

연구 proposal

2017. 4. 13. 강선아

과제명: 딥러닝을 이용하여 기후변화에 따른 전염성 질환 발생 패턴 분석

■ 연구의 필요성 및 목적

□ 지구온난화등 기후변화로 인한 이상기후 발생

- 기후변화(climate change)는 지구 규모의 기후 또는 지역적 기후의 시간에 따른 변화를 말하며 일반적으로 10년에서부터 수백만 년의 기간 동안의 대기의 평균적인 변화를 의미함.
- 기후변화로 인해 이상기후가 발생하고 있는데 특히 지구온난화는 심각한 사회 문제로 대두됨. 지난 30년 동안 지구의 온도는 약 0.6°C 상승하였으며, United Nations Intergovernmental Panel on Climate Change가 발표한 내용에 의하면 2099년 현재 추세가 유지된다면 4.0°C의 온도가 상승하고 그 결과 해수면이 9~88cm 상승할 것이라고 보고함.
- 국내에서도 2016년 대부분의 월 기온이 전년보다 1°C 이상 높았고, 연평균 기온 역시 크게 상승함.

□ 지구온난화로 인해 질병발생 등 직, 간접적으로 인간의 건강에 영향을 미칠 것으로 예상(Patz, J. A., Frumkin, H., Holloway, T., Vimont, D. J., & Haines, A. (2014). Climate change: challenges and opportunities for global health. *Jama*, 312(15), 1565-1580.)

- IPCC(Intergovernmental Panel on Climate Changes)는 기후 변화가 인간의 건강에 위협적인 요소가 될 것이라고 보고하였으며, 홍수, 폭풍과 같은 기상 재해를 통해 직접적인 건강 영향과 수인성 전염병, 수질, 대기질 등 간접적인 건강 영향이 있을 것으로 예상.
- 기후변화의 영향은 모든 인구집단에서 동일하게 나타날 것으로 예상되지 않고 있는데, 인구밀집도, 경제발전수준, 식량의 가용성, 소득수준, 국지적 환경조건, 질병유무, 공공의료의 수준 및 가용성에 따라 특히 더 위험한 집단이 존재할 가능성이 있음.
- 따라서 기후변화에 따른 질병발생 영향을 파악하기 위해서 지역별 기후조건 등을 고려할 필요가 있음.

□ 최근 딥러닝 기술은 새로운 사회현상을 예측하는 데 우수한 성능을 보임

- 딥러닝은 신경망의 발전된 형태로 단순한 신경망 모형에서 은닉층을 늘려 모형의 성능을 향상시킨 심층 신경망 모형임.
- 컴퓨팅 파워 향상과 데이터의 폭발적인 증가로 딥러닝의 성능이 비약적으로

발전하면서 사회현상을 예측하는데 딥러닝을 활용하는 사례가 증가하고 있음.
또한 딥러닝을 질병예측, 기후변화 예측 분야에 활용하려는 움직임이 일고 있음.

- 다수의 연구를 통해 비선형 패턴을 보이는 데이터의 경우 인공신경망 모형이 적합한 것으로 나타남이 증명되고 있으므로, 딥러닝 모델을 이용하여 질병의 패턴분석 및 발병에 대한 모델링을 하고자 함.

□ 본 연구에서는 딥러닝 기술을 이용하여 기후변화에 따른 전염성 질환의 지역별 발생 패턴을 추적하고 예측하고자 함

- 딥러닝 기술을 활용하여 기후 변화 정도, 기후의 극성값 등으로 야기되는 전염성 질환을 예측

■ 선행연구

1. Artificial neural networks for infectious diarrhea prediction using meteorological factors in shanghai

- 출처 : Wang, Y., Gu, J., & Zhou, Z. (2014). Artificial neural networks for infectious diarrhea prediction using meteorological factors in Shanghai. In SEKE (pp. 501-506).
- 배경 : 전염병은 사회, 경제, 인간의 유출입 등 다양한 요인에 의해 발생하며, 이러한 전염병의 매개체는 기후변화에 민감하게 반응함.
- 목적 : ECDC에서 출간된 연구 결과를 토대로 전염성에 대한 환경 및 역학 네트워크의 청사진을 개발함으로써 기후변화와 전염병의 관계에 대한 이해도를 높이고자 함.
- 내용 :

전염병 구분	매개체	기후변화와와의 관계
vector-borne disease	mosquito-borne disease	<ul style="list-style-type: none"> - 기후가 매개체의 지리적 서식지의 결정요인이긴 하나, 최근의 기후변화가 매개체 전염병 증가에 영향을 미친다고 결론을 내릴 수 없음. - 하지만 지난 십년동안의 기후변화는 진드기의 서식지를 높은 위도 쪽으로 확장시킴.
	sandfly-borne disease	
	tick-borne disease	
rodent-borne disease	plague	<ul style="list-style-type: none"> - 많은 강수량은 렙토스피라증과 같은 설치류 병균의 번식에 직, 간접적으로 영향을 줌
	hantavirus	
water-borne disease	recreational water use	<ul style="list-style-type: none"> - 노후화된 수질처리 및 배급 시스템은 기상이변에 취약하며, 식수공급에 심각한 취약성을 초래함
	drinking water	
air-borne disease		<ul style="list-style-type: none"> - 습도는 하부기도감염과 연관이 있음
food-borne disease	campylobacteriosis	<ul style="list-style-type: none"> - 기온과 살모넬라와 식중독 발생과의 관계에는 밀접한 영향을 미치는 것으로 나타남
	salmonellosis	

