

Real-time probabelistic forecasts of the 2018 – 2019 Ebola outbreak in the North-Eastern Democratic Republic of Congo

CANDIDATE NUMBER

September 22, 2019

Word counts:

Abstract: XXX

Report: XXXX

Abstract

Objectives: Construct and evaluate probabilistic forecasting models for the 2018–2019 Ebola outbreak in north-eastern DRC on both a national and sub-national level.

Design: Using publicly available data on daily numbers of reported Ebola cases we constructed a set of semi-mechanistic models based on a branching process with a time varying reproduction number. We included models with both Poisson and negative binomial offspring distributions and two methods for forecasting the reproduction number. These methods used the estimated historic evolution of reproduction number and then either forecasted a constant reproduction number or used a forecast based on a Bayesian state space model with a semi-local trend. We compared forecasts up to 28 days ahead from the models using proper scoring rules and checked the calibration using Probability Integral Transformations.

Results: On the national level we found that a negative binomial distribution was needed for calibrated one-day ahead forecasts, but that the model with a Poisson offspring distribution and a semi-local trend for the reproduction number was the best model for longer forecasts and could provide calibrated forecasts for a forecasting horizon of nine days. The semi-local trend models could also provide calibrated forecasts for some of the health zones, but for most of the health zones with a large number of cases the models did at best provide a few days of calibrated forecasts. At both the national and the sub-national level we found that the models with constant reproduction number would underestimate the uncertainty in the process and that semi-local trend models are likely to overestimate the uncertainty of the process.

Conclusions:

Flexible semi-mechanistic models based on a branching process can be used to give calibrated short-term forecasts of the Ebola outbreak in northern DRC on the national level and partly at the sub-national level. The forecasts can be provided in real-time to support the outbreak response and can provide insight to increase our understanding of the outbreak.

Keywords: Ebola, DRC, probabilistic forecasts, mathematical modelling, semi-mechanistic model, calibration, proper scoring rules.

Contents

1	Introduction	4
2	Background	5
2.1	2018–2019 Ebola outbreak in north-eastern DRC	5
2.2	Mathematical Modelling of Ebola	6
3	Methods	8
3.1	Data Sources	8
3.2	Model	8
3.2.1	Offspring distribution	10
3.2.2	Reproduction number	12
3.2.3	Forecasting incidence	14
3.3	Assessing probabilistic forecasts	14
3.3.1	Calibration	15
3.3.2	Sharpness and Bias	16
3.3.3	Proper scoring rules	17
3.4	Implementation	18
4	Results	19
4.1	National level forecasts	21
4.2	Sub-national forecasts	23
5	Discussion	26
6	Conclusions	33
Appendices		42

1 Introduction

Real time modelling and forecasting of infectious diseases can play a crucial role during disease outbreaks like the ongoing Ebola outbreak in the north-eastern Democratic Republic of Congo (DRC) [1, 2]. The use of models for ebola response was pioneered during the previous large Ebola outbreak in West Africa in 2014–2015 [3]. The models, that were developed during and after this outbreak, were used both to increase our understanding of the outbreak and to predict the spread of the outbreak. Estimates were used in the official response planning [4] and for example to estimate needed bed capacity [5]. Disease modelling has also been successfully implemented and used to inform response plans for many other infections diseases including Zika [6] and influenza [7]. To effectively use results from mathematical modelling in an outbreak it is of key importance that modelling outputs are integrated into the outbreak response and decision making [2]. This means that forecasts need to be available in real-time and be integrated into the outbreak surveillance information flow. In addition improved communication and embedding of modelling and modellers into the outbreak response is needed.

To provide useful information in an outbreak situation models should be probabilistic such that they assess and incorporate uncertainty into their predictions [8, 9, 10]. An understanding of the range of possible outcomes from a model and their associated probabilities is essential for effective use of modelling results in an outbreak situation since there are both large uncertainties and randomness associated with the disease transmission process and in the necessary reduction of this process to a mathematical model. Once a model has been formulated, it needs to be critically evaluated before use. Evaluations for probabilistic forecasts requires us to not just assess the accuracy of point predictions, but to assess if the model correctly assess it's own uncertainty [11, 12]. A model that successfully assesses it's own uncertainty is said to be calibrated. We should strive to create well calibrated models that can be updated in real-time and that that are interpretable and can be used in the outbreak response.

The aim of this report is to describe a framework for flexible probabilistic forecasts for the ongoing Ebola outbreak in north-eastern DRC and to assess the calibration of the resulting models. We will compare different models using proper scoring rules to choose the model that gives the best forecasts and use these models to explore the epidemiology of the current outbreak.

2 Background

2.1 2018–2019 Ebola outbreak in north-eastern DRC

Ebola virus disease (EVD) is an often fatal disease that is transmitted through close contact with bodily fluids from infected persons [13]. The symptoms start 2–21 days after infection and include fever, vomiting, diarrhoea and in some cases internal and external bleeding. The risks of contracting the disease depends on close contact with or handling of the sick person or their diseased body. It has therefore been found that health-care workers are at high risk of infection and that burial practises that include direct contact with the dead body is a source of many infections [14]. There have been multiple Ebola outbreaks since the first recorded outbreak in 1976. The case fatality rate has varied between 10 and 90% and apart from the large outbreak in Western Africa 2014–2015 that had over 28,000 cases all of the previous outbreaks have had less than 500 cases [13] making the current outbreak the second largest in history.

The current outbreak in north-eastern DRC, started on the 1st of August 2018, when the Ministry of Health in the Democratic Republic of Congo notified the WHO about a new Ebola outbreak in the North Kivu province [15]. North Kivu and the surrounding provinces Ituri and South Kivu constitute a populous region in north-eastern DRC that borders both Rwanda and Uganda. The area has seen both the Second Congo War (1998–2003) and the ongoing “Kivu Conflict”. The Kivu conflict has seen local rebel forces fight against the armed forces of the government. In addition to local forces, a large UN peacekeeping force

has been heavily involved in the region. This makes the area very difficult to operate in due to security concerns and the high levels of distrust from local population to both the central government and international organisations. By the 27th of September the outbreak had continued spreading and the WHO assessed that the outbreak constituted a very high national and regional risk [16].

The security and distrust from the local community has made the outbreak response very challenging. This has lead to difficulty in tracing contracts, giving vaccinations and treating patients. There have been many episodes of violence directed towards the Ebola responders which has led to fatalities and multiple stops in the response activities [17, 18]. Even with these difficulties the outbreak has mainly been contained in the region with only imported cases and no transmission seen in the large city of Goma or across the borders in Uganda and Rwanda.

During the outbreak, the experimental vaccination rVSV-ZEBOV-GP has been used in a ring vaccination strategy. The vaccine was given under the compassionate use regime and have been evaluated during the outbreak. Preliminary data shows positive signs that this vaccine will provide protection from Ebola [19]. The vaccines has also been given to health-care workers and children. This outbreak has also seen a randomised trial for four proposed medicines to treat Ebola. The trial was terminated at the 12th of August due to findings that two of the four drugs that were included showed a clear effect of reducing mortality from Ebola, especially with early treatment [20].

On the 17th of July the Director General of the WHO declared the Ebola outbreak a Public Health Emergency of International Concern [21] and as of September 2019 the outbreak is still ongoing with recent new cases from 17 health zones [22].

2.2 Mathematical Modelling of Ebola

A large number of different mathematical models have been proposed for modelling and forecasting of Ebola, see e.g [3, 23]. The models can be divided into categories based on

the structure of the models; mechanistic, semi-mechanistic, phenomenological and hybrid models. Mechanistic models are based on a specified mechanism for disease transmission and includes standard compartmental models, while phenomenological aim to describe empirical relationships and do not include any specified mechanism for disease transmission. For Ebola prediction an example of a mechanistic model is the SEIR model in [24]. Many different phenomenological models have been developed for ebola predictions. They include logistic growth models [25] and the class of autoregressive integrated moving average models used for many applications of time-series forecasting [26]. A semi-mechanistic model include some mechanistic assumptions, together with some parts of the model that are more empirical or phenomenological. The aim of these models is to model some important parts of the disease transmission process, while giving enough flexibility to fit real life outbreaks. For Ebola forecasting examples include SEIR models with a random walk describing the transmission rate or reproduction number [8, 27]. Based on a review of models for the West African Ebola outbreak it was found that it was important with models that actively modelled the uncertainty in the disease transmission process [3, 28] and that community standards for developing, reporting and comparing models are put in place.

The most recent evaluation of Ebola models was the RAPIDD Ebola Forecasting challenge in 2018 [23]. The challenge presented a range of different scenarios of an Ebola outbreak that was similar to the West African outbreak and invited different groups to submit forecasts and compared them. Eleven different models with a large range of different modelling structures were submitted to the competition. When comparing the models they found that for short-term forecasting the accuracy was not related to model complexity and that “light” non-parametric models achieved similar forecasting accuracy as larger mechanistic models. The more complex mechanistic models would on the other hand be better suited to understand the effects of interventions.

In this thesis we will present a semi-mechanistic modelling framework that includes some mechanistic assumptions about serial intervals and forces of infection combined with flexible

time-series methods. The aim is a flexible model with parameters that are easily interpretable and useful during an outbreak response situation. This also requires that the models can be run in real-time and be integrated with the response surveillance.

3 Methods

3.1 Data Sources

We will use daily reports of new cases from the Ebola outbreak in north-eastern DRC as reported by the Ministry of Health and the World Health Organisation as the data source for our models. This data has been made publicly available and has been distributed via the Humanitarian Data Exchange [29]. We use data from the beginning of the outbreak until the 1st of September 2019. The data gives the number of incident cases both at the national and at the health-zone level. For our models we will use confirmed cases so that estimates of the reproduction number and other epidemiological parameters correspond only to Ebola and not Ebola and Ebola-like illnesses. Due to the nature of the outbreak and the nature of the data some confirmed cases are later found to not be Ebola. This can lead to days with negative incidence. We have chosen to set any negative incidence to zero as we do not have any information about which of the earlier cases has been removed. In July and August 2018 there were some days without any reported data, for these days we equally distribute the number of new cases seen at the end of the interval of missing data to each missing day.

3.2 Model

In our aim to construct a flexible semi-mechanistic framework for modelling of Ebola we will consider models where the daily incidence, I_t , follows a modified branching process [30]. This class of models have been considered for infectious disease modelling previously in [31, 32, 33] and were used to model Ebola during the 2014–2016 Ebola outbreak in Western Africa [4, 34]. The main idea is that each infected person gives rise to ν new infections,

where ν is a random variable distributed with an offspring distribution with an expected value given by the reproduction number R . This gives rise to the standard interpretation of the reproduction number as the average number of secondary cases caused by one primary case. Using random variables for the number of secondary cases gives a stochastic model that allows us to model the inherent uncertainty in the infectious transmission process. To model the time between the initial case and the secondary cases we will use the serial interval. This gives a process where the expected number of new cases at time t is given by the force of infection, λ_t , which is a product of previous incidence weighted by the discretised serial interval, w_τ , and the instantaneous reproduction number, R_t . The expected value of the incidence is given by

$$E(I_t) = \lambda_t = R_t \sum_{s=1}^{t-1} I_s w_{t-s}. \quad (1)$$

This equation gives a general and flexible modelling framework where an individual model is specified by the probability distribution for I_t , the serial interval and the reproduction number as a function of time. Due to the semi-mechanistic nature of this model, it does not depend on many assumptions about the disease transmission process. The main assumptions include that the force of infection is given by the weighted sum of incidence and serial interval and that the reproduction number and that the offspring distribution and time varying reproduction number can adequately describe the outbreak. For all the models in this study we will use a gamma-distribution with mean 15.3 days and standard deviation of 9.3 days as the serial interval as fitted to data from the West-Africa Ebola outbreak [4]. Our general approach will be to use a few different specifications for the offspring probability distribution and for estimating the time-varying reproduction number and then assess which model gives the best fit to the data.

3.2.1 Offspring distribution

The offspring distribution gives the distribution of the number new infections from each infected person. By definition the expected value of this distribution is the reproduction number. The offspring distribution needs to be a discrete probability distribution over the non-negative integers. The simplest such distribution is the Poisson distribution, where the probability of observing k cases, $P(X = k)$ is given by

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!},$$

where the λ is both the mean and variance of the distribution. This distribution implies that the number of new infections over a small time interval is described by a constant rate where the chance of a new infection in each small sub-interval is independent.

For a number of diseases it has been shown that the offspring distribution has larger variance than that given by the Poisson distribution and has so called super-spreaders [32]. One common way of modelling this over-dispersed distribution is by the negative binomial distribution. The negative binomial distribution describes a process where we have a series of independent opportunities to infect a new person each with probability p , where the process continues until we have had r infection opportunities without an infection. The distribution is given by

$$P(X = k) = \binom{k + r - 1}{k} (1 - p)^r p^k,$$

with $E(X) = \frac{pr}{1-p}$ and $Var(X) = \frac{pr}{(1-p)^2}$. We will use a parameterisation of the negative binomial distribution where we use the mean, μ , and the dispersion parameter k . In this formulation the variance is given by $\mu(1 + \frac{1}{k})$. For $k \rightarrow \infty$ we get a Poisson distribution. The smaller k is, the more the outbreak is dominated by a few super-spreaders and the larger k is the more similar the impact of each infected person is. For SARS it has been found that $k = 0.16$ [32]. For the Ebola outbreak in West Africa they found $k = 0.53$ using

conservative assumptions and data from the transmission tree of the infections [34]. The different offspring distributions can be seen in Figure 1

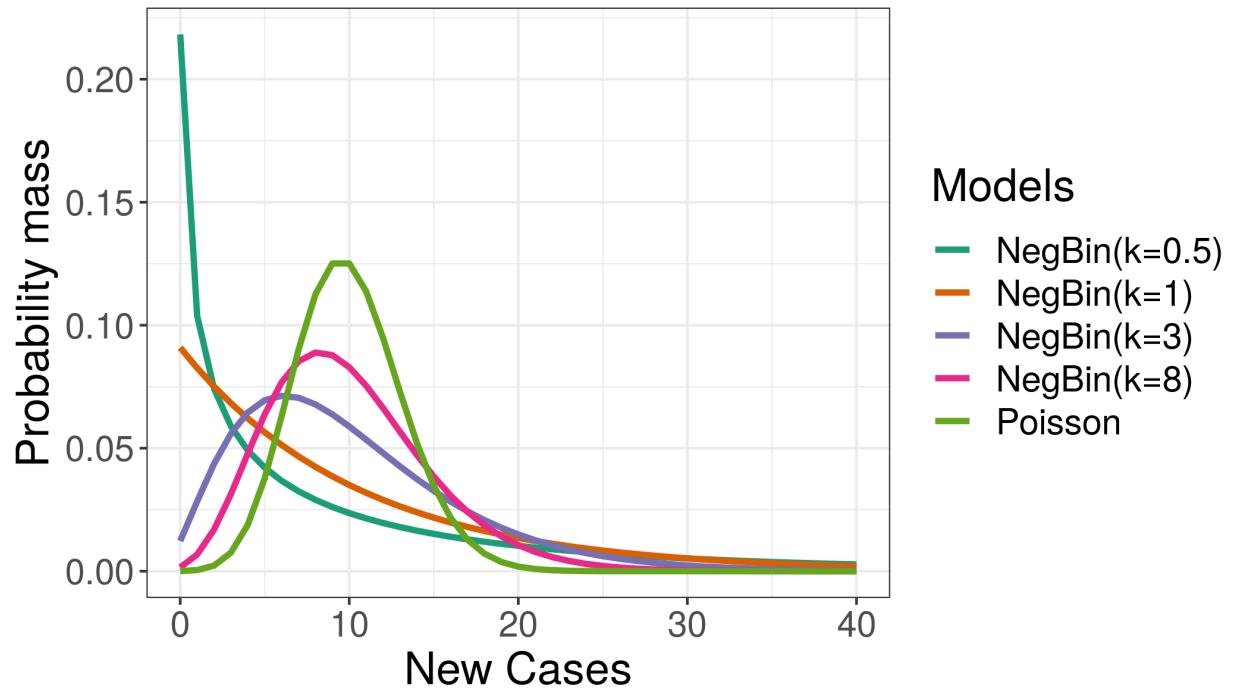


Figure 1: Poisson and negative binomial offspring distributions with varying dispersion parameter k . All the distributions have a mean of 10.

The distribution for total incidence is given by the sum of the offspring distributions for each infected person weighted by the serial distribution. For both the Poisson and negative binomial distribution the sum of independent realisations of the distributions will give back the same distribution with an expected value given by the sum of the expected values. For the negative binomial distribution the dispersion parameter k will remain the same. Therefore the probability distribution for incidence is also given by either the Poisson or the negative binomial distribution if the individual offspring distribution is. We will investigate models with both offspring distributions to assess which models fit the current outbreak better.

3.2.2 Reproduction number

The final ingredient needed to completely specify the dynamics of the model is the evolution of the reproduction number with time. The function for R_t could be some parametric function where we could fit the parameters to data or we can use non-parametric methods to allow a fairly unconstrained evolution with time. This free function for the reproduction number is what gives the model it's flexibility to fit a large number outbreak types. As is common for other non-parametric methods we will impose conditions on the rate of change of the reproduction number with time to ensure that the evolution of the reproduction number is sensible and not jumping around too much.

To forecast I_t , we first need to forecast R_t from data prior to t . In this study we will use an approach where to forecast R_t , we first estimate the values of the reproduction number prior to t using the method described below. We then use these estimated values to predict the future evolution of R_t . We will use different methods for forecasting R_t to compare how well they work.

To estimate the reproduction number from the historical incidence data we use the method developed in [31, 35]. This approaches used the same formulation as in Equation 1 to estimate the reproduction number from daily incidence and the same serial interval as in our model. If the reproduction number is estimated daily it is likely to vary too much from day to day in a manner that it is unrealistic. Therefore the method averages over the last 7 days to get more stable estimates. A Bayesian procedure is used to estimate both the best fit value and the uncertainty of the estimate. We use the R-package EpiEstim [36] to estimate the reproduction number using a parametric gamma distribution for the serial interval with the same mean of 15.3 days and standard deviation of 9.3 days [4].

Once we have calculated the historical values of R_t we can use them to calculate a forecast for R_t that can be used for predictions of new incidence. We will use two different procedures to forecast the reproduction number. The first method is a simple method were we assume that reproduction number remains constant from the last historical value. We

use the uncertainties given by the method in [31] as a prediction interval. In practice this means that to predict the value of R_t we use the method above to find the parameters of the gamma posterior distribution for R_{t-1} . We then get realisations of R_t by drawing random numbers from this posterior distribution.

The second approach to forecast the reproduction number is based on Bayesian structural time series fitted to the historical values of R_t . We use a model with a semi-local linear trend to allow the estimation of a trend in the recent data. To ensure that predicted values of R_t are between 0 and 15 we fit the time series on a transformed scale

$$r_t^* = \log \left(\frac{R_t}{15 - R_t} \right). \quad (2)$$

We use the R-package BSTS [37] to fit the Bayesian structural time series with the following model:

$$r_{t+1}^* = r_t^* + \delta_t + \epsilon_t, \epsilon_t \sim N(0, \sigma_\mu),$$

$$\delta_{t+1} = D + \phi(\delta_t - D) + \eta_t, \eta_t \sim N(0, \sigma_\delta).$$

δ_t is the semi-local trend that we model as an auto-regressive process of order one. This gives a trend process that can oscillate around the level D with smaller random-walk excursions. Inverse gamma-priors are used for the standard deviation parameters σ_μ and σ_δ , a Gaussian prior on D and a $N(0, 0.1)$ prior for ϕ . The ϕ parameter determines how much the trend behaves as a random-walk. For $\phi = 1$ the trend follows a random walk, while for $\phi = 0$ the trend is just constant with Gaussian noise. We use a prior for small values of ϕ to keep the model for having very rapid growth in the variance of r^* that would make the model very over-dispersed. A Markow Chain Monte Carlo (MCMC) algorithm is used to estimate the parameters in the model which then allows us to forecast future values of r^* . After the model is fitted for all the values up to r_{t-1}^* we can draw samples of the predictive distribution for r_t^* . After using the inverse transform of Equation 2 to get predictions for R_t

we can use these values to forecast incidence.

3.2.3 Forecasting incidence

Once we have specified our model by specifying the offspring distribution and the method for forecasting R_t we can use the model to generate probabilistic forecasts. As in many other application, the forecast is not available in a closed form, but we can simulate the model to get a sample of values from the predictive distribution. If we want to generate a forecast for I_t we first use all the data up until and including time $t - 1$ to estimate the history of the reproduction number until time. Our probabilistic forecast will be based on sampling possible outcomes to generate a distribution of outcomes. We therefore first draw sample for R_t from the predictive distributions of the models discussed above, then we combine this with the historical incidence data to calculate the total force of infection λ_t . We then sample I_t from the specified offspring distribution with an expected value given by λ_t . If we want to forecast over multiple time-steps we follow the same procedure by sampling values for the reproduction number, now for both R_t and R_{t+1} . When calculating λ_{t+1} we use the sampled value for I_{t+1} together with the historical data I_s for $s \leq t$. This allows us to generate probabilistic forecasts for any forecasting horizon.

3.3 Assessing probabilistic forecasts

The aim of probabilistic forecasts is not only to predict the correct average value, but to forecast and the correct probabilities for all the different possible outcomes. Therefore, when evaluating such forecasts it is not sufficient to use metrics that only depend on a point estimate like the mean or median. To assess the forecasts of the models discussed in the previous section we will follow the paradigm of maximising sharpness of the predictive distribution subject to calibration [11]. In addition we will consider proper scoring rules for comparing probability distributions. We follow the approach taken in [8], where probabilistic forecasts for the West African Ebola outbreak were assessed using similar methods.

3.3.1 Calibration

A forecasting model is calibrated if the forecasted distribution is equal to the actual distribution of observed values. This means that the model can accurately assess its own uncertainty. For example if we predict that it will rain with a 60% chance and we find that over time it does rain 60% of days when we predicted a 60% chance of rain, the model would be well calibrated. Mathematically, if we assume that real distribution of outcomes in nature is given by a cumulative density function, G_t , and our model predicts a cumulative density function, F_t , we say that the forecast is ideal and perfectly calibrated if $F_t = G_t$. To assess calibration we will use a randomised Probability Integral Transformation (PIT) [38]. We consider a forecasted cumulative distribution, F_t , compared to an observed value k_t . For each such pair of forecast and observed value we calculate:

$$u_t = F_t(k_t) + \nu(F_t(k_t) - F_t(k_t - 1)),$$

where ν is a standard uniform random variable. If the prediction is ideal, the u_t s will be distributed as a standard uniform distribution. Once we have calculated a set of u_t values we can then use the Anderson-Darling test of uniformity to assess if the distribution of u_t is uniform. If the p-value for this test is less than 0.1 we can reject that the model is calibrated. We implement this using the goftest [39] R-package. It is important to note that uniform PIT values is a necessary, but not sufficient condition for an ideal forecast and since we only have limited data that we might not be able to reject calibration even if the model is not well calibrated. In addition to assessing calibration, a histogram of PIT values can tell us if the forecast is under or over-dispersed [38]. If the forecast is too dispersed the PIT values are likely to cluster in the centre of the PIT histogram, while if the forecast is under-dispersed they are likely to cluster along the edges of the histogram. We introduce a simple measure of centrality, which is equal to the fraction of u_t values that are between 0.25 and 0.75 as a

way to assess if the forecasts are under or over dispersed if they are not calibrated

$$\text{centrality} = \frac{\text{Number where } 0.25 < u_t < 0.75}{\text{Total number}} - 0.5.$$

When the centrality scores is less than 0, most of the PIT values are outside of central region suggesting that the forecasts are underestimating the real uncertainty. If the centrality score is larger than 0 then the PIT scores are mainly in the central region and this indicates that the forecasts are overestimating the amount of uncertainty. Since the PIT values include a random component, we will for each set of forecasts repeat the calculations of the PIT values 10 times and take averages to average out the effect of the randomness in the definition of u_t [8].

3.3.2 Sharpness and Bias

Sharpness is defined as the range of values in the forecast. The sharper a forecast, the more certain we are of the predicted values. Sharpness depends only on the forecast and not on the observed values. We will use the normalised absolute deviation about the median of y to quantify sharpness

$$S_t(F_t) = \frac{1}{0.675} \text{median}(|I_t - \text{median}(I_t)|),$$

where the normalisation factor means that S_t is equal to the standard deviation if F_t is normal.

It is also of importance to assess the bias of the forecast. Are we more likely to predict too large or too small values? We will quantity bias as

$$B_t(F_t, k_t) = 1 - (F_t(k_t) - F_t(k_t - 1)).$$

If $B_t = 0$ half the probability mass is above and half below the observed value, and the

forecast is unbiased. B_t is always between -1 and 1, with both extreme values signifying a completely biased forecast where all the forecasted values are either smaller or larger than the observed value.

3.3.3 Proper scoring rules

Proper scoring rules have been developed to compare and rank probabilistic forecasts [40]. To be a proper scoring rule, the scoring rule has to minimised when the forecast is ideal. These scoring rules combine calibration and sharpness and give a consistent ranking of forecasts. We will use the continuously ranked probability score (CRPS) and the Dawid-Sebastiani score (DSS) as implemented in the scoringRules R-package [41] for samples from the predictive distribution. The CRPS score is given by

$$CRPS(F_t, k_t) = \int_R (F_t(z) - \mathbf{1}_{k_t \leq z})^2 dz.$$

The CRPS can be viewed as an extension of the mean absolute error for a deterministic forecast. The DSS score only depends on the mean, μ_p and standard deviation, σ_p of the predictive distribution

$$DSS(F_t, k_t) = \left(\frac{k_t - \mu_p}{\sigma_p} \right)^2 + 2 \log \sigma_p.$$

The DSS allows an intuitive understanding of the proper scoring rules. The first term tells us about how far away the observed value is from the mean of the predictive distribution in units of the standard deviation and the second term gives information about the sharpness of the predictive distribution. For our models we will only have samples from the predictive distribution. To calculate the CRPS a kernel density estimate is used to estimate F_t , and for the DSS we use the mean and standard deviation of the sample.

3.4 Implementation

All the models and assessments were implemented in the R programming language [42] and are available as open-source code at http://github.com/gulfa/msc_ebola. We will consider four different models to assess which model fits the data best. The models are:

1. Model 1 (Poisson Latest): Constant reproduction number and Poisson offspring distribution
2. Model 2 (NegBin Latest): Constant reproduction number and negative binomial offspring distribution
3. Model 3 (Poisson Semi-local): Varying reproduction number with a semi-local trend and Poisson offspring distribution
4. Model 4 (NegBin Semi-local): Varying reproduction number with a smell-local trend and negative binomial offspring distribution

For the dispersion parameter k for the negative binomial models we fit a value for the simple negative binomial model (Model 4) by minimising the continuous ranked probability score for the one day ahead predictions for the model on the level of the whole outbreak.

We will run the models both on the national level using data from the entire outbreak and on the sub-national level. For sub-national predictions on the health zone level we run the model on each health zone separately using only data from that health zone. This corresponds to an assumption that once a case has been imported into a health zone the outbreak in the health zone is independent of the rest of the outbreak.

We will assess how well the models work both for the entire epidemic and for each health zone. To evaluate a model we will estimate the calibration, sharpness, bias, DSS score and CRPS for daily forecasts with forecasting horizons up to 28 days. To do this we start 17 days¹ after the first case of the outbreak in the given location and calculate the d -days ahead

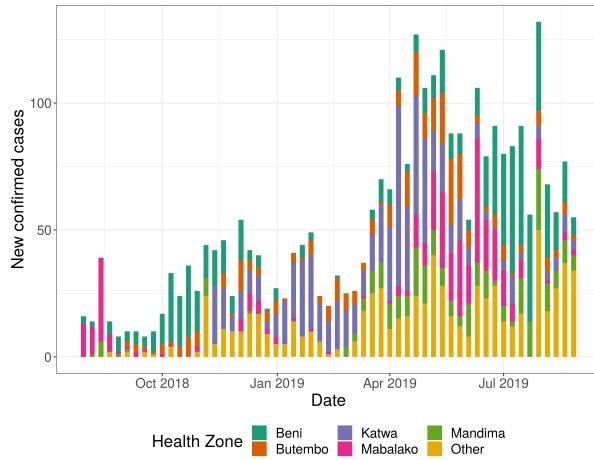
¹The 17 days are due the fact that method for estimating the reproduction number needs 17 days to give reasonable values for the reproduction number

prediction for all historically available data. For the calibration we use all forecasts for the d -days ahead forecasts compared to the observed values the values to assess if PIT values are uniformly distributed, while for all the other metrics we average them over all the available start days for the forecast.

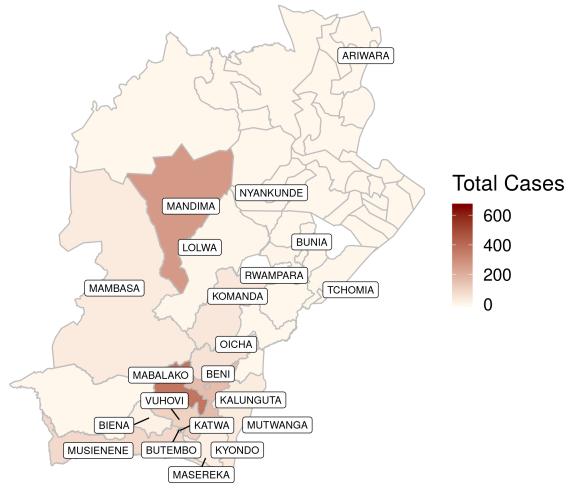
4 Results

A total of 2,926 confirmed Ebola cases and 1,926 confirmed Ebola deaths have been reported in the ongoing outbreak in DRC as of the 31st of August, giving a case fatality ratio of 66%. Figure 2a shows the weekly number of confirmed cases and the map in Figure 2b shows the total number of cases from each health zone. From Figure 2a and the estimate of the instantaneous reproduction number in Figure 2c we can see the evolution of the outbreak. After the initial period with varying weekly case numbers, there was an increase in cases from October 2018 with a large reproduction number, followed by a more varied period where the reproduction number varied around one with some waves of larger weekly incidence around December 2018 and February 2019. From late March and early April 2019 the number of cases per week increased significantly as the reproduction number rapidly increased to over two. The larger number of weekly cases continued in June and July even as the reproduction number decreased. From July 2019 the reproduction has varied around one, indicating that the outbreak is still not under control, but that the intensity of the outbreak is less than in May and June. There are large regional differences with Beni and Katwa health zones being hardest hit with close to half the cases between them. From the epi-curve there seems to be some connection between the large increases in incidence and introduction or reintroduction into specific health zones. The outbreak started mainly in Mabalako, then in October 2018 the peak was due to increased transmission in Beni, the peak of incidence in February 2019 was mainly dominated by cases in the health zone of Katwa. The large increase of cases starting around March/April saw an increasing number of health zones having many new

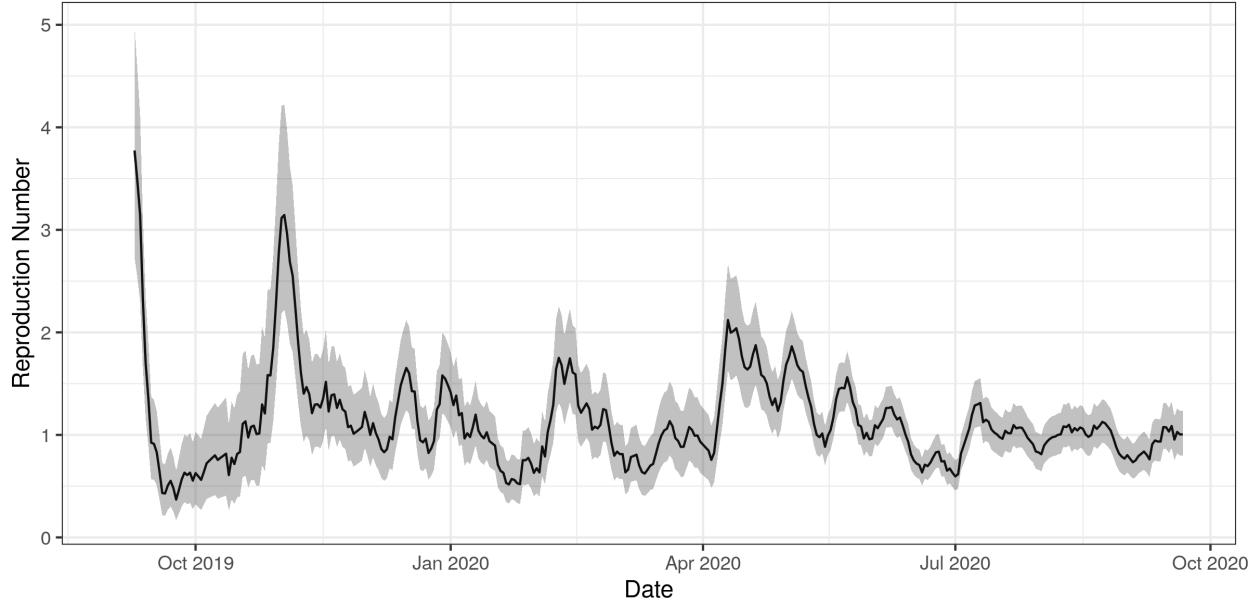
cases. From this it is clear that the hot-spot of the disease is moving from health zone to health zone.



