

Metody předvídání volatility s využitím realizované volatility a tržních cen opcí

MILAN FIČURA¹

Abstrakt: Tato práce testuje předpovědní schopnosti a informační obsah 8 modelů běžně používaných pro předvídání volatility (EWMA, GARCH, FIGARCH, ARIMA-RV, ARFIMA-RV, HAR-RV, Black-Scholesova implikovaná volatilita a model-free volatilita). Modely jsou testovány na 5 letech denních dat vývoje měnového kurzu EUR/USD a to pro předvídání denní realizované volatility v horizontu 1 den, 5 dní a 20 dní. Nejlepších výsledků dosáhly modely vycházející z tržních cen opcí (volatilita implikovaná z Black-Scholesova modelu a Model-free volatilita), následovány modely časových řad, vycházejícími z realizované volatility, využívajícími zároveň dlouhou paměť (ARFIMA-RV a HAR-RV), z nichž ARFIMA dominovala zejména na delších předpovědních horizontech, kde překonala i oba opční modely. Testy informačního obsahu dále ukázaly, že opční modely neobsahují veškeré informace obsažené v ryze ekonometrických modelech (ARFIMA a HAR). Z tohoto důvodu bylo pokusně zkonstruováno několik hybridních modelů (kombinujících opční a ekonometrické předpovědi), jejichž výsledky byly na všech horizontech v průměru výrazně lepší než výsledky jednotlivých dílčích modelů samostatně.

AMS/JEL klasifikace: C22, C53, G14

Klíčová slova: Předvídání volatility, realizovaná volatilita, implikovaná volatilita, model-free volatilita

1 Úvod

Předvídání volatility cen finančních aktiv dnes hraje nezastupitelnou roli v mnoha oblastech finanční teorie a praxe - počínaje risk managementem, přes investiční rozhodování, až po oceňování derivátů. Nelze se tedy divit tomu, že v posledních dekáдах vznikla celá řada různých metod k tomu jak volatilitu předvídat. Tyto metody lze rozdělit do dvou základních skupin - na metody vycházející z historických časových řad a metody vycházející z aktuálních tržních cen opcí. Zatímco první skupina je založena na ryze ekonometrických postupech a využívá pouze informací obsažených v minulém vývoji dané časové řady (je backward-looking), druhá skupina metod se pokouší na základě bezarbitrážních vztahů z tržních cen opcí vyvodit, jakou úroveň volatility očekávají účastníci trhu do budoucna (je forward-looking). Opční předpovědi volatility by tak dle teorie efektivních trhů měly obsahovat jak všechny historické informace o vývoji dané časové řady, tak také všechny ostatní relevantní informace, které jsou účastníkům trhu k dispozici, a měly by tak dosahovat vždy lepších výsledků, než předpovědi vycházející jen z historických časových řad (Figlewski 2004). Praktickým problémem spojených s jejich použitím však je, že obvykle nejsou splněny všechny předpoklady, z nichž jednotlivé opční modely vycházejí. Obzvlášť problémová je v tomto ohledu existence tzv. volatilitní premie, která je v cenách opcí obsažena a která vede k

¹Vysoká škola ekonomická v Praze, Náměstí Winstona Churchilla 4, Praha 3, E-mail: xficm03@vse.cz
Výzkum byl podpořen z grantu "Pokročilé modelování finančních rizik" IGA VŠE F1/6/2012

systematickému nadhodnocování volatility vyvozované z tržních cen opcí oproti následně realizované volatilítě (Bakshi a Kapadia 2003).

To vede k tomu, že i přesto, že ve většině empirických studií dosahují opční modely zpravidla lepších výsledků než modely časových řad (podrobný přehled viz v Poon a Granger 2003), zejména v porovnání s modernějšími ekonometrickými postupy (především pak těmi využívajícími informací z vysoko-frekvenčních dat) již jejich náskok není zdaleka tak jednoznačný (viz Pong a kol. 2003). Významnou otázkou též je, do jaké míry opční modely skutečně obsahují ve svých předpovědích všechny informace obsažené v ekonometrických modelech, na což se názory vesměs různí (viz Pong a kol. 2003, Taylor a kol. 2007, atd.).

Cílem této práce je porovnat mezi sebou několik v současné době nejpoužívanějších ekonometrických a opčních metod předvídání volatility, zhodnotit jejich předpovědní schopnosti, a zjistit, zda opční předpovědi zahrnují všechny informace z ryze ekonometrických modelů, či nikoliv. Dále se též zaměříme na to, zda by bylo možné využitím kombinovaných předpovědí na základě opčních i ekonometrických modelů dosáhnout lepších výsledků, než při použití kterékoliv z obou metod samostatně.

Pro porovnání jednotlivých modelů byla vybrána časová řada kurzů EUR/USD a to z toho důvodu, že na rozdíl od časových řad mnoha jiných finančních aktiv disponuje určitými velice žádoucími vlastnostmi, které značně ulehčují praktické použití některých metod (dostupnost kvalitních intradenních dat, obchodování 24 hodin denně, absence leverage efektu a s ním spojených asymetrií, atd.). Modely jsou testovány na 5 letech denních dat vývoje měnového kurzu EUR/USD, přičemž předpovědi volatility jsou generovány v horizontu 1 den, 1 týden (5 dní) a 1 měsíc (30 dní) a to s využitím překrývajících se i nepřekrývajících se předpovědních intervalů. Jako měřítko volatility, které se pokoušíme předvídat, byla zvolena realizovaná volatilita agregovaná na základě kombinace 15minutových a 30minutových časových intervalů. Mezi testované modely spadají ekonometrické modely EWMA, GARCH, FIGARCH, ARIMA-RV, ARFIMA-RV, HAR-RV a dále dva opční modely, a to Black-Scholesova implikovaná volatilita a Model-free volatilita.

Modely EWMA, GARCH a FIGARCH se vyznačují tím, že pro předvídání volatility na určité frekvenci využívají čtverců logaritmičkových výnosů počítaných na té samé frekvenci. Ty představují sice nevychýlený, ale jen značně nepřesný způsob odhadu skutečné volatility, ze které výnosy na dané frekvenci vycházejí (Andersen a kol. 2005). Daleko přesnějšího odhadu volatility je možno dosáhnout s využitím informací z vysoko-frekvenčních dat (jsou-li tyto data v dostatečné kvalitě k dispozici). Toho využívá tzv. realizovaná volatilita (Andersen a Bollerslev 1998), definovaná pro určitou frekvenci jako součet čtverců logaritmičkových výnosů na nějaké vyšší frekvenci. Takto definovaná realizovaná volatilita přitom (při splnění určitých podmínek) podává nestranný, jakož i konzistentní odhad integrované volatility pro danou frekvenci a je tak možno ji chápat jako přímo pozorovatelnou míru volatility a aplikovat na ni standardní modely časových řad jako jsou ARIMA, ARFIMA či HAR.

