



NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

**ANALYZING THE EFFECTS OF URBAN COMBAT ON
DAILY CASUALTY RATES**

by

Hakan Yazilitas

June 2004

Thesis Advisor:
Second Reader:

Samuel E. Buttrey
Saverio Manago

Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.				
1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE June 2004	3. REPORT TYPE AND DATES COVERED Master's Thesis		
4. TITLE AND SUBTITLE: Analyzing the Effects of Urban Combat on Daily Casualty Rates				5. FUNDING NUMBERS
6. AUTHOR(S) Hakan Yazilitas				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000				8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A				10. SPONSORING/MONITORING AGENCY REPORT NUMBER
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.				
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited				12b. DISTRIBUTION CODE
13. ABSTRACT (maximum 200 words) This study explores whether the attacker's daily casualty rate (DCR) changes according to the terrain. The data set is a part of a larger database, Division Level Engagement Database from the Dupuy Institute. There are data on 253 battles, 96 of which occurred in urban areas. All the engagements are selected from European Theater of Operation (ETO) in World War II. The available data set contains measurements about the battles like initial strengths, daily casualties, terrain, front width, linear density, attacker's and defender's country, and armor losses. Hypothesis tests are used to find if the DCR is different in urban operations. A linear regression model is constructed to predict outcomes of similar engagements and to see the effect of each variable. It is concluded that the attacker's daily casualty rate is, on average, lower in urban operations. Terrain and force ratio are the most effective drivers of the daily casualty rate. In addition, it is seen that allied forces (U.S., U.K. and Canada) had a different approach to Military Operations on Urban Terrain than Soviet and German forces. The Allies used extensive combat power in urban operations.				
14. SUBJECT TERMS Casualty Rate, Linear Regression, MOUT, Urban Warfare				15. NUMBER OF PAGES 95
				16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UL	

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

**ANALYZING THE EFFECTS OF URBAN COMBAT ON DAILY CASUALTY
RATES**

Hakan Yazilitas
First Lieutenant, Turkish Army
B.S., Turkish Military Academy, 1998

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

**NAVAL POSTGRADUATE SCHOOL
June 2004**

Author: Hakan Yazilitas

Approved by: Samuel E. Buttrey
Thesis Advisor

Saverio Manago
Second Reader

Jim N. Eagle
Chairman, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

This study explores whether the attacker's daily casualty rate (DCR) changes according to the terrain. The data set is a part of a larger database, Division Level Engagement Database from the Dupuy Institute. There are data on 253 battles, 96 of which occurred in urban areas. All the engagements are selected from European Theater of Operation (ETO) in World War II. The available data set contains measurements about the battles like initial strengths, daily casualties, terrain, front width, linear density, attacker's and defender's country, and armor losses. Hypothesis tests are used to find if the DCR is different in urban operations. A linear regression model is constructed to predict outcomes of similar engagements and to see the effect of each variable. It is concluded that the attacker's daily casualty rate is, on average, lower in urban operations. Terrain and force ratio are the most effective drivers of the daily casualty rate. In addition, it is seen that allied forces (U.S., U.K. and Canada) had a different approach to Military Operations on Urban Terrain than Soviet and German forces. The Allies used extensive combat power in urban operations.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

I.	INTRODUCTION.....	1
	A. GENERAL INTRODUCTION.....	1
	B. BACKGROUND.....	2
	1. The Dupuy Institute Study (TDI).....	3
	2. Bracken’s Attrition Models of the Ardennes Campaign and Lanchester Equations.....	5
	3. Faruk Yigit’s Thesis.....	6
	C. THE IMPACT OF URBAN TERRAIN ON OPERATIONS.....	7
	D. OBJECTIVES.....	11
	E. LIMITATIONS AND ASSUMPTIONS.....	11
	F. RESEARCH QUESTIONS.....	12
II.	SUMMARY STATISTICS.....	15
	A. INTRODUCTION.....	15
	B. DATA DESCRIPTION.....	16
	1. Independent Variables.....	17
	a. Attacker’s Strength (<i>XStrength</i>).....	18
	b. Defender’s Strength (<i>YStrength</i>).....	18
	c. Attacker’s and Defender’s Country (<i>XCountry, YCountry</i>)..	19
	d. Attacker’s Front Width in Kilometers (<i>XFrontWidth</i>).....	20
	e. Linear Density (<i>XMenFront</i>).....	21
	f. Total Strength Ratio (<i>Force Ratio</i>).....	21
	2. Dependent Variable.....	22
	C. DESCRIPTIVE ANALYSIS.....	23
	1. Parametric Two-Sample Tests.....	24
	2. Nonparametric Tests.....	25
	3. Independent Variables and Terrain.....	26
	a. Attacker’s Strength and Terrain.....	26
	b. Defender’s Strength and Terrain.....	27
	c. Attacker’s Front Width and Terrain.....	29
	d. Linear Density and Terrain.....	30
	e. Force Ratio.....	31
	4. Dependent Variables and Terrain.....	32
	a. Attacker’s Daily Casualty Rate and Terrain.....	32
	b. Defender’s Daily Casualty Rate and Terrain.....	34
	D. DISCUSSION.....	34
	E. CONCLUSIONS.....	35
III.	MODELS.....	37
	A. INTRODUCTION.....	37
	B. PREPARATION OF DATA.....	40
	1. Transformation of Data.....	40

2.	Interaction Terms	40
3.	Singularity	41
C.	ATTACKER'S DAILY CASUALTY MODEL	42
1.	The Model without Interactions	42
2.	The Model with Interactions.....	44
3.	Effect of the Terrain on the DCR	46
4.	The Effect of the Attacker's Country on the DCR	48
5.	The Effect of the Attacker's Strength, the Defender's Strength and the Force Ratio on the DCR	50
6.	The Effect of the Attacker's Front Width on the DCR.....	56
7.	The Effect of the Linear Density on the DCR	58
D.	CONCLUSIONS	61
IV.	CONCLUSIONS	63
APPENDIX A.	ACRONYMS.....	65
APPENDIX B.	SUMMARY STATISTICS.....	67
A.	INDEPENDENT VARIABLES	67
1.	Numeric Variables	67
2.	Categorical Variables	67
B.	DEPENDENT VARIABLES.....	67
APPENDIX C.	TRANSFORMATION OF VARIABLES.....	69
APPENDIX D.	ATTACKER'S DAILY CASUALTY RATE MODELS.....	71
A.	MODEL WITHOUT INTERACTIONS	71
B.	MODELS WITH TWO-WAY INTERACTIONS	71
	LIST OF REFERENCES.....	73
	INITIAL DISTRIBUTION LIST	75

LIST OF FIGURES

Figure 1.	The Daily Casualty Rates (DCR) for Attacker and Defender [After Ref. 7].....	7
Figure 2.	Vertical Nature of the Battlefield in Urban Areas [From Ref. 8]	9
Figure 3.	Content of Data	16
Figure 4.	Explanatory Plots of Attacker’s Strength	18
Figure 5.	Explanatory Plots of Defender’s Strength	19
Figure 6.	Explanatory Plots of Attacker’s Front Width	20
Figure 7.	Dispersion of Force Ratio	22
Figure 8.	Variability of Attacker’s Casualty Rate	23
Figure 9.	Histogram of Attacker’s Strength Given Terrain.....	27
Figure 10.	Histogram of Defender’s Strength Given Terrain	28
Figure 11.	Histograms of Attacker’s Front Width (in km) Given Terrain	29
Figure 12.	Histograms of Attacker’s Linear Density Given Terrain.....	31
Figure 13.	Boxplot of Force Ratio Given Terrain.....	32
Figure 14.	Boxplot of Attacker’s Daily Casualty Rate Given Terrain and Country	33
Figure 15.	Defender’s Daily Casualty Rate Given Terrain	34
Figure 16.	Aachen After Battle [From Ref. 12]	35
Figure 17.	Plot of Residuals vs. Fitted Values in the Model with Interactions.....	44
Figure 18.	Plot of Actual Transformed Values vs. Predicted Values.....	46
Figure 19.	The Effect of Terrain on the DCR	47
Figure 20.	The Effect of the Attacker’s Strength on the DCR in Urban Areas.....	52
Figure 21.	The Effect of Attacker’s Strength on DCR in Non-Urban Area.....	53
Figure 22.	The Effect of the Defender’s Strength on the DCR in Non-Urban Areas	54
Figure 23.	The Effect of the Force Ratio on the DCR (Terrain/Varying Force).....	55
Figure 24.	The Effect of the Front Width in Urban Areas	57
Figure 25.	The Effect of the Front Width in Non-Urban Areas.....	58
Figure 26.	The Effect of the Linear Density in Urban Areas	60
Figure 27.	The Effect of the Linear Density in Non-Urban Areas.....	60
Figure 28.	Transformed Attacker’s Daily Casualty Rate Plots	69
Figure 29.	Transformed Defender’s Daily Casualty Rate Plots.....	69

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

Table 1.	Content of Data	16
Table 2.	Variability of Data	17
Table 3.	Battlefield Density Throughout the Ages [From Ref. 10]	21
Table 4.	Test Results	26
Table 5.	Difference in the Daily Casualty Rates Between Countries	50
Table 6.	Summary Statistics for Numeric Independent Variables	67
Table 7.	Summary Statistics for Categorical Independent Variables	67
Table 8.	Summary Statistics for Dependent Variables	67
Table 9.	Summary Statistics for the Model without Interactions	71
Table 10.	Summary Statistics for the Model with Interactions	72
Table 11.	Coefficients Given the Attacker's Country and Terrain	72

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

I am grateful to Turkish Armed Forces for giving me such a chance. Also, I would like to thank my advisor, Profesor Samuel E. Buttrey and my second reader, LTC Saverio Manago for their support, guidance and help throughout the research.

Special thanks to Major Joseph Burger (USA) and Mr. Chris Lawrence (TDI) for helping me to gather the data set.

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

Urban warfare has always been a great concern for military planners. Although, Military Strategy books recommend avoiding urban operations, history records numerous occasions where important battles occurred in urban areas. These occasions include some unforgettable battles such as the battles for Troy, Istanbul, Berlin, Stalingrad, Suez City, Grozny and many others.

As urbanization grows, it is unlikely to find a campaign that does not include Military Operations in Urban Terrain (MOUT). MOUT is becoming increasingly important. Rapid urbanization and the resources that urban areas have, such as communication and transportation nodes, make MOUT more likely. We are more likely to see defending forces operating in urban area when they are outnumbered, in order to because they will want to use the advantages of cover and concealment and man-made construction in urban areas.

Until recently, the question regarding MOUT has been the outcome of battles. As the public concern over the number of casualties increased, the focus shifted to trying to predict the number of casualties. The days of “war of attrition” have gone for the modern militaries. In modern wars, a general is not considered successful unless he wins the battle with minimum casualties.

This thesis analyzes the effects of urban terrain on the attacker’s daily casualty rates. The main focus is to distinguish differences between urban and non-urban operations in terms of attacker’s daily casualty rate (DCR). Also, analyses are conducted to find what causes the difference in DCR according to the terrain.

The data set used in the thesis was compiled by the Dupuy Institute (TDI). It consists of 96 urban and 156 non-urban engagements. Each data point shows a division-level engagement. All the observations are selected from the WW II era. The opposing

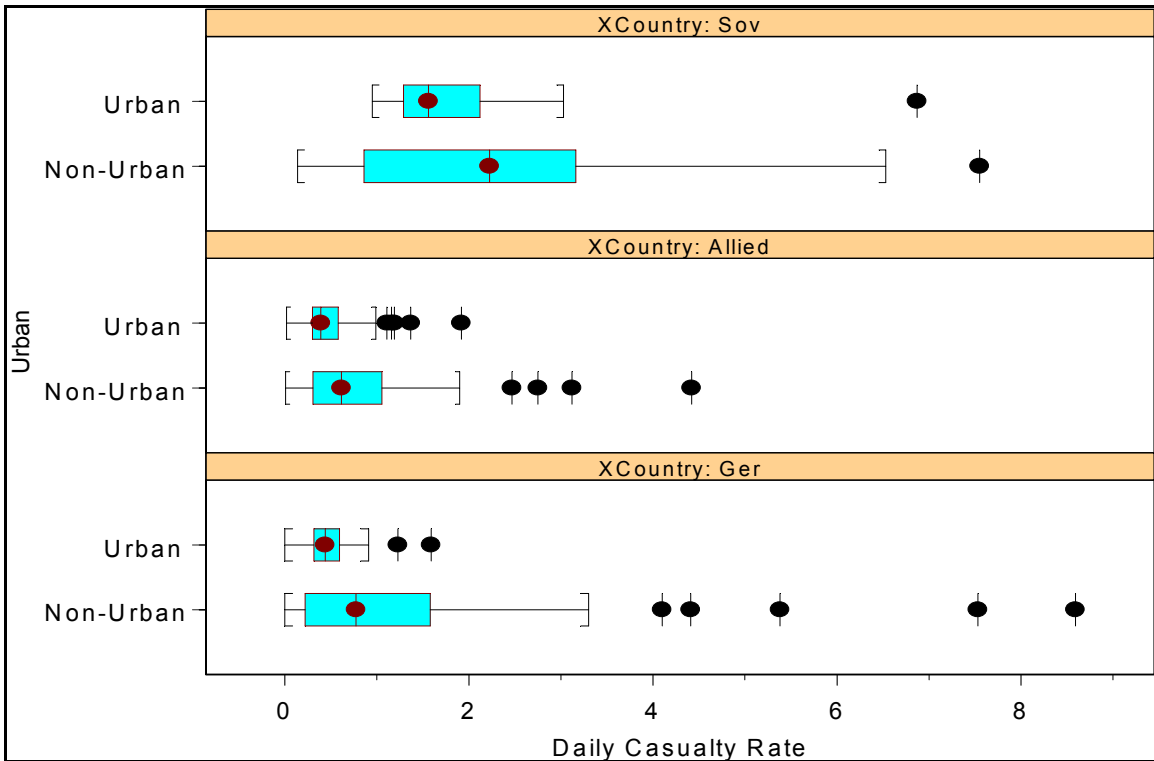
forces are Canada, U.K., U.S., the Soviet Union and Germany. Since there are limited number of observations for Canada and U.K., they are combined with U.S. engagements under the name of “Allied.”

This data set is selected for a couple of different reasons. First of all, this data set includes a sufficient number of urban and non-urban engagements to conduct the analysis. Also each observation was verified by checking the reports of the both sides; thus it presumably is reliable. Furthermore, the battles occurred at about the same time, in similar locations and between similar forces.

The “Capture Rate Study, Phase I and II” of TDI concluded that for casualty effectiveness, Allied and German forces were almost the same and German forces, more effective than Soviet forces. However, it is assumed that these differences remained constant throughout the period of the data set. Also, it is assumed that there is no technological difference between the forces. These assumptions enable the analysis to claim that any differences between the DCR’s in urban and non-urban operations are actually the result of the terrain.

The model to predict the DCR is constructed by using linear regression. Attacker’s strength, front width, linear density, country as well as force ratio, terrain and defender’s strength are used to build the model. It is concluded that power transformations and two-way interactions are necessary in the model to increase its ability to explain the variability of the data set and to comply with the assumptions of linear regression.

The hypotheses tests revealed that DCR in urban operations is *lower* than in rural operations. This result seemed counterintuitive. However, the variable which is lower in urban operations is the daily casualty rate, not the total number of casualties. Since urban operations take more time, even though the daily casualty rate is lower, the total number of casualties might be higher in urban operations. The effect of Terrain on DCR appears below.



The analyses resulted that terrain and the attacker's strength are the most important factors for determining DCR. Defender's strength does not significantly change the DCR. The German and Allied forces had almost the same daily casualty rates, somewhat lower than those of the Soviet forces. Although the analysis indicates that Allied forces had a different understanding of MOUT, their casualty rates were almost the same as those of the German forces. In addition, the effect of each variable is different for Allies and the others, depending in the terrain.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. GENERAL INTRODUCTION

The need for understanding the nature of war can be seen throughout history. Scholars and military planners alike have long looked for a way to understand the physics of war. Despite modern developments in warfare, Sun Tzu remains one of the best-known warrior-philosophers, whose 25-century-old studies still affect every phase of war planning and execution.

Most of the early studies gave priority to the outcome of the battle, because human casualties were not considered as important as they are today. Improvements in technology, communication and welfare make human loss in battle more sensational today than it has ever been. Nowadays losing even one soldier in combat seems to affect the prestige of a nation. Thus, determining a method to estimate the number of casualties is becoming more important every day.

While early studies relied on pure observation and data analysis, contemporary studies also benefit from simulations that can be used to estimate the number of casualties. Once a simulation is set up, many replications and iterations can be made in a relatively short time. However, setting up a simulation can be costly in both money and time. Most simulations need results from data analysis to feed the model requiring reliable data, which may be as costly to produce as obtaining results from the simulations themselves.

Consider the following key finding from a 2001 study on world urbanization conducted by the United Nations that points out the increasing rate of urbanization:

Half the world population is expected to live in urban areas in 2007. The urban population reached 2.9 billion in 2000 and is expected to rise to 5 billion by 2030, whereas 30 per cent of the world population lived in urban areas in 1950 and the proportion of urban dwellers rose to 47 per cent by 2000 and is projected to attain 60 per cent by 2030. [Ref. 1]

Given the increasing ratio of urban population to rural population, it is nearly certain that future wars will involve urban combat. Almost every operation involving U.S. forces after WWII has required combat in urban terrain. The basic reason is the superiority of

U.S. forces in technology and training. This dominance has forced inferior forces to use populated areas as a shelter. To minimize civilian casualties, technologically advanced forces have had to change their rules of engagement (ROE), in a way that reduced their technological advantage.

As a result of dense civilian populations and the complex structure of urban terrain, Military Operations in Urban Terrain (MOUT) have aspects different from non-urban operations. For this reason, combat in cities has tended to be avoided by commanders unless absolutely necessary. Thus, in studying casualties one must also consider the effects of terrain. However, most of the studies concerned with predicting casualties either used data on non-urban operations or did not make any discrimination between urban and non-urban operations.

This thesis will focus on finding the differences between casualties in urban and non-urban operations by means of data analysis tools. The Dupuy Institute constructed the data through recent research and by extracting related observations from the existing Division Level Engagement Database (DLEDB). Data are selected such that battles occurred close together in time and space. The reliance on the conclusions of “Capture Rate Study” renders spatial effects unimportant. [Ref. 2]

Considering the aforementioned factors, the most reliable and available data comes from World War II (WWII). The data set includes 253 division-level engagements from early 1943 to the beginning of 1945 executed in the European Theater of Operations (ETO) between allied and axis forces. Of 253 engagements, 97 are urban engagements and 156 are non-urban engagements.

B. BACKGROUND

After the devastation of WWII, the international community, especially developed countries, became more sensitive about the loss of human life in wars. They started questioning whether material gains were worth losses of human lives. In addition, people started watching the destructive effects of wars on civilians as well as on their own soldiers on television.

During the Vietnam War, people started demonstrating their concern about war's effects more effectively. This forced decision makers to find new ways to decrease

casualties. The low casualty rates in the Gulf War set a new standard [Ref. 3]. However, even with this new standard, community concern for casualties has not stopped. Although the number of casualties in an entire contemporary war is less than the number of losses suffered in one day in WWII, war still poses a great risk for decision makers who do not want to lose their positions. This fact led to new studies for predicting the number of casualties and developing new technologies for reducing the number of casualties. The latter caused the invention of smart, highly technological weapon systems that cause less collateral damage and acceptable personnel protection, so as to increase the probability of survival given a hit.

Governments now sponsor institutions to find models to predict the numbers of casualties. Thus, they have an analytical way to compare the positives and negatives of their actions. The Israel Armed Forces Medical Corps were the first to use these models. After studying the Arab-Israeli war, they established models to predict casualties. With the help of these models, they decreased the numbers of casualties by assigning more medical units to the places where more casualties were expected. Although medical planners were the first to use these models, today mostly force planners and decision makers use these models.

1. The Dupuy Institute Study (TDI)

The Dupuy Institute (TDI) and its founding father, Col. Trevor N. Dupuy, have issued many books, data and researches on military history. Col. Dupuy, who wrote over 80 books about military history and tactics, tried to increase the effectiveness of models during his analysis of military history in order to find patterns and trends. With the help of these findings, he hoped to find a basis to win the next war. As a non-profit corporation the Dupuy Institute followed this path. It has authored numerous reports and studies for the U.S. Army.

The study “Measuring the Effects of Combat in Cities” [Ref. 4] is one of the recent studies of the differences between urban and non-urban operations. TDI collected its own data from the WWII European Theater of Operations (ETO), and verified the accuracy of the data by looking at the records of the two sides. This thesis will employ these data as well.

The data were selected in such a way that each army involved in these battles (U.S., U.K., Canadian, Soviet and German) had doctrines similar to those of the U.S. Army; no significant difference exists among them [Ref. 4]. The human factors are not of much concern according to an initial study [Ref. 2]. In that study, TDI compared combat effectiveness, primarily the casualty effectiveness, and concluded that human factors remained the same throughout the time frame this data concerns. They also concluded that the level of engagement is an important factor and they selected division-level operations as the basis of their study.

Since the selected forces were similar and the engagements happened during the same time period, the analysts assumed that there were no differences in technology and performance, or at least, that the differences remained constant. This allowed the conclusion that differences between urban and non-urban engagements are indeed dependent on terrain. TDI analyzed the differences between the force ratios, outcome, casualty rates, duration of combat, width of advance and linear density for each type of terrain and attacker.

One of the conclusions of the TDI study, which is important for the scope of this thesis, is:

...the attacker casualties in the urban engagements are less than in the non-urban engagements and the casualty exchange ratio favors the attacker as well. Because of the selection of the data, there is some question whether these observations can be extended beyond this data, but it does not provide much support to the notion that urban combat is a more intense environment than non-urban combat.

This conclusion may look counter-intuitive. However, there may be several factors which lead to that conclusion. One is the difference in the manner in which the opposing forces reported casualties. Although Americans reported the severely wounded as well as the lightly and the moderately wounded, Germans normally reported only those sent to army-level hospitals, thus excluding most of the lightly wounded and some of the moderately wounded in the casualty reports.

One hypothesis is that the defenders did not use the advantages of the urban area. Those forces defended their cities in a conventional style as if they were defending a non-urban area. That might also lead to the lower number of casualties for the attacker.

