Do we really need Foundation Models for multi-step-ahead Epidemic Forecasting?

Mrinmoy Dey*
IIT Kharagpur
mdey3368@kgpian.iitkgp.ac.in

Aprameyo Chakrabartty*
IIT Kharagpur
aprameyo@kgpian.iitkgp.ac.in

Dhruv Sarkar*
IIT Kharagpur
dhruv.sarkar@kgpian.iitkgp.ac.in

Tanujit Chakraborty
Sorbonne University
tanujitisi@gmail.com

Abstract

The emergence of Foundation Models has radically transformed the Deep Learning scene and also accelerated its adoption in other domains. In particular, Large Language Models (LLMs) are being used in many time series forecasting tasks including Epidemic Forecasting. While the adoption of a new technology is generally a good sign, we must be scientific in analysing the benefits of doing so. We try two LLMs used in time series forecasting and show that on average, they perform almost similar or marginally better to the very popular classical statistical method ARIMA when applied to epidemic forecasting. We have performed extensive experiments on many Epidemic Forecasting datasets and thoroughly validated our conclusion that we need Foundation Models like LLMs for Epidemic Forecasting for the growth of the field even if the benefits are not proportionate to the costs immediately.

1 Introduction

Despite the recent surge in interest and investment into Foundation Models like LLMs, their application in domains outside of natural language processing remains an open question. LLMs have demonstrated remarkable abilities in language generation, translation, and understanding, but transferring these strengths to other domains such as time series forecasting, specifically epidemic forecasting, presents several challenges. Foremost among these are the domain-specific nuances that LLMs may fail to capture, particularly when handling epidemiological data that follows complex and irregular patterns unlike textual sequences.

Classical statistical models, such as ARIMA [1], and machine learning-based approaches, including Long Short-Term Memory (LSTM) [2]and Temporal Convolutional Networks (TCNs)[3], have been traditionally applied to time series forecasting problems. These models have proven robust and efficient, even with comparatively lower computational costs. In contrast, the computational resources required to fine-tune and deploy LLMs on such tasks are substantial. Not only do these models demand vast amounts of data and computing power, but their interpretability and reliability in the context of epidemic forecasting remain questionable. All these factors naturally provoke some cynicism regarding the applicability of LLMs in domain-specific tasks.

In this work, we aim to rigorously assess whether LLMs offer any real advantages over traditional statistical and machine learning models for epidemic forecasting. We believe that despite the

*Equal contribution

NeurIPS 2024 Workshop on Time Series in the Age of Large Models.

growing allure of using sophisticated models for every task, there are instances where simplicity can outperform complexity. Through a series of comprehensive experiments across multiple epidemic datasets, we evaluate the efficacy of popular LLMs when compared to baseline statistical models.

Our findings reveal that LLMs are marginally better than traditional methods of epidemic forecasting based on evaluations across many different datasets and using different metrics. Ultimately, this study seeks to answer the fundamental question: Do we really need Foundation Models for Epidemic Forecasting, or are traditional methods more suitable for this critical task?

The remainder of this paper is organized as follows: In Section 2, we review the background and related work on epidemic forecasting and the role of deep learning models, particularly LLMs. Section 3 details our experimental setup, including datasets, models, and evaluation metrics. Section 4 presents our results and analysis. Finally, Section 5 concludes the paper with our recommendations and potential future directions for epidemic forecasting research.

2 Related Work

2.1 Epidemic Forecasting: Challenges and Methods

Traditional approaches to epidemic forecasting have relied heavily on statistical models[4]. These models include auto-regressive integrated moving average (ARIMA)[5], exponential smoothing[6], and compartmental models such as the SIR (Susceptible-Infectious-Recovered) model [7] and its variations. While these methods have proven to be useful in short-term forecasting, they often struggle to capture long-range dependencies and the intricate, nonlinear nature of disease transmission in real-world scenarios[8].

Recent advances in machine learning and deep learning have led to the application of more sophisticated models to epidemic forecasting. Techniques such as Long Short-Term Memory (LSTM) networks [9], Temporal Convolutional Networks (TCNs)[10], and hybrid models combining statistical and machine learning approaches have shown promise in capturing the temporal dependencies and complex patterns present in epidemic data. These models have provided notable improvements over traditional methods, especially in medium- to long-term forecasts where patterns in the data are more intricate[8].

2.2 The Rise of Large Language Models (LLMs) and Foundation Models

The central idea behind the use of LLMs in time series tasks is that sequences of data, whether they be words or numerical measurements, share some structural similarities. Like sentences in a language, time series data points can be viewed as temporally dependent, following certain patterns. Consequently, LLMs, with their strong sequence modelling capabilities, have been adapted for tasks beyond NLP, including financial forecasting[11], traffic accident forecasting[12], agricultural meteorological recommendations[13], and epidemic forecasting.

Despite their versatility, LLMs come with several limitations when applied outside of their native NLP domain[14]. They require large amounts of data for training and fine-tuning, making them resource-intensive. Additionally, their black-box nature raises concerns regarding interpretability, which is a crucial factor in high-stakes fields such as public health. While LLMs have shown promise in general time series forecasting tasks[15], their performance in specific, highly nuanced domains such as epidemic forecasting remains underexplored.

