ERROR! UNKNOWN DOCUMENT PROPERTY NAME.
Edited by M. Trotman *et al*.

The Use of Seasonal Adjustment Software within the Office for National Statistics

Fida Hussain, Craig H. McLaren and Nigel Stuttard

Abstract

The Office for National Statistics (ONS) is responsible for the derivation and production of a large number of time series estimates. The office has recently developed a software system which enables time series functions to be applied consistently across a wide range of time series. The system uses the X-12-ARIMA software package to carry out seasonal adjustment and forecasting. A number of other time series functions, including interpolation, are carried out by the system using SAS®. This paper gives an overview of the software system and highlights the benefits of an integrated approach that allows time series estimates to be derived in a consistent, timely and transparent way. We also consider future directions for the analysis of time series at the ONS.

Keywords

Seasonal adjustment; forecasting; interpolation

1. Introduction

The Office for National Statistics (ONS) produces a wide range of statistics which are used by policy makers, institutions, academics and the public. These outputs include estimates of Gross Domestic Product, unemployment and price and earnings indices. Published data are usually provided in two forms: non-seasonally adjusted estimates and seasonally adjusted estimates. It is important that the seasonal adjustment process is well understood and that published seasonally adjusted estimates are of high quality.

The ONS has previously announced plans to streamline and modernise software systems for the production of official statistics. The modernisation will improve the quality of published seasonally adjusted outputs by building a common platform that supports state of the art time series packages. Currently, a variety of different time series packages and methodologies are used for carrying out the same tasks across the ONS. The approach to modernisation of statistical systems within the ONS has focused on a number of re-engineering projects. This paper looks at the development of the new software system focusing on the time series functions for seasonal adjustment, forecasting and interpolation used within the National Accounts system.

2. Current time series system for National Accounts

National Accounts is the central framework for the presentation and measurement of the stocks and flows within the economy. Estimates are published in key publications: the United Kingdom Economic Accounts (published quarterly) and the annual United Kingdom National Accounts (the Blue Book). In addition, monthly estimates of output for production and services are also published. Seasonal adjustment is performed on a regular basis each month or quarter. A review of the seasonal adjustment parameter options is performed on an annual basis to ensure the quality of all time series. A centralised database system is currently used to coordinate and manage seasonal adjustment within National Accounts. This system was developed in 1991 and performs a number of functions, including:

1. Serving as a time series database which contains all National Accounts data;
2. Providing the functionality necessary to produce National Accounts, for example aggregation, benchmarking, chain-linking and balancing;
3. Providing a variety of specialist time series tools, including seasonal adjustment (X-11-ARIMA, Dagum 1988), forecasting based on the Holt-Winters modelling approach (Winters, 1960) and interpolation based on an internally developed splining approach (Baxter, 1991).

The broad steps in the current process for reviewing and implementing seasonal adjustment of National Accounts are:

1. Time Series Analyst (TSA) specialists meet with the client to discuss dataset and any potential issues related to seasonal adjustment.
2. TSA, download time series and associated prior adjustments from the central database.
3. TSA seasonally adjust individual time series using X-12-ARIMA to obtain optimal parameter settings and provide the settings to the client.
4. Client area manually sets up analogous X-11-ARIMA settings in central database.
5. Client area performs a production run using concurrent seasonal adjustment.

A similar process is used for other production areas in the ONS. Generally, X-11-ARIMA is used to produce the seasonally adjusted estimates, although the parameter settings used in the seasonal adjustment process are derived by the Time Series Analysis area using the more advanced seasonal adjustment package X-12-ARIMA (Findley, et al. 1998). These parameters include the time series decomposition model, ARIMA model, and prior adjustments to account for any possible outliers and structural breaks. One limitation of the current production framework is that the production area have to manually apply these settings to X-11-ARIMA. This process can be streamlined by the consistent use of a single seasonal adjustment package. This will reduce the risk of human error as there will be less manual manipulation, for example, in uploading prior adjustment factors for individual time series.

The current database system is due to be decommissioned in 2009. A modernised replacement database and system has now been developed to ensure an internally consistent National Accounts dataset from 1997 onwards. Pre-1997 data will be available at the published level.

3. Methodological change: improving time series functions

Improved methodology for time series issues are now available for implementation into ONS production systems. The modernisation program will enable sophisticated methodologies to be implemented so that the quality of the published statistical output is improved.

