



Budapesti Műszaki és Gazdaságtudományi Egyetem
Természettudományi Kar

DIPLOMAMUNKA

Impedancia jelek fluktuációanalízise:
Sejtek mikromigrációjának matematikai és
biológiai karakterizálása ECIS
(Electric Cell-Substrate Impedance Sensing)
készülékben

Készítette

Schrempf Dóra

Témavezető

Dr. Láng Orsolya

Semmelweis Egyetem,

Genetikai, Sejt- és Immunbiológiai Intézet

Konzulens

Dr. Farkas Lóránt

Budapesti Műszaki és Gazdaságtudományi Egyetem,

Analízis Tanszék

2018.

Tartalomjegyzék

Rövidítések jegyzéke	3
1. Bevezetés	4
2. Irodalmi áttekintés	6
2.1. ECIS készülék és teória	6
2.2. Zajanalitikai módszerek	8
2.2.1. Variancia számítás	8
2.2.2. Standard deviáció számítás	9
2.2.3. Inkrement számítás	9
2.2.4. Teljesítménysűrűség-spektrum	10
2.2.5. Hurst-analízis	11
2.3. A Hurst exponens története	12
2.4. A Hurst exponens alkalmazásai	13
2.4.1. Pénzügyben	13
2.4.2. Szívverés és érzelmi minőség kapcsolatában	14
2.4.3. Csapadék mennyiségének mérésében	15
2.4.4. Geomagnetikus mezőben	15
2.4.5. Egyéb alkalmazásai	15
2.5. A Hurst exponens kiszámítási módszerei	16
2.5.1. R/S-analízis algoritmusa	16
2.5.2. Anis és Lloyd módszere	17
2.5.3. GM1 módszer	18
2.5.4. GM2 módszer	18
3. Célkitűzés	20
4. Anyag és módszer	21
4.1. Adatsorok előállítása	21
4.1.1. Generált adatsorok	21
4.1.2. Biológiai adatsorok	21
4.1.2.1. Sejtek és tenyésztésük	21

4.1.2.2.	Mikromigráció mérés ECIS készülékben	22
4.1.2.3.	Kezelés és vizsgált vegyületek	23
4.2.	Software-es rutinok	23
4.2.1.	Variancia, inkrement kiszámolására	24
4.2.2.	Teljesítménysűrűség-spektrum kiszámolására	25
4.2.3.	Hurst exponens meghatározására	26
4.3.	Statisztika	26
4.3.1.	Normalitás teszt	27
4.3.1.1.	Shapiro-Wilk teszt	27
4.3.2.	F-próba	28
4.3.3.	Kétmintás t-próba	28
5.	Eredmények	30
5.1.	Számítógéppel generált adatsorok	30
5.2.	Biológiai adatsorok	34
5.3.	Sejtmentes minták	34
5.3.1.	Sejtes minták	36
5.4.	Sejtes és sejtmentes adatsorok statisztikai összehasonlítása	37
5.5.	Tumorelleses szerekkel kezelt sejtek mikromigrációja	38
6.	Megbeszélés	39
7.	Következtetések	41
8.	Összefoglalás	43
	Függelék	45
	Irodalomjegyzék	56

Rövidítések jegyzéke

ECIS - Electric Cell-Substrate Impedance Sensing
H - Hurst exponens
Z - Impedancia
R - Rezisztancia
C - Kapacitív ellenállás
EKG - Elektrokardiográfia
 $V_{fázis}$ - Fázis feszültsége
 $V_{ellenfázis}$ - Ellenfázis feszültsége
I - Áramerősség
D - Fraktáldimenzió
FFT - Gyors Fourier-transzformáció
DFT - Diszkrét Fourier-transzformáció
RR - Újraskálázott terjedelelem-analízis, Rescaled Range Analysis
MRS - Módosított újraskálázott terjedelelem-analízis
MLE - Maximum Likelihood becslés
AVT - Aggregált eltérési idő módszer
GM1 - Első geometriai módszer a Hurst exponens kiszámítására
GM2 - Második geometriai módszer a Hurst exponens kiszámítására

1. fejezet

Bevezetés

Gimnáziumi tanulmányaim során a matematika mellett az egyik kedvenc tantárgyam a biológia volt. Már akkor is nagyon érdekesnek találtam az emberi test felépítését és működését. Sajnos, mint oly sokan mások, én is megtapasztaltam közvetlen környezetemben, hogy milyen nagy baj, ha egy betegség megtámadja az emberi test tökéletes rendszerét. Így nem volt kérdés, hogy amennyiben lehetőségem adódik, tanulmányaimat az orvostudomány segítésére használjam.

Diplomamunkám egy három szemeszteres kutató munka eredménye, rengeteg sejt kísérlettel, melyek kiértékelésében és az eredmények értelmezésében nagy segítségünkre volt a matematika tudománya. A tumor-sejtek vizsgálatára ebben a kutatásban a legfontosabb eszközünk az úgynevezett ECIS (Electric Cell-Substrate Impedance Sensing) készülék volt, amely az idő függvényében, váltóáram alkalmazása mellett képes a sejtek rezisztencia értékét detektálni szigetelő tulajdonságuknak köszönhetően. Ezen értékek tulajdonságai szerint kerestünk magyarázatot a grafikonban jól látható fluktuációra, a grafikon "szöszösségére", azaz az adatok zajjal való terhelésére, hiszen a zaj tartalmazhat információt a biológiai jelenségekről, mely összefüggésben van az ECIS készülék elektródáját borító sejtek mozgásával, az ún. mikromigrációval. Lo és munkatársai kimutatták, hogy I. Giaver és C. R. Keese által a XX. században megalkotott ún. Electric Cell-Substrate Impedance Sensing (ECIS) készülék által mért fluktuáció, nagy különbséget mutat daganatos és egészséges sejtek között. A mérés során a gépben található arany elektródok felületén kitapadó sejtek morfológiai változásai, adhéziója és migrációja is valós időben nyomon követhető. A készülékben külön mérhető az impedancia (Z), a rezisztencia (R) és a kapacitív ellenállás (C), mely értékeket az idő függvényében kapunk meg.

A kutató munka legfőbb célja a Semmelweis Egyetem Genetikai, Sejt- és Immunbiológiai Intézet kutatócsoportja segítségével végzett mérések során kapott idősorok jellemzésére alkalmas fontos paraméterek vizsgálata volt, úgy, mint a szórás, vari-

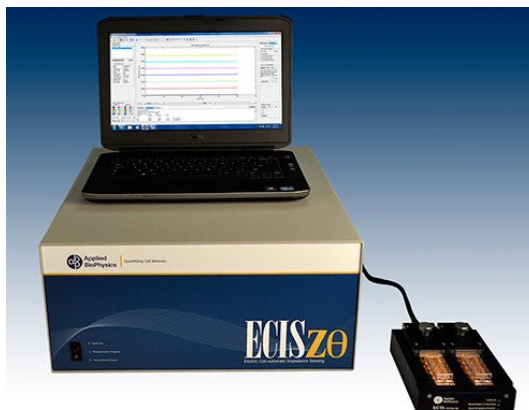
ancia, inkrement (növekmény), teljesítménysűrűség-spektrum és a Hurst exponens. Ezek közül számunkra a Hurst exponens volt a legérdekesebb, mivel korábbi vizsgálatok nem tudták egyértelműen bizonyítani az alkalmazhatóságát mikromigráció eredmények értékelésében, illetve a sejtes és sejtmentes minták elkülönítésében. Annak ellenére, hogy számos területen, például a kardiológiában (EKG), hidrológiában, pénzügyben és a meteorológiában is hasznos paraméternek bizonyult. Az irodalomban megtalálható a Hurst exponens kiszámítására alkalmas módszerek közül ötöt vizsgáltunk, és hasonlítottunk össze, majd ezek MATLAB-ban való beprogramozása után EXCEL fájlban gyűjtöttük össze a kapott eredményeket. Az egyes számítási módszerek alkalmazásakor akadályokba ütköztünk, míg végül megtaláltuk a számunkra legmegfelelőbb módszert a Hurst exponens becslésére: az általunk GM2-nek nevezett módszert. Feladatom a sejtmentes és sejtes minták vizsgálatán túl, a tumor ellenes vegyülettel kezelt sejt mintákból származó eredmények elemzése és tanulmányozása is volt. Végül nem volt más dolgunk, mint a matematikai statisztika segítségével megvizsgálni, hogy a GM2 módszerrel meghatározott Hurst exponens ismeretében eldönthető-e egy adatsorról, hogy az sejtes vagy sejtmentes mintából származik-e, illetve, hogy a kezelt minták esetén milyen változás figyelhető meg az adatsorokban.

2. fejezet

Irodalmi áttekintés

2.1. ECIS készülék és teória

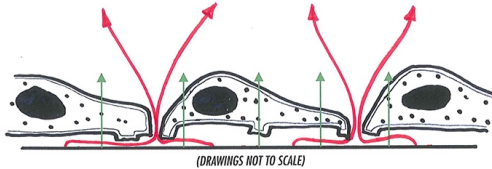
Az ECIS (Electric Cell-Substrate Impedance Sensing) egy valós idejű, elektromos impedancia mérésen alapuló bioszenzor, amely képes folyamatosan követni a hozzá csatlakozó arany elektródok felszínén lévő sejtek viselkedését, tevékenységeit. Ez utóbbi közé tartozik a sejtek morfológiai változásai, illetve a sejtek vertikális irányú mozgásai. A mérés elvét, illetve az első készüléket Ivar Giaever és Charles R. Keese [1] alkották meg (2.1 ábra).



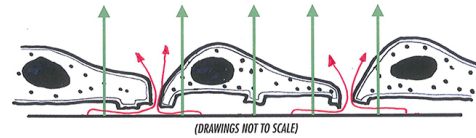
2.1. ábra. ECIS készülék

Az ECIS készülék impedancia spektroszkópia elvén működik. Az áramkör táplálásáról egy váltóáramú generátor gondoskodik, mely különböző frekvenciájú (25 Hz-től 100 kHz), 1V feszültség amplitúdójú jelet szolgáltat [2]. A váltóáram amplitúdója elég kicsi ahhoz, hogy a sejtekben ne hozzon létre detektálható biológiai változásokat. A készülék lelke egy olyan speciális plate, melynek alját arany elektródák borítják. Ebbe pipettázzuk a sejteket, és amikor a sejtek letapadnak, és elterülnek az elektródákon, akkor a váltóáram ellenállása megváltozik, hiszen a sejtek hidrofób sejtmembránjuknak köszönhetően szigetelő részecskékként viselkednek elektromos térben. Az impedancia érték változása fontos információt mond el a sejtek számá-

ról, méretéről, alakjáról és mozgásáról. A sejtek borította elektródákon az áram áthaladása a különböző frekvenciákon eltérő módon történik. Viszonylag alacsony frekvencia esetén (< 2000 Hz) az áram nagy része a szomszédos sejtek alatt és között folyik. Míg magasabb frekvencia esetén (> 40000 Hz) az áram nagy része közvetlenül a szigetelő sejt membránokon keresztül folyik [3].



2.2. ábra. Alacsony frekvencián

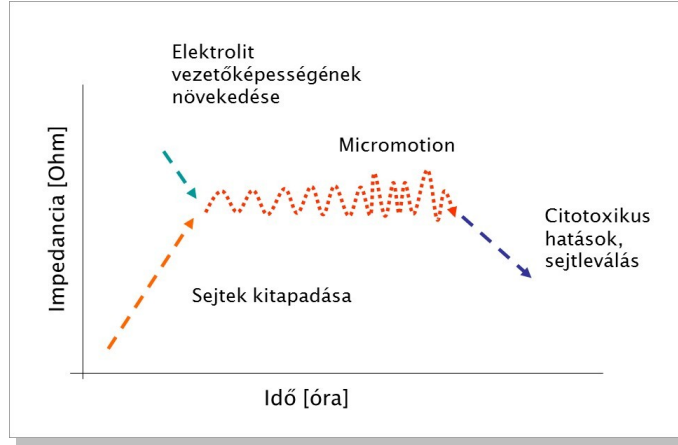


2.3. ábra. Magas frekvencián

Az impedancia értékeken túl a gép képes meghatározni az áramkör két feszültségkomponensét, az Ohmikus (R) és a kapacitív (C) ellenállását:

$$R = \frac{V_{fazis}}{I}, \quad C = \frac{V_{ellenfazis}}{I}$$

Ahol V_{fazis} jelöli a fázis feszültségét, $V_{ellenfazis}$ az ellenfázis feszültségét, és I az áramerősséget. Az újabb, ECIS Z Θ készülékkel már ellenállást (R) és kapacitást (C) is tudunk mérni különböző váltóáram frekvencia értékek mellett az idő függvényében [2]. Az egyes paraméterek eltérő biológia jelenségek jellemzésére használhatók, például 4 kHz frekvencia érték mellett az Ohmikus ellenállás értéke korrelál az elektród lefedettségével, így a sejtek adhéziójával, míg alacsonyabb frekvenciánál (pl. 400 Hz) inkább a barrier funkcióról, sejt-sejt kapcsolatokról nyújt információt. A 2.4 ábrán egy sematikus görbe mutatja be az impedancia érték változását egy kísérlet során. A sejtek kitapadásával párhuzamosan az értéke növekszik. Miután a sejtek teljesen elterültek, a mért impedancia elkezd fluktuálni. Ebben az esetben a fluktuáció a sejtek vertikális mozgásából, a mikromigrációból ered [1]. Feladatunk, hogy egy olyan matematikai paramétert keressünk, ami segít a jel fluktuációjának, azaz a görbe "szöszösségének" leírásában. Hiszen a fluktuáció egyedi a különböző sejt típusokra, azaz úgy határozza meg, mint egy embert az ujjlenyomata [4].



2.4. ábra. Az impedancia érték időfüggése, és változásának lehetséges okai
 Napjainkban az ECIS készüléket a biológia számos területén alkalmazzák. Többek között megfigyelhetjük a sejtek viselkedését, kémiai vegyületek hatását, mérhetjük a sejtek pusztulását, illetve a daganatos sejtek invazív jellegét [4],[5], valamint alkalmazható az endothel sejtek barrier funkciójának vizsgálatára is [6], [7]. Mivel a sejtek mozgási és morfológiai változásai nagy mértékben befolyásolják a készülékben mérhető impedancia értékét.

2.2. Zajanalitikai módszerek

Hosszú biológiai adatsorok vizsgálatára több különböző zajanalitikai módszer is ismert. A következő fejezetekben pár rutinszerűen alkalmazott módszert mutatunk be, és kitérünk ezek biológiai alkalmazhatóságára.

2.2.1. Variancia számítás

Különböző eloszlások megkülönböztetésére, valamint a matematikában az elméleti számításoknál gyakran használják a varianciát (más néven szórásnégyzetet). Azaz a variancia a valószínűségszámításban az eloszlásokat jellemző egyik fontos paraméter, illetve az elméleti statisztikában a szétszóródást megadó mérték [8]. A valószínűségszámításban is használt abszolút eltérésnél jobban jellemző paraméter, hiszen érzékenyebb a nagy eltérésekre, mely a definíciójából adódik:

$$Var(X) = Cov(X, X) = \mathbb{E}[(X - \mu)(X - \mu)] = \mathbb{E}[(X - \mu)^2]$$

ahol X egy valószínűségi változó, $\mathbb{E}(X) = \mu$ ezen változó várható értéke és $Cov(X, X)$ a változó önmagával vett kovarianciája. Ebből

$$Var(X) = \mathbb{E}(XX) - \mathbb{E}(X)\mathbb{E}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$$

azaz a szórásnégyzet nem más, mint a változó négyzetének várható értéke és a változó várható értékének négyzete közötti különbség.

A statisztikában a variancia torzítatlan becslése a következő:

$$Var(X) = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

ahol X egy valószínűségi változó, x_i ennek elemei, \bar{x} az átlaga, n pedig az elemszáma.

A varianciaanalízis eljárások a biológiában a legszélesebb körben alkalmazott technikák közé tartoznak az adatok elemzésére. Széleskörű felhasználása mind a kísérleti, mind a mintavételi tanulmányokban, azt mutatja, hogy a varianciaanalízis bizonyítottan könnyen érthető, értelmezhető módszertan. Mégis az irodalomban a kutatók véleménykülönbsége jelentős a felhasználásáról és az értelmezéséről [9].

2.2.2. Standard deviáció számítás

A statisztikában az adatok változékonyságának általánosan használt mérőszáma a standard deviáció, vagy más néven a minta szórása. Az előző fejezetben említett variancia másik megnevezése nem hiába szórásnégyzet, hiszen értéke megegyezik a szórás négyzetével. Azaz a szórás nem más, mint a variancia négyzetgyöke:

$$S(X) = \sqrt{Var(X)} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$$

2.2.3. Inkrement számítás

Az inkrement szintén az elméleti statisztikában egy valószínűségi változó vizsgálatára alkalmas paraméter. A paramétervektor egyes értékeit úgy kapjuk, hogy a változó vektorban az előtte és utána levő értékek különbségét képezzük.

$$I_i = R_{i+1} - R_{i-1} \quad i = 2, \dots, n - 1$$

ahol n a változó vektor hossza. Az inkrement vektor első és utolsó eleme a definíciója miatt megegyezik a változó vektor első és utolsó elemével.

$$I_1 = R_1 \quad I_n = R_n$$

Az inkrement standard deviációja alkalmas a mikromigráció karakterizálására, azaz arra, hogy mérjük vele a sejtek motilitását, és a rezisztencia jelek zaját. Láng és munkatársai [4] tanulmányukban megmutatták, hogy emelkedett érték a sejtek nagyobb motilitását jelenti.

