

Cholera seasonality, rainfall, and fadeouts: a geostatistical approach

Abstract: The seasonality of cholera and its spatial variability remain unexplained. Uncovering the role of environmental drivers in these seasonal patterns is critical for understanding temporal variability over longer time scales, including trends and variability between years. Rainfall has been proposed as a key driver of the seasonality of cholera. To address this hypothesis, we examine the association between rainfall and cholera is both time and space.

We show that historical data for the districts of former British India (1900-1940) cluster into two main regions that differ not just in the effect of rainfall but also in their seasonal pattern and frequency of fadeouts. Results support a previously proposed model of cholera seasonality with two different routes of transmission and a dual nature of rainfall, with both positive and negative effects. We define “endemic” and “epidemic” cholera regions based on these findings.

Introduction

The seasonality of cholera remains a mystery, it exhibits robust regularity but with important geographic variation. Furthermore, its environmental drivers are poorly defined because the seasonal pattern varies in space, with different lags relative to rainfall, one major potential driver. For example, two peaks per year are the typical pattern described for cholera in Bangladesh and former Bengal, with a decline in the summer during the monsoons, but only one peak is present in other regions of former British India and current Brazil, which coincides with the rainy season (see review by Pascual et al., 2002; and Codeço, 2001).

A better understanding of cholera’s seasonality is key to identify the regional mechanisms behind the described effect of the El Niño Southern Oscillation (ENSO) (Pascual et al., 2000; Koelle and Pascual, 2004). It is also fundamental to build scenarios

for cholera with global change. Both climate variability (ENSO) and climate change are likely to act on infectious diseases through the modulation of the seasonal cycle and the crossing of environmental thresholds (Pascual and Dobson, 2005).

Studies in volunteers confirmed that the ingestion of a dosage between 10^7 and 10^{11} bacteria, depending on the method (for example, neutralizing stomach acidity, or with different foods), develop in an infection (Kaper et al., 1995); consequently, bacteria density is considered a major indicator of potential outbreaks. Moreover, the existence of two possible routes of transmission for cholera includes rainfall in the seasonal cycle of cholera cases. Primary transmission presumably occurs from a reservoir of the pathogen *Vibrio cholerae* in the aquatic environment. Brackish water and estuaries, has been shown to provide adequate environmental conditions for the bacterium to survive outside the human host (Colwell et al., 1977). Secondary or “human-to-human” transmission occurs via the ingestion of fecally contaminated water, and some times through food (Glass et al., 1991). The relative importance of these two routes of transmission is highly debated; however, if a strong feedback exists from infected hosts to the natural environments, the distinction between the two is blurred.

Based on these two transmission routes, Dobson et al. (in prep.) have previously proposed the following mechanisms behind the bimodal seasonal pattern of cholera in Bangladesh. The first peak occurs in the spring, during the dry season when temperature warms up, because the bacterium thrives in the environment where it is also more highly concentrated, and humans’ interaction with water bodies increases (i.e. primary transmission). The monsoon leads to a decline in cholera in the summer, as heavy rainfall dilutes the concentration of the pathogen in the environment and both salinity and pH, favorable conditions for the bacteria, decay. This dilution effect of rainfall represents a

negative influence. However, a positive effect would follow with an increase in cholera as humans concentrate in the flooded landscape and existing sanitary conditions break down (i.e. secondary transmission).

This seasonal model provides predictions for both endemic and epidemic areas. In endemic regions, cholera should exhibit a negative association with rainfall at zero lag (dilution effect) and a positive correlation at positive lags reflecting the increase in secondary transmission after the rains. However, in regions with long and sustained periods of rainfall, and consequently with low concentration of the pathogen in the aquatic reservoirs, there should be an increase in the local extinction of the disease. Hence, frequent fade-outs and an erratic behavior should be favored in regions with lower human populations.

To examine these predictions, we analyze the association of cholera and rainfall in space and time, investigating also the notion of a Critical Community Size for cholera (Keeling, 1997).

Methods and data

The predictions conceived from the hypothesis about the temporal dynamics of primary and secondary transmission can be evaluated studying historical records of cholera disease.

