Scopus

Documents

Jayaramu, V.[a] , Zulkafli, Z.[a] , De Stercke, S.[b] , Buytaert, W.[b] , Rahmat, F.[c] , Abdul Rahman, R.Z.[c] , Ishak, A.J.[c] , Tahir, W.[d] , Ab Rahman, J.[e] , Mohd Fuzi, N.M.H.[f]

Leptospirosis modelling using hydrometeorological indices and random forest machine learning
(2023) *International Journal of Biometeorology*, 67 (3), pp. 423-437.

DOI: 10.1007/s00484-022-02422-y

[a] Department of Civil Engineering, Universiti Putra Malaysia, Serdang, Malaysia
[b] Department of Civil and Environmental Engineering, Imperial College London, London, United Kingdom
[c] Department of Electrical and Electronic Engineering, Universiti Putra Malaysia, Serdang, Malaysia
[d] Flood Control Research Group, Faculty of Civil Engineering, Universiti Teknologi Mara, Shah Alam, Malaysia
[e] Department of Community Medicine, Kulliyyah of Medicine, International Islamic University Malaysia, Kuantan, Malaysia
[f] Kelantan State Health Department, Ministry of Health Malaysia, Kelantan, Kota Bharu, Malaysia

Abstract
Leptospirosis is a zoonosis that has been linked to hydrometeorological variability. Hydrometeorological averages and extremes have been used before as drivers in the statistical prediction of disease. However, their importance and predictive capacity are still little known. In this study, the use of a random forest classifier was explored to analyze the relative importance of hydrometeorological indices in developing the leptospirosis model and to evaluate the performance of models based on the type of indices used, using case data from three districts in Kelantan, Malaysia, that experience annual monsoonal rainfall and flooding. First, hydrometeorological data including rainfall, streamflow, water level, relative humidity, and temperature were transformed into 164 weekly average and extreme indices in accordance with the Expert Team on Climate Change Detection and Indices (ETCCDI). Then, weekly case occurrences were classified into binary classes “high” and “low” based on an average threshold. Seventeen models based on “average,” “extreme,” and “mixed” indices were trained by optimizing the feature subsets based on the model computed mean decrease Gini (MDG) scores. The variable importance was assessed through cross-correlation analysis and the MDG score. The average and extreme models showed similar prediction accuracy ranges (61.5–76.1% and 72.3–77.0%) while the mixed models showed an improvement (71.7–82.6% prediction accuracy). An extreme model was the most sensitive while an average model was the most specific. The time lag associated with the driving indices agreed with the seasonality of the monsoon. The rainfall variable (extreme) was the most important in classifying the leptospirosis occurrence while streamflow was the least important despite showing higher correlations with leptospirosis.

Author Keywords
Cross-correlation analysis; Feature selection; Hydrometeorological indices; Leptospirosis; Random forest; Variable importance

Index Keywords
car driving, human, leptospirosis, random forest, season, temperature; Automobile Driving, Humans, Leptospirosis, Random Forest, Seasons, Temperature

References

- Adler, B., de la Peña Moctezuma, A.
Leptospira and leptospirosis
(2010) *Vet Microbiol*, 140 (3-4), pp. 287-296.
(,),., (,):., https://doi.org/10.1016/j.vetmic.2009.03.012

- Ahangarcani, M., Farnaghi, M., Shirzadi, M.R., Pilesjö, P., Mansourian, A.
Predictive risk mapping of human leptospirosis using support vector machine classification and multilayer perceptron neural network
(2019) *Geospatial Health*, 14 (1).
(,),., (,)., https://doi.org/10.4081/gh.2019.711

- (2017) *Chapter 23 - Leptospirosis*, pp. 336-344.
Elsevier Inc., In The Travel and Tropical Medicine Manual (Fifth Edition)

- **Random forest and deep learning performance on the Malaria DREAM sub challenge one random forest and deep learning performance**

(2020) *Res Comput Sci*, 149 (5), pp. 163-170.

- Benacer, D., Thong, K.L., Min, N.C., Verasahib, K.B., Galloway, R.L., Hartskeerl, R.A., Souris, M., Zain, S.N.M.
  Epidemiology of human leptospirosis in Malaysia, 2004–2012
  (2016) *Acta Tropica*, 157, pp. 162-168.

- Barcellos, C., Sabroza, P.C.
  The place behind the case: leptospirosis risks and associated environmental conditions in a flood-related outbreak in Rio de Janeiro
  (2001) *Cad Saúde Pública / Ministério Da Saúde, Fundação Oswaldo Cruz, Escola Nacional De Saúde Pública*, 17, pp. 59-67.

