

Analysis of Survivability for Combatants during Offensive Operations at the Tactical Level

Jaeoh Kim^a · HyungJun Cho^{a,1} · GakGyu Kim^b

^aDepartment of Statistics, Korea University;

^bCenter for Army Analysis and Simulation, ROK Army

(Received June 29, 2015; Revised July 9, 2015; Accepted August 10, 2015)

Abstract

This study analyzed military personnel survivability in regards to offensive operations according to the scientific military training data of a reinforced infantry battalion. Scientific battle training was conducted at the Korea Combat Training Center (KCTC) training facility and utilized scientific military training equipment that included MILES and the main exercise control system. The training audience freely engaged an OPFOR who is an expert at tactics and weapon systems. It provides a statistical analysis of data in regards to state-of-the-art military training because the scientific battle training system saves and utilizes all training zone data for analysis and after action review as well as offers training control during the training period. The methodologies used the Cox PH modeling (which does not require parametric distribution assumptions) and decision tree modeling for survival data such as CART, GUIDE, and CTREE for richer and easier interpretation. The variables that violate the PH assumption were stratified and analyzed. Since the Cox PH model result was not easy to interpret the period of service, additional interpretation was attempted through univariate local regression. CART, GUIDE, and CTREE formed different tree models which allow for various interpretations.

Keywords: scientific battle training, Cox proportional hazard model, CART, GUIDE, CTREE

1. 서론

‘평화를 원한다면 전쟁을 준비하라’는 말과 같이 전쟁은 문명의 생성과 소멸로부터 국가의 흥망성쇠에 관한 문제를 결정해 왔으며 특히, 남북한이 군사적 대치 상황에 직면한 대한민국의 경우 군사학에 관한 깊은 연구가 필요하다. 군사학(military science)은 본질적으로 전쟁과 군사현상에 관한 연구이며 군법, 군사행정, 참모학, 통솔법 등의 일반 군사학, 제대별 공격 및 방어작전 등 전술에 대해 연구하는 전술학, 무기체계에 대한 각종 연구를 포함하는 화포학 및 무기체계학 등으로 구분할 수 있다. 이와 같이 군사학은 매우 광범위하여 일종의 종합과학적 성격이 가지며 다양한 학문에서 연구가 진행되고 있다. 그러나 창끝 전투력이 강한 부대가 전투에서 결국 승리한다는 최근 경향만큼 대표적 창끝부대인 보병대대급 전투현상을 연구하는 노력은 다른 관련 연구에 비해 부족한 실정이다. 이는 대대급 제대의 주요 구성요소가 개인 및 개인화기를 비롯한 소구경 화기가 대부분이고 전투현장 역시 대항군을 포함하여 약 1000여

Supported by a Korea University Grant(K1507651).

¹Corresponding author: Department of Statistics, Korea University, 145 Anam-ro, Sungbuk-Gu, Seoul 136-701, Korea. E-mail: hj4cho@korea.ac.kr

명의 사람과 수많은 무기체계가므로 이를 모형화하기 매우 어렵기 때문이다. 반면 우리 군은 첨단화된 육군 과학화 전투훈련장(Korea combat training center; KCTC)를 운영하여 보병대대의 실전과 유사한 전투훈련을 지원하고 있으며, 훈련간 발생한 모든 상황은 데이터 형태로 중앙 서버에 저장되어 훈련 종료후 대대장을 비롯한 지휘관 및 참모의 지휘판단 및 부대지휘절차의 적절성, 소/분대장 등 소부대 지휘자의 지휘통솔 능력 등을 사후검토(after action review; AAR)하고 있다.

과학화 전투훈련에 관한 주요 연구는 다음과 같다. Kim과 Kim (2014)에서는 다중회귀분석을 통하여 과학화 전투훈련 결과로부터 공격작전의 승리요인을 찾고자 하였으며, Kim 등 (2013), Hodson과 Baldwin (2009) 등은 과학화 전투훈련을 구현하기 위한 이론인 국방 M&S(modeling & simulation) 방법 및 발전 방향을 제시하였다. Herl 등 (2005)은 부대의 기동 방해요소 패턴을 예측하기 위해 미군의 과학화 전투훈련장인 CMTC-LF(combat maneuver training center-live fire) 데이터를 이용하였다. 그러나 앞선 연구는 대부분 과학화 전투훈련에 대한 공학 측면에서 발전방향 제시가 주를 이루고 있으며 Kim과 Kim (2014) 연구도 대부대 공격교리에 관한 것으로 전술제대 군사교리(military doctrine) 발전 및 검증에 과학화 전투훈련 자료로부터 분석하는 연구는 많지 않다. 과학화 전투훈련 자료가 전술제대급 군사교리를 연구하는데 매우 효과적이고 방대한 자료가 축적되어 있음에도 불구하고 많은 연구가 없었던 이유는 군사자료이므로 일반인의 접근이 어려움과 더불어 로그 파일형태의 원 데이터를 가공하기 쉽지 않기 때문으로 판단된다. 본 연구에서는 과학화 전투훈련 데이터 전문가로부터 원 데이터를 통계 분석에 용이하게 전환하여 분석을 하였다. 한편 생존자료분석(survival data analysis) 병원의 임상 자료분석이 많은 부분을 차지하며, Barney (2002)의 생존분석방법에 특히출현 이후 소멸을 적용한 연구, Schmidt과 Witte (1989)의 재소자 출소 후 재범에 관한 연구 등 다양한 분야에 적용이 있었다.

여기서 증강된 보병대대의 공격작전 과학화 전투훈련 결과를 Cox PH(proportional hazard) 모형과 다양한 생존분석을 위한 나무모형에 적용하여 전술제대 공격작전간 전투원의 생존성 관련 사항을 분석 및 해석한다. 과학화 전투훈련 자료는 중도절단(censoring)이 발생하였으며 생존시간(survival time)을 반응변수로 설정할 수 있는 점에서 Cox PH 모형이 적절하며, 생존나무 모형을 적용함으로써 다양한 전술적 임무를 가진 여러 부대들로 구성된 증강된 보병대대의 생존성 분석을 위한 해석이 더욱 풍부해 진다. 본 연구의 주요 특징은 우선 분석하고자 하는 과학화 전투훈련 자료는 통계적 측면에서 학술연구에 적용된 최초 사례라는 점이다. 둘째, Grambsch and Therneau 검정을 통해 Cox PH 모형의 비례성 가정을 확인하고 이를 위배하는 변수를 적절히 처리하여 왜곡된 결과를 방지하였다. 셋째, 일반적으로 알려진 CART(classification and regression trees) 기반의 생존나무(survival tree) 모형 뿐 아니라 생존나무 모형에서 변수 선택의 편향(bias)문제가 극복된 GUIDE(generalized, unbiased, interaction detection and estimation)와 CTREE(conditional inference trees) 기반의 모형을 추가 적용하였다.

