

**MATEMATICKO-FYZIKÁLNÍ
FAKULTA**
Univerzita Karlova

DIPLOMOVÁ PRÁCE

Bc. Barbora Kubišová

**Ekonometrické modelování
a předpovídání spotových cen zemního
plynu**

Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové práce: RNDr. Radek Hendrych, Ph.D.

Studijní program: Matematika

Studijní obor: Pravděpodobnost, matematická statistika
a ekonometrie

Praha 2018

Prohlašuji, že jsem tuto diplomovou práci vypracovala samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

V Praze dne 17. července 2018

Barbora Kubišová

Na tomto mieste by som sa rada poďakovala vedúcemu práce RNDr. Radkovi Hendrychovi, Ph.D. za množstvo jeho odborných a cenných rád, ochotu konzultovať, priateľský prístup a trpezlivosť. Ďakujem tiež spoločnosti innogy Slovensko za poskytnutie vstupných dát. Vďaka patrí aj mojej rodine za ich podporu počas celého štúdia.

Název práce: Ekonometrické modelování a předpovídání spotových cen zemního plynu

Autor: Bc. Barbora Kubišová

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí diplomové práce: RNDr. Radek Hendrych, Ph.D., Katedra pravděpodobnosti a matematické statistiky

Abstrakt: Práca sa venuje problematike modelovania a predpovedania spotovej ceny zemného plynu, jeho spotreby a priemernej dennej teploty. O týchto troch veličinách predpokladáme, že sú navzájom ovplyvňované, keďže so znižujúcou sa teplotou rastie spotreba, čím následne so zvyšujúcim sa dopytom rastie cena. Preto v práci navrhujeme modelovať tieto premenné pomocou vektorovej autoregresie. Tento model porovnávame s jednorozmernými modelmi, kde pre každú radu zostavujeme model z triedy ARMA-GARCH. Modely odhadujeme na historických hodnotách a následne navrhnuté modely použijeme na simuláciu scenárov vývoja. Práve analýza scenárov poskytuje spoločnostiam dodávajúcich plyn informácie o odhade spotreby portfólia a finančných tokoch súvisiacich s nákupom zemného plynu.

Klíčová slova: VARX, BEKK, ARMA, GARCH, zemný plyn

Title: Econometric Modelling and Forecasting of Natural Gas Spot Prices

Author: Bc. Barbora Kubišová

Department: Department of Probability and Mathematical Statistics

Supervisor: RNDr. Radek Hendrych, Ph.D., Department of Probability and Mathematical Statistics

Abstract: The thesis deals with modeling and forecasting of natural gas spot prices, consumption of natural gas and average daily temperature. We assume that these three variables are influenced by each other, because as temperature decreases, consumption increases, which in turn increases the price with the increasing demand. Therefore, we propose to model these variables by vector autoregression. We compare this model with one-dimensional models where for each one we build a model from the ARMA-GARCH class. Models are estimated using historic values and then designed models are used to simulate scenarios. Analysis of scenarios provides information to gas supply companies estimates of portfolio consumption and financial flows related to the purchase concerning natural gas.

Keywords: VARX, BEKK, ARMA, GARCH, natural gas

Obsah

| | |
|---|-----------|
| Úvod | 2 |
| 1 Teória časových rád | 3 |
| 1.1 Jednorozmerné časové rady | 6 |
| 1.1.1 Tvorba modelu Boxovej-Jenkinsovej metodológie | 8 |
| 1.1.2 Predpovede v modeli ARMA | 14 |
| 1.2 Viacrozmerné časové rady | 15 |
| 1.2.1 Výstavba modelu | 18 |
| 1.2.2 Predpoveď v modeli VAR(p) | 22 |
| 1.3 Modelovanie volatility | 22 |
| 1.3.1 Jednorozmerné modely volatility | 23 |
| 1.3.2 Viacrozmerné modely volatility | 25 |
| 2 Trh so zemným plynom | 30 |
| 2.1 Základné pojmy z plynárenstva | 30 |
| 2.2 Analýza vstupných dát | 31 |
| 3 Jednorozmerné modely | 36 |
| 3.1 Model spotreby plynu | 36 |
| 3.2 Model teploty | 38 |
| 3.3 Model spotovej ceny | 40 |
| 4 Viacrozmerný model | 43 |
| 4.1 Odhad modelu VARX | 43 |
| 4.2 Odhad modelu diagonálny BEKK | 47 |
| 5 Analýza scenárov | 51 |
| 5.1 Jednorozmerné modely | 51 |
| 5.2 Jednorozmerný model - bootstrap | 53 |
| 5.3 Viacrozmerný model | 54 |
| 5.4 Viacrozmerný model - bootstrap | 55 |
| 5.5 Porovnanie výsledkov | 57 |
| 6 Praktická úloha | 61 |
| Záver | 66 |
| Zoznam použitej literatúry | 67 |
| Zoznam obrázkov | 69 |
| Zoznam tabuliek | 70 |
| Zoznam použitých skratiek | 71 |
| Prílohy | 72 |

Úvod

Predložená práca sa venuje problematike modelovania a predpovedania spotovej ceny zemného plynu. Zemný plyn je jednou zo základných surovín, ktorá sa využíva vo výrobe, v petrochemickom priemysle a tiež na vykurovanie, varenie a výrobu teplej vody. Okrem modelovania spotovej ceny sa v práci zaoberáme tiež modelovaním teploty a spotreby plynu. Keďže plyn je primárne využívaný na vykurovanie, jeho spotreba výrazne závisí na teplote. Podobne cena plynu so zvyšujúcim sa dopytom, hlavne v zimnom období rastie, a tak predpokladáme, že tieto tri veličiny spolu súvisia.

Cieľom tejto práce je navrhnúť spoločný model spotreby, teploty, ceny a následne simulovať vývoj denných hodnôt pre nasledujúce obdobie. Motiváciou je presnejšie odhadovanie nákladov spojených s nákupom a predajom plynu na spotovom trhu.

Práca je štruktúrovaná nasledovne. Prvá kapitola zhrňuje teoretický základ použitý v praktickej časti práce. Popisuje teóriu jednorozmerných časových rád a modely Boxovej-Jenkinsovej metodológie. Podrobne opisujeme výstavbu modelu, jeho identifikáciu a verifikáciu. Štatistické testy formulujeme postupne podľa použitia pri tvorbe modelu. V ďalšej časti popisujeme modely viacrozmerných časových rád a definujeme pojmy používané vo viacrozmernej analýze. V tejto časti sa zameriavame hlavne na model vektorovej autoregresie. Súčasťou analýzy je tiež modelovanie volatility. Diskutujeme jednorozmerné modely ARCH a GARCH, ako aj ich rozšírenia do viacrozmernej verzie. V závere opisujeme modely podmienenej kovariančnej matice VECH a BEKK.

V druhej kapitole uvádzame základné pojmy plynárenstva nevyhnutné k priblíženiu danej problematiky. Vysvetľujeme rôzne spôsoby nakupovania plynu, ako je forwardový nákup a spotové balansovanie pozície. Ďalej v nej predstavujeme dáta, ktoré poskytla spoločnosť innogy Slovensko s.r.o., ktorá je druhý najväčší dodávateľ zemného plynu, elektriny a energetických riešení na Slovensku. K dispozícií sú denné merania spotreby portfólia domácností a maloobderu od roku 2014, priemerné denné teploty z územia Slovenska a spotové ceny indexu NCG.

Tretia kapitola obsahuje jednotlivé modely pre spotrebu, teplotu a cenu, ich odhad a verifikáciu. Každá z rovníc má modelový tvar ARMA-GARCH.

Štvrtá kapitola je venovaná praktickej implementácii spoločného viacrozmerného modelu VAR s exogénnymi premennými. Odhadnutý model je verifikovaný a vzhľadom na porušenie predpokladu konštantného rozptylu reziduí je nutné pristúpiť k viacrozmernému modelovaniu volatility. Konštruujeme model diagonálny BEKK.

Kapitola 5 sa venuje analýze scenárov a grafickému zobrazeniu výstupov simulácií. Scenáre vývoja generujeme na obdobie jedného roka podľa modelov navrhnutých v kapitole 3 a 4. Výsledky simulácie jednorozmerných modelov a modelu VARX porovnáme a následne aplikujeme na úlohu z praxe, ktorej je venovaná posledná kapitola 6. Záver zhrňuje výsledky práce a diskutuje možné vylepšenia. V prílohách prikladáme dodatočné informácie o odhadnutých modeloch a ďalšie grafické výstupy.

1. Teória časových rád

Úvodná kapitola práce sa venuje teoretickému základu analýzy časových rád, ktorý je použitý neskôr v praktickej časti. Priekopníkmi v systémovom spracovaní časových rád boli Box a Jenkins, ktorí v roku 1970 vydali monografiu s popisom a metodologickým postupom (Box a Jenkins, 1970). Základom tejto analýzy je práca s reziduálnou zložkou a vyšetovanie autokorelačnej štruktúry časových rád. Týmto prístupom sú často analyzované časové rady so silno závislými pozorovaniami.

Definície náhodného procesu, autokovariancie a stacionarity definujeme nasledovne (Prášková, 2016).

Definícia 1 (Náhodný proces). *Nech (Ω, \mathcal{A}, P) je pravdepodobnostný priestor, (S, \mathcal{E}) merateľný priestor a $T \subset \mathbb{R}$. Rodina náhodných veličín $\{y_t, t \in T\}$ definovaných na (Ω, \mathcal{A}, P) s hodnotami v S sa nazýva náhodný (stochastický) proces.*

Ak je indexová množina $T = \{0, \pm 1, \pm 2, \dots\}$, alebo $T \subset \mathbb{Z}$, potom $\{y_t, t \in T\}$ nazývame proces s diskrétnym časom, alebo časová rada. Ďalej definujeme autokovariančnú funkciu a stacionaritu.

Definícia 2 (Autokovariančná funkcia). *Ak $\{y_t, t \in T\}$ je proces s konečnými druhými momentami, tj. $E |y_t|^2 < \infty, \forall t \in T$, potom (všeobecne komplexná) funkcia definovaná na T predpisom*

$$\gamma_k = E [(y_t - E y_t) (\bar{y}_{t-k} - E \bar{y}_{t-k})] = \text{cov} (y_t, y_{t-k})$$

sa nazýva autokovariančná funkcia procesu $\{y_t, t \in T\}$ pre spozdenie k . Hodnota γ_0 sa nazýva rozptyl procesu v čase t .

Podmienkou pre použitie Boxovej-Jenkinsovej metodológie je stacionarita časovej rady, pričom všeobecne vnímame stacionaritu ako stochasticky ustálené chovanie rady. Rozlišujeme dva typy stacionarity. Nech y_1, \dots, y_n , je konečná, n prvková množina náhodných veličín z náhodného procesu $\{y_t, t \in T\}$.

Definícia 3 (Striktná stacionarita). *Hovoríme, že náhodný proces $\{y_t, t \in T\}$ je striktné stacionárny, ak pre ľubovoľné $n \in \mathbb{N}$, pre ľubovoľné reálne y_1, \dots, y_n , pre ľubovoľné t_1, \dots, t_n a h také, že $t_k \in T, t_k + h \in T, 1 \leq k \leq n$, platí*

$$F_{t_1, \dots, t_n}(y_1, \dots, y_n) = F_{t_1+h, \dots, t_n+h}(y_1, \dots, y_n).$$

V tejto definícii označuje $F_{t_1, \dots, t_n}(y_1, \dots, y_n)$ združenú distribučnú funkciu konečnej, n prvkovej, množiny náhodných veličín y_{t_1}, \dots, y_{t_n} z náhodného procesu $\{y_t, t \in T\}$. Striktná stacionarita teda hovorí o tom, že rozdelenie náhodného vektoru je invariantné voči posunom v čase. O niečo slabšie podmienky kladie na náhodný proces slabá stacionarita, ktorá hovorí o invariancii prvého a druhého momentu.

Definícia 4 (Slabá stacionarita). *Náhodný proces $\{y_t, t \in T\}$ s konečnými druhými momentami sa nazýva slabo stacionárny, ak má konštantnú strednú hodnotu a jeho autokovariančná funkcia $\gamma_k = \text{cov}(y_t, y_{t-k})$ je iba funkciou k .*

V ekonometrii sa najčastejšie pracuje so slabou stacionaritou, ktorú odteraz budeme nazývať jednoducho stacionarita. Nasledujúce časti čerpáme hlavne z knihy Cipra (2008).

Ako odhad autokovariančnej funkcie, označme $\hat{\gamma}_k = c_k$, pre stacionárnu radu používame

$$c_k = \frac{1}{n} \sum_{t=k+1}^n (y_t - \bar{y})(y_{t-k} - \bar{y}), \quad k = 0, 1, \dots, n-1, \quad \bar{y} = \frac{1}{n} \sum_{t=1}^n y_t.$$

Potom ďalej definujeme autokorelačnú funkciu pre spozdenie k .

Definícia 5 (Autokorelačná funkcia). Autokorelačná funkcia pre spozdenie k procesu $\{y_t, t \in T\}$ s kladnými rozptylmi je definovaná ako

$$\rho_k = \frac{\gamma_k}{\gamma_0}, \quad k = \dots, -1, 0, 1, \dots$$

Pre autokorelačnú funkciu platí, že $|\rho_k| \leq 1$, $\rho_0 = 1$. Autokovariančná funkcia je párna funkcia, teda platí, že $\gamma_k = \gamma_{-k}$ a z toho zrejme aj $\rho_k = \rho_{-k}$. Preto stačí, keď sa obmedzíme na kladné hodnoty $k \geq 0$. Odhad autokorelačnej funkcie má tvar

$$r_k = \frac{c_k}{c_0}, \quad k = 0, 1, \dots, n-1.$$

Okrem autokorelačnej funkcie (ozn. ACF) pracujeme tiež s parciálnou autokorelačnou funkciou (ozn. PACF). Parciálna autokorelačná funkcia pre oneskorenie $k \in \mathbb{N}$ je definovaná ako parciálny korelačný koeficient medzi y_t a y_{t-k} pri pevných hodnotách $y_{t-k+1}, \dots, y_{t-1}$. Odhad PACF získame pomocou vhodného odhadu koeficientu φ_{kk} v modeli

$$y_t = \delta + \varphi_{k1}y_{t-1} + \varphi_{k2}y_{t-2} + \dots + \varphi_{kk}y_{t-k}.$$

Potom $\hat{\varphi}_{kk}$ označuje hľadaný odhad parciálnej autokorelácie y_t a y_{t-k} . Tento odhad $\hat{\varphi}_{kk}$ budeme označovať r_{kk} . Platí, že $r_{11} = r_1$ a $r_{00} = 1$. V softwaroch sa na odhad PACF používa rekurentný vzorec, ktorý sa riadi použitím vzorcov:

$$\begin{aligned} r_{11} &= r_1 \\ r_{kk} &= \frac{r_k - \sum_{j=1}^{k-1} r_{k-1,j} \cdot r_{k-j}}{1 - \sum_{j=1}^{k-1} r_{k-1,j} \cdot r_j}, \quad \text{pre } k > 1, \end{aligned}$$

kde

$$r_{kj} = r_{k-1,j} - r_{kk} \cdot r_{k-1,k-j} \quad \text{pre } j = 1, \dots, k-1.$$

Graf autokorelačnej funkcie ρ_k pre jednotlivé k nazývame korelogram. Korelogram autokorelačnej, alebo parciálnej autokorelačnej funkcie sa používa na overenie, či je daná rada biely šum. Je v ňom tiež vykreslená krátkodobá dynamika, kde vidíme silu lineárnej závislosti medzi pozorovaniami časovej rady.

Testy na jednotkový koreň

Stacionaritu časových rád testujeme zväčša dvojicou testov, a to ADF a KPSS testami. Tieto testy sa používajú spoločne, pretože majú opačné nulové hypotézy, a tak sa dopĺňajú pri potvrdení, alebo zamietnutí stacionarity. Ak je výsledkom

testovania, že H_0 pri ADF teste je zamietnutá a H_0 pri KPSS teste nemôže byť zamietnutá, potom je potvrdená stacionarita časovej rady. Ak naopak H_0 pri ADF teste nemôže byť zamietnutá a H_0 pri KPSS teste je zamietnutá, potom je potvrdená nestacionarita. Vo zvyšných dvoch prípadoch sme nič nepreukázali. Testy formulujeme podľa knihy Cípra (2008).

Ako prvý uvedieme Dickey-Fullerov test, ktorého hypotéza a alternatíva sú:

$$\begin{aligned} H_0 : & \quad \beta = 0 \leftrightarrow \text{rada nie je stacionárna,} \\ H_1 : & \quad \beta < 0 \leftrightarrow \text{rada je stacionárna,} \end{aligned}$$

kde koeficient $\alpha - 1 = \beta$ plynie z rovnice

$$\begin{aligned} y_t &= \alpha y_{t-1} + u_t \\ y_t - y_{t-1} &= (\alpha - 1)y_{t-1} + u_t \\ \Delta y_t &= \beta y_{t-1} + u_t. \end{aligned}$$

V tejto rovnici značí u_t biely šum a symbolom Δy_t označujeme prvú diferenciu a teda $y_t - y_{t-1}$. Operátor Δ nazývame diferenčný operátor. Testová štatistika Dickey-Fuller testu je vlastne klasický t-test významnosti parametru. Kritický obor na hladine významnosti α je

$$DF = \frac{\hat{\beta}}{s.e.(\hat{\beta})} \leq t_{1-\alpha}^*(n).$$

Avšak testová štatistika nemá za platnosti H_0 a ani asymptoticky t-rozdelenie ako je to v prípade klasického testovania významnosti regresných parametrov. Rozdelenie DF je neštandardné, má ťažšie konce ako t-rozdelenie a kritické hodnoty bolo nutné napočítať simulačne, čo urobili autori pri publikovaní testu, preto používame zvláštnu značku t^* . Pôvodne autori formulovali 3 varianty tohto testu, kedy do rovnice pre Δy_t vložili konštantnú zložku (intercept) a následne aj trend.

Rozšírený Dickey-Fuller test, ADF (Augmented Dickey-Fuller) test, vznikol rozšírením Dickey-Fullerovho testu, ktorý je použiteľný iba v prípade, že u_t je nezávislý biely šum. Tento test umožňuje testovať jednotkový koreň v autoregresných procesoch vyššieho rádu. Testujeme nulovú hypotézu $\psi = 0$ oproti alternatíve $\psi \neq 0$ v rovnici

$$\Delta y_t = \psi y_{t-1} + \sum_{i=1}^p \gamma_i \Delta y_{t-i} + u_t.$$

Rád p môžeme určiť pomocou informačných kritérií. Rozšírený test má rovnakú testovú štatistiku DF a aj kritické hodnoty ako pôvodný Dickey-Fuller test. Podobne existujú tiež varianty testu s interceptom, či trendom.

Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test má opačné hypotézy:

$$\begin{aligned} H_0 : & \quad \text{rada je stacionárna,} \\ H_1 : & \quad \text{rada nie je stacionárna.} \end{aligned}$$

Tento test bol práve navrhnutý opačne ako predošlý ADF. Princíp testu je, že radu y_t , u ktorej chceme overiť jej stacionaritu, rozkladáme na súčet deterministického trendu, náhodnej prechádzky a stacionárnej chyby

$$y_t = \xi t + t_t + \varepsilon_t,$$

kde $r_t = r_{t-1} + u_t$ a u_t je biely šum s rozptylom σ_u^2 . Testujeme nulovú hypotézu $\sigma_u^2 = 0$. Bližšie podrobnosti môžeme čerpať z Kwiatkowski (1992). Po tom, ako sme overili stacionaritu rady, alebo sme pristúpili k prvým diferenciám, či nejakej inej transformácii, môžeme aplikovať Boxovu-Jenkinsovu metodológiu pri výstavbe modelu.

1.1 Jednorozmerné časové rady

Základné modely spomenutej Boxovej-Jenkinsovej metodológie sú lineárny proces, MA proces, AR proces a zmiešaný ARMA proces, ktoré si predstavíme v nasledujúcej časti.

Lineárny proces definujeme nasledovne

$$y_t = \varepsilon_t + \psi_1\varepsilon_{t-1} + \psi_2\varepsilon_{t-2} + \dots,$$

kde $\{\varepsilon_t\}$ je postupnosť nekorelovaných náhodných veličín s nulovou strednou hodnotou a konštantným kladným konečným rozptylom σ^2 , tj. biely šum. Ekvivalenčne môžeme lineárny proces prepísať pomocou operátora spätného posunu B , ktorý oneskorí veličinu v čase o jednu časovú jednotku

$$By_t = y_{t-1}.$$

Podobne definujeme j -tu mocninu operátora časového posunu B^k , ktorý spozdí veličinu o j časových jednotiek

$$B^k y_t = y_{t-k}.$$

Zápis lineárneho procesu pomocou spätných operátorov má tvar

$$y_t = (1 + \psi_1 B + \psi_2 B^2 + \dots)\varepsilon_t = \psi(B)\varepsilon_t.$$

Predpokladáme, že mocninná rada $\psi(z) = (1 + \psi_1 z + \psi_2 z^2 + \dots)$ konverguje pre $|z| \leq 1$, tj. že korene polynómu ležia von z jednotkového kruhu v komplexnej rovine. Táto podmienka je tiež postačujúca pre existenciu lineárneho procesu s nulovou strednou hodnotou, ktorý je stacionárny. Hovoríme, že lineárny proces je invertibilný, ak ho môžeme zapísať v tvare

$$y_t = \pi_1 y_{t-1} + \pi_2 y_{t-2} + \dots + \varepsilon_t, \text{ resp. } \varepsilon_t = \pi(B)y_t.$$

Postačujúcou podmienkou invertibility lineárneho procesu je podobne ako pri overení stacionarity to, že mocninná rada $\pi(z) = (1 - \pi_1 z - \pi_2 z^2 - \dots)$ konverguje pre $|z| \leq 1$. Takéto vyjadrenie je výhodné najmä pri konštrukcii predpovedí. Ostatné procesy, ktoré budeme následne prezentovať sú špeciálnymi prípadmi lineárneho procesu.

Proces kĺzavých súčtov

Proces kĺzavých súčtov rádu q ($q \in \mathbb{N}$), ktorý označujeme $MA(q)$ vzniká useknutím lineárneho procesu v bode oneskorenia q a zapisujeme ho nasledovne

$$y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}.$$

Rada ε_t je biely šum a $\theta_i \in \mathbb{R}$, $i = 1, \dots, q$, $\theta_q \neq 0$ sú parametre modelu. Proces MA(q) je vždy stacionárny s nulovou strednou hodnotou. Jeho rozptyl je daný vzťahom

$$\sigma_y^2 = (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2) \sigma^2.$$

Proces je invertibilný, ak všetky korene z_1, z_2, \dots, z_q polynómu $\theta(z) = (1 + \theta_1 z + \theta_2 z^2 + \dots + \theta_q z^q)$ ležia von z jednotkového kruhu v komplexnej rovine. Autokorelačná funkcia MA(q) procesu je v tvare

$$\rho_k = \begin{cases} \frac{\theta_k + \theta_1 \theta_{k+1} + \dots + \theta_{q-k} \theta_q}{1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2} & \text{pre } k = 1, 2, \dots, q, \\ 0 & \text{pre } k > q. \end{cases}$$

Vidíme že ACF má bod useknutia a to v bode q . Naopak parciálna autokorelačná funkcia ρ_{kk} nemá bod useknutia, ale je obmedzená lineárnou kombináciou geometricky klesajúcich postupností sínusoid s geometricky klesajúcimi amplitúdami.

Autoregresný proces

Autoregresný proces rádu p ($p \in \mathbb{N}$), označme AR(p), má tvar

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p} + \varepsilon_t,$$

kde $\varphi_i \in \mathbb{R}$, $i = 1, \dots, p$, $\varphi_p \neq 0$. V niektorých softwaroch sa však využíva iný zápis, ktorý vyjadríme pomocou operátora spätného posunu. AR(p) proces môžeme tiež zapísať rovnicou

$$y_t - \varphi_1 y_{t-1} - \varphi_2 y_{t-2} - \dots - \varphi_p y_{t-p} = \varphi(B) y_t = \varepsilon_t,$$

kde $\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p$ nazývame autoregresný operátor. Tento proces je stacionárny, ak korene z_1, z_2, \dots, z_p polynómu $\varphi(z)$ ležia von z jednotkového kruhu v komplexnej rovine. Proces AR(p) je vždy invertibilný a rozptyl tohto procesu s nulovou strednou hodnotou je v tvare

$$\sigma_y^2 = \frac{\sigma^2}{(1 + \varphi_1 \rho_1 + \varphi_2 \rho_2 + \dots + \varphi_p \rho_p)}.$$

Autokorelačnú funkciu stacionárneho autoregresného procesu môžeme odvodiť pomocou diferenčných rovníc. Tento postup krátko popíšeme. Najskôr je potrebné postupne vynásobiť rovnicu AR(p) členmi $y_t, y_{t-1}, \dots, y_{t-k}$ pre $k > 0$ a následne prejdeme k stredným hodnotám:

$$\mathbf{E}(y_t y_{t-k}) = \varphi_1 \mathbf{E}(y_{t-1} y_{t-k}) + \varphi_2 \mathbf{E}(y_{t-2} y_{t-k}) + \dots + \varphi_p \mathbf{E}(y_{t-p} y_{t-k}) + \mathbf{E}(\varepsilon_t y_{t-k}).$$

Keďže je AR(p) proces centrovaný, platí, že $\mathbf{E}(y_t y_{t-k}) = \gamma_k$ je autokovariančná funkcia procesu pre spozdenie k . Člen $\mathbf{E}(\varepsilon_t y_{t-k})$ je rovný 0 pre $k > 0$ vzhľadom k možnosti vyjadriť stacionárny AR(p) proces ako lineárny proces. Rovnice so strednými hodnotami podelíme rozptylom procesu a dostaneme diferenčné rovnice

$$\rho_k = \varphi_1 \rho_{k-1} + \varphi_2 \rho_{k-2} + \dots + \varphi_p \rho_{k-p} \text{ pre } k > 0.$$

Pripomíname, že uvedená metodológia sa aplikuje na slabo stacionárne rady. Respektíve ak rada nie je slabo stacionárna, nultým krokom je jej stacionarizácia pomocou vhodnej transformácie rady. Ak model v poslednej fáze neprejde diagnostickými testami, je nutné model upraviť a opakovať postup. Jednotlivé kroky výstavby modelu popíšeme bližšie v nasledujúcich odstavcoch, kde čerpáme opäť z knihy Cipra (2008).

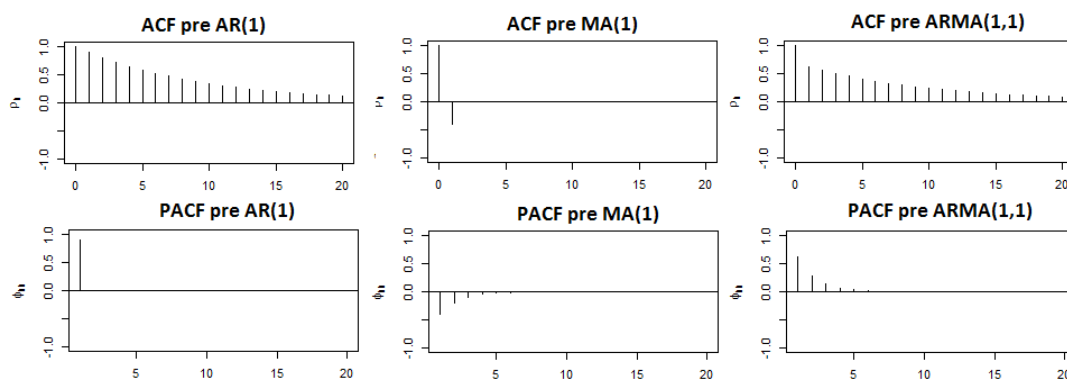
Identifikácia modelu

Pred prvým krokom pri tvorbe modelu Boxovej-Jenkinsovej metodológie je kontrola stacionarity časovej rady. Na časovú radu aplikujeme testy na jednotkový koreň. Ak je potvrdená stacionarita, môžeme pristúpiť k identifikácii rádu modelu, inak by sme museli radu vhodne transformovať (diferencovaním či systematickou dekompozíciou). Základné informácie o type modelu nám poskytuje tvar odhadnutej autokorelačnej a parciálnej autokorelačnej funkcie. Zvyčajne nám postačuje grafické znázornenie pomocou korelogramu s prvými približne 15 hodnotami. Pri definovaní procesov $AR(p)$, $MA(q)$, $ARMA(p, q)$ sme uvádzali aj tvary ACF a PACF funkcií, ktorých zhrnutie uvádzame v tabuľke 1.1 a grafické zobrazenie na obrázku 1.1.

| | $AR(p)$ | $MA(q)$ | $ARMA(p, q)$ |
|------|---------------------|---------------------|---|
| ACF | v tvare krivky U | bod useknutia v q | nemá bod useknutia: v tvare krivky U po prvých $q - p$ hodn. |
| PACF | bod useknutia v p | obmedzená krivkou U | nemá bod useknutia: obmedzená krivkou U po prvých $q - p$ hodn. |

Poznámka: U označuje krivku v tvare lineárnej kombinácie klesajúcich geometrických postupností a sínusoid s klesajúcimi amplitúdami

Tabuľka 1.1: Tvar ACF a PACF funkcií základných modelov



Obr. 1.1: Tvar ACF a PACF funkcií základných modelov

Z tvaru odhadnutého korelogramu a parciálneho korelogramu a tiež na základe bodu useknutia priradíme danej časovej rade vhodný typ modelu. Okrem subjektívneho posúdenia bodu useknutia, ozn. k_0 sa v praxi používa Bartlettova aproximácia pre ACF s asymptotickým kritickým oborom na hladine významnosti $\alpha \in (0, 1)$:

$$|r_k| \geq u_{1-\alpha/2} \sqrt{\frac{1}{n} \left(1 + 2 \sum_{j=1}^{k_0} r_j^2\right)} \text{ pre nejaké } k > k_0,$$

kde $u_{1-\alpha/2}$ je príslušný kvantil štandardizovaného normálneho rozdelenia. Rovnakú hypotézu, avšak teraz pre PACF a to, že $\rho_{kk} = 0$ pre $k > k_0$ môžeme testovať pomocou Quinellovej aproximácie s kritickým oborom na hladine významnosti $\alpha \in (0, 1)$:

$$|r_{kk}| \geq u_{1-\alpha/2} \sqrt{\frac{1}{n}} \text{ pre nejaké } k > k_0,$$

kde $u_{1-\alpha/2}$ je opäť príslušný kvantil štandardizovaného normálneho rozdelenia.

Modernejší prístup k výberu vhodného modelu je použitie informačných kritérií. K identifikácii modelu ARMA(p, q) pristupujeme ako k problému odhadu parametrov p, q na základe optimalizácie

$$(\hat{p}, \hat{q}) = \arg \min_{(k,l)} A(k, l),$$

kde $A(k, l)$ je vhodné informačné kritérium, na ktorého výpočet potrebujeme odhadnúť model ARMA(k, l). Minimalizácia prebieha cez vopred zvolenú sieť hodnôt $k = 0, 1, \dots, K$ a $l = 0, 1, \dots, L$. V súčasnosti sú v softwaroch najviac využívané tieto informačné kritériá: Akaikeho informačné kritérium (ozn. AIC) v tvare

$$AIC(k, l) = \ln \sigma_{k,l}^2 + \frac{(k + l + 1) \cdot 2}{n}$$

a Bayesovo informačné kritérium (ozn. BIC), tiež nazývané Schwartzovo, ktoré viac penalizuje za počet koeficientov, definujeme v tvare

$$BIC(k, l) = \ln \sigma_{k,l}^2 + \frac{(k + l + 1) \cdot \ln n}{n}.$$

V oboch informačných kritériách vystupuje $\sigma_{k,l}^2$, čo je odhadnutý rozptyl bieleho šumu. V softwaroch sa však tieto definície môžu líšiť. Napríklad v EViews (2015) je namiesto prvého členu v obidvoch definíciách $-2LL/n$, kde LL je hodnota logaritmickej vierohodnosti odhadnutého modelu. Druhý člen má v čitateli $k + l + 1$, čo je počet parametrov, ktoré je nutné odhadnúť, pričom v tomto prípade počítame s modelom ARMA(k, l) s interceptom. Ak v modeli vystupujú ešte iné, napríklad exogénne premenné, alebo trend, je nutné ich počet a teda počet parametrov zohľadniť v tomto člene.

Odhad modelu

Model vybraný v predošlom kroku odhadneme. Počiatočné hrubé odhady parametrov modelu, tzv. momentové odhady, môžeme získať zo vzťahov medzi parametrami modelu a autokoreláciami. Tieto odhady slúžia ako počiatočné odhady,

z ktorých štartuje iteračný algoritmus. Odvodené vzorce momentových odhadov môžeme čerpať z knihy Cipra (2008). Napríklad v softwari EViews je možné vybrať si, aké počiatočné hodnoty chceme použiť. Na výber sú rôzne možnosti podľa výberu metódy odhadu. Parametre môžeme odhadovať metódou maximálnej vierohodnosti (ML), zobecných najmenších štvorcov (GLS), alebo podmienenou metódou najmenších štvorcov (CLS). K dispozícii sú numerické optimalizačné algoritmy: BFGS, OPG-BHHH, Kohn-Ansley a Newton-Raphson. V softwari R rôzne knižnice poskytujú rôzne metódy. Najčastejšie sa však stretávame s metódou maximálnej vierohodnosti, alebo podmienenou metódou najmenších štvorcov. Uvedené metódy odhadu sú podrobnejšie opísané v literatúre, pozri napríklad Brockwell a Davis (2002).

Verifikácia modelu

Adekvátnosť modelu $ARMA(p, q)$ kontrolujeme v rámci Boxovej-Jenkinsovej metodológie nasledovnými nástrojmi. Jednou z vlastností vhodného modelu je stacionarita. Model je stacionárny ak korene autoregresného polynómu neležia vo vnútri jednotkového kruhu v komplexnej rovine, tj. pre korene z_1, \dots, z_p rovnice

$$1 - \varphi_1 z - \varphi_2 z^2 - \dots - \varphi_p z^p = 0$$

platí $|z_i| > 1$. Ekvivalentne by sme mohli hovoriť o tom, že prevrátené hodnoty koreňov ležia vo vnútri jednotkového kruhu v komplexnej rovine. Pri zavedení základných modelov ARMA sme spomínali aj invertibilitu, avšak tá nie je dôležitá z hľadiska diagnostiky. Stacionaritu môžeme tiež overiť odozvou na impulz. Jedná sa o analýzu reakcie na jednorázový, resp. opakovaný impulz vo výške nejakého násobku odhadnutej smerodajnej odchýlky bieleho šumu. Tento impulz následne ovplyvňuje hodnoty procesu ARMA a my sledujeme, ako rýchlo a či vôbec tento impulz vyšumí resp. sa stabilizuje. Ak impulz postupne odznie až na nulovú hodnotu, resp. sa odozva po opakovanom impulze stabilizuje na nejakej úrovni, rada je stacionárna.

Ďalšou dôležitou vlastnosťou je nekorelovanosť vypočítaného bieleho šumu. Na overenie nekorelovanosti existuje viacero spôsobov. Jedným z nich je grafický záznam odhadnutého korelogramu a parciálneho korelogramu, v ktorom sledujeme významnosť jednotlivých autokorelácií. Test je založený na Bartlettovej aproximácii, kde nulová hypotéza nekorelovanosti vypočítaného bieleho šumu má na hladine významnosti $\alpha \in (0,1)$ kritický obor:

$$|r_k| \geq u_{1-\alpha/2}/\sqrt{n} \text{ pre } k = 1, 2, \dots,$$

kde $u_{1-\alpha/2}$ je príslušný kvantil štandardizovaného normálneho rozdelenia. Tento test sme spomínali v súvislosti určenia bodu useknutia ACF, avšak jeho využitie je možné aj v tomto prípade. Alternatívny spôsob, ktorý sa používa pri testovaní nekorelovanosti sú Portmanteau testy, niekedy nazývané Q-testy. Tieto testy okrem nekorelovanosti taktiež overujú správnosť použitej štruktúry ARMA.

