A Study of Easter Proximity Regressors for REGARIMA Modelling

Frank Masci
Version 1.0, 25/10/2005

1. Introduction and Goals

This document expands on previous work of finding an appropriate Easter Proximity (EP) regressor to account for the pattern of activity leading up to and during Australia's Easter holiday period. Since Easter is a moving holiday, it can impact time series data aggregated on a regular calendar basis in an irregular but systematic manner. Movement of the Easter holiday period needs to be taken into account in the seasonal adjustment process of monthly or quarterly series (particularly those related to the retail trade) to avoid biasing trend estimates. The effect is most noticeable when the Easter holiday period falls close to or on the boundary of March and April, i.e., when the increase in retail activity leading up to Easter falls in March and the four-day holiday (and hence inactivity period) falls in April. In this case, March will be systematically increased, while April activity will be diminished. This is illustrated in the schematic of Figure 2.

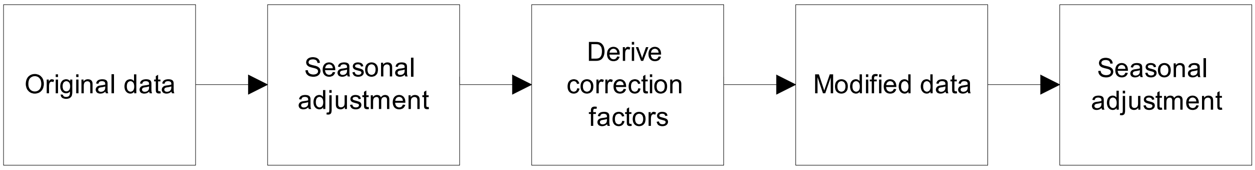
The intent here is to break some new ground on exploring a generic class of models that can correct for the EP effect in a REGARIMA framework. REGARIMA is an extension of ARIMA modelling which allows for regressors to be defined by the user to correct for systematic calendar effects. The TSA production team plan to utilise the predictive powers of ARIMA modelling to greater extent in the near future. This calls for a proper treatment of moving holiday effects specific to the Australian calendar. For an overview of REGARIMA, see paper by W. Bell, 1998 [ ]. In summary the main goals of this and future work are:

- Explore whether there is a generic class of EP models that can be applied to a wide range of series under a REGARIMA framework or;
- Explore robust criteria for selecting the best (appropriate) EP regressor for a given series (or group).
- Are the D13 irregulars currently used in production adequate for diagnosing the EP effect?
- Improve the statistical tests for detecting the EP effect, in particular exploring the sensitivity from outliers on the significance of an EP effect.
- How to implement these improvements in the SEASABS suite.

The work presented here explores the first two bullet points.

The current method for correcting the EP effect in the ABS seasonal adjustment package SEASABS involves using the final X11 irregulars from the D13 table. The EP effect causes the final irregulars to systematically deviate from one, the neutral line of irregulars (see EP plots in section 4). The irregulars are regressed on a model parameterized in terms of the fraction of "before" and "during" Easter holiday days that fall in March. A set of coefficients for these two periods are computed and tested for non-zero significance to ascertain whether an EP effect was present. The major drawback of using the X11-D13 irregulars directly is that they may have been distorted by the initial seasonal adjustment process that may or may not be related to Easter. Also, outliers are not corrected for prior to diagnosis of an EP effect. The D13 irregulars only provide a qualitative diagnostic to assess the impact of the EP effect, and any corrections derived therefrom may not be robust in the general sense, although they have sufficed for government work until now. For an overview and in-depth benchmark study of methods used by other statistical agencies, see the paper by C. McLaren & C. Leung, 1999 [📄].

D13 Method



The REGARIMA method estimates the EP effect from the original data by regressing the data against a model for the EP effect in conjunction with an ARIMA model for the series. Correction factors are then derived from the EP regression parameters and the original data is modified for input into the X11 or X12 iterative seasonal adjustment procedure (see Figure 1). This work is based on the X12-ARIMA package developed by the U.S. Census Bureau and the enhanced X11- SEASABS suite of software developed by the ABS.

Regression-ARIMA Method

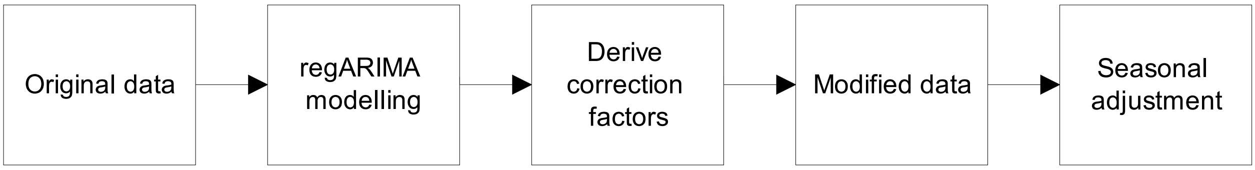


Figure 1. Schematic of regARIMA modelling and seasonal adjustment under X12 ARIMA

2. A Toy Model for the EP Effect

Extending the work of C. Mclaren & C. Leung, 1999 [📄], we adopt the generic model illustrated in Figure 2. There are two components to this model whose positions in the calendar move between the months of March and April according to the date of Easter. The first component represents an increase in the daily activity rate leading up to Easter and is parameterised as a power-law in time with index p-1, and the second component is assumed to represent a constant daily rate (=c_2) over the holiday period:

$$\frac{dA}{dt}=\begin{cases} c_1 t^{p-1} & Good\,Fri-w<t<Good\,Fri \\ c_2 & Good\,Fri\leq t\leq Easter\,Mon \end{cases} \tag{1}$$

where w is the window length in days over which the power-law is applicable. The window length of the holiday period (Good Friday to Easter Monday inclusive), h, is nominally set to 4. The reason why "p-1" is used for the power law index is that when the daily activity rate is integrated over the window length w, the power becomes "p" (see below).

If we define March as our reference (base) month, we can find the fraction of total activity that falls only in March by defining new variables n and m to represent the number of w, and number of h days that fall in March respectively (see Figure 2). On integrating Equation (1) with these variables as limits, we find the *total relative March activities in the* "before" *and* "during" Easter periods:

$$x_b(\mathrm{MAR})=\frac{\int_0^n (dA/dt)\,dt}{\int_0^w (dA/dt)\,dt}=\left(\frac{n}{w}\right)^{\mathrm{p}} \quad \text{(before Easter)} \tag{2}$$

$$x_d(\mathrm{MAR})=\frac{\int_0^m (dA/dt)\,dt}{\int_0^h (dA/dt)\,dt}=\left(\frac{m}{h}\right) \quad \text{(during Easter)}$$

From this construction, the *total relative April activities* can be wrtten:

$$x_b(\mathrm{APR})=1-\left(\frac{n}{w}\right)^{\mathrm{p}} \quad \text{(before Easter)} \tag{3}$$

$$x_d(\mathrm{APR})=1-\left(\frac{m}{h}\right) \quad \text{(during Easter)}$$

These represent our generic Easter-regressor variables, where p, w are treated as free parameters and *m, n* depend on the date of Easter. More specifically, we refer to this model as the "p-linear" model since it consists of a power-law dependence with index p (e.g., p = 2 implies a quadratic model) in the before Easter period, and a linear dependence during the Easter holiday period. In this work, h is fixed at 4.

A general EP regressor matrix for input into REGARIMA can be set-up using the explanatory variables parameterised by Equation (2). If E_b is the "before" Easter activity parameter and E_d is the "during" Easter Holiday parameter, the regression mean function used to remove the source of Easter proximity non-stationarity from the input series y_t is defined as:

$$\mu_t = E_b x_b + E_d x_d .$$

where the residuals $z_t = y_t - \mu_t$ can now be modelled using a generalised ARIMA model:

$$\phi(B)\Phi(B^s)(1-B)^d(1-B^s)^D\left[y_t - \mu_t\right] = \theta(B)\Theta(B^s)\varepsilon_t , \qquad \varepsilon_t \sim N(0,\sigma^2)$$

x_b and x_d are the corresponding "explanatory" variables defined by a monthly regression matrix:

$$\begin{aligned} x_b &= \left(\frac{n}{w}\right)^p ; \quad x_d = \left(\frac{m}{h}\right) \text{ for MARCH} \\ x_b &= -\left(\frac{n}{w}\right)^p ; \quad x_d = -\left(\frac{m}{h}\right) \text{ for APRIL} \\ x_b &= 0; \quad x_d = 0 \text{ otherwise} \end{aligned}$$

Thus, March is the reference month defining the regressor, and the symmetry $x_{b,d}$ (APR) = -$x_{b,d}$ (MAR) ensures that the correction for the EP effect has no net effect on the final adjusted series over a year. Since Easter usually falls in April, we balance out the EP effect by throwing its "inverse" (or negation) into April so that the process of seasonal adjustment can deal with it. For a log-transformed "stationary" series (described by an additive model), "no net effect" is equivalent to writing:

$$\sum_{month} x_b = \sum_{month} x_d = 0,$$

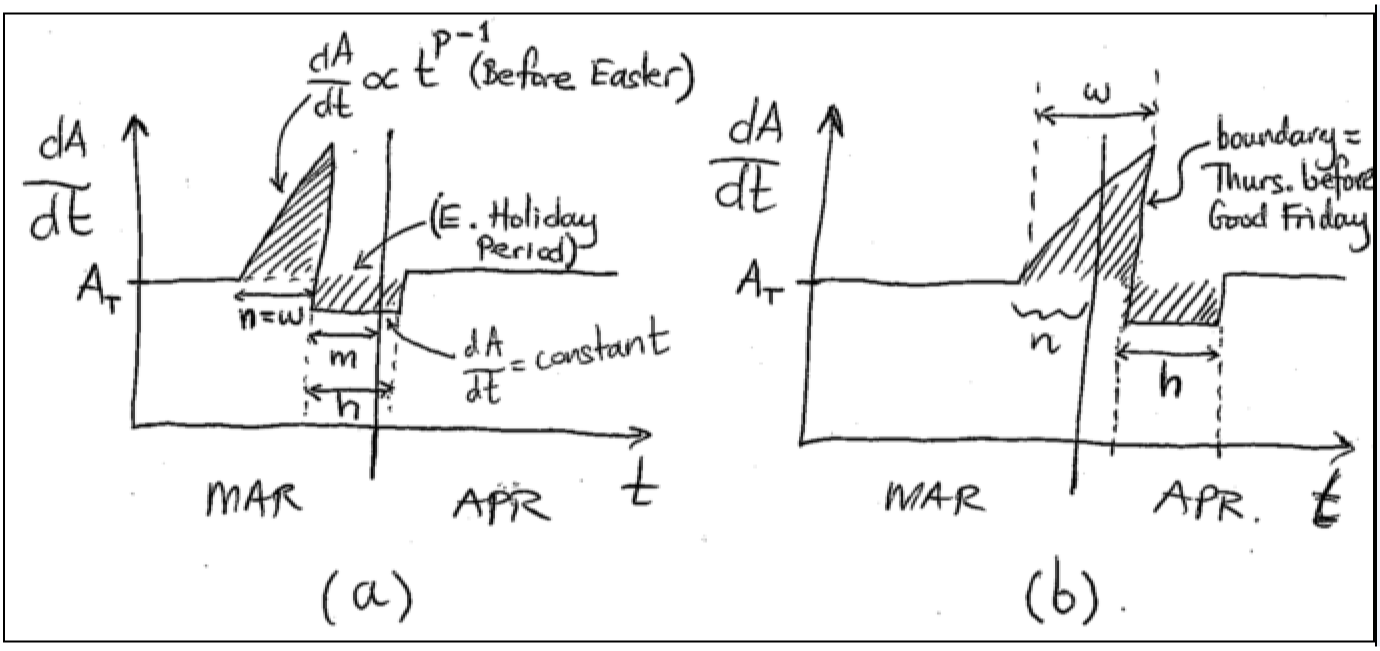
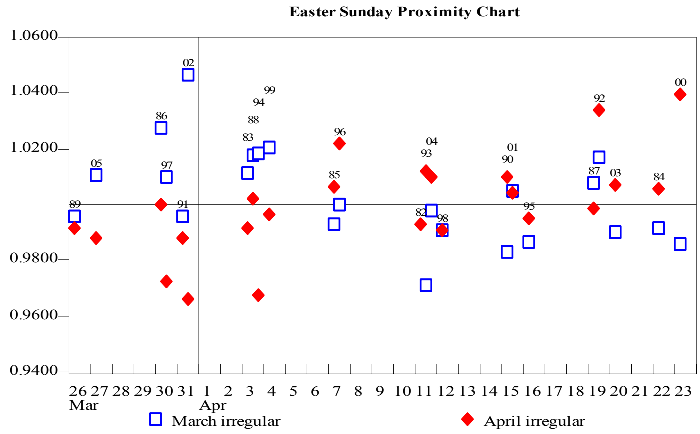


