6. Extending the ARMA model: Seasonality and trend

Edward Ionides

2018-01-30

Contents

Objectives

- Monthly time series often exhibit seasonal variation. January data are similar to observations at a different January, etc.
- Many time series exhibit a trend.
- We wish to extend the theoretical and practical elegance of the ARMA framework to cover these situations.

Seasonal autoregressive moving average (SARMA) models

- A general SARMA$(p, q) \times (P, Q)_{12}$ model for monthly data is

$$[S1] \qquad \phi(B)\Phi(B^{12})(Y_n - \mu) = \psi(B)\Psi(B^{12})\epsilon_n,$$

  Here, $\phi(B)$ is the monthly polynomial and $\Phi(B^{12})$ is the seasonal (annual) polynomial

  where $\{\epsilon_n\}$ is a white noise process and

$$\begin{aligned}
\mu &= \mathbb{E}[Y_n] && (1)\\
\phi(x) &= 1 - \phi_1 x - \cdots - \phi_p x^p, && (2)\\
\psi(x) &= 1 + \psi_1 x + \cdots + \psi_q x^q, && (3)\\
\Phi(x) &= 1 - \Phi_1 x - \cdots - \Phi_p x^P, && (4)\\
\Psi(x) &= 1 + \Psi_1 x + \cdots + \Psi_q x^Q. && (5)
\end{aligned}$$

- We see that a SARMA model is a special case of an ARMA model, where the AR and MA polynomials are factored into a **monthly** polynomial in B and an **annual** polynomial in B^{12}. The annual polynomial is also called the **seasonal** polynomial.
- Thus, everything we learned about ARMA models (including assessing causality, invertibility and reducibility) also applies to SARMA.

- One could write a SARMA model for some **period** other than 12. For example, a SARMA$(p,q)\times(P,Q)_4$ model could be appropriate for quarterly data. In principle, a SARMA$(p,q)\times(P,Q)_{52}$ model could be appropriate for weekly data, though in practice ARMA and SARMA may not work so well for higher frequency data.
- Consider the following two models:

$$[S2] \qquad Y_n = 0.5Y_{n-1} + 0.25Y_{n-12} + \epsilon_n,$$

$$[S3] \qquad Y_n = 0.5Y_{n-1} + 0.25Y_{n-12} - 0.125Y_{n-13} + \epsilon_n,$$

Question: Which of [S2] and/or [S3] is a SARMA model?

Model S3 is a SARMA$(1,0)\times(1,0)_{12}$. $(1-\frac{1}{2}B)(1-\frac{1}{4}B^{12})Y_n = \epsilon_n$. Thus, $Y_n = \frac{1}{2}Y_{n-1} + \frac{1}{4}Y_{n-12} - \frac{1}{8}Y_{n-13} + \epsilon_n$. Model S2 is *not* a SARMA model.

Question: Why do we assume a multiplicative structure in [S1]?

This gives a convenient class of models, but there is not particular scientific motivation for it.

- What theoretical and practical advantages (or disadvantages) arise from requiring that an ARMA model for seasonal behavior has polynomials that can be factored as a product of a monthly polynomial and an annual polynomial?

