

Adaptive Mortality Drift for Short-Horizon Forecasting of Under-Five Mortality in West Africa: Model Development, Monte Carlo Validation, Robustness, and a 16-Country Application

Mumini Idowu Adarabioyo

adarabioyomi@abuad.edu.ng

Afe Babalola University

Research Article

Keywords: under-five mortality, West Africa, forecasting, adaptive drift, time series, Monte Carlo simulation, rolling-origin validation, child survival

Posted Date: October 12th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10711146/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Adaptive Mortality Drift for Short-Horizon Forecasting of Under-Five Mortality in West Africa: Model Development, Monte Carlo Validation, Robustness, and a 16-Country Application

Mumini Idowu ADARABIOYO

Department of Mathematical and Physical Science, Afe Babalola University, Ado-Ekiti, Nigeria

Corresponding email: adarabioyomi@abuad.edu.ng

Abstract

Accurate short-term (less than 5 years) forecasts of under-five deaths are useful for tracking child-survival outcomes and limited health resources. The work presented here builds an Adaptive Mortality Drift (AMD) model for the annual under-five mortality rates in a set of 16 countries in West Africa. AMD models the annual mortality change as a continuous process, and estimates instead of fixing the proportion of the most recent change which is carried over to the next year. Pooled persistence estimator is calculated using the least squares method, recursive multi-step forecasts and model-based prediction intervals are calculated. For 2018 to 2024, AMD has RMSE of 0.177, MAE of 0.130, MAPE of 0.208%, and MASE of 0.034, outperforming local linear drift, ARIMA(1,1,1), damped Holt, ridge regression, random forest, gradient boosting and a lag-1 naive benchmark for strict rolling-origin evaluation. The regional persistence estimated was 0.966, which suggests strong, yet slightly muted continuity of the changes in mortality observed in the region. The Monte Carlo experiments used 1000 replications per scenario, and resulted in small bias, RMSE of the parameters decreasing with the length of the history and nominal 95% cover of the baseline design. AMD remained stable under heavy-tailed innovations and moderate missingness, but pooled parameter recovery deteriorated under country heterogeneity and structural breaks. Targeted shrinkage and exponentially weighted variants improved those stress scenarios but did not outperform AMD on the observed West African test period. In 2024, mortality ranged from 11.14 per 1,000 live births in Cabo Verde to 115.65 in Nigeria; AMD forecasts continued declines through 2030, although large cross-country disparities persist. AMD is therefore best interpreted as a parsimonious regional short-horizon forecasting model rather than a causal or universally country-optimal model.

Keywords: under-five mortality; West Africa; forecasting; adaptive drift; time series; Monte Carlo simulation; rolling-origin validation; child survival

1. Introduction

Under-five mortality rate (U5MR) is the probability that a child born in a specified period dies before reaching age five, expressed per 1,000 live births. It is a widely used indicator of child survival and of the broader social, economic, environmental, and health-system conditions in which children live (World Health Organization [WHO], n.d.). Despite major long-run progress, child mortality remains concentrated in sub-Saharan Africa. The 2025 United Nations Inter-agency Group for Child Mortality Estimation (UN IGME) report estimated 4.9 million under-five deaths globally in 2024 and identified sub-Saharan Africa and Southern Asia as the regions bearing the heaviest burden (UN IGME, 2026).

The forecasting problem is particularly important in West Africa, where mortality levels and rates of improvement differ sharply across countries. Recent UN estimates place Western Africa's 2024 U5MR at about 94 deaths per 1,000 live births, while the Economic Community of West African States remains among the highest-burden regional groupings. These regional averages conceal large national differences and therefore motivate models that can borrow strength across countries while retaining transparent interpretation.

A substantial literature already exists on child-mortality forecasting. Mejía-Guevara et al. (2019) modelled age patterns and forecast under-five mortality in 31 sub-Saharan African countries. Adeyinka and Muhajarine (2020) compared artificial neural networks, ARIMA, and Holt-Winters methods for Nigeria. More recently, Adama et al. (2025) applied ARIMA models to Ghana, Niger, Nigeria, and Sierra Leone. Thus, neither multi-country mortality forecasting nor the use of conventional or machine-learning forecasting methods is itself new.

Machine-learning studies provide a complementary literature, although many address individual-level mortality classification rather than annual national-rate forecasting. Samuel et al. (2024) used the 2018 Nigeria Demographic and Health Survey to compare several supervised learning algorithms, while Asnake et al. (2026) analysed recent DHS data from 23 sub-Saharan African countries. These studies demonstrate the value of flexible prediction, but they answer a different question from forecasting a smooth annual national mortality-rate series.

The present study therefore focuses on a narrower methodological problem: whether the very strong persistence in annual mortality changes can be represented by a simple, estimable drift mechanism that improves genuinely out-of-sample forecasts. We develop the Adaptive Mortality Drift (AMD) model, in which the latest annual change is multiplied by a persistence parameter estimated from historical changes across the regional panel. The contribution is the combination of this pooled adaptive-drift formulation, rolling-origin evaluation, Monte Carlo parameter-recovery assessment, targeted robustness experiments for heavy tails, missingness, heterogeneity and structural change, and a 16-country West African application. The model is intentionally parsimonious and is not presented as an unrelated new stochastic family: in first differences it is closely related to an AR(1) process and hence to an ARIMA(1,1,0)-type representation.

2. Related Literature and Positioning

Forecasting approaches to mortality range from demographic models to conventional time-series and machine-learning methods. Mejía-Guevara et al. (2019) developed a demographic forecasting framework for 31 sub-Saharan African countries and showed substantial heterogeneity in mortality trajectories and expected attainment of child-survival targets. Their work establishes that regional forecasting must explicitly acknowledge country differences rather than infer uniform progress from a regional average.

In Nigeria, Adeyinka and Muhajarine (2020) did a comparative study of an artificial neural network (ANN) based on GMDH compared to ARIMA and Holt-Winters methods using the U5MR for the country. They demonstrate that the benefit of nonlinear methods over traditional models is possible in a single country setting, but also highlight the importance of explicit out-of-sample accuracy comparisons. After forecasting with time series in four West African Countries, Adama et al. (2025) proposed country specific ARIMA specifications using information criteria and forecast diagnostics. The present work is guided by these studies.

The literature on machine learning is wider, but generally focuses on the probability or determinants of child death at the individual child level. For instance, Samuel et al. (2024) applied supervised learning to ANN, KNN, SVM, naive Bayes, random forest and logistic regression on Nigerian DHS data; and Asnake et al. (2026) applied supervised learning over 23 countries in sub-Saharan Africa. This type of model may be very useful for risking stratification and determinant analysis, but it is fundamentally different from the annual country-level U5MR forecasting.

In this study, the forecasting evaluation is based on the accepted principles of forecasting accuracy. The RMSE and MAE give a summary of scale dependent error, MAPE gives a relative percentage summary and MASE allows comparison with a naive scaling benchmark. In particular, Hyndman and Koehler (2006) argued for the use of scaled errors like MASE, since percentage and scale dependent errors may have undesirable properties in certain forecasting situations. The concept of structural-change robustness is also applicable because the time-series relationship may change after policy, epidemiological, conflict or economic shocks and Aue and Horváth (2013) highlight the need to identify the instability and not presume constant parameters indefinitely.

3. Materials and Methods

3.1 Data and study region

Under-five mortality rates are derived from the annual estimates of the UN IGME/WHO data system up to 2024 for 16 West African countries including Benin, Burkina Faso, Cabo Verde, Côte d'Ivoire, The Gambia, Ghana,

Guinea, Guinea-Bissau, Liberia, Mali, Mauritania, Niger, Nigeria, Senegal, Sierra Leone, and Togo. U5MR is expressed as the number of deaths before the age of 5/1000 live births. The outcome series are smooth annual national estimates rather than individual child-level observations. Contextual indicators were explored during model development, but the final short-horizon comparison showed that recent mortality dynamics dominated the contextual predictors for one-year-ahead forecasting; consequently, the final AMD specification is intentionally parsimonious.

