

Pénzügyi idősorok hiányzó adatainak kezelése – afrikai devizaárfolyamok példáján*

Sávai Marianna,

a Szegedi Tudományegyetem
PhD-hallgatója

E-mail: savai.marianna@eco.u-szeged.hu

Kiss Gábor Dávid,

a Szegedi Tudományegyetem
egyetemi adjunktusa

E-mail: kiss.gabor.david@eco.u-szeged.hu

A tanulmány célja a pénzügyi idősorokban megjelenő adathiányok kezelésére alkalmazott főbb eljárások összehasonlítása, azok momentumokra, volatilitásmodellezésre és value-at-risk jelzésekre gyakorolt hatásain keresztül. Tekintettel arra, hogy a hiányzó adatok kezelését elsősorban kérdőíves lekérdezésekből származó adathiányok esetében tárgyalja a szakirodalom, szükséges a pénzügyi idősorokon történő összehasonlításuk. A szerzők a listaszerű adattörlést, átlaggal pótlást és a likelihood-becsléseken alapuló általános várakozásmaximalizációs eljárásokat hasonlítják össze napi záró devizás idősorokon. A vizsgált minta az afrikai lebegő devizákat tartalmazza 2000. március 8. és 2015. március 6. között dollárban denominálva, kiegészítve az euróval és az ahhoz kötött CFA frankkal. Az elvégzett számítások eredményei alapján az EM-eljárás alkalmazását nem javasolják, annak a volatilitásra, korrelációra és extrém elmozdulásokra gyakorolt hatásai miatt.

TÁRGYSZÓ:
Hiányzó adat.
Idősor.

DOI: 10.20311/stat2016.07.hu0736

* A szerzők köszönetet mondanak a Pallas Athéné Domus Scientiae Alapítványnak a pénzügyi támogatásért, amely hozzájárult a tanulmány elkészüléséhez.

A többváltozós idősorok elemzése során szinkronizált és folytonos adatokra van szükségünk – azaz minden kereskedési nap és vizsgált piaci eszköz esetében rendelkezünk kell egy árfolyammal. Mindazonáltal lehetnek olyan speciális alkalmak, amikor egy vagy több adat hiányzik a kereskedési aktivitás hiánya miatt. Tanulmányunkban a hiányzó adatok pótlására alkalmazott főbb módszerek GARCH- (generalized autoregressive conditional heteroskedasticity – általánosított autoregresszív feltételes heteroskedaszticitás), VaR- (value at risk – kockázatosított érték) eljárások paramétereire gyakorolt hatását hasonlítjuk össze, különösképpen kitérve a volatilitás perzisztenciájára és aszimmetriájára, valamint az eredeti és javított idősorok momentumaira.¹

A hiányzó adat (vagy hiányzó érték) definíciója a következő: „azon adatok összessége, melyek nem találhatók meg a vizsgálatra kiválasztott mintában” (Kang [2013] 402. old.). Adathiány az adatgeneráló-folyamat átmeneti felfüggesztése miatt jön létre idősorok esetében (Graham [2012]), melyek kapcsán komoly elvárásokat szokás megfogalmazni: a hiányzó adatok kezelésekor sem „illik” torzítani az ARIMA- (autoregressive integrated moving average – autoregresszív integrált mozgóátlag-folyamat) és GARCH-paramétereket, rontani a reprezentativitást vagy az átlag, szórás értékét és az autokorreláltság fokát (Juan Carlos *et al.* [2010], Kang [2013]).

Munkánk fő hozzáadott értéke a napi felbontású, pénzügyi idősorok hiányzó adatainak kezelése során alkalmazott főbb eljárások (a listaszerű vagy páronkénti törlés, átlaggal vagy a rendelkezésre álló historikus mediánnal történő pótlás és a likelihood-alapú várakozásmaximalizáció) összehasonlítása, miután a szakirodalom elsősorban a kérdőív típusú hiányzó adatok problémájának kezelését tárgyalja.

A vizsgált mintát az afrikai – az IMF [2013] besorolása szerinti – lebegő árfolyam-politikát folytató országok devizái adják: a kenyai shilling (KES), a ghánai cedi (GHS), a dél-afrikai rand (ZAR), a tanzániai shilling (TZS), ugandai új shilling (UGX), gambiai dalasi (GMD), madagaszkári ariary (MGA) és mozambiki metical (MZN) fizetőeszközök, amerikai dollárban kifejezve. Kontrollcsoportként a CEMAC (Commission of the Economic and Monetary Community of Central Africa – Középfrikai Gazdasági és Monetáris Közösség) által alkalmazott, euróhoz rögzített CFA frank 1 (XAF)² is a mintába került, valamint a devizák közötti korreláció teszteléséhez az EUR/USD árfolyamot is elemeztük. Vizsgálatunk a Bloomberg adatbázisából

¹ Az adathiány kezelése nem kizárólag a pénzügyi idősorokra vonatkozik, nemzetközi folyamatok összehasonlítására is alkalmas lehet, ahol nagyon gyakran az adathiány miatt kell egy-egy országot vagy változót kihagyni az elemzésből (lásd például Udvari–Urbáné [2015] vagy Udvari–Pontet [2015]).

² Középfrikai valutaközösségi frank (Csád, Egyenlítői Guinea, Gabon, Kamerun, Kongó, Középfrikai Köztársaság).

letöltött, 2000. március 8. és 2015. március 6. közötti időszak napi záró adatain alapszik. A mintaválasztást az idősorok egzotikus jellege indokolja, amellet, hogy e gazdaságok strukturális gyengése miatt a pénz- és tőkepiaci idősorokra hagyományosan jellemző, eltérő nemzeti ünnepek miatti kereskedési szünetek mellett, további kereskedési szüneteket is találhatunk.

Tanulmányunk felépítése: az első fejezetben foglaljuk össze a hiányzó adatokkal kapcsolatos elméleti feltételezéseket, rámutatva a kérdőíves és idősoros adatok közötti különbségekre. Ezt követi a kezelésükre kidolgozott főbb eljárások bemutatása a módszertani második fejezetben, valamint az eredmények összehasonlítására szolgáló GARCH- és DCC-GARCH- (dynamic conditional correlation – dinamikus feltételes korreláció) modellek és az alapértelmezett VaR-eljárás ismertetése. A harmadik fejezetben előbb bemutatjuk a nyers, szinkronizálatlan devizás idősorok tulajdonságait, majd összevetjük őket a szinkronizált, hiányzó adatoktól megtisztított idősorok teszteredményeivel.

1. Elméleti háttér

A pénzügyi idősorok, különösen a napi záróárfolyamok esetében előfordul adathiány a kereskedési adatokban, hiszen egyes piacok zárva tarthatnak, míg a többi piacon javában zajlik a kereskedés. Ez egy érdekes többváltozós jelenség. A kereskedés hiánya egyrészt a nemzeti sajátosságokra vezethető vissza (például az ünnepekre, a szombati munkavégzésre), másrészt a piaci erők okozta alacsonyforgalmú helyzetekre (ez leggyakrabban a kis tőzsdecégek részvényeivel³ fordul elő), továbbá az árak hirtelen összeomlása miatti felfüggesztésre. Óriási irodalma van az árképzést és a piaci hatékonyságot alacsony jelenségeknek, amelyek közül a legtöbbet idézett a „hétvégehatás” (weekend effect) (*Keim–Stambaugh* [1984], *Robins–Smith* [2015], *Shahid–Mehmood* [2015]).

A szakirodalom három típusát különbözteti meg az adathiánynak (*Graham* [2012], *Junger–Leon* [2015], *Oravec* [2008]):

– MCAR (missing completely at random – teljesen véletlenszerű adathiány): az adathiány nem függ az adatok vagy más, különösen a megfigyelt változó értékeitől. Kihagyásuk nem torzítja a becslésünket a homogenitásuk miatt (*Enders* [2010], *Junger–Leon* [2015], *Kang* [2013]).

