

PERIÓDUS VIZSGÁLAT: ALKALMAZÁSI FELTÉTELEK ÉS NEHÉZSÉGEK

**KOVÁCS JÓZSEF¹ – KOVÁCSNÉ SZÉKELY ILONA² –
KÖNCZÖL ANDRÁS³**

¹ELTE, Földrajz- és Földtudományi Intézet, 1116 Budapest PázmányPéter sé-
tány1/C, kevesolt@geology.elte.hu

²BGF, KVIFK, Módszertani Intézet, 1054 Budapest, Alkotmány u. 9-11,
iszekely@geology.elte.hu

³REPÉT Környezetvédelmi Kft., 1118 Budapest, Brassó út 169-179. G. ép.,
repet.metallo@t-online.hu

Abstract: At present stage there are certain methods for time period estimation which can also be applied for time series sampled in a non equidistance manner. With the help of these methods one can estimate at a certain significance level the time period of the given time series. If the time series is not stationary, it is useful to remove different trends. Then, one should consider those results which are not depending on the removal of trends and which give the same result.

Bevezetés

Számos természeti folyamat periodikus tulajdonságokat mutat. A periódus idő becslésére több módszer ismert. Ezek közül a legegyszerűbbek szubjektív elemeket is hordoznak, mivel mozgóátlagolással történő simítás után a vizsgált paraméter vizuális megjelenítése alapján becsülhető a periódus idő (*REIMANN–V. NAGY* 1984). Vannak objektív módszerek is, ezek régóta ismertek és folyamatosan javulnak. Az új eljárások két irányban javítottak az eddigi alkalmazási lehetőségeken. Egyrészt, a szubjektív elemeket háttérbe szorították, másrészt mivel nem minden esetben biztosítható az adott idősor azonos időközönkénti megfigyelése, olyan új eljárás is kidolgozásra került, ami ennek a szigorú feltételnek a létét nem követeli meg.

Egy modern periodogram becslési eljárás

Gyakran fellépő igény idősorok vizsgálata során, hogy milyen periodikus változások lehettek fel az adott folyamatban. A válasz spektrálemeléssel adható

meg. A spektrum fogalom Newtonnál jelent meg, amikor a fehér fényt színes alkotóelemeire bontotta. Egy periodikus determinisztikus függvény különböző frekvenciájú sinus-és cosinusfüggvények végtelen sok elemből álló lineáris kombinációjára bontható fel. A spektrálanalízis ennek a felbontásnak egy általánosítását kínálja sztochasztikus folyamatokra. (TUSNÁDY–ZIERMANN 1986) Gyakorlatban egy idősor spektrális összetevői az úgynevezett periodogrammal becsülhetők, amely szinte bármely statisztikai software-csomagban elérhető. A számítások elvégzéséhez szükséges feltétel, hogy a vizsgált folyamatról egyenletes időközönként rendelkezünk mérési adattal. Ezt a földtudományi szakterületen roppant nehéz biztosítani. Gondoljunk egy 50 éve folyamatosan, heti rendszerességgel ugyanazon a napon mért karsztvízszint idősorra. További nehézséget jelent a kapott eredmény kiértékelése. Ugyanis a kapott periodogram nem teszi lehetővé, hogy a periódus idő megállapítása mellé megbízhatósági szintet rendeljünk, mert a tradicionális teljesítménysűrűség becslések [power spectral density (psd), estimations] során kapott eredménynek nem határozható meg az eloszlása.

A közelmúltban kidolgozott Lomb-Scargle periodogram (LOMB 1976, PRESS 1992), olyan psd becslő eljárás, amelynek teljesítményértékeihez egzakt módon megadható az eloszlás típusa, így konfidencia intervallum is számítható, továbbá pontosan megadható az is, hogy mely periódusok jelentősek a vizsgált idősorban egy adott szignifikancia szinten.