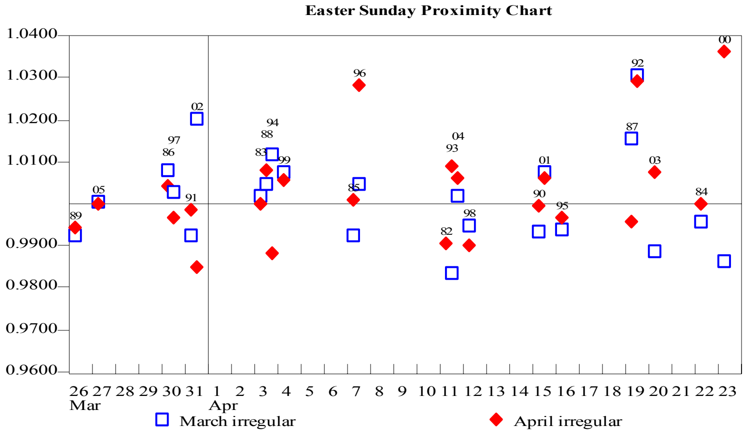
Figure 2. Schematic of a simple model for the EP effect. Vertical axes represent the daily activity rate and horizonal axes are time with the March and April months shown. "w" is the window length in days of the pre-Easter activity period and "h" is the window length of the Easter holiday period (nominally set to 4 to represent Good Friday to Easter Monday). "n" is the number of "w" days, and "m" is the number of "h" days that fall in March. (a) represents a scenario where the complete pre-Easter activity period, and a fraction of the during-holiday period fall in March and (b) is where a fraction of the pre-Easter period but the complete holiday period fall in April. The "level" A_T is the mean-constant activity rate for those days that don't fall within the "w" or "h" windows.

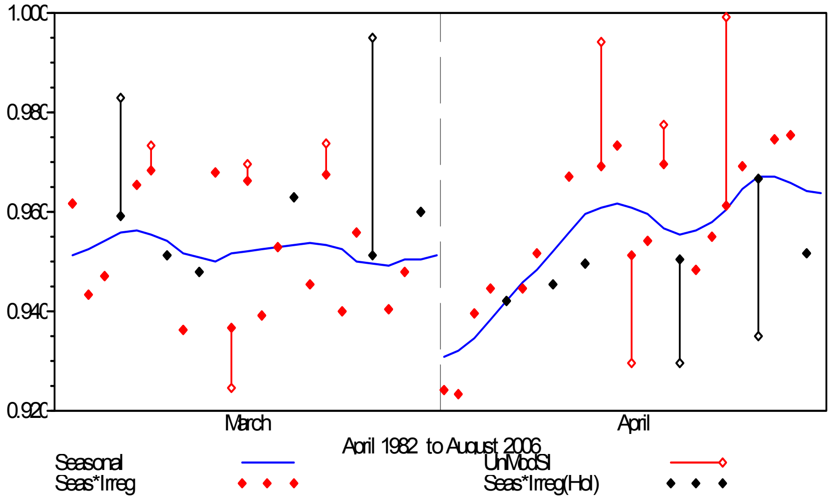
The following are examples of EP proximity and corresponding S*I charts before and after correction using a quadratic (p=2)--linear model:

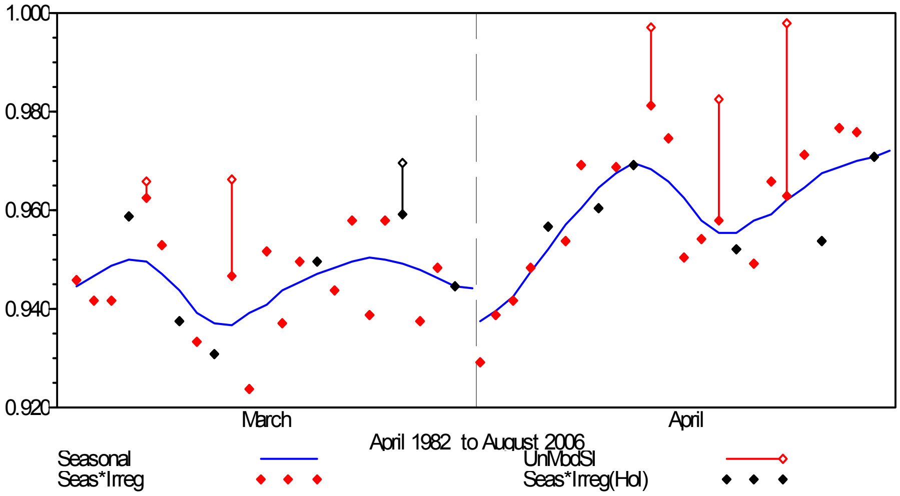
Without EP correction

With EP correction









From these plots, it's important to note that:

- In the no EP correction S*I plot (left), SEASABS has treated (and corrected) the large March irregular values as outliers for estimating the final seasonal factors. When the EP correction is turned on, the outlying "EP irregulars" are corrected using the EP algorithm.

- The EP model is only used to correct for systematics in the irregulars for those years when Easter falls in March or Early April (black dots).
- Since Easter usually falls in April (red dots), it can be treated as a seasonal effect and thus its effects can be captured in the seasonal adjustment process.
- From the S*I sharts, the March S*I values are brought down, and the April S*I values are pushed up on average after the EP correction is applied by virtue of modifying the irregulars.
- For a bi-monthly series, i.e. if data were only available every two months (thus sampled at a rate 6 times/year), would you expect to have an EP effect?

3. An Empirical Estimation of *p* and *w*

I have attempted to find the best values of the model parameters p and w that can be used to correct for the EP effect in a broad range of Australian retail series. An empirical approach was used where seven retail series were averaged together and a model constructed from Equations (2) and (3) above was fitted to the data. See below.

The model is based on using the original series data and not the D13 irregulars which may have been corrupted by the process of seasonal adjustment. We have modelled the ratio of total March to April *original* activity, R. This ratio is expected to change in a predictable manner as the date of Easter changes according to the above model. In terms of the original values and its decompositions under a multiplicative model: T (average trend), S (seasonal component), I (irregular-systematic component) and ε (irregular-random noise component), the ratio R can be written:

$$\begin{aligned} R &= \frac{O_M}{O_A} \\ &= \frac{T_M S_M I_M \varepsilon_M}{T_A S_A I_A \varepsilon_A} \\ &= \frac{T_M\left(1+\left[I'_M / T_M\right]\right)\varepsilon_M}{T_A\left(1+\left[I'_A / T_A\right]\right)\varepsilon_A} \\ &= \frac{T_M\left(1+E_M\right)\varepsilon_M}{T_A\left(1+E_A\right)\varepsilon_A}, \qquad (4) \end{aligned}$$

where we assumed the seasonal components for March (subscript *M*) and April (subscript *A*) are negligible, so that $S_M \sim S_A \sim 1$, and that any residual seasonality or noise due to Easter or otherwise can be absorbed into effective irregular terms: $I'_M$and I'_A . Therefore, the I' terms (the measurable part of the irregularity) include the effects of Easter, and the ε terms contain the remaining irregularity, uncorrelated with the I' components, e.g, $\varepsilon \sim N(1,\sigma^2)$. The initial multiplicative decomposition has therefore been transformed into an additive one to describe the EP effect, where the E_M and E_A variables now represent the contribution of Easter relative to the overall trend levels for March and April respectively. These are expected to lie in the interval

$$\begin{aligned} -\delta \le E_M \le \delta \\ -\delta \le E_A \le \delta \end{aligned},$$

where δ is the maximum possible contribution from Easter and can be determined from observations (see below). Over a long span of years, the EP effect for either March or April is expected to average out to zero if the date of Easter is approximately random over this span. This implies that the above intervals will be nearly symmetric.

We can simplify Equation (4) by taking the logarithm of both sides:

$$Log\,R = Log\left(\frac{T_M}{T_A}\right) + Log\left(\frac{\varepsilon_M}{\varepsilon_A}\right) + E_M - E_A, \qquad (5)$$

where we used the approximation *Log(1 + E) ~ E* since as seen in the March/April ratio for a number of retail series in Figure 4 (and their averages in Figure 5), the magnitude of either E_M or E_A is expected to be small for those years when Easter falls in March or early-April (red vertical lines in Figure 5). In other words, the effects of Easter are expected to be no larger than δ ~6% relative to the average trend ratio expected for March and April:

$$\frac{T_M}{T_A} = \frac{31\,\text{days}}{30\,\text{days}} \approx 1.03333... \qquad (6)$$

The relative March and April activities (E_M and E_A) can be parameterized in terms of our "toy" model Equations (2) and (3) as follows:

$$\begin{aligned} E_M &= \phi_b\left(\frac{n}{m}\right)^p - \phi_d\left(\frac{m}{h}\right) \\ E_A &= \phi_b\left[1-\left(\frac{n}{m}\right)^p\right] - \phi_d\left(1-\frac{m}{h}\right) \\ \Rightarrow E_M - E_A &= \phi_b\left[2\left(\frac{n}{w}\right)^p - 1\right] - \phi_d\left[2\left(\frac{m}{h}\right)-1\right] \end{aligned}, \qquad (7)$$

where the coefficients $\{\phi_b, \phi_d\} \geq 0$, represent "before" and "during" Easter holiday coefficients and will be determined from a regression fit (see below). Combining Equations (5), (6), (7), the logarithm of the total March to April activity can be written:

$$Log\,R = C + Z_\varepsilon + \phi_b\left[2\left(\frac{n}{w}\right)^p - 1\right] - \phi_d\left[2\left(\frac{m}{h}\right) - 1\right] \quad , \qquad (8)$$

where

$$C = Log\left(\frac{T_M}{T_A}\right) = Log\left(\frac{31\,\text{days}}{30\,\text{days}}\right) \approx 0.033,$$

$$Z_\varepsilon = Log\left(\frac{\varepsilon_M}{\varepsilon_A}\right). \qquad (9)$$

A schematic of the March/April ratio (R not *Log R*) as parameterized by Equation (8) using the fiducial and nominal "quadratic-linear" model values: $p = 2;\ w = 7;\ h = 4;\ \phi_b = 0.2;\ \phi_d = 0.05;\ Z_\varepsilon = 0$ is shown in Figure 3.

