EPICASTING

Tanujit @Sorbonne

EPICASTING

Tanujit @Sorbonne

EPICASTING

Tanujit @Sorbonne


EPICASTING
Tanujit @Sorbonne

About Me

2014

- ✓ Indian Statistical Institute (MS)
- ✓ Industrial Internship

STATISTICS
+
DATA SCIENCE
+
MACHINE LEARNING

2016

- ✓ Indian Statistical Institute (PhD)
- ✓ Theoretical & Applied ML
- ✓ Three best paper awards

2020

- ✓ Defended PhD Thesis
- ✓ Consultant at Bajaj Finserv
- ✓ Postdoc Fellow at IIITD

2021

- ✓ Mphasis Fellow at IIITB
- ✓ A/Prof at Sorbonne Uni.
- ✓ Adjunct Prof at Woxsen Uni.

2023

- ✓ Consultant at PharmaACE
- ✓ Research Advisor at SCAI
- ✓ Member of Franco-Indo Digital Health Campus
- ✓ Visiting Duke-NUS

FORECASTING
+
EPIDEMICS
+
GEOPHYSICS

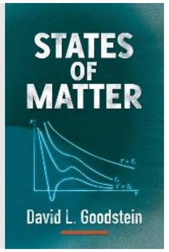
Outline

- Past of Forecasting
- Basics of Forecasting
- Epidemic Forecasting using EWNet
- Appendix

Outline

- Past of Forecasting

“Ludwig Boltzmann, who spent much of his life studying statistical mechanics, died in 1906, by his own hand. Paul Ehrenfest, carrying on the work, died similarly in 1933. Now it is our turn to study statistical mechanics.”



- Basics of

- Epidemic Forecasting using EWNet

- Appendix

Past of Forecasting

Forecasting by maggots: Clay model of sheep's liver, stored in British Museum.

- Forecasting has fascinated people for thousands of years, sometimes being considered a sign of divine inspiration, and sometimes being seen as a criminal activity.

- The Jewish prophet Isaiah wrote in about 700 BC "Tell us what the future holds, so we may know that you are gods." (Isaiah 41:23).

- One hundred years later, in ancient Babylon, forecasters would foretell the future based on the distribution of maggots in a rotten sheep's liver.

Past of Forecasting

Forecasting by hallucination.

- Beginning in 800 BC, a priestess known as the Pythia would answer questions about the future at the Temple of Apollo on Greece's Mount Parnassus.

- It is said that she, the Oracle of Delphi, dispensed her wisdom in a trance - caused, some believe, by the hallucinogenic gases that would seep up through natural vents in the rock.

Forecasters are to blame!

- Forecasters had a tougher time under the emperor Constantius, who issued a decree in AD357 forbidding anyone "to consult a soothsayer, a mathematician, or a forecaster -- May curiosity to foretell the future be silenced forever."

- News report on 16 August 2006: A Russian woman is suing weather forecasters for wrecking her holiday. A court in Uljanovsk heard that Alyona Gabitova had been promised 28 degrees and sunshine when she planned a camping trip to a local nature reserve, newspaper Nowyje Iswestija said.

- But it did nothing but pour with rain the whole time, leaving her with a cold. Gabitova has asked the court to order the weather service to pay the cost of her travel.

Reputations can be made and lost

Some Misconceptions (Low Expectations): Our forecasts will always be inaccurate, so we should concentrate our efforts elsewhere.

"I think there is a world market for maybe five computers. (Chairman of IBM, 1943)

"There is no reason anyone would want a computer in their home." (President, DEC, 1977)

"There's no chance that the iPhone is going to get any significant market share. No chance."
(Steve Ballmer, CEO Microsoft, April 2007)

"We're going to be opening relatively soon …The virus … will go away in April."
(Donald Trump, February 2020)

"Prediction is very difficult, especially if it's about the future!" - Niels Bohr

Reputations can be made and lost

Some Misconceptions (High Expectations): If only we had the latest forecasting technology, then all our problems could be solved.

Don't be fooled: Covid won't be cured by a panacea

Philip Ball

The Guardian

'Cure-alls' such as vitamin D and ivermectin seem appealing. But the truth is, specific diseases demand specific medicines

ELSEVIER

International Journal of Forecasting

Volume 38, Issue 2, April–June 2022, Pages 423-438

Forecasting for COVID-19 has failed

John P.A. Ioannidis [a], Sally Cripps [b], Martin A. Tanner [c]

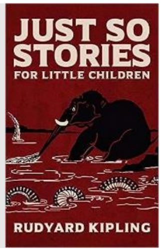
- Poor data input
- Wrong modeling assumptions
- Lack of incorporation of epidemiological features
- Poor past evidence on effects of available interventions

- Lack of transparency
- Consideration of only one or a few dimensions of the problem at hand
- Lack of expertise in crucial disciplines
- Groupthink and bandwagon effects

Outline

- Past of Forecasting
- Basics of Forecasting
- Epidemic Foreca
- Appendix

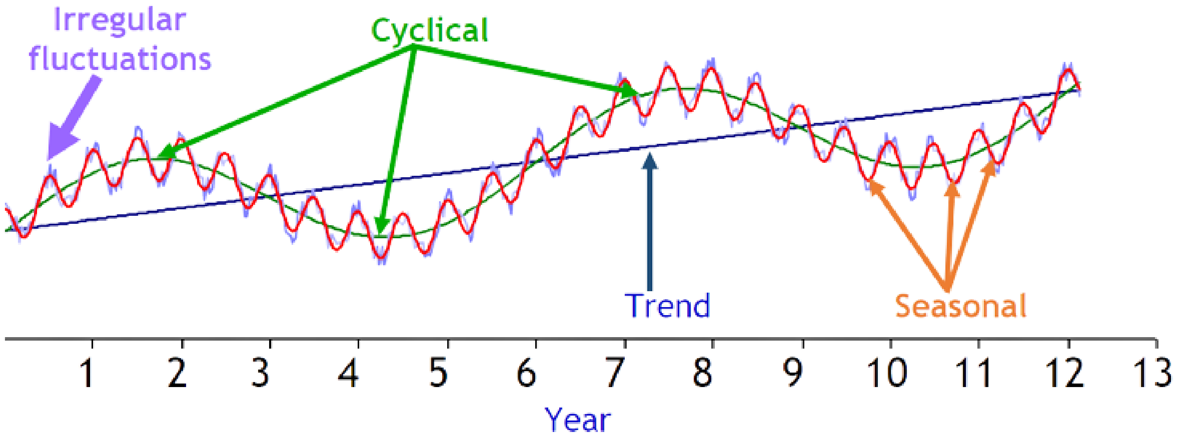
"I keep six honest serving-men
(They taught me all I knew);
Their names are What and Why and When
And How and Where and Who."



What is time series?

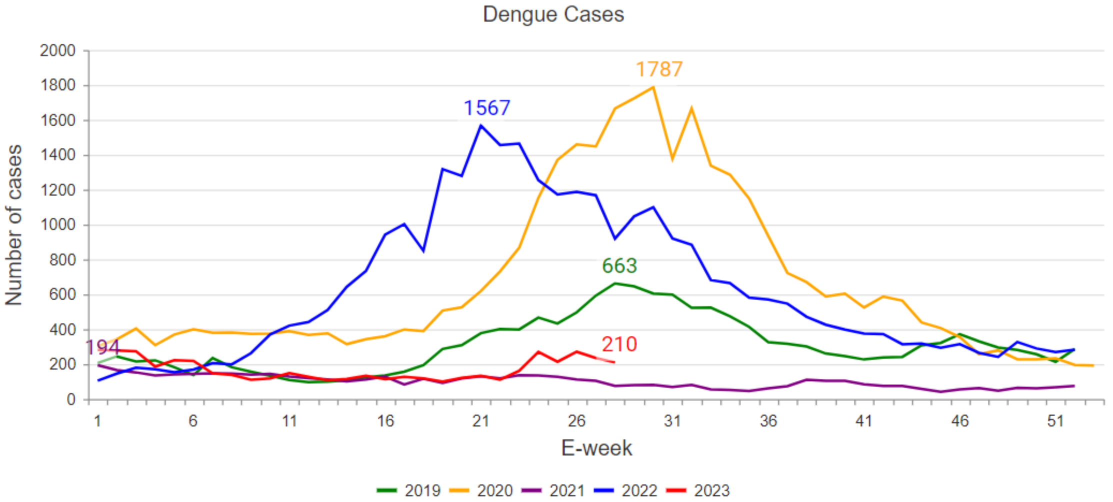
Time series is a set of observations, each one being recorded at a specific time.

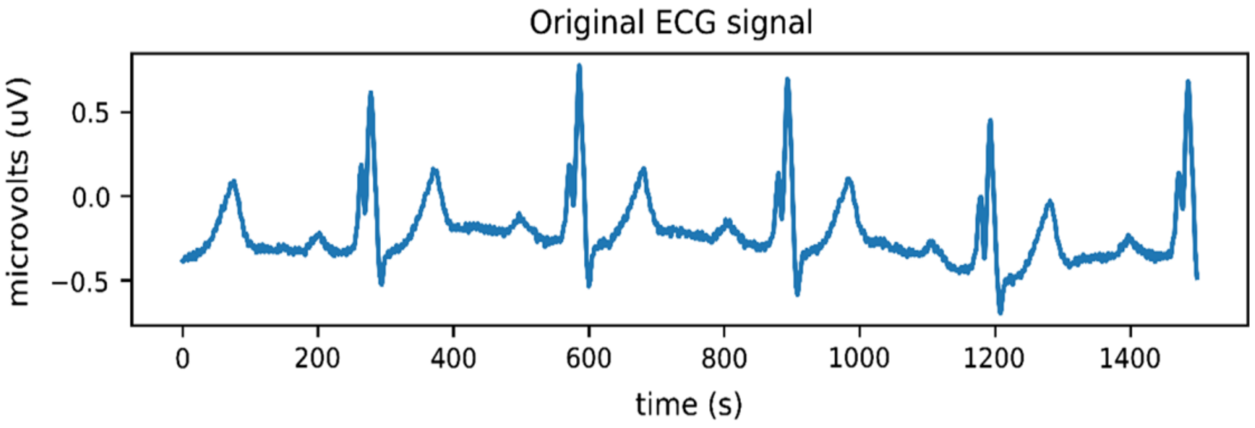
Stationary time series is roughly horizontal, constant variance and no patterns predictable in the long-term.