– MAR (missing at random – véletlenszerű adathiány): a hiányzás attól a változótól független, amely adatsorában felmerül (*Kang*

³ Erről a jelenségről bővebben ír például *Giovanni* [2005].

[2013]), de valamilyen mechanizmust feltételezhetünk a hiányzás mögött (*Graham* [2012]). Kihagyásuk károsíthatja az olyan időbeli struktúrákat, mint az autokorreláció, a trendek vagy a szezonálisitás (*Junger–Leon* [2015]).

– MNAR (missing no at random – nem véletlenszerű adathiány) akkor következik be, ha a hiány nem független az azt tartalmazó változótól (*Oravecz* [2008]), de amikor lehetséges, akkor torzítatlan becslést illeszthetünk a hiányzó adatokra (*Graham* [2012]).

Előfordul, hogy az adathiányt a kutató nem tudja kezelni, mert az eloszlása ismeretlen, a MAR esetén csak feltételezéssel lehet élni (*Graham* [2012]).

A következőkben három különféle adathiány kezelésére használható, *Baraldi et al.* [2015] által is leírt gyakorlatot mutatunk be. Az első, az adathiányt tartalmazó esetek listaszerű vagy páronkénti törlése, amikor eltávolítjuk azokat az intervallumokat, ahol legalább egy hiányzó adatunk van. A törlés az idősor töredezettségét okozhatja vagy torzítóvá válhat a paraméter becslése, ezért csak MCAR-adathiánynál alkalmazható (*Kang* [2013]). A második, az adathiány átlaggal vagy a rendelkezésre álló historikus mediánal történő pótlása, amit *Junger–Leon* [2015] is javasol aszimmetrikus adatoknál. Ennek a módszernek is hasonló hatása lehet, mint az előzőnek a logaritmikusan hozamokra, melynek idősora zéró átlagú és móduszú. Ezt a megoldást *Graham* [2012] nem javasolja az MCAR-adathiány esetén, az eloszlások magasabb átlag körüli koncentrációja, valamint a standard hibák és variancia alulbecslése miatt (*Junger–Leon* [2015], *Enders* [2010]). A harmadik és egyben legmodernebb eljárás, amikor a hiányzó adatokat megpróbálják rekonstruálni, minimalizálva a hibafüggvényt, a szórás deriválásával vagy valószínűségi hányadossal (*Baraldi et al.* [2015], *Ceylan et al.* [2013], *Juan Carlos* [2010]). Az EM- (expectation maximization – várakozásmaximalizáció) modellek közül a leggyakrabban alkalmazott a maximum likelihood becslés, de neurális hálózatokra építő és genetikusan alapuló megközelítések is elterjedtek (*Ceylan et al.* [2013], *Juan Carlos* [2010]). Az EM-eljárások hátránya, hogy alkalmazásuk több időt vehet igénybe, mert az algoritmusuk és a likelihood függvény nehezen számítható ki (*Ruud* [1991]), valamint szükség lehet további adatgeneráló modellre is (*Horari et al.* [2013]). Az EM-eljárások nem okoznak problémát MCAR-adathiány esetén, ilyenkor jól használhatók az EM eljárások. A legjobb választási lehetőség a maximum likelihood becslés a hiányos többváltozós normális eloszlású adatok kezelésére, ugyanis ezek az eljárások kevésbé torzítanak, mint az adathiány törlése vagy az átlaggal való pótlás. A maximum likelihood alkalmazhatósága a többi eljáráshoz képest függ a hiányzó adatok arányától, a mintanagyságtól és adatstruktúra kovarianciájától (*Wothke* [1998]).

A hiányzó adatok miatti problémák befolyásolhatják olyan, a napi idősorokon történő többváltozós alkalmazások használatát, mint a volatilitás tovagyrűzését, az

extrém ingadozások, illetve fertőzések modellezését, ahol a feltételes varianciáról, a kovarianciáról és a korrelációról azt feltételezzük, hogy a vizsgálat szempontjából döntő jelentőségűek.

2. Módszertan

A tanulmány három különféle adathiányt kezelő módszert hasonlít össze a legfontosabb centrális momentumok, az autokorreláció, a volatilitás tartóssága és az extrém elmozdulások alapján.

Tegyük fel, hogy a vizsgált n külföldi deviza közül kiválasztott i -edik deviza ($1 \leq i \leq n$) árfolyamából vett P_i minta /1/ minden y kereskedési napon v mintamérettel a következők szerint írható le:

$$P_i = \begin{bmatrix} y_1 & P_{i,1} \\ \dots & \dots \\ y_v & P_{i,v} \end{bmatrix}. \quad /1/$$

Emellett kiválasztunk még egy másik k -edik ($1 \leq k \leq n$ és $k \neq i$) devizát is /2/ w minta és z ($z \neq y$) időindexszel:

$$P_k = \begin{bmatrix} z_1 & P_{k,1} \\ \dots & \dots \\ z_w & P_{k,w} \end{bmatrix}. \quad /2/$$

A minta mérete rendre megegyezik a kereskedési napok számával.

Amennyiben az előbbieket alapján leírható $P_{1, \dots, i, k, \dots, n}$ mátrixokat egyesíteni kell egy többváltozós elemzés céljából, akkor az időindexek összehangolására (szinkronizálására) van szükségünk, ami az Y és Z oszlopvektorok közötti közös és a diszjunkt halmazok elemeinek kezelését jelenti.

A listaszerű adattörlésnél /3/ kizárunk minden olyan esetet, amikor legalább egy érték hiányzik – ez az Y és Z halmazok közös metszetét (T) jelenti:

$$T = Y \cap Z. \quad /3/$$

Az átlaggal való pótlás /4/ alkalmazása különösen abból a szempontból praktikus, mert a logaritmikus hozamok jellemzően nulla közeli átlaggal és módusszal rendel-

keznek. Az LOCF- (last observation carried forward – utolsó elvégzett megfigyelés értékelése) módszer alkalmazása szintén erre az eredményre vezet, nulla logaritmi- kus hozamot produkálva (feltételezve, hogy az o időindexnél hiányzik az adat, míg az $o - 1$ esetében van).⁴

$$\begin{aligned} T &= (Y \cup Z), \quad p_{i,o} = p_{i,o-1} \text{ és} \\ T &\notin (Y \cap Z), \text{ illetve } r_{i,o} = p_{i,o} - p_{i,o-1}. \end{aligned} \quad /4/$$

Az általános EM-algoritmus alapja egy iterált lineáris regressziós elemzés, de ezt helyettesítettük egy gauss eloszlású regressziós paraméter feltételes maximum likelihood becslésével /5/ Schneider [2001] tanulmánya alapján. Egyes $p_{t,i} \notin P$ hiányzó értékkel rendelkező mátrix esetén, a hiányzó (kereskedési napok) és a rendelkezésre álló értékekkel megadott árak közötti kapcsolatra lineáris regressziós modell írható fel:

$$p_{NaN} = \mu_{NaN} + (p_a - \mu_a)B + \varepsilon, \quad /5/$$

ahol a a meglévő adatot jelenti, $B \in \mathfrak{R}^{n_a \times n_{NaN}}$ pedig a regressziós együtthatók mátrixa a hiányzó és meglévő értékek kovarianciamátrixával, az n számú összes mintából. Az $\varepsilon \in \mathfrak{R}^{1 \times n_{NaN}}$ reziduumból feltételezzük, hogy nulla átlagú és $C \in \mathfrak{R}^{n_{NaN} \times n_{NaN}}$ egy ismeretlen kovarianciamátrix-vektor. Az EM-algoritmus iterációiban a $\mu \in \mathfrak{R}^{1 \times n}$ átlagát és a $\Sigma \in \mathfrak{R}^{n \times n}$ kovarianciamátrix becslését adottnak tekintjük, és ezek becsléséből számoljuk ki a \mathbf{B} mátrix regressziós együtthatóira és a \mathbf{C} kovarianciamátrix reziduumaikra vonatkozó feltételes maximum likelihood becsléseket minden hiányzó értéket tartalmazó bejegyzésre. Mindezt úgy, hogy az algoritmus minden hiányzó érték helyére imputált értéket helyettesítsen a teljes μ vektor és Σ mátrix újraszámítását megelőzően. Ezek alapján a becsült regressziós együttható két (hiányzó-hiányzó és meglévő-hiányzó) becsült kovarianciamátrixból adódik:

$B = \sum_{aa}^{-1} \sum_{aNaN}$, amely a reziduum kovarianciamátrixának becslésére használható a későbbiekben. Azonban, az általános EM-algoritmus minden hiányzó értékkel

⁴ Könnyen belátható, hogy nulla várható érték mellett az átlaggal való pótlás és az LOSCF-eljárás azonos eredményre vezet, amennyiben a javított idősornak a továbbiakban az $r_{i,o}$ differenciáltjaival számolunk: $r_{i,o} = p_{i,o} - p_{i,o-1}$ a $p_{i,o} = p_{i,o-1}$ esetben $r_{i,o} = 0$ eredménnyel zárul. Logaritmi- kus hozamoknál ellenben problémákba ütközünk, hiszen $e^x \neq 0$, ami miatt érdemes egy kellően kicsi $\varepsilon = 10^{-6}$ számmal kiegészíteni: $p_{i,o} = p_{i,o-1} + \varepsilon$, ekkor $r_{i,o} = \ln(p_{i,o} - p_{i,o-1} - \varepsilon) = \ln(\varepsilon) \approx 0$ lesz az eredmény.

rendelkező bejegyzés esetén a $B = \left(\sum_{aa} + h^2 \text{Diag} \left(\sum_{aa} \right) \right)^{-1} \sum_{aNaN}$ becslést használ-

ja, ahol a h szabályozó paraméter a diagonális elemeket $1 + h^2$ tényezővel bővíti.

Az említett eljárásokkal szinkronizált többváltozós idősorokon végrehajtott érzékenységvizsgálat a modellben szereplő bizonytalan input torzításának, a középértékek, az autokorreláció vizsgálatát érinti, valamint meghatározza az adathiány százalékos súlyát a mintában (Kang [2013], Graham [2012]). A varianciamodellre hatással lehet az adathiány, a modellválasztás és a paraméterek torzítása. Különböző GARCH-modelleket illesztettünk az adatokra a volatilitás tartósságának vizsgálatához, követve *Cappeillo–Engle–Sheppard* [2006] munkájában leírtakat. Az alkalmazott (/6/-/10/) modellek hasznosak a volatilitás és időbeli csoportosulásának (a heteroszkedaszticitás) megragadásához.

– GARCH (p, q):

$$\sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2, \quad /6/$$

ahol σ_t^2 a jelenbeli variancia, ω a konstans tényező, p a késleltetés száma az α_i paraméterű ε_{t-1}^2 innovációnak, valamint q a késleltetés száma a β_j paraméterű σ_{t-j}^2 varianciának a volatilitás tartósságának kimutatásához. Aszimmetrikus GARCH-modellek alapján:

$$\begin{cases} S_{t-i}^- = 1, \text{ ha } \varepsilon_{t-i} < 0 \\ S_{t-i}^- = 0, \text{ ha } \varepsilon_{t-i} \geq 0 \end{cases} \quad /7/$$

egy jelzés a csökkenő hozamokra adott aszimmetrikus reakciókra.

– GJR- (Glosten–Jagannathan–Runkle) GARCH (p, o, q):

$$\sigma_t = \omega + \sum_{i=1}^p \alpha_i |\varepsilon_{t-i}| + \sum_{i=1}^o \gamma_i S_{t-i}^- |\varepsilon_{t-i}| + \sum_{j=1}^q \beta_j \sigma_{t-j}. \quad /8/$$

– TGARCH (p, o, q) (threshold GARCH – küszöb GARCH):

$$\sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^o \gamma_i S_{t-i}^- \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2. \quad /9/$$

– APARCH (p, o, q) (asymmetric power ARCH – aszimmetrikus teljesítmény ARCH):

$$\sigma_t^\delta = \omega + \sum_{i=1}^p \alpha_i \left(|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i} \right)^\delta + \sum_{j=1}^q \beta_j \sigma_{t-j}^\delta, \quad /10/$$

ahol $\alpha_i > 0$ ($i = 1, \dots, p$), $\gamma_i + \alpha_i > 0$ ($i = 1, \dots, o$), $\beta_i \geq 0$ ($i = 1, \dots, q$), $\alpha_i + 0,5\gamma_i + \beta_k < 1$ ($i = 1, \dots, p; j = 1, \dots, o; k = 1, \dots, q$) és δ index paraméter 1 és 2 közötti értéket vehet fel.

A modellválasztás során a homoszkedasztikus hibatagokra koncentráltunk (2-vel késleltetett ARCH–LM-tesztet futtattunk), és megkerestük a legalacsonyabb BIC-cel (Bayesian information criterion – Bayes-féle információs kritérium) rendelkezőt. Engle [2002] alapján tanulmányunk a GARCH-alapú DCC-modellt⁵ alkalmazza a kiválasztott piacok napi együttmozgásának vizsgálatára.

A hiányzó értékek hatással vannak az adatokból számolt logaritmusos hozamok sűrűségfüggvényére, a listaszerű adattörlesztés feltételezhetően több adatot eredményezhet az eloszlás szélein a csonkított idősor esetében, míg az átlaggal történő imputáció növeli a 0 elmozdulás gyakoriságát a kiegészített idősoránál. Az EM-nek az átlag és az extrém értékek közötti adatokat kellene produkálnia. Az adatok extrém ingadozását közönséges VaR-moddellel /11/ vizsgáltuk, ahol az extrém adatok súlya és a nem extrém adatok kurtózisa volt a változó.

$$r = r_n + r_x^- + r_x^+, \quad r_x^- < \mu - 1,65 * \sigma, \quad r_x^+ > \mu + 1,65 * \sigma, \quad /11/$$

ahol r a logaritmusos hozam, μ a feltétel nélküli átlag, σ a feltétel nélküli szórás, r_x^- jelzi az extrém negatív, r_x^+ az extrém pozitív hozamokat és r_n mutatja a nem extrém adatok részhalmazát (Madura [2008]).

A DCC-GARCH-modell definíciója Engle [2002] alapján:

$$r_t = \mu_t + \alpha_t, \quad \alpha_t = H_t^{1/2} z_t, \quad H_t = D_t R_t D_t, \\ R_t = Q_t^{*-1} Q_t Q_t^{*-1}, \quad Q_t = (1-a-b)\bar{Q} + a\varepsilon_{t-1}\varepsilon_{t-1}^T + bQ_{t-1}, \quad /12/$$

ahol r_t a hozamok logaritmusát, α_t a korrigált hozamot jelöli ($E[\alpha_t] = 0$ és $Cov[\alpha_t] = H_t$), μ_t a várható értéke a feltételes r_t -nek, H_t az α_t feltételes varianciamátrixa, $H_t^{1/2}$ adódik a Cholesky-felbontás után, D_t feltételes

⁵ A becslés a Kevin Sheppard által fejlesztett Oxford MFE- és UCSD-toolboxokon alapul: https://www.kevinsheppard.com/MFE_Toolbox

szórása az R_t feltételes korrelációs mátrixnak, z_t a független azonos eloszlású hibatagok vektora, Q_t feltétel nélküli kovarianciamátrixa a standardizált ε_t hibatagnak (Cappiello *et al.* [2006]).

A tanulmányunk a három modellt a következő alapján vizsgálja: eltérés a nyers és finomított adatok között, középértékek, autokorreláció, heteroszkedaszticitás, normális eloszlás, gyenge stacionaritás, GARCH-modell és paraméterválasztás, VaR-súlyok és kurtózis, valamint a dinamikus feltételes korreláció eredményeit hasonlítjuk össze.

