

Forecasting Rate of Decline in Infant Mortality in South Asia Using Random Walk Approximation

Tapan Kumar Chakrabarty*

Department of Statistics, North Eastern Hill University, Shillong – 793022, Meghalaya, India

Abstract: The Millennium Development Goal 4 (MDG 4) of United Nations had set the target of reducing high rates of under-five and infant mortality (IMR) by two thirds to be reached by 2015 using 1990 as the benchmark year. By the availability of time series data on IMR from United Nations Inter-agency Group for Child Mortality Estimation (UN IGME, 2012), led by UNICEF, WHO, the World Bank and United Nations, it has become possible to track the rate of progress towards this goal. Using the UN IGME 2012 data for all the South Asian Countries, I have considered three specific issues in this article. (1) How does the South Asian Countries fair in reducing the IMR towards this MDG target? Although the time series data exhibit declining trends for all the countries in South Asia, to what extent such trends are attributed by their average annual progress trajectory over the period for which data are available? (2) Whether deterministic or stochastic trend can attribute the IMR decline in South Asian countries and what alternative time series models be used to forecast the decline in Infant Mortality? Can we find a serviceable representative model for the entire region? (3) In case, a satisfactory representative model for the entire region exists, how do we assess the forecast accuracy for this model and quantify the propagation of forecast error?

Keywords: Infant mortality, ARIMA model, random walk, MDGs, demographic forecast, unit root test.

1. INFANT AND CHILD MORTALITY IN SOUTH ASIA

It is well established that, if development is to be measured using a comprehensive and inclusive assessment, it is necessary to appraise human as well as economic progress. From UNICEF's point of view, there is a need for an agreed method of measuring the level of child well-being and its rate of change.

The United Nations Millennium Development Goals (MDGs) set the stage for developing countries not only to reduce extreme poverty, but also the problems that accompany it, such as high rates of infant, child and maternal mortality [1-3].

Each goal has a specific target level for progress, such as halving poverty or reducing infant mortality rates by two thirds. All goals are to be reached by 2015, using 1990 as the benchmark year. By setting a time frame and specific levels of reductions for a variety of indicators, progress towards the goals is measurable, if data on indicators is available. Tracking progress is an essential step towards meeting the goals, as problem areas can be identified only through monitoring and evaluation, and interventions and strategies can then be developed to target them.

South Asia as a whole, though, likely to halve poverty levels by 2015, situation is not as promising in

other areas, such as reduction of high rates of infant, child and maternal mortality [4, 5]. World has witnessed a substantial progress in Millennium Development Goal 4 (MDG 4) during the period with reference to the decline in the number of under-five deaths worldwide. Since 1990 the global under-five mortality rate has dropped 41 percent—from 87 (85, 89) deaths per 1,000 live births in 1990 to 51 (51, 55) in 2011 [6, 7]. Eastern Asia, Northern Africa, Latin America and the Caribbean, South-eastern Asia and Western Asia have reduced their under-five mortality rate by more than 50 percent. The annual rate of reduction in under-five mortality has accelerated—from 1.8 (1.7, 2.1) percent a year over 1990–2000 to 3.2 (2.5, 3.2) percent over 2000–2011—but remains insufficient to reach MDG 4, particularly in Oceania, Sub-Saharan Africa, Caucasus and Central Asia, and Southern Asia [6, 8].

The highest rates of child mortality are still in Sub-Saharan Africa—where 1 in 9 children dies before age five, more than 16 times the average for developed regions (1 in 152)—and Southern Asia (1 in 16) [6, 9]. As under-five mortality rates have fallen more sharply elsewhere, the disparity between these two regions and the rest of the world has grown. Under-five deaths are increasingly concentrated in Sub-Saharan Africa and Southern Asia, while the share in the rest of the world dropped from 31 percent in 1990 to 17 percent in 2011 [6]. Table 1 provided below show the rate of progress towards MDG target for the countries belonging to South Asian Region.

*Address correspondence to this author at the Department of Statistics, North Eastern Hill University, NEHU Permanent Campus, Shillong 793022, Meghalaya, India; Tel: (364)2722905; Fax: (364)2550076; E-mail: tapankumarchakrabarty@gmail.com

Table 1: Rate of Progress Towards MDG 2015 Target (IMR) for South Asia

Country	Infant Mortality Rate				Average Annual Rate of Reduction(%)				Reduction Since 1990(%)	Reduction Since 2000(%)	Extrapolated Value (2015)	MDG Target Value (2015)
	1970	1990	2000	2010	1970-1990	1990-2000	2000-2010	1990-2010				
Afghanistan	207.2	129.4	94.9	74.5	2.31	3.07	2.4	2.74	42.4	21.5	60.8	43
Bangladesh	150.6	96.5	62	38.6	2.1	4.22	4.65	4.41	60.0	37.7	16.6	32
Bhutan	191	96.3	65	43.6	3.38	3.83	3.91	3.87	54.7	32.9	24.5	32
India	127.6	81	64.2	48.6	2.25	2.26	2.73	2.49	40.0	24.3	36.2	27
Maldives	161.1	75.7	41.3	10.6	3.69	5.6	12.5	8.99	86.0	74.3	**	25
Nepal	158.2	93.5	61.8	40.6	2.54	4.05	4.12	4.08	56.6	34.3	20.2	31
Pakistan	136	94.6	75.9	60.4	1.76	2.14	2.26	2.2	36.2	20.4	49.4	32
Sri Lanka	56.6	24.2	16.4	10.8	4.1	3.79	4.09	3.94	55.4	34.1	*	8
South Asia	132.1	84.8	65.9	49.7	2.18	2.44	2.78	2.6	41.4	24.6	30	28

