Is Aggregation Necessarily Bad? (With apologies to Grunfeld and Griliches)

Moonyoung Baek, P. Geoffrey Allen, and Bernard J. Morzuch
Department of Resource Economics, University of Massachusetts Amherst

With high frequency data (e.g., hourly), when decisions are based on lower frequency aggregates (e.g., four-hourly intervals) the possibilities are to aggregate the data then directly forecast the aggregates, or indirectly to estimate the disaggregate series then aggregate the forecasts. No clear principle has emerged concerning this choice, and past empirical work has produced conflicting results with no indication as to what circumstances might favor one approach over the other. Aggregating the data amounts to throwing away information. On the other hand, if the amount of noise in the disaggregate data swamps any additional signal, parameter estimates are more difficult to make. The added uncertainty leads to less accurate forecasts.

For the two series examined, hourly arrivals at a hospital emergency room, and hourly electricity load data, results are fairly consistent. More accurate forecasts are obtained from estimating and forecasting the disaggregate data and aggregating the forecasts compared with estimating and forecasting using aggregate data. The variation over a typical day and the day-to-day variation through the week are greater for the electricity data than for the emergency room arrivals. The differences between the series are apparently not large enough to affect the preferred strategy of disaggregate estimation and forecasting.

Is Aggregation Necessarily Bad? (With apologies to Grunfeld and Griliches)

Introduction

Forecasting is about managing uncertainty. In a practical setting this requires including information that you know for sure and leaving out of the analysis relationships that are difficult to quantify. Because it is relatively parsimonious, the Holt-Winters method is a popular and usually effective way of making forecasts, especially when causal relationships are either unclear or not estimable with available data.

One area where uncertainty is an issue is in situations where data are available at a micro-level but forecasts are needed at a higher level of aggregation. Is it better to "forecast first then aggregate" or "aggregate first then forecast?" The question arises in contemporaneous aggregation, for example across countries, industries, government expenditure categories, and in temporal aggregation, for example from months to quarters. The latter is of interest here.

Temporal aggregation has been studied for many years, though the majority of the literature concerns theoretical issues, not practical applications. When the data generating process (DGP) is known, the result is straightforward: better forecasts result from using the disaggregate information. For an ARIMA process, the order of integration of the aggregated process is the same as the order of integration of the original series, as is the order of the autoregressive process, but this is not true for the moving average component. Generalizing to VARMA processes, Lütkepohl (2005) shows that forecasts from disaggregated flow variables generally have lower MSE than forecasts from the aggregate, although the difference disappears as the forecast horizon goes to infinity. The two forecasts will be equal for small horizons if and only if the disaggregate series has a finite moving average structure of order less than the length of season. This argues for the strategy "forecast first, then aggregate."

In practice, the DGPs have to be specified either by assumption as with the Holt-Winters structure or on the basis of limited sample information. This adds uncertainty. Model parameters must be estimated again using limited sample information, adding further uncertainty. In that case quite different results may be obtained and, in particular, forecasts based on disaggregate processes may be inferior to those based on the aggregate directly. If the sample size is large enough, the part of the forecast MSE due to estimation uncertainty will eventually be so small that the forecast based on the disaggregate is again superior to the corresponding aggregate. If estimated instead of known processes are used, it is possible that forecasts based on disaggregates are worse because the MSE part due to estimation may be larger for the former than for the latter. This situation is in particular likely to occur if the DGPs are such that effciency gains from disaggregation do not exist or are small and the aggregated process has a simple structure which can be captured with a parsimonious model (Lütkepohl, 2005).

Man (2004) performed a detailed analysis of a relatively simple periodic ARMA model. He considered when a misspecified (or "adapted") model on aggregate data gave better forecasts in a mean squared error sense than an adapted model on disaggregated data. He compared this difference relative to the MSE of the true model forecasts. The aggregate approach is preferred when the adapted disaggregate model is badly misspecified, when the aggregate data has weak dependency (or is close to noise), when the aggregate data has little variability, and when the disaggregate moving average model has smaller noise variance at the seasonal lag than at shorter lags. A weak seasonal pattern and badly misspecified disaggregate model favor an aggregate model with simpler structure.