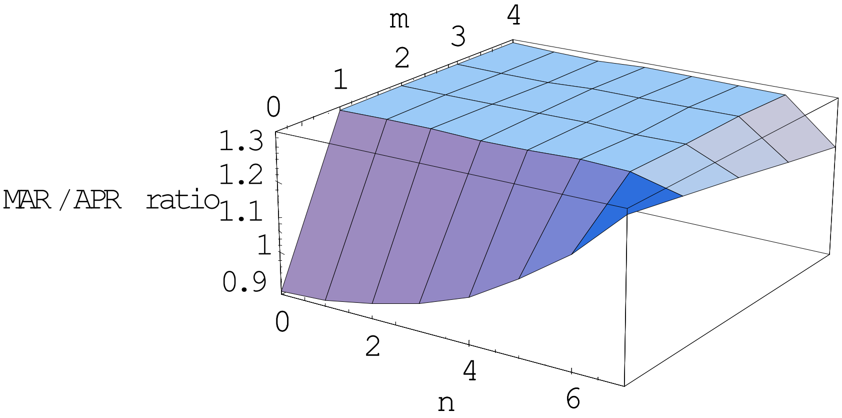


Figure 3. Ratio of total March-to-April activities as a function of the *m* and *n* variables as parameterised by Equation (8). These variables were defined in Figure 2.

We have estimated the model parameters $\{\phi_b, \phi_d, p, w\}$ by minimising the Mean Squared Error (MSE) of the model described by Equation (8), i.e.,

$$\left\langle Z_\varepsilon^{\ 2} \right\rangle = \left\langle \left(Log\,R - C - \phi_b\left[2\left(\frac{n}{w}\right)^p - 1\right] - \phi_d\left[2\left(\frac{m}{h}\right) - 1\right] \right)^2 \right\rangle \qquad (10)$$

using *averaged Log R* series data compiled from seven retail series (see Figure 5). This amounts to a non-linear regression and our primary goal here is just to obtain the best fit estimates of $\{\phi_b, \phi_d, p, w\}$. An estimation of the parameter confidence region will require proper treatment of the error distribution function for Z_ε through Monte Carlo simulation since it is not necessarily Gaussian.

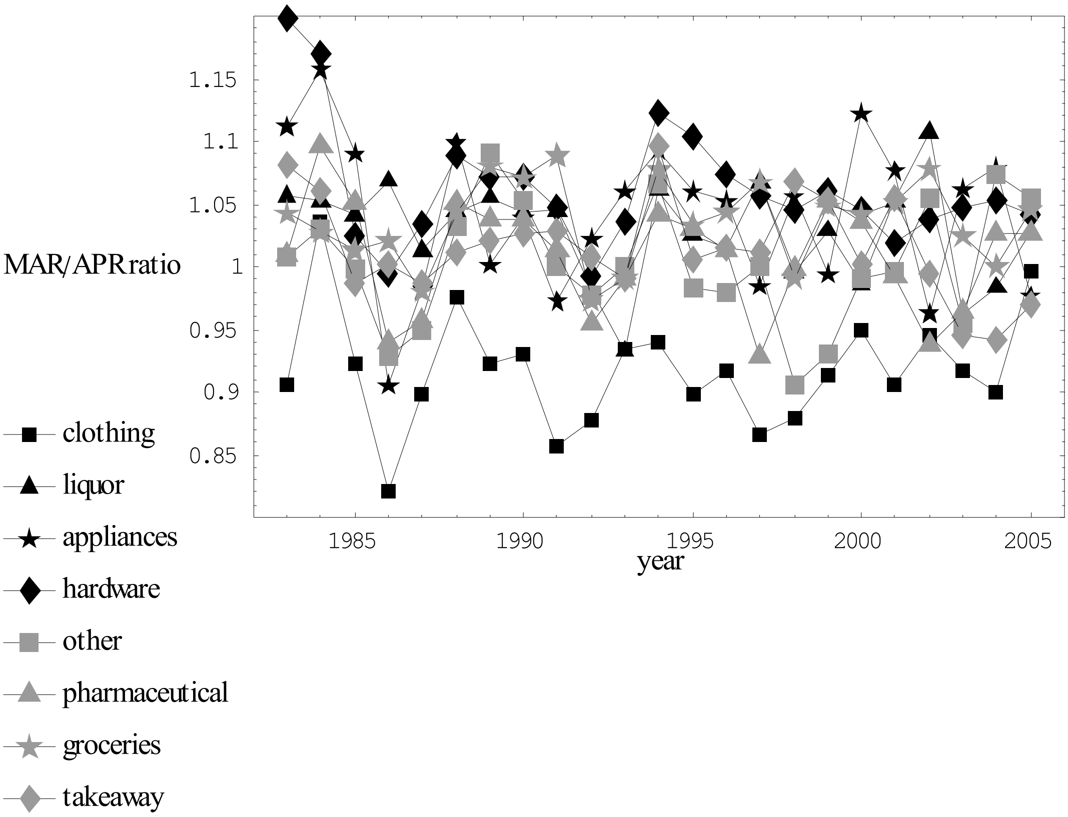


Figure 4. The ratio of March-to-April raw values from eight retail series whose types are shown in the legend.

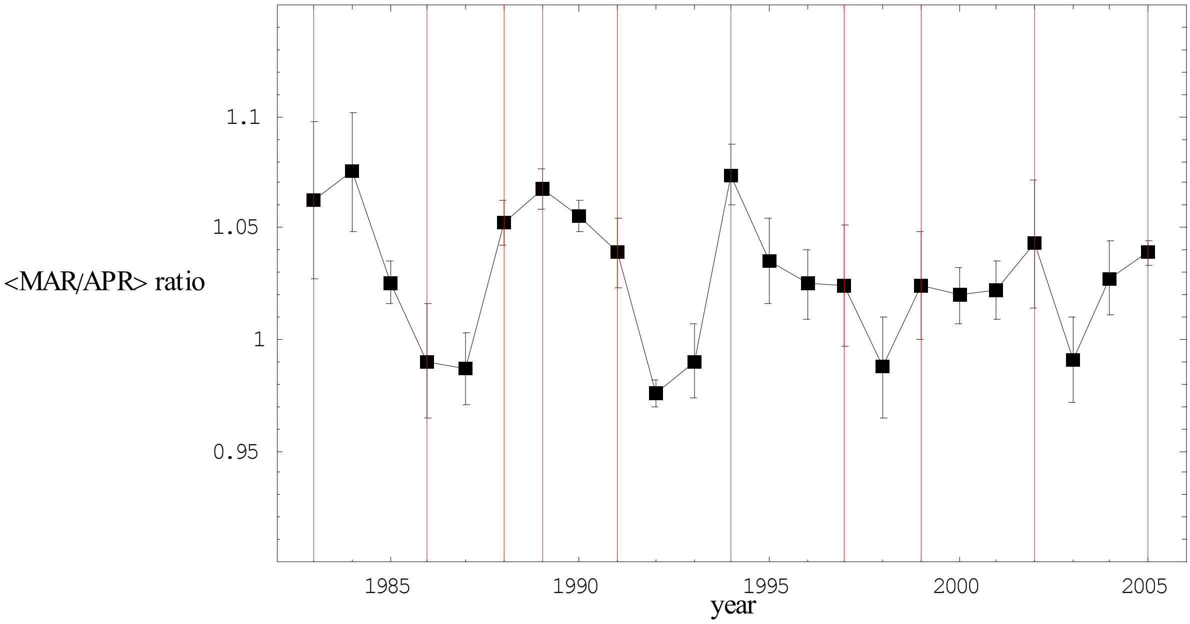


Figure 5. The average ratio of March-to-April raw values from seven retail series shown in Figure 4 (*the outlying "clothing" series was omitted*). The red vertical lines represent those years where Easter fell in March or early April (i.e. on or before Apr. 4). Note that the March values are ~3% larger (i.e. ~31/30 as expected). The error bars are sample-variances/sqrt(N) where N is the number of samples averaged for each year.

A summary of the non-linear least squares regression on the observed March/April ratios is as follows:

- The best fit values for p and w are p=4.01 and w=3, which implies a strong dependence in the retail activity as a function of time leading up to Easter (compared to the quadratic p=2 model), and a relatively short pre-Easter window compared to the nominal w=7 currently used in SEASABS and previous studies. These estimates assume the nominal value h=4.
- A comparison of the March/April activity ratio predicted by the best model fit to the average values observed is shown in Figure 6 (red curve). Also shown (blue curve) is the best nominal "quadratic-linear" model fit with p=2 and w=7. The MSE for this latter model is minimised if the "during" Easter holiday coefficient, ϕ_d, is very close to zero. In other words, it requires no reduction in retail activity during the Easter holiday period.
- A fit with p and w treated as free parameters yields a MSE ~ 0.014, and that of a fixed quadratic-linear model yields a MSE ~ 0.019, i.e., the fit is not as good.

However, this finding is not consistent with the results of REGARIMA modelling shown in Table (i) of section 5. In other words, the free-parameter model with p=4.01 and w=3 does a significantly poorer job at correcting for the EP effect in time series than the quadratic-linear model based on an analysis of final D13 irregulars. This leaves one to believe that the use of observed March/April ratios to constrain an EP model is not robust. Both the blue and red curves in Figure 6 provide qualitatively the expected behaviour for the EP effect, but the dependence of the March/April ratio on the date of Easter provides a poor discriminator between various "toy" models. This may be because:

- Random errors in the March/April ratios are actually larger than those represented by the dispersion amongst several series. Note that the random errors (error bars in Figure 6) could well be under-estimated. These ratios are just inherently noisy so that any "random" toy model would just fit equally well. In other words, the "global minimum" (if any) in the parameter space of the MSE cost function is very broad.
- There is not a large enough dynamic range in these ratios to unequivically constrain a good model.
- Not enough time series were averaged to get sufficient signal-to-noise (SNR) in these ratios so to exacerbate the pattern from the EP effect.
- The toy-model is conceptually wrong or too simple! e.g., we may need to allow for an evolving pre-Easter period window length, or an evolving activity rate index (p). This however will introduce too many degree's of freedom and make the model exceedingly difficult to constrain.

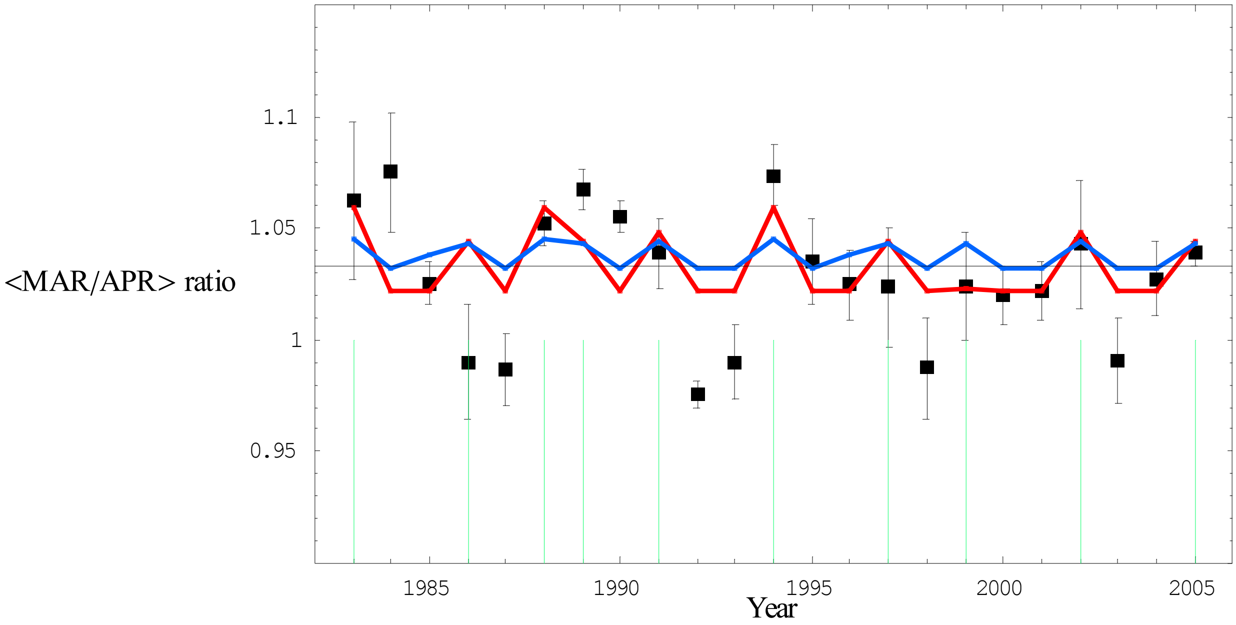


Figure 6. Average observed ratio of March-to-April values (boxes with error bars), best fit of Equation (8) with parameters: $p=4.01;\ w=3;\ \phi_b=0.018, \phi_d=0.007$ **(red curve) and the "best" quadratic-linear model fit:** $p=2.0;\ w=7;\ \phi_b=6\times10^{-5}, \phi_d=0.0$ **(blue curve). Both cases assume h=4. Note that the best quadratic-linear model requires a "during" Easter holiday correction of zero, in other words, with no reduction in activity during the Easter holiday period. Green vertical lines indicate the March/early-April Easter years.**

4. REGARIMA Case Studies

We explored seven regressor models parameterised by six different combinations of the parameters [w, h, p]. These are listed in the first column of Table (i) in Section 5. These were drawn from a testing pool of 13 models, chosen to give very similar results for different combinations of [w, h, p]. Except for the "My model fit" [w, h, p = 3, 4, 4] derived from the March/April ratio fit above, the best models (those giving the least variance in final irregulars) appeared to require p ~ 2.

