

기계학습을 이용한 노면온도변화 패턴 분석

Analysis of Road Surface Temperature Change Patterns using Machine Learning Algorithms

양충헌	Yang, Choong Heon	정회원 · 한국건설기술연구원 도로연구소 연구위원 (E-mail : chyang@kict.re.kr) 과학기술연합대학원대학교 교통물류 및 ITS공학과 교수
김승범	Kim, Seoung Bum	국립경상대학교 건축도시토목공학부 교수 (E-mail : kimsb@gnu.ac.kr)
윤천주	Yoon, Chun Joo	정회원 · 한국건설기술연구원 도로연구소 수석연구원 (E-mail : cjyoon@kict.re.kr)
김진국	Kim, Jin Guk	정회원 · 한국건설기술연구원 도로연구소 연구원 (E-mail : jingukkim@kict.re.kr)
박재홍	Park, Jae Hong	정회원 · 한국건설기술연구원 도로연구소 연구원 (E-mail : jhpark@kict.re.kr)
윤덕근	Yun, Duk Geun	정회원 · 한국건설기술연구원 도로연구소 수석연구원 · 교신저자 (E-mail : dkyun@kict.re.kr) 과학기술연합대학원대학교 교통물류 및 ITS공학과 교수

ABSTRACT

PURPOSES : This study suggests a specific methodology for the prediction of road surface temperature using vehicular ambient temperature sensors. In addition, four kind of models is developed based on machine learning algorithms.

METHODS : Thermal Mapping System is employed to collect road surface and vehicular ambient temperature data on the defined survey route in 2015 and 2016 year, respectively. For modelling, all types of collected temperature data should be classified into response and predictor before applying a machine learning tool such as MATLAB. In this study, collected road surface temperature are considered as response while vehicular ambient temperatures defied as predictor. Through data learning using machine learning tool, models were developed and finally compared predicted and actual temperature based on average absolute error.

RESULTS : According to comparison results, model enables to estimate actual road surface temperature variation pattern along the roads very well. Model III is slightly better than the rest of models in terms of estimation performance.

CONCLUSIONS : When correlation between response and predictor is high, when plenty of historical data exists, and when a lot of predictors are available, estimation performance of would be much better.

Keywords

Surface Road Temperature Change, Machine Learning, Road Weather, Thermal Mapping, Ambient Temperature

Corresponding Author : Yun, Duk Geun, Senior Researcher
Highway & Transportation Research Institute, Korea Institute of Civil
Engineering and Building Technology, 283 Goyangdae-ro, Ilsanseo-gu,
Goyang-si, Gyeonggi-do, 10223, Korea
Tel : +82.31.910.0159 Fax : +82.31.910.0746
E-mail : dkyun@kict.re.kr

International Journal of Highway Engineering

http://www.ksre.or.kr/

ISSN 1738-7159 (print)

ISSN 2287-3678 (Online)

Received Jan. 19, 2017 Revised Jan. 23, 2017 Accepted Mar. 31, 2017

1. 서론

1.1. 연구의 배경 및 목적

동절기에는 노면결빙으로 인한 교통사고 발생건수와

그에 따른 부상자 수가 많이 발생한다. 해외 연구에서도 동절기와 동절기의 기간에 발생한 각종 자료를 비교 분석한 결과, 동절기 교통사고는 약 19%, 부상자수는 약 13%

증가하는 것으로 나타났다(Black and Mote 2015).

이와 같이 동절기 노면결빙은 도로이용자에게 매우 큰 위협이 될 수 있다. 노면결빙으로 인한 미끄럼 사고는 기상상태, 제설작업 상태, 운전자의 부주의와 같은 외생적 요소들이 복합적으로 작용하여 발생하지만 가장 근본적인 원인은 노면온도가 낮아짐에 따라 차량의 타이어와 노면사이의 마찰계수가 감소하기 때문이다. 따라서 이러한 노면온도에 대한 적절한 모니터링과 이에 따른 도로 관리 및 교통운영전략이 알맞게 수립된다면 동절기 미끄럼 사고로 인한 인적·물적 피해를 상당히 감소시킬 수 있을 것으로 판단된다. 현재의 노면온도 파악 방법은 크게 3가지로 구분될 수 있다. 먼저, 교통량 검지기처럼 도로 포장면에 노면온도센서를 설치하여 온도데이터를 측정하는 방법이 있다. 이는 가장 일반적인 방법으로 널리 활용되기는 하나, 도로전체 중 특정 지점에 대한 온도데이터만 취득할 수 있고, 포장면 보수 시 쉽게 파손될 수 있는 단점이 있다. 둘째는 적외선 방식으로 노면온도를 측정하는 것인데, 이는 앞서 언급한 포장에 설치하는 방법 보다는 비용이 저렴하나, 수집데이터에 대한 신뢰도에 있어서 다소 열악한 면이 있다고 알려져 있다. 마지막으로 이동식 노면온도측정시스템(Thermal Mapping)은 차량에 부착하여 주행하면서 도로구간의 노면온도데이터를 취득하는 장비로서, 이것도 적외선 방식을 활용하고 있다. 언급된 3가지 방식은 모두 장·단점을 가지고 있는데, 특히 전체적인 도로구간의 노면온도를 측정하기 어렵고, 측정한다고 해도 비용 대비 효과적이지 않다. 이전 연구에서는 Thermal Mapping 장비를 활용해 수집한 노면온도의 신뢰성을 통계적 기법을 통해 평가하고, 평가된 노면온도를 기준으로 도로를 분할하는 방법을 제시한 바 있다(양충현 외, 2016).

본 연구에서는 이전연구를 기반으로 비용·효과적이면서도 대량의 온도데이터를 수집·활용·예측할 수 있는 새로운 접근법을 시도하였다. 다시 말하면, 본 연구에서는 일반 차량에 내재되어 있는 외기온도센서로부터 수집되는 온도데이터를 활용하여 도로노면온도변화 패턴을 추정하는 모형을 기계학습(Machine Learning)을 통해 구축하였다.

