International Journal of Mathematics and
Computer Science, **17**(2022), no. 1, 243–253

The Risk Area Assessment on Re-emerging Diseases in Elderly People by Using Extreme Value Theory

Nipaporn Chutiman[1], **Pannarat Guayjarernpanishk**[2], **Monchaya Chiangpradit**[1], **Piyapatr Busababodhin**[1], **Saowanee Rattanawan**[1], **Tossapol Phoophiwfa**[1], **Butsakorn Kong-ied**[1]

[1]Data Science and Sustainable Agriculture Research Unit
Department of Mathematics
Faculty of Science
Mahasarakham University
Maha Sarakham 41150, Thailand

[2]Faculty of Interdisciplinary Studies
Nong Khai Campus
Khon Kaen University
Nong Khai 43000, Thailand

email: nipaporn.c@msu.ac.th, panngu@kku.ac.th, monchaya.c@msu.ac.th, piyapatr.b@msu.ac.th, saowanee.r@msu.ac.th, tossapol.pho@msu.ac.th, butsakorn.k@msu.ac.th

(Received July 5, 2021, Accepted August 17, 2021)

Abstract

In this paper, we apply the Extreme Value Theory (EVT) in the analysis for finding an appropriate distribution pattern by using the Generalized Extreme Value (GEV) with the monthly data of the food poisoning rate in elderly people to forecast the chance of food poisoning in elderly people in Khon Kaen Province, Maha Sarakham Province, and Roi Et Province in the middle northeastern part of Thailand. Analyzing our results revealed that the GEV was appropriate according to the goodness of fit test with Kolmogorov-Smirnov

Key words and phrases: Re-emerging Diseases, Risk Area Assessment, Extreme Value Theory.
AMS (MOS) Subject Classifications: 60G70, 62P10, 62P25.
Corresponding author email: butsakorn.k@msu.ac.th
ISSN 1814-0432, 2022, http://ijmcs.future-in-tech.net

Statistics (KS Test) and the estimation of the return levels in the annual return period of 2, 5, 10, 20, 25, and 50 years. Our results showed the area with the maximum food poisoning rate and the findings provided information to make decisions in planning to cope with the risk areas and for setting guidelines to plan public health affairs to cope with possible next outbreaks.

1 Introduction

The current population structure of Thailand has changed into an aging society in 2020 with 11,627,130 elderly people at the age of 60 years or above. That constitutes 17.57 percent of the total population in the country whereas the largest number of the elderly people is in the northeastern part at 3,684,395. The Ministry of Social Development and Human Security estimates that in the next 20 years, the number of elderly people at the age of 60 years or above in Thailand will rise to 30 percent. On the other hand, it is expected that, due to climate change such as high temperature and humidity, seasonal contagious diseases may occur. In addition, the spread of non-carrier diseases such as diarrhea and food poisoning is sometimes affected by climate change and is usually found more in the elderly. Boon et al. [1] found that human mechanism has to adapt to climate change and natural disasters which are expected to be more frequent. These disasters have both direct and indirect effects on the groups of poor and elderly people which are directly affected by disasters. Therefore, adaptation to such effects need cooperation and data from all sectors of the government and civil societies. The assessment on life quality of the elderly affected by global warming and disasters is therefore necessary to maintain their well-being and quality of life.

Forecasting on diseases and health threats is very important and the forecasting results enable epidemiological surveillance to estimate the chance of emergence of a future diseases. The results are beneficial to plan resources for effective prevention and treatments. Some models may have weaknesses with high errors of the forecasting results. Such errors cause decision-making incompatible with future situations. As a result, appropriate models for forecasting disease emergence and appropriate distribution of epidemiology data are very important in explaining a disease outbreak. In time-series forecasting with public health data, problems of errors are usually found with incompatible with the assumptions. When an outbreak occurs resulting in a large number of dead people, such data are regarded as abnormal and so must be excluded from the analysis. However, the data of infection rates from the maximum epidemics are very important for planning and preparing resources to cope with them. To obtain high effectiveness in planning, applying the Extreme Value Theory (EVT) explains the maximum rates of infection and death.