2.2.4. Teljesítménysűrűség-spektrum

Egy jel teljesítménysűrűség-spektrumán (más néven autospektrumán) a teljesítmény frekvenciák szerinti eloszlását értjük [10]. Azaz a fluktuáló idősor autospektruma megadja a harmonikus összetevők hozzájárulását a jel teljesítményében. Egy periodikus viselkedés a teljesítménysűrűség-spektrumban csúcsként jelenik meg. Ha az autospektrumban egy bizonyos frekvencián jelentős csúcsot tapasztalunk, akkor ez azt jelenti, hogy a megfelelő összetevőnek nagy az amplitúdója, azaz ez az összetevő jelentős súlyt képvisel a jel teljesítményében. Amennyiben egyáltalán nem tartalmaz csúcsot, az a periodikus viselkedés hiányát jelenti az idősorban [2].

A vizsgált adatsor autospektrumát a következőképp tudjuk kiszámolni a Fourier-transzformáció segítségével:

$$S(0) = \frac{1}{N^2} |C_0|^2$$

$$S(k) = \frac{1}{N^2} (|C_k|^2 + |C_{N-k}|^2) \quad k = 1, 2, \dots, \frac{N}{2} - 1$$

$$S\left(\frac{N}{2}\right) = \frac{1}{N^2} |C_{N/2}|^2$$

A fenti képletekben N jelöli a minta hosszát és C_k ($k = 0, 1, \dots, N-1$) pedig a minta Fourier-transzformáltját. Esetünkben a minta diszkrét Fourier-transzformáltját (DFT) a következő módon lehet meghatározni:

$$C_k = \sum_{j=0}^{N-1} c_j \exp \frac{2\pi i j k}{N} \quad k = 0, 1, \dots, N-1$$

ahol a c_j ($j = 0, 1, \dots, N/2$) értékek a mintavételezett jel időtartománybeli értékei. Ezt az algoritmust arra használjuk, hogy az időtartománybeli mintavételezett jel értékeket frekvencia tartományba transzformáljuk. Abban az esetben, ha a minta hossza kettő hatvány, akkor gyors Fourier-transzformáltat (FFT) lehet használni. Esetünkben a lehetséges mintahosszok miatt érdemesebb volt a diszkrét Fourier-transzformáltat használni.

Végül az $S(i)$ ($i = 0, 1, \dots, N/2$) értékeket osztjuk a mintavételi frekvenciával és szorozzuk a minta hosszával. Így megkapjuk a vizsgált adatsor teljesítménysűrűség-spektrumát [2].

A fluktuáló jelek teljesítménysűrűség-spektrumának segítségével jól vizsgálható a görbe "szöszössége". Ha azt látjuk, hogy a jel amplitúdója hosszú idő alatt nagyot változik és kicsi a fluktuáció, akkor az autospektrum alacsony frekvenciás összetevői nagyok. Ellenkező esetben az alacsony frekvenciás összetevők kicsik.

2.2.5. Hurst-analízis

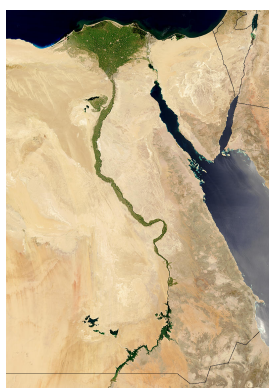
A Hurst exponens egy numerikus becslést ad az idősor megjósolhatóságára. Az idősorok statisztikai elemzésének egyik célja, hogy a múltbeli adatokból következtetni tudjunk ezen adatok jövőbeli változására. A módszer alapja, hogy feltételezzük, a múltban tapasztalt trend várhatóértékben a jövőben is folytatódni fog. Jelöljük a vizsgált paramétert H -val, melynek értéke 0 és 1 közé esik. A Hurst exponens matematikai definíciójának tekinthetjük, hogy $H = 2 - D$, ahol D jelöli a fraktáldimenziót. Az értékek különböző jelentésekkel bírnak. Amennyiben $0 < H < 0,5$, akkor a görbe viselkedését antiperzisztens viselkedésnek nevezzük, másnéven negatív autokorrelációnak. Ami számunkra azt jelenti, hogy ha t_{i-1} és t_i időpillanatok között a görbén növekedés történt, akkor annak nagyobb a valószínűsége, hogy a következő időpillanatban t_i és t_{i+1} között csökkenés lesz, és fordítva, ha a múltban csökkent, akkor a jövőben növekedni fog. A $H = 0,5$ véletlen bolyongást jelöl. Ekkor nincs korreláció az idősor jelenlegi és jövőbeli értékei között, azaz nulla az autokorreláció. Amennyiben a paraméter értéke $0,5$ és 1 közé esik, a görbe úgynevezett perzisztens viselkedést követ, másnéven ezt pozitív autokorrelációnak nevezzük. Ennek jelentése ellentétes a 0 és $0,5$ közötti viselkedéssel, vagyis a görbe növekedését nagy valószínűséggel növekedés fogja követni, és csökkenés esetén hasonló viselkedés figyelhető meg [2], [11].

Az eddig felsorolt paraméterek mellett dolgozatomban a Hurst exponens fontosságát, alkalmazhatóságát vizsgáltuk biológiai minták esetén.

2.3. A Hurst exponens története

A Hurst exponens az alkalmazott matematika számos területén használt paraméter, úgy, mint a fraktálok, a káosz elmélet, a hosszú távú folyamatok, illetve a spektrál analízis. A Hurst exponens becslését a matematikusokon kívül a biofizikusok, meteorológusok, pénzügyi szakemberek és az informatikusok is használják. Azonban eredete a hidrológiához kötődik, mégpedig az egyiptomi Nílus gátszabályozásához [12].

Edwin Harald Hurst britt hivatalnok 1906-ban került Kairóba, majd ott tartózkodása alatt beleszeretett a Nílusba, melynek tanulmányozásával, vizsgálatával teltek napjai. Láttá, hogy az alsó gát nem tölti be a kívánt szerepét és szükség van egy víztárolóval kombinált nagyobb gátra. Számításaihoz szüksége volt az előző évi víz-állásokra, ami hatalmas kutató munkát eredményezett. Megfigyelte, hogy a nagyobb áradásokat nagyobb valószínűséggel követnek nagyobb áradások és fordítva, kisebbeket pedig kisebb áradások. Tehát nem optimális a csapadék mennyiség átlagát venni és ez alapján kiszámítani a gát, illetve víztároló szükséges kapacitását, hanem nagyobb gátat kell építeni, mint ahogy azt a normális eloszlás adná. Ez a megfigyelés vezetett a róla elnevezett Hurst módszerhez [13].



2.5. ábra. Nílus

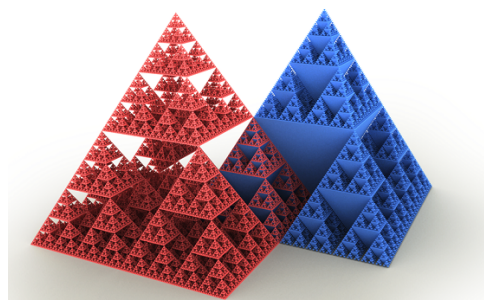


2.6. ábra. Edwin Harald Hurst

Később egy újfajta geometriai objektumnál, a fraktáloknál játszott fontos szerepet a Hurst exponens modern technikákkal való becslése. Ezen objektumok népszerűsítése egy amerikai matematikus, Benoit Mandelbrot nevéhez fűződik. A fraktál geometria a tudomány szinte minden területén fellelhető, amely segít a körülöttünk lévő világ más szemszögből való szemlélésében. Fraktáloknak két csoportját különböztetjük meg, úgy, mint a szabályos és a véletlen fraktálok. Utóbbi csoportba tartoznak a természetben fellelhető fraktálok, melyekkel már mindenki találkozhatott. Ilyen fraktál például a páfrány levél és a Sierpinski piramis.



2.7. ábra. Páfrány



2.8. ábra. Sierpinski piramis

Ezek fő tulajdonsága az önhasnolóság, melynek jelentése, hogy az adott objektum egy része hasonló magához az egész objektumhoz. Egy új, úgynevezett fraktáldimenzió bevezetésére volt szükség, ami a következő összefüggésben áll az általunk vizsgált Hurst exponenssel: $D = 2 - H$, ahol H jelöli a Hurst exponenst, D pedig a fraktáldimenziót. Miszerint egy nagyobb Hurst exponenssel rendelkező idősor kisebb fraktáldimenzióval, azaz simább felszínnel rendelkezik. Míg egy kisebb Hurst exponenssel rendelkező idősor nagyobb fraktáldimenzióval, azaz érdekesebb felszínnel rendelkezik [14], [2].

2.4. A Hurst exponens alkalmazásai

Nem meglepő, hogy a Hurst exponens kiszámítására több módszer is ismert, hiszen a biológián és orvostudományon túl számos területen használják ezt a paramétert. Ilyen területek például a pénzügy, kardiológia, hidrológia, geomagnetika, fizika, meteorológia. A következő fejezetekben ezek közül néhányról részletesebben is olvashatunk.

2.4.1. Pénzügyben

A Hurst exponens pénzügyi, tőzsdei vonatkozásban alkalmas annak vizsgálatára, hogy a mérsékelt kockázatú befektetési alapok, pénzpiaci-, hosszú kötvény-, és ingatlanalapok árfolyamának változása rendelkezik-e hosszú emlékezettel egy adott időszakban.

Annak ismerete, hogy az adott alap rövid- vagy hosszú emlékezettel rendelkezik olyan információt biztosít a befektetők számára, mely segíthet a jó befektetési stratégia kialakításában, illetve a megfelelő belépési pont megtalálásában [15]. A Hurst exponens segítségével egyfajta mérőszámot kapunk a memóriáról: 1 jelöli, hogy teljesen emlékszik a múltra, 0,5 a véletlen bolyongást, azaz csak arra emlékszik, hogy most hol van. 0 pedig, hogy még a mostani pozíciójára sem emlékszik.

Az idősorok statisztikai elemzésének egyik célja, hogy a múltbéli adatokból tudjunk következtetni az adatok jövőbeli alakulására. Így valamilyen mintázat beazonosításából, és e mintázatnak a jövőre történő kiterjesztéséből áll. Vagyis az a feltételezésünk, hogy a múltban felismert trend a jövőben is folytatódni fog.

A Hurst exponens becslésére az első ismert módszer az R/S-analízis volt. Az R/S-analízis különböző időperiódusokra kiszámolja a kumulált adatok átlag körüli ingadozásainak R terjedelmét, majd ezt az adatok S szórásával elosztva standardizálja, ezért is nevezik újraskálázott terjedelem-analízisnek (angolul: Rescaled Range Analysis, röviden R/S-analízis). Feltetésünk szerint $\frac{R}{S}(l) \sim l^H$ [11], ahol l jelöli az egyes időperiódusok hosszát.

2.4.2. Szívverés és érzelmi minőség kapcsolatában

Tommaso Costa és munkatársai [16] megvizsgálták a pulzus idősorának Hurst exponensét, és összefüggését az érzelmi minőséggel és az izgatottság mértékével. A különböző érzelmi filmingerek (boldogság, szomorúság, düh, félelem) hatása alatt elektrokardiogrammot (EKG) készítettek. Az EKG görbe alapján számított pulzus idősorok Hurst exponense különbséget mutatott boldog és szomorú érzelmi állapotok között, de a negatív érzelmek esetén (szomorúság, düh, félelem) nincs különbség. Ami számunkra azt jelzi, hogy a Hurst exponens a szubjektív érzelmi minőség indikátora. Azaz a Hurst exponens ismeretében a boldogság és szomorúság érzéseket meg tudjuk különböztetni, még negatív érzelmi állapotok esetén nincs különbség a Hurst exponens értékében. Ez azt jelzi, hogy a strukturális változások a fiziológiai jelekben kapcsolatban vannak az érzelmi élmények szubjektív érzelmi minőségével. Az eredmények alátámasztják azt a feltételezést, mi szerint a szív jeleinek szerkezeti felépítése egy diszkrét tulajdonság, amely vált a negatív és pozitív érzelmek között. Pontosabban, nagy fraktáldimenzió esetén (ekkor az ismert $H = 2 - D$ összefüggés miatt a Hurst exponens kicsi) a pozitív érzelmek dominálnak, míg a negatív érzelmek nagy Hurst értéket adnak. Larsen, Diener és Russel [17] [18] szerint a félelem és a düh jellemezhető hasonló érzelmi minőséggel, míg ez a minőség különböző boldogság és szomorúság esetén. Ez a megállapítás összhangban van Barrett és Russel [19] eredményeivel, akik megmutatták, hogy az izgatottság független az érzelmi minőségtől. Továbbá, Barrett [20] megmutatta, hogy az érzelmi minőség alapvető, invariáns építőköve az emberi elme képességéből eredő érzelmi életnek. Ezek az eredmények azt sugallják, hogy a vegetatív folyamatok hozzájárulhatnak a kódoláshoz és az érzelmi ingerek kitapasztalásához. Costa, Galati és Rognoni tanulmányában elfogadott mérték a Hurst exponens, amely látszólag képes érzékelni az érzelmi folyamatokban előforduló változásokat. Így a Hurst exponens kapcsolatot teremt a kardiológiai folyamatok és az érzelmi élmények között. Ezért több vizsgálatot kell

végezni az érzelmi állapotok fiziológiai jeleiről új elemzési módszereket alkalmazva. És nem utolsó sorban ebben a tanulmányban megmutatták, hogy a Hurst exponens vizsgálata az ún. wavelet transzformációs módszerrel képes volt feltárni a kardiológiai jelek sajátos tulajdonságait az érzelmi minőséggel kapcsolatban.

2.4.3. Csapadék mennyiségének mérésében

A Hurst hatás a hidrológiában jól megalapozott és általánosan elismert, mint egy fontos tulajdonsága a természet működésének [21], [22]. A csapadék Hurst exponensének vizsgálatát az motiválja, hogy az aszályok és árvizek ciklusai magyarázhatók a Hurst jelenség véletlen hatásaival [23], [24]. Jamal Munshi [25] tanulmányában elemzi Anglia és Wales 1766 és 2016 közötti havi csapadékmennyiségét éves időskálán mind a 12 naptári hónapban. Eredményeiben nem található bizonyíték függőségre, hosszú távú memóriára vagy perzisztenciára. A tanulmányban azt kapta, hogy mind a 12 vizsgált hónap csapadékmennyiségének Hurst exponense 0,5-höz közeli, mely Gauss véletlenszerűsége utal. Tehát a havi átlagos csapadékmennyiség idősorra úgy tekinthetünk, mint ha Gauss sorozat lenne, és feltehető, hogy teljesül a függetlenségi feltétel.

2.4.4. Geomagnetikus mezőben

Paola De Michelis és Giuseppe Consolini [26] tanulmányukban megpróbálták jellemezni az északi félteken elhelyezett 45 földi geomágneses megfigyelő rendszerből származó rövid időskálájú mágneses tér ingadozások skálázási jellemzőinek térbeli eloszlását. A geomagnetikus mező fluktuációjának skálázási tulajdonságainak változásait vizsgálták a lokális Hurst exponens becslésével és rekonstruálták ezt az indexet, mint a geomagnetikus aktivitási szint függvénye. Ez a leképezés lehetővé tette számukra, hogy lokalizálják a zavarokért felelős és az ionoszférikus áramrendszerekhez kapcsolódó különböző szélességi struktúrákat. Megállapították, hogy a különböző ionoszférikus áramrendszerekhez kapcsolódó geomágneses tér fluktuációinak különböző méretezési jellemzői vannak, amit a lokális Hurst exponens bizonyít. Azt is belátták, hogy a lokális Hurst exponens általában csendes magnetoszférikus periódusokra nagyobb, mint az aktívabb időszakokban, ami arra utal, hogy a megzavart periódusok folyamán aktivált dinamikai folyamatok felelősek a geomágneses tér fluktuációjának változásaiért a természetben.

2.4.5. Egyéb alkalmazásai

A hidrológiai vizsgálatokban gyakran használják a Hurst analízist, ezért a Hurst exponens kiszámítására számos különböző becslési módszer ismert. Byung-Sik Kim,

Hung-Soo Kim és Sun-Hong Min [27] tanulmányában hét különböző módszert használnak erre a becslésre, úgy mint, a korrigált terjedelem analízist, az újraskálázott terjedelem analízist (RR), a módosított újraskálázott terjedelem analízist (MRR), az $\frac{1}{f}$ autospektrum analízist, a maximum Likelihood becslést (MLE), a DFA módszert, és az aggregált eltérési idő módszert (AVT). Megállapították, hogy a DFA módszer a legmegfelelőbb a Hurst exponens becslésére, mind a rövidtávú memória, mind a hosszútávú memória esetén. Azt találták, hogy a SOI (déli oszcillátorok indexe) sorozatok -melyek nem lineáris idősorok, és hosszútávú memóriával rendelkeznek- Hurst exponense 0,92, míg a hasonló tulajdonságú fa gyűrű sorozatok Hurst exponense 0,97.

2.5. A Hurst exponens kiszámítási módszerei

A Hurst exponens kiszámítására több módszer is ismert, melyek a következő alfejezetekben olvashatók időrendi sorrendben.

2.5.1. R/S-analízis algoritmus

A Hurst exponens kiszámítására alkalmazott első és legnépszerűbb módszer az R/S -analízis volt, mely Mandelbrot és Wallis [28] nevéhez fűződik, és Hurst [29] tanulmányán alapszik. Az R/S módszer -melyet először Mandelbrot [30],[31] vezetett be a közgazdaságtanban- különböző időperiódusokra kiszámolja a kumulált adatok átlag körüli ingadozásainak R terjedelmét, majd ezt az adatok S szórásával elosztva standardizálja, ezért is nevezik újraskálázott terjedelemnek (Rescaled Range, RR). Feltehető, hogy a feldolgozandó idősorunk n darab adatot tartalmaz. Ezt az adatsort szeretnénk felosztani különböző hosszúságú szakaszokra. Jelölje l a szakaszok hosszát és s a szakaszok számát. Egy lépésben a szakaszok hossza legyen azonos és legalább 10-20 adatot tartalmazzon egy szakasz, azaz a hosszuk legyen minimum 10-20. A j -edik szakaszon a következőképp számoljuk ki RR-t. Jelöljük E_i -vel az ezen a szakaszon mért ellenállások értékeit ($i = 1, 2, \dots, l$).

