

POSUDEK OPONENTA DIPLOMOVÉ PRÁCE

Název: Ekonometrické modelování a předpovídání spotových cen zemního plynu
Autor: Bc. Barbora Kubišová

SHRNUTÍ OBSAHU PRÁCE

Řešitelka se ve své práci věnuje modelování spotřeby a ceny zemního plynu, a to pomocí jedno- a vícezměrných i vícezměrných modelů časových řad typu GARCH v kombinaci s regresním modelem. V první kapitole jsou přehledově popsány použité modely a testy, jmenovitě testy stacionarity, jedno- a vícezměrné ARMA modely a jejich odhady a verifikace, jednorozměrné i vícezměrné GARCH modely. Následně je představen použitý datový soubor a odhadnuté modely. V kapitole 5 jsou pak simulovány různé budoucí vývoje analyzovaných časových řad. Kapitola 6 je věnována analýze nákladů na riziko změny poptávky.

CELKOVÉ HODNOCENÍ PRÁCE

Téma práce. Téma práce je rozhodně zajímavé. Z matematického hlediska se však bohužel práce omezuje pouze na poměrně stručné a povrchné shrnutí existujících metod a testů a jejich následnou aplikaci na jeden datový soubor.

Vlastní příspěvek. Vlastní příspěvek práce spočívá především v provedení praktické analýzy datového souboru spotřeby a ceny zemního plynu a teploty na Slovensku.

Matematická úroveň. V práci jsou využívány netriviální matematické modely a pokročilé statistické testy. Většina těchto testů a modelů je však představena jen velmi stručně a jejich vlastnosti a vzorce jsou převzaty z literatury. Práce neobsahuje žádný matematický důkaz ani samostatné odvozování.

Práce se zdroji. Veškeré použité zdroje jsou řádně citovány. U článku Kwiatkowski a kol. (1992) však chybí spoluautoři.

Formální úprava. Po formální stránce je práce na velmi dobré úrovni. Jedinou výtku bych měla k popiskům některých obrázků, které jsou velmi těžko čitelné (obr. 5.13, 6.1 a 6.3).

PŘIPOMÍNKY A DOPLŇUJÍCÍ OTÁZKY

1. V práci jsou dost často zaměňovány pojmy „heteroskedasticita“ a „podmíněná heteroskedasticita“. Prosím o vysvětlení těchto dvou pojmů a jejich odlišností pro jednorozměrný ARCH (nebo GARCH) model. Dále pak prosím o bližší vysvětlení a diskuzi aplikace Whiteova testu homoskedasticity pro model z části 3.3 a jakým způsobem nám problém závislosti rozptylu reziduí na regresorech řeší GARCH model.
2. Některé drobné chyby/překlepy (obsahové): str. 5: poslední řádek t_t by mělo být r_t (vzhledem k 1.řádku na str. 6); str. 6: zpoždění o k časových jednotek namísto j ; str. 13: test se aplikuje na $\hat{\varepsilon}_t$, jelikož ε_t nepozorujeme (podobně pak na str. 20-21 pro mnohorozměrný případ); str. 20: výraz \mathbf{a}_t^2 nedává smysl, jelikož se jedná o vektor; str. 23, poslední odstavec: rozdělení ε_t je normální, nikoliv e_t ; str. 27, 4. řádek: \mathbf{C} není čtvercová matice, ale vektor; str. 17: definice maticového autoregresního polynomu $\Psi(\mathbf{z})$ evokuje představu, že by \mathbf{z} měl být vektor, což není, a není jasně uvedeno, co je myšleno pod pojmem „kořen“ tohoto polynomu (jedná se o polynom s maticovými koeficienty); str. 72: počet stupňů volnosti t-rozdělení zjevně není $\nu \in \mathbb{R}$.

3. Popis Q-Q grafu na str. 13 (poslední odstavec) je hodně zvláštní. Co jsou „body z empirického rozdělení“ a co „úsečka znázorňující kvantily normálního rozdělení“?
4. Pojem „rezidua“ je v práci používán poměrně volně — místy jsou pod tímto pojmem myšleny náhodné chyby (bílý šum vstupující do modelu), jinde pak vypočtená rezidua z modelu.
5. Na str. 17, 4. odstavec, je zmíněna stacionarita VARX modelu. Neměli bychom zde předpokládat, že řada $\{\mathbf{x}_t\}$ je stacionární?
6. Inovace $\varepsilon_{t,j}$, $j = 1, 2, 3$, v kapitole 3 mají t-rozdělení nebo standardizované t-rozdělení?
7. Hodnota 17 stupňů Celsia v definici hdd_t je volena subjektivně. Není možné tento bod zvratu odhadnout na základě dat?
8. U modelů v kapitole 3 není jasné, zda prezentované modely jsou finální modely získané postupnou výstavbou modelu, nebo jde o převzaté struktury. Proč např. pro modelování ceny musíme uvažovat jednotlivé dny v týdnu a nestačí odlišit víkend od všedních dnů jako u spotřeby?
9. V modelu pro teplotu v kapitole 3.2 vychází statisticky významný záporný lineární teplotní trend. Jedná se tímto o důkaz, že dochází ke globálnímu ochlazení? Jak to koresponduje s obrázkem 2.3 a s diskuzí vlivu globálního oteplování v závěru práce?
10. Nevím, zda nebylo vhodnější v kapitole 4 trochu sjednotit řád jednotlivých analyzovaných proměnných (změnou jednotek). Takto vychází parametry pro spotřebu řádově několikrát nižší než pro ostatní proměnné a na str. 45 jsou pak uvedeny jako 0.000. Chce-li čtenář zjistit znaménko tohoto koeficientu, musí v tabulkách v příloze zkoumat znaménko t-statistiky, což není velmi pohodlné.
11. Testy stacionarity uvedené v tabulce 4.2 jsou aplikovány na neupravená data nebo nějakým způsobem očištěná data? Jaká verze ADF testu je využita? (Na str.5 je zmíněna existence více variant.)
12. Jak je možné logicky vysvětlit „jednosměrnou závislost teploty na zpožděných hodnotách log výnosů“ ze str. 46?
13. Nevím, zda je úplně ideální kreslit společný histogram s odhadnutou hustotou pro lednová nasimulovaná data (obr. 5.13), jelikož se vzhledem k týdenní sezónnosti, trendu atd. jedná o nestejně rozdělená data. Dále pak mi není jasné, jak je posuzována „úzkost“ a „šířkost“ histogramu pro veličiny měřené v rozličných jednotkách.

ZÁVĚR

Doporučuji uznat předloženou práci jako diplomovou.

RNDr. Šárka Hudecová, PhD.

KPMS MFF UK

27.8.2018