MATEMATICKO-FYZIKÁLNÍ FAKULTA
Univerzita Karlova

Igor Böhm

Detekce a přisuzování změn klimatu pomocí ekonometrických metod

Francisco Estrada, Pierre Perron

15. května 2024

Úvod

- Článek je kompilací různých metod pro modelování globální teploty s důrazem na jejich historický vývoj.
- Cílem je najít metody pro modelování globální teploty, které by dobře detekovaly změny v teplotách a případně mohly odpovědět na otázku, zda a kolik se na změně teploty podílí člověk.
- Zde si uvedeme hlavní metody a myšlenky použité v článku.

Procesy pro modelování

- Předpokládáme časovou řadu $y_t = \tau_t + z_t$, kde τ_t je trendová složka a z_t je náhodný šum. Budeme chtít, aby časová řada byla slabě stacionární.
- Proces se nazývá integrovaný stupně d ($I(d)$), pokud je d-tá diference šumu stacionární ($\Delta^d z = (1 - L)^d z_t$, L je zpětný operátor).
- Pro proces ARMA (kde $A(L)z_t = B(L)e_t$, A, B jsou polynomy, $e_t \sim i.i.d.(0, \sigma^2)$) platí, že pokud jsou všechny kořeny polynomu A vně jednotkového kruhu, je tento proces stacionární a kauzální. Pokud pro chybovou složku d-té diference procesu $\{y_t\}$ platí, že má A právě d jednotkových kořenu a ostatní za jednotkovým kruhem, pak je $I(d)$.

Jednotkové kořeny

- předpokládejme model AR(1):

$$z_t = \alpha z_{t-1} + e_t, \; e_t \sim i.i.d(0, \sigma^2).$$

  Příkladem procesu s jednotkovým kořenem je $\alpha = 1$. Zde je $\Delta z_t = e_t$ *i.i.d.* proces. V tomto procesu má každá náhoda e_t permanentní vliv na podobu řady v budoucnosti $z_t = z_0 + \sum_{j=1}^{t} e_j$.
- Toto je proces klasické háhodné procházky.

Diferenčně a trendově stacionární modely

- Nezávislost v náhodné procházce je dost restriktivní. V náhodné složce by se mohla uvažovat korelace mezi blízkými časovými momenty $\Delta z = v_t$, kde v_t je náhodný proces projevující nějakou korelaci.
- Jako příklad může posloužit v_t jako MA(1) model $v_t = e_t + \theta e_{t-1}$. Pak $z_t = z_{t-1} + e_t + \theta e_{t-1}$. Předpokládáme-li jednotkový skok $e_t = 1$, pak jeho vliv v čase $t + 1$ bude zvýšení hodnoty o $(1 + \theta)$.
- Pro v_t obecný stacionární a kauzální ARMA proces platí $A(L)v_t = B(L)e_t$, $e_t \sim i.i.d.(0, \sigma^2)$. Z kauzality jej lze vyjádřit ve tvaru MA(∞): $v_t = A(L)^{-1}B(L)e_t = \psi(L)e_t$. Pak dlouhodobý vliv jednotkového skoku e_t na z_t je dán jako $\psi(1) = \sum_{i=0}^{\infty} \psi_i$.

Náhodný proces s driftem a stochastickým trendem

- Mnohé řady mají tendenci prokazovat linearitu v trendu, budeme uvažovat $\tau_t = \mu + \beta t$. β je směrnice trendu (také drift). Pak $\Delta y_t = \beta + v_t$ pro $A(L)v_t = B(L)e_t$ a $y_t = \mu + \beta t + \sum_{j=1}^{t} v_j$.
- Trend se skládá z deterministické složky, která je dána driftem, a z náhodné složky, která je dána permanentním skokem každého e_t.
- Trend se dá získat tzv. Beveridgeovým-Nelsonovým rozkladem. Předpokládejme MA reprezentaci Δy_t: $\Delta y_t = \beta + \psi(L)e_t$. Rozložme $\psi(L) = \psi(1) + (1 - L)\psi^*(L)$, kde $\psi^*(L) = -\sum_{j=0}^{\infty} \psi_i^* L^i$, $\psi_i^* = \sum_{j=i+1}^{\infty} \psi_j$.
- Pak $y_t = \mu + \beta t + \psi(1)\sum_{j=1}^{t} e_t + w_t$, kde $w_t = \psi^*(L)e_t$ je stacionární náhodný proces.