Another reason could be the devastating examples of defending cities for the city itself. Defending forces might have given up fighting when they felt that they could not hold the city and withdrew to avoid the destruction of the city and its resources.

The last reason, which maybe the most important, is the concept that attacking is the ultimate way to gain objectives. The main principle for conducting defense is to provide enough time for friendly forces to gather and reorganize forces for attack. Following this principle, a defender's main effort should be to delay an attacker's advance to gain time instead of destroying attacking forces.

2. Bracken's Attrition Models of the Ardennes Campaign and Lanchester Equations

Bracken's study [Ref. 5] focuses on the Ardennes Campaign for which there is reliable and complete daily data available. Bracken's study differs from TDI's studies in the manipulation of data. Unlike the authors of the TDI study, Bracken computes the combat power of each side by giving weight to each weapon system according to its combat effectiveness. He uses this weighted combat power to calculate the force ratio.

Like TDI, Bracken assumes all American soldiers are the same, which means they all have the same properties, but includes weapons in his calculations. Since he calculates the combat power by weighting each element (e.g. tank, soldier, carrier), all his models are homogenous.

This study reveals that although the personal effectiveness of an Allied soldier is greater than that of a German soldier, the total German force was as effective as Allied forces. This is consistent with the Dupuy Institute's conclusions that none of the forces is different. Moreover, Bracken proposes that the attacker always has an advantage over the defender.

Bracken's study concludes that a Lanchester linear model fits the Ardennes campaign. The Lanchester linear model assumes that the casualty rate of a force is proportional to the product of its force size and enemy's force size. A famous maxim of

Lanchester's linear law is that the size of the forces matter a great deal as if to confirm Voltaire's famous saying, "God is always on the side of big battalions." As a deterministic model, Lanchester linear model inputs into to the formula the size and attrition rate for each side to predict who will win.

Since it is a deterministic model, it always gives the same result, which does not reflect real life. As opposed to a stochastic model, which provides results with measures of uncertainty, a deterministic model provides only a point estimate, which presents diminished utility when applied to actual force variables. It does not reflect the high variability of the number of casualties, especially in urban warfare. In an urban engagement, given the same conditions, one side may encounter no casualties or lose all its forces [Ref. 6]. The data used by this thesis exhibit the same kind of variability in that there are cases in which the attacker suffered no casualties and there are cases in which the defender lost all combat power.

3. Faruk Yigit's Thesis

In his thesis, Yigit [Ref. 7] studied 3-to-1 ratio, which is considered to be the minimum ratio of attacking forces to defending forces to make victory likely and evaluated the historical changes in dispersion and casualty rates. He used a variety of data, from the Netherlands' War of Independence in 1600 to the Israel-Lebanon War in 1982. He examined 532 battles, at varying force levels and tried to determine trends for force ratio, dispersion and casualty rates.

For force ratio (attacker's manpower divided by defender's manpower), he concludes that historical data supports the 3-to-1 rule of thumb. He found that the higher the force ratio, the more likely the attacker wins the battle. Using force ratio as an estimator is useful as a gross measure for campaign planning and forecasting the battle outcomes. This study also showed that there is no time trend in force ratio.

In his analysis of dispersion (area per soldier), Yigit realized that dispersion exhibited a trend across time. The trend shows that area per man has steadily increased. That is most probably the effect of the invention of lethal weapons. "As lethality increased, tactics, such as increasing the dispersion of combat forces, were adopted to minimize the effectiveness of the enemy's weapons."

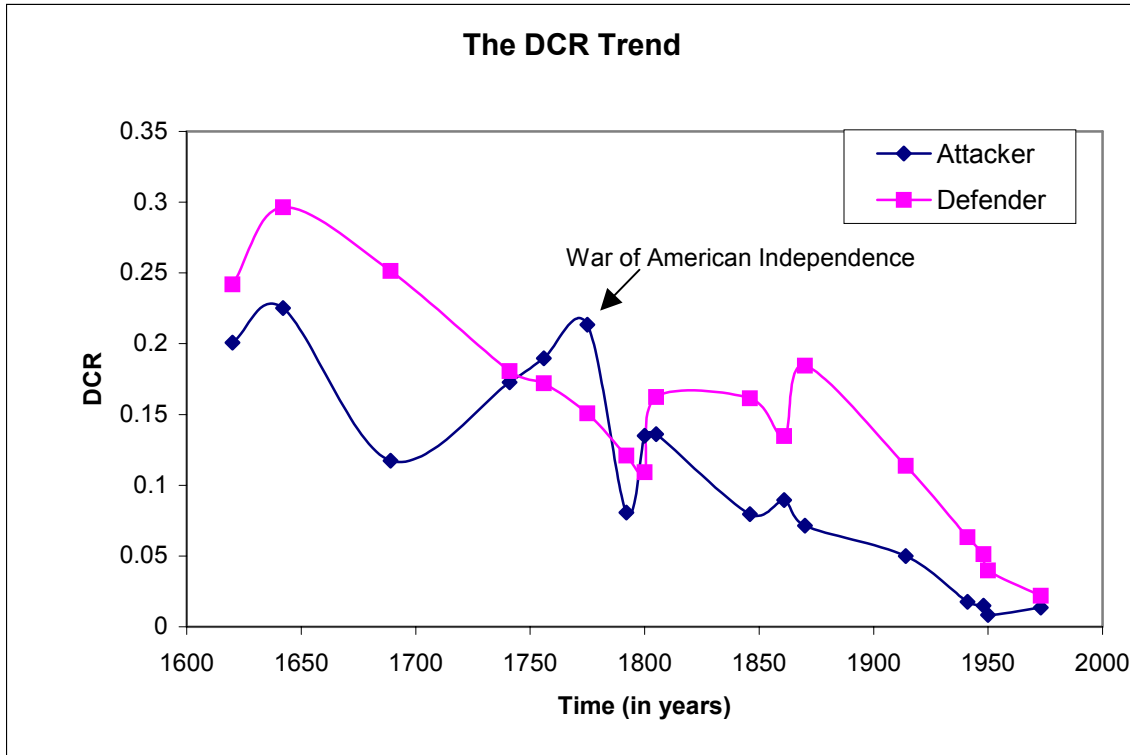


Figure 1. The Daily Casualty Rates (DCR) for Attacker and Defender [After Ref. 7]

For the daily casualty rate, he went further and looked not only at the historical trends but also at the size of units and battles for each campaign in time consequences.

The conclusions are:

- Casualty rates have declined over time.
- The attacker's casualty rates were usually lower than the defender's.
- The daily casualty rate increases as unit size decreases.
- Battle outcomes are more probabilistic than deterministic.

C. THE IMPACT OF URBAN TERRAIN ON OPERATIONS

The rule is, not to besiege walled cities if it can possibly be avoided.

Sun Tzu, The Art of War

Cities have always been the centers of social, economical, industrial and political powers. Governmental departments, resources, key communication nodes, dense industry and labor, most of which are vital to maintaining the functioning of the state and the

armed forces, reside in urban areas. It is important for the attacking forces to capture these assets not just to break the will of the defender but also to use these assets for themselves. Thus, war planners have always thought of cities as “centers of gravity.”

Additionally, gaining control of a city has a huge psychological impact for people because of historical and sentimental ties. One striking example is the German defense of Aachen in the late 1944's. Although Aachen was the first German city attacked, the already over-extended German forces defended the city at all costs suffering large numbers of casualties. If the city in question is a country's capital or a major city, thus more valuable and meaningful, then its defense is more important. One apparent example is the German attack on Stalingrad. Although the Germans knew that attacking Stalingrad was not the best course of action, they took the risk of attacking that city because of its psychological and strategic importance to each side.

Since major cities have vital resources, communication nodes and psychological importance, they have always been an arena for wars. However, every conventional force operating in an urban area has learned that fighting in an urban area poses a set of difficulties and challenges usually different from the conventional battle space in which they were trained to operate.

Military Operations in Urban Terrain (MOUT) differs from operations in non-urban areas because of man-made obstacles. Buildings channel attacking forces towards possible planned kill zones or ambushes. Buildings also offer good cover and concealment for the defender. Snipers can use this cover very effectively. Also, city structure interrupt the command, control and communication chains, which might lead to a separation of forces or a lessening of the synergistic effect of the operations.

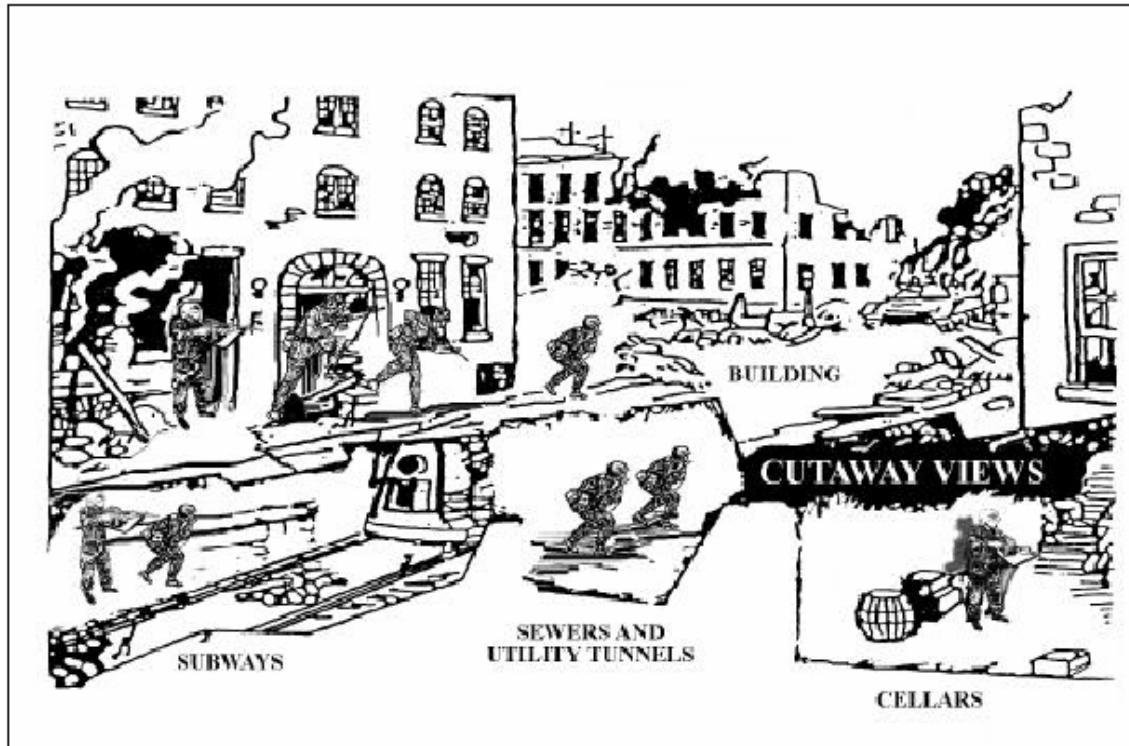


Figure 2. Vertical Nature of the Battlefield in Urban Areas [From Ref. 8]

The “vertical” nature of a city adds another dimension to the difficulty of conducting MOUT. In addition to above-ground constructions such as multi-story buildings, underground constructions, such as sewer systems, pose extra threats and demand added attention from each soldier performing the operation, contributing combat fatigue. High buildings grant good observation and fire positions. Also, they provide antitank weapons the ability to attack armored vehicles from above, their usual point of weakness. An attack on the rear or flank, which is usually more effective, can come from the underground systems.

The complex nature of streets provides advantages to the defending forces who are generally more familiar with the city. City structure also provides an excellent environment for the execution of hit-and-run types of guerilla warfare. Blocking of streets is more effective in an urban area than it is in a non-urban area, since in a city it generally is not possible to circumnavigate the obstacles.

The high density of civilians in cities is another factor affecting MOUT and normally restricts the operating forces' Rules of Engagement (ROE). Forces cannot behave as if they are executing in the open area. They have to consider every move they make to avoid unexpected consequences. The international community's growing concern about non-combatant casualties, perhaps particularly in the case of non-governmental organizations and United Nations workers, restricts available combat options. Given today's immediate media broadcasts from within combat areas, military planners attempt to minimize collateral damage that could affect public opinion negatively toward that force.

Using weapons of mass destruction in cities might allow a tactical victory but it might also cause a strategic defeat through the loss of public and international support. Inconsiderate damage to civilian property and life can alienate the local population and prevent the preservation of facilities for future use by attacking forces [Ref. 8].

Urban areas can be controlled by clearing every street, every building and every cellar, but this makes urban operations manpower-intensive. Also, the denser the forces, the greater the casualties. Thus, during the planning phase, it is vital to calculate the optimum number of soldiers to achieve objectives while minimizing casualties.

One other factor that affects MOUT is the variation of tactics used by the forces. If one force uses guerilla tactics it becomes harder for the other force to differentiate combatants from non-combatants. On the other hand, guerilla warfare is often the only method for forces inferior in number and technology to defend themselves. Experience has verified that the hit-and-run technique is not only good in non-urban areas but also in urban areas.

Combat in cities also affects the organization of forces. For example, armored vehicles need strong infantry support to operate in urban areas. Urban areas also restrict the use of some types of ammunitions and equipment, such as flamethrowers and artillery, which are in the organic structure of the conventional forces. Instead, forces may need bulldozers or guided munitions.

It is not wise to use non-urban tactics in urban operations. MOUT needs special training and tactics. If forces do not practice these tactics during peacetime, they have to learn the tactics through experience on the battlefield.

D. OBJECTIVES

A number of questions arise for data analysts as the analysis of urban operations becomes more important. One obvious question is how to identify the differences between urban and non-urban areas. Differences between tactics, equipment, force organizations and psychological effects are the first ones to consider. This thesis examines if daily casualty rates for the attacker forces in urban and non-urban engagements differ. The next step will be to identify the important factors and their effects on the attacker's daily casualty rate (DCR).

Another objective of this thesis is to find a model that could be used to predict the casualty rate for an engagement under the same assumptions. Due to rapid development and employment of technology, especially in military technology, models based on historical data may not produce reliable estimates for contemporary use. However, these results may provide bounds for the attacker's DCR, and thus, help in planning their human needs and the organization of their forces.

E. LIMITATIONS AND ASSUMPTIONS

Due to the multi-dimensional nature of war and a lack of complete data, studies cannot go further without some assumptions. This study is no exception.

It is assumed that each force has the same doctrine as the others or at least that the differences are constant. Bracken's study also makes this assumption. This assumption is supported by the Capture Rate Study of The Dupuy Institute [Ref. 2]. This study concludes that the combat effectiveness of U.S. and German forces were equal and the German forces were 20 to 50 percent more effective than those of the U.K. and Canada. German forces were also rated better than the Soviets in combat effectiveness. Due to the small number of cases where U.K. and Canadians fought against the Germans, the Germans advantage in combat effectiveness can be ignored. However, there are many data points for engagements between Soviets and Germans. The assumption is that the difference in the combat effectiveness between Germans and Soviets remained the same.

Although another assumption is that the combat effectiveness of all countries is equal or the difference is constant, the analysis uses a nationality factor to capture any differences among nations.

Reliance on WWII data to represent all the engagements occurring before or later, and in different areas, is another big assumption. However, considering that the main principles of war have hardly changed for centuries, the results of this thesis will still be enlightening.

One limitation of this study is the absence of technological factors. Due to the aforementioned assumption, technological advances will not be considered. Hence, this requires excluding the effects of airplanes, missiles and so forth.

It is not possible to infer any conclusion about the effectiveness of the Navy and Air Force in MOUT because of their lack of experience in urban operations.

All the engagements are selected from division-level engagements. Thus, the conclusion may not fit for lower-level engagements, such as those of battalion and company levels. It is likely that casualty rates would be greater in the lower-level units than the higher-level units as the latter contains a higher proportion of combat support and combat service support elements. Yigit's thesis [Ref. 7] also supports this finding. However, this thesis will not seek to discover if the conclusions change according to the level of operations.

Since a validated model does not exist to compare a Soviet T-34, German Panther and an American Sherman tank, this thesis did not use any weighting techniques to calculate combat power for attackers and defenders nor the force ratio. Thus, the attacker and defender strength and force ratio is calculated using the number of personnel, excluding the number or quality of armor, artillery and aircraft in the model.

F. RESEARCH QUESTIONS

Unlike most studies, this thesis will not try to ascertain the “outcome” of the battle. It will not answer the question, “who will win?” This study will differ from the previous studies in two aspects: it will discriminate between urban and non-urban engagements and it will try to discover an answer to the question, “what is the expected attacker's daily casualty rate?”

The main purpose of this study is to determine if the daily casualty rates are different for urban and rural engagements in division-level engagements. The following questions are relevant to this determination:

- Which factors are statistically significant for predicting casualties?
- Who suffers more casualties?
- What is the best method to estimate the daily casualty rate?
- Do urban engagements result in more casualties?

Answers to these questions already exist in some form. They can be found in military manuals and other studies. However, they could give different answers to the same question. This is normal because of the different assumptions and high number of variables and high variances of these variables.

THIS PAGE INTENTIONALLY LEFT BLANK

II. SUMMARY STATISTICS

A. INTRODUCTION

This thesis utilizes the data used for the study “Measuring the Effects of Combat in Cities: Phase I and II.” The data were collected for the two-phased study completed in June 2003. These data consist of 253 engagements between the Soviet Union, the United States, Canada and the U.K. against Germany in WWII between 1943 and 1944. The availability of numerous reliable data resulted in the selection of these years and places.

These data are enhanced by 111 engagements recently added to the existing DLEDB. The Dupuy Institute added these engagements to the study after a two-year research. The Dupuy Institute claims that all data come from original unit records, the U.S. National Archives, the British Public Records Office, the Federal German Archives, the Canadian Military Headquarters Historical Section and the Russian Military Archives [Ref. 4: 7-10, Ref. 9:6-10].

The data consist of both continuous variables and categorical variables with up to nine levels. Most of the variables are objective results derived from the archives of the forces. Nevertheless, there are also subjective variables derived from the researchers’ judgments, though only one of those subjective variables will be used in this thesis. Precisely speaking, “Terrain” which is judged according to a set of rules created beforehand, will be the only subjective element in the analysis. The “data description” section provides more details.

Figure 3 represents a quick look at the data. It shows the number of engagements for each force and the terrain. One notable issue is the absence of data for Germans attacking Allied forces in urban areas. Thus, to analyze German operations on urban terrain, it is necessary to rely on German attacks on Russian forces. As Soviet, U.S., U.K. and Canadian forces were allies during WWII, no data appear for engagements among them in Figure 3.

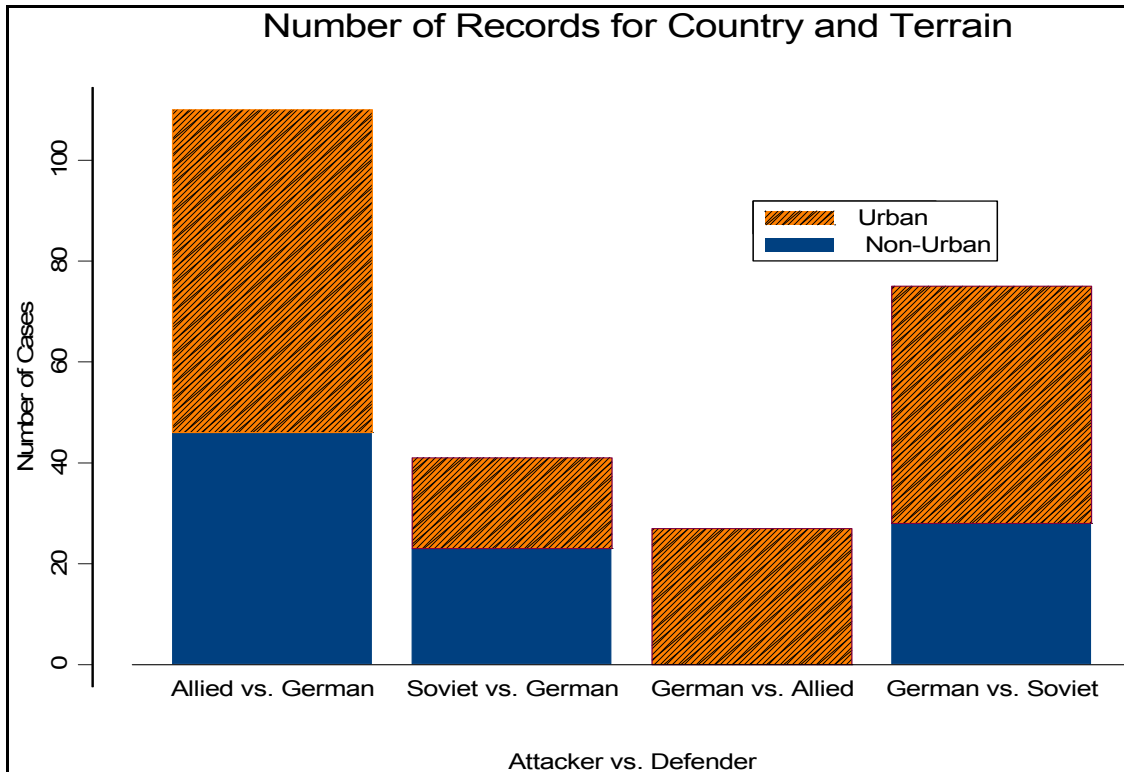


Figure 3. Content of Data

Another significant aspect of the graph is that in German versus Allied engagements, the Germans were usually the defender. Conversely, among engagements between Germans and Soviets, the Germans were usually the attacker.

Attacker	Defender	Urban	Rural	Total
Allied	German	46	64	110
Soviet	German	23	18	41
German	Allied	0	27	27
German	Soviet	28	47	75
Total		97	156	253

Table 1. Content of Data

Out of 253 engagements, 156 occurred in non-urban areas and the remaining 97 occurred in urban areas.

B. DATA DESCRIPTION

The original data consist of 26 columns. The first six columns provide information about who fought against whom, where, when and for how long. The next two columns are based on the outcome of the engagement, a subjective determination.

The rest of the data presents the facts of the engagement, giving the number of personnel, advance rate, total casualties, daily casualty rates, attacker’s front width and linear density (number of attacking soldiers per kilometer in the front). The last part of the data consists of the total and daily armor losses for each side.

Some data are missing on the daily advance rate, attacker’s front width and linear density. For armor loss, most data are missing.

In addition to the high number of variables, the high variability of each variable affects any outcome of a model using these variables. For this reason, it is difficult to predict the outcome of a battle with confidence. Table 2 shows the range and variability of some variables present in the data.