(a) Number of new confirmed ebola cases by week and by health zone.



(b) Total number of confirmed cases by health zone



(c) Mean estimated reproduction number as a function of time with the shaded region showing the 95% credible interval.

Figure 2: The three figures show the number of cases and the reproduction number as a function of time and health-zone.

4.1 National level forecasts

The first step in modelling the outbreak on the national level was to find the best fitting value of the dispersion parameter for the negative binomial models. We found that the dispersion parameter that gives the smallest CRPS score for one day ahead predictions for the negative binomial models was $k \approx 8$, but a fairly large range of parameters gave a reasonable fit. We therefore chose to use $k = 8$ for the negative binomial distributions both for the national and health zone forecasts.

All four of the models were successfully fitted on incidence data from 395 days of national data and then used to provide short term forecasts. The calibration and proper scoring rules scores for the four models for different forecasting horizons can be seen in Table 1 for weekly forecasting horizons up to four weeks. We also show the evolution of calibration, bias, CRPS and the centrality of PIT values as a function of the forecasting horizon in Figure 3.

Model	Horizon	Sharpness	Bias	CRPS	DSS	Centrality	Calibration
NegBin Semilocal	1	4.08	-0.03	1.87	3.30	0.04	0.20
NegBin Latest	1	4.14	-0.00	1.82	3.26	0.06	0.16
Poisson Semilocal	1	2.93	0.01	1.87	3.32	-0.07	0.00
Poisson Latest	1	2.97	0.05	1.82	3.24	-0.08	0.00
NegBin Semilocal	7	4.83	-0.07	1.80	3.41	0.16	0.00
NegBin Latest	7	4.33	0.01	2.36	3.94	-0.03	0.00
Poisson Semilocal	7	4.19	-0.02	1.72	3.23	0.09	0.01
Poisson Latest	7	3.08	0.06	2.47	4.31	-0.14	0.00
NegBin Semilocal	14	5.98	-0.10	2.56	4.62	0.13	0.00
NegBin Latest	14	4.62	0.02	2.82	4.74	-0.08	0.00
Poisson Semilocal	14	5.78	-0.06	2.52	4.54	0.10	0.00
Poisson Latest	14	3.28	0.09	2.99	5.32	-0.22	0.00
NegBin Semilocal	21	7.55	-0.13	3.24	5.69	0.15	0.00
NegBin Latest	21	5.12	0.01	4.07	5.57	-0.19	0.00
Poisson Semilocal	21	7.67	-0.10	3.26	5.88	0.15	0.00
Poisson Latest	21	3.53	0.06	4.54	6.85	-0.26	0.00
NegBin Semilocal	28	9.14	-0.16	4.77	7.43	0.17	0.00
NegBin Latest	28	5.87	0.03	5.30	7.38	-0.20	0.00
Poisson Semilocal	28	9.64	-0.13	4.89	7.12	0.16	0.00
Poisson Latest	28	3.88	0.08	6.00	9.08	-0.28	0.00

Table 1: Model evaluations for predictions with different forecasting horizons for the national level.

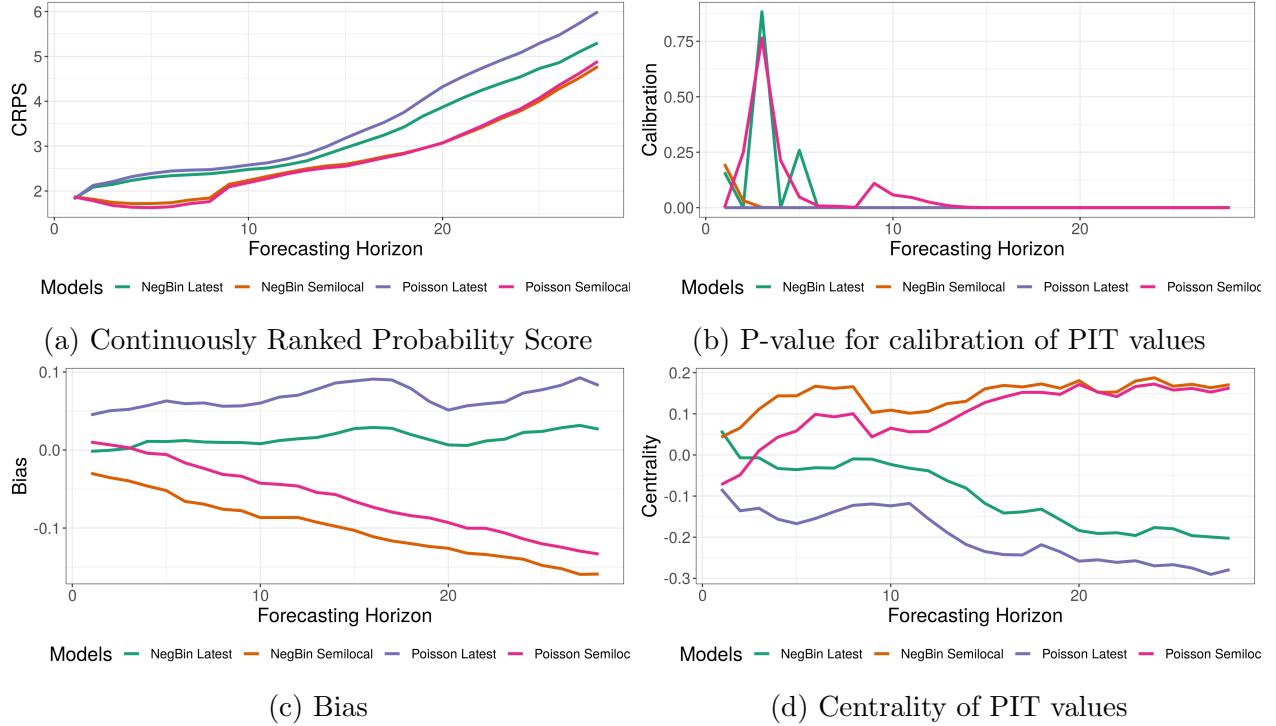


Figure 3: Evaluation of the four models as a function of forecasting horizon at the national level.

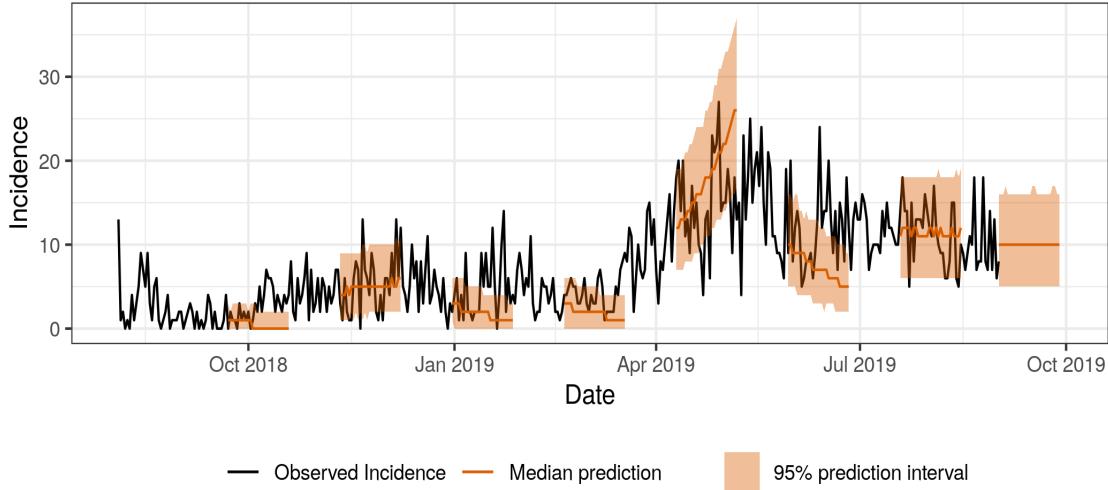
Together they show that for next day forecasts, the negative binomial offspring distribution is needed for a calibrated forecast as only the two negative binomial models are calibrated with a p-value larger than 0.1. We can also see that there are no big differences for one-day ahead forecasts based on the method of forecasting the reproduction number. For longer forecasting horizons the Poisson distribution with the semi-local time series prediction for R_t is the best model and is calibrated up until a maximum of nine days even if it is not well calibrated for all the days before. At very short forecasting horizons it is uncalibrated as it is underestimating its own uncertainty. This is the only model that is calibrated for longer forecasting horizons and it also has the lowest CRPS and DSS scores. The Poisson model with constant reproduction number could not provide any calibrated forecasts, the negative binomial model with constant reproduction number was calibrated during the first 5 days and the negative binomial model with semi-local trend in the reproduction number was only calibrated for the first day.

The centrality of the PIT scores in Figure 3d give us extra information about why the models become uncalibrated. A negative centrality means that the model has too little uncertainty and we can see that this is the case for the two models based on using the latest reproduction number. The two models using the semi-local trend for the reproduction number are closer to zero centrality, but they have too much uncertainty since their centrality score is above zero. For the same method of forecasting the reproduction number, the negative binomial models always have larger uncertainty as we can also see from the sharpness figures. This is why the Poisson model is preferred at larger forecasting horizons. After about one week the sharpness of the Poisson semi-local model has grown to be the same as the negative binomial latest model. We can see how the semi-local models quickly become less sharp, while the sharpness for the constant models does change as much with forecasting horizon. We can also see that the two semi-local models have a small negative bias indicating that on average they underestimate the incidence.

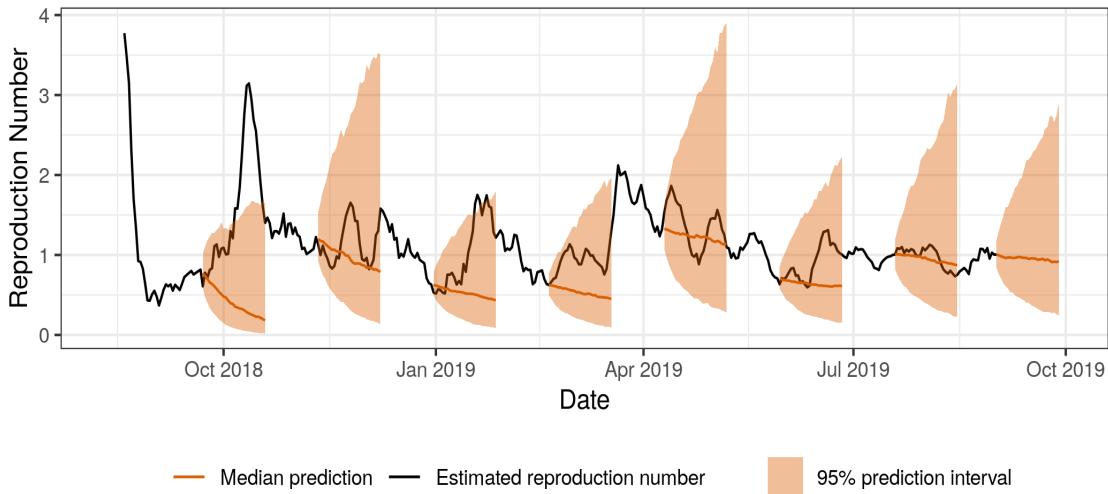
In Figure 4 we show 28 day forecasts every 50 days for the semi-local Poisson model and the corresponding forecasts for the reproduction number. These results show the same pattern as the more formal evaluations. In the short term the model works quite well, but that the amount of uncertainty often seems a bit too large for the reproduction number forecasts. We also forecast the incidence for the 28 days following the 1st of September 2019 in Figure 4. Here the model predicts a slightly decreasing median reproduction number and a fairly steady incidence through the month of September. We expect between 10–15 cases per day over the next month.

4.2 Sub-national forecasts

We fitted all four models separately for each health zone that had more than one case of Ebola. The performance of the models varied significantly from health zone to health zone. Table 2 shows the best model(s) for each health zone, how large a forecasting horizon we can have without ruling out calibration and the total number of confirmed cases in the health



(a) Forecasted and predicted incidence for the semi-local Poisson model



(b) Forecasted and predicted reproduction numbers for the semi-local Poisson model

Figure 4: Median forecast with 95% prediction intervals and observed values for incidence and reproduction number for the semi-local Poisson model

zone. There is a large difference from health zones where the forecasts are calibrated for 28 days, to health zones where not even the one day ahead predictions are calibrated. Even if the model is calibrated at 28 days, it does not mean that it is calibrated for all forecasting horizons less than 28 days.

For all health zones apart from two, we find that the best model is a model with a semi-local trend in the reproduction number. Sometimes the Poisson models are better and

Location	Largest Horizon	Best model	Cases
Kayna	28	NegBin Semilocal	22
Nyiragongo	28	NegBin Semilocal	3
Tchomia	28	NegBin Semilocal	2
Lolwa	26	All models	3
Mwenga	18	All models	6
Bunia	17	NegBin Semilocal	4
Alimbongo	12	NegBin Semilocal	5
national	9	Poisson Semilocal	2942
Kayina	7	Poisson Semilocal	10
Rwampara	6	NegBin Latest	8
Lubero	4	NegBin Semilocal	31
Musienene	4	Poisson Semilocal	84
Kyondo	3	Poisson Semilocal	22
Vuhovi	3	Poisson Semilocal	103
Beni	2	Poisson Semilocal	661
Biena	2	Poisson Latest	16
Butembo	2	NegBin Semilocal	279
Mabalako	2	NegBin Semilocal	371
Kalunguta	NA	No calibrated model	164
Katwa	NA	No calibrated model	647
Mangurujipa	NA	No calibrated model	20
Masereka	NA	No calibrated model	50
Mutwanga	NA	No calibrated model	31
Oicha	NA	No calibrated model	55
Komanda	NA	No calibrated model	43
Mambasa	NA	No calibrated model	32
Mandima	NA	No calibrated model	264

Table 2: For each health zone we show the maximal forecasting horizon where we can not exclude calibration at the $p=0.1$ level. If multiple models are equally calibrated we choose the one with smallest CRPS. For some health zones there were no calibrated forecasts

sometimes the negative binomial models are better.

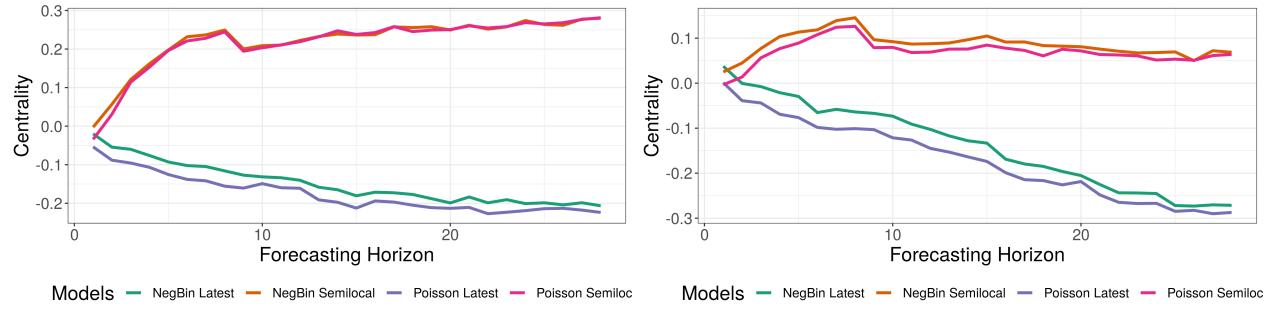
The health zones where we get very well calibrated models at 28 days are health zones with small bursts of Ebola activity followed by long periods of no cases. For the larger health zones where there is more sustained transmission the models work for some health zones, but not as well on the national level. There are also some health zones with many cases where the models do not work well. We show the predictions and the centrality measures for the Beni and Katwa health zones in Figure 5. The models and evaluations for all the

other health zones can bee seen in the Appendix. For both Beni and Katwa the main reason the models are not well calibrated is that they overestimated their own uncertainty. This is especially the case in Katwa and explains why even the one-day ahead prediction is not well calibrated.

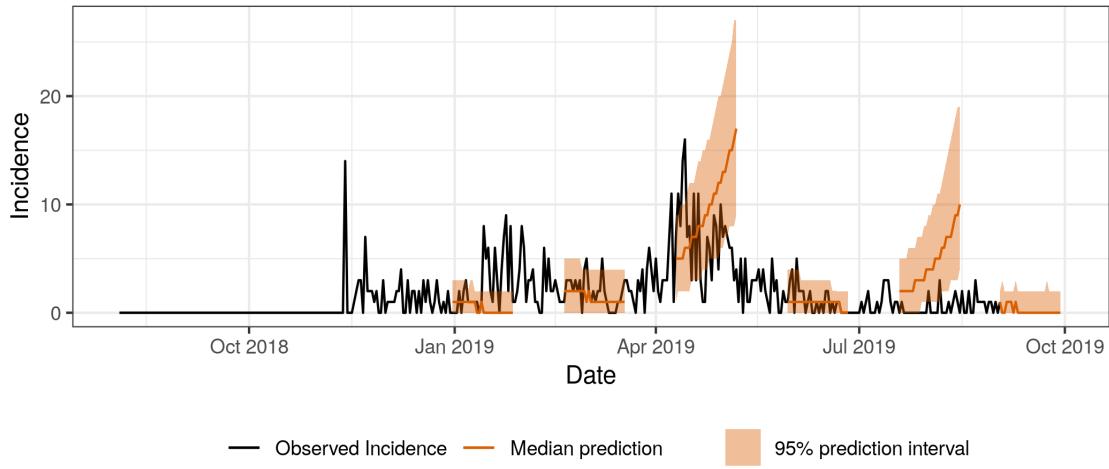
As a final check of the health zones predictions, we compared the one-day ahead predictions of the national level model with the one-day ahead predictions of a sum of the semi-local Poisson model for all the health zones in Figure 6. The predictions are fairly close, but when looking at the calibration we find that the combined model is not calibrated and is also biased to predict too large incidence. The centrality scores for the combined models show that it actually underestimates the uncertainty at large forecasting horizons.

5 Discussion

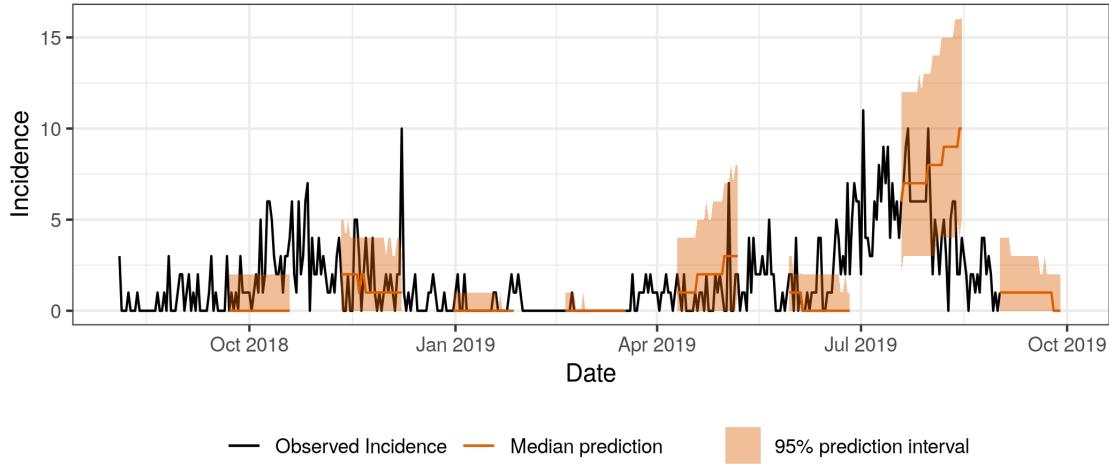
The ongoing Ebola outbreak in northern DRC is the second largest Ebola outbreak in history and is of significant regional and international concern. As of the 1st of September 2019, the outbreak is not under control as we can see both from the fact the reproduction number is not below one and that there are still plenty of new cases not on contact lists and community deaths [43]. In addition to the large number of confirmed Ebola cases, this outbreak has seen a high CFR of 66% compared to 51% in the west African outbreak [44]. Both the long duration, difficulty of control and high CFR are likely due to the difficult context for the outbreak response. This part of DRC has a history of violence and mistrust of the central government and international organisations. This makes all parts of the response from finding cases, tracking contacts, ensuring early treatment and safe burials much more difficult. Based on the models investigated in this paper it seem like the large increase of cases seen from March - June 2019 is over and that there is hope that the outbreak can be controlled, but the large uncertainty associated with the outbreak which is also reflected in the large uncertainty in the models make long term predictions very difficult if not impossible.



(a) Centrality of PIT values for the models in (b) Centrality of PIT values for the model Beni
Katwa health zone Beni health zone



(c) Median forecast with 95% prediction intervals and observed values for incidence for the Poisson semi-local model in Katwa



(d) Median forecast with 95% prediction intervals and observed values for incidence for the Poisson semi-local model in Beni

Figure 5: Evaluations and prediction for Katwa and Beni

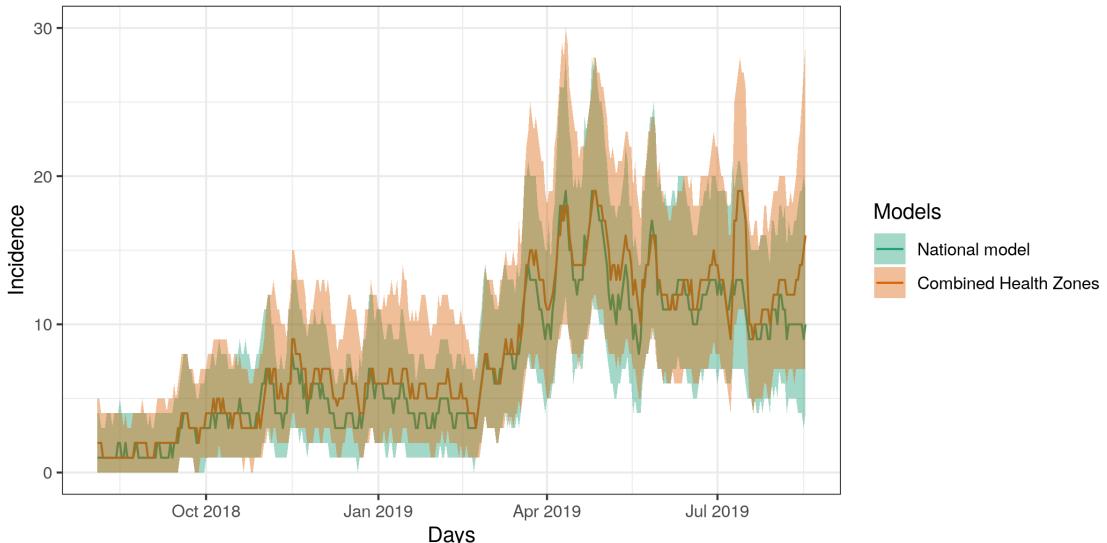


Figure 6: Median and 95% prediction intervals for one-day ahead predictions from the Poisson semi-local model for national level compared with the same model when summing up predictions from all health zones

Compared to estimates of the reproduction number in the West-African Ebola outbreak [45] the reproduction number in the DRC outbreak has seen larger variations on short time-scales and less clear patterns over longer time-scales. In the previous outbreak, both Guinea and Liberia saw that the reproduction number peaked around two after a few months followed by a long and slow decline. The current outbreak has been characterised by multiple short term peaks or waves of higher reproduction number with the peaks reaching over three at some points. It is possible that the different patterns from the different outbreaks can be explained by the characteristics of the outbreak response. The West African response was slower to get properly started, but once it did the response was fairly uninterrupted. In the current outbreak the response had a faster start, but due to the security situation the response has been interrupted and this could be why we see the short term peaks. The range of reproductive values seen is also similar to other previous Ebola outbreaks [46].

We find that on the national level the models with a semi-local trend for the reproduction number were calibrated out to about a forecasting horizon of 9 days and significantly outperformed the simpler models based on a non-changing reproduction number. The length

of the forecasting horizon is similar, but a bit shorter than the 2 weeks found in [8] where a semi-mechanistic model based on a SEIR model that included reporting delays. That model was also semi-mechanistic as the transmission rate was allowed to vary as a random walk. While direct comparisons are not possible due to the different outbreaks, it would be of interest if our model with much less deterministic structure can give fairly similar calibrated forecasting horizons. Other models have been used to forecast the current Ebola outbreak, e.g [47, 48]. They use very different modelling techniques and a proper comparison between the models would require comparing proper scoring rules score using the same data. All the models presented in this paper can easily be run daily on a normal laptop to provide real-time predictions in an outbreak setting.

The best model for the national outbreak is uncalibrated when forecasting more than 9 days forward due to overestimating the amount of uncertainty in the incidence. Depending on how the forecasts are used, it is possible that uncalibrated but over-dispersed forecasts still can be useful for planning purposes since they are conservative. An upper limit for an over-dispersed forecasts will give a conservative upper limit for future incidence. Forecasts based on under-dispersed models, require a lot more care and should be treated with caution. Models that only provide point-predictions or do not actively try to quantify the uncertainty are by definition underdispersed. The main reason for the overdispersion in the semi-local models is likely due to the forecasting of the reproduction number. This model allows large variation in the reproduction number over fairly short time, but does not have any information about what range of reproduction numbers are actually feasible. In this outbreak we see fairly fast changes of the reproduction number, but the reproduction number almost never is larger than three. A model for reproduction number than allows this fast variation and short peaks, but that keeps the reproduction number within a more narrow band, might provide better forecasts.

A probabilistic modelling approach with tests for calibration and using proper scoring rules to see which forecast is better is a crucial step to compare forecasts and to evaluate

their validity. The proper scoring rules are mainly useful when comparing models, while the calibration gives an absolute measure for how well the forecast captures the uncertainty seen in the process. In this study we found these tools crucial both for model selection and for the final interpretation of our model. Once the forecasting horizon is longer than 9 days we can not trust that the model adequately predicts the uncertainty in the process. Even when considering a well calibrated model, it can make large forecasting errors if the underlying process changes in a way not seen before during the time period used to evaluate the model. Therefore it is always crucial to critically evaluate both the models and the metrics used to evaluate the models. We find that both the CRPS and DSS scores give similar performance, but that calibration using the PIT values was especially valuable as an absolute measure of calibration. The introduction of the centrality measure of the PIT histograms give additional valuable information and allows us to understand better how our models are failing if they are not well calibrated. The centrality is a fairly crude measure of this and a measure that also includes other departures from uniformity of the PIT values could give additional information.

We find that for one-day ahead forecasts the negative binomial distribution is needed for a calibrated forecast. This indicates that a Poisson offspring distribution where everyone is always equally likely to transmit the disease in every small time step does not adequately describe the transmission process. We found a relatively large value for the k parameter which gives a fairly modest increase in variance that is much smaller than found when investigating transmission chains in the West African Ebola outbreak [34]. It is possible that variance is smaller in this outbreak, but this should be taken as fairly weak evidence of that and a proper investigation of transmission chains would be needed.

We could provide forecasts for all health-zones with more than one cases, but the quality was varying. For some health zones where there were very few cases and then longer period with no transmission, the model managed fairly well, but for most health zones with a large number of cases the models were again too over-dispersed to provide calibrated predictions.

Once our model has seen no new infections in a few weeks they will predict close to zero new cases with essentially zero sharpness irrespective of the current reproduction number. This behaviour works well when there are no new imported cases as we can see in some of these health zones. Some of the well calibrated health zone forecasts only saw their first ebola case quite late in the outbreak, giving many fewer data points to test the model. This can make it hard to reject calibration even if the model is uncalibrated. For almost all health zones the models with semi-local trend were better than the models with constant reproduction number.

The models and forecasts in this report have multiple limitations. Firstly, using a data source that only includes the reported number of new cases every day has many disadvantages as compared to a line list of all cases. A line list would give a much better idea of the date of onset of disease, it would remove the problem with negative incidence rates after corrections and any problems with having to interpolate incidence rates. It is a strength of the modelling approach that it can be used on less than ideal data to still give reasonable short-term forecasts. The stringent tests of calibration can give us confidence in the forecasts even without perfect data. This is of key importance in outbreak situations where having up-to-date line lists of cases can be difficult or impossible. In these situation we could still hope that the approaches in this report could give reasonable forecasts.

The models in their current form also have several limitations. As previously discussed, the semi-local trend models seem to predict too much variance in the reproduction number which leads to over-dispersed predictions. There are also clear limitations when the models are used on the health-zone level since the spread between health zones is not modelled. Each health zone is modelled independently as if it were the whole outbreak. We only model the spread of the disease after it has been introduced to a health zone. For the national level we do not have this limitation, but it makes it very difficult to use the current model to study the spatial spread of the disease. Due to the flexibility of the model structure it would be fairly straight-forward to improve the models by incorporated an additional force

of infection term that gives the force of infection from outside the health zone. There are multiple options for how to model the geographic dependence that would all easily fit into our model structure, this includes spread from adjacent health zones, using a gravity model where the amount of spread between health zones is based on the populations [49] or, if available, data on inter health-zone travel could be used. Without this spread term, the model can not be used to forecast probabilities for the spread of Ebola to new health zones. In addition the current way of estimating the time-varying reproduction number requires at least 17 days of data, so we need at least 17 days of data in a health zone to be able to forecast.

In addition to the spatial spread, it would be beneficial to combine the separate parts of the model together into one generative model. This means including the estimation of the reproduction number into the modelling of the incidence. One promising way of doing this is a full Bayesian generative model. One of the main benefits of a generative model is that it models the whole data generation process and would allow us to simulate whole outbreaks. For future work it would also be very interesting to investigate different models for the time evolution of the reproduction number. A full Bayesian treatment would also allow us to model uncertainty in the parameters of the offspring distribution, uncertainty about serial interval and could even include imputation for missing data using the general modelling framework presented here.

In this report we used a semi-mechanistic model with fairly few assumptions about the exact mechanism of disease transmission. The flexibility of the reproduction number in this model allows us to fit complicated incidence histories that could not have been easily fit by traditional compartmental models. The benefit of the structure we do have in the model is that certain basic facts about how disease spread that are likely to apply to a large range of diseases. The model used in this report could easily be adopted to other diseases and other contexts without any real changes in the model structure. We would need a different serial interval, but the rest of the model could stay the same. The benefit of adding more

structure to a model, for example a SEIR structure is that it is easier to model the results of interventions like vaccination ². One key aspect to consider when choosing models for an outbreak situation is that maybe more than one model should be used during the outbreak. At the start of the outbreak when the important mechanisms of disease transmission and population behaviours are less well known, maybe a model with less mechanistic structure is the most useful. Once we learn more about the outbreak we can add more structure to the model. In many forecasting challenges it has been shown that an ensemble of models perform better than individual models [50], this was also found in the RAPPID modelling competition for Ebola [23]. So combining multiple models with different strengths could give a better overall forecast. For example, a combination of the negative binomial model with a constant reproduction number which did well for early forecasts with the Poisson semi-local model that did better for slightly longer forecasts could give an even better model.

Theoretically it should be possible to forecast infectious disease outbreaks on the timescale of individual outbreaks [51]. The models in this report and other disease forecasting models still have a long way to go to achieve this goal. For the current Ebola outbreak many additional sources of data exists that are not used in models, this includes the deaths, community deaths versus deaths in treatment centres, contacts followed-up, vaccinations, data on response activities etc. It is possible that combining these additional data sources together with improved modelling techniques we could significantly increase forecasting horizon where we can provide calibrated predictions.

6 Conclusions

We presented a probabilistic disease modelling and forecasting framework based on branching process with a time varying reproduction number. When modelling the reproduction number with a semi-local trend, the models could give well calibrated forecasts for the 2018–2019

²Using the fact that $R_t = R_o * s(t)$ in a compartmental model of the SEIR type, it would likely be possible to also model some simple interventions in the models considered in this report.

Ebola outbreak in DRC for forecasting horizons of nine days. The models provide easy to interpret forecasts of incidence and reproduction number and could also provide some reasonable forecasts for incidence at the sub-national level.

