

Comparing Models for Early Warning Systems of Neglected Tropical Diseases

Luis Fernando Chaves*, Mercedes Pascual

Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, Michigan, United States of America

Abstract

Background: Early warning systems (EWS) are management tools to predict the occurrence of epidemics of infectious diseases. While climate-based EWS have been developed for malaria, no standard protocol to evaluate and compare EWS has been proposed. Additionally, there are several neglected tropical diseases whose transmission is sensitive to environmental conditions, for which no EWS have been proposed, though they represent a large burden for the affected populations.

Methodology/Principal Findings: In the present paper, an overview of the available linear and non-linear tools to predict seasonal time series of diseases is presented. Also, a general methodology to compare and evaluate models for prediction is presented and illustrated using American cutaneous leishmaniasis, a neglected tropical disease, as an example. The comparison of the different models using the predictive R^2 for forecasts of “out-of-fit” data (data that has not been used to fit the models) shows that for the several linear and non-linear models tested, the best results were obtained for seasonal autoregressive (SAR) models that incorporate climatic covariates. An additional bootstrapping experiment shows that the relationship of the disease time series with the climatic covariates is strong and consistent for the SAR modeling approach. While the autoregressive part of the model is not significant, the exogenous forcing due to climate is always statistically significant. Prediction accuracy can vary from 50% to over 80% for disease burden at time scales of one year or shorter.

Conclusions/Significance: This study illustrates a protocol for the development of EWS that includes three main steps: (i) the fitting of different models using several methodologies, (ii) the comparison of models based on the predictability of “out-of-fit” data, and (iii) the assessment of the robustness of the relationship between the disease and the variables in the model selected as best with an objective criterion.

Citation: Chaves LF, Pascual M (2007) Comparing Models for Early Warning Systems of Neglected Tropical Diseases. *PLoS Negl Trop Dis* 1(1): e33. doi:10.1371/journal.pntd.0000033

Editor: Juerg Utzinger, Swiss Tropical Institute, Switzerland

Received: April 12, 2007; **Accepted:** June 4, 2007; **Published:** October 22, 2007

Copyright: © 2007 Chaves, Pascual. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: LFC was supported by a summer fellowship from the Department of Ecology and Evolutionary Biology, and grants from: The International Institute, The Rackham Graduate School at University of Michigan, and Fundación Polar (Caracas, Venezuela). MP was supported by The Graham Institute for Environmental Sustainability at University of Michigan, the National Oceanic and Atmospheric Administration (Oceans and Health Program NA 040 AR 460019) and NSF-NIH (Ecology of Infectious Diseases Grant EF 0430 120). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

*E-mail: lfchaves@umich.edu

Introduction

One of the best documented patterns in the dynamics of vector-transmitted diseases is their periodicity at seasonal and interannual temporal scales [1–7]. These periodicities are the basis for the proposal that early warning systems (EWS) are feasible and useful tools for planning and decision making [2]. EWS are alert systems whose objective is to predict either epidemic outbreaks in regions where disease transmission is unstable or large outbreaks where the disease is endemic. From the early 1910s, when Captain S. R. Christophers of the British army developed a system to predict malaria in India using climatic and socioeconomic data [8,9], to present times when systems are based on indoor resting densities of vectors [10], climate, land use, and satellite imagery [11], EWS have been regarded as useful tools to help the development of poor and disease-stricken nations [2,11]. The early experience by Christophers was highly successful, and his system was in use until the 1940s, when the importance of malaria as a public health issue in the Indian subcontinent diminished [9,11]. However, recent

results have demonstrated that the blind use of EWS can lead to unreliable forecasts, especially when models are used in regions where the connection between climate and disease is not well understood [12].

Despite the possible caveats of climate-based EWS, especially because of the complexity of human diseases for which social components can be as important as natural forces [13–15], there are successful examples of prediction of “out-of-fit” data based on the known association between climate and disease [6]. Although most of the effort in developing EWS has been focused on malaria [1,2,16], similar efforts would be valuable for neglected tropical diseases, which represent a large burden for developing countries and whose transmission is sensitive to climate variability [6,17]. The leishmaniasis in particular represent the fourth most important neglected tropical disease, with a burden of at least 2.1 million infected people per year, second to malaria in terms of the number of people affected by a protozoan vector-transmitted disease [17,18]. Like many other diseases, the infections are caused by protozoa, belonging to any of several different species of *Leishmania*

