

VIZUÁLNÍ NELINEÁRNÍ REKURENTNÍ ANALÝZA A JEJÍ APLIKACE NA ČESKÝ AKCIOVÝ TRH*

Jan Kodera, Tran Van Quang, Centrum základního výzkumu pro dynamickou ekonomii a ekonometrii¹, Vysoká škola ekonomická v Praze

1. Úvod

Myšlenka, že složitý vývoj finančních řad může být generován relativně jednoduchým nelineárním dynamickým systémem, není příliš stará. Přibližně dvacet let se vědci snaží popsat složitý vývoj finančních řad relativně jednoduchými deterministickými systémy. To však neznamená, že se lineární stochastické systémy přestaly používat pro popis složité dynamiky finančních časových řad (Trešl, Blatná, 2007). Oba přístupy mají rovnocenné místo v odborné literatuře. U deterministických nelineárních systémů bychom rádi zdůraznili jednu jejich vlastnost. Touto důležitou vlastností je, že jsou schopny vytvářet složitější dynamiku časových řad, která se vyznačuje aperiodickými oscilacemi, tedy je to dynamika vytvářená relativně jednoduchými deterministickými systémy (existence nelineárních prvků je nezbytná). Tato dynamika bývá relativně složitá (aperiodické oscilace), takže spíše připomíná časové řady generované stochastickými dynamickými systémy. Vzniká tedy problém, jak odlišit případy, kdy řada s relativně složitou dynamikou, je generována nelineárním deterministickým systémem a kdy se jedná o řadu stochastickou. Přispět k řešení tohoto problému je jedním z cílů tohoto článku.

Objev, že dynamické nelineární systémy mohou generovat aperiodické oscilace podobné stochastickým procesům je připisován meteorologovi Edwardu Lorenzovi (Lorenz 1963), který vytvořil zjednodušený systém tří nelineárních diferenciálních rovnic popisující pohyb vodních částic v atmosféře zahříváných zezdola a ochlazených shora. Stopa, kterou tyto částice opisují v třírozměrném abstraktním prostoru, se ustálí na tzv. Lorenzově atraktoru. Záměrně používáme slovo abstraktní, protože se nejedná o prostor, ve kterém se pohybujeme, ale o prostor, kde jeden rozměr je vyhrazen rychlosti rotace, druhý rozměr je rozdíl teploty klesající a stoupající částice a třetí rozměr zahrnuje distorzi od linearity vertikálního teplotního profilu.

Není to poprvé, kdy zásadní objev v oblasti přírodních věd je aplikován a úspěšně rozvíjen v ekonomii. Možnost analyzovat finanční řady pomocí nelineárních metod se v současné době detailně zkoumá v řadě odborných článků v prestižních časopisech. Kdyby se v průběhu doby naplněné náročným zkoumáním možností deterministického popisu finančních časových řad skutečně prokázalo, že to možné je, pak by bylo možné dynamiku časových řad krátkodobě předvídat. Z hlediska dlouhodobého by to ovšem tak jednoduché nebylo. Důvodem je exponenciální rozbíhavost dostatečně blízkých

* Stať byla zpracována za finanční podpory GAČR v rámci grantů č. 402/06/0209 a 402/06/0990.

1 Centrum základního výzkumu pro dynamickou ekonomii a ekonometrii je zřízeno Ministerstvem školství, mládeže a tělovýchovy České republiky prostřednictvím projektu LC 06075.

trajektorií nelineárního dynamického systému spolu s nepřesností měření. V některých případech je možno sestavit atraktor a ten použít pro dlouhodobou předpověď.

Modelování časových řad užitím nelineárních deterministických modelů je poměrně složitý úkol, který navíc vyžaduje velké množství vstupních dat. V realitě tento požadavek není vždy splnitelný. Určitým východiskem z tohoto problému je technika rekurentních grafů, se kterou přišli Eckmann, Kamphorst a Ruelle (1987). Tato technika nám umožní poměrně snadno a rychle rozhodnout, zda předemná časová řada je generována určitým deterministickým procesem. Podstata této techniky spočívá v tom, že rekurenční graf zkoumá, jestli jednotlivé stavy příslušného dynamického systému rekonstruované z dané časové řady jsou sobě blízké případně od sebe vzdálené podle určitého předem zvoleného prahu. Grafickou interpretaci těchto vzdáleností nazýváme rekurenční graf. Výsledný vzor výskytu těchto stavů, charakterizovaných vzdálenostmi a zaznamenaných na rekurenčním grafu, nám poskytuje důležité informace o chování zkoumaného systému.

Interpretace rekurenčního grafu se může mnohdy jevit jako příliš subjektivní. Proto Zbilut a Webber (1992) navrhli postup, kterým se snaží kvantifikovat informace obsažené v rekurenčních grafech, aby získali přesnější výsledky. Jedná se o několik důležitých statistik, které podávají cenné informace o charakteru generujícího procesu. V tomto případě hovoříme o kvantitativní rekurenční analýze, která je hojně využívána v biomedicině a geofyzice. Aplikace tohoto postupu na finančních datech je však poměrně vzácná. Lze s ní setkat v práci Antonioua a Vorlowa (2000), Fabrettiho a Ausloose (2005) a Schwabeho (2007).

V této práci bychom se rádi pokusili o aplikaci kvantitativní rekurenční analýzy na datech získaných z českého akciového trhu jako alternativní přístup na zjištění deterministického chování finančních řad k přístupu využitému v pracích Trana (2007) a Trešla (2008)². Determinace přítomnosti případného deterministického procesu bude ověřena na řadách akciového indexu PX a nejlépe viditelných akciových titulů. Jedná se o akcie Komerční banky (KB), akcie Českých energetických závodů (ČEZ) a akcie Telefonica O₂ (dřívější Český Telecom) v období cca 12 let od roku 1995 do roku 2007.

Struktura práce je následující. Po ukončení této úvodní části, zahájíme část druhou, kde se stručně věnujeme technice rekonstrukce fázového prostoru dynamického systému. V další, třetí části bude vysvětlena podstata kvantitativní rekurenční analýzy a ve čtvrté části budou prezentovány výsledky aplikace rekurenční analýzy na finančních časových řadách z českého akciového trhu. V poslední části budou shrnuty závěry.

2. Rekonstrukce trajektorie dynamického systému ve fázovém prostoru

Rekurenční graf je grafický nástroj zkoumající dynamiku systému a je založen na rekonstrukci jeho trajektorie ve fázovém prostoru. Na rekurenčním grafu nejsou znázorněny vzdálenosti jednotlivých bodů z dané časové řady, ale jsou tam zaznamenány vzdálenosti jednotlivých stavů zkoumaného systému ve fázovém prostoru. Protože

2 O využití teorie katastrof při modelování náhlých změn dynamiky kurzů se pokusili Vošvrda a Baruník (2008). Dalším přístupem v této oblasti je využití *waveletů*, který můžeme nalézt např. v práci autorů Váchy a Vošvrdy (2007).