1. lépés: Szükségünk van a rezisztencia értékek átlagára: $m = \frac{E_1 + E_2 + \dots + E_l}{l}$.
2. lépés: Számoljuk ki az értékek eltérését a kapott átlagtól: $D_i = |E_i - m|$.
3. lépés: Képezzük a következő összegeket: ($t = 1, 2, \dots, l$)

$$P_t = \sum_{i=1}^t D_i$$

4. lépés: Ezen P_t értékek maximumának és minimumának különbsége adja idősor

R -rel jelölt terjedelmét: $R(l) = \max(P_1, \dots, P_l) - \min(P_1, \dots, P_l)$.

5. lépés: A standardizáláshoz határozzuk meg az S -sel jelölt E_i értékek szórását, más néven a standard deviációját:

$$S(l) = \sqrt{\frac{1}{l} \cdot \sum_{i=1}^l (E_i - m)^2}$$

6. lépés: Az úgynevezett Rescaled Range-t pedig úgy kapjuk meg, hogy az R -et elosztjuk a standard deviációval: $\frac{R}{S} = \frac{R(l)}{S(l)}$.

7. lépés: Számoljuk ki az összes szakaszra az R/S értéket és vegyük átlagukat.

Mivel a következő összefüggés teljesül: $(R/S)_l \approx c \cdot l^H$, ezért a H -val jelölt Hurst exponens lineáris regresszióval megkapható: $\log(R/S)_l = \log c + H \cdot \log l \Rightarrow H = \frac{\log(R/S)_l - \log c}{\log l}$ [2], [32].

Sanchez Granero és munkatársai az R/S analízist tesztelték több Brown mozgásból származó adatsorra, és azt figyelték meg, hogy az l érték, azaz a szakaszok hosszának a megválasztása nagy mértékben hatással van a kapott átlagok, illetve standard deviációk értékeire [32]. Amennyiben l elég nagy, az átlag közelebb kerül a valós 0,5 értékhez. Kis l értékek esetén az átlag nő. Tehát ahhoz, hogy megfelelően használjuk az R/S analízist, nagy l értéket kell választanunk, ez azonban nem lehetséges rövid adatsorok esetén. Így az R/S analízis használatakor problémába ütközhetünk, ha nem elég hosszú az adatsorunk. Továbbá még egy probléma, hogy kvantált adatokkal dolgozunk, ezért kis l értékre előfordulhat, hogy a standard deviáció nulla, illetve a standard deviáció miatt a módszer érzékeny a kerekítésre.

2.5.2. Anis és Lloyd módszere

A Nílus gátszabályozásakor működő R/S módszer sajnos a gyakorlatban, más idősorok esetén nem adott mindig pontos, értelmezhető értékeket. Ezért más módszerek is születtek H kiszámítására. Például Anis és Lloyd [33] egy új módszerrel állt elő. Először a következő formulával kiszámítjuk az $\mathbb{E}(R/S)_l$ értéket:

$$\mathbb{E}(R/S)_l = \begin{cases} \frac{l-\frac{1}{2}}{l} \cdot \frac{\Gamma(\frac{l-1}{2})}{\sqrt{\pi} \cdot \Gamma(\frac{l}{2})} \cdot \sum_{i=1}^{l-1} \sqrt{\frac{l-i}{i}} & \text{ha } l \leq 340 \\ \frac{l-\frac{1}{2}}{l} \cdot \frac{1}{\sqrt{l \cdot \frac{\pi}{2}}} \cdot \sum_{i=1}^{l-1} \sqrt{\frac{l-i}{i}} & \text{ha } l \geq 340 \end{cases}$$

Majd a H -val jelölt Hurst exponenst úgy kapjuk, hogy 0,5-höz hozzáadjuk az $(R/S)_l - \mathbb{E}(R/S)_l$ görbe meredekségét. Azonban ekkor Sanchez Granero és munkatársainak [32] cikkében (és nálunk is) a kapott H érték negatív lehet, aminek

nincs értelme. Ezt a problémát elkerülve $\mathbb{E}(R/S)_l$ meghatározása után a következő számítást javasolják cikkükben:

$$\log H_l = \log (R/S)_l - \log (\mathbb{E}(R/S)_l) + \log (l)/2$$

Végül lineáris regresszióval meghatározzuk a Hurst exponenst:

$$\log H_l = \log c + H \cdot \log l$$

Ez a módszer abból indul ki, hogy normális eloszlású a zaj. Amennyiben ez nem teljesül (nálunk ez a kvantálás miatt vitatható), akkor nem garantált, hogy helyes H értéket kapunk.

2.5.3. GM1 módszer

A Hurst exponens becslésére a geometriában a következő formula [32] is ismert:

$$\overline{\Delta B} \propto l^H$$

ahol $\overline{\Delta B} = \overline{|B(t+l) - B(t)|}$, l jelöli a szakaszok hosszát, H a Hurst exponenst, és α egy arányszám. Először felosztjuk az n hosszú adatsort s szakaszra, melyek hossza l és minden $m = 1, \dots, s$ -re végezzük el a következő számításokat:

1. $S_m = X_{ml} - X_{(m-1)l+1}$
2. $H_l = \text{mean}\{S_m : m = 1, \dots, s\}$.

Ekkor lineáris regresszióval a Hurst exponens kiszámítható:

$$\log H_l = \log c + H \log l$$

2.5.4. GM2 módszer

Az előző módszer már helyes H értéket ad eredményül, azonban a GM1 módosításával kapott GM2 módszer pontosabb megoldást adhat:

$$\overline{\text{range}(B)} \propto l^H$$

ahol $\overline{\text{range}(B)} = \overline{\max\{B(S) : t \leq S \leq t+l\} - \min\{B(S) : t \leq S \leq t+l\}}$.

Több véletlen bolyongásból származó adatsorra tesztelve a GM1 módszert és ennek módosítását, a GM2 módszert, Sanchez Granero és munkatársainak [32] cikke alapján azt állíthatjuk, hogy mindkettő megfelelően működik, hiszen Brown

mozgásra mindkettő 0,5-höz közeli Hurst exponenst ad eredményül. A cikk [32] szerint a standard deviáció kisebb volt a GM2 módszernél, mint a GM1-nél. Ezért, amennyiben lehetőségünk van a GM2 módszert alkalmazni, ez a módszer pontosabb értéket ad.

3. fejezet

Célkitűzés

A diplomamunka készítése során az alábbi célokat tűztük ki:

1. A Hurst exponens biológiai fontossága, alkalmazhatósága és az irodalomban megtalálható számítási módszerek beprogramozása, illetve összehasonlítása.
 - (a) Minta adatsorok generálása különböző Hurst értékekre és ezek tesztelése az irodalomban található módszerekkel.
 - (b) Biológiai mintából származó adatok vizsgálata a módszerekkel.
 - (c) Sejtek kezelésének hatásai a Hurst exponensre.
2. Egyéb zajanalízisre alkalmas paraméterek alkalmazása és összevetése.
3. Az eredmények elemzése és a Hurst exponens számításának korlátai.

4. fejezet

Anyag és módszer

4.1. Adatsorok előállítása

4.1.1. Generált adatsorok

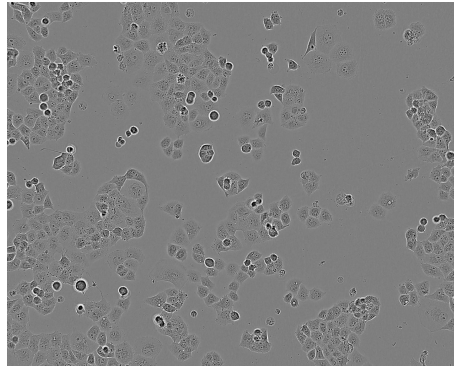
Először különböző H értékekre generáltunk adatsorokat az R programban beépített függvény segítségével. A függvény két bemeneti paramétere n , azaz a minta elemszáma (ami a mi esetünkben 3600, mivel ennyi a biológiai minták hossza), és H , a Hurst exponens ($H = 0; 0, 1; 0, 2; 0, 3; 0, 4; 0, 5; 0, 6; 0, 7; 0, 8; 0, 9; 1$).

4.1.2. Biológiai adatsorok

Majd a Semmelweis Egyetem Genetikai, Sejt- és Immunbiológiai Intézet kutató csoportja által, az ECIS készülék segítségével generált adatokat teszteltük. Ezen adatok egy része sejtmentes mintából, míg másik része sejtes mintából származott. A statisztikai elemzés során legfőbb célunk, hogy a két minta a Hurst exponens ismeretében szignifikánsan megkülönböztethető legyen.

4.1.2.1. Sejtek és tenyésztésük

A biológiai kísérletekhez egy 56 éves hasnyálmirigy tumoros betegből izolált, PANC1 sejtvonalat használtunk, amelyet standard sejtenyésztési körülmények között tenyésztettünk. A sejteket steril környezetben, 37 Celsius fokon, 5% CO₂ tartalom mellett növesztettük, DMEM médiumban, amit kiegészítettünk 2mg/ml glutaminnal, 10% FBS-sel és 1% penicillin-Streptomycin komplexszel. A tenyészeteket 3 naponta, 90%-os konfluenciánál passzáltuk a 4.1.2.2. fejezetben említett lépéseknek megfelelően.



4.1. ábra. PANC1 sejtek morfológiája fáziskontraszt mikroszkópban ([34])

4.1.2.2. Mikromigráció mérés ECIS készülékben

A mérés előkészületeit a steril munkavégzést lehetővé tevő, úgynevezett lamináris boxban végeztük. A biológiai kísérlet menete röviden a következő volt.

Először csak a sejtenyésztésre használt médiumot teszteltük az ECIS készülékben. Ehhez az arany elektródokat tartalmazó ECIS 96W1E típusú plate well-jeibe $200\mu\text{l}$ sejtenyésztő oldatot pipettáztunk. Ezután a plate-et behelyeztük a gép mérőegységébe, amely az inkubátorban volt található, így biztosítva a sejtek számára a megfelelő közeget, ami 37°C Celsius fokot és 5% szén-dioxidot jelent. Egy rövid egyensúlyozást (kb. 1 h) követően az impedancia mérésre az alábbi beállítások mellett került sor:

1. AC frekvencia = 4000 Hz
2. Mintavételi frekvencia 2 minta másodpercenként
3. Mérés időtartama fél óra (így 3600 hosszú adatsorokat kaptunk).

Az alapvonal vagy sejtmentes minta mérését követően sejteket adtunk a rendszerhez. A sejtoldat előkészítése passzálás során történt, melynek főbb lépései a következők. Először a sejtekről leöntöttük a kondicionált médiumot, majd foszfát pufferben ($\text{pH } 7,4$) mostuk őket, így eltávolítva a halott sejteket, illetve a sejtörmelékeket. Ahhoz, hogy megszüntessük a sejtek kitapadását a tenyésztésre használt flaskában, úgynevezett trypszin enzimet pipettáztunk a sejtekre. A sejtek leválását mikroszkópban ellenőriztük, majd a reakciót friss médiummal leállítottuk. A sejtoldatot lecentrifugáltuk 200g -vel majd médiumot cseréltünk a sejteken. Neubauer hemocytometer segítségével, mikroszkópban megszámoltuk a sejteket majd hígítottuk 10^5 sejt/ml végkoncentrációra. Végül a plate-et kivettük az ECIS készülék mérőfejéből és steril körülmények között a well-ekbe pipettáztuk. Az alkalmazott sejtdenzitás 10^4 sejt/well volt. A korábban beállított paraméterek mellett folytattuk a mérést. A sejtek

kiültetését követő 24 órában csak a kitapadást monitoroztuk, majd a kívánt adhézió beállta után a korábban említett beállítások mellett elvégeztük a mikromigráció mérését. A kapott adatok közül (Z, C, R) a valós ellenállás értékeket használtuk a matematikai elemzésnél.

4.1.2.3. Kezelés és vizsgált vegyületek

Végeztünk olyan méréseket is, ahol a korábban bemutatott mikromigráció méréseket követően (24 – 26h-nál), a sejteket tumor ellenes szerekekkel megkezelve is monitoroztuk a mikromigrációt (5.6).

4.2. Software-es rutinok

Ahhoz, hogy a generált és a biológiai adatsorokat az előző fejezetekben bemutatott zajanalitikai módszerek segítségével elemezni tudjuk, MATLAB-ban rutinokat írtunk ezek meghatározására.

4.2.1. Variancia, inkrement kiszámolására

A variancia, inkrement és ennek standard deviációjának kiszámolására MATLAB rutint írtunk, melynek bemenete egy \underline{R} vektor, mely tartalmazza a mért rezisztencia adatokat (R értékek), és az adatsor hossza, azaz a vektor elemszáma.

```
1 function [Rstd,Rvar,Rincrement,Rincsd]=variancia(R,hossz)
2
3 for i=1:hossz:length(R)-hossz
4     Ratlag(ceil(i/hossz))=mean(R(i:i+hossz-1));
5 end
6
7 Rstd=std(Ratlag);
8 Rvar=var(Ratlag);
9
10 Rstd
11 Rvar
12
13 for i=2:length(Ratlag)-1
14     Rincrement(i)=Ratlag(i+1)-Ratlag(i-1);
15 end
16
17 Rincrement(1)=Ratlag(1);
18 Rincrement(length(Ratlag))=Ratlag(length(Ratlag));
19
20 Rincsd=std(Rincrement);
21
22 Rincsd
23
24 end
```

MATLAB rutin - Variancia és inkrement számítás

4.2.2. Teljesítménysűrűség-spektrum kiszámolására

Az irodalmi áttekintésben leírt számítás alapján, a Fourier-transzformáció beépített függvényének segítségével egy idősor autospektrumának kiszámítására írt rutin a következő. A függvény bemenetként a vizsgált rezisztencia értékeket kapja. Ahhoz, hogy meghatározza az autospektrum vektor egyes elemeit, szüksége van az adatsor hosszára (N), és az adatsor Fourier-transzformáltjára (C). Végül kiszámolja a görbére illesztett egyenes meredekségét.

```
1 function [p,Aspec]=autospec(X)
2
3 N=length(X);
4 C=fft(X);
5 C=C(1:N);
6 C;
7
8 S(1)=(1/N^2)*(abs(C(1)))^2;
9 for k=2:N/2-1
10     S(k)=(1/N^2)*((abs(C(k)))^2+(abs(C(N-k)))^2);
11 end
12 S(N/2)=(1/N^2)*(abs(C(N/2)))^2;
13
14 for i=1:N/2
15     Aspec(i)=(S(i)/2)*N;
16 end
17
18 p = polyfit(1:1800,10*log10(Aspec),1);
19 x1 = linspace(0,1800);
20 f1 = polyval(p,x1);
21 plot(1:1800,10*log10(Aspec))
22 hold on
23 plot(x1,f1,'r—')
24
25 Aspec
26
27 end
```

MATLAB rutin - Teljesítménysűrűség-spektrum számítás

4.2.3. Hurst exponens meghatározására

A Hurst exponens kiszámítására alkalmas, előző fejezetekben bemutatott öt módszerre írt MATLAB rutin a következő. A rutin bemenetként kap egy \underline{X} vektort, mely tartalmazza a mért rezisztencia adatokat (R értékek), majd eredményül megadja az öt különböző módszerrel meghatározott H értéket (H_{cl} , H_{al1} , H_{al2} , H_{geo1} , H_{geo2}).

```
1 rangek=200:floor(length(X)/5);
2 [rs,ers]=RSana(X,rangek,'Hurst');
3 Filter=~isnan(rs);
4 lrs=log10(rs(Filter));
5 lrangek=log10(rangek(Filter));
6 Hcl=polyfit(lrangek,lrs,1)
7 figure(1);
8 plot(lrangek,lrs,lrangek,lrangek*Hcl(1));
9 lH=lrs-log10(ers(Filter))+lrangek/2;
10 Hal2=polyfit(lrangek,lH,1)
11 figure(2);
12 plot(lrangek,lH,lrangek,lrangek*Hal2(1));
13 Hal1=polyfit(rangek(Filter),rs(Filter)-ers(Filter),1)
14 figure(3);
15 plot(lrangek,lH,lrangek,lrangek*Hal1(1));
16 rs=RSana2(X,rangek,'Geo1');
17 lH=log10(rs);
18 lrangek=log10(rangek);
19 Hgeo1=polyfit(lrangek,lH,1)
20 figure(4);
21 plot(lrangek,lH,lrangek,lrangek*Hgeo1(1));
22 rs=RSana2(X,rangek,'Geo2');
23 lH=log10(rs);
24 lrangek=log10(rangek);
25 Hgeo2=polyfit(lrangek,lH,1)
26 figure(5);
27 plot(lrangek,lH,lrangek,lrangek*Hgeo2(1));
```

MATLAB rutin - Hurst exponens kiszámításai

Ez a rutin az RSana2 függvényt használja, ami a függelékben megtalálható.

4.3. Statisztika

A matematikai statisztika segítségével megvizsgáltuk, hogy az előbbiekben meghatározott (GM2 módszerrel számolt) Hurst exponens ismeretében egy adatsorról szignifikánsan eldönthető-e, hogy az sejtmentes vagy sejtes mintából származik-e.