Author Summary

Early Warning Systems (EWS) are management tools to predict the occurrence of epidemics. They are based on the dependence of a given infectious disease on environmental variables. Although several neglected tropical diseases are sensitive to the effect of climate, our ability to predict their dynamics has been barely studied. In this paper, we use several models to determine if the relationship between cases and climatic variability is robust—that is, not simply an artifact of model choice. We propose that EWS should be based on results from several models that are to be compared in terms of their ability to predict future number of cases. We use a specific metric for this comparison known as the predictive R^2 , which measures the accuracy of the predictions. For example, an R^2 of 1 indicates perfect accuracy for predictions that perfectly match observed cases. For cutaneous leishmaniasis, R^2 values range from 72% to 77%, well above predictions using mean seasonal values (64%). We emphasize that predictability should be evaluated with observations that have not been used to fit the model. Finally, we argue that EWS should incorporate climatic variables that are known to have a consistent relationship with the number of observed cases.

spp. (Kinetoplastida: Trypanosomatidae), transmitted by sand flies (Diptera: Psychodidae). The clinical manifestation encompasses visceral and cutaneous/mucocutaneous cases, and is associated with a certain parasite species [6]. Our previous results indicate that American cutaneous leishmaniasis (ACL) is a good candidate for the use of climate-based EWS, because predictions with seasonal autoregressive (SAR) models can have an accuracy of over 75% [6]. Our objective in this paper is to illustrate a protocol for the development of EWS, including the evaluation of different linear and non-linear techniques for time series modeling and prediction, as well as assessment of the robustness of the relationship between the disease and climate that is the basis for building EWS.

Methods

Data

Leishmaniasis. Monthly records of ACL cases from January 1991 to December 2001 were obtained from the epidemic surveillance service Vigilancia de la Salud, of Costa Rica. The data were normalized using a square root transformation.

Climatic Covariates. The temperature (T) data are those used in [6] consisting of the average temperature in the $0.5^\circ \times 0.5^\circ$ grids composing the Costa Rica land surface [http://www.cru.uea.ac.uk,19]. The monthly average of these temperature records, T , and the multivariate ENSO index, MEI , [http://www.cdc.noaa.gov/people/klaus.wolter/MEI, 20] were used as predictors for modeling the transformed ACL cases. For all the models below, except for the non-linear forecasting (NLF) and the basic structural model (BSM), the lags for the introduction of climate covariates T and MEI were chosen based on our previous results using cross-correlation functions [6], with a fixed delay (i.e., months preceding the cases series) of 13 months for MEI and 4, 8, and 20 months for T . All time series are shown in Figure 1. Other climatic covariates, precipitation and relative humidity, were ignored since they did not show a strong association with the case data using non-stationary tools like wavelet cross-coherence [6].

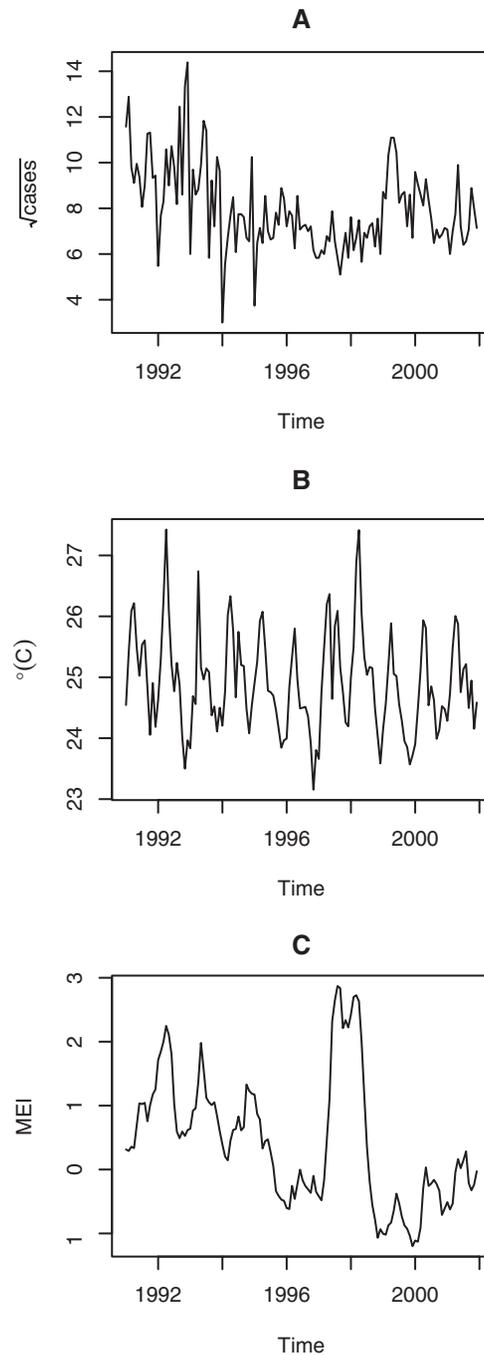


Figure 1. Time Series. (A) Square root Transformed ACL Cases in Costa Rica. (B) Mean Temperature in Costa Rica. (C) MEI. doi:10.1371/journal.pntd.0000033.g001

Statistical Analyses

Forecasting models. Several linear and non-linear models were fitted to the square root transformed case data. Brief descriptions follow of: (1) the approach to handle seasonality, (2) the types of models used, and (3) their classification as linear or non-linear.