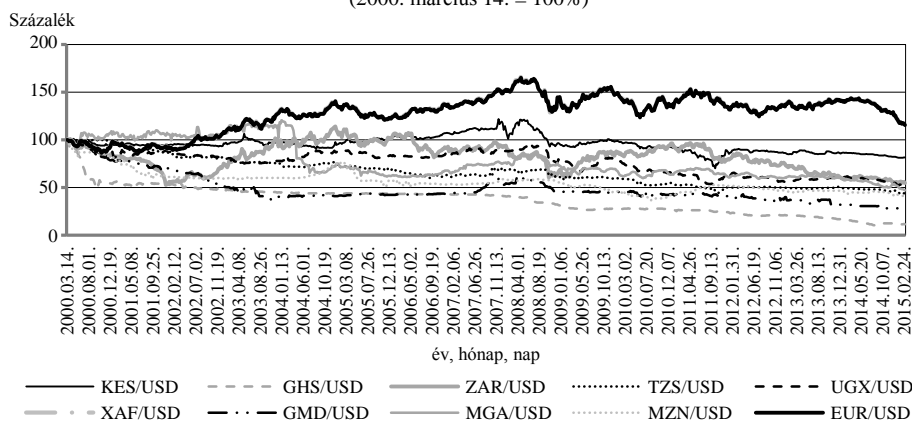
3. Eredmények és adatok

A nyers, szinkronizálatlan idősorok statisztikáit, valamint a szinkronizálás és a háromféle hiányzóadat kezelési eljárás nyomán kapott eredményeket hasonlítjuk össze ebben a fejezetben, bemutatva a pénzügyi idősorok vizsgálatára használt GARCH-, DCC-GARCH- és VaR-modellek hiányzó adatpótlási eljárásokkal szembeni érzékenységét.

3.1. Eredeti idősor, egyváltozós tesztek

Az ábrán látható, hogy az euróhoz rögzített XAF szorosan követte e kulcsvaluta dollárral szembeni erősödését, majd 2008-at követő oldalazását. Emellett a KES és a ZAR értékelődött fel a subprime-válság (jelzáloghitel-válság) 2008-as kiéleződése előtt. Ezeket az eseteket leszámítva a minta többi részére a leértékelődés volt jellemző.

A mintában szereplő afrikai devizák heti árfolyamának változása a bázishoz képest 2000 és 2015 között
(2000. március 14. = 100%)



Forrás: Bloomberg.

A nyers idősor logaritmikus hozamai nulla átlaggal és alacsony szórással rendelkeztek, míg szimmetria csupán az EUR és a GMD esetében valósult meg. (Lásd az 1. táblázatot.) Kiugró csúcosság volt jellemző a mintára, ami a normális eloszlásnál az elvárthoz képest több extrém árfolyamváltozást jelzett – csupán a rögzített XAF-nál és az EUR-nál találtunk ideális, háromhoz közeli értékeket. A normális eloszlás hipotézisét valamennyi deviza esetében elvetettük, míg két késleltetés mellett a minta jelentős hányadára az autokorreláltság (kivéve az EUR-t) és a heteroszkedaszticitás (kivéve a KES-t, a ZAR-t és az EUR-t) volt jellemző. Az adatok gyenge stacionaritást mutattak.

1. táblázat

Az egyváltozós idősorok leíró és alapstatisztikái

Deviza	Átlag	Szórás	Aszimmetria	Csúcosság	Normális eloszlás Jarque–Bera (<i>p</i>)	Autokorre- láció Ljung–Box (<i>p</i>)	Heteroszkedaszticitás ARCH-LM (<i>p</i>)	Stacionaritás ADF (<i>p</i>)
KES/USD	0,00	0,01	-0,32	20,51	0,00	0,00	0,10**	0,00
GHS/USD	0,00	0,01	-1,23	33,77	0,00	0,00	0,00	0,00
ZAR/USD	0,00	0,01	-1,05	15,74	0,00	0,01	0,14**	0,00
TZS/USD	0,00	0,01	0,82	39,78	0,00	0,00	0,00	0,00
UGX/USD	0,00	0,01	-0,47	16,76	0,00	0,00	0,02	0,00
XAF/USD	0,00	0,01	0,13	5,14	0,00	0,00	0,00	0,00
GMD/USD	0,00	0,02	0,05	169,41	0,00	0,00	0,03	0,00
MGA/USD	0,00	0,01	-1,63	54,41	0,00	0,00	0,00	0,00
MZN/USD	0,00	0,01	-0,80	42,11	0,00	0,00	0,00	0,00
EUR/USD	0,00	0,01	-0,02	4,39	0,00	0,52*	0,59**	0,00

* Nem autokorrelált 2 késleltetés mellett.

** Homoszkedasztikus 2 késleltetés mellett.

Forrás: Saját számítás a Kevin Sheppard-féle MFE-toolbox segítségével.

Négy különböző (GARCH-, TARCH-, GJR-GARCH-, APARCH-) modell 13 különböző késleltetése mellett vizsgáltuk az idősorokat, keresve azt az eljárást, amely homoszkedasztikus hibatagokat eredményez abszolút értékben a legalacsonyabb BIC mellett. A minta felére az aszimmetrikus varianciát leíró modellek illeszkedtek jól (kivéve a GHS-t, a TZS-t, az UGX-t, az EUR-t), továbbá megjelent a volatilitás perzisztenciája (az 1-hez közeli béták). Sajátos eredményre jutottunk a GMD esetében, ahol az innovációk (vagy sokkok) kaptak komolyabb súlyozást. (Lásd a 2. táblázatot.)

2. táblázat

GARCH-modell az egyváltozós idősorokon

Deviza	Modell	ω	α_1	α_2	γ	β_1	β_2	BIC-
		paraméter						
KES/USD	TARCH(1,1,2)	0,00	0,26		-0,03	0,46	0,29	-4,20
GHS/USD	GARCH(1,2)	0,00	0,11			0,37	0,51	-3,70
ZAR/USD	GJR-GARCH(1,1,1)	0,00	0,12		-0,10	0,93		-3,21
TZS/USD	GARCH(1,1)	0,00	0,21			0,79		-3,91
UGX/USD	GARCH(1,1)	0,00	0,20			0,80		-3,82
XAF/USD*								
GMD/USD	TARCH(2,1,1)	0,00	0,30	0,30	-0,16	0,48		-2,81
MGA/USD	GJR-GARCH(1,1,2)	0,00	0,02		0,03	0,47	0,49	-3,18
MZN/USD	GJR-GARCH(1,1,2)	0,00	0,26		-0,11	0,29	0,51	-3,32
EUR/USD	GARCH(1,1)	0,00	0,04			0,96		-3,71

* Egy modell sem volt képes normális eloszlású homoszkedasztikus hibatagok létrehozására.

Forrás: Saját számítás a Kevin Sheppard-féle UCSD-toolbox segítségével.

VaR segítségével lehetőség nyílt arra, hogy az extrém elmozdulások leválogatásával létrehozott csonka idősor csúcossága 5 alá, jellemzően 3 közelébe csökkenjen. (Lásd a 3. táblázatot.) Az extrém árfolyammozgások súlya 10 százalék alatti (kivéve a XAF-ra és EUR-ra jellemző 11 százalékot), így az eljárás alkalmas volt azon ritka ingadozások kiszűrésére, amelyek az idősor vastag farkúságáért felelnek.

3. táblázat

VaR-tulajdonságok az egyváltozós idősorokon

Mutató	Deviza									
	KES	GHS	ZAR	TZS	UGX	XAF	GMD	MGA	MZN	EUR
Átlag	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Szórás	0,00	0,01	0,01	0,00	0,00	0,00	0,01	0,01	0,01	0,00
Aszimmetria	0,07	-0,14	-0,04	-0,02	-0,02	0,02	-0,04	0,04	0,05	0,01
Csúcosság	3,48	4,34	2,51	3,16	2,84	2,52	4,07	3,68	4,24	2,48
Extrém hozam (-) küszöb	-0,01	-0,02	-0,02	-0,01	-0,01	-0,01	-0,03	-0,02	-0,02	-0,01
Extrém hozam (+) küszöb	0,01	0,02	0,02	0,01	0,01	0,01	0,03	0,02	0,02	0,01
Extrém hozam (-) (db)	148	121	179	136	149	190	105	100	123	189
Extrém hozam (+) (db)	138	128	145	121	127	178	95	90	107	207
Nem extrém hozam (db)	3 606	3 573	3 589	3 578	3 563	3 455	3 589	3 424	3 598	3 517

Forrás: Saját számítás Madura [2008] nyomán.