2. 문헌고찰

동절기 노면결빙으로 인한 미끄럼 사고를 예방하기 위해 핀란드의 VTT 기술연구센터에서는 도로의 결빙 상태를 운전자에게 사전에 경고하기 위해 Fig. 1과 같이 블랙

아이스 검지 및 경고 기술을 개발하고 있다. 결빙 여부는 기상정보와 다수의 차량들로부터 수집되는 차량의 구동축과 회전축의 속도 변화를 기반으로 결정된다(VTT Technical Research Center of Finland, 2013).



Fig. 1 Slipperiness Detection System in Finland

미국의 경우 블랙아이스 발생 지점을 예측하고 위험 구간 정보를 제공하기 위해 블랙아이스 발생 예측 모형을 개발하고 이에 필요한 데이터 수집은 도로기상정보시스템 및 각종 기상자료를 활용하고 있다(Bukkapatnam et. al, 2014). 또한, Remote sensing 기술이나 스테레오 카메라와 같은 영상 장비를 활용하여 노면의 건조, 습윤, 결빙상태 정도를 구분하는 기술도 실용화 단계에 있다(Jonsson, 2011; Omer and Liping, 2010).

국내의 경우 결빙방지시스템을 적용한 도로포장체에 대한 연구가 진행되고 있으며, 일부 도로 구간에 적용된 사례도 있다. 한국도로공사에서는 어는비(Freezing rain)예측 시스템을 개발하여 영동고속도로 구간에서 시범운영하고 있다. 위에서 개발하거나 개발 중인 기술들은 다소 복잡한 모형을 활용하고 대부분 공공부문관점에서 도로관리 목적으로 활용된다.

본 연구에서는 불특정 다수차량으로부터 외기온도데이터를 수집하고 기계학습 방법 적용을 통해 노면온도를 추정하기 때문에 많은 비용이 필요하지 않고, 도로망 전체의 추정이 가능하다는 장점이 있다. 도로교통분야에서는 이미 오래 전부터 기계학습기법을 적용하여 다양한 패턴을 추정하는데 활용되어 왔다. 기계학습(Machine Learning)은 지식습득 절차를 자동화하는 계산방식에 대한 연구 분야이다. 일반적인 전략은 학습데이터군 안에서 패턴을 발견하는 것이다. 이러한 패턴은 새로운 데이터의 행태를 예측하는데 사용된다. 기계학습 기법의 주요 분류는 다음과 같다(Mahapatra and Bose, 2001).

- 규칙 유도(Rule induction)
- 신경망(Neural networks)
- 사례기반 추론(Case-based reasoning)
- 유전자 알고리즘(Genetic algorithm)
- 귀납적 논리 프로그래밍(Inductive logic programming)

Antoniou and Koutsopoulos(2006)의 연구에서는 기계학습방법을 적용하여 교통속도와 밀도의 관계를 규명하였다. 이 연구에서 기계학습 방법이 기존의 다른 모형보다 속도 추정에 있어서 정확도를 현저하게 향상할 수 있음을 입증하였다.

Chen and Xie(2015)의 연구에서는 대형 상업용 화물차량 운전자의 운전행태와 관련한 패턴을 파악하는데 기계학습 기법을 적용하였다. 이를 통해 대형 상업용 화물차량관련 교통사고의 위험도를 예측하였다.

국내에서도 전우혁 외(2016)은 대량의 날씨 정보와 교통소통 데이터를 분석하고 예측하기 위해서 기계학습 방법을 적용하였다. 본 연구는 분산 시스템 환경 하에서 작동하는 기계학습 알고리즘을 이용하여 기상과 이와 관련한 교통 정체 구간 예측 시스템 개발을 제안하였다.

카이스트에서는 빅데이터 분석 및 기계학습 기법으로 미래 교통상황을 예측하고 그 결과를 공유할 수 있는 방법을 연구하고 있다. 고속도로에 설치된 VDS, DSRC, TCS 등의 인프라 데이터 등을 기반으로 속도, 교통류, 밀도 등과 같은 상태를 예측할 수 있다. 이를 위해 개발된 '리얼트래픽(REALTraffic)' 엔진은 기존 데이터 기반의 추측과 시뮬레이션을 통한 예측을 통해 가장 최적화된 결과를 보여준다(IT동아, 2016).

이와 같이 기계학습알고리즘을 활용하여 도로교통분야에 적용한 사례가 증가하는 추세에 있으며, 본 연구의 목적과 같이 단순한 수학적 모형을 통해 자세히 묘사하지 못하는 경우에 매우 적절한 접근법인 것으로 판단된다.

3. 자료수집

본 연구의 분석지역은 자유로 성동 IC~이산포 IC 왕복 34km 구간이며 2015년 10월에 1차적으로 5차레에 걸쳐 외기온도와 노면온도데이터를 수집하였다. 2016년 10월에도 동일구간에 대해 총 4차레에 걸쳐 동일한 종류의 데이터를 수집하였다. 이때는 외기온도와 더불어 대기온도와 습도 데이터를 추가적으로 수집하였다.

Table 1은 2015년과 2016년에 수집한 데이터와 날씨 조건을 요약한 것이다. Table 1에 A는 맑고 바람이 없는 조건을, B는 구름과 바람이 조금 있는 조건을 C는 구름, 습도, 바람이 모두 있는 조건을 나타낸다.

Table 1. Data Overview

Date	Origin	Destination	Weather condition
15.10.16	SungDong IC	Isanpo IC	B
15.10.23	Isanpo IC	SungDong IC	B
15.10.23	SungDong IC	Isanpo IC	B
15.10.26	Isanpo IC	SungDong IC	C
15.10.28	SungDong IC	Isanpo IC	A
16.10.12	Isanpo IC	SungDong IC	B
16.10.25	Isanpo IC	SungDong IC	B
16.10.26.	Isanpo IC	SungDong IC	B
16.10.26.	Isanpo IC	SungDong IC	C

데이터 수집을 위해서 노면온도, 대기온도, 그리고 습도를 복합적으로 수집할 수 있는 Thermal Mapping System을 활용하였다. 이는 차량에 탑재 가능한 것으로서 특정 장소에 제약 없이 설치할 수 있는 이동형 센서라는 장점이 있다. 본 연구에서 활용된 Thermal Mapping System의 센서는 한국기상산업진흥원의 기상측기검사 방법에서 제시하는 기상장비와 오차범위를 만족한다. 기계식 온도계의 경우 최대 오차범위가 $\pm 1.0^{\circ}\text{C}$ 이며, 전자식은 $\pm 0.3^{\circ}\text{C}$ 이하로 제시되었다. 본 연구에서 사용한 온도계는 전자식이므로, 오차범위 $\pm 0.3^{\circ}\text{C}$ 를 만족하는 센서를 선정하였다.

