

A Survey of Methods to Interpolate, Distribute and Extrapolate Time Series*

Jose Manuel Pavía-Miralles

Department of Applied Economics, Universitat de Valencia, Valencia, Spain.
E-mail: Jose.M.Pavia@uv.es

Received July 22nd, 2010; revised September 10th, 2010; accepted October 18th, 2010.

ABSTRACT

This survey provides an overview with a broad coverage of the literature on methods for temporal disaggregation and benchmarking. Dozens of methods, procedures and algorithms have been proposed in the statistical and economic literature to solve the problem of transforming a low-frequency series into a high-frequency one. This paper classifies and reviews the procedures, provides interesting discussion on the history of the methodological development in this literature and permits to identify the assets and drawbacks of each method, to comprehend the current state of art on the subject and to identify the topics in need of further development. It would be useful for readers who are interested in the techniques but are not yet familiar with the literature and also for researchers who would like to keep up with the recent developments in this area. After reading the article the reader should have a good understanding of the most important approaches, their shortcomings and advantages, and be able to make an informed judgment on which methods are most suitable for his or her purpose. Interested readers, however, will not find much detail of the methods reviewed. Due to the broadness of the subjects and the large number of studies being referenced, it is provided some general assessments on the methods revised without great detailed analysis. This review article could serve as a brief introduction to the literature on temporal disaggregation.

Keywords: Adjustment, Benchmarking, High Frequency, Quarterly Accounts, Reconciliation, Temporal Disaggregation

1. Introduction

Time series modeling is commonly used by private and public economic and financial analysts in order to forecast future values, analyze the properties of a series, characterize its salient features and monitoring the current status of the economy [1,2]. However, despite social and economic life having quickened and becoming more turbulent, it is not unusual that some relevant variables are not available with the desired timing and frequency. The extra costs associated with collecting data more frequently, some practical limitations to obtain certain variables with a higher regularity along with delays in the process of handling and gathering more frequent records deprive analysts, more often than desired, of the valuable help that high frequency data would provide to perform closer and more accurate short-term analyses. In fact, when temporally aggregated data are used to model, study and investigate the relationship among variables, individuals, and/or entities, it is quite possible that a distorted view of parameters' values, lag structures and sea-

sonal components could be reached [3] and that, as a consequence, poor models and/or forecasts could be obtained and wrong decisions taken [4]. In fact, in econometric modeling, when some of the relevant series are only available at lower frequencies, obvious improvements in model selection, efficiency in parameters' estimates and prediction quality are usually obtained if previously the frequency of the low-frequency time series is increased [5-11].

It is not surprising, therefore, that great quantities of methods, procedures and algorithms have been proposed in the literature in order to increase the frequency of some critical variables, especially in the economic area where it would facilitate a lower delay and a more precise analysis of the condition that an economy or company is experiencing at a given time, making easier to anticipate changes and to react to them.

Obviously, business and economics are not the only fields where it would be useful. Other areas also use these approaches and take advantage of these techniques to improve the quality of their analysis. This paper, however, will focus on the contributions made and used

* Supported by the Spanish MICINN project CSO2009-11246.

within economics; since the problems faced by short-term analysts and quarterly account producers have been the main catalysts for enlargement and improvement of most of the procedures proposed in the literature. For instance, they have been historically linked to the need of solving problems related to the production of coherent quarterly national and regional accounts see, e.g., [12] and [13]. Actually, it is quite probable that some of the future fruitful developments in this subject come to solve the new challenges posed in these areas. Additionally to the methods specifically proposed to estimate quarterly and monthly accounts, another significant number of methods suggested to estimate missing observations have been also adapted to this issue. Thus, classifying the large variety of methods proposed emerges as a necessary requirement in order to perform a systematic and ordered study of the numerous alternatives suggested. Furthermore, due to differing algorithms can give rise to series with different cyclical, seasonal and stochastic properties - even though they have been derived from the same research field using the same basic information [14], a classification would be in itself a proper tool for a suitable selection of the technique.

Depending on the criteria adopted alternative classifications could be reached. This paper will follow the basic division proposed in [15,16], which divide the temporal disaggregation problems according to the basic information available for estimation. Although other divisions have been suggested in the literature, as for instance in [17] who divide the literature according to the nature of the problems, no categorization avoids the problem of deciding where to place some procedures or methods (which could belong to different groups and whose location turns out as extremely complicated) and it is the belief of the author that the former approach clarifies and makes simpler the exposition. Furthermore, to make quicker and easier to follow the paper the mathematical expressions have been omitted using a narrative style to inform about the logic of each procedure.

Previously to introduce any categorization, the different problems that can appear when disaggregating a time series must be set up. The next section is devoted to it. In particular, the rest of the paper is organized as follows. Section 2 introduces the different types of problems faced on this framework. Section 3 exposes the criteria followed to classify the different alternatives and details the main characteristics of each group. Sections 4 to 7 present the methods and section 8 discusses and concludes.

2. Interpolation, Distribution and Extrapolation Problems

In general, inside this framework and depending on the kind of variable handled (either, flow or stock) two dif-

ferent main problems can emerge: the distribution problem and the interpolation problem. On the one hand, the distribution problem appears when the observed values of a flow low-frequency (LF) series of length N must be distributed among kN values (where k is the number of sub-periods in which each period of low-frequency is divided: For instance, $k = 3$ if a monthly series must be estimated from a quarterly observed series, $k = 4$ if quarterly estimates for yearly data are desired and $k = 12$ if monthly data are required for an annually observed series), such that the temporal sum of the estimated high-frequency (HF) series fits the values of the LF series. It is, if $\mathbf{y} = [y_1, \dots, y_N]'$ represents the $(N \times 1)$ vector of observed LF values and $\mathbf{z} = [z_1, \dots, z_T]'$ the $(T \times 1)$ vector of missing HF values, with $T = kN$; the vectors \mathbf{y} and \mathbf{z} can be related by $\mathbf{y} = \mathbf{Bz}$, with $\mathbf{B} = \mathbf{I}_N \otimes \mathbf{u}$, where \mathbf{I}_N is the identity matrix of order N , \otimes stands for Kronecker's product and \mathbf{u} is a $(1 \times k)$ vector of ones. On the other hand, the interpolation problem consists in generating a HF series with the values of the new series being the same as the ones of the LF series for those temporal moments where the latter is observed. It is \mathbf{u} is equal to either a vector $(1 \times k)$ of zeroes except for the first or the last component that is one, depending on the corresponding high-frequency point time where the LF series is observed.

In both cases, when estimates are extended out of the period covered by the low-frequency series, the problem is called extrapolation. Extrapolation is used, therefore, to forecast values of the high-frequency series when no temporal constraints from short series are available; although, nevertheless, in some cases (especially in multivariate contexts), other different forms of constraints can exist. In these cases, the matrix \mathbf{B} is replaced by a proper design matrix [18].

Furthermore, related to distribution problems, they can be found benchmarking and balancing problems [17] which are mainly used in management and by official statistical agencies to adjust the values of a HF series of ballpark figures (usually obtained employing sample techniques) to a more accurate LF series and other temporal disaggregation problems where the temporal aggregation function is different from the sum function. Anyway, despite the great quantity of procedures for temporal disaggregation of time series being proposed in the literature, the fulfillment of the constraints derived from the observed LF series is the norm in the subject.

3. Criteria of Classification

As has been stated on the introduction many different criteria could be proposed to divide the great amount of proposals that can be found in this framework. A first division could arise attending to the plane from which

the problem is faced, either the frequency or the temporal plane. This division, however, is not well-balanced, since the temporal perspective has been by large more popular. On the one hand, the procedures that deal with the problem from a temporal perspective will be analyzed in Sections 4, 5, and 6. On the other hand, the methods that try to solve the problem from the frequency domain will be introduced in Section 7.

Another possible criterion of division attends to the use or not of related series, usually called *indicators*. Economic events tend to made visible in different ways and to affect many dimensions. The economic series are therefore correlated variables that do not evolve in an isolate way. Consequently, it is not unusual that some variables available in HF could display similar fluctuations than those (expected) for the target series. Some methods try to take advantage of this fact to temporally distribute the target series. Thus, the use or not of indicators has been considered as another criterion of classification. The procedures which deal with the series in an isolated way and compute the missing data of the disaggregated HF series from the temporal plane taking into account only the information given by the objective series are presented in Section 4.

Complementary to the previous methods appear the set of procedures which exploit the economic relationships between indicators and objective series. This group, composed for an extent and varied collection of methods widely used, has an enormous success. As example, it could be cited that the use of procedures based on indicators is the rule among the governmental statistic agencies that use indirect methods to estimate quarterly and monthly accounts [16,19]. These procedures are presented in Section 5.

Finally, a last group has been considered: the Kalman filter approaches. The methods that use the state space representation to estimate the non available values have been grouped in Section 6. The great flexibility that offers the representation of the temporal processes in the state space and the enormous possibilities that these representations present to properly deal with log-trans- formations and dynamic approximations justify its own section, despite all the approaches of this section being classifiable in the previous two sections.

4. Non-Indicators Algorithms

Different approaches and strategies have been employed without using related variables. The first methods proposed were developed without any theoretical justification for the elaboration of quarterly (or monthly) national accounts as mere instruments. They were quite mechanical and generated the HF series imposing some *a priori* desired conditions. Gradually, nevertheless, new methods-theoretically founded on the Autoregressive

Integrated Moving Average (ARIMA) representation of the series to be disaggregated-were progressively appearing, introducing more flexibility in the process.

