



PÉCSI TUDOMÁNYEGYETEM
KÖZGAZDASÁGTUDOMÁNYI KAR



Etikus statisztikai modellezés

KSH és Gazdaságmodellezési Társaság
„Statisztika és modellezés” tematikus ülés

2012. február 16.



Statisztikai modell általános formája

Tényleges jelenség

Modellezett érték

Reziduum

$$y_i = \hat{y}_i + \varepsilon_i$$





Statisztikai modellezői „torzítások” tipizálása

„Bűn”

Szándékosan rossz adat (hazugság)

Hibás módszerválasztás

Eredmények fals interpretálása

Jó adat – torz megjelenítés

Etikai „vétség”

Információáradat szisztematikus szelektálása

Bizonytalanság (mérési hiba, szóródás) elhallgatása

Modellezendő jelenség túlzott egyszerűsítése

Túlságosan összetett módszertan („parasztvakítás”)





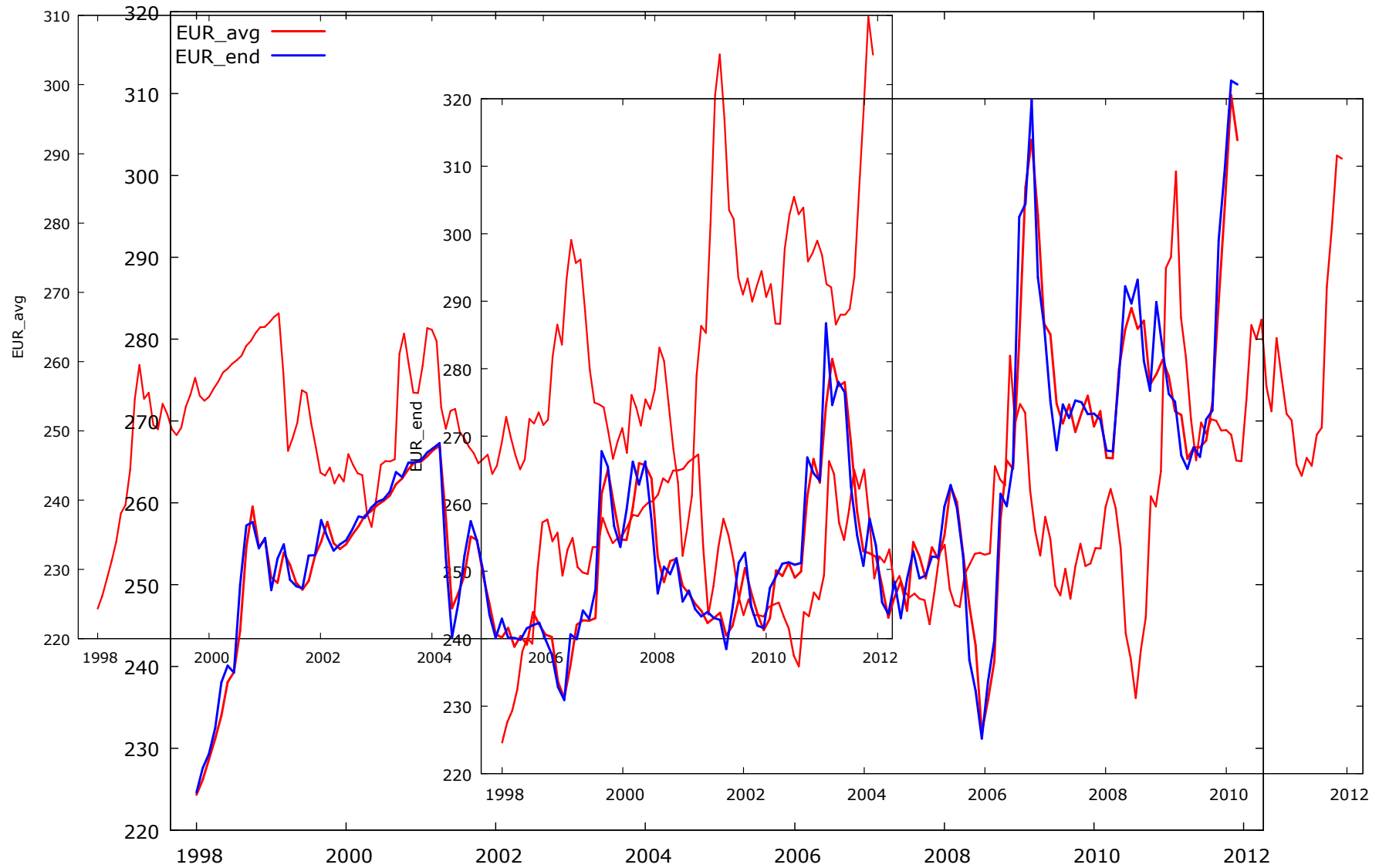
Néhány „tipikus” modellezői vétség

1. Idősorok helytelen aggregálása (folyamatok frekvenciájának inadekvát megválasztása)
2. Első-, illetve másodfajú hiba (hipotézisellenőrzés) veszélyességének helytelen megítélése, esetleg próbák felesleges összekötése (omnibusz-tesztek)
3. Outlier-szűrés (csonkolt mintán alapuló következtetés) nem elég körültekintő végrehajtása
4. „Mintajavító” technikák (monte carlo, bootstrap), illetve mintán kívüli információ beépítésének (Bayes-i módszertan) átgondolatlan alkalmazása
5. Túlságosan gyakori adat-, illetve modell-korrekció, előrejelzések feleslegesen sűrű újraszámítása





€-árfolyam (havi bontású, 1998. január - 2011. december)





Stacionárius volt-e az €-árfolyam?