Ebi et al. [2] studied the effects of climate change on health in the United States and found that the trends increasing illness and thus the risk of death have both direct and indirect effects on human health. Chen et al. [3] estimated the epidemics of influenza in Zhejiang Province, China by using the data from April 2009 until November 2013 and the EVT and found that this model is appropriate in forecasting the influenza epidemics there. Thomas et al. [4] estimated the weekly death rate of pneumonia and influenza in France by using the EVT and found that the Generalized Extreme Value (GEV) was appropriate in estimating the return level and some risk criteria were more than the existing observation values. This method could be used for public health planning with the extreme values of the events, and it can be used with various issues in epidemiology. However, future studies are needed on some important issues.

The objective of our paper is a study on forecasting the food poisoning rate of the elderly people in Khon Kaen Province, Maha Sarakham Province and Roi Et Province where the food poisoning rates of the elderly people were on the high ranks of the middle northeastern part of Thailand. In this study an area-based analysis is performed by applying the EVT with the GEV and the estimation of the return level in the annual return period of 2, 5, 10, 20, 25 and 50 years. The estimation of the return level can inform the situations and the maximum return level in each area. The method is useful for effectively planning in resource preparation and prevention measures.

2 Data preparation

This research used the monthly data of the food poisoning rate per 10,000 elderly population from January 2013 until December 2020; i.e., 96 months in total with classification according to the districts with complete data in Khon Kaen Province, Maha Sarakham Province, and Roi Et Province. In the area-based analysis, the EVT was used with the GEV.

3 Methodology

3.1 Generalized Extreme Value

The Generalized Extreme Value (GEV) is suitable for analyzing the extreme value in the interested time period such as in years, months, quarters, or weeks. The data on extreme value in each studied period are selected. This selection is called the Block Maxima Method. In GEV, there are 3 parameters: μ referring to location, σ referring to scale, and ξ referring to shape [5].

Let x be a random variable represented by $x \sim \mathrm{GEV}(\mu, \sigma, \xi)$ with the function of cumulative distribution [6] as follows:

$$F(x) = exp\left\{ -\left(1 + \xi\frac{x-\mu}{\sigma}\right)^{-1/\xi} \right\}, \tag{3.1}$$

where $1 + \xi\left(\frac{x-\mu}{\sigma}\right) > 0$ is represented by μ as a location parameter, σ as a scale parameter, and ξ as a shape parameter. The GEV can be classified into three types according to the shape parameter. The case of $\xi \longrightarrow 0$ is called Gumbel Distribution, the case of $\xi > 0$ is called Fréchet Distribution, and the case of $\xi < 0$ is called Weibull Distribution.

3.2 Tests on the model appropriateness with Kolmogorov-Smirnov Statistics

Kolmogorov-Smirnov Statistics (KS Test) is a statistical test on appropriateness of distribution. Let X be a continuous random variable with distribution function $F(x)$ and let $X_1, X_2, X_3, \ldots, X_n$ be a random sample from X with order statistics $X_{(1)}, X_{(2)}, X_{(3)}, \ldots, X_{(n)}$. We wish to test the null hypothesis $H_0 : F(x) = F_0(x)$, for all $x \in (-\infty, \infty)$ against the general alternative $H_1 : F(x) \neq F_0(x)$, for some $x \in (-\infty, \infty)$, where $F_0(x)$ is a hypothesized distribution function to be tested.

$$
\begin{aligned}
KS^2 &= \left\{ \sup_{t \in (-\infty,\infty)} \left| F_n(t) - F_0(t) \right| \right\}^2 \\
&= \left(\max_{1 \leq i \leq n} \left[\max \left\{ \frac{i}{n} - F_0(X_{(i)}), F_0(X_{(i)}) - \frac{i-1}{n} \right\} \right] \right)^2,
\end{aligned}
$$

where KS the best-known statistics for goodness-of-fit-tests [7].

4 Results

In the data analysis using the Extreme Value Theory (EVT), the appropriate model was identified by the Generalized Extreme Value (GEV), the analysis of the extreme value in a monthly period, and the analysis of the return level. The results are presented in Tables 13. The extreme values of the food poisoning rate in elderly people in different return levels are transformed into contour graphs by using Geographic Information System (GIS) Kreiging interpolation in GIS, and the results are illustrated in Figure 1.

Table 1: The GEV distribution of the food poisoning rate in the elderly people classified according to districts in Khon Kaen Province, and the return levels in 2, 5, 10, 20, 25 and 50 years.