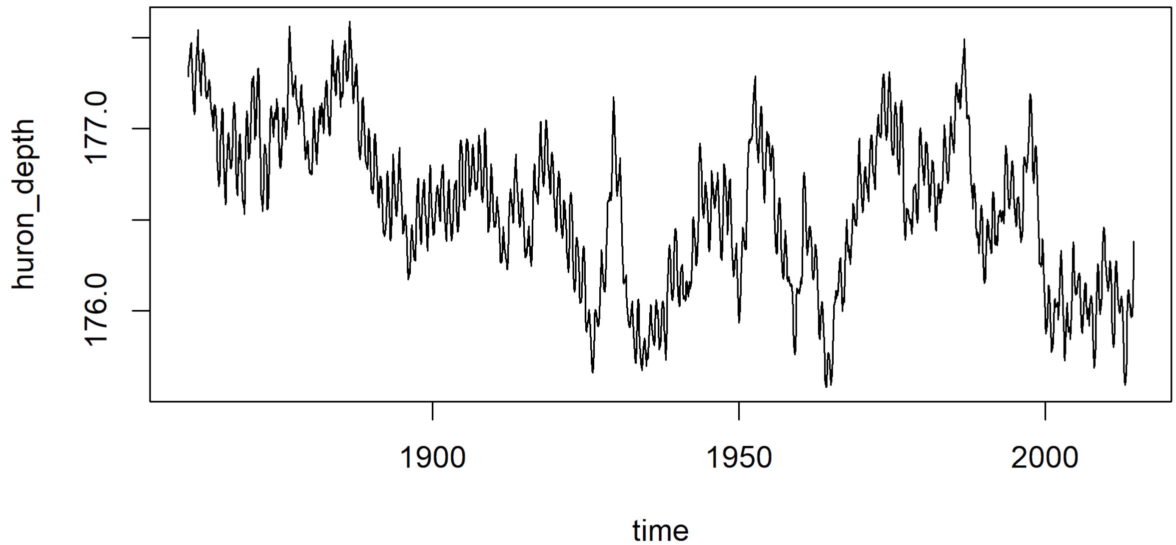
Fitting a SARMA model

- Let's do this for the full, monthly, version of the Lake Huron depth data described in Section 5.5.
- The data were read into a dataframe calles `dat`

```
head(dat)
```

```
##         Date Average year month
## 1 1860-01-01 177.285 1860     1
## 2 1860-02-01 177.339 1860     2
## 3 1860-03-01 177.349 1860     3
## 4 1860-04-01 177.388 1860     4
## 5 1860-05-01 177.425 1860     5
## 6 1860-06-01 177.461 1860     6
```

```
huron_depth <- dat$Average
time <- dat$year + dat$month/12 # Note: we treat December 2011 as time 2012.0, etc
plot(huron_depth~time,type="l")
```



- Now, we get to fit a model. Based on our previous analysis, we'll go with AR(1) for the annual polynomial. Let's try ARMA(1,1) for the monthly part. In other words, we seek to fit the model

$$(1 - \Phi_1 B^{12})(1 - \phi_1 B)Y_n = (1 + \psi_1 B)\epsilon_n.$$

Left-most part, i.e. $(1 - \Phi_1 B^{12})$ is the AR(1) annual polynomial - there's no corresponding annual poly on right b/c AR(1) doesn't include a moving-average (MA) term.

Poly $(1 - \phi_1 B)$, along with poly on right, i.e. $(1 + \psi_1 B)$, make up the ARMA(1,1) polynomials for the monthly part of the model. (See notes4 The General ARMA Model)

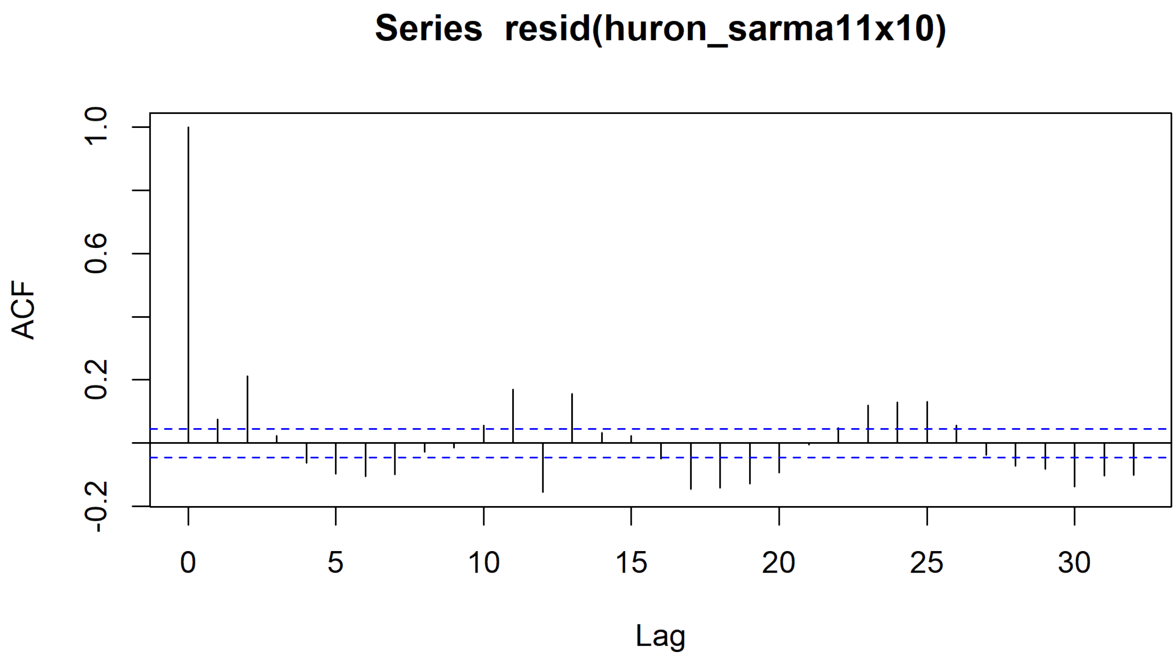
```
huron_sarma11x10 <- arima(huron_depth,
   order=c(1,0,1),
   seasonal=list(order=c(1,0,0),period=12)
)
huron_sarma11x10
```