- Kiterjesztett Dickey-Fuller próba (ADF)
H₀: egységgyököt tartalmaz (nem stacioner)
H₁: stacioner

$$y_t = \mu + \rho y_{t-1} + \beta t + \varepsilon_t$$
$$\Delta y_t = \mu + (\rho - 1) y_{t-1} + \beta t + \gamma_1 \Delta y_{t-1} + \dots + \varepsilon_t$$
$$H_0 : \rho - 1 = 0$$
$$H_0 : \rho - 1 < 0$$

Idősor	Paraméter	Próbafüggvény	ass. p-érték
eur_avg	-0,1272	-4,152	0,0052
eur_end	-0,1419	-2,923	0,1549





Első- vagy másodfajú hibától féljek-e?

Az ismeretlen valóságban	A döntés során a nullhipotézist	
	Elfogadjuk	Elvetjük
H_0 igaz	<i>Helyes döntés ($1-\alpha$)</i>	<i>I. fajú hiba (α)</i>
H_0 hamis	<i>II. fajú hiba (β)</i>	<i>Helyes döntés ($1-\beta$)</i>





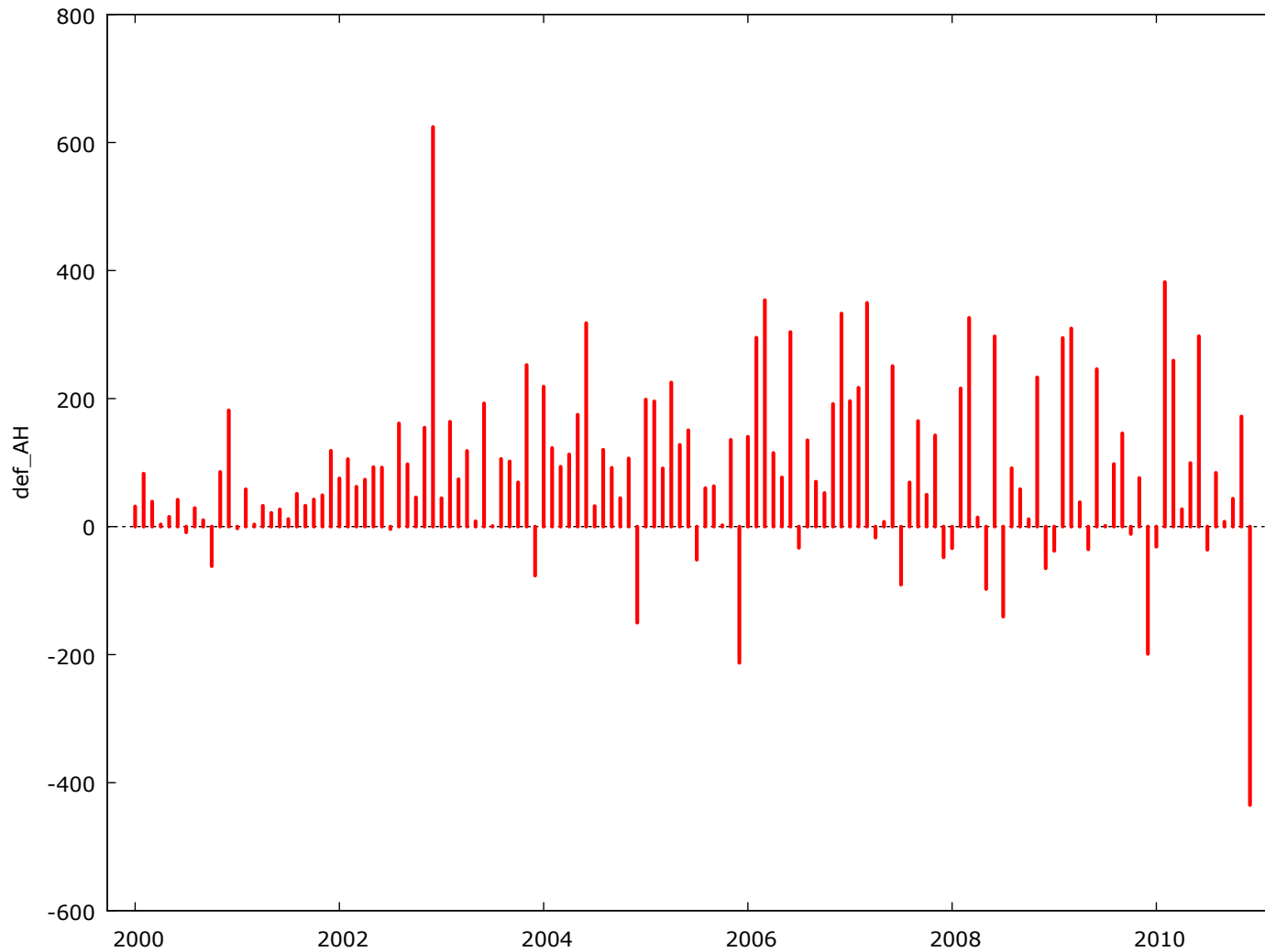
Döntési szituációk stacionaritás-tesztekben

Valóság (nullhipotézisében ezt feltételező teszt)	Döntésünk szerint	
	Stacioner	Nem stacioner
Stacioner (KPSS)	<i>Helyes döntés</i>	P_{KPSS}
Nem stacioner (ADF)	P_{ADF}	<i>Helyes döntés</i>





Államháztartás hiánya Magyarországon (deficit szemlélet, Mrd Ft 2000-2010, havi bontás)



$p_{ADF} = 0,005$
 $p_{KPSS} = 0,027$





Önmérséklet szükségessége

- Adatállomány bővítésének átgondolása
- Informatikai támogatás igénybevétele
- „Optimális” modell nem pusztán technikai értelmezése
- Feltalálni, vagy felfedezni dilemmája
- A hír információértékének és valóságtartalmának paradoxonja