- Batchelor, T.W.K., Stephenson, T.S., Brown, P.D., Amarakoon, D., Taylor, M.A.
  Influence of climate variability on human leptospirosis cases in Jamaica
  (2012) *Climate Res*, 55 (1), pp. 79-90.

- **Random forests**
  (2001) *Mach Learn*, 45 (1), pp. 5-32.

- Campbell, A.M., Racault, M.F., Goult, S., Laurenson, A.
  Cholera risk: a machine learning approach applied to essential climate variables
  (2020) *Int J Environ Res Public Health*, 17 (24), pp. 1-24.

- Cann, K.F., Thomas, D.R., Salmon, R.L., Wyn-Jones, A.P., Kay, D.
  Extreme water-related weather events and waterborne disease
  (2013) *Epidemiol Infect*, 141 (4), pp. 671-686.

- Carvajal, T.M., Viacrusis, K.M., Hernandez, L.F.T., Ho, H.T., Amalin, D.M., Watanabe, K.
  Machine learning methods reveal the temporal pattern of dengue incidence using meteorological factors in metropolitan Manila Philippines
  (2018) *BMC Infect Dis*, 18 (1), pp. 1-15.

- Chadsuthi, S., Modchang, C., Lenbury, Y., Iamsirithaworn, S., Triampo, W.
  Modeling seasonal leptospirosis transmission and its association with rainfall and temperature in Thailand using time-series and ARIMAX analyses
  (2012) *Asian Pac J Trop Med*, 5 (7), pp. 539-546.

- Cunha, M., Ribeiro, G.S., Carvalho, M.S., Reis, R.B., Pischel, L., Gouveia, E.L., Santos, A.C., Ko, A.I.
  Rainfall and other meteorological factors as drivers of urban transmission of leptospirosis
  (2022) *Plos Negl Trop Dis*, 16 (4).

- (2017),
  [DID], Flood management - programme and activities, Accessed 27 Nov 2020

- Desvars, A., Jégo, S., Chiroleu, F., Bourhy, P., Cardinale, E., Michault, A.
  Seasonality of human leptospirosis in Reunion Island (Indian Ocean) and its association with meteorological data
  (2011) *PLoS ONE*, 6 (5).

- Dhewantara, P.W., Lau, C.L., Allan, K.J., Hu, W., Zhang, W., Mamun, A.A., Soares Magalhães, R.J.
  Spatial epidemiological approaches to inform leptospirosis surveillance and control: A systematic review and critical appraisal of methods
  (2019) *Zoonoses and Public Health*, 66 (2), pp. 185-206.
  (,),., (,):., https://doi.org/10.1111/zph.12549

- Ding, G., Li, X., Li, X., Zhang, B., Jiang, B., Li, D., Xing, W., Hou, H.
  A time-trend ecological study for identifying flood-sensitive infectious diseases in

Guangxi, China from 2005 to 2012
(2019) *Environ Res*, 176 (July), p. 108577.

- Ehelepola, N.D.B., Ariyaratne, K., Dissanayake, W.P.
The correlation between local weather and leptospirosis incidence in Kandy district, Sri Lanka from 2006 to 2015
(2019) *Global Health Action*, 12 (1), p. 1553283.

- Ghizzo Filho, J., Nazário, N.O., Freitas, P.F., Pinto, G.D.A., Schlindwein, A.D.
Temporal analysis of the relationship between leptospirosis, rainfall levels and seasonality, Santa Catarina, Brazil, 2005–2015
(2018) *Rev Inst Med Trop Sao Paulo*, 3154 (1), pp. 17-18.

- Glaros, A.G., Kline, R.B.
Understanding the accuracy of tests with cutting scores: The sensitivity, specificity, and predictive value model
(1988) *J Clin Psychol*, 44 (6), pp. 1013-1023.
(,),., (,):., https://doi.org/10.1002/1097-4679(198811)44:6%3C1013::AID-JCLP2270440627%3E3.0.CO;2-Z

- Gómez, A.A., López, M.S., Müller, G.V., López, L.R., Sione, W., Giovanini, L.
Modeling of leptospirosis outbreaks in relation to hydroclimatic variables in the northeast of Argentina
(2022) *Heliyon*, 8 (6).
(,),., (,):,., https://doi.org/10.1016/j.heliyon.2022.e09758

- Guo, P., Liu, T., Zhang, Q., Wang, L., Xiao, J., Zhang, Q., Luo, G., Ma, W.
Developing a dengue forecast model using machine learning: a case study in China
(2017) *PLoS Negl Trop Dis*, 11 (10).