Table 1 summarizes the relatively small number of studies discovered that compare aggregate forecasts on real time series. Most aggregate from monthly to quarterly values and use an ARIMA model. None were found that specifically compared Holt-Winters estimation under temporal aggregation, though of course, HW with no or additive seasonality can be expressed as an equivalent ARIMA model. In addition, there are one or two simulation studies (e.g., Hotta and Neto, 1993, Souza and Smith, 2004). Results are clearly mixed with somewhat more support for the “forecast then aggregate” strategy and no clear indication of when the alternative strategy is preferred. Lütkepohl (1986) considered both temporal and contemporaneous aggregation simultaneously. He showed that, under a broad range of values for number of contemporaneous series and amount of aggregation, temporal aggregation reduces forecasting efficiency while contemporaneous aggregation may improve it. In addition, Koreisha and Fang (2004) reworked Butter’s (1976) data and on differential yield rates and in addition compared updating the quarterly forecast with the aid of the most recent monthly observation. This was highly useful for one-quarter-ahead forecasts but the value of updating tailed off rapidly at longer horizons.

Data

For the hourly emergency-room arrivals, the estimation sample runs from midnight April 3, 2000 to 2300 hours on December 31, 2000 (39 weeks, 6552 observations) and the post-sample period runs from midnight January 1, 2001 to 2300 hours on April 1, 2001 (13 weeks, 2184 observations).

Hourly electricity loads for the five New England states of the United States (Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island and Vermont) are reported by the Independent Service Operator on its website (http://www.iso-ne.com) as the second data set. The estimation sample and post-sample periods are identical to those of the emergency room data. Figures 1 and 2 show the first four weeks of each series. Comparing the two data series, both show strong daily patterns. Electricity loads show stronger weekly patterns than do the emergency room arrivals. Emergency room arrivals sometimes have zero values while electricity loads do not.

In each series, hourly data are aggregated to four-hour blocks. In contrast to the situation analyzed by Man (2004) who considered annual aggregates of quarterly data, we expect the aggregate data to display seasonal patterns in much the same way as did the disaggregate data.

Models

Ord, et al. (1997) devised, and Koehler, et al. (2001), and Hydman, et al. (2002) further developed, a complete state-space form for the standard Holt-Winters structure. With additive seasonal factors and additive error in the observation equation the specification is

$$y_t = l_{t-1} + b_{t-1} + s_{t-m} + \varepsilon_t \tag{1a}$$

$$l_t = l_{t-1} + b_{t-1} + \alpha_1 \varepsilon_t \tag{1b}$$
$$b_t = b_{t-1} + \alpha_2 \varepsilon_t \tag{1c}$$
$$s_t = s_{t-m} + \alpha_3 \varepsilon_t \tag{1d}$$

where $\varepsilon_t \sim N(0, \sigma^2)$; y_t is an observation at time t of the time series of interest; l_t, b_t, s_t are the unobservable level, trend, and season (length m) components; and α_1, α_2, and α_3 are unknown parameters in the state equations.

Under the special assumption of perfect correlation among the error terms, this *innovations state-space model* can be estimated by one of the standard Holt-Winters algorithms. The two benchmark results are based on this standard single seasonality model, with season lengths m =24 and m = 168. While there are several commercial software packages that accommodate season length of 24 we found none that accommodated 168 periods. Results are computed by a custom program written in R that was checked for m = 24 against the output from the Time Series Module in SAS®.

Using m = 168 permits each day to have a different pattern, though the updating of the seasonal factors occurs only once per week. If in addition to a daily seasonal pattern there is also a weekly pattern (for example where weekend days have a different pattern from weekday days) this can be modelled as multiple (in this case double) seasonality. Taylor (2003) was the first to propose such a development, by adding a second set of seasonal factors to the standard H-W structure. In state-space form with additive seasonality this appears as (Gould, et al. 2005):