- Vhodnou definicí trendu je podmíněná střední hodnota $\mathrm{E}\left(y_{t+k}|y_s, s \le t\right)$. Pak dostáváme

$$\mathrm{E}\left(y_{t+k}|y_s, s \le t\right) = \mu + \beta(t+k) + \psi(1)\sum_{j=1}^{t} e_j,$$

  jelikož w_t je stacionární.
- Protože je trend náhodnou fluktuací permanentně ovlivněn, říká se takovým modelům modely se stochastickým trendem (označují se DS).

Trendově stacionární modely

- Alternativou jsou modely, kde je nárůst trendu čistě deterministický, tedy $y_t = c + \beta t + w_t$, kde je $C(L)w_t = D(L)e_t$ pro $e_t \sim i.i.d.(0, \sigma^2)$ se všemi kořeny $D(L)$ a $C(L)$ mimo jednotkový kruh.
- Takovýmto modelům se říká TS modely (*transitory*), neboť náhodná fluktuace e_t je pouze přechodný jev.
- Rozlišení mezi TS a DS procesy se říká problém testování jednotkového kořene.

Trendová funkce se skoky

- Široce rozšířenou třídou trendově stacionárních procesů v literatuře o modelování klimatu jsou jsou modely se změnami tvaru trendové funkce. Obecně parametry trendu a jejich strukturální změny nemusí být deterministické.
- Jedním z takových modelů může být

$$y_t = \mu_t + \beta_t t + z_t,$$

  kde $\mu_t = \mu_{t-1} + v_t$ a $\beta_t = \beta_{t-1} + t_t$.
- Šumová komponenta se dá modelovat jako kombinace dvou normálně rozdělených veličin, kde jedna má vysoký rozptyl a druhá nízký nebo nulový. Toto lze popsat například tímto způsobem: $u_t = \lambda_t \gamma_{1t} + (1 - \lambda_t)\gamma_{2t}$ a $v_t = \kappa_t \delta_{1t} + (1 - \kappa_t)\delta_{2t}$, kde $\gamma_{it} \sim N(0, \sigma^2_{\gamma i})$, $\delta_{it} \sim N(0, \sigma^2_{\delta i})$, přičemž λ_t a κ_t jsou alternativně rozdělené veličiny s pravděpodobnostmi úspěchu α_λ, resp. α_δ.

Frakcionálně integrované procesy

- V případě, kdy pro stacionaritu potřebujeme parametr diference d neceločíselný, využívají se tzv. frakcionálně integrované procesy. Nejčastějším modelem je ARFIMA, což je zobecnění ARIMA modelu pro neceločíselné d. V takovém případě je

$$(1-L)^d = 1 - dL + \frac{d(d-1)}{2!}L^2 + \frac{d(d-1)(d-2)}{3!}L^3 + \cdots =$$
$$= \sum_{k=0}^{\infty} \frac{\Gamma(k-d)L^k}{\Gamma(-d)\Gamma(k+1)}.$$

- Autokorelační funkce takového procesu je

$$\rho_t = \frac{\Gamma(1-d)\Gamma(\tau+d}{\Gamma(d)\Gamma(k+1)} \approx c\tau^{2d-1}$$

pro nějaké c.

- Autokorelační funkce klesá pomalu hyperbolicky pro $0 < d < 0,5$. Pro tento typ procesů je y_t stacionární a kauzální pro $|d| < 0,5$. Pro $0 < d < 0,5$ vykazuje proces dlouhodobou paměť (mnohem delší než I(0) s podobnou ARMA strukturou).
- Pro $0,5 < d < 1$ je proces nestacionární, ale vliv jednotkového náhodného skoku konverguje k 0.