neznáme explicitně celou soustavu řídicích rovnic systému, je nutné z jedné zkoumané časové řady, která je vlastně jakousi projekcí celé trajektorie do jedné osy fázového prostoru, rekonstruovat celý původní systém v prostoru. Nejčastější metodou rekonstrukce trajektorie dynamického systému ve fázovém prostoru je metoda časových zpoždění. V rekonstruovaném prostoru je stavový vektor $y(t)$ vyjádřen v závislosti na hodnotách analyzované řady $x(t)$ následujícím způsobem:

$$y(t) = (x(t), x(t-\tau), \dots, x(t-(m-2)\tau), x(t-(m-1)\tau)). \quad (1)$$

V tomto vztahu figurují dvě důležité konstanty: časové zpoždění τ a dimenze vnoření m , kde τ udává časovou vzdálenost mezi sousedními prvky a m udává celkovou dimenzi atraktoru. Veličiny τ a m jsou tedy parametry vnoření, na kterém je rekonstrukce fázového prostoru založena (Takens, 1985). Teorie vnoření předpokládá nekonečně dlouhou řadou měřenou s absolutní přesností. Ve skutečnosti však pracujeme s časovými řadami, které mají konečný počet pozorování a jsou zatíženy určitými chybami při jejich měření. Proto určení vhodných hodnot těchto parametrů je velice důležité pro rekonstrukci fázového prostoru, přičemž při rekonstrukci chceme vytvořit takovou strukturu, která popisuje úplně dynamiku systému bez nadbytečných parametrů. Myšlenka úplného popisu systému s minimálním počtem parametrů je velmi důležitá. Podrobnější informace o rekonstrukci fázového prostoru a jejích principech najdeme v publikacích Kantz (1997), Sprott (2003) a Horák (2003).

2.1. Volba vhodné délky časového zpoždění

Pro zvolení správné délky časového zpoždění se doporučují v literatuře různé přímé postupy (Horák, 2003) a zvolená hodnota má přímý vliv na následnou rekonstrukci fázového prostoru. Pokud je tato hodnota příliš malá, rozdíl mezi jednotlivými stavy rekonstruovaného prostoru bude nepatrný a význam informací získaných tímto způsobem je pak malý. Na druhé straně, pokud tato hodnota je příliš velká, může se stát, že jednotlivé stavy rekonstruovaného systému budou vnímány jako nezávislé. Jinak řečeno, volbou dlouhé časové vzdálenosti mezi jednotlivými stavy systému se ztratí souvislost jeho dynamiky a jako důsledek se takto rekonstruovaná trajektorie jeví jako náhodný proces.

Ve starší odborné literatuře se doporučovalo, aby za délku časového zpoždění byla zvolena taková hodnota, kdy autokorelační funkce původní časové řady klesá poprvé pod určitou mez, např. $1/e$, $1/10$ nebo 0 (Horák, 2003).³ Autokorelační funkce však zohledňuje pouze lineární závislost ve zkoumané časové řadě, což znamená, že nebere v úvahu možnost nelineárních procesů. Z tohoto důvodu vhodnější kritérium pro optimální hodnotu časového zpoždění je míra vzájemné informace. Míra vzájemné informace představuje průměrné množství informace, kterou lze z hodnot náhodné veličiny X získat o hodnotách náhodné veličiny Y . Vyjádříme ji následujícím vztahem

³ Kde e je Eulerovo číslo.

$$I(X, Y) = \sum_{x_i, y_i} p(x_i, y_i) \log_2 \left(\frac{p(x_i, y_i)}{p(x_i) p(y_i)} \right), \quad (2)$$

kde $I(X, Y)$ je míra vzájemné informace dvou náhodných veličin X a Y , $p(x_i)$ je pravděpodobnost, že $X = x_i$, $p(y_i)$ je pravděpodobnost, že $Y = y_i$ a $p(x_i, y_i)$ je pravděpodobnost, že současně $X = x_i$ a $Y = y_i$. Obdobně jako vztah autokorelačního koeficientu ke korelačnímu koeficientu můžeme definovat časově zpožděnou míru vzájemné informace k míře vzájemné informace ve tvaru

$$I(\tau) = \sum_{x_i, x_{i+\tau}} p(x_i, x_{i+\tau}) \log_2 \left(\frac{p(x_i, x_{i+\tau})}{p(x_i) p(x_{i+\tau})} \right), \quad (3)$$

kde τ je délka časového zpoždění. Pokud jednotlivá pozorování jsou navzájem nezávislá, hodnota míry vzájemné informace mezi nimi je rovna nule. Jako vhodná délka časového zpoždění pro rekonstrukci fázového prostoru se doporučuje taková hodnota τ , kdy míra vzájemné informace dosáhne prvního minima. Toto minimum znamená, že pozorování $x(t_i + \tau)$ v čase $t_i + \tau$ přináší v průměru nejvyšší informační příspěvek k informaci, kterou máme k dispozici z pozorování $x(t_i)$ v čase t_i a redukuje redundanci na minimum.

2.2. Volba vhodné dimenze vnoření

Vhodnou dimenzi vnoření při rekonstrukci fázového prostoru je možné volit několika způsoby. Jedním z nich je metoda nejbližších falešných sousedů (Kennel a kol., (1992)). Podstata této metody spočívá v tom, že při projekci trajektorie systému z originálního fázového prostoru do prostoru s nižší dimenzí dochází k tzv. překřížení této trajektorie se sebou samou. Kvůli tomuto překřížení vznikají tzv. falešní sousedé. Při zvyšování dimenze vnoření se počet těchto falešných sousedů postupně snižuje. Zcela vymizí v případě, že dimenze vnoření použitá při rekonstrukci je rovna skutečné dimenzi originálního fázového prostoru. Skuteční sousedé jsou blízko sebe při jakékoli dimenzi vnoření, samozřejmě i v původním fázovém prostoru. Metoda falešných nejbližších sousedů je potom založena na detekci jejich počtu při zvýšení dimenze vnoření.

Předpokládejme, že jsme v m -dimensionálním prostoru, kde je rekonstruován stav $y_m(t_i)$

$$y_m(t_i) = (x(t_i), x(t_i - \tau), \dots, x(t_i - (m-2)\tau), x(t_i - (m-1)\tau)). \quad (4)$$

kde τ je vhodná délka zpoždění. Ke stavu $y_m(t_i)$ je nejbližší soused označený jako

$$y_m^{NN}(t_i) = (x^{NN}(t_i), x^{NN}(t_i - \tau), \dots, x^{NN}(t_i - (m-2)\tau), x^{NN}(t_i - (m-1)\tau)). \quad (5)$$

Pokud je $y_m^{NN}(t_i)$ falešným sousedem, při zvýšení dimenze vnoření dojde ke zvětšení vzdálenosti mezi těmito stavy natolik, že již nelze je považovat za sousedy. Zvýšení dimenze vnoření o jednu se souřadnice těchto dvou stavů změní na

$$y_{m+1}(t_i) = (x(t_i), x(t_i - \tau), \dots, x(t_i - (m-1)\tau), x(t_i - m\tau)) \quad \text{a}$$

$$y_{m+1}^{NN}(t_i) = (x^{NN}(t_i), x^{NN}(t_i - \tau), \dots, x^{NN}(t_i - (m-1)\tau), x^{NN}(t_i - m\tau)).$$

Euklidovská vzdálenost těchto dvou stavů ve fázovém prostoru o dimenzi vnoření m je:

$$d_m(t_i) = \sqrt{\|y_m(t_i) - y_m^{NN}(t_i)\|^2} = \left(\sum_{k=1}^m (x(t_i - (k-1)\tau) - x^{NN}(t_i - (k-1)\tau))^2 \right)^{\frac{1}{2}}, \quad (6)$$

a v $(m+1)$ -rozměrném fázovém prostoru je

$$d_{m+1}(t_i) = \sqrt{\|y_{m+1}(t_i) - y_{m+1}^{NN}(t_i)\|^2} = \left(\sum_{k=1}^{m+1} (x(t_i - (k-1)\tau) - x^{NN}(t_i - (k-1)\tau))^2 \right)^{\frac{1}{2}}. \quad (7)$$