The design of these primary methods, nevertheless, was already in those first days influenced for the need of solving one issue that appears recurrently in the subject and whose solution must tackle every method suggested to temporal disaggregate a time series: the spurious step problem. To prevent series with undesired discontinuities from one year to the next, the pioneers proposed to estimate quarterly values using figures from several consecutive years. The disaggregation methods proposed by [20-22] were devised to estimate the quarterly series corresponding to year t as a weighted average of the annual values of periods $t-1$, t and $t+1$. They estimate the quarterly series through a fix weight structure. The difference among these methods lies on the election of the weight matrix. [20] calculated the weight matrix by requiring that the estimated series verify some *a priori* "interesting" properties. [21] assumed that the curve of the quarterly estimates is located upon a second degree polynomial that passes by the origin. And [22] extended [21]'s proposal to polynomials with other degrees. Furthermore, following pioneers' ideas, [23] expanded [20] to the case of distributing quarterly or annual series into monthly ones and later [24] provided, in an econometric computer package G, a method to convert annual figures into quarterly series, assuming that a cubic polynomial is fitted to each successive set of two points of the low-frequency series. All these methods, however, are univariate and it was necessary to wait two more decades to find inside this approach a solution for the multivariate problem. Just recently, [25,26] have extended, for both stock and flow variables, [24]'s univariate polynomial method to the multivariate case.

A different approach was proposed in [27] who, using also an ad-hoc mathematical procedure, build the quarterly series by solving an optimization problem. In particular, their method builds the quarterly series as solution of the minimization of the sum of squares of either the first or the second differences of the (unknown) consecutive quarterly values, under the condition that the annual aggregation of the estimated series adds up the available annual figures. Although [27] algorithms mostly reduced the subjective charge of the preceding methods, their way of solving the problem of spurious steps was still a bit subjective and therefore it was not free of criticism. On the particular, it could be consulted, among others [6,14,15,28].

In the same way that polynomial procedures were generalized, [27] approach was also extended in both flexibility and in the number of series to be handled. [29] extended [27] work introducing flexibility in the process

in a double way. On one hand, they dealt with any pair of possible combination of high and low frequencies. And, on the other hand, they considered the minimization of the sum of the squared of the i th differences between successive sub-period values (not only first and second differences). The multivariate extension, nevertheless, was introduced later, in [30].

Additionally to the abovementioned ad-hoc mathematical algorithms, many other methods could be also classified within this group of techniques. Among them, it will be highlighted [31] and [32]. [31] assumed that the target series is observed in its higher frequency in a part of the sample period (generally at the end of the sample) and proposed to use this subsample to estimate the non-available values employing the temporal characteristics of the series obtained from it. This strategy, however, as proved by [33], yields inefficient estimates. [32], on the other hand, proposed to obtain the target series minimizing a squared function defined by the inverse covariance matrix associated to the quarterly stationary ARMA(p,q) process obtained taking differences on the non-stationary ARIMA(p,d,q) one (an ARIMA process with autoregressive order p , integrated order d , and moving average order q). In particular, they suggested adjusting an ARIMA model to the low-frequency series, selecting a model for the quarterly values compatible with the annual model and minimizing the d th differences of the objective series by the loss function using the annual series as constraint. Unfortunately, according to [34], [32] method only performs well when the series are long enough to permit a proper estimation of the ARIMA process. [32], however, made possible to reassess [27]. They showed [27]'s algorithm being equivalent to use their procedure assuming that the series to be estimated follow either a temporal integrated process of order one or two (*i.e.*, I(1) or I(2)).

To know the advantages and disadvantages of many of the methods introduced in this section and decide which to use under what circumstances, it could be consulted [34,35]. They performed a simulation and a real-data exercise in which [20,21,27,36]—in its variant without indicators [32,37] and Lagrange polynomial interpolators were compared.

5. Based Method Indicators

Procedures based on related variables have been the most popular and the most widely used and successful. Thus, a great number of procedures can be found within this category. Comparing with the algorithms not-using indicators, related variable procedures have been assigned two main principal advantages: (i) they present better foundations in the construction hypothesis (which can comparatively affect the results validation); and, (ii) they

make use of relevant economic and statistical information, being more efficient. Although, in return, the resulting estimates depend crucially on the indicators chosen and, as [38] observed, they hide an implicit hypothesis according to which the LF relationship is accepted to be maintained in the HF.

The first drawback implies that a special care should be taken in selecting indicators. This problem, however, far from being closed has been opened during decades. Already in 1951, [39] tried to establish some criteria the indicators should fulfill, and this issue has been repeatedly treated again and again [38,40,12,30]. It, however, does not yield that some universal accepted sound criteria have been proposed. Indeed, although the use of indicators to predict the probable evolution of key series throughout the quarters of a year is the rule among the countries that estimate quarterly accounts by indirect methods there are no apparent fixed criteria to select them. A summary of the indicators used by the different national statistics agencies can be found in [12] and a list of features that a variable should fulfill to be selected as indicator is pointed out in [13].

In order to better manage the large quantity of procedures classified in this category, this section has been divided in three subsections. The first subsection is devoted to those procedures, called benchmarking techniques, which given an initial estimated HF series adjust their values using some penalty function in order to fulfill the LF constraints. Subsection 2 presents the procedures that take advantage of econometric models to approximate the incompletely observed variables. Some techniques which use dynamic regression models in the identification of the relationship linking the series to be estimated and the (set of) related time series have been also included in this subsection. Finally, subsection three shows those methods, named optimal procedures, which jointly estimate both parameters and HF series combining target LF series and HF indicators and incorporating the LF constraints in the process of estimation, basically [41] and its extensions.

5.1. Benchmarking/Adjusting Algorithms

As a rule, benchmarking and adjusting methods are composed of two stages. In the first step an initial approximation of the objective series is obtained. In the second step the first estimates are adjusted by imposing the constraints derived from the available and more reliable annual series. The initial estimates are reached using either sample procedures or some kind of relationship among indicators and target series. Among those options that use related variables to obtain initial approximations both non-correlated and correlated strategies could be found. The non-correlated proposals (the former from an

historical perspective) do not explicitly take into account the existing correlation between target series and indicators and rapidly lost analysts' favor. [42] can be consulted for a wide summary of these algorithms. The correlation strategies, on the other hand, usually assume a lineal relationship between the objective series and the indicators from which an initial HF series is obtained. Once the initial approximation is available, it is adjusted to make it congruent with the observed LF series. The discrepancies between both LF series (the observed series and the series obtained by aggregation of the initial estimates) are then removed.

A great quantity of benchmarking procedures can be found in the literature. Bassie [43] proposed to distribute annual discrepancies by a structure of fixed weights calculated taking into account the discrepancies corresponding to two successive years and assuming that the weights function follows a third degree polynomial. [43] has been historically applied to series of the French and Italian economy [44-46] and currently Finland and Denmark use variants of this method to adjust their quarterly GDP series [12]. However, it lacks of theoretical justification and, according to [43], the method spawns series with irregularities and cyclical components different to the initial approximations when the annual discrepancies are too big [12].

Vangrevelinghe [47] planned a different approach. His proposal (primary suggested to estimate the French quarterly familiar consumption series) consists of (1) applying [20] to both objective annual series and indicator annual series to obtain, respectively, an initial approximation and a control series, to then (2) modifying the initial estimate by aggregating the discrepancies between the observed quarterly indicator and the control series, using as scale factor the Ordinary Least Squares (OLS) estimator of the linear observed annual model. Later, minimal variations of [46] were proposed by [28] and [49]. [28] suggested obtaining the initial estimates using [27], instead of [20], and [48] proposed generalizing [88] by allowing the weight structure to be different for each quarter and year, with the weight structure obtained, using annual constraints, from a linear model.

One of the most successful methods in the area (not only among benchmarking procedures) is the approach proposed by [36] in 1971. The great attractiveness of methods such as [36] - and also [41] - among analysts and statistical agencies [17], despite using more sophisticated procedures generally yielding better estimates [13], can be explained because as [49] pointed out short-term analysis in general and quarterly accounts in particular need disaggregation techniques being "...flexible enough to allow for a variety of time series to be treated easily, rapidly and without too much interven-

tion by the producer;" and where "the statistical procedures involved should be run in an accessible and well known, possibly user friendly, and well sounded software program, interfacing with other relevant instruments typically used by data producers (i.e. seasonal adjustment, forecasting, identification of regression models,...)".

Denton [36] suggested adjusting the initial estimates minimizing a loss function defined by a square form. Therefore, the choice of the symmetrical matrix determining the specific square form of the loss function is the crucial element in [36]. Denton concentrated on the solutions obtained minimizing the h th differences between the to-be-estimated series and the initial approximations and found [27] as a particular case of his algorithm. Later on, [50] proposed a slight modification to this function family in order to avoid dependence on the initial conditions. The main extensions of [36] nevertheless were reached by [51-55] who made more flexible the algorithm and extended it to the multivariate case.