- Haake, D.A., Levett, P.N.
Leptospirosis in humans
(2015) *Curr Top Microbiol Immunol*, 387, pp. 65-97.

- Hacker, K.P., Sacramento, G.A., Cruz, J.S., De Oliveira, D., Nery, N., Lindow, J.C., Carvalho, M., Costa, F.
Influence of rainfall on leptospira infection and disease in a tropical urban setting Brazil
(2020) *Emerg Infect Dis*, 26 (2), pp. 311-314.

- Hayati, K.S., Sharifah Norkhadijah, S.I., Salmiah, M.S., Edre, M.A., Khin, T.D.
Hot-spot and cluster analysis on legal and illegal dumping sites as the contributors of leptospirosis in a flood hazard area in Pahang, Malaysia
(2018) *Asian J Agric Biol*, 5 (2), pp. 56-59.

- Hu, H., Wang, H., Wang, F., Langley, D., Avram, A., Liu, M.
Prediction of influenza-like illness based on the improved artificial tree algorithm and artificial neural network
(2018) *Sci Rep*, 8 (1), pp. 1-8.

- Ismail, W.R., Haghroosta, T.
Extreme weather and floods in Kelantan state, Malaysia in December 2014
(2018) *Res Mar Sci*, 3 (1), pp. 231-244.

- Jamaludin, N., Mohammed, N.I., Khamidi, M.F., Wahab, S.N.A.
Thermal comfort of residential building in Malaysia at different micro-climates
(2015) *Proc Soc Behav Sci*, 170, pp. 613-623.

- Joshi, Y.P., Kim, E.H., Cheong, H.K.
The influence of climatic factors on the development of hemorrhagic fever with renal syndrome and leptospirosis during the peak season in Korea: an ecologic

study
(2017) *BMC Infect Dis*, 17 (1).
(,),. (,)., https://doi.org/10.1186/s12879-017-2506-6

- Khan, S., Ullah, R., Khan, A., Sohail, A., Wahab, N., Bilal, M., Ahmed, M.
Random forest-based evaluation of Raman spectroscopy for dengue fever analysis
(2017) *Appl Spectrosc*, 71 (9), pp. 2111-2117.

- Kuhn, M.
Building predictive models in R using the caret package
(2008) *J Stat Soft*, 28 (5), pp. 1-26.

- Kumar, V., Minz, S.
Feature selection: A literature review
(2014) *Smart Comput Rev*, 4 (3), pp. 211-229.

- Kupek, E., de Sousa Santos Faversani, M.C., de Souza Philippi, J.M.
The relationship between rainfall and human leptospirosis in Florianópolis, Brazil, 1991–1996
(2000) *Braz J Infect Dis: Off Publ Braz Soc Infect Dis*, 4 (3), pp. 131-134.
COI: 1:STN:280:DC%2BD3M%2FhsFWrsw%3D%3D

- Lau, C.L., Clements, A.C.A., Skelly, C., Dobson, A.J., Smythe, L.D., Weinstein, P.
Leptospirosis in American Samoa - estimating and mapping risk using environmental data
(2012) *PLoS Negl Trop Dis*, 6 (5).

- Lau, C.L., Smythe, L.D., Craig, S.B., Weinstein, P.
Climate change, flooding, urbanisation and leptospirosis: fuelling the fire?
(2010) *Trans R Soc Trop Med Hyg*, 104 (10), pp. 631-638.

- Lipovetsky, S., Conklin, M.
Analysis of regression in game theory approach
(2001) *Appl Stoch Model Bus Ind*, 17 (4), pp. 319-330.

- López, M.S., Müller, G.V., Lovino, M.A., Gómez, A.A., Sione, W.F., Aragonés Pomares, L.
Spatio-temporal analysis of leptospirosis incidence and its relationship with hydroclimatic indicators in northeastern Argentina
(2019) *Sci Total Environ*, 694.
(,),., https://doi.org/10.1016/j.scitotenv.2019.133651

- Mayfield, H.J., Lowry, J.H., Watson, C.H., Kama, M., Nilles, E.J., Lau, C.L.
Use of geographically weighted logistic regression to quantify spatial variation in the environmental and sociodemographic drivers of leptospirosis in Fiji: a modelling study
(2018) *Lancet Planet Health*, 2 (5), pp. e223-e232.

- Mohammadinia, A., Alimohammadi, A., Saeidian, B.
Efficiency of geographically weighted regression in modeling human leptospirosis based on environmental factors in Gilan province Iran
(2017) *Geosci*, 7 (4), p. 136.