Z opčních modelů volatility byly testovány dva nejpoužívanější a to implikovaná volatilita z Black-Scholesova modelu oceňování opcí a dále Model-free volatilita, která na rozdíl od B-S implikované volatility neuvažuje žádný konkrétní proces pro vývoj ceny podkladového aktiva a očekávanou volatilitu vyvozuje pouze s využitím bez-arbitrážních vztahů (B-S implikovaná

volatilita naproti tomu uvažuje, že cena aktiva koná geometrický Brownův pohyb s konstantní volatilitou). Model-free volatilita též pracuje s celou sérií opcí a využívá tak úplnějšího souboru informací než B-S implikovaná volatilita, která pracuje jen s jednou opcí. Na druhou stranu však Model-free volatilita pracuje s kontinuem tržních cen opcí při všech variantách různých strike-price, což je v praxi nereálný požadavek, který vede k nutnosti využití interpolací a extrapolací pro dopočet cen opcí které nejsou k dispozici, což předpovědní sílu tohoto modelu značně snižuje.

V následující kapitole se podrobněji zaměříme na základní míry volatility, jakož i na charakteristiku jednotlivých modelů, které budou v empirické části práce použity. Závěrečná kapitola obsahuje empirický výzkum, kde jsou jednotlivé modely aplikovány na reálná data, za účelem srovnání jejich předpovědních schopností a informačního obsahu.

2 Metodologie a charakteristika jednotlivých modelů

Pod pojmem volatilita v této práci rozumíme směrodatnou odchylku logaritmických výnosů určitého aktiva, u nichž uvažujeme, že v diskrétním čase konají proces:

$$r_t = \mu_t + \varepsilon_t = \mu_t + \sigma_t z_t$$

Kde $r_t = p_t - p_{t-1}$ představuje logaritmický výnos (p_t je logaritmus ceny aktiva v čase t), μ_t podmíněnou střední hodnotu v čase t , ε_t náhodnou složku procesu, σ_t podmíněnou volatilitu a z_t bílý šum, pro který platí $z_t \sim i. i. d.$, $E(z_t) = 0$, $Var(z_t) = 1$.

Podmíněnou střední hodnotu budeme dále považovat za zanedbatelnou, takže platí:

$$r_t = \sigma_t z_t$$

Čtverce logaritmických výnosů r_t^2 tak v tomto případě poskytují nevychýlený odhad podmíněného rozptylu σ_t^2 , jsou však zároveň zatíženy chybou v podobě z_t^2 .

Ve spojitém čase budeme dále uvažovat, že logaritmus ceny aktiva vykonává proces:

$$dp(t) = \mu(t)dt + \sigma(t)dW(t)$$

Kde $dp(t)$ značí diferenciál logaritmu ceny aktiva, $\mu(t)$ okamžitou střední hodnotu (tzv. drift rate), $\sigma(t)$ okamžitou volatilitu (spotovou volatilitu) a $dW(t)$ diferenciál Wienerova procesu.

Zanedbáme-li drift rate, získáme proces:

$$dp(t) = \sigma(t)dW(t)$$

Pohyb logaritmu ceny v jednom dni lze tak vyjádřit jako:

$$r(t) = \int_{t-1}^t \sigma(s)dW(s)$$

Z čehož lze pomocí kvadratické variace definovat integrovaný rozptyl pro daný den jako:

$$IV(t) = \int_{t-1}^t \sigma^2(s)ds$$

Integrovaná denní volatilita (odmocnina z integrovaného rozptylu) je přitom kvantitou, kterou se budeme pokoušet předvídat. Vzhledem k tomu, že integrovaný rozptyl je ze své podstaty nepozorovatelný, využíváme pro jeho odhad realizovaný rozptyl, který je definován jako:

$$RV(t, \Delta) = \sum_{j=1}^{\frac{1}{\Delta}} r^2(t - 1 + j\Delta, \Delta)$$

Kde $r(t, \Delta)$ je logaritmický výnos za období $t - \Delta$ až t . Jedná se tedy o součet čtverců logaritmických výnosů na určité vyšší frekvenci, pro odhad integrované volatilita na nějaké nižší frekvenci (Andersen a Bollerslev 1998).

Realizovaná volatilita představuje nestranný a konzistentní odhad integrované volatilita. S poklesem délky intervalu Δ tak realizovaná volatilita konverguje k integrované volatilitě ($RV(t, \Delta) \rightarrow IV(t)$ když $\Delta \rightarrow 0$) a chyba odhadu klesá k nule. V praxi však vznikají problémy s využitím příliš krátkých frekvencí pro výpočet realizované volatilita a to zejména kvůli efektům fungování tržní mikrostruktury a s nimi spojenou negativní autokorelací výnosů na příliš vysokých frekvencích (zejména patrný je tento efekt na minutových či tick datech). To vede k nadhodnocování realizované volatilita oproti zkoumané integrované volatilitě (Andersen a kol. 2005). V praxi je tak třeba buď dané logaritmické výnosy o výše zmíněné efekty tržní mikrostruktury upravit, nebo používat nižší frekvence, na nichž je toto zkreslení stále ještě zanedbatelné (a chyba odhadu je již dostatečně malá). V našem případě byly použity 15 minutové a 30 minutové výnosy, na nichž se u zkoumané časové řady již prakticky žádná statisticky významná autokorelace nevyskytovala.

První skupina modelů testovaných v této práci vychází z výše definovaného procesu vývoje logaritmických výnosů v diskrétním čase a pro předvídaní volatilita využívá čtverců logaritmických výnosů počítaných na té samé frekvenci pro jakou je volatilita předvídána. Nebudeme přitom uvažovat tzv. modely stochastické volatilita, které pracují s náhodnou složkou v procesu vykonávaném volatilitou, čímž se rovnice pro cenový proces redukuje na:

$$r_t = \sigma_{t|t-1} Z_t$$

Kde $\sigma_{t|t-1}^2 = \text{Var}[r_t | \mathcal{F}_{t-1}]$ a podmíněný rozptyl v čase t je tak plně určen informační množinou v čase $t - 1$.

Prvním z těchto modelů se kterým budeme pracovat je model EWMA (Exponentially Weighted Moving Average), spadající do skupiny filtrů s nekonečnou impulzní odezvou. Tento model předpovídá budoucí volatilitu na základě váženého průměru z předchozích čtverců logaritmických výnosů, přičemž váhy starších pozorování exponenciálně klesají k nule. Matematická formulace modelu vypadá následovně:

$$\hat{\sigma}_t^2 = \gamma r_t^2 + (1 - \gamma) \hat{\sigma}_{t-1}^2 = \gamma \sum_{i=1}^{\infty} (1 - \gamma)^{i-1} r_{t-i}^2$$

Kde faktor γ determinuje rychlost exponenciálního poklesu vah starších pozorování. Tento parametr je buď možné optimalizovat, nebo jej zvolit expertně. V našem případě byl zvolen

expertní odhad ve výši $\gamma = 0,06$, který vychází z metodologie RiskMetrics publikované bankou J.P.Morgan v roce 1994.

Výhodou modelu EWMA je jeho jednoduchost a snadná použitelnost. Nevýhodou je, že na změny volatility reaguje často příliš pomalu, nedokáže se přesně přizpůsobit autokorelační struktuře logaritmických výnosů (rezidua modelu tak zpravidla stále vykazují autokorelaci) a nedokáže ani zachytit tzv. mean-reverzní vlastnost volatility, neboli tendenci navracet se ke své nepodmíněné střední hodnotě (Andersen a kol. 2005).