Porovnání modelů

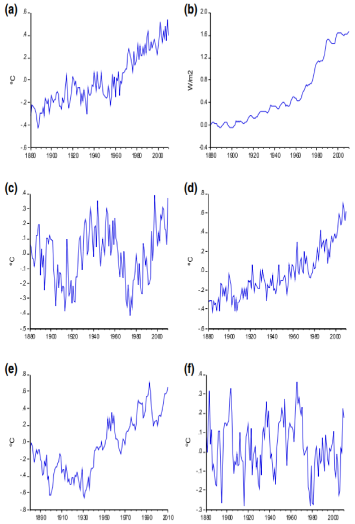

Fig. 1 Global temperatures (filtered; **a**), TRF (**b**), AMO (**c**), Temp_TSB (TS plus break in slope; **d**), Temp_DS (**e**), and Temp_S (**f**)

- **(a)** Graf globální teploty.
- **(b)** Globální teplota s odfiltrovaným radiačním působením (*total radiative forcing*).
- **(c)** Globální teplota s odfiltrovanou nízkofrekvenční oscilací podmínek v severním Atlantiku (*AMO*).
- **(d)** Simulovaný TS proces se skoky v driftu.
- **(e)** Simulovaný DS.
- **(f)** Simulovaný STACIONÁRNÍ PROCES.

- Pokud ke změně dojde jednou nebo pouze v několika případech v dlouhém časovém horizontu, je obtížné modelovat změnu pomocí stochastické struktury kvůli malému počtu pozorovaných změn. Většinou se tyto fenomény vysvětlují přítomností exogenních faktorů a nejsou explicitně modelovány parametrickou stochastickou strukturou.
- Takové situace jsou většinou způsobeny jevy jako zvýšená sluneční aktivita, malé změny v orbitě země nebo zvýšenou produkcí skleníkových plynů.
- V následující části jsou proto uvedeny metody testování jednotkového kořene.

Testování jednotkového kořene

- Nejčastěji používaným testem je test ADF. Předpokládejme AR(p), kde p je známé:

$$y_t = \sum_{i=1}^{p} a_i y_{t-i} + e_t,$$

neboli $A(L)y_t = e_t$, kde $A(L) = 1 - a_1 L - \ldots - a_p L^p$. Obsahuje-li proces jednotkový kořen, pak $A(L) = (1 - L)A^*(L)$, kde A^* je polynom stupně $p - 1$ se všemi kořeny mimo komplexní jednotkový kruh. Proto $A(1) = 0$ a $\sum_{i=1}^{p} a_i = 1$.

- Po reparametrizaci získáme

$$y_t = \alpha y_{t-1} + \sum_{i=1}^{p-1} d_i \Delta y_{t-i} + e_t,\ \alpha = \sum_{i=1}^{p} a_i$$

Test je založen na odhadu předchozí rovnosti, kde se při použití t-statistiky pro testování nulové hypotézy, že $\alpha = 1$.

- Je-li přítomná deterministická část trendu x_t, pak se použije model

$$y_t = \beta x_t + \alpha y_{t-1} + \sum_{i=1}^{p-1} d_i \Delta y_{t-i} + e_t.$$

- Které deterministické komponenty zahrnout záleží na konkrétní časové řadě.

Testování jednotkového kořene při strukturálních změnách funkce trendu

- Pokud se struktura stochastického trendu mění v čase, jsou standardní testy velice slabé (nejsou ani asymptoticky konzistentní).
- Do teď byl diskutován rozdíl mezi DS procesy (trendová funkce se mění stále) a TS procesy (trendová funkce je pořád stejná). Lepší pohdled by byl se zeptat, jestli se mění trend pořád nebo jen občas. Případně jaká je frekvence prenamentních skoků?
- Pokud bychom tyto změny v trendu modelovali jako náhodné veličiny, těžko by se odhadovalo jejich rozdělení.
- Proto je lepší si položit otázku, jestli data ukazují na občasnou změnu v trendu a pokud ano, tak jestli je přítomný jednotkový kořen v náhodné komponentě?