Portmanteau test

Diagnostický Portmanteau test pre zistenie adekvátnosti $ARMA(p, q)$ modelu zaviedli Box a Pierce v roku 1970. Tento test je založený na asymptotickej distribúcii autokorelácií reziduí. Testom kontrolujeme nekorelovanosť vypočítaného

bieleho šumu a to tak, že testujeme významnosť jeho prvých K autokorelácií. Hypotézy Portmanteau testu sú:

$$\begin{aligned} H_0 : & \quad \rho_1 = \dots = \rho_K = 0, \\ H_1 : & \quad \rho_i \neq 0 \quad \text{pre aspoň jedno } i = 1, \dots, K, \end{aligned}$$

kde $\rho_i = \text{corr}(\varepsilon_t, \varepsilon_{t-i})$, $i = 1, \dots, K$, predstavuje autokorelácie reziduí. Testová štatistika, nazývaná v tomto prípade Q-štatistika, ktorú navrhli Box-Pierce má tvar

$$Q = n \sum_{k=1}^K r_k^2 \stackrel{H_0}{\sim} \chi^2(K - p - q).$$

Druhou možnosťou je Ljung-Boxova Q-štatistika, ktorá je v tvare

$$Q^* = n \cdot (n + 2) \sum_{k=1}^K \frac{r_k^2}{n - k} \stackrel{H_0}{\sim} \chi^2(K - p - q),$$

kde n je počet pozorovaní a $r_k = \text{corr}(\hat{\varepsilon}_t, \hat{\varepsilon}_{t-i})$ sú odhadnuté autokorelácie vypočítaného bieleho šumu. Za nulovej hypotézy majú obidve Q-štatistiky χ^2 rozdelenie s rovnakým stupňom voľnosti $K - p - q$. Kritický obor má na hladine významnosti α tvar

$$Q \geq \chi_{1-\alpha}^2(K - p - q), \quad \text{resp. } Q^* \geq \chi_{1-\alpha}^2(K - p - q).$$

Niekedy sa používa n^2 namiesto $n \cdot (n + 2)$. Konštantu K je nutné zvoliť vopred. Odporúča sa voliť K vo výške \sqrt{n} . Tento test je možné taktiež použiť pri testovaní podmienenej heteroskedasticity. V tomto prípade testujeme významnosť prvých K autokorelácií kvadrátu reziduí.

Whiteov test

Nasleduje test heteroskedasticity, ktorý skúma konštantnosť rozptylu reziduálnej zložky $\text{var}(\varepsilon_t) = \sigma^2 < \infty$. Princíp testu popíšeme na jednoduchom príklade, kedy chceme testovať nulovú hypotézu Whiteovho testu, že rozptyl reziduí v modeli

$$y_t = \beta_1 + \beta_2 x_{t2} + \beta_3 x_{t3} + \varepsilon_t, \quad t = 1, \dots, n$$

je konštantný (homoskedasticita). Pri testovaní sa vytvorí pomocný model

$$\varepsilon_t^2 = \alpha_1 + \alpha_2 x_{t2} + \alpha_3 x_{t3} + \alpha_4 x_{t2}^2 + \alpha_5 x_{t3}^2 + \alpha_6 x_{t2} x_{t3} + u_t,$$

alebo model bez bilineárnych členov $x_{t2} x_{t3}$ podľa výberu, pretože výpočtový software typicky ponúka dve možnosti (with cross terms / without cross terms). V tomto pomocnom modeli značí u_t biely šum. V pomocnom modeli potom testujeme súhrnným F-testom významnosť koeficientov

$$H_0 : \alpha_2 = \dots = \alpha_6 = 0.$$

Kritický obor na hladine významnosti α je

$$\frac{n - 6}{5} \cdot \frac{RRSS - URSS}{URSS} \geq F_{1-\alpha}(5, n - 6),$$

kde $k = 6$ je počet koeficientov pomocného modelu a $m = 5$ počet lineárnych obmedzení. Neobmedzený reziduálny súčet štvorcov $URSS$ je z pomocného modelu a obmedzený reziduálny súčet štvorcov $RRSS$ je z modelu $\varepsilon_t^2 = \alpha_1$. Existujú aj iné testy na homoskedasticitu, napríklad Breusch-Pagan-Godfrey test, avšak žiadny z nich nenaznačuje, čo robiť v prípade zamietnutia nulovej hypotézy.

ARCH LM test

Ďalším možným nástrojom na testovanie prítomnosti podmienenej heteroskedasticity v reziduálnej zložke je ARCH LM test, ktorý navrhol Engle v roku 1982. Už z názvu vyplýva, že sa jedná o test založený na Lagrangeových multiplikátoroch. Testujeme, či rada reziduí $\{\varepsilon_t\}$ nevykazuje dodatočnú ARCH štruktúru, ktorá doposiaľ nebola zohľadnená. Testová štatistika ARCH LM testu sa počíta z pomocnej regresie

$$\varepsilon_t^2 = \beta_0 + \left(\sum_{s=1}^q \beta_s \varepsilon_{t-s}^2 \right) + u_t.$$

Nulová hypotéza testu je, že v kvadrátoch štandardizovaných reziduí nie je žiadna autokorelácia do rádu q . Výstupom môžu byť dve testové štatistiky a to F-štatistika, ktorá testuje spoločnú významnosť všetkých spozdených ε_{t-s}^2 a Englova LM testová štatistika v tvare $n \cdot R^2$, kde n , počet pozorovaní a R^2 , koeficient determinácie, sú z pomocnej regresie. Rozdelenie F-štatistiky za platnosti nulovej hypotézy nepoznáme, avšak $n \cdot R^2$ má za platnosti H_0 asymptoticky χ^2 rozdelenie s q stupňami voľnosti.

Testovanie normality

Predpoklad normality bieleho šumu môžeme overovať pomocou známeho Jarque-Bera testu. Tento štatistický test normality je založený na koeficiente šikmosti μ_3 a špicatosti μ_4 normálne rozdelenej náhodnej veličiny X . Tieto dva koeficienty sú v prípade $X \sim N(\mu_X, \sigma_X^2)$ nasledovné:

$$\mu_3 = \frac{\mathbf{E}(X - \mu_X)^3}{\sigma_X^3} = 0, \quad \mu_4 = \frac{\mathbf{E}(X - \mu_X)^4}{\sigma_X^4} = 3.$$

V praxi testujeme normalitu náhodného výberu pomocou výberového koeficientu šikmosti $\hat{\mu}_3$ a špicatosti $\hat{\mu}_4$. Testová štatistika Jarque-Bera testu definovaná v tvare

$$JB = n \left(\frac{\hat{\mu}_3}{6} + \frac{\hat{\mu}_4 - 3}{24} \right)$$

má za platnosti nulovej hypotézy normálneho rozdelenia asymptoticky χ^2 rozdelenie s 2 stupňami voľnosti. Kritický obor na hladine významnosti α je v tvare

$$JB \geq \chi_{1-\alpha}^2(2).$$

Alternatívou k štatistickému testu je grafický výstup v podobe Q-Q diagramu (ozn. QQ plot), alebo histogramu. V Q-Q diagrame porovnávame kvantily normálneho rozdelenia oproti empirickým kvantilom odhadnutých reziduí z modelu. Ak sú body z empirického rozdelenia blízko úsečky znázorňujúcej kvantily normálneho rozdelenia, značí to zhodu rozdelení. Ak sú na krajoch body empirického rozdelenia ďalej od tejto úsečky, môže sa jednať o rozdelenie s ťažšími chvostami, resp. iný typ rozdelenia. V Q-Q diagrame môžeme samozrejme porovnávať zhodu rozdelenia aj s iným ako normálnym rozdelením.

1.1.2 Predpovede v modeli ARMA

Postup predpovede ukážeme pre stacionárny a invertibilný proces ARMA(p, q) s nulovou strednou hodnotou. Tento proces, keďže je stacionárny, môžeme vyjadriť dvomi spôsobmi:

$$\begin{aligned} y_t &= \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} \\ &= \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + \dots \end{aligned}$$

Označme $\hat{y}_{t+k}(t)$ predpoveď hodnoty y_{t+k} v čase t o k krokov dopredu. Chceme skonštruovať lineárnu predpoveď s najmenšou strednou štvorcovou chybou MSE

$$MSE = \mathbf{E} [y_{t+k} - \hat{y}_{t+k}(t)]^2.$$

Linearitou predpovede chceme dosiahnuť, aby $\hat{y}_{t+k}(t)$ bola lineárnou funkciou hodnôt y_t, y_{t-1}, \dots , alebo $\varepsilon_t, \varepsilon_{t-1}, \dots$, čo je ekvivalentné pri predpoklade stacionarity a invertibility procesu. Z linearity môžeme predpoveď písať v tvare

$$\hat{y}_{t+k}(t) = \pi_k \varepsilon_t + \pi_{k+1} \varepsilon_{t-1} + \dots$$

Z druhej vlastnosti potom hľadáme koeficienty π_k, π_{k+1}, \dots také, aby minimalizovali MSE:

$$\begin{aligned} \mathbf{E} [y_{t+k} - \hat{y}_{t+k}(t)]^2 &= \mathbf{E} [\varepsilon_{t+k} + \psi_1 \varepsilon_{t+k-1} + \dots + \psi_{k-1} \varepsilon_{t+1} + (\psi_k - \pi_k) \varepsilon_t \\ &\quad + (\psi_{k+1} - \pi_{k+1}) \varepsilon_t + (\psi_{k+2} - \pi_{k+2}) \varepsilon_{t-1} + \dots]^2 \\ &= \sigma^2 (1 + \psi_1^2 + \dots + \psi_{k-1}^2 + (\psi_k - \pi_k)^2 + \dots) \end{aligned}$$

Minimalizáciu dosiahneme práve vtedy, keď $\psi_j = \pi_j$ pre všetky $j = k, k+1, \dots$. Hľadaná predpoveď z času t o k krokov dopredu je teda v tvare

$$\hat{y}_{t+k}(t) = \psi_k \varepsilon_t + \psi_{k+1} \varepsilon_{t-1} + \dots$$

Pre chybu predpovede z času t o k krokov dopredu, ozn. $e_{t+k}(t)$, platí

$$e_{t+k}(t) = y_{t+k} - \hat{y}_{t+k}(t) = \varepsilon_{t+k} + \psi_1 \varepsilon_{t+k-1} + \dots + \psi_{k-1} \varepsilon_{t+1}.$$

Stredná hodnota a rozptyl chyby majú tvar

$$\mathbf{E} [e_{t+k}(t)] = 0, \quad \text{var}[e_{t+k}(t)] = 1 + \psi_1^2 + \dots + \psi_{k-1}^2.$$

V praxi konštruujeme predpovede vždy o jeden krok dopredu a následne o ďalší krok atď. Vieme, že platí vzťah

$$y_{t+k} = \varphi_1 y_{t+k-1} + \dots + \varphi_p y_{t+k-p} + \varepsilon_t + \theta_1 \varepsilon_{t+k-1} + \dots + \theta_q \varepsilon_{t+k-q}.$$

Z toho vzhľadom k linearite predpovede platí

$$\hat{y}_{t+k}(t) = \varphi_1 \hat{y}_{t+k-1}(t) + \dots + \varphi_p \hat{y}_{t+k-p}(t) + \varepsilon_t + \theta_1 \hat{\varepsilon}_{t+k-1}(t) + \dots + \theta_q \hat{\varepsilon}_{t+k-q}(t).$$

Stačí, aby sme si uvedomili, že platí

$$\begin{aligned} \hat{y}_{t+j}(t) &= y_{t+j} \text{ pre } j \leq 0, \\ \hat{\varepsilon}_{t+j}(t) &= \begin{cases} \varepsilon_{t+j} = y_{t+j} - \hat{y}_{t+j}(t+j-1) & \text{pre } j \leq 0, \\ 0 & \text{pre } j > 0. \end{cases} \end{aligned}$$

Problematikou predpovedí sme ukončili časť venovanú jednorozmerným časovým radám.

1.2 Viacrozmerné časové rady

Viacrozmernou analýzou časových rád sa snažíme v jednom modeli spoločne modelovať niekoľko časových rád a ich vzájomné väzby. Uvedený prístup je zobecnením jednorozmerných lineárnych procesov a ich špeciálnych prípadov modelov AR, MA a ARMA. Viacrozmerné modely pre časové rady sa v súčasnosti často používajú v analýze ekonomických vzťahov svetových ekonomík, alebo finančných časových rád práve kvôli ich flexibilita a jednoduchosti. Pri písaní tejto časti čerpáme hlavne z kníh Cipra (2008), Lütkepohl (2005) a Tsay (2014).

Nech \mathbf{y}_t je k -rozmerná časová rada, tj. $\mathbf{y}_t = (y_{1t}, y_{2t}, \dots, y_{kt})^T$. Stacionaritu v mnohorozmernom prípade chápeme, podobne ako v jednorozmernom tak, že prvé a druhé momenty časovej rady \mathbf{y}_t sú invariantné voči posunu v čase, za predpokladu, že druhé momenty zložiek existujú a sú konečné. Rozptylovú maticu pre stacionárnu časovú radu definujeme v tvare

$$\mathbf{\Gamma}_0 = \text{var}(\mathbf{y}_t) = \text{E}[(\mathbf{y}_t - \boldsymbol{\mu})(\mathbf{y}_t - \boldsymbol{\mu})^T],$$

kde $\boldsymbol{\mu}$ je vektor stredných hodnôt \mathbf{y}_t . Na diagonále matice $\mathbf{\Gamma}_0$ sú rozptyly zložiek vektoru \mathbf{y}_t a mimo diagonálu kovariance jeho zložiek. Maticu vzájomných kovariancií (cross covariance matrix) pre spozdenie l definujeme v tvare

$$\mathbf{\Gamma}_l = \text{cov}(\mathbf{y}_t, \mathbf{y}_{t-l}) = \text{E}[(\mathbf{y}_t - \boldsymbol{\mu})(\mathbf{y}_{t-l} - \boldsymbol{\mu})^T].$$

Označme $\gamma_{ij}(l)$ prvok (i, j) matice $\mathbf{\Gamma}_l$. Z definície vieme, že tento prvok je kovariancia medzi $y_{i,t}$ a $y_{j,t-l}$. Pre l kladné môžeme $\gamma_{ij}(l)$ chápať ako mieru lineárnej závislosti i -teho komponentu $y_{i,t}$ na l -tej spozdenej hodnote j -teho komponentu $y_{j,t-l}$. Všeobecne platí, že $\gamma_{ij}(l) \neq \gamma_{ji}(l)$ pre $i \neq j$ a $l > 0$ a teda matica $\mathbf{\Gamma}_l$ nie je symetrická. Pre l záporné platí

$$\begin{aligned} \mathbf{\Gamma}_l &= \text{E}[(\mathbf{y}_t - \boldsymbol{\mu})(\mathbf{y}_{t-l} - \boldsymbol{\mu})^T] \\ &= \text{E}[(\mathbf{y}_{t+l} - \boldsymbol{\mu})(\mathbf{y}_t - \boldsymbol{\mu})^T] \quad (\text{zo stacionarity}) \\ &= (\text{E}[(\mathbf{y}_t - \boldsymbol{\mu})(\mathbf{y}_{t+l} - \boldsymbol{\mu})^T])^T \quad (\text{pretože } C = (C^T)^T) \\ &= (\text{E}[(\mathbf{y}_t - \boldsymbol{\mu})(\mathbf{y}_{t-(-l)} - \boldsymbol{\mu})^T])^T \\ &= (\mathbf{\Gamma}_{-l})^T \\ &= (\mathbf{\Gamma}_{-l}^T). \end{aligned}$$

Na rozdiel od jednorozmerného prípadu, kde je autokovariačná funkcia párna, je pre l záporné potrebné transponovať maticu vzájomných kovariancií, pretože platí $\mathbf{\Gamma}_{-l} = \mathbf{\Gamma}_l^T$. Niektorí autori môžu definovať maticu vzájomných autokovariancií obrátene, preto je dobré, vzhľadom na rôznu interpretáciu prvkov mimo diagonálu vedieť rozlíšiť, o ktorú definíciu sa jedná.

Pre stacionárnu radu \mathbf{y}_t definujeme vzájomnú korelačnú maticu (Cross correlation matrix CCM) pre spozdenie l v tvare

$$\boldsymbol{\rho}_l = \mathbf{D}^{-1} \mathbf{\Gamma}_l \mathbf{D}^{-1} = [\rho_{ij}(l)],$$

kde $\mathbf{D} = \text{diag}\{\sigma_1, \sigma_2, \dots, \sigma_k\}$ je diagonálna matica smerodajných odchýlok časových rád (y_{1t}, \dots, y_{kt}) . Teda j -ty prvok na diagonále matice \mathbf{D} je $\sigma_j = \sqrt{\text{var}(y_{jt})}$.

Matica $\boldsymbol{\rho}_l$, podobne ako matica $\boldsymbol{\Gamma}_l$ nie je symetrická. Prvok (i, j) vzájomnej korelačnej matice môžeme vyjadriť vzťahom

$$\rho_{ij}(l) = \frac{\text{cov}(y_{i,t}, y_{j,t-l})}{\sqrt{\text{var}(y_{i,t}) \text{var}(y_{j,t})}}.$$

Odhad maticovej autokovariančnej funkcie pre spozdenie l je v tvare

$$\hat{\boldsymbol{\Gamma}}_l = \mathbf{C}_l = \frac{1}{n} \sum_{t=l+1}^n (\mathbf{y}_t - \bar{\mathbf{y}})(\mathbf{y}_{t-l} - \bar{\mathbf{y}})^T, \quad l = 0, 1, \dots, n-1, \quad \bar{\mathbf{y}} = \frac{1}{n} \sum_{t=1}^n \mathbf{y}_t.$$

Výberovú maticovú autokorelačnú funkciu pre spozdenie l definujeme v tvare

$$\mathbf{R}_l = \hat{\mathbf{D}}^{-1} \mathbf{C}_l \hat{\mathbf{D}}^{-1}, \quad l = 0, 1, \dots, n-1,$$

kde

$$\hat{\mathbf{D}} = \text{diag}\{\sqrt{\mathbf{C}_0}\} = \text{diag}\{\sqrt{\widehat{\text{var}}(y_{1t}), \dots, \sqrt{\widehat{\text{var}}(y_{kt})}\}$$

a prvky $\widehat{\text{var}}(y_{i,t})$, $i = 1, \dots, k$ sú výberové rozptyly, teda diagonálne prvky matice $\hat{\boldsymbol{\Gamma}}_0$. Na základe matice $\boldsymbol{\Gamma}_l$ môžeme odhaliť rôzne lineárne vzťahy v časovej rade \mathbf{y}_t . Hovoríme, že

- Rady $y_{i,t}$ a $y_{j,t}$ sú vzájomne nekorelované, ak $\rho_{ij}(l) = \rho_{ji}(l) = 0 \quad \forall l \geq 0$, tj. medzi radami neexistuje žiadna lineárna závislosť.
- Rady $y_{i,t}$ a $y_{j,t}$ sú súčasne korelované, ak $\rho_{ij}(0) \neq 0$.
- Rady $y_{i,t}$ a $y_{j,t}$ sú nesúčasne nekorelované, ak $\rho_{ij}(l) = \rho_{ji}(l) = 0 \quad \forall l > 0$.
- Ak $\rho_{ij}(l) = 0 \quad \forall l > 0$, ale $\rho_{ji}(l) \neq 0$ pre nejaké $l > 0$, potom existuje jednosmerná závislosť rady $y_{j,t}$ na $y_{i,t}$, tj. $y_{i,t}$ nezávisí na žiadnej minulej hodnote rady $y_{j,t}$, ale $y_{j,t}$ závisí na nejakých minulých hodnotách $y_{i,t}$.
- Ak $\rho_{ij}(l) \neq 0$ pre nejaké $l > 0$ a $\rho_{ji}(m) \neq 0$ pre nejaké $m > 0$, potom medzi radami $y_{j,t}$ a $y_{i,t}$ existuje spätná väzba.

Hovoríme, že k -dimenzionálny reálny náhodný proces $\{\mathbf{a}_t, t \in \mathbb{Z}\}$ je viacrozmernej biely šum, ak splňuje nasledovné vlastnosti:

$$\begin{aligned} \mathbb{E}(a_{it}^2) &< \infty, & i = 1, \dots, k, & \quad \forall t, \\ \mathbb{E}(\mathbf{a}_t) &= \mathbf{0}, & & \quad \forall t, \\ \mathbb{E}(\mathbf{a}_s \mathbf{a}_t^T) &= \boldsymbol{\Sigma}_a, & s = t, & \\ &= \mathbf{0}_k, & s \neq t, & \end{aligned}$$

kde $\boldsymbol{\Sigma}_a$ je pozitívne definitná matica.

Vektorová autoregresia

Model vektorovej autoregresie rádu p ($p \in \mathbb{N}$), ozn. VAR(p), pre mnohorozmernú časovú radu \mathbf{y}_t definujeme v tvare

$$\mathbf{y}_t = \boldsymbol{\Phi}_0 + \boldsymbol{\Phi}_1 \mathbf{y}_{t-1} + \dots + \boldsymbol{\Phi}_p \mathbf{y}_{t-p} + \mathbf{a}_t,$$

kde $\{\mathbf{a}_t\}$ je viacrozmerný biely šum s pozitívne definitnou kovariančnou maticou $\Sigma_a = \text{var}(\mathbf{a}_t)$. Symbolom $\Phi_i, i = 1, \dots, p$, označujeme štvorcové matice koeficientov VAR modelu, Φ_0 je k -dimenzionálny vektor predstavujúci intercept. Rozdelenie vektoru \mathbf{a}_t sa zvyčajne uvažuje mnohorozmerné normálne, alebo mnohorozmerné t-rozdelenie.

Matice koeficientov Φ_i modelu VAR(p) merajú dynamickú závislosť zložiek rady \mathbf{y}_t . Tým myslíme, že ak v modeli VAR(1) je prvok (1, 2) matice Φ_1 rovný 0, potom y_{1t} nezávisí na y_{2t} , pretože koeficient pri spozdení premennej v prvej rovnici je nulový. Ak však prvok (1, 2) matice Φ_1 je 0 a súčasne prvok (2, 1) je rôzny od 0, znamená to, že existuje jednosmerná závislosť a hovoríme, že rada y_{2t} závisí na y_{1t} . Ak sú oba prvky (1, 2) a (2, 1) matice Φ_1 rôzne od 0, hovoríme, že existuje spätná väzba medzi radami y_{1t} a y_{2t} . Matica Σ_a vyjadruje súčasnú závislosť medzi jednotlivými radami. Ak táto matica nie je pozitívne definitná, bolo by možné zredukovať dimenziu vektoru \mathbf{y}_t vynechaním nejakej rady.

Nutná a postačujúca podmienka stacionarity \mathbf{y}_t v modeli VAR(p) je, že maticový autoregresný polynóm $\Phi(z) = \mathbf{I}_k - \Phi_1 z - \dots - \Phi_p z^p$ má všetky korene von z jednotkového kruhu v komplexnej rovine. Vzhľadom na jednoduchšie rozhodovanie používame obrátené korene, ktoré majú byť vo vnútri jednotkového kruhu.

V tejto časti prejdeme ešte k opisu modelu VAR s exogénnymi premennými, ktorý sa nazýva VARX(p, s), kde X vo všeobecnosti označuje prítomnosť exogénnych premenných, p je rád autoregresného procesu a s určuje spozdenie exogénnych premenných. Nech \mathbf{y}_t je k -dimenzionálna časová rada a \mathbf{x}_t je m -dimenzionálna rada exogénnych premenných. Všeobecná forma modelu vektorovej autoregresie s exogénnymi premennými VARX(p, s) je tvaru

$$\mathbf{y}_t = \Phi_0 + \sum_{i=1}^p \Phi_i \mathbf{y}_{t-i} + \sum_{j=0}^s \beta_j \mathbf{x}_{t-j} + \mathbf{a}_t,$$

kde \mathbf{a}_t je viacrozmerný biely šum s pozitívne definitnou kovariančnou maticou Σ_a , $\Phi_i, i = 1, \dots, p$ sú štvorcové matice koeficientov VAR modelu, Φ_0 je k -dimenzionálny vektor predstavujúci intercept a β_j sú $k \times m$ matice koeficientov príslušné exogénnym premenným. Stupne p, s môžeme určiť, podobne ako v modeli VAR(p), na základe informačných kritérií. Podmienka stacionarity je zhodná s podmienkou stacionarity VAR(p) modelu.

Vektorové kľzavé súčty

Viacrozmerný model kľzavých súčtov rádu q ($q \in \mathbb{N}$), nazývaný VMA(q), definujeme v tvare

$$\mathbf{y}_t = \Theta_0 + \mathbf{a}_t + \sum_{i=1}^q \theta_i \mathbf{a}_{t-i},$$

kde Θ_0 je vektor stredných hodnôt \mathbf{y}_t , θ_i sú matice parametrov procesu $k \times k$ a $\{\mathbf{a}_t\}$ je viacrozmerný biely šum s pozitívne definitnou kovariančnou maticou Σ_a . Tento proces je stacionárny, ak existuje rozptylová matica Σ_a . Podobne aj v tomto prípade je možné pridať do modelu exogénne premenné.

Vektorový zmiešaný proces

Jednorozmerné modely ARMA môžeme taktiež zobecniť do vektorového modelu, ktorý nazývame VARMA model. V praxi sa tieto modely využívajú typicky pre nižšie rády p, q ($p \in \mathbb{N}, q \in \mathbb{N}$). Modely VARMA(p, q) definujeme rovnicou

$$\mathbf{y}_t = \Psi_0 + \sum_{i=1}^p \Phi_i \mathbf{y}_{t-i} + \mathbf{a}_t + \sum_{i=1}^q \theta_i \mathbf{a}_{t-i},$$

kde $\{\mathbf{a}_t\}$ je viacrozmerný biely šum s pozitívne definitnou kovariančnou maticou Σ_a . V tomto modeli predpokladáme, že polynómy

$$\Phi(z) = I - \sum_{i=1}^p \Phi_i z^i, \quad \theta(z) = I + \sum_{i=1}^q \theta_i z^i$$

nemajú žiadne spoločné korene. Pri konštrukcii modelu VARMA sa často vyskytuje problém s nejednoznačnosťou. Na jeho vyriešenie sa preto využívajú štrukturálne špecifikácie. Exogénne premenné je možné pridať aj do zmiešaného vektorového procesu, ktorý potom nazývame VARMAX.

1.2.1 Výstavba modelu

Podobne ako sme v jednorozmernom prípade opisovali výstavbu modelu Boxovej-Jenkinsovej metodológie, prebieha aj tvorba modelu viacrozmerných časových rád. Konštrukcia modelu VARMA a jej modifikácií sa skladá z troch fáz: identifikácia, odhad a verifikácia modelu.

Identifikácia rádu modelu

V prípade viacrozmerných modelov by bolo veľmi náročné identifikovať rád modelu pomocou korelogramu, alebo parciálneho korelogramu. Preto sa využívajú iné spôsoby, ktoré sú zvládnuteľné softwarovo a to informačné kritériá, alebo štatistické testy. Podobne ako v jednorozmernom prípade, aj tu má najvhodnejší model najnižšie informačné kritérium. Akaikeho informačné kritérium v m -rozmernej verzii má tvar

$$AIC(k) = \ln |\hat{\Sigma}_k| + \frac{2 \cdot k^*}{n},$$

kde $|\hat{\Sigma}_k|$ je determinant odhadnutej rozptylovej matice reziduálnej zložky a k^* je počet parametrov, ktoré je nutné odhadnúť. Bayesovo informačné kritérium je analogicky ako v jednorozmernom prípade oproti AIC rozdielne v druhej zložke, kde v čitateli nenásobíme 2, ale $\ln(n)$. Jeho tvar je nasledovný

$$BIC(k) = \ln |\hat{\Sigma}_k| + \frac{\ln(n) \cdot k^*}{n}.$$

Definujeme ďalšie kritérium a to Hannah-Quinnovo (HQ) v tvare

$$HQ(k) = \ln |\hat{\Sigma}_k| + \frac{\ln(\ln n) \cdot k^*}{n}.$$

V niektorých softwaroch je ešte okrem troch uvedených klasických informačných kritérií dostupné tiež Akaikeho kritérium založené na chybe odhadu (final prediction error, ozn. FPE), ktorého definíciu nájdeme napríklad v Lütkepohl (2005).

Identifikovať rád modelu môžeme aj pomocou štatistických testov. Ponúka sa na to test pomerov vierohodností, likelihood ratio test (ozn. LR test), ktorý porovnáva dva modely a určuje, ktorý z nich je lepší. Pre určenie rádu modelu sa používa tento test sekvenčne. Princíp sekvenčného testovania v krátkosti opíšeme v nasledujúcom odstavci.

Predpokladajme, že M je horná hranica rádu VAR modelu. LR testom testujeme nasledovné hypotézy a alternatívy:

$$\begin{array}{ll}
H_0^1: & \Phi_M = \mathbf{0}_k \quad \text{vs.} \quad H_1^1: \quad \Phi_M \neq \mathbf{0}_k, \\
H_0^2: & \Phi_{M-1} = \mathbf{0}_k \quad \text{vs.} \quad H_1^2: \quad \Phi_{M-1} \neq \mathbf{0}_k \mid \Phi_M = \mathbf{0}_k, \\
& \vdots \\
H_0^i: & \Phi_{M-i+1} = \mathbf{0}_k \quad \text{vs.} \quad H_1^i: \quad \Phi_{M-i+1} \neq \mathbf{0}_k \mid \Phi_M = \dots = \Phi_{M-i+2} = \mathbf{0}_k, \\
& \vdots \\
H_0^M: & \Phi_1 = \mathbf{0}_k \quad \text{vs.} \quad H_1^M: \quad \Phi_1 \neq \mathbf{0}_k \mid \Phi_M = \dots = \Phi_2 = \mathbf{0}_k.
\end{array}$$

V každom kroku sa nulová hypotéza testuje podmienene na tom, že predošlé nulové hypotézy neboli zamietnuté. Ak je niektorá z nulových hypotéz zamietnutá, napr. H_0^i , končíme testovanie a rád $M-i+1$ označíme za vhodný odhad rádu VAR modelu. Testová štatistika LR testu na testovanie i -tej nulovej hypotézy má kritický obor

$$LR(i) = n \left(\ln |\hat{\Sigma}_a(M-i)| - \ln |\hat{\Sigma}_a(M-i+1)| \right) \geq \chi_{1-\alpha}^2(k^2),$$

kde $\ln |\hat{\Sigma}_a(M-i)|$ je prirodzený logaritmus z determinantu rozptylovej matice odhadnutých rezidií príslušného modelu VAR($M-i$).

Spomenutými informačnými kritériami dostaneme návrhy rôzneho rádu modelu, ktoré slúžia hlavne orientačne. V praxi je však dôležité následné overenie splnenia predpokladov modelu.

Odhad modelu

Model vektorovej autoregresie môžeme odhadovať metódou maximálnej vierohodnosti, alebo metódou najmenších štvorcov.

Za predpokladu mnohorozmerného normálneho rozdelenia \mathbf{a}_t je odhad metódou maximálnej vierohodnosti asymptoticky ekvivalentný s odhadom metódou najmenších štvorcov. Na odhad modelu VMA sa používa metóda podmienenej maximálnej vierohodnosti, alebo presná metóda maximálnej vierohodnosti. Zo skúsenosti pri tvorbe práce odporúčame ako vhodný software pre odhad najširšieho modelu VARMAX SAS, ktorý môžeme následne upraviť na ľubovoľný podmodel (VAR, VARX, VMA, VMAX). EViews, alebo Rko sú v tomto ohľade obmedzenejšie a v nich je možný odhad len modelov VAR, VARX, alebo VMA. Proces VARMA sa zvyčajne odhaduje metódou maximálnej vierohodnosti. Podrobný opis spomenutých metód na odhady modelov je možné nájsť v knihe Lütkepohl (2005), alebo Tsay (2014).

Verifikácia modelu

V nasledujúcej časti uvedieme nástroje na verifikáciu predpokladov odhadnutého modelu a to stacionaritu a vlastnosti bieleho šumu. Zameriame sa pritom

na verifikáciu modelu VAR(p). Na overenie stacionarity potrebujeme vypočítať korene inverzného polynómu k odhadnutému autoregresnému polynómu. Model je stacionárny, ak všetky tieto korene ležia vo vnútri jednotkového kruhu v komplexnej rovine.

Stacionaritu, podobne ako v jednorozmernom prípade môžeme skúmať odozvou na impulz. Vyšetrujeme tým reakciu zvolenej vysvetľovanej premennej na impulz v rovniciach modelu VAR. V každej z k rovníc vyšetrujeme k odoziev na impulz od vysvetľovaných premenných. Ak je daný model VAR(p) stacionárny, odozvy na impulz odznejú na nulovú hodnotu vo všetkých k^2 prípadoch. Dôležité je tiež sledovať rýchlosť, ktorou impulz vyšumí.

Jedná sa o analýzu reakcie na jednorázový, resp. opakovaný impulz vo výške nejakého násobku odhadnutej smerodajnej odchýlky bieleho šumu. Tento impulz následne ovplyvňuje hodnoty procesu ARMA a my sledujeme, ako rýchlo a či vôbec tento impulz vyšumí resp. sa stabilizuje. Ak impulz postupne odznie až na nulovú hodnotu, resp. sa odozva po opakovanom impulze stabilizuje na nejakej úrovni, rada je stacionárna. Popis testovania odozvy na impulz čerpáme z Cipra (2008).

Mnohorozmerný Portmanteau test

Portmanteau test pre overenie nekorelovanosti k -rozmernej rady testuje nulovosť prvých K vzájomných korelačných matíc $\hat{\Gamma}_i$ vypočítaných reziduí. Jeho nulová hypotéza a alternatíva sú:

$$H_0 : \hat{\Gamma}_1 = \dots = \hat{\Gamma}_K = \mathbf{0}_k \text{ vs. } H_0 : \hat{\Gamma}_j \neq \mathbf{0}_k \text{ pre nejaké } 1 \leq j \leq K.$$

Kritický obor Portmanteau testu s Ljung-Boxovou testovou štatistikou je v tvare

$$Q_k(K) = n^2 \sum_{l=1}^K \frac{1}{n-l} \text{tr}(\hat{\Gamma}_l^T \hat{\Gamma}_0^{-1} \hat{\Gamma}_l \hat{\Gamma}_0^{-1}) \geq \chi_{1-\alpha}^2(k^2(K-p)),$$

kde symbolom $\text{tr}(\mathbf{A})$ označujeme stopu (trace) matice \mathbf{A} , teda súčet diagonálnych prvkov matice \mathbf{A} . Konštantu K sa odporúča voliť vo výške \sqrt{n} .

Portmanteau testy slúžia okrem testovania nekorelovanosti aj na testovanie konštantného rozptylu reziduálnej zložky. Uvedený viacrozmerný test s Ljung-Boxovou testovou štatistikou použijeme na vzájomné korelačné matice \mathbf{a}_t^2 . Avšak vo viacrozmernej verzii máme viac prístupov k testovaniu homoskedasticity, pričom niektoré z nich sú dobre použiteľné pri predpoklade normálneho rozdelenia, avšak pri rozdelení s ťažkými chvostami môžu byť tieto testy zlé, „*may fare poorly in finite samples when \mathbf{a}_t has heavy tails*“ Tsay (2014).

