

A Lee-Carter módszer magyarországi alkalmazása

Baran Sándor,
Gáll József,
Ispány Márton,
Pap Gyula

Alkalmazott Matematika és Valószínűségszámítás Tanszék,
Debreceni Egyetem, Informatikai Kar

Feladatok: hazai mortalitási adatokra a Lee-Carter modell illesztése

előrejelzés,

Lee-Carter variánsok vizsgálata,

CI, (becslési) hibák

Jelölések:

$m_{x,t}$: halálozási ráta (Demográfiai Évkönyv)

x : életkor (1, ..., 100)

t : év (1949-2003)

Lee-Carter módszer:

Lee, R. D., and Carter, L. R. 1992. *"Modelling and forecasting the time series of U.S. mortality"*, Journal of the American Statistical Association 87, no. 419 (September): 659-671.

Lee, R. D. 2000. *"The Lee-Carter method for forecasting mortality, with various extensions and applications"*, North American Actuarial Journal 4, no. 1: 80-91.

Booth, H., Maindonald, J., and Smith, L. 2002. *"Age-time interactions in mortality projection: applying Lee-Carter to Australia"*, Working Paper, The Australian National University

Renshaw, A. E., and Haberman, S. 2003. *"Lee-Carter mortality forecasting with age-specific enhancement"*, Insurance: Mathematics and Economic 33: 255-272.

Lee-Carter modell:

$$\ln(m_{x,t}) = a_x + b_x k_t + \varepsilon_{k,t}$$

a_x : életkori „fő” összetevő

k_t : halálozási szint a t évben,

b_x : érzékenység az x életkorban

$$\frac{\Delta \ln(m_{x,t})}{\Delta t} = b_x \frac{\Delta k_t}{\Delta t},$$

$\varepsilon_{x,t} \sim \mathcal{N}(0, \sigma_\varepsilon)$: hiba

Nem egyértelmű paraméterek:

$$a_x + b_x k_t = (a_x - c b_x) + b_x (k_t + c),$$

$$a_x + b_x k_t = a_x + (c b_x) (k_t / c),$$

Feltételek:

$$\sum_{x=1}^N b_x = 1, \quad \sum_{t=1}^T k_t = 0$$

Paraméterek becslése:

$$\hat{a}_x = \frac{1}{T} \sum_{t=1}^T \ln(m_{x,t})$$

$$M_{x,t} := \ln(m_{x,t}) - \hat{a}_x$$

$$M = (M_{x,t})$$

\hat{b}_x és \hat{k}_t : SVD (szinguláris felbontás) az M -re, azaz: $M = UDV$,

$$\hat{b}_x = \frac{1}{c} U_{x,1}, \quad \hat{k}_t = c D_{1,1} V_{1,t},$$

ahol $c = \sum_{x=1}^N U_{x,1}$

Illesztés és előrejelzés:

\hat{k}_t idősorra illesztés (ARIMA), $\Rightarrow \hat{\hat{k}}_t$, majd mortalitás előrejelzés:

$$\hat{m}_{x,t} = \exp(\hat{a}_x + \hat{b}_x \hat{\hat{k}}_t)$$

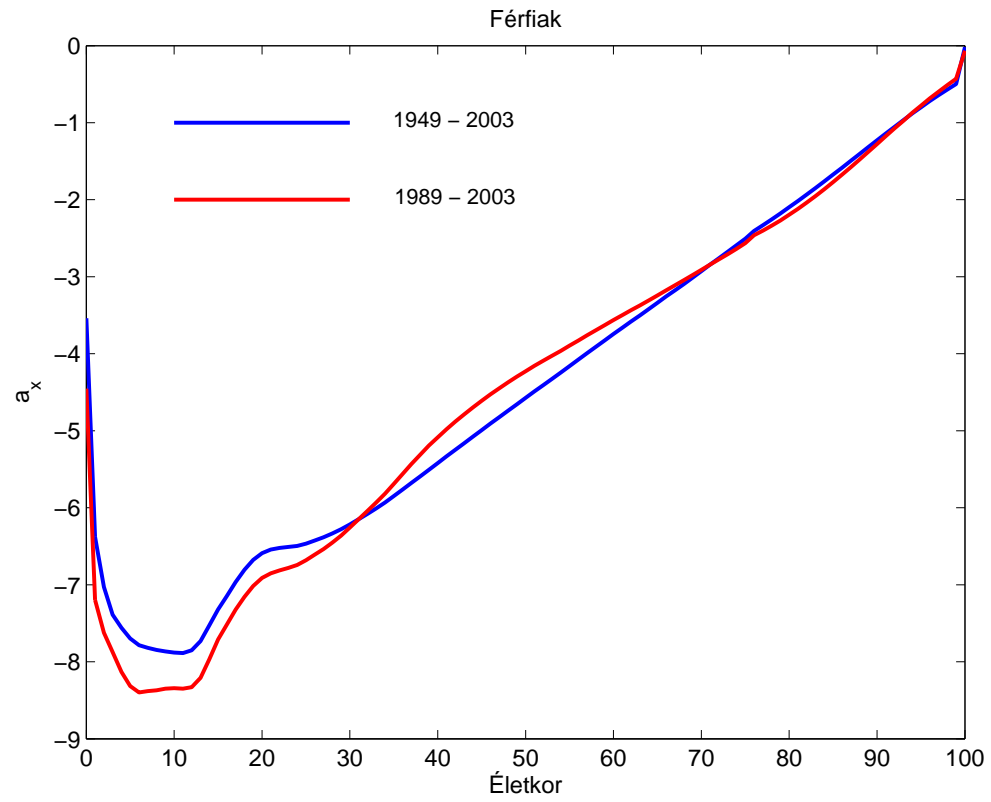
Megjegyzés: $\sum_{t=1}^T \hat{k}_t = 0$,