Periodogram becslést a Diszkrét Fourier Transzformációra (DFT):

$$DFT(\omega) = \sum_{n=1}^N x(t_n) e^{-i\omega t_n}$$

alapozva is lehet definiálni, de a kapott periodogram nem invariáns az időbeli eltolással szemben. Ezért Lomb módosította a periodogram definícióját a következő módon:

$$P_N(\omega) \equiv \frac{1}{2\delta^2} \left[\frac{\left[\sum_j (y_j - \bar{y}) \cos \omega_i (t_j - \tau_i) \right]^2}{\sum_j \cos^2 \omega_i (t_j - \tau_i)} + \frac{\left[\sum_j (y_j - \bar{y}) \sin \omega_i (t_j - \tau_i) \right]^2}{\sum_j \sin^2 \omega_i (t_j - \tau_i)} \right]$$

ahol az átlag és variancia a szokásos módon kerül kiszámolásra:

$$\bar{y} \equiv \frac{1}{N} \sum_{i=1}^N y_i$$

$$\delta^2 \equiv \frac{1}{N-1} \sum_{i=1}^N (y_i - \bar{y})^2$$

és az i -edik szögfrekvencia érték: $\omega_i = 2\pi f_i$. A τ_i konstans a következő módon számítható:

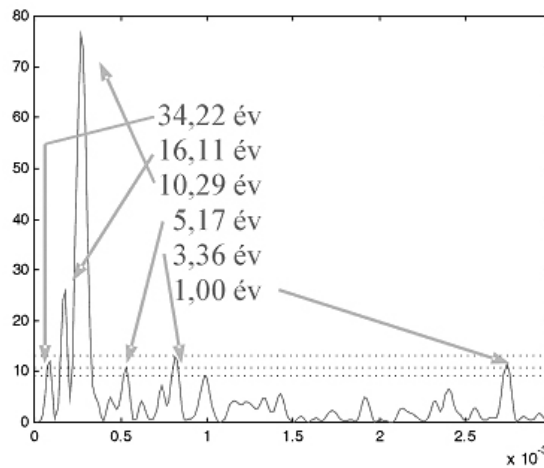
$$\tau_i \equiv \frac{1}{2\omega_i} \arctan \left[\frac{\sum_j \sin 2\omega_i t_j}{\sum_j \cos 2\omega_i t_j} \right]$$

A Lomb-Scargle periodogram számítása ekvivalens az ω szögfrekvenciájú szinusz jel — legkisebb négyzetek elvével történő — illesztésével. A számításokhoz PRESS et al. (1992) gyors algoritmusát használtuk. A Lomb-Scargle eljárásnak további előnye, hogy a nullhipotézisnek megfelelő eloszlás egzakt módon megadható, így konfidencia intervallum adható meg a teljesítmény értékekhez.

$$\text{Prob}(P_N > z) = 1 - (1 - e^{-z})^M$$

azaz a P_N teljesítmény exponenciális eloszlásból származtatható. A független frekvencia összetevők száma (M), jó közelítéssel megegyezik az adatok számával (N).