Two case series were used for the model comparisons - a "Liquor retail series" and a "SuperMarket/grocery-store retail series", which are state aggregate series. These series were specifically chosen for the presence of a strong EP effect (see EP charts below). The March/April ratio plots for these series is as follows. Note the "excesses" (peaks) at the green vertical lines which indicate the March/early-April Easter years.

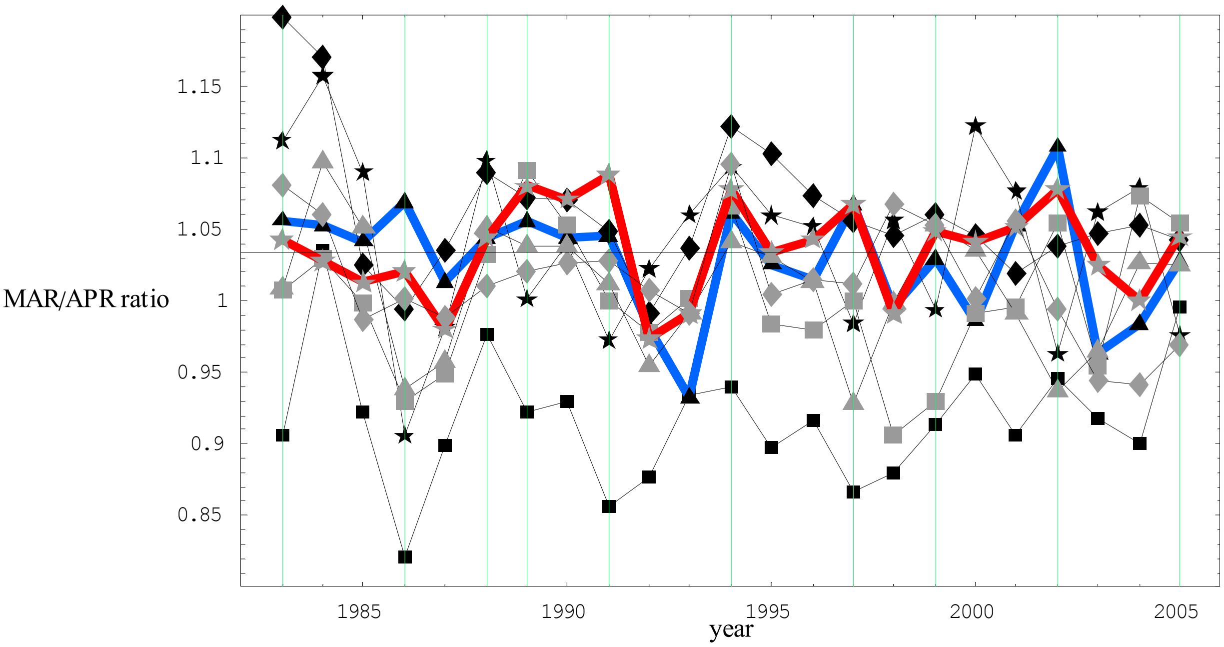


Figure 7. March/April ratios of original retail series with *Liquor retail series* (in blue) and *SuperMarket/grocery retail series* (in red). The symbols defining other retail series were labelled in Figure 4. The green vertical lines indicate the March/early-April Easter years.

i. Liquor-retail series

The EP (D13-irregular) chart for this series, before correcting for any EP effect is as follows.

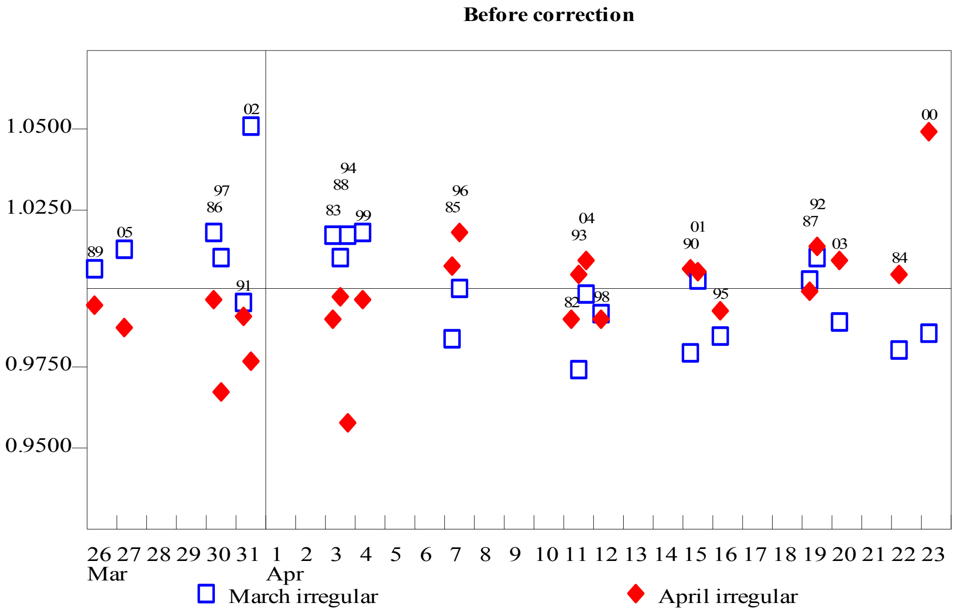


Figure 8. Easter Proximity chart for the state-aggregate "Liquor retail series" (before correction).

Listed below are the EP (D13-irregular) charts after correcting for the EP effect using each of the seven [w, h, p] models listed in the first column of Table (i) in the REGARIMA modelling. The result from the SEASABS model correction is shown first. Beware that the vertical scale on each of these charts (autoscaling from SEASABS) is different.

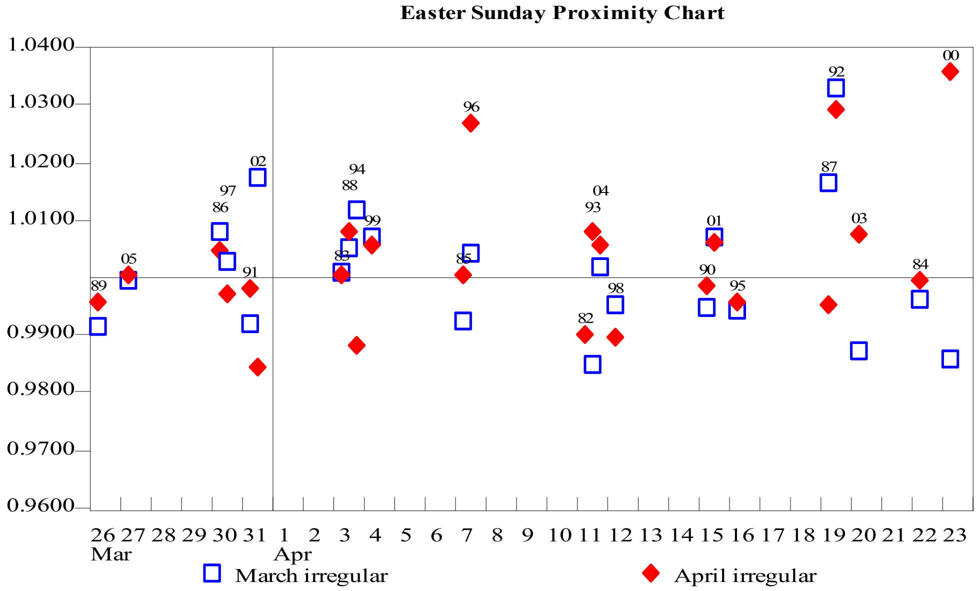


Figure 9. Easter proximity chart after correction using the SEASABS algorithm.

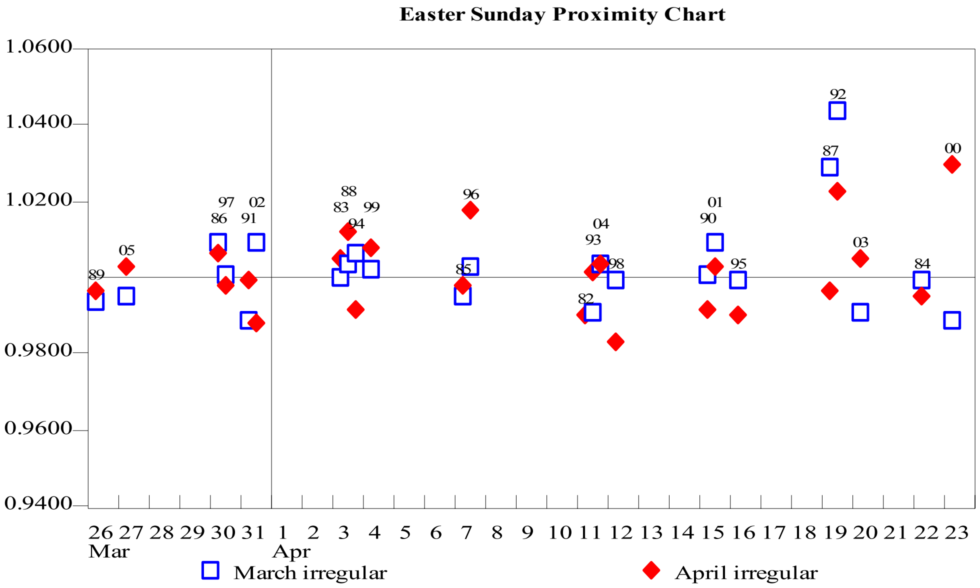


Figure 10. Easter proximity chart after correction with REGARIMA[w, h, p] = [7, 4, 2] (quad-linear regressor).