Variable	Min	Max	Max/Min
Attacker Strength	5600	67829	12.11
Defender Strength	358	57089	159.47
Attacker's Front Width (Km.)	1	203.65	203.65
Linear Density ((men in Km.)	203.65	19322.08	94.88
Force Ratio	0.4	45.53	113.83

Table 2. Variability of Data

For convenience, some abbreviations are created in order to understand the codes and variables. Throughout the entire study, variables beginning with “X” denote the attacker. Thus, XCasualty refers to the attacker’s casualty rate. “Y” denotes the defender, and YStrength indicates the defender’s strength.

1. Independent Variables

The independent variables are those which are assumed to be known by the attacker at the beginning of the engagements. Thus, the strengths of forces, length of the front line, force ratio, terrain and the nationality of forces are considered independent variables. Note that all numeric independent variables are greater than zero.

a. Attacker's Strength ($X_{Strength}$)

The attacker's strength is the sum of the number of troops subject to enemy fire including combat support and combat service support elements at the beginning of the engagement. This number does not include the number of tanks, artillery and so forth.

As seen from Figure 4, the distribution of attacker's strength is right skewed. The mean is 24257 and the median is 19032. The great difference between the mean and median is expected due to the number of extreme outliers. Extreme outliers are observations which are more distant from the median than 1.35 times the difference between the first and third quartile.

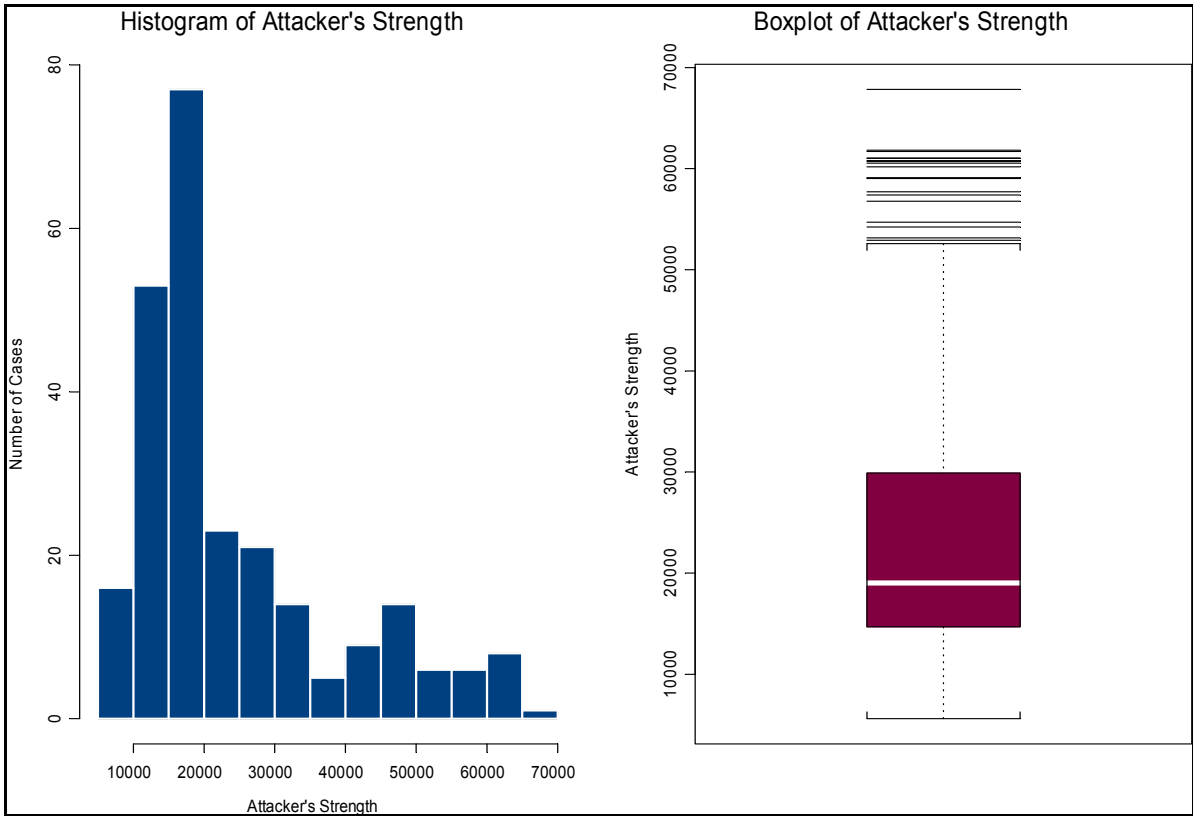


Figure 4. Explanatory Plots of Attacker's Strength

b. Defender's Strength ($Y_{Strength}$)

This is the number of defending personnel in the field. It does not include the number of artillery, tanks or planes. The distribution of defender's strength is also right skewed with extreme outliers, which causes the mean to be larger than the median.

However, the defender's strength is not as dispersed as the attacker's strength. Since all elements were considered homogenous, no weights were given to any particular forces while calculating the attacker and defender's strength. See Figure 5.

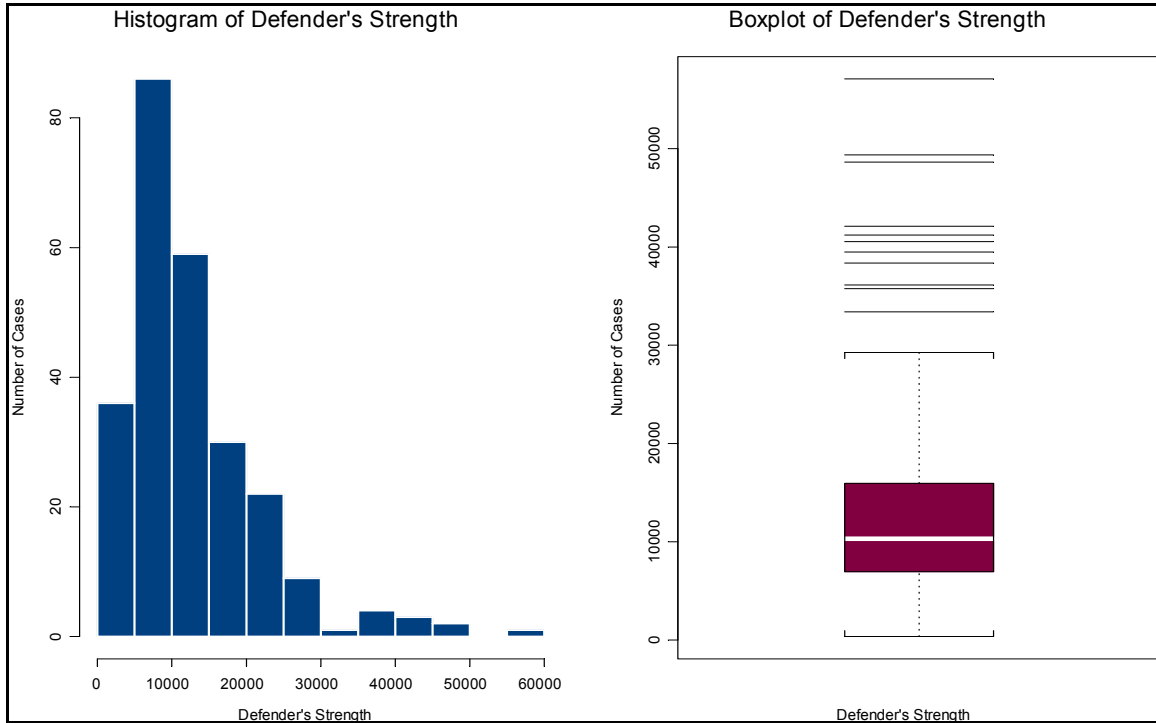


Figure 5. Explanatory Plots of Defender's Strength

c. Attacker's and Defender's Country (XCountry, YCountry)

XCountry and YCountry are categorical variables with three levels: German, Allied and Soviet. Due to the limited number of U.K. and Canadian engagements, these countries are included with U.S. engagements. The name "Allied" forces encompasses all the engagements consisting of U.S., U.K., and Canadian forces.

By including country variables, the expectation was to increase the ability to explain their effectiveness as well as to decrease the variability. As mentioned in TDI Study [Ref. 4], Allied forces usually attacked urban objectives with more troops and the Soviets defended their cities with more troops. Thus, it was plausible to expect a country and a terrain interaction term in this analysis.

Some additional facts must be mentioned. No information illustrating German attack on Allied forces in urban area exists. Additionally, while Germans were

generally the defender against allied forces, they were generally the attacker against the Soviets. However, these results purely rely on the available data. It might not be generalized, since it includes a specific time period, namely, 1943 and 1944.

d. Attacker's Front Width in Kilometers (*XFrontWidth*)

This is the length of the front line for the attacking force measured in kilometers. These measurements are generally derived from the operations plans of that force. Thus, it might be slightly different from the actual due to changes during the execution of the operation. The term “Attacker’s Front Width” makes it easier to understand the density of the forces.

As with other numerical variables, this variable is also right skewed. However, compared to the previous variables, this one has fewer extreme outliers. Attacking forces are generally assigned 12 km. of the front line. However, because the variable is skewed to the right, the average is higher.

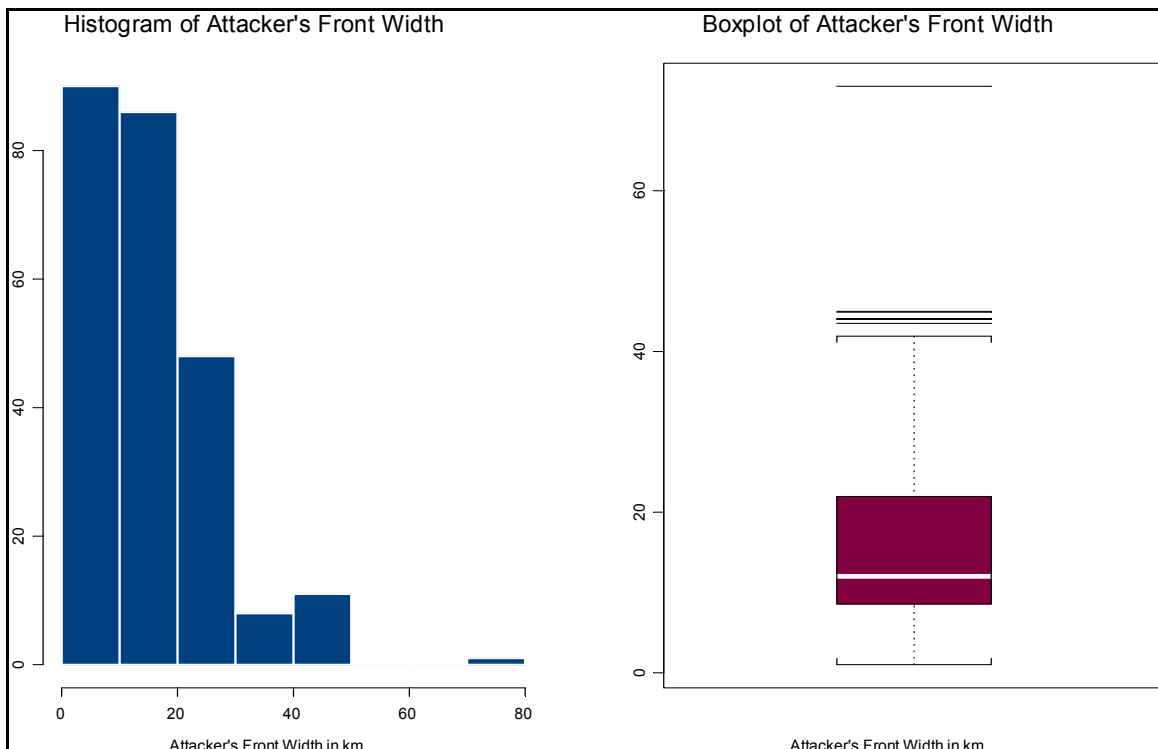


Figure 6. Explanatory Plots of Attacker's Front Width

e. Linear Density (*XMenFront*)

Linear density is a derived continuous variable indicating the number of personnel relative to the front width. It is calculated by dividing the attacker’s total strength (*XStrength*, defined above) by the attacker’s front width (*XFrontWidth*, defined above), where $XMenFront = \frac{XStrength}{XFrontWidth}$. This concept is known as dispersion in the military. It is one of the main considerations when planning a tactical movement, especially when selecting the movement technique.

After more accurate and lethal weapons were employed on the battlefield, military planners emphasized the dispersion of troops to decrease the effectiveness of these weapons. Thus, from a cost-effectiveness perspective, using expensive highly technological weapons against a highly dispersed enemy is not recommend, especially if the campaign is expected to be long-lived, all other things being equal.

Table 3 clearly demonstrates the trend in the density. The linear density has decreased steadily over time, however, the slope of decrease in linear density is steeper in non-urban areas, indicating greater dispersion.

Period	Antiquity	Napoleonic Wars	U.S. Civil War	World War II	October War
Density in all Terrain	100000	4970	3883	32	25
Density in Urban Terrain	16300	46400	11600	1300	1100

Table 3. Battlefield Density Throughout the Ages [From Ref. 10]

f. Total Strength Ratio (Force Ratio)

Force ratio is another derived continuous variable. It is calculated by dividing the attacker’s strength (*XStrength*) by the defender’s strength (*YStrength*), where, $ForceRatio = \frac{XStrength}{YStrength}$, indicates the relative strength of attacking forces.

Therefore, the attacker outnumbered the defender if the force ratio is bigger than 1. Force

ratio is one of the most dispersed variables in the data set as seen in Figure 7. Its range is [0.4, 45.53]. A close look at the data shows that most of the extreme outliers are from allied forces attacking Germans in urban terrain.

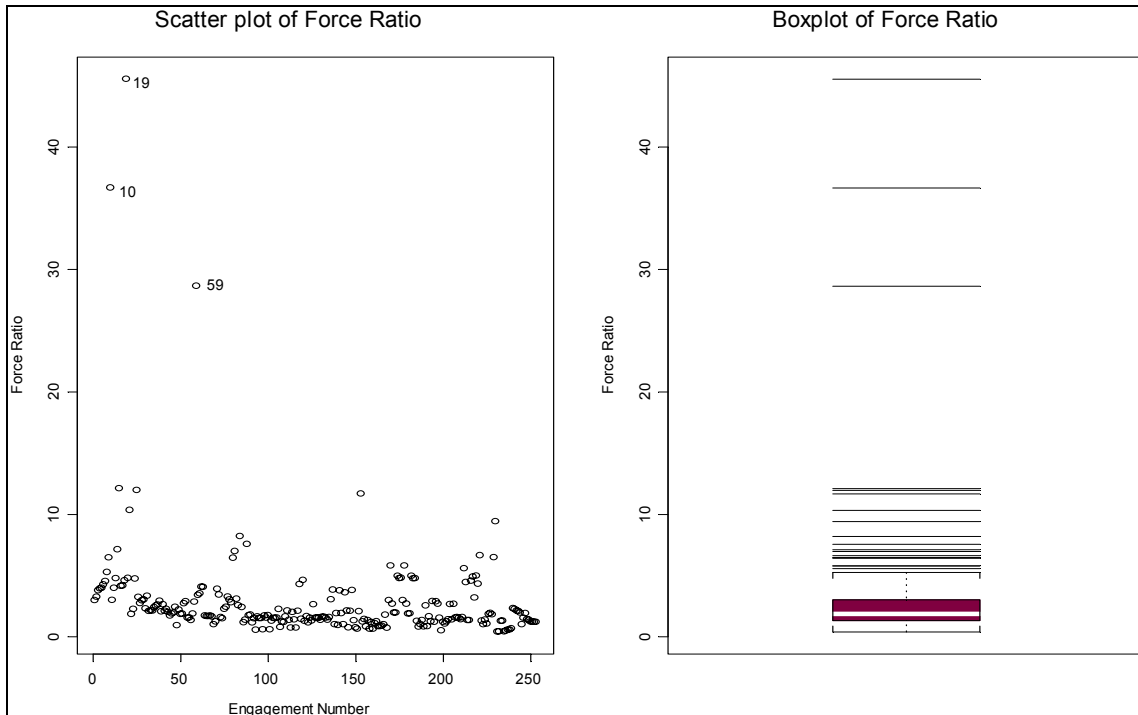


Figure 7. Dispersion of Force Ratio

The variation of force ratio should be related to the principle of force economy and establishing a main effort area. It is normal to see a higher ratio for a division, which is the main effort of a higher command. It is also normal to see lower force ratios for divisions whose orders are to disguise themselves and separate the enemy.

2. Dependent Variable

The dependent variable, also called the response variable, is the one desired to be predicted using independent variables. In this case, the dependent variable is the attacker's casualty rates.

The attacker's daily casualty rate is a continuous dependent variable. It is calculated by dividing the total number of attacker's casualties by the product of total

strength at the beginning and the duration of the engagement. It is also right-skewed with a range between 0 and 8.59 percent. Although on average, the attacker lost 1.07 percent of its strength daily, the median loss was 0.65 percent.

Figure 8 shows the variability of the attacker's casualty rate. It implies that attackers could suffer occasional unexpected or high casualties. Attacker's lost at most 8.59 percent of their beginning strength. This variable, like the others, is right-skewed. One important point is the presence of some observations with a zero casualty rate.

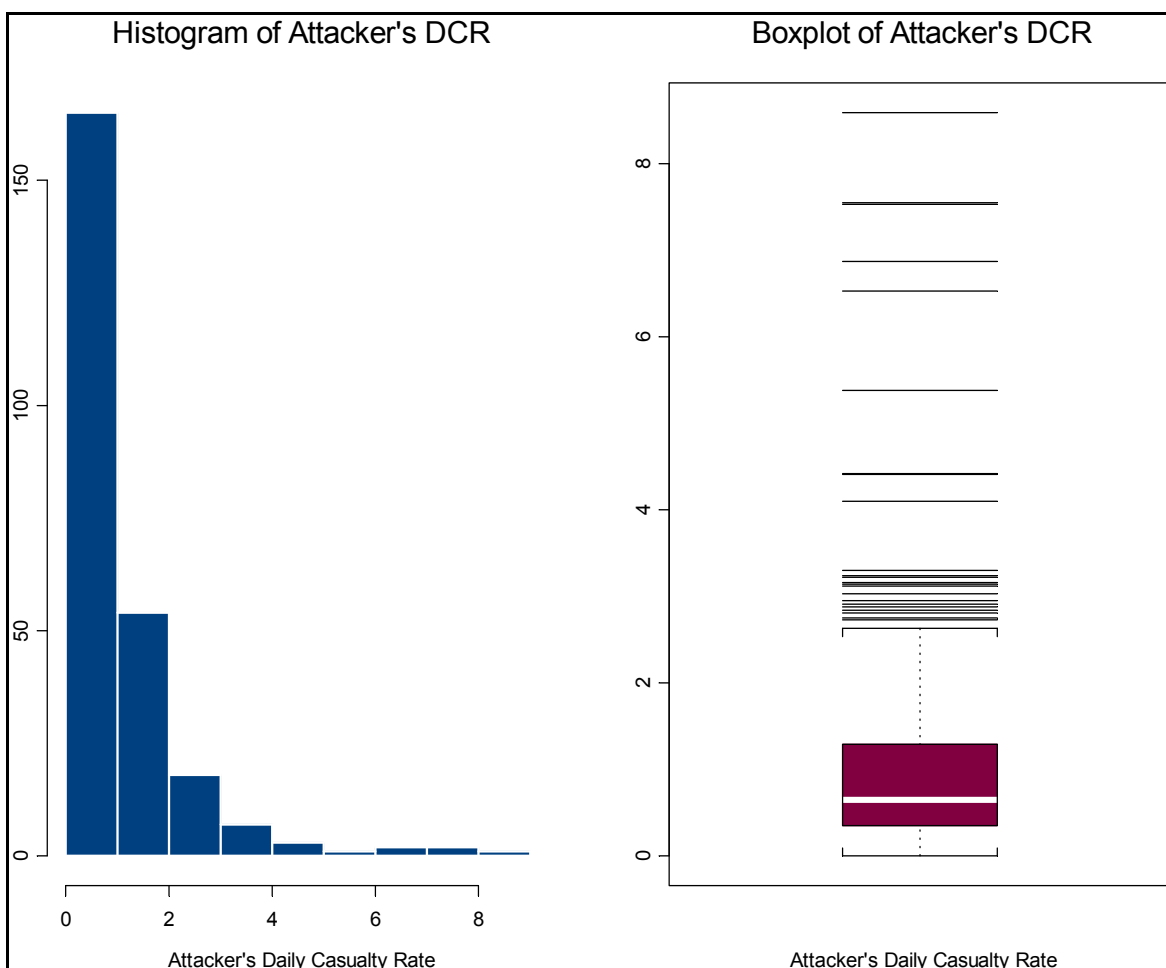


Figure 8. Variability of Attacker's Casualty Rate

C. DESCRIPTIVE ANALYSIS

This section focuses on the effects of terrain for each independent variable. We will treat the available data as if it constitutes a random sample from a hypothetical

population of WWII battles. Conclusions will be drawn from graphs and two-sample tests. The values for urban engagements will compose one sample and values for the non-urban the other sample. There are two methods for making inferences based on two-samples: parametric and nonparametric tests. Nonparametric tests are also known as distribution-free tests.

1. Parametric Two-Sample Tests

Parametric two-sample tests look for the difference between the means of the two samples. Parametric tests assume that samples come from a specific distribution and are used to test hypotheses and to determine confidence intervals. Tests of hypotheses consist of a null hypothesis (the prior belief) H_0 and an alternative hypothesis H_a . Only if the sample evidence strongly contradicts the null hypothesis is an alternative hypothesis accepted.

The null hypothesis states that the difference between the means of attacker's daily casualty rate in urban and non-urban areas equals zero. The populations are indistinguishable if the test results fail to reject the null hypothesis. Thus, it is possible to conclude that the mean of the samples are different with the rejection of the null hypothesis.

$$H_0 : \mu_1 - \mu_2 = 0$$

$$H_a : \mu_1 - \mu_2 \neq 0$$

If the distributions are normal and the variances of the population known, inferences about the sample and a confidence interval are determined by using z tests.