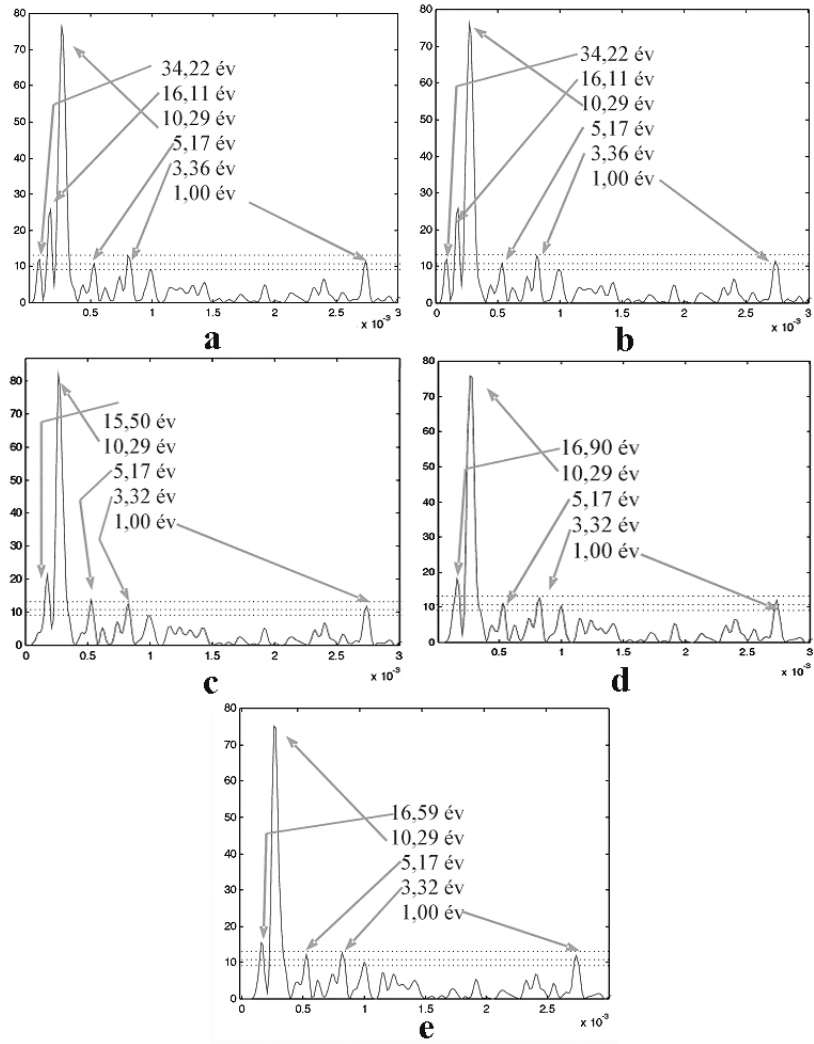
Példák a karsztvízszint idősorok periódusainak vizsgálatára



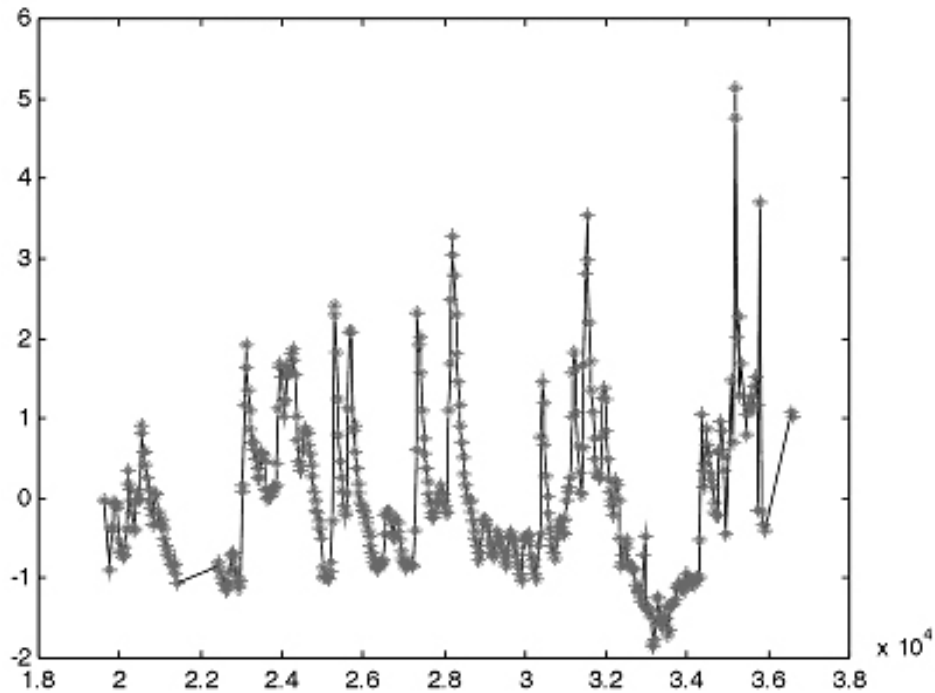
I. ábra: Periodogram trendfüggvény illesztése nélkül
 Fig. 1: Periodogram after the removal of the first order trend, calculated on its residual

