

Weather Impact on Heat-Related Illness in a Tropical City State, Singapore

Hai-Yan Xu¹, Xiuju Fu¹, Chin Leong Lim², Stefan Ma³, Tian Kuay Lim⁴, Paul Anantharajah Tambyah⁵, Mohd Salahuddin Habibullah¹, Gary Kee Khoon Lee⁶, Lee Ching Ng⁷, Kee Tai Goh⁸, Rick Siow Mong Goh¹, Lionel Kim Hock Lee²

¹Institute of High Performance Computing, Singapore

²Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore

³Ministry of Health, Singapore

⁴Urban Redevelopment Authority, Singapore

⁵National University of Hospital, Singapore

⁶Rolls-Royce Singapore Pte Ltd., Singapore

⁷National Environment Agency, Singapore

⁸Saw Swee Hock School of Public Health, National University of Singapore, Singapore

Email: *fuxj@ihpc.a-star.edu.sg, *xuh@ihpc.a-star.edu.sg

How to cite this paper: Xu, H.-Y., Fu, X.J., Lim, C.L., Ma, S., Lim, T.K., Tambyah, P.A., Habibullah, M.S., Lee, G.K.K., Ng, L.C., Goh, K.T., Goh, R.S.M. and Lee, L.K.H. (2018) Weather Impact on Heat-Related Illness in a Tropical City State, Singapore. *Atmospheric and Climate Sciences*, 8, 97-110.

<https://doi.org/10.4236/acs.2018.81007>

Received: September 10, 2017

Accepted: January 21, 2018

Published: January 24, 2018

Copyright © 2018 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

In this article we propose a novel hurdle negative binomial (HNB) regression combined with a distributed lag nonlinear model (DLNM) to model weather factors' impact on heat related illness (HRI) in Singapore. AIC criterion is adopted to help select proper combination of weather variables and check their lagged effect as well as nonlinear effect. The process of model selection and validation is demonstrated. It is observed that the predicted occurrence rate is close to the observed one. The proposed combined model can be used to predict HRI cases for mitigating HRI occurrences and provide inputs for related public health policy considering climate change impact.

Keywords

Distributed Lag Nonlinear Model, Heat-Related Illness, Hurdle Model, Negative Binomial distribution, Weather Factors

1. Introduction

Heat related illness (HRI) is a well-known hazard for workers under high temperature environments and people exercising in extreme heat. The understanding of the association of weather factors on HRI is important for making preventive measures and related policy in workplace and schools, or during sports ac-

tivities.

Several modeling studies had investigated the association between weather conditions and HRI [1] [2] [3]. These studies were based mainly on summer data from temperate countries. An additive linear model was applied to investigate the association between meteorological variables and the logarithm of HRI rate in these studies. Temperature was found to have adverse effect on the occurrence of HRI and there is no consensus on the impact of other meteorological variables.

Previous studies on HRI in Singapore have been documented [4] [5] [6] [7] [8]. Preventive measures have been implemented in Singapore to manage heat injuries, especially in the Singapore Armed Forces [9] [10]. The annual incidence of HRI cases admitted to hospital had decreased from over 200 in the early 1990's to fewer than 100 in the 2000's (see **Figure 1**). However, the incidence has been increasing since 2008, with 55, 79, 88 and 106 reported HRI cases admitted to hospital in 2007, 2008, 2009, and 2010, respectively. This reversal in HRI trends requires a thorough understanding of the factors impacting HRI in Singapore.

Factors that impact HRI may include weather, social activity, and individual health status. In our study, we focused on modeling weather impact on HRI. Other factors were also considered by categorized parameters like holidays and weekdays.

For modeling the impact of weather on HRI in Singapore, we employed hurdle negative binomial (HNB) models together with distributed lag non-linear models (DLNMs). As many zero counts were found in our HRI data, a hurdle was assumed when the count of HRI cases turns from zero to non-zero. For the HNB model, a censored Poisson (right-censoring at 0) governed the binary outcome of the impact of weather on the occurrence and non-occurrence of HRI cases. If the counts of HRI cases were not zero, the conditional distribution of

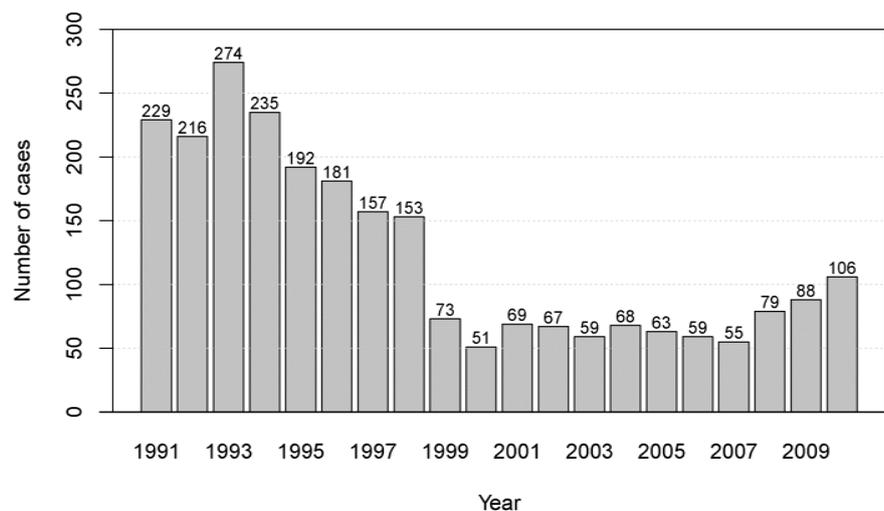


Figure 1. Annual counts of hospital admissions for heat related illness in Singapore, 1991-2020.

the counts was governed by a zero truncated negative binomial model [11] [12] [13].