3.2 Adaptive Mortality Drift model

Let $M_{i,t} > 0$ denote the under-five mortality rate (U5MR), measured as deaths before age five per 1,000 live births, for country $i = 1, \dots, N$ in year t . Define the annual mortality change by

$$\Delta M_{i,t} = M_{i,t} - M_{i,t-1}.$$

The Adaptive Mortality Drift (AMD) model assumes that the most recent annual mortality change persists into the next period at an estimable rate λ :

$$\Delta M_{i,t} = \lambda \Delta M_{i,t-1} + \varepsilon_{i,t},$$

with

$$E(\varepsilon_{i,t}) = 0, \quad \text{Var}(\varepsilon_{i,t}) = \sigma^2.$$

Substituting the definition of the first difference gives the equivalent level representation

$$M_{i,t} = M_{i,t-1} + \lambda(M_{i,t-1} - M_{i,t-2}) + \varepsilon_{i,t}.$$

Consequently, the conditional one-step-ahead forecast based on information through year $t - 1$ is

$$\hat{M}_{i,t|t-1} = M_{i,t-1} + \hat{\lambda}(M_{i,t-1} - M_{i,t-2}).$$

The parameter λ has a direct mortality-dynamics interpretation. When $\lambda = 1$, the preceding annual change is carried forward completely and AMD reduces to ordinary local linear drift. When $\lambda = 0$, the forecast reduces to the lag-1 naive forecast $\hat{M}_{i,t|t-1} = M_{i,t-1}$. Values $0 < \lambda < 1$ imply persistence with attenuation, $\lambda > 1$ implies acceleration of the most recent change, and $\lambda < 0$ implies reversal. Dynamic stability of the change process is obtained when $|\lambda| < 1$.

3.2.1 Conditional moments

Conditional on the two most recent mortality levels,

$$E(M_{i,t} | \mathcal{F}_{t-1}) = M_{i,t-1} + \lambda(M_{i,t-1} - M_{i,t-2}),$$

and

$$\text{Var}(M_{i,t} | \mathcal{F}_{t-1}) = \sigma^2.$$

Thus, λ governs the conditional mean trajectory, whereas σ^2 governs short-run innovation uncertainty.

3.3 Estimation and derivative structure

Write

$$X_{i,t} = \Delta M_{i,t-1}, \quad Y_{i,t} = \Delta M_{i,t}.$$

Then the AMD change equation is a pooled zero-intercept regression,

$$Y_{i,t} = \lambda X_{i,t} + \varepsilon_{i,t}.$$

3.3.1 Least-squares criterion

For the set $\mathcal{J}_T$ of country-year pairs available at a forecast origin T , the pooled residual sum of squares is

$$S_T(\lambda) = \sum_{(i,t) \in \mathcal{J}_T} (Y_{i,t} - \lambda X_{i,t})^2.$$

The first derivative with respect to λ is

$$\frac{\partial S_T(\lambda)}{\partial \lambda} = -2 \sum_{(i,t) \in \mathcal{J}_T} X_{i,t} (Y_{i,t} - \lambda X_{i,t}).$$

Expanding the derivative gives

$$\frac{\partial S_T(\lambda)}{\partial \lambda} = -2 \sum X_{i,t} Y_{i,t} + 2\lambda \sum X_{i,t}^2.$$

The stationary condition is obtained by setting the derivative equal to zero:

$$-2 \sum X_{i,t} Y_{i,t} + 2\lambda \sum X_{i,t}^2 = 0.$$

Hence the closed-form pooled estimator is

$$\hat{\lambda}_T = \frac{\sum_{(i,t) \in \mathcal{J}_T} \Delta M_{i,t-1} \Delta M_{i,t}}{\sum_{(i,t) \in \mathcal{J}_T} (\Delta M_{i,t-1})^2}.$$

The second derivative is

$$\frac{\partial^2 S_T(\lambda)}{\partial \lambda^2} = 2 \sum_{(i,t) \in \mathcal{J}_T} X_{i,t}^2.$$

Provided that at least one preceding annual change is non-zero,

$$\frac{\partial^2 S_T(\lambda)}{\partial \lambda^2} > 0,$$

so the criterion is strictly convex in λ and the stationary point is the unique least-squares minimum.

3.3.2 Gaussian likelihood and score derivatives

If the innovations are conditionally Gaussian,

$$\varepsilon_{i,t} \stackrel{\text{ind}}{\sim} N(0, \sigma^2),$$

then, apart from constants, the log-likelihood is

$$\ell(\lambda, \sigma^2) = -\frac{n_T}{2} \log(2\pi) - \frac{n_T}{2} \log(\sigma^2) - \frac{1}{2\sigma^2} \sum_{(i,t) \in \mathcal{J}_T} (Y_{i,t} - \lambda X_{i,t})^2,$$

where $n_T = |\mathcal{J}_T|$. The score for λ is

$$\frac{\partial \ell}{\partial \lambda} = \frac{1}{\sigma^2} \sum_{(i,t) \in \mathcal{J}_T} X_{i,t} (Y_{i,t} - \lambda X_{i,t}).$$

Setting this score to zero reproduces the least-squares estimator $\hat{\lambda}_T$. The score for σ^2 is

$$\frac{\partial \ell}{\partial \sigma^2} = -\frac{n_T}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{(i,t) \in \mathcal{I}_T} (Y_{i,t} - \lambda X_{i,t})^2.$$

Setting this derivative equal to zero yields the Gaussian maximum-likelihood variance estimator

$$\hat{\sigma}_{\text{ML},T}^2 = \frac{1}{n_T} \sum_{(i,t) \in \mathcal{I}_T} \hat{\varepsilon}_{i,t}^2,$$

where

$$\hat{\varepsilon}_{i,t} = Y_{i,t} - \hat{\lambda}_T X_{i,t}.$$

For unbiased residual-variance estimation in the one-parameter zero-intercept regression, the denominator may instead be taken as $n_T - 1$:

$$\hat{\sigma}_T^2 = \frac{1}{n_T - 1} \sum_{(i,t) \in \mathcal{I}_T} \hat{\varepsilon}_{i,t}^2.$$

The relevant second derivative of the log-likelihood with respect to λ is

$$\frac{\partial^2 \ell}{\partial \lambda^2} = -\frac{1}{\sigma^2} \sum_{(i,t) \in \mathcal{I}_T} X_{i,t}^2 < 0,$$

which confirms concavity of the Gaussian log-likelihood in λ . Conditional on the regressors, the usual model-based variance of $\hat{\lambda}_T$ is

$$\text{Var}(\hat{\lambda}_T | X) = \frac{\sigma^2}{\sum X_{i,t}^2},$$

and the plug-in standard error is

$$\text{SE}(\hat{\lambda}_T) = \sqrt{\frac{\hat{\sigma}_T^2}{\sum X_{i,t}^2}}.$$

A nominal 95% Wald interval is therefore

$$\hat{\lambda}_T \pm 1.96 \text{SE}(\hat{\lambda}_T).$$

All estimation in the empirical rolling-origin experiment is recursive: a forecast for year $T + 1$ uses only observations dated T or earlier. This restriction prevents information from the evaluation year from entering $\hat{\lambda}_T$.

3.4 Multi-step forecasting, forecast-error variance, and prediction intervals

At a forecast origin T , the predicted first future change is

$$\widehat{\Delta M}_{i,T+1|T} = \hat{\lambda}_T \Delta M_{i,T}.$$

Recursive substitution gives

$$\widehat{\Delta M}_{i,T+j|T} = \hat{\lambda}_T^j \Delta M_{i,T}, \quad j = 1, 2, \dots, h.$$

Therefore, the h -step level forecast is

$$\widehat{M}_{i,T+h|T} = M_{i,T} + \Delta M_{i,T} \sum_{j=1}^h \hat{\lambda}_T^j.$$

For $\hat{\lambda}_T \neq 1$, the geometric sum gives

$$\widehat{M}_{i,T+h|T} = M_{i,T} + \Delta M_{i,T} \frac{\hat{\lambda}_T(1 - \hat{\lambda}_T^h)}{1 - \hat{\lambda}_T}.$$

When $|\hat{\lambda}_T| < 1$, the forecasted annual changes decay geometrically and

$$\lim_{h \rightarrow \infty} \widehat{M}_{i,T+h|T} = M_{i,T} + \frac{\hat{\lambda}_T}{1 - \hat{\lambda}_T} \Delta M_{i,T}.$$

This limiting form explains why AMD does not project the most recent annual decline indefinitely at an unchanged slope.