본 논문의 구성은 다음과 같다. 제 2절에서 과학화 전투훈련의 개념을 소개하며 관련된 분석변수와 분석방법인 생존분석 및 생존분석을 위한 나무모형에 대해 소개한다. 제 3절에서는 Cox PH 모형 결과 및 가정 위배사항에 대해 상세히 설명하며 CART, GUIDE, CTREE모형을 이용한 나무모형을 이용하여 분석 및 해석한다. 본 연구에서는 CART와 CTREE는 R의 'rpart'와 'party' 패키지를, GUIDE는 위스콘신 대학에서 제공하는 guide 프로그램(<http://pages.stat.wisc.edu/~loh/guide.html>)을 사용하였다. 마지막 제 4절에서 결론을 제시한다.

2. 연구 방법

2.1. 과학화 전투훈련의 개념

군의 과학화 전투훈련은 컴퓨터 공학인 M&S기술을 바탕으로 비약적 발전을 거듭하고 있다. 과학화 전

투훈련은 각개병사의 개인훈련부터 제대별 부대훈련 및 기동훈련 등에 이르기까지 다양한 목적으로 활용되며 통합된 과학화 훈련체계를 LVC(live, virtual, constructive simulation)체계라 부른다. 이 중 Live 체계란 증강된 보병대대인 훈련부대가 MILES(multiple integrated laser engagement system)와 증양통제장비체계 등 과학화된 훈련체계를 운용하여 전문 대항군과 쌍방 자유기동 형태로 실시하는 훈련이다. 훈련에 참여한 모든 인원 및 장비의 교전활동과 피해결과는 실시간 디지털 정보로 훈련통계본부로 전송 및 처리된다. 다시 말하면 강원도 홍천에 위치한 KCTC 전투 훈련장에서 MILES를 착용하고 기동하는 병력 및 장비의 전투행동에 관한 모든 데이터가 실시간 훈련통계본부로 수집되어 처리되고, 실제훈련이 제한되는 상급 및 인접부대 기동, 화학전, 항공전력, 곡사화기 등의 사격 및 피해정보는 실전 데이터를 근거로 사이버 공간에서 가상 모의되며 교전에 의한 피해결과가 훈련 병력에 직접 전송되어 훈련을 통제하는 첨단 훈련체계이다.

MILES 시스템은 광학 디지털 기술을 이용해 화기의 사격 및 피격효과를 실제와 유사하게 묘사해 주는 장비로서 각종 화기에 레이저 발사기를 부착하고 표적에는 레이저 광선 감지기를 부착하여 사격시 레이저 발사를 통해 표적의 명중여부를 판별하여 훈련하는 실전적인 다중통합레이저 훈련체계이다. 대대급 부대의 주요 화기인 K-1, K-2 소총을 비롯하여 TOW 대전차화기, 크레모아, 전차 등을 비롯하여 5,600여종의 장비들이 MILES 시스템에 포함된다. 또한 훈련에 참가한 장병들은 개인용 감지기를 팔에 부착하며 전투현장에 설치된 중계탑을 통해 무선데이터 통신으로 훈련통제본부에 전투간 발생한 각종 상황을 전송하며 피해결과 등을 전송받는다. 또한 전문 대항군(opposing force) 대대는 적의 편제 및 화기와 유사하게 조직되어 있으며 인민군과 유사한 복장을 착용하고 적 전술로 전투를 실시하여 훈련 효과를 극대화한다.

2.2. 분석 자료

본 연구에서는 전술제대 공격작전간 장병의 생존성 보장과 관련된 분석을 위해 실제 사람과 장비의 움직임과 전투행동 결과를 모두 기록한 KCTC 자료를 분석한다. KCTC에서는 우리 육군의 보병 대대를 대상으로 과학화 훈련을 실시하고 있으며, 1개 대대가 공격 및 방어작전을 포함한 전술훈련을 약 2주간 진행한다. 본 연구에서는 당해년 최우수 부대로 선정된 2010년 8번째로 훈련한 증강된 보병 대대의 훈련 자료 중 공격작전 자료를 대상으로 분석한다. 이 부대의 훈련에 참여한 전투원은 783명이다.

원 자료는 소속 제대정보, 개인별 정보(신장, 몸무게, 시력, 복무기간 등), 화력지원 요청 및 결과 등 157개 주기 훈련 schema, 무기 및 탄약 현황, 지뢰 및 화생방 현황 등 115개 schema, 각종 장비의 정비 현황을 포함하는 34개 schema로 구성되어 있다. 원 자료는 우리 육군 및 U.S. Army (2008)에서 강조하는 공격작전 중요 변수를 추출하기 위하여 각 전투원의 시점별 위치 정보, 경사 및 중상 등 전투피해 평가 원인 및 결과, 상급부대 화력지원 정도 등 다양한 정보로 가공이 필요하다. 원 자료는 종류도 다양하지만 그 형태로 일반적이지 않아 과학화 훈련 시스템 전문가에 의해 가공이 이루어졌다. 이렇게 가공한 정보로부터 다음 절에서 설명하는 설명 및 반응변수를 설정하였다.