The area studied was the region of India known as Madras Presidency. The Madras Presidency was a former province of the British Empire that included several districts in the southern region of India between latitudes 20° and 8°N and longitudes 74° and 86°E. We digitized the historical maps of the province and its 26 administrative sub-province divisions or districts, where the included historical districts were: (1) Annapur,

(2) Bellary, (3) Chingleput, (4) Chittoor, (5) Cuddapah, (6) Ganjam, (7) Godivari East, (8) Godivari West, (9) Guntur, (10) Kistna, (11) Kurnool, (12) neighborhood of Madras, the former capital city, (13) Malabar, (14) Nellore, (15) Nilgiris, (16) North Arcot, (17) Ramnad, (18) Salem, (19) South Arcot, (20) South Kanara, (21) Tanjore, (22) Tinnevely, (23) Vizagapatam, (24) Trichinopoly, (25) Coimbatore and (26) Madua (Fig. 1).

For each district, monthly cholera mortality data were collected from January 1892 until December 1940 and, during the same period, information about population size for some of the districts was obtained. In addition, from January 1901 to December 1970, several meteorological stations located in the region collected daily rainfall data. A monthly estimation of rainfall during that period was calculated averaging the stations and taking in consideration the historical borders of the districts. Because of the partial overlapping, analyses considering both data only included the period from January 1901 to December 1940. On the other hand, whenever cholera mortality or rainfall was analyzed independently from each other, the full time series was used (1892-1940 and 1901-1970 respectively).

A spatial correlogram (Bailey and Gatrell, 1995, Fortin et al., 2002) was performed to detect up to which amplitude, or distance, the area of influence of each district extends when cholera mortality is the variable under study. With the distance obtained (200 Km.), proximity matrices, that determinate which districts are considered as neighbors, were defined. Once that a neighborhood was defined spatial autocorrelation analyses were performed considering both cholera and rainfall dataset. The well-known (global) spatial autocorrelation index, Moran's I (Cliff and Ord, 1973), and its modified version, which incorporated the computation for *local* variations, Local Indicators of

Spatial Association (LISA) (Anselin, 1995, Fortin et al., 2002), was obtained for the mean values of both time series.

In order to study the interactions between cholera and rainfall the coefficient of correlation between the time series was calculated including a lag between the series ranging from 0 to 12 months (with rainfall preceding cholera). Then, both global and local spatial autocorrelation indexes, using the coefficients of correlation as input, were calculated. Whereas Moran's I evaluates the degree of global clustering, LISA index permits local variations and hence the identification of “hot-spots” and “cold-spots” or clusters where high and low values for the interaction cholera-rainfall were observed is possible.

Spatial autocorrelation shows a static picture of the distribution of cholera, however, dynamical measures are necessary to have a complete perspective. Two of such measures were considered here: the Critical Community Size and the seasonal variability of rainfall. The Critical Community Size (CCS), the population size below which a disease dies out in the troughs between epidemics, reflects the dynamics underlying outbreaks (Keeling, 1997). Hence, a qualitative measure of the CCS was obtained for the districts where population data was available. For this analysis, a period of at least two consecutive months without mortality was considered a *die out* of the disease, known as fadeout. In addition, the spatial autocorrelation of the amount of fade-outs also was analyzed. Finally, the incorporation of the dynamic pattern manifested by rainfall fulfills the previous analysis. All the districts exhibited a strong peak in rainfall due to the so-called southwest monsoon (for a detailed explanation see Krishnamurthy and Kinter III, 2002). Some districts exhibit high precipitation values can last for even 6 months, and

after such a long rain season the rainfall diminish to irrelevant levels during the rest of the year, whereas in others districts the monsoon effect is typically shorter (3 or 4 months) and a second peak of rainfall occurs within an interval of few months. Exceptions to this second peak are the districts on the Arabian Sea. Determining how many times and for how long the monthly-accumulated rainfall exceeded a threshold value—fixed at a value of 25% over the expected rainfall per district—allowed to discriminate districts with two rainfall peaks per year from those with only one. This characterization of districts was later analyzed considering the variability of cholera observed in conjunction with the rainfall patterns.

Results

The methods described in the previous section provide evidence to describe the variability manifested in rainfall, mortality due to cholera, and their interconnection in light of the situations expected by our hypotheses.