3.1 Improving methodology for seasonal adjustment

X-12-ARIMA is a leading seasonal adjustment package commonly used by official government agencies for the purposes of seasonal adjustment. For example, X-12-ARIMA is used within the ONS, Bank of England, Statistics Canada, Australian Bureau of Statistics and the United States Census Bureau. Both X-12-ARIMA and TRAMO-SEATS (Gomez and Maravall, 1997) have been endorsed by Eurostat as acceptable seasonal adjustment approaches for official statistics within the European Union.

X-12-ARIMA has a number of features that are improvements over X-11-ARIMA the current production used in the ONS. They include:

1. The use of user-defined regression matrices which allow complicated effects to be estimated and removed. This aids in the identification and estimation of outliers, structural breaks and calendar effects as well as improving the estimation of seasonality.
2. Various methodological improvements, including automatic decomposition and model selection methods, improved statistical diagnostics and more flexible options to determine quality and stability of seasonally adjusted estimates.
3. New functionality including a more flexible benchmarking tool, developed by Statistics Canada (Quenneville, et al. 2004).

See Findley et al., (1998) for more details.

The consistent and integrated use of X-12-ARIMA for all aspects of seasonal adjustment will provide a coherent interface between the Time Series Analysis area that conducts the annual seasonal adjustment review and the client who is responsible for production. Each time series will continue to be reviewed by the Time Series Analysis area annually, and the X-12-ARIMA specification file updated to take into account any changes. Production areas will then use these specification files to seasonally adjust series, individually or in groups, during the production process. Seasonally adjusted aggregates using component time series will be calculated in a separate framework. On occasion the client area may require time series, which do not have a specification file, to be seasonally adjusted. In these circumstances a default specification file will be used. The Time Series Analysis area will retain overall responsibility for the quality assurance of all published seasonally adjusted estimates. Seasonal adjustment reviews will continue to be held on an annual basis, with emerging issues addressed on an as needed basis.

3.2 Improving methodology for forecasts and backcasts

Forecast and backcast estimates are used for a variety of purposes including imputation of missing values, benchmarking, and interpolation (see Section 3.3). In National Accounts, forecasts are needed to impute missing values particularly for the preliminary Gross Domestic Product estimates. In the current time series system, the Holt-Winters methods (Winters, 1960) are usually used to forecast most National Accounts series. Some restricted ARIMA (Autoregressive Integrated Moving Average) models are also available.

The Holt-Winters approach is a specialised case of the more general ARIMA model framework. Cryer and Miller (1991) and Taylor (2003) discuss how the Holt method can be replicated by Box Jenkins ARIMA models. For example, the double exponential smoothing Holt model is equivalent to the ARIMA (0,2,2) model which is given by:

$$y_t = 2y_{t-1} - y_{t-2} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2}$$

where y_t is an observation at time t, θ_i is a parameter estimated from the time series, and ε_t is a white noise process with mean 0 and variance σ_ε^2. See Sclove (2002) for more details.

X-12-ARIMA will not generate exactly the same estimates for the parameters as the current central database because the method of parameter estimation is different. However these parameters within X-12-ARIMA, unlike the parameters used within the current system, are statistically optimal having been derived using maximum likelihood estimation as opposed to a generic grid-search technique using initial start up values used in the current system. Simple extrapolation techniques are available to be used for short times series of less than three years in length.

There is potential for significant improvements in using X-12-ARIMA for all forecasting and backcasting. For example:

1. Tailored forecasting models can be generated for important series allowing the flexibility to obtain an adequate ARIMA model for forecasting purposes only.
2. ARIMA modelling within an X-12-ARIMA framework provides important model diagnostics: whether the model is stationary and invertible, standard errors, confidence intervals and autocorrelation measures. These statistics can aid the user in assessing whether the model chosen is adequate and gauge the uncertainty of a chosen ARIMA model. These diagnostics cannot be generated for the existing Holt-Winters approach within the current forecasting system.
3. X-12-ARIMA's ARIMA functionality enables greater flexibility. In addition to a variety of ARIMA models, the current Holt method and Holt-Winters method (for additive series) can be replicated.
4. ARIMA models are used in the seasonal adjustment process, hence there will be a consistent approach in the use of models.

Figure 1 outlines the proposed approach for forecasting. X-12-ARIMA will automatically identify an ARIMA model that best fits the existing data and then will use the model to forecast future data. The initial step in the proposed system will be to assess whether the time series has an X-12-ARIMA specification file. If so, the file will be used and the pre-selected ARIMA model will be used as the forecasting model. This will ensure consistency of model choice. Where no specification file exists, the means of choosing a model will depend on the periodicity of the series. Annual time series, for which no seasonal adjustment specification file will exist, will be forecast using the ARIMA (0,2,2) model which is equivalent to the non-seasonal Holt model. The quarterly or monthly time series will be forecast using a default forecasting specification file which will call the adapted TRAMO (Time series Regression with ARIMA noise, Missing values and Outliers) forecasting routine of X-12-ARIMA and will automatically construct an appropriate ARIMA model.