2.3. 변수 설명

본 연구의 반응변수는 전투불능 상태인 전사 및 중상까지 생존시간이다. 경사 판정을 받는 경우 정상적으로 전투활동을 지속할 수 있으므로 이러한 경우는 포함하지 않는다. 즉, 생존분석의 반응변수인 사건의 발생시간(time-to-event)이 훈련에 참가한 각개 장병의 전사 및 중상까지 시간이고, 공격작전에 관한 총 훈련시간이 약 70시간인 점을 고려하여 단위는 분이다. 중도절단은 장병 개인에게 부착된 센서의 고장으로 훈련 참여가 중단되거나 훈련간 위규 행동 등으로 훈련 통제관에 의해 강제 훈련탈퇴를 조치

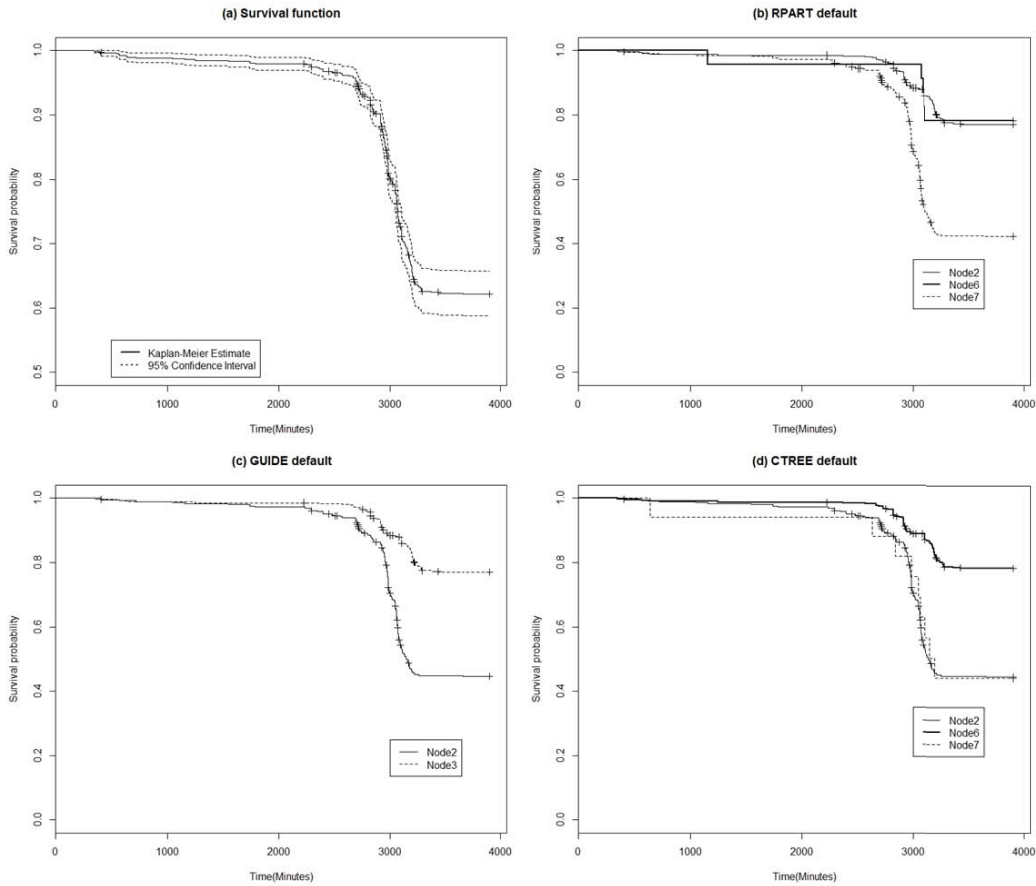


Figure 2.1. (a) Kaplan-Meier estimate and pointwise 95% confidence intervals of the survival function, (b) Survival function by RPART for each terminal nodes, (c) Survival function by GUIDE for each terminal nodes, (d) Survival function by CTREE for each terminal nodes.

를 받는 경우를 포함하며 아울러 공격작전 종료시까지 생존하는 경우이다. 중도절단된 관측치는 497개로 약 63.47%이다. Figure 2.1 (a)는 Kaplan-Meier 방법으로 실선의 생존함수와 점선의 95% 신뢰구간을 추정하였다. Figure 2.1 (a)로부터 공격작전간 가장 치열했던 전투가 대략 2200분부터 3300분 사이에 있었음을 알 수 있다.

본 연구에서는 설정한 설명변수는 복무기간(POS), 시력(Sight), BMI, 상급부대 사격지원량(SA), 분대원간 평균거리(Save), 중대원간 최대 거리(Cmax)인 6개의 연속형 변수와 부대임무(Mission), 개인화기종류(IW)인 2개의 범주형 변수이다. 복무기간은 훈련부대에 전입간 시점을 기준으로 산정하였으며 상급부대 사격지원량은 공격작전간 상급부대 포병사격지원 및 항공기에 의해 소속 중대가 지원받은 사격횟수를 의미한다. 대대급 작전의 최소 전술제대인 분대에 대해 작전간 분대원과 시점별 평균 거리와 전술적 임무를 수행하는 중대에 대해 무전기 교신 및 작전적 고립 등을 고려하여 타 중대원과 공격작전간 매 시점별 최대 거리의 평균 거리를 설명변수로 선정하였다. KCTC 자료는 모든 전투원이 사격, 기동 등 이벤트가 발생할 때마다 또는 몇 초 단위로 위치를 중앙통제본부로 전송하므로 개인별 수집에서 수천

개의 시점별 위치, 발생한 이벤트에 대한 정보를 갖기 때문에 이를 전체적으로 고려한 Save와 Cmax는 전술적으로 의미있는 변수로 볼 수 있다. Mission은 전술적 임무에 따라 지휘부, 주공 및 적지중심작전 부대, 조공 및 지원중대, 포병부대와 의무, 공병, 정비 등 기타 부대의 5개 범주로 구분하고, IW는 권총, K-1 및 K-2 소총, K-3 기관총의 3개 범주로 구분한다.

2.4. Cox PH 모형 및 나무모형 소개

본 절에서는 본 연구의 이론적 배경이 되는 생존분석 기법 중 Cox PH 모형과 중도절단을 포함하는 자료에 대해 나무모형을 구성하는 대표적인 방법인 CART, GUIDE, CTREE에 대해 간략히 설명한다.