3 Experiments

The experiments were conducted on 15 epidemic forecasting datasets. Four of these datasets comprised monthly data, while the remaining 11 were composed of weekly data points. The models were evaluated on short-term, medium-term, and long-term forecasting tasks. For the short-term forecast, the evaluation period was set to 12 weeks for weekly and 6 months for monthly datasets. For medium-term forecasting, this extended to 26 weeks and 12 months, respectively. For long-term forecasting, the models were tested over 52 weeks for weekly and 24 months for monthly datasets. We compared the performance of two large language models (LLMs), Chronos[16] and GTT[17], against the ARIMA model across these datasets. The tables of the experiments and the definitions of the performance metrics (MSE, NRMSE, sMAPE, MASE) are given in the appendix below.

4 Results

In all the plots, a lower value in the y-axis indicates better performance, as all these metrics quantify error.

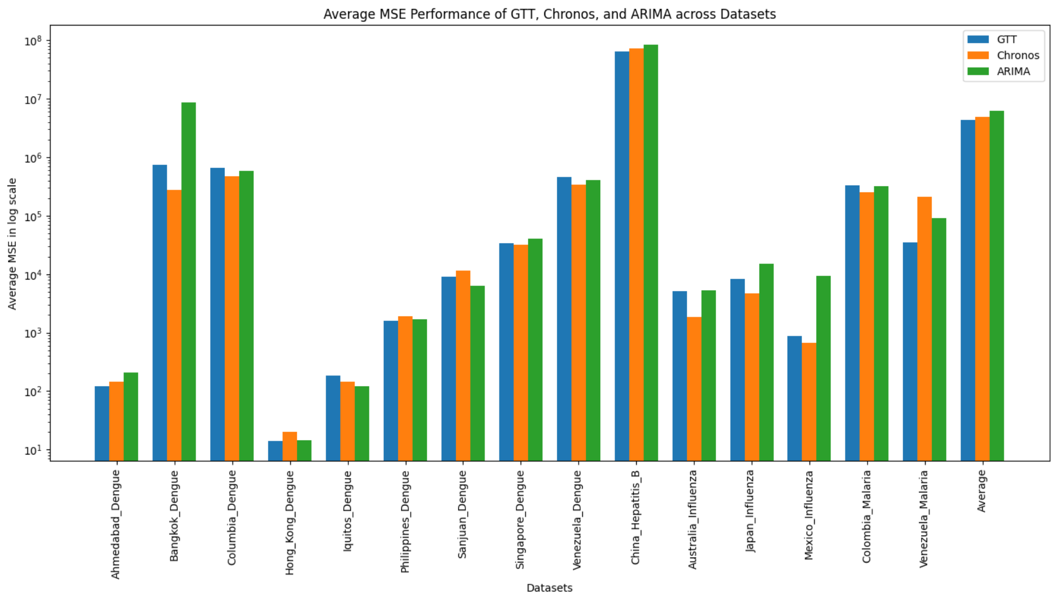

Figure 1: Average MSE performance

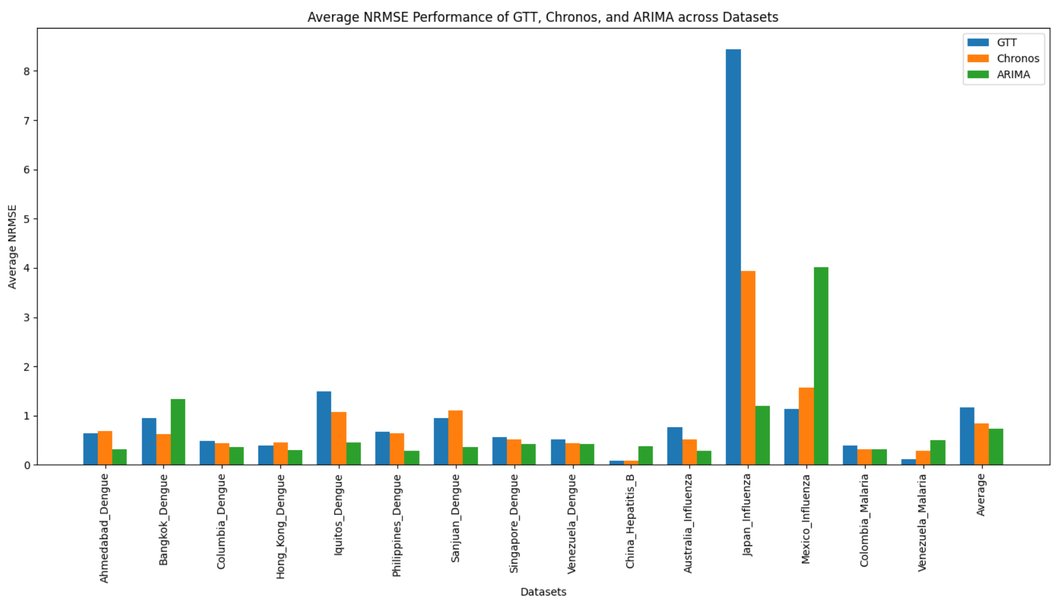

Figure 2: Average NRMSE performance

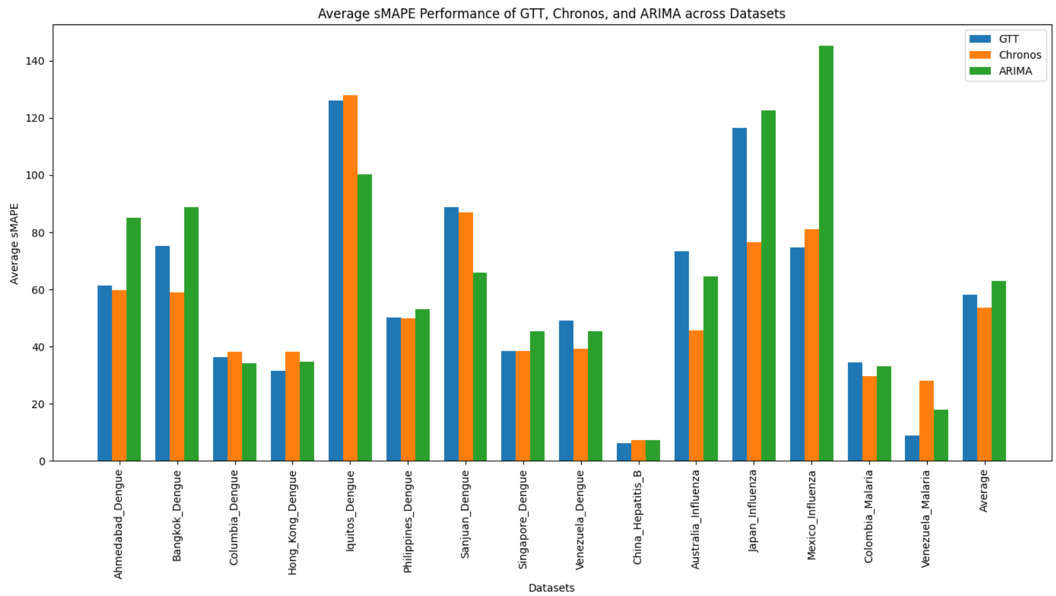

Figure 3: Average sMAPE performance

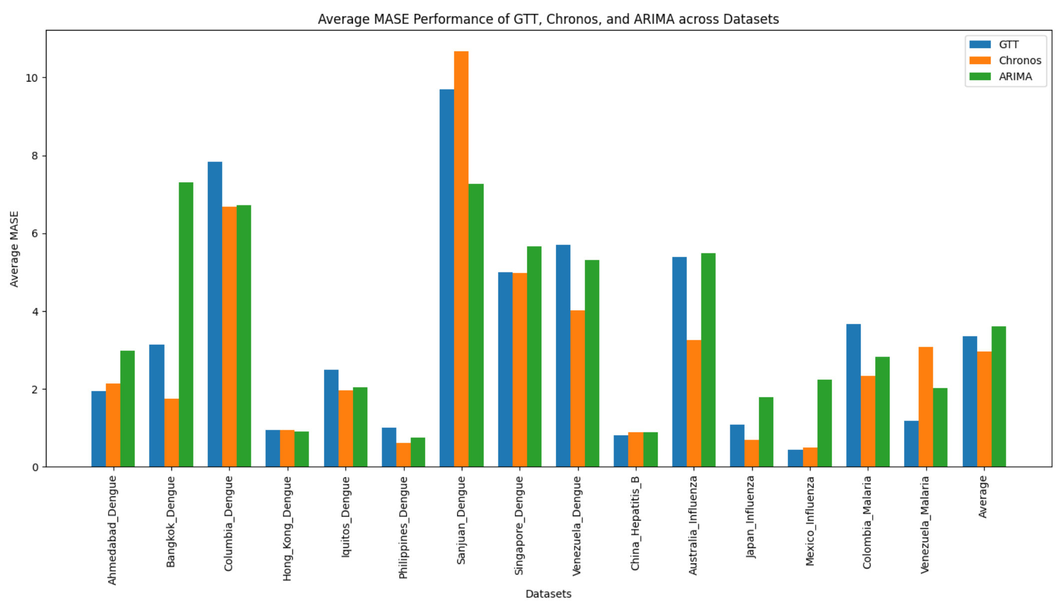

Figure 4: Average MASE performance

5 Conclusion

In this paper, we explored whether Foundation Models like LLMs are necessary for Epidemic Forecasting. Contrary to our initial expectations, we found that LLMs outperformed traditional statistical methods and machine learning models. While concerns about their computational cost are valid, the improved accuracy and generalization across diverse epidemic datasets suggest that the benefits of LLMs can justify their use in many cases. Our results demonstrate that LLMs hold significant potential for advancing epidemic forecasting, especially in scenarios demanding high precision and adaptability.