In code above, order args are p, diff, q (no difference here), still stationary model.

```
##
## Call:
## arima(x = huron_depth, order = c(1, 0, 1), seasonal = list(order = c(1, 0, 0),
##     period = 12))
##
## Coefficients:
##          ar1     ma1    sar1  intercept
##       0.9641  0.3782  0.5104   176.5714
## s.e.  0.0063  0.0203  0.0218     0.0909
##
## sigma^2 estimated as 0.002592:  log likelihood = 2884.36,  aic = -5758.72
```

- Residual analysis is similar to what we've seen for non-seasonal ARMA models.
- We look for residual correlations at lags corresonding to multiples of the period (here, 12, 24, 36, ...) for misspecified annual dependence.

```r
acf(resid(huron_sarma11x10))
```



Question: What do you conclude from this residual analysis? What would you do next?

The current model isn't capturing all the dependence in the data. Maybe add AR(2) term to fit the lag residual correlation. (Adding more AR terms could be good b/c we see an oscillatory behavior in the residual ACF)

ARMA models for differenced data

- Applying a difference operation to the data can make it look more stationary and therefore more appropriate for ARMA modeling.
- This can be viewed as a **transformation to stationarity**
- We can transform the data $y^*_{1:N}$ to $z^*_{2:N}$

$$z^*_n = \Delta y^*_n = y^*_n - y^*_{n-1}.$$

- Then, an ARMA(p,q) model $Z_{2:N}$ for the differenced data $z^*_{2:N}$ is called an **integrated autoregressive moving average** model for $y^*_{1:N}$ and is written as ARIMA(p,1,q).
- Formally, the ARIMA(p,d,q) model with intercept μ for $Y_{1:N}$ is

$$[S4] \qquad \phi(B)\big((1-B)^d Y_n - \mu\big) = \psi(B)\epsilon_n,$$

  where $\{\epsilon_n\}$ is a white noise process; $\phi(x)$ and $\psi(x)$ are the ARMA polynomials defined previously.
- It is unusual to fit an ARIMA model with $d > 1$.
- We see that an ARIMA(p,1,q) model is almost a special case of an ARMA(p+1,q) model with a **unit root** to the AR(p+1) polynomial.

Question: why "almost" not "exactly" in the previous statement?

This is almost true b/c it treates the mean slightly differently.

Why fit an ARIMA model?

- There are two reasons to fit an ARIMA(p,1,q) model

1. You may really think that modeling the differences is a natural approach for your data. The S&P 500 stock market index analysis in Section 3.5 is an example of this, as long as you remember to first apply a logarithmic transform to the data.
2. Differencing often makes data look "more stationary" and perhaps it will then look stationary enough to justify applying the ARMA machinery.

- We should be cautious about this second reason. It can lead to poor model specifications and hence poor forecasts or other conclusions.
- The second reason was more compelling in the 1970s and 1980s. With limited computing power and the existence of computationally convenient (but statistically inefficient) method-of-moments algorithms for ARMA, it made sense to force as many data analyses as possible into the ARMA framework.
- ARIMA analysis is relatively simple to do. It has been a foundational component of time series analysis since the publication of the influential book "Time Series Analysis" by Box and Jenkins (1st edition, 1970) which developed and popularized ARIMA modeling. A practical approach is:

1. Do a competent ARIMA analysis.
2. Identify potential limitations in this analysis and remedy them using more advanced methods.
3. Assess whether you have in fact learned anything from (2) that goes beyond (1).

Question: What is the trend of the ARIMA(p,1,q) model?

- Hint: recall that the ARIMA(p,1,q) model specification for $Y_{1:N}$ implies that $Z_n = (1-B)Y_n$ is a stationary, causal, invertible ARMA(p,q) process with mean μ. Now take expectations of both sides of the difference equation.