Seasonality. To introduce seasonality, the strategy for all models was to include lags 12 and 13 of the transformed case data. This approach was chosen because the autoregressive treatment of seasonality is known to be the best approximation to the

asymptotic cyclical structure of a time series [21]. This approach specifically allows a better minimization of the error variance when compared to a fixed seasonality implemented with a standard cyclical function (such as sines or cosines) that leads to a symmetrical cyclical structure [21].

Linear. In this class of models, parameters have a linear relationship with the response variable [22], in this case the transformed number of cases. This definition should not be confused with that of a linear dynamical system where the relationship of the dependent variables or covariates is linear with that of the independent variable [23]. In fact, linear models can be used to fit the parameters of non-linear dynamical systems, provided that the relationship between a response (which can be a transformation of the independent variable in the non-linear dynamical system) and the covariates (which also can be transformed) is linked by a parameter linearly. Linear models used in this paper include SAR and BSM.

Non-Linear. In these models, the relationship between the response and the parameters for the predictors is not constrained to be linear. Models include NLF, generalized additive models (GAM), and feed-forward neural networks (FNN). A description of the methods (linear and non-linear) and of the fitted models can be found in Protocol S1.

Forecasts. For all models, forecasts were obtained for prediction time intervals of 1, 3, 6, and 12 months ahead for a total of 24 months each. Each model was refitted recurrently before computing the next prediction by including all the previous months in the series [6]. The accuracy of the forecast was measured using the predictive R^2 , which has an interpretation similar to the R^2 of a linear regression by definition [23] and is obtained as $R^2 = 1 - (\text{mean square error} / \text{variance of the series})$. Thus, the errors are normalized by the variance of the time series; an R^2 of 1 indicates perfect forecasts while a value close to 0 or negative indicates poor predictability. Forecasting accuracy was tested for all the fitted models. To establish a baseline for comparison, the predictive R^2 was also computed when the prediction is the monthly mean value of the transformed time series.

Robustness of the exogenous forcing by climate

Once the best modeling approach was selected, the robustness of the association between the cases and the exogenous forces T and MEI was assessed with a non-parametric bootstrap approach based on 10,000 randomizations. The idea of the non-parametric bootstrap is to reconstruct an experimental dataset based on the fitted values of a model plus the residuals sampled with replacement from such a model [24]. To generate the bootstrap samples, the model with the highest predictive R^2 was used. The bootstrap was initially used to see the frequency (%) of times that the model from which we generated the bootstrap samples was actually selected as the best model, using the Akaike Information criterion [25,26]. Then, using the sub-sample of models selected as best that also have the highest probabilities in the above bootstrap test, we constructed confidence intervals for the parameters. We further refitted the model without the last 24 points to make forecasts and get the predictive R^2 confidence intervals.

Results

Figure 2 shows the square root transformed cases plotted against their lagged values (1, 12 and 13 months) and the lagged covariates T (4 months) and MEI (13 months). In all cases, no obvious non-linearity is apparent in the relationship among the three variables. As expected, all models but FNN were most

successful for predictions of 1 month ahead. However, for prediction steps larger than one month only NLF, SAR and GAM models with environmental covariates, MEI and T (4 months lag), did better than predictions based on the average of the time series (Table 1). The models with the worst performance were FNNs, followed by BSM and the null SAR (i.e., without covariates). For NLF, the best results were found with $E = 2$ and $E = 3$, with the latter embedding dimension providing slightly better results for a 12 months ahead prediction.

The predictive R^2 was highest for the SAR model with T (4 months lag) and MEI (13 months lag) as covariates. Thus, the fitted values and residuals used for the bootstrap were those of the model in the first equation of (1) in Protocol S1. The bootstrap results show that the best model is the one used to generate the data (for 67.40% of the simulated time series, the model was selected as best). The confidence intervals for this model show that the parameters for T and MEI are statistically significant, a result that holds even if the intervals are constructed using the values for this parameter when the model was not selected as best (Figure 3A). The autoregressive terms, however, are not significant as the confidence intervals include 0. The variance of the residuals obtained from the real data is significantly shorter than the one from the simulations, probably because of the destruction of the autoregressive structure by the re-sampling of residuals [25]. Finally, the results also show (Figure 3B) that the maximum forecasting ability for these models is 80%, and can be as low as 50% probably because of the sensitivity of the models to a lack of a well-defined SAR structure.