3.4.1 Forecast-error decomposition

Treating λ as fixed, the one-step error is

$$e_{i,T+1} = \varepsilon_{i,T+1}.$$

For two steps,

$$e_{i,T+2} = (1 + \lambda)\varepsilon_{i,T+1} + \varepsilon_{i,T+2}.$$

More generally,

$$e_{i,T+h} = \sum_{m=1}^h \left(\frac{1 - \lambda^m}{1 - \lambda} \right) \varepsilon_{i,T+h-m+1}.$$

Under independent innovations with common variance σ^2 ,

$$\text{Var}(e_{i,T+h}) = \sigma^2 \sum_{m=1}^h \left(\frac{1 - \lambda^m}{1 - \lambda} \right)^2.$$

The corresponding plug-in forecast standard error is

$$\widehat{\text{SE}}_{i,T+h} = \hat{\sigma}_T \sqrt{\sum_{m=1}^h \left(\frac{1 - \hat{\lambda}_T^m}{1 - \hat{\lambda}_T} \right)^2}.$$

An approximate two-sided $(1 - \alpha)$ prediction interval is

$$\widehat{M}_{i,T+h|T} \pm z_{1-\alpha/2} \widehat{\text{SE}}_{i,T+h}.$$

Because mortality rates are non-negative, the reported lower endpoint is truncated at zero:

$$L_{i,T+h} = \max\{0, \widehat{M}_{i,T+h|T} - z_{1-\alpha/2} \widehat{\text{SE}}_{i,T+h}\}.$$

3.4.2 Sensitivity of the multi-step forecast to persistence

For fixed $M_{i,T}$ and $\Delta M_{i,T}$, the derivative of the h -step forecast with respect to λ is

$$\frac{\partial \hat{M}_{i,T+h|T}}{\partial \lambda} = \Delta M_{i,T} \sum_{j=1}^h j \lambda^{j-1}.$$

For $\lambda \neq 1$, this derivative may also be written as

$$\frac{\partial \hat{M}_{i,T+h|T}}{\partial \lambda} = \Delta M_{i,T} \frac{1 - (h+1)\lambda^h + h\lambda^{h+1}}{(1-\lambda)^2}.$$

This sensitivity grows with the forecast horizon when λ is close to one, showing why uncertainty in the persistence parameter becomes more consequential for long-horizon forecasts. The empirical intervals used in this study primarily propagate innovation uncertainty; they are therefore interpreted as model-based approximate intervals rather than exact intervals that fully incorporate persistence-parameter uncertainty.

3.5 Benchmark models and evaluation design

AMD was compared with local linear drift, ARIMA(1,1,1), damped Holt, ridge regression, random forest, gradient boosting, and a lag-1 naive benchmark. Random forest and gradient boosting were included as canonical nonlinear machine-learning comparators following Breiman (2001) and Friedman (2001), respectively. The final evaluation used strict rolling one-step forecasts for 2018-2024, producing 112 country-year predictions. Model accuracy was summarized by RMSE, MAE, MAPE, and MASE where available. The development/evaluation separation was retained throughout the analysis; second-stage variants motivated by residual inspection are reported as exploratory robustness checks rather than as a new untouched confirmatory test.

3.6 Monte Carlo design

Monte Carlo validation examined recovery of λ and one-step forecast performance under controlled data-generating mechanisms. Each scenario used 16 countries and $R=1,000$ replications. Baseline persistence values were $\lambda \in \{0.50, 0.80, 0.966\}$, with historical lengths $T \in \{20, 40, 60\}$. Innovations were initially Gaussian with unit variance. For each replication, λ was estimated from pooled first-difference pairs, and bias, RMSE, empirical standard deviation, 95% Wald coverage, and one-step forecast RMSE/MAE were recorded. The local-drift benchmark fixed $\lambda=1$.

The student-t innovations were then scaled to unit variance and country-specific heterogeneous persistence was added, while the mid-scenario of a structural break ($\lambda=0.966$ to $\lambda=0.80$) was added halfway through the history, and 10% randomly missing annual observations were also considered in robustness experiments. Two variants of this form were tested: exponentially weighted AMD (EW-AMD), whereby the adaptations after structural change are faster, and shrinkage AMD (S-AMD), whereby the country specific persistence estimates are pulled toward the pooled estimate.4. Results

4.1 Baseline Monte Carlo parameter recovery

Table 1. Monte Carlo recovery and one-step forecast performance of AMD.

True λ	T	Mean $\hat{\lambda}$	Bias	RMSE($\hat{\lambda}$)	95% coverage	AMD forecast RMSE	Local-drift RMSE
0.500	20	0.4950	-0.0050	0.0415	0.951	1.0005	1.1549
0.500	40	0.4997	-0.0003	0.0323	0.941	1.0000	1.1533
0.500	60	0.4978	-0.0022	0.0260	0.955	1.0040	1.1537
0.800	20	0.7974	-0.0026	0.0291	0.954	1.0009	1.0559
0.800	40	0.7975	-0.0025	0.0222	0.951	0.9914	1.0434
0.800	60	0.7987	-0.0013	0.0183	0.954	1.0026	1.0563
0.966	20	0.9635	-0.0025	0.0175	0.954	1.0139	1.0175
0.966	40	0.9635	-0.0025	0.0115	0.958	0.9951	1.0024
0.966	60	0.9646	-0.0014	0.0092	0.952	0.9988	1.0072

Parameter recovery was strong across all baseline scenarios. Bias was small, parameter RMSE declined as the historical length increased, and empirical 95% coverage remained close to the nominal level. At the empirically relevant $\lambda=0.966$, parameter RMSE declined from 0.0175 at $T=20$ to 0.0092 at $T=60$, while coverage remained between 0.952 and 0.958. AMD's forecast advantage over local drift was largest when true persistence was well below one, which is expected because local drift fixes $\lambda=1$.

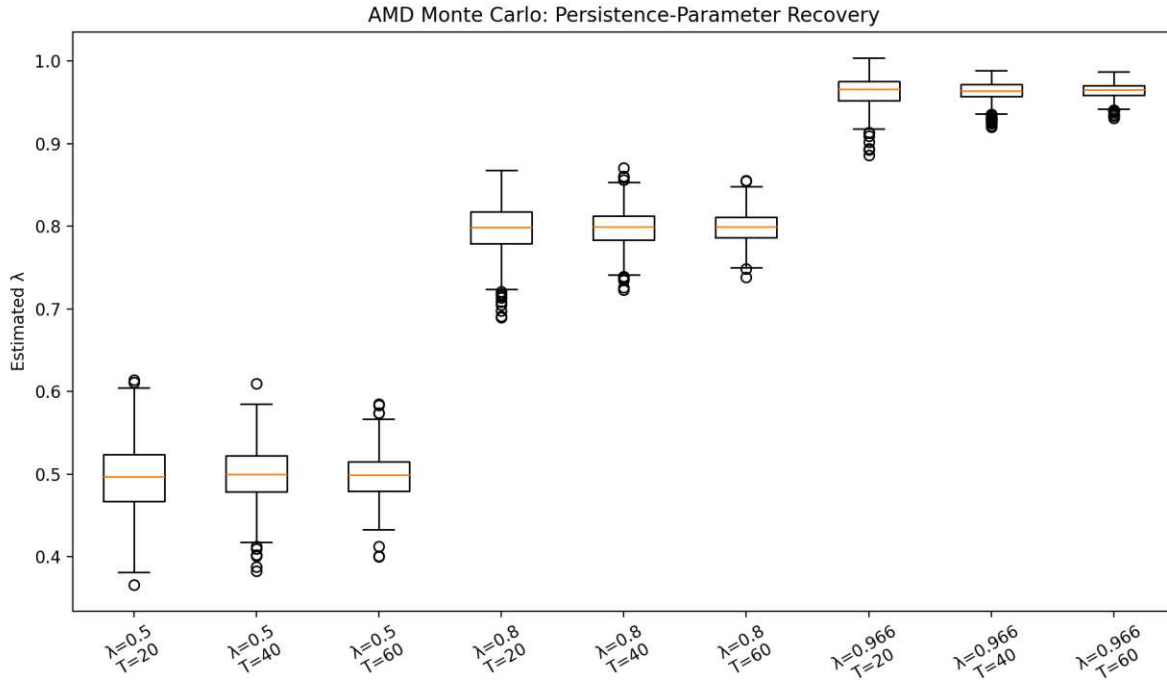


Figure 1. Monte Carlo box plots of the AMD persistence estimator across true persistence values and historical lengths.

The box plots demonstrate the finite-sample behaviour of the AMD persistence estimator, $\hat{\lambda}$, across different true values of λ and historical lengths (Figure 1). The distributions are generally centred close to the true parameter values, while their dispersion decreases as the historical length increases from $T = 20$ to $T = 60$. This provides graphical evidence of improving precision and consistency of the estimator with increasing information.