The DLNMs were developed and had been used, simultaneously, to estimate the nonlinear and delayed effects of temperature (or air pollution) on mortality [14] [15] [16]. In this study, we used the DLNMs to investigate delayed and non-linear effect of weather variables on the daily incidence of hospital admission for HRI cases.

2. Materials and Methods

2.1. Data Description

Singapore is a city-state in Southeast Asia and located at 01°22'N and 103°48'E. Because of its geographical location and maritime exposure, the climate of Singapore is tropical with daily ambient temperature ranging from 25°C to 33°C and daily humidity between 60% and 95%. There are two main monsoon seasons, the northeast monsoon (December-early March) and the southwest monsoon season (June-September), separated by two relatively short inter-monsoon periods (late March-May and October-November).

The HRI data, provided by the Singapore Ministry of Health, recorded daily incidence of hospital admission for HRI cases (categorized as 992.0, 992.3, 992.5 under ICD-9 CM codes) from 1 January 1991 to 31 December 2010 in Singapore. The summary of the data is listed in **Table 1**. Daily Singapore weather data, including maximum temperature (MaxT), mean temperature (MeanT), minimum temperature (MinT), relative humidity (RH) and wind speed (Wdsp), were provided by the Singapore National Environment Agency for the years 1991 to 2010. The daily weather conditions and daily incidence of hospital admission for heat related illness cases are shown in **Table 2**.

Table 1. Summary of the daily incidence of hospital admissions for heat-related illness cases by each day in a week, and public holiday, Singapore, 1991-2010.

	Number of daily HRI												Proportion of days with HRI (%)
	0	1	2	3	4	5	6	7	8	9	12	15	
Counts of days corresponding to the number of illness cases													
Overall	5626	1222	289	88	40	23	5	5	2	3	1	1	22.98
Public holiday	236	16	3	–	–	–	–	–	–	–	–	–	7.46
Sunday	874	121	21	8	6	5	2	1	1	3	1		16.20
Monday	830	142	47	16	6	1	1	–	–	–	–	–	20.42
Tuesday	785	192	46	14	3	3	–	1	–	–	–	–	24.81
Wednesday	792	193	37	15	3	4	–	–	–	–	–	–	24.14
Thursday	769	194	52	12	8	5	2	2	–	–	–	–	26.34
Friday	767	202	46	15	8	5	–		1	–	–	–	26.53
Saturday	809	178	40	8	6	–	–	1	–	–	–	1	22.44

Table 2. Summary statistics of daily weather conditions and daily incidence of hospital admission for heat related illness cases in Singapore, 1991-2010.

	MinT	MeanT	MaxT	RH	Wdsp	WBGT_MeanT	HRI
Min.	20.2	23.1	23.6	66.0	0.1	26.1	0
1st Qu.	24.0	26.9	30.8	79.8	1.1	28.8	0
Median	24.8	27.7	31.8	83.1	1.7	29.5	0
3rd Qu.	25.7	28.6	32.6	86.8	2.5	30.1	0
Max.	29.1	30.9	36.0	99.5	5.9	32.3	15
Mean (sd)	24.9 (1.2)	27.7 (1.2)	31.6 (1.6)	83.4 (5.0)	1.9 (1.0)	29.4 (0.9)	0.34 (0.80)

1st Qu. and 3rd Qu. represent the 25th and 75th percentiles, respectively.

Wet-bulb globe temperature index (WBGT) is a common index for measuring environmental heat stress and is used widely as an assessment for the risks of HRI [17] [18]. In this study, no actual WBGT data was available, and theoretical WBGT data was only available for the years 2000-2001 and 2003-2006. We used in our model approximated WBGT, which was calculated based on ambient temperature and relative humidity. This is an indicator to reflect the compound effect of weather variables on HRI. The approximated WBGT used in this study, followed that used in Australia [19], which was successfully employed during the 2000 summer Olympics in Sydney, Australia [20], and was found to “most reliably” predict environmental heat stress [21]. High correlation (the Pearson’s correlation coefficient was 0.93) was found between theoretical WBGT and approximated WBGT based on available data in the years including 2000-2001 and 2003-2006. The expression of the approximated WBGT is shown as the following,

$$\text{WBGT} \approx 0.567 \times Ta + 0.393 \times e + 3.94 \quad (1)$$

where e is environmental water vapor pressure, calculated from temperature and relative humidity using $e = (RH/100) \times 6.105 \times \exp[17.27 \times Ta / (273.7 + Ta)]$, RH is the relative humidity and Ta is dry bulb temperature ($^{\circ}\text{C}$). In this study, Ta was the daily mean temperature, and the corresponding approximated WBGT based on Formula (1) was denoted WBGT_MeanT. Formula (1) was first proposed by [22], and subsequently it was recommended by the American College of Sports Medicine [23].