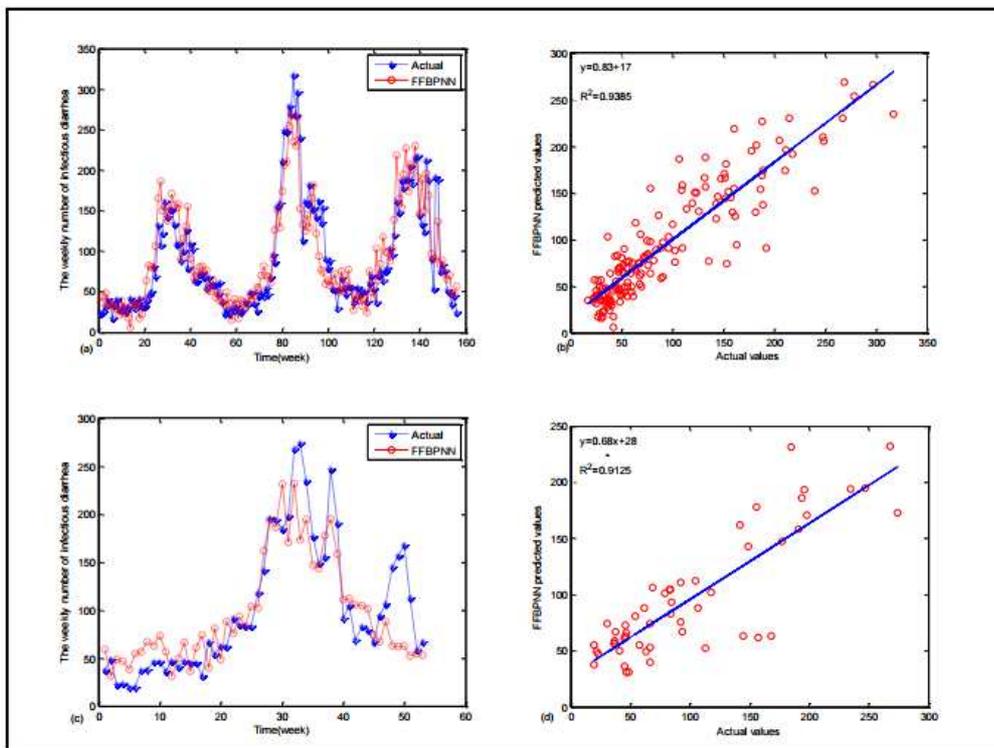
2. Artificial neural networks for infectious diarrhea prediction using meteorological factors in shanghai

- 출처 : Wang, Y., Gu, J., & Zhou, Z. (2014). Artificial neural networks for infectious diarrhea prediction using meteorological factors in Shanghai. In SEKE (pp. 501-506).
- 연구 방법론 :

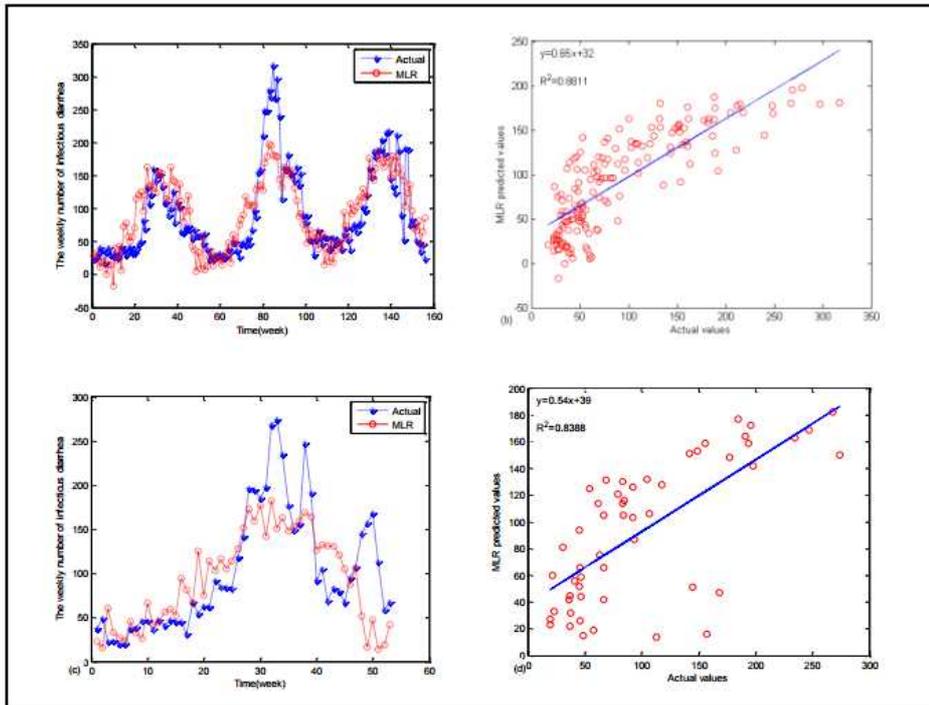
구분	내용
Type	실험 연구
데이터	<ul style="list-style-type: none"> - meteorological factors: 최고기온, 최저기온, 평균기온, 최저 상대습도, 평균 상대습도, 평균 기압, 일조량, 평균 풍속, 강수량 - infectious diarrhea: 주별 질병 건수
데이터 출처	<ul style="list-style-type: none"> - National disease supervision information management system - shanghai meteorological bureau of city environmental meteorological center
공간 단위 정보	상하이
시간 단위 정보	2005년 1월 3일~2009년 1월 4일(209주)
분석 방법 및 프로그램	<p>Matlab을 이용하여 FFBPNN(feed-forward back-propagation artificial neural network model) 구현</p> <ul style="list-style-type: none"> - 은닉층 : 4개 - learning rate: 0.74 - momentum rate: 0.9 - 1500 epoch
예측율 정확도	<p>FFBPNN model의 성능을 검증하기 위해 multivariate regression model과 성능을 비교</p> <ul style="list-style-type: none"> - multivariate regression model: $WNID = -1972.7903 - 10.9619T_{\max} + 20.8158T_{\min} - 2.6208T_{\text{avg}} - 1.6506RH_{\min} + 0.2993RH_{\text{avg}} + 2.0902AP_{\text{avg}} + 5.7734SD - 15.7205WS_{\text{avg}} + 1.6048R$ - 예측모델 성능