Výše zmíněné problémy modelu EWMA se pokouší řešit model GARCH (Generalized ARCH), představující zobecnění známého modelu ARCH (Autoregressive conditional heteroskedasticity). Model ARCH (Engle 1982) předpovídá volatilitu na základě lineární kombinace zpožděných čtverců logaritmických výnosů:

$$\sigma_{t|t-1}^2 = \alpha_0 + \alpha_1 r_{t-1}^2 + \alpha_2 r_{t-2}^2 + \dots + \alpha_q r_{t-q}^2 = \alpha_0 + \sum_{i=1}^q \alpha_i r_{t-i}^2$$

Kde pro parametry platí, že $\alpha_0 > 0$ a pro $i > 0$ je $\alpha_i \geq 0$.

Nevýhodou tohoto modelu je, že pro přesné zachycení autokorelační struktury logaritmických výnosů často potřebuje velké množství členů (zpožděných výnosů) a je pak nutné odhadovat příliš mnoho parametrů. Tento problém řeší model GARCH (Bollerslev 1986), který do rovnice přidává zpožděné hodnoty podmíněných rozptylů.

$$\begin{aligned} \sigma_{t|t-1}^2 &= \alpha_0 + \alpha_1 r_{t-1}^2 + \dots + \alpha_q r_{t-q}^2 + \beta_1 \sigma_{t-1|t-2}^2 + \dots + \beta_p \sigma_{t-p|t-p-1}^2 = \\ &= \alpha_0 + \sum_{i=1}^q \alpha_i r_{t-i}^2 + \sum_{i=1}^p \beta_i \sigma_{t-i|t-i-1}^2 \end{aligned}$$

Kde opět platí, že $\alpha_0 > 0$ a pro $i > 0$ je $\alpha_i \geq 0$ a $\beta_i \geq 0$.

K relativně přesnému zachycení autokorelační struktury zkoumaného procesu pak často stačí jednoduchá specifikace v podobě modelu GARCH(1,1), kterou budeme používat i v této práci. Ta má tvar:

$$\sigma_{t|t-1}^2 = \omega + \alpha r_{t-1}^2 + \beta \sigma_{t-1|t-2}^2$$

Jednou z hlavních nevýhod modelu GARCH je, že se jedná o takzvaný model s krátkou pamětí, který pracuje s exponenciálním poklesem autokorelační funkce zkoumaného procesu (volatility). Četné empirické výzkumy přitom ukazují, že autokorelační funkce čtverců logaritmických výnosů (jakož i realizované volatility) klesá daleko pomaleji než exponenciálně a to hyperbolicky (vykazuje tzv. dlouhou paměť, viz. Ding a kol. 1993). Tu je možno v některých případech do určité míry substituovat nekonečnou pamětí, kterou disponuje model IGARCH (Integrated GARCH), který vznikne (v podobě modelu IGARCH(1,1,1)) z modelu GARCH(1,1), pokud zvolíme parametry tak, aby platilo že $\alpha + \beta = 1$. Nevýhodou tohoto modelu je, že je nestacionární a neobsahuje ani vlastnost mean-reversion, což je v rozporu s empiricky pozorovanými vlastnostmi volatility.

Za účelem přesnějšího zachycení dlouhé paměti v autokorelační struktuře volatilit, jakož i současném zachování stacionarity a mean-reverze, byl proto Bailem a kol. (1996) vyvinut model FIGARCH (Frakcionally Integrated GARCH), představující zobecnění modelu IGARCH, ve kterém frakcionální parametr d může nabývat i neceločíselných hodnot. Model FIGRACH(1,d,1) má podobu:

$$\sigma_{t|t-1}^2 = \omega + \beta\sigma_{t-1|t-2}^2 + [1 - \beta L - (1 - \alpha L - \beta L)(1 - L)^d]r_{t-1}^2$$

Kde L představuje operátor zpoždění a d frakcionální parametr. Při nastavení $d = 0$ pak vzniká model GARCH(1,1) s krátkou pamětí, zatímco při nastavení $d = 1$ model IGARCH(1,1,1) s nekonečnou pamětí. Při $0 < d < 1$ vzniká model s dlouhou pamětí, přičemž pro $d < 0,5$ je v daném procesu stále zachována stacionarita a mean-reverze, zatímco pro $0,5 < d < 1$ se již jedná o nestacionární proces, který je však stále mean-reverzní a to až do $d \geq 1$, kdy mean-reverzní být přestává.

Všechny výše zmíněné modely vycházejí při předvídání volatilit z čtverců logaritmičeských výnosů počítaných na stejné frekvenci, pro jakou je volatilita daným modelem předvídána. Jak již bylo řečeno výše, tento způsob odhadu volatilit ve sledovaném období je značně nepřesný a jsou-li dostupné intradenní data, je vhodnější použít pro odhad volatilit realizovanou volatilitu. S tou je následně možno pracovat jako s přímo pozorovatelnou mírou integrované volatilit a její budoucí hodnoty předvídát pomocí modelů časových řad jako jsou modely ARIMA, ARFIMA či HAR.

Modely ARIMA a ARFIMA se mezi sebou liší opět pouze v tom, že zatímco v modelu ARIMA (Box a Jenkins 1970) může parametr d nabývat jen celočíselných hodnot (a model je tak schopný zachytit jen krátkou paměť, či nekonečnou paměť), v modelu ARFIMA (Granger a Joyeux 1981) může d nabývat i hodnot neceločíselných (a model je tak schopný zachytit i dlouhou paměť). Oba modely lze úsporným způsobem zapsat následovně:

$$\phi(L)(1 - L)^d(RV(t, \Delta) - \mu_{RV}) = \psi(L)z_t$$

Kde $\phi(L)$ a $\psi(L)$ představují polynomiální AR a MA operátory:

$$\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p$$

$$\psi(L) = 1 + \psi_1 L + \psi_2 L^2 + \dots + \psi_q L^q$$

A kde L představuje operátor zpoždění a d frakcionální parametr. Při volbě $d = 0$ se jedná o stacionární model ARMA(p,q) s krátkou pamětí, zatímco při $d = 1$ se jedná o nestacionární model ARIMA(p,1,q) s nekonečnou pamětí. Při volbě neceločíselné hodnoty d se pak jedná o model ARFIMA, kdy pro $-1 < d < 1/2$ je daný model stacionární a mean-reverzní, pro $1/2 < d < 1$ již je nestacionární, avšak stále ještě mean-reverzní a pro $d \geq 1$ již není ani stacionární ani mean-reverzní (viz Ishida a Watanabe 2008). V empirických výzkumech vychází d pro časové řady realizované volatilit obvykle okolo 0,4 (Andersen a kol. 2005), což poukazuje na stacionární, avšak vysoce perzistentní proces, kde i velice stará pozorování (často až 100 period) ovlivňují statisticky významně současnou volatilitu. V této práci jsou testovány dva základní modely výše zmíněného typu a to ARMA(1,1) a ARFIMA(1,d,1).