Hillmer and Trabelsi [51,52] worked on the problem of adjusting a univariate HF series using data obtained from different sampling sources, and found [36] and [50] as particular cases of their proposal. In particular, they relaxed the requirements about the LF series permitting it to be observed with error; although, as compensation, they had to suppose known the temporal structure of the errors caused by sampling the LF series. When benchmarks are observed without error, the problem transforms into minimizing the discrepancies between the initial estimates and the LF series according to a loss function of the square form type [52]. In these circumstances, they showed that the method of minimizing the h th differences proposed by [36] and [50] implies to implicitly admit: (1) that the rate between the variances of the observation errors and the ARIMA model errors of the initial approximation tends to zero; and, (2) that the observation errors follow an $I(h)$ process with either null initial conditions, in [36], or with the initial values of the series of observation errors begin in a remote past, in [50].

In sample survey most time series data come from repeated surveys whose sample designs usually generate autocorrelated errors and heteroscedasticity. Thus, [53] introduced a regression model to take into account it explicitly and showed that the gain in efficiency of using a more complex model varies with the ARMA model assumed for the survey errors. In this line, [56] showed, through a simulation exercise and assuming that the survey error series follows an AR(1) process, that [53] and [57] have great advantages over [36] and that they are robust to misspecification of the survey error model. The multivariate extension of [36], on the other hand, was

introduced under a general accounting constraint system in [54] and [55]. They assumed a set of linear relationships among target variables and indicators from which initial estimates are obtained, to then, applying the movement preservation principle of Denton approach subject to the whole set of contemporaneous and temporal aggregation relationships, reach estimates of all the series verifying all the constraints.

Although [36] and also [55] do not require any reliability measurement of survey error series to be applied, their need in many other [36]'s extensions [51] led [58] to propose an alternative approach. In particular, to overcome some arbitrariness in the choice of the stochastic structure of the high frequency disturbances, [59] and [60] developed a new adjustment procedure assuming that the initial approximation and the objective series share the same ARIMA model. They combined an ARIMA-based approach with the use of high frequency related series in a regression model to obtain the Best Linear Unbiased Estimate (BLUE) of the objective series verifying LF constraints. This approach permits an automatic (which takes a recursive form in [60]) 'revision' of the estimates with each new observation, what introduces an important difference with the other procedures where the estimates obtained for the periods relatively far away from the last period of the sample are in practice 'fixed'. On the other hand, the multivariate and extrapolation extensions of Guerrero's approach were likewise provided by Guerrero and colleagues [58,61]. [61] suggested a procedure for estimating unobserved values of multiple time series whose temporal and contemporaneous aggregates are known using vector autoregressive models; while [58] proposed a recursive approach to estimate current disaggregated values of the series and a method to predict future disaggregated values. In these cases, nevertheless, it should be noted that even though the problem can be cast using a state-space formulation, the Kalman filter approach cannot be applied directly in these circumstances since the usual assumptions underlying Kalman filtering are not fulfilled.

When dealing with economic variables, it is not uncommon to use logarithms or other transformations of original data (for example, most time series become stationary after applying first differences to their logarithms) to achieve better models of time series due to, as [141] showed, "...the failure to account for data transformations may lead to serious errors in estimation". A very interesting variant in this framework therefore emerges when log-transformations are taken. The problem of dealing with log-transformed variables in the temporal distribution framework was first considered by [63], and later treated, among others, in [64] and [1]. However, due to the logarithmic transformation not being additive,

the LF aggregation constraint cannot be directly applied. When only log-transformation are taken, [64] proposed to obtain initial estimates applying the exponential function to the approximations reached using a linear relationship between log-transformation of the target series and indicators, to then in a second step adopt [36] algorithm to get the final values; although according to [47] this last step could be unnecessary as "*the disaggregated estimates present only negligible discrepancies with the observed aggregated values.*" On the other hand, when the linear relationship is expressed in terms of rate of change of the target variable (*i.e.*, using the logarithmic difference), [47] and [65] suggested to obtain initial estimates for the non-transformed values of the objective variable using [66] and further benchmark (using Denton's formula) for either flow or index variable to exactly fulfill the temporal aggregation constraints.

5.2. Econometric Model Approaches

Economic theory stands for functional relationships among variables. The econometric models express those relations by means of equations. Models based on annual data conceal higher frequency information and are not considered sufficiently informative to policy makers. Building quarterly and monthly macro-econometric models is therefore imperative. Sometimes, the frequency of the variables taking part in the model is not homogeneous and expressing the model in the lower common frequency almost never offers an acceptable approximation, due to, as [67,68] showed, it is preferable to estimate the missing observations simultaneously with the econometric model rather than previously interpolated the unavailable values to directly handle the high-frequency equations. Thus, putting the model in the desired frequency and use the same model, not only to estimate the unknown parameters but also to estimate the non-observed values of the target series, represents in general a good alternative.

There are a lot of econometric models susceptible of being formulated. And therefore, many strategies may be adopted to estimate the missing observations [7,10, 57,69-79].

Drettakis [69] formulated a multiequation dynamic model about the United Kingdom economy with one of the endogenous variables observed only annually for a part of the sample and obtained estimates for the parameters and the unobserved values by Maximum Likelihood (ML) with complete information. [70] extended [69] to the case in which the number of unobserved series is higher than one and introduced an improvement to reduce the computational charges of the estimation procedure. The use of ML was also followed in [74], [75] and [10]. As example, [10] derived the ML estimator

when data are subject to different temporal aggregations and compared its sample variance with those obtaining after applying to [74,75] estimators Generalized Least Squares (GLS) and Ordinary Least Squares (OLS). On the other hand, GLS estimators were employed by [71,76,78] for models with missing observations in the exogenous variables and, therefore, probably with a heteroscedastic and serially correlated disturbance term.

In the extension to dynamic regression models, the ML approach was again used in Palm and Nijman's works. [79] considered a simultaneous equations model, not completely specified, about the Dutch labor market with some variables only annually observed and proposed to obtain initial estimates for those variables using the univariate quarterly ARIMA process derived from the observed annual series. These initial estimates were used to estimate the model parameters by ML. [77] studied the problem of parameters identification and [79] and [9] the estimation one. To estimate the parameters they proposed two alternatives based on ML. The first one consisted of building the likelihood function from the forecast errors, using the Kalman filter. The second alternative consisted of applying the EM algorithm adapted to incomplete samples. This adaptation was developed in a wide and long paper by [73]. [57], on the other hand, presented a general dynamic stochastic regression model, which permits to deal with the most common short-term data treatment (including interpolation, benchmarking, extrapolation and smoothing), and showed that the GLS estimator is the minimum variance linear unbiased estimator [17]. Additionally, although other temporal disaggregation procedures based on dynamic models have been proposed [49,80,81], they will be considered in the next subsection as they can be observed as dynamic extensions of [41]. It must be noted however that they might be also placed on the previous subsection due to they could be seen under the perspective of the classical two-step approach of adjusting algorithms.

5.3. Optimal Procedures

Optimal methods get their name from the estimation strategy they adopt. Such procedures directly incorporate the restrictions derived from the observed annual series into the estimation process to jointly obtain the BLUE of both parameters and quarterly series. To do that, a linear relationship between target series and indicators is usually assumed. This group of methods is one of the most widely used and in fact its root proposal, [41], has served as basis for many statistical agencies [16,82, 83,84] and analysts [85,86,87,88] to quarterly distribute annual accounts and to provide flash estimates of quarterly growth, among other tasks.

There are many links between benchmarking and op-

timal procedure; however, according to [14], "(1) ... compared to optimal methods, adjustment methods make an inefficient (and sometimes, biased) use of the indicators; (2) the various methods have a different capability of providing statistically efficient extrapolation". In contrast, the solution of optimal methods crucially depends on the correlations structure assumed for the disturbances of the linear relationship. In fact, many optimal proposals are only different in that point. All of them, nevertheless, pursue to avoid spurious steps in the estimated series.

Friedman [42] was the first one in applying this approach. He solved the stock variable case obtaining the BLUE of both coefficients and objective series. Nevertheless, [41] were who solved, using a common notation, the interpolation, distribution, and extrapolation problems and wrote the paper probably most influential and cited in this subject. They focused on the case of converting a quarterly series into a monthly one and assumed an AR (1) hypothesis for the disturbances in order to avoid unjustified discontinuities in the HF estimated series. Under this hypothesis, the covariance matrix is governed by the autoregressive coefficient of order one of the HF disturbance series, which is unknown. Hence, to apply the method it has to be previously estimated. [41] suggested exploiting the functional relationship between autoregressive coefficients of order one of the LF and the HF errors to estimate it. Specifically, they proposed an iterative procedure to estimate the monthly AR (1) coefficient from the rate between elements (1, 2) and (1, 1) of the quarterly error covariance matrix.

The [41] strategy of relating the autoregressive coefficients of order one of the high and low error series, however, cannot be completely generalized to any pair of frequencies [89] and consequently several other strategies were followed to solve the issue. In line with [41], [55] obtained for the annual-quarterly case a function between the two autoregressive coefficients. The relation reached by [55], unfortunately, only has unique solution for non-negative annual autoregressive coefficients. Despite it, [90,91] took advantage of such a relation to suggest two iterative procedures to handle the Chow-Lin method with AR (1) errors in the quarterly-annual case. [90] even provided a solution to apply the method when an initial negative estimate of the autoregressive coefficient is obtained. Although, to handle the problem of the sign, [40] had already proposed to estimate the autoregressive coefficient through a two-step algorithm in which, in the first step, the element (1, 3) of the covariance matrix of the annual errors is used to determinate the sign of the autoregressive coefficient. In addition to the above possibilities, strategies based on the maximum likelihood (with the hypothesis of normality for the errors) have been also tried. Examples of this approach can

be found in [46,92,93].