PECs	Models			
	FFBPNN		MLR	
	Training	Testing	Training	Testing
MAE	20.7628	27.7547	29.8077	35.3774
RMSE	28.3007	36.0526	39.3739	48.9395
MAPE	0.2727	0.3841	0.4337	0.4182
R	0.8783	0.8490	0.8089	0.6968
R ²	0.9213	0.9125	0.8811	0.8388

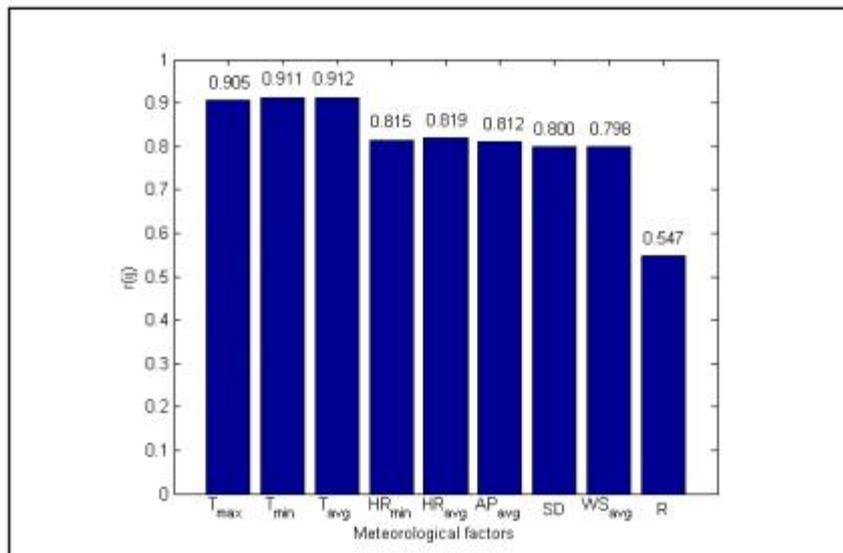
- 실제 데이터와 FFBPNN 예측 결과 값의 비교



- 실제 데이터와 MLR 예측 결과 값의 비교



- 종속변수와 독립변수간의 영향정도 분석(CAM; Cosine Amplitude Method)

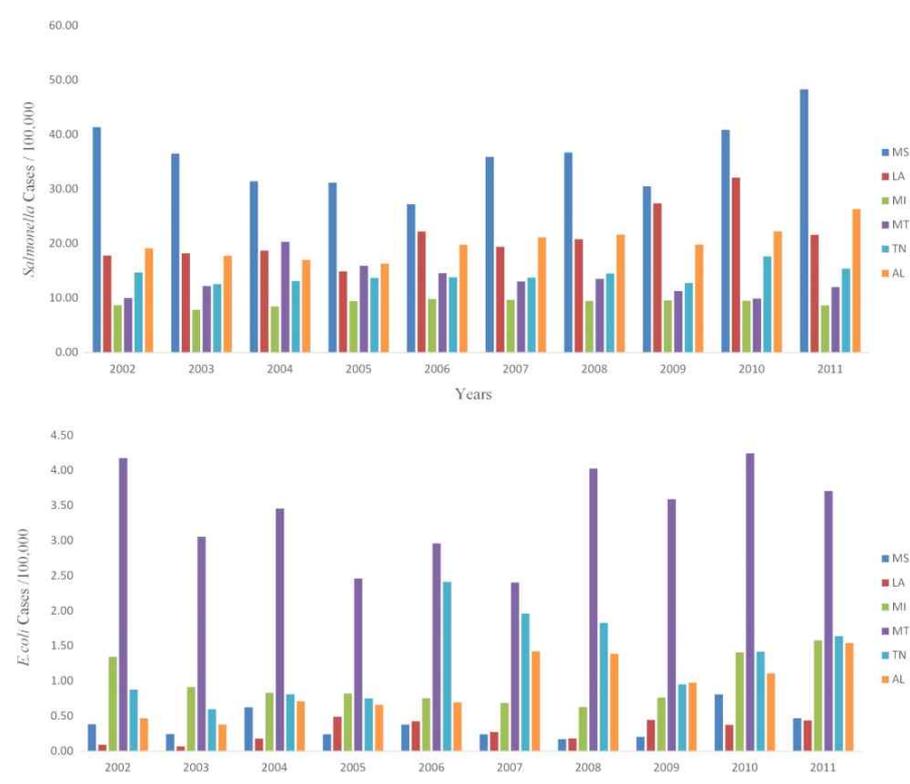


- 분석결과: FFBPNN이 다변량 분석보다 설사병 예측에 더 적합한 알고리즘으로 나타났다으며, 기온이 기후적 요소에서 가장 영향을 많이 미치는 변수인 것으로 드러남.