The low values obtained for the spatial autocorrelation, measured by the global index Moran's I, indicate that there is not significant spatial clustering when both cholera mortality and rainfall are independently considered; however, when Moran's I is evaluated for the correlation between them, the index reaches significance (Table 1). Hence, a notable spatial association emerged when the interconnection between cholera and rainfall was explored.

Considering the locality for cholera mortality, the results provided by the LISA index delimited one small-size cluster (districts number 17, 24 and 26) with high values in the central southern part of the region under analysis, and a second cluster close to the

first, also small in size (districts 15 and 18) and with low values compared to that in the high-value cluster. In contrast, for rainfall, a medium-size cluster (districts number 4, 5, 11 and 14) of low values appears in the central northern region. As well as the global clustering changes when the correlation between cholera and rainfall is analyzed, the pattern exhibited locally is also dramatically different (Fig. 2): one big-size cluster with high values is formed in the northeast region (districts 7, 8, 9, 10, 11 and 23, denoted as HIGH-HIGH in the figure); a second big-size cluster with low values is found in the central southern area (districts 15, 16, 17, 18, 19, 21, 24, 25 and 26, LOW-LOW); and three one-district clusters with relative low values, respecting their neighborhood, are also delimited (districts 12, 13 and 20). This pattern could be read as follows: In the northeast area, the correlation between rainfall and cholera without considering any temporal lag between the series is reach positive significant values, whereas in the central southern area the values are negative. When the local clustering evaluated with LISA index is studied introducing time lags between cholera and rainfall time series, with rainfall preceding cholera, a similar pattern is obtained when the lag is 10, 11 and 12 months, whereas the pattern is inverted—negative correlation in the northeast region and positive correlation in the central southern area— for values ranging from 3 to 7, and finally the pattern is not clear for 1, 2, 8 and 9 months (Fig. 3 shows some of the lags).

Considering the dynamic aspects of this study, the CSS plotted in Figure 4 shows that districts with higher density have a lower number of fadeouts, while in the less dense districts the presence of the disease is frequently interrupted. It is very important to notice the high variability present in the dataset, but in these low populated districts, cholera can be seen as a disease with irregular outbreaks, whereas in the former districts cholera exhibits an endemic behavior. The Moran's I value considering the amount of fadeouts

for each district, Table 1, indicates that there is not spatial clustering. Local clustering analysis considering the amount of fadeouts does not reach statistical significance; however, mapping the amount of fadeouts more fadeouts occurs in the northern districts (Fig. 5). A similar geographical pattern is obtained when the duration of fadeouts is considered, with the northern districts having longer fadeouts than the southern ones.

Finally, a classification considering the mean annual duration of the rain season determined three different regions (Fig. 5): a northern area with long rain season (districts 6, 7, 8, 9, 10, 13, 15, 20 and 23); a Central area with a moderate length of rainfall (districts 1, 2, 3, 4, 5, 11, 12, 14, 16, 19 and 21); and a southern region with two short wet season (districts 17, 18, 22, 24, 25 and 26).

Discussion

The results obtained in this work allow splitting the Madras Presidency into two main regions, Northeastern and Southern, with different cholera seasonality and different patterns of association between cholera and rainfall. In particular, the southern region exhibits a pattern which is similar with the one described in the literature for Bangladesh (Bouma and Pascual, 2001), with the whole seasonal pattern shifted in time in accordance with the earlier dominant monsoon season. This seasonal pattern is then characteristic of “endemic” regions, with regular and persistent infection, and contrasts with the stochastic nature of “epidemic” regions with only one sporadic peak coincident with rains and recurrent fadeouts.

The northeastern region comprehends the following districts: Godivari East (7), Godivari West (8), Guntur (9), Kistna (10), Kurnool (11), and Vizgapatam (23). In this area, a positive correlation between cholera and rainfall is shown and in general, one

epidemic peak of cholera mortality per year is presented. Moreover, these districts display longer and more frequent fadeouts, suggesting a non-permanent presence of the disease.