References

- [1] H. Heesterbeek, R. M. Anderson, V. Andreasen, S. Bansal, D. D. Angelis, C. Dye, K. T. D. Eames, W. J. Edmunds, S. D. W. Frost, S. Funk, T. D. Hollingsworth, T. House, V. Isham, P. Klepac, J. Lessler, J. O. Lloyd-Smith, C. J. E. Metcalf, D. Mollison, L. Pellis, J. R. C. Pulliam, M. G. Roberts, C. Viboud, and I. N. I. I. Collaboration, “Modeling infectious disease dynamics in the complex landscape of global health,” *Science*, vol. 347, p. aaa4339, Mar. 2015.
- [2] C. Rivers, J.-P. Chretien, S. Riley, J. A. Pavlin, A. Woodward, D. Brett-Major, I. M. Berry, L. Morton, R. G. Jarman, M. Biggerstaff, M. A. Johansson, N. G. Reich, D. Meyer, M. R. Snyder, and S. Pollett, “Using outbreak science to strengthen the use of models during epidemics,” *Nature Communications*, vol. 10, pp. 1–3, July 2019.
- [3] J.-P. Chretien, S. Riley, and D. B. George, “Mathematical modeling of the West Africa Ebola epidemic,” *eLife*, vol. 4.
- [4] W. E. R. Team, “Ebola Virus Disease in West Africa The First 9 Months of the Epidemic and Forward Projections,” *New England Journal of Medicine*, vol. 371, pp. 1481–1495, Oct. 2014.
- [5] A. Camacho, A. Kucharski, Y. Aki-Sawyerr, M. A. White, S. Flasche, M. Baguelin, T. Pollington, J. R. Carney, R. Glover, E. Smout, A. Tiffany, W. J. Edmunds, and S. Funk, “Temporal Changes in Ebola Transmission in Sierra Leone and Implications for Control Requirements: a Real-time Modelling Study,” *PLoS Currents*, vol. 7, Feb. 2015.
- [6] P.-Y. Kobres, J. P. Chretien, M. A. Johansson, J. Morgan, P.-Y. Whung, H. Mukundan, S. D. Valle, B. M. Forshey, T. M. Quadelacy, M. Biggerstaff, C. Viboud, and S. Pollett, “A systematic review and evaluation of Zika virus forecasting and prediction research

during a public health emergency of international concern,” *bioRxiv*, p. 634832, May 2019.

- [7] J.-P. Chretien, D. George, J. Shaman, R. A. Chitale, and F. E. McKenzie, “Influenza Forecasting in Human Populations: A Scoping Review,” *PLOS ONE*, vol. 9, p. e94130, Apr. 2014.
- [8] S. Funk, A. Camacho, A. J. Kucharski, R. Lowe, R. M. Eggo, and W. J. Edmunds, “Assessing the performance of real-time epidemic forecasts: A case study of Ebola in the Western Area region of Sierra Leone, 2014–15,” *PLOS Computational Biology*, vol. 15, p. e1006785, Feb. 2019.
- [9] W. Wei and L. Held, “Calibration tests for count data,” *TEST*, vol. 23, pp. 787–805, Dec. 2014.
- [10] T. Gneiting, “Editorial: Probabilistic forecasting,” *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, vol. 171, no. 2, pp. 319–321, 2008.
- [11] T. Gneiting, F. Balabdaoui, and A. E. Raftery, “Probabilistic forecasts, calibration and sharpness,” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 69, no. 2, pp. 243–268, 2007.
- [12] C. Czado, T. Gneiting, and L. Held, “Predictive Model Assessment for Count Data,” *Biometrics*, vol. 65, no. 4, pp. 1254–1261, 2009.
- [13] World Health Organisation, “Ebola virus disease fact sheet.”
- [14] J. Brainard, L. Hooper, K. Pond, K. Edmunds, and P. R. Hunter, “Risk factors for transmission of Ebola or Marburg virus disease: a systematic review and meta-analysis,” *International Journal of Epidemiology*, vol. 45, pp. 102–116, Feb. 2016.
- [15] World Health Organization, “Ebola Outbreak in DRC 01: 7 August 2018,” Aug. 2018.

- [16] World Health Organisation, “Ebola Outbreak in DRC 09: 04 October 2018,” Oct. 2018.
- [17] World Health Organization, “Ebola Outbreak in DRC 16: 21 November 2018,” Nov. 2018.
- [18] World Health Organization, “Ebola Outbreak in DRC 38: 24 April 2019,” Apr. 2019.
- [19] W. H. Organization, “Preliminary results on the efficacy of rVSV-ZEBOV-GP Ebola vaccine using the ring vaccination strategy in the control of an Ebola outbreak in the Democratic Republic of the Congo: an example of integration of research into epidemic response.,” tech. rep., Apr. 2019.
- [20] National Institute of Allergy and Infectious Diseases, “Independent Monitoring Board Recommends Early Termination of Ebola Therapeutics Trial in DRC Because of Favorable Results with Two of Four Candidates,” Aug. 2019.
- [21] World Health Organization, “Ebola Outbreak in DRC 51: 23 July 2019,” July 2019.
- [22] World Health Organization, “Ebola Outbreak in DRC 58: 10 September 2019,” July 2019.
- [23] C. Viboud, K. Sun, R. Gaffey, M. Ajelli, L. Fumanelli, S. Merler, Q. Zhang, G. Chowell, L. Simonsen, and A. Vespignani, “The RAPIDD Ebola Forecasting Challenge: Synthesis and Lessons Learnt,” *Epidemics*, vol. 22, pp. 13–21, Mar. 2018.
- [24] R. H. Gaffey and C. Viboud, “Application of the CDC EbolaResponse Modeling tool to disease predictions,” *Epidemics*, vol. 22, pp. 22–28, Mar. 2018.
- [25] B. Pell, Y. Kuang, C. Viboud, and G. Chowell, “Using phenomenological models for forecasting the 2015 Ebola challenge,” *Epidemics*, vol. 22, pp. 62–70, Mar. 2018.
- [26] M. M, V. A, Z. Singh, A. J. Purty, J. Bazroy, and S. Kannan, “Forecasting the trend in cases of Ebola virus disease in West African countries using auto regressive integrated

- moving average models,” *International Journal Of Community Medicine And Public Health*, vol. 3, pp. 615–618, Feb. 2017.
- [27] J. Asher, “Forecasting Ebola with a regression transmission model,” *Epidemics*, vol. 22, pp. 50–55, Mar. 2018.
- [28] A. A. King, M. Domenech de Cells, F. M. G. Magpantay, and P. Rohani, “Avoidable errors in the modelling of outbreaks of emerging pathogens, with special reference to Ebola,” *Proceedings of the Royal Society B: Biological Sciences*, vol. 282, p. 20150347, May 2015.
- [29] Hummanitarian Data Exchange, “Ebola Cases and Deaths in the North Kivu Ebola Outbreak in the Democratic Republic of the Congo (DRC) - Humanitarian Data Exchange.”
- [30] C. Jacob, “Branching Processes: Their Role in Epidemiology,” *International Journal of Environmental Research and Public Health*, vol. 7, pp. 1186–1204, Mar. 2010.
- [31] A. Cori, N. M. Ferguson, C. Fraser, and S. Cauchemez, “A new framework and software to estimate time-varying reproduction numbers during epidemics,” *American Journal of Epidemiology*, vol. 178, pp. 1505–1512, Nov. 2013.
- [32] J. O. Lloyd-Smith, S. J. Schreiber, P. E. Kopp, and W. M. Getz, “Superspreading and the effect of individual variation on disease emergence,” *Nature*, vol. 438, p. 355, Nov. 2005.
- [33] P. Nouvellet, A. Cori, T. Garske, I. M. Blake, I. Dorigatti, W. Hinsley, T. Jombart, H. L. Mills, G. Nedjati-Gilani, M. D. Van Kerkhove, C. Fraser, C. A. Donnelly, N. M. Ferguson, and S. Riley, “A simple approach to measure transmissibility and forecast incidence,” *Epidemics*, vol. 22, pp. 29–35, Mar. 2018.

- [34] International Ebola Response Team, J. Agua-Agum, A. Ariyarajah, B. Aylward, L. Bawo, P. Bilivogui, I. M. Blake, R. J. Brennan, A. Cawthorne, E. Cleary, P. Clement, R. Conteh, A. Cori, F. Dafae, B. Dahl, J.-M. Dangou, B. Diallo, C. A. Donnelly, I. Dorigatti, C. Dye, T. Eckmanns, M. Fallah, N. M. Ferguson, L. Fiebig, C. Fraser, T. Garske, L. Gonzalez, E. Hamblion, N. Hamid, S. Hersey, W. Hinsley, A. Jambei, T. Jombart, D. Kargbo, S. Keita, M. Kinzer, F. K. George, B. Godefroy, G. Gutierrez, N. Kannan-garage, H. L. Mills, T. Moller, S. Meijers, Y. Mohamed, O. Morgan, G. Nedjati-Gilani, E. Newton, P. Nouvellet, T. Nyenswah, W. Pereira, D. Perkins, S. Riley, G. Rodier, M. Rondy, M. Sagrado, C. Savulescu, I. J. Schafer, D. Schumacher, T. Seyler, A. Shah, M. D. Van Kerkhove, C. S. Wesseh, and Z. Yoti, “Exposure Patterns Driving Ebola Transmission in West Africa: A Retrospective Observational Study,” *PLoS medicine*, vol. 13, p. e1002170, Nov. 2016.
- [35] R. N. Thompson, J. E. Stockwin, R. D. van Gaalen, J. A. Polonsky, Z. N. Kamvar, P. A. Demarsh, E. Dahlqwist, S. Li, E. Miguel, T. Jombart, J. Lessler, S. Cauchemez, and A. Cori, “Improved inference of time-varying reproduction numbers during infectious disease outbreaks,” *Epidemics*, p. 100356, Aug. 2019.
- [36] A. Cori, *EpiEstim: EpiEstim: a package to estimate time varying reproduction numbers from epidemic curves*. 2013.
- [37] S. L. Scott, *bsts: Bayesian Structural Time Series*. 2019.
- [38] C. Czado, T. Gneiting, and L. Held, “Predictive Model Assessment for Count Data,” *Biometrics*, vol. 65, no. 4, pp. 1254–1261, 2009.
- [39] J. Faraway, G. Marsaglia, J. Marsaglia, and A. Baddeley, *goftest: Classical Goodness-of-Fit Tests for Univariate Distributions*. 2017.

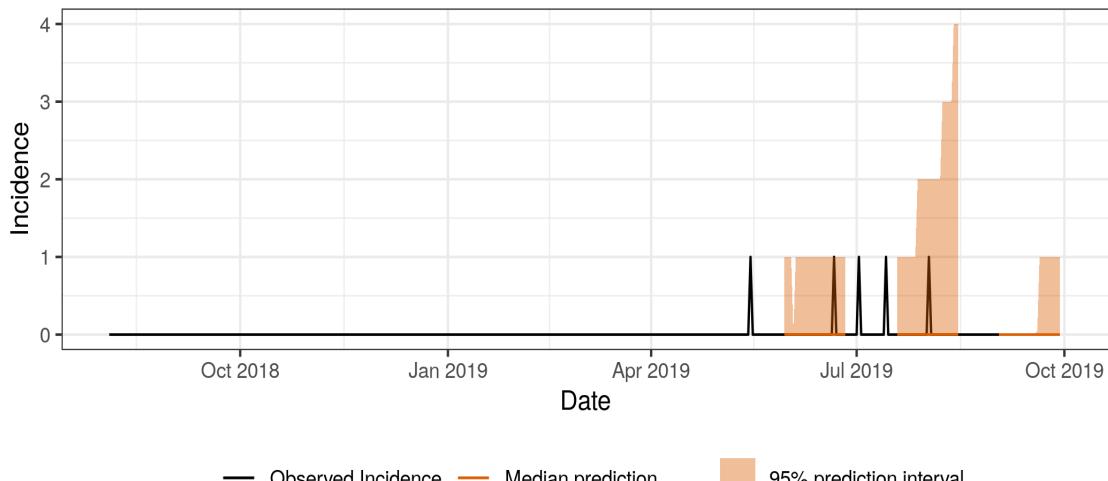
- [40] T. Gneiting and A. E. Raftery, “Strictly Proper Scoring Rules, Prediction, and Estimation,” *Journal of the American Statistical Association*, vol. 102, pp. 359–378, Mar. 2007.
- [41] A. Jordan, F. Krueger, and S. Lerch, “Evaluating Probabilistic Forecasts with scoringRules,” *Journal of Statistical Software*, 2018.
- [42] R Core Team, *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2018.
- [43] World Health Organisation, “Disease Outbreak News 12th of September : Ebola virus disease Democratic Republic of the Congo.”
- [44] A. M. Rojek, A. Salam, R. J. Ragotte, E. Liddiard, A. Elhussain, A. Carlqvist, M. Butler, N. Kayem, L. Castle, L. Odondi, K. Stepniewska, and P. W. Horby, “A systematic review and meta-analysis of patient data from the west Africa (2013-16) Ebola virus disease epidemic,” *Clinical Microbiology and Infection*, vol. 0, July 2019.
- [45] “West African Ebola Epidemic after One Year Slowing but Not Yet under Control,” *New England Journal of Medicine*, vol. 372, pp. 584–587, Feb. 2015.
- [46] J. Legrand, R. F. Grais, P. Y. Boelle, A. J. Valleron, and A. Flahault, “Understanding the dynamics of Ebola epidemics,” *Epidemiology & Infection*, vol. 135, pp. 610–621, May 2007.
- [47] J. D. Kelly, J. Park, R. J. Harrigan, N. A. Hoff, S. D. Lee, R. Wannier, B. Selo, M. Mossoko, B. Njoloko, E. Okitolonda-Wemakoy, P. Mbala-Kingebeni, G. W. Rutherford, T. B. Smith, S. Ahuka-Mundeke, J. J. Muyembe-Tamfum, A. W. Rimoin, and F. P. Schoenberg, “Real-time predictions of the 2018-2019 Ebola virus disease outbreak in the Democratic Republic of the Congo using Hawkes point process models,” *Epidemics*, p. 100354, July 2019.

- [48] A. R. Akhmetzhanov, H. Lee, S.-m. Jung, T. Kayano, B. Yuan, and H. Nishiura, “Analyzing and forecasting the Ebola incidence in North Kivu, the Democratic Republic of the Congo from 201819 in real time,” *Epidemics*, vol. 27, pp. 123–131, June 2019.
- [49] K. E. Haynes and A. S. Fotheringham, “Gravity and Spatial Interaction Models,” in *Book Chapters*, pp. 1–48, Regional Research Institute, West Virginia University, 1985.
- [50] T. G. Dietterich, “Ensemble Methods in Machine Learning,” in *Multiple Classifier Systems*, Lecture Notes in Computer Science, pp. 1–15, Springer Berlin Heidelberg, 2000.
- [51] S. V. Scarpino and G. Petri, “On the predictability of infectious disease outbreaks,” *Nature Communications*, vol. 10, pp. 1–8, Feb. 2019.

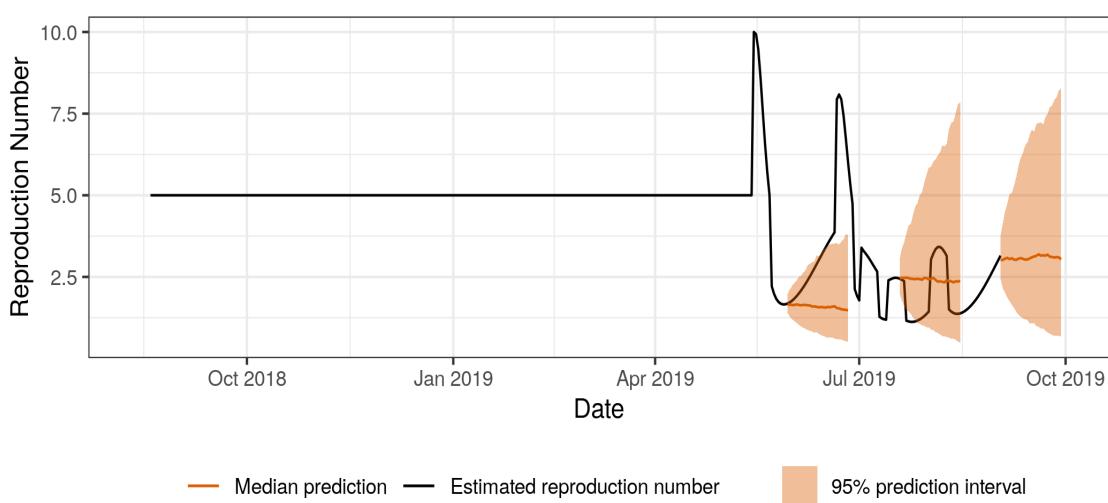
Appendices

This appendix shows the best fitting model and evaluation score for each health zone.

A Alimbongo



(a) Forecasted and predicted incidence for the best fitting model



(b) Forecasted and predicted reproduction numbers for the semilocal poisson model

Figure 7: Median forecast with 95 % prediction intervals and observed values for incidence and reproduction number for the semilocal poisson model for Alimbongo.

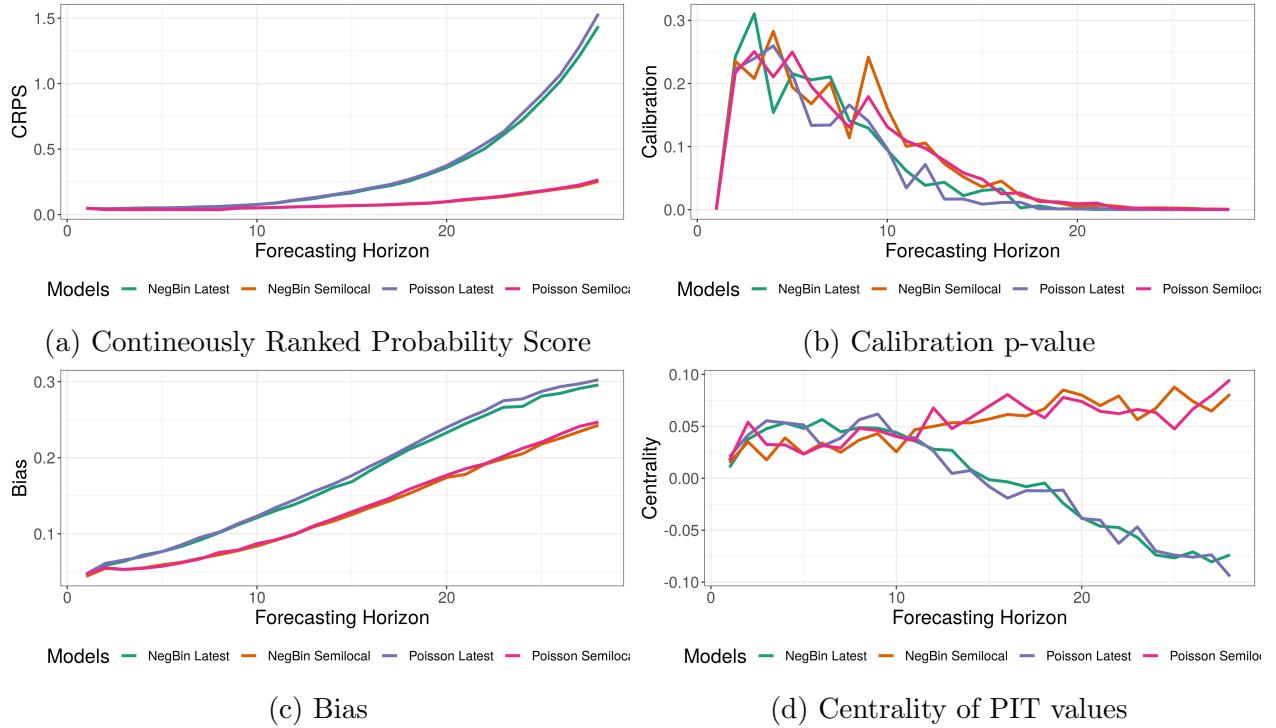
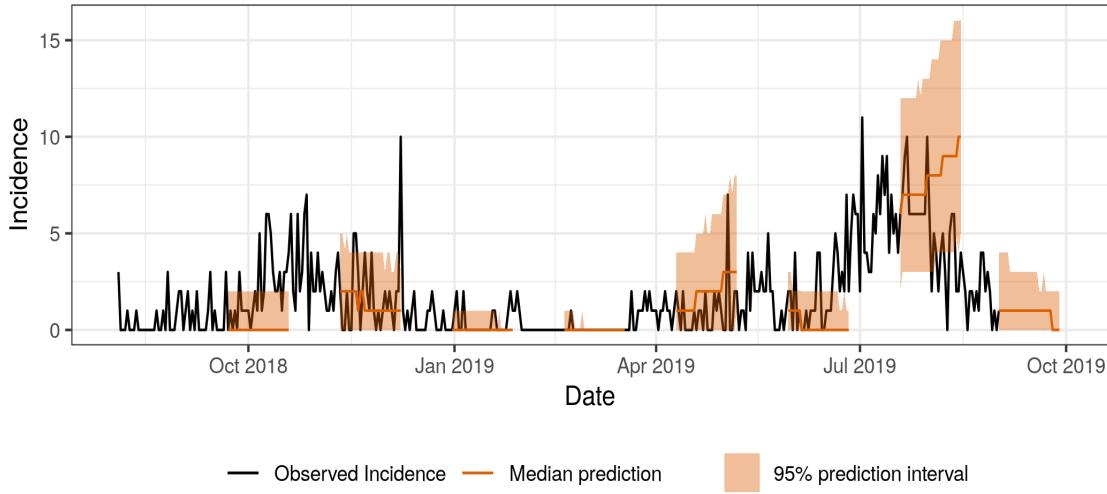
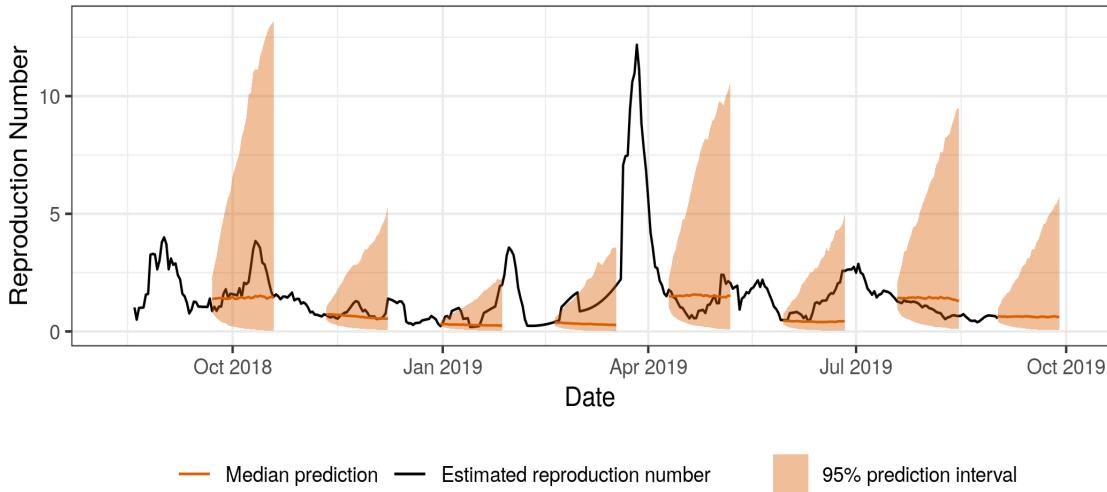


Figure 8: Scores for Alimbongo as a function of the forecasting horizon.

B Beni



(a) Forecasted and predicted incidence for the best fitting model



(b) Forecasted and predicted reproduction numbers for the semilocal poisson model

Figure 9: Median forecast with 95 % prediction intervals and observed values for incidence and reproduction number for the semilocal poisson model for Beni.

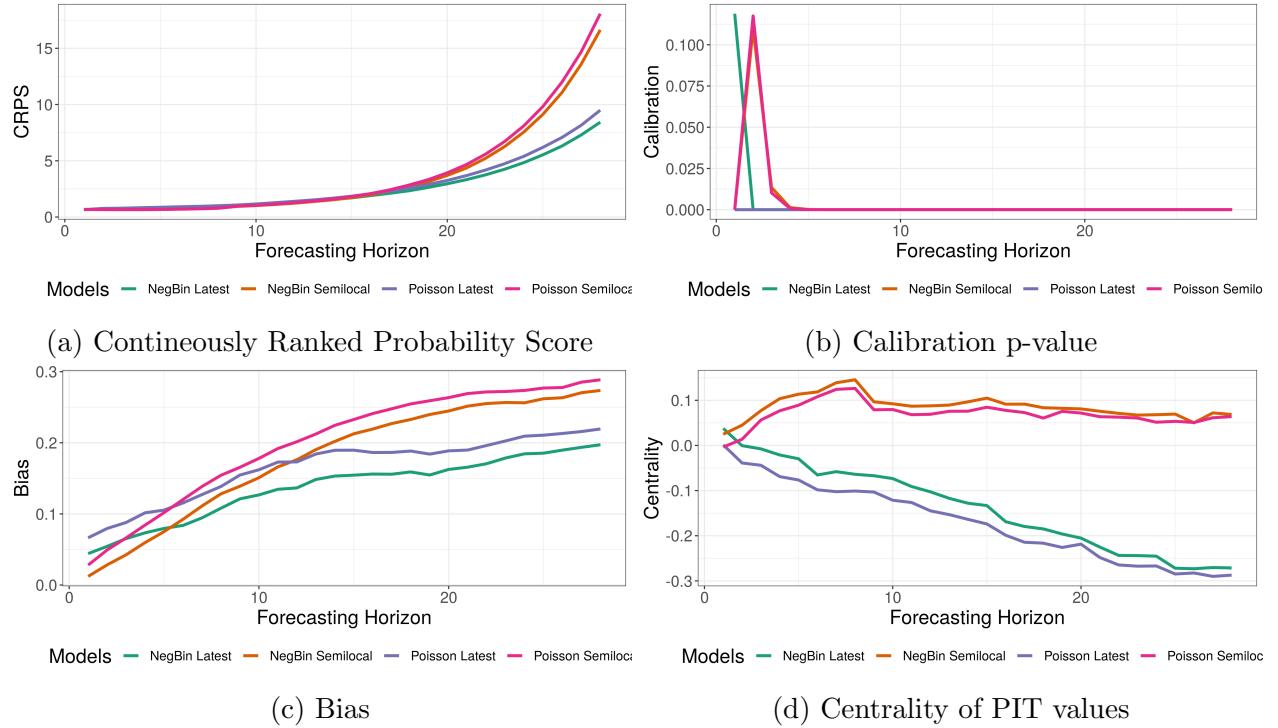
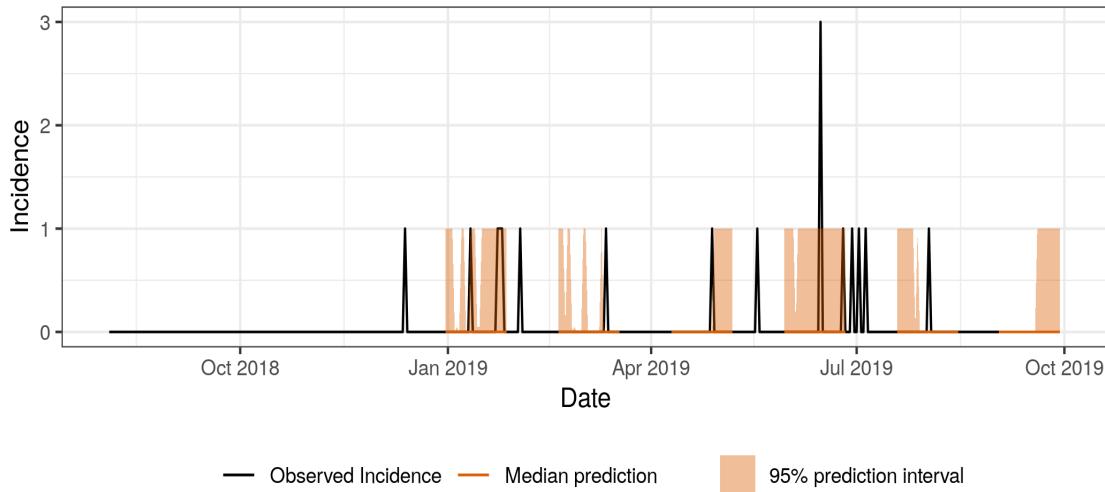
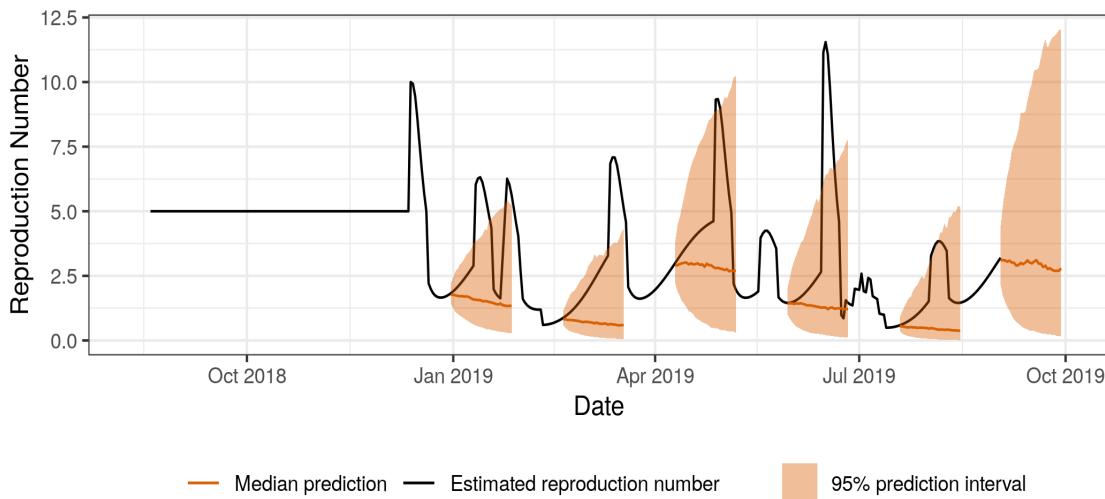


Figure 10: Scores for Beni as a function of the forecasting horizon.

C Biena



(a) Forecasted and predicted incidence for the best fitting model



(b) Forecasted and predicted reproduction numbers for the semilocal poisson model

Figure 11: Median forecast with 95 % prediction intervals and observed values for incidence and reproduction number for the semilocal poisson model for Biena.