Discrete time series is one in which the set of time points at which observations are made is a discrete set (e.g., weekly dengue incidence cases)

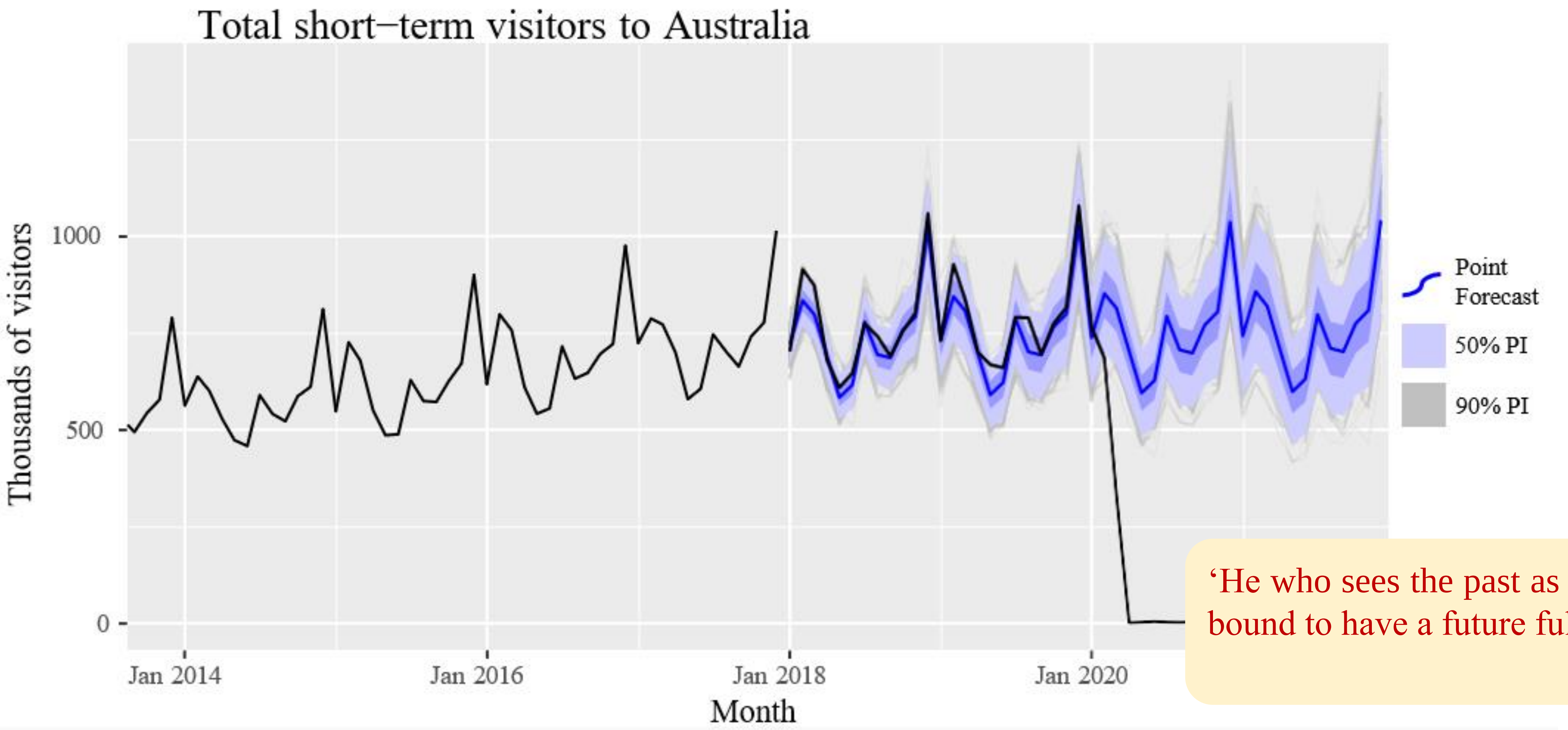
Continuous time series are obtained when observations are made continuously over some time intervals (e.g., ECG graph)





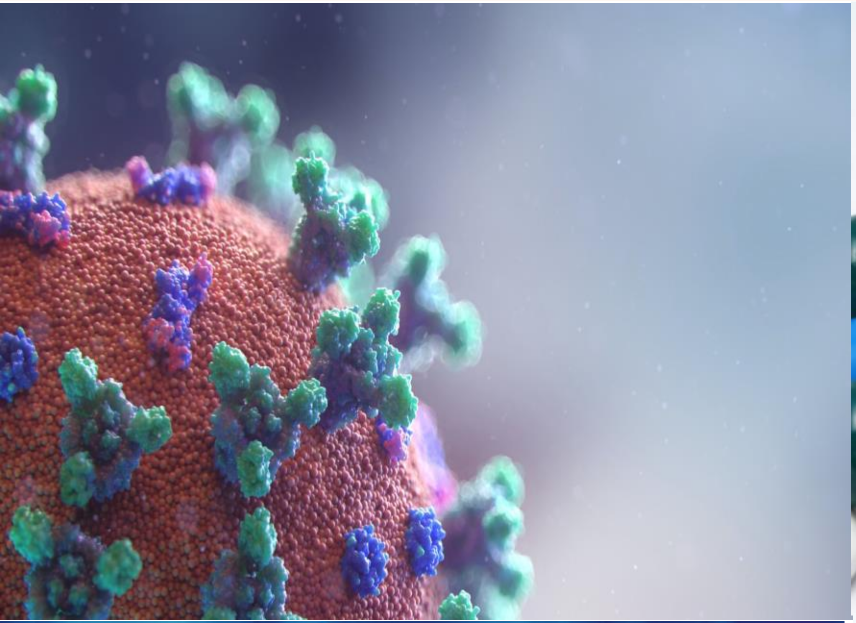
What is a forecast?

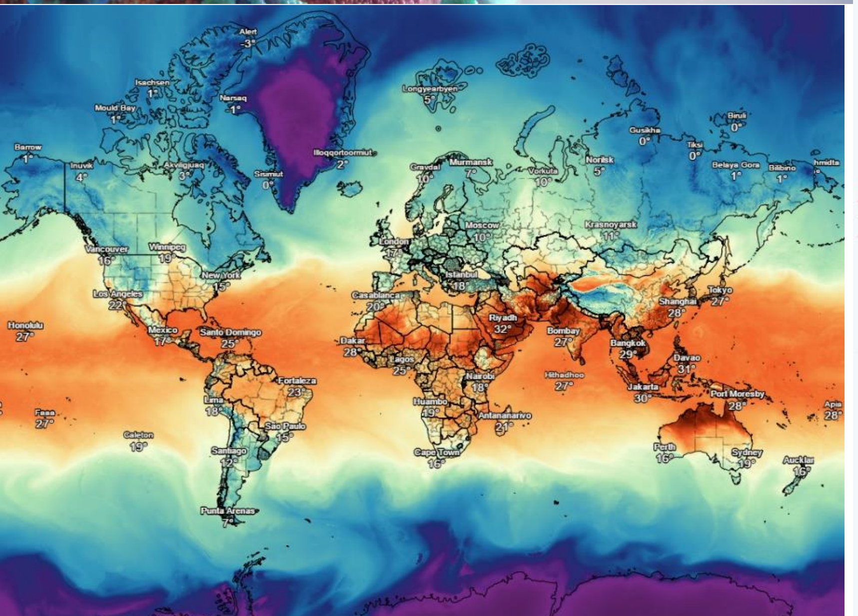
Forecasting is estimating how the sequence of observations will continue into the future.

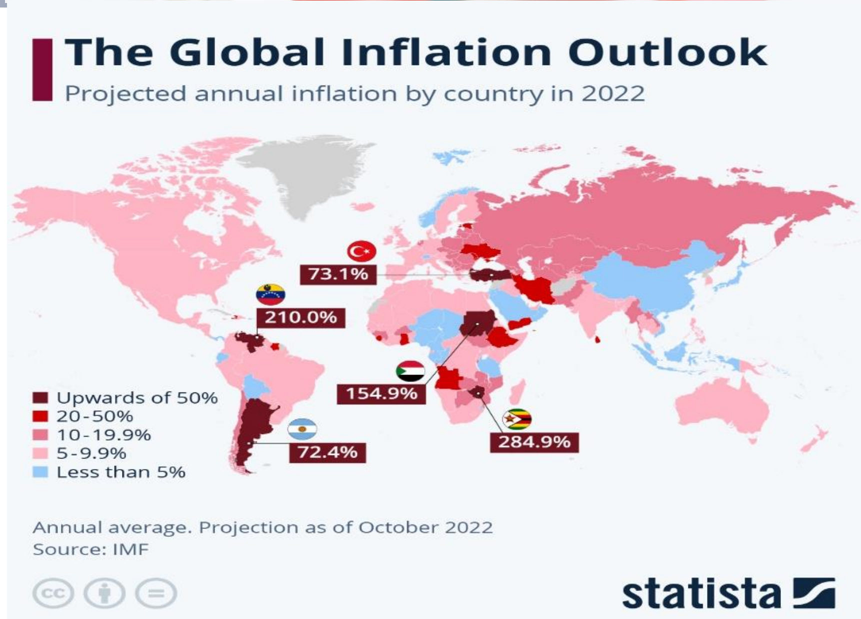


'He who sees the past as surprise-free is bound to have a future full of surprise.'
- Amos Tversky

What can we forecast?







Forecasting approaches

Types of forecasting approaches

- Judgemental
  - Role playing
  - Expert opinion (Delphi method)
- Scientific
  - Statistical Methods (ES, ARIMA, GP)
  - ML Methods (LSTM, ARNN)

Consideration:

- How far-ahead? (the forecast ‘horizon’)
- What level of aggregation? (Types/ Which disease?)
- Number and frequency of forecasts required.
- Availability of historical data.
- Relative accuracy of options.

PLOS COMPUTATIONAL BIOLOGY

OPEN ACCESS PEER-REVIEWED

RESEARCH ARTICLE

A human judgment approach to epidemiological forecasting

David C. Farrow, Logan C. Brooks, Sangwon Hyun, Ryan J. Tibshirani, Donald S. Burke, Roni Rosenfeld

Published: March 10, 2017 • https://doi.org/10.1371/journal.pcbi.1005248

Science

Forecasting the Global AIDS Epidemic: Good computer models might help persuade officials of developing countries to institute anti-AIDS strategies, but modeling has proven easier said than done

ELIZABETH CULOTTA Authors Info & Affiliations

SCIENCE • 23 Aug 1991 • Vol 253, Issue 5022 • pp. 852-854 • DOI: 10.1126/science.1876845

BMC Infectious Diseases

Chimeric forecasting: combining probabilistic predictions from computational models and human judgment

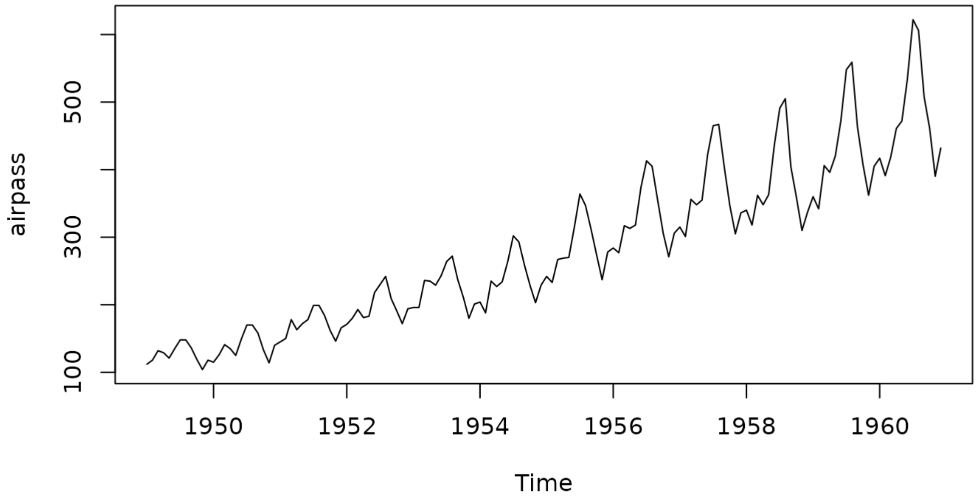
Thomas McAndrew, Allison Codi, Juan Cambeiro, Tamay Besiroglu, David Braun, Eva Chen, Luis Enrique Urtubey De Césaris & Damon Luk