Discussion

The need for forecasts by policy makers goes well beyond the development of EWS for diseases. Due to large-scale, rapid changes, from increased average temperatures to extensive land use changes, major alterations in biogeochemical cycles, water availability, food production, biodiversity and diseases are already occurring and likely to be exacerbated in the future [27,28]. Although the imperative need for predictions that can inform policy has been repeatedly emphasized [11,28], the common practice regarding diseases is to evaluate models by their ability to fit the data [29–35] and only in very few instances have tests been conducted based on data that have not been used to fit the models [6]. Consideration of “out-of-fit” data is critical if we are to evaluate the ability of the models to predict the future.

In this paper, we have presented several methods to study seasonal time series, and used a simple measure, the predictive R^2 , to compare models based on their ability to predict future dynamics and not their goodness of fit of the past. By comparison with modeling results for other infectious diseases on the predictability of NLF methods [36], our results demonstrate a very high predictability for ACL. An important element that might explain this difference is the association of this disease to climate, since models that incorporated climatic covariates performed generally better than those that only considered previous disease levels. Another explanation might be the robustness of the association between the disease and climatic covariates as demonstrated by the bootstrap results. While the parameters for the covariates are statistically significant, the autoregressive parameters are not consistently so, and the variance of the residuals significantly increases.

One of the main lessons from the study of populations is that non-linear dynamics are common in nature but often satisfactorily captured by linear approximations [37,38]. This has been demonstrated by the analysis of time series from a wide variety

