

Rövid távú olajár-előrejelzések teljesítményének stabilitása*

Uliha Gábor,

az MVM Partner Energiakereskedelmi Zrt. piacelemzési szakértője, a Budapesti Corvinus Egyetem PhD-hallgatója

E-mail: gabor.uliha@gmail.com

Az olajárak alakulását, illetve az ezt meghatározó adatgeneráló folyamatot régóta kutatják közgazdászok és piaci elemzők. A 2014 nyarán kezdődött trendszerű csökkenés újra a figyelem középpontjába helyezte az ármozgásokat, valamint azok rövid távú előrejelezhetőségét. Jelen tanulmány célja, hogy a West Texas Intermediate olajfajta példáján keresztül betekintést nyújtson a használatban lévő módszerekbe, tesztelje az általuk adott valós idejű előrejelzések pontosságát és relatív teljesítményének időbeli változását. A cikk az alkalmazott modellek egy meglehetősen széles körét dolgozza fel, a szokásos ökonometriai, idősorelemzési eljárások mellett bemutatja a gépi tanulás területéhez tartozó, a közgazdasági alkalmazásokban csak az utóbbi években teret nyerő, és az olajár egy, két- és háromnapos előrejelzésére potenciálisan alkalmazható módszereket is. Az eredmények tükrében megállapítható, hogy a határidős termékek, illetve egyes olajszármazékok árának felhasználása a vizsgált esetek mindegyikében javítja az előrejelzés pontosságát. Ugyanakkor a modellek relatív teljesítménye az előrejelzési horizonttól, a vizsgált időszaktól és az alkalmazott hibamutatótól egyaránt függ, így a legjobb modellek halmaza meglehetősen instabil, nincs univerzálisan használható eljárás az olajárak rövid távú előrejelzésére.

TÁRGYSZÓ:

Olajár-előrejelzés.

Ökonometria.

Gépi tanulás.

* A szerző köszönetet mond *Vincze János*nak (Budapesti Corvinus Egyetem, Magyar Tudományos Akadémia Közgazdaság- és Regionális Tudományi Kutatóközpont Közgazdaság-tudományi Intézet), *Vonnák Balázs*nak (Magyar Nemzeti Bank), *Badics Milán*nak (Budapesti Corvinus Egyetem) és *Kucsera Henrik*nek (Magyar Nemzeti Bank) a tanulmányhoz kapcsolódó értékes megjegyzéseikért.

A kőolaj tőzsdei jegyzésében 2014-ben bekövetkezett trendszerű elmozdulások ismét a figyelem középpontjába helyezték a fosszilis energiahordozó piacát. Ennek megfelelően a médiában is egyre gyakrabban találkozunk a különböző bankok, elemzőházak, kereskedők által adott előrejelzésekkel, melyek a rövid (pár napos) és hosszú távú (akár több éves) változások „jóslására” is vállalkoznak. Az előrejelzések sokszínűsége, illetve a pozitív/negatív kockázatok hosszú listája azonban jelzi, hogy a kivetítés nem triviális probléma, és nagy biztonsággal senki sem képes megmondani, mi várható a jövőben. Nem egyértelmű ugyanis, hogy a komplex adatgeneráló-folyamat mellett egyáltalán megvalósítható-e az árak „kellően pontos” előrejelzése. Az eredmények ellentmondásosak, és sok függ attól, mennyire részletes idősorokkal dolgozunk. *Hamilton* [2009] például kétségbe vonja a reál olajárak előrejelezhetőségét még havi vagy negyedéves szinten is. E hipotézist mások ugyan cáfolták (lásd *Baumeister–Kilian* [2013]), a rövid távú, napi bontású előrejelzéseknél a probléma ettől még fennállhat. Bár számos tanulmány született a napi spot árak előrejelzésének témakörében, ezek gyakran nem valós idejű projekciókat takarnak (lásd többek között *Haidar–Kulkarni* [2009], *Shabri–Samsudin* [2014], valamint *Yu–Wang–Lai* [2008]), illetve jellemzően kevés modell összehasonlítását végzik, így nehéz megállapítani, hogy a javasolt módszerek egymáshoz képest hogyan teljesítenek, valamint ez a teljesítménybeli különbség más időszakok során is fennáll-e. Ennek egyik oka alighanem a nagy mintákkal járó jelentős számítási kapacitásigény, ami szűkíti a kutatók lehetőségeit. Ezzel szemben havi vagy negyedéves adatok feldolgozása mellett már gyakoribb, hogy egyetlen tanulmányon belül akár több tucat modell eredményeit is összevetik (*Alquist–Kilian* [2010], *Alquist–Kilian–Vigfusson* [2011]). Jelen tanulmányban néhány napos előrejelzések esetén is elvégzem ezt a vizsgálatot, és az alkalmazott modellek meglehetősen széles körét tesztelem. Tudomásom szerint a cikkben bemutatott valós idejű olajár-előrejelzéseket adó eljárásokat együtt még nem vizsgálták. A tesztelés során a következő kérdésekre keresem a választ.

- Javít-e az előrejelzés pontosságán a határidős termékek vagy a különböző olajszármazékok árainak bevonása?
- Mely modellek teljesítenek a legjobban a tényadatokra való illeszkedés és az árváltozás irányának előrejelzése szempontjából?
- Az előrejelzési horizont tekintetében mennyire stabilak az eredmények, azaz vannak-e olyan modellek, amelyek minden időhorizonton (egy-, két- és háromnapos előrejelzésnél is) a legjobban teljesítő eljárások közé tartoznak?

– Mennyire stabil a halmaz az előrejelzési időszak tekintetében, az az időben állandó marad-e a legjobb modellek halmaza, vagy folyamatosan változik?

– Egyáltalán sikerül-e egy egyszerű benchmarknál számottevően pontosabb modellt építeni?

A kérdésfeltevésből következik, hogy a kutatás során az induktív megközelítést alkalmazom, nincs előre rögzített, tesztelhető elméletem a helyes válaszokkal kapcsolatban, a végső konklúziók csupán a modellek és időszakok azon szűk halmazára vonatkoznak, melyeket az alkalmazás során áttekintek. Ezekből az egyedi megfigyelésekből általános érvényű összefüggések nem, csak sejtések szűrhetők le.

A tanulmány felépítése a következő. Előbb áttekintem a kapcsolódó szakirodalmat, ismertetem a dolgozatban használt eljárásokat és adatokat, majd megvizsgálom az egyes modellek teljesítőképességét. Végül összefoglalom a főbb tanulságokat, a potenciális továbblépési lehetőségeket.

1. A szakirodalom áttekintése

Bashiri–Manso [2013] összefoglaló tanulmánya kvantitatív és kvalitatív kategóriákba sorolja az ismert eljárásokat, előbbit pedig tovább bontja ökonometriai, és általában „nem standard” eljárásoknak nevezett csoportra, ami valójában a gépi tanulós modelleket takarja. A kvantitatív modelleken belül jellemzően az ökonometriai eljárásokat használják, de az utóbbi időben gyorsuló ütemben terjednek a gépi tanuláson (machine learning) alapuló megközelítések, azon belül is az SVM,¹ illetve a neurális háló alkalmazása. A következőkben *Bashiri–Manso* [2013] felosztásának megfelelően mutatom be az egyes eljárásokat és azok néhány alkalmazását.

1.1. Ökonometriai modellek

Az ökonometriai modellek csoportján belül két megközelítés lehetséges. Az első kizárólag az olajár múltbeli értékeiből jelez előre, más változókat egyáltalán nem használ. Ezek az egyváltozós idősormodellek mintázatot, szabályszerűséget keresnek a folyamat dinamikájában, a fundamentumokat figyelmen kívül hagyják. A másik

¹ SVM (support vector machine): a tanulmányban gyakran elő fog fordulni, hogy az eredeti angol elnevezést használom. Ezt sajnos elkerülhetlenné teszi, hogy egyes kifejezéseknek nincs magyar megfelelője, illetve ha van, akkor sem terjedt el (közgazdasági) szakmai körökben. Az SVM például tartóvektorgép lenne, de ezzel a kifejezéssel ritkán találkozunk.

megközelítés ezzel szemben a közgazdasági elméletet hívja segítségül, és a releváns árfolyásoló tényezők felhasználásával próbál javítani az előrejelzés pontosságán.

1.1.1. Egyváltozós idősorelemzés

Az egyváltozós idősorelemzési modellekben egyetlen magyarázóváltozó, a függő változó késleltetett értékei szerepelnek. Az ARIMA-modell² (*Wang–Yu–Lai* [2004], *Xie et al.* [2006]) népszerűségét nemcsak az adja, hogy könnyen érthető és implementálható eljárásnak számít, hanem az is, hogy a belőle nyert előrejelzés számos esetben viszonyítási pontként szolgál a szofisztikáltabb eljárások számára (lásd például *Yu–Wang–Lai* [2008]). Gyakran találkozni olyan cikkel, ahol a szerzők valamilyen ARIMA projekcióhoz és/vagy a véletlen bolyongás feltevéséhez³ viszonyítják saját eredményeiket. *Alquist–Kilian* [2010] például havi adatokkal dolgozva megmutatták, hogy a WTI⁴ olajtípusnál a futures árak rosszabb előrejelzői a jövőbeli spot áraknak, mint a véletlen bolyongás feltevése melletti kivetítés. Ennek lehetséges okaként az árak magas volatilitásból eredő kockázatot jelölik meg, ami beépül az árakba.⁵

1.1.2. Pénzügyi modellek

A pénzügyi modellek valójában továbbra is idősorelemzési megközelítéseket takarnak, az előbb bemutatott eljárásokhoz képest egy fő különbséggel, mégpedig a futures árak mint magyarázóváltozók bevonásával. Összességében elmondható, hogy a futures árak nem tekinthetők a jövőbeli spot árak torzítatlan előrejelzéseinek (*Alquist–Kilian* [2010], *Yousefi–Weinreich–Reinarz* [2005]), sőt az sem teljesen egyértelmű, hogy használatuk javítja-e a modellek előrejelző képességét, *Alquist–Kilian–Vigfusson* [2011] legalábbis nem találtak erre utaló jelet.⁶

1.1.3. Strukturális modellek

A strukturális modellek lényege, hogy a futures árak mellé újabb magyarázóváltozókat vonnak be. Ezek köre felhasználási területtől és az idősorok részletezettségétől (napi, havi, negyedéves) függ. *Tang–Hammoudeh* [2002] például egy „target

² ARIMA (autoregressive integrated moving average): autoregresszív integrált mozgóátlag-folyamat.

³ Amikor a holnapi, holnaputáni stb. ár előrejelzése a legutolsó ismert értékkel egyenlő (eltolás nélküli véletlen bolyongás).

⁴ WTI (West Texas Intermediate): nyersolajfajta elnevezése, amit az olajkereskedelemben referenciaként használnak.

⁵ A volatilitás modellezésére ARCH/GARCH típusú eljárásokat használnak (*Kang–Kang–Yoon* [2009]), mivel sem ez, sem a szintén gyakran előkerülő hibakorrekciós eljárás (*Lanza–Manera–Giovannini* [2005]) nem kapcsolódik szervesen a dolgozat témájához, bővebb ismertetésüktől eltekintek.

⁶ Igaz, ők sem a napi adatokra fókuszáltak. Ellenben *Haidar–Kulkarni* [2009] úgy találták, hogy egy neurális háló előrejelző képességét már javítják a futures adatok (a WTI napi árára vonatkozóan). Erre a cikkre még visszatérek a gépi tanulásról szóló fejezetben.

zone” modellt használnak (*Krugman* [1991] módszerére építve), amelynek lényege, hogy bizonyos alsó és felső árküszöbök mellett az OPEC⁷ beavatkozik. Ennek oka, hogy túlságosan alacsony árszintek mellett a tagországok költségvetése nyomás alá kerül, míg magas árak esetén a szervezet befolyásán kívüli termelők a felhozatal növelésére, a vásárlók pedig a fogyasztás optimalizálására vannak ösztönözve. Konklúziójuk szerint a küszöbszintek közelében ez a modell valóban segíti az előrejelzést, de a köztes, illetve külső szakaszokon ez már nem mondható el.⁸ Más szerzők (*Merino-Ortiz* [2005]; *Ye-Zyren-Shore* [2005], [2006]) a készleteket használták magyarázóváltozóként. E tanulmányok lényeges következtetése (összhangban *Tang-Hammoudeh* [2002] modelljével), hogy az áralakulás folyamata nem lineáris, egész más reakció figyelhető meg a készletek szintjétől függően. Ezen eredmények ismeretében az elemzés során én is kitérek különböző nem lineáris modellek használatára.

1.2. Gépi tanulás

A számítási kapacitások rohamos növekedése következtében mára lehetővé vált a gépi tanulásnak (machine learning) nevezett modellek széleskörű használata. Ez a terület valahol a statisztika és a számítástudomány határmezsgyéjén található. Lényege, hogy nem egy konkrét, előre megadott függvényforma paramétereit becsli, hanem a modellező által specifikált tanuló algoritmus révén alakítja ki a végső függvényformát és kalibrálja annak paramétereit.⁹ A gépi tanulás előbb a műszaki és természettudományos területeken jelent meg, mára azonban a társadalomtudományokban is elterjedt. Előnye, hogy a standard ökonometriában használatos, többnyire lineáris megközelítéseknel jóval komplexebb összefüggések leképezésére alkalmas. Ebben a formában azt mondhatjuk, hogy a gépi tanulás nem korlátozza a vizsgálati spektrumot egyetlen formulára, hanem számtalan lehetőség közül választja ki a legjobbat: nem a felhasználó, hanem az adatok döntenek. Ez az előny azonban a legnagyobb hátrány is egyben. Az ökonometriai eljárásokkal ellentétben a gépi tanulás ritkán alkalmas az ok-okozati kapcsolatok kimutatására (*Statnikov-Hardin-Aliferis* [2006]), és gyakori probléma a túlillesztés, amikor az adatokban mutatkozó zajra, egyedi mintákra tanul rá az algoritmus. Ennek következtében a gépi tanulás alkalma-

⁷ OPEC (Organization of the Petroleum Exporting Countries): Kőolaj-exportáló Országok Szervezete.

⁸ Megjegyzendő, hogy az 1988–1999-es időszakot vizsgálták, amikor hivatalosan nem volt ilyen jellegű ár-zóna az OPEC-nél. A szerzők ugyanakkor megállapították, hogy a kartell viselkedése, konkrét döntései nem mondanak ellent az általuk alkalmazott elméletnek, azaz valószínűleg voltak olyan informális küszöbszintek, amelyek elérésekor emelkedett a közvetlen beavatkozás valószínűsége.

⁹ E megfogalmazás kissé leegyszerűsítő, mert például neurális hálók esetén a függvényformát a modellező specifikálja, de az eljárás alapja épp az, hogy olyan struktúrát használ, amivel a függvények meglehetősen széles köre közelíthető. Így valójában az elemző/programozó nem a változók közötti konkrét összefüggést adja meg parametrikus formában, csupán (közvetve) a szóba jöhető függvények halmazát, és a tanulás módját specifikálja.

zásának legfontosabb – vagy legalábbis a kapcsolódó szakirodalomban mindig kiemelt helyen szereplő – eleme a túlillesztés elkerülése. Bár nincs univerzális, minden modellre és adathalmazra érvényes megoldás, számos lehetőség közül lehet választani. A téma feldolgozása meghaladná e tanulmány kereteit, ezért a későbbiekben csak azokat a módszereket fogom áttekinteni, melyeket az előrejelzés során is használok. Az olvasó a fontosabb eljárásokról átfogó leírást talál Berk [2008] és Vapnik [1995] (SVM), valamint Kriesel [2007] munkáiban.

A nyersolaj piaci árának előrejelzésére is mind gyakrabban használják e módszereket. Fernandez [2010] SVM és NN¹⁰ segítségével készített előrejelzést napi adaton. Következtetése szerint rövid távon ezek nem teljesítenek jobban egy egyszerű ARIMA-modellnél, de hosszabb távon „legyőzik” azt. Alapvetően gyakori jelenség az irodalomban, hogy több modell összevetése során nem találunk olyan eljárást, amely minden időhorizonton jobbnak bizonyul (lásd Baumeister–Kilian [2013]). Szintén érdemes figyelembe venni Haidar–Kulkarni [2009] eredményeit, akik a WTI napi spot árait modellezték (előrecsatolt) neurális hálók segítségével. Megállapításuk szerint az adatgeneráló-folyamatban levő nem lineáris függvényforma¹¹ miatt a szokásos lineáris megközelítések alkalmatlanok az előrejelzésre, emellett az adatok megfelelő transzformációja, zajtól való előzetes szűrése is fontos lépés, ugyanis ezzel lehet elkerülni a túlillesztés (a zajra való rátanulás) csapdáját.¹² További következtetésük, hogy a piaci információk gyorsan beépülnek a futures árakba, így azok változása felhasználható az egy-, két- és háromnapos spot ár előrejelzésekben.

A gépi tanuláson alapuló módszerrel alkalmanként rendkívül magas, akár 80 százalékos feletti arányban is sikerült eltalálni a néhány napos változások irányát (Haidar–Kulkarni [2009], Yu–Wang–Lai [2008]), ám ezek az eredmények erős fenntartással kezelendők. Yu–Wang–Lai [2008] például egy többlépcsős eljárással modellezi a Brent és WTI olajárak napi bontású idősorait. Az első lépcsőben EMD-módszerrel¹³ bontják fel az ár idősorokat különböző komponensekre, majd ezekre egyenként végznek előrejelzést előrecsatolt neurális hálók (feed-forward neural network) segítségével, végül az egyes komponensekre kapott előrejelzésekből egy adaptív lineáris neurális hálóval (adaptive linear neural network) becsülik meg az ár idősorának alakulását. A megközelítés hátránya, hogy az eredmények nem adják vissza, milyen lenne a modell teljesítménye valós időben. Egyrészt az EMD-eljárás végponti problémákkal küzd (lásd Ma–Liu–Zhou [2013]), így a szerzők által (1986. január 1-jétől 2006. szeptember 30-ig) elvégzett EMD-felbontás belső pontjai nem feltétlenül es-

¹⁰ NN (neural network): neurális háló.

¹¹ A nem linearitás formális tesztelésére azonban nem tér ki a tanulmány, azt Haidar–Wolff [2011] igazolja.

¹² Shabri–Samsudin [2014] is azt találják, hogy megfelelő szűréssel javítható az előrejelzés pontossága. Ők ugyanakkor nem kifejezetten a zaj szűrésére fókuszálnak, hanem wavelet-transzformáció segítségével bontják elemekre az olajár historikus idősorát, majd egy szelekciós eljárással eltávolítják az idősor azon részeit, melyek együttlomozása az eredeti idősorral nem elég erős. Az így nyert idősorokat egy neurális háló input adataként használják.

¹³ EMD (empirical mode decomposition): empirikus dekompozíció.

nek egybe azzal az eredménnyel, mint amit valós idejű feldolgozás esetén kaptak volna. Másrészt az algoritmus maga dönti el, hány komponensre osztja az idősort, tehát akár az is lehetséges, hogy a WTI-re kapott 11 komponens egy korábbi időpontban nem is állt volna elő (például 10 vagy 12 lett volna belőle).

Ahogy az említett példa is mutatja, komoly nehézségekbe ütközik az egyes modellek – kizárólag az irodalomra hagyatkozó – rangsorolása, mivel különböző időszakokra, eltérő olajfajtákra és alkalmanként össze nem vethető módszerekkel állították elő őket. *Yu–Wang–Lai* [2008] cikkének eredményeit például nem lehet összehasonlítani egy olyan modellével, ahol minden pontban kizárólag az akkor már rendelkezésre álló adatokat használták fel. További probléma, hogy a modelleket általában egy egyszerű benchmarkhoz (véletlen bolyongás, ARIMA) viszonyítják, így nem derül ki, hogy a mások által elért eredményekhez képest miként teljesít az új módszer. A később ismertetendő empirikus eredmények tükrében úgy tűnik, hogy a gyakorlatban egy-, két és háromnapos távlatban az 55–60 százalékos találati arány már igencsak kedvezőnek tekinthető. Feltéve persze, hogy nem egyszeri kiugrásról, hanem tartósan ebbe a sávba eső teljesítményről van szó.

2. Módszertan

A tanulmányban alkalmazott módszerek három csoportba sorolhatók:

- benchmarkként szolgáló heurisztikus;
- ökonometriai;
- gépi tanuláson alapuló modellek.

1. táblázat

Az előrejelzés során használt modellek

Módszertan		Előrejelzés	Frissítés	Hivatkozás
neve	típusa			
Véletlen bolyongás	heurisztikus	ár, irány	naponta	–
ARIMA	ökonometriai	ár, irány	évente/naponta	<i>Hamilton</i> [1994]
ARX ¹⁴	ökonometriai	ár, irány	naponta	<i>Wang–Jain</i> [2003]
LPM ¹⁵ /logit	ökonometriai	irány	naponta	<i>Wooldridge</i> [2012]
Neurális háló	gépi tanulás	ár, irány	évente	<i>Kriesel</i> [2007]
SVM	gépi tanulás	irány	évente	<i>Cristianini–Shawe-Taylor</i> [2000]

¹⁴ ARX (autoregressive exogenous model): exogén változókkal bővített autoregresszív modell.

¹⁵ LPM (linear probability model): lineáris valószínűségi modell.

A konkrét eljárásokat, azok típusát, az előrejelzés tárgyát (ár/irány), a paraméterek frissítésének gyakoriságát és a modellek részletes bemutatásának irodalmát az 1. táblázat foglalja össze. A módszertani leírást a következő alfejezetek tartalmazzák.

2.1. Heurisztikus modellek – Véletlen bolyongás

A modell feltevése szerint a spot olajárak alakulása rövid távon eltolás nélküli véletlen bolyongást követ, így a holnapi, holnaputáni stb. árak legjobb előrejelzése a mai érték. Ennek oka, hogy amennyiben az ár eltolás nélküli véletlen bolyongást követ, akkor a folyamat várható értéke bármely jövőbeli időpontban a legutolsó ismert tényadattal egyezik. Ez a megközelítés persze (egy-két kivételes alkalomtól eltekintve) soha nem fogja eltalálni a változás irányát. A modell formálisan:

$$p_t = p_{t-1} + \varepsilon_t, \quad /1a/$$

$$\hat{p}_{t-1+m} = p_{t-1}, \quad /1b/$$

ahol p_t az t . időszak spot ára, \hat{p}_t az erre vonatkozó előrejelzés, ε_t a hibtag, m pedig az előrejelzési időszak hossza (jelen esetben $m = 1, 2, 3$).

Egy másik lehetőség az eltolásos véletlen bolyongás használata:

$$p_t = \infty + p_{t-1} + \varepsilon_t, \quad /2a/$$

$$\hat{p}_{t-1+m} = m \cdot \hat{\infty} + p_{t-1}, \quad /2b/$$

ahol ∞ az eltolás mértéke (drift), $\hat{\infty}$ pedig ennek becsült értéke.

2.2. Ökonometriai modellek

Az elemzés során olyan standard ökonometriai eljárásokkal dolgozom, amelyek felhasználási területe meglehetősen széles, így részletekbe menő ismertetésüktől ezúttal eltekintek, csak az eredmények értelmezéséhez feltétlenül szükséges pontok bemutatására szorítkozom.

2.2.1. ARIMA

Az ARIMA(p, d, q)-modellnél p jelöli az autoregresszív tagok számát, d az idősor integráltságának fokát (hányszori differenciaképzés után lesz stacionárius), q pedig a mozgó átlagolású tagok számát:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \dots + \beta_p y_{t-p} + \varepsilon_t + \gamma_1 \varepsilon_{t-1} + \dots + \gamma_q \varepsilon_{t-q}, \quad /3/$$

ahol y_t a függő változó értéke a t . periódusban, ε_t pedig a hibatag. Mivel az eljárás csak stacionárius idősorokra érvényes, y_t nem a spot árat, hanem annak transzformált alakját, a logaritmikusan differenciált jelöli (ez a módosítás az összes, később ismerttetendő előrejelzésnél is megtörténik):

$$y_t = \ln(p_t) - \ln(p_{t-1}). \quad /4/$$

Az egyenlet paramétereit likelihood eljárással becslöm a MATLAB R2013a programcsomag Econometrics Toolboxának segítségével. A kiértékeléskor természetesen nem a transzformált változó dinamikus előrejelzéseit vizsgálom, hanem az abból nyert, az eredeti változóra (a nominális spot árra) vonatkozó projekciókat. Az ARIMA-modellezés során két megközelítést alkalmazok:

- minden év elején az addig rendelkezésre álló adatokon megbecsülöm az ARIMA($p, 1, q$)-modelleket ($p, q = 1, \dots, 10$) majd a Schwarz információs kritérium¹⁶ alapján kiválasztom a becslési mintán legjobban teljesítő beállítást, és a következő 12 hónapban végig ezzel a modellel végzem az egy-, két- és háromnapos dinamikus előrejelzést;¹⁷
- minden nap elején az akkor rendelkezésre álló adatokon megbecsülöm az ARIMA($p, 1, 0$)-modellt ($p = 1, \dots, 20$), majd ebből készítek – szintén 1–3 napos, dinamikus – kivetítést.¹⁸

2.2.2. ARX

Az exogén változókkal bővített autoregresszív modell az ARIMA($p, d, 0$)-modell kiterjesztésének tekinthető. Az előző pontban ismerttetthez képest annyi változás figyelhető meg, hogy újabb magyarázóváltozókkal bővül az egyenlet:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \dots + \beta_p y_{t-p} + \gamma_1 x_{t-m} + \dots + \gamma_s x_{t+1-m-s} + \varepsilon_t, \quad /5a/$$

¹⁶ $SIC = -2 \cdot \ln(L) + k \cdot \ln(n)$, ahol L a loglikelihood függvény értéke, k a becslendő paraméterek, n pedig a megfigyelések száma.

¹⁷ Az év elején a rendelkezésre álló adatok alapján kiválasztom az alkalmazandó modell formáját (ez egyébként minden alkalommal az ARIMA(0,1,1)-alakot jelenti), megbecsülöm a paramétereiket, majd az újonnan érkező adatokat ebbe a modellbe helyettesítve készítek előrejelzéseket. A következő év elején megismétlem az eljárást, és frissítem a paramétereiket.

¹⁸ Az ARIMA-modellekhez tartozó számítási igény meglehetősen nagy, míg az AR(p)-modellek közvetlenül becsülhetők a legkisebb négyzetek módszerével, így hamar lefutnak, és utólag is könnyen előállítható a valós idejű, napi szinten frissülő előrejelzések időszora. A mozgó átlag tagokat ezért a későbbiek során is kihagyom az elemzésből.

ahol x_i a bevont magyarázóváltozó i . időszaki értéke, γ_j pedig a megfelelő késleltetéshez tartozó koefficiens. Látható, hogy minden esetben az x változónak csak az előrejelzés időpontjában már ismert (legalább m nappal korábbi) értékei kerülnek az egyenletbe, míg a spot olajár esetén már az egy nappal korábbi érték is (dinamikus előrejelzés). A modell paraméterei naponta frissülnek.

Ahogy a felírásból is kitűnik, az exogén magyarázóváltozó maximális késleltetése nem feltétlenül esik egybe a függő változó hasonló paraméterével. Emellett az alkalmazások során előfordul, hogy nemcsak egy, hanem több magyarázóváltozó szerepel a modellben, ekkor az /5a/ egyenlet értelemszerűen újabb x változókkal bővül (az egyszerűség kedvéért itt is felteszem, hogy a késleltetések száma s , azaz minden exogén változó esetén ugyanannyi). Az egyes modelleknél ez a következő változókat takarja.

1. A futures árak és/vagy olajszármazékok (benzin, fűtőolaj) dlog idősorainak késleltetett értékei:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \dots + \beta_p y_{t-p} + \gamma_1 \text{dlog}(p_{t-m}^x) + \dots + \gamma_p \text{dlog}(p_{t+1-m-p}^x) + \varepsilon_t, \quad /5b/$$

ahol p_t^x a bevont futures termék vagy olajszármazék tőzsdei jegyzése az alsó index szerinti időpontban, $p = s$ és $p = 1, \dots, 20$.

A futures árak késleltetett értékeinek szerepeltetését *Haidar–Kulkarni* [2009] eredményei indokolják. Ez a megoldás azért javíthat az előrejelzésen, mert egyes kutatások szerint a futures piacok nagyobb likviditása miatt az információk itt hamarabb beépülnek az árakba, mint a spot piacon (*Silvapulle–Mossa* [1999]). Valószínűsíthető ugyanakkor, hogy a hatás erősebben jelentkezik a napon belüli, órás vagy még részletesebb bontású kereskedési adatokon (*Brooks et al.* [2001]), így csak a konkrét eredmények ismeretében lehet eldönteni, hogy napi bontásnál is érdemes-e bevonnai az új változókat.

Ehhez hasonlóan megvizsgálom, hogy a benzin és fűtőolaj árak változásának felhasználásával javítható-e a modell előrejelző képessége. Az olajszármazékok ára ekkor a végtermék iránti kereslet proxyjaként funkcionál, ami később a nyersolaj piacon is megjelenik. Problémát jelent azonban, hogy az árváltozás gyakran a termelési költség növekedéséből ered, ilyenkor pedig az ok-okozati szerepek felcserélődnek (a nyersolaj drágulása a benzin áremelkedését eredményezi), így szintén csak az eredmények ismeretében lehet dönteni e változók szükségességéről.

Bizonyos esetekben nem egyetlen plusz időszorral bővül az egyenlet. Hat változó kilenc kombinációját teszteltem, azaz y késleltetett értékei mellett a következő tagok szerepelnek az egyenlet jobb oldalán (a változók pontos leírását később ismertetem):

– futures árak: egy-, két-, három, négyhónapos futures árak szerepeltetése együtt és külön-külön (5 eset);

- olajszármazékok: benzin és fűtőolaj spot árak szerepeltetése együtt és külön-külön (3 eset);
- teljes modell: a négy olaj futures és a két olajszármazék spot ár együttes szerepeltetése (1 eset).

Az alkalmazott módszertantól függetlenül mindig ezt a kilenc esetet vizsgálom, azaz a spot olajár késleltetett értékei mellé a korábbi idősorok megfelelő (dlog vagy spread) transzformációi kerülnek az elemzésbe, függetlenül attól, hogy milyen modellről van szó. A későbbiekben ezért eltekintek a változók körének részletes felsorolásától.

2. A spot olajár logaritmusának és az olaj futures vagy benzin/fűtőolaj árak logaritmusainak különbsége (spread), ahol $p = 1, \dots, 20$, x -ek esetén pedig kizárólag az utolsó ismert spread érték szerepel az egyenletben (tehát m napos előrejelzésnél $\gamma_{t-m} \neq 0$, minden más – spread – koefficiens nulla):

$$y_t = \beta_0 + \beta_1 y_{t-1} + \dots + \beta_p y_{t-p} + \gamma \left[\ln(p_{t-m}^x) - \ln(p_{t-m}) \right] + \varepsilon_t. \quad /5c/$$

Baumeister et al. [2013] havi adatokon megmutatták, hogy hosszú távon az olaj reál árának előrejelzését javítja a benzin és fűtőolaj spread értékek használata. Javulás alatt azt értették, hogy a változatlan áras előrejelzéshez képest akár 20 százalékkal is csökkenteni lehetett az előrejelzés átlagos négyzetes hibáját. Kérdés, hogy ez 1–3 napos viszonylatban is teljesül-e, amit szintén megvizsgálók az elemzési szakaszban. Ennek mintájára a futures árakra is elvégzem az elemzést, ezek ugyanis a jövőbeli spot árakra vonatkozó piaci várakozást tükrözik.¹⁹

2.2.3. Bináris függő változós modellek

Amennyiben a modellező célja nem az árak hanem az árváltozás irányának minél pontosabb előrejelzése, akkor bináris függő változós modellek segítségével lehetőség van ennek közvetlen becslésére. A bemutatásra kerülő eljárásokat kizárólag egynapos projekciókhoz használom, mivel a később ismertetendő profitfüggvény maximalizálásához nincs szükség ennél hosszabb kitekintésre. A korábbiakhoz hasonlóan ezeknél a modelleknél is napi szinten frissítem a paramétereket.

Lineáris valószínűségi modell (LPM). Az LPM-regresszió egy bináris (0/1) függő változós modell OLS-becslése²⁰, ahol a függő változó várható értéke a bekövetkezési valószínűséggel egyenlő.

¹⁹ Bár mint korábban szóba került, a futures árak nem tekinthetők a jövőbeli spot árak torzítatlan előrejelzéseinek.

²⁰ OLS (ordinary least squares): legkisebb négyzetek módszere.

Az előrejelzés során a következő egyenletet használom:

$$D_t = \beta_0 + \beta_1 y_{t-1} + \dots + \beta_p y_{t-p} + \gamma_1 x_{t-m} + \dots + \gamma_s x_{t+1-m-s} + \dots + \varepsilon_t, \quad /6/$$

ahol D_t a növekedés dummy a t . időszakban, melynek értéke 1, ha t -ben magasabb az ár, mint $t-1$ -ben, különben 0.²¹ A változók köre – és a késleltetési struktúra – pontosan ugyanaz, mint a korábbi modellekben, azaz kifejezhetik:

- a spot olajár logaritmikus differenciájának késleltetett értékeit;
- a futures olajárak logaritmikus differenciájának késleltetett értékeit vagy a spot árhoz képesti spread-et (ezúttal is az árak logaritmusai közötti különbségként definiálva);
- a különböző olajszármazékok hasonló transzformációt (dlog és spread).

Mivel előrejelzésként nem egy konkrét kimenetet (növekedés vagy csökkenés), hanem egy bekövetkezési valószínűséget ad a modell, az előrejelzéshez meg kell adni egy c küszöbértéket, ami alapján létrejön a projekció:

$$D_t^{\text{pred}} = \begin{cases} 1, & \text{ha } \hat{D}_t > c \\ 0, & \text{különben} \end{cases}, \quad /7/$$

ahol D_t^{pred} a t . időszakra vonatkozó előrejelzés (1, ha növekedés, 0, ha csökkenés), \hat{D}_t pedig az LPM-modellből nyert becslt valószínűség. A konkrét alkalmazások során $c = 0,5$ (50 százalékos küszöbszint) feltevéssel élek.

Logisztikus regresszió (logit). Az LPM-moddellel kapcsolatban gyakori kritika, hogy a becslt valószínűség nem feltétlenül esik a (0;1) intervallumba (kivéve, ha kizárólag *dummy* változókkal dolgozunk), illetve a hibatag heteroszkedaszticitása miatt a szokásos tesztek érvénytelenek. Ugyanakkor az olajárváltozás irányának előrejelzése szempontjából az LPM legnagyobb hátulütője, hogy nehezen boldogul a nem lineáris dinamikák leírásával. A logit modell ennél szofisztikáltabb, és több területen is sikerrel alkalmazott módszernek számít. Az eljárás egy látens – nem megfigyelhető – változós formulából indul ki:

$$D_t^* = \beta_0 + \beta_1 y_{t-1} + \dots + \beta_p y_{t-p} + \gamma_1 x_{t-m} + \dots + \gamma_s x_{t+1-m-s} + \dots + \varepsilon_t. \quad /8/$$

²¹ Formálisan: $D_t = \text{sign}(p_t - p_{t-1})$.

Amennyiben ez a látens (D^*) változó eléri a kritikus értéket (meghaladja a nullát), akkor az esemény bekövetkezik (az olajár nő), különben a függő változó 0 értéket vesz fel (az olajár csökken):

$$D_t = \begin{cases} 1, & \text{ha } D_t^* > 0 \\ 0, & \text{különben} \end{cases} \quad /9/$$

Az előrejelzés elvégzéséhez tehát a jobb oldali változók és a β paraméterek ismeretére van szükség. Utóbbiak kalibrálása *likelihood* becsléssel történik, mégpedig azzal a feltevessel, hogy a hibatag standard logisztikus eloszlást követ. Megmutatható (lásd *Wooldridge* [2012]), hogy ebben az esetben a bekövetkezés valószínűsége:

$$P(D_t = 1 | Y, X) = G(\beta_0 + \beta_1 y_{t-1} + \dots + \beta_p y_{t-p} + \gamma_1 x_{t-m} + \dots + \gamma_s x_{t+1-m-s}), \quad /10/$$

ahol $G(\cdot)$ a hibatag valószínűségi eloszlásához tartozó eloszlásfüggvény.²² Az innen nyert előrejelzések ismét bekövetkezési valószínűségek lesznek, így most is egy küszöbszint (továbbra is $c = 0,5$) alapján kell eldönteni, hogy növekedés vagy csökkenés várható a következő napon.

Itt érdemes megjegyezni, hogy a logit modellt ritkán használják idősoros elemzésekhez, mivel sérül a megfigyelések függetlenségének feltétele. Jelen esetben viszont a fellépő autokorreláció kevésbé számít, mert a cél nem a magyarázóváltozók parciális hatásának vagy a komplett adatgeneráló folyamatnak a feltárása, hanem az árváltozás irányának előrejelzése.

2.3. Gépi tanulás

Az ökonometriai modellekhez képest jóval általánosabb, komplexebb függvényformákkal dolgoznak a most bemutatásra kerülő gépi tanulási módszerek. Így viszont a paraméterek kalibrációja is nehezebb, és – főleg a neurális hálónál – könnyen előfordulhat, hogy az iterációs eljárás lokális optimumba konvergál. Ez nem csupán az előrejelzési képességet rontja, de az eredmények reprodukcióját is ellehetetleníti. Ezért – bár a paraméterek végleges értékeit külön nem közlöm – a leírásban minden olyan beállítást megnevezek, amivel a szimulációs eredmények megismételhetők.

²² Standard logisztikus eloszlásnál: $G(a) = 1 / (1 + e^{-a})$.

2.3.1. Neurális háló

A neurális hálóknak számos típusa létezik, itt most kizárólag a közgazdasági alkalmazásokban leggyakrabban használt, háomrétegű előrecsatolt neurális hálóval foglalkozom. Az eljárás során a bemeneti réteg neuronjai az input adatokat jelentik, ahonnan az információ a rejtett rétegbe kerül, melyben tetszőleges számú neuron található. A neuronok mindegyike egy-egy függvényt jelöl, ami a kapcsolódó input neuronokból származó változók értékeit transzformálja:

$$s_{kt} = f \left(\sum_{i=1}^p w_{ik}^y y_{t-i} + \sum_{j=1}^s w_{jk}^x x_{t+1-m-j} + c_k \right), \quad /11/$$

ahol s_{kt} a rejtett rétegben lévő k . neuron által adott függvényérték a t . időszakban, w_{ik}^y és w_{jk}^x az input változók k . neuronhoz tartozó súlyai (koefficiensei), c_k a neuronhoz tartozó torzítás (konstans tag), $f(\cdot)$ pedig az alkalmazott (jelen esetben szigmoid) függvényforma.

A rejtett réteg így J darab (J a rejtett rétegben található neuronok száma) értéket ad eredményül. Ezek az „outputok” a kimeneti rétegben található neuron inputjaiként szolgálnak, ahol egy függvény transzformálja őket egyetlen számmá, a függő változó előrejelzésévé:

$$y_t = g \left(\sum_{k=1}^J w_k s_{kt} + c \right), \quad /12/$$

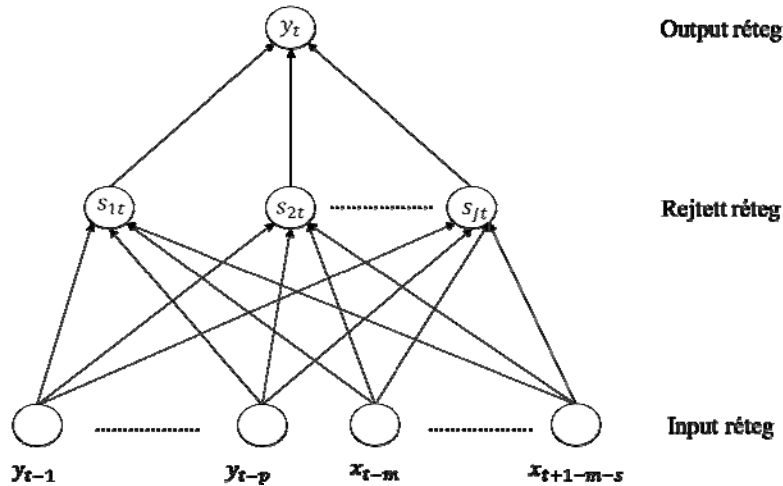
ahol w_j a rejtett réteg j . neuronjából érkező input súlyát (koefficiensét), c a torzítási együtthatót, $g(\cdot)$ pedig a transzformációs függvényt (identitásfüggvény) jelöli. A háló felépítését az 1. ábra mutatja:

Az input és output adatok a korábban is alkalmazott változókat takarják, de a tanulás gyorsítása, illetve a szigmoid függvény korlátos értéktartománya miatt előbb a $[0;1]$ intervallumba transzformálom őket a következő képlet segítségével:

$$x = \frac{x_{\text{eredeti}} - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}}. \quad /13/$$

Arra is figyelni kell azonban, hogy valós idejű előrejelzés esetén az idősor minimum és maximum értékei folyamatosan változhatnak, ezért azokat mindig az előrejelzés pillanatában már rendelkezésre álló, a tanuló vagy validációs mintában is szereplő időszak szélsőértékeivel azonosítom.

1. ábra. A használt előrecsatolt neurális háló sematikus ábrája



A kísérletek során egy rejtett réteget használok 10 neuronnal, míg a kimeneti rétegben értelemszerűen egyetlen neuronnal dolgozom. A túlillesztés elkerülése érdekében a keresztvalidációs eljárást alkalmazom, és a mintát tanuló, validációs, valamint teszt részekre bontom. A tanuló algoritmus arra törekszik, hogy a célfüggvényt, jelen esetben az átlagos négyzetes hibát minimalizálja a tanuló mintán, de a tanulási folyamat során úgy változtatja a súlyok és a torzítási tényezők értékeit, hogy az új értékeket csak akkor tartja meg, ha azok a validációs mintán is javítják az illeszkedést. Azaz a tanuló minta határozza meg, hogy milyen irányban és mértékben változnak a paraméterek, de az így nyert függvény teljesítményét a validációs mintán is teszteli, és elveti a módosításokat, ha ott a hibamutató növekedését tapasztalja. A tesztminta csupán ellenőrzésre szolgál, a tanulás során ezek az információk nem kerülnek felhasználásra. A keresztmetszeti alkalmazásokkal ellentétben, idősoroknál a minta felosztását nem célszerű véletlenszerűen elvégezni, ezért jelen esetben a tanuló minta mindig megelőzi az egyéves validációs mintát, azt pedig a szintén egyéves tesztminta követi (ennek pontos specifikációjára később még visszatérek). Ez egyben azt is jelenti, hogy a korábbiakkal ellentétben a modell évente csak egyszer frissül, és a következő 12 hónapban beérkező input adatok nem befolyásolják a kalibrációt (hasonlóan az ARIMA-modellhez). A paraméterek becslése a Levenberg–Marquardt-féle *backpropagation* módszerrel történik, a szimulációkat a MATLAB Neural Network Toolbox-ával végzem.

Mivel az iterációs folyamat a kezdőértékek függvényében eltérő eredményre vezethet, ezért minden alkalommal 50 szimuláció történik, és az így nyert előrejelzések illeszkedési mutatóit átlagolom, ezzel csökkentve annak lehetőségét, hogy

egy-egy kiugróan jó/rossz előrejelzés torzítsa a modell teljesítményének megítélését.²³ Hangsúlyozandó, hogy ezzel a módszerrel nem egy konkrét előrejelzés, hanem az eljárás hatékonyságát mérem, és könnyen elképzelhető, hogy nincs is olyan kivetítés, melynek teljesítménye épp a használt illeszkedési mutatók átlagával egyezne meg.

Az elemzés során az ökonometriai részben ismertetett felírásokat használom:

- NAR-modell²⁴ p késleltetéssel ($p = 1, 5, 10, 15, 20$);
- nem lineáris autoregresszív modell exogén (dlog és spread) változókkal $p = 1$ késleltetéssel.²⁵

A NARX-modell²⁶ esetén a dinamikus előrejelzés módszere némileg eltér a korábban bemutatott megoldástól, mivel a két-, illetve háromnapos előrejelzések során a becsléshez kizárólag az előrejelzések pillanatában már rendelkezésre álló adatok kerülnek felhasználásra. Ez azt jelenti, hogy minden egyes futás során három különböző (egy-, két-, illetve háromnapos) előrejelzés keletkezik az árváltozás nagyságára vonatkozóan ($\hat{y}_{t+1}, \hat{y}_{t+2}, \hat{y}_{t+3}$), amikből már visszaszámolható az árszintre vonatkozó kivetítés.

Neurális háló az irány előrejelzésére. A NAR- és NARX-modellek mintájára építhető olyan hálózat, ahol a bemeneti és a rejtett réteg szerkezete megegyezik a korábbival, de a kimeneti rétegben két neuron van, melyek a csökkenés/növekedés kategóriákat jelölik. Az LPM- és a logit modellek mintájára ezzel a konstrukcióval is elvégzem az egynapos árváltozás irányának előrejelzését.

Az eredeti háléhoz képest mutatkozó különbségek:

- a kimeneti rétegben két neuron van, mindkettő egy-egy szigmoid függvényt takar;
- az output két szám lesz, ezek közül a nagyobb érték mutatja meg, hogy melyik kimenet (növekedés vagy csökkenés) a valószínűbb;
- amikor kizárólag a spot olajár késleltetett dlog értékei szerepelnek a modellben (a NAR mintájára), akkor a $p = 1, \dots, 20$, míg a másik két esetben (a NARX mintájára) a $p = 1, 5, 10, 15, 20$ késleltetéseket vizsgálom;

²³ A futásszám optimális értékének meghatározására nincs univerzális szabály, jelen esetben azért esett 50-re a választás, mert az eredmények vizsgálatok kiderült, hogy bár többnyire 20-30 ismétlés is elég a konvergenciához, alkalmanként ennél tovább kell menni. Az 50-es beállítás ezekben a helyzetekben is elégnek bizonyult, így az összes esetben ezzel az értékkel dolgoztam.

²⁴ NAR (non-linear autoregression): nem lineáris autoregresszív.

²⁵ $p = 1$ -re azért esett a választás, mert a NAR-modellek közül NAR(1) adta a legpontosabb előrejelzéseket.

²⁶ NARX (nonlinear autoregressive exogenous): nem lineáris autoregresszív exogén.

– a benzin és a fűtőolaj változókat együtt vagy külön-külön bevonó modelleknél a paraméterek változtatása az SCG-módszerrel²⁷ történik, míg minden más esetben a Levenberg–Marquardt-eljárást használok.²⁸

A célfüggvény továbbra is az átlagos négyzetes hiba, a paraméterek évente frissülnek, és ezúttal is 50 futás eredményeit átlagolom.

2.3.2. Support vector machine (SVM)

Az SVM-eljárás során az algoritmus egy olyan hipersíkkal osztja két részre a teret, ami a két csoport „széleitől” egyenlő, és a lehető legnagyobb távolságra van. A kategóriák „szélei” azokat a megfigyeléseket, ún. tartóvektorokat (support vectors) jelentik, amelyek hipersíktól vett távolsága a legkisebb. Mivel gyakorlati alkalmazásoknál meglehetősen ritka a teljes szeparálhatóság, ezért a célfüggvényben büntetni kell a rosszul klasszifikált megfigyeléseket.²⁹

$$\min_{w,b,s} \left(\frac{1}{2} \langle w, w \rangle + c \sum_{t=1}^n s_t \right), \quad /14/$$

$$s.t.: \forall t - \text{re} D_t (\langle w, x_t \rangle + b) \geq 1 - s_t, \quad s_t \geq 0,$$

ahol w és b a hipersíkot meghatározó paraméterek ($\langle w, x \rangle + b = 0$), $D_t \in \{-1; 1\}$ a kiemeneti változó értéke (csökken vagy nő az ár), x_t a mintában található t . megfigyelés vektora (k . eleme mutatja, hogy a t . megfigyelés esetén mekkora a k . változó értéke), s_t egy nemnegatív segédváltozó, ami akkor és csak akkor vesz fel nulla értéket, amikor az t . megfigyelés klasszifikációja helyes.³⁰ A hibás besorolás relatív költségét szabályozó c paraméter értékét 1-re állítom, a bevont változókat pedig standardizálom. Ezen túlmenően egy Kernel-függvényt (a Gaussi radiális bázisfüggvényt) vezetek be, ami lehetővé teszi az eredeti térben történő nem lineáris szeparációt:

$$K(a, b) = \exp \left[\frac{-\langle (a-b), (a-b) \rangle}{2\sigma^2} \right]. \quad /15/$$

²⁷ SCG (scaled conjugate gradient): skálázott konjugált grádiens.

²⁸ Futures árak használata mellett az SCG-eljárás még ezer iteráció után sem találta meg az optimumot, és jellemzően a tesztminta minden napján az árak csökkenését tartotta valószínűbbnek. A Levenberg–Marquardt-módszerrel sikerült kiküszöbölni a hibát.

²⁹ Az itt ismertetett L^1 -norma felírás alternatívája az L^2 -norma használata, ahol az s_t segédváltozó négyzete szerepel a célfüggvényben.

³⁰ Tartóvektor esetén pedig $s_t = 0$ mellett az is igaz, hogy a feltétel egyenlőségre teljesül.

A σ^2 paraméter értékét szintén 1-re állítom a futások során, de ahogyan a neurális hálónál a neuronok számára, úgy itt erre és c -re is igaz, hogy változtatásuk révén optimalizálható az eljárás. A korábban felírt (primál) feladat egyetlen ponton változik, a feltételben $\langle w, x_t \rangle$ helyett $\langle w, \varphi(x_t) \rangle$ szerepel, ahol $K(a, b) = \langle \varphi(a), \varphi(b) \rangle$. A paraméterek becslése a duál feladat megoldásával történik SMO-eljárással³¹.

Az előrejelzések során a kategóriaváltozó továbbra is az olajár növekedését ($D = 1$) vagy csökkenését ($D = -1$) jelöli egynapos távlatban, míg a szétválasztás a korábban is alkalmazott változók és $p = 1, \dots, 20$ késleltetés szerint történik. A modell becslését és a szeparáló hipersík kialakítását itt is évente végzem: az év elején az addig rendelkezésre álló adatokon megkeresem a szeparáló hipersíkot, a következő 12 hónapban pedig attól függően adok növekvő/csökkenő előrejelzést az olaj árára, hogy a bevont változók alapján a hipersík mely oldalára esik az új megfigyelés.³² Az SVM-klasszifikáció a MATLAB Statistics Toolbox-ával történik.

2.4. Eredmények kiértékelése

A modellek teljesítményének mérésére nem létezik univerzális mutató. Könnyen előfordulhat, hogy egy kutató közgazdász számára a „pontosság” egészen más jelent, mint egy kereskedőnek, és még e kategóriák is tovább bonthatók. Elképzelhető, hogy az egyik kereskedőnek olyan eljárásra van szüksége, amely nagy biztonsággal jelzi, hogy milyen irányba mozdulnak az árfolyamok, míg a másiknak elég, ha az elmozdulás nagyságának abszolút értékét ismeri. A példákat sokáig lehetne sorolni, de a lényeg, hogy a modellek kiértékelése nem történhet egyetlen mutató alapján, így a következőkben több lehetőséget is megvizsgálók.

A pontbecslés jóságát a MAE-val³³ és az RMSE-vel³⁴ mérem:

$$\text{MAE} = \frac{\sum_{t=1}^n |p_t - \hat{p}_t|}{n}, \quad /16/$$

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^n (p_t - \hat{p}_t)^2}{n}}, \quad /17/$$

³¹ SMO (sequential minimal optimization): szekvenciális minimális optimalizálás.

³² A paraméterek 12 havonta történő frissítése ugyan nem a legoptimálisabb eljárás, azonban a becslések futási ideje miatt ennél gyakoribb ismétlésre nem volt mód. Könnyen elképzelhető, hogy a modellek havi/negyedéves kalibrációja pontosabb előrejelzésekhez vezetne.

³³ MAE (mean absolute error): átlagos abszolút hiba.

³⁴ RMSE (root mean squared error): átlagos négyzetes hiba gyöke.

ahol p_t a nyersolaj piaci ára a t . periódusban, \hat{p}_t az erre adott előrejelzés, n pedig az előrejelzések száma. A kiértékelés 2011 januárjától évente történik, így az első három évben rendre $n = 252$, míg 2014-ben $n = 211$ napnyi előrejelzésre számolom ki a /16/-/17/ mutatókat (egy-, két- és háromnapos kivetítésre egyaránt). Míg az abszolút hiba arányosan bünteti az eltéréseket, addig az RMSE nagyobb hibához nagyobb súlyt rendel. A kereskedés során fontos, hogy egy szerencsétlenül megválasztott pozíció esetén se legyen túl nagy a keletkező veszteség, ezért ilyenkor érdemesebb RMSE szerint értékelni a versenyző modellek teljesítményét.

Ami legalább ilyen fontos a tőzsdei szereplőknek, az a változás iránya: ha pontosan tudnák, hogy milyen irányba fog változni az olaj ára, akkor tulajdonképp az árváltozás mértékét sem kellene ismerniük ahhoz, hogy nyereséges stratégiát építsenek. Ezért a modellek azon képességét, hogy milyen jól jelzik előre a változás irányát, a találati rátával jellemzem (az esetek hány százalékában sikerült eltalálni az árváltozás előjelét):

$$DA = \frac{\sum_{t=1}^n \text{sign}(\Delta p_t) \cdot \text{sign}(\Delta \hat{p}_t)}{n} . \quad /18/$$

A találati arány használatának másik előnye, hogy segítségével összevethető a pont- és az irányelőrejelzést adó modellek teljesítménye, míg a MAE és RMSE szempontjából utóbbiak nem értelmezhetők.

Mivel a találati ráta sosem lesz 100 százalék, ezért korántsem biztos, hogy egy magasnak tűnő arány elég a sikeres kereskedéshez. Ha például valaki az esetek 70 százalékban a jó irányra „tesz”, de ezzel minden esetben csak mérsékelt hozamot realizál, miközben a maradék 30 százalékban többször is nagyot bukik, akkor könnyen lehet, hogy mínuszban zár. Ezért szokás nem csupán a találati arányt vizsgálni, hanem a fals pozitív/negatív előrejelzések arányát, illetve azok költségvonzatát figyelembe venni. Ezért az eredmények kiértékelése egy hipotetikus profitfüggvénnyel is megtörténik, mégpedig a következő feltevések mellett:

- a modell egynapos előrejelzésétől függ, hogy vétel (várhatóan emelkedés) vagy eladás (várhatóan csökkenés) történik;
- nincsenek adók és tranzakciós költségek;
- mind a vétel, mind az eladás a napi spot záróáron történik;
- a kereskedés során a teljes (a kezdetben rendelkezésre álló összeg hozamokkal korrigált értékének megfelelő) összeg felhasználásra kerül;
- vásárlás esetén a hozam megegyezik a spot olajárak változásával, eladás esetén a felhasználható összeg egészen addig változatlan, amíg nem történik újabb vásárlás (azaz csökkenő árak előrejelzése esetén a

meglévő készlet értékesítése történik a mai záró áron, és csak akkor kerül felhasználásra, ha a kivétítés áremelkedést jósol);
 – nincs alternatív befektetési lehetőség (vagy ami ezzel ekvivalens: kizárólag 0 százalékos hozamú egynapos betétben vagy olajban lehet tartani a pénzt).

Ennek megfelelően a t . kereskedési napig elért teljes hozam képlete a következő:

$$\text{PROFIT}_t = \left[I(\hat{p}_t > p_{t-1}) \cdot (1 + \text{PROFIT}_{t-1}) \cdot \frac{p_t}{p_{t-1}} - 1 \right] + \quad /19/ \\ + \left[1 - I(\hat{p}_t > p_{t-1}) \right] \cdot \text{PROFIT}_{t-1}$$

ahol PROFIT_t a t . időszakra elért hozam ($\text{PROFIT}_0 \equiv 0$), $I(\cdot)$ pedig egy indikátorfüggvény: ha a t . időszakra vonatkozó előrejelzés meghaladja az utolsó ($t-1$. napi) tényadatot, akkor $I(\cdot) = 1$, azaz vétel történik, különben $I(\cdot) = 0$ és a készletek értékesítésre kerülnek. Bár a valós kereskedés eredménye ennél jóval összetettebb formulával határozható meg, ráadásul az áresésből profitáló pozíciókat is lehet építeni, a /19/ mutató elemzéséből több hasznos megállapítás következik majd.

3. Adatok

A modelleket a WTI olajtípus napi spot árain futtatom. Az adatok forrása az U.S. EIA³⁵ honlapja. A spot árak között a következő változók szerepelnek (a dátum minden esetben a kereskedési napra vonatkozik):

- WTI spot, napi bontás, USD/hordó (2002.01.02.–2014.11.04.);
- New York Harbor Conventional Gasoline spot, napi bontás, USD/gallon (2002.01.02.–2014.11.04.);
- New York Harbor Heating Oil No. 2 spot, napi bontás, USD/gallon (2002.01.02.–2014.11.04.).

A benzin és a fűtőolaj esetén az 1 hordó = 42 gallon összefüggés alapján konvertáltam azonos mértékegységre (USD/hordó) a változókat. Emellett a WTI egy-, két-,

³⁵ U.S. EIA (Energy Information Administration): az Egyesült Államok Energiaügyi Hivatala.

három és négyhónapos, napi bontású futures árait használom (USD/hordó, 2002.01.02.–2014.11.04.).³⁶

Az idősorok nem teljesek, mivel nincs minden nap kereskedés (hétvége, ünnepnapok). Ezért azokat a napokat, ahol nem ismert a WTI spot ár, kihagyom az elemzésből (így például a hétfői ár előrejelzésénél az AR(1) változó a pénteki spot árra vonatkozik). Bár többnyire azonos napokon vannak zárva a piacok, előfordul, hogy míg a WTI spot piacon zavartalanul folyik a kereskedés, addig egy másik piacról nincs információ. A hiányzó megfigyelések okát az adattáblák nem nevezik meg, de viszonylag ritkák, így nem valószínű, hogy jelentősen befolyásolnák az eredményeket. Amennyiben a modell által becsült egyenlet jobb oldalán található változó értéke ismeretlen, az adott napot hiányzó megfigyelésként kezelem, és kihagyom az elemzésből. A tesztelési mintán (2011 januárjától 2014 októberéig) nincs ilyen probléma, azaz minden kereskedési napra elvégezhető az előrejelzés.

További probléma, hogy a spread, illetve dlog idősoroknál mi történjen abban az esetben, amikor nincs WTI, de van benzin vagy fűtőolaj spot ár. Ezeket az eseteket figyelmen kívül hagyom, azaz, ha egy pénteki napra van benzin, de nincs WTI-érték, akkor a hétfői előrejelzésnél a csütörtöki nap jelenti az utolsó időszakot. Az alacsony esetszám miatt itt megint arra lehet számítani, hogy a becslési eredményt érdemben nem befolyásolják a hiányzó értékek.

Mivel az árváltozás mögötti dinamikák (vagy akár a fundamentális változók köre) időben változnak, az előrejelzések során nem használok fel minden rendelkezésre álló adatot:

– a tanuló minta minden esetben 2002.01.02-án indul (*Kilian* [2009] szerint a világgazdasági reál aktivitás ettől az évtől kezdve gyorsult fel, amit az ázsiai – azon belül is elsősorban a kínai – keresletnövekedés táplált, és az olajpiaci folyamatokat is nagyban befolyásolta);

– a tanuló minta első végpontja 2010.12.31-ére esik, utána folyamatosan tolódik (ökonometriai modelleknél – kivéve ARIMA – mindig egy-egy napot, gépi tanuláshoz és ARIMA-nál pedig egy-egy évet ugorva), a választást az indokolja, hogy a neurális hálókhoz használt validációs mintában így nincs 2010 előtti adat, ez pedig mérsékli annak veszélyét, hogy a korábban tapasztalt trendszerűen emelkedő/csökkenő dinamikára tanul rá az algoritmus;

– SVM esetén az „éves” előrejelzési periódusok (tesztminták) kezdete rendre 2011.01.03., 2012.01.03., 2013.01.02. és 2014.01.02., míg

³⁶ Bár szokás egy-, két-, három, illetve négyhónapos futures árakról beszélni, valójában ez csak durva közelítésnek tekinthető. A pontos definíciókat lásd itt: http://www.eia.gov/dnav/pet/TblDefs/pet_pri_fut_tbldef2.asp.

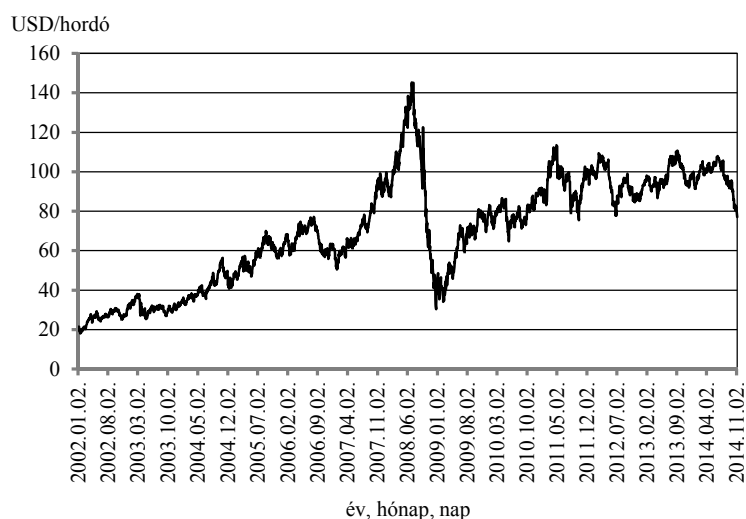
az ezekhez tartozó utolsó olyan nap, ahonnan még előrejelzés kezdődik rendre 2011.12.30., 2012.12.31., 2013.12.31 és 2014.10.31.³⁷

– neurális hálók esetén ez a felosztás annyiban módosul, hogy az eddigi tanuló minta utolsó egy évét validációs mintának használok, rendre 2010.01.04., 2011.01.03., 2012.01.03. és 2013.01.02. kezdési időpontokkal (a validációs minta utolsó eleme pedig a tesztminta első napját megelőző kereskedési nap lesz);

– a kiértékelés az „éves” tesztmintákon külön-külön történik.

A vizsgált időszakban a WTI alakulását a 2. ábra mutatja.

2. ábra. Napi WTI spot ár, 2002.01.02.–2014.11.04.



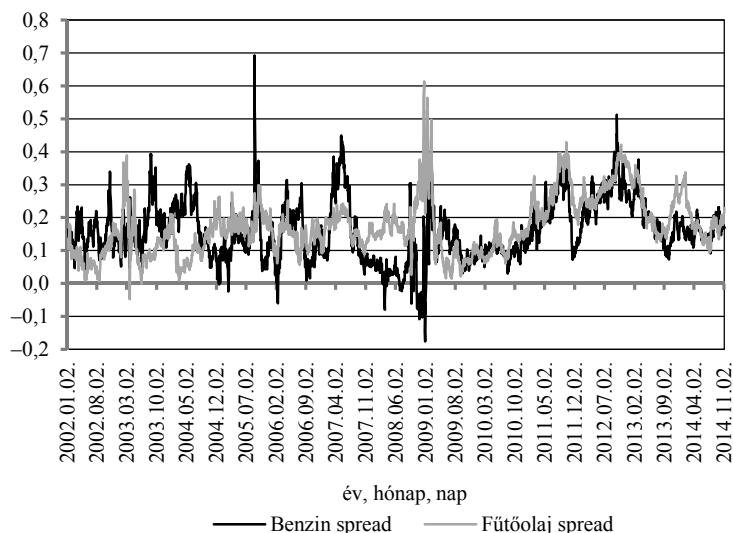
Forrás: Itt és a 3. ábránál U.S. EIA (<http://www.eia.gov/petroleum/data.cfm#prices>).

Látható, hogy a 2002-ben indult, a dinamikusan bővülő ázsiai kereslet táplálta ár-növekedés kisebb visszaesést követően 2007-től meredeken emelkedni kezdett, majd a pénzügyi válság eszkalálódása nyomán a 2008.07.03-án mért 145,31 USD/hordós csúcstról nagyot zuhant. Később a jegyzés korrigált, és a 80–110 USD/hordós sávban mozgott.³⁸ A különböző olajszármazékok árai 2011-ig ezzel nagyjából együtt mozogtak, onnantól kezdve azonban növekedtek a különbségek. (Lásd a 3. ábrát.) A futures áraknál nem történt hasonló törés.

³⁷ Ebből következően egy 2014.10.31-én induló háromnapos előrejelzés október 31-re, illetve november 3-ára és 4-ére is ad becslést.

³⁸ Ezt követően bőven 80 dollár alá is süllyedt a WTI hordonkénti ára, de az elemzés ezt az időszakot már nem tartalmazza.

3. ábra. Benzin és fűtőolaj spreadek, 2002.01.02.–2014.11.04.



A 2. táblázatban az egyes változók dlog idősorainak leíró statisztikái láthatók a teljes időszakra vonatkozóan. Minden esetben nulla közelében vannak az átlag- és mediánértékek, és az interkvartilis terjedelem (ahová az értékek középső 50 százaléka esik) is kivétel nélkül 3 százalékpont alatt marad. Ugyanakkor alkalmanként előfordulnak jelentős, változótól függően 10–25 százalékos elmozdulások is.

2. táblázat

A dlog idősorok leíró statisztikái, 2002.01.02.–2014.11.04.

Mutató	WTI spot	Benzin	Fűtőolaj	WTI F1	WTI F2	WTI F3	WTI F4
Átlag	0,0004	0,0004	0,0004	0,0004	0,0004	0,0004	0,0004
Medián	0,0010	0,0014	0,0000	0,0010	0,0009	0,0009	0,0010
Szórás	0,0233	0,0269	0,0215	0,0229	0,0210	0,0201	0,0196
Minimum	-0,1519	-0,1789	-0,1104	-0,1307	-0,1143	-0,1038	-0,1260
Maximum	0,1641	0,2351	0,1102	0,1641	0,1278	0,1212	0,1147
Terjedelem	0,3160	0,4140	0,2206	0,2948	0,2421	0,2250	0,2407
25. perc	-0,0114	-0,0139	-0,0115	-0,0117	-0,0111	-0,0106	-0,0101
75. perc	0,0128	0,0151	0,0128	0,0126	0,0120	0,0115	0,0110
IQR ³⁹	0,0243	0,0290	0,0244	0,0243	0,0231	0,0221	0,0211

Forrás: Itt és a további táblázatoknál saját számítás.

³⁹ IQR (interquartile range): interkvartilis terjedelem.

A spread idősorok hasonló értékeit a 3. táblázat mutatja. A benzin és a fűtőolaj esetén tapasztalt többletek abból erednek, hogy nem ugyanannak a terméknek a spot és futures árai kerülnek összevetésre, így a spread mutatók egyensúlyi értéke – feltéve, hogy egyáltalán létezik ilyen – a pozitív tartományba esik. Érdekes megfigyelés továbbá, hogy a futures árak esetén a horizont (hány „hónapos” határidő) növekedésével párhuzamosan (de nem arányosan) távolodnak az eloszlás kitüntetett pontjai a nullától, így az eloszlás terjedelme is folyamatosan növekszik. Ez megfelel az előzetes várakozásoknak, hiszen minél távolabb tekintünk a jövőbe, annál nagyobb elmozdulások lehetségesek. Az is látható azonban, hogy az egy-, illetve kéthónapos futures árak között van a legnagyobb ugrás, utána a változás üteme mérséklődik.

3. táblázat

A spread $[\ln(\text{változó}) - \ln(\text{WTI spot})]$ idősorok leíró statisztikái, 2002.01.02.–2014.11.04.

Mutató	Benzin	Fűtőolaj	WTI F1	WTI F2	WTI F3	WTI F4
Átlag	0,1590	0,1730	0,0007	0,0054	0,0073	0,0074
Medián	0,1490	0,1604	0,0001	0,0047	0,0079	0,0109
Szórás	0,0909	0,0864	0,0100	0,0244	0,0370	0,0474
Minimum	-0,1755	-0,0472	-0,1780	-0,2335	-0,2488	-0,2615
Maximum	0,6922	0,6139	0,2526	0,3279	0,3705	0,4028
Terjedelem	0,8677	0,6611	0,4306	0,5614	0,6193	0,6643
25. perc	0,0963	0,1117	-0,0006	-0,0048	-0,0132	-0,0220
75. perc	0,2112	0,2180	0,0015	0,0124	0,0215	0,0285
IQR	0,1148	0,1063	0,0022	0,0172	0,0347	0,0504

4. Eredmények

A különböző eljárások teljesítményének mérésére a következő benchmarkokat használom:

- RMSE: eltolás nélküli véletlen bolyongás;
- MAE: eltolásos véletlen bolyongás;
- találati arány: eltolásos véletlen bolyongás, ami mindig növekedést jelez előre, de abban az esetben, ha ez 50 százalék alatti eredményt ad (ez történik 2014-ben), akkor a szigorúbb 50 százalékos küszöböt alkalmazom;

– profit: a „buy and hold” stratégia, tehát az éves hozam megegyezik az időszak végén és elején fennálló árak közötti különbséggel (mivel 2012-ben és 2014-ben ez negatív hozamot eredményezne, ezekben az években a 0 százalékos értéket használom viszonyítási alapnak).

A küszöbszintek megválasztása az előrejelzés jósága alapján történt. A teljes vizsgált – 2011–2014-es – periódusban RMSE tekintetében az eltolás nélküli, míg az átlagos abszolút hibánál az eltolásos véletlen bolyongás teljesített jobban. Mivel 2014 kivételével az árváltozás iránya az esetek legalább felében pozitív volt, ezért itt a benchmark a mindig növekedést előrejelző heurisztikus modell lett, aminek eredménye egybeesik az eltolásos véletlen bolyongással. A profitmutatónál a szokásos „buy and hold” stratégiával dolgozom, de mivel ez két évben (2012, 2014) is negatív hozamot eredményez, ezért egy szigorúbb korlátot használok, és az ominózus periódusokban 0 százalékra állítom a küszöbszintet. A viszonyítási alapok konkrét értékeit a 4–7. táblázatok mutatják.

4. táblázat

RMSE benchmark értékek (eltolás nélküli véletlen bolyongás)
(USD)

Előrejelzés	2011.	2012.	2013.	2014.
	év			
Egynapos	2,0207	1,4897	1,1291	1,2359
Kétnapos	2,9043	1,9518	1,6409	1,6548
Háromnapos	3,5494	2,3409	2,0658	1,9636

5. táblázat

MAE benchmark értékek (véletlen bolyongás eltolással)
(USD)

Előrejelzés	2011.	2012.	2013.	2014.
	év			
Egynapos	1,5038	1,1155	0,8911	0,8929
Kétnapos	2,1716	1,5016	1,2592	1,2269
Háromnapos	2,7825	1,7523	1,6315	1,5122

6. táblázat

Találati arány (DA) benchmark értékek (véletlen bolyongás eltolással, kivéve 2014)
(százalék)

Előrejelzés	2011.	2012.	2013.	2014.
	év			
Egynapos	52,78	53,17	52,78	50,00
Kétnapos	52,38	50,00	53,97	50,00
Háromnapos	56,35	52,38	53,17	50,00

7. táblázat

Profit benchmark értékek („buy and hold”, kivéve 2012 és 2014)
(százalék)

Megnevezés	2011.	2012.	2013.	2014.
	év			
Éves hozam	8,15	0,00	6,90	0,00

8. táblázat

A benchmark értékeknél jobban teljesítő modellek aránya az összes modellhez képest
(százalék)

Modell	Előrejelzés	2011.	2012.	2013.	2014.
		év			
RMSE	Egynapos	2,72	9,90	5,45	48,27
	Kétnapos	2,72	0,99	14,11	33,91
	Háromnapos	12,13	0,00	36,39	23,76
MAE	Egynapos	11,88	15,59	9,16	53,22
	Kétnapos	4,95	23,27	16,09	57,43
	Háromnapos	58,66	3,22	44,80	26,73
DA	Egynapos	16,38	16,26	45,53	82,04
	Kétnapos	55,94	68,07	13,61	63,12
	Háromnapos	34,16	5,20	9,90	29,70
PROFIT		47,28	26,54	53,39	13,72

A 8. táblázatban az egyes cellák azt fejezik ki, hogy az oszlop szerinti évben, a sor szerinti hibamutató/előrejelzési horizontpár alapján a vizsgált modellek hány

százaléka adott a vonatkozó benchmarknál jobb eredményt. Megjegyzendő, hogy míg az egynapos találati arány és a profit mutatók esetén összesen 1654 előrejelzés kerül az elemzésbe, addig a többi esetben csupán 404 modellel dolgozom. A különbséget az egynapos irány-előrejelzésre épített LPM-, logit, SVM- és neurális háló modellek okozzák. Az eredményekből megállapítható, hogy az egyes módszerek relatív – a benchmarkhoz viszonyított – teljesítménye mind az előrejelzési horizont (egy-, két- vagy háromnapos előrejelzés), mind a vizsgált időszak (év) tekintetében erősen volatilis. Az egynapos projekciók RMSE-értékei például évtől függően az esetek 2,5–48 százalékban haladják meg a küszöbszintet, és hasonló nagyságrendű eltérések tapasztalhatók a MAE és a találati arány tekintetében is. Az előrejelzési horizont változása is nagyban befolyásolja a felülteljesítő modellek arányát: RMSE és MAE mellett is igaz, hogy míg 2013-ban az előrejelzési horizont növekedésével párhuzamosan bővült a benchmarkot legyőző modellek halmaza, ez a trend 2014-ben megfordult.

9. táblázat

*A benchmark értékeknél 0, 1, 2, 3, illetve 4 alkalommal (évben)
is jobban teljesítő modellek aránya az összes modellen belül
(százalék)*

Modell	Előrejelzés	0	1	2	3	4
		alkalommal				
RMSE	Egynapos	44,06	46,78	7,92	1,24	0,00
	Kétnapos	61,63	26,98	9,41	1,98	0,00
	Háromnapos	57,18	21,78	12,62	8,42	0,00
MAE	Egynapos	39,60	39,36	12,62	8,42	0,00
	Kétnapos	34,16	38,86	18,56	7,92	0,50
	Háromnapos	20,79	38,37	28,96	10,40	1,49
DA	Egynapos	9,13	32,35	48,79	8,65	1,09
	Kétnapos	2,48	23,02	48,02	24,26	2,23
	Háromnapos	40,35	41,58	16,83	1,24	0,00
PROFIT		13,18	43,53	33,62	8,52	1,15

Az eredmények alapján tehát kijelenthető, hogy a legtöbb modell relatív teljesítménye az előrejelzési horizont és a vizsgált periódus függvényében változik. Kérdés azonban, hogy létezik-e a modelleknek olyan szűk részhalmaza, amely a többséggel ellentétben mentes az efféle robusztusságbeli hiányosságoktól, és képes folyamatosan legyőzni a benchmarkot. A 9. táblázat elemei mutatják, hogy a sor szerinti hibamutató/előrejelzési horizont pár tekintetében a modellek hány százaléka teljesített

jobban 0, 1, 2, 3, illetve 4 alkalommal (évben) a heurisztikus eljárásnál. Az RMSE benchmarkot – főleg a 2012-es kivetítések pontatlansága miatt – egyetlen módszer sem tudta minden évben felülteljesíteni, és három alkalommal is csak egy meglehetősen szűk körük (1–8%) volt rá képes. A nagy eltéréseket kevésbé büntető MAE-mutatónál már valamivel jobb a helyzet, de itt is csak a két-, illetve háromnapos előrejelzéseknél sikerült néhány olyan modellt találni, amely minden évben felülteljesített (az összes modell 0,5–1,5 százaléka). Találati arány esetén ez az érték előrejelzési horizonttól függően 0–2 százalék, míg a profitmutatónál 1,15 százalék.

Bizonyos esetekben tehát léteznek olyan eljárások, amelyek hosszabb távon is elérik/meghaladják a küszöbszintet, azonban bőven találunk kivételt. További kérdés, mennyire stabil e halmaz az előrejelzési horizont dimenziójában, azaz jellemző-e, hogy az egynapos viszonylatban legtöbbször felülteljesítő modell két- és háromnapos kivetítésnél is megőrzi e kedvező tulajdonságát. Ennek eldöntéséhez a 10. táblázat nyújt segítséget, amelynek elemei azt fejezik ki, hogy az oszlop szerinti mutató esetén hány modell esik a benchmarkot legtöbbször legyőzők halmazába egy-, két- és háromnapos projekciónál, illetve mekkora e halmazok metszete.

10. táblázat

A benchmark értékeknél a legtöbb alkalommal jobban teljesítő modellek halmazának közös része

Előrejelzés	RMSE-modell			MAE-modell			DA-modell		
	Egynapos	Kétnapos	Háromnapos	Egynapos	Kétnapos	Háromnapos	Egynapos	Kétnapos	Háromnapos
	előrejelzés								
Egynapos	5	0	3	34	2	2	18	0	0
Kétnapos	–	8	8	–	2	2	–	9	0
Háromnapos	–	–	34	–	–	6	–	–	5

RMSE-nél például 5 eljárás esik a legjobb egynapos előrejelzők csoportjába (kivétel nélkül 5 késleltetést alkalmazó ARX-modellek, amelyek a futures és a spot árak közti spread értékeket használják), míg hosszabb távon 8 (kétnapos) és 34 (háromnapos) modell tartozik ugyanebbe a halmazba. Érdekes megfigyelés, hogy kivétel nélkül futures árakat használó ARX spread modellekről van szó, azaz RMSE tekintetében jellemzően a határidős áráktól vett eltérés bevonásával javítható az előrejelzés pontossága. Ugyanakkor az előrejelzési horizonttól függően változik e modellek köre: az egy- és kétnapos kivetítéseknek nincs közös részhalmaza, de a kétnapos távlatban legjobb modellek kivétel nélkül a háromnapos felülteljesítők kategóriájába esnek, azaz a két nap alatt megszerzett „előny” a következő 24 órában is kitart. Az

eltolás nélküli véletlen bolyongáshoz képest egyébként nem túl nagy a javulás mértéke, jellemzően 1 százalék alatt marad a csökkenés, ami alól csak a 2014-es év jelent kivételt, amikor a negatív trend miatt könnyebb megverni a változatlan árak kivetítését (akár 2-3 százalékkal is). Ennek tudatában nem meglepő, hogy az utolsó időszakot leszámítva a benchmark és a modellalapú előrejelzések „azonosságának” nullhipotézise csak ritkán utasítható el.⁴⁰

11. táblázat

Az olajszármazék és futures spread modellek találati arányainak jellemző értékei (átlag, medián, 90. percentilis) (százalék)

Előrejelzés	Mutató	2011.		2012.		2013.		2014.	
		év							
		Olaj-származék	Futures spread modell	Olaj-származék	Futures spread modell	Olaj-származék	Futures spread modell	Olaj-származék	Futures spread modell
Egynapos	Átlag	52,78	51,65	51,87	49,79	50,64	49,86	51,41	55,65
	Medián	51,98	51,19	52,38	49,60	50,79	49,60	51,18	55,92
	90. perc	56,75	53,57	53,57	51,75	52,86	53,21	54,50	58,34
Kétnapos	Átlag	55,11	53,48	51,96	50,70	51,51	49,67	49,28	53,89
	Medián	55,95	54,17	52,18	50,60	51,39	48,81	48,82	53,79
	90. perc	58,73	57,14	53,97	53,17	55,16	53,21	53,55	55,92
Háromnapos	Átlag	57,29	55,63	51,00	49,58	50,99	50,56	46,41	52,13
	Medián	57,94	56,35	51,59	49,80	50,40	50,79	46,21	52,37
	90. perc	59,17	57,58	52,78	52,38	55,56	52,38	48,39	54,98

Megjegyzés. A félkövér kiemelés jelzi azokat az eseteket, amikor a futures modellek DA-értékei magasabbak.

MAE esetén a benzin spread-et használó ARX-eljárások jelentik a három halmaz közös részét (egy, illetve két késleltetéssel),⁴¹ és továbbra is jellemző, hogy a legjobbba közé kevés kivételtől eltekintve ARX spread modellek kerültek, de míg RMSE-nél kizárólag a határidős árak szerepeltek, itt már többször előjönnek a különböző

⁴⁰ A tesztelés során *Clark–West* [2006] eljárását alkalmaztam, és a nullhipotézis szerint a két modell által adott előrejelzések négyzetes hibáinak különbsége várható értékben minden nap nulla.

⁴¹ Ugyanakkor az eltolásos véletlen bolyongáshoz képest minimális a javulás mértéke. Sosem haladja meg az 1 százalékot, és a *Diebold–Mariano*-teszthez (*Diebold–Mariano* [1995]) is végig 10 százalék feletti *p*-értékek tartoznak. A DM-teszt nullhipotézise szerint a két modell által adott előrejelzések abszolút hibáinak különbsége várható értékben minden napon nullával egyenlő.

olajszármazékokat használó eljárások. Ez arra enged következtetni, hogy az olajszármazékok és a WTI spot árak különbségét használó előrejelzések relatíve gyakran vétettek nagy hibát. Ez különösen a 2014-es évre igaz, ahol a trendforduló utáni csökkenő árakat a futures modellek sokkal jobban le tudták követni: míg a 2011–2013-as időszakban az olajszármazék spread-re építő ARX-becslések DAMutatóinak átlaga, mediánja és 90. percentilise szinte minden év/előrejelzési horizont pár esetén magasabb volt, addig a 2014-es évben jóval a futures csoport hasonló mutatói alá estek (lásd a 11. táblázatot), az ezzel járó nagyobb előrejelzési hibák következtében pedig az RMSE-értékek is megugrottak. Mivel az első három évben az olajszármazékos csoport jobb találati arányai ellenére is jellemzőbb volt a magasabb RMSE-mutató, gyanítható, hogy a nagyobb előrejelzési hibák eleve gyakrabban előfordulnak ebben a körben, a trendtörés pedig csak ráerősített e tulajdonságra. Bár konkrét bizonyíték nincs rá, könnyen elképzelhető, hogy a futures árak gyorsabban reagálnak a megjelenő hírekre, így felhasználásukkal csökkenthető a nagy előrejelzési hibák előfordulásának valószínűsége.

Visszatérve a 10. táblázathoz, megállapítható, hogy találati arány szempontjából az egynapos kivetítéseknél túlnyomórészt bináris függő változós modellek kerülnek a halmazba, és az SVM révén itt már a gépi tanulás is képes a folyamatos felülteljesítésre. Mind itt, mind a két- és háromnapos projekcióknál jellemző, hogy a modellek között egyaránt van olajszármazék vagy határidős árat (esetleg mindkettőt) használó eljárás, azaz nincs olyan magyarázóváltozó halmaz, amely előrejelzés tekintetében dominálná a másikat. Ennek következtében a három halmaz metszete üres, ami abban a tekintetben nem meglepetés, hogy az egynapos irány-előrejelzéseknél a korábbanál jóval bővebb, körülbelül négyszer akkora elemszámú csoportból lettek kiválogatva a módszerek. Megjegyzendő továbbá, hogy minden év/előrejelzési horizont pár esetén valahol 55–60 százalék környékén található a legnagyobb találati aránnyal bíró modell, azaz a korábban hivatkozott (*Haidar–Kulkarni* [2009], *Yu–Wang–Lai* [2008]) tanulmányokban található 80 százalék feletti értékek valós idejű előrejelzésnél nem tarthatóak hosszú távon.⁴²

Az egyes bináris függő változós modellekhez tartozó, éveken át a viszonyítási alap fölött teljesítő irány-előrejelzések pontossága alighanem hozzájárul ahhoz, hogy döntően e módszerekből kerülnek ki a legnagyobb profitot eredményező eljárások is. Bár az összes vizsgált módszer mintegy kétharmada bináris függő változót használ, a legjobban teljesítők között ennél jóval nagyobb az arányuk (84%), csak elvétve találni egy-két folytonos függő változójú modellt. Itt ugyanakkor érdemes kitérni a „legjobban teljesítő” definíciójára. Ha ugyanis az eddigiekhez hasonlóan a teljesítményt abban mérjük, hogy hányszor tudta felülmúlni az előrejelzés a benchmarkot, akkor a jó relatív

⁴² Annál is inkább, mivel a legjobb modellek halmazába egyetlen olyan eljárás sem került, amelynél mind a négy évben el lehetett volna utasítani a Pesaran–Timmermann-teszt nullhipotézisét (*Pesaran–Timmermann* [1992]).

teljesítmény robusztusságáról beszélünk, míg egy kereskedőnek adott esetben fontosabb lehet a hosszú távon, tehát a teljes időszak során realizált profit. Az eredményekből kitűnik, hogy a két halmaz nem esik egybe: míg a folyamatos felülteljesítők között elsősorban bináris függő változós ökonometriai (LPM- és logit) modellek szerepelnek, amelyek mind az olajszármazékok spot, mind az olaj határidős árait felhasználják, addig a közel négy év egésze alatt begyűjtött legmagasabb hozamok az SVM-eljáráshoz tartoznak.

12. táblázat

*A teljes előrejelzési periódus (2011–2014) alatt
a 20 legmagasabb profitmutatót eredményező modell*

Modell	Késletelés ideje (nap)	Profit (százalék)
SVM_S_F2	18	80,35
SVM_S_F3	18	76,38
LOGIT_D_MINDEN*	10	72,79
LOGIT_D_MINDEN*	12	70,48
SVM_S_F4	18	69,34
LOGIT_D_F1234*	12	66,87
LOGIT_D_F1234*	10	65,08
LOGIT_S_F1	5	63,07
LOGIT_D_MINDEN	8	62,99
LPM_S_F1	5	62,78
LPM_S_F1*	4	62,44
LOGIT_S_F1*	4	61,34
LPM_D_F1	9	60,67
LPM_D_MINDEN*	11	60,34
LPM_D_F1234	9	60,28
SVM_S_F2	5	60,20
LPM_D_MINDEN*	12	60,20
LOGIT_D_MINDEN*	19	59,55
LOGIT_D_F1234*	9	59,08
LOGIT_S_F1	7	58,31

Megjegyzés. Az első oszlopban a modell neve (módszertan + típus (S: spread, D: dlog) + magyarázóváltozók (FX – X „hónapos” futures, MINDEN – 4 futures, illetve benzin és fűtőolaj spot)), a másodikban a késletelések száma (p) szerepel. A *-gal jelölt modelleken alapuló kereskedési stratégiák mind a négy év során a küszöbszintet meghaladó profitot eredményeztek.

A 12. táblázatból leolvasható, hogy a futures spread-eket használó SVM-ek akár 70–80 százalékos profitot is hozhattak, ám egyikük sem fér bele az eddig használt

„legjobb modellek” definíciójába. A 18 késleltetéssel dolgozó, „kéthónapos” futures spread-et használó SVM például hiába eredményezi a legnagyobb bővülést, ennek felhalmozása döntően az időszak elejére esik, és az utolsó két évben már nem éri el a küszöbszintet.⁴³ Ez az eredmény megerősíti azt a következtetést, amely szerint az eljárások rangsora attól függ, hogy az időszak, előrejelzési horizont és hibamutató „paramétereknek” milyen kombinációját használjuk.

Végül érdemes megvizsgálni az ökonometriai és gépi tanulós modellek egymáshoz viszonyított teljesítményét. Az egyszerűség kedvéért az elemzés csupán az egyes csoportokra jellemző jósági mutatók medián értékeire fókuszál.

13. táblázat

A gépi tanulós modellek ökonometriai modellekhez mért relatív teljesítménye (mediánok hányadosa) (százalék)

Modell	Előrejelzés	2011.	2012.	2013.	2014.
		év			
RMSE	Egynapos	99,66	99,14	99,61	100,69
	Kétnapos	100,08	100,64	99,83	100,10
	Háromnapos	100,01	100,20	100,12	99,41
MAE	Egynapos	99,75	99,01	99,39	100,84
	Kétnapos	100,45	100,30	99,89	100,88
	Háromnapos	100,06	100,13	100,27	99,39
DA	Egynapos	99,22	100,78	95,56	95,68
	Kétnapos	94,43	98,12	102,11	99,04
	Háromnapos	96,37	103,41	102,46	104,00
PROFIT		0,71	-0,19	-4,99	-4,84

Megjegyzés. Az RMSE, MAE és DA esetén a medián hibamutatók hányadosa, profitnál a hozamok különbsége szerepel. Félkövéren kiemelve azok az esetek, ahol a gépi tanuláson alapuló modellek jobb „medián” előrejelzőnek bizonyultak.

A 13. táblázatban a gépi tanuláson alapuló modellekhez tartozó RMSE-, MAE- és DA-hibaértékek mediánjainak, valamint az ökonometriai eljárások hasonló mutatóinak hányadosai szerepelnek. A hipotetikus profitfüggvény esetén a medián hozamok közti különbséget („gépi tanulás mínusz ökonometria”) vizsgálok. Ennek megfelelően a 100 százalékos alatti RMSE és MAE, illetve a 100 százalékos feletti DA, valamint a pozitív profit mutatók jelölik a gépi tanulás jobb medián teljesítményét (ezeket az eseteket a táblázatban is kiemeltem). Az eredmények alapján nem lehet egyértelmű

⁴³ Míg 2011-ben közel 65 százalékos többletet eredményezett, utána rendre 6,5, 5,3, illetve -2,3 százalék volt a modell hozadéka.

összefüggést találni, hol az egyik, hol a másik csoport szerepel jobban. Ez rávilágít arra, hogy bár a gépi tanuláson alapuló modellek ritkán tudták elérni a kijelölt küszöbszinteket, medián teljesítményük még így sem maradt el az ökonometriai módszerekétől. Ezért a tanuló algoritmusok paramétereinek optimalizálása, a megfelelő késleltetési struktúra megtalálása, illetve a kalibrálás – jelenleg éves – gyakoriságának növelése révén akár az is elképzelhető, hogy a hiányolt robusztusság elérhető.

Az előbbieket ismeretében a következő konklúziók vonhatók le:

1. Függetlenül attól, hogy az időszak, előrejelzési horizont és hibamutató „paraméterek” milyen kombinációját alkalmazzuk, az olajszármazék spot ill. az olaj futures árak bevonásával javítható az előrejelzés pontossága. Míg RMSE tekintetében egyértelműen a futures spread-ekre épülő modellek használata javallott, addig az átlagos abszolút hiba, a találati arány vagy a bemutatott hipotetikus profitfüggvény esetén mindkét változócsalád alkalmazható.

2. Az ökonometriai modellek robusztusság tekintetében ugyan jobban teljesítenek a gépi tanulásnál, azonban összehasonlításuk több szempontból is problémás. Egyrészt a machine learning eljárásoknál évente csak egyszer frissülnek a modellek, másrészt a tanuló algoritmusok paramétereinek kioptimalizálása nem történt meg, azok egyetlen beállítás mellett futnak. Könnyen előfordulhat, hogy a megfelelő módosításokkal ez a modelles család is versenyképesé válna. Ugyanakkor már most is lehet olyan mutatót (profit) találni, amelynél az SVM kifejezetten jó teljesítményt nyújt (legalábbis a teljes időszak során elért hozamok tekintetében), ráadásul a gépi tanuláson alapuló modellek medián teljesítménye egyáltalán nem rosszabb az ökonometriai megközelítésekenél.

3. A legjobban teljesítő modellek halmaza mind az előrejelzési horizont, mind a vizsgált időszak függvényében változik, és csak néhány olyan eljárás van, ami hosszabb távon végig felülteljesít. Mivel összesen négy év vizsgálata történt meg, ezért elképzelhető, hogy ez a teljesítmény csupán a véletlen műve, és hosszabb tesztelési mintával dolgozva már nem lennének képesek folyamatosan túlteljesíteni a küszöbszintet. Ennek azért is nagy az esélye, mert sem az RMSE-, sem a MAE-, sem pedig a DA-mutatóknál nem volt olyan modell, ami minden évben szignifikánsan jobb lett volna a benchmarknál.

4. Ebből az is következik, hogy bár az eredmények alapján úgy tűnik, lehetséges az egyszerű heurisztikus szabályoknál jobb modellt építeni, ennek időbeli robusztussága megkérdőjelezhető, és sosem lehetünk biztosak abban, hogy egy éveken át felülteljesítő modell a következő hónapokban/években is megőrzi e képességét.

5. Összefoglalás

Tanulmányom célja az olajár előrejelzésére használható módszerek egy igen széles körének bemutatása, a modellek teljesítőképességének összehasonlítása volt. A WTI napi spot idősorát ökonometriai modellekkel és gépi tanulással jeleztem előre, felhasználva a WTI futures, illetve két olajszármazék (benzin és fűtőolaj) spot referencia árait. Az elemzés során egy-, két- és háromnapos valós idejű kivetítéseket generáltam a 2011 januárjától 2014 novemberéig terjedő tesztidőszakra, így a kapott eredmények megmutatják, milyen hatékonysággal működtek a modellek a gyakorlatban. A kiértékelés során több alternatív mutatót (RMSE, MAE, találati arány, profit) is használtam, és a „jó modellek” halmazát a mutatóként eltérő benchmark modell eredményeihez képest határoztam meg. Konklúzióm szerint a „legjobb” modellek halmaza erősen függ az előrejelzési horizont nagyságától, valamint a tesztidőszak és a hibamutató megválasztásától. A fundamentális változók bevonása ugyanakkor minden esetben preferált, a legtöbb évben felülteljesítő modellek köre szinte kivétel nélkül felhasználja az olaj futures és/vagy az olajszármazékok spot árait. Hogy pontosan milyen modellt, és milyen változókkal, késleltetési struktúrával érdemes használni, már a felhasználó konkrét céljától (a hibamutató megválasztásától) függ. Arra utaló eredmények is születtek, hogy bizonyos esetekben (a hipotetikus profitfüggvény maximalizálásakor) érdemes lehet átlépni az ökonometriai modellek határát, és a gépi tanulás eszközeihez folyamodni. Azonban itt is szembetűnő a teljesítménybeli robusztusság hiánya, előfordul, hogy az egyik évben kiválóan teljesítő modell már a következő időszakban a heurisztikus szabálynál rosszabb előrejelzést produkál. Figyelembe véve a legjobb modellek halmazának erős változékonyságát, nem szabad egyetlen eljárásra támaszkodni, és ígéretesebbnek tűnik a különböző kivetítések között szelektálni illetve azokat kombinálni. A gépi tanuláson alapuló modellek paramétereinek optimalizálása, új módszerek kipróbálása bővítené az ehhez használható előrejelzések körét. Jelen kutatás természetes továbblépési irányának tűnik e szelektáló, kombináló eljárás(ok) és az újonnan bevonandó modellek halmazának megtalálása.

Irodalom

- ALQUIST, R. – KILIAN, L. [2010]: What Do We Learn from the Price of Crude Oil Futures? *Journal of Applied Econometrics*. Vol. 25. No. 4. pp. 539–573.
- ALQUIST, R. – KILIAN, L. – VIGFUSSON, R. J. [2011]: *Forecasting the Price of Oil*. CEPR Discussion Paper. No. 8388. Centre for Economic Policy Research. London.
- BASHIRI, B. N. – MANSO, P. J. R. [2013]: *Crude Oil Price Forecasting Techniques: A Comprehensive Review of Literature*. <http://ssrn.com/abstract=2275428>
- BAUMEISTER, C. – KILIAN, L. [2013]: *Forecasting the Real Price of Oil in a Changing World: A Forecast Combination Approach*. Working Paper. No. 2013-28. Bank of Canada. Ottawa.

- BAUMEISTER, C. – KILIAN, L. – ZHOU, X. [2013]: *Are Product Spreads Useful for Forecasting? An Empirical Evaluation of the Verleger Hypothesis*. Working Paper. No. 2013-25. Bank of Canada. Ottawa.
- BERK, R. A. [2008]: *Statistical Learning from a Regression Perspective*. Springer. New York.
- BROOKS, C. – REW, G. – ALISTAIR, S. – RITSON, S. [2001]: A Trading Strategy Based on the Lead-lag Relationship between the Spot Index and Futures Contracts for the FTSE 100. *International Journal of Forecasting*. Vol. 17. No. 1. pp. 31–44.
- CLARK, T. E. – WEST, K. D. [2006]: Using Out-of-Sample Mean Squared Prediction Errors to Test the Martingale Difference Hypothesis. *Journal of Econometrics*. Vol. 135. No. 1–2. pp. 155–186.
- CRISTIANINI, N. – SHAWE-TAYLOR, J. [2000]: *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. First Edition*. Cambridge University Press. Cambridge.
- DIEBOLD, F. X. – MARIANO, R. S. [1995]: Comparing Predictive Accuracy. *Journal of Business and Economic Statistics*. Vol. 13. No. 3. pp. 253–263.
- FERNANDEZ, V. [2010]: Forecasting Energy Spot Prices. In: *Castro, R. – Emery, X. – Kuyvenhoven, R.* (eds.): *Minin 2010. Proceedings of the 4th International Conference on Mining Innovation*. Gecamin. Santiago. pp. 271–278. http://www.minin2010.com/evento2010/images/stories/docs/front_pages.pdf
- Haidar, I. – Kulkarni, S. [2009]: Forecasting Model for Crude Oil Price Using Artificial Neural Networks and Commodity Futures Prices. *International Journal of Computer Science and Information Security*. Vol. 2. No. 1. pp. 81–88.
- Haidar, I. – Wolff, C. R. [2011]: *Forecasting Crude Oil Price*. Proceedings of the 30th USAEE Conference. 9–12 October. Washington, D.C.
- HAMILTON, J. D. [1994]: *Time Series Analysis*. Princeton University Press. Princeton.
- HAMILTON, J. D. [2009]: Understanding Crude Oil Prices. *The Energy Journal*. Vol. 30. No. 2. pp. 179–206.
- KANG, S. H. – KANG, S. M. – YOON, S. M. [2009]: Forecasting Volatility of Crude Oil Markets. *Energy Economics*. Vol. 31. No. 1. pp. 119–125.
- KILIAN, L. [2009]: Not All Oil Price Shocks Are Alike: Disentangling Demand and Supply Shocks in the Crude Oil Market. *American Economic Review*. Vol. 99. No. 3. pp. 1053–1069.
- KRIESEL, D. [2007]: *A Brief Introduction to Neural Networks*. <http://www.dkriesel.com>
- KRUGMAN, R. [1991]: Target Zones and Exchange Rate Dynamics. *The Quarterly Journal of Economics*. Vol. 106. No. 3. pp. 669–682.
- LANZA, A. – MANERA, M. – GIOVANNINI, M. [2005]: Modeling and Forecasting Cointegrated Relationships among Heavy Oil and Product Prices. *Energy Economics*. Vol. 27. No. 6. pp. 831–848.
- MA, T. – LIU, Z. – ZHOU, Z. [2013]: *Inhibition of Envelope Correction Method to EMD End Effect*. Proceedings of the 2nd International Conference on Computer Science and Electronics Engineering. Atlantis Press. Paris. http://www.atlantis-press.com/php/download_paper.php?id=4959
- MERINO, A. – ORTIZ, A. [2005]: Explaining the So-called ‘Price Premium’ in Oil Markets. *OPEC Energy Review*. Vol. 29. No. 2. pp. 133–152.
- PESARAN, M. H. – TIMMERMANN, A. [1992]: A Simple Nonparametric Test of Predictive Performance. *Journal of Business and Economic Statistics*. Vol. 10. No. 4. pp. 461–465.

- SHABRI, A. – SAMSUDIN, R. [2014]: *Daily Crude Oil Price Forecasting Using Hybridizing Wavelet and Artificial Neural Network Model*. *Mathematical Problems in Engineering*. <http://www.hindawi.com/journals/mpe/2014/201402/#B33>
- SILVAPULLE, P. – MOSSA, A. [1999]: The Relation between Spot and Future Prices: Evidence from the Crude Oil Market. *The Journal of Futures Markets*. Vol. 19. No. 2. pp. 175–193.
- STATNIKOV, A. – HARDIN, D. – ALIFERIS, C. F. [2006]: *Using SVM Weight-Based Methods to Identify Causally Relevant and Non-Causally Relevant Variables*. *Proceedings of the Neural Information Processing Systems 2006 Workshop on Causality and Feature Selection*. http://ccdlab.org/paper-pdfs/NIPS_2006.pdf
- TANG, L. – HAMMOUDEH, S. [2002]: An Empirical Exploration of the World Oil Price under the Target Zone Model. *Energy Economics*. Vol. 24. No. 6. pp. 577–596.
- VAPNIK, V. N. [1995]: *The Nature of Statistical Learning Theory. Second Edition*. Springer. New York.
- WANG, S. Y. – YU, L. – LAI, K. K. [2004]: A Novel Hybrid AI System Framework for Crude Oil Price Forecasting. *Lecture Notes in Computer Science*. No. 3327. pp. 233–242. http://link.springer.com/chapter/10.1007%2F978-3-540-30537-8_26#page-1
- WANG, G. C. S. – JAIN, C. L. [2003]: *Regression Analysis – Modeling and Forecasting*. Graceway Publishing Company. Great Neck.
- WOOLDRIDGE, J. M. [2012]: *Introductory Econometrics: A Modern Approach. Fifth Edition*. South-Western Cengage Learning. Mason.
- XIE, W. – YU, L. – XU, S. – WANG, S. [2006]: *A New Method for Crude Oil Price Forecasting Based on Support Vector Machines*. *Lecture Notes in Computer Science*. No. 3994. Springer. New York. http://link.springer.com/chapter/10.1007%2F11758549_63#page-1
- YE, M. – ZYREN, J. – SHORE, J. [2005]: A Monthly Crude Oil Spot Price Forecasting Model Using Relative Inventories. *International Journal of Forecasting*. Vol. 21. No. 3. pp. 491–501.
- YE, M. – ZYREN, J. – SHORE, J. [2006]: Forecasting Short-run Crude Oil Price Using High and Low Inventory Variables. *Energy Policy*. Vol. 34. No. 17. pp. 2736–2743.
- YOUSEFI, S. – WEINREICH, I. – REINARZ, D. [2005]: Wavelet-Based Prediction of Oil Prices. *Chaos, Solitons and Fractals*. Vol. 25. No. 2. pp. 265–275.
- YU, L. – WANG, S. Y. – LAI, K. K. [2008]: Forecasting Crude Oil Price with an EMD-Based Neural Network Ensemble Learning Paradigm. *Energy Economics*. Vol. 30. No. 5. pp. 2623–2635.

Summary

Oil price fluctuations and the relevant data generating process have long been researched by economists and market analysts. Lately, the recent downward trend has put them and their short-run predictability in the focus again. The study provides insight into the main oil price forecasting techniques and examines the accuracy of these real time predictions via West Texas Intermediate spot prices. It presents not only standard econometric tools but also some machine learning models to predict nominal spot prices for a 1-3 day interval. The results suggest that the inclusion of oil futures prices and gasoline/heating oil spot prices in the model can reduce forecast errors, but the relative performance of the models is affected by the projection horizon, the test set and the way we measure prediction errors. That means, there is no universal model for short-run prediction.