$$y_t = l_{t-1} + b_{t-1} + s_{1,t-m_1} + s_{2,t-m_2} + \varepsilon_t \tag{2a}$$

$$l_t = l_{t-1} + b_{t-1} + \alpha_1 \varepsilon_t \tag{2b}$$
$$b_t = b_{t-1} + \alpha_2 \varepsilon_t \tag{2c}$$
$$s_{1,t} = s_{1,t-m_1} + \alpha_3 \varepsilon_t \tag{2d}$$
$$s_{2,t} = s_{2,t-m_2} + \alpha_4 \varepsilon_t \tag{2e}$$

where $s_{1,t}$ and $s_{2,t}$ are the seasonal indices with seasonal lengths m_1 (= 24) for $s_{1,t}$ and m_2 (=168) for $s_{2,t}$ ($m_1 \le m_2$). Although more heavily parameterized than HW(168), requiring one additional smoothing parameter and 24 initial values for the shorter set of seasonal factors, Taylor found that his double seasonal exponential smoothing approach outperformed the single seasonal Holt-Winters method, double multiplicative seasonal ARIMA model (Taylor, 2003), an artificial neural network model, a principal-component-analysis based regression model (Taylor, et al., 2006), and a periodic ARMA model. For some data series, it appears that the HW(168) approach suffers by using week-old seasonal factors in the forecasts, rather than day-old.

Gould, et al. (2005) have developed an additive error version of the innovations state-space model with multiple seasonal patterns. The most general version contains seven sets of seasonal factors, each of length 24, one factor for each day of the week.

$$y_t = l_{t-1} + b_{t-1} + \sum_{i=1}^{r} x_{it}\ s_{i,t-m_1} + \varepsilon_t \tag{3a}$$

$$l_t = l_{t-1} + b_{t-1} + \alpha_1 \cdot \varepsilon_t \tag{3b}$$

$$b_t = b_{t-1} + \alpha_2 \cdot \varepsilon_t \tag{3c}$$

$$s_{it} = s_{i,t-m_1} + (\sum_{j=1}^{r} \gamma_{i,j} \cdot x_{jt})\varepsilon_t, \qquad (i = 1, \ldots, r) \tag{3d}$$

There is a 7 x 7 matrix of seasonal smoothing parameters, $\gamma_{i,j}$ each row referring to a the seasonal factors corresponding to a particular day of the week, and each column referring to a day of the week. Only one day's smoothing parameters are used at once, so that, for example, on Monday, the first day of the week, smoothing parameters in column 1 are used to update each set of seasonal factors. Use of the appropriate column is controlled by the dummy variable x_{it} where

$$x_{it} = \begin{cases} 1 & \text{if time period t occurs when sub - cycle i is in effect} \\ 0 & \text{otherwise} \end{cases}$$

In contrast with the previous models, it is now essential to keep track of the day of the week for the current observation in order to apply the correct smoothing parameter to each set of seasonal factors. In the most general case there are $k = \frac{m_2}{m_1} = 7$ sets of seasonal factors. Where different days have a common seasonal pattern, they can be grouped into r common subcycles where $r \le k$. For example r = 2 if all the weekdays are grouped together and the weekend days are in a separate group. This would reduce the number of seasonal factors to 48 and the matrix of smoothing parameters would be 2 x 2.

Independently, restrictions can be placed on the matrix of smoothing parameters. Gould et al. (2005) propose a common diagonal, $\gamma_{ii} = \gamma_1^*$ for all i, and common off-diagonal elements $\gamma_{ij} = \gamma_2^*$ for all i ≠ j. If the off-diagonal parameters are non-zero, then updating occurs in all seasonal factors including those applying to other days of the week. A further restriction is $\gamma_1^* = \gamma_2^*$ and a final one is $\gamma_2^* = 0$. With the last restriction, seasonal factors for all except the current day are carried forward without updating, so that updating occurs once per week, corresponding identically to HW(168).

Multiplicative seasonality versions of all of the above models can be constructed similarly. Multiplicative error versions are also possible, but do not alter the point forecasts, only prediction intervals and statistical tests.