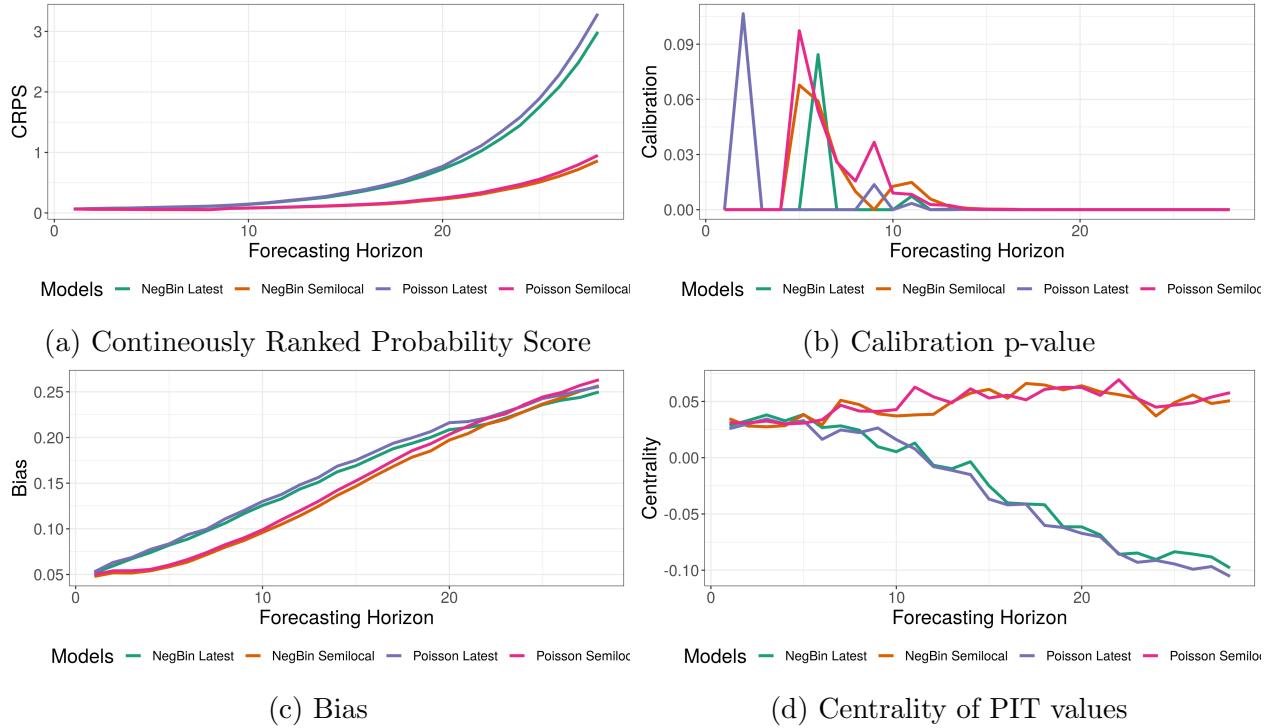
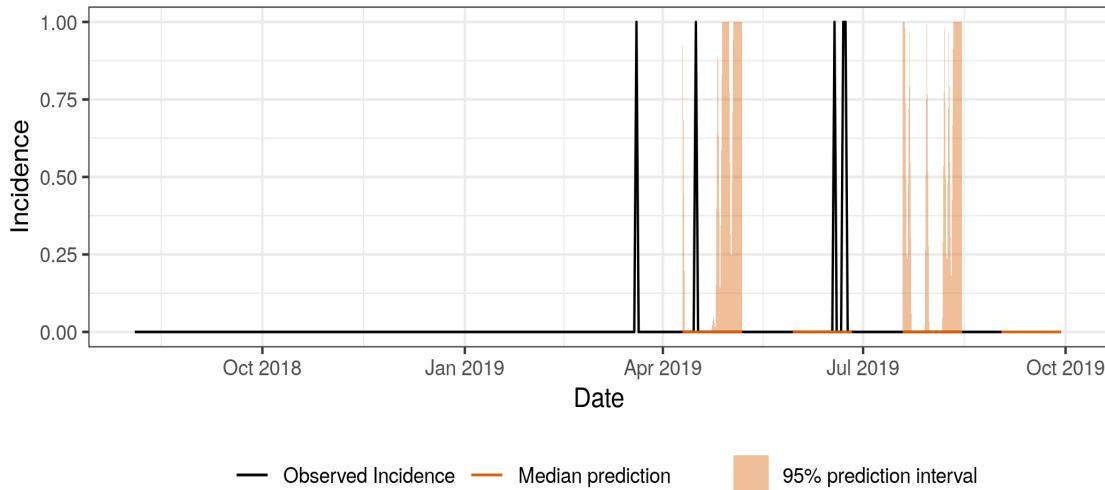
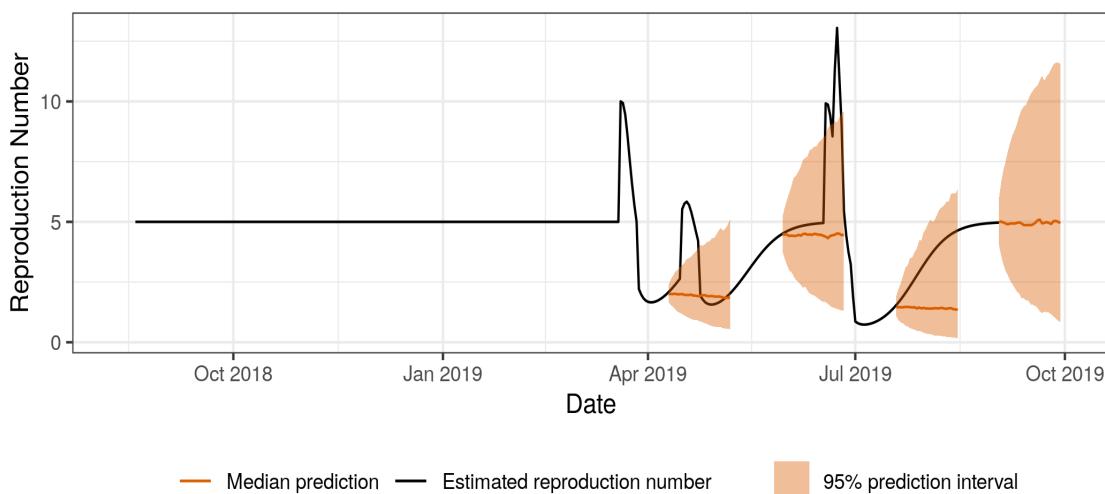


Figure 12: Scores for Biena as a function of the forecasting horizon.

D Bunia



(a) Forecasted and predicted incidence for the best fitting model



(b) Forecasted and predicted reproduction numbers for the semilocal poisson model

Figure 13: Median forecast with 95 % prediction intervals and observed values for incidence and reproduction number for the semilocal poisson model for Bunia.

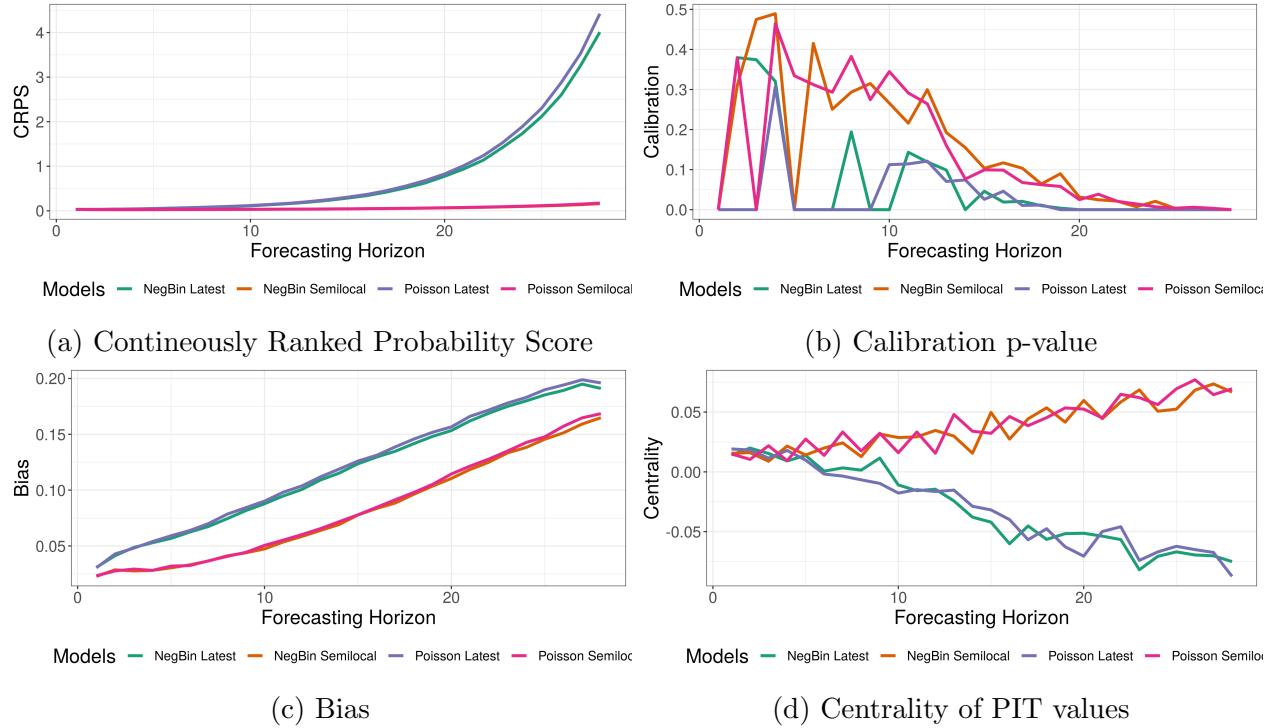
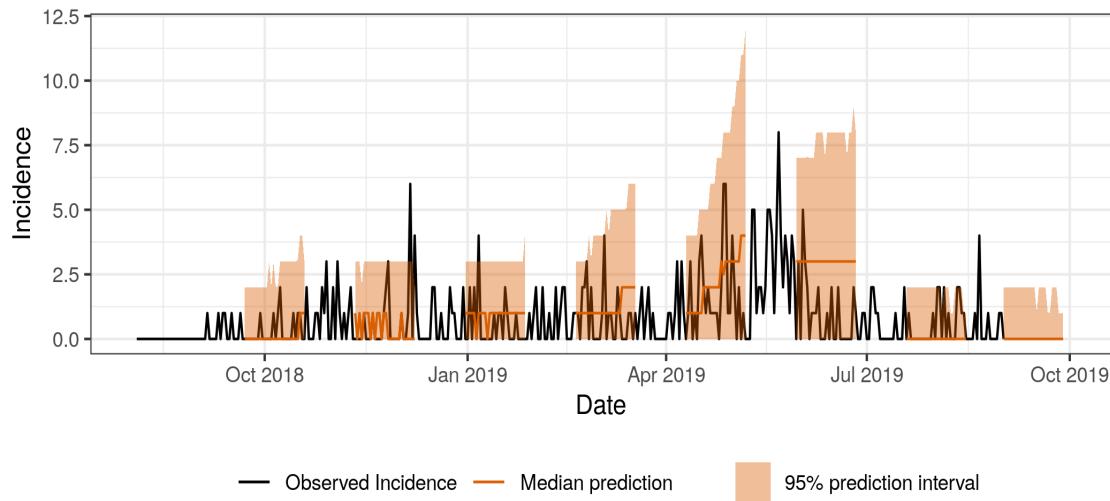
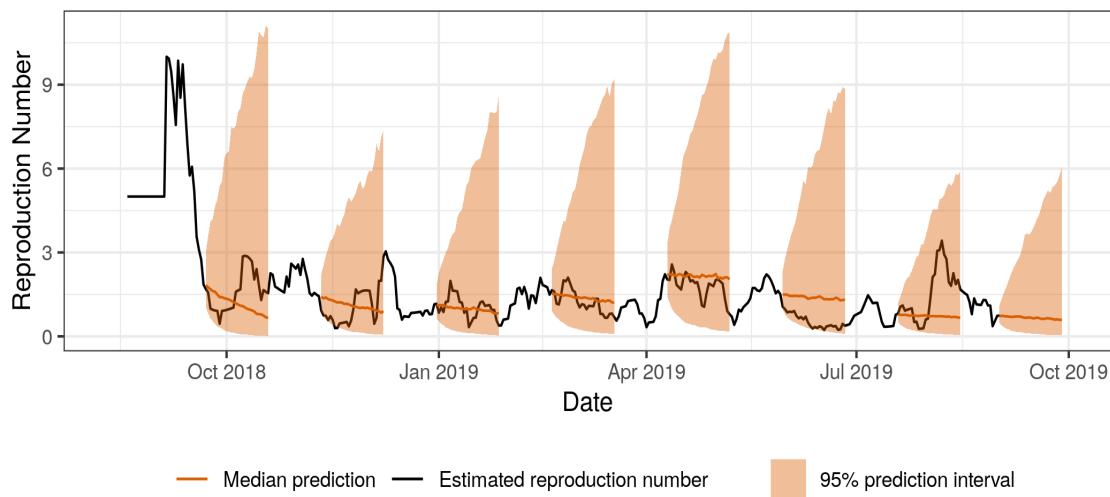


Figure 14: Scores for Bunia as a function of the forecasting horizon.

E Butembo



(a) Forecasted and predicted incidence for the best fitting model



(b) Forecasted and predicted reproduction numbers for the semilocal poisson model

Figure 15: Median forecast with 95 % prediction intervals and observed values for incidence and reproduction number for the semilocal poisson model for Butembo.

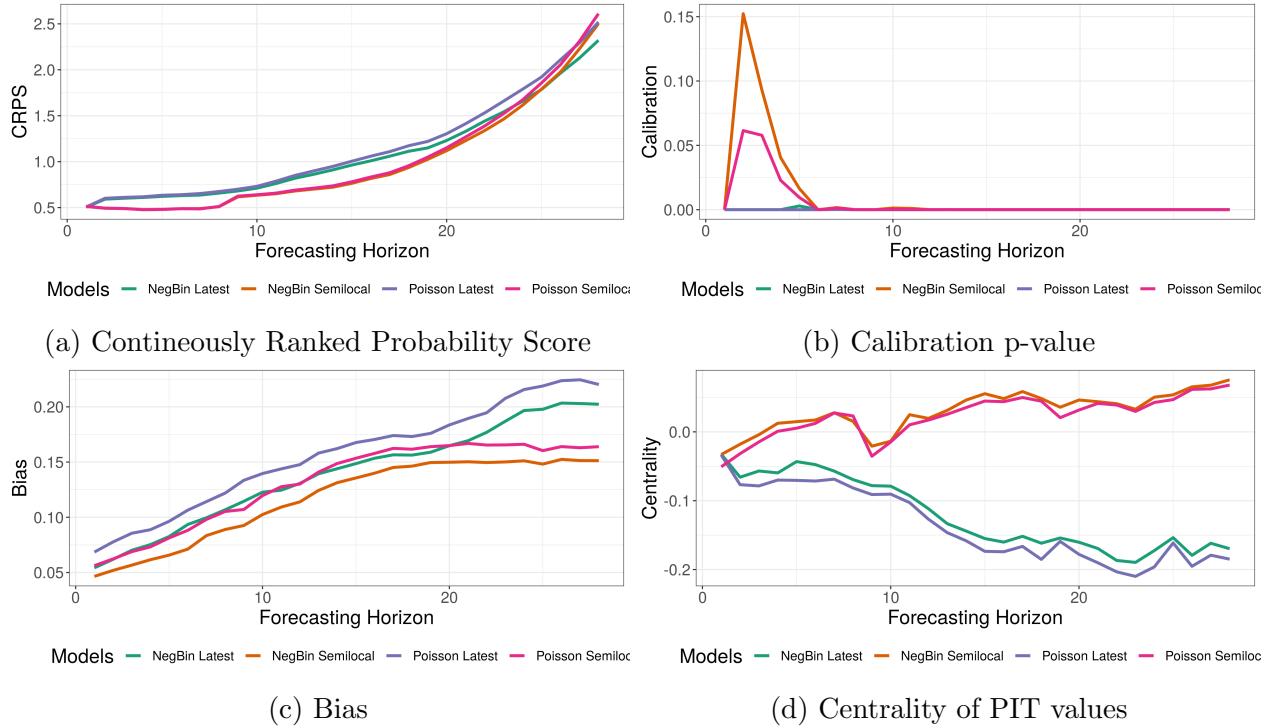
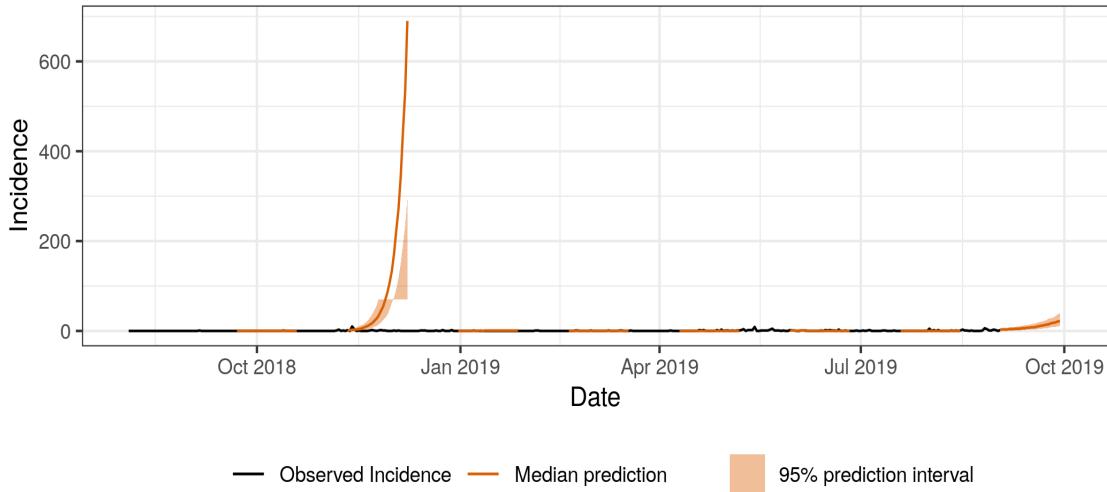
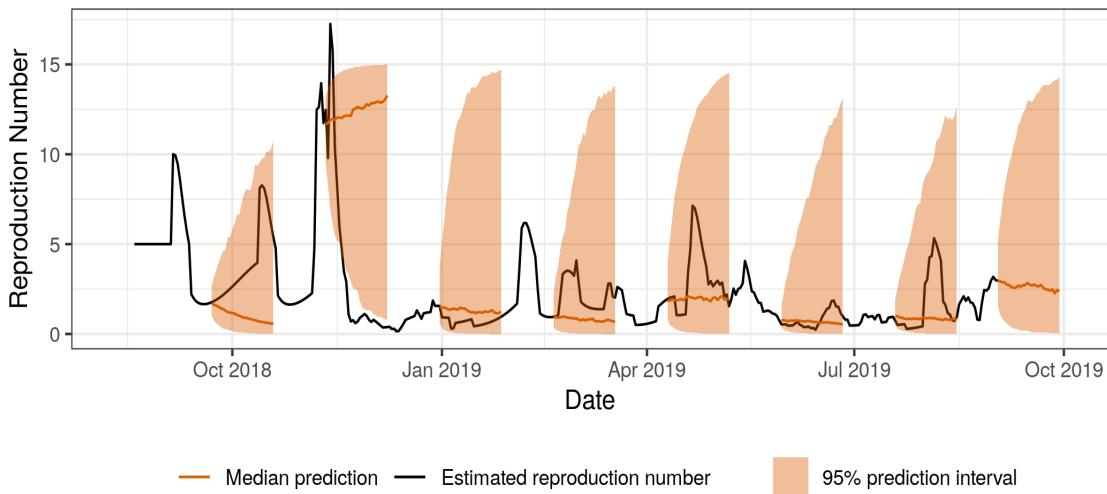


Figure 16: Scores for Butembo as a function of the forecasting horizon.

F Kalunguta



(a) Forecasted and predicted incidence for the best fitting model



(b) Forecasted and predicted reproduction numbers for the semilocal poisson model

Figure 17: Median forecast with 95 % prediction intervals and observed values for incidence and reproduction number for the semilocal poisson model for Kalunguta.

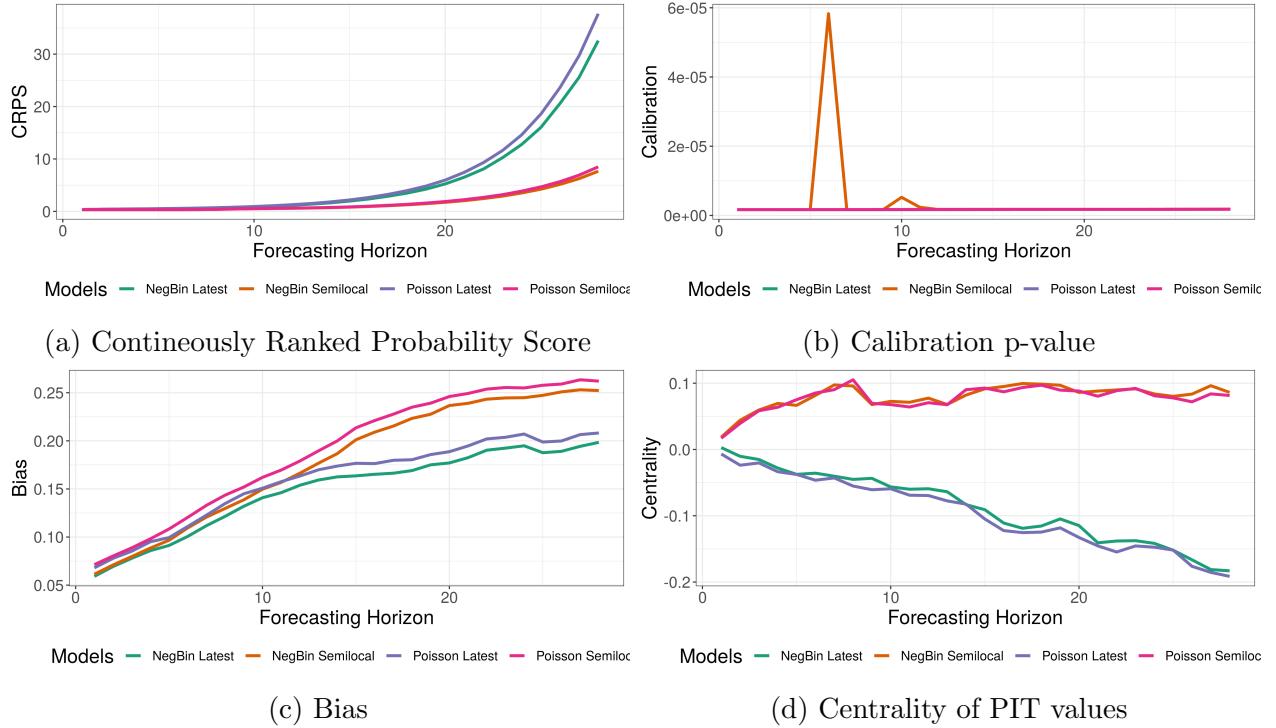
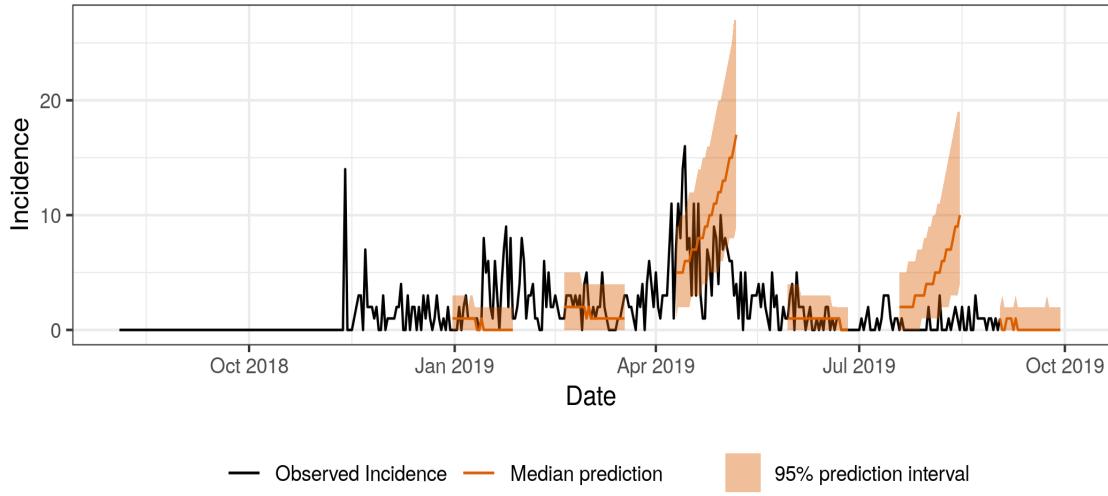
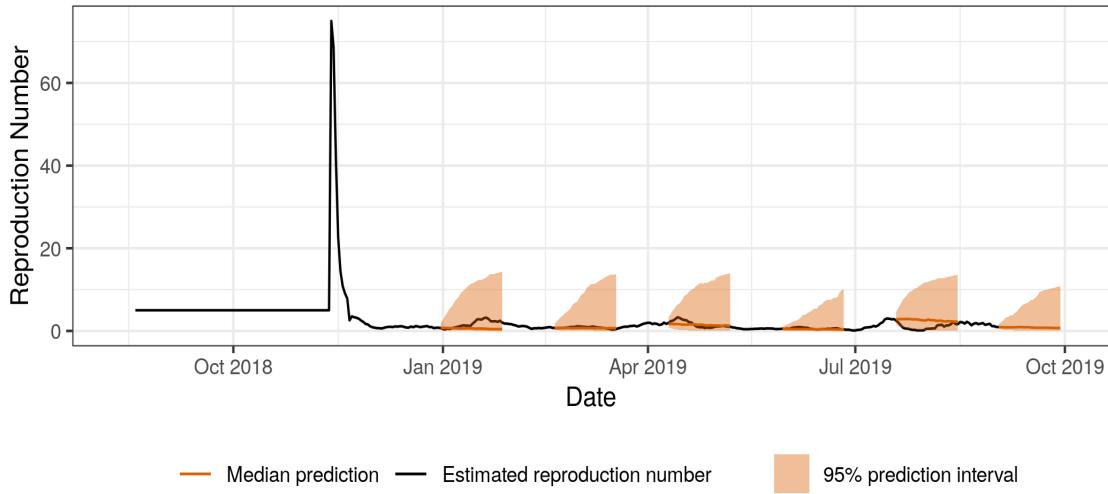


Figure 18: Scores for Kalunguta as a function of the forecasting horizon.

G Katwa



(a) Forecasted and predicted incidence for the best fitting model



(b) Forecasted and predicted reproduction numbers for the semilocal poisson model

Figure 19: Median forecast with 95 % prediction intervals and observed values for incidence and reproduction number for the semilocal poisson model for Katwa.

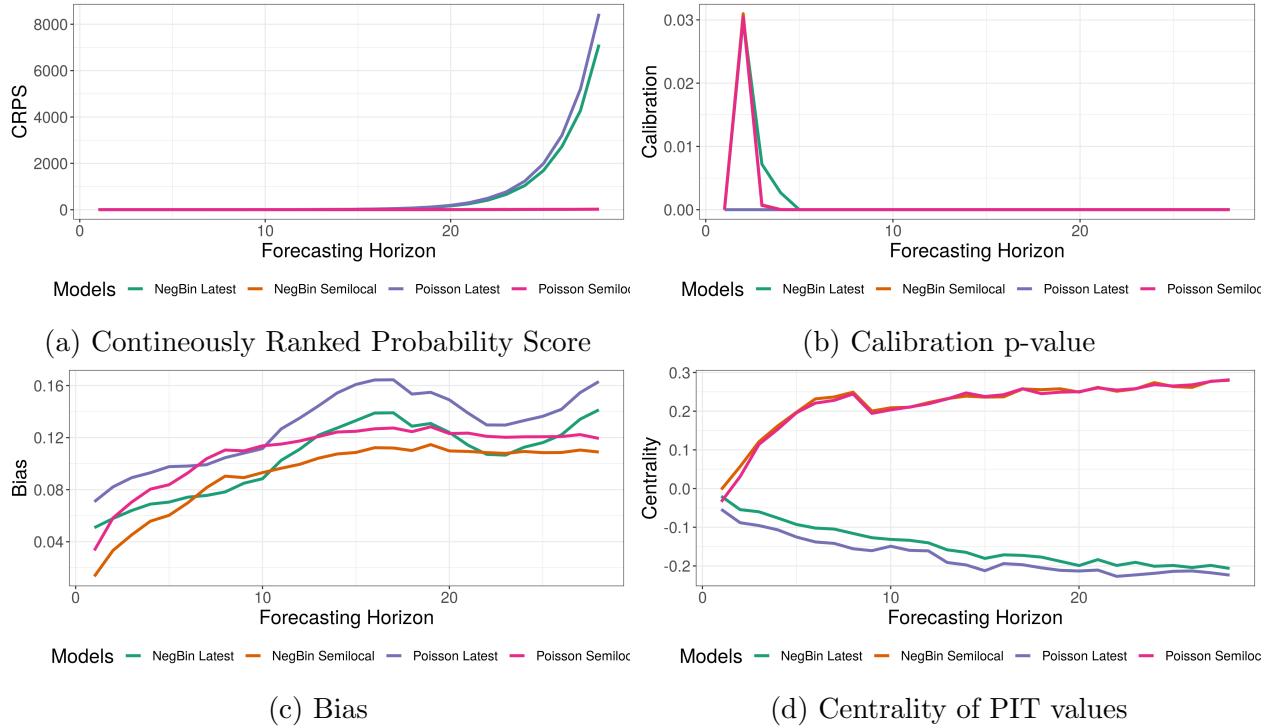
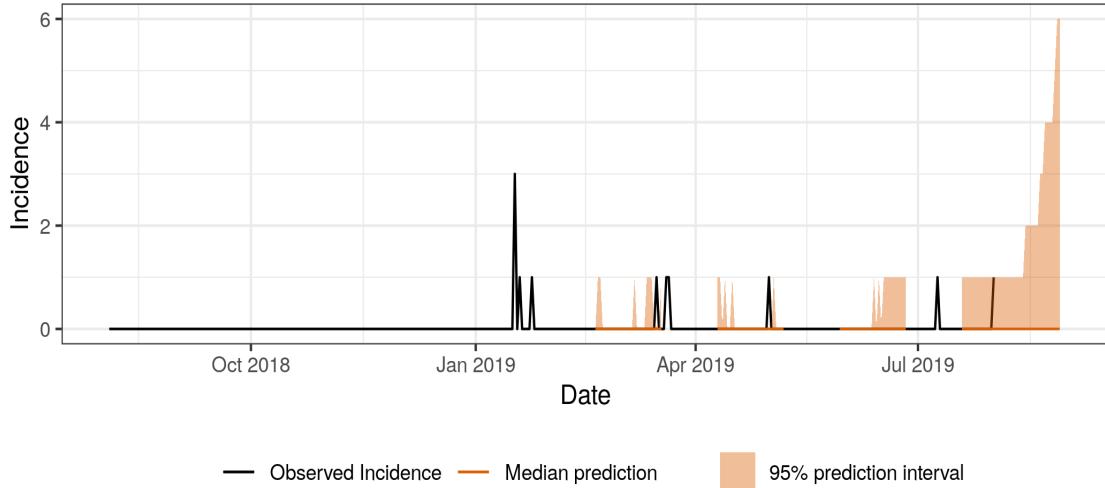
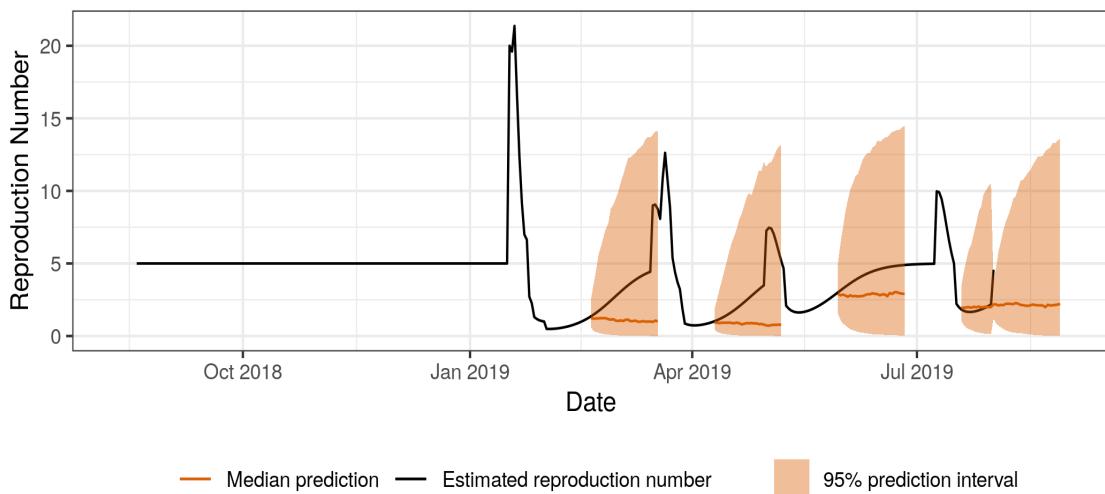


Figure 20: Scores for Katwa as a function of the forecasting horizon.

H Kayina



(a) Forecasted and predicted incidence for the best fitting model



(b) Forecasted and predicted reproduction numbers for the semilocal poisson model

Figure 21: Median forecast with 95 % prediction intervals and observed values for incidence and reproduction number for the semilocal poisson model for Kayina.