References

[1] Soheila Mehrmolaei and Mohammad Reza Keyvanpour. Time series forecasting using improved arima. In *2016 Artificial Intelligence and Robotics (IRANOPEN)*, pages 92–97, 2016.

[2] Sima Siami-Namini, Neda Tavakoli, and Akbar Siami Namin. A comparison of arima and lstm in forecasting time series. In *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 1394–1401, 2018.

[3] Pedro Lara-Benítez, Manuel Carranza-García, José M Luna-Romera, and José C Riquelme. Temporal convolutional networks applied to energy-related time series forecasting. *applied sciences*, 10(7):2322, 2020.

[4] Igor Nesteruk. Statistics-based predictions of coronavirus epidemic spreading in mainland china. 2020.

[5] Alok Kumar Sahai, Namita Rath, Vishal Sood, and Manvendra Pratap Singh. Arima modelling & forecasting of covid-19 in top five affected countries. *Diabetes & metabolic syndrome: clinical research & reviews*, 14(5):1419–1427, 2020.

[6] Ana Keuly Luz Bezerra and Éthynna Marina Corrêa Santos. Prediction the daily number of confirmed cases of covid-19 in sudan with arima and holt winter exponential smoothing. *International Journal of Development Research*, 10(08):39408–39413, 2020.

[7] Agus Kartono, Savira Vita Karimah, Setyanto Tri Wahyudi, Ardian Arif Setiawan, and Irmansyah Sofian. Forecasting the long-term trends of coronavirus disease 2019 (covid-19) epidemic using the susceptible-infectious-recovered (sir) model. *Infectious Disease Reports*, 13(3):668–684, 2021.

[8] Sima Siami-Namini, Neda Tavakoli, and Akbar Siami Namin. A comparison of arima and lstm in forecasting time series. In *2018 17th IEEE international conference on machine learning and applications (ICMLA)*, pages 1394–1401. Ieee, 2018.

[9] Osman Tayfun Bişkin. Multi-step forecasting of covid-19 cases in european countries using temporal convolutional networks. *Mugla Journal of Science and Technology*, 7(1):117–126, 2021.

[10] Naoki Masuda and Petter Holme. Predicting and controlling infectious disease epidemics using temporal networks. *F1000prime reports*, 5, 2013.

[11] Huaqin Zhao, Zhengliang Liu, Zihao Wu, Yiwei Li, Tianze Yang, Peng Shu, Shaochen Xu, Haixing Dai, Lin Zhao, Gengchen Mai, et al. Revolutionizing finance with llms: An overview of applications and insights. *arXiv preprint arXiv:2401.11641*, 2024.

[12] Irene de Zarzà, Joachim de Curtò, Gemma Roig, and Carlos T Calafate. Llm multimodal traffic accident forecasting. *Sensors*, 23(22):9225, 2023.

[13] Ji-jun Park and Soo-joon Choi. Llms for enhanced agricultural meteorological recommendations. *arXiv preprint arXiv:2408.04640*, 2024.

[14] Muhammad Usman Hadi, Rizwan Qureshi, Abbas Shah, Muhammad Irfan, Anas Zafar, Muhammad Bilal Shaikh, Naveed Akhtar, Jia Wu, Seyedali Mirjalili, et al. A survey on large language models: Applications, challenges, limitations, and practical usage. *Authorea Preprints*, 2023.

[15] Mingyu Jin, Hua Tang, Chong Zhang, Qinkai Yu, Chengzhi Liu, Suiyuan Zhu, Yongfeng Zhang, and Mengnan Du. Time series forecasting with llms: Understanding and enhancing model capabilities. *arXiv preprint arXiv:2402.10835*, 2024.

[16] Abdul Fatir Ansari, Lorenzo Stella, Caner Turkmen, Xiyuan Zhang, Pedro Mercado, Huibin Shen, Oleksandr Shchur, Syama Sundar Rangapuram, Sebastian Pineda Arango, Shubham Kapoor, Jasper Zschiegner, Danielle C. Maddix, Hao Wang, Michael W. Mahoney, Kari Torkkola, Andrew Gordon Wilson, Michael Bohlke-Schneider, and Yuyang Wang. Chronos: Learning the language of time series, 2024.

[17] Cheng Feng, Long Huang, and Denis Krompass. Only the curve shape matters: Training foundation models for zero-shot multivariate time series forecasting through next curve shape prediction, 2024.

A Appendix

Performance Metrics

Mean Squared Error (MSE)

The Mean Squared Error (MSE) is defined as:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

where y_i is the actual value and $\hat{y}_i$ is the predicted value.

Normalized Root Mean Squared Error (NRMSE)

The Normalized Root Mean Squared Error (NRMSE) is defined as:

$$NRMSE = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}}{range(y)}$$

where $range(y) = \max(y) - \min(y)$ is the range of the actual values, y_i is the actual value and $\hat{y}_i$ is the predicted value.