2.4.1. Cox PH 모형 중도절단 자료의 포함을 큰 특징으로 하는 생존분석자료를 분석하는 방법은 생명표(life-table) 방법과 Kaplan-Meier 추정량을 이용하는 비모수적(non-parametric) 방법, 생존 시간이 지수분포, 와이블 분포 등 모수분포를 갖는다고 가정하고 분석하는 모수적(parametric) 방법이 있다. 한편 Cox (1972)의 생존시간과 설명변수의 관계를 위험함수(hazard function)를 통해 표현하는 Cox PH 모형이 있다. 이 모형은 분포에 대한 가정을 하지 않지만, β 라는 모수를 추정하는 점에서 준모수적(semi-parametric)방법으로 분류된다. 식 (2.1)에서 $\lambda_0(t)$ 는 분포 가정없는 기저위험함수(baseline hazard function)이며, $\beta = (\beta_1, \dots, \beta_p)$ 는 회귀계수로 공변량 Z_{ik} 가 한 단위 증가할 때마다 $\exp(\beta_k)$ 만큼 위험률이 증가함을 의미한다.

$$\lambda(t; Z_i) = \lambda_0(t) \exp(Z_i \beta). \quad (2.1)$$

두 개체 i 와 j 의 사건 발생 위험비는 시간에 의존하지 않고 항상 일정하게 유지되는 성질이 있어 비례 위험모형이라 한다. 이러한 성질은 모형의 모든 설명변수가 모든 시점에서 위험함수에 일정하게 승법적으로 작용한다는 비례위험 가정으로 Cox PH 모형의 주요 가정이다. Altman 등 (1995)은 130개 논문을 포함하는 5개의 임상 종양학 저널을 검토한 결과, Cox PH 모형을 적용한 43개 논문 중 비례위험 가정의 증명을 언급한 논문은 단 2개임을 확인하였다. 국내의 경우 KOMCI(Korean medical citation index) 웹사이트(<http://komci.org>)에서 2006년 1월부터 2010년 11월까지 5년간 Cox PH 모형을 적용한 114개 논문 중 비례위험 가정의 증명을 언급한 논문은 6개이며 모두 그래프만을 이용하여 확인하였고 그 중 하나는 단지 생존함수가 교차하는지를 평가하였다. 그러나 Cox PH 모형을 통한 분석시 왜곡된 결과를 방지하기 위해서 가정에 대한 확인이 필요하며 비례위험 가정을 위배하는 경우 적절히 모형을 개선해야한다. 비례위험 가정을 위배하는 경우 적절한 모형 전략이 필요하고, 그 대안으로 시간-의존 공변량(time-dependent covariates)을 적용한 모형과 변수를 층화(stratification)한 모형과 같은 확장된 Cox 모형을 이용할 수 있다 (Lee와 Wang, 2003).

2.4.2. 생존분석 자료의 나무모형 나무모형은 의사결정규칙(decision rule)을 도표화하여 관심대상이 되는 집단을 몇 개의 소집단으로 분류(classification)하거나 예측(prediction)을 수행하는 분석방법이다 (Choi와 Seo, 1999). 나무모형은 분류 또는 예측을 목적으로 하는 경우에도 사용될 수 있으며 공변량간 상관관계가 있거나 분석의 정확도보다 분석과정의 설명이 필요한 경우에 더 유용하게 사용된다 (Choi와 Seo, 1999; Ahn과 Loh, 1994). 본 절에서는 서로 특성이 상이한 세 가지 나무모형 및 생존나무모형에 관하여 일반적 사항을 설명한다.

먼저 Breiman 등 (1984)의 CART는 범주형 반응변수에 대해 Gini index를, 연속형 반응변수에 대해 분산의 감소량을 이용하여 이분화(binary split)하는 과정을 반복하여 나무모형을 형성한다. 최대

나무모형을 형성한 이후 과적합(overfitting)문제 방지를 위해 비용-복잡도 척도(cost-complexity measure)를 적용하여 가지친 나무모형(pruned tree)을 만든다. CART는 전체를 탐색하는 방법으로 계산량이 많으며, 변수 선택시 편향(bias)이 발생하는 문제가 있다 (Loh, 2009). CART를 구현한 것이 R의 `rpart`이며, `rpart`는 best prediction error 또는 표준오차(standard error) 1이내 중 가장 복잡도가 높은 나무모형을 가지친다는 점에서 CART와 약간 다른 알고리즘을 가진다. Loh (2002)의 GUIDE는 Bonferroni-adjusted level의 유의성을 검정하여 나무모형을 형성하고 매 분지마다 2단계 아래 노드에서 불순도(impurity)를 감소시키는지 확인한다. 또한 GUIDE는 계산량을 줄이고 변수 선택의 편향이 거의 없을 뿐 아니라 기존의 나무모형이 최종노드(terminal node)에서 반응변수의 평균값을 예측치로 두는 constant 모형에 국한된 것에 반하여 단계적 선형회귀(stepwise linear)모형, 분위수 회귀(quantile regression)모형, 로지스틱 회귀(logistic regression)모형 등을 적합할 수 있는 장점이 있다 (Chang, 2010). Hothorn 등 (2006)의 CTREE는 나무모형 분지를 위한 변수 선택시 순열 검정(permutation test)의 유의확률(p -value)을 이용하는 방법이다. 이 방법은 가지치기를 하지 않으며 정지규칙(stopping rule)은 Bonferroni-adjusted 유의확률에 기초한다.

생존분석 자료의 나무모형인 생존나무(survival tree)모형은 중도절단자료를 다루는데 기존 회귀나무(regression tree)모형과 차이가 있다. `rpart`에 구현된 생존나무모형은 Poisson 회귀모형에 기초하며 분지(splitting) 규칙을 두 Poisson 그룹간 likelihood ratio test로 하고, 노드별 오차를 노드내 이탈도(deviance)로 둔다. 이러한 방법은 모형의 CV(cross-validation) 결과 이탈도가 무한대값을 갖는 문제를 가지며, 이는 노드별 사건발생율(event rate)의 Bayes 추정을 통해 해결이 가능하다. 특히 CART 생존나무모형은 생존시간에 log변환을 하고 중도절단된 자료의 기저위험함수를 선형으로 가정하는 등의 적절한 사전 변환(pre-scaled)을 통해 구현되었다. Therneau 등 (2015)와 LeBlanc과 Crowley (1992)에 `rpart` 생존나무모형에 대해 자세히 연구되어 있다. GUIDE는 노드별로 서로 다른 누적 기저위험함수를 가지므로 노드별로 Cox PH 모형의 적합은 제한되는 문제를 가진다. 그러나 Poisson 회귀모형 반응값으로 Cox PH 모형의 event indicator를, offset 변수로 기저위험함수에 log변환하는 수정을 통해 GUIDE 생존나무모형 구현이 가능하다 (Loh 등, 2015). CTREE 생존나무모형은 log-rank test결과 유의확률을 이용하는 점에서 CART, GUIDE 생존나무모형과 다르다 (Hothorn 등, 2006).