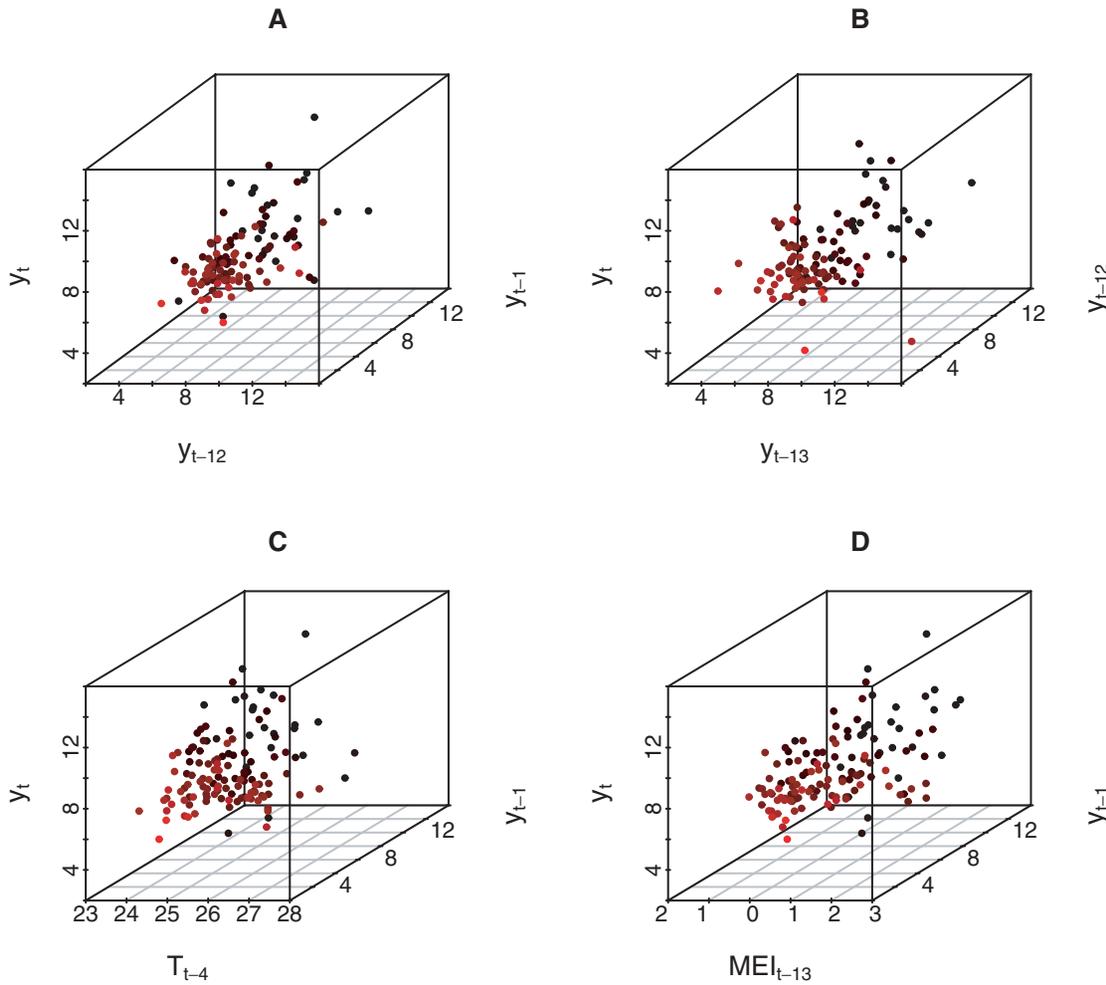


Figure 2. Multidimensional plots for the square root transformed ACL cases (y_t) as function of lagged components and climatic covariates. (A) Autoregressive (y_{t-1}) and Seasonal (y_{t-12}) components. **(B)** Seasonal (y_{t-12}) and Autoregressive Seasonal (y_{t-13}) components. **(C)** Autoregressive component (y_{t-1}) and Temperature (lag 4, T_{t-4}). **(D)** Autoregressive component (y_{t-1}) and MEI (lag 13, MEI_{t-13}). doi:10.1371/journal.pntd.0000033.g002

Table 1. Models and predictive R^2

Model	1 month	3 months	6 months	12 months
NLF (E=2)	0.69	0.62	0.61	0.66
NLF (E=3)	0.67	0.60	0.59	0.67
NLF (E=4)	0.66	0.59	0.58	0.66
FNN (2 Layers)	0.55	0.53	0.44	0.44
FNN (3 Layers)	0.62	0.58	0.61	0.60
SAR (null)	0.71	0.64	0.62	0.57
SAR (MEI)	0.73	0.67	0.67	0.66
SAR (MEI+T)	0.77	0.73	0.73	0.72
BSM	0.69	0.59	0.52	0.65
GAM (MEI)	0.66	0.59	0.56	0.57
GAM (MEI+T)	0.73	0.68	0.67	0.68
MEAN	0.64	0.64	0.64	0.64

For model identification, see common abbreviations. Months indicate the number of months predicted ahead. Mean indicates the results that could be obtained by just using the monthly average number of cases. doi:10.1371/journal.pntd.0000033.t001

of animals and diseases. While chaos is present in a small sample of the populations considered, periodicities are common, particularly in infectious diseases, that can be explained by either the effect of exogenous forces, like climate, or endogenous ones, like recruitment of new individuals and the concurrent changes in densities [39–42]. Our results indicate that ACL is another example of a population phenomenon whose dynamics can be satisfactorily described by linear statistical models, provided that appropriate covariates and transformations of the data are used. Thus, though linear models do best, functional forms underlying the influence of covariates are likely to be non-linear as indicated by the transformations used. This result is further supported by the observation that the predictive R^2 for NLF with $E=3$ does not vary with the prediction time step, while this value for the SAR model without covariates decreases abruptly, as expected in systems where the dynamics are non-linear [36–43]. Linear models were also used successfully for other vector-borne diseases, malaria [43] and Ross river virus [4,35], and for cutaneous leishmaniasis in other regions of the new world [45]. For ACL, the usefulness of linear models (after appropriate transformation) might also follow from the fact that humans are only sinks for the pathogen and, therefore, provide no feedback to transmission [46,47]. This conjecture would not necessarily apply to other

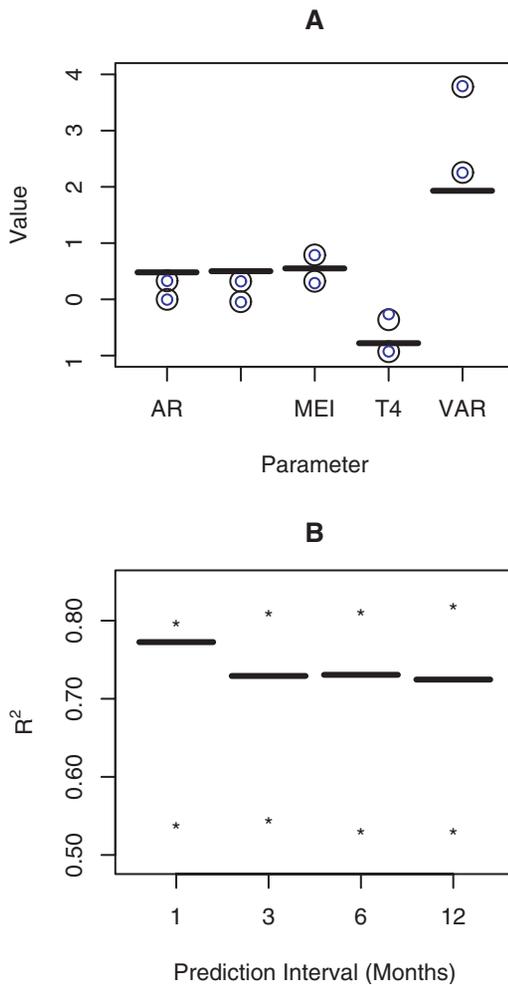


Figure 3. Bootstrap Experiment. (A) 95% Confidence intervals for the parameters of the best model. AR stands for the autoregressive component of the model (ϕ_1); ARseas for the seasonal autoregressive component of the model (ϕ_{12}); VAR for the variance of the residuals (σ_e^2); MEI and T4 for the parameter for MEI at lag 13 (α) and Temperature at lag 4 (γ), respectively. Black signs are 95% confidence intervals using values from the sub-sample when the model is selected as best, and blue including all the bootstrap samples. The structure of the best model can be seen in Protocol S1. **(B)** Predictive R^2 and the 95% confidence intervals, indicated by stars, for the bootstrapped best model and prediction interval.