Backcasts of time series are required as part of the interpolation and benchmarking processes. The model used for backcasting a series is the same as that used for forecasting.

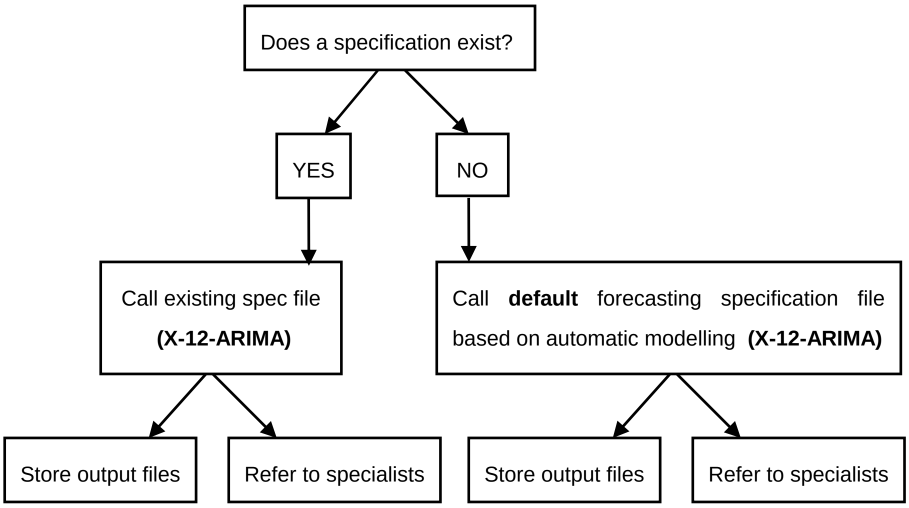

Figure 1. Proposed forecasting approach

3.3 Improving methodology for interpolation

Interpolation is a technique used to produce estimates for indicators that are needed at a higher frequency than are available from the source data, for example, when quarterly or monthly data are required and only annual data are available. The situations in which interpolation is used fall into two separate classes: a) temporal disaggregation where a total or average covering the whole time period needs to be distributed over a shorter time period or b) where a point in time estimate is known for a particular instant in the time period and estimates are required for other instances during that time period. Temporal disaggregation is typically used for flow series whereas point in time interpolations are required for stock series.

In general, the process of interpolation involves initially forecasting and backcasting the low frequency series and then interpolating by fitting a set of continuous curves to the expanded series. When the input data are the total or average for a period and we are interested in temporal disaggregation, then the areas under the curve over the input intervals are constrained to equal the given total. When the input data relate to a point in time, the curve is constrained to pass through the given data points. High frequency time series are then obtained with the additional step of truncating the series to the length of the original series.

The current approach is to use a custom ONS developed package for temporal disaggregation and point in time data (Baxter, 1991). The proposed approach will use an in-built function available within SAS®. version 9.1.3 for both situations. The use of the SAS®. package has a number of advantages including: ability to deal with both temporal disaggregation as well as interpolation in a single function, being methodologically flexible with a range of sophisticated options, complying with the modular design aspects of the modernised system and ensuring production of high quality outputs.

4. Results: Empirical analysis

4.1 Evaluating a default X-11 specification file using X-12-ARIMA

Default settings for X-12-ARIMA specification files were derived from past experience of seasonal adjustment. The Time Series Analysis area has gained expertise from undertaking the seasonal adjustment re-analysis of predominantly ONS data series. This has involved building upon past experience with X-11-ARIMA, and by working with beta versions of X-12-ARIMA supplied by the United States Census Bureau. For example, seasonal adjustment settings have been refined for ONS time series by extending the default list of ARIMA models.

We compared three different versions of X-12-ARIMA: version 0.3 Build 135, version 0.3 Build 169 and X-13A-S version 0.3 Build 128. The most recent release of X-12-ARIMA is available for download from: http://www.census.gov. For X-13A-S, the SEATS component was not used so as to analyse the model based X-11 functionality. Analyses of the final values for the irregular component from the X-11 process showed no difference between the three software versions indicating that, under the chosen default specification and the use of the X-11 process, that all three versions produce identical results. Table 1, shows the range of ARIMA models selected using the available default automatic modelling (AUTOMDL) routine within X-12-ARIMA. A variety of time series of differing lengths were used. This shows that the commonly used airline model is a reasonable fit to a significant proportion of ONS monthly and quarterly time series data.