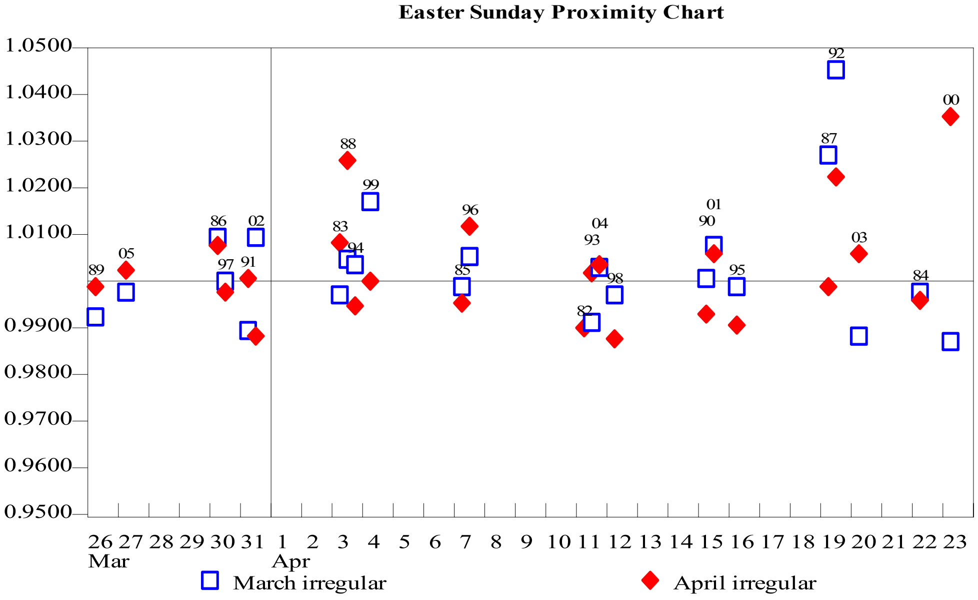


Figure 11. Easter proximity chart after correction with REGARIMA[w, h, p] = [3, 4, 4] (My model fit).

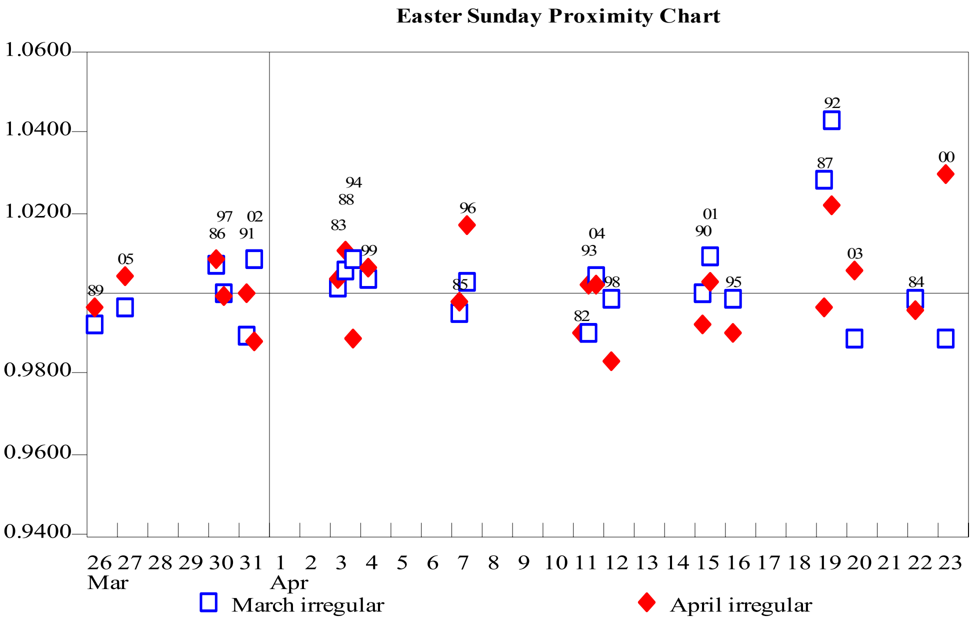


Figure 12. Easter proximity chart after correction with REGARIMA[w, h, p] = [7, 0, 2] (quad. model with <u>No</u> "during" Easter holiday regressor).

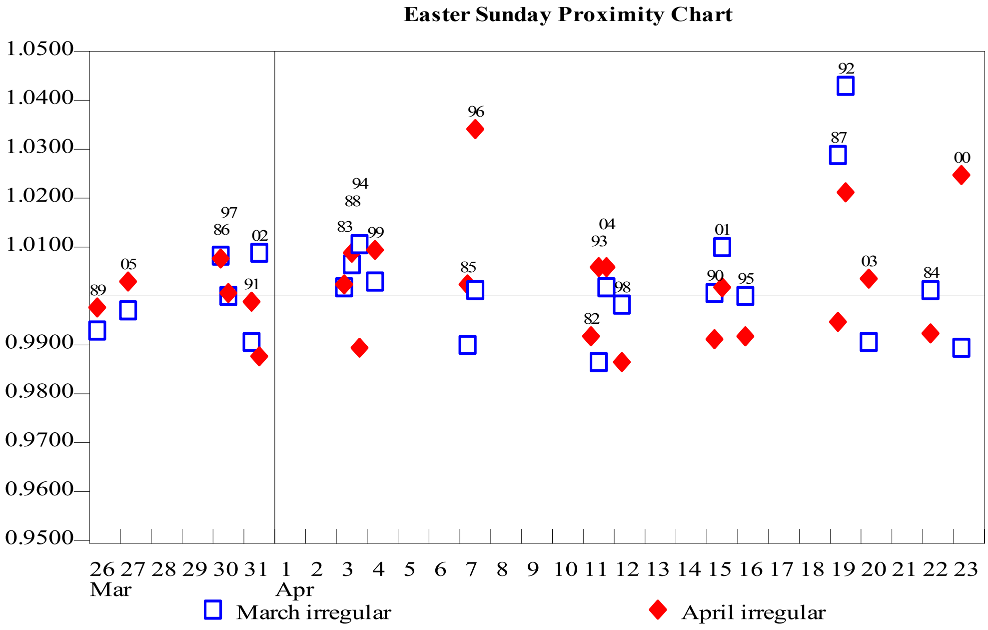


Figure 13. Easter proximity chart after correction with REGARIMA[w, h, p] = [14, 4, 2] (quad-linear regressor with big pre-Easter window).

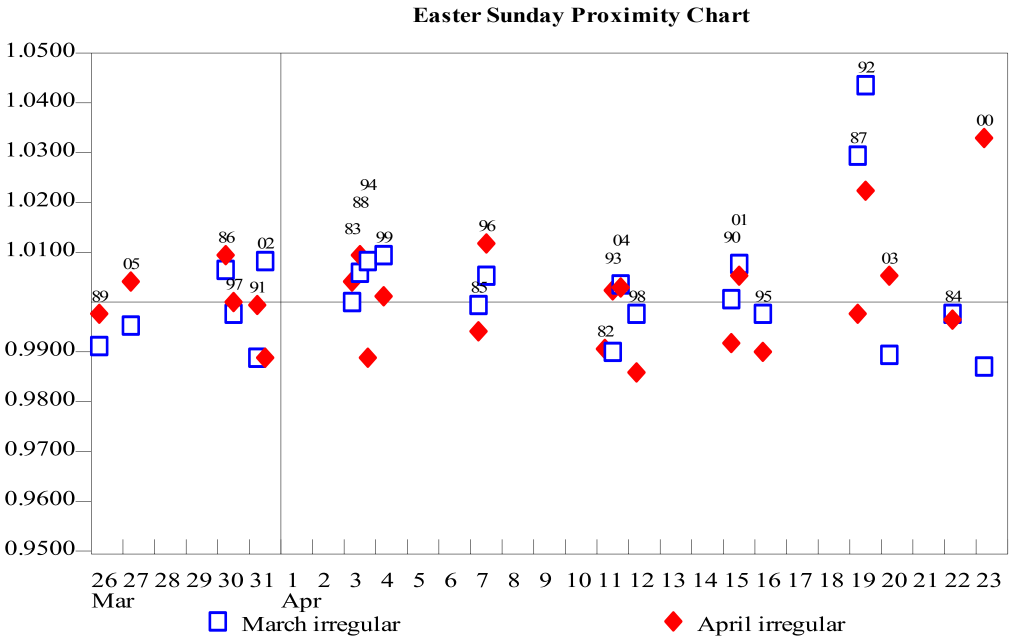


Figure 14. Easter proximity chart after correction with REGARIMA[w, h, p] = [3, 0, 2] (quad. model with No "during" Easter holiday regressor and 3-day pre-Easter window).

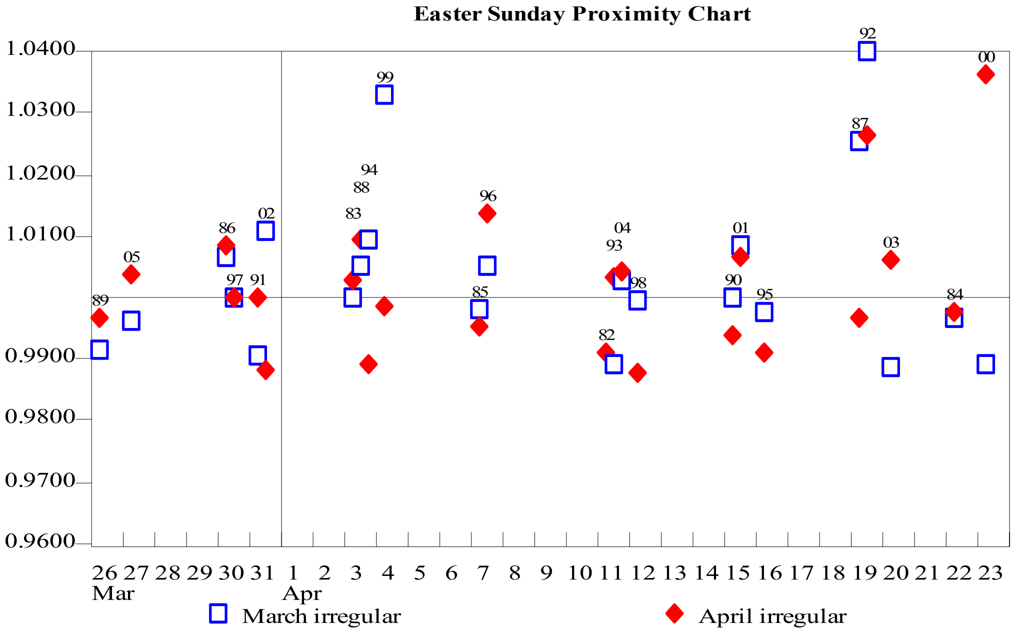
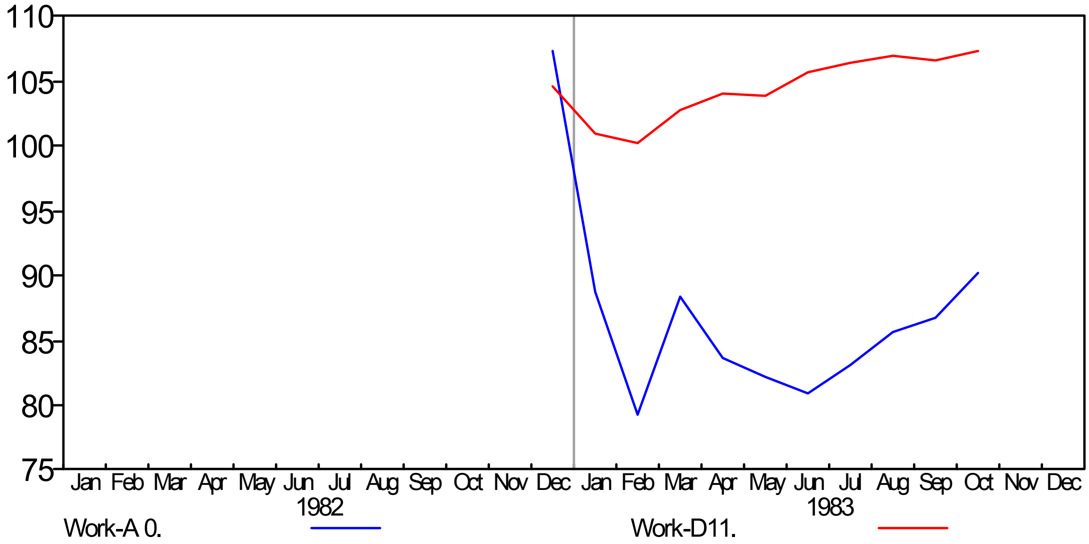
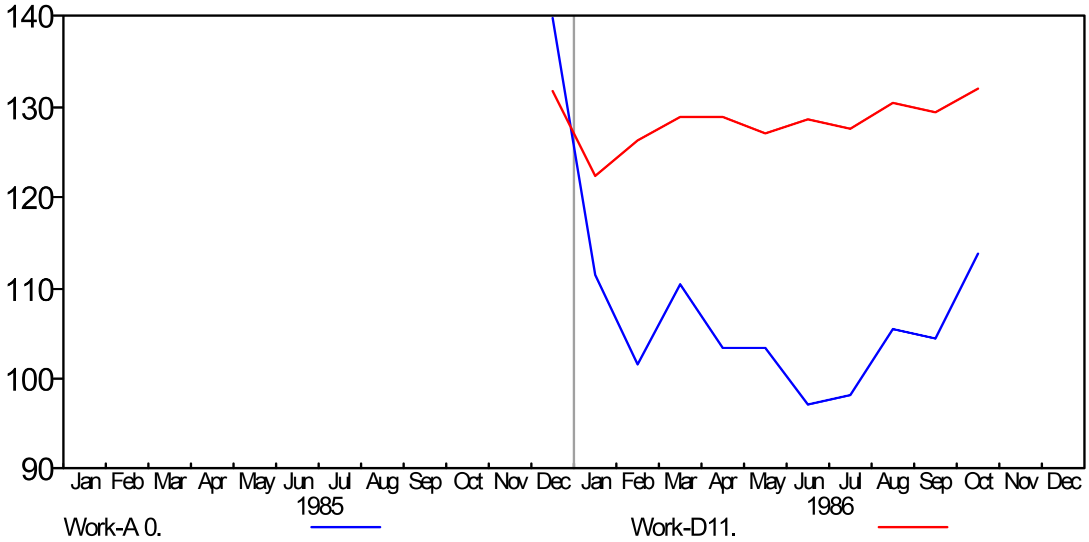