Achieved or on target	On the track	Out of track
-----------------------	--------------	--------------

Source: UN IGME 2012 Estimates, www.childmortality.org

2. ANALYZING THE RATE OF PROGRESS IN INFANT MORTALITY REDUCTION IN SOUTH ASIA

Let M_t denote the value of the observed infant mortality rate in the year t . Annual rate of reduction (ARR) from year $t-1$ to year t is defined as

$$ARR_t = \frac{M_{t-1} - M_t}{M_{t-1}} * 100$$

so that Average ARR(AARR) over a period can be obtained as just the arithmetic mean of these ARR_t . For instance, AARR over the period 1970-1990 can be obtained as

$$AARR_{1970-90} = \frac{\sum_{t=1971}^{1990} ARR_t}{20}$$

Table 1 calculates some important statistics including AARR over various periods. UN IGME estimates of infant mortality show that South Asia is well short of its MDG target [1]. The countries for which the rate of progress in reducing IMR decline far behind the required are India, Pakistan and Afghanistan. However, the amount of deficit (25.3%) till 2011 for the entire region is quite close to that of India (26.7%) as shown in the Figure 1a. Figure 1b displays the trajectory of decadal average of annual rate of reduction in IMR during 1970-2010. Here too, the average line for South Asia closely matches to that of India. Sharp decline are observed for Bangladesh and Nepal, closely followed by relatively flat lines for Sri Lanka and Bhutan which have started with a very much higher rates of reduction during 1970-1980. Pakistan remains far below the regional average throughout the

period under consideration matching Afghanistan at the start and end of the period. Summarily, out of 8 countries in the region, Maldives has already achieved its target. A linear extrapolation shows that at 1990-2010 average annual rate of reduction the rate of progress is on the track for countries such as Bangladesh, Bhutan and Nepal. Whereas, Afghanistan and Pakistan are completely out of the track, India's position is far behind its set target. Overall, the time series of IMR in the region exhibit variant degrees of departures from one another with respect to their rates of decline.

3. MODELING RATE OF DECLINE IN INFANT MORTALITY

In this section, we consider logarithm of decline rate of infant mortality as a time series and fit alternative parametric time series models. Fitted models thereafter shall be used for forecasting the values till 2015, say, and the results will be analyzed to assess the shortfall in achieving MDG target. The data for this purpose have been downloaded from the site of <http://www.childmortality.org> [10]. We have used data for the South Asian countries, Afghanistan, Bangladesh, Bhutan, India, Maldives, Nepal, Pakistan and Sri Lanka over the period 1960-2010.

As earlier, if M_t denote the value of the observed infant mortality rate in the year t , then,

$$Z_t = \log_e \left(\frac{M_{t-1}}{M_t} \right)$$

is defined as the rate of decline in infant mortality from the year $t-1$ to t . We took Z_t as the variable to be

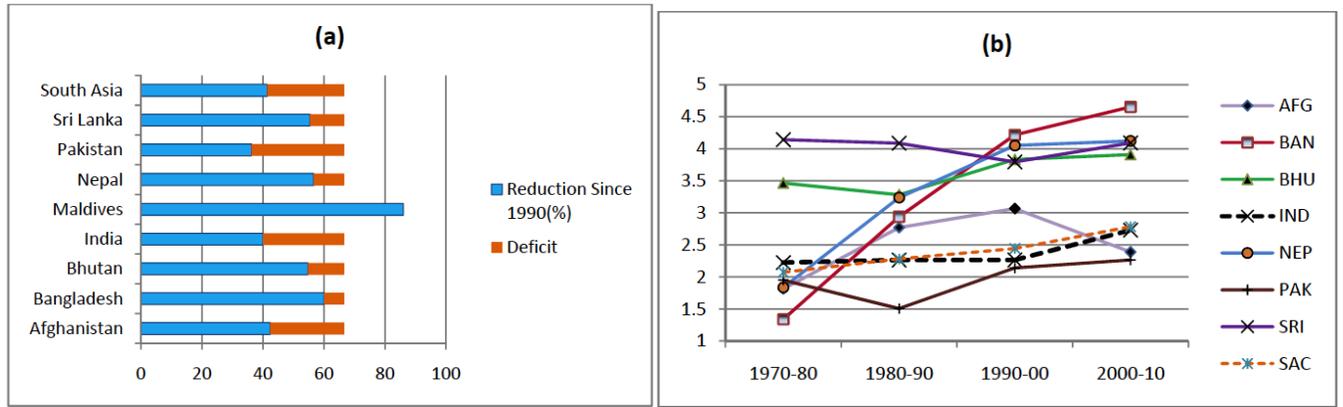


Figure 1: Rate of Progress in IMR Reduction in South Asia. (a) Deficit as on 2011 for MDG target in percentage. (b) Decadal Average of Annual Rate of Reduction in IMR.

analyzed. This guarantees the positivity of all results, but, more importantly, it transforms changes into relative scale, which seems reasonable given the large variation in the level of the infant mortality rates. The trend in the rate of decline in infant mortality since 1970 is depicted in the Figure 2 below for all the countries of South Asia except the Maldives.