Vo viacrozmernom prípade, ak majú reziduá rozdelenie s ťažkými chvostami, by pri testovaní konštantného rozptylu bolo vhodné prejsť k robustnej modifikácii testu heteroskedasticity. Jedným zo spôsobov ako odstrániť efekt ťažkých chvostov je orezať časť pozorovaní. Pri testovaní volatility sa používa jednoduchá procedúra, ktorá odstráni 5% horného chvosta. Pre výpočet testovej štatistiky sa tak použije $0,95 \cdot n$ pozorovaní a to tie, ktoré sú menšie ako 95% kvantil. Tretou verziou testovania podmienenej heteroskedasticity je transformácia k -rozmerného reziduálneho vektoru \mathbf{a}_t na jednorozmerný vektor e_t :

$$e_t = \mathbf{a}_t^T \Sigma_a^{-1} \mathbf{a}_t - k,$$

kde Σ_a je matica nepodmienených rozptylov vektoru \mathbf{a}_t , ako bolo uvedené v definícii viacrozmerného bieleho šumu $\mathbf{E}(\mathbf{a}_t \mathbf{a}_t^T) = \Sigma_a$. Následne testujeme autokorelácie nových e_t jednorozmerným Portmanteau testom s Ljung-Boxovou testovou štatistikou. Ďalšou uvedenou možnosťou testovania homoskedasticity v Tsay (2014) je test založený na kategóriách (rank-based test), ktorý navrhli Dufour a Roy pre e_t . Normalitu reziduálnej zložky testujeme pomocou viacrozmerného Jarque-Bera testu. Bližšie k testovaniu viacrozmernej normality sa môžeme dočítať napríklad v Lütkepohl (2005).

Testovanie kauzality

Príčinnosť medzi jednotlivými blokmi vyšetrovaných premenných testujeme pomocou Grangerových testov kauzality. Kauzalitu podľa Grangera chápeme ako existenciu korelovanosti medzi súčasnou hodnotou jednej premennej a minulými hodnotami iných premenných. Tento vzťah nás zaujíma hlavne kvôli predikcii, kedy jedna rada, ktorá príčinne ovplyvňuje druhú radu pomáha zlepšiť jej predpoveď. V procese VAR je testovanie kauzality prevedené na testovanie nulovosti blokov určitých parametrov. Za predpokladu stacionarity môžeme na vyšetrovanie nulovosti blokov parametrov použiť F-testy, alebo Waldove testy. Nulová hypotéza Grangerovho testu znie:

$$H_0 : \quad \text{spozdené hodnoty } i\text{-tej premennej } y_i \text{ nepôsobia podľa Grangera} \\ \text{na súčasné hodnoty } j\text{-tej premennej } y_j.$$

Označme symbolom $y_i \not\rightleftharpoons y_j$ nulovú hypotézu definovanú vyššie. Keďže chceme skúmať vzťahy medzi premennými, testujeme vždy dvojicu nulových hypotéz:

$$H_{01} : \quad y_i \not\rightleftharpoons y_j \quad (y_i \text{ nepôsobí podľa Grangera na } y_j \text{ v rovnici pre } y_j), \\ H_{02} : \quad y_j \not\rightleftharpoons y_i \quad (y_j \text{ nepôsobí podľa Grangera na } y_i \text{ v rovnici pre } y_i).$$

Existujú štyri možnosti výsledku tohto testovania:

1. Zamietame hypotézy H_{01} a H_{02} tj. y_i pôsobí podľa Grangera na y_j a y_j pôsobí podľa Grangera na y_i . Medzi veličinami existuje spätná väzba (feedback).
2. Zamietame iba H_{01} tj. y_i pôsobí podľa Grangera na y_j , ale y_j nepôsobí podľa Grangera na y_i . Existuje jednosmerná závislosť y_i na y_j a teda spozdené hodnoty y_i sú významné v rovnici pre y_j .
3. Zamietame iba H_{02} tj. y_i nepôsobí podľa Grangera na y_j , ale y_j pôsobí podľa Grangera na y_i . Existuje jednosmerná závislosť y_j na y_i a teda spozdené hodnoty y_j sú významné v rovnici pre y_i .
4. Nezamietame hypotézy H_{01} a H_{02} tj. premenné y_i a y_j sú nezávislé podľa Grangera.

Kauzalitu skúmame ako pre dvojicu premenných, tak aj pre bloky premenných. Nulová hypotéza v tejto variante testu je, že spozdené hodnoty bloku premenných sú nevýznamné a teda model je obmedzený oproti alternatíve, že blok spozdených premenných je významný a nemôže byť zredukovaný.

Uvedme ešte jednoduchý príklad testovania kauzality, ako je uvedené v EViews (2015). Skúmame kauzálny vzťah medzi dvojicou premenných (y_i, y_j) z nejakého bloku premenných. V rovniciach

$$\begin{aligned} y_{j,t} &= \alpha_0 + \alpha_1 y_{j,t-1} + \dots + \alpha_l y_{j,t-l} + \beta_1 y_{i,t-1} + \dots + \beta_l y_{i,t-l} + u_t \\ y_{i,t} &= \alpha_0 + \alpha_1 y_{i,t-1} + \dots + \alpha_l y_{i,t-l} + \beta_1 y_{j,t-1} + \dots + \beta_l y_{j,t-l} + e_t \end{aligned}$$

testujeme pomocou F-testu nulovosť parametrov

$$\beta_1 = \beta_2 = \dots = \beta_l = 0$$

pre každú rovnicu. Nulová hypotéza z prvej rovnice, H_{01} je, že y_i nepôsobí podľa Grangera na y_j v rovnici pre y_j . Za platnosti nulovej hypotézy z druhej rovnice H_{02} platí, že y_j nepôsobí podľa Grangera na y_i v rovnici pre y_i . Počet oneskorení l ktoré majú byť použité v regresných modeloch volíme na základe vlastného uváženia, pričom sa odporúča uprednostniť vyššie l , ako nižšie.

1.2.2 Predpoveď v modeli VAR(p)

V tejto časti sa zameriame len na predpovede v modeli VAR(p), vzhľadom na analogický postup predikcie s jednorozmernými predpoveďami. Podobne ako v jednorozmernom modeli, je predpoveď o jeden krok dopredu z času t , $\hat{\mathbf{y}}_{t+1}(t)$, daná rovnicou

$$\hat{\mathbf{y}}_{t+1}(t) = \hat{\Phi}_0 + \hat{\Phi}_1 \mathbf{y}_t + \dots + \hat{\Phi}_p \mathbf{y}_{t+1-p} + \hat{\mathbf{a}}_{t+1}(t).$$

Predpovede vždy konštruujeme jednokrokovu a teda najprv predikujeme hodnotu $\hat{\mathbf{y}}_{t+1}(t)$ o jeden krok dopredu, z ktorej následne spočítame $\hat{\mathbf{y}}_{t+2}(t)$ atď. Pre predpoveď o j krokov dopredu platí

$$\hat{\mathbf{y}}_{t+j}(t) = \hat{\Phi}_0 + \hat{\Phi}_1 \hat{\mathbf{y}}_{t+j-1}(t) + \dots + \hat{\Phi}_p \hat{\mathbf{y}}_{t+j-p}(t) + \hat{\mathbf{a}}_{t+j}(t).$$

Ďalej platí

$$\begin{aligned} \hat{\mathbf{y}}_{t+j}(t) &= \mathbf{y}_{t+j} \quad \text{pre } j \leq 0, \\ \hat{\mathbf{a}}_{t+j}(t) &= \begin{cases} \mathbf{a}_{t+j} = \mathbf{y}_{t+j} - \hat{\mathbf{y}}_{t+j}(t+j-1) & \text{pre } j \leq 0, \\ 0 & \text{pre } j > 0. \end{cases} \end{aligned}$$

Ak je rada \mathbf{y}_t stacionárna, predpoveď o j krokov dopredu so zväčšujúcim sa j konverguje k vektoru jej stredných hodnôt.

1.3 Modelovanie volatility

Volatilita je dôležitou súčasťou analýzy dát, kedy v dnešnej dobe skúmame veľa finančných časových rád, ktoré majú často nekonštantný rozptyl. Aj keď volatilita nie je priamo pozorovateľná, má typické charakteristiky, ktoré vidíme v časovej rade. Typické rysy finančných dát sú zhlukovanie volatility, kedy sa striedajú obdobia vysokej a nízkej volatility, ďalej pákový efekt, kedy je volatilita vyššia pri poklese cien ako pri ich náraste o rovnakú hodnotu a leptokurtické rozdelenie tj. užší pás okolo stredu, ťažšie konce a väčšia špicatosť oproti normálnemu

rozdeleniu $\mathcal{N}(0,1)$. Modelovanie volatility zohráva dôležitú úlohu v štatistických a finančných aplikáciách a schopnosť odhadnúť a predpovedať volatilitu vedie k lepšiemu pochopeniu súčasných a budúcich finančných rizík.

Najstarší prístup odhadoval volatilitu ako výberový rozptyl, alebo ako smerodajnú odchýlku za určité historické obdobie. O trochu sofistikovanejší bol EWMA model volatility (exponentially weighted moving average), v ktorom váhy pozorovaní klesajú exponenciálne do minulosti. Prelomový prístup k modelovaniu volatility priniesol Engle v roku 1982, ktorý navrhol model ARCH (Autoregressive conditional heteroskedasticity). Neskôr prichádzali rôzne modifikácie tohto modelu, ktoré ho vylepšovali tak, aby dokázal lepšie zachytiť špecifické vlastnosti finančných časových rád.

Jednorozmerné modely však modelujú volatilitu nezávisle na ostatných rádoch čo je nedostatočné pri prepojeniach a nadväznostiach, ktoré medzi veličinami existujú. Na finančných trhoch a medzi rôznymi aktívami dochádza k rozlianiu volatility (spillover) a tiež pri konštrukcii diverzifikovaných portfólií je potrebné poznať korelácie medzi jednotlivými zložkami. Modelovanie mnohorozmernej volatility má tiež významné postavenie v rámci riadenia rizika napríklad v bankách a poisťovniach.

V nasledujúcej časti uvedieme základné typy jednorozmerných a viacrozmerých modelov volatility, ktoré čerpáme z literatúry Cipra (2008) a Tsay (2014).

1.3.1 Jednorozmerné modely volatility

Uvedieme základné typy modelov volatility, a to ARCH a GARCH model. Autoregresný model s podmienenou heteroskedasticitou (Autoregressive conditional heteroskedasticity model, ozn. ARCH model), vychádza z dvoch axiómov a to, že finančné časové rady sú heteroskedastické a volatilita je jednoduchou kvadratickou funkciou minulých predpovedných chýb. Tento prístup je priamou aplikáciou Boxovej-Jenkinsovej metodológie pre volatilitu. Autoregresný model s podmienenou heteroskedasticitou rádu m (ozn. ARCH(m)), $m \in \mathbb{N}$, môžeme zapísať v tvare

$$y_t = \mu_t + e_t, \quad e_t = \sigma_t \varepsilon_t, \quad \sigma_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \dots + \alpha_m e_{t-m}^2, \quad \varepsilon_t \stackrel{iid}{\sim} (0,1),$$

kde parametre $\alpha_i \in \mathbb{R}$, $i = 0, \dots, m$. Ďalej μ_t predstavuje model podmienenej strednej hodnoty a σ_t^2 model podmieneného rozptylu. Zvyčajne je μ_t modelovaná lineárnym regresným modelom, interceptom, alebo ARMA modelom.

Postačujúce podmienky kladných hodnôt σ_t^2 sú v tomto prípade

$$\alpha_0 > 0, \quad \alpha_1 \geq 0, \quad \dots \quad \alpha_m \geq 0.$$

Postačujúcou podmienkou existencie kladného konečného rozptylu a slabej stacionarity rady e_t je

$$\sum_{i=1}^m \alpha_i < 1.$$

Rozdelenie procesu e_t zvyčajne uvažujeme normálne, Studentovo t-rozdelenie, alebo GED rozdelenie. Tento model však má niekoľko nedostatkov, ktoré sa snažia vylepšiť ďalšie modely, viď napríklad Cipra (2008).

Jeden z nich je, že na dostatočné zachytenie vývoja volatility je nutné použiť vysoký rád modelu, čím rastie počet parametrov, ktoré musia spĺňať podmienky pre kladný rozptyl a stacionaritu. V modeli ARCH tiež nie je zohľadnený pákový efekt.

V roku 1986 prišiel Bollerslev so zobecneným modelom ARCH (Generalized ARCH, ozn. GARCH model), ktorý odstraňuje niektoré nedostatky modelu ARCH. V rovnici volatility vystupujú okrem oneskoreného bieleho šumu e_{t-i}^2 taktiež spozdené hodnoty podmieneného rozptylu σ_{t-j}^2 . Model GARCH(m, s), $m \in \mathbb{N}$, $s \in \mathbb{N}$, má tvar

$$y_t = \mu_t + e_t, \quad e_t = \sigma_t \varepsilon_t, \quad \sigma_t^2 = \alpha_0 + \sum_{i=1}^m \alpha_i e_{t-i}^2 + \sum_{j=1}^s \beta_j \sigma_{t-j}^2, \quad \varepsilon_t \stackrel{iid}{\sim} (0,1),$$

kde parametre $\alpha_i \in \mathbb{R}$, $i = 0, \dots, m$, $\beta_j \in \mathbb{R}$, $j = 1, \dots, s$. V modeli môžeme voliť rôzne rozdelenia, napríklad normálne, t-rozdelenie alebo GED rozdelenie. Rovnica strednej hodnoty môže opäť nadobúdať rôzny tvar. Postačujúce podmienky kladných hodnôt σ_t^2 sú v tomto prípade

$$\alpha_0 > 0, \quad \alpha_i \geq 0, \quad \beta_j \geq 0.$$

Postačujúcou podmienkou existencie kladného konečného rozptylu a slabej stacionarity rady e_t v modeli GARCH(m, s) je

$$\sum_{i=1}^m \alpha_i + \sum_{j=1}^s \beta_j < 1.$$

Vyjadrime rovnicu volatility pomocou spätných operátorov

$$\begin{aligned} \sigma_t^2 &= \alpha_0 + \sum_{i=1}^m \alpha_i e_{t-i}^2 + \sum_{j=1}^s \beta_j \sigma_{t-j}^2 \\ &= \alpha_0 + \alpha(B) e_t^2 + \beta(B) \sigma_t^2. \end{aligned}$$

Prevedením posledného člena na druhú stranu dostaneme výraz

$$\sigma_t^2 - \beta(B) \sigma_t^2 = (1 - \beta(B)) \sigma_t^2 = \alpha_0 + \alpha(B) e_t^2,$$

kde $\alpha(B) = \alpha_1 B + \alpha_2 B^2 + \dots + \alpha_m B^m$ a $\beta(B) = \beta_1 B + \beta_2 B^2 + \dots + \beta_s B^s$. Ak korene polynómu $1 - \beta(z)$ ležia von z jednotkového v komplexnej rovine a $\alpha(z), \beta(z)$ nemajú žiadny spoločný koreň, rovnicu volatility môžeme vyjadriť v tvare

$$\sigma_t^2 = \omega^* + \sum_{i=0}^{\infty} \delta_i e_{t-i}^2,$$

kde koeficienty ω^* a δ_i sú určené vzťahmi

$$\omega^* = \frac{\alpha_0}{1 - \beta(z)}, \quad \delta(z) = \sum_{i=0}^{\infty} \delta_i z^i = \frac{\alpha(z)}{1 - \beta(z)}.$$

Z tohto zápisu vidíme, že proces GARCH(m, s) môžeme vyjadriť vo forme procesu ARCH(∞). Jedná sa o analogický postup, ako vyjadrenie procesu ARMA(p, q)

ako kauzálny lineárny proces $MA(\infty)$. Postup získania koeficientov kauzálneho lineárneho procesu a práca s operátormi je podrobne opísaná v Prášková (2016). Ak sú koeficienty ω^* a δ_i nezáporné, je zaručená aj nezápornosť σ_t^2 . Autori v článku Neslon a Cao (1992) pomocou tohto vyjadrenia ukázali, že na to, aby bol zaručený kladný podmienený rozptyl σ_t^2 , nemusia byť splnené všetky podmienky 1.1. V prípade modelu GARCH(1, 2) k zaisteniu kladného σ_t^2 stačí, ak parametre splňujú

$$\alpha_0 > 0, \quad \alpha_1 \geq 0, \quad \beta_1 \geq 0, \quad \alpha_1 \cdot \beta_1 + \alpha_2 \geq 0.$$

Parameter α_2 teda môže byť záporný. Túto definíciu používa napríklad software EViews, ktorý povoľuje zápornosť vybraných koeficientov.

Okrem dvoch spomenutých existuje veľké množstvo iných modelov podmienenej heteroskedasticity, ako napr. integrovaný GARCH model (ozn. IGARCH), prahový GARCH (ozn. TGARCH), exponenciálny GARCH (ozn. EGARCH), GARCH in mean (ozn. M-GARCH) a iné, viď napríklad Cipra (2008).

1.3.2 Viacrozmerné modely volatility

Rovnako ako bolo potrebné použitie viacrozmerých modelov ARMA, rozšírením jednorozmerných GARCH modelov vznikli viacrozmerne GARCH modely (Multivariate GARCH, ozn. MGARCH). Teóriu k tejto časti čerpáme z článku Silvennoinen a Teräsvirta (2008). Nech vektor \mathbf{a}_t s dimenziou $k \times 1$ je náhodný proces taký, že $\mathbf{E} \mathbf{a}_t = \mathbf{0}$. Nech \mathcal{F}_{t-1} je informačný set generovaný pozorovanou radou \mathbf{a}_t do času $t - 1$ vrátane. Predpokladajme, že \mathbf{a}_t je podmienene heteroskedastická rada

$$\mathbf{a}_t = \mathbf{H}_t^{1/2} \mathbf{z}_t,$$

podmienená informáciou \mathcal{F}_{t-1} do času $t - 1$. Štvorcová matica \mathbf{H}_t typu $n \times n$ je podmienená kovariančná matica vektoru \mathbf{a}_t :

$$\text{cov}(\mathbf{a}_t | \mathcal{F}_{t-1}) = \mathbf{H}_t.$$

V tejto rovnici sa na výpočet odmocninovej matice \mathbf{H}_t , ktorá je pozitívne definitná a ktorú označujeme $\mathbf{H}_t^{1/2}$, môže použiť napríklad Choleského rozklad. Druhý člen, \mathbf{z}_t , je postupnosť nezávislých, rovnako rozdelených náhodných veličín s nulovou strednou hodnotou a jednotkovou kovariančnou maticou. Zostáva teda bližšie špecifikovať podmienenú kovariančnú maticu \mathbf{H}_t . Existuje množstvo prístupov, ako túto pozitívne definitnú maticu definovať. Podľa toho môžeme MGARCH modely rozdeliť do 4 kategórií:

- **Modely podmienenej kovariančnej matice:** Vznikli zobecnením jednorozmerných GARCH modelov a to v roku 1988 prácou Bollersleva, Engla a Wooldridgea Bollerslev a kol. (1988), ktorí navrhli model VECH. V tejto modelovej triede sa jedná sa o priame modelovanie matice \mathbf{H}_t , ktorá vzniká súčtom jednotlivých matíc, ktoré reprezentujú ARCH a GARCH zložku. Následne po modeli VECH predstavili v roku 1990 autori Baba, Engle, Kraft a Kroner model BEKK Baba a kol. (1990), ktorý má podobnú štruktúru ako model VECH, avšak navyše zaručuje pozitívnu definitnosť matice \mathbf{H}_t .

- **Faktorové modely:** Motiváciou na vytvorenie týchto modelov bolo zredukovanie počtu parametrov modelu a tiež ekonomická teória, v ktorej predpokladáme, že výnosy sú generované množstvom spoločných, nepozorovaných komponentov, alebo faktorov. Prvý faktorový GARCH model navrhol Engle, Ng a Rotschild v roku 1990 Engle a kol. (1990). V tejto triede predpokladáme, že proces \mathbf{a}_t vzniká malým počtom nepozorovaných heteroskedastických faktorov, ktoré majú štruktúru typu GARCH. Tento prístup je výhodný v tom, že redukuje dimenzionalitu tým, že počet faktorov je nízky vzhľadom k dimenzionalite vektoru \mathbf{a}_t .
- **Modely podmienených rozptylov a korelácií:** Namiesto priameho modelovania matice \mathbf{H}_t sú v tomto prípade modelované najprv jednorozmerné podmienené rozptyly a korelácie, ktoré sú následne použité pri získavaní podmienenej kovariančnej matice. V tejto modelovej triede maticu \mathbf{H}_t rozkladáme na tri matice a to

$$\mathbf{H}_t = \mathbf{D}_t \mathbf{R}_t \mathbf{D}_t.$$

Matica \mathbf{D}_t značí diagonálnu maticu podmienených smerodajných odchýlok \mathbf{a}_t na diagonále a \mathbf{R}_t korelačnú maticu. Podľa toho, či je korelačná matica \mathbf{R}_t konštantná, alebo premenná v čase, rozlišujeme (Constant Conditional Correlation) CCC-GARCH, alebo (Dynamic Conditional Correlation) DCC-GARCH.

- **Neparametrické a semiparametrické metódy:** Tieto modely sú alternatívou ku parametrickým modelom. Kompenzujú stratu efektívnosti parametrických modelov, ktorá môže nastať kvôli zlej špecifikácii štruktúry podmienenej kovariančnej matice. Taktiež sú oproti parametrickým modelom lepšie v efektívnosti, ak reziduá nie sú normálne rozdelené. Semiparametrické modely kombinujú výhody parametrického modelu v tom, že sú konzistentné a interpretovateľné, a neparametrického modelu, ktorý je robustný pri chybných špecifikáciách distribučnej funkcie.

V nasledujúcej časti postupne predstavíme modely prvej modelovej triedy, a teda modely podmienenej kovariančnej matice.

VECH model

Model VECH navrhnutý v práci Bollerslev a kol. (1988) zobecňuje jednorozmerné GARCH modely, kedy podmienené rozptyly a kovariancie sú funkciou nielen spozdených podmienených rozptylov a kovariancií, ale tiež spozdených kvadrátov reziduí. Symbol $VECH(\cdot)$ predstavuje operátor, ktorý usporiada prvky nad a na hlavnej diagonále danej štvorcovej matice do stĺpcového vektoru. Pre maticu \mathbf{H} je jej $VECH(\mathbf{H})$ nasledovný vektor:

$$\mathbf{H} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} \quad VECH(\mathbf{H}) = (h_{11}, h_{22}, h_{33}, h_{12}, h_{13}, h_{23})^T.$$

Zo všeobecnej štvorcovej matice \mathbf{H} typu $k \times k$ tak tento operátor urobí vektor s $k(k+1)/2$ zložkami. Model VECH(p, q), $p \in \mathbb{N}$, $q \in \mathbb{N}$ definujeme rovnicou

$$VECH(\mathbf{H}_t) = \mathbf{C} + \sum_{i=1}^p \mathbf{A}_i \cdot VECH(\mathbf{a}_{t-i} \mathbf{a}_{t-i}^T) + \sum_{j=1}^q \mathbf{B}_j \cdot VECH(\mathbf{H}_{t-j}),$$

kde \mathbf{C} , \mathbf{A}_i , \mathbf{B}_j sú štvorcové matice parametrov modelu. Keďže je matica \mathbf{H}_t symetrická, vektor $VECH(\mathbf{H}_t)$ obsahuje všetky jedinečné prvky matice. Odhadujeme matice \mathbf{C} , \mathbf{A}_i , \mathbf{B}_j , ktoré majú vo všeobecnom modeli celkom

$$\frac{k(k+1)}{2} + (p+q) \cdot \left(\frac{k(k+1)}{2} \right)^2$$

parametrov. Pozitívne na modeli je jeho flexibilita avšak negatívom je množstvo parametrov, ktorých odhad je náročný. Nešťastné je tiež, že máme iba postačujúce podmienky pre pozitívnu definitnosť matice \mathbf{H}_t . V najjednoduchšom prípade má model VECH(1, 1) tvar

$$VECH(\mathbf{H}_t) = \mathbf{C} + \mathbf{A} \cdot VECH(\mathbf{a}_{t-1} \mathbf{a}_{t-1}^T) + \mathbf{B} \cdot VECH(\mathbf{H}_{t-1}),$$

ktorý môžeme rozpísať do tvaru

$$VECH(\mathbf{H}_t) = \begin{pmatrix} c_{11} \\ c_{22} \\ c_{33} \\ c_{12} \\ c_{13} \\ c_{23} \end{pmatrix} + \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} & a_{26} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} & a_{46} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} & a_{56} \\ a_{61} & a_{62} & a_{63} & a_{64} & a_{65} & a_{66} \end{pmatrix} \cdot \begin{pmatrix} a_{t-1,1}^2 \\ a_{t-1,2}^2 \\ a_{t-1,3}^2 \\ a_{t-1,1} \cdot a_{t-1,2} \\ a_{t-1,1} \cdot a_{t-1,3} \\ a_{t-1,2} \cdot a_{t-1,3} \end{pmatrix} \\ + \begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14} & b_{15} & b_{16} \\ b_{21} & b_{22} & b_{23} & b_{24} & b_{25} & b_{26} \\ b_{31} & b_{32} & b_{33} & b_{34} & b_{35} & b_{36} \\ b_{41} & b_{42} & b_{43} & b_{44} & b_{45} & b_{46} \\ b_{51} & b_{52} & b_{53} & b_{54} & b_{55} & b_{56} \\ b_{61} & b_{62} & b_{63} & b_{64} & b_{65} & b_{66} \end{pmatrix} \cdot \begin{pmatrix} h_{t-1,11} \\ h_{t-1,22} \\ h_{t-1,33} \\ h_{t-1,12} \\ h_{t-1,13} \\ h_{t-1,23} \end{pmatrix}$$

V tomto prípade, keď $k = 3$, odhadujeme celkom 78 parametrov. Model je stacionárny práve vtedy, keď sú všetky vlastné čísla matice $\mathbf{A} + \mathbf{B}$ vo vnútri jednotkového kruhu.

V prípade diagonálneho VECH modelu sú štvorcové matice \mathbf{A} , \mathbf{B} diagonálne, čím sa zredukuje počet parametrov v prípade trojrozmerného GARCH modelu z pôvodných 78 na 18.

Zjednodušený model VECH, kde sú matice \mathbf{A}_i , \mathbf{B}_j diagonálne zaviedli pôvodní autori VECH modelu v roku 1988 v snahe získať pozitívne definitnú maticu \mathbf{H}_t a zredukovať počet parametrov, pričom chceli zachovať flexibilitu a interpretovateľnosť modelu. Model je možné odhadovať rovnicu po rovnici, čo sa však zdá už príliš obmedzujúce, pretože nie je povolená interakcia medzi podmienenými rozptylmi a kovarianciami.

Model diagonálny VECH(p, q), $p \in \mathbb{N}$, $q \in \mathbb{N}$ definujeme v tvare

$$\mathbf{H}_t = \mathbf{C} + \sum_{i=1}^p \mathbf{A}_i \odot (\mathbf{a}_{t-i} \mathbf{a}_{t-i}^T) + \sum_{j=1}^q \mathbf{B}_j \odot \mathbf{H}_{t-j},$$

kde matice $\mathbf{C}, \mathbf{A}_i, \mathbf{B}_j$ sú symetrické matice parametrov $k \times k$ a symbolom \odot označujeme Hadamardov súčin matíc, tj. násobenie po zložkách. V tomto modeli odhadujeme celkom

$$(1 + p + q) \cdot \frac{k(k + 1)}{2}$$

parametrov. V rôznych softwaroch je možné použiť rôzne definície pre matice $\mathbf{C}, \mathbf{A}_i, \mathbf{B}_j$. V EViews sú na výber tri možnosti definície pre matice $\mathbf{C}, \mathbf{A}_i, \mathbf{B}_j$ a to konštanta, tj. $\mathbf{C} = c \cdot \mathbf{I}$, $c > 0$, diagonálna matica a symetrická matica.

BEKK model

Ďalší model podmienenej kovariančnej matice je model BEKK, pomenovaný podľa inciálov autorov, navrhnutý v článku Baba a kol. (1990). Tento model je akýsi ďalší vývojový stupeň po diagonálnom modeli VECH. Jeho dôležitou vlastnosťou je však zaručenie pozitívnej definitnej matice \mathbf{H}_t už z definície modelu. Model BEKK(p, q, M), $p \in \mathbb{N}$, $q \in \mathbb{N}$, $M \in \mathbb{N}$ v jeho všeobecnej podobe definujeme rovnicou

$$\mathbf{H}_t = \mathbf{C} \cdot \mathbf{C}^T + \sum_{i=1}^p \sum_{m=1}^M \mathbf{A}_{mi}^T \cdot \mathbf{a}_{t-i} \mathbf{a}_{t-i}^T \cdot \mathbf{A}_{mi} + \sum_{j=1}^q \sum_{m=1}^M \mathbf{G}_{mj} \cdot \mathbf{H}_{t-j} \cdot \mathbf{G}_{mj}^T,$$

kde \mathbf{C} je horná trojuholníková a $\mathbf{A}_{mi}, \mathbf{G}_{mj}$ sú štvorcové matice parametrov typu $k \times k$. V tomto neobmedzenom modeli BEKK(p, q, M) máme

$$\frac{k(k + 1)}{2} + M(p + q)k^2$$

parametrov. Niekedy sa v literatúre definuje model BEKK s opačným poradím transpozície matíc, $\mathbf{C}^T \cdot \mathbf{C}$, čo však nič nemení na jeho vlastnostiach. Model BEKK(p, q, M) je stacionárny práve vtedy, keď všetky vlastné čísla matice

$$\sum_{i=1}^p \sum_{m=1}^M \mathbf{A}_{mi} \otimes \mathbf{A}_{mi} + \sum_{j=1}^q \sum_{m=1}^M \mathbf{G}_{mj} \otimes \mathbf{G}_{mj},$$

sú vo vnútri jednotkového kruhu. Táto matica vznikne Kroneckerovým súčinom matíc (ozn. \otimes), ktorý je definovaný v prílohe práce, časť A.1.

Z každého modelu BEKK je možné skonštruovať jedinečný model VECH, ktorý dáva pozitívne definitnú maticu \mathbf{H}_t . V článku Engle a Kroner (1995) sú uvedené postačujúce podmienky pre ekvivalenciu medzi modelmi BEKK a VECH. Ak je $M > 1$ príliš veľké, môže nastať problém s identifikáciou, preto sa zvyčajne v aplikáciách používa $M = 1$. Taktiež sú v praxi typické nízke rády modelu p, q .

Najjednoduchší model BEKK(1, 1) s $M = 1$ je definovaný v tvare

$$\mathbf{H}_t = \mathbf{C} \cdot \mathbf{C}^T + \mathbf{A}^T \cdot \mathbf{a}_{t-1} \mathbf{a}_{t-1}^T \cdot \mathbf{A} + \mathbf{G} \cdot \mathbf{H}_{t-1} \cdot \mathbf{G}^T.$$

Počet parametrov v tomto modeli pri $k = 3$ rozmernom vektore \mathbf{a}_t a pri všeobecnej definícii matíc $\mathbf{C}, \mathbf{A}, \mathbf{G}$ je 24. Ak by sme chceli znížiť počet parametrov a následne znížiť tiež všeobecnosť modelu, môžeme zaviesť diagonálny BEKK model, kde sú matice \mathbf{A}, \mathbf{G} diagonálne. V diagonálnom modeli BEKK by sme pre $k = 3$ odhadovali už len 13 parametrov. Vidíme, že sa ponúka veľa možností modifikácií uvedených modelov.

Odhad a verifikácia modelov volatility

Jednorozmerné modely volatility odhadujeme softwarovo, metódou maximálnej vierohodnosti. Rád modelu určíme buď z bodu useknutia parciálnej autokorelačnej funkcie kvadrátu reziduí, alebo postupným výberom modelu. Začíname najjednoduchším modelom a postupne zvyšujeme rád, kým nemôžeme zamietnuť nulovú hypotézu ARCH LM testu. Odhadnutú rovnicu volatility verifikujeme použitím testov na štandardizované reziduá $\tilde{\varepsilon}_t$ dané vzťahom:

$$\tilde{\varepsilon}_t = \frac{\varepsilon_t}{\hat{\sigma}_t}.$$

Portmanteau Q-testami na kvadráty štandardizovaných reziduí $\tilde{\varepsilon}_t^2$ overujeme nekorelovanosť. Pravdepodobnostné rozdelenie rady $\{\tilde{\varepsilon}_t\}$ môžeme posudzovať pomocou Q-Q diagramu, histogramu, alebo normálne rozdelenie pomocou Jarque-Bera testu.

Adekvátnosť odhadnutých viacrozmerných modelov volatility môžeme testovať podobne ako v jednorozmerných ARCH, či GARCH modeloch, avšak musíme pristúpiť k viacrozmerným verziám testov. V Lütkepohl (2005), str. 557-584 sú uvedené testy založené na Lagrangeových multiplikátoroch pre overenie použitia modelu MGARCH. Taktiež je možné použiť mnohorozmerný Portmanteau test na kvadráty reziduí (viď 1.2.1 Mnohorozmerný Portmanteau test, alebo Tsay (2014)) a normalitu môžeme testovať pomocou viacrozmerného Jarque-Bera testu.

V prílohe A.1 tejto práce definujeme použité pojmy a to Hadamarov a Kroneckerov súčin, ako aj základné pravdepodobnostné rozdelenia používané v práci, normálne, t-rozdelenie a ich viacrozmerné verzie.

2. Trh so zemným plynom

V nasledujúcej kapitole ozrejmime základné pojmy oblasti plynárenstva a priblížime problematiku nákupu plynu, ktorá je dôležitá pre plánovanie finančných nákladov dodávateľa plynu. Ďalej predstavíme hlavné analyzované premenné z dátového súboru a to spotreba plynu, priemerná denná teplota a spotová cena plynu.

2.1 Základné pojmy z plynárenstva

Zemný plyn je jednou zo základných fosílnych surovín. Nachádza v podzemných náleziskách spolu s ropou, odkiaľ sa ťaží a prepravuje až ku koncovým zákazníkom. Zemný plyn sa využíva na rôzne účely. Slúži ako zdroj energie na zabezpečenie výrobných technických procesov, ako základná surovina pre výrobu syntetických polymérov a iných chemických produktov (dusíkaté hnojivá). Tak tiež sa využíva na vykurovanie, varenie a prípravu teplej vody v domácnostiach. Počiatok vzniku plynárenstva ako odvetvia sa datuje do 18. a 19. storočia, kedy Londýn prichádza s osvetlením ulíc a domov sviatplynom. Úlohou dodávateľov plynu je zabezpečiť dodávku plynu koncovým zákazníkom. Dodávateľia nakupujú plyn cez dlhodobý kontrakt podľa očakávanej spotreby portfólia. Pričom nakupujú vždy mesačné baseloady, a to rovnaké množstvo plynu na každú hodinu v mesiaci. Jedná sa o tzv. forwardový nákup, pričom môžu nakupovať ročné, kvartálne, alebo mesačné objemy. Forwardový kontrakt je dohoda medzi dvoma stranami o nákupe a predaji aktíva v určitom čase v budúcnosti za cenu stanovenú v súčasnosti. Tieto obchody sa uzatvárajú over-the-counter. Spoločnosť si tak zaistuje (hedguje) dodávku plynu. Denné nakúpené množstvo plynu v daný mesiac je teda podiel nakúpeného plynu počtom dní v mesiaci. Prirodzene nekonštantnú spotrebu plynu v jednotlivých dňoch, spôsobenú napríklad aj výkyvmi počasia, odstávkami výroby zákazníkov, pracovnými, nepracovnými dňami, musia dodávateľia kompenzovať buď skladovaním plynu v zásobníkoch, alebo flexibilitou dlhodobého kontraktu. Pričom flexibilita dlhodobého kontraktu znamená dokup, alebo odpredaj plynu za spotové ceny na trhu. V tejto diplomovej práci sa sústreďujeme iba na portfólio domácností a maloodberu (ozn. MD), ktorých spotreba plynu je primárne využívaná na vykurovanie, prípravu teplej vody a varenie. Ako sme stanovili v úvode, motiváciou práce je presnejšie odhadovanie nákladov spojených s nákupom a predajom plynu na spotovom trhu. Dosiahnuť to chceme tým, že vylepšíme a rozšírime doteraz v innogy používané jednorozmerné modely a navrhujeme spoločný model troch veličín, tj. spotreby, teploty, ceny a následne simulujeme vývoj denných hodnôt pre nasledujúce obdobie.