BMC Infectious Diseases 22, Article number: 833 (2022) | Cite this article

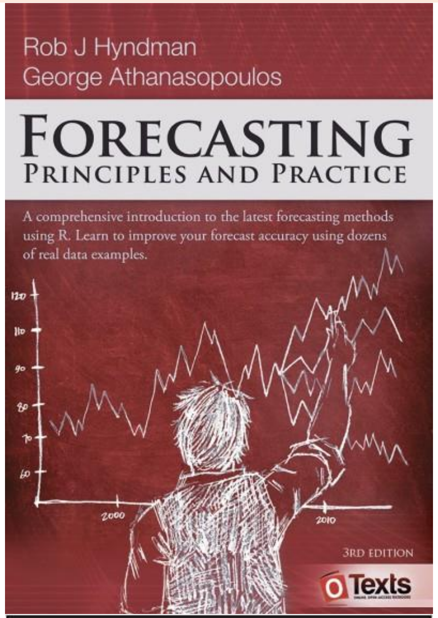
Forecastability factors for scientific methods

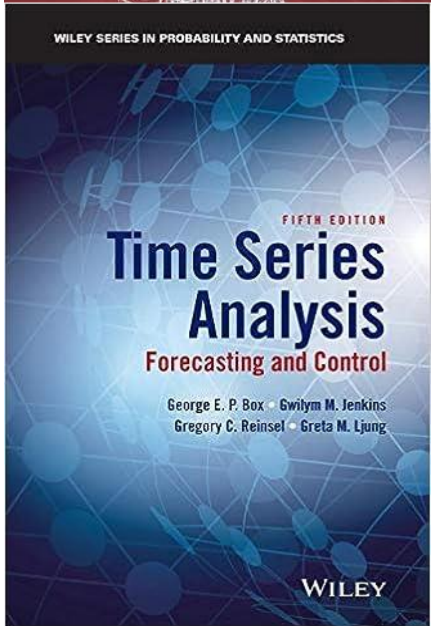
Something is easier to forecast if:

- we have a good understanding of the factors that contribute to it, and can measure them.
- there is lots of data available;
- the future is somewhat similar to the past
- the forecasts cannot affect the thing we are trying to forecast.
- Structural break over data history (pandemic)



- Clear trend
- clear seasonal patterns
- good length of data history
- short forecasting horizon.





Process of forecasting

nature communications

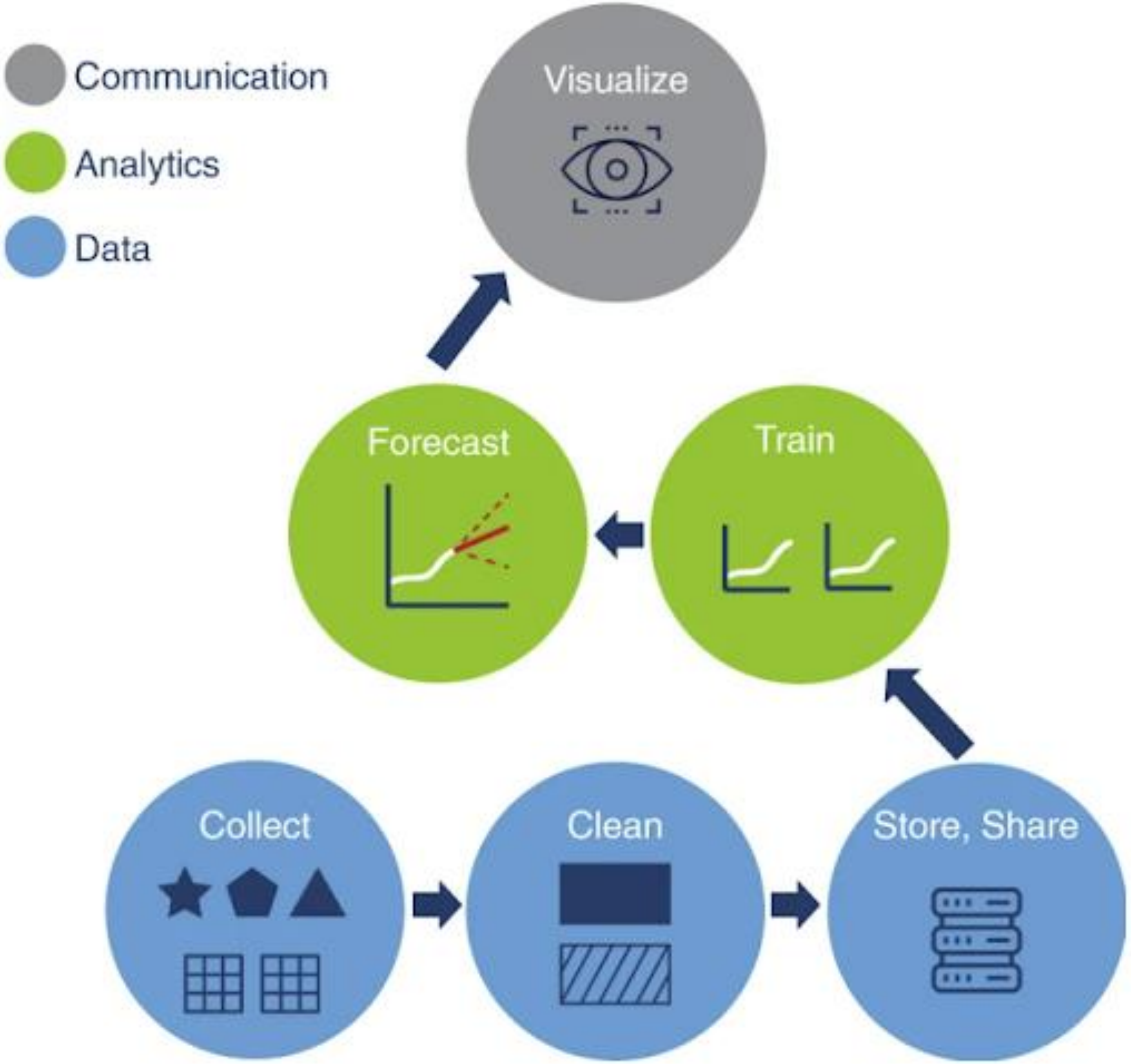
Technology to advance infectious disease forecasting for outbreak management

Dylan B. George, Wendy Taylor, Jeffrey Shaman, Caitlin Rivers, Brooke Paul, Tara O'Toole, Michael A. Johansson, Lynette Hirschman, Matthew Biggerstaff, Jason Asher & Nicholas G. Reich

Nature Communications **10**, Article number: 3932 (2019) | Cite this article

Forecasting is beginning to be integrated into decision-making processes for infectious disease outbreak response. We discuss how technologies could accelerate the adoption of forecasting among public health practitioners, improve epidemic management, save lives, and reduce the economic impact of outbreaks.

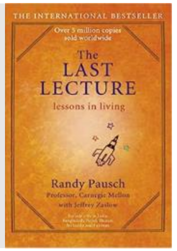
'Data gaps undermine our ability to target resources, develop policies and track accountability. Without good data, we're flying blind. If you can't see it, you can't solve it.' - Kofi Annan (Nature, 2018)



Outline

- Past of Forecasting
- Basics of Forecasting
- Epidemic Forecasting using EWNet
- Appendix

“When there's an elephant in the room introduce him.”



Epidemics: A serious problem

- Epidemic: a (usually rapid) rise in a prevalence of a disease or condition.

- Pandemic: ‘an epidemic that spreads globally’

- In practice: an epidemic that is driven by a new pathogen
  - Very little is known about it
  - Population has no pre-existing immunity against it

- Singular Events:
  - Atom Bomb on Hiroshima: ~ 100,000 deaths.
  - 1918 ‘Spanish Flu’ pandemic: 50,000,000 deaths.

- Ongoing Events:
  - Syrian civil war, 2011-present 100,000 deaths/year
  - Seasonal (non-pandemic) Flu: 250,000 deaths/year
  - Dengue: 50,000,000 - 500,000,000 cases/year
  - Dengue: 25,000 death/year, growing

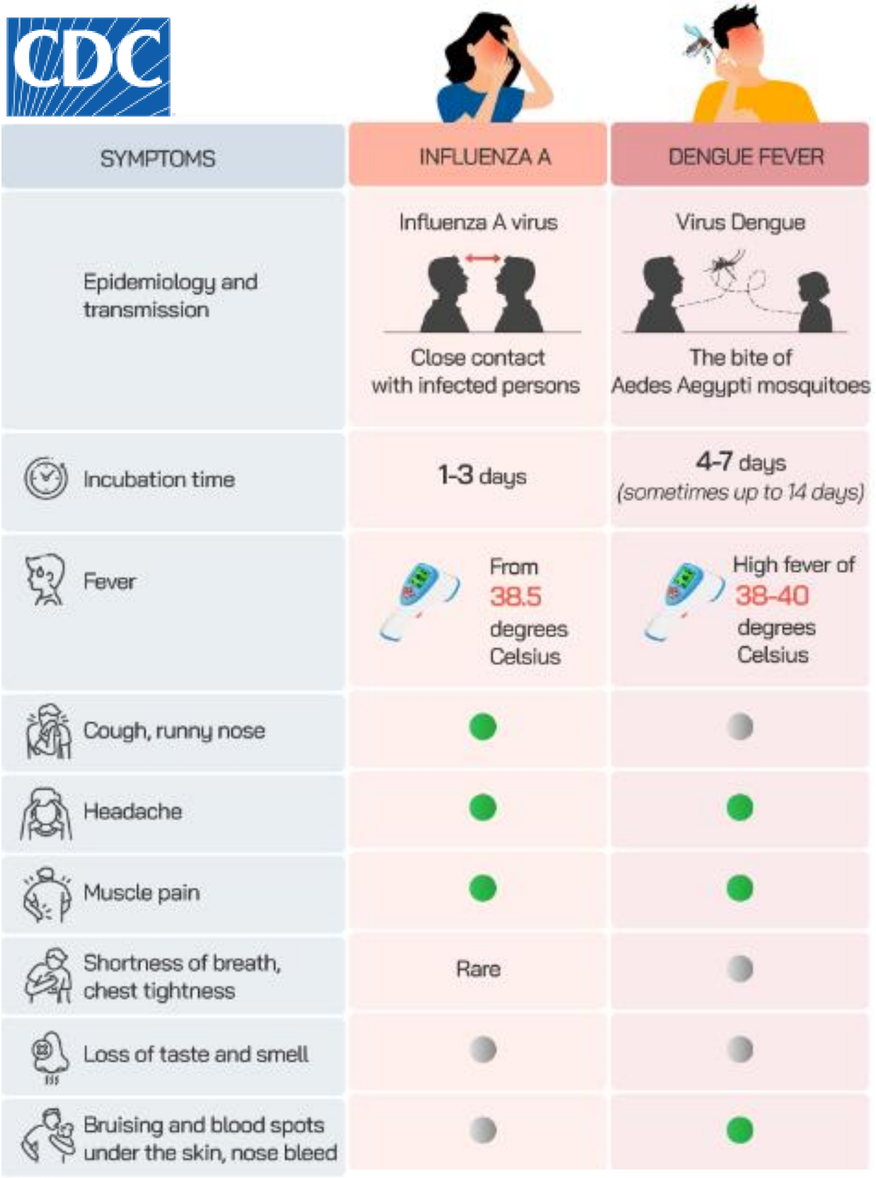
Epidemic

Pandemic

Epidemics are regular

- Flu:
  - In tropics: year round, erratic
  - In subtropics: semi-regular, 1-2 epidemics a year
  - In temperate zones: every winter, but otherwise irregular in both timing and intensity.
  - Fast moving, but not simultaneous