2.2. Data Analysis

A hurdle negative binomial (HNB) regression model combined with a distributed lag non-linear model (DLNM) was applied here to explore the association between daily HRI counts with weather conditions. The hurdle model was suggested by [11] [12] [13] to deal with counts data with extra zeros. In this study, we employed a hurdle model as, in about 77.02% (100% - 22.98%) days of the studied period, no HRI cases were observed in our data (see Table 1). The hurdle model was aimed to explore the impact of weather factors on binary result of

the presence of HRI or not, as well as the impact on the count of HRI cases if the count is not zero.

Actually, the occurrence of HRI is a comprehensive impact of weather factors and other factors, such as physical activities, the lack of environmental acclimatization, poor physical fitness, and illness [17] [24]. In consensus, if the count of HRI is abnormally high, it implies a high possibility that other than weather factors have stepped in. For example, according to our collected HRI counts (2002-2010) on the dates when Standard Chartered Marathon was held in Singapore (see **Table 3**), the counts of HRI cases were in a range from 4 to 12 compared to the overall mean of 0.34 cases (see **Table 2**). We need to arrange different models to reflect the different scenarios. Here, the separation line was assumed to be at the location where the counts of HRI cases were zero or more as more than normal zeros were observed. Therefore, the hurdle model was employed where the zero counts and the positive counts of HRI cases were modeled by different distributions.

On the other hand, the positive count of HRI cases could be accounted for different origins. Hence, zero-truncated Negative binomial (ZNB) distribution was applied to describe the count, as a NB distribution is a mixture of Poisson distribution [25]. Furthermore, according to **Table 1**, HRI counts exhibited a long and flat tail. Again, the NB distribution was adopted, as it was similar to a Poisson model but with a longer, fatter tail to the extent that the variance exceeds the mean [26].

A DLNM was used to check if weather conditions had delayed adverse effect on the occurrence of HRI. At the same time, linear or nonlinear effect was also investigated through DLNMs.

For the zero hurdle part, a censored Poisson (right-censoring at zero) was employed to model the impact of weather on the occurrence and non-occurrence of HRI cases. The Poisson mean on day t was denoted as μ_{0t} . On the other hand, the positive counts part was modeled by a ZNB model with the negative binomial mean of μ_{1t} and shape parameter of θ on day t . Modeling processes for the two parts could be conducted separately according to the work by [27]. In this study, μ_{kt} was assumed to satisfy:

$$\begin{aligned} \log(\mu_{kt}) = & \alpha_k + \sum_j s_k(\text{Weather}_{jt}) + \delta_k \cdot \text{Dow}_t + \gamma_k \cdot \text{Holiday}_t \\ & + \log(\text{Pop}_t) + \text{trend}_{kt} \end{aligned} \quad (2)$$

where α_k was the intercept, $s_k(\cdot)$ was a smoothing function. It was noted that the default knots were equally-spaced percentiles. $s_k(\text{Weather}_{jt})$ were obtained by applying a DLNM to Weather_{jt} (*i.e.*, MinT, MeanT, MaxT, RH,

Table 3. Count of HRI on the day of Standard Chartered Marathon Singapore (2002-2010).

Date	12/8/2002	12/7/2003	12/5/2004	12/4/2005	12/3/2006	12/2/2007	12/7/2008	12/6/2009	12/5/2010
HRI	4	9	12	5	5	6	5	4	9

Wdsp and WBG_T_MeanT). Dow_t was day t of the week, and δ_k was vector of coefficient. $Holiday_t$ was a binary variable that was “1” if day t was a holiday, and γ_k was the coefficient. Dow_t and $Holiday_t$ were used to reflect the impact of social activities related with heat illness. For the offset term, Pop_t was the total population of Singapore on day t . As only the mid-year population was obtained (from the Department of Statistics, Singapore), daily population was approximated by linear interpolation. $trend_{kt}$ was determined through n -phase piecewise linear functions. n -phase piecewise linear function means that the piecewise linear function has n phases.

R software (version 3.1.1; R Development Core Team, 2014) was used to fit all models. The “dlnm” package was used to create the DLNM [28]. For the zero hurdle part, let π_t denote the probability of at least one case of HRI observed on day t , then $\pi_t = 1 - \exp(-\mu_{0t})$, *i.e.*, $\log(\mu_{0t}) = \log(-\log(1 - \pi_t))$. Hence, the censored Poisson was modeled by a binomial regression model with a “pois” link in the “glm” function. For the positive counts part, a ZNB model with a “log” link was used. “countreg” package was employed to create the ZNB model and “pscl” package to create the final HNB model.

The modeling process was as follows. The first step was to determine the trend of HRI cases through n -phase piecewise linear functions. The intercept and slope might differ for different phases. Corresponding knots were chosen among a sequence of the last day of dual months: “02/28/1991”, “04/30/1991”, ..., “10/31/2010”, “12/31/2010”. The Akaike information criterion (AIC) was used to determine the best trend. For the zero hurdle part, a three phase piecewise linear function with knots “2/28/1994” and “6/30/2002” was detected. On the other hand, for the positive counts part, a two phase piecewise linear function with the knot “10/31/1998” was detected, *i.e.*, a change point was detected and it happened during the period between year 1998 and 1999 (see **Figure 2**).