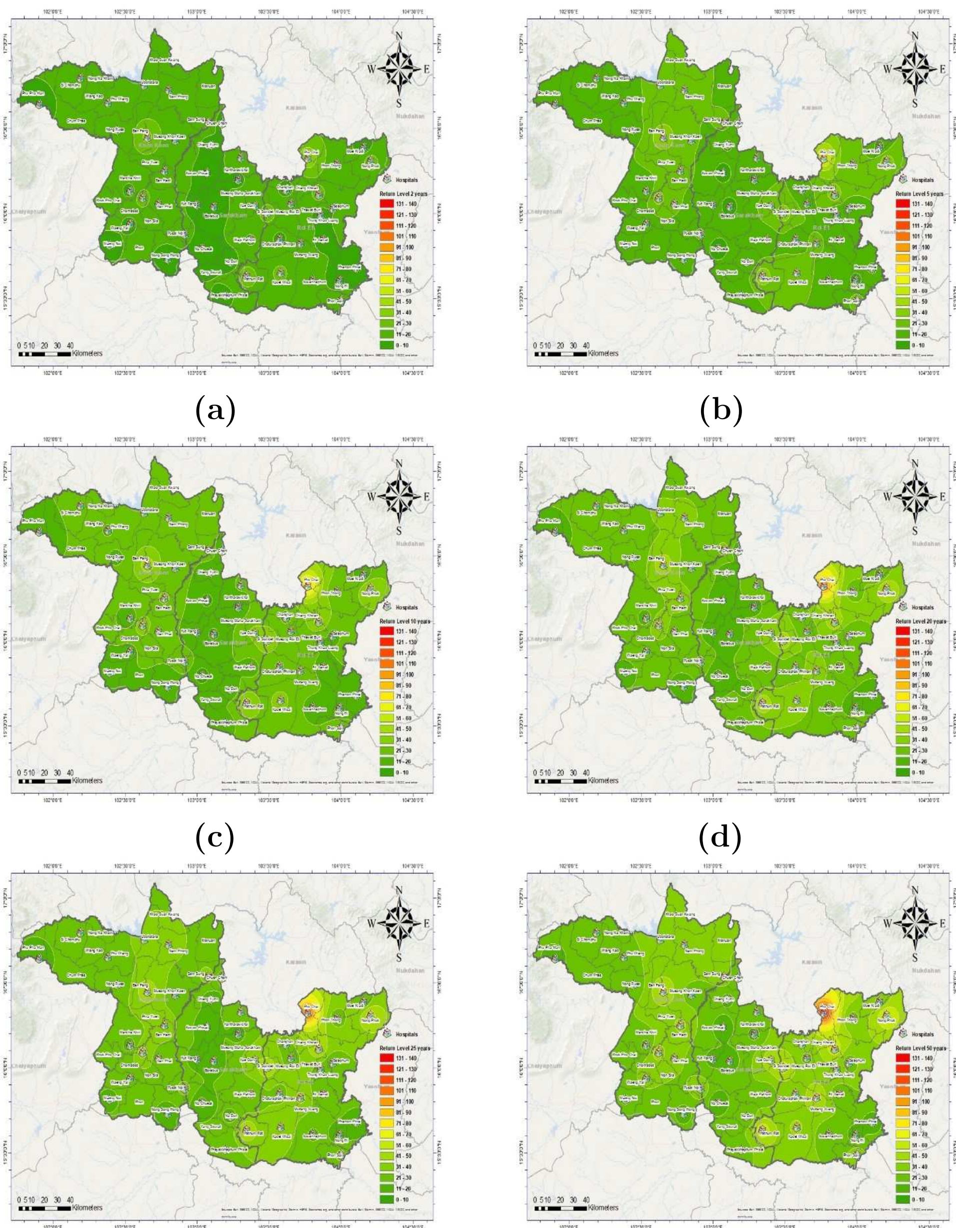
Districts	distribution	Return levels (years)					
		2	5	10	20	25	50
Ban Fang	**Weibull**	**33.03**	**41.23**	**45.71**	**49.42**	**50.49**	**53.50**
Chonnabot	**Gumbel**	**27.40**	**35.96**	**40.66**	**44.56**	**45.69**	**48.85**
Sam Sung	Gumbel	20.83	26.11	29.91	33.81	35.1	39.24
Nam Phong	Gumbel	19.08	25.15	29.16	32.98	34.19	37.92
Chum Phae	Gumbel	18.10	25.83	29.77	32.86	33.72	36.06
Ban Haet	Gumbel	17.48	26.25	31.82	36.98	38.59	43.43
Meuang	Weibull	16.57	20.74	22.86	24.52	24.98	26.23
Ban Phai	Gumbel	15.18	20.48	23.71	26.61	27.5	30.1
Ubonratana	Gumbel	14.02	21.2	26.39	31.72	33.49	39.19
Phu Wiang	Weibull	13.51	17.55	19.56	21.09	21.51	22.64
Pueai Noi	Gumbel	13.08	21.55	27.16	32.55	34.25	39.52
Nong Roea	Weibull	11.64	15.98	18.42	20.49	21.09	22.81
Waeng Yai	Frchet	8.41	13.82	18.3	23.43	25.25	31.52
Waeng Noi	Gumbel	8.04	11.87	14.26	16.45	17.12	19.14
Mancha Khiri	Gumbel	6.94	10.47	12.47	14.16	14.66	16.07
Nong Song Hong	Gumbel	6.91	9.56	11.39	13.2	13.79	15.63
Phu Pha Man	Gumbel	6.45	10.09	12.18	13.97	14.5	16.03
Si Chompu	Frchet	5.49	9.37	12.74	16.72	18.16	23.25