It can be seen that the trend for the entire region South Asian Countries as a whole is much similar to that of India over the whole period, in other words, the trend of the entire region is being driven mostly by India. The post 1990 episode has witnessed a much sharper decline for countries like Bangladesh, Bhutan and Nepal. In contrast, the data support that the pace of decline in infant mortality rates is slowing down in India and Sri Lanka as well. The trend line for Pakistan lies much below the other countries with no clear increase or decrease in the pace.

Many vital processes appear non-stationary [11, 12]. We have analysed the logarithmic rate of decline

Z_t , over 1960-2010 for all the South Asian countries (SAC). Sample autocorrelation function (ACF) and sample partial autocorrelation function (PACF) were used to assess stationarity status of all the series. Persistent patterns of sample ACF for all the countries were evident. For the sake of illustrations, we provide the plots of autocorrelations and partial autocorrelations for India and for the SAC as a region in Figure 3.

The autocorrelations did not approach zero, as they should for a linear process. Initial impressions were further confirmed using the augmented Dickey-Fuller test of unit root [13, 14]. The results are presented in Table 2 which confirm that logarithmic rate of decline is non-stationary for all the South Asian countries at all common significance levels. We then looked for the smallest d such that the d^{th} difference both looked stationary in a plot and had an ACF that did approach zero fairly quickly. In order to remove the persistent pattern from the sample ACF, we had to take the first differences of the series Z_t . The sample first-

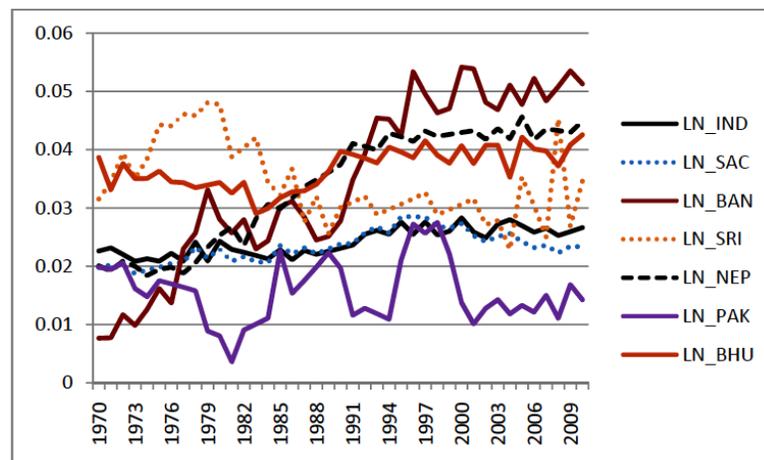


Figure 2: The trend of decline in infant mortality rate of South Asian Countries, 1970-2010.