Before this point, the pattern of HRI monthly counts was an apparent

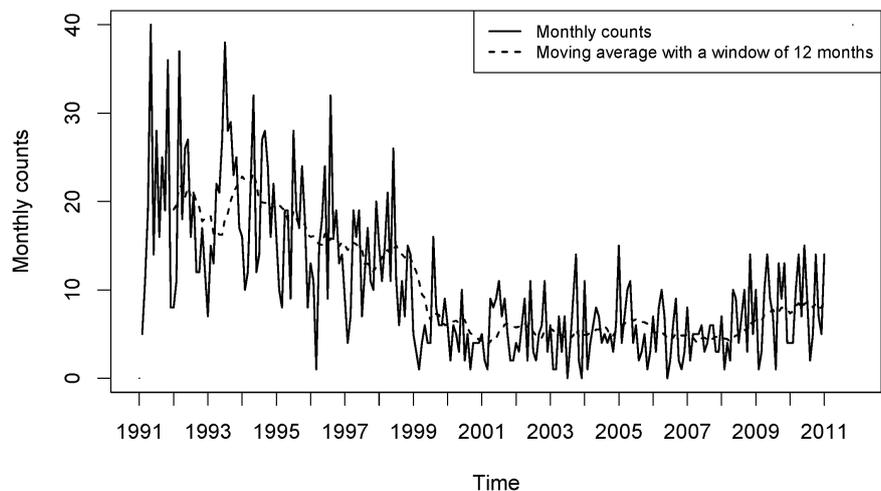


Figure 2. Monthly counts of hospital admission for heat related illness cases, Singapore, 1991-2010.

decreasing trend with higher number of cases and higher variance. After this point, the pattern showed a relatively stable trend with relatively lower number of cases, lower variance.

The second step was to explore the weather predictors that best predicted the occurrence of HRI after adjusting for trend, holiday and specific days in a week. A DLNM was used to find the best weather predictors. A natural splines (ns) function, with degree of freedom (df) of 3 and knots placed at equally-spaced quantiles, was applied to reflect the nonlinear effect of each weather predictor.

In the censored Poisson model, both single lag effect and distributed lag effect were investigated with a maximum lag number of 14 days. For a single lag model, AIC of the single lag models was used to determine the best lag. In our study, the best lag numbers found for weather predictors were all 0s. For the distributed lagged effect model (DLNM), a different ns function with df of 2 was used to describe the effect of lags. Nevertheless, constrained DLNM does not exist when the maximum lag number is less than 2. Hence unconstrained DLNM was used for the lag number 0 and 1. During this process, we also found that the best lags for weather predictors were all 0's.

In the TZBN model, similar model selection procedure was conducted as described above for the binomial model. For a single lag model, the current effect (lag number = 0) of all considered weather parameters was found with smaller AIC values than other lagged effect. For the distributed lagged effect model, under AIC, the best maximum number of lags of all considered weather predictors were all 0's except for approximated WBGT, in which the maximum number of lag 3 was associated with the smallest AIC values. We then applied Vuong test (in "pscl" package, R software, version 3.1.1), which was designed to do model selection between two non-nested models [29], to test the significance of the improvement of the model only considering current effect to the model cumulating up to 3 lagged effects. It was interesting to observe that the model with only current effect was significantly better than the other model (p -value < 0.01). The reason is because in the Vuong test, there is more punishment on the number of parameter than an AIC does. On the other hand, the AIC value of the current-effect model was slightly higher than the cumulative-lag-effect model. Hence, only current effect of approximated WBGT was considered in the following step of the modeling process.

The third step was to explore the linear or nonlinear effect of the weather predictors on HRI. The ns functions with df from 1 to 5 was compared to reflect the impact of the weather predictors. When $df = 1$, the ns function was actually a linear function. AIC was also used to determine the best model.

The fourth step was to examine the best combination of weather predictors. The first weather predictor was selected if it was associated with the smallest AIC value among all candidate weather predictors. After the first weather predictor was included, a second predictor variable was included if the AIC value of the new model was reduced. The process continued until no new entry could reduce

the AIC value or all weather predictors were included in the model. During the selection procedure, the best ns function was selected and used for each new entry.

During the process, if a complicated model was first selected according to AIC values, the significance of its superiority against a corresponding simple model was further tested. If it was not significant, the simple model would be chosen.

3. Results

There were 2474 HRI cases admitted to hospital from year 1991 to 2010 (7035 days) with a daily mean of 0.34 cases (see **Figure 1**). The daily incidence of HRI was sparse and the proportion of days with HRI was only 22.98% (1679 days, see **Figure 1** and **Table 1**). The proportion of days with HRI cases was the lowest (7.46%) in the “public holiday” category, followed by the “Sunday” category (16.2%). However, high incidence of HRI occurred mostly on Sundays (32.5% of days with 5 and more cases).

