

# Tímaraðalíkön fyrir langt minni

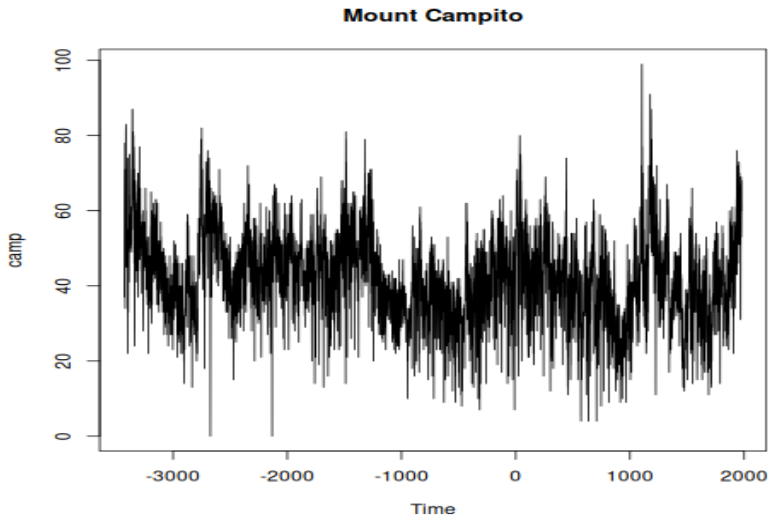
Helgi Tómasson  
helgito@hi.is

24. október 2023

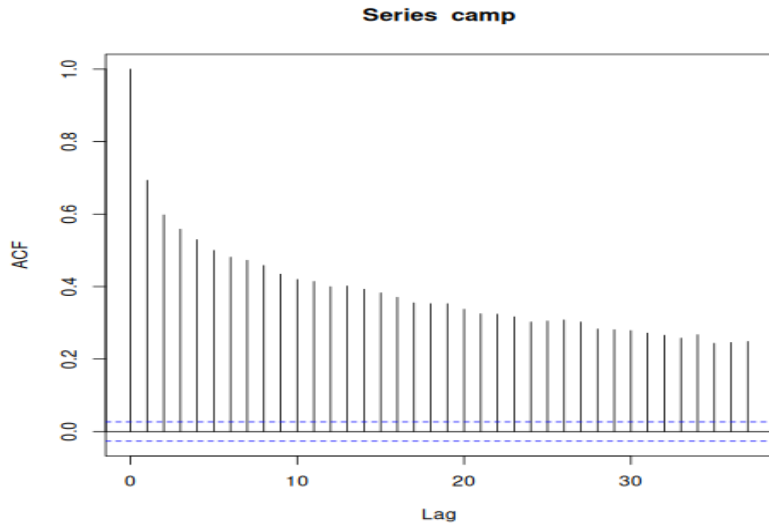
# Skipulag fyrirlestrar

- Hvað er tímaraðalíkan? Hvernig á að lýsa þeim?
- Hvað er langt minni?
- Sýnishorna af röð. Hvað á að gera?

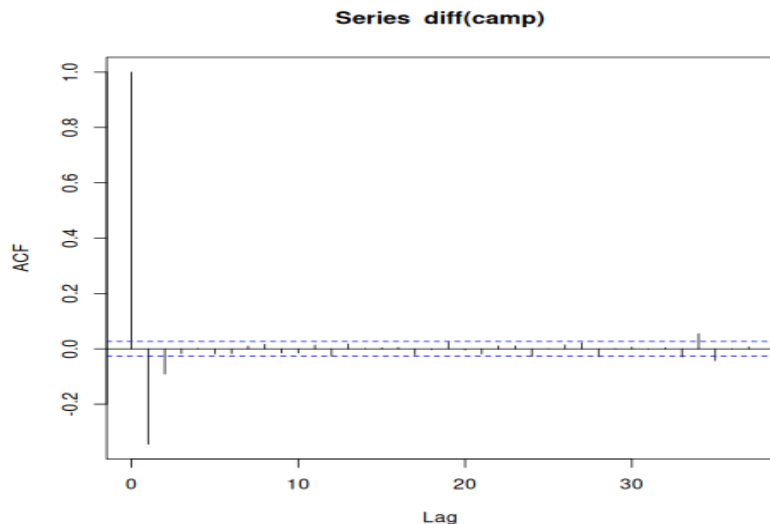
# Trjáhringir úr Mount Campito



# Sjálffylgnifall



## Sjálffylgnifall fyrsta mismunar



Box/Jenkins myndu álykta að þetta sé  $ARIMA(0,1,2)$ . Hvers vegna er það rangt?