If those assumptions are not true, the next thing is to look at the sample sizes. For sufficiently large samples, the Central Limit Theorem (CLT) guarantees that the difference between the mean has approximately a normal distribution. In practice, the CLT is often used when both sample sizes are greater than 40 [Ref. 11]. If one of the sample sizes is less than 40 or if the standard deviation of the population is unknown, assuming they are normal distributed, a two-sample t test should be used to make inferences. The following formula calculates the test statistic for large sample tests:

$$t = \frac{\bar{x} - \bar{y} - \Delta_0}{\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}}$$

where Δ_0 is the difference between the mean daily casualty rate, which is

zero in this case, and s_1 is the standard deviation of the sample.

As seen in “data description,” all variables are right-skewed. Thus, it is not appropriate to use z tests. With 96 observations for urban and 153 for non-urban engagements, it is then feasible to assume that CLT assures that the difference between the means of the two samples is normally distributed. Thus, the analysis of the thesis’s hypotheses will use t tests.

2. Nonparametric Tests

Unlike t or z tests, nonparametric tests do not require the assumption of normal distribution. Since they do not require that the data come from any parametric distribution family, they are called distribution-free tests.

If it is not possible to use a t test because the sample size is small or the distribution is not normal, the Wilcoxon Rank-Sum test is used to analyze hypotheses [Ref. 11:p. 659]. To use this test, each sample is assumed to come from a continuous distribution with the same shape and spread. However, the means might be different. In null hypothesis, the assumption is that the difference between the means (μ_1, μ_2) of the samples is zero. The alternative hypothesis will be that the difference is not zero. In this analysis μ_1 will denote the samples from the urban operations and μ_2 will denote the samples from the non-urban operations.

The following section will attempt to ascertain whether urban operations are different from non-urban operations by analyzing each independent variable in each terrain type. The null hypothesis is rejected if the calculated test statistic is less than 0.05.

Table 4 shows the results of the t and Wilcoxon test and the mean of each predictor variable in each terrain type. The results indicate that attacker’s strength, linear density and attacker’s front width are not statistically significant. That is to say, in the presence of the other variables, including these variables are not necessary. The mean of each variable exhibits the characteristics of urban warfare. On average, each country used

bigger force size and higher force ratio and linear density in urban operations. Defending country used less force in urban areas or they defended in urban areas when they did not have enough force to defend in non-urban area.

	p-values		Estimated Mean in	
	t Test	Wilcoxon Test	Urban	Non-Urban
XStrength	0.3317	0.1299	25370	23565
YStrength	0.0004	0.0213	10561	14119
XMenFront	0.4295	0.8935	2347	2095
XFrontWidth	0.1779	0.7981	14.30	15.98
ForceRatio	0.0294	0.0005	3.747	2.339
XCasualty	0.0144	0.1162	0.844	1.212
YCasualty	0.0023	0.0012	14.36	4.988

Table 4. Test Results

3. Independent Variables and Terrain

a. Attacker's Strength and Terrain

The extreme outliers in the urban section of the plots are caused by the overwhelming power the U.S. VIII Corps used in the engagements of Brest in September 1944. In these engagements, Allied forces used 60,000 personnel on average, which is almost three times higher than the mean of all engagements in an urban area.

For non-urban operations, the outliers are also mostly attributed to U.S. attacks on German forces. Thus, it is possible to conclude that Allied forces tended to use extensive force more frequently than both German and Soviet forces, possibly because of available manpower or the effective use of what the military calls as the “main effort principle.”

The histogram in Figure 9 shows that the attacker's strength in urban and non-urban operations does not come from a normal distribution. The heavy tail on the right side supports this assumption. However, since there are more than 40 data points in each sample, it is possible to assume that the difference between the means of the attacker' strength in urban areas and non-urban areas follows a normal distribution.

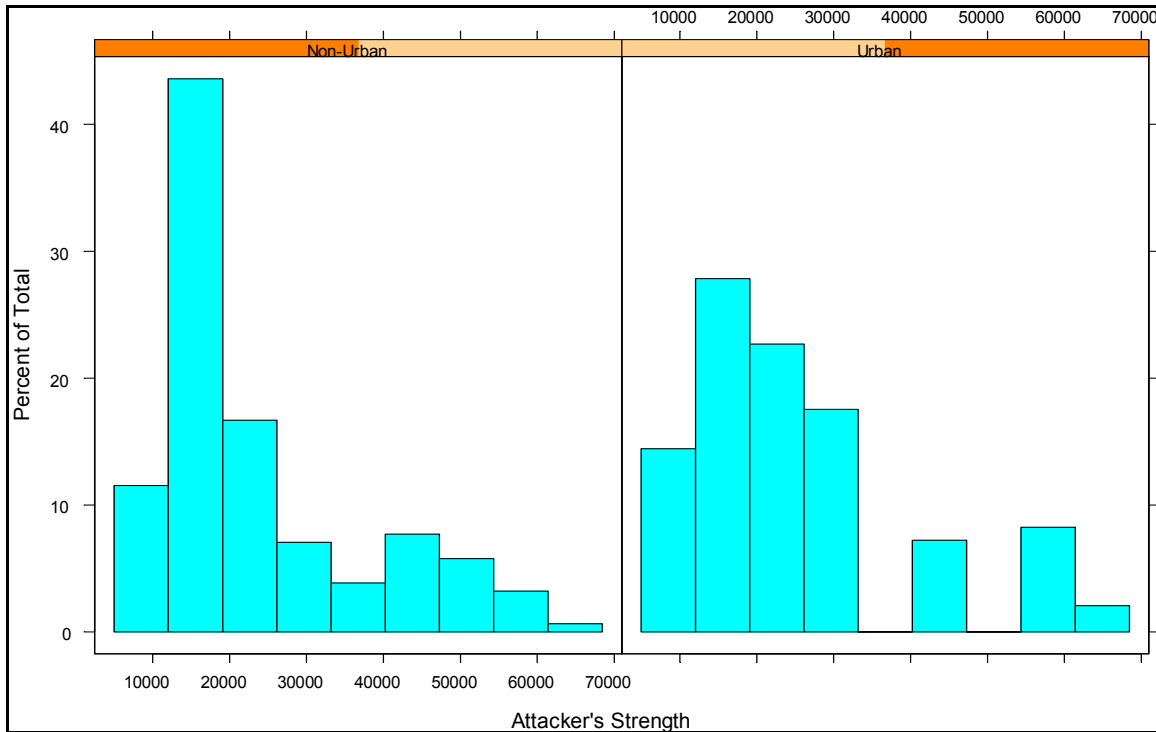


Figure 9. Histogram of Attacker's Strength Given Terrain

Since the sample size is higher than 40, the large sample test is used. The result of the test rules out rejecting H_0 , which means that there is no discernible difference between the attacker's average strength in urban and non-urban areas and this is confirmed by the Wilcoxon test. Since the data consist of division-level engagements, it is difficult to confirm or refute the assumption that urban operations need more manpower simply by assessing the attacker's strength. Only the Allied forces used more strength in urban operations compared to non-urban engagements. While the force strengths of Allies and Germans were nearly equivalent, the Soviets used more strength for non-urban engagements.

b. Defender's Strength and Terrain

It is often assumed that defenders use urban areas as a battleground if they assess that they do not possess enough forces to defend non-urban area. This might be considered in order to avoid the destruction of the cities, civilian lives and resources. It was also previously mentioned that inferior forces use urban areas as a battleground because of its advantages.

Figure 10 supports this idea. The average number of defenders in urban engagements seems to be smaller than the number of defenders in non-urban engagements. The urban engagements contain no outliers. However, a careful look at the data indicates that the Soviet defense of Kharkov and the German defenses in Cherbourg and in Aachen were implemented with a higher number of personnel than other urban engagements. Since those cities were critical to the flow of the entire campaign, the defenders defended by any means possible.

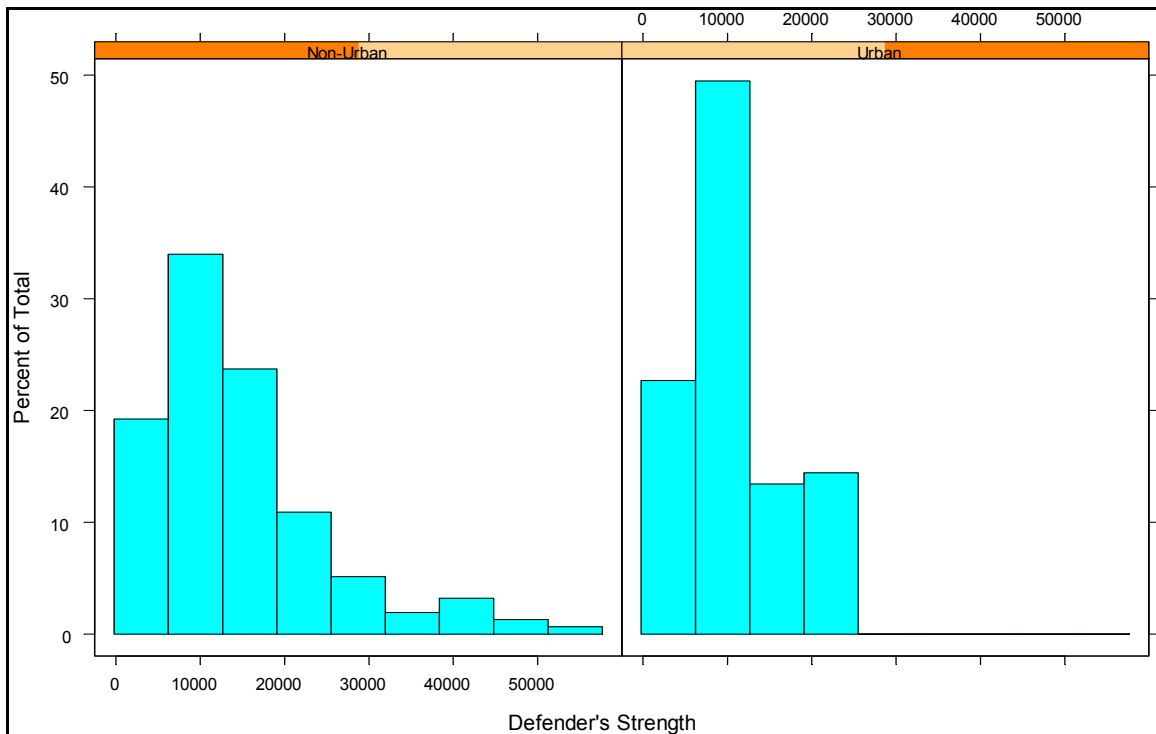


Figure 10. Histogram of Defender's Strength Given Terrain

In non-urban operations, the number of defending personnel varies much more. The extreme outliers are primarily Soviet defenses in non-urban engagements. It is plausible that the Soviets defended themselves with a large number of men because of available personnel.

A two-sided large sample test shows that the mean number of defenders in urban engagements is significantly different from the mean number of defenders in non-urban engagements. The Wilcoxon rank-sum test also supports that analysis.

It is generally known that urban areas include many resources. Since neither defender nor attacker wants to destroy the resources of urban areas, both tend to avoid urban engagements. However, when the defender did not have enough forces to defend a non-urban area, they used urban areas as a battleground.

c. Attacker's Front Width and Terrain

As urban areas are generally rich in industrial and personnel resources, they are considered centers of gravity. Thus, one might expect to see less front width for the forces attacking urban areas. Figure 11 shows the distribution of attacker's front width in each type of terrain. Although the distributions of both samples are almost the same, non-urban engagements include an extreme outlier. This extreme outlier is the Soviets attack on the German forces north of Kharkov. All the extreme outliers for non-urban engagements are German attacks on Allied and Soviet forces.

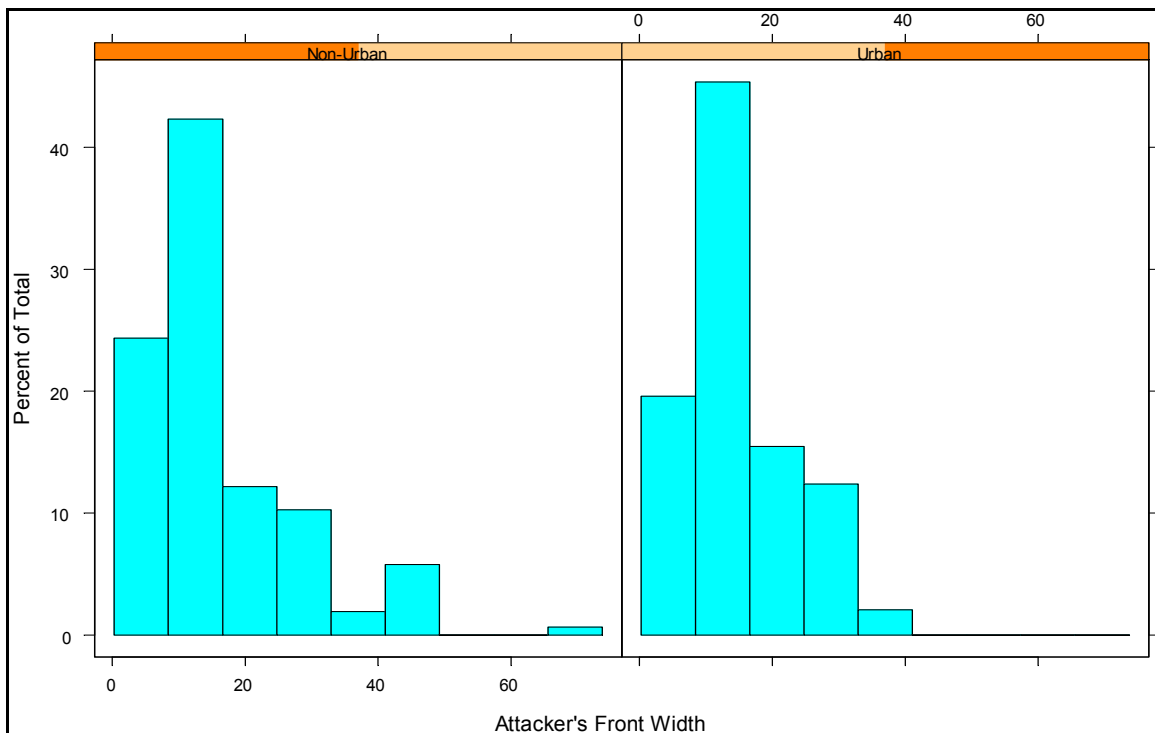


Figure 11. Histograms of Attacker's Front Width (in km) Given Terrain

A large sample two-sided test shows that the difference between the means of the samples can not be distinguished from zero. Thus, the front width does not change according to the terrain. The Wilcoxon test agrees with this result.

d. Linear Density and Terrain

There are two opposing ideas about linear density in operations. The first one proposes that since urban areas provide good concealment and cover for the defender, it is a good idea to disperse more when attacking urban areas. The opposite posits that a force is not strong anywhere if its forces are dispersed everywhere. Military planners normally consider the balance between these as an important concern for all operations.

The last two pages concluded that the attacker's strength and front width do not change for each type of terrain. The expectation is that linear density, the ratio between attacker's strength to the width of front line, does not change according to the terrain.

Figure 12 shows that each sample comes from a similar distribution, with a couple of extreme outliers on the right-hand tail. The only extreme outlier for urban engagements is from a Soviet attack at Prudyanka. All the extreme outliers in non-urban engagements are German attacks.

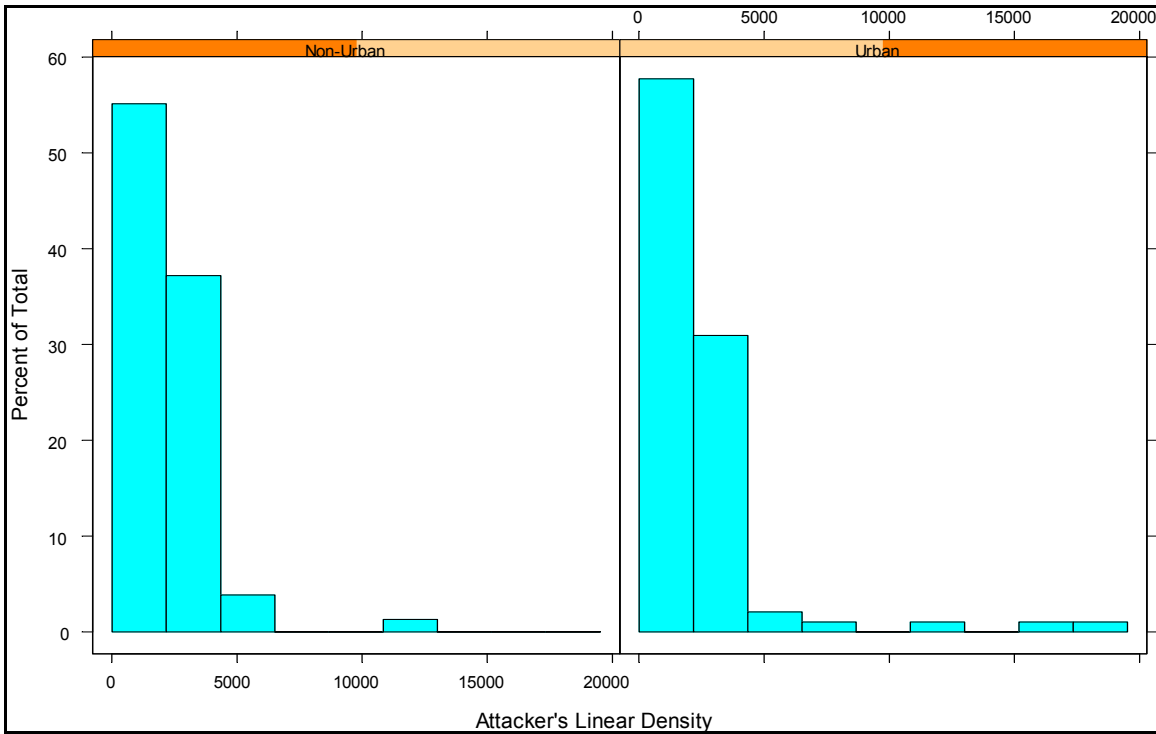


Figure 12. Histograms of Attacker's Linear Density Given Terrain

A large sample test fails to reject the null hypothesis. This means that the difference between the population means might be zero. In other words linear density did not change according to the terrain. Partitioning the above graph indicated that the linear density of the Soviet forces was almost the same for every attack they conducted, construed as Soviets being more consistent when planning their operations.

e. Force Ratio

The force ratio is the ratio of the attacker's strength to the defender's strength. Previous analyses showed that while the attacker's strength has no relationship to the terrain, the defender's strength is related. Thus, it is expected that the force ratio change according to the terrain.

Figure 13 shows that the mean and median of the two samples are different and possibly higher in urban areas. There is more than one severe outlier in both samples. The severe outliers in urban operations are attributed to the Allied forces attacks on Germans. Although there are limited number of Canadian and U.K. operations, most of these severe outliers are the result of their engagements. German forces also attacked

with high force ratios that make up some of the outliers in urban engagements. However, all outliers in non-urban engagements are attributed to German attacks on the Soviets. The Soviets did not use force ratios higher than 6:1.

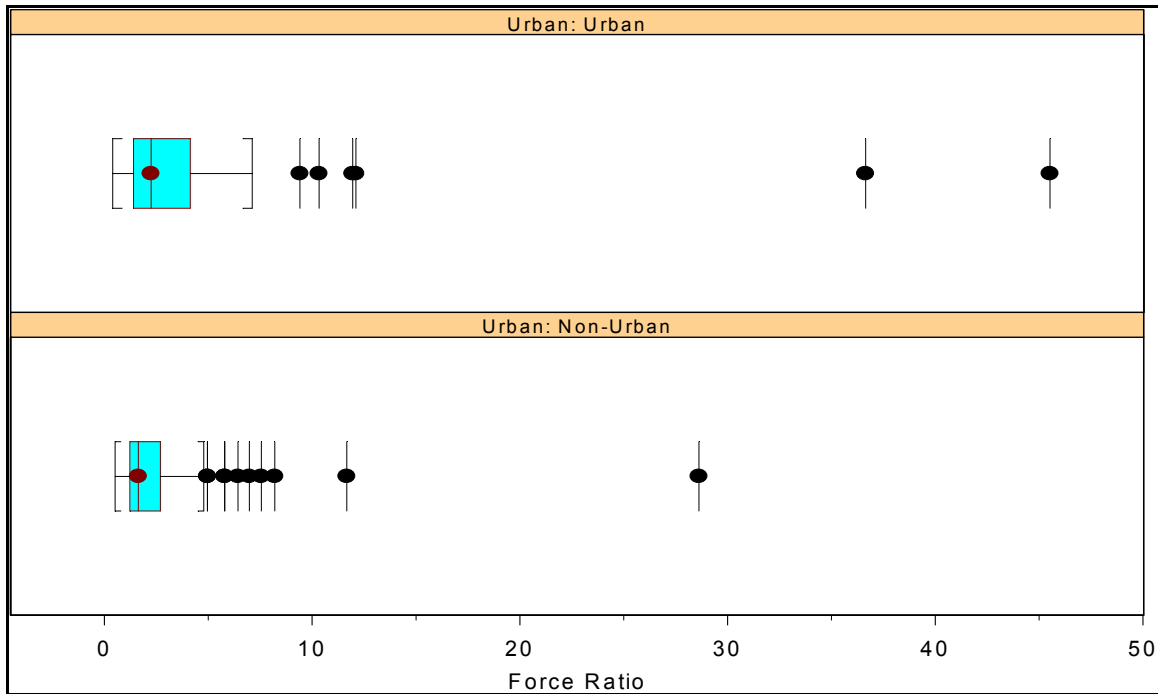


Figure 13. Boxplot of Force Ratio Given Terrain

A one-sided large sample test where the alternative hypothesis means that force ratio in urban operations is higher than it is non-urban operations rejects the null hypothesis. In other words, the mean of the force ratio in urban operations is higher than non-urban operations. The Wilcoxon test supports this finding. This conclusion also agrees with the common interpretation of urban operation in the way that these operations need more manpower.

4. Dependent Variables and Terrain

a. Attacker's Daily Casualty Rate and Terrain

Military Strategy books typically mentioned that planners must avoid urban operations because they consume a great number of resources including manpower. Moreover, urban operations are normally expected to beget higher casualty rates.

However, the boxplot contradicts these findings. It indicates that non-urban casualty rates are possibly higher and exhibit more variation. The German and Soviet forces suffered a high number of casualties when attacking in non-urban engagements. However, only the Soviets experienced extreme losses in urban operations. See Figure 14.