In the study period, the average daily maximum temperature (MaxT) was 31.6°C, mean temperature (MeanT) was 27.7°C, minimum temperature (MinT) was 24.9°C, mean relative humidity was 83.4%, and mean wind speed was 1.9 m/sec (see **Table 2**). The correlation was 0.53 between MaxT and MinT, 0.81 between MaxT and MeanT, and 0.80 between MeanT and MinT (see **Table 4**). The correlation between relative humidity and the three types of temperatures were all negative. The highest negative correlation was -0.76 between RH and MeanT (see **Table 4**). The correlation between WBGT_MeanT and MeanT was the highest (0.91). Hence, in the model selection procedure, we avoided putting both WBGT_MeanT and MeanT in the same model.

Mean temperature and approximated WBGT were associated with lower AIC values (see **Table 5**). In the binomial model for the occurrence of HRI or not, MeanT was associated with the lowest AIC. Hence, MeanT was first selected into the model. On the other hand, approximated WBGT was associated with the lowest AIC in the ZTNB model for the non-zero counts part of HRI. Nevertheless, the AIC associated with approximated WBGT was only slightly lower than that associated with MeanT (**Table 5**). A Vuong test was further applied to test

Table 4. Pearson’s correlation coefficients among weather conditions in Singapore, 1991-2010.

	MeanT	MaxT	RH	Wdsp	WBGT_MeanT
MinT	0.80	0.53	-0.53	0.24	0.77
MeanT		0.81	-0.76	0.27	0.91
MaxT			-0.66	0.17	0.71
RH				-0.50	-0.41
Wdsp					0.05

All the correlation coefficients are statistically significant with $p < 0.001$.

Table 5. Hurdle binomial models associated with a single weather predictor.

Weather predictor	MinT	MeanT	MaxT	RH	Wdsp	WBG_T_meT
Zero hurdle part (censored Poisson model)						
AIC ^a	7155.9	7041.6	7090.4	7205.1	7319.0	7047.5
Best df ^b	3	1	3	2	3	1
Non-zero counts part (ZTNB model)						
AIC ^a	2958.3	2948.8	2952.2	2967.1	2974.8	2947.7
Best df ^b	1	1	1	2	2	1

^aThe AIC values were obtained for anns function with the best df (df: from 1 to 5) and only current effect was included. ^bThe best df was selected in from 1 to 5.

the significance of the difference. It showed that the model with approximated WBG_T was not significantly better than the model associated with MeanT (p-value = 0.4). Hence, MeanT was first included in the model.

In the censored Poisson model for the occurrence of HRI or not, mean temperature was first chosen, following which the model further selected relative humidity and MaxT. It was noted that linear model (df of ns is 1) showed the best prediction capability (AIC value was the lowest) among all candidate ns functions (df: from 1 to 5) for all the selected weather predictors. The estimated coefficients and results were presented in **Table 6**.

In the ZTNB model for positive counts of HRI cases, MeanT with linear effect (df of ns is 1) was selected, as it was associated with the smallest AIC value. After that, wind speed with linear affect was selected. However, the effect of wind speed was observed not significant with p-value = 0.11. Hence, only MeanT was included in the ZTNB model.

For the zero hurdle part of HRI (**Table 6**) after adjusting for relative humidity and MaxT, a 1°C increase in MeanT was associated with a 43% (95% CI: 31%, 57%) increase in the occurrence rate of HRI, which was consistent with the effect estimated by the single weather variable model. After adjusting for MeanT and MaxT, every unit increase in relative humidity was associated with a 3% (95% CI: 1%, 5%) increase in the occurrence rate of HRI. Every degree increase in MaxT was associated with an 11% (95% CI: 5%, 18%) increase in the occurrence rate of HRI after adjusting MeanT and relative humidity. For the positive counts part of HRI (**Table 6**), every degree increase in MeanT was associated with a 31% (95% CI: (18%, 46%)) increase in the occurrence rate of HRI.

The high temperature is the key factor leading to the occurrence of HRI. Considering the concern of global climate change, we may foresee the increase of temperature in both tropical and temperate regions. In IPCC Fourth Assessment Report [30], it is projected that the Mean T may increase at least 1.5 degree (2.7°F) under “all scenarios except the one representing the most aggressive mitigation of greenhouse gas emissions”. With such an increase, the relative risk for the occurrence of HRI increases 72% on average in Singapore with the

Table 6. Summary of hurdle negative binomial model associating daily weather predictors and the counts of hospital admission for heat related illness, Singapore, 1991-2010 (Sunday is the reference group).