Jiný model pro zachycení dlouhé paměti v autokorelační struktuře volatilitu než model ARFIMA je model HAR (Heterogenous Autoregressive) představený Corsim (2004). Ten vychází z teorie heterogenních trhů, která uvažuje, že na trhu figurují různé subjekty rozhodující se na základě různých kritérií a v různých časových horizontech. To mimo jiné způsobuje, že se volatilita kaskádním způsobem přelévá z nižších frekvencí směrem k vyšším, jelikož obchodníci s krátkodobým horizontem mají tendenci sledovat volatilitu i na vyšších frekvencích, zatímco obchodníci s dlouhodobým horizontem krátkodobé změny volatilitu příliš nesledují. A právě tak vzniká dle této teorie v autokorelační struktuře volatilitu dlouhá paměť. Model HAR se pokouší reflektovat toto vysvětlení dlouhé paměti a předvídá proto volatilitu na základě kombinace tří AR procesů, které odrážejí investiční horizonty různých typů obchodníků. Využívá přitom realizované volatilitu agregované vždy za poslední den, za poslední týden (5 dní) a za poslední měsíc (20 dní). Model je možno zapsat následovně:

$$\sigma_{t+1d}^{(d)} = c + \beta^{(d)}RV_t^{(d)} + \beta^{(w)}RV_t^{(w)} + \beta^{(m)}RV_t^{(m)} + \omega_{t+1d}^{(d)}$$

Kde indexy d , w a m označují období (den, týden, či měsíc), za které je v daném případě realizovaná volatilita agregována.

Výhoda tohoto modelu je zejména jeho jednoduchost. Nevýhoda je, že pracuje s určitou konkrétní podobou dlouhé paměti v autokorelační struktuře volatilitu a není tak v tomto ohledu tak flexibilní jako model ARFIMA, který je schopný se pomocí frakcionálního parametru d přímo přizpůsobit autokorelační struktuře volatilitu pro daný konkrétní případ.

Dále jsou v této práci použity dva modely vycházející z tržních cen opcí, a to Black-Scholesova implikovaná volatilita Model-free volatilita. Jak již bylo zmíněno, výhodou těchto modelů je, že pracují s očekáváním trhu a měly by tak krom historických informací o vývoji volatilitu v minulosti reflektovat i všechny ostatní relevantní informace pro její předvídání, které jsou účastníkům trhu k dispozici (podrobněji viz Figlewski 2004). Nevýhodou těchto modelů jsou některé jejich předpoklady, které v praxi zpravidla nebývají splněny.

B-S implikovaná volatilita vychází ze známého Black-Scholesova modelu oceňování opcí (Black a Scholes 1973). Ten má v případě evropské call opce tvar:

$$C(S, t) = Se^{-q(T-t)}N(d_1) - Ke^{-r(T-t)}N(d_2)$$

$$d_1 = \frac{\ln\left(\frac{S}{K}\right) + (r - q)(T - t)}{\sigma\sqrt{T - t}} + \frac{\sigma\sqrt{T - t}}{2}$$

$$d_2 = d_1 - \sigma\sqrt{T - t}$$

Kde $C(S, t)$ je cena opce, S cena podkladového aktiva, K je strike price, r bezriziková úroková míra, q výnosnost podkladového aktiva, $T - t$ doba do splatnosti opce a σ volatilita očekávaná do doby splatnosti. Implikovaná volatilita je z modelu vyvozována tak, že se za $C(S, t)$ dosadí tržní cena opce a volatilita σ se dopočte jako neznámá. Bohužel, implikovanou volatilitu není možné z modelu vyjádřit přímo analyticky a je tak nutné ji počítat pomocí numerických metod (metoda půlení intervalu, Newton-Raphsonova metoda, Brentova metoda, atd.). V našem případě byly pro předvídání volatilitu použity již vypočtené hodnoty

implikovaných volatilit z at-the-money opcí, získané prostřednictvím aplikace Reuters3000 Xtra a nebylo tak nutné provádět vlastní výpočet.

Hlavní nevýhodou B-S modelu je, že předpokládá konstantní volatilitu do doby splatnosti opce a geometrický Brownův pohyb pro vývoj ceny podkladového aktiva. To jsou v praxi značně nereálné předpoklady, které podhodnocují pravděpodobnost extrémních cenových pohybů, jelikož implikují normální rozdělení výnosů, přičemž empirické rozdělení výnosů je daleko špičatější a má tlusté konce. To vede k tomu, že má Black-Scholesův model tendenci in-the-money a out-of-the-money opce podhodnocovat a at-the-money opce nadhodnocovat (přinejmenším na měnových trzích). Při použití různých opcí (s různými strike price) pro výpočet implikované volatility je pak volatilita vypočtená z in-the-money a out-of-the-money opcí typicky vyšší než volatilita vypočtená z at-the-money opcí (vzniká zde tzv. volatility smile). Objevuje se zde tedy otázka, které opce pro výpočet implikované volatility použít? V praxi jsou nejčastěji používány at-the-money opce, jelikož bývají zpravidla jedny z nejlikvidnějších (Figlewski 2004), což zvyšuje informační hodnotu jejich cen. Jinou možností je použití opcí s nejnižší vega, jejichž implikovaná volatilita by teoreticky měla být zatížena nejmenší chybou odhadu (Wang a kol. 2009), případně je též možno použít kombinace několika různých opcí (pro shrnutí jednotlivých metod viz Wang a kol. 2009).

Některé výše zmíněné problémy B-S implikované volatility se pokouší řešit Model-free volatilita (Neuberger a Britten-Jones 2000), která nepředpokládá žádný konkrétní proces pro pohyb ceny podkladového aktiva, nýbrž vyvozuje trhem očekávanou volatilitu jen na základě bez-arbitrážních vztahů. Využívá přitom zpravidla všech dostupných cen opcí s daným datem splatnosti a vychází tak z většího množství informací než B-S implikovaná volatilita využívající jen jedné opce.

Východiskem Model-free volatility je zjištění, že očekávaná integrovaná volatilita do času T při rizikově neutrálních pravděpodobnostech je plně určena kontinuem cen opcí s maturitou v čase T . To bylo dokázáno pro všechny difúzní procesy (Neuberger a Britten-Jones 2000) a následně zobecněno i na procesy obsahující skoky (Jiang a Tian 2005). Tuto závislost, vycházející z bez-arbitrážních vztahů, lze zapsat následovně:

$$E_0^F \left[\int_0^T \left(\frac{dF_t}{F_t} \right)^2 \right] = 2 \int_0^\infty \frac{C^F(T, K) - \max(0, F_0 - K)}{K^2} dK$$

Kde $C^F(T, K) = C(T, K)/B(T, t)$ značí forwardovou cenu opce s maturitou v čase T , $B(T, t)$ značí cenu bondu v čase t , který vyplatí 1 dolar v čase T , F_t značí forwardovou cenu aktiva s maturitou v čase T a E_0^F značí očekávání při forwardové pravděpodobnostní míře.

Očekávaná volatilita je tak určena integrálem rozdílů mezi forwardovými cenami opcí a jejich vnitřními hodnotami vypočteném přes celé kontinuum možných strike price. Bohužel, v praxi je obvykle dostupných jen několik málo aktivně obchodovaných opcí a pro dopočet tržních cen zbylých opcí, které model využívá, je nutno použít interpolaci a extrapolaci z těch několika málo cen opcí, které známe. Často používaný postup, vypracovaný Jiangem a Tianem (2005), využívá cen dostupných at-the-money a out-of-the-money call a put opcí (out-of-the-money opce jsou upřednostňovány před in-the-money opcemi, jelikož bývají zpravidla likvidnější). Z těchto cen jsou nejprve spočítány B-S implikované volatility, které jsou

interpolovány kubickou spline funkcí pro výpočet volatility smilu. Ze získaného volatility smilu jsou pak zpětně dopočteny ceny opcí, při všech variantách různých strike price, které jsou modelem požadovány pro dostatečně přesný výpočet finálního integrálu, který je počítán numericky (v našem případě lichoběžníkovou metodou). Jiang a Tian též při konstrukci volatility smilu provádějí extrapolaci implikovaných volatilit na okraji smilu konstantní funkcí. My jsme vyzkoušeli 4 možné postupy pro výpočet požadovaného smilu a to lineární a kubickou spline interpolaci, vždy s a bez použití extrapolace konstantní funkcí. Použití extrapolace přitom v našem případě vždy vedlo k výraznému poklesu předpovědních schopností modelu. K nejlepším výsledkům na in-sample datech vedla kubická-spline interpolace (bez extrapolace) a ta byla také použita ve zbytku této práce.

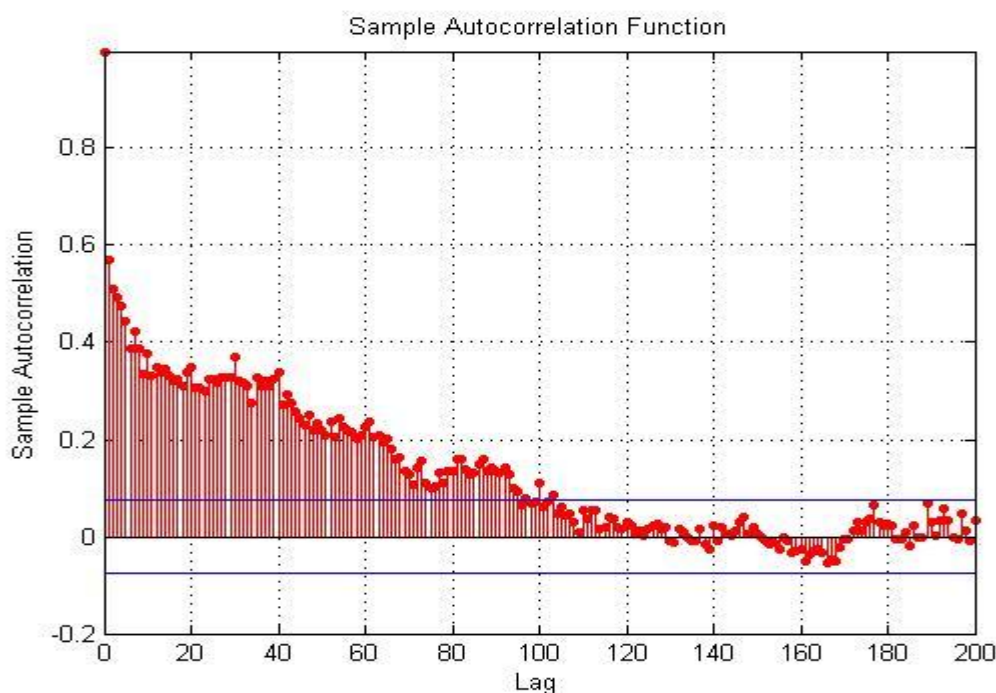
S oběma výše zmíněnými opčními modely je spojena ještě další komplikace, způsobená tím, že nezohledňují vliv volatilitní prémie, která je v tržních cenách opcí obsažena, a která způsobuje, že opční předpovědi volatility realizovanou volatilitu systematicky nadhodnocují (Bakshi a Kapadia 2003). Pro úspěšné předvídaní realizované volatility je proto nutné předpovědi těchto modelů o vliv volatilitní prémie upravit. Toho lze docílit zkonstruováním jednoduché regresní rovnice, kdy za vysvětlovanou proměnnou dosazujeme realizovanou volatilitu v in-sample období a za vysvětlující proměnnou vždy opční předpověď pro daný den. Odhadnuté koeficienty regresní rovnice nám pak sdělují, jakým způsobem daná opční předpověď realizovanou volatilitu nadhodnocuje a je též možné jich využít pro úpravu budoucích předpovědí pro out-of-sample období.

3 Empirický výzkum

Výše popsaných 8 modelů volatility (EWMA, GARCH, FIGARCH, ARIMA, ARFIMA, HAR, B-S implikovaná volatilita a Model-free volatilita) bylo aplikováno na vývoj časové řady měnového kurzu EUR/USD pro posouzení jejich předpovědních schopností. Celkové období o délce 1395 dní (od 4.12.2007 do 18.4.2012) bylo rozděleno na development sample o délce 698 dní (od 4.12.2007 do 13.8.2010) a testing sample o délce 697 dní (od 13.8.2010 do 18.4.2012). Na development sample byly realizovány veškeré odhady parametrů jednotlivých modelů, načež byly již odhadnuté modely aplikované na testing sample pro posouzení jejich out-of-sample předpovědních schopností. Cílem výzkumu bylo předvídat denní realizovanou volatilitu v horizontu 1 den, 5 dní a 20 dní, přičemž realizovaná volatilita byla agregována z 15minutových a 30minutových časových intervalů (z ryze praktických důvodů souvisejících s nedostupností potřebných dat). Výpočet ukázal, že ani na jedné z těchto frekvencí logaritmické výnosy nevykazovaly prakticky žádnou statisticky významnou autokorelaci (na rozdíl od 1minutových a 5minutových frekvencí, kde autokorelace stále přetrvávala) a je tedy možné je označit za vhodné pro výpočet realizované volatility.

Dále byly vypočteny autokorelační funkce čtverců logaritmických výnosů, které vykazovaly statisticky vysoce významnou autokorelaci na všech zkoumaných frekvencích (od 1 minuty po 1 měsíc). Ještě významnějších autokorelačních vztahů pak dosahovala realizovaná volatilita (viz obrázek 1), kde odhad frakcionálního parametru d v programu Cronos dosáhl hodnoty 0,451713, což poukazuje na stacionární, avšak vysoce persistentní proces, blízký se hranici nestacionarity (0,5).

Obrázek 1 - Autokorelační funkce denní realizované volatility počítané z 15min výnosů



Parametry všech ekonometrických modelů byly odhadnuty pomocí programu Matlab (s využitím metod založených na maximalizaci věrohodnostní funkce), s výjimkou modelu ARFIMA, u kterého byl použit program Cronos (odhad opět pomocí maximum likelihood).

Opční modely volatility byly v prvním kroku aplikovány na in-sample období ve své původní podobě, bez úpravy o vliv volatilní prémie. Hodnoty R-squared dosahovaly v tomto případě 0,5592 u B-S implikované volatility a 0,5864 u Model-free volatility.

Následně byla na předpovědi všech modelů aplikována Mincer-Zarnowitz regression (Mincer a Zarnowitz 1969), pro posouzení toho, nakolik jsou tyto předpovědi zkreslené (tedy v případě opčních modelů zejména pro posouzení vlivu volatilní prémie). Použitá regresní rovnice vypadala následovně:

$$y_{t+1} = \alpha + \beta \hat{y}_{t+1|t} + \epsilon_{t+1}$$

Kde y_{t+1} značí realizovanou volatilitu v čase $t+1$ a $\hat{y}_{t+1|t}$ předpověď této volatility uskutečněnou v čase t , některým z testovaných modelů. Pokud by předpovědi generované daným modelem byly nezkreslené, pak by mělo platit, že $\alpha \cong 0$ a $\beta \cong 1$. To bylo v našem případě splněno u všech ekonometrických modelů, s výjimkou modelu GARCH, kde β dosáhla překvapivě hodnoty 1,15. U obou opčních modelů pak byla β výrazně nižší než 1, a to $\beta = 0,808$ u B-S implikované volatility a $\beta = 0,859$ u Model-free volatility (koeficienty α byly v obou případech statisticky nevýznamně odlišné od nuly). To poukazuje na to, že oba opční modely realizovanou volatilitu nadhodnocují, pravděpodobně v důsledku existence volatilní prémie. Oba modely byly proto o vliv tohoto zkreslení upraveny, vynásobením jimi generovaných předpovědí koeficientem β . Výsledky modelů po této úpravě zobrazuje následující tabulka.

Tabulka 1 - Výsledky jednotlivých modelů v in-sample a out-of-sample období

IN-SAMPLE	EWMA	GARCH	FIGARCH	ARIMA	ARFIMA	HAR	IV	MFV
SSR	1,64E-06	1,70E-06	1,86E-06	1,45E-06	1,40E-06	1,41E-06	1,27E-06	1,26E-06
R-Squared	0,4974	0,4772	0,4300	0,5557	0,5705	0,5673	0,6099	0,6126
OUT-SAMPLE	EWMA	GARCH	FIGARCH	ARIMA	ARFIMA	HAR	IV	MFV
SSR	3,04E-07	3,15E-07	3,46E-07	2,63E-07	2,58E-07	2,56E-07	2,36E-07	2,45E-07
R-Squared	0,2786	0,2514	0,1789	0,3745	0,3877	0,3913	0,4390	0,4180

Z tabulky je patrné, že nejlepších výsledků dosáhly v obou obdobích modely vycházející z tržních cen opcí, což koresponduje s tím, že vycházejí z širšího souboru informací, než ryze ekonometrické modely. Z porovnání výsledků B-S IV a MFV zároveň vyplývá, že modernější MFV (využívající vícero cen opcí a realističtějších předpokladů ohledně vývoje ceny podkladového aktiva) nedosáhla lepších výsledků než tradiční B-S IV. K podobným závěrům dospěly i některé dřívější studie (Muzzioli 2008), přičemž jedním z možných vysvětlení je, že Model-free volatilita má tendenci být citlivější na změny volatilitní prémie než B-S implikovaná volatilita (viz Andersen a Bondarenko 2007).

Na dalším místě se umístily modely vycházející z konceptu realizované volatility, z nichž modely s dlouhou pamětí (ARFIMA a HAR) dosáhly v obou obdobích lepších výsledků než model ARIMA obsahující jen krátkou paměť. To je v souladu s tvarem empiricky pozorované autokorelační funkce volatility, jakož i s dříve provedenými studiemi (Corsi 2004).

Nejhorších výsledků dosáhly modely předvídající denní volatilitu na základě denních čtverců logaritmických výnosů. Kupodivu z nich přitom dosáhl nejlepších výsledků model EWMA, který je teoreticky nejméně sofistikovaný a nebyl v našem případě ani nijak přizpůsoben datům (v souvislosti s čímž je zářející tak výrazný propad mezi jeho výsledky in-sample a out-of-sample). Za ním se umístil populární model GARCH a až na posledním místě model FIGARCH, který by teoreticky měl podávat lepší výsledky než model GARCH, jelikož využívá dlouhou paměť (často tomu však tak není, viz třeba Calvet a Fisher (2001)).

Dále byla vypočtena dvourozměrná Minzer-Zernowitz regression, pro posouzení toho, zda opční modely ve svých předpovědích zahrnují veškeré informace obsažené v ekonometrických modelech, či nikoliv. Z důvodu úspory času a místa jsme zde porovnali jen 4 nejúspěšnější modely z předchozích analýz a to B-S implikovanou volatilitu, Model-free volatilitu a dále ekonometrické modely ARFIMA a HAR. Sestrojená regresní rovnice vypadala následovně:

$$y_{t+1} = \alpha + \beta \hat{y}_{t+1|t} + \beta_A \hat{y}_{A,t+1|t} + \epsilon_{t+1}$$

Kde $\hat{y}_{t+1|t}$ představuje odhad realizované volatility spočtený jednou ze zkoumaných metod, přičemž $\hat{y}_{A,t+1|t}$ je odhad získaný nějakou alternativní metodou. Pokud by původní metoda ve svých odhadech již zahrnovala veškeré informace, které jsou obsaženy v předpovědích z dané alternativní metody, pak by se koeficient $\beta = 1$ a koeficient β_A by vyšel statisticky nevýznamně odlišné od nuly ($\beta_A \cong 0$). Tabulka 2 zobrazuje koeficienty vypočtené pro různé kombinace výše zmíněných metod.

Tabulka 2 - Porovnání informačního přínosu jednotlivých modelů na in-sample datech

	IV & ARFIMA	MFV & ARFIMA	IV & HAR	MFV & HAR
α coeff	1,73E-06	3,32E-07	3,30E-06	1,83E-06
α t-stats	0,623035	0,119815	1,236595	0,681968
α p-value	0,533476	0,904666	0,216676	0,495497
β coeff	0,758781	0,758983	0,788042	0,787536
β t-stats	9,364567	9,303291	9,387728	9,334174
β p-value	1,17E-19	1,96E-19	9,67E-20	1,51E-19
β_A coeff	0,275871	0,275275	0,226673	0,226985
β_A t-stats	3,213376	3,185871	2,711140	2,702194
β_A p-value	0,001376	0,001511	0,006879	0,007065

Z tabulky vidíme, že i přesto, že se předpovědi opčních modelů jeví jako významnější, vysoce statisticky významné jsou ve všech případech i zkoumané ekonometrické modely. To znamená, že opční modely nezahrnují ve svých předpovědích všechny informace z minulého vývoje dané časové řady a že by tak potenciálně mohlo být výhodné oba přístupy nějakým vhodným způsobem kombinovat. Jedním z možných způsobů je využití té samé regresní rovnice, se kterou jsme pracovali výše. Výsledky několika hybridních modelů (tentokrát i se zapojením modelu ARIMA) zobrazuje tabulka 3.

Tabulka 3 - Výsledky několika hybridních modelů v in-sample a out-of-sample období

IN-SAMPLE	ARIMA & IV	ARIMA & MFV	HAR & IV	HAR & MFV	ARFIMA & IV	ARFIMA & MFV
SSR	1,26E-06	1,25E-06	1,25E-06	1,24E-06	1,24E-06	1,24E-06
R-Squared	0,6140	0,6154	0,6172	0,6182	0,6205	0,6205
OUT-SAMPLE	ARIMA & IV	ARIMA & MFV	HAR & IV	HAR & MFV	ARFIMA & IV	ARFIMA & MFV
SSR	2,31E-07	2,41E-07	2,29E-07	2,36E-07	2,30E-07	2,40E-07
R-Squared	0,4507	0,4281	0,4553	0,4400	0,4535	0,4313

Z tabulky je vidět, že modely kombinující opční a ekonometrické předpovědi volatility dosáhly ve všech případech lepších výsledků, než jednotlivé dílčí modely samostatně.

Dále byly všechny zkoumané modely aplikovány pro konstrukci týdenních (5 dní) a měsíčních (20 dní) předpovědí volatility a to s překrývajícími se i nepřekrývajícími se předpovědními intervaly. Výsledky jsou v tabulkách 4 až 7.