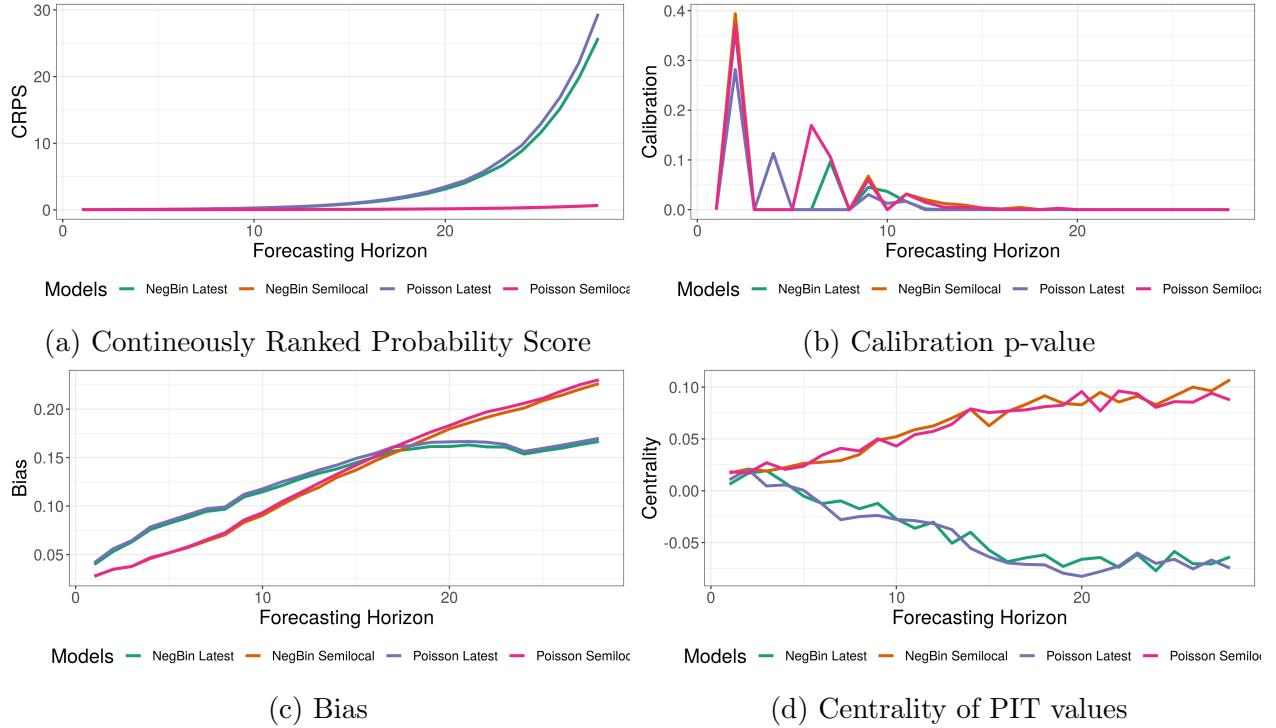
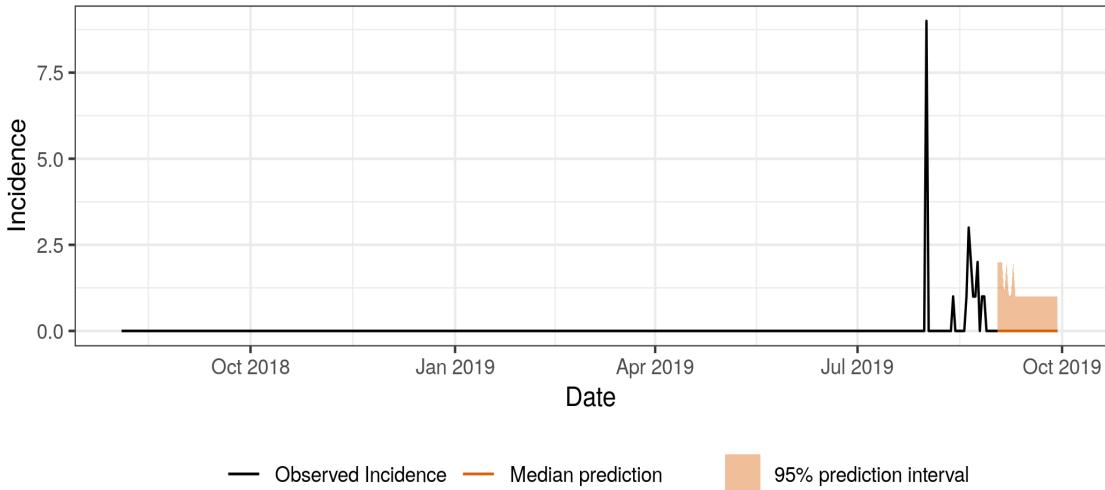
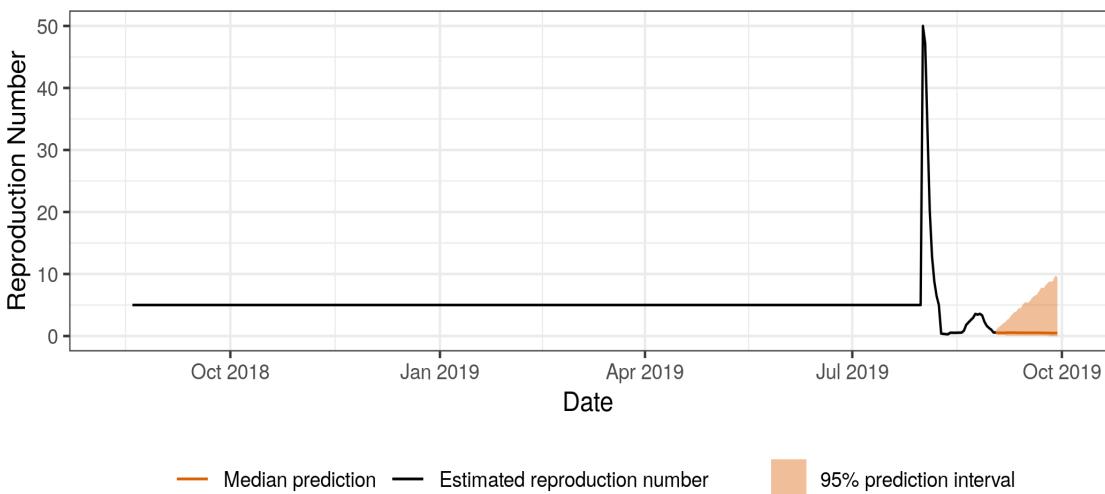


Figure 22: Scores for Kayina as a function of the forecasting horizon.

I Kayna



(a) Forecasted and predicted incidence for the best fitting model



(b) Forecasted and predicted reproduction numbers for the semilocal poisson model

Figure 23: Median forecast with 95 % prediction intervals and observed values for incidence and reproduction number for the semilocal poisson model for Kayna.

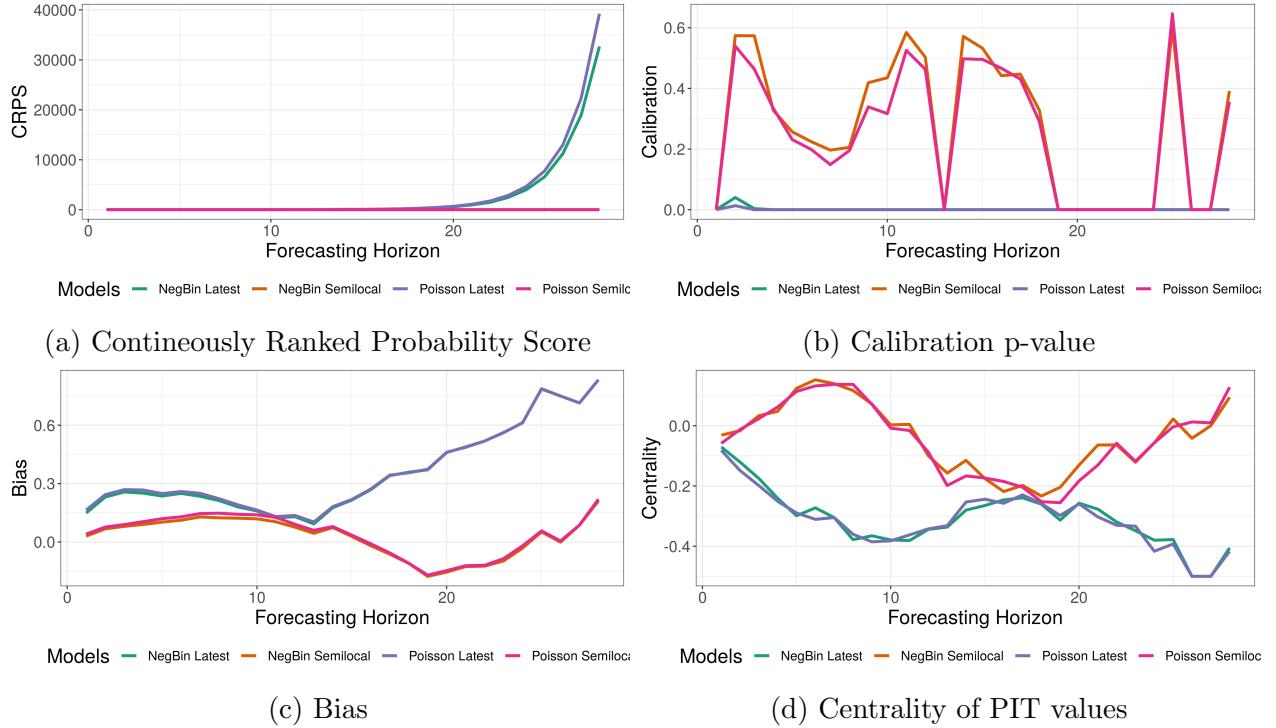
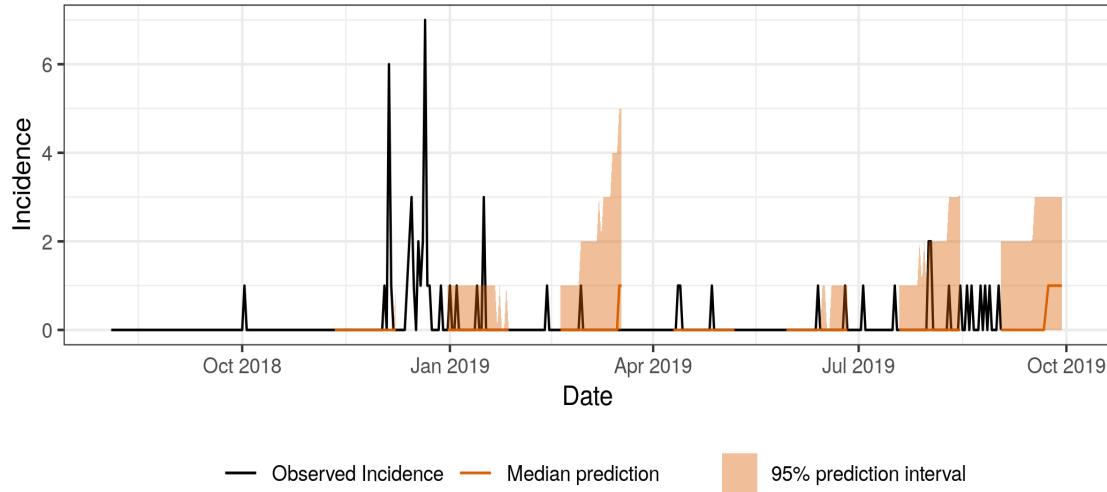
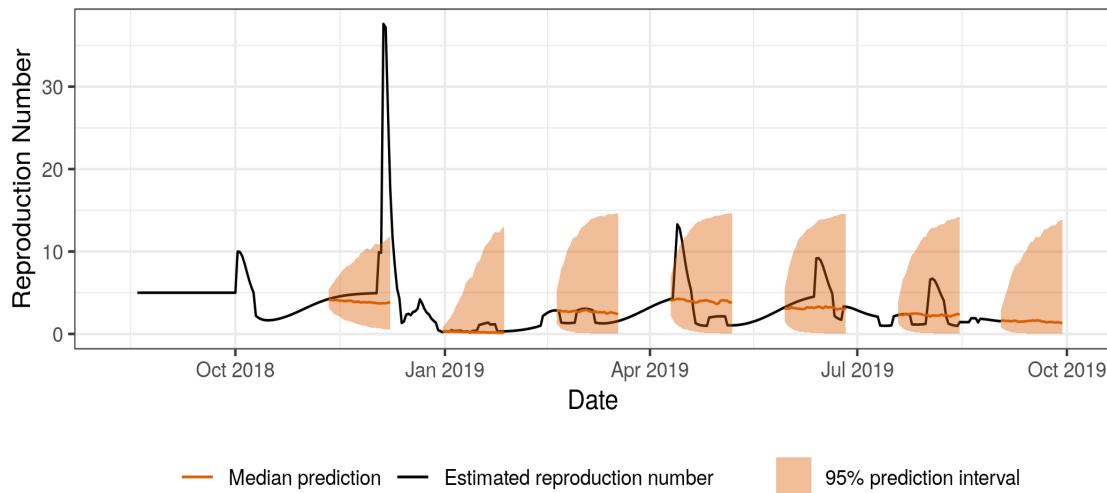


Figure 24: Scores for Kayna as a function of the forecasting horizon.

J Komanda



(a) Forecasted and predicted incidence for the best fitting model



(b) Forecasted and predicted reproduction numbers for the semilocal poisson model

Figure 25: Median forecast with 95 % prediction intervals and observed values for incidence and reproduction number for the semilocal poisson model for Komanda.

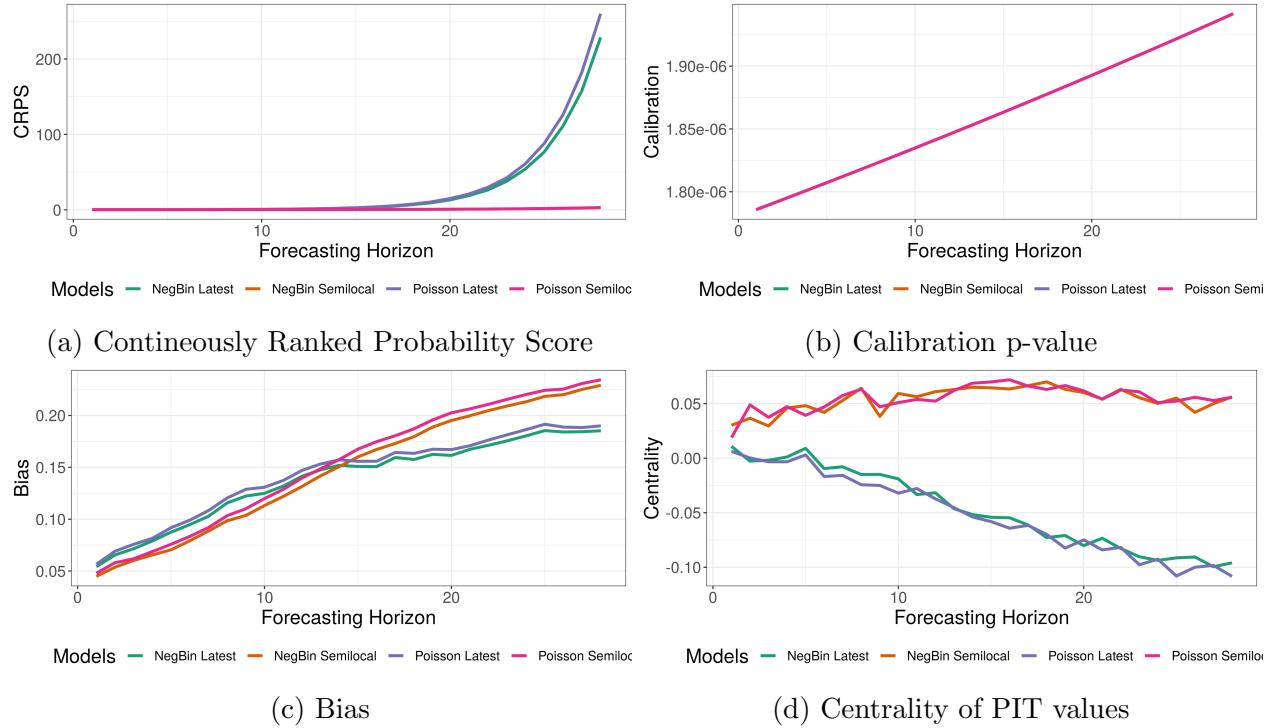
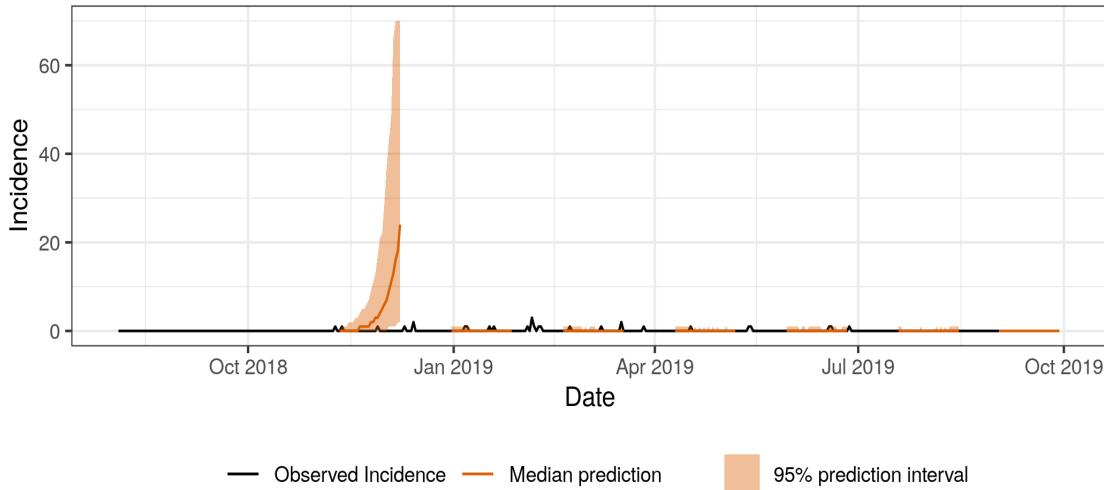
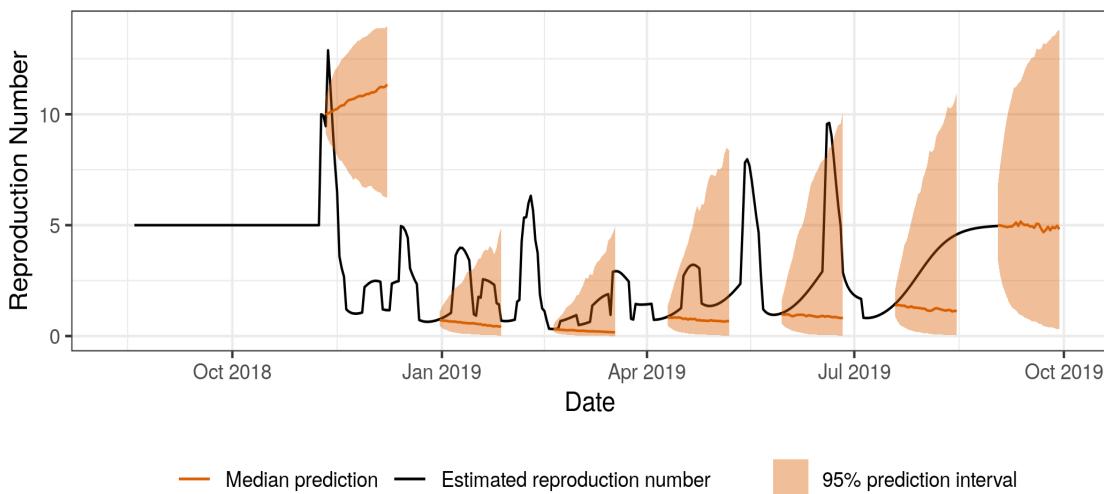


Figure 26: Scores for Komanda as a function of the forecasting horizon.

K Kyondo



(a) Forecasted and predicted incidence for the best fitting model



(b) Forecasted and predicted reproduction numbers for the semilocal poisson model

Figure 27: Median forecast with 95 % prediction intervals and observed values for incidence and reproduction number for the semilocal poisson model for Kyondo.

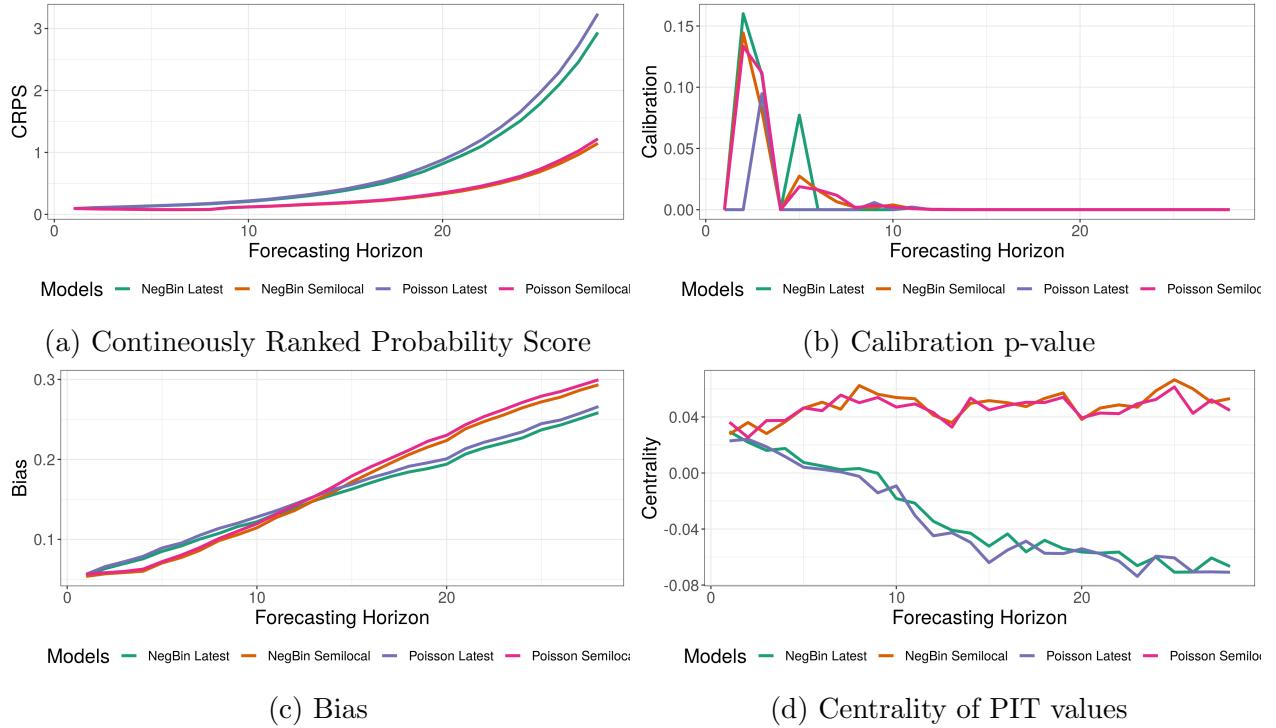
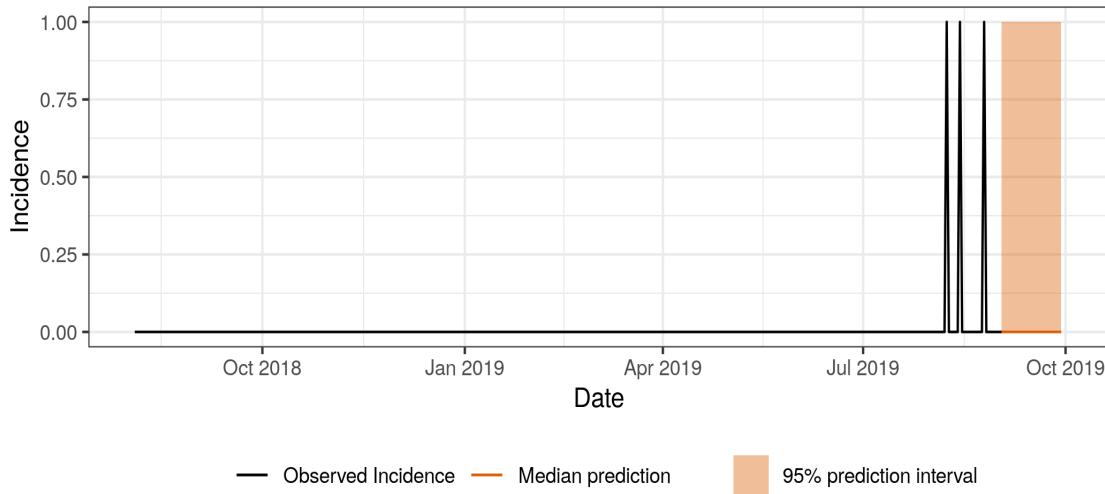
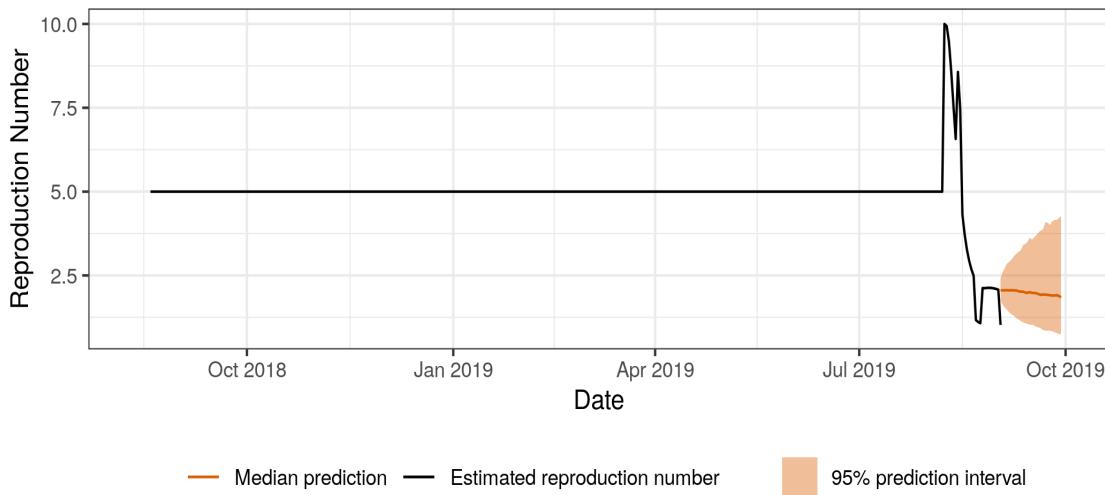


Figure 28: Scores for Kyondo as a function of the forecasting horizon.

L Lolwa



(a) Forecasted and predicted incidence for the best fitting model



(b) Forecasted and predicted reproduction numbers for the semilocal poisson model

Figure 29: Median forecast with 95 % prediction intervals and observed values for incidence and reproduction number for the semilocal poisson model for Lolwa.

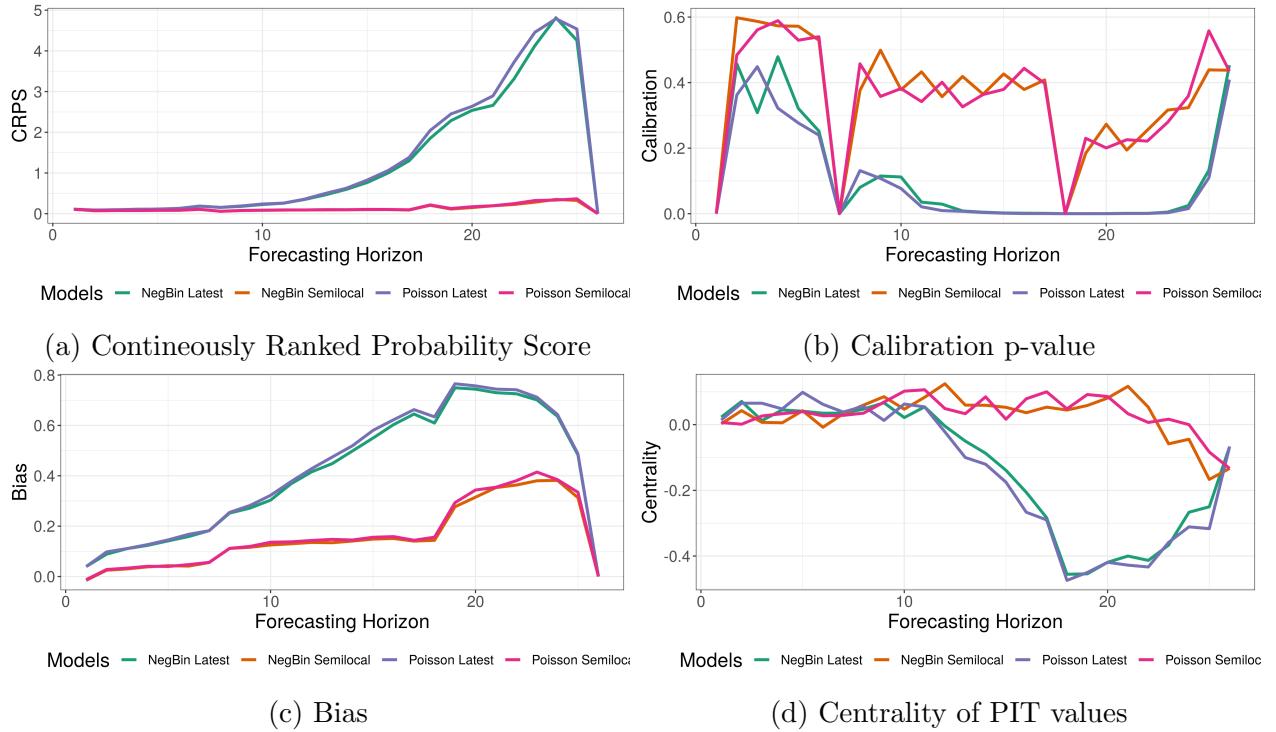
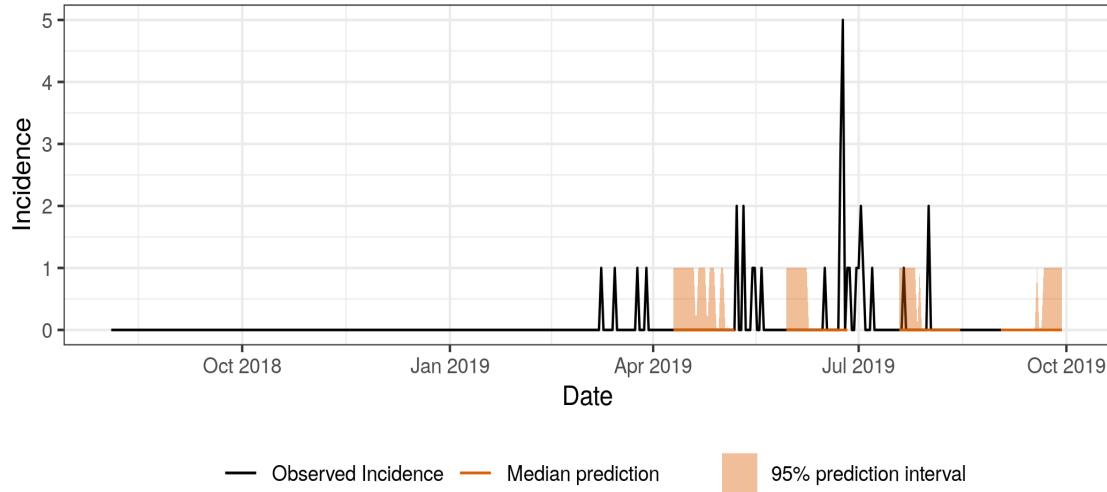
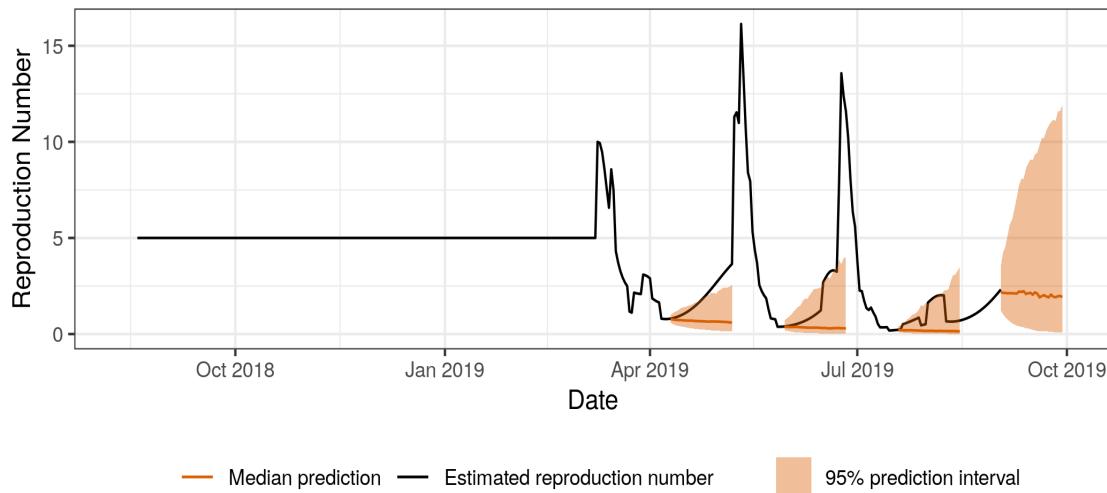


Figure 30: Scores for Lolwa as a function of the forecasting horizon.