Zero hurdle model coefficients (censored Poisson with log link, AIC = 7022.5)					
	Estimate	Std. Error	z value	Pr (> z)	RR ^a (95%CI ^b)
(Intercept)	-26.6893	1.6537	-16.1388	< 0.001	-----
MeanT	0.3599	0.0459	7.8486	< 0.001	1.43 (1.31, 1.57)
RH	0.0286	0.0083	3.4361	< 0.001	1.03 (1.01, 1.05)
MaxT	0.1046	0.03	3.4892	< 0.001	1.11 (1.05, 1.18)
Holi	-1.2769	0.2316	-5.5131	< 0.001	0.28 (0.18, 0.44)
Monday	0.3625	0.1041	3.484	< 0.001	1.44 (1.17, 1.76)
Tuesday	0.5271	0.0997	5.2849	< 0.001	1.69 (1.39, 2.06)
Wednesday	0.5246	0.1008	5.2021	< 0.001	1.69 (1.39, 2.06)
Thursday	0.605	0.0988	6.122	< 0.001	1.83 (1.51, 2.22)
Friday	0.6205	0.0989	6.2769	< 0.001	1.86 (1.53, 2.26)
Saturday	0.3838	0.1018	3.7714	< 0.001	1.47 (1.20, 1.79)
v1 ^a	1.3e-04	2.17e-5	6.0064	< 0.001	-----
v3 ^b	-7.8e-04	4.04e-5	-17.7376	< 0.001	-----
Count model coefficients (ZTNB with log link, AIC = 2948.8)					
	Estimate	Std. Error	z value	Pr (> z)	
(Intercept)	-30.4068	510.4238	-0.0596	0.9525	-----
MeanT	0.2728	0.0544	5.013	< 0.001	1.31 (1.18, 1.46)
Holi	-0.9665	0.7309	-1.3224	0.186	0.38 (0.09, 1.59)
Monday	-1.1671	0.2459	-4.7469	< 0.001	0.31 (0.19, 0.50)
Tuesday	-1.3346	0.2379	-5.6093	< 0.001	0.26 (0.17, 0.42)
Wednesday	-1.5246	0.2455	-6.211	< 0.001	0.22 (0.13, 0.35)
Thursday	-1.1226	0.2335	-4.8083	< 0.001	0.33 (0.21, 0.51)
Friday	-1.2291	0.235	-5.2291	< 0.001	0.29 (0.18, 0.46)
Saturday	-1.366	0.2438	-5.6021	< 0.001	0.26 (0.16, 0.41)
a1 ^c	1.8163	0.1978	9.1823	< 0.001	-----
a3 ^d	-3.97e-04	1.20e-04	-3.3139	< 0.001	-----
Log (theta)	-14.8195	510.4217	-0.029	0.9768	-----

^arelative risk. ^bconfidence interval. ^cv1 = t - 4199. ^dv3 = (t - 1155)*I (1155 < t < 4199) - 3044*I (t < 4199) ^ea1 = I (t < 2861). ^fa3 = t*I (t < 2861). where t is the number of days starting from 1st Jan. 1991, in which t = 1, 1155, 2861 and 4199 imply respectively, 1st Jan. 1991, 28th Feb. 1994, 31st Oct. 1998 and 30th Jun. 2002. I (x) is an indicator function which is equal to 1 when x is true and 0 otherwise.

assumption that the current demographic, social and environmental settings remain similar. *Sensitivity analysis* To evaluate if the modeling on the impact of weather predictors is stable, we examined the estimated coefficients of weather predictors when our model adopted different trends. For example, we exchanged

the trend functions used in the binomial regression model and in the ZTNB model. The model AIC values increased slightly, but results for the impact of weather predictors were still similar. In addition, when we included the information of HRI of the Standard Chartered Marathon Singapore from 2002-2010 (see **Table 3**) in the model, the AIC values of the model decreased, but it gave similar estimation for the impact of weather predictors. These further indicated that the models built in this study had adequately captured the main effects of weather predictors on the occurrence of HRI.

Model validation To validate the model, we used data from 1991 to 2008 as training data to estimate model parameters, then applied the trained model to predict HRI cases in 2009 and 2010 given weather variables in these two years. The results were shown in **Figure 3**. It was observed that the predicted frequency of daily counts of HRI cases was close to the observed frequency.

4. Discussion

To our best knowledge, this study is the first mathematical modeling study to demonstrate the combined effect of weather conditions on the incidence of hospital admissions for HRI over 20 years in Singapore. This study provides an extensive investigation on the association between weather conditions and occurrence of HRI. The model developed can be used to predict HRI occurrences to enable the mitigation of HRI and preparation of onsite medical cares. Further, based on the model constructed, we also estimate the future HRI occurrences in Singapore under climate change impact to provide references for future related policy making.

In our study, a hurdle model showed that the impacts of weather factors on the occurrence and non-occurrence of HRI cases and on the positive counts of HRI cases were different and extra zeros existed. The extra zeros might be

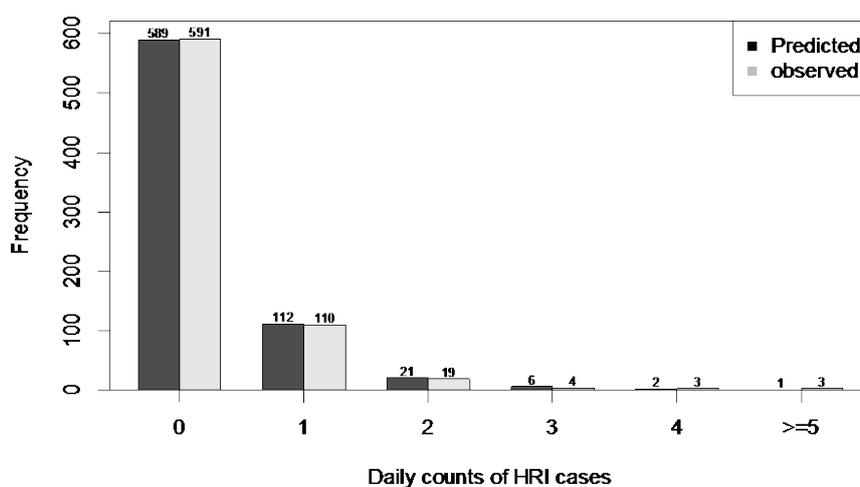


Figure 3. Observed and predicted frequency of daily counts of HRI cases in 2009 and 2010 based on hurdle negbin model with parameters estimated according to data from 1991 to 2008.