Na první pohled je poněkud zarážející, že s růstem předpovědního intervalu u mnoha modelů narůstala hodnota koeficientu determinace. Podobný efekt lze však pozorovat ve výzkumech zabývajících se předvídaním volatility relativně často (viz třeba Pong a kol. 2003, Blair a kol. 2000, či Hseu a kol. 2007) a souvisí pravděpodobně s její vlastností vykazovat mean-reverzní chování na všech existujících frekvencích (více o této vlastnosti viz Calvet a Fisher 2001).

Z tabulky je dále patrný relativně výrazný rozdíl mezi výsledky modelů ARFIMA, HAR, BS-IV a MFV na straně jedné a ostatními modely na straně druhé, což mluví jednoznačně ve prospěch modernějších přístupů oproti tradičně používaným metodám, jako jsou GARCH či

EWMA. Zajímavý je též výrazný nárůst rozdílu mezi výsledky modelů ARIMA a ARFIMA při měsíčním předpovědním horizontu, což jasně ukazuje na význam dlouhé paměti, který se zvyšuje při konstrukci dlouhodobých předpovědí (to dokumentoval též Corsi 2004).

Celkově lze za nejúspěšnější z modelů na delších horizontech označit model ARFIMA, který ve většině situací předčil jak model HAR, jakožto svého hlavního konkurenta mezi ekonometrickými modely, tak také i oba opční modely, které na denním předpovědním horizontu jasně dominovaly.

Tabulky 8 až 11 dále zobrazují výsledky kombinovaných modelů při týdenních a měsíčních horizontech předpovědí. Je z nich opět vidět, že kombinace ekonometrických a opčních modelů vedla ve většině případů k nezanedbatelnému zlepšení výsledků. Výjimkou jsou zde jen kombinace využívající model ARIMA, které na dlouhých horizontech zcela selhaly.

4 Závěr

V této práci jsme testovali předpovědní schopnosti 8 populárních modelů volatility při předvídání realizované volatility měnového páru EUR/USD. Nejlepších výsledků při denním horizontu předpovědi dosahovaly modely vycházející z tržních cen opcí (Black-Scholesova implikovaná volatilita a Model-free volatilita) následované modely využívajícími informací z intradenních dat (ARIMA-RV, ARFIMA-RV a HAR-RV), z nichž o něco lepších výsledků dosahovaly modely s dlouhou pamětí (ARFIMA a HAR), než model ARIMA s krátkou pamětí. Nejhorších výsledků dosáhly při denním horizontu (jakož i při všech ostatních horizontech) modely vycházející z čtverců denních logaritmických výnosů (EWMA, GARCH a FIGARCH). Při delších předpovědních horizontech (týden a měsíc) pak dominoval zejména model ARFIMA, následovaný oběma opčními modely a dále modelem HAR. S růstem předpovědního horizontu též narůstal rozdíl mezi výsledky modelů s dlouhou pamětí (ARFIMA a HAR) a modelů s krátkou pamětí (ARIMA).

Dále bylo prokázáno, že opční předpovědi neobsahují všechny informace obsažené v modelech časových řad (ARFIMA a HAR) a že by tak mělo být potenciálně výhodné oba přístupy navzájem kombinovat. Bylo proto zkonstruováno 6 jednoduchých hybridních modelů, které následně dosáhly nezanedbatelně lepších výsledků na všech předpovědních horizontech než jednotlivé dílčí modely samostatně.

Tabulka 4 - Výsledky modelů pro týdenní horizont a překrývající se intervaly předpovědí

IN-SAMPLE	EWMA	GARCH	FIGARCH	ARIMA	ARFIMA	HAR	IV	MFV
SSR	2,34E-05	2,40E-05	2,47E-05	1,97E-05	1,94E-05	1,99E-05	1,91E-05	1,86E-05
R-Squared	0,5978	0,5876	0,5746	0,6608	0,6654	0,6574	0,6718	0,6799
OUT-SAMPLE	EWMA	GARCH	FIGARCH	ARIMA	ARFIMA	HAR	IV	MFV
SSR	4,12E-06	4,31E-06	4,72E-06	3,39E-06	2,87E-06	3,15E-06	3,06E-06	2,93E-06
R-Squared	0,3474	0,3172	0,2526	0,4625	0,5445	0,5013	0,5156	0,5361

Tabulka 5 - Výsledky modelů pro týdenní horizont a nepřekrývající se intervaly předpovědí

IN-SAMPLE	EWMA	GARCH	FIGARCH	ARIMA	ARFIMA	HAR	IV	MFV
SSR	4,87E-06	4,98E-06	5,07E-06	4,51E-06	4,36E-06	3,96E-06	4,57E-06	4,46E-06
R-Squared	0,5821	0,5729	0,5651	0,6129	0,6260	0,6603	0,6076	0,6173
OUT-SAMPLE	EWMA	GARCH	FIGARCH	ARIMA	ARFIMA	HAR	IV	MFV
SSR	8,47E-07	8,61E-07	9,33E-07	6,03E-07	5,43E-07	5,81E-07	6,45E-07	6,15E-07
R-Squared	0,3126	0,3008	0,2424	0,5105	0,5594	0,5286	0,4768	0,5004

Tabulka 6 - Výsledky modelů pro měsíční horizont a překrývající se intervaly předpovědí

IN-SAMPLE	EWMA	GARCH	FIGARCH	ARIMA	ARFIMA	HAR	IV	MFV
SSR	3,59E-04	3,42E-04	3,62E-04	2,73E-04	3,04E-04	3,56E-04	3,05E-04	3,10E-04
R-Squared	0,5335	0,5547	0,5289	0,6455	0,6046	0,5373	0,6038	0,5964
OUT-SAMPLE	EWMA	GARCH	FIGARCH	ARIMA	ARFIMA	HAR	IV	MFV
SSR	4,78E-05	5,24E-05	6,27E-05	6,26E-05	3,50E-05	3,68E-05	3,59E-05	3,48E-05
R-Squared	0,3423	0,2795	0,1375	0,1387	0,5189	0,4938	0,5059	0,5220

Tabulka 7 - Výsledky modelů pro měsíční horizont a nepřekrývající se intervaly předpovědí

IN-SAMPLE	EWMA	GARCH	FIGARCH	ARIMA	ARFIMA	HAR	IV	MFV
SSR	1,98E-05	1,85E-05	1,96E-05	1,52E-05	1,70E-05	1,75E-05	1,66E-05	1,68E-05
R-Squared	0,4886	0,5239	0,4939	0,6070	0,5607	0,5476	0,5709	0,5663
OUT-SAMPLE	EWMA	GARCH	FIGARCH	ARIMA	ARFIMA	HAR	IV	MFV
SSR	2,39E-06	2,63E-06	3,13E-06	2,38E-06	1,44E-06	1,99E-06	1,51E-06	1,53E-06
R-Squared	0,3353	0,2667	0,1289	0,3375	0,6002	0,4454	0,5786	0,5748

Tabulka 8 - Výsledky kombinovaných modelů (týdenní horizont a překrývající se intervaly)

IN-SAMPLE	ARIMA & IV	ARIMA & MFV	HAR & IV	HAR & MFV	ARFIMA & IV	ARFIMA & MFV
SSR	1,75E-05	1,73E-05	1,71E-05	1,69E-05	1,70E-05	1,69E-05
R-Squared	0,6989	0,7026	0,7061	0,7090	0,7078	0,7093
OUT-SAMPLE	ARIMA & IV	ARIMA & MFV	HAR & IV	HAR & MFV	ARFIMA & IV	ARFIMA & MFV
SSR	2,57E-06	2,60E-06	2,69E-06	2,62E-06	2,54E-06	2,59E-06
R-Squared	0,5919	0,5880	0,5740	0,5848	0,5969	0,5890