\hat{k}_t tipikusan csökkenő,

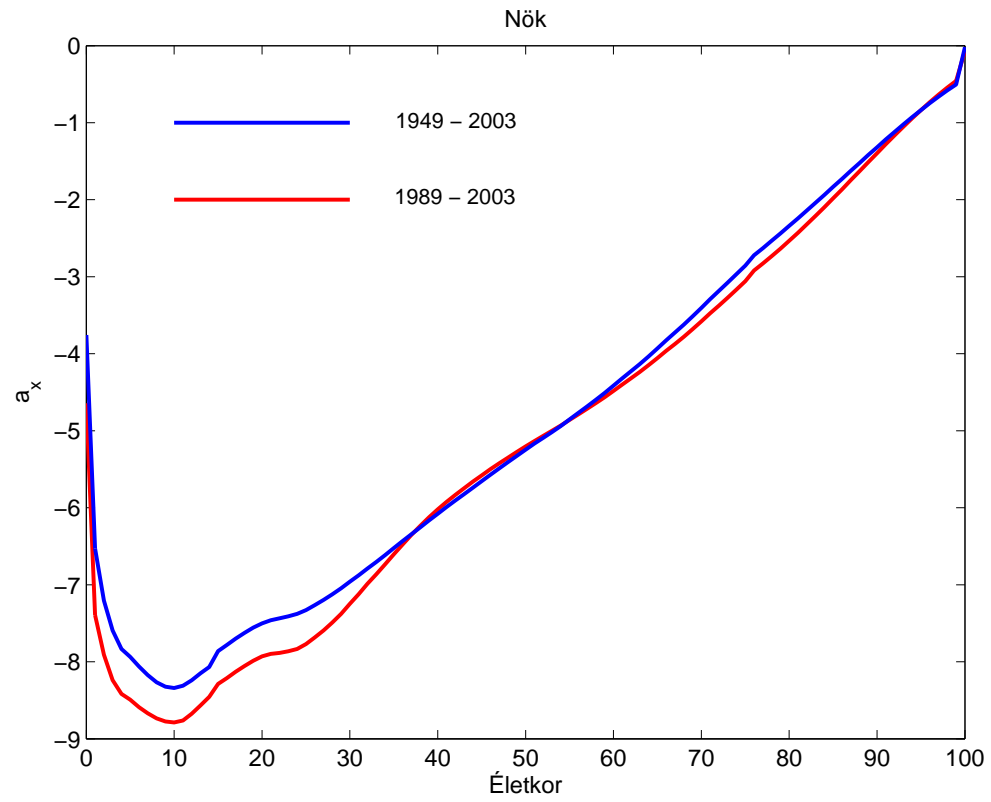
ha $\hat{b}_x < 0$ akkor a ráta növekvő!

Két illesztés: 55 év (1949–2003) és 15 év (1989–2003) alapján

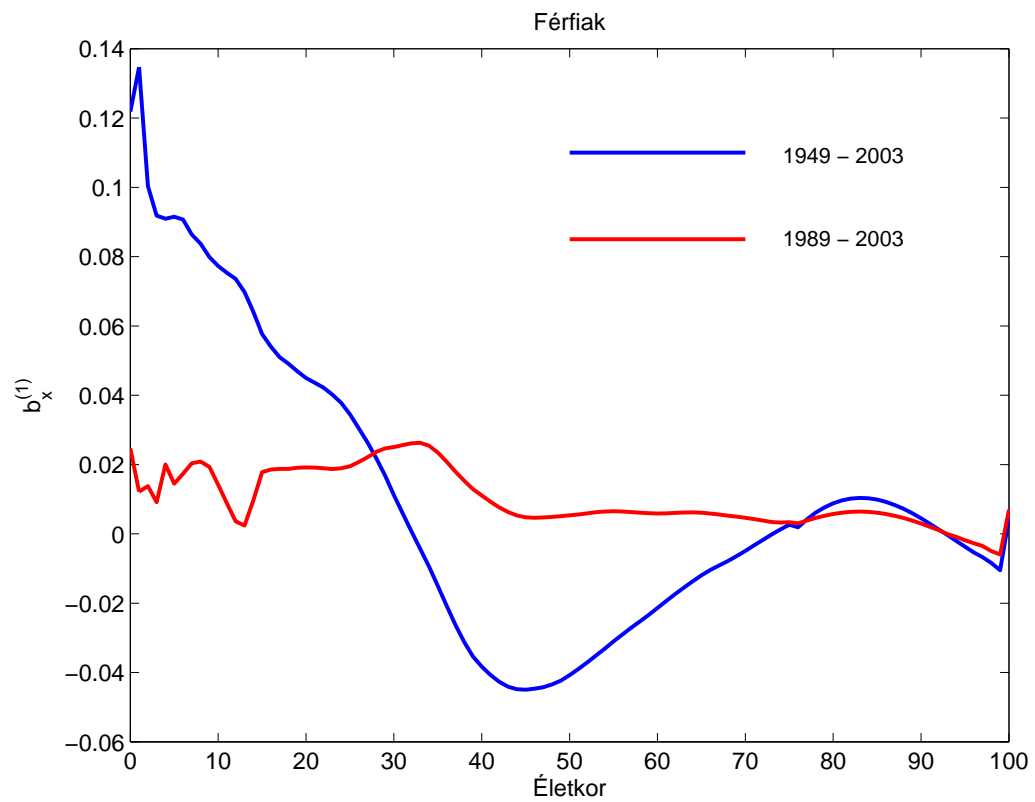
\hat{a}_x , férfiak



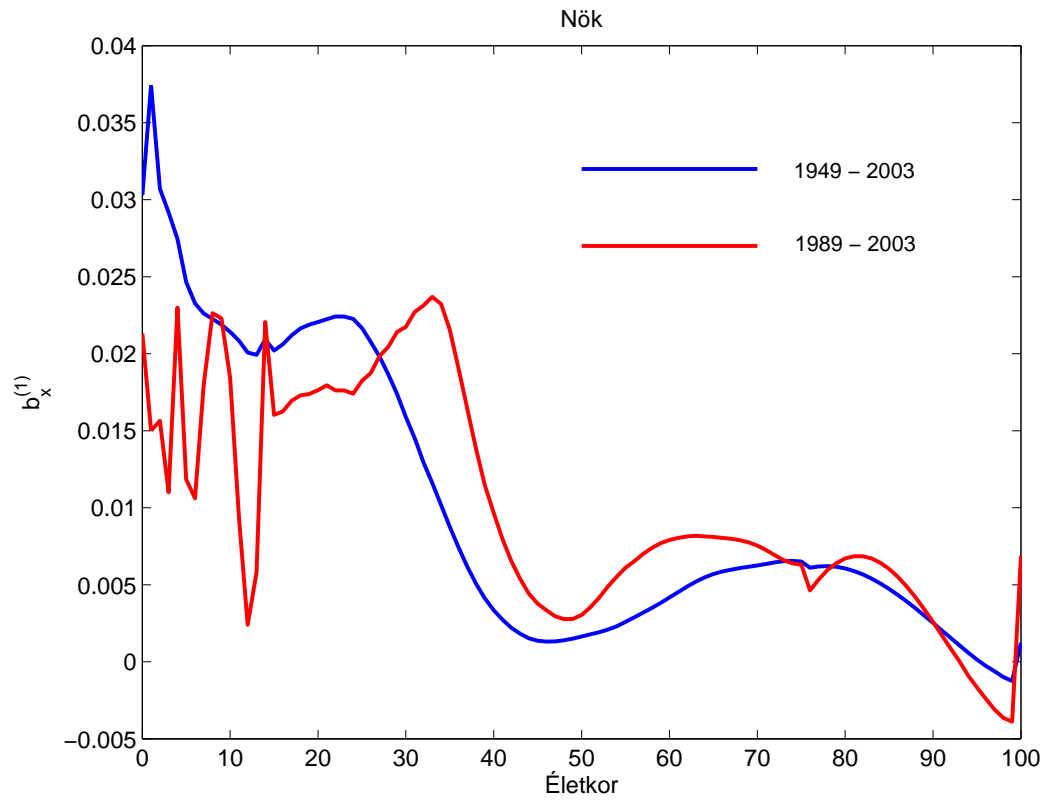
\hat{a}_x , nők



\hat{b}_x , férfiak



\hat{b}_x , nők



További szinguláris értékek használata: $M = UDV$

$$b_x^{(1)}, b_x^{(2)}, b_x^{(3)}, \dots$$

$$k_t^{(1)}, k_t^{(2)}, k_t^{(3)}, \dots$$

becsléseik:

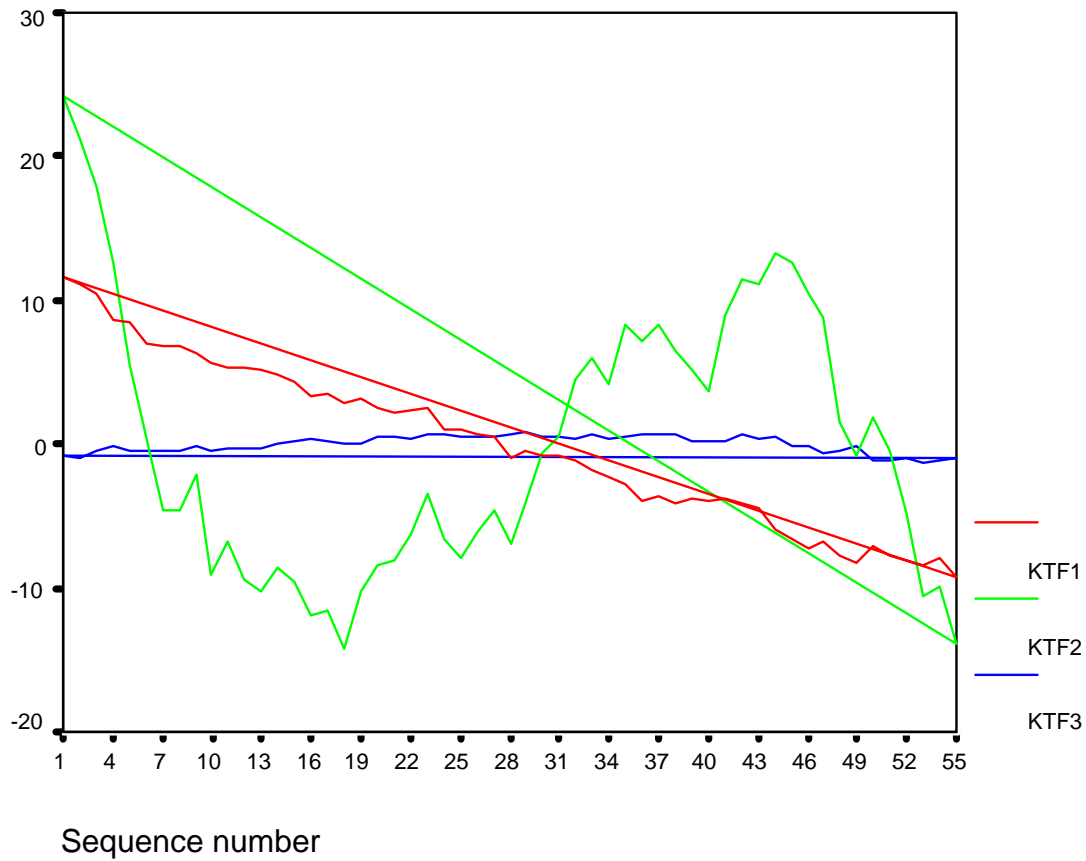
$$\hat{b}_x^{(i)} = \frac{1}{c_i} U_{x,i}, \quad \hat{k}_t^{(i)} = c_i D_{i,i} V_{i,t},$$

ahol $c_i = \sum_{x=1}^N U_{x,i}$

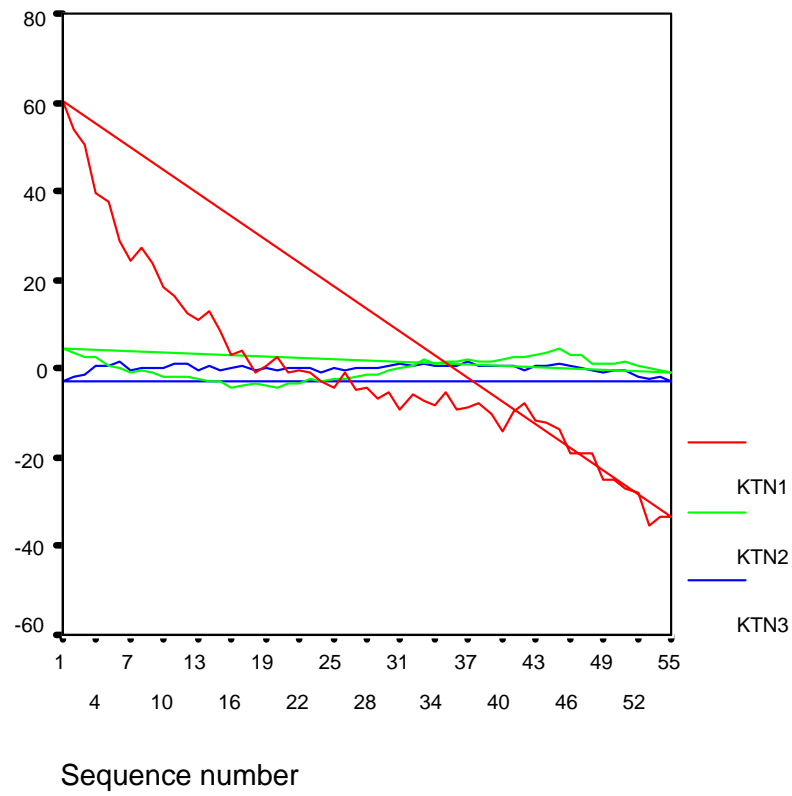
előrejelzés:

$$\hat{m}_{x,t} = \exp \left(\hat{a}_x + \sum_i \hat{b}_x^{(i)} \hat{k}_t^{(i)} \right)$$

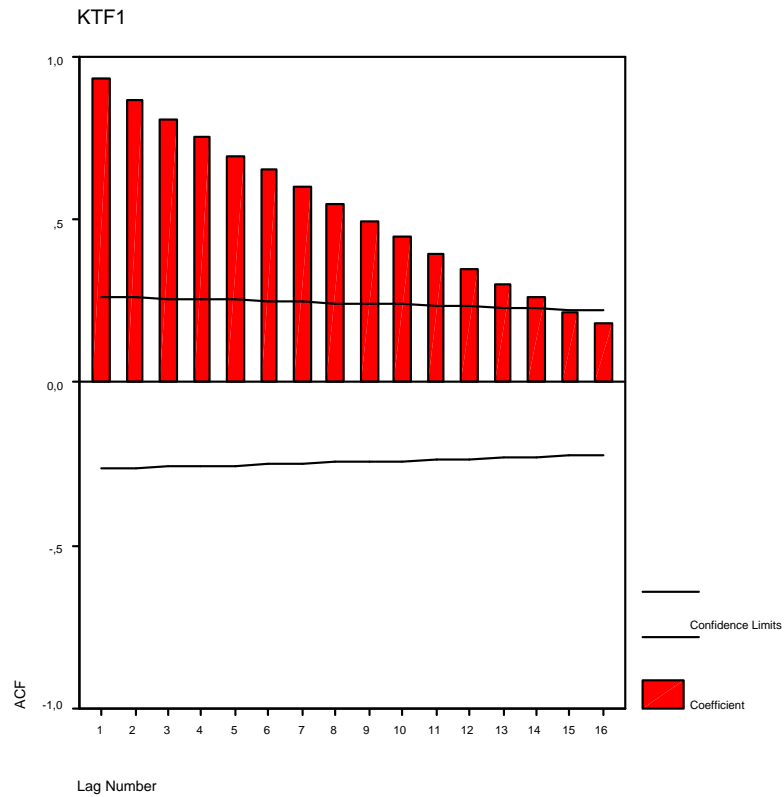
Halálzási szintek (indexek): férfiak ($k_t^{(1)}$, $k_t^{(2)}$, $k_t^{(3)}$)