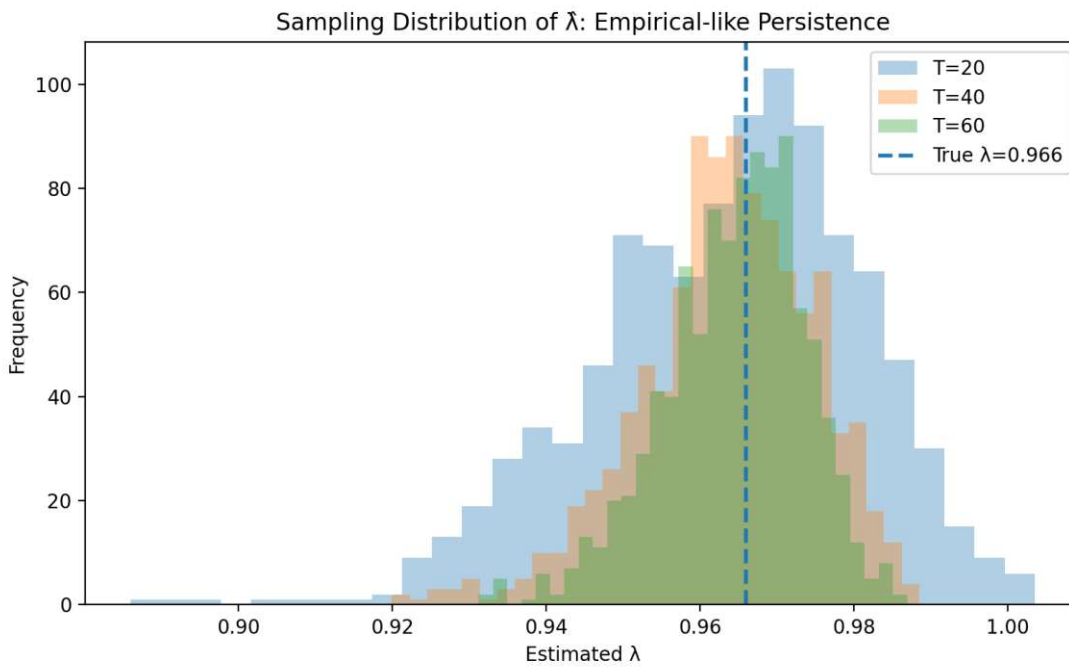


Figure 2. Sampling distributions of $\hat{\lambda}$ for the empirically relevant $\lambda=0.966$ scenario.

The sampling distributions of the $\hat{\lambda}$ are centered around the true value of 0.966, as indicated by the histograms. As the length of history increases, the distribution gets narrower, which means there is less sampling variability (Figure 2). This agrees with the numerical Monte Carlo results that show small bias and decreasing RMSE.

4.2 Robustness experiments

Table 2. Robustness results at T=60 (R=1,000 per scenario).

Scenario	Mean $\hat{\lambda}$	Bias	RMSE($\hat{\lambda}$)	Coverage	AMD RMSE	Local drift RMSE
Normal baseline	0.9645	-0.0015	0.0092	0.951	1.0047	1.0109
Heavy-tailed t(3)	0.9652	-0.0008	0.0085	0.956	1.0722	1.0778
Country-heterogeneous lambda	0.9765	0.0180	0.0219	0.410	1.0064	1.0058
Structural break	0.9243	0.1243	0.1251	0.000	1.0130	1.0475
10% missing observations	0.9643	-0.0017	0.0107	0.944	0.9957	1.0035

The above results indicate that the innovations were heavy tailed and random missingness was moderate, and in this case AMD was not affected. Under t(3) innovations at T=60, mean $\hat{\lambda}$ was 0.9652 and coverage was 0.956. With 10% missing observations the parameter RMSE rose marginally to 0.0107 and the coverage was 0.944. There were two stress conditions that were more of a problem. In the country-specific heterogeneous persistence, the pooled estimator is upwardly biased and a coverage of 0.410 was obtained. The full-history estimator was designed to change "slowly" to this new $\lambda=0.80$ regime, and it did so, yielding mean $\hat{\lambda}\approx 0.924$ and no nominal coverage for the post-break target.

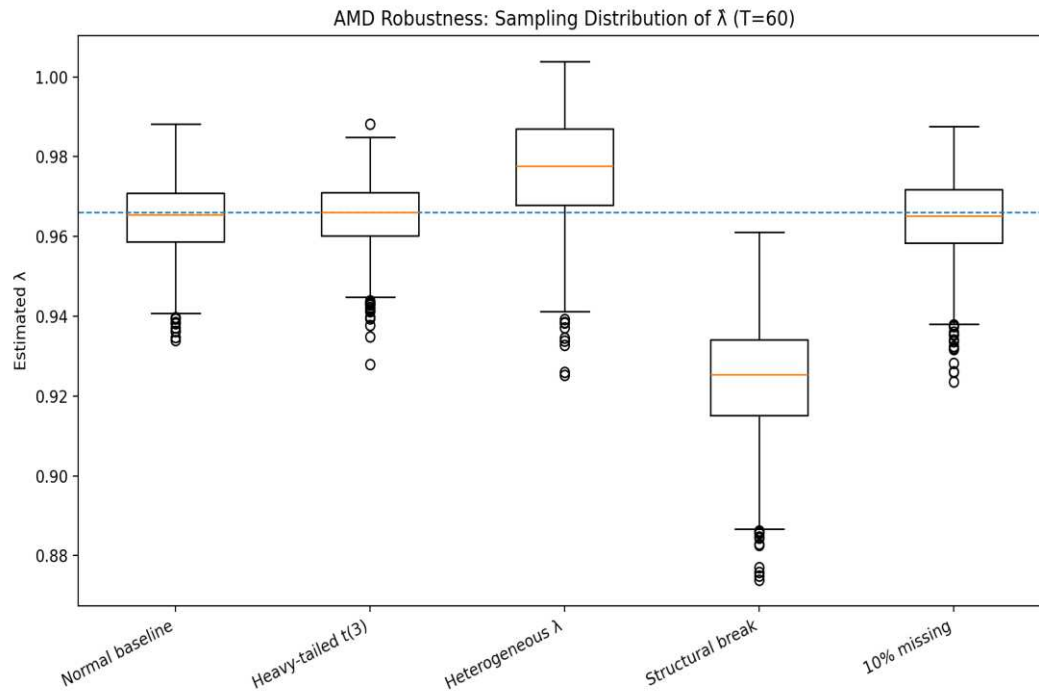


Figure 3. Box plots of $\hat{\lambda}$ under baseline and robustness scenarios (T=60).

Figure 3 compared parameter recovery under normal conditions with heavy-tailed innovations, country heterogeneity, structural change, and missing observations. AMD remains relatively stable under heavy-tailed errors and moderate missingness. Greater departures are evident under country heterogeneity and, particularly, structural breaks, identifying these as the principal challenges to the pooled AMD specification.

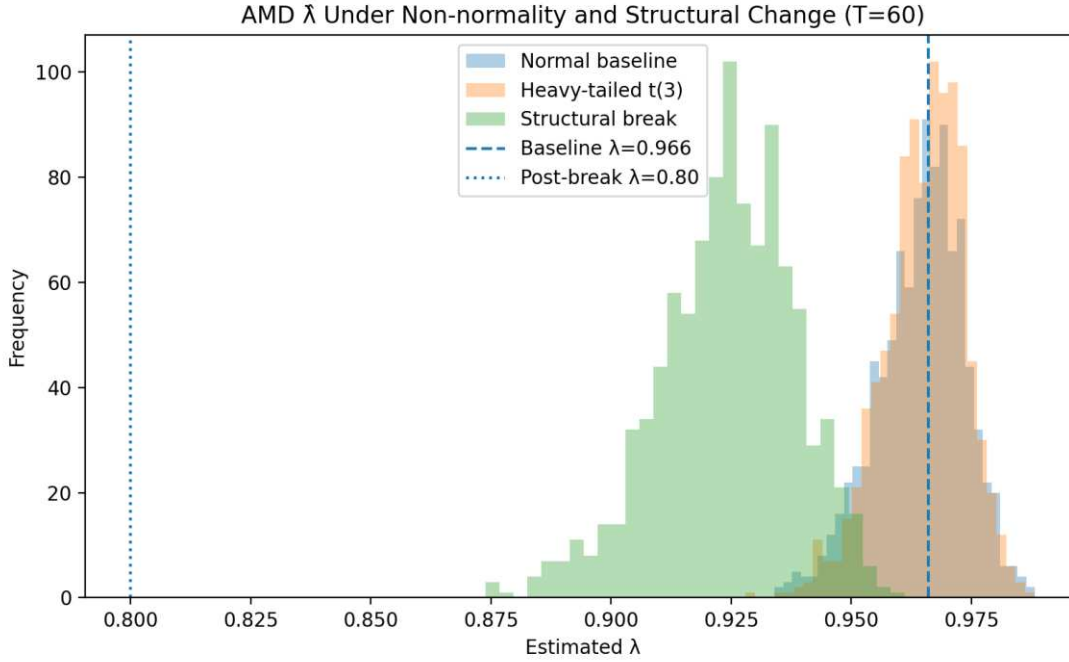


Figure 4. Histograms of $\hat{\lambda}$ under normal, heavy-tailed, and structural-break conditions.