Finally, the assumption of a single source of error can be relaxed to give an unobserved components (UC) form of the HW structure. First, the basic structural model with a dummy variable type single seasonality appears as:

$$y_t = \mu_t + \gamma_t + \varepsilon_t \tag{4a}$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \tag{4b}$$
$$\beta_t = \beta_{t-1} + \zeta_t \tag{4c}$$
$$\gamma_t = -\sum_{j=1}^{s-1} \gamma_{t-j} + \omega_t \tag{4d}$$

where $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$, $\eta_t \sim N(0, \sigma_\eta^2)$, $\zeta_t \sim N(0, \sigma_\zeta^2)$, and $\omega_t \sim N(0, \sigma_\omega^2)$, and s is the season length ($s = 24$ for a daily pattern or $s = 168$ for a weekly pattern).

The fully unrestricted version of the UC model with double seasonality is:

$$y_t = \mu_t + \gamma_{1,t} + \gamma_{2,t} + \varepsilon_t \tag{5a}$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \tag{5b}$$
$$\beta_t = \beta_{t-1} + \zeta_t \tag{5c}$$
$$\gamma_{1,t} = -\sum_{j=1}^{s_1-1} \gamma_{1,t-j} + \omega_{1,t} \tag{5d}$$
$$\gamma_{2,t} = -\sum_{j=1}^{s_2-1} \gamma_{2,t-j} + \omega_{2,t} \tag{5e}$$

where s_1 (e.g., $s_1 = 24$) and s_2 (e.g., $s_2 = 168$)are seasonal lengths ($s_1 < s_2$), $\omega_{1,t} \sim N(0, \sigma_{\omega_1}^2)$, and $\omega_{2,t} \sim N(0, \sigma_{\omega_2}^2)$. This version of the UC model is the most general specification of a structural time series model with double seasonality because two independent state equations of each seasonal pattern are independently included, and both state equations capture dynamic changes in each of two seasonal patterns. Like Taylor's double seasonality approach, this form of the UC model contains 192 seasonal factors. Unfortunately, although this structure can be specified in PROC UCM of SAS®, the program appears unable to estimate its parameters, possibly because of multicollinearity when a short seasonal pattern is repeated within a longer one.

A considerably more parsimonious specification, and one that can be estimated by PROC UCM is the model with blocked seasonality. It restricts each day to have a common daily pattern but allows the scale to change from day to day.

$$y_t = \mu_t + \gamma_{1,t} + \gamma_{2,t} + \varepsilon_t \tag{6a}$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t \tag{6b}$$
$$\beta_t = \beta_{t-1} + \zeta_t \tag{6c}$$
$$\gamma_{1,t} = -\sum_{j=1}^{s-1} \gamma_{1,t-j} + \omega_{1,t} \tag{6d}$$

$$\gamma_{2,t} = -\sum_{i=1}^{k}[\delta_i \cdot \sum_{j=0}^{s-1}\gamma_{1,t-j}] + \omega_{2,t} \qquad (6e)$$

where δ_i's are dummies for the blocked seasonal patterns, $\delta_i = 1$ for the i^{th} block and $\delta_i = 0$ otherwise, k is the number of the blocks, s is the identical block size per block ($s = 24$, here). The blocked seasonal UC model has 31 (= 24 + 7) seasonal factors.

Models are estimated once with the original hourly data and the forecasts aggregated to give a four-hour forecast, corresponding to one step ahead. Second, the data are aggregated into four-hour blocks and the models re-estimated. These correspond to the "forecast first then aggregate" and "aggregate first then forecast" approaches.

Results

We examined both additive and multiplicative seasonality for the electricity load data. Results are almost universally better for multiplicative seasonality and only those results are presented. For the emergency room arrivals data, the presence of zeroes in the original data series precludes the use of multiplicative seasonality.

Results for both data series are clear cut. Regardless of the model compared, the better strategy is to forecast first using the original hourly data, then aggregate the forecasts.

With the electricity load series, as shown in Table 2, the loss of performance from working with aggregate data is striking. Forecast errors, measured by root mean squared error practically double when only aggregate data are available.

There are differences caused by the use of different algorithms. In general the Multiple Seasonal model algorithm leads to larger errors compared with equivalent models estimated differently. Unobserved Components models are the best approach with disaggregate data, yet with aggegate data their performance is mixed. Double seasonality models, while never best, perform consistently well with both aggregated and disaggregated data. The strong daily and weekly cycles in the hourly data appear to be maintained in the aggregate data. A surprisingly poor performer is the Multiple Seasonality model with common weekday and separate common weekend cycles (r=2), Although the data display this pattern visually, this parsimonious approach for some reason fails to benefit.

