Mathematics and Statistics: Open Access
Received: Dec 05, 2015, Accepted: Jan 21, 2016, Published: Jan 23, 2016
Math Stat, Volume 1, Issue 1
http://crescopublications.org/pdf/msoa/MSOA-1-003.pdf
Article Number: MSOA-1-003

Research Article Open Access

Coherent Forecasting of Fertility in Urban and Rural Areas of Pakistan

Farah Yasmeen*

University of Karachi, 75270, Pakistan

***Corresponding Author:** Farah Yasmeen, University of Karachi, 75270, Pakistan, E-mail: riazfarah@yahoo.com

Citation: Farah Yasmeen (2016) Coherent Forecasting of Fertility in Urban and Rural Areas of Pakistan. Math Stat 1: 003.

Abstract

In this paper, we intend to obtain coherent forecasts of age-specific fertility rates of Pakistan for the two broader regions, namely, Urban and Rural areas. This is an application of coherent functional models of Hyndman RJ, Booth H & Yasmeen F [1] on the disaggregation of fertility forecasts by regions. On the basis of previous studies [2, 3], we found that fertility rates in rural areas are higher than those of urban areas. Here, we can assume that the future fertility rates of rural areas will remain higher than those of urban areas, for all age groups. Hence, the fertility forecasts thus obtained are assumed to be coherent or more realistic with the historical fertility rates.

We first describe the concept of coherence in the context of age-specific fertility and discuss some problems with using independent functional time series models of Hyndman RJ & Ullah MS [4] for urban and rural regions. Then, we apply the coherent functional model to the age-specific fertility data. An empirical comparison of the independent and coherent models based on the fertility rates has been made. The purpose here is to see the performance of coherent forecasting models for the fertility rates of these two regions of Pakistan and to find the more realistic fertility forecasts by this application.

Keywords: Fertility; Fertility levels; Age-specific fertility rates; Rural-urban fertility gap; Functional time series; Forecasting; Coherent forecast.

1. Introduction

The study of fertility levels and differentials are key factors to understand population growth. In the past, striking differences have been found in fertility rates of rural and urban areas of Pakistan. It has been documented that rural populations are more fertile than urban populations[5, 6], specifically in developing countries like Pakistan. Although, regardless of the overall level of fertility, rural-urban differences have been observed in most societies, which are more significant in less developed societies and their interpretation and meaning have been eagerly questioned [7]. Some pollsters contended that rural-urban differences are typically small and before the onset of fertility transition arise solely due to lower proportions of married amongst women in urban areas as a consequence of more job opportunities for urban women [6]. However, despite the considerable attention on this topic, little work has been done to relate rural-urban fertility differences systematically to socio-economic development.

To describe the reproductivity and fertility pattern in Pakistan a variety of mathematical models have been proposed in the past [8-11]. However to study the fertility, the existing literature was lacking the modelling approach. Also, neither of these studies had considered the forecasting of fertility curve for future.

To obtain the fertility forecasts Yasmeen F and Mahmood Z[2, 12] applied Functional Time Series (FTS) models to the age-specific fertility data of entire population of Pakistan and to the fertility rates of Urban and Rural areas respectively. Those pieces of work mainly concentrate on the modelling and forecasting age-specific fertility rates using independent functional time series (FTS)models.
However, on the basis of past studies, the figures indicate that the urban fertility is quite lower than the rural one, as the awareness about family planning is less in the rural areas of Pakistan. Moreover, rural areas population consist of around 64%and urban is on 36%, so it obvious to have higher fertility rate in rural areas. Hence, it is desirable to obtain the fertility forecast of urban and rural areas in such a way that the resultant forecast will exhibit relationship in the historical fertility rates.
In this paper, we apply the coherent functional time series models of Hyndman RJ, Booth H & Yasmeen F [1] to the age-specific fertility rates of urban and rural areas of Pakistan. The main objectives of this paper are:

- To observe the components of FTS models applied independently to the urban and rural regions.
- To apply the coherent functional model of Hyndman RJ, Booth H & Yasmeen F [1] based on the product and the ratios of the fertility rates of two regions
- To study the components of product and the ratios of fertility rates
- To obtain coherent forecasts of fertility curve for the next twenty years

The paper is divided into six sections. Section 1 is introductory; with fertility patterns for the urban and rural areas of Pakistan are discussed in section 2. Some basic concepts of functional time series (FTS) and the coherent functional model are explained in section 3 and 4 respectively. The results of the statistical analysis are discussed in section 5, and finally some concluding remarks are given in section 6.