- Dengue:
  - Tends to follow the wet season (mosquitos)
  - In South America: Usually January-April
  - In S.E. Asia: two seasons/year, large & small
  - Intensity is highly variable (10x or more)
  - Highly local outbreaks, but global diffusion



Source: Ministry of Health, U.S. CDC

Benefits of epicasting

- To Governments:
  - Timing and focus of communications (e.g., vaccination campaigns)
  - Antiviral policy
  - Mosquito control, door-to-door campaigns

- To Health Care Providers:
  - Staffing, vacations
  - Elective surgery
  - Equipment pre-positioning

- To Individuals:
  - Protect our families (old and comorbidity)

Vision:

1. Building an epidemic forecasting model to handle the data irregularities
2. State-of-the-art performance
3. Understanding the theoretical and computational aspect
4. Building software for public use (similar to weather forecasts)
5. Making an impact in healthcare

Epicasting approaches

EPICASTING METHODS

Data-Centric Epidemic Forecasting: A Survey

Alexander Rodríguez*, Harshavardhan Kamarthi*, Pulak Agarwal, Javen Ho, Mira Patel, Suchet Sapre, and B. Aditya Prakash†

College of Computing, Georgia Institute of Technology, USA

Mechanistic Models

Less data · Describes epidemic states · Surveillance DGP

Compartmental Models (SIR, SEIR)

Agent-based Models (Stochastic ABM)

Phenomenological Models

Data centric · Describes surveillance data · Stochastic

Classical Time Series Models (ARIMA, ETS, SETAR)

Regression & ML Models (DT, LR, RF, DL)

THE LANCET
Nowcasting and forecasting the potential domestic and internatior spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study
Prof Joseph T Wu, PhD • Kathy Leung, PhD • Prof Gabri
Published: January 31, 2020 • DOI: https://doi.org/10.1016/S0140

nature medicine
A stochastic agent-based model of the SA epidemic in France

The Thirty-Seventh AAAI Conference on Artificial I
EINNs: Epidemiologically-Informed Neu
Alexander Rodríguez[1], Jiaming Cui[1], Naren Ramakrishnai B. Aditya Prakash[1]

Journal of the Royal Statistical Society
Series A: Statistics in Society
An Ensemble Method for Early Predicti Outbreak

PLOS COMPUTATIONAL BIOLOGY
A hybrid stochastic-deterministic approa explore multiple infection and evolution HIV
Jesse Kreger[1,2]*, Natalia L. Komarova[2], Dominik Wodarz[2,3]

PNAS
Evaluation of individual and ensemble probabilistic forecasts of COVID-19 morta in the United States
Estee Y. Cramer, Evan L. Ray, Velma K. Lopez, and Nicholas G. Reich
April 8, 2022 | 119 (15) e2113561119 | https://doi.org/10.1073/pnas.2113561119

nature communications
Forecasting influenza activity using machine-learned mobility map
Srinivasan Venkatramanan, Adam Sadilek, Arindam Fadikar, Christopher L. Barrett, Matthew Biggerstaff, Jiangzhuo Chen, Xerxes Dotiwalla, Paul Eastham, Bryant Gipson, Dave Higdon, Onur Kucuktunc, Allison Lieber, Bryan L ... ing Wang & Madhav Marathe

scientific reports
D-19 activity in Australia to support se: May to October 2020

PHYSICA A
Forecasting dengue epidemics using a hybrid methodology
Tanujit Chakraborty, Swarup Chattopadhyay, Indrajit Ghosh

Common approaches to modeling epidemics

- Mechanistic: Compartmental models, e.g., S-I-R (understanding epidemics)
  - Oversimplified assumptions (e.g., perfect mixing – every person is interacting equally often with every other persons)
  - Hard to estimate parameters in real-time

- Mechanistic: Agent-Based Models (individual-level simulation)
  - Many parameters, hard to fit/validate

- Non-Mechanistic: Statistical/machine learning e.g., SARIMA, ARNN
  - Extrapolate trend, seasonal, and auto-correlation effect into the near future (relies on explanatory variables)
  - Assumptions don't always hold (flu is annual but not periodic!)
  - Needs historical data, less suitable for novel (e.g., pandemic) situations

- Non-Mechanistic: Deep learning e.g., LSTM
  - Often computationally expensive, low-test accuracy
  - Less interpretable (e.g., estimating the effect of explanatory variable)

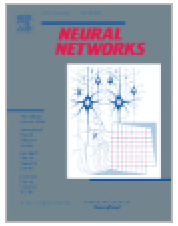
- More recently: Data assimilation methods from weather forecasting (e.g., Kalman Filters, Particle Filtering)
  Time frequency domain tools from signal and image processing (e.g., Wavelet and Fourier decomposition)

- Finally – assess forecast accuracy on test set and assess the impact the accuracy is having.

Epicasting using EWNet

ELSEVIER

Neural Networks
Volume 165, August 2023, Pages 185-212



Epicasting: An Ensemble Wavelet Neural Network for forecasting epidemics

Madhurima Panja [a 1], Tanujit Chakraborty [b a c 1], Uttam Kumar [a], Nan Liu [d]

R epicasting

iiit-b ज्ञानमुत्तमम

PARIS SORBONNE UNIVERSITÉ

DukeNUS Medical School

Types of epicasting, targets & metrics

- Across seasons
- Within season
- NearCasting
- NowCasting
- BackCasting

- Targets:
  What? (flu or dengue)
  How Bad? (season's peak intensity)
  How Long? (epidemic duration)
  Nearcasting: expected cases in next few months

- Metrics:
  Point predictions ("what is the most likely outcome?"):
  Error Metrics (RMSE, MAE, MASE)
  Distributional predictions ("how likely is each outcome?")

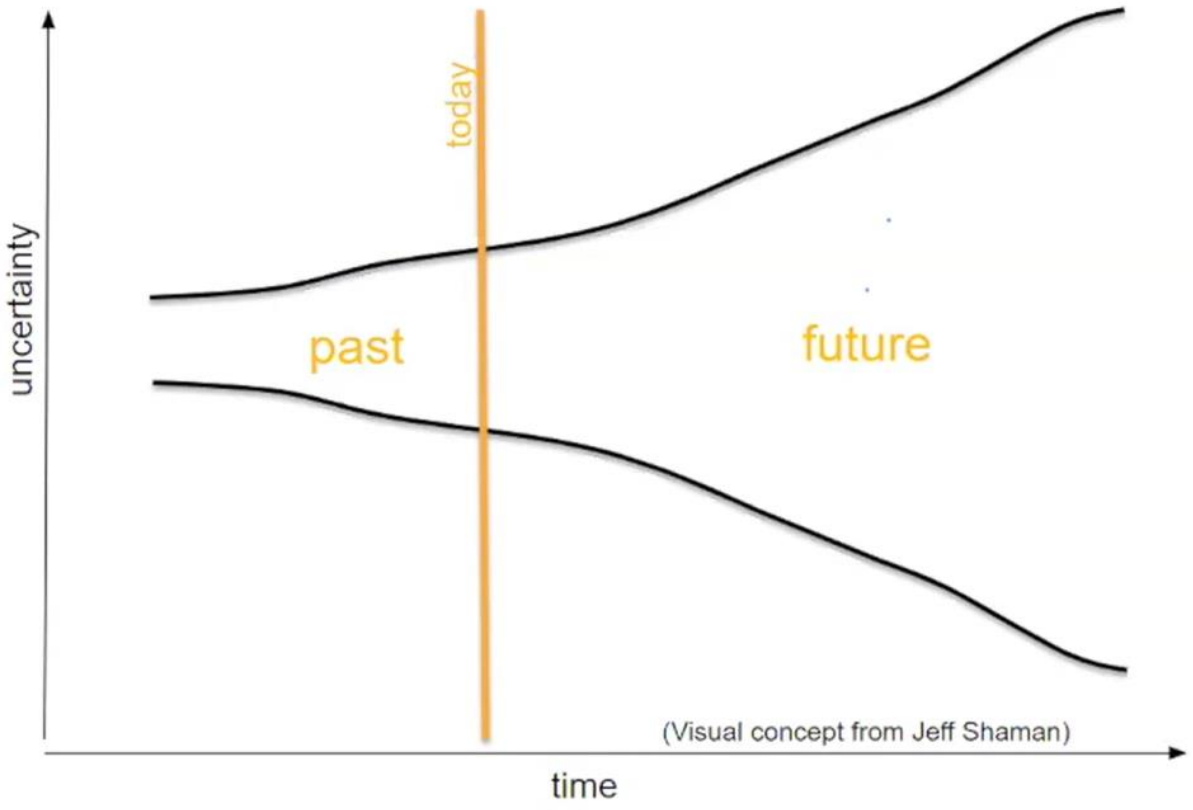
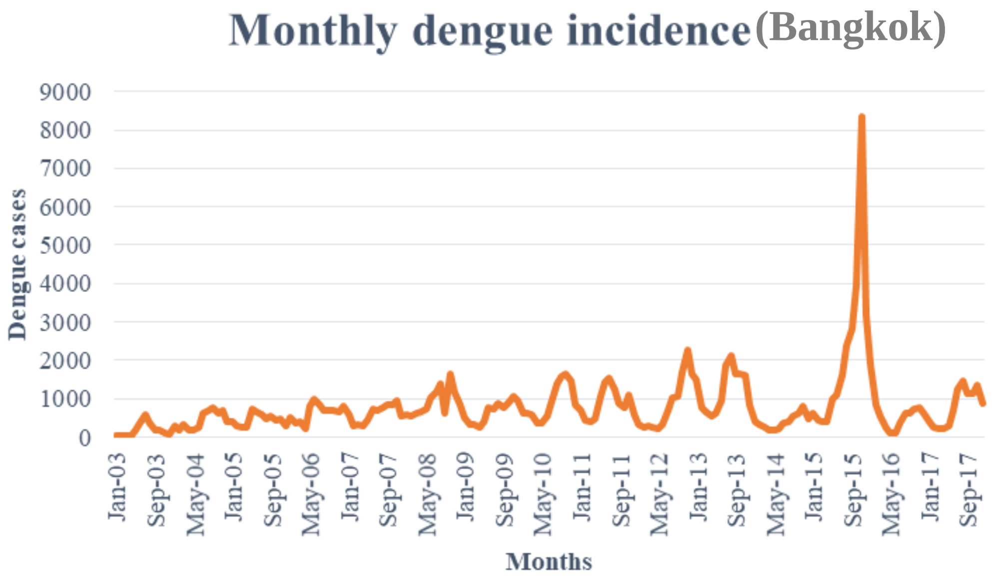


Fig: Uncertainty in past and present results in highly uncertain future

Common features of epidemic data

In general, epidemic datasets are complex and noisy in nature.
They represents the following behavioral characteristics

- ✓ Non-stationarity - The statistics of the epidemic time series changes over time.
- ✓ Non-linear - The process generating the epidemic incidence over time do not follow a linear pattern.
- ✓ Long-term dependent - The analysis of the epidemic time series suggests that they possess long memory and the rate of decay of statistical dependence of two points in the series is slower than an exponential decay.
- ✓ Seasonal - Another essential characteristic of an epidemic time series is its tendency of repeating its patterns at subsequent time intervals.