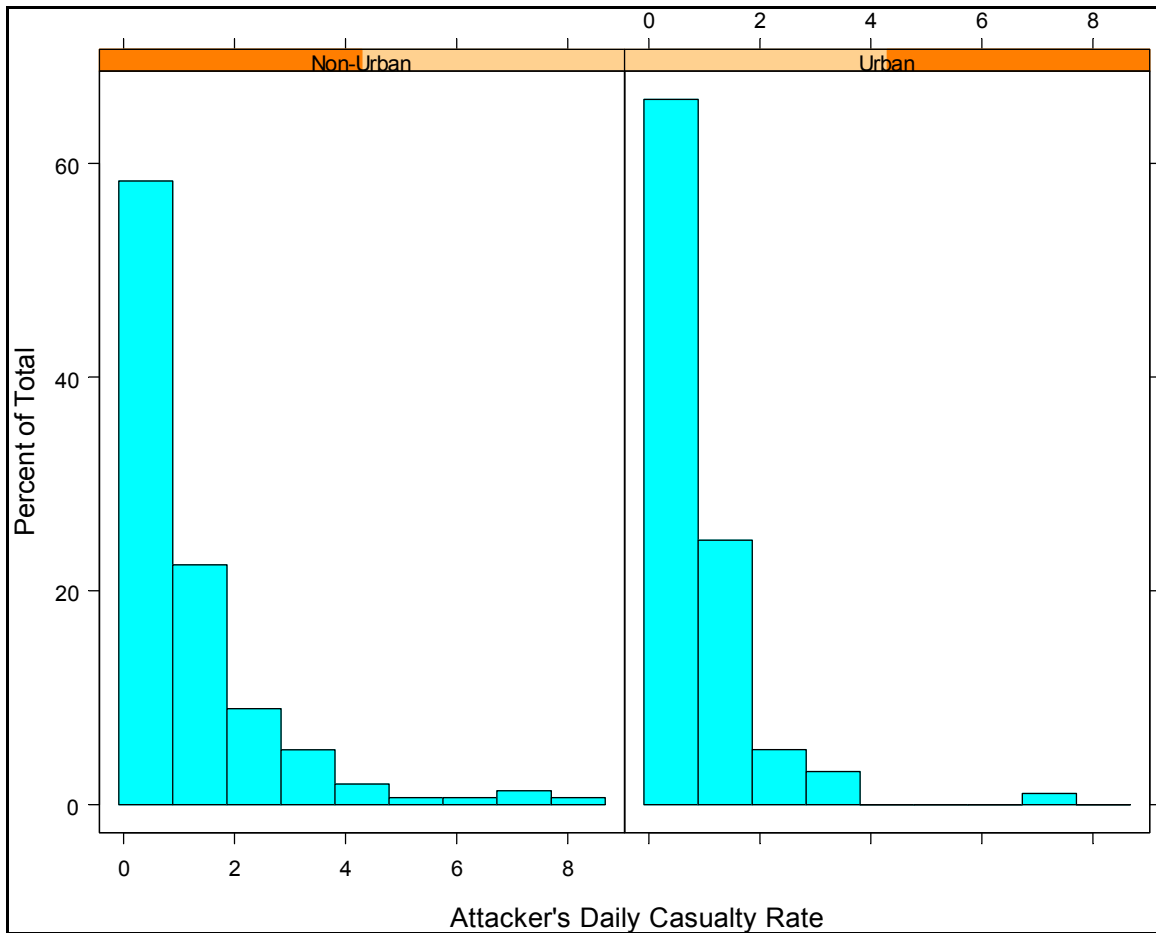


Figure 14. Boxplot of Attacker's Daily Casualty Rate Given Terrain and Country

A one-sided large sample test shows that the difference between the means of two samples is not the same. It means the attacker's urban casualty rate is lower, on average, than the non-urban casualty rate. On the other hand, the Wilcoxon test does not support this conclusion. However, since the number of sample sizes is at least twice the size of the rule-of-thumb point, the results of the large test will be used.

b. Defender's Daily Casualty Rate and Terrain

Because of the excellent concealment and cover opportunities in urban operations, fewer casualties for the defender are expected. However, in summary statistics (Appendix II), it appears that the defender's daily casualty rate in urban engagements is actually higher than in non-urban engagements. This is true for both the mean and medians of the samples.

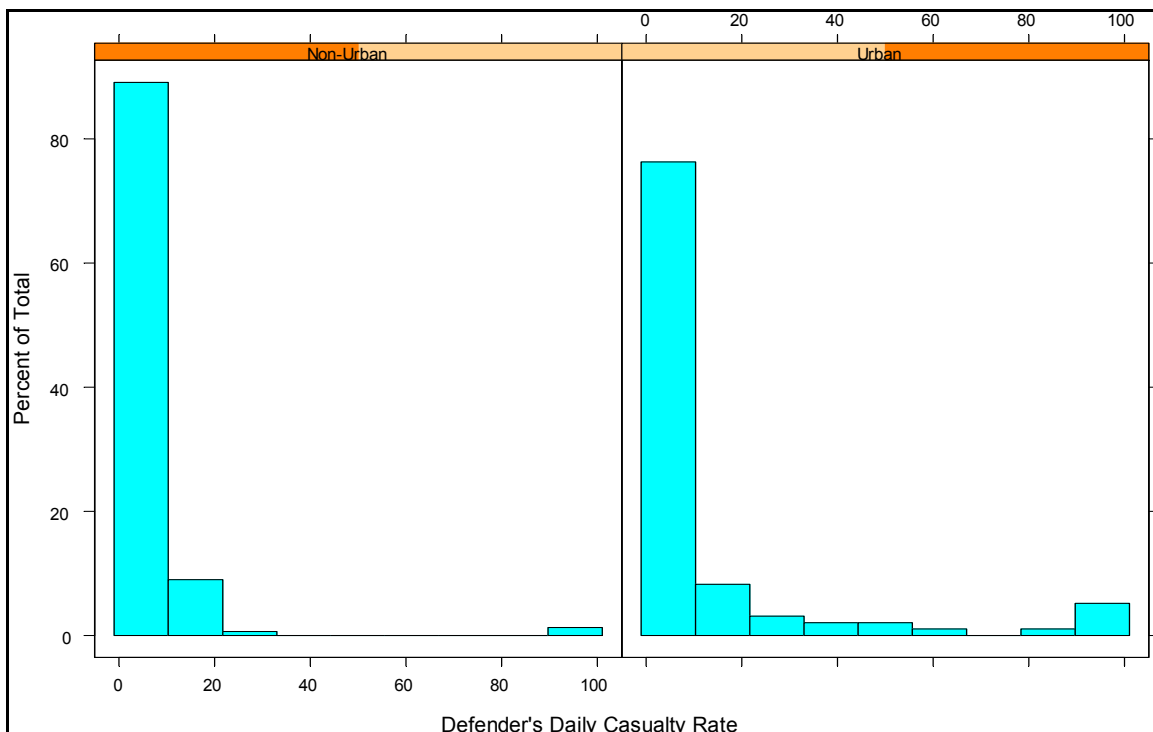


Figure 15. Defender's Daily Casualty Rate Given Terrain

A large one-sided two-sample test shows the differences in means are not zero. In other words the mean daily casualty rate of defender in urban engagements is higher. The Wilcoxon test is consistent with that analysis.

D. DISCUSSION

One concern arises after these analyses: why are the defender's daily casualty rates higher in urban operations given the apparent advantages of fighting in urban areas?

One factor leading that result is the higher force ratios of the attackers in those engagements. Actually, these high force ratios are not the result of a higher number of attackers but the fewer number of defenders.

Another factor affecting the higher daily casualty ratios in urban areas is the extensive use of non-line-of-sight weapons such as artillery and mortars. Forces had more leeway in implementing ROE in WWII than they do today.

Aerial bombardment against cities contributed to higher destruction and casualties as depicted in Figure 16. Bombers turned the cities into rubble, killing many civilians in addition to military personnel [Ref. 13]. Thus, it is possible that a high daily defender casualty rate and low attacker casualty rate could be a result of the use of indirect fire platforms.



Figure 16. Aachen After Battle [From Ref. 12]

E. CONCLUSIONS

The type of terrain does not affect the attacker's strength. However, relative to the other forces, Allied forces used more power in urban engagements. For non-urban

operations, the Soviet forces had more strength relative to the others. These balanced the differences between the attacker's force strengths in urban and non-urban operations. These differences will be analyzed more deeply in the next chapter.

The defender's force strength is generally less in urban operations. This supports the idea that inferior forces prefer using urban areas. On the other hand, Allied forces were the only forces that defended with more personnel in urban engagements.

In general, the attacker's front width does not change according to the terrain. Unlike the Germans and Soviets, the Allied forces used a wider front line in urban engagements. Linear density is not different for each type of terrain. However, Allied forces were the only forces possessing a higher linear density in urban engagements.

There is a statistically significant difference in force ratios when we compare urban and non-urban engagements. Force ratios in urban operations are higher than in non-urban operations. Allied forces had higher force ratios in urban engagements relative to German and Soviet force ratios.

The attacker's daily casualty rate is higher for all forces in non-urban operations, but the defender's daily casualty rate is higher in urban operations. However, the Allied forces had higher casualty ratios in urban areas relative to the other forces' casualty ratios.

These findings yield one other conclusion. The manner in which Allied forces behaved in urban operations is different from the behavior of German and Soviet forces. When attacking in an urban environment, Allied forces used more forces and a wider front line resulting in a higher linear density and force ratio. While defending in urban operations, unlike the others, the Allied forces used more troops and suffered more daily casualties than in non-urban operations. Thus, *the Allied forces differ from the Germans and Soviets in their understanding of the urban operations.*

III. MODELS

A. INTRODUCTION

This chapter uses “Multivariate Linear Regression” to predict the daily casualty rates for the attacker. In order to achieve this result, first, linear regression is explained. Second, the reason for selecting linear regression is discussed. Next, the models are constructed, and finally, the conclusions presented.

Linear regression assumes that a linear relation exists between a continuous response variable and the predictor variables plus noise. In other words, the response variable is a linear combination of predictor variables. If only one predictor variable exists, the model is called “simple linear regression.” It is called multiple regression if there is more than one predictor variable.

Linear regression gives an *expected* value for the response variable by using the given values of predictor variables. Since the expected value is given, it might differ from the actual value. It is assumed that this difference is random. A general representation of multiple linear regressions appears below.

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon \text{ [Ref. 11]}$$

Y represents the actual value of the observation. If the random error (residual) ε is not included, then the expected value of the response variable is calculated. Assume that the random errors are normally distributed with a mean of 0 and a constant variance σ^2 . Although the researcher knows the values of x_i s, the coefficients of this variable, β_i 's, are unknown.

Linear regression tries to fit a line through the data by using Ordinary Least Squares (OLS). OLS finds the β_i s which minimizes the sum of the squared residuals. OLS return the optimal line in this sense.

Although linear regression gives the optimal linear combination of predictor variables for estimating the response variables, the OLS assumptions must be checked to ensure the results are useful. First, the analyst verifies if the residuals come from a

normal distribution with a constant variance (homoscedasticity) σ^2 . OLS guarantees that the mean will be zero. The best way to check for normality is to look at the residual vs. fitted values plot and quantile to normal plot of residuals. There should not be a pattern in the plots.

It is important to notice that we don't require the response to be normally distributed; instead we assume that the errors, the differences between actual values and predicted values, are normally distributed.

Linear regression is also vulnerable to the effects of outliers. Even one outlier may change the intercept and slope of the line. In our case, the outliers were examined and checked for any inaccuracies. None of the outliers were deleted.

The original data has problems with the assumptions, for example, heteroscedasticity. To overcome this problem of varying variances, we used power transformations. Since all the variables are right-skewed, variables are transformed by raising to the powers of less than one. Hence, the attacker's daily casualty is raised to the power of 0.3, and the resulting variable called *XC3*. Although it is not required, it is a good idea to transform the predictor variables to reduce influence and heteroscedasticity [Ref. 14]. Attacker's front width and linear density is raised to the power of 0.2 and they are named *XFrontWidth2* and *XMenFront2*.

It is not always appropriate to include all the predictor variables in the model. There are search strategies to find the best model given the available predictor variables [Ref. 14]. Examining every possible subset of predictor variables takes too much calculation time. There are some shortcuts to overcome this problem, such as stepwise regression, which possesses two main methods: forward inclusion and backward elimination.

In forward inclusion, a model starts with one variable with the highest correlation to the response variable. Next, at each step, it adds another predictor variable, until it can no longer decrease the values of Akaike's criterion, *AIC* significantly. *AIC* shows how well the model explains the data with a penalty for complexity [Ref. 14].

In backward elimination, a model starts with all the predictor variables included. Then, it tries to delete that variable that reduces AIC maximally. The normal stepwise method is a combination of the above two methods. In stepwise regression method, a model starts with all predictor variables. It removes the predictor variable, which decreases the value of the AIC most. Thereafter, it reintroduces the deleted variables in the model and calculates which option decreases the AIC maximally. This loop keeps going until adding or dropping a term no longer significantly decreases the value of AIC . These models stop when all the variables included in the model are “significant”. In our case, we used the stepwise regression by both forward inclusion and backward elimination.

Linear regression developed prior to the computer age. Once the power of computers could be harnessed, the applicability of linear regression increased by combining it with other methods, such as using transformed data, adding interaction terms, which are very hard to do by hand. However, contrary to simple linear regression, those new models are mostly complicated and difficult to interpret. On the other hand, linear models predict better in situations with small amounts of data and low signal-to-noise data [Ref. 14].

Since it is possible to get information about the n 'th level of a categorical variable given that the subsequent $n-1$ levels are known, each categorical variable was represented as $n-1$ levels in the model instead of n levels. In other words, attacker's country was represented by *Allied* and *Soviet* excluding *German*.

The building of the models and graphs is done with the S-Plus @ 6.1 for Windows application tool. Data rows with missing values for variables were not included in the model. Significant variables in a model were selected by the method `stepAIC()` which can be attached to the S-Plus @ 6.1 from the MASS library for Venables & Ripley (2002) version 7.0. Microsoft @ Excel 2000 was used to constructing tables.

B. PREPARATION OF DATA

1. Transformation of Data

The previous chapter demonstrated that all the variables are right-skewed including the response variables. An appropriate power transformation changes the response variable to a nearly normal, or at least to a symmetric shape [Appendix B]. It is concluded that raising the attacker's casualty rate to the power of 0.3 results in the most symmetric shapes. In fact, a better transformation for the attacker's daily casualty rate is achieved by taking the natural logarithm of the column. However, since the data has zero values (no casualty), computing the logarithm results in “- infinity”. Thus, some observations would be lost using this approach.

Among the predictor variables, only the number of men at the front and the size of attacker's front width are transformed. Although this transformation is not necessary, it decreases heteroscedasticity and the number of influential points. After several trials, the attacker's casualty rate was raised to the power of 0.2.

$$\text{transformed variable} = \text{variable}^{\text{power}}$$

Attacker's casualty rate was raised to the power of 0.3 and named *XC3*, and the attacker's front width and linear density were raised to the power of 0.2 and named *XFrontWidth2* and *XMenFront2* respectively.

2. Interaction Terms

Linear regression allows users to add interactions that result from the product of predictor variables. Thus, they take into account the effects of one predictor variable depending linearly on another predictor variable. This may result in adding a non-significant variable to the model only because its interaction with another variable is significant. During the initial descriptive analyses, it is ascertained that the attacker's strength does not change according to the terrain. However, further analyses showed that the attacker's strength changes for different countries in different terrain. Interaction terms capture these details.

Since it was ascertained that differences existed for the attacker's strength for each country, it is possibly necessary to include the two-way interactions in the model. Adding a three-way interaction should provide better results. It is possible that more than

two variable interactions affect the attacker's daily casualty rate. However, including more than two-way interaction cause a large decrease in the degrees of freedom and the interpretability. Hence, we used two-way interactions in our model.

Categorical variables in a model affect the regression line differently. Categorical variables as the main effect change the intercept. On the other hand, categorical variables used as interaction terms affect the slope. Also, the interaction of two categorical variable changes the intercept. If the categorical variables are used as the main and interaction terms, the slope and the intercept change. Thus, it is possible to find different lines for each level of the categorical variables.

3. Singularity

Singularity problems occur when one column is the same as another column or a linear combination of more than one column. In the case of singularity, it is impossible to find the coefficients of the variables. The solution is to find that column and remove it from the model. It is possible to believe that deleting an entire column would negatively affect the model. This is not correct because the linear combination of the other columns provides the same information to the model.

In the case of this thesis, it is also possible to think that including the attackers and defender's strengths and force ratios cause a singularity problem. Recalling that force ratio is the ratio between the other two, it is possible to believe that the likelihood of removing the force ratio and obtaining the information from the other two exists. However, this is not true. Since force ratio is not a linear combination of the other two, it is impossible to get the effects of these variables will be different and it will not cause a singularity problem.

Unexpectedly, singularity occurs when the attacking forces and defending force's countries are evaluated together in the full model with all variables including the two-way interactions. This happens because of many zero values in the matrix. For example, no country attacks its own forces. Thus, the model does not include attacking and defending countries together. For simplicity and interpretability reasons, only the attackers' country has been used.

However, it is still possible to obtain the information about the defending force by only using the attacking forces' information. Recall that each of the two country variables has the same levels, *Allied*, *German*, *Soviet* and there is no data about Allied forces attacking the Soviets [Figure 3]. If the attacking country is known, then in two out of three cases, the defending country is definitely identified. If the attacking country is known to be the Allied or the Soviet, then it is definite that the Germans are the defending country. On the other hand, if the Germans are attacking, the identity of the defender is unclear.

Even then, it is still possible to acquire some information about the defender during on a German attack. Since the Germans did not attack the Allies in urban areas (Figure 3), it is possible to obtain more insight if the terrain is known. If the Germans have attacked in an urban area, then it is definite that the Soviets were the defenders. It is indeterminate if Germans attacked in non-urban areas, which represents almost 29% of the data.

C. ATTACKER'S DAILY CASUALTY MODEL

The focus of this section is to construct a linear model able to predict the attacker's daily casualty rate by using the variables. The predictor variables are those that should be available before the beginning of the engagement. Among the columns in the available data set, the attacker's nationality, strength, the width of the front line, the number of men per km. in front (linear density), the terrain, the defender's strength and force ratio are selected as the predictor variables. Both a model without interactions and a model with interactions will be analyzed.

1. The Model without Interactions

To find the best model, first, a model with all the transformed variables is used to fit the model. Then, using Akaike's criterion, the preferred model was selected. Residuals vs. fitted values, actual values vs. fitted values, Cook's distance and Quantile-Normal help analyze the assumptions.

The model without the interactions (Appendix D, part A) suggests that the inclusion of the defender's strength and the number of men in front is not significant. This is actually not surprising, because in the presence of other variables, these two can

be calculated. Thus, the model can derive almost the same results without the defender's strength and the number of men in front. For example, if the force ratio and the attacker's strength is known, then it is very easy to find the defender's strength.

There are two influential points (49 and 173) in the model. These two points are different from the others. While the first one has higher casualty rates and density, the latter has lower casualty rates. Although removing these two points results in a 2% improvement in R^2 , the coefficients of the model does not change noticeably. In addition, Cook's distances for these are less than 0.05, which means that they do not greatly influence the model. The presence of 5 data points with the attacker's casualty rate equal to zero gives the highest error terms which threatens the normality assumption. Otherwise, it is possible to assume that the model without interactions agrees with the assumptions of OLS.

However, the model without interactions only explains 26% of the variability (Appendix D). The implications of the best model without interactions are, considering the other variables remain constant:

- Given the same terrain, force ratio and front width, the Germans had lower casualty rates than the Allies and Soviets. The Soviets had the highest casualty rate
- The attacker's daily casualty rates are smaller in urban operations than in the non-urban ones;
- The higher the force ratio, the fewer the attacker's casualties;
- The higher the front width, the fewer the attacker's casualties;
- Among all, the most important variables are the attacker's country, terrain, force ratio, attacker's strength and front width.

Although the implications of this model seem reasonable, the noise level is high. However, given that this data is real-life data, the R^2 seems acceptable.

The low capability to explain the data also suggests that there might be other factors that are not included in the model. Since the data set used all the available variables, adding new variables needs detailed work and data collection. However,

adding interactions can help the model. At the end of Chapter II, it is concluded that although some variables do not differ according to the terrain, they differ according to the terrain **and** the attacker's country. Thus, two-way interactions are included in the model.

2. The Model with Interactions

The same path is followed to find the best fit for the data set including the interaction variables. First, the model started with all predictor variables including the two-way interactions. Then, the significant variables were selected by the Akaike criterion. Afterwards, the graphs were analyzed to validate the assumptions.

The model with interactions has 35 variables. By using the Akaike criterion, 11 are eliminated. Nine observations are deleted because of missing values. Thus, the candidate for the best model has 24 predictor variables and an intercept. There are 244 total degrees of freedom, 219 for residuals. This model explains 40% percent of the variability of the data. This is a 50% improvement in exchange of losing 18 degrees of freedom. Since the data set has 244 available data rows, losing 18 degrees of freedom is acceptable. The residual standard error is 0.2669 (Appendix D part B).