Tedy:

$$d_{m+1}^2(t_i) = d_m^2(t_i) + (x(t_i - m\tau) - x^{NN}(t_i - m\tau))^2. \quad (8)$$

Pro to, abychom mohli rozhodnout, zda $y_m(t_i)$ a $y_m^{NN}(t_i)$ jsou skuteční nebo falešní sousedé, definujeme poměr $d(t_i)$ ve tvaru

$$d(t_i) = \sqrt{\frac{d_{m+1}^2(t_i) - d_m^2(t_i)}{d_m^2(t_i)}} = \frac{\|x(t_i - m\tau) - x^{NN}(t_i - m\tau)\|}{d_m(t_i)}. \quad (9)$$

Jestliže jeho hodnota překročí určitou prahovou hodnotu, budou sousedé považovány za falešné sousedy. Za dostatečnou dimenzi vnoření bereme hodnotu, která snižuje počet falešných sousedů na nulu. Další možnost je určit, zda přírůstek vzdálenosti mezi $y_m(t_i)$ a $y_m^{NN}(t_i)$ při přechodu z dimenze vnoření m na dimenzi $m+1$ nepřekročí nějakou předem zvolenou prahovou hodnotu d_p , neboli

$$\|x(t_i - m\tau) - x^{NN}(t_i - m\tau)\| \geq d_p, \quad (10)$$

kde d_p volíme jako

$$d_p = \frac{1}{N} \sum_{i=1}^N (x(t_i) - \bar{x}) \quad \text{a} \quad \bar{x} = \frac{1}{N} \sum_{i=1}^N x(t_i). \quad (11)$$

Volba výše uvedené prahové hodnoty představuje určitý problém. Proto Cao (1997) navrhol následující zlepšení. Místo vztahů (9) a (10) definuje poměr euklidovských vzdáleností dvou sousedních stavů v m - a $(m+1)$ -rozměrném prostoru

$$a(t_i, m) = \frac{\|y_{m+1}(t_i) - y_{m+1}^{NN}(t_i)\|}{\|y_m(t_i) - y_m^{NN}(t_i)\|}, \quad (12)$$

a zavádí veličinu $E(m)$ jako aritmetický průměr všech hodnot, přičemž

$$E(m) = \frac{1}{N - m\tau} \sum_{i=1}^{N-m\tau} a(t_i, m). \quad (13)$$

Při dostatečně velké hodnotě m nebude $E(m)$ dále klesat, protože počet falešných sousedů bude redukován na minimum.

3. Rekurenční graf a kvantitativní rekurenční analýza

3.1. Rekurenční graf

Fázový prostor daného systému obvykle nemusí být pouze dvojrozměrný nebo trojrozměrný, ale také vícerozměrný, což znemožňuje grafické zobrazení. Vyšší dimenze lze vizualizovat pouze projekcí do dvoj- nebo trojrozměrného prostoru. Eckmann a kol. (1987) vynalezli nový postup, jak vizualizovat rekurenci stavů ve fázovém prostoru a nazvali jej rekurenční graf (recurrence plot). Tento postup umožňuje zkoumat trajektorii m -rozměrného fázového prostoru prostřednictvím její rekurence ve dvojrozměrném zobrazení. Rekurence určitého stavu v čase i v jiném čase j je zaznamenána ve čtvercové matici jedničkou, jeho nerekurence nulou, v rekurenčním grafu tmavou a světlou barvou, kde obě osy jsou časové. Tento vztah můžeme formálně zapsat jako

$$R_{i,j} = \Theta(\varepsilon - \|y_i - y_j\|), \text{ pro } i, j = 1, 2, \dots, N, \quad (14)$$

kde $\Theta(\cdot)$ je Heavisideova funkce, která nabývá hodnotu 1, když $\varepsilon - \|y_i - y_j\| > 0$ a 0 v ostatních případech, ε je prahová vzdálenost, $\|\cdot\|$ je vzdálenost dvou stavů v m -rozměrném fázovém prostoru, (obvykle je to euklidovská vzdálenost, ale euklidovská nemusí být vždycky), y_i, y_j jsou jednotlivé stavy a N je celkový počet stavů.

Rekurenční graf je z podstaty symetrický podle diagonály, tedy linie se sklonem 45° a každý rekurenční graf by obvykle měl zahrnovat jednu až všechny z těchto základních struktur: samostatné body, diagonální linie a horizontální nebo vertikální linie, přičemž tyto struktury mají následující význam:

- jednotlivé samostatné body znamenají, že stavy ve fázovém prostoru jsou jedinečné, systém nesetrvává na nich dlouho nebo silně fluktuuje,
- diagonální linie $L_{i+k, j+k} = l$, pro $k = 1, 2, \dots, l$, kde l je délka jedné diagonální linie se vyskytují v případě, když část trajektorie běží paralelně s jinou částí trajektorie, neboli trajektorie se vrací do stejné oblasti v různých časech. Tyto linie indikují existenci nestabilních periodických orbit a jsou tedy charakteristickým znakem přítomnosti determinismu. Pokud na rekurenčním grafu existují pouze ty-

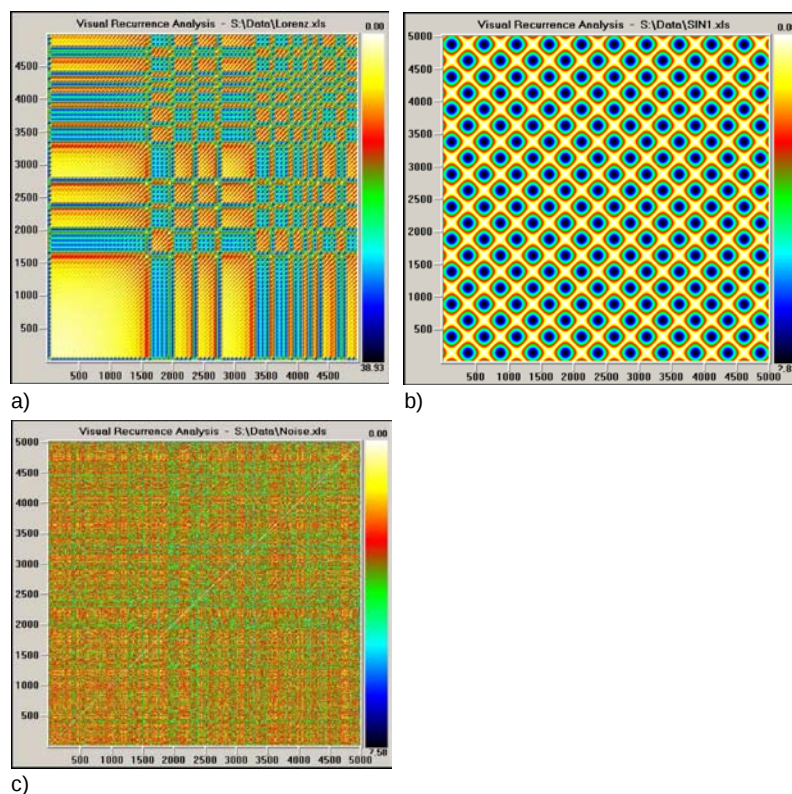
to diagonální linie, jedná se pak o periodický signál (vzdálenosti od jednoho stavu se pravidelně opakují). Délka tohoto úseku je určena dobou, jak dlouho jednotlivé části trajektorie setrvává v těchto oblastech,

- vertikální nebo horizontální linie $L_{i,j+k} = l$ pro $k = 1, 2, \dots, l$ znamenají, že stavy setrvávají v jednom bodě nebo se jen málo mění z tohoto bodu, systém je tedy uvězněn v těchto bodech.

