KARSZTFEJLŐDÉS XIII.
Szombathely, 2008. pp. 87-102.

A BÜKKI KARSZTVÍZÉSZLELŐ RENDSZER HIÁNYZÓ MÉRÉSI ADATAINAK PÓTLÁSA; VÍZSZINT ELŐREJELZÉS

NÉMETH ÁGNES - LÉNÁRT LÁSZLÓ - KOVÁCS BALÁZS

Miskolci Egyetem, Hidrogeológiai-Mérnökgeológiai Tanszék, H-3515 Miskolc-Egyetemváros nemeth.agnes85@gmail.com; hgll@uni-miskolc.hu; modflow@gmail.com

Abstract: In 1992, the water companies concerned in the karst water exploitation of the Bükk had asked the University of Miskolc to work out a system for karst water level prognosis. During monitoring of the measurement points, instrument and measurement error have led to data gaps in the time-series of the water level, which correlations (between the different measurement points) are very important, however until now any reliable "replacement solution" has not been found. Knowing all values of the entire time-series, we can predict the changes of the karst water level caused by the precipitation, for example the flood or drought. Thus, we propose corresponding protective measures. In addition knowing the relationship between the time-series, in one hand less measurement could also be sufficient and in the other hand it can eliminate the gaps in the time series. Using 15 years time-series of 6 monitoring places (these wells had the more comprehensive and detailed time series available), and a monitoring station of precipitation, firstly I investigate the correspondence between the water level data and the precipitation. In some cases we can see, that the diagrams are very similar, and their development over time too. Another solution of the missing data of the time-series is the self-learning algorithm, based on artificial intelligence learning. Time series were supposed like multiple functions by the self-learning algorithm, which can transform simpler forms (function). Both methods provided very appropriate solution of the replacement of the missing data and prediction, even so maybe using the sophisticated method we can get more reliable result in the future.

1. Bevezetés

A Miskolci Egyetem Hidrogeológia - Mérnökgeológia Tanszékét 1992-ben bízták meg a Bükkben érintett vízmű Rt-k a bükki karsztvízszint előrejelzés kidolgozásához szükséges sokévi karsztvízszint adatsor kimérésének munkálataival. Ennek egyik kiváltó oka az volt, hogy a korábbi években igencsak veszélybe került a terület vízellátása. Az elmúlt 47 évben tartósan is előfordult rendkívüli aszály és rendkívüli csapadékbőség (*LÉNÁRT 2005, 2008*).

A mérések során műszerhibákból, mérési hibákból adathiányok álltak elő, melyek pótlására még nem született megnyugtató megoldás. A karsztvízszint adatok feltételezett korrelációja viszont több okból is meghatározó: teljes adatsorból ugyanis előre jelezhetők a nem kívánatos csapadék okozta vízszintváltozások (aszály és árvíz), és ezzel együtt a megfelelő óvintézkedések is előre megtervezhetők. Az összefüggések ismeretében hosszú távon kevesebb mérés is elegendő, és az előforduló mérési hibák is könnyebben kiküszöbölhetőek.

Hat megfigyelőhely és egy csapadékmérőhely legutóbbi 15 éves adatsorát felhasználva az adatpótláshoz első megközelítésben az egyes megfigyelőhelyeken mért vízszintadatok csapadékkal való kapcsolatát vizsgáltuk.

A karsztvízszint - adathiányok pótlására két lehetőséggel éltünk. Abban az esetben, ha több, hasonló lefutású görbénk ismert rövidebb időszakra, akkor ezek meglévő értékeihez viszonyítva az adathiányok pótlása egyszerűbb ,módon lehetséges, melynek alapját a regressziós vizsgálat (lineáris korreláció) adja. Hosszabb időszakokban azonban már összetettebb a feladat. Mind a geológiai mind a hidrológiai, mind az éghajlati tényezők máshogy jelentkeznek a különböző mérési helyeken, különböző időben, így ezek megnehezítik az esetleges korrelációk megtalálásának a lehetőségét.

A Szegedi Tudományegyetem Számítógépes Algoritmusok és Mesterséges Intelligencia Tanszékén évek óta folynak kutatások idősorok elemzésének, pótlásának és előrejelzésének új matematikai, mesterséges intelligencia alapú megoldásaira. A háromfős kutatócsoport részére eljuttattuk a bükki karvízszint-megfigyelőrendszer adatait, akik a módszer fejlesztése és tesztelése érdekében elvégezték a futtatásokat, aminek hidrogeológiai - karsztvízföldtani kiértékelését mi végeztük el.

2. Az adatpótlás lehetőségei

Az eddigi kutatási eredmények alapján a mérési rendszer legfontosabb tagja a Nagymezőn lévő Nv-17 figyelőkút s ebből történt a Bükkre általánosítható összefüggések megfogalmazása.

A vízszint adatsorok korrelációs vizsgálatát először lineáris, majd egyéb, magasabb fokú regresszióval kezdték. Az adatsorok naponkénti lépcsőkben történő eltolásával megkeresték a legszorosabb kapcsolatot, és az így kapott eltolási időtartam megadta a két mérőhely közti késleltetési időt. Legegyszerűbben a lineáris korreláció értelmezhető, ezért a továbbiakban ez került felhasználásra: két észlelőhely korábbi kapcsolatának súlyozott átlagából létrehoztak egy korrigált adatsort.

Kétféle lehetőséget tételeztünk fel a karsztvízszint-adatok hiányának pótlására. Az első lehetőség az egyszerűbben kivitelezhető, lényegében hasonló a korábbi módszerhez: amennyiben az adatok egy emelkedési, vagy csökkenési szakasz részeként, nem túl nagy időintervallumban hiányoznak, azok pótlása interpolációval, vagy regressziós vizsgálattal lehetséges. Mi a regressziós módszert használjuk.