나무모형의 특성 및 성능의 차이는 Loh (2014)와 Lim 등 (2000)에 상세히 연구되어 있다. 본 연구에서 대상으로 하는 나무모형의 큰 차이는 분지규칙으로 CART는 전체 탐색(exhaustive search)를 하는 반면 GUIDE는 주효과(main effect)와 상호작용(interaction) 검정에 의해 비편향(unbiased) 변수를 선택하고, CTREE는 permutation test를 하여 비편향 변수를 선택하므로 변수 선택에 관한 편향 문제가 거의 없다는 것이다. 또한 CART와 CTREE는 Wasserstein Metrics를 이용하여 Kaplan-Meier 생존함수간 측정된 거리를 최대화하도록 불순도 함수를 정의하여 나무모형을 형성한다. 반면 GUIDE는 piecewise Cox PH 모형을 적용하여 나무모형을 형성한다 (Loh, 2014; Bou-Hamad 등, 2011).

3. 분석 결과 및 해석

3.1. Cox PH 모형 결과 및 해석

Table 3.1은 제 2절에서 제시한 변수에 대해 df(degrees of freedom)와 AIC(Akaike Information Criteria)를 최소화하도록 단계적(stepwise) 변수선택법을 적용한 결과를 나타낸 것이다. 이 결과와 군사작전 전문가 의견을 고려하여 POS, BMI, Save, Cmax, Mission을 설명변수로 하는 Model 4를 기본 모형으로 선정하였다.

Table 3.2는 선정한 변수에 대해 Cox PH 모형을 적용한 결과 및 Cox PH 모형의 비례성 가정에 관한

Table 3.1. Result for Stepwise variable selection

Model	Covariate	df	AIC	Difference
1	POS+Sight+BMI+Save+Cmax+SA+IW+Mission	12	3578.13	-
2	POS+Sight+BMI+Save+Cmax+SA+Mission	10	3574.95	3.18
3	POS+BMI+Save+Cmax+SA+Mission	9	3573.84	1.11
4	POS+BMI+Save+Cmax+Mission	8	3573.81	0.03

Table 3.2. Result for Cox PH model and assumption test

Covariate	β	$\exp(\beta)$	S.E.	p -value 1	ρ	χ^2	p -value 2
POS	0.1454	1.1565	0.0519	0.0051	-0.0420	5.20e-01	0.6830
BMI	0.0795	1.0828	0.0556	0.1525	0.0267	0.1704	0.6800
Save	-0.1631	0.8495	0.0800	0.0416	0.0145	0.0536	0.8170
Cmax	0.1651	1.1795	0.0737	0.0252	-0.1694	7.2911	0.0069
Mission (Main attack)	0.8772	2.4041	0.1985	9.88e-06	0.1809	9.6614	0.0019
(Sub attack)	0.7471	2.1109	0.2079	0.0003	0.1711	8.6322	0.0033
(Artillery)	-0.4688	0.6257	0.2285	0.0402	0.2601	19.8840	8.23e-06
(Support)	-0.7444	0.4750	0.3167	0.0187	0.1145	3.8385	0.0501
Global	-	-	-	-	-	29.2691	2.84e-04

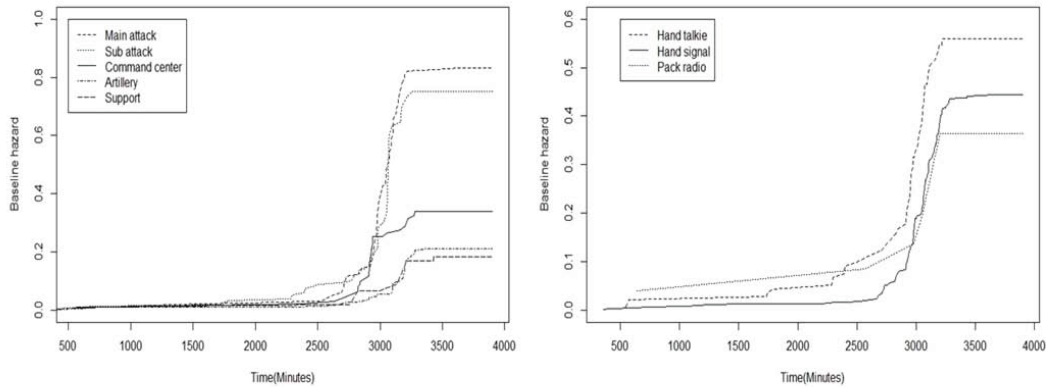


Figure 3.1. Baseline hazard functions for Mission and Cmax.

Grambsch and Therneau 검정 결과이다 (Grambsch와 Therneau, 1994). p -value 1로 볼 때 POS와 Mission이 매우 유의함을 알 수 있다. ρ 는 scaled Schoenfeld 잔차와 시간 변수간 상관계수이며 χ^2 통계량은 시간에 대해 scaled Schoenfeld 잔차의 기울기가 ‘0’이라는 귀무가설에 대한 검정 통계량과 이에 대한 유의확률인 p -value 2로 볼 때 Cmax와 Mission은 시간에 따라 공변량의 효과가 고정됨을 가정하는 귀무가설을 기각할 증거가 충분하므로 Cox PH 모형의 가정을 위배하는 시간-의존 공변량이다.

먼저 시간-의존 공변량 중 Mission은 범주형 변수이므로 층화하여 Cox PH 모형의 변수로 적용하였다. 이 모형에서는 Mission에 따른 각 층(stratum)들이 서로 다른 기저 위험함수를 가진다고 가정하여 위험함수를 추정할 수 있다. 그러나 Mission은 관심있는 변수로 층화 비례모형을 적용할 경우 단지 층화된 부대 임무에 따라 위험비(hazard ratio)가 시간에 따라 변화하며 다른 변수에 대한 영향만을 해석할 수 있으므로 추가적인 방법을 통한 해석이 필요하다. 따라서 본 연구에서는 Figure 3.1 좌측과 같이 Mission의 각 층화된 변수별로 기저위험함수를 추정하여 추가적인 해석을 하였다. 또한 연속형 변수인

Table 3.3. Final Result for Cox PH model

Covariate	β	$\exp(\beta)$	S.E.	<i>p</i> -value
POS	0.0131	1.0130	0.0048	0.0066
BMI	0.0313	1.0320	0.0218	0.1516
Save	-44.6400	4.103e-20	25.5700	0.0809
Cmax(hand talkie 1)	-1.4760	0.2284	0.5584	0.0082
Cmax(hand talkie 2)	-1.2430	0.2884	0.5541	0.0248
Cmax(pack radio)	-1.6080	0.2004	0.5717	0.0049