U takto konstruovaného rekurenčního grafu musí být vždy předem zvolena prahová hodnota ε . V poslední době díky pokroku ve výpočetní technice je možné obejít tento problém tím, že se používá barevný rekurenční graf, na kterém jsou zobrazeny všechny vzdálenosti od jednoho stavu (který je na diagonále) a na pravé straně rekurenčního grafu je tzv. mapa barev, jež definuje tyto vzdálenosti v odpovídající barvě. Příklady rekurenčních grafů některých systémů jsou uvedeny na obrázku 1. Tyto příklady není bohužel možné na tomto místě uvést v barvách, pouze v odstínech, což poněkud ztěžuje orientaci. Čtenáře, který se zajímá o barevnou podobu obrázku, odkazujeme na webové stránky Centra základního výzkumu pro dynamickou ekonomii a ekonometrii (<http://czvdee.utia.cas.cz/odkazy>).

Obrázek 1

Rekurenční grafy a) Lorenzova systému, b) periodického sinusoidálního signálu, a c) bílého šumu



Na rekurenčním grafu Lorenzova systému u sinusoidálního signálu je jasné vidět diagonální linie prokazující jeho periodicitu a na rekurenčním grafu bílého šumu nelze pozorovat žádnou jasnou strukturu.

3.2. Kvantitativní rekurenční analýza

Rekurenční graf je poměrně užitečný nástroj pro rychlou detekci přítomnosti možné deterministické nelinearity ve zkoumaných časových řadách. Má však jednu nevýhodu. Jeho interpretace závisí pouze na vizuální schopnosti člověka rozlišovat různé vzory na těchto grafech a mohou existovat určité vzory, které člověk nemusí být schopen rozlišit nebo může podléhat tzv. optickým klamům. Abychom předešli tento nedostatek a mohli dostávat z rekurenčních grafů jednoznačnou interpretaci výsledků analýzy, byla vytvořena tzv. kvantitativní rekurenční analýza (Zbilut a Webber (1992), Marvan (2002)). Kvantitativní rekurenční analýza kvantifikuje vlastnosti rekurenčních grafů několika statistikami, které budou vymezeny níže. Dříve, než budeme definovat tyto statistiky, nadefinujeme některé pomocné statistiky. Jsou to

$$R_{ij} = \begin{cases} 1 & \text{pokud stavy } (i, j) \text{ jsou rekurentní} \\ 0 & \text{jinak} \end{cases}$$

$$D_{ij} = \begin{cases} 1 & \text{pokud } (i-1, j-1), (i, j) \text{ a } (i+1, j+1) \text{ jsou rekurentní} \\ 0 & \text{jinak} \end{cases}$$

$$V_{ij} = \begin{cases} 1 & \text{pokud } (i-1, j-1), (i, j) \text{ a } (i+1, j+1) \text{ jsou rekurentní} \\ 0 & \text{jinak} \end{cases}$$

$$RR_i = \sum_{j=1}^N R_{ij}.$$

Stavy i a j jsou rekurentní, pokud vzdálenost mezi nimi je menší než prahová hodnota. Statistika R_{ij} určí, zda stavy i a j jsou rekurentní, D_{ij} je délka diagonálního úseku složeného ze tří stavů a podobně V_{ij} je délka vertikálního úseku složeného ze tří stavů a RR_i je počet rekurentních stavů vzhledem ke stavu i . Necht' d je délka jednoho diagonálního úseku a v je délka jednoho vertikálního úseku, pak nadefinujeme ještě dvě pomocné statistiky

$$P(d) = \sum_{i,j=1}^N (1-R_{i-1,j-1})(1-R_{i+1,j+1}) \prod_{k=0}^{d-1} R_{i+k,j+k}$$

respektive

$$P(v) = \sum_{i,j=1}^N (1-R_{i,j})(1-R_{i,j+v}) \prod_{k=0}^{v-1} R_{i,j+k},$$

kde $P(d)$ a $P(v)$ jsou pravděpodobnosti výskytu diagonálního nebo vertikálního úseku délky d respektive v v rekurenčním grafu. Pro kvantitativní rekurenční analýzu jsou vypočteny tyto statistiky:

- rekurentní míra RR , která je definována jako podíl rekurentních stavů v rekurenčním grafu a představuje pravděpodobnost, že se určitý stav znovu objeví. Je velmi podobná korelačnímu integrálu s jedním rozdílem, že se u korelačního integrálu⁴ nepočítá stav, kdy $i = j$, neboli

$$RR = \frac{1}{N^2} \sum_{i,j=1}^N R_{ij} \quad (15)$$

- míra determinismu DET , která je definována jako podíl stavů tvořících diagonální úseky na celkovém počtu rekurentních stavů. Míra determinismu je spojena s předvídatelností dynamického systému. Čím vyšší je hodnota míry determinismu, tím vyšší je možnost předvídat chování systému.

$$DET = \frac{\sum_{d=d_{\min}}^N d P(d)}{\sum_{i,j=1}^N R_{ij}} \quad (16)$$

- laminarita, která je definována jako podíl stavů tvořících vertikální úseky na celkovém počtu rekurentních stavů. Laminarita je spojena s počtem ojedinelých stavů v celém systému, neboť

$$LAM = \frac{\sum_{v=v_{\min}}^N v P(v)}{\sum_{i,j=1}^N R_{ij}}, \quad (17)$$

trend,⁵ který je regresní koeficient lineárních závislostí mezi hustotou rekurentních stavů v jedné linii paralelní s diagonálou a její vzdáleností k diagonále a poskytuje informace o stacionaritě systému. Pro rekurentní míru v diagonální linii, která je para-

4 Korelační integrál (suma) nějakého dynamického systému $C(N, m, \varepsilon)$ s N pozorovanými stavy při dimenzi vnoření m se počítá vztahem

$$C(N, m, \varepsilon) = \frac{1}{N(N-1)} \sum_{1 \leq i < j \leq N} \Theta(\varepsilon - \|y_i - y_j\|),$$

kde ε je nějaké předem zvolené malé číslo, $\Theta(\cdot)$ je Heavisideova funkce, která nabývá hodnotu 1, když její argument je kladný a 0 když její argument je nekladný. Symbol $\|\cdot\|$ značí vzdálenost mezi dvěma zvolenými stavy. Korelační integrál je tedy podíl stavů systému y_i a y_j takových, že jsou blízké sobě, blízkost znamená, že vzdálenost mezi nimi je menší než předem zvolená malá hodnota ε .

5 $TREND$ je tedy směrnice regresní přímky $\hat{RR}_i = \hat{\beta}_0 + \hat{\beta}_1 i$, a tudíž $TREND = \hat{\beta}_0$.

lelní s diagonální linií se sklonem 45° ve vzdálenosti i , tedy RR_i , je *TREND* definován vztahem

$$TREND = \frac{\sum_{i=1}^{\bar{N}} \left(i - \frac{\bar{N}}{2} \right) (RR_i - \overline{RR_i})}{\sum_{i=1}^{\bar{N}} \left(i - \frac{\bar{N}}{2} \right)^2}, \quad (18)$$

kde $\overline{RR_i}$ a $\bar{N}$ jsou průměrné hodnoty všech RR_i a $\bar{N}$, kde $\bar{N} < N$. Tato podmínka se vyžaduje, aby se vyloučily stavy s nižší hustotou, které se vyskytují na okraji rekurvenčního grafu.