$\mathbb{E}[(1-B)Y_n] = \mu \implies (1-B)\mathbb{E}(Y_n) = \mu \implies \mathbb{E}(Y_n) - \mathbb{E}(Yn-1) = \mu \implies \mathbb{E}(Y_n) = \mu + \mathbb{E}(Y_{n-1})$. Thus, $\mathbb{E}(Y_n) = n\mu + constant$, i.e. the difference z_n has constant mean μ, but Y_n has mean increasing/decreasing with n (linear mean, non-stationary).

Question: What is the trend of the ARIMA(p,d,q) model, for general d?

Note. $d=2 \implies$ quadratic trend, $d=3 \implies$ cubic trend, etc.

The SARIMA$(p,d,q) \times (P,D,Q)$ model

- Combining integration of ARMA models with seasonality, we can write a general SARIMA$(p,d,q) \times (P,D,Q)_{12}$ model for nonstationary monthly data, given by

$$[S5] \qquad \phi(B)\Phi(B^{12})\big((1-B)^d(1-B^{12})^D Y_n - \mu) = \psi(B)\Psi(B^{12})\epsilon_n,$$

  where $\{\epsilon_n\}$ is a white noise process, the intercept μ is the mean of the differenced process $\{(1-B)^d(1-B^{12})^D Y_n\}$, and we have ARMA polynomials $\phi(x)$, $\Phi(x)$, $\psi(x)$, $\Psi(x)$ as in model [S1].

- The SARIMA$(0,1,1) \times (0,1,1)_{12}$ model has often been used for forecasting monthly time series in economics and business. It is sometimes called the **airline model** after a data analysis by Box and Jenkins (1970).

Modeling trend with ARMA noise.

- A general **signal plus noise** model is

$$[S6] \qquad Y_n = \mu_n + \eta_n,$$

  where $\{\eta_n\}$ is a stationary, mean zero stochastic process, and μ_n is the mean function.

- If, in addition, $\{\eta_n\}$ is uncorrelated, then we have a **signal plus white noise** model. The usual linear trend regression model fitted by least squares in Section 2.4 corresponds to a signal plus white noise model.

- We can say **signal plus colored noise** if we wish to emphasize that we're not assuming white noise.

- Here, **signal** and **trend** are used interchangeably. In other words, we are assuming a deterministic signal.

- At this point, it is natural for us to consider a signal plus ARMA(p,q) noise model, where $\{\eta_n\}$ is a stationary, causal, invertible ARMA(p,q) process with mean zero.
- As well as the $p+q+1$ parameters in the ARMA(p,q) model, there will usually be unknown parameters in the mean function. In this case, we can write

$$\mu_n = \mu_n(\beta)$$

where β is a vector of unknown paramters, $\beta \in \mathbb{R}^K$.

- We write θ for a vector of all the $p+q+1+K$ parameters, so

$$\theta = (\phi_{1:p}, \psi_{1:q}, \sigma^2, \beta).$$

Linear regression with ARMA errors

- When the trend function has a linear specification,

$$\mu_n = \sum_{k=1}^{K} Z_{n,k}\beta_k,$$

the **signal plus ARMA noise** model is known as **linear regression with ARMA errors**.

Here, μ_n encodes the trend for the data.

- Writing Y for a column vector of $Y_{1:N}$, μ for a column vector of $\mu_{1:N}$, η for a column vector of $\eta_{1:N}$, and Z for the $N \times K$ matrix with (n,k) entry $Z_{n,k}$, we have a general linear regression model with correlated ARMA errors,