because some people were vulnerable to HRI incidence while others were not. In [1], many zero counts were also observed, but this phenomenon was ascribed to the small sample size in that study. A zero inflated Poisson model was applied in [3], but only results on positive counts of HRI were reported. It was also noticed that there was a change point between year 1998 and 1999. The HRI monthly counts before this point were much higher than after. The phenomenon of less vulnerable people than before might get benefit from the popular of air-conditioners.

Besides weather factors, the impact of day-of-the-week (**Table 6**) on HRI cases also showed some interesting results. Compared to Sunday, in other days of a week, the possibility of the occurrence of HRI cases was significantly high. However, once HRI cases occurred, the number of HRI cases in these days was significantly smaller than in Sundays. The reason might be because in Sundays, people were usually off from work. Hence, in general, less people than other days were exposed to HRI incidence. However, some activities, which were more likely to trigger HRI and evolved a large amount of people, were held in some Sundays. For example, Standard Chartered Marathon Singapore is held on very first Sunday of December in each year and attracted about 65,000 runners in 2011. On these days when the activity was held, an abnormally high number of HRI cases observed (see **Table 3**).

The positive counts of HRI were described by a zero-truncated negative binomial (ZTNB) model, as NB distribution was a mixture of several Poisson distributions [25]. The NB distribution was selected instead of a Poisson distribution because the origins of HRI cases might be also attributed to other factors besides weather, to name a few, physical activities, the lack of environmental acclimatization, poor physical fitness, dehydration, medications, drugs supplements, high body mass, specific genotypes, muscle injury, and illness [17] [24].

The tropical climate in Singapore will continue to pose a health risk on HRI, especially among those who undertake strenuous outdoor physical activity. With the global climate change, the temperature is with increasing trend, which will impact more to the health risk of HRI in tropical country, like Singapore. This will alert the public health policy makers to think about corresponding policy to mitigate the increase of HRI cases and utilize HRI prediction model to forecast HRI occurrences to prevent HRI cases or prepare for onsite medical cares.

5. Conclusion

Heat related illness is an event with low occurrence rates in Singapore since the late of 1990s. To reveal the impact of weather factors on HRI and the future public health risk of HRI, we proposed a hurdle negative binomial regression combined with a distributed lag nonlinear model considering that hurdle models could answer simultaneously the possibility of occurrence of HRI and the count when it occurred. The model constructed could be used to predict the occurrence of HRI cases and evaluate the future public health burden of HRI under climate change trend, which can provide useful inputs for policy makers for related policy making. The proposed modeling approach can also be applied to

model and evaluate lag and nonlinear effect of environmental factors on other public health concerns based on count data with excess zeros.

References

- [1] Shanks, N.J. and Papworth, G. (2001) Environmental Factors and Heat Stroke. *Occupational Medicine*, **51**, 45-49. <https://doi.org/10.1093/occmed/51.1.45>
- [2] Piver, W.T., Ando, M., Ye, F. and Portier, C.J. (1999) Temperature and Air Pollution as Risk Factors for Heat Stroke in Tokyo, July and August 1980-1995. *Environmental Health Perspectives*, **107**, 911-916. <https://doi.org/10.1289/ehp.99107911>
- [3] Bassil, K.L., Cole, D.C., Moineddin, R., Lou, W., Craig, A.M., Schwartz, B., *et al.* (2011) The Relationship between Temperature and Ambulance Response Calls for Heat-Related Illness in Toronto, Ontario, 2005. *Journal of Epidemiology and Community Health*, **65**, 829-831. <https://doi.org/10.1136/jech.2009.101485>
- [4] Kumar, S. (2001) Heat Disorders amongst Military Recruits in the Singapore Armed Forces: A Descriptive and Case-Control Study. Ph.D. Dissertation, National University of Singapore, Singapore.
- [5] Lim, M.K. (1989) Heat Stroke. *SMJ*, **30**, 127-128.
- [6] Lim, M.K. (1994) Occupational Heat Stress. *Annals Academy of Medicine Singapore*, **23**, 719-724.
- [7] Seth, P. and Juliana, P. (2011) Exertional Heat Stroke in a Marathon Runner with Extensive Healed Deep Burns: A Case Report. *International Journal of Emergency Medicine*, **4**, 12. <https://doi.org/10.1186/1865-1380-4-12>
- [8] Tham, M.K., Cheng, J. and Fock, K.M. (1989) Heat Stroke: A Clinical Review of 27 Cases. *SMJ*, **30**, 137-140.
- [9] Lee, L., Chia, H.P. and Tan, E.H. (1991) Prevention of Heat Disorders in the Singapore Armed Forces—1984-1989. *Annals Academy of Medicine Singapore*, **20**, 347-350.
- [10] Lee, L., Fock, K., Lim, C., Ong, E., Poon, B., Pwee, K., *et al.* (2010) Singapore Armed Forces Medical Corps-Ministry of Health Clinical Practice Guidelines: Management of Heat Injury. *SMJ*, **51**, 831-835.
- [11] Mullahy, J. (1986) Specification and Testing of Some Modified Count Data Models. *Journal of Econometrics*, **33**, 341-365. [https://doi.org/10.1016/0304-4076\(86\)90002-3](https://doi.org/10.1016/0304-4076(86)90002-3)
- [12] King, G. (1989) Event Count Models for International Relations: Generalizations and Applications. *International Studies Quarterly*, **33**, 123-147. <https://doi.org/10.2307/2600534>
- [13] Rose, C.E., Martin, S.W., Wannemuehler, K.A. and Plikaytis, B.D. (2006) On the Use of Zero-Inflated and Hurdle Models for Modeling Vaccine Adverse Event Count Data. *Journal of Biopharmaceutical Statistics*, **16**, 463-481. <https://doi.org/10.2307/2600534>
- [14] Armstrong, B. (2006) Models for the Relationship between Ambient Temperature and Daily Mortality. *Epidemiology*, **17**, 624-631. <https://doi.org/10.1097/01.ede.0000239732.50999.8f>
- [15] Gasparrini, A., Armstrong, B. and Kenward, M.G. (2010) Distributed Lag Non-Linear Models. *Statistics in Medicine*, **29**, 2224-2234. <https://doi.org/10.1002/sim.3940>
- [16] Guo, Y., Barnett, A.G., Pan, X., Yu, W. and Tong, S. (2011) The Impact of Temper-