Spotová cena je omnoho volatilnejšia ako cena dlhodobých kontraktov a odráža bezprostredný vývoj dopytu a ponuky na trhu s touto komoditou, teda aj vývoj takých faktorov, ako je napríklad menový kurz (USD/EUR), cena ropy, uhlia, počasie, živelné pohromy (zemetrasenie, silný vietor), výbuch zásobníku, prerušenie dodávky, prípadne odstávky ťažobných zariadení, vývoj ekonomiky a hospodárstva, politické zmeny, vojnové konflikty atď. Zmena v týchto faktoroch sa najviac prejaví na cene krátkodobých kontraktov ako je spotová cena, cena nasledujúceho mesiaca, alebo kvartálu. Spotové trhy s plynom v Európe sú zväčša málo likvidné, jedná sa totiž o regionálne trhy, čo vyplýva z finančne náročnej-

šieho transportu a limitovaných možností uskladňovania plynu. Tieto trhy sú tak oveľa menšie ako napríklad svetový trh ropy, ktorý je ovplyvňovaný hlavne vývojom globálneho dopytu. Hlavné dodávky zemného plynu pre trh strednej Európy pochádzajú z Ruska a Nórska.

V súčasnosti existuje niekoľko energetických trhov, kde je spotová cena tvorená tržnými mechanizmami, tj. založená na vývoji dopytu a ponuky. Ako príklad uvedieme burzy APX Endex, CEGH, ICE, Powernext. Nami študované spotové ceny pochádzajú z energetickej burzy EEX (Energy Exchange), ktorá je v súčasnosti zlúčená cez platformu Pegas s burzou Powernext. Pegas je členom skupiny EEX Group a je centrálnou platformou pre obchody s plynom v Európe. Majetkové zloženie tejto burzy je zložitá a preto budeme ďalej ako zdroj spotových cien označovať burzu EEX. V tejto práci pracujeme so spotovými cenami zemného plynu pre index NCG (NetConnectGermany - virtuálny obchodný bod). Keďže na Slovensku neexistuje virtuálny obchodný bod pre nákup a predaj plynu na báze spotového trhu, za referenčný bod sa považuje práve NCG. Podľa Vyhlášky je virtuálnym obchodným bodom bod lokalizovaný v prepravnej sieti medzi vstupnými bodmi a výstupnými bodmi, v ktorom je možná zmena vlastníctva plynu, viď Vyhláška č. 24/2013 Z. z. (2013). Na platforme PEGAS sú verejne dostupné informácie aj o iných obchodných bodoch, ich spotové a forwardové ceny, ako aj objemy obchodovania a to napríklad pre holandský virtuálny bod TTF, rakúsky CEGH, nemecký NCG a Gaspool, britský NBP, belgický Zeebrugge.

Spotová cena je počítaná podľa metodiky pre Európsky plynový spotový index (European Gas Spot Index EGSi). Tento výpočet je založený na obchodovaných cenách a objemoch kontraktov Day Ahead a Weekend pre daný obchodný bod. Cena je vypočítaná ako objemovo vážený priemer uskutočnených obchodov medzi 8:00 a 18:00 a je publikovaná každý obchodný deň o 18:45 stredoeurópskeho času.

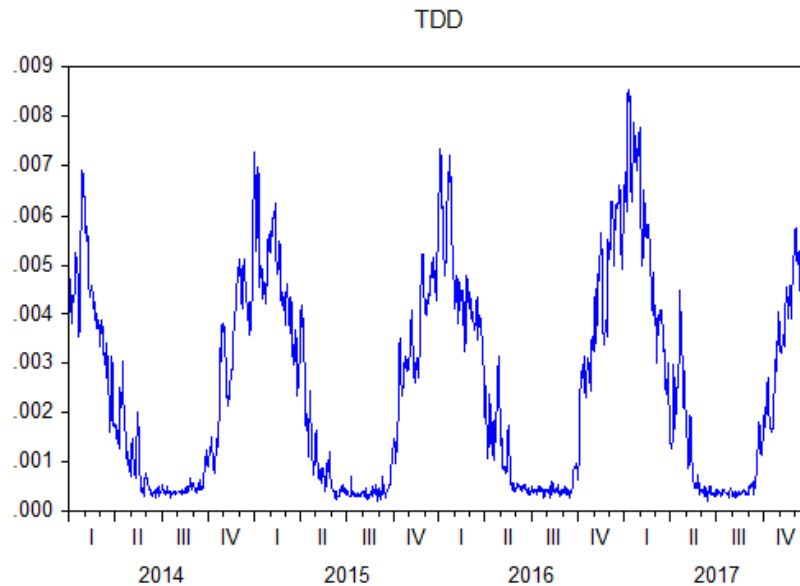
2.2 Analýza vstupných dát

Spotreba plynu je v tomto prípade vyjadrená premennou teplotný diagram dodávky tdd . Hodnota tdd_t je percentuálne vyjadrenie spotreby portfólia pre daný deň t z celkovej ročnej spotreby, platí $0 < tdd_t < 1$. V ideálnom prípade by malo platiť, že súčet tdd_t za všetky dni v roku je 1. Avšak v skutočnosti portfólio odoberie za celý rok menej ako je ročné zazmluvnené množstvo. Vysčítanie všetkých tdd_t v danom roku sa pohybuje zvyčajne okolo 0,9 až 0,95. Hodnotu tdd_t pre deň t spočítame tak, že celkovú spotrebu portfólia MD pre deň t vydáme aktuálnym ročným kontrahovaným množstvom CYC. Hodnota CYC sa môže meniť každý mesiac, vzhľadom na počet získaných a stratených zákazníkov, ktorých prevod dodávateľa plynu je vždy k 1. dňu v mesiaci. Preto je dobré modelovať spotrebu bezrozmernou premennou tdd_t , ktorá je nezávislá na veľkosti portfólia, teda CYC. Spotreba portfólia MD je publikovaná každý deň distribučnou spoločnosťou SPP-D. ¹ Teplotný diagram dodávky má presne opačný vývin ako teplota a teda so stúpajúcou teplotou klesá tdd_t (pozri obrázok 2.1). V letných mesiacoch je tdd_t skoro konštantné, keďže väčšina zákazníkov v tomto období plyn nevyužíva na kúrenie, ale skôr na varenie, ohrev vody.

¹Prístupné na: <https://selfcare.spp-distribucia.sk/SelfCare/logon/?redirect=%2fselfcare>, navštívené 3.7.2018

| Min | Max | Mean | Median | St. dev. |
|-------|-------|-------|--------|----------|
| 0,018 | 0,855 | 0,250 | 0,199 | 0,207 |

Tabuľka 2.1: Deskriptívne štatistiky teplotného diagramu dodávky tdd v %



Obr. 2.1: Denné hodnoty spotreby tdd_t

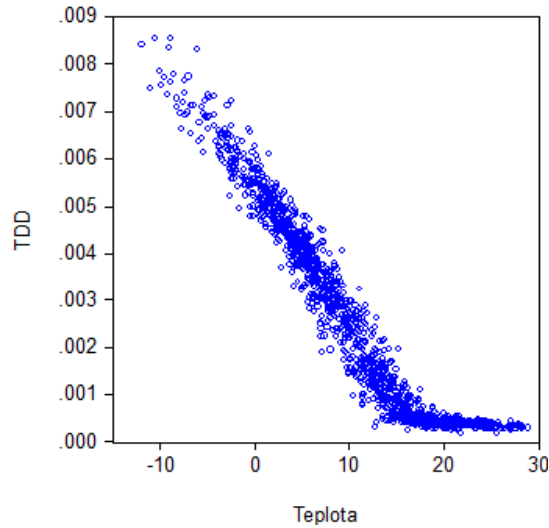
Najvyššia spotreba plynu za pozorované obdobie bola zaznamenaná 8.1.2017 vo výške 0,855 % a najnižšia spotreba koncom leta, 31.8.2015 s hodnotou 0,018 %. Ostatné deskriptívne štatistiky uvádzame v tabuľke 2.1. Závislosť spotreby na teplote môžeme pozorovať taktiež na obrázku 2.2. Vidíme, že do istej teploty (približne 17 °C) je vzťah spotreby na teplote lineárny, a teda so vzrastajúcou teplotou spotreba klesá. Po dosiahnutí dostatočnej teploty už je spotreba plynu konštantná a na teplote nezávisí. Toto pozorovanie viedlo k zavedeniu pojmov hdd (heating degree day) a cdd (cooling degree day). Označme hdd_t je hodnota vykurovania za deň t , ktorú spočítame ako

$$hdd_t = \max(17 - temp_t, 0),$$

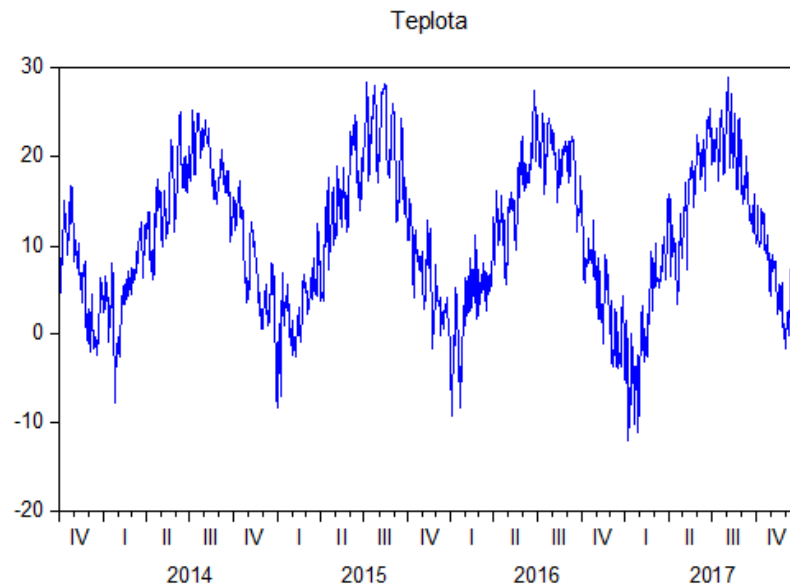
kde 17 °C je prahová hodnota, ktorú sme zvolili subjektívne. Touto veličinou hdd_t môžeme merať energiu potrebnú na vykurovanie. Ak je totiž priemerná vonkajšia teplota nižšia ako istá prahová hodnota, je potrebné začať s vykurovaním budovy. Táto hranica sa môže pre jednotlivé oblasti sveta líšiť. My sme zvolili 17 °C práve kvôli závislosti spotreby, na teplote, ktorá sa viditeľne na obrázku 2.2 láme v tejto hodnote.

Veľmi dôležitou premennou v tejto práci je teplota. Konkrétne priemerná denná teplota na Slovensku, ktorú spočítame ako aritmetický priemer teploty z dvoch meracích staníc, a to letisko Bratislava a letisko Košice. Aj keď je Slovensko členitá krajina a tieto dva meracie body sú približne v rovnakej nadmorskej výške, priemerovanie viacerých miest neprináša výrazné zlepšenie pri modelovaní spotreby plynu.

Na obrázku 2.3 je vykreslený vývoj priemernej teploty za obdobie posledných 4 rokov a v tabuľke 2.2 uvádzame jej základné popisné charakteristiky. Minimálna



Obr. 2.2: Závislosť spotreby plynu na teplote



Obr. 2.3: Vývoj priemernej dennej teploty na Slovensku v °C

priemerná denná teplota za pozorované obdobie nastala 7.1.2017 kedy bolo nameraných $-11,96$ °C a maximálna priemerná denná teplota bola nameraná dňa 4.8.2017 s hodnotou $28,90$ °C.

Spotová cena S_t označená v datasete ako DA je denná referenčná cena plynu z burzy EEX indexu NCG. Ako bolo spomenuté v úvode, spotovú cenu zemného plynu ovplyvňuje množstvo faktorov od cien iných komodít, cez vývoj počasia až po geopolitické zmeny. Pozrime sa tiež na vývoj spotovej ceny zemného plynu od 1.10.2013 až do 31.12.2017 (pozri obrázok 2.4). Základné popisné charakteristiky

| Min | Max | Mean | Median | St. dev. |
|----------|---------|---------|---------|----------|
| $-11,96$ | $28,90$ | $10,79$ | $10,73$ | $8,28$ |

Tabuľka 2.2: Deskriptívne štatistiky priemernej dennej teploty v °C



Obr. 2.4: Spotová cena zemného plynu v EUR/MWh

spotovej ceny NCG uvádzame v tabuľke 2.3.

Z vývoja spotovej ceny je patrný sezónny charakter časovej rady, kedy v letnom období sú ceny nižšie ako v zime. V nasledujúcom odstavci bližšie popíšeme situácie, ktoré ovplyvnili vývoj spotovej ceny zemného plynu. Informácie sme čerpali z výročných správ spoločnosti innogy Slovensko.

Koncom roku 2013 a začiatkom roku 2014 boli ceny najvyššie za pozorované obdobie. Svetové trhy so zemným plynom poznačila v roku 2013 situácia na Ukrajine a zníženie dodávok zemného plynu do Európy, čo zapríčinilo nárast a vyššiu volatilitu spotovej ceny. Maximum DA NCG bolo dosiahnuté koncom roku 2013 a to presne 5.11.2013, kedy bola cena na úrovni 30,25 EUR/MWh. V nasledujúcom roku však ceny klesali v dôsledku teplej zimy na prelome rokov 2014 a 2015 a klesajúcej cene ropy na veľkoobchodných trhoch. V roku 2016 pozorujeme vysokú volatilitu ceny DA NCG. Ponuka plynu do Európy prevažovala nízky dopyt po komodite z dôvodu nadpriemerných teplôt začiatkom roku 2016, čo zapríčinilo prepád cien plynu na trhu. Cena dosiahla svoju najnižšiu hodnotu 20.6.2016 na úrovni 10,74 EUR/MWh. Zmena v cenovom vývoji nastala po rade neplánovaných i plánovaných odstávok na európskom importe s kombináciou ochladenia. V druhej polovici roku 2016 zdvihlo ceny na energetických trhoch v Európe aj viacero tém, ako vystúpenie Británie z EÚ, dohoda krajín OPEC o obmedzení ťažby ropy, či robustne rastúca cena elektriny z dôvodu zníženej produkcie jadrových elektrární vo Francúzsku. Studená zima 2016/2017 zvýšila úroveň spotreby plynu (viditeľné tiež na obrázku 2.1), ktorá spôsobila rast cien. Zvýšenie DA ceny v zimnej sezóne tiež podporilo odstavenie britského zásobníku Rough. Menšie objemy plynu v zásobníkoch spôsobené studenou zimou podporovali zvýšenú volatilitu

| Min | Max | Mean | Median | St. dev. |
|-------|-------|-------|--------|----------|
| 10,74 | 30,25 | 18,73 | 18,43 | 4,06 |

Tabuľka 2.3: Deskriptívne štatistiky spotovej ceny plynu v EUR/MWh

cien, ktoré tak citlivejšie reagovali na zmeny počasia. V závere roka 2017 stúpajúce ceny ropy a uhlia, ako aj viaceré neplánované obmedzenia dodávok v Európe spôsobili rast DA cien plynu. Výraznou udalosťou v decembri 2017 bol výbuch zásobníku plynu v rakúskom Baumgarten, ktorý spôsobil skokovitý nárast DA ceny, ktoré sa ale po obnovení prevádzky skonsolidovali na svoju doterajšiu úroveň.

3. Jednorozmerné modely

V ďalšej časti práce uvedieme 3 osobitné modely a to pre spotrebu plynu, priemernú dennú teplotu a pre logaritmické výnosy spotovej ceny.

3.1 Model spotreby plynu

V predchádzajúcej kapitole sme už ukázali základné popisné charakteristiky týkajúce sa spotreby plynu. V tejto časti práce uvedieme model pre spotrebu plynu. Presnejšie sa jedná o model tdd , avšak vynásobením hodnoty tdd a CYC získame očakávanú spotrebu portfólia. Modelom spotreby sa zaoberal článok Göncü a kol. (2013), kde autori modelovali osobitne dennú spotrebu plynu a teplotu. Na model spotreby použili lineárny regresný model s vysvetľujúcimi premennými hdd , dummy premennou indikujúcou víkend, trendovou zložkou a spotovou cenou.

Ako bolo ukázané vyššie, spotreba závisí na teplote, a preto by mala byť významnou súčasťou modelu. Niekedy sa môže stať, že portfólio nereaguje na teplotu v daný deň a reakcia môže byť posunutá, alebo sa nárast teploty vôbec nemusí prejavíť, preto do modelu vkladáme tiež spozdenú teplotu z predchádzajúceho dňa. Dôležitým faktorom je tiež týždenná sezónnosť, ktorú do modelu vkladáme cez premennú wh . Táto premenná nadobúda hodnoty 0 a 1 ak:

$$wh_t = \begin{cases} 1, & \text{ak je daný deň sviatkov, alebo víkend,} \\ 0, & \text{inak.} \end{cases}$$

Avšak k týždennej sezónnosti pristupujeme inak v letnom a zimnom období, kedy počas leta sa zmena spotreby počas víkendu neprejaví. Tak vznikne nová premenná a to $wh_t \cdot hdd_t$. Do modelu tiež vkladáme goniometrické funkcie na modelovanie ročnej sezónnosti, s dĺžkou sezóny 365,25 kvôli prestupným rokom. Model tdd_t má tvar

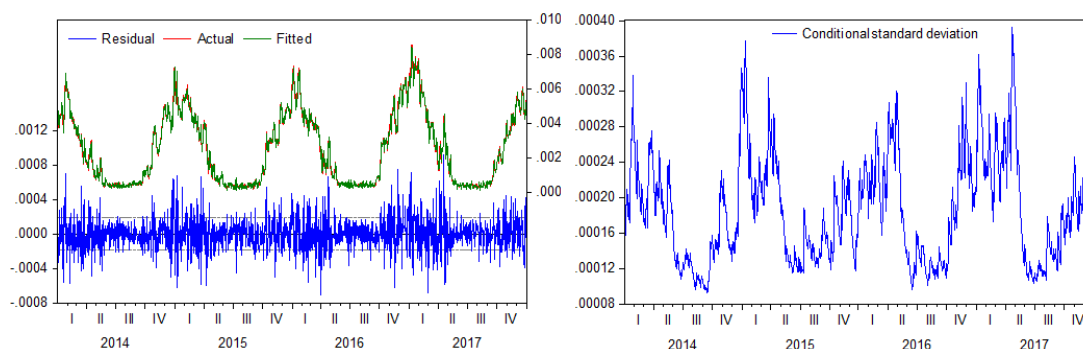
$$\begin{aligned} tdd_t = & a_0 + a_1 \cdot temp_t + a_2 \cdot temp_{t-1} + a_3 \cdot \sin(2\pi t/365,25) \\ & + a_4 \cdot \cos(2\pi t/365,25) + a_5 \cdot wh_t \cdot hdd_t + e_{t,1}, \end{aligned} \quad (3.1)$$

kde $\{e_{t,1}\}$ je biely šum s konštantným konečným rozptylom $\sigma_1^2 > 0$. Model sme odhadli metódou najmenších štvorcov. V tomto jednoduchom modeli bolo zistené porušenie predpokladu autokorelácie a konštantného rozptylu reziduí, kedy Whiteov test homoskedasticity s krížovými členmi s p-hodnotou 0,00 zamietá nulovú hypotézu. Vzhľadom na tieto skutočnosti sme prešli k ARMA(2,0) modelu reziduí, aby sme opravili autokorelovanosť a ku GARCH(1,1) modelu kvôli heteroskedasticite. Vybrali sme tento model s ohľadom na najmenšie AIC. ADF test na jednotkový koreň zamietá nulovú hypotézu a KPSS test nezamietá nulovú hypotézu stacionarity reziduí.

Metódou maximálnej vierohodnosti odhadujeme model

$$\begin{aligned}
 tdd_t &= \underset{(0,00)}{0,004} - \underset{(0,00)}{7,8 \times 10^{-5}} \cdot temp_t - \underset{(0,00)}{1,9 \times 10^{-5}} \cdot temp_{t-1} \\
 &\quad + \underset{(0,00)}{3,7 \times 10^{-4}} \cdot \sin(2\pi t/365,25) + \underset{(0,00)}{0,002} \cdot \cos(2\pi t/365,25) \\
 &\quad - \underset{(0,00)}{7,3 \times 10^{-6}} \cdot wh_t \cdot hdd_t + e_{t,1}, \\
 e_{t,1} &= \underset{(0,00)}{1,157} \cdot e_{t-1,1} - \underset{(0,00)}{0,206} \cdot e_{t-2,1} + u_{t,1}, \\
 u_{t,1} &= \varepsilon_{t,1} \cdot \sigma_{t,1}, \quad \varepsilon_{t,1} \stackrel{iid}{\sim} t_{3,993}, \\
 \sigma_{t,1}^2 &= \underset{(0,10)}{2,3 \times 10^{-9}} + \underset{(0,00)}{0,050} \cdot u_{t-1,1}^2 + \underset{(0,00)}{0,899} \cdot \sigma_{t-1,1}^2,
 \end{aligned}$$

kde v zátvorkách je p-hodnota testu významnosti parametrov.

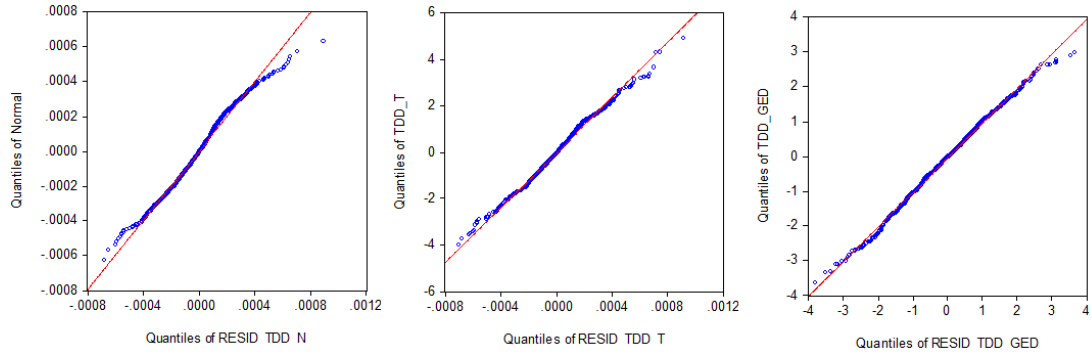


Obr. 3.1: Odhad modelu spotreby

Na obrázku 3.1 sú vykreslené odhadnuté hodnoty spolu s reziduami a graf odhadnutej podmienenej smerodajnej odchýlky. ARCH LM test s počtom oneskorení $q = 7$ s p-hodnotou 0,90 nezamieta nulovú hypotézu, a teda v odchýlkach nebola nájdená žiadna dodatočná ARCH štruktúra. Invertované korene AR polynómu s hodnotami 0,93 a 0,24 ležia vo vnútri jednotkového kruhu, čo stacionaritu nevyvracia. Korelogram štandardizovaných reziduí $\tilde{e}_t = \hat{e}_{t,1}/\hat{\sigma}_{t,1}$ je zobrazený v tabuľke 3.1. V korelograme sú všetky p-hodnoty Portmanteau testu vyššie ako 0,05,

| | ACF | PACF | Q-štat | p-hod | | ACF | PACF | Q-štat | p-hod |
|----|--------|--------|--------|-------|----|--------|--------|--------|-------|
| 1 | 0,010 | 0,010 | 0,157 | | 11 | -0,026 | -0,033 | 12,472 | 0,188 |
| 2 | -0,044 | -0,044 | 2,968 | | 12 | 0,001 | 0,000 | 12,475 | 0,254 |
| 3 | -0,012 | -0,011 | 3,173 | 0,075 | 13 | 0,011 | 0,010 | 12,653 | 0,317 |
| 4 | -0,005 | -0,007 | 3,210 | 0,201 | 14 | 0,048 | 0,043 | 16,124 | 0,186 |
| 5 | -0,030 | -0,031 | 4,567 | 0,206 | 15 | -0,025 | -0,027 | 17,035 | 0,198 |
| 6 | -0,051 | -0,051 | 8,411 | 0,078 | 16 | 0,006 | 0,011 | 17,098 | 0,251 |
| 7 | 0,028 | 0,026 | 9,558 | 0,089 | 17 | 0,056 | 0,051 | 21,683 | 0,116 |
| 8 | -0,020 | -0,026 | 10,123 | 0,120 | 18 | 0,001 | 0,002 | 21,683 | 0,154 |
| 9 | -0,029 | -0,028 | 11,390 | 0,122 | 19 | 0,012 | 0,020 | 21,892 | 0,189 |
| 10 | 0,007 | 0,004 | 11,455 | 0,177 | 20 | 0,080 | 0,084 | 31,423 | 0,026 |

Tabuľka 3.1: Korelogram štandardizovaných reziduí modelu spotreby



Obr. 3.2: QQ plot štandardizovaných reziduí modelu spotreby

okrem p-hodnoty pre lag 20, to je však už vysoký lag, čo značí nekorelovanosť štandardizovaných reziduí. Reziduá sme modelovali t-rozdelením, kde počet stupňov voľnosti bol odhadnutý na 3,99. Na Q-Q diagrame sme overili vhodnosť tohto rozdelenia (pozri obrázok 3.2). Na tomto obrázku je tiež Q-Q diagram ak by sme reziduá modelovali normálnym rozdelením, alebo GED rozdelením. Za najvhodnejšie rozdelenie z uvedených možností považujeme t-rozdelenie.

3.2 Model teploty

Teplota vysvetľuje najväčšiu časť spotreby (a ceny) plynu. V literatúre sa uvádza množstvo modelov teploty ako napríklad ARFIMA model s dlhou pamäťou Caballero a kol. (2002), Ornsteinov-Uhlenbeckov stochastický model Göncü a kol. (2013), aditívny model s deterministickou a náhodnou zložkou Stoll a Wiebauer (2010) atd. My sa v tejto práci zameriame na dekompozíciu časovej rady na systematické zložky a následné pridanie modelovej triedy ARMA-GARCH, podobne ako je uvedené v článku Stoll a Wiebauer (2010). Ročný sezónny charakter teploty modelujeme goniometrickými funkciami. Model teploty má tvar

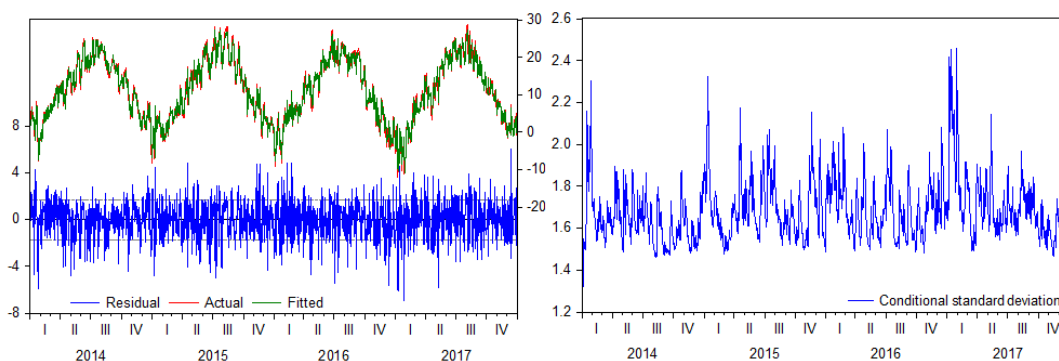
$$temp_t = b_0 + b_1 \cdot t + b_2 \cdot \sin(2\pi t/365,25) + b_3 \cdot \cos(2\pi t/365,25) + e_{t,2}, \quad (3.2)$$

kde $\{e_{t,2}\}$ je biely šum s konštantným konečným rozptylom $\sigma_2^2 > 0$. Model 3.2 sme odhadli OLS metódou, avšak pri verifikácii predpokladov lineárneho regresného modelu bolo zistené porušenie nekorelovanosti reziduí a heteroskedasticita. P-hodnoty Portmanteau testu boli 0,00 čo značí autokorelovanosť. Whiteov test s krížovými členmi s p-hodnotou 0,00 zamietá nulovú hypotézu o konštantnom rozptyle. Testy na jednotkový koreň potvrdzujú stacionaritu reziduí, a tak nie je nutné použiť diferencovanie. Reziduá tak budeme modelovať pomocou triedy ARMA-GARCH a to AR(3) - GARCH(1,1) modelom, ktorý bol najvhodnejší z hľadiska minimalizácie Akaikeho informačného kritéria a predpokladov na kladné parametre modelu rozptylu. Metódou maximálnej vierohodnosti

sme odhadli nasledovný model pre teplotu $temp_t$:

$$\begin{aligned}
 temp_t &= \underset{(0,00)}{12,119} - \underset{(0,03)}{0,001} \cdot t - \underset{(0,00)}{2,522} \cdot \sin(2\pi t/365,25) \\
 &\quad - \underset{(0,00)}{10,384} \cdot \cos(2\pi t/365,25) + e_{t,2}, \\
 e_{t,2} &= \underset{(0,00)}{1,111} \cdot e_{t-1,2} - \underset{(0,00)}{0,408} \cdot e_{t-2,2} + \underset{(0,00)}{0,109} \cdot e_{t-3,2} + u_{t,2}, \\
 u_{t,2} &= \varepsilon_{t,2} \cdot \sigma_{t,2}, \quad \varepsilon_{t,2} \overset{iid}{\sim} t_{15,131}, \\
 \sigma_{t,2}^2 &= \underset{(0,28)}{0,296} + \underset{(0,08)}{0,053} \cdot u_{t-1,2}^2 + \underset{(0,00)}{0,844} \cdot \sigma_{t-1,2}^2,
 \end{aligned}$$

kde v zátvorkách je p-hodnota testu významnosti parametrov.

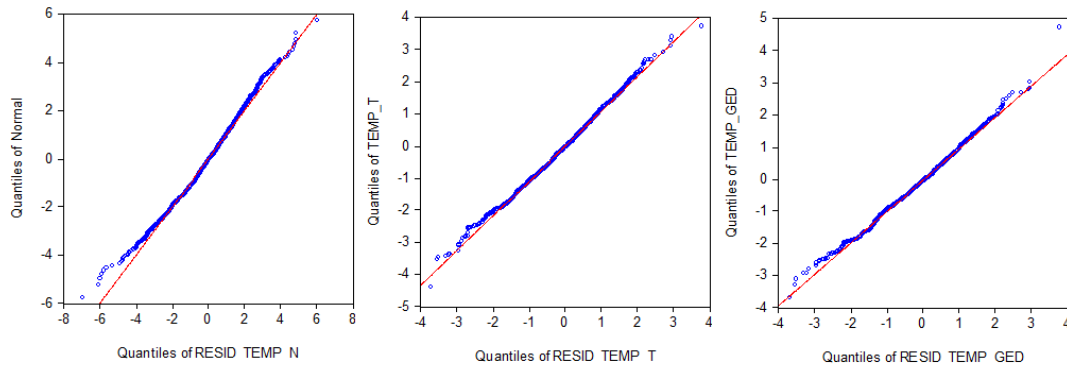


Obr. 3.3: Odhad modelu teploty

Na obrázku 3.3 sú reziduá spolu s odhadnutými a skutočnými hodnotami modelovanej premennej a na obrázku vpravo graf odhadnutej podmienenej smerodajnej odchýlky. ARCH LM test s p-hodnotou 0,76 a s počtom oneskorení 7, nezamieta nulovú hypotézu a teda v odchýlkach nebola týmto testom nájdená žiadna dodatočná ARCH štruktúra. V korelograme v tabuľke 3.2 sú všetky p-hodnoty Portmanteau testu vyššie ako 0,05 čo značí nekorelovanosť štandardizovaných reziduí a verifikovanie správnej rovnice volatility. Reziduá sme modelovali t-rozdelením, kde počet stupňov volnosti bol odhadnutý na 15,13. Na Q-Q diagrame sme overili adekvátnosť tohto rozdelenia (pozri obrázok 3.4). Na tomto

| | ACF | PACF | Q-štat | p-hod | | ACF | PACF | Q-štat | p-hod |
|----|--------|--------|--------|-------|----|--------|--------|--------|-------|
| 1 | 0,000 | 0,000 | 0,000 | | 11 | -0,017 | -0,015 | 6,822 | 0,556 |
| 2 | -0,005 | -0,005 | 0,034 | | 12 | -0,012 | -0,010 | 7,052 | 0,632 |
| 3 | -0,001 | -0,010 | 0,126 | | 13 | 0,029 | 0,029 | 8,311 | 0,599 |
| 4 | 0,036 | 0,036 | 2,113 | 0,146 | 14 | -0,015 | -0,015 | 8,655 | 0,654 |
| 5 | -0,001 | -0,001 | 2,115 | 0,347 | 15 | -0,025 | -0,026 | 9,552 | 0,655 |
| 6 | 0,003 | 0,003 | 2,129 | 0,546 | 16 | 0,035 | 0,035 | 11,408 | 0,577 |
| 7 | -0,036 | -0,035 | 4,043 | 0,400 | 17 | 0,016 | 0,011 | 11,776 | 0,624 |
| 8 | -0,034 | -0,035 | 5,711 | 0,335 | 18 | -0,008 | -0,010 | 11,876 | 0,688 |
| 9 | -0,006 | -0,006 | 5,759 | 0,451 | 19 | 0,040 | 0,041 | 14,253 | 0,580 |
| 10 | -0,021 | -0,022 | 6,398 | 0,494 | 20 | 0,039 | 0,038 | 16,521 | 0,487 |

Tabuľka 3.2: Korelogram štandardizovaných reziduí modelu teploty



Obr. 3.4: QQ plot štandardizovaných rezidií modelu teploty

obrázku je tiež Q-Q diagram, ak by sme reziduá modelovali normálnym a GED rozdelením.

3.3 Model spotovej ceny

Nasleduje model spotovej ceny indexu NCG, ktorý sme bližšie opísali v kapitole 2. K dispozícii máme celkom 1553 denných pozorovaní a to od 1.10.2013 do 31.12.2017. Model je však postavený na pozorovaniach až od 1.1.2014 vzhľadom na model *tdd*, kde máme pozorovania až od tohto dátumu.

Model spotovej ceny sa vyskytuje v literatúre veľmi často. Uvedieme niekoľko prístupov tohto modelovania. Klasický, jednofaktorový model ceny komodity uviedol Schwartz (1997), ktorý prirodzene rozšíril geometrický Brownov pohyb na proces s vracaním sa k určitej úrovni na tzv. mean-reverting process, keďže ceny komodít majú takúto vlastnosť. O niečo širší model priniesli Schwartz a Smith (2000), ktorí logaritmus spotovej ceny modelovali ako súčet dvoch náhodných faktorov a to krátkodobý a dlhodobý efekt. Keďže tieto dva efekty nie sú priamo pozorovateľné, na odhad faktorov použili techniku Kalmanovho filtra. Prvý, krátkodobý efekt je modelovaný ako Ornstein-Uhlenbeckov proces s vracaním sa k nulovej úrovni a dlhodobý efekt je modelovaný ako Brownov pohyb. Ďalšie práce nadväzovali na predchádzajúce, pričom sa snažili vylepšiť ich nedostatky pridaním tretieho faktoru do modelu, alebo pridaním sezónnych členov. Matematický model v práci Stoll a Wiebauer (2010) bol postavený na dekompozícii časovej rady na deterministické a náhodné zložky. V deterministickej zložke bol zahrnutý trend, týždenná, ročná sezónnosť a vplyv zásobníkov, zatiaľ čo v náhodnej zložke bol modelovaný krátkodobý efekt modelom AR(1) a dlhodobý efekt geometrickým Brownovým pohybom. Reakciou na tento model bola práca Müller a kol. (2015), kde bola do modelu navyše pridaná závislosť na cene ropy. V tejto práci budeme modelovať logaritmické výnosy, na čo použijeme modely z triedy ARMA-GARCH.

Ako bolo spomenuté vyššie, jednou z hlavných hnacích síl spotového trhu je dopyt, ktorý je nižší počas víkendu hlavne z dôvodu nižšej priemyselnej spotreby. Taktiež sa na burze počas víkendu neobchoduje. Významným faktorom je tak týždenná sezónnosť avšak nielen výskyt víkendu, ale jednotlivé dni týždňa.