The southern region including the districts of Nilgiris (15), North Arcot (16), Ramnad (17), Salem (18), South Arcot (19), Trichinopoly (24), Coimbatore (25), and Madua (26). In this region, a negative association between cholera and rainfall is observed and two peaks of cholera mortality are commonly shown. Also in this region, shorter and infrequent (sporadic) fadeouts are observed, probably indicating a permanent presence of the disease.

The remaining districts (numbers 1, 2, 3, 4, 5, 6, 12, 13, 14, 20, 21, 22 and 23 in Fig. 1), located in the central and west regions exhibited an intermediate behavior between the ones presented in the two regions described above.

The bimodal Southern pattern is consistent with the predictions of our seasonal model (Dobson et al., in prep). This result provides the first quantitative evidence for both a positive and negative influence of rainfall on the seasonality of cholera, through its dilution effect on the pathogen and its enhancing effect of “human-to-human” or secondary transmission. While the first effect has been described in the literature, the second is novel for the bimodal cholera pattern.

The unimodal monsoonal Northern pattern combined with its stochastic nature suggests that in places where secondary transmission cannot be sustained over time, an environmental reservoir of pathogenic *Vibrio cholerae* is not effectively maintained. Thus this human feedback from infected individuals to aquatic reservoirs appears critical to sustain the so-called primary transmission, which underscores the importance of secondary transmission itself. Sustained periods of rain further dilute the pathogen in

aquatic reservoirs and epidemics occur during the rainy season, presumably through immigration of infected individuals and the consequent secondary route of transmission.

Traditionally, the explanation for the occurrence of outbreaks after rainfall periods only considers the environment driven infection processes and basically states that moderate rainfalls are able to create optimal (or quasi optimal) environmental conditions for the spread of *Vibrio cholerae* (Lipp et al., 2002). The findings presented here supported the hypothesis that a long wet season is able to create a dilution effect avoiding the bacteria to anchor in the environment (Pascual et. al, 2002), but due to the aggregation of people, probably because of floods, which disproportionally raises the human-to-human oral-fecal mechanism of contagion and drastically reduces the quality of sanitary conditions, the disease is able to settle when is introduced from an endemic area. Consequently, rainfall plays two different roles concerning cholera dynamics: short wet seasons favor cholera leading to an environment driven disease, whereas long wet seasons entail an infectious dynamics.

In conclusion, the dynamics of cholera can be defined by two very different regions inside the Madras Presidency: the southern endemic region and the northeast region that exhibits an epidemic behavior. A possible explanation for the re-emergence of the disease in epidemic zones can be the migration of infective agents (food, water or people) from the endemic zones. Although more information is necessary to evaluate explicitly such scenario, our findings may be used to design health care policies in order to avoid massive mortality situations.

References

Anselin, L. (1995) "Local indicators of spatial association - LISA," Geographical

- Analysis 27:93–115.
- Bailey, T. C., and Gatrell, A. C. (1995) “Interactive spatial data analysis,” Pearson Education Limited, Harlow, England.
- Bouma, M. J., and Pascual, M. (2001) “Seasonal and interannual cycles of epidemic cholera in Bengal 1891-1940 in relation to climate and geography,” *Hidrobiologia* 460:147-156.
- Cliff, A. D., and Ord, J. K. (1973) “Spatial Autocorrelation,” Pion, London, England.
- Codeço, C. T. (2001) “Endemic and epidemic dynamics of cholera: the role of the aquatic reservoir,” *BMC Infectious Diseases* 1:1.
- Colwell, R. R., Kaper, J., and Joseph, S. W. (1977) “*Vibrio cholerae*, *Vibrio parahaemolyticus* and other vibrios: Occurrence and distribution in Chesapeake Bay” *Science* 198:394-396.
- Fortin, M.-J., Dale, M. R. T., and ver Hoef J. (2002) “Spatial analysis in Ecology” Vol 4:2051-2058 in Encyclopedia of Environmetrics (ISBN 0471 899976) Ed. Abdel H. El-Shaarawi and Walter W. Piegorsch. John Wiley & Sons, Ltd, Chichester.
- Glass, R. I.; Claeson, M.; Blake, P. A.; Waldman, R. J. and Pierce, N. F. (1991) “Cholera in Africa: lessons on transmission and control for Latin America” *The Lancet* 338:791-795
- Kaper, J. B., Morris, J. G. Jr., and Levine, M. M. (1995) “Cholera” *Clinical Microbiology Reviews* 8(1):48-86.
- Keeling, M. (1997) “Modelling the persistence of measles,” *Trends in Microbiology* 5:513-518.
- Krishnamurthy, V. and Kinter III, J. L. (2002) “The Indian Monsoon and its relation to