Figure 15. Easter proximity chart after correction with REGARIMA[w, h, p] = [1, 0, 2] (quad. model with No "during" Easter holiday regressor and 1 day pre-Easter window).

The following are plots of original series (A0 series in X11: blue) and seasonally adjusted series (D11 series in X11: red) that have been prior-corrected for the EP effect using the REGARIMA[w, h, p] = [7, 0, 2] model. Only snapshots of some March/early-April Easter years are shown. The noteworthy feature here is the degree of correction for the systematic EP effect seen in the original March/April irregulars (blue curves).





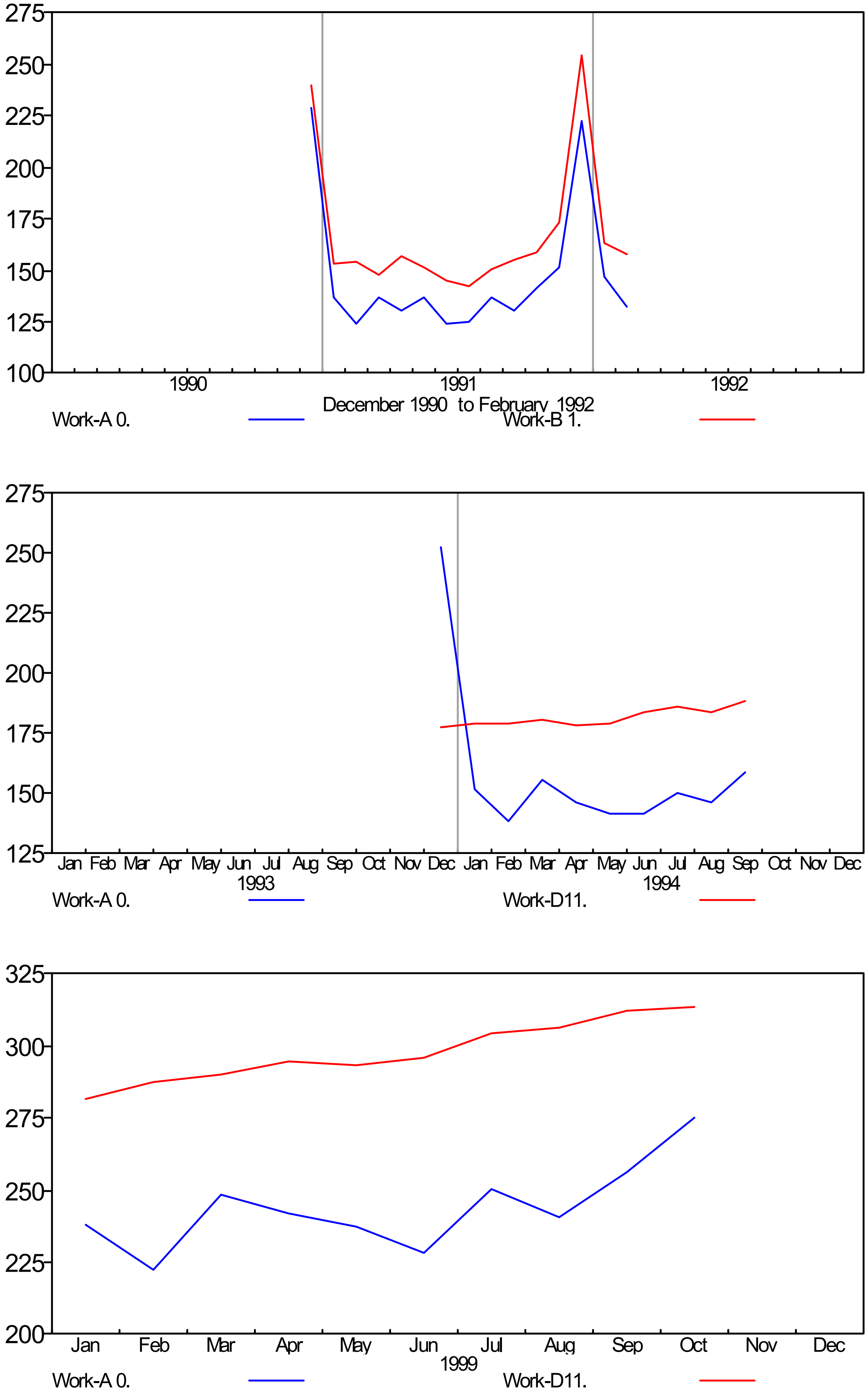
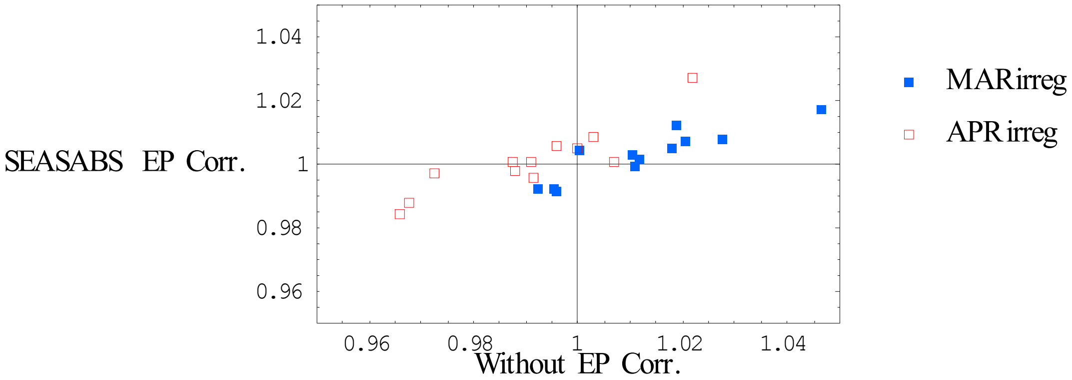
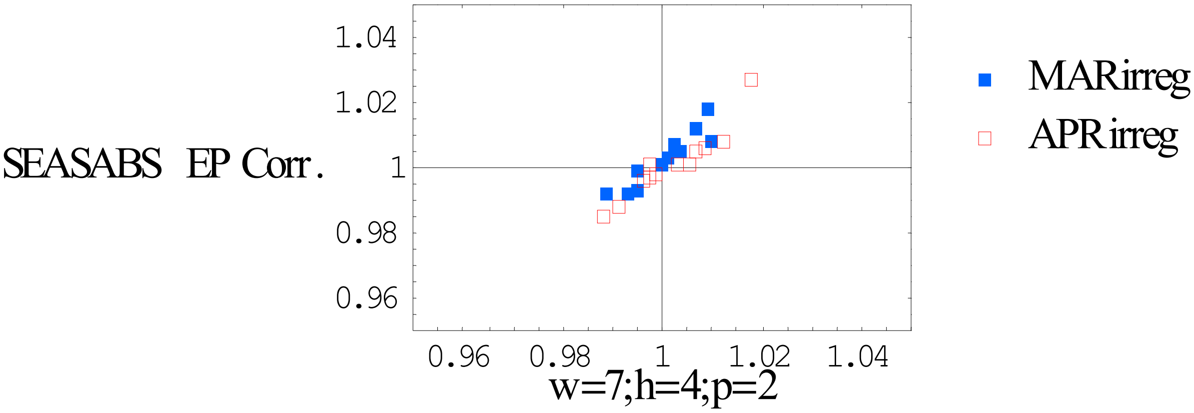


Figure 16. Series snapshots of some March/early-April Easter years

before (blue) and after (red) EP correcting the original series with prior factors from REGARIMA model [w, h, p] = [7, 0, 2].

The following charts are "model-model" plots where the X11 D13-irregulars from one model are plotted against the irregulars of another. Only the irregulars for March/early-April Easters are shown. This allows for easier comparison between competing models. The X and Y axes have the same scale to facilitate a better comparison than the EP irregular charts above. Here, the SEASABS EP correction model is compared to three REGARIMA models.