또한, 습도센서의 경우도 오차범위를 기계식은 $\pm 5\%$, 전자식은 $\pm 3\%$ 기준을 제시했으며, 본 허용범위를 만족하기 위하여 대기 습도 측정용 장비의 오차범위가 $\pm 3\%$ 이하인 센서를 선정하였다. 조사차량에는 자료 측정 위치 파악을 위한 GPS 장치와 자료를 수집·표시 및 저장하는 컴퓨터와 이들을 처리하기 위한 각종 소프트웨어들이 탑재되어 있다.

차량의기온도센서의 경우, 2개의 센서를 조사차량의 전면부와 후면부에 각각 설치하였다. 본 센서는 USB 연결단자를 통해 현재 측정된 온도, 습도 정보를 실시간으로 전송해준다.

노면온도 데이터는 2015년 총 4,665개, 2016년 3,187개 수집되었으며, 분석을 위해서 이 중 절반에 해당하는 데이터를 임의로 추출(Random Sampling)하여 기계학습을 통한 모형 구축에 사용하였다. 나머지 절

반의 데이터는 구축된 모형을 평가하는데 사용하였다.

4. 분석방법론

기계학습의 핵심은 표현(Representation)과 일반화(Generalization)이다. 표현은 데이터의 평가이며, 일반화는 아직 알 수 없는 데이터에 대한 처리를 의미한다. 기계학습의 기본원리는 예측변수(Predictor)와 반응변수(Response)와의 관계를 모르는 상태에서 주어진 예측변수들과 그에 상응하는 반응변수와의 관계를 반복적으로 학습시킨 후 새로운 예측변수에 대한 반응변수를 추정하는 것이다.

양충현 외(2016) 연구에서는 제한된 환경 하에서 동일노선, 동일 기상조건 하에서의 노면온도변화 패턴의 유사성을 확인하였다. 대표적인 통계적 기법인 상관분석(Correlation Analysis)을 적용하였는데, 분석결과 주어진 조건 하에서 노면온도변화 패턴에 분명한 통계적 유의성이 있음을 증명하였다. 이 연구에서 가장 중요한 부분인 노면온도변화와 관련한 예측 변수로는 조사차량에 부착된 2개의 차량센서(전방외기온도센서와 후방외기온도센서)로부터 수집된 외기온도였다. 따라서 모델링 시 고려한 주요 예측 변수는 이 두 가지 외기온도데이터를 기본으로 활용하고, 추가적으로 Thermal Mapping System에서 수집한 습도와 대기온도데이터 일부를 모형 구축에 활용하였다.

4.1. 2015년 노면온도와 외기온도 비교

2015년에 수집된 노면온도데이터 중 절반을 무작위 추출(Random Sampling)하였는데, 이는 조사구간 및 환경이 다소 제한적이어서 온도데이터 자체가 기계학습을 통해 구축된 모형의 성능 역시 제한적일 수 밖에 없

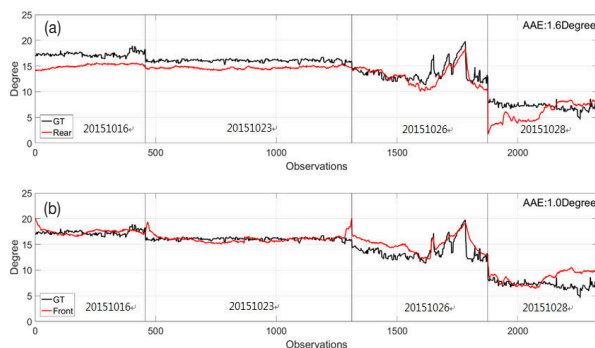


Fig. 2 (a) Ambient Temp. of Rear, (b) Ambient Temp. of front and Road Surface Temp. for Development of Model I

을 것으로 판단하였기 때문이다. 이를 학습용으로 사용하여 모형 I을 구축하였다. 모형 I의 반응변수는 노면온도이며, 예측변수는 조사차량의 전·후방 외기온도 데이터를 사용하였다. Fig. 2에서는 노면온도와 전·후방 외기온도 데이터를 각각 비교한 예시를 나타낸다. 그림 안의 수직선은 데이터 수집일 경계를 나타낸다.

두 개의 시계열 데이터간 유사도를 측정하기 위해 본 연구에서는 절대평균오차(Average Absolute Error, AAE)를 계산하여, Fig. 2의 오른쪽 위에 나타내었다. 절대평균오차는 다음과 같이 계산된다.

$$\text{절대평균오차(AAE)} = \frac{\sum_{i=1}^n \text{시계열}1_i - \text{시계열}2_i}{n}$$

여기서, 시계열 1_i : 시계열 1의 i 번째 데이터

시계열 2_i : 시계열 2의 i 번째 데이터

조사차량의 후방외기온도의 경우 AAE가 1.6°C인 반면 전방외기온도는 1.0°C로 전방외기온도가 실제 노면온도에 더 근접하는 것을 알 수 있다. 전반적으로 전방외기온도는 노면온도보다 높은 경향을 보이는데 이는 전방외기온도센서가 차량 전면에 장착되어 엔진룸 내부 온도, 차량 자체의 열 등으로부터 영향을 받기 때문인 것으로 추정된다. 반면 후방외기온도는 실제 노면온도보다 낮은 경향을 나타낸다. 조사구간의 노면온도는 대체적으로 큰 변화없이 안정적인 상태를 보이지만 10월 26일, 28일 조사 데이터처럼 짧은 순간에 갑작스러운 온도변화를 보이거나 이 때 외기온도센서는 그 변화에 민감하게 반응하지 않는 것으로 판단된다.