3.2. Eljárások összehasonlítása

A hiányzások természetének jobb megértéséhez előbb teszteltük a két hiányzó adat felbukkanása között eltelt kereskedési napok számának normális eloszlását, illetve a hiányzó adatok mintabeli súlyát. A 4. táblázatban látható, hogy leginkább az MGA, illetve a GMD, a GHS és az XAF volt terhelt hiányzó adatokkal, továbbá a legkevésbé kíméletesnek a listaserű törlés bizonyult a magasabb hiányzó százalékos értékeivel. Az adatok hiánya azonban nem véletlenszerűen jelenik meg, miután az adathiányok között eltelt idő adott devizán belül nem követ normális eloszlást. A hiányok létrejöttének időbeli eloszlása tehát nem a véletlentől függ, hanem egyéb gazdasági hatásoktól (ünnepektől, piaci viszonyoktól).

4. táblázat

*A hiányzó adatok százalékos aránya
és a közöttük eltelt idő normális eloszlásának tesztelése*

Deviza	Adathiány százalékos súlya (százalék)		Hiányzó adatok között eltelt idő normális eloszlása Kolmogorov–Smirnov-teszt (p)
	Listaserű adattörlés	Átlaggal pótlás	
KES/USD	1	1	0,00
GHS/USD	3	2	0,00
ZAR/USD	0	0	nincs adat
TZS/USD	2	2	0,00
UGX/USD	2	2	0,00
XAF/USD	3	2	0,00
GMD/USD	4	3	0,00
MGA/USD	9	8	0,00
MZN/USD	2	2	0,00
EUR/USD	0	0	nincs adat

Forrás: Saját számítás.

A Függelék F1. táblázata tartalmazza a háromféle megközelítéssel kezelt idősorok logaritmikus differenciáltjainak momentumait, a normális eloszlásának, az autokorreláltságnak és heteroszkedasztikusságnak és gyenge stacionaritásnak tesztjeit. Az első momentum nullához közeli maradt az eredeti idősorokhoz hasonlóan, azonban a szórás megduplázódott-triplázódott az esetek 60 százalékában az EM-eljárásnál. Az idősorok aszimmetriájának jellegét mindegyik módszer befolyásolta. A csúcosság az esetek 40, illetve 50 százalékában nőtt a listaserű adattörlés és

átlaggal pótlás mellett, míg akár nagyságrendi ugrásokat is mutatott az EM-eljárás mellett. A logaritmikus hozamok továbbra sem voltak normál eloszlásúak, viszont gyengén stacionerek igen, továbbá nem volt érdemi változás sem autokorreláltságuk, sem heteroszkedaszticitásuk tekintetében.

A VaR-eljárás esetében (lásd az F2. táblázatot) hasonló eredményre jutottunk, mint a csúcosságnál: az első két eljárás nem mutatott érdemi változásokat, ellenben az EM-et tekintve a VaR kevesebb extrém mozgást jelzett, míg a „nem extrém” halmaz kurtóziisa az esetek 80 százalékában megnőtt (kivéve a ZAR-nál és az EUR-nál). Az EM-eljárás alatt, a hiányzó adatok pótlása során a többváltozós idősor kovariancia-mátrixának stabilizálására tett lépések ezek szerint torzíthatják a feltétel nélküli szórás értékét.

A listaszerű adattörlesztés a volatilitás modelljeinél (lásd az F3. táblázatot) csekély változást idézett elő a paraméterek körében (az innovációs paraméterek kismértékben növekedtek, a múltbéli volatilitás paraméterei csökkentek), és csupán az MGA és MZN devizáknál javasolt más modell kiválasztása, amellelt, hogy ezúttal már volt olyan GARCH-modell, amely homoszkedasztikus hibatagokat eredményezett az XAF esetében. Az átlaggal pótlás módszere az egyébként is aszimmetrikus volatilitású MGN és GHS devizákra az APARCH-modellt illesztette, emellett a GHS szimmetrikusból aszimmetrikus volatilitásúvá vált. A volatilitás paraméterei úgy mutattak emelkedést, hogy közben a BIC-értékek változatlanok maradtak. Az EM-eljárás átrendezte a korábbi, volatilitás szimmetriájával vagy épp aszimmetriájával kapcsolatos elképzeléseket, négy deviza (a KES, a GHS, a TZS, az UGX) vált szimmetrikusból aszimmetrikussá, míg három (a GMD, az MGA, az MZN) aszimmetrikusból szimmetrikus lett. A deviza árfolyamának csökkenése és a volatilitás növekedése közötti kapcsolatról alkotott képet tehát nagyban befolyásolta a hiányzó adatok kezelésére alkalmazott eljárás. Mindezt úgy, hogy az EM mellett kapott optimális GARCH BIC-értékek abszolút értékben alacsonyabbak a másik két eljárásénál, tehát jobban illeszkedő modellel van dolgunk. A ZAR és az EUR volatilitás-modelljei annak ellenére, hogy a listaszerű törlés miatt rengeteg adatot veszítettek (és emiatt töredezettebbek lettek), nem változtak érdemben, míg az egyébként számos hiányzó nappal terhelt GMD és MGA GARCH-modelljei és azok paraméterezése az adatkezelési eljárások függvényében alakultak.

Dacára az Európai Unió és az Egyesült Államok által a mintaországok számára nyújtott számos külkereskedelmi, segélyezési és egyéb, fizetési mérleg javítását is magában foglaló, programjának (*Udvari–Pontet* [2015]), a listaszerű adattörléssel és átlaggal pótlással szinkronizált többváltozós idősorok korrelálatlanok voltak az EUR/USD árfolyamával szemben – leszámítva a feltörekvő ZAR és a rögzített XAF esetét. (Lásd az F4. táblázatot.) A devizák izoláltságát jelzi, hogy a mintából csak a ZAR és az XAF mutatott szorosabb feltétel nélküli korrelációt. Ehhez képest az EM-eljárás esetében számos, komolyabb feltétel nélküli együttmozgást sejtető eredmény-

re juthatunk, ilyen a GHS és az UGX között mért 0,64-es, valamint a GHS és a KES, a TZS, a GMD, az MGA és az MZN között tapasztalt 0,3–0,5 közötti értékek, amirehhez hasonlókat találunk még az UGX–KES, a TZS–UGX, a GMD–MZN párok esetében is. Az idősorok időbeli változásának lehetősége miatt teszteltük az idősor dinamikus feltételes korrelációját (DCC-GARCH) is, ahol ismét csupán a ZAR és az XAF mutatott valamiféle együttmozgást. (Lásd az 5. táblázatot.) Bár a listaszerű adattörlés mellett magasabbnak tűnik az átlagos feltételes korreláció, kétféle t -próbával csupán a ZAR/USD–EUR/USD páros esetében találni szignifikáns különbséget az átlaggal pótláshoz és EM-hez képest ($p = 0,55$ mellett, míg a többi esetben $p = 0,00$ volt az eredmény).

