated as follow;

$$\text{On-Time Production Rate} = (1 - ((\text{Planned Production} - \text{Actual Production}) / \text{Planned items})) * 100$$

Order Fill Rate. Order fill rate is a measurement of delivery performance and is expressed as a percentage. It is also known as DIFOT (Delivery In Full On time). This metric evaluates the percentage of an order completed and shipped on time in the first shipment. Measuring order fill rate or DIFOT explains if the company meets the customer expectation. Zinn defines order fill rate as "orders filled completely as a fraction of the total number of orders" (Zinn et al., 2002) Higher-order fill rate equate to better customer service and satisfaction. The main constraints for the performance excellence of order fill rate would be quality-related and it formulated as follow;

$$\text{Order Fill Rate} = (1 - ((\text{full order quantity} - \text{shipped quantity}) / \text{full order quantity})) * 100$$

3.4. Deliver

The delivery function is one of the most important tasks of logistics. Making precise planning and controlling product and service flow as well as storing the goods are the steps of the logistic process Lummus, et al said. (Lummus, et. Al 2001). The delivery concept is not only involved with transportation but also includes warehouse operations, inventory management, and order

management. Warehouse management systems are broadly used in the delivery process to increase productivity.

3.4.1. Proposed KPIs and related Constraint

On-time Shipping Rate. It is the percentage of orders/items completed before or at the requested ship date. According to Chae customer service and sales can be improved by on-time and consistent delivery. However, organizations that are moving into the global market face challenges and complexity of the fast product delivery more overwhelmingly (Chae, 2009). Higher rates of on-time shipping increase customer satisfaction. The main constraints for the performance excellence of on-time shipping rate would be time-related and it formulated as follow;

$$\text{On Time Shipping Rate} = (\text{Total on time orders} / \text{Total orders}) * 100$$

Distribution Cost. Total distribution cost includes inbound transportation cost, outbound freight cost, and material/packaging cost. Figure 3 shows the visual relation between three cost factors.

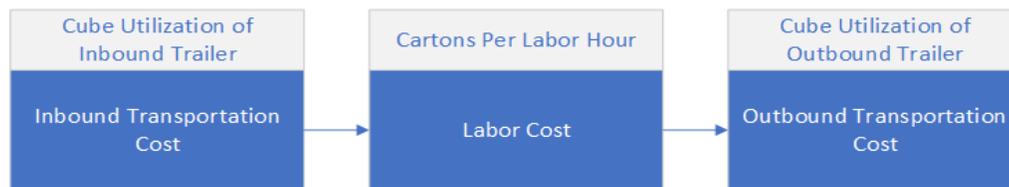


Figure 3. Factors of the distribution cost

The main constraints for the performance excellence of distribution cost would be time-related and it is formulated as follows.

$$\text{Inbound transportation cost per piece} = \text{total transportation cost} / \text{number of pieces received}$$

or

$$\text{Outbound freight cost per piece} = \text{total freight cost} / \text{number of pieces shipped.}$$

In supply chain management, the typical cost of performed distribution and transportation, operation costs can range from 6% to 8% of overall revenues for consumer-packaged-goods (Ruffing, Shehorn, & Banerjee, 2018) as shown in Figure 4.

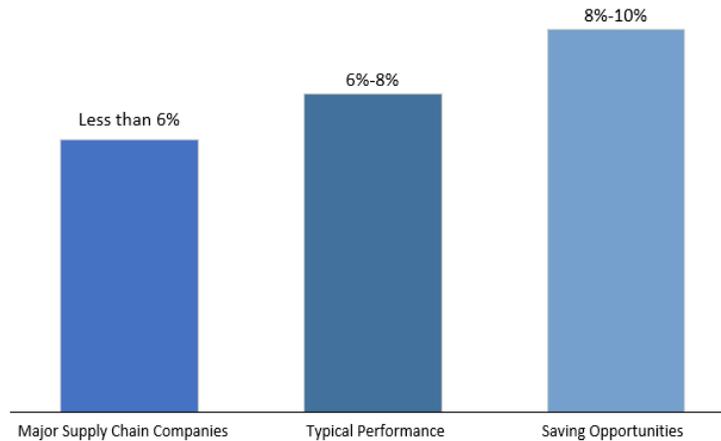


Figure 4. Cost of Distribution and transportation over revenue.

3.5. Return

Customers could return purchased goods because of various reasons such as receiving incorrect items, damaged goods, customers do not want items anymore, the wrong size, etc. Managing the return process and inventory is called reverse supply chain or reverse logistics. The reverse supply chain also includes the flow of products, information, and cash from the customer end to the manufacturer. It is a challenging process for the supply chain companies to manage the distribution and handling of the return products at a low cost. (Hildago et. Al 2019). As soon as the returned product arrives back at the warehouse, inventory carrying cost and warehouse space are affected negatively. It also consumes the warehouse labor force that causes inefficiency in the operations. Companies need to leave room in the operations to process the return products and monitor the progress and uncertainty with the following KPIs to make predictions for upcoming events.

3.5.1. Proposed KPIs and related Constraint

Product Return Rate. This metric offers insight information about why shipped items are returned and what is the percentages of the return items over the total shipped products. This metric provides the breakdown of the reasons why the customer returns items. Specifically, e-commerce and online retail stores are facing the challenge of handling return product and cost associated subject matters, Richey said (Richey et al., 2005). In addition, as Mollenkopf et al., stated, “The significance of product returns in internet retailing is derived from its contributions to customer satisfaction and long-term customer value “(Mollenkopf et al., 2011). A less product return rate increase customer satisfaction, boost profits and improve cash flow as a result. The main constraints for the performance excellence of product return rate would be quality-related and it formulated as follow;

$$\text{Product Return Rate} = \text{Total Number of Items Return} / \text{Total number of Items Sold}$$

Product Damage Rate. This metric identifies quality issues within categories that are potentially caused by either, supplier, production, or packaging. This needs to be broken down into three categorical reason codes such as received damage from suppliers, poor manufacturing, or inappropriate packaging. Data collection is crucial to identify the root cause of the damaged products. Supply chain analytics must study which items have the most significant damage rate, what are the average quantities, how do they pack, etc. To pinpoint the issues, constraints for the product damage rate would be quality-related and it can be formulated as follow;

$$\text{Product Damage Rate} = \text{Total Number of Items Damage} / \text{Total number of Items Sold}$$

4. Business Units Responsibility Distribution

Overall, proposed performance metrics are evaluated, reviewed, and communicated on regular basis between the responsible business units to execute successful supply chain operations. As Chae mentioned in his article on developing key performance indicators for supply chains; companies must set the KPI targets and adjust according to the improvement. In addition, designated team members such as Operation Strategy or S&OP team need to set up meetings to evaluate overall operations in the supply chain including demand planning to supply network designs (Chae, 2009). Closely monitored KPIs help each business unit to identify problematic or low-performance areas of the supply chain operations. It also supports the team to find the real cause of the identified problems and communicate effectively with the responsible business unit and department within the whole supply chain enterprise. Table 1 and Figure 5 Summarize the corresponding business units' distribution between related KPI as well as the impact of each KPI in terms of cost, time, and quality that support performance excellence of the organizations.

Table 1. KPI responsibility distribution.

KPI Responsibility Distribution	
Supply Chain Management KPIs	Responsible Business Units/Departments
Forecasting Accuracy	Strategic Planning
Inventory Turnover Rate	Sales and Marketing
Supplier Lead Time	Purchasing
Supplier Defect Rate	Purchasing
On-Time Production Rate	Production
Order Fill Rate	Production
On-time Shipping Rate	Production
Distribution Cost	Strategic Planning
Product Return Rate	Purchasing- Production
Product Damage Rate	Production

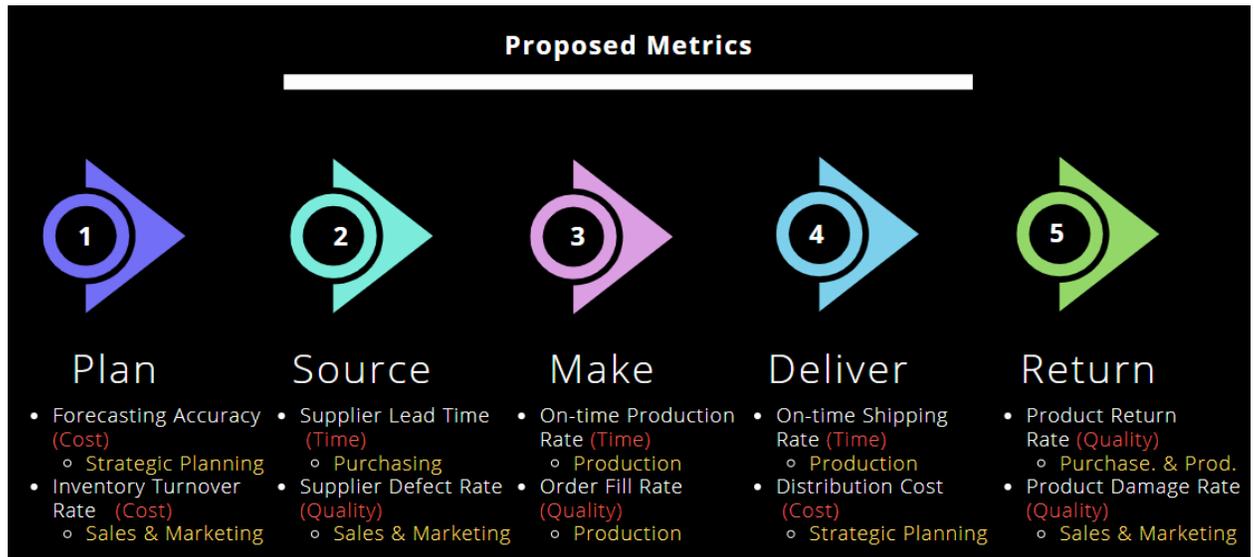


Figure 5. Proposed KPIs with the business units' responsibility distribution.

5. Adapting Immediate Changes- Dynamic Supply Chain

Proposed performance metrics- KPIs presented in this paper are the milestones of dynamic supply chain concepts. Nowadays, change and agility are inevitable in the industry. Companies with high flexibility and are open to innovative mindsets advance in the competitive market. Monitoring supply chain operations diligently helps the companies assessing critical situations and adapt to the immediate changes that are needed in the process. Especially, the latest urge for the change in the supply chain due to the pandemic caused by Covid-19 is becoming imperative for the companies and put a big spotlight on their operations to be more effective. Companies such as Home Depot, Target, Walmart, Kohl's, Best Buy, Bed Bath & Beyond, and Dick's Sporting Goods have initiated curbside pickup operations to meet their customer demand. According to adobe analytics result that is reported on CNBC news, numbers of online orders placed by customers to pick up at stores went up 208% between April 1st, 2020, and April 20th, 2020, compared to last year's sales. (Lauran, 2020). The curbside pickup concept helps most of the big players in the supply chain and retail business to reach out to their customers in a short period with low prices by eliminating shipping costs. In addition, the idea helped prevent thousands of customers exposed to the Covid-19 virus. Consequently, factors like having a diverse supplier portfolio, keeping safety stock, taking advantage of innovation, effective collaboration between business units, and being flexible become game-changer subject matters in the decision-making process. So that, using innovative tools, and techniques, as well as monitoring performance metrics, are becoming the center of interest to increase the effectiveness of these factors.