# Sjálffylgniföll og spektrur

- Random ferli (*stochastic process*),  $X_t$ , er ferli þar sem  $X_t$  er háð  $X_s$ .
- Tímaraðgreining gengur út á að gögn séu mælingar á random ferli.
- Ef tilraun er ekki endurtakanleg, þarf að gera ráð fyrir einhvers konar stöðugleikahutaki. T.d. að meðaltal og staðalfrávik  $X_t$  séu fyrrsjáanleg föll (af  $t$ , t.d. leitni, árstíð, skali) ásamt því að tengsl yfir tíma séu föst.
- Tengslum yfir tíma er lýst með sjálffylgni (*autocorrelation*), falli,  $\rho(k)$ , eða sköluð Fourier-formi sjálffylgni fallsins, spektri (*spectrum*).

$$E(X_t) = 0, \quad E(X_t X_{t-k}) = \gamma(|k|), \quad \rho(k) = \frac{\gamma(k)}{\gamma(0)}$$

$$f(\lambda) = \frac{\sigma^2}{2\pi} \sum_{k=-\infty}^{\infty} \exp(-ik)\rho(k), \quad \text{spektrur,} \quad \sigma^2 = \gamma(0).$$

# Hefðbundin ARMA líkön



$$Y_t = \phi_1 Y_{t-1} + \dots + \phi_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q}$$
$$E(\varepsilon_t) = 0, \text{ when } t \neq s \quad E(\varepsilon_t \varepsilon_s) = 0, E(\varepsilon_t^2) = \sigma_\varepsilon^2.$$

- Það sem einkennir þau er að nútíminn ( $Y_t$ ) er útskýrður með nýliðnum ( $Y_{t-k}$ ) gildum og nýliðnum skellum ( $\varepsilon_{t-j}$ ). Skellur nútímans ( $\varepsilon_t$ ) hefur enga fylgni við fortíðina.
- Þetta er „línulegt“ í þeim skilningi að besta spá línuleg samantekt fortíðar.
- Ef AR(1),  $Y_t = \phi Y_{t-1} + \varepsilon_t$  er skoðað:  
 $\Delta Y_t = Y_t - Y_{t-1} = (\phi - 1) Y_{t-1} + \varepsilon_t$   
að það líkist línulegri diffurjöfnu  $y' = ay$ .
- Ef  $-1 < \phi < 1$ , er AR(1) ferlið sístætt og það svarar til þess að  $a < 0$  í diffurjöfnunni og lausnarferlið stefnir þá á 0 með veldishraða.
- Sjálffylgniföll sístæðra ARMA ferla stefna á 0 með veldishraða og hafa því skammtímaminni.

## Notkun í hagrannsóknum ofl.

- Box & Jenkins (1970) stungu upp á tæknilegri nálgun við gerð spálíkana með tímaraðagögnum. Kjarninn var ARIMA( $p,d,q$ ) líkan:  
$$(1 - L)^d \Phi(L) Y_t = \Theta(L) \varepsilon_t$$
- Gangurinn var ca. 1) ákveða varíans-stabilizing transform (log í hagfræði), finna hve oft ætti að taka mismun  $d$  og að ákveða dýnamískar tafir,  $p$  og  $q$ . 2) meta parametra,  $\phi_i$ ,  $\theta_i$  og  $\sigma$ . 3) skoða metna afgangslíði, þeir eiga að líkjast white-noise og 4) ef afgangslíðir virðast skikkanlegir reikna spá spámörk.
- Hversu oft var tekinn mismunur,  $d$ , fór eftir því hversu ósístæður ferlið virtist. ARIMA(0,1,0) er random-walk.
- Hagfræðistéttin (1970+ ) var ekki impressed.