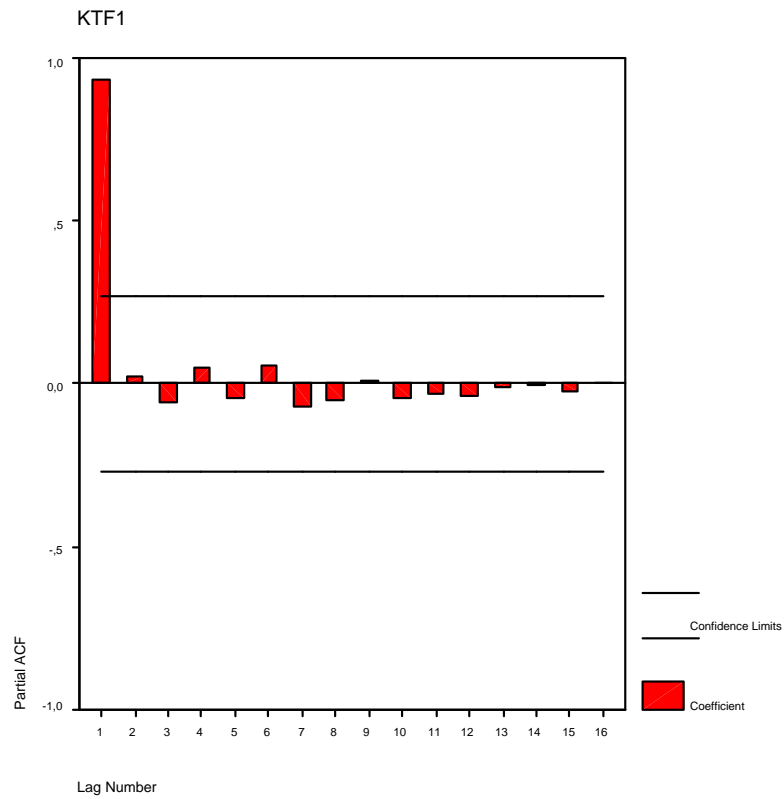
Halálózási szintek (indexek): nők ($k_t^{(1)}$, $k_t^{(2)}$, $k_t^{(3)}$)



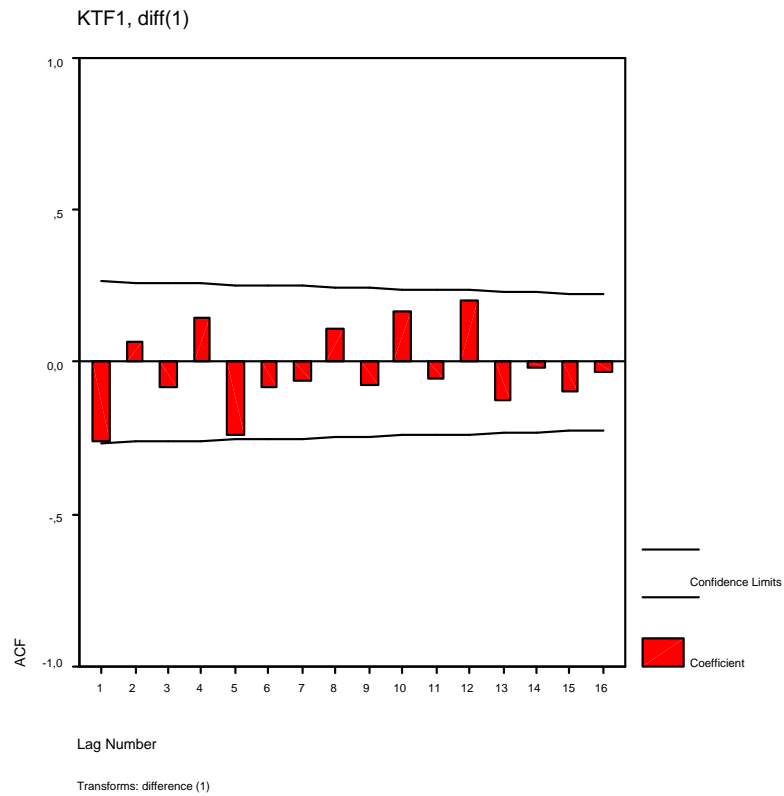
Idősorok illesztése a halálzási szintekre,
példa: $k_t^{(1)}$ esete, autokorreláció:



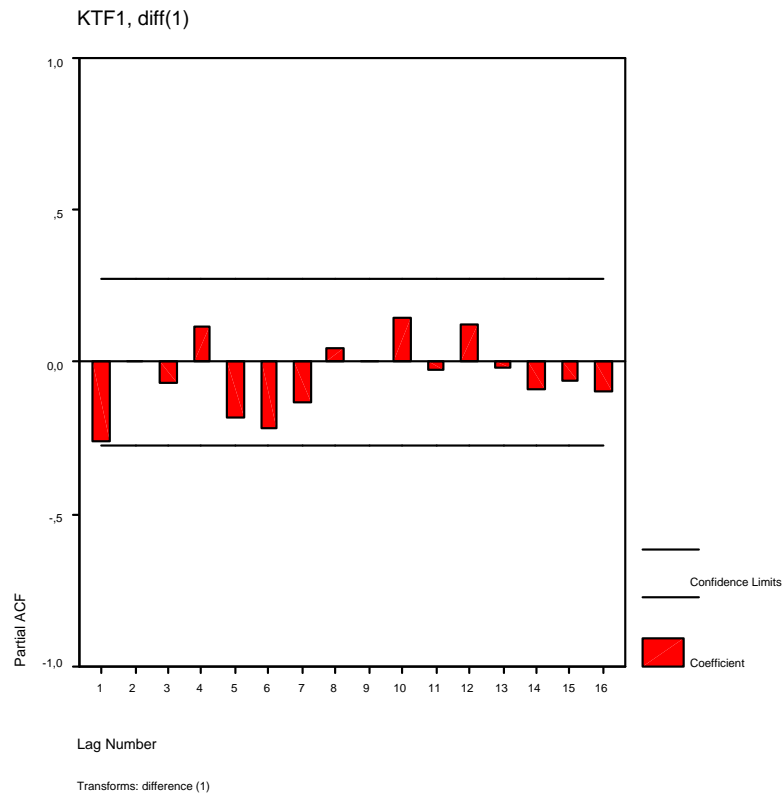
$k_t^{(1)}$ esete, parciális autokorreláció:



$k_t^{(1)}$ elsőrendű differenciái, autokorreláció:



$k_t^{(1)}$ elsőrendű differenciái, parciális autokorreláció:



$k_t^{(1)}$ -re illesztett néhány (legjobb) modell összefoglalása:

ARIMA	AIC	SBC	P (konstans)	P (paraméter)
(0,1,0)	97,63	99,62	,000015	–
(1,0,0)	124,45	126,45	–	,000
(0,1,1)	95,91	99,89	,000	,0617
(1,1,0)	95,67	99,65	,000	,0529

Az elfogadott modellek:

55 év esetén

$\hat{k}^{(1)}$: ARIMA(0,1,0), negatív konstanstag,

$\hat{k}^{(2)}$: az elsőrendű diff. fehér zaj,

$\hat{k}^{(3)}$: ARIMA(0,1,0) (férfiak), ARIMA(1,0,0) (nők)

15 év esetén

$\hat{k}^{(1)}$: ARIMA(0,1,0), negatív konstanstag,

$\hat{k}^{(2)}$: ARIMA(1,0,0) (férfiak), fehér zaj (nők),

$\hat{k}^{(3)}$: fehér zaj,

További kérdések, megjegyzések:

intervallum becslés,

hibák:

$$\ln(m_{x,t+s}) = \hat{a}_x + \alpha_x + (\hat{b}_x + \beta_x)(\hat{k}_{t+s} + u_{t+s}) + \varepsilon_{x,t+s}$$

β_x szórásbecslése: bootstrap, kevés adat, M véletlenítése, majd

SVD

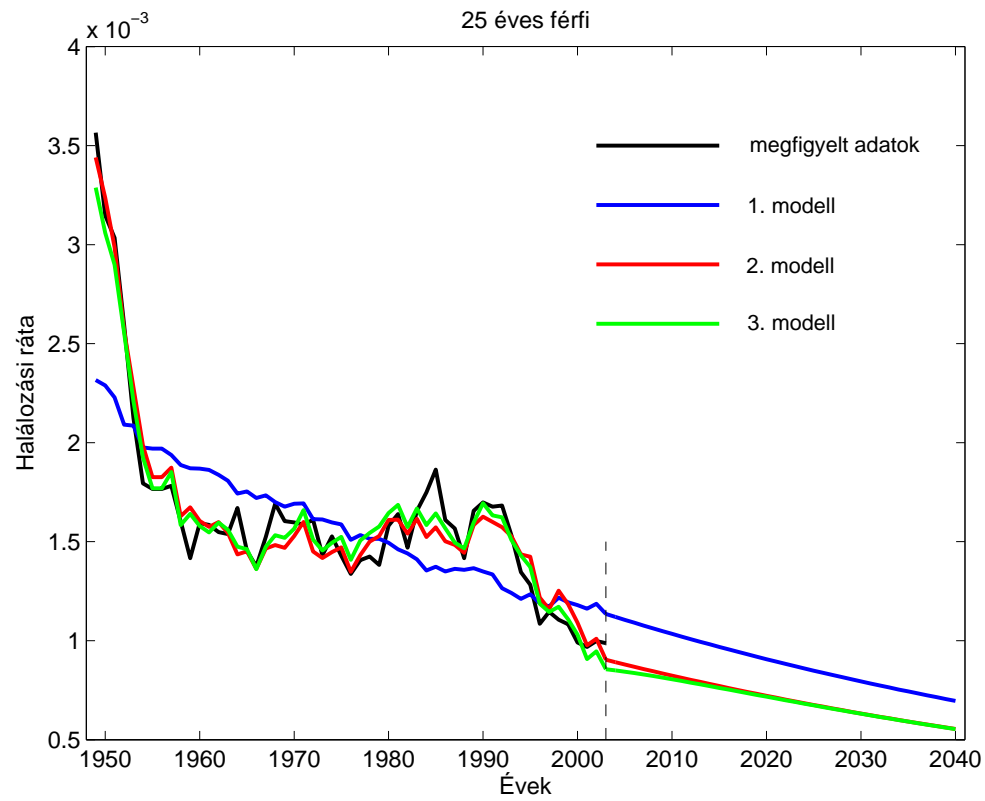
kicsi mintaméret,

nagy szórások!!!

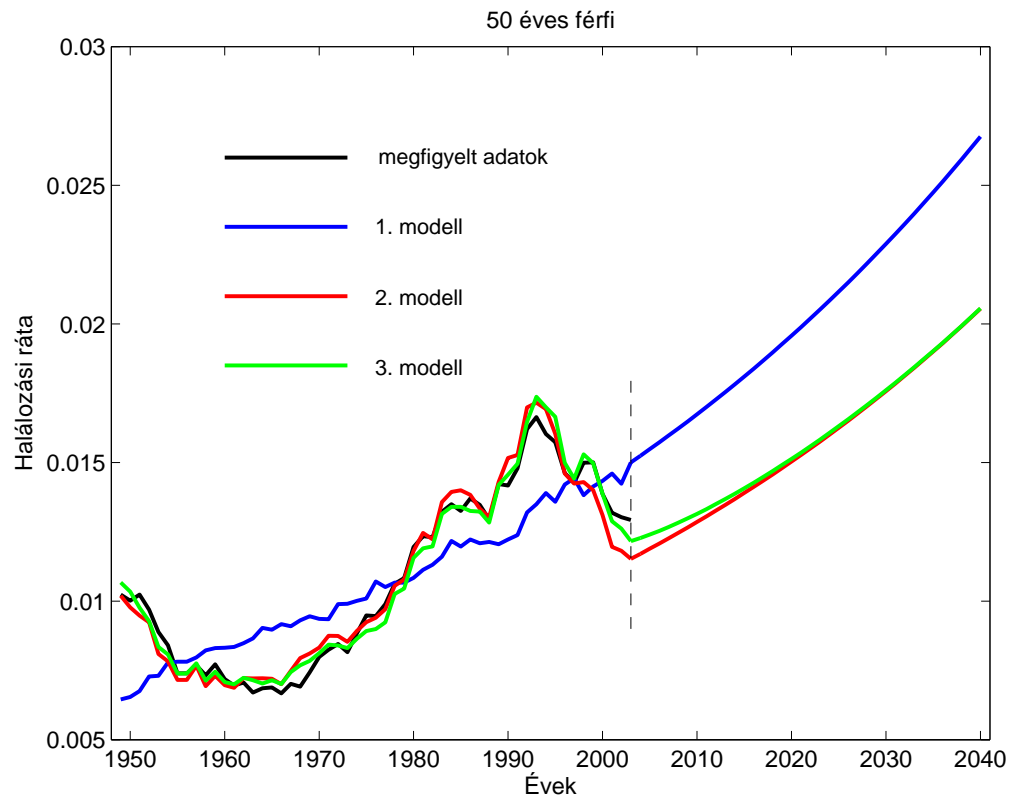
Előrejelzések:

1. modell: az eredeti Lee-Carter (1 szinguláris érték használata),
2. modell: 2 szinguláris érték használata,
3. modell: 3 szinguláris érték használata