A másik lehetőség már némileg összetettebb. A következő vizsgálatokat végeztük:

• Először a vízszintadatok hatékony csapadékcsoportokkal való kapcsolatára próbáltunk összefüggéseket keresni. Megnéztük, hogy a különböző megfigyelőhelyek a hatékony csapadékcsoportra való reagálása hány nap után következik be, majd azt is, hogy ugyanezen csapadékcsoport hatására mikor érik el a vízszintek a maximum értékeiket. Ezekből az adatpárokból meghatároztuk a görbék emelkedési szögét (iránytangensét) Megjegyeznénk, hogy a téli hónapokban a reagálási idő, és ebből kifolyólag a maximális vízszint eléréséhez szükséges idő több mint 10 nap (előfordult 45 nap is), ezt a késői, hirtelen felmelegedést követő olvadás okozhatja, ezért valószínűleg ezen adatokat a későbbiekben külön kell vizsgálni.

• Ezután megnéztük a csapadékcsoport kezdete és a reagálási idő, a csapadékcsoport kezdete és a maximum elérése között eltelt idő, és az emelkedés nagyságának kapcsolatát. A tapasztalat azt mutatja, hogy a csapadékcsoportokra való reagálások és a maximumok elérése minden egyes megfigyelőhelyen hasonlóan alakulnak, a legtöbb esetben pár napos eltérések figyelhetők meg.

Nagyon jó korrelációt mutat az Szinva-forrás és Garadna-forrás: néhány esettől eltekintve (ezek a téli hóolvadást követő időpontok) a csapadékra történő reakcióik (főleg a reagálási idők) majdnem ugyanazon a napokon következnek be. Ez azért is lesz fontos, mert az egyikből becsülni tudjuk a másik hiányosságát. Mivel ismerjük a csapadékcsoport kezdetét, 1-2 napos eltéréssel jó közelítést tudunk adni mind a görbék emelkedésének kezdeti időpontjaira, mind a maximális vízszint eléréséhez.

Ugyanezt állapíthatjuk meg a Felsőtárkány Sz-5 és az Miskolctapolca 2. sz. megfigyelőhelyek esetében is, de itt is fenntartással kell élnünk a téli hónapokban. Azt is tudjuk azonban, hogy a Felsőtárkány Sz-5-re jellemző magas vízállás csak a 2000-es évtől tapasztalható, így a pótlás csak ettől az időponttól kezdve lehetséges. Mindezek ellenére a Miskolctapolca 2. sz. termelés hatására történő heves reagálásai miatt sem a regressziós, sem az összetett módszerrel nem pótolhatók az adathiányok.

Ha ugyancsak eltekintünk a hóolvadás és a kora tavaszi esőzések okozta vízszintemelkedésektől, N-17-re és a Tebepuszta-1-re (Tbp-1) is ráillik a fenti megállapítás, kiegészülve még azzal, hogy a Tbp-1 reagálásai általában korábban kezdődnek. Ha a megállapítást felhasználjuk, és átlagoljuk a követési időket, akkor:

- a reagálási idők közötti különbség: a Szinva-forrás és Garadna-forrás között 0 nap, a Felsőtárkány Sz-5 és Miskolctapolca 2.sz. figyelőkút között 3 nap, és a Nagyvisnyó-17 és a Tebepuszta-1 között pedig 1 nap lesz az eltérés értéke.

- a maximum eléréséhez szükséges időtartam közötti különbség: a Szinva-forrás és a Garadna-forrás között 0 nap, Nagyvisnyó-17 és a Tebepuszta-1 között 5 nap, Felsőtárkány Sz-5 és Miskolctapolca 2. sz. figyelőkút között 1 nap.

3. Az egyszerűen elvégezhető pótlások

Az egyszerűen elvégezhető adatpótlások csoportjába azokat az adathiányokat soroltuk, melyek vagy egy emelkedési, vagy egy lecsengési időszak nem túl hosszú időtartama alatt adódtak.

Első lépésben megvizsgáltuk a különböző görbék regressziós kapcsolatát. Ez alapján választottuk ki a legszorosabb korrelációt mutató görbepárokat, s a továbbiakban ezeket együtt használtuk az adathiányok kiegészítésére. Azonban a feladatot megnehezítette az a tény, hogy csaknem minden görbénél az adathiányok néhány esetben közel azonos időszakokban léptek fel – feltehetően szisztematikus mérési- vagy műszerhiba miatt – , ezért ilyenkor a második legszorosabb korrelációt mutató adatrendszer került felhasználásra. Így előfordult, hogy egy összefüggő adathiány pótlása két másik megfigyelőhely adataiból készült el.

4. Az összetett adatpótlás

Összetett adatpótlás alatt azok a vizsgálatok értendők, melyek nemcsak egy emelkedési vagy lecsengési szakasz intervallumába esnek, hanem több hónapot, esetleg évet átölelnek. A korrekciók:

● A hatékony csapadékcsoportra történő legszorosabb reagálást illetve maximumot mutató megfigyelőhelyeket továbbiakban páronként vizsgáltuk. A kiszámított átlag értékek segítségével megkaptuk azoknak a megfigyelőhelyeknek is a reagálási, illetve maximum elérési időpontjait, melyek még hiányoztak. (Megjegyzendő, hogy bizonyos esetben mindkét figyelőhely vízszintadata hiányzott, így ilyenkor más évek átlagából pótoltuk a hiányokat.)

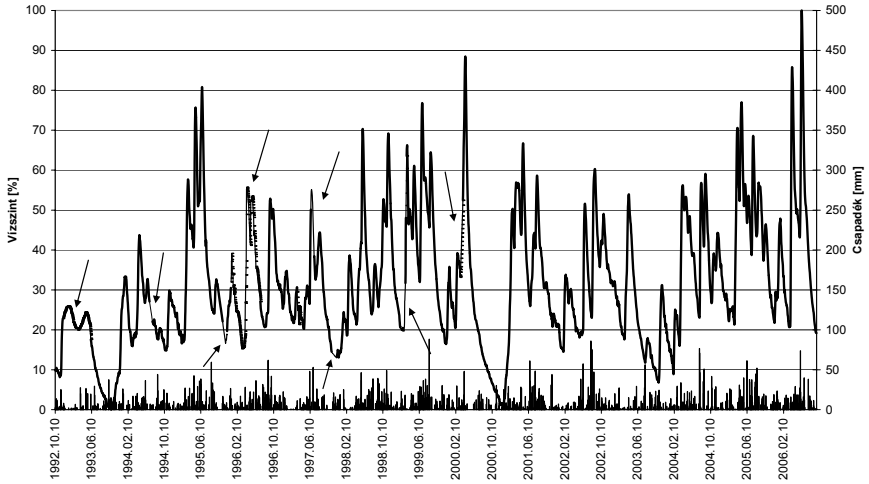
● Következő lépésben a hiányzó iránytangenseket pótoltuk az egyes megfigyelőhelyek esetében. Három alaphelyzet fordult elő:

- Ha csak a reagálási időpontot kellett pótolni, akkor az árhullám emelkedésének, vagy lecsengésének (esetenként mindkettőnek) az iránytangensét véve könnyen vissza tudtuk számolni azt. Ugyanígy működik ellenkező esetben is, amikor a maximum időpontja hiányzik.
- Ha reagálási és maximum időpont sincs, akkor a többi év átlagából megnéztük a szóba jöhető időpontokat, hiszen sok esetben közel azonos nap

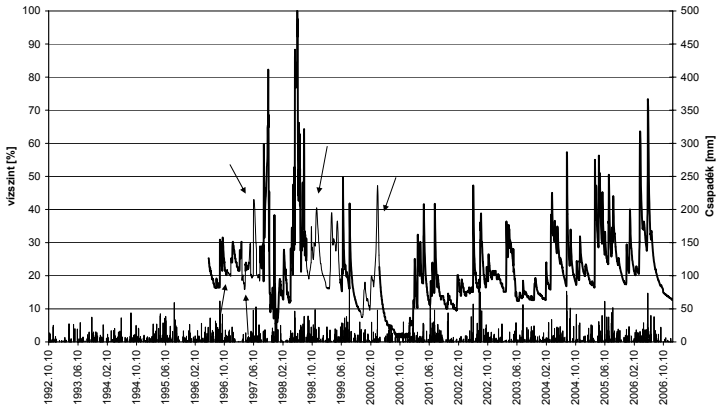
után következik be a csapadékra történő reagálás. Ezek után már csak a megfelelő csapadékcsoporthoz kellett illeszteni az eredményt.
- Ha ugyancsak hiányzik a reagálási időpont, és a maximum időpontja is, akkor a további megfigyelőhelyek időpontjaira támaszkodva az előző fejezetben számolt átlagreagálásokat és átlagmaximumokat használtuk a hiányok pótlására. (A pontosabb eredmény érdekében érdemes minél több esetet megvizsgálni)

Miután az iránytangenseket, és az emelkedési és lecsengési szakaszok kezdeteit és maximumait megkaptuk, az előző fejezetben készített regressziós összefüggéseket felhasználva, a késleltetési idő függvényében néhányszor lehetőség nyílt a vízszintek pótlására. Előfordult az is (pl.: Szinvaforrásnál), hogy semmilyen regressziós összefüggés nem hozott megnyugtató megoldást, ilyen esetekben a megfigyelőhely további emelkedési vagy csökkenési szakaszainak eltolásával történt a pótlás.

Az *1. ábrán* az Nv-17 adatsora található, megkülönböztetve az egyszerű, az összetett, és a felülvizsgálat után el nem fogadott pótlásokat. Ez utóbbi esetben csak 40-50 cm-es különbséget tapasztaltunk a korábbi és az általunk kapott pótlások között.

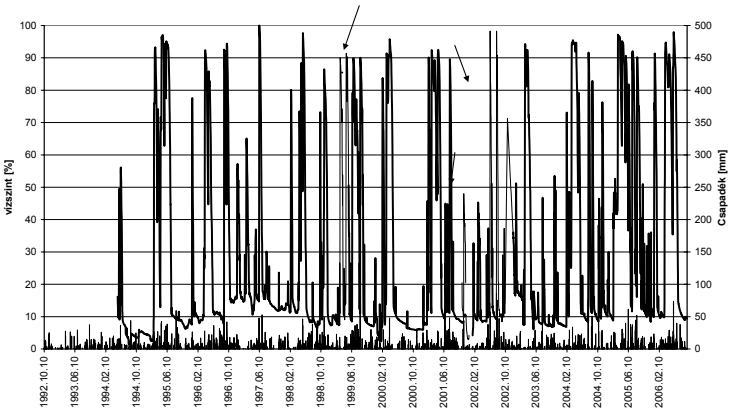


1. ábra: Az Nv-17 egyszerű, és felülvizsgált pótlásai, vékony vonal:pótlások, vastag vonal: mért értékék
Fig. 1: The simple and multiple completion of the Nv-17, thick lines: measured values, thin lines: completioned values



2. ábra: A Garadna-forrás egyszerű, és összetett pótlásai, vékony vonal:pótlások, vastag vonal: mért értékék
Fig.2: The simple and multiple completion of the Garadna-spring, thick lines: measured values, thin lines: completioned values

A *2. ábrán* a Garadna-forrás regresszióval, illetve összetett vizsgálattal történő pótlásai láthatók. A legtöbb esetben az Nv-17 vízszintészlelőkúttal mutatott regressziós kapcsolata alapján történt a hiányzó értékek pótlása.

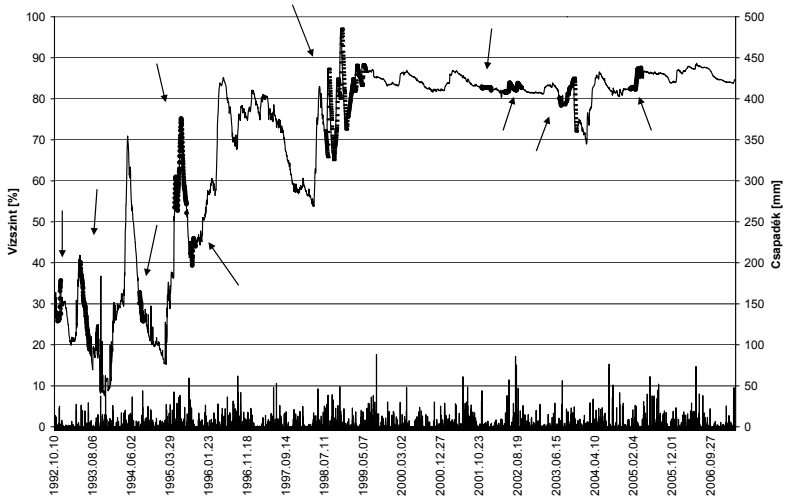


3. ábra: A Szinva-forrás egyszerű, és összetett pótlásai, vékony vonal pótlások, vastag vonal: mért értékék
Fig.3: The simple and multiple completion of the Szinva-spring, thick lines: measured values, thin lines: completioned values