Protože délka diagonálních linií je spojena s dobou, kdy části trajektorie systému běží vedle sebe, považuje se (Marvan, 2002) převrácená hodnota nejdelšího diagonálního úseku jako přibližný odhad nejvyššího Ljapunovova exponentu dynamického systému. Nejdelší diagonální úsek je definován jako:

$$MAXLINE = \max(\{d_i; i = 1, \dots, N_d\}) \quad (19)$$

Pravděpodobnost, že diagonální úsek bude mít přesně délku d , je $p(d)$, kde

$$p(d) = \frac{P(d)}{\sum P(d)}.$$
 Entropie,⁶ která obecně poskytuje informace o komplexnosti

deterministické struktury tohoto systému, je definována jako

$$ENTR = - \sum_{d=d_{\min}}^N p(d) \ln(p(d)) \quad (20)$$

dobu v pasti *TT* (trapping time) je definována jako průměrná délka všech vertikálních linií a představuje průměrnou dobu, kterou se systém zdržuje v jednom ojedinělém stavu, tj.

$$TT = \frac{\sum_{v=v_{\min}}^N v p(v)}{\sum_{v=v_{\min}}^N p(v)}. \quad (21)$$

6 Existuje několik typů entropie. Kolmogorovova entropie je důležitým kvantifikátorem determinismu dynamického systému a udává, jakou rychlostí ztrácíme informace o chování našeho systému. Představuje tedy míru jeho předpověditelnosti a tato míra je úměrná převrácené hodnotě entropie. Pokud dynamický systém má nulovou entropii, je periodický. V případě, že má konečnou entropii, je chaotický a je-li entropie nekonečná, systém je pak čistě náhodný. Zde se však jedná o Shannonovu entropii pocházející z oblasti informační teorie aplikovanou na dynamický systém, konkrétně na pravděpodobnost výskytu linie určitého typu. Je patrné, že se tento vztah velmi podobá vztahu pro výpočet míry vzájemné informace. Čím vyšší je hodnota entropie (6-8), tím je systém uspořádanější (periodičtější) a tím méně informací se potřebuje k jeho rekonstrukci. Naopak nižší hodnota entropie indikuje větší chaotičnost a tudíž je potřeba více informací k jeho rekonstrukci.

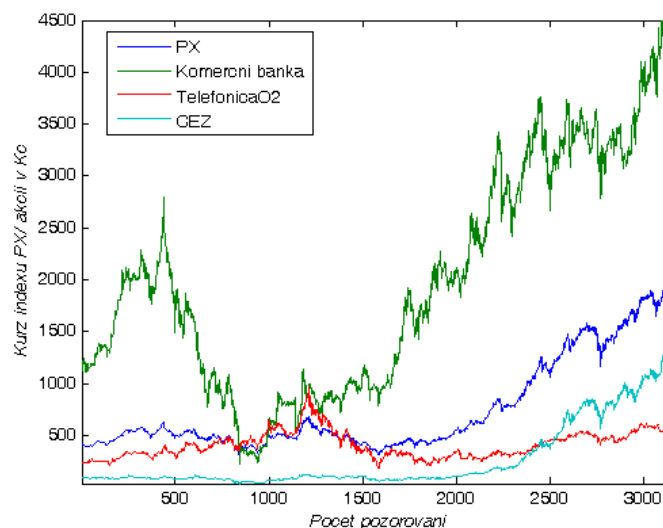
4. Aplikace vizuální rekurenční analýzy na datech z českého akciového trhu

Pro aplikaci kvantitativní rekurenční analýzy jsme zvolili časové řady výnosností akciového indexu PX a tří nejlikvidnějších akciových titulů na českém akciovém trhu. Jsou to řady výnosností akcií Telefonica O₂ (O2), Komerční banky (KB) a Českých energetických závodů (CEZ). Jedná se o řady denních uzavíracích cen v období od ledna 1995 do listopadu 2007, tj. období cca 12 let. Počet pozorování je tedy dostatečně veliký. Vývoj denních kurzů zkoumaných akciových titulů a indexu PX je zobrazen v obrázku 2. Na tomto obrázku je patrný silný pokles cen akcií Komerční banky, který se zastavil až u hodnoty 229 Kč na konci roku 1998. Pak následoval prudký růstový trend, který trvá až do současnosti. u kurzů akciového indexu PX a akcií ČEZ lze pozorovat poměrně silný růstový trend v posledních letech. u hodinových kurzů, které byly zaznamenány až od roku 1998, silný sestupný trend cen akcií Komerční banky již nenastal a až na lokální maxima dosažená v roce 2000 je u všech akcií a indexu PX patrný růstový trend.

Data z výše vybraných časových řad byla následně transformována na odpovídající řady logaritmických výnosových měr. Touto transformací se odstraňuje vlastnost nezápornosti cen akcií a indexu a získáváme nové časové řady s relativně symetrickým rozdělením s akceptovatelnou úrovní stacionarity. Na zkoumaných časových řadách denních kurzů akcií a akciového indexu PX a na řadách jejich výnosností byla nejdříve provedena analýza jejich základních statistických vlastností v programu Eviews5. Základní deskriptivní statistiky časových řad denních kurzů a výnosností zkoumaných akciových titulů a indexu PX jsou uvedeny v tabulce 1 a 2.

Obrázek 2:

Vývoj denních kurzů indexu PX a zkoumaných akcií



Tabulka 1
Základní deskriptivní statistiky zkoumaných časových řad

	PX	KB	O2	CEZ
Průměr	742,88	1960,42	412,03	263,69
Medián	523,40	1810,42	386,75	100,09
Maximum	1936,90	4500,00	967,30	1410,00
Minimum	316,00	229,17	184,80	38,55
Std odchylka	441,29	1081,66	126,63	316,79
Šikmost	1,30	0,41	0,85	1,71
Špičatost	3,23	1,97	3,67	4,64
Jarque-Bera	906,70	229,24	433,72	1918,41
P-value	0,00	0,00	0,00	0,00
Počet pozorování	3124	3124	3124	3124

Tabulka 2
Základní deskriptivní statistiky časových řad výnosností

	PX	KB	CEZ	O2CR
Průměr	0,0004	0,0004	0,0007	0,0003
Medián	0,00	0,00	0,00	0,00
Maximum	0,0705	0,2327	0,1671	0,1134
Minimum	-0,0708	-0,2409	-0,1983	-0,1617
Std odchylka	0,0119	0,0255	0,0211	0,0210
Šikmost	-0,29	-0,42	-0,28	-0,19
Špičatost	5,79	13,93	9,74	7,56
Jarque-Bera	1052,93	15629,40	5946,23	2718,77
P-value	0,00	0,00	0,00	0,00
Počet pozorování	3123	3123	3123	3123