- Spektralfall ARMA ferlis er:

$$f(\lambda) = \frac{\sigma^2 |\Theta(\exp(-i\lambda))|^2}{2\pi |\Phi(\exp(-i\lambda))|^2},$$

$$\Phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p,$$

$$\Theta(z) = 1 - \theta_1 z - \dots - \theta_q z^q.$$

- Athugið að  $V(Y_t) = \int_{-\pi}^{\pi} f(\lambda) d\lambda$ .
- Ekkert vandamál ef  $f(\lambda)$  er endanlegt allstaðar.

## Langt minni

- Beran (1994) er lykilheimild varðandi langtímaminni. Þar segir að ferli sýni langtíma minni ef spektralfallið er óendanlegt í að minnsta kosti einni tíðni. Það jafngildir því að sjálffylgnifallið sé ekki tölugildissummanlegt.
- Meðfærilegur flokkur líkana sem nær þeim eiginleika eru ARFIMA(p,d,q)  
 $(1 - L)^d \Phi(L) Y_t = \Theta(L) \varepsilon_t$ ,  
þar sem  $d$  er leyft að vera brot.
- Hugmyndin er að vinna í and Box & Jenkins (1970) og láta  $\phi_i$  og  $\theta_i$  ná skammtímasveiflum og nota  $d$  sem mælikvarða á minnið.
- Sjálffylgnifall ARFIMA deyr út polynomial, þ.e. mun hæggar en með veldishraða. Áhrif fortíðar deyja mjög hægt út.
- Stærðfræðilega eru svona ferli ekki Markov né martingale.
- Spectralfall ARIMA(p,d,q) er á forminu:

$$f(\lambda) = \frac{\sigma^2}{2\pi} \underbrace{|(1 - \exp(-i\lambda))|^{-2d}}_{\text{long-memory}} \frac{|\Theta(\exp(-i\lambda))|^2}{|\Phi(\exp(-i\lambda))|^2}$$

# Ýmsir eiginleikar

- ARFIMA( $p,d,q$ ) með  $0 < d < 1/2$  hefur þann eiginleika að  $f(0) = \infty$ ,  $\int f(\lambda)d\lambda = \text{VAR}(Y(t))$  og auto-covariancar  $\gamma(k)$
- Vegna þess hve mælingar eru háðar verða ályktanir um t.d. meðaltal  $\mu = E(Y(t))$  og þýðingu ytri stærða frábrugðnar hefðbundinni ályktanafræði.
- Flóknar formúlur gilda um dreifingu úrtaksmeðaltals. Staðalfrávik þess er stundum ekki  $\propto T^{-1/2}$  heldur  $\propto T^{-1/2+d}$ . Þ.e. það er erfitt að álykta um meðaltal.
- Með því að sigta 0-tíðni frá og nota spektral aðferðir getur verið auðveldara að álykta um  $d$  en  $\mu$ .

# Ýmsar útvíkkarir og hagnýtingar

- Baillie & Chung (2002) skoða áhringi úr trjám.
- Chan & Palma (1998) sýna state-space nálgun.
- Comte & Renault (1996) sýna hvernig megi skilgreina fractional  $W$  í samfelldum tíma í mörgum víddum.
- Dagsvik, Fortuna & Moen (2020) meta long-memory raðir fyrir mælt hitastiga á ca. 100 veðurstöðvum í 200 ár ásamt öðrum hitastigsgögnum í ca. 2000 ár.
- Maddanu (2023) lýsir stuttminnis-ferli (sést líka hjá Hosking (1981)), FerARMA, með eiginleika sem minna á long-memory,  $\Phi(L)^{\delta_1} Y_t = \Theta(L)^{\delta_2} \varepsilon_t$ .
- Johansen (2008) tekur fyrir fractional vektorkerfi og í seinni greinum hvernig skuli álykta um fractional co-integration.

- Hurst (1951) tók saman nokkur hundruð ára gögn um flóðahæð í Níl og sá að sjálffylgnifallið dó hægt út. Hann fann upp hugtakið Hurst exponent= $d+1/2$  (fyrir normaldreifingu??).
- Granger (1980) rekur hvernig long-memory ferli getur myndast sem summa margra AR(1) ferla.