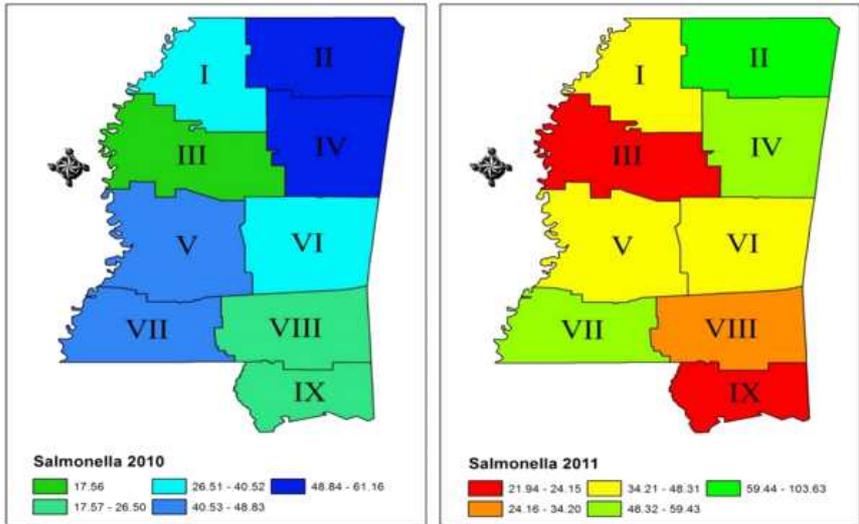
3. salmonella infections modelling in Mississippi using neural network and

geographical information system

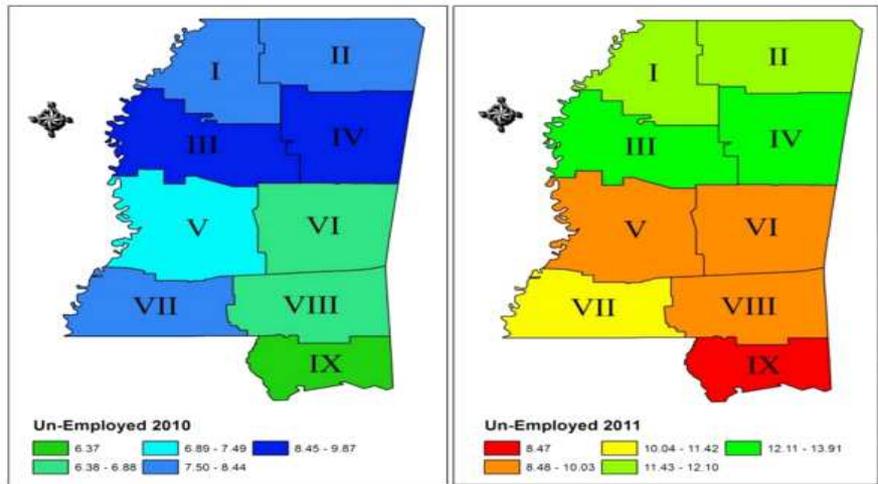
- 출처 : Akil, L., & Ahmad, H. A. (2016). Salmonella infections modelling in Mississippi using neural network and geographical information system (GIS). *BMJ open*, 6(3), e009255.
- 배경 및 목적 : 미시시피주는 식품매개질병(foodborne disease)의 발생이 높은 지역으로 salmonella와 escherichia coli 감염의 정도를 살펴보고, GIS와 신경망 모델을 이용하여 salmonella 감염과 사회경제적 지위 간의 상관관계를 분석
- 연구 방법론 :

구분	내용
Type	실험 연구
데이터	- 사회경제적 지위 데이터: 실업률, PCP(primary care provider; 일차 의료 제공자) - salmonella, E. coli: 월별 질병건수(2002~2012)
데이터 출처	- 기후 데이터: Centers for Disease Control and Prevention, MS state Department of Health and another states Department of Health.
공간 단위 정보	미시시피주를 비롯한 미국의 5개주 (Alabama, Tennessee, Louisiana, Montana, Michigan)
분석	- GIS 분석 

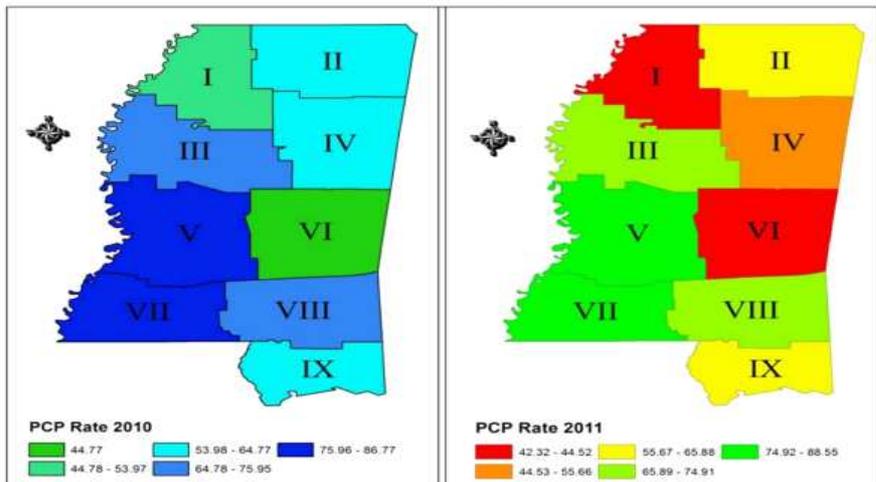
MS Salmonella Percent by Districts



MS Un-Employed Percent by Districts



MS PCP Percent by District



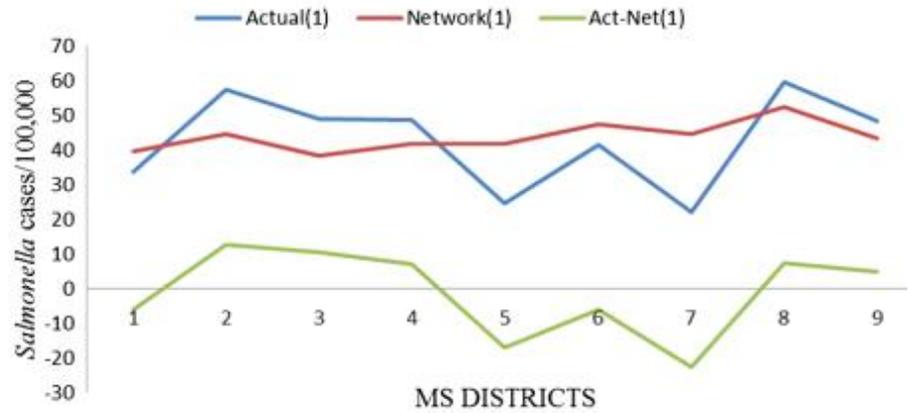
- NN 분석

Table 1 Results of neural network model

Output	GRNN
R^2	0.4169
r^2	0.4169
Mean squared error	175.872
Mean absolute error	11.545
Min. absolute error	4.381
Max. absolute error	22.137
Correlation coefficient r	0.6457

Results of GRNN model for *Salmonella* and socioeconomic factors.