Despite the AR (1) temporal error structure being the most extensively analyzed, other disturbance structures has been proposed. Among the stationary ones, [94] held MA (1), AR (2), AR (4), and a mix between AR (1) and AR (4) as reasonable possibilities to deal with the annual-quarterly case. These complexities, however, were probed as unnecessary in [95], who through a Monte Carlo experiment showed that, although disturbances actually follow other stationary structures, assuming an AR (1) hypothesis on the disturbance term does not significantly influence the quality of the estimates. In regard to the extensions towards no stationary structures, [66] and [96] can be cited. On the one hand, [66], based on [97,98] results, recommended to use [36] and showed that such an approach to the problem is equivalent to use [41] with a random walk hypothesis for the errors. And, on the other hand, [96] studied the problem of monthly disaggregating a quarterly series and extended [41] for the case in which the residual series followed a Markov random walk. [96], however, did not solve the problem of estimating the parameter of the Markov process for small samples. Fortunately, [95] found a solution to this problem and extended [99] to the case of annual series and quarterly indicators.

Despite [55] abovementioned words about the superiority of optimal methods over adjustment procedures, all the previous methods can be, in fact, obtained as solutions of a quadratic-linear optimization problem [63], where the metric matrix that defines the loss function is the inverse of the high-frequency covariance error matrix. Theoretically, therefore, other structures for the disturbances could be easily managed. Thus, for example, in line with [37], the HF covariance error matrix could be estimated from the ARIMA structure of the LF covariance matrix. Despite it, low AR order models are still systematically chosen in practice due to (1) the covariance matrix of the HF disturbances cannot be, in general, uniquely identified from the LF ones and (2) the typical sample sizes occurring in economics usually provide poor LF error matrix estimates [64, 49]. In fact, the Monte Carlo evidence presented in [100] showed that this approach would likely perform comparatively badly when the LF sample size is lower than 40 (a really non infrequent size in economics).

The estimates obtained following [41] are, however, only completely satisfactory when the temporal aggregation constraint is linear and there are no lagged dependent variables in the regression. Thus, to improve accuracy of estimates taking into account dynamics specifications, usually encountered in applied econometrics works, several authors [80,81,101,102] have proposed to generalize [41] (including [66] and [96]) by the use of linear

dynamic models, which permit to perform temporal disaggregation problems providing more robust results in a broad range of circumstances. In particular, [80] following the works initiated by [81,101,102] which were very difficult to implement in a computer program proposed an extension of [41] that is particularly adequate when the series used are stationary or co-integrated [103]. They besides solve the features of the estimation of the first low-frequency period and produce disaggregated estimates and standard errors in a straightforward way. In [104] a MATLAB library to perform it is provided. This library completed the MATLAB libraries, also detailed in [105], that to run [27,36,41,54,66,96,106] offers the Spanish Statistical Official Institute (INE) free of charge [107]. On the other hand, empirical applications of these procedures could be consulted in [49,65], where furthermore a panoramic revision of [80]'s procedure is offered.

The [41] approach and its abovementioned extensions are all univariate, thus to handle problems with more than J (>1) series to-be-estimated, multivariate extensions are required. In this situations, apart from the low-frequency temporal constraints, some additional cross-section, transversal or contemporaneous aggregates among the HF target series are usually available. To deal with this issue, different procedures [41] have been proposed in the literature. [108] was the first who faced this problem. [108] assumed that the contemporaneous HF aggregate of the J series is known and proposed to apply an estimation procedure in two steps. In the first step, he suggested applying [41], in an isolated way, to each one of the J series imposing only the corresponding LF constraint and assuming white noise residuals. In the second step, he proposed to run again [41] procedure, imposing as constraint the observed contemporaneous aggregated series under a white noise error vector, to simultaneously estimate the J series using as indicators the series estimated in the first step. This strategy, however, as [106] pointed out, does not guarantee the fulfillment of the temporal restrictions.

[102], attending to [108] limitation, generalized the [41] estimator and got the BLUE of the J series, fulfilling simultaneously the temporal and the transversal restrictions. Similar to [41], [106] again obtained that the estimated series crucially depend on the structure assumed for the disturbances. Nevertheless, he only offered a practical solution under the hypothesis of errors temporally uncorrelated. That hypothesis unfortunately is inadequate due to it can produce spurious steps in the estimated series. In order to solve it, [109] and [110] introduced a structure for the disturbances in which each one of the J error series follow either an AR (1) process or a random walk with shocks only contemporaneously

correlated. [110], additionally, extended the estimator obtained in [106] to situations with more general contemporaneous aggregations and provided an algorithm to run such so complex disturbance structure in empirical works. Algorithm that is unnecessary under [54]'s simplification, which proposes assuming a multivariate random walk structure for the error vector. Finally, the multivariate extrapolation issue was introduced in [13] extending [109] proposal.

6. The Kalman Filter Strategies

In the study of time series, one of the approaches is to consider the series as a realization of a stochastic process with a particular model generator (e.g., an ARIMA process), which depends on some parameters. In order to predict how the series will behave in a future or to rebuild the series estimating the missing observation it is necessary to know the model parameters. The Kalman filter permits to take advantage of the temporal sequence of the series to implement through a set of mathematical equations a predictor-corrector type estimator, which is optimal in the sense that it minimizes the estimated error covariance when some presumed conditions are met. In particular, it is an efficient recursive filter that estimates the state of a dynamic system from a series of incomplete and noisy measurements. This approach appears very promising within the temporal disaggregation problem due to its great versatility. Moreover, for economists, it presents the additional advantage of making possible that both unadjusted and seasonally adjusted series can be simultaneously estimated.

Among the different approaches to approximate the population parameters of the data generating process it stands out ML. The likelihood function of the stochastic process can be calculated in a relatively simple and very operative way by the Kalman filter. The density of the process, under a Gaussian distribution assumption for the series, can be easily derived from the forecast errors. Prediction errors can be computed in a straightforward way by representing the process in the state space, and the Kalman filter can then be used. In general, the pioneers methods based on the representation in the state space supposed an ARIMA process for the objective series and computed the likelihood of the process through the Kalman filter by employing the smooth point-fixed algorithm (see, e.g., [111] for details in both univariate and multivariate cases) to estimate the not available values.

Despite the representation of a temporal process in the state space not being unique, the majority of the proposals to adapt Kalman filter to manage missing observations can be reduced to the one proposed by [112]. [112] suggested building the likelihood function excluding the

prediction errors associated to those temporal points where no observation exist and proposed to use forecasts obtained in the previous instant to go on running the Kalman filter equations. Among others, this pattern was followed by [111-117]. Additionally to [112] approach, other approaches can be found. [118] developed a new filter and some smooth algorithms which allow interpolating the non observed values with simpler computational and analytical expressions. [119] used state-space models to adjust a monthly series obtained from a survey to an annual benchmark. And, [120] followed the strategy of estimating missing observations considering them as outliers, while [121] introduced a prescribed multiplicative trend in the problem of quarterly disaggregating an annual flow series using its state-space representation.

Jones [112], pioneer in the estimation of missing observations from the state space representation, treated the case of a stock variable which is assumed to follow a stationary ARMA process. Later on, [113], also dealing with stationary series, extended [112] proposal for the case of flow variables. Besides, they adapted the algorithm to the case in which the target series follows a regression model with stationary residuals and dealt with the problem of working with logarithms of the variable. What's more, [113] also extended the procedure to the case of stock variables following non stationary ARIMA processes; although in this case, they compelled the target variable being available in HF for a large enough sample subperiod. In the non stationary case, however, when [113]'s hypothesis is not verified, building the likelihood of the process becomes difficult. Problems in converting the process into stationary and in defining the initial conditions arise. In order to solve it, [114] proposed to consider a diffuse initial distribution in the pre-sample, while [115] suggested transforming the observations in order to define the likelihood of the process. [115]'s transformation made possible to generalize the previous results (including those reached by [113]), although at the cost of destroying the sequence of the series, altering both smoothing and filtering algorithms. Fortunately, [117] went beyond this difficulty and solved it making possible to use the classical tools to deal with the problem of non stationary processes whatever the structure of the missing observations. However, although [115] and [117] extended the issue to the treatment of regression models with non stationary residuals (allowing related variables to be included in this framework), they did not deal with the case of flow variables in an explicit way. [116] was who handled such a problem and extended the solution to non stationary flow series. [116], moreover, suggested using the Kalman filter for the recursive estimate of the non-observed values as a tool to overcome the problem of the change of the estimates due

to the increasing of the available sample. In this line, [122] used information contained in related series to estimate the monthly Swiss GDP from the quarterly series, while [123] estimated a monthly US GDP series from quarterly values after testing several state-space representations to, through a MonteCarlo experiment, identify which variant of the model gives the best estimates. They found the more simple representations did almost as well as more complex ones.