4.2. 2015년 데이터 기반 노면온도 추정 결과

모형 I을 평가하기 위해 무작위 추출된 데이터의 나머지 데이터를 활용하였다. Fig. 3은 모형 I 평가를 위해 사용할 데이터와 노면온도를 비교한 것이다.

Fig. 2의 결과와 유사한 AAE 그리고 10월 26일, 28일에 관측되었던 온도 간 차이 등이 Fig. 3에서도 관측된다. 따라서 모형 구축과 평가에 사용한 전·후방외기온도와 노면온도의 유사도는 크게 다르지 않음을 알 수 있다. Fig. 4는 모형 I을 이용한 노면온도변화 패턴 추정치와 실제 노면온도변화 패턴을 비교한 결과를 나타낸다.

추정된 노면온도변화 패턴과 실제 노면온도변화 패턴

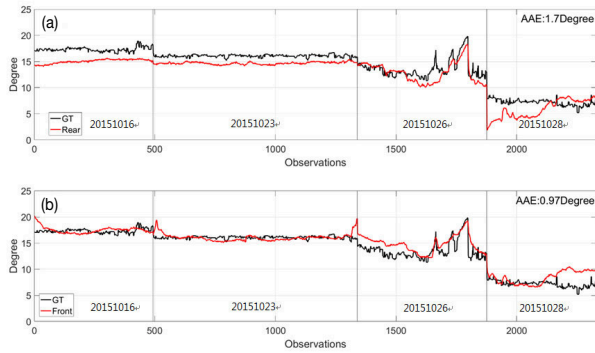


Fig. 3 (a) Ambient Temp. of Rear, (b) Ambient Temp. of front and Road Surface Temp. for Evaluation of Model I

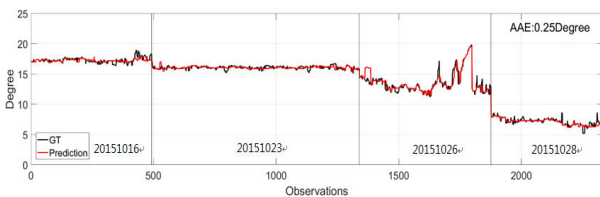


Fig. 4 Estimated Road Surface Temperature using Model I

을 비교했을 때 AAE는 0.25℃로 전·후방외기온도보다 평균적으로 실제 노면온도와 더 유사함을 알 수 있다. 또한 10월 26일과 28일에 갑작스런 온도변화에 대해서도 추정치는 민감하게 반응하는 것으로 나타났다.

총 수집된 노면온도데이터 중 절반인 2,333개 데이터를 활용하여 모형 I을 구축하였고, 나머지 절반의 노면온도데이터를 평가에 활용하여 노면온도변화 패턴 추정치와 실제 노면온도변화 패턴을 비교하여 모형 I의 정확도를 검증하였다. 만약 모형 적용없이 단순히 수집된 외기온도를 이용하여 노면온도변화 패턴을 추정한다면, 절대평균오차를 크게는 1.7℃, 적게는 1.0℃ 기대할 수 있으나, 구축된 모형을 이용할 경우 0.25℃정도 됨을 확인하였다. 결과적으로 노면온도변화를 추정하기 위해 외기온도 데이터를 이용할 때 보다 75%~85%정도 향상된 유사도를 구축된 모형으로부터 기대할 수 있을 것으로 판단된다.

4.3. 2016년 노면온도와 외기온도 비교

2016년에 수집한 데이터를 바탕으로 두 가지 모형을 구축·평가하였다. 2015년과 마찬가지로 2016년에 수집한 노면온도 데이터를 바탕으로 무작위 추출(Random Sampling)을 수행하여 데이터의 절반을 추출하고 이를 바탕으로 구축된 것이다. 모형 II는 외기온도, 습도,

그리고 대기온도가 고려되었지만 모형 III에서는 이들을 고려하지 않았다. 즉, 모형 II는 노면온도변화 패턴 추정치에 대한 습도와 대기온도의 영향을 분석하기 위해 구축되었다. Fig. 5는 노면온도와 전·후방외기온도를 비교한 것을 나타내며, Fig. 6은 2016년에 추가로 수집된 습도와 대기온도 데이터의 예를 나타낸다.

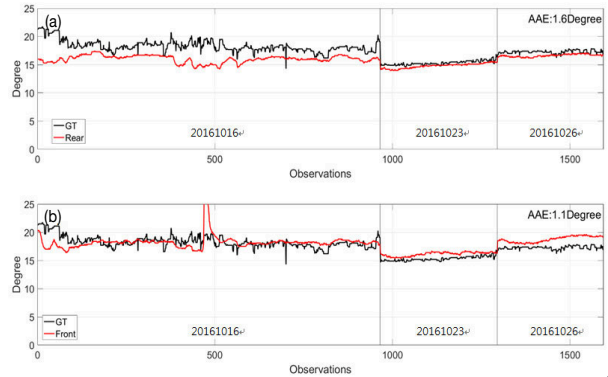


Fig. 5 (a) Ambient Temp. of Rear, (b) Ambient Temp. of front and Road Surface Temp. for Development of Model II

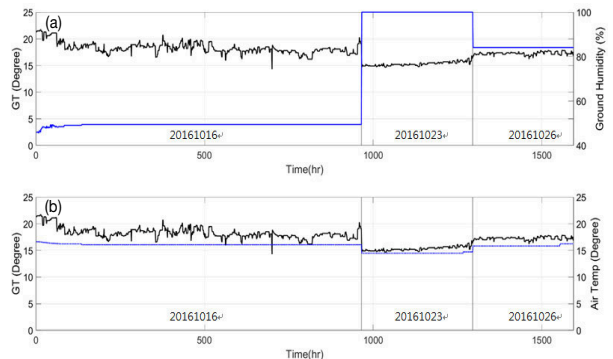


Fig. 6 (a) Humidity, (b) Ambient and Road Surface Temp. for Development of Model II

외기온도와 노면온도의 유사도 측면에서 2016년 데이터는 2015년과 크게 다르지 않음을 알 수 있다. 다만, 전방외기온도의 AAE는 2015년 때보다 0.1℃ 상승한 것을 알 수 있다.