doi:10.1371/journal.pntd.0000033.g003

vector-transmitted diseases in which infected humans provide sources of new infections within the population.

This result also highlights two open questions that need to be addressed when modeling infectious diseases transmitted by vectors: first, the appropriate functional form to introduce climate variables into the dynamics [46,48]; second, the best approach for modeling seasonality [8,49]. Mathematically the relationship between climatic co-variables and the numbers of the disease can be non-linear, described by simple non-linear functions, like those of the functional responses in consumer-resource interactions (e.g., hyperbolic functions) [50] or modeled by linear models with self-excited thresholds [51]. This is especially relevant, since a saturating non-linear functional form can lead to very different scenarios in the dynamics of the disease under altered environmental conditions. In the case of ACL, however, no apparent need

for non-linear functions describing the relationship to climate was evident. In general, seasonality has been modeled using fixed structures, i.e., values are assumed to be constant [e.g., 9,49] or approximated by the sum of sine and cosine functions [e.g., 41,52]. The introduction of SAR seasonality in mechanistic models should be further investigated.

A factor that deserves further consideration in developing EWS is the understanding of the role of space. Predictability at more local scales was not addressed here because half of the series was only available at levels below that of the whole country, and because Costa Rica encompasses a small area for which temperature variability is quite homogenous, as seen in the very small variability between temperature grids. However, for larger spatial scales heterogeneities in the landscape for disease transmission would need to be considered [53].

Conclusions

EWS are a feasible ecological application for neglected tropical diseases, as illustrated for ACL. Available models have good levels of predictability up to one year ahead for the number of cases. Predictability strongly depends on the use of an appropriate structure for the different components of the model, including seasonality and exogenous drivers such as climatic variables. Depending on the model, predictability can range from poor, with approximately 50% accuracy, to high, with 80% accuracy, significantly better than that of seasonal averages (about 65%). Forecasts can be useful in planning services for the populations affected, allowing estimates of approximate number of hospital beds, vaccine shots, drug doses and vector control measures. If EWS need to incorporate the spatial spread of the disease, they should do so dynamically and in relation to different landscapes, such as the geopolitical unit of this study or regions with similar climatic patterns [53]; otherwise, predictions are likely to fail, as illustrated by [12]. While there is no unique early warning system for a given disease, there should be a general approach for the development of EWS. Our work illustrates three key components of such an approach for vector-borne diseases: (i) the evaluation of predictability with “out-of-fit” data and not simply goodness of fit [6,40,41]; (ii) the comparison of a suite of possible models in terms of predictability [55,56], and (iii) the robustness of the relationship with covariates in the selected model. Here, robustness is used following [55], to identify covariates that are useful to predict disease numbers even when the skeleton of the model changes. Finally, none of these efforts are possible without the invaluable role of sustained surveillance and monitoring efforts. A historical retrospective reinforces this point: the success of Christophers was possible because of data availability and a deep knowledge of malaria biology, from parasites to mosquitoes and humans, realizing the influence of factors as diverse as weather and wheat prices in rendering the epidemics of malaria predictable [8]. Time series sufficiently long for developing and evaluating forecasting models around the world are countable; their number pales by comparison to the data available for weather forecasting. It is imperative that ongoing efforts are sustained and new ones are initiated whose long-term planning includes EWS as a specific goal.

Supporting Information

Protocol S1 Linear and non-linear models for time series forecasting.

Found at: doi:10.1371/journal.pntd.0000033.s001 (0.07 MB DOC)

Acknowledgments

We thank Justin M. Cohen, Inés Széliga, Andy Dobson and Mark Wilson for their comments and suggestions during early stages of this work. Suggestions by a referee also helped in strengthening the discussion.