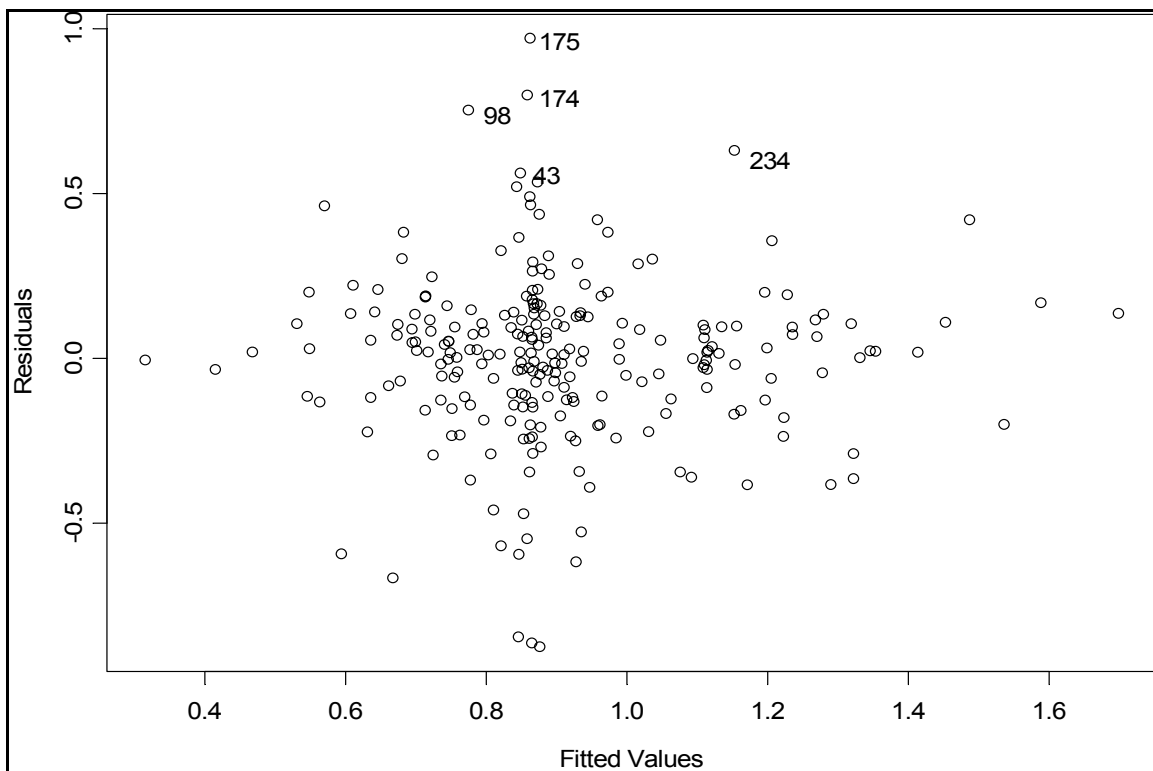


Figure 17. Plot of Residuals vs. Fitted Values in the Model with Interactions

The plot of residuals versus fitted values (Figure 17) does not show any pattern and heteroscedasticity. All data points are gathered around zero. Most of the data points are within two standard deviations from the mean. There are five points (labeled) outside this range. For a 95% confidence level, the expectation is that more than 12 points are outside this range. Thus, these five outliers are not considered a violation of OLS's normality assumption.

The graph shows that the model has a tendency to underestimate occasionally. For outliers, 48 and 107, the reason is the attacker's success in achieving a low casualty rate with a low force ratio. The other outliers 174,175,234 (rows 183,184 and 243 in the data set) are the result of the attacker's high casualty rates although they had the advantage of a higher force ratio and small front width. Thus, these unusual engagements appeared as outliers. However, these do not affect the assumption that errors are independently and identically distributed. The graph also supports the assumption that the variance remained constant.

Figure 18 shows that the assumption of this thesis, that linear regression can be used to predict the response variable, is plausible. The graph shows that the fitted values are related almost linearly to the actual values. The most influential point is point 228, whose Cook's distance is 0.14. This engagement has the biggest effect on the regression line. This engagement occurred between forces with fewer attackers and defenders than usual. In addition, the attacker did not lose any personnel in this particular engagement. However, deleting this observation does not significantly affect the regression line.

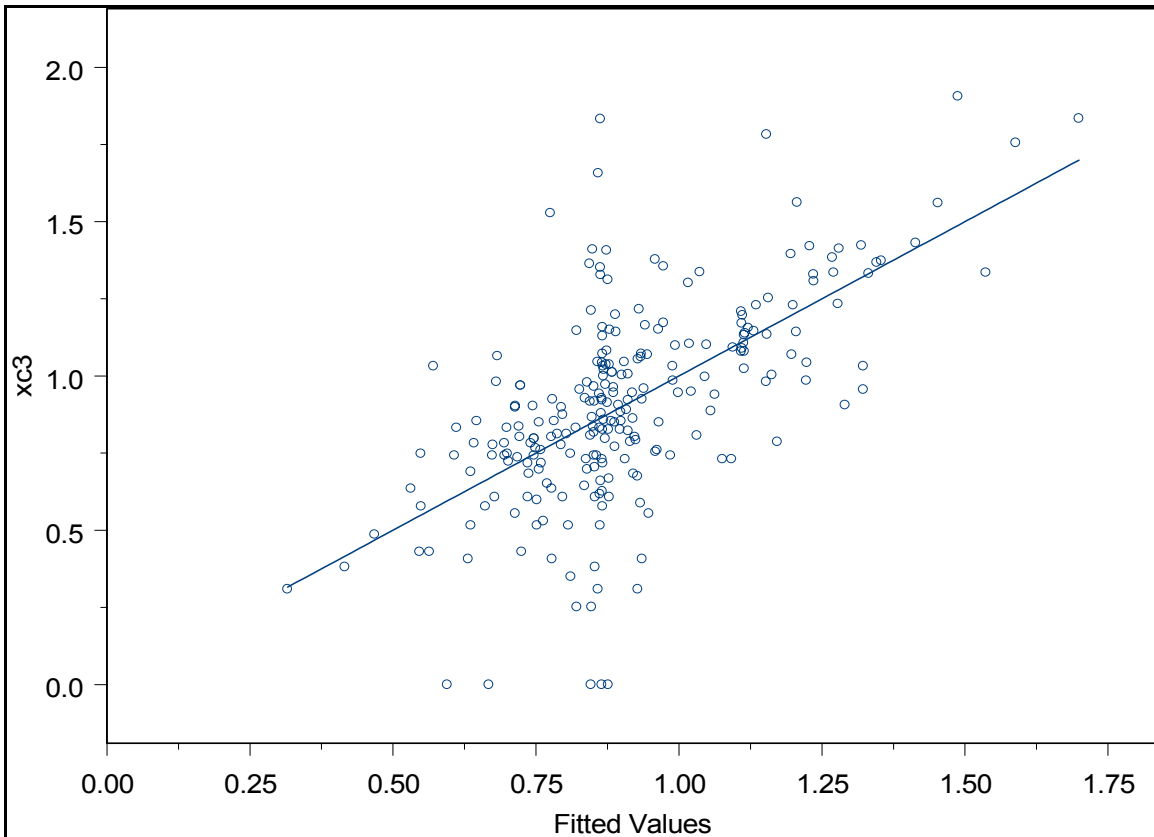


Figure 18. Plot of Actual Transformed Values vs. Predicted Values

This model, as opposed to the model without interactions, includes all the predictor variables as the main effects. Although some are not significantly important, they are included in the model because of their effects in the interactions.

This model provides more insight into the operations but it should be viewed carefully. It is difficult to make conclusions in a model with interactions and transformations. In order to obtain some insight, each variable will be analyzed one by one including the variables with which they interact while holding the others constant.

The primary concern of the analysis in each part is the effects of terrain on DCR given the particular variable. Since the terrain interacts with all the numerical predictor variables, for each one, it is possible to analyze the effects of terrain.

3. Effect of the Terrain on the DCR

In order to analyze the effect of terrain, it is crucial to consider the coefficients of the variables with which it is interacting, namely, the force ratio, the attacker's men in front per km. and the width of the front line. Considering that the other variables remain

constant, the effect of urban areas will be a number for each level of the other categorical variable, that of the attacker's country. Since *Urban* interacts with every predictor variable but the defender's strength [Appendix D], the equation that would exhibit the effect of the urban terrain is the same as the output at Appendix D except for now the value of the intercept is 9.7827 instead of 10.3621. There is also a slight change for the coefficient of *XFrontWidth2*, but it is negligible

To give the idea about how the terrain affects DCR, values are calculated by using the mean of every variable. Figure 19 shows the results. It is obvious that, the difference between each terrain type is not a constant number for each country. The DCR is higher in non-urban engagements than in urban operations. In addition the difference between terrain types for Soviet forces is more significant than it is for Allied and German forces. The effect of Urban for Allied and German forces is almost the same.

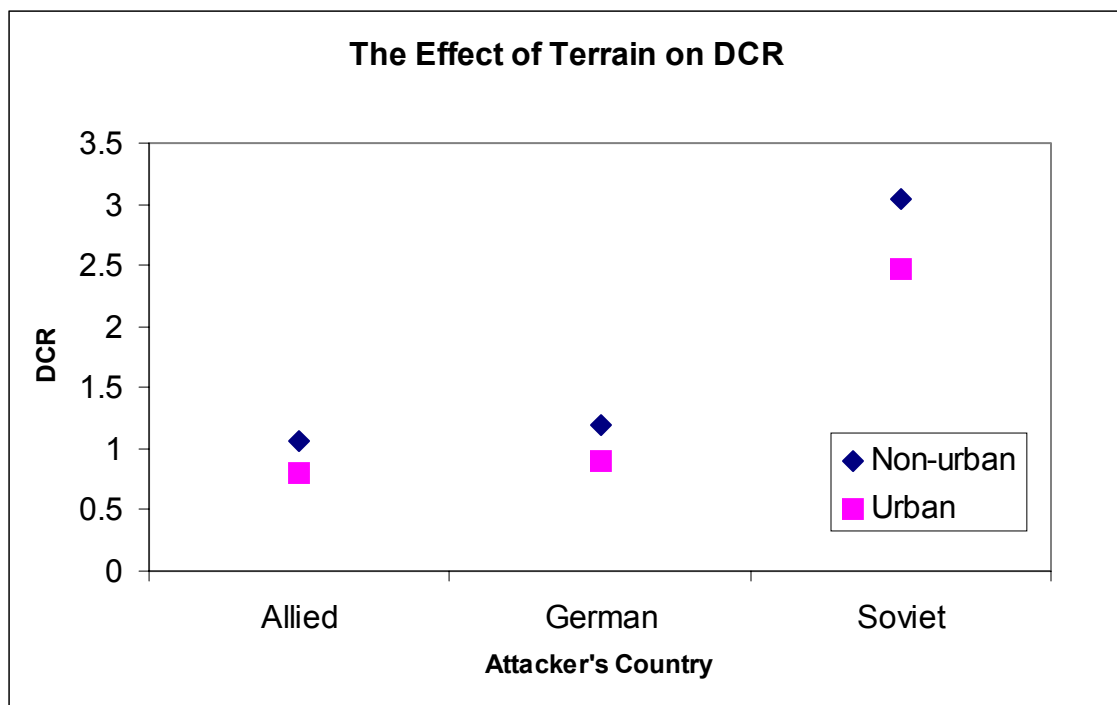


Figure 19. The Effect of Terrain on the DCR

Further analysis presenting a different a perspective is possible. For each country, it is possible to find which variables affect this difference. The subtraction of the equation for the non-urban engagement from the urban engagement yields the following equation

that demonstrates which variables affect the difference due to terrain in DCR for Allied forces. The variables that affect the difference for the other forces is the same but their coefficients are different.

$$XC3_{non-urban} - XC3_{urban} = -9.2886 + 0.0234ForceRatio - 3.5611XFrontWidth2 - 1.2877XMenFront2$$

The equation shows that force ratio, front width and linear density have an impact on the difference. The difference between the daily casualty rates (DDCR) decreases as the front width and linear density increases. However, the effect of these variables cannot be examined by fixing the other variables because of the derivation of the data. Using the derivation of data as an advantage, the aforementioned equation becomes the following:

$$XC3_{non-urban} - XC3_{urban} = -9.2886 + 0.0234 \frac{XStrength}{YStrength} - 3.5611XFrontWidth2 - 1.2877 \frac{XStrength^{0.2}}{XFrontWidth2}$$

It is easier to interpret this equation. If the attacker and the defender's strength are fixed to their means, the effect of front width becomes clear. As front width increases, the difference between the daily casualty rates in terrain decreases. A one-unit increase in the defender's strength increases the difference between the casualty rates. In other words, increasing the number of defenders has more impact in the non-urban areas. A similar analysis holding the defender's strength and front width constant gives the following; an increase in the attacker's strength increases the difference between the non-urban casualty rate and the urban casualty rate. This also means that increasing the attacker's strength has more effect in non-urban areas.

Another way to determine whether *Urban* is significantly important in the model is to conduct an *F* test. An analysis of variance (ANOVA) test compares the full model with the model without urban-related terms. In this case, the null hypothesis implies that *Urban* is not needed in the model. This also indicates that the attacker's DCR does not change according to the terrain. The result of the ANOVA test leads to the rejection of null hypothesis. Thus, *Urban* is a statistically significant variable in the model and the DCR changes according to the terrain.

4. The Effect of the Attacker's Country on the DCR

According to the model, the type of terrain does not make any difference on the effect of the attacker's country to the DCR since an interaction term is not present. The

attacker's strength, front width, and number of men per km. are associated with the daily casualty rates. The other variables have the same effect for each country. Fixing the variables that $XCountry$ is not interacting made it possible to formulate the relationship between the daily casualty rates of Allied, Germans and Soviets forces and the other predictor variables. The value of the c is the same for each country. The variables $XFrontWidth2$, $XMenFront2$ and $XStrength$ produces the differences in DCR for each country.

$$XC3_{Allied} = 1.6351 + 1.8657XFrontWidth2 + 0.8533XMenFront2 - 0.00048XStrength + c$$

$$XC3_{Germans} = 10.3621 - 1.2336XFrontWidth2 - 0.0974XMenFront2 + 0.0005XStrength + c$$

$$XC3_{Soviets} = 14.2323 - 2.3686XFrontWidth2 - 0.5504XMenFront2 + 0.00053XStrength + c$$

where,

$$\begin{aligned} c = & -10.9237Urban + 4.8 * 10^{-5} YStrength - 0.49ForceRatio \\ & - 3.22 * 10^{-5} Urban : XStrength - 0.0234Urban : ForceRatio \\ & + 3.5611Urban : XFrontWidth2 + 1.2877Urban : XMenFront2 \\ & - 0.0001XStrength : XFrontWidth2 - 4.39 * 10^{-5} XStrength : XMenFront \\ & + 2.83 * 10^{-5} YStrength : XFrontWidth2 + 0.1591ForceRatio : XFrontWidth2 \\ & + 0.0444ForceRatio : XMenFront2 - 1.309XFrontWidth2 : XMenFront2 \end{aligned}$$

Remembering that front width and linear density are inversely proportional, increasing the front width decreases the linear density. However, since both coefficients have the same sign, the effect of changing one does not change the conclusion. For allied forces, it is expected to see higher casualty rates in higher front widths. On the other hand for Germans and Soviets, the daily casualty rate decreases if the front width increases. The attacker's strength has the same effect on all of them. An increase in the attacker's strength increases the DCRs.

The above equations do not answer the question: which country has, on average, a higher casualty rate? However, it is possible to compare the casualty rates among each country. Table 5 exhibits the 95% confidence levels for the differences in the DCR. The difference between the German and Allied DCR is not significant. In other words, we expect to see the very similar casualty rates for German and Allied forces. On the other

hand, German and Allied forces had lower casualty rates relative to Soviet forces. On average, Soviet forces had 0.05% more casualty rates than German forces and 0.15% more casualty rates than Allied forces. The difference between the Allied and German DCRs is not statistically significant. This conclusion agrees with Figure 19.

	Estimate	Lower Bound	Upper Bound
German-Allied	0.00171	-0.00003	0.02770
German-Soviet	-0.05246	-0.19996	-0.00533
Allied-Soviet	-0.14475	-0.68320	-0.00724

Table 5. Difference in the Daily Casualty Rates Between Countries

5. The Effect of the Attacker’s Strength, the Defender’s Strength and the Force Ratio on the DCR

The effect of the attacker’s and defender’s strength and force ratio will be examined together because of the way force ratio derived. Usually, the best way to see the effect of one variable is to fix the other variables and change the value of the variable concerned. However, in this case, it is not possible. Although there is no interaction term between the attacker’s strength and force ratio, a change in the attacker’s strength will also change the force ratio. Thus, in the analysis for this thesis, it is mandatory to analyze those relationships, which will also make the analysis more complicated. For example, although the model does not specify an interaction between force ratio and terrain, there is an interaction between attacker’s strength and terrain. Therefore it is useful to examine the force ratio separately for each terrain type.

The method to analyze the effect of these variables is as follows. First, the relationship between each variable will be defined. Then, all the variables that do not have any interaction or relationship with the variable under examination will be set to their mean values. Finally, for each level of categorical variables, the expected daily casualty rate, DCR, will be calculated. Graphs using these values will help in making conclusions and comparisons.

In order to analyze the effect of the attacker’s strength to the DCR, the defender’s strength and attacker’s front width are held constant at their mean values of 12755 and 15.35 Km.. The DCR is calculated by changing the attacker’s strength within its range in

the data set. Since there is only one data point under 10000, the value of the attacker's strength varied between 10000 and 60000. Figure 20 shows the results of these calculations. Vertical lines show the first and third quantiles of the attacker's strength to explain this idea better.

The effect of the attacker's strength is almost the same for each country in urban areas. An increase in the attacker's strength up to 35000 for Allied and German forces and up to 50000 for Soviet forces increases the DCR. As the size of attacker's strength becomes greater than these numbers, their DCR start decreasing. In order to achieve the lowest casualty rates, attackers should use either a higher or lower number of soldiers in urban operations. Although using fewer soldiers might result in lower casualty rates, it does not necessarily result in a victory.

The decrease in the DCR for the lower values of the attacker's strength is counterintuitive. The expectation is to have more casualties as strength decreases while other variables remain constant. The lower number of troops means a lower force ratio and lower linear density. In the case of a lower force ratio, higher casualty rates are expected. On the other hand, lower casualty rates are expected with lower linear densities. Thus, the conclusion might be that in *urban operations, the effect of linear density is more important than the effect of force ratio.*

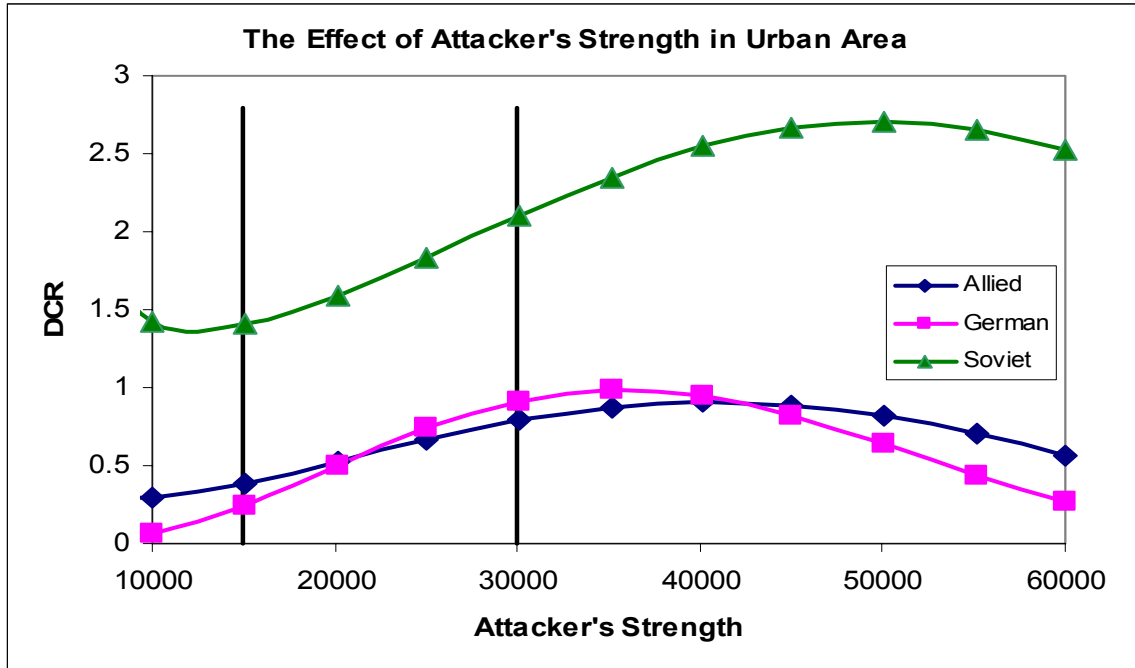


Figure 20. The Effect of the Attacker's Strength on the DCR in Urban Areas

The effect of the attacker's strength is different in non-urban areas than in urban areas (Figure 21). In this graph, the attacker's strength is varied between 5000 and 60000 because of the existence of more than one data point where the attacker's strength is less than 10000. The effect of the attacker's strength is almost the same for each country. An increase in the attacker's strength up to its median value decreases the DCR. Any increase over the median no longer decreases the DCR. The DCR starts decreasing where the attacker's strength is more than the median value. An increase in the attacker's strength means a both higher force ratio and a higher linear density. These two variables are expected to have an opposite effect on the DCR. However, Figure 21 implies that their effects are somewhat balanced. Thus, it is possible to conclude that *the effect of force ratio and linear density is almost the same in non-urban areas.*

In order to achieve a lower casualty rate, it would be more advantageous for an attacker to use a higher number of soldiers in urban operations. In non-urban areas, it would be more advantageous for attackers to use more than 15000 soldiers. Any attack on non-urban areas with a force size less than 15000 could be costly in casualty rate terms.

Both Figures 20 and 21 support the conclusion that under the same circumstances, Soviet forces had the highest DCR and German and Allied forces had almost the same DCR as one another in each terrain type.

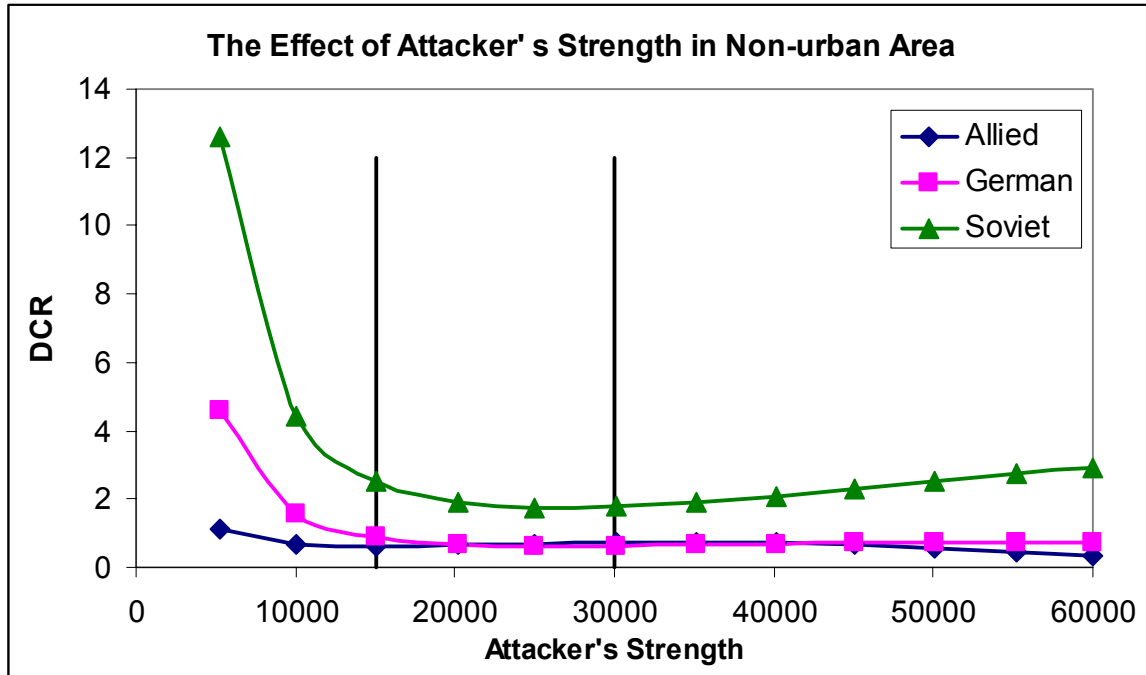


Figure 21. The Effect of Attacker's Strength on DCR in Non-Urban Area

In order to analyze the effect of the defender's strength on the DCR, the attacker's strength and front width are held constant at their mean values of 24257 and 15.35 Km. For values in the range of the defender's strength, the DCR is calculated. Figure 22 is the result of these calculations. The first and third quartiles of defender's strength are also shown to explain the idea better.