습도와 대기온도의 경우 수집기간 동안에 큰 변화없이 일정한 값을 나타내고 있다. 다만 대기온도의 경우 노면온도 대비 낮은 수치를 보이고 있지만 일별 변화에는 민감한 것으로 보아 노면온도변화 패턴 추정에 유용할 것으로 판단된다. 모형 II의 경우 Fig. 4와 Fig. 5에 모든 노면온도 데이터를 사용하여 구축하고, 모형 III의 경우는 Fig. 5의 데이터만을 이용하여 구축하였다.

4.4. 2016년 데이터 기반 노면온도 추정 결과

2015년과 마찬가지로 무작위 추출하고 남은 절반의 데이터를 토대로 모형 II와 모형 III을 통해 노면온도 변화 패턴 추정치와 실제 노면온도변화 패턴을 비교·평가하였다. 평가에 사용하는 데이터는 Fig. 7과 Fig. 8과 같다.

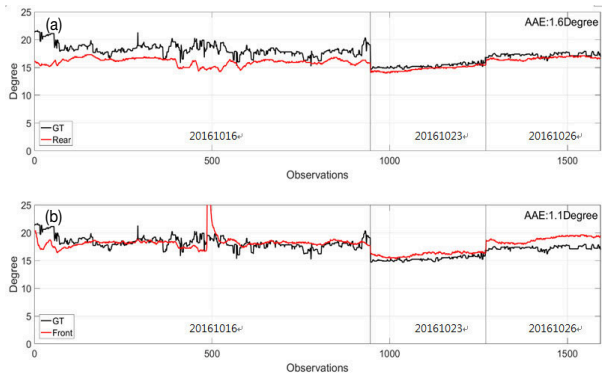


Fig. 7 (a) Ambient Temp. of Rear, (b) Ambient Temp. of front and Road Surface Temp. for Evaluation of Model II

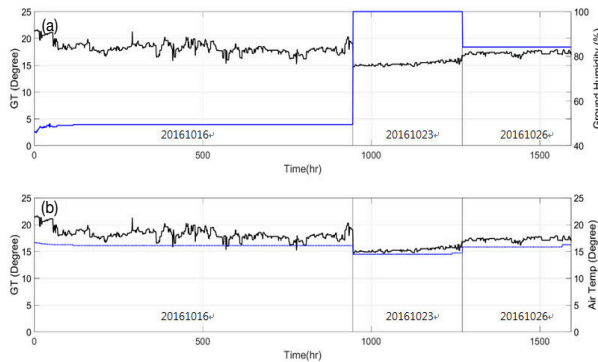


Fig. 8 (a) Humidity, (b) Ambient and Road Surface Temp. for Evaluation of Model II

조사차량의 전·후방외기온도 Fig. 7과 습도 및 대기 온도(Fig. 8)를 모형 II에 입력 후 노면온도변화 패턴을 추정하였으며 실제 노면온도변화 패턴과의 비교 결과는 Fig. 9와 같다.

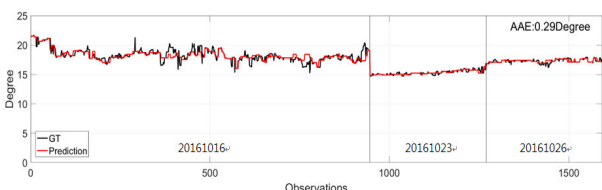


Fig. 9 Estimated Road Surface Temperature using Model II

Fig. 9에서 10월 16일은 실제 노면온도변화폭이

$\pm 2.5^{\circ}\text{C}$ 내외로 같은 해 다른 날에 비해 큰 것으로 보인다. 이것은 추정된 노면온도변화 폭 역시 비슷한 패턴인 것으로 판단된다. 하지만 Fig. 7에서 알 수 있듯이 외기온도 역시 유사한 문제점을 지니고 있다. 추정된 노면온도의 AAE는 0.29°C 이며, 외기온도로 노면온도를 추정했을 때 보다 유사도 측면에서 모형을 활용하는 것이 보다 바람직한 것으로 판단된다. Fig. 9는 습도와 대기온도를 제외한 외기온도만으로 노면온도를 추정한 결과를 보여주고 있다.

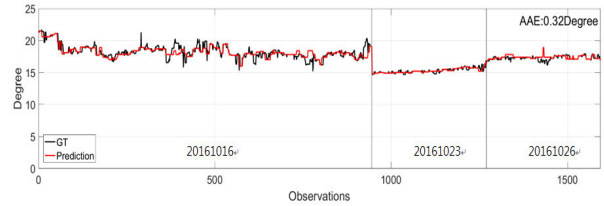


Fig. 10 Estimated Road Surface Temperature using Model III

Fig. 10을 살펴보면, Fig. 9와 비교했을 때 유사성 측면에서 모형 II의 AAE는 0.29°C 인 반면 모형 III은 0.32°C 로 차이가 분명하지는 않지만 모형 II가 모형 III보다 상대적으로 설명력이 높은 것을 알 수 있었다. 즉, 모형 구축 시 습도와 대기온도를 예측 변수에 추가하면 노면온도변화 패턴 추정에 긍정적인 영향을 주는 것으로 판단된다.

2016년에 수집한 노면온도 데이터 중 절반인 1,594개의 데이터를 바탕으로 모형 II와 모형 III을 구축하였다. 모형 II와 모형 III의 차이는 2016년부터 수집하기 시작한 습도와 대기온도를 예측 변수로 추가 유·무의 차이이다. 나머지 절반의 노면온도 데이터를 평가에 활용하기 위해 모형 II와 모형 III에 대입하고 노면온도변화 패턴 추정치와 실제 노면온도변화 패턴을 비교·분석하였다. 전체적인 결과는 2015년의 결과와 크게 다르지 않으나, 습도와 대기온도를 예측 변수로 추가한 모형 II가 모형 III보다 더 나은 유사도를 나타내었다.