Mean Absolute Scaled Error (MASE)

The Mean Absolute Scaled Error (MASE) is defined as:

$$MASE = \frac{\frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|}{\frac{1}{n-1}\sum_{i=2}^{n}|y_i - y_{i-1}|}$$

where the denominator is the mean absolute error of a naive forecast, y_i is the actual value and $\hat{y}_i$ is the predicted value.

Symmetric Mean Absolute Percentage Error (SMAPE)

The Symmetric Mean Absolute Percentage Error (SMAPE) is defined as:

$$SMAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|y_i - \hat{y}_i|}{\frac{|y_i|+|\hat{y}_i|}{2}} \times 100\%$$

where y_i is the actual value and $\hat{y}_i$ is the predicted value.

Tables

The tables of data are given here

Tables of average of short, medium and length predictions

Datasets	GTT (Avg)	Chronos (Avg)	ARIMA (Avg)
Ahmedabad_Dengue	122.6004	143.6623	207.1183
Bangkok_Dengue	740825.7504	275728.3536	8550124.604
Columbia_Dengue	652858.6193	475264.3818	580755.9641
Hong_Kong_Dengue	13.8348	19.9225	14.446
Iquitos_Dengue	184.3089	145.5153	119.4936
Philippines_Dengue	1596.7497	1881.6994	1706.9289
Sanjuan_Dengue	9047.1206	11499.1600	6382.6534
Singapore_Dengue	33718.4547	31537.8512	40336.3723
Venezuela_Dengue	464906.4173	338874.0402	401445.7626
China_Hepatitis_B	63755587.85	71747238.59	83331279.89
Australia_Influenza	5151.4435	1850.9421	5287.3691
Japan_Influenza	8167.124	4725.788	15093.8094
Mexico_Influenza	862.9881	662.4614	9334.4113
Colombia_Malaria	327205.35	253106.5358	321406.2418
Venezuela_Malaria	34424.3747	212854.3003	90574.99003

Table 1: MSE Average Performance of GTT, Chronos, and ARIMA on Different Datasets

Datasets	GTT (Avg)	Chronos (Avg)	ARIMA (Avg)
Ahmedabad_Dengue	0.64263	0.69180	0.31840
Bangkok_Dengue	0.95577	0.62730	1.33147
Columbia_Dengue	0.48413	0.44290	0.36627
Hong_Kong_Dengue	0.39377	0.45913	0.30797
Iquitos_Dengue	1.49687	1.06730	0.45653
Philippines_Dengue	0.66927	0.64290	0.29087
Sanjuan_Dengue	0.95823	1.11163	0.37060
Singapore_Dengue	0.55843	0.51633	0.42817
Venezuela_Dengue	0.51537	0.43363	0.41827
China_Hepatitis_B	0.08113	0.08517	0.38250
Australia_Influenza	0.76350	0.51673	0.28390
Japan_Influenza	8.44523	3.93940	1.20380
Mexico_Influenza	1.13933	1.56350	4.00750
Colombia_Malaria	0.39090	0.32340	0.31303
Venezuela_Malaria	0.11833	0.27893	0.50407

Table 2: NRMSE (Normalized Root Mean Square Error) Averages for Short Term, Medium Term, and Long Term Predictions across Datasets

Datasets	GTT (Avg)	Chronos (Avg)	ARIMA (Avg)
Ahmedabad_Dengue	61.35187	59.77397	85.16460
Bangkok_Dengue	75.18923	58.94240	88.70867
Columbia_Dengue	36.47127	38.32477	34.31167
Hong_Kong_Dengue	31.55783	38.20170	34.70473
Iquitos_Dengue	126.07760	127.92887	100.13343
Philippines_Dengue	50.17337	49.82253	53.09113
Sanjuan_Dengue	88.93157	86.92853	65.88140
Singapore_Dengue	38.38880	38.52190	45.42720
Venezuela_Dengue	49.06920	39.37093	45.52933
China_Hepatitis_B	6.25870	7.26393	7.32417
Australia_Influenza	73.28930	45.68120	64.65940
Japan_Influenza	116.47000	76.62087	122.62970
Mexico_Influenza	74.68967	81.07277	145.35950
Colombia_Malaria	34.59950	29.76523	33.27630
Venezuela_Malaria	9.01937	28.04663	17.96730

Table 3: sMAPE Average Values of Short Term, Medium Term, and Long Term Predictions across Datasets

Datasets	GTT (Avg)	Chronos (Avg)	ARIMA (Avg)
Ahmedabad_Dengue	1.94060	2.14977	2.98450
Bangkok_Dengue	3.13973	1.74447	7.31587
Columbia_Dengue	7.83080	6.67217	6.71217
Hong_Kong_Dengue	0.93920	0.95553	0.90003
Iquitos_Dengue	2.50020	1.96197	2.04223
Philippines_Dengue	1.01543	0.61387	0.75357
Sanjuan_Dengue	9.70020	10.67870	7.26073
Singapore_Dengue	4.99743	4.96997	5.65803
Venezuela_Dengue	5.69360	4.01090	5.31507
China_Hepatitis_B	0.81690	0.88130	0.89020
Australia_Influenza	5.38260	3.25907	5.48367
Japan_Influenza	1.09017	0.69210	1.78370
Mexico_Influenza	0.43457	0.49617	2.24700
Colombia_Malaria	3.66647	2.33897	2.82943
Venezuela_Malaria	1.18080	3.07593	2.02100