M Lubero



(a) Forecasted and predicted incidence for the best fitting model



(b) Forecasted and predicted reproduction numbers for the semilocal poisson model

Figure 31: Median forecast with 95 % prediction intervals and observed values for incidence and reproduction number for the semilocal poisson model for Lubero.

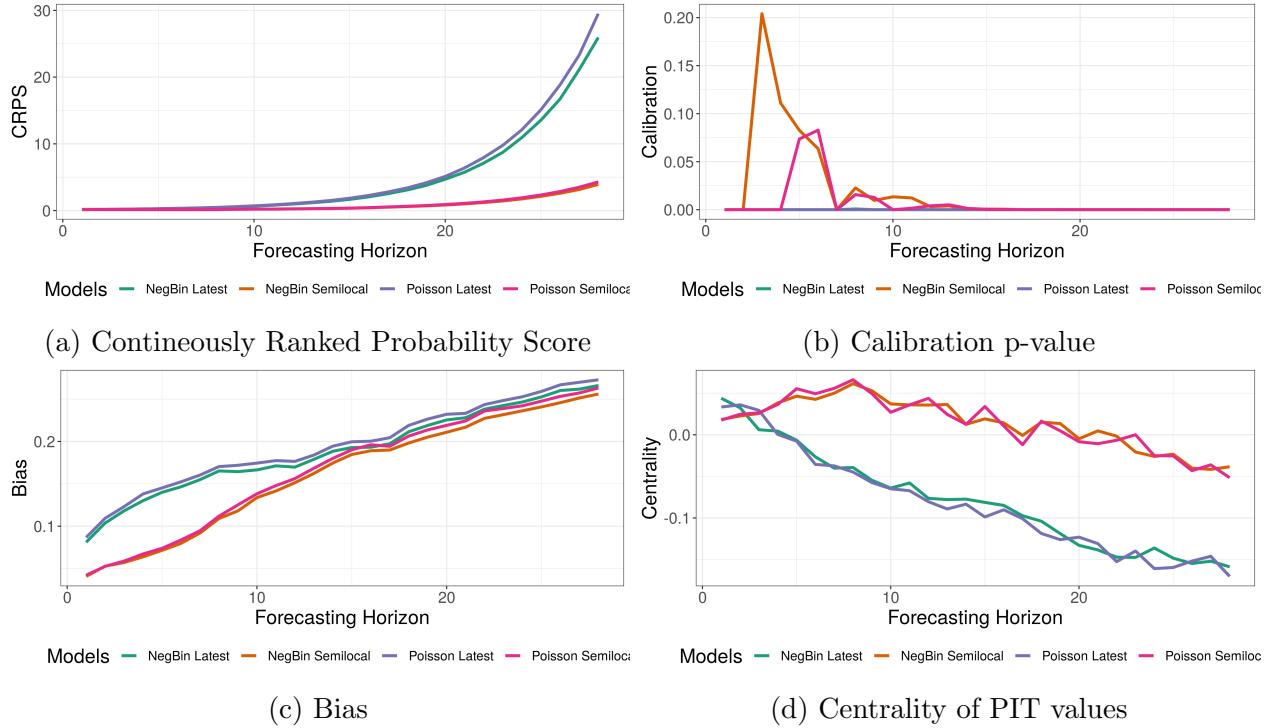
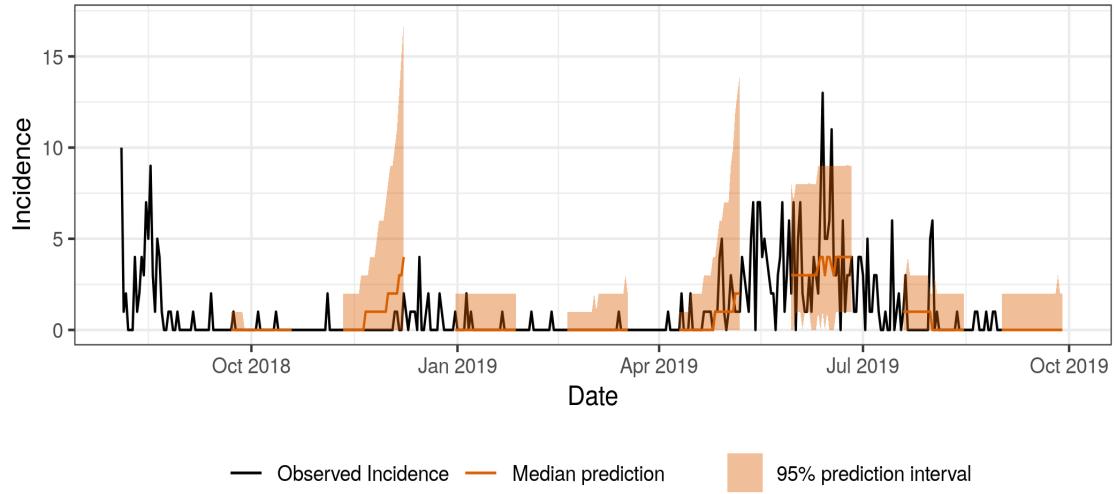
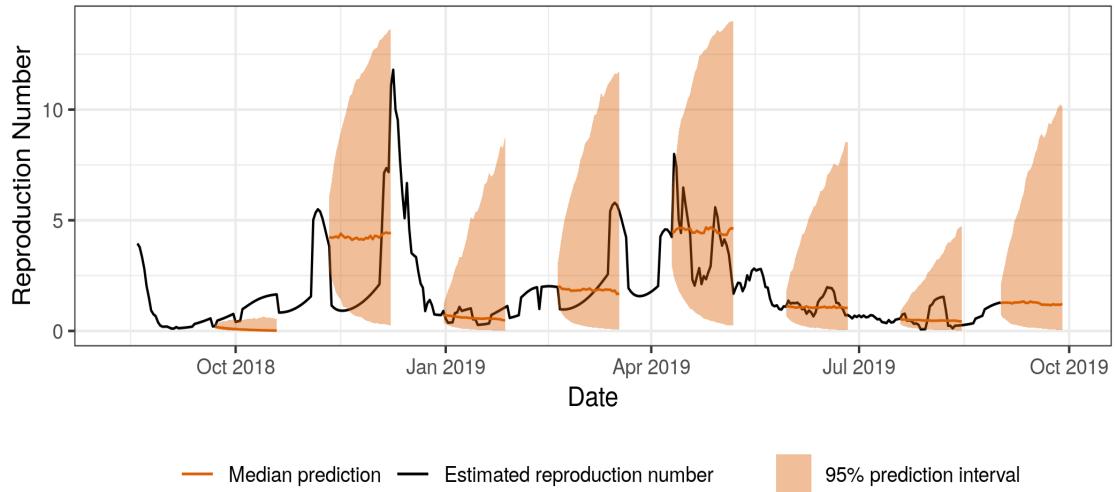


Figure 32: Scores for Lubero as a function of the forecasting horizon.

N Mabalako



(a) Forecasted and predicted incidence for the best fitting model



(b) Forecasted and predicted reproduction numbers for the semilocal poisson model

Figure 33: Median forecast with 95 % prediction intervals and observed values for incidence and reproduction number for the semilocal poisson model for Mabalako.

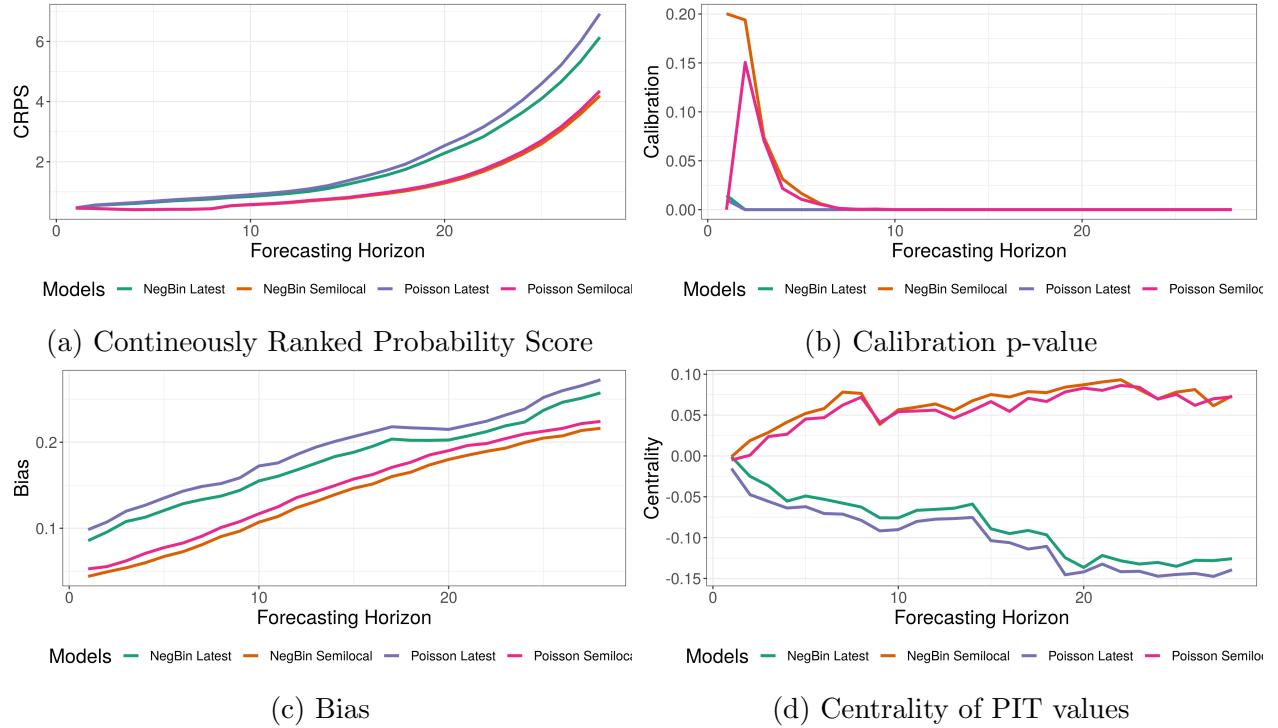
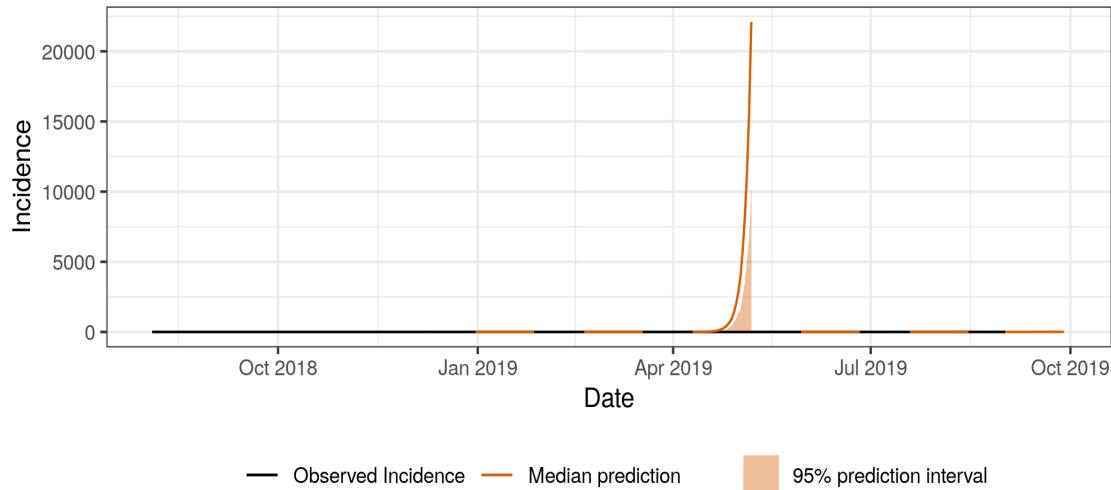
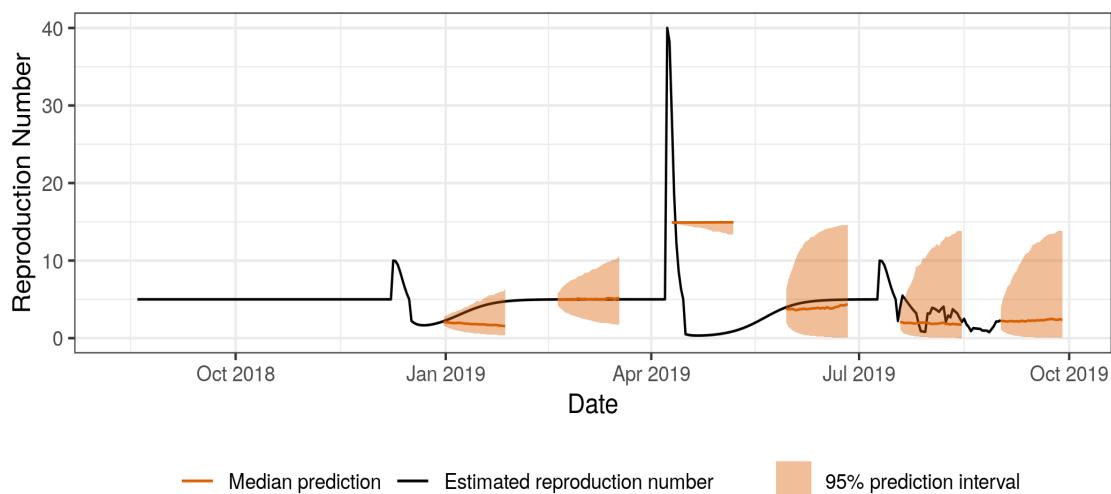


Figure 34: Scores for Mabalako as a function of the forecasting horizon.

O Mambasa



(a) Forecasted and predicted incidence for the best fitting model



(b) Forecasted and predicted reproduction numbers for the semilocal poisson model

Figure 35: Median forecast with 95 % prediction intervals and observed values for incidence and reproduction number for the semilocal poisson model for Mambasa.

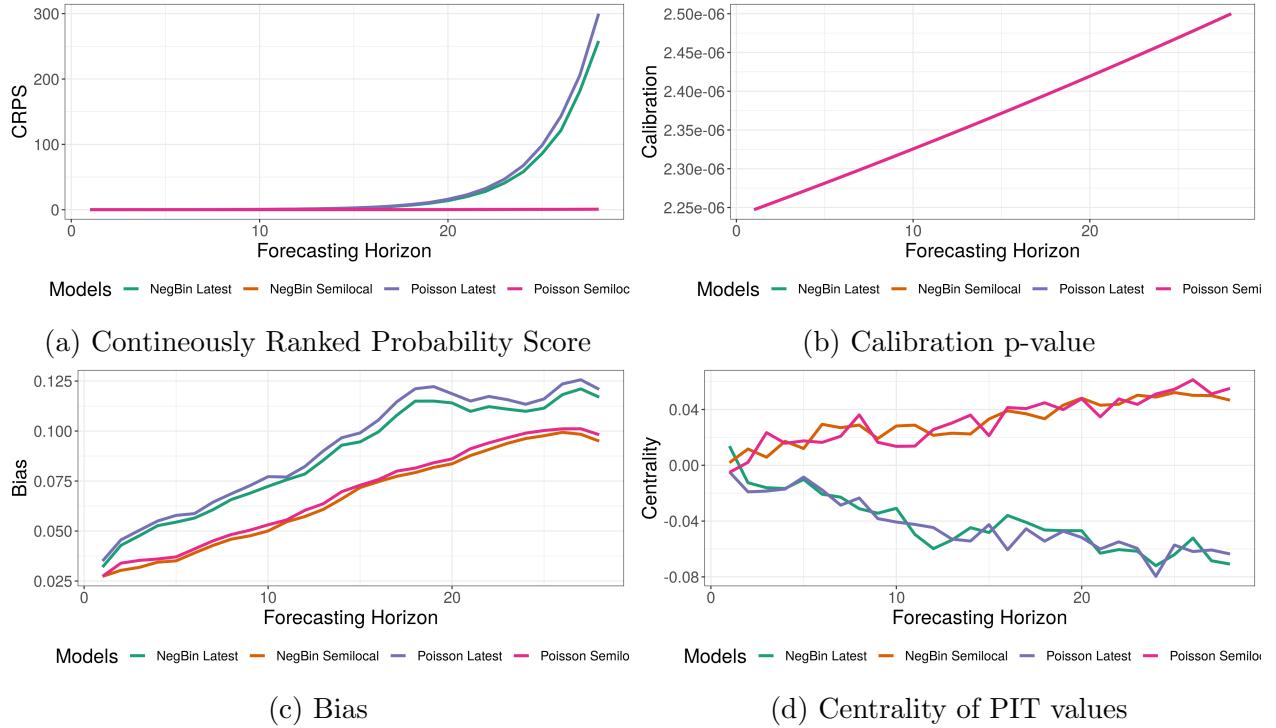
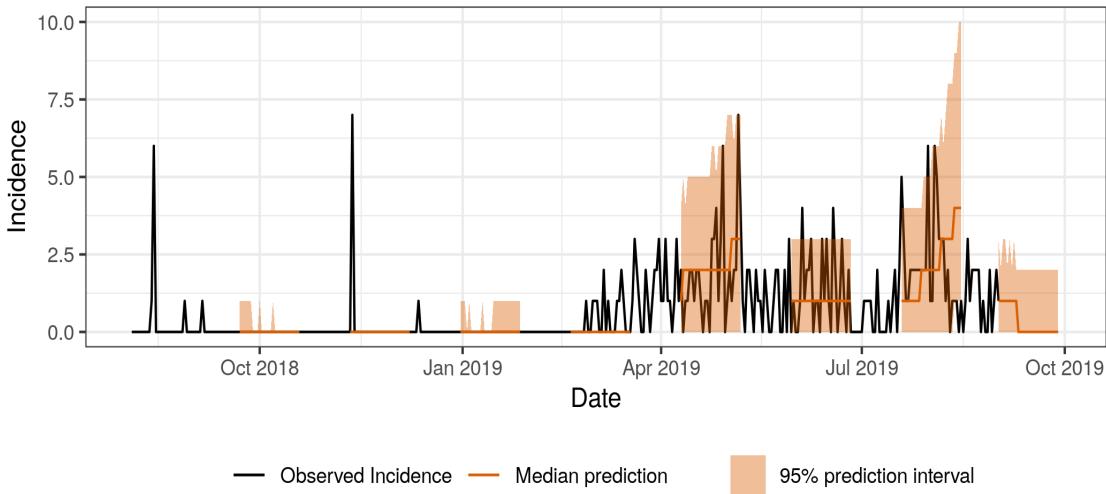
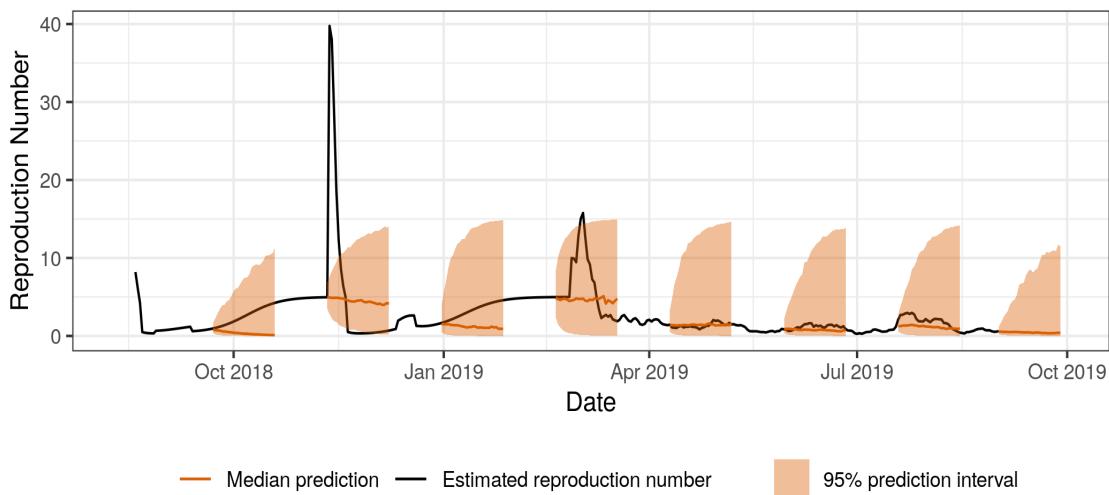


Figure 36: Scores for Mambasa as a function of the forecasting horizon.

P Mandima



(a) Forecasted and predicted incidence for the best fitting model



(b) Forecasted and predicted reproduction numbers for the semilocal poisson model

Figure 37: Median forecast with 95 % prediction intervals and observed values for incidence and reproduction number for the semilocal poisson model for Mandima.

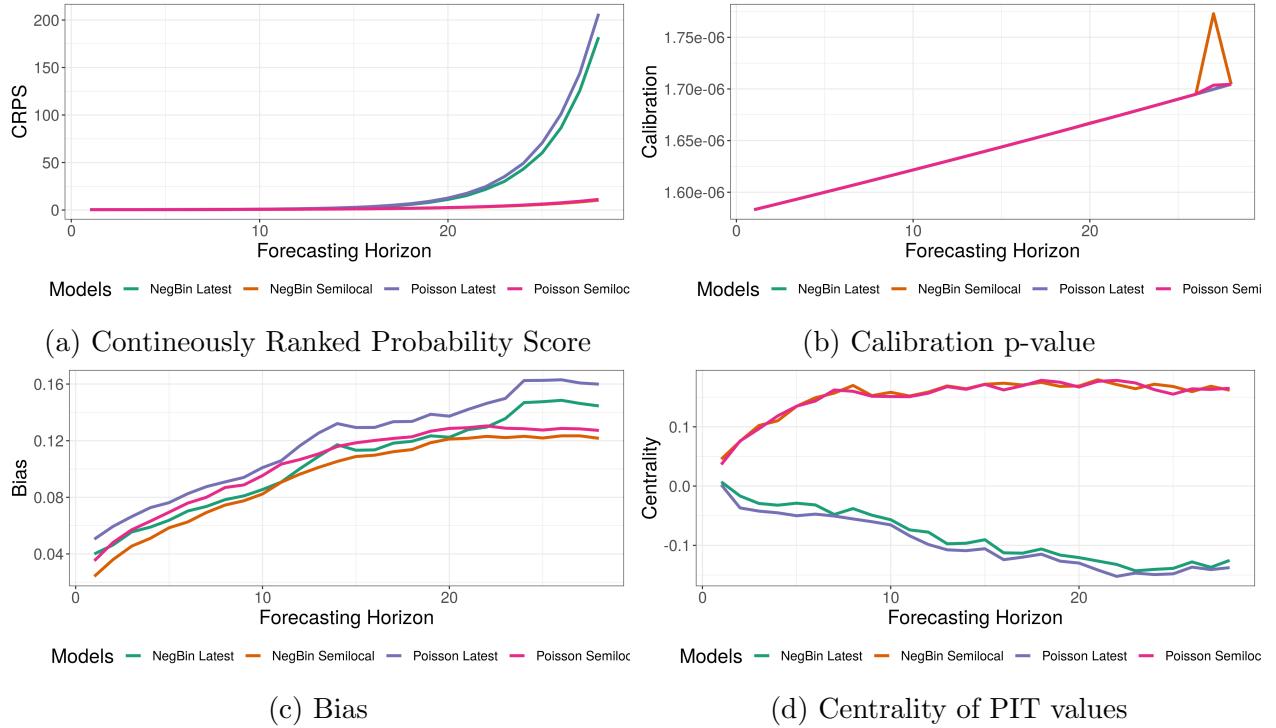
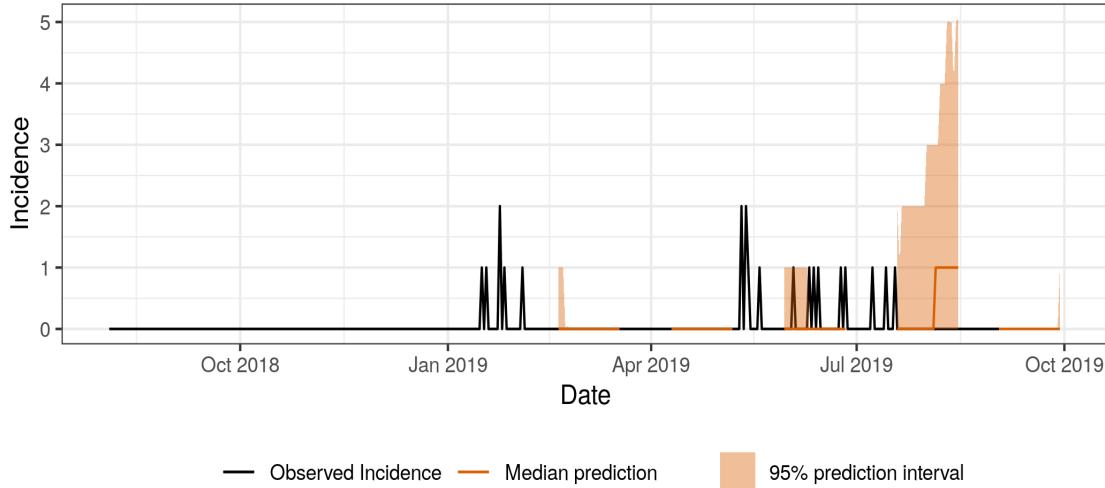
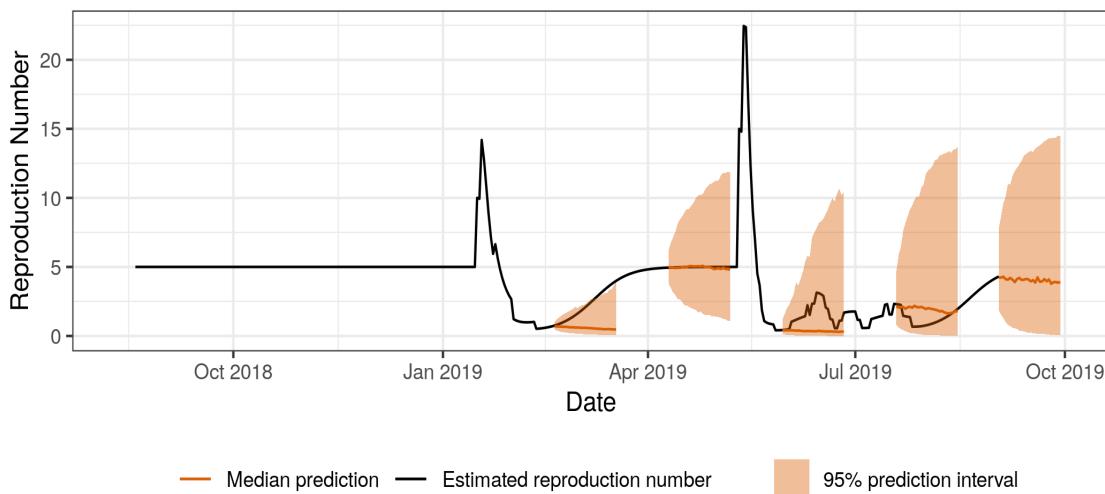


Figure 38: Scores for Mandima as a function of the forecasting horizon.

Q Mangurujipa



(a) Forecasted and predicted incidence for the best fitting model



(b) Forecasted and predicted reproduction numbers for the semilocal poisson model

Figure 39: Median forecast with 95 % prediction intervals and observed values for incidence and reproduction number for the semilocal poisson model for Mangurujipa.

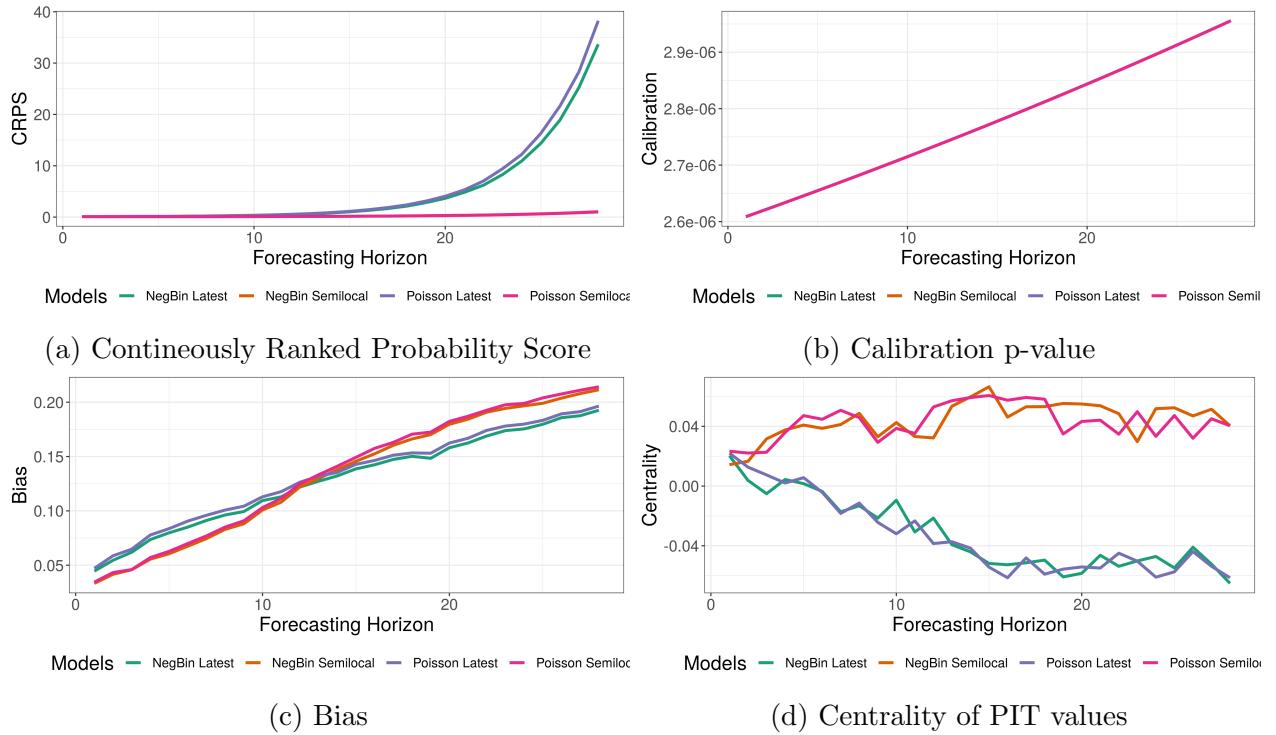
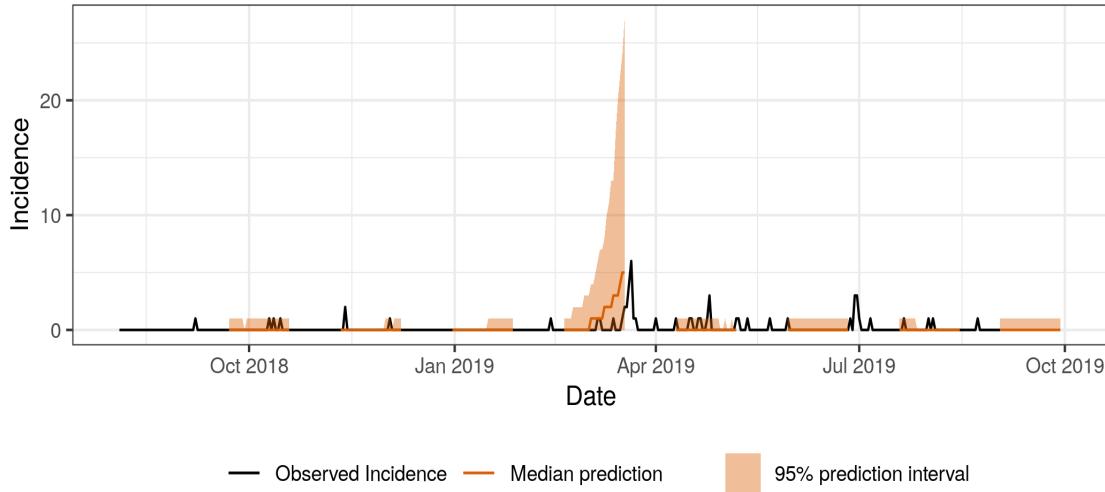
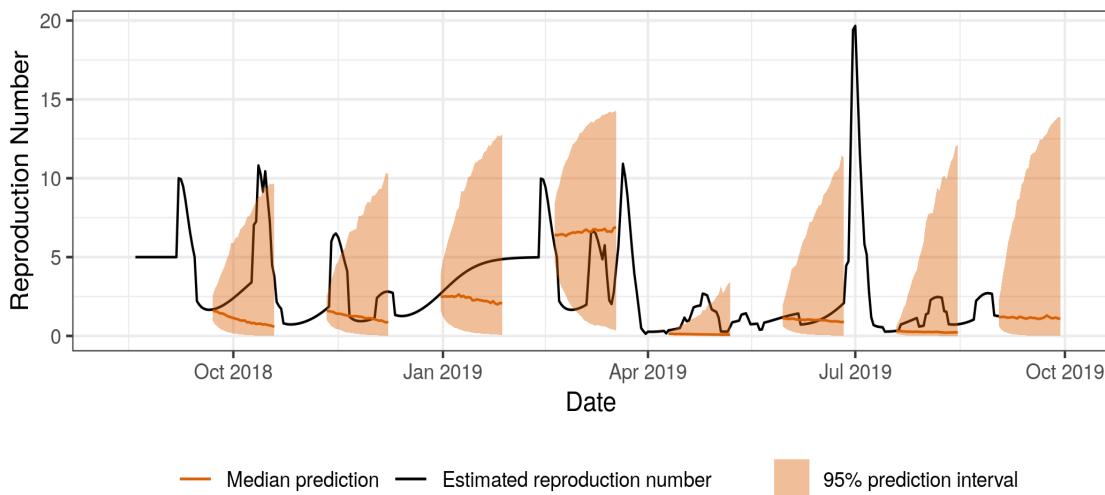


Figure 40: Scores for Mangurujipa as a function of the forecasting horizon.