Cmax는 수신호 확인거리(hand signal), 분대급 무전기 운용거리(hand talkie 1), 소대급 무전기 운용거리(hand talkie 2), 중대급 무전기 운용거리(pack radio)로 구분하여 범주형 변수로 변환하고 Figure 3.1 우측과 같이 층화된 변수별 기저위험함수를 추정하였다. 범주형 변수로 변환한 Cmax의 Grambsch and Therneau 검정 결과 유의확률은 각각 0.3553, 0.4908, 0.4113으로 비례위험 가정을 만족하였다. 최종 모형에서는 Mission은 층화된 변수로 포함하고, Cmax는 전술적 의미를 부여할 수 있는 범주형 변수로 변환하였다.

Table 3.3의 Cox PH 모형 최종 결과에 따른 주요 해석은 다음과 같다. POS가 한 단위 증가할 때 전투불능이 될 위험률이 약 1.013배 증가하고, Cmax(hand talkie 1)가 증가할 때 위험률이 약 0.22배 감소한다. 이는 전투원의 생존성 보장을 위해 제대 및 개인 간격을 적절히 유지해야함을 의미한다. 공격 작전에 관한 군사교리에서 강조하는 전투력 집중이 제대 및 개인 간격을 감소하는 것으로 해석이 된다. 또한 Mission에 따라 전투원의 생존확률에 큰 차이가 있음을 층화된 변수별 기저위험함수로 알 수 있다. Mission을 층화된 변수별로 Cox PH 모형 적용 주요 결과는 지휘부의 경우 타 제대인 중대급 부대와 간격이 증가할 수록 위험률이 증가하는 것으로 볼 때 지휘부 방호체계 및 전투부대와 적절한 간격이 중요함을 알 수 있고, 최선두에서 공격작전을 수행하는 주공 및 적지중심작전부대는 개인 및 제대 간격보다 복무기간과 BMI가 증가할 수록 위험률이 증가하는 결과를 보였다. 주공 및 적지중심작전부대는 소부대 지휘자의 지휘통솔이 중요한 부대로 복무기간이 길수록 위험한 임무를 수행하고 있음을 알 수 있다. 포병부대는 복무기간이 긴 전투원의 생존성이 적으로부터 노출이 많기 때문에 낮으며 야전포간 적절한 간격유지가 전투원 생존에 중요한 요소임을 알 수 있었다.

일반적으로 복무기간이 길수록 전투행위에 대한 숙련도와 전장(combat field)에 대한 익숙함이 높아 생존확률이 높을 것으로 생각된다. 그러나 본 연구에서 적용한 Cox PH 모형의 결과 복무기간이 길수록 위험률이 증가하는 경향을 보였다. Figure 3.2는 복무기간에 대한 생존시간을 중도절단자료를 고려한 local 회귀분석한 결과이다. 대부분의 관측치가 20개월 미만에 집중되어 있으며 군사 전문가 의견에 따르면 그 중 복무기간이 상대적으로 긴 전투원들이 분대장, 기관총 사수, 침병 등 어려운 임무를 수행하고 있어 생존시간이 감소한 것으로 해석된다.

3.2. 나무모형 결과 및 해석

Figure 3.3은 CART, GUIDE, CTREE 방법을 통하여 가지치기가 없는 최대나무모형과 각 프로그램에서 제공하는 최적계수인 디폴트 옵션으로 만든 나무모형을 나타낸 것이다. 모든 모형에서 최초 분지는 Mission을 전투임무와 비전투임무로 구분하는 것으로 동일하게 적용되었다. CART의 최대나무모형에서 비전투임무 부대의 생존성은 BMI가 중요한 변수이며, 전투부대의 경우 Save가 중요한 변수이다. GUIDE의 최대나무모형을 통하여 전투임무를 갖는 부대의 경우 노드 2로 분류되며 이러한 부대는 제대 간격 및 개인 간격이 중요한 변수로 선정됨을 알 수 있다. 노드 3의 비전투임무 부대의 생존확률은 우선

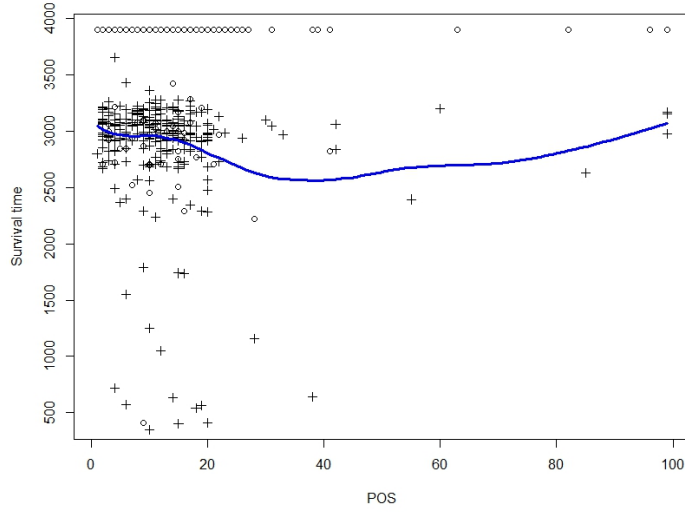


Figure 3.2. Local regression superimposed on the plot of survival time against POS.

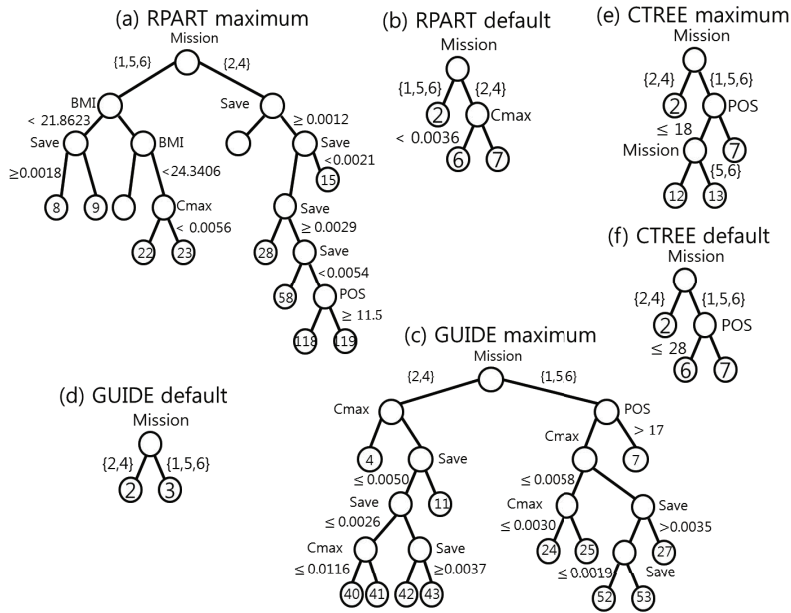


Figure 3.3. Survival tree models.