On the other hand, the next biggest challenge would be the distribution of the supply chain due to the pandemic. Customer demand for low-cost products continues to grow. Strategically assessed centralized warehouses, distribution centers, and storage spaces aim to meet this demand. Plan, source, and delivery elements of the supply chain management get the majority of the impact through these uncertain times. The biggest logistics and transportation companies such as FedEx, UPS, and DHL are preparing their operations accordingly. Recently, Covid-19 vaccine distribution has become the hottest subject for these organizations and challenge their existing shipping strategies to ship vaccines safely wherever needed in the world. Such as cold chain distribution, packaging utilization, inventory management, traceability, and technology integration are the major subject of these distributions. In this regard, UPS has started to build centralized warehouses in Louisville, Kentucky, and Venlo, the Netherlands for their cold chain distribution. In addition, FedEx has already added 10 and DHL has 8 cold chain facilities to meet the need of the covid-19 vaccine distribution. The U.S. Department of Health & Human Services Secretary Alex Azar said, "As part of Operation Warp Speed, we have been laying the groundwork for months to distribute and administer a safe and effective COVID-19 vaccine as soon as it meets FDA's gold standard." He mentioned also, "This in-depth, round-the-clock planning work with our state and local partners and trusted community organizations, especially through CDC, ensure that Americans can receive a safe and effective vaccine in record time." (HHS, 2020). According to his brief, supply chain constraints such as Time, Cost, and Quality take essential place for the distribution of the Covid-19 vaccine. Consequently, companies such as FedEx, UPS, and DHL who signed distribution contracts with the government must monitor their operations with essential key performance metrics and communicate outcomes accordingly to conduct effective vaccine distribution.

6. Conclusions

This paper represented essential measurable supply chain operational metrics, which are considered as key milestones to support performance excellence. Measurable metrics are reviewed in terms of time, cost, and quality for each global supply chain key components/processes; plan, source, make, deliver and return.

Clearly identified performance metrics contribute to the organization's strategic goals by synchronizing business units and departments together within the supply chain organization. Sharing essential and accurate KPIs eliminates miscommunication between business units, increases visibility and accessibility of the information, and eliminates the gaps between operations in the supply chain. It provides instant feedback for the current situation of the existing operations and helps the team to identify potential problems before they cause a larger negative impact in terms of time, cost, and quality for the company. The collaboration of time, cost, and quality metrics are considered to improve operational efficiency and risk/return ratio. From a practical perspective, proposed performance metrics are assigned to the related business units/department to keep them responsible and accountable for their performance. Weaknesses and strengths in the operation's output provide a simplistic approach to make business decisions.

Overall, supply chain metrics are crucially important tools and need to be monitored

against the organization's strategic goal as well as industry needs between business units and team members to support performance excellence. Successful controlled supply chain operations provide positive metrics, KPIs, and organization with the right item in the right quantity at the right time at the right place for the right price in the right condition to the right customer.

7. References

- Jayaram. (2016). Lean Six Sigma Approach for Global Supply Chain Management Using Industry 4.0 and IIoT doi:10.1109/IC3I.2016.7917940
- Anand, N., & Grover, N. (2015). Measuring retail supply chain performance. *Benchmarking: An International Journal*.
- Angerhofer, B.J. and Angelides, M.C. (2006), "A model and a performance measurement system for collaborative supply chains", *Decision Support Systems*, Vol. 42, pp. 283-301.
- Aviv, Y. (2007), "On the benefits of collaborative forecasting partnerships between retailers and manufacturers", *Management Science*, Vol. 53 No. 5, pp. 777-94.
- Ben-Daya, M., Hassini, E., & Bahroun, Z. (2019). Internet of things and supply chain management: a literature review. *International Journal of Production Research*, 57(15-16), 4719-4742. Carmignani, G. (2009). Supply chain and quality management. *Business Process Management Journal*.
- Bahl, H.C., & Ritzman L. P. (1984). Acyclical scheduling heuristic for lot sizing with capacity constraints. *International Journal of Production Research*, 22, 791-800.
- Ballou, R. H. (2000). Evaluating inventory management performance using a turnover curve. *International Journal of Physical Distribution & Logistics Management*.
- Bhuvaneshwaran, S. (2019). 11 Essential Procurement KPIs You Can't Ignore. Kissflow. Retrieved from <https://kissflow.com/procurement-process/11-essential-procurement-kpis/>
- Bressolles, G., & Lang G. (2020). KPIs for performance measurement of e-fulfillment systems in multi-channel retailing: An exploratory study. *International Journal of Retail & Distribution Management*, 48(1).
- Chae, B. (2009), "Developing key performance indicators for supply chain: an industry perspective", *Supply Chain Management*, Vol. 14 No. 6, pp. 422-428.
- Chang, G. (2009). Total quality management in the supply chain. *International Business Research*, 2(2), 82-85.
- Duanmu, J., & Taaffe, K. (2007). Measuring manufacturing throughput using takt time analysis and simulation. In *2007 Winter Simulation Conference* (pp. 1633-1640). IEEE.
- Durcevic, S. (2021). 15 Supply Chain Metrics & KPIs You Need for A Successful. Datapine Retrieved from <https://www.datapine.com/blog/supply-chain-metrics-and-kpis/>
- Gunasekaran, A., Patel, C. and Tirtiroglu, E. (2001), "Performance measures and metrics in a supply chain environment", *International Journal of Operations & Production Management*, Vol. 21 No. 1, pp. 71-87
- Handfield, R. B., & Nichols, E. L. (2002). *Supply chain redesign: Transforming supply chains into integrated value systems*. Ft Press.

- HHS- U.S. Department of Health & Human Services. (2020). Trump Administration Releases COVID-19 Vaccine Distribution Strategy. Retrieved from <https://www.hhs.gov/about/news/2020/09/16/trump-administration-releases-covid-19-vaccine-distribution-strategy.html>
- Kuei, C. H., Madu, C. N., & Lin, C. (2011). Developing global supply chain quality management systems. *International Journal of Production Research*, 49(15), 4457-4481.
- Lauren, Thomas. (2020). Curbside pickup at retail stores surges 208% during coronavirus pandemic. Retrieved from <https://www.cnbc.com/2020/04/27/coronavirus-curbside-pickup-at-retail-stores-surges-208percent.html>
- Lee, H.L., Padmanabhan, V., and Whang, S. (1997), "The bullwhip effect in supply chains", *Sloan Management Review*, Vol. 38 No. 3, pp. 93-102
- Lorentz, H., Shi, Y., Hilmola, O. P., Srari, J. S., Ramanathan, U., Gunasekaran, A., & Subramanian, N. (2011). Supply chain collaboration performance metrics: a conceptual framework. *Benchmarking: An international journal*.
- Meixell, M. J., Kenyon, G. N., & Westfall, P. (2014). The effects of production outsourcing on factory cost performance: an empirical study. *Journal of Manufacturing Technology Management*.
- Mollenkopf, D.A., Frankel, R. and Russo, I. (2011), "Creating value through returns management: exploring the marketing-operations interface", *Journal of Operations Management*, Vol. 29 No. 5, pp. 391-403.
- Randall, W., and Theodore Farris, M. (2009), "Supply chain financing: using cash-to-cash variables to strengthen the supply chain", *International Journal of Physical Distribution & Logistics Management*, Vol. 39 No. 8, pp. 669-689.
- Richey, R.G., Chen, H., Genchev, S.E. and Daugherty, P.J. (2005), "Developing effective reverse logistics programs", *Industrial Marketing Management*, Vol. 34 No. 8, pp. 830-840.
- Ruffin, R., Shehorn, M., Banerjee, D., & Lapin, J. (2018). Are Your Distribution and Transportation Costs Out of Control? Bain & Company. Retrieved from <https://www.bain.com/insights/are-your-distribution-and-transportation-costs-out-of-control/>
- Tan, K. C., Kannan, V. R., & Handfield R.B (1998). Supply Chain Management: Supplier Performance and Firm Performance, *International Journal of Purchasing & Materials Management*, 34 (3)
- Thürer, M., Pan, Y. H., Qu, T., Luo, H., Li, C. D., & Huang, G. Q. (2019). Internet of Things (IoT) driven kanban system for reverse logistics: solid waste collection. *Journal of Intelligent Manufacturing*, 30(7), 2621-2630.
- Vollmann, T., Berry, B., Whybark, D. and Jacobs, F. (2005), *Manufacturing Planning and Control for Supply Chain Management*, McGraw-Hill, Boston, MA.
- W. Zinn, J.T. Mentzer, K.L. (2002) Croxton Customer-based measures of inventory availability *Journal of Business Logistics*, 23 (2), pp. 19-43
- Wilson, R. (2004). 15th Annual State of Logistics Report. Council of Supply Chain Management Professionals. Available online: <http://www.cscmp.org/>
- Wu, L., Yue, X., Jin, A., & Yen, D. C. (2016). Smart supply chain management: a review and implications for future research. *The International Journal of Logistics Management*, 27(2), 395-417.

Building Design, Resident's Behavior, and Indoor Radon in the Commonwealth of Kentucky

Osama E. Mansour¹

Bashar Haddad¹

¹Western Kentucky University

Corresponding E-mail; oe.mansour@wku.edu

Abstract

This research aims to identify the impact of foundation type on average indoor radon levels in houses in the Commonwealth of Kentucky. It compares the averages of indoor Radon in 4 houses in Warren County, Kentucky, an area classified as zone 1 by the US Environmental Protection Agency. The predicted averages of indoor Radon greater than four pCi/L. The four houses are wood-frame structures; two of them are built on slabs-on-grade, the other two are raised on crawlspaces. A questionnaire was used to gather data about the occupant's behavior towards the indoor air quality in their houses during the test period. A T-test was performed to analyze the radon readings. The data analysis shows a significant difference in indoor Radon between houses of different foundation types. Unlike the previous studies, the houses built on slabs-on-grade show significantly lower averages of indoor Radon than those built on crawlspaces. Also, the results confirm earlier studies regarding the impact of the user's behavior on indoor air quality. Natural ventilation can significantly reduce the averages of indoor Radon. A limitation of this study is that it was carried out in only four houses. More surveys of indoor Radon in buildings with different foundation types are required for results to be generalizable.

1. Introduction

Radon is an odorless radioactive gas in the soil in areas rich in Radium and Uranium (Fuente et al., 2019). Radon is the second cause of lung cancer in the United States (Hahn et al., 2020; Stanifer et al., 2020). Testing one's home for Radon is necessary to determine exposure risk. However, despite Radon's public awareness, the proportion of people in the United States who have tested their homes remains low, ranging from 3% to 15% (Eheman et al., 1996; Stanifer et al., 2020; Wang et al., 2000). There is a level of uncertainty regarding the classification of areas with potential indoor Radon hazards (Friedmann et al., 2017). The doubt is originated from the inconsistency of Radon concentration and its progeny in the indoor environment (Barros-Dios et al., 2007). The indeterminate time of exposure to Radon and its progeny can potentially harm the human body (Spiegelman et al., 2016) and lack of evidence in the literature regarding the impact of current mitigation systems on avoiding Radon's health hazards (Vogeltanz-Holm & Schwartz, 2018). However, a cohort of researchers proved a dose-response correlation between residential exposure to Radon and lung cancer risk (Baskaran, 2016; Gillmore et al., 2018; Krewski et al., 2005). This research is aligned with the past three decades' efforts towards developing mitigation systems to reduce Radon's concentration in the indoor environment. Researchers in this field consider radon concentration as one of the major indoor air pollutants in vast areas of North America and Europe.

Radon's primary sources in houses are infiltration of radon gas from the soil and drinking water (Bruno, 1983). However, radon gas's infiltration from the soil is prioritized as the main reason for elevated Radon indoors.

There are two practical approaches to improve indoor air quality (Spengler et al., 2001). The first one is based on source control, which reduces the number of pollutants coming from a specific source by alternating the source or leaving it out of the indoor environment. For instance, to overcome the harmful impact of inhaling volatile organic compounds (V.O.C.), building designers use materials that have low V.O.C. emissions (Schieweck & Bock, 2015). The second approach is based on improving indoor air quality by increasing the ventilation rate to reduce unwanted indoor air substances. This process includes air filtration to reduce the particulate matter and the outdoor air pollutants introduced to the indoor environment with fresh air (Stabile et al., 2017).

The indoor radon concentration is increased by either high source magnitudes under the buildings or low ventilation rates (Nero, 1981). Current radon mitigation systems are primarily a source control approach. They are built on the concept of making a depressurization system underneath the building being mitigated. This system collects the gas in a depressurized chamber and uses a pipe with a fan to flush it out to the air. This system is branched into two types based on how this fan operates; Active system, in which the fan is running mechanically using electricity to keep the system working, passive system, in which the same components of the depressurization system are used. However, the system operates passively; the fan is powered by the wind movement above the house (Denu et al., 2019). Active and passive depressurization Mitigation systems are commonly used across North America and Europe. It costs an average of 2000 USD in the United States and requires an invasive installation process to cut a hole in the foundation, install a pipe and a fan inside or outside the building to get the radon gas from the soil to the outdoor air. Then the Radon dilutes the fresh air to a level that is safer for humans.