- global climate variability,” in *Global Climate*, Ed. Rodó, X. Springer-Verlag.
- Koelle, K. and Pascual, M. (2004) “Disentangling extrinsic from intrinsic factors in disease dynamics: A nonlinear time series approach with an application to cholera,” *The American Naturalist* 163 (6): 901-913.
- Lipp, E. K., Huq, A., and Colwell, R. R. (2002) “Effects of global climate on infectious disease: the cholera model” *Clinical Microbiology Reviews* 15(4):757-770.
- Pascual, M., Bouma, M. J., Dobson, A. P. (2002) “Cholera and climate: revisiting the quantitative evidence,” *Microbes and Infection* 4:237-245.
- Pascual, M. and Dobson, A. (2005) “Seasonal patterns of infectious diseases,” *PLoS Medicine* 2(1):e5.
- Pascual, M., Rodó, X., Elner, S., Colwell, R., Bouma, M. (2000) “Cholera Dynamics and El Niño-Southern Oscillation,” *Science* 289:1766-1769.

Tables and Figures

	Mean Cholera	Mean Rainfall	Correlation between Cholera and Rainfall	Amount of Fadeouts
Moran's I	0.04 (0.187)	0.15 (0.018)	0.49 (0.001)	0.10 (0.062)

Table 1: Moran's Index for the different variables. The p-values are indicated in parentheses.