R Masereka



(a) Forecasted and predicted incidence for the best fitting model



(b) Forecasted and predicted reproduction numbers for the semilocal poisson model

Figure 41: Median forecast with 95 % prediction intervals and observed values for incidence and reproduction number for the semilocal poisson model for Masereka.

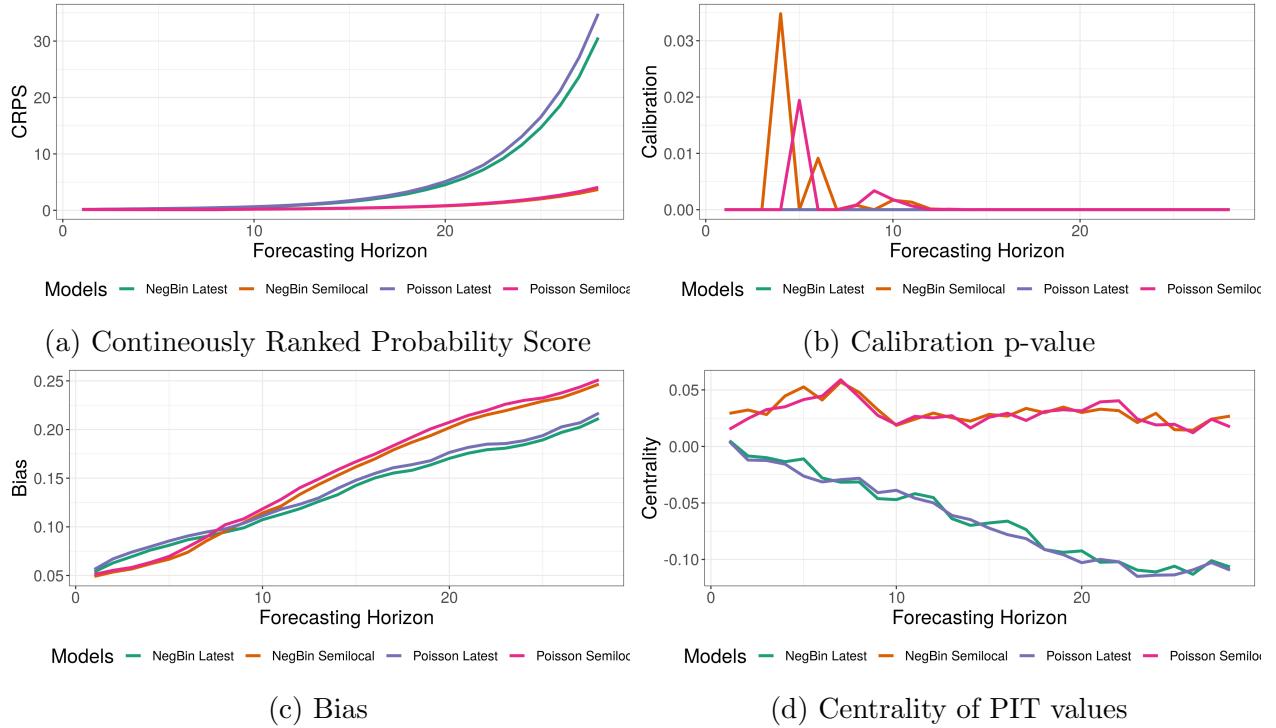
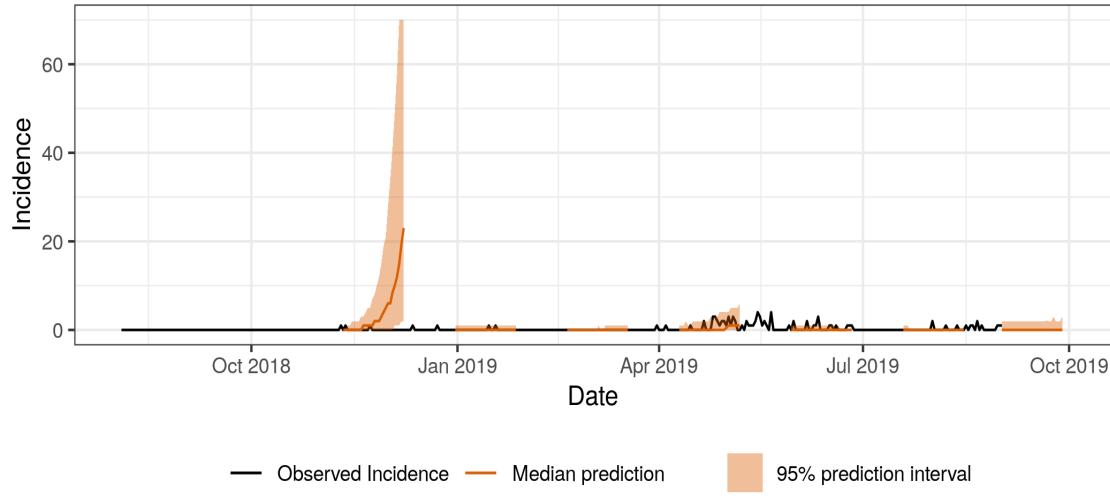
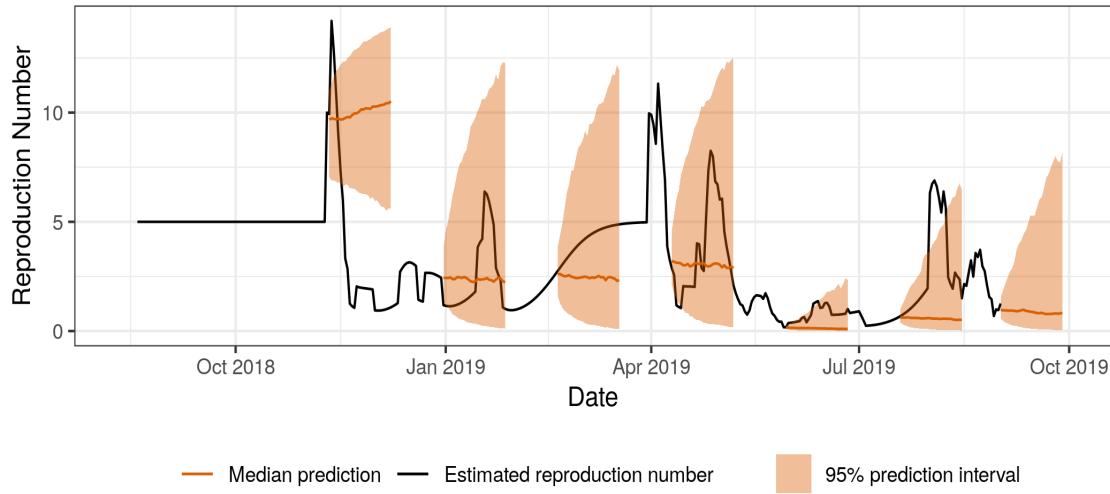


Figure 42: Scores for Masereka as a function of the forecasting horizon.

S Musienene



(a) Forecasted and predicted incidence for the best fitting model



(b) Forecasted and predicted reproduction numbers for the semilocal poisson model

Figure 43: Median forecast with 95 % prediction intervals and observed values for incidence and reproduction number for the semilocal poisson model for Musienene.

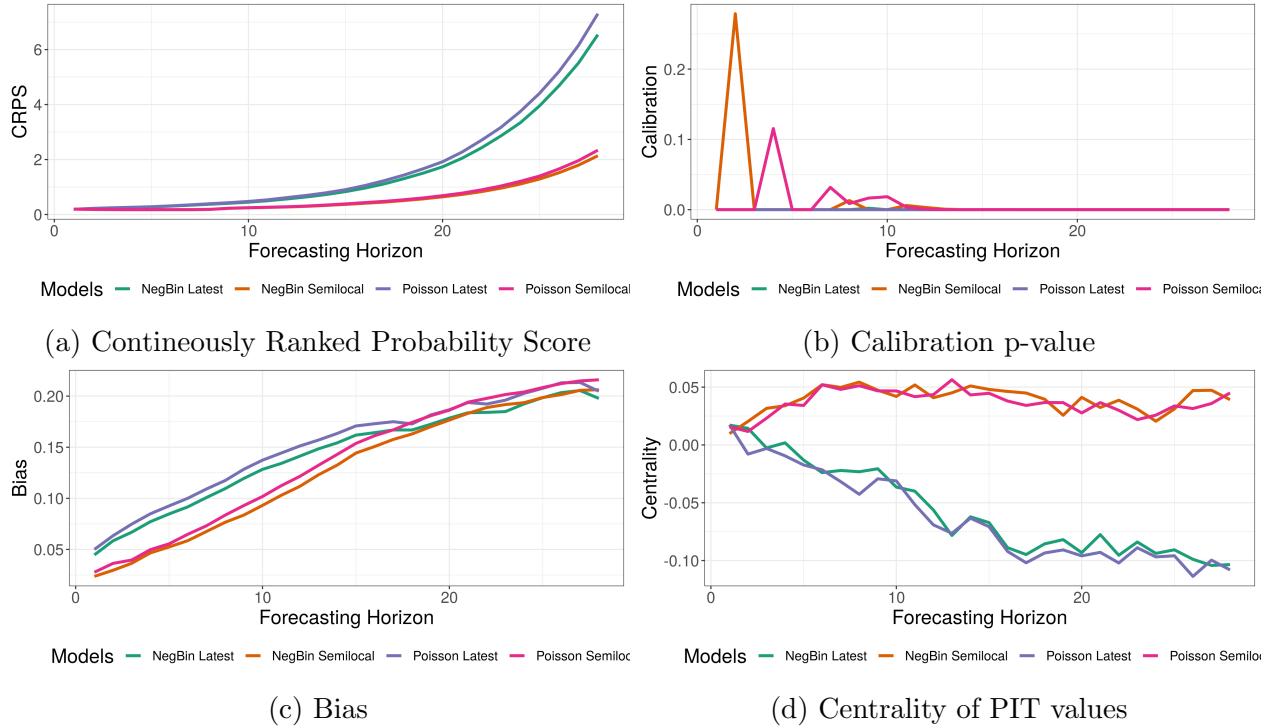
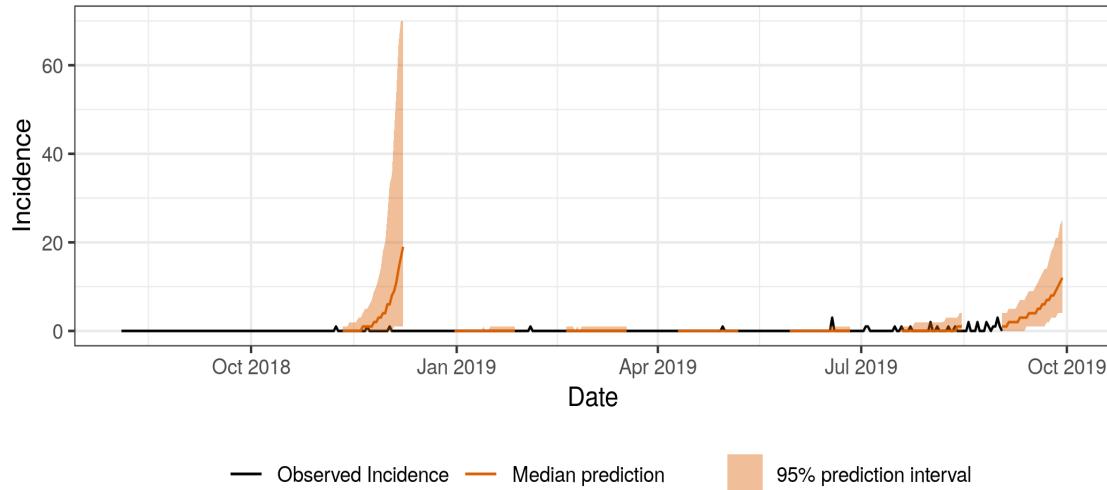
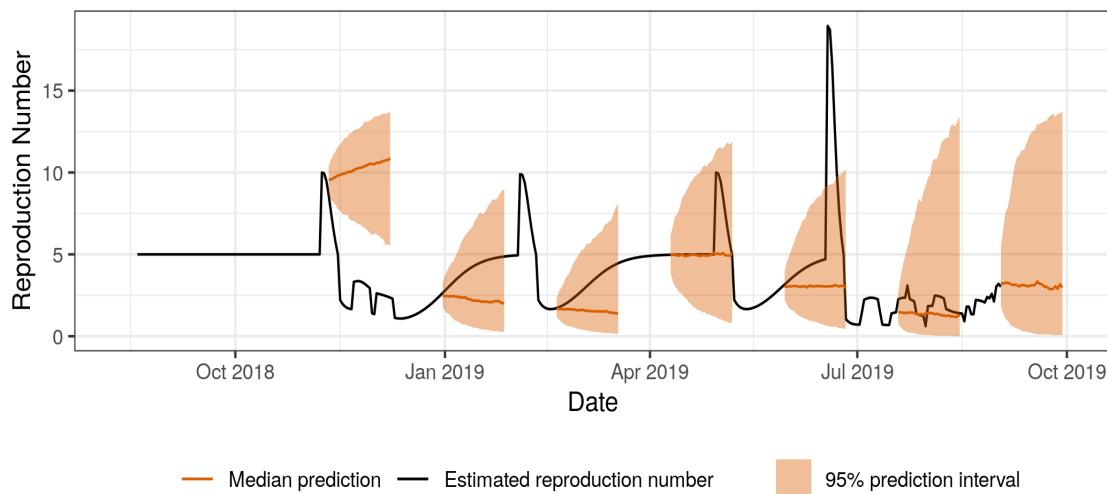


Figure 44: Scores for Musienene as a function of the forecasting horizon.

T Mutwanga



(a) Forecasted and predicted incidence for the best fitting model



(b) Forecasted and predicted reproduction numbers for the semilocal poisson model

Figure 45: Median forecast with 95 % prediction intervals and observed values for incidence and reproduction number for the semilocal poisson model for Mutwanga.

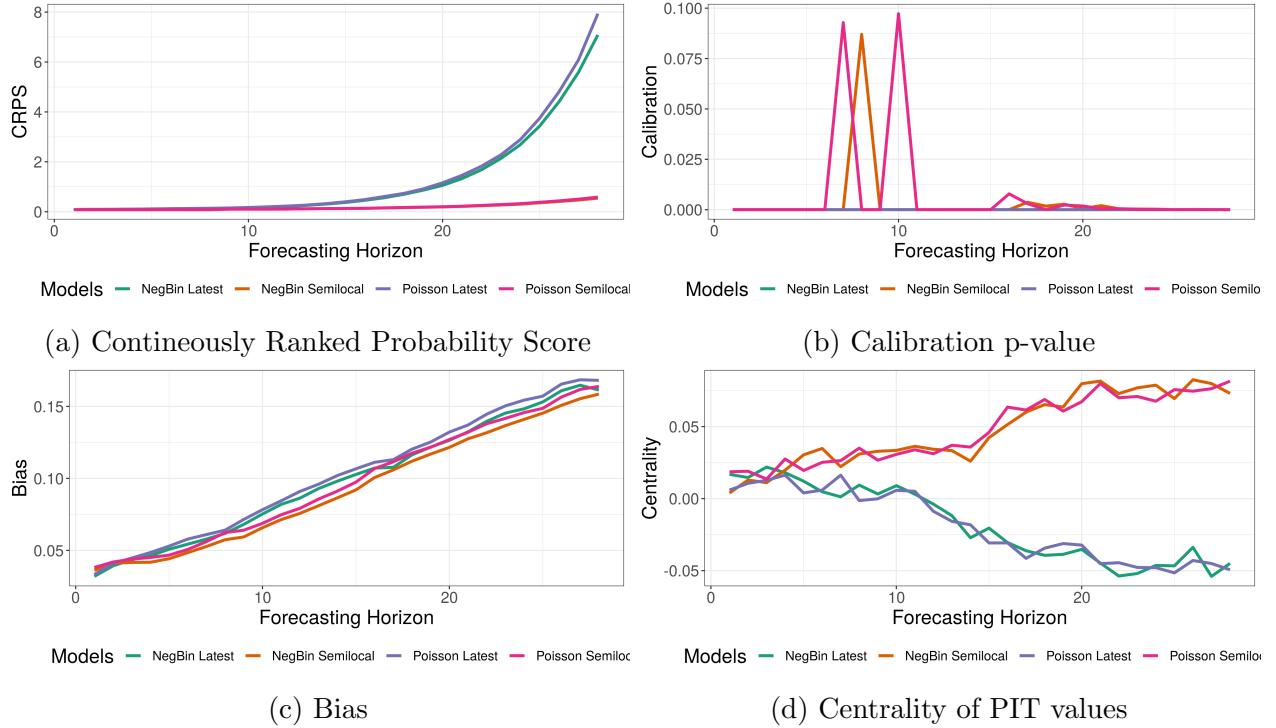
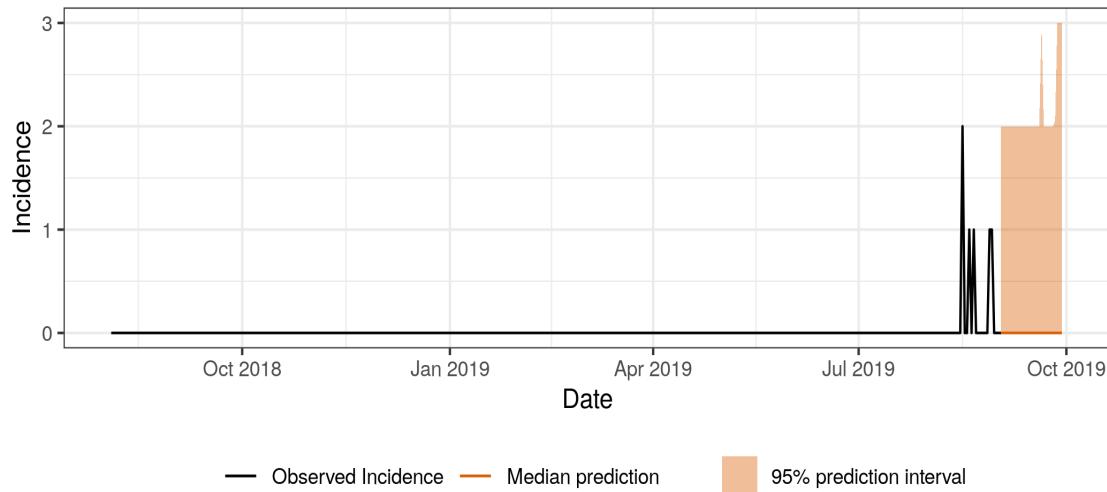
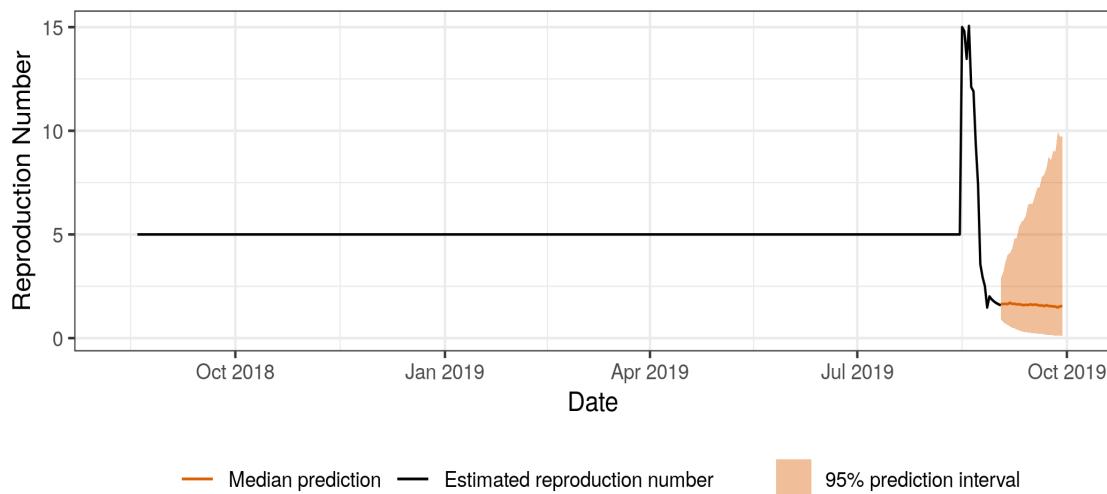


Figure 46: Scores for Mutwanga as a function of the forecasting horizon.

U Mwenga



(a) Forecasted and predicted incidence for the best fitting model



(b) Forecasted and predicted reproduction numbers for the semilocal poisson model

Figure 47: Median forecast with 95 % prediction intervals and observed values for incidence and reproduction number for the semilocal poisson model for Mwenga.

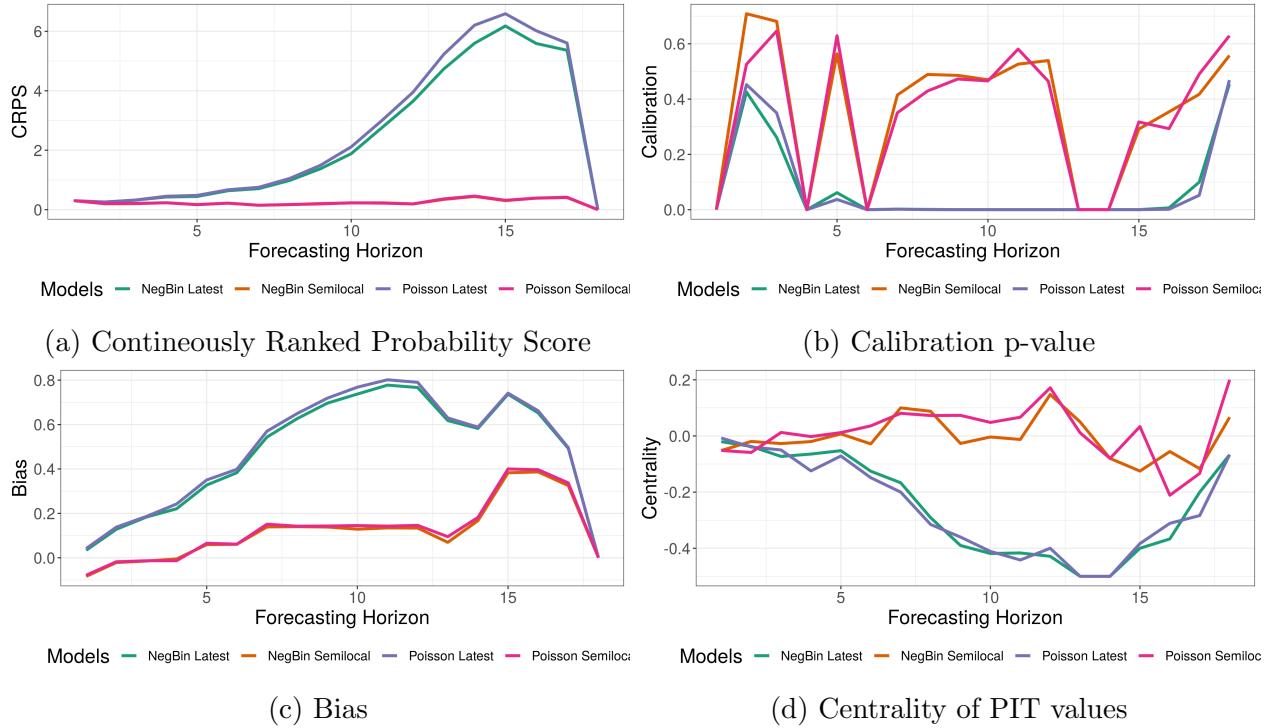
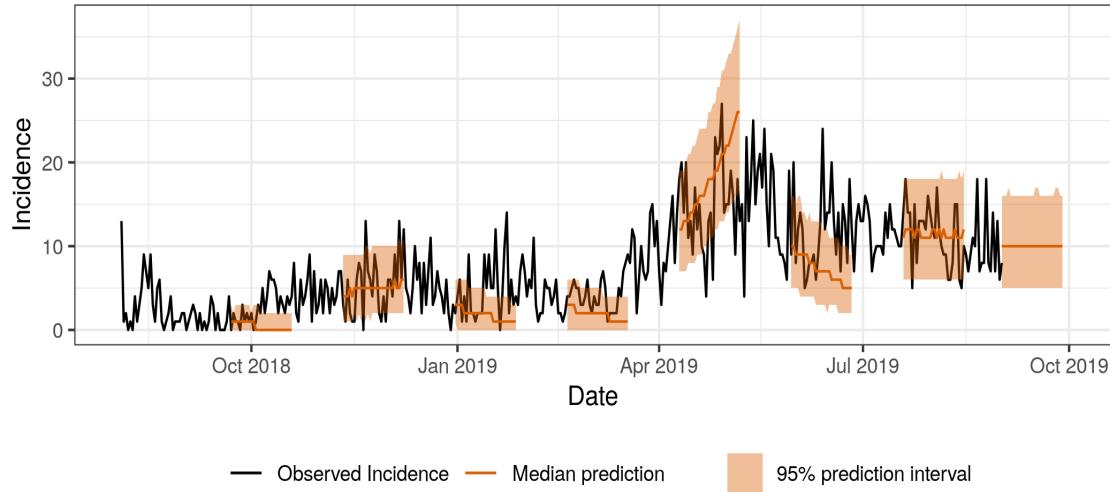
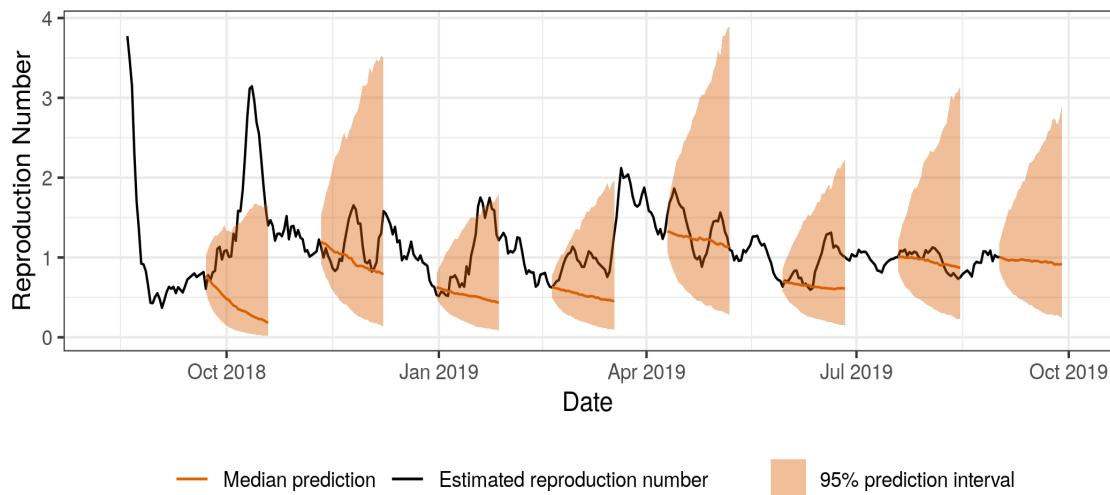


Figure 48: Scores for Mwenga as a function of the forecasting horizon.

V national



(a) Forecasted and predicted incidence for the best fitting model



(b) Forecasted and predicted reproduction numbers for the semilocal poisson model

Figure 49: Median forecast with 95 % prediction intervals and observed values for incidence and reproduction number for the semilocal poisson model for national.

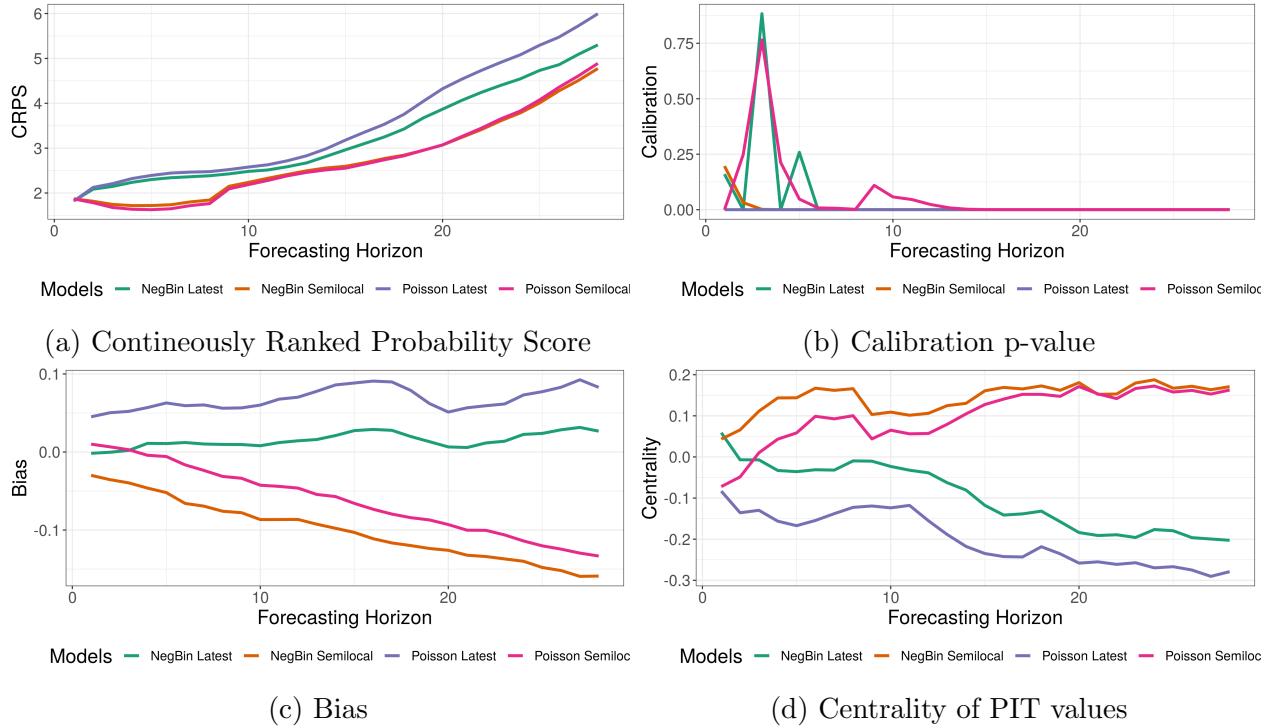
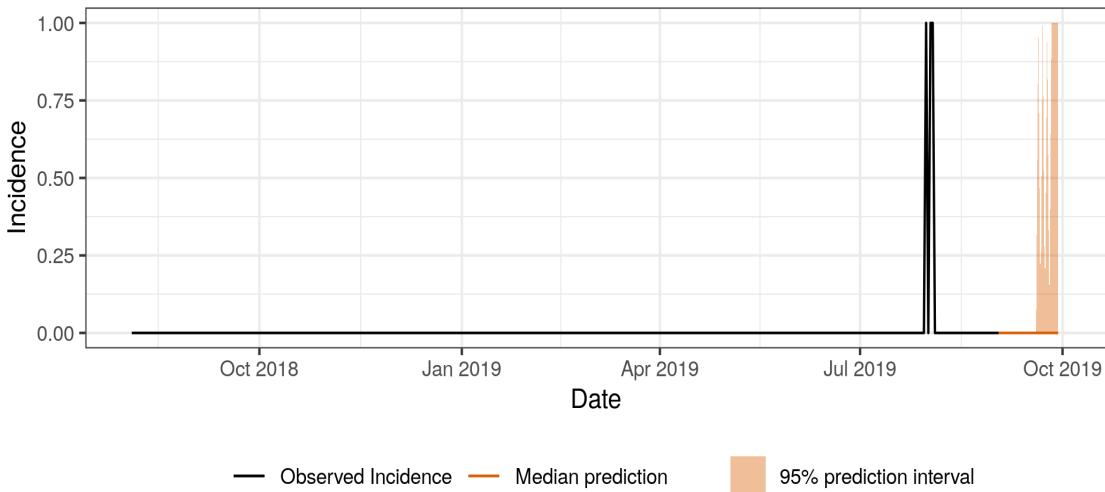
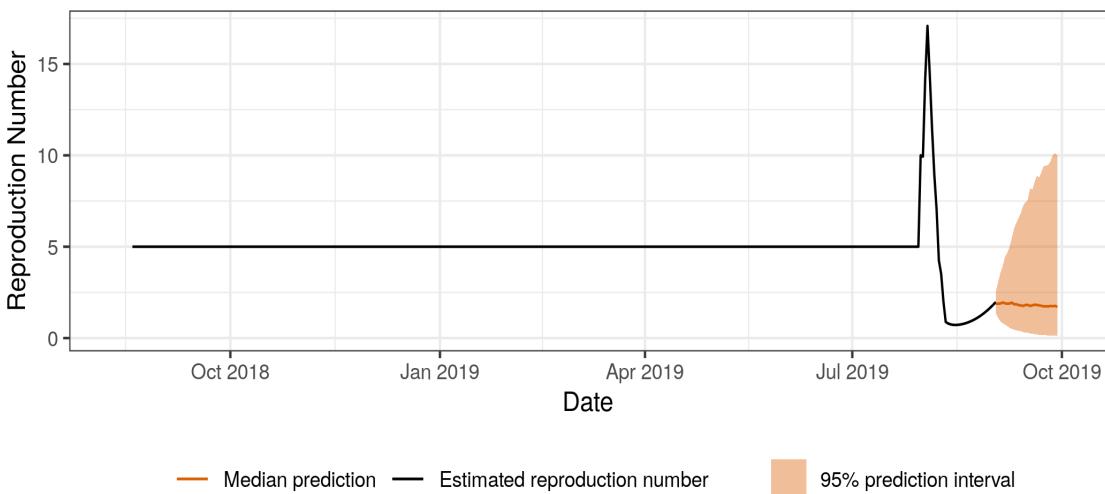


Figure 50: Scores for national as a function of the forecasting horizon.