Monthly		Quarterly	
Chosen model	Percentage	Chosen model	Percentage
$(011)(011)_{12}$	50.4	$(011)(011)_4$	38.7
$(100)(011)_{12}$	11.3	$(100)(011)_4$	9.7
$(010)(011)_{12}$	5.0	$(010)(011)_4$	9.7
$(110)(011)_{12}$	3.2	$(110)(011)_4$	6.5
$(210)(011)_{12}$	2.7	$(210)(011)_4$	3.2
Other models	27.5	Other models	32.2

Table 1: Chosen models for seasonal adjustment runs using automatic modelling in X-12-ARIMA

4.2 Assessing Forecasting and Backcasting

Our primary objective was to ensure that appropriate forecasts could be generated by using the available automatic modelling (AUTOMDL) routine within X-12-ARIMA. The automatic model routine is closely based on the TRAMO methodology of Gomez and Maravall (1997).

Table 2 gives the results of the application of the default specification file and failure rates of the automatic modelling routine against selected ONS time series. These results show that, for the monthly and quarterly series, AUTOMDL builds a suitable ARIMA model using an automatic process, and successfully fits a model in all but a small proportion of cases. However, it fails to produce an output in over one fifth of cases when dealing with annual series. This is because of the difficulty of fitting a model with a limited number of observations. An alternative procedure available within X-12-ARIMA is the use of pick model (PICKMDL) which selects an ARIMA model based on a pre-defined set of ARIMA models. Table 3 shows that the AUTOMDL procedure is preferred when compared to the PICKMDL procedure. The comparison is based on an adequacy criteria which assesses the stationarity and invertibility of the ARIMA model and that there is no serial correlation in the irregular component. The comparison is a summary over 200 time series where each individual series has been fitted using the alternative procedures.

As a result of our trials, we recommend the use of AUTOMDL for model fitting for monthly and quarterly series and the use of a fixed model, ARIMA (0,2,2) for annual series.

Njimi and Melard (2007) conducted a study analysing two automatic forecasting packages which are based on TRAMO methodology, TRAMO-SEATS and TSE-AX. The TRAMO methodology is equivalent to the AUTOMDL procedure available within X-12-ARIMA: They looked at nearly 3000 (annual, quarterly and monthly) series and make the point that even though TRAMO-SEATS was not designed primarily as a forecasting software package, it appears as a "very satisfactory forecasting solution”. Maravall (Pena et al. 2000) applied TRAMO on 35 series, including 22 simulated series, and found that TRAMO identified "either the same model as the one identified by the time series expert or an also acceptable, sometimes better, model".

Periodicity	**Time series used**	**AUTOMDL results**	**Recommendation**
Monthly	265	3 failed to produce output	Use AUTOMDL
Quarterly	175	6 failed	Use AUTOMDL
Annual	370	80 failed	Use ARIMA (0,2,2)

Table 2: Default specification tests using National Accounts data

		Monthly: AUTOMDL		Quarterly: AUTOMDL	
		Adequate	Not adequate	Adequate	Not adequate
PICKMDL	Adequate	59%	4%	84%	9%
	Not adequate	14%	23%	2%	5%

Table 3: Comparison between AUTOMDL and PICKMDL within X-12-ARIMA for selected time series in terms of adequacy of model fit.

4.3 Assessing the seasonal adjustment process

Having established that the differing versions of X-12-ARIMA and the model-based X-11 part of X-13A-S, as outlined in Section 4.1, gave similar results we investigate whether the SEATS seasonal adjustment is at least as good as the model-based X-11 adjustment. Analysis is conducted on default specifications for both X-11 and SEATS so the quality of seasonal adjustment is not assessed on an individual series by series basis. This is on going research.

Revision measures are useful in comparing different seasonal adjustment methods. The basic revision is the difference between the initial seasonal adjustment (concurrent adjustment) and the seasonal adjustment with all the data available at the time of the analysis (i.e. the final adjustment). The percent revision for the seasonal adjustment at time t is defined as $R_t = (A_{t|T} - A_{t|t}) / A_{t|t}$ where $A_{t|t}$ is the seasonal adjustment at time t for data up to time t and $A_{t|T}$ is the seasonal adjustment at time t for the data up to the end of the series at time T. Similar revisions can be generated for the month-to-month changes in the seasonally adjusted data, and the trend estimates.