The non-normal innovations have relatively small impact on the parameter recovery, with the normal and heavy-tailed distributions staying close to the baseline persistence value. The structural-break distribution, however, moves away from the post-break parameter value, indicating that an estimator based on the full historical record will not react quickly to a large change in the persistence of the mortality process (Figure 4).

4.3 Targeted extensions for structural change and heterogeneity

Table 3. Targeted AMD extensions under stable, heterogeneous, and structural-break scenarios.

Scenario	Model	Mean estimate	Parameter RMSE	Forecast RMSE	Forecast MAE
Stable	AMD	0.9644	0.0097	0.9979	0.7973
Stable	EW-AMD	0.9639	0.0155	0.9989	0.7982
Stable	S-AMD	0.9461	0.0230	1.0048	0.8042
Country heterogeneity	AMD	0.9771	0.0224	0.9985	0.7985
Country heterogeneity	EW-AMD	0.9805	0.0275	0.9989	0.7983
Country heterogeneity	S-AMD	0.9443	0.0177	0.9960	0.7967
Structural break	AMD	0.9242	0.1251	1.0182	0.8124
Structural break	EW-AMD	0.8233	0.0356	0.9984	0.7963
Structural break	S-AMD	0.8982	0.0996	1.0156	0.8097

EW-AMD successfully mitigated the structural-break stress scenario, by having a mean persistence estimate of 0.8233, which was close to the post-break target of 0.80, compared to 0.9242 for the original AMD. Also, its forecast RMSE decreased from 1.0182 to 0.9984. S-AMD slightly enhanced forecast RMSE in the country heterogeneity model and parameter RMSE. But the targeted extensions did not have a positive impact on the observed rolling forecasts in West Africa. On 2018-2024 data, AMD RMSE was 0.1771, compared with 0.1880 for S-AMD and 0.2343 for EW-AMD. In this sense, the extensions are kept as variants to check the robustness of the main model

diagnostics rather than replacing the main model.

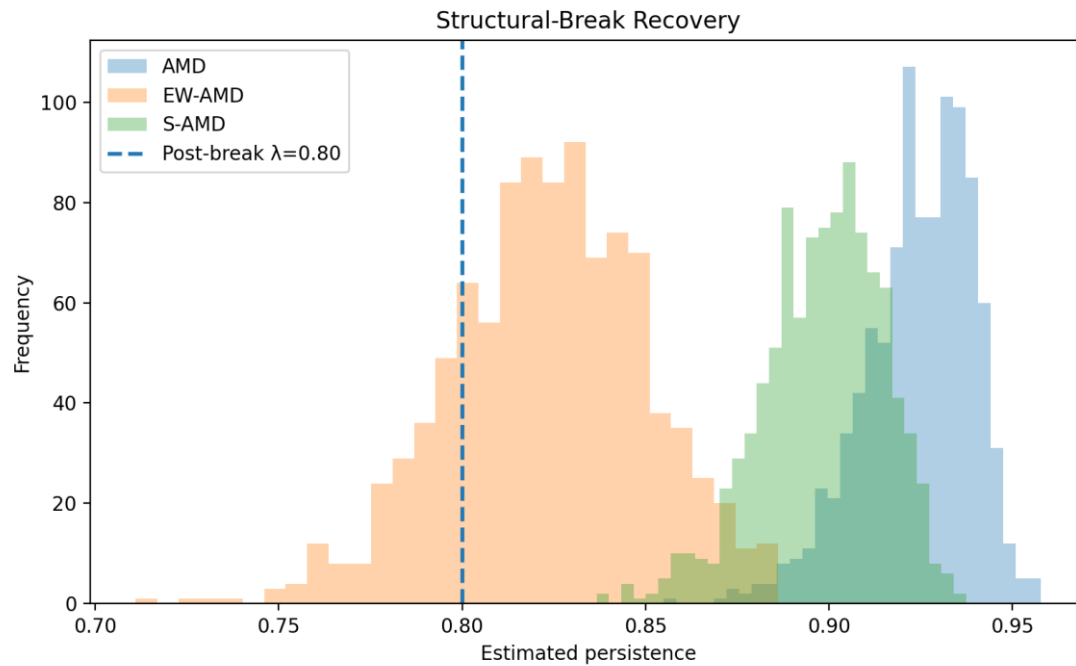


Figure 5. Structural-break sampling distributions for AMD, EW-AMD, and S-AMD.

When there is a structural break, the advantage of EW-AMD is illustrated by the histograms. Compared to the former's persistence regime, the conventional AMD estimator continues to be influenced by the earlier one, whereas EW-AMD gives more weight to recent observations and hence the calculated value climbs to the post-break value significantly. S-AMD offers less improvement since it is developed to deal with the cross-country heterogeneity, not with the temporal structural change (Figure 5).

4.4 Empirical model comparison in West Africa

Table 4. Final rolling-origin predictive comparison, 2018-2024.

Rank	Model	N	RMSE	MAE	MAPE (%)	MASE
1	Adaptive Drift	112	0.1771	0.1301	0.2081	0.0345
2	Local Linear Drift	112	0.1874	0.1474	0.2482	0.0393
3	S-AMD	112	0.1880	0.1405	0.2188	
4	EW-AMD	112	0.2343	0.1854	0.2596	
5	ARIMA(1,1,1)	112	0.3324	0.2076	0.2996	0.0523
6	Damped Holt	112	0.4044	0.2098	0.2875	0.0517
7	Ridge	112	1.2371	0.9461	2.0286	0.2525
8	Random Forest	112	1.7862	1.2976	3.2106	0.4073
9	Naive Lag-1	112	2.2649	2.0178	3.1687	0.5196
10	Gradient Boosting	112	2.3154	1.5792	4.9520	0.5421

AMD ranked first on the final rolling evaluation, with RMSE 0.1771, MAE 0.1301, MAPE 0.2081%, and MASE 0.0345. Compared to the local linear drift, the RMSE, MAE and MAPE were lowered by about 5.5%, 11.7% and 16.2% respectively by the use of AMD. The other two models; ARIMA(1,1,1) and damped Holt, failed to perform as well, and the machine-learning models performed significantly poorer for the 1-year-ahead prediction. This should not be interpreted as an indication that machine learning is worse in general, but rather as an indication that these annual U5MR series are extremely smooth, and that the dynamics in recent mortality data have a very large

amount of information that could be used for short-horizon forecasting.

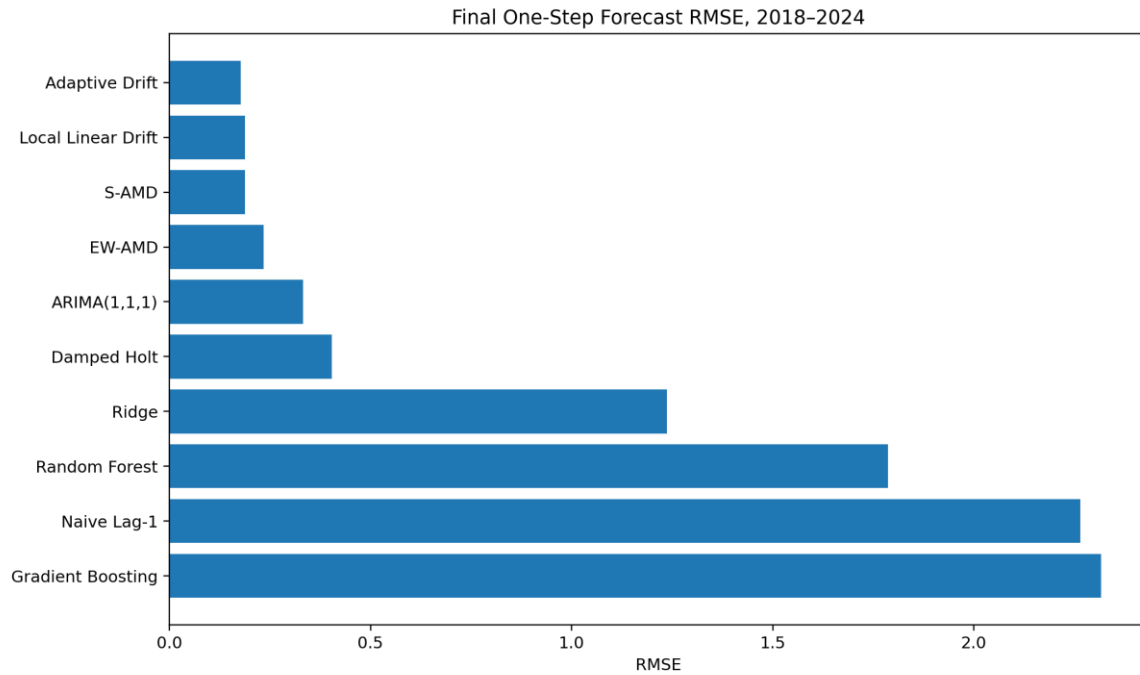


Figure 6. Final one-step forecast RMSE comparison across predictive models, 2018-2024.