4.5. 통합 데이터 분석

기계학습을 통해 구축된 모형의 성능은 학습된 데이터의 양과 학습된 데이터의 질(반응변수와 예측변수간의 적절한 상관관계)에 의존된다. 따라서 본 절에서는 2015년과 2016년에 공통적으로 수집된 데이터인 조사차량의 전·후방외기온도 데이터만을 활용하여 모형 IV를 구축하고 해당 모형을 평가하였다. Fig. 11은 모형 IV를 구축하기 위해 사용한 전·후방외기온도와 노면온

도를 나타낸다. 또한, Fig. 12는 구축된 모형을 평가할 전·후방외기온도와 노면온도를 나타낸다.

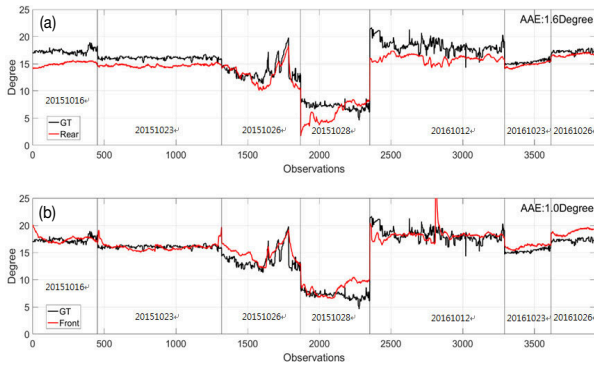


Fig. 11 (a) Ambient Temp. of Rear, (b) Ambient Temp. of front and Road Surface Temp. for Development of Model IV

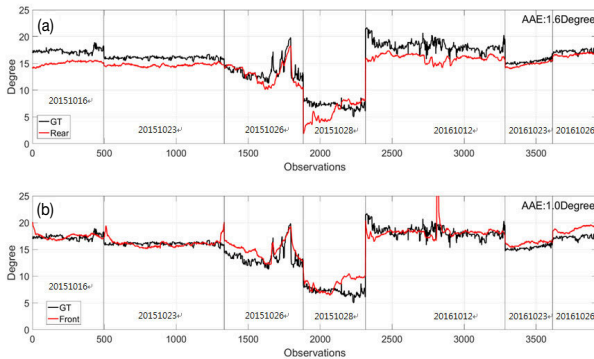


Fig. 12 (a) Ambient Temp. of Rear, (b) Ambient Temp. of front and Road Surface Temp. for Evaluation of Model IV

Fig. 12는 Fig. 11의 데이터를 기반으로 구축한 모형 IV에 Fig. 12의 전·후방외기온도를 이용하여 추정된 노면온도변화 패턴과 실제 노면온도변화 패턴과의 비교 분석결과를 나타낸다. 전반적으로 2015년과 2016년의 분석 결과를 통합한 것과 크게 다르지 않다. 하지만 AAE를 기반으로 유사도를 살펴보면 모형 IV의 AAE가 다른 3개 보다 큰 것으로 나타났다.

2015년과 2016년의 노면온도변화 패턴을 비교 분석해보면 대부분의 노면온도가 15℃~20℃ 사이에 분포된 것을 알 수 있다. 또한, 이에 상응하는 외기온도변화 패턴을 비교 분석해보면, 2015년의 경우 후방외기온도는 노면온도보다 2℃~3℃ 낮지만, 비교적 노면온도변화 패턴을 외기온도가 잘 반영하고 있다. 그러나 2016년의 경우, 노면온도변화 패턴과 상이한 외기온도변화 패턴이 나타나며, 노면온도와 차이를 수집날짜에 따라 다른 것을 알 수 있다. 즉, 2015년과 2016년의 데이터를 통합한 결과 학습데이터의 절대량은 증가했지만 특정

노면온도(15℃~20℃ 사이)를 추정하기 위해 사용한 학습데이터간 이질성이 크기 때문에 오히려 추정 성능이 저하된 것으로 판단된다.

결론적으로 주어진 모형으로부터 일정한 추정 성능을 보장하기 위해서는 학습데이터의 양 자체도 중요하지만 온도스펙트럼을 설명할 수 있는 양질의 데이터가 필수적일 것으로 판단된다.

4.6. 분산도에 따른 노면온도추정 성능분석

수집된 노면온도변화폭은 대체적으로 크지 않지만 일부 짧은 구간 내에서 온도변화폭이 다소 큰 경우(예: 2016년 10월 12일)가 나타난다. 실제 이러한 온도변화는 외부환경(터널 또는 교량 등)에 의한 것으로 추정할 수 있다. 이에 노면온도변화 패턴 추정 모형이 이러한 특징들을 적절히 반영할 수 있는지 평가할 필요가 있다. 급격한 노면온도변화 패턴에 대한 모형의 설명력을 정량적으로 평가하기 위해 모형 IV를 대상으로 분산도가 큰 노면온도표본에서 노면온도추정치와 유사도(AAE 기반)를 추가적으로 분석하였다. 분석을 위한 단계는 다음과 같다.

- 단계 1) 5개의 노면온도데이터에서 표본데이터를 추출
- 단계 2) 해당 표본에 대해서 분산(variance)을 계산
- 단계 3) 분산도가 일정 수준 이상인 표본이라면 노면추정온도와 전·후방외기온도에 대해서 각각 AAE를 계산
- 단계 4) 단계1~3을 거친 다음 표본데이터에 인접한 또 다른 5개의 표본데이터에 대해서 반복

4단계를 거치는 이유는 표본 내에서 노면온도의 분산 정도가 큰 경우(즉, 온도가 증가 또는 감소추세에 있거나 변화의 폭이 큰 경우)에 해당 노면추정온도의 AAE를 통해 유사도를 검증하고, 전·후방외기온도의 AAE와 비교하기 위함이다. Fig. 13은 각 표본으로부터 계산된 분산에 대한 결과를 나타낸다. 본 분석에서는 Fig. 13을 근거로 노면온도변화의 폭이 큰 표본의 선정 기준을 0.5℃²로 정의하였다. 몇 가지 대표표본을 살펴본 결과, 분산도가 0.5℃² 이상인 표본데이터들은 온도가 급격히 낮아지거나 높아지는 패턴이 뚜렷한 반면, 0.5℃² 이하인 경우, 이러한 패턴이 거의 나타나지 않는 것을 볼 수 있었다. 따라서 경험적 추론에 의해 이 기준을 선정하였다.