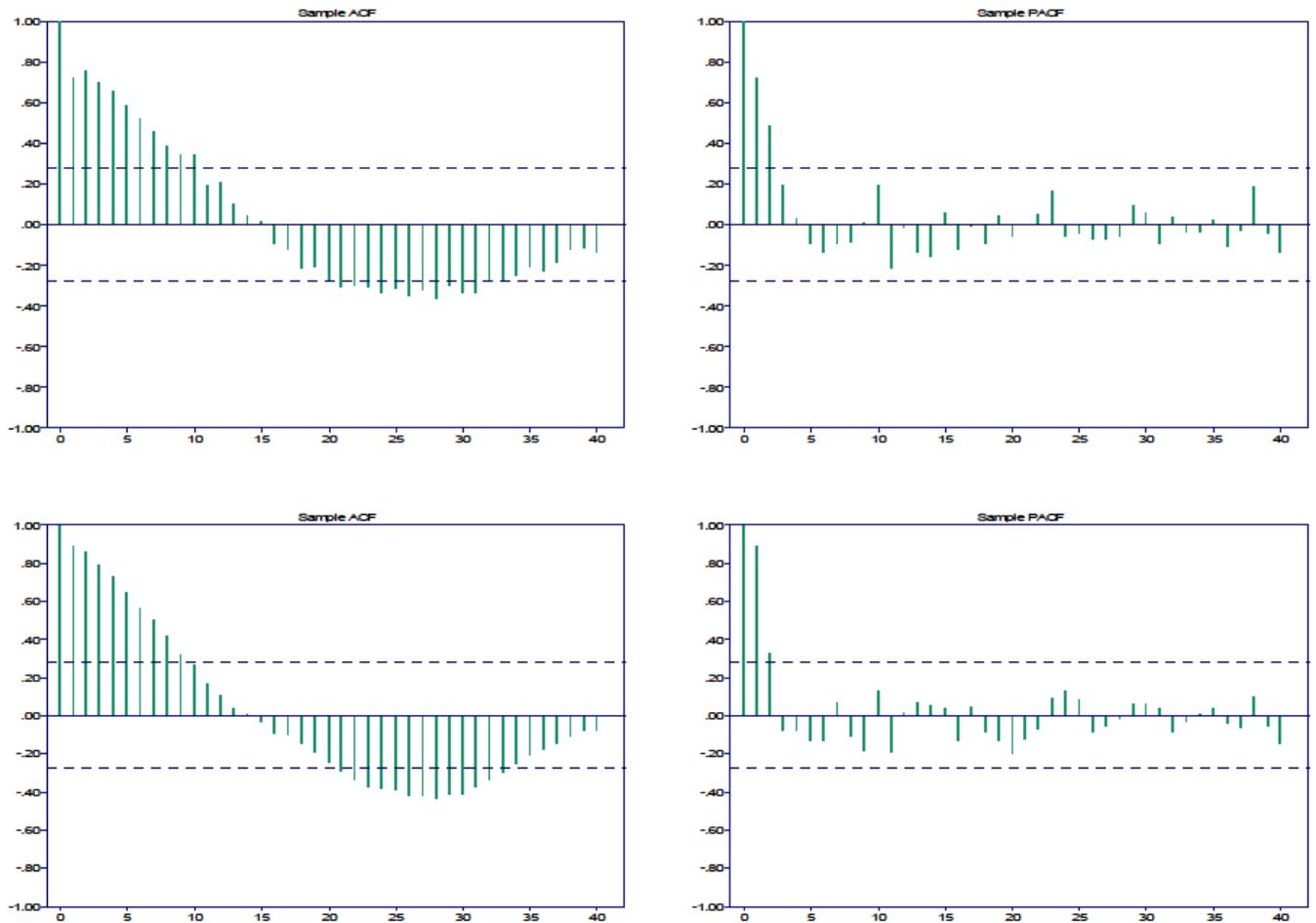


Figure 3: Top panel: sample ACF and PACF for India. Bottom Panel: sample ACF and PACF for SAC.

autocorrelations r_1 of the first differences of the series Z_t varied from -0.57 to -0.04 and for second autocorrelations r_2 , we got, $-0.12 \leq r_2 \leq 0.21$. It was found that sample autocorrelations at lag = 1 were statistically significant for all the countries, including the SAC region, except for Bangladesh, Pakistan and Afghanistan. All the latter values of sample autocorrelation, at lag = 2 or more were not significantly different from zero. The analysis indicates that there are opportunities for ARMA modeling of the first differences of these series, the representations may be approximate only [15, 16].

4. ALTERNATIVE TIME SERIES MODELS OF INFANT MORTALITY DECLINE

Various ARIMA(p,1,q) models were fitted. Based on residual checks, ARIMA(1,1,1), ARIMA(2,1,0), ARIMA(1,1,2) etc. were not acceptable. A random walk model and/or a ARIMA(1,1,0) fitted well to most of these series. In few cases, ARIMA(0,1,2) fitted marginally better than ARIMA(1,1,0). However, we restricted our choice to ARIMA(1,1,0) as adding more

parameters does not help. We provide the results in Table 3, to substantiate our findings.

Consider the estimated model for Bangladesh

$$Z_t - Z_{t-1} = \hat{\phi}(Z_{t-1} - Z_{t-2}) + \epsilon_t \tag{1}$$

with $\hat{\phi} = -0.1251$, if the mean of the differences is assumed to be zero. The estimated standard error of the autoregressive parameter is 0.1572 and the p-value is 0.431. This amounts to saying that $\hat{\phi}$ is not significantly different from zero. Thus, a random walk model is perhaps a better choice, since putting $\hat{\phi} = 0$ in model (1) leads to a random walk model. Therefore, a random walk model was fitted with innovation variance estimated as 0.000034. Figure 4 has a plot of the decline rate of Bangladesh infant mortality for 1961-2010. It also shows three forecasts of the series. Stationary ARMA(p, q) models do not seem appropriate for the series based on the unacceptable fit, but we have included a forecast made with an AR(1) model to show the effect of using a stationary model for a series that obviously is non-stationary. The other two