2. Fertility Patterns in Urban and Rural Areas of Pakistan

The estimated population of Pakistan was over 187 million [13] in 2011, and it is estimated to be increased by over fourfold during 1950–2011. Estimates from different sources imply decline in fertility in Pakistan particularly after 1990s [3].
The graph of ASFR of Pakistan (Figure 1) [12] shows that during 1984-2005, the fertility rates have been decreased for all age-groups. A greater decline occurred in the middle age group i.e. 20-24, 25-29 and 30-34, and relatively slower decline in the other age groups (15-19, 35-39, 40-44 and 45-49). Another import point is that the pattern is almost the same for the study period (1984-2005), which is to be considered as the reciprocal of V-shape, and we did not find any significant change in this pattern. By plotting ASFR of Pakistan, it is found that their distribution has a typical shape. For the fertility data of Pakistan, this non-linear pattern of fertility curve can be considered as the reciprocal of the V-shapes. In literature, we found that some polynomial models are fitted to the ASFR. However, little attention has been paid to forecast these rates. As stated before, Yasmeen Fatima and Mahmood [2] applied independent FTS models to the urban and rural fertility rates and obtained 20-year forecasts. In these forecasts the least reduction has been seen in the rural areas as compared to urban region because of lack of educational facilities, lack in exposure of mass media, early marriages, poor health facilities and lacking of extracurricular activities. If we compare the forecast values of the fertility rates for the two regions of Pakistan (Figure 5) [12], the maximum reduction was expected in the age-group (25-35) and relatively small in the other age groups. However, the future fertility rates were expected to decline slowly in the rural region than those for the urban region.

3. Functional Time Series (FTS) Models

Functional time series (FTS) models based on functional data analysis (FDA) have attained substantial development in recent years. Such time series refers to the data in the form of curves that are observed at regular intervals in time. These models are first introduced by Hyndman RJ &Ullah MS [4], with further applications on breast cancer incidence, lung cancer incidence and age-specific fertility forecasting [14-17]. Yasmeen F [18] described the common principal component (CPC) approach to the FTS models.
A functional time series model proposed by Hyndman RJ & Ullah MS [4] was first used for demographic forecasting. Their technique is generally known as Hyndman-Ullah technique. Ramsay JO and Silverman BW [19] described the functional data paradigm more generally. The Hyndman-Ullah technique uses several principal components obtained by applying singular value decomposition (SVD) to the matrix of log mortality/fertility rates. The main model is

$$y_t(x) = f_t(x) + \sigma_t(x)\varepsilon_t, \qquad (1)$$

Where $y_t(x)$ denotes the observed log fertility rate at age 'x' in year 't'. We assume that there is a smooth function of 'x', allows the amount of noise to vary with 'x', and ε_t, are considered to be independent and identically distributed random variables with zero mean and unit variance. The first step is to estimate these smooth functions from the discrete noisy data. This is possible using nonparametric smoothing methods. Hence the smoothed curves are decomposed using

$$f_t(x) = \mu(x) + \sum_{j=1}^{J} \beta_{t,j}\phi_j(x) + e_t(x) \qquad (2)$$

Where $\mu(x)$ is the mean log fertility rate across years and $\{\phi_j(x)\}$ is a set of orthogonal basis functions.

The values $\{\beta_{j,1},\ldots,\beta_j,n\}$ form a univariate time series for $j=1,\ldots,J$. The basic functions $\{\phi j(x)\}$ are computed using functional principal components applied to the smooth curves $ft(x)$.

The h-step ahead forecast of $\hat{y}_{,n+h,}(x)$ can be obtained as

$$\hat{y}_{n+h}(x) = \hat{\mu}(x) + \sum_{j=1}^{J} \hat{\beta}_{j,n+h}\hat{\phi}_j(x) \qquad (3)$$

Where $\hat{\mu}(x)$ and $\hat{\phi}_j(x)$ are the estimates of the mean function and the basic functions, respectively and denotes the h-step ahead forecast of $\beta_{j,n+h}$. Assuming independence, the forecast variance can be obtained by adding the variances of all individual terms in equations (1) and (2).

4. Coherent Functional Models

Hyndman RJ, Booth H & Yasmeen F[1] proposed the coherent functional model that provides forecasts of the mortality rates of two or more groups within a given population by simultaneously modeling the geometric mean of the age-specific mortality rates and the ratio of the group-specific rates to the geometric mean of the mortality rates of the population as a whole. However, these models are used for mortality forecasting, in this paper, we tried to apply them for the first time on fertility rates of two groups.

Coherent Functional Model for Urban and Rural Fertility Rates

Let $f_{t,u}(x)$ denote the age-specific fertility rate for urban women of age x in year t, t = 1. . , n. We model the log fertility rate, $y_{t,u}(x) = log[f_{t,u}(x)]$, and assume that there is an underlying smooth function that we are observing with error. Thus,

$$y_{t,u}(x) = log[f_{t,u}(x)] + \sigma_{t,u}(x_i)\,\epsilon_{t,w,i} \quad (4)$$

Where xi is the center of age group i $(i = 1, \ldots, p)$, $\epsilon_{t,w,i}$ is an independent and identically distributed standard normal random variable and $\sigma_{t,u}(x_i)$ allows the amount of noise to vary with age x. Analogous notation can be used for rural fertility rates.

Define the square roots of the products and ratios of the smoothed rates for each group:

$$p_t(x) = f_{t,u}(x).\,f_{t,r}(x)$$

and

$$r_t(x) = f_{t,u}(x) /\, f_{t,r}(x) \qquad (5)$$

5. Fitting Coherent FTS Models

Data

The secondary data of age-specific fertility rates (ASFR) from 1984 to 2007, obtained from Pakistan Demographic Surveys [19] are used. These data are available for 1984-1986, 1988-1992, 1995-1997, 1999-2001, 2003 and 2005-2007. Age-specific fertility rates are missing for rest of the years, so we estimate those using interpolating splines. For each age-group, we consider a fixed number of knots as available ASFR and estimate the function for missing years (The reader is referred to Ref. [12] for details).