Definujeme kategoriálnu premennú deň v týždni dow (day of the week) v tvare

$$dow_t = \begin{cases} 1, & \text{ak je daný deň pondelok,} \\ 2, & \text{ak je daný deň utorok,} \\ \vdots & \\ 6, & \text{ak je daný deň víkend, alebo sviatok.} \end{cases}$$

Kde jednotlivé sviatky odpovedajú sviatkom Veľkej Británie, na ktoré je naviazané obchodovanie na danej burze. ¹ Ďalšia zložka v modeli predstavuje závislosť ceny na spotrebe. Ak je spotreba vysoká v dôsledku nízkej teploty, prejaví sa to aj v cene. Významné sú odchýlenia od dlhodobého priemeru spotreby, kedy je na trhu veľký dopyt po plyne, čím stúpne jeho cena. Vzhľadom na prítomnosť jednotkového koreňa pri priamom modelovaní ceny, modelujeme logaritmické výnosy indexu NCG spočítané nasledovne

$$r_t = \log \left(\frac{P_t}{P_{t-1}} \right),$$

kde P_t je spotová cena indexu NCG. Model logaritmických výnosov má tvar

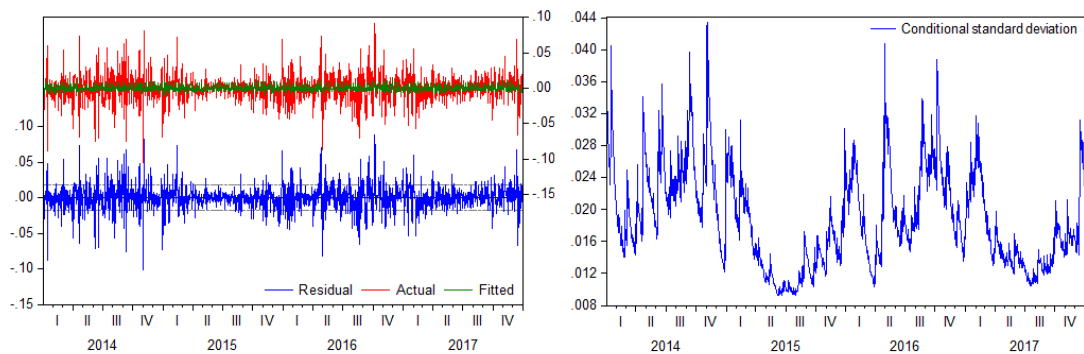
$$r_t = c_0 + c_1 \cdot \mathbb{I}\{dow_t = 2\} + c_2 \cdot \mathbb{I}\{dow_t = 3\} + c_3 \cdot \mathbb{I}\{dow_t = 4\} + c_4 \cdot \mathbb{I}\{dow_t = 5\} + c_5 \cdot \mathbb{I}\{dow_t = 6\} + c_6 \cdot tdd_t + c_7 \cdot tdd_{t-1} + e_{t,3} \quad (3.3)$$

kde $\{e_{t,3}\}$ je biely šum s konštantným konečným rozptylom $\sigma_3^2 > 0$. Tento model 3.3 sme odhadli opäť metódou najmenších štvorcov. Pri verifikácii predpokladov sme testovali reziduá na konštantný rozptyl, avšak Whiteov test s p-hodnotou 0,00 zamietá nulovú hypotézu homoskedasticity. Prikročili sme teda ku GARCH modelu. Rád modelu GARCH sme určili minimalizáciou Akaikeho informačného kritéria. Vyskúšaním rôznej voľby parametrov nám vyšiel ako vhodný model GARCH(1,2). Tento model dostatočne pokrýva heteroskedasticitu a dostaneme už nekorelované kvadráty štandardizovaných reziduí. ARCH-LM test s p-hodnotou 0,30 a s počtom oneskorení 7 nezamietá nulovú hypotézu, a teda v odchýlkach nebola nájdená dodatočná ARCH štruktúra.

Odhadnuté parametre modelu spolu s p-hodnotami testu významnosti uvádzame v nasledujúcej sústave rovníc:

$$\begin{aligned} r_t &= \underset{(0,00)}{0,008} - \underset{(0,00)}{0,006} \cdot \mathbb{I}\{dow_t = 2\} - \underset{(0,00)}{0,007} \cdot \mathbb{I}\{dow_t = 3\} - \underset{(0,00)}{0,007} \cdot \mathbb{I}\{dow_t = 4\} \\ &\quad - \underset{(0,00)}{0,009} \cdot \mathbb{I}\{dow_t = 5\} - \underset{(0,00)}{0,011} \cdot \mathbb{I}\{dow_t = 6\} + \underset{(0,00)}{6,172} \cdot tdd_t \\ &\quad - \underset{(0,00)}{6,273} \cdot tdd_{t-1} + e_{t,3}, \\ e_{t,3} &= \varepsilon_{t,3} \cdot \sigma_{t,3}, \quad \varepsilon_{t,3} \stackrel{iid}{\sim} t_{3,671}, \\ \sigma_{t,3}^2 &= \underset{(0,03)}{4,6 \times 10^{-6}} + \underset{(0,03)}{0,080} \cdot e_{t-1,3}^2 + \underset{(1,00)}{4,2 \times 10^{-8}} \cdot \sigma_{t-1,3}^2 + \underset{(0,00)}{0,919} \cdot \sigma_{t-2,3}^2. \end{aligned}$$

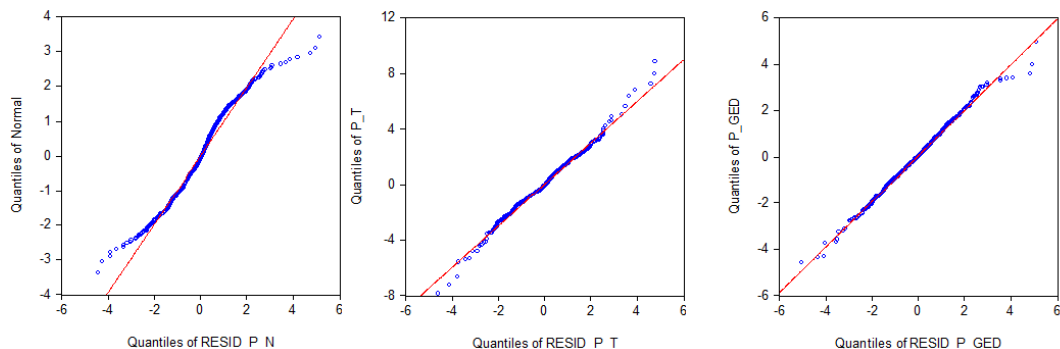
Ďalej prikladáme odhadnutý model v podobe grafu a vývoj odhadnutej smerodajnej odchýlky (pozri obrázok 3.5). V korelograme v tabuľke 3.3 vidíme p-hodnoty Portmanteau testu na štandardizované reziduá, ktoré sú vyššie ako 0,05.



Obr. 3.5: Odhad modelu logaritmickej výnosov

| | ACF | PACF | Q-štat | p-hod | | ACF | PACF | Q-štat | p-hod |
|----|--------|--------|--------|-------|----|--------|--------|--------|-------|
| 1 | 0,014 | 0,014 | 0,2883 | 0,591 | 11 | -0,011 | -0,009 | 7,9512 | 0,718 |
| 2 | -0,015 | -0,015 | 0,6006 | 0,741 | 12 | 0,005 | 0,002 | 7,9946 | 0,786 |
| 3 | 0,051 | 0,051 | 4,3783 | 0,223 | 13 | 0,011 | 0,013 | 8,1748 | 0,832 |
| 4 | 0,000 | -0,002 | 4,3783 | 0,357 | 14 | 0,094 | 0,096 | 21,151 | 0,098 |
| 5 | -0,028 | -0,026 | 5,4878 | 0,359 | 15 | -0,019 | -0,022 | 21,688 | 0,116 |
| 6 | -0,015 | -0,017 | 5,8283 | 0,443 | 16 | -0,032 | -0,032 | 23,169 | 0,109 |
| 7 | 0,005 | 0,005 | 5,8613 | 0,556 | 17 | 0,006 | -0,003 | 23,216 | 0,142 |
| 8 | -0,012 | -0,010 | 6,0662 | 0,640 | 18 | 0,013 | 0,014 | 23,465 | 0,173 |
| 9 | 0,027 | 0,029 | 7,1062 | 0,626 | 19 | -0,004 | 0,006 | 23,487 | 0,217 |
| 10 | -0,021 | -0,024 | 7,7789 | 0,650 | 20 | 0,013 | 0,016 | 23,746 | 0,254 |

Tabuľka 3.3: Korelogram štandardizovaných reziduí modelu výnosov



Obr. 3.6: QQ plot štandardizovaných reziduí modelu logaritmickej výnosov

Nekorelovanosť kvadrátu štandardizovaných reziduí nezamietame. Na obrázku 3.6 vidíme Q-Q diagramy na posúdenie zhody rozdelenia reziduí.

Odhadnuté koeficienty všetkých troch jednorozmerných modelov spolu so smerodajnými odchýlkami, t-štatistikou a p-hodnotami uvádzame v prílohe, časť A.2.

¹Prístupné na: <https://www.gov.uk/bank-holidays>, navštívené dňa 3.7.2018

4. Viacrozmerný model

Tri základné veličiny vystupujúce v tejto práci sme modelovali samostatne v predchádzajúcej kapitole. Pri odhade sme použili modely typu ARMA-GARCH a Studentovo t-rozdelenie reziduí. Zhrnutie jednorozmerných modelov uvádzame v tabuľke 4.1.

Vzhľadom na očakávané prepojenie jednotlivých premenných, kedy je spotreba výrazne ovplyvnená teplotou, ktorá následne pôsobí na cenu považujeme za vhodné prejsť k viacrozmernému modelovaniu. Nasleduje kapitola, ktorá sa venuje modelovaniu viacrozmerných časových rád, kde čerpáme z kníh Cípra (2008), Tsay (2014), Lütkepohl (2005). V tomto kontexte je tiež vhodné vyšetriť kointegráciu časových rád, ktorá je často prítomná vo finančných časových radoch. Kointegráciu chápeme ako vzťah určitej dlhodobej rovnováhy medzi ekonomickými veličinami, kedy jednotlivé časové rady sú nestacionárne, ale ich spoločný (kointegračný) pohyb v čase smeruje k určitému rovnovážnemu stavu. Prvý krok v konštrukcii mnohorozmerného modelu je testovanie jednotkového koreňa pre jednotlivé časové rady (pozri tabuľka 4.2). Podobne ako v jednorozmernom prípade, aj tu prejdeme k transformácií časovej rady spotovej ceny a budeme pracovať s logaritmickými výnosmi.

Na základe ADF testu je nulová hypotéza o jednotkovom koreni zamietnutá pre radu $\{temp_t\}$ a $\{r_t\}$, zatiaľ čo nulovú hypotézu o stacionarite časovej rady KPSS testu nezamietame pre žiadnu radu. Prvé diferencie veličín už sú stacionárne. Testy nepreukázali prítomnosť jednotkového koreňa, a tak nie je potrebné zavádzať EC (error correction) modely a prejdeme ku konštrukcii modelu VAR, respektíve modelu VAR s exogénnymi premennými. Kointegrácií je venovaná napríklad kapitola 12.5 v knihe Cípra (2008), odkiaľ čerpáme.

4.1 Odhad modelu VARX

V úvodnej kapitole sme uviedli všeobecný model vektorovej autoregresie, ktorý budeme aplikovať na uvedené vysvetľované premenné. Ako exogénne premenné budú do modelu vstupovať uvažované vysvetľujúce premenné použité v jednorozmernom prípade, okrem premennej $wh_t \cdot hdd_t$, ktorá by nebola exogénnou premennou, keďže hdd_t je odvodené od teploty. Taktiež z modelu vylúčime *trend*, ktorý nie je významný v modeli VARX. V modeli vystupujú goniometrické funkcie $\sin(2\pi t/365,25)$, $\cos(2\pi t/365,25)$ kedy vzhľadom na multikolinearitu nie je možné použiť ich oneskorené hodnoty. Model je odhadovaný metódou najmenších štvorcov, ktorá nedovoľuje mať lineárne závislé stĺpce. Je možné však použiť oneskorené hodnoty veličiny *dow*. Pred odhadnutím modelu tak definujeme nulové hodnoty pre oneskorené zložky \sin , \cos a pridáme iba oneskorené členy *dow*.

| | |
|----------|-------------------|
| tdd_t | AR(2)-GARCH(1, 1) |
| $temp_t$ | AR(3)-GARCH(1, 1) |
| r_t | GARCH(1, 2) |

Tabuľka 4.1: Jednorozmerné modely veličín

| Premenná | ADF test (p-hodnota) | | KPSS test (test. štatistika) | |
|----------|----------------------|---------------|------------------------------|---------------|
| | level | 1. diferencie | level | 1. diferencie |
| tdd_t | 0,0517 | 0,0000 | 0,0941 | 0,1145 |
| $temp_t$ | 0,0009 | 0,0000 | 0,0789 | 0,0791 |
| r_t | 0,0000 | 0,0000 | 0,1775 | 0,0131 |

Poznámka: Kritická hodnota KPSS testu na 5% hladine spoľahlivosti je 0,4630

Tabuľka 4.2: Testy stacionarity časovej rady

| Model | AIC | BIC |
|------------|---------|---------|
| VARX(3, 0) | -24,171 | -23,986 |
| VARX(3, 1) | -24,206 | -23,967 |
| VARX(3, 2) | -24,120 | -23,906 |

Tabuľka 4.3: Informačné kritériá modelov VARX

V tabuľke 4.3 zhrňujeme výsledky informačných kritérií pri rôznej voľbe oneskorenia parametrov. Parameter autoregresného rádu $p = 3$ volíme preto, že pre nižší rád je evidentne porušený predpoklad nekorelovanosti. Parameter s pre rád oneskorenia exogénnych premenných volíme na základe tabuľky 4.3.

V tabuľke 4.3 vidíme informačné kritériá odhadnutých modelov, kde je iba malý rozdiel medzi spozdovaním premenných. Podľa informačných kritérií sú najlepšie modely VARX(3, 0), alebo VARX(3, 1). Vzhľadom na lepšie výsledky verifikácie modelu sa prikloníme k širšiemu modelu so spozdovaním premenných. Budeme odhadovať model VARX(3, 1), ktorý môžeme zapísať takto

$$\mathbf{y}_t = \Phi_0 + \Phi_1 \mathbf{y}_{t-1} + \Phi_2 \mathbf{y}_{t-2} + \Phi_3 \mathbf{y}_{t-3} + \beta_0 \mathbf{x}_t + \beta_1 \mathbf{x}_{t-1} + \mathbf{a}_t,$$

kde

$$\mathbf{y}_t = \begin{pmatrix} tdd_t \\ temp_t \\ r_t \end{pmatrix}, \quad \mathbf{x}_t = \begin{pmatrix} \sin(2\pi t/365,25) \\ \cos(2\pi t/365,25) \\ \mathbb{I}\{dow_t = 2\} \\ \mathbb{I}\{dow_t = 3\} \\ \mathbb{I}\{dow_t = 4\} \\ \mathbb{I}\{dow_t = 5\} \\ \mathbb{I}\{dow_t = 6\} \end{pmatrix}$$

a \mathbf{a}_t je viacrozmerný biely šum, teda postupnosť nekorelovaných náhodných vektorov s nulovou strednou hodnotou a konštantnou, pozitívne definitnou rozptylovou maticou. Upozorňujeme, že sa jedná o model s apriórными obmedzeniami, kedy vo vektore \mathbf{x}_{t-1} neuvažujeme prvé dve zložky a preto v β_1 nastavujeme prvé dva koeficienty na 0.

V nasledujúcej časti uvedieme odhadnutý model VARX s diagnostikou a následne model MGARCH. Všetky modely boli spracované v softwari R, v knižniciach MTS (Tsay, 2015), vars (Pfaff a Stigler, 2013), mgarch (Ghalanos, 2015), rmgarchBEKK (Schmidbauer a kol., 2016) a súbežne v softwari EViews hlavne kvôli MGARCH modelom. Nápomocnou pri viacrozmernom modelovaní a hlavne pri knižnici MTS bola kniha Tsay (2014). V prílohe práce uvádzame okrem odhad-

nutých koeficientov tiež ich smerodajné odchýlky, t-štatistiky a p-hodnoty (A.3). Matrice nižšie zobrazujú odhadnuté koeficienty modelu VARX(3,1).

$$\begin{aligned}\Phi_0 &= \begin{pmatrix} 0,000 \\ 2,456 \\ 0,008 \end{pmatrix} & \Phi_1 &= \begin{pmatrix} 1,249 & 0,000 & 0,001 \\ -1169,363 & 0,995 & -5,245 \\ -0,666 & 0,000 & 0,049 \end{pmatrix} \\ \Phi_2 &= \begin{pmatrix} -0,401 & 0,000 & 0,001 \\ 1253,607 & -0,330 & -4,214 \\ -1,042 & 0,000 & -0,046 \end{pmatrix} & \Phi_3 &= \begin{pmatrix} 0,094 & 0,000 & 0,001 \\ -267,520 & 0,113 & -3,499 \\ 0,660 & 0,000 & 0,039 \end{pmatrix} \\ \beta_0 &= \begin{pmatrix} 0,000 & 0,000 & -0,000 & -0,000 & 0,000 & 0,000 & -0,000 \\ -0,372 & -1,895 & 0,436 & 2,102 & -1,376 & -0,594 & 0,066 \\ -0,001 & 0,001 & -0,003 & 0,005 & -0,009 & -0,006 & -0,009 \end{pmatrix} \\ \beta_1 &= \begin{pmatrix} 0,000 & 0,000 & 0,000 & 0,000 & 0,000 & 0,000 & 0,000 \\ 0,000 & 0,000 & -1,557 & 1,822 & 1,065 & 0,389 & 0,310 \\ 0,000 & 0,000 & -0,007 & 0,004 & 0,001 & -0,008 & 0,005 \end{pmatrix}\end{aligned}$$

Odhadnutá kovariančná a korelačná matica reziduí \mathbf{a}_t má tvar:

$$\Gamma_0 = \begin{pmatrix} 0,000 & 0,000 & 0,000 \\ 0,000 & 2,826 & -0,002 \\ 0,000 & -0,002 & 0,000 \end{pmatrix}, \quad \rho_0 = \begin{pmatrix} 1,000 & -0,665 & 0,083 \\ -0,665 & 1,000 & -0,063 \\ 0,083 & -0,063 & 1,000 \end{pmatrix}.$$

Korelačná matica chýb má vysoké korelácie, čo nás utvrdzuje v správnosti prechodu od jednorozmerných k viacrozmerným modelom. Korelácia medzi spotrebou a *tdd* je podľa očakávania záporná, a teda čím vyššia teplota, tým nižšia spotreba. Podobne, aj keď nie už tak významne, cena rastie so spotrebou a so stúpajúcou teplotou cena klesá. Môžeme si všimnúť, napríklad pri modeli teploty, že odhadnuté koeficienty pri *tdd*(-1), *tdd*(-2) sú podobné, majú však opačné znamienko. Toto, spoločne so vzájomným prepojením rovníc znemožňuje zistiť, aký vplyv má zmena danej premennej na budúce hodnoty iných premenných v systéme. Na zistenie tohto vplyvu sa používajú štatistiky ako block significance test, odozva na impulz, alebo rozklad rozptylu. Viacrozmerný ARCH LM test na kvadráty reziduí s p-hodnotou 0,00 zamieta konštantný rozptyl a teda by bolo vhodné prejsť ku GARCH modelu reziduí. Stacionaritu overujeme klasickým spôsobom. Prevrátené korene autoregresného polynómu, zoradené zostupne, sú:

$$\left(0,917 \quad 0,759 \quad 0,408 \quad 0,408 \quad 0,358 \quad 0,358 \quad 0,277 \quad 0,277 \quad 0,275 \right)^T.$$

Všetky korene sú vo vnútri jednotkového kruhu a odhadnutý model tým splňuje podmienku stacionarity a je stabilný.

Na obrázku 4.1 uvádzame graf významnosti autokorelácií a vzájomných korelácií reziduálnej zložky (cross-correlation matrix CCM), ktorú sme teoreticky opísali v časti 1.2 a vedľa vpravo p-hodnoty Portmanteau testu s Ljung-Boxovou testovou štatistikou. Tento test aplikujeme v prípade VARX(3, 1) modelu pre rád *m* vyšší ako 3. Všetky jeho p-hodnoty sú vyššie ako 0,05, čo značí nekorelovanosť reziduí. Číselne sú výsledky Portmanteau testu uvedené v prílohe práce (A.3). Reziduá, ktoré budeme neskôr modelovať pomocou MGARCH modelu sú vykreslené na obrázku 4.2.

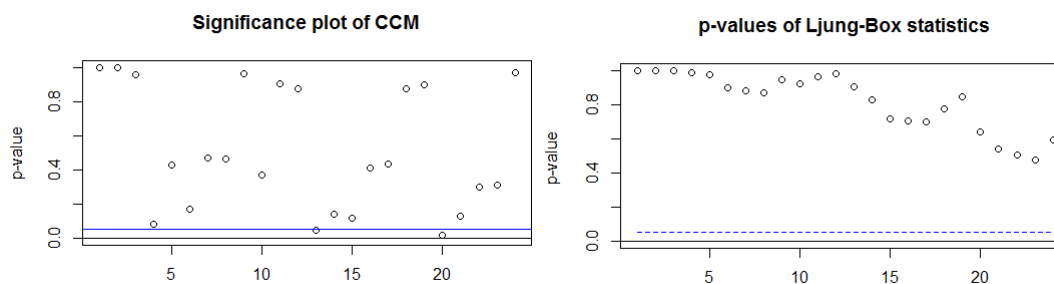
Ďalej pristúpime k testovaniu kauzality. Grangerov test kauzality hovorí o korelácií medzi súčasnými a minulými hodnotami vybraných premenných a to tak, že testuje nulovosť blokov určitých parametrov (viď kapitolu 1.2.1, Testovanie kauzality). Z tabuľky 4.4 je zrejme, že na 5% hladine významnosti:

- existuje obojsmerná závislosť (spätná väzba) medzi *tdd* a *temp*,
- existuje jednosmerná závislosť *tdd* na oneskorených logaritmickejých výnosoch, teda spozdené hodnoty *r* sú významné v rovnici pre *tdd*,
- existuje jednosmerná závislosť *temp* na oneskorených logaritmickejých výnosoch, teda spozdené hodnoty *r* sú významné v rovnici pre teplotu,
- na *tdd* kauzálne pôsobia obidve zostávajúce premenné, blok parametrov je významný v rovnici spotreby,
- na *temp* kauzálne pôsobia obidve zostávajúce premenné, blok parametrov je významný v rovnici teploty.

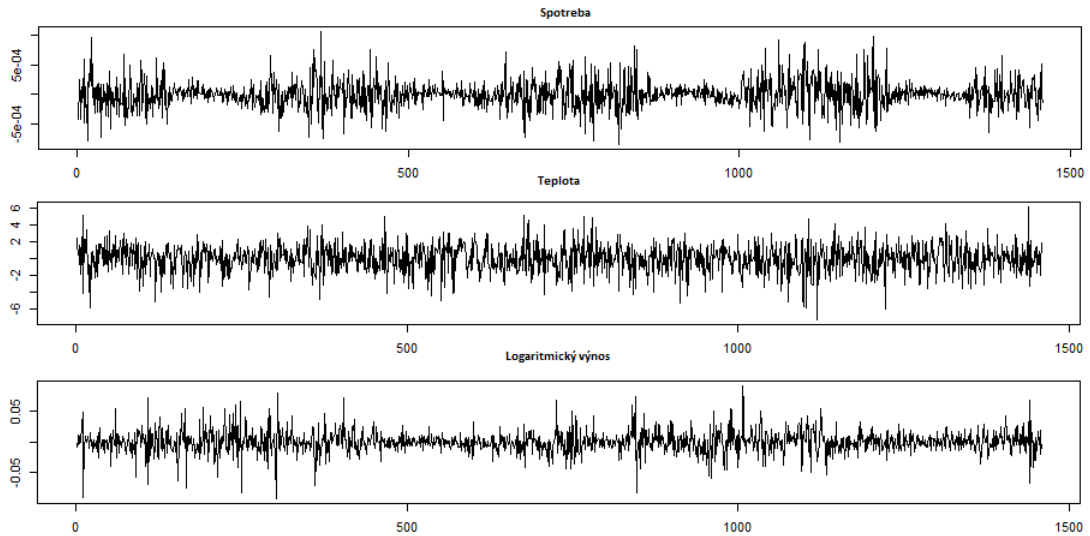
| <i>tdd</i> | | | <i>temp</i> | | | <i>r</i> | | |
|-------------|----------|-------|-------------|----------|-------|-------------|----------|-------|
| Vybraná | χ^2 | p-val | Vybraná | χ^2 | p-val | Vybraná | χ^2 | p-val |
| <i>temp</i> | 14,657 | 0,002 | <i>tdd</i> | 23,572 | 0,000 | <i>tdd</i> | 1,467 | 0,690 |
| <i>r</i> | 9,708 | 0,021 | <i>r</i> | 9,553 | 0,023 | <i>temp</i> | 0,381 | 0,944 |
| všetky | 24,406 | 0,004 | všetky | 34,795 | 0,000 | všetky | 2,520 | 0,866 |

Tabuľka 4.4: Grangerov test kauzality

Uvedené výsledky Grangerovho testu kauzality môžeme interpretovať tiež nasledovne. Obojsmerný kauzálny vzťah *tdd* a *temp* hovorí o tom, že spozdené informácie o teplote vysvetľujú súčasnú hodnotu spotreby, čo platí aj opačne. Túto informáciu sme očakávali vzhľadom na výraznú závislosť spotreby na teplote. Jednosmerná závislosť *temp* na *r* hovorí o tom, že pridanie oneskorených členov výnosov prispieva k vysvetleniu teploty. Analogicky aj ďalší výsledok testovania, ktorý vyplýva z obojsmernej závislosti *temp* a *tdd*. Podobne aj blok parametrov *temp* a *r* pôsobí kauzálne na *tdd*, čo vyplýva zo spojenia predchádzajúcich výsledkov testovania, obojsmernej závislosti *temp* a *tdd* a jednosmernej závislosti *tdd* na *r* a *temp* na *r*.



Obr. 4.1: Významnosť vzájomných korelácií a autokorelácií modelu VARX(3,1)



Obr. 4.2: Reziduá modelu VARX(3,1)

| | a_{1t} | a_{2t} | a_{3t} |
|-----------|----------|----------|----------|
| šíkmosť | 0,274 | -0,286 | 0,024 |
| špicatosť | 4,720 | 3,593 | 6,998 |

Tabuľka 4.5: Šíkmosť a špicatosť reziduí modelu VARX(3,1)

Na obrázku 4.3 je zobrazený korelogram pre viacrozmerný model, ktorý má tvar matice dimenzie $p \times p$ vzájomných autokorelácií. V tomto korelograme vidíme, že v niektorých prípadoch hodnoty autokorelácií prekračujú hranicu významnosti, čo by mohlo nasvedčovať autokorelovanosti, čo sa však nepotvrdilo v súhrnnom Portmanteau teste, ktorý nezamieta nulovú hypotézu o nekorelovanosti.

Ďalej bude nasledovať Jarque-Bera test normality. V tabuľke 4.5 uvádzame hodnoty šíkmosti a špicatosti jednotlivých reziduí. Viacrozmerný Jarque-Bera test s p -hodnotou 0,00 zamieta nulovú hypotézu mnohorozmerného normálneho rozdelenia.

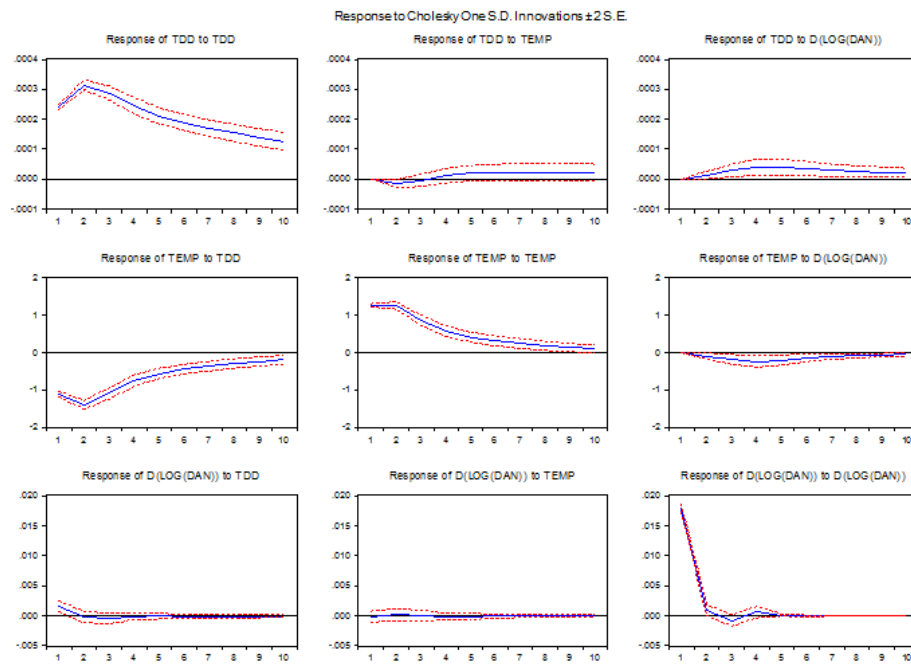
Posledným uvedeným diagnostickým nástrojom modelu VARX(3,1) je odozva na impulz. Na obrázku 4.4 pozorujeme, že vo všetkých troch rovniciach vplyv impulzu rýchlo vyšumí, čo hovorí o tom, že model je stacionárny. Impulz tdt v prvej rovnici však odznieva najdlhšie, čo je dôsledok koreňa vo výške 0,917, a preto trvá dlhšie, kým impulz vyšumí na nulovú úroveň.

4.2 Odhad modelu diagonálny BEKK

Po diagnostike modelu strednej hodnoty pre vektor \mathbf{y}_t uvedieme model podmienenej volatility. V teoretickej časti práce sme uviedli štyri rôzne prístupy MGARCH modelovania. Pre účely tejto práce sa zameriame na prvú kategóriu, ktorá je priamym zobecnením jednorozmerných GARCH modelov. V tejto triede máme k dispozícii variácie modelov typu VECM a BEKK, pričom software EViews nám ponúka odhad diagonálneho VECM, alebo diagonálneho BEKK mo-



Obr. 4.3: Korelogram reziduí modelu VARX(3,1)



Obr. 4.4: Odozva na impulz modelu VARX(3,1)

delu s dvomi typmi rozdelenia a to normálne a t-rozdelenie.

Vzhľadom na menší počet parametrov a záruku pozitívne definitnej matice \mathbf{H}_t sme sa rozhodli pre diagonálny BEKK model, kde je podľa informačných kritérií najlepší model BEKK(2,1) s mnohorozmerným Studentovým t-rozdelením chybovej zložky. Tento model zapíšeme v tvare

$$\begin{aligned} \mathbf{y}_t &= \boldsymbol{\Phi}_0 + \boldsymbol{\Phi}_1 \mathbf{y}_{t-1} + \boldsymbol{\Phi}_2 \mathbf{y}_{t-2} + \boldsymbol{\Phi}_3 \mathbf{y}_{t-3} + \boldsymbol{\beta}_0 \mathbf{x}_t + \boldsymbol{\beta}_1 \mathbf{x}_{t-1} + \mathbf{a}_t, \\ \mathbf{a}_t &= \mathbf{H}_t^{1/2} \mathbf{z}_t, \quad \mathbf{z}_t \sim mvt_{p,v}(\boldsymbol{\Sigma}_a), \\ \mathbf{H}_t &= \mathbf{C}\mathbf{C}^T + \mathbf{A}_1^T \mathbf{a}_{t-1} \mathbf{a}_{t-1}^T \mathbf{A}_1 + \mathbf{A}_2^T \mathbf{a}_{t-2} \mathbf{a}_{t-2}^T \mathbf{A}_2 + \mathbf{G}_1 \mathbf{H}_{t-1} \mathbf{G}_1^T. \end{aligned} \quad (4.1)$$

V tomto zápise označujeme symbolom $mvt_{p,v}(\boldsymbol{\Sigma}_a)$ p -rozmerné t-rozdelenie s počtom stupňov voľnosti v a $p \times p$ rozmernou, pozitívne semidefinitnou škálovou maticou $\boldsymbol{\Sigma}$. Keďže sa jedná o model diagonálny BEKK, matice \mathbf{C} , \mathbf{A}_1 , \mathbf{A}_2 , \mathbf{G}_1 sú diagonálne. Model odhadujeme v softwari EViews, kde je zavedené iné značenie pre rozklad matice \mathbf{H}_t . Matica \mathbf{M} označuje súčin matíc $\mathbf{C}\mathbf{C}^T$. V súlade s týmto značením tak uvedieme model pre maticu \mathbf{H}_t v tvare

$$\mathbf{H}_t = \mathbf{M} + \mathbf{A}_1^T \mathbf{a}_{t-1} \mathbf{a}_{t-1}^T \mathbf{A}_1 + \mathbf{A}_2^T \mathbf{a}_{t-2} \mathbf{a}_{t-2}^T \mathbf{A}_2 + \mathbf{G}_1 \mathbf{H}_{t-1} \mathbf{G}_1^T.$$

Odhadnuté koeficienty modelu sú:

$$\begin{aligned} \mathbf{C} &= \begin{pmatrix} 0,000 & 0,000 & 0,000 \\ 0,000 & 0,024 & 0,000 \\ 0,000 & 0,000 & 0,000 \end{pmatrix}, \quad \mathbf{A}_1 = \begin{pmatrix} 0,224 & 0,000 & 0,000 \\ 0,000 & 0,185 & 0,000 \\ 0,000 & 0,000 & 0,252 \end{pmatrix}, \\ \mathbf{A}_2 &= \begin{pmatrix} 0,211 & 0,000 & 0,000 \\ 0,000 & 0,200 & 0,000 \\ 0,000 & 0,000 & -0,127 \end{pmatrix}, \quad \mathbf{G}_1 = \begin{pmatrix} 0,951 & 0,000 & 0,000 \\ 0,000 & 0,961 & 0,000 \\ 0,000 & 0,000 & 0,949 \end{pmatrix}. \end{aligned}$$

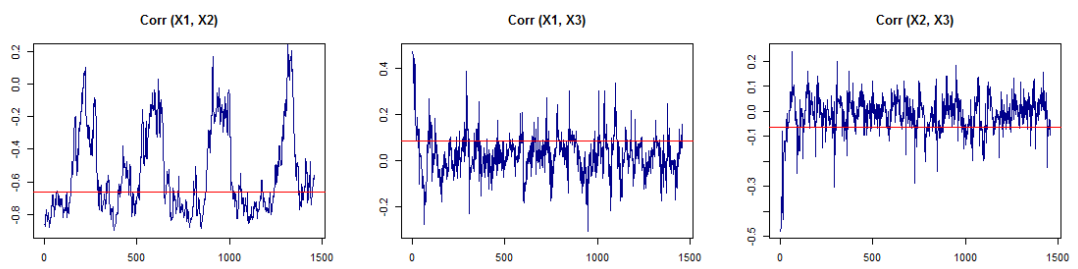
Stupeň voľnosti mnohorozmerného t-rozdelenia bol odhadnutý na hodnotu 8,63. Aby sme zistili, či je model kovariančne stacionárny, musíme overiť podmienku na vlastné čísla matice \mathbf{K} , ktorá vznikne súčtom jednotlivých Kroneckerových súčinov matíc \mathbf{A}_1 , \mathbf{A}_2 , \mathbf{G}_1 :

$$\mathbf{K} = \mathbf{A}_1 \otimes \mathbf{A}_1 + \mathbf{A}_2 \otimes \mathbf{A}_2 + \mathbf{G}_1 \otimes \mathbf{G}_1$$

Vlastné čísla matice \mathbf{K} , ktoré získame vyriešením rovnice $|\mathbf{K} - \lambda \mathbf{I}| = 0$ sú:

$$\left(0,999 \quad 0,997 \quad 0,997 \quad 0,997 \quad 0,980 \quad 0,933 \quad 0,933 \quad 0,932 \quad 0,932 \right)^T.$$

Všetky vlastné čísla ležia vo vnútri jednotkového kruhu a teda podmienka je splnená, model je kovariančne stacionárny. Robustná verzia viacrozmerného ARCH testu na kvadráty reziduí s p-hodnotou 0,268 nezamieta nulovú hypotézu konštantného rozptylu a teda modelom dostatočne pokrývame heteroskedasticitu. Robustnú verziu testu sme uviedli v kapitole 1.2.1, časť Mnohorozmerný Portmanteau test, uvedené tiež v Tsay (2014). Pre doplnenie ešte prikkladáme obrázky 4.5 podmienených korelácií spolu s nepodmienenou koreláciou, ktorú je možné číselne vidieť v matici $\boldsymbol{\rho}_0$. Z týchto grafov usudzujeme, že vzájomná korelácia je premenlivá, a preto je dobré pristupovať k reziduám modelom MGARCH. Nielen kvôli výraznej korelácií reziduí prvej a druhej rovnice v zimnom období, ale aj kvôli premenlivým koreláciám medzi reziduami 1 a 3 rovnice ako aj 2 a 3.