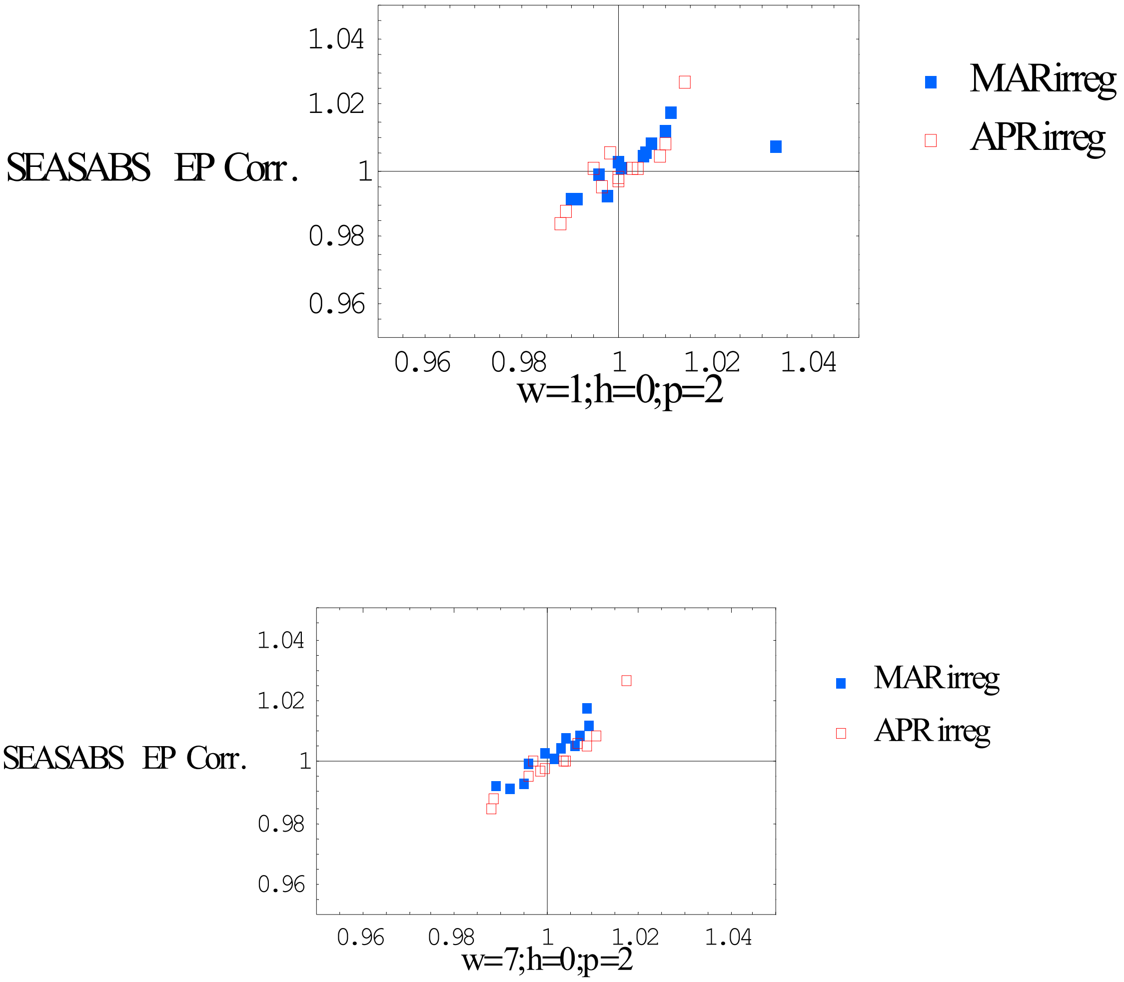


Figure 17. SEASABS-model versus REGARIMA-model comparison plots of March/early-April Easter year irregulars after correction was applied.

ii. Super-market/grocery retail series

The EP (D13-irregular) chart for this series, before correcting for any EP effect is as follows.

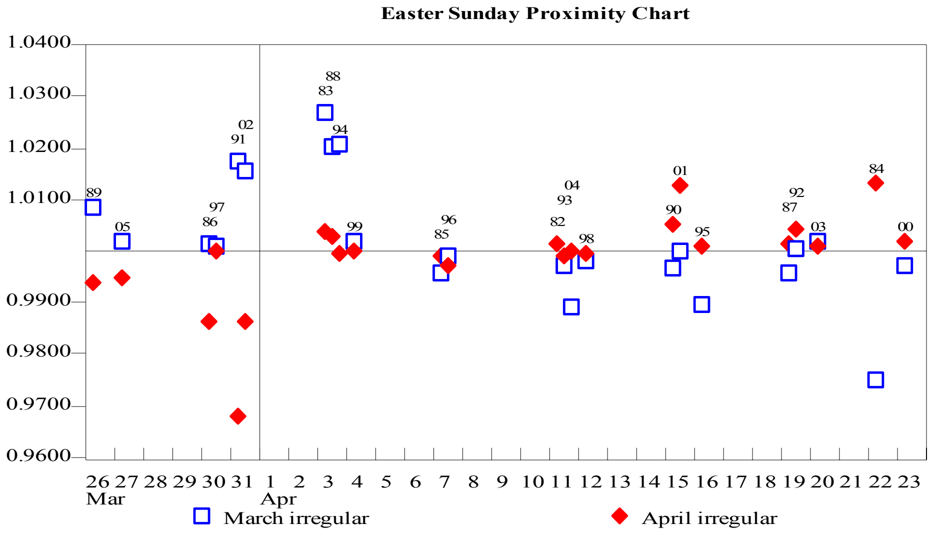


Figure 18. Easter Proximity chart for the state-aggregate "SuperMarket/grocery store retail series" (before correction).

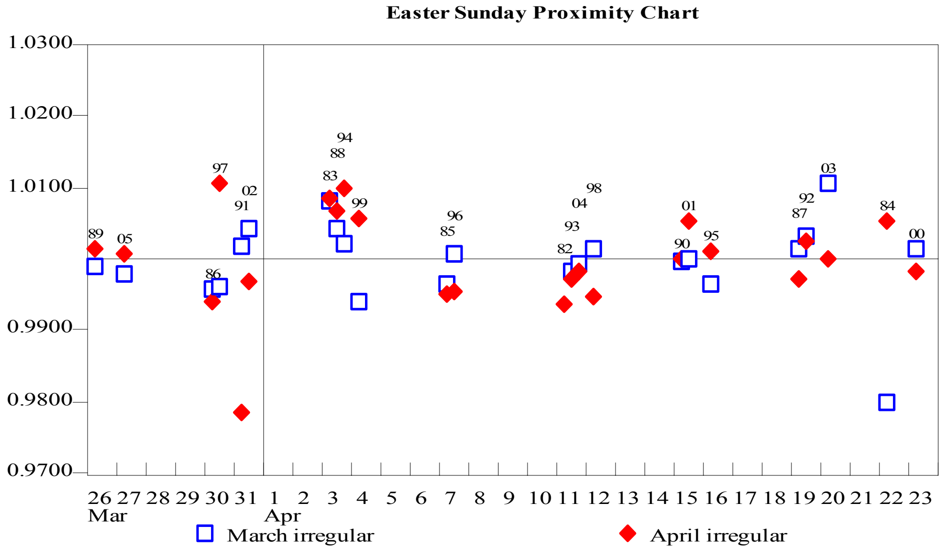


Figure 19. Easter proximity chart after correction using the SEASABS algorithm.

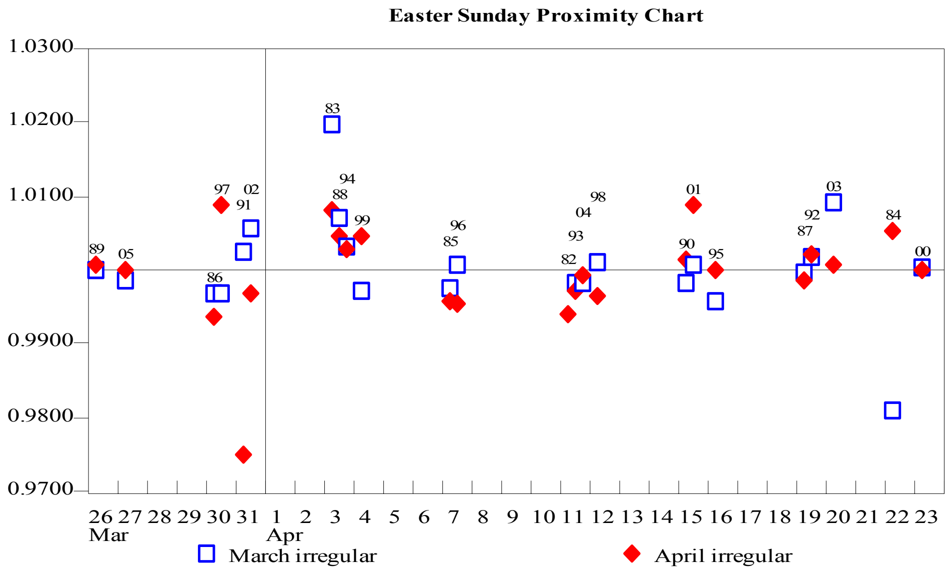


Figure 20. Easter proximity chart after correction with REGARIMA[w, h, p] = [7, 4, 2] (quad-linear regressor).

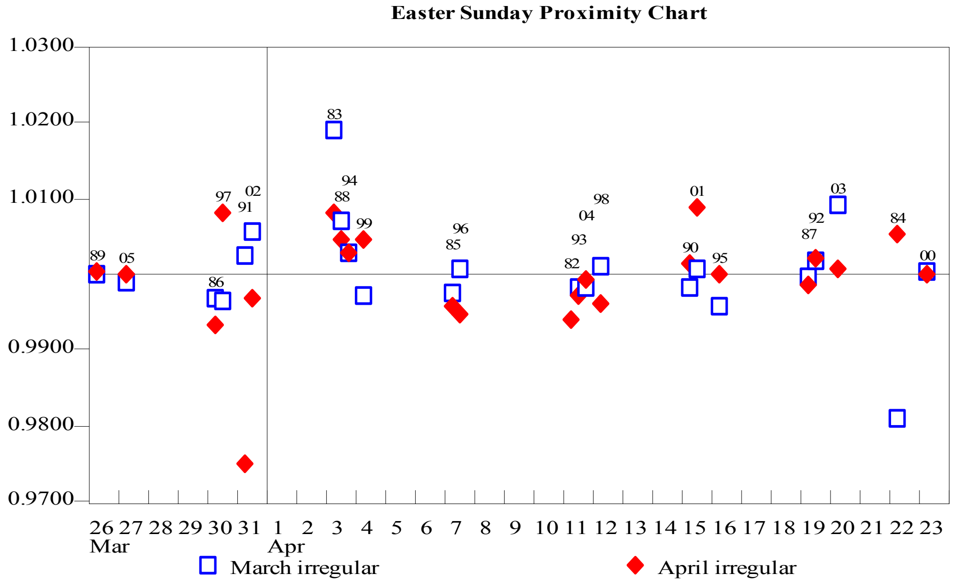


Figure 21. Easter proximity chart after correction with REGARIMA[w, h, p] = [7, 0, 2] (quad. model with No "during" Easter holiday regressor).

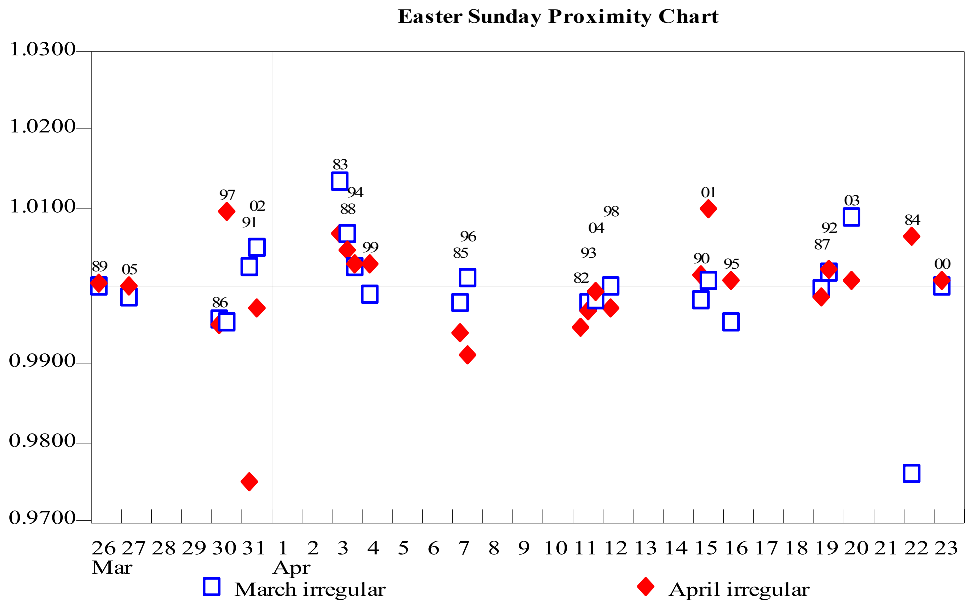


Figure 22. Easter proximity chart after correction with REGARIMA[w, h, p] = [3, 0, 2] (quad. model with <u>No</u> "during" Easter holiday regressor and 3-day pre-Easter window).