The RMSE comparison shows that the model of AMD is the best overall model for short horizon forecasting, closely followed by Local Linear Drift and S-AMD (Figure 6). The machine-learning models do much worse and ARIMA and Holt tend to give large forecast errors. This implies that the smooth and well-preserved annual mortality patterns support parsimonious models that explicitly make use of the recent temporal variation.

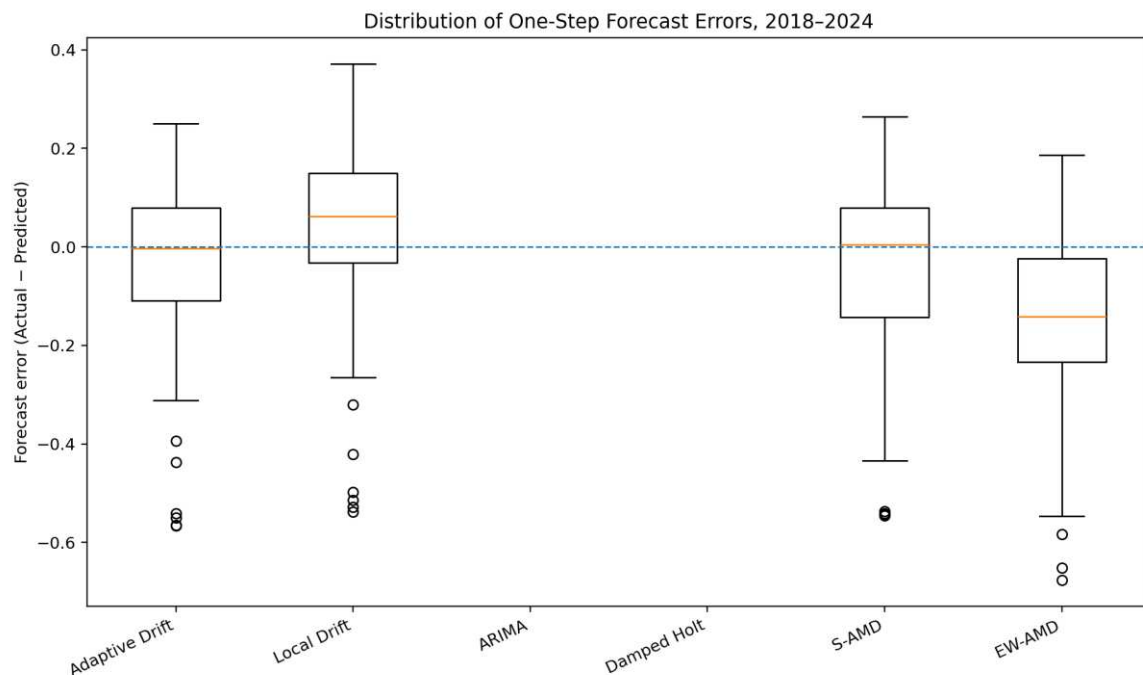


Figure 7. Observed rolling forecast errors for selected alternative models shown as box plots.

The forecast-error box plots are used to complement the results of aggregate RMSEs, to represent the distribution of forecast error. AMD errors are very close to zero and have relatively small standard deviation, which shows a small bias and high predictive accuracy (Figure 7). Several models have a higher frequency of extreme errors and wider distributions, which is a reason for their lower aggregate forecasting ability.

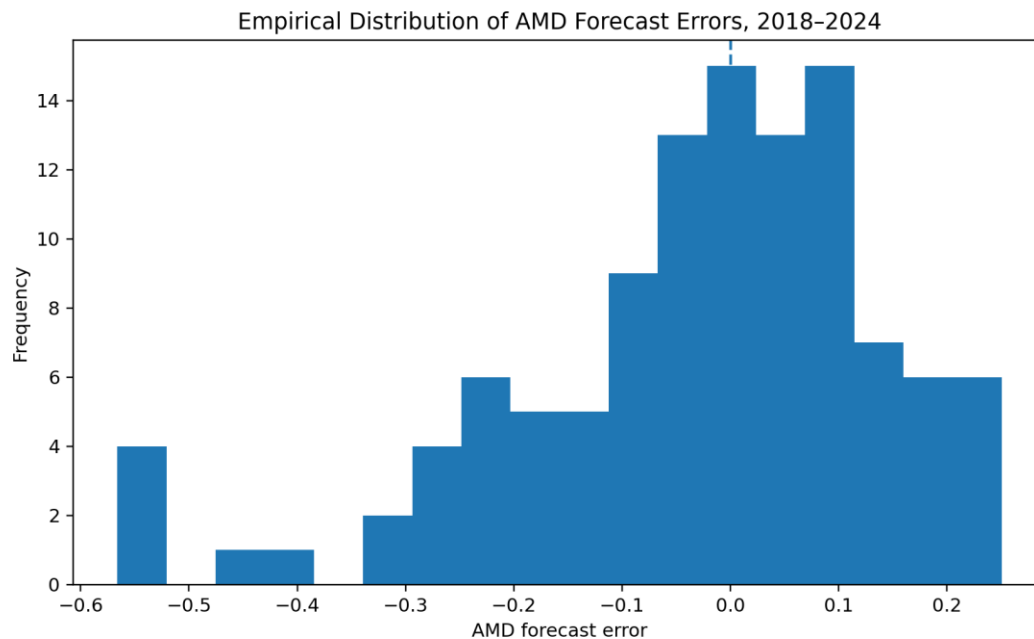


Figure 8. Empirical rolling forecast errors, 2018-24, histogram

The histogram indicates that the forecast errors from AMD are well clustered around zero, and are similar to the very small MAE and RMSE values obtained from the rolling validation (Figure 8). There is still some asymmetry and tail behaviour though, which suggests that not all short-term fluctuations in mortality are captured by AMD. This observation reinforces the use of prediction intervals and the need to be aware of the limitations of using a point prediction as an estimate of the future mortality rate.

4.5 Country-level accuracy and heterogeneity

There was no country-level superiority for AMD that was implied by the pooled superiority. AMD reduced RMSE compared to local linear drift in 11 of the 16 countries, in particular in Sierra Leone, The Gambia, Mali, Burkina Faso, Togo, Ghana and Guinea-Bissau. Local drift was superior in Benin, Guinea, Niger, Senegal and marginally in Nigeria. This is a point of heterogeneity that is crucial: AMD is the recommended regional pooled short-horizon model, not the model for all countries.

Table 5. Accuracy of country-specific AMD forecasts 2018-2024

Country	RMSE	MAE	MAPE (%)
Mauritania	0.0539	0.0429	0.1083
Togo	0.0542	0.0466	0.0764
The Gambia	0.0756	0.0630	0.1311
Ghana	0.0910	0.0737	0.1907
Burkina Faso	0.0982	0.0865	0.1044
Guinea-Bissau	0.1051	0.0829	0.1127
Sierra Leone	0.1088	0.0980	0.0958
Cabo Verde	0.1125	0.1002	0.7488
Mali	0.1287	0.1043	0.1294
Benin	0.1379	0.1115	0.1333
Côte d'Ivoire	0.1487	0.1148	0.1571
Senegal	0.1890	0.1738	0.4029
Guinea	0.2344	0.2176	0.2160
Liberia	0.2510	0.2235	0.2546
Nigeria	0.3231	0.2673	0.2290
Niger	0.3476	0.2753	0.2393

4.6 Mortality levels, historical progress, and 2030 forecasts

Table 6. Country comparison: observed mortality, historical reduction, and AMD 2030 forecasts.