Obr. 4.5: Grafy podmienených korelácií modelu BEKK(2,1) spolu s konštantnou nepodmienenou koreláciou

5. Analýza scenárov

Odhadnuté a verifikované modely z predchádzajúcich kapitol použijeme pri simulácii hodnôt na nasledujúce obdobie. Exogénne premenné sú dané deterministicky, náhodnú zložku budeme generovať pomocou príslušnej rovnice. Porovnávať budeme celkom 4 simulácie, kde v jednorozmernom a viacrozmernom modeli použijeme na generovanie reziduí dva spôsoby. Prvý spôsob je klasické generovanie náhodnej zložky z určeného rozdelenia, v tomto prípade t-rozdelenia a druhý spôsob bude generovanie reziduí metódou bootstrap. Predpokladáme, že modelované dáta sa vyvíjajú spôsobom navrhnutým v kapitole 3 a 4. Skúmať budeme obdobie nasledujúceho roku a keďže sa jedná o denné dáta, generujeme 365 hodnôt.

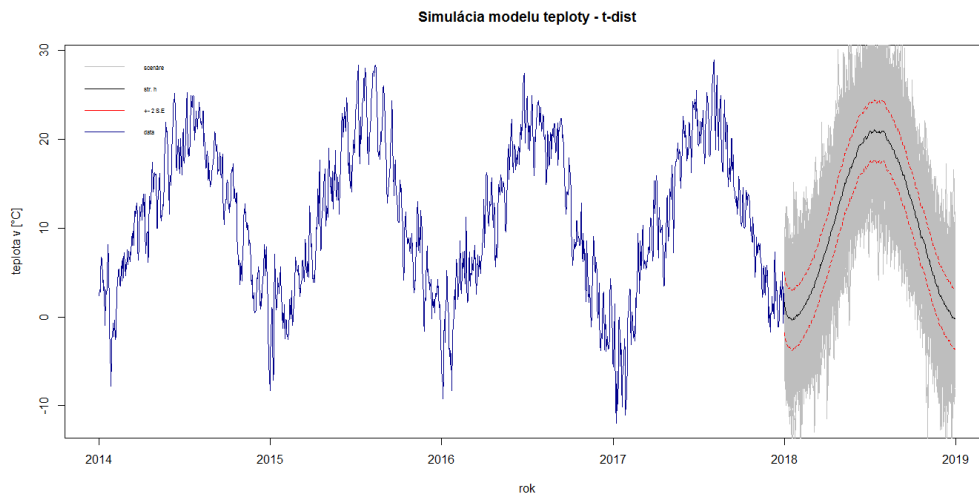
5.1 Jednorozmerné modely

Kapitola 3 obsahuje jednorozmerné modely spotreby, teploty a logaritmických výnosov spotovej ceny. Pri modelovaní jednotlivých rovníc potrebujeme ako prvú simulovať teplotu, ktorá je následne použitá pri simulácii spotreby. Ako poslednú radu generujeme výnosy, ktorých vstupom je spotreba. Generujeme 500 scenárov v softvare R a to v knižnici rugarch (Ghalanos, 2018). Priložený kód obsahuje podrobnosti zvolených parametrov pri odhade koeficientov modelu a následnej simulácii.

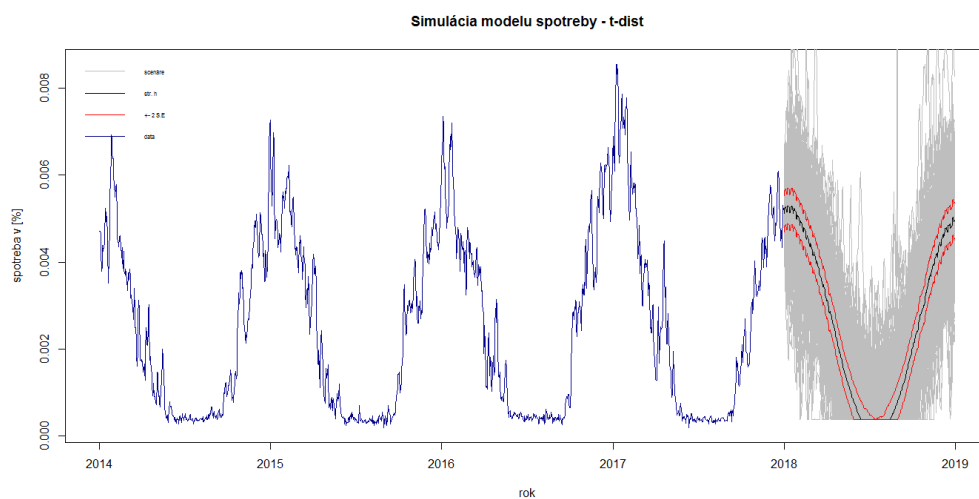
Proces simulácie hodnôt teploty (*temp*) popisujú odhadnuté rovnice z kapitoly 3.2, kde $t = n + 1, \dots, n + 365$. Pri generovaní použijeme ako štartovacie hodnoty posledné odhadnuté hodnoty reziduí $u_{n,2}$ a posledný člen GARCH modelu $\sigma_{n,2}$. Náhodná zložka $\varepsilon_{t,2}$ je generovaná zo Studentovho t-rozdelenia so 14,83 stupňami voľnosti. Na obrázku 5.1 sú vykreslené namerané hodnoty teploty z minulých období, ktoré sme mali k dispozícii pri odhade modelu a 500 scenárov budúceho vývoja. Zvýraznená je priemerná hodnota scenárov a odchýlka od priemernej hodnoty vo výške ± 2 krát smerodajná odchýlka.

Ďalšou premennou, ktorú sme generovali je spotreba *tdd*. V modeli spotreby plynu (pozri podkapitolu 3.1) vystupuje teplota, ktorú sme v predošlej časti nasimulovali. V súvislosti s premennou *wh·hdd* sme nasimulovanú teplotu ešte upravili podľa vzorca na $hdd_t = \max(17 - temp_t, 0)$. Analogickým postupom ako v modeli teploty sme generovali 500 scenárov pre obdobie 365 dní. Náhodná zložka $\varepsilon_{t,1}$ bola v tomto prípade generovaná zo Studentovho t-rozdelenia so 4 stupňami voľnosti a pri generovaní sme použili štartovacie hodnoty $u_{n,1}$ a $\sigma_{n,1}$. Vzhľadom na to, že *tdd* je obmedzené zdola hodnotou 0, bolo nutné nasimulované hodnoty v letnom období orezať. Za minimálnu výšku spotreby sme zvolili najnižšiu nameranú hodnotu za posledné 4 roky. Tohto nedostatku modelu sme si vedomí, ale vzhľadom na to, že spotreba je v letných mesiacoch konštantná, model sme neupravovali. Namerané hodnoty *tdd* spolu so simuláciami uvádzame na obrázku 5.2.

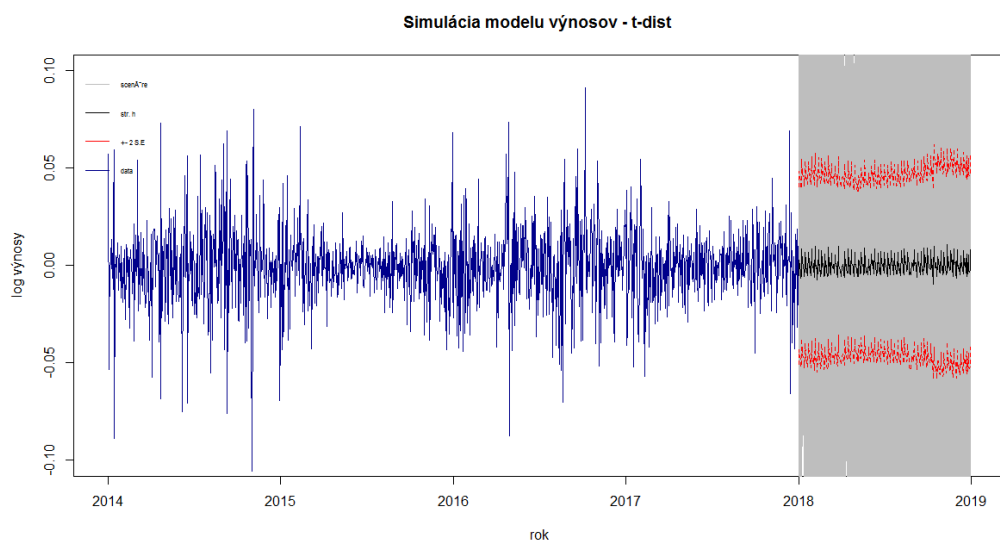
Logaritmické výnosy spotovej ceny *r* budeme generovať na základe jednorozmerného modelu uvedenom v podkapitole 3.3. Vstupujúci regresor *tdd* máme nagenovaný z predošlého kroku. Náhodnú zložku $\varepsilon_{t,3}$ generujeme zo Studentovho t-rozdelenia s 3,7 stupňami voľnosti a postupujeme podobne ako v predošlých dvoch modeloch. Historicky namerané výnosy a simulácia sú zobrazené na obrázku 5.3.



Obr. 5.1: Namerané hodnoty teploty a simulácie jednorozmerného modelu



Obr. 5.2: Namerané hodnoty spotreby a simulácie jednorozmerného modelu



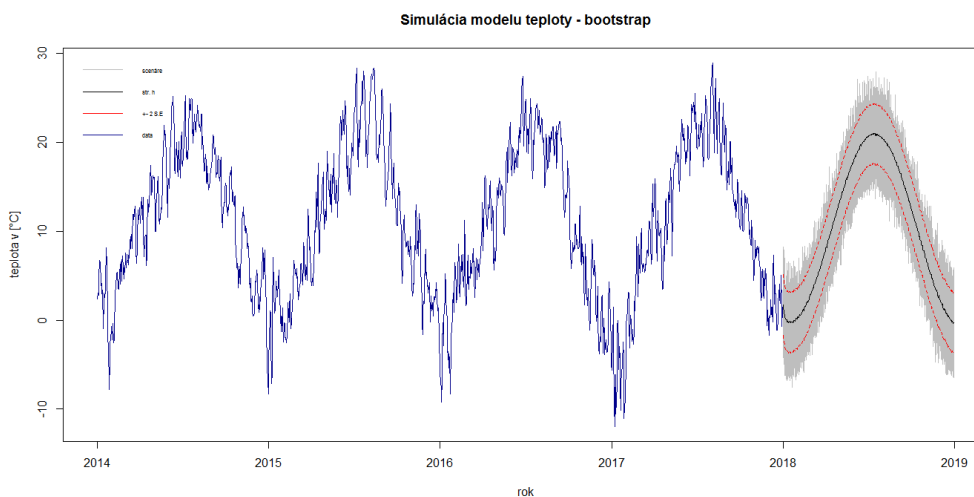
Obr. 5.3: Namerané hodnoty výnosov a simulácie jednorozmerného modelu

5.2 Jednorozmerný model - bootstrap

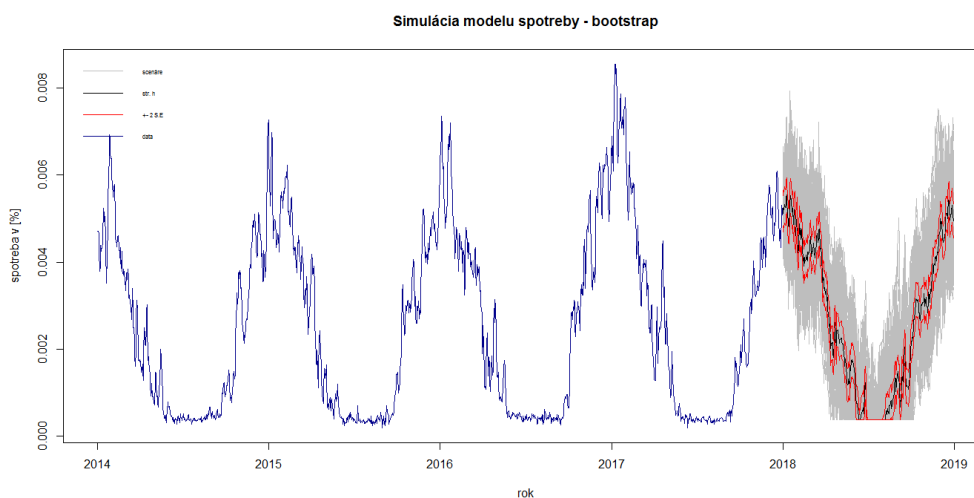
Reziduá jednorozmerných modelov sme navrhli modelovať pomocou modelov podmienenej heteroskedasticity. V tomto modeli je ale dôležité tiež určiť rozdelenie reziduí. Software nám dáva na výber tri možnosti a to normálne, t-rozdelenie a GED rozdelenie. Na obrázkoch 3.2, 3.4 a 3.6 vidíme porovnanie týchto možností pre dané 3 modely. Za najvhodnejšie považujeme Studentovo t-rozdelenie, avšak ani toto rozdelenie nie je pre dáta ideálne. Preto navrhujeme na reziduá použiť metódu reziduálny bootstrap. Bližšie k metóde reziduálny bootstrap je možné nájsť v článku Efron (1979), v ktorom bola metóda prvý krát publikovaná.

Pri simulácii metódou bootstrap použijeme hodnoty $\varepsilon_{i,1}, \varepsilon_{i,2}, \varepsilon_{i,3}, i = n + 1, \dots, n + 365$, ktoré negenerujeme z t-rozdelenia s odhadnutým počtom stupňov voľnosti, ako v predošlom prípade, ale robíme náhodný výber s vracaním z odhadnutých hodnôt $\hat{\varepsilon}_{i,1}, \hat{\varepsilon}_{i,2}, \hat{\varepsilon}_{i,3}, i = 1, \dots, 1461$, ktoré sme následne ešte centrovali.

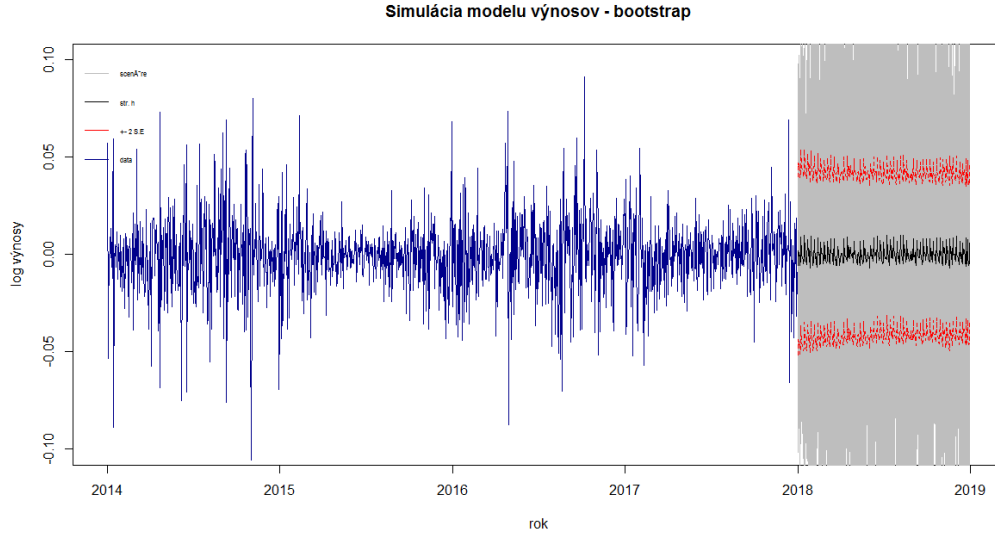
Obrázky 5.4, 5.5 a 5.6 zhrňujú namerané hodnoty, ktoré boli použité pri odhade modelu, a simulácie spolu s priemernou hodnotou.



Obr. 5.4: Namerané hodnoty teploty a simulácie modelu - bootstrap



Obr. 5.5: Namerané hodnoty spotreby a simulácie modelu - bootstrap



Obr. 5.6: Namerané hodnoty výnosov a simulácie modelu - bootstrap

Pri simulácii spotreby vidíme, že oproti simulácii z t-rozdelenia je krivka priemerných hodnôt veľmi členitá čo je spôsobené práve reziduálnou bootstrapovou metódou. Ak porovnáme predchádzajúce roky spotreby, vidíme, že podobné skoky sa vyskytujú pomerne často. V letnom období sme simulované hodnoty \widehat{td}_t opäť upravili a obmedzili zdola minimálnou nameranou hodnotou za posledné 4 roky pozorovaní.

5.3 Viacrozmerný model

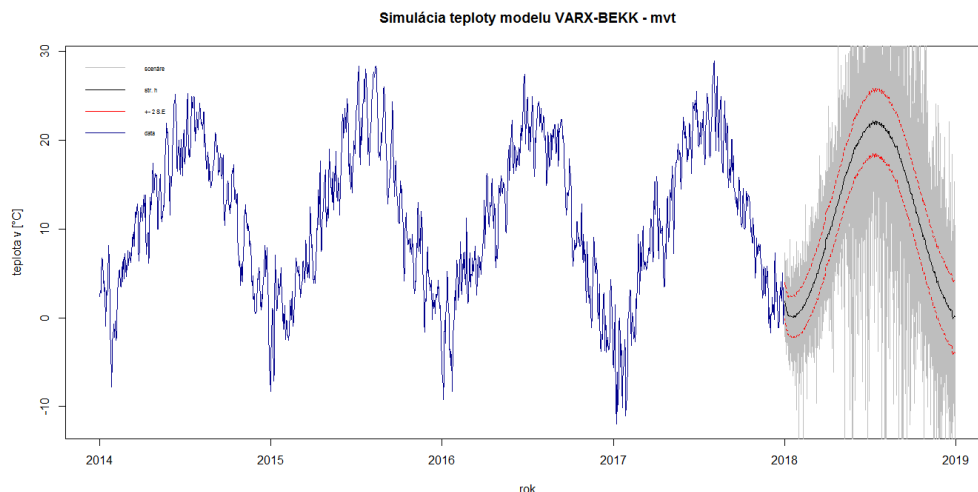
Generovanie scenárov v navrhnutom modeli VARX(3,1)-BEKK(2,1) prebieha analogicky ako v jednorozmernom prípade. Generujeme hodnoty $\hat{\mathbf{y}}_{n+i}(n)$ podľa odhadnutého modelu VARX(3,1)-BEKK(2,1) (viď kap. 4.1 a 4.2):

$$\hat{\mathbf{y}}_{n+i}(n) = \hat{\mathbf{\Phi}}_0 + \hat{\mathbf{\Phi}}_1 \hat{\mathbf{y}}_{n+i-1}(n) + \hat{\mathbf{\Phi}}_2 \hat{\mathbf{y}}_{n+i-2}(n) + \hat{\mathbf{\Phi}}_3 \hat{\mathbf{y}}_{n+i-3}(n) + \hat{\beta}_0 \mathbf{x}_{n+i} + \hat{\beta}_1 \mathbf{x}_{n+i-1} + \hat{\mathbf{a}}_{n+i}(n),$$

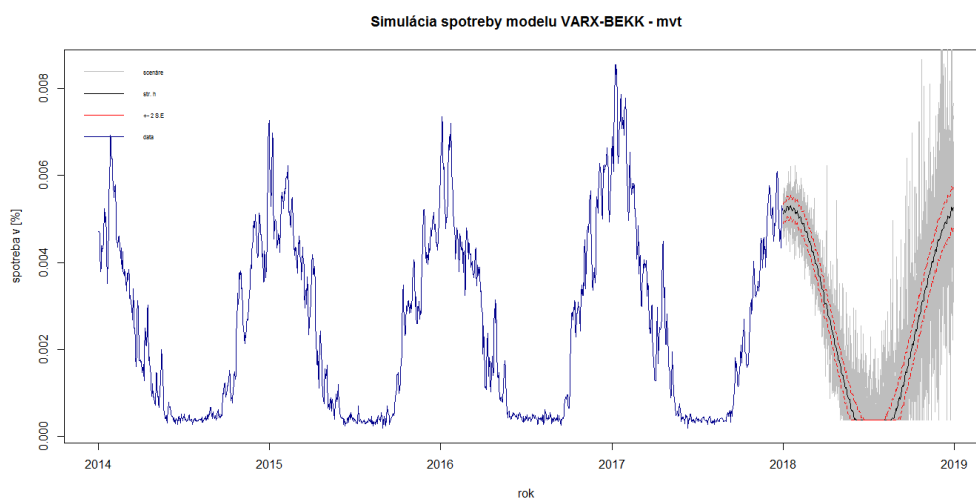
kde $i = 1, \dots, 365$ a rovnica pre $\hat{\mathbf{a}}_{n+i}$ je uvedená v kapitole 4.2, vzorec 4.1. Na generovanie zložky $\hat{\mathbf{a}}_{n+i}(n)$ modelom BEKK(2,1) použijeme posledné odhadnuté hodnoty $\hat{\mathbf{a}}_n$, $\hat{\mathbf{a}}_{n-1}$, a maticu \mathbf{H}_n . Náhodnú zložku \mathbf{z}_{n+i} generujeme z viacrozmerného Studentovho t-rozdelenia s počtom stupňov voľnosti 8,63. Takto dostaneme 500 scenárov vývoja $\hat{\mathbf{y}}_{n+i}$. Odhad viacrozmerného GARCH modelu prebiehal v softwari EViews. V tomto softwari je možné odhadovať modely diagonálny VECM a diagonálny BEKK s chybovou zložkou z mnohorozmerného normálneho, alebo t-rozdelenia, zatiaľ čo v Rku je možné odhadovať iba BEKK model a to iba s náhodnou zložkou z mnohorozmerného normálneho rozdelenia. Následná simulácia bola vykonaná v softwari R, kde sme prispôbili naprogramovanú funkciu *simulateBEKK*¹ z knižnice mgarchBEKK na t-rozdelenie. Upravenú časť kódu, ako aj celý proces odhadu a simulácie modelu prikladáme v prílohe práce.

¹Prístupné na: <https://github.com/cran/mgarchBEKK/blob/master/R/BEKK.R>, navštívené dňa 3.7.2018

Podobne ako v jednorozmernom prípade, aj teraz sme spotrebu v letných mesiacoch orezali na minimálnu historicky nameranú za posledné 4 roky. Nasledujúce obrázky 5.7, 5.8 a 5.9 zhrňujú výsledky simulácie jednotlivých premenných. Opäť uvádzame namerané hodnoty, priemernú simulovanú hodnotu a odchýlku od priemeru vo výške ± 2 krát smerodajná odchýlka.



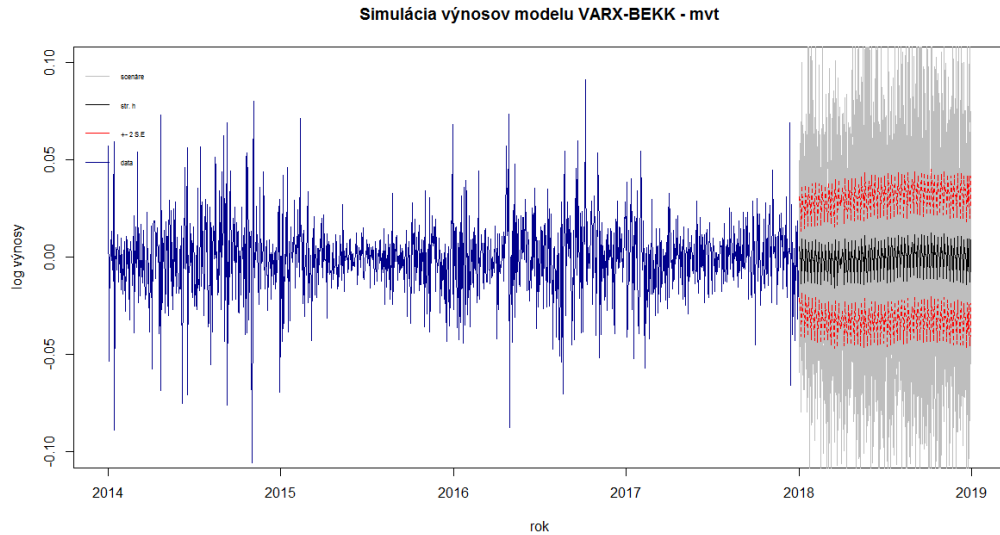
Obr. 5.7: Namerané hodnoty teploty a simulácie modelu VARX-BEKK



Obr. 5.8: Namerané hodnoty spotreby a simulácie modelu VARX-BEKK

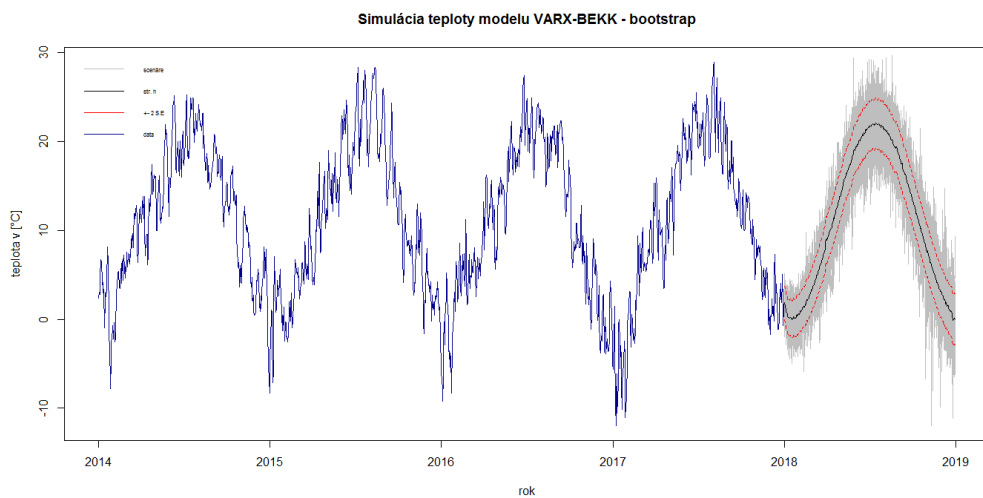
5.4 Viacrozmerný model - bootstrap

Rozdelenie zložky z_t môžeme s ohľadom na daný software (EViews) vo viacrozmernom prípade odhadovať pomocou mnohorozmerného normálneho rozdelenia, alebo viacrozmerným t-rozdelením. Vhodnejšie sa javilo t-rozdelenie, ktoré sme aj použili v predošlej simulácii. Ani to však nemusí byť správny typ rozdelenia, preto pri generovaní reziduí modelu VARX-BEKK opäť použijeme metódu reziduálny bootstrap, podobne ako v jednorozmernom prípade. Hodnoty z_{n+i} , $i = 1, \dots, 365$ vyberáme náhodne s vracaním z vypočítaných centrovanych

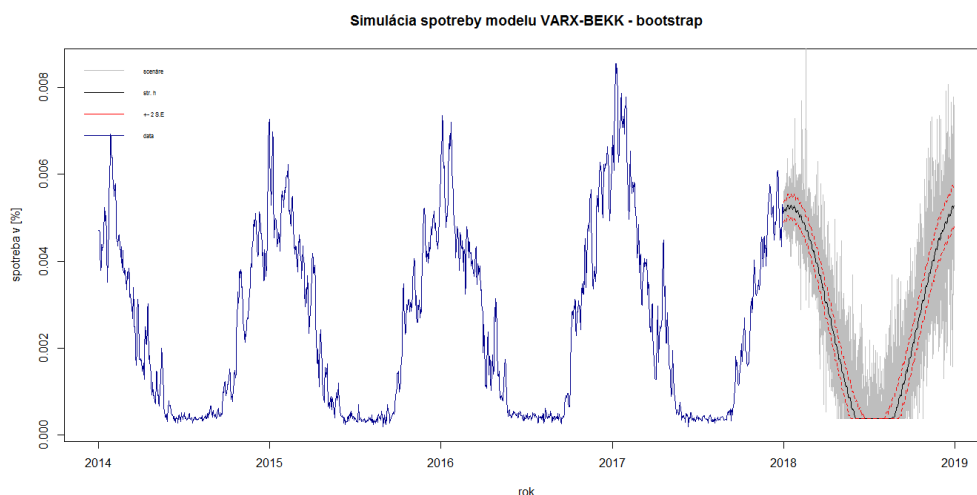


Obr. 5.9: Namerané hodnoty výnosov a simulácie modelu VARX-BEKK

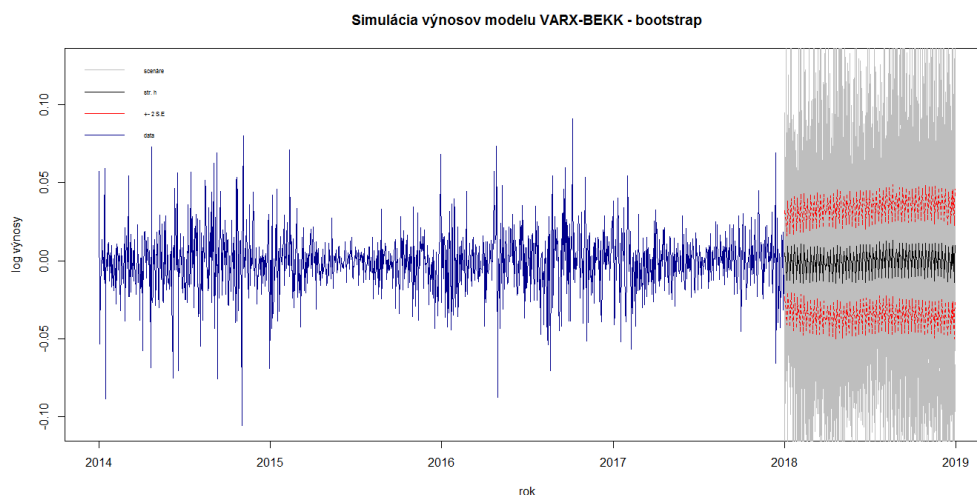
reziduí \hat{z}_i , $i = 1, \dots, 1461$, Generujeme 500 scenárov vývoja \hat{y}_{n+i} , $i = 1, \dots, 365$ na základe odhadnutého modelu VARX(3,1)-BEKK(2,1). Na obrázkoch 5.10, 5.11 a 5.12 sú vykreslené namerané hodnoty teploty, spotreby a logaritmickej výnosov z minulých období, ktoré sme mali k dispozícii pri odhade modelu a scenáre budúceho vývoja.



Obr. 5.10: Hodnoty teploty a simulácie modelu VARX-BEKK - bootstrap



Obr. 5.11: Hodnoty spotreby a simulácie modelu VARX-BEKK - bootstrap



Obr. 5.12: Hodnoty výnosov a simulácie modelu VARX-BEKK - bootstrap

5.5 Porovnanie výsledkov

Výsledné hodnoty simulácií porovnáme v nasledujúcej časti. Pri jednotlivých obrázkoch simulácie výnosov sme si mohli všimnúť odlišný tvar simulácií jednorozmerného a viacrozmerného modelu pri výnosoch, kedy vo viacrozmernom prípade máme oveľa väčšie medzidenné rozdiely, čo je spôsobené premennou *dow*.

K porovnaniu prístupu jednorozmerného a viacrozmerného modelovania použijeme miery presnosti predpovedí a to strednú štvorcovú chybu (MSE) a strednú absolútnu chybu (MAE), ktoré sme spočítali podľa vzťahov

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_t - \hat{y}_t)^2, \quad MAE = \frac{1}{n} \sum_{i=1}^n |y_t - \hat{y}_t|,$$

na dáta z obdobia rokov 2014-2017 na ktorých sme modely odhadovali. V tabuľke 5.1 vidíme, že viacrozmerný model má lepšiu presnosť premennej *temp* a výnosy *r*, zatiaľ čo jednorozmerný model má menšiu chybu v modelovaní *tdd*.

Zaujímavé bude tiež pozrieť sa na prierez nameraných hodnôt za nejaké obdobie. My vyberáme prvých 31 generovaných pozorovaní a teda mesiac január, ktorý

| | Jednorozmerný model | | | Viacrozmerný model | | |
|-----|---------------------|----------------------|----------------------|--------------------|----------------------|----------------------|
| | temp | tdd | r | temp | tdd | r |
| MSE | 2,86 | $3,3 \times 10^{-8}$ | $3,3 \times 10^{-4}$ | 2,78 | $5,7 \times 10^{-8}$ | $3,2 \times 10^{-4}$ |
| MAE | 1,32 | $1,4 \times 10^{-4}$ | $1,3 \times 10^{-2}$ | 1,31 | $1,7 \times 10^{-4}$ | $1,2 \times 10^{-2}$ |

Tabuľka 5.1: Porovnanie modelov podľa presnosti predpovedí

sa považuje za najchladnejší z celého roka. Z týchto približne 1500 pozorovaní pre každú premennú urobíme histogram a vykresľujeme ho s jadrovým odhadom hustoty. Na obrázku 6.3 sú histogramy zo všetkých simulovaných modelov.

Na nich môžeme vidieť, že nasimulovaná denná teplota sa v januári pohybuje v jednotlivých modeloch v rôznych rozmedziach. Najširšie pole hodnôt pozorujeme v prvom modeli, zatiaľ čo najužší pás je v modeli VARX s reziduami z t-rozdelenia. Hodnoty *tdd* sú rozmiestnené v užšom páse ako teploty. Najširší je pás *tdd* v simulácii jednorozmerného modelu s t-rozdelením reziduí. Logaritmičné výnosy sa pohybujú v najužšom páse z pozorovaných troch veličín.