One noticeable feature of the electricity load series is that the smoothing parameters for level tend to be high, often exceeding 0.9. Electricity consumption does display an annual pattern and this possibility was not allowed for, and this possibly accounts for the near non-stationarity of the series.

The first striking difference between electricity loads and emergency room arrivals is that while forecasting with the disaggregate data is still desirable, the accuracy using aggregate data is hardly much worse, and for two of the methods is actually slightly better. Table 3 shows the details. Double seasonality approaches perform the best, followed by long-cycle single seasonality, with short cycle seasonality the worst. The difference between most and least accurate is small, about 12 percent, compared with the much wider range for the electricity data.

Neither the Unobserved Components model nor the two-cycle version of Multiple Seasonality perform well. Finally, the within-sample performance is a good guide to choice of model for the electricity load series, but much less reliable for the emergency room arrivals.

Conclusions

The two hourly series examined here provide additional support to the principle: if you have disaggregate data, use it, even when the forecast is needed for a more aggregated time period. The principle was more strongly supported in the series that displayed stronger daily and weekly patterns.

Taylor's (2003) double seasonal modification to the standard Holt-Winters approach was remarkably robust. The two series did not provide conclusive evidence concerning the value of permitting multiple sources of error when specifying state-space models for the Holt-Winters structure. In one case Unobserved Components models were superior, but not in the other.

Rather surprisingly, the principle of parameter reduction was not supported. In particular, the Multiple Seasonal models developed by Gould et al (2005) gained nothing from imposing common cycle restrictions, even when visual inspection of the data suggested that the restriction was appropriate.

References

Abraham, B. (1982). Temporal aggregation and time series. *International Statistical Review, 50*, 285-291.

Butter, F. (1976). The use of monthly and quarterly data in an ARMA model. *Journal of Econometrics, 4*, 311-324.

Gould, P.G., Koehler, A.B., Vahid-Araghi, V., Snyder, R. D., Ord, J. K. & Hyndman, R. J. (2005). *Forecasting time-series with multiple seasonal patterns*. Monash University, Dept. of Econometrics and Business Statistics, Working Paper 28/05, revised October 2005. (Forthcoming as Gould, P.G., Koehler, A.B., Ord, J. K., Snyder, R. D., Hyndman, R. J., &Vahid-Araghi, V. (2008). Forecasting time series with multiple seasonal patterns. *European Journal of Operational Research ,191*, 207–222.)

Grunfeld, Y., & Griliches, Z. (1960). Is aggregation necessarily bad? *Review of Economics and Statistics, 42*, 1-13.

Hotta L. K. & Neto, J. C. (1993). The effect of aggregation on prediction in autoregressive integrated moving- average models. *Journal of Time Series Analysis, 14*, 261-269.

Hyndman, R. J., Koehler, A. B., Snyder, R. D., & Grose, S. (2002). A state space framework for automatic forecasting using exponential smoothing methods. *International Journal of Forecasting, 18*, 439-454.

Koehler, A. B., Snyder, R. D., & Ord, J. K. (2001). Forecasting models and prediction intervals for the multiplicative Holt-Winters method. *International Journal of Forecasting, 17*, 361-377.

Koreisha, S.G. & Fang, Y. (2004). Updating ARMA predictions for temporal aggregates. *Journal of Forecasting, 23*, 275-296.

Kumar, V., Leone, R. P., & Gaskins, J. N. (1995). Aggregate and disaggregate sector forecasting using consumer confidence measures. *International Journal of Forecasting*, *11*, 361-377.

Lütkepohl, H. (1986). Comparison of predictors for temporally and contemporaneously aggregated time series. *International Journal of Forecasting*, *2*, 461-475.

Lütkepohl, H. (2006). Forecasting with VARMA models, chapter 6 in G. Elliott, C. Granger and A. Timmerman (eds.) *Handbook of Economic Forecasting*, Elsevier North-Holland.
Ord, J. K., Koehler, A. B., & Snyder, R. D. (1997). Estimation and prediction for a class of dynamic nonlinear statistical models. *Journal of the American Statistical Association, 92*, 1621-1629.