Table 4: MASE Average Values of Short Term, Medium Term, and Long Term Predictions across Datasets

Datasets	**Range**	**GTT**	**Chronos**	**ARIMA**
Ahmedabad_Dengue	Short Term	49.9611	79.3344	222.9468
Ahmedabad_Dengue	Medium Term	213.7265	248.0489	272.7149
Ahmedabad_Dengue	Long Term	104.1137	103.6036	127.6932
Bangkok_Dengue	Short Term	546469.3153	261267.6635	255629.9703
Bangkok_Dengue	Medium Term	123656.7008	226621.6852	192769.0608
Bangkok_Dengue	Long Term	1552351.235	337295.7122	25039824.78
Columbia_Dengue	Short Term	34576.6152	69252.4099	30045.2777
Columbia_Dengue	Medium Term	903659.5308	615275.3524	717175.1625
Columbia_Dengue	Long Term	1007339.712	743265.383	1013047.452
Hong_Kong_Dengue	Short Term	5.0303	8.2889	8.436
Hong_Kong_Dengue	Medium Term	14.5035	10.6096	15.1997
Hong_Kong_Dengue	Long Term	19.9706	38.8689	19.7023
Iquitos_Dengue	Short Term	153.8279	27.1052	84.244
Iquitos_Dengue	Medium Term	226.5892	270.2082	171.7012
Iquitos_Dengue	Long Term	174.5096	139.2326	102.5357
Philippines_Dengue	Short Term	670.0413	786.2015	483.9639
Philippines_Dengue	Medium Term	1602.4613	922.6166	1768.4554
Philippines_Dengue	Long Term	2515.7464	2936.2802	2868.3674
Sanjuan_Dengue	Short Term	7462.0006	13676.6681	952.1177
Sanjuan_Dengue	Medium Term	8476.4303	9409.2027	9476.8718
Sanjuan_Dengue	Long Term	11204.9309	11413.6093	8718.9706
Singapore_Dengue	Short Term	46268.1056	48427.2396	58239.5684
Singapore_Dengue	Medium Term	35179.4115	34621.0287	36844.9782
Singapore_Dengue	Long Term	19709.8469	11563.2852	25924.5703
Venezuela_Dengue	Short Term	629291.0722	479475.1294	557141.3027
Venezuela_Dengue	Medium Term	242586.9309	148171.6894	234704.0395
Venezuela_Dengue	Long Term	520841.2489	386977.3017	412491.9455
China_Hepatitis_B	Short Term	36264831.43	39770696.5	24107439.35
China_Hepatitis_B	Medium Term	95853695.84	123859614.3	136064376.8
China_Hepatitis_B	Long Term	59148236.27	53011404.96	89955023.53
Australia_Influenza	Short Term	192.9541	1434.0061	435.9705
Australia_Influenza	Medium Term	8640.1844	1548.5851	9841.2444
Australia_Influenza	Long Term	6621.0921	1568.235	5584.8924
Japan_Influenza	Short Term	11515.5003	4553.492	7938.3993
Japan_Influenza	Medium Term	8653.3359	652.7386	41.649
Japan_Influenza	Long Term	5282.5359	8969.1334	37301.38
Mexico_Influenza	Short Term	26.3165	243.1802	4723.1524
Mexico_Influenza	Medium Term	249.9592	429.3543	10556.6482
Mexico_Influenza	Long Term	2312.6887	1314.8496	12754.4334
Colombia_Malaria	Short Term	66999.9089	85176.543	92912.1633
Colombia_Malaria	Medium Term	252286.8395	58833.5159	270917.3322
Colombia_Malaria	Long Term	664329.3017	616310.5485	601389.2299
Venezuela_Malaria	Short Term	23921.4229	39633.5184	34978.9183
Venezuela_Malaria	Medium Term	43205.3772	238473.1437	174636.977
Venezuela_Malaria	Long Term	36146.324	358456.2389	62607.0748