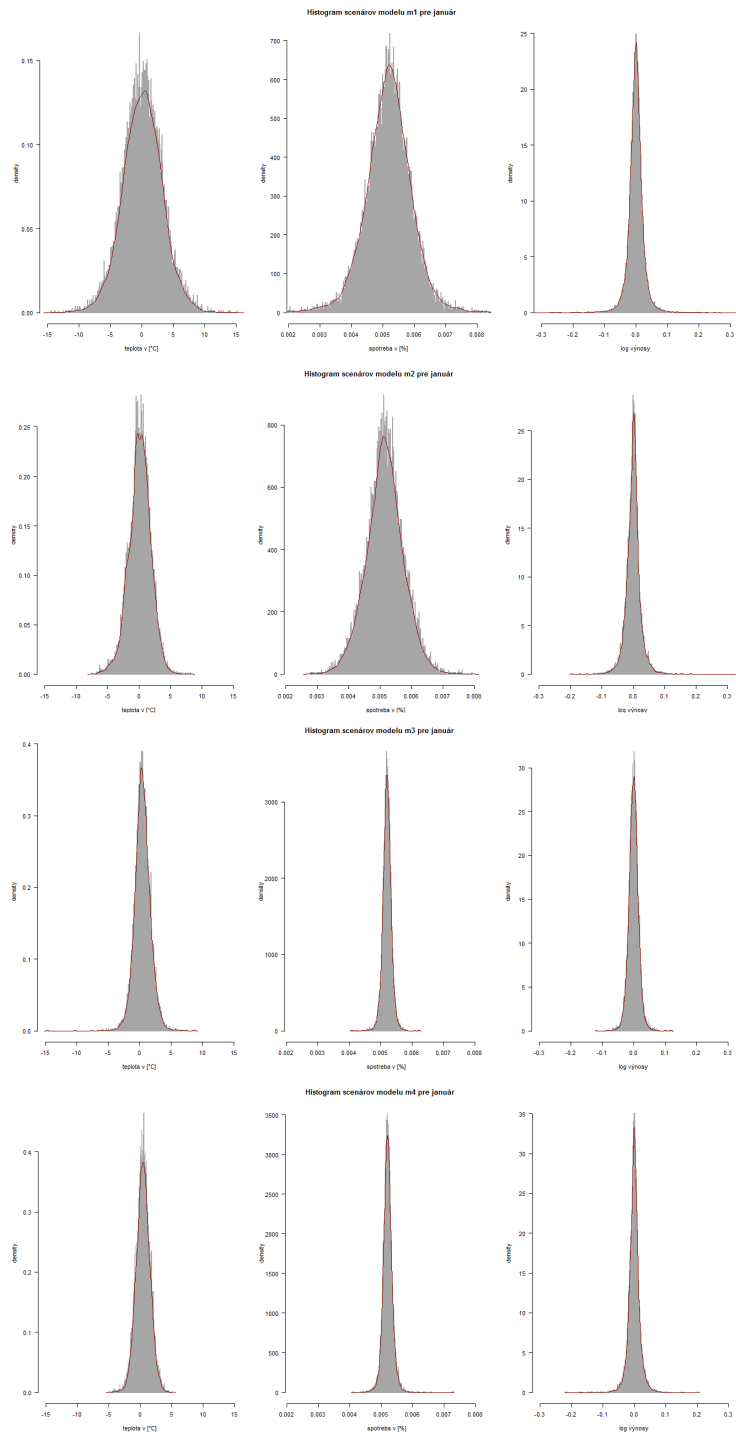
Modely pre lepšiu prehľadnosť označíme. Jednorozmerný model s generovaním reziduí z t-rozdelenia označíme m1, jednorozmerný model s reziduálnym bootstrapom označíme m2, model VARX-BEKK s generovaním z mnohorozmerného t-rozdelenia označíme m3 a posledný VARX-BEKK model s reziduálnym bootstrapom označíme m4.

V tabuľke 5.2 uvádzame hodnoty ktoré nám približujú grafické zobrazenie januárových hodnôt simulácií a to 5. a 95. percentil (ozn. P5 a P95). Tieto hodnoty hovoria o tom, že 5 % pozorovaní sa nachádza pod, respektíve nad danou hodnotou. Hodnoty P5 a P95 sa v jednotlivých modeloch značne líšia. Jednorozmerné modely majú viac ako dvojnásobné väčšie rozmedzie medzi týmito hodnotami.

Histogramy pre ďalšie vybrané mesiace, konkrétne pre mesiace koncov kvartálov, marec, jún, september, december, uvádzame v prílohách, kapitola A.4. Zo všetkých štyroch modelov budeme ďalej porovnávať priemerné hodnoty simulácií za dané mesiace v roku. Sústreďme sa preto na hodnoty „čiernej farby“ obrázkov 5.1 až 5.12. V tabuľke 5.3 uvádzame tri bloky priemerných hodnôt. Prvé informácie sú o priemernej teplote za daný mesiac. Počítame jednoduchý aritmetický priemer scenárov teploty za dané obdobie. Druhý stĺpec sú priemerné denné *tdd* v mesiaci. Teda percentuálna časť celkovej ročnej spotreby, ktorá má byť spotrebovaná priemerne za 1 deň v daný mesiac. Tretí stĺpec ukazuje priemernú spotovú cenu prepočítanú z priemerných simulovaných logaritmičných výnosov. Z 500 scenárov spočítame scenáre spotových cien, ktoré následne priemerujeme.

| Model | Teplota [°C] | | Spotreba [%] | | Log výnosy | |
|-------|--------------|------|--------------|--------|------------|------|
| | P5 | P95 | P5 | P95 | P5 | P95 |
| m1 | -4,96 | 5,18 | 0,3993 | 0,6331 | -0,04 | 0,04 |
| m2 | -2,82 | 2,85 | 0,4176 | 0,6119 | -0,04 | 0,04 |
| m3 | -1,54 | 2,54 | 0,4991 | 0,5428 | -0,03 | 0,02 |
| m4 | -1,35 | 2,26 | 0,4982 | 0,5447 | -0,03 | 0,03 |

Tabuľka 5.2: Rozmedzie simulácií modelov pre mesiac január



Obr. 5.13: Histogramy scenárov modelov pre mesiac január

Z tabuľky 5.3 pozorujeme, že priemerné hodnoty jednotlivých modelov sú pomerne blízko seba. Najbližšie sú si vždy dvojice, jednorozmerné modely m1 a m2 majú veľmi podobné priemerné hodnoty všetkých troch veličín a aj viacrozmerné modely m3 a m4 majú priemerné hodnoty tesne pri sebe. Najväčší rozdiel teploty, ktorý medzi modelmi pozorujeme je v letných mesiacoch, kedy sa modely líšia o 1 °C. Spotreba sa medzi modelmi líši maximálne o 0,07 %, zatiaľ čo spotovú cenu má najväčší rozdiel až 3,7 EUR/MWh a to medzi modelmi m2 a m3 v letných mesiacoch.

| Mesiac | Priem. teplota [°C] | | | | Priemerné denné TDD [%] | | | | Priemerná cena [Eur/MWh] | | | |
|--------|---------------------|-------|-------|-------|-------------------------|--------|--------|--------|--------------------------|-------|-------|-------|
| | m1 | m2 | m3 | m4 | m1 | m2 | m3 | m4 | m1 | m2 | m3 | m4 |
| Jan | 0,09 | 0,09 | 0,45 | 0,44 | 0,5189 | 0,5156 | 0,5209 | 0,5210 | 18,94 | 19,01 | 18,96 | 18,94 |
| Feb | 1,44 | 1,33 | 1,74 | 1,76 | 0,4685 | 0,4309 | 0,4815 | 0,4815 | 18,89 | 18,99 | 18,17 | 18,27 |
| Mar | 5,17 | 5,12 | 5,72 | 5,72 | 0,3806 | 0,4204 | 0,3805 | 0,3806 | 18,86 | 18,99 | 17,29 | 17,47 |
| Apr | 10,35 | 10,40 | 11,22 | 11,20 | 0,2562 | 0,2511 | 0,2440 | 0,2442 | 18,63 | 18,74 | 16,46 | 16,62 |
| Máj | 15,68 | 15,65 | 16,63 | 16,64 | 0,1273 | 0,1742 | 0,1088 | 0,1087 | 18,51 | 18,88 | 15,90 | 16,01 |
| Jún | 19,45 | 19,48 | 20,44 | 20,48 | 0,0459 | 0,0742 | 0,0388 | 0,0388 | 18,45 | 18,95 | 15,55 | 15,79 |
| Júl | 20,81 | 20,77 | 21,82 | 21,81 | 0,0382 | 0,0382 | 0,0382 | 0,0382 | 18,55 | 19,37 | 15,58 | 16,04 |
| Aug | 19,26 | 19,22 | 20,26 | 20,27 | 0,0479 | 0,0695 | 0,0418 | 0,0417 | 18,64 | 19,76 | 16,06 | 16,70 |
| Sep | 15,29 | 15,23 | 16,21 | 16,20 | 0,1323 | 0,1333 | 0,1242 | 0,1251 | 18,89 | 19,82 | 16,71 | 17,46 |
| Okt | 10,02 | 9,90 | 10,71 | 10,69 | 0,2607 | 0,2823 | 0,2629 | 0,2628 | 19,25 | 20,05 | 17,37 | 18,21 |
| Nov | 4,64 | 4,62 | 5,32 | 5,33 | 0,3909 | 0,3640 | 0,3975 | 0,3975 | 19,46 | 20,19 | 17,87 | 18,87 |
| Dec | 0,86 | 0,80 | 1,37 | 1,40 | 0,4750 | 0,4948 | 0,4937 | 0,4934 | 19,49 | 20,36 | 17,93 | 19,10 |

Tabuľka 5.3: Priemerné hodnoty scenárov modelov

6. Praktická úloha

Uvedené výsledky simulácie použijeme na reálnu úlohu, ktorou sa zaoberajú firmy dodávajúce plyn danému portfóliu zákazníkov ako sú domácnosti a malo-odberatelia. V tomto prípade sa budeme zaujímať hlavne o výsledky simulácie spotreby a logaritmických výnosov, ktoré sme previedli na spotové ceny zemného plynu. Firmy dodávajúce plyn nakupujú na veľkoobchodnom trhu v mesačnej granularite tzv. mesačné baseloady. Jedná sa teda o konštantné množstvo plynu na každú hodinu v mesiaci. My v tomto prípade zanedbáme zmenu času a predpokladáme, že všetky dni majú rovnaký počet hodín. V zimnom období je väčšia spotreba plynu ako v letných mesiacoch, preto sa na rôzne mesiace v roku nakupujú rôzne objemy. Potrebné množstvo plynu počítame na základe mesačných TDD^M . Výšku mesačného TDD^M získame vysčítaním priemerných denných hodnôt tdd_t v danom mesiaci. Hodnota TDD^M vyjadruje koľko % ročného množstva plynu sa spotrebuje v mesiaci M . Výsledný objem plynu potrebný na daný mesiac, označme V^M , získame vynásobením mesačného TDD^M a CYC , teda ročným kontrahovaným množstvom. Hodnota CYC , udávajúca veľkosť portfólia, je konštanta, ktorá je daná vopred. Matematicky môžeme zapísať objem V^M potrebný na daný mesiac M takto:

$$V^M = TDD^M \cdot CYC, \quad TDD^M = \sum_{t \in M} tdd_t.$$

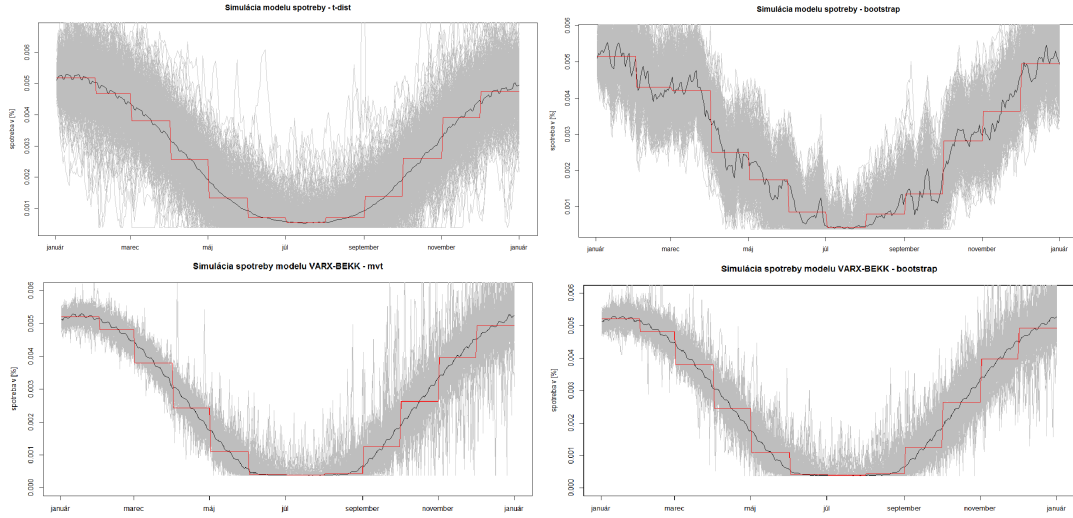
Výšku dennej spotreby V_t dostaneme vynásobením denného tdd_t a CYC . V tabuľke 6.1 uvádzame výšku mesačných TDD^M , $M = 1, \dots, 12$, z modelov m1, m2, m3, m4. V tabuľke vidíme, že hodnoty TDD^M modelov m1, m3 a m4 sú veľmi podobné a ich celkový súčet je približne 95 %. Model m2 má trochu iné rozloženie spotreby a súčet jednotlivých TDD^M je až 99 %.

| Mesiac | m1 | m2 | m3 | m4 |
|--------|-------|-------|-------|-------|
| Jan | 16,09 | 15,98 | 16,15 | 16,15 |
| Feb | 13,12 | 12,07 | 13,48 | 13,48 |
| Mar | 11,80 | 13,03 | 11,80 | 11,80 |
| Apr | 7,69 | 7,53 | 7,32 | 7,32 |
| Máj | 3,95 | 5,40 | 3,37 | 3,37 |
| Jún | 1,38 | 2,23 | 1,16 | 1,16 |
| Júl | 1,19 | 1,19 | 1,19 | 1,19 |
| Aug | 1,48 | 2,15 | 1,30 | 1,29 |
| Sep | 3,97 | 4,00 | 3,73 | 3,75 |
| Okt | 8,08 | 8,75 | 8,15 | 8,15 |
| Nov | 11,73 | 10,92 | 11,93 | 11,93 |
| Dec | 14,72 | 15,34 | 15,31 | 15,29 |
| Suma | 95,19 | 98,59 | 94,88 | 94,89 |

Tabuľka 6.1: Mesačné TDD^M jednotlivých modelov v %

Celkový súčet TDD^M sa nesčíta na 1, čo hovorí o tom, že portfólio domácností spotrebováva menej ako je hodnota ročného kontrahovaného množstva. Objem

plynu V^M , sa nakupuje podľa určenej nákupnej cesty približne rok pred dodávkou plynu na veľkoobchodnom trhu. Podľa daného % však nenakupujeme 100 % CYC, ale podľa vyššie uvedených výsledkov iba 95 %. Tento tzv. forwardový nákup je rozložený rovnomerne do 12 mesiacov, z ktorého následne vyjde istá forwardová cena. Označme FW^M je forwardová cena a teda vážený priemer cien nákupov pre mesiac M . Vychýlenia od nasimulovaných denných tdd a konštantnej priemernej hodnoty v daný mesiac môžeme vidieť na obrázku 6.1 pre všetky štyri uvažované modely.



Obr. 6.1: Mesačné TDD^M a priemerné denné tdd_t zo simulácií

Rozdiely medzi nakúpeným konštantným množstvom a skutočnou spotrebou v ten daný deň sú vysporiadúvané za aktuálne tržné ceny. Nakúpený mesačný objem V^M je rovnomerne rozpočítaný na denné objemy. Tento denný konštantný objem označme V_t^M . Označme S_t je spotová cena zemného plynu pre deň t .

Matematicky môžeme náklady na riziko zmeny dopytu (Demand variation risk, DVR) pre deň t v mesiaci M zapísať v tvare

$$DVR_t^M = (V_t^M - V_t) \cdot (S_t - FW^M),$$

kde

$$V_t^M = V^M / \text{počet dní v mesiaci } M.$$

Jedná sa teda o dodatočný nákup (alebo predaj) plynu, pričom cena tohto nákupu je ohodnotená rozdielom aktuálnej spotovej ceny a ceny forwardového nákupu. Tieto náklady sú neoddeliteľnou súčasťou podnikania v danej oblasti z dôvodu nákupu konštantného množstva plynu a rôznej spotreby v jednotlivé dni. Dodávateľov plynu zaujímajú očakávané náklady spojené s týmto rizikom zmeny dopytu. Tieto náklady spočítame pre všetky 4 modely.

Predpokladajme, že veľkosť portfólia je $CYC = 3$ TWh. Spočítame množstvo plynu V_t^M pre jednotlivé dni a tiež V_t podľa modelov simulácií. Rozdiely ohodnotíme rozdielom spotovej a forwardovej ceny, čím dostaneme DVR_t .

V tejto úlohe však v našom prípade nastáva zjednodušenie a to, že

$$\sum_{t \in M} (V_t^M - V_t) = 0, \text{ pretože } TDD^M = \sum_{t \in M} tdd_t.$$

Rozdiely objemov v jednotlivých mesiacoch sa sčítajú na 0, keďže očakávame priemerný scenár. Forwardová cena FW^M je v mesiaci M konštantná, a tak keď počítame riziko DVR_M pre jednotlivé mesiace, jeho výška nezávisí na FW^M , platí

$$DVR^M = \sum_{t \in M} (V_t^M - V_t) \cdot S_t - \sum_{t \in M} (V_t^M - V_t) \cdot FW^M$$

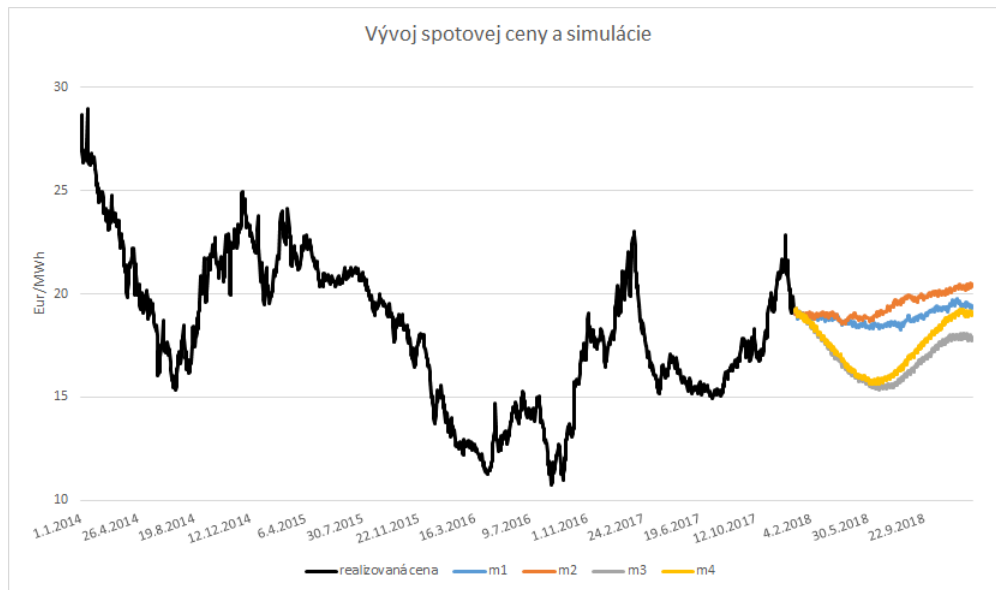
$$DVR^M = \sum_{t \in M} (V_t^M - V_t) \cdot S_t.$$

Tabuľka 6.2 zhrňuje náklady na DVR za jednotlivé mesiace. Spotová cena S_t , ktorou hodnotíme rozdiely v spotrebe je priemerná cena vypočítaná z 500 scenárov daného modelu. V modeli m1 sme narazili na scenár, ktorého výnosy boli veľmi vysoké, čím aj vypočítaná cena vyskočila na enormné hodnoty (milióny EUR/MWh). Scenár 343 sme preto nahradili priemerným scenárom výnosov bez inkriminovaného scenáru. V modeloch m1 a m2 vyšli náklady podobné, čo je spôsobené veľmi podobnou krivkou spotovej ceny. Rozdiely v nákladoch v rôznych mesiacoch však spôsobuje mesačné % nákupu TDD^M , respektíve denná spotreba V_t^M , ktorá je v prípade m2 výrazne skokovitá. Najvyššie náklady očakávame v mesiaci marec a tiež v jesenných mesiacoch pri prechode na zimné obdobie. V decembri očakávame výnosy, okrem modelu m2, čo je hlavne spôsobené odpredajom objemu za vysoké ceny začiatkom decembra. Celkové náklady na DVR očakávame v jednorozmerných modeloch vo výške približne 10 000 Eur. Vo viacrozmernom modeli vyšli náklady skoro štvornásobné oproti jednorozmerným modelom a to vo výške 46 000 Eur. Spôsobené je to výrazným rozdielom v modelovaní spotovej ceny, ktorá má vo viacrozmernom modeli prehnutejší tvar a líši sa od jednorozmerných modelov približne o 2 EUR/MWh. Priemerná spotreba medzi jednotlivými modelmi je podobná, tj. rozdiely $(V_t^M - V_t)$ sú na podobnej úrovni. Vysoké rozdiely medzi jednorozmernými a viacrozmernými modelmi preto spôsobuje spotová cena.

| Mesiac | m1 | m2 | m3 | m4 |
|--------|--------|---------|---------|---------|
| Jan | -550 | -737 | -338 | -293 |
| Feb | -270 | 43 | -4 399 | -4 124 |
| Mar | 1 085 | -2 205 | -10 564 | -8 511 |
| Apr | -631 | 387 | -7 871 | -8 960 |
| Máj | -445 | -1 492 | -4 149 | -4 043 |
| Jún | 288 | 1 728 | -59 | -11 |
| Júl | 0 | 0 | 0 | 0 |
| Aug | -1 020 | -1 464 | -895 | -1 176 |
| Sep | -2 115 | -1 796 | -6 206 | -7 852 |
| Okt | -2 164 | -885 | -8 382 | -7 631 |
| Nov | -3 905 | -3 704 | -3 242 | -4 488 |
| Dec | 619 | -850 | 194 | 487 |
| Suma | -9 108 | -10 975 | -45 910 | -46 602 |

Tabuľka 6.2: Priemerné mesačné náklady na DVR jednotlivých modelov v EUR

Uvedené náklady na DVR , ktoré sme spočítali sú však len minimálne, ktoré by nastali v ideálnom prípade, kedy by bola spotreba a cena priemerná a platilo



Obr. 6.2: Realizované a priemerné spotové ceny jednotlivých modelov

by, že $\sum_{t \in M} (V_t^M - V_t) = 0$. V skutočnosti však dochádza k rôznym výkyvom v spotrebe a náklady na DVR sa preto vyšplhajú vyššie.

Vývoj priemerných spotových cien jednotlivých modelov je vykreslený na obrázku 6.2. Modely vektorovej autoregresie majú prehnutejší tvar oproti jednorozmernému modelu výnosov. Bližšie hodnoty sú uvedené v tabuľke 5.3.

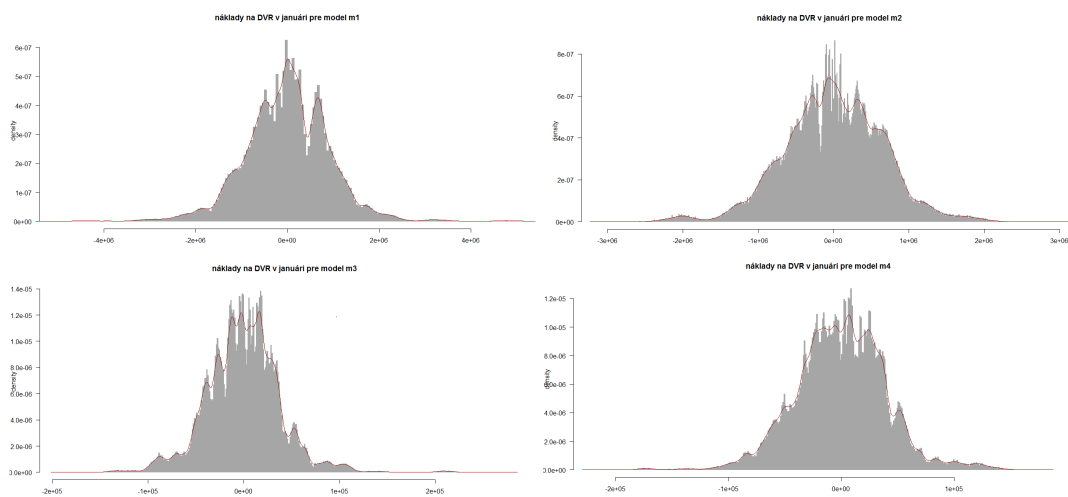
Úloha modelovania 3 veličín, ktoré sú navzájom prepojené, prináša okrem predikcií daných veličín tiež predikciu tzv. „spoločnej“ veličiny, ktorou je DVR a do ktorej vstupuje spotreba a spotová cena. Firmu dodávajúcu plyn v skutočnosti zaujíma nielen priemerné riziko DVR , ktoré, ako sme videli je veľmi nízke, ale aj celkové rozdelenie DVR^M , odkiaľ je možné následne spočítať extrémne náklady na riziko DVR . My sa v tejto práci konkrétne zameriame opäť iba na mesiac január, vzhľadom na výpočtovú zložitosť. Najprv spočítame denné hodnoty DVR_d . Z 500 scenárov spotreby spočítame $(V_t^1 - V_t)$, ktoré následne ohodnotíme 500 scenármi spotovej ceny. Celkovo tak máme maticu scenárov januárových hodnôt DVR_d dimenzie $31 \times 250\,000$. Denné hodnoty scenárov sčítame a dostaneme hodnotu mesačného DVR^1 .

Tabuľka 6.3 zhrňuje náklady na DVR pre 250 000 scenárov v každom zo 4 modelov. V nej je vidieť markantné rozdiely jednorozmerného a viacrozmerného modelovania. Viacrozmerné modelovanie prináša oveľa menšie náklady, pretože z obrázkov 5.8, 5.11 je patrné oveľa užšie rozdelenie ako v jednorozmernom prípade 5.2, 5.5. Viď tiež obrázok 6.1, kde je šedá časť scenárov spotreby širšia pre jednorozmerné modely. V prípade priemerného DVR , je situácia opačná, viď tabuľku 6.2, kedy jednorozmerný model hovorí o násobne menších nákladoch. Tento rozdiel, ako sme spomínali, je spôsobený rôznou priemernou spotovou cenovou krivkou.

V histórii posledných 4 rokov spoločnosť innogy pozorovala najvyššie mesačné náklady na DVR vo výške približne 1 200 000 EUR. Bežne sa však mesačné náklady pohybujú vo výške 300 000 EUR v zimných mesiacoch. Preto považujeme simulácie DVR z jednorozmerného modelu za bezpečnejší odhad nákladov.

| | m1 | m2 | m3 | m4 |
|-----|-----------|------------|---------|---------|
| P5 | 1 439 192 | -1 055 972 | -64 449 | -60 741 |
| P95 | 1 394 962 | 1 018 779 | 60 033 | 59 219 |

Tabuľka 6.3: Náklady na *DVR* pre mesiac január v EUR



Obr. 6.3: Histogram nákladov na *DVR* pre mesiac január

Záver

Cieľom tejto práce bolo nájsť vhodný model ceny plynu a tiež jeho spotreby a priemernej dennej teploty.

Jednorozmerné modely veličín, ktoré sa pôvodne používali v spoločnosti inogy, sme v tejto práci vylepšili pridaním sofistikovanejších modelových nástrojov ako sú autoregresné procesy a modely podmienenej heteroskedasticity chýb. Jednotlivé rovnice sme prepojili do viacrozmernej sústavy, kde sme spočítali korelačnú maticu chýb, ktorá preukázala silnú koreláciu medzi jednotlivými rovnicami čo odôvodňuje správnosť modelovať premenné v rámci jednej sústavy. Podmienenu variačnú maticu sme modelovali diagonálnym BEKK modelom, ktorý zaručil jej pozitívnu semidefinitnosť. Po odhade sme pristúpili k verifikácii modelov, kde sme nezistili žiadne porušenie predpokladov modelu. Následne sme simulovali scenáre vývoja na obdobie jedného roku. Okrem klasického prístupu generovania reziduí zo známeho rozdelenia sme použili pri simulácii bootstrapovú metódu na reziduá. Dostali sme tak 4 rôzne simulácie scenárov, ktoré sme následne porovnávali a použili v praktickej úlohe.

V súčasnosti od začiatku roku 2018 pozorujeme stále rastúce ceny plynu, čo je hlavne ovplyvnené rastúcou cenou ropy. Preto by bolo možné vylepšenie modelu ceny pridaním závislosti na cene ropy. Vzhľadom na globálne otepľovanie sme tiež svedkami rôznych teplotných anomálií, kedy zaznamenávame extrémne ovla častejšie ako v minulých rokoch. Napríklad veľmi studený marec roku 2018 s teplotami 5 °C pod priemernými hodnotami vystriedalo po krátkom čase letné počasie už v apríli. Na tieto zmeny rýchlo reaguje aj spotreba a tiež cena. V budúcnosti bude preto dôležité zamerať sa aj na extrémne výkyvy. Zakomponovať premennú, ktorá by ukazovala stav naplnenosti zásobníkov v kontexte modelovania ceny by bolo tiež možné vylepšenie. V práci Stoll a Wiebauer (2010) autori definovali premennú normalizované kumulované vykurovacie dni, ktorá mala vyjadrovať stav zásobníku na základe teploty. V našom modeli však táto premenná vyšla nevýznamne, čo však môže byť spôsobené tým, že teploty sú namerané na Slovensku a my sme ako referenčnú cenu používali nemecký index NCG. Celkovo by mohlo byť zaujímavé pridanie nemeckej teploty do modelu ceny.

Pri prehľade literatúry sa pri modelovaní spotovej ceny často spomína tiež modelovanie pomocou Kalmanovho filtra. Stavové modelovanie by mohlo byť taktiež možným prínosom a rozšírením modelu.

Súčasťou práce sú tiež súbory obsahujúce použité vstupné dáta, zdrojový kód a výstupný súbor. Uvedené modelovanie je dobrým nástrojom pre výpočet očakávanej spotreby dlho dopredu, avšak veličinu ako je spotová cena ovplyvňuje množstvo faktorov a preto jej dlhodobé modelovanie nie je spoľahlivé. Na základe charakteristík MSE a MAE sme sa presvedčili o tom, že viacrozmernej prístup k skúmanej problematike modelovania troch premenných, ktoré navzájom spolu interagujú je vhodnejší ako jednorozmerné modely. Pri porovnaní priemerných hodnôt scenárov boli oba modely porovnateľné.

Zjednotenie modelov do ucelenej sústavy, ktorá zohľadňuje väzby medzi veličinami považujeme za prínosné riešenie problému. Model VARX má výhodu v tom, že jeho odhad je jednoduchý, má bohatšiu štruktúru ako jednorozmerné modely a pomocou MGARCH tiež vieme prispôsobiť model volatility.

Zoznam použitej literatúry

- BABA, Y., ENGLE, R., KRAFT, D. F. a KRONER, K. F. (1990). Multivariate Simultaneous Generalized ARCH. *Department of Economics, University of California, San Diego*.
- BOLLERSLEV, T., ENGLE, R. F. a WOOLDRIDGE, J. M. (1988). A Capital-Asset Pricing Model with Time-Varying Covariances. *Journal of Political Economy*, **96**, 116–131.
- BOX, G. E. P. a JENKINS, G. M. (1970). *Time Series Analysis, Forecasting and Control*. Holden-Day, San Francisco. ISBN 978-0-470-27284-8.
- BROCKWELL, P. J. a DAVIS, R. A. (2002). *Introduction to Time Series and Forecasting*. Second Edition. Springer, New York. ISBN 0-387-95351-5.
- CABALLERO, R., JEWSON, S. a BRIX, A. (2002). Long memory in surface air temperature: detection, modeling, and application to weather derivative valuation. *Climate Research*, **21**(2), 127–140.
- CIPRA, T. (2008). *Finanční ekonometrie*. 1. vydání. Ekopress, Praha. ISBN 978-80-86929-43-9.
- EFRON, B. . (1979). Bootstrap methods: another look at the jackknife. *The Annals of Statistics*, **7**(7), 1–26.
- ENGLE, R. a KRONER, K. F. (1995). Multivariate Simultaneous Generalized ARCH. *Economic Theory*, **11**(1), 122–150.
- ENGLE, R. F., NG, V. K. a ROTSCCHILD, M. (1990). Asset pricing with a factor ARCH covariance structure: Empirical Estimates for Treasury Bills . *Journal of Econometrics*, **45**, 213–237.
- EViews (2015). *EViews 9 User's Guide II*. IHS Global Inc., Irvine, California. ISBN 978-1-880411-27-8.
- GHALANOS, A. (2015). *rmgarch: Multivariate GARCH Models*. R package, version 1.3-0.
- GHALANOS, A. (2018). *rugarch: Univariate GARCH Models*. R package, version 1.4-0.
- GÖNCÜ, A., KARAHAN, M. O. a KUZUBAS, T. U. (2013). Forecasting Daily Residential Natural Gas Consumption: A Dynamic Temperature Modelling Approach. *Working papers, Bogazici University, Department of Economics*.
- KWIATKOWSKI, E. A. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, **54**(2), 159–178.
- LÜTKEPOHL, H. (2005). *New Introduction to Multiple Time Series Analysis* . Springer, Berlin. ISBN 3-540-40172-5.

- MÜLLER, J., HIRSCH, A. a MÜLLER, A. (2015). Modeling the Price of Natural Gas with Temperature and Oil Price as Exogenous Factors. In Glau, K., Schrer, M. Zagst, R. (Eds.), *Innovations in Quantitative Risk Management, Springer Proceedings in Mathematics and Statistics*, New York: Springer, pages 109–128.
- NESLON, D. B. a CAO, C. Q. (1992). Inequality Constraints in the Univariate GARCH Model. *Journal of Business and Economic Statistics*, **10**(2), 229–235.
- PFAFF, B. a STIGLER, M. (2013). *vars: VAR Modelling*. R package, version 1.5-2.
- PRÁŠKOVÁ, Z. (2016). *Základy náhodných procesů II*. 2. upravené vydanie. Karolinum, Praha. ISBN 978-80-246-3516-3.
- SCHMIDBAUER, H., ROESCH, A. a TUNALIOGLU, V. S. (2016). *rmgarchBEKK: Simulating, Estimating and Diagnosing MGARCH (BEKK and mGJR) Processes*. R package, version 0.0.2.
- SCHWARTZ, E. S. (1997). The stochastic behavior of commodity prices: implications for valuation and hedging. *Journal of Finance*, **52**(3), 923–973.
- SCHWARTZ, E. S. a SMITH, J. E. (2000). Short-term variations and long-term dynamics in commodity prices. *Management Science*, **46**(7), 893–911.
- SILVENNOINEN, A. a TERÄSVIRTA, T. (2008). Multivariate GARCH models. In: Handbook of Financial Time Series (T. G. Andersen, R. A. Davis, J.-P. Kreiss and T. Mikosch, eds.) . pages 201–229.
- STOLL, S. O. a WIEBAUER, K. (2010). A spot price model for natural gas considering temperature as an exogenous factor with applications. *The Journal of Energy Markets, Series B*, **3**(3), 113–128.
- TSAY, R. S. (2015). *MTS: All-Purpose Toolkit for Analyzing Multivariate Time Series (MTS) and Estimating Multivariate Volatility Models*. R package, version 0.33.
- TSAY, S. R. (2014). *Multivariate Time Series Analysis: With R and Financial Applications*. 1st Edition. Wiley, Chicago. ISBN 978-1-118-61790-8.
- VYHLÁŠKA Č. 24/2013 Z. z. (2013). Vyhláška Úradu pre reguláciu sieťových odvetví, ktorou sa ustanovujú pravidlá pre fungovanie vnútorného trhu s elektrinou a pravidlá pre fungovanie vnútorného trhu s plynom. Prístupné na: <https://www.slov-lex.sk/pravne-predpisy/SK/ZZ/2013/24/>, navštívené 3.7.2018.