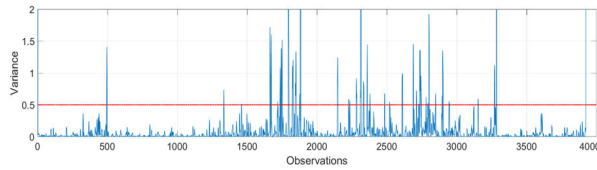


Fig. 13 Variation of Road Surface Temperature from Sample

Fig. 13에서 분산이 0.5°C^2 이상인 표본을 대상으로 추정된 노면온도변화 패턴과 실제노면온도변화 패턴과의 AAE를 계산하고 전·후방외기온도를 대상으로 같은 계산을 반복한 결과를 Fig. 14와 같이 시각화하였다. 여기서 “○”은 추정노면온도 기반 AAE가 외기온도 기반 AAE보다 낮은 표본이며 “■”은 반대로 외기온도 기반 AAE가 높을 때를 말한다. 즉, 실제 노면온도 변화폭이 큰 표본에서 추정노면온도가 외기온도보다 그 변화를 잘 반영한다면 ○의 개수가 ■의 개수보다 많을 것으로 기대할 수 있다.

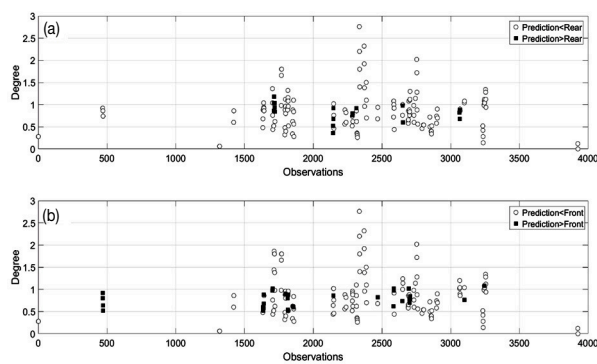


Fig. 14 Comparison AAE at highest Degree of Variance from Sample Data

Fig. 14(a)에서 후방외기온도와 추정노면온도를 비교해보면 ○의 개수는 145개인 반면 ■의 개수는 19개로 126개의 차이를 보이고 있다. Fig. 14(b)의 경우 ○의 개수는 129개인 반면 ■의 개수는 35개로 96개의 차이를 나타내고 있다. 즉, 전반적으로 분산도가 높은 노면온도변화 패턴을 보일 때 모형으로부터 노면온도변화 패턴을 추정하는 것이 외기온도를 직접적으로 사용하는 것보다 적합함을 알 수 있다. 또한 외기온도 중에서도 전방 외기온도가 후방 외기온도에 비해 노면온도변화 패턴을 더 잘 반영함을 간접적으로 확인하였다.

2015년과 2016년에 공통적으로 수집된 노면온도와 외기온도 데이터만을 통합 활용하여 모형 IV를 구축하였다. 그 결과 모형 IV의 성능은 다른 3개의 모형에 비해 추정 성능이 다소 떨어지는 것으로 분석되었다.

2015년과 2016년의 데이터의 통합은 궁극적으로 학습 데이터의 양은 증가했지만 특정 노면온도를 추정하기 위해 학습데이터로 사용되는 외기온도의 편차가 커 모형의 추정 성능은 오히려 저하되었다. 다양한 환경에서 수집된 데이터가 통합될 시 개별 환경에 특수성을 적절하게 반영하기 위해서는 환경변수를 고려할 필요가 있어서 추가적으로 실제 노면온도 변화가 큰 상황에서 모형의 성능을 평가해 보았다.

실제 노면온도변화가 큰 상황에서도 추정된 노면온도가 외기온도보다 민감하게 변화함을 정량적으로 분석하였다.

5. 결론

본 연구에서는 노면 결빙에 대해 선제적으로 대응하고 궁극적으로 동절기 교통사고를 줄이기 위한 전략 수립의 기초자료가 될 수 있는 노면온도변화 패턴을 추정하기 위한 방법론과 모형을 제시하였다. 노면온도변화 패턴을 추정하기 위하여 외기온도를 주된 예측 변수로 사용하였으며 습도와 대기온도 역시 추가적인 예측 변수로 분석에 활용하였다. 반응변수인 노면온도는 Thermal Mapping System을 이용하여 취득하였다. 노면온도변화 패턴 추정 모델링을 위해서 MATLAB에서 제공하는 기계학습 toolbox인 Classification Learner를 사용하였다. MATLAB에서 제공되는 기계학습 알고리즘은 약 20여개 이며, 본 연구에서는 이러한 알고리즘을 모두 실행 후 가장 정확도가 높은 weighted KNN을 선택하였다. 노면온도 수집데이터는 모형을 구축하기 위한 데이터와 모형을 평가하기 위한 데이터로 구분한 뒤, 전자에 상응하는 예측변수와 반응변수를 기계학습에 사용하여 모형을 구축하고 후자에 상응하는 예측변수와 반응변수를 이용하여 모형을 평가

Table 2. Define Variables for Each Model

Variables	Model I	Model II	Model III	Model IV
Response	Surface Temp.	Surface Temp.	Surface Temp.	Surface Temp.
Predictor	Front Temp.	Front Temp.	Front Temp.	Front Temp.
	Rear Temp.	Rear Temp.	Rear Temp.	Rear Temp.
	-	Humidity	-	-
	-	Temp.	-	-
Data	2015	2015	2016	2015+2016

하였다. 모형구축과정에서 적합하지 않은 데이터는 (예: Outlier) 전처리 과정을 거쳐 모두 제거하였다. 본 연구를 통해 구축된 총 4개의 모형은 Table 2와 같다.