GRNN, General Regression Neural Network.



- 분석결과: 연구는 사회경제적 지위 데이터와 salmonella 질병 발생 간에 유의미한 상관관계가 있음을 보여주었으며, LSES(lower socioeconomic status)와 salmonella 발생 간에 상관관계가 있음을 밝힘.

4. 국내 기후변화 관련 감염병과 기상요인간의 상관성

- 출처 : 김시현, & 장재연. (2010). 국내 기후변화 관련 감염병과 기상요인간의 상관성. 예방의학회지, 43(5), 436-444.
- 배경: 기상요인들이 감염병 발생에 영향을 주는 주요요인이며, 적절한 예방과 적응을 위해 기후변화와 감염병에 관한 연구의 필요성이 대두됨
- 목적: 기후변화와 관련성이 높다고 알려진 감염병들을 대상으로 기상요인과 상관성을 분석함으로써 국내에서도 감염병 발생에 기후변화가 영향을 주고 있거나 향후 줄 수 있는지를 평가해보고자 함
- 연구 방법론 :

구분	내용
Type	실험 연구
데이터	<ul style="list-style-type: none"> - 질병 데이터: 2001년부터 2008년까지 신고된 법정전염병 개인발병 자료 - 기후 데이터: 같은 기간 기상청과 기상대에서 측정한 자료(평균기온, 최고기온, 최저기온, 상대습도, 강수량) - 연도별, 지역별 인구자료
데이터 출처	<ul style="list-style-type: none"> - 질병 데이터: 질병관리본부 - 기후 데이터: 기상청 - 인구자료: 통계청
분석	<ul style="list-style-type: none"> - 성, 연령, 직업별 발생건수와 지역에 따른 발생률 분석 - 감염병 발생률은 연평균 인구 백만명당 누적 발생률로 조사기간동안 연동별 각 시·도에서 발생한 총 환자수를 같은 해 그 지역의 평균 인구수로 나눈 값으로 계산 - 국내의 연중 기후 양상과 상관성이 있는지 확인하기 위하여, 8년간 전국에서 발생한 월별 총 사례수와 각 월에 해당하는 기상요인들의 전국 평균값을 이용하여 상관분석

- 분석결과:

연구대상 감염병의 월별 발생건수를 표시한 것으로 쓰쯔가무시증, 신증후군출혈열, 렙토스피라증은 가을철에, 말라리아와 비브리오패혈증은 여름철에 주로 발생하는 경향을 보임 (계절성이 매우 강한 것으로 나타남)

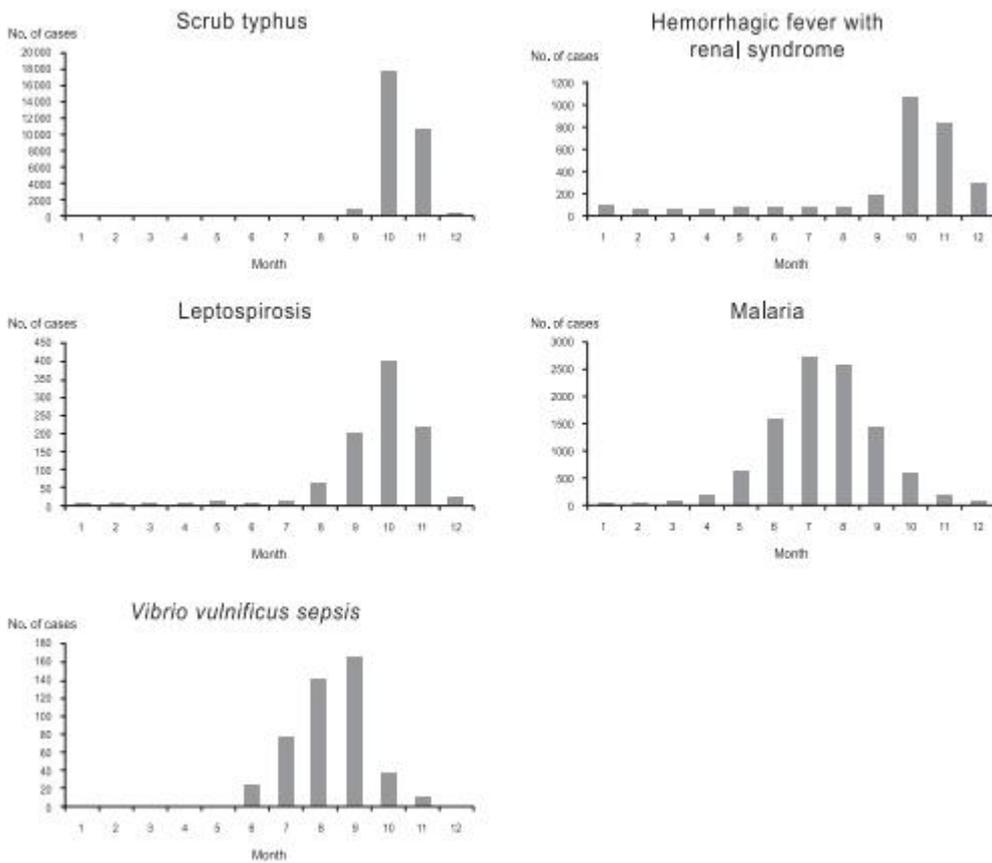


그림 1 연중 감염병의 계절성

말라리아 환자발생건수는 연중 기온, 습도, 강수량 변화 패턴과 양의 상관관계를 보이며, 비브리오패혈증 역시 유의미한 상관관계를 보임. 반면 쓰쯔가무시증, 신증후군출혈열, 렙토스피라증은 월별 환자발생건수와 동일한 시기의 기상요인과 유의한 상관관계를 보이지 않았음.