Results

At first, we obtain ASFR for the years 1984-2007 and consider them as functional observation. We smooth these rates using non-parametric methods. We use weighted regression B spline to obtain smooth curves. In applying the coherent functional model to the ASFR of urban and rural areas, the first step is to get the product component $p_t(x)$ and ratio component $r_t(x)$ using equation (5). Figures 1 and 2 depict these components of the smoothed log fertility rates during the years 1984 – 2007. The product component represents the overall effect of the whole population (irrespective of groupings), and it shows that overall, the fertility rates for women of both group increased with age till age 27 years, and then they declined. This decline was sharper in the age group (25-35) and relatively slower in the early ages (below 22 years) and older ages(above 37 years). Also, during the study period (1984-2007), the fertility rates of all Pakistani women decreased and the pattern was not change since 1984. The rate declined continually and the decline in rates appeared relatively faster for the young women in the most recent years.

The ratio component shows the square root of the ratio of the fertility rates of urban to the fertility rates of rural defined in equation (5), as displayed in Figure 2. This figure suggests that the urban to rural fertility ratio was less than 1.0 for all age groups for all available years. The ratio decreased continually since 1984, and in most recent years, the ratio has consistently been less than 1 for all age groups. This suggests that urban women have lower fertility rates than rural women for all ages during the study period. Also, the ratio is decreasing for the women aged below 25 years and over 35 years, where the ratio of women between ages 25-35 is relatively stable, it means that the major differences in the fertility rates of urban and rural populations occur in younger and the older women and relatively stable in the women of middle age-group.

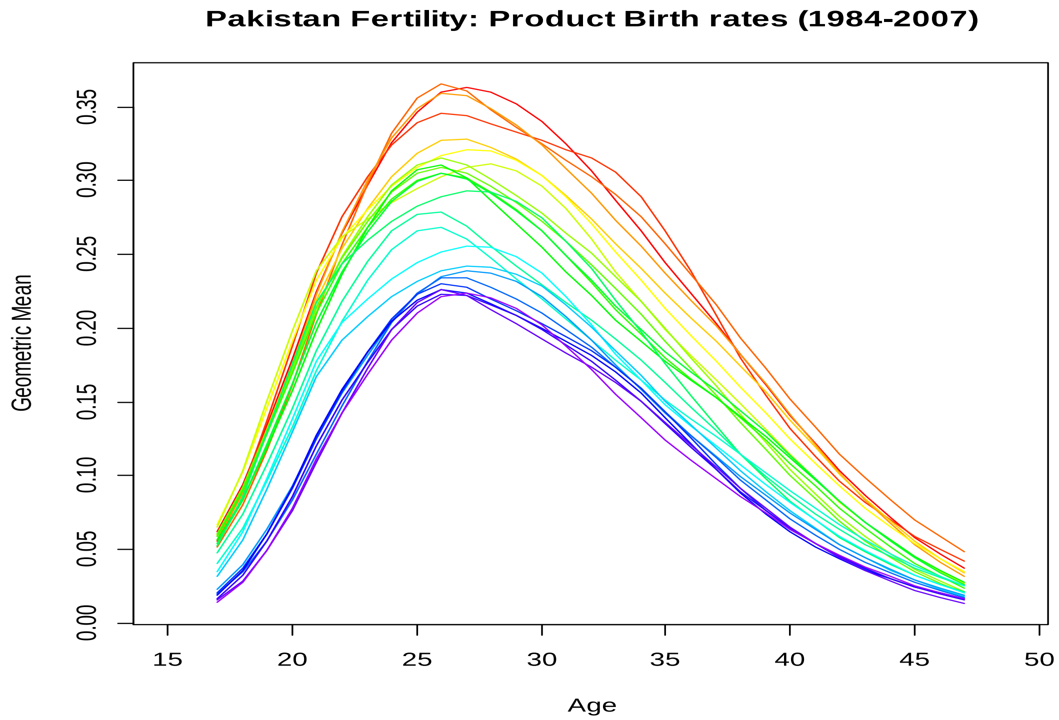

Figure 1: Product Component of fertility rates of Pakistan for urban and rural regions.

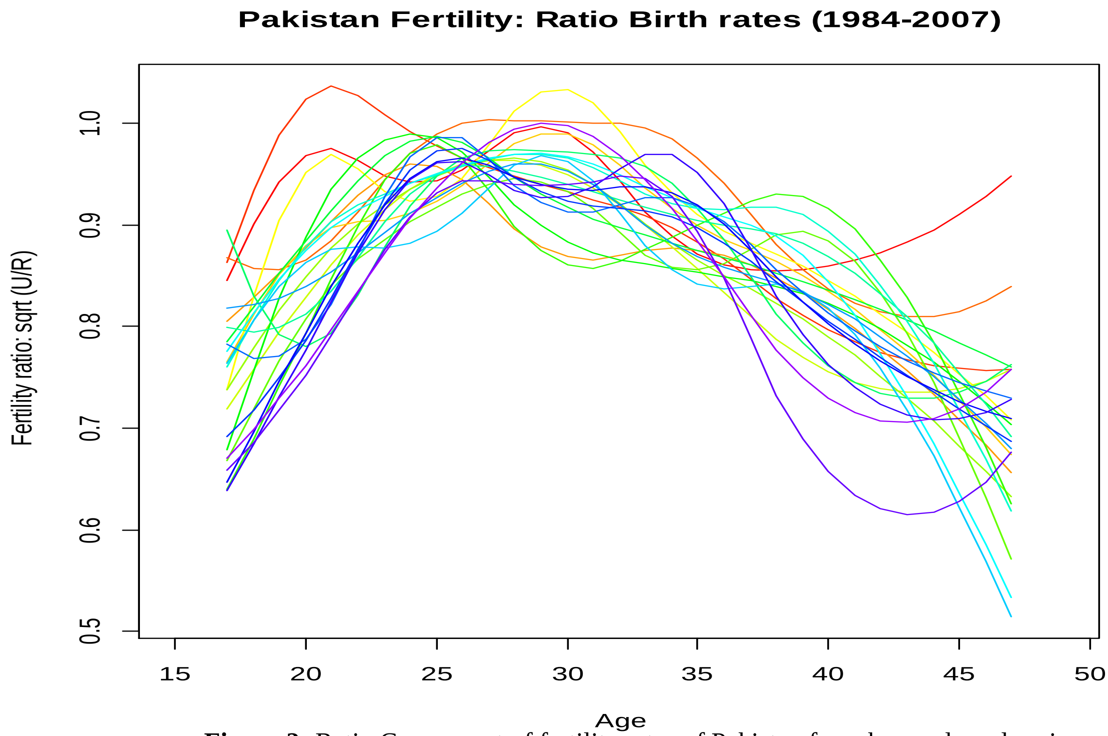

Figure 2: Ratio Component of fertility rates of Pakistan for urban and rural regions.

The next step is to apply the basic functional time series model [4] to the product and ratio components. Figures 3 and 4 show basis functions and the respective time series coefficients obtained by applying the functional principal components decomposition to the product and ratio curves, along with twenty-year forecasts of the time series coefficients. These plots represent the various sources of variation in the two components. From Figure 3, it is clear that the mean function of the product model is an increasing till age 27 years and then it is declining. The first basis function of the product component shows that the age-specific fertility rates for the two groups of women (urban and rural) were decreasing for younger women(under 25-years). Similarly, the second basis function is a contrast between very young (15 years) and very old (45 years). Its respective coefficient shows that the rates for older women were initially decreasing (from 1984 till 2000) and then they started to increase slightly from year 2000. These rates are expected to decline slowly in the next 20-years. We use ARIMA models to forecast the product component coefficient, using automatic forecasting algorithm of Government of Pakistan, Population Demographic Survey-2005 [20] to select the best possible ARIMA model.

The mean function of the ratio component in Figure 4 shows that on average, the urban to rural fertility ratio increases with age till 25 years and then decreasing.

The first basis function of the ratio component shows that most of the variation in the urban to rural fertility-ratio is in the age groups of women aged 17 years and under, and above 45years and the ratio of their fertility rates is decreasing since 1984. While the second basis function represent ages 40 years and above, with corresponding rates are expected to decrease in the next 20-years. For forecasting the ratio coefficients, we use stationary ARFIMA *(p, d, q)* $(-0.5 < d < 0.5)$ models and obtain forecasts using the algorithm developed by Hyndman RJ, Booth H & Yasmeen F [1]. The twenty-year forecasts of the first ratio coefficient show that urban to rural fertility ratio may remain constant in the future. This implies that the forecasts for the entire fertility curves for urban and rural women will not diverge in the long run. The other basis functions and coefficients have little effect on the ratio forecast, as the forecast values of the coefficients are nearly zero.

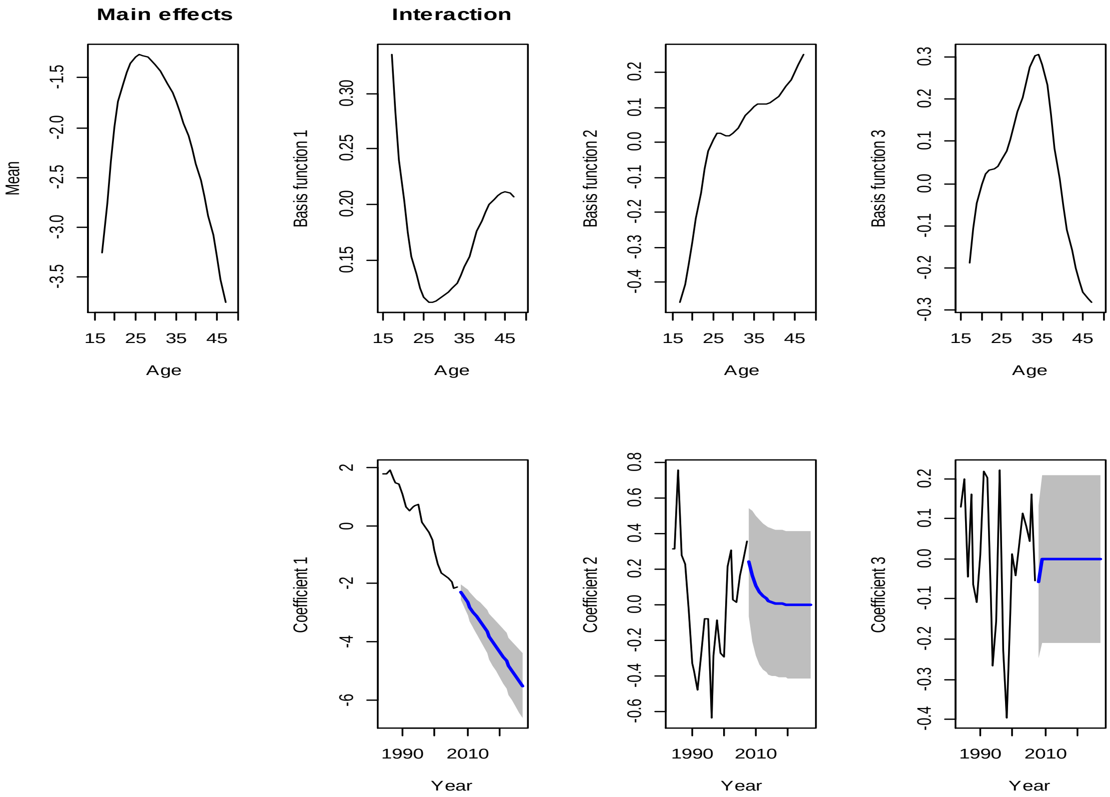

Figure 3: Forecast for the first three coefficients of FTS model for product component.

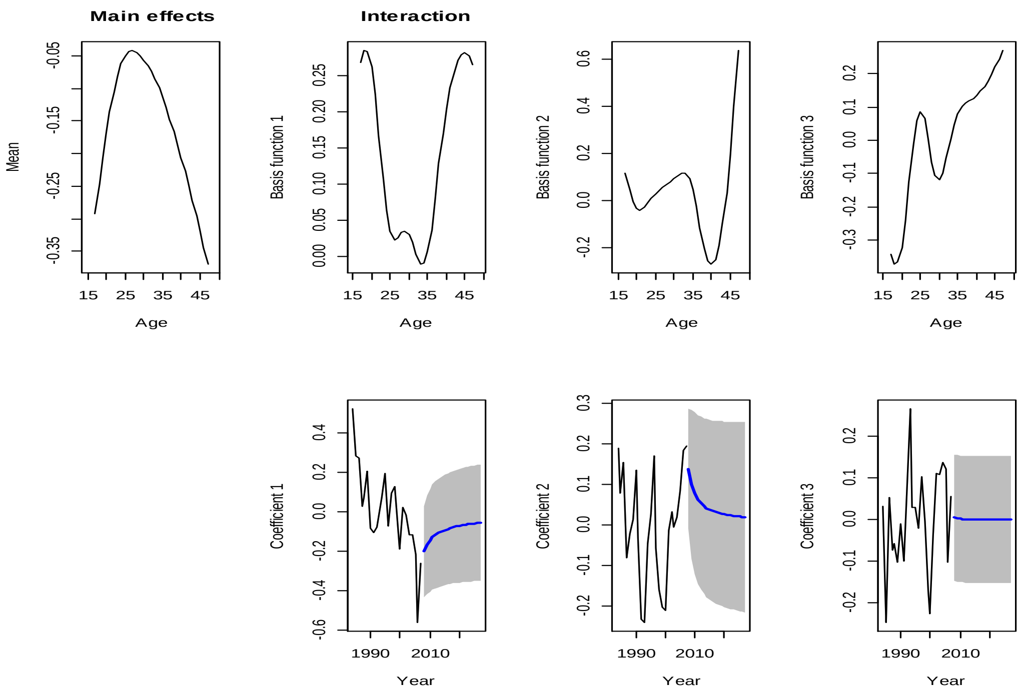

Figure 4: Forecast for the first three coefficients of FTS model for ratio component.

If we compare the figures 3 and 4, it can be observed that the coefficients of product models are non-stationary, whereas the coefficients of ratio models are stationary with narrow prediction intervals. Hence, the stationary condition makes the forecasts of the two groups to be non-divergent.

Figures 5 and 6 show twenty-year forecasts of age-specific fertility for urban and rural areas of Pakistan. These graphs show a decline in fertility rates for both urban and rural women over the next twenty years. This decline is steeper among urban women of middle age (25-35 years), and relatively slower among younger women (under 25 years of age) and older women(above 35 years). However, for rural areas, future fertility rates are expected to decline substantially for women aged between 22 and 38 years, and relatively slowly for women of other ages.

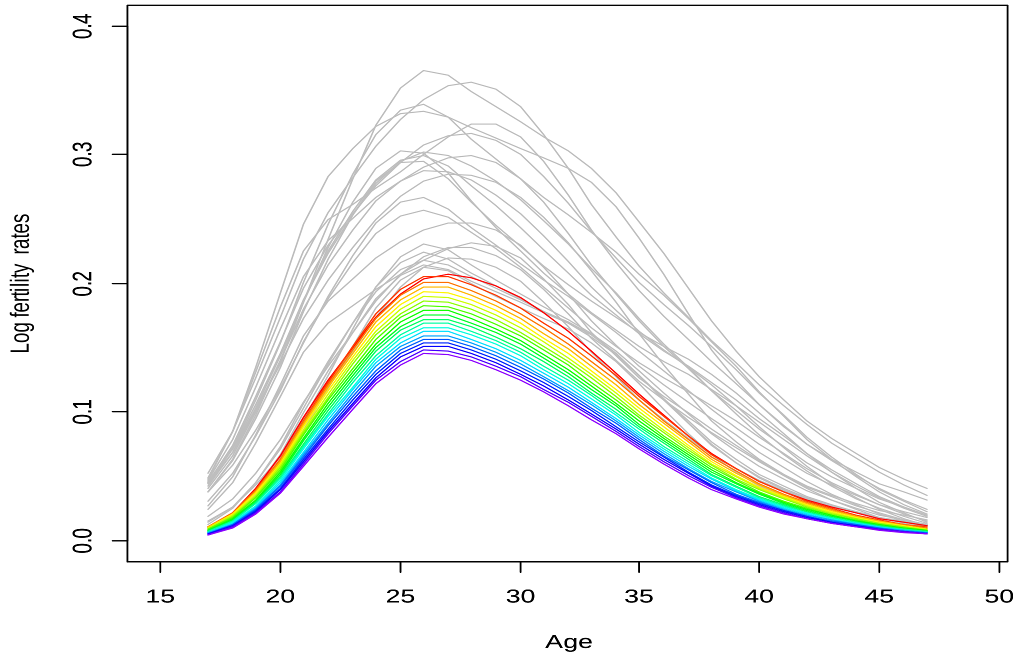

Figure 5: 20-year Fertility Forecasts of Pakistan (urban region) using Coherent Functional Models

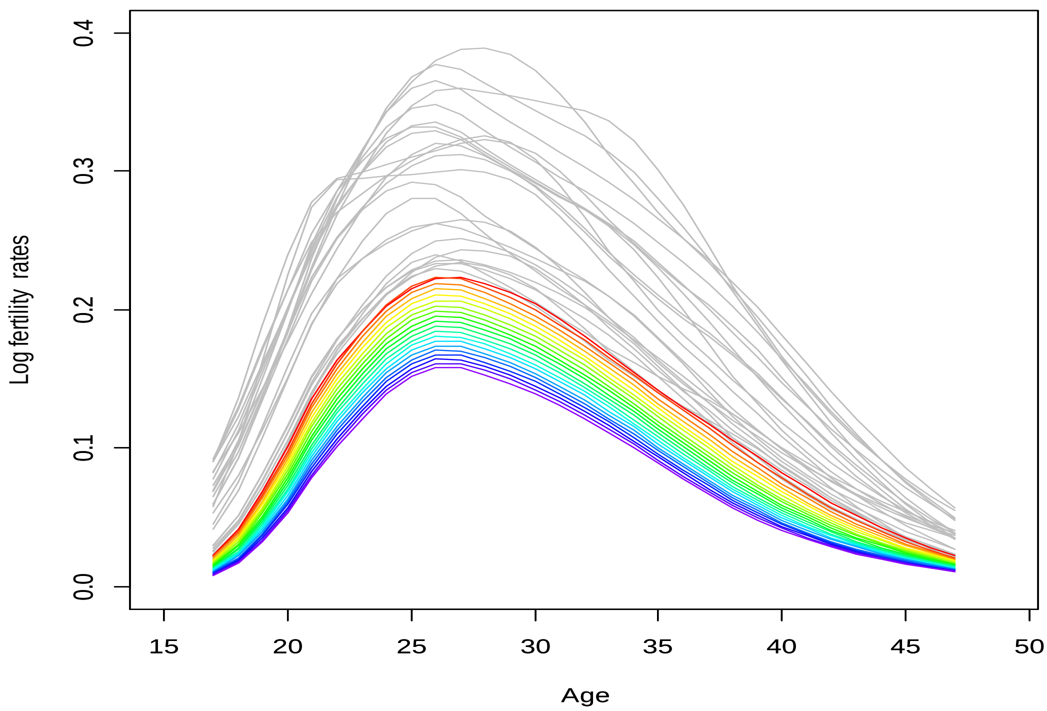

Figure 6: Fertility forecasts of Pakistan (rural region) using Coherent Functional Models

Empirical Comparison of Independent and Coherent Models

In this section, we will compare the results obtained from the application of independent and coherent functional models to the fertility rates of urban and rural areas of Pakistan. Figure 7 depicts the first basis functions for the two polar groups obtained from independent functional models. It is clear that the first basis function of the urban population (which is more responsible for the overall variation in the fertility rates) represents women of ages 20 years and 40 years, whilst the first basis function for rural represent very young women(15 years of age) and 35 years. The values are relatively smaller for the age group 25-35 years for urban women, as compared to rural (Figure 7) for the same age-group. This causes the divergence in the fertility forecasts for the two groups, as shown in Figure 8.

Using independent models, the future fertility rates in urban population of Pakistan are expected to decline relatively faster in ages 25-35 years, than their rural counterpart. In rural areas, fertility rates are expected to decline faster for 35-years of age and it seems that the future fertility rates will become lower for rural areas than the urban rates for this age-group. However, it is not expecting using coherent models, as the forecasts are more realistic with the historical fertility rates.

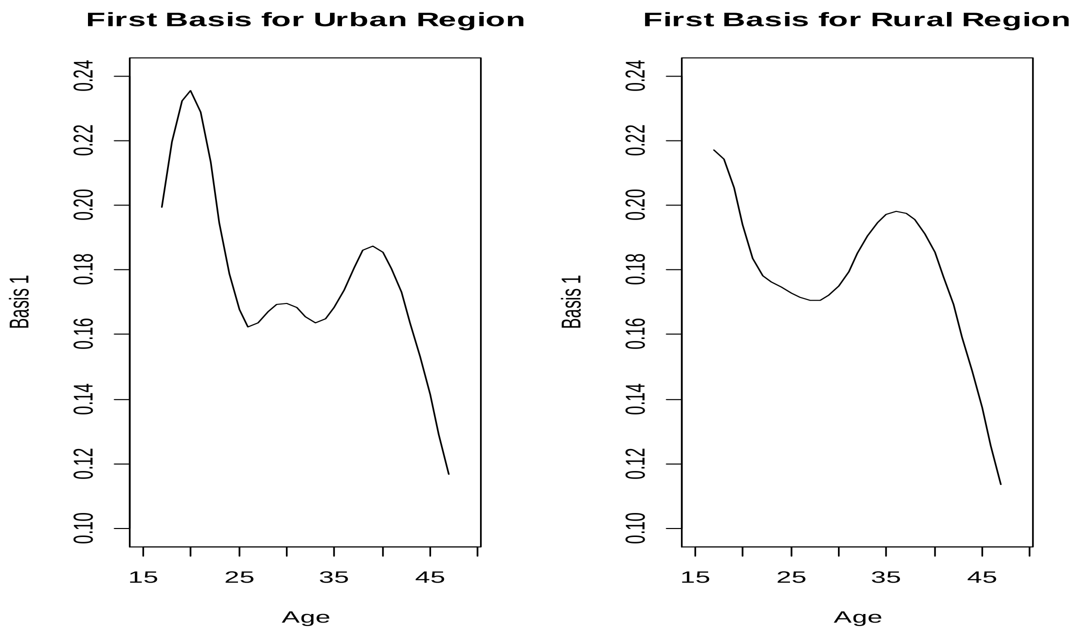

Figure7: First Basis Function of Independent FTS for Urban and Rural Fertility in Pakistan

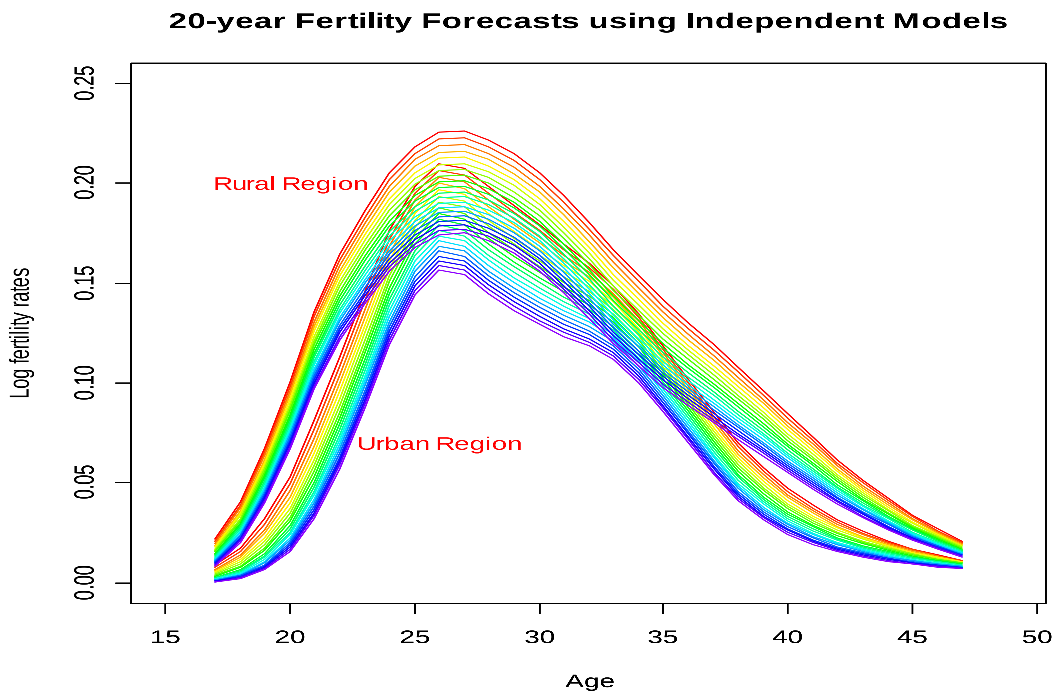

Figure 8: 20-years Forecasts of fertility rates of Pakistan for urban and rural regions using independent FTS models

6. Conclusion

In this paper, we applied coherent functional time series (FTS) models to the age-specific fertility rates of Pakistan for the two broader areas, urban and rural. We have shown that, on average, rural women have had higher fertility rates than their urban counter-parts in all age-groups since 1984 (from where the data of fertility rates are available). Significant inverse correlation between proportion of urban population and fertility rate leads to arrive the conclusion that difference in reproductive behavior between the women of two polar type community stems mainly from differences in their socio-economic conditions, specifically from differences in the literacy rate, level of education, and age at marriage. Our forecasts of future fertility rates show that these differences are expected to be maintained over the next twenty years.

Estimates from different sources imply decline in fertility in Pakistan particularly after 1990s [22]. Our results also confirm the results of the previous research. In addition, we obtained 20-years predictions for the ASFR. These plots exhibit that the future fertility rates will decline for all ages, but that the decline will be greatest for the middle age women, and relatively slower for the youngest and oldest women. The results obtained here are consistent with the findings of Yasmeen Fatima and Mahmood [2], Yasmeen F and Mahmood Z[12] andHyndman R & Khandakar Y[21] that the overall fertility rates are decreasing.

The study also suggests that rural areas should be the focus of policy formulation and program implementation because major part of the population of Pakistan is still inhabited there. Currently almost two-third of its population live in rural areas showing high fertility rate than that of urban population and is playing major part in overall rapid population growth in the country. Resultantly, the quality of population in terms of education, jobs, provision of health and other facilities is adversely affected; hence this issue needs realistic response.

References

[1] Hyndman RJ, Booth H &Yasmeen F (2013) The product-ratio method with functional time series models, Demography 50(1), 261–283.

[2] Yasmeen Fatima and Mahmood (2014)An FDA Approach to Forecast Age-Specific Fertility Rates of Pakistan Region-Wise. Computer Science and Applications 1(6), 341-348, www.ethanpublishing.com

[3] Nasir JA, Akhtar M and Tahir MH (2009) Reproductivity and age-specific fertility rates in Pakistan after 1981.Pakistan Journal of Statistics 25(3), 251-263.

[4] Hyndman RJ &Ullah MS (2007) Robust forecasting of mortality and fertility rates: a functional data approach, Computational Statistics and Data Analysis 51(10), 4942–4956.

[5] UN (1973) The Determinants and Consequences of Population Trends, New Summery of Findings on Interaction of Demographic, Economic and Social Factors. Population Studies. New York: Department of Economic and Social Affairs. (50)1.

[6] Mosk C (1980) Rural-Urban Fertility Differences and the Fertility Transition, JSTOR: Population Studies. (34)1, pp. 77-90.

[7] Kuznets S (1974) Rural-Urban Differences in Fertility: An International Comparison. Proceedings of the American Philosophical Society 118, 1-19.

[8] Pollard AH, Yusuf F and Pollard AH (1990) Demographic Techniques. Pergamon Press, New York, Inc.

[9] Peristera P and Kostaki A (2007) Modelling fertility in modern populations. Demographic Research, 16, 141-194.

[10] Sathar ZA and Casterline JB (1998) The onset of fertility transition in Pakistan. Population and Development Review, 24, 773-796.

[11] Sathar ZA and Kazi S (1990) Women, work and reproduction in Karachi. International Family Planning Perspectives, 16, 66-80.

[12] YasmeenF and Mahmood Z (2012) Forecasting age-specific fertility rates of Pakistan using functional time series models, in: ICCS-12, Qatar.

[13] CIA (2011): The World Fact Book, Population Growth Rate.

[14] Yasmeen F, Hyndman RJ &Erbas B (2010) Forecasting age-related changes in breast cancer mortality among white and black US women: A functional data approach, Cancer Epidemiology 34(5), 542-549.

[15] Yasmeen F and Zaheer S (2014) Functional time series models to estimate future age-specific breast cancer incidence rates for women in Karachi, Pakistan, Journal of Health Science, 2 (2014) 213-221.

[16] YasmeenF and Mughal S (2014) Functional Time Series Models and the APC models: A comparative study on the lung cancer incidence rates in Denmark Journal of US-China Medical Science, David Publishing Company, 11 (3) pages. 121-128.

[17] Yasmeen F (2015a) Coherent Forecasts of US Age-specific Breast Cancer Mortality, Donnish Journal of Cancer Research and Experimental Oncology, Vol 1(1), 001-008. http://www.donnishjournals.org/djcreo

[18] Yasmeen F (2015b) The Common Principal Component (CPC) Approach to Functional Time Series (FTS) Models, Science Journal of Applied Mathematics and Statistics, 3(1): 22-26.

[19] Ramsay JO and Silverman BW (2005) Functional Data Analysis, Springer.

[20] Government of Pakistan (2005) Population Demographic Survey- 2005.Federal Bureau of Statistics, Statistics Division, Islamabad, Pakistan.
[21] Hyndman R & Khandakar Y (2008) Automatic time series forecasting: the forecast package for R, Journal of Statistical Software 27(3). URL: http://www.jstatsoft.org/v27/i03
[22] Hakim MA (2000) Fertility Transition and Differentials in Pakistan, Pakistan's Population Issues in the 21st Century. Conference Proceedings, 24-26 October Karachi. Islamabad: Population Association of Pakistan, 611-633.