복무기간이 중요한 변수로 선정되었다. 특히, 복무기간이 17개월 이하는 대부분 의무복무를 하는 병사들로 노드 6으로 분류되며, 비전투임무 부대라도 병사의 경우 제대 및 개인 간격이 중요한 변수임을 알 수 있다. 지휘부 등 비전투임무 부대에서 복무기간이 17개월 이상인 노드 7은 대부분 경험많은 기간 간부이다. CTREE의 최대나무모형에서는 최초 분지이후 복무기간 18개월 이상, 이하로 분지하였다. 특히, 비전투임무 부대 중 복무기간이 18개월 이하인 그룹은 다시 부대 임무를 세분화하여 분류되었다.

RPART는 전체 탐색으로 인한 변수 선택간 편향(selection bias)이 발생하는 문제가 있으며, 이러한 편향은 분지 지점(split point)을 많이 갖는 경우 발생한다 (Loh, 2014). RPART에서 주로 선정된 변수는 각각 783개의 분지 지점을 갖는 Save, Cmax이다. 42개의 분지 지점을 갖는 POS는 Cox PH 모형 결과 중요한 변수임에도 RPART에서는 거의 선정되지 않았다. 반면, GUIDE와 CTREE는 변수 선택간 편향이 거의 없는 방법으로 POS를 중요 변수로 선정하여 나무모형을 구성하였다. 잔차 분석에 기반한 GUIDE와 순열 검정에 기반한 CTREE는 편향이 없는 변수가 선정되는 것으로 알려져 있으며 본 연구에서 일부 상이한 결과를 얻었는데, 나무모형의 알고리즘에 따른 결과의 다양성을 확인할 수 있다.

Figure 2.1 (b), (c), (d)는 CART, GUIDE, CTREE의 디폴트 옵션으로 구성한 나무모형의 최종노드에서 추정된 Kaplan-Meier 생존함수이다. 먼저 RPART는 전투임무 부대를 Cmax로 분지하여 노드 6과 7로 구분하였다. RPART 결과에 따르면 전투임무 부대일 지라도 제대간 거리가 가깝게 유지되면 생존성이 보장됨을 알 수 있다. 다음 CTREE에서는 비전투임무 부대라도 복무기간이 긴 전투원의 생존성은 전투임무 전투원과 유사함을 알 수 있다. 이것은 본부, 의무부대와 같이 비전투임무 부대의 복무기간이 긴 전투원은 분대장을 비롯한 정찰, 첩병 등 위험한 임무를 수행하고 있기 때문으로 해석된다.

4. 결론

본 연구에서는 우리 육군이 보유한 과학화 전투훈련 자료를 생존분석 기법과 나무모형을 응용하여 분석하고 군사교리 측면에서 유용한 결과를 얻기 위해 다양한 측면에서 해석을 시도하였다. 군사교리는 첨단 무기체계의 발전과 더불어 지속적으로 발전해야 할 중요한 분야임에도 불구하고 컴퓨터 기반의 시뮬레이터, 위게임 모델을 이용하여 검증 및 발전만을 해온 것이 사실이다. 물론 전략제대의 군사교리 및 군사적 의사결정 문제 등은 그러한 방법의 적용이 더 합리적이지만, 창끝으로 표현되는 전술제대의 군사교리는 실 기동훈련 자료에 대한 분석이 가장 효과적이고 현실적으로 유일한 방법이다. 이는 전투원 개개인이 명령을 받지만 자율적 행동에 기반하여 기동 및 사격하는 전투행동을 컴퓨터로 묘사하기에 매우 제한되기 때문이다.

본 연구는 군사교리에 명시된 장병 생존성 보장을 위해 취해야 할 전투행동에 대한 통계적 관점의 검증 및 과학적 근거를 제시한 것에 의미가 있다. Cox PH 모형을 통해 복무기간이 증가할 수록 위험한 임무를 부여받고 적에게 노출될 상황이 증가하여 생존성이 감소되며, 전투력 집중을 위해 제대 및 개인 간격을 줄일 때 생존성이 증가한다. 또한 부대임무 변수를 총화하여 적용한 결과 지휘부는 복무기간이 길수록, 다른 중대급 제대와 간격이 증가할 수록 생존성 감소되며, 주공 및 적지중심작전부대의 경우 복무기간과 BMI가 중요한 요소임을 확인하였다. 특히, 복무기간이 길수록 위험도가 증가되는 것은 일반적인 상식과 다른 것으로 군사적으로 중요한 의미를 지닌다고 할 수 있다. 나무모형을 통해 더 상세한 해석을 시도할 수 있었으며, 전투 및 비전투로 구분되는 부대임무는 본 연구의 자료를 설명하는 가장 중요한 변수였다. 그러나 동일한 전투부대라 할지라도 제대간 거리에 따라 생존함수가 상이하며, 동일한 비전투 부대에서도 복무기간에 따라 생존함수가 다름을 통해 부대임무 뿐 아니라 본 연구에서 설정한 다른 변수들이 전투원 생존성에 중요함을 알 수 있다.

군사적으로 중요한 요소를 나타내는 METT+TC(mission, enemy, terrain and weather, troops and support available, time available, civil considerations) 중 본 연구에서 분석한 대대급 전술제대 공격 작전의 설명변수로 포함하지 못한 온도지수, 각종 지원 및 배속부대 능력 등과 공격작전 교리에는 있으나 본 연구에서 포함하지 못한 기동속도 유지, 주요 부대간 이격정도 등 다양한 설명변수는 추가적인 훈련부대 자료 확보를 통해 모형의 보완이 필요하다. 특히, BMNT(beginning morning nautical twilight), EENT(end of evening nautical twilight), 산악지역, 도로건부지역 등 군사적으로 중요한 의미를 갖는 시공간 효과를 모형에 포함하는 것은 추후 연구로 남긴다.