References

- Connor SJ, Thomson MC, Molyneux DH (1999) Forecasting and prevention of epidemic malaria: new perspectives on an old problem. *Parassitologia* 41: 439–448.
- Thomson MC, Connor SJ (2001) The development of malaria early warning systems for Africa. *Trends Parasitol* 17: 438–445.
- Hay SI, Myers MF, Burke DS, Vaughn D, Endy T, et al. (2000) Etiology of interepidemic periods of mosquito-borne disease. *Proc Natl Acad Sci USA*, 97: 9335–9339.
- Tong SL, Hu WB, McMichael AJ (2004) Climate variability and Ross River Virus transmission in Townsville region, Australia, 1985–1996. *Trop Med Int Health* 9: 298–304.
- Patz JA, Campbell-Lendrum D, Holloway T, Foley JA (2005) Impact of regional climate change on human health. *Nature* 438: 310–317.
- Chaves LF, Pascual M (2006) Climate cycles and forecasts of cutaneous leishmaniasis, a nonstationary vector-borne disease. *PLoS Med* 3: e295.
- Woodruff RE, Guest CS, Garner MG, Becker N, Lindsay M (2006) Early warning of Ross virus epidemics: combining surveillance data on climate and mosquitoes. *Epid* 17: 569–575.
- Gill CA (1928) The genesis of epidemics and the natural history of disease. An introduction to the science of epidemiology based upon the study of epidemics of malaria, influenza and plague. London: Bailliere, Tindall and Cox.
- Gill CA (1938) The seasonal periodicity of malaria and the mechanism of the epidemic wave. London: J. & A. Churchill Ltd.
- Lindblade KA, Walker ED, Wilson ML (2000) Early warning of malaria epidemics in African highlands using *Anopheles* (Diptera: Culicidae) indoor resting density. *J Med Entomol* 37: 664–674.
- Rogers DJ, Randolph SE, Snow RW, Hay SI (2002) Satellite imagery in the study and forecast of malaria. *Nature* 415: 710–715.
- Hay SI, Were EC, Renshaw M, Noor AM, Ochola SA, et al. (2003) Forecasting, warning, and detection of malaria epidemics: a case study. *Lancet* 361: 1705–1706.
- Wilson ML (1994) Developing paradigms to anticipate emerging diseases – transmission cycles and a search for pattern. *Ann N Y Acad Sci* 740: 418–422.
- Levins R, Lopez C (1999) Toward an ecosocial view of health. *Int J Health Serv* 29: 261–293.
- Chaves LF (2007) *Casas Muertas* and *Oficina No.1* : internal migrations and malaria trends in Venezuela 1905–1945. *Parasitol. Res.* 101: 19–23.
- Thomson MC, Mason SJ, Phindela T, Connor SJ (2005) Use of Rainfall and sea surface temperature monitoring for malaria early warning system in Botswana. *Am J Trop Med Hyg* 73: 214–221.
- Hotez PJ, Molyneux DH, Fenwick A, Ottesen E, Erlich-Sachs S, et al. (2006) Incorporating a rapid impact package for neglected tropical diseases with programs for HIV/AIDS, tuberculosis and malaria. *PLoS Med* 3: e102.
- Lainson R, Shaw JJ (1978) Epidemiology and ecology of leishmaniasis in Latin-America. *Nature* 273: 595–600.
- New M, Lister D, Hulme M, Makin I (2002) A high-resolution data set of surface climate over global land areas. *Clim Res* 21: 1–25.
- Wolter K, Timlin MS (1998) Measuring the strength of ENSO—How does 1997/98 rank? *Weather* 53: 315–324.
- Priestley MB (1988) Non-linear and non-stationary time series analysis. London: Academic Press.
- Faraway JJ (2005) Linear models with R. Boca Raton: Chapman & Hall. CRC.
- Kaplan M, Glass D (1995) Understanding Nonlinear Dynamics New York: Springer.
- Efron B, Tibshirani R (1993) An introduction to the bootstrap. London: Chapman & Hall.
- Brockwell PJ, Davis RA (2002) Introduction to time series and forecasting. 2nd ed. New York: Springer.
- Shumway RH, Stoffer DS (2000) Time series analysis and its applications. New York: Springer.
- Levins R, Awerbuch T, Brinkmann U, Eckardt I, Epstein P, et al. (1994) The emergence of new diseases. *Am. Scien.* 82: 52–60.
- Clark JS, Carpenter SR, Barber M, Collins S, Dobson A, et al. (2001) Ecological forecasts: an emerging imperative. *Science* 293: 657–661.
- Abeku TA, De Vlas SJ, Boorsboom GJJM, Tadege A, Gebreyesus Y, et al. (2004) Effects of meteorological factors on epidemic malaria in Ethiopia : a statistical modeling based approach based on theoretical reasoning. *Parasitol* 128: 585–593.

Author Contributions

Conceived and designed the experiments: LC. Analyzed the data: LC. Wrote the paper: MP LC.