A Dunántúli-középhegységben, a Balatonfelvidéken elhelyezkedő Balatonfüred-1 megfigyelő kútban 1953.09.-2000.04. között mérték a karsztvízszintet. A Lomb-Scargle módszerével elvégzett periódusvizsgálatok eredményeit az *I, 2.a,b,c,d,e. ábrák* mutatják be. A vízszintes tengelyeken az egyes csúcsokhoz tartozó frekvencia határozható meg. Azokat az értékeket kell figyelembe venni, amelyek a felhasználó által megszabott szignifikancia szint felett vannak. A mintavételezési gyakoriságot természetesen figyelembe kell venni a periódusidő kiszámításánál. Az ábrákon három féle (5%, 2%, 1%) szignifikancia szint mellett a feltüntetett periódusidők fogadhatók el. Mindegyik periódus becslés ugyanarra az idősorra készült. A különbség annyi, hogy a mért időszori realizáció nem stacionárius, trendet tartalmaz, aminek eltávolítása különböző fokszámú polinomokkal történt. Az így kapott maradékokra különböző periódus értékek becsülhetők. Amint látható, az 1; 3,36; 5,17 és 10,29 év periódus idejű folyamatok mindegyik esetben megtalálhatók voltak, míg a hosszabb periódus idők az illesztett polinom fokszámának növelésével módosultak illetve eltűntek.

Az eredmény kapcsán több kérdés is felmerül. Az első, hogy miért kell trendet levenni? A választ az adja meg, hogy a trend jelenlétét a periódus számítás nagy hullámhosszú folyamatként érzékeli, ebből következően – szélsőséges esetben - becsülhető olyan periódus hossz, ami nagyobb, mint a realizáció hossza (példaként lásd *I. táblázat* adatait)! Esetünkben a 34,2 év ilyennek tekinthető. Az idősoros vizsgálatokkal foglalkozó matematikai irodalom éppen ezért azokat a periódus időket tartja elfogadhatónak, amelyek a körülbelül ötször *“beleférnek”* a megfigyelési idő tartományába. Ez hidrogeológiai szakterületen, a hosszú periódusok meghatározásánál szinte teljesíthetetlen feladat. Egy 11 éves periódus megbízható kimutatásához ebben az esetben már legalább 55 évre lenne szükség. Ezért esetünkben ezt a szigorú feltételt nehéz, vagy esetek túlnyomó részében lehetetlen betartani.