It is likely to see a higher DCR while attacking a stronger defender in urban operations. However, the slope of the increase goes to zero, as the defender's strength gets closer to its third quartile value, 15943. Thus, adding one more person to the defender where its strength is less than 15943 increase the attacker's DCR. This increase is higher when the defender's strength is relatively small. While defending against an attacker who

has 24257 men and using a 15.35 km. front width, the optimal number of soldiers for the defender is 15943. Adding additional personnel to that force does not increase the attacker's DCR, but might change the outcome of the battle.

The effect of the defender's strength is the same for each country. It is more likely that the Germans caused higher daily casualty rates to the Soviet forces than to the Allied forces. For example, if German forces were defending with 15000 soldiers in an urban area, the DCR for the Allied would be expected to be 0.97% while it would be expected to be 2.21% for Soviet forces.

The effect of the defender's strength in non-urban areas is the same as in urban areas. This agrees with the results of the regression model. In the model, there is no interaction term between the defender's strength and terrain. It is possible to say that, in general, when the defender's strength is between its first and third quartile, increasing the defender's strength increases the attacker's casualty rate. Finding this conclusion would be impossible if the effects of force ratio were not considered. The regression model suggests that the effect of force ratio where the front width is 15.35 km., is almost zero. Those effects are the same in urban areas as well.

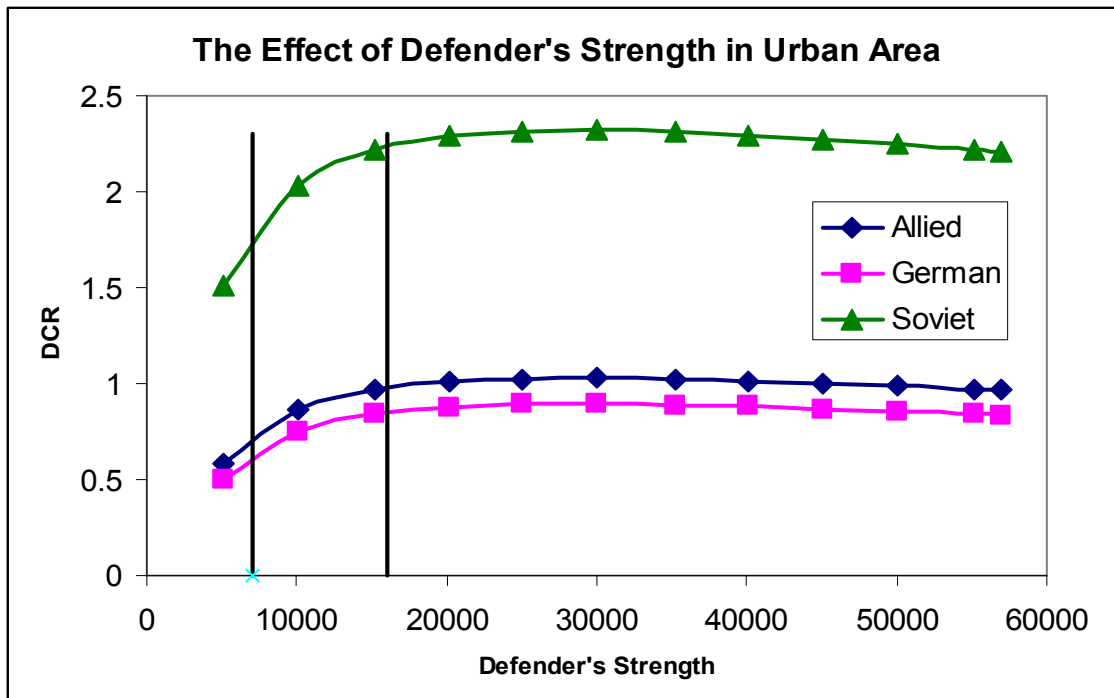


Figure 22. The Effect of the Defender's Strength on the DCR in Non-Urban Areas

Since both the attacker and defender's strength affect the value of the force ratio, it is very difficult to find the effects of force ratio on the DCR. An increase in the force ratio comes out as a result of an increase in the attacker's strength or a decrease in the defender's strength or both. To overcome this problem, first, the force ratio will be changed by holding the defender's strength constant and changing the attacker's strength and second vice versa. The values of the first and third quartile of the force ratio will be added to see the general picture.

Figure 23 shows that changing the force ratio by means of changing the defender's strength does not at all affect the DCRT. On the other hand, changing it by varying the attacker's strength does affect the DCRT. *As the force ratio increases, the daily casualty rate decreases.* Although the effect is different for each type of terrain for values of force ratio outside the first and third quartile, it is possible to state that terrain does not change the effect of the force ratio except for an attack in an urban area by changing the attacker's strength. These conclusions agree with the findings for the attacker's strength and the defender's strength.

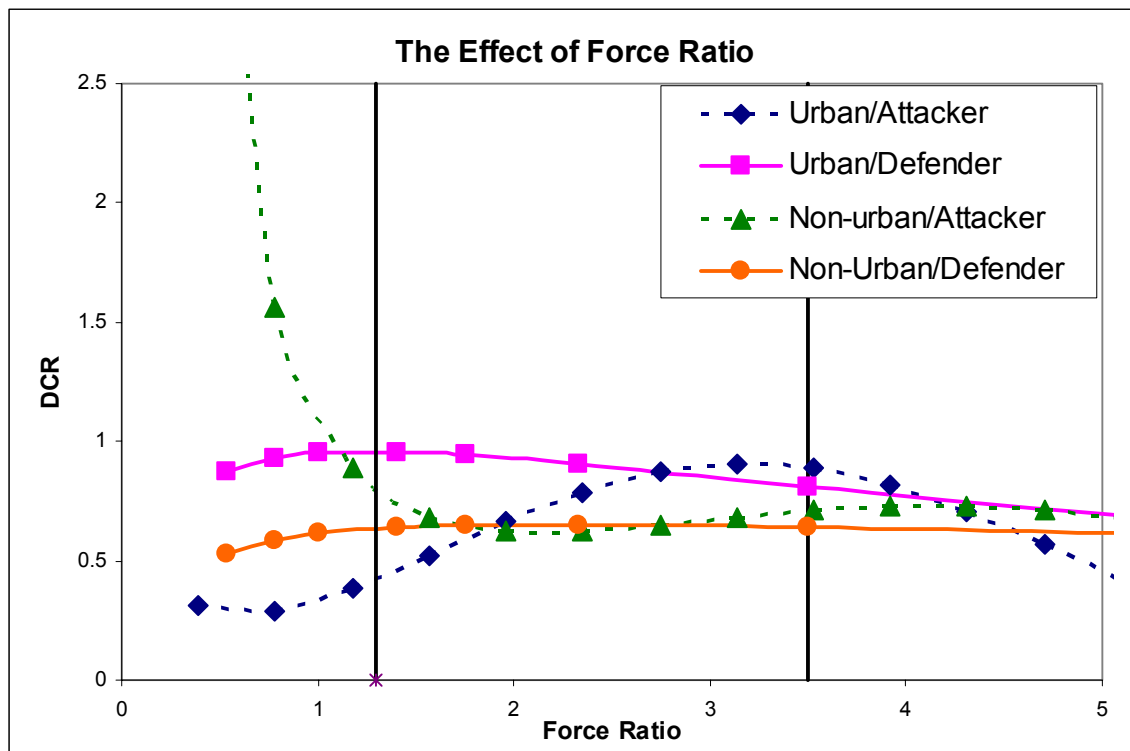


Figure 23. The Effect of the Force Ratio on the DCR (Terrain/Varying Force)

6. The Effect of the Attacker's Front Width on the DCR

The effect of the attacker's front width to its casualty rate can be analyzed by fixing the variables with which the attacker's front width does not interact. Fixing these variables results in the following equation:

$$\begin{aligned}XC3 = & 10.3621 - 8.727XC_{CountryAllied} + 3.8701XC_{CountrySov} - 10.9237Urban \\ & + 0.0005XStrength - 0.4902ForceRatio - 0.0974XMenFront2 \\ & - 1.2336XFrontWidth2 + 3.0993XC_{CountryAllied}XFrontWidth2 \\ & - 1.135XC_{CountrySov}XFrontWidth2 + 3.5611Urban : XFrontWidth2 \\ & - 0.0001XStrength : XFrontWidth2 + 0.1592ForceRatio : XFrontWidth2 \\ & - 1.309XMenFront2 : XFrontWidth2\end{aligned}$$

The coefficient of front width and its interactions cause complexity in the analysis. A figure in which all the above variables are set to their mean makes the comparison easy. The attacker's strength is set to 24257 and the defender's strength to 12755. The force ratio is not set to its mean; rather, it is set to the ratio of the mean attacker's strength over the mean defender's strength. Although the maximum front width is 73, there is only one point larger than 45. The attacker's front width is therefore varied between 5 and 45 Km.. The author assumes that even at the maximum value, the trend does not change.

Figure 24 shows the effect of front width on the DCR in urban operations. The effect is different for Allied forces than the other two forces. Front width and DCR are inversely related for Soviet and Germans forces. Increasing the front width results in lower casualty rates for Soviet and German forces. On the other hand, the effect of front width on Allied forces is different. Although its effect is the same in the first quantile, its effect is the opposite for the rest. For values of front width greater than 10 Km., the effect is proportional. As the front width increases, the casualty rate increases.

The front width is extremely important for the Allies in urban areas. Under the same conditions, the higher the front width, the higher the daily casualty rate for the Allies. This is not true for the Germans and Soviets.

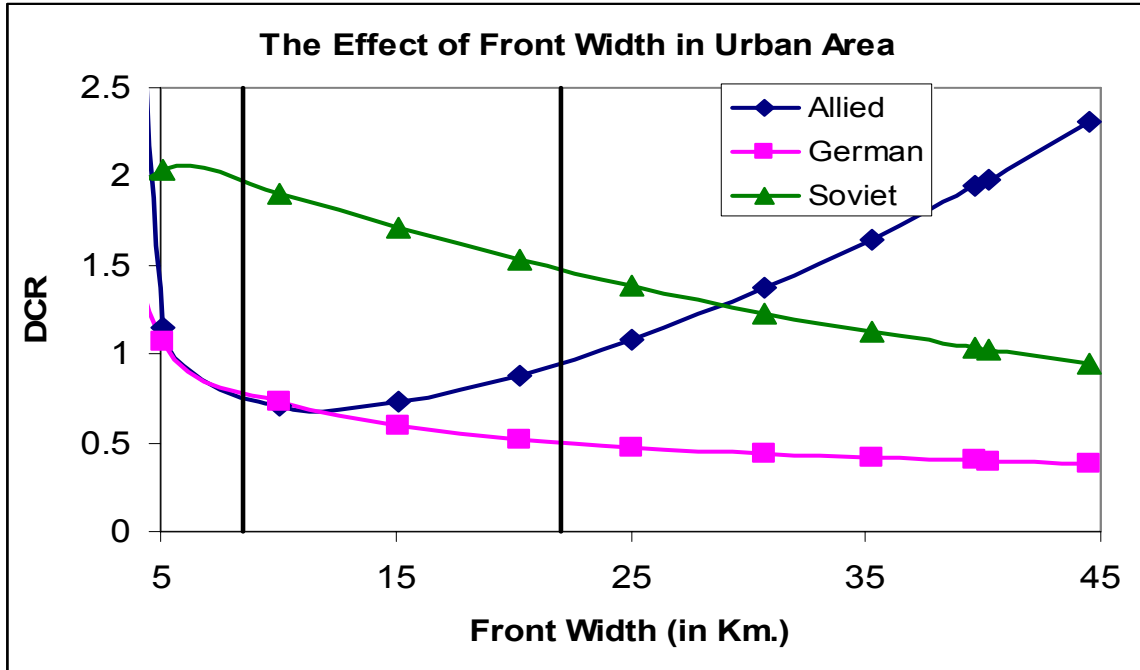


Figure 24. The Effect of the Front Width in Urban Areas

To analyze the effect of front width on the attacker’s DCR in non-urban operations, the coefficients related to *Urban* are set to zero. The attacker and defender’s strength is set to their mean. Linear density is calculated accordingly.

A graph where every variable but the front width changes exhibits the trends clearly. Figure 25 explains the effect of front width on the attacker’s DCR. The effect of front width can be looked at in two parts, the effects where the front width is less than 8.5 Km. and where it is higher. In the first part, the effect of increasing the front width increases the DCR of each country. The increase for Soviet forces is the highest. The effect for the Allied and Germans forces is almost the same.

In the second part, it is apparent that the effect of front width on the Allies is different than the others. While a one-unit increase in front width decreases the DCR of the Germans and Soviets, it is constant for the Allies. In order to achieve low casualty rates, German and Soviet forces should use either a front width higher than the third quantile or lower than the first quantile. Allied forces should use the mean front width, 9 Km..

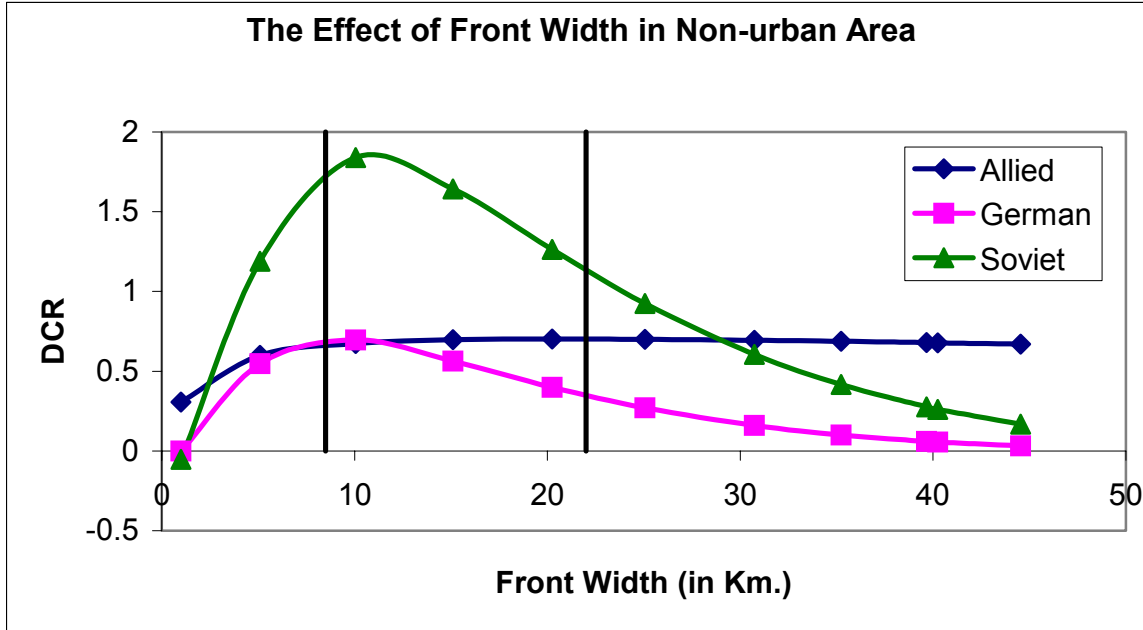


Figure 25. The Effect of the Front Width in Non-Urban Areas

Comparing Figures 24 and 25 reveal the following conclusions. For values of front width higher than its first quantile, the effect of front width is different for Allied forces. For the values of front width greater than the first quantile, 78% of the occurrences, increasing the front width decreases the DCR. It is possible to state that front width has a significant effect on the DCR. The effect of front width is different for Allied forces than for the others.

7. The Effect of the Linear Density on the DCR

The effect of the number of men in front per km (linear density) depends on the terrain, the attacking country, the force ratio, the front width and the attacker’s strength. Although there is no direct relation to the defender’s strength, because of its interactions with the force ratio, it is imperative to add the defender’s strength as well. The following equation shows the relationship:

$$\begin{aligned}
 XC3 = & -8.727XC_{CountryAllied} + 3.8701XC_{CountrySov} - 10.9237Urban + 0.0005XStrength \\
 & - 0.4902ForceRatio - 1.2336XFrontWidth^2 - 0.0974XMenFront^2 \\
 & + 0.9507XC_{CountryAllied}XMenFront^2 - 0.453XC_{CountrySov}XMenFront^2 \\
 & + 1.2877Urban : XMenFront^2 + 0.0444ForceRatio : XMenFront^2 \\
 & - 1.309XMenFront^2 : XFrontWidth^2
 \end{aligned}$$

Due to the way the variables derived, it is impossible to fix the attacker's strength and front width to their mean values. An increase in linear density either results from a change in the attacker's strength or a change in the attacker's front width or both. Evaluating the third option is extremely difficult. Thus, there are two ways to see the effect of linear density on the attacker's DCR. One is to change the linear density by changing the attacker's strength, and the second is to change it by changing the attacker's front width.

However, changing the attacker's front width does not produce the range of linear density that the data set possesses. Thus, the linear density's effect will be analyzed by changing the attacker's strength. The defender's strength and the attacker's width are set to their mean values. The other variables are calculated according to these values. The attacker's strength is changed within its range. Figure 26 shows the relationship between the linear density and the attacker's daily casualty rate.

It is obvious that the DCR is a monotonic increasing function of the linear density up to the values of 3000. The attacker's DCR increases as the linear density increases up to 3000 men per km. for German and Allied forces and up to 4000 per km. for Soviet forces. Values higher than these numbers decrease the DCR. The higher rates of the Soviet DCR at lower linear densities can be ignored because there are no data points for these values in the data set for the Soviets. In other words, trying to estimate the DCR for the Soviet forces at densities lower than 1000 per Km is beyond the scope of this model.

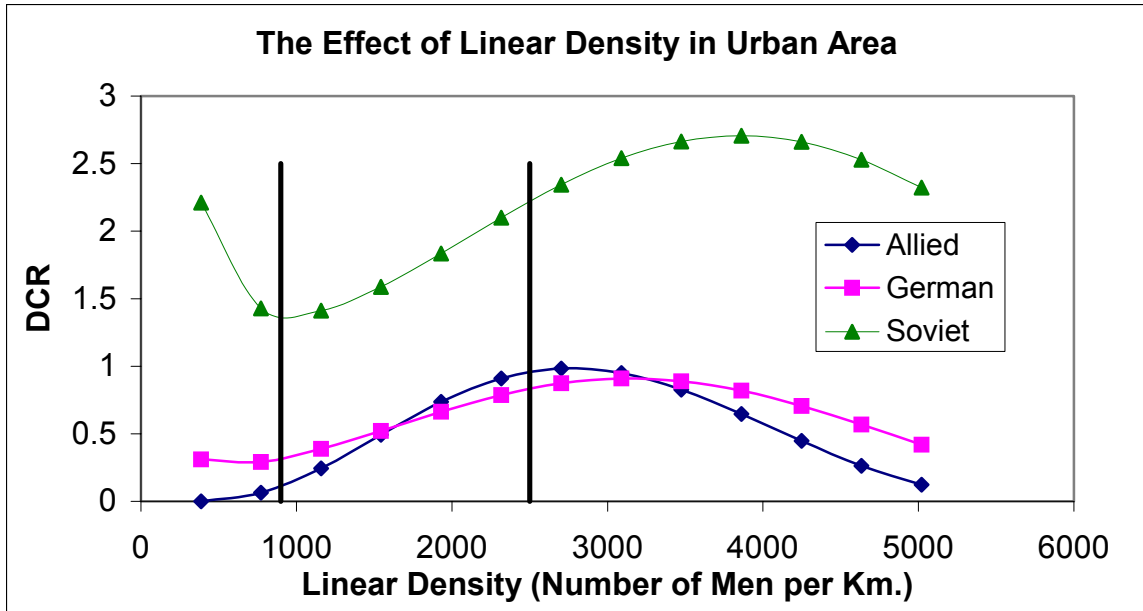


Figure 26. The Effect of the Linear Density in Urban Areas

The effect of the linear density in non-urban areas is the same for each country. As the linear density increases in non-urban operations, the DCR decreases. This might be the effect of the “concentration of forces.”