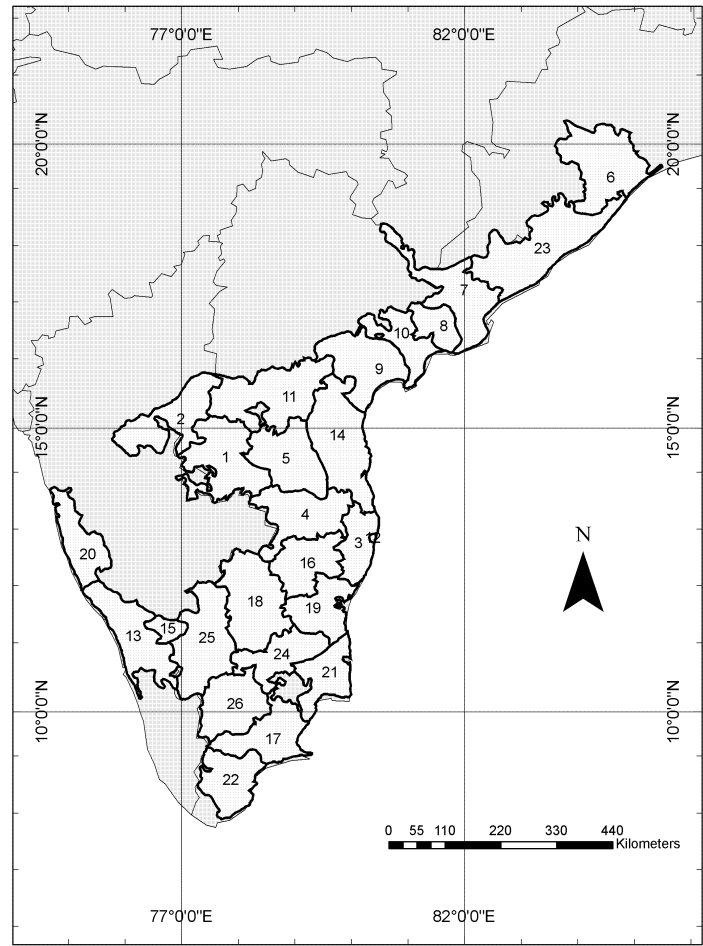


Figure 1: The Madras Presidency with its 26 districts

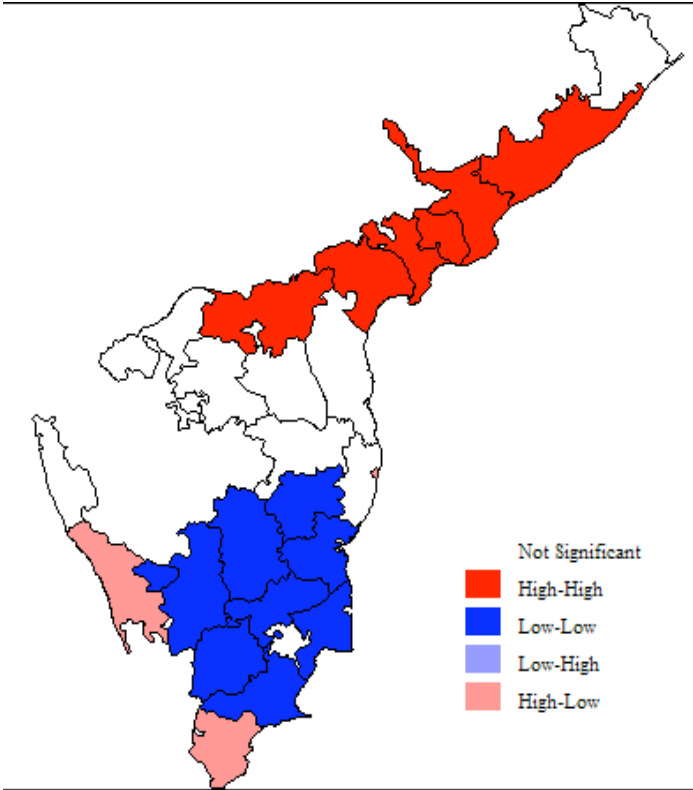


Figure 2: LISA index applied to the coefficient of correlation between cholera and rainfall (without lag).

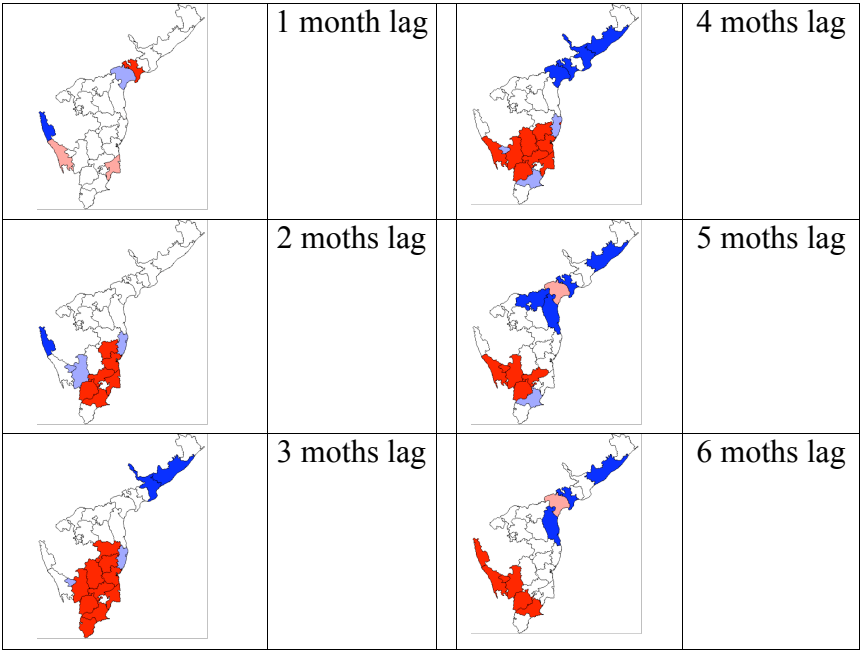


Figure 3: LISA index applied to the coefficient of correlation between cholera and rainfall considering different time lags.