W Nyiragongo



(a) Forecasted and predicted incidence for the best fitting model



(b) Forecasted and predicted reproduction numbers for the semilocal poisson model

Figure 51: Median forecast with 95 % prediction intervals and observed values for incidence and reproduction number for the semilocal poisson model for Nyiragongo.

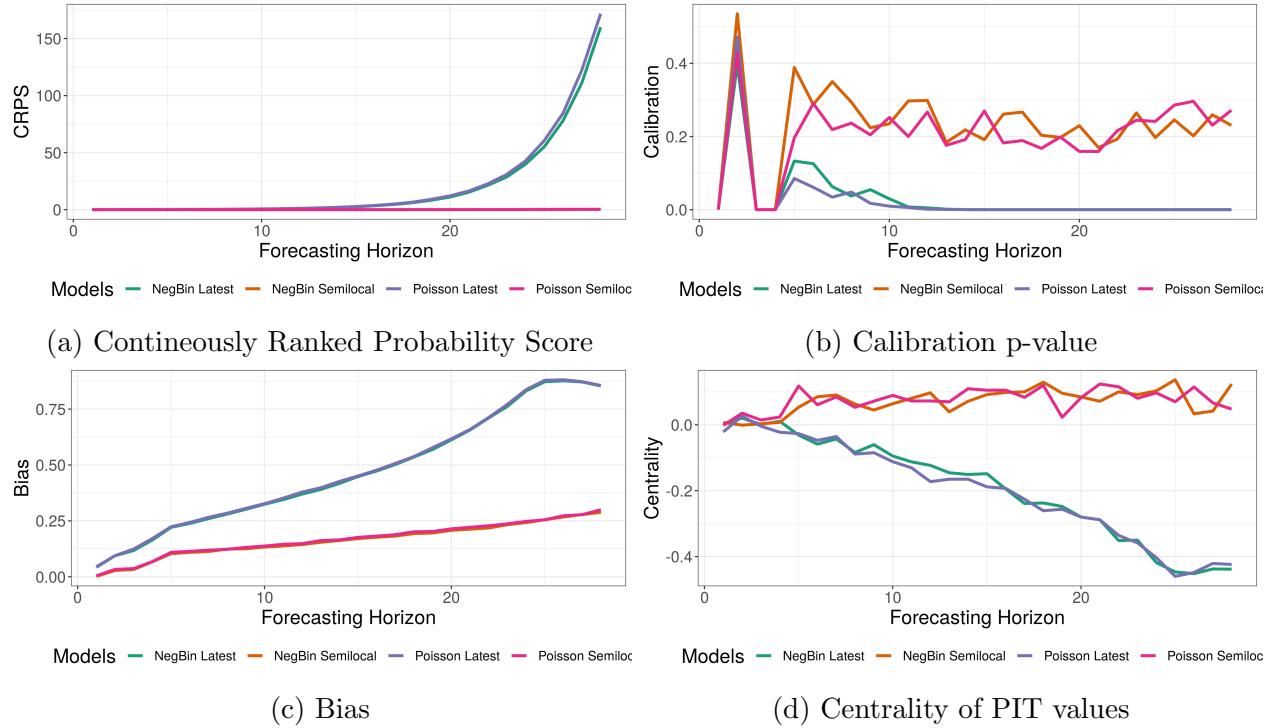
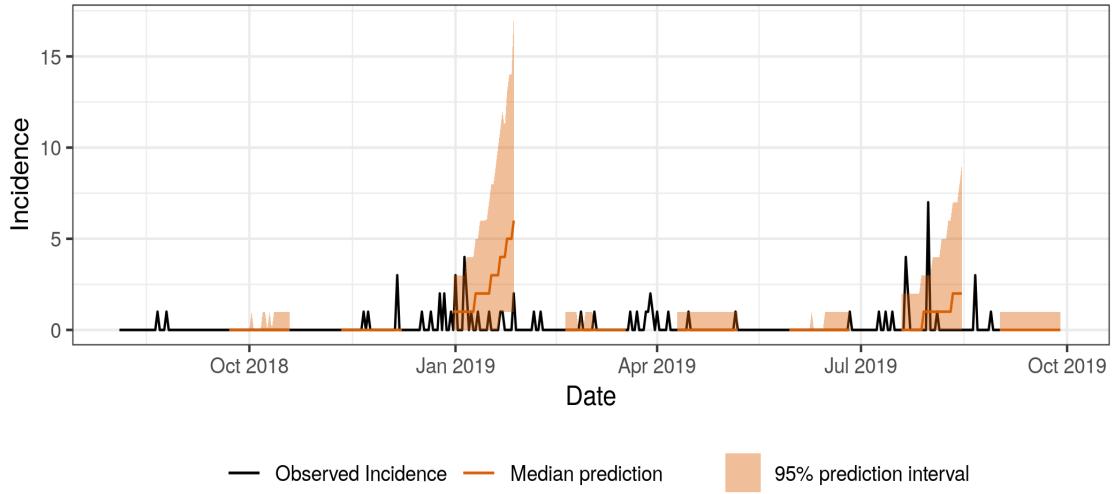
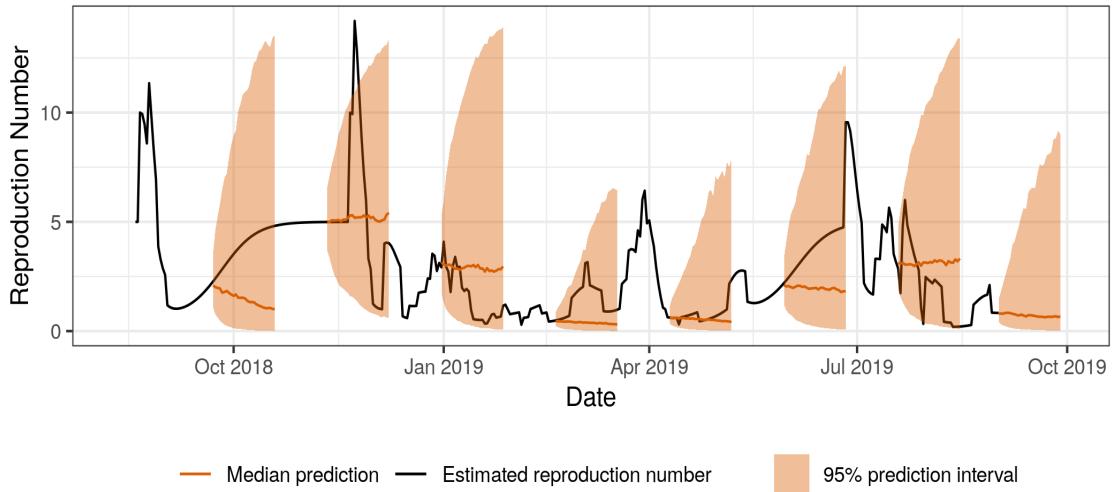


Figure 52: Scores for Nyiragongo as a function of the forecasting horizon.

X Oicha



(a) Forecasted and predicted incidence for the best fitting model



(b) Forecasted and predicted reproduction numbers for the semilocal poisson model

Figure 53: Median forecast with 95 % prediction intervals and observed values for incidence and reproduction number for the semilocal poisson model for Oicha.

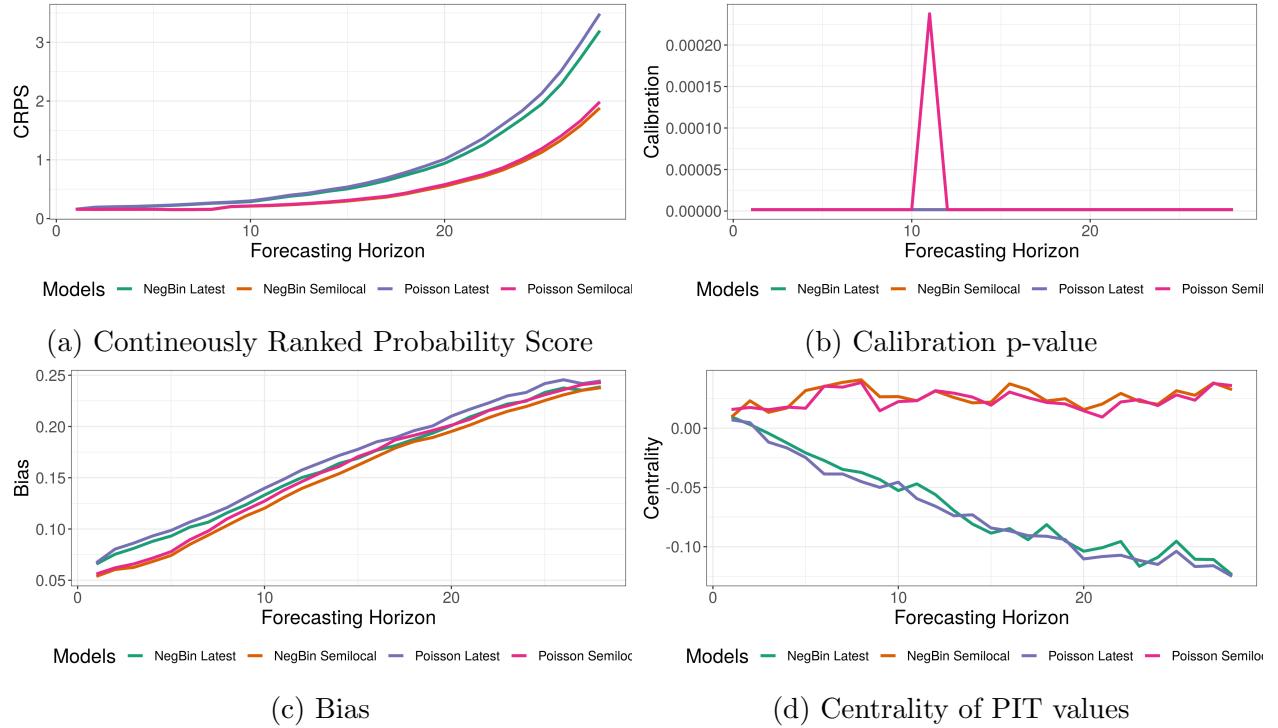
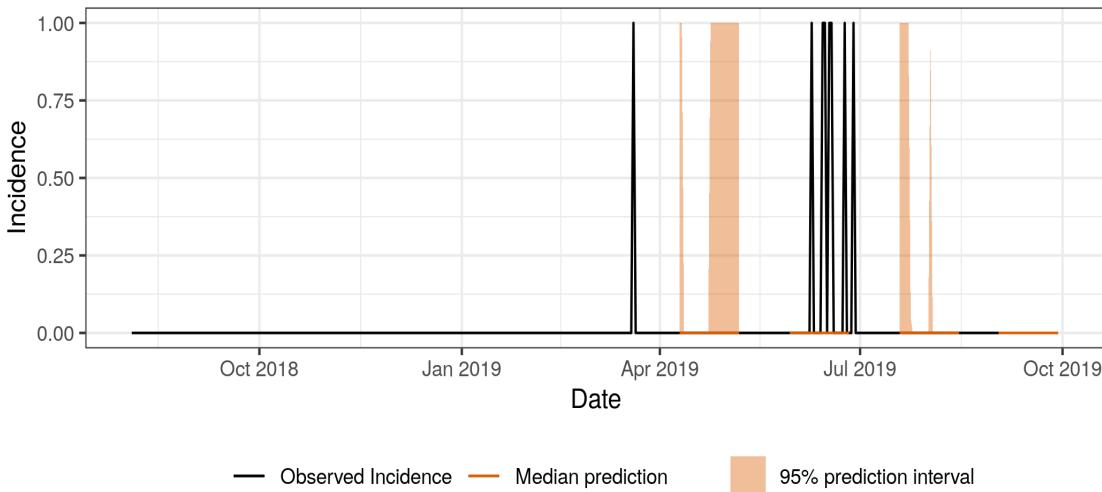
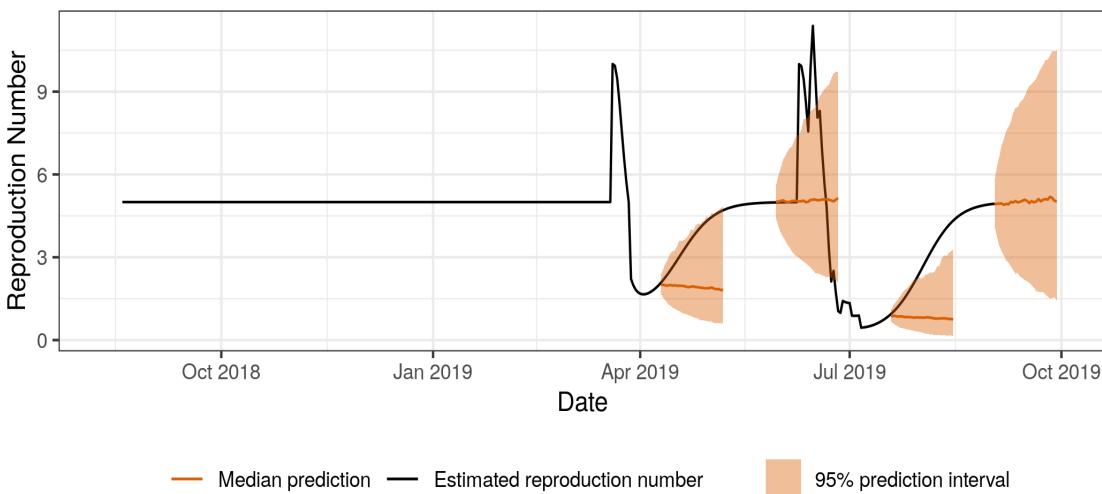


Figure 54: Scores for Oicha as a function of the forecasting horizon.

Y Rwampara



(a) Forecasted and predicted incidence for the best fitting model



(b) Forecasted and predicted reproduction numbers for the semilocal poisson model

Figure 55: Median forecast with 95 % prediction intervals and observed values for incidence and reproduction number for the semilocal poisson model for Rwampara.

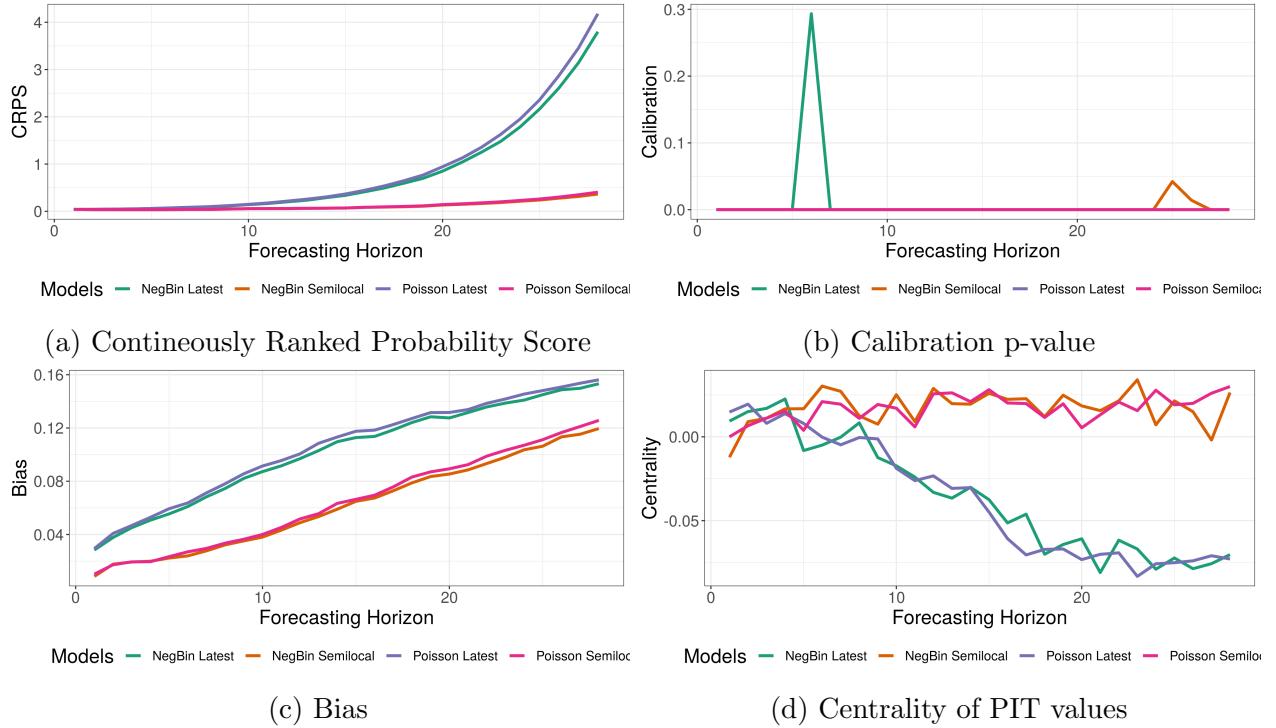
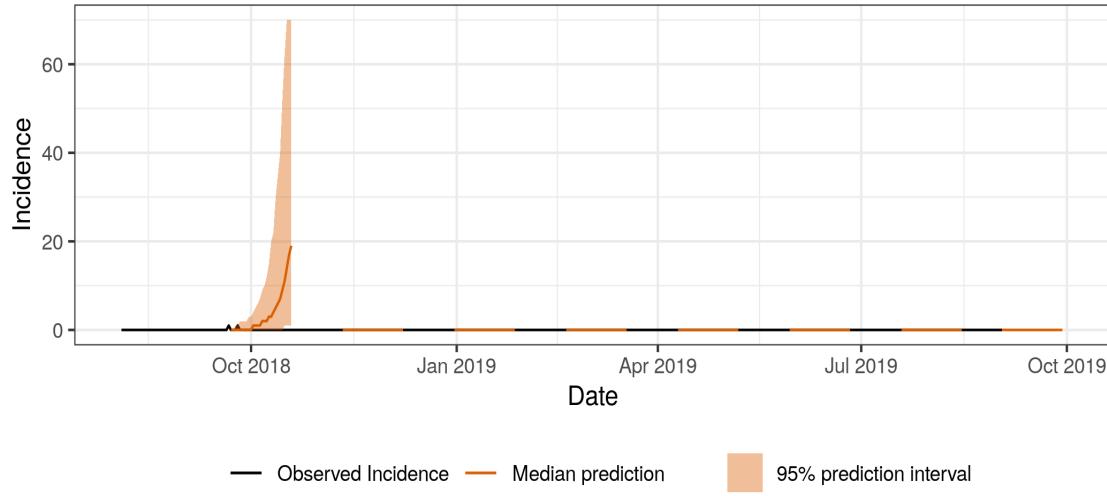
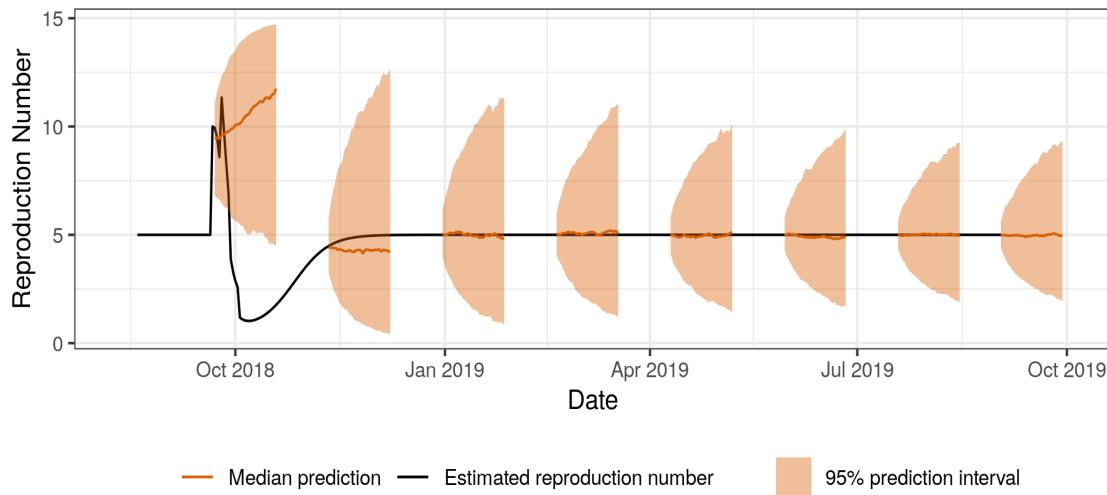


Figure 56: Scores for Rwampara as a function of the forecasting horizon.

Z Tchomia



(a) Forecasted and predicted incidence for the best fitting model



(b) Forecasted and predicted reproduction numbers for the semilocal poisson model

Figure 57: Median forecast with 95 % prediction intervals and observed values for incidence and reproduction number for the semilocal poisson model for Tchomia.

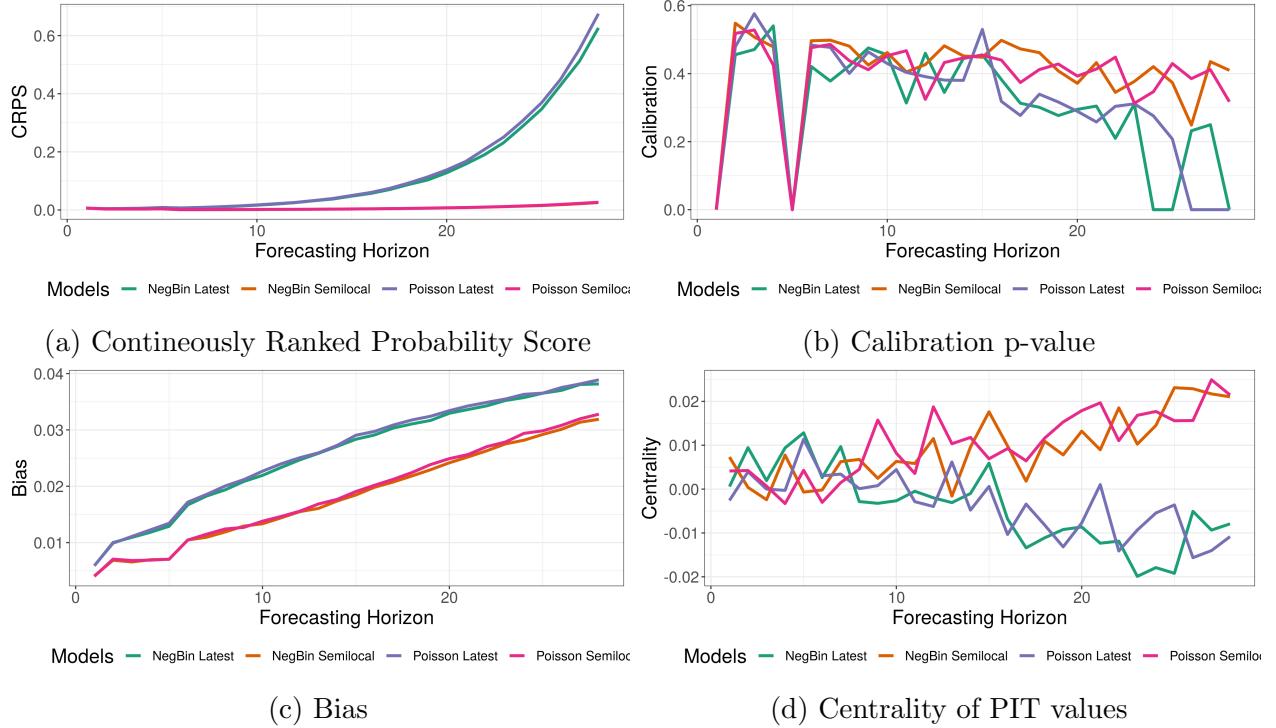
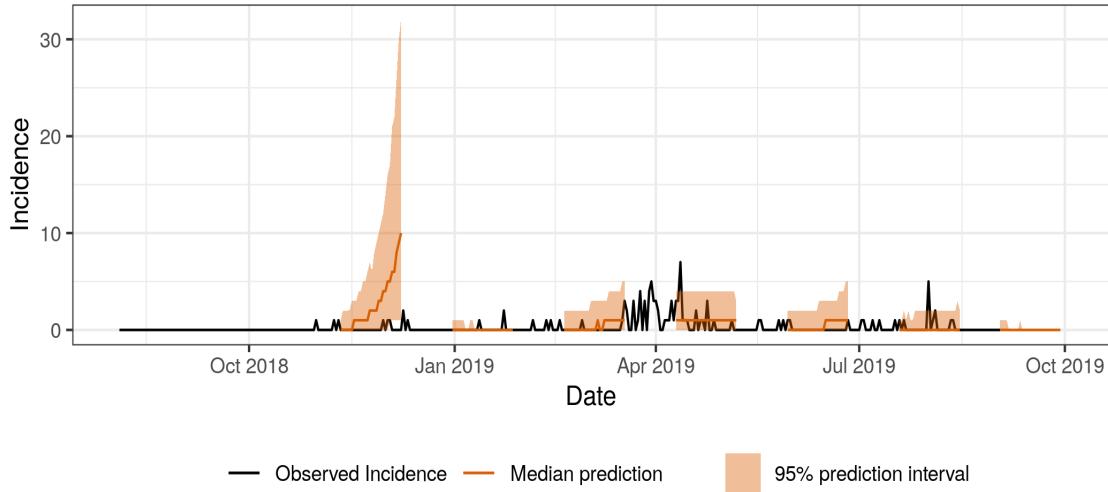
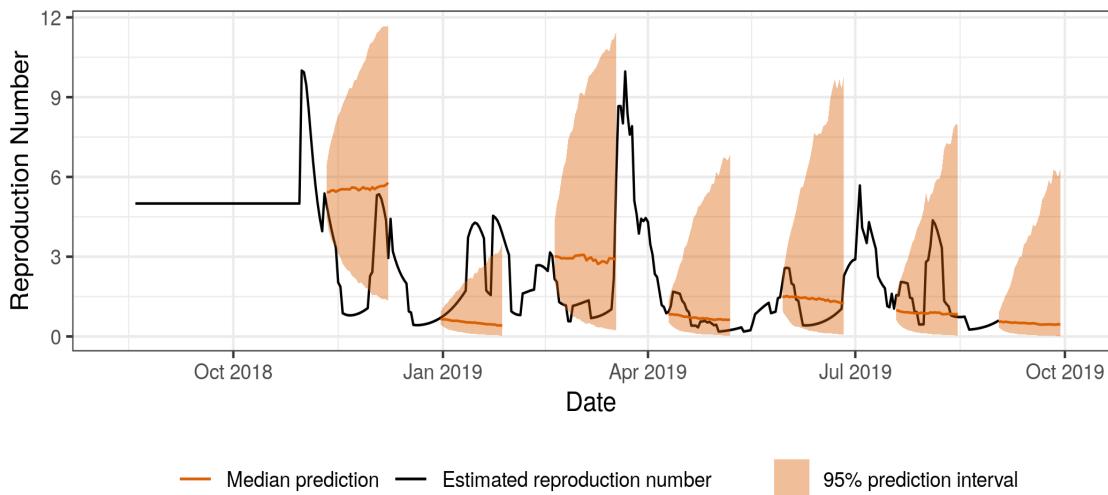


Figure 58: Scores for Tchomia as a function of the forecasting horizon.

Vuhovi



(a) Forecasted and predicted incidence for the best fitting model



(b) Forecasted and predicted reproduction numbers for the semilocal poisson model

Figure 59: Median forecast with 95 % prediction intervals and observed values for incidence and reproduction number for the semilocal poisson model for Vuhovi.

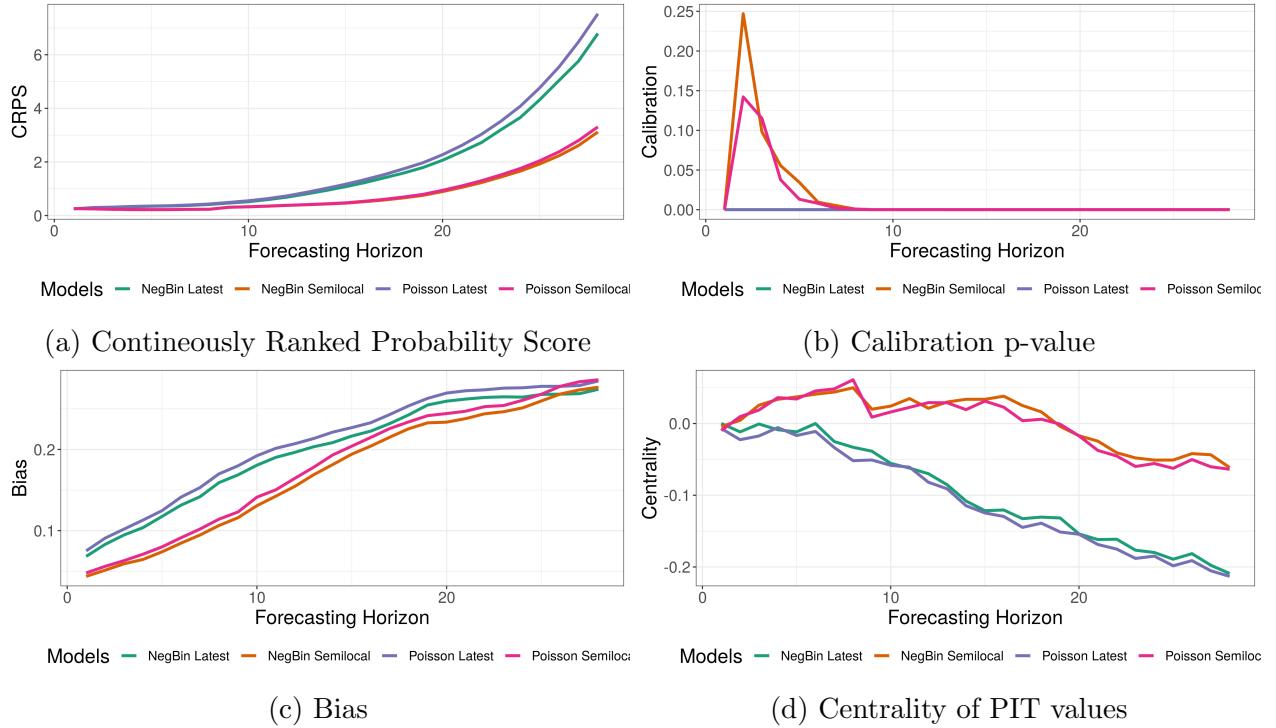


Figure 60: Scores for Vuhovi as a function of the forecasting horizon.