- Mohd Radi, M.F., Hashim, J.H., Jaafar, M.H., Hod, R., Ahmad, N., Nawi, A.M., Baloch, G.M., Ayub, N.I.F.
Leptospirosis outbreak after the 2014 major flooding event in Kelantan, Malaysia: A spatial-temporal analysis
(2018) *Am J Trop Med Hyg*, 98 (5), pp. 1281-1295.

- Péres, W.E., Russo, A., Nunes, B.
The association between hydro-meteorological events and leptospirosis

hospitalizations in Santa Catarina Brazil
(2019) *Water*, 11 (5), p. 1052.

- Peterson, T.C., Folland, C.C., Gruza, G., Hogg, W., Mokssit, A., Plummer, N.
Report on the activities of the working group on climate change detection and related rapporteurs 1998–2001
(2001) *Rep. WCDMP-47, WMO-TD 1071, Geneve, Switzerland, March*, p. 143..
Accessed 10 Oct 2022

- Picardeau, M.
Diagnosis and epidemiology of leptospirosis
(2013) *Med Et Mal Infect*, 43 (1), pp. 1-9.

- Rahayu, S., Adi, M.S., Saraswati, L.D.
Mapping of leptospirosis environmental risk factors and determining the level of leptospirosis vulnerable zone in Demak District using remote sensing image
(2018) *In: E3S Web of Conferences 2018.*, 31, p. p 06003.

- Rahmat, F., Ishak, A.J., Zulkafli, Z., Yahaya, H., Masrani, A.
Prediction model of leptospirosis occurrence for Seremban (Malaysia) using meteorological data
(2019) *Int J Integr Eng*, 11 (4), pp. 61-69.

- Rahmat, F., Zulkafli, Z., Ishak, A.J., Mohd Noor, S.B., Yahaya, H., Masrani, A.
Exploratory data analysis and artificial neural network for prediction of leptospirosis occurrence in Seremban, Malaysia Based on Meteorological Data
(2020) *Front Earth Sci*, 8, p. 377.

- Sánchez-Montes, S., Espinosa-Martínez, D.V., Ríos-Muñoz, C.A., Berzunza-Cruz, M., Becker, I.
Leptospirosis in Mexico: epidemiology and potential distribution of human cases
(2015) *PLoS ONE*, 10 (7).

- Santos, M.S., Soares, J.P., Abreu, P.H., Araujo, H., Santos, J.
Cross-validation for imbalanced datasets: avoiding overoptimistic and overfitting approaches [research frontier]
(2018) *IEEE Comput Intell Mag*, 13 (4), pp. 59-76.
(,),., (,):., https://doi.org/10.1109/MCI.2018.2866730

- Schneider, M.C., Nájera, P., Aldighieri, S., Bacallao, J., Soto, A., Marquiño, W., Altamirano, L., Espinal, M.
Leptospirosis outbreaks in Nicaragua: identifying critical areas and exploring drivers for evidence-based planning
(2012) *Int J Environ Res Public Health*, 9 (11), pp. 3883-3910.

- Schober, P., Boer, C., Schwarte, L.A.
Correlation coefficients: appropriate use and interpretation
(2018) *Anesth Analg*, 126 (5), pp. 1763-1768.

- Sehgal, S.C., Sugunan, A.P., Vijayachari, P.
Outbreak of leptospirosis after the cyclone in Orissa
(2002) *Natl Med J India*, 15 (1), pp. 22-23.
COI: 1:STN:280:DC%2BD387itFOntQ%3D%3D

- Soo, Z.M.P., Khan, N.A., Siddiqui, R.
Leptospirosis: increasing importance in developing countries
(2020) *Acta Tropica*, 201.
(,),.,:,., https://doi.org/10.1016/j.actatropica.2019.105183

- Strobl, C., Boulesteix, A.L., Zeileis, A., Hothorn, T.
Bias in random forest variable importance measures: illustrations, sources and a

solution
(2007) *BMC Bioinforma*, 8 (1), pp. 1-21.

- Sumi, A., Telan, E.F.O., Chagan-Yasutan, H., Piolo, M.B., Hattori, T., Kobayashi, N.
Effect of temperature, relative humidity and rainfall on dengue fever and leptospirosis infections in Manila, the Philippines
(2017) *Epidemiol Infect*, 145 (1), pp. 78-86.

- Suwanpakdee, S., Kaewkungwal, J., White, L.J., Asensio, N., Ratanakorn, P., Singhasivanon, P., Day, N.P.J., Pan-Ngum, W.
Spatio-temporal patterns of leptospirosis in Thailand: is flooding a risk factor?
(2015) *Epidemiol Infect*, 143 (10), pp. 2106-2115.