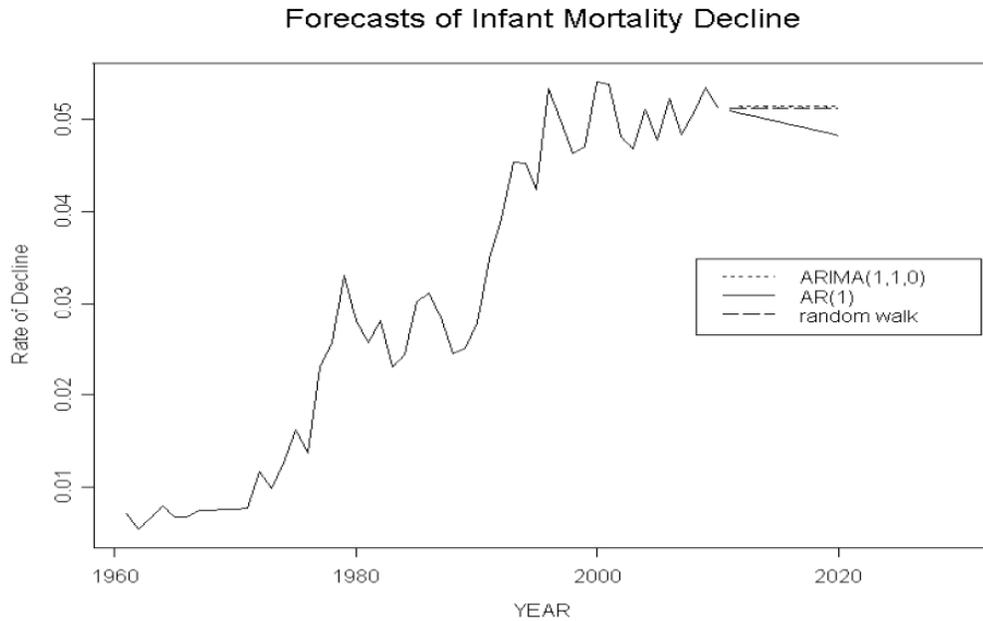


Figure 4: ARIMA, AR and random walk forecasts of infant mortality decline.

Table 2: Augmented Dickey-Fuller Test for Unit Root

Country	Test Statistic Z(t)	Interpolated Dickey-Fuller			MacKinnon approximate p-value for Z(t)
		1% Critical Value	5% Critical Value	10% Critical Value	
Afghanistan	-1.994	-4.187	-3.516	-3.190	0.6049
Bangladesh	-2.473	-4.187	-3.516	-3.190	0.3414
Bhutan	-2.027	-4.270	-3.552	-3.211	0.5866
India	-1.977	-4.187	-3.516	-3.190	0.6139
Maldives	-2.475	-4.196	-3.520	-3.192	0.3404
Nepal	-1.551	-4.187	-3.516	-3.190	0.8110
Pakistan	-3.547	-4.196	-3.520	-3.192	0.0346
Sri Lanka	-1.970	-4.187	-3.516	-3.190	0.6177
South Asia	-1.317	-4.187	-3.516	-3.190	0.8838

Source: Author's own computation.

forecasts are based on model(1) and a random walk model without drift. We see that the AR(1) based forecast continues smoothly from the last observed value to the historical (1990-2010) mean. The ARIMA(1,1,0) and Random Walk models produce essentially the same forecast.

5. ASSESSMENT OF FORECAST ACCURACY

The objective of this section is to provide a mathematical framework for the assessment of forecast accuracy of a serviceable representative ARIMA(1,1,0) model for forecasting logarithmic rate of infant mortality decline in South Asia.

Let $Y_{-1}, Y_0, Y_1, Y_2, \dots$ be a (doubly infinite) sequence of random variables. A particular realization $y_{-1}, y_0, y_1, y_2, \dots$ of the process is called a sample path. Suppose the iid sequence $\epsilon_{-1}, \epsilon_0, \epsilon_1, \epsilon_2, \dots$ with $E[\epsilon_t] = 0$ and $Var[\epsilon_t] = \sigma_\epsilon^2$ is white noise. Assume that each Y_t can be written in the form

$$Y_t = \psi_0 \epsilon_t + \psi_1 \epsilon_{t-1} + \psi_2 \epsilon_{t-2} + \dots \dots \dots (2)$$

where $\psi_0 = 1$, and the series of the absolute values of ψ_j 's converges. It can be shown that, the variance of Y_t is finite, and of the form

$$Var(Y_t) = \sigma_\epsilon^2 \sum_{j=0}^{\infty} \psi_j^2 (3)$$

Table 3: Estimated Time Series Models for Infant Mortality Decline

Country Name	ARIMA(1,1,0)		Random Walk Approximation
	AR Coefficient Estimate (Std. error)	p-value	Estimate of White Noise variance
India	-0.6553 (0.1193)	0.000	.0000044
Bangladesh	-0.1251 (0.1572)	0.431	.000034
Sri Lanka	-0.8647 (0.1461)	0.000	.000043
Nepal	-0.5510 (0.1351)	0.000	.0000076
Pakistan	-0.0569 (0.1602)	0.724	.000017
Maldives	-0.0392 (0.1593)	0.807	.00013
Bhutan	-0.5693 (0.1507)	0.001	0.000012
Afghanistan	-0.2721 (0.1524)	0.082	0.000016
South Asia	-0.5143 (0.1356)	0.000	0.0000046

Source: Author's own computation.

for all t. More generally, we have that

$$Cov(Y_t, Y_{t+k}) = \sigma_\epsilon^2 \sum_{i=0}^{\infty} \psi_i \psi_{i+k} \tag{4}$$

for all t and k ≥ 0. Suppose we make a forecast for Y_{t+k} at time t for lead time k = 1, 2, ... etc.