In a study of indoor radon concentration in residential buildings in Spain, factors such as building materials, age of the building were found to influence radon concentration in bivariate analysis (Barros-Dios et al., 2007). Also, ventilating buildings with fresh air can significantly reduce indoor Radon (Li et al., 2018). An experiment that examined the influence of air filtration on Radon's level in houses found that air cleaners significantly decrease Radon's decay products (Li & Hopke, 1991). In areas with a high Radon level, natural ventilation can reduce radon concentration significantly than forced ventilation (Cavallo et al., 1996). Moreover, indoor Radon and its progeny in a Japanese dwelling changed based on changing air conditioning system and lifestyle of inhabitants (Pornnumpa et al., 2015).

In the United States, foundation type is usually selected based on consumer preference, type of soil, and climate zone. There is no documented evidence for selecting foundation type to support the indoor air quality in areas prone to a high level of indoor Radon. According to the National Association of Home Builders (N.A.H.B.), the Survey of Construction data shows that using slabs single-family homes in the East South-Central region declined 48.6% from 2004 to 2019. Slabs were used in 61% of new homes built in 2004. The percentage of slabs used went down to 31.3% of new homes built in 2018. For the same period, using crawl spaces in new homes increased 166% from 2004 to 2018. Zha stated that crawl spaces were used 17.6% of new homes built in 2004. The percentage went up to 46.9% of new homes built in 2018.

Warren County is one of the areas classified by E.P.A. as zone 1. Predicted average levels of indoor Radon

exceed 4PCi/L. This study examines the impact of the two most common foundation types in the region, crawlspace, and slab-on-grade on the average levels of indoor Radon in single-family dwellings of Bowling Green, Kentucky.

Radon Testing and Initial Investigation

Radon gas is naturally changing every day based on various factors. Air pressure, temperature, humidity, wind, and precipitation can significantly affect the radon concentration indoors (Khan & Chreim, 2019; Yarmoshenko et al., 2016). The highest indoor levels are often found during the heating season. Weather conditions, operation of furnaces and fireplaces, and opening/closing of windows and doors are among the factors that cause these patterns.

Most of the research into indoor Radon's behavior involves measurements of the indoor environment's radon level through one of two ways (Yarmoshenko et al., 2016). Short-term testing using passive devices or active devices that usually take between 2 and 90 days. The passive devices include alpha track detectors, charcoal canisters, or charcoal liquid scintillation detectors. After using passive devices for testing, they are usually sent to a laboratory for analysis. On the other hand, active devices require power to operate. Continuous monitors and Alpha trackers are commonly used for long-term testing (George, 2008). A long-term test gives a reading that is more likely to reflect the building's year-round average radon level than a short-term test (Barros et al., 2014). In the present study, long-term testing with an experimental design approach offers greater potential for generalization.

2. Research Problem

Building design attributes such as materials and structure systems are not used effectively as mitigation approaches for indoor Radon. A building foundation is the first defense line against radon gas seeping from the soil. Dai and others found that a crawlspace foundation lowers the averages of indoor Radon (Dai et al., 2019). Other studies found that building on a slab significantly lowers indoor Radon's averages than building on crawlspaces (Andersen et al., 1997; Curado et al., 2017; Majborn, 1992). No documented evidence is available to support a specific foundation type to lower the averages of indoor Radon in areas prone to Radon's high level in the Commonwealth of Kentucky.

3. Research Objectives

The present research investigates the impact of the building design attributes along with the occupants' behavior on indoor Radon in the Commonwealth of Kentucky. It aims to identify which foundation type, crawlspace, or slab-on-grade, positively impacts the averages of indoor Radon. The purpose is to explore a potential role for the current design and construction practices to mitigate Radon throughout the building's lifetime.

4. Methodology

To address the research objectives, data for this study were collected from two groups of houses. Group A included houses that were built on a crawlspace. Group B included houses that were built on slab on grade (SOG). The following null hypotheses were tested:

H₀₁: For group A houses; there was no statistically significant difference in radon level readings between

the non-control house and the control house.

H₀₂: For group B houses, there was no statistically significant difference in radon level readings between the non-control house and the control house.

H₀₃: There was no statistically significant difference in radon level readings between group A non-control houses and group B non-control houses.

H₀₄: There was no statistically significant difference in radon level readings between group A control houses and group B control houses.

H₀₅: There was no statistically significant difference in radon level readings between group A houses and group B houses.

The unit used in measuring radon levels is Picocuries per liter of air, or pCi/L, which is one of the preferred measurements for the speed of decay in radon, which is equal to one trillionth of a curie, abbreviated as pCi. The pCi unit is used in the United States because it is required by federal law. Just about everywhere else that uses the metric system, including the World Health Organization, measures in Becquerels; 1 pCi/L is equal to 37 Bq/m³.

Preliminary data was collected from two different buildings located in two different locations in Warren County, Kentucky. Radon digital sensors were installed to detect indoor radon concentration averages for six months, from September 2018 to the end of January 2019. The data shows inconsistent levels of concentration higher than the threshold set by the U.S. Environmental Protection Agency (E.P.A.) 4 pCi/L. Also, some alterations were introduced to the spaces during the test period. It has been found that natural ventilation while using ceiling fans for 2 hours a day reduced the indoor radon concentration significantly. However, the reduction changed with other factors such as air temperature and precipitation over the test period. The results of this pilot study conform with previous studies in Warren County. It shows that the average level of indoor Radon is higher than what is set by the environmental protection agency E.P.A. 4 pCi/L (Webster, 1990).

The present study identifies the foundation type's impact on indoor Radon in four houses in Warren County, Kentucky. The houses were designated into two groups; group A includes two houses built on a crawlspace. Group B consists of 2 houses built on slab-on-grade. Due to the significant change of Radon concentration from one location to another, the two groups were selected in the same area of Warren County. Each group consists of two adjacent houses to minimize the impact of location on the measurements; see the houses' characteristics in Table 1. Airthings Corentium Professional Radon detectors were used to measure and record the radon concentrations. Each detector samples indoor air through a passive diffusion chamber using alpha spectrometry to calculate the radon level precisely. The detector took hourly readings throughout the testing period.

Table 1. Characteristics of the samples

Group	House	Year built	Gross Area (sqft)	Foundation	Floors	Walls	HVAC System	Air Filter size/Rating
A	1	2009	1972	Slab-on-grade	Wood flooring	Brick veneer on light wood frame	Central air conditioning - gas furnace with outdoor electric heat pump	20"X20"X1" /MERV 11
A	2	2011	1700	Slab-on-Grade	Wood flooring	Brick veneer on light wood frame	Central air conditioning - gas furnace with outdoor electric heat pump	20"X20"X1" /MERV 11
B	3	2002	2300	Crawlspace	Wood flooring	Brick veneer on light wood frame	Central air conditioning - gas furnace with outdoor electric heat pump	20"X20"X1" /MERV 11
B	4	2006	2200	Crawlspace	Wood flooring	Brick veneer on light wood frame	Central air conditioning - gas furnace with outdoor electric heat pump	20"X20"X1" /MERV 11

The houses were selected based on specific criteria: 1) to be similar in the physical characteristics such as structure system, ventilation system, and wall design, see table 1 and fig. 1&2; 2) to be around the same size, maximum of 300 sqft difference; 3) to be around the same age (the maximum difference is less than a decade) to accommodate the same technology and to minimize the impact of wear and tear; 4), to have around the same number of occupants. The selection criteria were set to mitigate the effects of factors such as materials, building systems, airtightness, and operation and maintenance practices on the results. The gathered data comprised the following two datasets.

1. Data were measured by digital radon monitors. The monitors sample indoor air through a passive diffusion chamber using alpha spectrometry to precisely calculate the radon level. The monitors were set to logging data around the clock for six months during the winter season, usually characterized by high averages of indoor Radon in houses in the area. The collected data includes Radon pCi/L, Temperature °F, Humidity %rH, and Pressure kPa.
2. Homeowners' responses, a questionnaire was used to gather information about the occupant's behavior towards the indoor air quality in their houses during the test period. The questionnaire was distributed at the end of the testing period to eliminate any impact from the questionnaire's narrative on respondents living in the control group houses.

T-test was used to analyze the difference of radon averages between the houses, the two houses

built on slabs (group A), and the two houses were raised on crawlspaces (group B). At-test statistical analysis is the hypotheses testing tool that aids researchers in determining if there is a significant difference between the means of two groups.

The researcher utilized Stata software version 13 for the statistical analysis of the data collected for the study. Stata is used by researchers due to its capabilities to perform data management, statistical analysis, and graphic simulations.

The second dataset was analyzed qualitatively to show the homeowner's practices towards indoor air quality in their houses and reveal any factors that could affect radon concentration readings.



Figure 1. The two houses representing group A



Figure 2. The two houses representing group B

In each group, one of the houses was designated as the control house, and the other one was the non-control (experimental) house. At the control house, no information nor instructions were given to building occupants. During the study period, residents lived their lives as usual and kept their daily routine without change. Due to the positive impact of natural ventilation (Andersen et al., 1997), the experimental (non-control group) were asked to open two windows across from each other for two hours a day to experiment with passive cross-ventilation as a mitigation technique to reduce radon

levels.

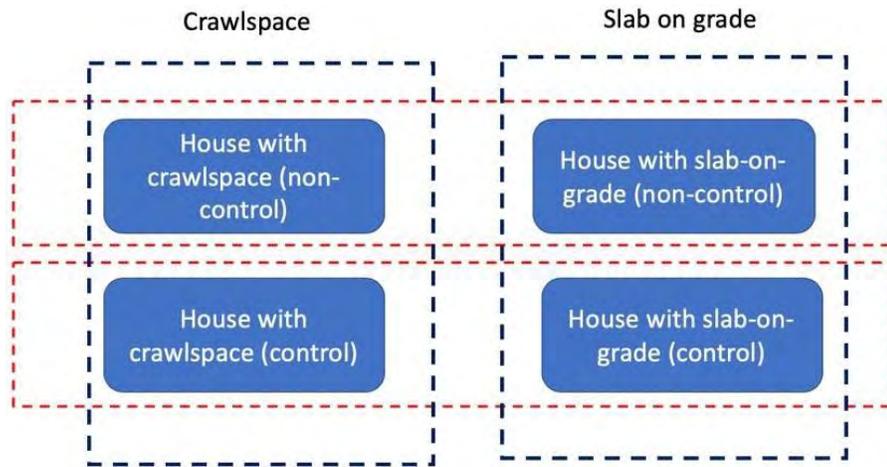


Figure 3. Designation of the houses in the experiment

5. Data Analysis and Results

The direct readings in indoor air radon were recorded hourly for each house. Apparent differences were noticed between the minimum, median, and maximum concentration levels of the control and non-control houses, as shown in Table 2.

Table 2. Radon concentrations during the test period

Houses	Radon concentration (pCi/L)			
	Minimum	Median	Maximum	Range
House 1 (crawlspac)	0.0	12.5	121.6	121.6
House 2 (crawlspac)	0.4	27.35	457.2	429.85
House 3 (slab-on-grade)	0.0	1.7	7.5	7.5
House 4 (slab-on-grade)	0.0	2.0	28	28

t-test was used to analyze the data as follows:

1. Test the significance between the averages of radon concentration in the control and non-control houses to identify the impact of user's behavior on indoor Radon when informed and given instructions to ventilate their houses.
2. Test the significance between the averages of radon concentration in the houses built on a slab and the houses raised on crawlspaces to measure the impact of foundation type on indoor Radon

Results from Null Hypothesis 1- Significance of user's behavior on Indoor Radon – control vs. non-control houses built on Crawlspaces

Table 3 reports the results of the independent sample t-test for radon readings in group A. The mean radon readings at the non-control house were 15.996 Pci/L, and at the control, the house was 37.527 Pci/l. The t-test analysis suggests a statistically significant difference between the radon readings obtained from the not control house and the non-control house in group A, $t_{(N-2)} = -17.4999$, $p < 0.01$. It suggests that occupants' awareness while imposing natural cross-ventilation for 2 hours into a house per day would significantly reduce the average indoor Radon in houses raised on crawl spaces. Therefore, we reject the null hypothesis of no difference in radon level readings between group A non-control house and the control house.