A Szinva-forrás pótlásait a *3. ábra* mutatja. Mint azt már korábban is megállapítottuk, annak ellenére, hogy a Garadna-forrás vízszint változásá-

nak jellegével mutatta a legszorosabb kapcsolatot, hiányzó vízszintjeinek pótlásai mégis az Nv-17 vízszintadatainak felhasználásával adódtak. Ennek oka az, maga az árhullám alakja sem növekvő, sem a csökkenő ágában nem mutat a Garadna-forráséval hasonlóságot.

A Felsőtárkány Sz-5 pótlásait mind a Nagyvisnyó-17, mind a Tebepuszta-1 vízszintadatainak alkalmazásával nyertük, melyek a *4. ábrán* találhatók.



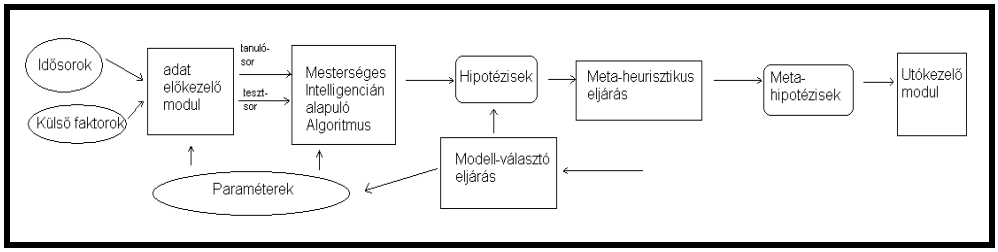
4. ábra: Felsőtárkány, Sz-5 figyelőkút pótlásai , vékony vonal: mért értékek, vastag vonal: pótlások
Fig.4: The simple and multiple completion of the Felsőtárkány Sz-5, thick lines: completioned values, thin lines: measured values

5. A neuron hálózaton alapuló megoldás

A vízszint-idősorok pótlásának és előrejelzésének matematikai jellegű értelmezése neuronhálózatok alapján történhet. A vízszint-idősorokat többváltozós függvényként tételezték fel, melyek egyenessé transzformálásához szükséges függvények megtalálásához öntanuló algoritmusokat használtak. Az algoritmusok úgy keresik a transzformált térben az idősort legjobban közelítő egyenest, hogy mérik a közelítés hibáját, majd ezt minimalizálni próbálják. A hiba nagyságát az n dimenziós térben vizsgálják, aminek lokális és abszolút minimumhelyei lehetnek. Annak érdekében, hogy lehetőleg a modell megtalálja az abszolút minimumhelyeket, ezért a hibaminimum keresés során lehetőség van a lokális minimumot jelentő térségekből való kilépésre (*Szarvas 2003).*

A Szegedi Tudományegyetem Számítógépes Algoritmusok és Mesterséges Intelligencia Tanszékén évek óta folynak kutatások különféle problémák mesterséges intelligencián alapuló megoldásaira. A Busa-Fekete Róbert, Ormándi Róbert, Szarvas György alkotta kutatócsoport idősorok elemzésével és pótlásával új matematikai megoldások kidolgozásával foglalkozik, akik részére Szanyi János hidrogeológus támogatásával számos vízföldtani észlelési adatsort, köztük a bükki karvízszint-megfigyelőrendszer adatait is eljuttattuk. A Szegedi Tudományegyetemen elvégezték a futtatásokat a módszer fejlesztése és tesztelése érdekében, aminek hidrogeológiai – karsztvízföldtani értékelését magunk végeztük el.

A neurális hálózatok működése két fő fázisra különíthető. Az első fázisban, melyet tanulásnak nevezünk, eltároljuk a háló struktúrájába a kívánt információ-feldolgozó eljárást, majd a másodikban a tárolt eljárás felhasználásával végezzük el az információ-feldolgozást *(5. ábra, JUHOS – SZARVAS 2003):*



5. ábra: A neuron hálózat működésének elvi rajza
Fig. 5: The time-series completion, prediction framework

Először a vízszint-idő függvény egy előzetes matematikai szűrésen megy keresztül, aminek célja a transzformált térbeli illesztések lehetségessé tétele. Ilyen szűrések során az eredeti adatrendszer kis mértékben módosul, amennyiben a függvényről valamilyen kiugró érték pl. mérési hiba kerül kiszűrésre akkor a minőség javul, de a szűrésnek hasznos információk is áldozatul eshetnek, ami a végén az adatpótlás vagy az előrejelzés minőségében jelentkezik. A szűrés során egyben megtörténik az adatsor normalizálása, a magas dimenziójú paraméterek dimenziójának csökkentése. Ennek megfelelően a vízszintváltozások abszolút nagysága a matematikai megoldás során az eredményeket nem befolyásolja.

Aztán a két részre osztott adatok a tanuló algoritmusba kerülnek, ahol a tanuló modul a rendszer tanítását végzi – a rendelkezésre álló adatokon, illetve megfelelő paraméter-beállítás mellett. Az ún. hipotézisek terében – többdimenziós tér, az input adatokon az összes szóba jöhető szabályszerűség – keressük a tanulási példánkra legjobban illeszkedőt. Ezután, - a

jobb tanulás érdekében – a megtanult hipotézis helyességét a teszthalmazon ellenőrizzük, és mérjük annak hibáját. A hiba egy „jósági” – fitness – függvényt határoz meg, melynek értékét heurisztikus kereséssel, a mi esetünkben szimulált hűtéssel (SA, Simulated Anheating) próbáljuk maximalizálni. A maximum érték meghatározásával a rendszer egy jó hipotézist állít fel a meglévő adatok segítségével. A keresés végén kapott fitness függvény értékét meghatározva döntünk arról, hogy optimum-e vagy nem, ha nem, akkor újraindítjuk a tanulókat a maradék (eredeti) tréninghalmazon, amivel a hipotézis folyamatosan javul. A modell- választó a legjobb hipotézis megtalálásáig finomítja az értéket, miközben a beállításokat az egész iteráció alatt megőrzi.

A megfelelő hipotézis megtalálása után a hipotéziseket egy ún. metaheurisztikus eljárással tovább finomítjuk, meta-hipotéziseket gyártva, melyek közül az egyik megadja a pótlás illetve az előrejelzés jövőbeli értékét. Egy ún. választó – decision – függvény segítségével kiválasztjuk azt a legjobb hipotézist, ami a teszt-sorozat legkisebb empirikus hibáját tartalmazza.

Végül, miután az első lépésben az adatokat normalizálták, a végső eredmény eléréséhez az utókezelő modulban az adatok denormalizálása történik.

A számítások során kétféle algoritmust, az ANN-t (Artifical Neural Network) és az SVR-t (Support Vector Regression) használtak fel, miközben az adatpótlások során egy vagy több egyidejű idősor adataiból is próbáltak számításokat végezni. Az előrejelzések során értelemszerűen további adatsorok felhasználása nem lehetséges, hiszen azokat sem ismerjük, ugyanakkor a transzformációs-függvények képzéséhez más közeli mérőpontok vízszint vagy csapadék-adatai is felhasználhatók.

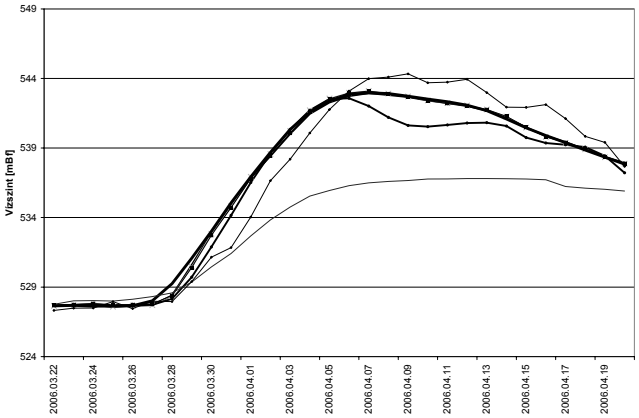
6. A neuron hálózatok alkalmazásával kapott eredmények

Az adatok pótlására a kiválasztott évet megelőző év ugyanazon hónapját használták fel. Minden egyes módszernél inputként kétféle bemeneti információval közelítettek: az egyik esetben kizárólag az Nagyvisnyó-17 adatait használták, másik esetben pedig az Nv-17 adataihoz hozzárendelték még a Tebepuszta-1 és a csapadék adatokat is, melyeket zajként, illetve beágyazott, magas dimenziójú adatként értelmeztek.

A vizsgálatok során kétféle adat-felhasználási metodikát alkalmaztak. Az egyik – amiket a továbbiakban „nem görgetett” vízszint adatsornak neveztek el – abból adódott, hogy az adatok pótlása csak a meglévő, mért adatok transzformációjával történt. A másik megoldásban – „*görgetett*” vízszintadatok – viszont a hiányzó vízszint adatokat már a korábban megjósolt

adatok felhasználásával, és azok transzformációjával pótolták. A módszer lényege az, hogy a mért vízszintekből megjósoljuk a következő napra várható vízszintet, majd az azt követő nap pótlását a mért és az előtte lévő, megjósolt nap értékeiből is számítjuk. Ebben az esetben azonban több a hibakeletkezés lehetősége, hiszen az előre jelzett (jósolt) adatok már önmagukban is tartalmazhatnak valamennyi hiba-paramétert, és ezek halmozása az újabb adatok pótlásában csak növeli az újonnan kapott értékek hibáit. Amint a *7. ábrán* is jól látszik, ez a megoldás semelyik vízszintváltozásnál nem adott elfogadható megoldást, a későbbiekben tehát a görgetett adatsorból történő pótlásokat kizártnak tekintjük (elkészítettük mind a három hónap görgetett adatsorral történő futtatások eredményeit, azonban helyhiány miatt nem részleteznénk – mindhárom eset alátámasztja a megállapítást).

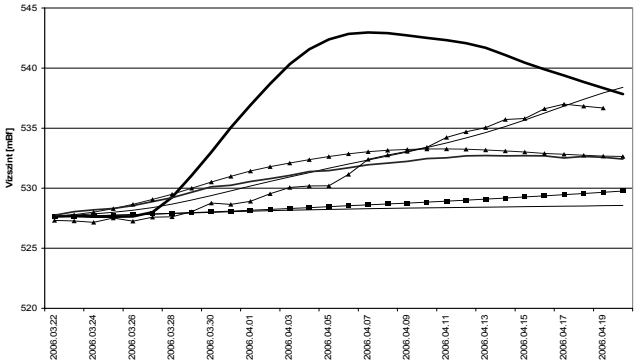
Azonban azt meg kell jegyezni, hogy más területeken, ahol ilyen hirtelen, ugrásszerű változások nem jellemzőek, pl. talajvízszintek előrejelzésénél ez a fajta hiba-halmozásos módszer megfelelő eredménnyel, szolgálhat.



6. ábra: 15 m-es vízszintváltozások összehasonlítása a különböző algoritmusok alkalmazásával, nem görgetett adatsorokkal

Fig.6: Comparison of the big water level changes whit the different algorithms using not rolled time-series, thick line: the measured values, thin lines: provided by all algorithms

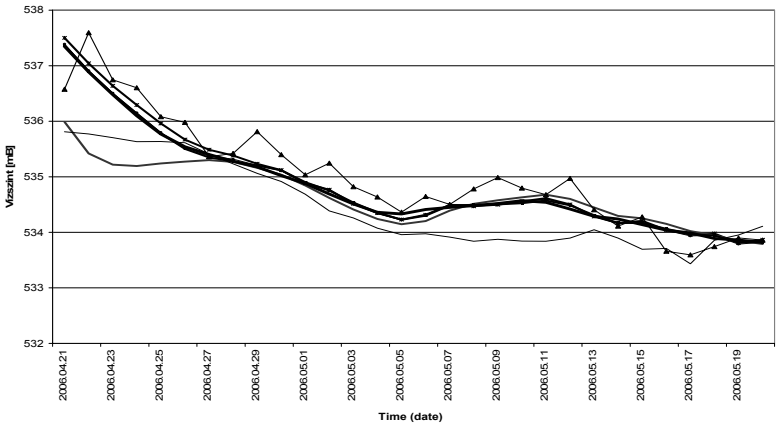
A *6. és 7. ábrán* a lehetséges öntanuló algoritmusok alkalmazásával nyert eredményeket ábrázoltuk a kiválasztott legnagyobb vízszintingadozást (15 m) mutató hónap lefuttatásával (2006.03.22 – 2006.04.20.). Az első csak a mért értékekből történő adatpótlást mutatja, a második pedig a mért értékek és a már megjósolt értékek további felhalmozásából adódó pótlást.



7. ábra: 15 m-es vízszintváltozások összehasonlítása a különböző algoritmusok alkalmazásával, görgetett adatsorokkal, vastag vonal: mért értékek, vékony vonal: az algoritmusok által szolgáltatott megoldások
Fig. 7: Comparison of the big water level changes whit the different algorithms using rolled time-series, *thick line: the measured values, thin lines: provided by all algorithms*

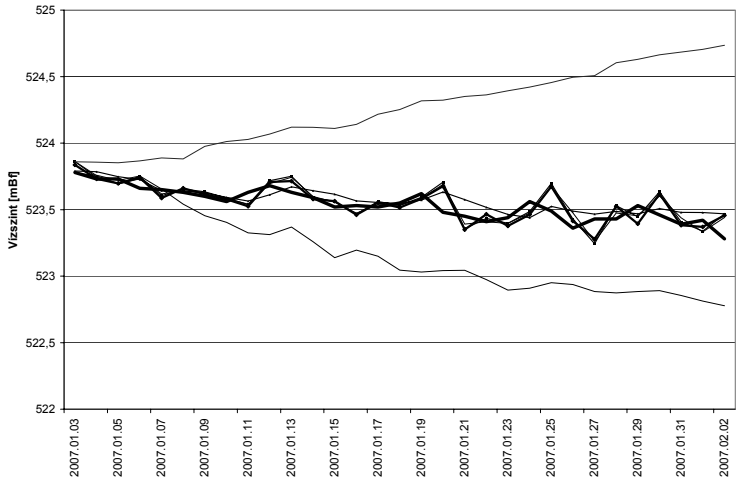
Legjobb illeszkedést az LS-SVR módszer, rbf transzformációjával kaptunk, zaj - és magas dimenziójú inputparaméterek nélkül.

A *8. ábra* szintén az összes lehetséges öntanuló algoritmus által kapott eredményt mutatja az eredeti mért adatok alkalmazásával, különböző input-paraméterek betáplálásával, azzal a különbséggel, hogy a vizsgált időszakban a vízszintváltozások 6-7 m közöttiek (2006.04.21 – 2006.05.21). Az SVR polynomiális illesztése, zaj beadásával adta a legjobb megoldást.



8. ábra: 6-7 m-es vízszintváltozások összehasonlítása a különböző algoritmusok alkalmazásával, nem görgetett adatsorokkal, vastag vonal: mért értékek, vékony vonal: az algoritmusok által szolgáltatott megoldások
Fig. 8: Comparison of the middle water level changes whit the different algorithms using not rolled time-series, thick line: the measured values, thin lines: provided by all algorithms

A két fenti adatpótlás eredményeihez hasonlóan a kis vízszintváltozású adatsor különböző algoritmusokon történő lefuttatása után a legjobb illeszkedést nyújtó módszer az SVR volt – szintén nem görgetett adatsorral *(9. ábra)* –, zaj betáplálásával.



9. ábra: 1-1,5 m-es vízszintváltozások összehasonlítása a különböző algoritmusok alkalmazásával,nem görgetett adatsorokkal, vastag vonal: mért értékek, vékony vonal: az algoritmusok által szolgáltatott megoldások
Fig.9: Comparison of the little water level changes whit the different algorithms using not rolled time-series, thick line: the measured values, thin lines: provided by all algorithms

A fentiek szerint tehát az SVR öntanuló algoritmus két különböző bemeneti információval/paraméterrel szolgáltatta a legjobb illeszkedést a három választott - különböző vízszintváltozású - adatsorra.

Érdemes kicsit részletesebben elemezni a nagy vízszintváltozású görbéknél kapott eredményt, hiszen ezen görbék közül a zaj, és magas dimenziójú paraméter betáplálása nélkül kapott görbe (függvény) adta a legjobban illeszkedő megoldást, annak ellenére, hogy a gyakorlatba azt feltételeznénk, hogy épp az ilyen nagy változásoknál szükséges a hatékony csapadékcsoportok ismerete. A másik két esetben (kicsi, illetve közepes vízszintváltozásnál) viszont a zaj input paraméterként való használatával nyertük a legszorosabb kapcsolatot a mért és pótolt értékek között. Ennek valószínűleg az az oka, hogy a tavaszi esőzések és hóolvadások hatására csaknem minden évben hasonló trendet mutatnak a vízszintértékek (nagy ugrású hónap, 2006. 03. 22 – 2006. 04. 20.), az árhullám megemelkedik. Az ezt kö-

vetkező időszakban (közepes ugrás, 2006.04.21 – 2006. 05. 20) viszont a csapadék mennyiségétől függően tovább emelkedik a karsztvízszint, vagy pedig leürül, ezért van szükség a külső tényezők ismeretére. A téli hónapban történt vizsgálatnál is hasonló indokból adhatott jobb megoldást a külső paraméterek felhasználása (2007. 01. 03 – 2007. 02. 02.), többször előfordult már ugyanis csapadékban bő, enyhe tél, amikor a talaj felső részét nem borította hó, és a víz akadály nélkül eljutott a karsztvíztárolókig. Nem kell azonban szem elől téveszteni azt a tényt sem, hogy a második legpontosabb illesztést mindhárom hónapban – a kicsi illetve közepes vízszintváltozású idősornál a különbséget csak a módszer zaj nélküli változata jelenti – az SVR módszer ugyanazon függvény-transzformációja (polynomiális illesztése) szolgáltatja, ugyanúgy hiba-paraméter hozzáadása nélkül.

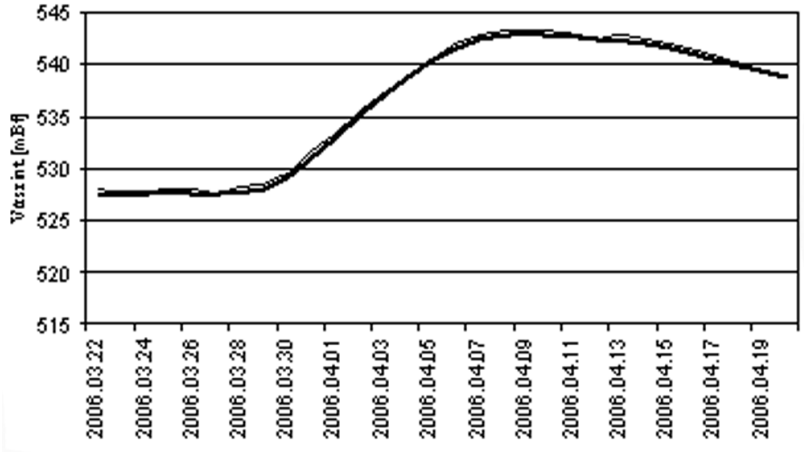
Összességében tehát megállapítható, hogy mindhárom idősornál az eredményként kapott legjobb és az utána következő hiba-érték közötti különbség 0,004-0,01 m közé esik, ami két dolgot jelent a jövőbeli pótlások alakulását tekintve. Az egyik, hogy nem feltétlenül kapunk rossz eredményt, ha nem ismerjük a csapadék- ill. éghajlati viszonyokat, vagy más megfigyelőhelyek vízszint-adatsorát. A másik pedig, ha 1 hónapnál is hosszabb időszakban kívánjuk a hiányzó értékeket pótolni – tehát előfordulhat 1-20 m közötti vízszintváltozás is -, akkor elfogadható megoldást kaphatunk, ha zaj nélküli inputokkal, csak az SVR (poly) algoritmust használjuk a hiányzó vízszintek pótlására.

7. Karsztvízszintek előrejelzésének eredményei neuronhálózatok felhasználásával

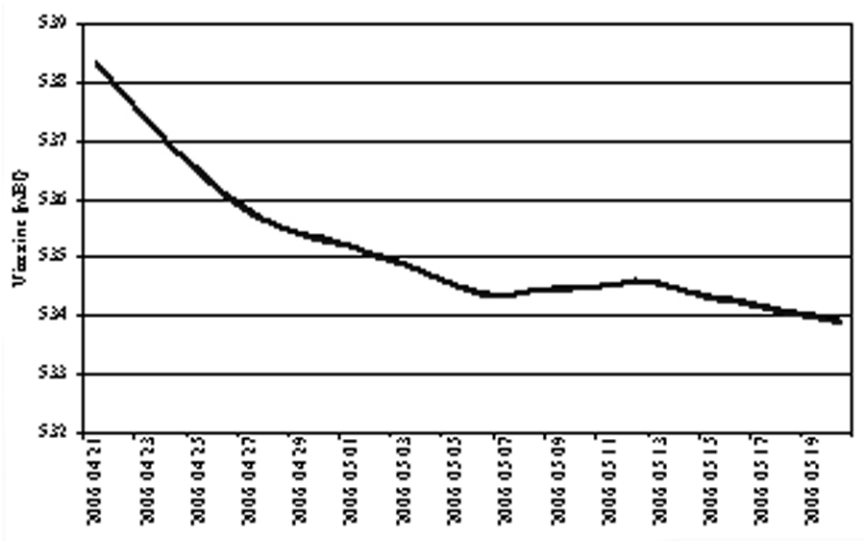
Karsztvízszint-adatsorok mesterséges intelligencián alapuló neuron hálózatokkal történő előrejelzése ugyanolyan elven működik, mint (ugyanazon) adatsorok hiányainak pótlására alkalmazott módszer, annyi különbséggel, míg az adatpótlások során egy vagy több (egyidejű) adatsor adatait is felhasználhattuk a számítások elvégzésére, addig az előrejelzések során értelemszerűen további adatsorok nem állnak rendelkezésre a feladat megoldása érdekében, hiszen azokat nem ismerjük. Ugyanakkor a transzformációs-függvények képzéséhez más közeli mérőpontok vízszint vagy csapadékadatai is felhasználhatók.

Az ANN és SVR (rbf és polinomos illesztés) mesterséges intelligencián alapuló algoritmusokkal kapott eredmények közül előrejelzésnél a tréning-sorozaton történő összehasonlítás után szintén az SVR módszer, rbf (rule based forecast/meghatározott szabályokon alapuló előrejelzés) típusú illesztése nyújtotta a legpontosabb illesztéseket, annyi különbséggel, hogy a

kiválasztott mindhárom hónapnál külső faktor betáplálásával történt a jövőbeli adatok becslése. Ennek főleg az lehet az oka, hogy előrejelzésnél jövőbeli értékeket nem ismerünk, ezért lehet szükség egy másik (hasonló) adatsor felhasználására is.