Most of above proposals, however, consider the temporal structure (the ARIMA process) of the objective series known. In practice, however, it is unknown and it is required to specify the orders of the process to deal with it. In order to solve it, some strategies have been followed. Some attempts have tried to infer the process of a HF series from the observed process of the LF one [6,60,116]; while many other studies have concentrated on analyzing the effect of aggregation over a HF process (e.g., among others, [32], [124-126], and, more recently, [127] and [3]) and on studying its effect over stock variables observed in fixed step times (among others, [128], or [11]). Fortunately, the necessary and sufficient conditions under which the aggregate and/or disaggregate series can be expressed by the same class of model was derived by [129].

Both multivariate and dynamic extensions have been also tackled from this framework, although they are just incipient. On the one hand, the multivariate approach started by [111] was continued in [130], who suggested a multivariate seemingly unrelated time series equations model to estimate the HF series using the Kalman filter when several constraints exits. The framework they proposed is flexible enough to allow for almost any kind of temporal disaggregation problems of both raw and seasonally adjusted time series. On the other hand, [131] offered a dynamic extension providing, among others contributions, a systematic treatment of [96], which permits to explain the difficulties commonly encountered in practice when estimating [96]'s model.

7. Frequency Domain Approaches

A great amount of energy has been devoted to deal with the matter from the temporal perspective. Similarly, great efforts have been also devoted from the frequency domain, although they have had less successful and have done less fruits. In this approach, the greatest efforts have been invested on estimating the spectral density function or spectrum of the series, the main tool of a temporal process in the frequency plane. The estimation of the spectrum of the series has been undertaken from both angles: the parametric and the non-parametric perspective.

Both Jones [132] and Parzen [133,134] were pioneers

in the study of missing observations from the frequency domain. They analyzed the problem under a systematic scheme for the observed (and therefore also for the unobserved) values. [132], one of the pioneers in studying the problem of estimating the spectrum, treated the case of estimating the spectral function of a stock stationary series sampled systematically. This problem was also faced by [134] who introduced the term of *amplitude modulation*, the key element in which later spectral developments were based on in their search for solutions. The amplitude modulation defines itself as a zeros and ones series in the sample period. The value of the amplitude modulation is one in those periods where the series is observed, whereas it is zero in case of not being observed.

Different schemes for the amplitude modulation have been considered in the literature. [135] studied the case in which the amplitude modulation followed a Bernoulli random scheme. This random scheme was extended to others by [136] and [137]. More recently, [138] obtained estimators of the spectral function for three types of modulation sequences: determinist, random and correlated random. On the other hand, [139], [140] and [141] followed a different approach; they assumed an ARIMA process and estimated its parameters with the help of the spectral approximation to the likelihood function.

Although the great majority of patterns for the missing observations can apparently be treated from the frequency domain, not all of them have a solution. This fact is a consequence of the impossibility of completely estimating the autocovariances of the process in many practical situations. In this sense, [142] studied the situations in which it is possible to estimate all the autocovariances. On the particular, it must be remembered [139] words: "... (the estimators) are *asymptotically efficient when compared to the Gaussian maximum likelihood estimate if the proportion of missing data is asymptotically negligible*". Hence, the problem of disaggregating an annual time series in quarterly figures is one of those that do not still have a satisfactory solution from this perspective. Likewise, the efforts made to employ the spectral tools to estimate the missing values using the information given by a group of related variables have required so many restrictive hypotheses that its use has not been advisable until now. Nevertheless, from a related approach, [143] have made some advances proposing a method to estimate (under some restrictive hypothesis and in a *continuous way*) a flow variable.

8. Discussion

The large list of the reference section clearly shows that a vast quantity of methods, procedures, and algorithms has been proposed in the literature to deal with the prob-

lem of transforming a low-frequency series (either annual or quarterly) into a high-frequency one (either quarterly or monthly). The first proposals, which built series using pure mathematical ad-hoc procedures, were progressively overcome and gradually strategies based on related indicators were gaining researchers' preferences, with the [41] method, and all its extensions, highlighting. Likewise, interesting solutions have been also suggested representing the series into the state-space and using the Kalman filter to handle the underlying dynamic system. In fact, in my opinion, the great flexibility of this strategy makes it a proper tool to deal with the future challenges to appear in the subject and to handle situations of non-systematic missing observations (non-treated in this paper). The advances made from the frequency domain however do not seem encouraging. None of the approaches, nevertheless, should be discarded rapidly due to, according to [144], pooling estimates obtained from different procedures can improve the quality of the disaggregated series.

An analysis of the historical evolution of the topic, nevertheless, points towards dynamic regression models and techniques using formulations in terms of unobserved component models/structural time series and the Kalman filter as the research lines that will hold a pre-eminent position in the future. On the one hand, the extension of the subject to deal with multivariate dynamic models is still waiting to be tackled; and, on the other hand, the state-space methodology offers the generality that is required to address a variety of inferential issues that have not been dealt with previously. In this sense, both approaches could be combined in order to solve one of the main open problems currently posed in the area: to jointly estimate some high-frequency series of rates when the low-frequency series of rates, some transversal constraints and several related variables are available. For example, it could be used to solve the problem of distributing among the regions of a country the national quarterly rate of growth, when the annual series of regional growth rates are known and several high-frequency regional indicators are available and, moreover, both the regional and sector structure of weights change quarterly and/or annually. In this line, Prioreti's [145] recent paper is a first step.

In addition, a new emerging approach that is taking into account the more recent developments in applied statistical literature (including data mining, dynamic common component analyses, time series models environment and Bayesian modeling) and that takes advantage of the continuous advances in computer hardware and software (making use of large dataset available) will likely turn up in the future as a main line in the subject. Indeed, as [146] point out: "*Existing methods ... are ei-*

ther univariate or based on a very limited number of series, due to data and computing constraints ...until the recent past. Nowadays large datasets are readily available, and models with hundreds of parameters are easily estimated". In this line, [147] dealt with a dynamic factor model using the Kalman filter to perform an index of coincident US economic indicators; [146] modeled a large datasets with a factor model and developed an interpolation procedure which clearly improves univariate approaches that exploits the estimated factors as summary of all the available information; and, [148] proposed a bivariate basic structural model that permits to carry out simultaneously the seasonal and calendar adjustment and the temporal disaggregation.

REFERENCES

- [1] J. G. de Gooijer and R. J. Hyndman, "25 Years of Time Series Forecasting," *International Journal of Forecasting*, Vol. 22, No. 3, 2006, pp. 443-473.
- [2] M. Marcellino and G. L. Mazzi, "Introduction to Advances in Business Cycle Analysis and Forecasting," *Journal of Forecasting*, Vol. 29, No. 1-2, 2010, pp. 1-5.
- [3] J. Casals, M. Jerez and S. Sotoca, "Modelling and Forecasting Time Series Sampled at Different Frequencies," *Journal of Forecasting*, Vol. 28, No. 4, 2008, pp. 316-341.
- [4] A. Zellner and C. Montmarquette, "A Study of Some Aspects of Temporal Aggregation Problems in Econometric Analyses," *The Review of Economic and Statistics*, Vol. 53, No. 4, 1971, pp. 335-342.
- [5] H. Lütkepohl, "Linear Transformations of Vector ARMA Processes," *Journal of Econometrics*, Vol. 4, No. 3, 1984, pp. 283-293.
- [6] Th. Nijman and F. C. Palm, "Series Temporelles Incompletes en Modelisation Macroeconomiques," *Cahiers Du Seminaire d'Econometrie*, Vol. 29, No. 1, 1985, pp. 141-168.
- [7] Th. Nijman and F. C. Palm "Efficiency Gains due to Missing Data Procedures in Regression Models," *Statistical Papers*, Vol. 29, 1988, pp. 249-256.
- [8] Th. Nijman and F. C. Palm, "Consistent Estimation of Regression Models with Incompletely Observed Exogenous Variables," *The Annals of Economics and Statistics*, Vol. 12, 1988, pp. 151-175.
- [9] Th. Nijman and F. C. Palm, "Predictive Accuracy Gain From Disaggregate Sampling in ARIMA Models," *Journal of Business and Economic Statistics*, Vol. 8, No. 4, 1990, 189-196.
- [10] F. C. Palm and Th. Nijman, "Linear Regression Using both Temporally Aggregated and Temporally Disaggregated Data," *Journal of Econometrics*, Vol. 19, No. 2-3, 1982, pp. 333-343.
- [11] A. A. Weiss, "Systematic Sampling and Temporal Aggregation in Time Series Models," *Journal of Econometrics*, Vol. 26, No. 3, 1984, 271-281.
- [12] OECD, "Sources and Methods Used by the OECD Member Countries, Quarterly National Accounts," Paris, OE- CD Publications, 1996.
- [13] J. M. Pavía-Miralles and B. Cabrer-Borrás, "On Estima-