Table 3. Independent sample t-test for radon readings in group A
Two-sample t test with equal variances

Variable	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
bradon~n	1332	15.99587	.3341232	12.19435	15.34041	16.65134
rradon~n	1332	37.52718	1.184128	43.21662	35.20422	39.85014
combined	2664	26.76152	.6494839	33.52243	25.48798	28.03507
diff		-21.53131	1.230365		-23.94387	-19.11874

diff = mean(bradonconcentr~n) - mean(rradonconcentr~n) t = -17.4999
 Ho: diff = 0 degrees of freedom = 2662

$$H_a: \text{diff} \neq 0$$

$$\Pr(|T| > |t|) < 0.01$$

Results from Null Hypothesis 2 - Significance of user's behavior on Indoor Radon – (control vs. non-control) Houses built on Slab-on-Grade

Table 4 shows the results of the independent sample t-test for radon readings in group B. The mean radon readings at the not control house were 2.230 Pci/L, and at the control, the house was 2.956 Pci/L. The t-test analysis suggested a statistically significant difference between the radon readings obtained from the not control house and the control house in group B, $t_{(N-2)} = -7.9222$, $p < 0.01$. Therefore, we reject the null hypothesis of no difference in radon level readings between group B non-control house and the control house.

**Table 4. Independent sample t-test for radon readings in group B
Two-sample t test with equal variances**

Variable	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
hradon~n	1333	2.230383	.0316048	1.1539	2.168382	2.292383
jradon~n	1317	2.956188	.0864433	3.137069	2.786607	3.12577
Combined	2650	2.591094	.0463387	2.385431	2.500231	2.681958
diff		-.7258057	.0916172		-.9054543	-.5461571
diff = mean(hradonconcentr~n) - mean(jradonconcentr~n)				t =	-7.9222	
Ho: diff = 0				degrees of freedom =	2648	

Ha: diff != 0

Results from Null Hypothesis 3 - Significance of Foundation Type on Indoor Radon – House with Crawlspace Vs. Slab-on-Grade (non-control)

Table 5 reports the results of the independent sample t-test for radon readings between group A, not the control house, and group B not the control house. The mean radon readings at the group A non-control house was 15.996 Pci/L and at group B not control house was 2.230 Pci/L. The t-test analysis suggested a statistically significant difference between the radon readings obtained from the non-control house in group A and the non-control house in group B $t(N-2) = 41.0309$, $p < 0.01$. Therefore, we reject the null hypothesis of no difference in radon level readings between group A non-control house and group B non-control house.

and group B houses.

**Table 6. Independent sample t-test for radon readings between group A control house and group B control house
Two-sample t test with equal variances**

Variable	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
rradon~n	1332	37.52718	1.184128	43.21662	35.20422	39.85014
jradon~n	1317	2.956188	.0864433	3.137069	2.786607	3.12577
combined	2649	20.33956	.6848844	35.24993	18.9966	21.68252
diff		34.57099	1.193953		32.22981	36.91216

diff = mean(rradonconcentr~n) - mean(jradonconcentr~n) t = 28.9551
 Ho: diff = 0 degrees of freedom = 2647 Ha: diff != 0
 Pr(|T| > |t|) < 0.01

Table 7. Statistical analysis of indoor radon averages between the crawlspace houses and slab-on-grade houses

Variable	Obs	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
radonc~n	2650	2.591094	.0463387	2.385431	2.500231	2.681958
var5	2664	26.76152	.6494839	33.52243	25.48798	28.03507
combined	5314	14.70815	.3660836	26.68647	13.99047	15.42582
diff		-24.17043	.6528356		-25.45026	-22.8906

diff = mean(radonconcentra~n) - mean(var5) t = -37.0238
 Ho: diff = 0 degrees of freedom = 5312
 Ha: diff != 0
 Pr(|T| > |t|) < 0.01

6. Responses to the questionnaire

A questionnaire was given to the homeowners of the four houses. In addition to gathering demographic information and description for the houses, the questionnaire included questions about their indoor air quality habits; how many hours they open the windows to seek fresh air per day? How many hours do they use their ceiling fans? How often do they replace the air filters? What temperature do they set the thermostat on throughout the different seasons? Were they aware of the potentially high level of Radon in their houses?

The responses showed that none of the houses' occupants were aware of Radon's potentially high level before this experiment. The four respondents are also unwilling to open the windows for natural ventilation in winter. However, the two experimental houses opened their windows for 2 hours a day during the testing period based on their instruction. The four houses used the same air filter type with the same Minimum Efficiency Reporting Value (M.E.R.V. 11), at least during the testing period; they also replace their filters every 2-3 months. None of the houses used the ceiling fans during the test nor were willing to use them in the future.

The questionnaire's responses confirmed similar behavior towards the indoor air in the four houses except what has been instructed to the two experimental houses' homeowners. This result excludes the human behavior factor on the readings of the indoor Radon.

7. Conclusion

While indoor Radon attributes to over twenty thousand lung cancer deaths every year in the United States (Sethi et al., 2012), it attributes to a higher lung cancer death rate in the commonwealth of Kentucky (Haneberg et al., 2020). Design attributes are an affordable and non-invasive approach to radon mitigation for new construction. It potentially lowers the indoor radon averages throughout the building's lifetime. In the present study, foundation type and occupants' behavior were investigated. Unlike Dai and others (Dai et al., 2019), the results show that houses built on slabs-on-grade have significantly lower averages of indoor Radon than the houses built on crawlspaces in Kentucky. The results match the findings of Curado and others (Curado et al., 2017; Majborn, 1992). Also, the results confirm Anderson's findings (Andersen et al., 1997), natural ventilation can be considered a mitigation factor for indoor Radon. The houses showed a significant reduction in indoor Radon concentration when the occupants were given instructions to cross-ventilate their house for two hours once a day.

As inhaling Radon and its progeny correlate with lung cancer, the present study suggests that building on slab-on-grade and educating building occupants about the importance of natural ventilation would significantly reduce the averages of indoor Radon in houses in the Commonwealth of Kentucky. A limitation of this study is that it was carried out in only four houses. More surveys of indoor Radon in buildings with different foundation types are required for results to be generalizable.

8. References

Andersen, C. E., Bergsøe, N. C., Majborn, B., & Ulbak, K. (1997). Radon and natural ventilation in newer Danish

single-family houses. *Indoor Air*, 7(4), 278-286.

Barros, N. G., Steck, D. J., & Field, R. W. (2014). A comparison of winter short-term and annual average radon measurements in basements of a radon-prone region and evaluation of further radon testing indicators. *Health physics*, 106(5), 535.

Barros-Dios, J., Ruano-Ravina, A., Gastelu-Iturri, J., & Figueiras, A. J. E. R. (2007). Factors underlying residential radon concentration: results from Galicia, Spain. 103(2), 185-190.

Baskaran, M. (2016). Radon: A Human Health Hazard in the Environment. In *Radon: A Tracer for Geological, Geophysical and Geochemical Studies* (pp. 229-254). Springer.

Bruno, R. C. (1983). Sources of indoor radon in houses: A review. *Journal of the Air Pollution Control Association*, 33(2), 105-109.

Cavallo, A., Gadsby, K., & Reddy, T. J. E. I. (1996). Comparison of natural and forced ventilation for radon mitigation in houses. 22, 1073-1078.

Curado, A., Silva, J., Carvalho, L., & Lopes, S. I. (2017). Indoor Radon concentration assessment in a set of single family houses: case study held in Barcelos, North of Portugal. *Energy Procedia*, 136, 109- 114.

Dai, D., Neal, F. B., Diem, J., Deocampo, D. M., Stauber, C., & Dignam, T. (2019). Confluent impact of housing and geology on indoor radon concentrations in Atlanta, Georgia, United States. *Science of The Total Environment*, 668, 500-511.

Denu, R. A., Maloney, J., Tomasallo, C. D., Jacobs, N. M., Krebsbach, J. K., Schmaling, A. L., Perez, E., Bersch, A. J., LeCaire, T. J., & Meiman, J. G. (2019). Survey of Radon Testing and Mitigation by Wisconsin Residents, Landlords, and School Districts. *WMJ*, 118(4), 169-176.

Eheman, C. R., Ford, E., Staehling, N., & Garbe, P. (1996). Knowledge about indoor radon in the United States: 1990 national health interview survey. *Archives of Environmental Health: An International Journal*, 51(3), 245-247.

Friedmann, H., Baumgartner, A., Gruber, V., Kaineder, H., Maringer, F., Ringer, W., & Seidel, C. J. J. o. e. (2017). The uncertainty in the radon hazard classification of areas as a function of the number of measurements. 173, 6-10.

Fuente, M., Rábago, D., Goggins, J., Fuente, I., Sainz, C., & Foley, M. (2019). Radon mitigation by soil depressurisation case study: Radon concentration and pressure field extension monitoring in a pilot house in Spain. *Science of The Total Environment*, 695, 133746.

George, A. C. (2008). World history of radon research and measurement from the early 1900s to today. *AIP Conference Proceedings*, Gillmore, G. K., Crockett, R. G., & Phillips, P. S. J. G. S., London, Special Publications. (2018). Radon as a carcinogenic built-environmental pollutant. 451(1), 7-34.

Hahn, E. J., Wilmhoff, C., Rayens, M. K., Conley, N. B., Morris, E., Larck, A., Allen, T., & Pinney, S. M. (2020). High School Students as Citizen Scientists to Decrease Radon Exposure. *International Journal of Environmental Research and Public Health*, 17(24), 9178.

Haneberg, W. C., Wiggins, A., Curl, D. C., Greb, S. F., Andrews Jr, W. M., Rademacher, K., Rayens, M. K., & Hahn, E. J. (2020). A geologically based indoor-radon potential map of Kentucky. *Geohealth*, 4(11), e2020GH000263.

- Khan, S. M., & Chreim, S. (2019). Residents' perceptions of radon health risks: a qualitative study. *BMC public health*, 19(1), 1-11.
- Krewski, D., Lubin, J. H., Zielinski, J. M., Alavanja, M., Catalan, V. S., Field, R. W., Klotz, J. B., Létourneau, E. G., Lynch, C. F., & Lyon, J. I. J. E. (2005). Residential radon and risk of lung cancer: a combined analysis of 7 North American case-control studies. 137-145.
- Li, C.-S., & Hopke, P. K. J. H. p. (1991). Efficacy of air cleaning systems in controlling indoor radon decay products. 61(6), 785-797.
- Li, Y., Fan, C., Xiang, M., Liu, P., Mu, F., Meng, Q., Wang, W. J. J. o. R., & Chemistry, N. (2018). Short-term variations of indoor and outdoor radon concentrations in a typical semi-arid city of Northwest China. 317(1), 297-306.
- Majborn, B. (1992). Seasonal variations of radon concentrations in single-family houses with different sub-structures. *Radiation Protection Dosimetry*, 45(1-4), 443-447.
- Nero, A. (1981). Radon concentrations and infiltration rates measured in conventional and energy-efficient houses.
- Pornnumpa, C., Tokonami, S., Sorimachi, A., & Kranrod, C. J. R. p. d. (2015). Characteristics of indoor radon and its progeny in a Japanese dwelling while using air appliances. 167(1-3), 87-91.
- Schieweck, A., & Bock, M.-C. (2015). Emissions from low-VOC and zero-VOC paints – Valuable alternatives to conventional formulations also for use in sensitive environments? *Building and Environment*, 85, 243-252. <https://doi.org/https://doi.org/10.1016/j.buildenv.2014.12.001>
- Sethi, T. K., El-Ghamry, M. N., & Kloecker, G. H. (2012). Radon and lung cancer. *Clin Adv Hematol Oncol*, 10(3), 157-164.
- Spengler, J. D., Samet, J. M., & McCarthy, J. F. (2001). *Indoor air quality handbook*.
- Spiegelman, D., Khudyakov, P., Samet, J., Wiggins, C., Liao, X., & Meisner, A. (2016). Survival analysis with measurement error in a cumulative exposure variable: radon progeny in relation to lung cancer mortality. *ISEE Conference Abstracts*.
- Stabile, L., Dell'Isola, M., Russi, A., Massimo, A., & Buonanno, G. (2017). The effect of natural ventilation strategy on indoor air quality in schools. *Science of The Total Environment*, 595, 894-902.
- Stanifer, S. R., Rayens, M. K., Wiggins, A., Gross, D., & Hahn, E. J. (2020). Home Radon Testing in Rural Appalachia. *The Journal of Rural Health*.
- Vogeltanz-Holm, N., & Schwartz, G. G. (2018). Radon and lung cancer: What does the public really know? *Journal of environmental radioactivity*, 192, 26-31.
- Wang, Y., Ju, C., Stark, A. D., & Teresi, N. (2000). Radon awareness, testing, and remediation survey among New York State residents. *Health physics*, 78(6), 641-647.
- Webster, J. W. (1990). Radon Contamination of Residences in a City Built Upon a Karst Landscape Bowling Green, Warren County, Kentucky.
- Yarmoshenko, I., Vasilyev, A., Malinovsky, G., Bossew, P., Žunić, Z. S., Onischenko, A., & Zhukovsky, M. (2016). Variance of indoor radon concentration: Major influencing factors. *Science of The Total Environm-160. (Science of the Total Environment)*.