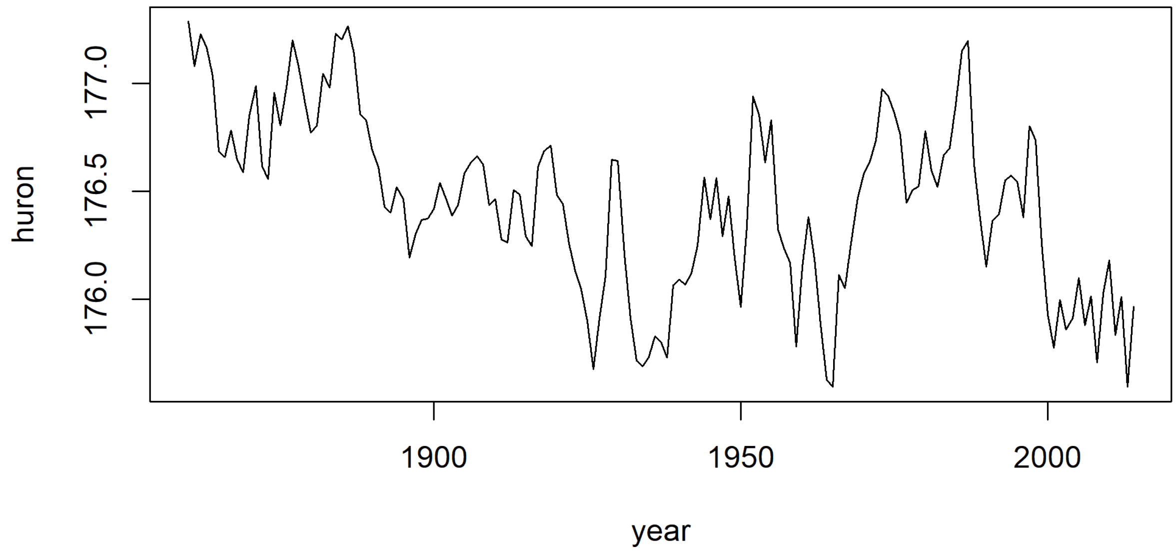
$$Y = Z\beta + \eta.$$

- Maximum likelihood estimation of $\theta = (\phi_{1:p}, \psi_{1:q}, \sigma^2, \beta)$ is a nonlinear optimization problem. Fortunately, `arima` in R can do it for us, though as usual we should look out for signs of numerical problems.
- Data analysis for a linear regression with ARMA errors model, using the framework of likelihood-based inference, is therefore procedurally similar to fitting an ARMA model.
- This is a powerful technique, since the covariate matrix Z can include other time series. We can evaluate associations between different time series. With appropriate care (since **association is not causation**) we can draw inferences about mechanistic relationships between dynamic processes.

Example: Looking for evidence of systematic trend in the depth of Lake Huron

- Let's restrict ourselves to annual data, say the January depth.

```
monthly_dat <- subset(dat, month==1)
huron <- monthly_dat$Average
year <- monthly_dat$year
plot(x=year,y=huron,type="l")
```



- Visually, there seems some evidence for a decreasing trend, but there are also considerable fluctuations.
- Let's test for a trend, using a regression model with Gaussian AR(1) errors. We have previously found that this is a reasonable model for these data.
- First, let's fit a null model.

```
fit0 <- arima(huron,order=c(1,0,0))
fit0
```

```
##
## Call:
## arima(x = huron, order = c(1, 0, 0))
##
## Coefficients:
##          ar1  intercept
##       0.8694   176.4588
## s.e.  0.0407     0.1234
##
## sigma^2 estimated as 0.04368:  log likelihood = 22,  aic = -38
```

- Now, we can compare with a linear trend model.

```
fit1 <- arima(huron,order=c(1,0,0),xreg=year)
fit1
```

```
##
## Call:
## arima(x = huron, order = c(1, 0, 0), xreg = year)
##
## Coefficients:
##          ar1  intercept     year
##       0.8240   186.0146  -0.0049
```

```
## s.e.  0.0451     3.7417   0.0019
##
## sigma^2 estimated as 0.0423:  log likelihood = 24.62,  aic = -41.25
```

In code above, β is coefficient associated with year (mean function or trend only a function of year)

- To talk formally about these results, we'd better write down a model and some hypotheses. Writing the data as $y^*_{1:N}$, collected at years $t_{1:N}$, the model we have fitted is

$$(1 - \phi_1 B)(Y_n - \mu - \beta t_n) = \epsilon_n,$$

  Here, our regression model is $Y_n = \mu + \beta t_n + e_n = \mu_n + e_n$ where e_n are the errors of the regression model. Then $(1 - \phi_1 B)e_n = \epsilon_n$ is an AR(1) model for the errors $e_n = Y_n - \mu - \beta t_n$.

  where $\{\epsilon_n\}$ is Gaussian white noise with variance σ^2. Our null model is

$$H^{\langle 0 \rangle} : \beta = 0,$$

  and our alternative hypothesis is

$$H^{\langle 1 \rangle} : \beta \neq 0.$$

Question: How do we test $H^{\langle 0 \rangle}$ against $H^{\langle 1 \rangle}$?

- Construct two different tests using the R output above.

  1) t-test for β (Fiher info CI), 2) likelihood ratio test (profile CI). Here, $\Delta loglik = 2.62$. The LRT compares $2\Delta loglik$ with χ^2. We obtain a p-value of $1 - pchisq(5.24, 1) = 0.02$. Thus, we reject the null model.

- Which test do you prefer, and why?
- How would you check whether your preferred test is indeed better?

Question: What other supplementary analysis could you do to strengthen your conclusions?

simulation under the null hypothesis