5. táblázat

Az afrikai devizák EUR/USD-vel szembeni dinamikus feltételes korrelációinak momentumai

Deviza	Listaszerű adattörlés				Átlaggal pótlás				EM-modell			
	1.	2.	3.	4.	1.	2.	3.	4.	1.	2.	3.	4.
	momentum											
KES/USD	0,0451	0,0489	0,1880	2,3352	0,0357	0,0459	0,5519	2,6715	0,0147	0,0354	-1,2180	3,4391
GHS/USD	-0,0010	0,0166	0,0417	6,9048	-0,0055	0,0094	0,1457	8,0598	0,0294	0,0157	0,7023	2,2311
ZAR/USD	0,4290	0,1775	-0,6407	3,4664	0,4119	0,1711	-0,4950	2,9529	0,4119	0,1711	-0,4950	2,9529
TZS/USD	-0,0037	0,0216	0,4166	10,2784	-0,0351	0,0088	-0,3345	1,9282	-0,0059	0,0023	-13,0668	201,9501
UGX/USD	0,0384	0,0214	0,2245	4,8841	0,0569	0,0167	0,2014	1,7215	0,0343	0,0228	0,4147	3,8064
XAF/USD	0,8373	0,1755	-1,3723	4,6628	0,7686	0,2313	-1,4152	5,3184	0,7642	0,2403	-1,4034	5,0187
GMD/USD	-0,0321	0,0151	-0,5513	2,1798	-0,0296	0,0156	-0,5877	2,2286	0,0060	0,0073	0,7218	20,5067
MGA/USD	0,0634	0,0322	-0,1573	2,7352	0,0531	0,0214	-0,0327	2,5063	0,0558	0,0117	0,5373	2,1534
MZN/USD	-0,0054	0,0191	-0,1388	2,1902	-0,0055	0,0390	-0,8071	2,6819	-0,0187	0,0350	-0,9365	3,1454

Forrás: Saját szerkesztés.

4. Összefoglalás

A likelihood-becsléseken alapuló általános várakozásmaximalizációs modellek kifejezetten népszerűnek számítanak napjainkban a kérdőívtípusú vizsgálatok hiányzó adatainak kezelése során, szerves részét képezve a leginkább elterjedt statisztikai programoknak (például az SPSS-nek). Mindazonáltal a pénzügyi idősorok esetében célszerűbb megmaradni az ott egyébként is előszeretettel alkalmazott listaszerű adat-

törlés vagy átlaggal pótlás eljárásainál, amennyiben a vizsgált idősorokban tömege-
sen megjelenik az adathiány. Ezek segítségével ugyanis kevésbé módosulnak a má-
sodik és negyedik momentumok, vagy csökkennek a VaR-jelzések, továbbá a nega-
tív logaritmikus hozamok is kevésbé hajlamosak a volatilitás növekedését okozni. A
piacok, illetve eszközök közötti fertőzések vizsgálata során pedig további torzításo-
kat kerülhetünk el általuk.

Függelék

F1. táblázat

A többváltozós idősorok leíró és alapstatisztikái

Deviza	Átlag	Szórás	Aszimmetria	Csúcsosság	Normális eloszlás Jarque–Bera (<i>p</i>)	Autokorre- láció Ljung–Box (<i>p</i>)	Heteroszke- daszticitás ARCH-LM (<i>p</i>)	Stacionaritás ADF (<i>p</i>)
Listaszerű adattörlés								
KES/USD	0,00	0,01	0,28	18,60	0,00	0,00	0,08	0,00
GHS/USD	0,00	0,01	1,78	39,88	0,00	0,00	0,00	0,00
ZAR/USD	0,00	0,01	1,07	17,89	0,00	0,22	0,50	0,00
TZS/USD	0,00	0,01	-0,87	30,73	0,00	0,00	0,00	0,00
UGX/USD	0,00	0,01	0,46	16,63	0,00	0,00	0,07	0,00
XAF/USD	0,00	0,01	-0,06	5,08	0,00	0,00	0,00	0,00
GMD/USD	0,00	0,02	-0,03	169,73	0,00	0,00	0,03	0,00
MGA/USD	0,00	0,01	1,77	58,07	0,00	0,00	0,00	0,00
MZN/USD	0,00	0,01	0,92	49,84	0,00	0,00	0,00	0,00
EUR/USD	0,00	0,01	-0,05	4,59	0,00	0,83	0,86	0,00
Átlaggal pótlás								
KES/USD	0,00	0,01	0,32	20,60	0,00	0,00	0,22	0,00
GHS/USD	0,00	0,01	1,25	34,57	0,00	0,00	0,00	0,00
ZAR/USD	0,00	0,01	1,05	15,74	0,00	0,01	0,14	0,00
TZS/USD	0,00	0,01	-0,84	40,52	0,00	0,00	0,00	0,00
UGX/USD	0,00	0,01	0,46	16,99	0,00	0,00	0,02	0,00
XAF/USD	0,00	0,01	-0,14	5,26	0,00	0,00	0,00	0,00
GMD/USD	0,00	0,02	-0,05	174,91	0,00	0,00	0,03	0,00
MGA/USD	0,00	0,01	1,61	57,71	0,00	0,00	0,00	0,00
MZN/USD	0,00	0,01	0,81	43,03	0,00	0,00	0,00	0,00
EUR/USD	0,00	0,01	-0,02	4,39	0,00	0,51	0,59	0,00

(A táblázat folytatása a következő oldalon.)

(Folytatás.)

Deviza	Átlag	Szórás	Aszimmetria	Csúcsosság	Normális eloszlás Jarque-Bera (p)	Autokorre- láció Ljung-Box (p)	Heteroszked- aszticitás ARCH-LM (p)	Stacionaritás ADF (p)
	EM-modell							
KES/USD	0,00	0,01	0,98	181,04	0,00	0,00	0,02	0,00
GHS/USD	0,00	0,03	-0,02	358,51	0,00	0,00	0,06	0,00
ZAR/USD	0,00	0,01	1,05	15,74	0,00	0,01	0,14	0,00
TZS/USD	0,00	0,02	0,04	198,98	0,00	0,00	0,00	0,00
UGX/USD	0,00	0,02	0,03	121,75	0,00	0,00	0,00	0,00
XAF/USD	0,00	0,01	-0,21	13,08	0,00	0,00	0,00	0,00
GMD/USD	0,00	0,03	-0,07	74,16	0,00	0,00	0,00	0,00
MGA/USD	0,00	0,04	0,11	24,16	0,00	0,00	0,00	0,00
MZN/USD	0,00	0,02	0,14	55,40	0,00	0,00	0,00	0,00
EUR/USD	0,00	0,01	-0,02	4,39	0,00	0,51	0,59	0,00

Forrás: Itt és a Függelék további táblázatainál saját számítás.

F2. táblázat

A VaR különböző eljárások mellett

Mutató	Deviza									
	KES	GHS	ZAR	TZS	UGX	XAF	GMD	MGA	MZN	EUR
	Listaszerű adattörzés									
Átlag	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Szórás	0,00	0,01	0,01	0,00	0,00	0,01	0,01	0,01	0,01	0,00
Aszimmetria	-0,05	0,09	0,06	0,06	0,00	-0,02	0,02	-0,02	-0,02	0,00
Csúcsosság	3,49	4,31	2,54	3,03	2,87	2,52	4,05	3,69	4,02	2,50
Extrém hozam (-) küszöb	-0,01	-0,02	-0,02	-0,01	-0,01	-0,01	-0,03	-0,02	-0,02	-0,01
Extrém hozam (+) küszöb	0,01	0,02	0,02	0,01	0,01	0,01	0,03	0,02	0,02	0,01
Extrém hozam (-) (db)	130	111	119	119	115	159	84	87	82	160
Extrém hozam (+) (db)	136	105	153	135	130	169	97	90	91	173
Nem extrém hozam (db)	3 146	3 196	3 140	3 158	3 167	3 084	3 231	3 235	3 239	3 079
	Átlaggal pótlás									
Átlag	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Szórás	0,00	0,01	0,01	0,00	0,00	0,00	0,01	0,01	0,01	0,00
Aszimmetria	-0,07	0,13	0,04	0,01	0,00	-0,03	0,05	-0,10	-0,02	0,00
Csúcsosság	3,51	4,47	2,51	3,25	2,88	2,59	4,37	3,95	4,38	2,48
Extrém hozam (-) küszöb	-0,01	-0,02	-0,02	-0,01	-0,01	-0,01	-0,03	-0,02	-0,02	-0,01
Extrém hozam (+) küszöb	0,01	0,02	0,02	0,01	0,01	0,01	0,03	0,02	0,02	0,01
Extrém hozam (-) (db)	138	128	145	121	129	188	97	102	111	189
Extrém hozam (+) (db)	148	124	179	135	149	192	107	105	123	207
Nem extrém hozam (db)	3 626	3 660	3 588	3 656	3 634	3 532	3 708	3 705	3 678	3 516

(A táblázat folytatása a következő oldalon.)