Table 4 highlights the results for a dataset of the Retail Sales Inquiry. The first two diagnostics are the percentage of months flagged for unstable seasonal factors (S %), and month-to-month changes in the seasonally adjusted estimates (M %). The other measures are revisions based and are the average absolute revisions from the initial to the final estimate of the seasonally adjusted data, change in seasonally adjusted data and trend. In terms of stability of seasonal adjustment, SEATS seasonal adjustment estimates are generally smoother than those for X11 estimates. Both SEATS and X11 adjustments are similar with respect to revisions. SEATS adjustments are preferred in terms of providing a more stable seasonal adjustment in terms of sliding span and revisions analysis. Monsell et al. (2003) assessed a single time series, the Retails Sales of Shoe Stores, using similar criteria and found that the SEATS adjustments were ‘slightly’ more stable than those for the default X-11 adjustment, in terms of revisions and sliding spans diagnostics.

Dataset: MONTHLY start date Jan 1986 (22 Retail Sales Inquiry series)		
Diagnostic	X-11	SEATS
Percentage of unstable seasonal factors	1.81 (*2.60*)	0.54 (*1.02*)
Percentage of unstable month-to-month changes in seasonally adjusted estimates	4.55 (*4.29*)	1.40 (*2.56*)
Average Absolute Revision: seasonally adjusted	0.45 (*0.23*)	0.52 (*0.28*)
Average Absolute Revision: change in seasonally adjusted	0.41 (*0.22*)	0.34 (*0.21*)
Average Absolute Revision: trend	0.89 (*0.36*)	0.89 (*0.41*)

Table 4: Seasonal adjustment diagnostics for X-11 and SEATS. standard deviation in brackets.

4.4 Assessing Interpolation

Approximately fifty National Accounts series were investigated and, on average, the new proposed method gave improved results, i.e. a smoother curve. Figure 2 shows a comparison of the current and proposed methods for an individual series (NHS Pharmaceutical Services). The dashed horizontal lines show the annual figure averaged over four quarters. The differences between the two approaches are small with an average discrepancy of 0.08%.

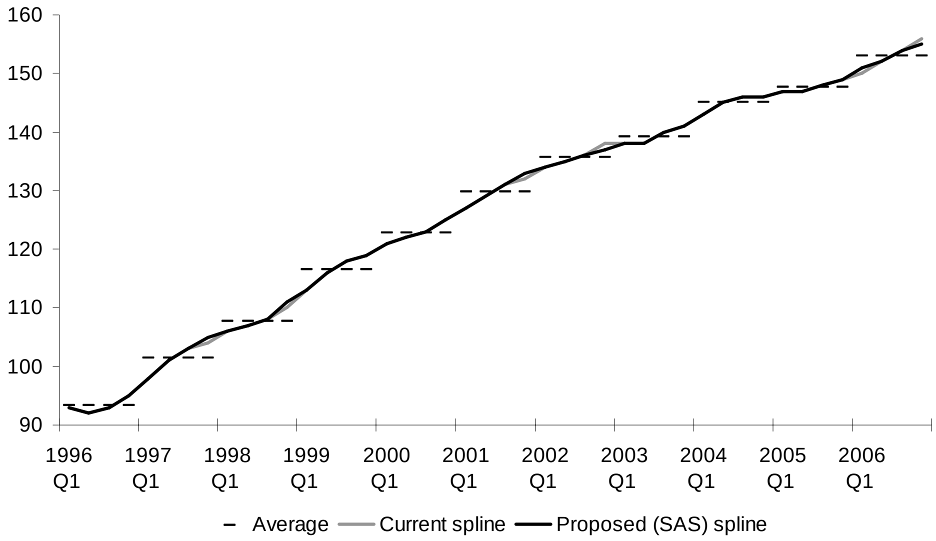

Figure 2: Comparison of Splining techniques (backcast period)

5. Discussion

The proposed modernised infrastructure for time series analysis will provide a consistent and coherent approach for the production of time series estimates. The quality of the final published seasonally adjusted outputs will be improved with the use of improved methodological approaches such as the consistent use of X-12-ARIMA for seasonal adjustment, forecasting and backcasting and the use of SAS for interpolation. The proposed system is planned to be implemented and operational by September 2008. The use of the proposed system will provide a number of challenges. For example:

1. The assessment and implementation of the most appropriate level for seasonal adjustment within National Accounts bearing in mind the need for high quality and the operational needs of National Accounts.
2. Use of default forecasting will need to be closely monitored within the production environment. An inadequate default specification may result in failure to generate forecasts and backcasts.
3. Provision of practical guidelines and training to ensure client areas have an appropriate level of knowledge, understanding and skills to address any time series issues as they occur, particularly in the initial use of the proposed system.