Rank	Country	2000	2024	Reduction 2000-24 (%)	2030 forecast	Forecast reduction (%)
1	Nigeria	177.28	115.65	34.8	109.42	5.4
2	Niger	227.85	110.72	51.4	96.07	13.2
3	Guinea	165.68	92.14	44.4	77.18	16.2
4	Sierra Leone	222.84	90.50	59.4	73.04	19.3
5	Liberia	194.86	86.39	55.7	81.56	5.6
6	Burkina Faso	176.01	74.89	57.4	61.72	17.6
7	Benin	134.97	74.75	44.6	61.77	17.4
8	Mali	189.93	72.98	61.6	55.77	23.6
9	Guinea-Bissau	174.83	67.34	61.5	54.75	18.7
10	Côte d'Ivoire	140.44	64.47	54.1	53.42	17.1
11	Togo	119.24	56.13	52.9	45.78	18.4
12	The Gambia	112.82	42.41	62.4	34.75	18.1
13	Mauritania	98.01	36.91	62.3	30.40	17.6
14	Senegal	128.87	36.53	71.7	27.40	25.0
15	Ghana	99.94	35.88	64.1	28.90	19.4
16	Cabo Verde	36.01	11.14	69.1	8.70	21.9

The 2024 data show very large regional disparities. Nigeria had the highest U5MR at 115.65 per 1,000 live births, followed by Niger (110.72), Guinea (92.14), Sierra Leone (90.50), and Liberia (86.39). The lowest value was in Cabo Verde (11.14). There was also uneven progress in the past: U5MR in Senegal declined by 71.7 per cent from 2000 to 2024, while U5MR in Cabo Verde declined by 69.1 per cent, Ghana declined by 64.1 per cent, The Gambia declined by 62.4 per cent, and Mauritania declined by 62.3 per cent; U5MR in Nigeria declined by 34.8 per cent.

With the above-estimated $\hat{\lambda} = 0.966459$ obtained from the available data prior to 2024, AMD predicts further decreases until 2030. The forecasts for the corresponding countries are Nigeria 109.42, Niger 96.07, Liberia 81.56, Guinea 77.18 and Sierra Leone 73.04. The lowest possible projection is Cabo Verde at 8.70. These are statistical extrapolations based on recent mortality trends and do not provide causal predictions of policy impacts or any

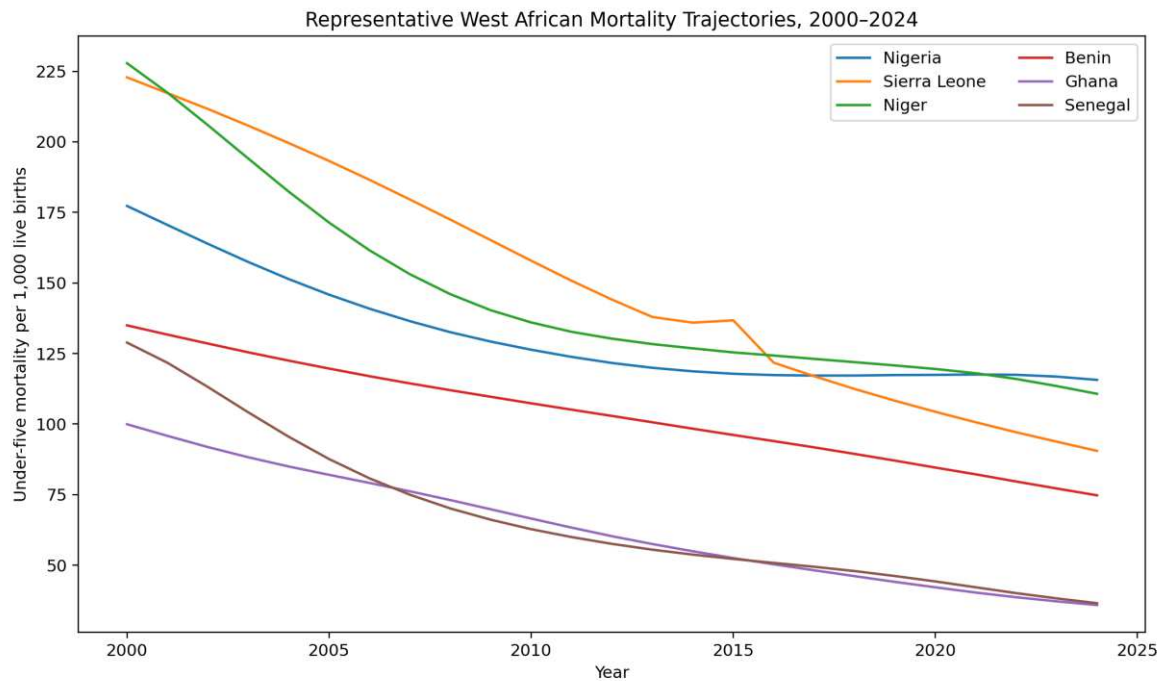


Figure 9. Representative West African U5MR trajectories, 2000-2024.

The historical trajectories reveal a broad downward movement in under-five mortality across the selected West African countries, although mortality levels and rates of improvement differ substantially. Some countries exhibit steep and sustained reductions, whereas others remain at comparatively high mortality levels. These differences demonstrate considerable regional heterogeneity despite the common long-term declining pattern (Figure 9).

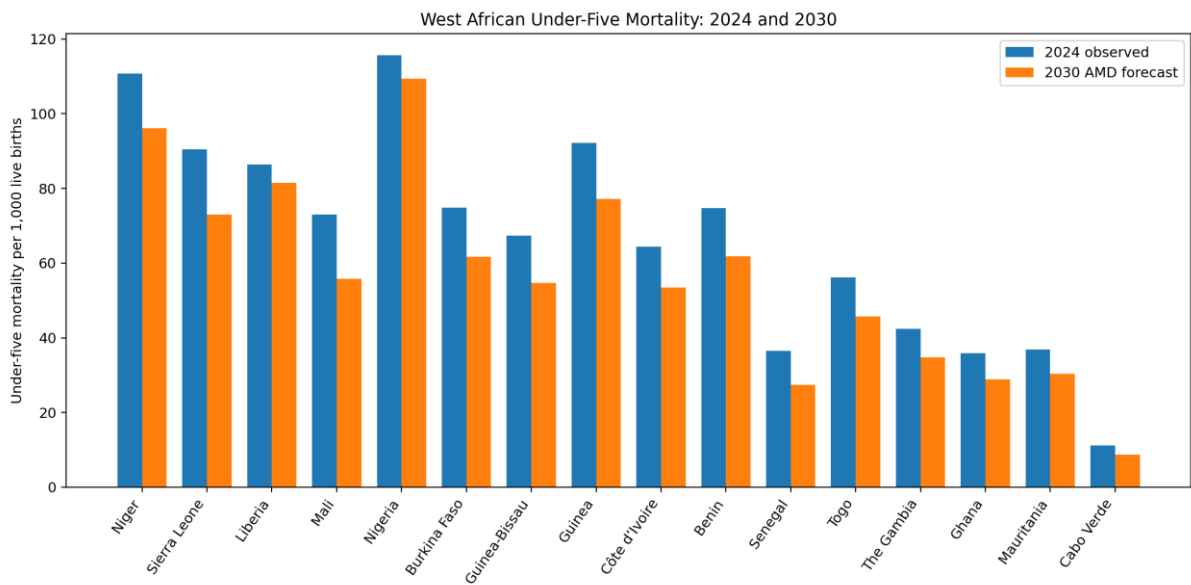


Figure 10. Projections of the U5MR and AMD for 2030 in the 16 countries.

The comparison shows that under-five mortality in the region will decline further in the future, although there will be large differences between countries (Figure 10). Countries with higher mortality rates in 2024 tend to stay higher in 2030 forecasts, and countries which are already relatively low will continue to be relatively low in 2030. The forecasts thus indicate that the region would continue to make progress, but also highlight the need for country-specific interventions in countries where mortality is still high.

Table 7. AMD forecasts and approximate prediction intervals for 2030.