- ature on Mortality in Tianjin, China: A Case-Crossover Design with a Distributed Lag Nonlinear Model. *Environmental Health Perspectives*, **119**, 1719-1725. <https://doi.org/10.1289/ehp.1103598>
- [17] Coris, E.E., Ramirez, A.M. and Van Durme, D.J. (2004) Heat Illness in Athletes: The Dangerous Combination of Heat, Humidity and Exercise. *Sports Medicine*, **34**, 9-16. <https://doi.org/10.2165/00007256-200434010-00002>
- [18] Parsons, K. (2006) Heat Stress Standard ISO 7243 and Its Global Application. *Industrial Health*, **44**, 368-379. <https://doi.org/10.2486/indhealth.44.368>
- [19] Bureau of Meteorology, Australia (2010) Thermal Comfort Observations. http://www.bom.gov.au/info/thermal_stress/
- [20] Teets, H.J. and Fehn, S. (2007) Improved Heat-Stress Algorithm. NASA Dryden Flight Research Center, DRC-05-01.
- [21] Howe, A.S. and Boden, B.P. (2007) Heat-Related Illness in Athletes. *The American Journal of Sports Medicine*, **35**, 1384-1395. <https://doi.org/10.1177/0363546507305013>
- [22] Yaglou, C.P. and Minard, D. (1957) Control of Heat Casualties at Military Training Centers. *Archives of Industrial Health*, **16**, 302-316.
- [23] American College of Sports Medicine (1984) Prevention of Thermal Injuries during Distance Running-Position Stand. *Medicine and Science in Sports and Exercise*, **16**, 9-14. <https://doi.org/10.1249/00005768-198410000-00017>
- [24] Cuddy, J.S. and Ruby, B.C. (2011) High Work Output Combined with High Ambient Temperatures Caused Heat Exhaustion in a Wildland Firefighter Despite High Fluid Intake. *Wilderness & Environmental Medicine*, **22**, 122-125. <https://doi.org/10.1016/j.wem.2011.01.008>
- [25] Zeileis, A., Kleiber, C. and Jackman, S. (2008) Regression Models for Count Data in R. *Journal of Statistical Software*, **27**, 1-25. <http://www.jstatsoft.org/v27/i08> <https://doi.org/10.18637/jss.v027.i08>
- [26] Hu, M.-C., Pavlicova and Nunes, E.V. (2011) Zero-Inflated and Hurdle Models of Count Data with Extra Zeros: Examples from an HIV-Risk Reduction Intervention Trial. *The American Journal of Drug and Alcohol Abuse*, **37**, 367-375. <https://doi.org/10.3109/00952990.2011.597280>
- [27] Xu, H.-Y., Xie, M. and Goh, T.N. (2014) Objective Bayes Analysis of Zero-Inflated Poisson Distribution with Application to Healthcare Data. *IIE Transactions*, **46**, 843-852. <https://doi.org/10.1080/0740817X.2013.770190>
- [28] Gasparrini, A. (2011) Distributed Lag Linear and Non-Linear Models in R: The Package DLNM. *Journal of Statistical Software*, **43**, 1-16. <https://doi.org/10.18637/jss.v043.i08>
- [29] Vuong, Q.H. (1989) Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses. *Econometrica*, **57**, 307-333. <https://doi.org/10.2307/1912557>
- [30] US EPA CCD (2017) Future of Climate Change. https://19january2017snapshot.epa.gov/climate-change-science/future-climate-change_.html#Temperature