Souza, L. R., & Smith, J. (2004). Effects of temporal aggregation on estimates and forecasts of fractionally integrated processes: a Monte Carlo study. *International Journal of Forecasting, 20*, 487-502.

Taylor, J. W. (2003). Short-term electricity demand forecasting using double seasonal exponential smoothing. *Journal of the Operational Research Society, 54*, 799-805.

Taylor, J. W., de Menezes, L. M., & McSharry, P. E. (2006). A comparison of univariate methods for forecasting electricity demand up to a day ahead. *International Journal of Forecasting, 22*, 1-16.

Wei, W. W. S. (1978). Some consequences of temporal aggregation in seasonal time series models, in A. Zellner (ed.), *Seasonal Analysis of Economic Time Series*, U.S. Department of Commerce, Bureau of the Census, pp. 433-444.

Figure 1: Electricity load data, first four weeks

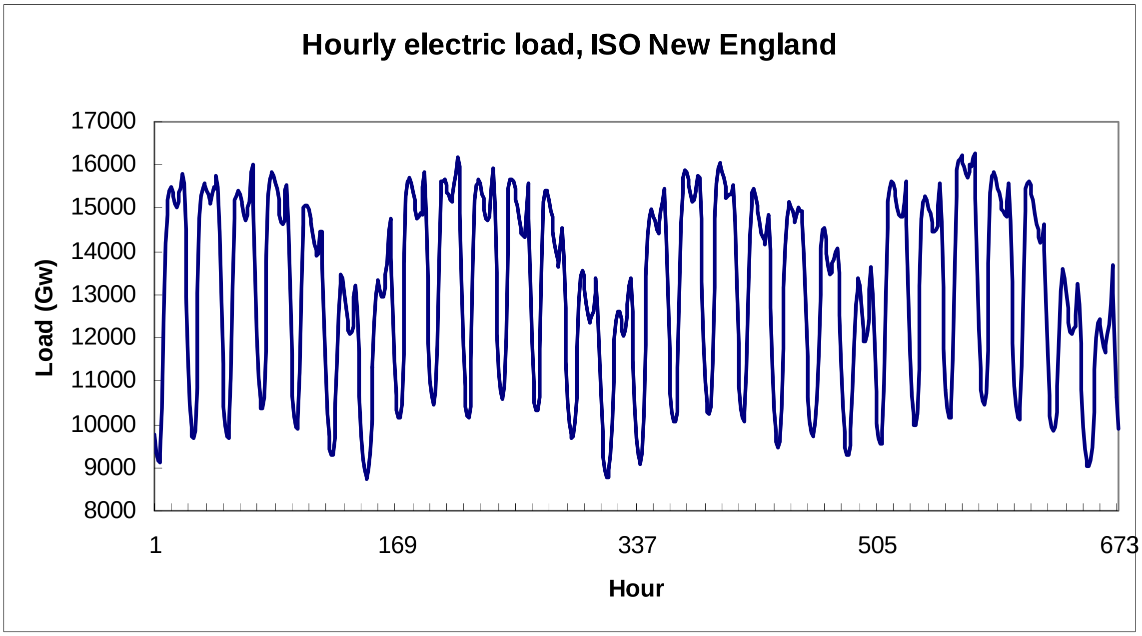

Figure 2: Emergency room arrivals data, first four weeks

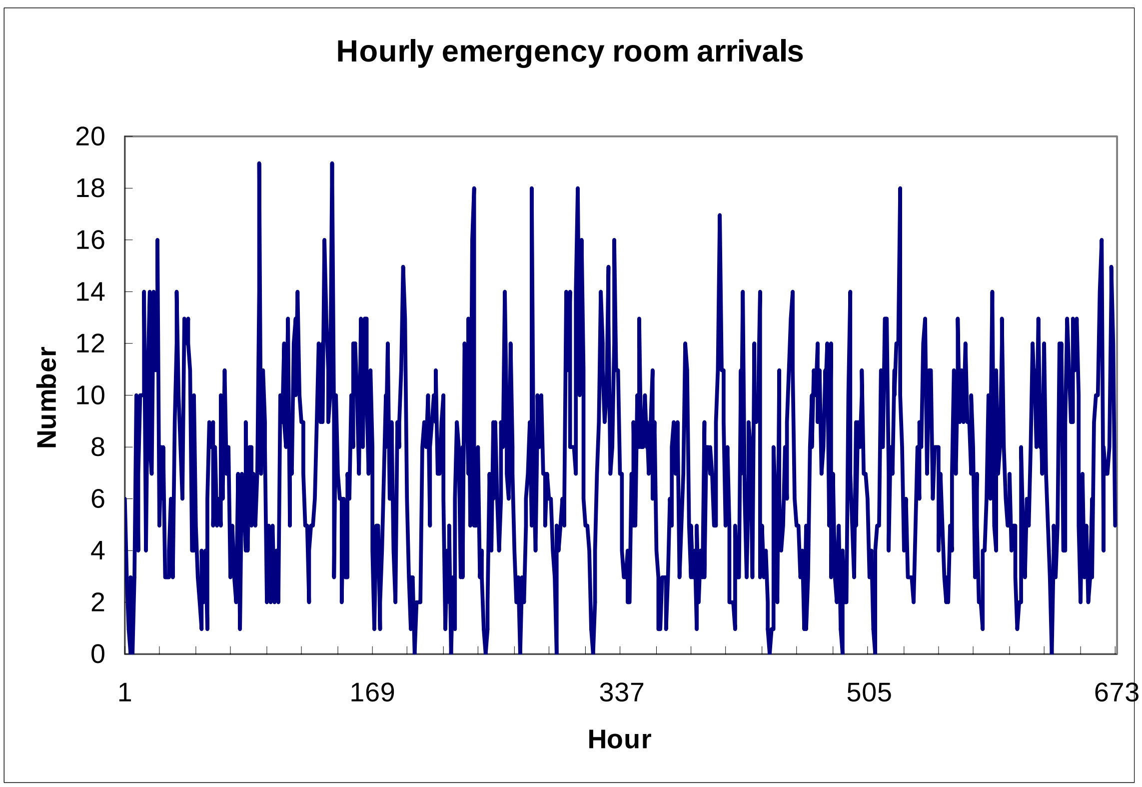

Table 1: Studies that compare forecast accuracy of "forecast then aggregate" and "aggregate then forecast" strategies.

Author/year	Variable	Estimation period	Method	Forecast period	Aggregation	Horizons	Result
Koreisha, Fang (2004)	Mortgage yield - govt loan yield (From Butter 1976)	1961.01-1974.12	ARMA	Within sample	Quarters	1	D