Table 5: Performance for Different Datasets and Prediction Ranges on MSE

Datasets	Range	GTT	Chronos	ARIMA
Ahmedabad_Dengue	Short Term	0.372	0.4687	0.3732
Ahmedabad_Dengue	Medium Term	0.6861	0.7391	0.3514
Ahmedabad_Dengue	Long Term	0.8698	0.8676	0.2306
Bangkok_Dengue	Short Term	0.6173	0.4268	0.8426
Bangkok_Dengue	Medium Term	0.4578	0.6197	0.3532
Bangkok_Dengue	Long Term	1.7922	0.8354	2.7986
Columbia_Dengue	Short Term	0.1919	0.2716	0.2575
Columbia_Dengue	Medium Term	0.7578	0.6253	0.4952
Columbia_Dengue	Long Term	0.5027	0.4318	0.3461
Hong_Kong_Dengue	Short Term	0.2588	0.3322	0.363
Hong_Kong_Dengue	Medium Term	0.448	0.3832	0.2998
Hong_Kong_Dengue	Long Term	0.4745	0.662	0.2611
Iquitos_Dengue	Short Term	2.1262	0.8925	0.706
Iquitos_Dengue	Medium Term	0.9933	1.0847	0.3743
Iquitos_Dengue	Long Term	1.3711	1.2247	0.2893
Philippines_Dengue	Short Term	0.5056	0.5476	0.3267
Philippines_Dengue	Medium Term	0.7517	0.5703	0.299
Philippines_Dengue	Long Term	0.7505	0.8108	0.2469
Sanjuan_Dengue	Short Term	1.1739	1.5893	0.241
Sanjuan_Dengue	Medium Term	0.6589	0.6941	0.4613
Sanjuan_Dengue	Long Term	1.0419	1.0515	0.4095
Singapore_Dengue	Short Term	0.5201	0.532	0.5509
Singapore_Dengue	Medium Term	0.5846	0.5799	0.4382
Singapore_Dengue	Long Term	0.5706	0.4371	0.2954
Venezuela_Dengue	Short Term	0.5602	0.489	0.5985
Venezuela_Dengue	Medium Term	0.4714	0.3684	0.2902
Venezuela_Dengue	Long Term	0.5145	0.4435	0.3661
China_Hepatitis_B	Short Term	0.0612	0.064	0.3301
China_Hepatitis_B	Medium Term	0.1006	0.1143	0.4848
China_Hepatitis_B	Long Term	0.0816	0.0772	0.3326
Australia_Influenza	Short Term	0.2187	0.5963	0.1221
Australia_Influenza	Medium Term	0.8591	0.3637	0.4221
Australia_Influenza	Long Term	1.2127	0.5902	0.3075
Japan_Influenza	Short Term	11.8139	7.4289	3.0723
Japan_Influenza	Medium Term	12.8649	3.5333	0.2225
Japan_Influenza	Long Term	0.6569	0.856	0.3166
Mexico_Influenza	Short Term	0.5005	1.5213	4.5816
Mexico_Influenza	Medium Term	1.7418	2.2828	6.8497
Mexico_Influenza	Long Term	1.1757	0.8864	0.5912
Colombia_Malaria	Short Term	0.2663	0.3002	0.3113
Colombia_Malaria	Medium Term	0.4226	0.204	0.3294
Colombia_Malaria	Long Term	0.4838	0.466	0.2984
Venezuela_Malaria	Short Term	0.0953	0.1226	0.4832
Venezuela_Malaria	Medium Term	0.1292	0.3034	0.7475
Venezuela_Malaria	Long Term	0.1305	0.4108	0.2814

Table 6: Performance on NRMSE for Different Datasets and Prediction Ranges

Datasets	**Range**	**GTT**	**Chronos**	**ARIMA**
Ahmedabad_Dengue	Short Term	39.0617	42.4409	72.075
Ahmedabad_Dengue	Medium Term	72.5782	68.2738	74.1195
Ahmedabad_Dengue	Long Term	72.4157	68.6072	109.2993
Bangkok_Dengue	Short Term	83.0279	51.8012	46.3884
Bangkok_Dengue	Medium Term	48.9045	54.846	59.2161
Bangkok_Dengue	Long Term	95.6353	70.18	158.5215
Columbia_Dengue	Short Term	13.2665	29.4646	11.5341
Columbia_Dengue	Medium Term	55.4624	47.4629	50.2788
Columbia_Dengue	Long Term	40.6849	38.0468	41.1221
Hong_Kong_Dengue	Short Term	23.2743	25.5224	28.0633
Hong_Kong_Dengue	Medium Term	35.3415	26.1447	38.1541
Hong_Kong_Dengue	Long Term	36.0577	64.938	35.8968
Iquitos_Dengue	Short Term	116.8591	111.6716	103.729
Iquitos_Dengue	Medium Term	119.0524	154.0943	91.4318
Iquitos_Dengue	Long Term	144.3213	116.0207	105.2395
Philippines_Dengue	Short Term	28.7104	46.8799	29.6495
Philippines_Dengue	Medium Term	62.3573	41.5399	65.163
Philippines_Dengue	Long Term	59.4524	59.0478	64.4609
Sanjuan_Dengue	Short Term	78.9867	88.3057	38.6054
Sanjuan_Dengue	Medium Term	70.9519	69.6806	79.5502
Sanjuan_Dengue	Long Term	116.8561	100.7993	77.4886
Singapore_Dengue	Short Term	41.6264	43.3445	49.1369
Singapore_Dengue	Medium Term	35.6501	39.8697	37.9576
Singapore_Dengue	Long Term	37.8899	32.3515	49.1871
Venezuela_Dengue	Short Term	62.5424	52.286	56.556
Venezuela_Dengue	Medium Term	40.5451	26.6095	39.5521
Venezuela_Dengue	Long Term	46.1201	39.2173	40.4799
China_Hepatitis_B	Short Term	5.687	5.651	3.7298
China_Hepatitis_B	Medium Term	8.9179	10.5759	10.4198
China_Hepatitis_B	Long Term	6.1712	5.5649	7.8229
Australia_Influenza	Short Term	25.6069	40.8236	33.2323
Australia_Influenza	Medium Term	80.0547	35.157	90.6562
Australia_Influenza	Long Term	116.2063	61.063	70.0897
Japan_Influenza	Short Term	132.8153	113.9108	164.6492
Japan_Influenza	Medium Term	128.3771	60.8674	67.7042
Japan_Influenza	Long Term	88.2176	55.0844	135.5356
Mexico_Influenza	Short Term	49.0266	70.541	138.0216
Mexico_Influenza	Medium Term	83.1781	84.4798	163.1871
Mexico_Influenza	Long Term	91.8643	88.1975	134.8698
Colombia_Malaria	Short Term	24.3532	30.0795	23.9966
Colombia_Malaria	Medium Term	34.6083	16.8536	35.3094
Colombia_Malaria	Long Term	44.837	40.3626	40.5229
Venezuela_Malaria	Short Term	6.8467	11.1651	10.3734
Venezuela_Malaria	Medium Term	10.9895	32.171	26.5742
Venezuela_Malaria	Long Term	11.2219	42.8038	14.9543