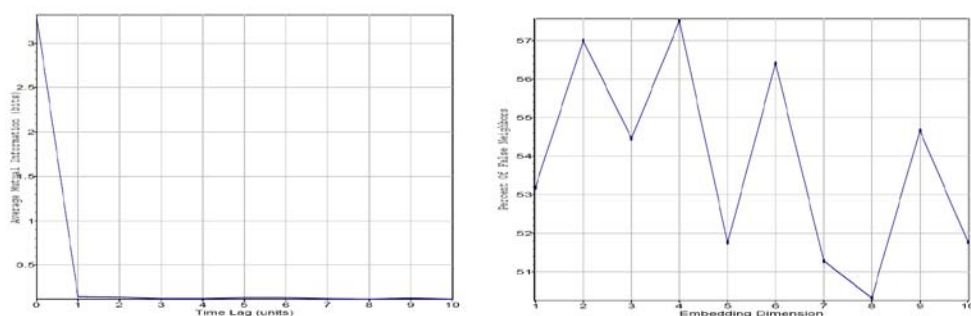
Z výsledků deskriptivních statistik řad denních výnosností zkoumaných akcií a akciového indexu PX je možné pozorovat, že se v sledovaném období výnosnosti akciového indexu PX pohybují v mnohem užším intervalu $(-0,07, 0,07)$ než výnosnosti individuálních akcií: u akcií Telefonica O₂ je to $(-0,16, 0,11)$, u akcií ČEZ je to $(-0,20, 0,18)$ a u akcií Komerční banky je tento interval nejširší $(-0,24, 0,23)$. Tomu odpovídají i hodnoty standardních odchylek denních výnosností. Výsledky Jarqueova-Berova testu ukazují, že rozdělení výnosností není normální, protože hodnoty šikmosti a špičatosti se statisticky značně liší od hodnot pro normální rozdělení (tabulka 2). Pro úplnost je třeba konstatovat, že se průměrná hodnota denních výnosností sledovaných akcií a akciového indexu PX pouze nepatrně liší od 0 v sledovaném období.

Abychom mohli rekonstruovat trajektorii systému generujícího naše řady výnosností ve fázovém prostoru, je zapotřebí nejdříve zjistit vhodné hodnoty parametrů vnoření τ a d . Optimální hodnoty délky časového zpoždění jsou zvoleny na základě míry vzájemné informace získané z řad denních výnosností zkoumaných akciových titulů

a indexu českého akciového trhu PX. Vhodné hodnoty dimenze vnoření jsou zvoleny podle metody nejbližších falešných sousedů. Tyto parametry jsou vypočteny v programovém balíku VRA⁷ a výsledky jsou uvedeny na obrázcích 3, 4, 6 a 7. Na těchto obrázcích nalevo jsou umístěny grafy závislosti míry vzájemné informace na délce časového zpoždění, kde na svislé ose je míra vzájemné informace a na vodorovné ose je délka časového zpoždění. Napravo jsou grafy znázorňující závislost podílu falešných sousedů na dimenzi vnoření, kde na vertikální ose je podíl falešných sousedů (v %) a na horizontální ose je dimenze vnoření. Je na nich patrný pomalý pokles průměrné míry informací získaných ze zpožděných pozorování. Vhodná délka časového zpoždění je zvolena hodnota, kdy vzájemná míra informace dosahuje prvního minima a v případě vhodné dimenze vnoření bude vždy zvolena hodnota, která snižuje počet falešných sousedů na minimum⁸ (viz výše uvedené obrázky).

Obrázek 3

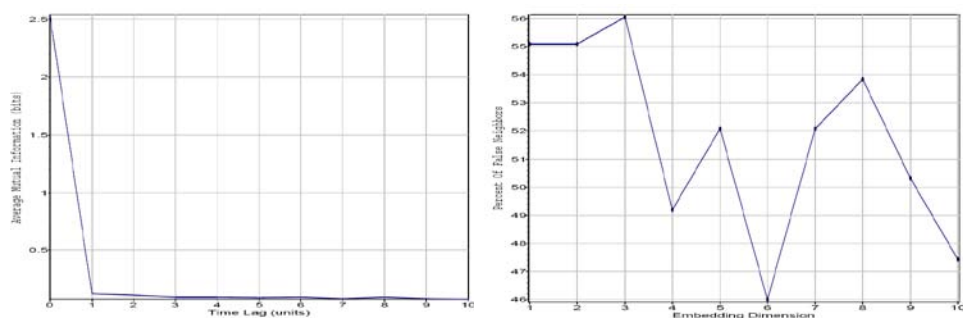
Časově zpožděná míra vzájemné informace a závislost počtu falešných sousedů na dimenzi vnoření řady výnosností akciového indexu PX



- 7 Visual Recurrence Analysis (VRA) je volně dostupný program vytvořený E. Korolovem v jazyku C++. VRA programový balík je řízen poměrně bohatou nabídkou různých funkcí pro lineární analýzu časových řad a je tedy velice uživatelsky komfortní. Lze jej využít jak ke kvalitativní a kvantitativní analýze, tak i k predikci nelineárních časových řad. Existuje rovněž komerční verze programu VRA, která na rozdíl od volně dostupné verze umožňuje uživateli ukládat predikované hodnoty.
- 8 Při zjištění, zda je nějaký stav falešný nebo ne, je třeba zvolit podle Kennelova postupu prahovou hodnotu. Ve VRA programovém balíku se tento problém obešel tímto postupem. V m -rozměrném prostoru ke každému stavu $y_m(t_i) = (x(t_i), x(t_i - \tau), \dots, x(t_i - (m-2)\tau), x(t_i - (m-1)\tau))$ najdeme jeho nejbližšího souseda $y_m^{NN}(t_i) = (x^{NN}(t_i), x^{NN}(t_i - \tau), \dots, x^{NN}(t_i - (m-2)\tau), x^{NN}(t_i - (m-1)\tau))$. Obdobně postupujeme v $(m+1)$ -rozměrném prostoru a vypočítáme vzdálenost $d = \|y_{m+1} - y_{m+1}^{NN}\|_2$. Tato vzdálenost současně je vzdálenost mezi obrazy y a y^{NN} . Když uvažujeme y_{m+1}^{NN} jako prediktor pro y_{m+1} , pak d je predikční chyba. Pokud je dimenze vnoření dostatečně velká, vzdálenost d mezi $(m+1)$ složkami stavů y a y^{NN} bude malá. Abychom určili, zda dva předmětní sousedi jsou falešní nebo ne, porovnáváme vzdálenost d s vzdáleností získanou triviálním prediktorem d_t . Jako triviální prediktor volíme jako y_m pro y_{m+1} . Pokud $d_t < d$, považujeme tento sousední stav za falešný. Formálně tedy: pokud $\|y_{m+1} - y_{m+1}^{NN}\|_2 \geq \|y_{m+1} - y_m\|_2$, je nalezený soused považován za falešného.

Obrázek 4

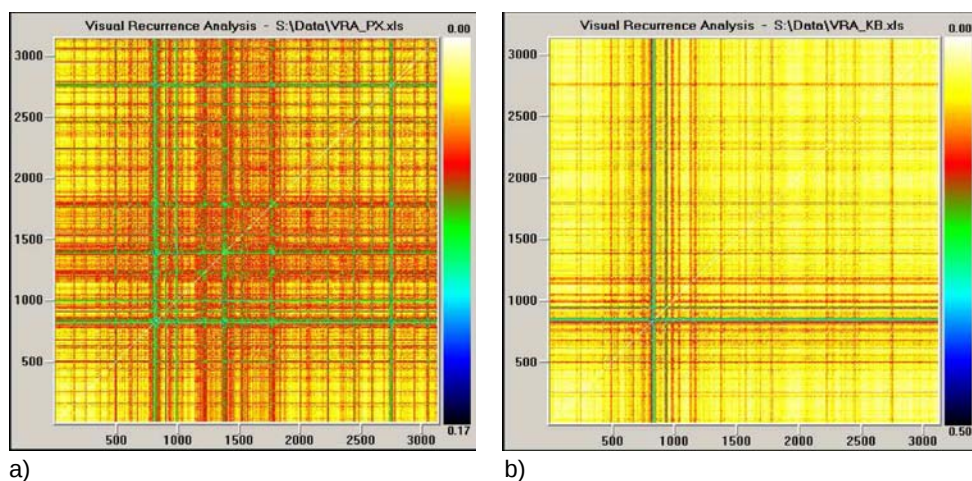
Časově zpožděná míra vzájemné informace a závislost počtu falešných sousedů na dimenzi vnoření řady výnosností akcií Komerční banky KB