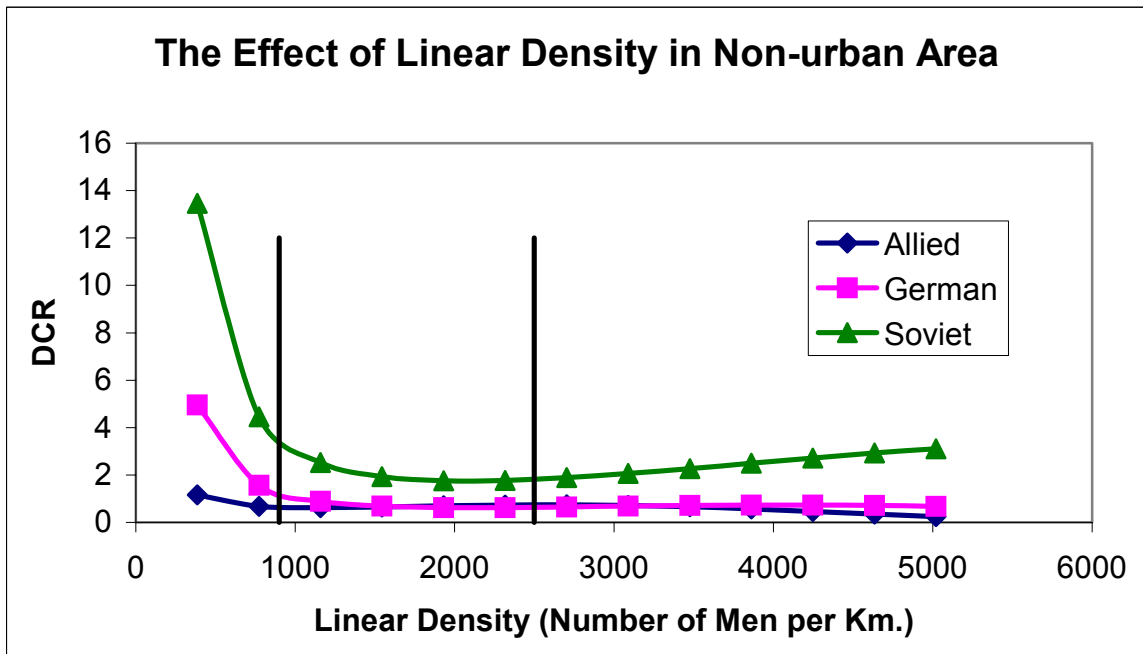


Figure 27. The Effect of the Linear Density in Non-Urban Areas

The effect of the linear density in each type of terrain is the same for each country. In general, while in urban areas, increasing linear density increases the DCR, it decreases the DCR in non-urban areas.

D. CONCLUSIONS

The best linear model contains all the available variables. Although some, such as the defender's strength are not significant alone, their interaction with other variables forces these variables to stay in the model.

An analysis of the plots showed that linear regression could be used to predict the attacker's daily casualty rate. However, the derivation of variables, transformation and interactions make this model difficult to interpret. By using other tools, it is still possible to show the effects of each variable. The terrain, the attacker's country and strength and the linear density affect the attacker's daily casualty rate. On the other hand, the defender's strength is not significant in most cases.

The effect of terrain is statistically significant. The DCR is higher in non-urban operations. This is a counter-intuitive result. However, it is known that urban operations take more time than non-urban operations [Ref. 15]. Thus, it is possible to state that the number of casualties may be higher in urban operations even if the *daily* casualty rate is smaller.

The attacker's country is a significant factor. Its effect did change according to the terrain. The casualty rate of the Allies was almost the same as the Germans, but possibly slightly smaller. The Soviets suffered the highest casualty rates.

The effect of the attacker's strength significantly changes according to terrain. While an increase in the attacker's strength increases the DCR in urban areas, it decreases the DCR in non-urban areas. However, with the same force size, it is more likely to have higher casualty rates in non-urban areas.

The defender's strength has the same effect in urban and non-urban areas. Thus, it does not change according to the terrain. This agrees with the absence of the term *Urban:YStrength* in the model. As the defender's strength increases, the attacker's casualty rate increases. Under the same circumstances, the Germans defense causes a higher casualty rate for the Soviets than the allies.

The effect of force ratio depends on the type of terrain. The force ratio has a significant effect on urban operations. However, its impact in non-urban operations is negligible. As expected, higher force ratios decrease the attacker's daily casualty rate, more in urban operations.

The relationship between the DCR and front width changes according to the terrain. Although the effect is similar for the Soviet and Germans in each type of terrain, the relationship is different for the Allies. In urban areas, the higher front width is better for the Soviet and German forces while it has the opposite impact on the Allies. In non-urban areas, the effect is similar between the Germans and Soviets. Front width higher than 9 Km. decreases the DCR of the German and Soviet forces.

The effect of linear density has a similar effect on the attacker's strength. In urban operations, the higher linear density translates into higher casualty rates. In non-urban operations, the effects are similar for each country. Higher linear density means lower DCR.

IV. CONCLUSIONS

Chapter I demonstrated that urban operations have always been of great concern for war planners. It has been recommended that Military Operations in Urban Terrain (MOUT) be avoided because MOUT causes an excess use of resources. The number of casualties is one of them. Since urban structures provide advantages to the defender, even a highly inferior number of defenders cause a large number of casualties to the attacker.

Despite the recommendation, it is no longer possible to avoid MOUT. As urbanization grows at a very high speed, it is impossible to imagine a campaign without MOUT. Especially if the attacker is overwhelmingly powerful, it is almost certain to conduct operations in urban areas.

Until recently, the main concern in military research was to predict the outcome of the battle. As public opinion started giving more importance to the number of casualties, some studies conducted research on the number of casualties. This thesis analyses the effect of terrain on the daily casualty rates in MOUT.

Descriptive statistics showed that, surprisingly, daily casualty rates tended to be higher in non-urban operations. However, this does not mean that MOUT produces more casualties. Remembering that MOUT takes more time, although the daily casualty rate is smaller in urban operations, the total number of casualties might still be higher.

Among all the other factors, force ratio and terrain have the most effect on the DCR. Commanders of attacking forces should gather as much force as possible if the casualty rate is of great concern. Allied forces successfully used this tactic in WW II. The answers to the research questions are as follows.

- In division-level engagements, the daily casualty rate is higher in non-urban operations. However, this does not mean that the total number of casualties is also higher.
- All the variables but the defender's strength are important factors in estimating the daily casualty rate. However, given the relationship between variables, terrain and attacker's strength are the most important factors. The low R^2 from the model implies that noise level is high and perhaps other variables should be included in the model.

- It is clear that the Soviets have the highest DCR. Allied forces have a lower daily casualty rate, which may be slightly lower than the Germans.
- Since the data results from daily engagements, it was not possible to ascertain if the number of casualties is higher in urban operations.
- Attackers have lower daily casualty rates than defenders.
- The effects of each variable but defender's strength change according to the terrain.

During the analysis, two interesting points arose. The manner in which the allies understand urban warfare is different from the Soviets and Germans. They had a higher front width, as well as more linear density, force ratio and strength than the others. The allies gave more weight to MOU. The other point is the Soviets' understanding of war. They generally suffered high casualty rates. It may be that they employed a strategy of attrition: since they possessed a large number of human resources, they were perhaps less concerned with the casualty rate.

Further research is possible on this subject. The defender's daily casualty rate can be analyzed as well. One recommendation for further study is to use logarithm to transform the response variable, if necessary. Also, adding more variables such as the number of tanks, aircrafts, missiles and ships will very likely provide better results. If available, The Dupuy Institute's DLEDB might help significantly.

One other topic should be the analysis of different levels of engagements, at the battalion level, for example. It is expected that the daily casualty rates are higher at lower levels, and could possibly be verified and compared with the conclusions of this thesis.

APPENDIX A. ACRONYMS

DCR:	Attacker's daily casualty rate
DDCR:	Difference in attacker's daily casualty rate
DCRT:	Transformed form of attacker's daily casualty rate
DLEDB:	Division Level Engagement Database
ETO:	European Theater of Operation
ForceRatio:	Force Ratio
TDI:	The Dupuy Institute
MOUT:	Military Operations in Urban Terrain
Urban:	Type of terrain "Urban" or "Non-Urban"
XCasualty:	Attacker's Daily Casualty Rate
XC3:	$XCasualty^{0.3}$
XFrontWidth:	Attacker's Front Width in kilometers
XMenFront:	Attacker's Number of Men in the Front Line, Linear Density
XStrength:	Attacker's Total Personnel Strength
YCasualty:	Defender's Daily Casualty Rate
YC2:	$YCasualty^{0.2}$
YStrength:	Defender's Total Personnel Strength

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX B. SUMMARY STATISTICS

A. INDEPENDENT VARIABLES

1. Numeric Variables

	XStrength	YStrength	ForceRatio	XFrontWidth	XMenFront
Min:	5600	358	0.4	1	203.65
1st Qu.:	14686	6960	1.32	8.575	1101.468
Mean:	24257.36	12755.01	2.88	15.35	2190.46
Median:	19032	10300	1.87	12	1789
3rd Qu.:	29895	15943	3.02	21.925	2497.74
Max:	67829	57089	45.53	73	19322.08
Total N:	253	253	253	253	253
NA's :	0	0	0	9	9
Std Dev.:	14348.55	8830.99	4.27	10.38	2141.08
LCL Mean:	22480.78	11661.59	2.35	14.04	1920.47
UCL Mean:	26033.95	13848.43	3.41	16.66	2460.46

Table 6. Summary Statistics for Numeric Independent Variables

2. Categorical Variables

XCountry	YCountry	Urban
Ger: 102	Allied: 27	Urban: 97
Allied:110	Ger: 151	Non-Urban:156
Sov: 41	Sov: 75	

Table 7. Summary Statistics for Categorical Independent Variables

B. DEPENDENT VARIABLES

	XCasualty	YCasualty
Min:	0	0
1st Qu.:	0.35	0.76
Mean:	1.07	8.58
Median:	0.65	2.4
3rd Qu.:	1.29	6.3
Max:	8.59	146.88
Total N:	253	253
NA's :	0	0
Std Dev.:	1.29	20.24
LCL Mean:	0.91	6.07
UCL Mean:	1.23	11.09

Table 8. Summary Statistics for Dependent Variables

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX C. TRANSFORMATION OF VARIABLES

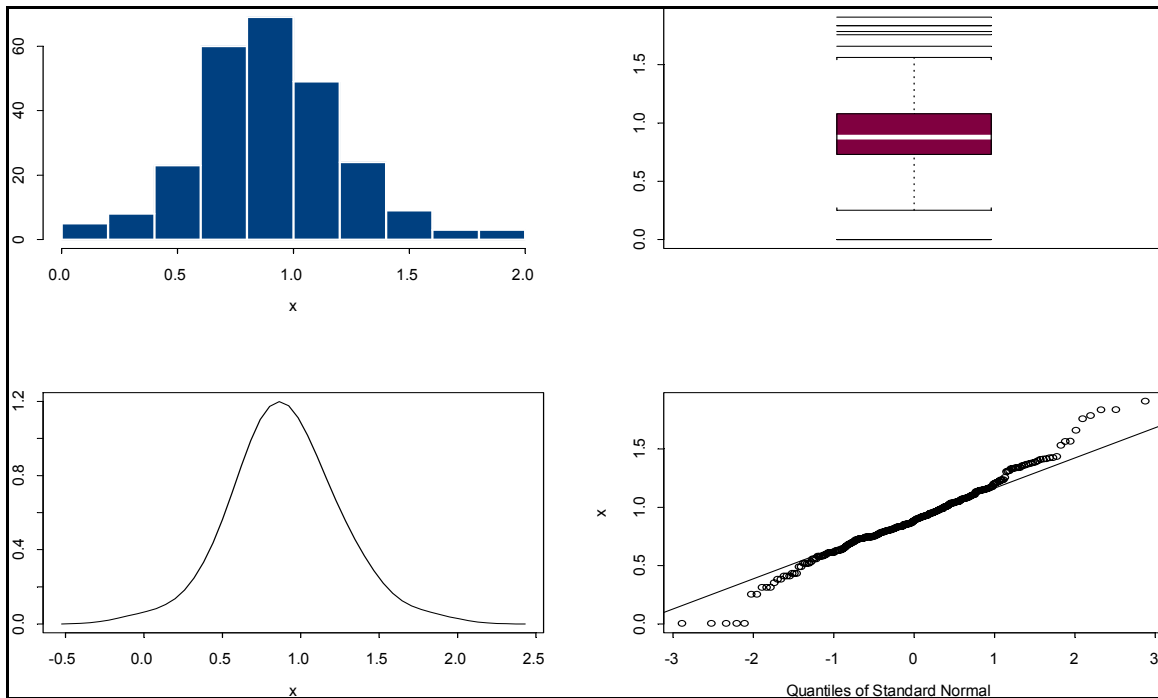


Figure 28. Transformed Attacker's Daily Casualty Rate Plots

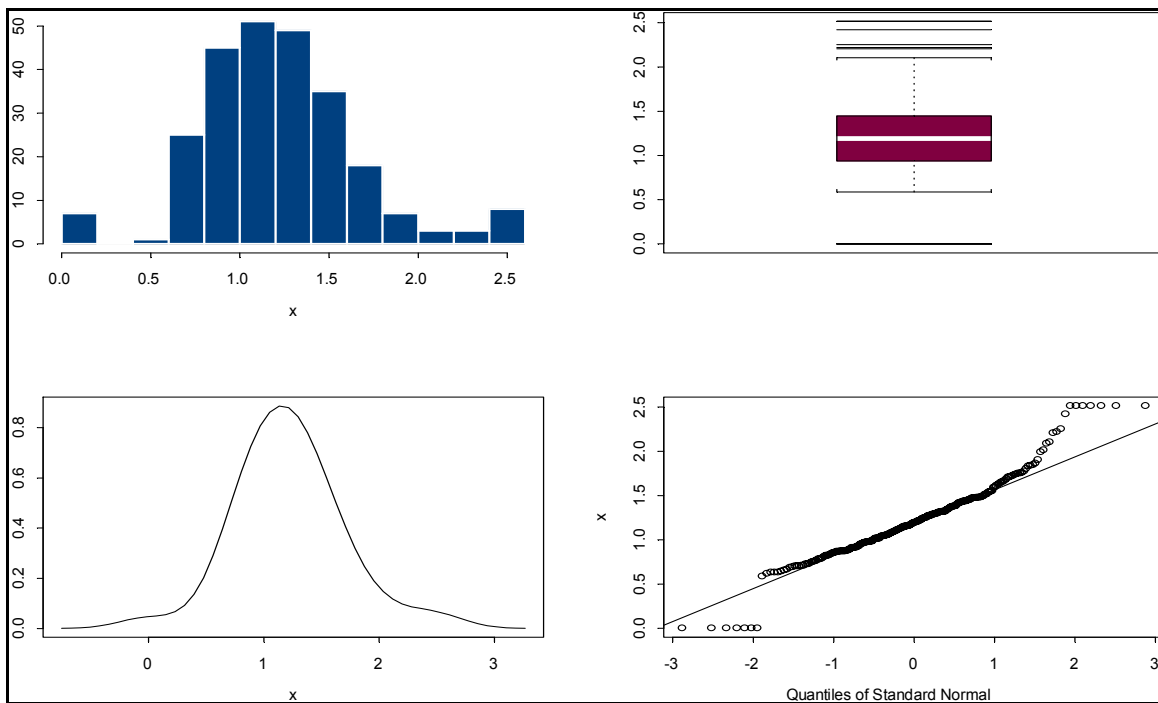


Figure 29. Transformed Defender's Daily Casualty Rate Plots

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX D. ATTACKER'S DAILY CASUALTY RATE MODELS

A. MODEL WITHOUT INTERACTIONS

	Value	Std. Error	t value	Pr(> t)
(Intercept)	1.4638	0.1492	9.8085	0.0000
XCountryAllied	0.0095	0.0406	0.2343	0.8149
XCountrySov	0.3884	0.0548	7.0858	0.0000
Urban	-0.0855	0.0389	-2.1970	0.0290
XStrength	0.0000	0.0000	2.8183	0.0052
ForceRatio	-0.0189	0.0046	-4.0864	0.0001
XFrontWidth2	-0.3795	0.0923	-4.1099	0.0001

Table 9. Summary Statistics for the Model without Interactions

Residual standard error: 0.2835 on 237 degrees of freedom

Multiple R-Squared: 0.2625

F-statistic: 14.06 on 6 and 237 degrees of freedom, the p-value is 1.109e-013

9 observations deleted due to missing values

B. MODELS WITH TWO-WAY INTERACTIONS

	Value	Std. Error	t value	Pr(> t)
(Intercept)	10.3621	1.7876	5.7967	0.0000
XCountryAllied	-8.7270	2.5764	-3.3872	0.0008
XCountrySov	3.8701	3.3341	1.1608	0.2470
Urban	-10.9237	2.5378	-4.3044	0.0000
XStrength	0.0005	0.0002	3.3077	0.0011
YStrength	0.0000	0.0000	-1.6828	0.0938
ForceRatio	-0.4902	0.1531	-3.2018	0.0016
XFrontWidth2	-1.2336	0.8208	-1.5028	0.1343
XMENFront2	-0.0974	0.2930	-0.3324	0.7399
XCountryAlliedXStrength	0.0000	0.0000	-2.6770	0.0080
XCountrySovXStrength	0.0000	0.0000	1.1782	0.2400
XCountryAlliedXFrontWidth2	3.0993	0.8966	3.4568	0.0007
XCountrySovXFrontWidth2	-1.1350	1.0728	-1.0579	0.2913
XCountryAlliedXMENFront2	0.9507	0.3035	3.1319	0.0020
XCountrySovXMENFront2	-0.4530	0.4283	-1.0577	0.2914
Urban:XStrength	0.0000	0.0000	-3.0437	0.0026

	Value	Std. Error	t value	Pr(> t)
Urban:ForceRatio	-0.0234	0.0167	-1.4049	0.1615
Urban:XFrontWidth2	3.5611	0.8778	4.0571	0.0001
Urban:XMenFront2	1.2877	0.3019	4.2648	0.0000
XStrength:XFrontWidth2	-0.0001	0.0000	-2.9513	0.0035
XStrength:XMenFront2	0.0000	0.0000	-3.3179	0.0011
YStrength:XFrontWidth2	0.0000	0.0000	1.7358	0.0840
ForceRatio:XFrontWidth2	0.1592	0.0506	3.1449	0.0019
ForceRatio:XMenFront2	0.0444	0.0158	2.8017	0.0055
XFrontWidth2:XMenFront2	-1.3090	0.4276	-3.0611	0.0025

Table 10. Summary Statistics for the Model with Interactions

Residual standard error: 0.2669 on 219 degrees of freedom

Multiple R-Squared: 0.3962

F-statistic: 5.988 on 24 and 219 degrees of freedom, the p-value is 5.496e-014
9 observations deleted due to missing values

	Urban			Non-Urban		
	Allied	Germans	Soviets	Allied	Germans	Soviets
(Intercept)	-9.2886	-0.5616	3.3085	1.6351	10.3621	14.2322
XStrength	0.0005	0.0005	0.0005	0.0005	0.0005	0.0005
YStrength	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
ForceRatio	-0.5136	-0.5136	-0.5136	-0.4902	-0.4902	-0.4902
XFrontWidth2	5.4268	2.3275	1.1925	1.8657	-1.2336	-2.3686
XMenFront2	2.1410	1.1903	2.1410	0.8533	-0.0974	-0.5504
XStrength:XFrontWidth2	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001	-0.0001
XStrength:XMenFront2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
YStrength:XFrontWidth2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
ForceRatio:XFrontWidth2	0.1592	0.1592	0.1592	0.1592	0.1592	0.1592
ForceRatio:XMenFront2	0.0444	0.0444	0.0444	0.0444	0.0444	0.0444
XFrontWidth2:XMenFront2	-1.3090	-1.3090	-1.3090	-1.3090	-1.3090	-1.3090

Table 11. Coefficients Given the Attacker's Country and Terrain

LIST OF REFERENCES

1. Population Division of the UN Department of Economic and Social Affairs, "World Urbanization Prospects: The 2001 Revision". Retrieved March 3, 2004, from UN Access:
<http://www.un.org/esa/population/publications/wup2001/wup2001dh.pdf>.
2. The Dupuy Institute, "Capture Rate Study, Phases I and II," prepared for U.S. Department of Army Center for Army Analysis, 6 March 2000.
3. Schmitt, Eric, "US Seeks to Cut Accidental War Death", New York Times, 9 December 1991.
4. The Dupuy Institute, "Measuring the Effects of Combat in Cities, Phase 1," prepared for U.S. Department of Army Center for Army Analysis, 11 January 2002.
5. Bracken, J., "Attrition Models of the Ardennes Campaign," vol. 42, pp. 559-577, Naval Research Logistics, 1995.
6. Personal Communication from Professor Moshe Kress, Operational Research Department, Naval Postgraduate School, Monterey, California, 19 April 2004.
7. Yigit, Faruk, "Finding the Important Factors in Battle Outcomes: A Statistical Exploration of Data from Major Battles," Master's Thesis, Naval Postgraduate School, Monterey, California, 2000
8. Military Operations on Urbanized Terrain (MOUT), Marine Corps Warfighting Publication (MCWP) 3-35.3, 16 April 1998.
9. The Dupuy Institute, "Measuring the Effects of Combat in Cities, Phase II (Part I)," prepared for U.S. Department of Army Center for Army Analysis, 30 June 2003.
10. Glenn, Russell W., "Heavy Matter: Urban Operations' Density of Challenges," RAND Corporation, October 2000.
11. Devore, Jay L., "Probability and Statistics for Engineering and the Sciences," Fifth edition, Duxbury Thomson Learning, 2000.
12. Ziemke, Earl F., "The U.S. Army in the Occupation of Germany," Chapter 10, p.145, Center of Military History, United States Army, Washington, D.C., 1990.

13. Rummel, Rudolph J., "Statistics of Democide: Genocide and Mass Murder Since 1900", Chapter 13, Center for National Security Law, School of Law, University of Virginia and Transaction Publishers, Rutgers University, 1997. Hamilton, Lawrence C., "Regression with Graphics, A Second Course in Applied Sciences," Duxbury Press, Belmont, California, 1992.
14. Hamilton, Lawrence C, "Regression with Graphics: A Second Course in Applied Statistics," Duxbury Press, 1991.
15. Military Field Manual FM 3-06 "Urban Operations," Chapter 4, Department of the Army Headquarters, June 2003.

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
Ft. Belvoir, Virginia
2. Dudley Knox Library
Naval Postgraduate School
Monterey, California
3. Genelkurmay Baskanligi
Personel Daire Baskanligi
Bakanliklar
Ankara, TURKEY
4. Kara Kuvvetleri Komutanligi
Personel Daire Baskanligi
Bakanliklar
Ankara, TURKEY
5. Kara Kuvvetleri Komutanligi
Kara Kuvvetleri Kutuphanesi
Bakanliklar
Ankara, TURKEY
6. Kara Harp Okulu Komutanligi
Kara Harp Okulu Kutuphasi
Bakanliklar
Ankara, TURKEY
7. U.S. Army Center for Army Analysis
Wilbur B. Payne Hall
Fort Belvoir, Virginia
8. Professor Samuel E. Buttrey
Department of Operations Research
Naval Postgraduate School
Monterey, California
9. LTC Saverio Manago
Department of Operations Research
Naval Postgraduate School
Monterey, California