- Jednou možností je použít rozšíření ADF testu na model, ve kterém předpokládáme, že mění pouze směrnici.

$$y_t = \mu_1 + \beta_1 t + (\beta_2 - \beta_1) DT_t^* + u_t,$$

  kde $DT_t^* = t - T_1$ pro $t > T_1$, 0 jinak. Nulová hypotéza zde je, že u_t je I(1), alternativní hypotéza, že je u_t I(0).
- Testování je založeno na následijící regresi

$$y_t = \mu + \beta t + \gamma DT_t^* + \widetilde{y}_t,$$

  kde $\gamma = \beta_2 - \beta_1$ a $\widetilde{y}_t$ je posloupnost s odstraněným trendem (detrending). Test je založen na hodnotě t-statistiky následující autoregresní posloupnosti:

$$\widetilde{y}_t = \alpha \widetilde{y}_{t-1} \sum_{i=1}^{k} a_i \Delta \widetilde{y}_{t-i} + e_t.$$

- Časové momenty změn se dají odhadnout minimalizací/maximalizací t-statistiky při určitých parametrech.

- Výsledný test jednotkového kořene testuje hypotézu, že $\alpha = 1$.
- Problémem v těchto úvahách je, že v závislosti, zda je proces I(0) nebo I(1), limitní rozdělení testů se budou lišit a pokud je proces špatně identifikován, testy budou mít špatné vlastnosti.
- Tento problém řeší tzv. Perronův-Yabuův testovací postup.

Table 2 Perron–Yabu test applied to G, TRF, AMO and to the simulated temperature series Temp_TSB, Temp_DS and Temp_S

Series	Exp-Wald statistic value
G	**12.34**
TRF	**7.92**
AMO	0.28
Temp_TSB	**19.50**
Temp_DS	0.15
Temp_S	−0.31

Bold figures denote statistical significance at the 1 % level

- Další metoda pro testování přítomnosti jednotkového kořene je Kimův-Perronův test. Ten je založen na na Perronově-Yabuově testu a pak na testování $\widetilde{\alpha} = 1$ pro

$$\widetilde{y}_t^n = \widetilde{\alpha}\widetilde{y}_t^n + \sum_{i=1}^{k} c_i \Delta \widetilde{y}_{t-i}^n + \widetilde{e}_t.$$

- Dále existují metody, které testují přítomnost nelineárního trendu se strukturálními změnami.

Perronova-Zhuova metodologie

- Jedná se o metodologii, která odhaduje interval spolehlivosti pro bod, ve kterém dojde ke změně parametrů tredni.
- Uvažujme lineární trend. Deterministická část trendu je dána

$$\tau_t = \mu + \beta t + \gamma DT_t^*,$$

kde $DT_t^* = t - T_1$ pro $t > T_1$, jinak 0 a $T_1 = \lambda T$. Taková funkce je spojitá.

- Metoda je založená na minimalizaci součtu čtverců reziduí, kde metodou nejmenších čtverců odhadneme model

$$_t = \mu + \beta t + \gamma DT_t^* + u_t.$$

- 95% interval spolehlivosti pro změnu trendu je v modelu G $(1946, 1966)$, pro TRF $(1956, 1964)$.