Ollech and Webel's combined seasonality test – Monthly seasonality of time series.

Mathematical transformations

Log transform

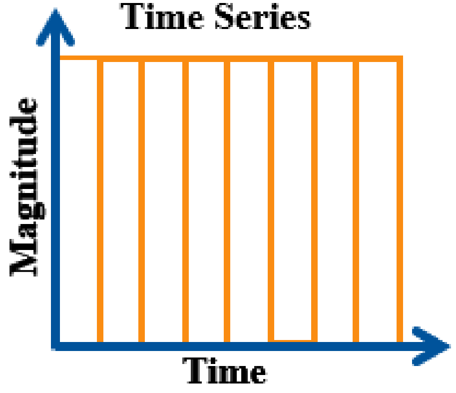
- Reduces the variability of skewed datasets.
- Highly impacted by outliers.
- Errors are symmetric on the original scale but asymmetric on the log scale.

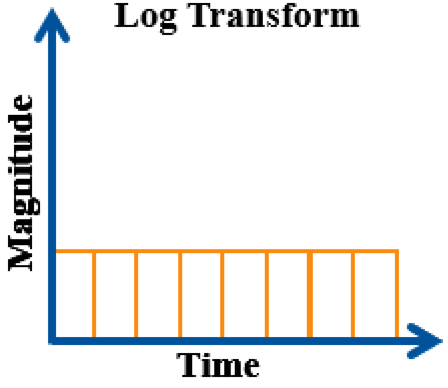
Fourier transform

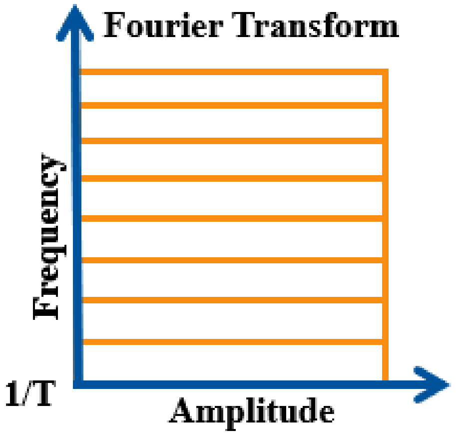
- Ideal for periodic signals.
- Represents a signal only in frequency domain
- For non-periodic signals with time-varying features, it gives averaged data, hence unsatisfactory.

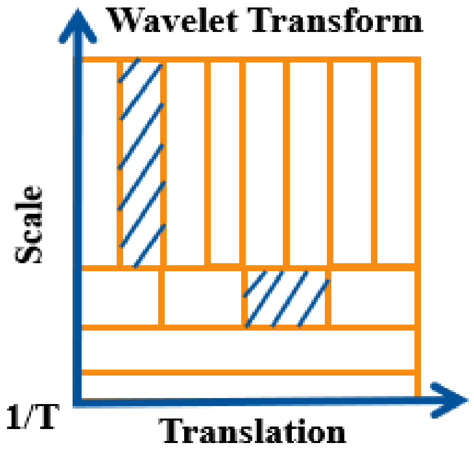
Wavelet transform

- Generalization of Fourier transform.
- It allows the independent choice of time and frequency resolution at different times and frequencies.









Proposed EWNet Model

- We propose an Ensemble Wavelet Neural Network (EWNet) that possesses the capabilities to handle the complex characteristics of epidemic datasets through its stable learning structure.

- EWNet combines wavelet decomposition (as a filtering stage) and autoregressive neural networks with exogenous variables to provide accurate forecasts of non-stationary and nonlinear time series.

- In the data pre-processing stage, Wavelet decomposition is used to generate a hierarchy of new time series from the original epidemic time series and makes them easier to model and forecast.

- MODWT decomposes the series into 'details' (contain dynamics of the epidemic systems at different scales) and 'smooth' (trend). This handles the seasonality and non-stationarity of the series.

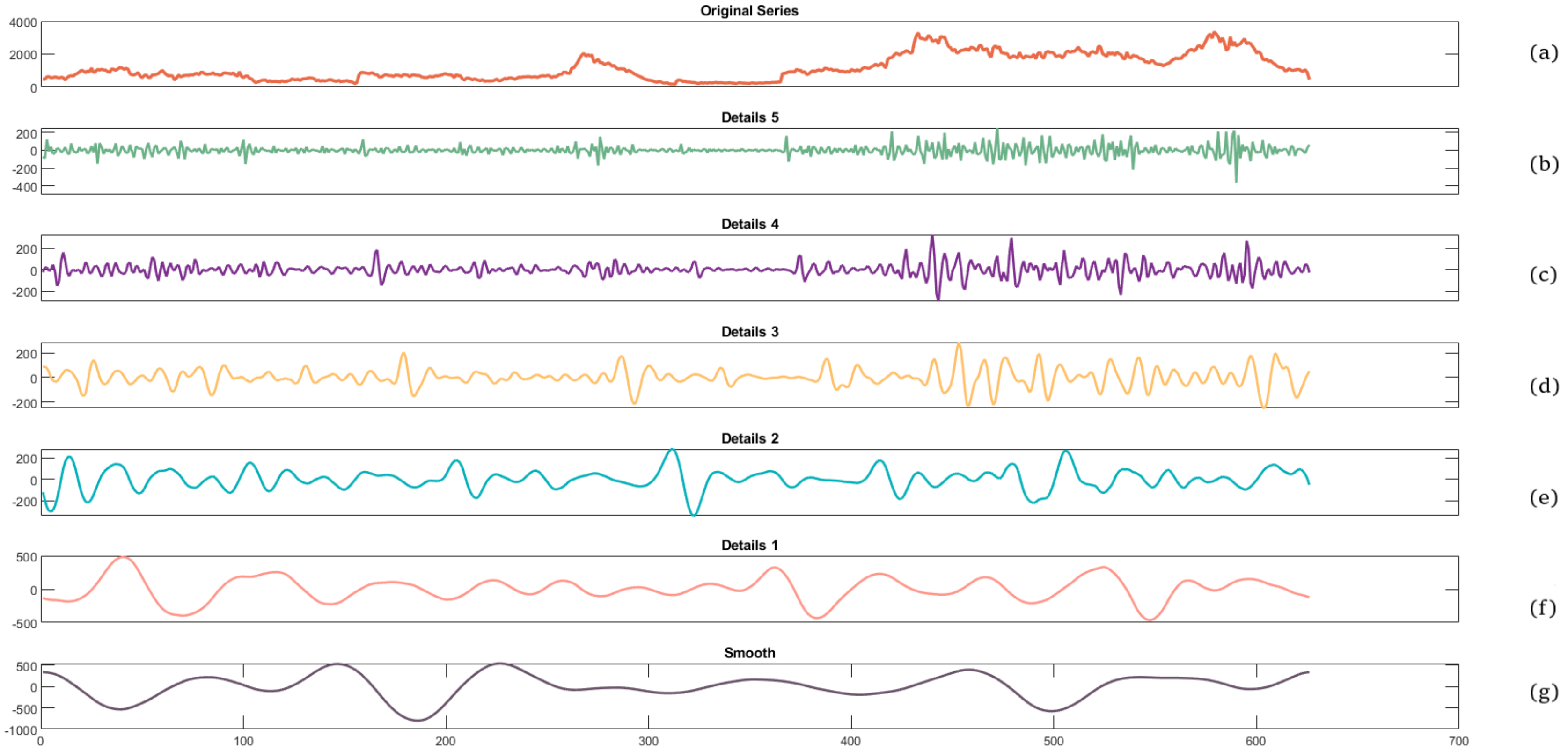
- Multi-resolution analysis (MRA) of the MODWT approach transforms non-stationary time series y_t into J details $D_{j,t}$ $(j = 1, 2, \cdots, J)$ and a smooth $S_{J,t}$ coefficients. Mathematically, it can be represented as follows:

$$y_t = \sum_{j=1}^{J} D_{j,t} + S_{J,t} \,,$$

where $D_{j,t}$ denotes the irregular fluctuations or high-frequency components at scale $(j = 1, 2, \cdots, J)$, and $S_{J,t}$ denotes the overall trend or low-frequency components of the original series y_t.

Proposed EWNet Model

MODWT decomposes the series into ‘details’ (contain dynamics of the epidemic systems at different scales) and ‘smooth’ (trend).



MRA-based MODWT decomposition of the Colombia dengue dataset with the original epidemic time series and its 6 levels. In Figure, (a) denotes the original time series in actual frequency scale; (b)-(f) denote the detail coefficients reproduced by the MODWT algorithm with haar filter, and (g) represents the scaling coefficients of the series generated by MODWT algorithm with haar filter. The figure depicts time-localized information on frequency patterns that are identified by wavelets.

Proposed EWNet Model

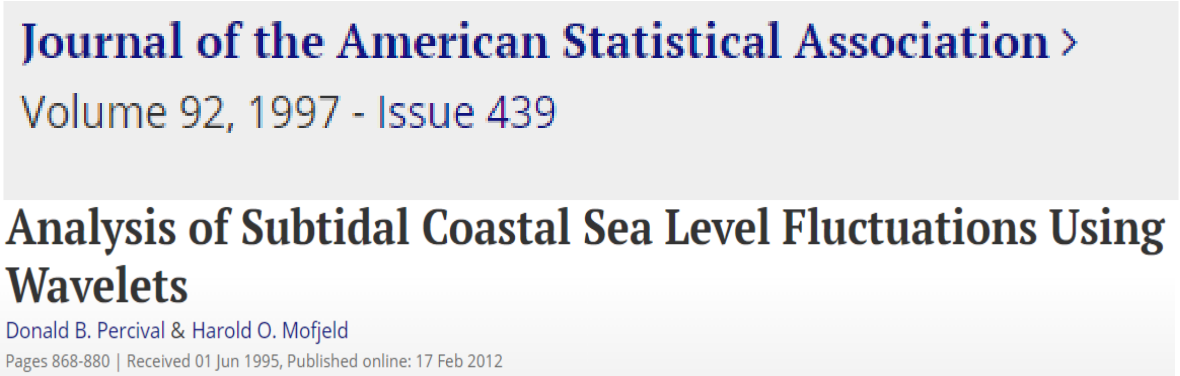
- Subsequently the ‘details’ and ‘smooth’ series are modeled using an autoregressive neural network (ARNN) with a pre-defined architecture in an ensemble setup.

- EWNet (p, k) model is a non-stationary and non-linear model which can be written as follows:

$$y_t = \sum_{j=1}^{J} f_j\left(D_{j,t}\right) + f_0\left(S_{J,t}\right),$$

where $J + 1$ ($[\log_e(\textit{length of training data})]$) is the number of wavelet levels, f_i $(i = 0, 1, 2, \cdots, J)$ is the one-hidden layered feedforward ARNN with p input nodes and k hidden nodes.

- Choice of p and k?
- Estimator complexity control!