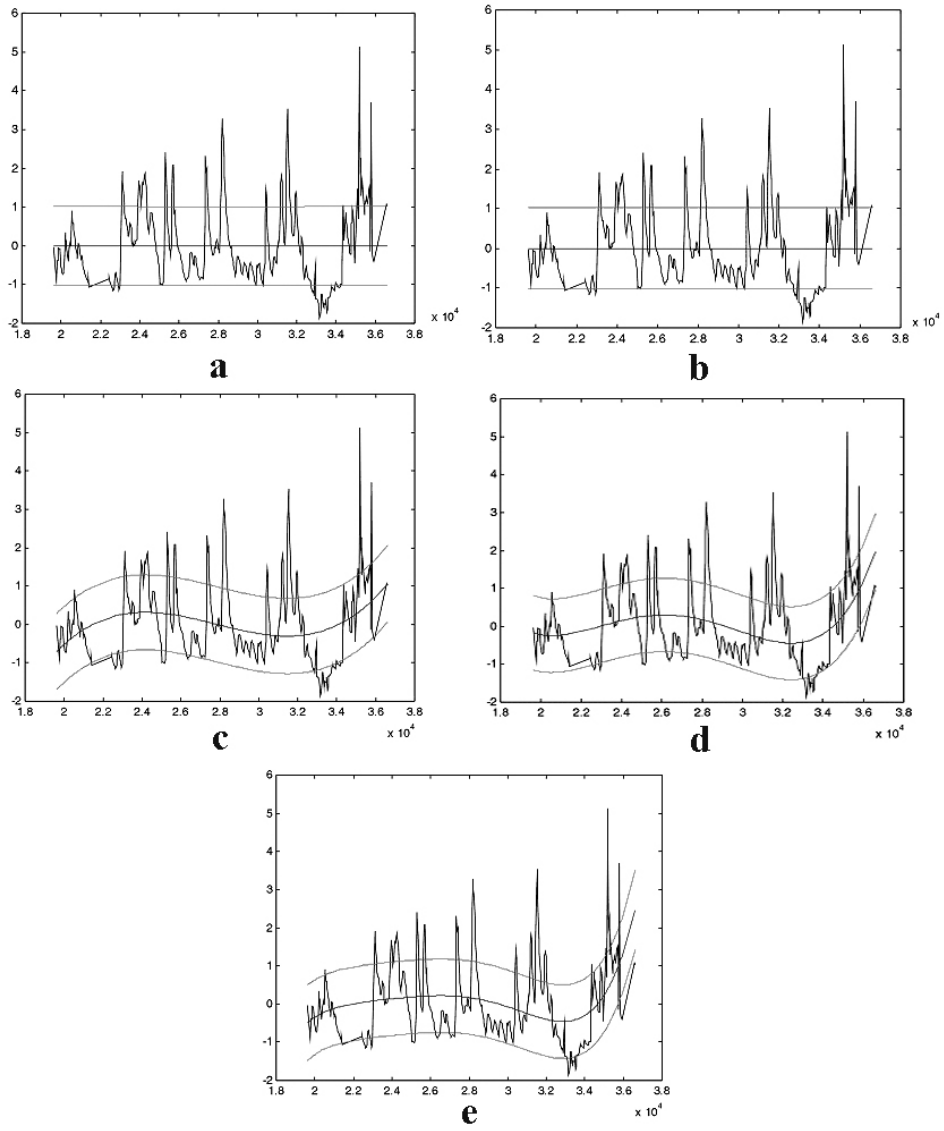


2. ábra: Periodogram a. I. fokú-, b. II. fokú-, c. III. fokú-, d. IV. fokú-, e. V. fokú trend maradékaira
 Fig. 2: Periodogram calculated on the residual of the a. first order- b. second order-, c. third order-, d. fourth order-, e. fifth order trend

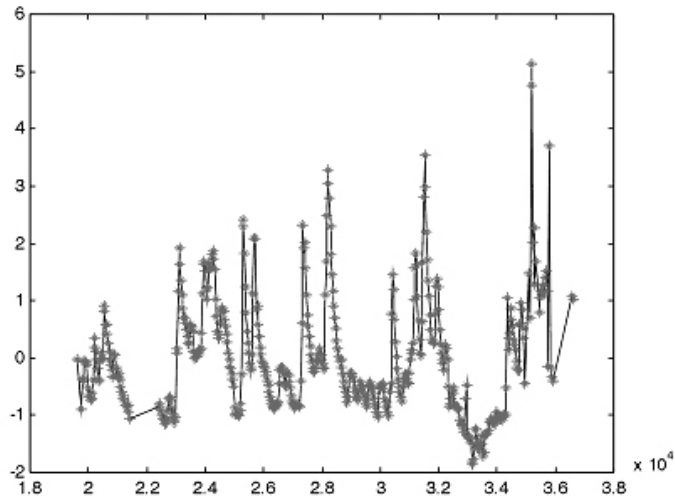


3. ábra: A mért vízszint adatok
 Fig. 3: The measured water levels

A második kérdés, hogy hányad fokú trendet kell eltávolítani? Az a tapasztalat, hogy a trend levétele nyomán némely periódusok eltűnnek, és újak jönnek „elő”. Ahhoz hogy mi is történik, tekintsük meg a 3, 4a,b,c,d,e, 5, 6.a,b,c,d,e. ábrsorozatot. A 3. ábra mutatja be, csillaggal jelölve, a mért értékek normált értékeit. (Ez a transzformáció a vizsgálatok helyességét nem érinti, viszont segít különböző nagyságrendű értékek egy skálára hozásában, ezért összehasonlítható ábrák készíthetők. A későbbiekben ezeket mint mért értékeket említjük.) A továbbiakban, a 4.a,b,c,d,e. ábrákon láthatóak azok a polinomiális trendek amelyeket a mért adatok értékeire illesztettünk az általunk választott 95% – os megbízhatósági szint mellett. A 5, 6.a,b,c,d,e. ábrák a maradékokat – az illesztett trendek és a mért értékek különbségeit – tüntetik fel.