- ting Contemporaneous Quarterly Regional GDP,” *Journal of Forecasting*, Vol. 26, No. 3, 2007, pp. 155-177.
- [14] T. DiFonzo and R. Filosa, “Methods of Estimation of Quarterly National Account Series: A Comparison,” unpublished, [Journée Franco-Italienne de Comptabilité Nationale (Journée de Statistique), Lausanne, 1987, pp. 1-69.
- [15] J. M. Pavía-Mirallas, “La Problemática de Trimestralización de Series Anuales,” Valencia, Universidad de Valencia, 1997.
- [16] Eurostat, “Handbook of Quarterly National Accounts,” Luxembourg, European Commission, 1999.
- [17] E. B. Dagum and P. A. Cholette, “Benchmarking, Temporal Distribution and Reconciliation Methods for Time Series,” New York, Springer Verlag, 2006.
- [18] J. M. Pavía and B. Cabrer, “On Distributing Quarterly National Growth among Regions,” *Environment and Planning A*, Vol. 40, No. 10, 2008, pp. 2453-2468.
- [19] A. M. Bloem, R. J. Dippelsman and N. Maehle, “Quarterly National Accounts Manual. Concepts, Data Sources, and Compilation,” Washington D.C., International Monetary Fund, 2001.
- [20] J. H. C. Lisman and J. Sandee, “Derivation of Quarterly Figures from Annual Data,” *Applied Statistics*, Vol. 13, No. 2, 1964, pp. 87-90.
- [21] S. Zani, “Sui Criteri di Calcolo Dei Valori Trimestrali di Tendenza Degli Aggregati della Contabilità Nazionale,” *Studi e Ricerche*, Vol. VII, 1970, pp. 287-349.
- [22] C. Greco, “Alcune Considerazioni Sui Criteri di Calcolo di Valori Trimestrali di Tendenza di Serie Storiche Annuali,” *Annali della Facoltà di Economia e Commercio*, Vol. 4, 1979, pp. 135-155.
- [23] H. Glejser, “Une Méthode d’Evaluation de Données Mensuelles à Partir d’Indices Trimestriels ou Annuels,” *Cahiers Economiques de Bruxelles*, No. 29, 1966, pp. 45-54.
- [24] C. Almon, “The Craft of Economic Modeling,” Ginn Press, Boston, 1988.
- [25] L. Hedhili and A. Trabelsi, “A Polynomial Method for Temporal Disaggregation of Multivariate Time Series,” Luxembourg, Office for Official Publications of the European Communities, 2005.
- [26] L. Zaier and A. Trabelsi, “Polynomial Method for Temporal Disaggregation of Multivariate Time Series,” *Communications in Statistics-Simulation and Computation*, Vol. 36, No. 3, 2007, pp. 741-759.
- [27] J. C. G. Boot, W. Feibes and J. H. Lisman, “Further Methods of Derivation of Quarterly Figures from Annual Data,” *Applied Statistics*, Vol. 16, No. 1, 1967, pp. 65-75.
- [28] V. A. Ginsburgh, “A Further Note on the Derivation of Quarterly Figures Consistent with Annual Data,” *Applied Statistics*, Vol. 22, No. 3, 1973, pp. 368-374.
- [29] K. J. Cohen, M. Müller and M. W. Padberg, “Autoregressive Approaches to Disaggregation of Time Series Data,” *Applied Statistics*, Vol. 20, No. 2, 1971, pp. 119-129.
- [30] J. M. Pavía, B. Cabrer and J. M. Felip, “Estimación del VAB Trimestral No Agrario de la Comunidad Valenciana,” Valencia, Generalitat Valenciana, 2000.
- [31] H. E. Doran, “Prediction of Missing Observations in the Time Series of an Economic Variable,” *Journal of the American Statistical Association*, Vol. 69, No. 346, 1974, pp. 546-554.
- [32] D. O. Stram and W. W. S. Wei, “Temporal Aggregation in the ARIMA Process,” *Journal of Time Series Analysis*, Vol. 7, No. 4, 1986, pp. 279-292.
- [33] G. C. Chow and A. Lin, “Best Linear Unbiased Estimation of Missing Observations in an Economic Time Series,” *Journal of the American Statistical Association*, Vol. 71, No. 355, 1976, pp. 719-721.
- [34] S. Rodríguez-Feijoo, A. Rodríguez-Caro and D. Dávila-Quintana, “Methods for Quarterly Disaggregation without Indicators: A Comparative Study Using Simulation,” *Computational Statistics and Data Analysis*, 2003, Vol. 43, No. 1, pp. 63-78.
- [35] B. Chen, “An Empirical Comparison of Methods for Temporal Disaggregation at the National Accounts,” 2007. <http://www.fcs.m.gov/07papers/Chen.V-A.pdf>
- [36] F. T. Denton, “Adjustment of Monthly or Quarterly Series to Annuals Totals: An Approach Based on Quadratic Minimization,” *Journal of the American Statistical Association*, Vol. 66, No. 333, 1971, pp. 99-102.
- [37] W. W. S. Wei and D. O. Stram, “Disaggregation of Time Series Models,” *Journal of the Royal Statistical Society, Ser. B*, Vol. 52, No. 3, 1990, pp. 453-467.
- [38] P. Nasse, “Le Système des Comptes Nationaux Trimestriels,” *Annales de L’Insee*, Vol. 14, 1973, pp. 127-161.
- [39] C. G. Chang and T. C. Liu, “Monthly Estimates of Certain National Product Components, 1946-49,” *The Review of Economics and Statistics*, Vol. 33, No. 3, 1951, pp. 219-227.
- [40] J. Bournay and G. Laroque, “Réflexions sur le Méthode d’Élaboration des Comptes Trimestriels,” *Annales de L’Insee*, Vol. 36, 1979, pp. 3-29.
- [41] G. C. Chow and A. Lin, “Best Linear Unbiased Interpolation, Distribution, and Extrapolation of Time Series By Related Series,” *The Review of Economics and Statistics*, Vol. 53, No. 4, 1971, pp. 372-375.
- [42] M. Friedman, “The Interpolation of Time Series by Related Series,” *Journal of the American Statistical Association*, Vol. 57, No. 300, 1962, pp. 729-757.
- [43] V. L. Bassie, “Economic Forecasting,” New York, Mc Graw-Hill, 1958, pp. 653-661.
- [44] ISCO, “L’Aggiustamento delle Stime nei Conti Economici Trimestrali,” *Rassegna del Valori Interni dell’Istituto*, Vol. 5, 1965, pp. 47-52.
- [45] OECD, “La Comptabilité Nationale Trimestrielle,” *Series Etudes Economiques*, Vol. 21, 1966.
- [46] ISTAT, “I Conti Economici Trimestrali dell’Italia 1970-1982,” *Supplemento al Bollettino Mensile di Statistica*, Vol. 12, 1983.
- [47] G. Vangrevelinghe, “L’Évolution à Court Terme de la Consommation des Ménages: Connaissance, Analyse et Prévision,” *Etudes et Conjoncture*, Vol. 9, 1966, pp. 54-102.
- [48] T. DiFonzo, “Temporal Disaggregation of Economic Time Series: Towards a Dynamic Extension,” Luxembourg, Office for Official Publications of the European Communities, 2003.
- [49] J. Somermeyer, R. Jansen and J. Louter, “Estimating Quarterly Values of Annually Known Variables in Quarterly Relationships,” *Journal of the American Statistical Association*, Vol. 71, No. 355, 1976, pp. 588-595.
- [50] P. A. Cholette, “Adjusting Sub-Annual Series to Yearly Benchmarks,” *Survey Methodology*, Vol. 10, No. 1, 1984,