Na obrázcích 5 a 8 jsou uvedeny rekurenční grafy zkoumaných časových řad výnosností vytvořené programem VRA (*Visual Recurrence Analysis*, Koronov, 2007). Na všech obrázcích není patrná žádná diagonální struktura, což značí, že neexistuje periodické chování v těchto řadách. Na rozdíl od původního přístupu týkajícího se rekurenčního grafu, v současnosti díky schopnosti soudobé výpočetní techniky na rekurenčním grafu lze zobrazovat vzdálenost jednoho stavu od ostatních a tyto vzdálenosti na diagramu vyznačit barvou, v případě černobílého obrázku odstínem. Hodnoty těchto vzdáleností je možné vyčíst z mapy odstínů umístěné napravo vedle rekurenčního grafu. Na získaných grafech jsou patrné svislé sloupce, což značí místa, kde se dané stavy systému odchylují od ostatních stavů na větší vzdálenost.

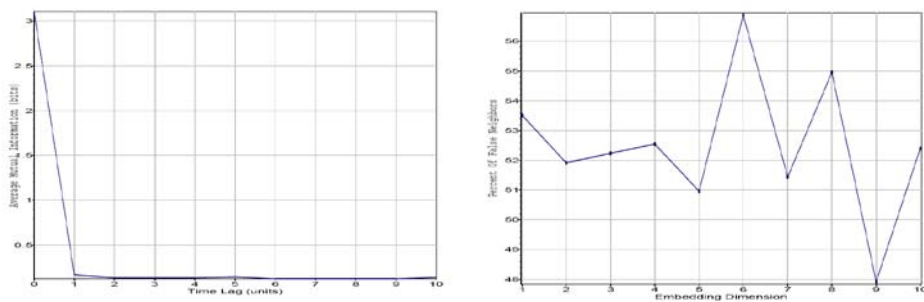
Obrázek 5

Rekurenční graf řady výnosností a) akciového indexu PX a b) Komerční banky KB.



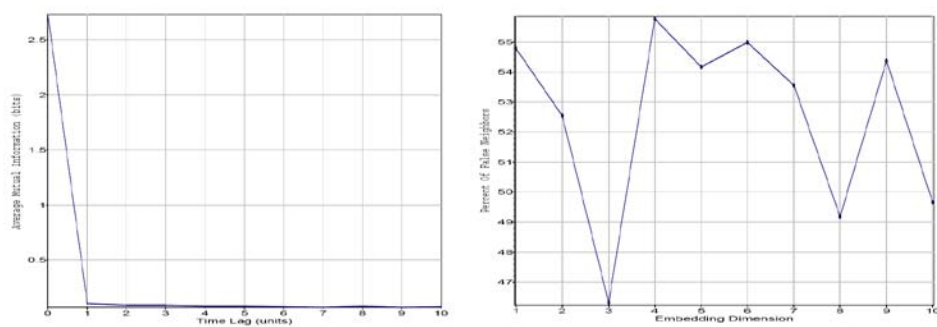
Obrázek 6

Časově zpožděná míra vzájemné informace a závislost počtu falešných sousedů na dimenzi vnoření řady výnosností akcií Telefonica O2



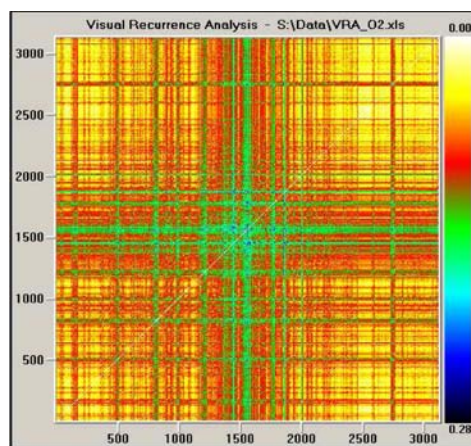
Obrázek 7

Časově zpožděná míra vzájemné informace a závislost počtu falešných sousedů na dimenzi vnoření řady výnosností akcií ČEZu

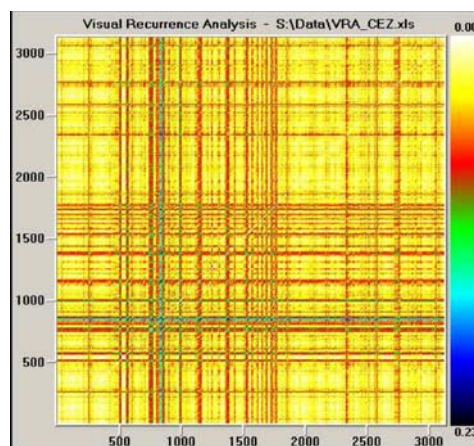


Obrázek 8

Rekurenční graf řady výnosností a) akcií Telefonica O2 a b) ČEZu.



a)



b)

Na všech rekurenčních grafech zkoumaných časových řad běžným pohledem nelze najít žádnou pravidelnou strukturu ani žádnou diagonální linii. Tyto rekurenční grafy se velmi podobají rekurenčnímu grafu bílého šumu s vertikálními liniemi různých barev. Jedná se o stavy, které se vzdalují od ostatních stavů s tím, že čím je barva těchto linií tmavší, tím je jejich vzdálenost větší. V původních časových řadách výnosností tomu odpovídají takové hodnoty výnosností, které silně vybočují z jejich průměru, a to na obou stranách hodnoty jejich průměru. Vizualně je tedy možné vyloučit přítomnost determinismu a periodicity ve zkoumaných řadách výnosností. Alternativně bychom tedy mohli konstatovat, že výnosnosti zkoumaných akcií a akciového indexu PX se chovají spíše náhodně než deterministicky.

V tabulce 3 jsou výsledky kvantitativní rekurenční analýzy. Pro srovnání jsou tam také uvedeny výsledky kvantitativní rekurenční analýzy pro Lorenzův systém. Kvantitativní rekurenční analýza byla opět provedena v programu Visual Recurrence Analysis (Koronov, 2007).