Kointegrace

- Integrované proměnné jsou kointegrované, pokud existuje jejich lineární kombinace, která je stacionární.
- Máme-li $x_t = \mu_{xt} + \epsilon_{xt}$ a $z_t = \mu_{zt} + \epsilon_{zt}$, kde μ_{xt} a μ_{zt} jsou procesy s jednotkovým kořenem, a ϵ_{xt}, ϵ_{zt} jsou stacionární procesy. Tyto procesy jsou kointegrované, pokud existuje lineární kombinace $\alpha_1 x_t + \alpha_2 z_t$ taková, že $\mu_{xt} - (\alpha_2/\alpha_1)\mu_{zt}$ je stacionární a $\alpha_1, \alpha_2 \neq 0$. Z toho plyne, že jsou μ_{xt} a μ_{zt} stejné až na konstantu.
- Pro n proměnných $y_t = (y_{1t}, \ldots, y_{nt})$ jsou kointegrované řádu d, b nebo $y_t \sim CI(d, b)$, pokud jsou všechny integrované řádu d a existuje vektor β, pro který $\beta \cdot y_t$ je integrovaný proces řádu $(d - b)$, $b > 0$.
- Máme-li dva procesy y_{1t}, který reprezentuje globální teplotu, a y_{2t}, který představuje nějaký jiný proces (například koncentraci skleníkových plynů), které jsou kointegrované, můžeme to připsat k tomu, že má lidská činnost skutečně vliv na globální (či hemisférickou) teplotu.

- DS_1 a DS_2 jsou simulované DS řady, které sdílí stochastický trends. DS_3 a DS_4 jsou jsou simulované DS řady s nezávislými stochastickými trendy.

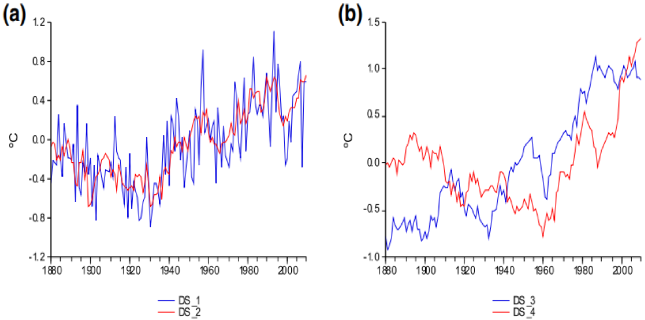

Fig. 2 Cointegrated (**a**) and not cointegrated (**b**) simulated difference stationary series

- Pro testování kointegrace se nejběžněji používají 3 různé testy. Tučně jsou vyjádřeny statisticky významné výsledky testů na hladině významnosti 1%.

Table 4 Cointegration tests applied to simulated variables DS_1 and DS_2, and DS_3 and DS_4

Series	Trace statistic	Max-eigenvalue statistic	Engle–Granger ADF test statistic
DS_1, DS_2	**27.91**	**27.11**	**−11.61**
DS_3, DS_4	4.25	3.77	−1.70

Bold numbers denote statistical significance at the 1 % level

Co-trending

- Zatím jsme řešili pouze lineární trendy, otázkou je, zda dvě či více časových řad sdílí trend?
- Nelineární trendově stacionární model lze zapsat jako $z_t = g(t) + u_t$, kde $g(t) = \beta_0 + \beta_1 t + f(t)$, kde z_t je k-rozměrná časová řada, u_t je k-rozměrný centrovaný stacionární proces a $f(t)$ je deterministická k-rozměrná nelineární trendová funkce, které zohledňují změny v trendu. Nelineární co-trending nastává, když existuje nenulový vektor θ takový, že $\theta \cdot f(t) = 0$.
- Nulová hypotéza je, že mnohorozměrná časová řada z_t obsahuje nelineární co-trending. To znamená, že existuje alespoň jedna lineární kombinace časových řad tak, že je stacionární kolem nějaké konstanty nebo lineárního trendu.

Závěr

Autoři závěrem konstatují, že je literatura zabývající se modelováním globálního oteplování velice široká. I přes to panuje většinová shoda na názoru, že se globální teplota na zemi skutečně zvyšuje a že na ni má lidská činnost zásadní vliv.

Zdroje

- Francisco Estrada, Pierre Perron. Detection and attribution of climate change through econometric methods. Sociedad Matemática Mexicana (2014).