- pp. 35-49.
- [51] S. C. Hillmer and A. Trabelsi, "Benchmarking of Economic Time Series," *Journal of the American Statistical Association*, Vol. 82, No. 400, 1987, pp. 1064-1071.
- [52] A. Trabelsi and S. C. Hillmer, "Bench-marking Time Series with Reliable Bench-Marks," *Applied Statistics*, Vol. 39, No. 3, 1990, pp. 367-379.
- [53] P. A. Cholette and E. B. Dagum, "Benchmarking Time Series With Autocorrelated Survey Errors," *International Statistical Review*, Vol. 62, No. 3, 1994, pp. 365-377.
- [54] T. DiFonzo, "Temporal Disaggregation of System of Time Series When the Aggregates is Known," Luxembourg, Office for Official Publications of the European Communities, 2003 [INSEE-Eurostat Quarterly National Accounts Workshop, Paris-Bercy, R. Barcellan and G. L. Mazzi, Eds., December 1994, pp. 63-78].
- [55] T. DiFonzo and M. Marini, "Benchmarking Systems of Seasonally Adjusted Time Series," *Journal of Business Cycle Measurement and Analysis*, Vol. 2, No. 1, 2005, pp. 84-123.
- [56] Z. G. Chen and K. H. Wu, "Comparison of Benchmarking Methods with and without a Survey Error Model," *International Statistical Review*, Vol. 74, No. 3, 2006, pp. 285-304.
- [57] E. B. Dagum, P. A. Cholette and Z. G. Chen, "A Unified View of Signal Extraction, Interpolation, Benchmarking, and Extrapolation of Time Series," *International Statistical Review*, Vol. 66, No. 3, 1998, pp. 245-269.
- [58] V. M. Guerrero, "Monthly Disaggregation of a Quarterly Time Series and Forecasts of Its Observable Monthly Values," *Journal of Official Statistics*, Vol. 19, No. 3, 2003, pp. 215-235.
- [59] V. M. Guerrero, "Temporal Disaggregation of Time Series: An ARIMA-Based Approach," *International Statistical Review*, Vol. 58, No. 1, 1990, pp. 29-46.
- [60] V. M. Guerrero and J. Martínez, "A Recursive ARIMA-Based Procedure for Disaggregating a Time Series Variable Using Concurrent Data," *Test*, Vol. 4, No. 2, 1995, pp. 359-376.
- [61] V. M. Guerrero and F. H. Nieto, "Temporal and Contemporaneous Disaggregation of Multiple Economic Time Series," *Test*, Vol. 8, No. 2 1999, 459-489.
- [62] D. M. Aadland, "Distribution and Interpolation using Transformed Data," *Journal of Applied Statistics*, Vol. 27, No. 2, 2000, pp. 141-156.
- [63] M. Pinheiro and C. Coimbra, "Distribution and Extrapolation of Time Series by Related Series Using Logarithms and Smoothing Penalties," *Economia*, Vol. 17, October 1993, pp. 359-374.
- [64] T. Proietti, "Distribution and Interpolation Revisited: A Structural Approach," *Statistica*, Vol. 58, No. 47, 1998, pp. 411-432.
- [65] T. DiFonzo, "Temporal Disaggregation Using Related Series: Log-Transformation and Dynamic Extensions," *Rivista Internazionale di Scienze Economiche e Commerciali*, Vol. 50, No. 2, 2003, pp. 371-400.
- [66] R. B. Fernández, "A Methodological Note on the Estimation of Time Series," *The Review of Economics and Statistics*, Vol. 63, No. 3, 1981, pp. 471-478.
- [67] W. R. Vanhonacker, "Estimating Dynamic Response Models when the Data are Subject to Different Temporal Aggregation," *Marketing Letters*, Vol. 1, No. 2, 1990, pp. 125-137.
- [68] J. Jacobs, "'Dividing by 4': A Feasible Quarterly Forecasting Method?" *CCSO Series 22. Groningen: Center for Cyclical and Structural Research*, 2004. <http://www.eco.rug.nl/ccso/CCSO series/ccso22.pdf>
- [69] E. G. Drettakis, "Missing Data in Econometric Estimation," *Review of Economic Studies*, Vol. 40, No. 4, 1973, pp. 537-552.
- [70] J. D. Sargan and E. G. Drettakis, "Missing Data in an Autoregressive Model," *International Economic Review*, Vol. 15, No. 1, 1974, pp. 39-59.
- [71] M. G. Dagenais, "The Use of Incomplete Observations in Multiple Regression Analysis: A Generalized Least Squares Approach," *Journal of Econometrics*, Vol. 1, No. 4, 1973, pp.317-328.
- [72] M. G. Dagenais, "Incomplete Observations and Simultaneous Equations Models," *Journal of Econometrics*, Vol. 4, No. 3, 1976, pp. 231-241.
- [73] A. P. Dempster, N. M. Laird and D. B. Rubin, "Maximun Likelihood from Incomplete Data via the EM Algorithm," *Journal of the Royal Statistical Society, Ser. B*, Vol. 39, No. 1, 1977, pp. 1-38.
- [74] C. Hsiao, "Linear Regression Using Both Temporally Aggregated and Temporally Disaggregated Data," *Journal of Econometrics*, Vol. 10, No. 2, 1979, pp. 243-252.
- [75] C. Hsiao, "Missing Data and Maximum Likelihood Estimation," *Economics Letters*, Vol. 6, No. 3, 1980, pp. 249-253.
- [76] C. Gourieroux and A. Monfort, "On the Problem of Missing Data in Linear Models," *Review of Economic Studies*, Vol. 48, No. 4, 1981, pp. 579-586.
- [77] F. C. Palm and Th. Nijman, "Missing Observations in the Dynamic Regression Model," *Econometrica*, Vol. 52, No. 6, 1984, pp. 1415-1435.
- [78] D. Conniffe, "Small-Sample Properties of Estimators of Regression Coefficients Given a Common Pattern of Missing Data," *Review of Economic Studies*, Vol. 50, No. 1, 1983, pp. 111-120.
- [79] Th. Nijman and F. C. Palm, "The Construction and Use of Approximations for Missing Quarterly Observations: A Model Approach," *Journal of Business and Economic Statistics*, Vol. 4, No. 1, 1986, pp. 47-58.
- [80] J. M. C. Santos Silva and F. N. Cardoso, "The Chow-Lin Method Using Dynamic Models," *Economic Modelling*, Vol. 18, No. 2, 2001, pp. 269-280.
- [81] S. Gregoir, "Propositions pour une Désagrégation Temporelle Basée sur des Modèles Dynamiques Simples," Luxembourg, Office for Official Publications of the European Communities, 2003.
- [82] INE, "Contabilidad Nacional Trimestral de España. Metodología y Serie Trimestral 1970-1992," Madrid, Instituto Nacional de Estadística, 1993.
- [83] ISTAT, "Principali Caratteristiche della Correzione per i Giorni Lavorativi dei Conti Economici Trimestrali," Rome, ISTAT, 2003.
- [84] INSEE, "Methodology of French Quarterly National Accounts," 2004. http://www.insee.fr/en/indicateur/cnat_trim/methodologie.htm
- [85] T. Abeyasinghe and C. Lee, "Best Linear Unbiased Disaggregation of Annual GDP to Quarterly Figures: The Case of Malaysia," *Journal of Forecasting*, Vol. 17, No. 7, 1998, pp. 527-537.

- [86] T. Abeyasinghe and G. Rajaguru, "Quarterly Real GDP Estimates for China and ASEAN4 with a Forecast Evaluation," *Journal of Forecasting*, Vol. 23, No. 6, 2004, pp. 33-37.
- [87] J. M. Pavía and B. Cabrer, "Estimación Congruente de Contabilidades Trimestrales Regionales: Una Aplicación," *Investigación Económica*, Vol. 62, No. 21, 2003, pp. 119-141.
- [88] D. Norman, and T. Walker, "Co-movement of Australian State Business Cycles," *Australian Economic Papers*, Vol. 46, No. 4, 2007, pp. 360-374.
- [89] L. R. Acosta, J. L. Cortigiani and M. B. Diéguez, "Trimestralización de Series Económicas Anuales," Buenos Aires, Banco Central de la República Argentina, 1977.
- [90] J. Cavero, H. Fernández-Abascal, I. Gómez, C. Lorenzo, B. Rodríguez, J. L. Rojo and J. A. Sanz, "Hacia un Modelo Trimestral de Predicción de la Economía Castellano-Leonesa. El Modelo Hispalink CyL," *Cuadernos Aragoneses de Economía*, Vol. 4, No. 2, 1994, pp. 317-343.
- [91] IGE, "Contabilidad Trimestral de Galicia. Metodología e Series Históricas 1980-1991," Santiago de Compostela, Instituto Galego de Estadística, 1997.
- [92] L. Barbone, G. Bodo and I. Visco, "Costi e Profitti in Senso Stretto: un'Analisi du Serie Trimestrali, 1970-1980," *Bolletino della Banca d'Italia*, Vol. 36, 1981, pp. 465-510.
- [93] E. Quilis, "Benchmarking Techniques in the Spanish Quarterly National Accounts," Luxembourg, Office for Official Publications of the European Communities, 2005.
- [94] J. R. Schmidt, "A General Framework for Interpolation, Distribution and Extrapolation of Time Series by Related Series," In: *Regional Econometric Modelling*, Boston, Kluwer Nighoff Pub, 1986, pp. 181-194.
- [95] J. M. Pavía, L. E. Vila and R. Escuder, "On the Performance of the Chow-Lin Procedure for Quarterly Interpolation of Annual Data: Some Monte-Carlo Analysis," *Spanish Economic Review*, Vol. 5, No. 4, 2003, pp. 291-305.
- [96] R. B. Litterman, "A Random Walk, Markov Model for Distribution of Time Series," *Journal of Business and Economic Statistics*, Vol. 1, 1983, pp. 169-173.
- [97] P. Nelson and G. Gould, "The Stochastic Properties of the Income Velocity of Money," *American Economic Review*, Vol. 64, No. 3, 1974, pp. 405-418.
- [98] R. B. Fernández, "Expectativas Adaptativas vs. Expectativas Racionales en la Determinación de la Inflación y el Empleo," *Cuadernos de Economía*, Vol. 13, No. 40, 1976, pp. 37-58.
- [99] J. L. Silver, "Two Results Useful for Implementing Litterman's Procedure for Interpolating a Time Series," *Journal of Business and Economic Statistics*, Vol. 4, No. 1, 1986, pp. 129-130.
- [100] W. Chan, "Disaggregation of Annual Time-Series Data to Quarterly Figures: A Comparative Study," *Journal of Forecasting*, Vol. 12, No. 8, 1993, pp. 677-688.
- [101] E. L. Salazar, R. J. Smith and R. Weale, "Interpolation using a Dynamic Regression Model: Specification and Monte Carlo Properties," *National Institute of Economic and Social Research*, No. 126, 1997.
- [102] E. L. Salazar, R. J. Smith and R. Weale, "A Monthly Indicator of GDP" *National Institute of Economic Review*, No. 161, 1997, pp. 84-89.
- [103] T. DiFonzo, "Constrained Retropolation of Highfrequency Data Using Related Series: A Simple Dynamic Model Approach," *Statistical Methods and Applications*, Vol. 12, No. 1, 2003, pp. 109-119.
- [104] E. Quilis, "Desagregación Temporal Mediante Modelos Dinámicos: El Método de Santos Silva y Cardoso," *Boletín Trimestral de Coyuntura*, No. 88, 2003, pp. 1-11.
- [105] A. Abad and E. Quilis, "Software to Perform Temporal Disaggregation of Economic Time Series," Luxemburg, Office for Official Publications of the European Communities, 2005.
- [106] T. DiFonzo, "The Estimation of M Disaggregate Time Series when Contemporaneous and Temporal Aggregates are Known," *The Review of Economics and Statistics*, Vol. 72, No. 1, 1990, pp. 178-182.
- [107] E. Quilis, "A MATLAB Library of Temporal Disaggregation Methods: Summary," Madrid, Instituto Nacional de Estadística, 2002.
- [108] N. Rossi, "A Note on the Estimation of Disaggregate Time Series when the Aggregate is Known," *The Review of Economics and Statistics*, Vol. 64, No. 4, 1982, pp. 695-696.
- [109] B. Cabrer and J. M. Pavía, "Estimating $J(>1)$ Quarterly Time Series in Fulfilling Annual and Quarterly Constraints," *International Advances in Economic Research*, Vol. 5, No. 3, 1999, pp. 339-350.
- [110] J. M. Pavía-Miralles, "Desagregación Conjunta de Series Anuales: Perturbaciones AR(1) Multivariante," *Investigaciones Económicas*, Vol. 24, No. 3, 2000, 727-737.
- [111] A. C. Harvey "Forecasting, Structural Time Series and the Kalman Filter," Cambridge, Cambridge University Press, 1989.
- [112] R. H. Jones, "Maximum Likelihood Fitting of ARMA Models to Time Series with Missing Observations," *Technometrics*, Vol. 22, No. 3, 1980, pp. 389-395.
- [113] A. C. Harvey and R. G. Pierse, "Estimating Missing Observations in Economic Time Series," *Journal of the American Statistical Association*, Vol. 79, No. 385, 1984, 125-131.
- [114] C. F. Ansley and R. Kohn, "Estimating, Filtering and Smoothing in State Space Models with Incompletely Specified Initial Conditions," *Annals of Statistics*, Vol. 13, No. 4, 1985, pp. 1286-1316.
- [115] R. Kohn and C. F. Ansley, "Estimation, Prediction, and Interpolation for ARIMA Models with Missing Data," *Journal of the American Statistical Association*, Vol. 81, No. 385, 1986, pp. 751-761.
- [116] M. Al-Osh, "A Dynamic Linear Model Approach for Disaggregating Time Series Data," *Journal of Forecasting*, Vol. 8, No. 2, 1989, pp. 85-96.
- [117] V. Gómez and A. Maravall, "Estimation, Prediction and Interpolation for Nonstationary Series with the Kalman Filter," *Journal of the American Statistical Association*, Vol. 89, No. 426, 1994, pp. 611-624.
- [118] P. De Jong, "Smoothing and Interpolation with the State-Space Model," *Journal of the American Statistical Association*, Vol. 84, No. 408, 1989, pp. 1085-1088.
- [119] J. Durbin and B. Quenneville, "Benchmarking by State Space Models," *International Statistical Review*, Vol. 65, No. 1, 1997, pp. 23-48.
- [120] V. Gómez, A. Maravall and D. Peña, "Missing Observations in ARIMA Models. Skipping Strategy versus Addi-