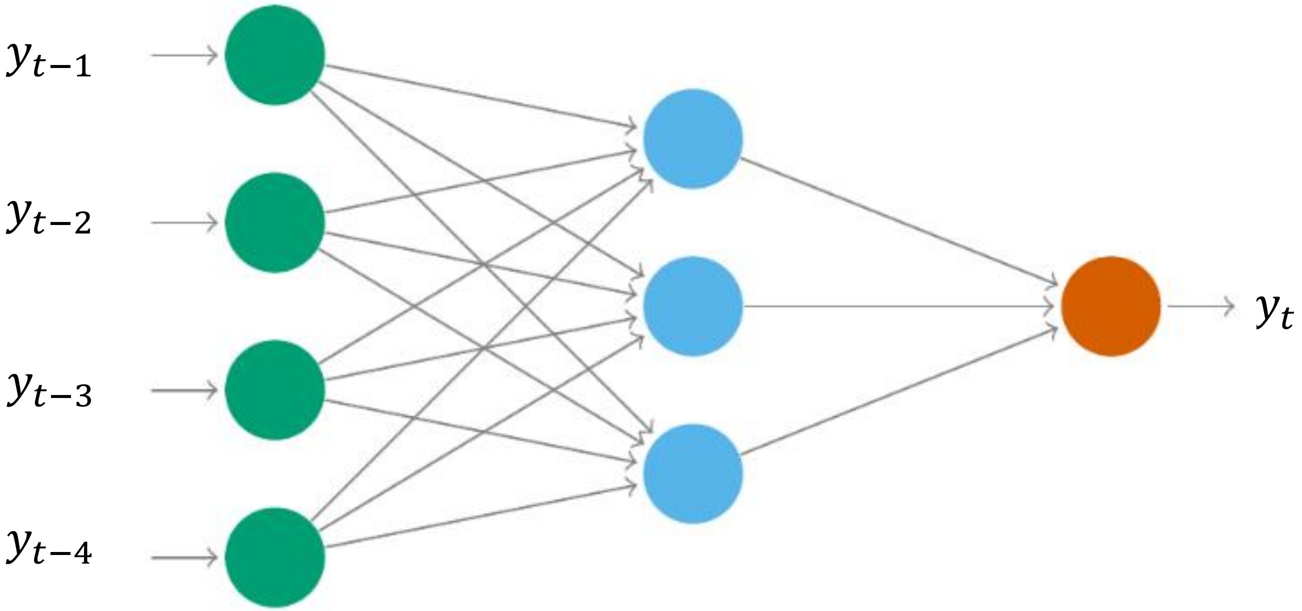

Journal of the American Statistical Association >

Volume 92, 1997 - Issue 439

Analysis of Subtidal Coastal Sea Level Fluctuations Using Wavelets

Donald B. Percival & Harold O. Mofjeld

Pages 868-880 | Received 01 Jun 1995, Published online: 17 Feb 2012



A neural network with four inputs and one hidden layer with three hidden neurons.

Proposed EWNet Model Architecture

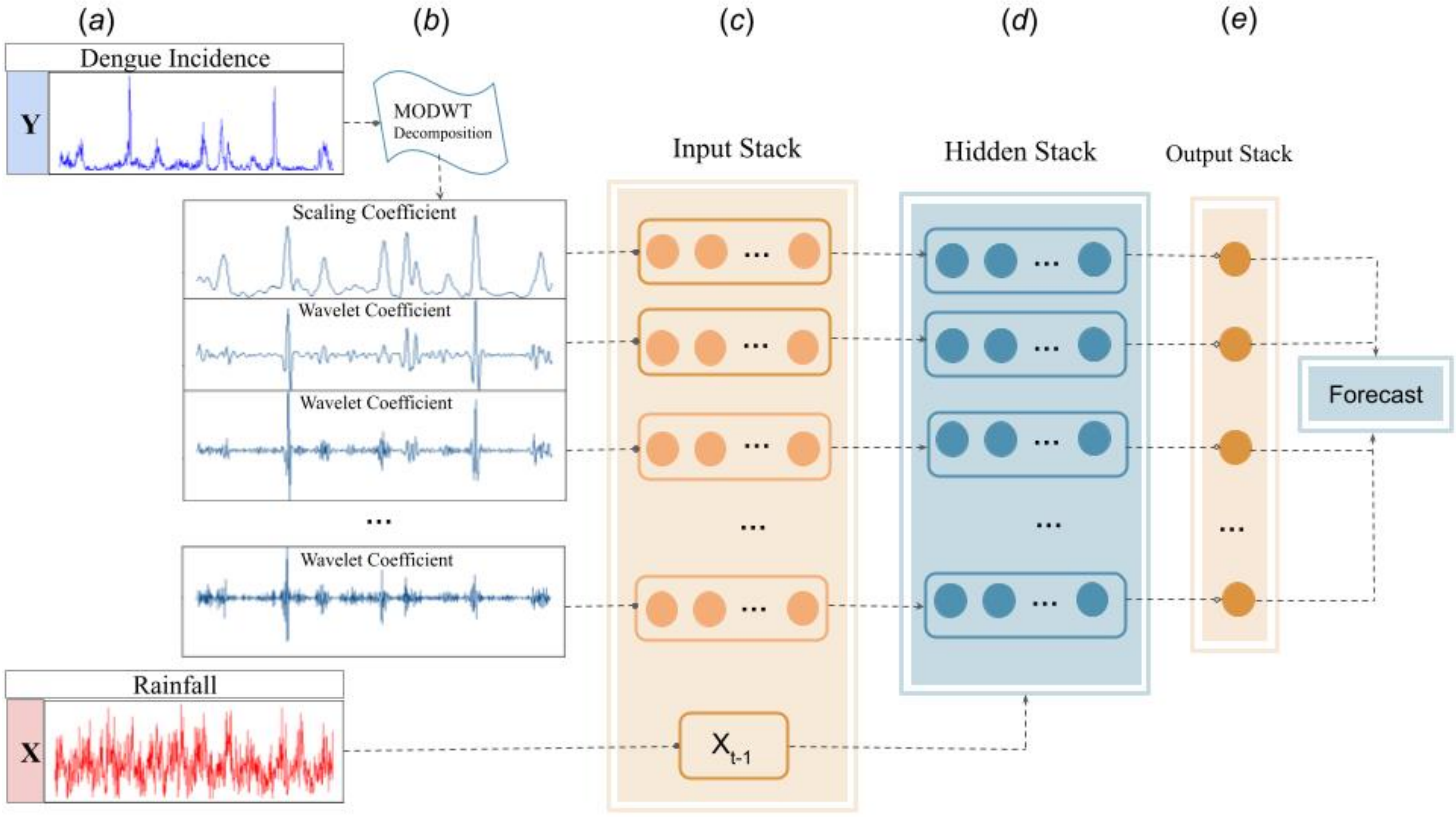


Figure: The proposed EWNet workflow:

a) To predict dengue incidence cases, we provide a weekly time series of dengue cases (Y_t) and rainfall (X_t) in the training period;
b) We perform a MODWT based MRA transformation on Y and generate multiple series of details and smooth coefficients;
c) We begin to train local auto-regressive neural networks to individually model the transformed series along with rainfall dataset in the input stack;
d) Each of the neural networks is trained with a single hidden layer having a pre-specified number of nodes inside the hidden stack;
e) The output stack comprises of one-step ahead forecast generated by individual neural networks. These predictions are combined to generate the final out-of-sample forecast.

Proposed EWNet Framework

- h-step ahead forecasts of Y_t based on N historical observations $(Y_1, Y_2, \dots, Y_N)$ can be generated by generating the simultaneous forecast for the details and smooth coefficients in an ensemble setup thus we have

$$\hat{Y}_{N+h} = \sum_{j=1}^{J} \hat{D}_{j,N+h} + \hat{S}_{J,N+h}$$

where, $\hat{D}_{j,N+h}$ and $\hat{S}_{J,N+h}$ are the h-step forecasts for the details and smooth coefficients.

From statistical point of view, we study the associated Markov Chain for the EWNet process.

- Stable learning
- Irreducibility
- Geometric ergodicity
- Asymptotic stationarity
- Empirical Risk Minimization
- Conformal Prediction

These theoretical results guarantee that EWNet model cannot show 'explosive' behavior or growing variance over time. Helps in model specification, estimation during simulation study. (See details in paper)

Stable learning in EWNet

Choice of k

From the boundary of stable learning in hidden and output neurons, we introduce a balancing equation as:

$$\alpha\eta\left(\frac{p}{k}\right) = \eta k,$$

where η *is the* learning rate and α *is a consistency constant with* $1 \leq \alpha \leq p$. *Thus, we initially choose* $k = \sqrt{\alpha p} \Rightarrow k \in \left(\sqrt{p}, p\right)$ *and using the AM-GM inequality we finally select* $k = \left\lfloor \frac{p+1}{2} \right\rfloor$.

Practical Significance

- The number of hidden nodes in the EWNet model is set to a fixed value depending on the number of lagged.
- Due to this, the running time of the EWNet model is minimal as compared to unstable neural networks in which the number of hidden nodes either becomes too large or too small.
- Thus, our proposed model does not face the problem of under-fitting or over-fitting.

Empirical Risk Minimization in EWNet

Proposition

Let the autoregressive neural network (of order k) is applied to the original data y_t by minimizing the risk $\mathfrak{R}_{Emp}$ and EWNet fits the ensemble model on the transformed data by minimizing the empirical risk $\mathfrak{R}_{Emp}^{W}$, then we have

$$\min \mathfrak{R}_{Emp}^{W} \leq \min \mathfrak{R}_{Emp}$$

Practical Significance

- In statistical learning theory, ERM defines a family of learning models and provides theoretical bound on their performances.
- This is useful since in practice we can't generalize how well a model will work (called true risk) due to not knowing the true data distribution. However, we can instead measure its performance on a known set of training data (called empirical risk).
- The above result show the robustness of the wavelet decomposed approach in EWNet from ERM perspective.

Model Evaluation

- **Forecast horizons**

To validate the efficiency of the proposed EWNet framework we check the forecast accuracy of the model for three different horizons namely, long-term forecast (52 weeks), medium-term forecast (26-weeks), and short-term forecast (13-weeks) for weekly datasets.

- **Performance Metrics**

This study adopts four popularly used key performance indicators, namely Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Symmetric Mean Absolute Percent Error (SMAPE), and Mean Absolute Scaled Error (MASE), to evaluate the deviation between the forecasts and the ground truth. By general convention, the model having the least values of these metrics is the *best* performing model.

- **Benchmark comparison**

Statistical models: Random Walk, ARIMA, ETS, Theta, WARIMA, SETAR, TBATS, BSTS.

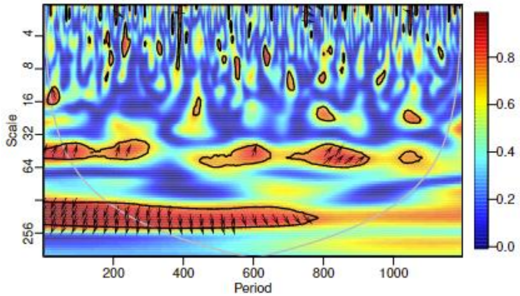
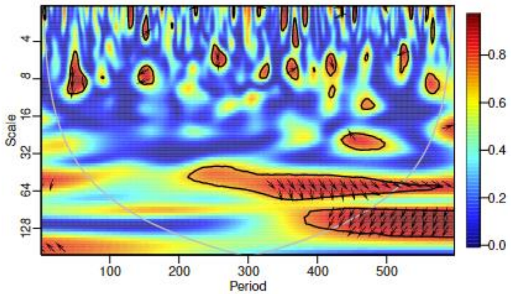
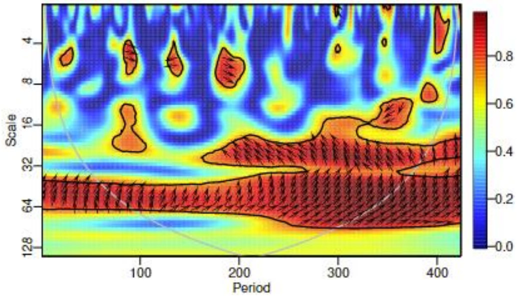
Machine learning models: MLP, ARNN, SVR, LSTM, NBeats, TCN, DeepAR, Transformer.