Feladatunk, hogy a H_0 -lal jelölt nullhipotézis és a H_1 -gyel jelölt alternatív hipotézis (ellenhipotézis) között döntsünk. Esetünkben a hipotézisek:

H_0 : Hurst exponens nem megfelelő paraméter

versus

H_1 : Hurst exponens megfelelő paraméter

ahol azt mondjuk, hogy a Hurst exponens akkor megfelelő, ha a sejtmentes adatok esetén számolt H érték nem azonos a sejtes adatok esetén számolt H értékkel. Ha megegyezik a kettő, akkor a Hurst exponens nem megfelelő.

Ezután csak egy számunkra megfelelő statisztikai próbát kellett találnunk, melynek feltételei teljesülnek és így használni is tudjuk. Olyan statisztikára volt szükségünk, amely azt vizsgálja, hogy két külön mintában egy-egy valószínűségi változó átlagai egymástól szignifikánsan különböznek-e. Ez nem más mint a két-mintás t-próba. Azonban mielőtt használtuk a tesztet, le kellett ellenőriznünk, hogy feltételei teljesülnek-e a vizsgált adatsorokra, azaz, hogy normális eloszlásúak-e, illetve, hogy szórásuk megegyezik-e.

4.3.1. Normalitás teszt

A statisztikában a normalitás tesztet arra használják, hogy eldöntsék egy adatsorról, hogy az normális eloszlású-e. Több erre alkalmas tesztet ismerünk. Razali és Wah [35] tanulmánya szerint a legnagyobb ereje a Shapiro-Wilk tesztnek van.

4.3.1.1. Shapiro-Wilk teszt

Shapiro-Wilk [36] teszt esetén adott egy n hosszú adatsor, melyről el szeretnénk dönteni, hogy normális eloszlást követ-e.

1. lépés: Rendezzük növekvő sorrendbe az adatokat: $x_1 \leq x_2 \leq \dots \leq x_n$
2. lépés: Számoljuk ki a következő összeget:

$$SS = \sum_{i=1}^n (x_i - \bar{x})^2$$

ahol \bar{x} jelöli az adatok átlagát.

3. lépés: Ha n páros, akkor $m = \frac{n}{2}$. Ha n páratlan, akkor $m = \frac{n-1}{2}$.
4. lépés: Számoljuk ki b -t a következőképp:

$$b = \sum_{i=1}^m a_i (x_{n+1-i} - x_i)$$

ahol az a_i értékeket a Shapiro-Wilk 1. táblából (8.1) olvashatjuk le. Megjegyezzük, hogy ha n páratlan, akkor a medián értékét nem használjuk b kiszámításában.

5. lépés: Számoljuk ki a teszt statisztikát:

$$W = b^2/SS$$

6. lépés: A kapott W érték és n ismeretében a Shapiro-Wilk 2. táblából (8.2) leolvasható a teszt p értéke.

Amennyiben a kapott p érték $> 0,05$, akkor a nullhipotézist fogadjuk el, miszerint az adatsor normális eloszlású.

4.3.2. F-próba

A kétmintás t-próba másik feltétele, két normális eloszlású változó szórásának egyenlősége F-próbával ellenőrizhető.

Legyen $X \sim \mathcal{N}(\mu_1, \sigma_1^2)$ és $Y \sim \mathcal{N}(\mu_2, \sigma_2^2)$ két ismeretlen paraméterű, normális eloszlású háttérváltozó. A szórások egyenlőségének tesztelésére vonatkozó hipotézis vizsgálat a következő:

$$H_0 : \sigma_1 = \sigma_2 \quad \text{versus} \quad H_1 : \sigma_1 \neq \sigma_2$$

Az n_1 elemű $X_1, \dots, X_{n_1} \sim \mathcal{N}(\mu_1, \sigma^2)$ független, azonos eloszlású és az $Y_1, \dots, Y_{n_2} \sim \mathcal{N}(\mu_2, \sigma^2)$ független, azonos eloszlású, egymástól is független minták alapján vizsgálódunk. Tudjuk, hogy $(n_1 - 1)S_X^{*2}/\sigma_1^2 \sim \chi^2(n_1 - 1)$ és $(n_2 - 1)S_Y^{*2}/\sigma_2^2 \sim \chi^2(n_2 - 1)$ függetlenek. Leosztva őket külön-külön a saját szabadságfokukkal, majd a hányadosukat véve $\mathcal{F}(n_1 - 1, n_2 - 1)$ -eloszlású valószínűségi változót kapunk. H_0 fennállása esetén a hányados:

$$F(\mathbf{X}, \mathbf{Y}) = \frac{S_X^{*2}}{S_Y^{*2}}$$

így ezt a próbastatisztikát vezetjük be. Ezután $1 - \epsilon$ szinten elutasítjuk H_0 -t, ha $F(\mathbf{X}, \mathbf{Y}) \geq F_{\epsilon/2}(n_1 - 1, n_2 - 1)$, ahol a megfelelő szabadságfokú F-eloszlás $(1 - \epsilon/2)$ -kvantilise a kritikus érték [37].

4.3.3. Kétmintás t-próba

A kétmintás t-próbát [37] kis mintákra szokták alkalmazni, melyek szükségképpen normális eloszlásúak, és szórásuk megegyezik. A próbát normális eloszlás várható értékének tesztelésére vagy két normális várható érték összehasonlítására hasz-

nálják. A kétmintás jelző arra utal, hogy két tetszőleges várható értékű, de azonos szórású háttérváltozót vizsgálunk. (Két normális eloszlású változó szórásának összehasonlítására az F-próbát használhatjuk.) Legyenek ezek: $X \sim \mathcal{N}(\mu_1, \sigma^2)$ és $Y \sim \mathcal{N}(\mu_2, \sigma^2)$. Ezen belül megkülönböztetünk egyoldali és kétoldali alternatívát. Nekünk most az utóbbira volt szükségünk. A hipotézis vizsgálat ekkor:

$$H_0 : \mu_1 = \mu_2 \quad \text{versus} \quad H_1 : \mu_1 \neq \mu_2$$

Az n_1 elemű $X_1, X_2, \dots, X_{n_1} \sim \mathcal{N}(\mu_1, \sigma^2)$ független, azonos eloszlású és az n_2 elemű $Y_1, Y_2, \dots, Y_{n_2} \sim \mathcal{N}(\mu_2, \sigma^2)$ független, azonos eloszlású, egymástól is független mintákból konstruált próbastatisztika:

$$t(\mathbf{X}, \mathbf{Y}) = \frac{\bar{X} - \bar{Y}}{\sqrt{(n_1 - 1)(S_X^*)^2 + (n_2 - 1)(S_Y^*)^2}} \cdot \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}}$$

és az $1 - \epsilon$ szignifikanciaszinthez konstruált kritikus tartomány:

$$\mathcal{X}_k = \{(\mathbf{x}, \mathbf{y}) : |t(\mathbf{x}, \mathbf{y})| \geq t_{\epsilon/2}(n_1 + n_2 - 2)\}$$

ahol $\frac{(n_1 - 1)(S_X^*)^2 + (n_2 - 1)(S_Y^*)^2}{\sigma^2} \sim \chi^2(n_1 + n_2 - 2)$.

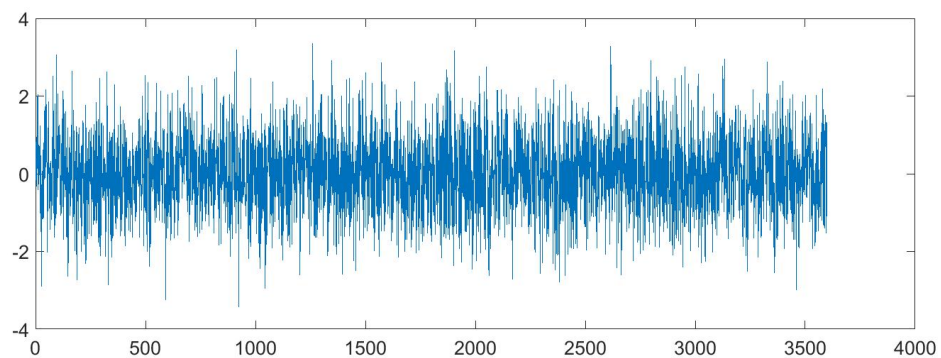
Tehát a nullhipotézisünket $1 - \epsilon$ szinten elfogadjuk, ha a mintarealizációból számolt $|t(\mathbf{x}, \mathbf{y})| < t_{\epsilon/2}(n_1 + n_2 - 2)$, különben pedig elutasítjuk.

5. fejezet

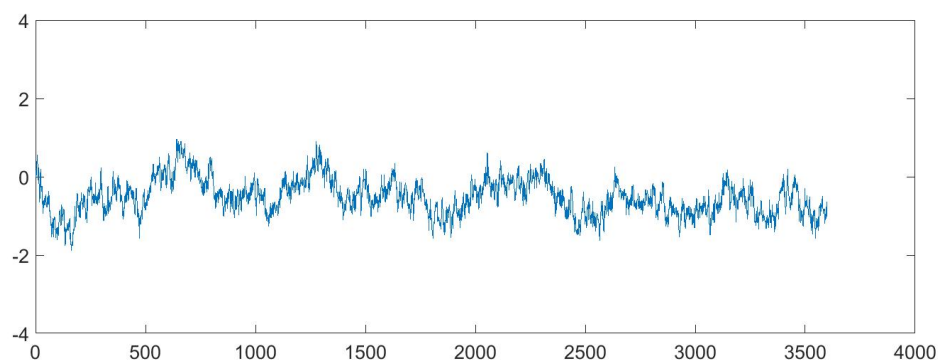
Eredmények

5.1. Számítógéppel generált adatsorok

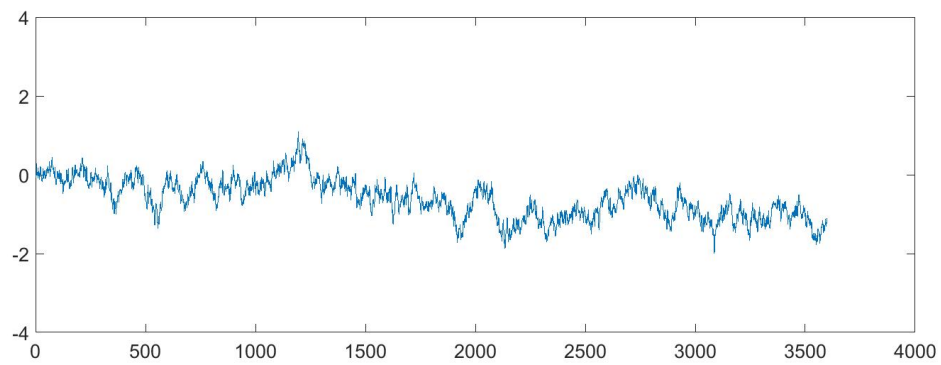
Az R program egy beépített függvénye (`fbmSim`) segítségével több különböző Hurst exponenssel rendelkező adatsort generáltunk, megadva az elemszámot (3600) és a H értéket (0; 0, 1; 0, 2; 0, 3; 0, 4; 0, 5; 0, 6; 0, 7; 0, 8; 0, 9; 1).



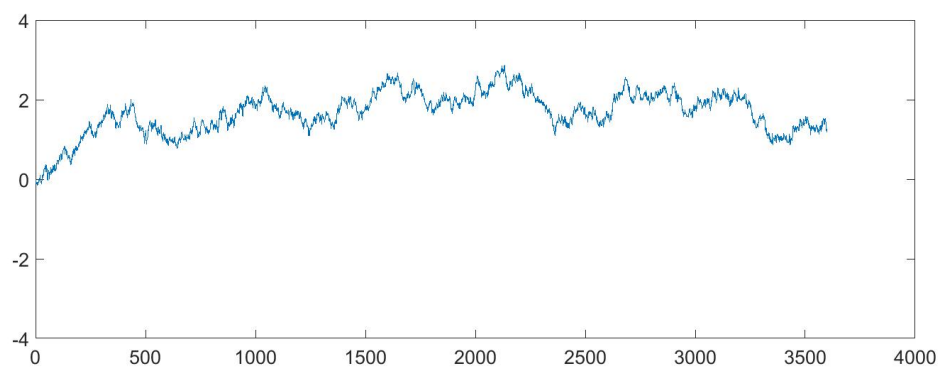
5.1. ábra. $H = 0$ - fehér zaj



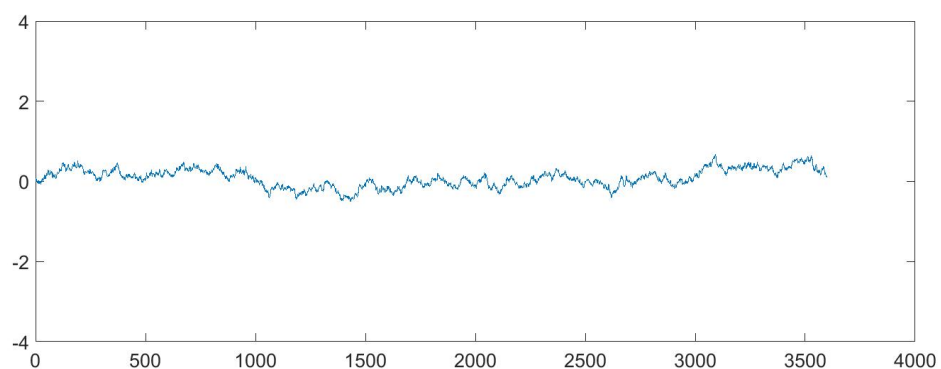
5.2. ábra. $H = 0,1$



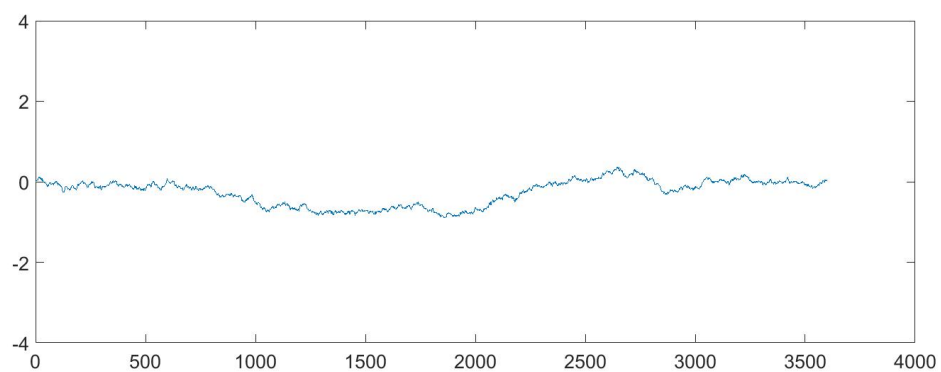
5.3. ábra. $H = 0,2$



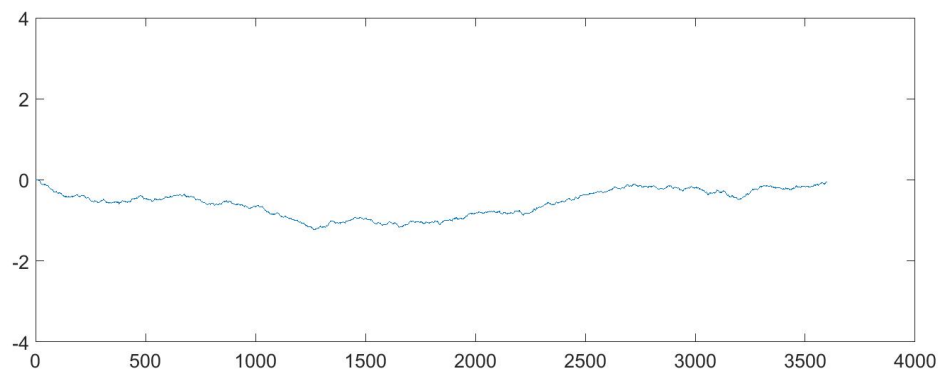
5.4. ábra. $H = 0,3$



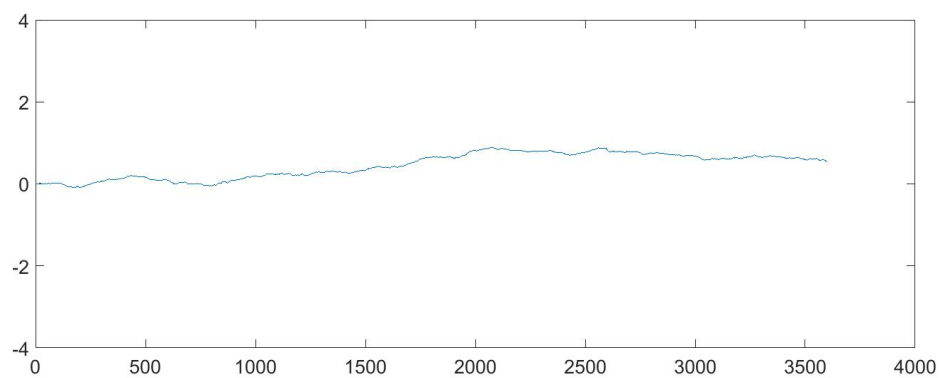
5.5. ábra. $H = 0,4$



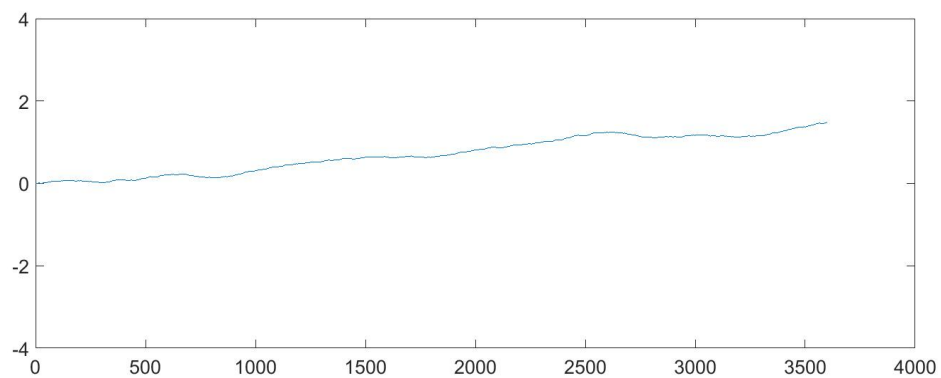
5.6. ábra. $H = 0,5$ - véletlen bolyongás