Country	Forecast	80% lower	80% upper	95% lower	95% upper
Benin	61.77	59.77	63.78	58.70	64.84
Burkina Faso	61.72	59.72	63.73	58.66	64.79
Cabo Verde	8.70	6.69	10.71	5.63	11.77
Côte d'Ivoire	53.42	51.41	55.43	50.35	56.49
The Gambia	34.75	32.75	36.76	31.68	37.82
Ghana	28.90	26.89	30.91	25.83	31.97
Guinea	77.18	75.18	79.19	74.11	80.25
Guinea-Bissau	54.75	52.74	56.75	51.68	57.82
Liberia	81.56	79.55	83.56	78.49	84.62
Mali	55.77	53.76	57.77	52.70	58.84
Mauritania	30.40	28.40	32.41	27.34	33.47
Niger	96.07	94.06	98.07	93.00	99.14
Nigeria	109.42	107.42	111.43	106.36	112.49
Senegal	27.40	25.39	29.40	24.33	30.47
Sierra Leone	73.04	71.04	75.05	69.98	76.11
Togo	45.78	43.77	47.78	42.71	48.84

5. Discussion and Conclusion

5.1 Discussion

The main empirical result is that a relatively minor change to the local drift substantially enhances the regional short-horizon forecast. Local drift is based on the assumption that the most recent annual mortality change is maintained entirely, while AMD estimates degree of maintenance. The regional estimate is approximately 0.966 indicating that nearly all of the previous change is carried forward, albeit with a bit of a dampening effect. This mild shrinkage is enough to lower the pooled final-test error as ordinary drift over-predicts the annual decline a bit too sharply.

The results are consistent with the wider forecasting literature in two ways: The results are consistent with the wider forecasting literature, in two ways: Firstly, they confirm the importance of using a strict out-of-sample comparison process highlighted by previous Nigeria and West Africa forecasts. Second, they demonstrate the importance of the match between model complexity and the information structure of the series. The annual national U5MR trends are smooth and positively persistent; this means that complex pooled machine-learning models do not automatically perform better in forecasting U5MRs one year ahead. This is not in contradiction with individual-level studies of ML, which use several child, household and community predictors and foresee a different prediction target.

The Monte Carlo experiments give insight about the reliability of AMD. The pooled estimator was seen to accurately recover λ and yield approximately nominal interval coverage under stable persistence and heavy-tailed innovations and moderate missingness. The model's reliability is reduced in situations where national persistence parameters are significantly different, or where a structural break alters the dynamics. These restrictions aren't just cosmetic. Pooled regional parameter is a common short run tendency, and if the national dynamics turns out to be quite different, the pooled estimator is not a country specific truth, but a compromise.

Such weaknesses are explained by the targeted extensions. EW-AMD improves recovery of parameters when dealing with heterogeneity and adapts much faster after a simulated break. However, both extensions degrade the forecast for West Africa in 2018-2024. This is helpful negative result: don't make model more complex because of a diagnostic hint for possible structure. The extra building should yield sustainable future benefits. According to the data available, parsimony is favoured by out-of-sample performance.

The country comparison shows that no single homogeneous child-survival environment exists in West Africa. Nigeria and Niger continue to have very high mortality rates and Cabo Verde is very low. Senegal, Ghana, The Gambia and Mauritania had large percentage reductions since 2000, whilst progress Nigeria was much slower. The

forecasts thus call for differentiated policy attention, with high-burden countries needing to continue scaling back mortality and fasten up the rate of progress, and lower-burden countries ensuring that their progress was not lost and that the growing share of residual and neonatal causes of death was addressed.

AMD is not a cause-based policy model. It does not name any intervention that would lead to a reduction in mortality and it does not explicitly include future vaccination coverage, conflict, epidemics, nutrition, macroeconomic shocks, health financing and health-system reforms. It has a more narrow policy purpose: to offer a transparent short horizon test of whether the observed pattern of mortality is continuing, slowing or diverging from the recently observed regional pattern. This type of benchmarking can be used in addition to, but not in lieu of, epidemiological and causal analyses.

5.2 Conclusion

This study developed and evaluated the Adaptive Mortality Drift model for annual under-five mortality forecasting in West Africa. AMD estimates the persistence of recent annual mortality changes rather than fixing full persistence as in ordinary local drift. In strict 2018-2024 rolling evaluation across 16 countries, AMD achieved the lowest pooled RMSE, MAE, MAPE, and MASE among the evaluated statistical and machine-learning benchmarks. Monte Carlo experiments supported accurate parameter recovery under stable conditions and showed robustness to heavy-tailed innovations and moderate missingness, while also identifying important weaknesses under structural breaks and country heterogeneity. Targeted variants addressed those simulated weaknesses but did not improve the observed West African forecasts, supporting retention of the simpler AMD model.

The empirical application documents large and persistent cross-country inequalities in child survival. Although AMD projects continued mortality decline through 2030, several high-burden countries remain far above low-mortality regional peers. The model should therefore be used as a transparent short-horizon forecasting benchmark and monitoring tool, alongside substantive public-health models that explain the determinants and causal mechanisms of mortality change.

Declarations

Ethics approval and consent to participate: Not applicable. This study used publicly available, country-level secondary data from the United Nations Inter-Agency Group for Child Mortality Estimation and the World Health Organization. It did not involve direct participation by human subjects, identifiable personal information, clinical interventions, or primary data collection. Therefore, ethical approval and informed consent were not required.

Consent for publication: Not applicable. The manuscript contains no identifiable personal data, images, or individual case information requiring consent for publication.

Availability of data and materials: The under-five mortality data analysed in this study are publicly available through the United Nations Inter-Agency Group for Child Mortality Estimation and the World Health Organization Global Health Observatory. All data used in the analysis were obtained from publicly accessible secondary sources.

Competing interests: The author declares that there are no competing interests.

Funding:

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Author's contribution: Mumini Idowu Adarabioyo conceived and designed the study, developed the methodology, acquired and analysed the data, conducted the simulations and forecasting experiments, interpreted the results, prepared the figures and tables, drafted and critically revised the manuscript, and approved the final version for publication. The author accepts full responsibility for the integrity and accuracy of the work.

Acknowledgement: The author acknowledges the United Nations Inter-Agency Group for Child Mortality Estimation and the World Health Organization for making the under-five mortality estimates used in this study publicly available.

References

- Adama, Z. K., Mettle, F. O., Baiden, B. M., & Bii, N. K. (2025). Forecasting progress: Analyzing the trajectory of under-five child mortality for Ghana, Niger, Nigeria, and Sierra Leone towards SDG3 using ARIMA time series model. *BMC Public Health*, 25, 1607. <https://doi.org/10.1186/s12889-025-22869-z>
- Adeyinka, D. A., & Muhajarine, N. (2020). Time series prediction of under-five mortality rates for Nigeria: Comparative analysis of artificial neural networks, Holt-Winters exponential smoothing and autoregressive integrated moving average models. *BMC Medical Research Methodology*, 20, 292. <https://doi.org/10.1186/s12874-020-01159-9>
- Asnake, A. A., Gebrehana, A. K., Asmare, Z. A., Seifu, B. L., Asebe, H. A., Bezie, M. M., Tsega, S. S., Negussie, Y. M., Melkam, M., & Fente, B. M. (2026). Prediction of under-five mortality using supervised machine learning algorithms in the 23 sub-Saharan African countries. *Scientific Reports*, 16, 9131. <https://doi.org/10.1038/s41598-026-40401-0>
- Aue, A., & Horváth, L. (2013). Structural breaks in time series. *Journal of Time Series Analysis*, 34(1), 1-16. <https://doi.org/10.1111/j.1467-9892.2012.00819.x>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5-32. <https://doi.org/10.1023/A:1010933404324>

- Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5), 1189-1232. <https://doi.org/10.1214/aos/1013203451>
- Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4), 679-688. <https://doi.org/10.1016/j.ijforecast.2006.03.001>
- Mejía-Guevara, I., Zuo, W., Bendavid, E., Li, N., & Tuljapurkar, S. (2019). Age distribution, trends, and forecasts of under-5 mortality in 31 sub-Saharan African countries: A modeling study. *PLOS Medicine*, 16(3), e1002757. <https://doi.org/10.1371/journal.pmed.1002757>
- Samuel, O., Zewotir, T., & North, D. (2024). Application of machine learning methods for predicting under-five mortality: Analysis of Nigerian demographic health survey 2018 dataset. *BMC Medical Informatics and Decision Making*, 24, 86. <https://doi.org/10.1186/s12911-024-02476-5>
- United Nations Inter-agency Group for Child Mortality Estimation. (2026). Levels & trends in child mortality 2025. UNICEF.
- World Health Organization. (n.d.). Under-five mortality rate (per 1000 live births) (SDG 3.2.1). Global Health Observatory.