Table 7: Performance on sMAPE for Different Datasets and Prediction Ranges

Datasets	Range	GTT	Chronos	ARIMA
Ahmedabad_Dengue	Short Term (Weekly)	1.3479	1.6204	3.3477
Ahmedabad_Dengue	Medium Term	2.8368	3.1959	3.4316
Ahmedabad_Dengue	Long Term	1.6371	1.6310	2.1742
Bangkok_Dengue	Short Term (Monthly)	2.6328	1.8479	1.7421
Bangkok_Dengue	Medium Term	1.1458	1.4177	1.4813
Bangkok_Dengue	Long Term	5.6406	1.9678	18.7242
Columbia_Dengue	Short Term (Weekly)	1.3202	2.8413	1.1234
Columbia_Dengue	Medium Term	10.2263	8.2198	8.9231
Columbia_Dengue	Long Term	9.9459	8.9554	10.0900
Hong_Kong_Dengue	Short Term (Monthly)	0.5897	0.6894	0.7362
Hong_Kong_Dengue	Medium Term	0.8737	0.6873	0.9563
Hong_Kong_Dengue	Long Term	1.3542	1.4899	1.0076
Iquitos_Dengue	Short Term	2.7141	1.0056	1.8842
Iquitos_Dengue	Medium Term	2.7717	3.1149	2.4476
Iquitos_Dengue	Long Term	2.0148	1.7654	1.7949
Philippines_Dengue	Short Term	0.3608	0.4879	0.3823
Philippines_Dengue	Medium Term	0.7726	0.5178	0.8298
Philippines_Dengue	Long Term	1.9129	0.8359	1.0486
Sanjuan_Dengue	Short Term	9.7748	12.6675	3.4368
Sanjuan_Dengue	Medium Term	9.2277	9.7126	10.0301
Sanjuan_Dengue	Long Term	10.0981	9.6560	8.3153
Singapore_Dengue	Short Term	6.5031	6.6971	7.3558
Singapore_Dengue	Medium Term	4.6416	4.9871	4.8633
Singapore_Dengue	Long Term	3.8476	3.2257	4.7550
Venezuela_Dengue	Short Term	6.9101	6.0295	6.4277
Venezuela_Dengue	Medium Term	4.2874	2.8156	4.1646
Venezuela_Dengue	Long Term	5.8833	5.1876	5.3529
China_Hepatitis_B	Short Term	0.7170	0.7119	0.4709
China_Hepatitis_B	Medium Term	1.0970	1.2798	1.2799
China_Hepatitis_B	Long Term	0.6367	0.6522	0.9198
Australia_Influenza	Short Term	1.2751	3.0448	1.7711
Australia_Influenza	Medium Term	8.2295	3.2642	9.1346
Australia_Influenza	Long Term	6.6432	3.4682	5.5453
Japan_Influenza	Short Term	1.3743	0.8516	1.6661
Japan_Influenza	Medium Term	0.9913	0.2269	0.0973
Japan_Influenza	Long Term	0.9049	0.9978	3.5877
Mexico_Influenza	Short Term	0.1166	0.3114	1.5686
Mexico_Influenza	Medium Term	0.3424	0.4197	2.4803
Mexico_Influenza	Long Term	0.8447	0.7574	2.6920
Colombia_Malaria	Short Term	1.3469	1.6643	1.3635
Colombia_Malaria	Medium Term	2.8422	1.1484	2.8980
Colombia_Malaria	Long Term	4.8103	4.2043	4.2268
Venezuela_Malaria	Short Term	0.8686	1.3838	1.2948
Venezuela_Malaria	Medium Term	1.3834	3.5803	3.0477
Venezuela_Malaria	Long Term	1.2904	4.2637	1.7205

Table 8: Performance on MASE for Different Datasets and Prediction Ranges