- Tassinari, W.S., Pellegrini, D.C., Sá, C., Reis, R.B., Ko, A.I., Carvalho, M.S.
Detection and modelling of case clusters for urban leptospirosis
(2008) *Tropical Med Int Health*, 13 (4), pp. 503-512.
(,),., (,):., https://doi.org/10.1111/j.1365-3156.2008.02028.x

- Togami, E., Kama, M., Goarant, C., Craig, S.B., Lau, C., Ritter, J.M., Imrie, A., Nilles, E.J.
A large leptospirosis outbreak following successive severe floods in Fiji, 2012
(2018) *American Journal of Tropical Medicine and Hygiene*, 99 (4), pp. 849-851.

- Ucar, M.K., Nour, M., Sindi, H., Polat, K.
The effect of training and testing process on machine learning in biomedical datasets
(2020) *Math Prob Eng*, 2020, p. 17.

- Uusitalo, R., Siljander, M., Dub, T., Sane, J., Sormunen, J.J., Pellikka, P., Vapalahti, O.
Modelling habitat suitability for occurrence of human tick-borne encephalitis (TBE) cases in Finland
(2020) *Ticks and Tick-borne Diseases*, 11 (5).
(,),., (,):,., https://doi.org/10.1016/j.ttbdis.2020.101457

- Van Stralen, K.J., Stel, V.S., Reitsma, J.B., Dekker, F.W., Zoccali, C., Jager, K.J.
Diagnostic methods I: sensitivity, specificity, and other measures of accuracy
(2009) *Kidney Int*, 75 (12), pp. 1257-1263.

- Vega-Corredor, M., Opadeyi, J.
Hydrology and public health: Linking human leptospirosis and local hydrological dynamics in Trinidad, West Indies
(2014) *Earth Perspectives*, 1, pp. 1-4.

- Weinberger, D., Baroux, N., Grangeon, J.P., Ko, A.I., Goarant, C.
El Nino southern oscillation and leptospirosis outbreaks in New Caledonia
(2014) *PLoS Negl Trop Dis*, 8 (4).
(,),., (,):,., https://doi.org/10.1371/journal.pntd.0002798

- **WHO recommended strategies for the prevention and control of communicable diseases (No. WHO/CDS/CPE/SMT/2001.13)**
(2001) *World Health Organization*,
Accessed 22 May 2021

- (2003),
Human leptospirosis: guidance for diagnosis, surveillance and control (No. WHO/CDS/CSR/EPH 2002.23). World Health Organization., Accessed 17 Aug 2021

- (2011) *Report of the Second Meeting of the Leptospirosis Burden Epidemiology Reference Group.*,
World Health Organization, Accessed 22 Aug 2021

- Zakharova, O.I., Korennoy, F.I., Iashin, I.V., Toropova, N.N., Gogin, A.E., Kolbasov, D.V., Surkova, G.V., Blokhin, A.A.
Ecological and Socio-economic determinants of livestock animal leptospirosis in the Russian arctic
(2021) *Front Vet Sci*, 8, p. 658675.

- Zhang, X., Alexander, L., Hegerl, G.C., Jones, P., Tank, A.K., Peterson, T.C., Trewin, B., Zwiers, F.W.
Indices for monitoring changes in extremes based on daily temperature and precipitation data
(2011) *Wiley Interdiscip Rev: Clim Chang*, 2 (6), pp. 851-870.

- Zhang, Z., Yang, Z., Ren, W., Wen, G.
Random forest-based real-time defect detection of Al alloy in robotic arc welding using optical spectrum
(2019) *J Manuf Process*, 42, pp. 51-59.

- Zhao, J., Liao, J., Huang, X., Zhao, J., Wang, Y., Ren, J., Wang, X., Ding, F.
Mapping risk of leptospirosis in China using environmental and socioeconomic data
(2016) *BMC Infect Dis*, 16 (1), pp. 1-10.

- Zhao, N., Charland, K., Carabali, M., Nsoesie, E.O., Maheu-Giroux, M., Rees, E., Yuan, M., Zinszer, K.
Machine learning and dengue forecasting: Comparing random forest and artificial neural networks for predicting dengue burden at national and sub-national scales in Colombia
(2020) *Plos Negl Trop Dis*, 14 (9).

Correspondence Address
Zulkafli Z.; Department of Civil Engineering, Malaysia; email: zeddiyana@upm.edu.my

Publisher: Springer Science and Business Media Deutschland GmbH

ISSN: 00207128
CODEN: IJBMA
PubMed ID: 36719482
Language of Original Document: English
Abbreviated Source Title: Int. J. Biometeorol.
2-s2.0-85147099122
Document Type: Article
Publication Stage: Final
Source: Scopus