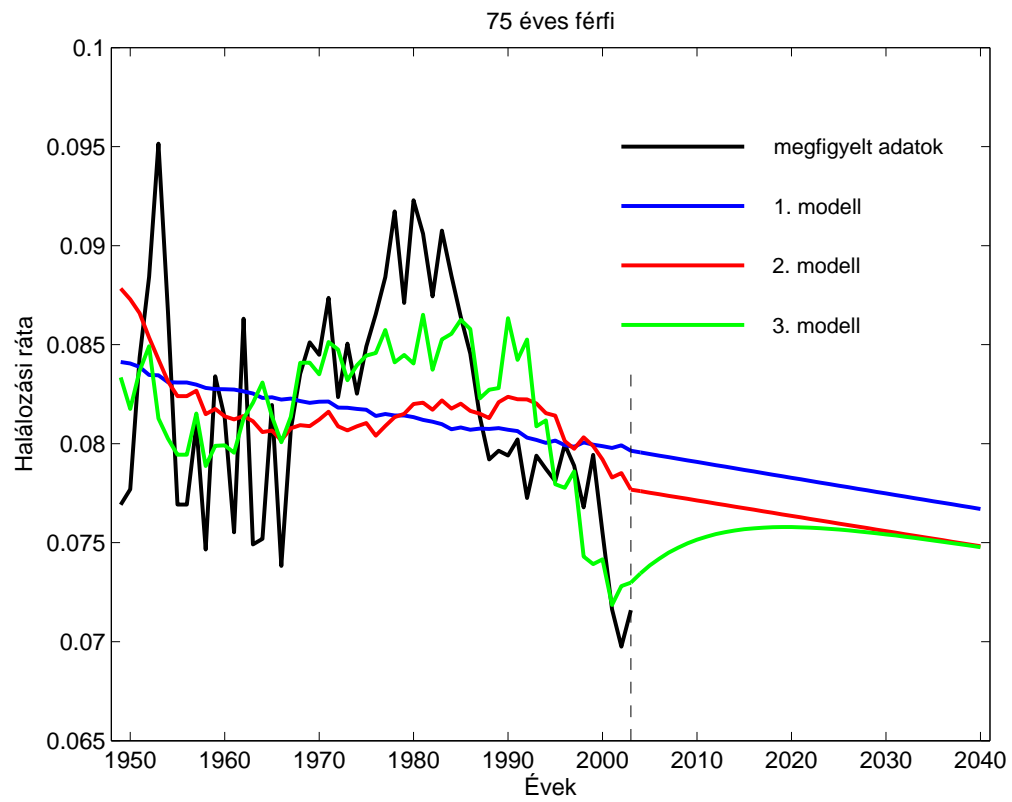
55 (minta)év esetén, 25 éves férfi:



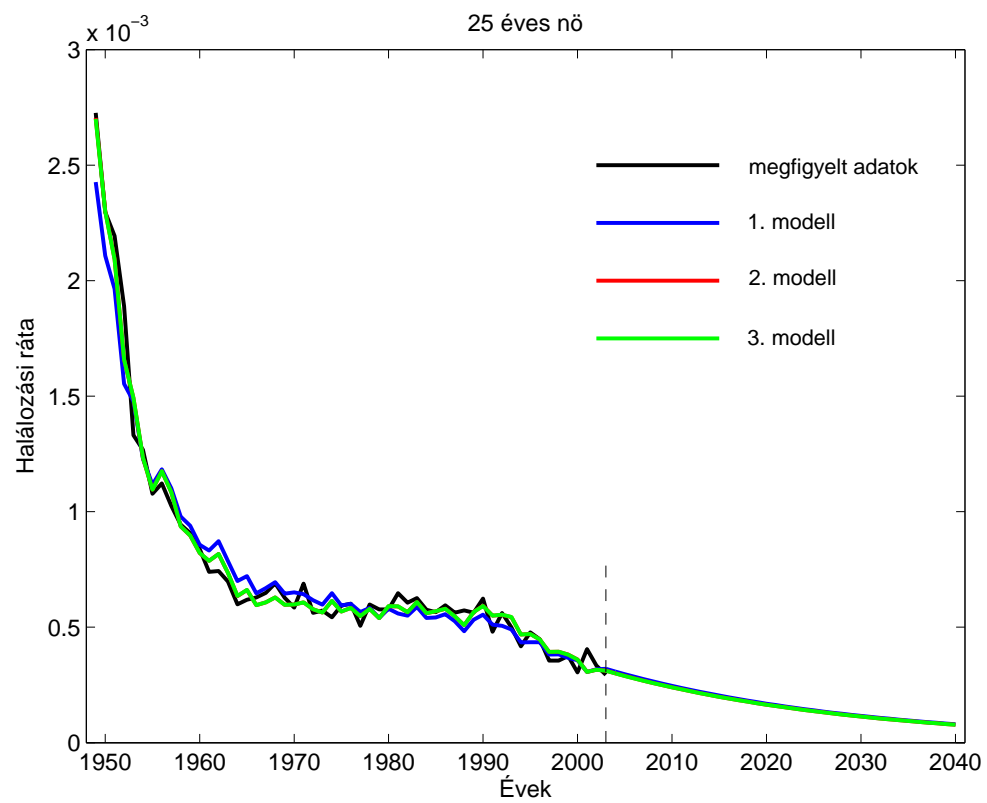
55 (minta)év esetén, 50 éves férfi:



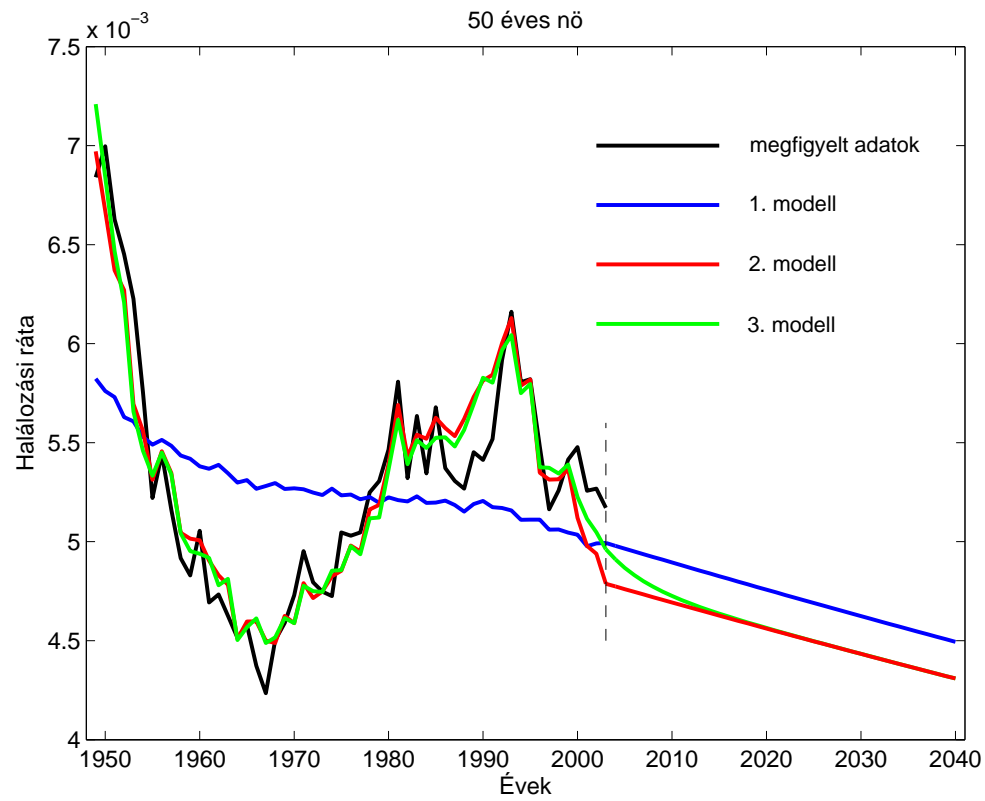
55 (minta)év esetén, 75 éves férfi:



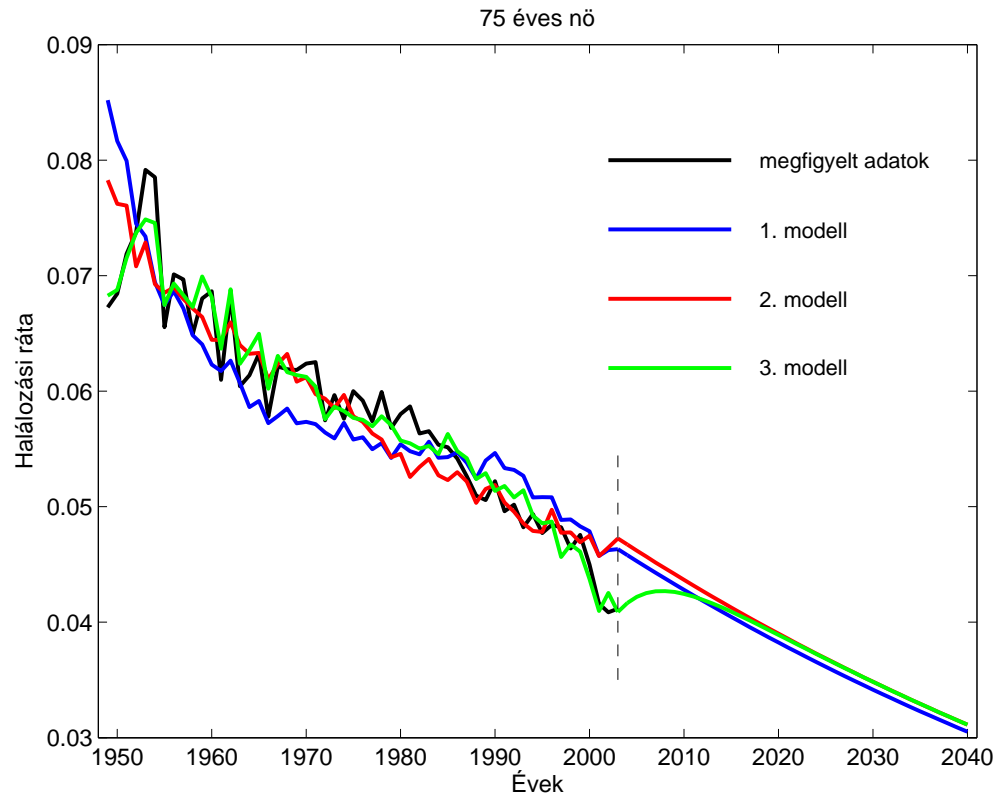
55 (minta)év esetén, 25 éves nő:



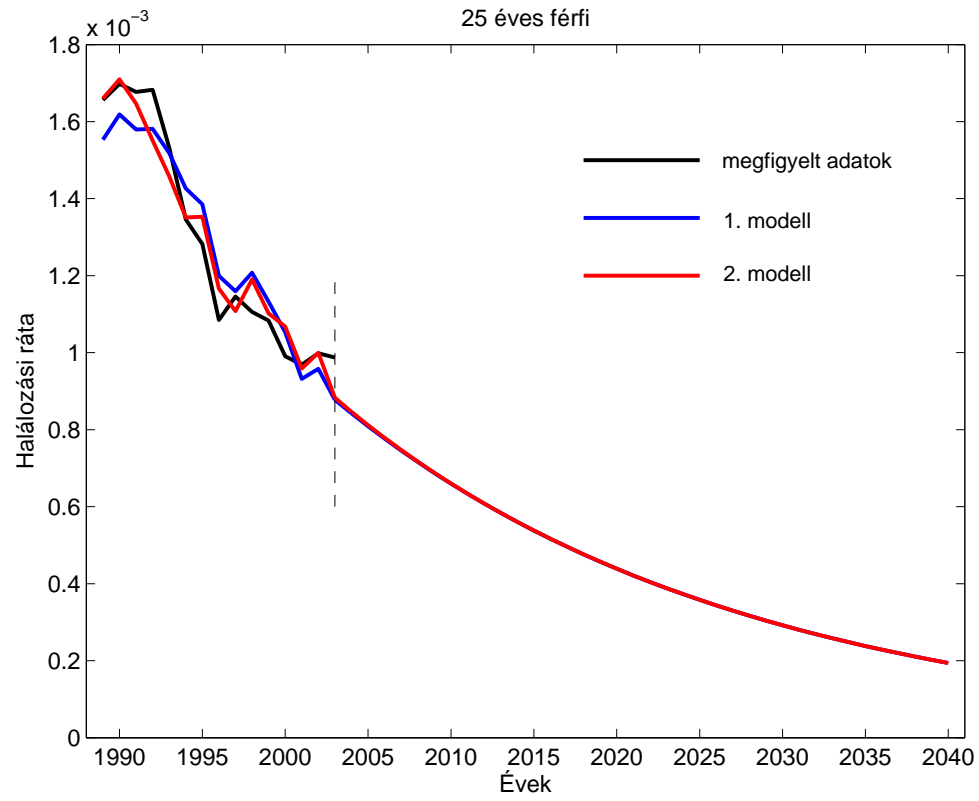
55 (minta)év esetén, 50 éves nő:



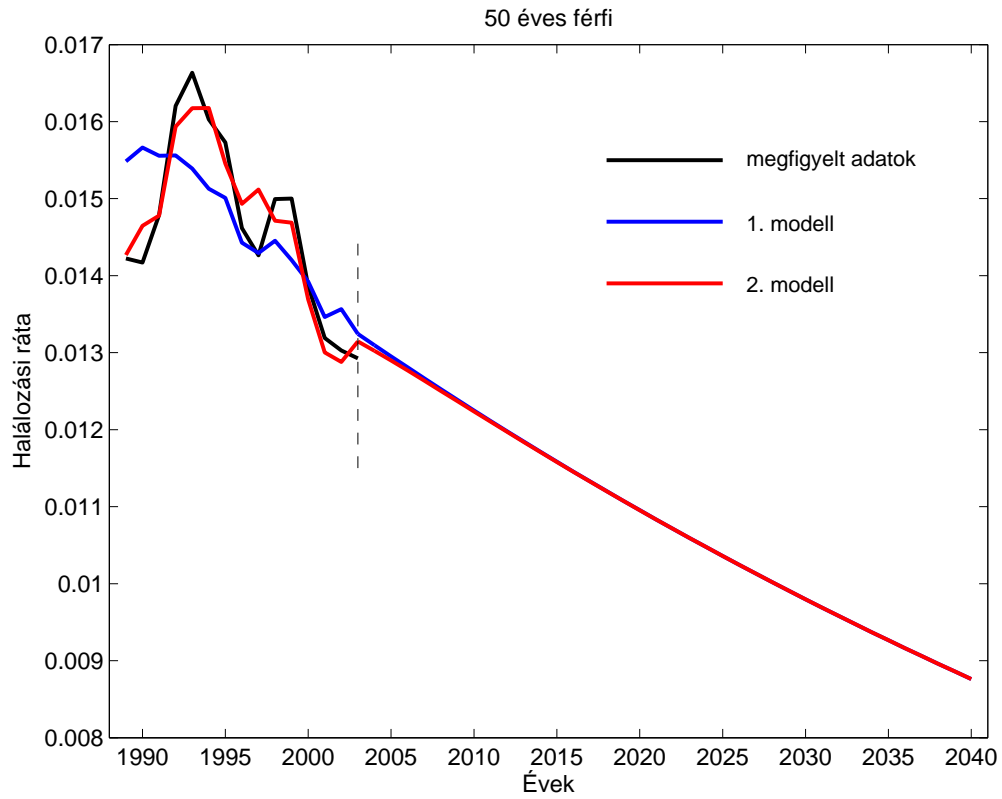
55 (minta)év esetén, 75 éves nő:



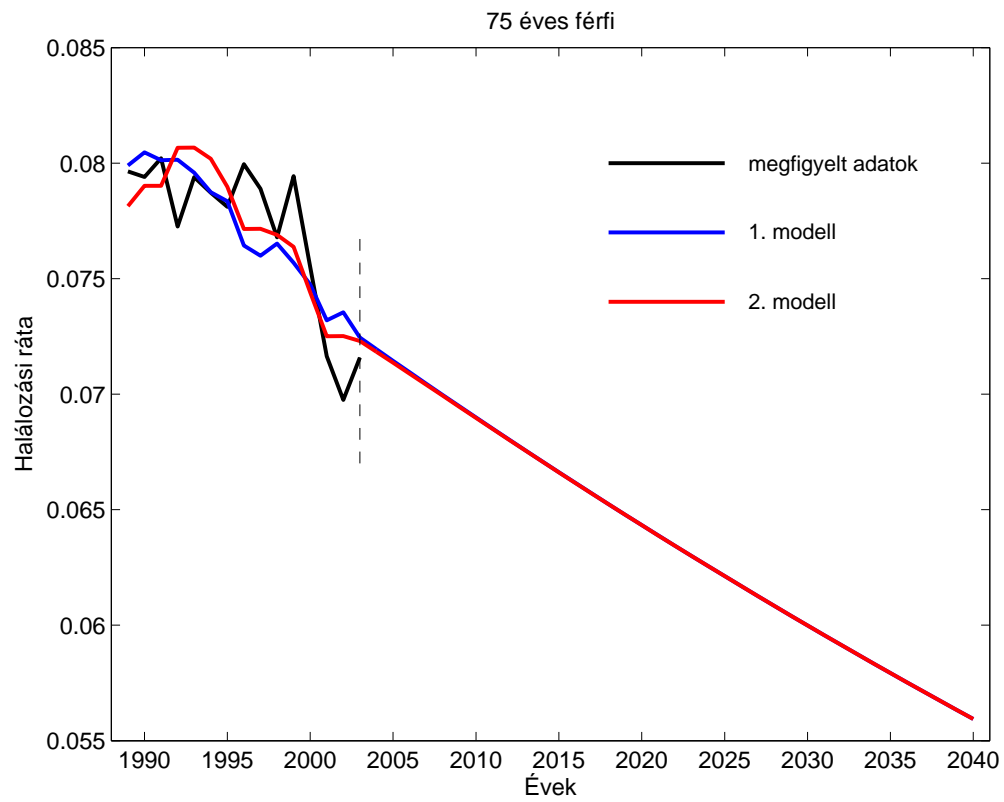
15 (minta)év esetén, 25 éves férfi:



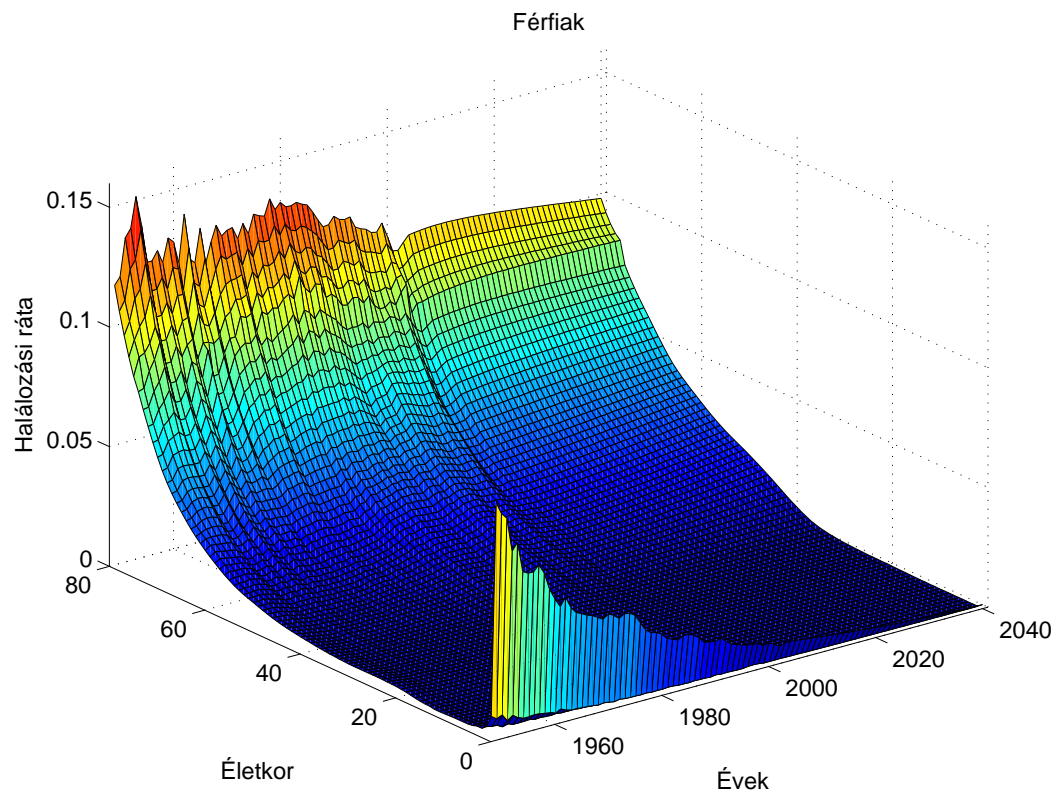
15 (minta)év esetén, 50 éves férfi:



15 (minta)év esetén, 75 éves férfi:



55 év esetén, minden életkorra



15 év esetén, minden életkorra