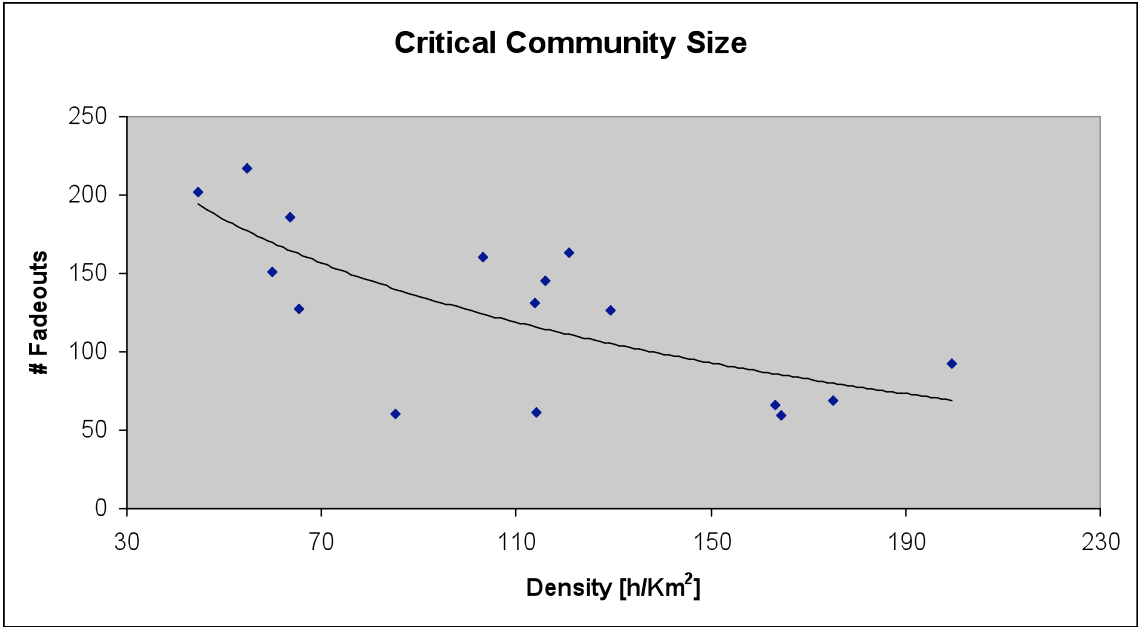


Figure 4: Critical Community Size

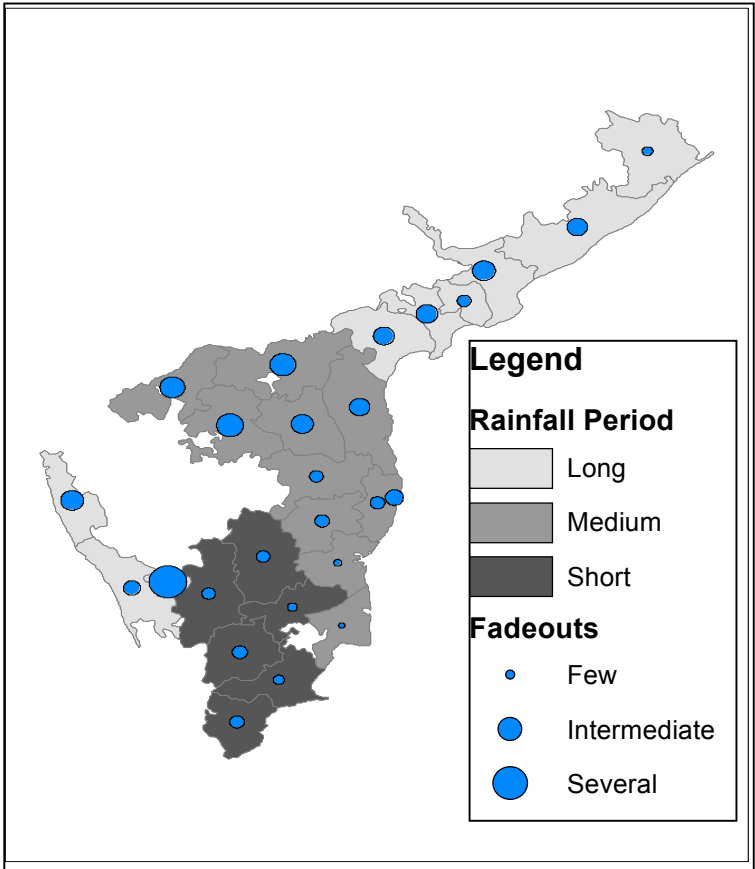


Figure 5: Spatial distribution of the amount of fadeouts and rainfall duration.