- Abeku TA, De Vlas SJ, Boorsboom GJJM, Teklehaimanot A, Kebede A, et al. (2004) Forecasting malaria incidence from historical morbidity patterns in epidemic-prone areas of Ethiopia: a simple seasonal adjustment method performs best. *Trop Med Int Health* 7: 851–857.
- Teklehaimanot HD, Lipsitch M, Teklehaimanot A, Schwartz J (2004) Weather-based prediction of *Plasmodium falciparum* malaria in epidemic-prone regions of Ethiopia I. Patterns of lagged weather reflect biological mechanisms. *Malar J* 3: 41.
- Teklehaimanot HD, Schwartz J, Teklehaimanot A, Lipsitch M (2004) Weather-based prediction of *Plasmodium falciparum* malaria in epidemic-prone regions of Ethiopia II. Weather-based prediction systems perform comparably to early detection systems in identifying times for interventions. *Malar J* 3: 44.
- Zhou G, Minakawa N, Githeko AK, Yan G (2004) Association between climate variability and malaria epidemics in the east African highlands. *Proc Natl Acad Sci USA* 101: 2375–2380.
- Hu WB, Nicholls N, Lindsay M, Dale P, McMichael AJ, et al. (2004) Development of a predictive model for Ross River virus disease in Brisbane, Australia. *Am J Trop Med Hyg* 71: 129–137.
- Hu WB, Tong SL, Mengersen K, Oldenburg B (2006) Rainfall, mosquito density and the transmission of Ross River: a time series forecasting model. *Ecol Model* 196: 505–514.
- Grenfell BT, Kleczkowski A, Ellner SP, Bolker BM (1994) Measles as a case study in nonlinear forecasting and chaos. *Phil Trans Roy Soc London A*. 348: 515–530.
- Royama T (1992) Analytical Population Dynamics. London: Chapman and Hall.
- Turchin P (2003) Complex Population Dynamics. Princeton: Princeton University Press.
- Hassell MP, Lawton JH, May RM (1976) Patterns of dynamical behavior in single species populations. *J. Anim. Ecol.* 45: 471–486.
- Turchin P, Taylor AD (1992) Complex dynamics in ecological time series. *Ecology*. 73: 289–305.
- Ellner S, Turchin P (1995) Chaos in a noisy world new methods and evidence from time series analysis. *Am. Nat.* 145: 343–375.
- Kendall BE, Prendergast J, Bjornstad ON (1998) The macroecology of population dynamics: taxonomic and biogeographic patterns in population cycles. *Ecol. Lett.* 1: 160–164.
- Sugihara G, May RM (1990) Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature*. 344: 734–741.
- Loevinsohn ME (1994) Climatic warming and increased malaria incidence in Rwanda. *Lancet*. 343: 714–718.
- Feliciangeli MD, Rabinovich J (1998) Abundance of *Lutzomyia ovallesi* but no *L. gomezi* (Diptera: Psychodidae) correlated with cutaneous leishmaniasis incidence in north-central Venezuela. *Med. Vet. Entomol.* 12: 121–131.
- Chaves LF, Hernandez MJ (2004) Mathematical modelling of American cutaneous leishmaniasis: incidental hosts and threshold conditions for infection persistence. *Acta Trop.* 92: 245–252.
- Chaves LF, Hernandez MJ, Dobson AP, Pascual M (2007) Sources and sinks: revisiting the criteria for identifying reservoirs for American cutaneous leishmaniasis. *Trends Parasitol.* 23: 311–316.
- Cazelles B, Hales S (2006) Infectious diseases, climate influences and nonstationarity. *PLoS Med.* 3: e328.
- Altizer S, Dobson AP, Hosseini P, Hudson PJ, Pascual M, et al. (2006) Seasonality and the dynamics of infectious diseases. *Ecol Lett.* 9: 467–484.
- Murdoch WW, Briggs CJ, Nisbet RM (2003) Consumer-Resource Dynamics. Princeton: Princeton University Press.
- Grenfell BT, Wilson K, Finkenstädt BF, Coulson TN, Murray S, et al. (2004) Noise and determinism in synchronized sheep dynamics. *Nature*. 394: 674–677.
- Pascual M, Rodo X, Ellner SP, Colwell R, Bouma MJ (2000) Cholera Dynamics and El Niño- Southern Oscillation. *Science*. 289: 1766–1769.
- Levin SA (1992) The problem of pattern and scale in ecology. *Ecology* 73: 1943–1967.
- Grover-Kopec E, Kawano M, Klaver RW, Blumenthal B, Ceccato P, et al. (2005) An online operational rainfall-monitoring resource for epidemic malaria early warning system in Africa. *Malar J.* 4: 6.
- Levins R (1966) The strategy of model building in population biology. *Am. Scien.* 54: 421–431.
- Levins R (2006) Strategies of abstraction. *Biol. Philos.* 21: 741–755.