Table 2: The GEV distribution of the food poisoning rate in the elderly people classified according to districts in Maha Sarakham Province, and the return levels in 2, 5, 10, 20, 25 and 50 years.

Districts	Fitted distribution	Return levels (years)					
		2	5	10	20	25	50
Kae Dam	**Gumbel**	**16.21**	**26.03**	**32.63**	**39.05**	**41.1**	**47.47**
Kantharawichai	**Gumbel**	**13.73**	**18.12**	**20.62**	**22.74**	**23.37**	**25.15**
Wapi Prathum	Gumbel	8.86	12.76	15.03	16.99	17.58	19.26
Mueang Maha Sarakham	Gumbel	8.50	11.33	13.00	14.47	14.91	16.19
Phayakkhaphum Phisai	Gumbel	7.62	11.91	14.98	18.11	19.14	22.44
Kut Rang	Gumbel	7.00	10.19	12.11	13.82	14.34	15.85
Na Dun	Gumbel	5.21	8.91	11.77	14.88	15.94	19.49
Chiang Yuen	Gumbel	4.72	7.91	10.28	12.78	13.62	16.35
Chuen Chom	Gumbel	3.72	7.92	11.22	14.84	16.10	20.31
Borabue	Gumbel	3.53	5.72	7.15	8.50	8.93	10.23
Kosum Phisai	Gumbel	2.92	4.97	6.56	8.30	8.90	10.89
Na Chueak	Gumbel	1.79	3.60	4.92	6.28	6.74	8.20

Table 3: The GEV distribution of the food poisoning rate in the elderly people classified according to districts in Roi Et Province, and the return levels in 2, 5, 10, 20, 25 and 50 years.

Districts	Fitted distribution	Return levels (years)					
		2	5	10	20	25	50
Pho Chai	**Frchet**	**40.26**	**63.54**	**81.78**	**101.76**	**108.66**	**131.78**
Si Somdet	**Gumbel**	**27.94**	**39.08**	**45.25**	**50.41**	**51.90**	**56.12**
Pathum Rat	**Gumbel**	**27.79**	**38.08**	**44.69**	**50.87**	**52.81**	**58.67**
Thawat Buri	Gumbel	24.62	37.62	44.94	51.12	52.93	58.07
Nong Phok	Gumbel	23.93	35.32	43.43	51.66	54.37	63.00
Kaset Wisai	Gumbel	23.38	30.43	34.50	38.00	39.03	42.01
Phon Thong	Gumbel	17.44	20.58	22.36	23.86	24.30	25.56
Phon Sai	Gumbel	15.71	22.65	26.86	30.63	31.77	35.14
Mueang Roi Et	Weibull	13.97	18.28	20.80	22.99	23.65	25.54
Mueang Suang	Gumbel	12.02	22.67	31.15	40.56	43.83	54.90
Suwannaphum	Weibull	11.07	14.38	16.11	17.48	17.86	18.91
Selaphum	Gumbel	9.63	14.04	16.81	19.37	20.17	22.55
At Samat	Gumbel	7.43	11.50	14.40	17.34	18.31	21.41
Chaturaphak Phiman	Gumbel	7.36	13.33	18.39	24.29	26.41	33.82
Phanom Phrai	Weibull	5.81	9.04	10.94	12.60	13.09	14.54
Changhan	Gumbel	5.07	10.86	16.37	23.42	26.10	36.04
Thung Khao Luang	Gumbel	4.56	10.75	16.52	23.77	26.50	36.49
Nong Hi	Gumbel	4.47	7.57	9.46	11.16	11.68	13.22
Moei Wadi	Gumbel	2.55	6.61	10.82	16.56	18.83	27.65