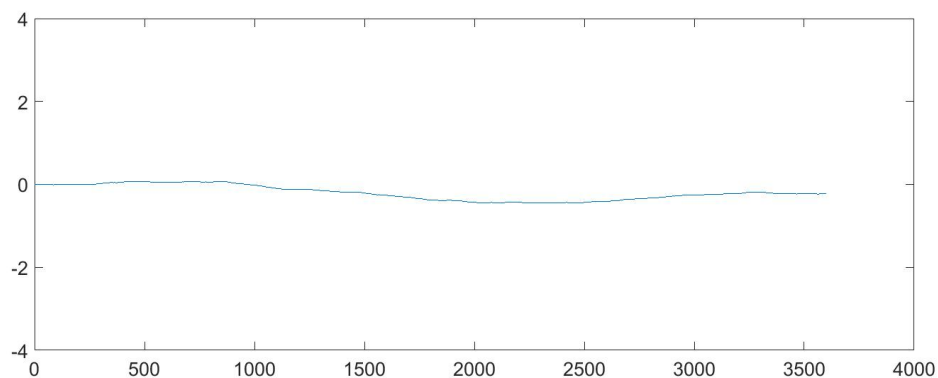
5.7. ábra. $H = 0,6$



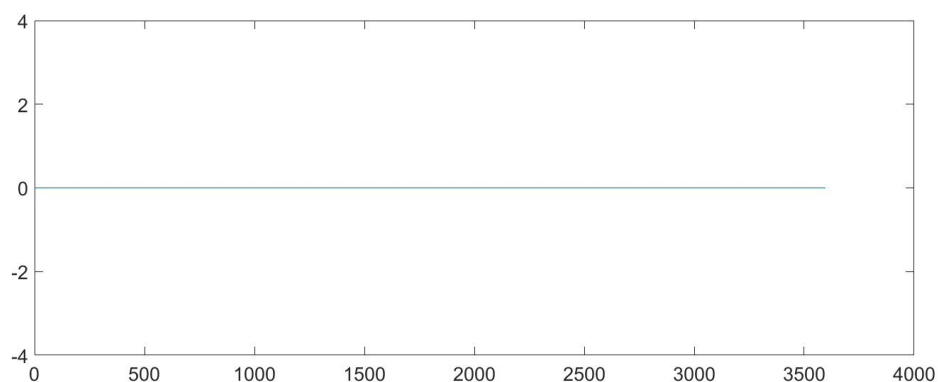
5.8. ábra. $H = 0,7$



5.9. ábra. $H = 0,8$



5.10. ábra. $H = 0,9$



5.11. ábra. $H = 1$ - lineáris görbe

Ezen adatsorok eredményeit tartalmazó táblázat a következő:

5.1. táblázat. Generált adatsorok

	H_{cl}	H_{al1}	H_{al2}	H_{gm1}	H_{gm2}
$H = 0$	0,6092	-0,7572	0,1084	-0,0813	0,0921
$H = 0,1$	1,0009	-0,4831	0,5001	-0,1365	0,2410
$H = 0,2$	0,9666	-0,5004	0,4658	0,2876	0,3203
$H = 0,3$	1,0226	-0,4129	0,5218	0,1660	0,4361
$H = 0,4$	0,9332	-0,4767	0,4324	0,4252	0,3522
$H = 0,5$	1,0030	-0,4362	0,5022	0,5754	0,5582
$H = 0,6$	0,9913	-0,4061	0,4905	0,7281	0,7006
$H = 0,7$	0,9849	-0,3978	0,4841	0,5889	0,6970
$H = 0,8$	0,9640	-0,3951	0,4632	0,8489	0,8061
$H = 0,9$	1,0478	-0,3465	0,5470	0,9565	0,9294
$H = 1$	0,9975	-0,3742	0,4967	0,9941	0,9606

A táblázatról leolvasható, hogy generált adatsorok esetén a klasszikus, az első Anis-Lloyd és a GM1 módszerrel kapott H értékek több esetben tartományon kívül estek, míg a GM2 módszerrel nem kaptunk rossz H értéket, ez a módszer adta a legjobban közelítő Hurst értéket.

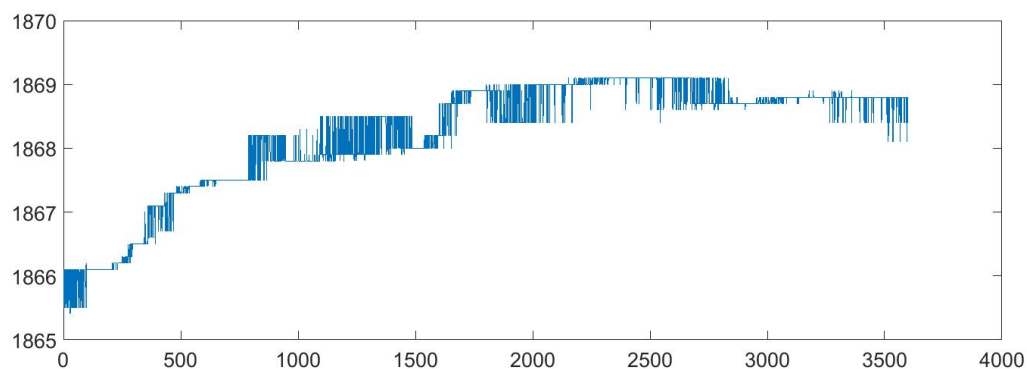
5.2. Biológiai adatsorok

Az ECIS készülékkel mért sejtes és sejtmentes minták különböző zajanalitikai módszerekkel kapott eredményeit tartalmazó EXCEL táblázatokat (8.4, 8.6) a következő fejezetekben kiértékeljük, elemeztük. Illetve mindkét esetben egy-egy konkrét példán keresztül szemléltettük, vizsgáltuk a kapott eredményeket.

5.3. Sejtmentes minták

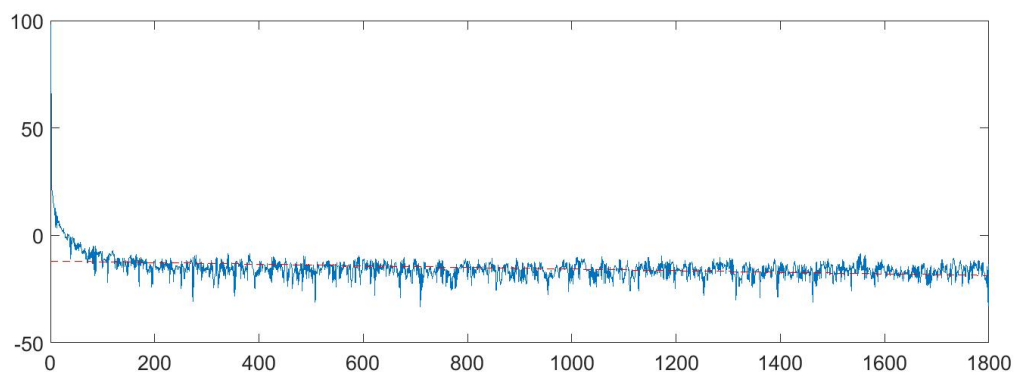
A sejtmentes minták huszonegy mérése közül (8.4) első lépésben kizártuk azokat, amelyek alkalmatlanok voltak a statisztikai elemzésre. Ebben az esetben két mérést kellett kizárnunk. Az első a hetedik mérés, hiszen ebben az esetben negatív GM1 módszerrel számolt Hurst exponenst kaptunk, amely érték nem felel meg a Hurst exponens definíciójának, miszerint $0 < H < 1$. Ennek oka a módszer érzékenysége a mesterséges rezgésre, hiszen nagy a variabilitás, azaz növelve az ablakot csökken a range (értékkészlet, bejárt tartomány). Míg a GM2 módszer ebben az esetben is jó H értéket ad, mivel kiszámításában range helyett a maximum és minimum különbsége szerepel, amely nem érzékeny erre. A tizenharmadik mérés kizárásának oka pedig az 1-nél nagyobb GM1 és GM2 módszerrel számolt Hurst exponens volt. Ez annak köszönhető, hogy a kapott rezisztencia értékek görbéje konvex. A varianciát és az inkrement standard deviációját különböző (4, 16, 64, 256 másodperc) időintervallumokra számoltuk ki. A táblázatból leolvasható, hogy sejtmentes adatsorok esetén a variancia csak legfeljebb tízes nagyságrendű, az inkrement szórása mindig ezernél kisebb minden időintervallumra. Míg az autospektrum $-0,001$ és $-0,01$ közötti érték. A különböző módszerekkel számolt H értékeket összehasonlítva jól látható, hogy a klasszikus (Hcl) és az első Anis-Lloyd ($Hal1$) módszer rossz H értékeket adnak, hiszen nem felelnek meg a definíciónak, tartományon kívüli értéket adnak. Előbbinél majdnem minden esetben 1-nél nagyobb Hurst exponens adódik, míg utóbbinál minden érték negatív. A másik három módszerrel számolt Hurst értékek ($Hal2$, $Hgeo1$, $Hgeo2$) eleget tesznek az előbbi feltételnek.

5.3.1. Példa. *Egy kiválasztott sejtmentes mérés eredményei:*



5.12. ábra. Rezisztencia értékek

A mért rezisztencia (R) értékek 1865 és 1870 között mozogtak.



5.13. ábra. Teljesítménysűrűség-spektrum

Az autospektrum görbéjére illesztett egyenes meredeksége $-0,0037$.

Ennél a példánál a variancia és az inkrement szórásának értékei:

5.2. táblázat. Variancia és inkrement szórása

	4	16	64	256
Var	0,7816	0,7803	0,7863	0,8103
Inc sd	88,0209	176,0431	349,6402	677,9757

Látható, hogy az időintervallumok növekedésével a variancia és az inkrement szórása is növekszik.

És végül a különböző módszerekkel számolt Hurst exponensek:

5.3. táblázat. Hurst exponens

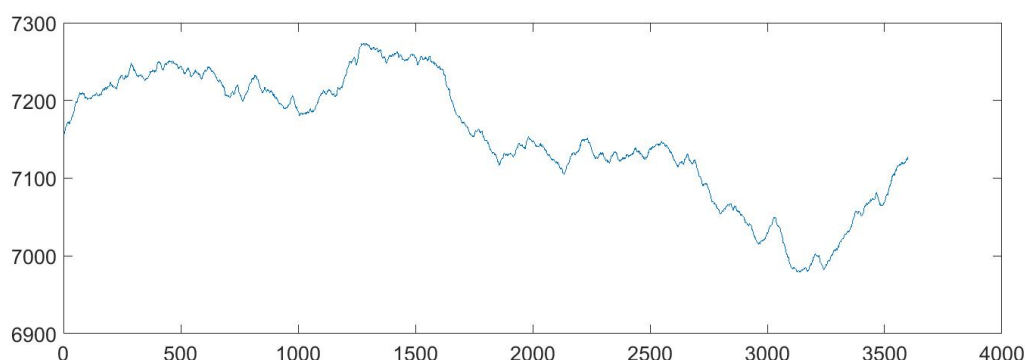
Hcl	Hal1	Hal2	Hgm1	Hgm2
1,2620	-0,4491	0,7612	0,8169	0,4953

Ennél a példánál is azt tapasztaltuk, hogy a klasszikus és az első Anis-Lloyd módszer tartományon kívüli H értéket adott. Egy sejtmentes idősortól azt várjuk, hogy annak Hurst exponense 0,5-höz közeli, amit legpontosabban a GM2 módszer közelít.

5.3.1. Sejtes minták

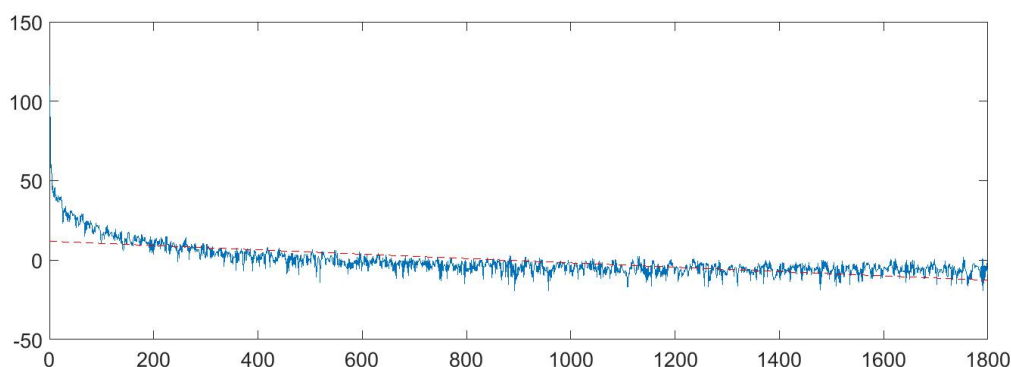
A sejtes minták húsz mérése közül (8.6) az előbbi szempontok szerint haladva most először kizártuk azokat, amelyek alkalmatlanok voltak az általunk vizsgált zajanalitikai módszerekre. Ebben az esetben csak egy mérést kellett kizárnunk, mert az adatsor kevesebb R értéket tartalmazott, mint a többi. Ennek oka a mérési idő rövidsége lehetett. A varianciát és az inkrement standard deviációját ebben az esetben is különböző (4, 16, 64, 256 másodperc) időintervallumokra számoltuk ki. A sejtes adatsorok esetén, a sejtmentes mintáktól eltérően a variancia értékek és az inkrement szórása száz-as-ezres nagyságrendűek minden időintervallumra, és a teljesítménysűrűség-spektrum $-0,010$ és $-0,015$ közötti érték. Ebben az esetben is vizsgáltuk a különböző módszerekkel számolt H értékeket, és hasonlóan az előzőhöz itt is rossz értékek adódtak a klasszikus és az első Anis-Lloyd módszerrel, míg a másik három módszerrel látszólag helyes értékek adódtak.

5.3.2. Példa. *Egy kiválasztott sejtes mérés eredményei:*



5.14. ábra. Rezisztencia értékek

A mért rezisztencia (R) értékek 6900 és 7300 között mozogtak.



5.15. ábra. Teljesítménysűrűség-spektrum

Az autospektrum görbéjére illesztett egyenes meredeksége $-0,0137$.

Ennél a példánál a variancia és az inkrement szórásának értékei:

5.4. táblázat. Variancia és inkrement szórása

	4	16	64	256
Var	6,32E+03	6,35E+03	6,39E+03	6,43E+03
Inc sd	336,6570	673,4113	1,3389e+03	2,6043e+03

Látható, hogy az időintervallumok növekedésével a variancia és az inkrement szórása is növekszik.

És végül a különböző módszerekkel számolt Hurst exponensek:

5.5. táblázat. Hurst exponens

Hcl	Hal1	Hal2	Hgm1	Hgm2
0,97	-0,42	0,47	0,51	0,6236

Ennél a példánál is azt tapasztaltuk, hogy az első Anis-Lloyd módszer tartományon kívüli H értéket adott.

5.4. Sejtes és sejtmentes adatsorok statisztikai összehasonlítása

A két táblázatot (8.4, 8.6) összehasonlítva szembetűnő, hogy sejtes és sejtmentes minták esetén a variancia és az inkrement szórása nagyságrendben eltér egymástól. A Hurst értékeknél már szemmel nem látható a különbség, így ezek összehasonlításában a matematikai statisztika volt segítségünkre. Mivel az irodalomban és az általunk generált adatsorok esetén is a GM2 módszer bizonyult a legalkalmasabb módszernek a Hurst exponens kiszámítására, ezért az így számolt értékek alapján hasonlítottuk össze a sejtes és sejtmentes adatsorokat.

Az R programban beépített függvények segítségével hajtottuk végre a statisztikai elemzést.

Elsőként a normalitás eldöntésére a sejtmentes és a sejtes adatsorok esetén kapott, GM2 módszerrel számolt Hurst exponensekre alkalmaztuk a Shapiro-Wilk tesztet (`shapiro.test(x)`). A sejtmentes esetben 0,825 lett a p értéke, ami lényegesen nagyobb, mint 0,05, így a nullhipotézist fogadjuk el, miszerint az adatsor normális eloszlást követ. A sejtes esetben pedig a kapott p érték 0,5061, azaz ebben az esetben is normális eloszlású az adatsor.

Ezután F-próbával (`var.test(x, y, alternative = "two.sided")`) vizsgáltuk, hogy a két adatsor szórása megegyezik-e. Mivel a kapott p érték ebben az esetben is jóval meghaladta a 0,05 értéket (0,6334), ezért szórásuk ugyanannyinak tekinthető. A feltételek teljesülése után már csak a kétmintás t-próbát kellett alkalmaznunk.

A tesztet ($t.test(x,y, var.equal=TRUE)$) elvégezve eredményül $p = 1,63e-11$ értéket kaptunk, azaz a H_1 hipotézist fogadjuk el. Tehát a Hurst exponens 99,99999999%-os szignifikancia szint mellett jó paraméter annak eldöntésére, hogy az adatsor sejtes vagy sejtmentes mintából származik-e.