From (2) we can write the future values as Y_{t+k} = F_k(t) + E_k(t) where

$$E_k(t) = \psi_0 \epsilon_{t+k} + \psi_1 \epsilon_{t+k-1} + \dots + \psi_{k-1} \epsilon_{t+1} \tag{5}$$

and

$$F_k(t) = \psi_k \epsilon_t + \psi_{k+1} \epsilon_{t-1} + \dots \tag{6}$$

If the ψ_j's are known, then we know the value of F_k(t) at time t for an invertible ARMA(p, q) process, but E_k(t) is independent of the past and has mean zero. It follows that F_k(t) is the minimum mean squared error forecast of Y_{t+k}. Since error = forecast – true value, E_k(t) is the negative of the forecast error.

Now, consider an integrated process Z_t that is related to a stationary process Y_t (as defined in (2)) via the first differences Y_t = Z_t – Z_{t-1}. Suppose we know the values of Z_t for t = ..., -1, 0, 1, 2, ... and we want to forecast Z_{t+k}, for k=1, 2, We can always write

$$\begin{aligned} Z_{t+k} &= Z_t + Y_{t+1} + \dots + Y_{t+k} \\ &= Z_t + F_1(t) + E_1(t) + \dots + F_k(t) + E_k(t) \end{aligned} \tag{7}$$

Therefore, if we ignore the estimation error in the ψ_j's, the optimal forecast is

$$\widehat{Z}_{t+k} = Z_t + F_1(t) + \dots + F_k(t),$$

and the negative of the forecast error is

$$Z_{t+k} - \widehat{Z}_{t+k} + E_1(t) + \dots + E_k(t) \equiv E_{(k)}, \text{ say.} \tag{8}$$

We see from (5) and (6) that the E_j's are all linear combinations of ε_{t+h}'s with h = 1, ..., k, so the forecast error is independent of the forecast. A direct calculation yields the result,

$$Cov(E_{(k)}, E_{(k+h)}) = \sigma_\epsilon^2 \sum_{i=1}^k \left(\sum_{j=0}^{k-i} \psi_j \right) \left(\sum_{l=0}^{k+h-i} \psi_l \right) \tag{9}$$

for k, h ≥ 0.

When the first differences follow an AR(1) process, we have F_k(t) = φ^kY_t,

where Y_t = Z_t – Z_{t-1}. Therefore, the forecast function is given as

$$\widehat{Z}_{t+k} = Z_t + \phi(Z_t - Z_{t-1})(1 - \phi^k) / (1 - \phi) \tag{10}$$

which tends to

$$\widehat{Z}_{t+k} = Z_t + \phi(Z_t - Z_{t-1}) / (1 - \phi)$$

as $k \rightarrow \infty$. In demographic forecasting $|\phi|$ is often small, so the asymptotic value tends to be close to the current value. Taking \widehat{Z}_{t+k} as in (10), from (9), we can derive the second moments of the forecast error as

$$\begin{aligned} Cov(E_{(k)}, E_{(k+h)}) &= \frac{\sigma_\epsilon^2}{(1 - \phi)^2} \times \\ &\left[k - (\phi + \phi^{h+1}) \frac{1 - \phi^k}{1 - \phi} + \phi^{h+2} \frac{1 - \phi^{2k}}{1 - \phi^2} \right]. \end{aligned} \tag{11}$$

Because the partial sums in (9) are all positive for AR(1) first differences, it is easy to show that the covariance is positive for $|\phi| < 1$. We see from (11) that the covariance of the forecast error is asymptotically proportional to the shorter lead time k . Also, as $\phi \rightarrow 0$, the autocorrelations $\rho(E_{(k)}, E_{(k+h)})$ tends to $(k/(k+h))^{1/2}$, which is the autocorrelation function of a random walk. We know that under an AR(1) model for the process increments, $Var(Y_t) = \sigma_\epsilon^2 / (1 - \phi^2)$. So for large k and $h = 0$, we shall approximately have,

$$Var(E_{(k)}) \approx k \times Var(Y_t) \times (1 + \phi) / (1 - \phi).$$

Thus, when the process increments follow an AR(1) model, an approximation for the variance of the forecast error can be obtained based on a simple

random walk model by multiplying the empirical variance of the process of increments by the quantity $(1 + \phi) / (1 - \phi)$.

We shall now illustrate the methodology discussed above for the series Z_t , the rate of decline in infant mortality for entire South Asia. From Table 3, we see that the estimated representative model for the South Asia is found as

$$Z_t - Z_{t-1} = \widehat{\phi}(Z_{t-1} - Z_{t-2}) + \epsilon_t$$

with $\widehat{\phi} = -0.5143$ and the estimated standard error of the autoregressive parameter is 0.1356.

The last three values of the infant mortality rate for South Asia were $M_{2009} = 51.1$, $M_{2010} = 49.7$ and $M_{2011} = 48.3$, so that $Z_{2010} = 0.02778$ and $Z_{2011} = 0.028573$. Therefore, the last observed difference was $Y_{2011} = -0.00079$. It follows that the point forecast of Z_{2011+k} is

$$\begin{aligned} \widehat{Z}_{2011+k} &= 0.028573 + (-0.00079) \{ -0.5143 \cdot (-0.5143)^2 \\ &+ \dots + (-0.5143)^k \} \end{aligned}$$

for $k = 1, 2, \dots, 10$. The variance of the forecast error $Var[E_{(k)}]$ has been calculated using formula (11) with $h = 0$, $\phi = -0.5143$, and $\sigma_\epsilon^2 = 0.0000034$. The 50% prediction intervals are of the form $\widehat{Z}_{2011+k} \pm 0.6745 \times Var(E_{(k)})^{1/2}$, based on a normal approximation for the distribution of \widehat{Z} . Figure 5 illustrates the logarithmic rate of infant mortality decline in South Asia along with its forecast for 2011 to 2020 with 50% prediction interval.