References

- Ahn, H. and Loh, W. Y. (1994). Tree-structured proportional hazards regression modeling, *Biometrics*, **50**, 471–485.
- Altman, D. G., De Stavola, B. L., Love, S. B. and Stepniewska, K. A. (1995). Review of survival analyses published in cancer journals, *British Journal of Cancer*, **72**, 511–518.
- Barney, J. A. (2002). Study of patent mortality rates: Using statistical survival analysis to rate and value patent assets, *AIPLA Quarterly Journal*, **30**, 317–352.
- Bou-Hamad, I., Larocque, D. and Ben-Ameur, H. (2011). A review of survival trees, *Statistics Surveys*, **5**, 44–71.
- Breiman, L., Friedman, J. H., Olshen, R. A. and Stone, C. J. (1984). *Classification and Regression Trees*, Wadsworth, Belmont.
- Chang, Y. (2010). The analysis of factors which affect business survey index using regression trees, *The Korean Journal of Applied Statistics*, **23**, 63–71.
- Choi, J. and Seo, D. (1999). Decision trees and its applications, *Journal of The Korean Official Statistics*, **4**, 61–83.
- Cox, D. R. (1972). Regression models and life-tables, *Journal of the Royal Statistical Society, Series B*, **34**, 187–220.
- Grambsch, P. and Therneau, T. (1994). Proportional hazards tests and diagnostics based on weighted residuals, *Biometrika*, **81**, 515–26.
- Herl, B. K., Doe, W. W. and Jones, D. S. (2005). Use of military training doctrine to predict patterns of maneuver disturbance on the landscape. I. Theory and methodology, *Journal of Terramechanics*, **42**, 353–371.
- Hodson, D. D. and Baldwin, R. O. (2009). Characterizing, measuring, and validating the temporal consistency of live virtual constructive environments, *Simulation*, **85**, 671–682.
- Hothorn, T., Hornik, K. and Zeileis, A. (2006). Unbiased recursive partitioning a conditional inference framework, *Journal of Computational and Graphical Statistics*, **15**, 651–674.
- Kim, G. G. and Kim, D. S. (2014). Development and application of effect measurement tool for victory factors in offensive operations using big data analytics, *Journal of the Korean Operations Research and Management Science Society*, **39**, 111–130.
- Kim, T. G., Kwon, S. J. and Kang, B. (2013). Modeling and simulation methodology for defense systems based on concept of system of systems, *Journal of the Korean Institute of Industrial Engineers*, **39**, 450–460.
- LeBlanc, M. and Crowley, J. (1992). Relative risk trees for censored survival data, *Biometrics*, **48**, 411–425.
- Lee, E. T. and Wang, J. (2003). *Statistical Methods for Survival Data Analysis*, John Wiley & Sons, New Jersey.
- Lim, T. S., Loh, W. Y. and Shih, Y. S. (2000). A comparison of prediction accuracy, complexity, and training time of thirty-three old and new classification algorithms, *Machine Learning*, **40**, 203–228.
- Loh, W. Y. (2002). Regression trees with unbiased variable selection and interaction detection, *Statistica Sinica*, **12**, 361–386.
- Loh, W. Y. (2009). Improving the precision of classification trees, *Annals of Applied Statistics*, **3**, 1710–1737.
- Loh, W. Y. (2014). Fifty years of classification and regression trees, *International Statistical Review*, **82**, 329–348.
- Loh, W. Y., He, X. and Man, M. (2015). A regression tree approach to identifying subgroups with differential treatment effects, *Statistics in Medicine*, **34**, 1818–1833.
- Schmidt, P. and Witte, A. D. (1989). Predicting criminal recidivism using ‘split population’ survival time models, *Journal of Econometrics*, **40**, 141–159.
- Therneau, T., Atkinson, B. and Ripley, B. (2015). Package ‘rpart’.
- U.S. Army (2008). *FM 3.0 Operations*, Headquarters Department of the Army.

전술제대 공격작전간 전투원 생존성에 관한 연구

김재오^a · 조형준^{a,1} · 김각규^b

^a고려대학교 통계학과, ^b육군본부 분석평가단

(2015년 6월 29일 접수, 2015년 7월 9일 수정, 2015년 8월 10일 채택)

요약

본 연구에서는 증강된 보병대대의 과학화 전투훈련 데이터 중 공격작전에 관한 장병들의 생존분석을 실시하였다. 과학화 전투훈련은 KCTC(Korea Combat Training Center)로 불리는 전투훈련장에서 MILES(Multiple Integrated Lazer Engagement System)와 중앙통제장비체계 등 과학화된 훈련장비와 체계 운용하 훈련부대가 적 전술 및 무기체계를 사용하는 전문 대항군과 실시하는 쌍방 자유기동훈련이다. 이는 훈련기간 동안 훈련지역의 모든 데이터가 저장되어 훈련통제 뿐 아니라 분석 및 사후검토를 할 수 있는 첨단화된 군사 훈련으로 통계적 분석이 가능한 데이터를 제공한다. 분석방법은 모수적 분포 가정이 필요하지 않은 Cox의 비례위험모형을 적용하였으며, 보다 풍부하고 용이한 해석을 위해 의사결정나무모형(CART(Classification and Regression Trees), GUIDE(Generalized, Unbiased, Interaction Detection and Estimation), CTREE(Conditional Inference Trees))을 활용하였다. Cox 비례위험모형의 비례성 가정을 확인하여 이를 위배하는 변수에 대해서 증화하여 분석하고, Cox 비례위험모형 결과 복무기간에 관한 해석이 용이하지 않아 단변량으로 local 회귀분석을 통해 추가적인 해석을 시도하였다. CART, GUIDE, CTREE는 모형의 특성별로 나무모형을 형성하며 이를 통하여 다양한 해석이 가능하다.

주요용어: 과학화 전투훈련, 공격작전, Cox의 비례위험모형, CART, GUIDE, CTREE

고려대학교 연구비에 의하여 수행되었음 (K1507651).

¹교신저자: (136-701) 서울특별시 성북구 안암로 145, 고려대학교 정경관 414호, 통계학과.

E-mail: hj4cho@korea.ac.kr