5.5. Tumorellenes szerekkel kezelt sejtek mikromigrációja

Az irodalomból ismert, hogy egyes kémiai szerek, például a citotoxikus gyógyszerek alacsony koncentrációban befolyásolják a sejtek mikromigrációját. Az ilyen, enyhe fokú toxicitásnál, az átlag impedancia, vagy ohmikus ellenállás értékben nem detektálható változás, csak a jel fluktuációjának csökkenése utal a sejtek enyhe károsodására. Kíváncsiak voltunk, hogy egy enyhe toxikus kezelés hatással van-e a Hurst exponensre, ezért felvettük a sejtek mikromigrációját közvetlenül a kezelés előtt, a kezelés után egy-két órával, és 24 óra elteltével is. Mivel a kezelés optimális koncentrációja nehezen határozható meg és több mérés sem volt alkalmas analízisre, mert vagy a túlzott toxikus hatás miatt a sejtek elpusztultak, vagy kis koncentrációban nem befolyásolta a kezelés a mikromigrációt, ezért a 16 minta adataiból csak egy reprezentatív példán keresztül vizsgáltuk a paraméterek változását a kezelés előtti, és utáni időpontokban. A sejteket az egyik mérésnél tumorellenes szerrel kezeltük, míg a másik mérésnél csak médiumot alkalmaztunk kontrollként.

5.6. táblázat. Kezelt mikromigráció

t	<i>Hcl</i>	<i>Hal1</i>	<i>Hal2</i>	<i>Hgeo1</i>	<i>Hgeo2</i>	Incsd(4)	Incsd(16)	Incsd(64)	Incsd(256)
0. kezelt	1,07	-0,36	0,57	0,47	0,58	325,46	650,81	1,29E+03	2,51E+03
0. kontroll	0,91	-0,44	0,41	0,67	0,59	322,94	646,13	1,28E+03	2,49E+03
1. kezelt	1,02	-0,36	0,52	1,01	1,01	3,66E+02	7,32E+02	1,46E+03	2,84E+03
1. kontroll	1,01	-0,40	0,51	0,49	0,58	317,80	635,65	1,26E+03	2,45E+03
24. kezelt	1,14	-0,34	0,64	1,03	0,87	102,36	204,71	406,67	789,10
24. kontroll	0,95	-0,42	0,45	0,37	0,55	376,77	753,63	1,50E+03	2,90E+03

A nulladik időpontban, azaz a kezelés előtt az inkrement standard deviációja megegyezett a kezelést követő egy-két órában mért értékével, ami arra utal hogy ilyen rövid távon nem hatott a szer. A 24 óra elteltével újramért értékek inkrementjének szórásában viszont jelentős csökkenés figyelhető meg, ami enyhe citotoxicitást jelent. Habár enyhe toxicitás mellett van értelme mérni a Hurst exponenst, rossz H értékek adódtak, még a GM2 módszerrel is. Így egyelőre a kezelés hatásának vizsgálatánál nem tudtuk egyértelműen igazolni a Hurst változását. Azonban érdemes további méréseket, vizsgálatokat végezni a kezelt mikromigrációról.

6. fejezet

Megbeszélés

Az előző fejezetekben leírtak alapján megállapítottuk, hogy az ismertetett Hurst exponens kiszámítására alkalmas módszerek közül úgy, mint az irodalomban, számunkra is a legpontosabban közelítő értéket adó módszer a GM2. Nem csak a számítógéppel generált adatsoroknál (5.1) bizonyult a legalkalmasabbnak, hanem a sejtmentes és sejtes idősorok (8.4, 8.6) esetén is. Majd az is kiderült számunkra, hogy az ezen módszerrel számolt H értékek ismeretében egy ECIS készülékben felvett ohmikus ellenállás idősróról szignifikánsan eldönthető, hogy az sejtmentes vagy sejtes mintából származik-e. Ennek bizonyítására a matematikai statisztika volt segítségünkre, és az erre használható R program. A statisztikai elemzésből azonban több adatsort is ki kellett zárunk, melynek okait a következő táblázatban foglaltuk össze:

6.1. táblázat. Rossz H értékek

Táblázatban	Matematikai ok	Biológiai magyarázat
Negatív GM1 módszerrel kapott H érték.	<ul style="list-style-type: none"> - Nagy a variabilitás, - Növelve az ablakot, csökken a range, - GM1 módszer érzékeny a mesterséges rezgésre 	<ul style="list-style-type: none"> - Pl.:Valaki kinyitotta az inkubátort VAGY - Az inkubátorba levegő befúvatás történt
1-nél nagyobb GM1 és GM2 módszerrel kapott H érték	<ul style="list-style-type: none"> - Semmi sztochasztikusság, - Majdnem lineáris görbe zajjal, - Növelve az ablakot lineárisan nő a range VAGY - Konvex a függvény 	<ul style="list-style-type: none"> - Túl erős toxikus hatás átlag impedancia csökkenést okoz, - Nem marad meg a mikromigráció
1-nél nagyobb GM1 módszerrel kapott H érték, a GM2 módszerrel kapott H érték 1-hez közeli	<ul style="list-style-type: none"> - Majdnem lineáris görbe zajjal, - Növelve az ablakot lineárisan nő a range, - Átlagban a range várható értékei a két módszerrel azonosak, de GM2 pontosabb 	<ul style="list-style-type: none"> - Túl erős toxikus hatás, - Nem marad meg a mikromigráció VAGY - Sejtmentes médium
Minden H_{cl} érték 1 feletti	Nem alkalmas a módszer	
Minden első Anis-Lloyd módszerrel kapott H érték negatív	Nem alkalmas a módszer	

Habár a második Anis-Lloyd módszer a definíciónak megfelelő értéket ad, ismeretében mégsem tudunk különbséget tenni a sejtmentes és sejtes minták közt. A sejtmentes adatsorok H értékeiről azt vártuk, hogy 0,5 körüliek legyenek, ezzel szemben ez a módszer ettől eltérő értékeket adott.

A kezelt minták esetén azt mondhatjuk, hogy a kezelés hatékony, ha enyhén toxikus. Ellenkező esetben, nagy toxikusság esetén a sejtek elpusztulnak. A táblázatot (5.6) tekintve ez úgy dönthető el, hogy az inkrement standard deviációjának változását figyeljük. Amennyiben a kezelést követően (24 óra elteltével) ez az érték jelentősen lecsökken, akkor hatékonynak mondható a kezelés, van értelme a H -kat összehasonlítani. Mivel, sajnos több számításnál is (köztük az előző fejezetben bemutatott példa esetén is) rossz H értéket kaptunk, ezért a kezelés hatását egyelőre nem tudjuk ilyen módon mérni.

7. fejezet

Következtetések

Diplomamunkámban a hosszú adatsorok elemzésére használható különböző zaj-analitikai módszerek alkalmazhatóságát vizsgáltuk, mind az irodalomban, mind esetünkben, az ECIS készülék által generált biológiai adatsorok esetén. Utóbbihoz MATLAB rutinokat írtunk, melyek eredményeit összefoglaló EXCEL táblázatok (8.4, 8.6) alapján összehasonlítottuk az általunk vizsgált módszereket (variancia, inkrement, teljesítménysűrűség-spektrum és Hurst exponens számítást). Legfőbb célunk az volt, hogy egy olyan paramétert találjunk, melynek ismeretében a sejtmentes és sejtes minták megkülönböztethetőek legyenek. Azt sejtettük, hogy erre a legalkalmasabb paraméter a Hurst exponens, mely számos területen (kardiológia, hidrológia, meteorológia, stb.) használt fontos érték.

Először számítógépen, az R program segítségével generált különböző Hurst exponenssel rendelkező adatsorokat (5.1 - 5.11 ábra) teszteltük az általunk írt H értékek kiszámítására alkalmas rutinokkal (5.1). Ezek az irodalomban talált különböző módszerek úgy, mint a klasszikus R/S-analízis, Anis és Lloyd nevéhez fűződő módszer és ennek módosítása, és két geometriai módszer (GM1, GM2). Generált adatsorok esetén láttuk, hogy a klasszikus és az első Anis-Lloyd módszer megbukott, mivel az általuk kiszámolt H értékek tartományon kívül estek. A második Anis-Lloyd nem tudott különbséget tenni a sejtmentes és sejtes minták közt. A két geometriai módszer közül pedig a második adta a legjobban közelítő értéket.

Ezután már a biológiai mintákat vizsgáltuk a fenti öt különböző módszerrel (8.4, 8.6). Eredményeink megegyeztek az irodalomban és a generált adatsorok esetén tapasztaltakkal. Ekkor is a GM2 bizonyult a legalkalmasabb Hurst exponens kiszámítási módszernek, ezért később erre végeztük el a statisztikai elemzést.

Dolgozatomban, az előbbieken kívül, citotoxikus vegyülettel kezelt sejtes adatsoro-

kat is vizsgáltunk a Hurst exponens tekintetéből (5.6). Azonban hatékony kezelés esetén sok esetben hibás H értékeket kaptunk, amelyből arra következtetünk, hogy a kezelés hatását nem tudjuk mérni a Hurst exponenssel. Azonban érdemes további méréseket, vizsgálatokat végezni a kezelt mikromigrációról.

Legfőbb célunkat, miszerint a sejtmentes és sejtes minták megkülönböztethetőek (H ismeretében) statisztikai elemzéssel bizonyítottuk. Pár hibás mérés kizárása után (melyeket a 6.1 táblázatban rögzítettünk), a megmaradt minták elemzése azt mutatta, hogy a Hurst exponens 99,99999999%-os szignifikancia szint mellett jó paraméter ennek eldöntésére.

8. fejezet

Összefoglalás

Diplomamunkám fő célkitűzése az volt, hogy megvizsgáljuk, egy ECIS készülékben kapott impedancia adatsorról valamilyen matematikai paraméter ismeretében eldönthető-e, hogy az sejtmentes vagy sejtes mintából származik-e.

A Hurst exponens egy olyan matematikai paraméter, melyet számos területen alkalmaznak, úgy, mint a pénzügyben, kardiológiában, hidrológiában, meteorológiában. Azt sejtettük, hogy a különböző zajanalitikai módszerek, például a Hurst-analízis alkalmas lesz a sejtmentes és sejtes minták elkülönítésére. Az irodalomban talált, a Hurst exponens kiszámítására alkalmas módszereket összehasonlítva vizsgáltuk, hogy melyik adja a legjobban közelítő H értéket.

Sejtésünk igazolásához először a számítógép, és azon belül is az R program segítségével adatsorokat generáltunk (lásd 5.1 - 5.11 ábra), majd a Semmelweis Egyetem Genetikai, Sejt- és Immunbiológiai Intézet kutatócsoportja segítségével biológiai adatsorokat vettünk fel impedancia mérésen alapuló ECIS készülékben. Ez az idő függvényében, váltóáram alkalmazása mellett képes a sejtek különböző rezisztencia értékeit mérni. A konfluens sejtréteg jelének fluktuációjára kerestünk magyarázatot, vizsgáltuk mennyire sejtspecifikus a görbe "szöszössége". Hiszen a zaj tartalmazhat információt a biológiai jelenségekről, ami összefüggésben van az ECIS készülék elektródáját borító sejtek mozgásával. Ahhoz, hogy a generált és a biológiai adatsorokat a különböző zajanalitikai módszerekkel vizsgálni tudjuk, MATLAB rutinokat írtunk ezek kiszámolására. A számunkra legfontosabb Hurst rutin eredménye az öt különböző módszerrel számolt H érték, melyeket összehasonlítva, mind a generált (5.1), mind a biológiai adatsorok (8.4, 8.6) esetén azt tapasztaltuk, hogy a klasszikus és az első Anis-Lloyd módszer megbukott, mivel az általuk kiszámolt H értékek tartományon kívül estek, a második Anis-Lloyd nem tudott különbséget tenni a sejtmentes és sejtes minták közt. A két geometriai módszer közül pedig a második (GM2) adta a legjobban

közelítő értéket. Azonban ahhoz, hogy eldöntsük, az ezen módszerrel számolt paraméter ismeretében a sejtes és sejtmentes minták megkülönböztethetők-e, statisztikai elemzést végeztünk. Az elemzést is az R programban hajtottuk végre. A normalitási feltétel és a szórások egyezésének feltétele után kétoldali t-próbát használtunk az adatsorok tesztelésére. Mivel eredményül 0,05-nél jóval kisebb p értéket kaptunk, ezért megállapítottuk, hogy a GM2 módszerrel számolt H érték ismeretében a sejtmentes és a sejtes minták szignifikánsan elkülöníthetők egymástól.

Végül a sejteket tumorellenes vegyülettel kezeltük és így is elvégeztük méréseinket (5.6). Sajnos azt tapasztaltuk, hogy hatékony kezelés esetén is sokszor rossz H értéket kapunk, és a kezelés hatását nem tudjuk jellemezni a Hurst exponens változásával.

Tehát a fő célkitűzésünket sikerült elérnünk, bebizonyítottuk, hogy a Hurst exponens 99,99999999%-os szignifikancia szint mellett jó paraméter annak eldöntésére, hogy az adatsor sejtes vagy sejtmentes mintából származik-e.

Függelék

Az RSana2 MATLAB rutint Sanchez Granero és munkatársai [32] cikke alapján írtuk meg.

```
1 function [RS,ERS]=RSana2(x,n,method)
2
3 if nargin<1 | isempty(x)==1
4     error('You should provide a time series.');
```

```
5 else
6     % x must be a vector
7     if min(size(x))>1
8         error('Invalid time series.');
```

```
9     end
10    x=x(:);
11    % N is the time series length
12    N=length(x);
13 end
14
15 if nargin<2 | isempty(n)==1
16     n=1;
17 else
18     % n must be either a scalar or a vector
19     if min(size(n))>1
20         error('n must be either a scalar or a vector.');
```

```
21     end
22     % n must be integer
23     if n-round(n)~=0
24         error('n must be integer.');
```

```
25     end
26     % n must be positive
27     if n<=0
28         error('n must be positive.');
```

```
29     end
30 end
31
32
33 for i=1:length(n)
34
35     % Calculate the sub-periods
```

```

36     a=floor(N/n(i));
37
38     % Make the sub-periods matrix
39     X=reshape(x(1:a*n(i)),n(i),a);
40
41     switch method
42     case 'Hurst'
43         % Hurst-Mandelbrot variation
44         % Estimate the mean of each sub-period
45         ave=mean(X);
46         % Remove the mean from each sub-period
47         cumdev=X-ones(n(i),1)*ave;
48         % Estimate the cumulative deviation from the mean
49         cumdev=cumsum(cumdev);
50         % Estimate the standard deviation
51         stdev=std(X);
52         % Estiamte the rescaled range
53         rs=(max(cumdev)-min(cumdev))./stdev;
54         clear stdev
55         % Take the logarithm of the mean R/S
56         RS(1,i)=mean(rs);
57
58     case 'Geol'
59         % First Geometric Implementation
60         RS(1,i)=mean(abs(X(n(i),:)-X(1,:)));
61     case 'Geo2'
62         % Second Geometric Implementation
63         RS(1,i)=mean(max(X)-min(X));
64     otherwise
65         error('You should provide another value for "method".');
66     end
67
68
69     if nargout>1
70         vect=1:n(i)-1;
71         if n(i)<=340
72             ERS(1,i)=(n-1/2)/n*gamma((n(i)-1)/2)/(sqrt(pi)*gamma(n(i)
↪ /2))*sum(sqrt(ones(1,n(i)-1)*n(i)-vect./vect));
73         else
74             ERS(1,i)=(n-1/2)/n*1/sqrt(n(i)*pi/2)*sum(sqrt(ones(1,n(i)
↪ -1)*n(i)-vect./vect));
75         end
76     end
77
78 end