Studying the Properties of Polymeric Composites of Metal Hydrides and Carbon Particles for Hydrogen Storage

¹Fenil J. Desai

¹Md. Nizam Uddin

¹Muhammad Mustafizur Rahman

¹Ramazan Asmatulu

¹Department of Mechanical Engineering, Wichita State University, Wichita, Kansas, USA

fjdesai@shockers.wichita.edu; enqrnizam02@gmail.com; muhammad.rahman@wichita.edu;

ramzan.asmatulu@wichita.edu

Abstract

Metal hydrides are promising hydrogen storage materials widely studied and accepted by many authorities, but still, it has not reached the set goal. In this work, polymer-based carbon particles along with metal hydride materials are proposed as a storage medium for hydrogen. Metal hydride particles were integrated into a polymeric matrix with carbon particles to improve the thermal stability and hydrogen storage capacity. Some physical properties, morphological effects, and thermal conductivity values of the polymeric composite were investigated using X-ray diffraction (XRD), Fourier Transform Infrared Spectroscopy (FTIR), and scanning electron microscopy (SEM). The test results indicated that metal hydrides and carbon particles were well integrated into the polymeric structure, which could drastically affect the hydrogen storage capacity of the polymeric composites for applications in the transportation industry.

1. Introduction

Energy consumption in the transportation industry is increasing, which is also causing the increased threat of greenhouse gas emissions (F. Desai, Atayo, et al., 2020; F. Desai, Seyedhassantehrani, et al., 2020). To reduce the usage of fossil fuels, the transportation industry needs to adopt new approaches to harness the inexhaustible sources of energy (Sunku Prasad et al., 2019). Hydrogen is a new generation of energy carrier which has the potential to be used as a clean, safe, reliable, and versatile fuel, and as a renewable and sustainable source of energy for transportation and other industries (F. Desai et al., 2021; F. J. Desai et al., 2013). Hydrogen can be stored in three ways: high pressure, cryogenic temperature, and materials-based storage. Among these, material-based hydrogen storage can be achieved at close to ambient pressure and temperatures conditions in sorbents, chemical hydrogen storage materials, and metal hydrides via a thermochemical reaction process. It works like a sponge where hydrogen (H_2) is absorbed in metal hydride (M) during an endothermic reaction and desorbed during an exothermic reaction process as shown in eq. 1. The metal hydrides can store hydrogen over a

wide operating range from low to high temperature but still do not meet the goal set by the US Department of Energy as shown in Table 1.



Table 1 Technical target for onboard hydrogen storage for vehicles [6]

Storage	Units	2020	2025	Ultimate
Usable, specific energy from H ₂ (net useful energy/max system mass)	kWh/kg (kg H ₂ /kg system)	1.5	1.8	2.2
Operating ambient temperature	°C	-40/85	-40/85	-40/85
Min/max delivery temperature	°C	-40/85	-40/85	-40/85
Min/max delivery pressure from storage system	bar	5/12	5/12	5/12

Metal hydrides such as intermetallic compounds (AB₅-, AB-, AB₂-, A₂B-, AB₃- type alloys), complex hydrides (alanates, borohydrides, nitrides), and magnesium-based alloys are the most researched materials and have the potential to meet the hydrogen storage density and thermodynamics requirements set by DOE (Rusman and Dahari, 2016). Intermetallic alloys have a wide range of applications in nickel hydride batteries, hydrogen storage, and purification systems, heat pumps and as hydrogen sensors and catalysts (Falahati and Barz, 2013). These intermetallic compounds are attractive and promising because of their excellent hydrogen-absorbing properties, and moderate pressure and temperature requirements (Rusman and Dahari, 2016). AB₅- type intermetallic alloys have gained wide attention in the recent decade as it operates on a low equilibrium pressure, and possess fast thermodynamics characteristics, high hydrogen storage capacities, and good cyclic stability (Park et al., 2009). LaNi₅- based AB₅ alloys have been one the most researched hydrides due to their faster reaction kinetics and hydrogen storage capability at ambient conditions (Chibani et al., 2020; Prigent et al., 2012). Also, various studies have reported 11 to 45.6 kJ/mol H₂ activation energy for absorption (Koh et al., 1989; Nahm et al., 1992). Therefore, LaNi₅ is an attractive and meaningful hydrogen storage material for sustainable development. In this work, the chemical and physical properties of the polymer and carbon-based materials used as catalysts together with metal hydride have been investigated for the first time. The surface functional groups were created using polymer to prevent agglomeration of the LaNi₅ nanoparticles (Uddin, Desai, & Asmatulu, 2020; Uddin, Desai, Rahman, et al., 2020). The carbon-based materials, graphene, and MWCNTs have been used to improve the thermal conductivity and hence the reaction kinetics.

2. Experiments

2.1 Materials

The raw LaNi₅ (<100 μm, 99 wt.% in purity) was purchased from Whole Win (Beijing) Materials Sci. & Tech. Co. Ltd. PMMA (molecular weight, M_w 120 000 g mol⁻¹, Glass transition temperature, T_g 99°C)

and PVDF (M_w 120 000 g mol⁻¹, T_g 99°C) were acquired from Sigma-Aldrich (St. Louis, MO, USA). Dimethylformamide (DMF) (M_w 73.09 g/mol, boiling point 153 °C) was used as a solvent and purchased from Sigma-Aldrich (St. Louis, MO, USA).

2.2 Preparation of encapsulated LaNi₅ composites

The encapsulated composites were prepared in an inert atmosphere. The samples were prepared by a solvent-based method as shown in Fig. 1. The polymer solution was initially prepared by dissolving PMMA/PVDF in Dimethylformamide (DMF) solvent with magnetic stirring. The polymeric matrix was then transferred along with carbon particles, graphene/MWCNTs, for sonication using a probe sonicator for 10 min. The LaNi₅ powder was then added to the polymer-carbon composite and sonicated again for 10 minutes. The composite mixture of LaNi₅-polymer-carbon particles was mixed using a shear mixture for 4 h. The homogenous composite mixture was transferred to form a pallet via the heat casting method in the oven at 150°C for 24 h. The ratio of LaNi₅-polymer-carbon particles was varied as shown in Table 2 to study the effect of two different polymers (PMMA/PVDF) and carbon particles (graphene/MWCNTs) on hydrogen storage properties of LaNi₅. As an example, four samples of LaNi₅ using a different concentration of PVDF (1, 2, 4, and 8 wt. %) and graphene (1, 2, 4, and 8 wt. %) were prepared. The encapsulated polymeric metal hydride composite will be called EPMC from here on as shown in Table 3.

Table 2 Ratio of polymer/graphene with LaNi₅

	PVDF wt.%	PMMA wt.%	Graphene wt.%	MWCNTs wt.%
LaNi₅	1, 2, 4, 8		1, 2, 4, 8	
	1, 2, 4, 8			1, 2, 4, 8
		1, 2, 4, 8	1, 2, 4, 8	
		1, 2, 4, 8		1, 2, 4, 8

Table 3 Composite samples reported in this paper

	PVDF wt.%	PMMA wt.%	Graphene wt.%	MWCNTs wt.%
Sample 3 (EPMC-1)	4		4	
Sample 7 (EPMC- 2)	4			4
Sample 11 (EPMC- 3)		4	4	
Sample 15 (EPMC-4)		4		4

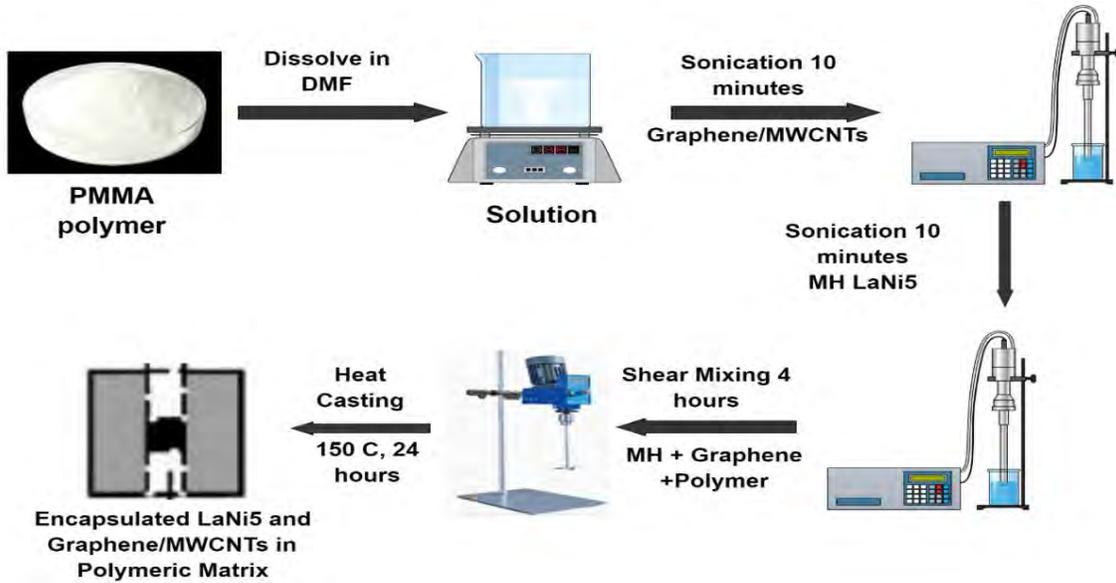


Figure 1 illustration of sample preparation of encapsulated EPMC

2.3 Characterizations

The structural phase constitution of the composite was explored by X-ray diffraction (XRD) in a PANalytical Empyrean Diffractometer using CuK α radiation within 2θ range of 10-80°. The qualitative estimations of surface functional groups of composites were studied by Fourier transform infrared spectroscopy (FTIR, Thermo Scientific™ Nicolet™ iN10 infrared microscope) over a range of 500-4000 cm⁻¹. The morphology of the composite was characterized by scanning electron microscopy (SEM) (FEI Nova Nano SEM 450).

3. Results and Discussion

3.1 Physical phase structure

3.1.1 XRD:

The XRD pattern of the composite with carbon and polymer-based materials on LaNi₅ is shown in Fig. 2 in the 2θ range of 20-80° to confirm structural refinement. The XRD graphs are reported as a function of different concentrations (S4, S8, S12, S16) of polymer and carbon particles in LaNi₅. The XRD patterns of all the samples (S4 to S16) directs a similar structure and good reliability of the results. The peak near 20° unveils the presence of the amorphous polymer. The XRD spectra of composites show main peak $2\theta = 30^\circ, 35.77^\circ, 42.59^\circ, 47.71^\circ, 63.47^\circ, 64.40^\circ, \text{ and } 68.89^\circ$ can be assigned to (011), (110), (111), (002), (202), (300) and (031) respectively to hexagonal LaNi₅.