## Nokkur R forrit

- longmemo arfima artfima LongMemoryTS fracdiff, rugarch, forecast, garma, vgam, waveslim, tsvge, o.s.frv.
- garma=Gegenbauer-ARMA, Gegenbauer er ákveðin fjöldskeylda af margliðum. Gegenbauer ARMA leyfir óendanlegt spectur í fleiri tíðnum en 0, t.d. vegna árstíða.
- Mjög tæknilegt bæði stærðfræðilega (ekki Markov og ekki Martingale) og reiknifræðilega. Ekki til endanleg state-space framsetning og því tímafrekt að reikna likelihoodfall.
- Samfelldur tími er áskorun. Í strjálum tíma þar sem engin gildi vantar má nota Durbin-Levinson algoritmann. Spektral aðferðir (nálganir) mikið notaðar.

## Að reikna spectral fall

- R-skipanir til að reikna spectral density fyrir ARFIMA(0,d,0).

```
lam=0.2
```

```
d=0.3
```

```
sigmax=1
```

```
fspec=function(lam){return(sigmax*1/(2*pi)*  
(2*abs(sin(lam/2)))^(-(2*d)))}
```

```
#
```

```
#
```

```
# Úr Beran, 1994, bls. 63
```

- Með ljótri algebru má finna auto-covariance fallið. R skipanir:

```
k=2
```

```
sigmax=1
```

```
gam=function(k){
```

```
return(sigmax*(-1)^k*gamma(1-2*d)/(gamma(k-d+1)*  
gamma(1-k-d))
```

```
}
```

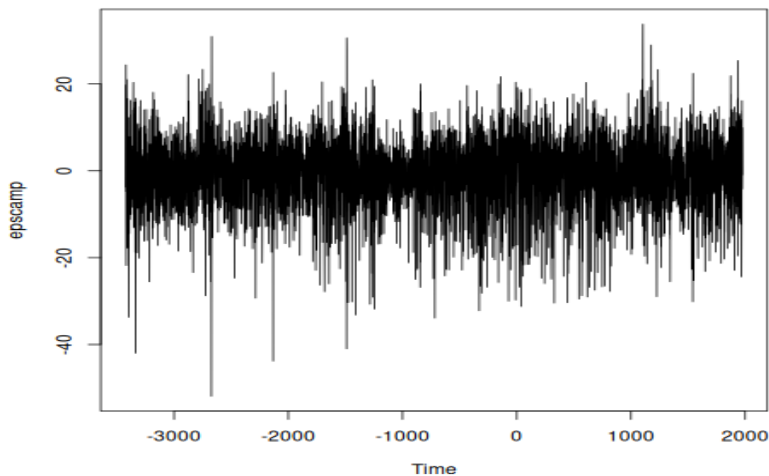
## Útkoma úr ARFIMA(0,d,0) fyrir trjáhringi

	Estimate	Std. Error
d.f	0.4468892	0.0103489
Fitted mean	43.9953123	9.1733309

- Úrtaksmeðaltalið er 42.3 og úrtaksstaðalfrávikði er 11.8. Ef gengið væri út frá random-úrtaki væri standard-error-of-the mean  $11.8/\sqrt{5405} = 0.16$  sem er miklu lægra en 9.17 sem er staðalvilla meðaltalsins úr ARFIMA líkaninu.
- Annað R-forrit (`longmemo`) gefur líka  $d = 0.44$ .
- Skoða metna afgangslíði  $\hat{\epsilon}$ . Varíans er 64 en varíans camper 138.