Diseases	Meteorological factors				
	Average temperature	Maximum temperature	Minimum temperature	Relative humidity	Precipitation
Scrub typhus	-0.024	0.010	-0.041	0.054	-0.306
HFRS	-0.091	-0.062	-0.101	0.040	-0.352
Leptospirosis	0.153	0.182	0.144	0.261	-0.174
Malaria	0.881*	0.846*	0.912*	0.942*	0.966*
<i>V. vulnificus</i> sepsis	0.697*	0.673*	0.725*	0.814*	0.656*

Each value is the Pearson's correlation coefficient between the number of cases and monthly mean value of meteorological factor.

HFRS: hemorrhagic fever with renal syndrome.

* $p < 0.05$.

그림 2 감염병과 기상요인과의 상관성

5. 기후변화와 전염병 질병 부담

- 출처 : 신호성, & 김동진. (2008). 기후변화와 전염병 질병 부담.
- 배경: 한반도의 기온이 높아지고 있으며, 이러한 기후변화는 인간의 건강에 영향을 미칠 가능성이 있음. 특히 기후의 변이성은 매개체에 의한 전염성 질환에 영향을 미칠 수 있으므로 기후 변화에 의한 전염병 발생영향을 분석할 필요성이 있음.
- 목적: 기후변화로 인한 전염병 발생양상 변화가 국민의 건강에 어떠한 영향을 미칠 것인지를 계량화하고, 이에 대비한 국가정책이 적절한지를 평가하고자 함.
- 연구 방법론 :

구분	내용
Type	실험 연구
데이터	<ul style="list-style-type: none"> - 질병 데이터: 2001년부터 2008년까지 신고된 법정전염병 개인발병 자료 - 기후 데이터: 같은 기간 기상청과 기상대에서 측정한 자료(평균기온, 최고기온, 최저기온, 상대습도, 강수량) - 연도별, 지역별 인구자료
데이터 출처	<ul style="list-style-type: none"> - 질병 데이터: 심평원 - 기후 데이터: 기상청
분석	<ul style="list-style-type: none"> - 기후요소의 영향- 3차원 스플라인을 적용하여 추정 $\log(\lambda_{ijt}) = \alpha + \sum_{k=1}^3 a_{0k} B_k(\text{Temp}_{ijt}) + \sum_{k=1}^3 b_{jk} B_k(\text{Temp}_{ijt}) + \text{rain}_{ijt} + \sin(x_i) + \cos(x_i) + L1 + L2 + L3 + LA + \text{year}$ <ul style="list-style-type: none"> - 전염병 발생 건수 $W_d = L1 + 2^d L2 + 3^d L3 + 4^d LA$

- 분석결과:

건강보험심사평가원으로부터 제공받은 2005~2007년 건강보험심사평가원 전산청구자료를 분석한 결과, 기후변화 관련 전염병 건수는 총 2,592,515건이었으며, 이중 본 연구의 대상이 되는 렙토스피라, 말라리아, 세균성이질, 장염 비브리오, 쯤쯤가무시의 수진건수는 2005~2007년 기간 동안 총 61,354건으로 나타났다.

질병별로는 쯤쯤가무시가 가장 많은 32,051건(52.2%)였고, 다음으로 말라리아 16,898건(27.5%), 세균성이질 8,975건(14.6%), 렙토스피라 1,756건(2.9%), 장염비브리오 1,674건(2.7%)의 순서로 분포하고 있었음. 이를 인구 10만명당 유행률로 환산하면, 쯤쯤가무시 22.1명, 말라리아 11.7명, 세균성이질 6.2명, 렙토스피라 1.2명, 장염비브리오 1.2명임.

질병별로는 말라리아의 유병률이 해마다 증가하고 있었는데 특히 남성집단에서의 유병률 증가를 확인할 수 있었고, 반대로 쓰쓰가무시의 경우 유병률이 낮아지고 있었음. 그러나 지난 3년간의 분석임으로 결과해석에 주의를 요함.

2005~2007년 3년 동안의 전염병 발생을 기준으로 온도변화에 따른 전염병 발생을 예측한 결과 우리나라의 온도가 섭씨 1도 상승할 경우 5가지 전염병의 평균 발생률은 4.27% 증가할 것으로 예측되었음.

5가지 전염병 중 쓰쓰가무시의 발생이 가장 높을 것으로 예측되었고(5.98%), 세균성이질(1.81%)의 경우가 가장 적은 영향을 받을 것으로 예측되었음. 기온 상승에 따른 전염병 발생 영향은 [쓰쓰가무시, 렙토스피라, 말라리아, 장염비브리오, 세균성이질의 순으로 나타났음.](#)

향후 인구구조의 변화를 감안하여 연령대별로 기온상승에 따른 질병부담을 산출한 결과, 온도가 상승함에 따라 질병부담도 높아졌음. 특히 35~64세, 65세 이상의 질병부담이 상대적으로 감소하고 저연령층의 질병부담이 높아지는 현상이 나타났음. 이는 인구구조의 변동에 의한 것으로 연령대별 질병발생 빈도에 기인한 것이 아니라 인구피라미드가 역삼각형 모양으로 변화된 것 때문임.