(Folytatás.)

Mutató	Deviza									
	KES	GHS	ZAR	TZS	UGX	XAF	GMD	MGA	MZN	EUR
	EM-modell									
Átlag	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Szórás	0,00	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,01	0,00
Aszimmetria	-0,05	0,04	0,04	-0,16	-0,03	-0,02	0,08	-0,20	-0,05	0,00
Csúcsosság	5,58	9,62	2,51	7,01	5,63	2,66	5,30	10,66	5,11	2,48
Extrém hozam (-) küszöb	-0,02	-0,05	-0,02	-0,03	-0,03	-0,01	-0,06	-0,07	-0,04	-0,01
Extrém hozam (+) küszöb	0,02	0,06	0,02	0,03	0,03	0,01	0,06	0,07	0,04	0,01
Extrém hozam (-) (db)	45	26	145	38	56	144	76	127	94	189
Extrém hozam (+) (db)	50	30	179	41	67	159	84	129	113	207
Nem extrém hozam (db)	3 817	3 856	3 588	3 833	3 789	3 609	3 752	3 656	3 705	3 516

F3. táblázat

A GARCH-modell a korrigált idősorokon

Deviza	Modell	ω	α_1	α_2	γ	β_1	β_2	δ	BIC-
		paraméter							
		Listaszerű adattörlesztés							
KES/USD	TARCH(1,1,2)	0,00	0,23		0,04	0,40	0,36		-4,12
GHS/USD	GARCH(1,2)	0,00	0,13			0,35	0,52		-3,68
ZAR/USD	TARCH(1,1,1)	0,00	0,05		0,07	0,91			-3,13
TZS/USD	GARCH(1,1)	0,00	0,25			0,75			-3,96
UGX/USD	GARCH(1,1)	0,00	0,21			0,79			-3,80
XAF/USD	GARCH(1,1)	0,00	0,04			0,95			-3,57
GMD/USD	TARCH(2,1,1)	0,01	0,15	0,38	0,10	0,42			-2,76
MGA/USD	APARCH(1,1,1)	0,00	0,02		0,05	0,94		3,98	-3,22
MZN/USD	GARCH(1,2)	0,00	0,23			0,30	0,47		-3,36
EUR/USD	GARCH(1,1)	0,00	0,05			0,95			-3,63
		Átlaggal pótlás							
KES/USD	TARCH(1,1,2)	0,00	0,23		0,03	0,44	0,32		-4,20
GHS/USD	APARCH(1,1,1)	0,00	0,05		-0,01	0,89		3,66	-3,69
ZAR/USD	GJR-GARCH(1,1,1)	0,00	0,02		0,10	0,93			-3,21
TZS/USD	GARCH(1,1)	0,00	0,20			0,80			-3,92
UGX/USD	GARCH(1,1)	0,00	0,19			0,81			-3,83
XAF/USD	GARCH(1,2)	0,00	0,06			0,02	0,91		-3,62
GMD/USD	TARCH(2,1,1)	0,00	0,14	0,28	0,17	0,49			-2,83
MGA/USD	APARCH(1,1,1)	0,00	0,02		0,07	0,95		3,44	-3,22
MZN/USD	TARCH(1,1,2)	0,00	0,14		0,12	0,29	0,52		-3,32
EUR/USD	GARCH(1,1)	0,00	0,04			0,96			-3,71

(A táblázat folytatása a következő oldalon.)

(Folytatás.)

Deviza	Modell	ω	α_1	α_2	γ	β_1	β_2	δ	BIC-
		paraméter							
		EM-modell							
KES/USD	GJR-GARCH(1,1,2)	0,00	0,09		-0,05	0,00	0,94		-3,72
GHS/USD	TARCH(1,1,2)	0,00	0,31		-0,10	0,00	0,74		-2,44
ZAR/USD	GJR-GARCH(1,1,1)	0,00	0,02		0,10	0,93			-3,21
TZS/USD	TARCH(1,1,2)	0,00	0,72		-0,43	0,30	0,20		-3,18
UGX/USD	APARCH(1,1,1)	0,00	0,16		-0,19	0,34		4,00	-3,04
XAF/USD		Egy modell sem adott homoszkedasztikus hibátgot							
GMD/USD	GARCH(1,1)	0,00	0,34			0,50			-2,28
MGA/USD	GARCH(1,2)	0,00	0,30			0,38	0,32		-2,35
MZN/USD	GARCH(1,1)	0,00	0,40			0,55			-2,75
EUR/USD	GARCH(1,1)	0,00	0,04			0,96			-3,71

F4. táblázat

Homoszkedasztikus hibátgokkal standardizált logaritmusos hozamokon számított feltétel nélküli korreláció

Deviza	Deviza									
	KES/USD	GHS/USD	ZAR/USD	TZS/USD	UGX/USD	XAF/USD	GMD/USD	MGA/USD	MZN/USD	EUR/USD
	Listaszerű adattörzés									
KES/USD	0,00	0,00	0,05	0,04	0,09	0,06	0,00	0,05	0,00	0,05
GHS/USD	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
ZAR/USD	0,05	0,00	0,00	0,00	0,00	0,38	0,00	0,00	0,00	0,43
TZS/USD	0,04	0,00	0,00	0,00	0,08	0,00	0,00	0,00	0,04	0,00
UGX/USD	0,09	0,00	0,00	0,08	0,00	0,04	0,05	0,00	0,00	0,04
XAF/USD	0,06	0,00	0,38	0,00	0,04	0,00	0,00	0,07	0,00	0,83
GMD/USD	0,00	0,00	0,00	0,00	0,05	0,00	0,00	0,00	0,00	0,00
MGA/USD	0,05	0,00	0,00	0,00	0,00	0,07	0,00	0,00	0,00	0,06
MZN/USD	0,00	0,00	0,00	0,04	0,00	0,00	0,00	0,00	0,00	0,00
EUR/USD	0,05	0,00	0,43	0,00	0,04	0,83	0,00	0,06	0,00	0,00
	Átlaggal pótlás									
KES/USD	0,00	0,00	0,03	0,05	0,08	0,06	0,00	0,00	0,00	0,04
GHS/USD	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
ZAR/USD	0,03	0,00	0,00	0,00	0,00	0,35	0,00	0,00	0,00	0,41
TZS/USD	0,05	0,00	0,00	0,00	0,08	0,00	0,00	0,00	0,00	0,00
UGX/USD	0,08	0,00	0,00	0,08	0,00	0,04	0,04	0,00	-0,05	0,00
XAF/USD	0,06	0,00	0,35	0,00	0,04	0,00	0,00	0,06	0,00	0,76
GMD/USD	0,00	0,00	0,00	0,00	0,04	0,00	0,00	0,00	0,00	0,00
MGA/USD	0,00	0,00	0,00	0,00	0,00	0,06	0,00	0,00	0,00	0,05
MZN/USD	0,00	0,00	0,00	0,00	-0,05	0,00	0,00	0,00	0,00	0,00
EUR/USD	0,04	0,00	0,41	0,00	0,00	0,76	0,00	0,05	0,00	0,00

(A táblázat folytatása a következő oldalon.)

(Folytatás.)