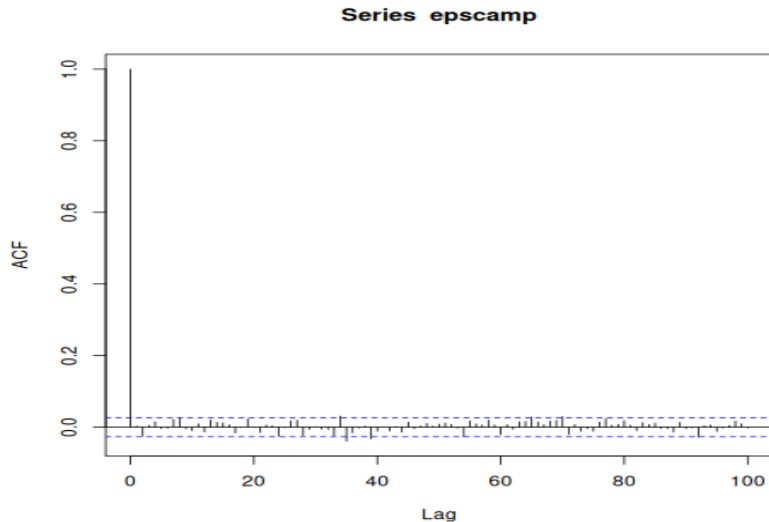


# Metnir afgangslíðir



Metnir afgangslíðir ARIMA(0,0.45,0) líkans fyrir Mount Campito.

# Sjálfylgnifall afgangslíða

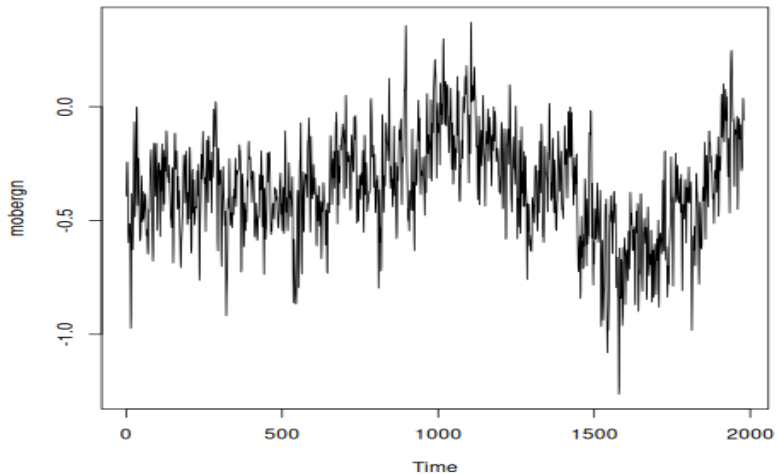


Sjálfylgnifall afgangslíða ARIMA(0,0.45,0) líkans fyrir Mount Campito.

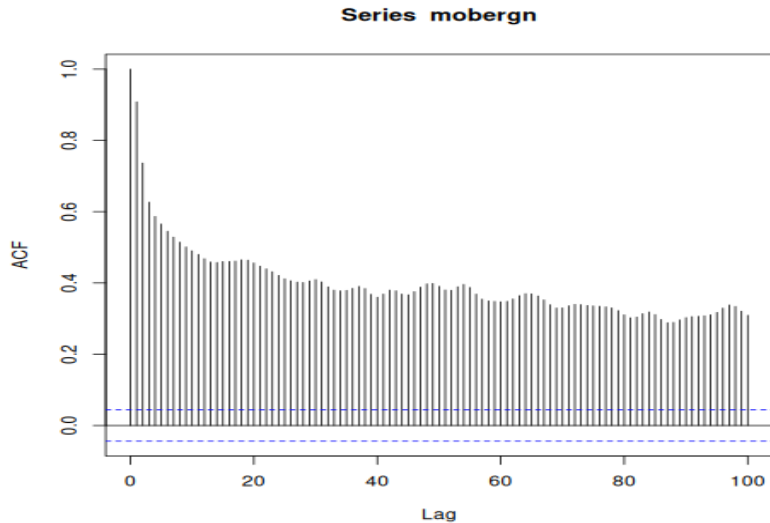
# Tvær aðrar náttúruvísindaraðir

- Moberg, Sonechkin, Holmgren, Datsenko & Karlén (Moberg et al.)  
Áætla lofthita jarðar ca. 2000 ár aftur í tímann byggt á ýmsum gögnum.
- Reynolds, Scourse, Halloran, Nederbragt, Wanamaker, Butler, Richardson, Heinemeier, Eiríksson, Knudsen & Hall (2016) skoða „árhringi“ í skeljum og reyna að giska á sjávarhita 1000 ár aftur í tímann.