연령대별 질병부담을 모두 합한 값으로 전염병 종류별 질병부담을 측정하면 쓰쓰가무시(5.98%), 렙토스피라(4.07%), 말라리아(3.40%), 장염비브리오(3.29%), 세균성이질(1.81%)의 순으로 나타났음. 현재의 인구구조를 그대로 적용한 모형에 비하여 말라리아, 세균성이질, 장염비브리오의 경우 인구구조 변동을 반영한 모형에서 상대적으로 질병부담이 증가하며 쓰쓰가무시의 경우 그 영향이 줄어드는 것으로 나타났음.

산출된 전염병 질병부담은 세계보건기구가 2008년에 발표한 질병별 질병부담(YLD)의 EME(Established Market Economics) 국가군과 비슷한 양상을 보였음.

■ 주요 연구내용

- 질병에 영향을 주는 요인 탐색 및 모형 변수 선정
 - 질병에 영향을 주는 기후요소의 도출 및 변수 선정
- 딥러닝을 이용한 질병예측알고리즘을 개발을 통해 기후변화에 따른 질병예측
 - 시/도 별 질병 발생과 변화 패턴파악
 - 질병에 영향을 미치는 기후데이터를 이용하여 시/도 별 질병 예측

■ 연구추진방법

- 관련 문헌 분석
 - 기후변화와 인간의 건강, 질병 분석과 관련된 자료 분석
 - 딥러닝 방법론을 적용하여 기후 및 질병을 분석한 자료 분석
- 데이터 수집 및 질병 감지 알고리즘 개발
 - 분석에 필요한 기후, 질병 데이터 수집
 - 기상청 기상자료개방포털에서 제공하는 기후 데이터 수집
 - 국민건강보험공단에서 제공하는 표본코호트DB 데이터 수집
 - 기후변화에 따른 질병예측을 위한 알고리즘 개발
 - GAM(Generalized Additive Models)를 이용하여 질병과 기후간의 지역별 패턴파악
 - 딥러닝의 다양한 방법론 중 RNN을 이용하여 분석 모델 개발

■ 데이터 수집 관련

□ 기후데이터

○ 기상자료개방포털

<https://data.kma.go.kr/data/grnd/selectAsosRltmList.do?pgmNo=36>

기상자료개방포털 데이터 기후통계분석 카탈로그 소통과 참여

데이터

기상관측 -

- 지상
 - 종관기상관측 >
 - 방재기상관측
 - 농업기상관측
 - 북한기상관측
- 고층
 - 레이즌데
 - 연직바람관측
- 해양
 - 해양기상부이
 - 등표기상관측
 - 파고부이
- 항공
 - 공항기상관측
- 활사
 - 부유분진측정

레이더 +

위성 +

수치예보 +

날씨예보 +

기후예보 +

자료조회-종관기상관측 Home > 데이터 > 기상관측 > 지상 > 종관기상관측 > 자료조회-종관기상관측

파일셋 조회 **자료조회** 자료설명

종관기상관측장비(ASOS)로부터 관측된 자료를 DB에서 조회하고 CSV파일로 다운로드 받으실 수 있습니다.
* 분자로 1일, 시간으로 1년, 일자로 10년까지 조회 및 다운로드 가능합니다.

자료조회

자료형태: 시간 자료 기간: 20170401 01 ~ 20170410 00

지점: 지점 명으로 선택 지도로 선택 요소: 선택

관악산, 서울, 부산, 대구, 대구(기), 강화, 백령도, 인천, 광주, 대전, 울산, 동두천, 수원, 양평, 이천, 파주, 강릉, 대관령, 동해, 북강릉, 북초전, 삼척, 속초, 영월, 원주, 인제, 경

기온, 강수량, 풍향, 풍속, 이슬점온도, 증기압, 습도, 일사, 일조, 3시간신적설, 적설

조회 다운로드

전체 20298건 10

지점	시간	기온(°C)	강수량(mm)	풍속(m/s)	풍향(16방향)	습도(%)	증기압(hPa)	이슬점온도
90	2017-04-02 17:00	11.7		4.7	160	63	8.6	4.8
90	2017-04-02 18:00	11.5		2	180	62	8.4	4.4
90	2017-04-02 19:00	10		0.2	0	71	8.7	4.9
90	2017-04-02 20:00	13.1		1.3	270	32	4.8	-3.2
90	2017-04-02 21:00	13.4		1.3	340	31	4.8	-3.4
90	2017-04-02 22:00	13.5		2.2	270	31	4.8	-3.3

- 지역: 총 101곳 - 서울(2), 부산광역시(1), 대구광역시(2), 인천광역시(3), 광주광역시(1), 대전광역시(1), 울산광역시(1), 경기도(5), 강원도(15), 충청북도(5), 충청남도(5), 전라북도(10), 전라남도(16), 경상북도(14), 경상남도(14), 제주도(6)
- 기상데이터: 기온(평균기온, 최저기온, 최고기온), 강수량(일강수량, 1시간 최대 강수량, 강수계속시간), 풍속(최대 순간 풍속, 최대 순간 풍속 풍향, 최대 순간 풍속 시간, 최대 풍속, 평균풍속), 습도(평균 이슬점 온도, 최소 상대습도, 평균 상대습도, 평균 현지기압, 최고 해면기압, 최저 해면기압, 평균 해면기압), 일사량(1시간 최대 일사량, 합계 일사량), 일조량(합계 일조시간), 적설 등
- 자료제공: 데이터는 csv 파일 형태로 제공하며, 지점, 일시, 요청데이터로 구성

□ 건강보험데이터

○ 국내 감염병(전염병) 발생현황과 예측변수 선정

표 1. 특정 감염성 및 기생충성질환 발생현황(국민건강보험공단 자료)

항목	2006	2007	2008	2009	2010	2011	2012	2013
특정 감염성 및 기생충성 질환(A00-B99)	305,114	400,872	420,755	414,097	449,744	495,914	530,560	540,261