```

MATLAB rutin - RSana2

8.1. ábra. Shapiro-Wilk 1. tábla - W értékek

n =	2	3	4	5	6	7	8	9	10	11	12	13	14
a1	0.7071	0.7071	0.6872	0.6646	0.6431	0.6233	0.6052	0.5888	0.5739	0.5601	0.5475	0.5359	0.5251
a2			0.1677	0.2413	0.2806	0.3031	0.3164	0.3244	0.3291	0.3315	0.3325	0.3325	0.3318
a3					0.0875	0.1401	0.1743	0.1976	0.2141	0.2260	0.2347	0.2412	0.2460
a4							0.0561	0.0947	0.1224	0.1429	0.1586	0.1707	0.1802
a5									0.0399	0.0695	0.0922	0.1099	0.1240
a6											0.0803	0.0539	0.0727
a7													0.0240

n =	15	16	17	18	19	20	21	22	23	24	25	26
a1	0.5150	0.5056	0.4968	0.4886	0.4808	0.4734	0.4643	0.4590	0.4542	0.4493	0.4450	0.4407
a2	0.3306	0.3290	0.3273	0.3253	0.3232	0.3211	0.3185	0.3156	0.3126	0.3098	0.3069	0.3043
a3	0.2495	0.2521	0.2540	0.2553	0.2561	0.2565	0.2578	0.2571	0.2563	0.2554	0.2543	0.2533
a4	0.1878	0.1939	0.1988	0.2027	0.2059	0.2085	0.2119	0.2131	0.2139	0.2145	0.2148	0.2151
a5	0.1353	0.1447	0.1524	0.1587	0.1641	0.1686	0.1736	0.1764	0.1787	0.1807	0.1822	0.1836
a6	0.0880	0.1005	0.1109	0.1197	0.1271	0.1334	0.1399	0.1443	0.1480	0.1512	0.1539	0.1563
a7	0.0433	0.0593	0.0725	0.0837	0.0932	0.1013	0.1092	0.1150	0.1201	0.1245	0.1283	0.1316
a8		0.0196	0.0359	0.0496	0.0612	0.0711	0.0804	0.0878	0.0941	0.0997	0.1046	0.1089
a9				0.0163	0.0303	0.0422	0.0530	0.0618	0.0696	0.0764	0.0823	0.0876
a10						0.0140	0.0263	0.0368	0.0459	0.0539	0.0610	0.0672
a11								0.0122	0.0228	0.0321	0.0403	0.0476
a12									0.0000	0.0107	0.0200	0.0284
a13											0.0000	0.0094

n =	15	16	17	18	19	20	21	22	23	24	25	26
a1	0.5150	0.5056	0.4968	0.4886	0.4808	0.4734	0.4643	0.4590	0.4542	0.4493	0.4450	0.4407
a2	0.3306	0.3290	0.3273	0.3253	0.3232	0.3211	0.3185	0.3156	0.3126	0.3098	0.3069	0.3043
a3	0.2495	0.2521	0.2540	0.2553	0.2561	0.2565	0.2578	0.2571	0.2563	0.2554	0.2543	0.2533
a4	0.1878	0.1939	0.1988	0.2027	0.2059	0.2085	0.2119	0.2131	0.2139	0.2145	0.2148	0.2151
a5	0.1353	0.1447	0.1524	0.1587	0.1641	0.1686	0.1736	0.1764	0.1787	0.1807	0.1822	0.1836
a6	0.0880	0.1005	0.1109	0.1197	0.1271	0.1334	0.1399	0.1443	0.1480	0.1512	0.1539	0.1563
a7	0.0433	0.0593	0.0725	0.0837	0.0932	0.1013	0.1092	0.1150	0.1201	0.1245	0.1283	0.1316
a8		0.0196	0.0359	0.0496	0.0612	0.0711	0.0804	0.0878	0.0941	0.0997	0.1046	0.1089
a9				0.0163	0.0303	0.0422	0.0530	0.0618	0.0696	0.0764	0.0823	0.0876
a10						0.0140	0.0263	0.0368	0.0459	0.0539	0.0610	0.0672
a11								0.0122	0.0228	0.0321	0.0403	0.0476
a12									0.0000	0.0107	0.0200	0.0284
a13											0.0000	0.0094

n =	27	28	29	30	31	32	33	34	35	36	37	38
a1	0.4366	0.4328	0.4291	0.4254	0.4220	0.4188	0.4156	0.4127	0.4096	0.4068	0.4040	0.4015
a2	0.3018	0.2992	0.2968	0.2944	0.2921	0.2898	0.2876	0.2854	0.2834	0.2813	0.2794	0.2774
a3	0.2522	0.2510	0.2499	0.2487	0.2475	0.2463	0.2451	0.2439	0.2427	0.2415	0.2403	0.2391
a4	0.2152	0.2151	0.2150	0.2148	0.2145	0.2141	0.2137	0.2132	0.2127	0.2121	0.2116	0.2110
a5	0.1848	0.1857	0.1864	0.1870	0.1874	0.1878	0.1880	0.1882	0.1883	0.1883	0.1883	0.1881
a6	0.1584	0.1601	0.1616	0.1630	0.1641	0.1651	0.1660	0.1667	0.1673	0.1678	0.1683	0.1686
a7	0.1346	0.1372	0.1395	0.1415	0.1433	0.1449	0.1463	0.1475	0.1487	0.1496	0.1505	0.1513
a8	0.1128	0.1162	0.1192	0.1219	0.1243	0.1265	0.1284	0.1301	0.1317	0.1331	0.1344	0.1356
a9	0.0923	0.0965	0.1002	0.1036	0.1066	0.1093	0.1118	0.1140	0.1160	0.1179	0.1196	0.1211
a10	0.0728	0.0778	0.0822	0.0862	0.0899	0.0931	0.0961	0.0988	0.1013	0.1036	0.1056	0.1075
a11	0.0540	0.0598	0.0650	0.0697	0.0739	0.0777	0.0812	0.0844	0.0873	0.0900	0.0924	0.0947
a12	0.0358	0.0424	0.0483	0.0537	0.0585	0.0629	0.0669	0.0706	0.0739	0.0770	0.0798	0.0824
a13	0.0178	0.0253	0.0320	0.0381	0.0435	0.0485	0.0530	0.0572	0.0610	0.0645	0.0677	0.0706
a14	0.0000	0.0084	0.0159	0.0227	0.0289	0.0344	0.0395	0.0441	0.0484	0.0523	0.0559	0.0592
a15			0.0000	0.0076	0.0144	0.0206	0.0262	0.0314	0.0361	0.0404	0.0444	0.0481
a16					0.0000	0.0068	0.0131	0.0187	0.0239	0.0287	0.0331	0.0372
a17							0.0000	0.0062	0.0119	0.0172	0.0220	0.0264
a18									0.0000	0.0057	0.0110	0.0158
a19											0.0000	0.0053

n =	39	40	41	42	43	44	45	46	47	48	49	50
a1	0.3989	0.3964	0.3940	0.3917	0.3894	0.3872	0.3850	0.3830	0.3808	0.3789	0.3770	0.3751
a2	0.2755	0.2737	0.2719	0.2701	0.2684	0.2667	0.2651	0.2635	0.2620	0.2604	0.2589	0.2574
a3	0.2380	0.2368	0.2357	0.2345	0.2334	0.2323	0.2313	0.2302	0.2291	0.2281	0.2271	0.2260
a4	0.2104	0.2098	0.2091	0.2085	0.2078	0.2072	0.2065	0.2058	0.2052	0.2045	0.2038	0.2032
a5	0.1880	0.1878	0.1876	0.1874	0.1871	0.1868	0.1865	0.1862	0.1859	0.1855	0.1851	0.1847
a6	0.1689	0.1691	0.1693	0.1694	0.1695	0.1695	0.1695	0.1695	0.1695	0.1693	0.1692	0.1691
a7	0.1520	0.1526	0.1531	0.1535	0.1539	0.1542	0.1545	0.1548	0.1550	0.1551	0.1553	0.1554
a8	0.1366	0.1376	0.1384	0.1392	0.1398	0.1405	0.1410	0.1415	0.1420	0.1423	0.1427	0.1430
a9	0.1225	0.1237	0.1249	0.1259	0.1269	0.1278	0.1286	0.1293	0.1300	0.1306	0.1312	0.1317
a10	0.1092	0.1108	0.1123	0.1136	0.1149	0.1160	0.1170	0.1180	0.1189	0.1197	0.1205	0.1212
a11	0.0967	0.0986	0.1004	0.1020	0.1035	0.1049	0.1062	0.1073	0.1085	0.1095	0.1105	0.1113
a12	0.0848	0.0870	0.0891	0.0909	0.0927	0.0943	0.0959	0.0972	0.0986	0.0998	0.1010	0.1020
a13	0.0733	0.0759	0.0782	0.0804	0.0824	0.0842	0.0860	0.0876	0.0892	0.0906	0.9190	0.0932
a14	0.0622	0.0651	0.0677	0.0701	0.0724	0.0745	0.0765	0.0783	0.0801	0.0817	0.0832	0.0846
a15	0.0515	0.0546	0.0575	0.0602	0.0628	0.0651	0.0673	0.0694	0.0713	0.0731	0.0748	0.0764
a16	0.0409	0.0444	0.0476	0.0506	0.0534	0.0560	0.0584	0.0607	0.0628	0.0648	0.0667	0.0685
a17	0.0305	0.0343	0.0379	0.0411	0.0442	0.0471	0.0497	0.0522	0.0546	0.0568	0.0588	0.0608
a18	0.0203	0.0244	0.0283	0.0318	0.0352	0.0383	0.0412	0.0439	0.0465	0.0489	0.0511	0.0532
a19	0.0101	0.0146	0.0188	0.0227	0.0263	0.0296	0.0328	0.0357	0.0385	0.0411	0.0436	0.0459
a20	0.0000	0.0049	0.0094	0.0136	0.0175	0.0211	0.0245	0.0277	0.0307	0.0335	0.0361	0.0386
a21			0.0000	0.0045	0.0087	0.0126	0.0163	0.0197	0.0229	0.0259	0.0288	0.0314
a22					0.0000	0.0042	0.0081	0.0118	0.0153	0.0185	0.0215	0.0244
a23							0.0000	0.0039	0.0076	0.0111	0.0143	0.0174
a24									0.0000	0.0037	0.0071	0.0104
a25											0.0000	0.0035

8.2. ábra. Shapiro-Wilk 2. tábla - p értékek

n \ P	0.01	0.02	0.05	0.1	0.5	0.9	0.95	0.98	0.99
3	0.753	0.756	0.767	0.789	0.959	0.998	0.999	1.000	1.000
4	0.687	0.707	0.748	0.792	0.935	0.987	0.992	0.996	0.997
5	0.686	0.715	0.762	0.806	0.927	0.979	0.986	0.991	0.993
6	0.713	0.743	0.788	0.826	0.927	0.974	0.981	0.986	0.989
7	0.730	0.760	0.803	0.838	0.928	0.972	0.979	0.985	0.988
8	0.749	0.778	0.818	0.851	0.932	0.972	0.978	0.984	0.987
9	0.764	0.791	0.829	0.859	0.935	0.972	0.978	0.984	0.986
10	0.781	0.806	0.842	0.869	0.938	0.972	0.978	0.983	0.986
11	0.792	0.817	0.850	0.876	0.940	0.973	0.979	0.984	0.986
12	0.805	0.828	0.859	0.883	0.943	0.973	0.979	0.984	0.986
13	0.814	0.837	0.866	0.889	0.945	0.974	0.979	0.984	0.986
14	0.825	0.846	0.874	0.895	0.947	0.975	0.980	0.984	0.986
15	0.835	0.855	0.881	0.901	0.950	0.975	0.980	0.984	0.987
16	0.844	0.863	0.887	0.906	0.952	0.976	0.981	0.985	0.987
17	0.851	0.869	0.892	0.910	0.954	0.977	0.981	0.985	0.987
18	0.858	0.874	0.897	0.914	0.956	0.978	0.982	0.986	0.988
19	0.863	0.879	0.901	0.917	0.957	0.978	0.982	0.986	0.988
20	0.868	0.884	0.905	0.920	0.959	0.979	0.983	0.986	0.988
21	0.873	0.888	0.908	0.923	0.960	0.980	0.983	0.987	0.989
22	0.878	0.892	0.911	0.926	0.961	0.980	0.984	0.987	0.989
23	0.881	0.895	0.914	0.928	0.962	0.981	0.984	0.987	0.989
24	0.884	0.898	0.916	0.930	0.963	0.981	0.984	0.987	0.989
25	0.888	0.901	0.918	0.931	0.964	0.981	0.985	0.988	0.989
26	0.891	0.904	0.920	0.933	0.965	0.982	0.985	0.988	0.989
27	0.894	0.906	0.923	0.935	0.965	0.982	0.985	0.988	0.990
28	0.896	0.908	0.924	0.936	0.966	0.982	0.985	0.988	0.990
29	0.898	0.910	0.926	0.937	0.966	0.982	0.985	0.988	0.990
30	0.900	0.912	0.927	0.939	0.967	0.983	0.985	0.988	0.990
31	0.902	0.914	0.929	0.940	0.967	0.983	0.986	0.988	0.990
32	0.904	0.915	0.930	0.941	0.968	0.983	0.986	0.988	0.990
33	0.906	0.917	0.931	0.942	0.968	0.983	0.986	0.989	0.990
34	0.908	0.919	0.933	0.943	0.969	0.983	0.986	0.989	0.990
35	0.910	0.920	0.934	0.944	0.969	0.984	0.986	0.989	0.990
36	0.912	0.922	0.935	0.945	0.970	0.984	0.986	0.989	0.990
37	0.914	0.924	0.936	0.946	0.970	0.984	0.987	0.989	0.990
38	0.916	0.925	0.938	0.947	0.971	0.984	0.987	0.989	0.990
39	0.917	0.927	0.939	0.948	0.971	0.984	0.987	0.989	0.991
40	0.919	0.928	0.940	0.949	0.972	0.985	0.987	0.989	0.991
41	0.920	0.929	0.941	0.950	0.972	0.985	0.987	0.989	0.991
42	0.922	0.930	0.942	0.951	0.972	0.985	0.987	0.989	0.991
43	0.923	0.932	0.943	0.951	0.973	0.985	0.987	0.990	0.991
44	0.924	0.933	0.944	0.952	0.973	0.985	0.987	0.990	0.991
45	0.926	0.934	0.945	0.953	0.973	0.985	0.988	0.990	0.991
46	0.927	0.935	0.945	0.953	0.974	0.985	0.988	0.990	0.991
47	0.928	0.936	0.946	0.954	0.974	0.985	0.988	0.990	0.991
48	0.929	0.937	0.947	0.954	0.974	0.985	0.988	0.990	0.991
49	0.929	0.939	0.947	0.955	0.974	0.985	0.988	0.990	0.991
50	0.930	0.938	0.947	0.955	0.974	0.985	0.988	0.990	0.991

8.1. táblázat. Összehasonlító táblázat (R/S analízis, Anis és Lloyd 1., Anis és Lloyd 2.)

	R/S analízis	Anis és Lloyd 1.	Anis és Lloyd 2.
Kiszámítása	l szakaszok hossza, s szakaszok száma, E_i mért rezisztencia értékek, $m = \frac{E_1+E_2+\dots+E_l}{l}$, $D_i = E_i - m$, $P_t = \sum_{i=1}^t D_i$ $(t = 1, 2, \dots, l)$, $R(l) = \max(P_1, \dots, P_l)$ $-\min(P_1, \dots, P_l)$, $S(l) = \sqrt{\frac{1}{l} \cdot \sum_{i=1}^l (E_i - m)^2}$, $H = \frac{\log(R/S)_l - \log c}{\log l}$	l szakaszok hossza, s szakaszok száma, E_i mért rezisztencia értékek, $m = \frac{E_1+E_2+\dots+E_l}{l}$, $D_i = E_i - m$, $P_t = \sum_{i=1}^t D_i$ $(t = 1, 2, \dots, l)$, $R(l) = \max(P_1, \dots, P_l)$ $-\min(P_1, \dots, P_l)$, $S(l) = \sqrt{\frac{1}{l} \cdot \sum_{i=1}^l (E_i - m)^2}$, $\mathbb{E}(R/S)_l = \begin{cases} \frac{l-\frac{1}{2}}{l} \cdot \frac{\Gamma(\frac{l-1}{2})}{\sqrt{\pi} \cdot \Gamma(\frac{l}{2})} \cdot \sum_{i=1}^{l-1} \sqrt{\frac{l-i}{i}} & \text{ha } l \leq 340 \\ \frac{l-\frac{1}{2}}{l} \cdot \frac{1}{\sqrt{l \cdot \frac{\pi}{2}}} \cdot \sum_{i=1}^{l-1} \sqrt{\frac{l-i}{i}} & \text{ha } l \geq 340 \end{cases}$ $\log H_l = \log(R/S)_l - \log(\mathbb{E}(R/S)_l) + \log(l)/2$, $H = \frac{\log H_l - \log c}{\log l}$	l szakaszok hossza, s szakaszok száma, E_i mért rezisztencia értékek, $m = \frac{E_1+E_2+\dots+E_l}{l}$, $D_i = E_i - m$, $P_t = \sum_{i=1}^t D_i$ $(t = 1, 2, \dots, l)$, $R(l) = \max(P_1, \dots, P_l)$ $-\min(P_1, \dots, P_l)$, $S(l) = \sqrt{\frac{1}{l} \cdot \sum_{i=1}^l (E_i - m)^2}$, $\mathbb{E}(R/S)_l = \begin{cases} \frac{l-\frac{1}{2}}{l} \cdot \frac{\Gamma(\frac{l-1}{2})}{\sqrt{\pi} \cdot \Gamma(\frac{l}{2})} \cdot \sum_{i=1}^{l-1} \sqrt{\frac{l-i}{i}} & \text{ha } l \leq 340 \\ \frac{l-\frac{1}{2}}{l} \cdot \frac{1}{\sqrt{l \cdot \frac{\pi}{2}}} \cdot \sum_{i=1}^{l-1} \sqrt{\frac{l-i}{i}} & \text{ha } l \geq 340 \end{cases}$ $\log H_l = \log(R/S)_l - \log(\mathbb{E}(R/S)_l) + \log(l)/2$, $H = \frac{\log H_l - \log c}{\log l}$
Hátránya	<ol style="list-style-type: none"> 1. Brown mozgás esetén kicsi l érték (szakaszhossz) esetén rossz H értéket ad. 2. Kis l értékre előfordulhat, hogy a standard deviáció nulla lesz. 3. Érzékeny a kerekítésre a standard deviáció miatt. 	H értéke negatív lehet.	Ha a zaj nem normális eloszlású (nálunk nem normális eloszlású), akkor nem garantált, hogy helyes H értéket kapunk.

8.2. táblázat. Összehasonlító táblázat (GM1 és GM2 módszer)

	GM1 módszer	GM2 módszer
Kiszámítása	l szakaszok hossza, s szakaszok száma, $\overline{\Delta B} \propto l^H$, $\overline{\Delta B} = \overline{ B(t+l) - B(t) }$, \propto arányszám, $\forall m = 1, \dots, s$ -re $S_m = x_{ml} - x_{(m-1)l+1}$ $H_l = \text{mean}\{S_m : m = 1, \dots, s\}$ $\log H_l = \log c + H \log l$ $H = \frac{\log H_l - \log c}{\log l}$	l szakaszok hossza, s szakaszok száma, $\frac{\text{range}(B)}{\text{range}(B)} \propto l^H$, $\frac{\text{range}(B)}{\text{range}(B)} =$ $\frac{\max\{B(S) : t \leq S \leq t+l\}}{\min\{B(S) : t \leq S \leq t+l\}}$, \propto arányszám
Hátránya	Helyes H értéket ad, de nem a legpontosabbat.	Ennél a módszernél a standard deviáció kisebb, mint a GM1 módszernél, ezért pontosabb eredményt ad.