						2-4	A
	International airline passengers (Series G of Box & Jenkins)	1949.01-1960.12	$(0,1,1)(0,1,1)^s$	Within sample	Quarters	1,2,15,20	D
						5	A
Wei (1978)	US employed civilian workers	1949.01-1973.12	$(0,1,1)(0,1,1)^s$	1974	Quarters	1-4	D
Abraham (1982)	Canada: new dwelling units	1948.01-1968.12	ARIMA	1969.01-1976.12	Quarters	1	D
	Canada: rail freight loaded	1954.01-1968.12	ARIMA	1969.01-1976.12	Quarters	1	D
	US: civilian employed workers	1949.01-1964.12	ARIMA	1965.01-1972.12	Quarters	1	D
	Canada: merchandise exports	1946.01-1968.12	ARIMA	1969.01-1976.12	Quarters	1	A
	US: employment, tobacco industry	1947.01-1964.12	ARIMA	1965.01-1972.12	Quarters	1	A
Lütkepohl (1986)	ΔUS investment, seasonally adj	1947.1-1968.4	AR(p) (by AIC)	1969.1-1972.4	Annual	1	A
Kumar et al. (1995)	US durable goods, refrigerators, furniture, appliances, automobile sectors	1980.01-1984.12	BVAR	1985.01-1986.12	Quarters	1,2,4	D
Man (2004)	UK total consumption, non-durables cons, workforce	1955.1-1983.4	ARIMA	1984.1-1988.4	Annual	1-5	D
	UK exports, imports of goods and services	1955.1-1983.4	ARIMA	1984.1-1988.4	Annual	1-5	A
	UK real total investment, gross domestic product	1955.1-1983.4	ARIMA	1984.1-1988.4	Annual	1 2-5	D A