Other hybrid and ensemble approaches.

Causality Test

In our study, we have identified the causal relationship between dengue incidence cases and rainfall using different statistical significance tests:

- Granger Causality test
- Wavelet Coherence plot

Statistical test	San Juan	Iquitos	Ahmedabad
Granger Causality test (Cases vs Rainfall)	Causality found (0.032*)	Causality found (0.030*)	Causality found (0.049*)
Wavelet Coherence plot			

MCB test results

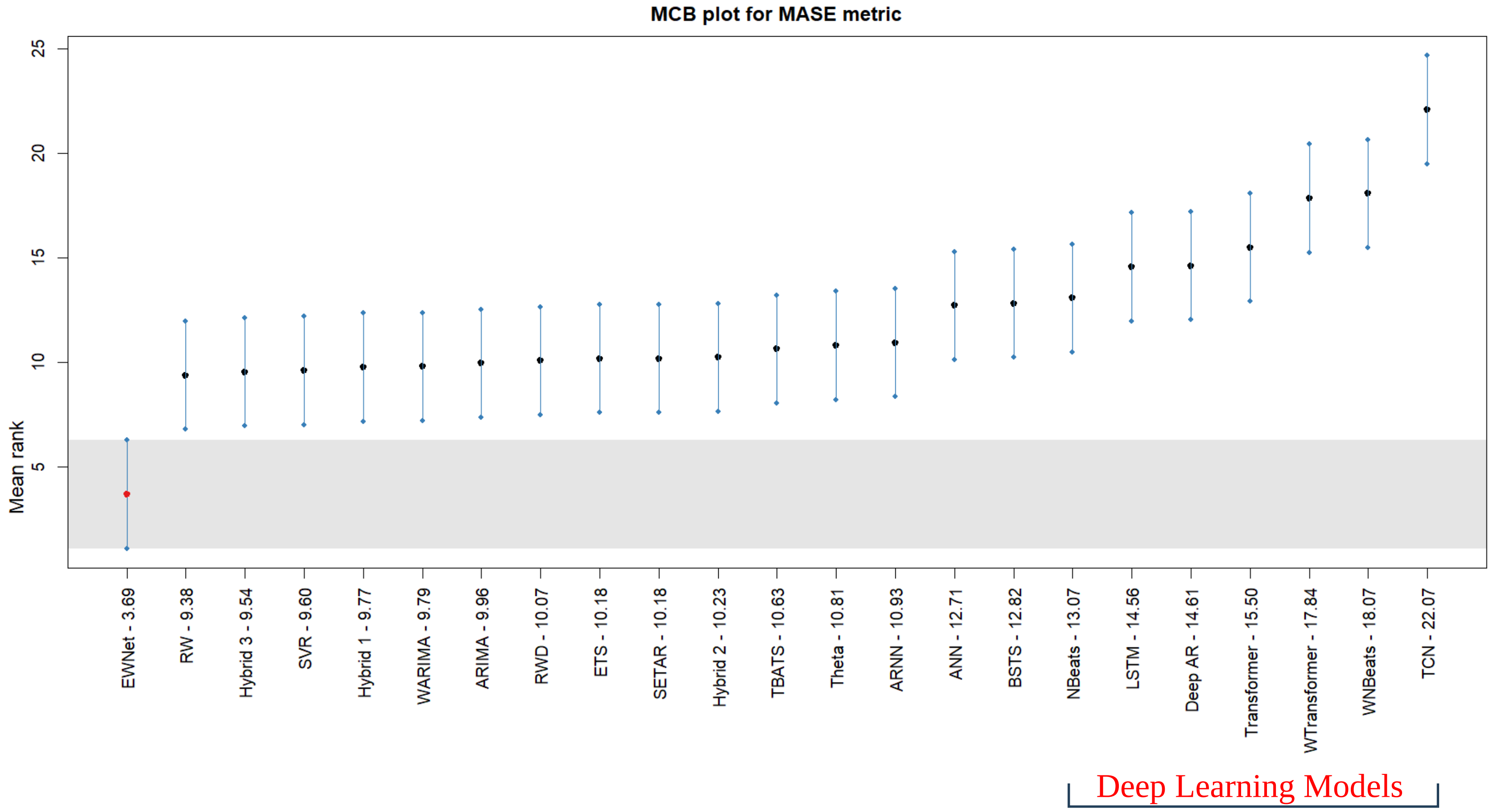


Figure: Schematic visualization of the multiple comparisons with the best (MCB) test. The plot provides the result for the MASE metric. For example, in the figure, EWNet - 3.69 specifies the rank of the EWNet model. The blue lines indicate the critical distance of the model, the middle point of this interval, which is denoted by black (significant) or red (not significant), represents the mean rank, and the shaded region marks the reference value.

Overall Performance Analysis

The overall experimental evaluation of the proposed model and the benchmark forecasters reveal some interesting observations.

- The traditional linear models like ETS and ARIMA fails to handle the irregularities of real-world dengue datasets. Although the exogenous variant of these models (ETSX and ARIMAX) marginally improves the forecast accuracy, the tendency of these methods to approximate the complex relationship between rainfall and dengue incidence cases by a linear function with a constant rate of change results in their failure.

- In case of the data-driven machine learning and deep learning frameworks the nonlinear relationship is modeled more precisely, however, their overall performance is unsatisfactory for long-term forecasts. Data set size creates a barrier to the performance of the deep learners

- The proposed EWNet approach can optimally model the complex non-linear relationship between the observed and covariate series, thus resulting in improved forecasts. Unlike the deep learning approaches, the stable architecture of our proposal limits the number of training parameters, hence restricting the model over-fitting.

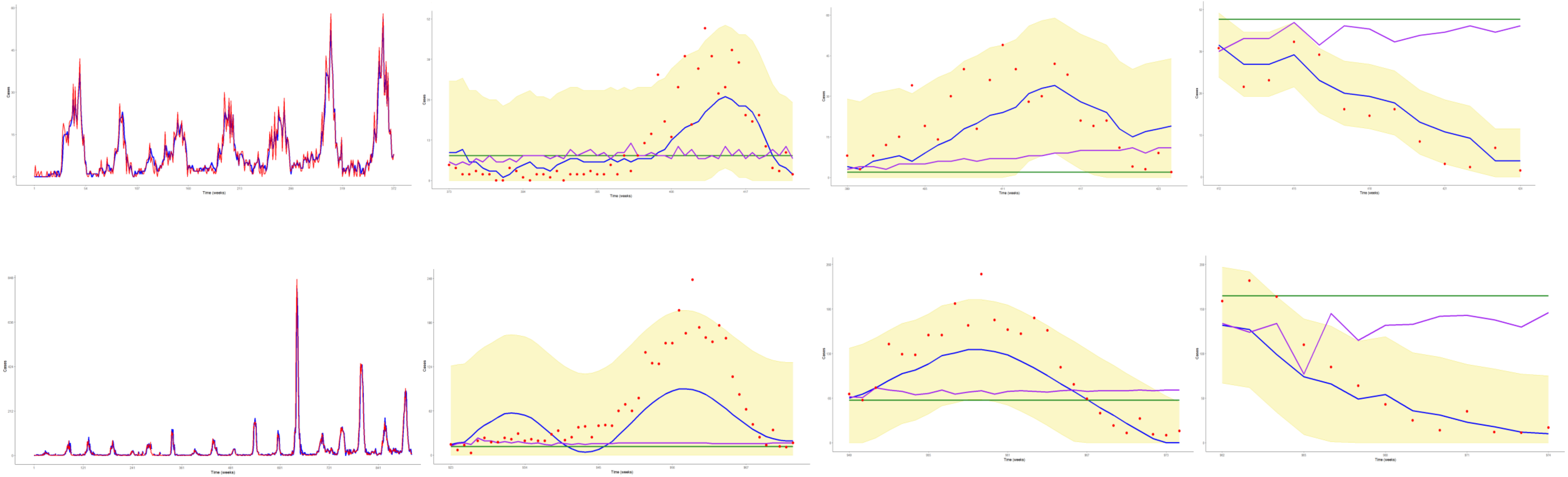
- Moreover, the use of MODWT-based MRA transformation generates the wavelet and scale coefficients that can overlook the signal through noise resulting in accurate long-term forecasts.

Actual vs Forecast Visualization

Training data and fitted values | Long-term forecast | Medium-term forecast | Short-term forecast

Ahmedabad Dengue

Australia Flu

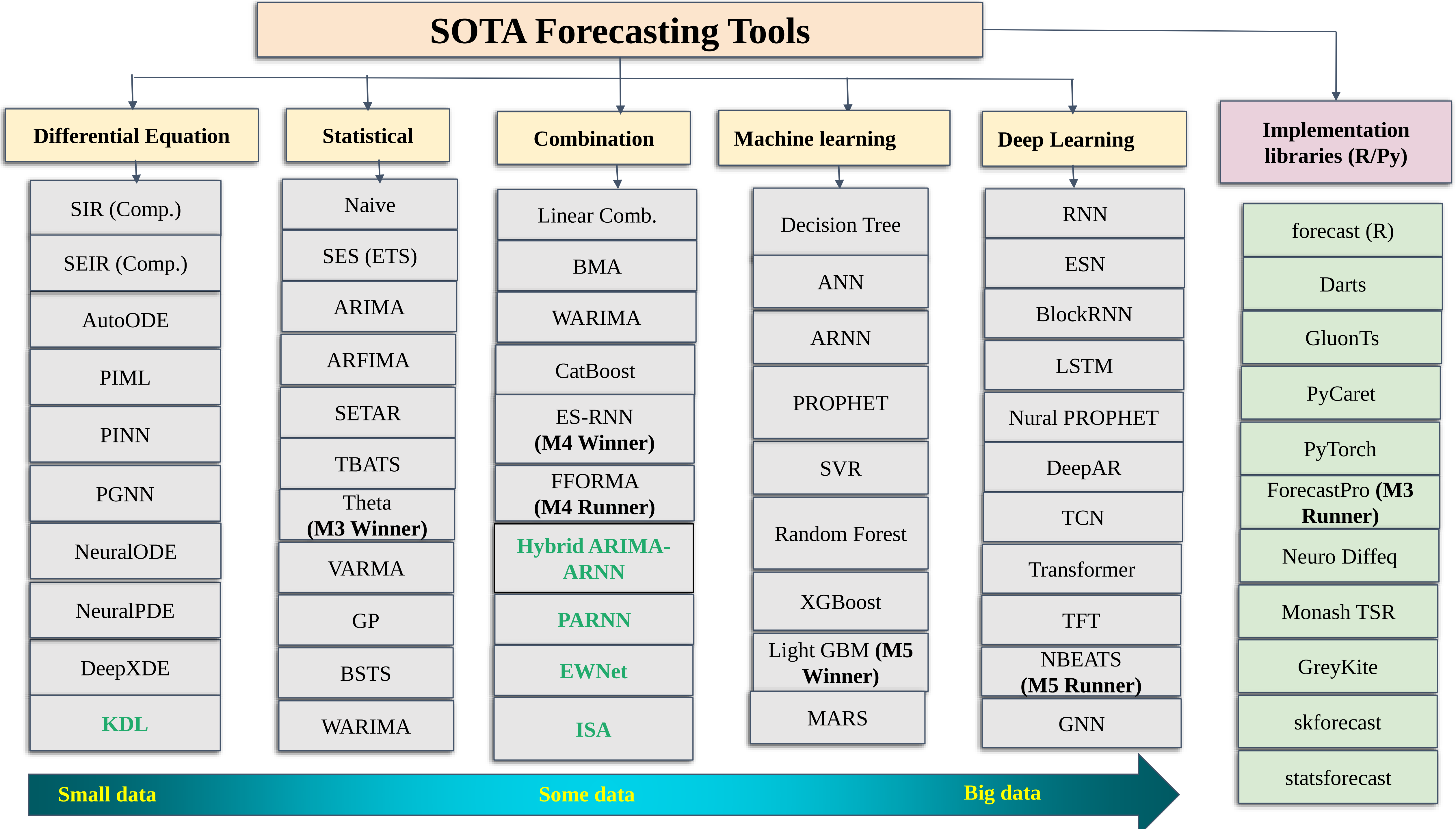


Train data
Test data
EWNet Fitted & Forecast
Random Walk Forecast
Hybrid ARIMA-ARNN Forecast
Prediction Interval