Further developments are planned to continually improve and enhance the ONS approach to seasonal adjustment. There has been a recent detailed review of the seasonal adjustment review process. This has identified areas to focus on which include: development of a dedicated seasonal adjustment analysis system to assist with the management of consistent application of seasonal adjustment parameters, and future consideration and evaluation of the benefits of the X-13A-S approach.

References

Baxter, M.A. (1991) Reduction of Annual Data to Monthly or Quarterly Data. ONS

Box, G.E.P. and Jenkins, G.M. (1970) *Time Series Analysis, Forecasting and Control.* San Francisco: Holden-Day (rev. edn publ. 1976).

Cryer, J.D., and Miller, R.B. (1991) *Statistics for Business: Data Analysis and Modelling.* PWS-Kent (Wadsworth), Boston.

Dagum, E.B. (1980) The X11ARIMA Seasonal Adjustment Method. Ottawa: Statistics Canada, cat. 12-564E.

Findley, D.F., Monsell, B.C., Bell, W.R., Otto, M.C. and Chen B. (1998) New Capabilities and Methods of the X-12-ARIMA Seasonal-Adjustment Program. *Journal of Business and Economic Statistics*, Vol. 16, No. 2., 127-177.

Gomez, V. & Maravall, A. (1997) Programs TRAMO and SEATS : Instructions for the user (beta version: June 1997), Working Paper 97001, Ministerio de Econonomia y Hacienda, Dirrection General de Analisis y Programacion Presupuestaria, Madrid.

Monsell, B.C., Aston, A.D., and Koopman, S.J. (2003) Toward X-13? ASA Proceedings.

Njimi, H., Melard G. (2006) An empirical comparison between two automatic procedures for time series. IAP Statistics Network systems. Brussels University.

Office for National Statistics, (2007) “United Kingdom Economic Accounts”, No. 56-59, 1350-4401, Palgrave Macmillan

Office for National Statistics, (2006) “United Kingdom National Accounts – The Blue Book”, 1-4039-9398-X, Palgrave Macmillan

Pena, D., Tiao, G.C., and Tsay, R.S. (2000) *A course in Time Series Analysis*. Wiley Press pp194

Quenneville, B., Cholette, G., Huot, K., Chiu, and T.D. Fonzo (2004) Adjustment of seasonally adjusted series to annual totals. Research report, Statistics Canada.

Sclove, S. (2002) Some remarks on exponential smoothing, University of Illinois, Chicago, Working Paper.

Scott, S. (2006) Evaluation of X-11 and Model-based Seasonal Adjustment Methods for U.S. Economic Time Series, *Conference on Seasonality, Seasonal Adjustment and their Implications for Short-Term Analysis and Forecasting Conference*, May 2006, Eurostat, Luxembourg.

Taylor, J. W. (2003) Short-term Electricity Demand Forecasting Using Double Seasonal Exponential Smoothing. Journal of Operational Research Society, 2003, Vol.54, pp 799-805.

Winters, P.R. (1960) Forecasting sales by exponentially weighted moving averages. *Man. Sci.***, 6**, 324-342.

About the Authors

Fida Hussain is a Methodologist within Time Series Analysis Branch (London). His work relates to seasonal adjustment, forecasting and interpolation issues. He can be contacted at ONS, Room D210, 1 Drummond Gate, London SW1V 2QQ; Tel: 0207 533 5978; Email: fida.hussain@ons.gsi.gov.uk.

Dr. Craig H. McLaren is Principle Methodologist within Time Series Analysis Branch (Newport). His work relates to seasonal adjustment, forecasting and interpolation issues. He can be contacted at ONS, Room A012, Government Buildings, Cardiff Road, Newport Gwent, NP10 8XG; Tel: 01633 655644; Email: craig.mclaren@ons.gsi.gov.uk.

Nigel Stuttard (corresponding author) is Principle Methodologist within Time Series Analysis Branch (London). His work relates to seasonal adjustment, forecasting and interpolation issues. He can be contacted at ONS, Room D211, 1 Drummond Gate, London SW1V 2QQ; Tel: 0207 533 6324; Email: nigel.stuttard@ons.gsi.gov.uk.