Zoznam obrázkov

| | | |
|------|--|----|
| 1.1 | Tvar ACF a PACF funkcií základných modelov | 9 |
| 2.1 | Denné hodnoty spotreby tdd_t | 32 |
| 2.2 | Závislosť spotreby plynu na teplote | 33 |
| 2.3 | Vývoj priemernej dennej teploty na Slovensku v °C | 33 |
| 2.4 | Spotová cena zemného plynu v EUR/MWh | 34 |
| 3.1 | Odhad modelu spotreby | 37 |
| 3.2 | QQ plot štandardizovaných reziduí modelu spotreby | 38 |
| 3.3 | Odhad modelu teploty | 39 |
| 3.4 | QQ plot štandardizovaných reziduí modelu teploty | 40 |
| 3.5 | Odhad modelu logaritmických výnosov | 42 |
| 3.6 | QQ plot štandardizovaných reziduí modelu logaritmických výnosov | 42 |
| 4.1 | Významnosť vzájomných korelácií a autokorelácií vo VARX(3,1) | 46 |
| 4.2 | Reziduá modelu VARX(3,1) | 47 |
| 4.3 | Korelogram reziduí modelu VARX(3,1) | 48 |
| 4.4 | Odozva na impulz modelu VARX(3,1) | 48 |
| 4.5 | Grafy podmienených korelácií modelu BEKK(2,1) spolu s konštantnou nepodmienenou koreláciou | 50 |
| 5.1 | Namerané hodnoty teploty a simulácie jednorozmerného modelu | 52 |
| 5.2 | Namerané hodnoty spotreby a simulácie jednorozmerného modelu | 52 |
| 5.3 | Namerané hodnoty výnosov a simulácie jednorozmerného modelu | 52 |
| 5.4 | Namerané hodnoty teploty a simulácie modelu - bootstrap | 53 |
| 5.5 | Namerané hodnoty spotreby a simulácie modelu - bootstrap | 53 |
| 5.6 | Namerané hodnoty výnosov a simulácie modelu - bootstrap | 54 |
| 5.7 | Namerané hodnoty teploty a simulácie modelu VARX-BEKK | 55 |
| 5.8 | Namerané hodnoty spotreby a simulácie modelu VARX-BEKK | 55 |
| 5.9 | Namerané hodnoty výnosov a simulácie modelu VARX-BEKK | 56 |
| 5.10 | Hodnoty teploty a simulácie modelu VARX-BEKK - bootstrap | 56 |
| 5.11 | Hodnoty spotreby a simulácie modelu VARX-BEKK - bootstrap | 57 |
| 5.12 | Hodnoty výnosov a simulácie modelu VARX-BEKK - bootstrap | 57 |
| 5.13 | Histogramy scenárov modelov pre mesiac január | 59 |
| 6.1 | Mesačné TDD^M a priemerné denné tdd_t zo simulácií | 62 |
| 6.2 | Realizované a priemerné spotové ceny jednotlivých modelov | 64 |
| 6.3 | Histogram nákladov na DVR pre mesiac január | 65 |
| 6.4 | Histogramy scenárov modelov pre mesiac marec | 78 |
| 6.5 | Histogramy scenárov modelov pre mesiac jún | 79 |
| 6.6 | Histogramy scenárov modelov pre mesiac september | 80 |
| 6.7 | Histogramy scenárov modelov pre mesiac december | 81 |

Zoznam tabuliek

| | | |
|------|--|----|
| 1.1 | Tvar ACF a PACF funkcií základných modelov | 9 |
| 2.1 | Deskriptívne štatistiky teplotného diagramu dodávky <i>tdd</i> v % . . . | 32 |
| 2.2 | Deskriptívne štatistiky priemernej dennej teploty v °C | 33 |
| 2.3 | Deskriptívne štatistiky spotovej ceny plynu v EUR/MWh | 34 |
| 3.1 | Korelogram štandardizovaných reziduí modelu spotreby | 37 |
| 3.2 | Korelogram štandardizovaných reziduí modelu teploty | 39 |
| 3.3 | Korelogram štandardizovaných reziduí modelu výnosov | 42 |
| 4.1 | Jednorozmerné modely veličín | 43 |
| 4.2 | Testy stacionarity časovej rady | 44 |
| 4.3 | Informačné kritériá modelov VARX | 44 |
| 4.4 | Grangerov test kauzality | 46 |
| 4.5 | Šikmost a špicatosť reziduí modelu VARX(3,1) | 47 |
| 5.1 | Porovnanie modelov podľa presnosti predpovedí | 58 |
| 5.2 | Rozmedzie simulácií modelov pre mesiac január | 58 |
| 5.3 | Priemerné hodnoty scenárov modelov | 60 |
| 6.1 | Mesačné TDD^M jednotlivých modelov v % | 61 |
| 6.2 | Priemerné mesačné náklady na <i>DVR</i> jednotlivých modelov v EUR | 63 |
| 6.3 | Náklady na <i>DVR</i> pre mesiac január v EUR | 65 |
| 6.4 | Odhadnuté parametre modelu spotreby | 73 |
| 6.5 | Odhadnuté parametre modelu teploty | 74 |
| 6.6 | Odhadnuté parametre modelu logaritmických výnosov | 74 |
| 6.7 | Odhadnuté koeficienty modelu VARX(3,1) | 75 |
| 6.8 | Pokračovanie: Odhadnuté koeficienty modelu VARX(3,1) | 76 |
| 6.9 | Portmanteau test reziduí modelu VARX(3,1) | 77 |
| 6.10 | Odhadnuté koeficienty modelu BEKK(2,1) | 77 |

Zoznam použitých skratiek

| | |
|-------|---|
| CYC | Ročné kontrahované množstvo (contracted yearly consumption) |
| DA | Day-Ahead (spotová cena) |
| DVR | Demand variation risk - riziko zmeny dopytu |
| HDD | Vykurovací deň (heating degree day) |
| MD | Domácnosti a maloobchod |
| NCG | NetConnetGermany virtuálny obchodný bod |
| NHDD | Normovaný vykurovací deň (normalized heating degree day) |
| SPP-D | Slovenský plynárenský priemysel - distribúcia |
| TDD | Teplotný diagram dodávky |

Prílohy

A.1

Hadamarov súčin

Nech \mathbf{A}, \mathbf{B} sú matice rovnakej dimenzie $m \times n$. *Hadamarov súčin* matíc \mathbf{A}, \mathbf{B} je matica rovnakého typu $m \times n$, označujeme $\mathbf{A} \odot \mathbf{B}$, s prvkami

$$(\mathbf{A} \odot \mathbf{B})_{ij} = (\mathbf{A})_{ij} \cdot (\mathbf{B})_{ij}.$$

Pre matice \mathbf{A}, \mathbf{B} typu 3×2 platí:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \odot \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} & a_{12}b_{12} \\ a_{21}b_{21} & a_{22}b_{22} \\ a_{31}b_{31} & a_{32}b_{32} \end{pmatrix}$$

Kroneckerov súčin

Nech \mathbf{A} je matica $m \times k$ a \mathbf{B} je matica $n \times l$. *Kroneckerov súčin* matíc \mathbf{A}, \mathbf{B} je matica typu $mn \times kl$, označujeme $\mathbf{A} \otimes \mathbf{B}$ a platí, že

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots & a_{1k}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & & a_{2k}\mathbf{B} \\ \vdots & & \ddots & \vdots \\ a_{m1}\mathbf{B} & a_{m2}\mathbf{B} & \cdots & a_{mk}\mathbf{B} \end{pmatrix}$$

Pravdepodobnostné rozdelenia

Normálne rozdelenie

Náhodná veličina $X \sim N(\mu, \sigma^2)$ s $\mu \in \mathbb{R}$ a $\sigma^2 > 0$ má hustotu v tvare

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)}, \quad x \in \mathbb{R}.$$

Studentovo rozdelenie

Náhodná veličina $X \sim t_v$ s počtom stupňov voľnosti $v \in \mathbb{R}$ má hustotu v tvare

$$f(x) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)\sqrt{v\pi}} \left(1 + \frac{x^2}{v}\right)^{-(v+1)/2}, \quad x \in \mathbb{R},$$

kde Gamma funkcia pre $p > 0$ je definovaná vzťahom

$$\Gamma(p) = \int_0^{\infty} x^{p-1} e^{-x} dx.$$

Mnohorozmerné normálne rozdelenie

Náhodná veličina $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ s $\boldsymbol{\mu} \in \mathbb{R}^k$ a pozitívne definitnou rozptylovou maticou $\boldsymbol{\Sigma}$ rádu $k \times k$ má hustotu v tvare

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{k/2} \sqrt{\det \boldsymbol{\Sigma}}} \exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right\}, \quad \mathbf{x} \in \mathbb{R}^k.$$

Mnohorozmerné Studentovo t-rozdelenie

Nech náhodná veličina $\mathbf{U} \sim N_k(\mathbf{0}, \boldsymbol{\Sigma})$, kde $\boldsymbol{\Sigma}$ je $k \times k$ pozitívne semidefinitná matica. Nech $V \sim \chi_v^2$ pre nejaké $v > 0$ a \mathbf{U} a V sú nezávislé. Potom hovoríme, že náhodný vektor

$$\mathbf{T} = \mathbf{U} \sqrt{\frac{v}{V}}$$

má k -rozmerné Studentovo t-rozdelenie s počtom stupňov volnosti v a škálovou maticou $\boldsymbol{\Sigma}$. Píšeme, že

$$\mathbf{T} \sim mvt_{k,v}(\boldsymbol{\Sigma}).$$

Ak je matica $\boldsymbol{\Sigma}$ regulárna, pozitívne definitná matica, potom hustota $mvt_{k,v}(\boldsymbol{\Sigma})$ má tvar

$$f(\mathbf{x}) = \frac{\Gamma(\frac{v+k}{2})}{\Gamma(\frac{v}{2}) v^{\frac{k}{2}} \pi^{\frac{k}{2}}} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \left\{1 + \frac{\mathbf{x}^T \boldsymbol{\Sigma}^{-1} \mathbf{x}}{v}\right\}^{-\frac{v+k}{2}}, \quad \mathbf{x} \in \mathbb{R}^k.$$

A.2

| Parameter | Odhad koef. | Sm. odchýlka | t-štatistika | p-hodnota |
|-------------|-------------|--------------|--------------|-----------|
| a_0 | 0,004 | 0,000 | 34,648 | 0,000 |
| a_1 | 0,000 | 0,000 | -16,898 | 0,000 |
| a_2 | 0,000 | 0,000 | -6,554 | 0,000 |
| a_3 | 0,000 | 0,000 | 2,872 | 0,004 |
| a_4 | 0,002 | 0,000 | 15,521 | 0,000 |
| a_5 | 0,000 | 0,000 | -8,585 | 0,000 |
| φ_1 | 1,157 | 0,027 | 43,681 | 0,000 |
| φ_2 | -0,206 | 0,026 | -7,754 | 0,000 |
| α_0 | 0,000 | 0,000 | 0,013 | 0,989 |
| α_1 | 0,050 | 0,016 | 3,229 | 0,001 |
| β_1 | 0,899 | 0,015 | 58,877 | 0,000 |
| v | 3,993 | 0,299 | 13,364 | 0,000 |

Tabuľka 6.4: Odhadnuté parametre modelu spotreby

| Parameter | Odhad koef. | Sm. odchýlka | t-štatistika | p-hodnota |
|-------------|-------------|--------------|--------------|-----------|
| b_0 | 12,119 | 0,432 | 28,028 | 0,000 |
| b_1 | -0,001 | 0,001 | -2,185 | 0,029 |
| b_2 | -2,522 | 0,318 | -7,926 | 0,000 |
| b_3 | -10,384 | 0,324 | -32,05 | 0,000 |
| φ_1 | 1,111 | 0,027 | 40,753 | 0,000 |
| φ_2 | -0,408 | 0,039 | -10,407 | 0,000 |
| φ_3 | 0,109 | 0,027 | 4,017 | 0,000 |
| α_0 | 0,296 | 0,274 | 1,078 | 0,281 |
| α_1 | 0,053 | 0,030 | 1,775 | 0,076 |
| β_1 | 0,844 | 0,120 | 7,031 | 0,000 |
| ν | 15,131 | 5,591 | 2,706 | 0,007 |

Tabuľka 6.5: Odhadnuté parametre modelu teploty

| Parameter | Odhad koef. | Sm. odchýlka | t-štatistika | p-hodnota |
|------------|-------------|--------------|--------------|-----------|
| c_0 | 0,008 | 0,001 | 8,243 | 0,000 |
| c_1 | -0,006 | 0,001 | -4,806 | 0,000 |
| c_2 | -0,007 | 0,001 | -5,809 | 0,000 |
| c_3 | -0,007 | 0,001 | -5,671 | 0,000 |
| c_4 | -0,009 | 0,001 | -6,901 | 0,000 |
| c_5 | -0,011 | 0,001 | -10,911 | 0,000 |
| c_6 | 6,172 | 1,440 | 4,287 | 0,000 |
| c_7 | -6,273 | 1,437 | -4,365 | 0,000 |
| α_0 | 0,000 | 0,000 | 1,014 | 0,311 |
| α_1 | 0,080 | 0,036 | 2,229 | 0,026 |
| α_2 | 0,000 | 0,052 | 0,000 | 1,000 |
| β_1 | 0,919 | 0,028 | 32,775 | 0,000 |
| ν | 3,671 | 0,442 | 8,299 | 0,000 |

Tabuľka 6.6: Odhadnuté parametre modelu logaritnických výnosov

A.3

| Parameter | | Odhad koef. | Sm. odchýlka | t-štatistika | p-hodnota |
|-----------|----|-------------|--------------|--------------|-----------|
| Φ_0 | 1 | 0,000 | 0,000 | 1,439 | 0,150 |
| | 2 | 2,456 | 0,609 | 4,031 | 0,000 |
| | 3 | 0,008 | 0,007 | 1,157 | 0,248 |
| Φ_1 | 11 | 1,249 | 0,035 | 35,329 | 0,000 |
| | 12 | 0,000 | 0,000 | -2,271 | 0,023 |
| | 13 | 0,001 | 0,000 | 2,125 | 0,034 |
| | 21 | -1169 | 247,1 | -4,733 | 0,000 |
| | 22 | 0,995 | 0,035 | 28,233 | 0,000 |
| | 23 | -5,245 | 2,464 | -2,128 | 0,033 |
| | 31 | -0,666 | 2,647 | -0,252 | 0,801 |
| | 32 | 0,000 | 0,000 | 0,327 | 0,744 |
| | 33 | 0,049 | 0,026 | 1,856 | 0,064 |
| Φ_2 | 11 | -0,401 | 0,052 | -7,632 | 0,000 |
| | 12 | 0,000 | 0,000 | 3,295 | 0,001 |
| | 13 | 0,001 | 0,000 | 1,543 | 0,123 |
| | 21 | 1254 | 366,9 | 3,417 | 0,001 |
| | 22 | -0,33 | 0,048 | -6,892 | 0,000 |
| | 23 | -4,214 | 2,475 | -1,702 | 0,089 |
| | 31 | -1,042 | 3,93 | -0,265 | 0,791 |
| | 32 | 0,000 | 0,001 | -0,338 | 0,736 |
| | 33 | -0,046 | 0,027 | -1,735 | 0,083 |
| Φ_3 | 11 | 0,094 | 0,034 | 2,721 | 0,007 |
| | 12 | 0,000 | 0,000 | -1,237 | 0,216 |
| | 13 | 0,001 | 0,000 | 1,667 | 0,096 |
| | 21 | -267,5 | 240,5 | -1,113 | 0,266 |
| | 22 | 0,113 | 0,034 | 3,293 | 0,001 |
| | 23 | -3,499 | 2,458 | -1,424 | 0,155 |
| | 31 | 0,666 | 2,576 | 0,259 | 0,796 |
| | 32 | 0,000 | 0,000 | -0,096 | 0,923 |
| | 33 | 0,039 | 0,026 | 1,475 | 0,140 |

Tabuľka 6.7: Odhadnuté koeficienty modelu VARX(3,1)

| Parameter | Odhad koef. | Sm. odchýlka | t-štatistika | p-hodnota | | |
|-----------|-------------|--------------|--------------|-----------|-------|-------|
| | 11 | 0,000 | 0,000 | 1,710 | 0,088 | |
| | 12 | 0,000 | 0,000 | 8,101 | 0,000 | |
| | 13 | 0,000 | 0,000 | -0,417 | 0,676 | |
| | 14 | 0,000 | 0,000 | -1,603 | 0,109 | |
| | 15 | 0,000 | 0,000 | 2,490 | 0,013 | |
| | 16 | 0,000 | 0,000 | 1,128 | 0,259 | |
| | 17 | 0,000 | 0,000 | -0,042 | 0,967 | |
| β_0 | 21 | -0,372 | 0,079 | -4,681 | 0,000 | |
| | 22 | -1,895 | 0,188 | -10,088 | 0,000 | |
| | 23 | 0,436 | 0,418 | 1,044 | 0,297 | |
| | 24 | 2,102 | 1,200 | 1,752 | 0,080 | |
| | 25 | -1,376 | 1,216 | -1,131 | 0,258 | |
| | 26 | -0,594 | 0,665 | -0,893 | 0,372 | |
| | 27 | 0,066 | 0,169 | 0,391 | 0,696 | |
| | 31 | -0,001 | 0,001 | -1,342 | 0,180 | |
| | 32 | 0,001 | 0,002 | 0,739 | 0,460 | |
| | 33 | -0,003 | 0,004 | -0,600 | 0,549 | |
| | 34 | 0,005 | 0,013 | 0,392 | 0,695 | |
| | 35 | -0,009 | 0,013 | -0,705 | 0,481 | |
| | 36 | -0,006 | 0,007 | -0,774 | 0,439 | |
| | 37 | -0,009 | 0,002 | -4,715 | 0,000 | |
| | | 11 | 0,000 | 0,000 | 0,000 | 0,000 |
| | | 12 | 0,000 | 0,000 | 0,000 | 0,000 |
| | | 13 | 0,000 | 0,000 | 1,465 | 0,143 |
| | 14 | 0,000 | 0,000 | -2,677 | 0,008 | |
| | 15 | 0,000 | 0,000 | -1,471 | 0,141 | |
| | 16 | 0,000 | 0,000 | -1,306 | 0,192 | |
| | 17 | 0,000 | 0,000 | -0,076 | 0,939 | |
| β_1 | 21 | 0,000 | 0,000 | 0,000 | 0,000 | |
| | 22 | 0,000 | 0,000 | 0,000 | 0,000 | |
| | 23 | -1,557 | 1,273 | -1,223 | 0,222 | |
| | 24 | 1,822 | 1,272 | 1,433 | 0,152 | |
| | 25 | 1,065 | 0,77 | 1,384 | 0,167 | |
| | 26 | 0,389 | 0,448 | 0,87 | 0,384 | |
| | 27 | 0,311 | 0,417 | 0,744 | 0,457 | |
| | 31 | 0,000 | 0,000 | 0,000 | 0,000 | |
| | 32 | 0,000 | 0,000 | 0,000 | 0,000 | |
| | 33 | -0,007 | 0,014 | -0,548 | 0,583 | |
| | 34 | 0,004 | 0,014 | 0,317 | 0,752 | |
| | 35 | 0,001 | 0,008 | 0,077 | 0,939 | |
| | 36 | -0,008 | 0,005 | -1,595 | 0,111 | |
| | 37 | 0,005 | 0,004 | 1,110 | 0,267 | |

Tabuľka 6.8: Pokračovanie: Odhadnuté koeficienty modelu VARX(3,1)

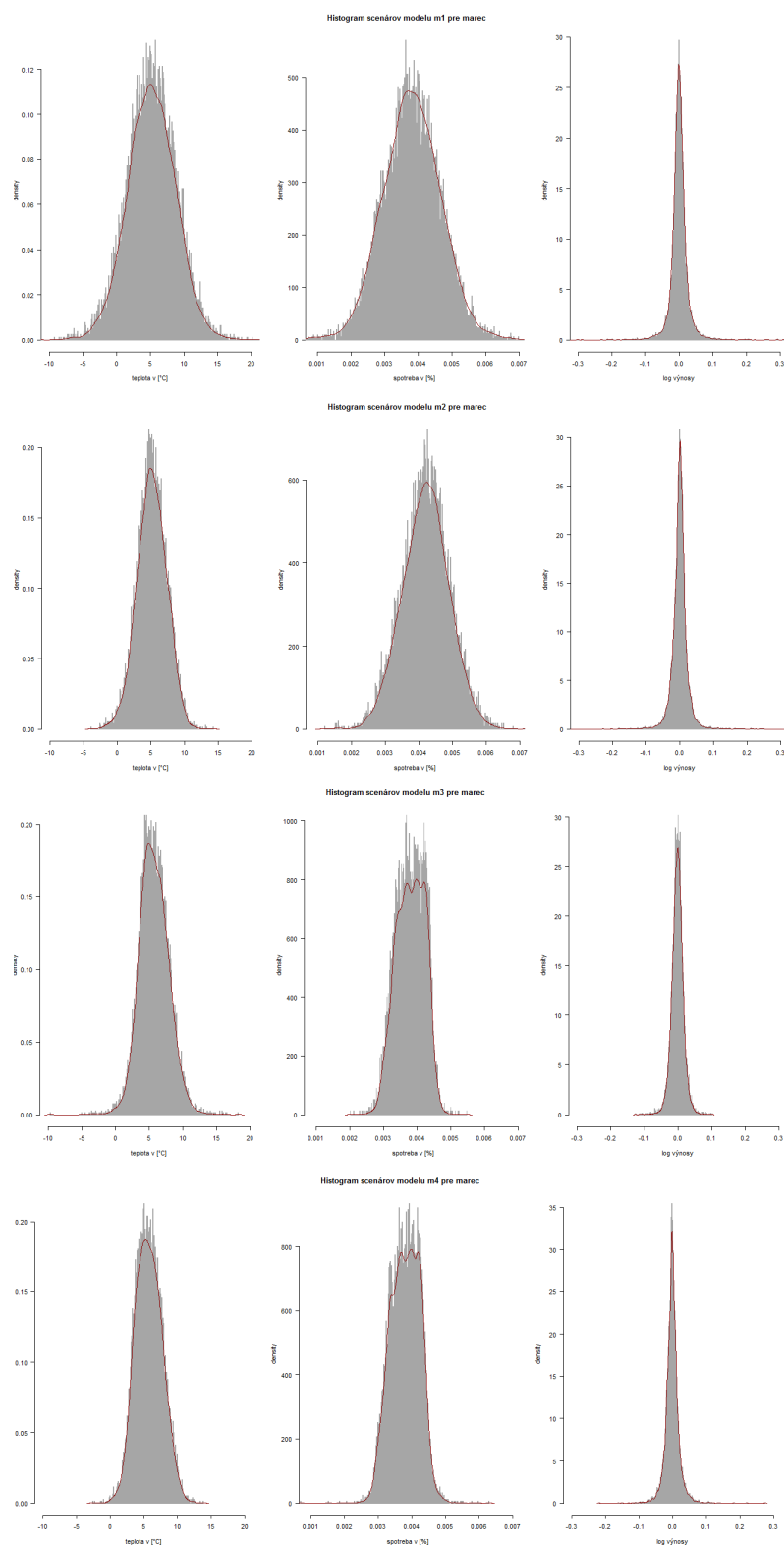
| m | Q(m) | df | p-hod | m | Q(m) | df | p-hod |
|----|--------|----|-------|----|---------|-----|-------|
| 1 | 0,487 | 9 | 1,00 | 11 | 84,518 | 99 | 0,85 |
| 2 | 0,835 | 18 | 1,00 | 12 | 88,634 | 108 | 0,91 |
| 3 | 4,061 | 27 | 1,00 | 13 | 102,477 | 117 | 0,83 |
| 4 | 20,949 | 36 | 0,98 | 14 | 124,22 | 126 | 0,53 |
| 5 | 30,394 | 45 | 0,95 | 15 | 140,921 | 135 | 0,35 |
| 6 | 44,354 | 54 | 0,82 | 16 | 149,898 | 144 | 0,35 |
| 7 | 55,734 | 63 | 0,73 | 17 | 158,779 | 153 | 0,36 |
| 8 | 67,848 | 72 | 0,62 | 18 | 162,331 | 162 | 0,48 |
| 9 | 70,469 | 81 | 0,79 | 19 | 166,259 | 171 | 0,59 |
| 10 | 79,752 | 90 | 0,77 | 20 | 185,124 | 180 | 0,38 |

Tabulka 6.9: Portmanteau test reziduí modelu VARX(3,1)

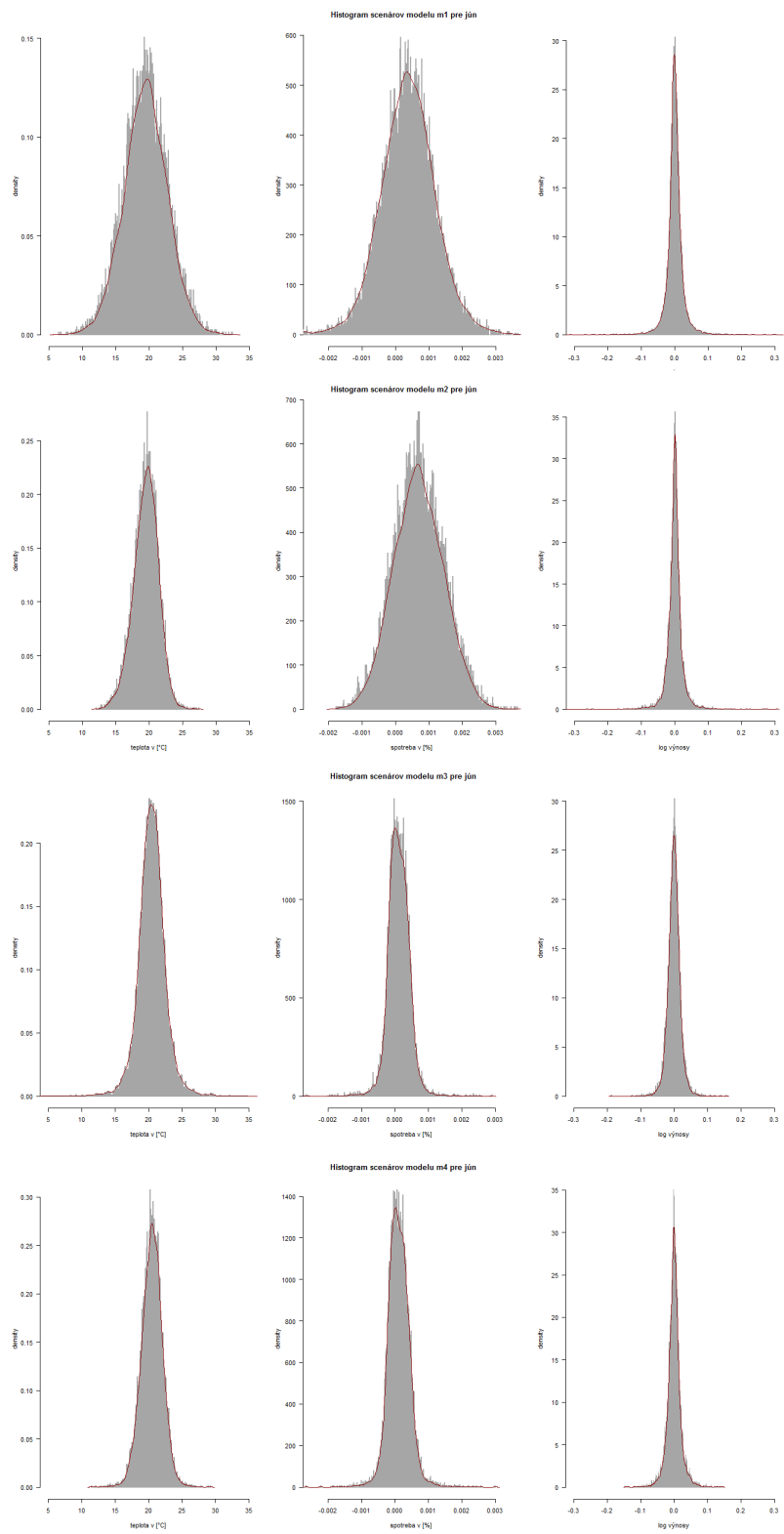
| Parameter | Odhad koef. | Sm. odchýlka | t-štatistika | p-hodnota | |
|----------------------|-------------|--------------|--------------|-----------|-------|
| C | 1,1 | 0,000 | 0,000 | 2,764 | 0,060 |
| | 2,2 | 0,024 | 0,012 | 2,012 | 0,044 |
| | 3,3 | 0,000 | 0,000 | 3,500 | 0,001 |
| A₁ | 1,1 | 0,224 | 0,038 | 5,890 | 0,000 |
| | 2,2 | 0,185 | 0,041 | 4,528 | 0,000 |
| | 3,3 | 0,252 | 0,025 | 10,049 | 0,000 |
| A₂ | 1,1 | 0,211 | 0,046 | 4,607 | 0,000 |
| | 2,2 | 0,200 | 0,044 | 4,586 | 0,000 |
| | 3,3 | -0,127 | 0,053 | -2,378 | 0,017 |
| G₁ | 1,1 | 0,951 | 0,006 | 172,031 | 0,000 |
| | 2,2 | 0,961 | 0,007 | 147,412 | 0,000 |
| | 3,3 | 0,949 | 0,008 | 124,188 | 0,000 |

Tabulka 6.10: Odhadnuté koeficienty modelu BEKK(2,1)

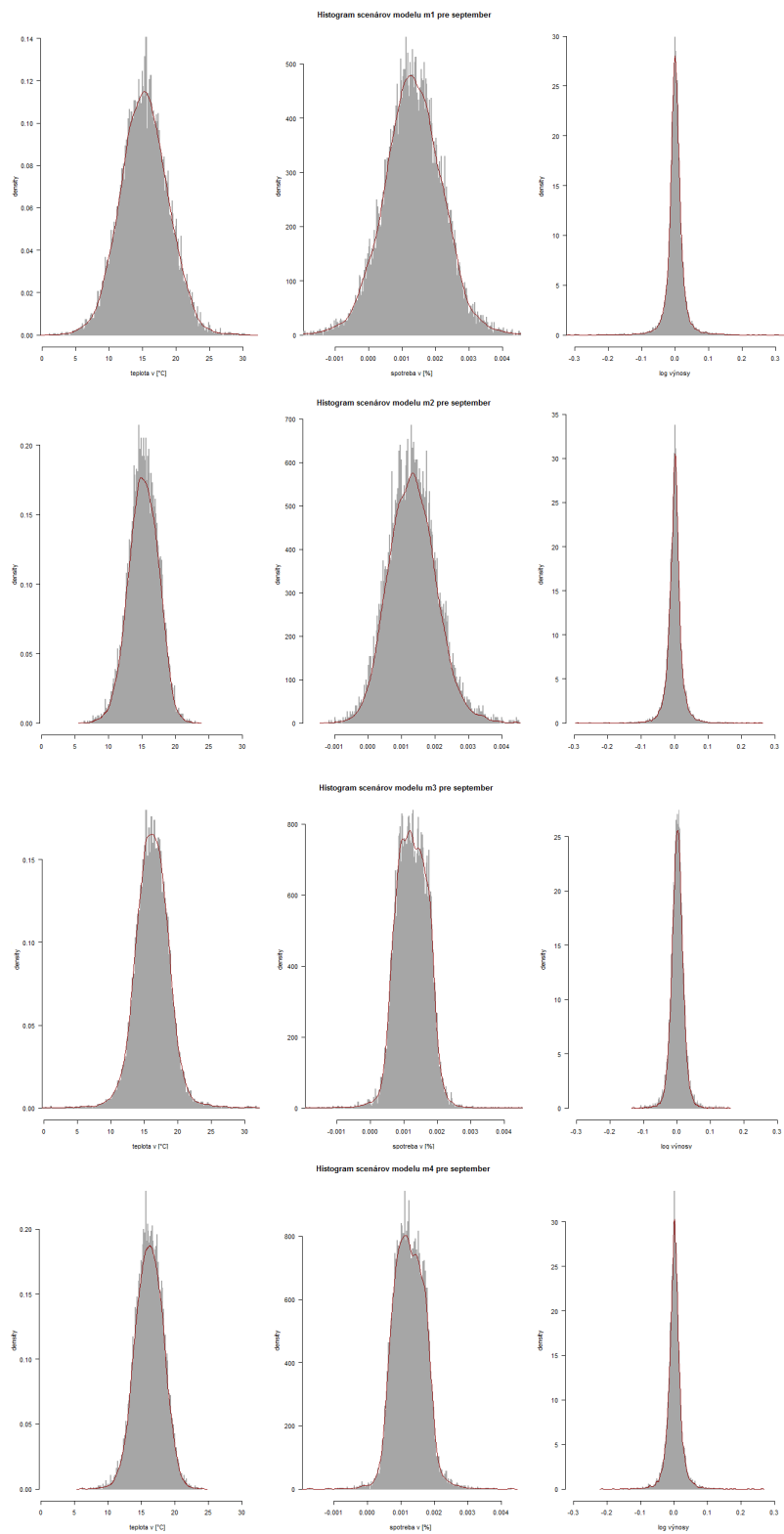
A.4



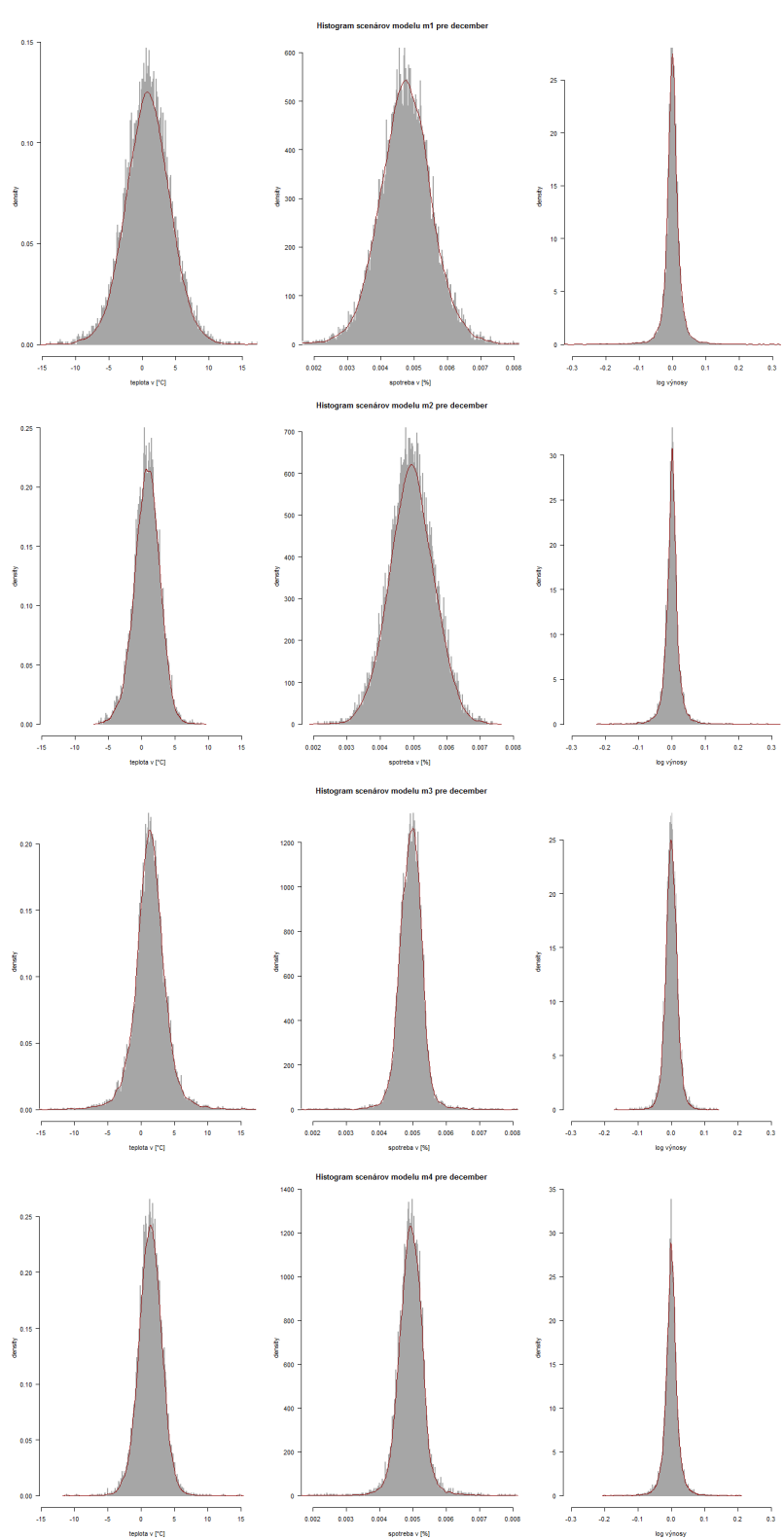
Obr. 6.4: Histogramy scénárov modelov pre mesiac marec



Obr. 6.5: Histograme scenárov modelov pre mesiac jún



Obr. 6.6: Histogramy scenárov modelov pre mesiac september



Obr. 6.7: Histogramy scenárov modelov pre mesiac december