- tive Outlier Approach,” *Journal of Econometrics*, Vol. 88, No. 3, 1999, pp. 341-363.
- [121] G. Gudmundsson, “Disaggregation of Annual Flow Data with Multiplicative Trends,” *Journal of Forecasting*, Vol. 18, No. 1, 1999, pp. 33-37.
- [122] N. A. Cuhe and M. K. Hess, “Estimating Monthly GDP in a General Kalman Filter Framework: Evidence from Switzerland,” *Economic and Financial Modelling*, Vol. 7, No. Winter, 2000, pp. 1-37.
- [123] H. Liu and S. G. Hall, “Creating High-Frequency National Accounts with State-Space Modelling: A Monte Carlo Experiment,” *Journal of Forecasting*, Vol. 20, No. 6, 2001, pp. 441-449.
- [124] T. Amemiya and R. Y. Wu, “The Effect of Aggregation on Prediction in the Autoregressive Model,” *Journal of the American Statistical Association*, Vol. 67, No. 339, 1972, pp. 628-632.
- [125] W. W. S. Wei, “Some Consequences of Temporal Aggregation in Seasonal Time Series Models,” In: *Seasonal Analysis of Economic Time Series*, A. Zellner, Ed., Washington DC, Government Printing Office, 1978, pp. 433-448.
- [126] W. W. S. Wei, “Effect of Systematic Sampling on ARIMA Models,” *Communications in Statistics A*, Vol. 10, No. 23, 1981, pp. 2389-2398.
- [127] R. J. Rossana and J. J. Seater, “Temporal Aggregation and Economic Time Series,” *Journal of Business and Economic Statistics*, Vol. 13, No. 4, 1995, pp. 441-451.
- [128] H. J. Werner, “On the Temporal Aggregation in Discrete Dynamical Systems,” in *System Modeling and Optimization*, R. F. Drenick and F. Kozin, Eds., New York: Springer-Verlag, 1982, pp. 819-825.
- [129] L. K. Hotta and K. L. Vasconcellos, “Aggregation and Disaggregation of Structural Time Series Models,” *Journal of Time Series Analysis*, Vol. 20, No. 2, 1999, pp. 155-171.
- [130] F. Moauro and G. Savio, “Temporal Disaggregation Using Multivariate Structural Time Series Models,” *The Econometrics Journal*, Vol. 8, No. 2, 2005, pp. 214-234.
- [131] T. Proietti, “Temporal Disaggregation by State Space Methods: Dynamic Regression Methods Revisited,” *The Econometrics Journal*, Vol. 9, No. 3, 2006, pp. 357-372.
- [132] R. H. Jones, “Spectral Analysis with Regularly Missed Observations,” *Annals of Mathematical Statistics*, Vol. 33, No. 2, 1962, pp. 455-461.
- [133] E. Parzen, “Mathematical Considerations in the Estimation of Spectra,” *Technometrics*, Vol. 3, No. 2, 1961, pp. 167-190.
- [134] E. Parzen, “On Spectral Analysis with Missing Observations and Amplitude Modulation,” *Sankhyä A*, Vol. 25, No. 4, 1963, pp. 383-392.
- [135] P. A. Scheinok, “Spectral Analysis with Randomly Missed Observations: The Binomial Case,” *Annals of Mathematical Statistics*, Vol. 36, No. 3, 1965, pp. 971-977.
- [136] P. Bloomfield, “Spectral Analysis with Randomly Missing Observations,” *Journal of the Royal Statistical Society, Series B*, Vol. 32, No. 3, 1970, pp. 369-380.
- [137] P. Bloomfield, “An Exponential Model for the Spectrum of a Scalar Time Series,” *Biometrika*, Vol. 60, No. 2, 1973, pp. 217-226.
- [138] C. M. C. Toloï and P. A. Morettin, “Spectral Analysis for Amplitude-Modulated Time Series,” *Journal of Time Series Analysis*, Vol. 14, No. 4, 1993, pp. 409-432.
- [139] W. Dunsmuir, “Estimation for Stationary Time Series When Data are Irregularly Spaced or Missing,” In: D. F. Findley Eds., *Applied Time Series Analysis II*, New York: Academic Press, 1981, pp. 609-649.
- [140] W. Dunsmuir and P. M. Robinson, “Parametric Estimators for Stationary Time Series with Missing Observations,” *Advances in Applied Probability*, Vol. 13, No. 1, 1981, pp. 129-146.
- [141] W. Dunsmuir and P. M. Robinson, “Estimation of Time Series Models in the Presence of Missing Data,” *Journal of the American Statistical Association*, Vol. 76, No. 375, 1981, pp. 456-467.
- [142] W. Clinger and J. W. VanNess, “On Unequally Spaced Time Points in Time Series,” *Annals of Statistics*, Vol. 4, No. 4, 1976, pp. 736-745.
- [143] G. Gudmundsson, “Estimation of Continuous Flows from Observed Aggregates,” *Journal of the Royal Statistical Society, Ser. D*, Vol. 50, No. 3, 2001, pp. 285-293.
- [144] M. Marcellino, “Pooling-Based Data Interpolation and Backdating,” *Journal of Time Series Analysis*, Vol. 28, No. 1, 2007, pp. 53-71.
- [145] T. Proietti, “Multivariate Temporal Disaggregation with Cross-sectional Constraints,” *Journal of Applied Statistics*, 2010, in-Press. DOI: 10.1080/02664763.2010.505952.
- [146] E. Angelini, J. Henry and M. Marcellino, “Interpolation and Backdating with a Large Information Set,” *Journal of Economic Dynamics and Control*, Vol. 30, No. 12, 2006, pp. 2693-2724.
- [147] T. Proietti and F. Moauro, “Dynamic Factor Analysis with Nonlinear Temporal Aggregation Constraints,” *Applied Statistics*, Vol. 55, No. 2, 2006, pp. 281-300.
- [148] T. Proietti, and F. Moauro, “Temporal Disaggregation and the Adjustment of Quarterly National Accounts for Seasonal and Calendar Effects,” *Journal of Official Statistics*, Vol. 24, No. 1, 2008, pp. 115-132.