Tabulka 9 - Výsledky kombinovaných modelů (týdenní horizont a nepřekrývající se intervaly)

IN-SAMPLE	ARIMA & IV	ARIMA & MFV	HAR & IV	HAR & MFV	ARFIMA & IV	ARFIMA & MFV
SSR	4,31E-06	4,25E-06	4,00E-06	3,96E-06	4,20E-06	4,17E-06
R-Squared	0,6304	0,6353	0,6566	0,6605	0,6396	0,6423
OUT-SAMPLE	ARIMA & IV	ARIMA & MFV	HAR & IV	HAR & MFV	ARFIMA & IV	ARFIMA & MFV
SSR	5,26E-07	5,25E-07	5,54E-07	5,38E-07	5,22E-07	5,27E-07
R-Squared	0,5734	0,5737	0,5506	0,5630	0,5760	0,5719

Tabulka 10 - Výsledky kombinovaných modelů (měsíční horizont a překrývající se intervaly)

IN-SAMPLE	ARIMA & IV	ARIMA & MFV	HAR & IV	HAR & MFV	ARFIMA & IV	ARFIMA & MFV
SSR	2,57E-04	2,56E-04	3,02E-04	3,07E-04	2,90E-04	2,93E-04
R-Squared	0,6659	0,6664	0,6071	0,6002	0,6230	0,6186
OUT-SAMPLE	ARIMA & IV	ARIMA & MFV	HAR & IV	HAR & MFV	ARFIMA & IV	ARFIMA & MFV
SSR	7,21E-05	6,92E-05	3,31E-05	3,27E-05	3,36E-05	3,20E-05
R-Squared	0,0083	0,0484	0,5448	0,5503	0,5384	0,5605

Tabulka 11 - Výsledky kombinovaných modelů (měsíční horizont a nepřekrývající se intervaly)

IN-SAMPLE	ARIMA & IV	ARIMA & MFV	HAR & IV	HAR & MFV	ARFIMA & IV	ARFIMA & MFV
SSR	1,51E-05	1,51E-05	1,64E-05	1,66E-05	1,64E-05	1,64E-05
R-Squared	0,6093	0,6111	0,5768	0,5719	0,5779	0,5758
OUT-SAMPLE	ARIMA & IV	ARIMA & MFV	HAR & IV	HAR & MFV	ARFIMA & IV	ARFIMA & MFV
SSR	2,94E-06	2,82E-06	1,44E-06	1,48E-06	1,38E-06	1,36E-06
R-Squared	0,1802	0,2148	0,5981	0,5887	0,6145	0,6220

Literatura

- [1] **ANDERSEN, T.G., BOLLERSLEV, T., (1998).** Answering the skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts, *International Economic Review*
- [2] **ANDERSEN, BOLLERSLEV, CHRISTOFFERSEN, DIEBOLD, (2005).** Volatility Forecasting, National Bureau of Economic Research
- [3] **ANDERSEN, BONDARENKO, (2007).** Construction and Interpretation of Model-free Implied Volatility, CREATES Research Paper
- [4] **BAILIE, BOLLERSLEV, MIKKELSEN, (1996).** Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*
- [5] **BAKSHI, KAPADIA, (2003).** Delta-hedged gains and the negative market volatility risk premium, *Review of Financial Studies*
- [6] **BLACK F., SHOLES, M., (1973).** The Pricing of Options and Corporate Liabilities, *Journal of Political Economy*
- [7] **BLAIR, POON, TAYLOR, (2000).** Forecasting S&P Volatility: The Incremental Information Content of Implied Volatilities and High Frequency Index Returns
- [8] **BOLLERSLEV, T., (1986).** Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*
- [9] **BOX, G., JENKINS, G., (1970).** Time series analysis: Forecasting and control
- [10] **BRITTEN-JONES, NEUBERGER, (2000).** Option Prices, Implied Price Processes and Stochastic Volatility, *Journal of Finance*, Vol. 55
- [11] **CALVET, FISHER, (2001).** Forecasting Multifractal Volatility, *Journal of Econometrics*
- [12] **CORSI, Fulvio, (2004).** A Simple Long Memory Model of Realized Volatility
- [13] **DAY, LEWIS, (1992).** Stock Market Volatility and the Information Content of Stock Index Options, *Journal of Econometrics*
- [14] **DING, GRANGER, ENGLE, (1993).** A long memory property of stock market returns and a new model, *Journal of Empirical Finance*
- [15] **ENGLE, Robert F., (1982).** Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica*
- [16] **FIČURA, Milan, (2013).** Metody předvídání volatility, Diplomová práce, Vysoká škola ekonomická v Praze
- [17] **FIGLEWSKI, STEPHEN, (2004).** Forecasting Volatility
- [18] **GRANGER, JOYEUX, (1980).** An introduction to long-memory time series models and fractional differencing, *Journal of Time Series Analysis*

- [19] **HSEU, CHENG, CHUNG, (2007)**. The Forecasting Performance of Model Free Implied Volatility: Evidence from an Emerging Market
- [20] **ISHIDA, WATANABE, (2008)**. Modeling and Forecasting the Volatility of the Nikkei 225 Realized Volatility Using the ARFIMA-GARCH Model, Institute of Economic Research
- [21] **JIANG, TIAN, (2005)**. The model-free implied volatility and its information content, Review of Financial Studies
- [22] **MINCER, ZARNOWITZ, (1969)**. The Evaluation of Economic Forecasts, Economic Forecasts and Expectations, New York: National Bureau of Economic Research
- [23] **MUZZIOLI, Silvia, (2008)**. Option based forecasts of volatility: An empirical study in the DAX index options market, CeFin working paper
- [24] **POON, GRANGER, (2003)**. Forecasting Volatility in Financial Markets : A Review, Journal of Economic Literature
- [25] **PONG, SHACKLETON, TAYLOR, XU, (2003)**. Forecasting Currency Volatility: A Comparison of Implied Volatilities and AR(FI)MA Models, Journal of Banking and Finance
- [26] **WANG J.W, YOUROUGOU P., WANG Y.D., (2009)**. Which Implied Volatility Provides The Best Measure of Future Volatility?, Springer Science

Abstract: In this paper we are testing the forecasting abilities and the information content of 8 popular models of volatility forecasting (EWMA, GARCH, FIGARCH, ARIMA-RV, ARFIMA-RV, HAR-RV, Black-Scholes implied volatility and model-free volatility). The models are applied to 5 years of daily data about the evolution of the EUR/USD exchange rate in order to forecast the realized volatility in 1 day, 5 day and 20 day horizon. The best forecasting results were achieved by the models based on option prices (Black-Scholes implied volatility and Model-free volatility), followed by the realized volatility models incorporating long memory (ARFIMA-RV and HAR-RV). From these models ARFIMA-RV has dominated the others especially on the longer horizons, where it surpassed even the option models. The tests of the information content showed that the option models do not subsume all of the information contained in the econometric models (ARFIMA and HAR). Because of that we created several hybrid models (using option as well as time series forecasts) that achieved on average better results than any of the basic models on their own.

AMS/JEL classification: C22, C53, G14

Keywords: Volatility forecasting, realized volatility, implied volatility, model-free volatility