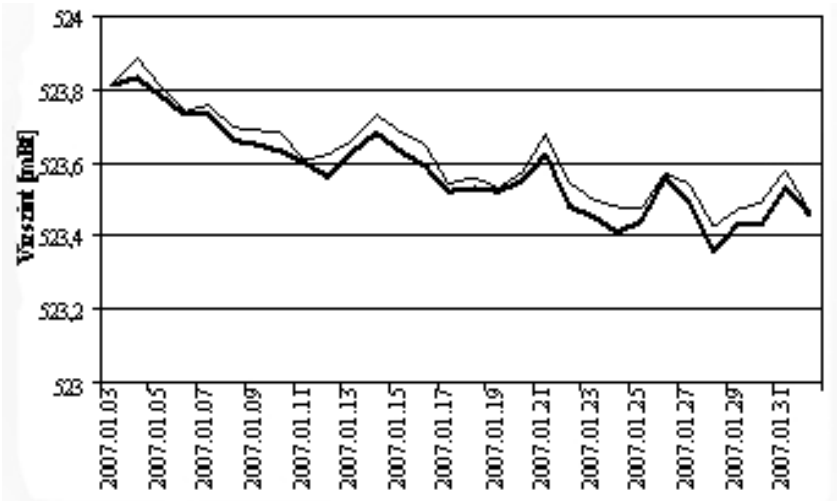


10. ábra: Előrejelzés LS-SVR módszer, rbf típusú illesztésével, zaj betáplálásával, nagy vízszintváltozású adatsoron, vastag vonal: mért értékek, vékony vonal: előre jelzett értékek
Fig. 10: Forecast of the big water level changes of LS-SVR (rbf) method using noise, thick line: measured values, thin line: the forecasted



11. ábra: Előrejelzés LS-SVR módszer, rbf típusú illesztésével, zaj betáplálásával, közepes vízszintváltozású adatsoron, vastag vonal: mért értékek, vékony vonal: előre jelzett értékek
Fig. 11: Forecast of the middle water level changes of LS-SVR (rbf) method using noise, thick line: measured values, thin line: the forecasted

A *10-12-es ábrák* a legpontosabb illesztések eredményeit mutatják a három különböző vízszintváltozás esetén. Mindhárom ábrán nagyon szépen követik egymást a becsült illetve a valódi értékek görbéi, mind a kisebb, nagyobb vízszintváltozások „*játékai*” megmutatkoznak. Hidrogeológiai szempontból mindenképpen elfogadható a közelítéssel szolgál az SVR öntanuló algoritmus, rbf típusú illesztése.



12. ábra: Előrejelzés LS-SVR módszer, rbf típusú illesztésével, zaj betáplálásával, kicsi vízszintváltozású adatsoron , vastag vonal: mért értékek, vékony vonal: előre jelzett értékek
Fig. 12.: Forecast of the little water level changes of LS-SVR (rbf) method using noise, thick line: measured values, thin line: the forecasted

Összefoglalás, javaslatok

A regresszió módszer alkalmazásával az adathiányok majdnem minden esetben pótolhatóak, vagy egy ugyanazon évben jelenlévő másik észlelőhely/észlelőhelyek vízszintadatai segítségével, vagy pedig az ugyanazon észlelőhely másik éveiben, ugyanerre az időszakra ismert értékei alapján. A valós, mért és „*becsült*” (számított) értékek közötti eltérés a legtöbb esetben cm-es nagyságrendűnek adódott, előfordult azonban dm-es nagyságrendű különbség is. Az Nv-17 nagyon sok esetben jó viszonyítási alapul szolgál, számos helyen – adathiányt kivéve – jó korrelációt mutat a vízszint-adatsorok nagy részével.

A neuron hálózaton alapuló megoldás eredményei alapján megállapítható, hogy karsztvízszint-adatsorok pótlására illetve előrejelzésére neuron hálózatok alkalmazása megfelelő megoldással szolgált, mind adatpótlásnál, mind előrejelzésnél a SVR (Support Vector Regression), rbf (rule based forecast) típusú illesztése adta a legpontosabb vízszintértékeket, azzal a kü-

lönbséggel – amit sejteni is lehetett –, hogy adatpótlásnál külső faktor alkalmazása nélkül, előrejelzésnél pedig külső faktor alkalmazásával nyertük a legjobb illesztéseket. Valószínűleg ennek az oka, hogy míg pótlásnál a vízszint-adatsoron „*korábbi és későbbi*" információk is rendelkezésünkre állnak, addig előrejelzésnél nem ismerjük az idősorok jövőbeli alakulást.

A vízszint idősorok pótlására, előrejelzésére mindkét módszer megoldása elfogadható –a valós vízszintértékekhez képest igen közeli – vízszintértékeket nyújtott. Regressziós vizsgálatnál cm-es és dm-es – de előfordult, hogy m-es - különbségek adódtak a mért és számított vízszintek között; az öntanuló algoritmusok által kapott megoldások pedig – nagy vízszintváltozások előrejelzésétől eltekintve, ami 2 m-es eltérést is mutatott – 20 cm alatti hibabecsléssel közelítették a karsztvízszinteket. Tehát megállapíthatjuk, hogy ezen vizsgálatok eredményei is alátámasztják a bükki karsztrendszer összefüggő voltát.

Bár hidrogeológiai szempontból a jövőben mindkét eljárással megfelelő pontossággal tudjuk becsülni a hiányzó vízszintértékeket, és előre jelezni a nem kívánatos csapadék okozta vízszintváltozásokat, a továbbiakban a matematikai alapon történő, öntanuló algoritmusok által kapott eredmények felhasználását javasoljuk. Bár igaz, hogy az idősorok futtatásához időre van szükség, de a későbbiekben hosszú távú előrejelzésekhez és pótlásokhoz mindenképpen pontosabb megoldás nyújthat.

IRODALOM

LÉNÁRT, L. (2005): Some aspects of the „3E's" (Economics-Environment-Ethics) model for sustainable water usage in the transboundary Slovakian and Aggtelek karst region based on some examples from the Bükk Mountains. - PhD disszertáció, Kassa/Kosice.

LÉNÁRT, L. (2008): A Bükkben keletkezett kitermelhető karsztvízkészlet folyamatos meghatározásának módszere XVII. - Miskolci Vízművek zRt, Észak-magyarországi Regionális Vízművek zRt, Heves megyei Vízművek zRt., Mezőkövesdi VG zRt, Smaragd GSH-Kft, Miskolci Egyetem megbízás.

SZARVAS, GY. (2003): Mesterséges neuronhálók felhasználása idősorok előrejelzésében. - SZTE, Informatikai Tanszékcsoport.

JUHOS, I. – SZARVAS, GY. (2003): Intelligent forecast with Dimension Reduction, Dep. of Computer Algorithms and Artificial Intelligence - Uni of SZ & Research Group on AI, Hungarian Acad. of Science