- 질병선정(한국표준질병·사인 기준)

표 2. 진료 DB 레이아웃

변수명	비고
개인일련번호	
청구일련번호	
요양기관식별 대체번호	
요양개시일자	
서식코드	
진료과목코드	
주상병	주된 상병분류기호 외의 추가 상병분류기호 한국표준질병사인분류코드 참고
부상병	주된 상병분류기호 외의 추가 상병분류기호 한국표준질병사인분류코드 참고

▶ '91~'00 년간 법정전염병 환자발생 추이(출처:IPCC 3차보고서 ('01) 등)

구분	증가추세	감소추세	불확실
기후변화와의 관련성	<ul style="list-style-type: none"> - 쯔쯔가무시증 - 말라리아 - 세균성이질 - 신증후군출혈열 - 렙토스피라증 - 발진열 - 덩기열 - 리슈마니아증 - 비브리오패혈증 	-	<ul style="list-style-type: none"> - 콜레라(2001년 이후 발생건수 급증) - 일본뇌염(2002년 발생건수 증가) - 페스트(발생사례없는 질병)
	<ul style="list-style-type: none"> - 유행성이하선염 	<ul style="list-style-type: none"> - 장티푸스 - 백일해 - 파상풍 - 결핵 - 한센병 - 성홍열 - 공수병 - 장출혈성대장균 - 풍진 - 레지오넬라증 - 브루셀라증 	<ul style="list-style-type: none"> - 홍역(2001년 이후 발생건수 급증) - 수막구균성수막염(2001년 이후 발생건수 급증) - 발생사례없는 질병 <ul style="list-style-type: none"> - 디프테리아 - 폴리오 - 발진티푸스 - 탄저

(* 출처 : 기후변화홍보포털)

제1군	2015	2014	2013
		2,128	1,816
콜레라	0	0	3
장티푸스	121	251	156
파라티푸스	44	37	54
세균성이질(시겔라증에 포함)	88	110	294
장출혈성대장균감염증	71	111	61
A형간염	1,804	1,307	867

제2군	2015	2014	2013
		73,957	74,477
디프테리아	0	0	0
백일해	205	88	36
파상풍	22	23	22
홍역	7	442	107
유행성이하선염	23,448	25,286	17,024
풍진	11	11	18
폴리오	0	0	0
B형간염(급성)	155	173	117
B형간염(산모)	3,468	3,912	3,211
B형간염(주산기)	43	30	59
일본뇌염	40	26	14
수두	40	26	14
B형헤모필루스인플루엔자	0	0	0
폐렴구균	228	36	-

제3군	2015	2014	2013
		52,050	52,128
말라리아	669	638	445
성홍열	7,002	5,809	3,678
수막구균성수막염	6	5	6
레지오넬라증	45	30	21
비브리오패혈증	37	61	56
발진티푸스	0	0	0
발진열	15	9	19
쯔쯔가무시증	9,513	8,130	10,365
렙토스피라증	104	58	50
브루셀라증	5	8	16
탄저	0	0	0
공수병	0	0	0
신증후군출혈열	384	344	527
매독(1기)	720	726	566
매독(2기)	253	258	210
매독(선천성)	33	31	22
크로이츠펠트-야콥병	33	65	34
결핵	32,181	34,869	36,089

한센병	2	6	7
후천성면역결핍증	1,018	1,081	1,013

제4군	2015	2014	2013
	561	245	314
페스트	0	0	0
황열	0	0	0
뎅기열	255	165	252
바이러스성출혈열	0	0	0
두창	0	0	0
보툴리눔독소증	0	1	0
중증급성호흡기증후군	0	0	0
동물인플루엔자 인체감염증	0	0	0
신종인플루엔자	0	0	0
야도병	0	0	0
큐열	27	8	11
웨스트나일열	0	0	0
신종감염병증후군	0	0	0
라임병	9	13	11
진드기매개뇌염	0	0	0
유비저	4	2	2
치쿤구니아열	2	1	2
중증열성혈소판감소증후군	79	55	36
중동호흡기증후군	185	-	-
리슈마니아증	-	-	-
바베시아증	-	-	-
크립토스포리디움증	-	-	-
주혈흡충증	-	-	-

출처: 보건복지부(2016), 법정감염병발생보고

- ▷ 제 1군: 마시는 물 또는 식품을 매개로 발생하고 집단 발생의 우려가 커서 발생 또는 유행 즉시 방역대책을 수립하여야 하는 감염병
- ▷ 제 2군: 예방접종을 통해 예방 및 관리가 가능하여 국가예방접종사업의 대상이 되는 감염병
- ▷ 제 3군: 간헐적으로 유행할 가능성이 있어 계속 그 발생을 감시하고 방역대책의 수립이 필요한 감염병
- ▷ 제 4군: 국내에서 새롭게 발생하였거나 발생할 우려가 있는 감염병, 해외 유행 감염
 - 국민건강보험공단 코호트 DB 세부내역
 - 기준: 2002년 자격 대상자(약 100만명)
 - 연도: 2002년~2013년(12개년)
 - 비용: 모든 데이터 요청 시 학술연구 목적이면 50% 할인 1,055,000원
모든 데이터 요청 시 정책연구 목적이면 80% 할인 422,000원

■ 진행상황

- 구글에서 제공하는 tensorflow를 python에 import
- 기상 데이터 수집 중

■ 연구계획

구분	3월	4월	5월	6월	7월	8월	9월	10월	11월
◦ 선행연구 분석 및 데이터 변수 선정									
◦ tensorflow설치 및 데이터 수집/학습									
◦ 기후데이터와 질병 간의 상관관계 분석 ◦ 기후와 질병 데이터 간의 사·도별 패턴분석(GAM 이용)									
◦ 질병 예측을 위한 RNN 알고리즘 구축 및 예측									
◦ 보고서 작성									수정