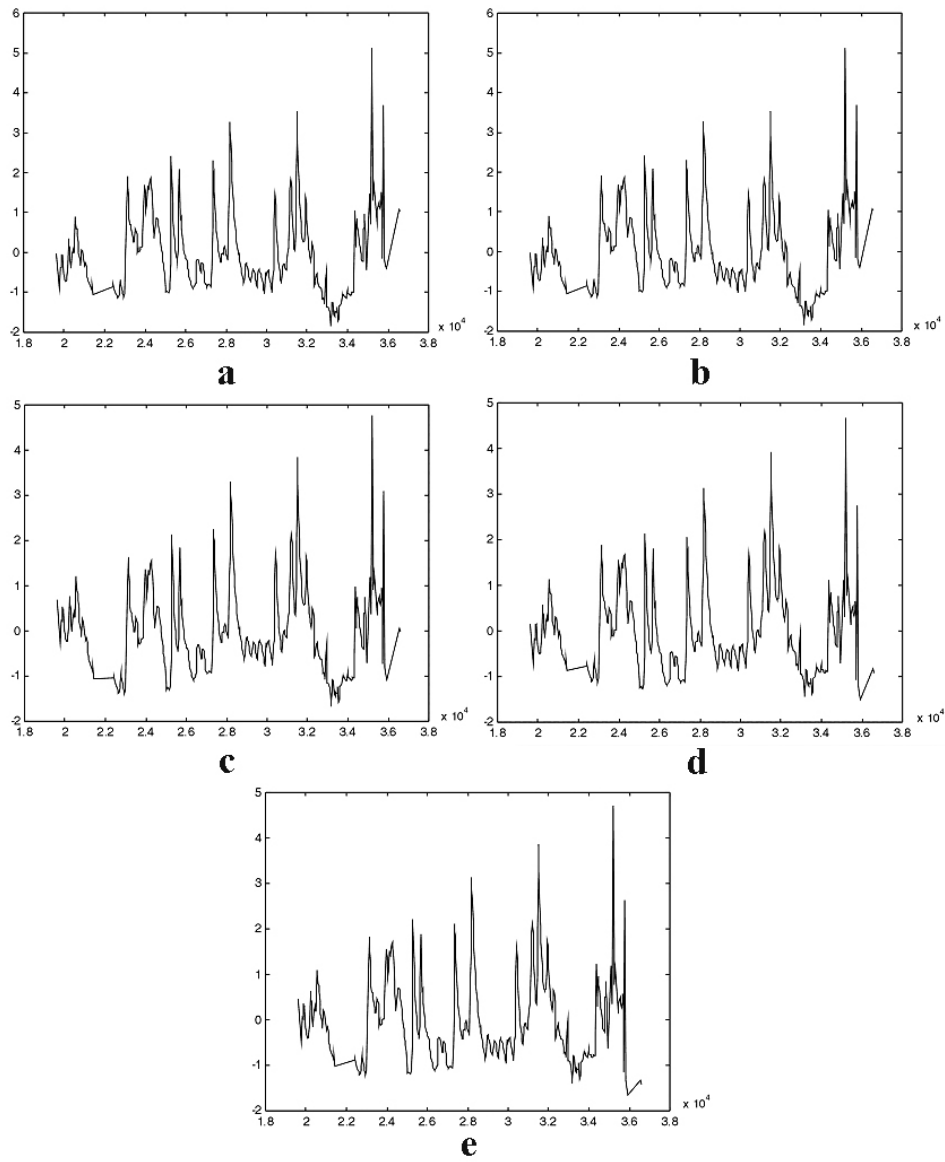
4. ábra: a. I. fokú-, b. II. fokú-, c. III. fokú-, d. IV. fokú-, e. V. fokú illesztett trend és a 95%-os konfidencia sávja
 Fig. 4: Estimated a. first order-, b. second order-, c. third order-, d. fourth order-, e. fifth order trend and its 95% confidence band