Deviza	Deviza									
	KES/USD	GHS/USD	ZAR/USD	TZS/USD	UGX/USD	XAF/USD	GMD/USD	MGA/USD	MZN/USD	EUR/USD
	EM-modell									
KES/USD	0,00	0,32	0,00	0,17	0,33	0,05	0,08	0,13	0,14	0,03
GHS/USD	0,32	0,00	0,00	0,46	0,64	-0,03	0,33	0,38	0,46	0,00
ZAR/USD	0,00	0,00	0,00	0,00	0,00	0,30	0,00	0,00	0,03	0,41
TZS/USD	0,17	0,46	0,00	0,00	0,41	0,00	0,29	0,29	0,38	0,00
UGX/USD	0,33	0,64	0,00	0,41	0,00	0,00	0,25	0,27	0,29	0,03
XAF/USD	0,05	-0,03	0,30	0,00	0,00	0,00	0,00	0,04	-0,03	0,66
GMD/USD	0,08	0,33	0,00	0,29	0,25	0,00	0,00	0,07	0,36	0,00
MGA/USD	0,13	0,38	0,00	0,29	0,27	0,04	0,07	0,00	0,22	0,00
MZN/USD	0,14	0,46	0,03	0,38	0,29	-0,03	0,36	0,22	0,00	0,00
EUR/USD	0,03	0,00	0,41	0,00	0,03	0,66	0,00	0,00	0,00	0,00

Megjegyzés. Csak a szignifikáns ($p < 0,05$) korrelációkat tartalmazza a táblázat. A szürkével jelzett értékek a nullától érdemben eltérő korrelációt jelölik.

Irodalom

- BARALDI, P. – DI MAIO, F. – GENINI, D. – ZIO, E. [2015]: Reconstruction of missing data in multidimensional time series by fuzzy similarity. *Applied Soft Computing Journal*. Vol. 26. No. 1. pp. 1–9. <http://dx.doi.org/10.1016/j.asoc.2014.09.038>
- BILMES, J. A. [1998]: A gentle tutorial of the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov models. *International Computer Science Institute*. Vol. 510. No. 4. pp. 1–13. <http://melodi.ee.washington.edu/people/bilmes/mypapers/em.pdf>
- CAPPELLO, L. – ENGLE, R. F. – SHEPPARD, K. [2006]: Asymmetric dynamics in the correlations of global equity and bond returns. *Journal of Financial Econometrics*. Vol. 4. Issue 4. pp. 537–572. <http://dx.doi.org/10.1093/jfinec/nbl005>
- CEYLAN, Y. – SIPAN, A. – CEM, I. – INCI, B. [2013]: Comparison of missing value imputation methods in time series: The case of Turkish meteorological data. *Theoretical and Applied Climatology*. Vol. 112. No. 1–2. pp. 143–167.
- ENGLE, R. F. [2002]: Dynamic conditional correlation – A simple class of multivariate GARCH models. *Journal of Business and Economic Statistics*. Vol. 20. No. 3. pp. 377–389. <http://dx.doi.org/10.2139/ssrn.236998>
- GIOVANNI, P. [2005]: Are Euro area small cap stocks an asset class? Evidence from mean-variance spanning tests. *European Financial Management*. Vol. 11. No. 2. pp. 229–253.
- GRAHAM, J. W. [2012]: *Missing Data – Analysis and Design*. Springer. New York, Heidelberg, Dordrecht, London.
- HOUARI, R. – BOUNCEUR, A. – KECHADI, T. – EULER, R. [2013]: A new method for estimation of missing data based on sampling methods for data mining. *Advances in Computational Science*,

- Engineering and Information Technology*. Vol. 225. pp. 89–100. http://dx.doi.org/10.1007/978-3-319-00951-3_9
- IMF (INTERNATIONAL MONETARY FUND) [2013]: *IMF Annual Report on Exchange Arrangements and Exchange Restrictions 2013*. Washington, DC.
- JUAN CARLOS, F. G. – DUSKO, K. – BELLO, L. – AMILCAR, C. [2010]: An evolutionary approach for imputing missing data in time series. *Journal of Circuits, Systems and Computers*. Vol. 19. Issue 1. pp. 107–121. <http://dx.doi.org/10.1142/S0218126610006050>
- JUNGER, W. L. – PONCE DE LEON, A. [2015]: Imputation of missing data in time series for air pollutants. *Atmospheric Environment*. Vol. 102. No. 2. pp. 96–104. <http://dx.doi.org/10.1016/j.atmosenv.2014.11.049>
- KANG, H. [2013]: The prevention and handling of the missing data. *Korean Journal of Anesthesiology*. Vol. 64. No. 5. pp. 402–406. <http://dx.doi.org/10.4097/kjae.2013.64.5.402>
- KEIM, D. B. – STAMBAUGH, R. F. [1984]: A further investigation of the weekend effect in stock returns. *The Journal of Finance*. Vol. 39. Issue 3. pp. 819–835. <http://dx.doi.org/10.1111/j.1540-6261.1984.tb03675.x>
- MADURA, J. [2008]: *International Financial Management*. Cengage Learning. Stamford.
- ORAVECZ B. [2008]: Hiányzó adatok és kezelésük a statisztikai elemzésekben. *Statisztikai Szemle*. 86. évf. 4. sz. 365–385. old.
- ROBINS, R. P. – SMITH, G. P. [2016]: No More Weekend Effect. *Critical Finance Review*. Vol. 5. No. 2. pp. 1–9. <http://cfr.ivo-welch.info/>
- RUUD, P. A. [1991]: Extensions of estimation methods using the EM algorithm. *Journal of Econometrics*. Vol. 49. No. 3. pp. 305–341. [http://dx.doi.org/10.1016/0304-4076\(91\)90001-T](http://dx.doi.org/10.1016/0304-4076(91)90001-T)
- UDVARI B. – PONTET J. [2015]: Költségvetés-támogatás: útban a hatékonyabb segélyezés felé (?) *Külgazdaság*. 59. évf. 1–2. sz. 134–163. old.
- UDVARI, B. – URBÁNNÉ MEZŐ J. [2015]: Az ifjúsági munkanélküliség és a munkaerő-piaci rugalmasság összefüggései az Európai Unióban. *Külgazdaság*. 59. évf. 11–12. sz. 52–78. old.
- SCHNEIDER, T. [2001]: Analysis of incomplete climate data: Estimation of mean values and covariance matrices and imputation of missing values. *Journal of Climate*. Vol. 14. No. 5. pp. 853–871. [http://dx.doi.org/10.1175/1520-0442\(2001\)014%3C0853:AOICDE%3E2.0.CO;2](http://dx.doi.org/10.1175/1520-0442(2001)014%3C0853:AOICDE%3E2.0.CO;2)
- SHAHID, M. N. – MEHMOOD, Z. [2015]: Calendar anomalies in stock market: A case of KSE 100 index. *International Journal of African and Asian Studies*. Vol. 7. pp. 16–23.
- WOTHKE, W. [1998]: Longitudinal and multi-group modeling with missing data. In: Little, T. D. – Schnabel, K. U. – Baumert, J. (eds): *Modeling Longitudinal and Multiple Group Data: Practical Issues, Applied Approaches and Specific Examples*. Lawrence Erlbaum Associates. Mahwah.

Summary

The multivariate time series analysis requires synchronized and continuous data for its models. However, there can be special occasions when one or some data is/are missing due to the lack of trading activity. This paper focuses on the impact of different missing data handling methods on GARCH and value-at-risk model parameters, namely the volatility persistence and asymmetry and

the fat-tailness of the corrected data. The main added value of the current paper is the comparison of the impact of different methods (like listwise deletion, mean-substitution, and maximum-likelihood-based expectation maximization) on daily financial time series because this subject has insufficient literature. The authors have tested daily closing data of floating currencies from Kenya, Ghana, South Africa, Tanzania, Uganda, Gambia, Madagascar and Mozambique in USD denomination against EUR/USD rate between 8 March 2000 and 6 March 2015 acquired from Bloomberg database. Current paper suggests the usage of mean substitution or listwise deletion for daily financial time series due to the biases of the expectation maximization approach on volatility, correlation and extreme fluctuations.