구축된 모형에 대한 전반적인 평가 결과는 다음과 같다.

- 구축된 4개의 모형을 이용하여 추정된 노면온도변화 패턴이 외기온도 자체를 그냥 활용하는 것 보다 실제 노면온도에 더 근접함을 확인함
- 2016년에 새롭게 수집된 습도와 대기온도를 예측변수로 추가하여 구축한 모형(모형 II)와 외기온도만을 사용한 모형(모형 III)을 비교한 결과 전자의 경우가 미미하게나마 설명력이 높은 것을 확인함
 - 즉, 외기온도 이외에 반응변수를 설명할 수 있는 추가적인 예측변수의 고려는 모형 성능 향상에 긍정적인 영향을 미칠 수 있음을 시사함
- 2015년과 2016년의 공통 예측변수인 전·후방외기온도를 예측변수로 구축된 모형(모형 IV)을 평가한 결과 다른 모형(모형 I, II, III)에 비해 미미한 추정 성능 저하를 볼 수 있으며 그 원인은 다음과 같음
 - 2015년과 2016년의 데이터를 통합한 결과 학습데이터의 절대량은 증가했지만 특정노면온도(15℃ ~ 20℃)를 추정하기 위해 사용한 학습데이터 간 이질성이 크기 때문에 오히려 추정 성능이 저하된 것으로 판단됨
- 유사한 노면온도에 상응하는 외기온도의 변화가 연도 별로 상이함

짧은 시간 안에 실제 노면온도의 유동폭이 큰 표본온도데이터를 대상으로 외기온도와 추정 노면온도(모형 IV를 이용)를 비교한 결과, 추정된 노면온도가 실제 노면온도 변화에 더 민감하게 변화함을 정량적으로 분석하였다.

향후 연구에서는 MATLAB에 내재되어 있는 Classification Learner를 이용한 노면온도변화 패턴 추정모형의 성능 향상을 위해서 변수의 레벨을 적절하게 집합화시킬 필요가 있다. 현재까지 노면온도 및 외기온도 단위는 소수점 첫째 자리까지 사용하였는데 활용 목적을 감안했을 때 0.5℃ 단위로 집합화시키는 것이 가능하다면 데이터 전처리과정을 통해 집합화시킨 후 모형을 재구축하여 사용하는 것이 필요할 것으로 판단된다.

또한, 물리적 변화(교량 및 터널) 또는 환경적 변화(구간 내 급격한 온도변화 또는 기상상태)에 따른 온도변화

에 민감한 모형 구축을 위해서는 다양한 환경변수의 고려가 필요하다. 특히 모형이 교량이나 터널과 같은 물리적 변화에 따른 온도변화를 추정하기 위해서는 “데이터 수집 위치”와 같은 dummy 예측변수를 모형에 추가할 필요가 있다.

결론적으로 모형 성능 향상을 위해서 가장 중요한 요소는 다양한 환경에서 수집된 양질의 학습데이터 양이다. 즉, 반응변수에 상응하는 예측변수의 조합이 다양하고 그 조합의 양이 많으면 많을수록 모형의 성능은 향상된다. 현재 수집된 반응변수와 예측변수의 조합을 살펴 보게 되면 대부분 13℃에서 18℃ 사이의 매우 좁은 온도 범위 내에서 온도의 변화가 비교적 안정적인 환경에서 수집된 데이터가 대부분을 차지하고 있다. 이러한 환경에서 구축된 모형의 경우 이질적인 환경의 노면온도 추정에 취약할 수밖에 없다. 따라서 궁극적으로 범용적 목적으로 활용할 모형을 구축하기 위해서는 제시한 방법론을 기반으로 다양한 환경 내에서 반복적으로 수집된 양질의 데이터 취득이 필수적일 것으로 판단된다.

2016년 12월부터 2017년 2월까지 수집한 동절기 노면온도 데이터를 통해 모형의 성능을 보다 향상하기 위한 노력도 현재 진행중이다.

감사의 글

본 연구는 한국건설기술연구원 주요사업인 “차량 센서를 이용한 도로노면온도 및 결빙 예측 기술 (3차년도)” 연구비 지원에 의해 수행되었습니다.

REFERENCES

Bose and Mahapatra (2001), Business Data Mining-a Machine Learning Perspective, Information and Management, 39, pp 211-225.

Black and Mote (2015), Effects of winter precipitation on automobile collisions, injuries, and fatalities in the United States. Journal of Transport Geography, Vol. 48, pp165~175.

Constantinos Antoniou and Haris Koutsopoulos (2006), Estimation of Traffic Dynamics Models with Machine-Learning Methods, Journal of the Transportation Research Record, Vol. 1965, pp 103-111.

Chen and Yuanchang Xie (2015), Machine Learning for Recognizing Driving Patterns of Drivers of Large Commercial Trucks, Journal of the Transportation Research Record, Vol. 2517, pp 18-27.

Jonsson (2011), Remote sensor for winter road surface status detection. Sensors, IEEE.

Jeon, W., J. Choi., K. Park., K. Kim (2016), Design of Traffic Congestion Predictive System with Machine Learning.

- Kari MaKela and Kimmo Erkkilla (2013), Slipperiness Detection System, VTT Technical Research Center of Finland.
- Liu, T., Wang, N., Yu, H., Basara, J., Hong, Y., and Bukkapatnam, S., (2014), Black ice detection and road closure control system for Oklahoma. Final Report, FHWA-OK-14-08.
- Omer and Liping (2010), An automatic image recognition system for winter road surface condition classification. Intelligent transportation systems (itsc), 13th international IEEE conference.
- Yang C, H., C.J, Yoon, J, Kim, J, Park, D, Yun (2016), Evaluation of Reliability for Road Surface Temperature Change Patterns and Road Segmentation based on Road Surface Temperature, International Journal of Highway Engineers, Vol. 18, No. 4, pp 103-111.
- IT Dong A Big Data and Traffic Prediction <http://it.donga.com/21145/>