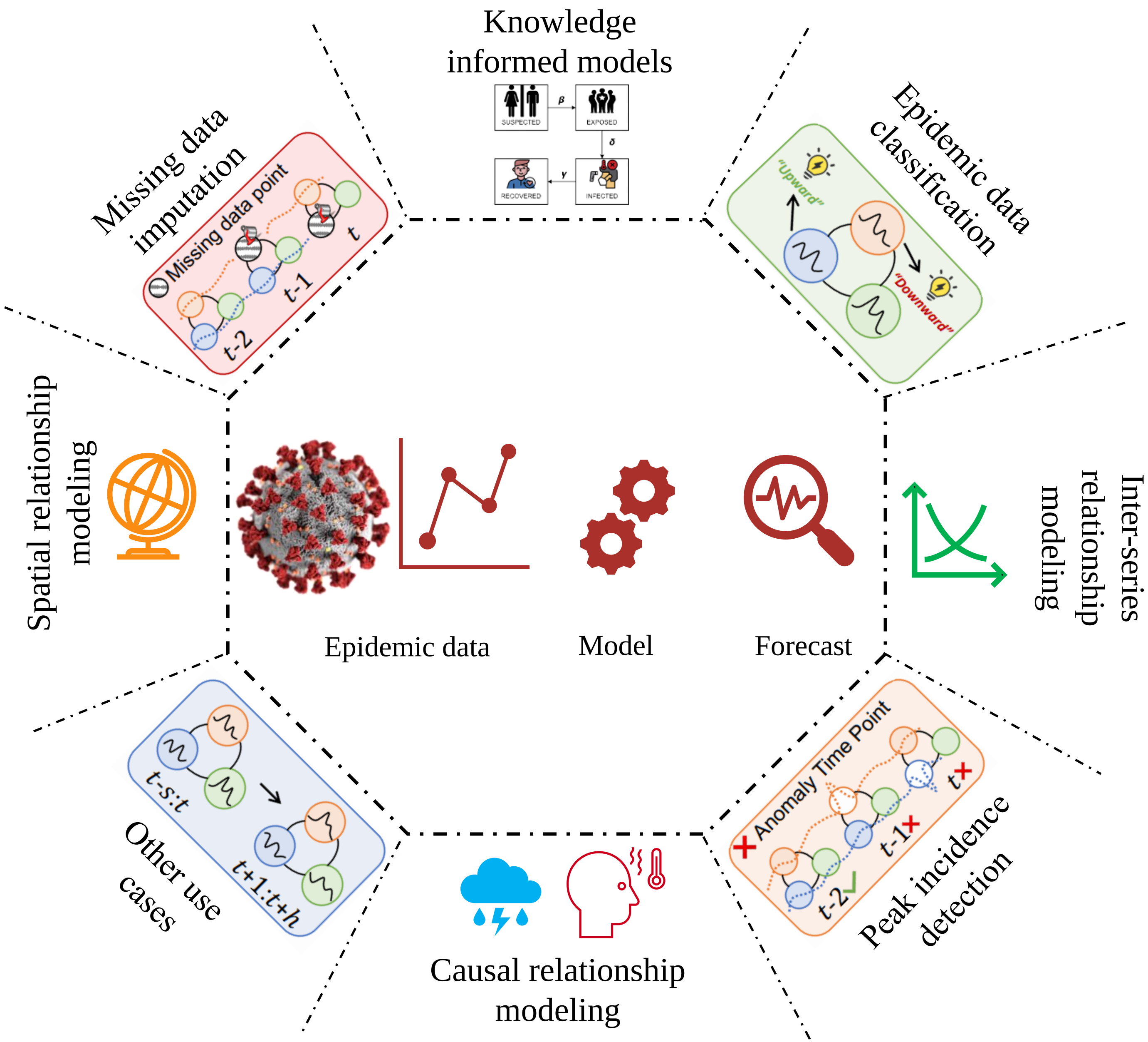
Data are collected from: DENGUE FORECASTING
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

Remarks

- Suitable for non-stationary, seasonal, and non-linear forecasting problems with limited historical data.
- Theoretical properties (learning stability, geometric ergodicity, and asymptotic stationarity) ensure the stability of model output.
- Simple and easily interpretable model, fast in implementation due to pre-defined architecture (multivariate set-up is yet to be explored).
- Experimental results suggest a significant improvement in long-range forecast accuracy owing to the wavelet decomposition.
- Epidemic dataset repository: https://github.com/mad-stat/Epicasting/tree/main/Datasets
- R package for implementation: https://cran.r-project.org/web/packages/epicasting/index.html
- Medium article on EWNet implementation: https://medium.com/@madhurima.panja/epicasting-package-in-r-epidemic-forecasting-made-easy-dcdaffd694b


SOTA Forecasting Tools
Differential Equation
SIR (Comp.)
SEIR (Comp.)
AutoODE
PIML
PINN
PGNN
NeuralODE
NeuralPDE
DeepXDE
KDL
Statistical
Naive
SES (ETS)
ARIMA
ARFIMA
SETAR
TBATS
Theta (M3 Winner)
VARMA
GP
BSTS
WARIMA
Combination
Linear Comb.
BMA
WARIMA
CatBoost
ES-RNN (M4 Winner)
FFORMA (M4 Runner)
Hybrid ARIMA-ARNN
PARNN
EWNet
ISA
Machine learning
Decision Tree
ANN
ARNN
PROPHET
SVR
Random Forest
XGBoost
Light GBM (M5 Winner)
MARS
Deep Learning
RNN
ESN
BlockRNN
LSTM
Nural PROPHET
DeepAR
TCN
Transformer
TFT
NBEATS (M5 Runner)
GNN
Implementation libraries (R/Py)
forecast (R)
Darts
GluonTs
PyCaret
PyTorch
ForecastPro (M3 Runner)
Neuro Diffeq
Monash TSR
GreyKite
skforecast
statsforecast
Small data
Some data
Big data

Bright future of Epicasting



Outline

- Past of Forecasting
- Basics of Forecasting
- Epidemic Forecasting using EWNet
- Appendix

Irreducibility of EWNet

Proposition 1

The sufficient conditions of forward accessibility for the control system y_t defined by

$$y_t = \upsilon + \psi_1 y_{t-1} + \sum_{i=1}^{k} \beta_i G\left(\phi_{i,1} y_{t-1} + \mu_i\right) + \varepsilon_t, \qquad (1)$$

where $\upsilon, \beta_i, \phi_{i,1}, \mu_i; (i = 1, 2, \cdots, k)$ are the weights and G is the activation function are stated as follows:

- *$G \in C^\infty$ is a bounded, non-constant, and asymptotically constant function (C^∞) (any function is C^∞ if derivatives of all orders are continuous).*
- *The linear part of R.H.S. of Eqn. (1) is controllable, i.e., $\psi_1 \neq 0$.*

The controllability of the linear components of the ARNN process as discussed in the Prop. 1 implies forward accessibility. But, the associated Markov chain is said to be irreducible when the support of the distribution of the noise process is sufficiently large.

Irreducibility of EWNet

Remark

Our proposed EWNet model can be thought of as a sum of $J+1$ *different ARNN* (p, k) *processes, where* $J+1$ *denotes the number of details and smooth coefficients obtained using the MODWT algorithm.*

Theorem 1 (Theorem of Irreducibility)

Suppose the distribution of ε_t *is absolutely continuous w.r.t. the Lebesgue measure* λ *and the probability distribution function (p.d.f.)* $\nu(\cdot)$ *of* ε_t *is positive everywhere in* $\mathcal{R}$ *and lower semi-continuous. Then under the condition prescribed in Prop. 1, the Markov chain in* $y_t = \psi_1 y_{t-1} + F(y_{t-1}) + \varepsilon_t$ *is irreducible on the state space* $(\mathcal{R}^2, \mathcal{B})$.

Theorem 1 shows the irreducibility property for the ARNN $(1, k)$ process and demonstrates its proximity to the concept of forward accessibility of a control system. However, we also showed that ARNN processes might not exhibit forward accessibility, and in such scenarios, inferring about the data-generating process from the observed data is impossible.

Ergodicity & Stationarity of EWNet

Theorem 2 (Main Theorem)

Suppose the Markov chain y_t of the ARNN $(1, k)$ process satisfies the conditions of Theorem 1 and $E|\varepsilon_t| < \infty$. Then, a sufficient condition for the geometric ergodicity (vis-a-vis asymptotic stationarity) of the Markov chain $\{y_t\}$ is that $|\psi_1| < 1$.

Theorem 2 states the sufficient condition for the geometric ergodicity of the ARNN $(1, k)$ process. Consider the following example: if $\psi_1 = 1$, then the long-term behavior of the ARNN $(1, k)$ process can be determined by the nonlinear part and the intercept term of the process. Moreover, the geometric convergence rate in Theorem 2 implies that the memory of the ARNN process vanishes exponentially fast. This means that the simplest version of the ARNN (p, k) process converges to a Wiener process. Also, theoretical results suggest that the shortcut weight corresponding to the autoregressive part determines whether the overall process is ergodic and asymptotically stationary.

Implications of the theoretical properties of EWNet

- In the ideal situation, when an irreducible ARNN process generates the data, the estimated weights are not too far from the true weights. Then, one can draw an indirect conclusion on the statistical nature of the estimated shortcut weight corresponding to the autoregressive part being less than one in absolute terms, and then the data generation process is said to be ergodic and stationary. But, if the conditions are not met, the model is likely to be unspecified, and the estimation procedure should be diligently done.

- The theoretical results of asymptotic stationarity and ergodicity for the EWNet (p, k) model would directly follow from the ARNN (p, k) process since the proposed EWNet is a simple aggregation of several ARNN models fitted after the Wavelet decomposition of the time series data. These theoretical results guarantee that the proposed EWNet model cannot show 'explosive' behavior or growing variance over time.

THANK YOU!

essentially,
all models are wrong,
but some are useful

George E. P. Box

Scan me
for Slides: