Economic Affairs, Vol. 62, No. 1, pp. 103-111, March 2017
DOI: 10.5958/0976-4666.2017.00033.X

Hierarchical Time-series Models for Forecasting Oilseeds and Pulses Production in India

Dipankar Mitra, Ranjit Kumar Paul and Soumen Pal

ICAR-Indian Agricultural Statistics Research Institute, New Delhi, India

Corresponding author: ranjitstat@iasri.res.in

Received: 01-11-2016 Revised: 06-01-2017 Accepted: 19-02-2017

ABSTRACT

Hierarchical time-series, which are multiple time-series that are hierarchically organized and can be aggregated at several different levels in groups based on geographical locations or some other features, has many practical importance. There are certain specialized strategies, viz. top-down, bottom-up, middle-out and optimal approaches which take care of predicting future values for such multi-level data. The top-down approach at first provides forecasting for the aggregated series at the top most level of the hierarchy, then disaggregating the forecasts in the lower levels based on historical and forecasted proportions. The bottom-up method provides forecasting for the most disaggregated series at the bottom level of the hierarchy and then aggregates these forecasts to obtain the forecasts at the top level of the hierarchy. The middle-out approach is a combination of bottom-up and top-down approaches. The optimal combination approach involves forecasting all series at all levels in the hierarchy, and then using a regression model to obtain the optimally combined forecasts. As an example, forecasting of oilseeds as well as pulses production in India is attempted using hierarchical time-series models.

Keywords: Bottom-up, forecasting, hierarchical time-series, middle-out, oilseeds production, pulses production, top-down

Time-series forecasting is an important statistical analysis technique used as a basis for manual and automatic planning in many application domains (Gooijer and Hyndman, 2006). Forecasts are calculated using mathematical models that capture a parameterized relationship between past and future values to express the behavior and characteristics of a historic time series. The parameters of these forecast models are estimated on a training data set to fit the specifics of the time-series by minimizing the forecast error. Time-series data collected in many situations are hierarchical in structure. A time-series can often be disaggregated in a hierarchical structure depending upon some attributes such as geographical location, product type etc. Thus, a hierarchical time-series is a collection of several time-series data that are correlated in a hierarchical manner. By contrast, a collection of time-series that are aggregated in a number of non-hierarchical ways, are called a grouped time-series. Zotteri *et al*. (2005) studied the impact of aggregation level on forecasting performance.

A cross-sectional hierarchical structure is an arrangement of items in which the items are ordered above, below or on the same level as others. For example, the national economic account is divided into production, income and outlay and capital transaction. Production is further disaggregated into production in India and production in the rest of world; income, outlay and capital transactions each further can be classified into persons, companies, public corporations, general government and rest of world. It is an example of hierarchical time-series since here the order of disaggregation is unique. In demographic forecasting, the infant mortality count in India can be grouped by gender; again, within each gender, mortality counts can be further classified according to geographic location, e.g. state. This second example is called a grouped time-series where the order of disaggregation is

not unique. The infant mortality counts in India can also be first disaggregated by states and then by genders. Therefore, the order is not important.

Existing approaches to hierarchical time-series forecasting usually involves top-down method, bottom-up method, middle-out method and optimal combination method. The top-down approach involves first generating forecasts for the aggregated series at the top most level of the hierarchy, then disaggregates the forecasts in the lower level series based on different types of proportions viz. historical proportion, proportions of historical averages and forecasted proportion. The bottom-up method follows the reverse approach, i.e., forecasting the lowest level series of the hierarchy and then aggregate the base forecasts to obtain the forecasts at the higher level of the hierarchy. Middle-out approach starts forecasting at an intermediate level of the hierarchy selected by the user and following bottom-up approach for the upper level forecasting and top-down approach for the lower level forecasting. Hyndman *et al.* (2011) proposed a statistical method to optimally combine hierarchical forecasting as well. The optimal combination approach provides forecasting for all the series in the hierarchy and then a regression model is used to obtain optimum (in terms of variance forecasts which is also unbiased. This method results the revised forecasts having some desirable properties which are not generally found in forecasts from other approaches. An application of above methodology can be found in Pal and Paul (2016) with respect to forecasting of sorghum production in India.

Indian agriculture has made considerable progress, particularly in respect of food crops such as wheat and rice in irrigated areas; however, performance has not been so good in case of other crops particularly oilseeds, pulses, and coarse cereals. Therefore, after achieving self-sufficiency in food grains the government is focusing attention on these agricultural commodities. The oilseed sector has been an important area of concern and interventions for Indian policy makers in the post-reforms period when India became one of the largest importers of edible oils in the world, importing about half of domestic requirement in the 1990's.

Oilseeds are grown in almost all the parts of India. India occupies a prominent position, both in acreage and production of oilseed crops and oilseed sector occupies an important position in the country's economy. India accounts for 12-15 per cent of world oilseed area, 6-7% of the global production of vegetable oil, and nearly 7 percent of protein meal. The main oilseed based products are edible vegetable oil and oil-cake. However, there are some restrictions on the export of oilseeds in order to meet the increasing demands of the country. In 1950-51, area and production under groundnut was 4.5 million hectares and 3.4 million tonnes respectively. By 1996-97 it increased up to 7.8 million hectares and 9 million tonnes respectively. Soybean is the most important oilseed crop followed by rapeseed-mustard and then groundnut. Major oilseed producing states are Andhra Pradesh, Madhya Pradesh, Maharashtra, Gujrat, Karnataka, Uttar Pradesh etc.

Pulses are the one of the most important agricultural crop in India. They are cultivated across all over the India. India is the world's largest producer and consumer of pulses accounting about 27% of total production and 30% of total consumption in the world. India occupies 35% area of the world and produce 27% of the world production in pulses. Pulse is a good source of protein, fiber and some essential nutrients. Pulses are mainly grown in the rabi season but can also be grown in kharif season also. Gram, lentil, green gram are some example of rabi pulses where arhar, urdbean, moongbean are some important khari pulse. The major pulse producing states are Andhra Pradesh, Bihar, Haryana, Karnataka, Madhya Pradesh, Maharashtra, Odisha, West Bengal etc.

METHODOLOGY

Consider a multi-level hierarchy, where level 0 (zero) denotes the completely aggregated series, level 1 the first level of disaggregation, down to level K containing the most disaggregated time series. A sequence of letters is used to identify the individual series and the level of hierarchy. For example: A denotes series A at level 1; AB denotes series B at level 2 within series A at level 1 and so on.

For example, let us assume that there are three levels in the hierarchy and each group at each level consists of three series. Therefore, in this case, K=3 and the structure of hierarchy is shown in Fig.1.

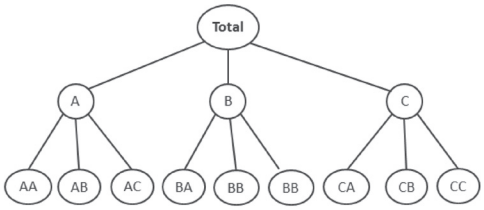

Fig. 1: A three structure hierarchical tree diagram

It is assumed that observations are recorded at times $t = 1,2, \ldots, n$, and one is interested in forecasting each series at each level at times $t = n + 1, n+2, \ldots, n+h$. It will sometimes be convenient to use the notation X to refer to a generic series within the hierarchy. Observation on series X are written as $y_{X,t}$. Thus, $y_{AB,t}$ is the value of series AB at time t. y_t denotes aggregate of all series at time t. Therefore,

$$y_t = \sum_i y_{ij,t},\ y_{i,t} = \sum_j y_{ij,t},\ y_{ij,t} = \sum_k y_{ijk,t},\ y_{ijk,t} = \sum_l y_{ijkl,t},$$

and so on. Thus, observations at higher levels are obtained by adding up the series below.

Let denotes the total number of series at level i(i=0,1,2,…,K) subject to the constraint, $m_i > m_{i-1}$, then the total number of series in the hierarchy is $m=m_0+m_1+\ldots+m_K$. In the above example m=13.

Let $\boldsymbol{y}_{i,t}$ denotes the vector of all observations at level i and time t and $\mathbf{y}_t = \left[y_t, \mathbf{y}_{1,t}, \ldots, \mathbf{y}_{k,t} \right]'$

$$\mathbf{y}_t = \mathbf{S}\mathbf{y}_{k,t} \tag{1}$$

where **S** is a "summing" matrix of order used to aggregate the forecasts of the lowest level series. In the above example, $\boldsymbol{y}_t = \left[y_t, y_{A,t}, y_{B,t}, y_{C,t}, y_{AA,t}, y_{AB,t}, \ldots, y_{CC,t} \right]'$ and the summing matrix is of order 13 × 9 and is given by

$$\boldsymbol{S} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The rank of **S** is m_k.

In hierarchical forecasting, the interest lies on working with forecasts rather than the actual observations of each series. Suppose $\hat{y}_{X,n}(h)$ that be the h-step-ahead forecasts for the series y_X. A sample of $t = 1,2, \ldots, n$ is used to generate the forecasts. Therefore, $\hat{y}_{AA,n}(h)$ denotes the h-step-ahead base forecast of series y_{AA} using the sample $y_{AA,1}, y_{AA,2}, \ldots, y_{AA,n}$. For level i, all h-step-ahead base forecasts will be represented by $\hat{\mathbf{y}}_{i,n}(h)$ and the h-step-ahead base forecasts for the whole hierarchy are given $\hat{\mathbf{y}}_n(h)$, which contains all of the base forecasts stacked in the same sequence as $\mathbf{y}_t$.

Using this notation, all existing hierarchical methods can be represented by the general form

$$\tilde{\mathbf{y}}_n(h) = \mathbf{SP}\hat{\mathbf{y}}_n(h) \tag{2}$$

where **S** is the summing matrix of order $m \times m_k$, as in Eq. (1), and is a matrix of order $m_k \times m$. The form of **P** differs depending on the hierarchical forecasting approach.

THE BOTTOM-UP APPROACH

The bottom-up method is one of the commonly used methods of hierarchical forecasting. The bottom-up method provides first independent base forecasts for most disaggregated series at the lowest level of the hierarchy and then aggregate these base forecasts upwards to obtain revised forecasts for rest of the series in the hierarchy. As an example, let us consider the hierarchy shown in the Fig. 1, after obtaining the h-step-ahead independent forecasts for each of the bottom level series namely $\hat{y}_{AA,n}(h), \hat{y}_{AB,n}(h), \ldots, \hat{y}_{CC,n}(h)$ aggregate these forecasts upwards to obtain the h-step-ahead forecasts for the whole hierarchy as follows:

$$\tilde{y}_A(h) = \tilde{y}_{AA}(h) + \tilde{y}_{AB}(h) + \tilde{y}_{AC}(h)$$
$$\tilde{y}_B(h) = \tilde{y}_{BA}(h) + \tilde{y}_{BB}(h) + \tilde{y}_{BC}(h)$$
$$\tilde{y}_C(h) = \tilde{y}_{CA}(h) + \tilde{y}_{CB}(h) + \tilde{y}_{CC}(h)$$
$$\tilde{y}(h) = \tilde{y}_A(h) + \tilde{y}_B(h) + \tilde{y}_C(h)$$

For the bottom–level series the revised forecasts are same as the base forecasts (i.e. $\tilde{y}_{AA}(h) = \hat{y}_{AA}(h)$).

The general form of this approach is represented as

$$\mathbf{P} = \left[\mathbf{0}_{m_k \times (m-m_k)} / \mathbf{I}_{m_k} \right] \tag{3}$$

where $\mathbf{0}_{i \times j}$ is the $i \times j$ null matrix. The role of **P** is to aggregate the revised forecasts of bottom level

series and to produce the revised forecasts for the whole hierarchy. Since the bottom-up method uses the most disaggregated bottom level series data for modeling, no information is lost due to aggregation.

TOP-DOWN APPROACH

The other commonly applied method in hierarchical forecasting is the top-down approach (Widiarta *et al*., 2007). The top-down approach first generates base forecasts of the "Total" series and then disaggregates these forecasts downwards based on the appropriate proportions of the data (Athanasopoulos *et al*., 2009). The general form of this approach is defined as

$$\mathbf{P} = \left[\mathbf{p} / \mathbf{0}_{m_k \times (m-1)} \right] \tag{4}$$

where $\mathbf{p} = \left[p_1, p_2, ..., p_{m_k} \right]'$ are a set of proportions for the bottom level series. $\mathbf{p}$ gives the distribution of the base forecasts of the "Total" series as revised forecasts for the bottom level of the hierarchy. So the role of $\mathbf{P}$ here is to disaggregate the top level forecasts to obtain the forecasts for the bottom-level series. In top-down approaches the top level revised forecasts is equal to the top level base forecasts, i.e., $\hat{y}_t(h) = \tilde{y}_t(h)$.

THE MIDDLE-OUT METHOD

Here, at the first step, forecasts are generated at an intermediate level of hierarchy and then disaggregated these forecasts to obtain revised forecasts at lower levels and aggregated for computing revised forecasts at higher levels of the hierarchy. Thus, the middle-out approach is a combination of both bottom-up and top-down approaches. In practice, production houses apply this method to study demand forecasting (Lo *et al.*, 2008). At first, base forecasts are produced for all the series of the selected middle level. Then, aggregate these base forecasts upwards to obtain the revised forecasts for the series above the middle level using bottom-up approach. And then, disaggregate the base forecasts of the middle level series downwards to obtain the revised forecasts for the series below the middle level using a top-down approach.

OPTIMAL FORECASTS USING REGRESSION

The optimal combination approach proposed by Hyndman *et al*. (2011), involves first generating independent base forecasts for each series in the hierarchy. It produces unbiased revised forecasts which are also consistent across the hierarchy, provided that the base forecasts are unbiased. According to this approach all h-step-ahead base forecasts can be expressed by the linear regression model as

$$\hat{\mathbf{y}}_n(h) = \mathbf{S}\boldsymbol{\beta}_n(h) + \boldsymbol{\varepsilon}_h \tag{5}$$

where $\beta_n(h) = E\left[\mathbf{y}_{K,n+h} / \mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_n \right]$ is the unknown mean of the bottom level K and $\boldsymbol{\varepsilon}_h$ has zero mean and covariance-variance matrix $Var(\boldsymbol{\varepsilon}_h) = \boldsymbol{\Sigma}_h$. Then $\boldsymbol{\beta}_n(h)$ the is estimated by considering Eq. (5) as a regression equation, and thereby obtain forecasts for all levels of the hierarchy. If $\boldsymbol{\Sigma}_h$ was known, one can use generalized least squares estimation to obtain the minimum variance unbiased estimate of $\boldsymbol{\beta}_n(h)$ as

$$\hat{\boldsymbol{\beta}}_n(h) = \left(\mathbf{S}'\boldsymbol{\Sigma}^{\dagger}{}_h \mathbf{S} \right)^{-1} \mathbf{S}'\boldsymbol{\Sigma}^{\dagger}{}_h \hat{\mathbf{y}}_n(h) \tag{6}$$

where $\boldsymbol{\Sigma}^{\dagger}{}_h$ is the Moore-Penrose generalized inverse of $\boldsymbol{\Sigma}_h$. A generalized inverse is used because $\boldsymbol{\Sigma}_h$ is often (near) singular due to aggregation involved in y_n. This leads to the following revised forecasts

$$\tilde{\mathbf{y}}_n(h) = \mathbf{S}\hat{\boldsymbol{\beta}}_n(h) = \mathbf{S}\mathbf{P}\hat{\mathbf{y}}_n(h), \text{ where } \mathbf{P} = \left(\mathbf{S}'\boldsymbol{\Sigma}^{\dagger}{}_h \mathbf{S} \right)^{-1} \mathbf{S}'\boldsymbol{\Sigma}^{\dagger}{}_h .$$

ILLUSTRATION

Rabi, kharif as well as total pulses and oilseeds production in different states along with at national level has been collected from Directorate of Economics and Statistics, Ministry of Agriculture, Government of India. The details of the hierarchy structure for oilseeds production data and pulses production data are presented in Table 1. Top level shows the all India oilseeds and pulses production and Level 1 shows production of oilseeds and pulses in kharif and rabi seasons separately in India. Level 2 shows state-wise seasonal oilseeds and pulses production in both the tables. The descriptive statistics of oilseeds and pulses production is reported in Table 2. A perusal of Table 2 indicates that the kharif oilseeds production is more than the rabi production for oilseeds production. The variability in production in two seasons are almost same i.e. just greater than 45% for oilseeds. But the variability is much more in the kharif season's oilseeds production in Maharashtra (CV=80.60) and Madhya Pradesh (CV=86.94) and rabi

(5)

season's oilseeds production of Rajasthan (CV=87.94) and Haryana (CV=75.29). The Table 2 also indicates that rabi pulses production is more than the kharif production. The coefficient of variation of pulses production is as high as 98.24% for Odisha in rabi season, and as low as 16.61% for Uttar Pradesh in kharif season. The variability (CV = 27.13%) in kharif season is also higher than rabi season (CV = 17.96). From the Table 2, it is also clear that most of the series are platykurtic and positively skewed. Fig. 2. and Fig. 3. illustrate the time-series data of area, production and yield of oilseeds and pulses in India, respectively. Fig. 4. depicts the oilseeds production of different series at Level 1 and Fig. 5. shows the pulses production of different series at Level 1. For each of the graphical representation, Y-axis represents production in '000 tonnes and X-axis indicates the time period (year).

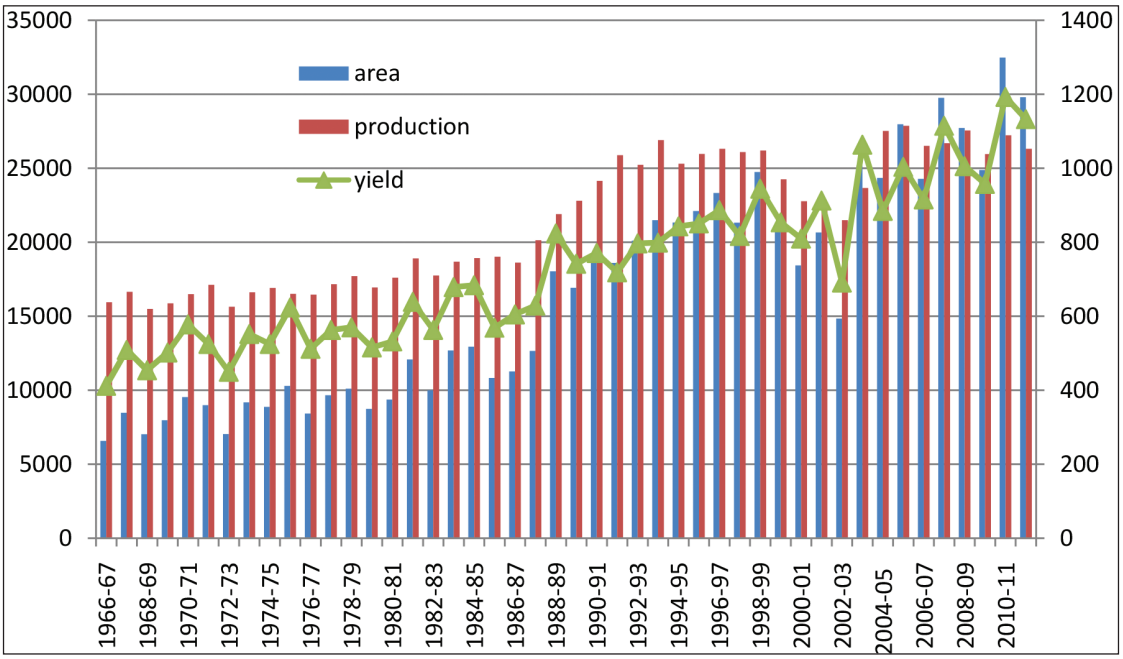

Fig. 1: Area, Production and Yield of Oilseeds in India

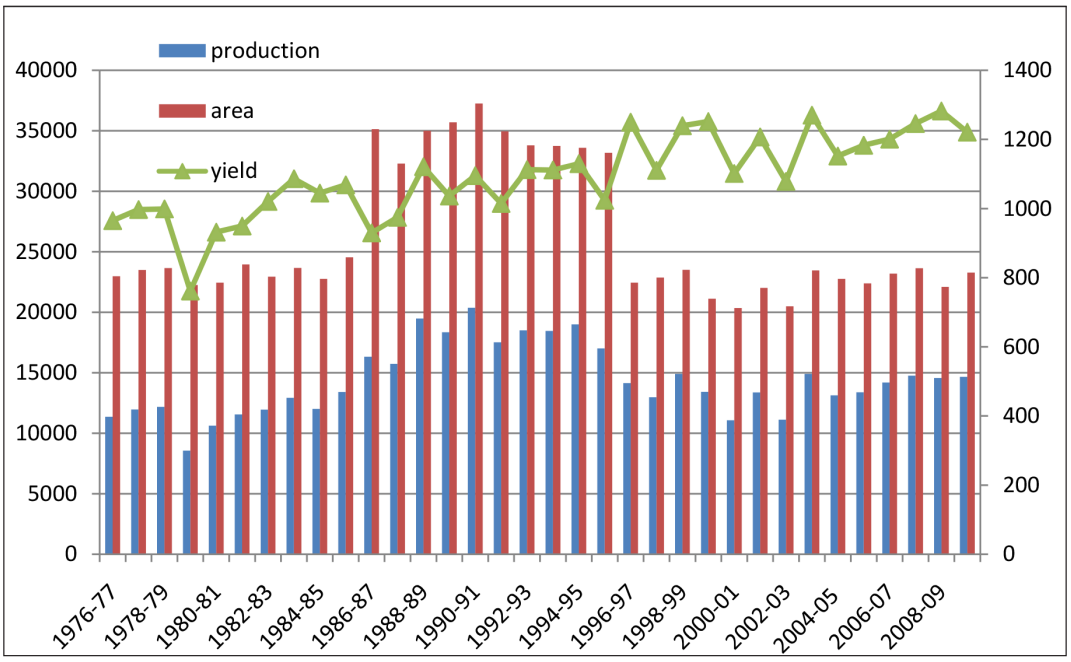

Fig. 2: Area, Production and Yield of Pulses in India

Table 1: Hierarchical structure of oilseeds and pulses production

Oilseeds			Pulses		
Top level			**Top level**		
1	Total	India	1	Total	India
Level 1: Season			**Level 1: Season**		
2	A	Kharif	2	A	Kharif
3	B	Rabi	3	B	Rabi
Level 2: State/Union territory			**Level 2: State/Union territory**		
4	AA	Kharif-Andhra Pradesh	4	AA	Kharif-Andhra Pradesh
5	AB	Kharif-Gujrat	5	AB	Kharif-Gujrat
6	AC	Kharif-Karnataka	6	AC	Kharif-Karnataka
7	AD	Kharif-Madhya Pradesh	7	AD	Kharif-Madhya Pradesh
8	AE	Kharif-Maharashtra	8	AE	Kharif-Maharashtra
9	AF	Kharif-Tamil Nadu	9	AF	Kharif-Odisha
10	AG	Kharif-Others*	10	AG	Kharif-Rajasthan
11	BA	Rabi- Andhra Pradesh	11	AH	Kharif-Uttar Pradesh
12	BB	Rabi-Gujrat	12	AI	Kharif-Others***
13	BC	Rabi-Haryana	13	BA	Rabi- Andhra Pradesh
14	BD	Rabi-Karnataka	14	BB	Rabi-Bihar
15	BE	Rabi-Madhya Pradesh	15	BC	Rabi-Haryana
16	BF	Rabi-Maharashtra	16	BD	Rabi-Odisha
17	BG	Rabi-Rajasthan	17	BE	Rabi-Madhya Pradesh
18	BH	Rabi-Uttar Pradesh	18	BF	Rabi-Maharashtra
19	BI	Rabi-Others**	19	BG	Rabi-Rajasthan
			20	BH	Rabi-Uttar Pradesh
			21	BI	Rabi-Others****
*(including Bihar, Kerala, Orissa, Punjab, Rajasthan, Uttar Pradesh), **(including Assam, Bihar, Odisha, Punjab, West Bengal).			*** (including Bihar, Haryana, Punjab, Tamil Nadu, West Bengal etc.), **** (including Assam, Gujrat, Karnataka, Punjab, Tamil Nadu, West Bengal etc.)		

Table 2: Descriptive statistics of oilseeds and pulses production

Oilseeds					Pulses				
Variable	**Mean**	**CV**	**Skewness**	**Kurtosis**	**Variable**	**Mean**	**CV**	**Skewness**	**Kurtosis**
Total	16531	45.22	0.38	-1.13	Total	14350	20.08	0.40	-0.51
Kharif(A)	10291	46.43	0.79	-0.35	Kharif(A)	5435	27.13	0.84	-0.40
Rabi(B)	6239	47.91	-0.10	-1.45	Rabi(B)	8915	17.96	0.06	-0.24
AA	1248	36.18	0.71	0.30	AA	359	33.54	-0.03	-0.92
AB	2079	54.58	0.73	0.06	AB	391	43.57	0.08	-0.57
AC	769	31.18	0.56	-0.66	AC	525	27.02	0.39	-0.08
AD	2546	86.94	0.56	-1.10	AD	649	31.25	0.63	-0.18
AE	1411	80.60	1.49	1.57	AE	1319	36.17	0.65	-0.14
AF	833	30.80	0.21	-0.54	AF	290	59.01	0.97	-0.16
AG	1402	40.78	1.79	3.50	AG	535	70.40	0.88	0.14
BA	468	62.78	-0.42	-1.08	AH	675	16.61	-0.20	-0.81

BB	373	69.50	-0.15	-1.33	AI	687	38.20	1.30	0.20
BC	451	75.29	0.18	-1.63	BA	611	52.08	0.07	-1.09
BD	349	61.02	-0.28	-1.23	BB	676	42.40	0.90	-0.73
BE	477	51.37	0.25	-1.18	BC	333	78.11	1.23	1.38
BF	413	57.61	0.68	-0.36	BD	481	98.24	0.94	-0.50
BG	1478	87.94	0.67	-0.58	BE	2434	29.31	-0.06	-0.90
BH	1046	21.31	0.23	-0.86	BF	466	57.97	1.19	1.08
BI	1181	44.28	-0.67	-1.06	BG	1022	41.94	0.88	0.61
					BH	1957	27.22	0.52	-0.53
					BI	932	34.82	1.25	0.53

Note: *CV: coefficient of variation*

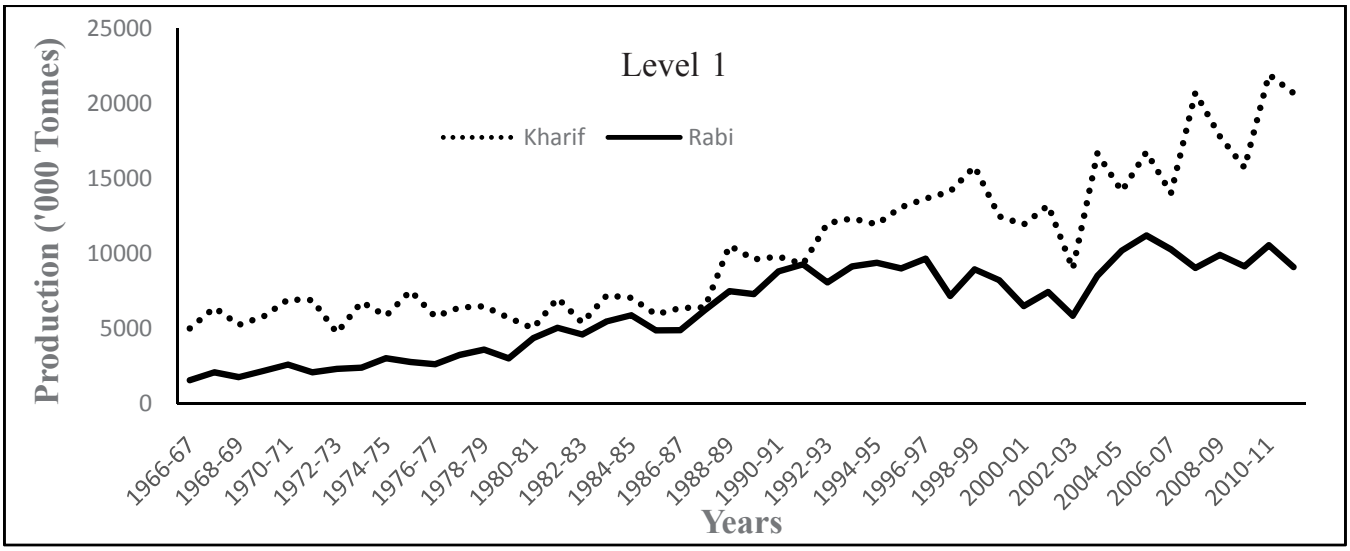

Fig. 3: Hierarchical time-series of oilseeds production at level 1

Fig. 4: Hierarchical time-series of pulses production at level 1

Using Eq. (7), Mean absolute prediction error (MAPE) is obtained using different approaches for each forecast horizon (h) and presented in the Table 3 for oilseeds and pulses production dataset. In Table 3, for a particular h, under any approach, each MAPE value is the average of total MAPE of all the series.

$$\text{MAPE}=\frac{1}{h}\sum_{i=1}^{h}\frac{\left|y_{t+i}-\hat{y}_{t+i}\right|}{y_{t+i}}\times 100 \qquad (7)$$

For each method, the final columns of Table 3 labeled "Average" show average MAPE, across all the forecast horizons. The bold entries identify the minimum among all the average values. The forecasts of oilseeds production and pulses at 1st two levels have been computed for the year 2012-2017 and are reported in Table 4.

Table 3: MAPE for each forecast horizon of oilseeds and pulses production

Methods	Oilseeds						
	Forecast horizon (h)						
	1	2	3	4	5	6	Average
Bottom-up	1.60	13.50	8.69	8.59	9.74	8.38	**8.42**
Top-down	2.86	13.78	10.19	7.78	9.12	7.52	8.54
Middle-out	3.37	12.14	11.63	9.85	10.91	7.42	9.22
Optimal	2.02	13.13	10.12	8.70	9.9	7.27	8.52
	Pulses						
Bottom-up	4.40	2.72	0.59	3.52	6.25	3.46	3.56
Top-down	0.94	1.00	4.94	7.64	5.53	4.93	3.49
Middle-out	0.93	0.62	4.20	6.03	4.76	4.86	4.16
Optimal	0.90	0.19	3.70	6.13	4.76	4.52	**3.37**
ARIMA	0.91	1.53	2.98	8.47	3.95	3.09	3.50

Table 4: Forecasting of oilseeds and pulses production ('000 tonnes) at level 0 and level 1 for 2012-2017

Series	Years					
	2012	2013	2014	2015	2016	2017
	Oilseeds					
Total	29546.24	30470.89	31018.74	31138.98	31238.57	31923.08
Kharif	19989.79	20890.98	21479.84	21393.62	21571.1	22158.17
Rabi	9556.46	9579.91	9538.90	9745.36	9667.47	9764.92
	Pulses					
Total	14468.52	14440.79	14426.01	14420.61	14419.31	14419.63
Kharif	4705.65	4766.45	4814.58	4854.33	4885.82	4910.04
Rabi	9762.87	9674.34	9611.43	9566.28	9533.48	9509.59

CONCLUSION

Investigation reveals that, bottom-up approach outperforms the other approaches of hierarchical time-series model as far as modeling and forecasting of oilseeds production in India, whereas for pulses production, optimum combination has come out to be the best. Two criterions namely MAPE and RMSE are used for validation purpose. For all the approaches, the percentage error is coming out to be less than 10% indicating good performance of the model. A traditional forecasting approach namely ARIMA model has also been performed for pulses production for each of the series. The residuals of the fitted modes were investigated for any presence of autocorrelations but it is seen that the residuals are independent and normally distributed. It ensures that the model building is right and there is no information left in the residuals which can be extracted by use of any statistical model. To this end it may be concluded that instead of using the traditional statistical models, one should use the recently developed most promising statistical models like hierarchical time-series model in order to increase the accuracy of the forecasts.

REFERENCES

Athanasopoulos, G., Ahmed, R.A. and Hyndman, R.J. 2009. Hierarchical forecasts for Australian domestic tourism. *Int. J. of Forecasting*. **25**: 146-166.

Gooijer, J.G.D. and Hyndman, R.J. 2006. 25 years of time-series forecasting. *Int. J. Forecasting*. **22**(3): 443-73.

Hyndman, R.J., Ahmed, R.A. and Shang, H.L. 2011. Optimal combination forecasts for hierarchical time-series. *Computational Statistics and Data Analysis*. **55**(9): 2579-2589.

Hyndman, R.J. and Khandakar, Y. 2008. Automatic time-series forecasting: The forecast package for R. *J. Statistical Software.*, **27**(3): 1-22.

Lo, S., Wang, F. and Lin, J. T. 2008. Forecasting for the LCD monitor market. *J. Forecasting*, **27**(4): 341-56.

Pal, S. and Paul, R.K. 2016. Modelling and Forecasting Sorghum Production in India using Hierarchical Time-Series Models. *Indian J. Agricultural Sciences*, **86**(6): 803–808.

Widiarta, H., Viswanathan, S. and Piplani, R. 2007. On the effectiveness of top-down approach for forecasting autoregressive demands. *Naval Res. Logistics*, **54**(2): 176-88.

Zotteri, G., Kalchmidt, M. and Caniato, F. 2005. The impact of aggregation level on forecasting performance. *Int. J. Prod. Eco.*, **93-94**: 479-91.