5. Analysis Results/Summary

Below is a summary of the case studies from section 4 with a number of diagnostics for choosing between appropriate models. Column descriptions are at the end of this section. The information in the first five columns are from X12-ARIMA output logs, while the results in the last three columns are from external computations performed with "personal" software - "personal" in the sense that no standard package was used. Results are discussed in section 6.

Table (i). Liquor retail series (1983 - 2005)

Model/test case [w, h, p]	**Parameter estimate(s)**	**standard error(s)**	**T-statistic for Regression (T_{crit} =1.97)**	**AICC statistic**	**σ^2 in Irreg. of MAR/early APR Easters**	**F-statistic for test: $\sigma^2 < \sigma^2_{NO\ CORR}$ (F_{crit} =0.495)**	**P-value for F-stat**
No correction	---	---	---	1134.687*	36.427e-05	1.000	0.500
SEASABS correction	---	---	---	1134.687*	8.728e-05	0.239	5.45e-04
Quad-Lin [7, 4, 2]	Eb=0.0357 Ed=-0.0053	0.00985 0.01194	3.62 -0.44	1114.272	5.796e-05	0.159	<1.4e-04
My model fit	Eb=0.03	0.01152	3.39	1116.05	7.825e-0	0.215	2.34e-04

[3, 4, 4]	90 Ed=-0.0109	0.01389	-0.79	8	5		
Quad only [7, 0, 2]	Eb=0.0323	0.00612	5.28	1112.339	5.763e-05	0.158	<1.4e-04
Quad-Lin [14, 4, 2]	Eb=0.0315 Ed=0.0019	0.00970 0.01112	3.25 0.17	1116.579	9.515e-05	0.261	1.05e-03
Quad only [3, 0, 2]	Eb=0.0322	0.00614	5.25	1112.939	5.381e-05	0.148	<1.4e-04
Quad only [1, 0, 2]	Eb=0.0299	0.00615	4.86	1116.356	9.377e-005	0.257	9.31e-04

(*) No EP regressors in the REGARIMA model were included (i.e. just plain ARIMA runs).

All models have ARIMA(p, d, q)(P, D, Q)$_s$ = (1,1,0)(0,1,1)$_{12}$

Table (ii). SuperMarket/Grocery retail series (1983 - 2005)

No correction	---	---	---	1666.847 *	15.983e-05	1.000	0.500
SEASABS correction	---	---	---	1666.847 *	5.110e-05	0.319	4.12e-03
Quad-Lin [7, 4, 2]	Eb=0.0130 Ed=0.0004	0.00344 0.00428	3.78 0.09	1638.018	6.364e-05	0.398	1.58e-02
Quad only [7, 0, 2]	Eb=0.0132	0.00212	6.25	1635.901	6.231e-05	0.389	1.39e-02
Quad only [3, 0, 2]	Eb=0.0134	0.00215	6.26	1635.586	5.553e-005	0.347	7.03e-03

(*) No EP regressors in the REGARIMA model were included (i.e. just plain ARIMA runs).

All models have ARIMA(p, d, q)(P, D, Q)$_s$ = (0,1,1)(0,1,1)$_{12}$

Column descriptions:

Model/test case:

The model or test case used for the Easter regressor: *w* is the window length (in days) up to and including the Thursday before Good Friday, *h* is the duration (days) of the Easter holiday period (nominally 4 days), and p is the power-law index of the cumulative retail activity leading up to Easter (See Figure 2 and derivation in section 1). In other words, the daily activity rate before and leading up to Easter scales as

$$\frac{dA}{dt} \propto t^{p-1},$$

where t is a time variable.

Parameter estimate(s):

User defined parameters estimated from the REGARIMA model. E_b is the "before" Easter activity parameter and E_d is the "during" Easter Holiday parameter. In general, the regression mean function used in the ARIMA model to ensure stationarity in the residuals, $y_t - \mu_t$, is defined as:

$$\mu_t = E_b x_b + E_d x_d .$$

where x_b and x_d are the corresponding "explanatory" variables defined by a monthly regression matrix:

$$x_b = \left(\frac{n}{w}\right)^p ; \quad x_d = \left(\frac{m}{h}\right) \text{ for MARCH}$$

$$x_b = -\left(\frac{n}{w}\right)^p ; \quad x_d = -\left(\frac{m}{h}\right) \text{ for APRIL}$$

$$x_b = 0; \quad x_d = 0 \text{ otherwise}$$

Thus, March is the reference month defining the regressor, and the symmetry $x_{b,d}$ (APR) = $-x_{b,d}$ (MAR) ensures that the correction for the Easter proximity effect has no net effect on the final adjusted series over a year, i.e.,

$$\sum_{month} x_b = \sum_{month} x_d = 0,$$

w and *h* were defined above and

$$n = \text{number of } w \text{ days in March,}$$

$$m = \text{number of } h \text{ days in April.}$$

When only one regression parameter is shown (e.g., just E_b), it means that the effect of *inactivity* from the second period (i.e., during the Easter holiday period) was intentionally ignored.

Standard errors:

Errors on each estimated model parameter corresponding to ~95% (2-sigma) confidence. These are from the X12-ARIMA output log derived as part of the REGARIMA model estimation. These facilitate as a diagnostic to judge the significance of each model parameter and/or the overall model.

T-statistic for Regression

From the X12-ARIMA output log. This assesses the statistical significance of each model parameter (E_b and E_d) . More specifically, it tests whether the estimated model parameter is significantly different from zero given its variance. The null hypothesis is that it is consistent with zero. The critical t-value for this "two-tailed" test corresponding to N - 1 = 166 degrees of freedom and P=0.05 is **t_{crit}=1.974.** T values above this can be considered significant.

AICC Statistic

From the X12-ARIMA output log. This statistic trades model fit (as measured by the log of the maximum likelihood for the best fit model), against model complexity (as measured by the total number of model parameters - both the ARMA and user-defined regression parameters). The lower this value, the more appropriate is the model for describing the series. The ARIMA component of the model was fixed at $(p, d, q)(P, D, Q)_s = (1,1,0)(0,1,1)_{12}$ for the *Liquor series* (Table i), and $(0,1,1)(0,1,1)_{12}$ for the *SuperMarket series* (Table ii). Therefore, differences in the AICC values are purely due to differences in the user-defined regressors.

σ^2 in Irreg. of MAR/early APR Easters

The sample variance of the X11-D13 irregulars for March and April (after correcting for the EP effect) across all years which have Easter Sunday falling in March or early April (before April 7). Between 1983 and 2005, there were 12 such years, so that $N = 24$. The sample variance is defined as:

$$\sigma^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(x_i - \langle x_i \rangle\right)^2,$$

where $\langle x_i \rangle$ is the sample mean.

F-statistic for test: $\sigma^2 < \sigma^2_{\text{NO CORR}}$

This is defined as the ratio of sample variances:

$$F = \frac{\sigma^2}{\sigma^2_{\text{NO CORR}}},$$

where $\sigma^2_{\text{NO CORR}}$ is the variance of the X11-D13 irregulars found before correcting for any EP effect (see first entry in tables above). This statistic is designed to test the hypothesis that the sample variance of the irregulars for a given EP model is less than that estimated using no correction. The null-hypothesis is that a specific model variance is consistent with no correction, or that any observed diminuition of the variance after correction with a specific model is just due to random chance. By examining how the Mean Squared Error (variance) in the irregular component after correction compares with the uncorrected (original) irregular pattern, we can use this as a further diagnostic on assessing the robustness of an EP model. The critical F-value for this "lower one-tailed" test corresponding to N - 1 = 23 degrees of freedom and P=0.05 is $\mathbf{F_{crit}=0.495}$. F-values below this can be considered significant, and as seen, all model variances are significantly lower than the "before correction" variance. These measures provide an alternative means of assessing the robustness of a model relative to the others.

P-value for F-stat (critical value =0.05)

The probability of obtaining an F-value greater than that quoted above by chance under the null-hypothesis.

6. Conclusions and Further Work?

- From an examination of the results in Tables (i) and (ii), we find that models with no reduced activity during the Easter holiday period appear to give best results (highlighted in red). In particular, models with the following parameter ranges give more-or-less the same minimal variance in the final irregulars, best quality of regression from T-stats, and overall AICC statistic:

$$3 \le w \le 8$$

$$h = 0$$

$$1.5 \le p \le 2.5$$

- Note that this result may just be specific to these two series. Need to analyse more series to explore the allowable parameter space (see below).

- In general, Easter proximity corrections using REGARIMA are better (i.e., in terms of the reducing the MSE in residuals) than the current SEASABS method.

- Constraints on [w, h, p] using the dependence of observed March/April ratios on Easter date do not appear to be robust. More series, covering longer time spans are needed to increase the Easter proximity (EP) effect signal-to-noise in the irregular pattern. Daily retail series data may also be used to constrain EP models. The data however are scarce.

- In the end, the best regressor is that which reduces the EP systematic pattern to a level consistent with the overall irregular behaviour in a series. You may be beating a dead horse by "over-tuning" a regressor model.

Where to from here?

- As a further and possibly more powerful diagnostic on choosing an appropriate EP regressor, we may need to look at the actual regARIMA model residuals from the original data. Statistical tests based on the X11-D13 irregulars (as analysed here) may not be robust enough to discriminate between models since they are the product of an iterative seasonal adjustment process. It is not yet understood how this process modifies the distribution of the underlying irregulars. It is inevitable that this process makes the irregular component noisier, therefore masquerading any residual systematic pattern from an EP effect that we are trying to measure or correct. In the end, this makes "regressor fine-tuning" difficult.

❒ Next step would be to use the ARIMA residuals to detect an EP effect since they will be more sensitive and robust on discriminating between different regressors.

❒ Outliers detection/removal before confirming the existence of an EP effect (not done on D13 method).

❒ In the end, we may want to produce EP charts using exclusively ARIMA residuals.

❒ Explore the coupling between the ARIMA and regressor-model parameters. Use of REGARIMA residuals will allow better separation of dependencies on the full parameter space.

❒ Can we accomodate these improvements in production work?

- The selection of an appropriate regressor may depend on the type of time series under analysis. For instance, even though retail-activity series (as a group) are expected to show the strongest EP effects, a variation between the series for different retail types is expected - e.g., liquor stores open over the Easter holiday period. Food/grocery retail however may be strongest in the period leading up to Easter. These is some hint of this in the above analysis. Another example is tourism which is likely to peak during the Easter holiday period.
    - ❐ Since the regressor model may be series specific, we may want to pick the best regressor (by cycling through a list of sensible regressors) under REGARIMA according to the ARIMA model acceptance criteria (e.g., AICC).
    - ❐ We may only want to correct for an EP effect if it really makes sense to have one for a series and if it's in the expected direction. e.g. any activity correlated with retail trade is expected to show a hint of the EP effect such that ratio of MAR/APR irregulars > 1 and activity correlated with tourism may also include ratios < 1 (e.g., when the holiday period falls in early April).

- Are evolutionay effects in the strength of an EP effect important? An examination of series prior to 1983 by Leung et al. (1999) shows evidence for this. For long series, we may need to split it up and apply different regressors. This is optimistic and it's best to keep it simple.

- Do we need to define a separate regressor model (for regARIMA) for quarterly series? The EP effect is expected to be less prevalent in a quarterly series, although it is still expected since Easter falling near the March/April boundary is equivalent to the 1st/2nd quarter boundary. Seasonally adjusted quarterly series derived from seasonally adjusted and EP-corrected monthly series may take care of this in the long term.