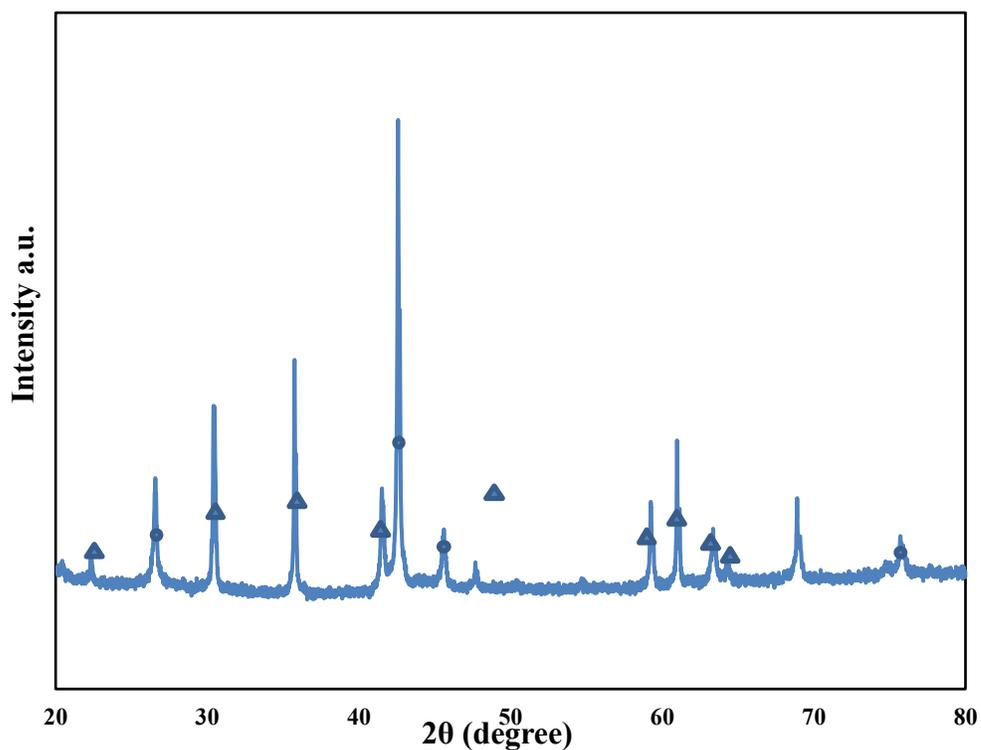


Figure 2 XRD for EPMC-1

3.1.2 FTIR:

The FTIR spectra of pure LaNi_5 , samples 3, 7, 11, and 15 of composites are given in Fig. 3 to identify the bonding of carbon and polymer with LaNi_5 . The FTIR spectra of composite show various surface functional groups which are close to individual LaNi_5 , polymer, and carbon-based materials. The LaNi_5 surface measured to be composed of Ni and La_2O_3 at 1910 and 1463 cm^{-1} respectively and it is in good agreement with previous research work by Sakaguchi et al., 1995. Therefore, the FTIR spectra evidently show that the addition of polymeric and carbon-based materials on LaNi_5 have created surface functional groups as expected. The addition of polymer (PPMA/PVDF) provides the porous, rough surface on the metal hydride surface which enhances the hydrophobicity and thus protects metal atoms from air exposure. Hence, it provides cyclic stability.

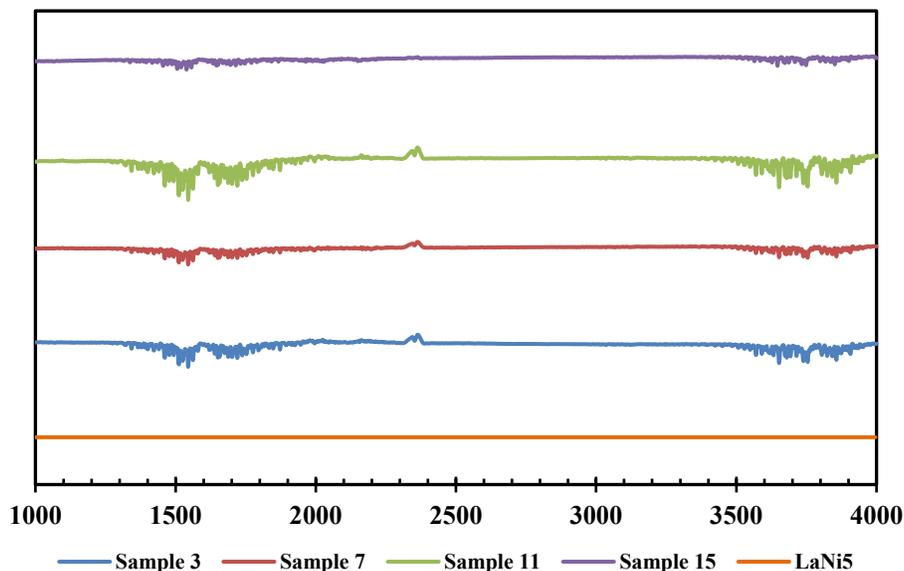


Figure 3 Infrared spectra of (a) EPMC-1 (b) EPMC-2 (c) EPMC-3 (d) EPMC-4 (e) LaNi₅

3.1.3 SEM:

The morphology of LaNi₅-polymer-carbon composite examined with FE-SEM is reported in Fig. 4. The Fe-SEM images of PVDF/LaNi₅/Graphene, PVDF/LaNi₅/MWCNTs, PMMA/LaNi₅/Graphene, PMMA/LaNi₅/MWCNTs indicated as (a, b), (c, d), (e, f) and (g, h) respectively. The SEM images reveal LaNi₅ particles are clearly embedded in randomly oriented porous polymeric and carbon matrix homogeneously. The hydrogen storage characteristics, especially kinetics in the composite will be hindered because of better dispersion of polymer and carbon particles in composites. The unique structure of MWCNTs (tubular structure) is expected to exhibit the most prominent influence as a catalyst for LaNi₅ and will help to enhance the hydrogen de-/rehydrogenation reactions. The SEM images clearly show the higher specific surface area created due to the addition of the polymeric and carbon particles. These will accommodate the heat generated during the endothermic absorption process; hence it provides cyclic stability for the metal hydride particles and improves sorption processes (absorption/desorption).

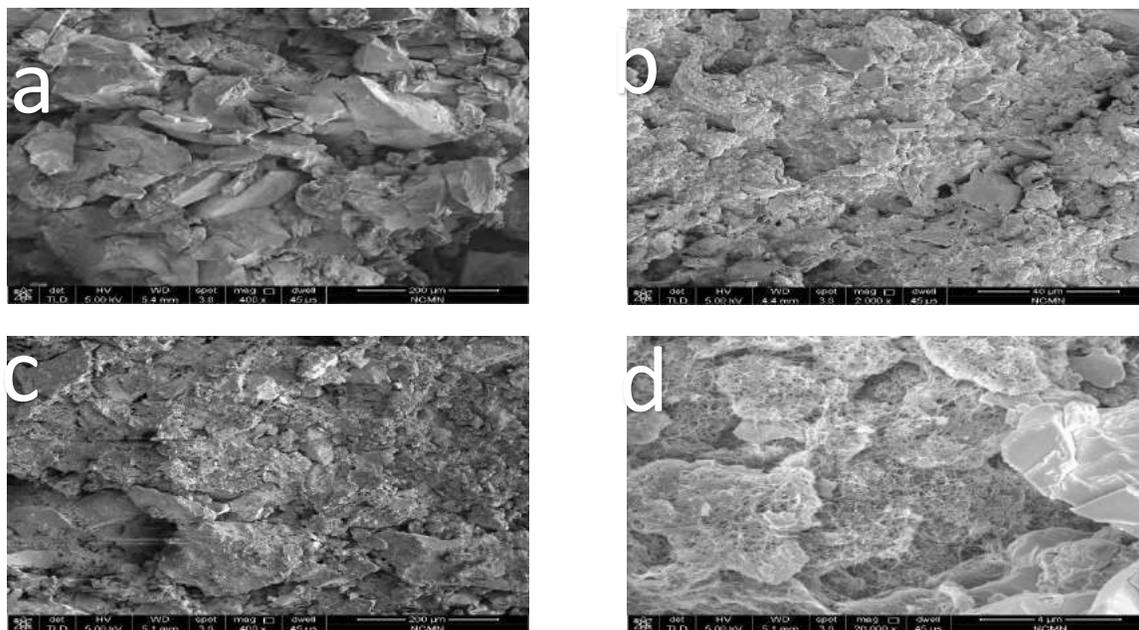


Figure 4 SEM images of (a) EPMC-1 (b) EPMC-2 (c) EPMC-3 (d) EPMC-4

4. Conclusion

In this study, innovative LaNi_5 /polymer/carbon composite was successfully prepared by the solvent-based method. The composite was developed to reach the department of energy goal for onboard hydrogen storage for fuel cell application. The addition of polymer (PPMA/PVDF) provides the porous, rough surface on LaNi_5 which increases the hydrophobicity and thus protects metal atoms from air exposure. The higher specific surface area of the polymer and carbon particles accommodates the heat generated during the endothermic absorption process and improves the cyclic stability during sorption processes (absorption/desorption).

5. References

- Chibani, A., Merouani, S., Bougriou, C., & Hamadi, L. (2020). Heat and mass transfer during the storage of hydrogen in LaNi_5 -based metal hydride: 2D simulation results for a large-scale, multi-pipes fixed-bed reactor. *International Journal of Heat and Mass Transfer*, 147, 118939.
<https://doi.org/10.1016/j.ijheatmasstransfer.2019.118939>
- Desai, F., Atayo, A., Prasad, J. S., Muthukumar, P., Rahman, M., & Asmatulu, E. (2020). Experimental Studies on Endothermic Reversible Reaction of Salts for Cooling. *Heat Transfer Engineering*, 1–13.
<https://doi.org/10.1080/01457632.2020.1777002>
- Desai, F. J., Kulashreshtha, D. B., & Sheth, N. R. (2013). Analytical and Experimental Analysis of Hydrogen Combustion. 23rd National Conference on I. C. Engine and Combustion (NCICEC 2013), Sardar Vallabhbhai National Institute of Technology, Surat, India.

- Desai, F., Seyedhassantehrani, N., Shagar, M., Gu, S., & Asmatulu, R. (2020). Preparation and characterization of KOH-treated electrospun nanofiber mats as electrodes for iron-based redox-flow batteries. *Journal of Energy Storage*, 27. <https://doi.org/10.1016/j.est.2019.101053>
- Desai, F., Sunku Prasad, J., Muthukumar, P., & Rahman, M. M. (2021). Thermochemical energy storage system for cooling and process heating applications: A review. *Energy Conversion and Management*, 229(July 2020), 113617. <https://doi.org/10.1016/j.enconman.2020.113617>
- DOE Technical Targets for Onboard Hydrogen Storage for Light-Duty Vehicles. (2018). Hydrogen and Fuel Cell Technologies Office. <https://www.energy.gov/eere/fuelcells/doe-technical-targets-onboard-hydrogen-storage-light-duty-vehicles>
- Falahati, H., & Barz, D. P. J. (2013). Evaluation of hydrogen sorption models for AB₅-type metal alloys by employing a gravimetric technique. *International Journal of Hydrogen Energy*, 38(21), 8838–8851. <https://doi.org/10.1016/j.ijhydene.2013.04.148>
- Koh, J. T., Goudy, A. J., Huang, P., & Zhou, G. (1989). A comparison of the hydriding and dehydriding kinetics of LaNi₅ hydride. *Journal of the Less-Common Metals*, 153(1), 89–100. [https://doi.org/10.1016/0022-5088\(89\)90536-5](https://doi.org/10.1016/0022-5088(89)90536-5)
- Nahm, K., Kim, W., Hong, S., & Lee, W. (1992). The reaction kinetics of hydrogen storage in LaNi₅. *International Journal of Hydrogen Energy*, 17(5), 333–338. [https://doi.org/10.1016/0360-3199\(92\)90169-W](https://doi.org/10.1016/0360-3199(92)90169-W)
- Park, I. S., Kim, J. K., Kim, K. J., Zhang, J., Park, C., & Gawlik, K. (2009). Investigation of coupled AB₅ type high-power metal hydride reactors. *International Journal of Hydrogen Energy*, 34(14), 5770–5777. <https://doi.org/10.1016/j.ijhydene.2009.05.012>
- Prigent, J., Joubert, J. M., & Gupta, M. (2012). Modification of the hydrogenation properties of LaNi₅ upon Ni substitution by Rh, Ir, Pt, or Au. *Journal of Alloys and Compounds*, 511(1), 95–100. <https://doi.org/10.1016/j.jallcom.2011.08.094>
- Rusman, N. A. A., & Dahari, M. (2016). A review on the current progress of metal hydrides material for solid-state hydrogen storage applications. *International Journal of Hydrogen Energy*, 41(28), 12108–12126. <https://doi.org/10.1016/j.ijhydene.2016.05.244>
- Sakaguchi, H., Tsujimoto, T., & Adachi, G. (1995). The confinement of hydrogen in LaNi₅ by poisoning of the hydride surface. *Journal of Alloys and Compounds*, 223(1), 122–126. [https://doi.org/10.1016/0925-8388\(94\)01489-2](https://doi.org/10.1016/0925-8388(94)01489-2)
- Sunku Prasad, J., Muthukumar, P., Desai, F., Basu, D. N., & Rahman, M. M. (2019). A critical review of high-temperature reversible thermochemical energy storage systems. *Applied Energy*, 254. <https://doi.org/10.1016/j.apenergy.2019.113733>