Notes: The frequency of the original disaggregated series is indicated by the number of digits after the year: one for quarterly, two for monthly. AR(I)MA means various autoregressive (integrated) moving average models estimated, the ones for aggregate data do not correspond to the model derived from the disaggregate estimated model. BVAR is Bayesian vector autoregression. The number of steps ahead refers to the aggregate units in the previous column. Result D means "forecast then aggregate" is more accurate, usually by mean squared error criterion, A means "aggregate then forecast" is more accurate.

Table 2: Electricity loads: multiplicative seasonality models, one-step-ahead forecast errors and rankings, based on root mean squared error

Method	Aggregate first			Forecast first		
	Within-sample	Post-sample	Post-sample Rank	Within-sample	Post-sample	Post-sample Rank
HW (short)	2172.5	2222.2	10	1002.9	964.3	8
MS(7)[R2a]	2197.2	2211.3	9	953.9	1026.0	9
UC (short)	2415.1	2186.3	8	408.0	360.8	1
HW (long)	1796.8	1848.1	5	1012.2	764.0	5
MS(7)[R1]	1932.8	1992.8	6	1046.7	1392.6	10
UC (long)	1966.9	1529.6	1	435.7	452.1	3
Double seasonal	1575.4	1562.7	3	685.9	743.5	4
MS(7)[R3]	1793.5	1550.0	2	784.9	805.6	7
MS(2)[R3]	1666.8	1590.3	4	737.7	791.5	6
UC(7xs)	2350.8	2127.2	7	410.3	361.6	2

HW = Holt-Winters method, MS(n)[R] = Multiple Seasonal model with n different seasonal cycles and R restriction on the seasonal smoothing parameters (see below), UC(m) = Unobserved Components model. Short = season length of one day (24 hours or 6 4-hour aggregates), long = season length of one week (168 hours or 42 4-hour aggregates), (7xs) = 7 seasonal blocks, controlled by 7 dummy variables for a single short season. For double seasonal see Taylor (2003)

The Multiple Seasonal model seasonal smoothing parameters form an rxr matrix. For r=2 the matrix with restriction R0 is $\begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix}$. Under R3 this becomes $\begin{bmatrix} \gamma_1^* & \gamma_2^* \\ \gamma_2^* & \gamma_1^* \end{bmatrix}$. Under R2 it is $\begin{bmatrix} \gamma_1^* & \gamma_1^* \\ \gamma_1^* & \gamma_1^* \end{bmatrix}$ and under R1 it is $\begin{bmatrix} \gamma_1^* & 0 \\ 0 & \gamma_1^* \end{bmatrix}$. According to Gould (2005) for r=7, R1 is equivalent to HW(long), R2 is equivalent to HW(short) and R3 is equivalent to double seasonality.

Table 3: Emergency room arrivals: additive seasonality models, one-step-ahead forecast errors and rankings, based on root mean squared error

Method	Aggregate first			Forecast first		
	Within-sample	Post-sample	Post-sample Rank	Within-sample	Post-sample	Post-sample Rank
HW (short)	6.221	5.316	8	5.885	5.059	7
MS(7)[R2a]	6.259	5.473	10	6.456	5.589	10
UC (short)	6.312	5.316	8	6.011	5.155	9
HW (long)	5.815	5.054	4	5.719	4.987	3
MS(7)[R1]	5.834	5.043	2	6.133	5.006	4
UC (long)	6.013	5.055	5	5.969	5.045	5
Double seasonal	5.838	5.044	3	5.812	4.986	2
MS(7)[R3]	5.810	5.037	1	5.995	4.979	1
MS(2)[R3]	5.806	5.078	6	5.963	5.050	6
UC(7xs)	5.950	5.151	7	5.917	5.154	8

See footnotes to table 1