Tabulka 3:

Výsledky kvantitativní rekurenční analýzy zkoumaných řad

	Index PX	Akcie KB	Akcie O2	Akcie ČEZu	Lorenzův systém
RR	0	0	0	0,003	0,039
DET	0	0	0	0	91,648
LAM	0	0	0	0	0
TREND	0	0	0	0,002	-0,023
MAXLINE	0	0	0	0	689
ENTR	0	0	0	0	4,522
TT	0	0	0	0	0

Výsledky kvantitativní rekurenční analýzy potvrzují to, co pozorujeme na rekurenčních grafech zkoumaných řad výnosností. Rekurentní míry u všech řad výnosností našeho zájmu jsou nulové. To znamená, že každý stav v rekonstruovaném fázovém prostoru je jedinečný a k němu se už nevrátí žádný další stav. Analýza také zjistila, že není přítomná v rekurentních grafech žádná diagonální linie ani vertikální úsek, hodnoty míry determinismu a laminarity jsou nulové. To potvrzují také hodnoty délky největšího diagonálního úseku, které jsou ve všech případech je nulové. Další hodnoty, které jsou předmětem naší kvantitativní rekurentní analýzy, jsou také nulové. Naproti tomu, u Lorenzova systému hodnota rekurentní míry je 3,9 %, tedy pravděpodobnost, že se určitý stav opakuje, je 3,9 %. Také míra determinismu je velmi vysoká (91,65 %) a nejdelší diagonální linie má 689 pozorování (ze vzorku s 5000 pozorováními). Nulová laminarita a doba v pastí (trapping time) svědčí o tom, že systém neuvízne v žádném ojedinělém bodě. Hodnota entropie, která je 4,522, není tedy příliš vysoká, a to jednak potvrzuje deterministický charakter Lorenzova systému, jednak také dokládá, že rekonstrukce Lorenzova systému je jednodušší a přesnější než rekonstrukce systémů, které generují námi zkoumané řady výnosností. Výsledky kvantita-

tivní rekurenční analýzy vypovídají nejspíše o neexistenci deterministického procesu generujícího výnosnosti zkoumaných akcií a akciového indexu PX na českém akciovém trhu. Pokud by byl v řadách výnosů přítomen deterministický proces, muselo by se jednat o proces vysoké dimenze, kde modelování finančních řad pomocí nelineárních metod není příliš efektivní.

5. Závěr

Modelování finanční časové řady tradičními lineárními metodami obvykle neposkytuje takové výsledky, které bychom potom mohli úspěšně využít k predikci jejich budoucího chování. Proto se snažíme aplikovat nové poznatky z jiných vědních disciplín. Nelineární metody založené na určení přítomnosti determinismu je jedním z možných východisek. Tyto metody jsou však náročné na vstupní data, která nejsou vždy dostupná. Pro rychlá zjištění skrytého deterministického procesu ve finančních časových řadách se proto jeví jako vhodná technika rekurenčních grafů založená na grafickém znázornění vzdáleností jednotlivých stavů příslušného dynamického systému zrekonstruovaného z dané časové řady, která na základě charakteristických vzorů výskytu v rekurenčním grafu poskytuje relevantní informace o chování zkoumaného systému.

Aby se předešlo nejednoznačné interpretaci rekurenčního grafu, byla později navržena kvantitativní rekurenční analýza, která se snaží získat exaktnější informace z rekurenčních grafů na základě několika statistik. Technika rekurenčních grafů a kvantitativní rekurenční analýza byla následně aplikována na denních datech získaných z českého akciového trhu k zjištění přítomnosti případného deterministického procesu. Pro tento účel byly vybrány kurzové řady akciového indexu PX a akcií Komerční banky (KB), akcií Českých energetických závodů (ČEZ) a akcií Telefonica O₂ (dřívejší Český Telecom, O2) v období od roku 1995 do roku 2007. Výsledky využití techniky rekurenčních grafů a kvantitativní rekurenční analýzy ukazují, že v řadách výnosností akciového indexu a jmenovaných akcií nebyla zjištěna přítomnost deterministického procesu.

Literatura

- ANTONIOU A.; VORLOW C. 2000. Recurrence plot and financial time series analysis. *Neural Network World*, 10, s. 131-145.
- CAO, L. 1997. Practical method for determining the minimum embedding dimensions of scalar time series. *Physica D*, 1997, 110.
- ECKMANN, J.; KAMPHORST, S.; RUELLE, D. 1987. Recurrence plot of dynamical system. *Europhysics Letters*, 1987, 4, s. 973-977.
- FABRETTI A.; AUSLOOSE M. 2005. Recurrence plot and recurrence quantification analysis technique for detecting a critical regime, example from financial market indices. *International Journal of Modern Physics C*, 2005, 16(5).
- HORÁK, J.; KRLÍN, L.; RAIDL, A. 2003. *Deterministický chaos a jeho fyzikální aplikace*. Praha : Academia, 2003.
- KANTZ, H.; SCHREIBER, T. 1997. *Nonlinear Time Series Analysis*. Cambridge : Cambridge University Press, 1997.
- KENNEL, M.; BROWN, R.; ABARBANEL, H. 1992. Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Physical Review A*, 45(6).
- KORONOV, E. 2007. Visual Recurrence Analysis. [software] The VRA package.

- LORENZ, E. N. 1963. Deterministic nonperiodic flow. *Journal of Atmospheric Science*, 1963, 20, str. 130-141.
- MARVAN, N.; WESSEL, N.; MEYERFELDT, U.; SCHIRDEWAN, A.; KURTHS, J. 2002. Recurrence plot based measures of complexity and its application to heart rate variability data. *Physical Review E*, 2002, 66(2).
- SCHWABE, T. 2007. Recurrence quantification analysis of financial data. [Přednáška na semináři o ekonometrii], Riezle, Německo, červen 2007.
- SPROTT, J. C. 2003. *Chaos and Time Series Analysis*. New York : Oxford University Press, 2003.
- TAKENS, F. 1985. On the numerical determination of the dimension of an attractor. In BRAAKSMA, B. L. J.; BROER, H. W.; TAKENS, F., *Dynamical Systems and Bifurcations. Lecture Notes in Math.* Heidelberg – New York : Springer, 1985.
- TRAN, VAN QUANG 2007. Testování slabé formy hypotézy efektivního trhu na českém akciovém trhu lineárními a nelineárními metodami. [Disertační práce] VŠE, Praha, 2007.
- TREŠL, J. 2008. Applications of Physics in Finance. In: *Book of Abstracts of 7th International Conference Aplimat 2008*. Bratislava : Slovak University of Technology, 2008. ISBN 978-80-89313-03-7.
- TREŠL, J.; BLATNÁ, D. 2007. Dynamic Analysis of Selected European Stock Markets. *Prague Economic Papers*. 2007, roč. 16, č. 4, s. 291-302.
- VÁCHA, L.; VOŠVRDA, M. 2007. Wavelet Decomposition of the Financial Markets. *Prague Economic Papers*. 2007, roč. 16, č. 1, s. 38-54.
- VOŠVRDA, M.; BARUNÍK, J. 2008. Modelování krachů na kapitálových trzích: aplikace teorie stochastických katastrof. *Politická ekonomie*. 2008, roč. 56, č. 6, s. 759-771.
- ZBILUT, J.; WEBBER, C. 1992. Embeddings and delays as derived from quantification of recurrence plots. *Physics Letters*, 171, str. 199-203.

VISUAL RECURRENCE ANALYSIS AND ITS APPLICATION ON THE CZECH STOCK MARKET

Jan Kodera, Tran Van Quang, Centre for Basic Research in Economic Dynamics and Econometrics, University of Economics Prague (kodera@vse.cz, tran@vse.cz)

Abstract

The aim of the article is to answer the question if the Czech stock market price dynamics is generated by non-linear deterministic dynamic process. To solve this complex problem requires using sophisticated computational operations to analyze huge amount of data input. To overcome this obstacle the visual recurrence analysis is applied in this article. This method enables visualization of the state space reconstructed from a time series in the so called recurrent plot. Further, it quantifies various geometric structures occurred in recurrent plots and gives us more exact information about the nature of the underlying process generating the time series. This analysis is then applied to the most liquid stock returns and the Czech stock market index PX series.

Keywords

nonlinear deterministic dynamics, Lorenz attractor, visual recurrence analysis, recurrence plot, quantitative recurrence analysis, finance time series, Czech stock market

JEL classification

C33, G19