Uddin, M. N., Desai, F. J., & Asmatulu, E. (2020). Biomimetic electrospun nanocomposite fibers from recycled polystyrene foams exhibiting superhydrophobicity. *Energy, Ecology and Environment*, 5(1), 1–11. <https://doi.org/10.1007/s40974-019-00140-7>

Uddin, M. N., Desai, F. J., Rahman, M. M., & Asmatulu, R. (2020). Highly Efficient Fog Harvester of Electrospun Permanent Superhydrophobic-Hydrophilic Polymeric Nanocomposite Fiber Mats. *Nanoscale Advances*. <https://doi.org/10.1039/d0na00529k>

3D Printing of Portland Cement Using a Binder Jetting System

Abdelhakim A. Al Turk¹

Gamal S. Weheba¹

¹Department of Industrial, Systems, and Manufacturing Engineering

Wichita State University

axalturk@shockers.wichita.edu

Abstract

This paper presents proof of the concept that Portland cement can be used as a material on a commercial binder jetting system. Sixteen cylindrical specimens were printed using different layer thicknesses and cured under different conditions. Both printing and curing followed a statistically designed experiment covering all possible combinations of the factors. The average values of the compressive strength and deviation from nominal height were used as response variables. Statistical analysis of the results supported the concept and suggested settings that would allow for maximum strength and minimum deviation from the nominal height. The concept appears to have several potential applications in restoration projects and the construction industry.

1 Introduction

Additive Manufacturing (AM) in recent years has become a common method for creating detailed parts or models from 3D CAD data. It relies on software to process the CAD data into thin layers that are used to manufacture the object in successive layers. All AM techniques have been successfully adopted for the manufacturing of functional end-use products. Among these techniques are Stereolithography, Binder Jetting, Material Jetting, Powder Bed Fusion, Material Extrusion, Sheet Lamination, and Directed Energy Deposition. In the past few years, AM technology has expanded into construction projects and is now leading to major breakthroughs in the industry. According to Al Turk and Weheba (2020), the constant growth of AM led to a wide range of applications in the construction industry. Design modeling, automated construction of full structures or structural elements, restoration of historical buildings, and repair of existing structures are examples of current applications. Significant benefits of applying the AM in construction projects have been reported. Construction is becoming faster and more accurate with the capability of producing complex geometries. AM has been reported to reduce labor costs, risk of injury, waste, the need for formworks.

2 Literature Review

Applications of AM in the construction industry are dominated by Contour Crafting (CC) and Binder Jetting (BJ). As noted by Khoshnevis (2003), Contour Crafting combines an extrusion system with filling process to build concrete objects. As the material is extruded through an extrusion nozzle, the top and side trowels work on the material to create a smooth outer edge of the rim. The filling process consists of pouring material to fill the area constructed by the extrusion nozzle. To achieve this, the system would need to use several different deposition heads to produce the structures required and manipulators to install fixtures. As an example, in 2015, WinSun company

(www.winsun3d.com) constructed a 3D printed structure. This project consisted of constructing a five-story apartment building. The total area of this project was about 1100 m², making it the tallest 3D printed structure in China.

Binder Jetting Technology was developed at the Massachusetts Institute of Technology (MIT) in the early 1990s. Originally, it was called 3D Printing (3DP) in which binder material is printed onto a powder bed to form one layer of the object. Once the first layer is constructed, a fresh layer of powder is spread over the bed, smoothed by a roller, and the binder is applied. These steps are repeated until the object is constructed. Early applications of the binder jetting process in construction were reported by Xia et al (2019). They used a Z-Corp 450 3DP system, with the standard Zb 63 binder, to study the effect of different powder parameters and binder droplet penetration behavior on the printability of geopolymers. They also investigated the effect of fly ash content on the compressive strength of the printed specimens. The results indicated that geopolymers can be used in the commercially available binder jetting systems. Another application of Binder Jetting is the D-shape project (dshape.wordpress.com) in the UK where a binder is ejected over a bed of sand. More powder is added, and subsequent layers are built up. The D-shape system has a build envelope of 212 ft³ (6 m³) with layer thicknesses ranging from 0.16" to 0.24" (4 to 6 mm).

3 Printing of Portland Cement

The Zprinter 450 is a binder jetting 3D printer that was originally offered by the Z-Corporation in 1994. The company was acquired by 3D Systems in 2012. The system utilizes calcium sulfate hemihydrate powder (ZP 151) and a water-soluble polymer binder (Zb 63). The system has full-color printing capability.

In this research, the standard powder ZP151 was replaced with Quick-Setting Cement material (Quikrete 1240). This type of commercial cement material is a Portland cement formulated for making structural repairs on vertical and horizontal surfaces. Quikrete is a blend of gray Portland cement, additional cementitious additives, and fine silica sand. The sand contains particles ranging from a maximum size of US Sieve No. 8 (making up approximately 3% of the product) through particles small enough to pass US Sieve No. 200 (75 micrometers). Typical physical properties of the Quikrete obtain from the material data sheet are shown in Table 1.

Table 1. Physical Properties of Quikrete 1240

Initial set	5 - 10 minutes
Final set	10 - 20 minutes
Compressive strength, ASTM C109	
24 hours	3000 psi (20.7 MPa)
7 days	5000 psi (34.5 MPa)
28 days	6000 psi (41.3 MPa)
Post-freeze/thaw Compressive strength	6430 psi (44.3 MPa)
Scaling resistance, ASTM C672	Excellent

The same binder recommended by 3D Systems is used. The Zb 63 binder is a commercial clear solution with a viscosity similar to pure water with 2-Pyrrolidone.

4 Experimental Factors

This research was focused on quantifying the effect of three factors. The curing time, curing solution, and layer thickness. Curing time may affect the properties of concrete such as durability, strength, wear resistance, volume stability, and resistance to freezing. According to the ASTM C 39 and C 192 standards, concrete cylinders must be cured for 28 days to ensure the best development

of compressive strength. According to ACI 318-19, the compressive strength of concrete will increase with time, gaining 70% of its strength after 7 days, 90% after 14 days, and 99% after 28 days. For these reasons, two curing durations were considered (14 and 28 days).

Also, two curing solutions were considered. One solution was the regular tap water and the other was an alkaline solution as recommended by Xia et al (2019). The alkaline solution used is composed of N Grade sodium silicate solution with a SiO₂/Na₂O ratio of 3.22 (71.4% w/w) and 8.0 M NaOH solution (28.6% w/w).

The ZPrinter 450 allows users to select four levels of the layer thickness (0.088, 0.100, 0.112, 0.125 mm). These values are used to control the platform downward movements following the construction of each layer. The smaller the layer thickness, the higher the amount of binder applied to connect the layers. As such, this factor was thought to have an important effect on the strength of the printed specimens. The three factors considered and their levels are described in Table 2. All other factors were held constant.

Table 2. Design Factors

Factor	Type	Low level (-)	High level (+)
A- Curing time (Days)	Numeric	14	28
B- Curing Solution (Type)	Categorical	Water	Solution
C- Layer Thickness (mm)	Numeric	0.088	0.125

5 Response Variables

The effects of the experimental factors were quantified by using two response variables. The height of the printed specimens and the average compressive strength. The height was measured using a Vernier Caliper and recorded in terms of the deviation from the nominal height (DNH). The Vernier scale allows much more precise readings to be taken, usually to the nearest 0.001" (0.02mm). Compressive strength was measured by utilizing the rebound hammer shown in Figure 1.



Figure 1. Concrete Test Hammer HT-20

This is a portable non-destructive tester used for evaluating the in-situ compressive strength of concrete structures. According to Carino (1994), the rebound hammer tester's underlying principle is that the rebound of an elastic mass depends on the hardness of the surface against which it strikes. The rebound hammer HT-20 with a measuring range of 1- 25 Mpa has been used. The test hammer was calibrated using a steel anvil with HRC 60. Repeatability and reproducibility of the tester were evaluated by collecting repeated measurements on a number of parts using different operators. In this research, 10 concrete tiles were tested 3 times each by 3 different operators. The data were collected as rebound numbers (RN) then converted to strength measurements (Mpa). Data were analyzed using STATGRAPHICS XVII (Statgraphics, Technologies, Inc. Version 18.1.13; 2021) and the results are shown in Table 3.

Table 3. Repeatability and Reproducibility Analysis

Measurement Unit	Estimated Sigma	Percent Total Variation	Estimated Variance	Percent Contribution	Percent of R&R
Repeatability	7.16318	84.8535	51.3111	72.0012	80.67
Reproducibility	3.50602	41.5316	12.2922	17.2487	19.33
R & R	7.97517	94.4722	63.6033	89.25	100.00
Parts	2.76783	32.7872	7.66091	10.75	
Total Variation	8.44181	100.0	71.2642		

Based on these results, it was decided to collect repeated measurements on each specimen to reduce the error variance. So, for each cement specimen, an average of five measurements was used to report the resulting strength. Also, to reduce the reproducibility component of the variance, one operator only was allowed to collect the RN measurements.

6 Experimental Design

As reported by Camacho et al (2018), an essential advantage of 3D printing is its ability to produce complex geometries that were hard or even impossible to construct with conventional construction methods. Large-scale 3D printing of end-use structures allows architects to build complex passageways, undercuts, and other interior and exterior designs. They are allowed to rethink their sketches and forms without affecting the productivity of the construction process. Architects do not have to worry about the constructability of each part anymore and can now focus on their design and functionality.

A replicated two-level factorial design was used to quantify the effect of selected factors on the strength and dimensional accuracy of the printed specimens. The procedures recommended by Coleman and Montgomery (1993) were followed. The data were analyzed using the Design Expert V13 software (Stat-Ease, Inc. Design-Expert. Version 13.0.2.0, 2021). The experiment required printing a total of 16 specimens (2X8) in a completely randomized order. All specimens had the same dimensions, 1.38" for the Height and 2.36" for the diameter. The design matrix shown in Table 4 indicates the run order for each of the factor level combinations and recorded values of the two response variables.