5. ábra: A mért vízszint adatok
Fig. 5: The measured water levels

A Balatonfüred-1. kút hidrográfja esetén az I. és II. illetve a III. – V. fokú trendek nem hoztak lényegesen különböző eredményeket. Egy olyan idősor esetén, ahol a nagymértékű változások, például folyamatos vízszint süllyedés tapasztalható a trend eltávolítása nyomán olyan periódusok válnak jelentőssé, amelyek a mért realizációban szinte észrevehetetlenek voltak. Ebből következik, hogy jelentős eltérések lehetnek a periódus idő becslésében. Erre mutat példát az 1964-2000 között mért Solymár-88 kút *I. táblázat*. Szemléletesen látható, hogy a trend miatt nagyobb a becsült periódus idő a mért adatokra, mint a mintavételezési idő tartománya. Az is látható, hogy a polinomiális trend fokszámának növelésével képzett maradékoknál a nagy periódus időket kivettük a jelből, így azok fokozatosan tűnnek el.

Felmerül a kérdés: hányad fokú trend levétele indokolt ahhoz, hogy a periódus idő becslése a lehető legjobban tükrözze a valóságot? Ennek megválaszolására egyetlen lehetőséget látunk. Statisztikailag elegendően nagy számú idősor esetén, el kell végezni a periódus idő becslését, különböző fokú trend függvények esetére. A meghatározásra került szignifikáns periódusok közül azok tekinthetők a valósághoz a legközelebb állóknak, amelyek a leghasonlóbb eredményt adják a trendeltávolítások függvényében. Természetesen ez a módszer csak akkor alkalmazható, egyrészt ha van kellő számú idősor a területről, másrészt mindegyik ugyanarról a jelenségről hordoz információt.



6. ábra: a. I. fokú-, b. II. fokú-, c. III. fokú-, d. IV. fokú-, e. V. fokú trend levétel utáni maradék
 Fig. 6: Residual of a. first order- b. second order-, c. third order-, d. fourth order-, e. fifth order trend after its removal

I. táblázat

Table I.

Különböző fokú trendfüggvények alkalmazása esetén becsült periódus idők, a Solymár-88 karsztvízszint megfigyelő kút idősorára

Estimated time periods by the application of different order trends. The time series were supplied by the Solymár-88 karst water level observation

<i>Polinomiális trend foka</i>	<i>A maradék adatokra becsült periódus [év]</i>
0. fokú trend	49,78
	18,08
1. fokú trend	19,56
	11,16
2. fokú trend	20,74
	11,16
	8,53
3. fokú trend	20,69
	11,15
	8,37
4. fokú trend	14,49
	7,94
	6,08
	5,18
	3,92
5. fokú trend	13,36
	8,30
	6,04
	5,18
6. fokú trend	3,86
	11,16
	7,63
	5,80
	4,42
	3,82

Összefoglalás

Napjainkban vannak olyan periódus becslő eljárások, amelyek időben nem egyenközűen mintavételezett idősorokra is alkalmazhatóak és segítségével adott valószínűségi szinten becsülhető az adott folyamat periódus ideje. Abban az esetben, ha az idősor nem stacionárius, célszerű többféle trendet eltávolítani, majd azokat az eredményeket kell figyelembe venni, amelyek a trendlevételtől függetlenek és ugyanazokat az eredményeket adják. Fontos megjegyezni, hogy a periódusok és a trendek megállapítása a múlt időszakokra érvényesek. Az eredményekből a jövőre vonatkozóan csak akkor vonhatunk le következtetéseket, ha az idősort befolyásoló tényezők fennállnak.

IRODALOM

LOMB, N.R. (1976): Least squares frequency analysis of unevenly spaced data - *Astronomical and Space Science*. 39., p. 447-462.

PRESS, W.H. – TEUKOLSKY, S.A. – VETTERLING, W.T. – FLANNERY, B.P. (2002): *Numerical recipes in C: The art of scientific computing*, Cambridge University Press, 1032 p.

REIMANN J. – V. NAGY I. (1984): *Hidrológiai statisztika*, Tankönyvkiadó, Budapest, 518 p.

TUSNÁDY G. – ZIERMANN M. (1986): *Idősorok analízise*, - Műszaki Könyvkiadó, Budapest, 340 p.