# Moberg röð

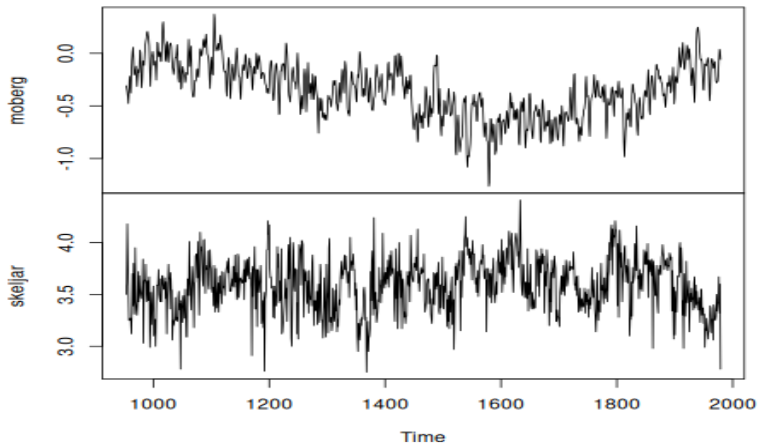


# Sjálffylgnifall Moberg raðar

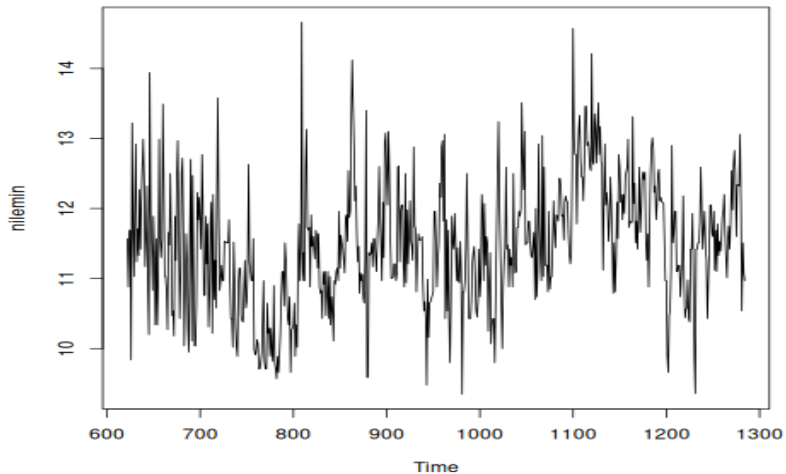


# Eru raðirnar sammála?

Skeljar og Moberg

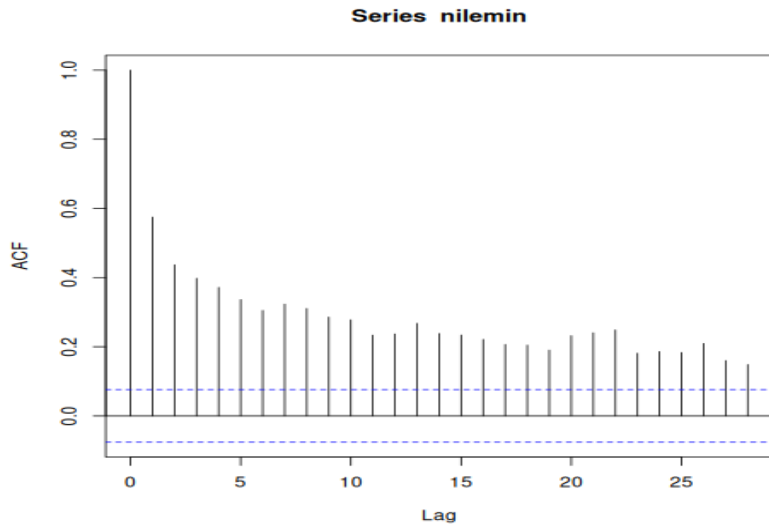


# Hvað sá Hurst (1951)?

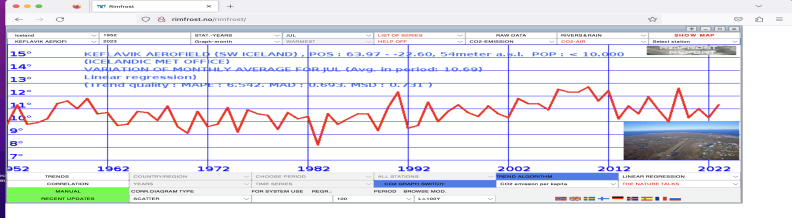


Lágmarksvatnsstaða í Níl.

