

## When the blind curve is finite... A NOTE ON ARFIMA MODEL SELECTION IN [R]

The *forecast* and *fracdiff* packages were used.

The strategy was derived based on suggestions found in:

- Reisen, V. A., & Lopes, S. (1999). Some simulations and applications of forecasting long-memory time-series models. *Journal of Statistical Planning and Inference*, 80(1-2), 269–287. doi:10.1016/S0378-3758(98)00254-7
- Silva, E. M., Franco, G. C., Reisen, V. a., & Cruz, F. R. B. (2006). Local bootstrap approaches for fractional differential parameter estimation in ARFIMA models. *Computational Statistics & Data Analysis*, 51(2), 1002–1011. doi:10.1016/j.csda.2005.10.007
- Stadnitski, T. (2012). Measuring fractality. *Frontiers in physiology*, 3(May), 127. doi:10.3389/fphys.2012.00127
- Wagenmakers, E.-J., Farrell, S., & Ratcliff, R. (2004). Estimation and interpretation of 1/f alpha noise in human cognition. *Psychonomic bulletin & review*, 11(4), 579–615.
- Xiu, J., & Jin, Y. (2007). Empirical study of ARFIMA model based on fractional differencing. *Physica A: Statistical Mechanics and its Applications*, 377(1), 138–154. doi:10.1016/j.physa.2006.11.030

The following strategy was used:

1. Estimate  $d$  using priodogram based estimate *dsperio*.
2. If  $d$  is significantly different from zero, difference the series by *dsperio*.
3. Find estimates for the order of  $p$  and  $q$  (max=2) using *auto.arima* possibly on the differenced data.
4. Fit an ARFIMA( $p,d,q$ ) model on the undifferenced data using the order found in step 3. If  $d$  in step 2 was significantly lower than 0 ARFIMA is fitted with a *drange=c(-1,0)* otherwise *drange=c(0,1)*
5. When an error is returned suggesting changing  $h$  in order to get an SE on  $d$ , the value of  $h$  is iteratively increased until a fit is possible or an upper limit of .001 is reached.
6. An ARIMA( $p,0,d$ ) is fitted of the same order found in 3.
7. The best fitting model is selected by selecting lowest AIC

This strategy allows fitting data with anti-persistent correlations ( $d$  will take on a negative value).

The modeling procedure prefers ARIMA(2,0,1) and ARIMA(2,0,2) over ARFIMA models of the same order. ARFIMA(2, $d$ ,2) appears to fit well to PSD slopes of -3 to -2 whereas

ARFIMA(2, $d$ ,1) mostly fits to PSD slopes around 0. A clear pattern is hard to find.