Table 4. Design Matrix

Std Order	Run Order	Curing Time (A)	Curing Solution (B)	Layer Thickness (C)	Compressive Strength (MPa)	DNH (mm)
8	1	28	Mix	0.088	14.3	0.0486
12	2	28	Water	0.125	1.0	0.0431
10	3	14	Water	0.125	0.3	0.0351
16	4	28	Mix	0.125	11.0	0.0460
11	5	28	Water	0.125	1.0	0.0446
13	6	14	Mix	0.125	9.4	0.0316
1	7	14	Water	0.088	0.5	0.0371
7	8	28	Mix	0.088	15.0	0.0706
14	9	14	Mix	0.125	8.6	0.0382
4	10	28	Water	0.088	1.1	0.0421

2	11	14	Water	0.088	0.5	0.0246
5	12	14	Mix	0.088	11.6	0.0780
6	13	14	Mix	0.088	13.2	0.0436
3	14	28	Water	0.088	1.2	0.0366
9	15	14	Water	0.125	0.3	0.0466
15	16	28	Mix	0.125	10.2	0.0416

6.1 Analysis of the Compressive Strength

Initial investigation of the average compressive strength indicated the need for applying a square root transformation to stabilize the error variance. The analysis of variance (ANOVA) for the transformed response is shown in Table 5. As shown, the curing time (A), curing solution (B), thickness (C), and the interaction involving both the curing solution and layer thickness (BC) have significant effects.

Table 5. ANOVA Results

Source	Sum of Squares	df	Mean Squares	F- value	P-value
Curing Time (A)	0.49	1	0.49	72.591	0.0001
Curing Solution (B)	26.50	1	26.50	3915.2	0.0001
Layer Thickness (C)	0.45	1	0.45	66.798	0.0001
Interaction (BC)	0.18	1	0.18	26.062	0.0003
Residual	0.07	11	0.01		
Lack of Fit	0.03	3	0.01	1.4336	0.3032
Pure Error	0.05	8	0.01		
Corrected Total	27.69	15			

The lack of fit is not significant indicating that all other interactions can be ignored. Diagnostic plots of the residuals indicated no violations of the underlying assumptions of the ANOVA procedure.

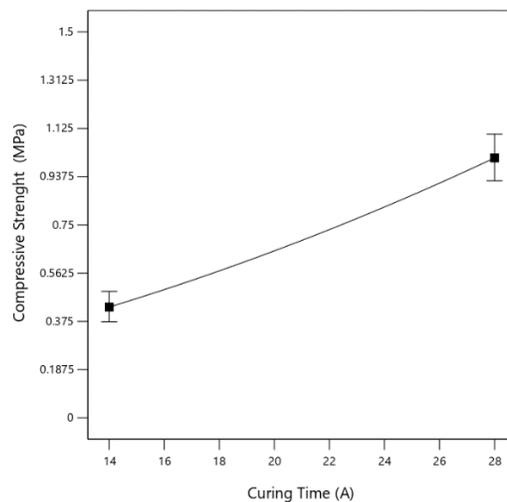


Figure 2. Effect of the Curing Time (A)

Figure 2 depicts the effect of curing time on the average compressive strength measured in MPa. As shown, the higher the curing duration the higher the average compressive strength. Using the high level of curing duration (28 days) resulted in a 22.9% increase in the average strength of the printed specimens.

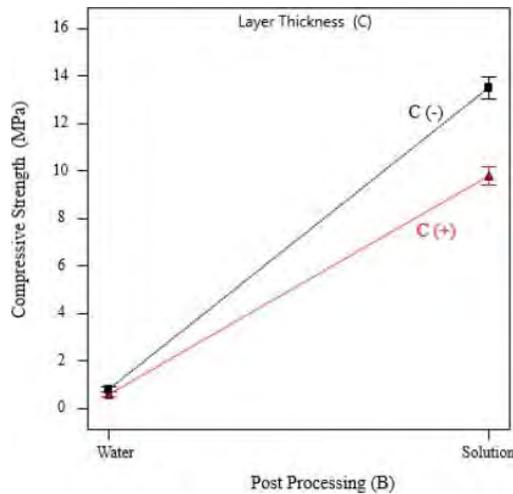


Figure 3. BC Interaction Plot

An examination of the interaction plot in Figure 3 indicates that when the alkaline solution is used, changes in the layer thickness appear to have a significant effect on the average compressive strength. The maximum strength is obtained when the specimens were printed using the low level of layer thickness (0.088 mm) and cured in the alkaline solution.

6.2 Analysis of the Height

A similar procedure was followed in analyzing changes in the specimens' height in terms of deviations from the nominal height (DNH). All reported values of the average DNH were positive indicating that curing (post-processing) increased height measurements. The ANOVA results are shown in Table 6. As shown, the effects of the curing solution (B), the layer thickness (C), and their interaction appear significant. All diagnostic plots of the residuals supported the underlying assumptions of the ANOVA.

Table 6. ANOVA Results

Source	Sum of Squares	df	Mean Squares	F- value	P-value
Model	0.0015	3	0.0005	4.93	0.0186
B-Curing Solution	0.0005	1	0.0005	4.94	0.0463
C-Layer Thickness	0.002	1	0.002	1.87	0.1967
BC Interaction	0.0008	1	0.0008	7.98	0.0153
Residual	0.0012	12	0.0001		
Lack of Fit	0.0002	4	0.000	0.3157	0.8598
Pure Error	0.0010	8	0.0001		
Corrected Total	0.0027	15			

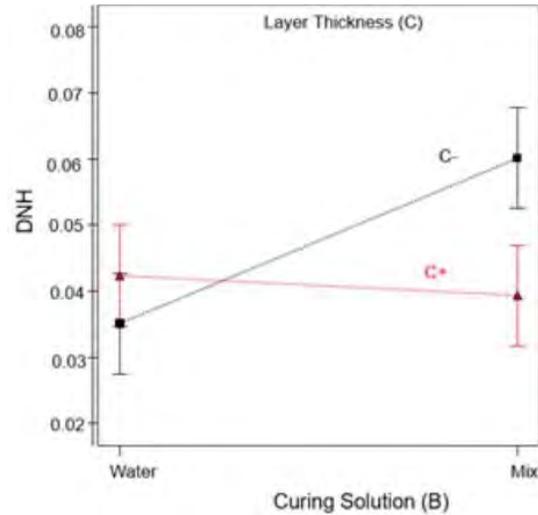


Figure 4. BC Interaction Plot

An examination of the interaction plot shown in Figure 4, reveals that changes in the layer thickness have a significant effect on the average deviation when the alkaline solution was used. While all measurements indicated an increase in the average height, the maximum deviation is observed when specimens were constructed using the low level of layer thickness (0.088) and cured using the alkaline solution. However, changes in the curing solution appear to have no significant effect on the average deviation when the specimens were constructed using the high level of layer thickness (0.125 mm)

7 Conclusions

This research was aimed at providing proof of the concept that the quick-setting cement can be used as a material on the ZPrinter 450. Sixteen specimens were constructed in a completely randomized order to protect against nuisance factors. Repeated measurements were collected by the researcher on the strength and height of each specimen to control the measurement errors. Statistical analysis of these results indicated that the maximum strength 1958 Psi (13.5 MPa) was obtained when the specimens were printed using the low level of layer thickness 0.0035" (0.088 mm) and cured in the alkaline solution. The overlay plot shown in Figure 5 was constructed by limiting the deviation from the nominal (DNH) to 0.05" (1.27mm) and setting the minimum strength at 1740.45 Psi (12 MPa). The plot highlights the recommended area for selecting levels of the layer thickness and curing time when the alkaline solution is used.

These results support the concept of printing cement objects using standard binder jetting systems. Cement objects can be 3D printed on the ZPrinter 450 subject to limitations of size 8" x 10" x 8" (203.2x254x203.2 mm) with a build speed of 2 to 4 layers per minute. The concept has several potential applications in the construction industry. It may encourage architects and designers to try new innovative ideas away from the restrictions of traditional construction methods. It may eliminate the need for traditional methods in restoration projects, where small tiles and sculptures need to be replaced. The method may be used to 3D print customized tiles with letter scripts like the house number plaque or the decorative calligraphy tiles found in Churches and Mosques. This may eliminate the need for carving, as objects can be designed using CAD software and printed directly.

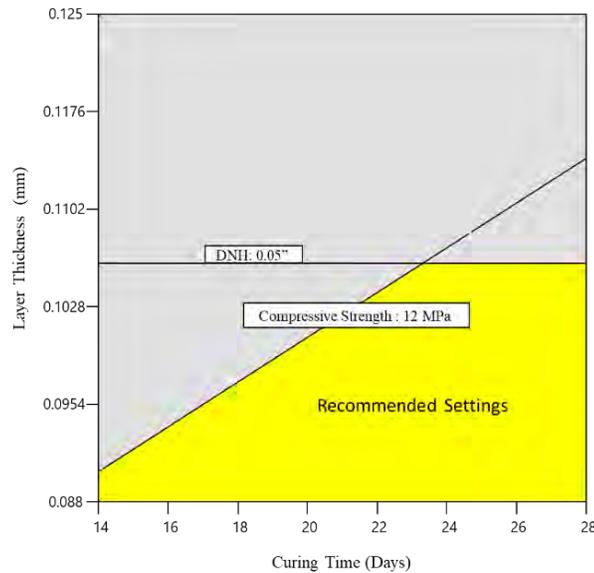


Figure 5. Overlay Plot with Recommended Settings with The Alkaline Solution

In addition, concrete printing lends itself to the construction of snow guards and fences used to prevent ice buildup on high-rise buildings. Snow guards customized to fit the building façades can be constructed to create friction between the roof surface and the snow to keep the accumulated ice on the roof. While cement printing has potential in several applications, there is a need for continuing research in the areas of materials, mechanical properties, and process control. Processes need to be reexamined to allow for the reinforcement of concrete. Mechanisms for integrating different reinforcement materials, patterns, and shapes during component printing need to be considered. These authors are considering different mechanisms for integrating reinforcement materials during component printing.

8 References

- ACI Committee, & International Organization for Standardization. (2008). Building code requirements for structural concrete (ACI 318-08) and commentary. American Concrete Institute.
- Al Turk, A. A., & Weheba, G. S. (2020). Applications of Additive Manufacturing in Construction Industry: A Literature Review. *Journal of Management & Engineering Integration*, 13(1), 32-39
- ASTM C39 / C39M-21, Standard Test Method for Compressive Strength of Cylindrical Concrete Specimens, ASTM International, West Conshohocken, PA.
- ASTM C192 / C192M-19, Standard Practice for Making and Curing Concrete Test Specimens in the Laboratory, ASTM International, West Conshohocken, PA, 2019, www.astm.org
- Camacho, D. D., Clayton, P., O'Brien, W. J., Seepersad, C., Juenger, M., Ferron, R., & Salamone, S. (2018). Applications of additive manufacturing in the construction industry—A forward-looking review. *Automation in construction*, 89, 110-119.
- Carino, N. J., (1994). Nondestructive Testing of Concrete: History and Challenges, ACI SP-44, Concrete Technology, Past, Present and Future, Ed. Mehta P. K, American Concrete Institute, 623-678.
- Coleman, D. E., & Montgomery, D. C. (1993). A systematic approach to planning for a designed industrial experiment. *Technometrics*, 35(1), 1-12.

D-Shape Enterprises L.L.C., Online <https://dshape.wordpress.com/> (accessed 25.01.20).

Lim, S., Buswell, R., Le, T., Austin, S., Gibb, A., and Thorpe, T. (2012). Developments in construction-scale additive manufacturing processes. *Automation in construction*, 21, pp.262--268.

Khoshnevis B., (2003). "Toward total automation of on-site construction an integrated approach based on contour crafting." Los Angeles: University of Southern California.

Quick-Setting Cement, Product No. 1240-11, -20, -50 Retrieved from: https://www.quikrete.com/pdfs/data_sheet-quick-setting%20cement%201240.pdf (accessed 15.02.21).

Samia, H., & Mohamed Nacer, G. (2012). Application of the combined method for evaluating the compressive strength of concrete on site. *Open Journal of Civil Engineering*, 2012.

WinSun China, online: <http://www.winsun3d.com/> (accessed 29.01.20).

Xia M, Nematollahi B, Sanjayan J., (2019). "Printability, accuracy, and strength of geopolymer made using powder-based 3D p

**SUMMER
2021**



AIEMS

AIEMS is committed to furthering the integration between business management, engineering, & industry. Our objective is to promote research, collaboration, & practice in these multidisciplinary areas. AIEMS seeks to encourage local, national, & international communication & networking via conferences & publications open to those in both academia & industry. We strive to advance professional interaction & lifelong learning via human & technological resources, and to influence and promote the recruitment and retention of young faculty and industrialists.

**AIEMS
VOL 14
NO 1**