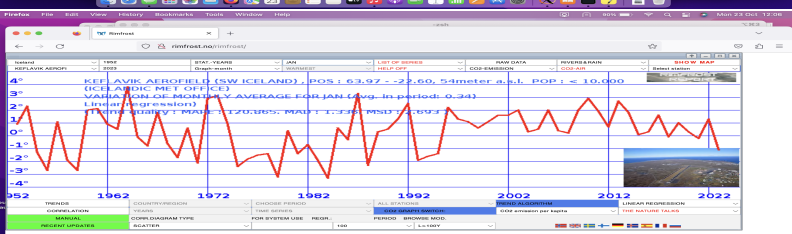
# Sjálffylgni í vatnsstöðu Nílar







Rimfrost is running in your browser thanks to [Webswina!](https://www.webswina.com/)



Rimfrost is running in your browser thanks to [Webswina!](https://www.webswina.com/)

Meðalmánaðarhiti í janúar og júlí á Keflavíkurflugvelli (rimfrost.no).

# Ályktanir Dagsvik, Fortuna & Moen (2020)

- Skoðaðar ca. 100 raðir af hitmælingum í 200 ár (til ca. 2010) ásamt Moberg-röð.
- Ályktun: Þróun hita föst miðað við long-memory líkan.
- Moen heldur úti síðunni rimfrost.no, Fortuna býr til R-pakka og gæðaleiðréttan gagnagrunn.
- Viðbót 2023, gögn til 2018, sama ályktun.
- Gæðaleiðrétt gögn ekki komin frá rimfrost.no í R-gagnagrunninn(tempFGN).
- ARFIMA tegund líkana notuð á breytileika ávöxtunar á fjármálamörkuðum.

- Baillie, R. & Chung, S.-K. (2002). Modeling and forecasting from trend-stationary long memory models with applications to climatology. *International Journal of Forecasting*, 18, 215–226.
- Beran, J. (1994). *Statistics for Long-Memory Processes*. Chapman & Hall.
- Box, G. E. P. & Jenkins, G. M. (1970). *Time Series Analysis, Forecasting and Control*. Holden Day, San Fransisco.
- Chan, N. H. & Palma, W. (1998). State space modeling of long-memory processes. *The Annals of Statistics*, 26(2), 719–740.
- Comte, F. & Renault, E. (1996). Long memory continuous time models. *Journal of Econometrics*, 73(1), 101–149.
- Dagsvik, J. K., Fortuna, M., & Moen, S. H. (2020). How does temperature vary over time?: evidence on the stationary and fractal nature of temperature fluctuations. *Journal of the Royal Statistical Society Series A*, 183(3), 883–908.
- Granger, C. (1980). Long memory relationships and the aggregation of dynamic models. *Journal of Econometrics*, 14, 227–238.
- Hosking, J. R. M. (1981). Fractional differencing. *Biometrika*, 68(1), 165–176.

- Hurst, H. E. (1951). Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116, 770–799.
- Johansen, S. (2008). A representation theory for a class of vector autoregressive models for fractional processes. *Econometric Theory*, 24(3), 651–676.
- Maddanu, F. (2023). Forecasting highly persistent time series with bounded spectrum processes. *Statistical Papers*, (64), 285–319.
- Moberg, A., Sonechkin, D. M., Holmgren, K., Datsenko, N. M., & Karlén, W. Highly variable northern hemisphere temperatures reconstructed from low- and high-resolution proxy data. *433(7026)*, 613–617.
- Reynolds, D. J., Scourse, J. D., Halloran, P. R., Nederbragt, A. J., Wanamaker, A. D., Butler, P. G., Richardson, C. A., Heinemeier, J., Eiríksson, J., Knudsen, K. L., & Hall, I. R. (2016). Annually resolved north atlantic marine climate over the last millennium. *Nature Communications*, 7(1), 13502.