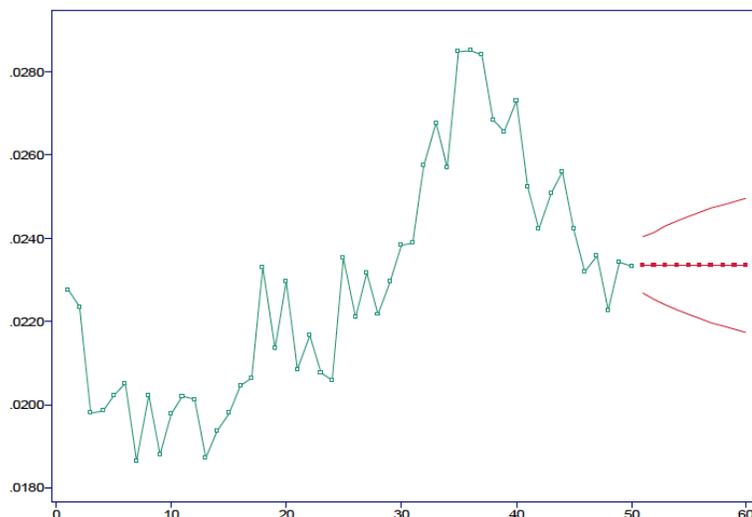


Figure 5: Logarithm of rate of Decline in Infant Mortality in South Asia and its Forecast for 2011-2020 with 50% Prediction Intervals.

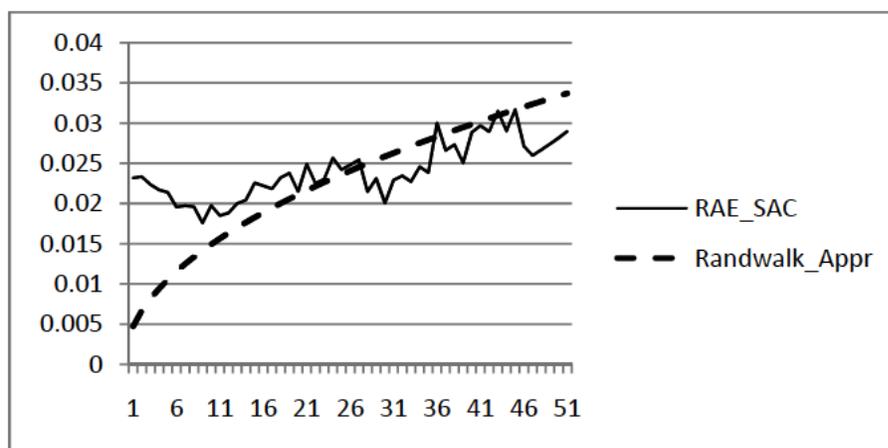


Figure 6: Absolute Relative Error of Infant Mortality Forecast as a Function of Lead Time for South Asia and a Random Walk approximation (dashed line).

6. ERROR ESTIMATE FOR INFANT MORTALITY FORECASTS IN SOUTH ASIA

If the decline in infant mortality rate were a random walk, using today's value for all future times would be optimal. Indeed, in countries, such as Bangladesh, Pakistan etc. a graph of these series often looks approximately like that of a random walk. Using this argument, suppose, X_{t+1} is a random variable representing a future value and let T be its forecast, which is based on past data. The absolute relative error of naive forecasts is defined as

$$RAE = \frac{|T - X|}{X}$$

Figure 6 displays empirical estimates of the absolute relative error of naive forecasts for the infant mortality rate in South Asia and the error of a random walk whose volatility closely matches the mean. Assuming the steps of the random walk are normal (Gaussian), the variance (volatility) estimated as 0.0047^2 ; then the absolute value of the error as a function of lead time is given by $0.0047 \times t^{1/2}$ and drawn as the dashed line in Figure 6. The order of magnitude of the estimated relative error at lead time 40 is approximately 0.03. This corresponds to an expected absolute error of about 3%.

Motivated by the developments in section 5, we now reconsider the forecasts of IMR in Table 4 for each country using our fitted models in section 4. From Table 4, it may be observed that linear extrapolation of IMR using the past rate may under or over estimate the shortfall in achieving MDG goal. For example, linear extrapolation of IMR for Bhutan at average annual rate

of reduction over 1990-2010 had suggested that it would over exceed the MDG 2015 target, however the ARIMA model indicates that there can be a shortfall to the extent of 11% with approximately 1% of error. So is the case, for the entire region of South Asia as indicated in the last row of Table 4. The shortfall was under estimated by about 28%.