(a) **(b)**

(c) **(d)**

(e) **(f)**

Fig. 1: The estimated return levels in (a) 2 years, (b) 5 years, (c) 10 years, (d) 20 years, (e) 25 years, and (f) 50 years with the GEV distribution of the food poisoning rate in the elderly people in 3 provinces

5 Conclusion and discussion

The analysis was performed with the generalized extreme value distribution on the maximum food poisoning rate per 10,000 elderly population, the goodness of fit test was used to examine the appropriateness of the distribution obtained from the analysis, and the Kolmogorov-Smirnov Statistics (KS Test) was used to illustrate the GEV distribution. According to the analysis on the distribution of the maximum food poisoning rate in the elderly people, in Khon Kaen Province, the Gumbel distribution were found in 12 districts, the Weibull distribution in 4 districts, and the Frchet distribution in 2 districts. In Maha Sarakham Province, the Gumbell distribution was found in the distribution of the maximum food poisoning rate in the elderly in all districts. In Roi Et Province, the Gumbel distribution were found in 15 districts, the Weibull distribution in 3 districts, and the Frchet distribution in only 1 district.

Regarding the return levels, the first rank was Pho Chai District in Roi Et Province, the second rank was Ban Fang District in Khon Kaen Province, and the third rank included 3 districts with similar return levels; i.e., Si Somdet District in Roi Et Province, Pathum Rat District in Roi Et Province, and Chonnabot District in Khon Kaen Province. Accordingly, the disease surveillance should be planned to cope with the epidemics as well as to find prevention measures in the risk areas of disease emergencies. Therefore, the application of the EVT and the epidemiological data can form an alternative method to estimate the chances of epidemic emergence.

Acknowledgment. This research project was financially supported by Thailand Science Research and Innovation Fund (TSRI) 2021 and Mahasarakham University. The authors would like to thank the editor and the referees.

References

[1] H. J. Boon, A. Cottrell, D. King, R. B. Stevenson, J. Millar, Bronfenbrenners bioecological theory for modelling community resilience to natural disasters, Nat Hazards, **60,** no.2, (2012), 381–408.

[2] K. L. Ebi, D. M. Mills, J. B. Smith, A. Grambsch, Climate change and human health impacts in the United States: an update on the results of the U.S. national assessment, Environ. Health Perspect., **114,** no. 9, (2006), 1318–1324.

[3] J. Chen, X. Lei, L. Zhang, B. Peng, Using Extreme Value Theory Approaches to Forecast the Probability of Outbreak of Highly Pathogenic Influenza in Zhejiang, China, PLoS ONE, **10,** no. 2, (2015), e0118521. doi:10.1371/journal.pone.0118521.

[4] M. Thomas, M. Lemaitre, M. L. Wilson, C. Viboud, Y. Yordanov, H. Wackernagel, F. Carratet, Applications of extreme value theory in public health, PLoS ONE, **11,** no. 7, (2016), e0159312.https://doi.org/10.1371/journal.pone.0159312.

[5] S. Gong, Estimation of hot and cold spells with extreme value theory, U. U. D. M. Project Report, (2012).

[6] S. Coles, S. Nadaraja, An Introduction to Statistical Modeling of Extreme Values, Great Britain: Springer-Verlag London, 2001.

[7] J. Zhang, Powerful goodness-of-fit- tests based on the likelihood ratio, Journal of the Royal Statistical Society, Series B, (Statistical Methodology), **64,** no. 2, (2002), 281–294.