8.3. táblázat. Sejtmentes adatsorok első része

	Szórás (4)	Szórás (16)	Szórás (64)	Szórás (256)	Hcl	Hal1	Hal2	Hgeo1	Hgeo2
1.	0.4208	0.4123	0.4093	0.4104	1.1591	-0.5538	0.6583	0.4541	0.3126
2.	0.4768	0.4682	0.4682	0.4731	1.1720	-0.5533	0.6713	0.3991	0.3155
3.	0.6303	0.6270	0.6294	0.6403	1.2766	-0.3844	0.7758	0.5472	0.4577
4.	0.3726	0.3662	0.3660	0.3728	1.1918	-0.6119	0.6909	0.6867	0.2331
5.	0.8841	0.8833	0.8868	0.9002	1.262	-0.4491	0.7611	0.8168	0.4953
6.	0.7934	0.7897	0.7939	0.8112	1.2461	-0.4788	0.7452	0.7697	0.4651
7.	0.4028	0.3952	0.3945	0.3917	1.1566	-0.5138	0.6556	-0.2051	0.2497
8.	0.4124	0.4062	0.4057	0.4113	1.1353	-0.5343	0.6344	0.6230	0.3563
9.	0.2661	0.2599	0.2592	0.2612	1.2939	-0.5191	0.7931	0.4733	0.3967
10.	0.4353	0.4312	0.4309	0.4348	1.1549	-0.5543	0.6541	0.4642	0.3260
11.	0.3917	0.3829	0.3806	0.3821	1.1951	-0.5326	0.6943	0.5467	0.3866
12.	0.4051	0.4027	0.4033	0.4096	1.1711	-0.5117	0.6703	0.2398	0.3831
13.	41.7453	41.7998	42.0714	43.0640	1.0217	-0.3603	0.5209	1.0336	1.0101
14.	1.4505	1.4314	1.4354	1.4630	1.3979	-0.3839	0.8971	0.6389	0.3549
15.	4.0050	4.0012	4.0254	4.1147	1.2315	-0.3513	0.7307	0.9433	0.6461
16.	0.1622	0.1527	0.1434	0.1158	0.9592	-0.6362	0.4583	0.5446	0.5206
17.	3.4712	3.4659	3.4867	3.5696	1.2656	-0.3633	0.7648	0.9313	0.6546
18.	2.2266	2.2256	2.2380	2.2839	1.3008	-0.3603	0.8000	0.7060	0.6073
19.	1.9182	1.9096	1.9181	1.9485	1.3689	-0.3604	0.8681	0.4655	0.4532
20.	1.3880	1.3849	1.3914	1.4226	1.1685	-0.5111	0.6677	0.9817	0.4806
21.	1.9722	1.9639	1.9730	2.0189	1.1662	-0.3894	0.6654	0.9345	0.5858
átlag					1.2045	-0.4721	0.7037	0.6188	0.4615
szórás					0.1017	0.0916	0.1017	0.2884	0.1748

8.4. táblázat. Sejtmentes adatsorok második része

	Var (4)	Var (16)	Var (64)	Var (256)	Inc sd (4)	Inc sd (16)	Inc sd (64)	Inc sd (256)	Autospektrum
1.	0.1771	0.1700	0.1676	0.1685	81.2077	162.4092	322.6005	625.6559	-0.0019
2.	0.2274	0.2193	0.2192	0.2239	95.6835	191.3544	380.0760	737.1007	-0.0022
3.	0.3975	0.3933	0.3962	0.4101	77.1891	154.3748	306.6248	594.6959	-0.0030
4.	0.1389	0.1341	0.1341	0.1390	87.9298	175.8610	349.2854	677.3286	-0.0019
5.	0.7816	0.7803	0.7863	0.8103	88.0209	176.04311	349.6402	677.9757	-0.0037
6.	0.6296	0.6237	0.6303	0.6580	76.3395	152.6793	303.2704	588.2197	-0.0041
7.	0.1624	0.1562	0.1557	0.1535	97.3881	194.7762	386.8674	750.3227	-0.0025
8.	0.1702	0.1651	0.1646	0.1691	79.6728	159.3409	316.4853	613.7940	-0.0027
9.	0.0708	0.0676	0.0672	0.0683	83.6296	167.2715	332.2477	644.3157	-0.0024
10.	0.1895	0.1860	0.1857	0.1891	76.6694	153.3656	304.6284	590.7605	-0.0037
11.	0.1535	0.1466	0.1449	0.1460	80.1654	160.3278	318.4496	617.6146	-0.0037
12.	0.1641	0.1622	0.1627	0.1678	84.7686	169.5451	336.7756	653.1490	-0.0028
13.	1.74E+03	1.75E+03	1.77E+03	1.85E+03	125.5290	251.1540	499.3141	971.6429	-0.0123
14.	2.1041	2.0489	2.0603	2.1402	116.3646	232.7446	462.3049	896.7545	-0.0025
15.	16.0399	16.0094	16.2035	16.9311	116.1356	232.2677	461.3731	895.1926	-0.0076
16.	0.0263	0.0233	0.0206	0.0134	113.1086	226.2053	449.3123	871.4300	-0.0033
17.	12.0489	12.0124	12.1572	12.7423	111.1590	222.3189	441.6230	856.8143	-0.0071
18.	4.9576	4.9532	5.0084	5.2161	112.0856	224.1714	445.3026	863.7328	-0.0053
19.	3.6796	3.6467	3.6792	3.7967	110.4746	220.9598	438.8901	851.2994	-0.0051
20.	1.9265	1.9179	1.9360	2.0238	109.1536	218.2987	433.5941	841.0460	-0.0032
21.	3.8898	3.8568	3.8928	4.0761	107.1995	214.4136	425.9020	826.1634	-0.0064

8.5. táblázat. Sejtes adatsorok első rész

	Szórás (4)	Szórás (16)	Szórás (64)	Szórás (256)	Hcl	Hal1	Hal2	Hgeo1	Hgeo2
1.	18.89891	18.9181	18.9670	18.6953	1.0063	-0.3926	0.5054	0.5402	0.7228
2.	120.5754	120.4479	121.166	123.8944	0.9700	-0.4043	0.4692	0.8139	0.7806
3.	88.8376	89.0370	89.4301	90.9501	1.0411	-0.3775	0.5403	0.7304	0.6981
4.	73.5739	73.7369	73.7159	72.9556	0.9750	-0.4123	0.4743	0.3080	0.6021
5.	74.1366	74.2947	74.3439	71.6747	0.9556	-0.4194	0.4547	0.4410	0.6307
6.	56.6550	56.6131	56.3489	54.6691	0.9839	-0.3895	0.4829	0.3235	0.5281
7.	98.5827	98.7356	99.2151	100.8682	0.9497	-0.4131	0.4490	0.7883	0.7248
8.	79.4716	79.6779	79.9444	80.1905	0.9728	-0.4159	0.4721	0.5079	0.6236
9.	64.2205	64.1548	63.9898	62.7918	0.9911	-0.4027	0.4904	0.4108	0.5344
10.	24.2608	24.221	23.8379	22.1294	0.9903	-0.4078	0.4895	0.1736	0.5474
11.	121.1915	121.1657	121.6437	122.5998	0.9404	-0.4349	0.4395	0.5592	0.6366
12.	73.8623	73.9448	73.7741	72.5010	0.9771	-0.4123	0.4763	0.6574	0.6753
13.	176.0801	176.3190	177.2323	180.2062	1.0129	-0.375	0.5121	0.5225	0.7153
14.	40.4834	40.4562	40.6963	41.6473	0.9572	-0.4161	0.4563	0.8025	0.7306
15.	56.2034	55.9921	56.3023	57.4181	0.9802	-0.4139	0.4794	0.8938	0.7524
16.	43.1443	43.2618	43.4288	44.1041	0.9491	-0.4223	0.4482	0.7230	0.6329
17.	24.913	24.5753	24.6404	24.7331	0.9055	-0.4517	0.4047	0.6109	0.5858
18.	28.6842	28.6431	28.7009	28.5991	1.1069	-0.3371	0.6060	0.5655	0.7147
19.	34.3485	34.3868	34.4988	34.9844	0.9979	-0.3992	0.4970	0.5193	0.6491
20.	31.8283	31.9116	31.9697	31.8096	0.9980	-0.3869	0.4971	0.7699	0.6777
átlag					0.9831	-0.4043	0.4823	0.5831	0.6582
szórás					0.0414	0.0242	0.0414	0.1918	0.0735

8.6. táblázat. Sejtes adatsorok második rész

	Var (4)	Var (16)	Var (64)	Var (256)	Inc sd (4)	Inc sd (16)	Inc sd (64)	Inc sd (256)	Autospektrum
1.	3.57E+02	3.58E+02	3.60E+02	3.50E+02	196.5766	393.2121	7.81E+02	1.52E+03	-0.0117
2.	1.45E+04	1.45E+04	1.47E+04	1.54E+04	265.5649	531.2925	1.06E+03	2.06E+03	-0.0124
3.	7.93E+03	7.93E+03	8.00E+03	8.27E+03	326.3116	652.2768	1.29E+03	2.50E+03	-0.0125
4.	5.41E+03	5.44E+03	5.43E+03	5.32E+03	425.2785	849.7361	1.69E+03	3.26E+03	-0.0141
5.	5.50E+03	5.52E+03	5.53E+03	5.14E+03	398.6423	796.6970	1.58E+03	3.06E+03	-0.0136
6.	3.21E+03	3.21E+03	3.18E+03	2.99E+03	445.6792	890.7912	1.77E+03	3.45E+03	-0.0129
7.	9.72E+03	9.75E+03	9.84E+03	1.02E+04	379.2741	758.3781	1.51E+03	2.91E+03	-0.0124
8.	6.32E+03	6.35E+03	6.39E+03	6.43E+03	336.6570	673.4113	1.34E+03	2.60E+03	-0.0137
9.	4.12E+03	4.12E+03	4.09E+03	3.94E+03	435.8667	872.1788	1.73E+03	3.38E+03	-0.0135
10.	5.89E+02	5.87E+02	5.68E+02	4.90E+02	301.1747	602.7510	1.20E+03	2.32E+03	-0.0124
11.	1.47E+04	1.47E+04	1.48E+04	1.50E+04	422.6456	844.7213	1.68E+03	3.24E+03	-0.0126
12.	5.46E+03	5.47E+03	5.44E+03	5.26E+03	423.0041	846.0850	1.68E+03	3.26E+03	-0.0129
13.	3.10E+04	3.11E+04	3.14E+04	3.25E+04	406.0202	811.6354	1.61E+03	3.15E+03	-0.0124
14.	1.64E+03	1.64E+03	1.66E+03	1.73E+03	299.5818	599.1416	1.19E+03	2.30E+03	
15.	3.16E+03	3.14E+03	3.17E+03	3.30E+03	395.4922	791.1182	1.57E+03	3.04E+03	-0.0123
16.	1.86E+03	1.87E+03	1.89E+03	1.95E+03	387.3223	774.3608	1.54E+03	2.98E+03	-0.0124
17.	6.21E+02	6.04E+02	6.07E+02	6.12E+02	402.7130	805.2970	1.60E+03	3.10E+03	-0.0123
18.	8.23E+02	8.20E+02	8.24E+02	8.18E+02	395.2316	790.3740	1.57E+03	3.05E+03	-0.0117
19.	1.18E+03	1.18E+03	1.19E+03	1.22E+03	375.9278	751.8498	1.49E+03	2.90E+03	-0.0114
20.	1.01E+03	1.02E+03	1.02E+03	1.01E+03	360.1342	720.2191	1.43E+03	2.78E+03	-0.0119

Irodalomjegyzék

- [1] Giaever and Keese: A morphological biosensor for mammalian cells, *Nature* 366(1993)591-2
- [2] Dávid Gergely, *Tumorsejtek elektromos jelének vizsgálata ECIS készülékben*, Szakdolgozat (2011.)
- [3] *Az ECIS készülék hivatalos honlapja*,
<http://www.biophysics.com/ecis-theory.php>
- [4] Láng, Orsolya, László Kőhidai, and Joachim Wegener. *Label-free profiling of cell dynamics: A sequence of impedance-based assays to estimate tumor cell invasiveness in vitro*. *Experimental cell research* 359.1 (2017): 243-250.
- [5] C.M. Lo, C.R. Keese, I. Giaever, *pH changes in pulsed CO₂ incubators cause periodic changes in cell morphology*, *Exp. Cell Res.* 213 (1994) 391?397.
- [6] D.C. Lovelady, J. Friedman, S. Patel, D.A. Rabson, C.M. Lo, *Detecting effects of low levels of cytochalasin B in 3T3 fibroblast cultures by analysis of electrical noise obtained from cellular micromotion*, *Biosens. Bioelectron.* 24 (2009) 2250?2254
- [7] C.-M. Lo, C.R. Keese, I. Giaever, *Monitoring motion of confluent cells in tissue culture*, *Exp. Cell Res.* 204 (1993) 102?109.
- [8] Montgomery, D.C. and Runger, G.C. (1994) *Applied statistics and probability for engineers*
- [9] Williams, Byron K. "The use of analysis of variance procedures in biological studies." *Applied Stochastic Models in Business and Industry* 3.4 (1987): 207-226.
- [10] *Autospektrum*,
<https://www.tankonyvtar.hu/hu/tartalom/tamop412A/2010-0013-merestechnika/9-4-autospektrum.html>
- [11] Gombos Kitti Kata, *Fraktálok a tőzsdén*, Szakdolgozat, Szegedi Tudományegyetem (2010.)
- [12] *Estimating the Hurst Exponent*,
<http://www.bearcave.com/misl/misltech/wavelets/hurst/index.html>
- [13] *Herpai Nándor Nílusról szóló blogja*,
<https://herpainandor.com/2013/06/22/nilus-es-a-tozsde/>

- [14] *Fraktálok, a fraktáldimenzió*,
<http://www.geo.u-szeged.hu/~joe/pub/Tamop/Jegyzet/ch09s03.html>
- [15] Somogyi Balázs István *Mérsékelt kockázatú befektetési alapok vagyonának és árfolyamának stabilitása és hosszú emlékezete* OTDK, Debreceni Egyetem Gazdálkodástudományi és Vidékfejlesztési Kar
- [16] Costa, Tommaso, Dario Galati, and Elena Rognoni. *The Hurst exponent of cardiac response to positive and negative emotional film stimuli using wavelet*. Autonomic Neuroscience: Basic and Clinical 151.2 (2009): 183-185.
- [17] Larsen, R.J., Diener, E., 1992. *Problems and promises with the circumplex model of emotion*. Rev. Person. Soc. Psychol. 13, 25?59.
- [18] Russell, J.A., 1980. *A circumplex model of affect*. J. Personal. Soc. Psychol. 39, 1161?1178.
- [19] Barrett, L.F., Russell, J.A., 1998. *Independence and bipolarity in the structure of current affect*. J. Personal. Soc. Psychol. 74 (4), 967?984.
- [20] Barrett, L.F., 2006. *Valence as a basic building block of emotional life*. J. Res. Pers. 40, 35?55.
- [21] Hurst, E. (1951). *Long-term storage capacity of reservoirs*. Trans. Amer. Soc. Civil Eng. , 116 (1951): 770- 808.
- [22] Koutsoyiannis, D. (2003). *Climate change, the Hurst phenomenon, and hydrological statistics*. Hydrological Sciences Journal , 48.1 (2003): 3-24.
- [23] Dimitriadis, P. (2010). *Three dimensional Hurst-Kolmogorov process for modeling rainfall*. EGU General Assembly Vienna Austria May 2010 (p. Session HS: 5.5). Vienna: EGU.
- [24] Papalexiou, S. (2008). *Probabilistic description of rainfall intensity at multiple time scales*. Role of Hydrology in Water Resource Management. Isle of Capri, Italy: UNESCO.
- [25] Munshi, Jamal. *The Hurst Exponent of Precipitation: England and Wales 1766-2016*. (2017).
- [26] Michelis, Paola De, and Giuseppe Consolini. *On the local Hurst exponent of geomagnetic field fluctuations: Spatial distribution for different geomagnetic activity levels*. Journal of Geophysical Research: Space Physics 120.4 (2015): 2691-2701.
- [27] Kim, Byung-Sik, Hung-Soo Kim, and Sun-Hong Min. *Hurst's Memory for Chaotic, Tree Ring, and SOI Series*. Applied Mathematics 5.1 (2014): 175.
- [28] Mandelbrot, Benoit B., and James R. Wallis. *Robustness of the rescaled range R/S in the measurement of noncyclic long run statistical dependence*. Water Resources Research 5.5 (1969): 967-988.
- [29] H. Hurst, *Long term storage capacity of reservoirs*, Transactions of the American Society of Civil Engineers 6 (1951) 770?799.

- [30] Mandelbrot, Benoit B. *When can price be arbitrated efficiently? A limit to the validity of the random walk and martingale models.* The Review of Economics and Statistics (1971): 225-236.
- [31] Mandelbrot, Benoit. *Statistical methodology for nonperiodic cycles: from the covariance to R/S analysis.* Annals of Economic and Social Measurement, Volume 1, number 3. NBER, 1972. 259-290.
- [32] M. A. Sánchez Granero, J. E. Trinidad Segovia, J. García Pérez, *Some comments on Hurst exponent and the long memory processes on capital markets*
- [33] Annis, A. A., and E. H. Lloyd. *The expected value of the adjusted rescaled Hurst range of independent normal summands.* Biometrika 63.1 (1976): 111-116.
- [34] *PANC1 sejtek morfológiája fáziskontraszt mikroszkópban*
<https://www.pheculturecollections.org.uk/products/celllines/generalcell/detail>
- [35] Razali, Nornadiah Mohd, and Yap Bee Wah. *Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests.* Journal of statistical modeling and analytics 2.1 (2011): 21-33.
- [36] Shapiro, Samuel Sanford, and Martin B. Wilk. *An analysis of variance test for normality (complete samples).* Biometrika 52.3/4 (1965): 591-611.
- [37] Marianna, Bolla. *Statisztikai következtetések elmélete.* Typotex Kft, 2005.