7. CONCLUDING REMARKS

Demographic forecasts [17] of vital rates are linked to public policy issues and can make major contribution to planning for future in important areas of economic and social life. If the vital rates closely follow their past trends, accurate forecasting is feasible, but increased fluctuations of the rates usually imply rapidly increasing forecast errors. Country level comparisons of forecasts will reflect varying degree of forecast accuracy as a result of varying degree of change in mortality, fertility, and migration etc. Although, it is difficult to find a unique method or a unified approach to make the comparisons possible even when modern theoretical models are available to the forecasters. At times, it may be possible to find a serviceable representative model for the countries under consideration. Following Harold Dorn (1950), it may be stated that producing population forecasts using the representative model that are fairly uncertain can still have value, as the forecasts may draw attention to looming public policy issues that would otherwise be neglected [18].

The present article demonstrates that it is possible to find a serviceable representative time series model for the countries in the entire region of South Asia for the purpose of forecasting IMR decline so that country level comparisons and monitoring the stagnation in their rate of progress is made possible. Although a

Table 4: Forecasts of IMR in South Asia

Name of the Country	Forecasts using ARIMA(1,1,0) models for the Year				MDG Target Value (2015)	Shortfall Estimates in % using ARIMA	Shortfall Estimates In % using linear extrapolation
	2012	2013	2015	2020			
Afghanistan	69	66	59	45	43	27.12	29.28
Bangladesh	35	33	30	23	32	-6.67	-92.77
Bhutan	40	39	36	29	32	11.11	-30.61
India	46	45	42	36	27	35.71	25.41
Maldives	8	7	5	3	25	-400.0	**
Nepal	37	36	33	27	31	6.06	-53.47
Pakistan	58	57	55	49	32	41.82	35.22
Sri Lanka	10	9.6	9	7	8	11.11	**
South Asia	47	46	43	37	28	34.88	6.67

Source: Author's own computation.

gross level comparisons may be indicative of success or otherwise for each country using descriptive measures such as average annual rate of reduction as used in this article, however, inappropriate choice of simple linear models or so may lead to significant magnitude of over or under estimation. Demographic processes are generally non-stationary containing a stochastic trend component. When such is the case, and an ARIMA(1,1,0) model provides a statistically accepted fit to vital time series data, the forecast errors can be approximated using a random walk process [11].

REFERENCES

- [1] United Nations. The Millennium Development Goals Report 2011, The United Nations: New York; 2011.
- [2] World Bank. Attaining the Millennium Development Goals in India: Role of Public Policy & Service Delivery. Washington DC USA; 2004.
- [3] Chatterjee A, Paily VP. Achieving Millennium Development Goals 4 and 5 in India. *BJOG* 2011; 118(suppl. 2): 47-59. <http://dx.doi.org/10.1111/j.1471-0528.2011.03112.x>
- [4] Sundaram K, Tendulkar S. Poverty Has Declined in the 1990s – A Resolution of the Comparability Problems in NSS Consumer Expenditure Data. *Economic and Political Weekly* 2003; 38: 25-31.
- [5] Sundaram K, Tendulkar S. Poverty in India in the 1990s – Revised Results for India and the 15 Major States for 1993-94. *Economic and Political Weekly* 2003; 38: 15-21.
- [6] UNICEF. Levels and Trends in Child Mortality Report 2012. United Nations Children's Fund: USA; 2012.
- [7] World Bank. World Development Indicators 2011. Washington, D.C., USA, 2011.
- [8] Bajpai N. India: Towards the Millennium Development Goals, Background Paper, Human Development Report, UNDP; 2003: Available from: http://hdr.undp.org/docs/publications/background_papers/2003/India/India_2003.pdf
- [9] Sullivan JM, Shea OR, George TB. Infant and Child Mortality. DHS Comparative Studies No. 15; Macro International Inc: Calverton Maryland 1994.
- [10] United Nations Inter-agency Group for Child Mortality Estimation. UN IGME Total under-5 Mortality Rate and Infant Mortality Rate database 2012; Available from: <http://www.childmortality.org/>
- [11] Alho JM, Spencer BD. Statistical Demography and Forecasting. Springer Science+ Business Media Inc: USA; 2005.
- [12] Bishai DM. Infant Mortality Time series are random walks with drift: Are they cointegrated with socioeconomic variables? *Health Economics* 1995; 4(3): 157-167. <http://dx.doi.org/10.1002/hec.4730040302>
- [13] Dickey DA and Fuller WA. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 1979; 74: 427-431
- [14] MacKinnon JG. Critical values for cointegration tests, in: Engle RF, Granjer CWJ, editors. *Long-Run Economic Relationships: Reading in Cointegration*. Oxford University Press 1991; p. 267-76.
- [15] Montgomery DC, Jennings CL, Kulahci M. *Introduction to Time Series Analysis and Forecasting*. Wiley: New York; 2008.
- [16] Shumway RH, Stoffer DS. *Time Series Analysis and its Applications With R Examples*. Springer Science + Business Media: LLC; 2011.
- [17] Girosi F and King G. *Demographic Forecasting*. 2008. Princeton University Press, New Jersey; 2008.
- [18] Dorn H. Pitfalls in population forecasts and projections. *Journal of the American Statistical Association* 1950; 45: 311-334. <http://dx.doi.org/10.1080/01621459.1950.10501127>